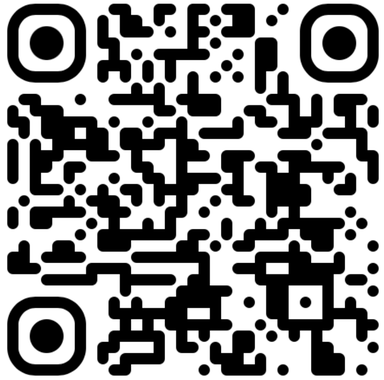
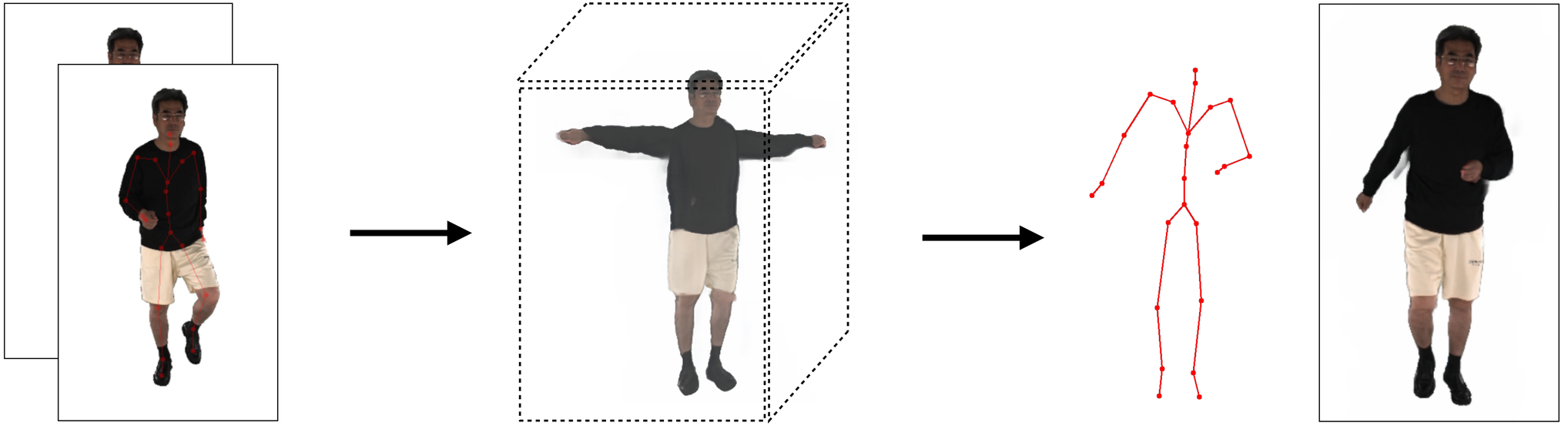


HumMorph: Generalized Dynamic Human Neural Fields from Few Views

Jakub Zadrozny, Hakan Bilen
University of Edinburgh

CV4Metaverse at CVPR 2025
Nashville | 11th June 2025

Dynamic Free-Viewpoint Human Rendering



Input: a few monocular views with pose parameters

Estimate: subject NeRF model in canonical pose

Input: target pose parameters

Output: render in target pose

Plenty of applications related to the Metaverse & beyond (movie production, immersive 3D communication, etc.).

Why *HumMorph*?

Subject-specific approaches

− Require test-time optimization

− Needs extensive observations (typically ca. 30 frames)

Other generalized approaches

− Assume accurate body shape and pose parameters (impractical)

***HumMorph* (ours, generalized)**

\+ Uses only feed-forward passes during inference

\+ Requires less observed views (1-4)

\+ Learns a prior, inpaints unobserved details

\+ Significantly more robust to errors in the noisy pose parameters

Estimated Body Shape and Pose

Accurate body shape and pose parameters are usually estimated from multi-view camera setups.

They should be directly estimated from the input views instead.

Accurate parameters.

Estimated parameters

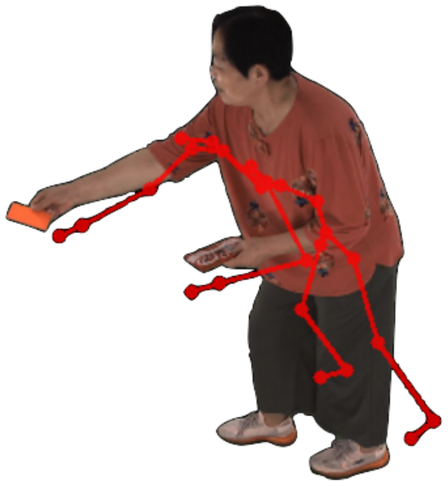
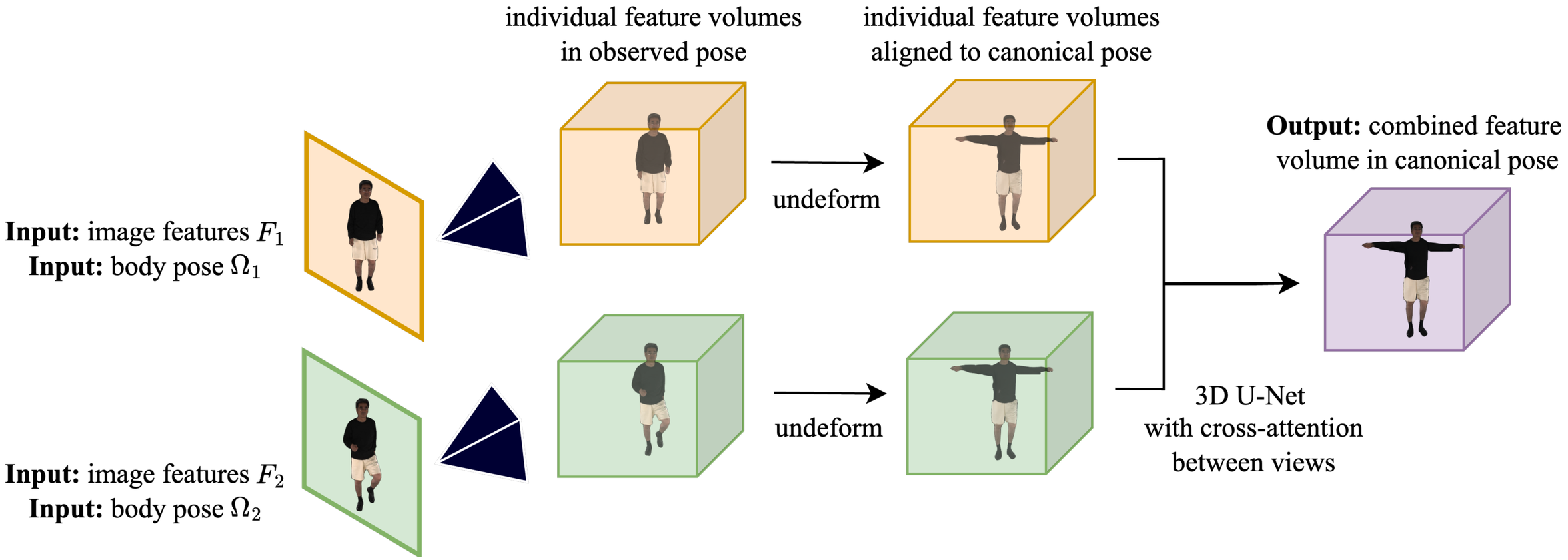


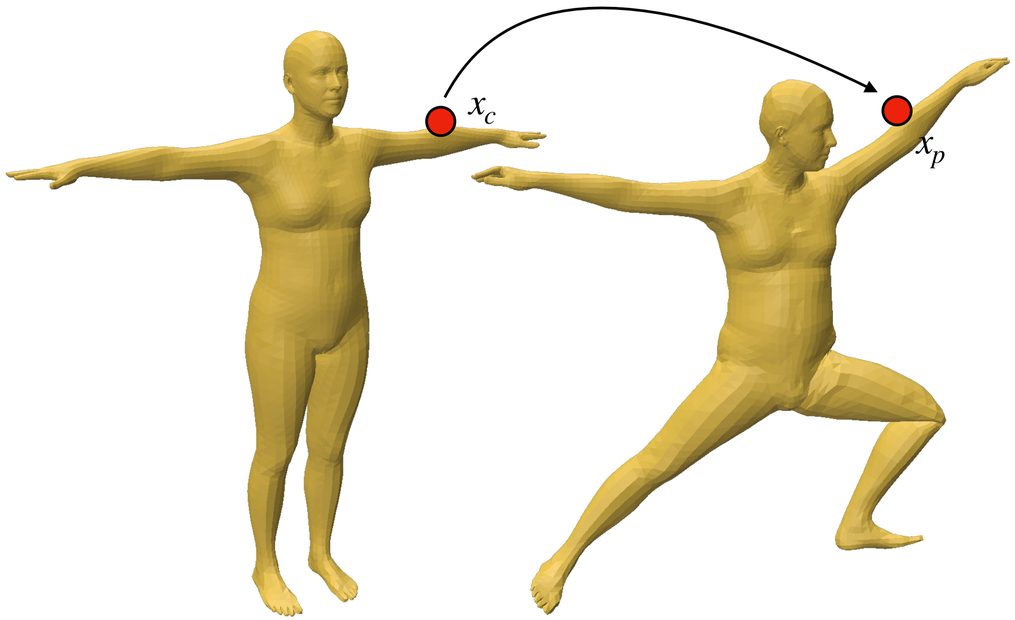
Fig.: Frames with skeleton annotated in red using *accurate* (left) and *estimated* (right) body shape and pose parameters. Parameters estimated using *HybrIK* (Li et al., CVPR ‘21).

The *VoluMorph* module

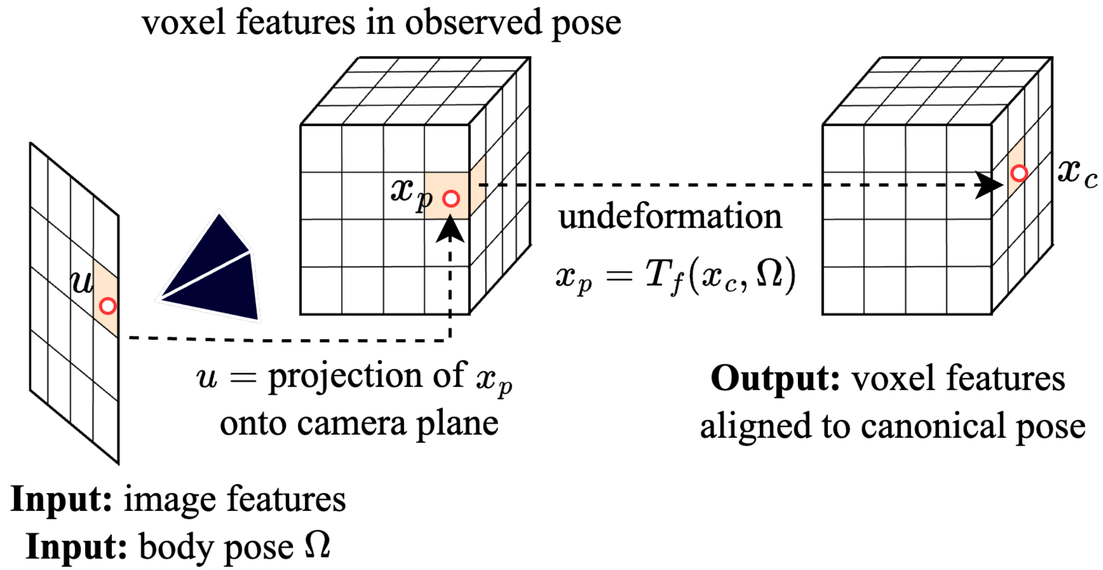


3D U-Net architecture from Szymanowicz et al., *Viewset Diffusion,* ICCV 2023.

Unprojection + Undeformation



canonical pose observed pose

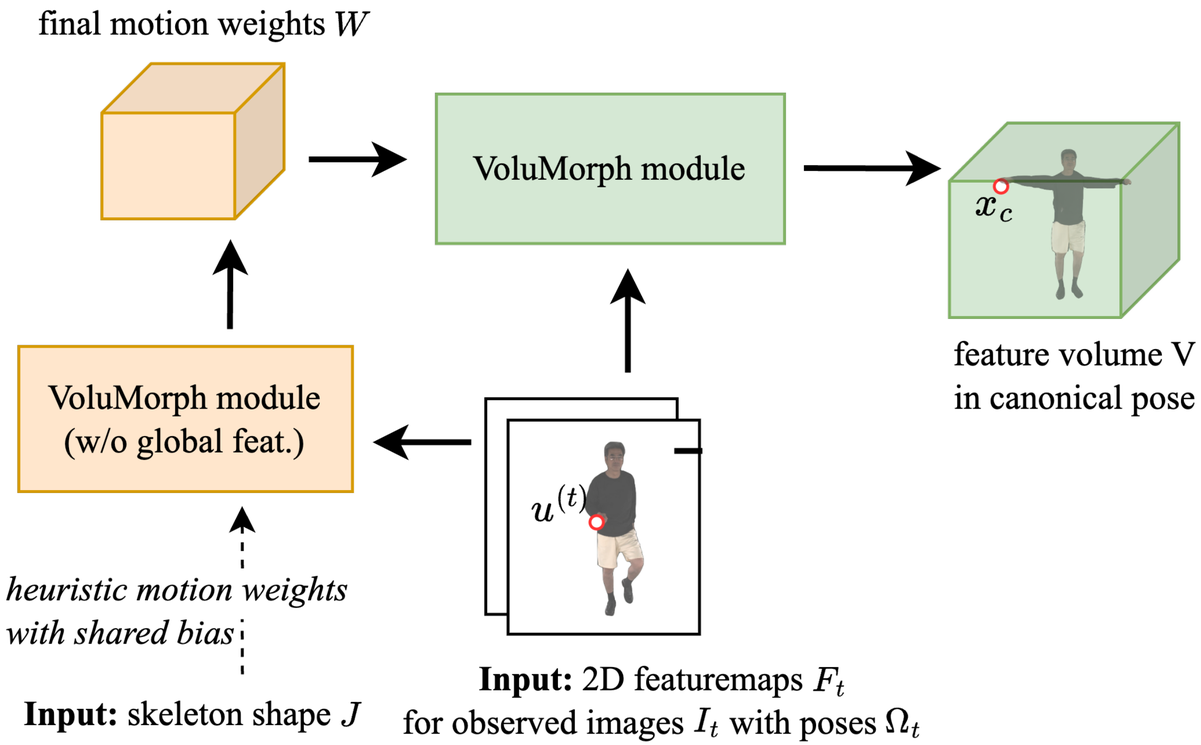


We use **linear blend skinning** for body deformations.

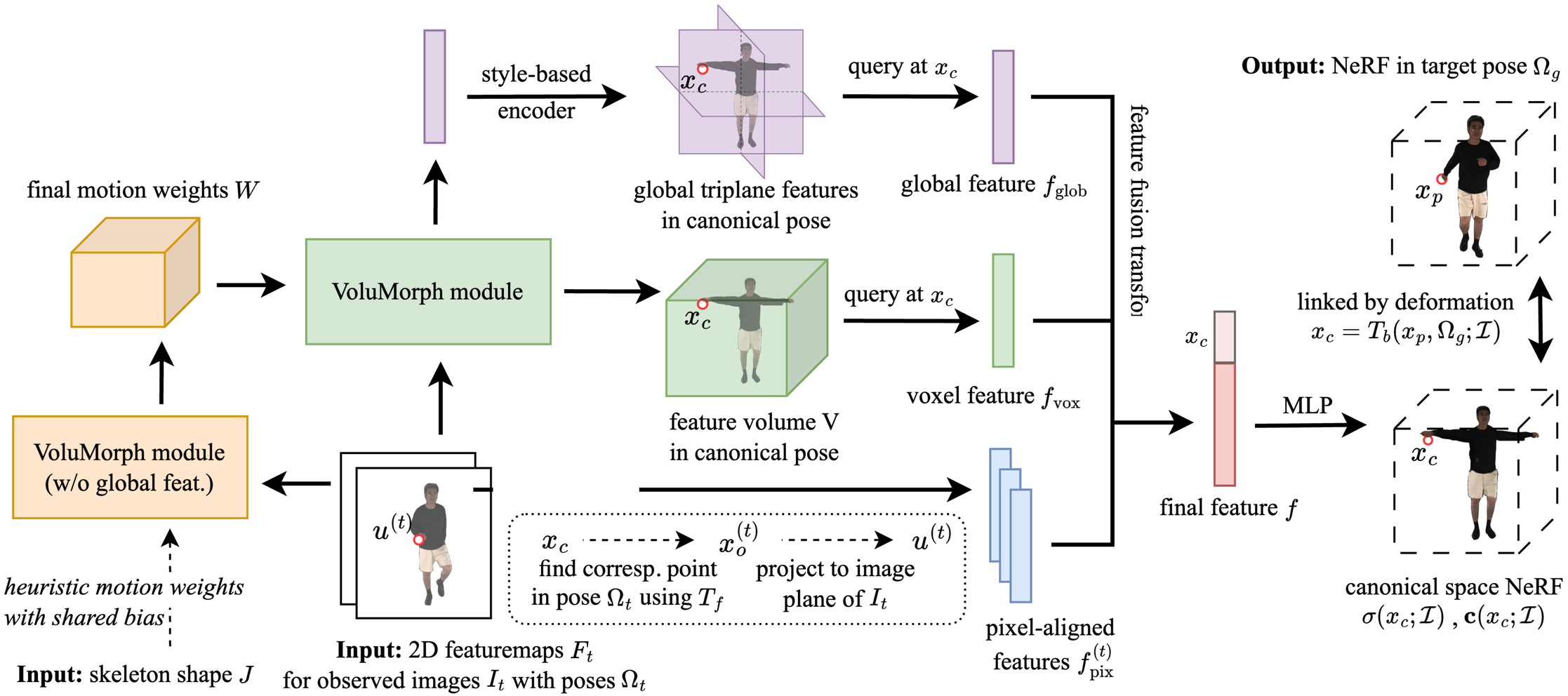
They identify 3D points (x_c and x_p) that correspond to the same body point in canonical and observed poses.

You can think of the *undeformation* as dragging voxels around according to the skinning deformation.

Zoom-Out to the Full Pipeline



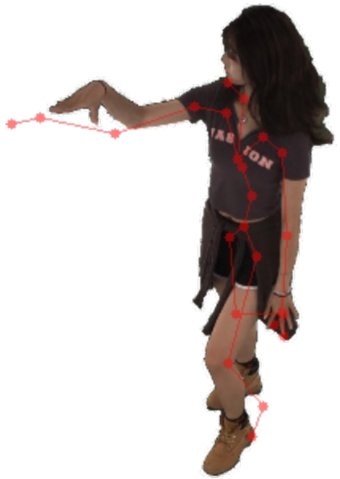
Zoom-Out to the Full Pipeline



Results with Estimated Poses

Observed 1

Observed 2

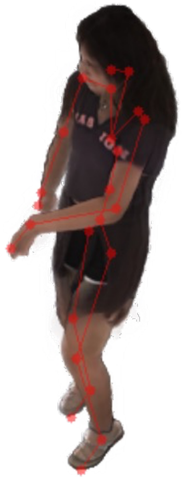


Observed 3

Observed 4

Ours (1)

Ours (1-2)



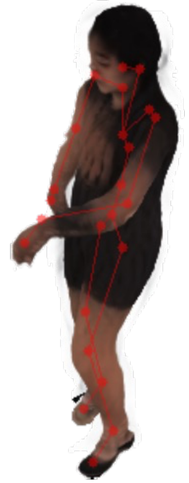
Ours (1-4)

Ground truth

SHERF (1)

SHERF (2)

GHuNeRF+ (1-4)



Numbers in parentheses indicate the range of observed views.

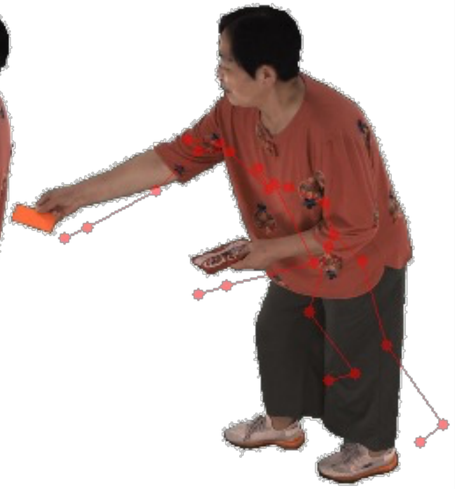
Li et al., *GHuNeRF,* 3DV 2024

Hu et al., *SHERF,* ICCV 2023

Results with Estimated Poses

Observed 1

Observed 2



Ours (1)

Ours (1-2)

Ours (1-4)

Ground truth

Observed 3

Observed 4

SHERF (1)

SHERF (2)

GHuNeRF+ (1-4)

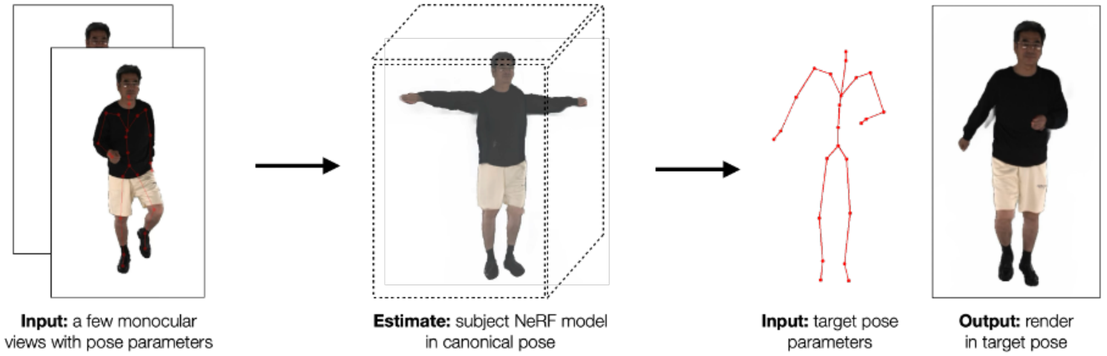
Numbers in parentheses indicate the range of observed views.

Li et al., *GHuNeRF,* 3DV 2024

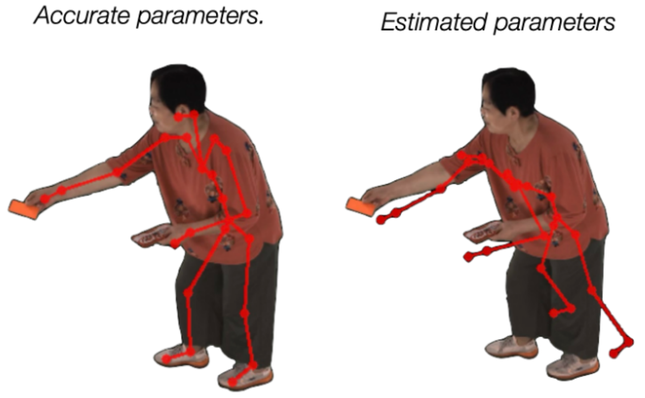
Hu et al., *SHERF,* ICCV 2023

Recap & thank you! Questions?

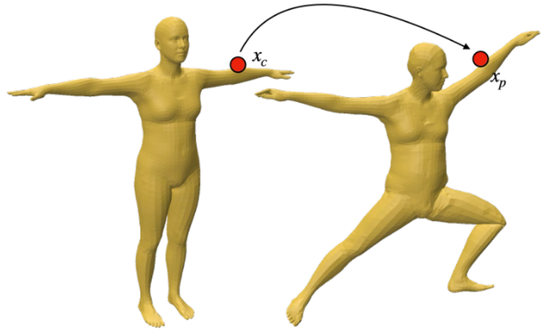
Project page:



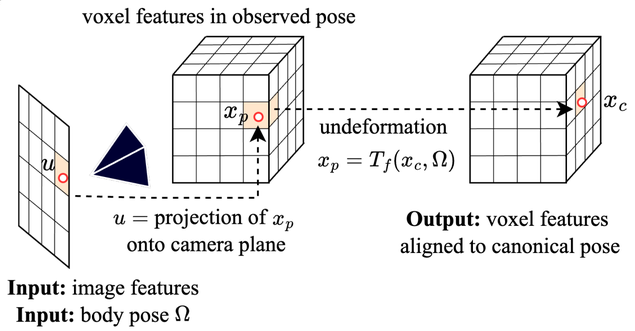
Dynamic free-viewpoint human rendering



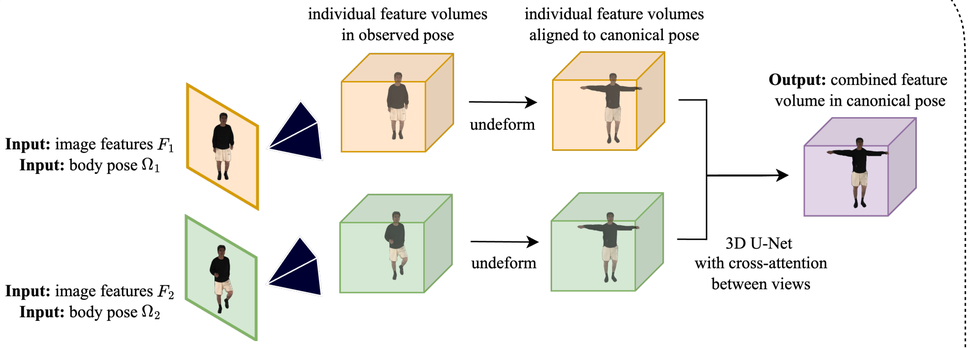
Realistic estimated poses



Skinning with learned weights



Unprojection + undeformation



The *VoluMorph* module

Limitations: modelling loose clothing and object interactions is not addressed properly, we use ground-truth camera poses.