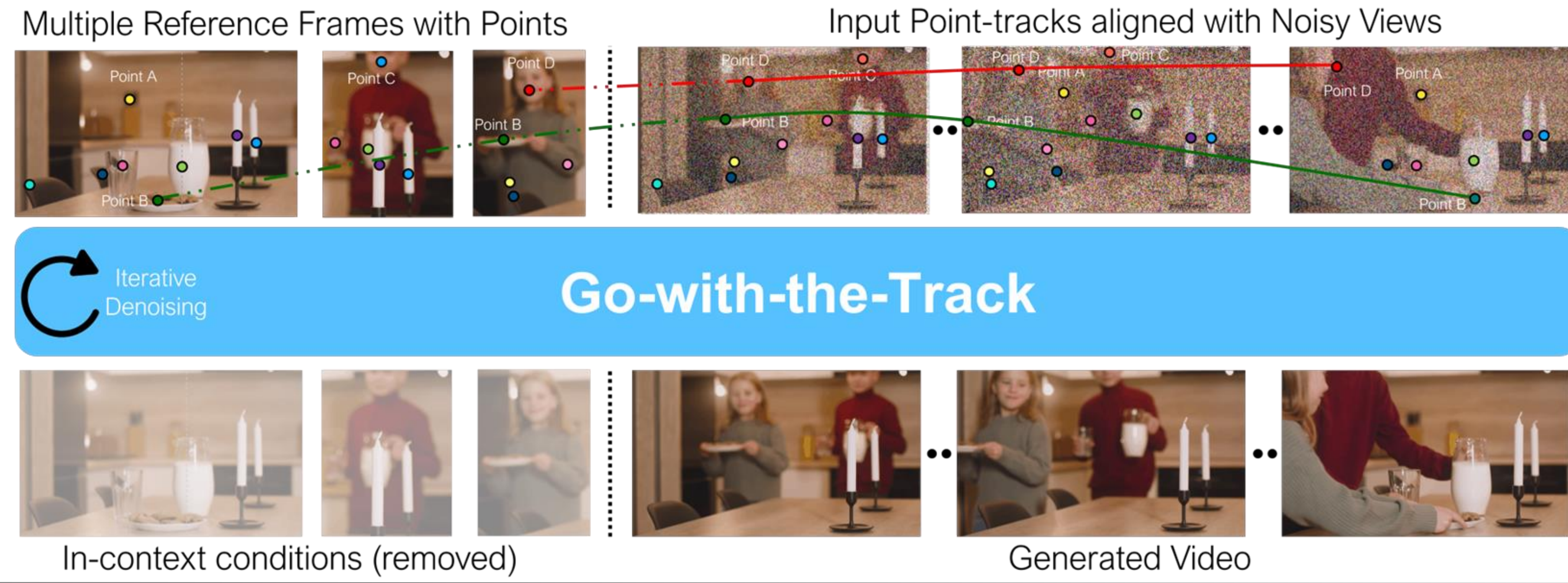


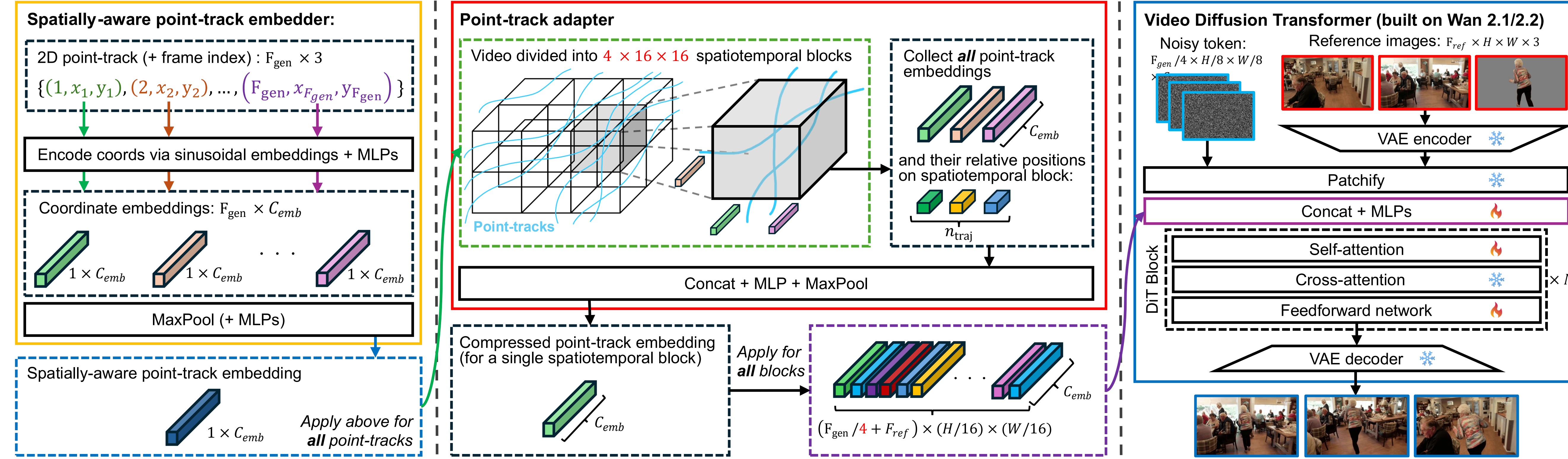


Overview

➤ Multiple reference images + Point-tracks → Composited video

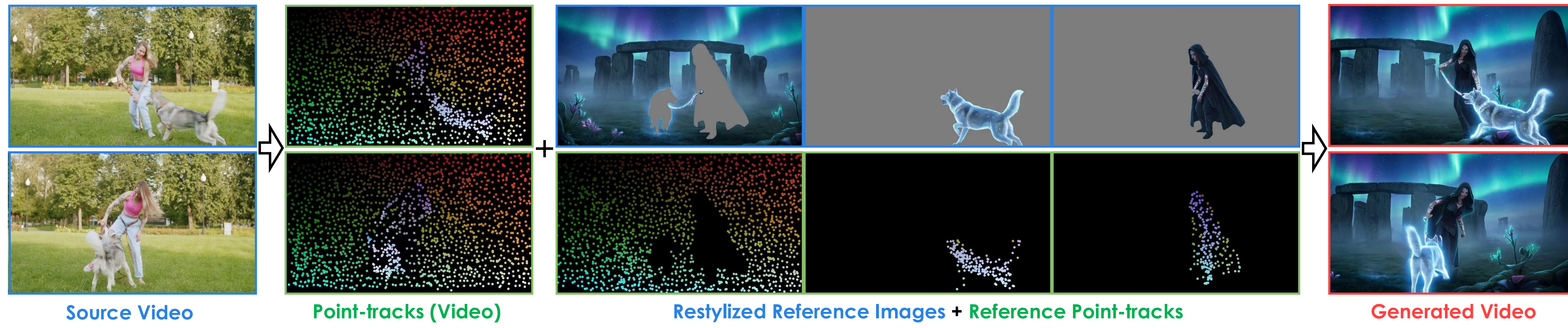


Pipeline

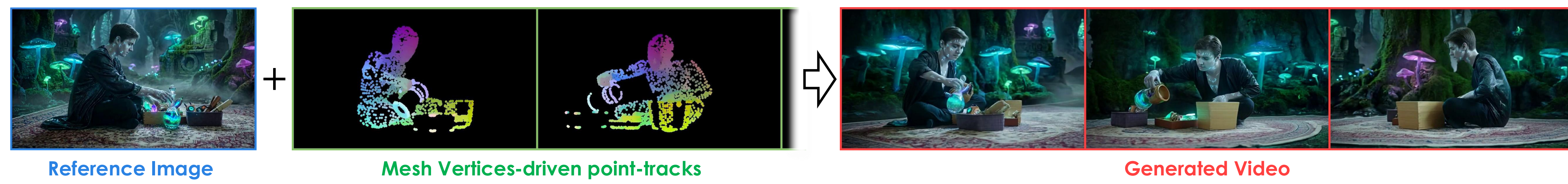


Qualitative Results

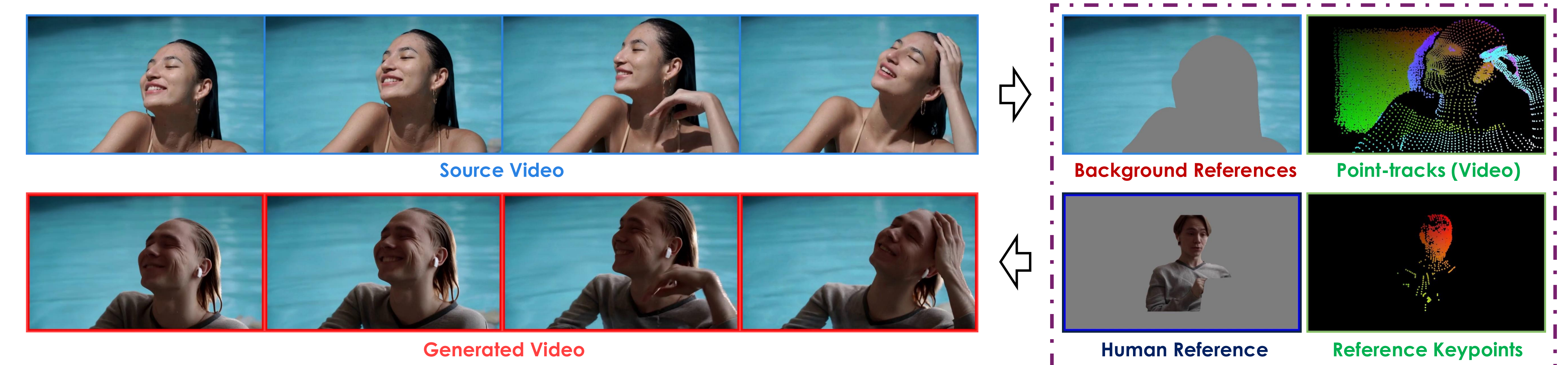
Multi-reference-driven Video Restylization



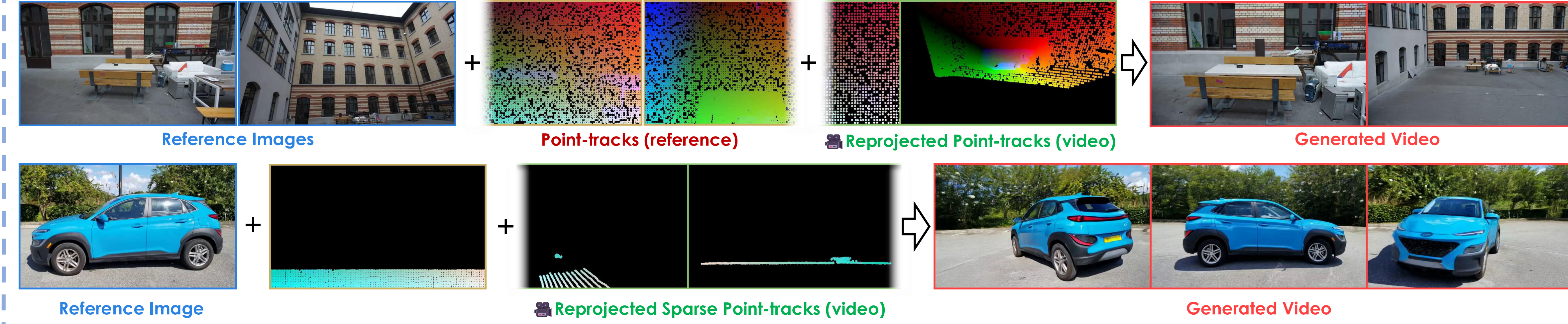
Mesh-driven Compositing



Keypoint-driven Compositing



Camera Motion Control in Static Scene

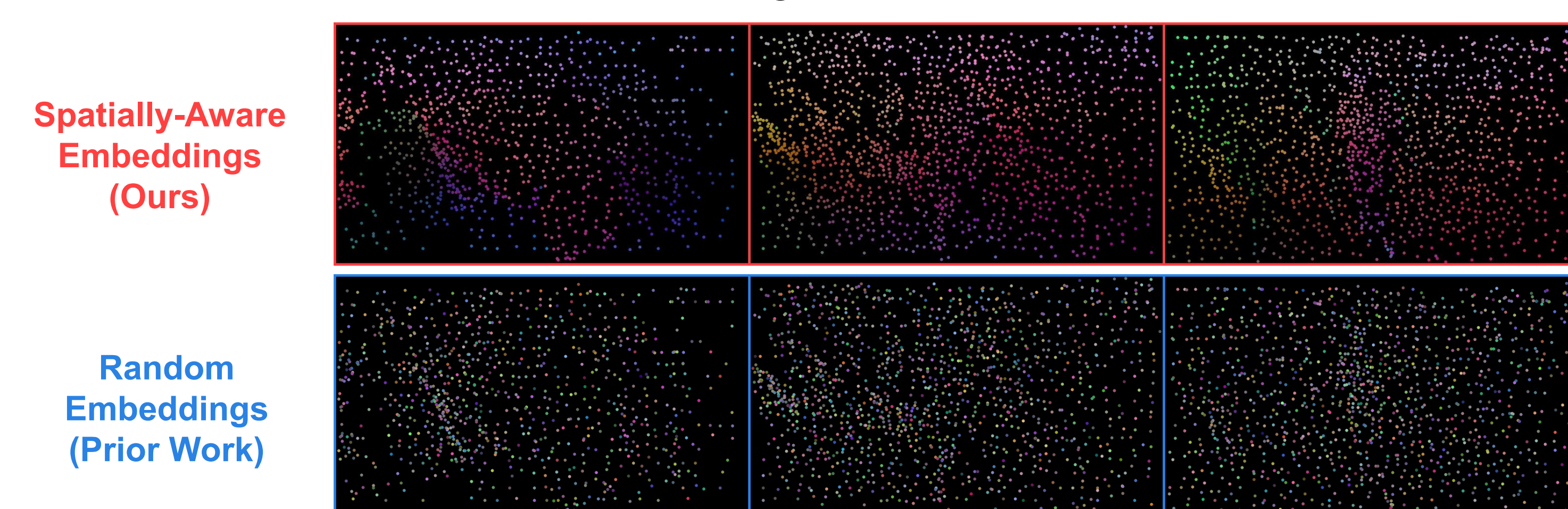


Camera Retargeting in Dynamic Scene

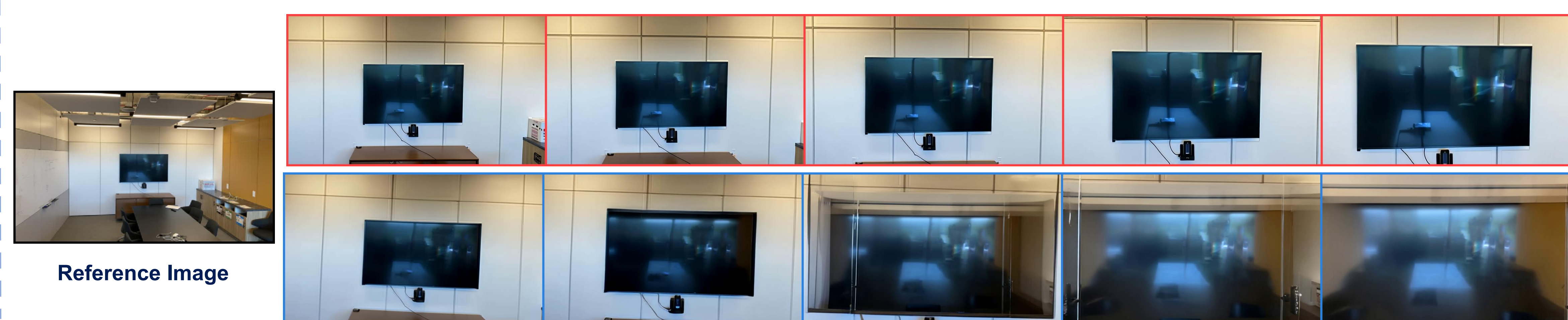


Ablation Study: Point-track Embeddings

Embedding Visualizations



Generated Video Frames



Our spatially-aware embeddings:

- Encodes the full sequence of point-track coordinates into an embedding.
- Ensures spatially close point-tracks exhibit similar embedding representations.
- Facilitates the model to associate points between the reference and generated frames despite being spatiotemporally discontinuous.