Data-driven Feature Tracking for Event Cameras

Nico Messikommer*, Carter Fang*, Mathias Gehrig, Davide Scaramuzza

TUE-PM-144

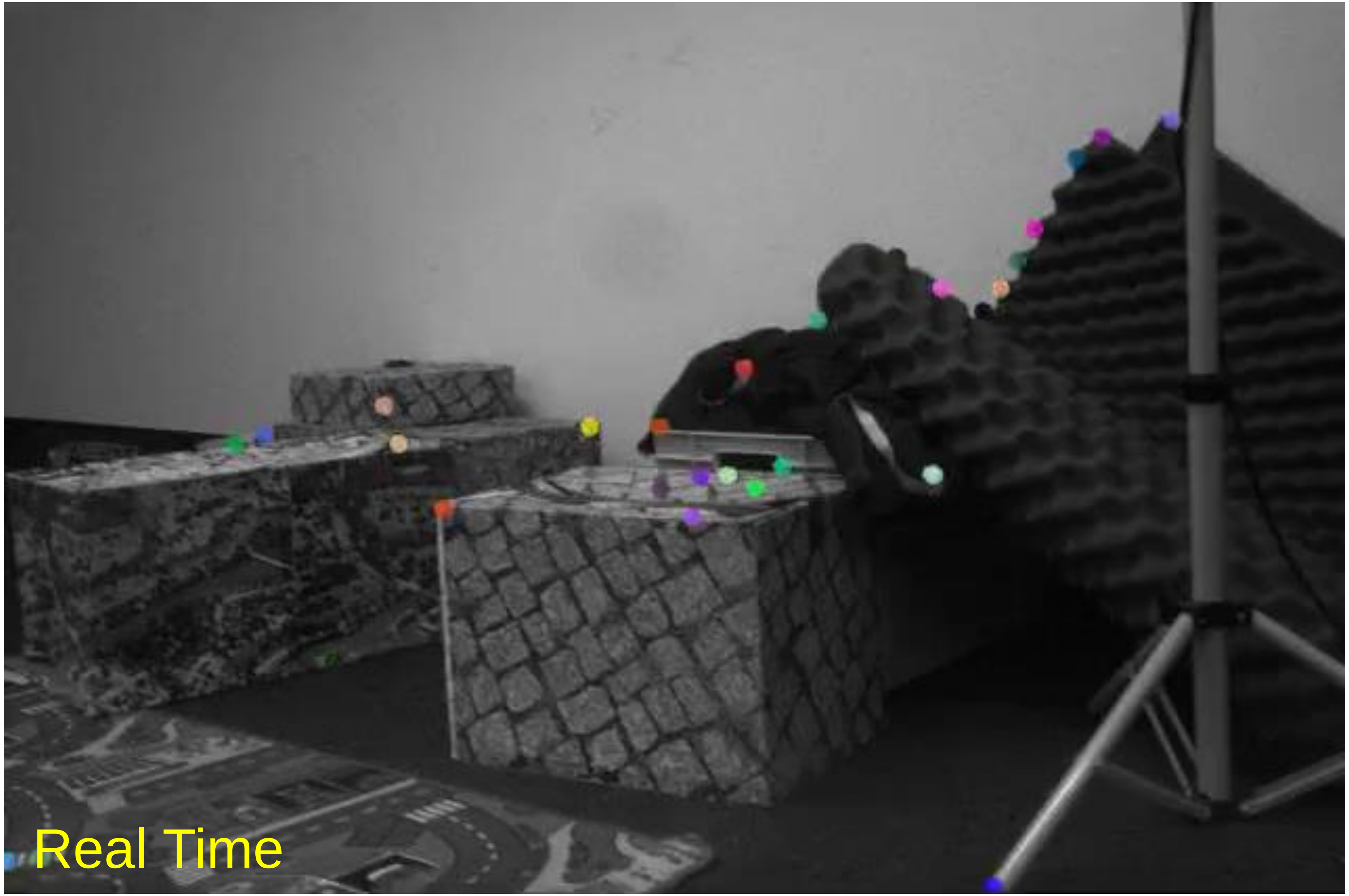
Source Code: https://github.com/uzh-rpg/deep_ev_tracker

University of Zurich[UZH]

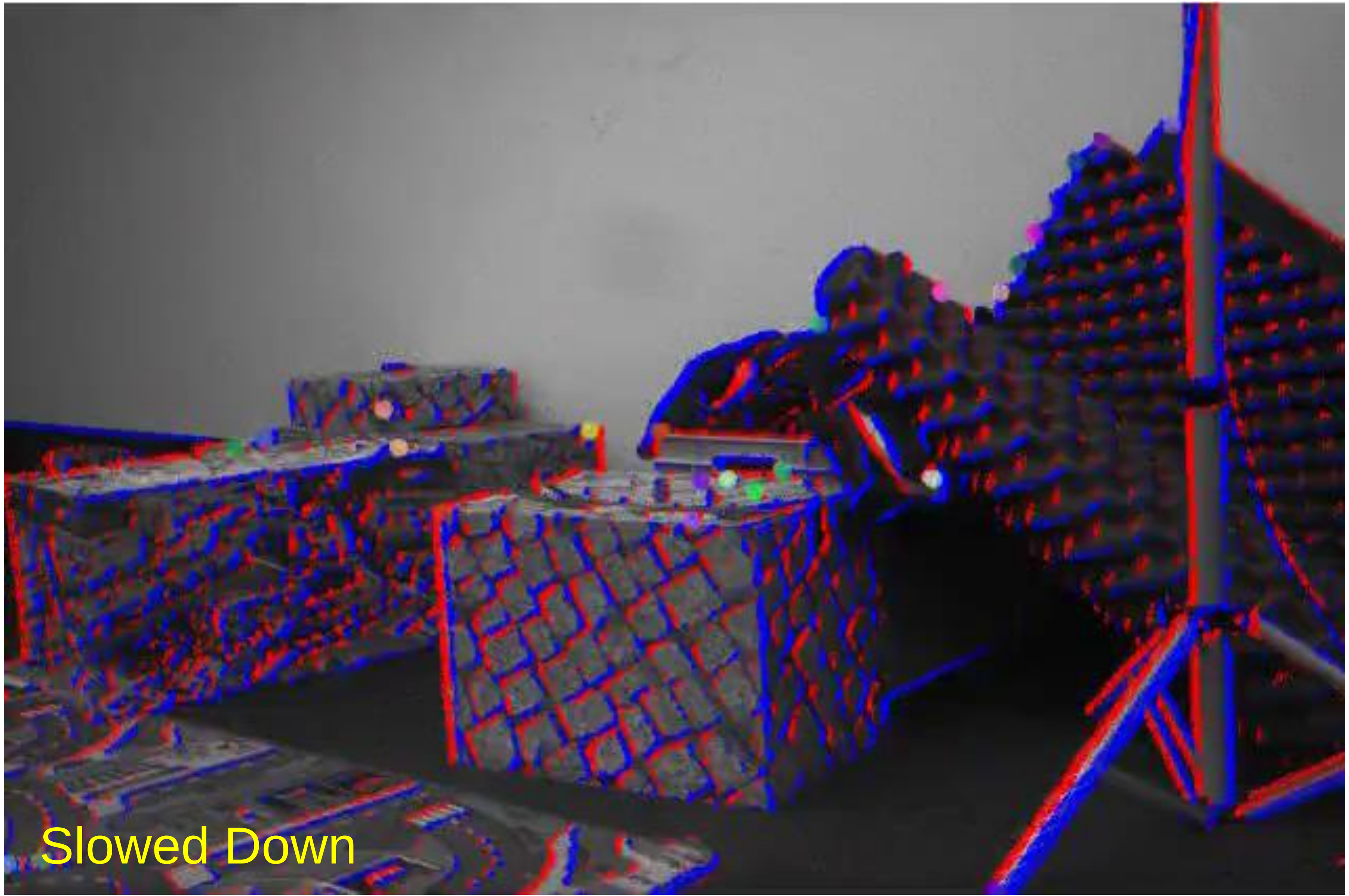
ROBOTICS & PERCEPTION GROUP

rpg.ifi.uzh.ch

We introduce the first data-driven feature tracker for event cameras



Our method predicts stable feature tracks in high-speed motion in which standard frames suffer from motion blur.



Existing feature trackers for event cameras rely on classical model assumptions

- Kueng et al., Low-latency visual odometry using event-based feature tracks. IROS, 2016
- Ni et al., Asynchronous event-based visual shape tracking for stable haptic feedback in microrobotics. IEEE Trans. Robot., 2012
- Zhu et al., Event-based feature tracking with probabilistic data association. ICRA, 2017
- Besl et al., A method for registration of 3d shapes. PAMI, 1992
- Dong et al., Standard and event cameras fusion for feature tracking. ACM, 2021
- Gehrig et al., EKLT: Asynchronous Photometric Feature Tracking Using Events and Frames. IJCV, 2020
- Seok et al., Robust feature tracking in dvs event stream using bezier mapping. WACV, 2020
- Alzugaray et al., ACE: An efficient asynchronous corner tracker for event cameras. 3DV, 2018
- Alzugaray et al., HASTE: multi-Hypothesis Asynchronous Speeded-up Tracking of Events. BMVC, 2020
- Hu et al., CDT: Event Clustering for Simultaneous Feature Detection and Tracking. IROS, 2020

Existing feature trackers for event cameras rely on classical model assumptions

- Kueng et al., Low-latency visual odometry using event-based feature tracks. IROS, 2016
- Ni et al., Asynchronous event-based visual shape tracking for stable haptic feedback in microrobotics. IEEE Trans. Robot., 2012
- Zh
- B
- D
- G
- Se
- A
- A
- Hu et al., CDT: Event Clustering for Simultaneous Feature Detection and Tracking. IROS, 2020

- Require extensive manual hand-tuning to adapt to different event cameras
- Difficulties to generalize to different scenarios due to unmodeled effects

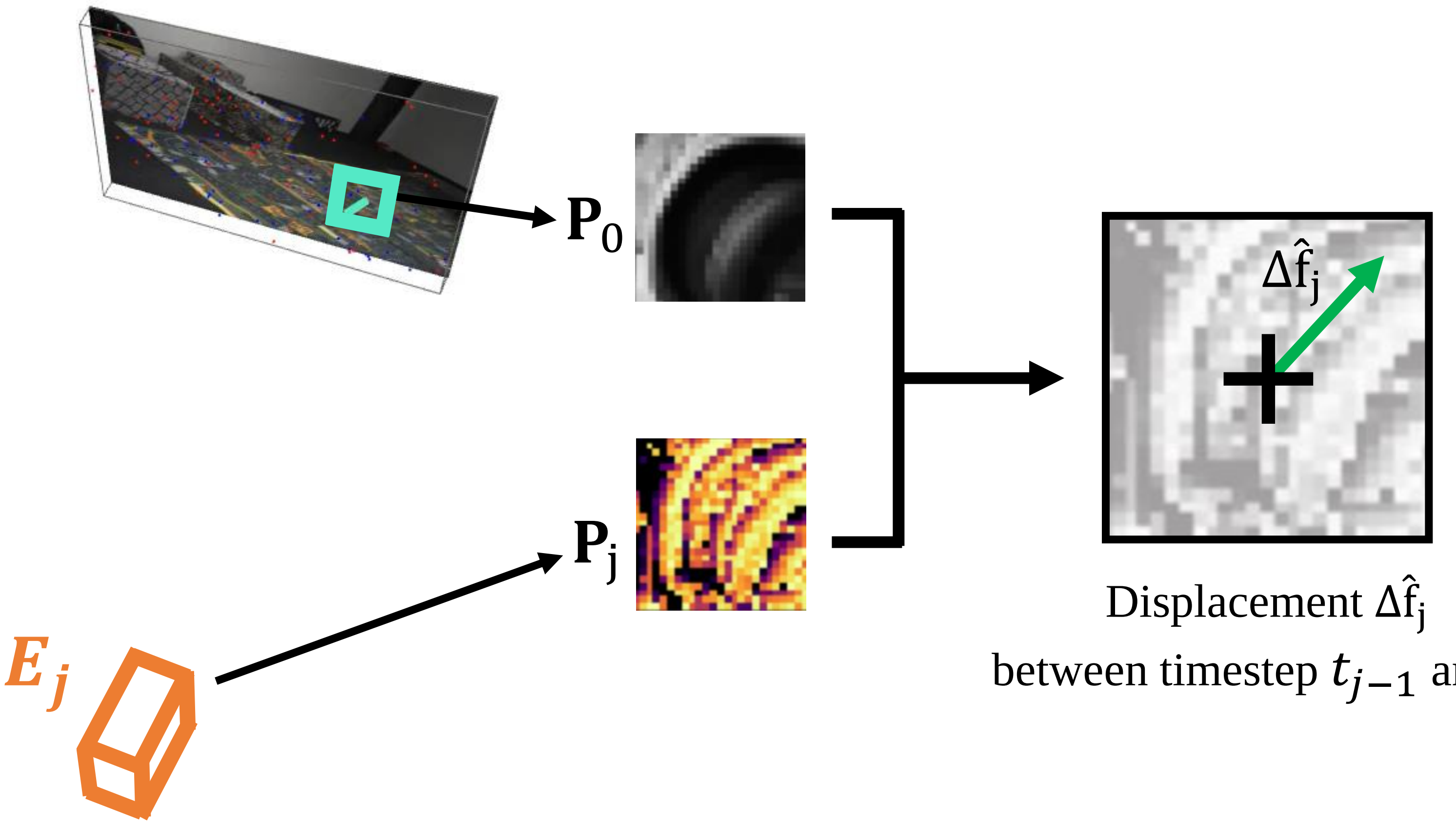
Existing feature trackers for event cameras rely on classical model assumptions

- Kueng et al., Low-latency visual odometry using event-based feature tracks. IROS, 2016
- Ni et al., Asynchronous event-based visual shape tracking for stable haptic feedback in microrobotics. IEEE Trans. Robot., 2012
- Zh
- B
- D
- G
- Se
- A
- A
- Hu et al., CDT: Event Clustering for Simultaneous Feature Detection and Tracking. IROS, 2020

- Require extensive manual hand-tuning to adapt to different event cameras
- Difficulties to generalize to different scenarios due to unmodeled effects

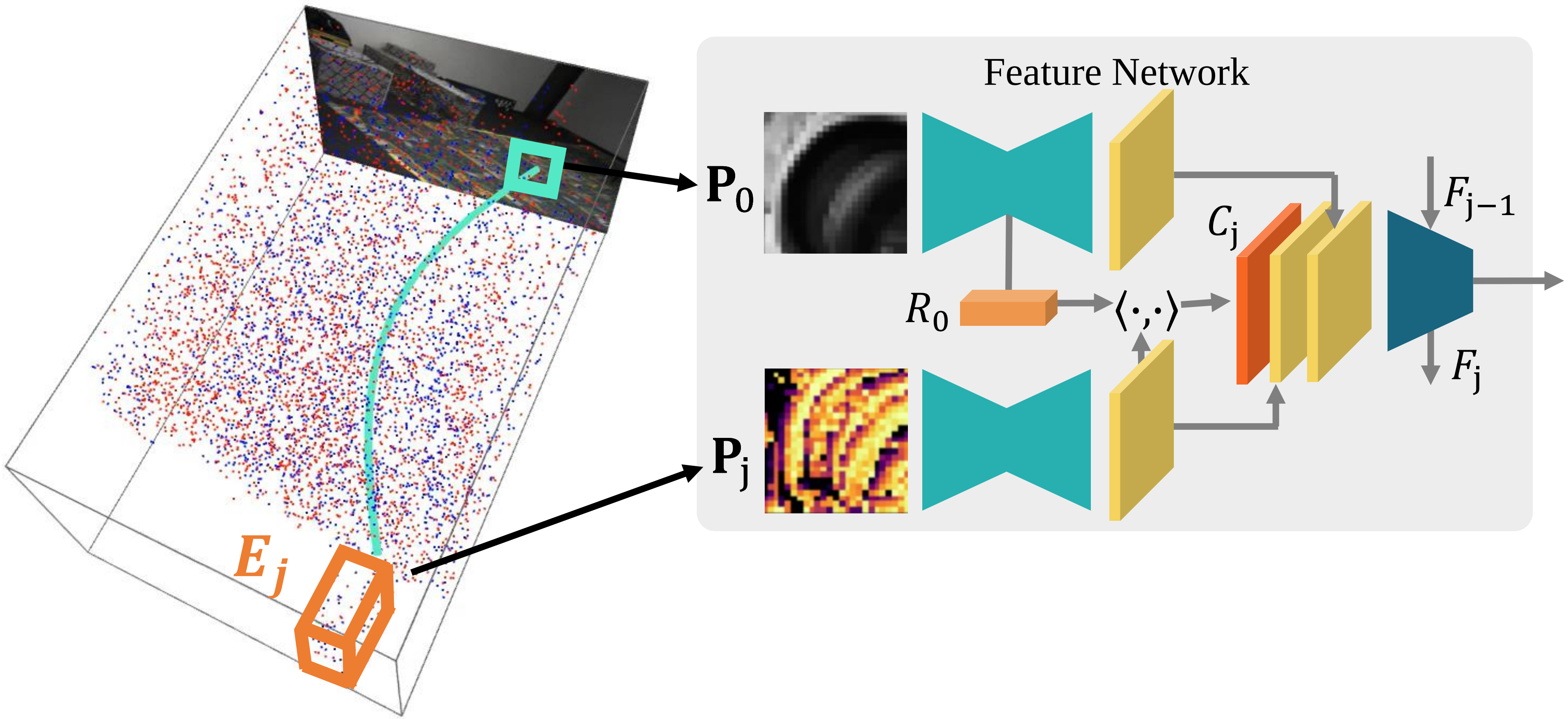
We propose the first data-driven feature tracker for event cameras

Our method predicts the displacement $\Delta\hat{\mathrm{f}}_{\mathrm{j}}$ of a feature by localizing a template patch $\mathbf{P}_0$ from a grayscale image $\boldsymbol{I}_0$ in subsequent event patches $\mathbf{P}_j$.

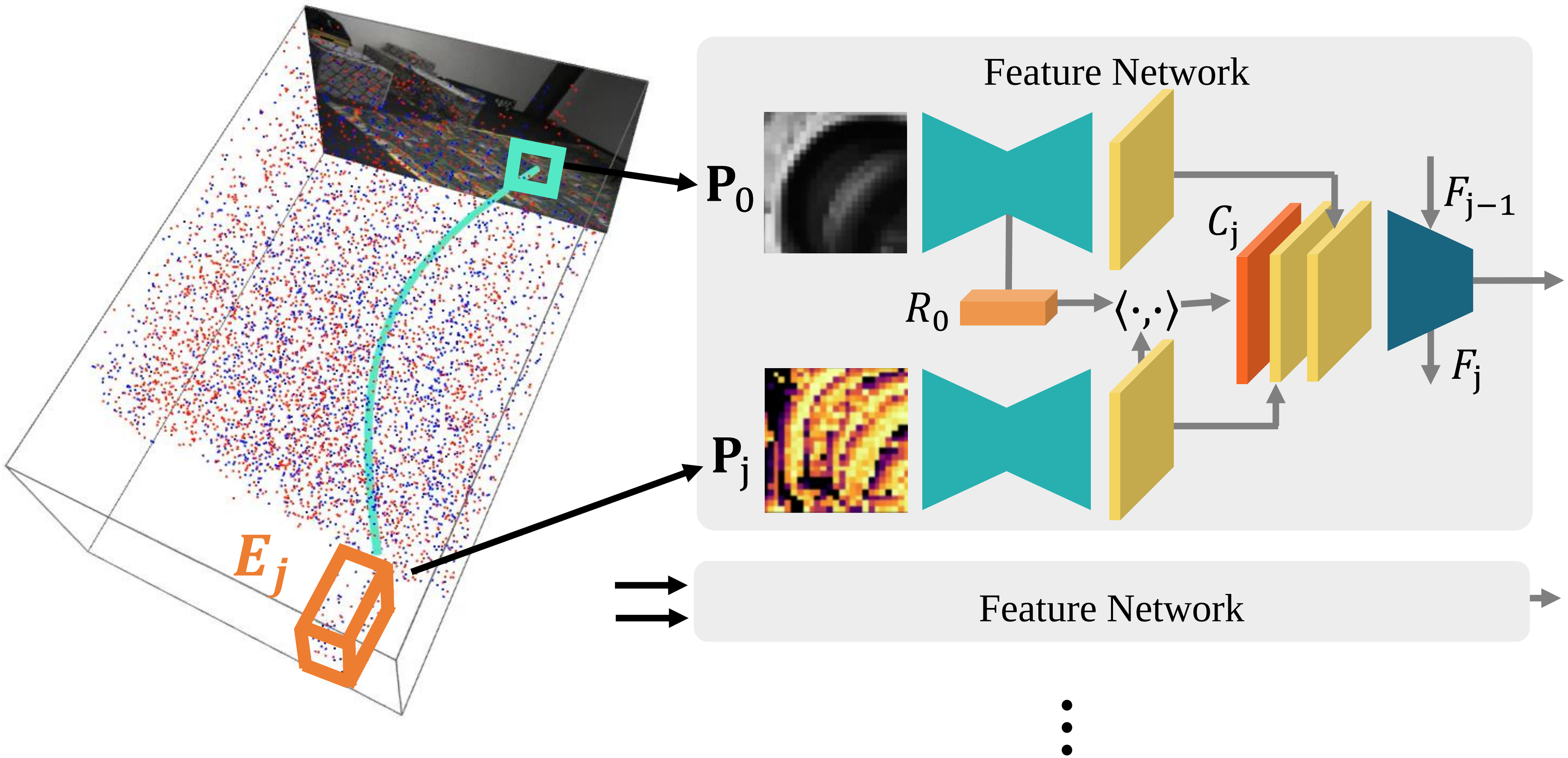


Displacement $\Delta\hat{\mathrm{f}}_{\mathrm{j}}$
between timestep t_{j-1} and t_j

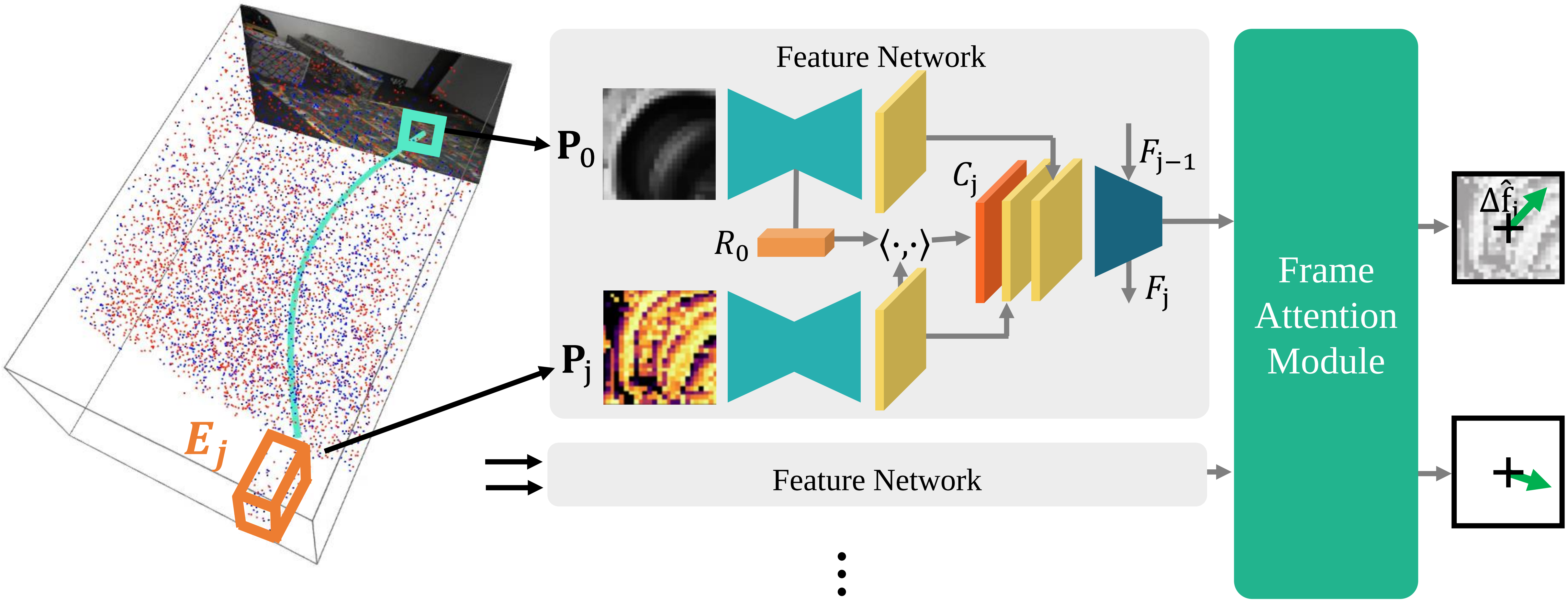
The feature network encodes both patches using a correlation and recurrent layers into a single feature vector with spatial dimension of 1×1.



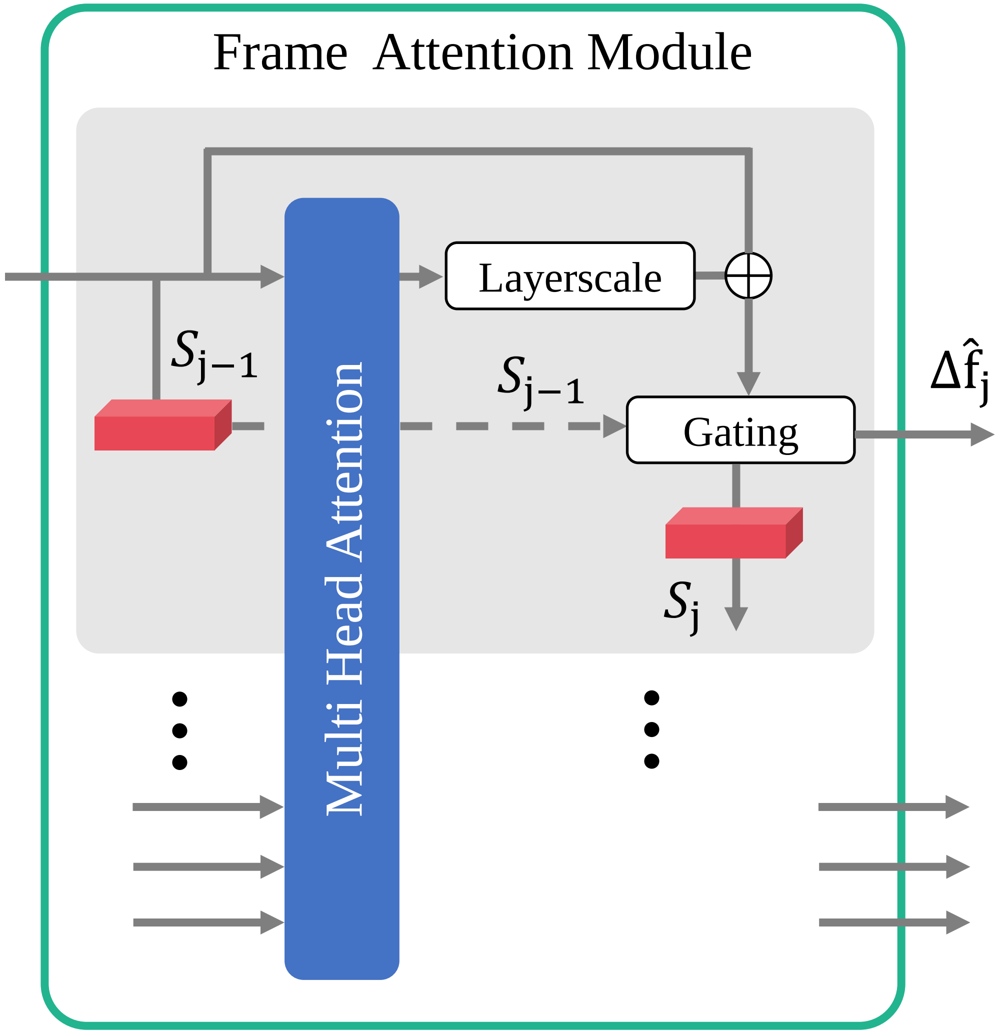
Each feature track is independently processed by the feature network.



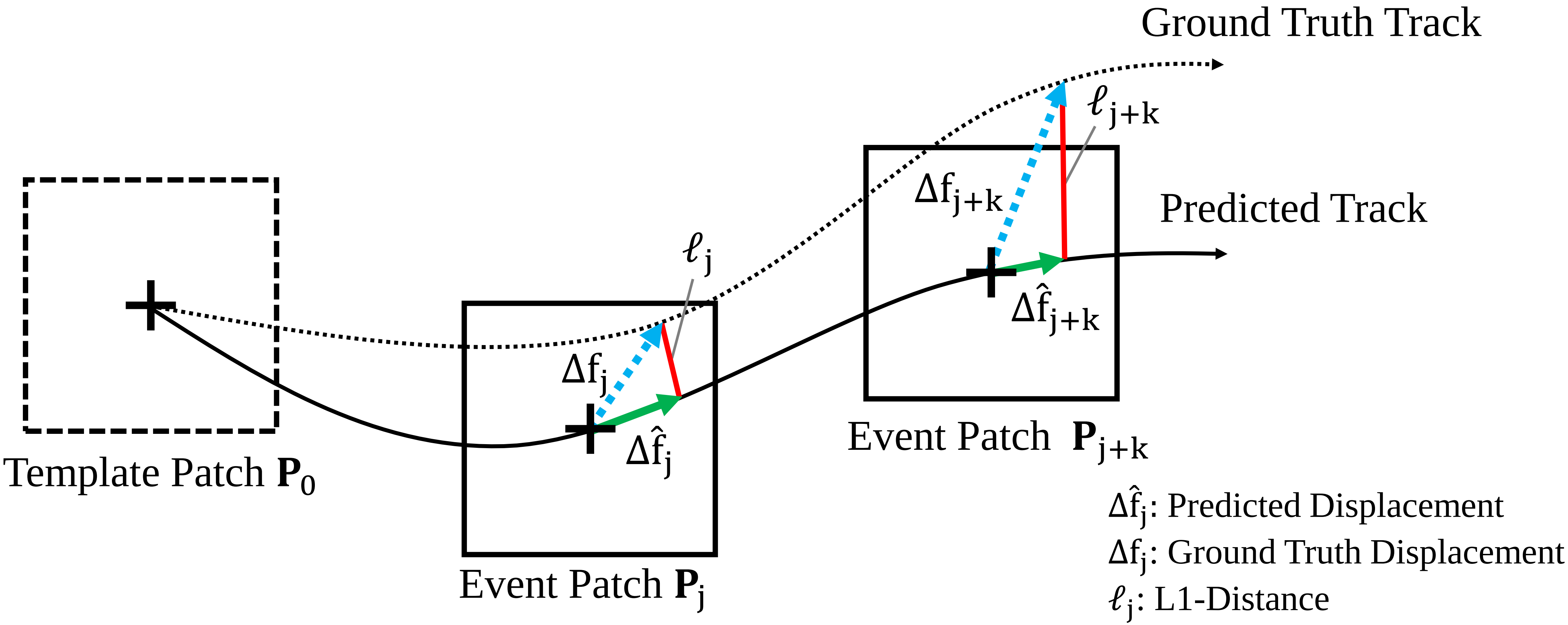
To share information between features in the same image, we introduce a novel frame attention module.



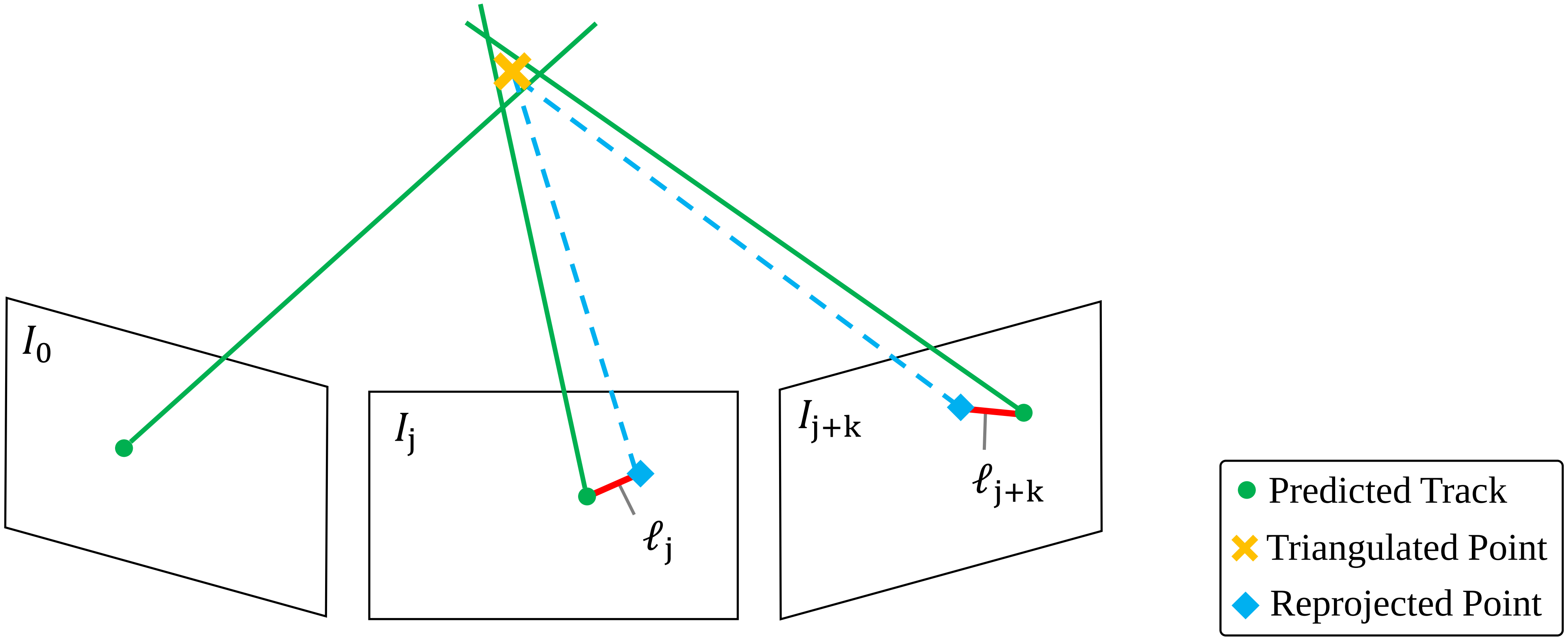
The frame attention module uses a self attention layer to share the information across the feature tracks and outputs the feature displacement $\Delta\hat{\mathrm{f}}_{\mathrm{j}}$.



We train our network on synthetic data by directly computing the L1-Distance between the predicted $\Delta\hat{f}_j$ and ground truth displacement Δf_j.



To close the gap between synthetic and real data, we introduce a fine-tuning strategy, which triangulates and reprojects a 3D point using camera poses.



By directly transferring zero-shot from synthetic to real data, our tracker outperforms existing approaches in relative feature age by up to 120%.

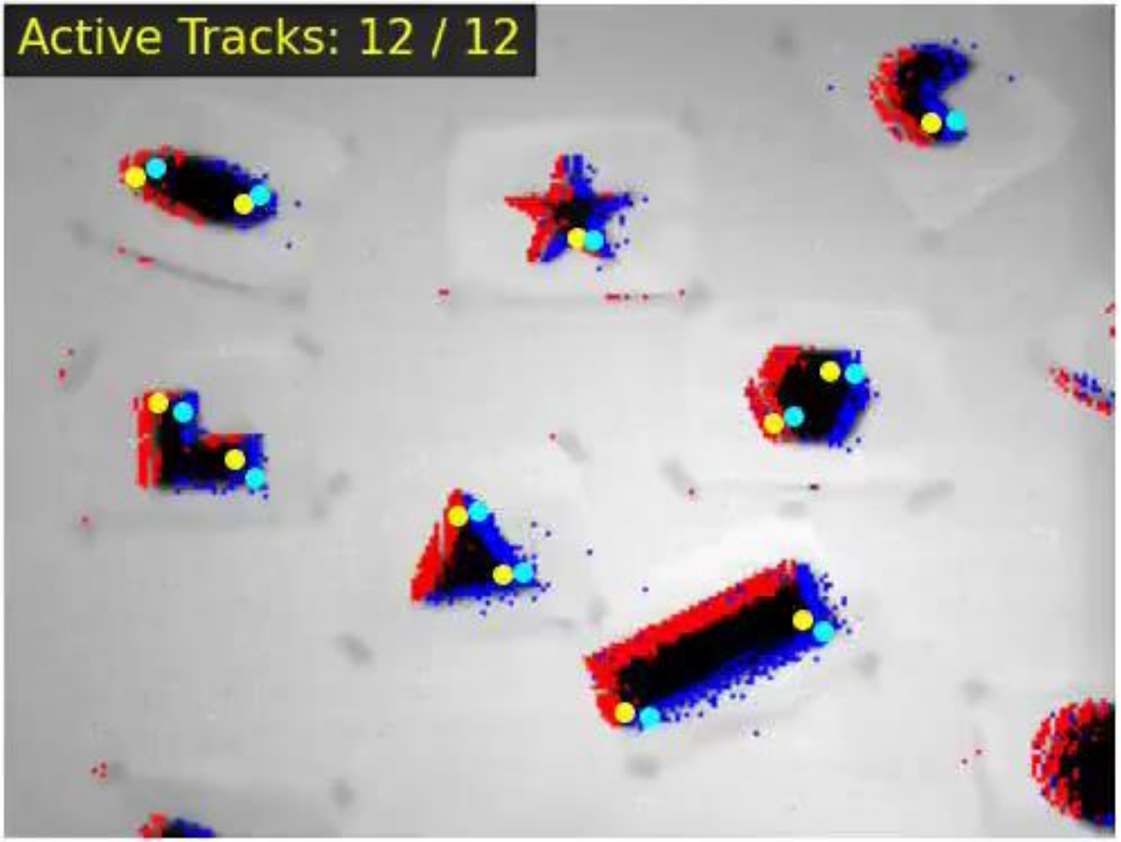
	EDS		EC	
Method	Feature Age (FA) ↑	Expected FA ↑	Feature Age (FA) ↑	Expected FA ↑
ICP [24]	0.060	0.040	0.256	0.245
EM-ICP [46]	0.161	0.120	0.337	0.334
HASTE [4]	0.096	0.063	0.442	0.427
EKLT [17]	0.325	0.205	0.811	0.775
Ours (zero-shot)	0.549	0.451	0.795	0.787

This performance gap is further increased to 130% by adapting our tracker to real data with a novel self-supervision strategy.

	EDS		EC	
Method	Feature Age (FA) ↑	Expected FA ↑	Feature Age (FA) ↑	Expected FA ↑
ICP [24]	0.060	0.040	0.256	0.245
EM-ICP [46]	0.161	0.120	0.337	0.334
HASTE [4]	0.096	0.063	0.442	0.427
EKLT [17]	0.325	0.205	0.811	0.775
Ours (zero-shot)	0.549	0.451	0.795	0.787
Ours (fine-tuned)	**0.576**	**0.472**	**0.825**	**0.818**

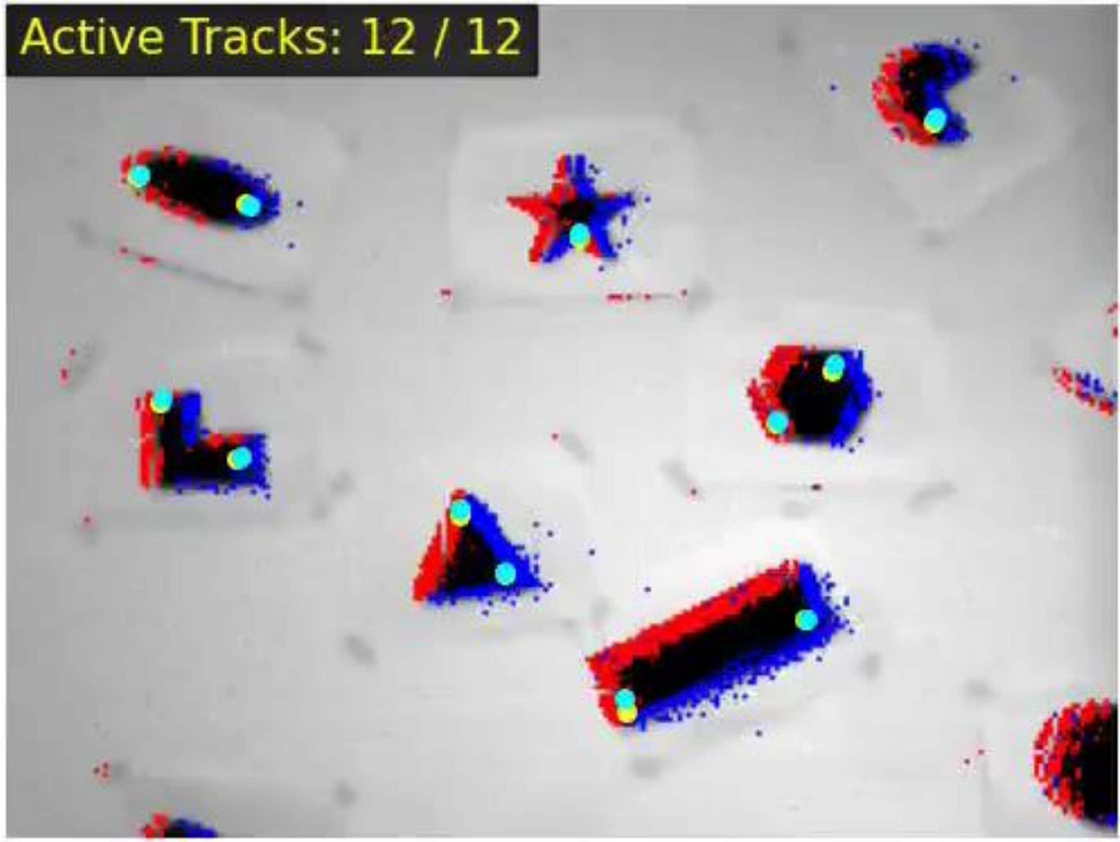
Qualitative Results EC

EKLT



Expected Feature Age: 0.696

Ours



Expected Feature Age: 0.882

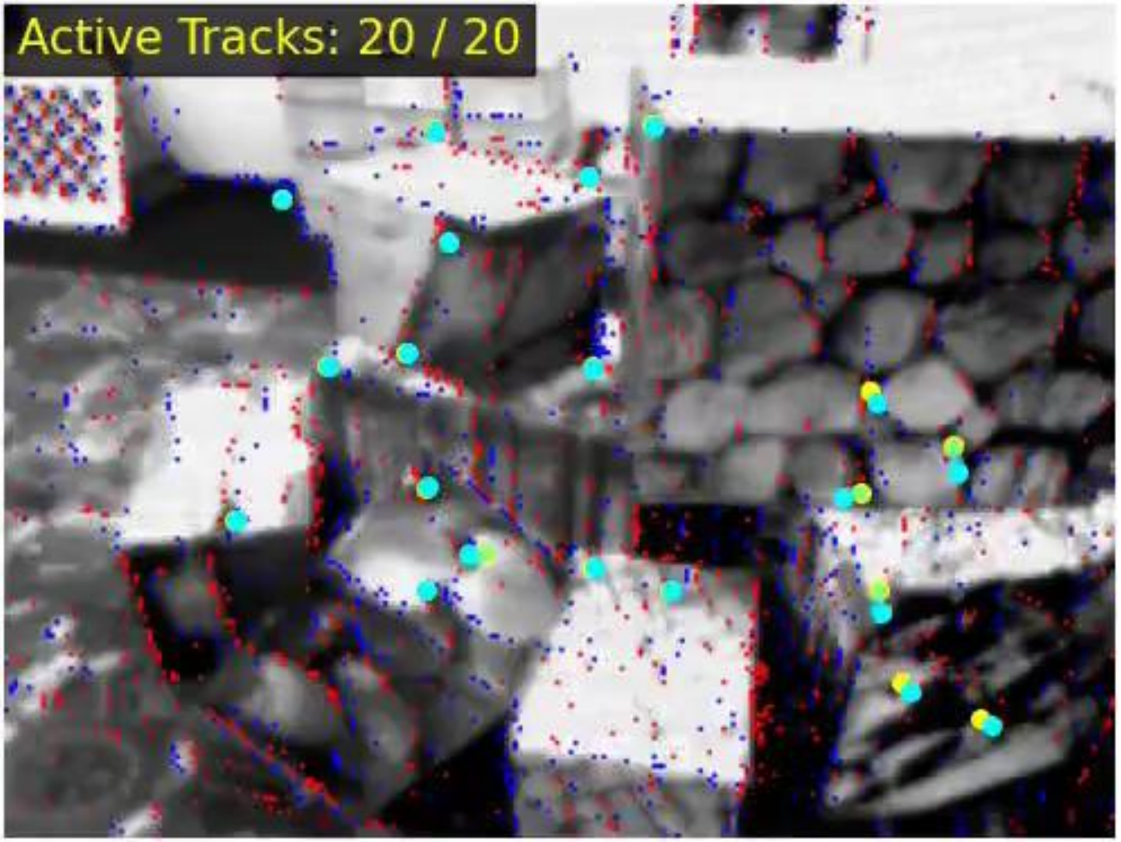
Slowed Down 0.1X

- Positive Events
- Negative Events

- Prediction
- Ground Truth

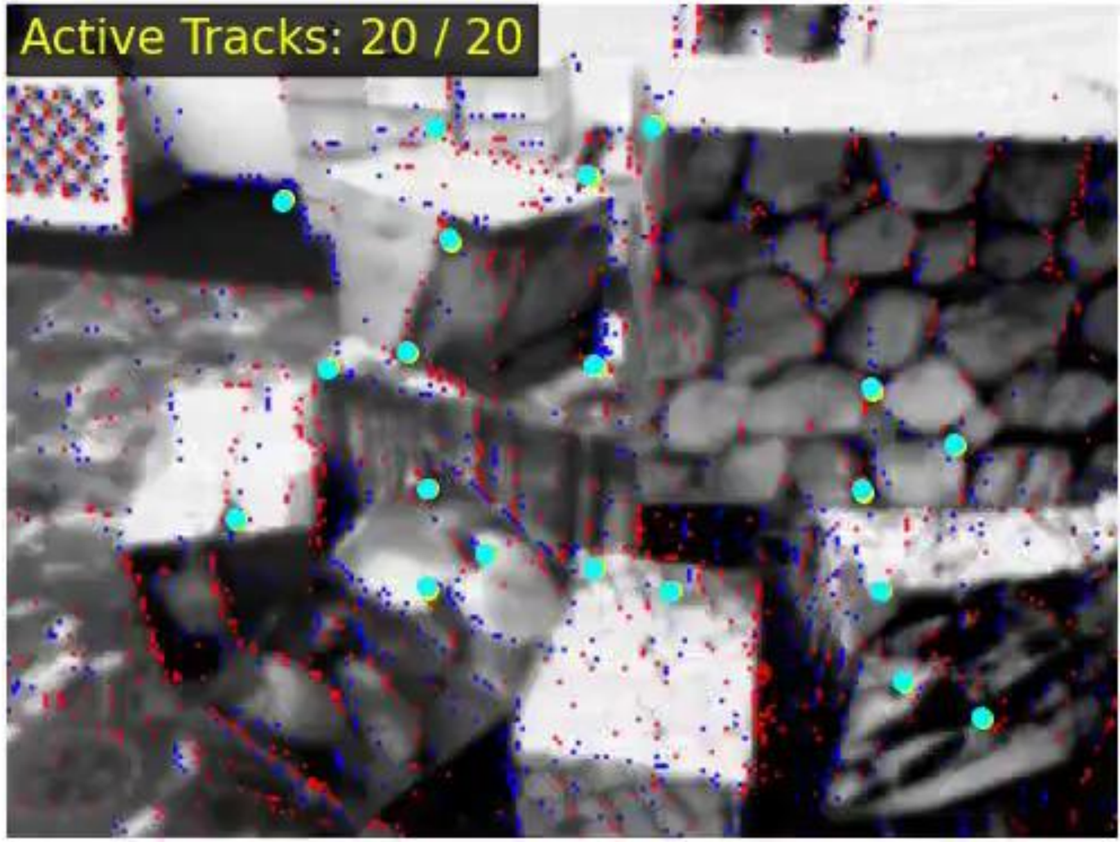
Qualitative Results EC

EKLT



Expected Feature Age: 0.644

Ours



Expected Feature Age: 0.869

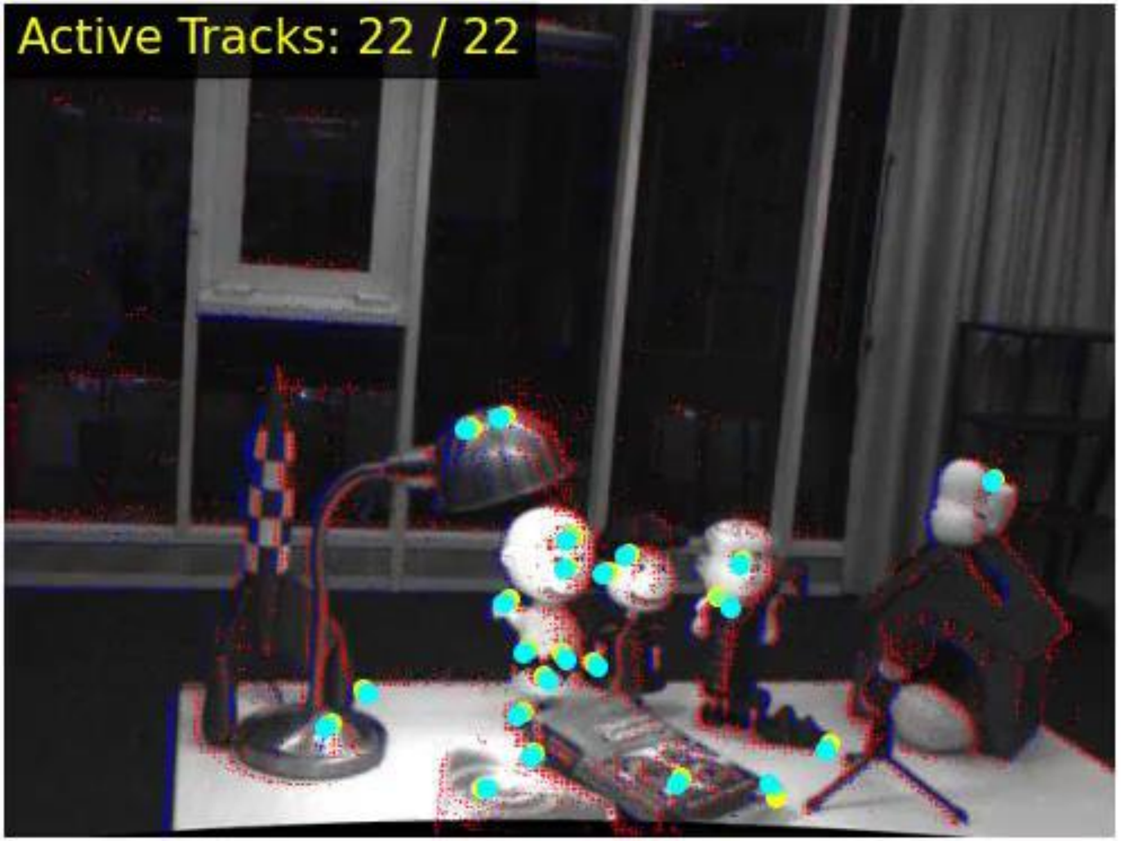
Slowed Down 0.1X

Positive Events
Negative Events

Prediction
Ground Truth

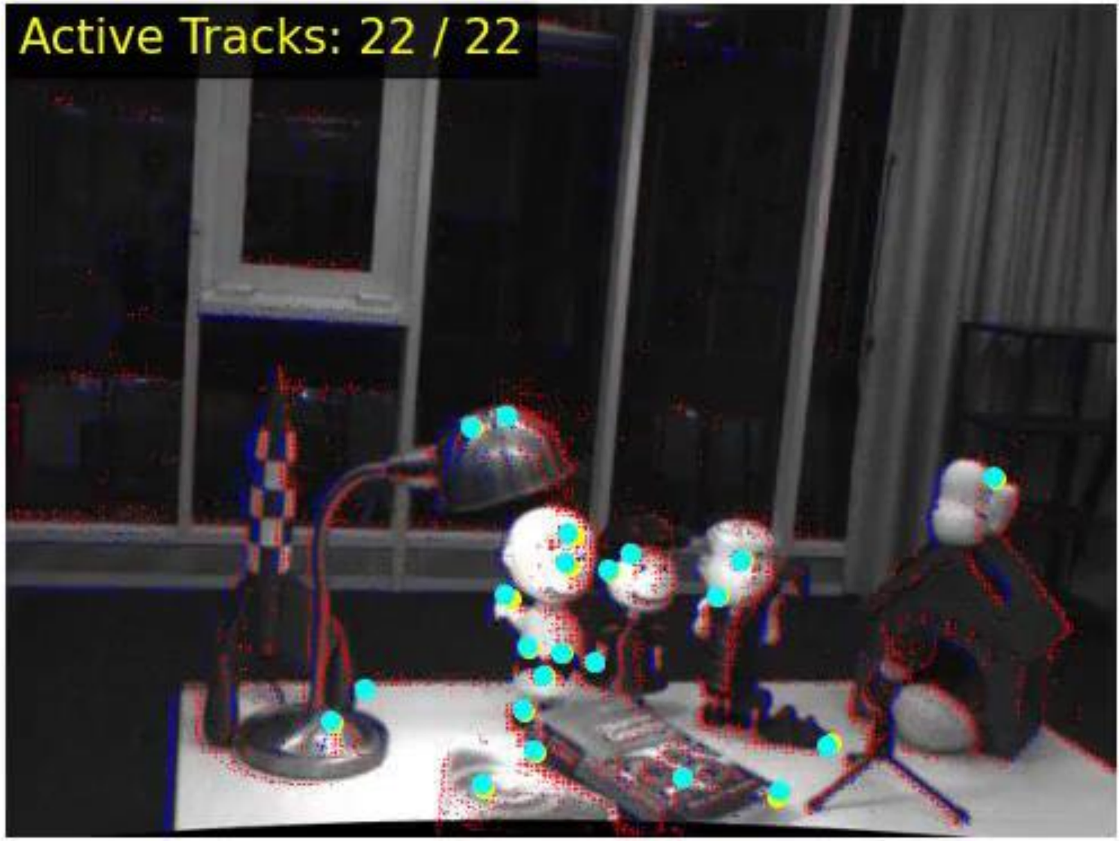
Qualitative Results EDS

EKLT



Expected Feature Age: 0.153

Ours



Expected Feature Age: 0.428

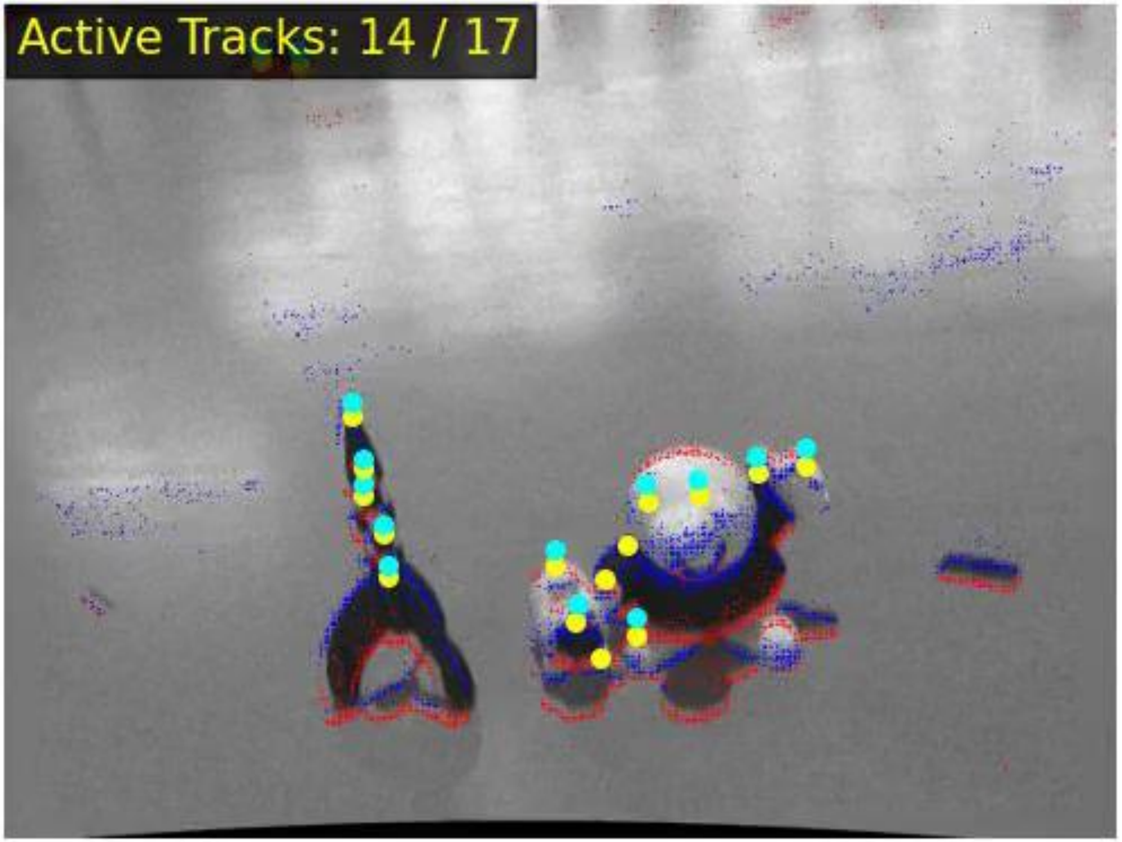
Slowed Down 0.3X

- Positive Events
- Negative Events

- Prediction
- Ground Truth

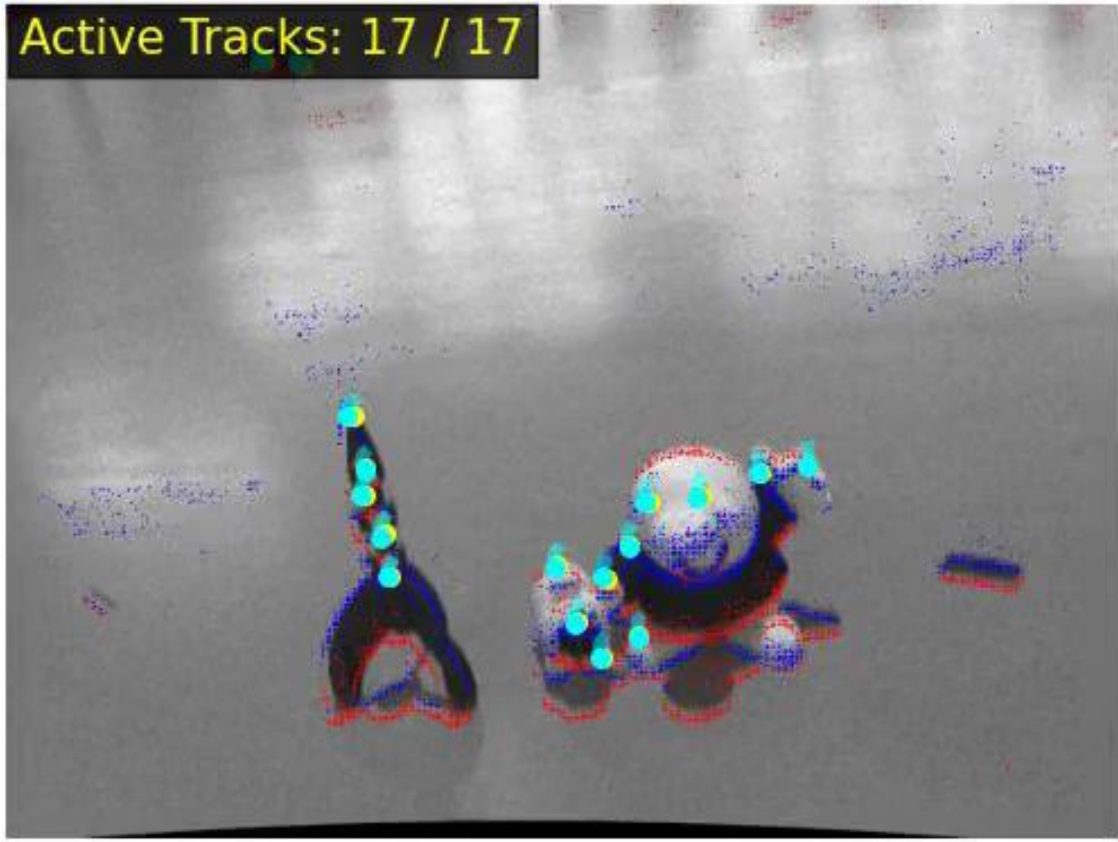
Qualitative Results EDS

EKLT



Expected Feature Age: 0.231

Ours



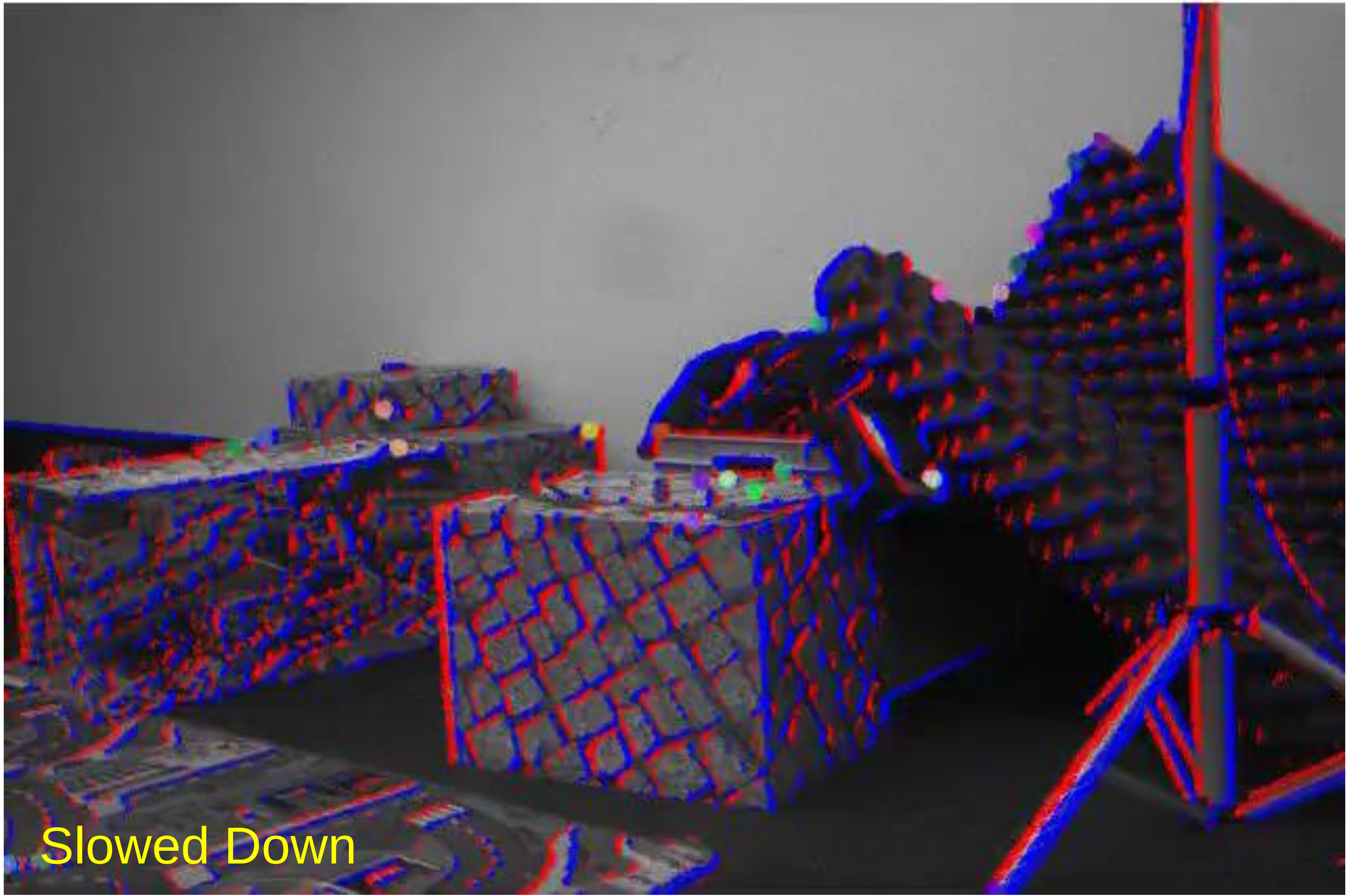
Expected Feature Age: 0.746

Slowed Down 0.3X

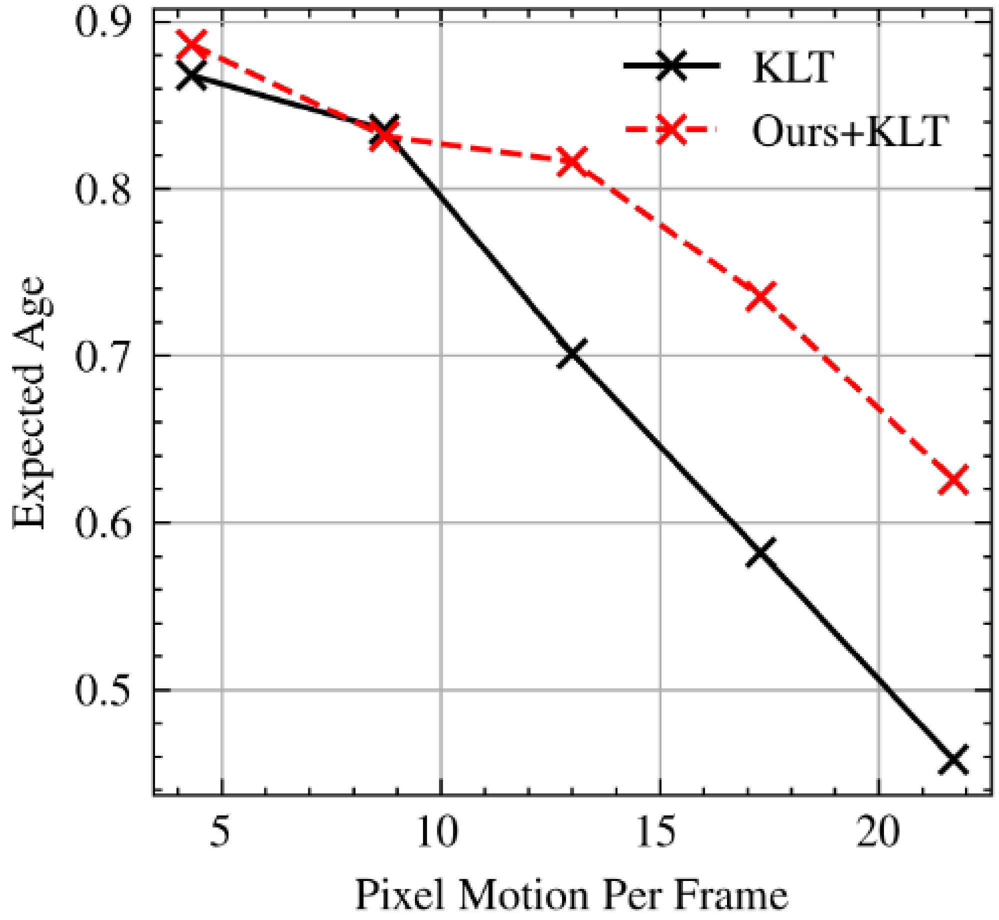
Positive Events
Negative Events

Prediction
Ground Truth

Finally, our method predicts stable feature tracks in high-speed motion in which standard frames suffer from motion blur.



Furthermore, we can combine our tracker with the frame-based KLT tracker increasing the robustness of feature tracks in high-speed motion.



Conclusion

- We introduce the first data-driven feature tracker for event cameras, which leverages low-latency events to track features detected in a grayscale frame.

- Our data-driven tracker outperforms existing approaches in relative feature age by up to 130 % while also achieving the lowest latency.

Source Code: https://github.com/uzh-rpg/deep_ev_tracker

ROBOTICS &
PERCEPTION
GROUP

rpg.ifi.uzh.ch