University of Zurich[UZH]

ROBOTICS & PERCEPTION GROUP

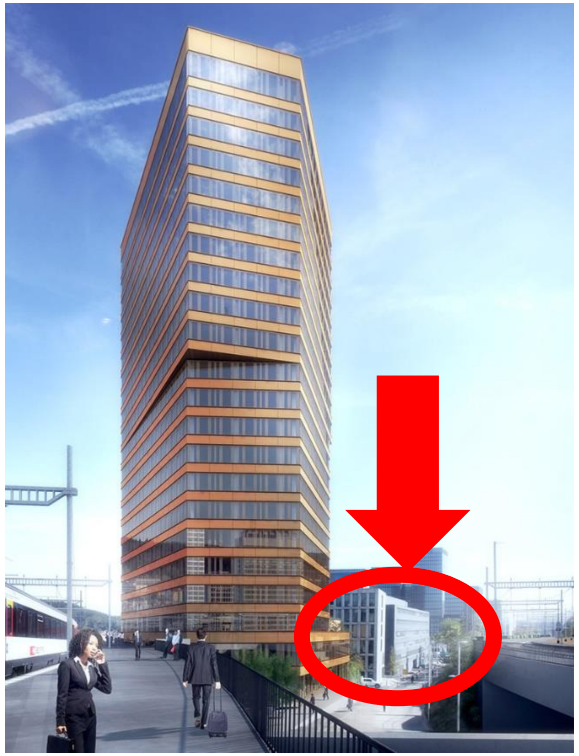
Event-based Imaging

Davide Scaramuzza

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

My lab: the Robotics and Perception Group

- **Address**: Andreasstrasse 15, 2nd floor, next to **Zurich Oerlikon** train station
- **Webpage**: http://rpg.ifi.uzh.ch



Our Research: Autonomous, Vision-based Navigation

Real-time, onboard computer vision, learning, and control for **autonomous agile robotics** using both **standard cameras** or neuromorphic **event-based** cameras

Event Cameras

http://rpg.ifi.uzh.ch/docs/EventVisionSurvey.pdf

Event-based Vision: A Survey

Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew Davison, Jörg Conradt, Kostas Daniilidis, Davide Scaramuzza

Abstract— Event cameras are bio-inspired sensors that work radically different from traditional cameras. Instead of capturing images at a fixed rate, they measure per-pixel brightness changes asynchronously. This results in a stream of events, which encode the time, location and sign of the brightness changes. Event cameras posses outstanding properties compared to traditional cameras: very high dynamic range ($140\,\mathrm{dB}$ vs. $60\,\mathrm{dB}$), high temporal resolution (in the order of $\mu\mathrm{s}$), low power consumption, and do not suffer from motion blur. Hence, event cameras have a large potential for robotics and computer vision in challenging scenarios for traditional cameras, such as high speed and high dynamic range. However, novel methods are required to process the unconventional output of these sensors in order to unlock their potential. This paper provides a comprehensive overview of the emerging field of event-based vision, with a focus on the applications and the algorithms developed to unlock the outstanding properties of event cameras. We present event cameras from their working principle, the actual sensors that are available and the tasks that they have been used for, from low-level vision (feature detection and tracking, optic flow, etc.) to high-level vision (reconstruction, segmentation, recognition). We also discuss the techniques developed to process events, including learning-based techniques, as well as specialized processors for these novel sensors, such as spiking neural networks. Additionally, we highlight the challenges that remain to be tackled and the opportunities that lie ahead in the search for a more efficient, bio-inspired way for machines to perceive and interact with the world.

Index Terms—Event Cameras, Bio-Inspired Vision, Asynchronous Sensor, Low Latency, High Dynamic Range, Low Power.

1 Introduction and Applications

"THE *brain is imagination, and that was exciting to me; I wanted to build a chip that could imagine something*[1]." that is how Misha Mahowald, a graduate student at Caltech in 1986, started to work with Prof. Carver Mead on the stereo problem from a joint biological and engineering per-

as well as new computer vision and robotic tasks. Sight is, by far, the dominant sense in humans to perceive the world, and, together with the brain, learn new things. In recent years, this technology has attracted a lot of attention from both academia and industry. This is due to the availability of prototype event cameras and the advantages that these

Open Challenges in Computer Vision

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

Latency

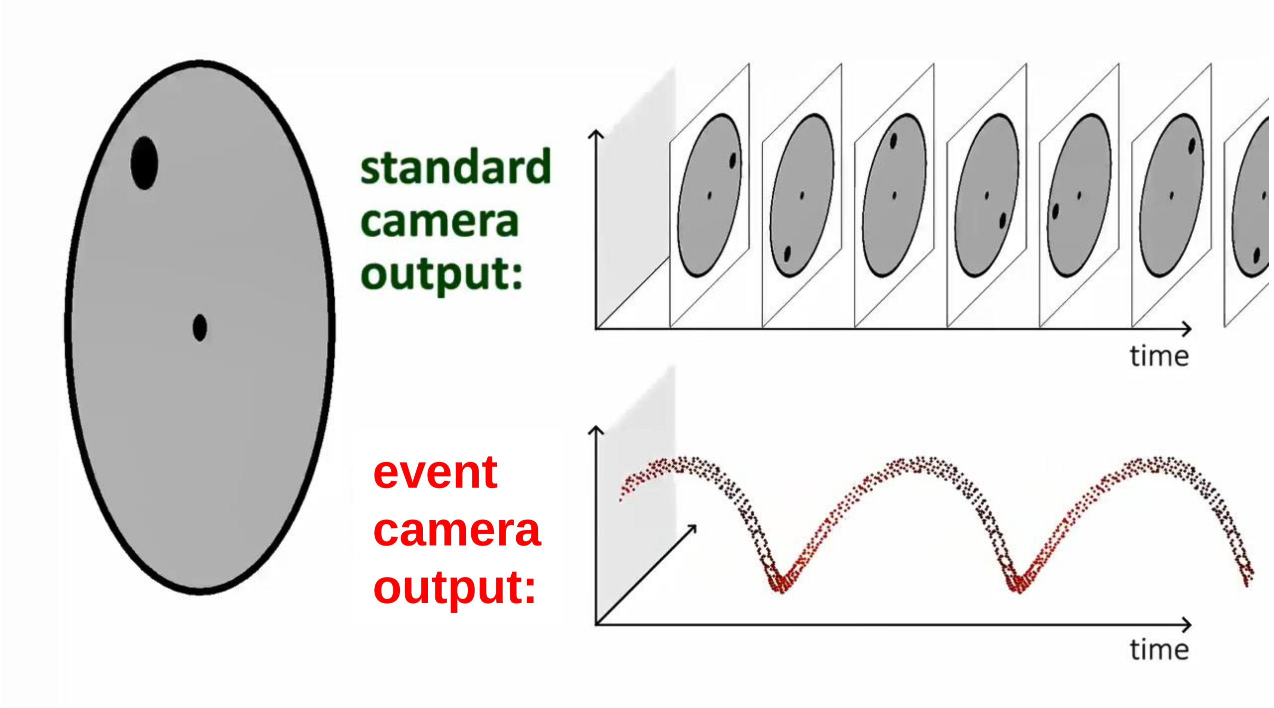
Motion blur

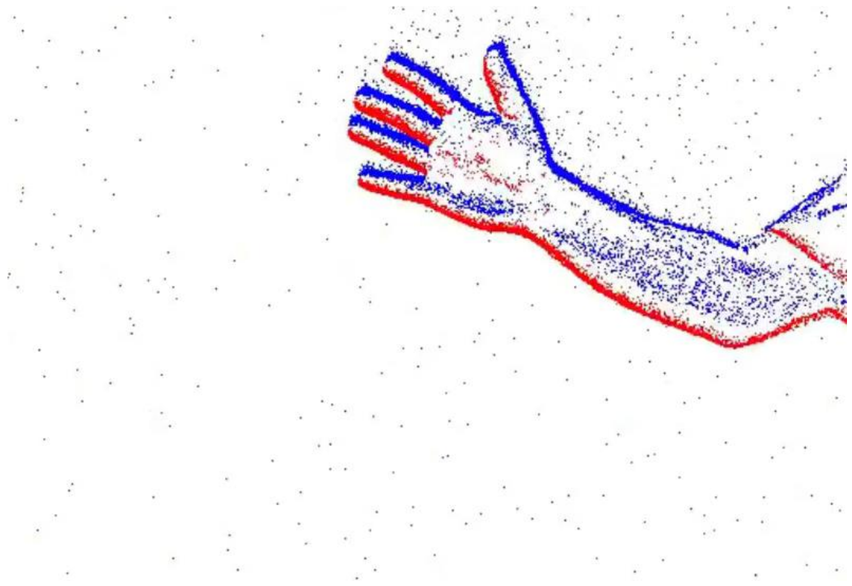
Dynamic Range

Event cameras do not suffer from these problems!

What is an Event Camera?

- It is camera that measures only **motion in the scene**
- **Key advantages:**
  1. **Low-latency** (~ 1 μs)
  2. **Ultra low power** (~1mW)
  3. **No motion blur**

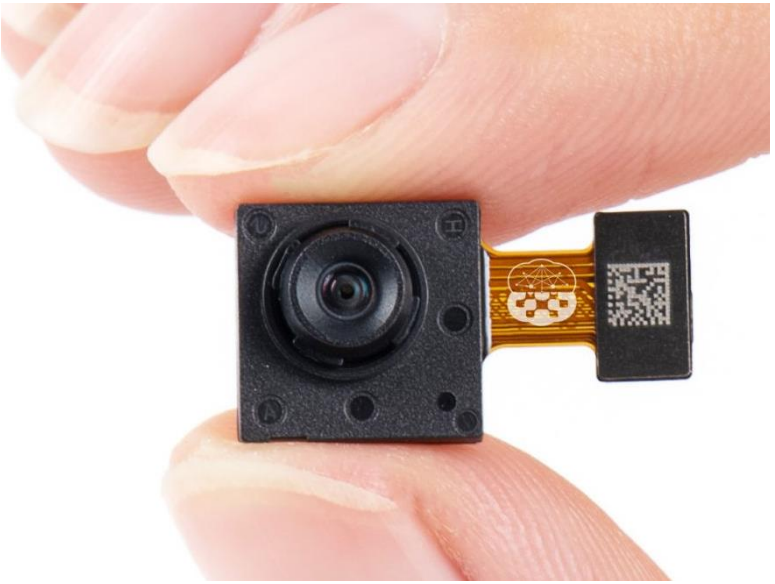




[1] Lichtsteiner, Posch, Delbruck, *A 128x128 120 dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor*, IEEE Journal of Solid-State Circuits, 2008. PDF
[2] Gallego et al., *Event-based Vision: A Survey*, T-PAMI, 2020. PDF.

Opportunities

- **Low latency**: AR/VR, robotics, automotive (<10ms)
- **Low power**: AR/VR, always-on devices (see Synsense)
- **No motion blur**: AR/VR, robotics, mobile devices



Event camera + Speck spiking-network neuromorphic processor from Synsense: can recognize faces at 50Hz while consuming < 1mW

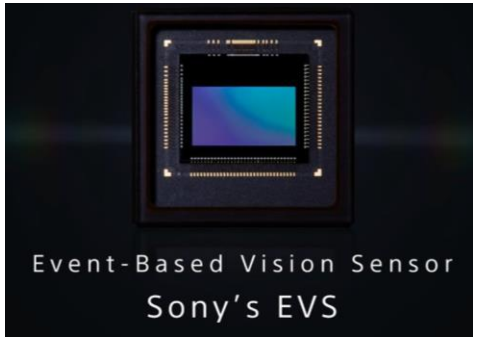
Who sells event cameras and how much are they?

- Prophesee & SONY:
  - **ATIS sensor: events**, **IMU**, absolute intensity at the event pixel
  - Resolution: **1M pixels**
  - **Cost: ~5,000 USD**

- Inivation & Samsung
  - **DAVIS sensor**: **frames**, **events**, **IMU**.
  - Resolution: **VGA** (640x480 pixels)
  - **Cost: ~5,000 USD**

- CelePixel Technology & Omnivision:
  - **Celex One: events, IMU**, absolute intensity at the event pixel
  - Resolution: **1M pixels**
  - **Cost: ~1,000 USD**

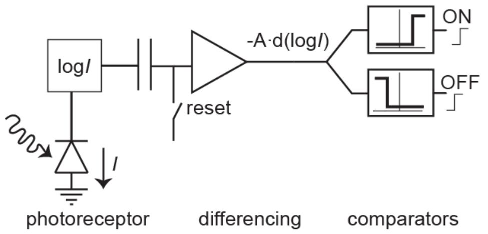
SONY



SAMSUNG

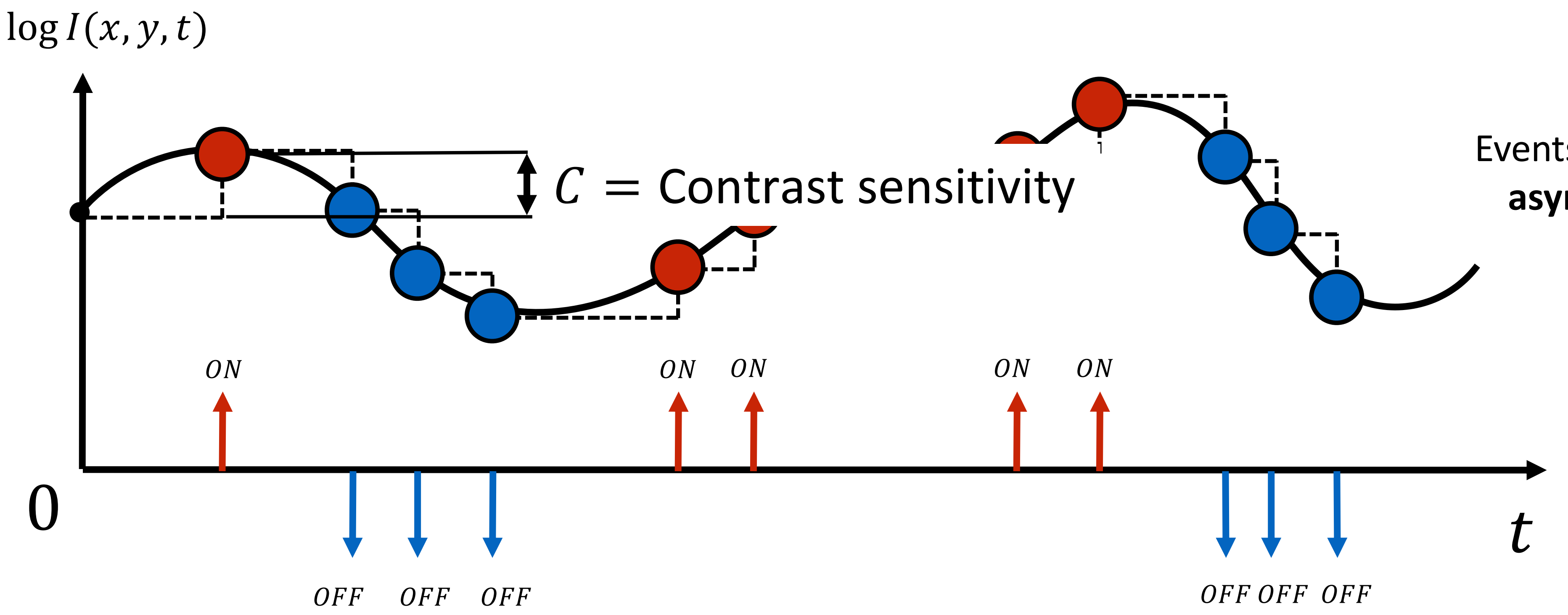
OmniVision

Generative Event Model



Consider the intensity at a **single pixel** (x, y). An event is generated when the following condition is satisfied:

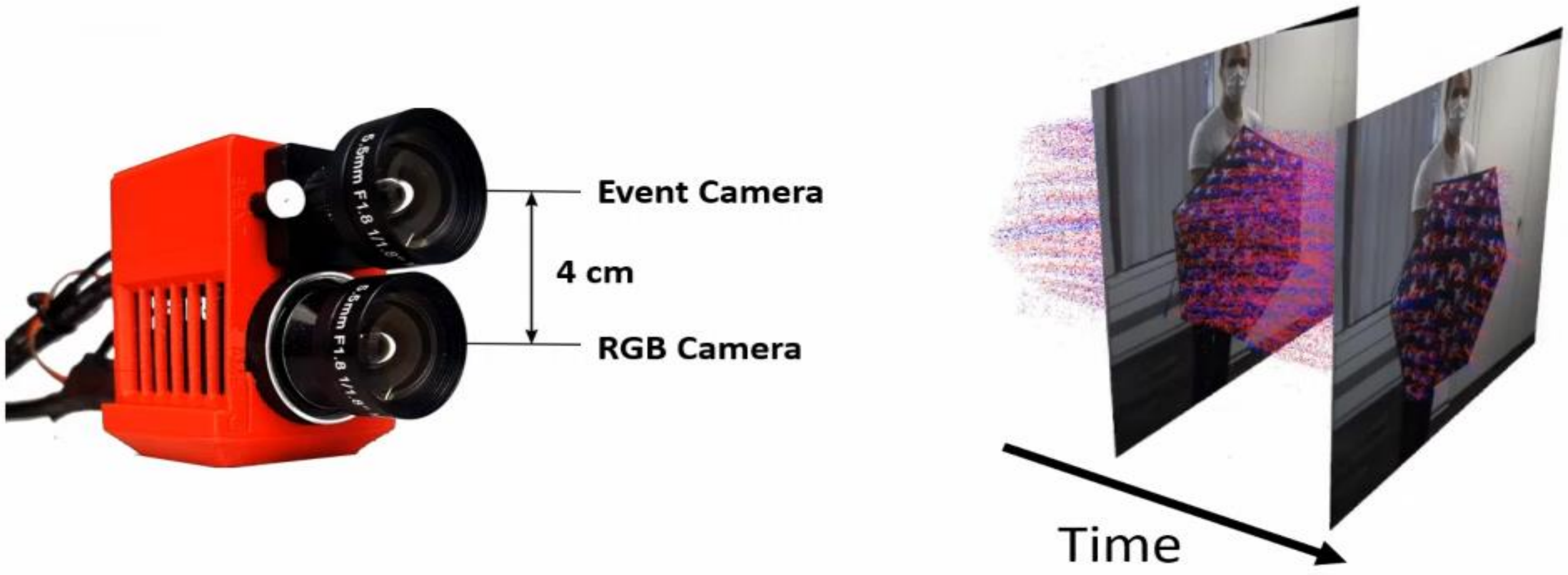
$$\log I(x, y, t + \Delta t) - \log I(x, y, t) = \pm C$$



Events are triggered **asynchronously**

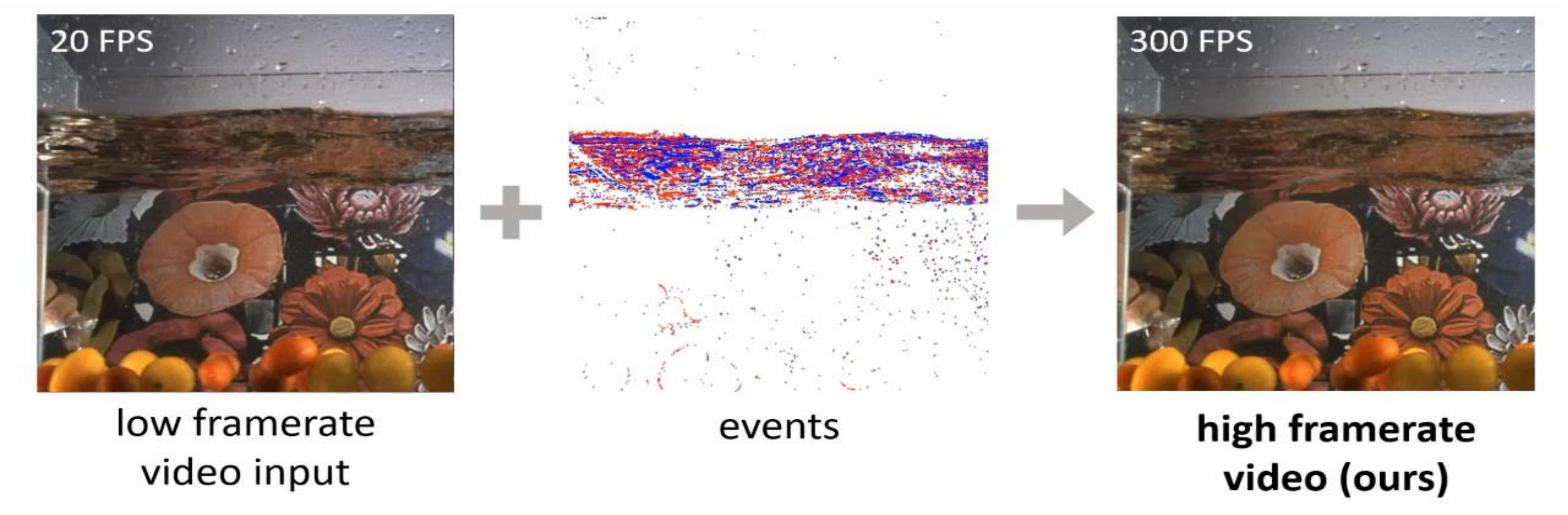
Can we reconstruct the pixel intensity? $\log(I(x, y, t)) = \log(x, y, 0) + \Sigma_{k=1}^{N_t} p_k C$

Combining Events and Frames for Ultimate Performance



Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR’21. PDF. Video. Code. Featured on Two-Minute-Papers: Video.

Combining Events and Frames for Ultimate Performance



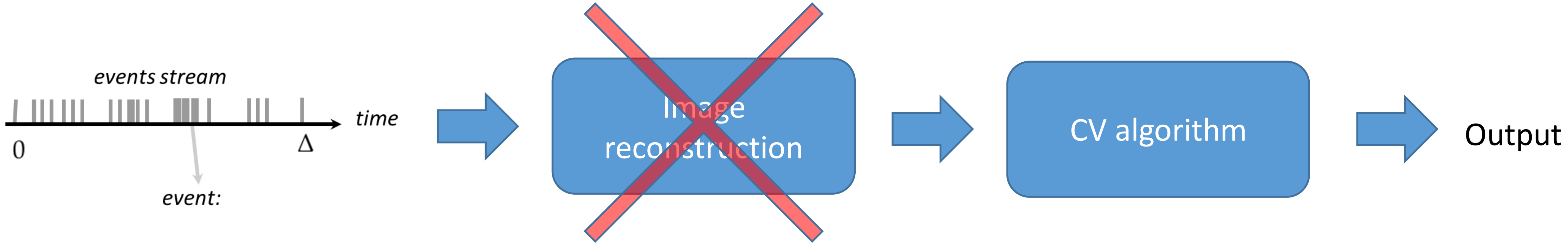
We use events to **upsample low-framerate video** by over **50 times** with **only 1/40th of the memory** footprint!

Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21. PDF. Video. Code. Featured on Two-Minute-Papers: Video.

Tulyakov et al., Time Lens++: Event-based Frame Interpolation with Parametric Flow and Multi-scale Fusion, CVPR 2022, PDF. Video. Code & Datasets

The Key Challenge

- The fact that we can reconstruct high quality video means that event cameras carry the same visual information as standard cameras
- So it must be possible to perform all vision tasks of standard cameras
- But we want to build **efficient and low energy algorithms** that compute the output **without passing through intermediate image reconstruction**

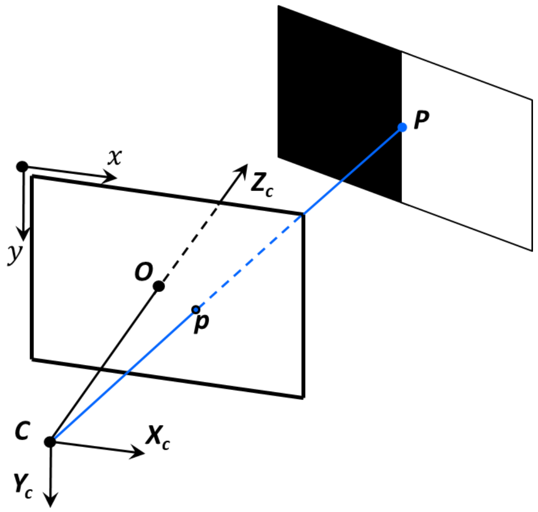


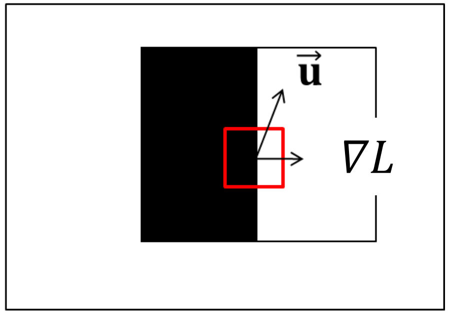
1st Order Approximation of the Generative Event Model

$$\pm C = \log I(\boldsymbol{x}, t) - \log I(\boldsymbol{x}, t - \Delta t)$$

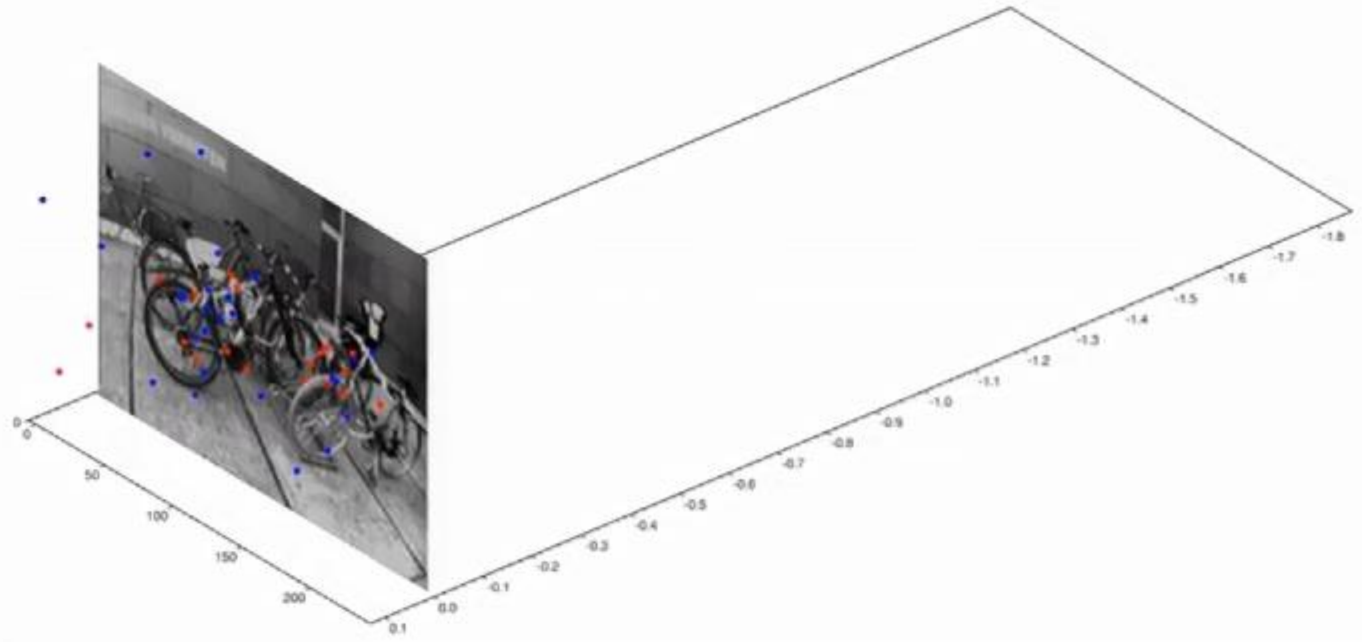
- Let $L(x, y, t) = Log(I(x, y, t))$
- Consider a given pixel $p(x, y)$ with gradient $\nabla L(x, y)$ undergoing the motion $\boldsymbol{u} = (u, v)$ in pixels, induced by a moving 3D point $\boldsymbol{P}$

$$\Rightarrow \ \pm C = -\nabla L \cdot \boldsymbol{u}$$





Application 1: Low-Latency & Low-Energy Tracking

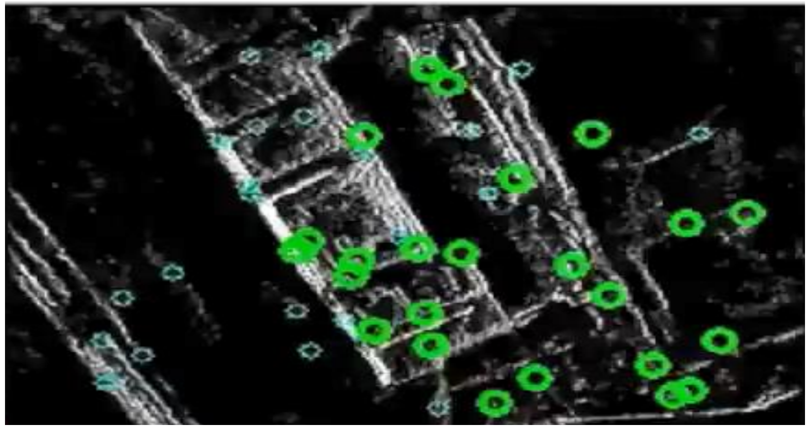


- [1] Gallego et al., *Event-based 6-DOF Camera Tracking from Photometric Depth Maps*, T-PAMI'18. PDF. Video.
- [2] Mueggler et al., *Continuous-Time Visual-Inertial Odometry for Event Cameras*, **TRO'18**. PDF
- [3] Rosinol et al., *Ultimate SLAM?,* **RAL'18 Best Paper Award finalist** PDF. Video. IEEE Spectrum.
- [3] Gehrig et al., *EKLT: Asynchronous, Photometric Feature Tracking using Events and Frames*, IJCV 2019. PDF, YouTube, Evaluation Code, Tracking Code

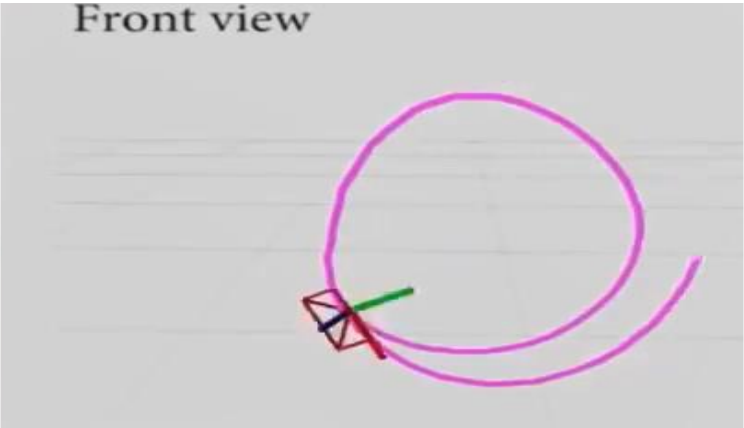
Application 1: Motion estimation in high-speed scenarios

Standard camera

Event camera



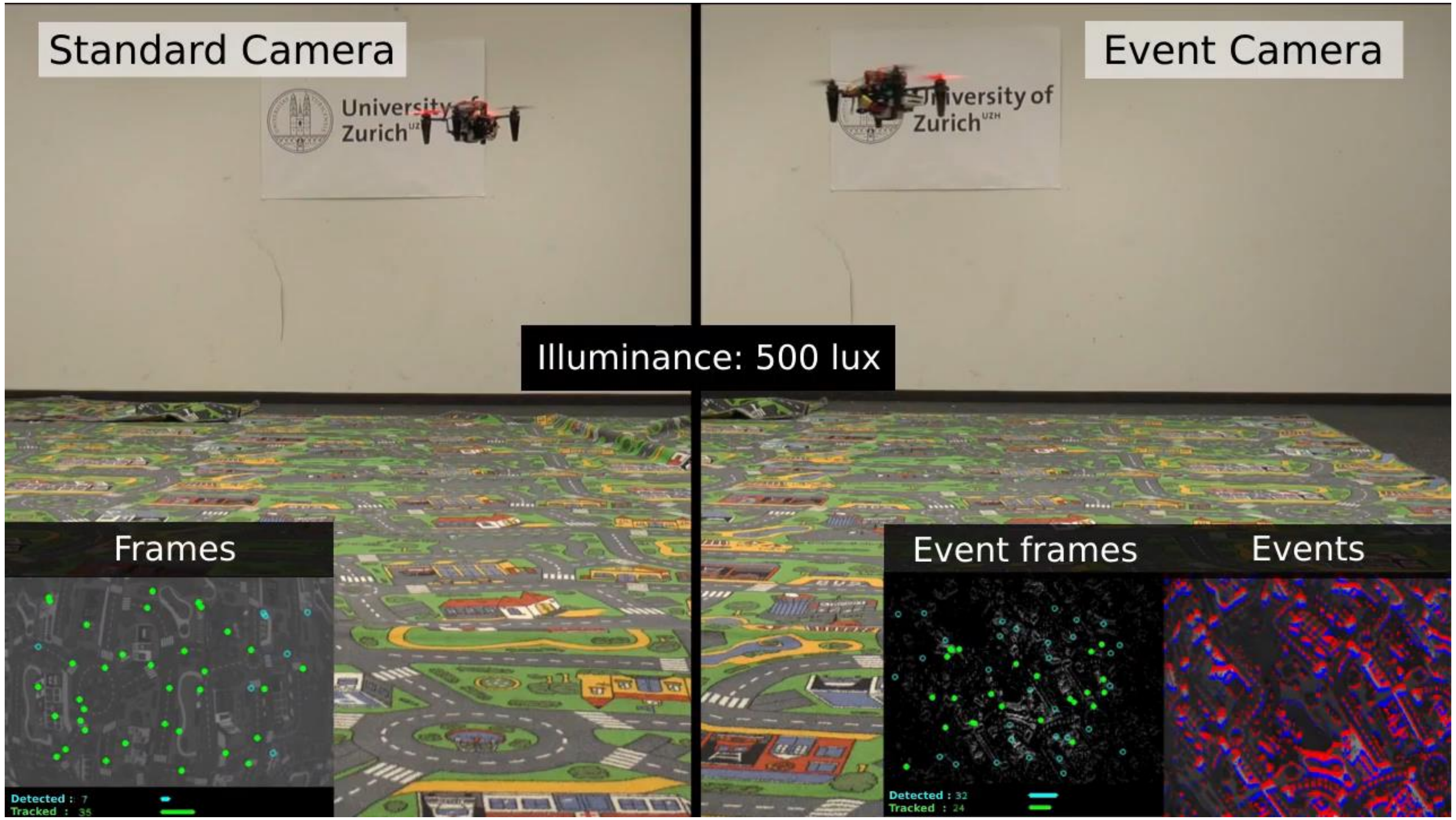
Estimated trajectory



Rosinol et al., *Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM,* **RAL'18 Best Paper Award Hon. Mention.** PDF. Video. Code.
Mueggler et al., *Continuous-Time Visual-Inertial Odometry for Event Cameras*, **T-RO'18**. PDF

Application 2: Keeping drones Flying when a **Rotor Fails**

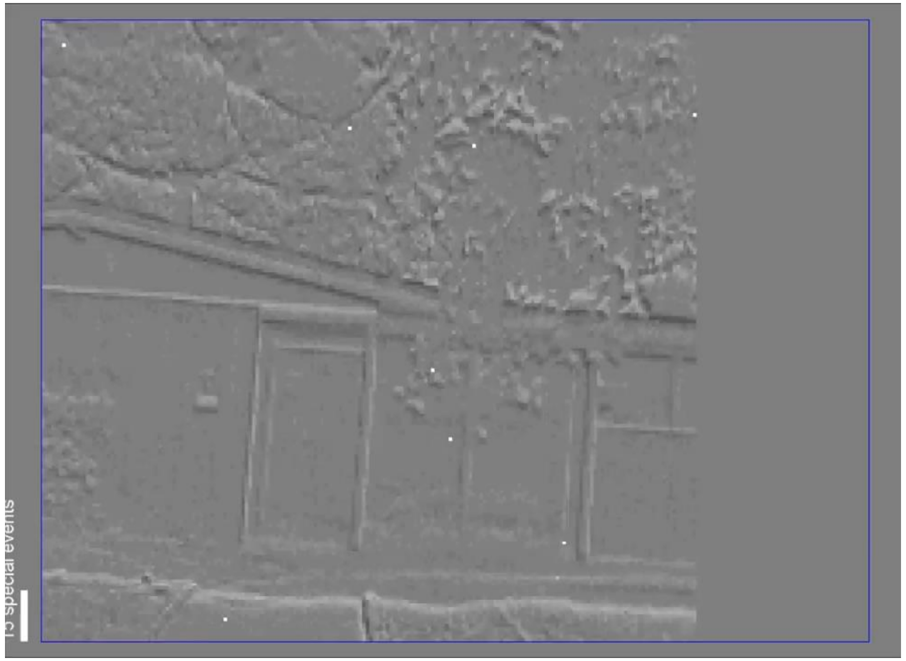
- Quadrotors subject to full rotor failure **require accurate position estimates** to avoid crashing
- Event cameras are **not affected by motion blur**



Sun, Cioffi, de Visser, Scaramuzza, *Autonomous Quadrotor Flight despite Rotor Failure with Onboard Vision Sensors: Frames vs. Events*, IEEE RAL'2021.
PDF. Video. Code. **1st place winner of the NASA TechBrief Award** (out of 700 participants)

Application 3: Video Stabilization

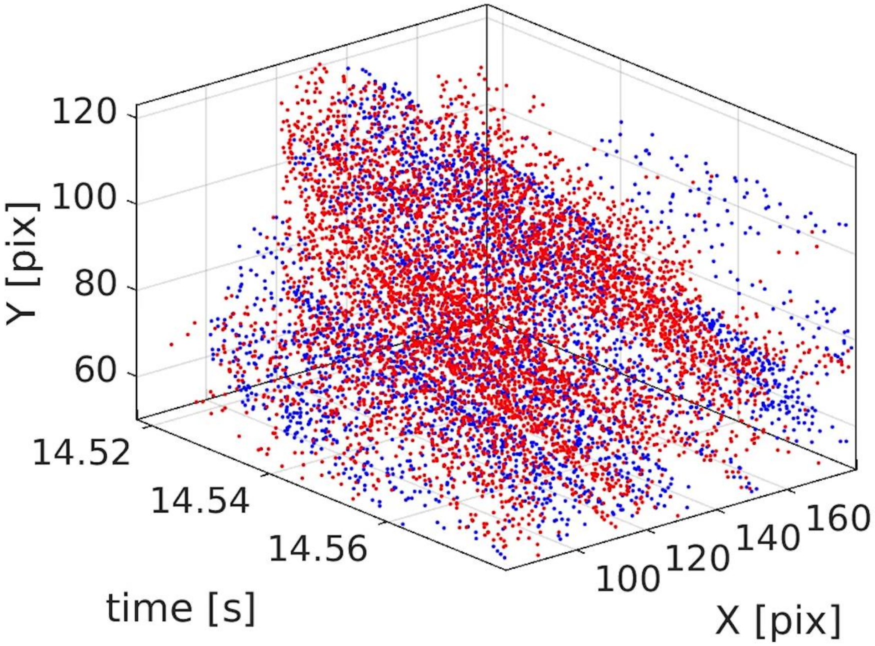
- Problem: **Estimate rotational motion** (3DoF) of an event camera
- **Can process millions of events per second** in **real time** on a **smartphone CPU** (OdroidXU4)



Gallego et al., Accurate Angular Velocity Estimation with an Event Camera, IEEE RAL'16. PDF. Video.

Application 3: Video Stabilization

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



Gallego et al., Accurate Angular Velocity Estimation with an Event Camera, IEEE RAL'16. PDF. Video.
Gallego et al., A Unifying Contrast Maximization Framework for Event Cameras, CVPR18, PDF, Video
Gallego et al., Focus Is All You Need: Loss Functions for Event-based Vision, CVPR19, PDF.

Contrast Maximization Framework

Input Events

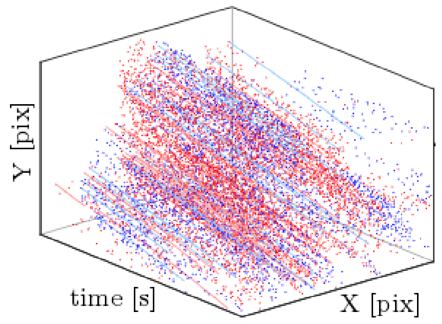
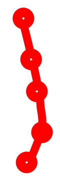


Image of

$I(\boldsymbol{x};\boldsymbol{\theta}) = \Sigma_{k=1}^{N_e} b_k \delta(\boldsymbol{x} - \boldsymbol{x}'_k)$

Optimize point trajectories

Focus score

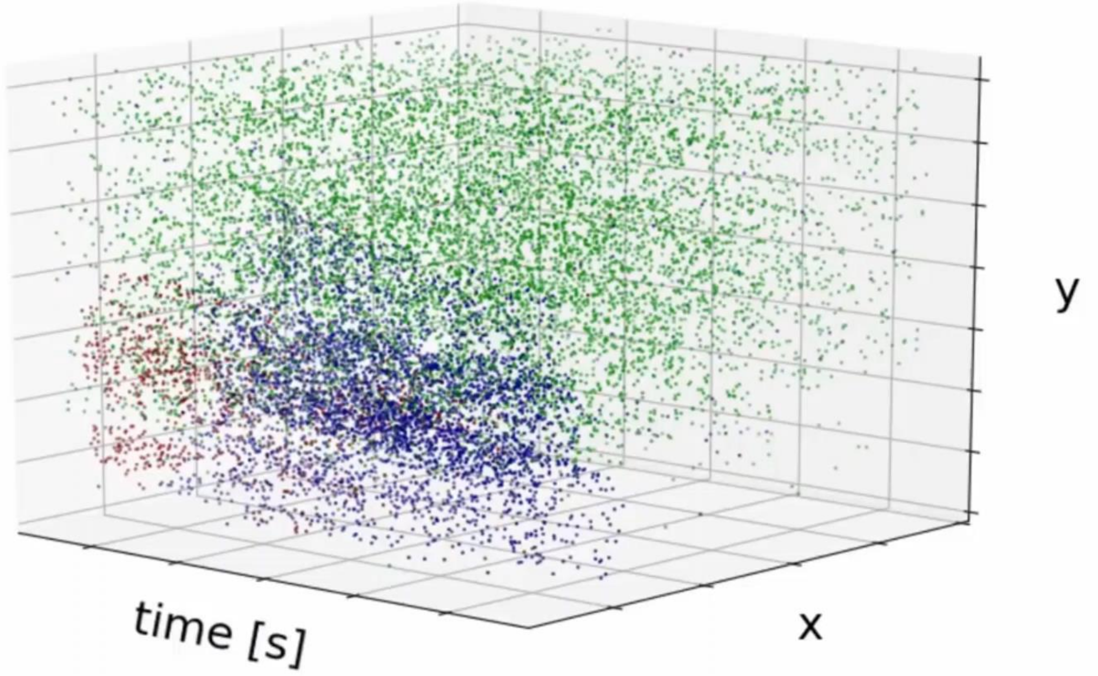


$\sigma^2(I(\boldsymbol{x};\boldsymbol{\theta}))$

Gallego et al., A Unifying Contrast Maximization Framework for Event Cameras, CVPR18, PDF, YouTube
Gallego et al., Focus Is All You Need: Loss Functions for Event-based Vision, CVPR19, PDF.

Application 4: Independent Motion Segmentation

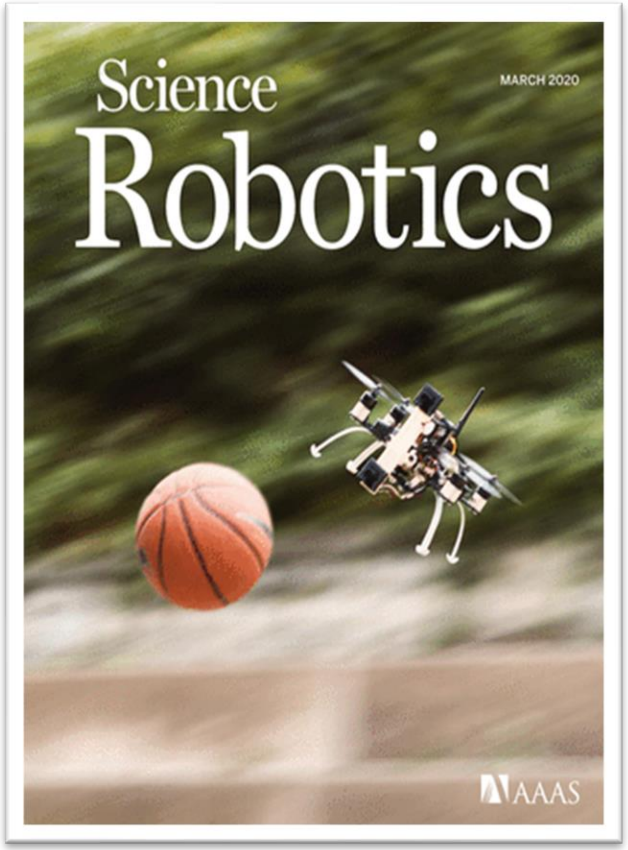
Idea: **warp spatio-temporal volume of events** to maximize focus (e.g., sharpness) of the resulting image



Stoffregen et al., Motion Segmentation by Motion Compensation, ICCV'19. PDF. Video.

Application 5: Dodging Dynamic Objects

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



Falanga et al., *Dynamic Obstacle Avoidance for Quadrotors with Event Cameras*, **Science Robotics,** 2020. PDF. Video. Featured in IEEE Spectrum

Application 6: Event-Guided Depth Sensing

- Standard depth sensors **sample depth** uniformly → **large power consumption** and **high latency**
- Idea: event cameras **guide depth sensing**: measure **more densely** areas with **high motion density**

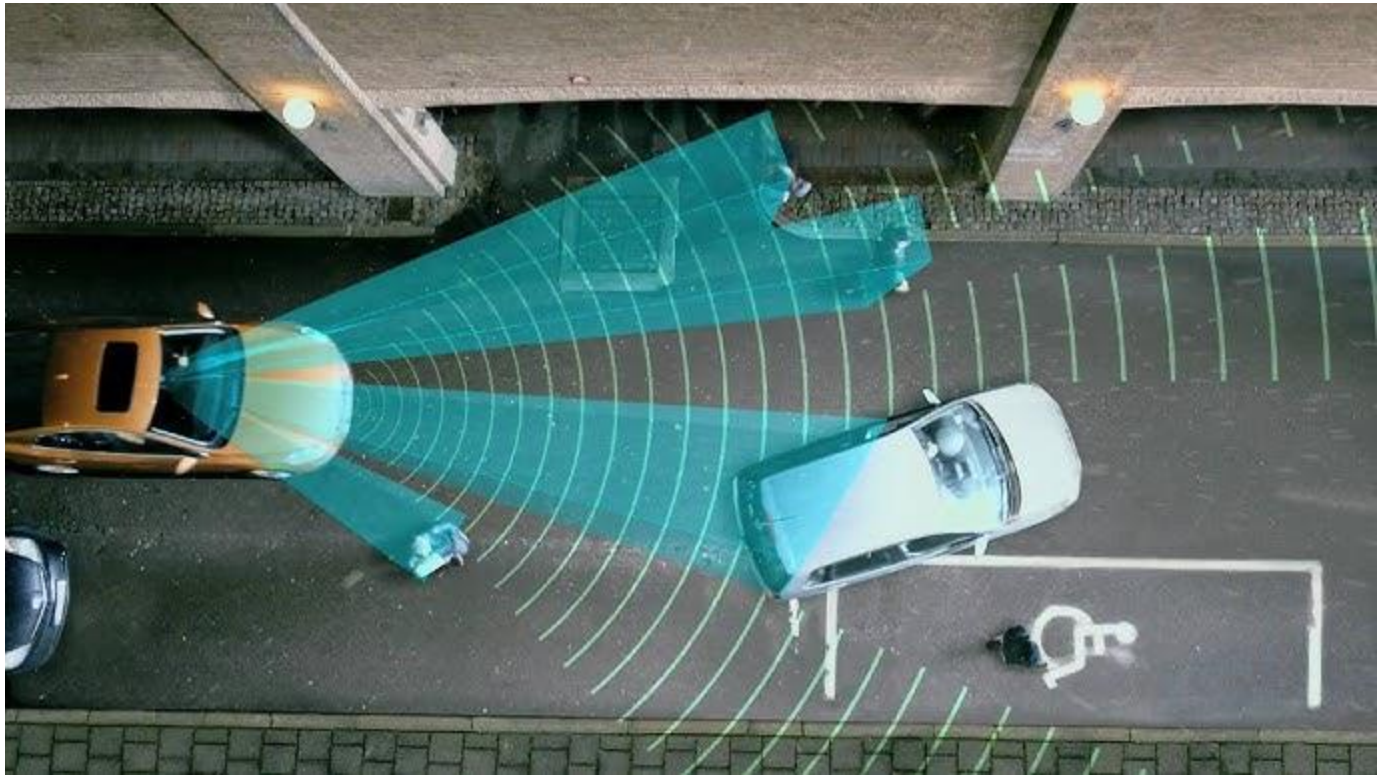
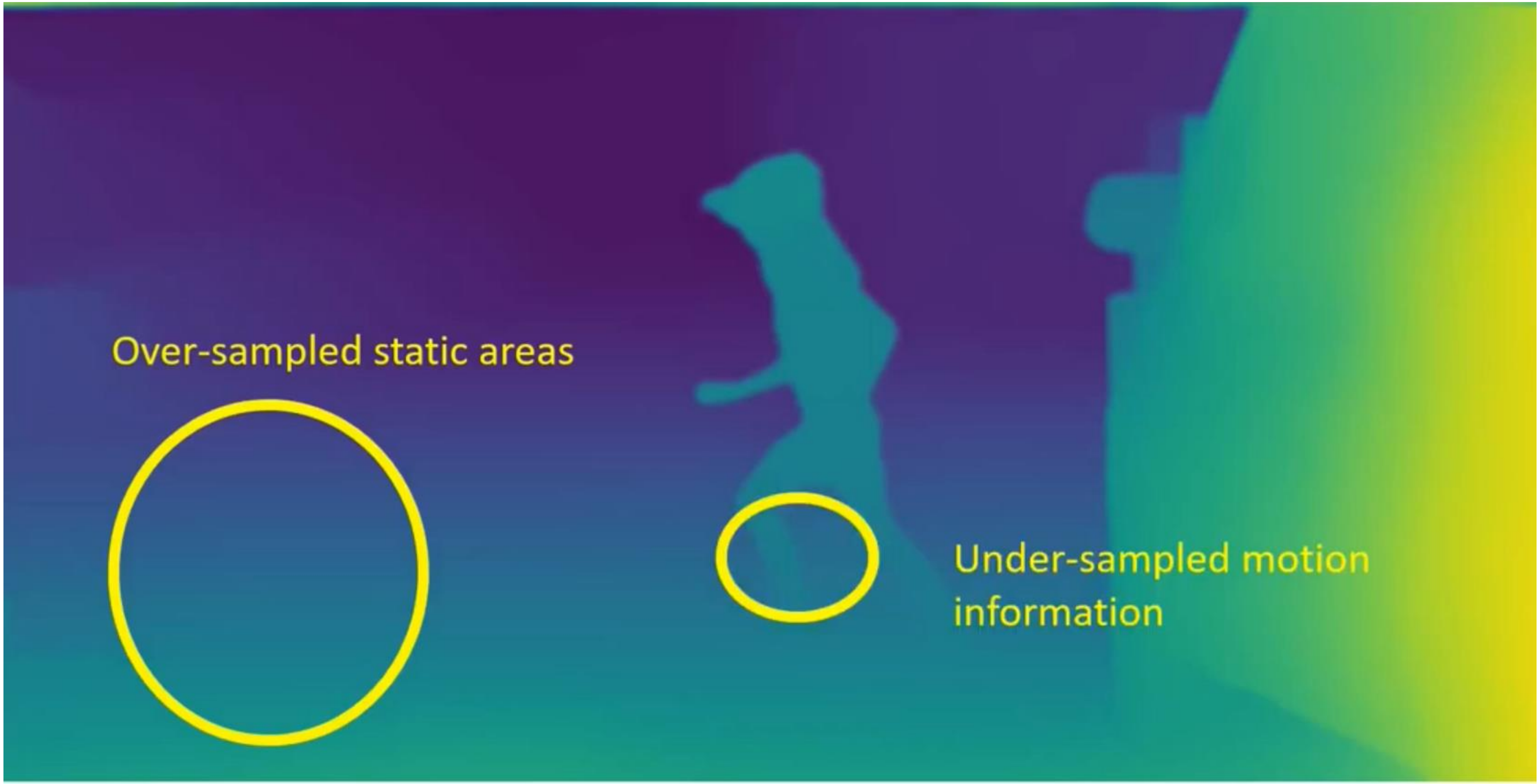


Image courtesy of Volvo

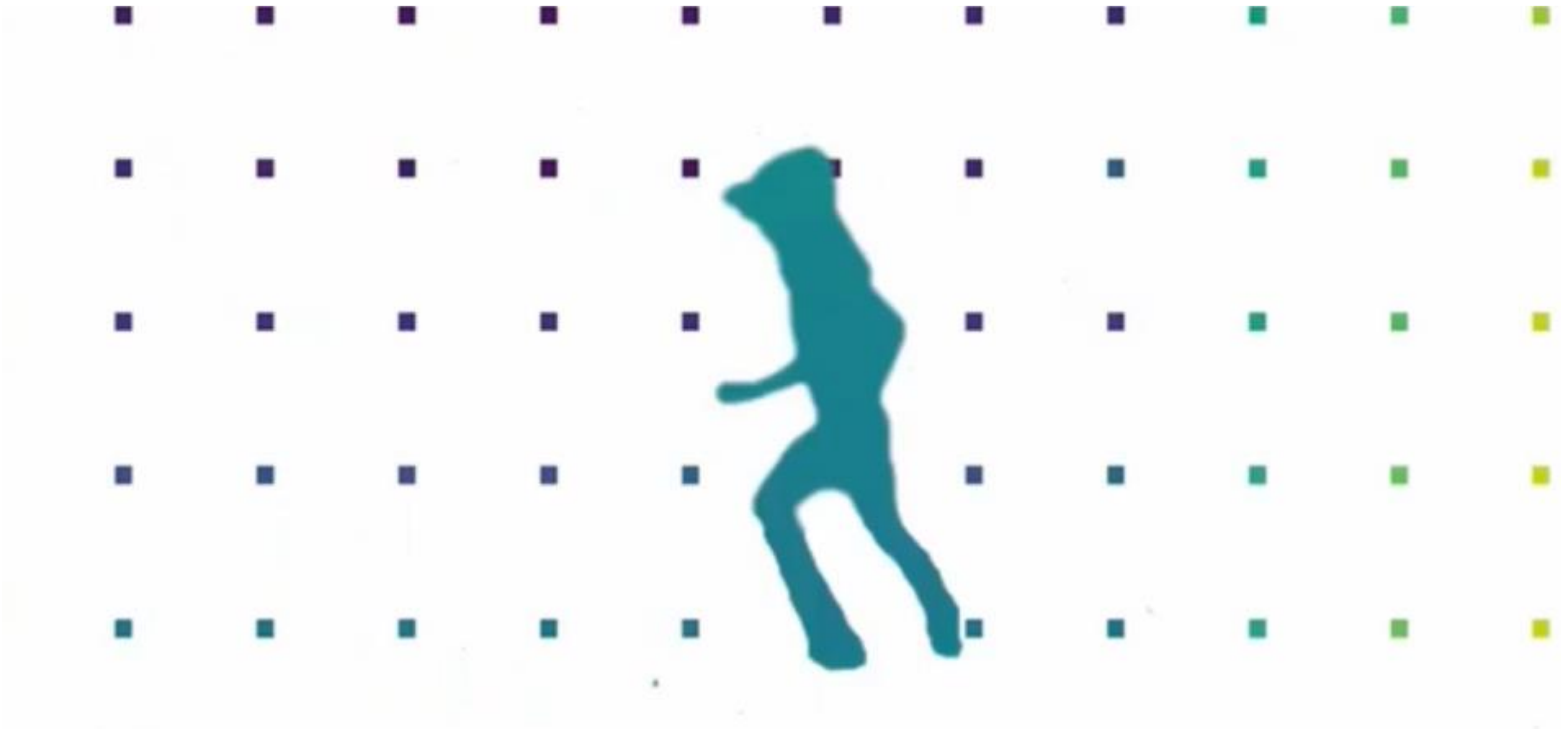
Muglikar, Gallego, Scaramuzza, *Event Guided Depth Sensing*, 3DV'21. PDF. Video

Motivation

Motivation

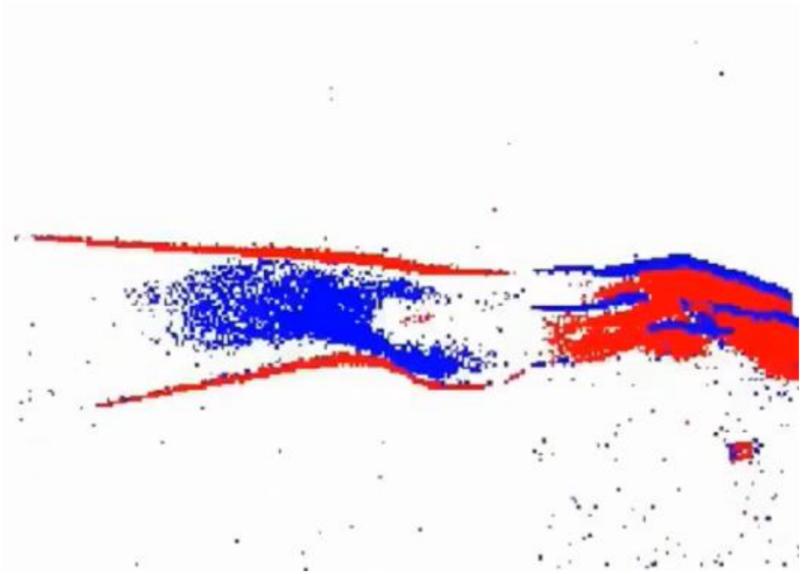

Over-sampled static areas
Under-sampled motion information

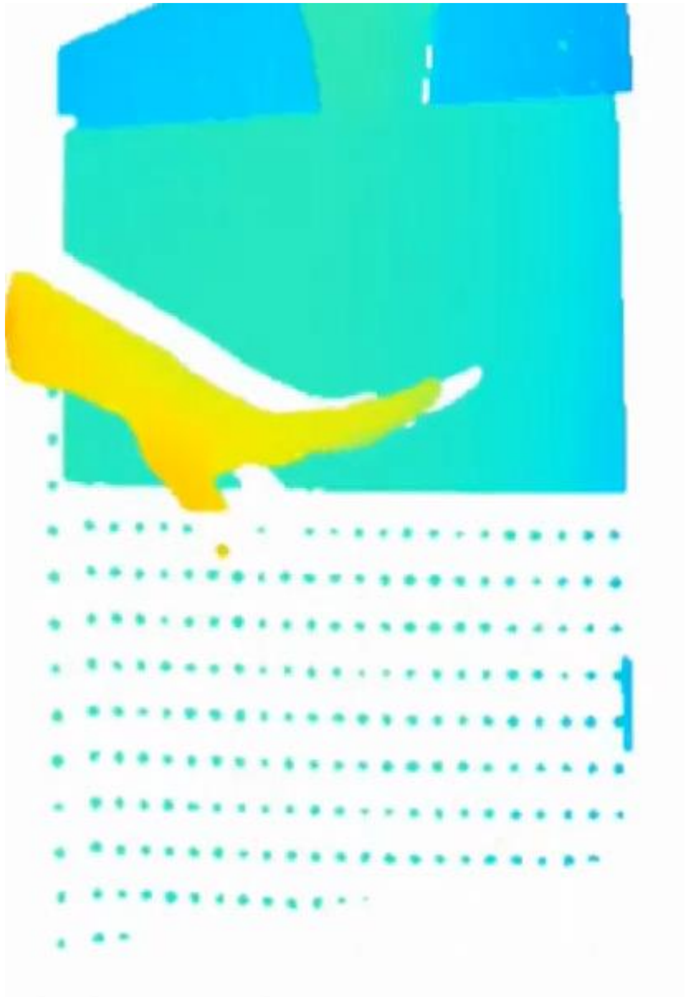
Muglikar, Gallego, Scaramuzza, *Event Guided Depth Sensing*, 3DV’21. [PDF](). [Video]()



Muglikar, Gallego, Scaramuzza, *Event Guided Depth Sensing*, 3DV’21. PDF. Video

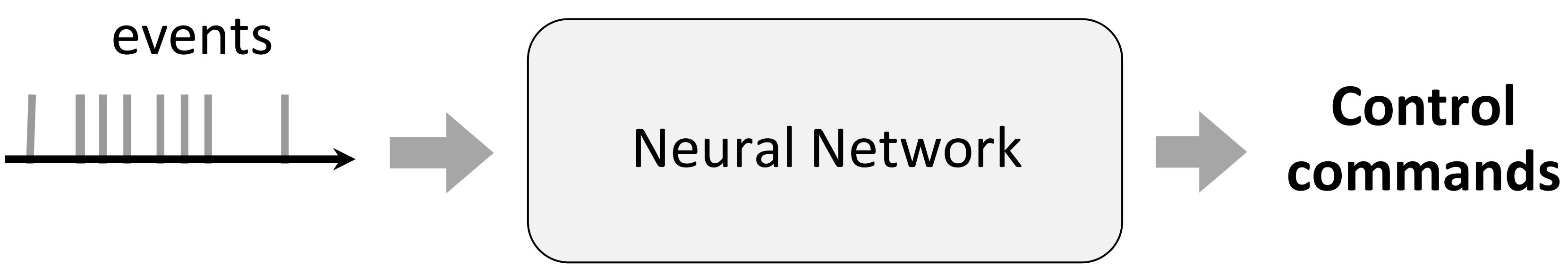
Results





Muglikar, Gallego, Scaramuzza, *Event Guided Depth Sensing*, 3DV’21. PDF. Video

Outlook: Event-driven Control on Neuromorphic Processors



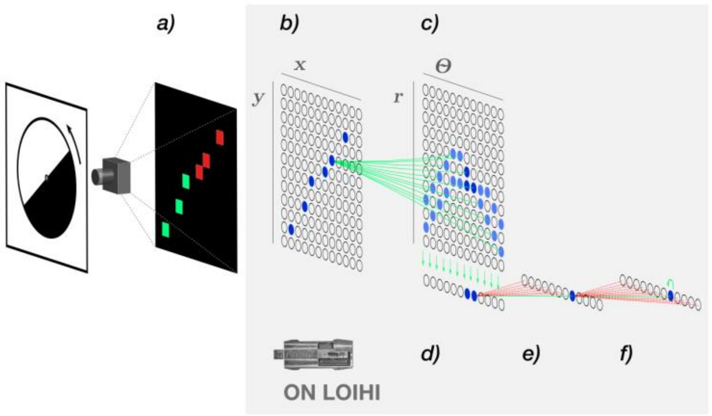
- Motivation: agile maneuvers require **low perception latency** and **high controller bandwidth**
- Goal: **map single events directly to control commands**
- Method: **Spiking Network** (SNN) running on **Intel Loihi neuromorphic chip**
- Advantage: **low-latency perception, high-bandwidth control**

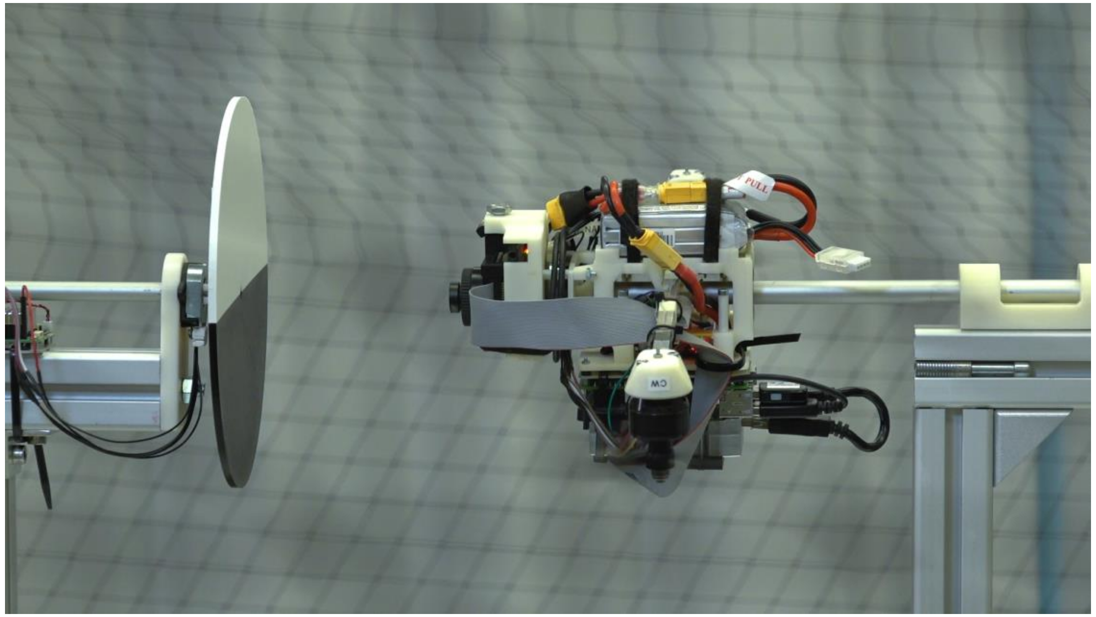
[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA'21. PDF. Video.
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA'20, PDF YouTube

Outlook: Event-driven Control on Neuromorphic Processors

events → **Neural Network** → **Control commands**

Reflex control task: mimic 1D rotation of a disc



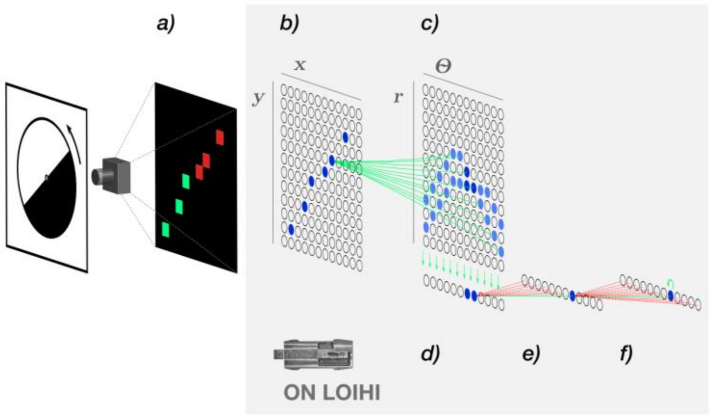


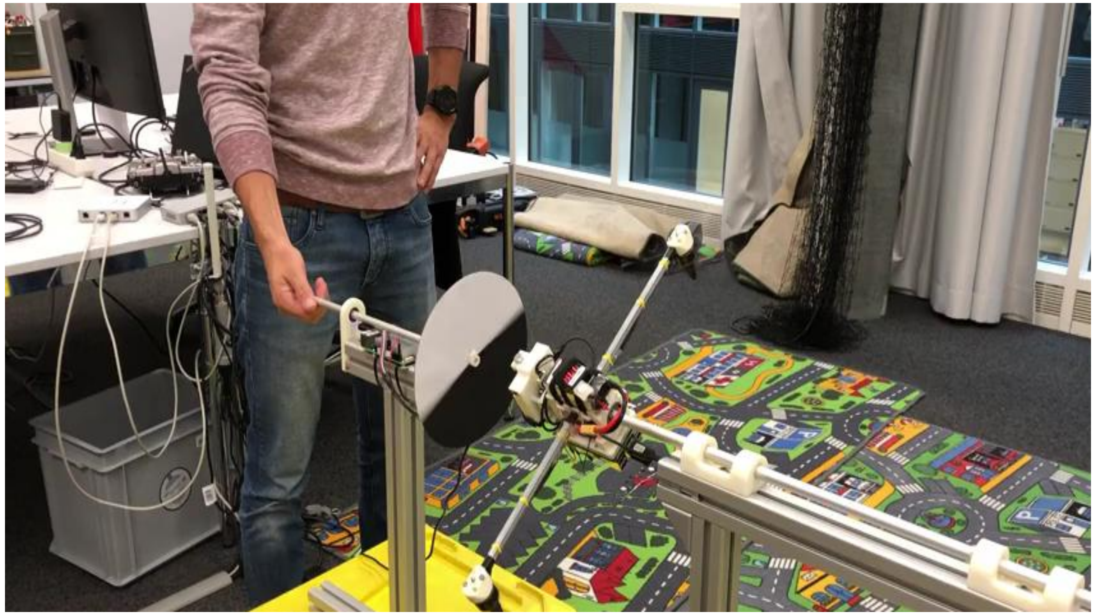
[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA'21. PDF. Video.
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA'20, PDF YouTube

Outlook: Event-driven Control on Neuromorphic Processors

events → **Neural Network** → **Control commands**

Reflex control task: mimic 1D rotation of a disc





[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA’21. PDF. Video.
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA’20, PDF YouTube

Thanks!

Code, Datasets, Simulator, Tutorials, Survey paper: http://rpg.ifi.uzh.ch/research_dvs.html

ailabRPG

@davsca1

@davidescaramuzza