

Code Released!

# AnyMoLe: Any Character Motion In-betweening Leveraging Video Diffusion Models

Kwan Yun, Seokyeon Hong, Chaelin Kim, Junyong Noh

CVPR2025 –Main paper  
Sunday, June 15

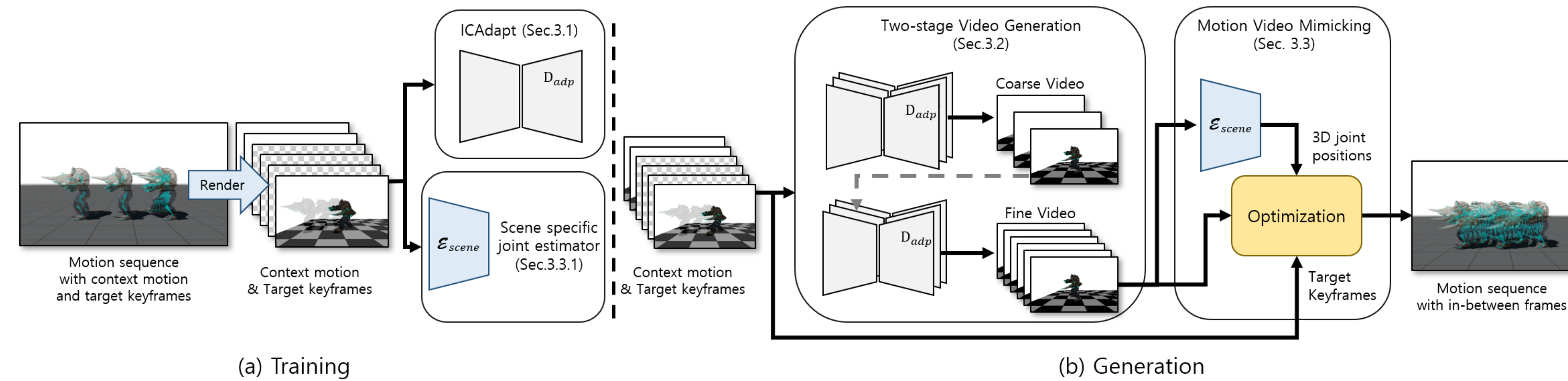
**TL;DR:** AnyMoLe performs motion in-betweening for arbitrary characters without training data.

### Motivation

- Motion In-betweening reduces the manual burden of artists by automatically generate the natural motion from sparse key frames.
- But, previous motion in-betweening methods require large motion data in one specific joint configuration, making another burden of making huge dataset for every characters.
- AnyMoLe leverages video diffusion model for arbitrary character without training data

### TakeAways:

- A video generation model can be used for an arbitrary 3D character animation.
- A few seconds of motion can effectively fine-tune a video diffusion model.
- Fitted to a specific scene and character, the few-shot model outperforms large general models.



**Method Summary:** Render the character, personalize video diffusion model, generate video, and optimize character’s motion using the generated video.

### Method:

- Given context motion and target key frames for a 3D character, AnyMoLe performs 3D motion inbetweening using a video diffusion model in following order
- Using the two seconds of context motion and target keyframes, render each frame from diverse views.
- From the multi-view rendered images, finetune the video diffusion model and train a scene-specific joint estimator concurrently using both 2D and 3D-aware features.
- The finetuned video diffusion model generates an in-between video in a two stage auto-regressive manner: first, generate coarse frames from the given context and key frames, second, fill in the remaining frames.
- Optimize character motion sequentially to align its motion with the generated video using a differentiable renderer and the trained joint estimator.

