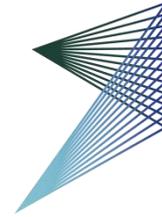




University of  
Zurich<sup>UZH</sup>

**ETH** zürich

Institute of Informatics – Institute of Neuroinformatics



ROBOTICS &  
PERCEPTION  
GROUP

# Event-based Cameras: Challenges and Opportunities

Davide Scaramuzza & Guillermo Gallego

Our research on Event cameras: [http://rpg.ifi.uzh.ch/research\\_dvs.html](http://rpg.ifi.uzh.ch/research_dvs.html)

Survey paper on Event cameras: <http://rpg.ifi.uzh.ch/docs/EventVisionSurvey.pdf>

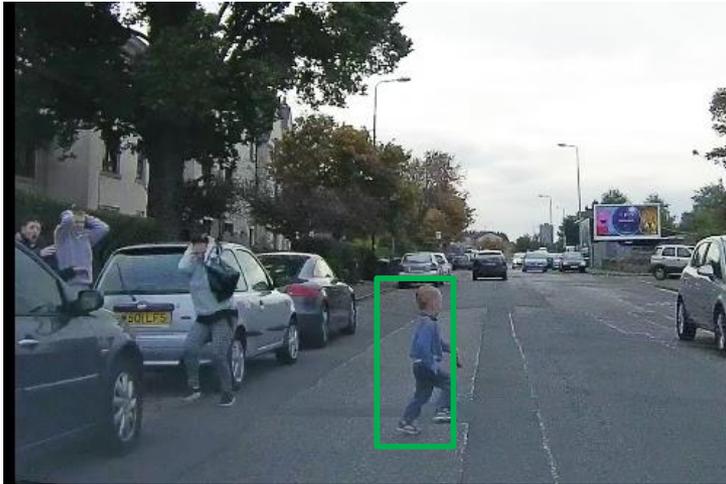
List of resources on event cameras (companies, code, datasets, drivers, etc.):

[https://github.com/uzh-rpg/event-based\\_vision\\_resources](https://github.com/uzh-rpg/event-based_vision_resources)

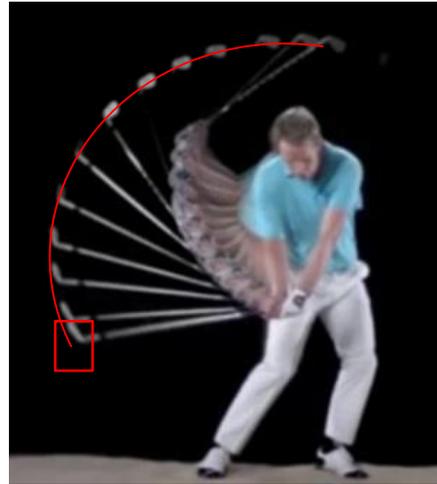
# Open Challenges in Computer Vision

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

**Latency**



**Motion blur**



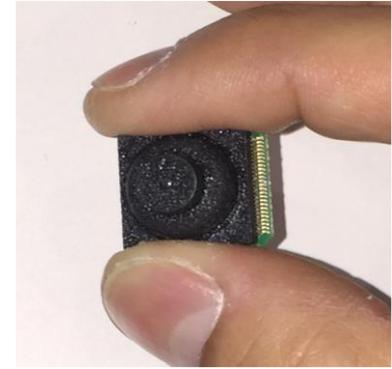
**Dynamic Range**



**Event cameras do not suffer from these problems!**

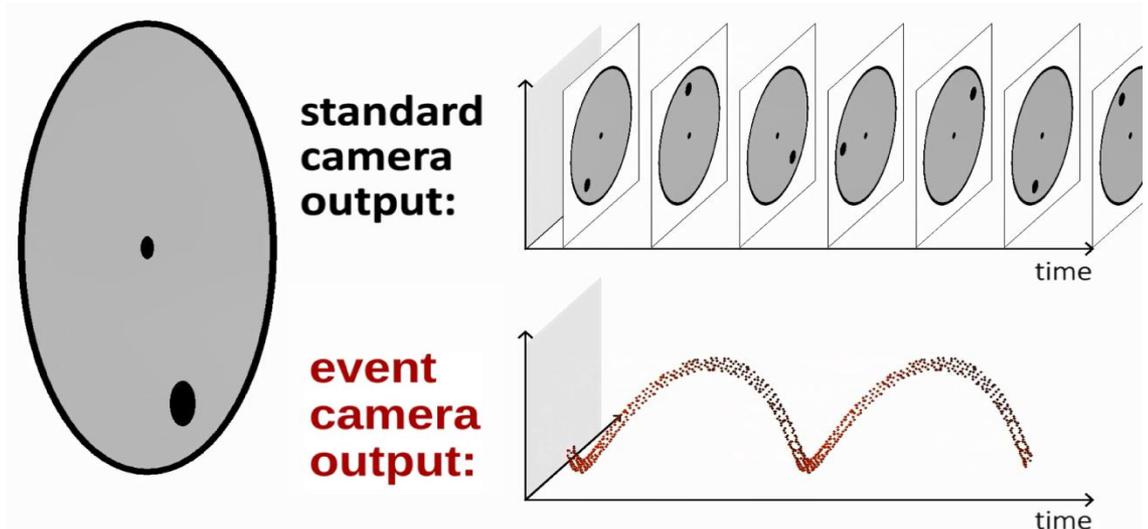
# What is an event camera?

- Novel sensor that measures only **motion in the scene**
- **First commercialized in 2008** by T. Delbruck (UZH&ETH) under the name of Dynamic Vision Sensor (DVS)
- **Low-latency** ( $\sim 1 \mu\text{s}$ )
- **No motion blur**
- **High dynamic range** (140 dB instead of 60 dB)
- **Ultra-low power** (10mW vs 1W)



Mini DVS sensor from IniVation.com

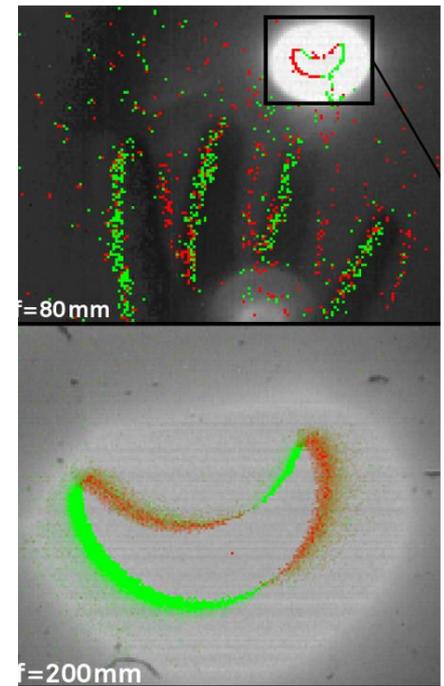
**Traditional vision algorithms cannot be used!**



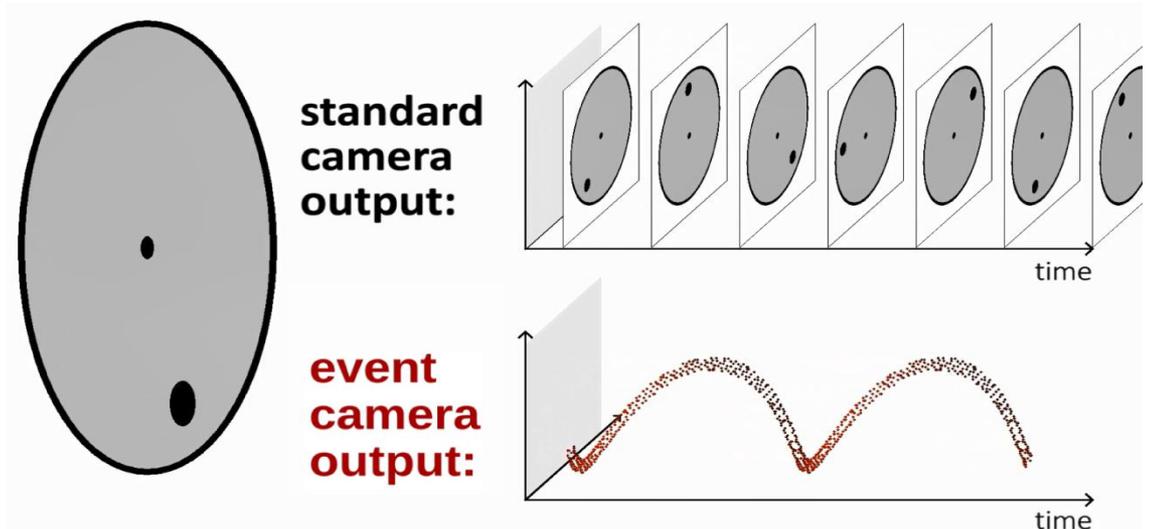
[Video animation here](#)

# What is an event camera?

- Novel sensor that measures only **motion in the scene**
- **First commercialized in 2008** by T. Delbruck (UZH&ETH) under the name of Dynamic Vision Sensor (DVS)
- **Low-latency** ( $\sim 1 \mu\text{s}$ )
- **No motion blur**
- **High dynamic range** (140 dB instead of 60 dB)
- **Ultra-low power** (10mW vs 1W)



Traditional vision algorithms cannot be used!



[Video animation here](#)

# The Challenge

**How can we unlock the capabilities of event cameras to design systems that are high-speed, high dynamic range, low latency while processing million events per second?**

# Event Camera Datasets & Simulator

- Datasets: [http://rpg.ifi.uzh.ch/davis\\_data.html](http://rpg.ifi.uzh.ch/davis_data.html)
- Event Camera Simulator (ESIM): <http://rpg.ifi.uzh.ch/esim.html>

Frame



Events



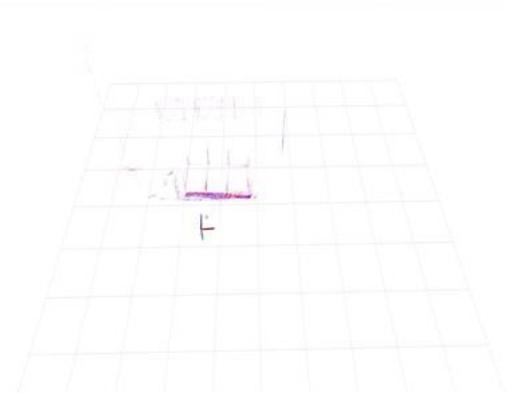
Color Frame



Depth



Optic Flow

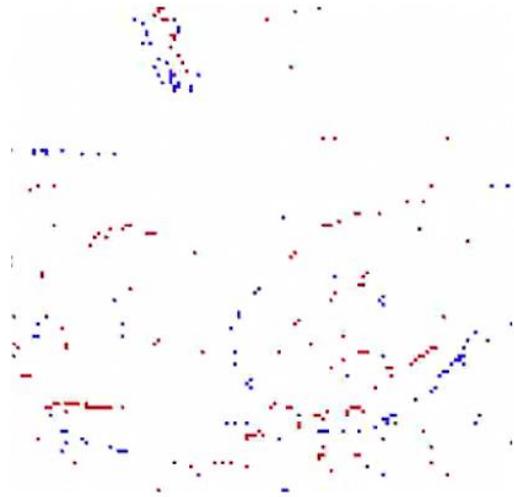


Point cloud

# Applications

# Low-latency, High-speed Camera Tracking

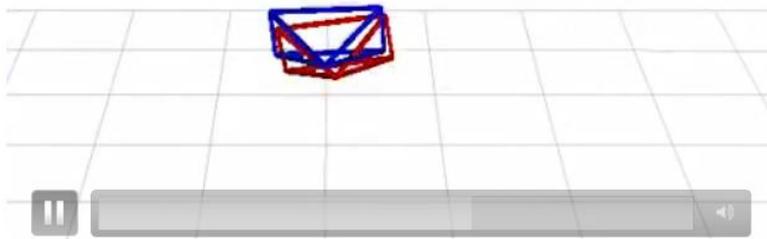
Event camera



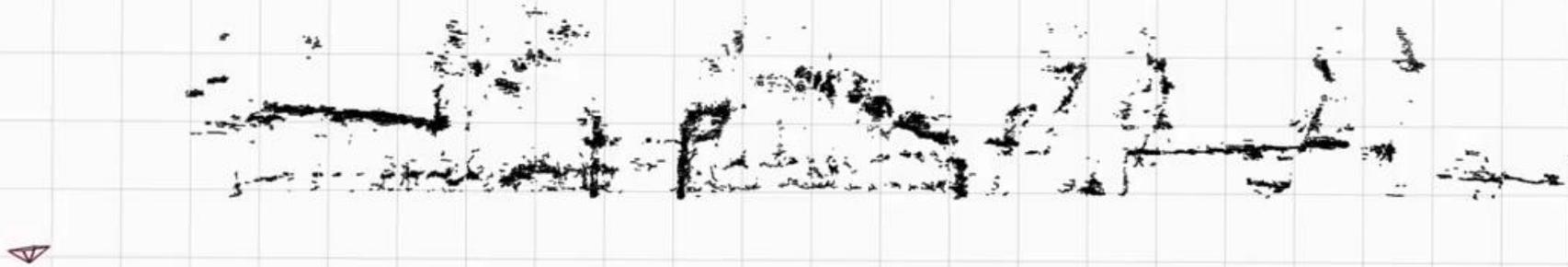
Standard camera



Event-based (EB)  
Frame-based (FB)

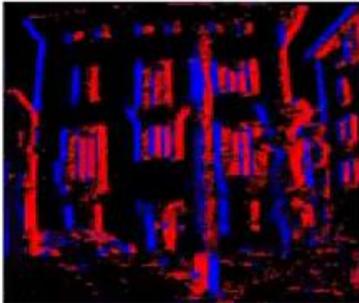


# 6-DOF, Event-based SLAM from a Train at 200km/h

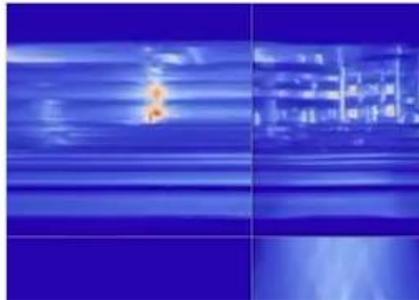


Video: <https://www.youtube.com/watch?v=fA4MiSzYHWA>

DVS  
Events



DSI



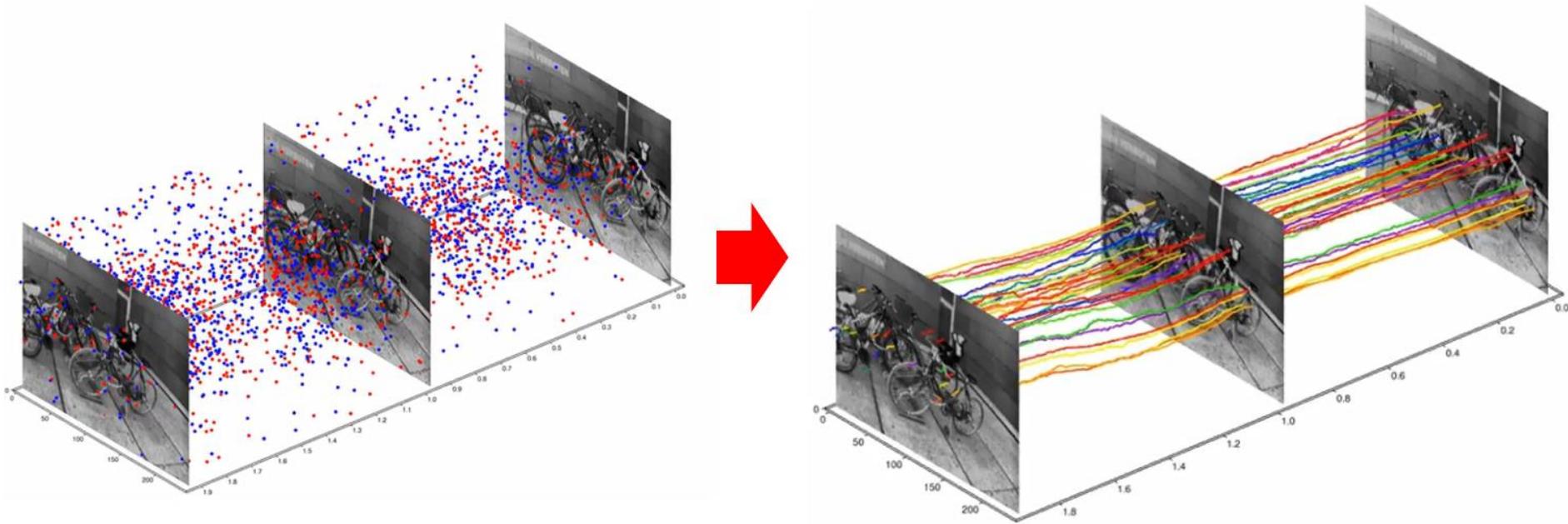
External  
camera



In collaboration with

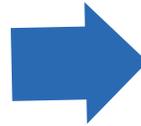
and Swiss train company, SBB

# Combining Standard and Event Cameras

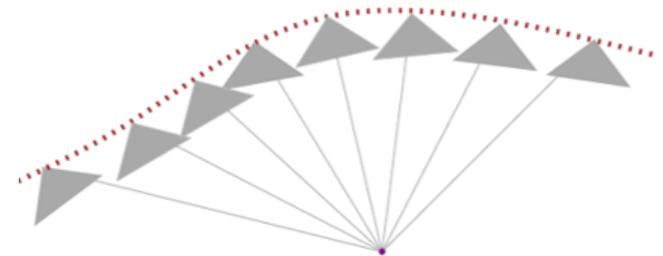
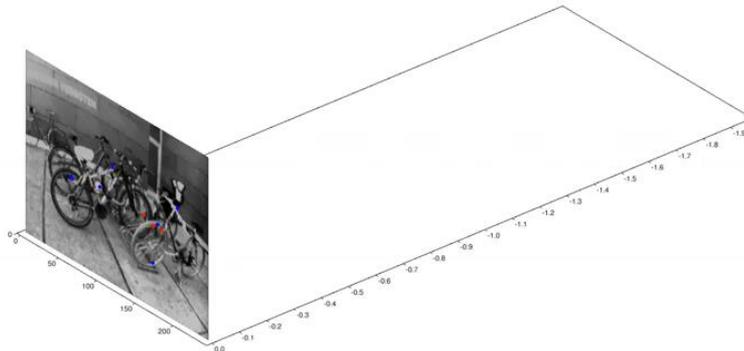


# UltimateSLAM: combines Events + Frames + IMU

**Front End:**  
Feature tracking from  
Events and Frames

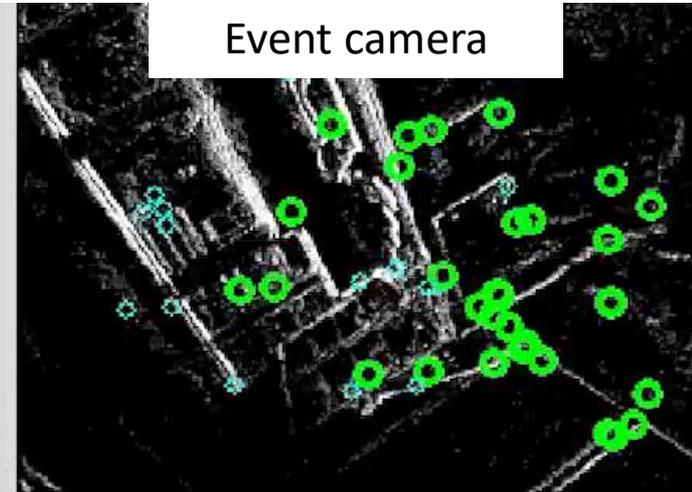


**Back-End**  
State-of-the-art  
Sliding-Window  
Visual-inertial Fusion

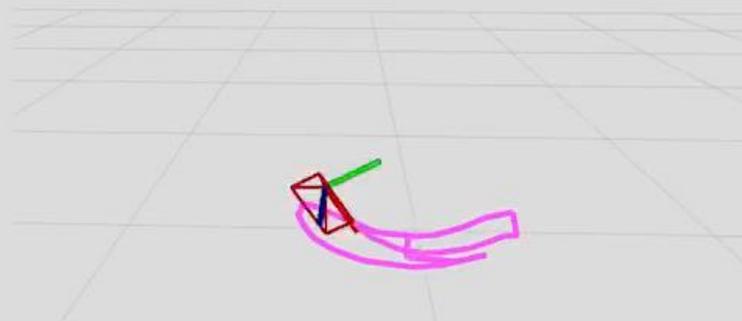


# UltimateSLAM: Frames + Events + IMU

85% accuracy gain over standard Visual-Inertial SLAM in HDR and high speed scenes



Front view



Top view

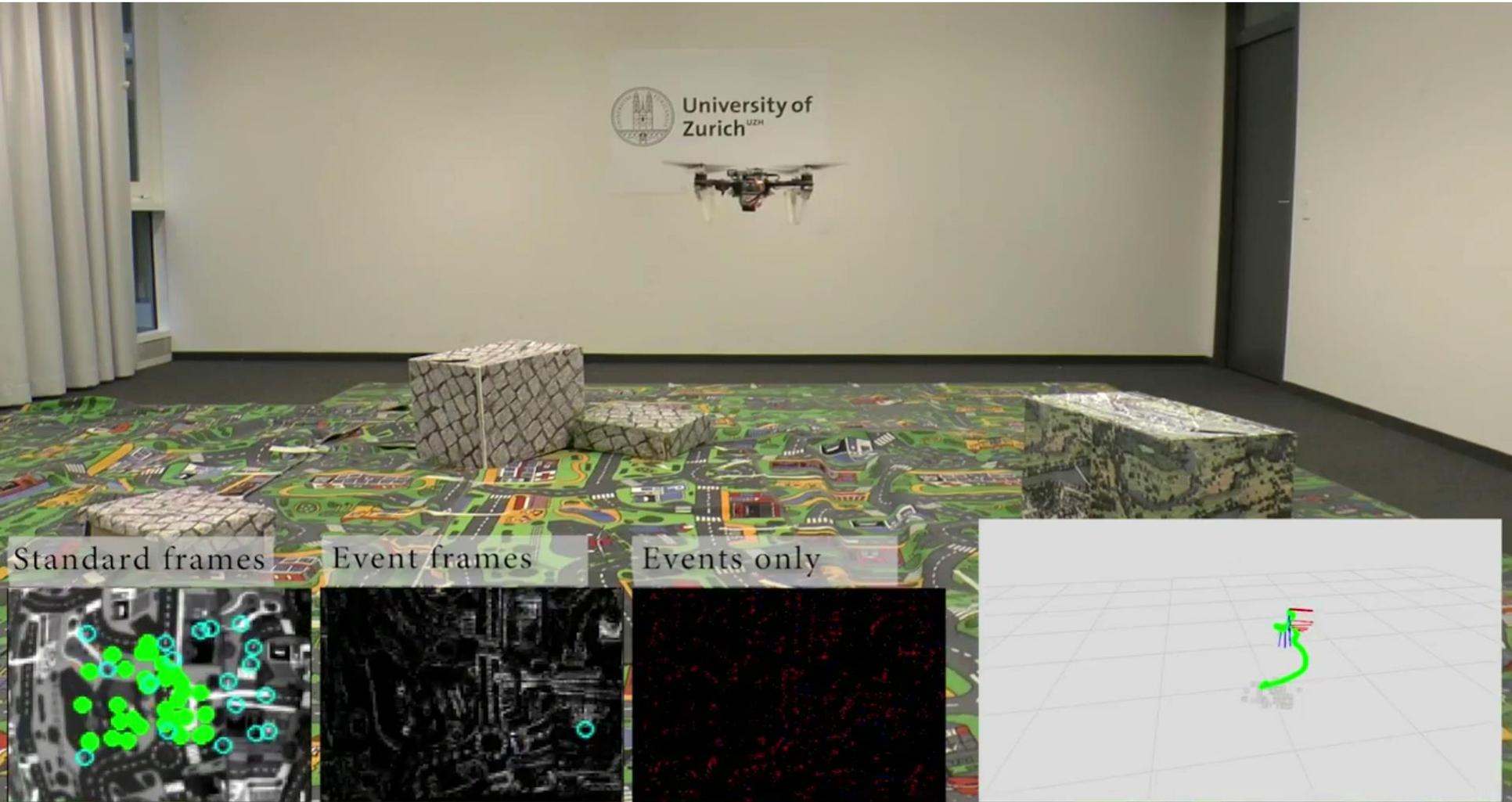


Rosinol et al., Ultimate SLAM? **RAL'18** – Best RAL'18 Paper Award Honorable Mention [PDF](#). [Video](#). [IEEE Spectrum](#).

Mueggler et al., Continuous-Time Visual-Inertial Odometry for Event Cameras, **TRO'18**. [PDF](#)

# Autonomous Drone Flight in Low Light

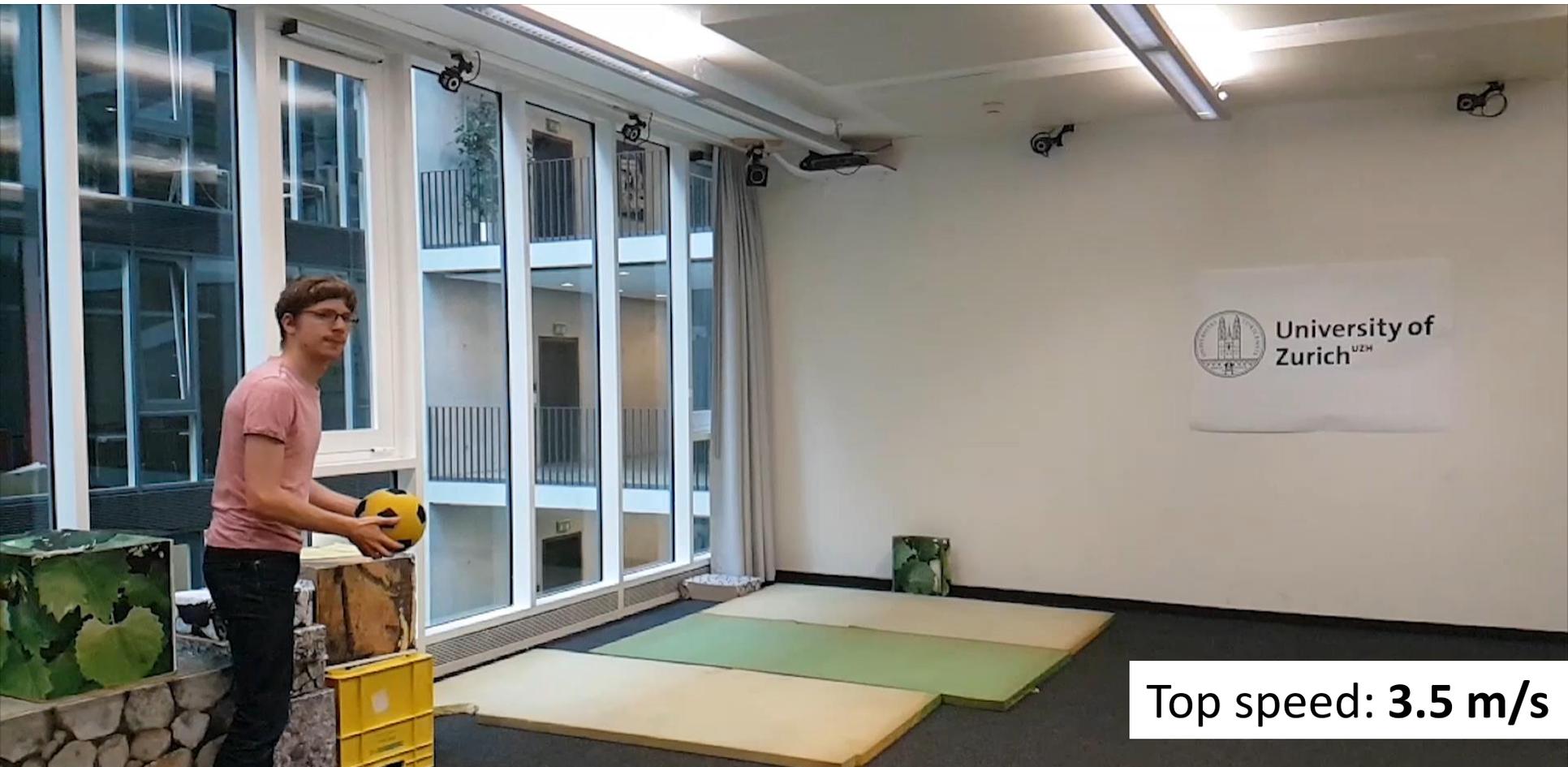
UltimateSLAM runs in real time on a smartphone processor.



# Dynamic Obstacle Detection & Avoidance

# Event-based Dynamic Obstacle Detection & Avoidance

- Top speed: **3.5 m/s**
- Object detection runs at 100Hz onboard



Top speed: **3.5 m/s**

Falanga et al. *How Fast is too fast? The role of perception latency in high speed sense and avoid*, RAL'19.

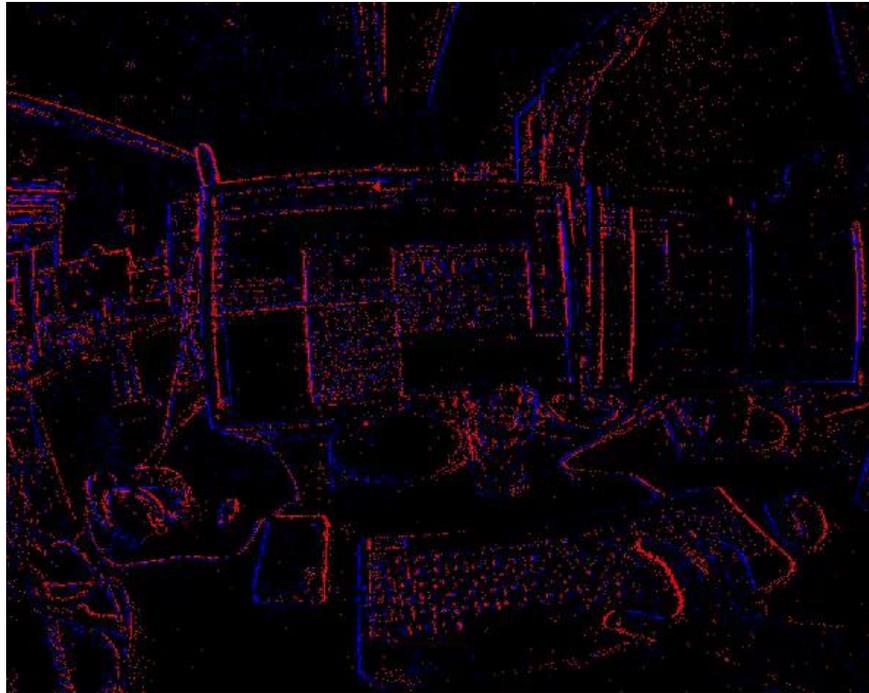
[PDF](#). [Video](#). Featured in [IEEE Spectrum](#).

# High Speed and High Dynamic Range Video Reconstruction with an Event Camera

Rebecq et al., “Events-to-Video: Bringing Modern Computer Vision to Event Cameras”, CVPR19. [PDF Video](#).  
Rebecq et al., “High Speed and High Dynamic Range Video with an Event Camera”, arXiv, 2019. [PDF Video Code](#)

# Image Reconstruction from Events

Events



Reconstructed image from events

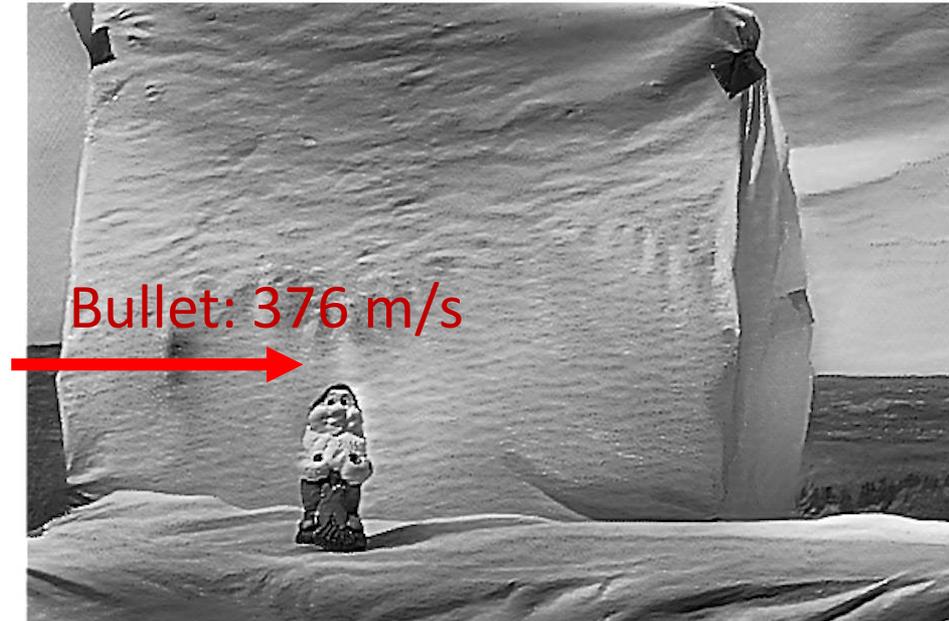


Rebecq et al., "Events-to-Video: Bringing Modern Computer Vision to Event Cameras", CVPR19. [PDF](#) [Video](#).  
Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", arXiv, 2019. [PDF](#) [Video](#) [Code](#)

# High Framerate Video (learned from simulation)



Huawei P20 Pro (240 FPS)



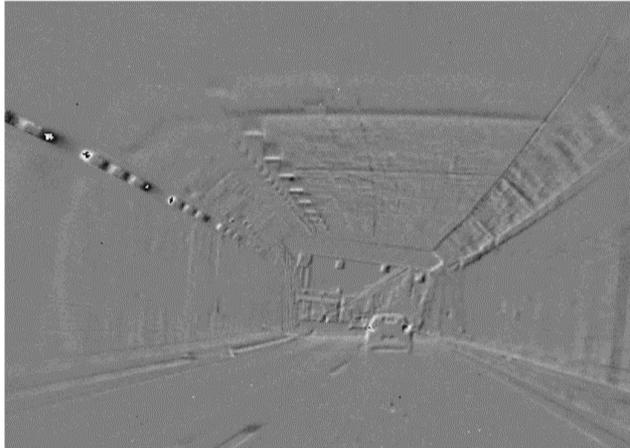
Our reconstruction (5400 FPS)

Real time

# HDR Video

## Driving out of a tunnel

### Driving out of a tunnel



**Events**



**Our reconstruction**



**Phone camera**

# Downstream applications (object detection)

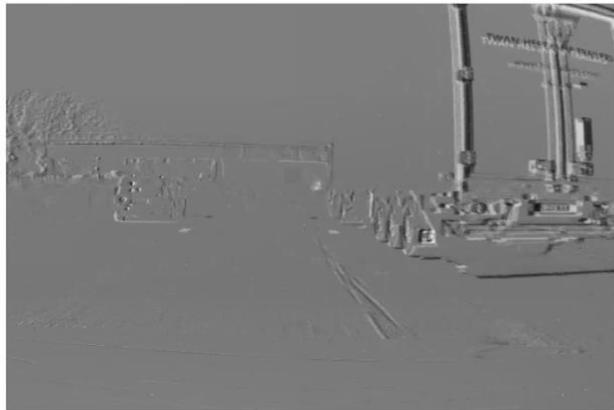


Events



Our reconstruction + object detections (YOLOv3)

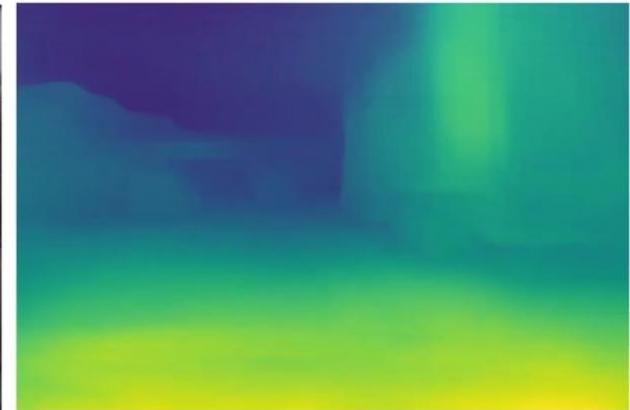
# Monocular Depth Estimation



Events



Our reconstruction



Monocular depth

# Conclusion

- **Event cameras** have many **advantages**: high dynamic range (HDR), high speed, low latency, low power.
- **Opportunities** in AR/VR, automotive, drones, IoT, ...

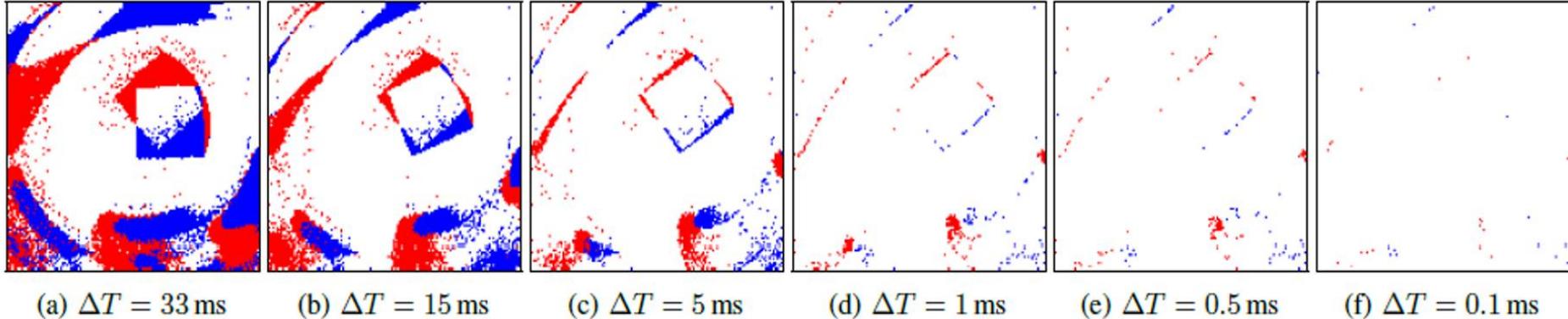
## Open Problems:

- How to **optimally process events** for a given task?
- **Noise modeling** and **sensor characterization**.
- How to perform **asynchronous learning** and inference?
- How to optimally combine events with **dedicated hardware**?
- **Low-latency estimation and control**.
- How to optimally **tune sensor parameters** for a given task?

How to Process the Events?

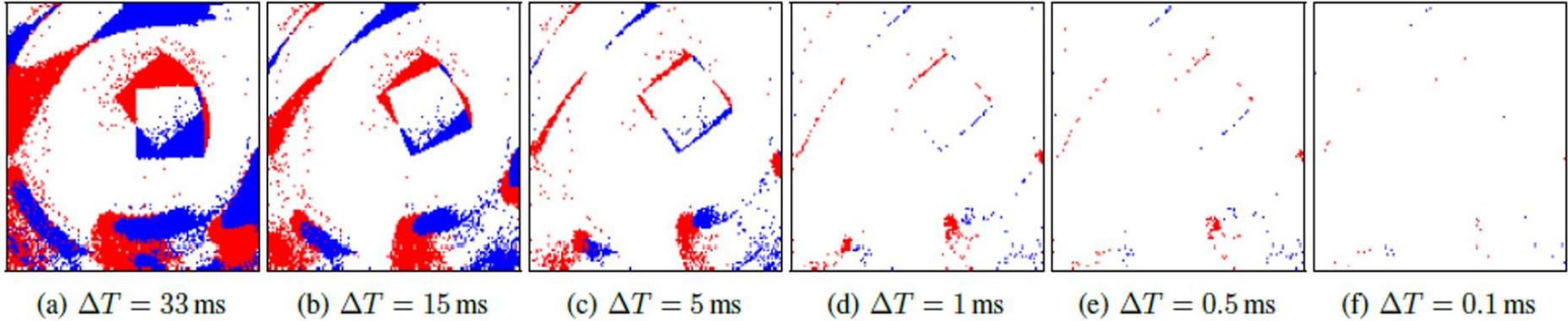
Event-by-event vs. event packets

# How many events should be used?

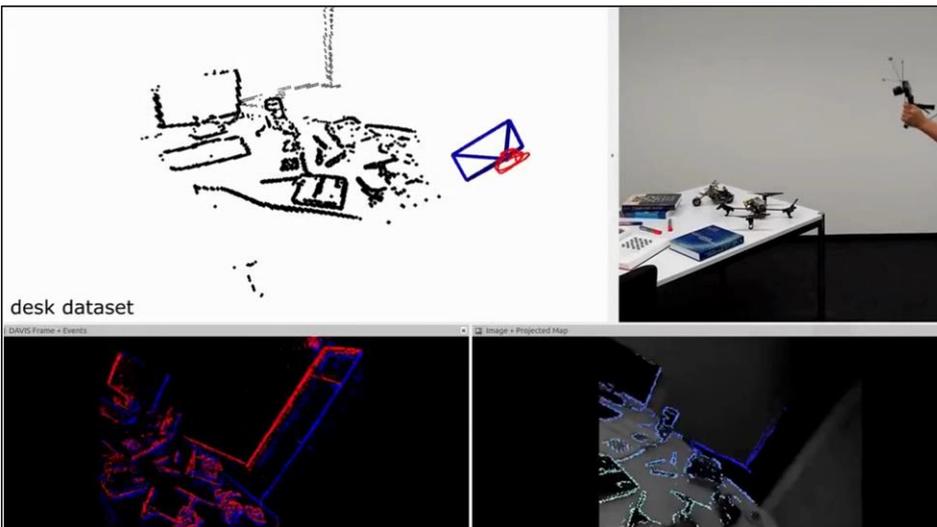


- **Event-by-event processing** (i.e., estimate unknown parameters **event by event**)
  - **Pros:** low latency (in principle down to microseconds)
  - **Cons:** with high-speed motion → millions of events per seconds → GPU
- **Event-packet processing** (i.e., process the last  $N$  events)
  - **Pros:**  $N$  can be tuned to allow real-time performance on a CPU
  - **Cons:** no longer microsecond resolution (when is it needed?)

# How many events should be used?

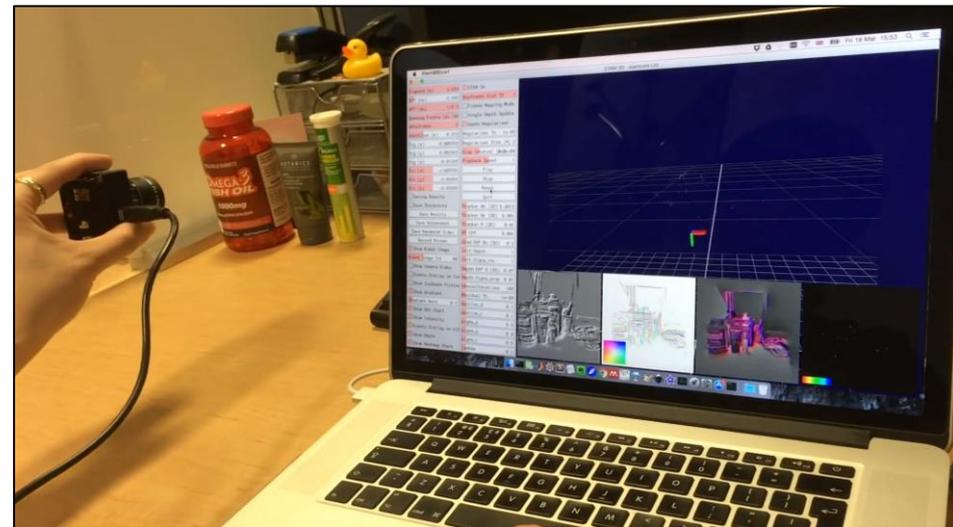


## Event-packet



Rebecq et al., RA-L 2017

## Event-by-event

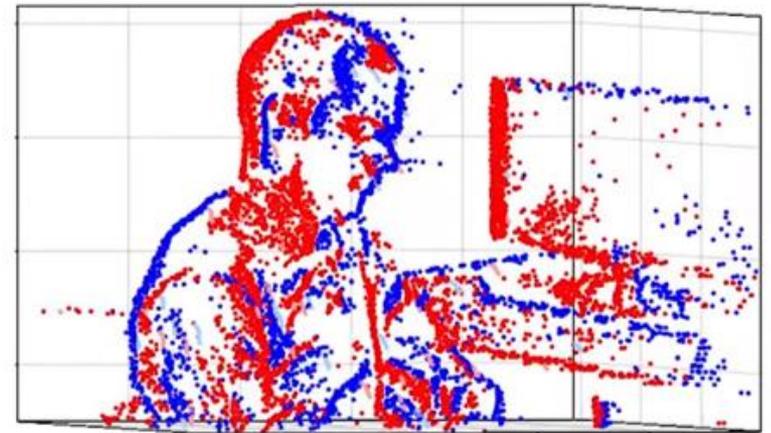
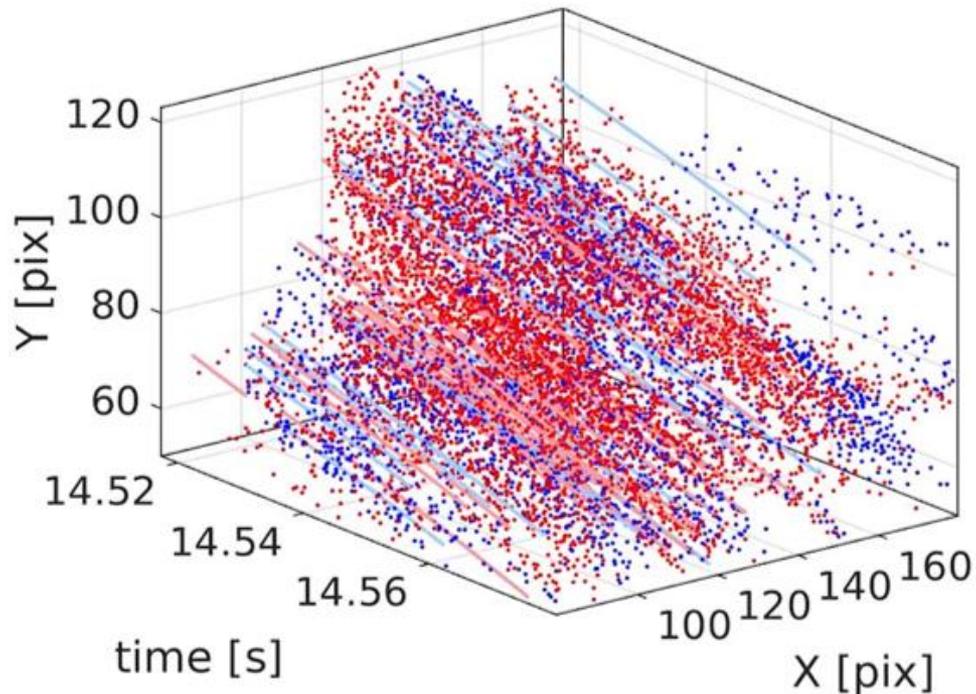


Kim et al., ECCV 2016

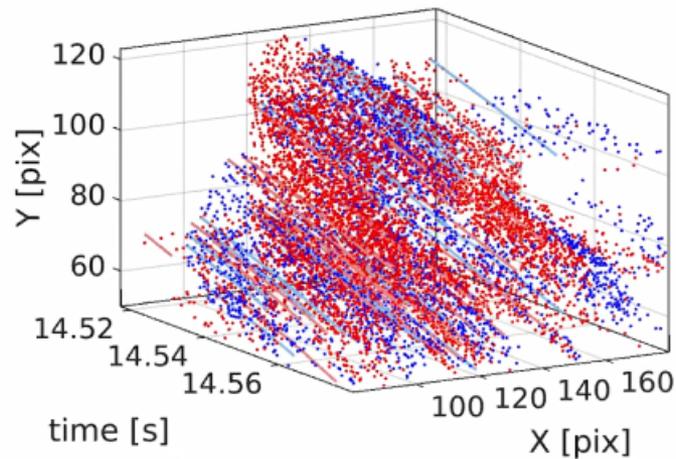
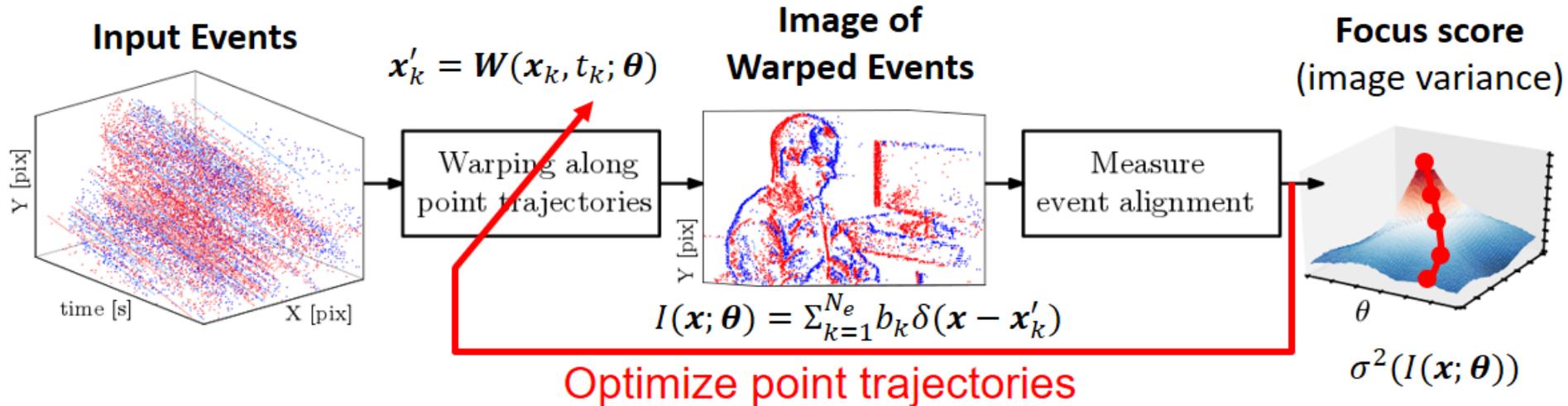
# Motion Estimation by Focus Maximization

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](#).

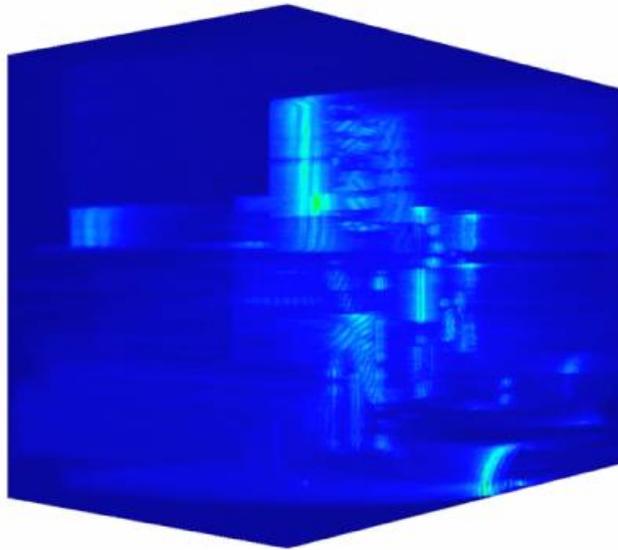
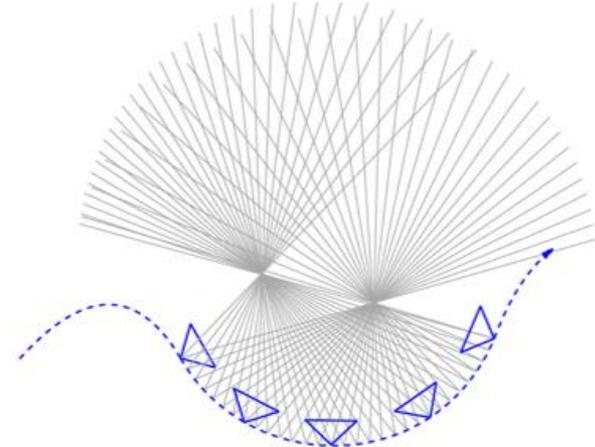
Idea: Warp spatio-temporal volume of events to **maximize focus** (e.g., sharpness) of the resulting image



# Focus Maximization Framework



# Example 1: Focus Maximization. 3D reconstruction



# Example 2: Motion Segmentation

Conventional Frames

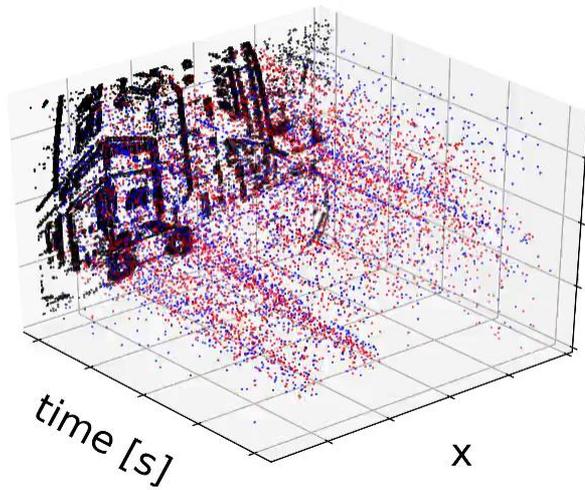


**Sequence: Buildings and car**

Motion-Compensated Segmented Events



Events



# Example 2: Motion Segmentation

Conventional Frames



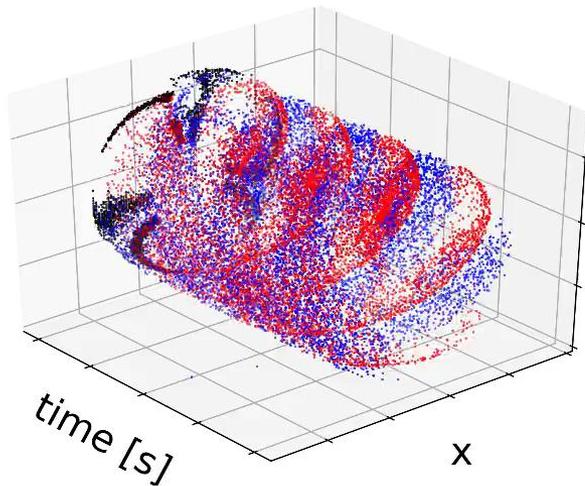
## Sequence: Fan and Coin

One motion model is used per cluster; one for the fan, modelling rotation, one for the coin, modelling optic flow

Motion-Compensated Segmented Events

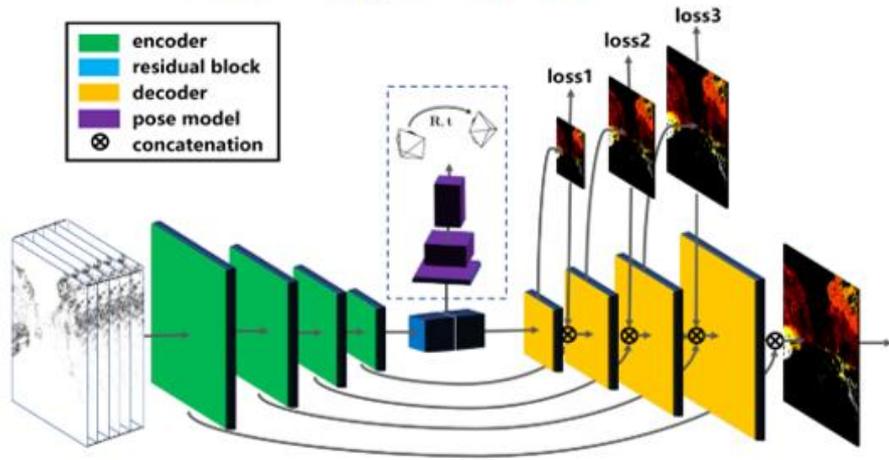


Events

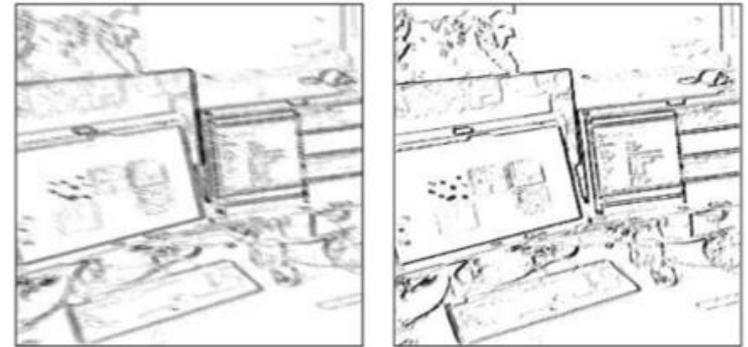


# Focus as Loss Function for Unsupervised Learning

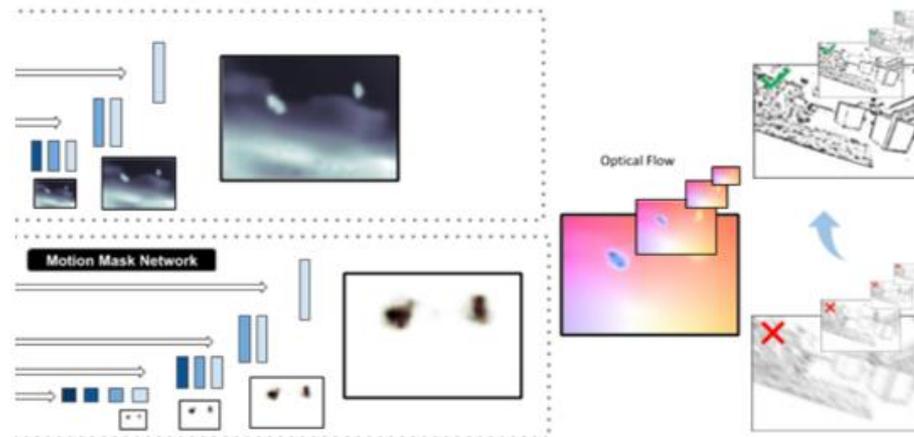
Zhu et al., CVPR 2019



Gallego et al., CVPR 2019

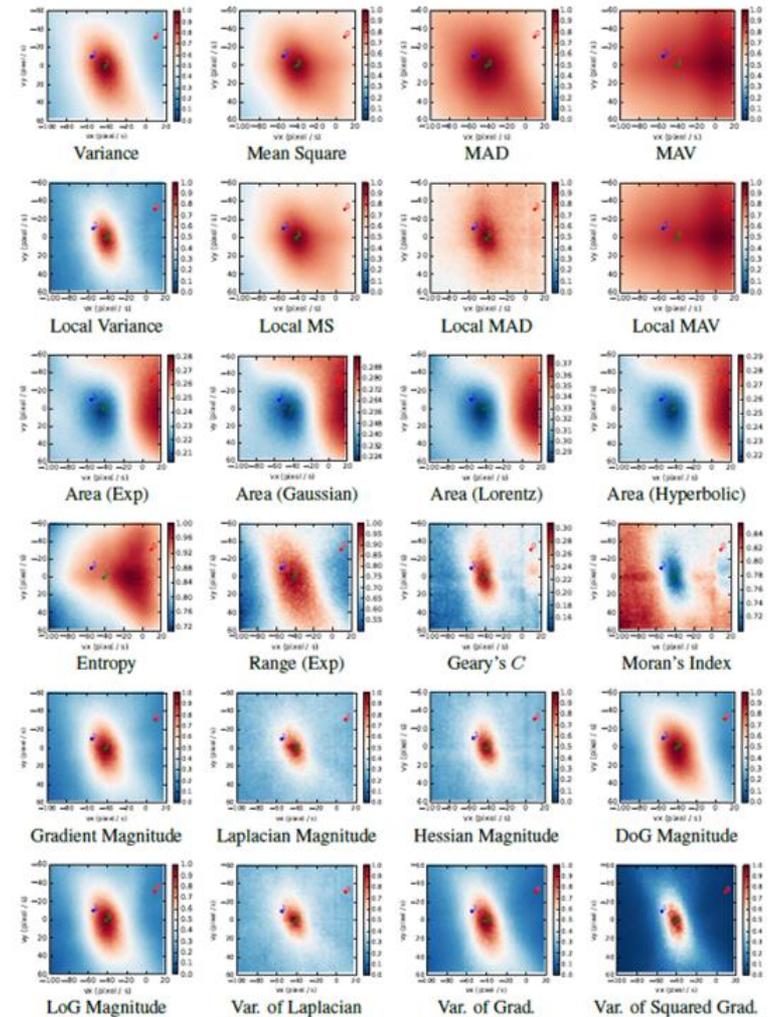


Mitrokhin et al., arXiv 2019



# Focus as Loss Function for Unsupervised Learning

- We proposed **22 focus loss functions**
- Focus is the “data fidelity” term



# To Recap

## ➤ **Focus Maximization** is:

- **Simple:** find point trajectories that achieve **event alignment**.
- **Effective:** solves **multiple computer vision tasks**.
- **Best Focus Metrics:** Variance, Magnitude Gradient and Magnitude of Laplacian.

# Research Challenges with Event Cameras

- Quantify the **trade-offs**:
  - Latency vs. power consumption and accuracy
  - Sensitivity vs. bandwidth and processing capacity
- **Active parameter adaptation**
- **Hardware**: pairing event cameras with dedicated hardware (SNN hardware, e.g., Intel Loihi, aiCTX Speck)
- **Learning** with event cameras:
  - How do we **exploit knowledge from image-based learning to event cameras**?
  - **Asynchronous** inference
  - **Where do we find learning data?** Event data is much more rare than frames. Potential solutions: unsupervised Learning, learning in simulation, transfer learning from frames to events

# Resources

- Our research on Event cameras: [http://rpg.ifi.uzh.ch/research\\_dvs.html](http://rpg.ifi.uzh.ch/research_dvs.html)
- Survey paper on Event cameras:  
<http://rpg.ifi.uzh.ch/docs/EventVisionSurvey.pdf>
- List of resources on event cameras (companies, code, datasets, drivers, etc.):  
[https://github.com/uzh-rpg/event-based\\_vision\\_resources](https://github.com/uzh-rpg/event-based_vision_resources)