

# Beyond the Frame:

## Generating 360° Panoramic Videos from Perspective Videos



Rundong Luo<sup>1</sup>



Matt Wallingford<sup>2</sup>



Ali Farhadi<sup>2</sup>



Noah Snavely<sup>1</sup>



Wei-Chiu Ma<sup>1</sup>

<sup>1</sup>Cornell University

<sup>2</sup>University of Washington

# Video Generation Models are powerful



Veo3 (proprietary model)



Wan2.2 (open-sourced model)



# Video Generation Models are powerful



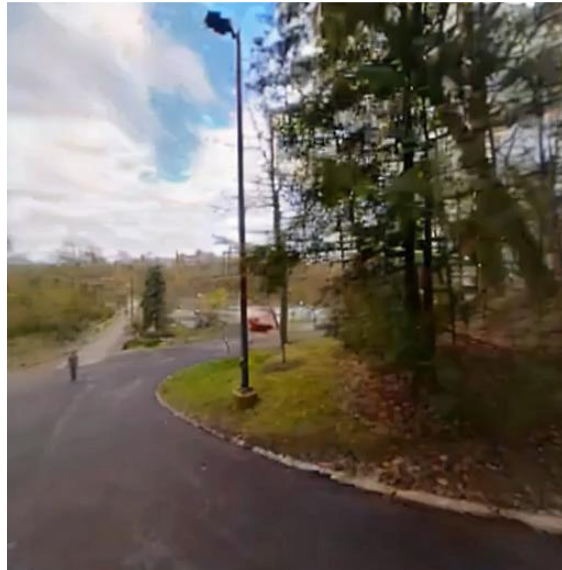
Veo3 (proprietary model)



Wan2.2 (open-sourced model)

## What if we want to explore beyond?

# Video outpainting?



SoTA video outpainting approach [1]

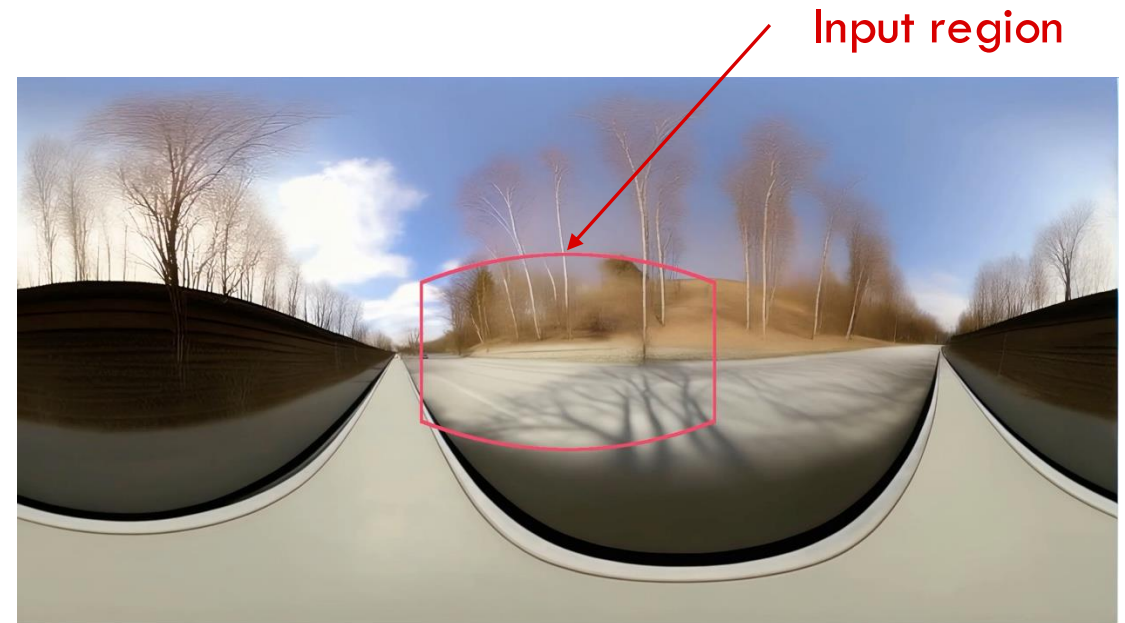
# Video-to-360°:

## Transform perspective video to 360° video



Input perspective video

Video-to-360°



Generated 360° video

Challenge: spatial layout & object dynamics



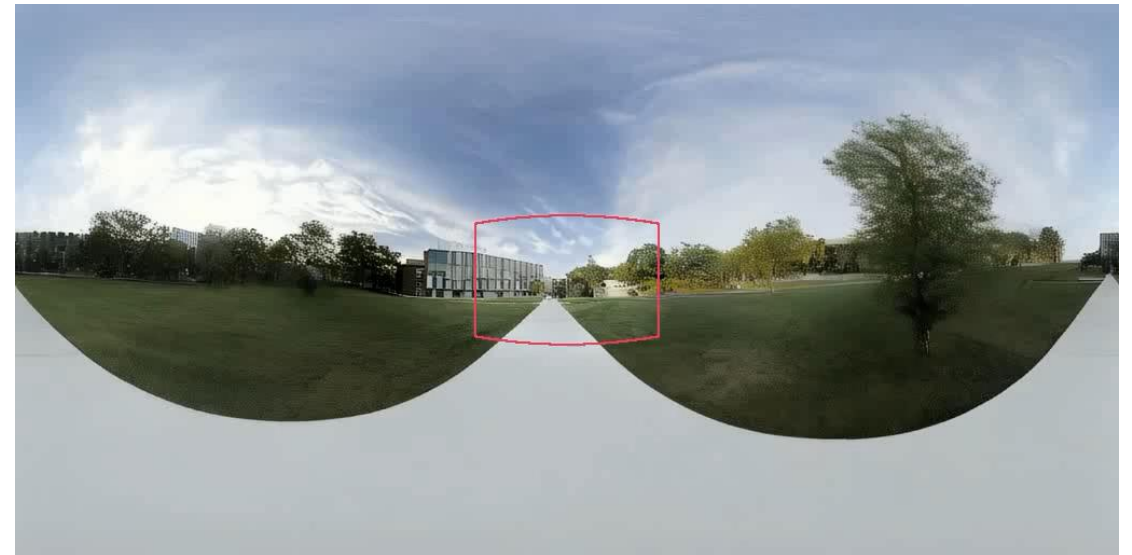
# Video-to-360°:

## Transform perspective video to 360° video



Input perspective video

Video-to-360°

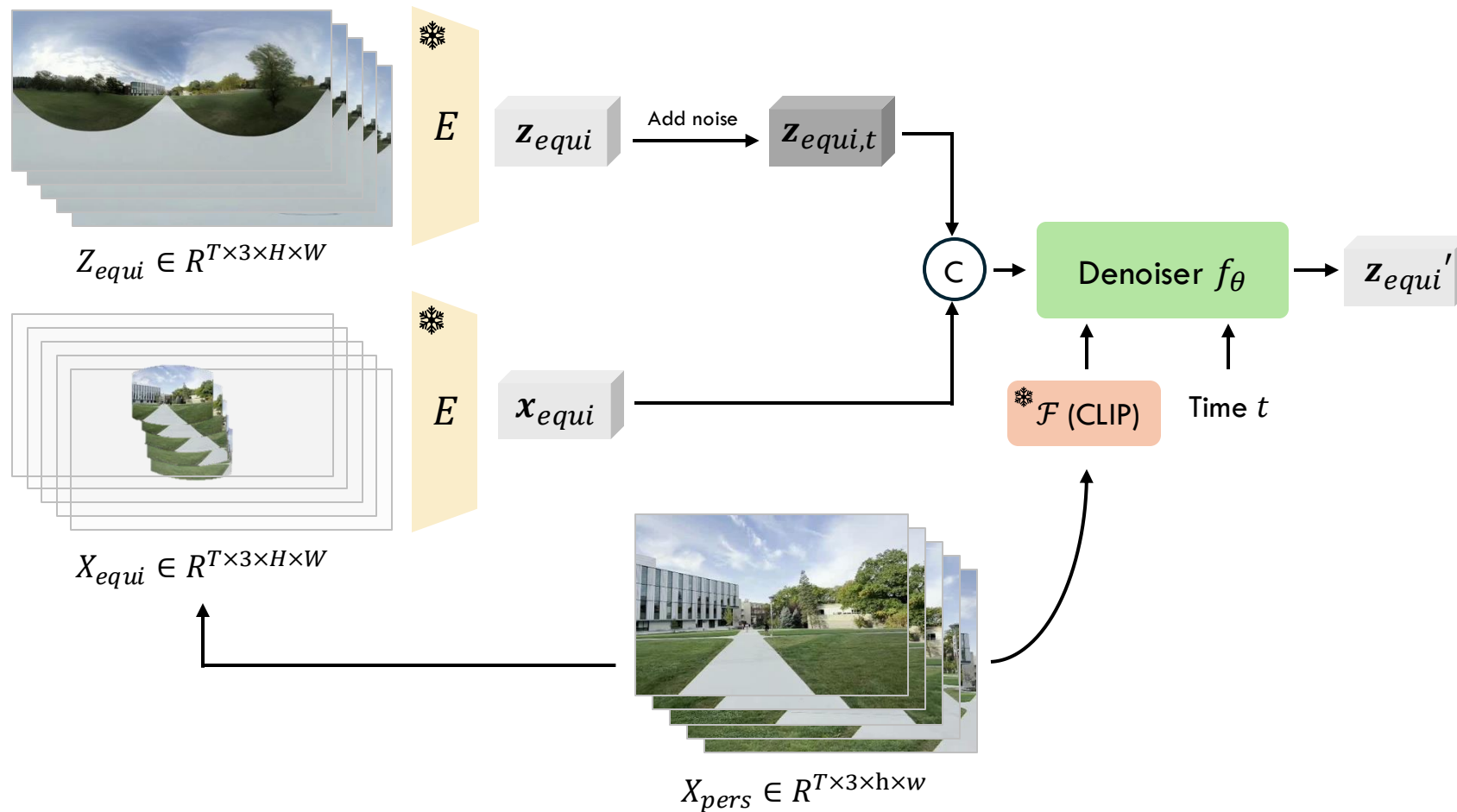


Generated 360° video

# Model Architecture (training)

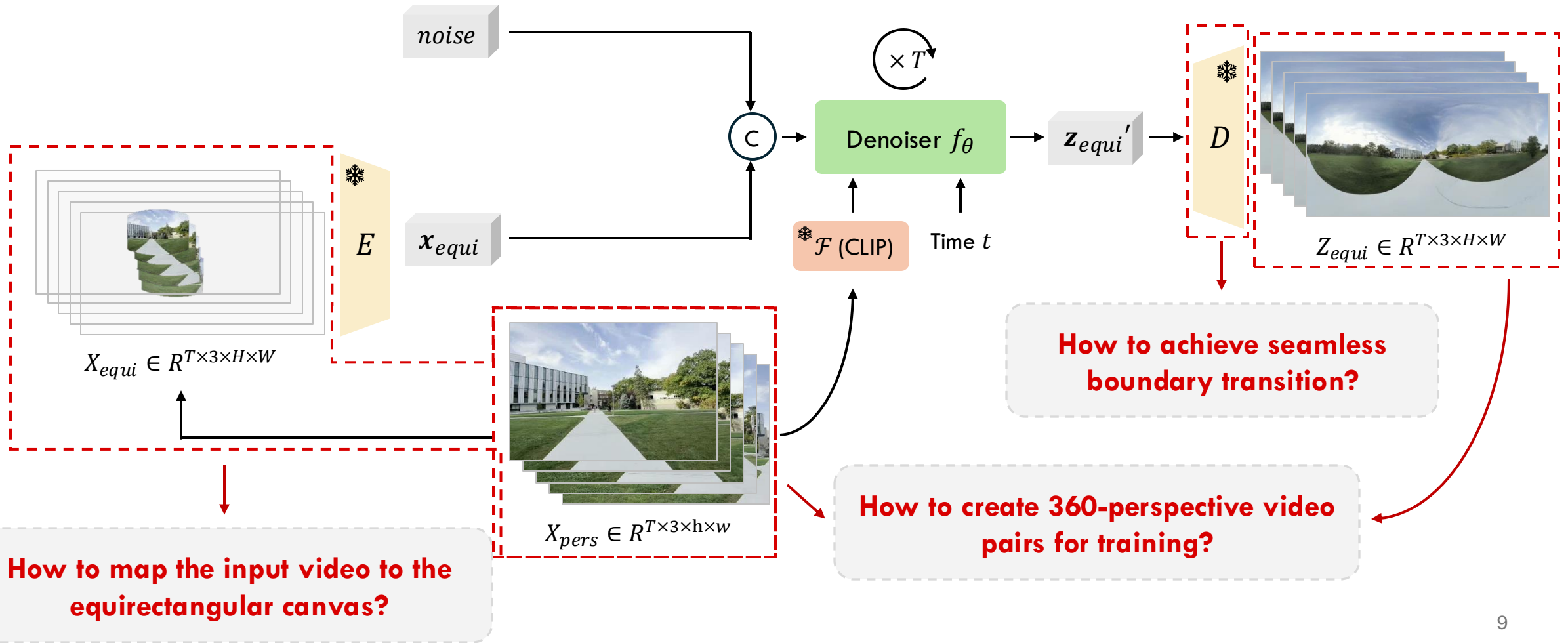


# Model Architecture (training)





# Model Architecture (inference)



# Challenge 1: Data Curation

- **Camera motion simulation:** extract perspective video from 360° video



360° video for training

Camera motion  
simulation  
→

Extracted perspective  
video

Masked equirectangular  
video

# Challenge 2: Map Input Video to the Canvas



Input perspective video

Map to the  
center

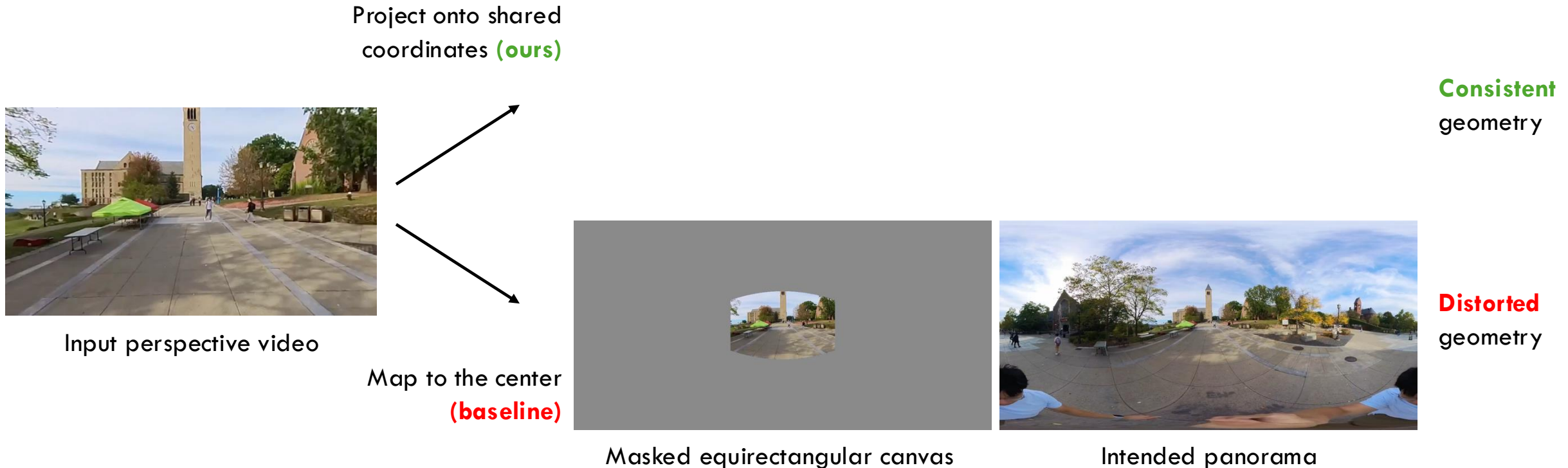


Masked equirectangular canvas

**Distorted**  
geometry

# Challenge 2: Map Input Video to the Canvas

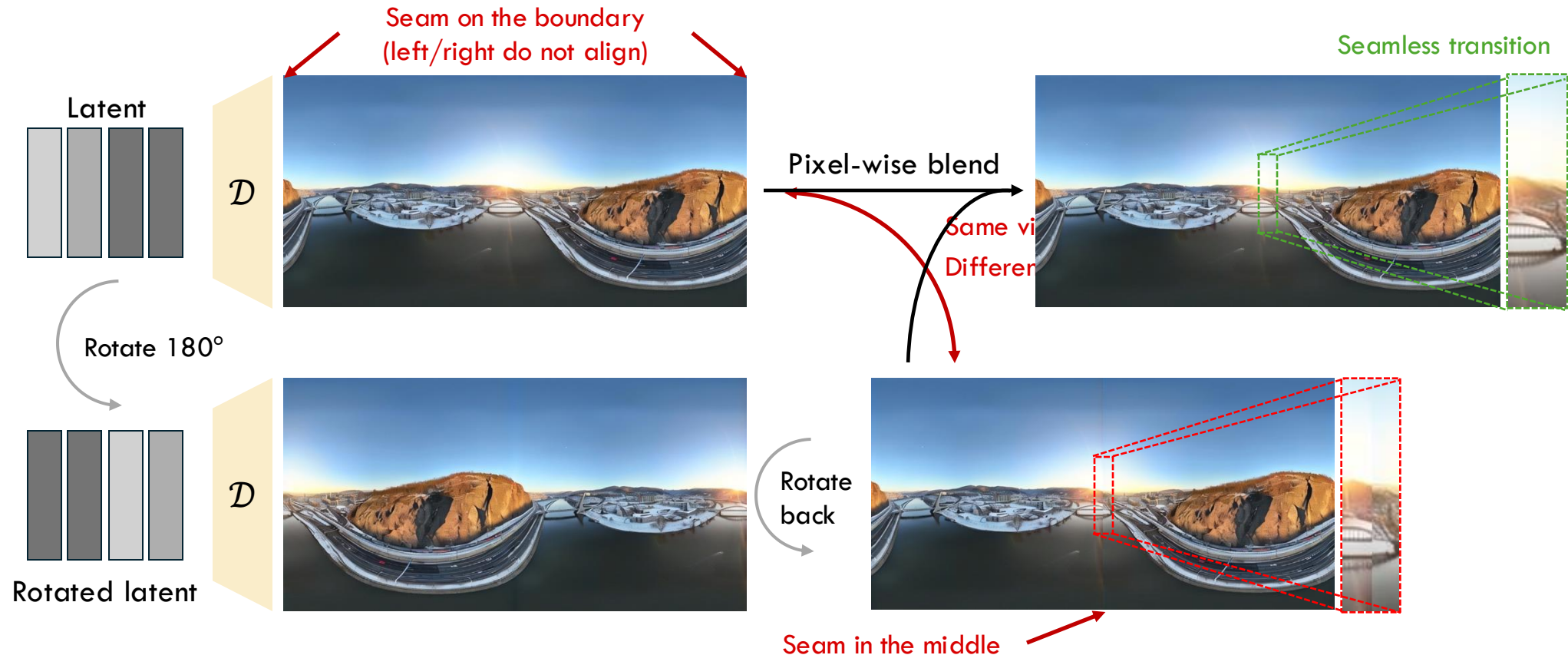
- **View-based frame alignment:** Ensure each part of the canvas corresponds roughly the same scene region across frames
- Inference: first calibrate the video with MegaSAM, then align to the first frame's coordinates





# Challenge 3: Seamless Boundary Transition

- **Blended decoding:** decode original latents and its 180°-rotated version, then blend together.



360DVD [1] blend results in latent space, while we blend in pixel space.

[1] Wang et al. 360DVD: Controllable Panorama Video Generation with 360-Degree Video Diffusion Model. In CVPR, 2024.

# Data

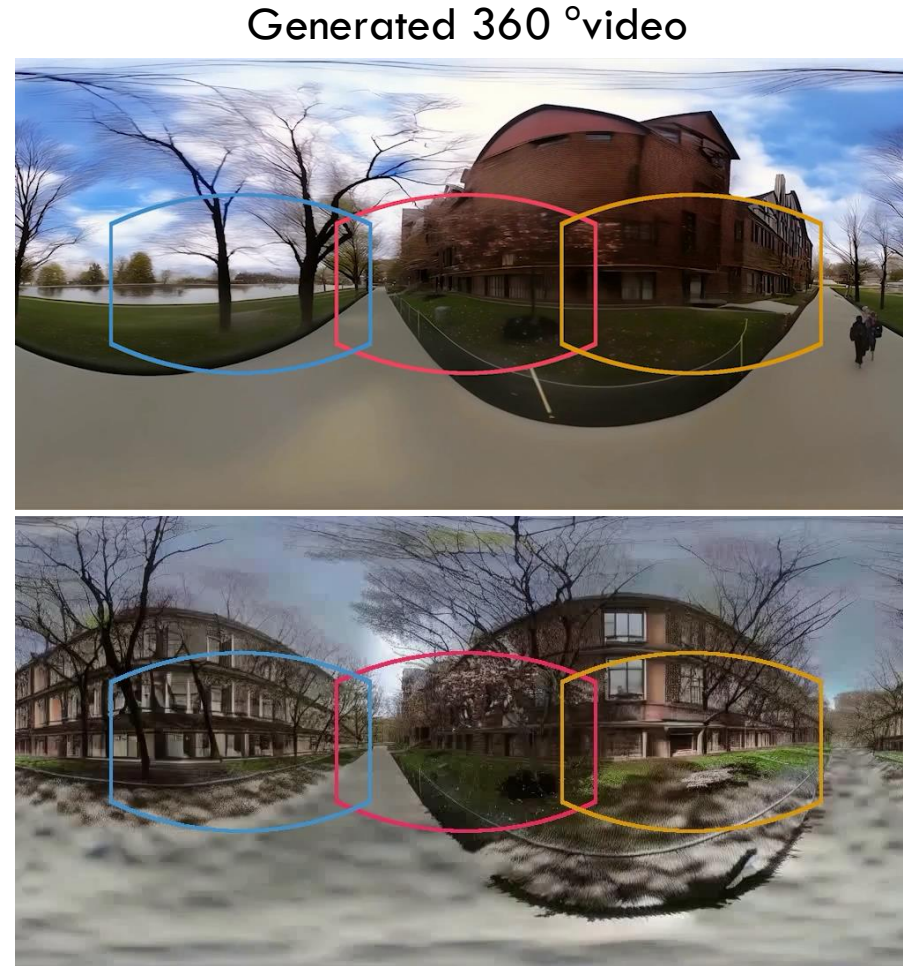
- We start from 360°-1M [1], which contains 1 million 360° videos YouTube of varying quality.
- Systematically filter to 8,686 videos and 283,863 10-second video clips



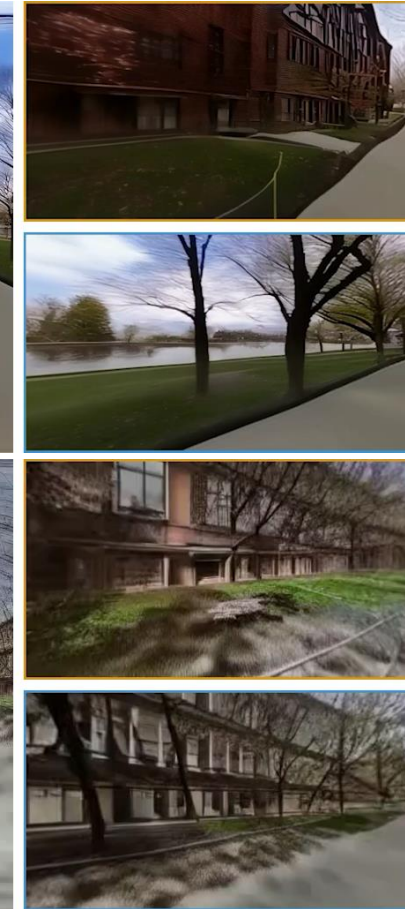


# Experiments

- Comparison with PanoDiffusion (image-to-360°) [1]



Unwrapped  
perspective views



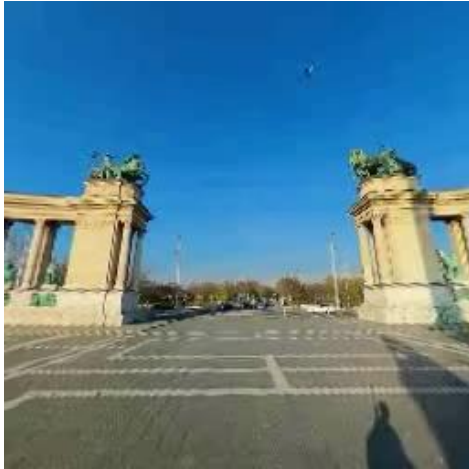
Ours

PanoDiffusion

# Experiments

- Comparison with Follow-Your-Canvas (video outpainting) [1]

Input video



Generated 360° video



Unwrapped  
perspective views



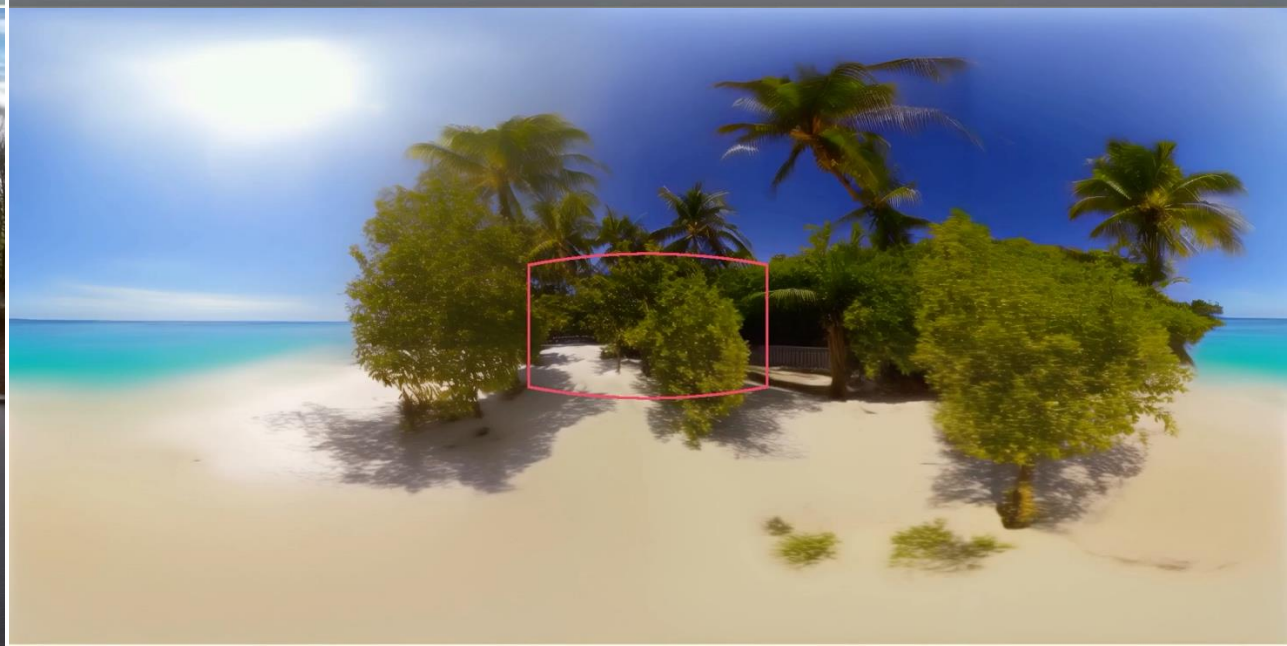
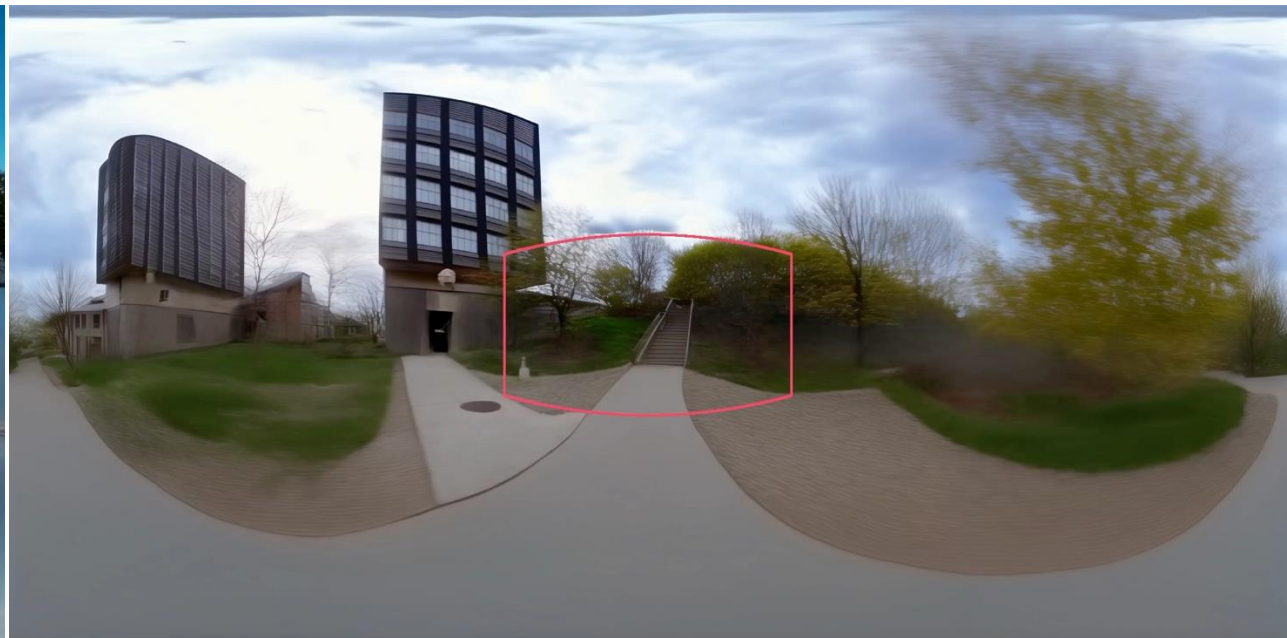
Ours



Follow-Your-Canvas







# Analysis: Interpret scene dynamics

**Question:** Can model predict the temporal dynamics of objects even they are beyond the frame?



Captured 360° video



60°-FoV perspective video



Generated 360° video

# Analysis: Interpret scene dynamics

**Question:** Can model predict the temporal dynamics of objects even they are beyond the frame?

---



60°-FoV perspective video



Generated 360° video



Extracted perspective video



# Analysis: Geometric consistency

- **Question:** Do the lines remain straight in the generated viewpoints?



Generated view (left)

w/ detected lines

Input view

w/ annotated lines

Generated view (right)

w/ detected lines

Method	PanoDiffusion	Imagine360	Ours
Line Consistency	0.6504	0.5623	<b>0.8506</b>



# Applications

# Applications : Video Stabilization

Input



Stabilization (ours)



Stabilization (reference)



# Applications: Camera control

Input



Clockwise rotate 30°



Clockwise rotate 45°





# Applications: Interactive VQA

**Question:** Determine **whether the red car overlaps** with the crosswalk in the second frame of a two-frame sequence extracted from a video.



Video of interest

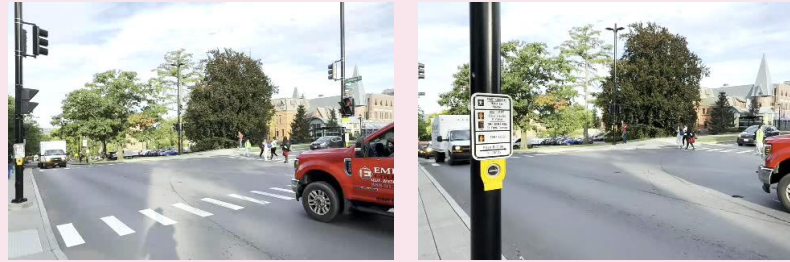


# Applications: Interactive VQA



Video of interest

**Question:** Determine **whether the red car overlaps** with the crosswalk in the second frame of a two-frame sequence extracted from a video.



**Answer:** In the second frame, the red car **does not overlap** with the crosswalk. It is positioned just beyond the crosswalk line.

**Question:** How should I **change the viewing direction** to verify your judgment?



**Answer:** To verify whether the red car overlaps with the crosswalk, you can **adjust your viewing direction by shifting slightly to the right** and lowering the camera angle. This perspective would give you a clearer line of sight along the crosswalk and the red car's position relative to it.



# Applications: Interactive VQA



**Question:** How should I **change the viewing direction** to verify your judgment?



**Answer:** To verify whether the red car overlaps with the crosswalk, you can **adjust your viewing direction by shifting slightly to the right** and lowering the camera angle. This perspective would give you a clearer line of sight along the crosswalk and the red car's position relative to it.



Our system



# Applications: Interactive VQA



Our system

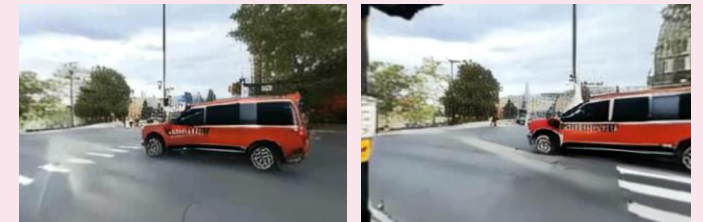


**Question:** How should I **change the viewing direction** to verify your judgment?



**Answer:** To verify whether the red car overlaps with the crosswalk, you can **adjust your viewing direction by shifting slightly to the right** and lowering the camera angle. This perspective would give you a clearer line of sight along the crosswalk and the red car's position relative to it.

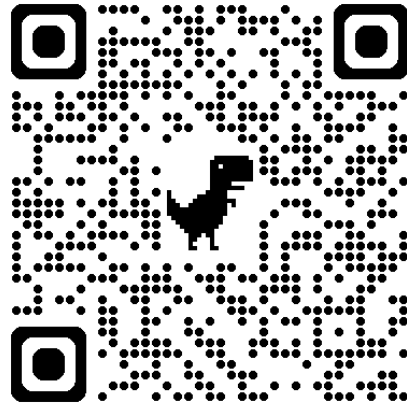
**Question:** These are the images captured after a 30-degree clockwise rotation. Re-evaluate and determine if the red car overlaps with the crosswalk in the second frame.



**Answer:** In the second frame after the 30-degree clockwise rotation, the red car does **indeed overlap with the crosswalk**. The front wheels and part of the car's front section extend over the crosswalk lines, making it clear that the vehicle intrudes into the crosswalk area from this new perspective.



Poster: 10/22 10:45-12:45 Session 3, #408



Thanks for listening!

