

TimeTracker: Event-based Continuous Point Tracking for Video Frame Interpolation with Non-linear Motion

Haoyue Liu¹, Jinghan Xu¹, Yi Chang^{1*}, Hanyu Zhou¹, Haozhi Zhao¹, Lin Wang², Luxin Yan¹
¹Huazhong University of Science and Technology ²Nanyang Technological University

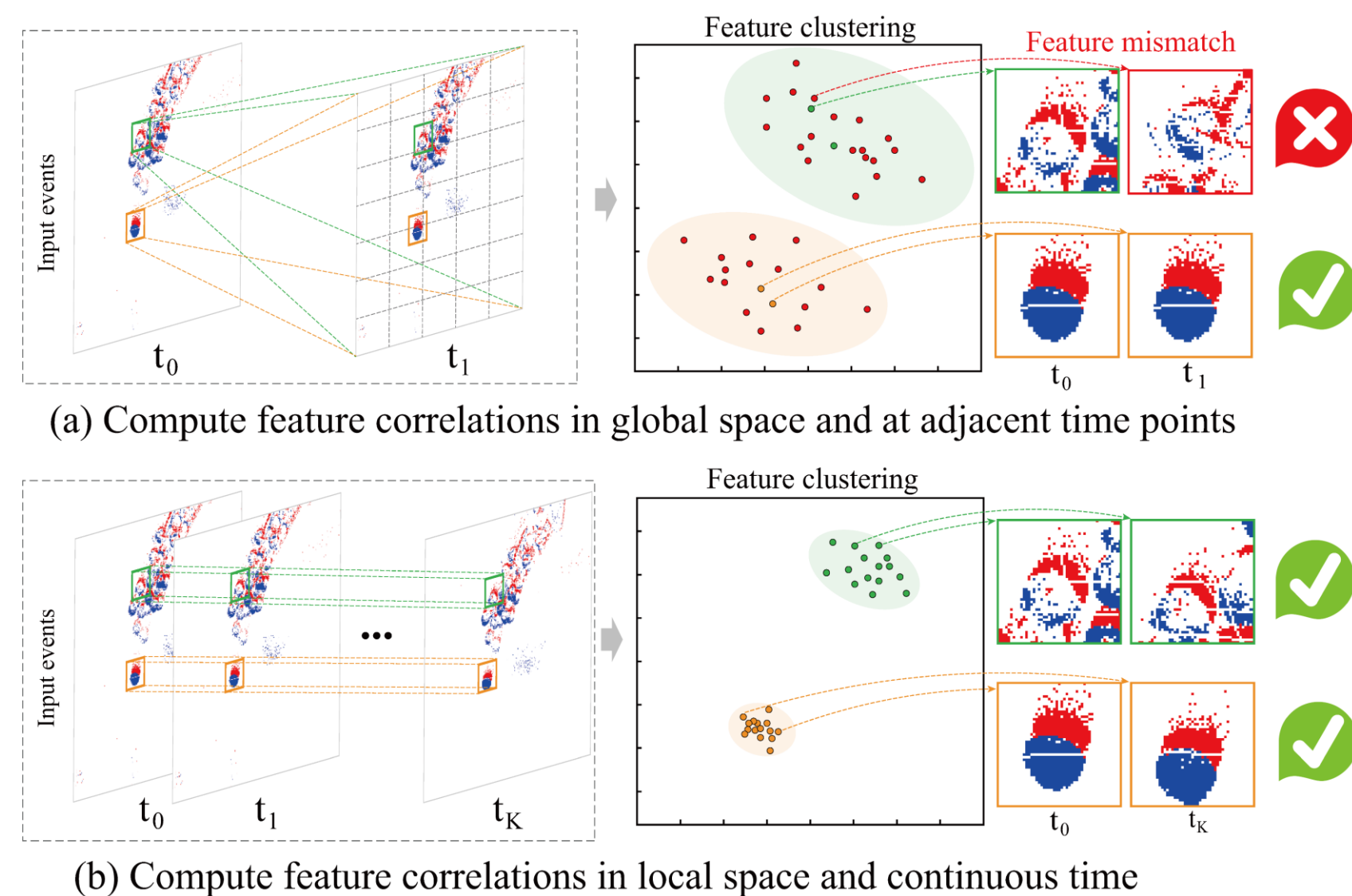
Introduction

Problem:

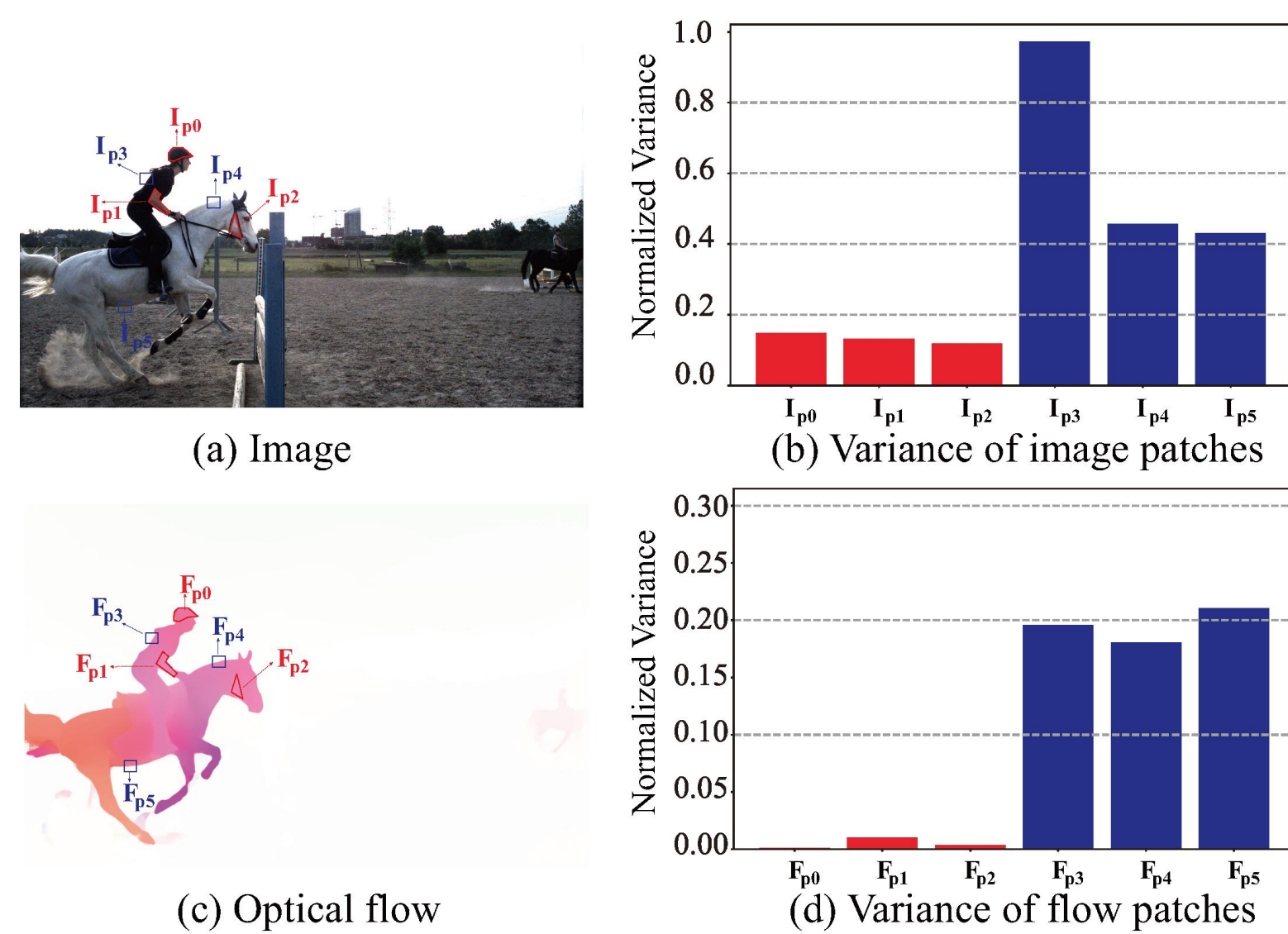
- The spatial distribution of events is sparse, making it difficult to provide accurate, dense global optical flow for the VFI task.

Motivation:

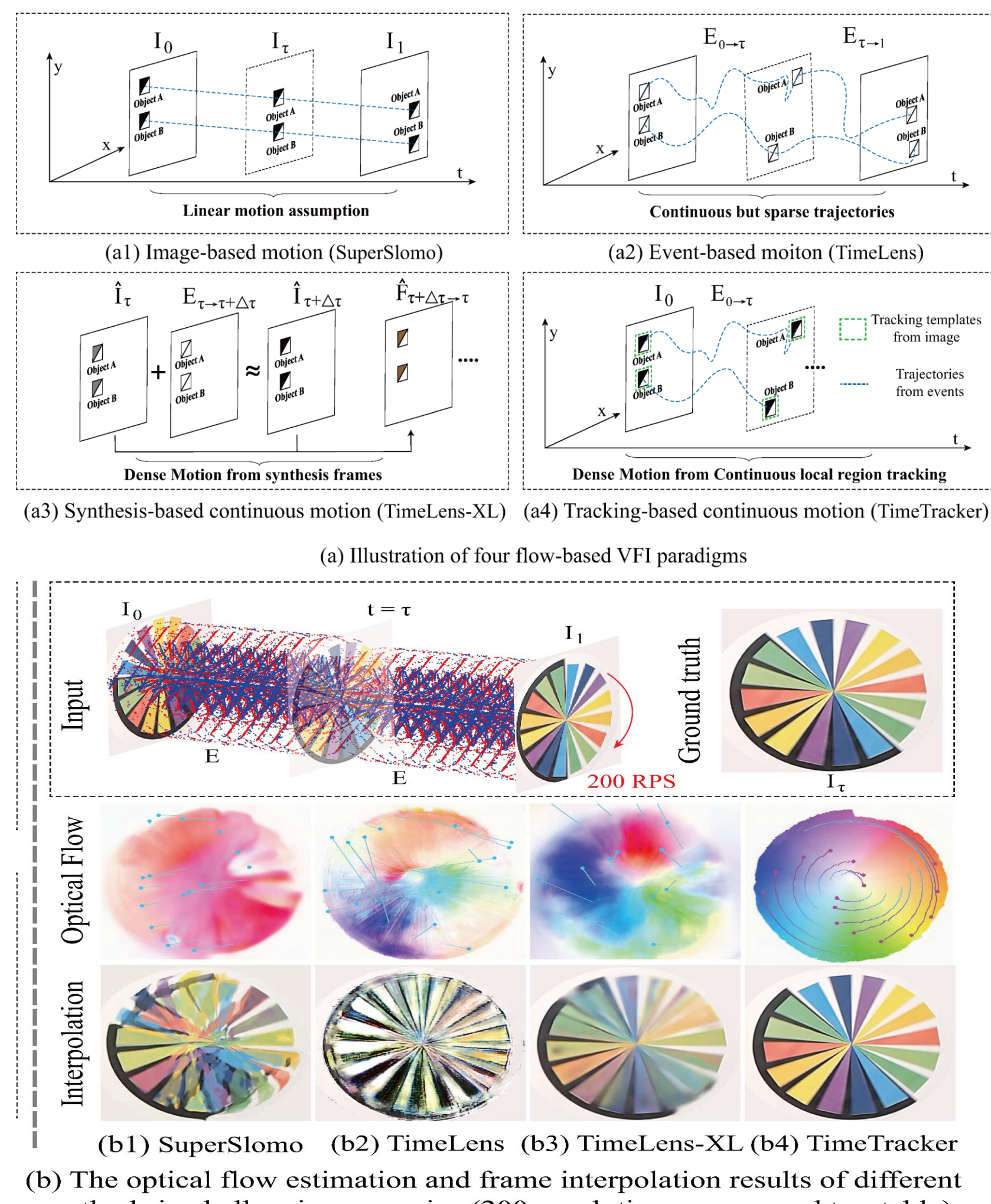
- Event data is spatially sparse but temporally dense, making it easier to find feature correspondences at a local spatial and continuous temporal scale.



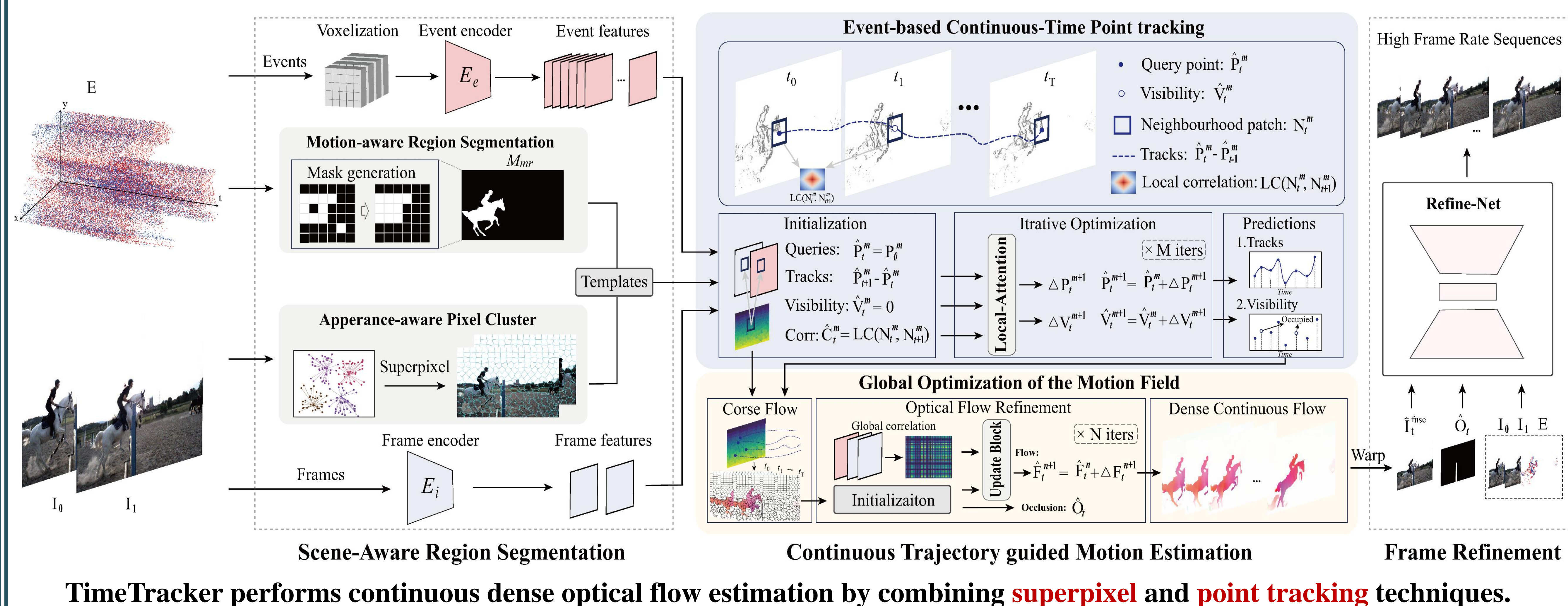
- Object boundaries are highly correlated with motion boundaries.



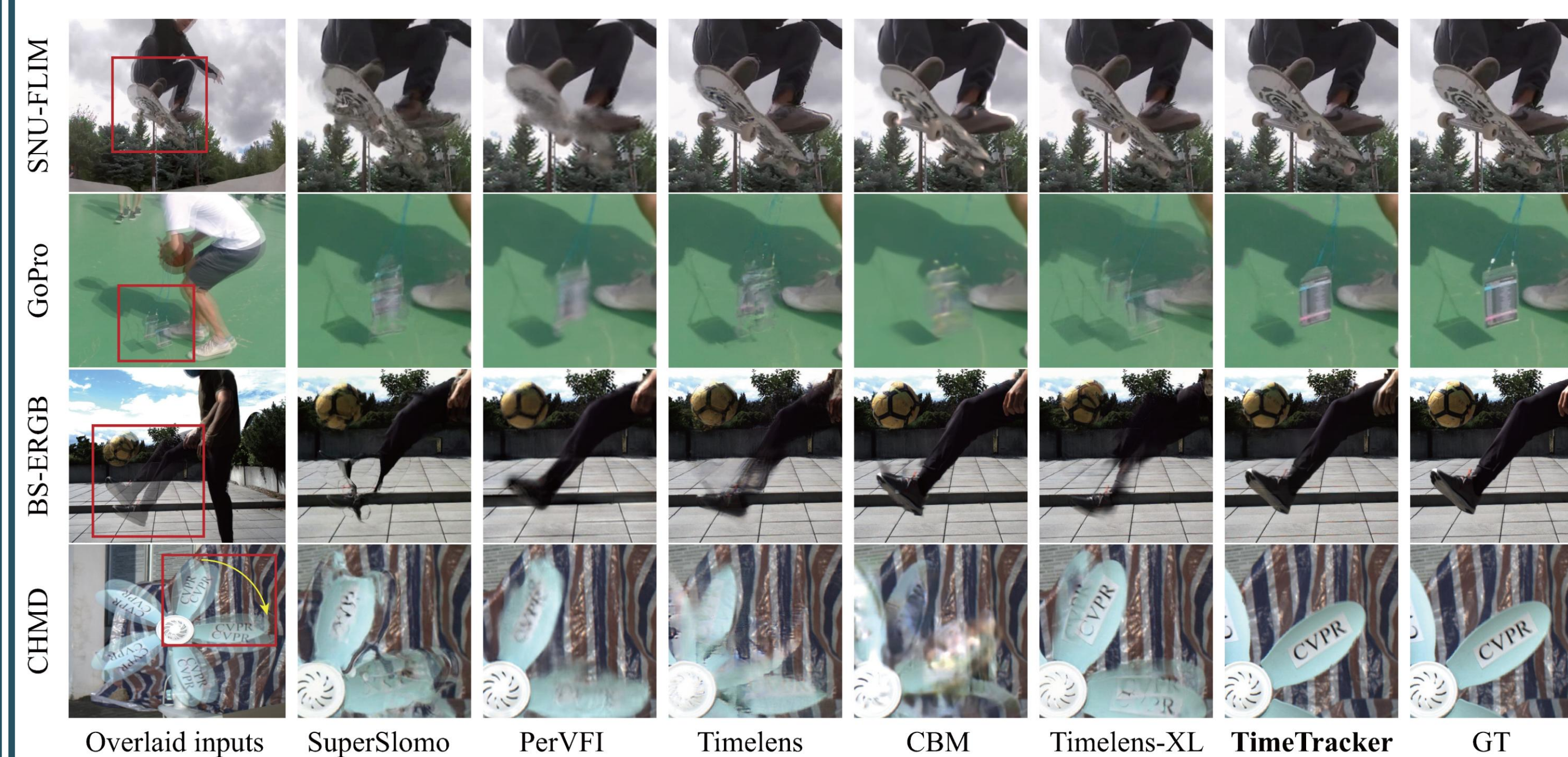
Flow-based VFI Comparison



Overall Architecture of TimeTracker



Qualitative and Quantitative Results



dataset	Gopro				SNU-FILM			
	7skips		15skips		Hard		Extreme	
SuperSloMo[3]	28.28	0.902	23.31	0.776	24.71	0.846	21.73	0.794
VFT-B[57]	30.80	0.912	26.10	0.836	26.34	0.883	25.49	0.852
PerVFI[56]	31.86	0.933	27.46	0.845	29.77	0.913	27.84	0.891
TimeLens[12]	34.42	0.948	33.31	0.928	31.45	0.928	28.73	0.897
SuperFast[17]	33.76	0.943	32.97	0.927	28.74	0.903	26.37	0.863
CBMNet[23]	36.86	0.955	35.32	0.947	30.87	0.918	27.56	0.884
TimeLens-XL[13]	37.02	0.959	36.19	0.949	30.95	0.920	27.93	0.894
TimeTracker (ours)	37.13	0.962	36.54	0.958	32.86	0.935	29.27	0.915

Results on synthetic datasets

dataset	BS-ERGB				Ours			
	1skip		3skips		7skips		15skips	
SuperSloMo[3]	23.33	0.734	22.43	0.716	21.17	0.705	20.26	0.674
VFT-B[57]	24.44	0.741	24.31	0.725	22.04	0.743	21.70	0.682
PerVFI[56]	27.72	0.761	26.07	0.763	24.82	0.768	21.65	0.702
TimeLens[12]	28.13	0.787	26.82	0.769	25.86	0.771	24.12	0.748
SuperFast[17]	27.87	0.768	26.77	0.758	22.79	0.762	20.59	0.722
CBMNet[23]	29.03	0.807	28.10	0.794	26.27	0.792	25.38	0.766
TimeLens-XL[13]	29.35	0.813	28.69	0.802	26.13	0.785	24.77	0.732
TimeTracker (ours)	29.85	0.823	29.14	0.807	28.45	0.814	27.69	0.805

Results on real-world datasets

Ablations and Discussion

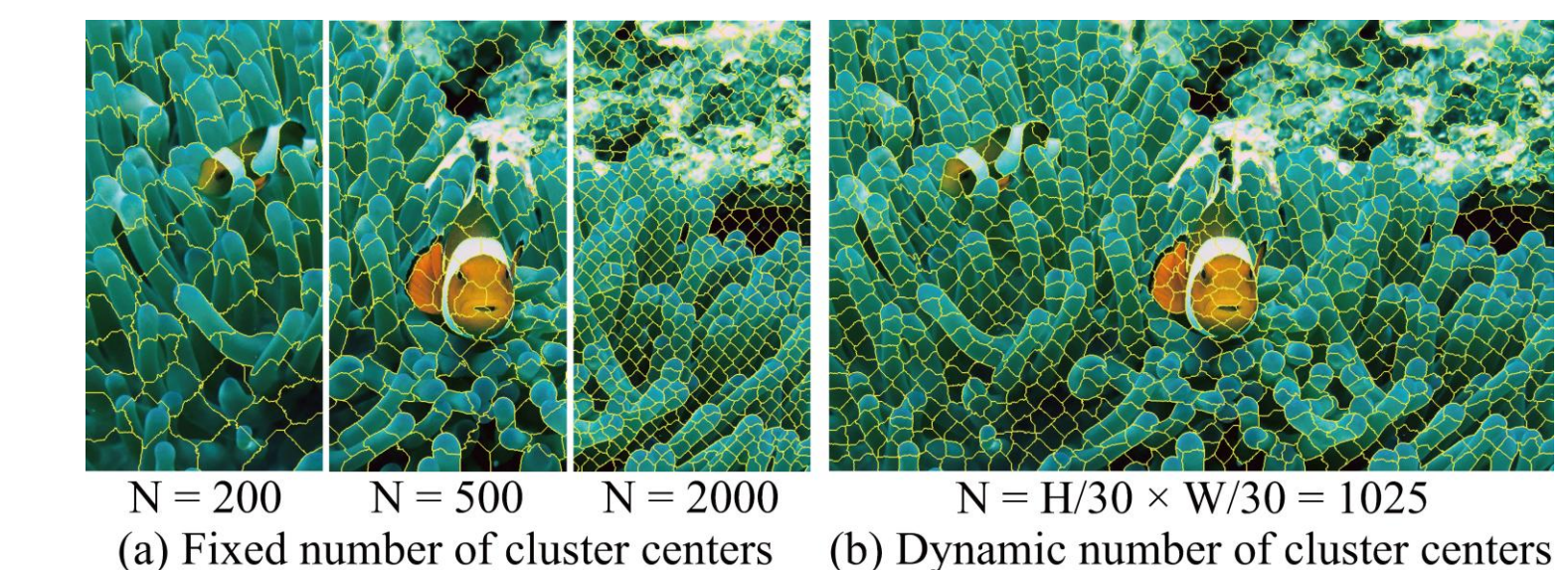
Ablation on Main Components

	w/o superpixel	w/o point tracking	w/o global optimization	w/o frame refinement	PSNR ↑	SSIM ↑
Superpixel	✓	✓	✓	✓	30.62	0.925
Point tracking	✓	✓	✓	✓	27.15	0.885
Global optimization	✓	✓	✓	✓	32.43	0.933
Frame refinement	✓	✓	✓	✓	35.29	0.942
TimeTracker	✓	✓	✓	✓	37.47	0.965

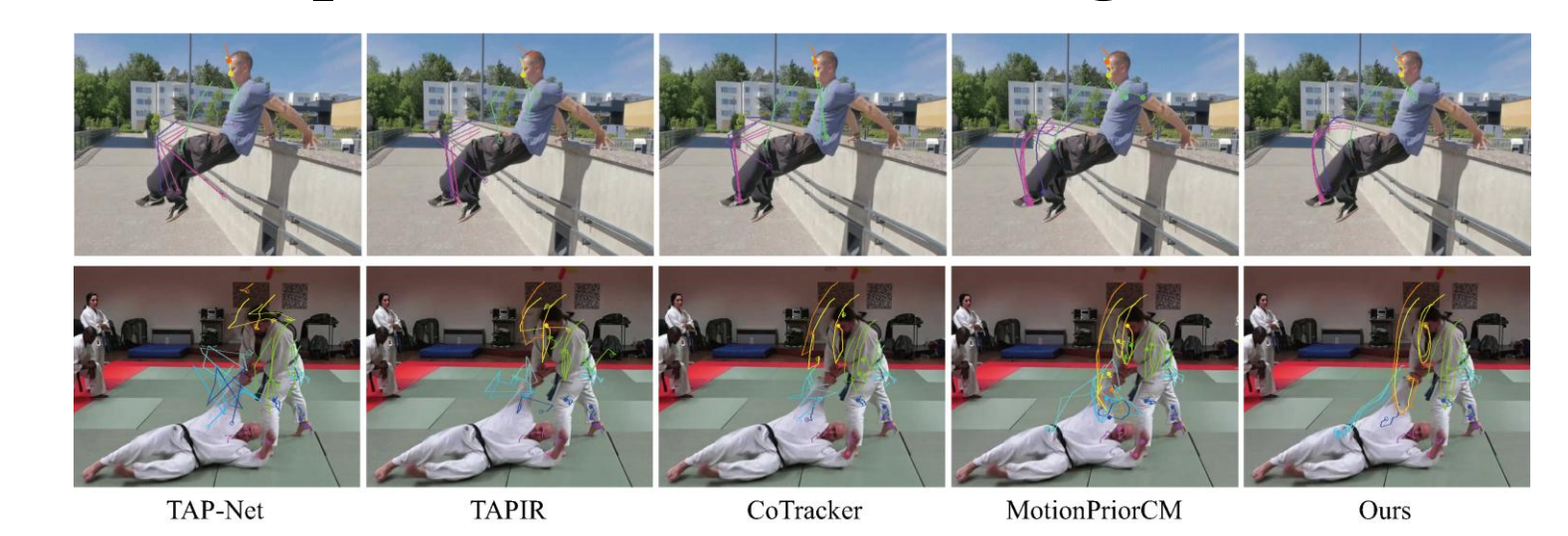
Comparison of Optical Flow Module

Methods	PerVFI	TimeLens	CBMNet	TimeLens-XL	TimeTracker
Data source	Image (I)	Event (E)	I+E	I+E	I+E
PSNR ↑	30.25	30.84	32.19	33.46	35.29
SSIM ↑	0.908	0.916	0.925	0.939	0.942

Number of SLIC Segmented Regions



Comparison of Point Tracking Module



Each input skip two frames to simulate high-speed motion