XLD: A Cross-Lane Dataset for Benchmarking Novel Driving View Synthesis

Supplementary Material

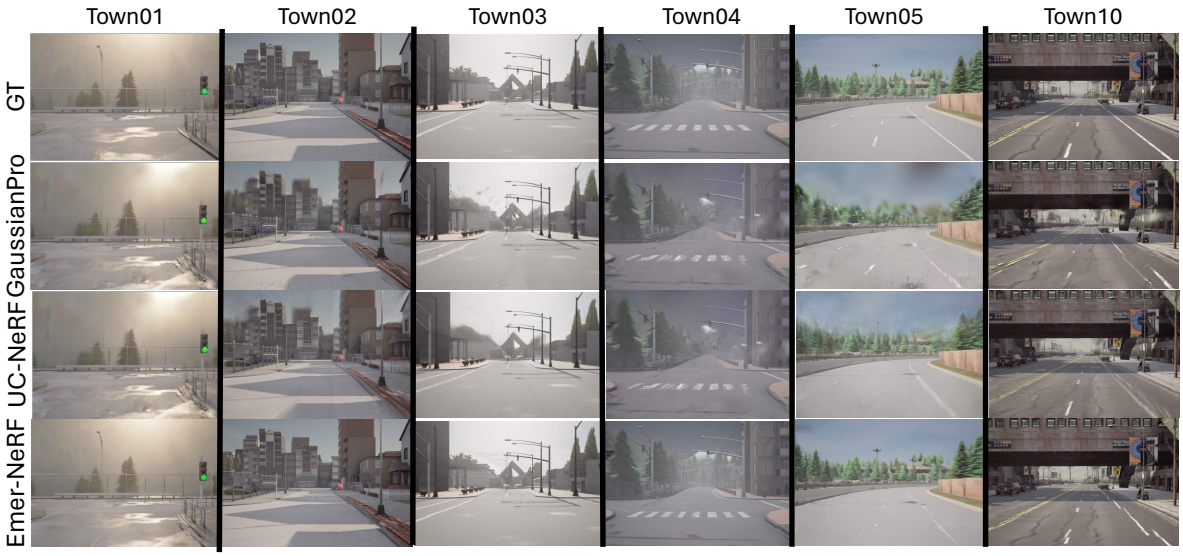

Figure 7. The ground truth and rendered results, obtained from various benchmarking methods, are compared across different scenes with a 1-meter offset.

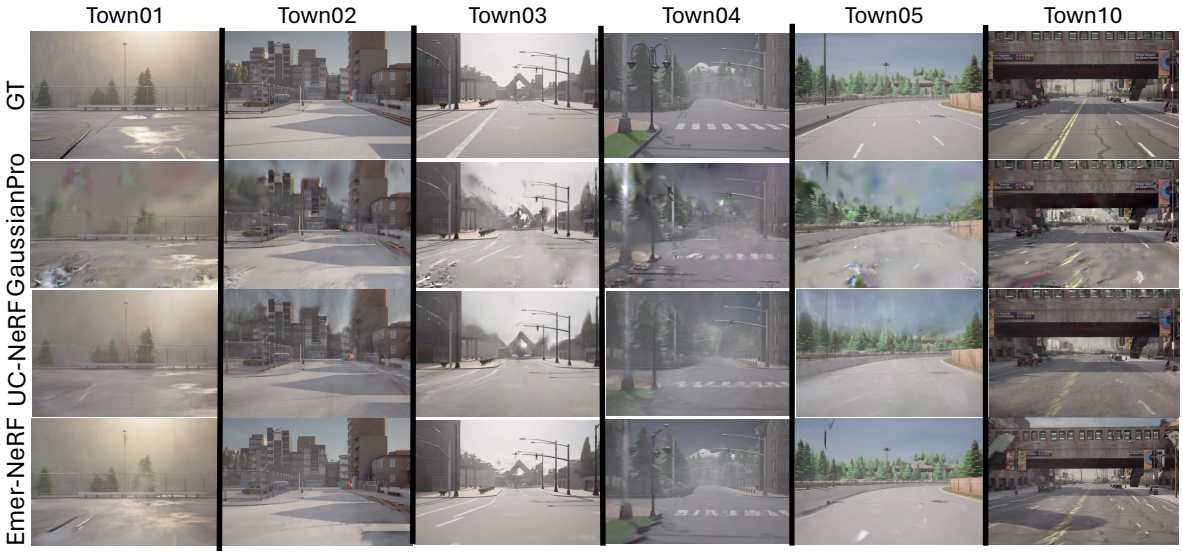

Figure 8. The ground truth and rendered results, obtained from various benchmarking methods, are compared across different scenes with a 4-meter offset.

In this supplementary, we present the evaluation results of novel view synthesis using both front-only and multi-camera settings. For the front-only camera setup, we detail the results with various offsets (e.g., 0m, 1m, 2m, 4m), as illustrated in Tables 4, 5, 6, and 7. Additionally, we provide comprehensive experimental results for the multi-camera configuration, which are shown in Table 12. The novel view synthesis results for this setting, also with different offsets (e.g., 0m, 1m, 2m, 4m), are detailed in Tables 8, 9, 10, and 11.

We also present additional results of various methods across different offsets for the Town-01 scene, as illustrated in Fig. 9. Furthermore, we showcase the visualization results for each scene using different methods. Specifically, Fig. 7 displays the results with a 1-meter offset, while Fig. 8 presents the results with a 4-meter offset.

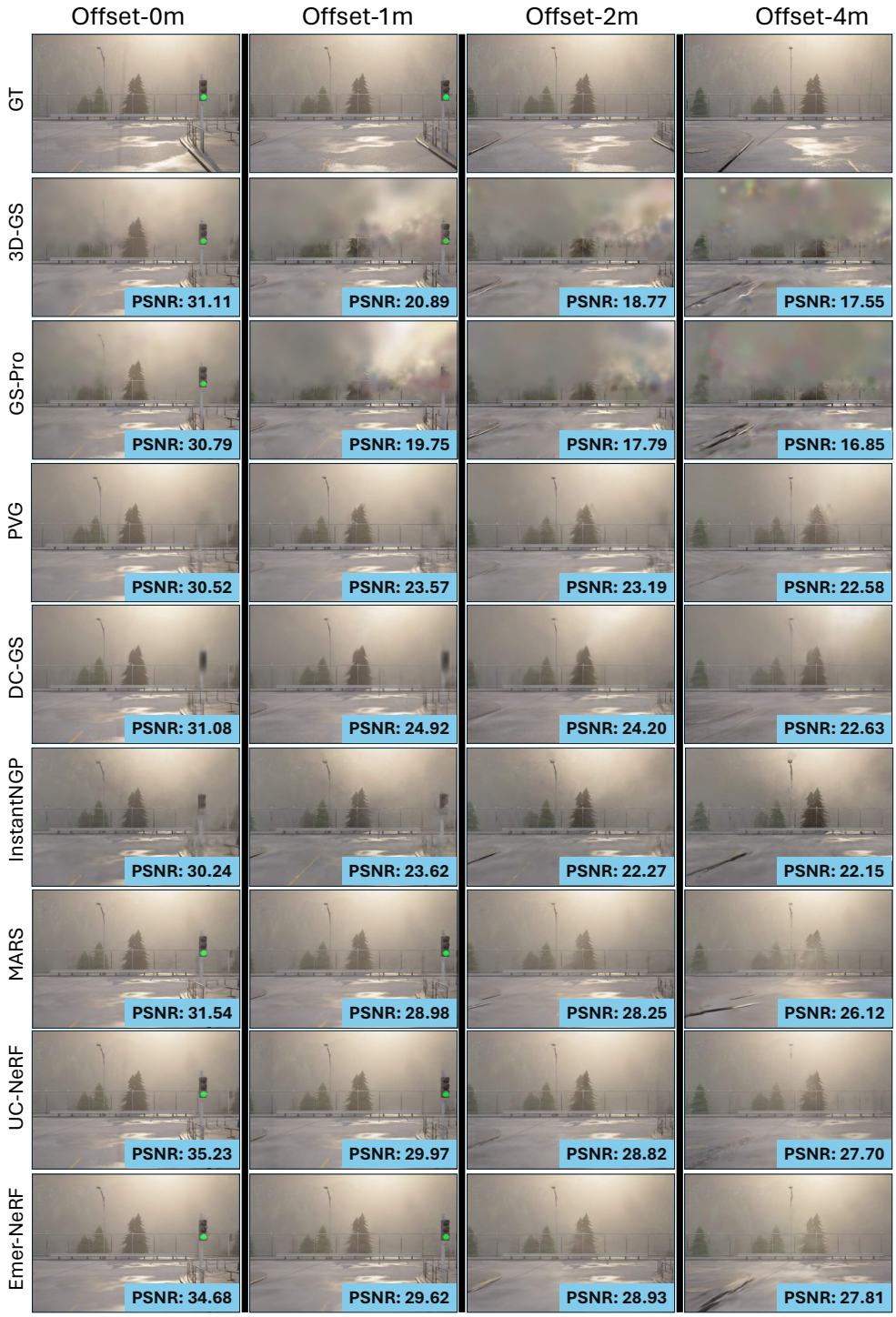

Figure 9. The ground truth and rendered results by different benchmarking methods from a sequence under different meters of offsetting.

Table 4. Results on our proposed benchmark with the 0 meter offset with 1 camera.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	31.64	0.919	0.214	30.95	0.881	0.265	32.13	0.928	0.203
UC-NeRF [15]	39.35	0.958	0.315	35.19	0.941	0.288	34.53	0.939	0.303
MARS [57]	32.70	0.896	0.140	31.87	0.921	0.100	31.76	0.894	0.125
NeRFacto [48]	28.74	0.905	0.253	28.40	0.901	0.224	31.25	0.924	0.190
EmerNeRF [60]	33.43	0.932	0.111	31.72	0.914	0.106	32.33	0.922	0.102
- Gaussian-based									
3D-GS [26]	32.12	0.937	0.326	29.09	0.908	0.295	29.62	0.916	0.248
PVG [13]	39.16	0.971	0.212	38.81	0.973	0.12	38.08	0.968	0.171
GaussianPro [16]	35.51	0.953	0.266	30.56	0.915	0.251	29.39	0.910	0.242
DC-Gaussian [53]	35.51	0.953	0.266	30.56	0.915	0.251	29.39	0.91	0.242

	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	29.58	0.921	0.221	27.37	0.892	0.241	26.86	0.820	0.372
UC-NeRF [15]	39.03	0.957	0.300	33.60	0.923	0.321	34.00	0.900	0.340
MARS [57]	36.68	0.957	0.084	26.31	0.857	0.188	30.44	0.875	0.151
NeRFacto [48]	26.09	0.899	0.265	23.64	0.871	0.261	26.23	0.826	0.316
EmerNeRF [60]	36.42	0.944	0.094	26.63	0.873	0.175	30.03	0.854	0.167
- Gaussian-based									
3D-GS [26]	33.60	0.946	0.280	25.95	0.872	0.380	28.45	0.841	0.390
PVG [13]	39.58	0.971	0.194	34.91	0.954	0.194	36.16	0.921	0.244
GaussianPro [16]	33.69	0.943	0.273	30.54	0.916	0.249	30.03	0.879	0.301
DC-Gaussian [53]	33.69	0.943	0.273	26.17	0.916	0.249	30.03	0.879	0.301

Table 5. Results on our proposed benchmark with the 1-meter offset with 1 camera.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	24.19	0.868	0.350	20.25	0.766	0.390	23.92	0.841	0.288
UC-NeRF [15]	29.60	0.925	0.355	28.68	0.894	0.338	30.11	0.911	0.336
MARS [57]	28.27	0.886	0.145	27.95	0.883	0.138	29.04	0.899	0.140
NeRFacto [48]	24.48	0.884	0.300	24.40	0.842	0.279	23.55	0.835	0.278
EmerNeRF [60]	28.81	0.913	0.123	27.82	0.876	0.143	29.39	0.901	0.126
- Gaussian-based									
3D-GS [26]	24.07	0.909	0.321	19.98	0.799	0.372	21.22	0.811	0.334
PVG [13]	25.62	0.917	0.282	27.25	0.912	0.267	28.15	0.907	0.253
GaussianPro [16]	23.97	0.902	0.329	20.71	0.811	0.360	21.92	0.840	0.322
DC-Gaussian [53]	27.34	0.923	0.272	24.87	0.873	0.303	26.35	0.891	0.275

	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	26.21	0.869	0.277	22.18	0.800	0.338	23.88	0.739	0.434
UC-NeRF [15]	34.10	0.938	0.338	29.42	0.890	0.353	28.51	0.820	0.410
MARS [57]	34.28	0.937	0.097	24.42	0.831	0.210	27.25	0.812	0.199
NeRFacto [48]	24.20	0.858	0.311	21.00	0.794	0.336	23.30	0.732	0.380
EmerNeRF [60]	34.15	0.930	0.102	24.83	0.851	0.194	26.96	0.797	0.211
- Gaussian-based									
3D-GS [26]	25.62	0.905	0.343	21.52	0.840	0.346	21.75	0.843	0.342
PVG [13]	29.19	0.926	0.286	25.33	0.864	0.293	25.51	0.764	0.397
GaussianPro [16]	25.56	0.907	0.341	21.52	0.840	0.346	21.92	0.841	0.335
DC-Gaussian [53]	30.67	0.936	0.266	26.17	0.883	0.297	25.51	0.796	0.376

Table 6. Results on our proposed benchmark with the 2-meter offset with 1 camera.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	24.21	0.861	0.377	20.14	0.767	0.405	21.61	0.802	0.344
UC-NeRF [15]	28.45	0.917	0.373	27.53	0.895	0.347	29.04	0.908	0.343
MARS [57]	26.58	0.875	0.165	27.29	0.886	0.132	27.37	0.886	0.151
NeRFacto [48]	23.73	0.869	0.334	22.95	0.818	0.312	21.29	0.795	0.333
EmerNeRF [60]	27.12	0.902	0.143	27.16	0.879	0.137	27.72	0.888	0.137
- Gaussian-based									
3D-GS [26]	22.00	0.881	0.369	19.28	0.796	0.394	21.91	0.840	0.322
PVG [13]	24.58	0.910	0.306	26.13	0.885	0.293	26.58	0.894	0.272
GaussianPro [16]	21.89	0.879	0.368	19.28	0.796	0.394	21.91	0.84	0.322
DC-Gaussian [53]	25.89	0.912	0.294	23.28	0.866	0.314	24.61	0.876	0.299

	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	24.71	0.841	0.327	21.03	0.776	0.387	22.5	0.694	0.474
UC-NeRF [15]	31.31	0.925	0.353	28.46	0.882	0.363	26.95	0.907	0.419
MARS [57]	31.57	0.924	0.120	24.19	0.819	0.222	26.12	0.795	0.214
NeRFacto [48]	24.54	0.873	0.301	20.25	0.772	0.381	21.97	0.685	0.419
EmerNeRF [60]	31.44	0.917	0.125	24.60	0.839	0.206	25.83	0.780	0.226
- Gaussian-based									
3D-GS [26]	23.35	0.878	0.384	20.52	0.817	0.381	20.53	0.806	0.458
PVG [13]	28.27	0.914	0.312	24.41	0.853	0.311	24.13	0.744	0.411
GaussianPro [16]	23.54	0.881	0.381	20.52	0.817	0.381	20.45	0.706	0.457
DC-Gaussian [53]	28.27	0.920	0.294	25.21	0.879	0.321	24.19	0.773	0.392

Table 7. Results on our proposed benchmark with the 4-meter offset with 1 camera.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	23.20	0.854	0.401	19.44	0.754	0.477	19.49	0.764	0.413
UC-NeRF [15]	26.12	0.903	0.410	24.34	0.872	0.400	27.23	0.888	0.373
MARS [57]	23.45	0.854	0.202	24.41	0.848	0.177	25.00	0.843	0.196
NeRFacto [48]	21.70	0.847	0.366	20.14	0.767	0.405	19.15	0.756	0.402
EmerNeRF [60]	23.89	0.876	0.184	24.31	0.843	0.181	25.39	0.862	0.180
- Gaussian-based									
3D-GS [26]	19.57	0.842	0.424	17.42	0.755	0.449	18.01	0.773	0.414
PVG [13]	22.59	0.893	0.336	22.74	0.862	0.327	23.65	0.866	0.310
GaussianPro [16]	19.50	0.840	0.420	17.39	0.753	0.447	17.96	0.767	0.411
DC-Gaussian [53]	23.56	0.897	0.329	22.35	0.835	0.362	22.06	0.854	0.338

	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	22.59	0.817	0.409	19.95	0.764	0.422	21.02	0.652	0.533
UC-NeRF [15]	27.96	0.890	0.409	26.27	0.852	0.397	26.07	0.755	0.473
MARS [57]	27.39	0.888	0.175	23.37	0.813	0.239	24.36	0.749	0.258
NeRFacto [48]	20.90	0.815	0.419	19.49	0.760	0.414	20.39	0.637	0.489
EmerNeRF [60]	27.35	0.886	0.177	23.66	0.814	0.227	24.17	0.739	0.266
- Gaussian-based									
3D-GS [26]	20.39	0.835	0.451	19.01	0.795	0.426	18.42	0.648	0.515
PVG [13]	25.11	0.891	0.359	22.90	0.831	0.338	22.00	0.700	0.447
GaussianPro [16]	20.39	0.835	0.451	18.91	0.783	0.43	18.33	0.647	0.511
DC-Gaussian [53]	24.26	0.893	0.345	23.18	0.850	0.356	21.99	0.737	0.427

Table 8. Results on our proposed benchmark with the 0-meter offset with 3 cameras.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	29.56	0.906	0.250	30.95	0.907	0.236	31.13	0.911	0.217
UC-NeRF [15]	35.43	0.936	0.375	31.40	0.904	0.365	31.95	0.921	0.344
MARS [57]	32.52	0.890	0.154	30.99	0.903	0.133	31.99	0.889	0.136
NeRFacto [48]	30.75	0.902	0.198	29.64	0.894	0.238	30.83	0.907	0.214
EmerNeRF [60]	33.25	0.926	0.125	30.84	0.896	0.139	32.56	0.917	0.113
- Gaussian-based									
3D-GS [26]	31.35	0.935	0.310	27.09	0.883	0.315	27.23	0.892	0.278
PVG [13]	34.59	0.955	0.258	33.77	0.948	0.223	34.08	0.948	0.219
GaussianPro [16]	31.07	0.934	0.211	27.23	0.883	0.215	26.97	0.888	0.283
DC-Gaussian [53]	33.95	0.948	0.258	29.19	0.910	0.262	28.58	0.914	0.251
	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	29.73	0.899	0.222	26.69	0.871	0.278	27.39	0.836	0.366
UC-NeRF [15]	36.29	0.942	0.332	32.51	0.909	0.340	31.15	0.859	0.403
MARS [57]	35.72	0.942	0.111	28.28	0.856	0.170	28.73	0.843	0.203
NeRFacto [48]	29.65	0.887	0.268	27.46	0.898	0.277	28.00	0.853	0.281
EmerNeRF [60]	35.46	0.929	0.121	28.60	0.872	0.157	28.32	0.822	0.219
- Gaussian-based									
3D-GS [26]	30.62	0.924	0.320	31.25	0.930	0.321	30.91	0.922	0.326
PVG [13]	35.41	0.956	0.243	30.79	0.916	0.263	31.36	0.874	0.329
GaussianPro [16]	30.45	0.923	0.322	25.69	0.871	0.328	27.39	0.836	0.366
DC-Gaussian [53]	32.74	0.947	0.251	27.91	0.898	0.275	29.01	0.857	0.327

Table 9. Results on our proposed benchmark with the 1-meter offset with 3 cameras.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	22.51	0.881	0.361	20.41	0.794	0.358	23.62	0.892	0.263
UC-NeRF [15]	29.60	0.925	0.355	28.68	0.894	0.338	30.11	0.911	0.336
MARS [57]	28.58	0.888	0.149	28.81	0.884	0.144	30.10	0.907	0.129
NeRFacto [48]	23.60	0.891	0.311	22.70	0.862	0.272	21.98	0.826	0.286
EmerNeRF [60]	29.12	0.915	0.127	28.68	0.877	0.149	30.448	0.909	0.115
- Gaussian-based									
3D-GS [26]	23.90	0.905	0.340	20.67	0.818	0.382	22.33	0.846	0.331
PVG [13]	25.48	0.92	0.290	25.82	0.884	0.295	28.11	0.907	0.270
GaussianPro [16]	23.37	0.902	0.349	20.81	0.817	0.353	22.19	0.0.740	0.424
DC-Gaussian [53]	28.01	0.928	0.288	25.08	0.869	0.315	25.92	0.891	0.289
	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	26.60	0.897	0.287	23.15	0.854	0.321	24.81	0.763	0.472
UC-NeRF [15]	34.10	0.938	0.338	29.42	0.890	0.353	28.51	0.820	0.410
MARS [57]	33.96	0.930	0.110	26.42	0.832	0.190	27.84	0.807	0.220
NeRFacto [48]	24.08	0.869	0.300	20.50	0.793	0.351	24.58	0.861	0.319
EmerNeRF [60]	33.83	0.923	0.115	26.83	0.852	0.174	27.55	0.792	0.232
- Gaussian-based									
3D-GS [26]	23.19	0.896	0.377	19.66	0.827	0.385	22.72	0.757	0.336
PVG [13]	29.08	0.926	0.305	25.54	0.864	0.319	25.69	0.769	0.428
GaussianPro [16]	23.45	0.896	0.379	19.50	0.819	0.389	22.69	0.757	0.443
DC-Gaussian [53]	30.27	0.935	0.289	25.85	0.878	0.306	25.33	0.794	0.402

Table 10. Results on our proposed benchmark with the 2-meter offset with 3 cameras.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	22.51	0.881	0.361	20.41	0.794	0.358	23.62	0.892	0.263
UC-NeRF [15]	29.60	0.925	0.355	28.68	0.894	0.338	30.11	0.911	0.336
MARS [57]	28.09	0.882	0.161	28.69	0.892	0.132	29.49	0.904	0.131
NeRFacto [48]	23.57	0.842	0.323	19.97	0.769	0.393	23.86	0.728	0.408
EmerNeRF [60]	28.63	0.909	0.139	28.56	0.885	0.137	29.84	0.906	0.117
- Gaussian-based									
3D-GS [26]	22.43	0.891	0.369	22.21	0.812	0.397	21.15	0.825	0.355
PVG [13]	24.31	0.912	0.313	24.74	0.885	0.298	26.63	0.894	0.289
GaussianPro [16]	22.39	0.891	0.368	20.06	0.812	0.396	20.76	0.819	0.362
DC-Gaussian [53]	26.98	0.921	0.303	24.02	0.871	0.315	24.51	0.878	0.306
	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	26.60	0.897	0.287	23.15	0.854	0.321	24.81	0.763	0.472
UC-NeRF [15]	34.10	0.938	0.338	29.42	0.890	0.353	28.51	0.820	0.410
MARS [57]	31.95	0.919	0.131	25.77	0.821	0.205	27.22	0.798	0.228
NeRFacto [48]	23.42	0.878	0.337	22.52	0.853	0.282	20.22	0.785	0.343
EmerNeRF [60]	31.82	0.912	0.136	26.18	0.841	0.189	26.93	0.783	0.240
- Gaussian-based									
3D-GS [26]	21.80	0.876	0.410	19.11	0.810	0.412	21.33	0.729	0.468
PVG [13]	28.10	0.917	0.325	24.65	0.855	0.333	24.58	0.756	0.435
GaussianPro [16]	21.69	0.876	0.408	18.70	0.804	0.426	21.48	0.732	0.465
DC-Gaussian [53]	28.25	0.919	0.313	25.12	0.862	0.324	24.32	0.778	0.411

Table 11. Results on our proposed benchmark with the 4-meter offset with 3 cameras.

	Scene001			Scene002			Scene003		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	22.04	0.866	0.414	19.95	0.781	0.405	18.33	0.745	0.412
UC-NeRF [15]	26.75	0.904	0.422	26.09	0.867	0.413	27.66	0.887	0.393
MARS [57]	25.53	0.871	0.185	26.97	0.868	0.158	28.04	0.870	0.155
NeRFacto [48]	22.41	0.862	0.367	21.51	0.817	0.326	18.33	0.745	0.412
EmerNeRF [60]	25.97	0.893	0.167	26.87	0.863	0.162	28.43	0.889	0.139
- Gaussian-based									
3D-GS [26]	20.73	0.872	0.402	18.86	0.786	0.436	19.24	0.797	0.393
PVG [13]	22.40	0.898	0.338	22.56	0.860	0.331	23.96	0.873	0.317
GaussianPro [16]	22.04	0.866	0.414	19.95	0.781	0.405	18.33	0.745	0.412
DC-Gaussian [53]	25.00	0.908	0.329	22.07	0.851	0.347	22.61	0.867	0.333
	Scene004			Scene005			Scene006		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- NeRF-based									
Instant-NGP [38]	23.23	0.835	0.380	19.70	0.771	0.434	23.03	0.700	0.522
UC-NeRF [15]	27.83	0.886	0.416	26.79	0.852	0.399	26.11	0.759	0.482
MARS [57]	28.38	0.888	0.177	23.90	0.818	0.227	26.10	0.768	0.256
NeRFacto [48]	22.38	0.815	0.382	19.04	0.756	0.427	22.53	0.681	0.446
EmerNeRF [60]	28.34	0.886	0.179	24.19	0.819	0.215	25.91	0.758	0.264
- Gaussian-based									
3D-GS [26]	20.31	0.845	0.460	17.90	0.785	0.453	19.87	0.693	0.499
PVG [13]	25.11	0.897	0.365	23.14	0.840	0.352	22.95	0.724	0.458
GaussianPro [16]	23.23	0.835	0.380	19.70	0.771	0.434	23.03	0.700	0.522
DC-Gaussian [53]	23.35	0.848	0.348	25.12	0.896	0.355	23.02	0.788	0.431

Table 12. Results on our proposed dataset with the different offsets using *left-front, front, right-front* cameras.

Method	*w/o* Offset			Offset-1m			Offset-2m			Offset-4m		
	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS	↑PSNR	↑SSIM	↓LPIPS
- *NeRF-based*												
Instant-NGP [38]	29.24	0.888	0.262	23.52	0.847	0.344	22.39	0.815	0.382	21.05	0.783	0.428
UC-NeRF [15]	33.12	0.912	0.360	30.07	0.896	0.355	28.81	0.881	0.373	26.87	0.870	0.421
MARS [57]	31.37	0.887	0.151	29.28	0.874	0.157	28.54	0.869	0.165	26.49	0.847	0.193
NeRFacto [48]	29.39	0.890	0.246	22.91	0.850	0.307	22.26	0.809	0.348	21.03	0.779	0.393
EmerNeRF [60]	31.51	0.894	0.146	29.41	0.878	0.152	28.66	0.873	0.160	26.62	0.851	0.188
- *Gaussian-based*												
3DGS [26]	29.74	0.914	0.312	22.08	0.842	0.359	21.34	0.824	0.402	19.47	0.796	0.441
PVG [13]	33.33	0.933	0.256	26.62	0.878	0.318	25.50	0.870	0.332	23.35	0.849	0.360
GaussianPro [16]	28.13	0.889	0.321	21.90	0.839	0.379	20.85	0.822	0.404	19.36	0.795	0.443
DC-Gaussian [53]	30.23	0.912	0.271	26.74	0.883	0.315	25.53	0.8715	0.329	23.53	0.860	0.357