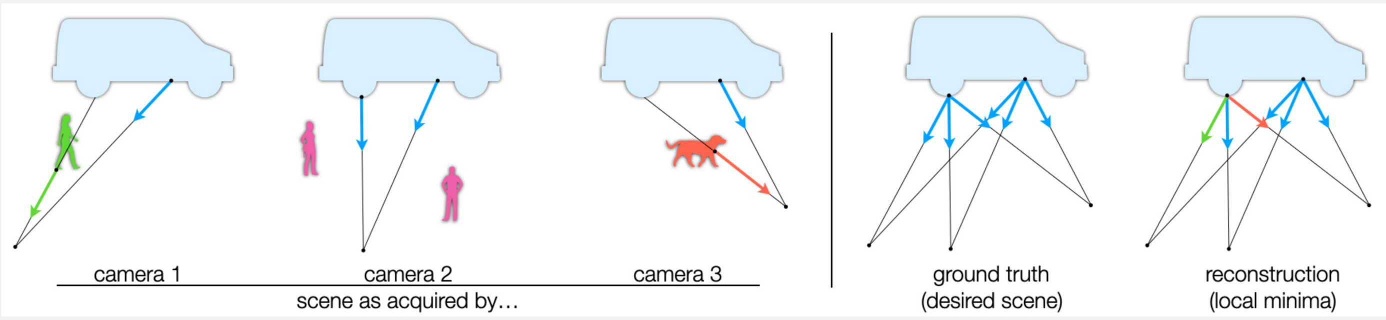
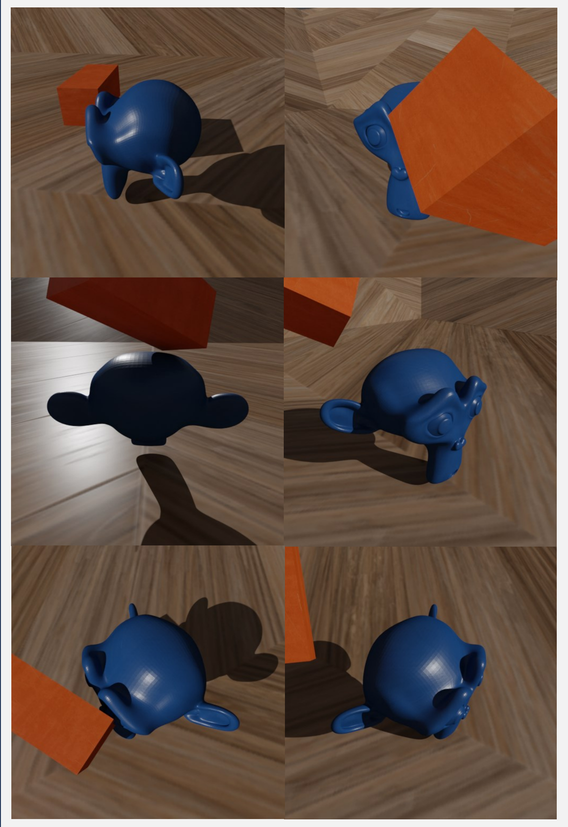
Robust Neural Surface Reconstruction

David Lennart Risch, Maximilian Oberle
Technical University Munich
TUM School of Computation, Information and Technology

Introduction



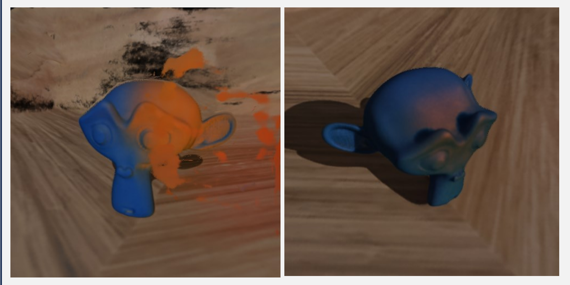
[2] Photometric inconsistencies lead to artifacts



Input views with transient distractor object

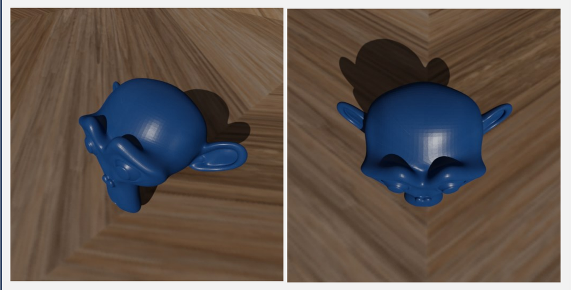
No robust loss

Distractor is interpreted as view dependent effect



Robust loss

Details preserved and distractor removed



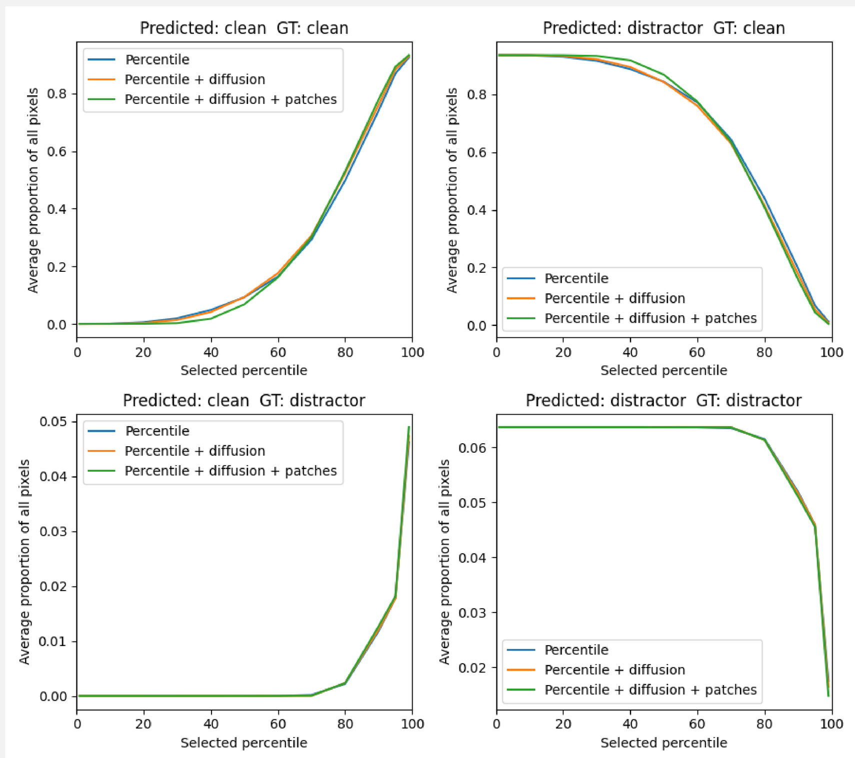
Novel views

Quantitative Results

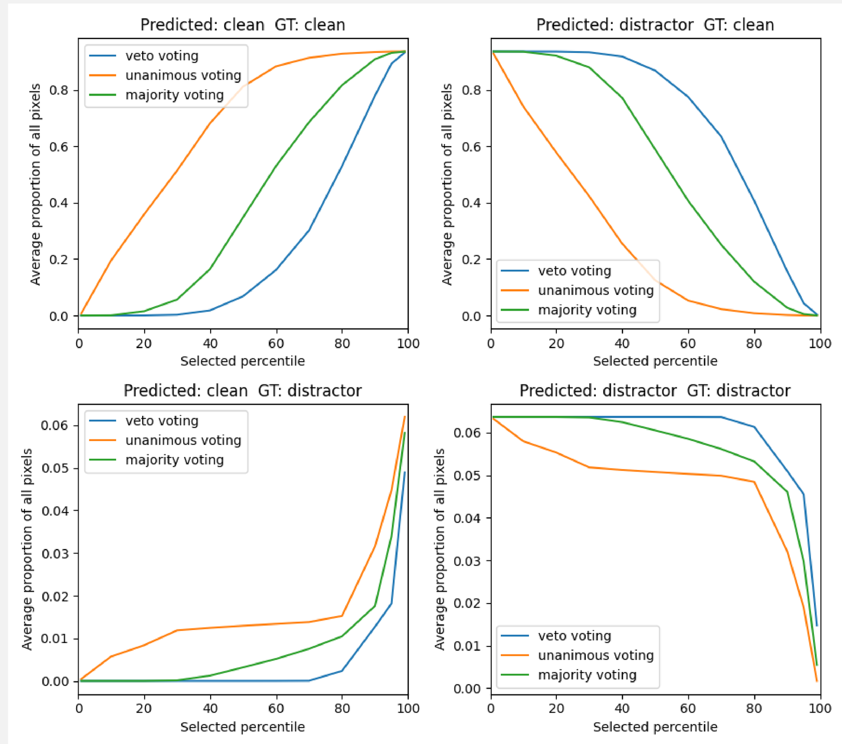
Dataset	Method \ Metric	PSNR ↑	SSIM ↑	LPIPS ↓	Chamfer distance ↓
Dataset w/o distractors	Baseline NeuS-facto [3]	**36.943**	**0.973**	**0.018**	**0.993**
	Ours: Robust loss with majority voting: percentile + diffusion + patches	33.253	0.952	0.048	1.609
Dataset with distractors	Baseline NeuS-facto [3]	29.630	0.949	0.093	**1.258**
	Ours: Robust loss per channel: percentile only	**34.722**	**0.969**	**0.021**	1.346
	Ours: Robust loss per channel: Percentile + diffusion + patches	33.362	0.954	0.051	1.424
	Ours: Robust loss with majority voting: percentile + diffusion + patches	32.679	0.950	0.048	1.423
	Ours: Robust loss with veto voting: percentile + diffusion + patches	32.406	0.951	0.054	1.483

Metric score for different models after training for 75k steps (5h on Nvidia RTX 3060)

Comparison of investigated methods

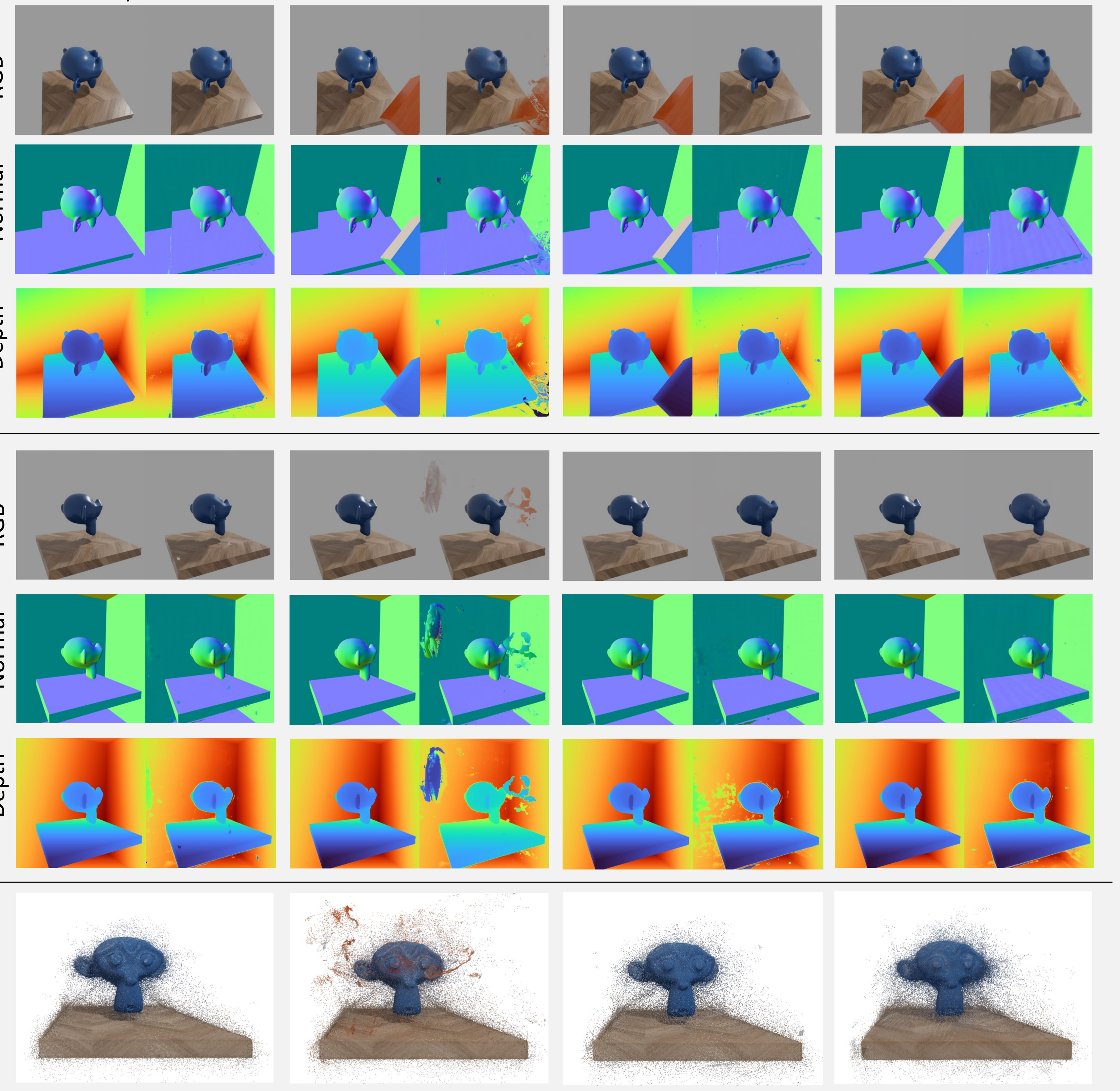


Comparison of combination methods of channels (rgb, normal, depth)



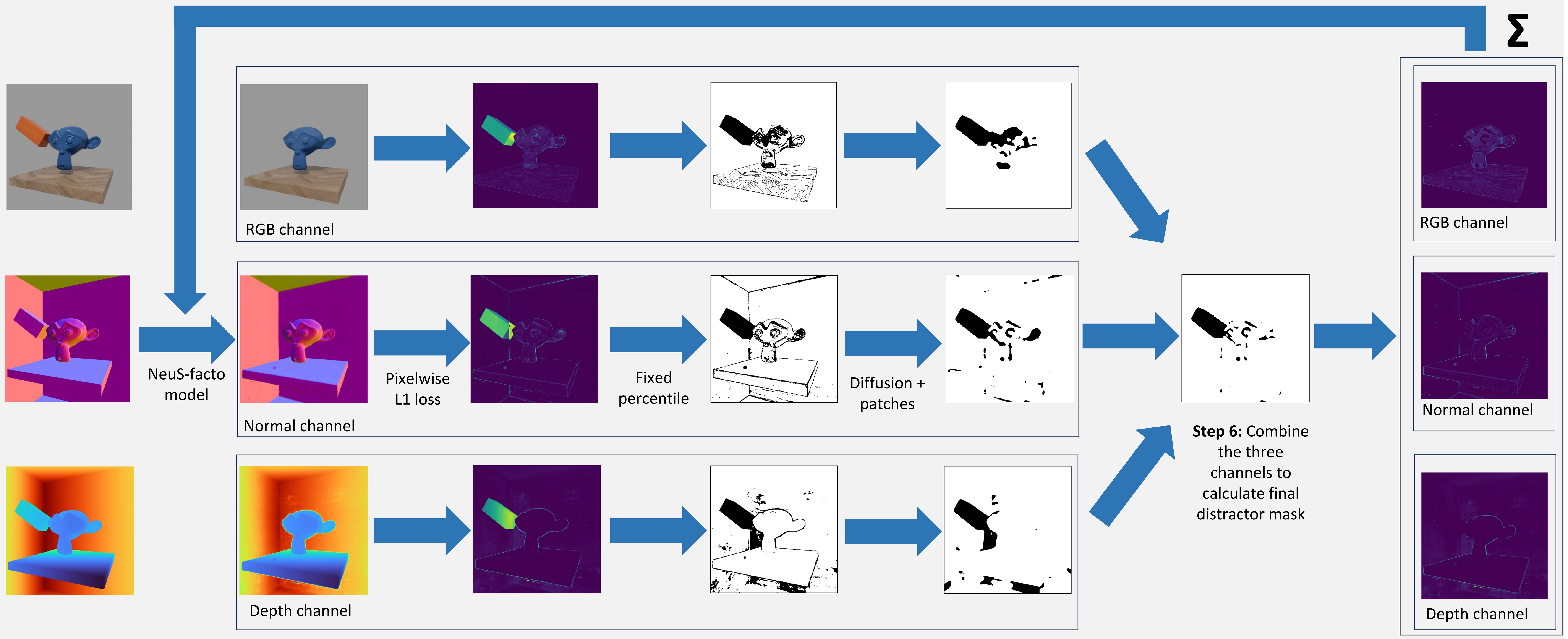
Qualitative Results

NeuS-facto baseline	NeuS-factor baseline	Percentile only (ours)	Full method* (ours)
Dataset w/o distractors	Dataset w distractors	Dataset w distractors	Dataset w distractors



*Robust loss with majority voting: percentile + diffusion + patches

Our Method



Step 1: Input rgb, normal, depth images

Step 2: Model predicts rgb, normals and depths

Step 3: Calculate pixelwise loss on each of the channels

Step 4: Estimate distractor mask using fixed percentile of loss

Step 5: Update distractor mask to improve spatially coherence

Step 7: Apply distractor mask to pixelwise loss

Summary

Improvements:
- Method improves robustness in presence of distractors while increasing reconstruction quality by using monocular-priors
- The estimate of which pixels are distractors is improved by considering all three channels (RGB, normal, depth)
- Robust loss function does not need foreground masks or knowledge about distractors
  → highly flexible
- Fine-grained details can still be learned

Limitations:
- Slower convergence as a large part of the loss is not used for backpropagation
- Robust loss requires adding a hyperparameter which needs to be tuned based on the approximate quantity of distractors in the dataset
- Scale and shift variant loss included in baseline causes artifacts which make evaluation difficult
  → currently not trained on omnidata depth and normal maps [1] but on ground truth

Links

Project page: code, poster and videos

References

[1] A. Eftekhar, A. Sax, J. Malik, and A. Zamir. Omnidata: A scalable pipeline form aking multi-task mid-level vision datasets from 3d scans. In Proc. of the IEEE International Conf. On Computer Vision (ICCV), 2021.

[2] Sara Sabour, Suhani Vora, Daniel Duckworth, Ivan Krasin, David J. Fleet, and Andrea Tagliasacchi. RobustNeRF: IgnoringDistractors with Robust Losses, Feb. 2023.

[3] Zehao Yu, Anpei Chen, Bozidar Antic, Songyou Peng Peng, Apratim Bhattacharyya, Michael Niemeyer, Siyu Tang, Torsten Sattler, and Andreas Geiger. Sdfstudio: A unified framework for surface reconstruction, 2022