

# LaVR: Scene Latent Conditioned Generative Video Trajectory

## Re-Rendering using Large 4D Reconstruction Models

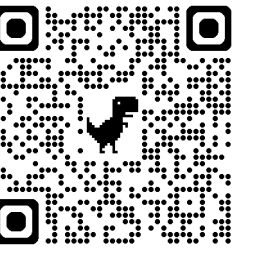
Mingyang Xie Numair Khan Tianfu Wang Naina Dhingra Seonghyeon Nam Haitao Yang Zhuo Hui  
Christopher Metzler Andrea Vedaldi Hamed Pirsiavash Lei Luo

From a **monocular** video of a **dynamic** scene,  
Generate a video of the same scene along a **new camera trajectory**.

Mingyang Xie WeChat ID: mingyangxie

Let's Chat!

5th year PhD Student @ UMD  
Email: mingyang@umd.edu  
Seeking Full time roles in 2027.



About LaVR,  
Research, or  
Anything!

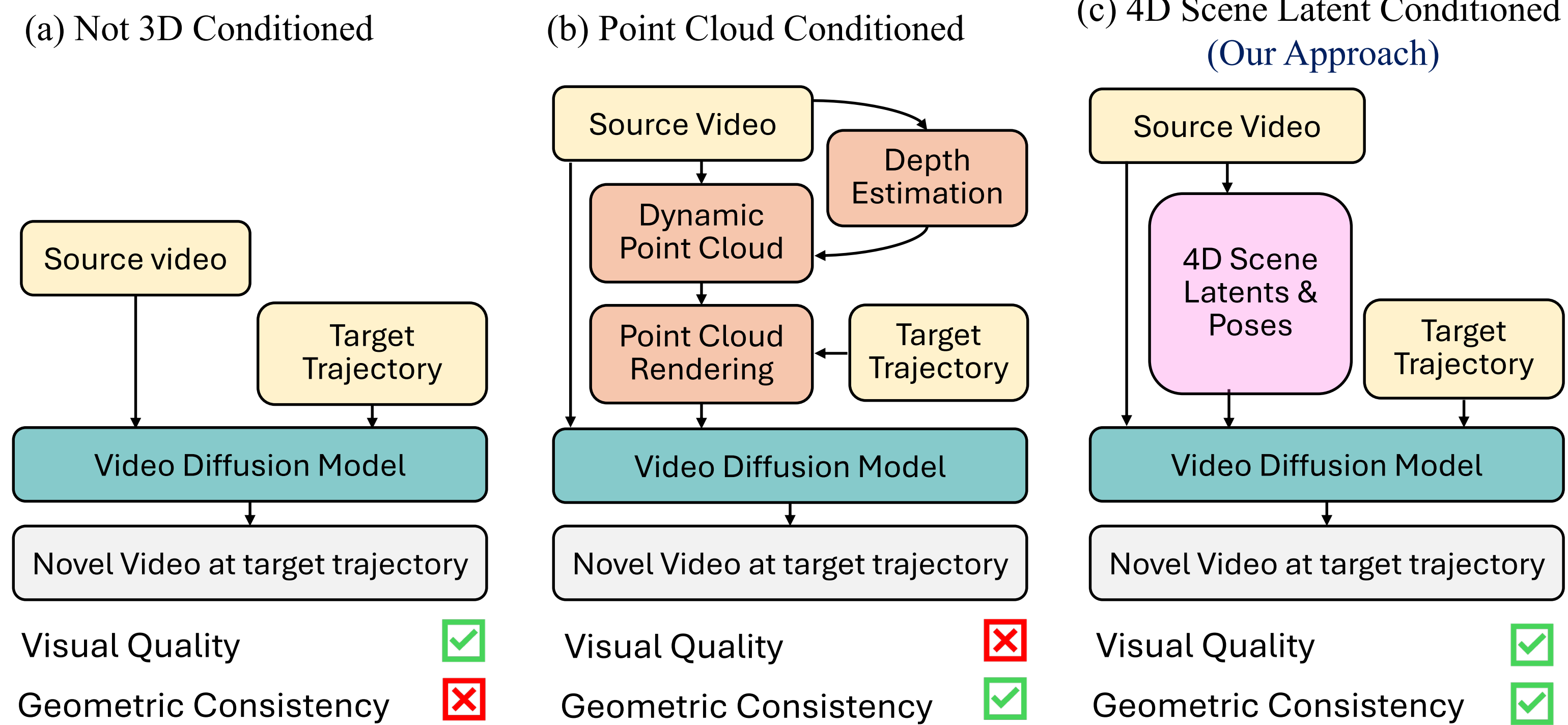
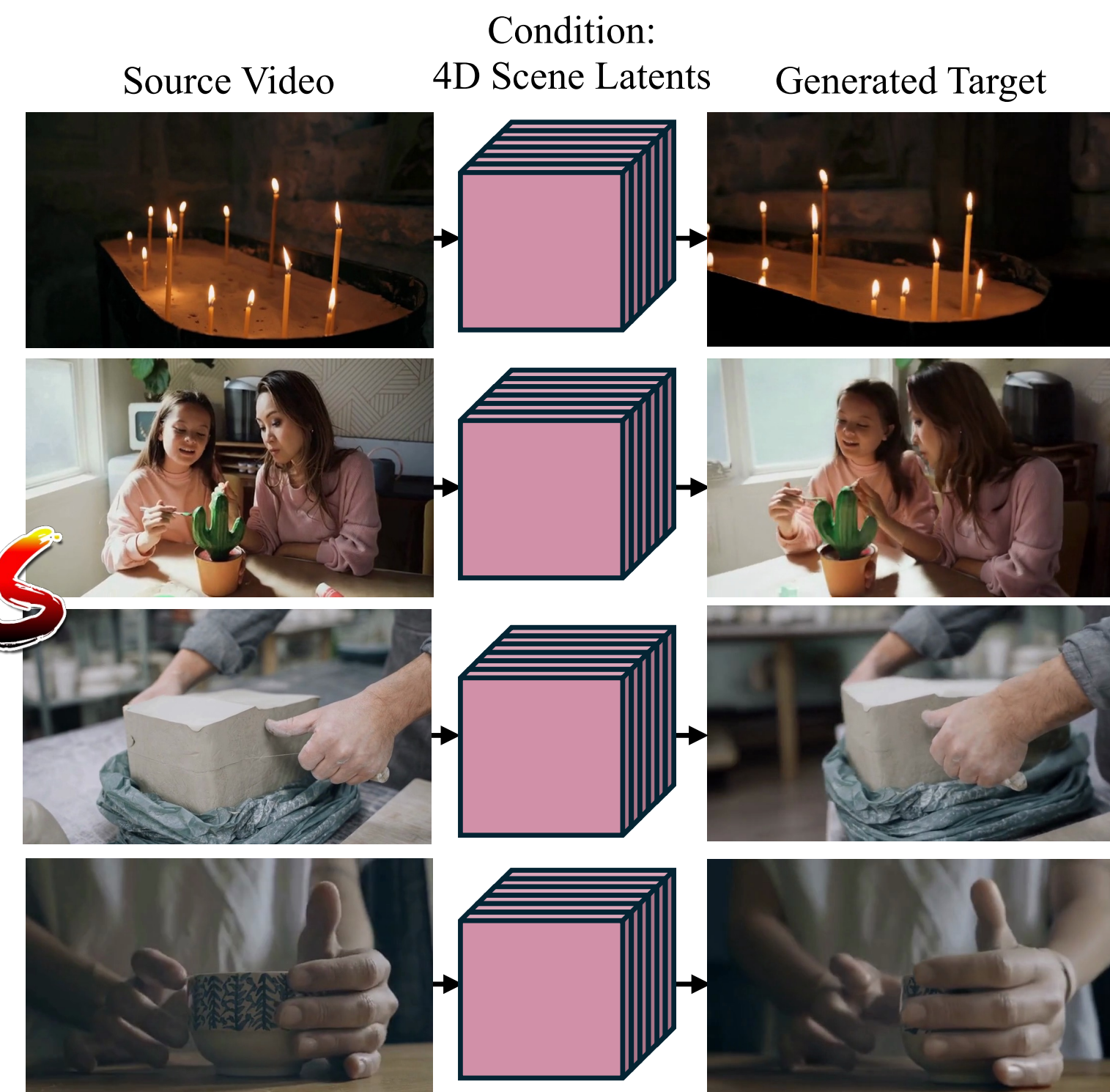
Tianfu Wang WeChat ID: Deefyfe

2nd year PhD Student @ UMD  
Webpage: <https://tianfwang.github.io/>  
Looking for Summer Internship in 2027.

### Point Cloud Based Methods



### Ours

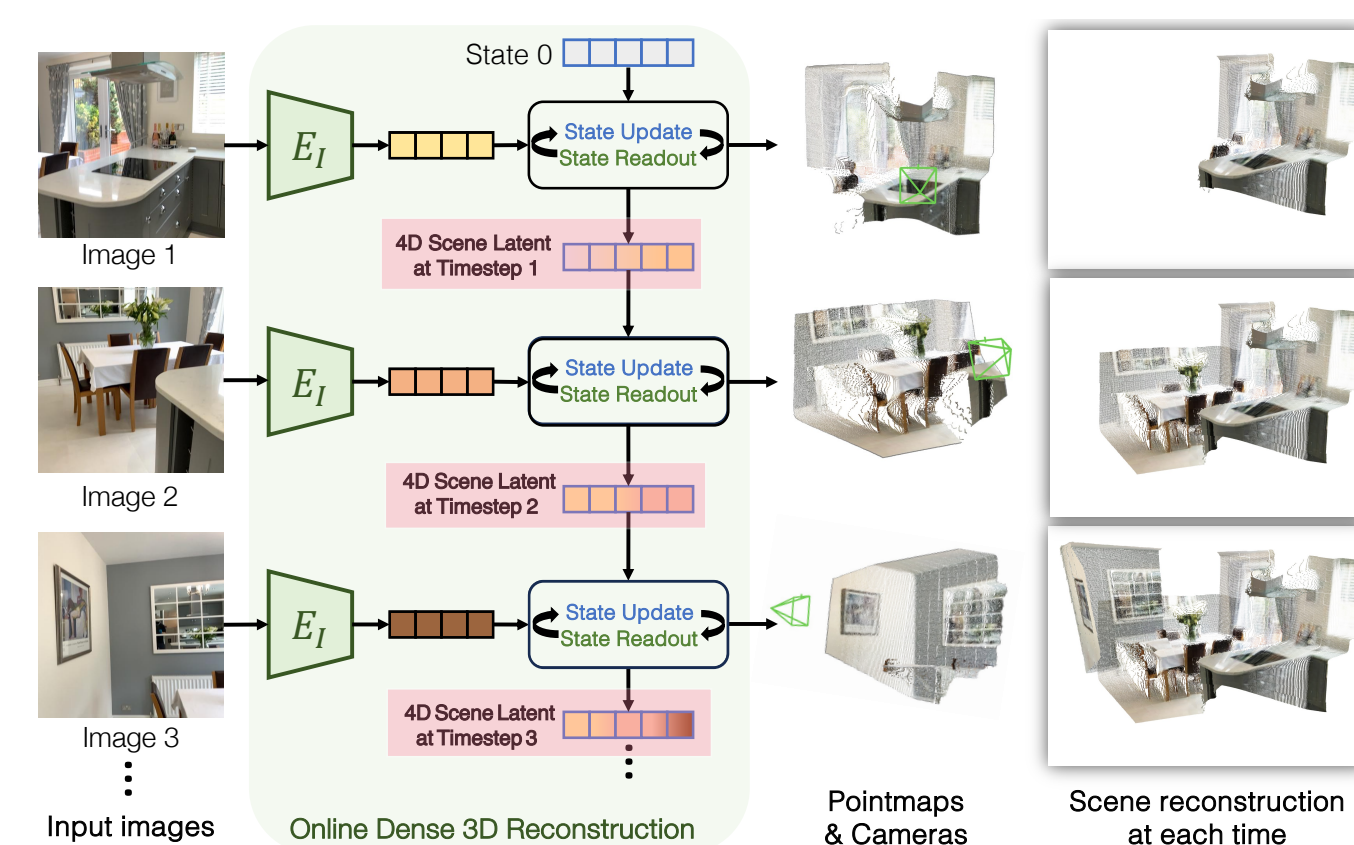


Our model is conditioned on scene latents from a large 4D reconstruction model.

### Our Motivations

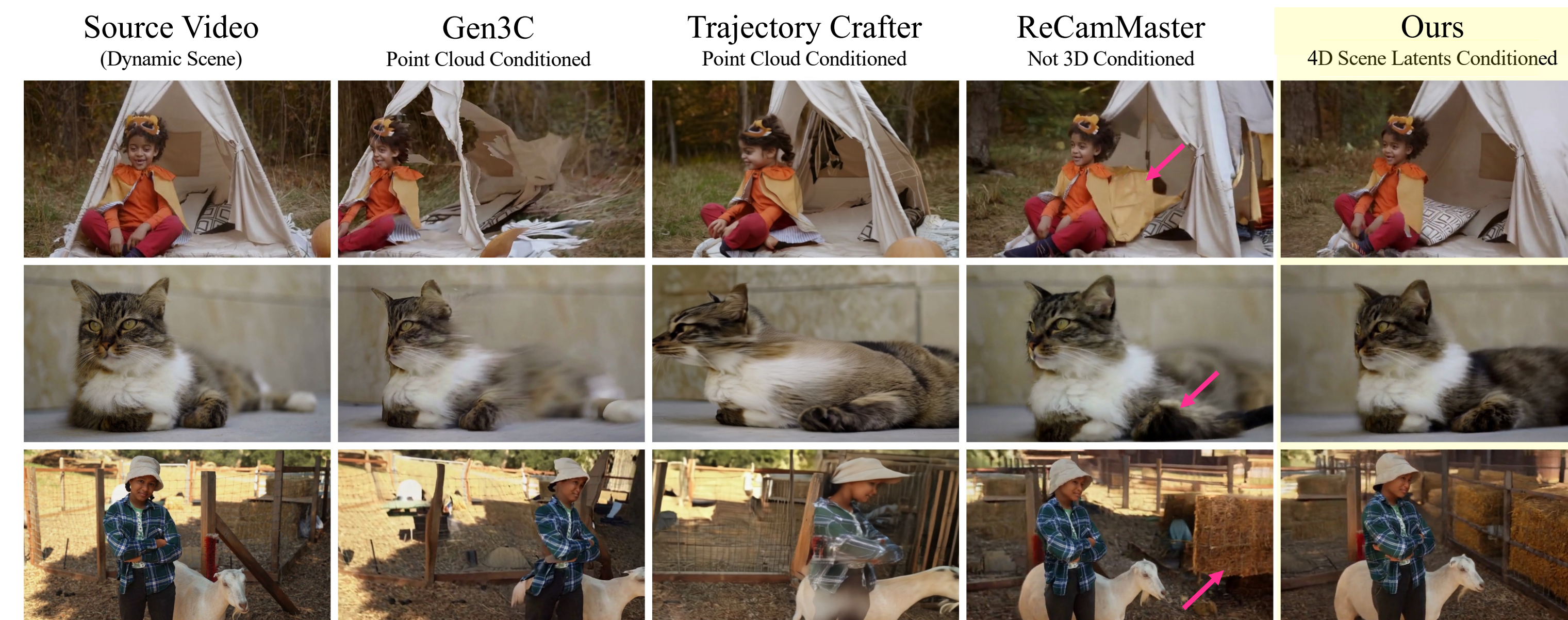
- Prior works guide video generation with re-rendered point clouds, which imposes an explicit and relatively hard geometric condition.
- However, even small depth-level errors in point clouds can appear as severe geometric artifacts under novel viewpoints, which then propagate into video generation.
- We instead use 4D scene latents from a Large 4D Reconstruction Model as a softer and more abstract geometric condition for video generation.
- This lets the pretrained video diffusion prior regularize imperfect geometry/depths into more realistic outputs.

### What is a Large 4D Reconstruction Model?

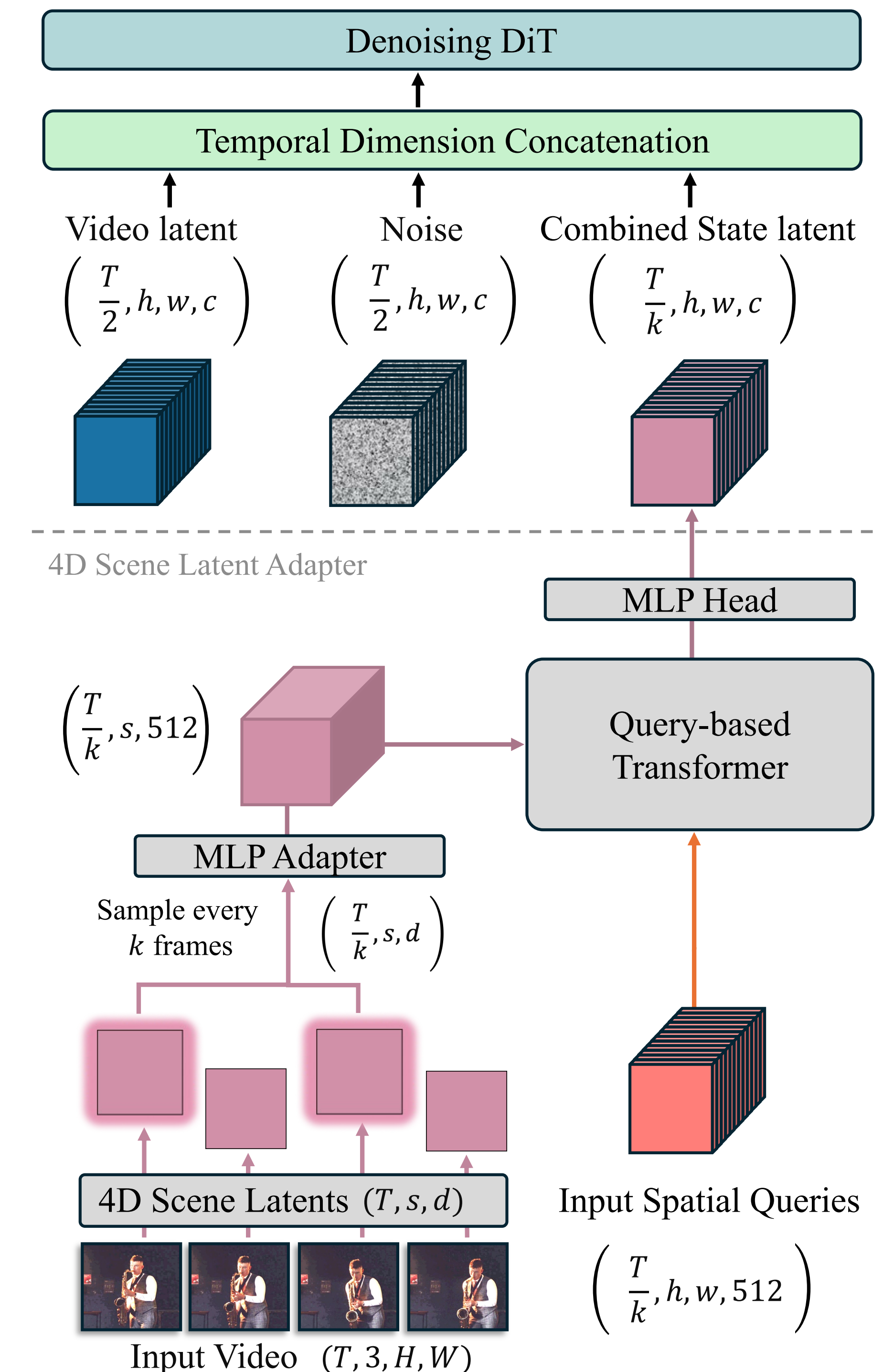


Such a model predicts the latent representation of a 4D scene at every timestep, from which it further predicts the point maps and camera poses.

\*Model architecture of CUT3R. Figure by Wang Qianqian et al. CVPR 2025. More at their webpage <https://cut3r.github.io/>.



Compared to ours, point cloud based methods are prone to depth estimation errors, methods with no 3D conditioning are prone to shape deformation.



Details of Our 4D Scene Latent Conditioning.