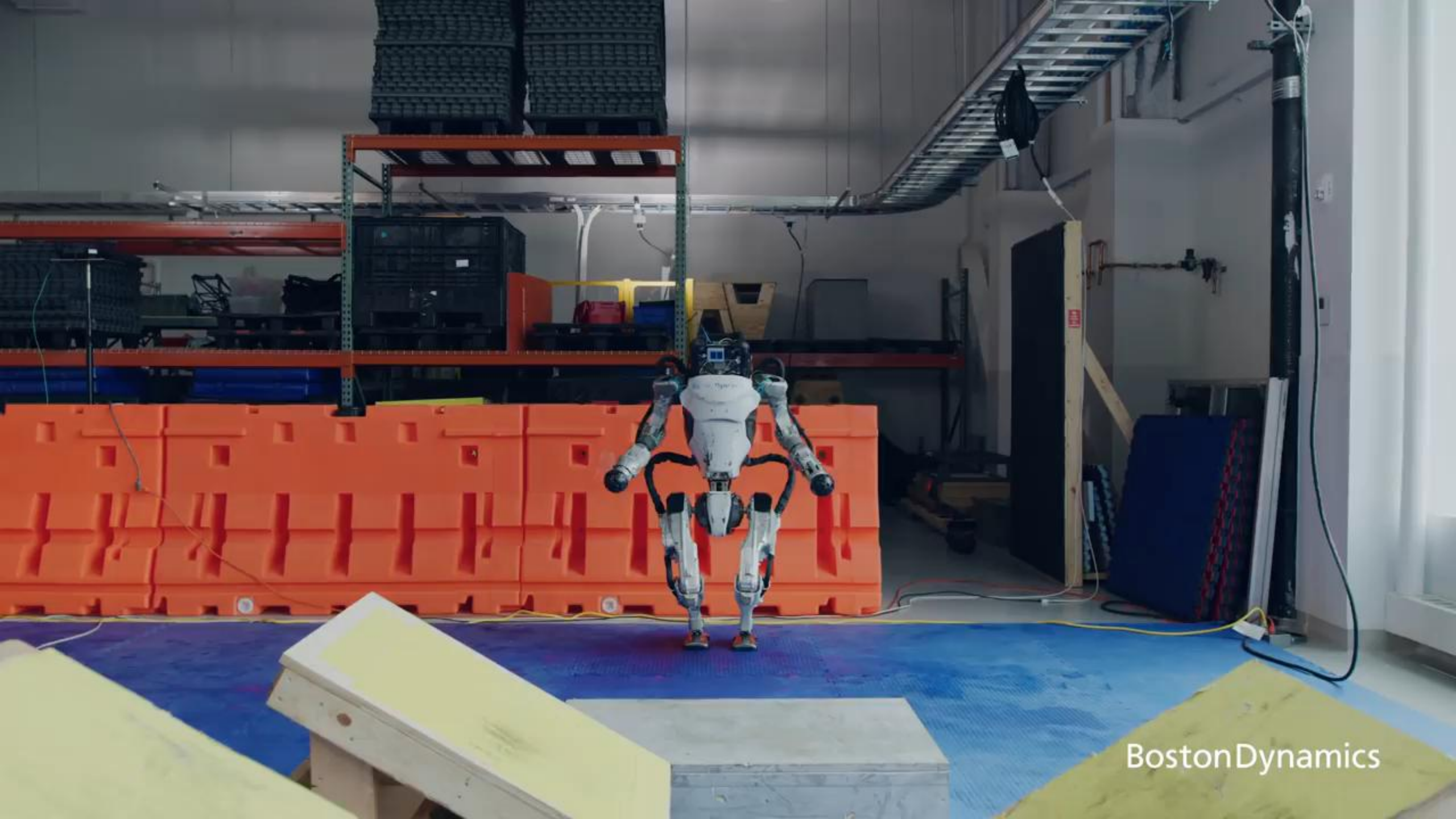
University of Zurich[UZH]

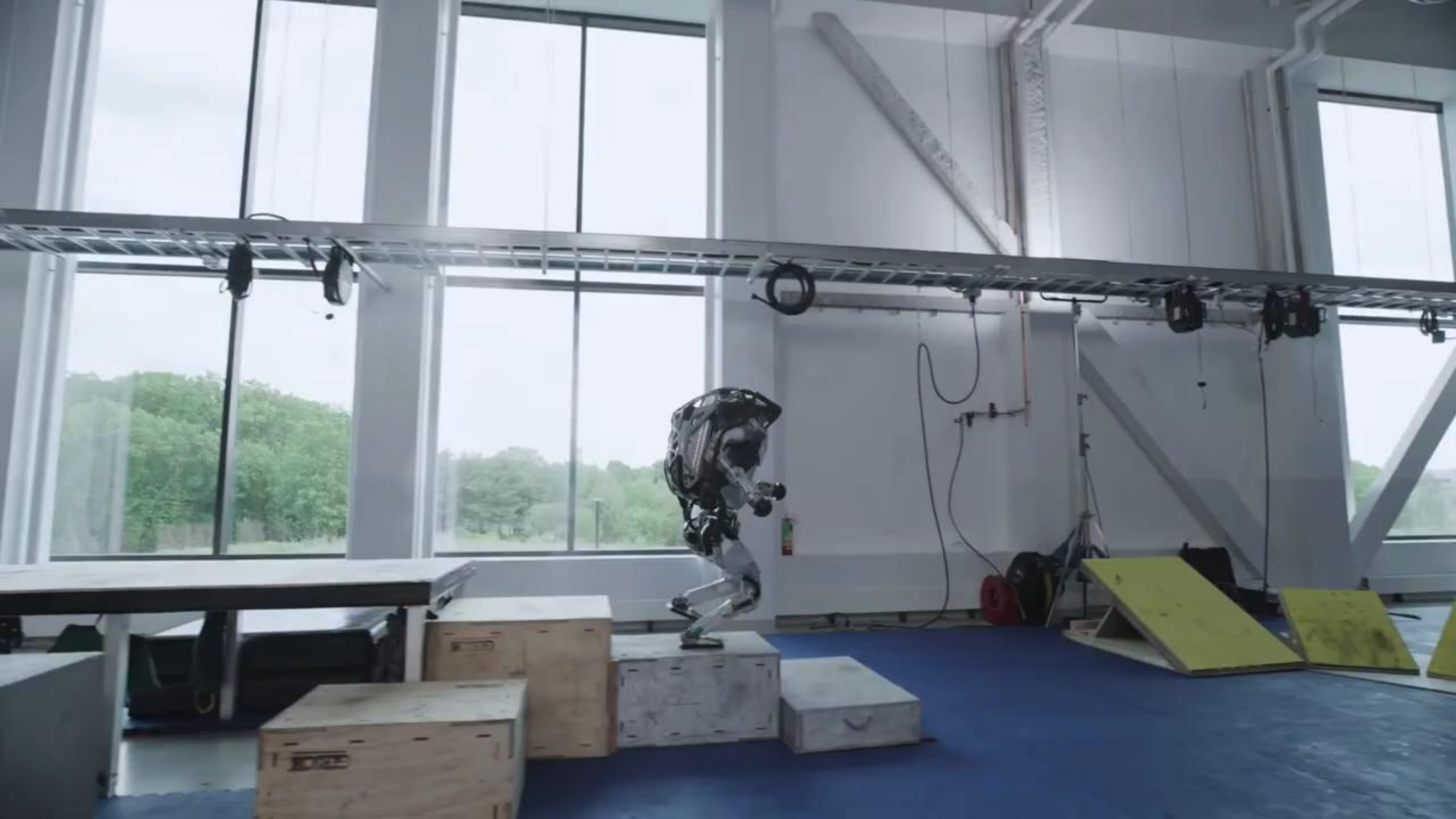
ROBOTICS & PERCEPTION GROUP

Event Cameras: a New Way of Sensing

Davide Scaramuzza

http://rpg.ifi.uzh.ch/

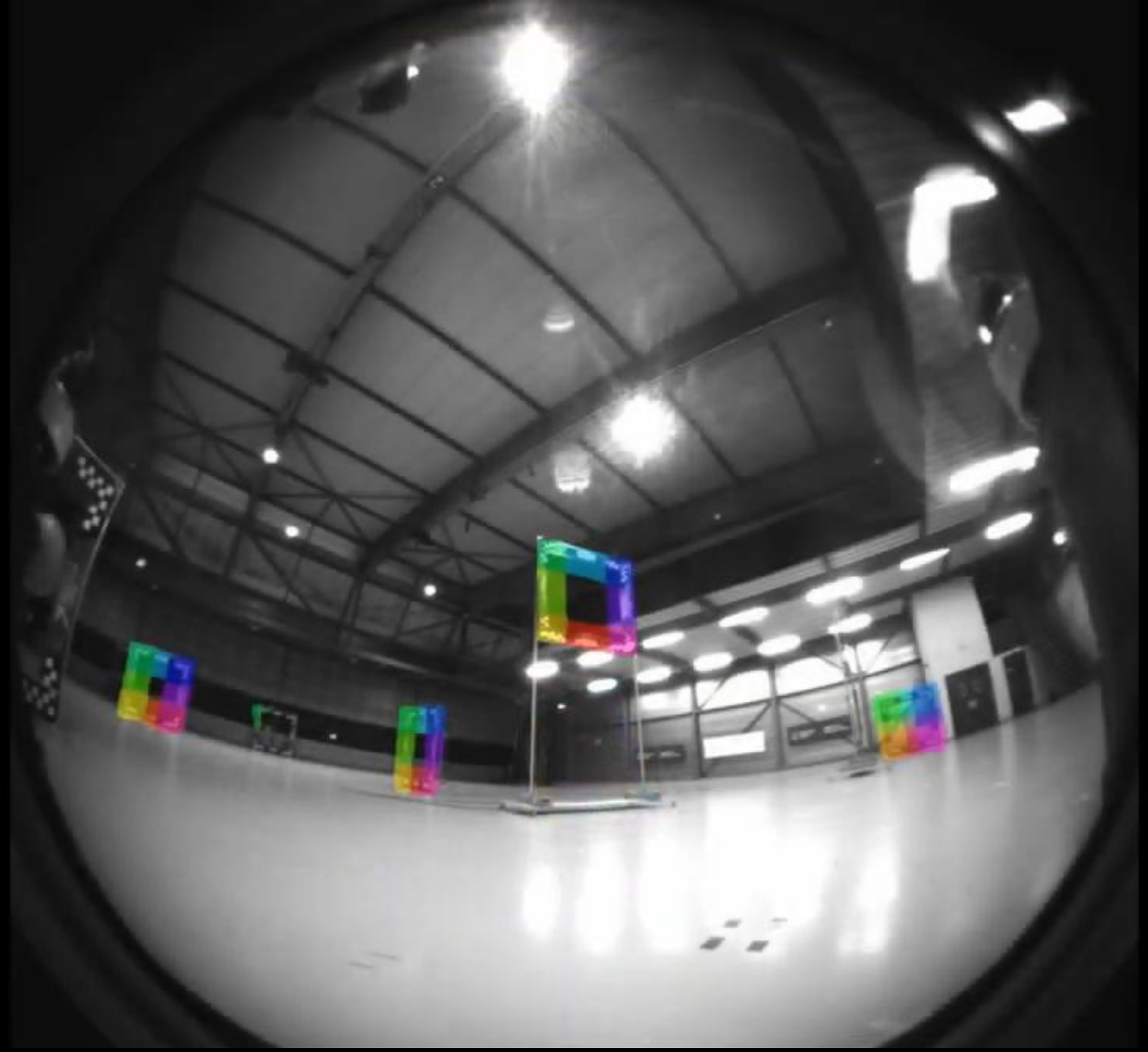

Boston Dynamics



Drone Racing

Autonomous Drone	World's Best Human Pilots
"Swift"	A. Vanover, T. Bitmatta, M. Schaepper

Kaufmann et al., *Champion-Level Drone Racing using Deep Reinforcement Learning*, **Nature, 2023**



Kaufmann et al., *Champion-Level Drone Racing using Deep Reinforcement Learning*, **Nature, 2023**


Universität Zürich
Robotics and Perception Group
intel robotics
ROBOTICS & PERCEPTION GROUP
UZH


Cook Shrimp
(autonomous)
3x speed
Stanford
University


Stanford
University

Open Challenges in Computer Vision

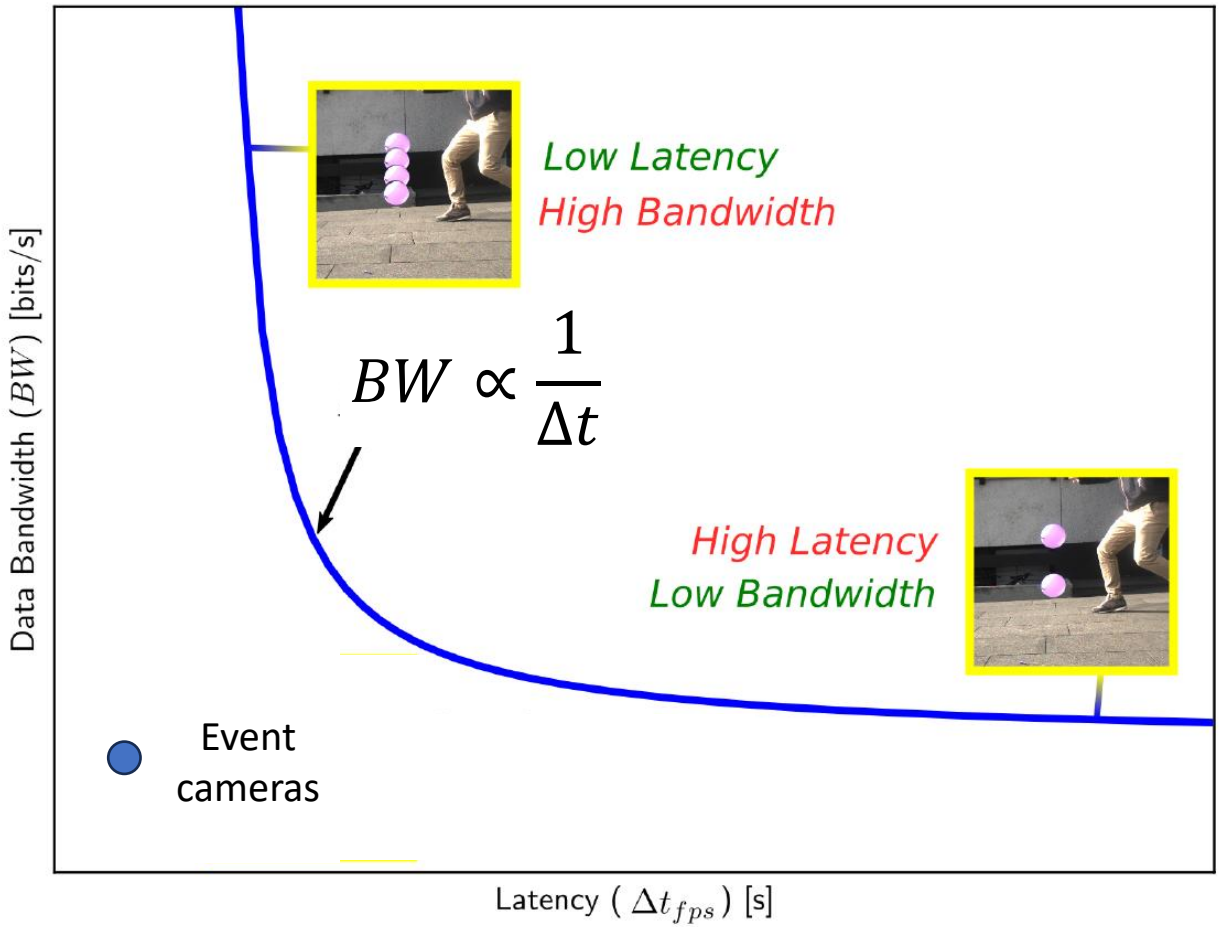
The past 60 years of research have been devoted to frame-based cameras but they are not good enough

Motion blur

Dynamic Range

Bandwidth-Latency tradeoff

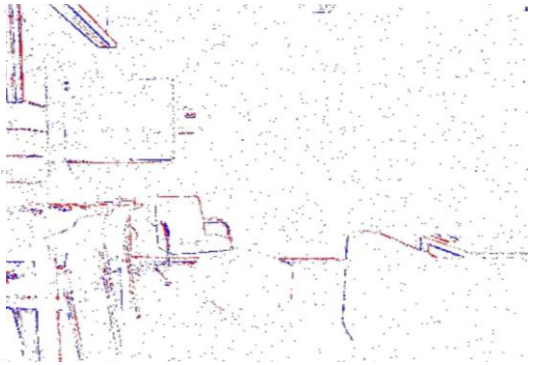
Standard cameras suffer from the bandwidth-latency tradeoff

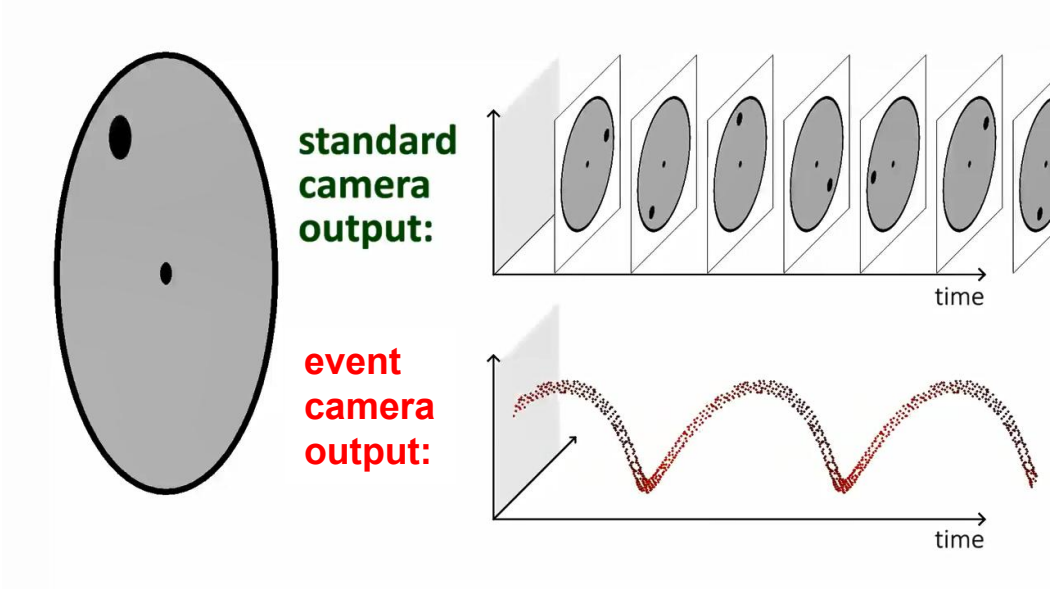


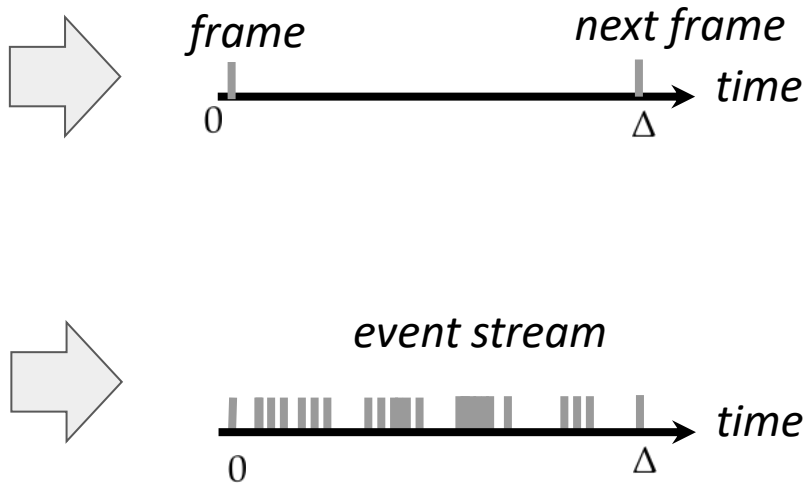
What is an Event Camera?

- It is camera that measures only **motion in the scene**
- **Key advantages:**
  1. Low-latency (~ 1 µs)
  2. Low bandwidth
  3. Negligible motion blur
  4. High dynamic range

Traditional vision algorithms cannot be directly applied!







[1] Lichtsteiner, Posch, Delbruck, *A 128x128 120 dB 15µs Latency Asynchronous Temporal Contrast Vision Sensor*, IEEE Journal of Solid-State Circuits, 2008
[2] Gallego et al., *Event-based Vision: A Survey*, **T-PAMI, 2020**

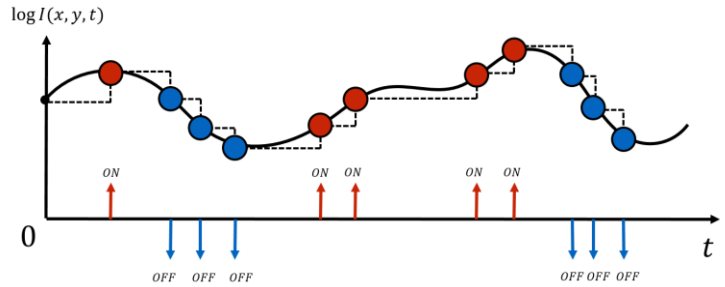
How do we apply event cameras to computer vision

without reconstructing the image?

Event-based Vision: Two Schools of Approaches

1. **Event-by-event** methods

$$\log I(x, y, t + \Delta t) - \log I(x, y, t) = \pm C$$



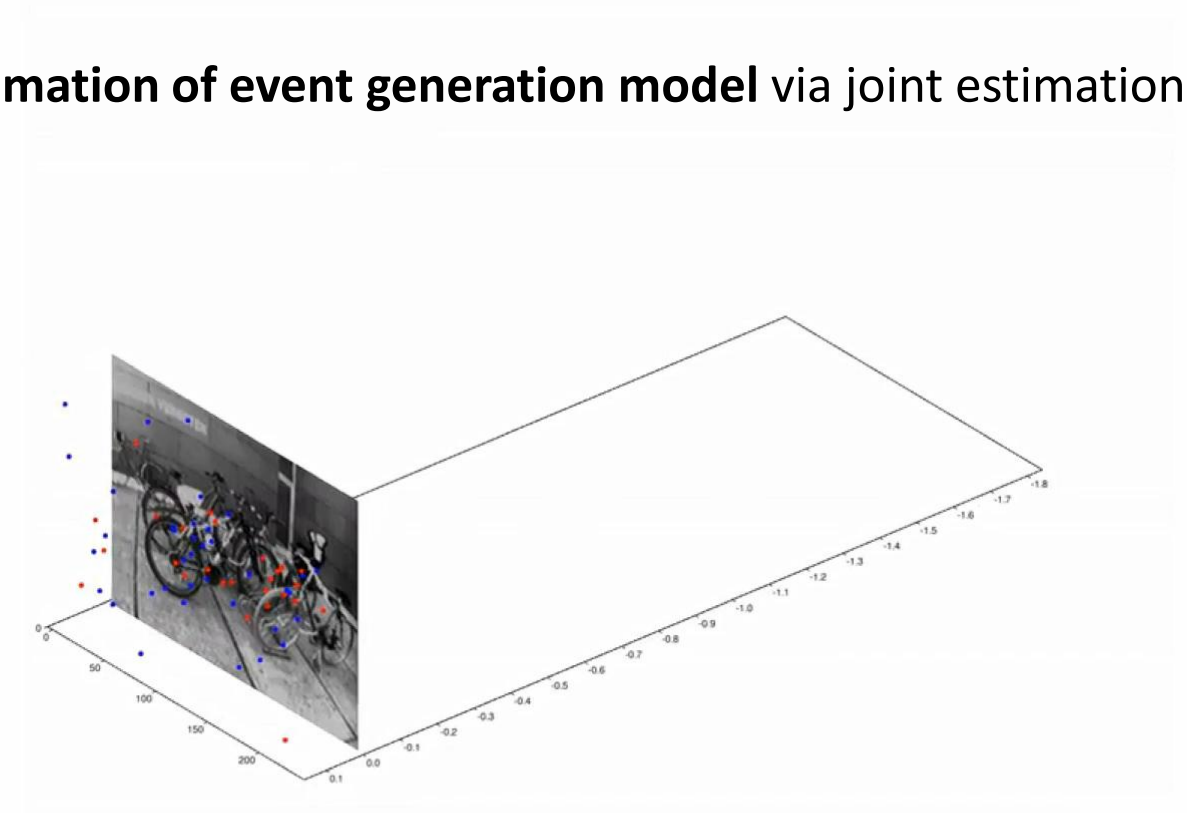
2. **Batch** methods (contrast maximization)

$$\boldsymbol{\theta} = argmax\ \sigma^2(I(\boldsymbol{x}; \boldsymbol{\theta}))$$

Gallego et al., *Event-based Vision: A Survey*, **T-PAMI, 2020**

Application 1: Event-based Feature Tracking

- **Goal**: Extract **features from standard frames** and track them using only **events** in the **blind time between two frames**
- Uses the **1st order approximation of event generation model** via joint estimation of patch warping and optic flow

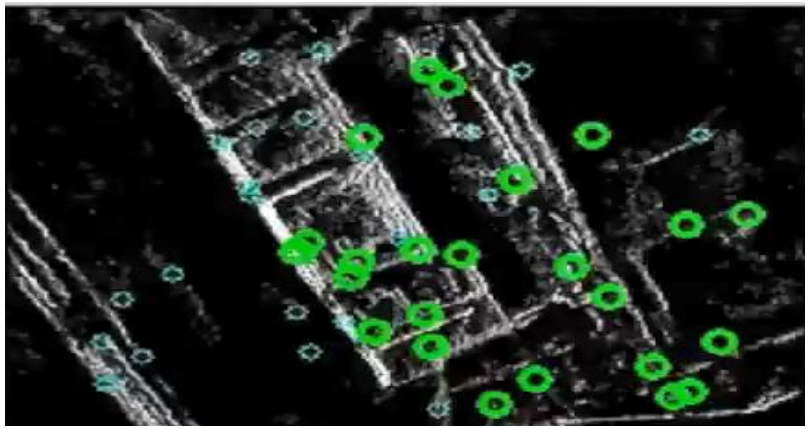


Gehrig, Rebecq, Gallego, Scaramuzza, *EKLT: Asynchronous, Photometric Feature Tracking using Events and Frames*, International Journal of Computer Vision (IJCV), 2019. PDF. Video. Code

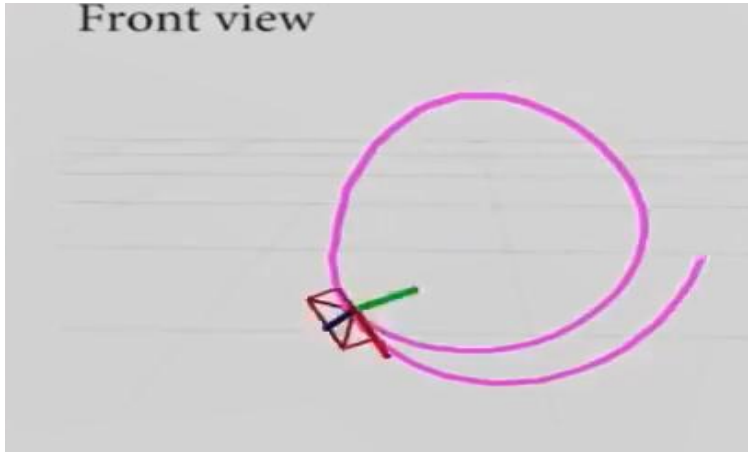
Application 2: “Ultimate SLAM”

Standard camera

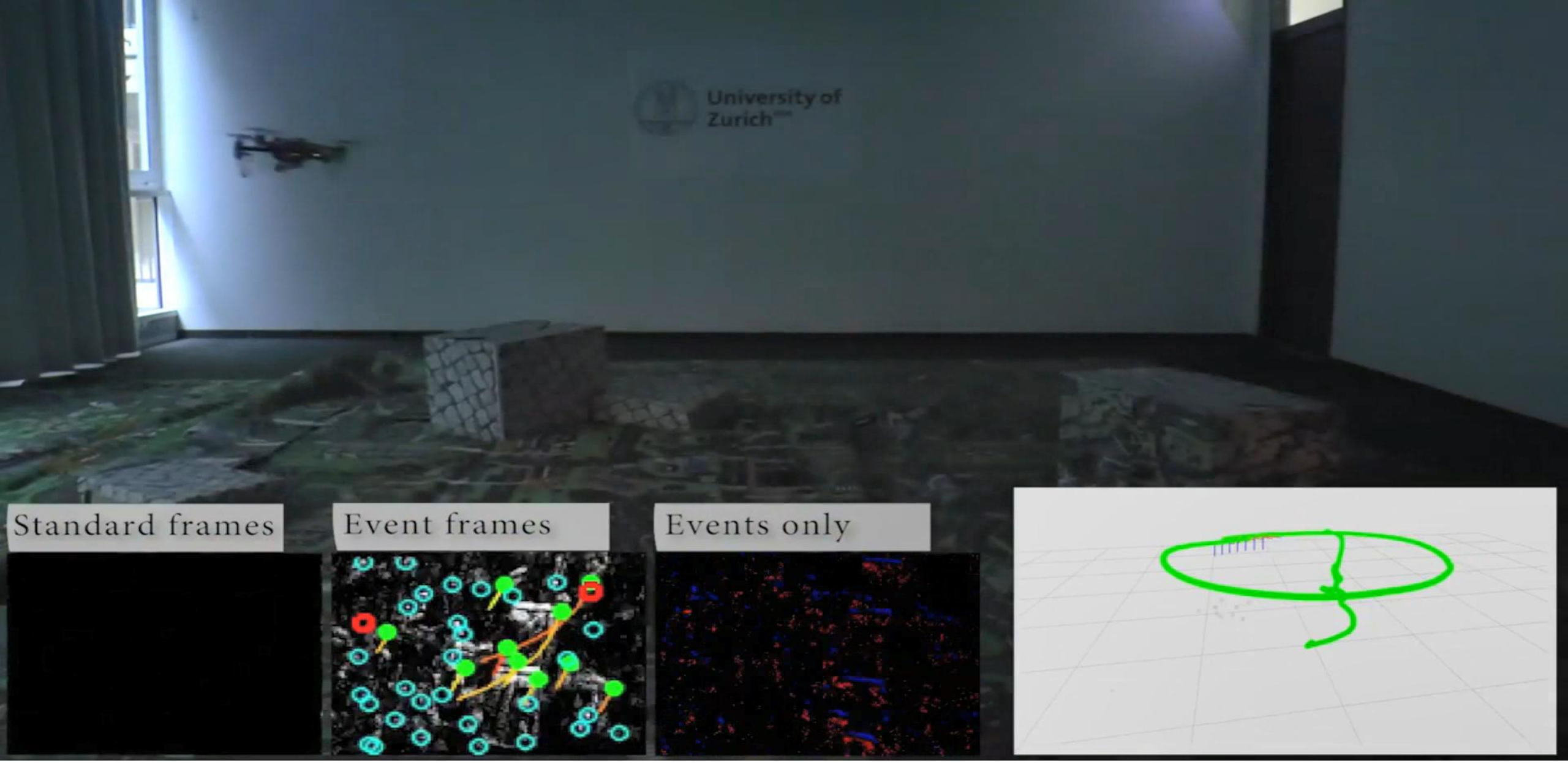
Event camera



Estimated trajectory



- Rosinol et al., *Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM,* **RAL’18 Best Paper Award Hon. Mention**
- Pellerito, Cannici, Gehrig, Belhadj, Dubois-Matra, Casasco, Scaramuzza, *Deep Visual Odometry with Events and Frames*, **IROS’24**
- Hidalgo-Carrió, Gallego, Scaramuzza, *Event-aided Direct Sparse Odometry*, **CVPR’22**, Oral



- Rosinol et al., *Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM,* **RAL'18 Best Paper Award Hon. Mention**
- Pellerito, Cannici, Gehrig, Belhadj, Dubois-Matra, Casasco, Scaramuzza, *Deep Visual Odometry with Events and Frames*, **IROS'24**
- Hidalgo-Carrió, Gallego, Scaramuzza, *Event-aided Direct Sparse Odometry*, **CVPR'22**, Oral

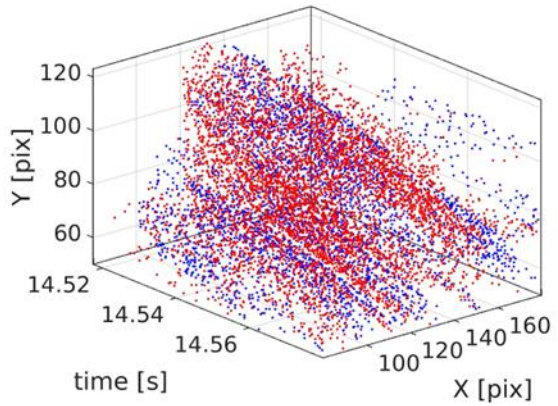
Event-based Vision: Two Schools of Approaches

1. **Event-by-event** methods

$$\log I(x, y, t + \Delta t) - \log I(x, y, t) = \pm C$$

2. **Batch** methods (contrast maximization)

$$\boldsymbol{\theta} = argmax\ \sigma^2(I(\boldsymbol{x}; \boldsymbol{\theta}))$$

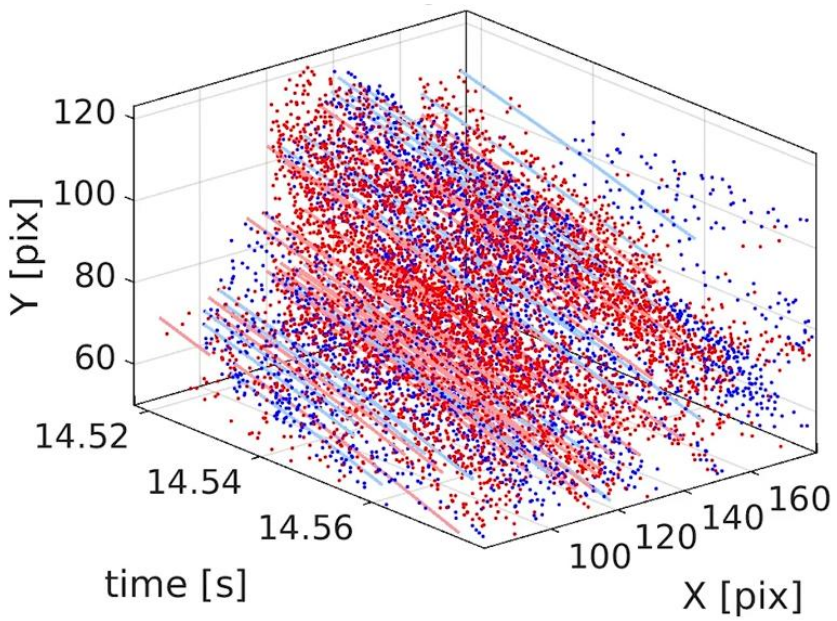


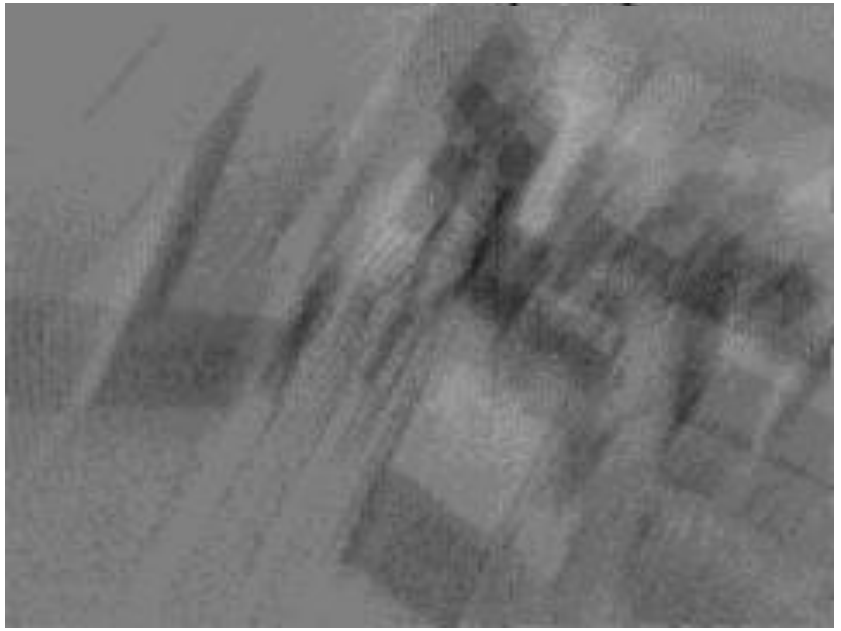
Gallego et al., *Event-based Vision: A Survey*, **T-PAMI, 2020**

Gallego, Rebecq, Scaramuzza, *A Unifying Contrast Maximization Framework for Event Cameras*, **CVPR'18**

Contrast Maximization Framework

Idea: Warp spatio-temporal volume of events to **maximize contrast** (e.g., sharpness) of the resulting image



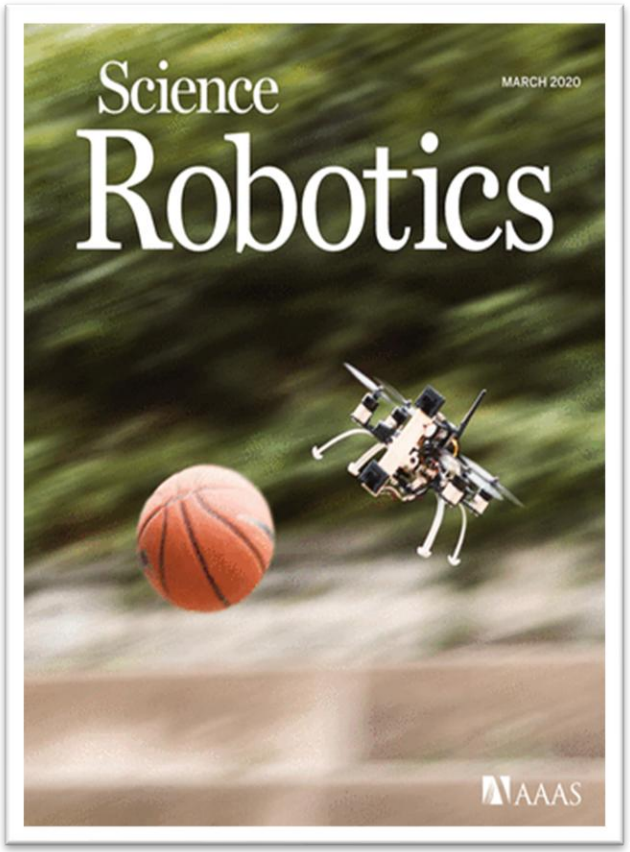


Aggregated image
without motion correction

Gallego, Rebecq, Scaramuzza, *A Unifying Contrast Maximization Framework for Event Cameras*, CVPR18, PDF, Video
Gallego, Gehrig, Scaramuzza, *Focus Is All You Need: Loss Functions for Event-based Vision*, CVPR19, PDF.

Application 2: Dodging Dynamic Objects

- Perception latency: **3.5 ms**
- Works with relative speeds of up to **10 m/s**



Falanga, Kleber, Scaramuzza, *Dynamic Obstacle Avoidance for Quadrotors with Event Cameras*, **Science Robotics,** 2020. PDF. Video

Application 3: High-Speed Inspection of Countersinks

Salah et al, Zweiri, *High speed neuromorphic vision-based inspection of countersinks in automated manufacturing processes*, **Journal of Intelligent Manufacturing**

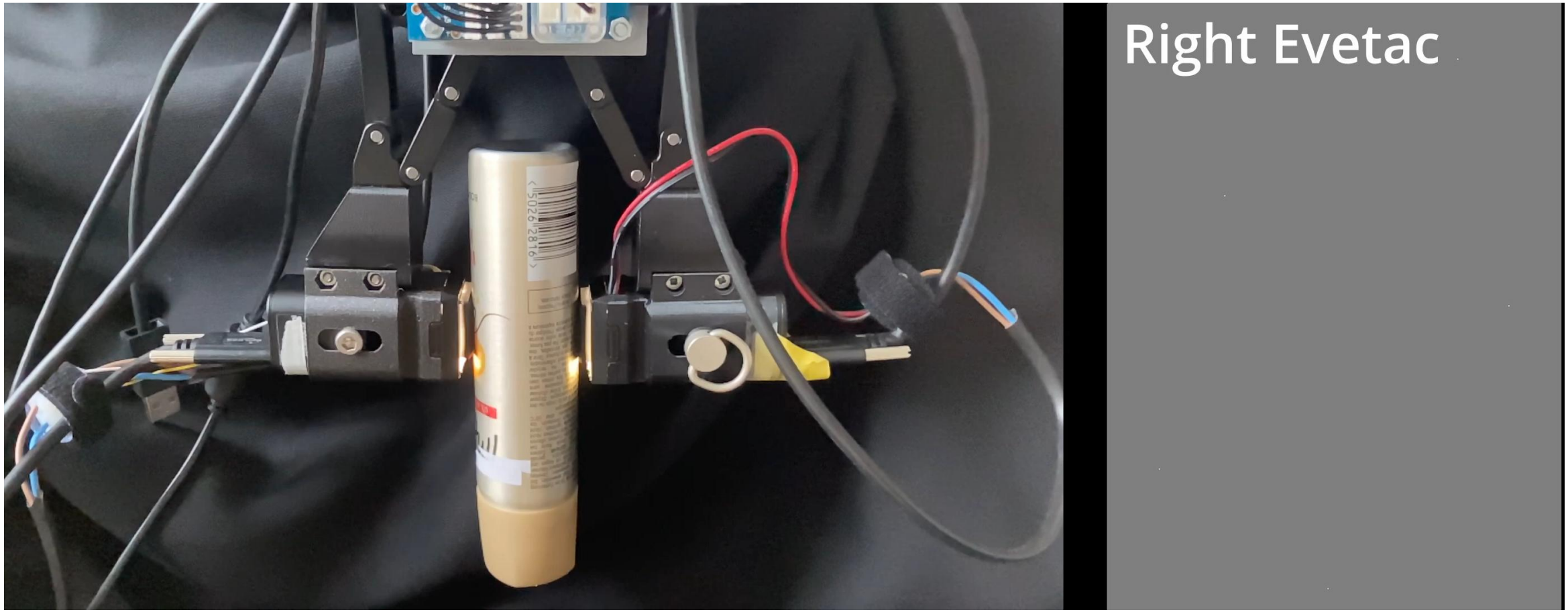
Event-based Optical Tactile Sensing

- Perception latency: **1 ms**
- 640x480 pixels, thereby offering both, human-like temporal and spatial resolution

Funk, Helmut, Chalvatzaki, Calandra, Peters, *Evetac: An Event-based Optical Tactile Sensor for Robotic Manipulation*, TRO'24

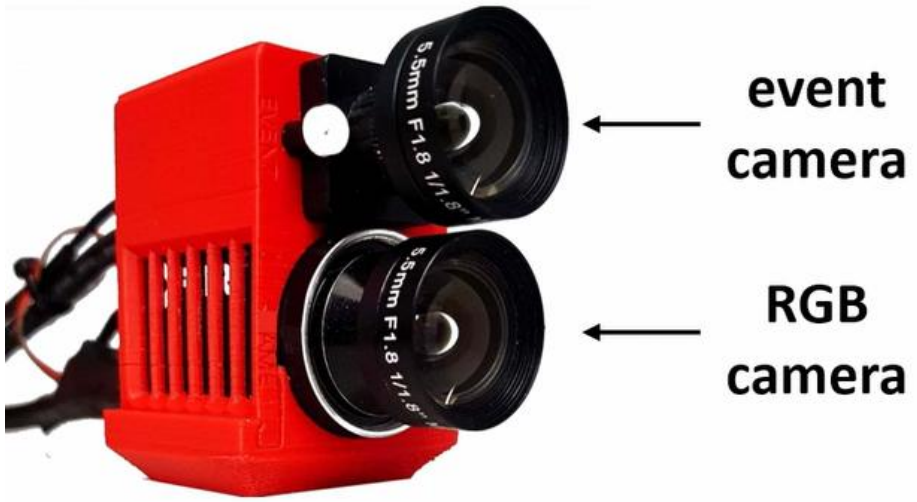
Event-based Optical Tactile Sensing

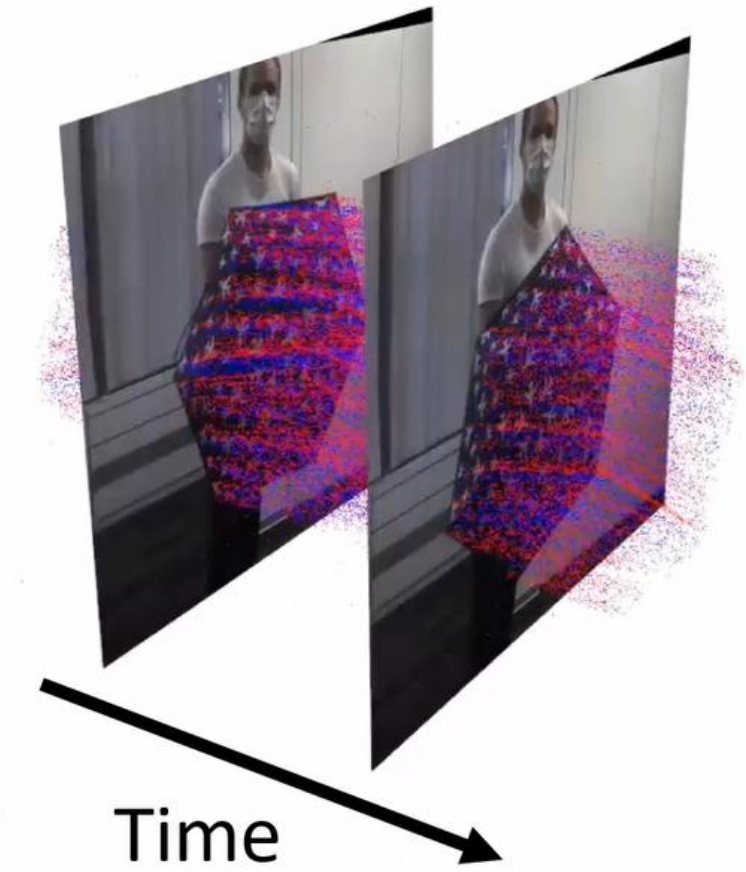
- Perception latency: **1 ms**
- 640x480 pixels, thereby offering both, human-like temporal and spatial resolution



Funk, Helmut, Chalvatzaki, Calandra, Peters, Evetac: An Event-based Optical Tactile Sensor for Robotic Manipulation, TRO'24

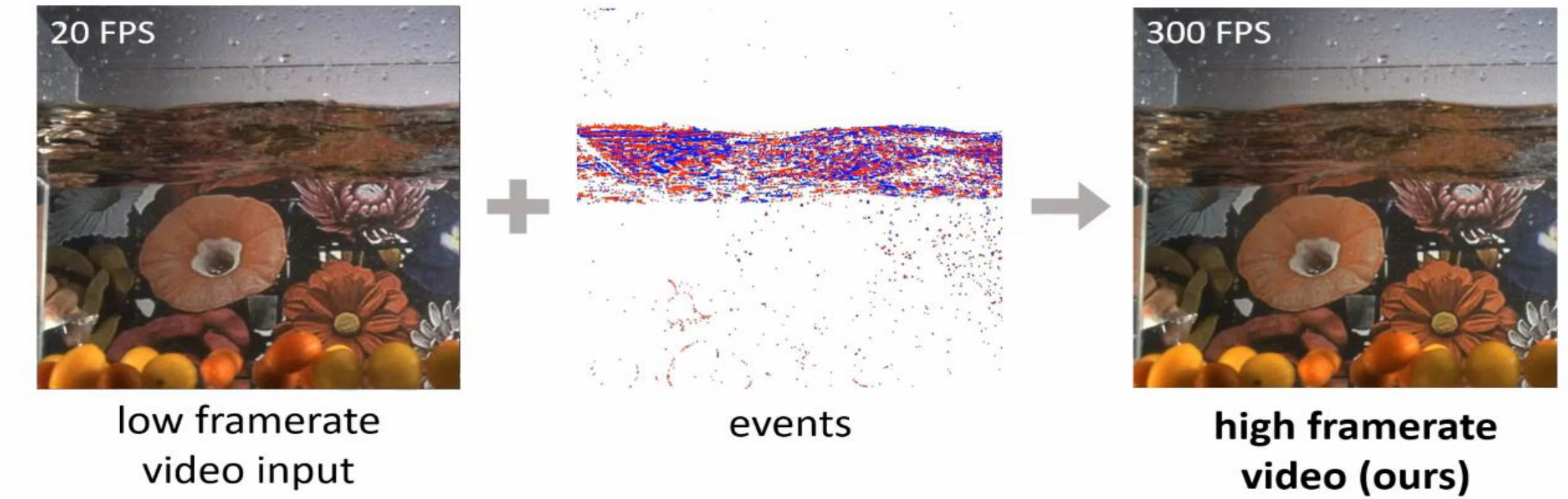
Combining Events and Frames for Ultimate Performance





Application 1: Slow Motion Video

- We can combine an event camera with an HD RG camera
- We use events to **upsample low-framerate video** by over **50 times** with **only 1/40th of the memory** footprint!



Code & Datasets: http://rpg.ifi.uzh.ch/timelens

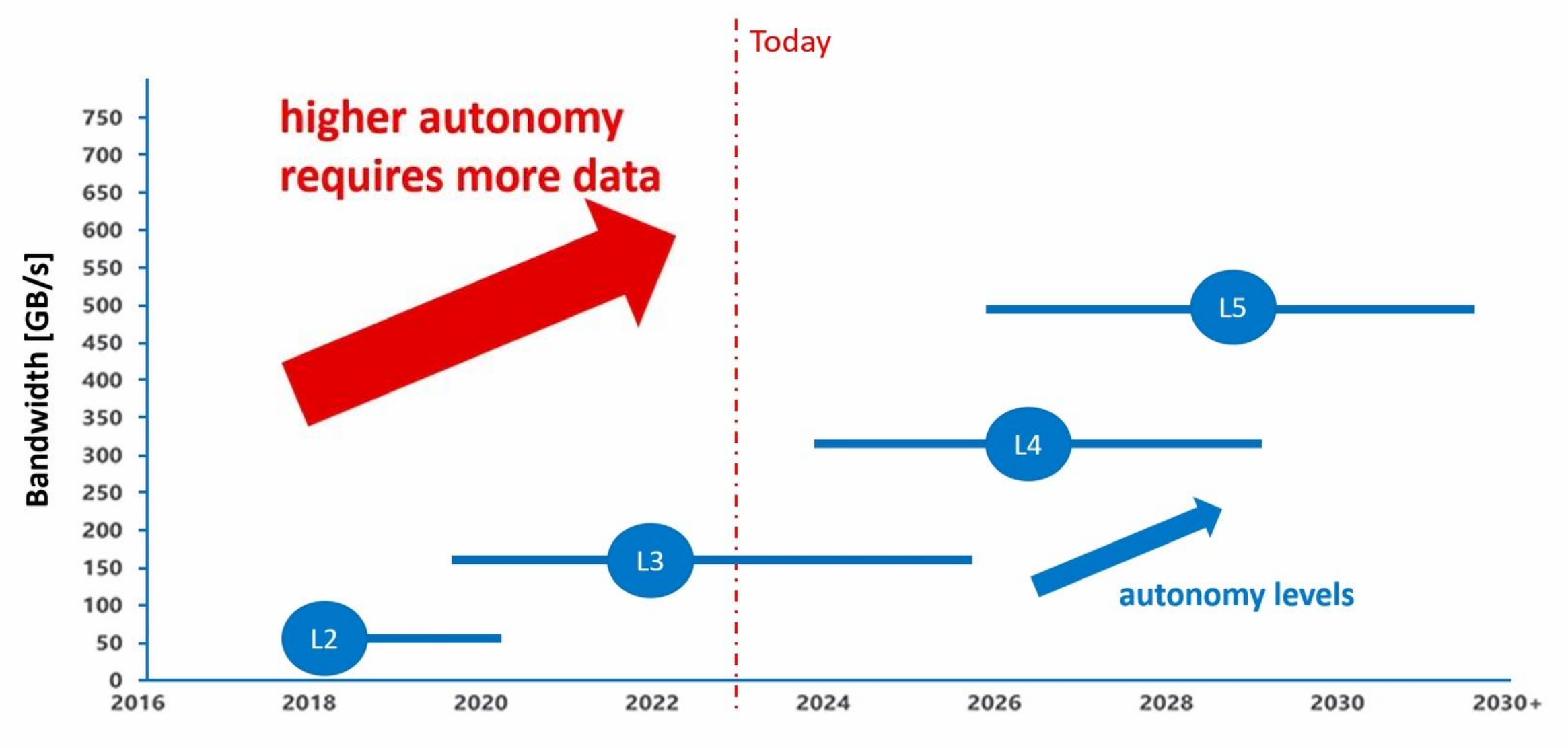
Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21. PDF. Video. Code.

- Tulyakov, Bochicchio, Gehrig, et al., Time Lens++: Event-based Frame Interpolation with Parametric Non-linear Flow and Multi-scale Fusion, CVPR'22
- Tulyakov, Gehrig, et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21

Advanced Driver Assistance Systems (ADAS)

Tesla Vision System

Memory Bandwidth Requirements by ADAS level

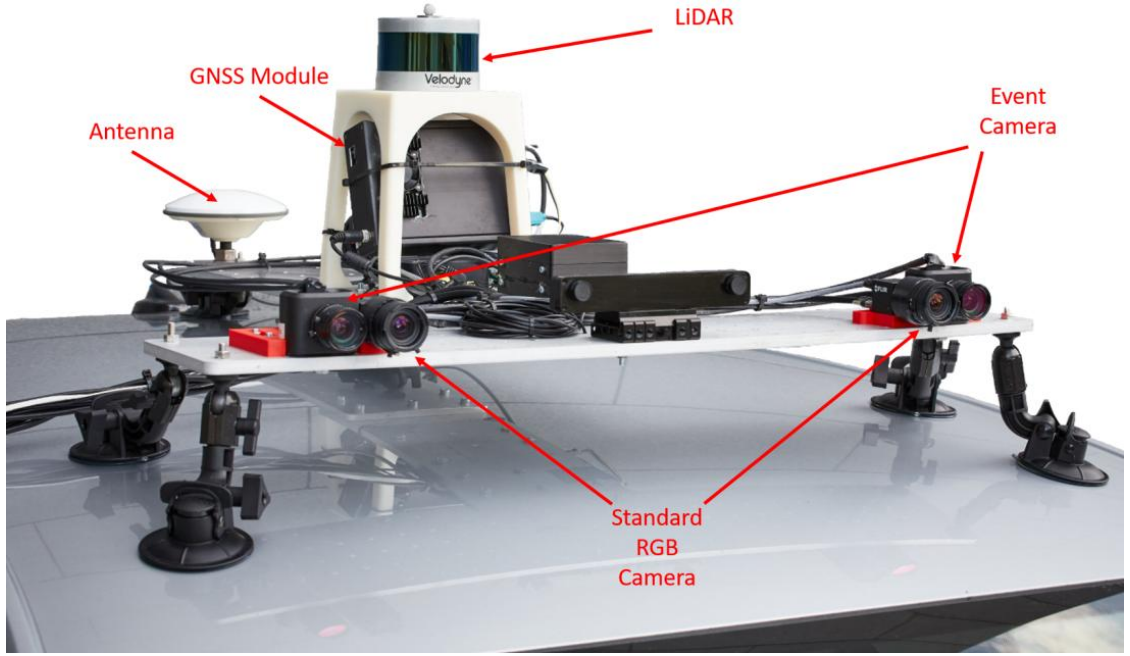


https://www.electronicspecifier.com/industries/automotive/pushing-the-envelope-for-adas-with-advanced-memory-technologies

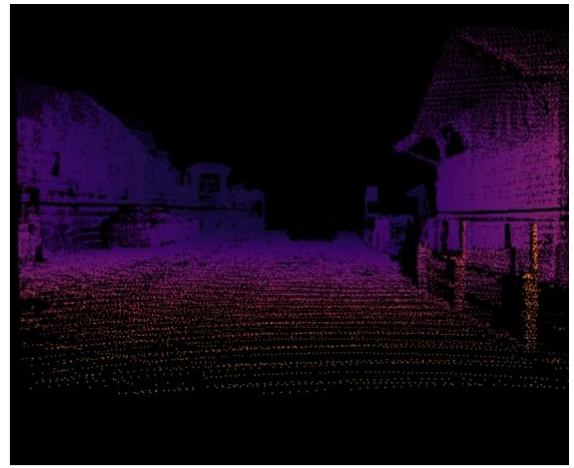
Can we build a **low-latency** and **low-bandwidth** navigation architecture?

Yes, by combining the complementary advantages of standard and event cameras!

DSEC Dataset: 40km of urban and rural driving across Switzerland



Events



Lidar

RGB Frames

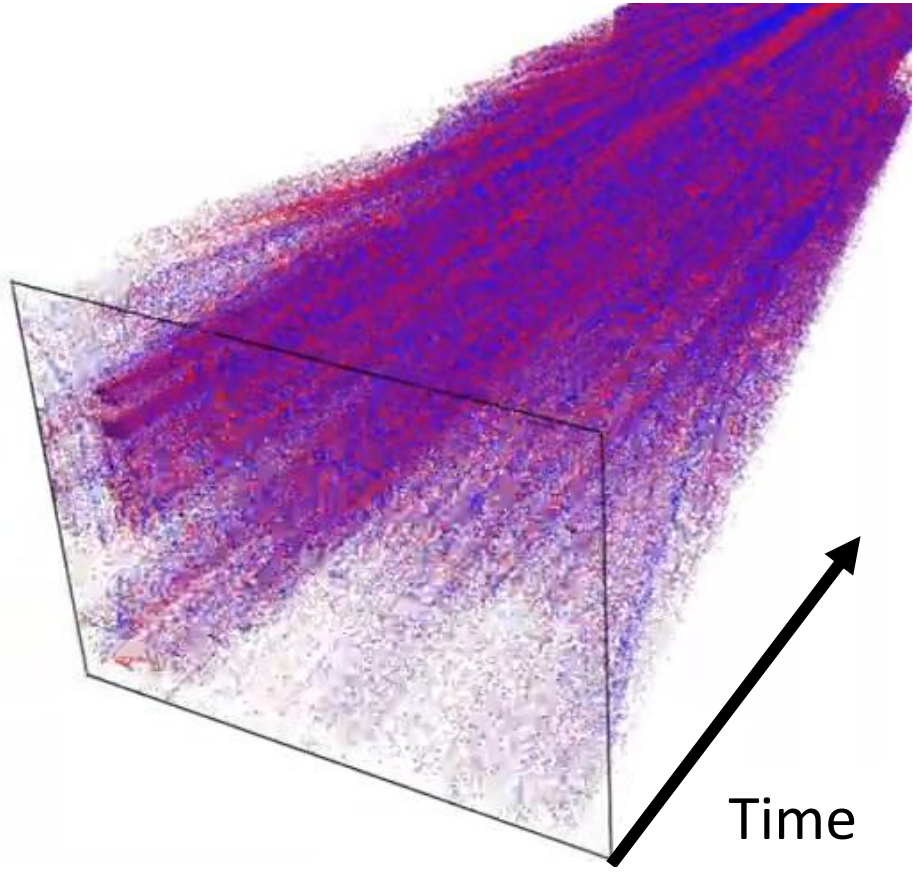
Driven route

Gehrig et al., *DSEC: A Stereo Event Camera Dataset for Driving Scenarios*, RAL'21: https://github.com/uzh-rpg/DSEC

Space-time visualization

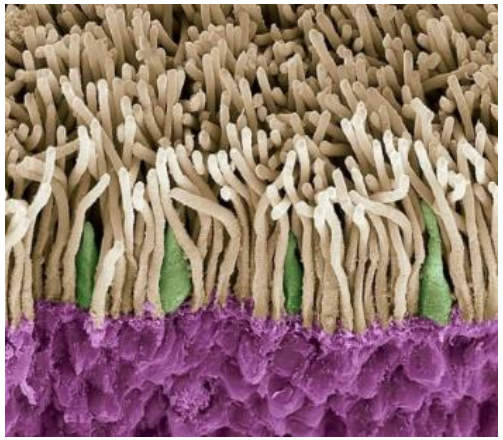
Standard camera

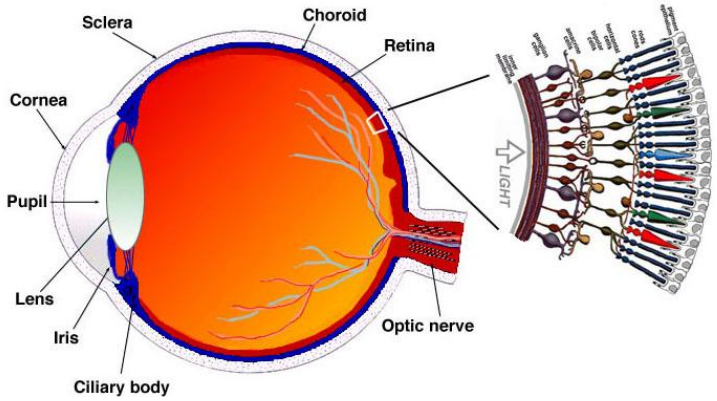
Event camera

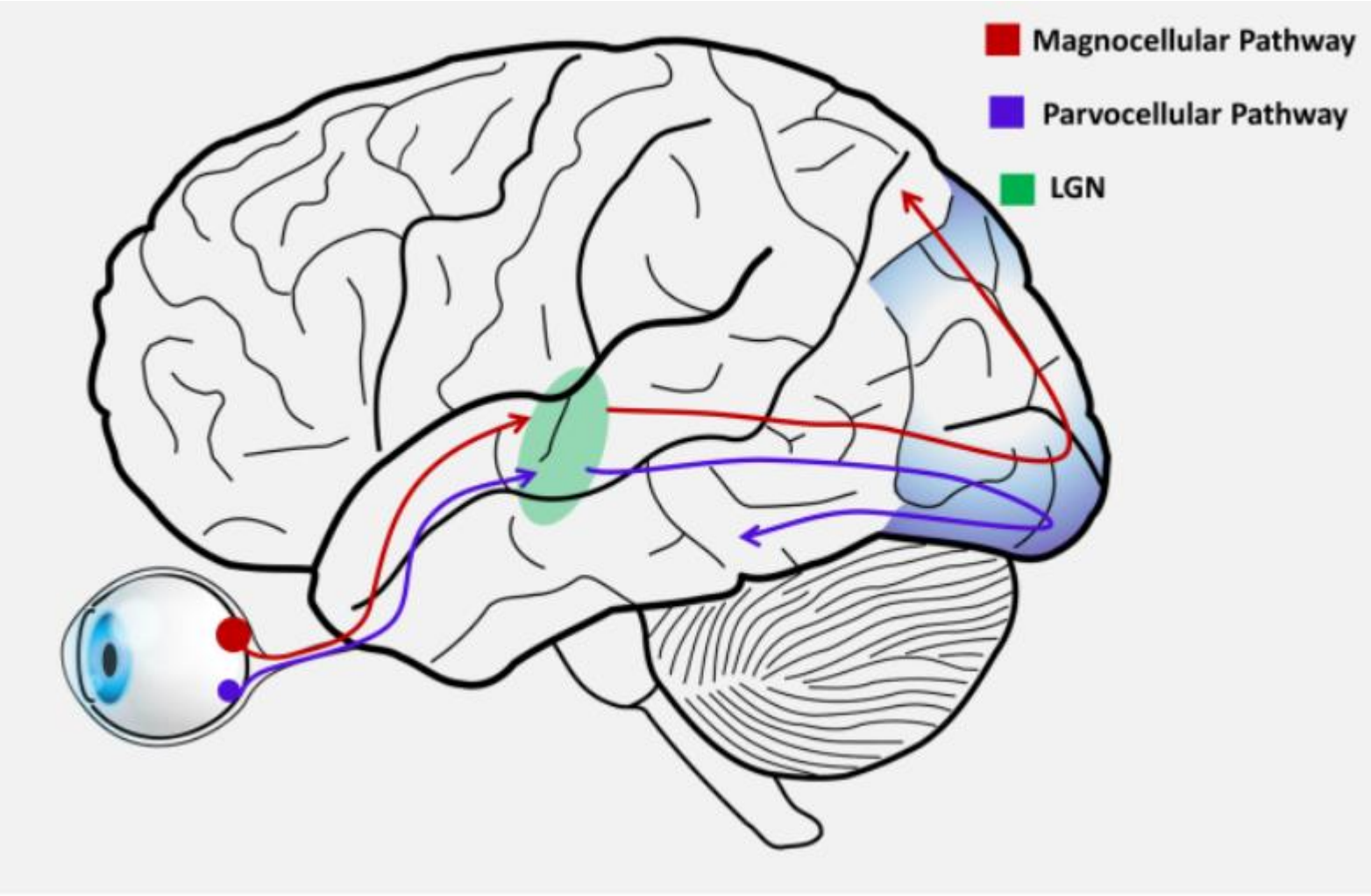


How do we combine the complementary advantages of standard and event cameras?

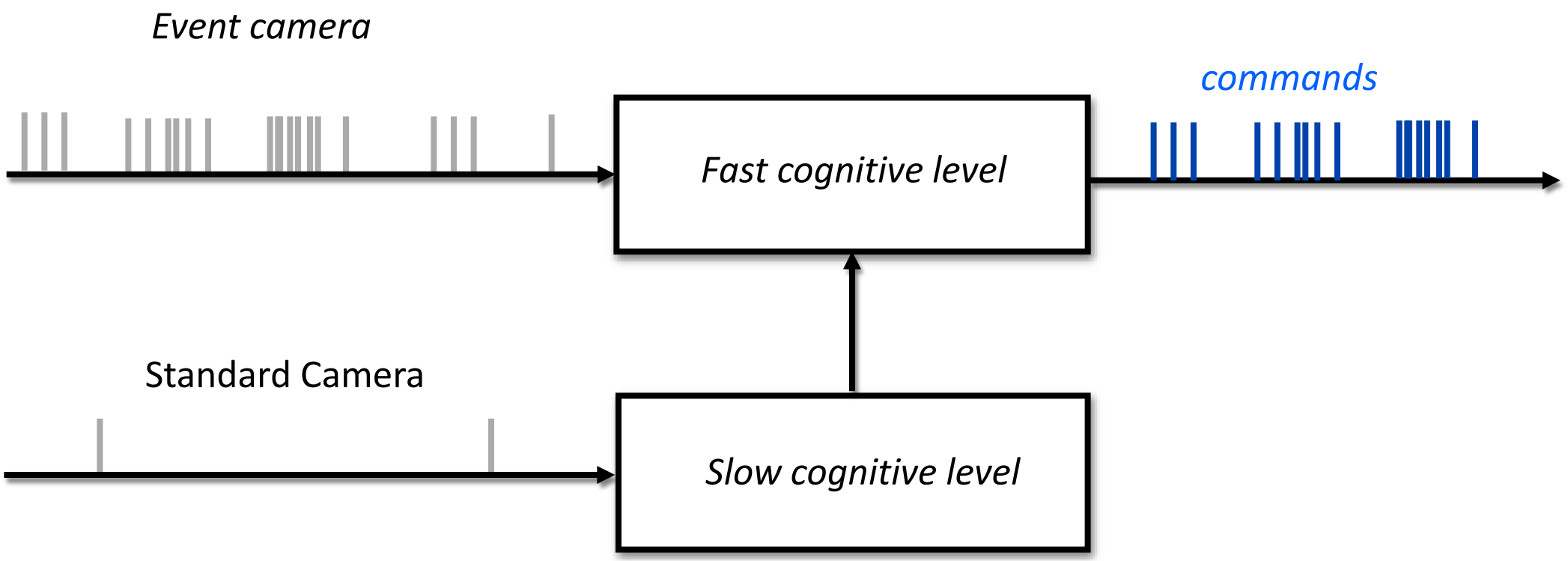
Magno and Parvo Pathways of the Primate Visual System





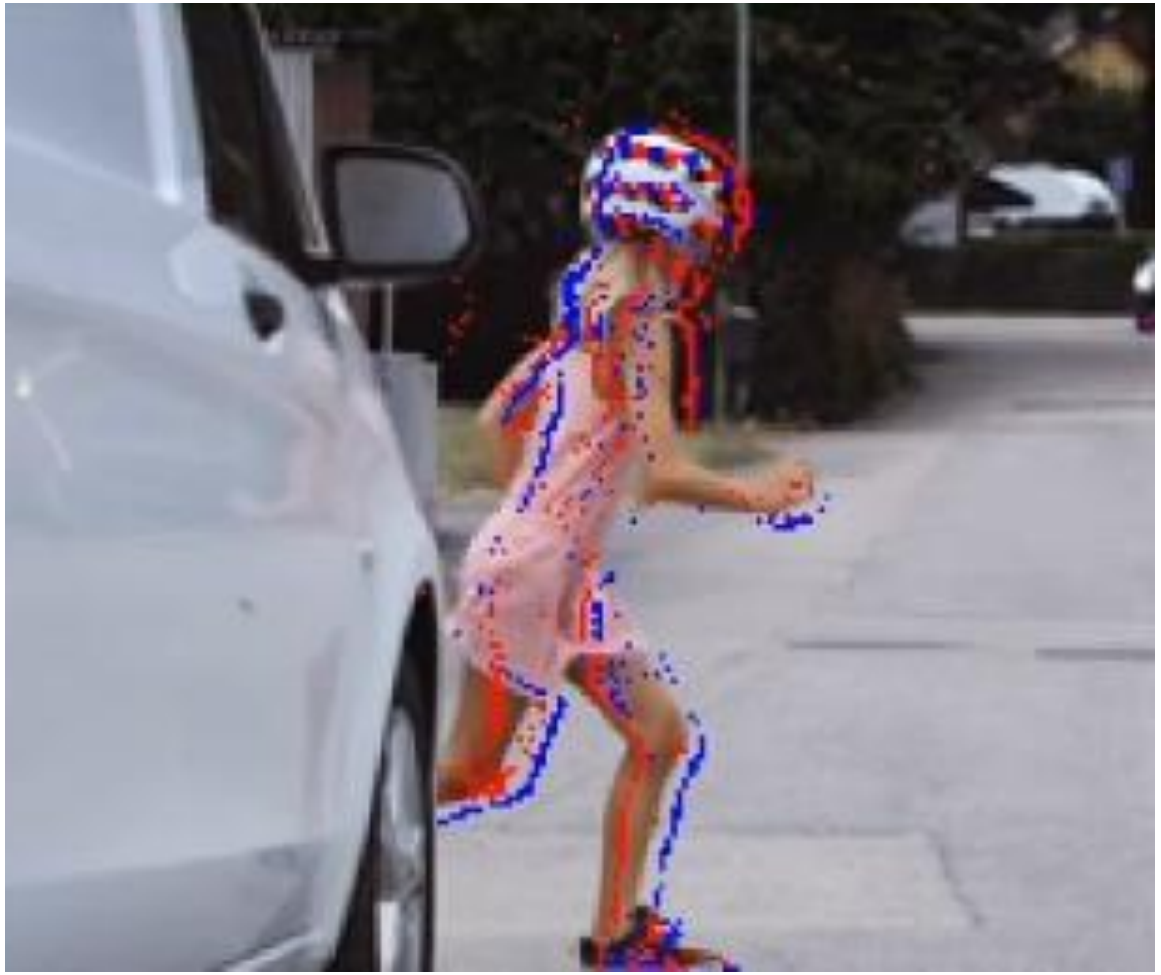


Hybrid Asynchronous Navigation Architecture



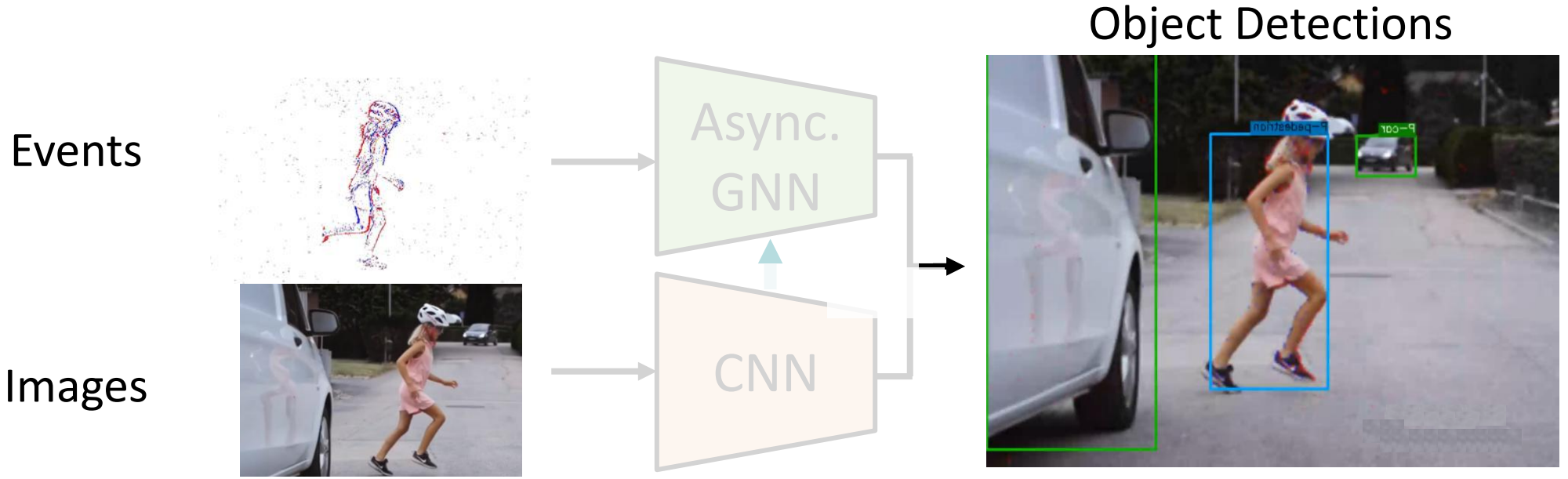
Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

Low Latency Automotive Vision



Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

Hybrid Asynchronous Object Detection

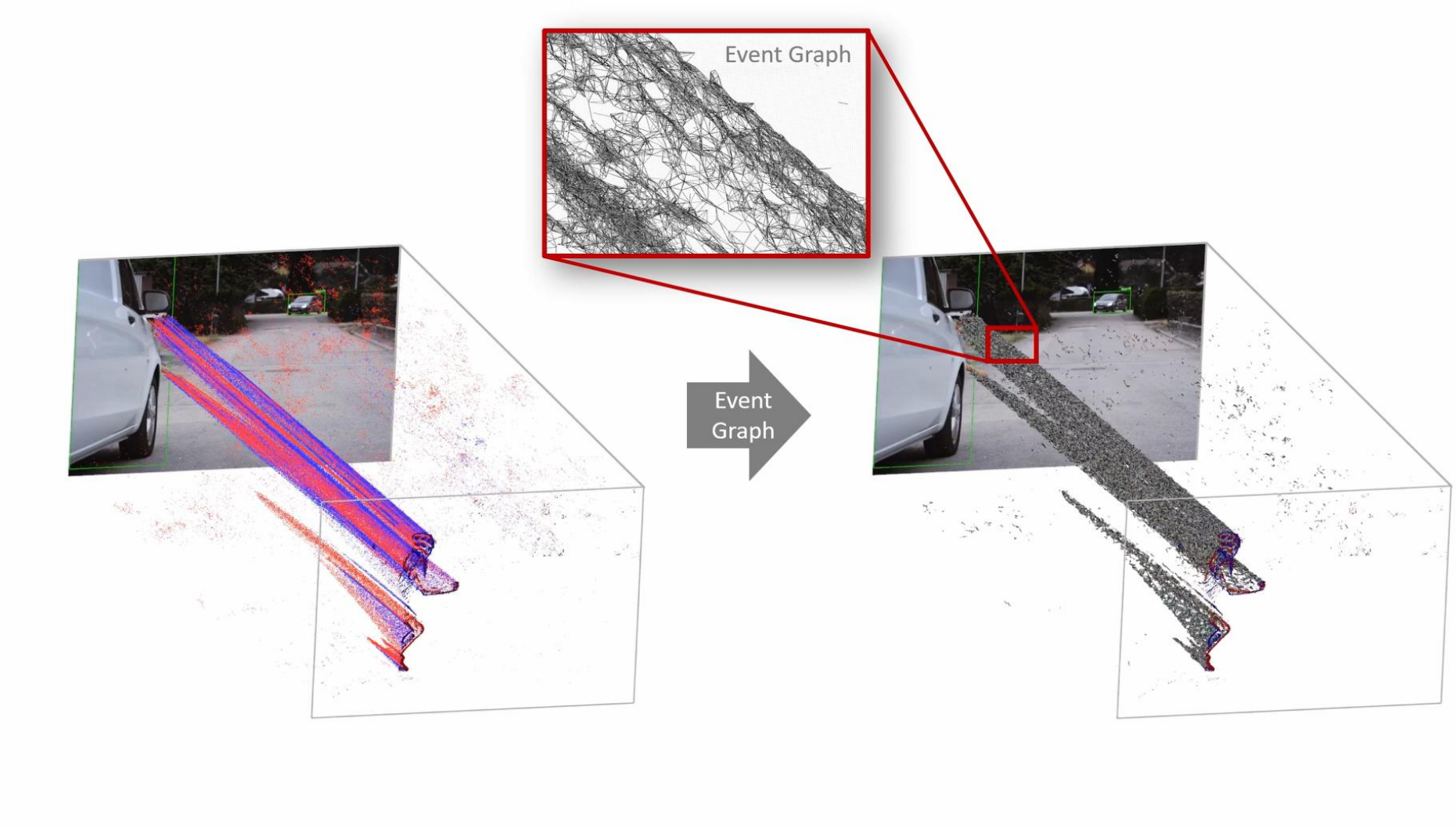


Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

Hybrid Asynchronous Object Detection

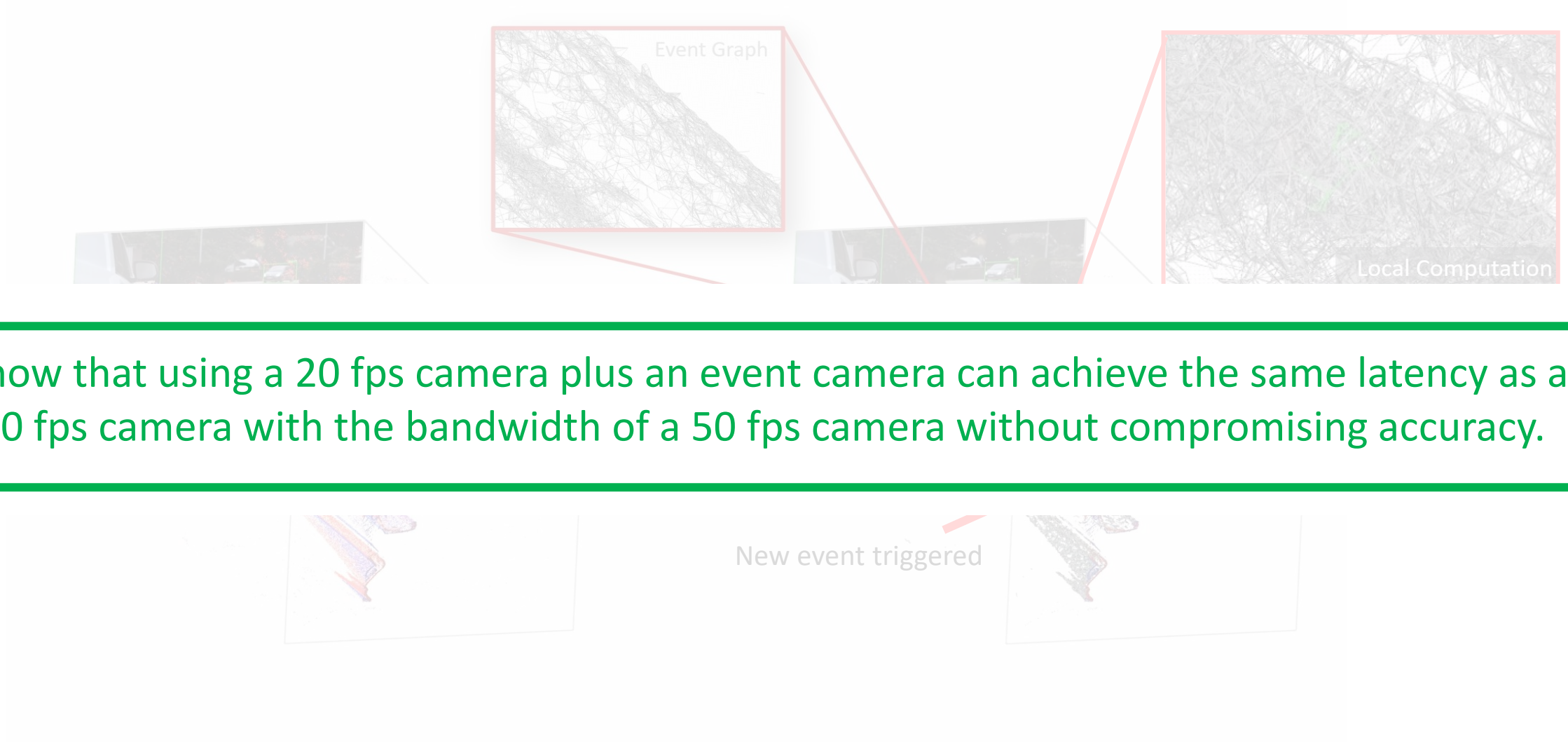
Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

Hybrid Asynchronous Object Detection



Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

Hybrid Asynchronous Object Detection



We show that using a 20 fps camera plus an event camera can achieve the same latency as a 5,000 fps camera with the bandwidth of a 50 fps camera without compromising accuracy.

Gehrig, Scaramuzza, *Low Latency Automotive Vision with Event Cameras*, **Nature**, 2024

The Evolution of Event Cameras

First event camera
by University of Zurich

Resolution: **128×128 pxl**
Pixel size: **40 microns**

First event camera
commercialized by IniVation

United States Patent

EVENT CAMERA FOR GENERATION OF EVENT-BASED IMAGES

Applicant: **Facebook Technologies**, LLC, Menlo Park, CA (US)

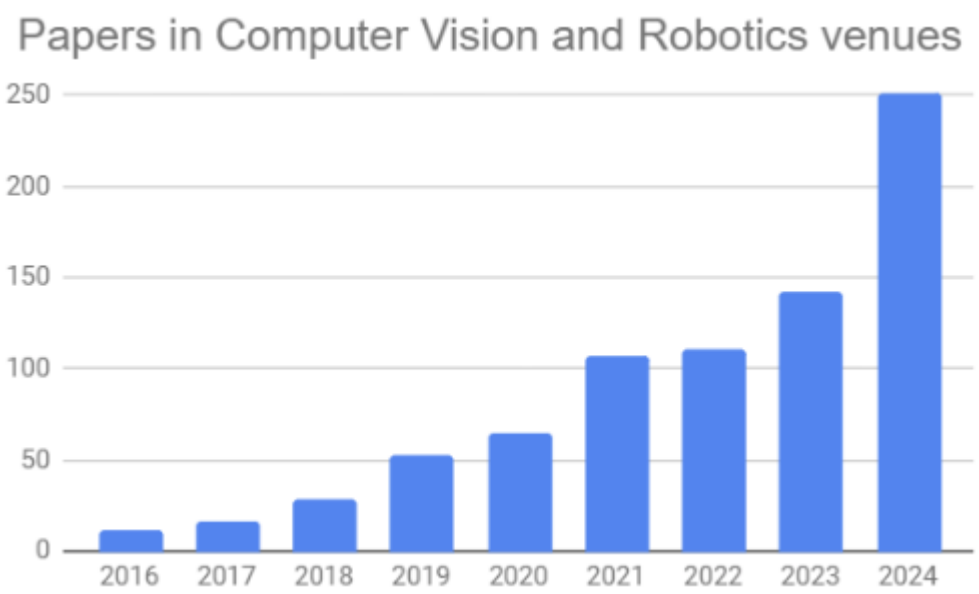
METHOD AND DEVICE FOR EYE TRACKING USING EVENT CAMERA DATA

Applicant: **Apple Inc.**, Cupertino, CA (US)

Meta opens
Event-based Sensing Lab

First event cameras in space

2008 — 2014 — 2019 — 2021 — 2022 — 2023



SONY

SAMSUNG

OmniVision

First Full-HD event sensors:

Resolution: **1280×720 pxl**
Pixel size: **5 microns**

REUTERS® World Business Markets Legal Breakingviews

Camera chip startup Prophesee and Qualcomm sign multi-year deal

By Jane Lee — February 27, 2023

PROPHESEE | Qualcomm

Thanks!

Previous and current sponsors