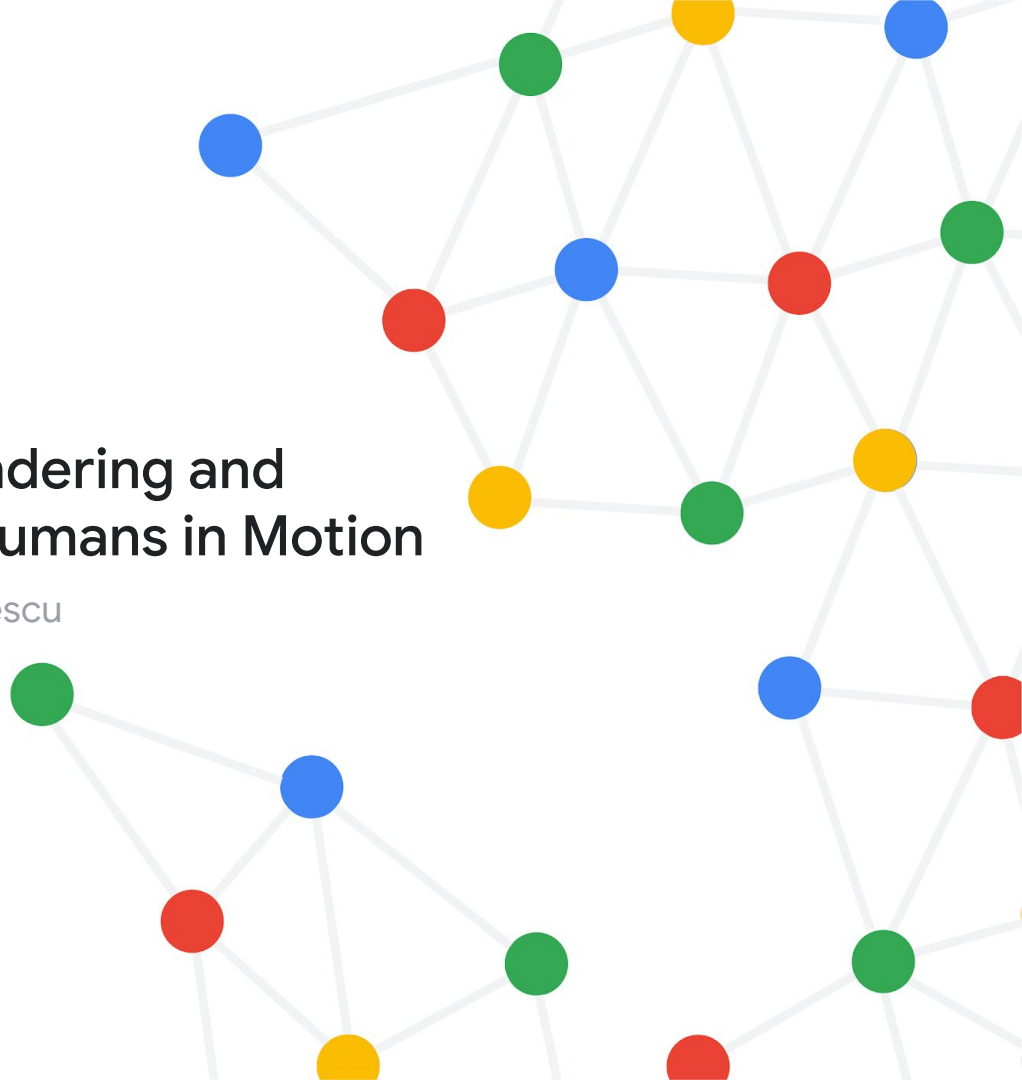
H-NeRF:
Neural Radiance Fields for Rendering and Temporal Reconstruction of Humans in Motion

Hongyi Xu, Thiemo Alldieck, Cristian Sminchisescu



Google Research

H-NeRF

Videos with novel views synthesized using H-NeRF

Google Research

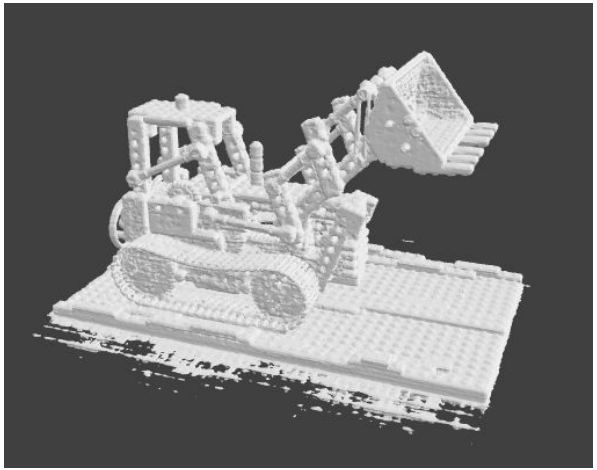
Neural Radiance Fields

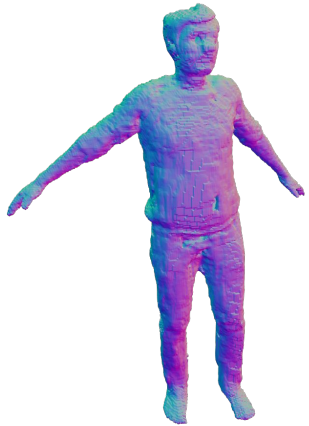
Static Scenes
[Mildenhall et al. 2020]

Human Motion
[Peng et al. 2021]

Google Research

Limitations of Neural Radiance Fields

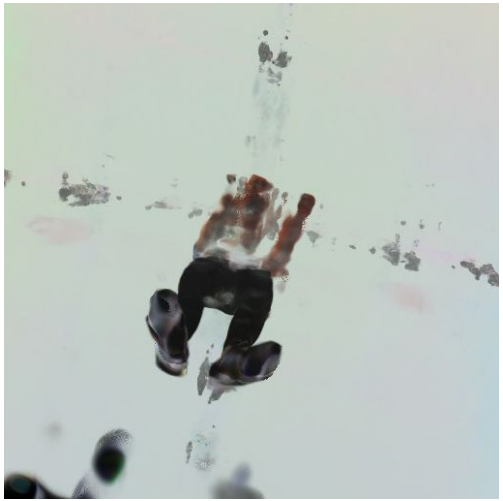




Neural rendering does not generalize well for viewpoints far from the training set

Google Research

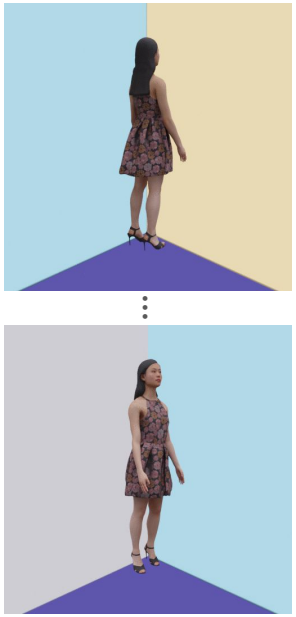
Limitations of Neural Radiance Fields



Neural rendering does not generalize well for viewpoints far from the training set, and for deformable and articulated structures

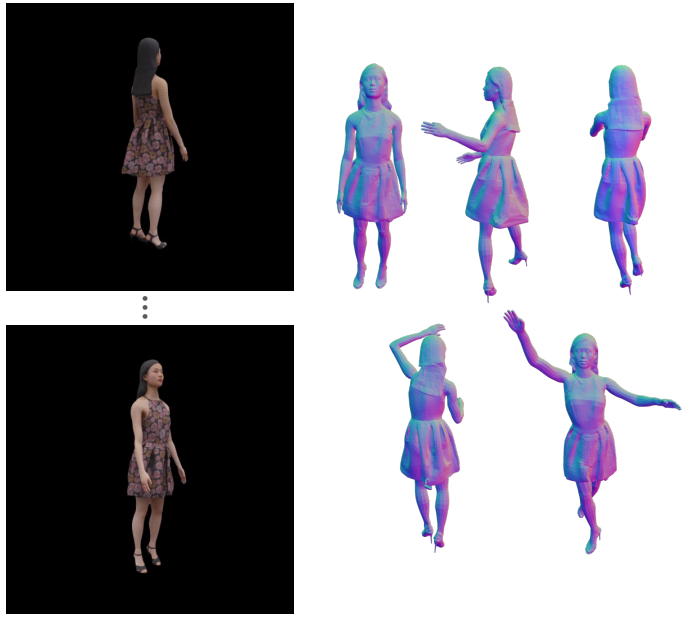
Google Research

H-NeRF



Input: monocular or
sparse multi-view videos

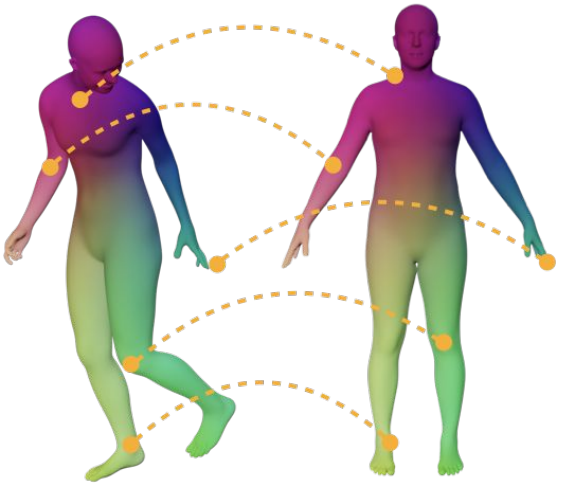
Neural radiance field



Output: rendering and temporal
reconstruction of a human in motion

Implicit 3D Body Model (imGHUM) for shape and appearance regularization

1. Rich geometric prior for occupancy

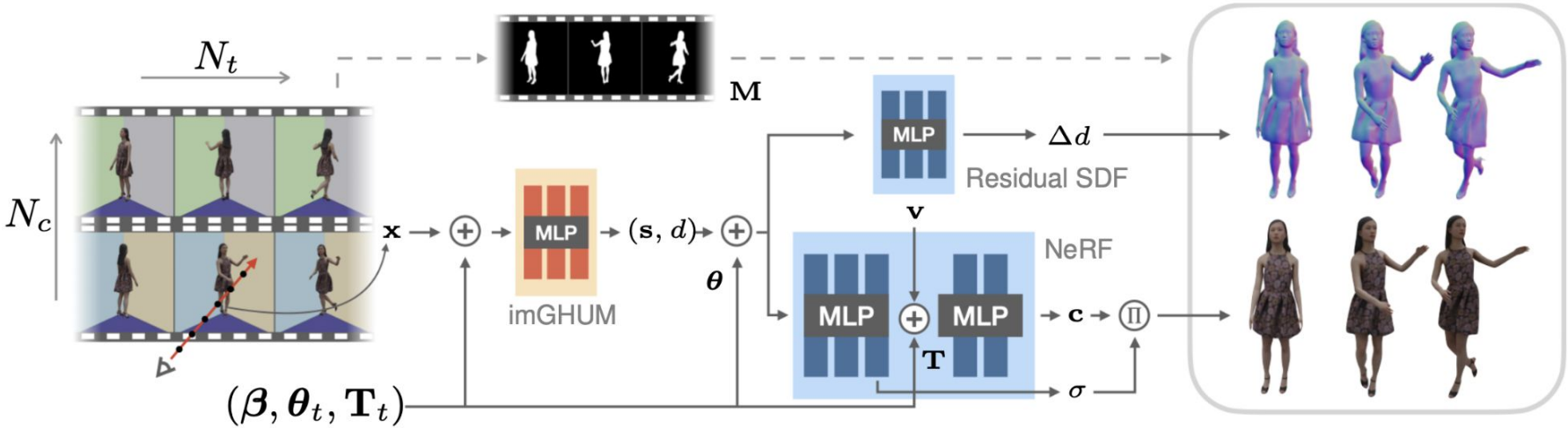


2. Semantic mapping function to integrate appearance over time into a common coordinate system

The implicit body model is an important building block of H-NeRF

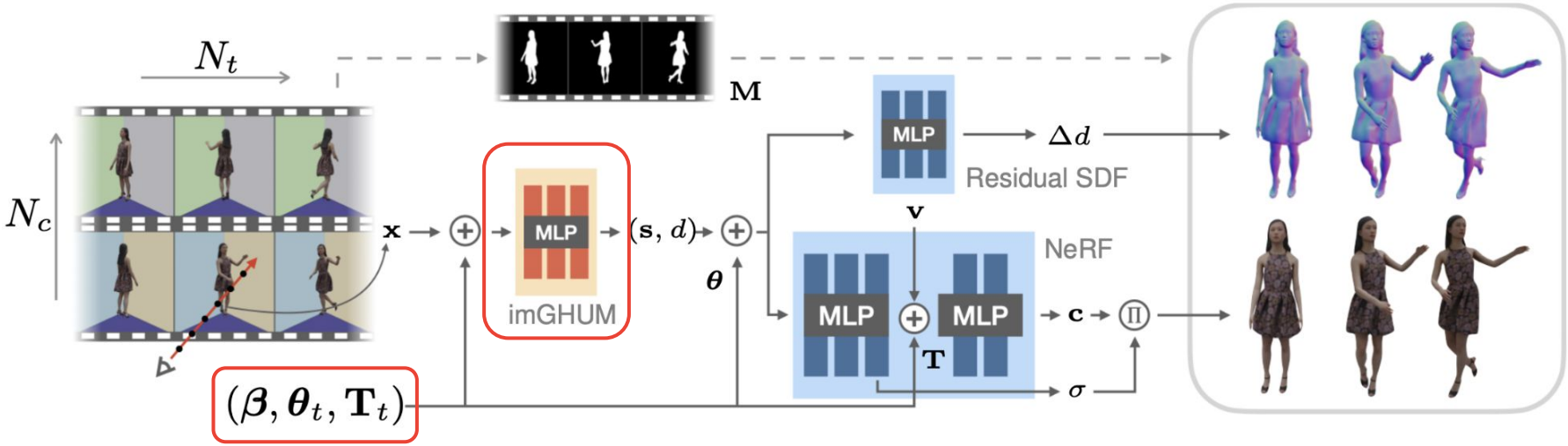
Google Research

Overview



Google Research

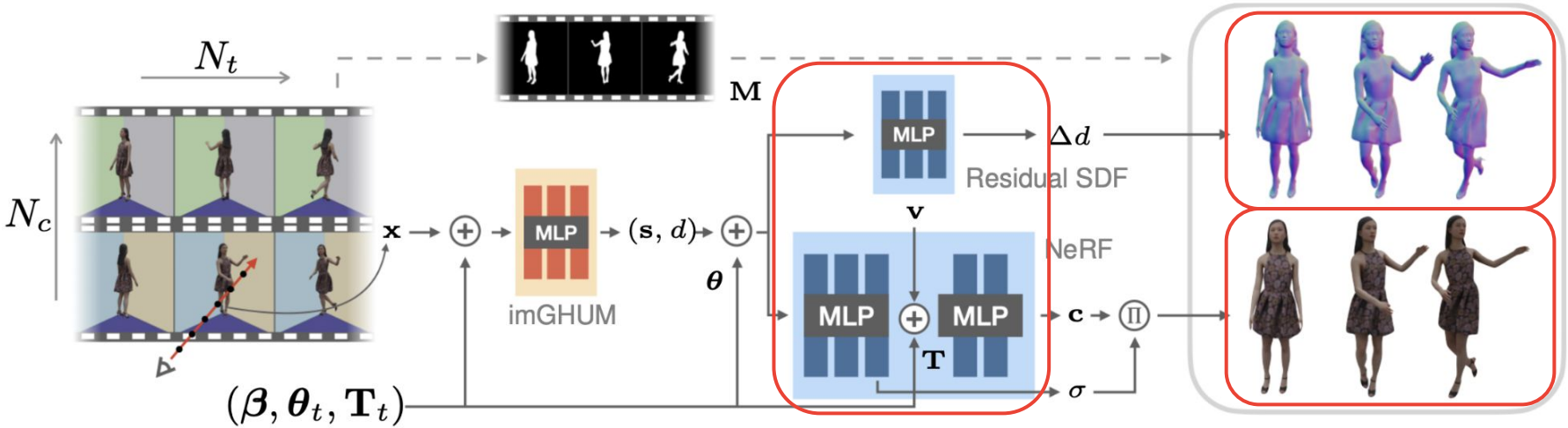
Overview



The implicit SDF-based body model imGHUM serves as a geometric prior and common reference frame

Google Research

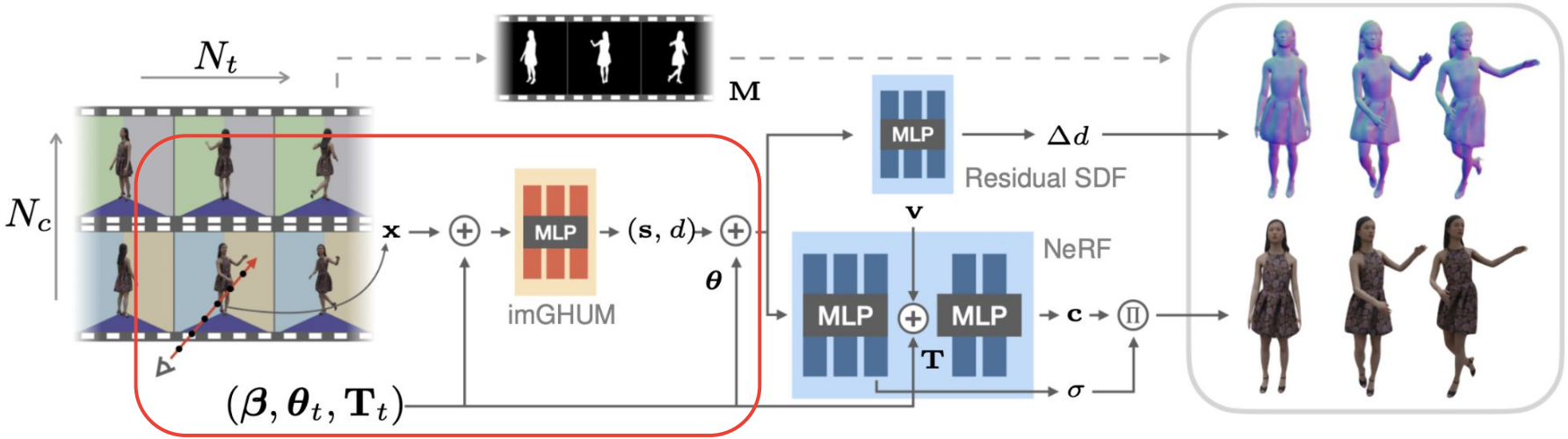
Overview



We co-train a neural radiance field and a signed distance function to enable high fidelity in **both** view synthesis **and** 3D reconstruction

Google Research

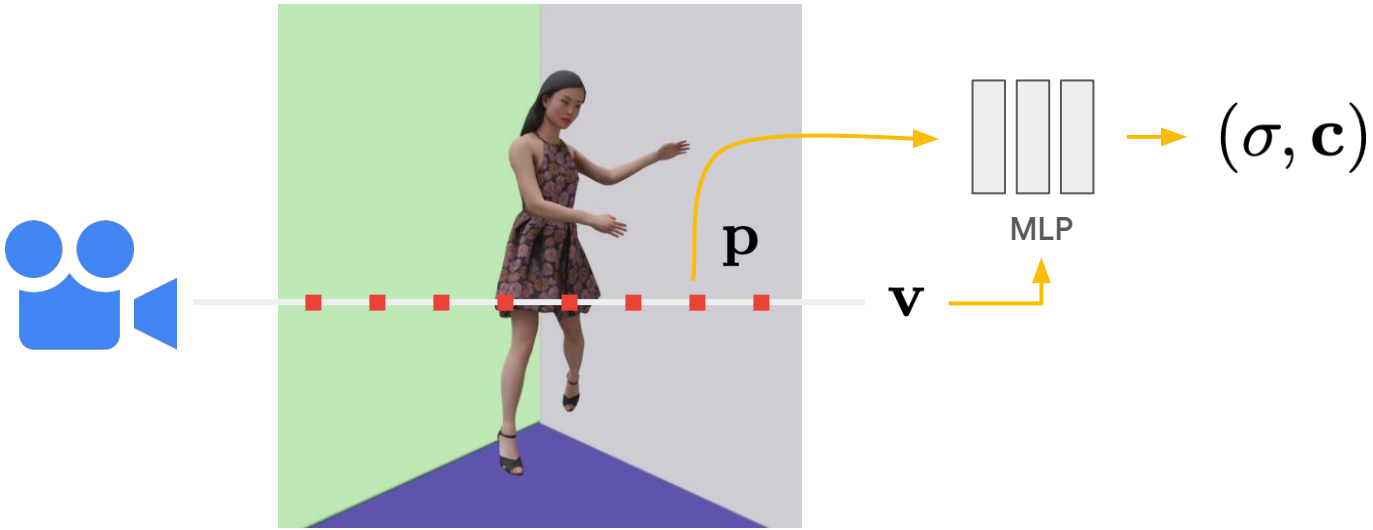
Overview



At test time, the subject can be animated with different motions, or even synthesized with extrapolated shape, by varying imGHUM's generative parameters

Google Research

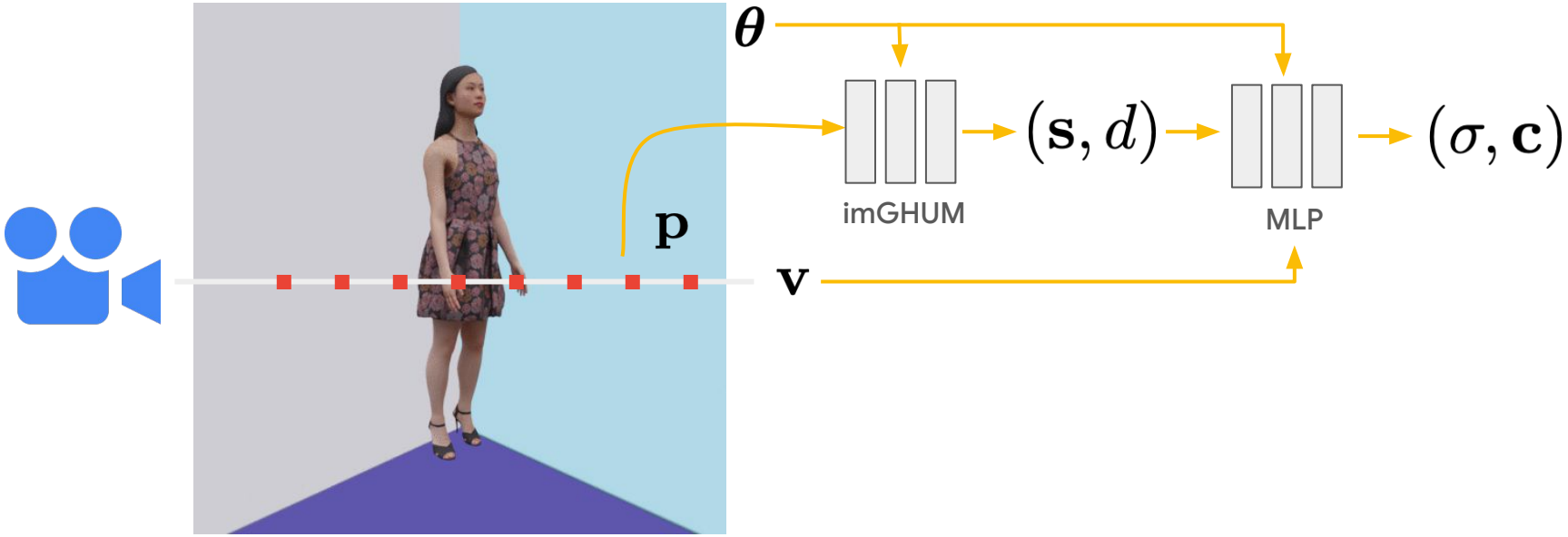
NeRF Formulation



The rendering component of our method builds upon the original NeRF formulation, and is trained using only image observations

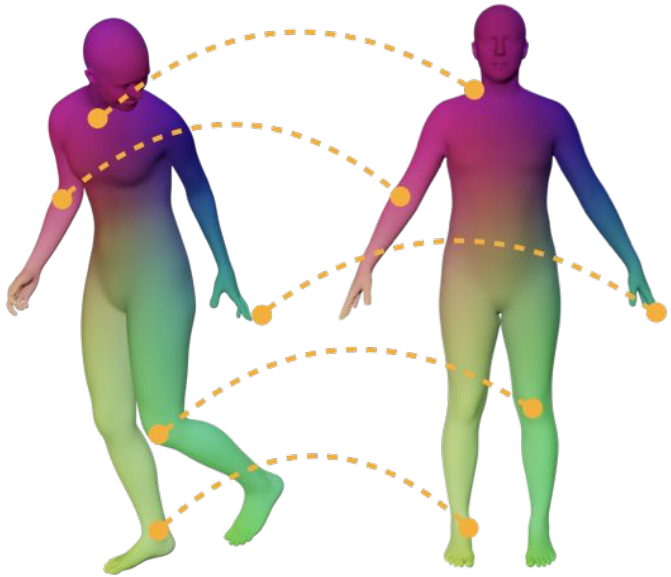
Google Research

Dynamic NeRF Formulation



Google Research

imGHUM as Spatial Warping Function

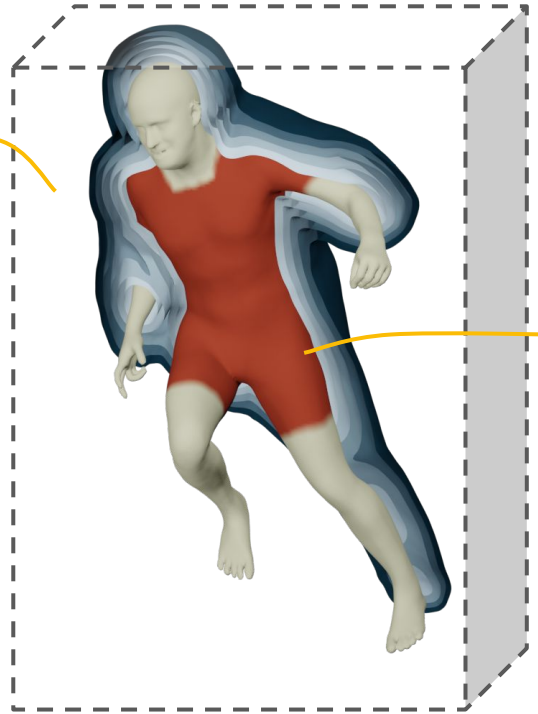


imGHUM returns distance and semantics that map every point in space into a common reference frame

Google Research

imGHUM as Geometric Prior

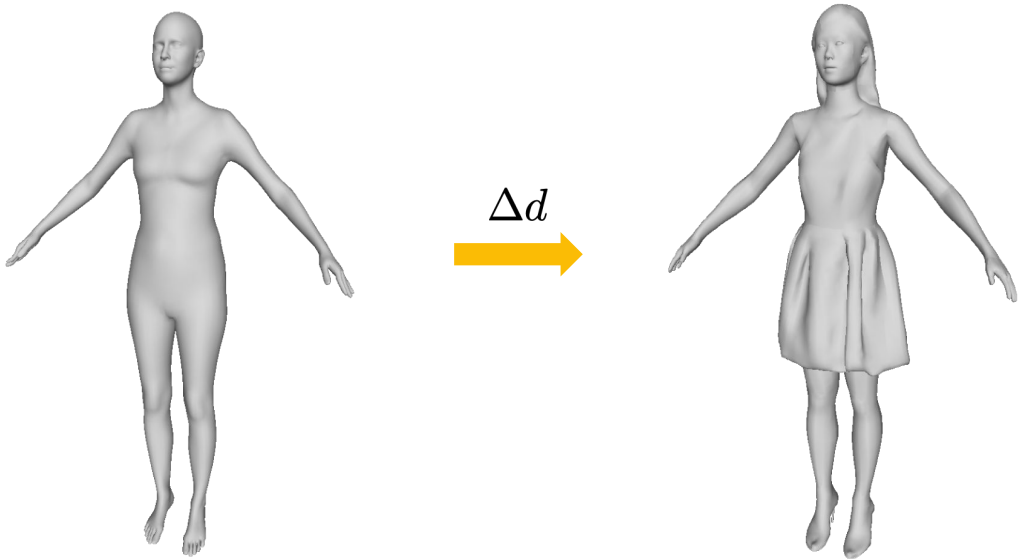
We expect the occupancy of the reconstructed subject to be inside imGHUM's relaxed bounding box.



We expect high NeRF density inside the imGHUM body estimate.

Google Research

Residual SDF



We co-learn a residual SDF to the (fixed) imGHUM model to be able to represent hair and clothing

Google Research

Results

Google Research

Novel View Synthesis

Ground truth H-NeRF

Ground truth H-NeRF

Training views

Test views (45° from training)

Google Research

Geometric Reconstruction

Each pair:
Ground truth (left) Ours (right)

Google Research

Comparison with SoTA

Ground truth

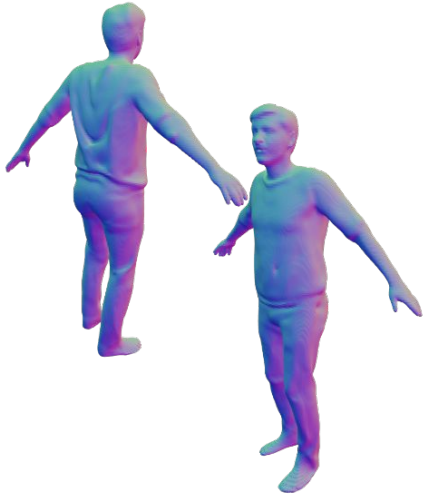
Ours

NeuralBody
[Peng et al. 2021]

Google Research

Comparison with SoTA

Ground-Truth



Ours

NeuralBody
[Peng et al. 2021]

Google Research

Rendering new Motions

Training sequence

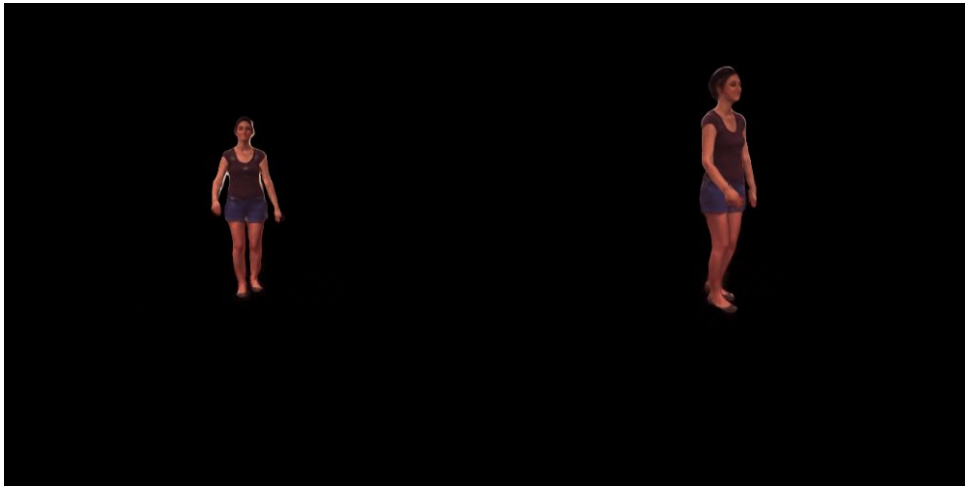
Training sequence

Synthesized **unseen motions** rendered using **testing** camera **views**

Google Research

Rendering new Motions based on Sparse Real-World Training Images

PeopleSnapshot (trained on 1 camera)



Human 3.6M (trained on 4 cameras)

Google Research

Statistical Body Shape Editing

Two subjects rendered with statistically modified body shapes,
both lower and higher BMI w.r.t. photographed (center)

Google Research

Generalization

Novel view synthesis and geometric reconstruction
with a novel shape and a novel pose sequence

Google Research

Thank You

Hongyi Xu
hongyixu@google.com

Thiemo Alldieck
alldieck@google.com

Cristian Sminchisescu
sminchisescu@google.com

Google Research