NeRFs are Mirror Detectors: Using Structural Similarity for Multi-View Mirror Scene Reconstruction with 3D Surface Primitives – *Supplementary Material* –

Leif Van Holland[1]   Michael Weinmann[2]   Jan U. Müller[1]   Patrick Stotko[1]   Reinhard Klein[1]
[1]University of Bonn   [2]Delft University of Technology

holland@cs.uni-bonn.de   m.weinmann@tudelft.nl   {muellerj,stotko,rk}@cs.uni-bonn.de

Hyperparameter	Value
Scoring	
SSIM window size	31
c_1	1×10^{-4}
c_2	9×10^{-4}
c	10
S	$0.5 \cdot \max_r s(r)$
Shape Fitting	
normal estimate k-neighbors	100
normal estimate search radius	0.5
clustering k-means #iterations	1×10^5
RANSAC distance threshold	5×10^{-3}
RANSAC #iterations	2×10^5
Joint Optimization	
τ_{init}	1×10^3
τ_{inc}	1×10^5
τ_{std}	2×10^5

Table 1. Hyperparameter choices.

Scene	IoU [%]	F_1 **score**
scene_1	98.8	0.994
scene_2	98.6	0.993
scene_3	98.6	0.993
scene_4	98.7	0.993
scene_5	98.6	0.993
scene_tram	0.0	0.000
scene_trumpet	98.2	0.991
scene_hallway	90.7	0.931
scene_library	96.7	0.983

Table 2. Mask overlap between the high quality annotations provided with the TraM-NeRF dataset and masks our automatic approach produces after training. Measured are intersection over union (IoU) and the F_1 score.

A. Hyperparameters

Tab. 1 shows the set of hyperparameters that were used throughout all synthetic and real world scenes in our evaluation, categorized into the different steps of the pipeline they are used in.

B. Additional Results

In Fig. 1, we show results on each of the scenes we selected from the TraM-NeRF dataset for evaluation.

Tab. 2 additionally reports the intersection over union (IoU) and F_1 score of the high-quality, manually annotated masks from the TraM-NeRF dataset compared to the masks produced by our automated approach after the final optimization. More specifically, the masks are rendered from all cameras in the test dataset and the metrics are evaluated on the resulting binary images and averaged over the views.

It can be seen that our approach gets close to the annotation data, but falls back to a standard NeRF reconstruction on "scene_tram", which corresponds to the failure case we have shown in the main paper.

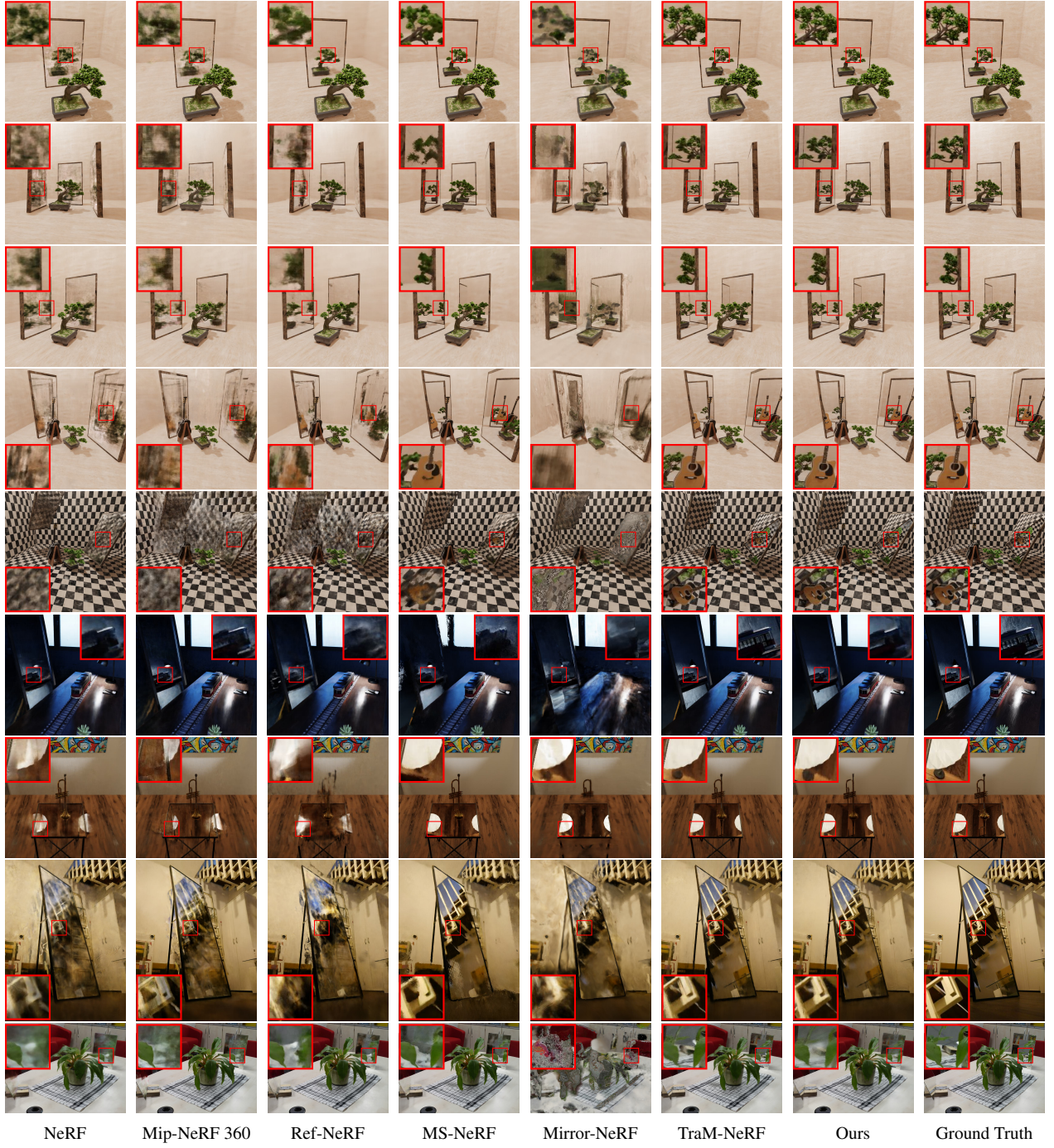|							
NeRF	Mip-NeRF 360	Ref-NeRF	MS-NeRF	Mirror-NeRF	TraM-NeRF	Ours	Ground Truth

Figure 1. Reconstruction results of our method compared to previous methods on the subset of the Tram-NeRF dataset we selected. Rows 1-7 show synthetic scenes, while rows 8 and 9 show real world scenes. The last column represents the ground truth image from the test set. The rows are in the same order as listed in Table 2.