SPARF:
Neural Radiance Fields from Sparse and Noisy Poses

CVPR 2023 - Highlight
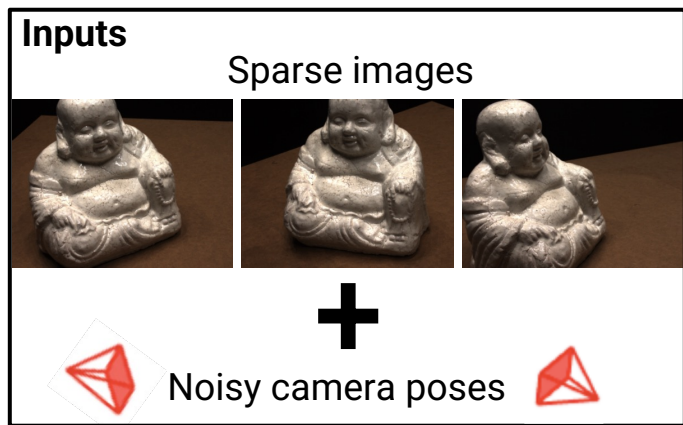TUE-PM-006

Prune Truong, Marie-Julie Rakotosaona,
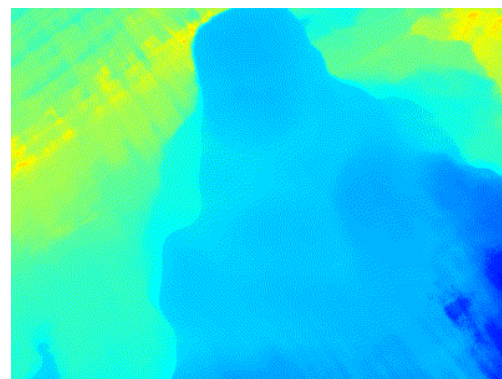Fabian Manhardt, Federico Tombari

Our approach SPARF: A joint pose-NeRF training strategy

Goal: Novel-view synthesis from **few images** with **noisy camera poses**

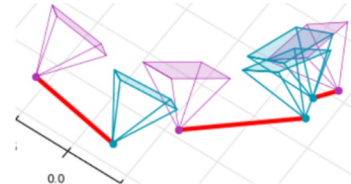- Exploits multi-view geometry principles

- 2 additional constraints

Google

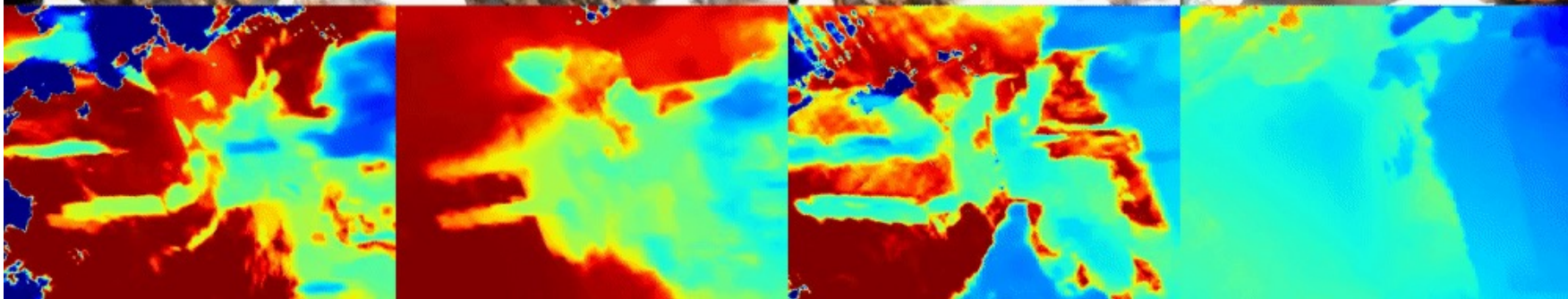**Inputs:**

Noisy camera poses

BARF RegBARF SCNeRF **SPARF (Ours)**

RGB

Depth

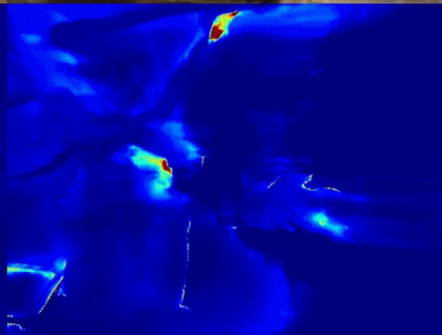**Inputs:**

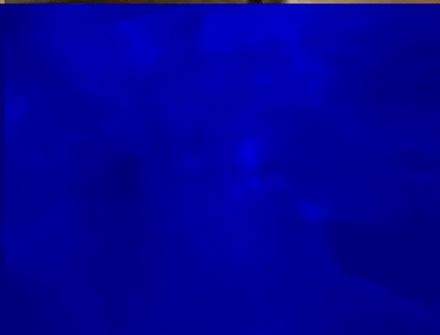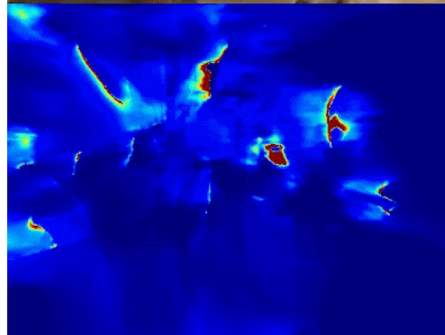Identity camera poses

BARF   RegBARF   SCNeRF   **SPARF (Ours)**

RGB

Depth

Inputs:

Noisy camera poses (obtained by COLMAP)

BARF   DS-NeRF   SCNeRF   **SPARF (Ours)**

RGB
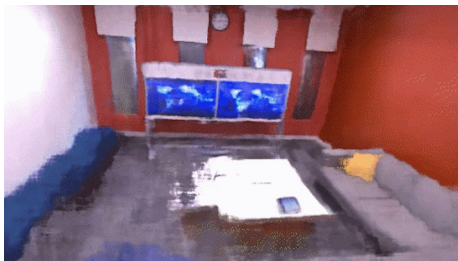
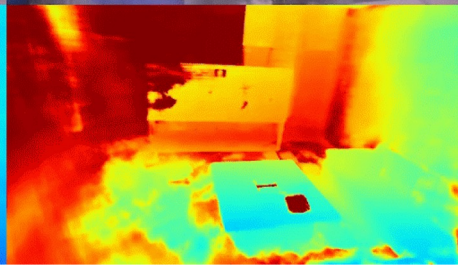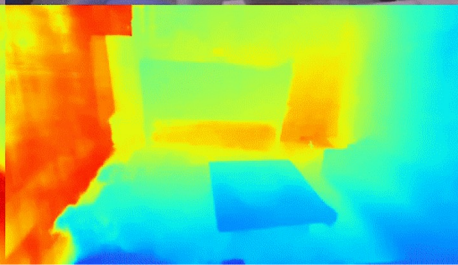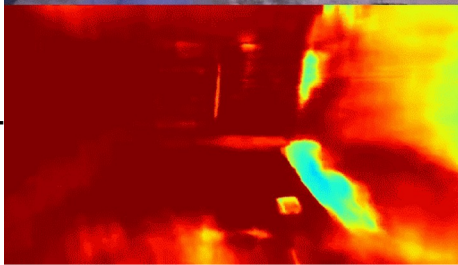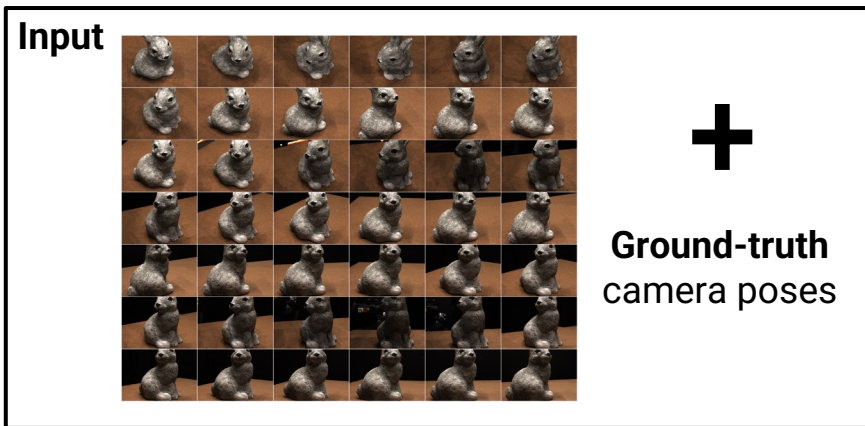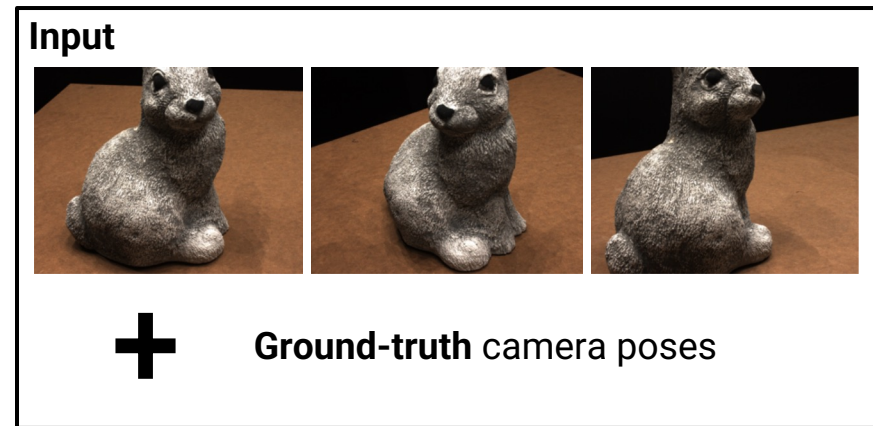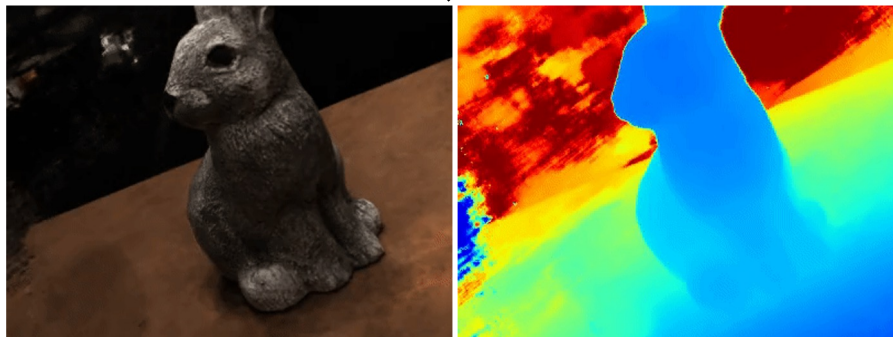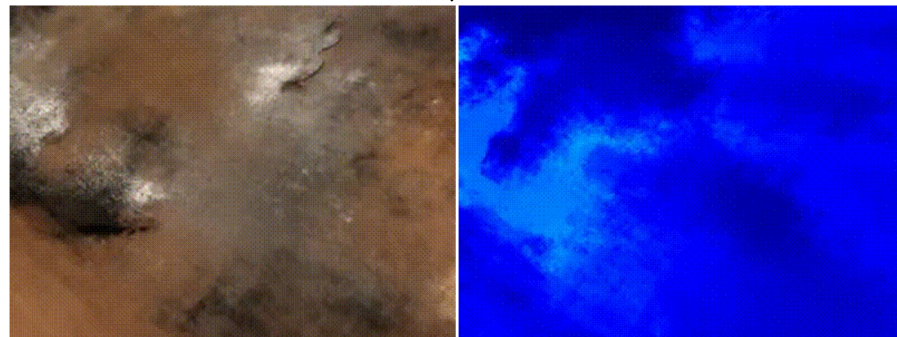Depth
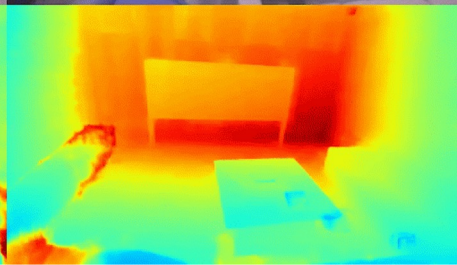
SPARF in more details
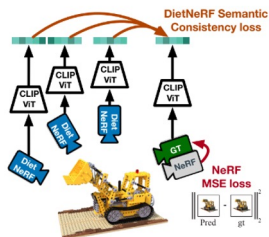
Google

NeRFs in the sparse-view setting

NeRFs in the sparse-view setting

Works that try to improve this

DietNeRF
(ICCV 2021)

RegNeRF
(CVPR 2022)

InfoNeRF
(CVPR 2022)

DS-NeRF (CVPR 2022)

They all assume ground-truth poses as input

This is an unrealistic assumption!
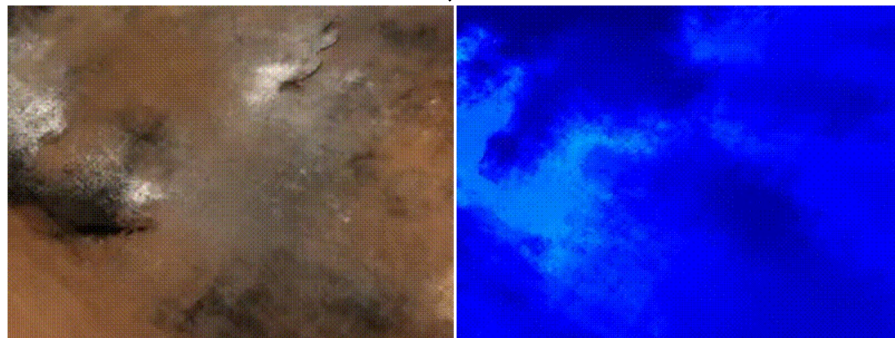
Input

+ **Ground-truth** camera poses

NeRF

NeRFs in the noisy-input-poses setting

How do we get the input poses?

We should assume access to only noisy camera poses as input!
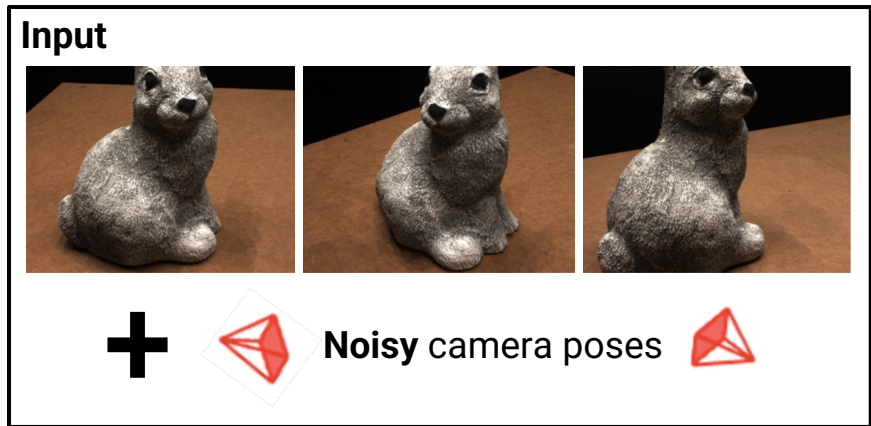
Works that train a NeRF from noisy input poses
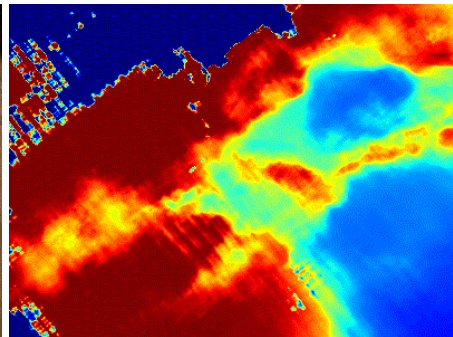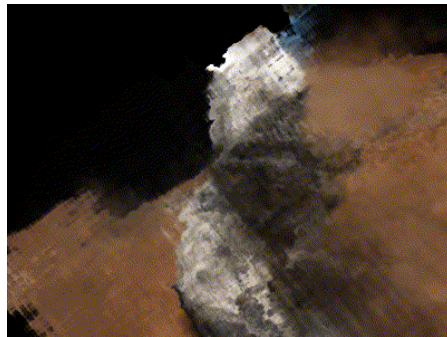
BARF (ICCV 2021) **+ Follow-ups**

SCNeRF (ICCV 2021)

They all assume ~~ense~~ input views

They fail with sparse input views

Input

+ **Noisy** camera poses

BARF

Our approach SPARF

Inputs

Sparse images

+

Noisy camera poses
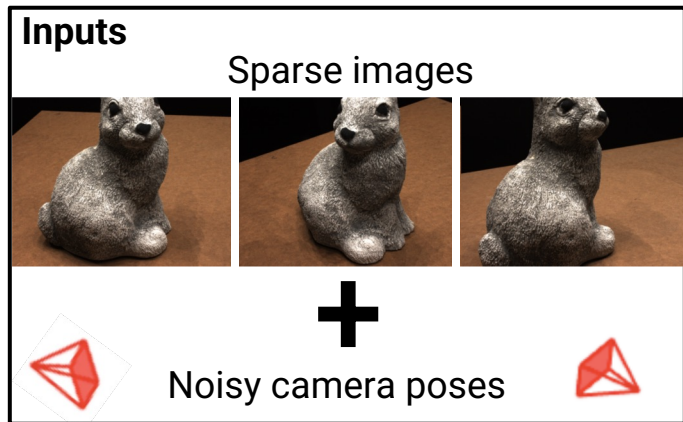
Our approach SPARF

- Joint pose-NeRF training strategy

- 2 additional constraints

- Exploits multi-view geometry principles

Google

SPARF: Multi-view correspondence loss

Sparse input images
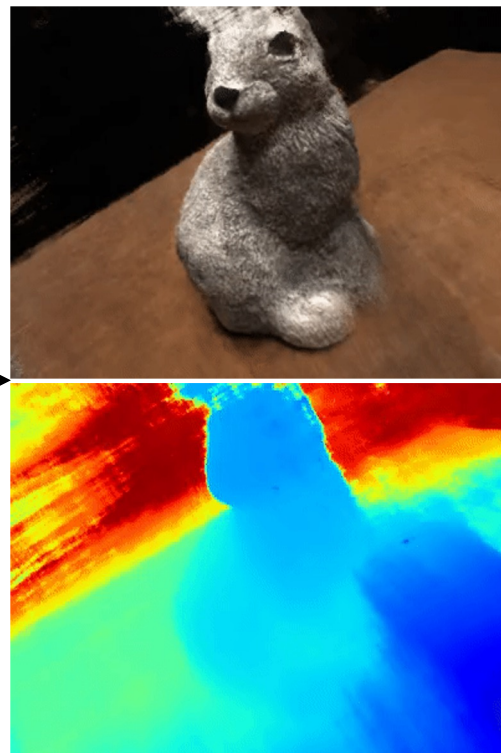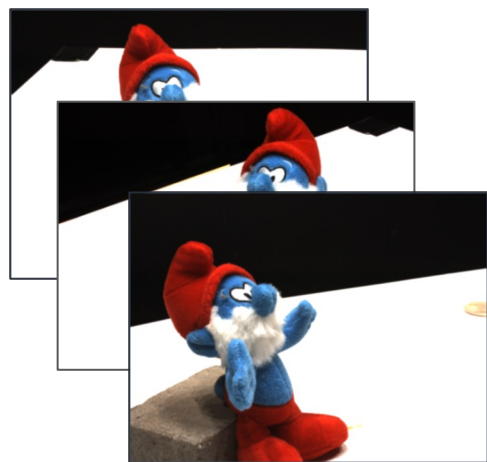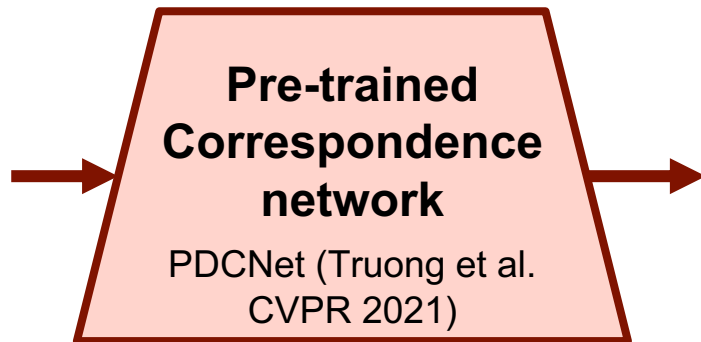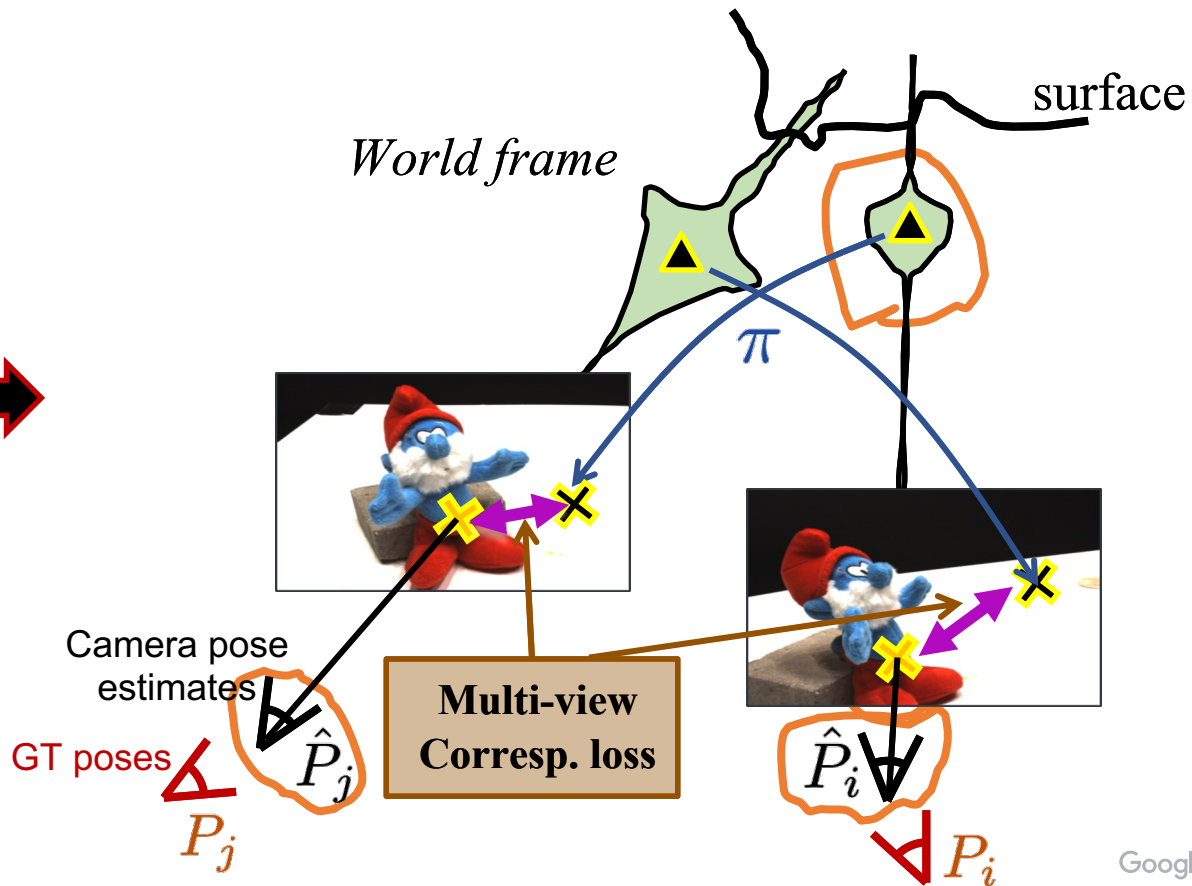
Pre-trained Correspondence network

PDCNet (Truong et al. CVPR 2021)

Google

SPARF: Multi-view correspondence loss

surface

World frame

π

Pre-trained
Correspondence
network
PDCNet (Truong et al.
CVPR 2021)

Sparse input images

Camera pose
estimates

GT poses

$\hat{P}_j$

P_j

**Multi-view
Corresp. loss**

$\hat{P}_i$

P_i

Google
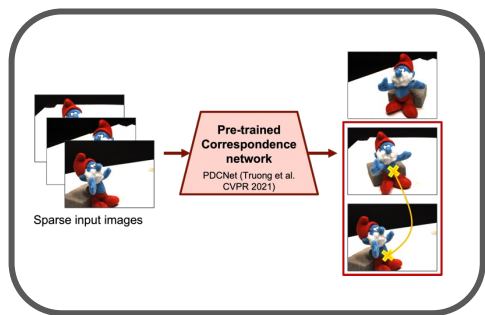
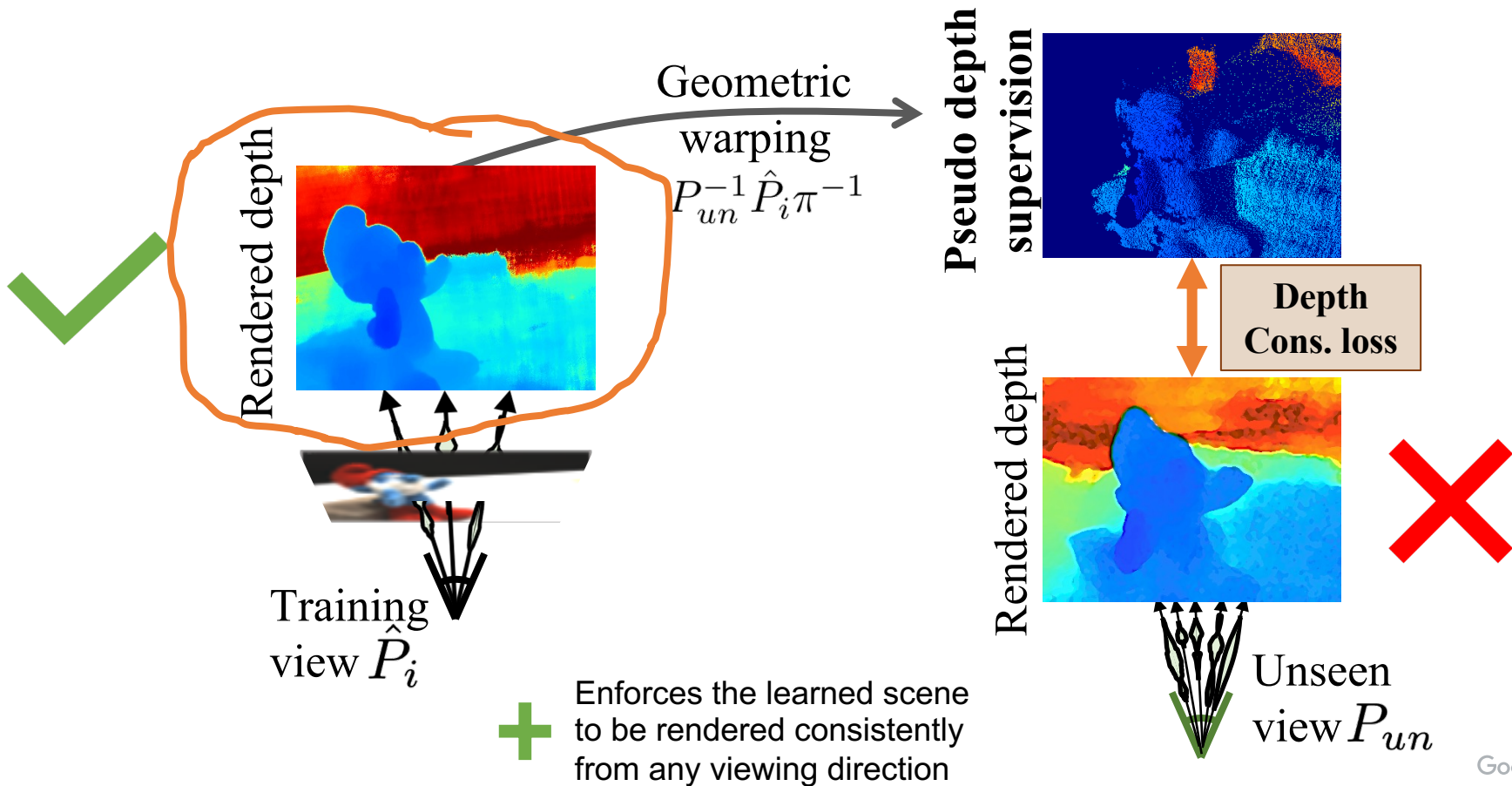# SPARF: Multi-view correspondence loss
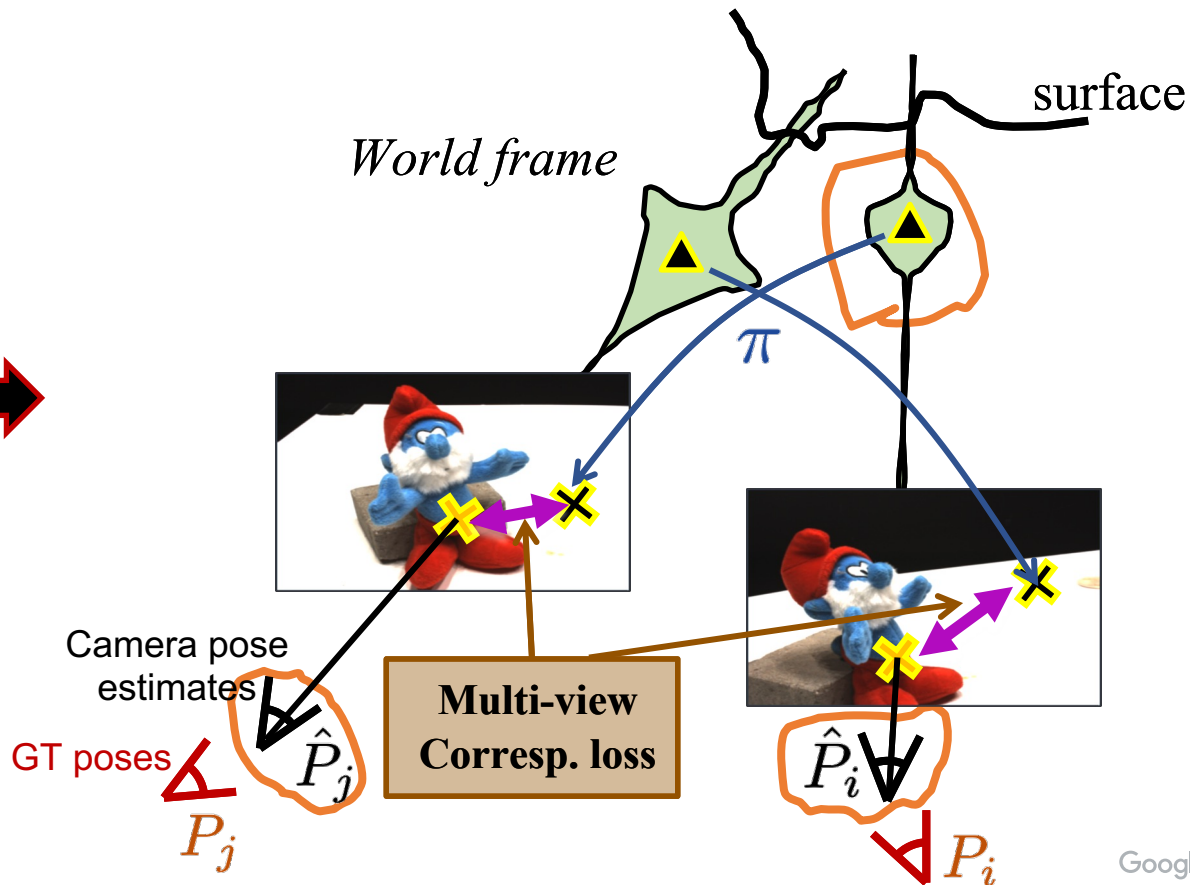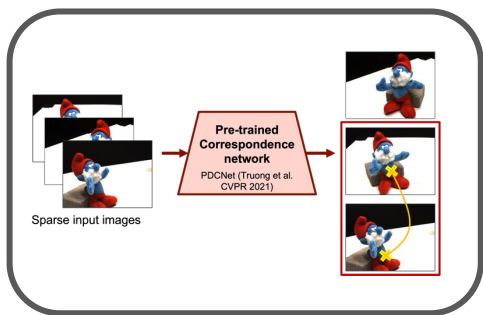
Enforces convergence to global and geometrically accurate solution

Enforces rendered depth to be close to the surface

surface

World frame

π

Pre-trained Correspondence network

PDCNet (Truong et al. CVPR 2021)

Sparse input images

Camera pose estimates

GT poses

Multi-view Corresp. loss

$\hat{P}_j$

P_j

$\hat{P}_i$

P_i

SPARF: Depth consistency loss

Geometric warping $P_{un}^{-1}\hat{P}_i\pi^{-1}$

Pseudo depth supervision

Rendered depth

Depth Cons. loss

Rendered depth

Training view $\hat{P}_i$

Unseen view P_{un}

Enforces the learned scene to be rendered consistently from any viewing direction

Google

Results: Experimental set-up

- Evaluation on multiple datasets: object-centered, forward-facing scenes, indoor scenes.
- Sparse-view scenario: only 3 available.
- Different 'noisy poses' initializations.
- **SPARF = State-of-the-art**

Poses obtained
by COLMAP

Synthetically perturbed
ground-truth poses with
Gaussian noise

Identity poses

DTU dataset
(DTU Informatics 2010, Aanaes et al.)

LLFF dataset
(BMVA 2021, Shafiei et al.)

Replica dataset
(2019, Straub et al.)
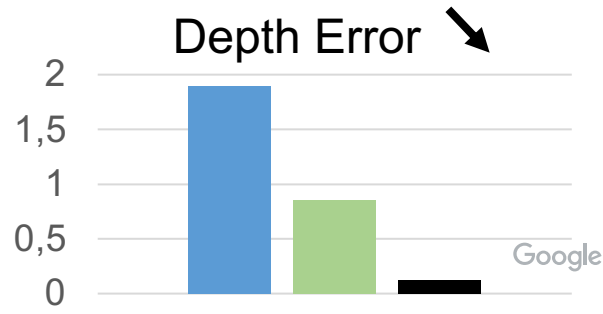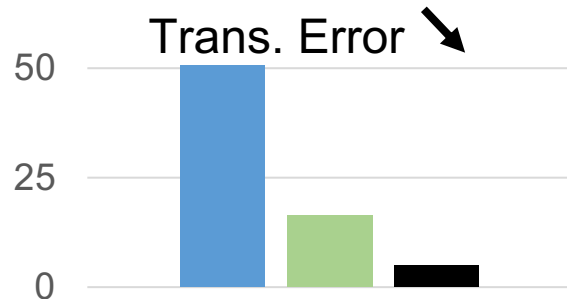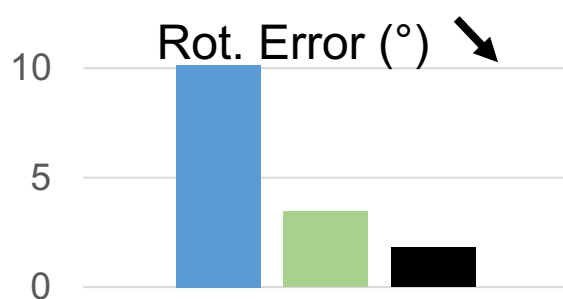
Results on DTU - 3 input views and noisy camera poses

Perturbed/optimized camera poses

Ground-truth camera poses

━━━ Translational error
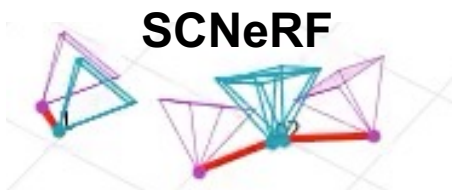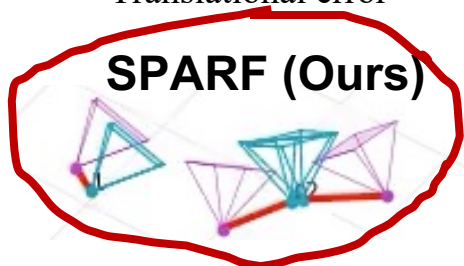
SPARF (Ours)

SCNeRF

BARF

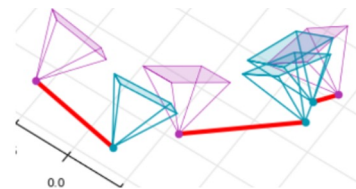■ BARF ■ SCNeRF ■ **SPARF (Ours)**

Rot. Error (°) ↘

Trans. Error ↘

PSNR ↗
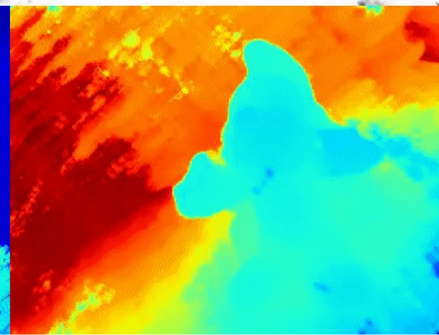
Depth Error ↘

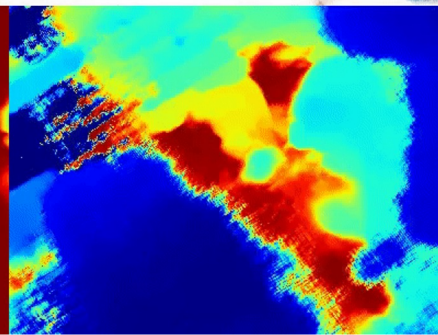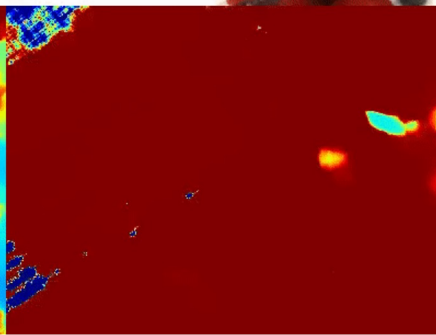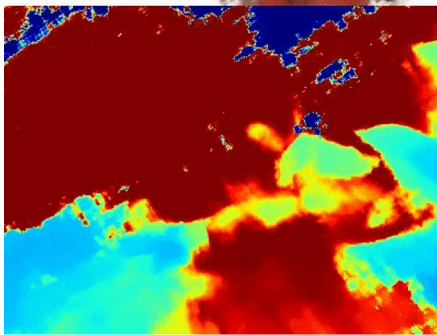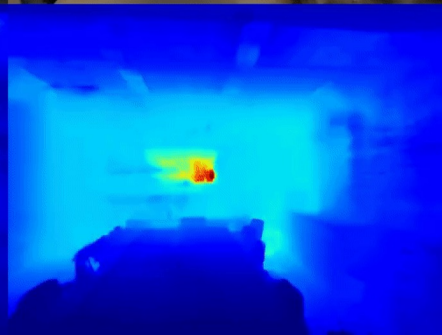Google

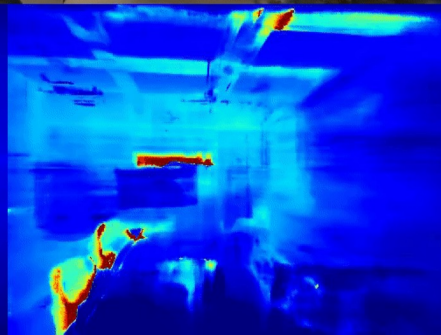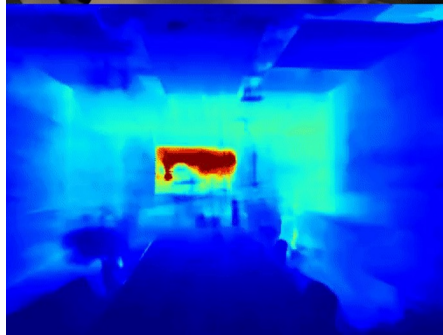Inputs:

Noisy camera poses

BARF | RegBARF | SCNeRF | **SPARF (Ours)**

RGB

Depth

Inputs:

Identity camera poses

BARF | RegBARF | SCNeRF | **SPARF (Ours)**

RGB

Depth

Inputs:

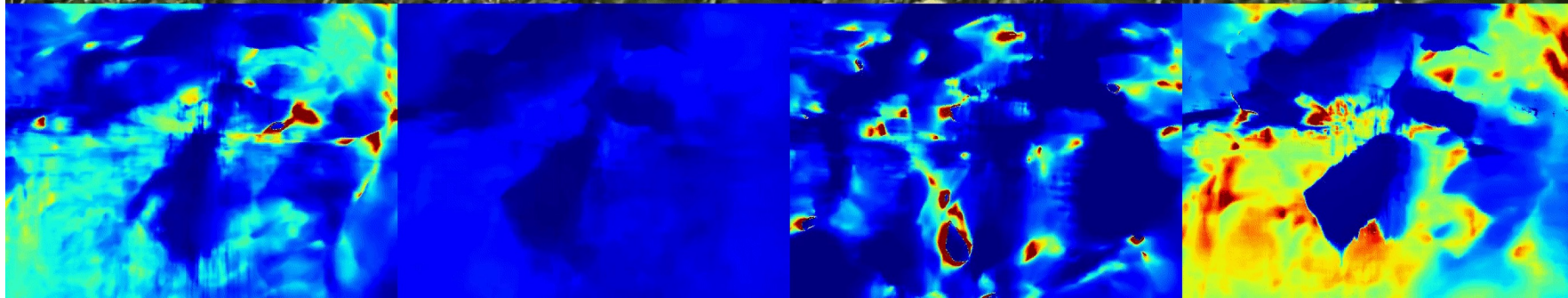**+** Identity camera poses

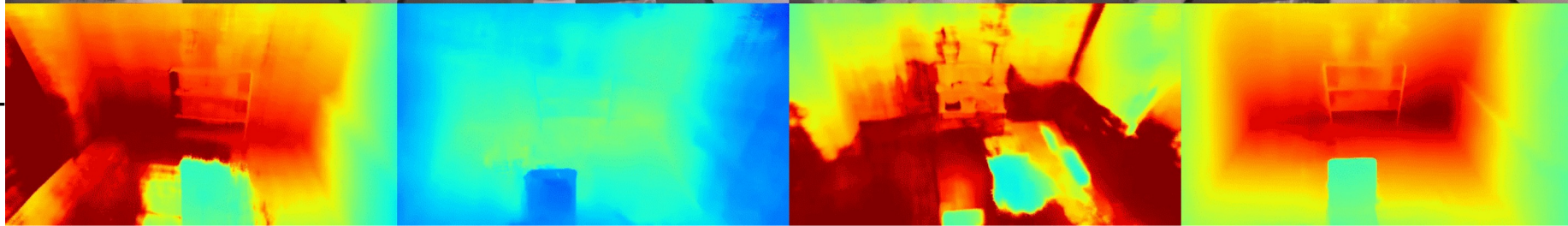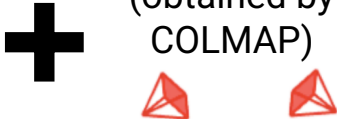| BARF | RegBARF | SCNeRF | **SPARF** (Ours) |

RGB

Depth

Inputs:

Noisy camera poses
(obtained by
COLMAP)

BARF   DS-NeRF   SCNeRF   **SPARF (Ours)**

RGB
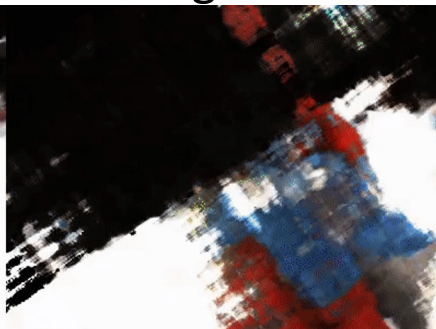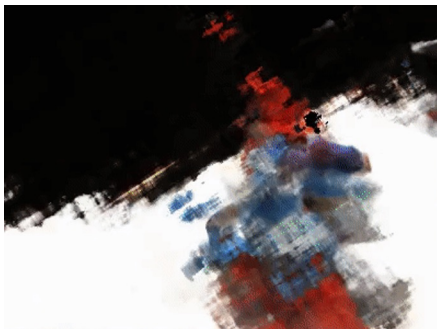
Depth

Inputs:

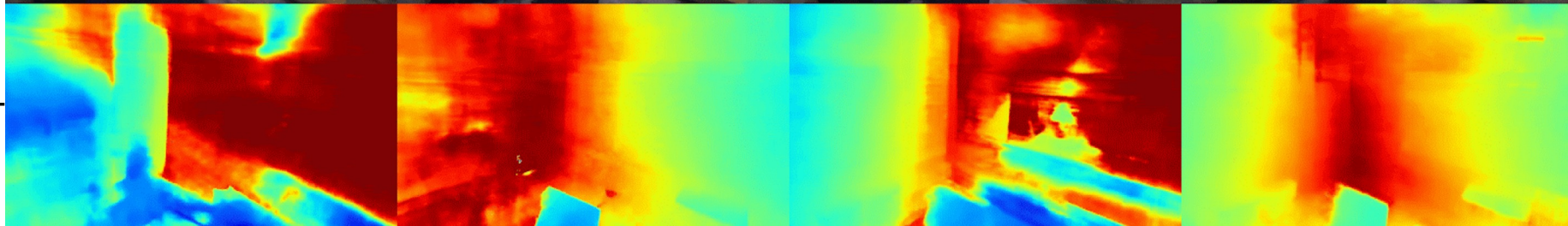Noisy camera poses (obtained by COLMAP)

BARF	DS-NeRF	SCNeRF	**SPARF (Ours)**

RGB

Depth

Thank you!

Code

Website