DualPrim: Compact 3D Reconstruction with Positive and Negative Primitives – Supplementary Material –

Xiaoxu Meng[1*] xiaoxumeng.cs@gmail.com
Zhongmin Chen[4,5*] chenzhongmin19@mails.ucas.ac.cn
Bo Yang[2] yangbo@waymo.com
Weikai Chen[1†] chenwk891@gmail.com
Weixiao Liu[3] weixiaoliu@lucidmotors.com
Lin Gao[4,5] gaolin@ict.ac.cn

[1]Independent Researcher [2]Waymo LLC [3]Lucid Motors
[4]Institute of Computing Technology, Chinese Academy of Sciences
[5]School of Advanced Interdisciplinary Science, University of Chinese Academy of Sciences

1. Supplementary Result

The supplementary material provides extended evaluations complementing the main paper. We compare DualPrim with a range of state-of-the-art methods across 5 categories: volume rendering-based methods RNb-NeuS (RNS) [2], tetrahedron-based method nvdiffrec (NVD) [11], primitive-based differentiable rendering method Differentiable Blocks World (DBW) [9], shape abstraction methods EMS [6], Marching Primitives (MP) [7], and PrimitiveAnything (PA) [14], and topology optimization methods Flexicubes (FC) [12] and MeshAnything (MA) [4].

Table 1 reports per-category Chamfer Distances on 12 ShapeNet [3] categories, with 15 objects randomly sampled per category to ensure coverage of diverse structures and geometries. DualPrim generally achieves lower reconstruction error, demonstrating its ability to recover compact and structured geometry from multi-view images.

Figure 1 and Figure 2 provides additional qualitative comparisons on the ShapeNet dataset and Figure **??** provides qualitative comparisons on objects from Objaverse and Sketchfab. Connected components are rendered in different colors for better visualization. RNS [2] and NVD [11] tend to produce over-tessellated surfaces due to grid-based Marching Cubes/Tetrahedrons extraction; FC [12] preserves global structure but struggles with fine details; MP [7], EMS [6], PA [14], and MA [4] generate compact meshes but the topologies are sensitive to SDF noise; and DBW [9] has difficulty modeling complex surfaces due to its limited primitive capacity. In contrast, DualPrim produces compact, structurally coherent meshes while preserving both global and fine-scale geometry.

These supplementary quantitative and qualitative results further validate the effectiveness of our method.

2. Implementation Details

Initialization. We initialize $K = 100$ super-primitives, with their parameters summarized in Table 2. To ensure overlap between PSQs and NSQs, the NSQs are initialized using the parameters of the PSQs.

The MLP used for lighting prediction contains 4 layers with a hidden width of 64, using ReLU activations for hidden layers and Sigmoid activation at the output to maintain bounded predictions.

Training Details. Each object is optimized independently and requires approximately 5 hours on an NVIDIA GeForce RTX 3090 GPU. A promising direction for future work is to accelerate training by reducing the number of ray samples, potentially achieving 45× speedups [13] through techniques such as occupancy-grid empty-space skipping (Instant-NGP [10], DVGO [13]) and hierarchical or proposal-based sampling (NeRF [8], Mip-NeRF 360 [1]).

References

[1] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5470–5479, 2022.

[2] Baptiste Brument, Robin Bruneau, Yvain Quéau, Jean Mélou, François Lauze, Jean-Denis Durou, and Lilian Calvet. Rnb-neus: Reflectance and normal-based multi-view 3d reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

*Equal contribution
†Corresponding author, this paper solely reflects the author's personal research and is not associated with the author's affiliated institution.

Category	CD ($\times 10^{-2}$)								
	Ours	RNS [2]	NVD [11]	FC [12]	PA [14]	MP [7]	MA [4]	EMS [6]	DBW [9]
airplane	**4.70**	4.73	4.72	5.76	6.09	8.38	7.54	10.81	8.39
bed	**9.27**	17.38	16.69	21.45	14.94	17.00	13.57	21.46	18.86
bench	**7.01**	16.13	15.80	20.32	11.52	21.61	11.31	21.98	21.00
bookshelf	9.25	12.51	13.04	13.23	15.89	12.99	**8.84**	17.63	15.46
bottle	**9.63**	10.89	11.45	12.25	13.09	11.63	9.72	12.68	11.51
camera	8.16	8.08	**7.40**	11.34	11.58	11.79	12.29	13.22	14.49
chair	**8.22**	12.88	12.25	16.01	10.47	14.34	9.16	22.21	22.41
clock	**8.59**	9.25	9.01	9.36	13.34	8.69	10.75	12.66	17.33
faucet	**5.11**	6.18	6.00	8.26	6.23	7.62	5.61	12.23	11.33
lamp	**5.71**	6.93	6.79	16.56	10.83	12.84	11.82	15.61	12.43
sofa	**10.10**	10.84	11.75	11.47	11.17	10.41	11.66	18.19	14.14
table	**8.77**	12.56	12.48	13.98	13.64	12.82	12.83	24.04	22.10

Table 1. Quantitative comparison of DualPrim (ours) with 8 representative methods on 12 categories.

Parameter Name	**Shape**	**Init Value**
(a_x, a_y, a_z)	$(2, 3)$	0.1
(ϵ_1, ϵ_2)	$(2, 2)$	1.0
α	(1)	1.0
θ	(1)	0.5
T	$(2, 3)$	$U(0, 1)$
R	$(2, 3)$	1
basic color c_{basic}	(3)	$U(0, 1)$

Table 2. The initial super-primitive parameter values for DualPrim.

[3] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An Information-Rich 3D Model Repository. Technical Report arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015.

[4] Yiwen Chen, Tong He, Di Huang, Weicai Ye, Sijin Chen, Jiaxiang Tang, Xin Chen, Zhongang Cai, Lei Yang, Gang Yu, Guosheng Lin, and Chi Zhang. Meshanything: Artist-created mesh generation with autoregressive transformers, 2024.

[5] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 Conference Papers*, pages 1–11, 2024.

[6] Weixiao Liu, Yuwei Wu, Sipu Ruan, and Gregory S. Chirikjian. Robust and Accurate Superquadric Recovery: a Probabilistic Approach . In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2666–2675, Los Alamitos, CA, USA, 2022. IEEE Computer Society.

[7] Weixiao Liu, Yuwei Wu, Sipu Ruan, and Gregory Chirikjian. Marching-primitives: Shape abstraction from signed distance function. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[8] 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. In *European conference on computer vision*, pages 405–421. Springer, 2020.

[9] Tom Monnier, Jake Austin, Angjoo Kanazawa, Alexei Efros, and Mathieu Aubry. Differentiable blocks world: Qualitative 3d decomposition by rendering primitives. *Advances in Neural Information Processing Systems*, 36:5791–5807, 2023.

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

[11] Jacob Munkberg, Jon Hasselgren, Tianchang Shen, Jun Gao, Wenzheng Chen, Alex Evans, Thomas Müller, and Sanja Fidler. Extracting Triangular 3D Models, Materials, and Lighting From Images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8280–8290, 2022.

[12] Tianchang Shen, Jacob Munkberg, Jon Hasselgren, Kangxue Yin, Zian Wang, Wenzheng Chen, Zan Gojcic, Sanja Fidler, Nicholas Sharp, and Jun Gao. Flexible isosurface extraction for gradient-based mesh optimization. *ACM Trans. Graph.*, 42(4), 2023.

[13] Cheng Sun, Min Sun, and Hwann-Tzong Chen. Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5459–5469, 2022.

[14] Jingwen Ye, Yuze He, Yanning Zhou, Yiqin Zhu, Kaiwen Xiao, Yong-Jin Liu, Wei Yang, and Xiao Han. Primitiveanything: Human-crafted 3d primitive assembly generation with auto-regressive transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[15] Fenggen Yu, Zhiqin Chen, Manyi Li, Aditya Sanghi, Hooman Shayani, Ali Mahdavi-Amiri, and Hao Zhang. Capri-net: Learning compact cad shapes with adaptive primitive assembly. In *Proceedings of the IEEE/CVF conference*

on computer vision and pattern recognition, pages 11768–11778, 2022.

[16] Fenggen Yu, Qimin Chen, Maham Tanveer, Ali Mahdavi Amiri, and Hao Zhang. D2csg: Unsupervised learning of compact csg trees with dual complements and dropouts. *Advances in Neural Information Processing Systems*, 36: 22807–22819, 2023.

[17] Haocheng Yuan, Adrien Bousseau, Hao Pan, Quancheng Zhang, Niloy J Mitra, and Changjian Li. Diffcsg: Differentiable csg via rasterization. In *SIGGRAPH Asia 2024 Conference Papers*, pages 1–10, 2024.

Figure 1. Qualitative comparison on ShapeNet objects (part 1). Each object occupies two rows. The first row shows input, Ours, PA, MA, MP, EMS, and NVD. The second row shows FC, RNS, CAPRI, D^{2}CSG, DiffCSG, 2DGS+MC, and 2DGS+CSG.

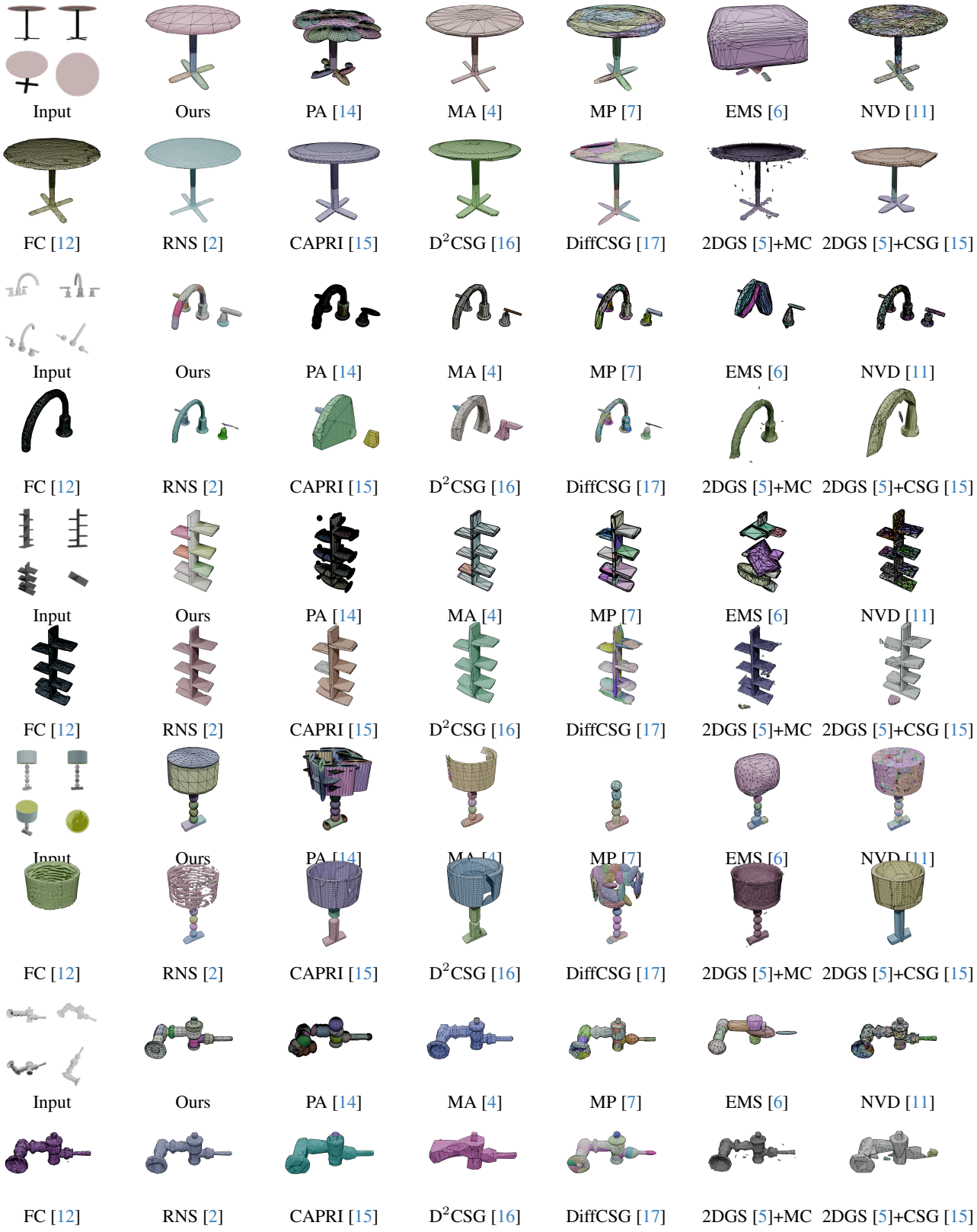

Figure 2. Qualitative comparison on ShapeNet objects (part 2). Each object occupies two rows. The first row shows input, Ours, PA, MA, MP, EMS, and NVD. The second row shows FC, RNS, CAPRI, D^{2}CSG, DiffCSG, 2DGS+MC, and 2DGS+CSG.

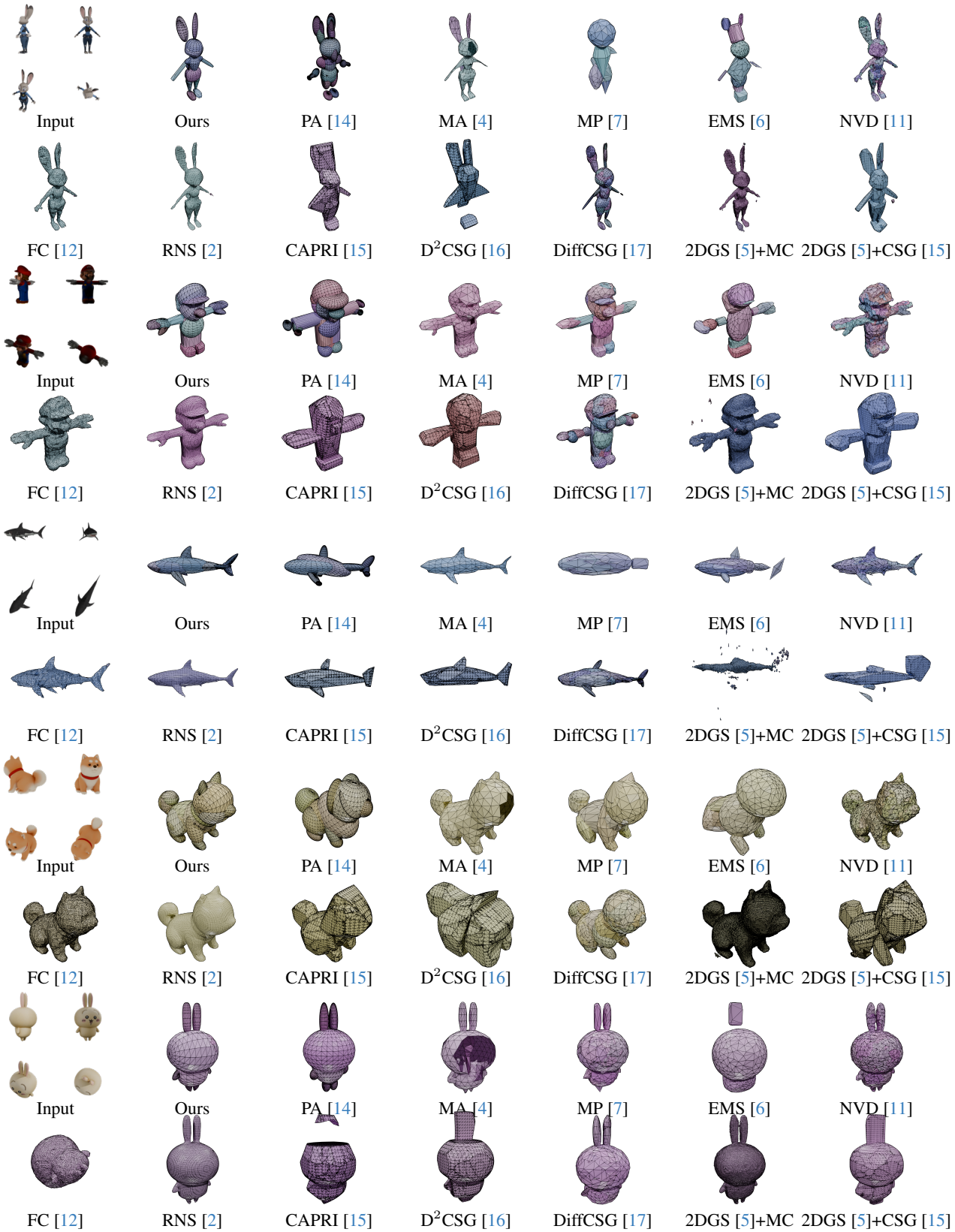

Figure 3. Qualitative comparison on in-the-wild objects. Each object occupies two rows. The first row shows input, Ours, PA, MA, MP, EMS, and NVD. The second row shows FC, RNS, CAPRI, D^{2}CSG, DiffCSG, 2DGS+MC, and 2DGS+CSG.