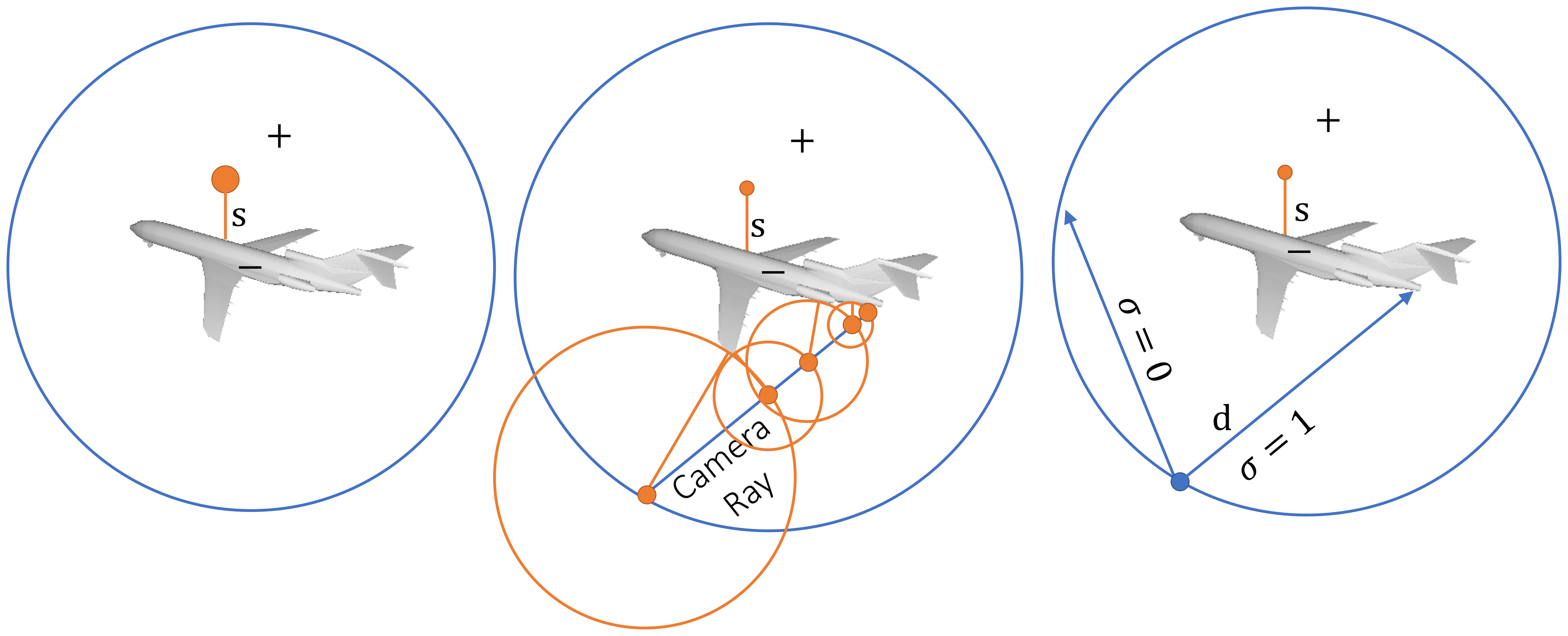
FIRe: Fast Inverse Rendering using Directional and Signed Distance Functions

Tarun Yenamandra[1], Ayush Tewari[2], Nan Yang[1], Florian Bernard[3], Christian Theobalt[4], Daniel Cremers[1]
[1]TU Munich, MCML [2]Massachusetts Institute of Technology [3]University of Bonn [4]MPI Informatics, SIC

Highlights

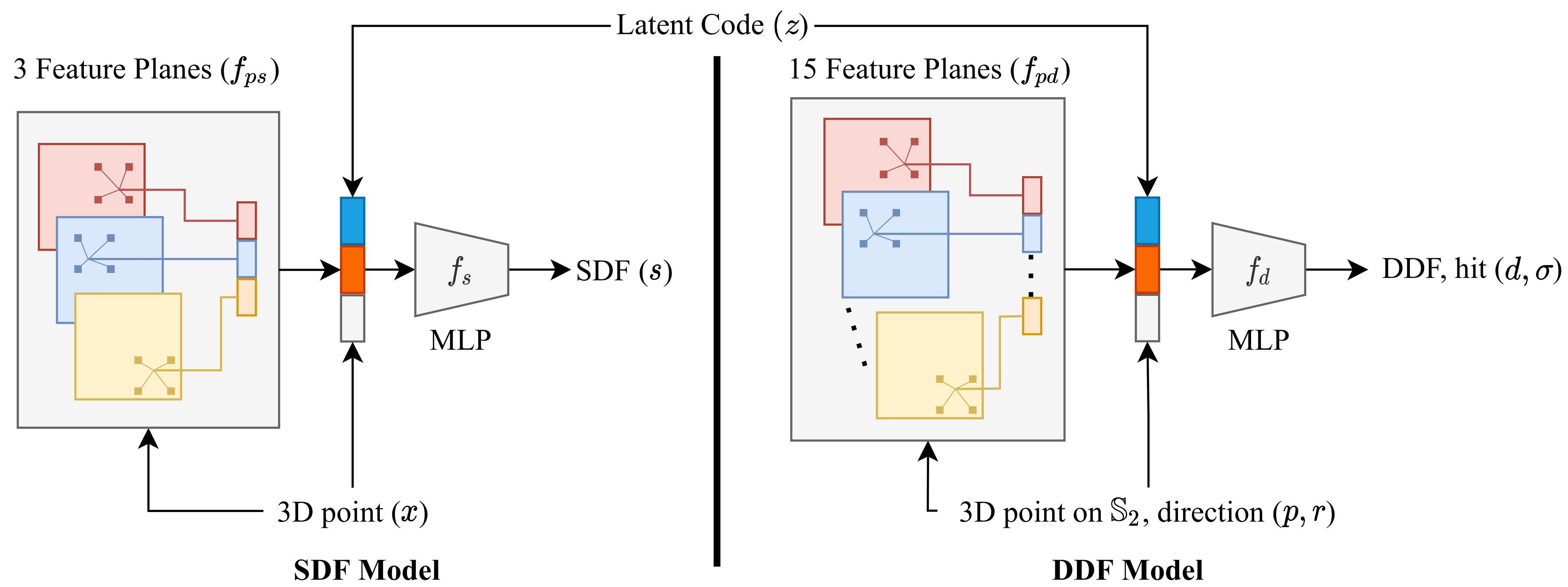
- We present a novel 3D neural scene representation called directional distance function (DDF)
- We learn a signed distance function (SDF) along with DDF to represent a class of shapes
- We propose a novel inverse rendering algorithm that is 15.5x (per iteration) faster than **the state of the art** while more accurately reconstructing 3D shapes from single-view depth maps

Motivation



SDF representation | Sphere tracing for rendering an object represented using SDF | Our representation combines DDFs and SDFs for faster rendering

Models



Training

Loss = L1 SDF Supervision	$\lVert SDF - SDF_{GT} \rVert_1$
+ L1 DDF Supervision	$\lVert DDF - DDF_{GT} \rVert_1$
+ DDF Mask	$BCE(RayHit, RayHit_{GT})$
+ TrackSDF	$\lVert SDF(p + DDF.r) \rVert_1$
+ TVL1 feature plane regularizer	
+ Latent code regularizer	

Trained per class of the ShapeNet[8] dataset

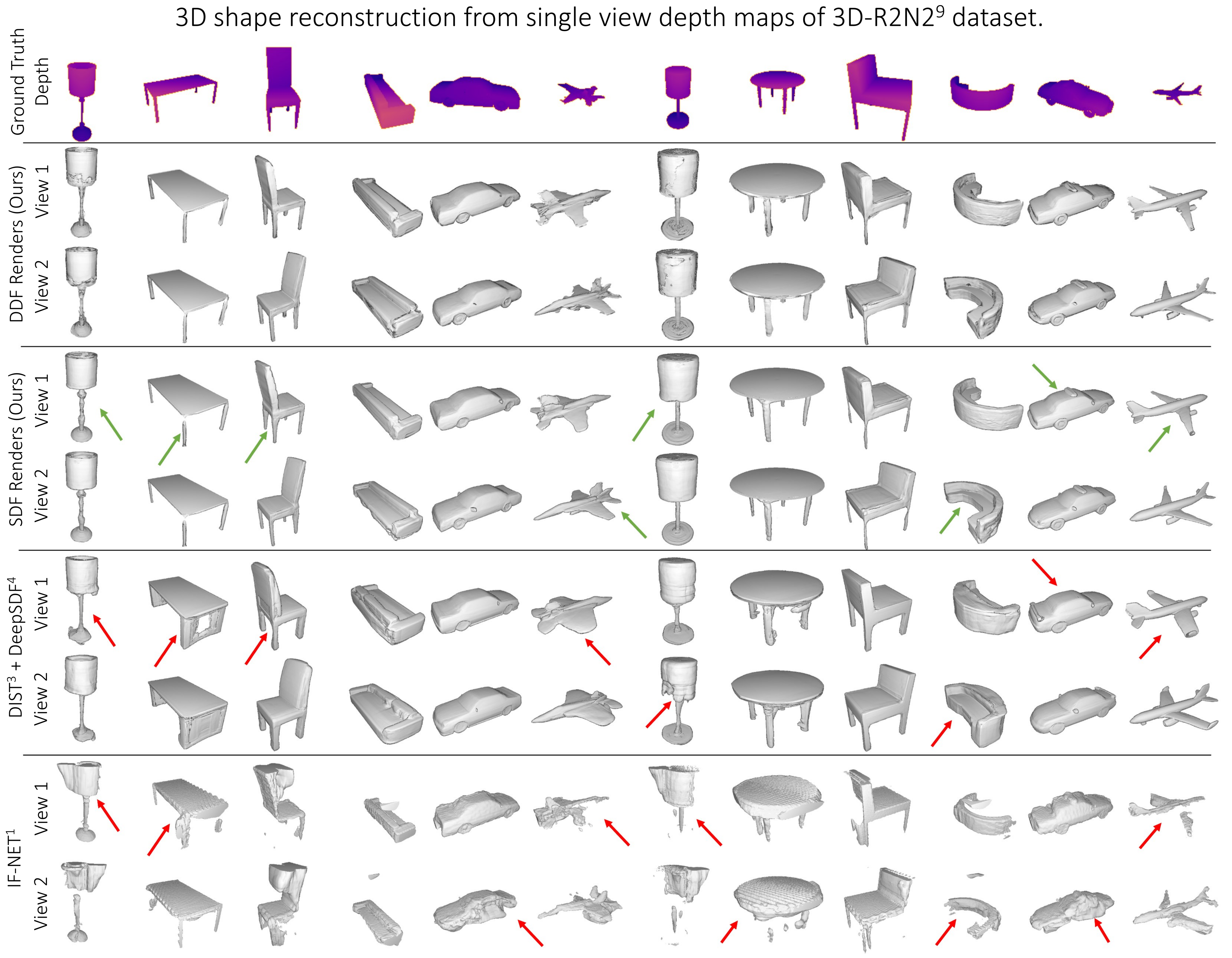
Inference: Optimization for latent code from a given depth map

Loss = Depth	$\lVert depth - depth_{GT} \rVert_1$
+ DDF Mask	$BCE(RayHit, RayHit_{GT})$
+ SDF Mask (hit)	$\lVert SDF(p + DDF.r) \rVert_1$
+ SDF Mask (miss) τ is SDF truncation thresh.	$\lVert SDF(p + DDF.r) - \tau \rVert_1$
+ Latent code regularizer	

Shapes are reconstructed from a given single-view depth map

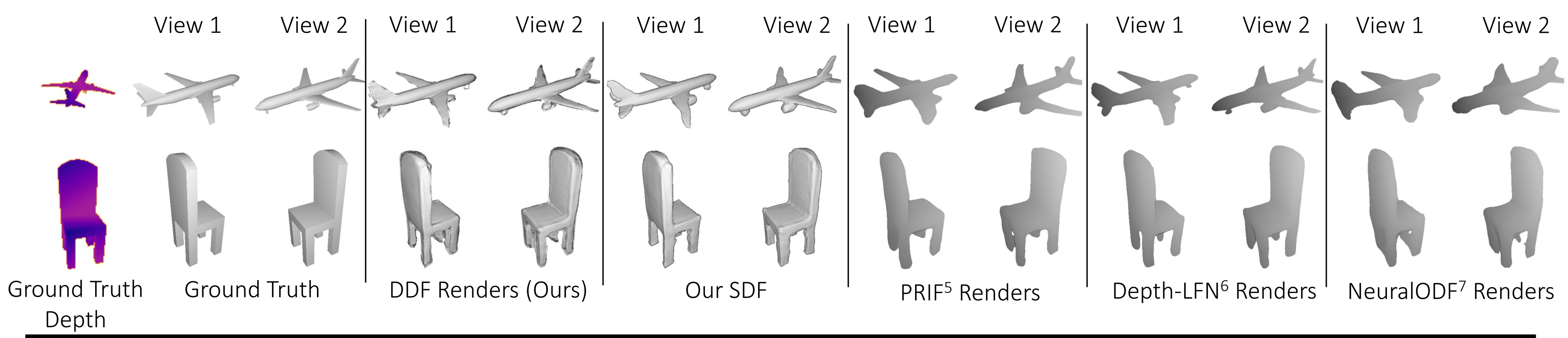
Shape from Depth

3D shape reconstruction from single view depth maps of 3D-R2N2[9] dataset.



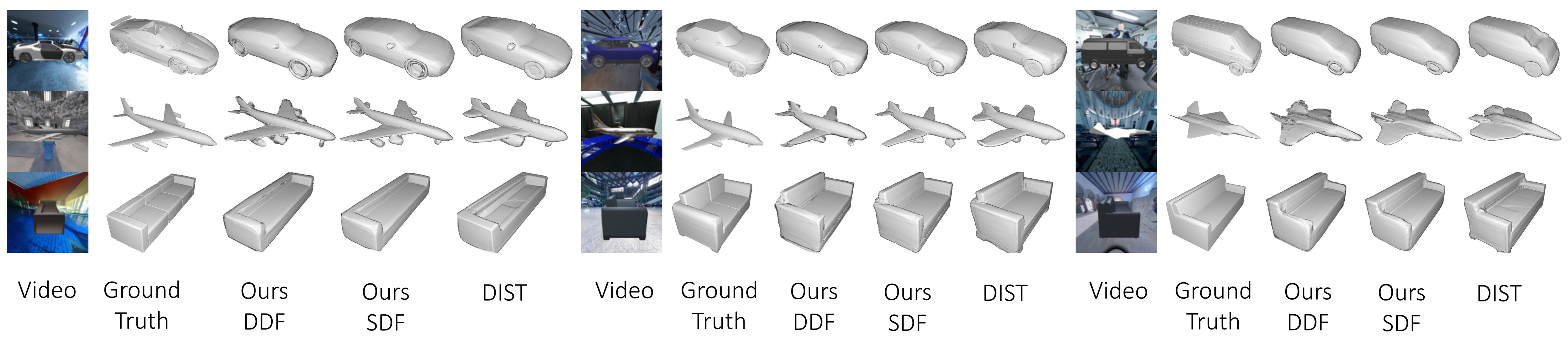
Shape from Depth

Comparison with other DDF models



Shape from Videos

3D reconstruction from videos from the PMO dataset[2].



Summary

We propose a novel neural scene representation, DDF defined on the unit sphere, for rendering images from our SDF model during inference with **1 forward pass** through the model. We present an algorithm to reconstruct 3D shapes from single view depth maps using our DDF and SDF models, which is **15.5× per iteration faster** than competing methods.

References

1. Chibane et al., "Implicit functions in feature space for 3d shape reconstruction and completion," CVPR'20.
2. Lin et al., "Photometric mesh optimization for video-aligned 3d object reconstruction," CVPR'19.
3. Liu et al., "Dist: Rendering deep implicit signed distance function with differentiable sphere tracing," CVPR'20.
4. Park et al., "Deepsdf: Learning continuous signed distance functions for shape representation," CVPR'19.
5. Feng et al., "Prif: Primary ray-based implicit function," ECCV'22.
6. Sitzmann et al., "Light field networks: Neural scene representations with single-evaluation rendering," NeurIPS'21.
7. Houchens et al., "Neuralodf: Learning omnidirectional distance fields for 3d shape representation," Arxiv'22.
8. Chang et al., "Shapenet: An information-rich 3d model repository," Arxiv'15.
9. Choy et al., "3d-r2n2: A unified approach for single and multi-view 3d object reconstruction," ECCV'16.