

VI. SUPPLEMENTAL MATERIALS

In this section, we use the ablation study to evaluate the compression efficiency of *NeRfCompressor* for static and dynamic scene datasets. We represent the model size of *NeRfCompressor* with different qp values in table IV. Figure 1 illustrates the overall design of *NeRfCompressor*. We present the details of each component next. The pipeline 4 in briefly explores the *NeRfCompressor* compression details. The total number of component features is $3(R_\sigma + R_c)$, where R_σ and R_c are 16 and 48. Therefore, there are 192 components in total. In table V rows 6-13, we also represent the sample result of the partial NeRF-synthetic dataset with different numbers of components, including '192', '120', '84', and '48'. We followed the components in TensorRF [5] for numbers of components in '192', '120', and '48', and the component numbers of '84' consist of R_σ and R_c , which are 16 and 12.

TABLE IV: **Model sizes after intra-scene compression (static scenes)** "density comp." and "appearance comp." represent the In the table, " qp " represents the quantization parameter we provided to the video codec during compression. $qp=0$ indicates lossless compression. The "model size" encompasses compressed model components and other essential information. The table shows a comparison of various model sizes after applying video codec compression to our quantized model. The model size combined with compressed plane size, grayscale vector size and additional information of the models(1014KB)

qp	density comp.	appearance comp.	model size
0	1.5MB	7.3MB	9.8MB
5	1.4MB	6.0MB	8.6MB
10	1.1MB	4.8MB	6.9MB
15	802KB	3.7MB	5.5MB
20	545KB	2.6MB	4.1MB
25	326KB	1.7MB	3.0MB
30	189KB	1.1MB	2.3MB
35	101KB	590KB	1.7MB
40	60KB	327KB	1.4MB

The second and third columns of the table indicate the output video size after quantizing the density component and appearance components, respectively, compositing them as frames in YUV444 pixel format, and compressing the videos using video codec in *NeRfCompressor*.

The last column indicates the size of the saved model file. This file contains all the necessary parameters saved after training and the component videos generated after quantization.

The presented results show the efficacy of applying intra-scene lossy compression on the static scene model. The model size was reduced to a minimum of 1.3MB from the original 68.8MB after compression, a significant reduction.

Table V shows the PSNR results for static scenes. Here, "192", "120", "84", and "48" represent the total number of feature components used (i.e., $3R_\sigma + 3R_c$). Note the JPEG compression result is based on the TensorRF VM-192 model. As the number of components decreases, the PSNR also decreases. The "jpeg10", "jpeg30", and "jpeg60" rows represent a baseline comparison where we directly leverage

JPEG [20] image compression for compressing the grayscale images. The numbers (i.e., 10, 30, and 60) represent the quantization parameter (qp) used in JPEG compression. A lower qp retains more details and leads to higher quality, with "jpeg10" resulting in the highest quality among the three. From our experiment results, the losslessly-compressed model size of *NeRfCompressor* is significantly smaller than the original TensorRF model. For instance, following the quantization process, our "lego" model occupies only 4.75 MB (compared to the original VM-48 model size of 18 MB), while the PSNR remains at 35.72 dB. When the model components are compressed as grayscale images using JPEG with $qp = 60$, the model size is slightly smaller than the compressed size of our model *NeRfCompressor*. However, the PSNR is significantly lower than the uncompressed TensorRF model(28.51 dB vs. 36.46 dB). We note that we only use the model size compressed by JPEG with $qp = 60$ as it is similar to the compressed model size of *NeRfCompressor*. While decreasing the qp of JPEG compression can improve the PSNR results, it results in larger compressed sizes, and the best results ("jpeg10") still underperform our *NeRfCompressor*.

A. Samples of static scene compression

Figure 5 further visualizes the reconstructed results step by step compared to the original and other methods. The figure illustrates the example results of static scenes from the synthetic-NeRF dataset. We present the density and appearance components after applying the quantization method. The grey image represents the density and appearance component of each object. These components describe the positional relationship of the detailed part of the object.

B. Quantitative results of NeRF-Synthetic dataset

In this section, Table VI explores the quantitative rendering result for NeRF-synthetic scenes. We use PSNR, SSIM and LPIPS to measure the quality of the reconstructed model. Table VII represents the model size after we apply *NeRfCompressor* on the NeRF-synthetic dataset.

TABLE V: **PSNR(dB) results of NeRfCompressor**. In this table, "*NeRfCompressor_xxx*" represents the quantized model with varying numbers of components. The term "plane only" indicates that our quantization method was only applied to plane components, not including lines.

	lego	chair	hotdog	drum
TensorRF (VM-192)	36.46	35.76	37.41	26.01
TensorRF-jpeg10	34.70	33.89	35.69	24.33
TensorRF-jpeg30	29.61	29.36	31.23	20.12
TensorRF-jpeg60	28.51	28.11	30.42	19.84
<i>NeRfCompressor_192</i>	36.59	35.70	37.65	25.95
<i>NeRfCompressor_120</i>	36.47	35.61	37.51	25.73
<i>NeRfCompressor_84</i>	36.17	35.48	37.20	25.55
<i>NeRfCompressor_48</i>	35.72	35.09	36.95	25.47
<i>NeRfCompressor_192</i> (plane only)	36.59	35.68	37.63	25.96
<i>NeRfCompressor_120</i> (plane only)	36.47	35.60	37.52	25.82
<i>NeRfCompressor_84</i> (plane only)	35.15	35.49	37.21	25.57
<i>NeRfCompressor_48</i> (plane only)	35.75	35.11	37.95	25.48

TABLE VI: Quantitative rendering Result with Lossy intra-scene compression for NeRF-Synthetic scenes. ”qp” donate the different compression rates.

QP value		lego	chair	hotdog	drum	ficus	materials	mic	ship
0	PSNR	36.43	35.71	37.39	25.99	33.99	30.12	34.61	30.77
	SSIM	0.985	0.984	0.983	0.936	0.982	0.952	0.988	0.895
	$LPIPS_{Vgg}$	0.017	0.021	0.030	0.071	0.022	0.058	0.015	0.138
	$LPIPS_{Alex}$	0.007	0.010	0.013	0.050	0.012	0.026	0.009	0.085
10	PSNR	36.42	35.71	37.37	25.98	33.97	30.11	34.59	30.76
	SSIM	0.985	0.984	0.983	0.936	0.982	0.952	0.988	0.895
	$LPIPS_{Vgg}$	0.017	0.021	0.030	0.071	0.022	0.058	0.015	0.138
	$LPIPS_{Alex}$	0.007	0.010	0.013	0.050	0.012	0.026	0.009	0.085
20	PSNR	36.29	35.32	37.18	25.92	33.85	30.09	34.43	30.68
	SSIM	0.984	0.982	0.981	0.930	0.981	0.952	0.987	0.894
	$LPIPS_{Vgg}$	0.020	0.027	0.036	0.083	0.024	0.059	0.017	0.141
	$LPIPS_{Alex}$	0.008	0.016	0.015	0.058	0.013	0.027	0.010	0.087
30	PSNR	35.85	34.74	37.01	25.74	32.60	29.79	33.17	29.86
	SSIM	0.981	0.979	0.979	0.930	0.977	0.949	0.983	0.885
	$LPIPS_{Vgg}$	0.026	0.039	0.043	0.100	0.034	0.065	0.018	0.155
	$LPIPS_{Alex}$	0.010	0.018	0.018	0.061	0.018	0.030	0.031	0.102
40	PSNR	31.23	31.44	34.22	22.73	27.70	27.27	29.28	26.67
	SSIM	0.971	0.965	0.966	0.896	0.955	0.939	0.968	0.842
	$LPIPS_{Vgg}$	0.019	0.063	0.061	0.127	0.058	0.042	0.057	0.211
	$LPIPS_{Alex}$	0.027	0.040	0.041	0.102	0.035	0.077	0.042	0.166

TABLE VII: Quantitative Model Size Result with Lossy Intra-scene Compression for NeRF-Synthetic scenes ”qp” represents the compression rate. ”den.” donates the size of density component plane after compression, while ”app.” represents the size of appearance component plane and ”other” indicate the size of related parameters of MLP network, masks and other necessary information.

qp	lego					chair				
	0	10	20	30	40	0	10	20	30	40
den.	1.5MB	1.1MB	545KB	189KB	60KB	1.2MB	793KB	395KB	147KB	56KB
app.	7.3MB	4.8MB	2.6MB	1.1MB	327KB	7.2MB	4.5MB	2.6MB	1.2MB	406KB
other	1MB	1MB	1MB	1MB	1MB	975KB	975KB	975KB	975KB	975KB
total	9.8MB	6.9MB	4.1MB	2.3MB	1.4MB	9.4MB	6.2MB	3.9MB	2.3MB	1.4MB
qp	hotdog					drum				
	0	10	20	30	40	0	10	20	30	40
den.	1.5MB	1MB	505KB	169KB	55KB	1.7MB	1.2MB	645KB	256KB	84KB
app.	7.5MB	4.9MB	2.5MB	979KB	276KB	6.5MB	4MB	2.2MB	1015KB	357KB
other	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1022KB	1022KB	1022KB	1022KB	1022KB
total	10.1MB	7MB	4.1MB	2.2MB	1.4MB	9.2MB	6.2MB	3.8MB	2.2MB	1.4MB
qp	materials					ship				
	0	10	20	30	40	0	10	20	30	40
den.	2.1MB	1.5MB	901KB	363KB	108KB	1.7MB	1.2MB	668KB	242KB	66KB
app.	7.7MB	5.1MB	2.9MB	1.2MB	379KB	7.2MB	4.7MB	2.8MB	1.2MB	328KB
other	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.2MB	1.2MB	1.2MB	1.2MB	1.2MB
total	10.9MB	7.7MB	4.9MB	2.7MB	1.6MB	10.1MB	7.1MB	4.7MB	2.6MB	1.6MB
qp	mic					ficus				
	0	10	20	30	40	0	10	20	30	40
den.	772KB	519KB	262KB	164KB	102KB	1.8MB	1.2MB	668KB	274KB	112KB
app.	4.5MB	2.3MB	1.2MB	523KB	205KB	6.7MB	3.9MB	2.2MB	988KB	388KB
other	1007KB	1007KB	1007KB	1007KB	1007KB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB
total	6.2MB	3.8MB	2.4MB	1.7MB	206.1MB	9.6MB	6.2MB	4MB	2.3MB	1.6MB

C. Samples of reconstructed scenes for NeRF-synthetic dataset

This section presents re-rendered sample images with varying values of qp , [0,10,20,30,40], for NeRF-synthetic dataset.

D. Quantitative result of model size and reconstruction for forward-facing dataset

In this section, Table VIII explores the quantitative rendering result for forward-facing scenes. We use PSNR, SSIM and LPIPS to measure the quality of the reconstructed model. Table IX represents the model size after we apply *NeRFCompressor* on the LLFF dataset.

E. Samples of reconstructed scenes for Forward-facing dataset

In this section, we display a re-rendered sample image with different values of qp for LLFF dataset.

F. Dynamic Result

The re-rendering sample videos are attached. We include three sub-videos in our sample video, including the first scene of the ’d-lego’ dataset in *NeRFCompressor_{SF}*, and the first scene of the ’d-pig’ dataset in both *NeRFCompressor_{SA}* and *NeRFCompressor_{SF}*. All three videos were reconstructed from $qp = 30$ while using the ffmpeg

TABLE VIII: Quantitative rendering Result with Lossy Intra-scene compression for forward-facing dataset "qp" donate the different compression rate.

QP value		flower	horns	T-rex	Room	Fern	Leaves	Fortress	Orchids
0	PSNR	28.07	28.29	27.55	32.33	25.27	21.30	31.30	19.87
	SSIM	0.857	0.883	0.909	0.952	0.814	0.752	0.895	0.649
	$LPIPS_{Vgg}$	0.178	0.182	0.201	0.167	0.237	0.217	0.149	0.278
	$LPIPS_{Alex}$	0.102	0.103	0.080	0.082	0.155	0.153	0.076	0.201
10	PSNR	28.07	28.29	27.54	32.33	25.27	21.30	31.30	19.87
	SSIM	0.857	0.883	0.909	0.952	0.814	0.752	0.895	0.649
	$LPIPS_{Vgg}$	0.178	0.182	0.201	0.167	0.237	0.217	0.149	0.278
	$LPIPS_{Alex}$	0.102	0.103	0.808	0.082	0.155	0.153	0.076	0.201
20	PSNR	28.03	28.1	27.48	32.25	25.26	21.29	31.27	19.87
	SSIM	0.850	0.879	0.905	0.951	0.814	0.750	0.893	0.648
	$LPIPS_{Vgg}$	0.191	0.199	0.209	0.173	0.245	0.221	0.155	0.280
	$LPIPS_{Alex}$	0.106	0.117	0.084	0.088	0.161	0.157	0.079	0.204
30	PSNR	27.87	27.42	26.46	31.33	25.04	21.15	30.79	19.86
	SSIM	0.856	0.841	0.881	0.939	0.798	0.727	0.871	0.638
	$LPIPS_{Vgg}$	0.202	0.253	0.262	0.119	0.285	0.255	0.203	0.303
	$LPIPS_{Alex}$	0.124	0.173	0.129	0.211	0.203	0.189	0.107	0.228
40	PSNR	22.92	21.19	20.78	28.31	23.02	19.51	28.00	19.21
	SSIM	0.732	0.702	0.741	0.899	0.699	0.575	0.739	0.554
	$LPIPS_{Vgg}$	0.428	0.416	0.455	0.298	0.413	0.398	0.385	0.405
	$LPIPS_{Alex}$	0.277	0.265	0.292	0.223	0.367	0.342	0.286	0.342

TABLE IX: Quantitative rendering Result with Lossy Intra-scene compression for forward-facing dataset "qp" represent the compression rate. "den." donate the size of density component plane after compression, while "app." represent the size of appearance component plane and "other" indicate the size of related parameters of MLP network, masks and other necessary information.

	flower					horns				
qp	0	10	20	30	40	0	10	20	30	40
den.	6.1MB	3.3MB	1.3MB	283KB	123KB	5.5MB	2.8MB	1MB	338KB	116KB
app.	17.7MB	10.1MB	4.3MB	1.3MB	365KB	15.6MB	8.5MB	3.5MB	1.1MB	208KB
other	1MB	1MB	1MB	1MB	1MB	1MB	1MB	1MB	1MB	1MB
total	24.8MB	14.4MB	6.6MB	2.6MB	1.5MB	22.1MB	12.3MB	5.5MB	2.4MB	1.3MB
	trex					orchids				
qp	0	10	20	30	40	0	10	20	30	40
den.	3.9MB	2.2MB	740KB	255KB	91KB	7.3MB	4.2MB	1.8MB	627KB	202KB
app.	12MB	8.9MB	2.6MB	886KB	257KB	22.4MB	13.8MB	6.6MB	2.3MB	609KB
other	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB
total	17MB	12.2MB	4.4MB	2.2MB	1.4MB	30.8MB	19.1MB	9.5MB	4MB	1.9MB
	fern					fortress				
qp	0	10	20	30	40	0	10	20	30	40
den.	5.1MB	2.9MB	1.3MB	455KB	137KB	3.9MB	1.6MB	517KB	162KB	60KB
app.	14.6MB	8.4MB	3.8MB	1.3MB	338KB	14.1MB	7.6MB	3.1MB	1MB	275KB
other	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB
total	20.8MB	12.4MB	6.2MB	2.8MB	1.6MB	19.1MB	10.3MB	4.7MB	2.3MB	1.4MB
	leaves					room				
qp	0	10	20	30	40	0	10	20	30	40
den.	5.4MB	3.2MB	1.5MB	508KB	157KB	3.3MB	1.3MB	451KB	155KB	57KB
app.	16.6MB	10.2MB	5MB	1.8MB	433KB	11.7MB	5.6MB	2.1MB	708KB	221KB
other	971KB	971KB	971KB	971KB	971KB	1.1MB	1.1MB	1.1MB	1.1MB	1.1MB
total	23MB	14.4MB	7.5MB	3.3MB	1.5MB	16.1MB	8MB	3.6MB	1.9MB	1.4MB

compression, named "dlego_CodecNeRF_sample.mp4", "dpig_NeRFCompressor_SF.mp4" and "dpig_NeRFCompressor_SA.mp4". These videos were generated independently for each scene. The PSNR results and model size have been explored in the main paper.

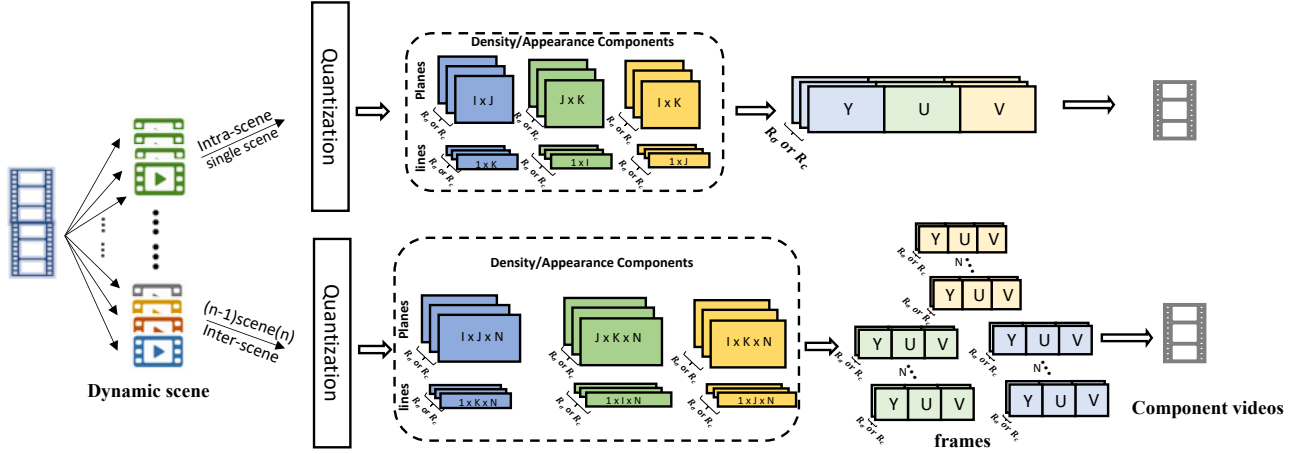


Fig. 4: **Overall design of *NeRFCompressor*** . The top part sketches our proposed idea on intra-scene compression that can be used for static scenes, similar to synthetic-NeRF scenes. The bottom part illustrates our proposed idea for inter-scene compression that can be used for dynamic scenes.

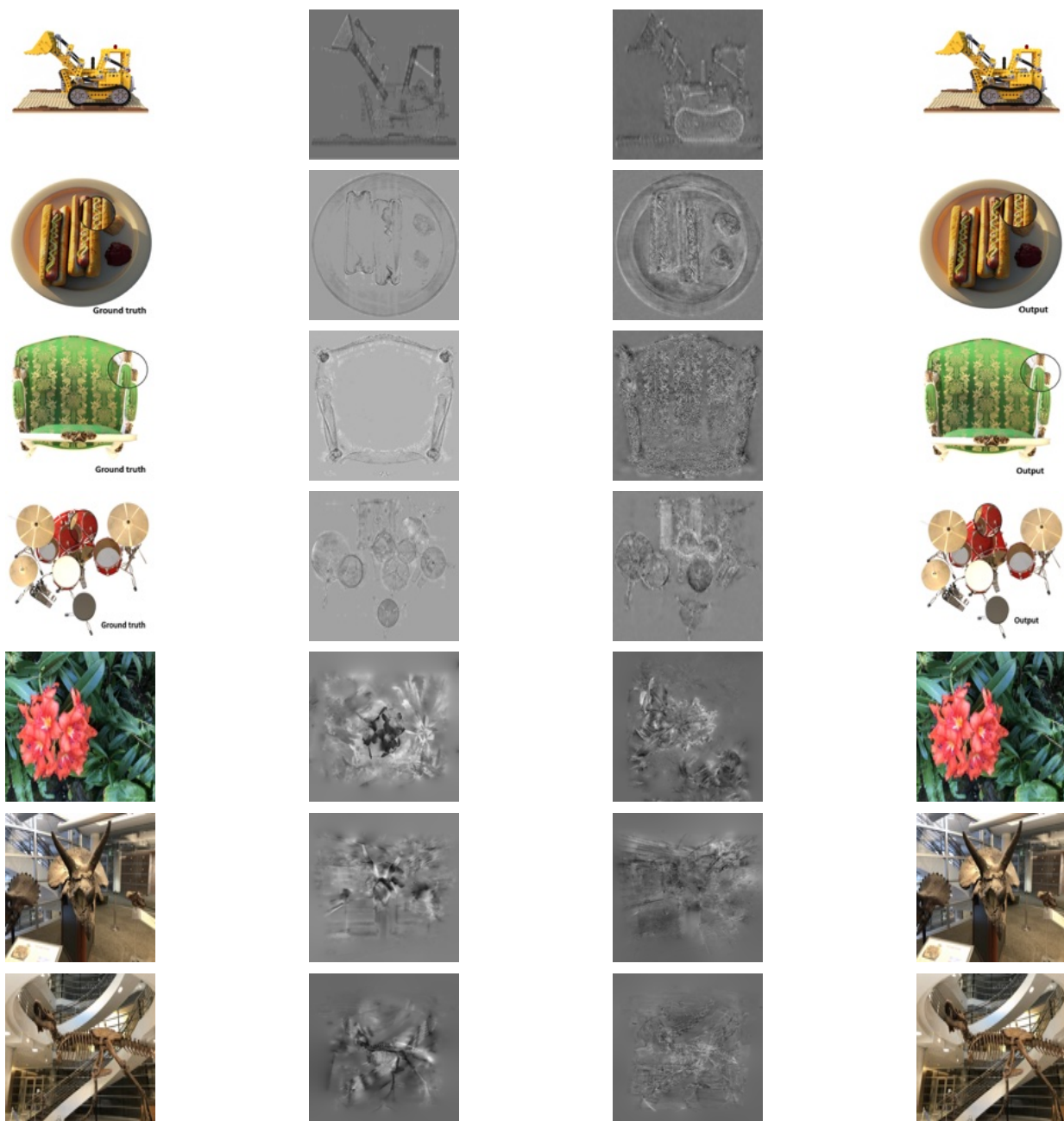


Fig. 5: **Visualization of samples in static scenes.** The column from left to right represent: "Ground truth", "Quantized density plane", "Quantized appearance plane" and "reconstructed result". It shows the sample visualized result from static scenes, from top to bottom, including the NeRF-synthetic dataset("lego", "hotdog", "drum", and "chair") and the Forward-facing dataset("flower", "horns" and "T-rex").

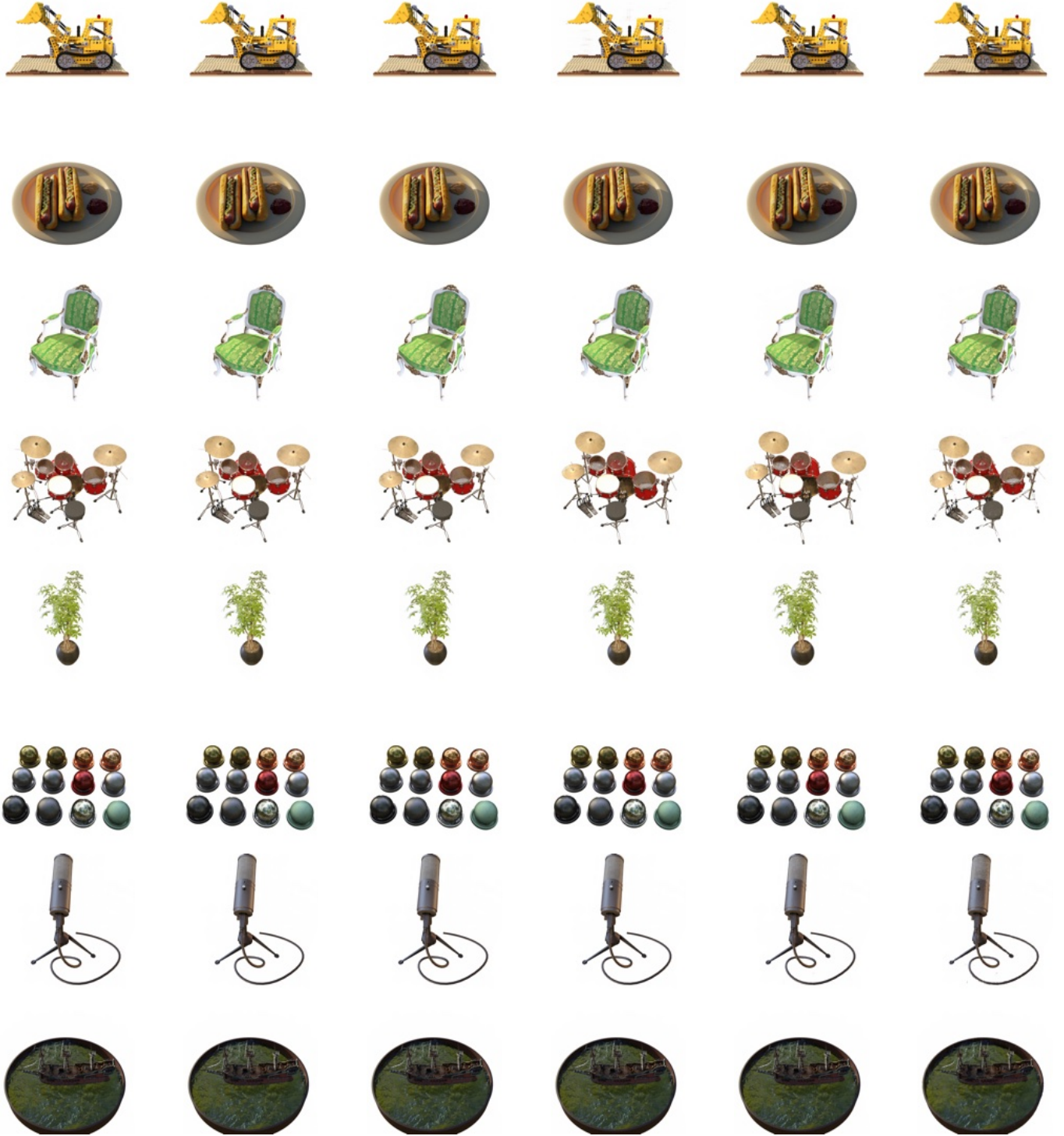


Fig. 6: **Visualization of samples in NeRF-Synthetic dataset.** The row from top to bottom represent eight scenes in NeRF-Synthetic dataset: "lego", "hotdog", "drum", "chair", "ficus", "materials", "mic" and "ship". Left column to Right explores the ground truth and the sample result of different compression rates(qp) with [0,10,20,30,40].

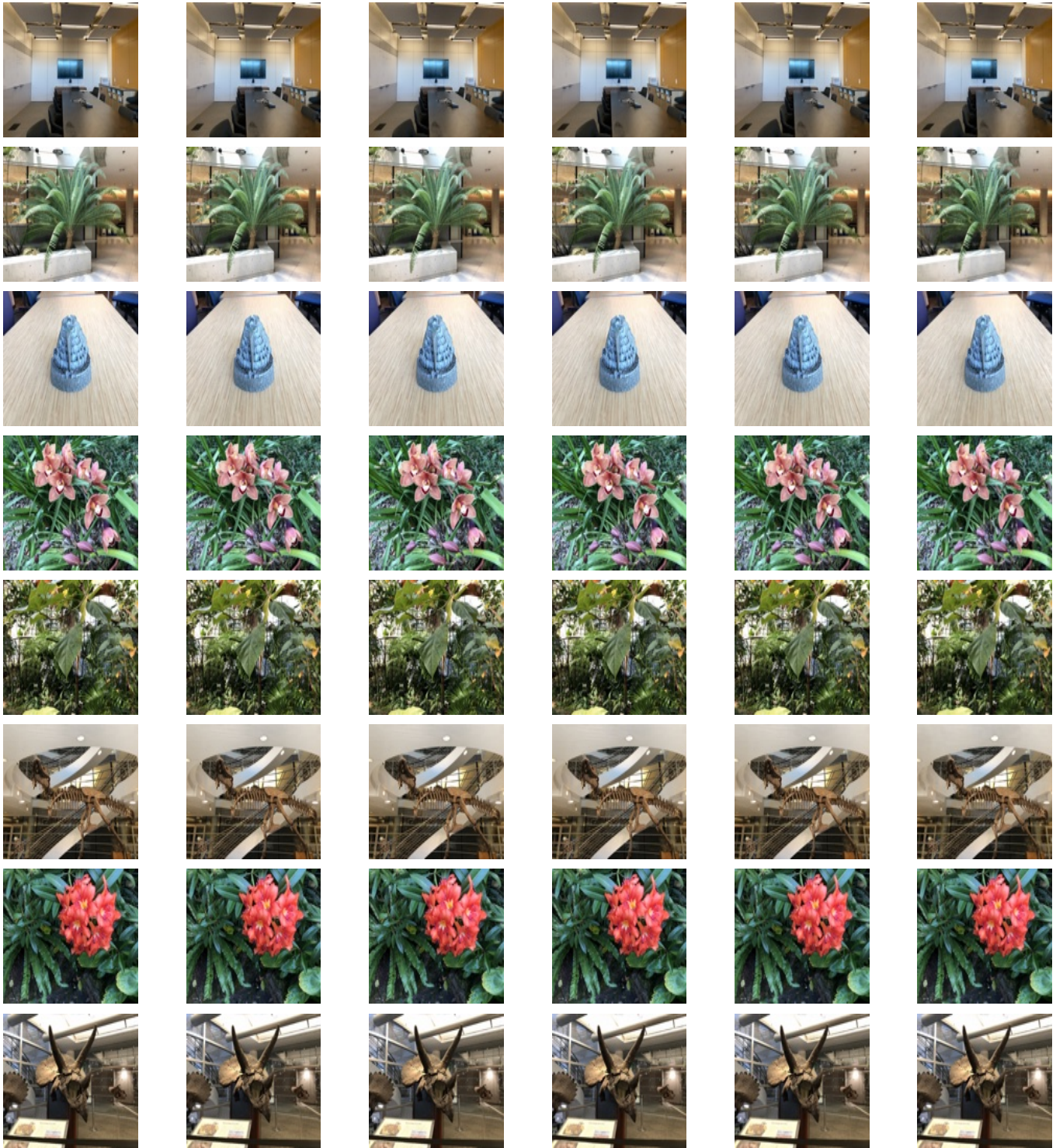


Fig. 7: **Visualization of samples in Forward-facing dataset.** The row from top to bottom represents eight scenes in NeRF-Synthetic dataset: room, fern, fortress, orchids,leaves, trex,flower and horns. The left column to the Right explores the ground truth and the sample result of different compression rates.