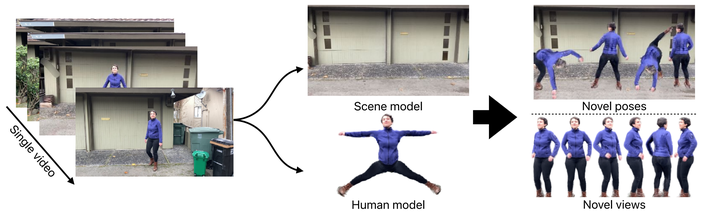
NeuMan: Neural Human Radiance Field from a Single Video

Wei Jiang[1,2], Kwang Moo Yi[1,2], Golnoosh Samei[1], Oncel Tuzel[1], and Anurag Ranjan[1]
[1]Apple, Inc. [2]The University of British Columbia

Abstract

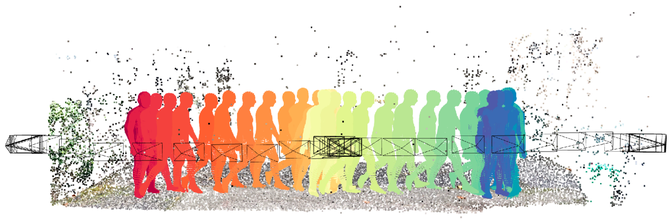
NeuMan reconstructs both the **background scene** and an **animatable human** from a **single** video using neural radiance fields.



Preprocessing

COLMAP ➡️ Camera poses, point cloud
ROMP ➡️ SMPL estimates
Detectron2 ➡️ Mask
mmpose ➡️ 2D joints
BoostingMonocularDepth ➡️ Monocular depth

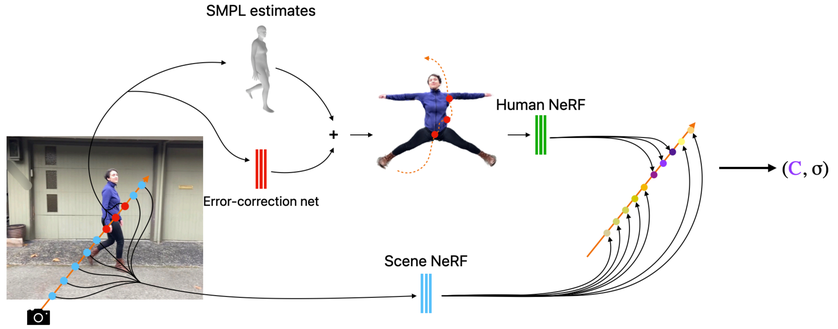
We use off-the-shelf methods to estimate the 2D or 3D geometry of the scene or the human:



Training

We first train the scene NeRF with the background pixels which are obtained from the predicted masks.

We then train the human NeRF in the canonical space. The ray is warped to the canonical space based on the SMPL mesh skinning.



We jointly optimize the poses during the training.

Results

Canonical human | Novel poses | Joint rendering

Multi-persons rendering

Code and Data

https://github.com/apple/ml-neuman

Try with your own videos!