

WarpRF: Multi-View Consistency for Training-Free Uncertainty Quantification and Applications in Radiance Fields

This document provides additional material for our WACV submission “WarpRF: Multi-View Consistency for Training-Free Uncertainty Quantification and Applications in Radiance Fields”. We provide results of uncertainty quantification for each scene in Sec. 1, report results of active view selection for each scene in Sec. 2, provide additional qualitative results in Sec. 3, present the efficiency analysis in Sec. 4 and ablate the design of our image-level uncertainty in Sec. 5.

1. Per-Scene Uncertainty Quantifications

Uncertainty quantification experiments are performed on the ETH3D and ScanNet++ datasets. For the ETH3D dataset, we use all scenes of the high-resolution set, having 14 to 76 images. We train 3DGS/NeRF on every other image and use the rest to evaluate uncertainty. For ScanNet++, we select the ‘27dd4da69e’, ‘3864514494’, ‘5eb31827b7’, ‘8b5caf3398’ ‘8d563fc2cc’, and ‘b20a261fdf’ scenes. Using the first 40 images of each scene, we train the radiance field on every other image and use the remaining ones for AUSE evaluation.

Tables 1 and 2 present per-scene AUSE evaluation results for each scene in ScanNet++ and ETH3D, respectively.

2. Per-Scene Active View Selection Results

Tables 3 to 5 present the per-scene quantitative results of active view selection on the Mip-NeRF360 dataset, evaluated using PSNR, SSIM, and LPIPS metrics. Similarly, Tables 6 to 8 report results for NeRF Synthetic scenes with 20 views, while Tables 9 to 11 provide results for 10-view settings. Finally, the results for Tanks and Temples dataset are provided in Tab. 12.

3. Additional Qualitative Results

Figures 1 and 2 present qualitative results for active view selection using 10 views, while Figures 3 and 4 show results using 20 views on the NeRF Synthetic dataset.

Method	27dd4da69e	3864514494	5eb31827b7	8b5caf3398	8d563fc2cc	b20a261fdf
NeRF + BayesRays	0.456	0.407	0.489	0.498	0.368	0.412
NeRF + WarpRF (ours)	0.487	0.365	0.485	0.418	0.345	0.437
3DGS + Manifold	0.569	0.474	0.514	0.466	0.505	0.533
3DGS + FisherRF	0.379	0.358	0.293	0.297	0.414	0.386
3DGS + WarpRF (ours)	0.364	0.300	0.291	0.331	0.308	0.430

Table 1. AUSE evaluation ↓ on ScanNet++ scenes.

Method	courtyard	delivery area	electro	facade	kicker	meadow	office	pipes	playground	relief	relief 2	terrace	terrains
NeRF + BayesRays	0.133	0.422	0.253	0.245	0.267	0.194	0.443	0.328	0.185	0.300	0.269	0.214	0.366
NeRF + WarpRF (ours)	0.130	0.333	0.394	0.205	0.288	0.104	0.294	0.195	0.229	0.217	0.271	0.292	0.301
3DGS + Manifold	0.539	0.524	0.514	0.509	0.532	0.474	0.531	0.432	0.606	0.533	0.550	0.520	0.524
3DGS + FisherRF	0.214	0.192	0.257	0.320	0.185	0.389	0.225	0.424	0.351	0.336	0.403	0.322	0.212
3DGS + WarpRF (ours)	0.196	0.173	0.227	0.211	0.221	0.253	0.192	0.243	0.352	0.193	0.264	0.243	0.180

Table 2. AUSE evaluation ↓ on ETH3D scenes.

Method	bicycle	bonsai	counter	flowers	garden	kitchen	room	stump	treehill
3DGS + Manifold	18.345	22.597	22.259	15.394	22.704	24.002	22.987	20.369	17.225
3DGS + FisherRF	18.006	22.597	21.682	14.847	22.017	23.559	22.552	20.058	17.341
3DGS + WarpRF (ours)	17.850	23.577	22.194	15.430	23.184	23.848	22.792	20.349	17.212

Table 3. PSNR \uparrow results for active view selection on Mip-NeRF360.

Method	bicycle	bonsai	counter	flowers	garden	kitchen	room	stump	treehill
3DGS + Manifold	0.415	0.809	0.776	0.321	0.681	0.836	0.796	0.476	0.436
3DGS + FisherRF	0.386	0.808	0.763	0.315	0.643	0.805	0.783	0.454	0.457
3DGS + WarpRF (ours)	0.376	0.830	0.775	0.333	0.706	0.829	0.808	0.480	0.418

Table 4. SSIM \uparrow results for active view selection on Mip-NeRF360.

Method	bicycle	bonsai	counter	flowers	garden	kitchen	room	stump	treehill
3DGS + Manifold	0.436	0.294	0.286	0.492	0.236	0.203	0.303	0.424	0.456
3DGS + FisherRF	0.447	0.295	0.300	0.499	0.261	0.232	0.318	0.440	0.444
3DGS + WarpRF (ours)	0.460	0.281	0.290	0.488	0.225	0.209	0.296	0.424	0.467

Table 5. LPIPS \downarrow results for active view selection on Mip-NeRF360.

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	30.920	23.228	30.734	29.464	28.864	25.887	30.387	25.699
3DGS + FisherRF	32.277	23.697	30.411	33.636	30.423	25.689	31.343	26.620
3DGS + WarpRF (ours)	32.150	23.638	30.731	33.344	30.691	26.254	30.571	26.772
SVRaster + Random	29.760	21.059	22.605	29.676	27.856	24.063	30.279	25.759
SVRaster + Farthest	30.394	21.462	22.736	30.613	27.722	22.849	30.620	26.040
SVRaster + WarpRF (ours)	30.519	22.156	24.034	31.454	28.536	25.283	30.973	26.325

Table 6. PSNR \uparrow results for active view selection on NeRF Synthetic with 20 views.

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	0.965	0.919	0.972	0.951	0.931	0.926	0.976	0.841
3DGS + FisherRF	0.974	0.927	0.972	0.969	0.951	0.923	0.980	0.844
3DGS + WarpRF (ours)	0.975	0.926	0.972	0.971	0.955	0.931	0.977	0.846
SVRaster + Random	0.949	0.871	0.902	0.933	0.921	0.882	0.973	0.810
SVRaster + Farthest	0.959	0.886	0.908	0.947	0.930	0.872	0.976	0.816
SVRaster + WarpRF (ours)	0.959	0.892	0.926	0.950	0.935	0.905	0.976	0.819

Table 7. SSIM \uparrow results for active view selection on NeRF Synthetic with 20 views.

4. Efficiency Analysis

In this section, we compare the computational efficiency of our method against baseline approaches for the active view selection task using 3DGS on the Mip-NeRF360 dataset (Sec. 4.2 of the main paper). Tab 13 reports the average memory usage and runtime across all scenes. Similar to FisherRF and Manifold, our method incurs only minimal overhead relative to the original 3DGS. Moreover, it requires less memory than both FisherRF and Manifold, while achieving faster runtime than FisherRF and nearly matching Manifold. We also report the average time each method needs to select the next best view (NBV Time), which involves computing uncertainty for every candidate view.

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	0.029	0.059	0.022	0.057	0.060	0.057	0.017	0.154
3DGS + FisherRF	0.021	0.054	0.023	0.038	0.043	0.061	0.014	0.145
3DGS + WarpRF (ours)	0.021	0.054	0.022	0.033	0.041	0.055	0.016	0.143
SVRaster + Random	0.052	0.124	0.082	0.085	0.069	0.100	0.023	0.167
SVRaster + Farthest	0.041	0.106	0.071	0.068	0.062	0.107	0.021	0.157
SVRaster + WarpRF (ours)	0.042	0.101	0.059	0.065	0.057	0.085	0.020	0.155

Table 8. **LPIPS ↓ results for active view selection on NeRF Synthetic with 20 views.**

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	27.006	21.220	26.152	25.240	24.330	21.898	26.825	20.600
3DGS + FisherRF	26.766	20.975	25.817	29.456	25.162	21.474	28.427	21.468
3DGS + WarpRF (ours)	27.875	21.269	26.859	27.971	24.842	21.682	28.446	20.714

Table 9. **PSNR ↑ results for active view selection on NeRF Synthetic with 10 views.**

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	0.929	0.887	0.943	0.916	0.873	0.879	0.957	0.746
3DGS + FisherRF	0.928	0.886	0.941	0.940	0.882	0.870	0.966	0.757
3DGS + WarpRF (ours)	0.940	0.891	0.949	0.939	0.891	0.871	0.966	0.753

Table 10. **SSIM ↑ results for active view selection on NeRF Synthetic with 10 views.**

Method	chair	drums	figus	hotdog	lego	materials	mic	ship
3DGS + Manifold	0.057	0.085	0.047	0.098	0.108	0.105	0.035	0.217
3DGS + FisherRF	0.056	0.090	0.048	0.065	0.099	0.111	0.025	0.199
3DGS + WarpRF (ours)	0.048	0.084	0.043	0.066	0.096	0.108	0.025	0.204

Table 11. **LPIPS ↓ results for active view selection on NeRF Synthetic with 10 views.**

Method	PSNR ↑		SSIM ↑		LPIPS ↓	
	Train	Truck	Train	Truck	Train	Truck
SVRaster + Random	13.449	19.818	0.437	0.710	0.477	0.239
SVRaster + Farthest	14.594	20.353	0.482	0.741	0.434	0.218
SVRaster + WarpRF (ours)	14.803	20.746	0.493	0.756	0.432	0.204

Table 12. **Active view selection results on Tanks and Temples dataset.**

Method	Runtime	Memory	NBV Time
3DGS + Random	18m32s	8.73 GB	≈0
3DGS + Manifold	20m13s	9.59 GB	2.36s
3DGS + FisherRF	21m43s	9.00 GB	11.62s
3DGS + WarpRF	20m02s	8.84 GB	5.69s

Table 13. **Efficiency analysis for Active View Selection on Mip-NeRF360.**

5. Ablation Study

We present an ablation study on the design of our image-level uncertainty. As outlined in Sec. 3.2.2 of the main paper, a straightforward approach is to sum the pixel-wise uncertainties derived from depth differences. In Tab. 14, we compare this baseline with two variants: (i) using the minimum difference across training views instead of averaging and (ii) replacing depth with RGB differences. The active view selection experiment is conducted on the Mip-NeRF360 dataset following the setup in Sec. 4.2 of the main paper. Results show that employing the minimum difference yields a substantial performance gain, and further improvements are achieved when using the RGB modality instead of depth.

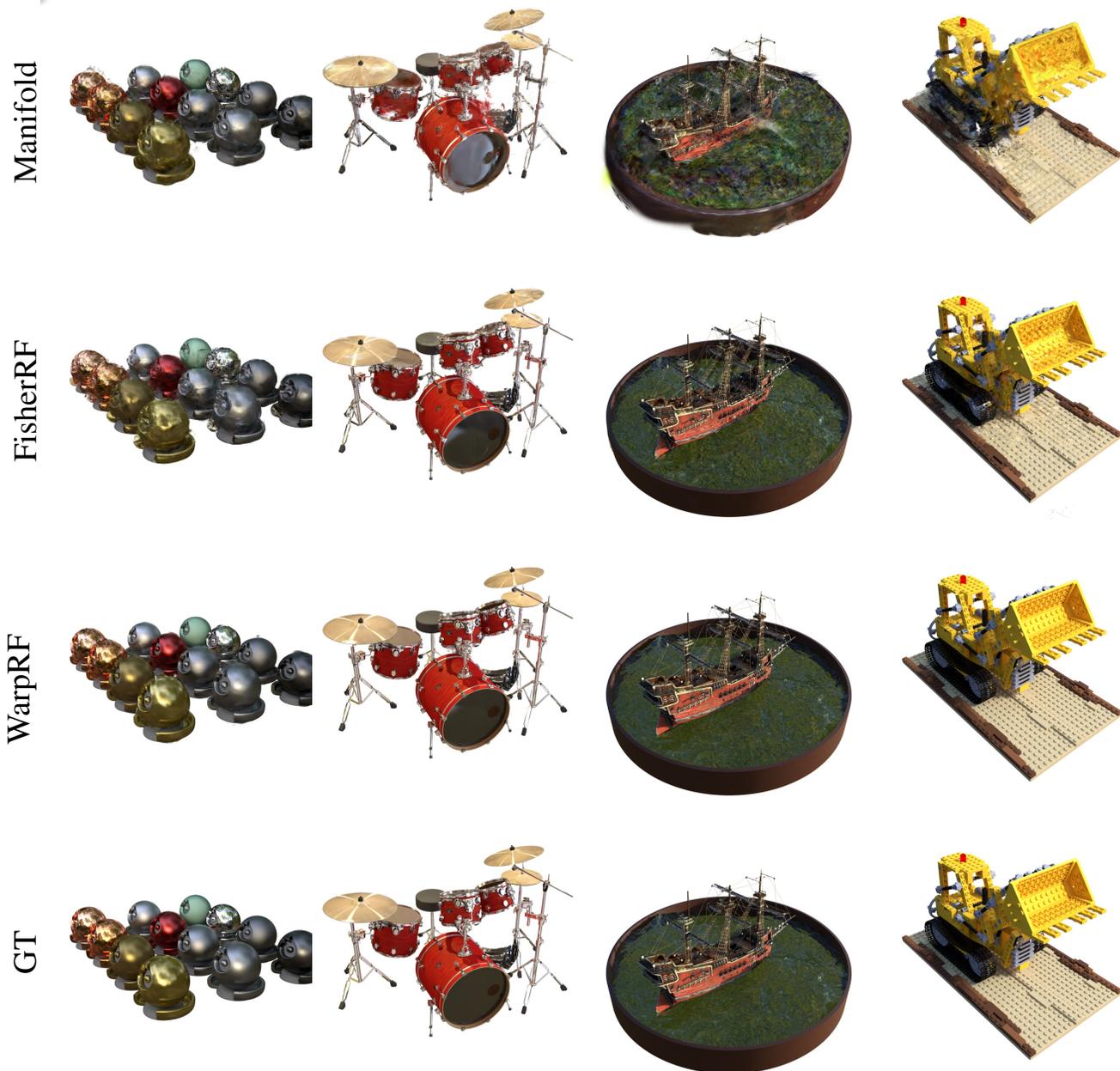


Figure 1. **Qualitative results of active camera selection on NeRF Synthetic dataset with 10 views.** From top to bottom, we collect rendered results by 3DGS trained with active view selection being performed by Manifold, FisherRF and WarpRF, followed by real images.

Aggregation	Modality	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
Average	Depth	19.086	0.5603	0.3931
Average	RGB	19.035	0.5604	0.3903
Minimum	Depth	20.086	0.5860	0.3719
Minimum	RGB	20.715	0.6174	0.3487

Table 14. **Ablation study on the design of the image-level uncertainty.**

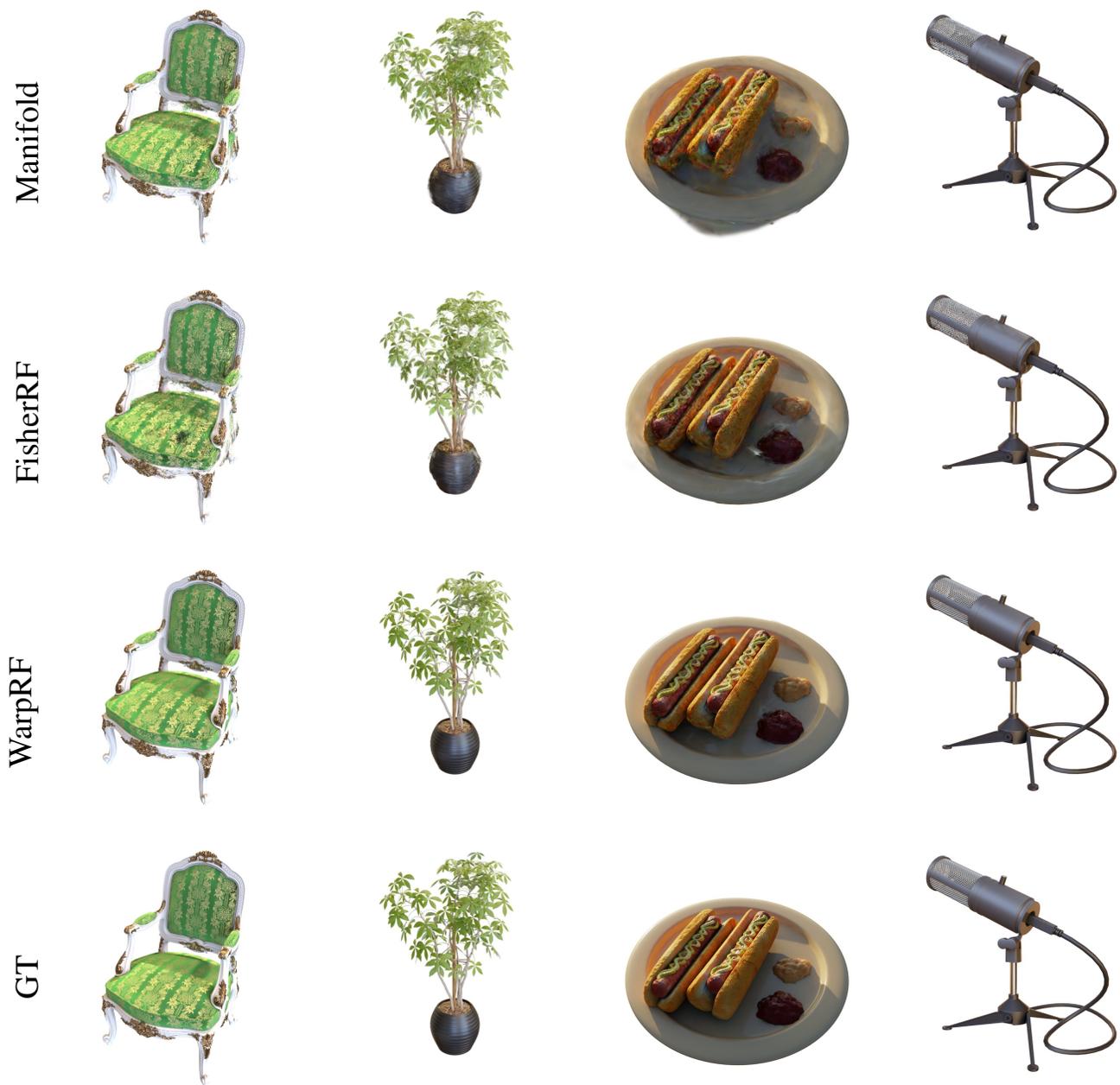


Figure 2. **Additional qualitative results of active camera selection on NeRF Synthetic dataset with 10 views.** From top to bottom, we collect rendered results by 3DGS trained with active view selection being performed by Manifold, FisherRF and WarpRF, followed by the real images.

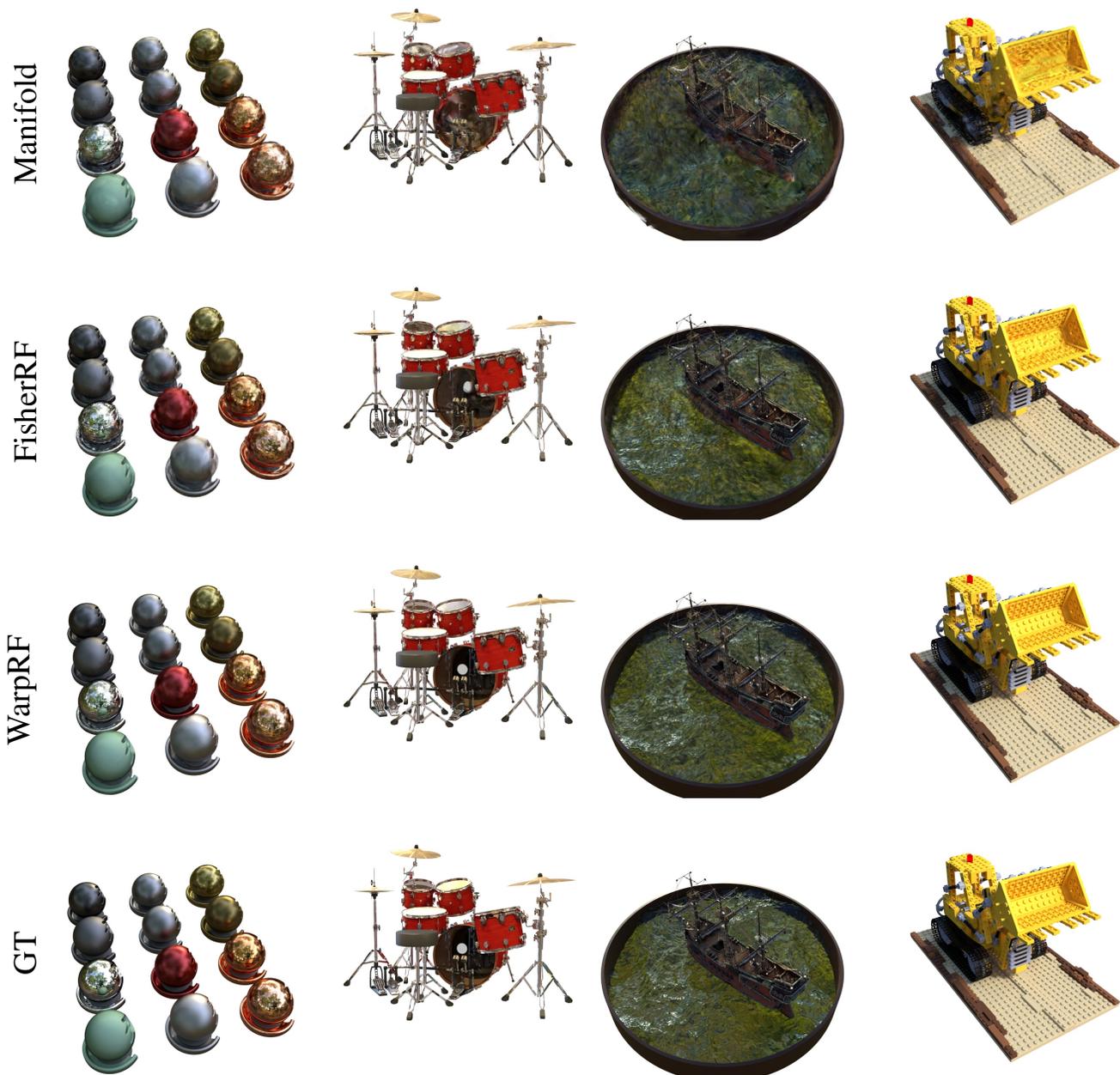


Figure 3. **Qualitative results of active camera selection on NeRF Synthetic dataset with 20 views.** From top to bottom, we collect rendered results by 3DGS trained with active view selection being performed by Manifold, FisherRF and WarpRF, followed by the real images.

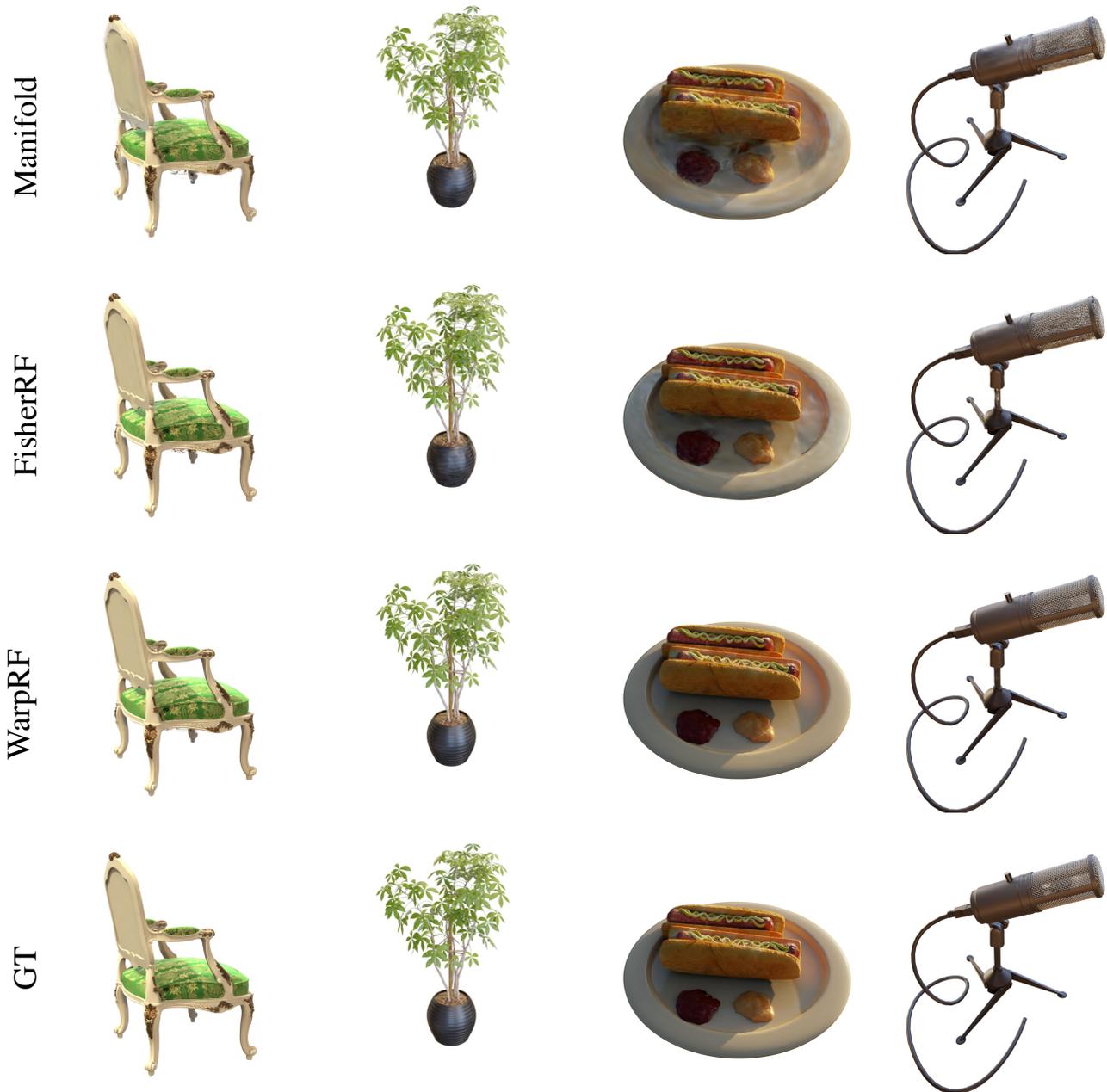


Figure 4. **Additional qualitative results of active camera selection on NeRF Synthetic dataset with 20 views.** From top to bottom, we collect rendered results by 3DGS trained with active view selection being performed by Manifold, FisherRF and WarpRF, followed by the real images.