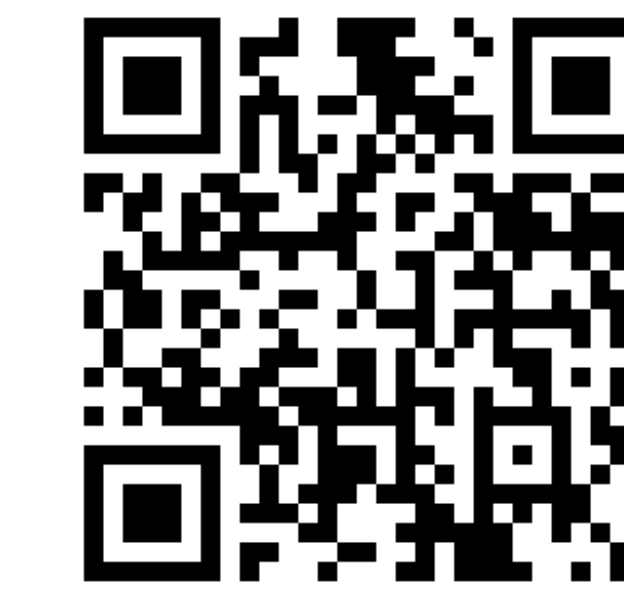


Making 360° Video Watchable in 2D: Learning Videography for Click Free Viewing

Yu-Chuan Su and Kristen Grauman
The University of Texas at Austin
<http://vision.cs.utexas.edu/projects/watchable360>

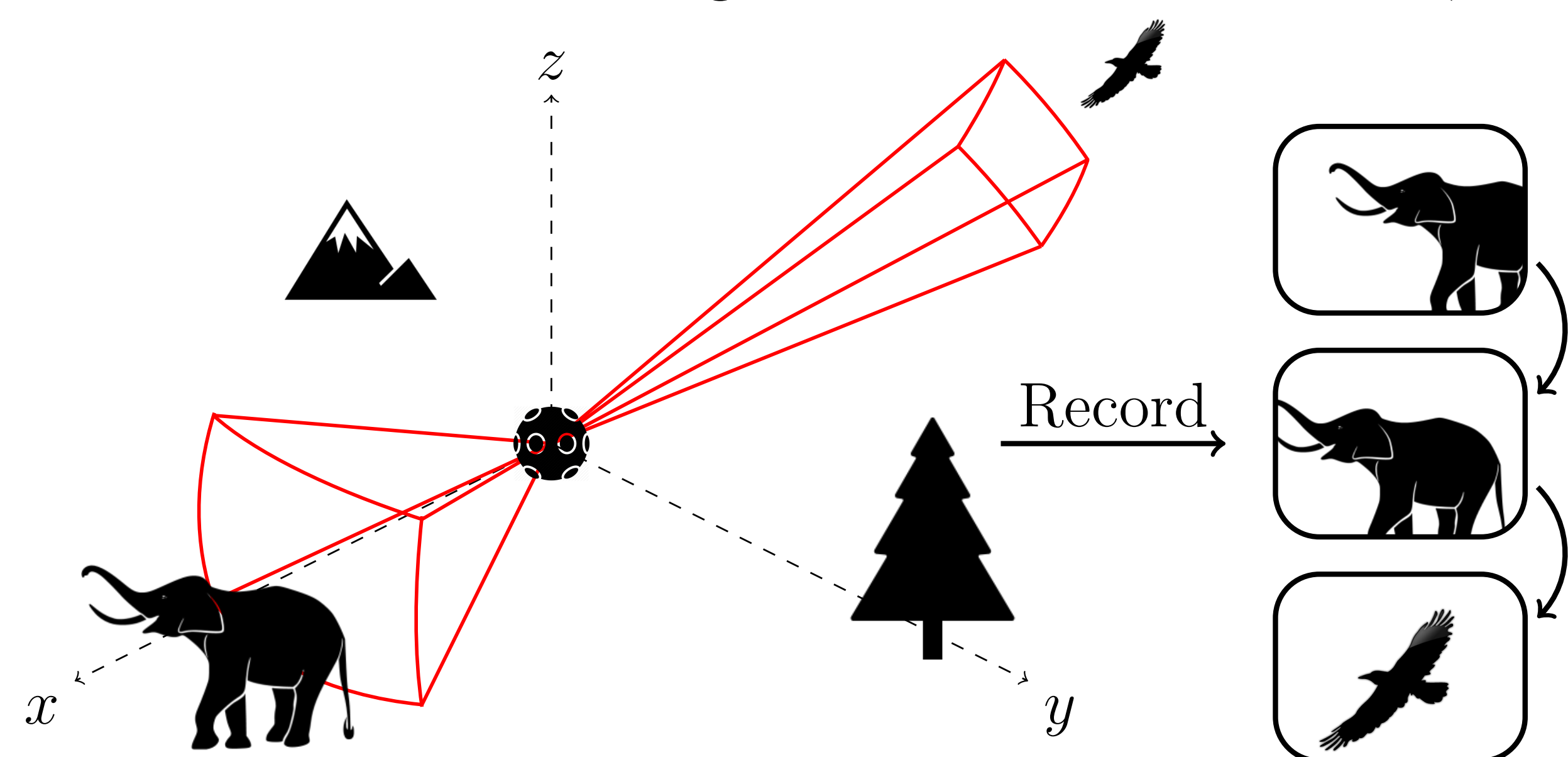


1. Pano2Vid Problem

Motivation: Help 360° video viewer determine where to look

Input: 360° video

Output: “Natural-looking” normal-field-of-view (NFOV) video

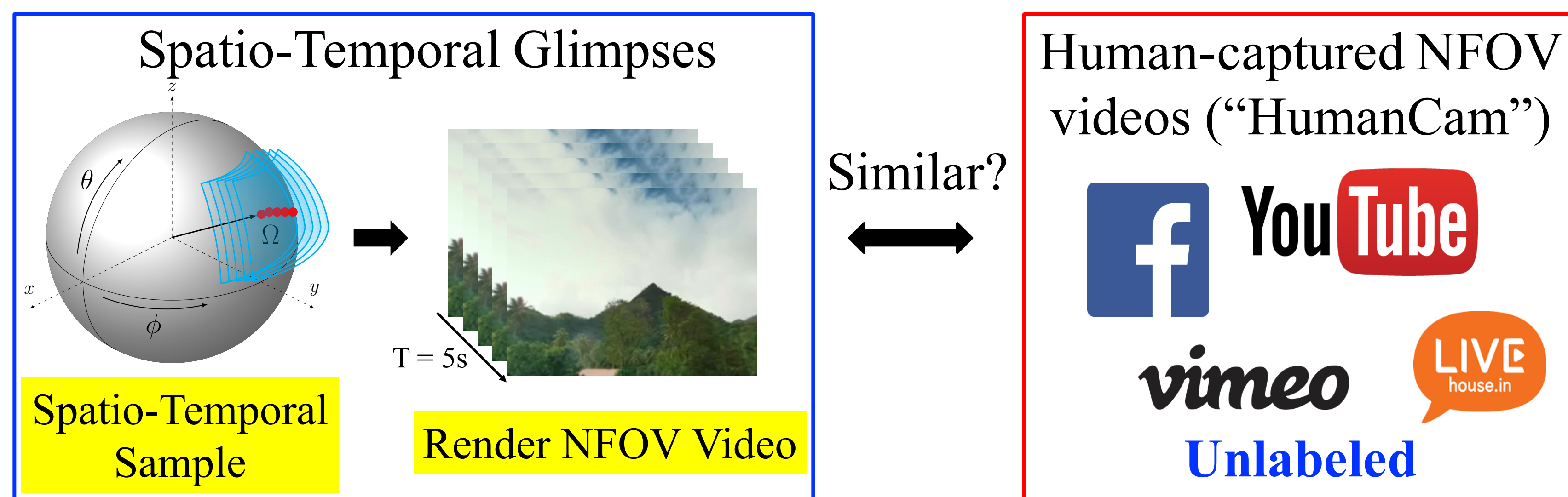


Task: Control the virtual camera direction & FOV

2. Our Approach: AutoCam

Idea: Learning *generic* virtual camera control from unlabeled human-captured NFOV video with *minimal* assumptions.

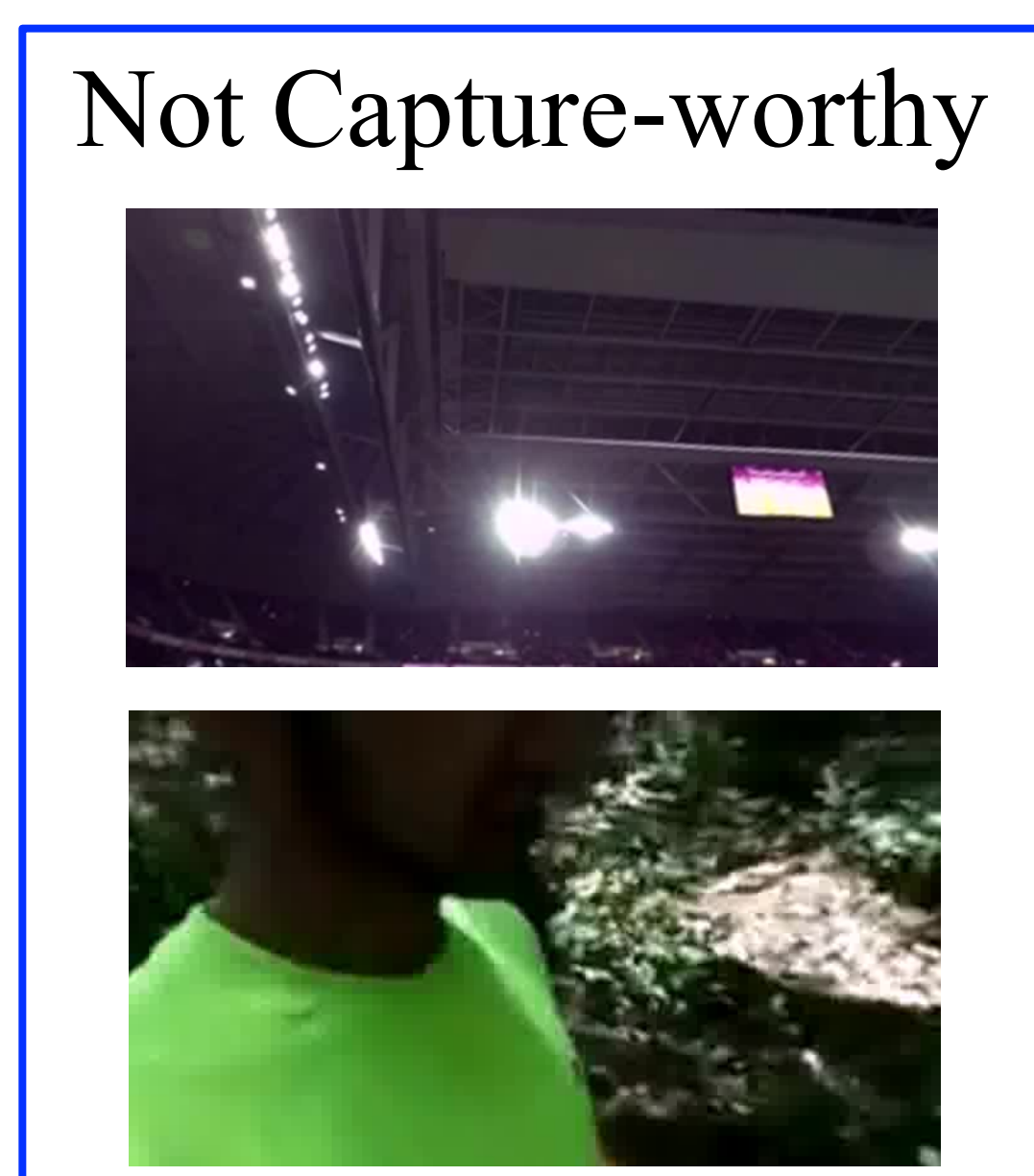
Learn Capture-worthiness (training)



Dataset

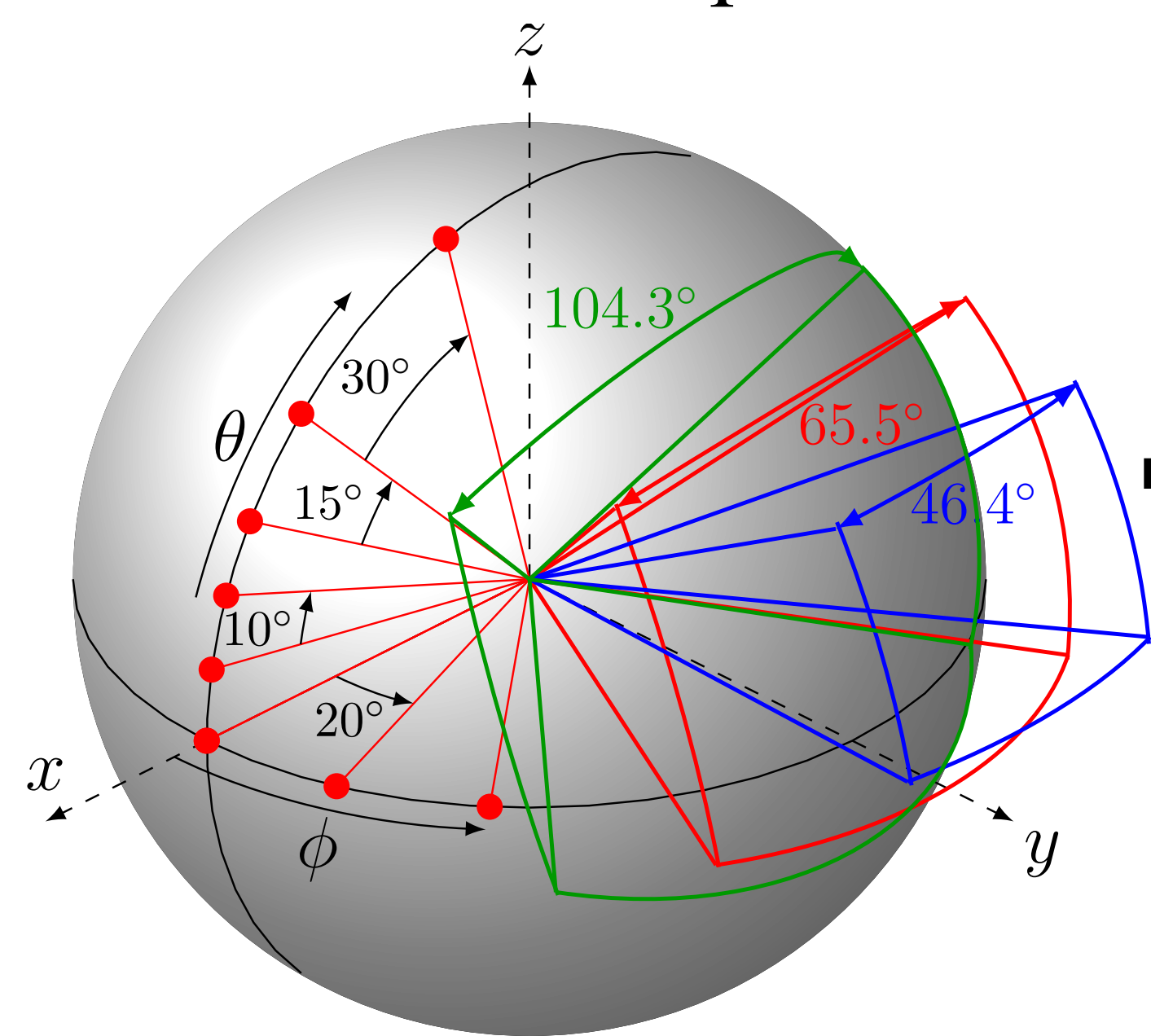
360° Videos	HumanCam	HumanEdit (for eval.)
86 videos / 7.3 hr	9,171 videos / 343 hr	480 trajectories / 12 hr

Example Glimpse Predictions

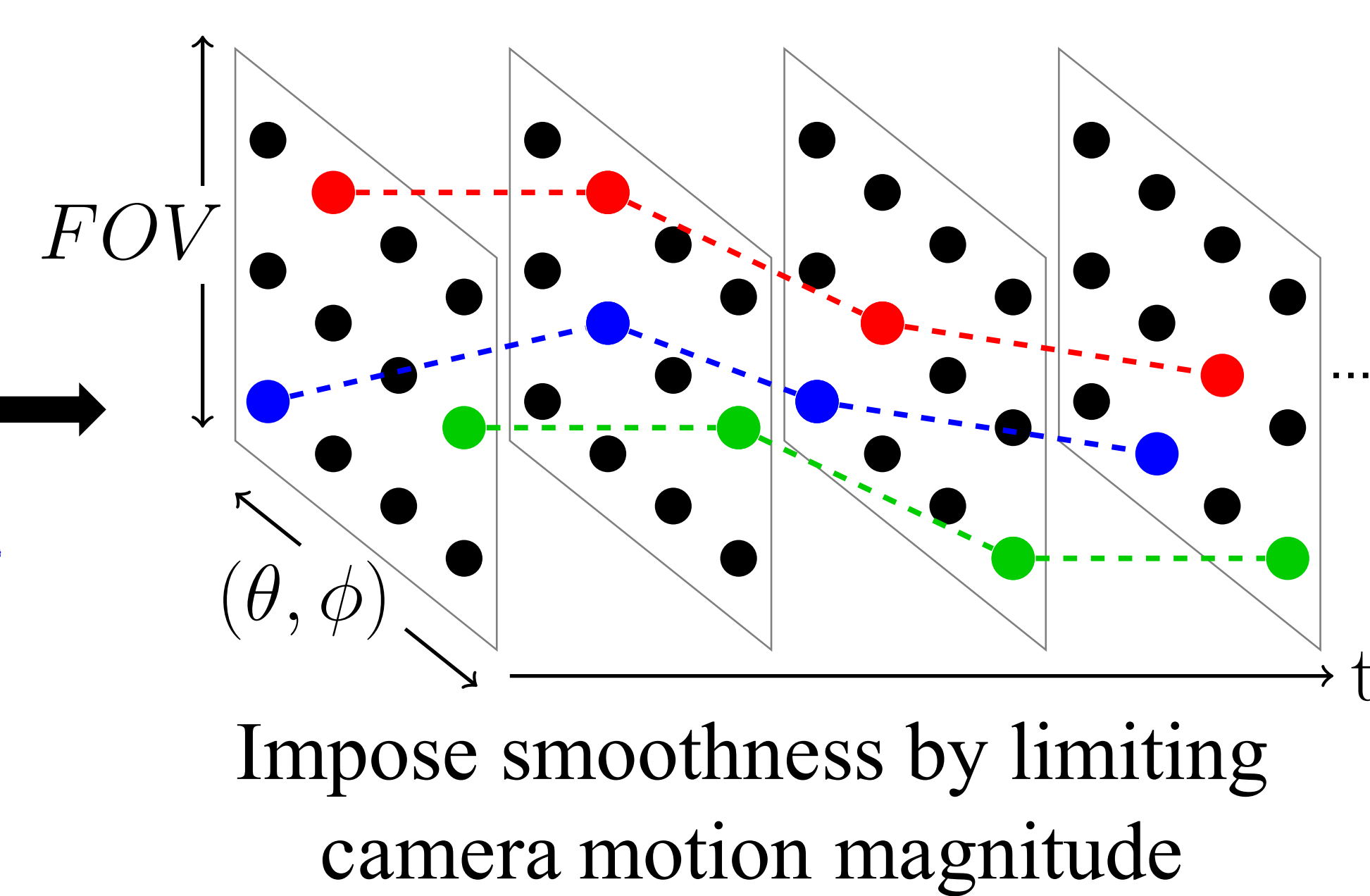


Construct Camera Trajectory (testing)

1. Densely Sample & Score ST-Glimpses



2. Find Smooth Trajectories with Maximum Accumulated Score



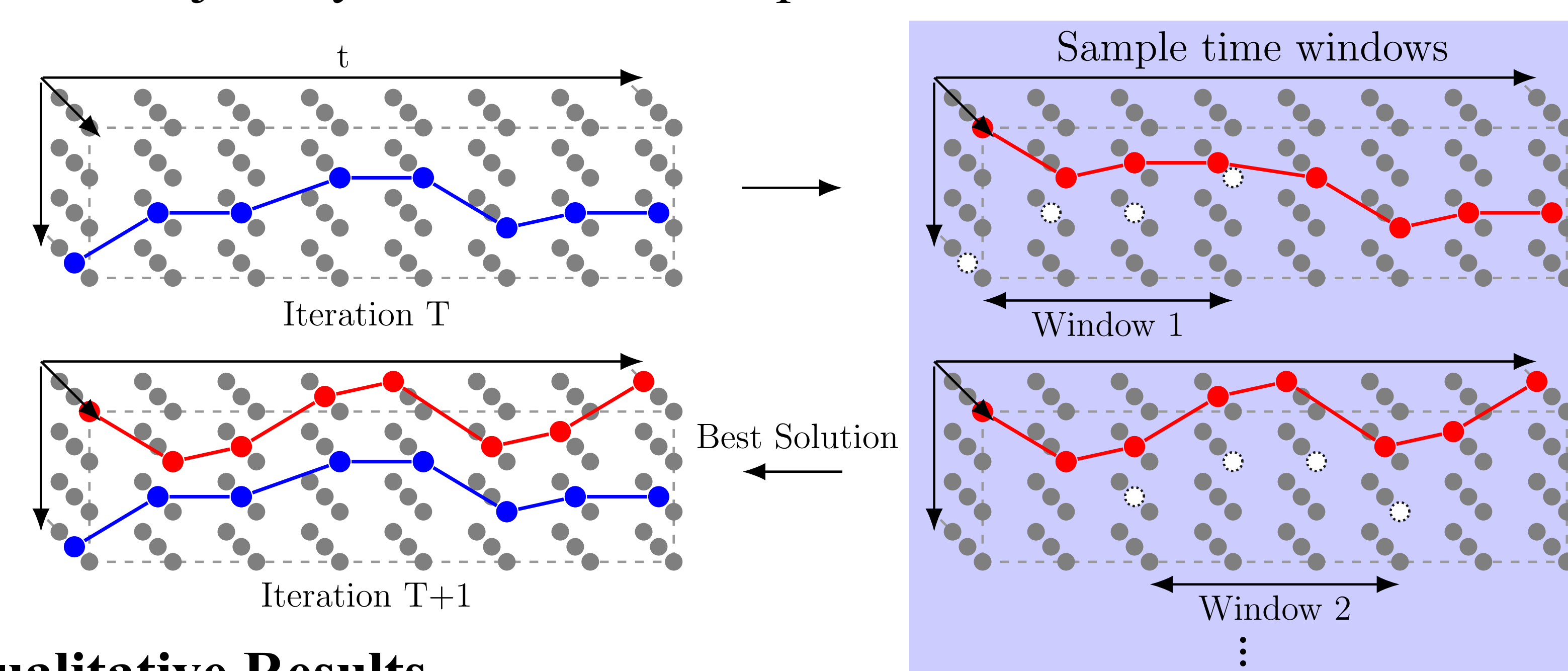
3. Generate Diverse Outputs for Each 360° Video

Multimodal Nature of Pano2Vid

- Multiple events / interesting objects in the same scene
- Personal preference of different viewers

Diverse Trajectory Search

- Generate trajectories iteratively
- New trajectory must differ from previous ones within a time window

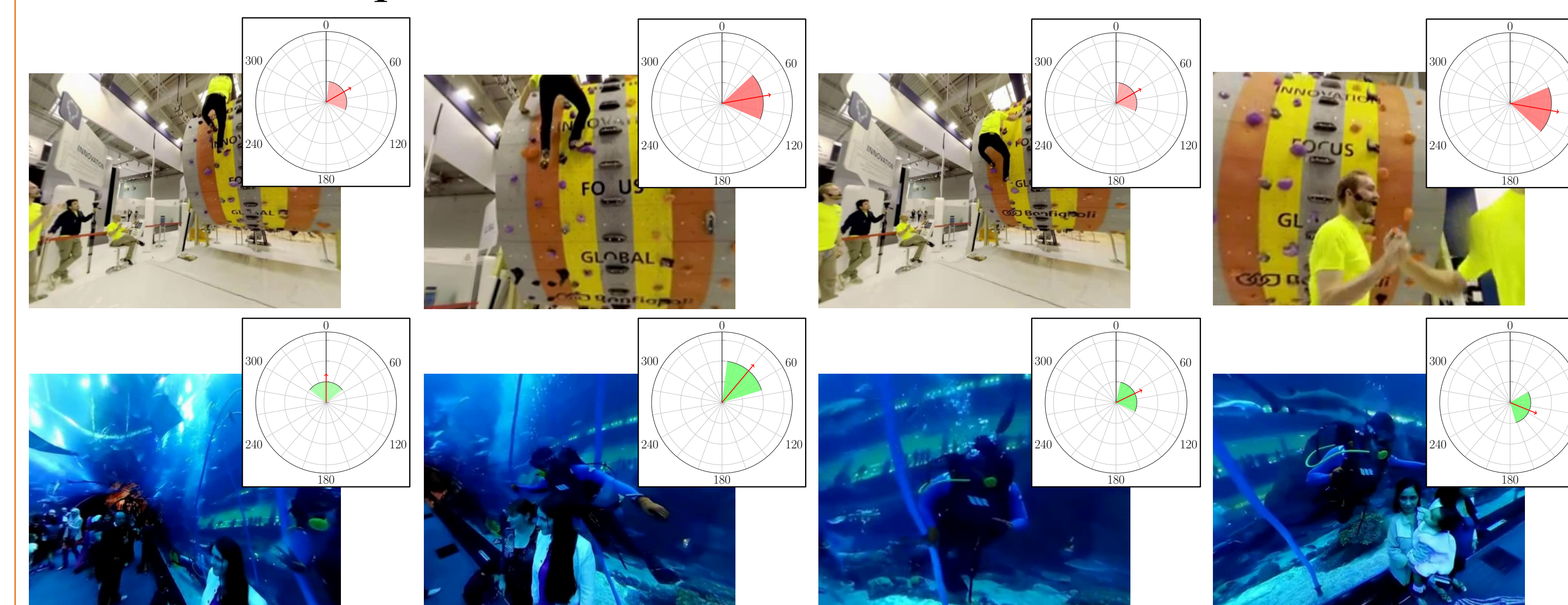


Qualitative Results



4. Enable Zooming in Virtual Camera Control

1. Better video presentation



2. Improve quantitative metrics by up to 43%

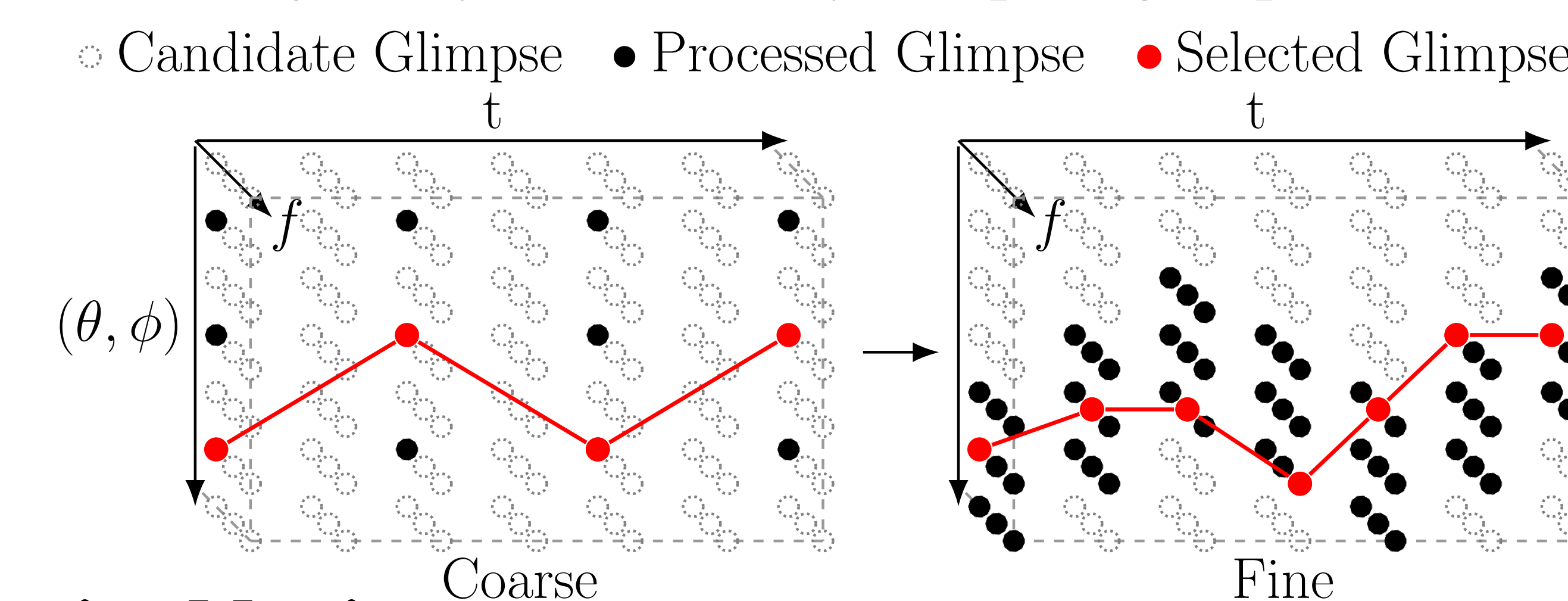
Problem – computational cost grows linearly with the zoom levels

5. Reduce Computational Cost

Bottleneck: Evaluating capture-worthiness for sampled glimpses

Coarse-to-fine Trajectory Search

- Avoid processing all candidate glimpses
- First construct trajectory over coarsely sampled glimpses
- Refine the trajectory over densely sampled glimpses



Evaluation Metrics

- HumanCam – does the video look human-captured?
- HumanEdit – are the algorithm choices similar to human editors’?

Quantitative Results

