Transientangelo: Few-Viewpoint Surface Reconstruction Using Single-Photon Lidar Supplemental Material

Weihan Luo[1]
weihan.luo@mail.utoronto.ca

Anagh Malik[1,2]
anagh@cs.toronto.edu

David B. Lindell[1,2]
lindell@cs.toronto.edu

[1]University of Toronto [2]Vector Institute
https://weihan1.github.io/transientangelo/

S1. Implementation Details

Network implementation. In this section, we provide a detailed description of the network architecture. We implement our implicit surface representation network using the open-sourced instant-nsr-pl [5], which is built on Pytorch Lightning [3]. We combine this with TransientNeRF's [11] neural transient renderer, which extends NerfAcc [8] and Instant-NGP [13].

All network hyperparameters are shared across our methods and the baseline methods unless stated otherwise. Our network architecture follows closely that of NeuS [15] and Neuralangelo [9]. In particular, the number of hash feature grids is set to 16 with a feature size of 2, and hash-map size 2^{19}. The base MLP consists of a single hidden layer with 64 neurons with ReLU activation, which predicts the latent vector and SDF. The color MLP consists of two hidden layers with 64 neurons each, mapping the normal vector, latent vector, and positionally encoded [12] viewing direction to color. We further employ the occupancy grid from NerfAcc with a resolution of 128^3, where the grid is binarized using an occupancy value of 10^{-3}. We set the bounding box radius to 1.5 for the simulated dataset and 0.4 for the captured dataset for all methods.

S1.1. Baseline Implementation Details

We compare our method against four baseline methods: Neuralangelo [9], RegNeRF [14], MonoSDF-M [16], and TransientNeRF [11]. To ensure a fair comparison, we implement those methods using NerfAcc and InstantNGP. As the first three methods are trained on images, we create intensity images by integrating the transients over the time dimension, and normalizing them between $[0, 1]$ using a scale factor, which is set per scene. Finally, the scaled images are gamma corrected with $\gamma = 2.2$. All image-based baseline methods (Neuralangelo, RegNeRF, MonoSDF-M) are trained using the conventional photometric loss [12].

$$\mathcal{L}_{\text{photo}} = \frac{1}{|\mathcal{R}_{\text{train}}|} \sum_{\mathcal{R}_{\text{train}}} ||\tilde{\boldsymbol{C}}(\boldsymbol{r}) - \boldsymbol{C}(\boldsymbol{r})||_2^2 \tag{S1}$$

where $\tilde{\boldsymbol{C}}$ and $\boldsymbol{C}$ denote the predicted and measured color, respectively. We compute this loss over the set of *active rays*, i.e. the set of rays with non-zero opacity values.

On the other hand, the transient-based method TransientNeRF is trained using the HDR-informed loss function:

$$\mathcal{L}_{\boldsymbol{\tau}} = \frac{1}{|\mathcal{R}_{\text{train}}|} \sum_{\mathcal{R}_{\text{train}}} \| \ln(\tilde{\boldsymbol{\tau}} + 1) - \ln(\boldsymbol{\tau}_f + 1)\|_1, \tag{S2}$$

where $\tilde{\boldsymbol{\tau}}$ is the measured transient and $\boldsymbol{\tau}_f$ is predicted transient.

Neuralangelo. We train Neuralangelo using only the Eikonal loss [4], as we did not notice any significant improvement by adding the curvature loss. The total objective is $\mathcal{L} = \mathcal{L}_{\text{photo}} + \lambda_{\text{eik}}\mathcal{L}_{\text{eik}}$. For both the simulated and captured datasets, we use $\lambda_{\text{eik}} = 0.1$

RegNeRF. For RegNeRF we extend Neuralangelo with an additional total variation regularizer on the depths rendered from the unseen cameras. We omit the color regularization as we do not have access to the weights of RealNVP [1]. We also implement Sample Space Annealing over the first 256 iterations. To apply their patch-based regularization, given an existing set of poses $\{\boldsymbol{P}\}_i = \{\boldsymbol{R}|\boldsymbol{t}\}_i$, we calculate the mean focus point by finding the point with shortest distances (in the least squares sense) to the optical axes of those cameras. Then, we calculate the bounding sphere of plausible cameras to sample from by computing the average of the distances of the focus point to the cameras. At each optimization step, we sample a new camera origin from that bounding sphere and align its rotation matrix such that it looks at the mean focus point. We then render an 8×8 patch over one of the training images and render the depth d of that patch following NeRF [12]. The final loss function extends Neuralangelo with the total variation regularizer on the rendered depth:

$$\mathcal{L}_{\text{ds}} = \sum_{\mathcal{R}_{\text{sampled}}} \sum_{i,j=1}^{8} (d(\boldsymbol{r}_{ij}) - d(\boldsymbol{r}_{i+1j}))^2 + d(\boldsymbol{r}_{ij}) - d(\boldsymbol{r}_{ij+1}))^2, \tag{S3}$$

where $\mathcal{R}_{\text{sampled}}$ are the rays from the sampled cameras. Therefore, the total loss is $\mathcal{L} = \mathcal{L}_{\text{photo}} + \lambda_{\text{eik}}\mathcal{L}_{\text{eik}} + \lambda_{\text{ds}}\mathcal{L}_{\text{ds}}$. For both the simulated and captured datasets, we use $\lambda_{\text{eik}} = \lambda_{\text{ds}} = 0.1$

MonoSDF w/ mask. MonoSDF has shown great surface reconstruction quality in few-view settings by introducing additional monocular depth and normal supervision. In particular, they leverage the pretrained Omnidata [2] model to estimate the depths and normals of training images. For a fair comparison, we instead estimate the depths from the measured transient using a log-matched filter [6]. As the estimated depths are noisy, and to provide additional supervision on the background, we used estimated masks to segment the object of interest, provide masked supervision on the depth and images [12].

From the estimated depth, we estimate normals by calculating the finite difference gradients of the depth point clouds. In particular, let $\tilde{d}$ be the estimated depth of a pixel in an image from a view, we use the camera parameters of that view to compute origins $\boldsymbol{o}$ and viewing directions $\boldsymbol{v}$ for each pixel. The point cloud $\boldsymbol{pc}$ for pixel (i,j) can be calculated as $\boldsymbol{pc}_{ij} = \tilde{d}\boldsymbol{v} + \boldsymbol{o}$. Next, we compute gradients in the horizontal (x) and vertical (y) directions, $g_x = \frac{\boldsymbol{pc}_{i+2j} - \boldsymbol{pc}_{ij}}{2}$ and $g_y = \frac{\boldsymbol{pc}_{ij+2} - \boldsymbol{pc}_{ij}}{2}$ and compute their normal as $\boldsymbol{n} = \frac{g_x \times g_y}{||g_x \times g_y||_2}$. Lastly, for fair comparison, we adopt the NeuS [15] conversion of the SDF to density. The training loss is $\mathcal{L} = \mathcal{L}_{\text{photo}} + \lambda_{\text{eik}}\mathcal{L}_{\text{eik}} + \lambda_{\text{depth}}\mathcal{L}_{\text{depth}} + \lambda_{\text{normal}}\mathcal{L}_{\text{normal}}$, where $\lambda_{\text{eik}} = \lambda_{\text{depth}} = 0.1$ and $\lambda_{\text{normal}} = 0.05$.

S1.2. Hyperparameters

For the simulated dataset, we set $\{\lambda_{\text{ref}}, \lambda_{\text{eik}}, \lambda_{\text{sc}}, \lambda_{\text{weight\_var}}, \lambda_{\text{sparse}}\} = \{3 \times 10^{-3}, 1 \times 10^{-5}, 7 \times 10^{-3}, 1 \times -10^{-3}, 3 \times 10^{-7}\}$. The sparsity scale α in the sparsity regularizer is set to 100. We do *not* tune these hyperparameters per scene. For the low photon experiments, since the overall signal of the transient becomes weaker, we need to enforce stricter supervision of the reflectivity loss. In particular, for photon experiments with $300, 150$ photons per pixel, we increase $\lambda_{\text{ref}} = 5 \times 10^{-3}$, for the experiment with 50ppp, we increase $\lambda_{\text{ref}} = 6 \times 10^{-3}$, and for the experiment with 10ppp, we set $\lambda_{\text{ref}} = 2 \times 10^{-2}$.

For the captured dataset, we set $\{\lambda_{\text{ref}}, \lambda_{\text{eik}}, \lambda_{\text{sc}}, \lambda_{\text{weight\_var}}, \lambda_{\text{sparse}}\} = \{7 \times 10^{-3}, 1 \times 10^{-5}, 1 \times 10^{-2}, 3 \times -10^{-2}, 1 \times 10^{-4}\}$ and increase the weight of the reflectivity loss λ_{ref} to 2×10^{-2} for the experiments with 10 ppp training transients, which improves convergence.

S1.3. Transient IoU Metric

In the main results of the paper, we compare our rendered transients against baseline methods, using the Transient IOU metric, which was previously introduced in Flying with Photons [10]. The metric measures the intersection over union of the predicted transient against the measured transient. Formally, let $\boldsymbol{\tau}_f$ and $\boldsymbol{\tau}$ be the predicted and measured transients, respectively. The Transient IoU is defined as:

$$\text{IoU}(\boldsymbol{\tau}_f, \boldsymbol{\tau}) = \frac{\sum \min(\boldsymbol{\tau}_f, \boldsymbol{\tau})}{\sum \min(\boldsymbol{\tau}_f, \boldsymbol{\tau})},$$

where the minimum and maximum are computed elementwise and the summation is summed across all dimensions of the tensor. The IoU metric ranges from 0 to 1 and measures the overlap between the predicted and measured transients. A score of 0 indicates no overlap while a score of 1 indicates perfect overlap.

S1.4. Low Photon Count Dataset

To create the low photon count transient dataset in simulation for each photon level $\{300, 150, 50, 10\}$, we integrate the transients along the time dimension to get images, and solve for a scale factor such that the ratio between the sum of pixels and the sum of the masked pixels of all images equals the desired photon level. After, we scale down the training transients by the scale factor and add a background level, which is set to 0.001 photons per 2850 scene photons to approximately match the captured results [11]. We then Poisson sample the resulting transient for each photon level.

To create the low photon dataset for the captured dataset, we again estimate the average number of photons per occupied pixel. To synthesize a transient with a desired average number of photons per pixel, we use subsample the photon arrivals, in a process called thinning [7].

S2. Additional Results

S2.1. Ablation Studies

Regularizers. In the following section, we show ablation studies of each of our regularizers on the *Lego* scene from the simulated dataset. We find that the method does not converge without the reflectivity loss, so we omit the quantitative results for that ablation. Our results are illustrated in Table S1.

In addition to the ablations shown in the main paper we also ablate the use of the Eikonal loss. The results obtained with and without the Eikonal loss have similar performance. As mentioned in the main paper, we tuned the Eikonal loss and the weight variance loss when ablating the reflectivity loss. In particular, we set $\lambda_{\text{eik}} = 0$ and $\lambda_{\text{weight\_var}} = 1 \times 10^{-5}$. Without those changes, training would diverge.

Argmax Depth vs. Convention NeRF Depth. To avoid multiple depth values skewing the expected ray termination, TransientNeRF [11] proposed to calculate the depth as the argmax of the weights along a ray:

$$d_{\text{ours}} = \arg\max_{t_i} w_i.$$

This is in contrast to the conventional depth calculation proposed in NeRF:

$$d_{\text{nerf}} = \sum_i w_i \frac{t_i + t_{i-1}}{2}.$$

We compare using each formulation in our weight variance regularizer in table S2.

S2.2. Simulated Results

Tables S4, S5, S6, and S7 provide a breakdown of the quantitative results in simulation across all scenes and photon levels. The results per scene are largely consistent with the general trends, averaged across scenes.

In Figures S9, S10, and S11 we provide additional qualitative results across varying photon levels. As can be seen, TransientNeRF [11] oftentimes fails to reconstruct the object in the case of 10 photons per pixel—we indicate such cases by omitting the mesh visualization in the figure. We attribute this failure to the lack of a reflectivity loss, which we observe helps our method converge robustly even at the lowest photon levels.

S2.3. Captured Results

Tables S8, S9, S10, S11 provide a breakdown of the quantitative results in the captured dataset across all scenes and photon levels. The results are largely consistent with general trends, averaged acorss scenes.

In Figures S6, S7, and S8 we provide additional qualitative results for scenes with an average of 1500 photons per pixel.

In Figures S12, S13, and S14 we provide additional qualitative results across varying photon levels. As can be seen, TransientNeRF [11] oftentimes fails to reconstruct the object for scenes with an average of 10 photons per pixel—we again indicate such cases by omitting the mesh visualization in the figure.

S2.4. Low-Photon Captured Results Without Masking

In Table S3, we show the results of our method and that of TransientNeRF in the low-photon regime when we are not applying ground-truth segmentation masks to filter out background noise. Compared with Table 2 in the main paper, our method drastically improves when supervised on segmented transients.

S2.5. Double Bounce

In Figure S1, we show on the left a rendered image of the *ficus* scene trained on 5 views and three selected pixel coordinates (right, blue, green) and their corresponding double-bounce transients on the right. From the transient figures, our method is fairly robust to double bounce returns, as the predicted transients match the ground-truth transients at those locations.

S2.6. Failure Case

In Figure S2, we show the extracted mesh from our optimized representation for the *boots* scene trained from two views across multiple photon levels (1500 ppp, 300 ppp, 150 ppp, 50 ppp). Multiple blob-like artifacts surrounding the mesh are present. We attribute this failure case to the ill-posed nature of the problem due to the sparse number of viewpoints.

Table S1. Ablation study on the proposed method in the *lego* scene. We present results while omitting the depth variance regularizer, without the Eikonal loss, without the sparsity loss, without the space carving loss, and without the reflectivity loss

	Chamfer Distance ↓			PSNR (dB) ↑			LPIPS ↓			L1 (depth) ↓			Transient IOU ↑		
Method	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
Ours	0.02	0.02	0.01	25.72	26.66	27.47	0.197	0.184	0.183	0.008	0.008	0.007	0.53	0.58	0.63
Ours w/o λ_{sc}	0.07	0.04	0.11	25.63	26.32	26.22	0.210	0.193	0.220	0.009	0.008	0.008	0.53	0.57	0.53
Ours w/o λ_{eik}	0.01	0.01	0.01	25.67	26.62	27.80	0.179	0.153	0.152	0.009	0.008	0.007	0.52	0.58	0.65
Ours w/o λ_{weight_var}	0.28	0.37	1.04	21.58	22.66	23.81	0.380	0.280	0.258	0.010	0.008	0.007	0.48	0.53	0.60
Ours w/o λ_{sparse}	0.02	0.03	0.08	25.43	26.87	26.58	0.218	0.193	0.194	0.009	0.008	0.007	0.53	0.58	0.63
Ours w/o λ_{ref}	5.50	2.91	0.53	22.38	24.92	26.22	0.239	0.193	0.153	0.010	0.017	0.007	0.49	0.55	0.65

Table S2. Ablation study comparing the effect of using our argmax depth formulation in the weight variance loss versus that of using NeRF's default depth calculation in the *lego* scene.

	Chamfer Distance ↓			PSNR (dB) ↑			LPIPS ↓			L1 (depth) ↓			Transient IOU ↑		
Method	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
$\mathcal{L}_{\text{weight\_var}}$ with d_{ours}	0.02	0.02	0.01	25.72	26.66	27.47	0.197	0.184	0.183	0.008	0.008	0.007	0.53	0.58	0.63
$\mathcal{L}_{\text{weight\_var}}$ with d_{nerf}	6.29	3.31	0.03	17.13	24.73	26.63	0.307	0.170	0.161	0.24	0.01	0.008	0.19	0.55	0.61

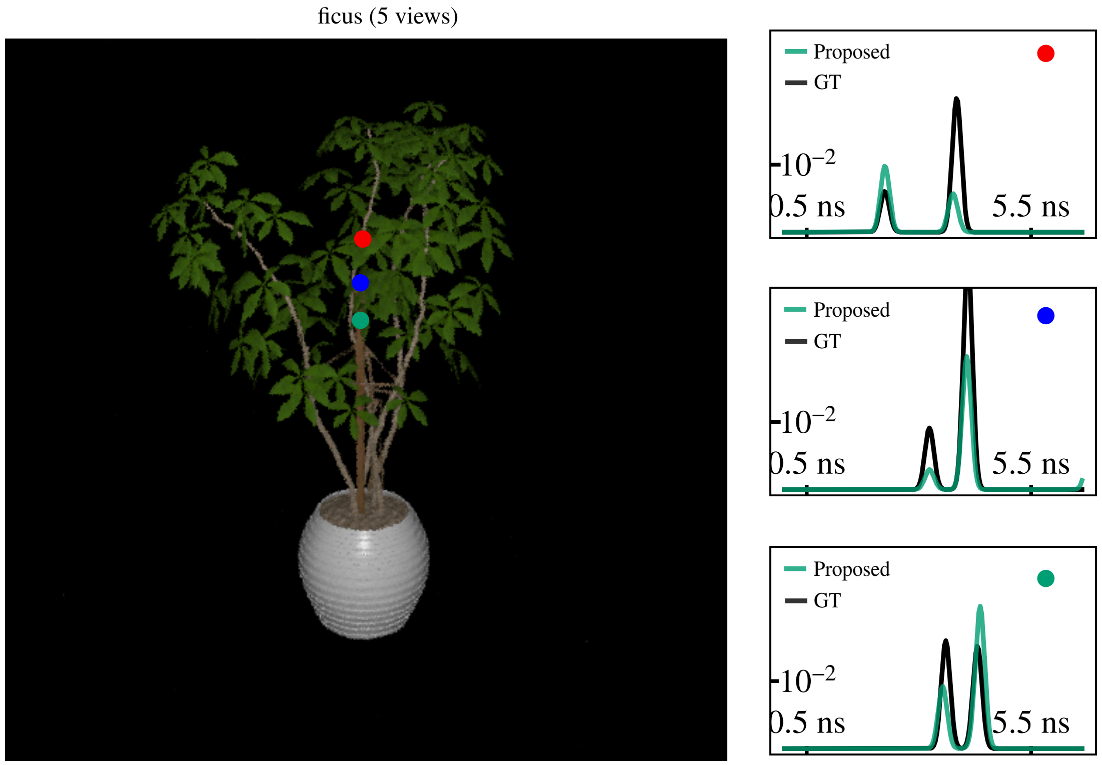

Figure S1. Predicted vs ground-truth double-bounce transients plotted versus time in (nanoseconds) for the selected pixels (red, blue, and green) for the *ficus* scene, trained on five views.

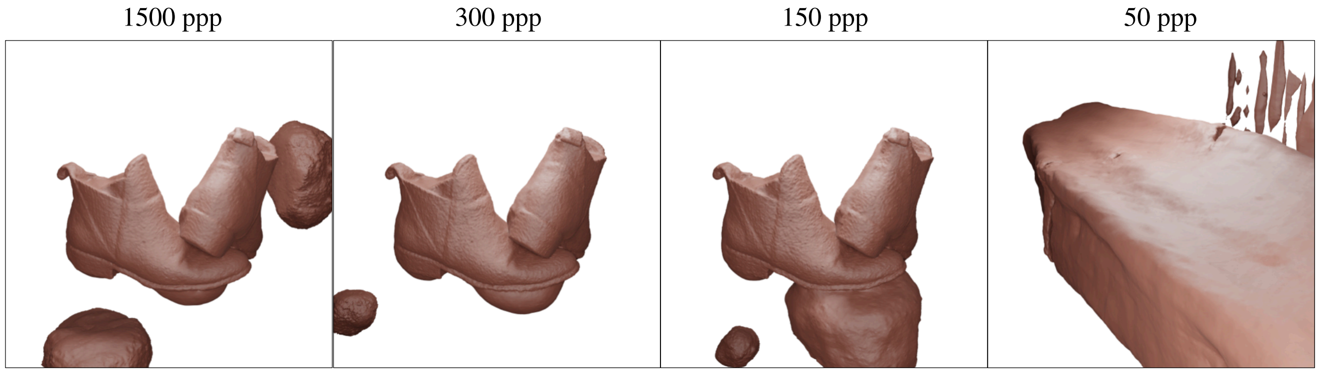

Figure S2. Failure case for our method. Rendered mesh for 1500 ppp, 300 ppp, 150 ppp, and 50 ppp where blob-like artifacts are surrounding the mesh.

Table S3. Captured results assessing image quality and depth accuracy without applying segmentation masks. The left column indicates average photons per pixel.

		PSNR (dB) ↑			LPIPS ↓			L1 (depth) ↓			Transient IOU ↑		
	Method	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
300	TransientNeRF	19.83	21.39	21.61	0.310	0.227	0.187	0.013	0.001	0.012	0.25	0.36	0.39
	Proposed	19.25	21.18	23.97	0.372	0.321	0.318	0.024	0.019	0.011	0.34	0.41	0.57
150	TransientNeRF	19.25	21.30	19.85	0.282	0.217	0.196	0.016	0.012	0.017	0.22	0.35	0.28
	Proposed	19.30	20.79	23.49	0.387	0.358	0.298	0.025	0.020	0.011	0.32	0.39	0.55
50	TransientNeRF	18.98	19.95	18.42	0.284	0.207	0.205	0.019	0.020	0.021	0.21	0.28	0.19
	Proposed	21.54	19.47	23.55	0.238	0.246	0.227	0.006	0.035	0.077	0.42	0.30	0.54
10	TransientNeRF	16.19	15.87	15.70	0.257	0.264	0.274	0.039	0.056	0.080	0.08	0.05	0.03
	Proposed	15.03	20.62	24.02	0.379	0.220	0.168	0.063	0.025	0.015	0.13	0.34	0.53

Table S4. Breakdown of Chamfer Distances across all 5 simulated scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
6000	Lego	4.03	3.02	3.31	2.95	2.12	2.76	1.50	0.28	0.07	0.05	0.08	0.05	0.02	0.02	0.01
	Chair	3.59	6.63	4.98	3.46	3.23	5.73	0.34	0.04	0.02	0.13	0.14	0.08	0.04	0.06	0.04
	Ficus	1.85	1.49	1.76	1.92	1.57	2.34	5.64	3.63	0.48	0.10	0.05	0.09	0.02	0.08	0.03
	Hotdog	7.71	3.54	2.92	12.59	3.90	1.94	1.48	1.02	0.29	0.66	1.45	1.06	0.22	0.06	0.03
	Bench	5.60	25.27	22.89	3.78	6.57	9.13	9.92	3.37	0.04	0.44	0.50	0.27	0.03	0.10	0.12
	Average	4.56	7.99	7.17	4.94	3.48	4.38	3.78	1.67	0.18	0.28	0.44	0.31	0.07	0.06	0.05
300	Lego	-	-	-	-	-	–	-	-	-	0.12	0.06	0.05	0.01	0.02	0.02
	Chair	-	-	-	-	-	-	-	-	-	0.17	0.17	0.10	0.06	0.03	0.07
	Ficus	-	-	-	-	-	-	-	-	-	0.11	0.46	0.57	0.05	0.12	0.08
	Hotdog	-	-	-	-	-	-	-	-	-	0.28	0.49	0.50	0.11	0.04	0.02
	Bench	-	-	-	-	-	-	-	-	-	0.60	0.47	0.28	0.11	0.10	0.12
	Average	-	-	-	-	-	-	-	-	-	0.26	0.33	0.30	0.07	0.06	0.06
150	Lego	-	-	-	-	-	–	-	-	-	0.09	0.15	0.04	0.02	0.05	0.04
	Chair	-	-	-	-	-	-	-	-	-	0.26	0.15	0.09	0.09	0.05	0.03
	Ficus	-	-	-	-	-	-	-	-	-	0.11	1.14	0.64	0.20	0.06	0.04
	Hotdog	-	-	-	-	-	-	-	-	-	0.20	1.18	0.25	0.05	0.05	0.04
	Bench	-	-	-	-	-	-	-	-	-	0.60	0.62	0.31	0.11	0.13	1.32
	Average	-	-	-	-	-	-	-	-	-	0.25	0.65	0.27	0.09	0.07	0.29
10	Lego	-	-	-	-	-	–	-	-	-	0.87	0.35	8.48	0.41	0.35	0.08
	Chair	-	-	-	-	-	-	-	-	-	0.97	0.30	5.56	0.09	0.13	0.25
	Ficus	-	-	-	-	-	-	-	-	-	3.91	0.36	3.20	0.18	0.36	0.21
	Hotdog	-	-	-	-	-	-	-	-	-	8.38	0.43	0.26	0.25	0.43	0.28
	Bench	-	-	-	-	-	-	-	-	-	0.15	0.12	10.02	0.15	0.12	0.07
	Average	-	-	-	-	-	-	-	-	-	2.86	0.31	5.50	0.22	0.28	0.18

Table S5. Breakdown of PSNR across all 5 simulated scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
6000	Lego	17.72	19.17	18.33	17.91	19.09	18.97	17.46	21.93	22.45	20.64	23.63	25.81	25.72	26.66	27.47
	Chair	17.29	16.92	24.54	14.87	17.04	24.64	21.38	21.44	30.86	20.75	21.99	34.48	25.70	26.98	31.14
	Ficus	23.26	23.63	25.82	22.97	23.41	25.32	23.83	23.75	26.22	24.57	26.10	27.70	28.14	28.35	32.52
	Hotdog	15.87	19.85	20.41	17.25	20.32	20.34	16.84	16.89	18.95	20.74	22.64	32.36	22.72	26.20	28.33
	Bench	18.72	17.65	21.04	18.58	19.57	19.08	19.30	19.68	24.76	20.20	23.06	21.57	24.10	25.24	27.99
	Average	18.57	19.44	22.03	18.32	19.89	21.67	19.76	20.74	24.65	21.38	23.48	28.38	25.28	26.69	29.49
300	Lego	-	-	-	-	-	–	-	-	-	20.42	23.97	25.04	26.04	26.79	26.75
	Chair	-	-	-	-	-	-	-	-	-	20.78	21.54	33.45	25.98	26.96	30.89
	Ficus	-	-	-	-	-	-	-	-	-	23.16	22.87	24.14	27.58	28.99	31.64
	Hotdog	-	-	-	-	-	-	-	-	-	21.55	22.64	21.89	24.42	25.36	26.70
	Bench	-	-	-	-	-	-	-	-	-	20.79	22.95	24.87	24.47	25.15	26.51
	Average	-	-	-	-	-	-	-	-	-	21.34	22.79	25.88	25.70	26.65	28.50
150	Lego	-	-	-	-	-	–	-	-	-	20.07	23.18	24.52	26.20	26.57	27.08
	Chair	-	-	-	-	-	-	-	-	-	20.68	21.47	31.46	25.77	26.89	30.75
	Ficus	-	-	-	-	-	-	-	-	-	22.32	22.01	23.63	27.20	29.13	31.98
	Hotdog	-	-	-	-	-	-	-	-	-	21.24	20.98	22.68	24.60	25.26	25.59
	Bench	-	-	-	-	-	-	-	-	-	20.87	23.28	24.48	24.03	24.76	25.57
	Average	-	-	-	-	-	-	-	-	-	21.04	22.18	25.35	25.56	26.52	28.19
10	Lego	-	-	-	-	-	–	-	-	-	16.75	23.32	15.90	21.85	23.32	23.72
	Chair	-	-	-	-	-	-	-	-	-	17.32	17.87	14.83	23.70	23.56	26.75
	Ficus	-	-	-	-	-	-	-	-	-	20.75	25.35	23.40	22.30	25.35	28.82
	Hotdog	-	-	-	-	-	-	-	-	-	16.98	23.81	15.79	23.03	23.81	23.29
	Bench	-	-	-	-	-	-	-	-	-	22.90	23.88	13.75	22.90	23.87	24.95
	Average	-	-	-	-	-	-	-	-	-	18.94	22.85	16.73	22.76	23.98	25.51

Table S6. Breakdown of LPIPS across all 5 simulated scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
6000	Lego	0.480	0.520	0.570	0.470	0.460	0.510	0.320	0.230	0.190	0.190	0.161	0.192	0.197	0.184	0.183
	Chair	0.400	0.530	0.430	0.320	0.460	0.410	0.150	0.140	0.080	0.138	0.138	0.037	0.140	0.136	0.110
	Ficus	0.220	0.230	0.200	0.230	0.240	0.220	0.170	0.150	0.110	0.094	0.079	0.069	0.114	0.106	0.105
	Hotdog	0.380	0.370	0.430	0.440	0.390	0.400	0.300	0.330	0.160	0.242	0.241	0.118	0.212	0.164	0.168
	Bench	0.460	0.560	0.540	0.420	0.510	0.550	0.260	0.370	0.120	0.194	0.139	0.159	0.216	0.186	0.183
	Average	0.388	0.442	0.434	0.376	0.412	0.418	0.240	0.244	0.132	0.172	0.152	0.115	0.176	0.155	0.150
300	Lego	-	-	-	-	-	–	-	-	-	0.221	0.149	0.167	0.261	0.207	0.200
	Chair	-	-	-	-	-	-	-	-	-	0.156	0.128	0.040	0.150	0.130	0.110
	Ficus	-	-	-	-	-	-	-	-	-	0.119	0.136	0.138	0.143	0.127	0.09
	Hotdog	-	-	-	-	-	-	-	-	-	0.277	0.206	0.159	0.204	0.171	0.142
	Bench	-	-	-	-	-	-	-	-	-	0.232	0.142	0.100	0.222	0.177	0.204
	Average	-	-	-	-	-	-	-	-	-	0.201	0.152	0.121	0.196	0.162	0.149
150	Lego	-	-	-	-	-	–	-	-	-	0.220	0.174	0.160	0.244	0.194	0.197
	Chair	-	-	-	-	-	-	-	-	-	0.191	0.132	0.048	0.157	0.138	0.105
	Ficus	-	-	-	-	-	-	-	-	-	0.134	0.147	0.140	0.131	0.118	0.09
	Hotdog	-	-	-	-	-	-	-	-	-	0.256	0.225	0.127	0.206	0.182	0.159
	Bench	-	-	-	-	-	-	-	-	-	0.215	0.129	0.109	0.228	0.192	0.151
	Average	-	-	-	-	-	-	-	-	-	0.203	0.161	0.117	0.193	0.165	0.151
10	Lego	-	-	-	-	-	–	-	-	-	0.290	0.248	0.243	0.308	0.248	0.268
	Chair	-	-	-	-	-	-	-	-	-	0.244	0.185	0.220	0.210	0.176	0.152
	Ficus	-	-	-	-	-	-	-	-	-	0.166	0.185	0.150	0.221	0.185	0.124
	Hotdog	-	-	-	-	-	-	-	-	-	0.358	0.250	0.205	0.267	0.250	0.208
	Bench	-	-	-	-	-	-	-	-	-	0.265	0.220	0.226	0.265	0.220	0.289
	Average	-	-	-	-	-	-	-	-	-	0.265	0.218	0.209	0.254	0.216	0.208

Table S7. Breakdown of SSIM across all 5 simulated scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
6000	Lego	0.590	0.550	0.480	0.580	0.590	0.550	0.730	0.830	0.850	0.860	0.897	0.899	0.900	0.920	0.930
	Chair	0.640	0.530	0.680	0.730	0.580	0.680	0.870	0.870	0.950	0.901	0.899	0.977	0.930	0.940	0.960
	Ficus	0.820	0.810	0.860	0.800	0.800	0.830	0.870	0.880	0.930	0.925	0.934	0.942	0.950	0.942	0.969
	Hotdog	0.640	0.670	0.650	0.570	0.630	0.660	0.810	0.780	0.870	0.882	0.875	0.963	0.900	0.942	0.960
	Bench	0.580	0.450	0.530	0.600	0.540	0.450	0.770	0.720	0.870	0.856	0.884	0.870	0.870	0.873	0.890
	Average	0.654	0.602	0.640	0.656	0.628	0.634	0.810	0.816	0.894	0.885	0.898	0.930	0.910	0.923	0.942
300	Lego	-	-	-	-	-	–	-	-	-	0.846	0.903	0.910	0.893	0.915	0.919
	Chair	-	-	-	-	-	-	-	-	-	0.891	0.892	0.974	0.925	0.936	0.961
	Ficus	-	-	-	-	-	-	-	-	-	0.895	0.879	0.883	0.935	0.943	0.962
	Hotdog	-	-	-	-	-	-	-	-	-	0.859	0.901	0.908	0.925	0.939	0.954
	Bench	-	-	-	-	-	-	-	-	-	0.842	0.886	0.896	0.861	0.872	0.878
	Average	-	-	-	-	-	-	-	-	-	0.867	0.892	0.914	0.907	0.921	0.935
150	Lego	-	-	-	-	-	–	-	-	-	0.840	0.882	0.896	0.895	0.913	0.920
	Chair	-	-	-	-	-	-	-	-	-	0.863	0.881	0.964	0.923	0.934	0.960
	Ficus	-	-	-	-	-	-	-	-	-	0.882	0.867	0.879	0.930	0.947	0.965
	Hotdog	-	-	-	-	-	-	-	-	-	0.874	0.855	0.911	0.930	0.934	0.946
	Bench	-	-	-	-	-	-	-	-	-	0.851	0.889	0.893	0.857	0.871	0.873
	Average	-	-	-	-	-	-	-	-	-	0.862	0.875	0.909	0.907	0.920	0.933
10	Lego	-	-	-	-	-	–	-	-	-	0.745	0.840	0.737	0.797	0.840	0.840
	Chair	-	-	-	-	-	-	-	-	-	0.801	0.816	0.778	0.872	0.876	0.876
	Ficus	-	-	-	-	-	-	-	-	-	0.850	0.877	0.866	0.871	0.877	0.877
	Hotdog	-	-	-	-	-	-	-	-	-	0.739	0.860	0.789	0.852	0.860	0.860
	Bench	-	-	-	-	-	-	-	-	-	0.823	0.844	0.762	0.822	0.844	0.844
	Average	-	-	-	-	-	-	-	-	-	0.792	0.847	0.786	0.843	0.859	0.859

Table S8. Breakdown of PSNR across all 6 captured scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
1500	Cinema	17.80	18.46	20.39	17.96	19.63	17.03	18.53	21.71	25.75	21.61	21.66	25.12	23.63	24.70	28.03
	Food	19.09	19.51	23.03	19.89	20.44	22.42	17.31	22.34	30.74	23.40	23.78	22.09	24.58	24.53	25.12
	Carving	19.36	21.44	22.17	19.63	21.28	21.47	19.27	21.85	27.00	23.52	24.20	24.70	20.62	25.65	27.92
	Boots	15.89	17.86	19.93	17.38	17.74	18.92	15.77	22.59	30.30	22.38	22.29	24.94	16.58	25.43	26.32
	Baskets	19.92	23.07	24.14	20.95	23.28	24.86	20.43	21.12	24.31	22.48	20.93	19.90	22.43	23.10	23.32
	Chef	14.69	15.32	16.76	15.20	15.15	17.03	14.47	17.35	25.38	19.27	18.14	19.55	20.02	20.30	20.61
	Average	17.79	19.28	21.07	18.50	19.59	20.29	17.63	21.16	27.25	22.11	21.83	22.72	21.31	23.95	25.22
300	Cinema	-	-	-	-	-	-	-	-	-	16.82	22.37	22.43	21.94	23.76	28.08
	Food	-	-	-	-	-	-	-	-	-	18.86	22.31	21.09	22.98	23.00	24.03
	Carving	-	-	-	-	-	-	-	-	-	23.48	23.83	25.39	21.73	25.48	27.82
	Boots	-	-	-	-	-	-	-	-	-	24.37	22.53	18.07	20.86	25.32	17.57
	Baskets	-	-	-	-	-	-	-	-	-	22.22	22.15	20.27	21.32	22.97	24.24
	Chef	-	-	-	-	-	-	-	-	-	15.09	18.72	20.53	20.78	20.66	20.98
	Average	-	-	-	-	-	-	-	-	-	20.14	21.98	21.30	21.60	23.53	23.79
150	Cinema	-	-	-	-	-	-	-	-	-	17.06	21.97	19.18	22.44	24.58	27.86
	Food	-	-	-	-	-	-	-	-	-	18.80	20.47	21.98	22.93	23.05	24.18
	Carving	-	-	-	-	-	-	-	-	-	18.36	24.14	25.19	18.96	25.39	27.75
	Boots	-	-	-	-	-	-	-	-	-	23.32	24.36	19.24	20.92	24.91	26.91
	Baskets	-	-	-	-	-	-	-	-	-	22.14	19.41	18.28	21.53	23.15	24.14
	Chef	-	-	-	-	-	-	-	-	-	15.07	16.02	15.80	20.61	20.28	21.13
	Average	-	-	-	-	-	-	-	-	-	19.12	21.06	19.94	21.23	23.56	25.33
10	Cinema	-	-	-	-	-	-	-	-	-	15.36	15.21	14.33	21.91	23.06	25.30
	Food	-	-	-	-	-	-	-	-	-	18.44	18.27	17.39	22.71	23.05	24.43
	Carving	-	-	-	-	-	-	-	-	-	17.58	19.59	17.71	23.07	24.76	27.37
	Boots	-	-	-	-	-	-	-	-	-	16.32	15.09	14.81	23.83	24.29	28.35
	Baskets	-	-	-	-	-	-	-	-	-	16.58	16.38	15.38	23.10	22.85	24.63
	Chef	-	-	-	-	-	-	-	-	-	14.13	13.63	12.79	19.79	20.26	22.76
	Average	-	-	-	-	-	-	-	-	-	16.40	16.36	15.40	22.40	23.04	25.47

Table S9. Breakdown of LPIPS across all 6 captured scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
1500	Cinema	0.424	0.387	0.350	0.390	0.330	0.470	0.330	0.260	0.180	0.281	0.245	0.178	0.223	0.163	0.138
	Food	0.384	0.368	0.247	0.400	0.330	0.270	0.370	0.230	0.130	0.286	0.205	0.154	0.183	0.176	0.141
	Carving	0.266	0.244	0.205	0.260	0.220	0.220	0.270	0.230	0.130	0.232	0.138	0.103	0.222	0.115	0.100
	Boots	0.289	0.369	0.261	0.300	0.320	0.290	0.380	0.160	0.110	0.221	0.182	0.155	0.302	0.128	0.111
	Baskets	0.318	0.298	0.268	0.320	0.270	0.250	0.320	0.220	0.160	0.269	0.165	0.164	0.211	0.179	0.176
	Chef	0.601	0.464	0.467	0.580	0.450	0.460	0.450	0.390	0.230	0.334	0.338	0.276	0.302	0.257	0.226
	Average	0.380	0.355	0.300	0.375	0.320	0.327	0.353	0.248	0.157	0.270	0.212	0.172	0.240	0.170	0.149
300	Cinema	-	-	-	-	-	-	-	-	-	0.356	0.204	0.189	0.238	0.182	0.148
	Food	-	-	-	-	-	-	-	-	-	0.256	0.191	0.186	0.180	0.169	0.140
	Carving	-	-	-	-	-	-	-	-	-	0.209	0.166	0.104	0.225	0.127	0.088
	Boots	-	-	-	-	-	-	-	-	-	0.223	0.200	0.171	0.241	0.135	0.153
	Baskets	-	-	-	-	-	-	-	-	-	0.222	0.183	0.177	0.184	0.163	0.163
	Chef	-	-	-	-	-	-	-	-	-	0.356	0.315	0.234	0.269	0.244	0.214
	Average	-	-	-	-	-	-	-	-	-	0.270	0.210	0.177	0.223	0.170	0.151
150	Cinema	-	-	-	-	-	-	-	-	-	0.356	0.232	0.219	0.212	0.183	0.144
	Food	-	-	-	-	-	-	-	-	-	0.236	0.171	0.170	0.188	0.178	0.134
	Carving	-	-	-	-	-	-	-	-	-	0.224	0.181	0.106	0.263	0.114	0.094
	Boots	-	-	-	-	-	-	-	-	-	0.247	0.113	0.168	0.216	0.123	0.117
	Baskets	-	-	-	-	-	-	-	-	-	0.261	0.210	0.193	0.192	0.182	0.166
	Chef	-	-	-	-	-	-	-	-	-	0.337	0.326	0.317	0.260	0.245	0.220
	Average	-	-	-	-	-	-	-	-	-	0.277	0.206	0.196	0.222	0.171	0.146
10	Cinema	-	-	-	-	-	-	-	-	-	0.266	0.243	0.333	0.305	0.238	0.193
	Food	-	-	-	-	-	-	-	-	-	0.225	0.218	0.227	0.205	0.202	0.162
	Carving	-	-	-	-	-	-	-	-	-	0.200	0.150	0.148	0.193	0.139	0.114
	Boots	-	-	-	-	-	-	-	-	-	0.213	0.396	0.270	0.183	0.150	0.124
	Baskets	-	-	-	-	-	-	-	-	-	0.279	0.237	0.290	0.243	0.217	0.197
	Chef	-	-	-	-	-	-	-	-	-	0.290	0.307	0.361	0.321	0.300	0.250
	Average	-	-	-	-	-	-	-	-	-	0.246	0.258	0.272	0.242	0.208	0.173

Table S10. Breakdown of SSIM across all 6 captured scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
1500	Cinema	0.600	0.690	0.710	0.600	0.650	0.480	0.730	0.810	0.910	0.850	0.812	0.879	0.879	0.897	0.925
	Food	0.650	0.760	0.840	0.500	0.610	0.720	0.610	0.830	0.950	0.826	0.879	0.873	0.890	0.892	0.912
	Carving	0.740	0.820	0.850	0.700	0.760	0.770	0.770	0.840	0.930	0.913	0.929	0.929	0.804	0.933	0.951
	Boots	0.740	0.670	0.800	0.670	0.640	0.690	0.640	0.900	0.960	0.912	0.909	0.914	0.706	0.933	0.940
	Baskets	0.700	0.760	0.790	0.630	0.700	0.750	0.690	0.810	0.900	0.846	0.850	0.826	0.854	0.860	0.870
	Chef	0.370	0.600	0.600	0.290	0.500	0.480	0.600	0.670	0.910	0.775	0.747	0.811	0.823	0.847	0.856
	Average	0.633	0.717	0.765	0.565	0.643	0.648	0.673	0.810	0.927	0.854	0.854	0.872	0.826	0.894	0.909
300	Cinema	-	-	-	-	-	-	-	-	-	0.585	0.851	0.859	0.854	0.882	0.922
	Food	-	-	-	-	-	-	-	-	-	0.813	0.877	0.840	0.884	0.889	0.908
	Carving	-	-	-	-	-	-	-	-	-	0.911	0.913	0.935	0.846	0.930	0.952
	Boots	-	-	-	-	-	-	-	-	-	0.923	0.899	0.851	0.809	0.932	0.869
	Baskets	-	-	-	-	-	-	-	-	-	0.859	0.860	0.824	0.845	0.867	0.885
	Chef	-	-	-	-	-	-	-	-	-	0.717	0.751	0.849	0.846	0.854	0.867
	Average	-	-	-	-	-	-	-	-	-	0.801	0.858	0.860	0.847	0.892	0.900
150	Cinema	-	-	-	-	-	-	-	-	-	0.663	0.842	0.810	0.863	0.889	0.920
	Food	-	-	-	-	-	-	-	-	-	0.816	0.855	0.878	0.882	0.884	0.912
	Carving	-	-	-	-	-	-	-	-	-	0.825	0.912	0.933	0.762	0.931	0.950
	Boots	-	-	-	-	-	-	-	-	-	0.899	0.930	0.878	0.834	0.932	0.942
	Baskets	-	-	-	-	-	-	-	-	-	0.851	0.804	0.791	0.851	0.865	0.886
	Chef	-	-	-	-	-	-	-	-	-	0.704	0.716	0.699	0.844	0.848	0.862
	Average	-	-	-	-	-	-	-	-	-	0.793	0.843	0.832	0.839	0.892	0.912
10	Cinema	-	-	-	-	-	-	-	-	-	0.744	0.741	0.457	0.802	0.841	0.883
	Food	-	-	-	-	-	-	-	-	-	0.808	0.789	0.652	0.868	0.871	0.907
	Carving	-	-	-	-	-	-	-	-	-	0.850	0.872	0.854	0.881	0.918	0.938
	Boots	-	-	-	-	-	-	-	-	-	0.833	0.163	0.519	0.897	0.905	0.929
	Baskets	-	-	-	-	-	-	-	-	-	0.765	0.761	0.517	0.845	0.844	0.874
	Chef	-	-	-	-	-	-	-	-	-	0.714	0.698	0.451	0.792	0.806	0.847
	Average	-	-	-	-	-	-	-	-	-	0.786	0.671	0.575	0.848	0.864	0.896

Table S11. Breakdown of L1 Depth across all 6 captured scenes

		Neuralangelo [9]			RegNeRF [14]			MonoSDF-M [16]			TransientNeRF [11]			Ours		
	Scene	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views	2 views	3 views	5 views
1500	Cinema	0.080	0.068	0.068	0.080	0.070	0.100	0.030	0.030	0.030	0.006	0.006	0.007	0.007	0.007	0.007
	Food	0.094	0.080	0.069	0.090	0.080	0.090	0.040	0.020	0.030	0.006	0.007	0.016	0.006	0.007	0.008
	Carving	0.054	0.060	0.045	0.070	0.060	0.060	0.030	0.030	0.020	0.006	0.005	0.006	0.007	0.006	0.006
	Boots	0.069	0.051	0.045	0.060	0.050	0.040	0.020	0.005	0.004	0.001	0.002	0.002	0.010	0.001	0.001
	Baskets	0.074	0.069	0.049	0.090	0.080	0.070	0.050	0.040	0.030	0.008	0.008	0.026	0.007	0.008	0.011
	Chef	0.086	0.100	0.090	0.120	0.100	0.100	0.030	0.030	0.010	0.003	0.006	0.006	0.002	0.004	0.005
	Average	0.076	0.071	0.061	0.085	0.073	0.077	0.033	0.026	0.021	0.005	0.006	0.010	0.006	0.006	0.006
300	Cinema	-	-	-	-	-	-	-	-	-	0.028	0.010	0.018	0.007	0.007	0.007
	Food	-	-	-	-	-	-	-	-	-	0.020	0.010	0.015	0.006	0.006	0.007
	Carving	-	-	-	-	-	-	-	-	-	0.007	0.007	0.007	0.007	0.006	0.006
	Boots	-	-	-	-	-	-	-	-	-	0.001	0.002	0.009	0.001	0.001	0.034
	Baskets	-	-	-	-	-	-	-	-	-	0.008	0.009	0.029	0.007	0.007	0.011
	Chef	-	-	-	-	-	-	-	-	-	0.010	0.005	0.006	0.002	0.004	0.004
	Average	-	-	-	-	-	-	-	-	-	0.012	0.007	0.014	0.005	0.005	0.012
150	Cinema	-	-	-	-	-	-	-	-	-	0.028	0.011	0.021	0.007	0.007	0.007
	Food	-	-	-	-	-	-	-	-	-	0.023	0.014	0.014	0.006	0.006	0.007
	Carving	-	-	-	-	-	-	-	-	-	0.024	0.007	0.006	0.014	0.006	0.006
	Boots	-	-	-	-	-	-	-	-	-	0.001	0.001	0.008	0.004	0.001	0.001
	Baskets	-	-	-	-	-	-	-	-	-	0.009	0.033	0.035	0.007	0.007	0.009
	Chef	-	-	-	-	-	-	-	-	-	0.012	0.022	0.022	0.001	0.005	0.004
	Average	-	-	-	-	-	-	-	-	-	0.016	0.015	0.018	0.006	0.005	0.006
10	Cinema	-	-	-	-	-	-	-	-	-	0.045	0.040	0.117	0.009	0.011	0.008
	Food	-	-	-	-	-	-	-	-	-	0.033	0.034	0.086	0.008	0.008	0.008
	Carving	-	-	-	-	-	-	-	-	-	0.028	0.014	0.023	0.010	0.008	0.007
	Boots	-	-	-	-	-	-	-	-	-	0.017	0.067	0.073	0.001	0.001	0.001
	Baskets	-	-	-	-	-	-	-	-	-	0.049	0.054	0.112	0.016	0.015	0.014
	Chef	-	-	-	-	-	-	-	-	-	0.049	0.037	0.131	0.006	0.005	0.004
	Average	-	-	-	-	-	-	-	-	-	0.037	0.041	0.090	0.008	0.008	0.00

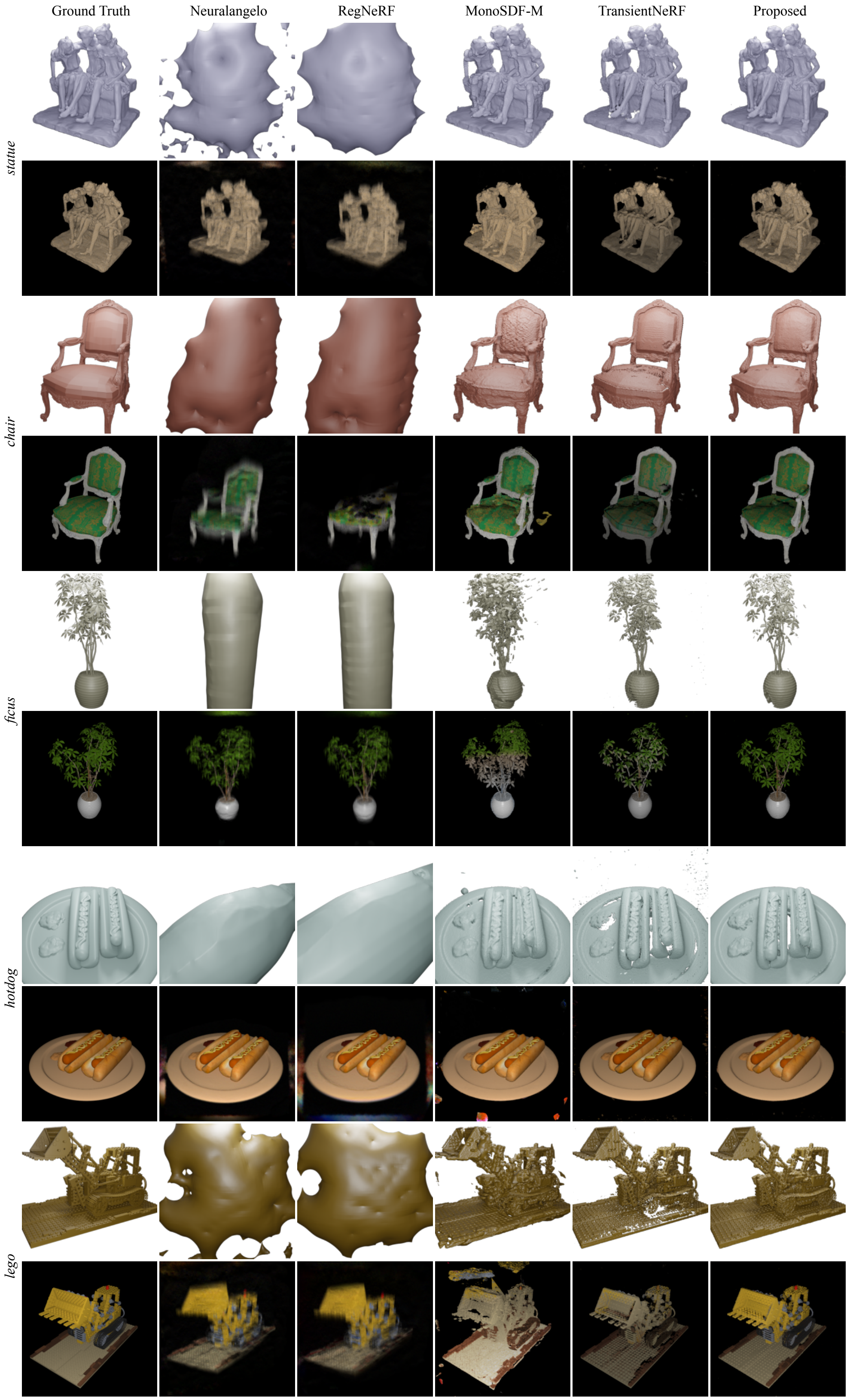

Figure S3. Rendered meshes and images on the simulated dataset for 2 views.

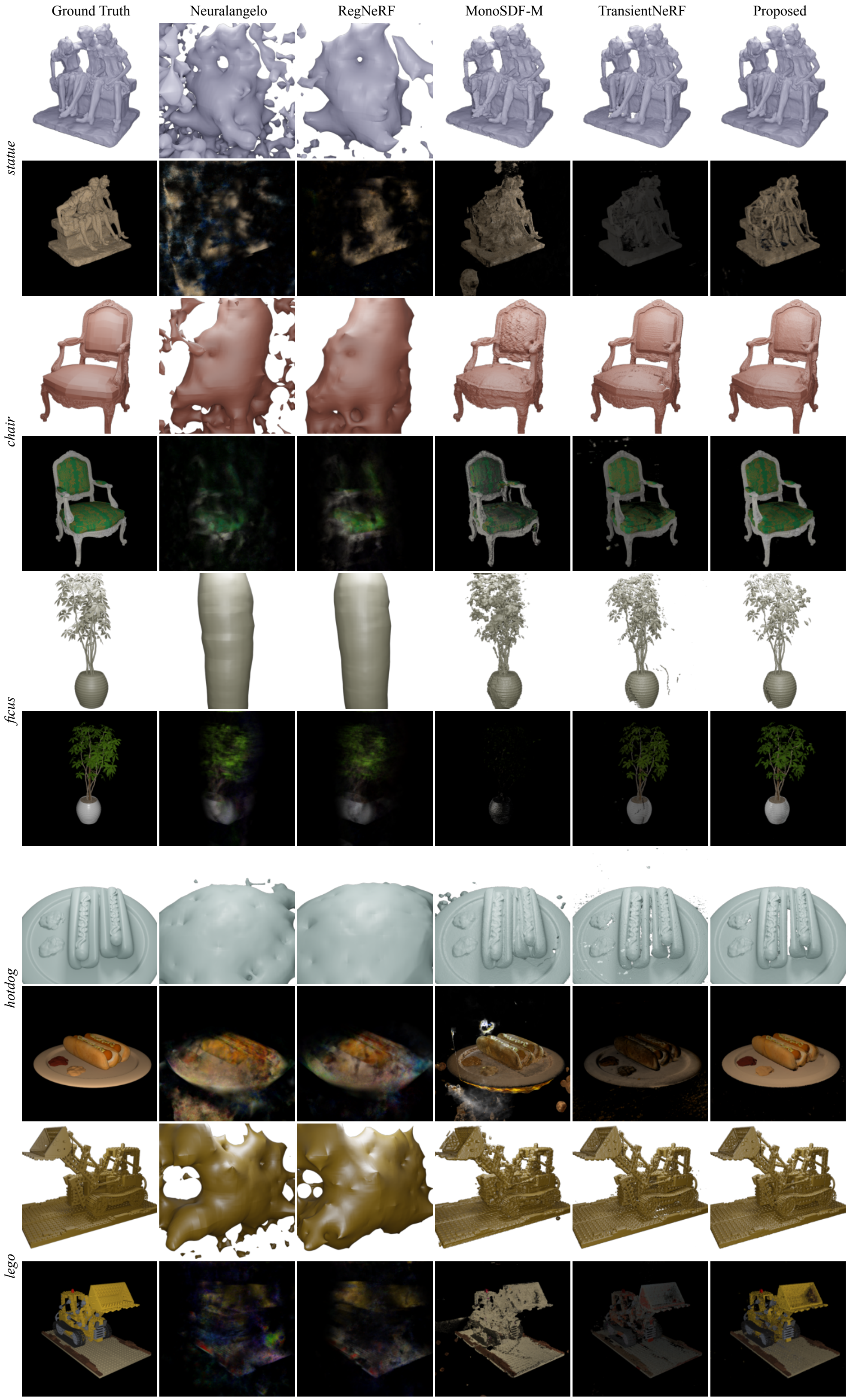

Figure S4. Rendered meshes and images on the simulated dataset for 3 views.

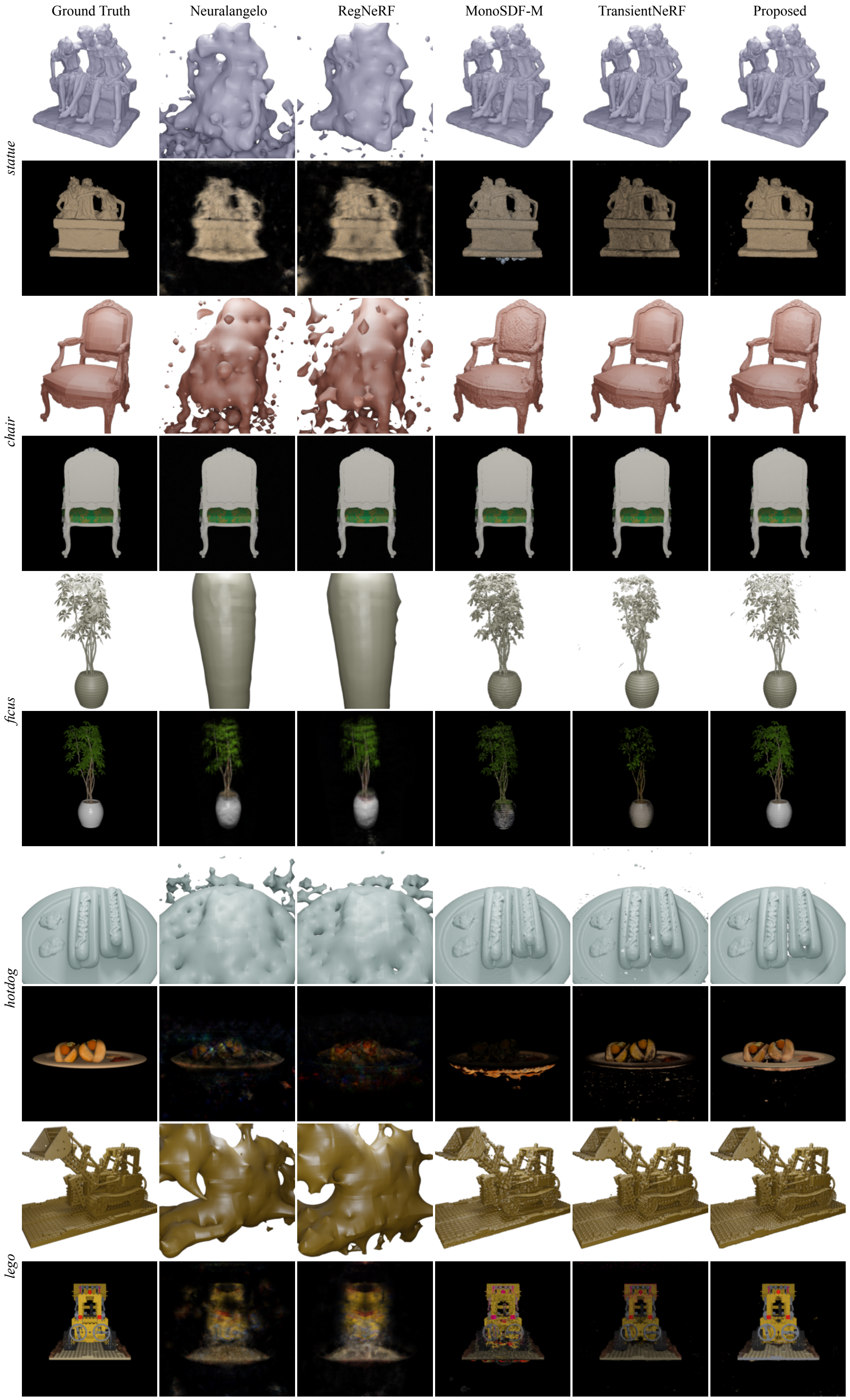

Figure S5. Rendered meshes and images on the simulated dataset for 5 views.

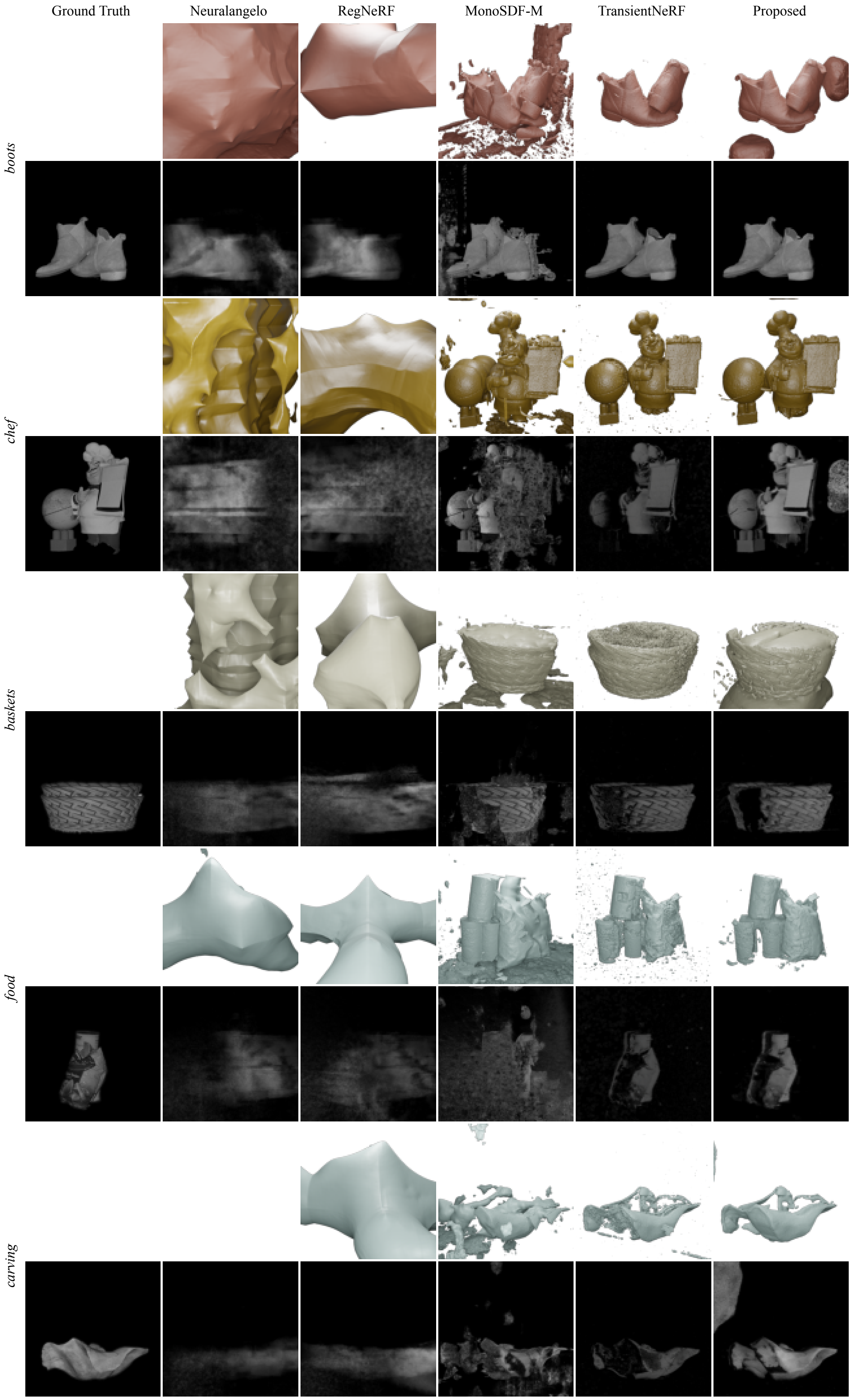

Figure S6. Rendered meshes and images on the captured dataset for 2 views.

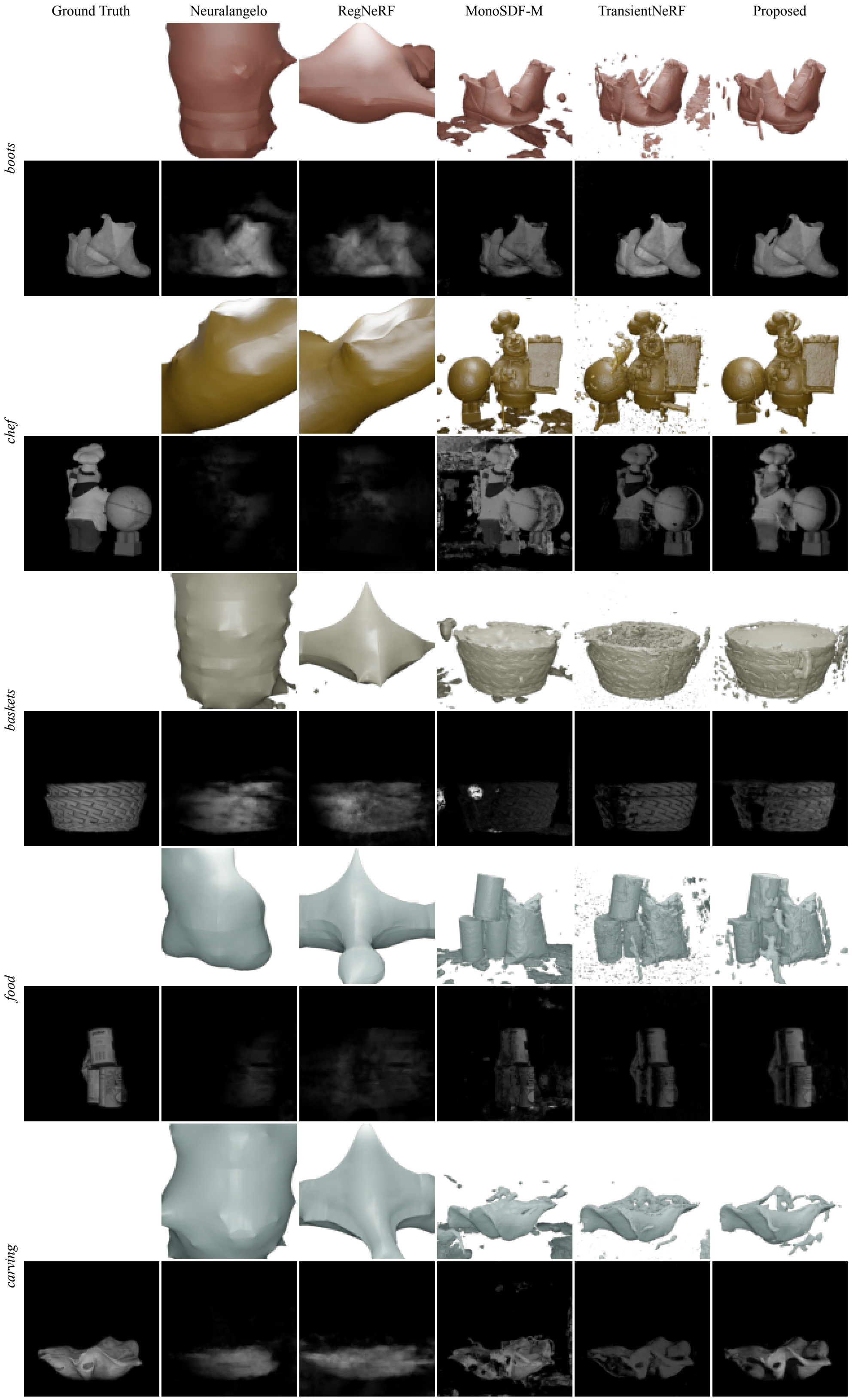

Figure S7. Rendered meshes and images on the captured dataset for 3 views.

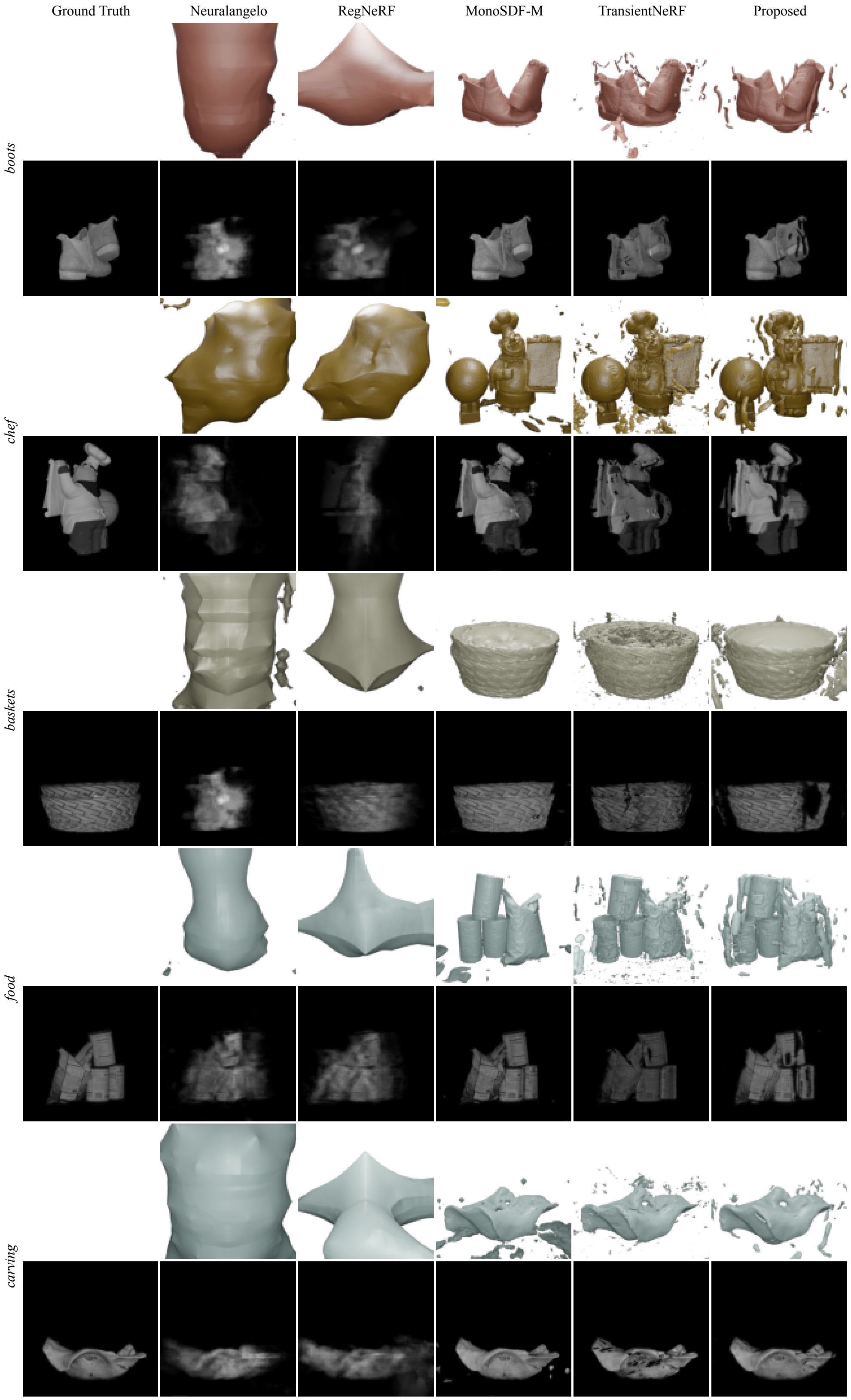

Figure S8. Rendered meshes and images on the captured dataset for 5 views.

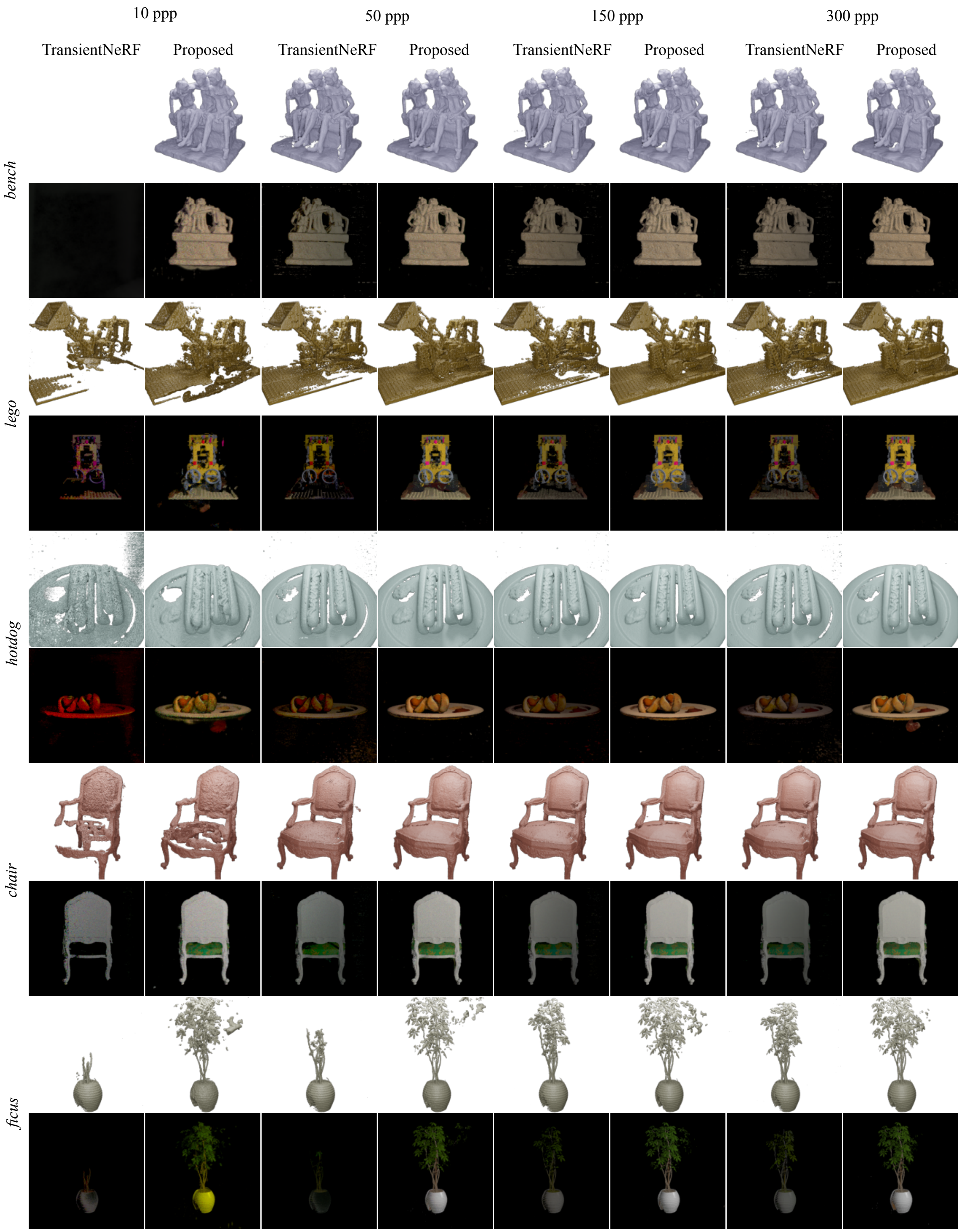

Figure S9. Rendered meshes and images for the low photon count experiments on the simulated dataset for 2 views.

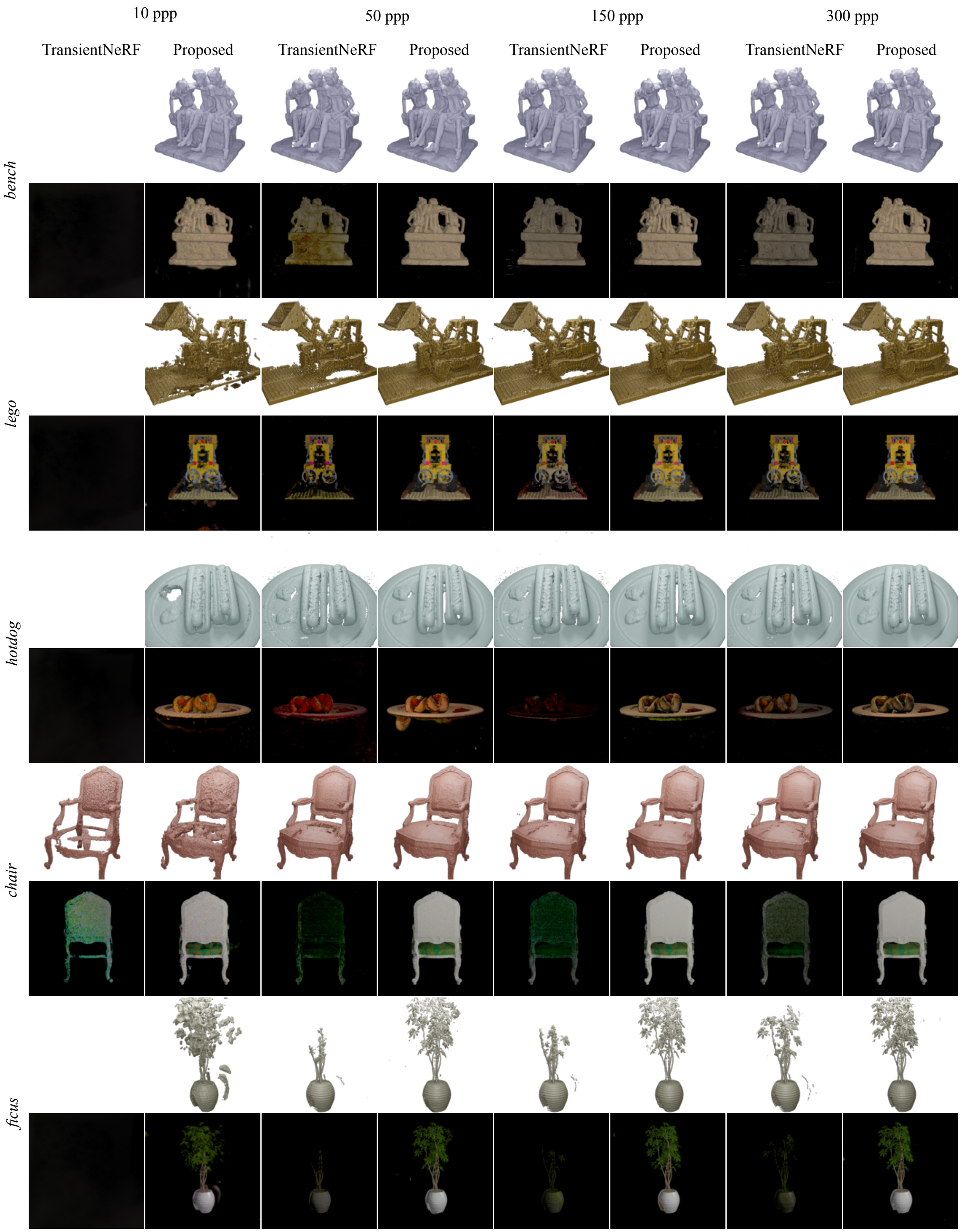

Figure S10. Rendered meshes and images for the low photon count experiments on the simulated dataset for 3 views.

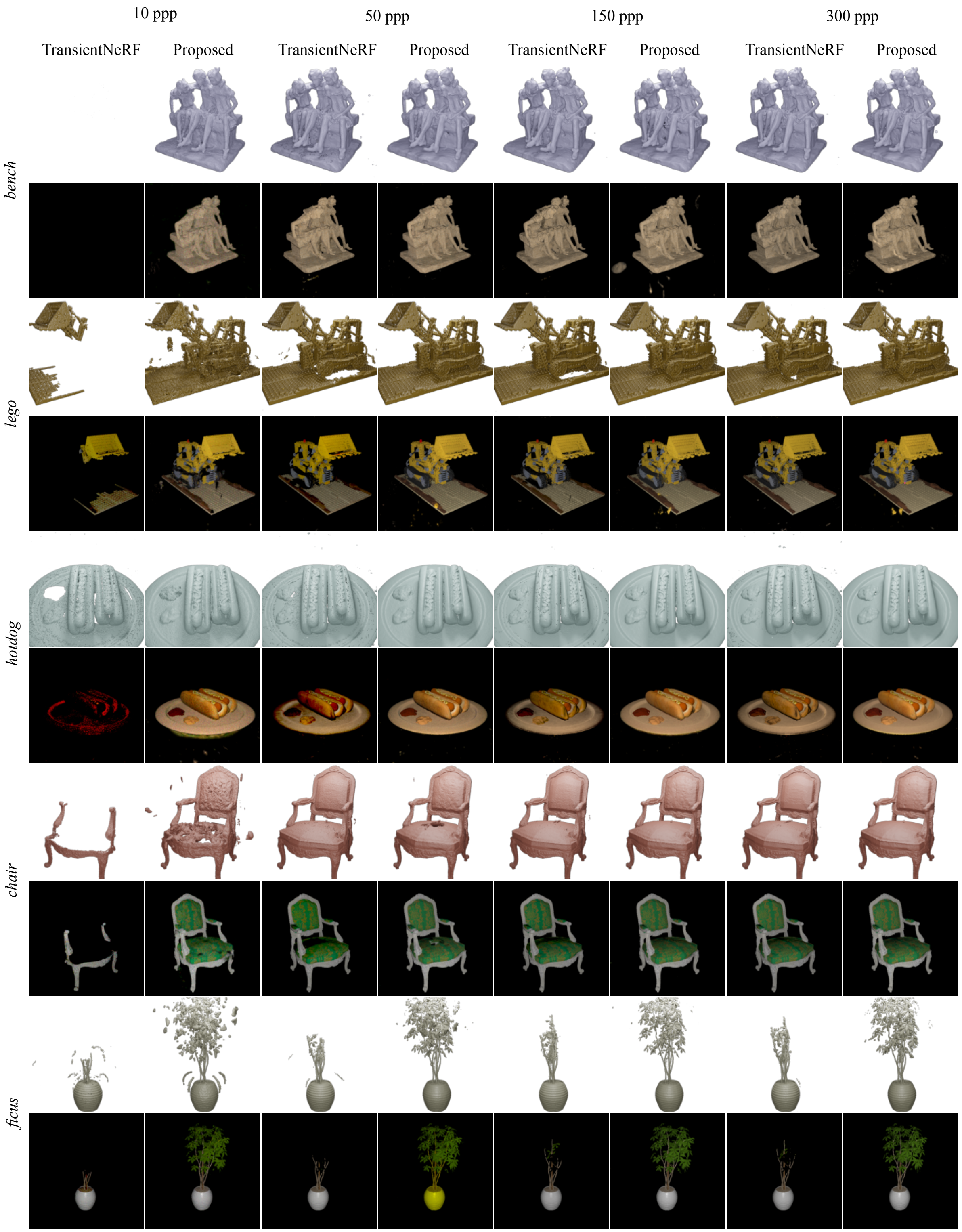

Figure S11. Rendered meshes and images for the low photon count experiments on the simulated dataset for 5 views.

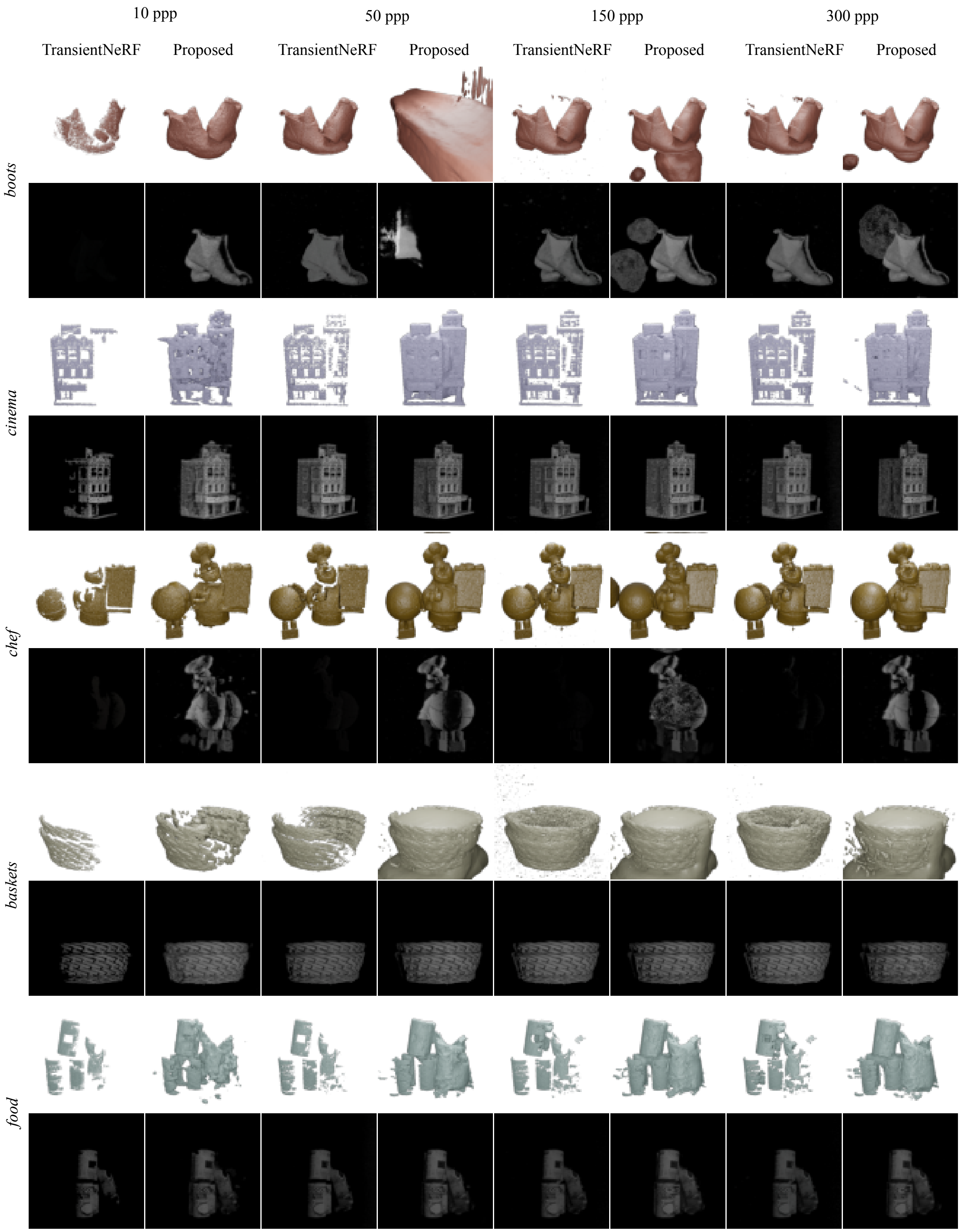

Figure S12. Rendered meshes and images for the low photon count experiments on the captured dataset for 2 views.

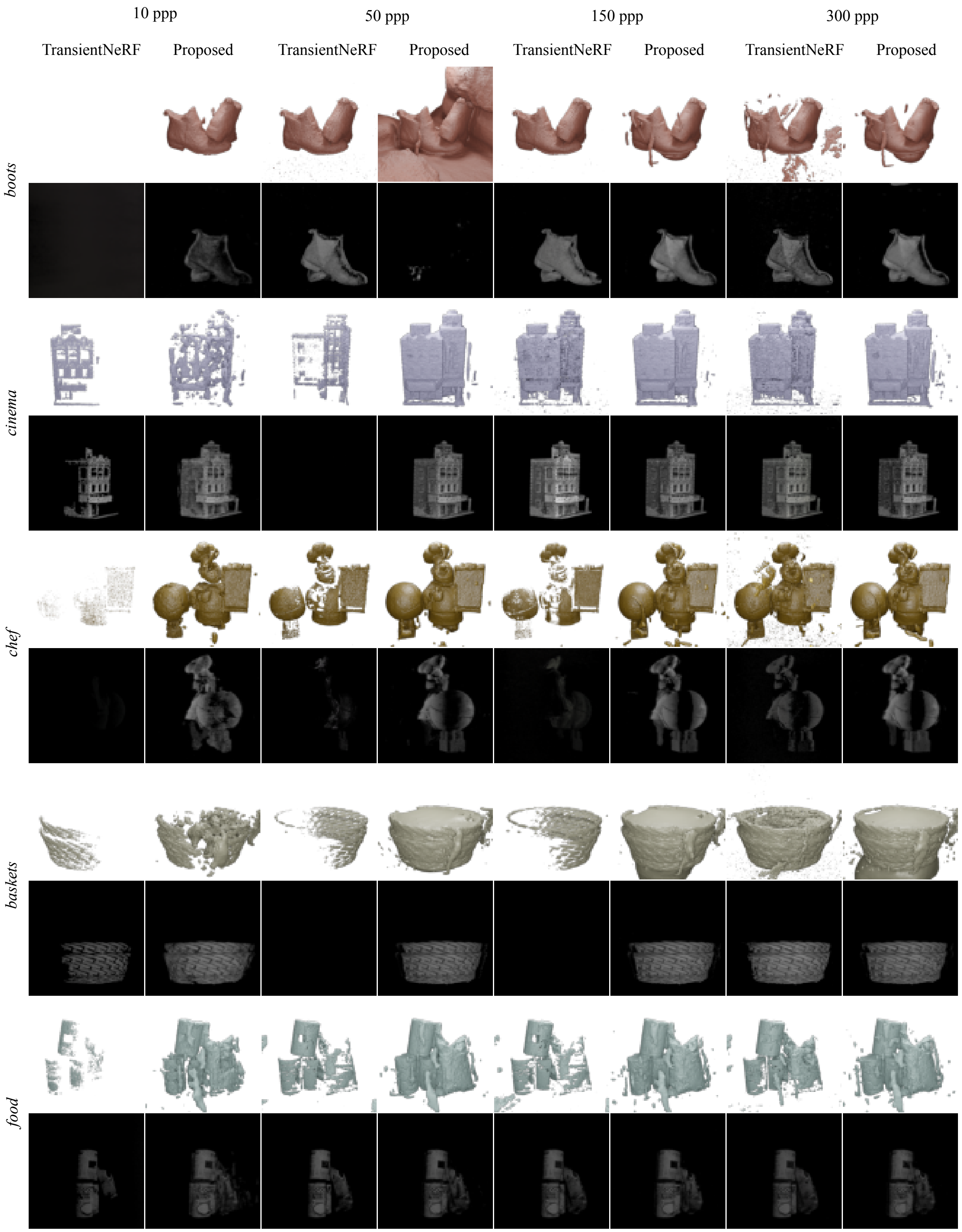

Figure S13. Rendered meshes and images for the low photon count experiments on the captured dataset for 3 views.

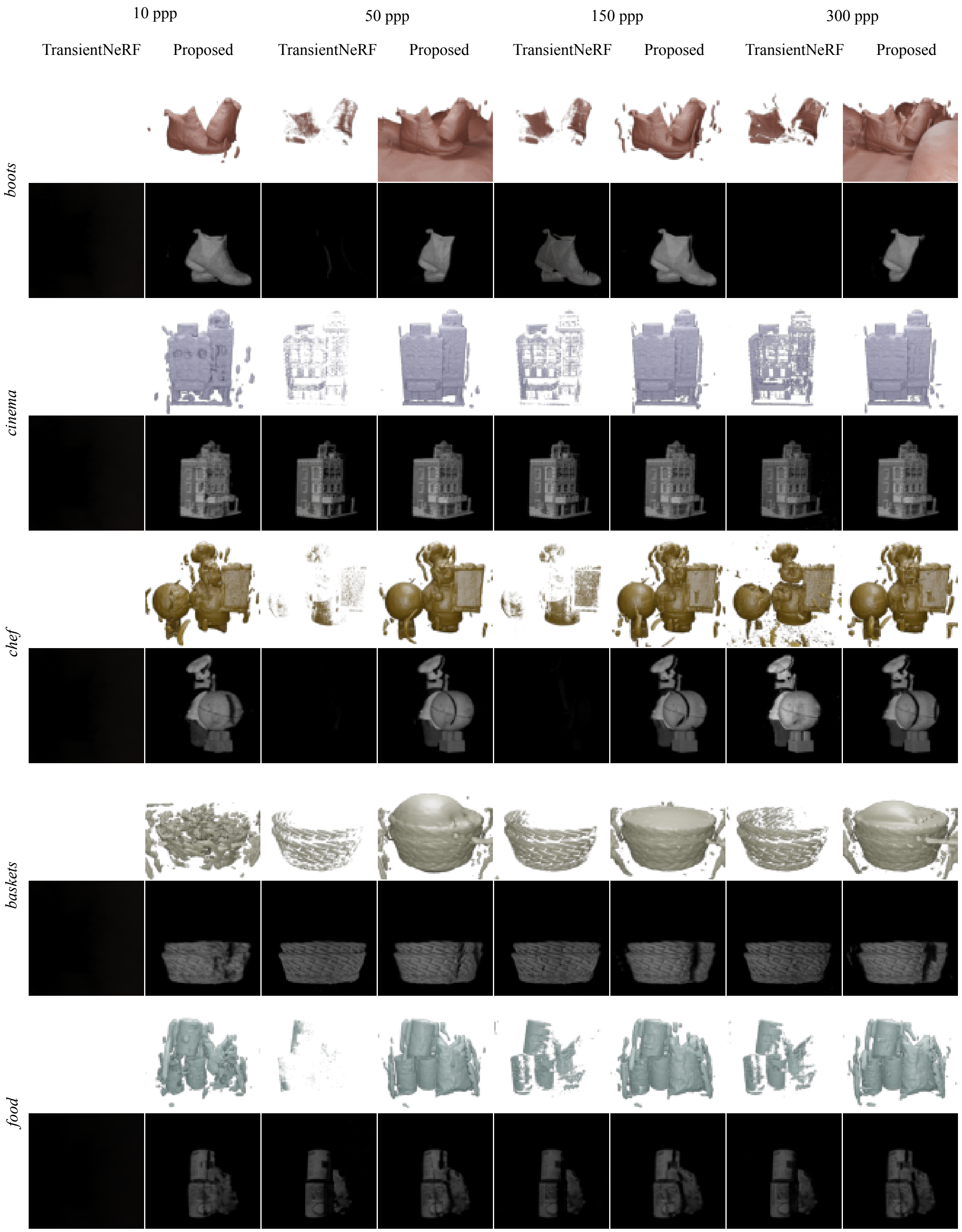

Figure S14. Rendered meshes and images for the low photon count experiments on the captured dataset for 5 views.

References

[1] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803*, 2016.

[2] Ainaz Eftekhar, Alexander Sax, Jitendra Malik, and Amir Zamir. Omnidata: A scalable pipeline for making multi-task mid-level vision datasets from 3d scans. In *Proc. ICCV*, 2021.

[3] William A Falcon. Pytorch lightning. *GitHub*, 3, 2019.

[4] Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. Implicit geometric regularization for learning shapes. *arXiv preprint arXiv:2002.10099*, 2020.

[5] Yuan-Chen Guo. Instant neural surface reconstruction, 2022. https://github.com/bennyguo/instant-nsr-pl.

[6] Ahmed Kirmani, Dheera Venkatraman, Dongeek Shin, Andrea Colaço, Franco NC Wong, Jeffrey H Shapiro, and Vivek K Goyal. First-photon imaging. *Science*, 343(6166):58–61, 2014.

[7] PA W Lewis and Gerald S Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval research logistics quarterly*, 26(3):403–413, 1979.

[8] Ruilong Li, Matthew Tancik, and Angjoo Kanazawa. Nerfacc: A general nerf acceleration toolbox. *arXiv preprint arXiv:2210.04847*, 2022.

[9] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *Proc. CVPR*, 2023.

[10] Anagh Malik, Noah Juravsky, Ryan Po, Gordon Wetzstein, Kiriakos N. Kutulakos, and David B. Lindell. Flying with photons: Rendering novel views of propagating light. In *Proc. ECCV*, 2024.

[11] Anagh Malik, Parsa Mirdehghan, Sotiris Nousias, Kiriakos N. Kutulakos, and David B. Lindell. Transient neural radiance fields for lidar view synthesis and 3d reconstruction. *Proc. NeurIPS*, 2023.

[12] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

[13] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):1–15, 2022.

[14] Michael Niemeyer, Jonathan T Barron, Ben Mildenhall, Mehdi SM Sajjadi, Andreas Geiger, and Noha Radwan. Regnerf: Regularizing neural radiance fields for view synthesis from sparse inputs. In *Proc. CVPR*, 2022.

[15] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *arXiv preprint arXiv:2106.10689*, 2021.

[16] Zehao Yu, Songyou Peng, Michael Niemeyer, Torsten Sattler, and Andreas Geiger. Monosdf: Exploring monocular geometric cues for neural implicit surface reconstruction. *Proc. NeurIPS*, 35, 2022.