

# **ActorsNeRF: Animatable Few-shot Human Rendering with Generalizable NeRFs**

Jiteng Mu<sup>1</sup>, Shen Sang<sup>2</sup>, Nuno Vasconcelos<sup>1</sup>, Xiaolong Wang<sup>1</sup>

<sup>1</sup>UCSD, <sup>2</sup>ByteDance

# Highlight: Few-shot Generalization

Category-level Prior



Few-shot Generalization to Novel Actors with Novel Poses



With only a few images from a monocular video, ActorsNeRF synthesizes **novel views** of a **novel person** with **novel poses**.

# **AIST++ Dataset**

# Animation



With only a few images (e.g., 30 shots) from a monocular video, ActorsNeRF synthesizes **novel views** of a **novel person** with **novel poses**.

# 100-shot Comparison

Ours



HumanNeRF



ActorsNeRF synthesizes novel views of the person with novel poses.

# 30-shot Comparison

Ours



HumanNeRF



ActorsNeRF synthesizes novel views of the person with novel poses.

# 10-shot Comparison

Ours



HumanNeRF

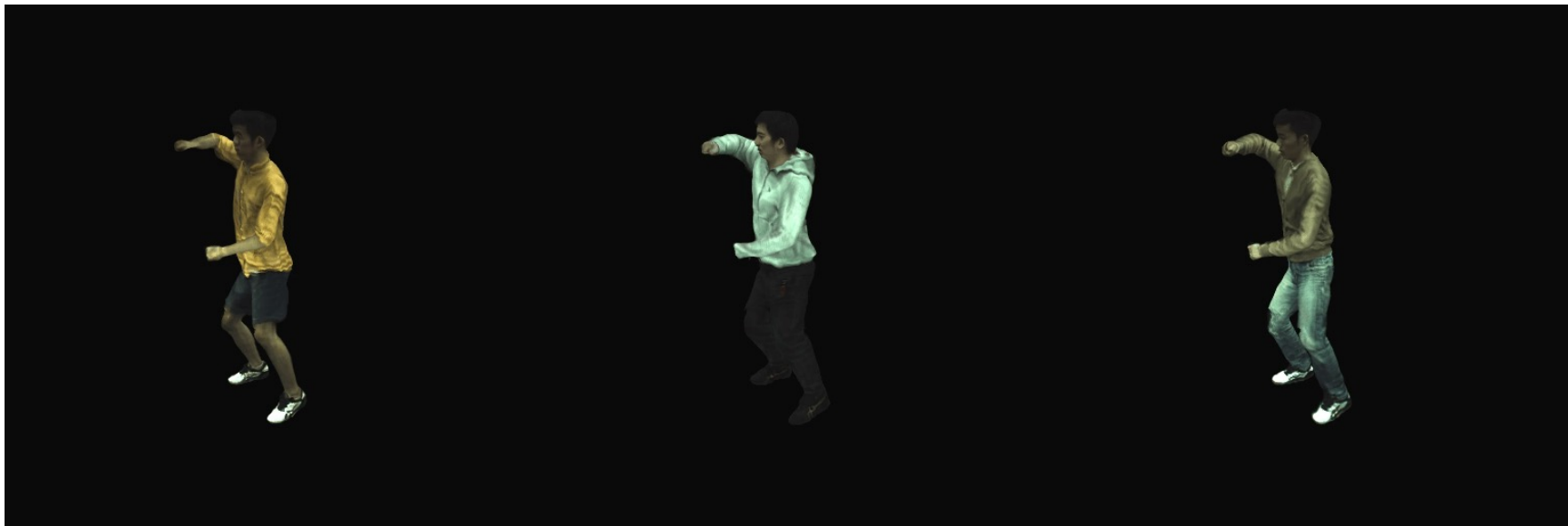


ActorsNeRF synthesizes novel views of the person with novel poses.

# **ZJU-MoCap Dataset**



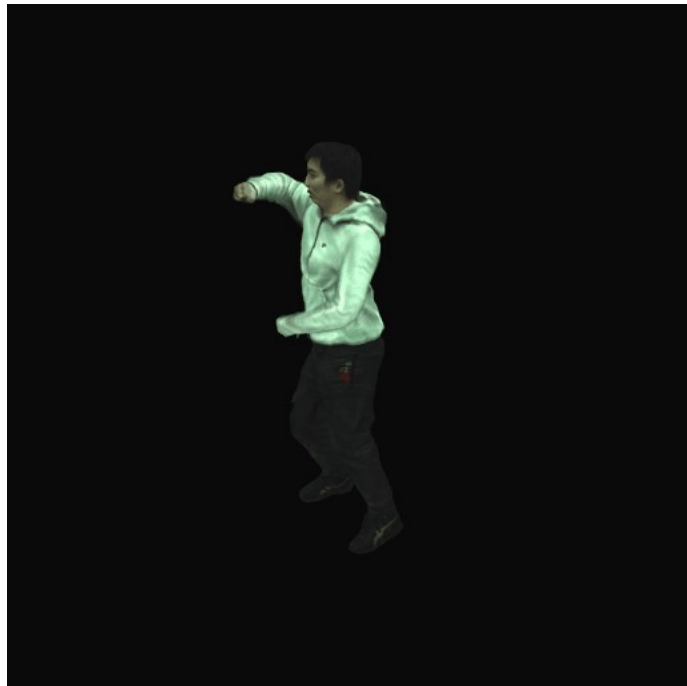
# Animation



With only a few images (e.g., 30 shots) from a monocular video, ActorsNeRF synthesizes **novel views** of a **novel person** with **novel poses**.

# 100-shot Comparison

Ours



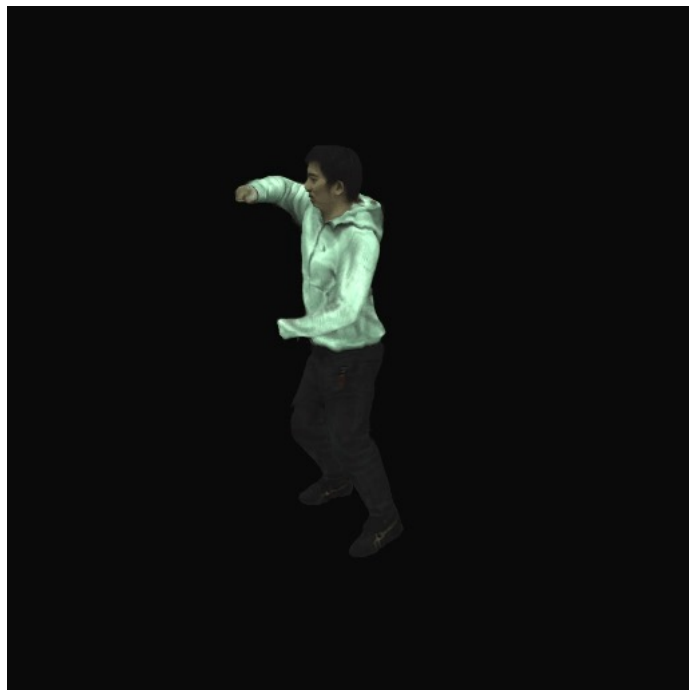
HumanNeRF



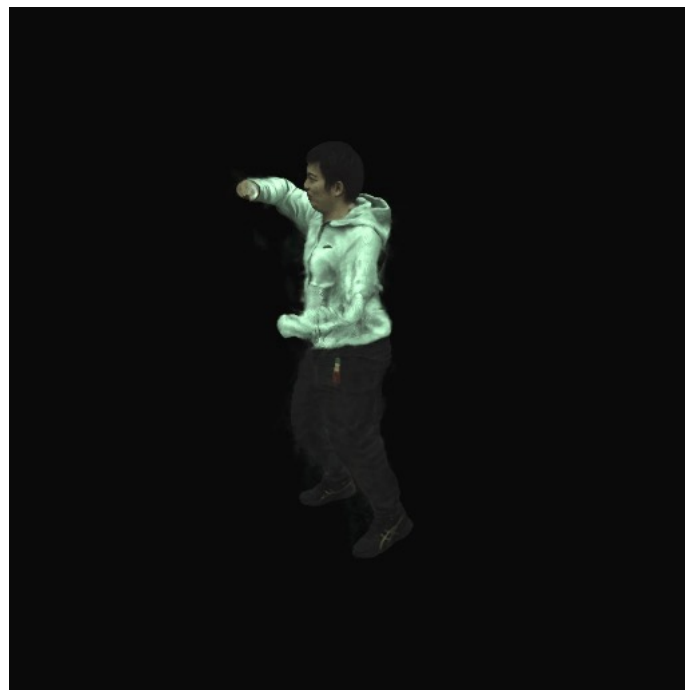
ActorsNeRF synthesizes novel views of the person with novel poses.

# 30-shot Comparison

Ours



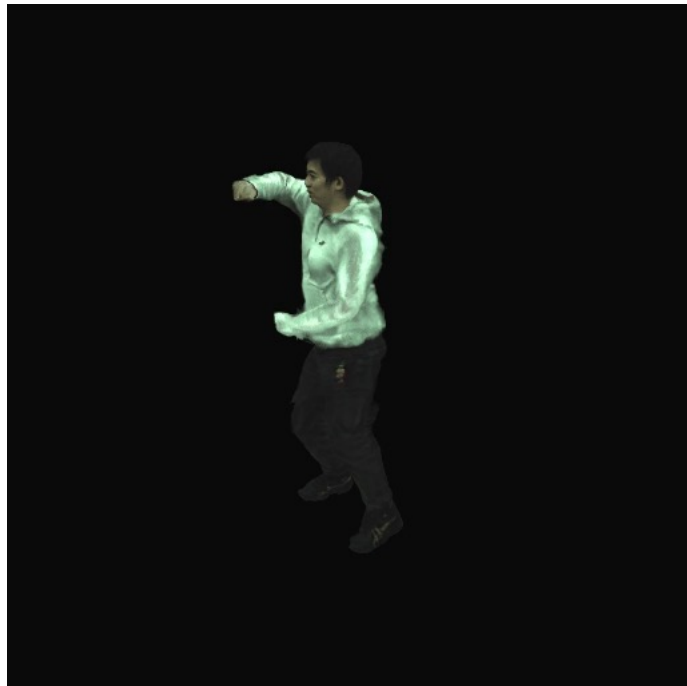
HumanNeRF



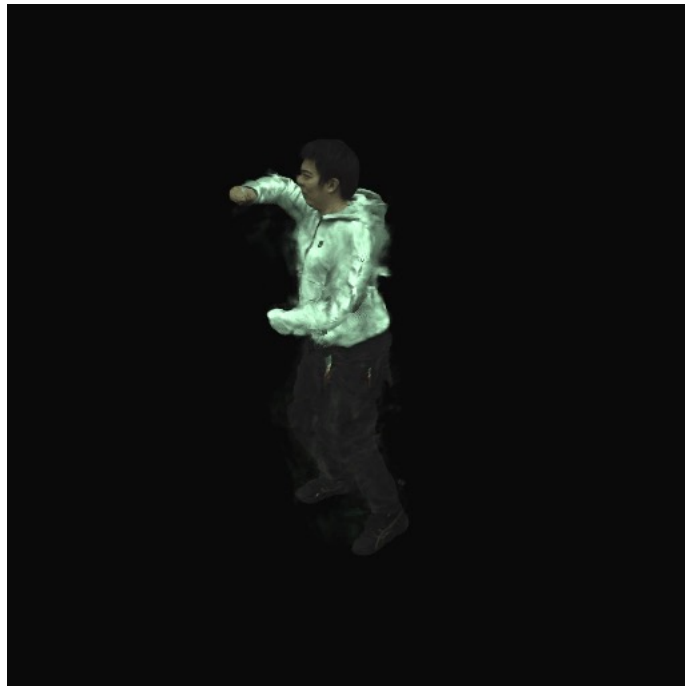
ActorsNeRF synthesizes novel views of the person with novel poses.

# 10-shot Comparison

Ours

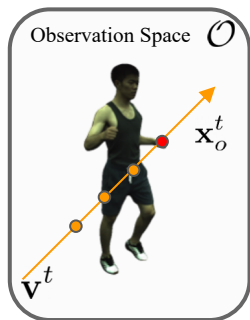


HumanNeRF

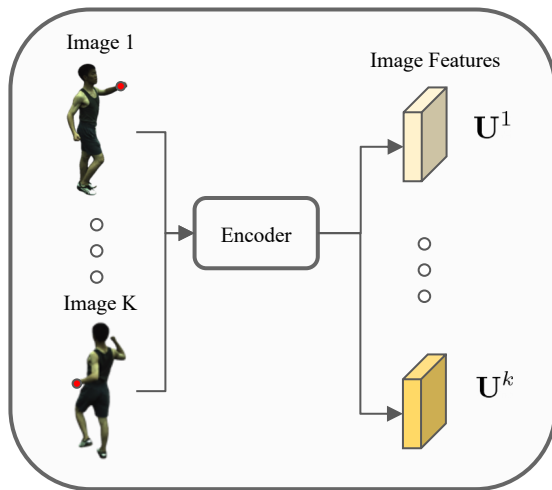
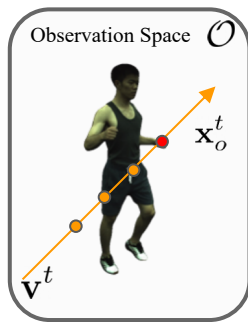


ActorsNeRF synthesizes novel views of the person with novel poses.

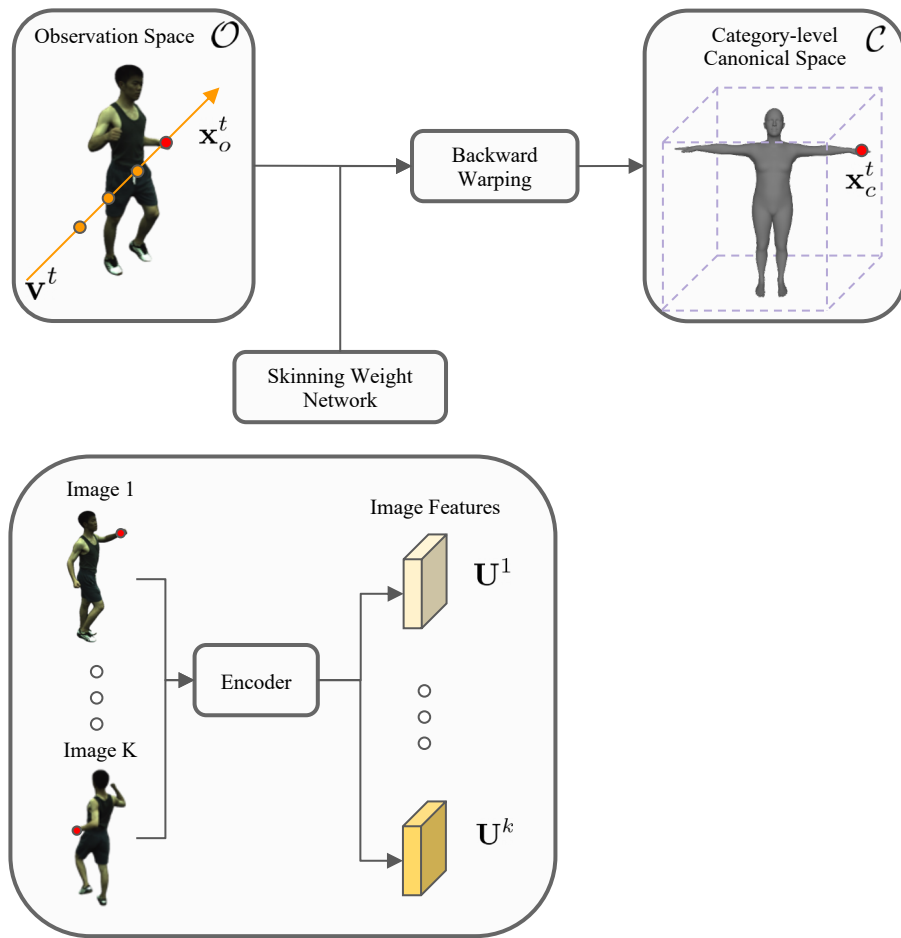
# **ActorsNeRF Architecture**



Sample 3D points in the observation space.

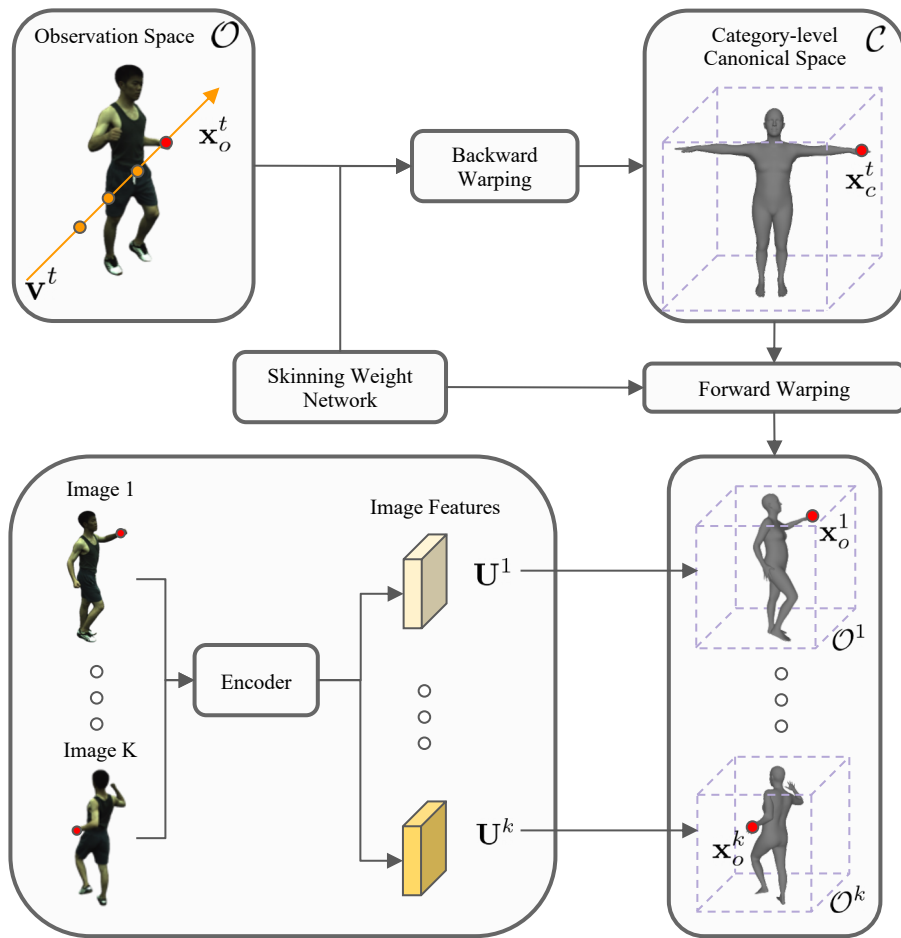


An encoder used to extract features from K images.

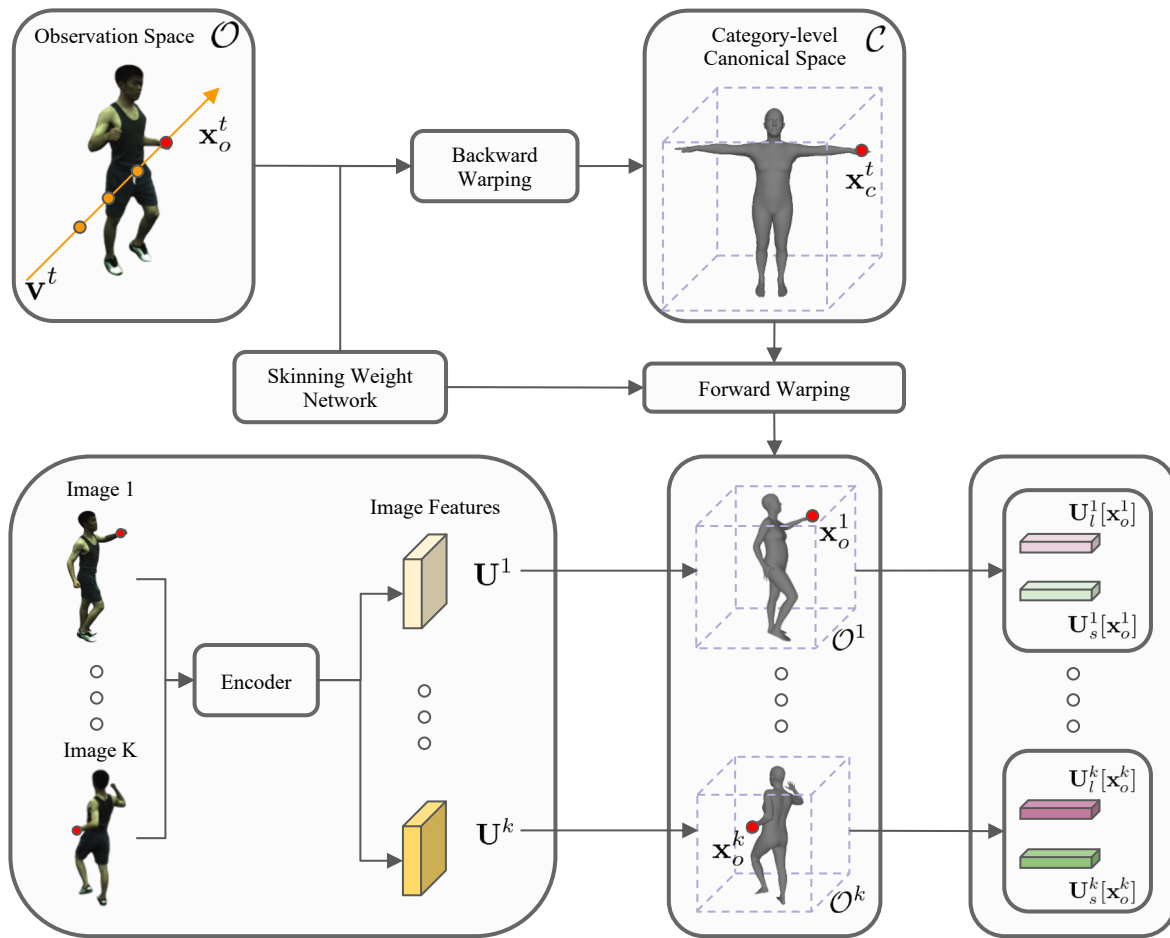


Transform to category-level canonical space.

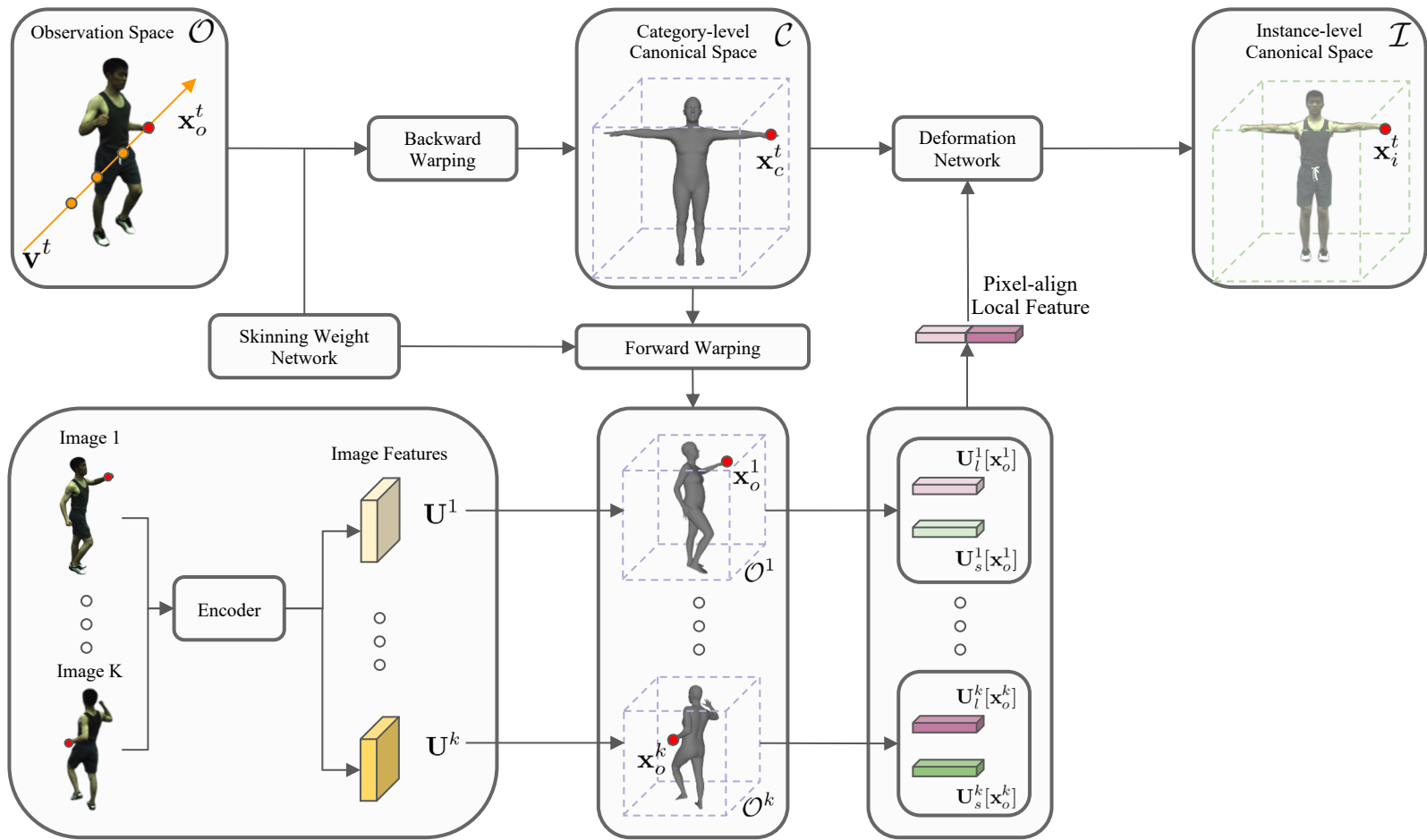




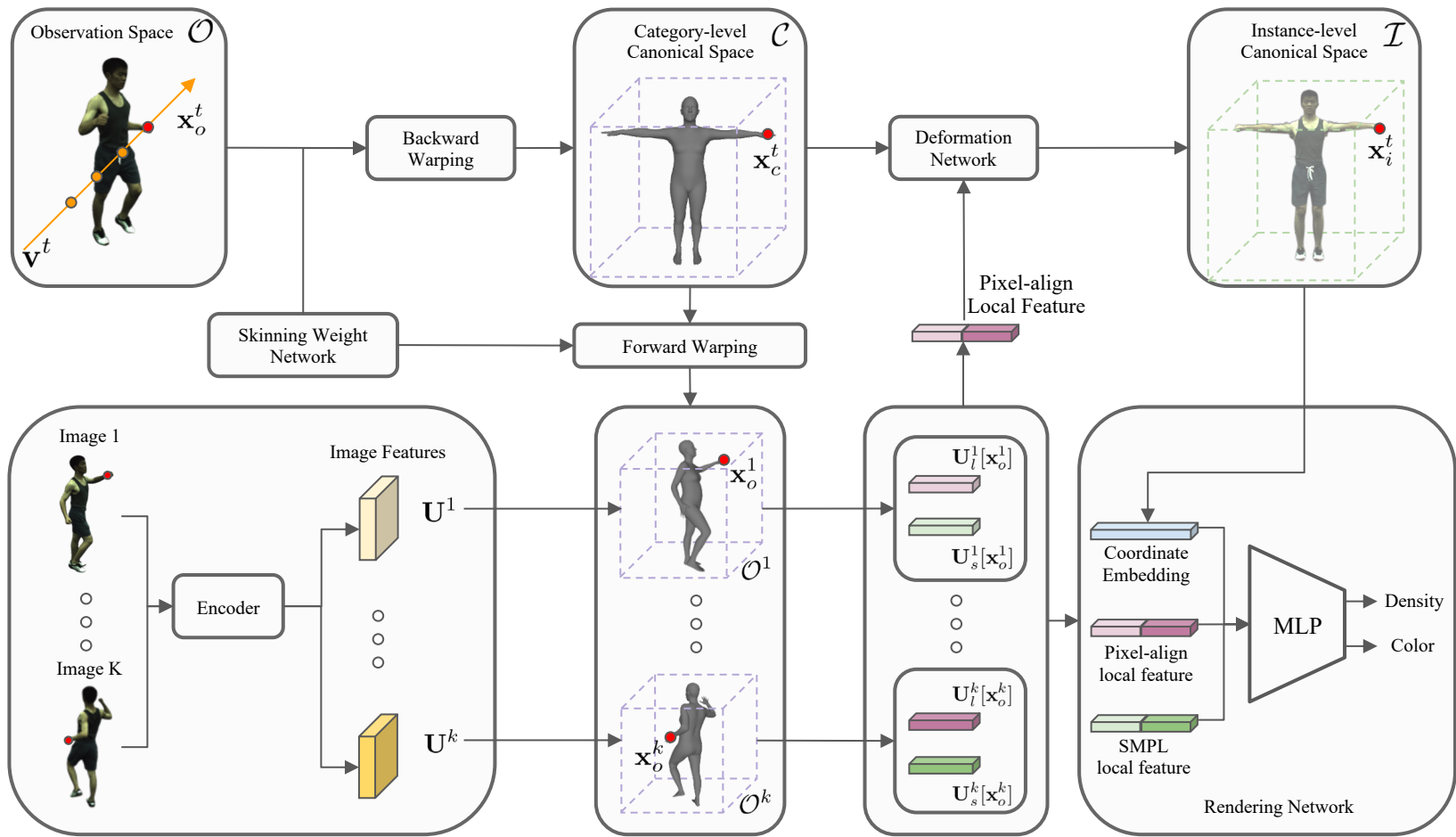
Transform to individual observation spaces through forward warping.



Query pixel aligned features from each image.



Transform to instance-level canonical space through deformation network.



Rendering from instance-level canonical space with pixel-aligned features.