SceNeRFlow: Time-Consistent Reconstruction of General Dynamic Scenes

Edith Tretschk  Vladislav Golyanik  Michael Zollhöfer  Aljaž Božič  Christoph Lassner  Christian Theobalt

Goal
General dynamic NeRF with time consistency/correspondences even for large motion

Problem Setting and Context
Input
General non-rigid scene captured with multi-view RGB videos (with known camera parameters and background images)

Output
Time-consistent reconstruction of geometry, appearance, and deformations

Prior Work: Either category-specific (e.g. humans) or only handles small motion (e.g. only consistent over short time windows) → Ours is first method to get correspondences for large general motion!

High-Level Method Idea:

- Static NeRF: No deformation, only geometry and appearance
- Deformable NeRF: Disentangle deformation from geometry and appearance
- Volumetric Video: Entangle deformation with geometry and appearance

Method

- Build static canonical model (i.e. geometry & appearance) at t=1
- Online optimization of deformations at t>1, regularized by as-rigid-as-possible deformation smoothness loss

- Decompose into coarse and fine deformations

Unexpected Challenge: Doing all this yields very strong artifacts!

Why? Backwards deformation models have bad initialization for large motion

Solution: Initialize surrounding space via deformation smoothness loss

Results

Time Consistency

Ablation: Letting the Canonical Model Vary Over Time

Application: Time-Consistent Editing

References:

Acknowledgments: All data capture and evaluation was done at MPI. Research conducted by Vladislav Golyanik and Christian Theobalt at MPI was supported in part by the ERC Consolidator Grant 652847 (775306). This work was also supported by a Meta Reality Labs research grant.

Code is available!
github.com/facialbaseresearch/SceNeRFlow

Video results:
vcai.mpi-inf.mpg.de/projects/scenerflow