

OmniSplat: Taming Feed-Forward 3D Gaussian Splatting for Omnidirectional Images with Editable Capabilities

Supplementary Material

A. Details on OmniSplat+opt

As briefly mentioned in Section 4.1, we add a small number of optimizations for each scene using the reference view images for fast performance improvement and call the method ‘OmniSplat+opt.’ Compared to ODGS [20], the optimization process of OmniSplat+opt is different in many aspects. First, while ODGS starts optimization from the sparse point cloud estimated by OpenMVG [26], our model starts from 3D Gaussians generated by OmniSplat, which contains much more information. Next, we set the number of optimization steps of OmniSplat+opt to 100, whereas ODGS optimizes 30,000 times. Since OmniSplat’s initial points are highly accurate, significant performance improvement can be achieved with a small number of iterations. Thus, the optimization time only takes 11 seconds, which is more than 150 times shorter than ODGS (32 minutes). Finally, we only optimize the color (sh coefficients) and opacity properties of 3DGS and keep the position and covariance fixed since changes in position or covariance can cause overfitting when optimizing with a few images.

Ablation Studies: number of optimizing iterations We measure the performance of test view images according to the number of optimizing iterations, starting from the original OmniSplat. As mentioned, optimization with reference images improves performance in test view images in the earlier stage. However, if we optimize 3DGS for a long time only with the reference views, the 3DGS may be overfitted, and the performance of the test views may be saturated or even reduced. Table A shows the changes of metrics (PSNR, SSIM, and LPIPS) as well as optimization time according to the number of optimization iterations (# opt.) in OmniBlender [7]. PSNR grows rapidly at the first 100 iterations but becomes saturated as optimization proceeds. For SSIM and LPIPS, the best value is achieved after 200 and 400 iterations of optimization and gets worse when the number of iterations gets larger. In terms of time, the optimization takes about 12 seconds for 100 iterations, and the time linearly grows as the number of optimizations increases. We find out that 100 or 200 iterations are optimal for additional optimization from OmniSplat, considering both time and performance. In the main manuscript, we use 100 times optimization, which takes a similar time to PixelSplat or LatentSplat.

We illustrate the PSNR-runtime trade-off, including the values of Table A in Figure A. Compared to existing feed-forward models, OmniSplat achieves higher PSNR with the

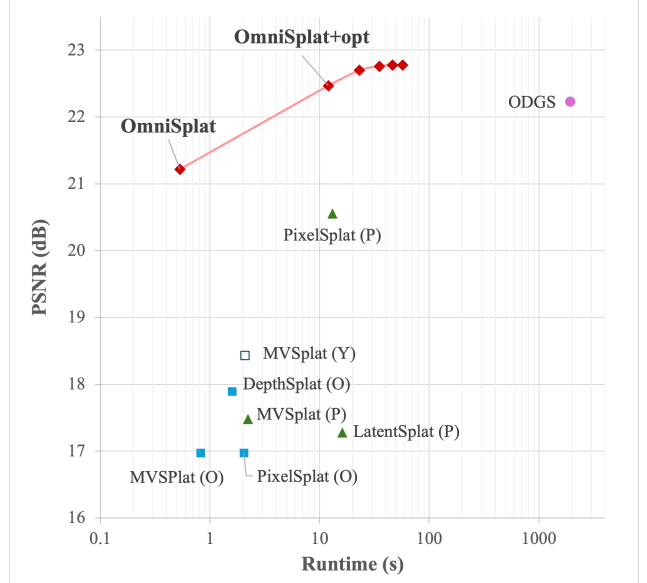


Figure A. **PSNR-runtime trade-off.** A chart of reconstruction performance-runtime trade-off in novel view image on OmniBlender [7] including the ablation according to the number of optimizations.

# opt.	0	100	200	300	400	500
PSNR	21.22	22.47	22.70	22.76	22.78	22.78
SSIM	0.6519	0.6960	0.6979	0.6962	0.6938	0.6914
LPIPS	0.3636	0.3209	0.3144	0.3125	0.3124	0.3129
Time (s)	0	12	23	35	46	57

Table A. Changes in performance according to the number of iterations for adaptation in OmniBlender [7].

fastest running speed. Additionally, the PSNR values get higher than ODGS within a minute of optimization, which are denoted as OmniSplat+opt. From the figure, we maintain that OmniSplat achieves the best PSNR-runtime trade-off for novel view synthesis.

B. Additional Gaussian Segmentation and Editing Results

To facilitate 3D editing in the result Gaussians, we first construct multiview consistent segmentation maps using the cross-attention scores. Figure B presents additional examples of multiview consistent segmentation. The selection operation allows users to choose regions by selecting Gaussians with the same label. Unlike optimization-based methods, pixel-aligned Gaussians provide cleaner boundary separa-

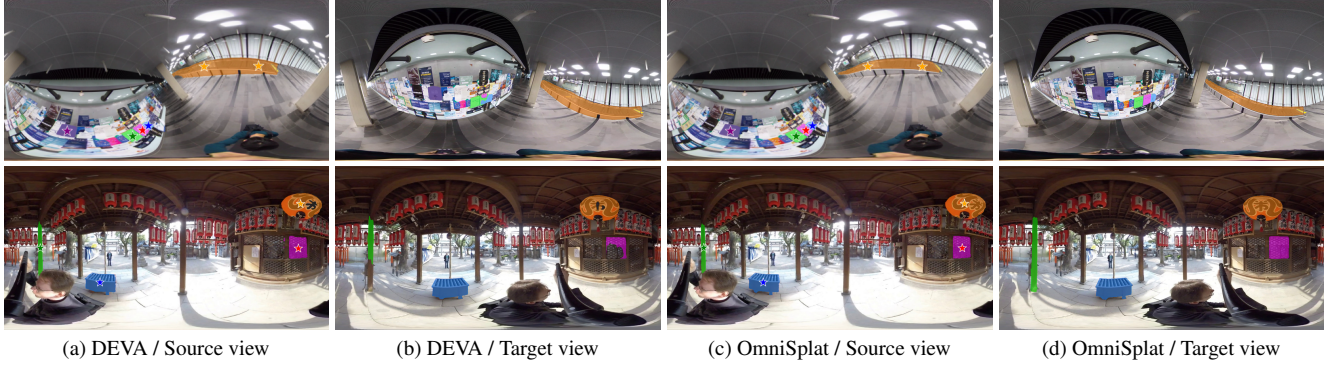


Figure B. **Visualization of segmentation matching.** We visualize the matched segments among the source and the target views. The stars in the image indicate query points for the user to segment objects containing the stars.

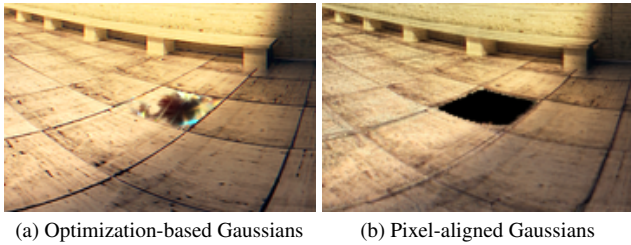


Figure C. **Example of Gaussian removal results in the object segment.** Compared to (a), where overlapping Gaussians result in incomplete removal, (b) pixel-aligned Gaussians offer clear edges that are advantageous for editing.

tion. Figure C illustrates the result of removing 3D Gaussians selected based on the multiview segmentation regions. The removal result with the Gaussians optimized as 3DGS is illustrated in Figure Ca. Since the optimized Gaussians tend to have elongated shapes [15, 33, 36], the selected regions are not cleanly removed, often resulting in needle-like artifacts. In contrast, our method with the pixel-aligned Gaussians provides clear boundaries, enabling a cleaner removal operation and establishing a solid foundation for subsequent operations such as image inpainting.

C. Implementation Details and Additional Quantitative Results

We write the indices for reference and test views for each scene in Tables B to G along with detailed quantitative results. For three datasets with relatively small camera motions (OmniBlender, Ricoh360 [7], OmniPhotos [2]), we select the two reference views with large pose distances and set the rest as test views. For the other three datasets (360Roam [13], OmniScenes [19], 360VO [12]), the pose distance between two adjacent cameras is large, and we selected two frames with 3 to 4 timestep intervals as two reference views. Then, the test view indices are composed of frames between the reference views. For example, if the reference view indices

are 1 and 5, the test view indices are 2, 3, and 4.

Also, we write quantitative metrics for every scene in Tables B to G, respectively. The best result for each metric is written in **bold**, and the second best result is written in underlined. As shown in the tables, OmniSplat+*opt* shows the best or the second best results in almost all metrics. Although ODGS shows conspicuous results, ODGS still shows optimal quality-runtime tradeoff, considering that ODGS requires scene-wise optimization.

D. Additional Qualitative Results

In Figures D to I, we provide additional qualitative comparison with more baselines on various datasets. Perspective feed-forward models, PixelSplat (P) [3] and MVSplat (P) [4], produce blurry images. For PixelSplat, Gaussians are generated for every pair among twelve perspective images, and since the positions of the Gaussians generated for each pair are slightly different, a blurry image is rendered in the novel viewpoints. For MVSplat, the encoded features and cross-attention values are averaged across all perspective images, which causes blur or ghost artifacts in novel view synthesized images. Omnidirectional feed-forward models, PixelSplat (O), MVSplat (O), and DepthSplat (O), produce dark images overall and striped patterns in the upper and lower areas. We attribute the phenomenon to the non-uniform sampling of the omnidirectional grid during rasterization. ODGS sometimes shows prominent image quality, but it requires over 30 minutes of optimization for each scene. Moreover, it often falls into overfitting, generating floater artifacts that severely degrade the quality of the image. Our method, OmniSplat, produces clearer and more accurate images than other feed-forward networks and ODGS. Also, the results can be further improved with a small number of optimization (OmniSplat+*opt*), which only takes 12 seconds, which is shorter than the execution time of PixelSplat (P).

We also provide a video sample that compares our models with PixelSplat (P), MVSplat (O), and ODGS. This video was rendered with a pose sequence created by interpolating

two reference views from several scenes. We can clearly see the floater artifact of ODGS, the blurriness of Pixelsplat (P), and the dark strip pattern of MVSplat (O) in this video. On the other hand, OmniSplat maintains consistently good image quality in all novel views.

Scene	Ref. view	Test view	ODGS	PixelSplat (P)	LatentSplat (P)	MVSP (P)	PixelSplat (O)	MVSP (O)	DepthSplat (O)	MVSP (Y)	OmniSplat	OmniSplat+opt
archviz-flat	14.86	2, 6, ..., 98	22.17/0.7511/0.3188	20.93/0.7128/0.3274	18.40/0.6859/0.3981	18.35/0.6704/0.3870	17.28/0.4285/0.6206	17.73/0.6818/0.4138	19.06/0.6812/0.4491	19.76/0.7013/0.3740	23.93/0.8010/0.2688	24.97/0.8147/0.3618
barbershop	14.86	2, 6, ..., 98	25.01/0.7683/0.3588	19.83/0.6243/0.3588	17.96/0.5627/0.4666	17.98/0.5636/0.4445	17.70/0.4262/0.6197	16.60/0.5214/0.4459	18.13/0.5510/0.5534	18.32/0.5615/0.4885	21.76/0.7067/0.3328	21.75/0.7067/0.3328
bistro bike	14.86	2, 6, ..., 98	21.05/0.6800/0.2687	20.20/0.5882/0.2981	15.43/0.4248/0.4343	17.29/0.4619/0.4241	16.50/0.3613/0.5527	16.40/0.5253/0.3354	16.97/0.4106/0.4817	17.54/0.4463/0.4007	20.20/0.6312/0.2844	21.75/0.6868/0.2459
bistro square	14.86	2, 6, ..., 98	19.30/0.6397/0.3030	20.11/0.5989/0.2904	15.43/0.4290/0.4427	16.41/0.4348/0.4376	14.93/0.2878/0.5746	13.76/0.4338/0.4070	15.36/0.3853/0.5050	16.11/0.4381/0.4085	18.43/0.5535/0.3185	18.98/0.5845/0.2844
classroom	14.86	2, 6, ..., 98	23.55/0.6436/0.3655	20.50/0.6803/0.3112	17.46/0.6150/0.4574	16.47/0.6022/0.4719	16.61/0.3884/0.6115	16.30/0.5223/0.4130	17.23/0.5620/0.5403	17.97/0.6018/0.4537	20.91/0.7213/0.3505	21.63/0.7254/0.3107
fisher hut	14.86	2, 6, ..., 98	21.07/0.6526/0.3516	22.75/0.7389/0.3253	19.91/0.6673/0.3773	21.08/0.7226/0.3660	18.89/0.4390/0.5900	20.90/0.6487/0.4265	21.83/0.6690/0.4156	21.79/0.6726/0.3644	25.08/0.7254/0.3155	26.60/0.7262/0.2743
lone monk	14.86	2, 6, ..., 98	20.21/0.6907/0.3402	16.89/0.5907/0.3402	14.12/0.4640/0.4077	15.23/0.3075/0.5029	14.93/0.3075/0.5029	13.09/0.5371/0.4259	15.69/0.4690/0.4417	16.51/0.5120/0.3855	19.04/0.6175/0.2978	19.28/0.6374/0.2945
LOU	14.86	2, 6, ..., 98	21.53/0.7332/0.2469	18.97/0.6801/0.3246	14.47/0.5251/0.4452	15.17/0.4348/0.4459	16.49/0.4040/0.5842	18.16/0.6678/0.3141	16.53/0.5327/0.4082	16.72/0.4991/0.3821	20.51/0.6438/0.3071	20.33/0.7420/0.2622
pavilion middle chair	14.86	2, 6, ..., 98	21.24/0.6541/0.3516	21.28/0.7035/0.3198	18.29/0.6050/0.4178	18.06/0.6347/0.3868	17.06/0.4832/0.5765	17.86/0.6136/0.4251	17.86/0.6136/0.4251	18.06/0.6856/0.3216	21.02/0.7026/0.3561	22.09/0.7026/0.3561
pavilion middle pond	14.86	2, 6, ..., 98	19.21/0.4918/0.4229	19.70/0.6058/0.3553	15.31/0.4486/0.5241	14.24/0.4348/0.6142	15.40/0.3021/0.6316	15.63/0.4752/0.4120	16.46/0.4528/0.4074	16.24/0.4902/0.4407	18.34/0.5609/0.3776	19.18/0.5749/0.3328
restroom	14.86	2, 6, ..., 98	30.21/0.8030/0.2220	24.95/0.6947/0.3132	23.33/0.6481/0.4063	21.71/0.6551/0.4263	21.54/0.4381/0.5999	21.43/0.6229/0.3305	21.67/0.5730/0.5103	23.16/0.6379/0.4304	23.02/0.7266/0.3792	28.17/0.7720/0.2845
average	-	-	22.23/0.6807/0.3124	20.56/0.6562/0.3222	17.28/0.5523/0.4361	17.48/0.5565/0.4385	16.97/0.3837/0.5967	16.97/0.5635/0.3949	17.89/0.5364/0.4753	18.43/0.5627/0.4107	21.02/0.6601/0.3231	22.33/0.6994/0.2879

Table B. Scene-wise quantitative results of 3D reconstruction on **OmniBlender** dataset.

Scene	Ref. view	Test view	ODGS	PixelSplat (P)	LatentSplat (P)	MVSP (P)	PixelSplat (O)	MVSP (O)	DepthSplat (O)	MVSP (Y)	OmniSplat	OmniSplat+opt
bricks	35.69	1, 3, ..., 99	16.41/0.5116/0.3738	16.28/0.5161/0.4249	13.38/0.4586/0.4781	17.86/0.5834/0.3820	15.89/0.4091/0.5296	15.01/0.5017/0.4070	13.34/0.3861/0.5950	16.76/0.5072/0.4053	19.25/0.5980/0.3374	19.91/0.6388/0.3052
bridge	33.57	1, 3, ..., 99	15.84/0.4353/0.4546	16.45/0.5264/0.3999	15.71/0.4367/0.5402	18.88/0.6020/0.2446	16.62/0.3937/0.5269	13.83/0.4450/0.4677	13.41/0.4371/0.5585	16.80/0.5111/0.3976	19.13/0.5947/0.3436	19.53/0.6198/0.3183
bridge under	23.77	1, 3, ..., 99	18.57/0.4841/0.3916	15.78/0.4519/0.4572	14.67/0.4772/0.4473	14.24/0.4492/0.4896	15.51/0.3123/0.5839	16.08/0.4179/0.4140	16.72/0.4394/0.4879	17.57/0.4607/0.4183	19.59/0.5598/0.3572	19.52/0.5794/0.3292
cat tower	3.83	1, 3, ..., 99	15.09/0.4641/0.4210	20.19/0.7243/0.3186	17.14/0.4167/0.6799	17.14/0.4167/0.6799	16.68/0.3654/0.5489	16.53/0.4640/0.4855	14.83/0.4511/0.5021	16.05/0.4910/0.4359	18.65/0.5348/0.4005	19.42/0.5507/0.3579
center	25.49	1, 3, ..., 99	19.56/0.6045/0.3659	21.61/0.6936/0.2948	17.28/0.3716/0.5863	21.32/0.7594/0.3107	17.97/0.4407/0.5556	17.01/0.6090/0.3886	16.38/0.5985/0.4985	19.15/0.6708/0.3459	20.66/0.7094/0.3210	21.57/0.7320/0.2887
farm	83.99	1, 3, ..., 99	18.05/0.5087/0.3431	18.31/0.5168/0.3605	16.31/0.4110/0.4937	19.73/0.6789/0.3447	16.49/0.3470/0.5068	12.93/0.3631/0.4969	17.22/0.4613/0.4171	17.76/0.4940/0.3723	18.29/0.5133/0.3646	18.56/0.5222/0.3409
flower	29.55	1, 3, ..., 99	15.37/0.4163/0.4084	17.02/0.5129/0.4260	16.46/0.4410/0.4992	16.03/0.5468/0.4202	15.65/0.3063/0.5376	14.57/0.3507/0.4954	15.21/0.3630/0.5423	14.83/0.3561/0.4838	16.90/0.4306/0.4350	17.46/0.4436/0.4120
gallery chair	5.23	1, 3, ..., 99	19.78/0.6557/0.3387	25.12/0.8192/0.2625	18.96/0.3871/0.7146	22.86/0.7280/0.3349	19.19/0.4497/0.5694	17.62/0.5742/0.4504	19.57/0.6800/0.4585	20.26/0.7888/0.4039	22.30/0.7374/0.3264	22.89/0.7427/0.3060
gallery park	33.99	1, 3, ..., 99	16.29/0.5766/0.3848	19.19/0.6971/0.3601	18.77/0.3470/0.7216	19.67/0.7169/0.3522	18.65/0.4561/0.5442	17.54/0.5931/0.4490	14.04/0.5440/0.5199	17.99/0.6154/0.4014	19.64/0.6568/0.3670	20.13/0.6653/0.3040
gallery pillar	21.39	1, 3, ..., 99	19.67/0.6461/0.3664	21.80/0.7497/0.3213	20.00/0.3621/0.7285	19.03/0.6873/0.3694	17.23/0.3913/0.5875	17.16/0.5931/0.4376	19.70/0.6556/0.4402	20.33/0.6451/0.3821	21.45/0.7052/0.3477	22.51/0.7282/0.3048
garden	31.59	1, 3, ..., 99	16.97/0.5338/0.4027	20.38/0.6525/0.3989	19.22/0.3886/0.6406	23.02/0.7193/0.3373	18.84/0.3943/0.5652	15.53/0.4060/0.4790	16.13/0.5098/0.5158	20.73/0.5723/0.4244	22.43/0.6335/0.3406	23.45/0.6544/0.3052
poster	51.85	1, 3, ..., 99	18.46/0.5342/0.4219	20.15/0.6964/0.3263	15.72/0.4675/0.5827	17.99/0.6179/0.4047	16.01/0.3879/0.5911	14.36/0.4897/0.4582	13.61/0.4881/0.5847	17.22/0.5780/0.4435	18.71/0.6517/0.3761	19.36/0.6674/0.3551
average	-	-	17.51/0.5309/0.3911	19.94/0.6486/0.3119	16.30/0.5389/0.4386	18.12/0.5947/0.3804	17.66/0.4356/0.5810	15.89/0.5478/0.4049	17.43/0.5470/0.4482	17.88/0.5482/0.4123	19.70/0.6104/0.3598	20.36/0.6288/0.3303

Table C. Scene-wise quantitative results of 3D reconstruction on **Ricoh360** dataset.

Scene	Ref. view	Test view	ODGS	PixelSplat (P)	LatentSplat (P)	MVSP (P)	PixelSplat (O)	MVSP (O)	DepthSplat (O)	MVSP (Y)	OmniSplat	OmniSplat+opt
Ballinloy	0.35	0, 5, ..., 90	19.89/0.5674/0.2157	21.78/0.7294/0.2729	17.89/0.6572/0.3928	20.46/0.7433/0.2816	18.49/0.4102/0.5756	18.45/0.7034/0.3039	20.34/0.7056/0.3548	20.12/0.7628/0.3308	19.07/0.7751/0.3469	21.46/0.8076/0.4315
BeihaiPark	0.55	0, 5, ..., 80	21.38/0.6417/0.2835	20.37/0.6286/0.3051	17.64/0.5379/0.4567	16.86/0.4977/0.4817	16.89/0.4015/0.5363	14.06/0.4564/0.4584	16.32/0.4822/0.4371	16.52/0.4948/0.4202	17.52/0.5324/0.3989	17.62/0.5432/0.3733
Cathedral	0.50	0, 5, ..., 80	19.54/0.5006/0.3809	17.08/0.5528/0.3270	13.50/0.3863/0.4850	15.29/0.4623/0.4196	13.22/0.2504/0.6097	13.79/0.4555/0.4124	14.64/0.4180/0.4833	15.12/0.4402/0.4537	15.71/0.5288/0.4045	16.24/0.5658/0.3672
Coast	0.45	0, 5, ..., 80	21.03/0.4943/0.4314	22.53/0.7179/0.2839	17.95/0.5629/0.4724	21.74/0.7544/0.2828	17.78/0.3252/0.5968	18.02/0.6409/0.3138	19.54/0.5807/0.4244	19.43/0.6072/0.3382	19.89/0.6807/0.3580	21.64/0.7268/0.2645
Field	0.45	0, 5, ..., 75	21.10/0.5530/0.4285	25.70/0.7509/0.2636	19.96/0.6498/0.4201	25.54/0.7584/0.2970	19.19/0.4497/0.5694	20.04/0.6931/0.3745	21.40/0.6279/0.4337	23.14/0.6644/0.3847	22.20/0.6903/0.4038	25.15/0.7116/0.2986
Numokiki2	0.50	0, 5, ..., 80	19.62/0.5883/0.3876	20.03/0.6641/0.3096	16.47/0.5805/0.4115	18.15/0.6152/0.3897	16.03/0.4017/0.5563	16.67/0.5549/0.4363	19.15/0.5970/0.4193	18.91/0.5965/0.3930	19.89/0.6267/0.3906	19.68/0.6519/0.3215
SecretGarden1	0.40	0, 5, ..., 75	20.91/0.6096/0.3430	19.57/0.6675/0.3021	17.00/0.5812/0.4043	17.85/0.6073/0.3741	18.50/0.4073/0.5424	15.13/0.5545/0.4298	17.44/0.5730/0.4345	17.85/0.6203/0.3968	17.71/0.6814/0.3370	20.71/0.7059/0.2946
Shrines1	0.45	0, 5, ..., 90	18.34/0.4264/0.3979	16.88/0.4936/0.4174	13.35/0.3260/0.5337	15.71/0.4179/0.5271	15.56/0.2643/0.5913	14.35/0.3434/0.4814	14.67/0.3491/0.5763	19.39/0.3655/0.5163	16.75/0.4375/0.4634	17.37/0.4712/0.4282
Temple3	0.25	0, 5, ..., 70	20.83/0.6062/0.3167	17.17/0.5869/0.3441	14.21/0.4620/0.4762	13.63/0.4758/0.4555	13.02/0.3183/0.6215	12.30/0.4879/0.4152	13.02/0.4896/0.5080	14.51/0.5221/0.4395	16.24/0.6164/0.3657	16.67/0.6164/0.3657
Wukongting	0.50	0, 5, ..., 95	19.88/0.6722/0.3403	18.31/0.6938/0.2926	15.01/0.6450/0.4135	16.04/0.6192/0.3652	15.56/0.3827/0.5816	16.10/0.6424/0.4236	16.79/0.6466/0.4108	16.59/0.6461/0.3760	19.13/0.7412/0.3253	19.76/0.7633/0.2773
average	-	-	20.25/0.5660/0.3730	19.94/0.6486/0.3119	16.30/0.5389/0.4386	18.12/0.5947/0.3804	17.66/0.4356/0.5810	15.89/0.5478/0.4049	17.43/0.5470/0.4482	17.76/0.5768/0.4051	18.50/0.6311/0.3794	19.63/0.6608/0.3273

Table D. Scene-wise quantitative results of 3D reconstruction on **OmniPhotos** dataset.

Scene	Ref. view	Test view	ODGS	PixelSplat (P)	LatentSplat (P)	MVSP (P)	PixelSplat (O)	MVSP (O)	DepthSplat (O)	MVSP (Y)	OmniSplat	OmniSplat+opt
bar	6.10	7, 8, 9	17.59/0.5680/0.3151	14.25/0.5383/0.4582	13.76/0.3802/0.5071	13.42/0.3278/0.5498	13.16/0.2092/0.6602	11.90/0.2472/0.5911	12.49/0.2652/0.6354	13.48/0.3211/0.5363	13.28/0.2862/0.5827	16.90/0.4730/0.4315
base	21.25	22, 23, 24	19.37/0.5900/0.3069	20.37/0.6286/0.3051	14.03/0.4357/0.4664	15.79/0.4301/0.4723	13.48/0.3134/0.6492	12.42/0.3149/0.4942	14.62/0.3606/0.5305	14.69/0.3652/0.5062	15.41/0.4231/0.4593	15.27/0.3996/0.4303
cafe	41.45	42, 43, 44	18.41/0.5789/0.3565	19.05/0.4919/0.4498	15.52/0.4934/0.4787	17.17/0.4946/0.4764	15.08/0.2912/0.6352	13.28/0.3129/0.4828	16.64/0.4424/0.5380	16.77/0.4486/0.5003	18.23/0.5444/0.4240	19.05/0.5758/0.3976
cantenna	74.78	75, 76, 77	16.42/0.4282/0.4944	13.37/0.4187/0.5467	12.94/0.4511/0.5565	13.39/0.4411/0.5745	13.44/0.2678/0.6525	10.99/0.2637/0.6211	13.55/0.4009/0.6010	14.38/0.4320/0.5675	15.39/0.4578/0.5203	15.16/0.4031/0.5318
center	52.56	53, 54, 55	21.07/0.6161/0.3905	17.17/0.5947/0.4146	16.70/0.6133/0.4874	16.91/0.5933/0.4656	17.05/0.4015/0.5175	17.15/0.5602/0.5346	18.11/0.6064/0.4883	20.48/0.6726/0.4130	20.88/0.6652/0.3642	21.05/0.6559/0.3215
center1	36.40	37, 38, 39	19.37/0.6232/0.4377	16.33/0.6594/0.4464	16.12/0.6832/0.4784	16.35/0.6534/0.4868	15.31/0.2966/0.6798	12.10/0.3938/0.5798	18.03/0.6464/0.5146	17.39/0.6279/0.5318	20.59/0.7127/0.4459	20.69/0.

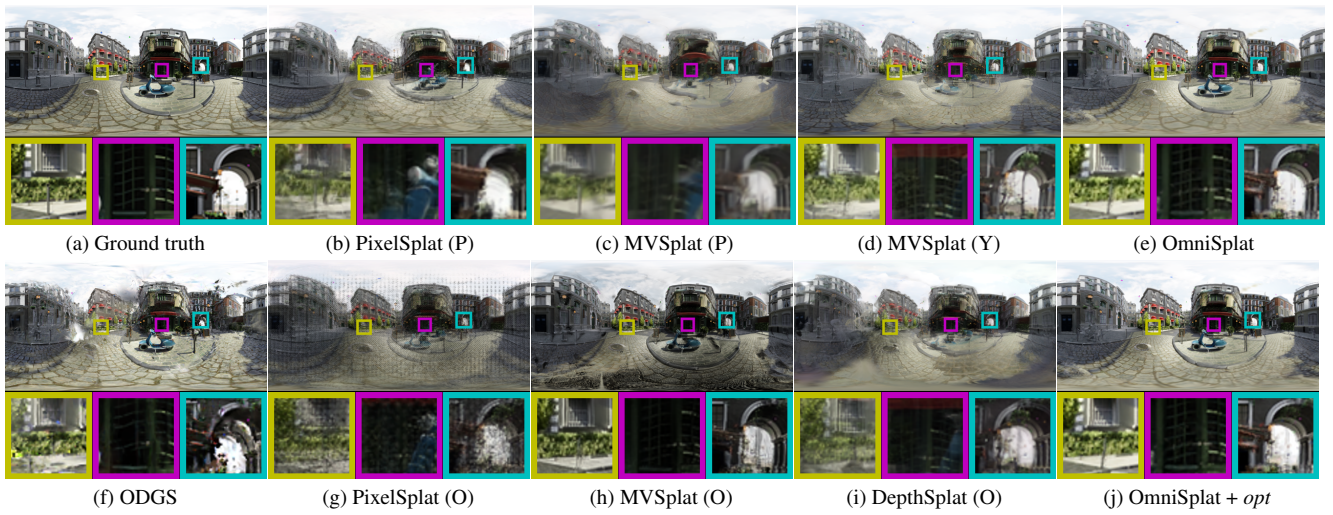


Figure D. Qualitative comparison on OmniBlender dataset.

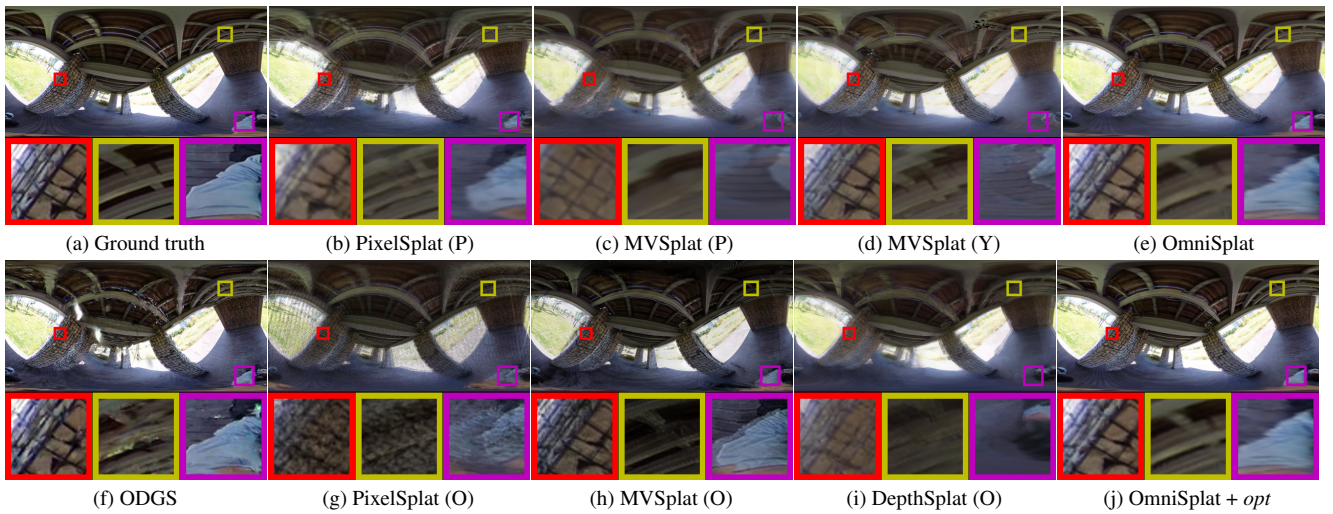


Figure E. Qualitative comparison on Ricoh dataset.

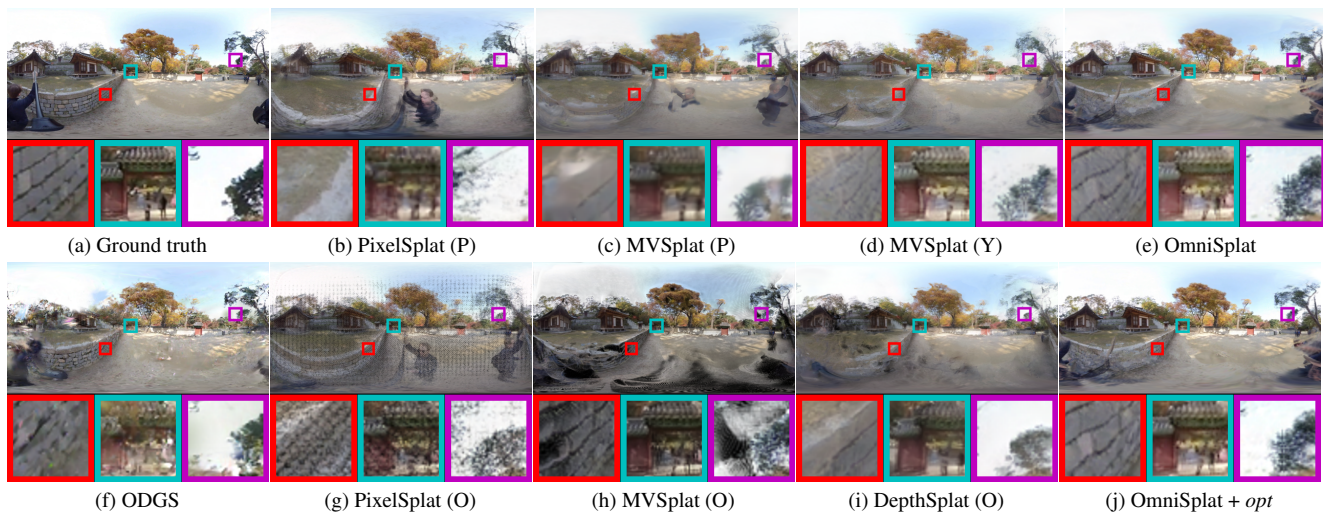


Figure F. Qualitative comparison on OmniPhotos dataset.

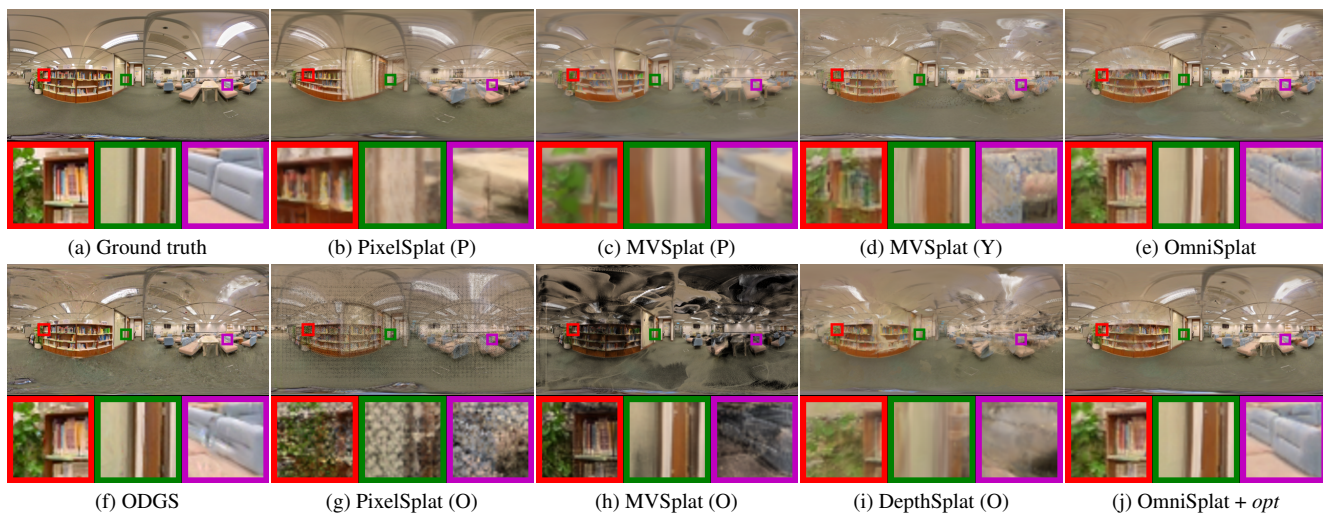


Figure G. Qualitative comparison on 360Roam dataset.

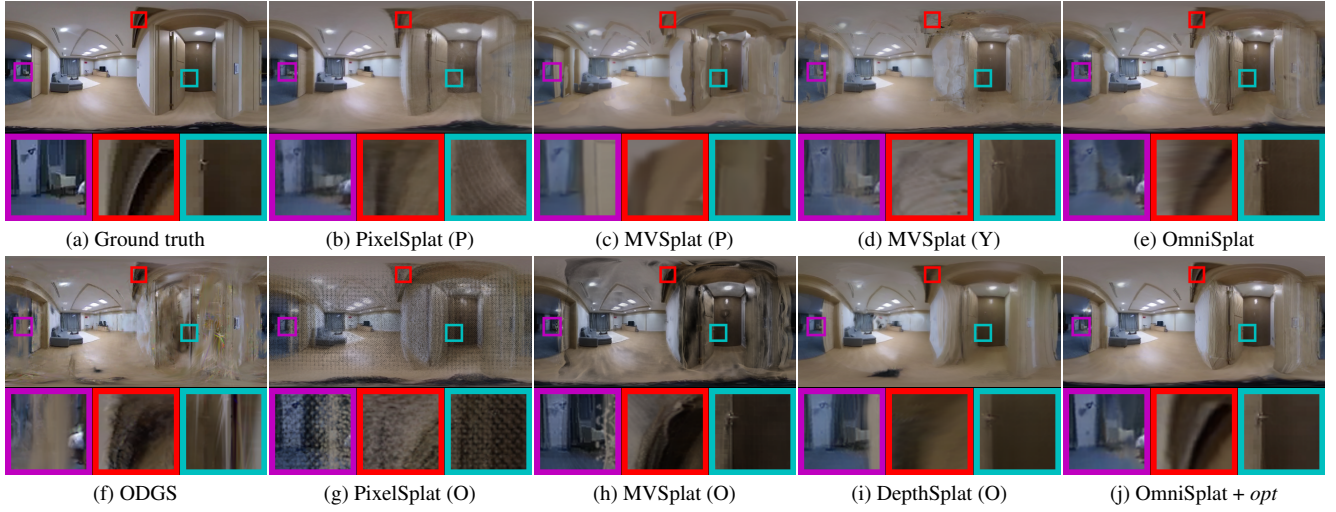


Figure H. Qualitative comparison on OmniScenes dataset.

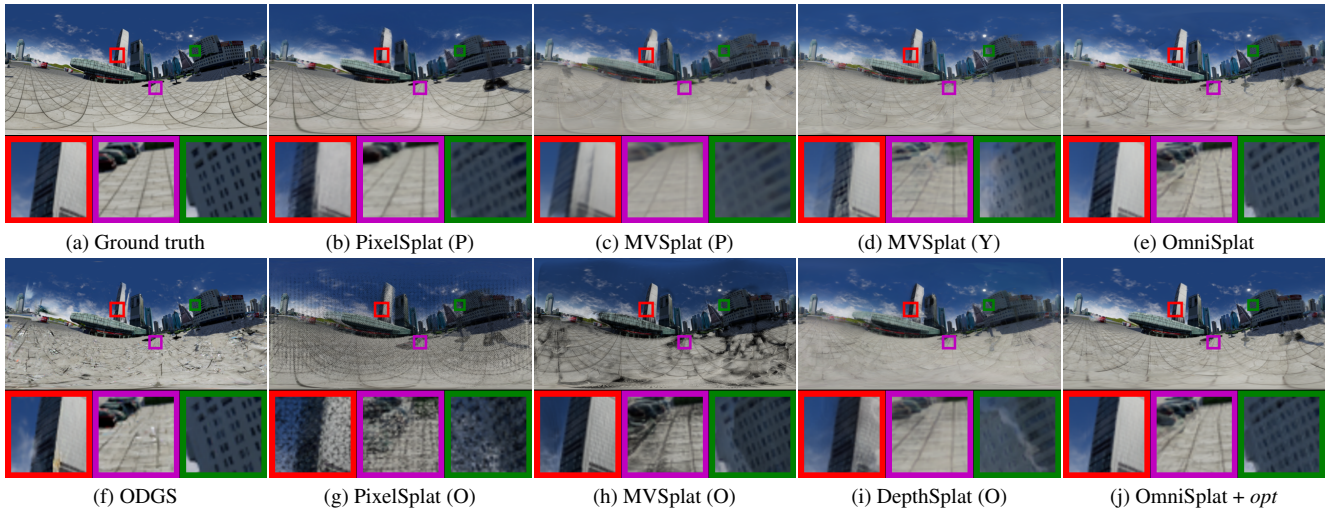


Figure I. Qualitative comparison on 360VO dataset.