

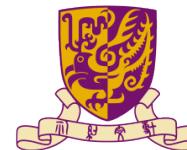


From Sim-to-Real: Toward General Event-based Low-light Frame Interpolation with Per-scene Optimization

SIGGRAPH Asia 2024

Ziran Zhang*, Yongrui Ma*, Yueting Chen, Feng Zhang,
Jinwei Gu, Tianfan Xue†, Shi Guo†

naturezhanghn@zju.edu.cn



香港中文大學
The Chinese University of Hong Kong



Video frame interpolation



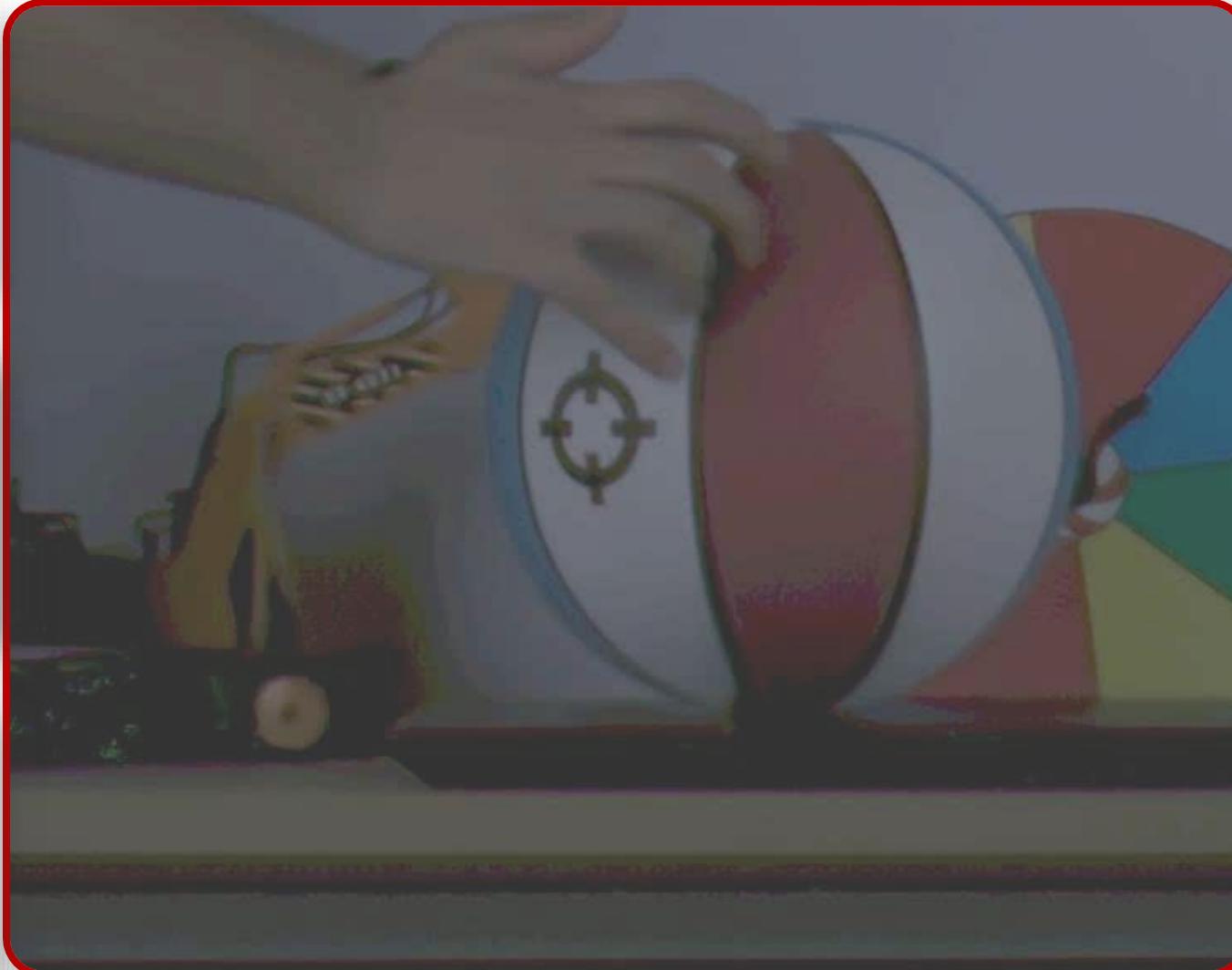
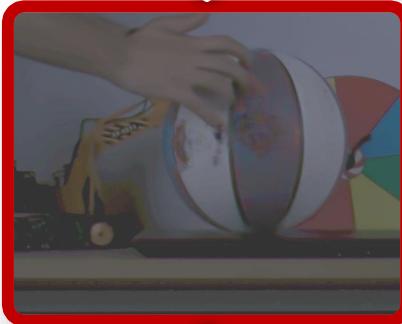
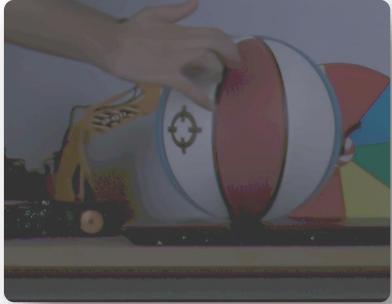
Sports Analysis



Physical Experiments

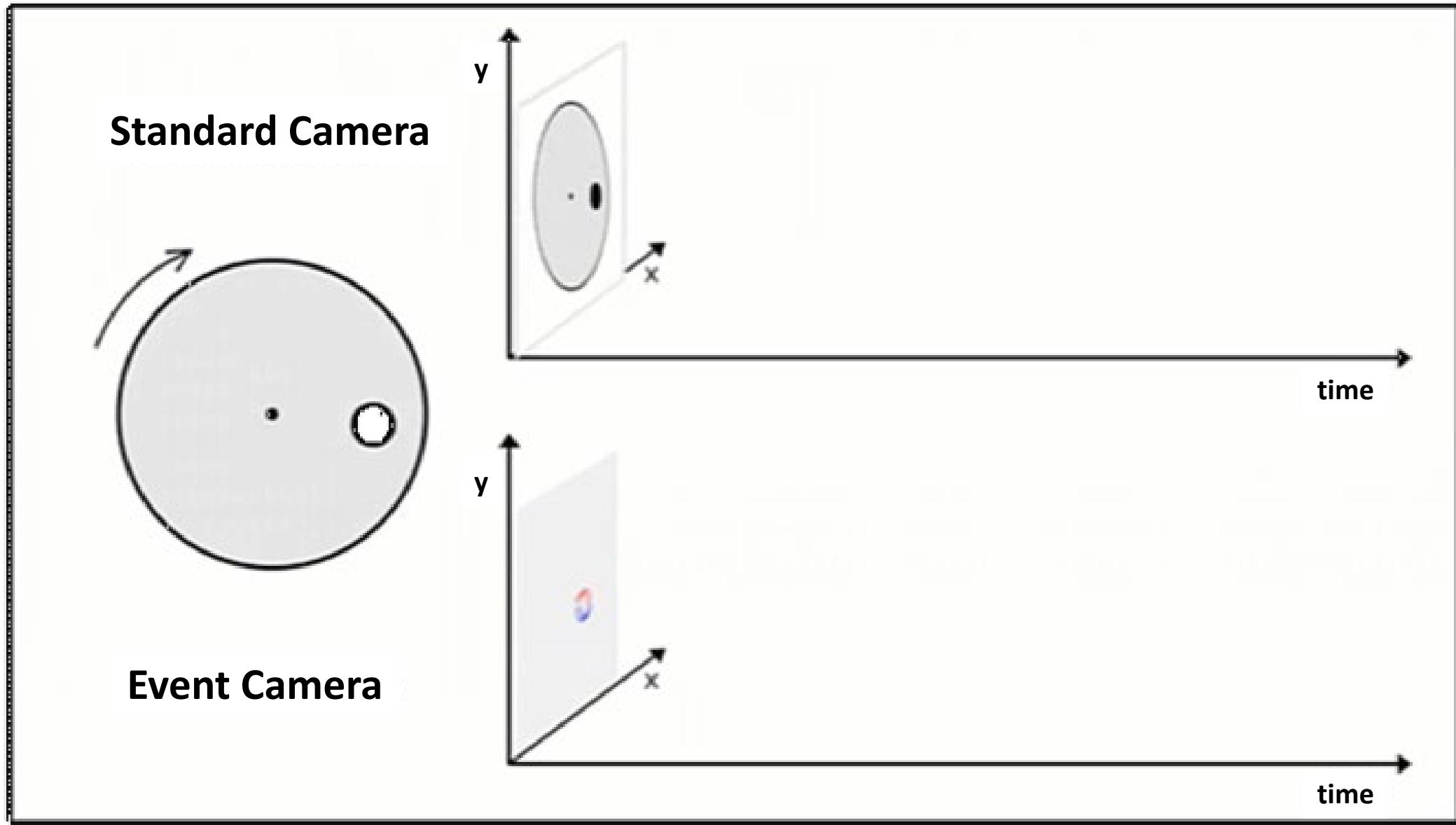


Lowlight frame interpolation



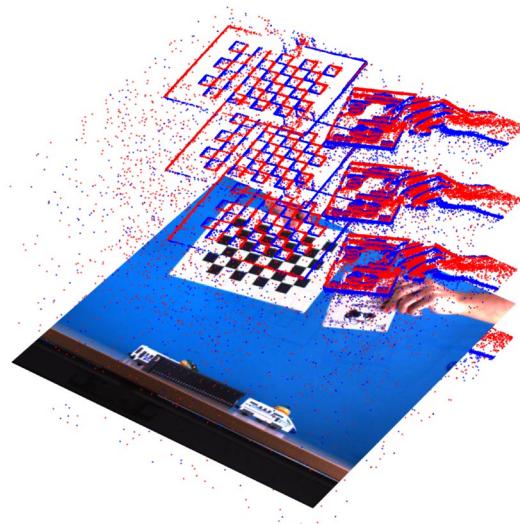
**Fast Nonlinear Low-Light
Scenes: Failed!**

Event camera

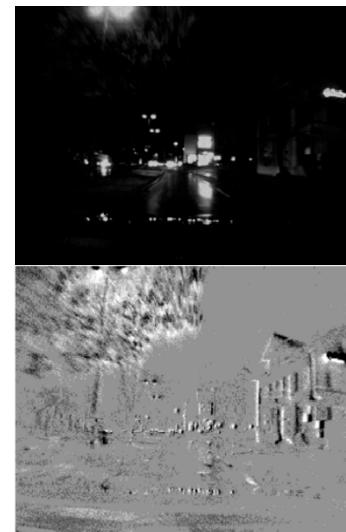




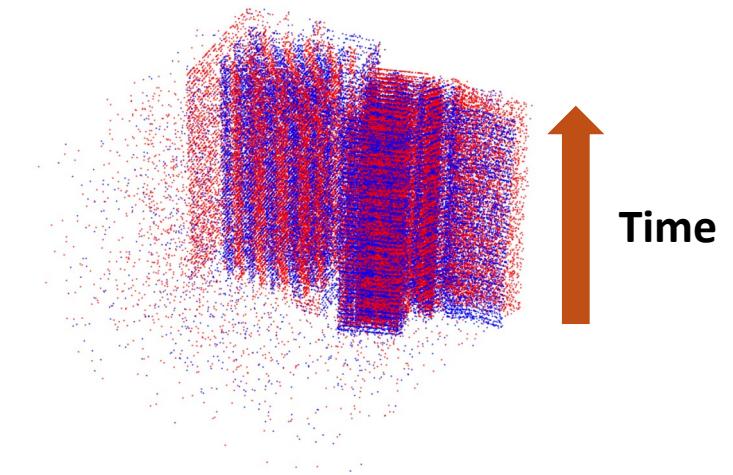
Event camera



Brightness changes



Large dynamic range

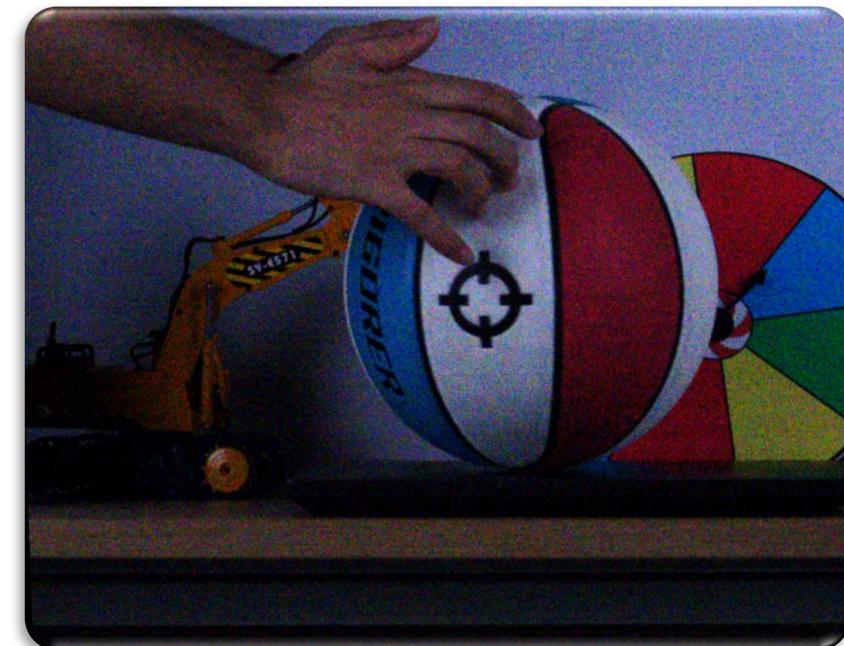


High temporal resolution

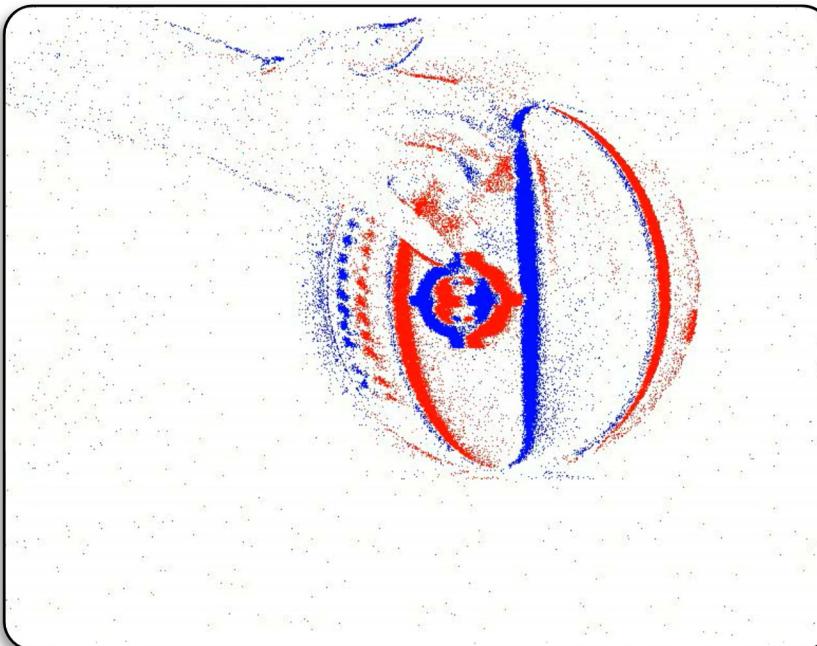


Lowlight frame interpolation

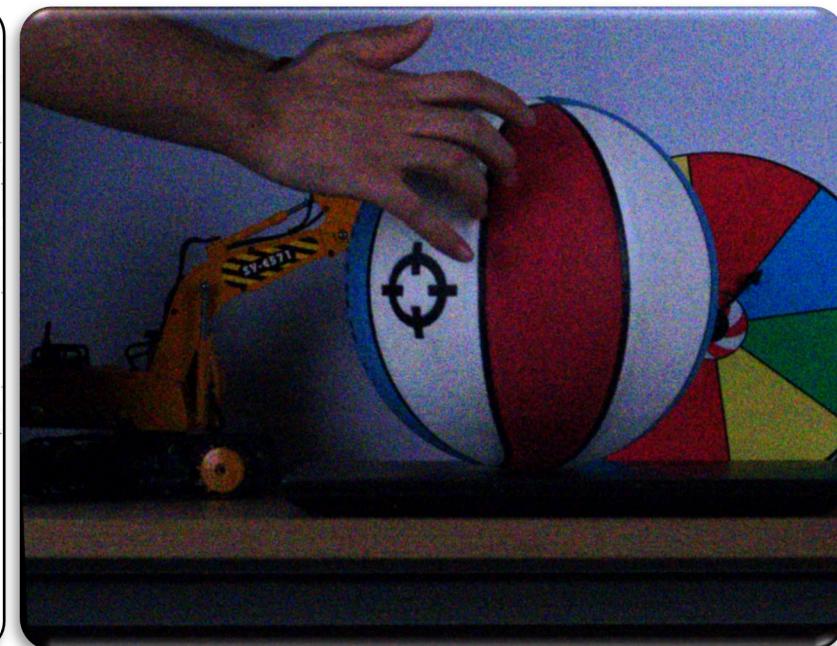
Events record **motion information**,
showing great potential in video frame interpolation.



Frame 0



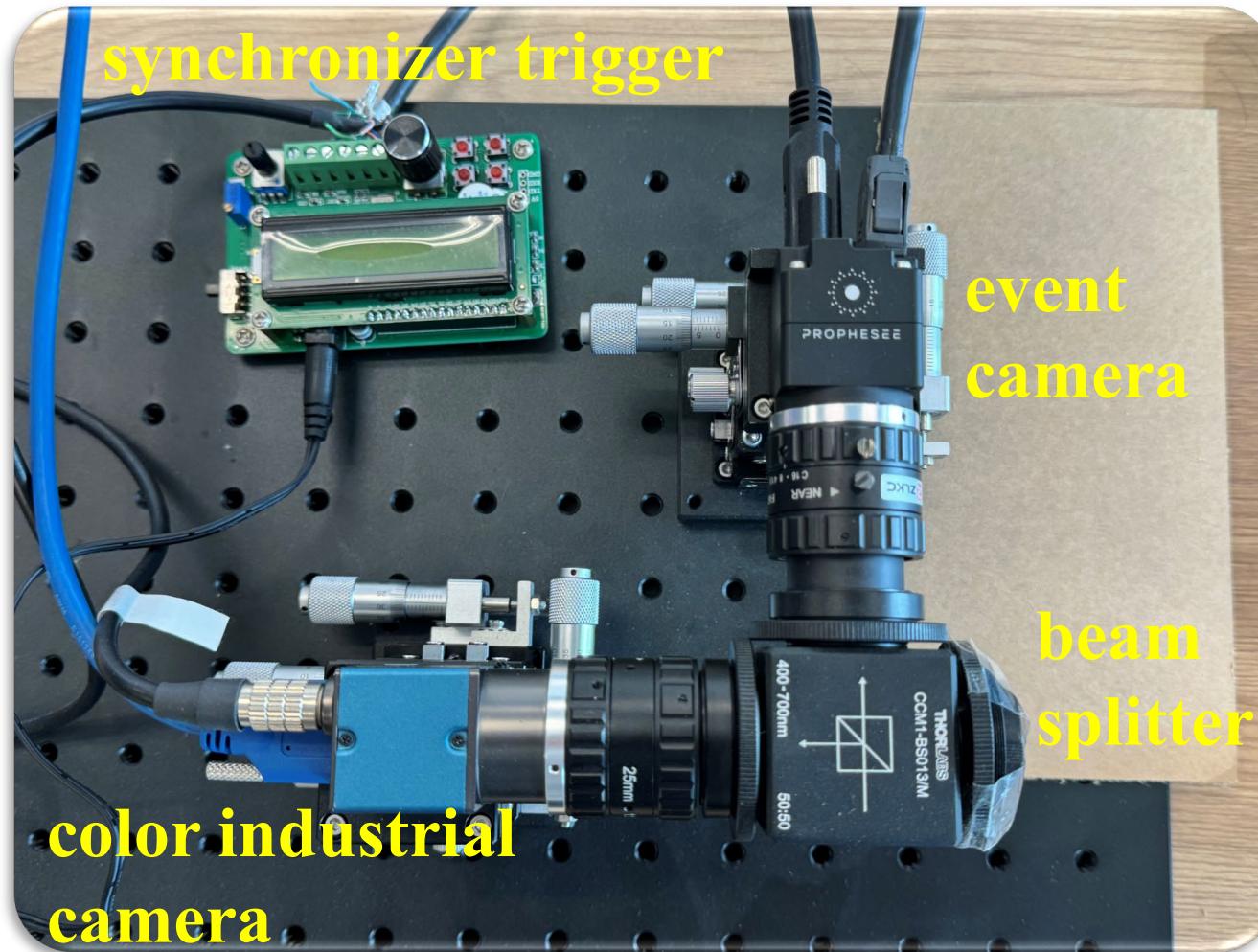
Events



Frame 1



EVS-RGB Beam Splitting Imaging System





Fast imaging demands in low light



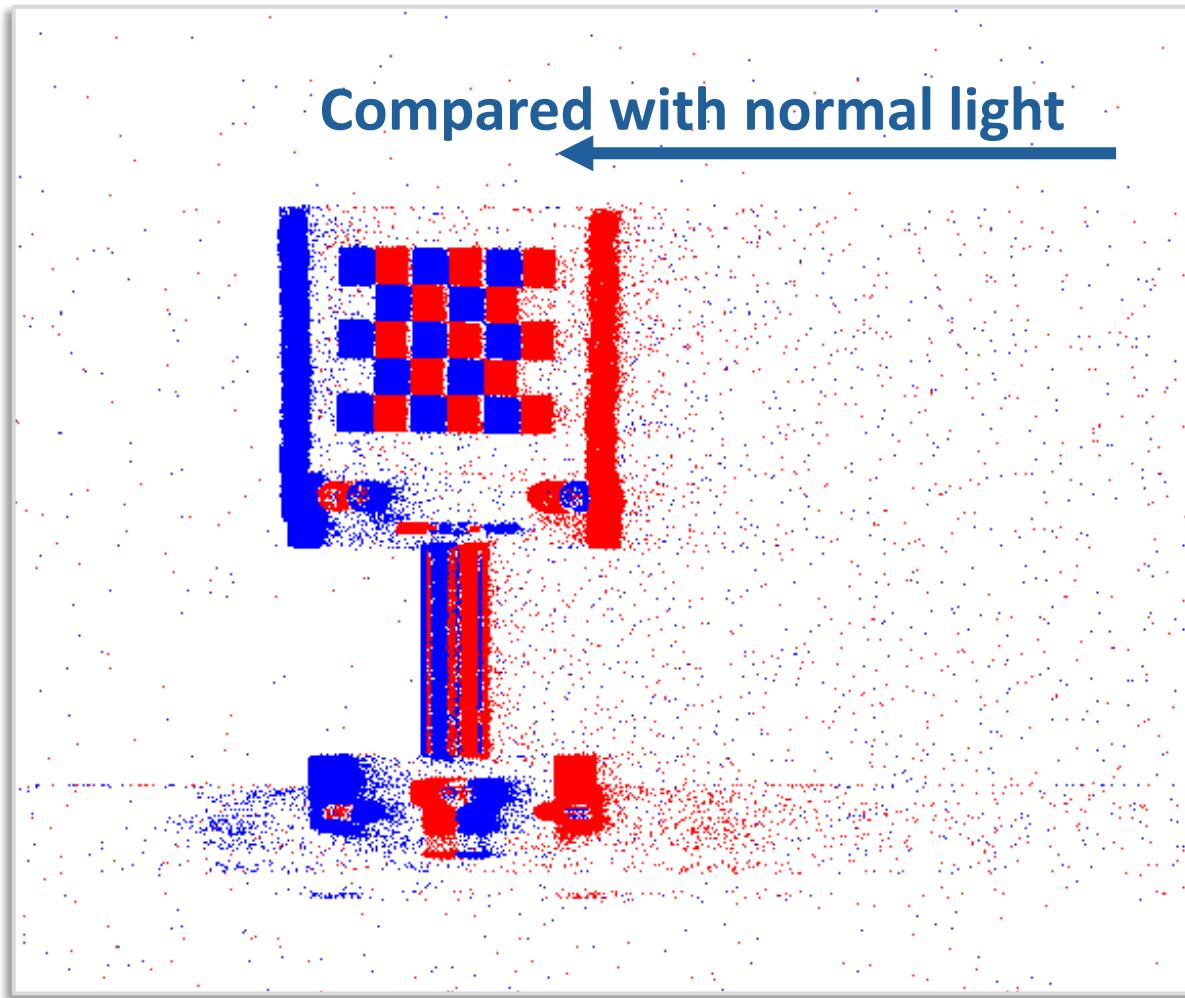
Autonomous Driving



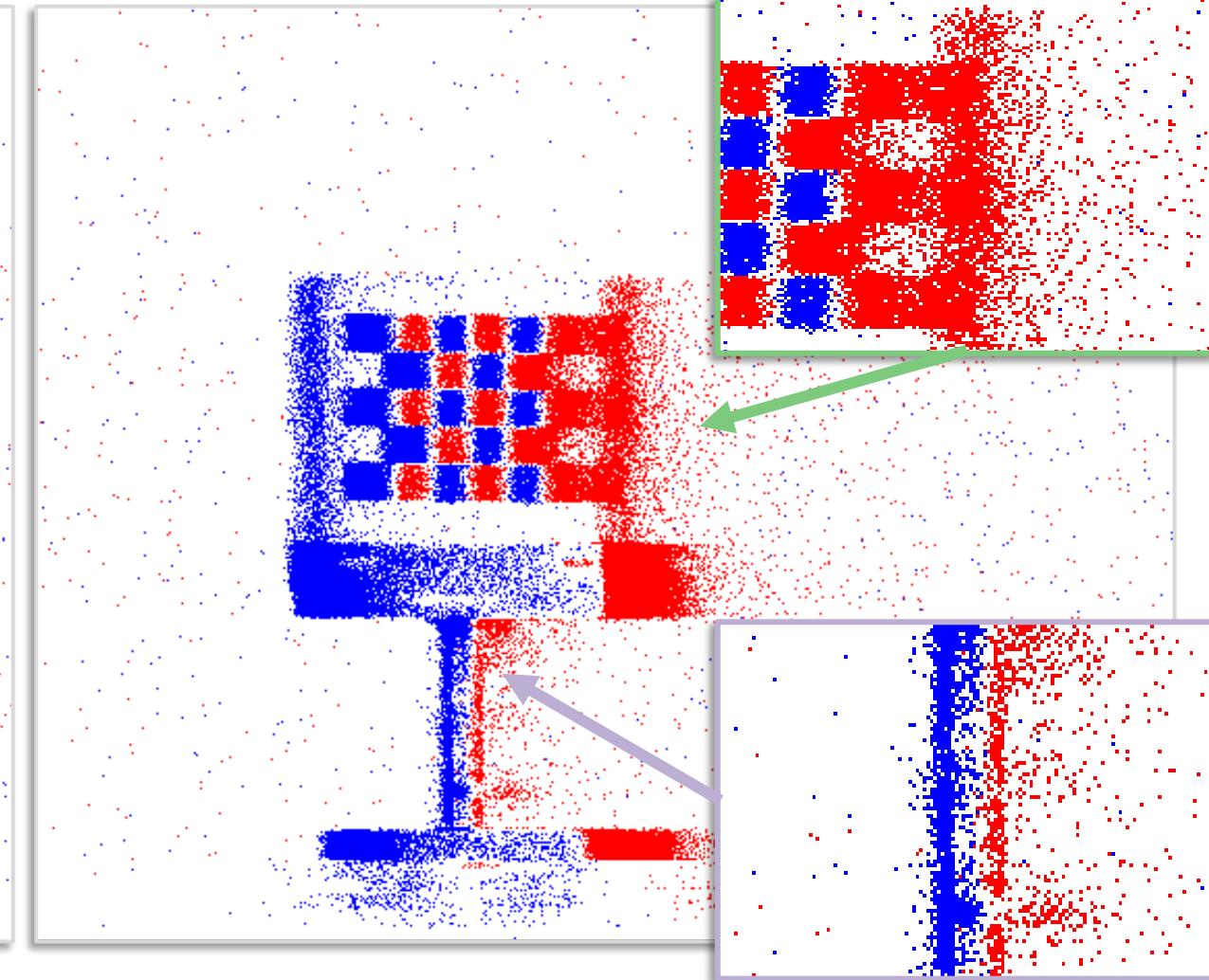
Drone Flight



Event Trailing Artifacts



Events under normal light

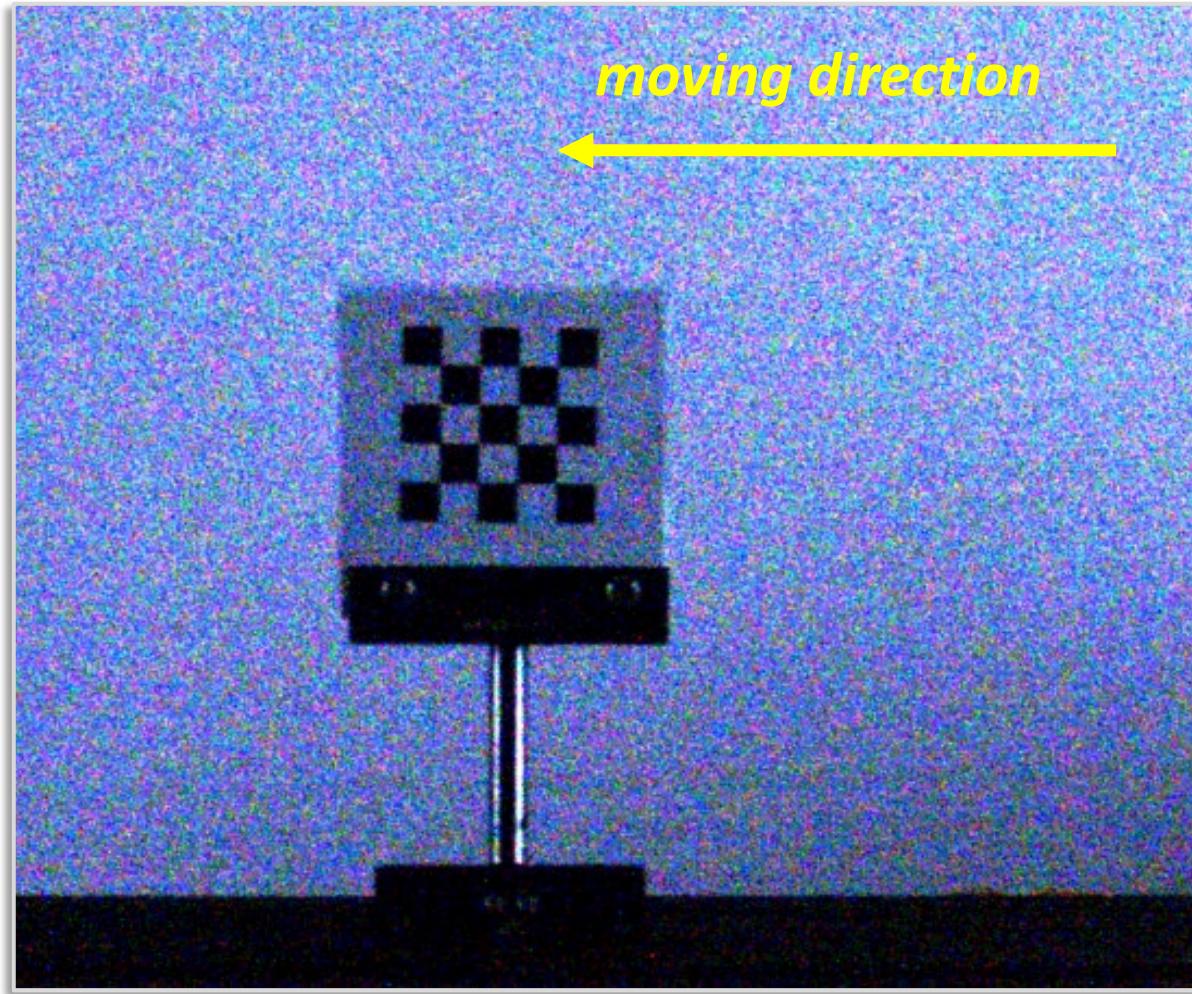


Events at 35lux

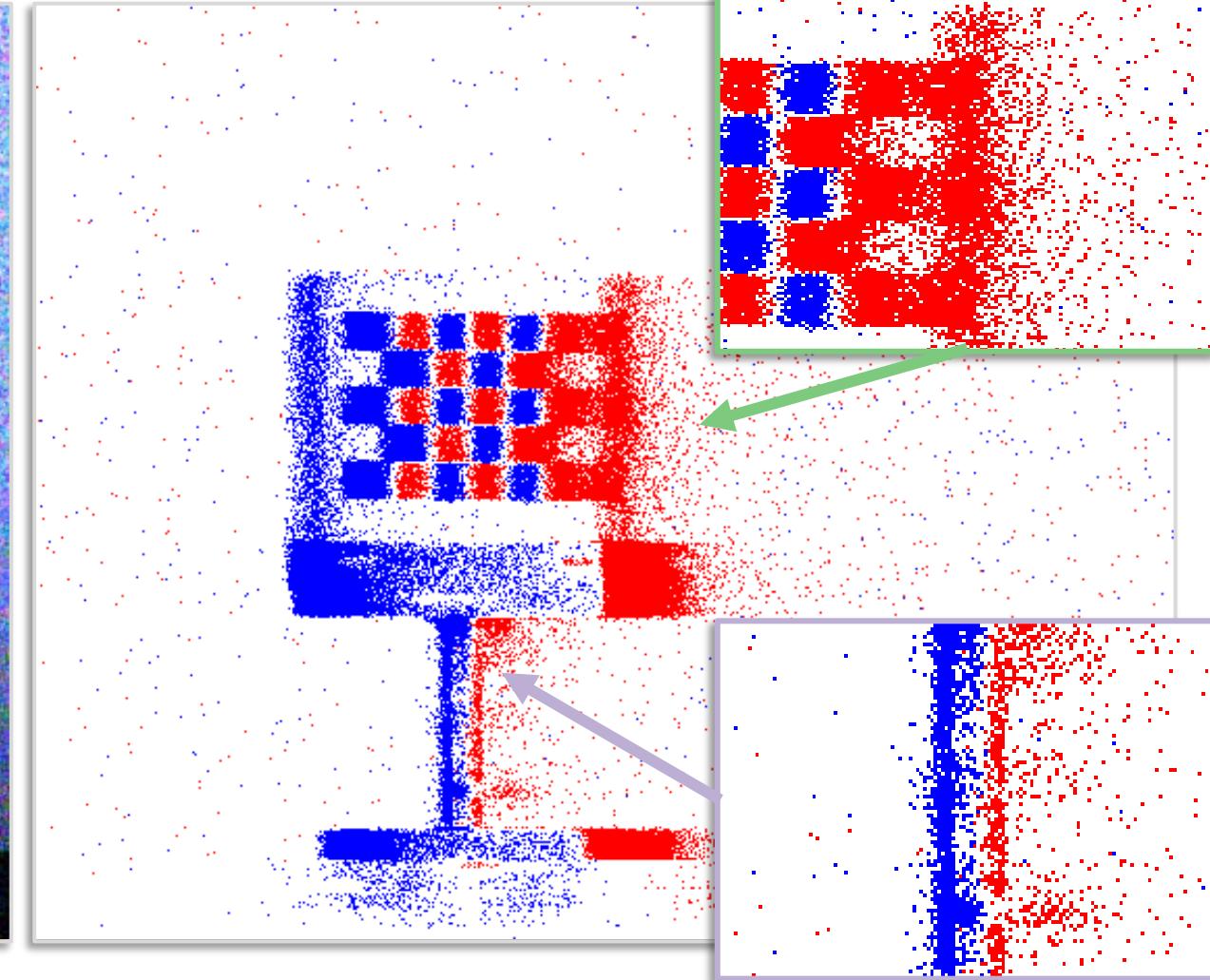
Trailing artifacts



Event Trailing Artifacts



Noisy RGB at 35lux

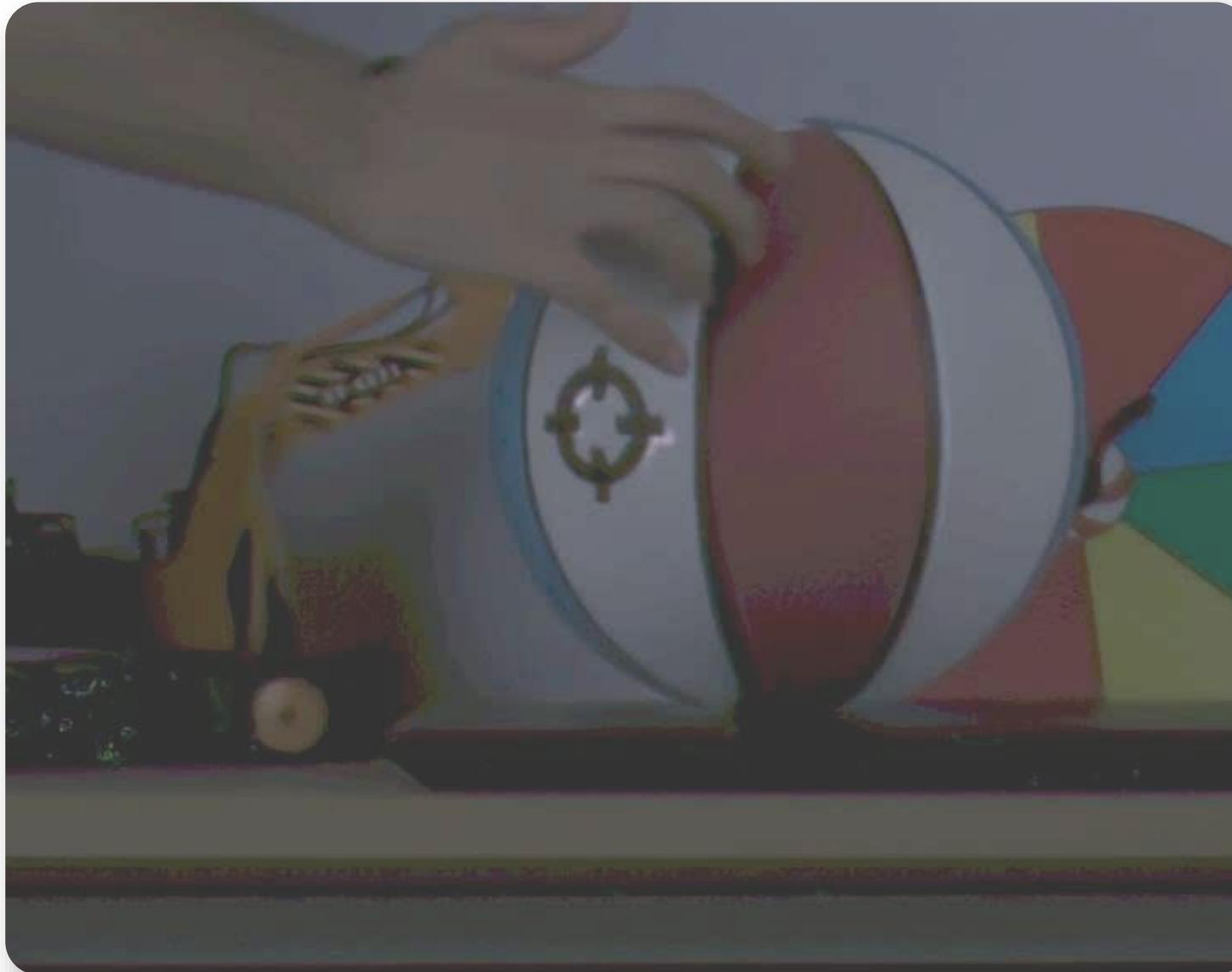


Events at 35lux

Trailing artifacts



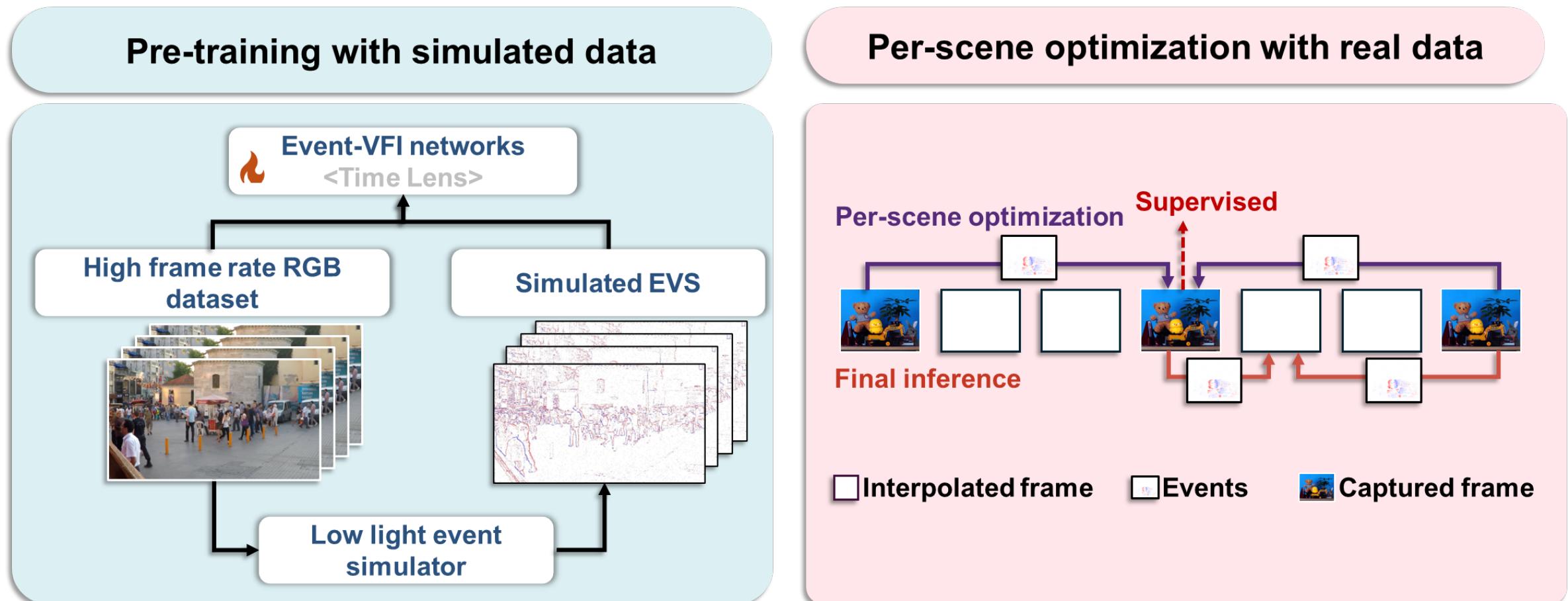
Trailing artifact



**Trailing artifact
in the final output.**

How can the Event-VFI method
be cost-effectively generalized to
real-world low-light conditions with
varying camera settings?

Our Pipeline

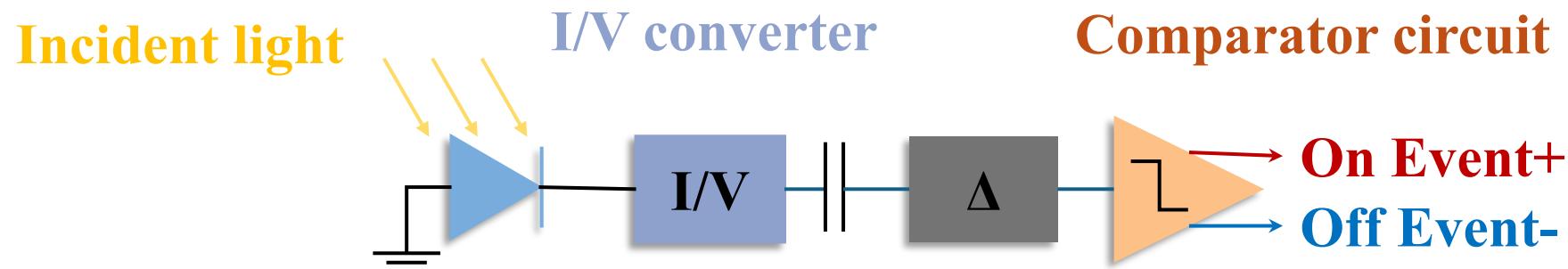


Events simulation in low light



Why events has trailing artifacts in lowlight?

Simplified Circuit Diagram



Photodiode

Capacitor:

$$\frac{dV_{\text{out}}}{dt} = \frac{1}{RC} (V_{\text{in}} - V_{\text{out}})$$

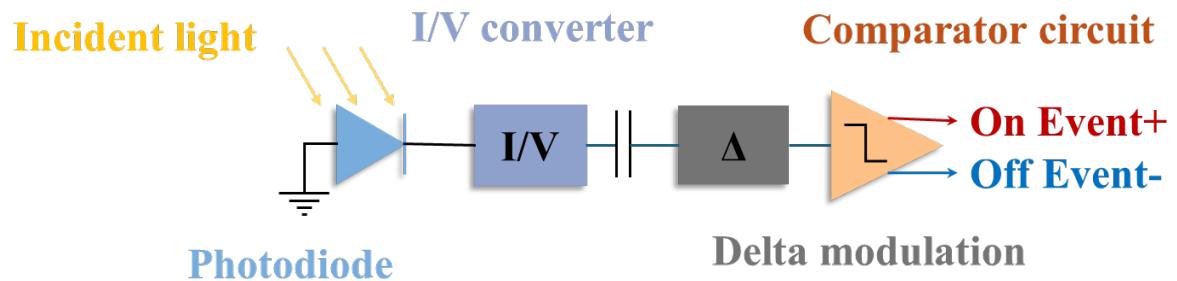
$$V(t) = V_{\text{in}}(1 - e^{-t/\tau}) + V(0)e^{-t/\tau}$$

Charge

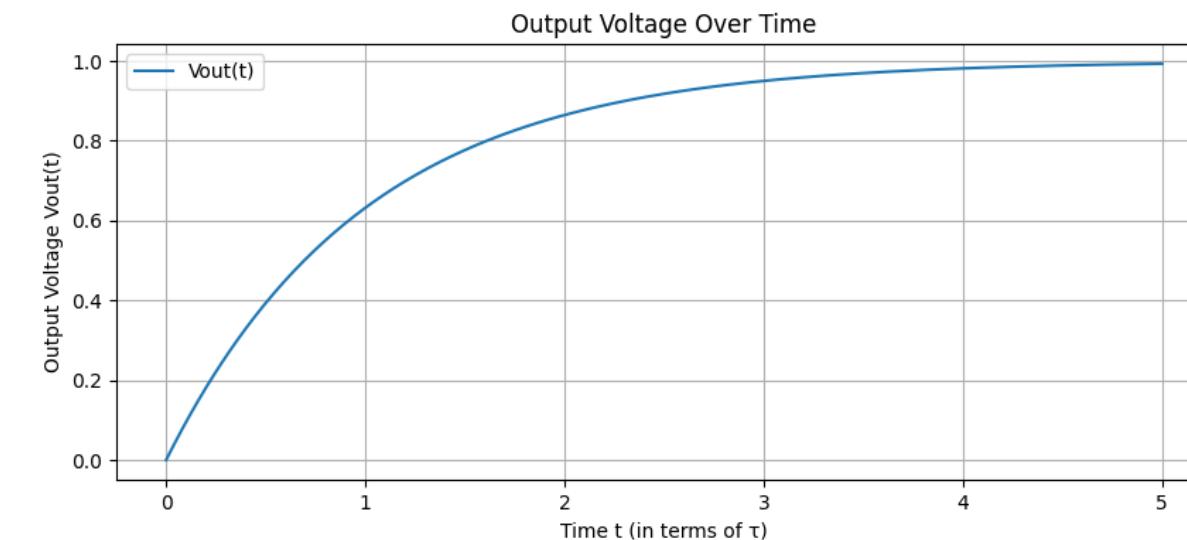
Discharge



Capacitor Charge/Discharge

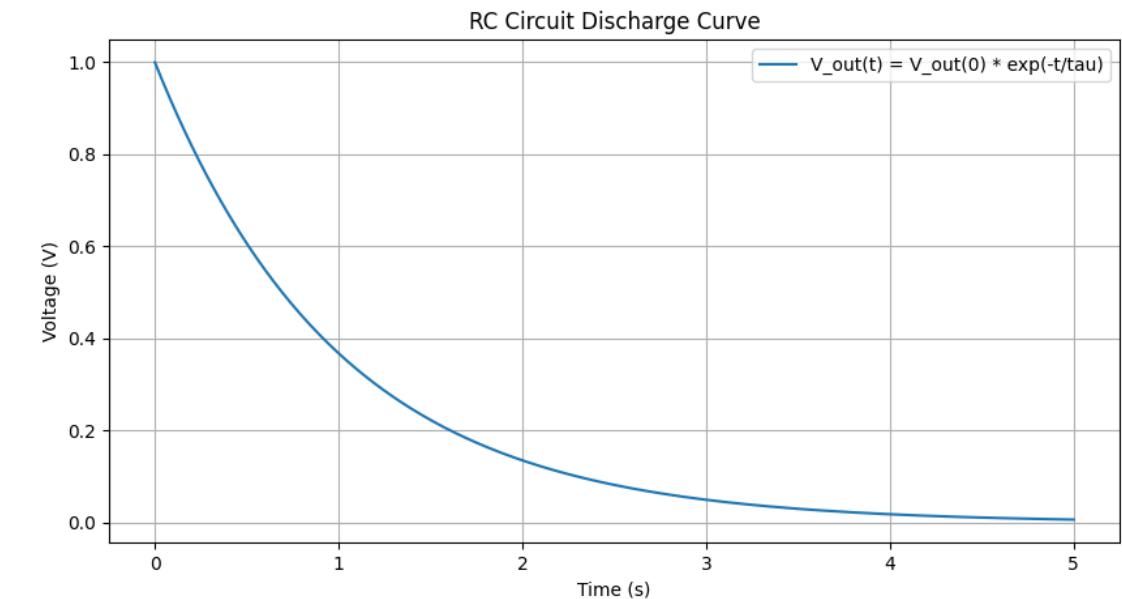


The charging process of the new signal



$$V(t) = V_{in}(1 - e^{-t/\tau})$$

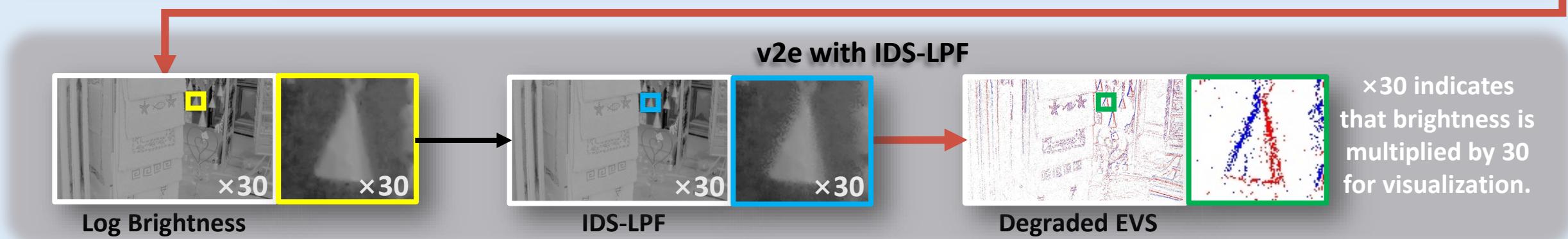
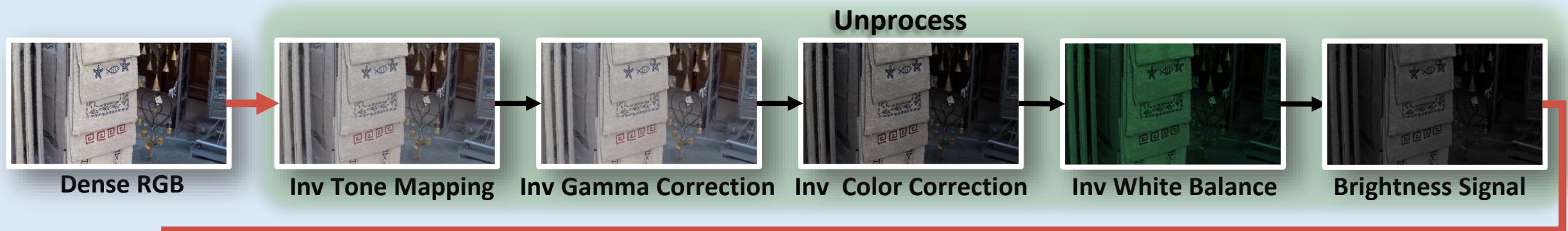
The discharging process of the residual signal



$$V(t) = V(0)e^{-t/\tau}$$



EVS Synthesis Pipeline





Model for Event Trailing Artifacts

- First order low-pass filter in v2e simulator^[1]

The diagram shows the mathematical expression for a first-order low-pass filter:

$$I'_t = \alpha I_t + (1 - \alpha) I_{t - \Delta t}$$

Annotations above the equation identify three components:

- Current signal**: Points to the term αI_t .
- New signal**: Points to the term $I_{t - \Delta t}$.
- Cached signal**: Points to the term $(1 - \alpha) I_{t - \Delta t}$.

Below the equation, two red boxes highlight the terms αI_t and $(1 - \alpha) I_{t - \Delta t}$. Below each highlighted term is the text "Capacitor Charging" in red.

[1] Hu, Y., Liu, S. C., & Delbrück, T. (2021). v2e: From video frames to realistic DVS events. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1312-1321).

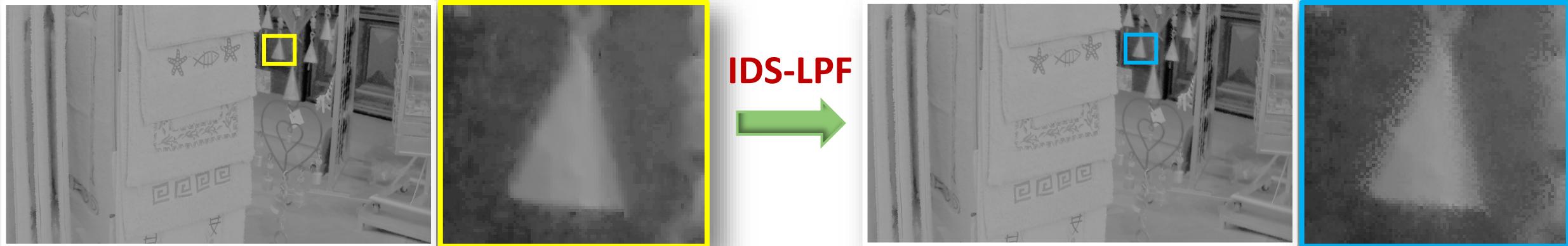


Model for Event Trailing Artifacts

- **IDS-LPF (Intensity-Dependent Stochastic Low-Pass Filter)**

$$I'_t = R(\alpha, p)I_t + (1 - R(\alpha, p))I_{t-\Delta t}$$

Exponential decay + Pixelwise random mask





Model for Event Trailing Artifacts

- Exponential decay + Pixelwise random mask

p
A random number between 0 and 1

$$\alpha = 1 - e^{-t/\tau}$$



$$R(\alpha, p) = \begin{cases} 1 - e^{-\Delta t/\tau} & \text{if } p > 0.5 \\ 1 & \text{if } p \leq 0.5 \end{cases}$$

Time constant

$$\tau = \frac{1}{2\pi f_c}$$

Cutoff frequency

$$f_c = \frac{1}{2\pi C} \cdot \frac{I_{ph}}{U_t}$$

Photocurrent

Photocurrent
(Light intensity)

Time constant

Capacitor charge/discharge

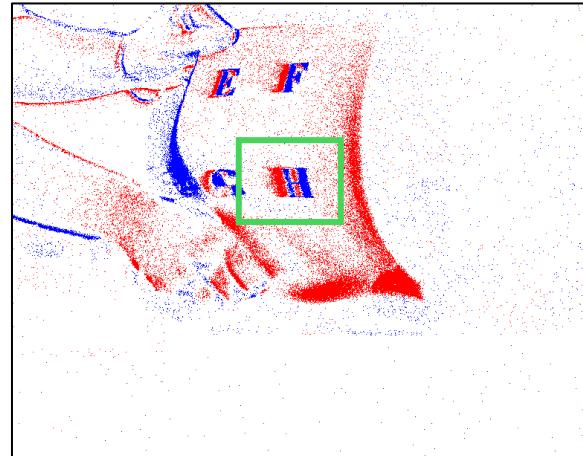




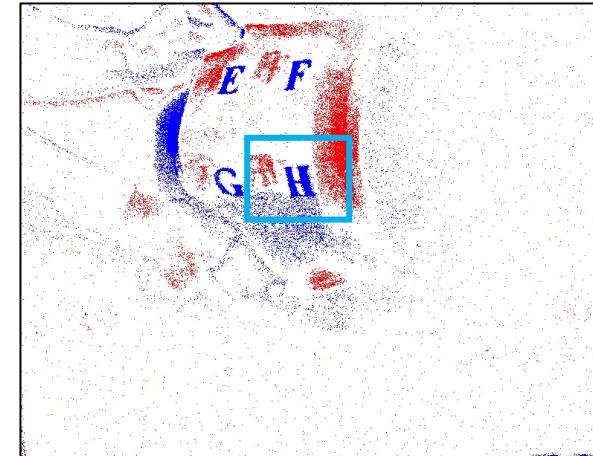
Simulation Comparison



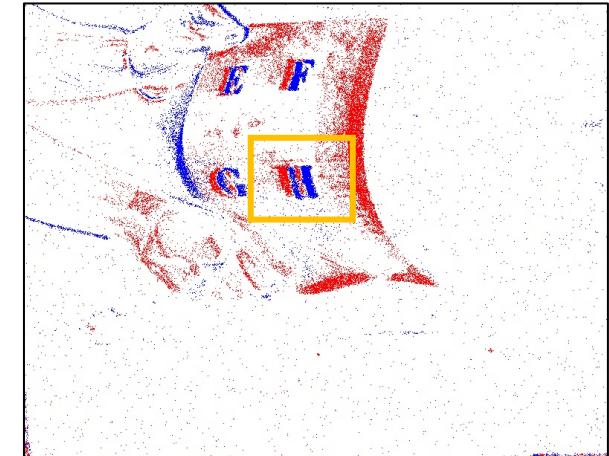
(a) Real RGB



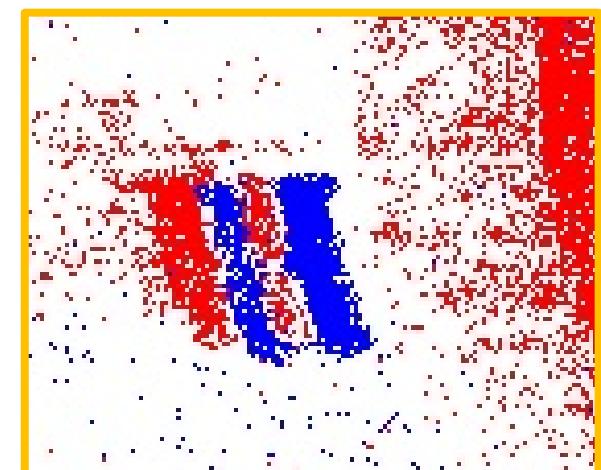
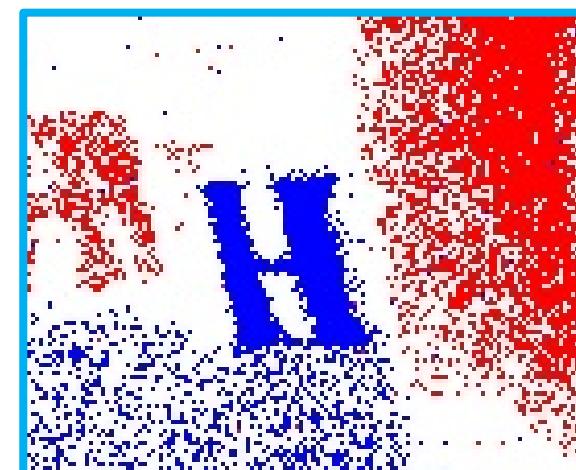
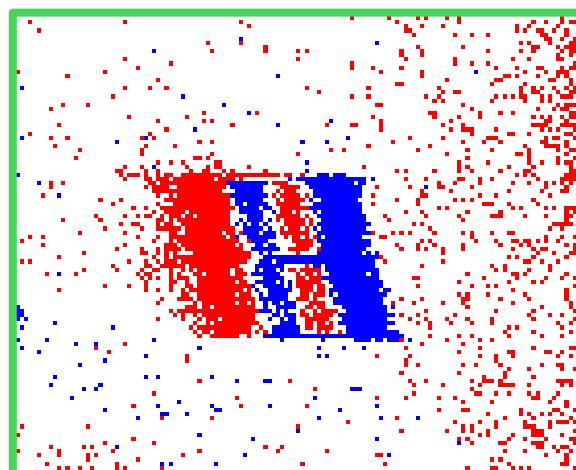
(b) Real Events



(c) v2e simulator



(d) Our simulator

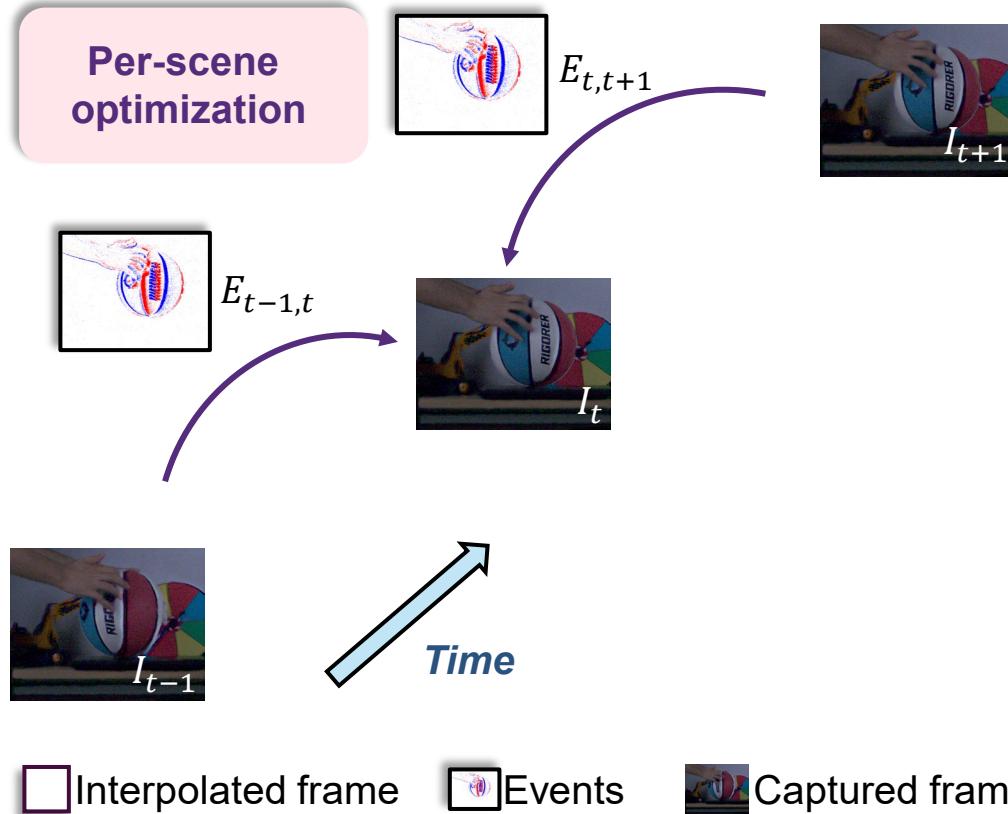




Per-scene Optimization



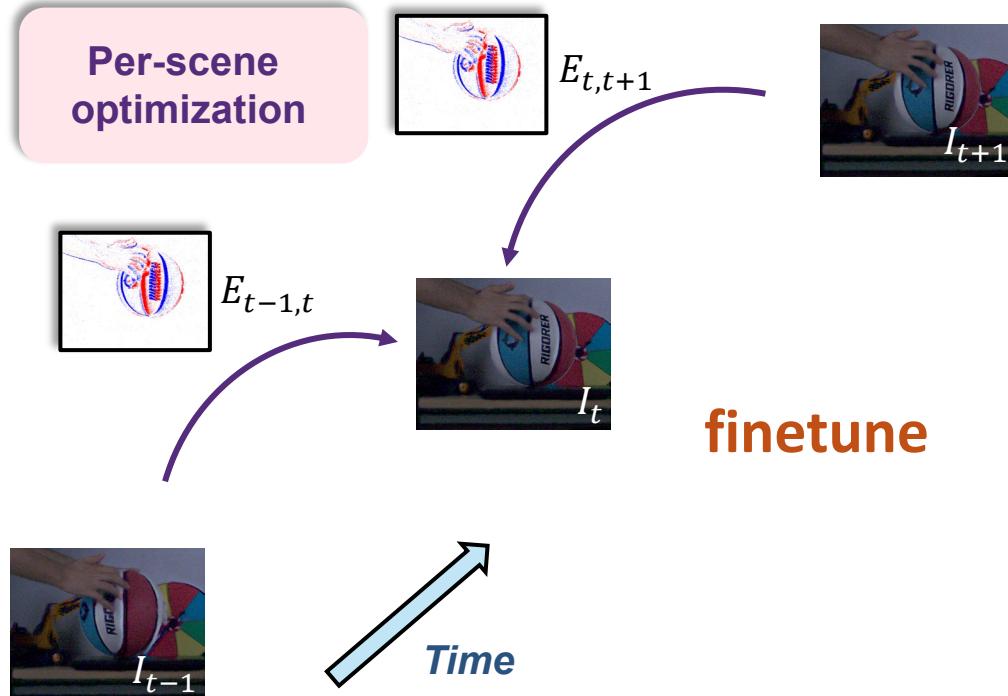
Per-scene Optimization



Noise and trailing artifacts are relatively constant in a single video.



Per-scene Optimization



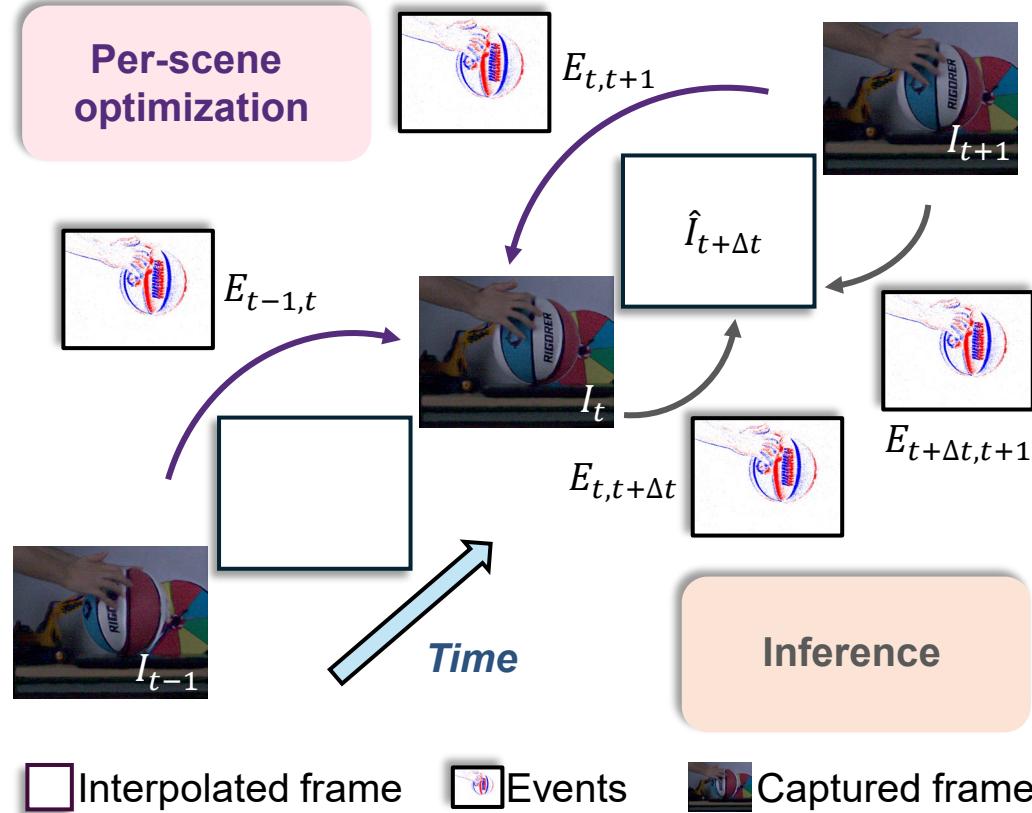
Interpolated frame

Events

Captured frame



Per-scene Optimization



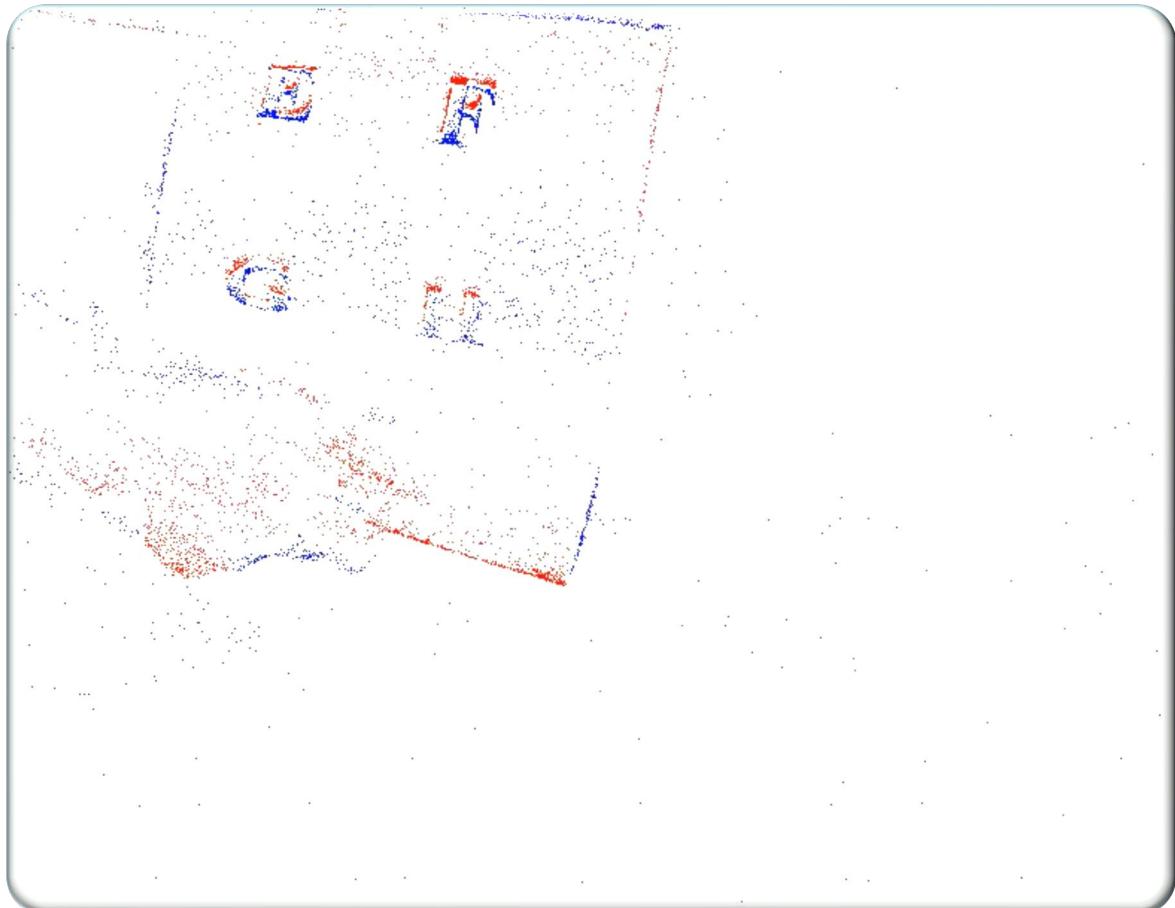
Per-scene optimization allows the pretrained model to generalize under different environments and settings.



Dataset : EVFI-LL



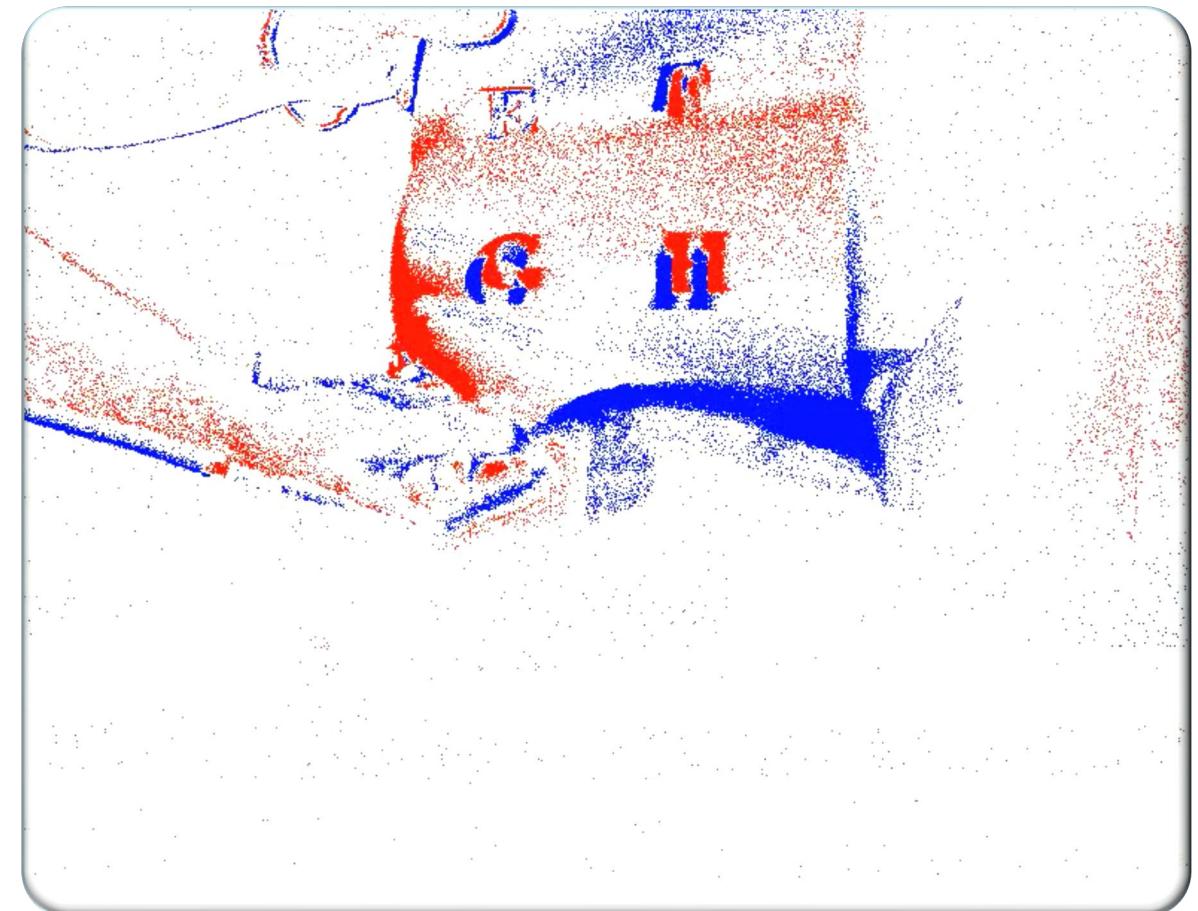
Nonlinear Low-Light Scenes
Event Camera Bias Settings +20 | 0 | -20
Lower bias → More events recorded





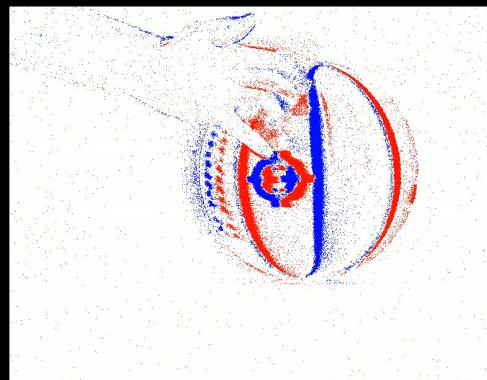
Dataset : EVFI-LL-C

Challenging track for large motion

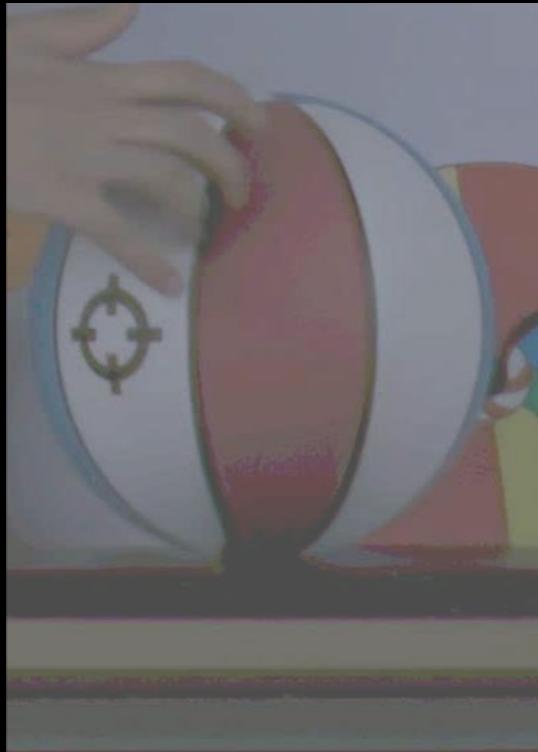




Method comparison



Event +4FPS
Video



RIFE
ECCV2022



CBMNet
CVPR2023



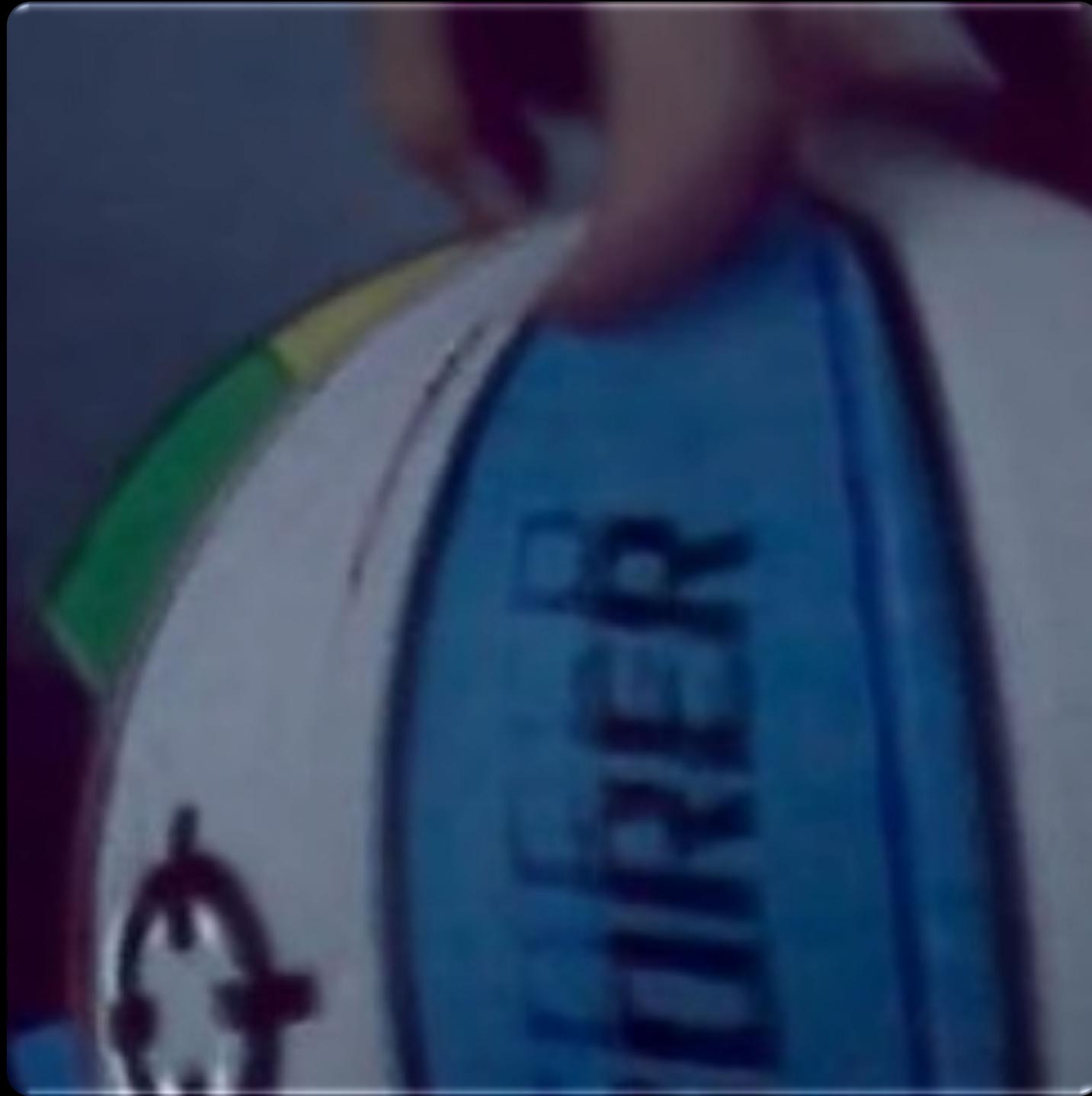
Ours

32FPS Video



Ablation Result

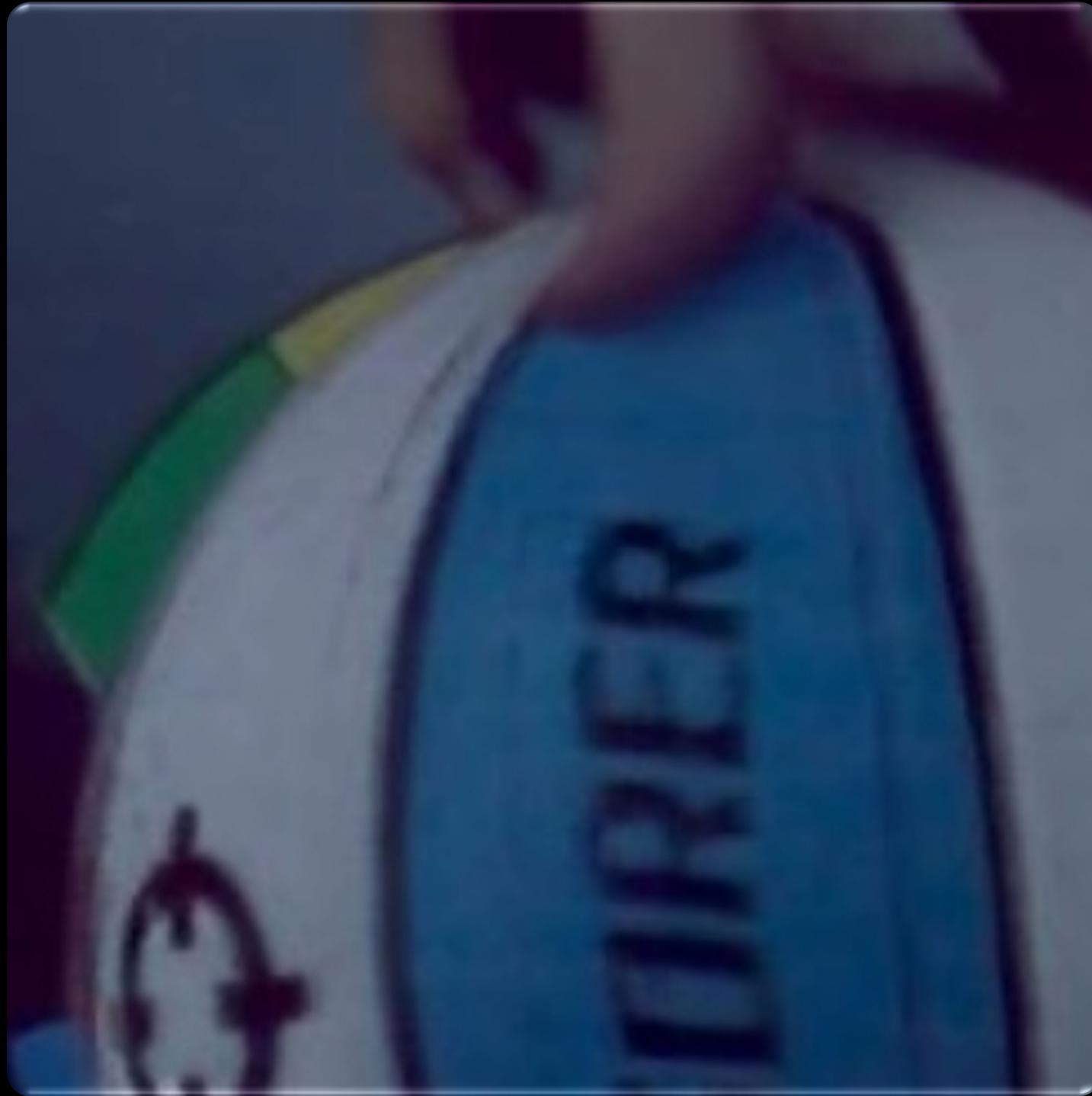
**Pretrained with
v2e simulator**





Ablation Result

**Pretrained with
our simulator**



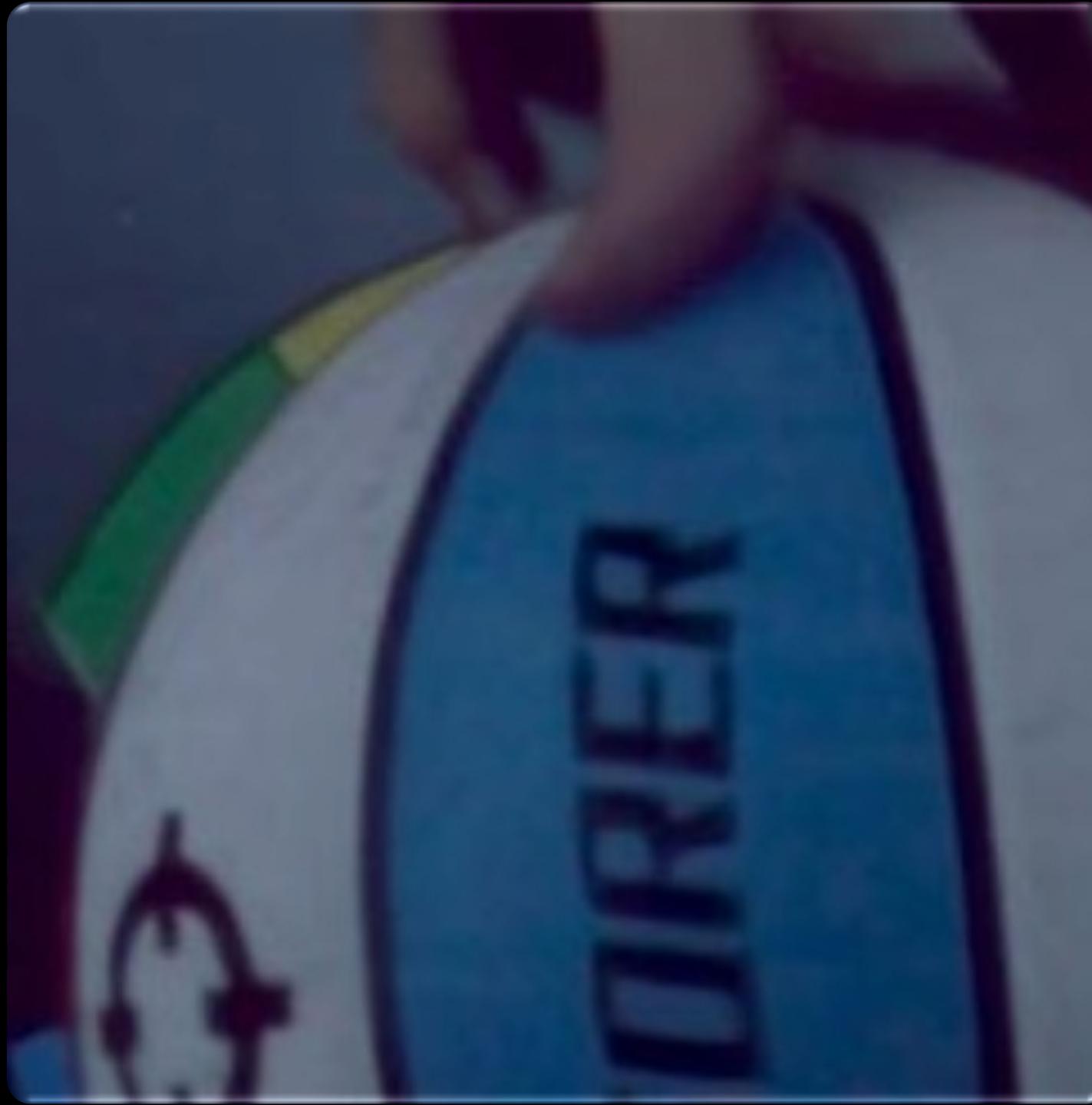


Ablation Result

Pretrained with
our simulator

+

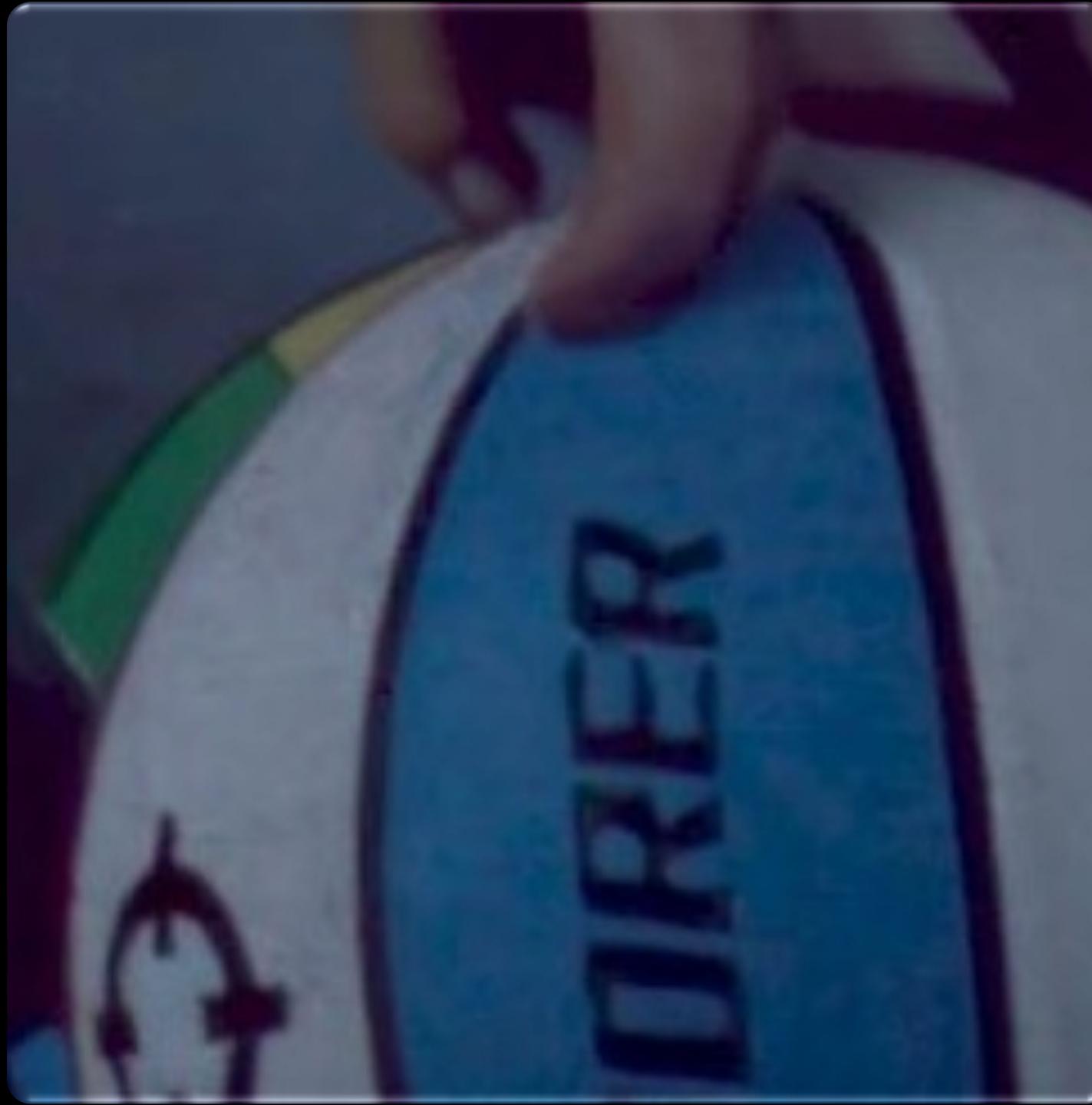
Per-scene
Optimization





Ablation Result

Ground truth



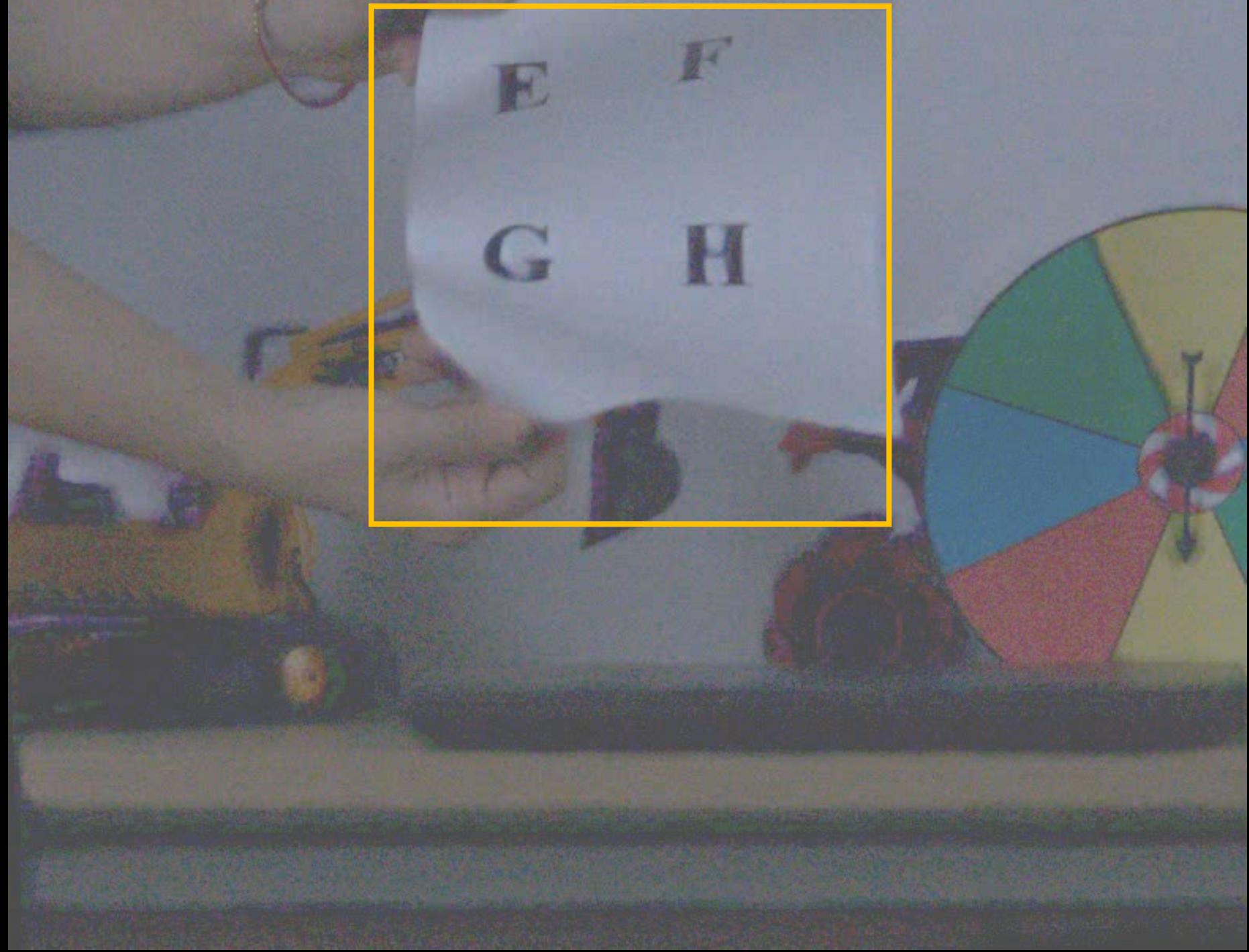


Video Results

Input video

&

Input events





Video Results

Output video



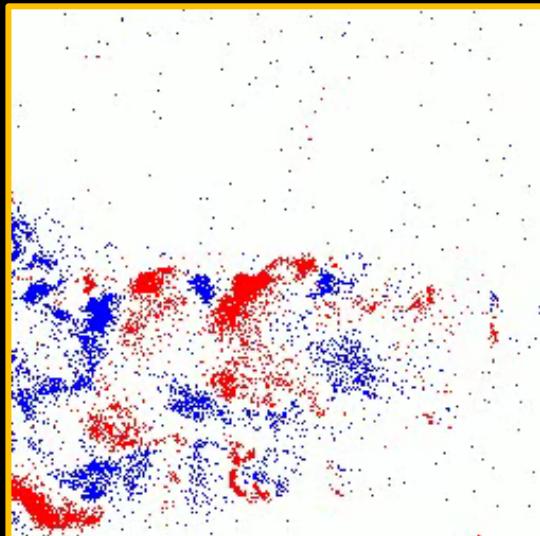


Video Results

Input video

&

Input events





Video Results

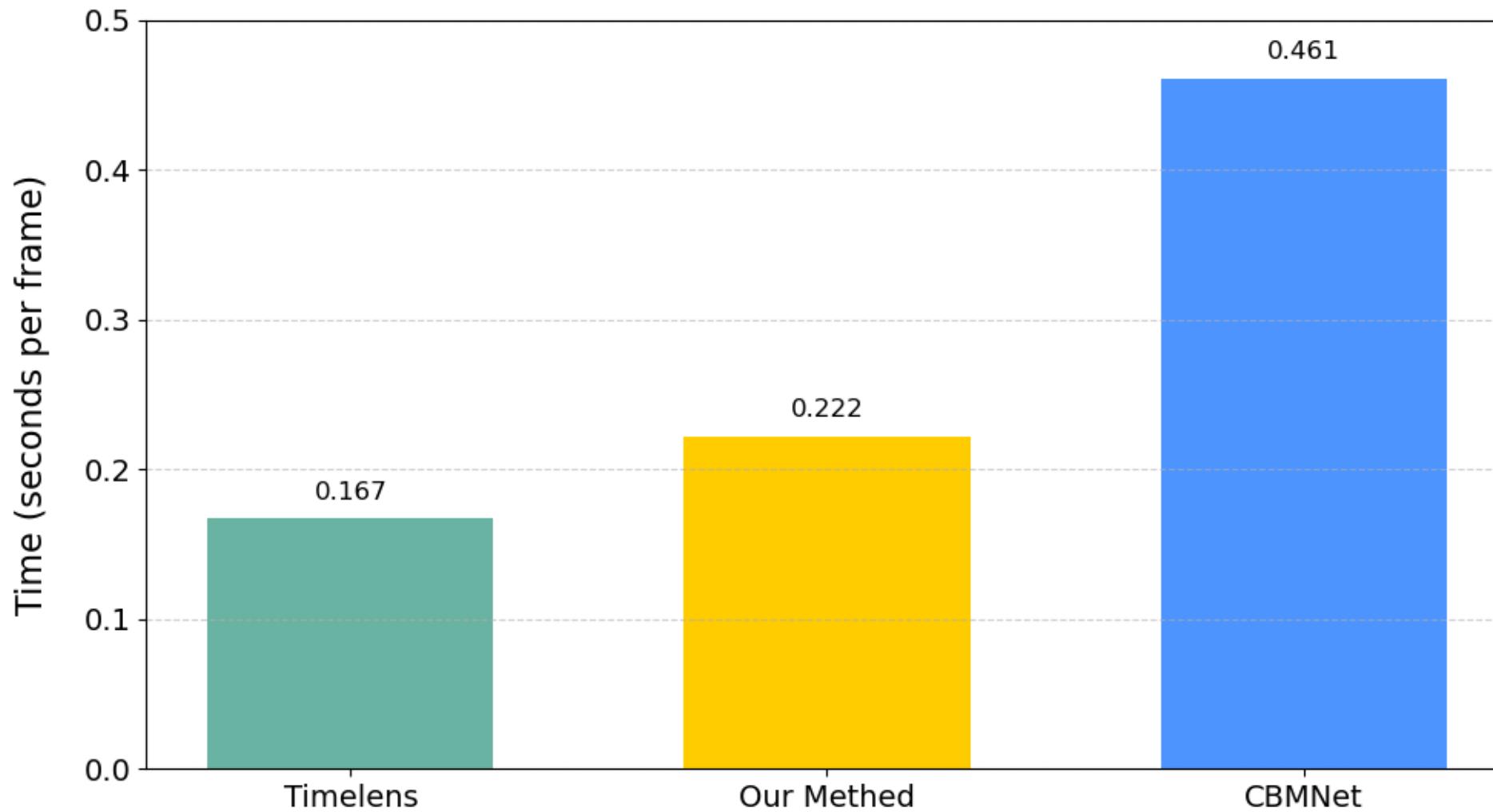
Output video





Inference Time

Comparison of Time per Frame for Different Methods.





Thanks!

naturezhanghn@zju.edu.cn

2024/12/05



香港中文大學
The Chinese University of Hong Kong