ProbeSDF: Light Field Probes For Neural Surface Reconstruction Supplementary material

Briac Toussaint[1]
briac.toussaint@inria.fr

Diego Thomas[2]
thomas@ait.kyushu-u.ac.jp

Jean-Sébastien Franco[1]
jean-sebastien.franco@inria.fr

[1]Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, France
[2]Kyushu University, Japan

1. Scheduling

The scheduling of the hyper-parameters is summarized in table 1. We used $\beta_1 = 0.9$, $\beta_2 = 0.995$ and $\lambda_{\text{photo}} = 40$ in all experiments, where β_1 and β_2 are ADAM's first and second moments. A blank cell indicates that the value is identical to that of the row above. Values between brackets are linearly interpolated based on the current training iteration of the given level of detail. Our general strategy is to start with a high learning rate and high regularizations that are progressively halved during training. We lower the learning rate in the final level of detail to stabilize convergence. The learning rate is also linearly increased starting from zero during the first 50 iterations of each level of detail as a warm-up. We use stronger regularizations at the coarsest level of detail on DTU and BMVS since the initialization from a sphere is poorer compared to the visual hull initialization of the other two datasets. The gradients are accumulated over 4 to 8 *complete* images before stepping the optimizer. The photometric loss is divided by the number of images in the batch as a normalization. We adjust the number of levels of detail based on the resolution of the input images. The grid resolution is manually chosen so that a voxel roughly projects to an area equivalent to that of a pixel. The MLP weights and probe features are trained with a lower learning rate compared to the spatial features and voxel sdf values, also for stability.

2. Visual ablation

A visual ablation of the spherical harmonic order l is presented in fig 1.

3. Detailed tables

Tables 7 to 19 contain detailed metrics from all our experiments. The dimensionality of the spatial features $\mathbf{F}_s$ is n_s and the dimensionality of the angular features $\mathbf{F}_a$ is n_a. We denote a training configuration with a triplet (n_s, n_a, l). A ✓-symbol denotes the training of per-camera bias vectors, an ✗-symbol is used otherwise.

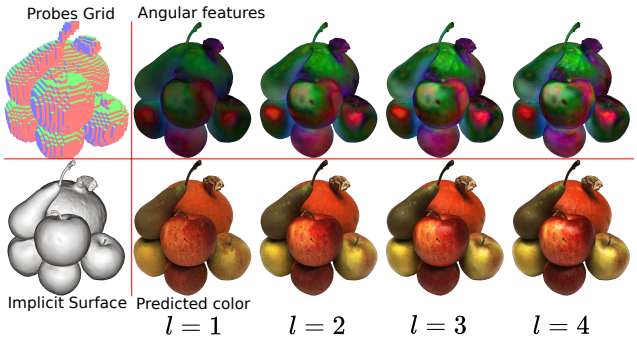

Figure 1. We train a scene with $l = 4$, that we then visualize with l varying from 1 to 4, left to right. The corresponding angular features are on the top row, and the predicted color is on the bottom row. We observe that the specularities can be removed simply by disabling the high order spherical harmonics coefficients.

4. Neural Architecture comparison

Table 5 gives a comparison of several neural architectures used in the context of implicit surface reconstruction. Thanks to our light field probes, our MLP is entirely agnostic to the surface orientation and position, hence we can reduce its size and obtain high quality renderings, in real-time.

5. Comparisons on ActorsHQ

Geometric comparisons on the ActorsHQ dataset [2] are shown in figures 3, 4, 5, 6. The dataset comes with meshes reconstructed by RealityCapture [1], a multi-view stereo reconstruction software. We cannot compute geometric met-

Level of detail	Learning rate (Voxels)	Learning rate (MLP)	λ_{Eik}	λ_{sdf}	$\lambda_{\text{features}}$	λ_{normal}	λ_{probes}	Images per batch	Iterations
LOD 4	[0.025, 0.01]	0.01	[1.0, 0.1, 0.1]	[2.0, 0.2, 0.2]	[0.5, 0.05, 0.05]	[0.5, 0.05, 0.05]	[2.0, 0.2, 0.2]	8	3000
LOD 3			[0.2, 0.1]	[0.4, 0.2]	[0.1, 0.05]	[0.1, 0.05]	[0.4, 0.2]		
LOD 2									
LOD 1									1000
LOD 0	[0.01, 0.001]	[0.01, 0.001]						4	500

Table 1. Schedule for DTU

Level of detail	Learning rate (Voxels)	Learning rate (MLP)	λ_{Eik}	λ_{sdf}	$\lambda_{\text{features}}$	λ_{normal}	λ_{probes}	Images per batch	Iterations
LOD 4	[0.1, 0.01]	0.01	[0.25, 0.025, 0.025]	[0.5, 0.05, 0.05]	[0.25, 0.025, 0.025]	[0.025, 0.0025, 0.0025]	[2.0, 0.2, 0.2]	8	4000
LOD 3			[0.05, 0.025]	[0.1, 0.05]	[0.05, 0.025]	[0.005, 0.0025]	[0.4, 0.2]		
LOD 2	[0.025, 0.01]								3000
LOD 1									1500
LOD 0	[0.01, 0.001]	[0.01, 0.001]						4	1000

Table 2. Schedule for BlendedMVS

Level of detail	Learning rate (Voxels)	Learning rate (MLP)	λ_{Eik}	λ_{sdf}	$\lambda_{\text{features}}$	λ_{normal}	λ_{probes}	Images per batch	Iterations
LOD 4	[0.025, 0.01]	0.01	[0.05, 0.025]	[0.2, 0.1]	[0.1, 0.05]	0	[0.4, 0.2]	4	1500
LOD 3									
LOD 2									
LOD 1									
LOD 0	[0.01, 0.001]	[0.01, 0.001]	0.0125	0.05	0.025	0.0063	0.1		1000

Table 3. Schedule for MVMannequins

Level of detail	Learning rate (Voxels)	Learning rate (MLP)	λ_{Eik}	λ_{sdf}	$\lambda_{\text{features}}$	λ_{normal}	λ_{probes}	Images per batch	Iterations
LOD 6	[0.05, 0.025]	0.01	[0.05, 0.025]	[0.2, 0.1]	[0.025, 0.0125]	0	[0.4,0.2]	8	500
LOD 5									1000
LOD 4									1500
LOD 3									
LOD 2						[0.025, 0.0125]			1000
LOD 1	[0.01, 0.0025]	[0.01, 0.001]	0.0125	0.05	0.0063	0.0125	0.1	4	
LOD 0	[0.01, 0.001]	0.001				0.025			500

Table 4. Schedule for ActorsHQ

Method	Grid Type	SDF MLP Layers / Neurons	Color MLP Layers / Neurons
NeuS	✗	8 / 256	4 / 256
NeuS2	hash-grid	1 / 64	2 / 64
Voxurf	dense	✗	4 / 192
Ours	sparse	✗	2 / 32

Table 5. Neural architectures in the literature.

rics since there is no ground truth obtained independently from the images. A qualitative comparison of the volume rendering quality is shown in figures 7, 8 and 9. We used the (4,4,4) configuration here. Note that the input images come with pre-baked segmentation masks, as shown in figure 2, that tend to have poorer accuracy on the arms and hands. This results in both geometric and photometric artifacts that are difficult to eliminate.

Figure 2. Imprecise segmentation example.

6. Comparisons on DTU

Figure 10 presents a comparison of some of the reconstruction results on DTU [3]. Close-ups of the volume rendered

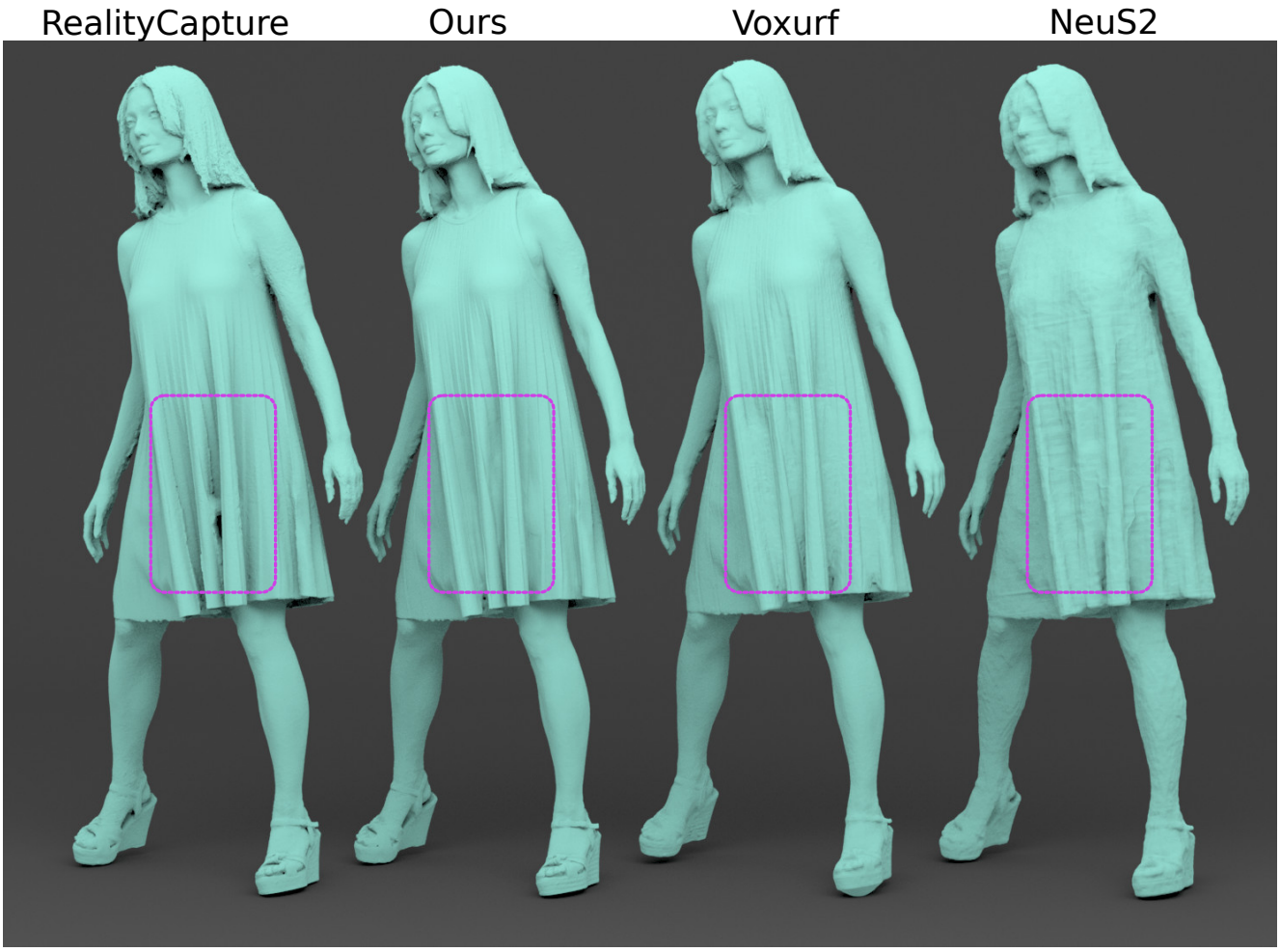

Figure 3. Reconstruction results on ActorsHQ. Left to right: RealityCapture, Ours, Voxurf, NeuS2.

images are shown in figures 11 and 12. Our results are obtained with the (4,4,4) configuration.

7. Comparisons on BlendedMVS

Geometric comparisons on the BlendedMVS dataset [4] are shown in figures 13 and 14. Qualitative comparisons of the volume rendering are shown in figure 15. We used the (8,8,4) configuration here as it performed a little better on this dataset (see table 16).

8. Performance Analysis

We record inference timings on the apples example (scan 63 of DTU, (4,4,4) configuration), with a window of resolution 1920×1163 and present the results in table 6. We lock the memory clock to 5001MHz and the gpu clock to 1500MHz to obtain stable performance measurements. The shading column corresponds to the assignment of a color to each voxel. The render column corresponds to the volume rendering kernel, which samples the SDF and color fields along rays to generate the final image. The first 4 rows correspond to the fully-fused color prediction kernel, with the computation of the spatial and angular features enabled or disabled. Thus, the 4th row corresponds to the MLP inference only whereas the 1st row corresponds to the full model. The last two rows correspond to the computation of $\mathbf{F}_s$ or $\mathbf{F}_a$ on their own, in separate kernels, and whose result is interpreted as a per-voxel color for visualization.

We observe that just evaluating the MLP (4th row) or computing the angular features on their own (5th row) roughly takes the same amount of time (1.80ms and 1.96ms) but that fusing the two operations together only takes 2.32ms, much less than the sum of the two ($1.80 + 1.96 = 3.76$ms). Including the computation of the spatial features gives the full model at 2.47ms.

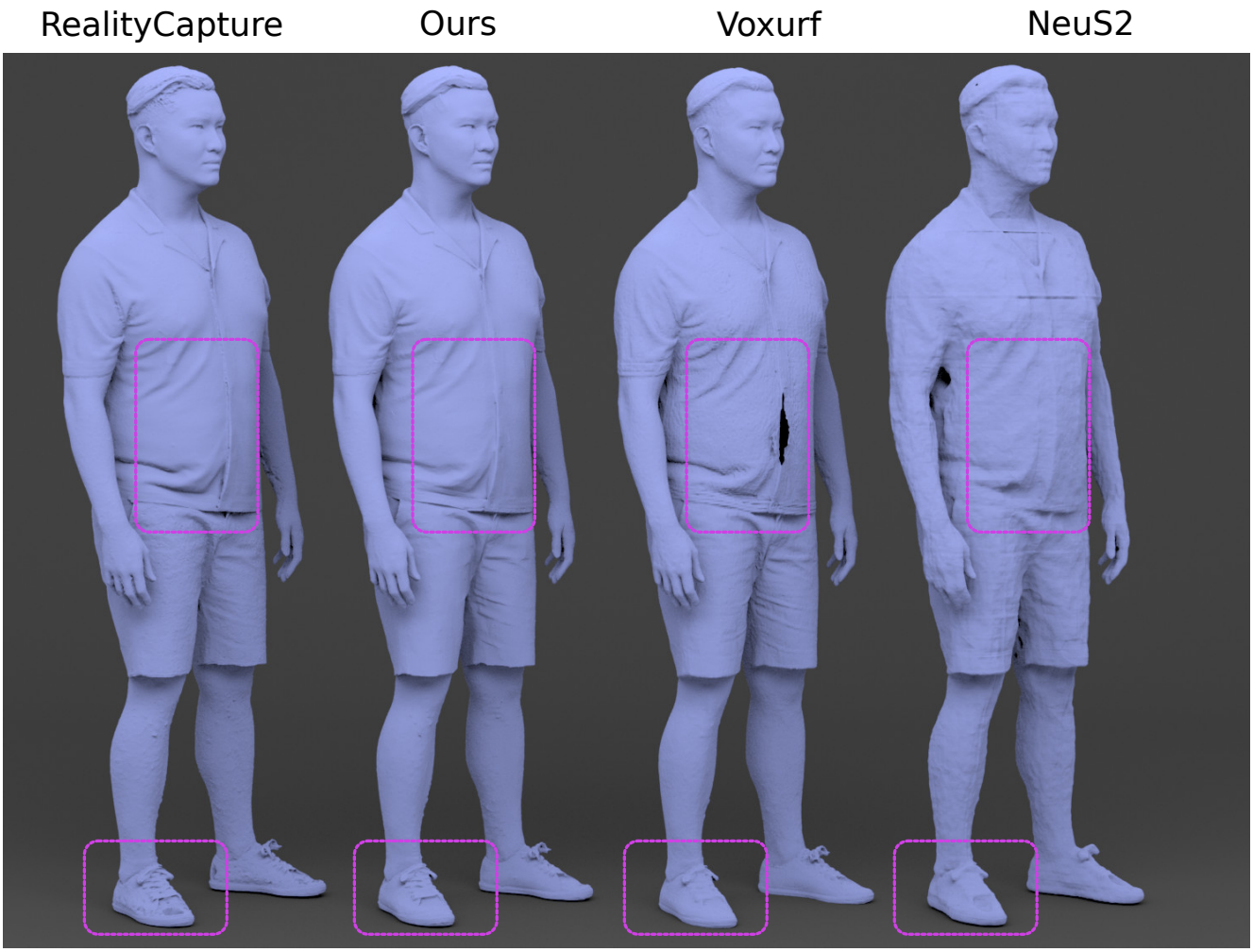

Figure 4. Reconstruction results on ActorsHQ. Left to right: RealityCapture, Ours, Voxurf, NeuS2.

Type	Shading	Render
MLP ✓, $\mathbf{F}_s$ ✓, $\mathbf{F}_a$ ✓	2.47	1.34
MLP ✓, $\mathbf{F}_s$ ✓, $\mathbf{F}_a$ ✗	1.92	1.34
MLP ✓, $\mathbf{F}_s$ ✗, $\mathbf{F}_a$ ✓	2.32	1.34
MLP ✓, $\mathbf{F}_s$ ✗, $\mathbf{F}_a$ ✗	1.80	1.34
$\mathbf{F}_a$ only	1.96	1.34
$\mathbf{F}_s$ only	0.82	1.34

Table 6. Timings in ms

References

[1] Realitycapture. https://www.capturingreality.com/. Accessed: 2024-11-21. 1

[2] Mustafa Işık, Martin Rünz, Markos Georgopoulos, Taras Khakhulin, Jonathan Starck, Lourdes Agapito, and Matthias Nießner. Humanrf: High-fidelity neural radiance fields for humans in motion. *ACM Transactions on Graphics (TOG)*, 42(4):1–12, 2023. 1

[3] Rasmus Jensen, Anders Dahl, George Vogiatzis, Engil Tola, and Henrik Aanæs. Large scale multi-view stereopsis evaluation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 406–413. IEEE, 2014. 2

[4] Yao Yao, Zixin Luo, Shiwei Li, Jingyang Zhang, Yufan Ren, Lei Zhou, Tian Fang, and Long Quan. Blendedmvs: A large-scale dataset for generalized multi-view stereo networks. *Computer Vision and Pattern Recognition (CVPR)*, 2020. 3

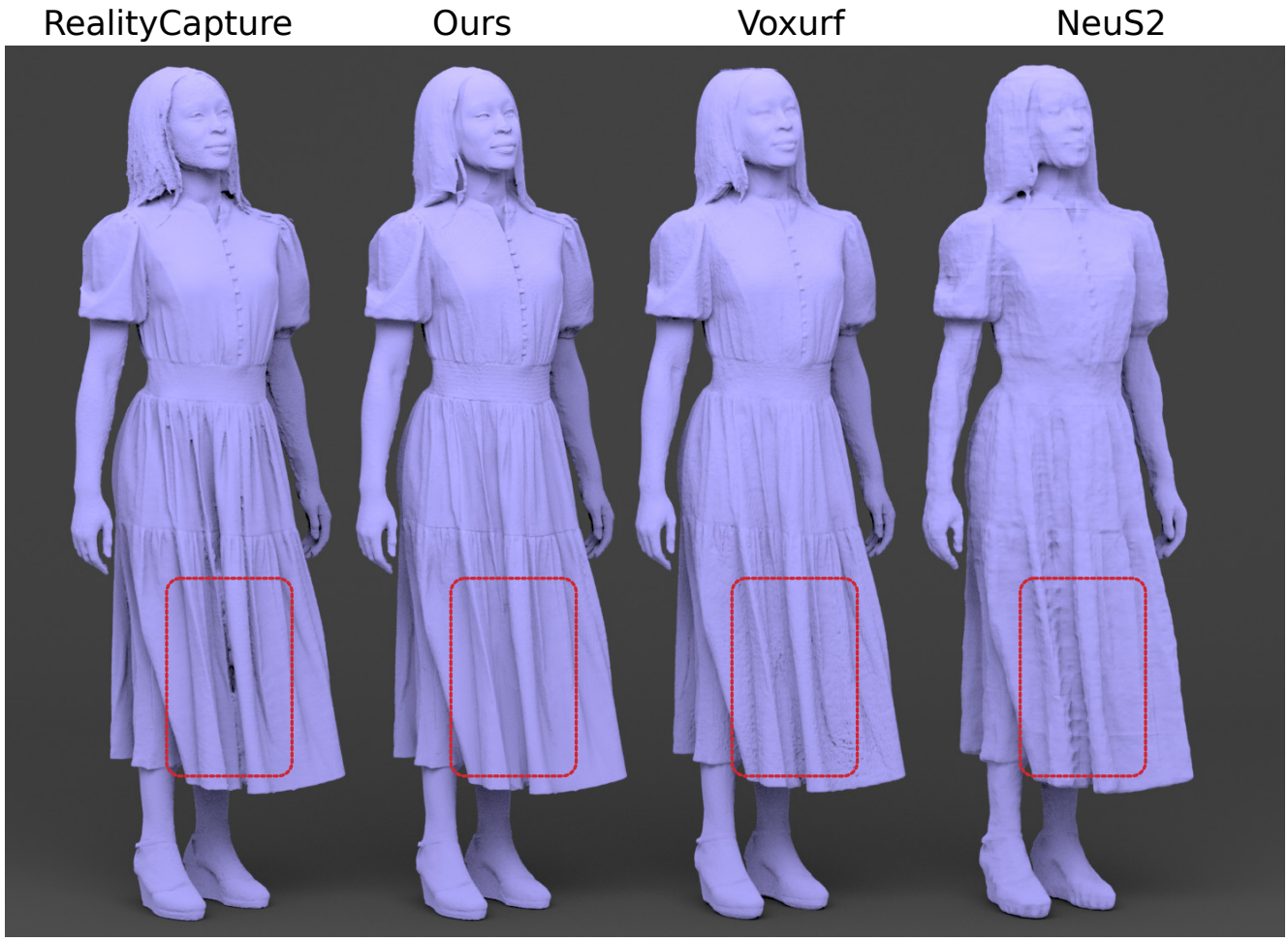

Figure 5. Reconstruction results on ActorsHQ. Left to right: RealityCapture, Ours, Voxurf, NeuS2.

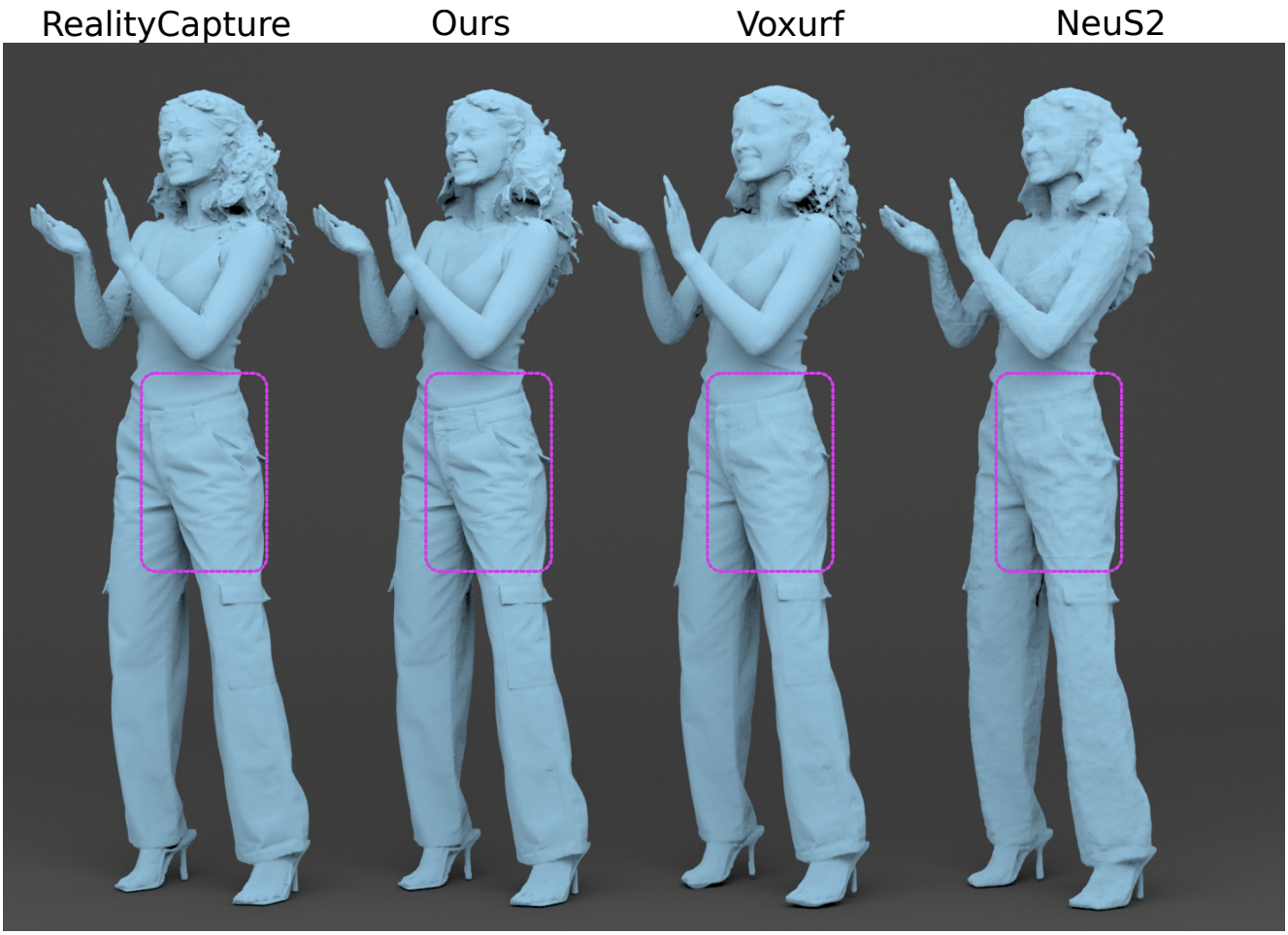

Figure 6. Reconstruction results on ActorsHQ. Left to right: RealityCapture, Ours, Voxurf, NeuS2.

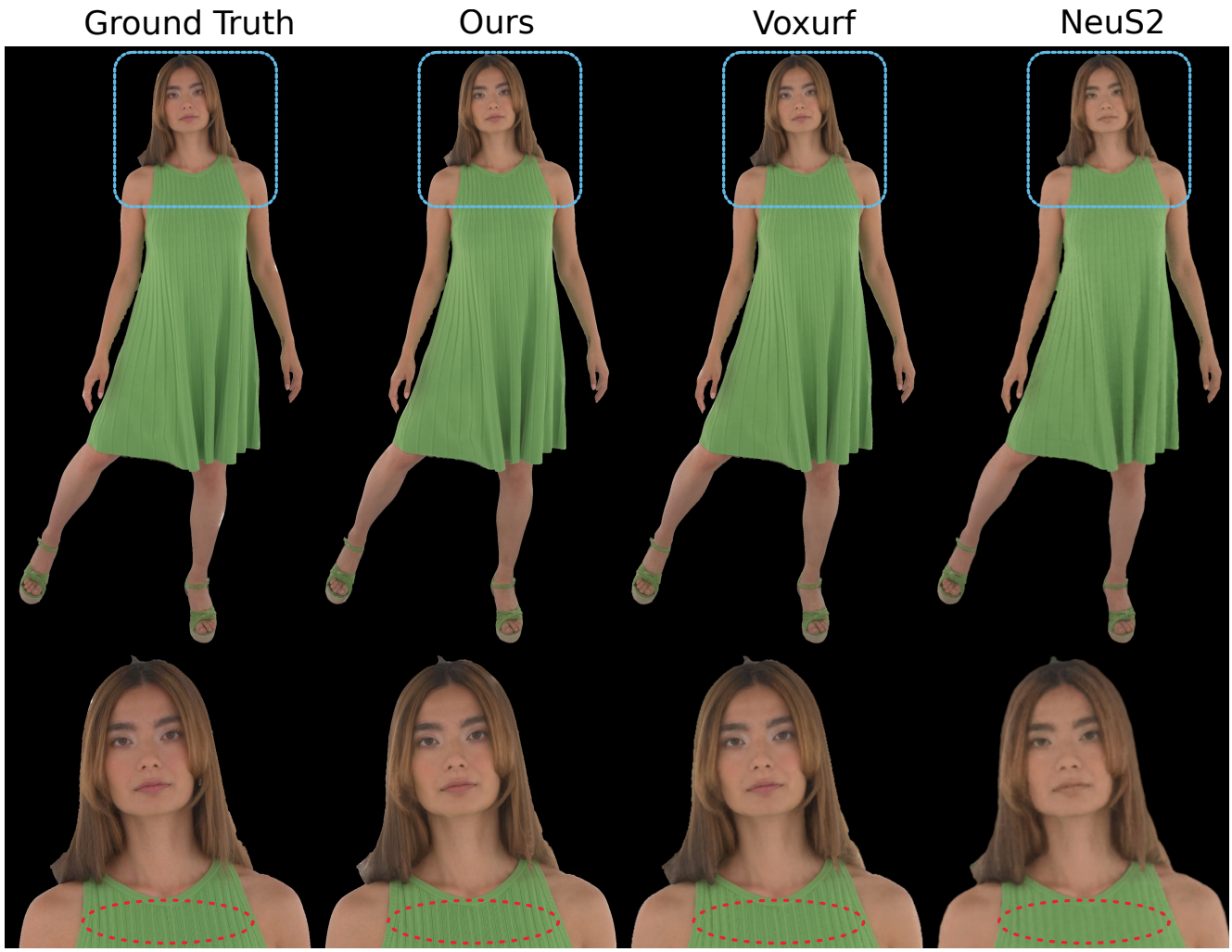

Figure 7. Qualitative comparison on ActorsHQ. Left to right: ground truth, Ours, Voxurf, NeuS2. Our method is able to handle the full resolution images, which enables to reconstruct the sewing patterns at a sub-millimetric scale.

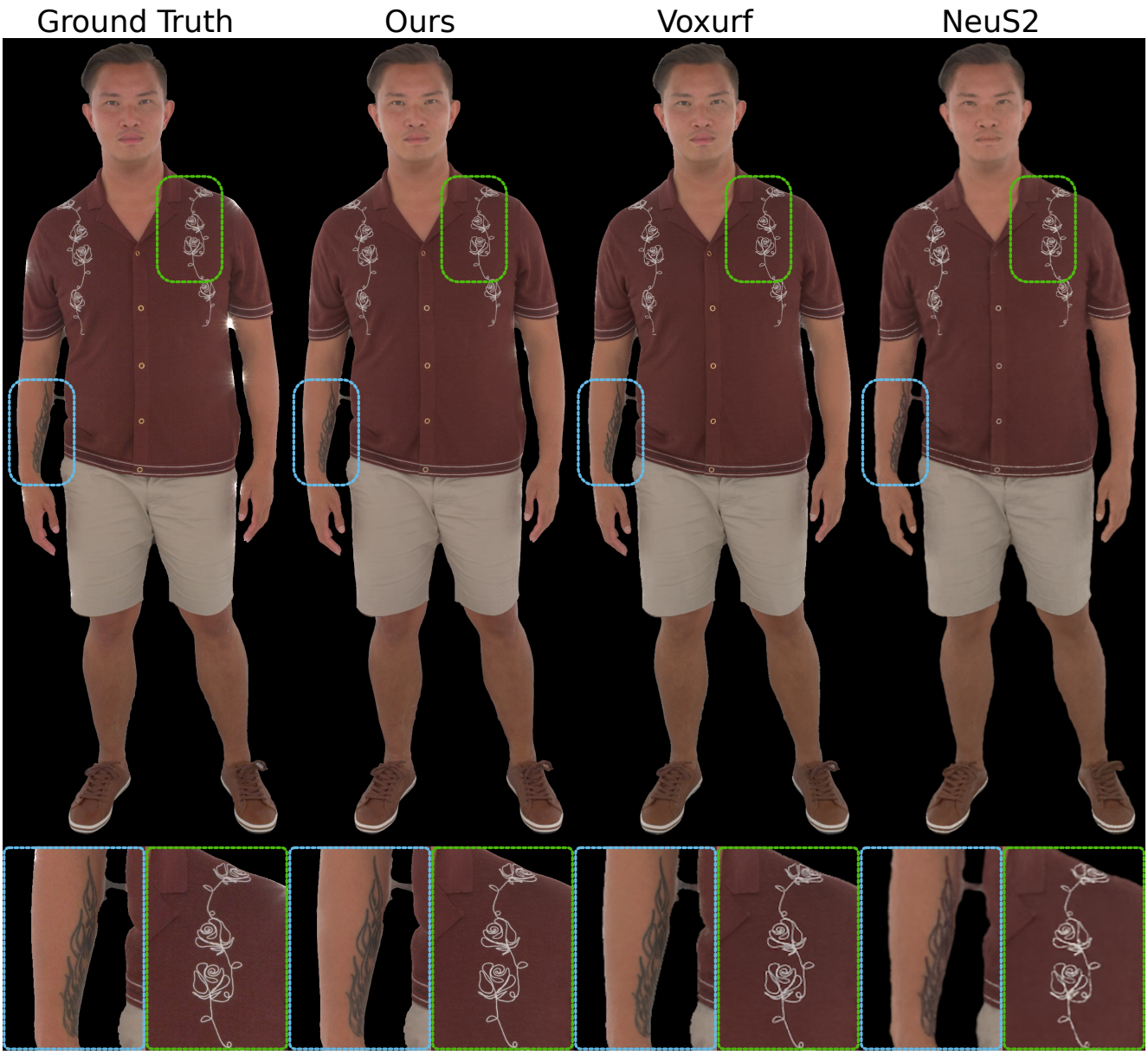

Figure 8. Qualitative comparison on ActorsHQ. Left to right: ground truth, Ours, Voxurf, NeuS2.

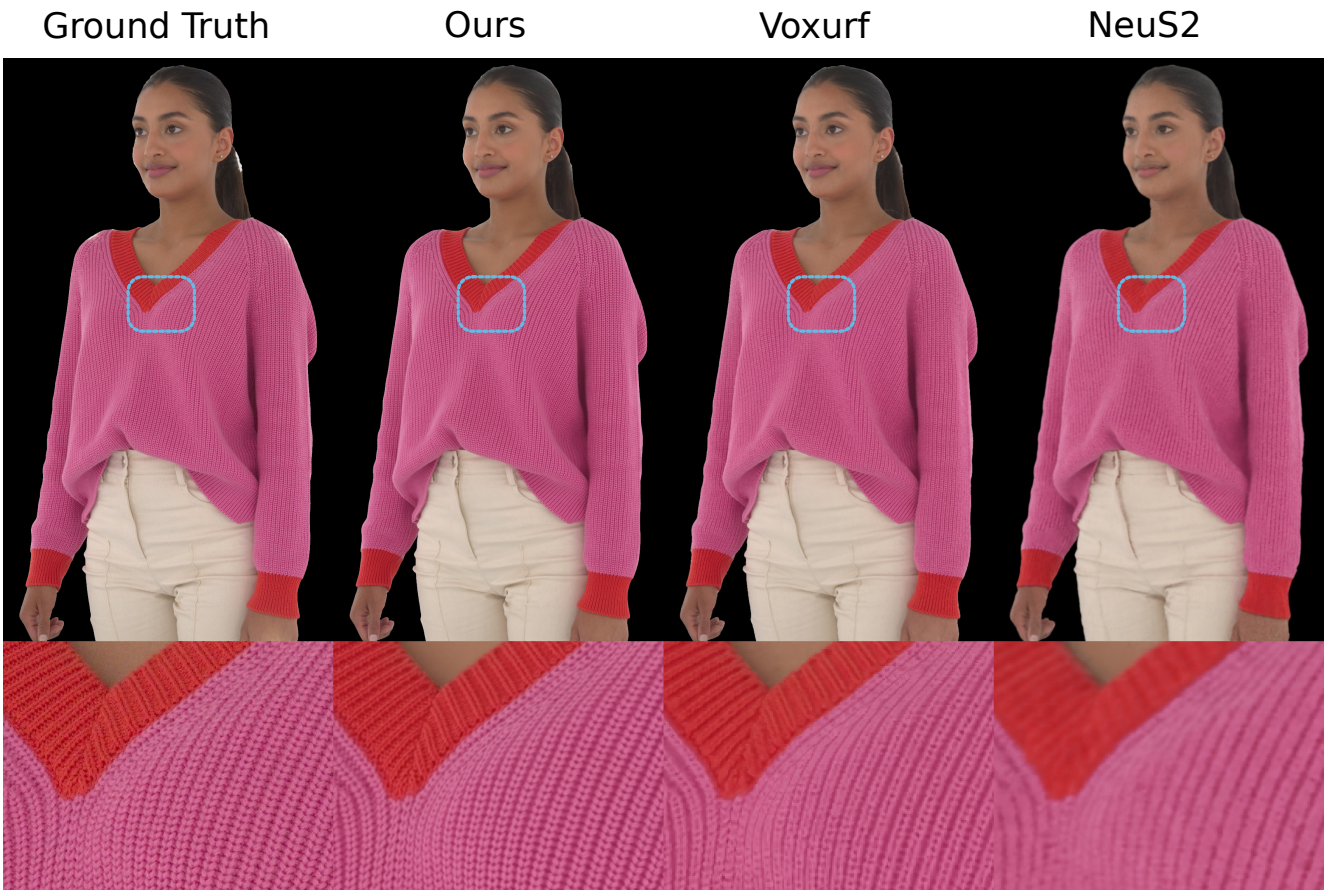

Figure 9. Qualitative comparison on ActorsHQ. Left to right: ground truth, Ours, Voxurf, NeuS2.

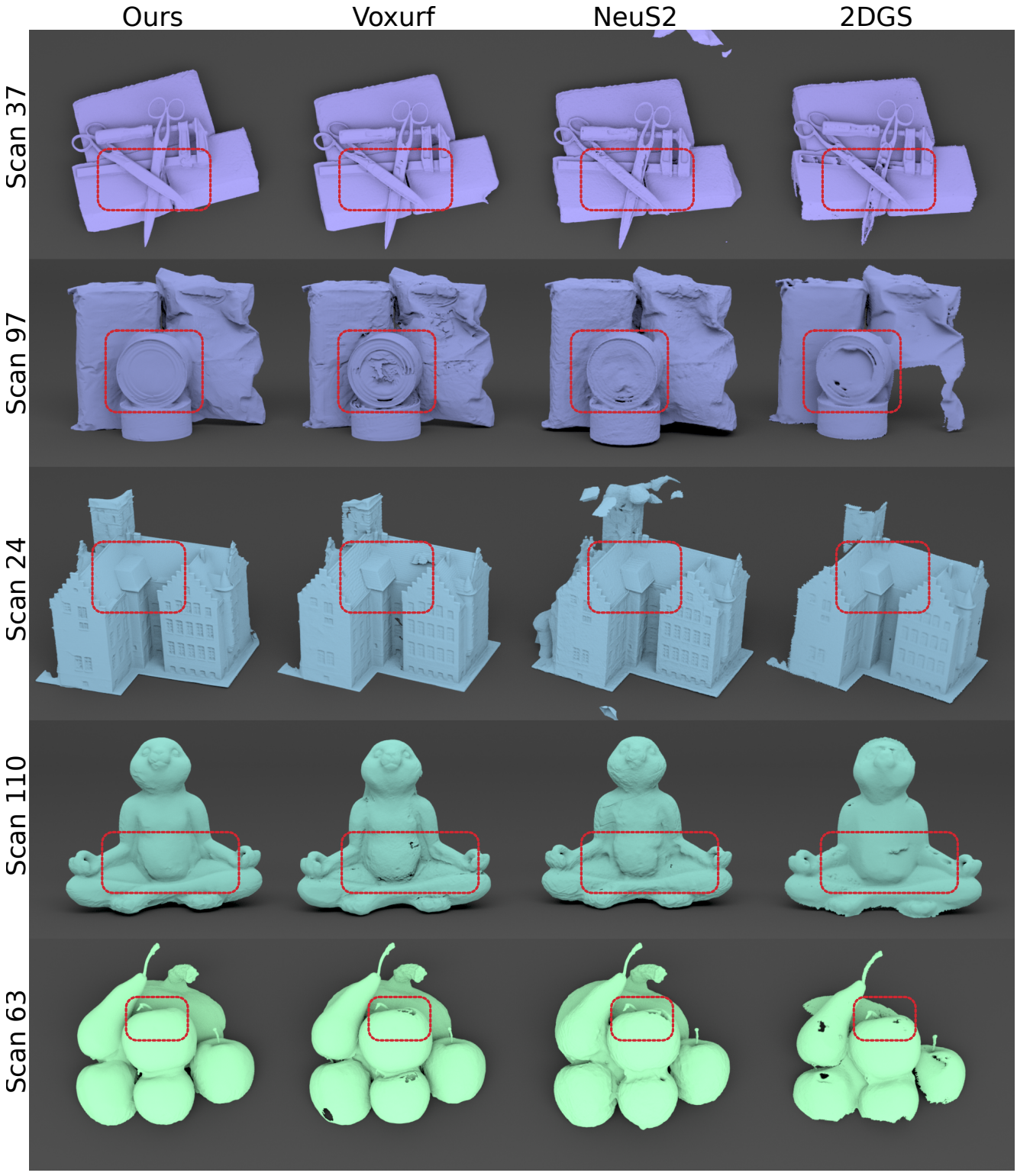

Figure 10. Reconstruction results on DTU. Left to right: Ours, Voxurf, NeuS2, 2DGS. We find that 2DGS excels at reconstructing flat surfaces (doll house roof) but tends to under-perform on reflective materials. 2DGS fails to extract geometry on some parts of the objects (scans 97 and 63). In contrast, our method recovers smooth surfaces even under strong specularities (metal scissors, tuna can and apples). Voxurf struggles on the most shiny materials despite its considerably larger MLP. NeuS2's reconstruction suffers from grid-aligned artifacts, possibly due to discontinuities in its hash-grid interpolation scheme (shoulder of the bunny in scan 110).

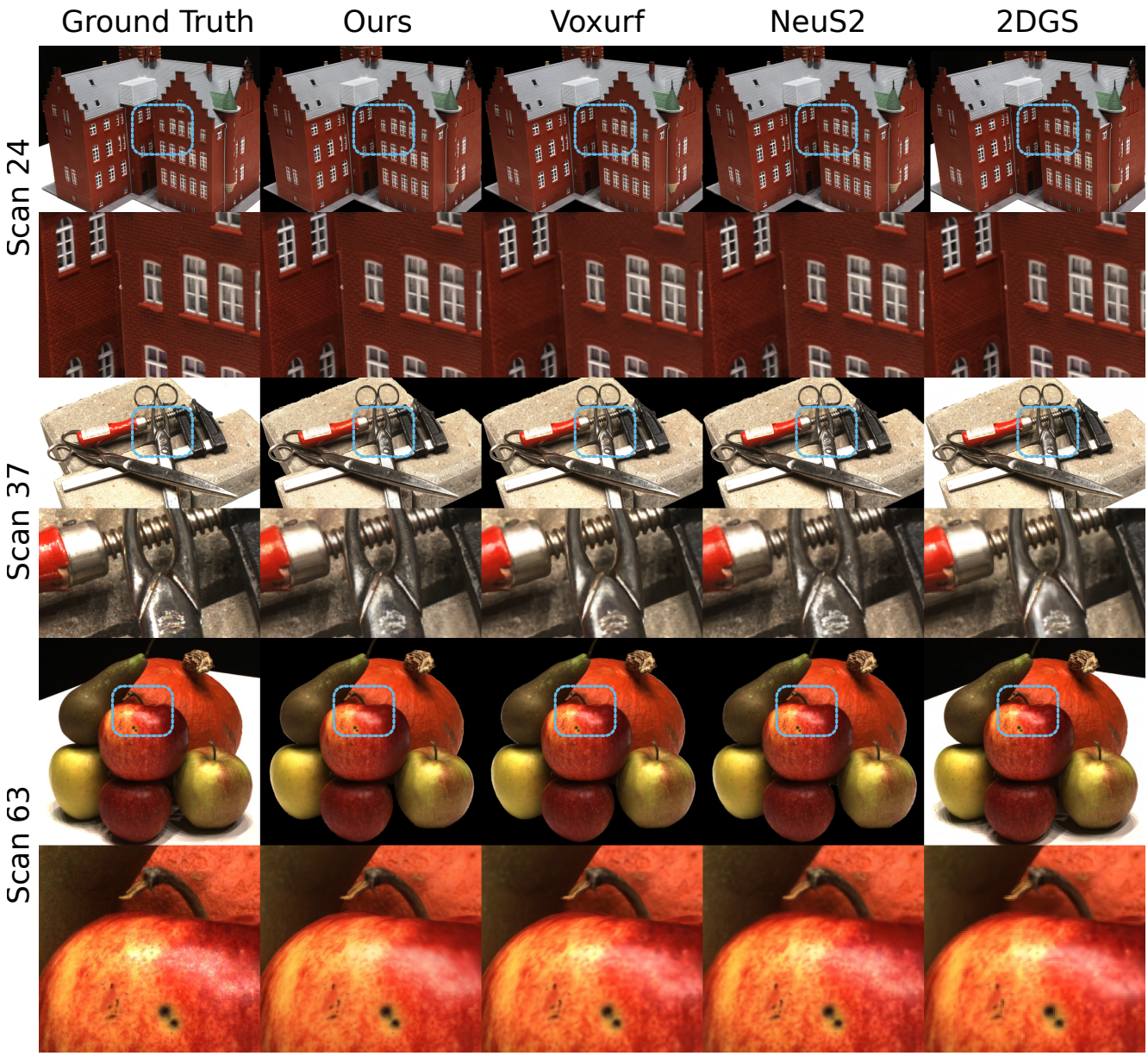

Figure 11. Qualitative comparison on DTU. Top: scan 24, middle: scan 37, bottom: scan 63. Left to right: ground truth, Ours, Voxurf, NeuS2, 2DGS.

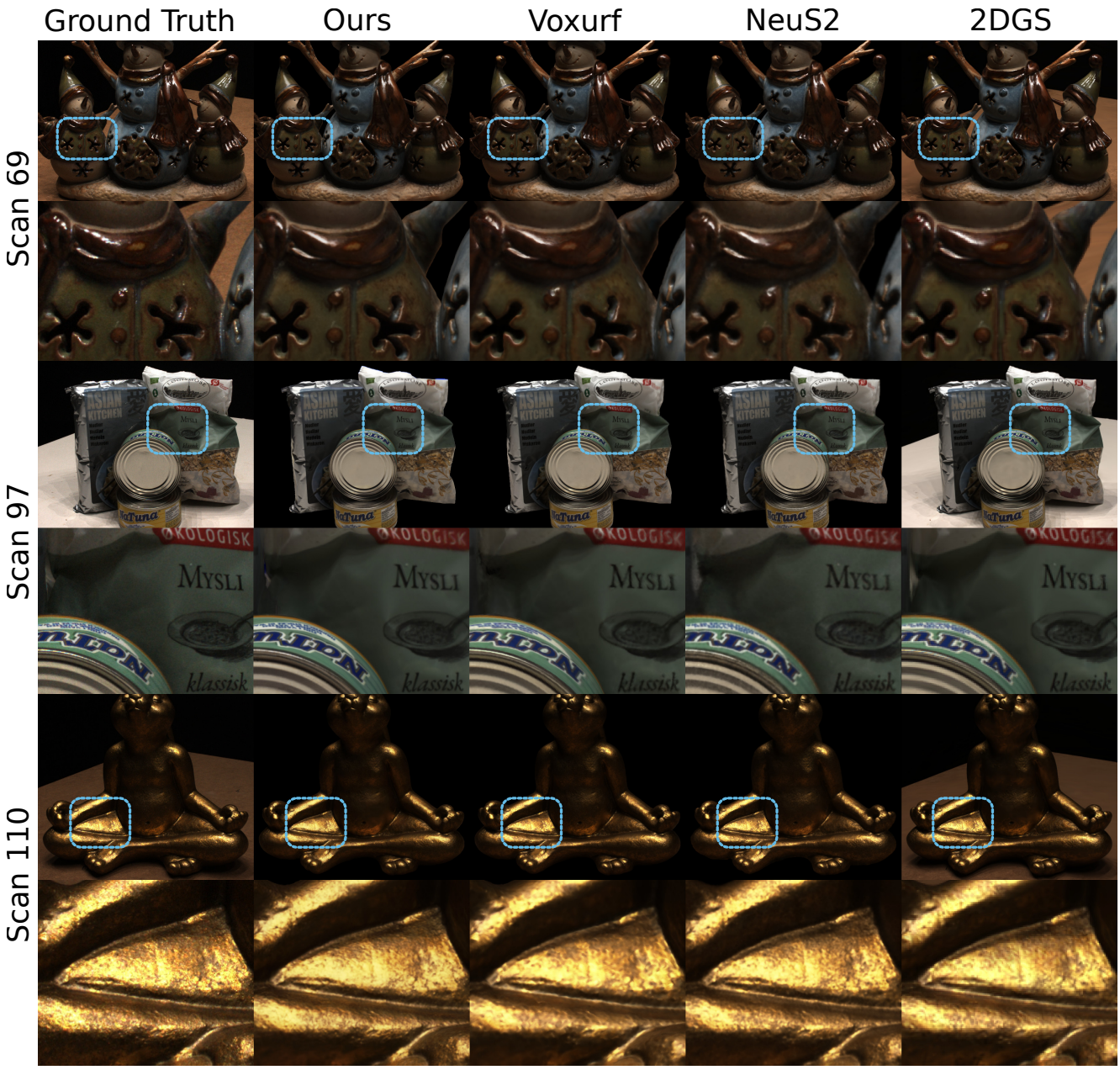

Figure 12. Qualitative comparison on DTU. Top: scan 69, middle: scan 97, bottom: scan 110. Left to right: ground truth, Ours, Voxurf, NeuS2, 2DGS.

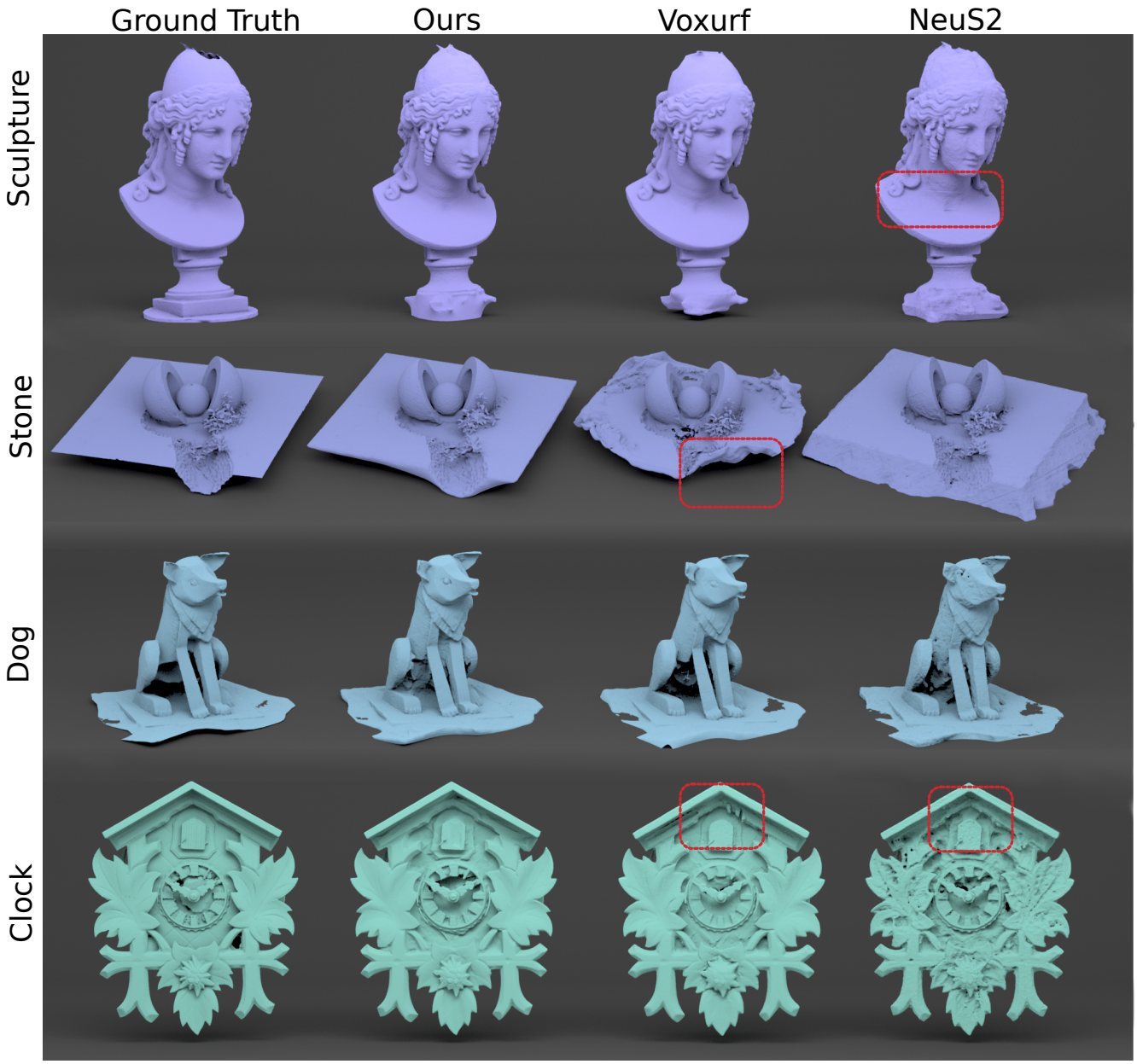

Figure 13. Reconstruction results on BlendedMVS. Left to right: Ground Truth, Ours, Voxurf, NeuS2.

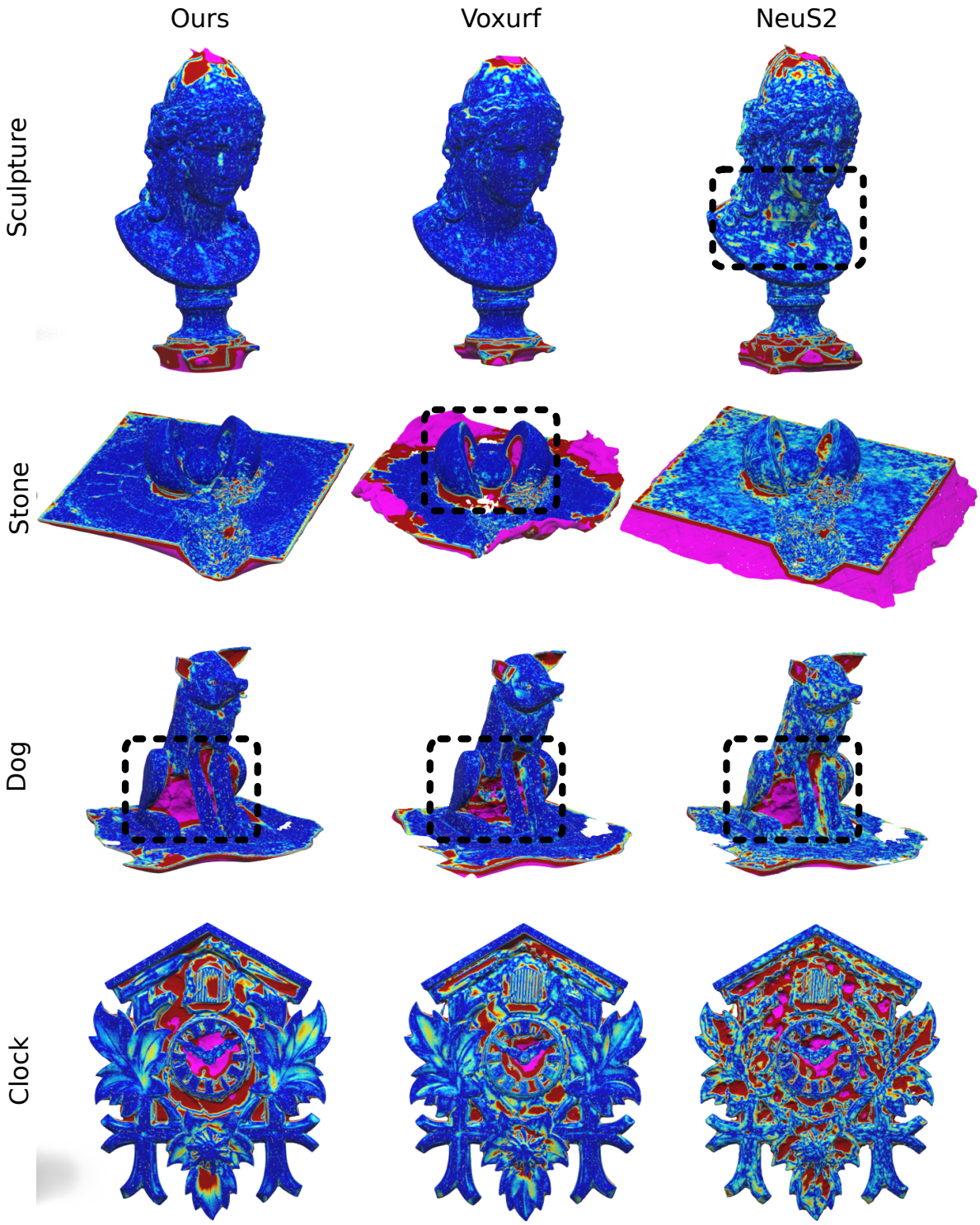

Figure 14. Accuracy heatmaps on BlendedMVS. The pink color denotes points too far away from the ground truth, which are ignored in the computation of the metrics. Left to right: Ours, Voxurf, NeuS2. Voxurf fails to carve inside the two hemispheres in the stone example and the corners of the base are missing. However, Voxurf is able to carve under the dog statue whereas both NeuS2 and our method fail on this example. NeuS2 is noticably more noisy on all examples.

Figure 15. Qualitative comparison on BlendedMVS. Left to right: ground truth, Ours, Voxurf, NeuS2.

		kinette								kino					
Chamfer (mm)	Mean	cos	naked	jea	opt1	opt2	opt3	sho	tig	cos	naked	jea	opt	sho	tig
(4,4,4)✗	**1.04**	**1.51**	**0.54**	**1.09**	**1.14**	**1.75**	**1.22**	**0.95**	**0.66**	**1.53**	0.60	**0.75**	**1.51**	**0.67**	**0.67**
MMH	1.15	1.55	**0.54**	1.10	1.16	2.12	1.45	1.07	0.68	1.77	**0.59**	0.80	1.72	0.75	0.75
Voxurf	1.59	1.70	1.28	1.62	1.62	2.06	1.61	1.48	0.98	2.60	1.30	1.27	2.54	1.12	1.06
Neus2	2.13	3.71	1.43	2.10	2.13	2.90	2	2.12	1.25	3.55	1.61	1.64	2.61	1.43	1.37
Colmap	3.51	2.69	4.27	2.99	4.59	4.18	2.81	3.71	2.50	4.11	4.20	2.48	4.34	3.09	3.23
2DGS	3.35	5.05	2.27	8.30	3.25	3.40	3.63	2.66	3.52	4.18	2.25	1.75	2.80	1.98	1.87

Table 7. MVMannequins per-scene chamfer (mm)

		kinette								kino					
PSNR (db)	Mean	cos	naked	jea	opt1	opt2	opt3	sho	tig	cos	naked	jea	opt	sho	tig
(4,4,4)✗	**36.81**	**29.48**	**40.61**	**30.73**	**40.48**	**38.69**	**34.13**	**37.70**	**31.74**	**39.64**	**41.57**	**35.80**	**39.29**	**37.50**	**37.99**
MMH	36.33	29.24	40.03	30.55	39.94	38.01	33.82	37.18	31.38	39.22	40.88	35.25	38.59	37.11	37.49
Voxurf	35.51	28.39	37.51	30.20	39.26	37.34	32.82	36.73	30.81	38.17	40.09	34.45	37.57	36.69	37.16
Neus2	34.22	28.09	37.23	29.44	36.98	35.55	32.58	34.98	30.14	36.65	38.05	33.39	35.78	35.01	35.23
2DGS	34.89	27.26	37.92	28.88	37.62	36.84	32.22	36.24	29.86	37.97	38.82	34.99	37.28	36.14	36.50

Table 8. MVMannequins per-scene PSNR

		kinette								kino					
Chamfer (mm)	Mean	cos	naked	jea	opt1	opt2	opt3	sho	tig	cos	naked	jea	opt	sho	tig
(8,8,4)✗	**1.03**	**1.48**	**0.54**	**1.08**	1.12	**1.70**	1.24	**0.93**	**0.65**	1.52	**0.59**	0.74	1.51	**0.66**	0.67
(12,12,4)✗	1.04	1.52	0.54	1.09	1.13	1.72	1.23	0.96	0.66	1.52	0.59	**0.74**	1.52	0.68	**0.67**
(4,4,1)✗	1.30	1.56	1.22	1.22	1.25	1.78	1.62	1.50	1.05	1.78	1.16	0.81	1.48	0.86	0.95
(4,4,2)✗	1.04	1.48	0.59	1.09	**1.10**	1.71	**1.22**	0.94	0.67	**1.51**	0.64	0.74	**1.45**	0.69	0.69
(4,4,3)✗	1.04	1.49	0.56	1.09	1.12	1.73	1.22	0.96	0.66	1.52	0.60	0.75	1.50	0.68	0.68
(4,4,4)✗	1.04	1.51	0.54	1.09	1.14	1.75	1.22	0.95	0.66	1.53	0.60	0.75	1.51	0.67	0.67

Table 9. Detailed Ablation Table. MVMannequins per-scene chamfer (mm)

		kinette								kino					
PSNR (db)	Mean	cos	naked	jea	opt1	opt2	opt3	sho	tig	cos	naked	jea	opt	sho	tig
(8,8,4)✗	36.90	29.56	40.76	**30.80**	40.62	38.80	34.23	37.76	**31.86**	39.76	41.66	35.88	39.35	37.56	38.07
(12,12,4)✗	**36.93**	**29.57**	**40.79**	30.79	**40.63**	**38.82**	**34.24**	**37.78**	31.83	**39.79**	**41.73**	**35.88**	**39.41**	**37.58**	**38.20**
(4,4,1)✗	35.84	29.24	38.61	30.52	39.03	37.47	33.55	36.24	31.24	38.81	39.46	35.37	38.68	36.58	36.98
(4,4,2)✗	36.53	29.31	40.09	30.62	40.07	38.38	33.95	37.42	31.55	39.35	41.06	35.67	39.01	37.26	37.71
(4,4,3)✗	36.68	29.38	40.39	30.66	40.30	38.54	34.05	37.56	31.62	39.50	41.37	35.72	39.14	37.41	37.86
(4,4,4)✗	36.81	29.48	40.61	30.73	40.48	38.69	34.13	37.70	31.74	39.64	41.57	35.80	39.29	37.50	37.99

Table 10. Detailed Ablation Table. MVMannequins per-scene PSNR

PSNR	Resolution	Mean	A1S1	A2S1	A3S1	A4S1	A5S1	A6S1	A7S1	A8S1	A1S2	A4S2	A5S2	A6S2	A8S2
(4,4,4)✗	r/1	**37.48**	**37.64**	**38.21**	**37.58**	**35.78**	**38.36**	36.80	**37.80**	**38.09**	**37.91**	**35.59**	**38.56**	36.59	**38.32**
(4,4,4)✗	r/2	36.62	36.86	36.90	36.59	34.55	37.54	36.39	37.03	37.28	37.19	34.32	37.75	36.24	37.47
(4,4,4)✗	r/4	34.75	35.43	34.05	35.21	32.08	35.59	34.96	35.41	35.16	35.67	32.14	35.94	34.75	35.34
Voxurf	r/2	36.56	37.04	36.93	36.17	34.58	36.96	**36.85**	36.69	36.99	37.12	34.33	37.53	**36.81**	37.31
Neus2	r/2	34.53	35.22	34.67	33.28	32.50	35.20	33.86	35.43	35.41	35.71	32.52	35.38	34.02	35.72

Table 11. ActorsHQ per-scene PSNR

Chamfer (mm)	Mean	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122
(4,4,4)✓	**0.68**	0.65	0.74	0.34	**0.34**	1.02	0.71	**0.62**	1.34	**0.94**	0.70	**0.53**	0.96	0.36	0.45	0.47
(8,8,4)✓	0.71	0.56	**0.71**	0.34	**0.34**	1.38	0.74	0.64	1.34	0.95	0.70	0.54	1.07	**0.35**	0.45	0.47
Voxurf	0.73	0.76	0.72	0.67	**0.34**	0.95	**0.62**	0.79	1.35	0.96	0.74	0.61	1.17	**0.35**	**0.44**	0.49
Neus2	0.80	0.55	0.81	1.66	0.38	**0.92**	0.72	0.79	1.31	1.07	0.80	0.61	**0.89**	0.46	0.52	0.58
2DGS	0.76	**0.47**	0.82	**0.32**	0.36	1.06	0.89	0.81	**1.30**	1.23	**0.66**	0.65	1.34	0.42	0.66	**0.46**

Table 12. DTU per-scene chamfer (mm)

PSNR (db)	Mean	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122
(4,4,4)✓	37.03	35.58	30.59	35.59	35.56	38.89	38.06	34.81	39.90	33.97	38.84	40.03	36.23	34.90	40.87	41.63
(8,8,4)✓	**37.74**	**36.73**	**31.53**	**36.31**	**36.74**	39.38	38.93	35.05	40.25	**34.69**	**39.51**	**40.45**	**36.79**	**35.60**	41.59	**42.49**
Voxurf	37.08	34.97	30.70	33.82	35.02	**39.45**	**39.22**	35.48	**41.03**	34.35	38.78	39.43	35.36	35.20	**41.79**	41.69
Neus2	36.00	34.57	29.82	34.30	34.64	37.93	36.87	33.79	38.95	32.79	38.13	38.37	35.14	34.37	39.93	40.32
2DGS	36.03	35.01	30.61	34.47	33.77	38.27	36.11	**35.84**	39.53	34.26	38.33	37.86	34.82	33.46	39.10	39.05

Table 13. DTU per-scene PSNR

Chamfer (mm)	Mean	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122
(4,4,4)✗	0.71	0.68	0.82	**0.34**	0.35	1.20	0.76	**0.59**	1.34	0.91	0.74	0.57	**0.91**	0.39	0.50	0.50
(8,8,4)✓	0.71	**0.56**	**0.71**	**0.34**	**0.34**	1.38	0.74	0.64	1.34	0.95	0.70	0.54	1.07	**0.35**	0.45	**0.47**
(12,12,4)✓	0.70	0.58	0.74	**0.34**	**0.34**	1.37	0.73	0.66	1.32	**0.87**	0.71	0.54	0.93	**0.35**	0.45	**0.47**
(4,4,1)✓	0.85	0.64	0.80	0.35	**0.34**	1.79	0.71	0.77	**1.30**	1.12	**0.69**	**0.52**	2.34	0.41	**0.43**	**0.47**
(4,4,2)✓	0.69	0.64	0.74	**0.34**	**0.34**	1.09	**0.69**	0.67	1.33	1.01	**0.69**	0.53	1.02	0.37	0.44	**0.47**
(4,4,3)✓	**0.67**	0.64	0.75	**0.34**	**0.34**	1.04	0.70	0.62	1.33	0.95	**0.69**	0.53	0.92	0.36	0.44	**0.47**
(4,4,4)✓	0.68	0.65	0.74	**0.34**	**0.34**	**1.02**	0.71	0.62	1.34	0.94	0.70	0.53	0.96	0.36	0.45	**0.47**

Table 14. Detailed Ablation Table. DTU per-scene chamfer (mm)

PSNR (db)	Mean	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122
(4,4,4)✗	36.18	34.73	29.81	34.87	34.52	38.23	36.38	34.52	39.35	33.17	38.34	38.80	36	34.03	39.62	40.38
(8,8,4)✓	37.74	36.73	31.53	36.31	36.74	39.38	38.93	35.05	40.25	34.69	39.51	**40.45**	36.79	35.60	41.59	42.49
(12,12,4)✓	**38.03**	**37.20**	**31.96**	**36.64**	**36.95**	**39.55**	**39.33**	**35.71**	**40.42**	**35.27**	**39.58**	40.32	**36.97**	**35.89**	**41.97**	**42.75**
(4,4,1)✓	35.92	35.13	29.81	34.97	34.66	36.40	36.77	32.90	38.86	32.29	38.03	39.43	34.36	34	40.36	40.92
(4,4,2)✓	36.44	35.25	29.97	35.31	35.27	37.66	37.32	33.62	39.48	32.98	38.53	39.71	35.49	34.39	40.53	41.11
(4,4,3)✓	36.76	35.36	30.27	35.40	35.46	38.46	37.66	34.40	39.71	33.59	38.67	39.80	35.95	34.67	40.69	41.32
(4,4,4)✓	37.03	35.58	30.59	35.59	35.56	38.89	38.06	34.81	39.90	33.97	38.84	40.03	36.23	34.90	40.87	41.63

Table 15. Detailed Ablation Table. DTU per-scene PSNR

Chamfer	Mean	dog	bear	clock	durian	man	sculpture	stone	jade
(4,4,4)✓	2.36	2.51	2.27	1.90	3.63	**1.79**	**1.61**	**1.30**	3.88
(8,8,4)✓	**2.21**	**2.24**	**2.03**	**1.69**	3.26	1.81	**1.61**	1.36	**3.71**
Voxurf	2.64	2.28	2.20	1.88	**2.98**	2.11	1.75	3.96	3.99
Neus2	2.93	2.78	2.71	2.63	4.23	2.25	2.50	1.91	4.43

Table 16. BlendedMVS per-scene chamfer

PSNR	Mean	dog	bear	clock	durian	man	sculpture	stone	jade
(4,4,4)✓	35.19	35.49	30.55	34.68	31.28	42.94	40.80	30.84	34.92
(8,8,4)✓	**35.89**	**36.30**	**30.96**	**35.42**	**31.65**	**43.72**	**41.34**	31.01	**36.69**
Voxurf	35.11	35.24	30.88	34.49	29.69	43.35	41.06	30.23	35.93
Neus2	33.62	34.56	29.99	31.04	29.21	40.88	39.10	**31.36**	32.79

Table 17. BlendedMVS per-scene PSNR

Chamfer	Mean	dog	bear	clock	durian	man	sculpture	stone	jade
(4,4,4)✗	2.47	2.18	2.71	2.05	3.81	1.97	1.75	1.36	3.95
(8,8,4)✓	**2.21**	2.24	2.03	**1.69**	**3.26**	1.81	1.61	1.36	**3.71**
(12,12,4)✓	2.31	**2.10**	2.31	1.86	3.58	1.96	**1.59**	1.34	3.74
(4,4,1)✓	2.58	3.27	**2.02**	2.44	3.88	1.92	1.88	1.33	3.93
(4,4,2)✓	2.35	2.59	2.03	2.04	3.66	1.80	1.62	**1.29**	3.75
(4,4,3)✓	2.36	2.56	2.35	1.94	3.59	**1.79**	1.64	1.31	3.74
(4,4,4)✓	2.36	2.51	2.27	1.90	3.63	**1.79**	1.61	1.30	3.88

Table 18. Detailed Ablation Table. BlendedMVS per-scene chamfer

(4,4,4)✗	34.76	35.71	30.35	33.76	31	42.65	40.08	30.84	33.72
(8,8,4)✓	35.89	36.30	30.96	35.42	31.65	43.72	41.34	31.01	36.69
(12,12,4)✓	**36.17**	**36.50**	**30.97**	**36.55**	**32.12**	**43.75**	**41.41**	**31.04**	**37.00**
(4,4,1)✓	34.44	34.48	30.34	32.81	31.12	42.44	39.89	30.75	33.72
(4,4,2)✓	34.86	35	30.53	33.85	31.38	42.71	40.40	30.82	34.19
(4,4,3)✓	35.05	35.18	30.64	34.34	31.36	42.90	40.69	30.82	34.48
(4,4,4)✓	35.19	35.49	30.55	34.68	31.28	42.94	40.80	30.84	34.92

Table 19. Detailed Ablation Table. BlendedMVS per-scene PSNR