



¹Snap Inc.



²UMass Amherst



³TU Crete

DELTA: Dense Efficient Long-range 3D Tracking for Any video

Tuan Duc Ngo^{1,2}, Peiye Zhuang¹, Chuang Gan², Evangelos Kalogerakis^{2,3},
Sergey Tulyakov¹, Hsin-Ying Lee¹, Chaoyang Wang¹



ICLR

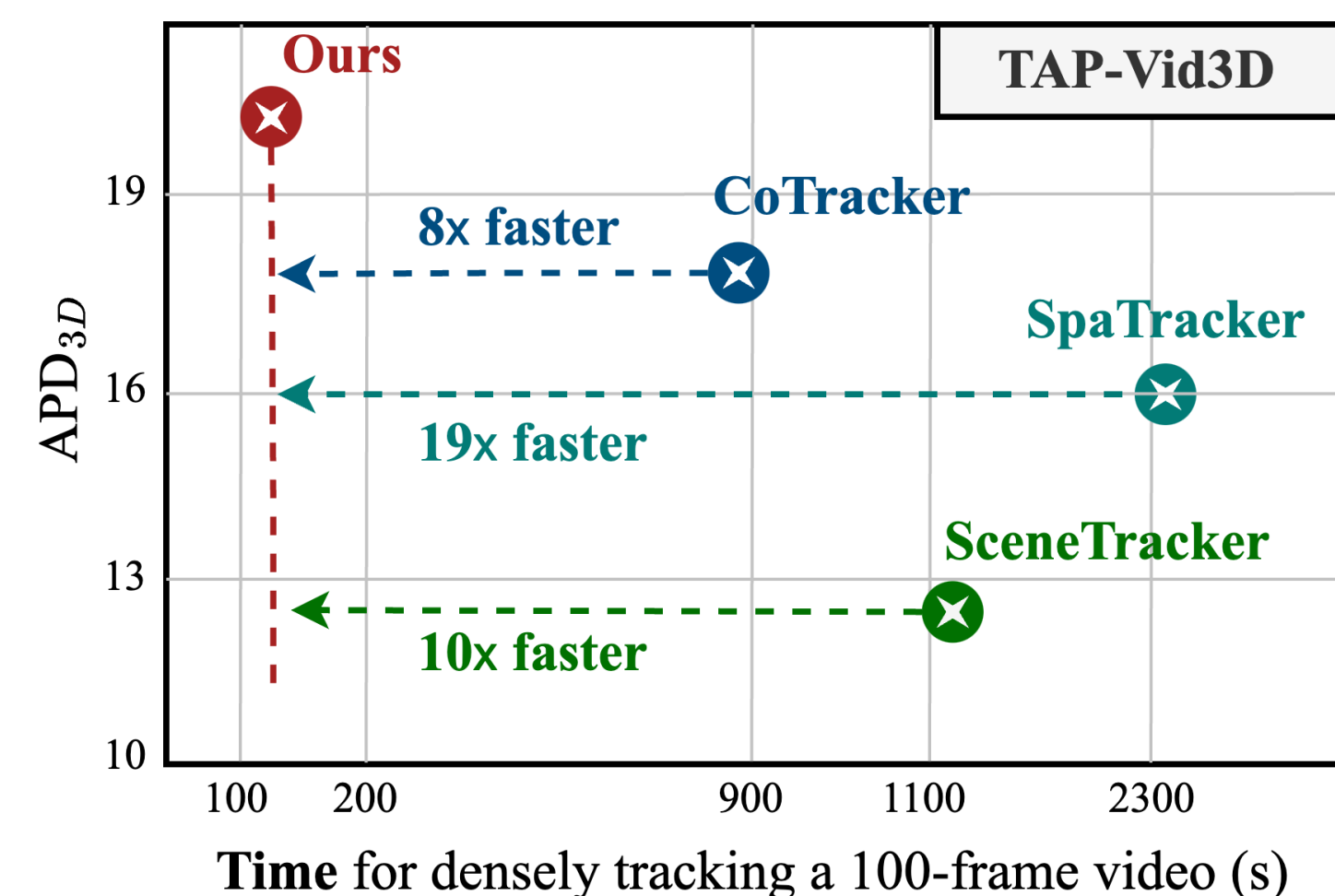


Webpage

Overview

DELTA is a dense 3D tracking approach that:

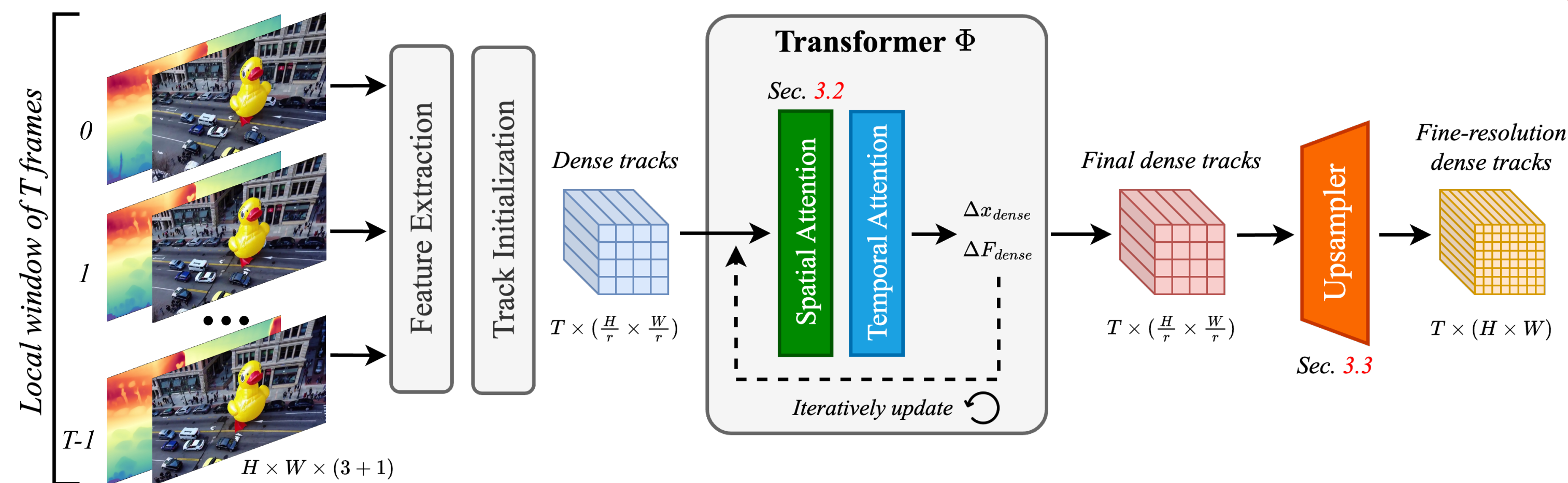
- Efficiently track **all pixels** from a long monocular video
- Obtain **consistent trajectories** in **3D space**
- Achieve **SoTA results** on 3D/2D tracking, long-range optical flow



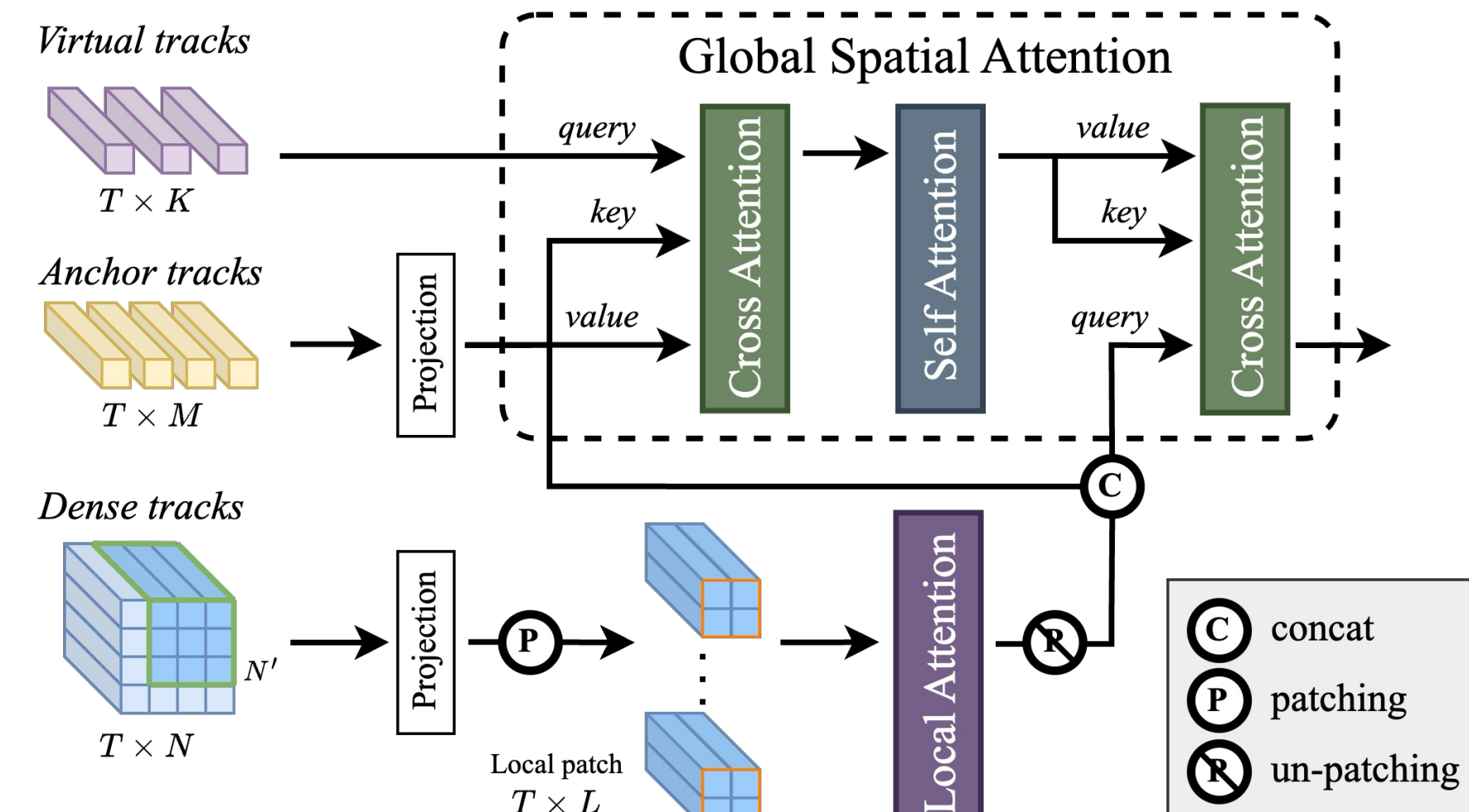
Method

A **coarse-to-fine** strategy:

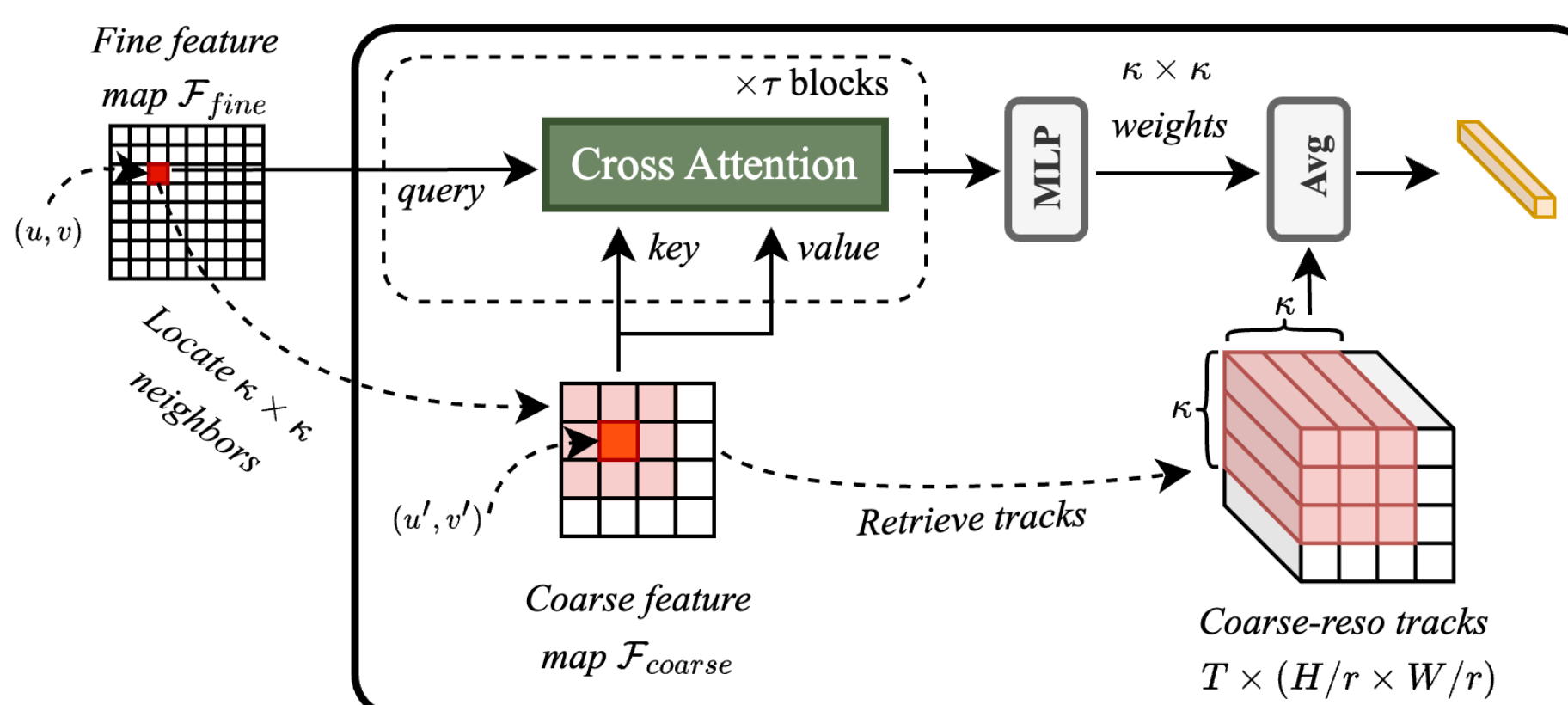
- Track dense pixels in reduced resolution
- Use a lightweight upsampler to obtain high-resolution dense track



Global-Lobal Spatial Attention



Attention-based Upsampler



Quantitative Results

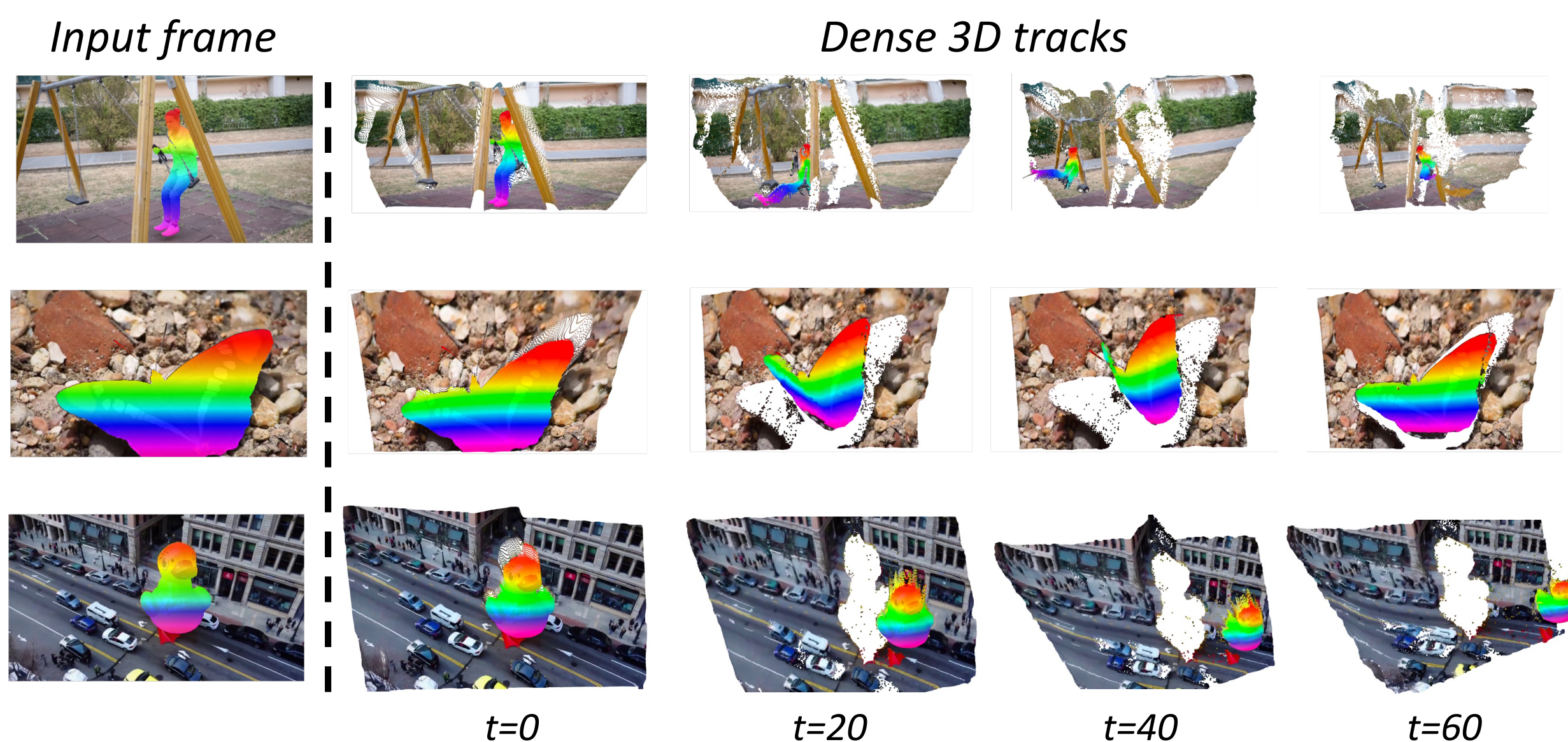
3D Point Tracking (TAP-Vid3D dataset)

Methods	Aria			DriveTrack			PStudio		
	AJ↑	APD _{3D} ↑	OA↑	AJ↑	APD _{3D} ↑	OA↑	AJ↑	APD _{3D} ↑	OA↑
TAPIR [†] + COLMAP	7.1	11.9	72.6	8.9	14.7	80.4	6.1	10.7	75.2
CoTracker [†] + COLMAP	8.0	12.3	78.6	11.7	19.1	81.7	8.1	13.5	77.2
BootsTAPIR [†] + COLMAP	9.1	14.5	78.6	11.8	18.6	83.8	6.9	11.6	81.8
CoTracker [†] + UniDepth	13.0	20.9	84.9	12.5	19.9	80.1	6.2	13.5	67.8
TAPTR [†] + UniDepth	15.7	24.2	87.8	12.4	19.1	84.8	7.3	13.5	84.3
LocoTrack [†] + UniDepth	15.1	24.0	83.5	13.0	19.8	82.8	7.2	13.1	80.1
SpatialTracker + UniDepth	13.6	20.9	90.5	8.3	14.5	82.8	8.0	15.0	75.8
SceneTracker + UniDepth	-	23.1	-	-	6.8	-	-	12.7	-
DOT-3D + UniDepth	13.8	22.1	85.5	11.8	17.9	82.3	3.2	5.3	52.5
Ours + UniDepth	16.6	24.4	86.8	14.6	22.5	85.8	8.2	15.0	76.4

Long-range 2D Optical Flow (CVO dataset)

Methods	CVO-Clean(7 frames)			CVO-Extended(48 frames)		
	EPE↓ (all/vis/occ)	IoU↑		EPE↓ (all/vis/occ)	IoU↑	
RAFT	2.48 / 1.40 / 7.42	57.6		21.80 / 15.4 / 33.4	65.0	
MFT	2.91 / 1.39 / 9.93	19.4		21.40 / 9.20 / 41.8	37.6	
TAPIR	3.80 / 1.49 / 14.7	73.5		19.8 / 4.74 / 42.5	68.4	
CoTracker	1.51 / 0.88 / 4.57	75.5		5.20 / 3.84 / 7.70	70.4	
DOT	1.29 / 0.72 / 4.03	80.4		4.98 / 3.59 / 7.17	71.1	
SceneTracker	4.40 / 3.44 / 9.47	-		11.5 / 8.49 / 17.0	-	
SpatialTracker	1.84 / 1.32 / 4.72	68.5		5.53 / 4.18 / 8.68	66.6	
DOT-3D	1.33 / 0.75 / 4.16	79.0		5.20 / 3.58 / 7.95	70.9	
Ours (2D)	0.89 / 0.46 / 2.96	78.3		3.63 / 2.67 / 5.24	71.6	
Ours (3D)	0.94 / 0.51 / 2.97	78.7		3.67 / 2.64 / 5.30	70.1	

Qualitative Results



Application: Non-rigid SfM

