University of Zurich[UZH] | ETH zürich

Institute of Informatics – Institute of Neuroinformatics

Event Cameras

Davide Scaramuzza

Slides, Publications, Videos, Code: http://rpg.ifi.uzh.ch/

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

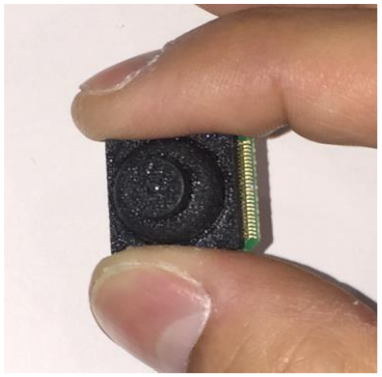
Motion blur

Dynamic Range

Event cameras do not suffer from these problems!

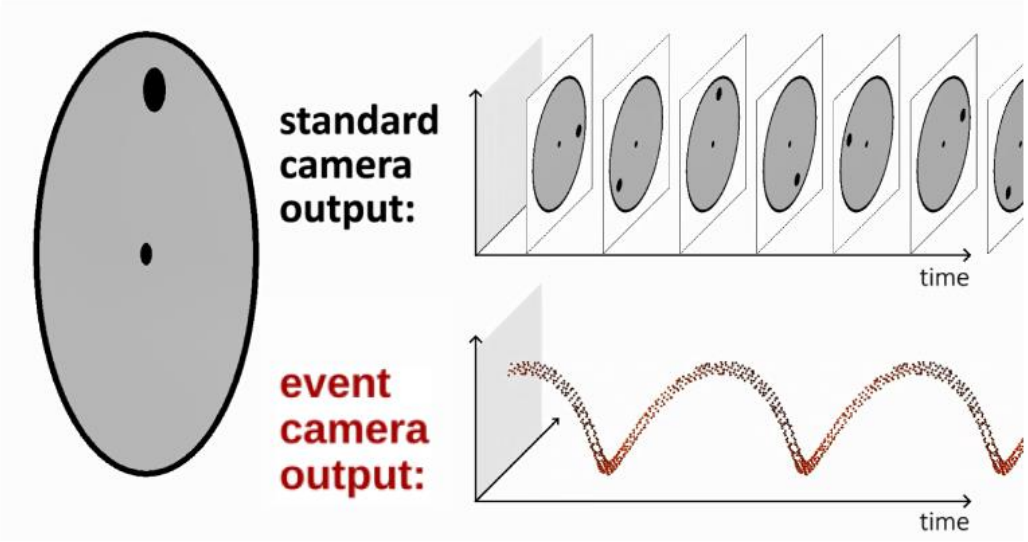
What is an event camera?

- Novel sensor that measures only **motion in the scene**
- First commercialized in 2010 by T. Delbruck (UZH&ETH) under the name of Dynamic Vision Sensor (DVS)
- **Low-latency** (~ 1 μs)
- **No motion blur**
- **High dynamic range** (140 dB instead of 60 dB)
- **Ultra-low power** (10mW vs 1W)



Mini DVS sensor from IniVation.com

Because the output is asynchronous, traditional vision algorithms cannot be used!



https://youtu.be/LauQ6LWTkxM?t=30

What is an event camera?

- Novel sensor that measures only **motion in the scene**
- First commercialized in 2010 by T. Delbruck (UZH&ETH) under the name of Dynamic Vision Sensor (DVS)
- **Low-latency** (~ 1 μs)
- **No motion blur**
- **High dynamic range** (140 dB instead of 60 dB)
- **Ultra-low power** (10mW vs 1W)

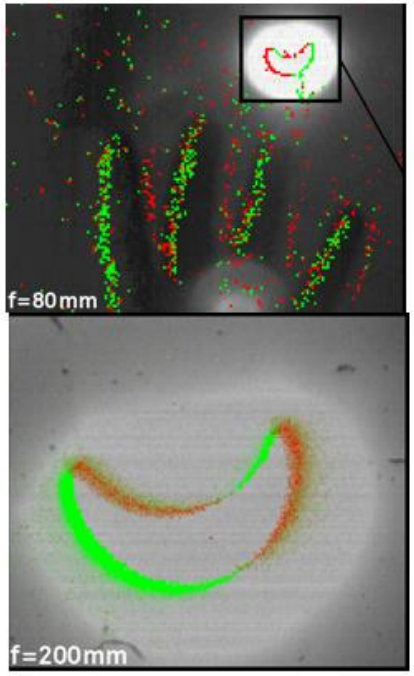
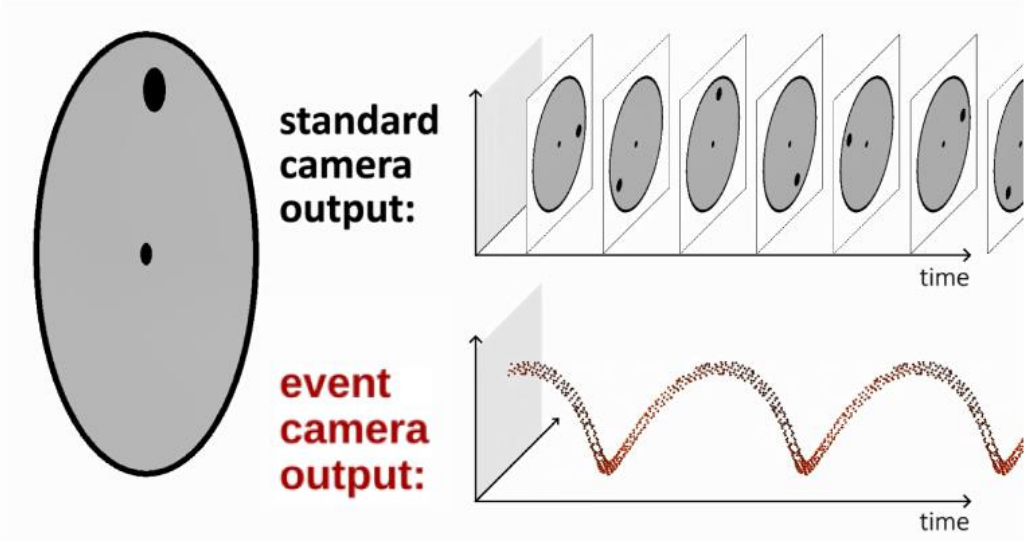


Image of the solar eclipse captured by a DVS

Because the output is asynchronous, traditional vision algorithms cannot be used!

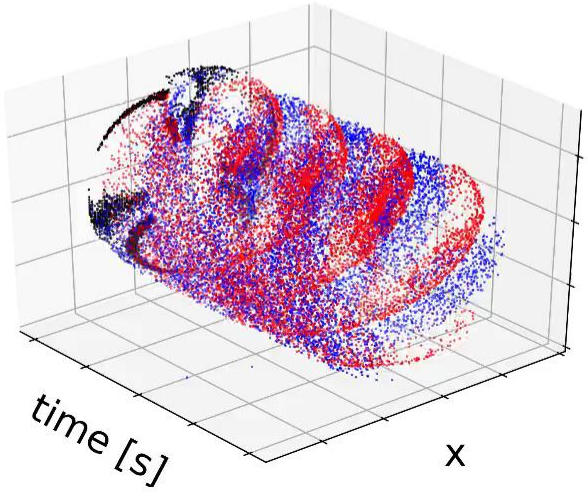


https://youtu.be/LauQ6LWTkxM?t=30

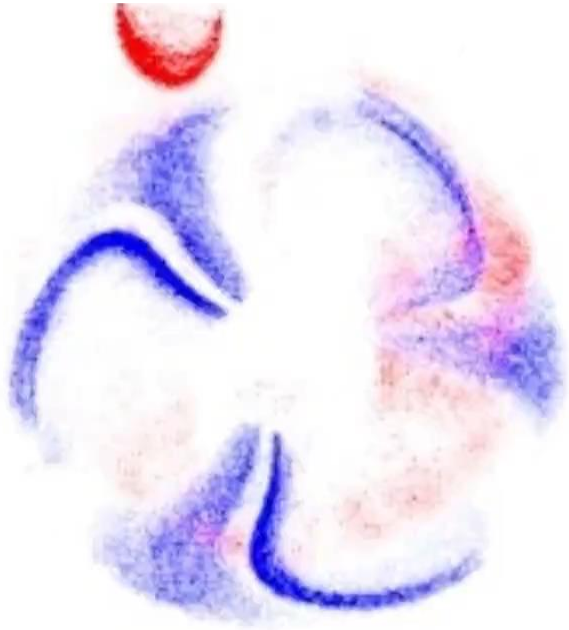
Sequence: Fan and Coin

Conventional Frames

Events



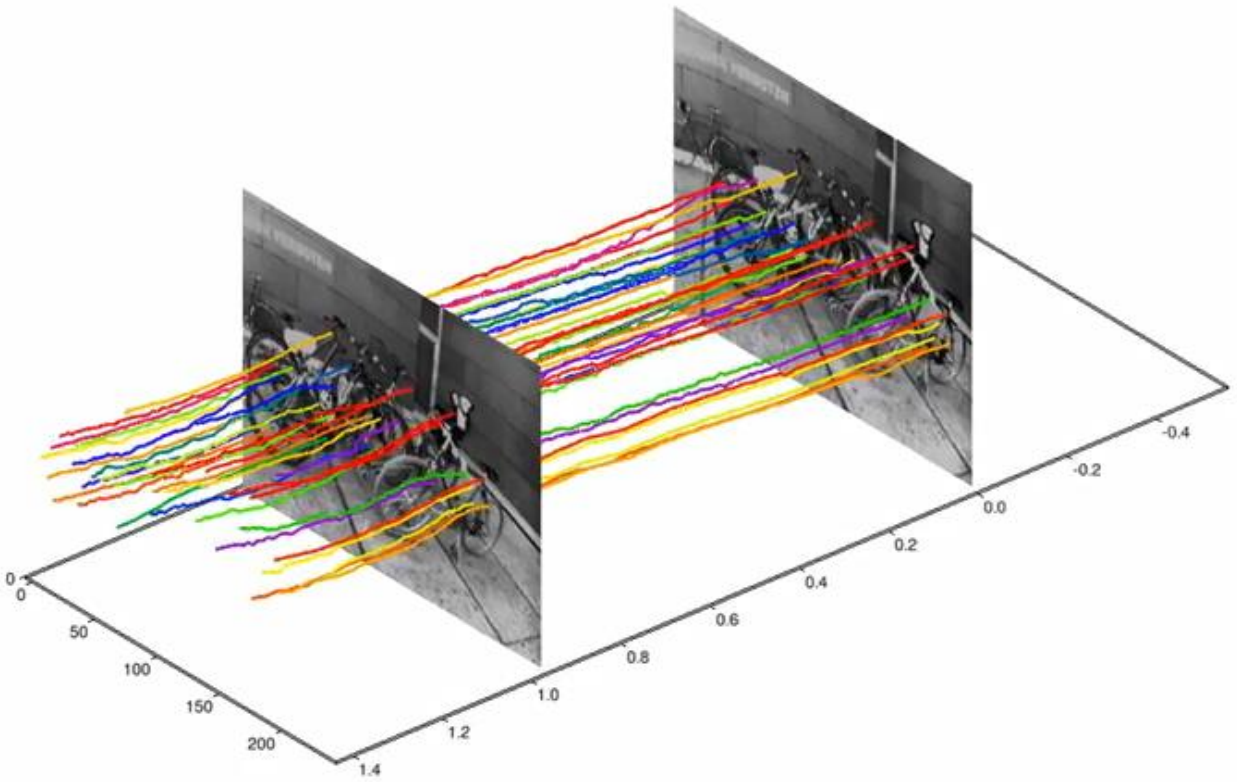
Events in the **space-time** domain (x, y, t)



Events in the **image domain** (x, y)
Integration time: ΔT = 10 ms

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

Application 1: Event-based Lucas-Kanade Tracking (E-KLT)



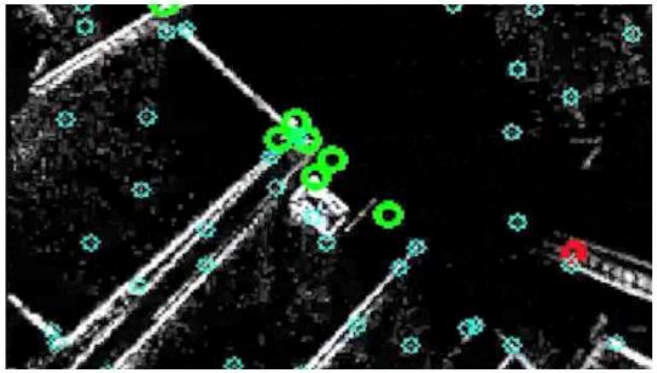
Gehrig et al., EKLT: Asynchronous, Photometric Feature Tracking using Events and Frames, IJCV'19**.** PDF. Video.

“UltimateSLAM”: Frames + Events + IMU

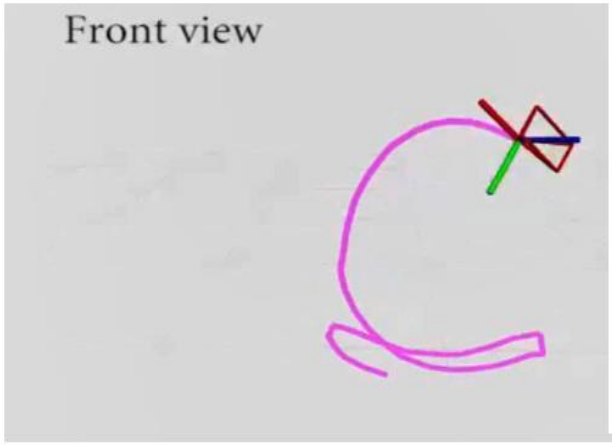
85% accuracy gain over standard visual-inertial SLAM in HDR and high speed scenes!

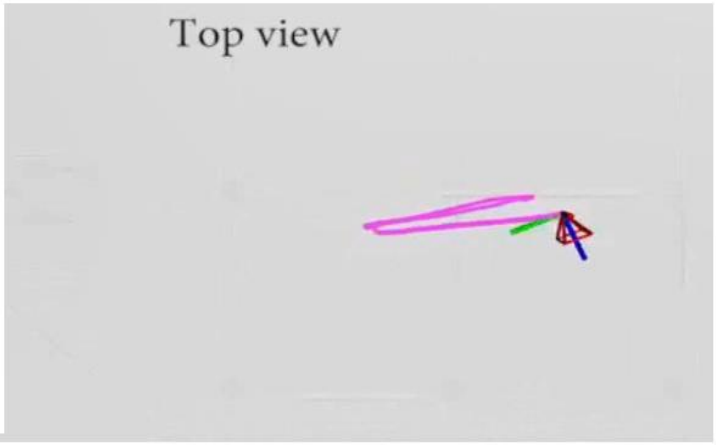
Standard camera

Event camera



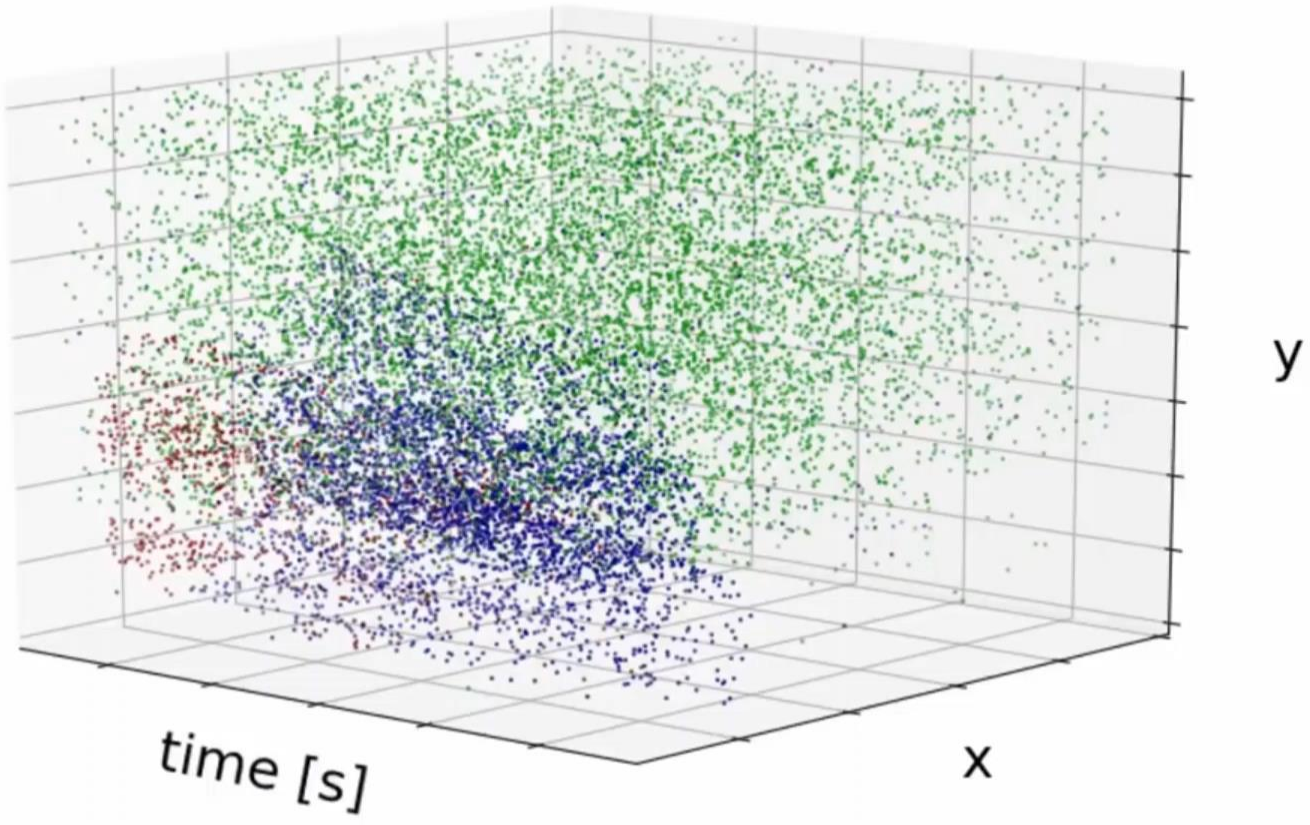
Estimated trajectory





Rosinol et al., *Ultimate SLAM?* **IEEE RAL’18 Best Paper Award Honorable Mention** PDF. Video. IEEE Spectrum.
Mueggler et al., *Continuous-Time Visual-Inertial Odometry for Event Cameras*, **IEEE T-RO’18**. PDF

Application 4: Motion Segmentation



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

Application 5: Dynamic Obstacle Avoidance

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.

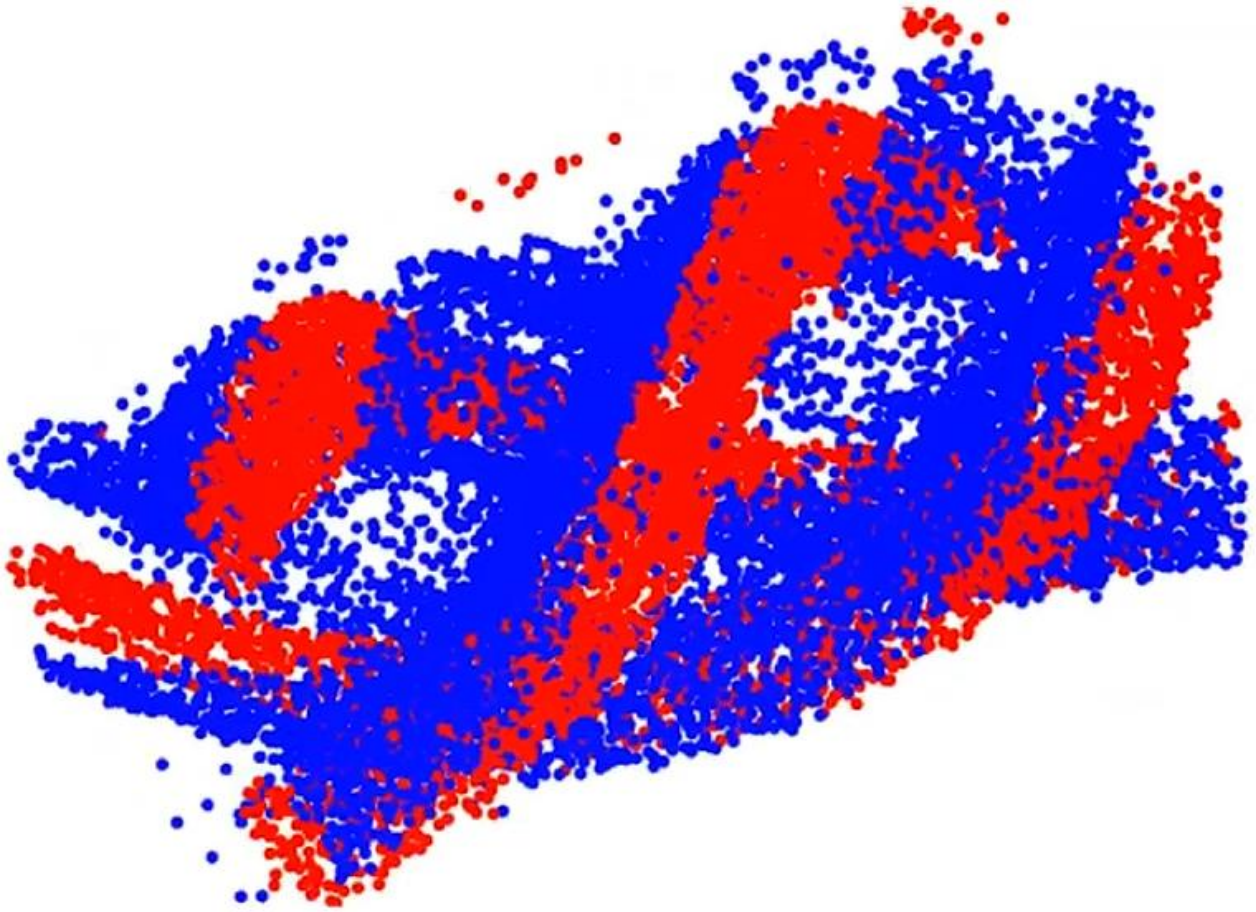
Learning with Event Cameras

- Approaches using synchronous, Artificial Neural Networks (ANNs) designed for standard images

- Approaches using asynchronous, Spiking neural networks (SNNs)

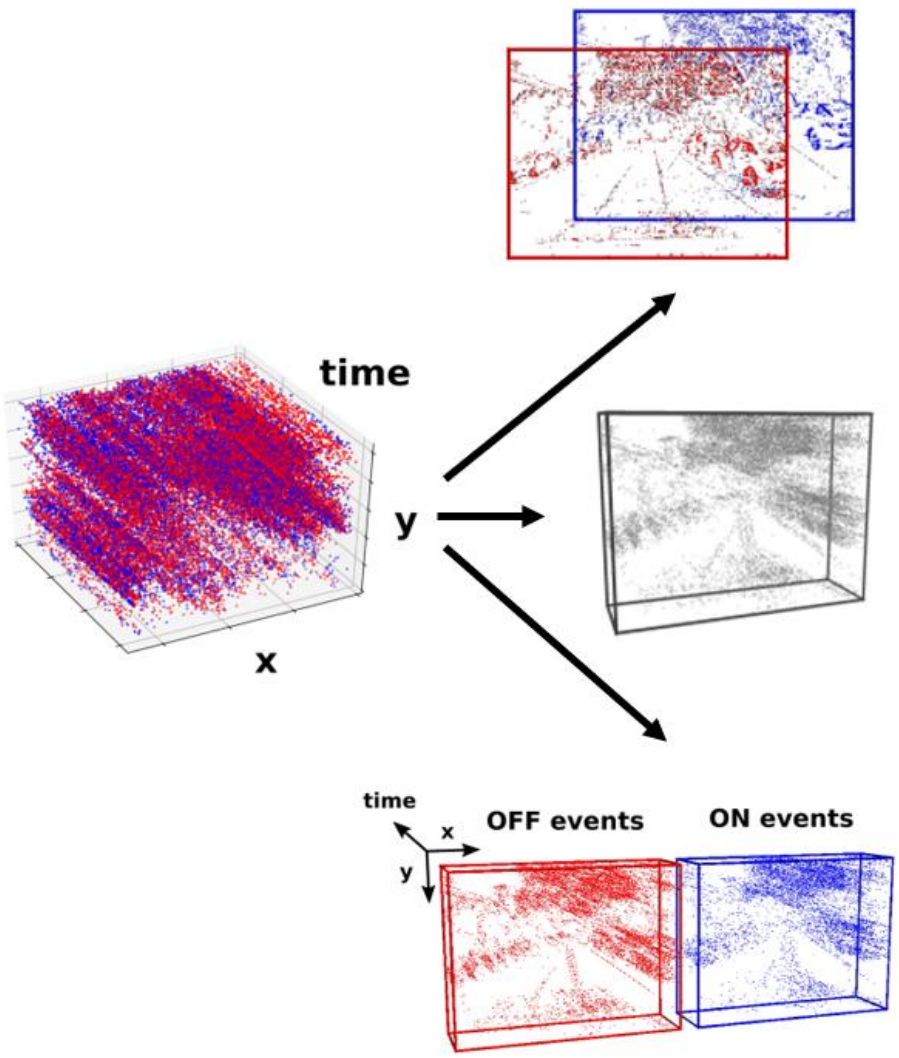
Input representation

How do we pass sparse events into a convolutional neural network designed for standard images?



Video from Zhu et al. (link)

Input representation



[Maqueda CVPR'18], [Zhu'RSS'18]

- **Aggregate positive and negative events into separate channels**
- Discards temporal information

[Zhu ECCVW'18], [Rebecq, CVPR'19], [Zhu, CVPR'19]

- **Represent events in space-time into a 3D voxel grid (x,y,t)**
- Each voxel contains sum of ON and OFF events falling within the voxel
- Preserves temporal information but discards polarity information

[Gehrig Arxiv'19]

- **Represent events in space-time as a 4D Event Spike Tensor (x,y,t,p)**
- Polarity information is preserved

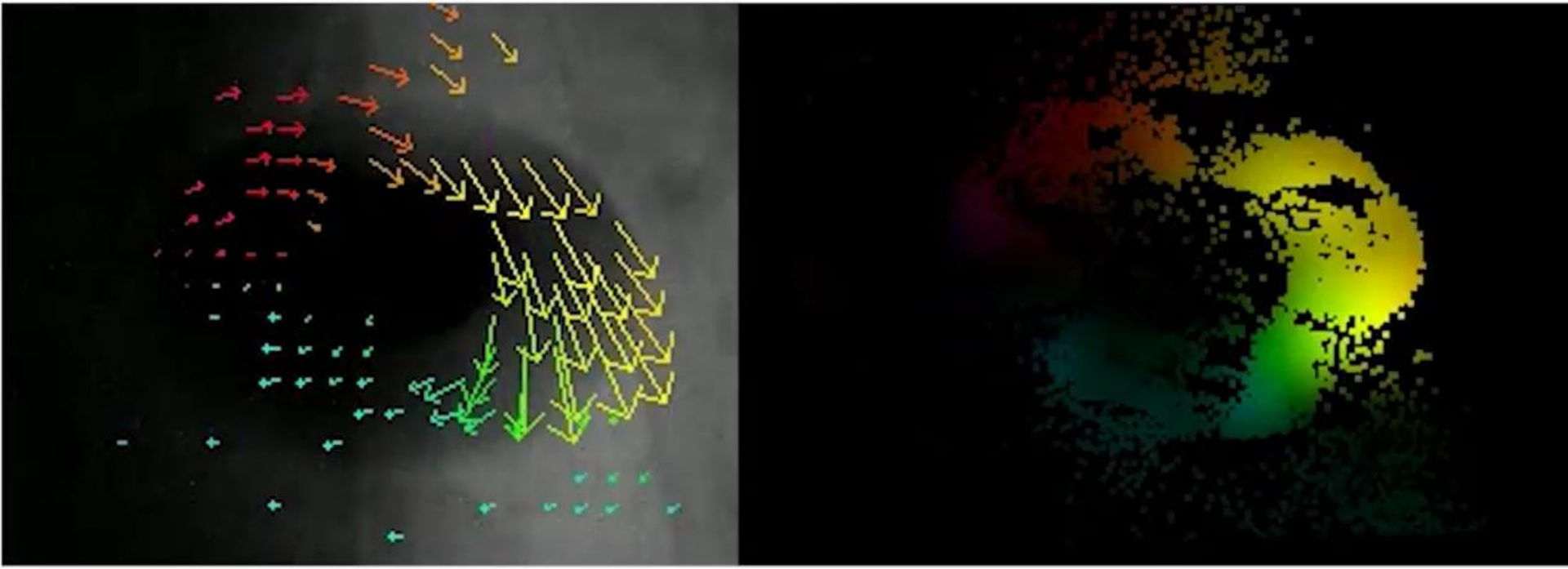
Gehrig et al., End-to-End Learning of Representations for Asynchronous Event-Based Data ICCV, 2019. PDF YouTube Code

Application: Unsupervised Learning of Optical Flow, Depth and Ego Motion

Focus used as loss: maximize sharpness of the aggregated event image.

Penn Engineering | GRASP Laboratory
General Robotics, Automation, Sensing & Perception Lab

Fidget Spinner w/ Challenging Lighting



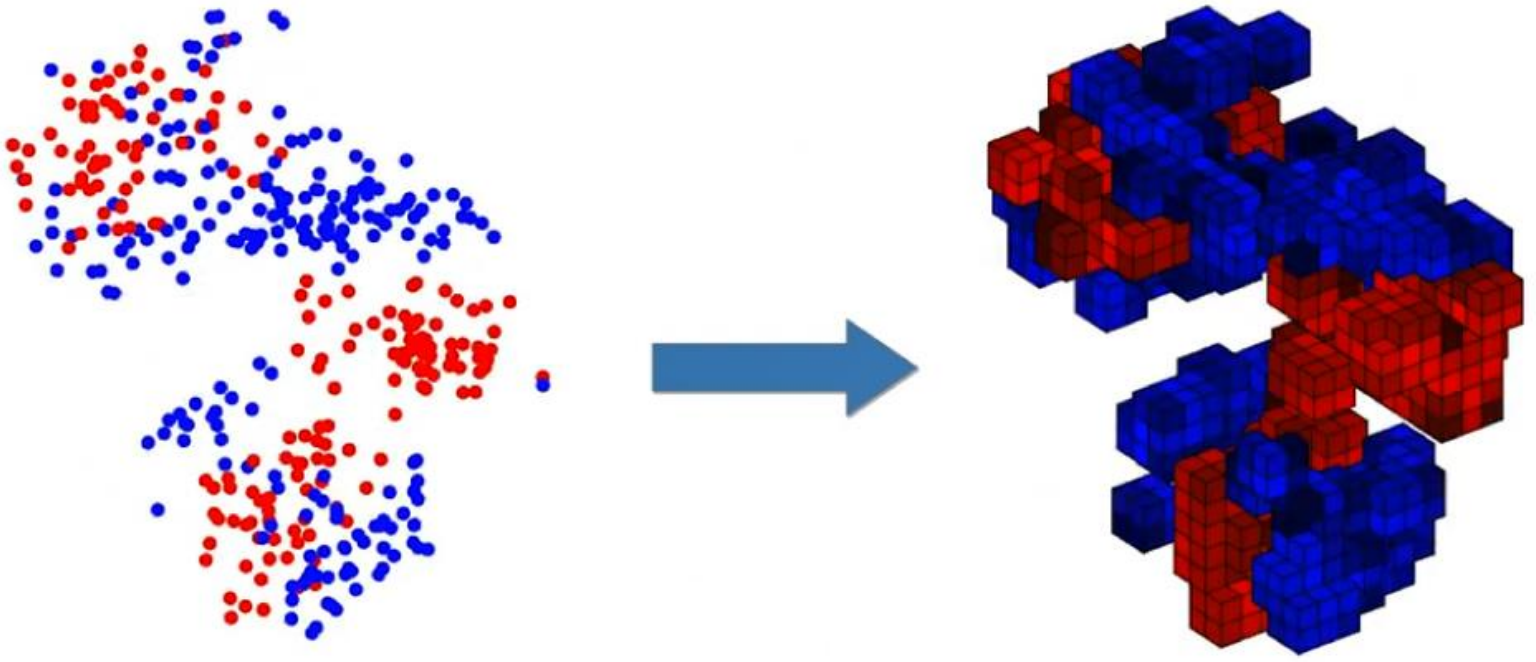
Grayscale Image w/ Sparse Flow Quiver

Dense Flow Output

Zhu et al., Unsupervised Learning of Optical Flow, Depth and Ego Motion, CVPR'19

Input representation

Discretized 3D volume (x,y,t): events are inserted into the volume with trilinear interpolation, resulting in minimal loss in resolution

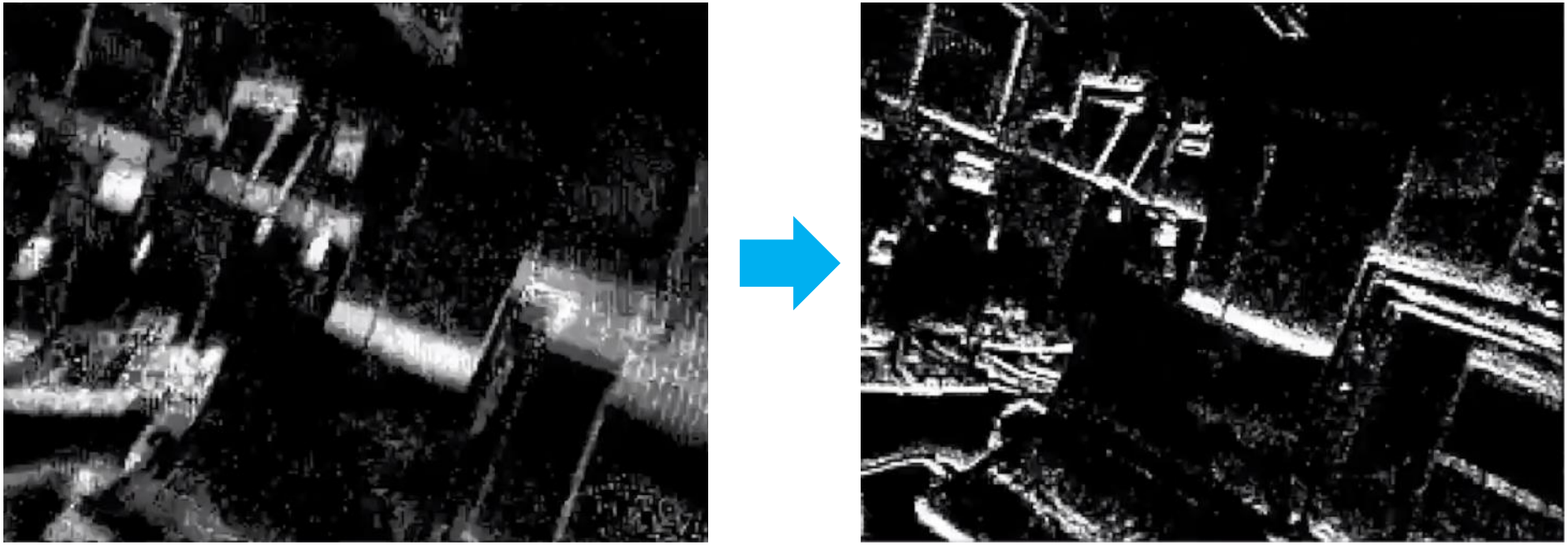


Video from [Zhu et all, CVPR'19]

[Zhu, ECCVW'18], [Zhu, CVPR'19], [Gehrig, ICCV'19], [Rebecq, CVPR'19]

Focus as Loss Function for Unsupervised Learning

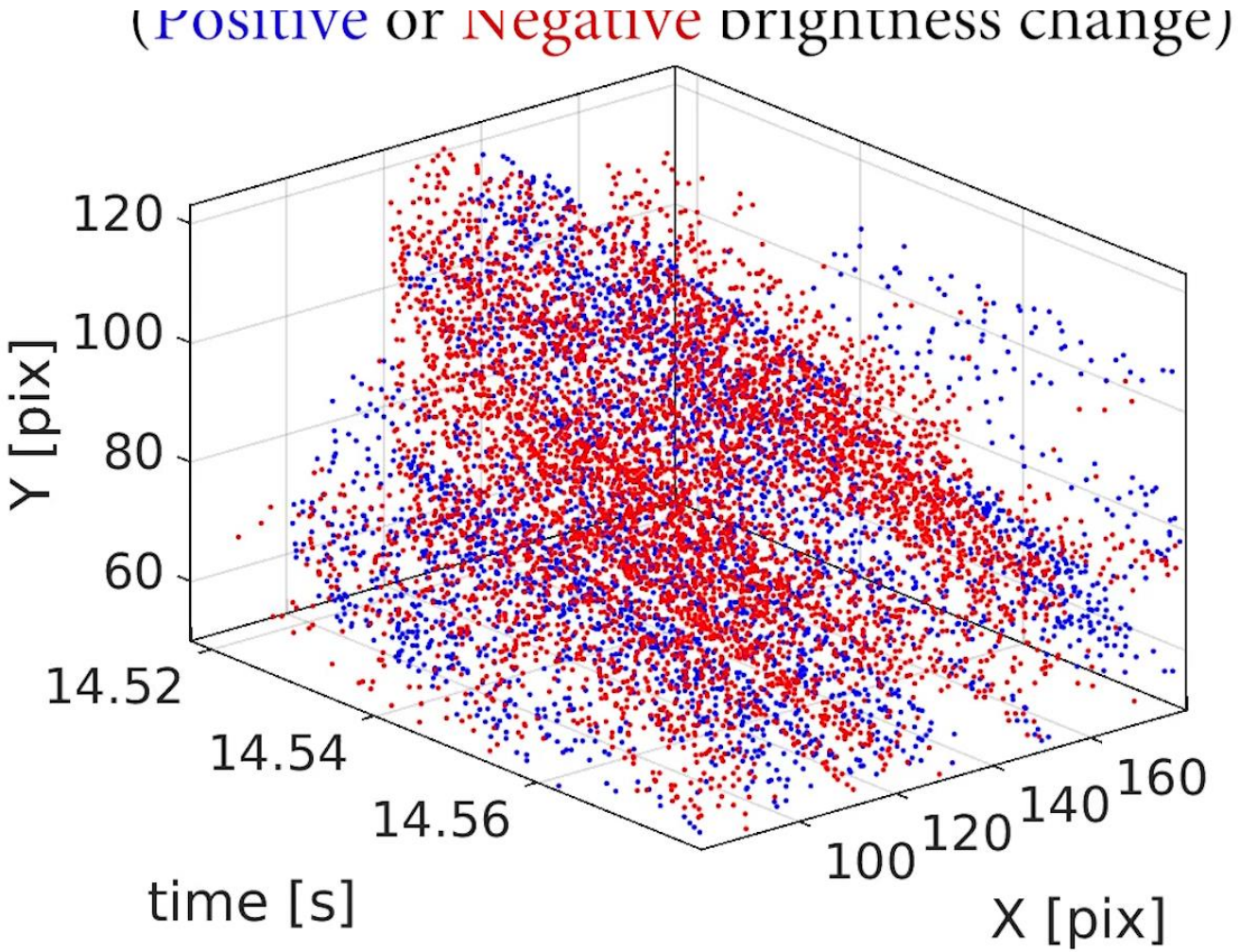
Focus used as loss: maximize sharpness of the aggregated event image.



Zhu, Unsupervised Event-based Learning of Optical Flow, Depth and Egomotion, CVPR 19

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

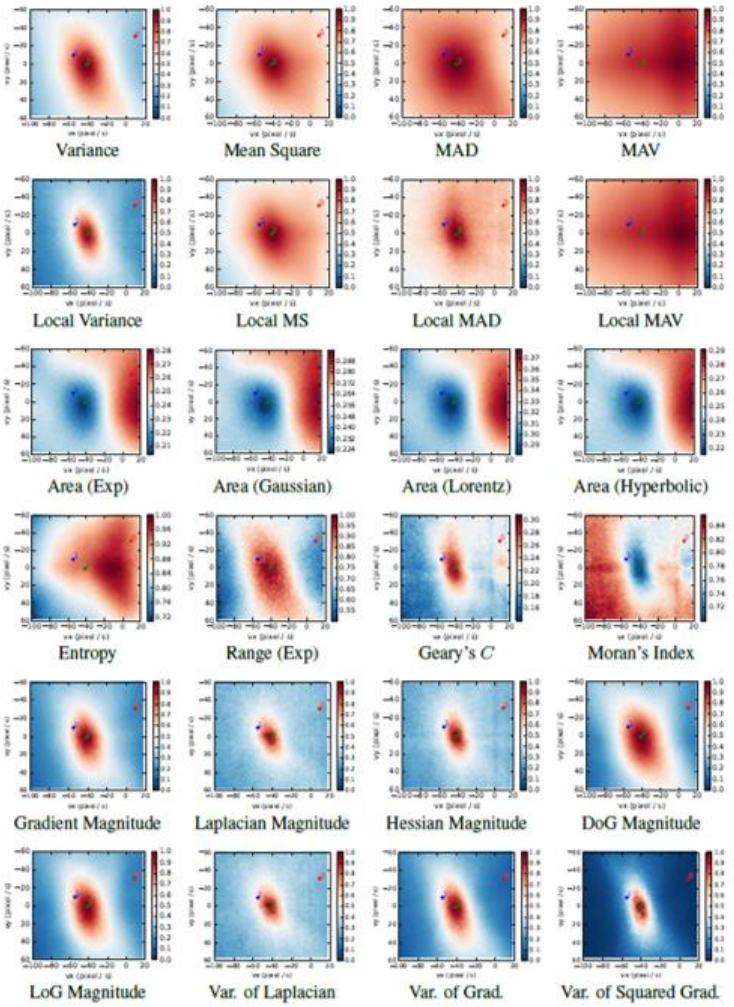


Gallego et al., A Unifying Contrast Maximization Framework for Event Cameras, CVPR18, PDF, YouTube
Gallego et al., Focus Is All You Need: Loss Functions for Event-based Vision, CVPR19, PDF.

Focus as Loss Function for Unsupervised Learning

- We proposed and benchmarked **22 focus loss functions**

- **Focus is the "data fidelity" term**

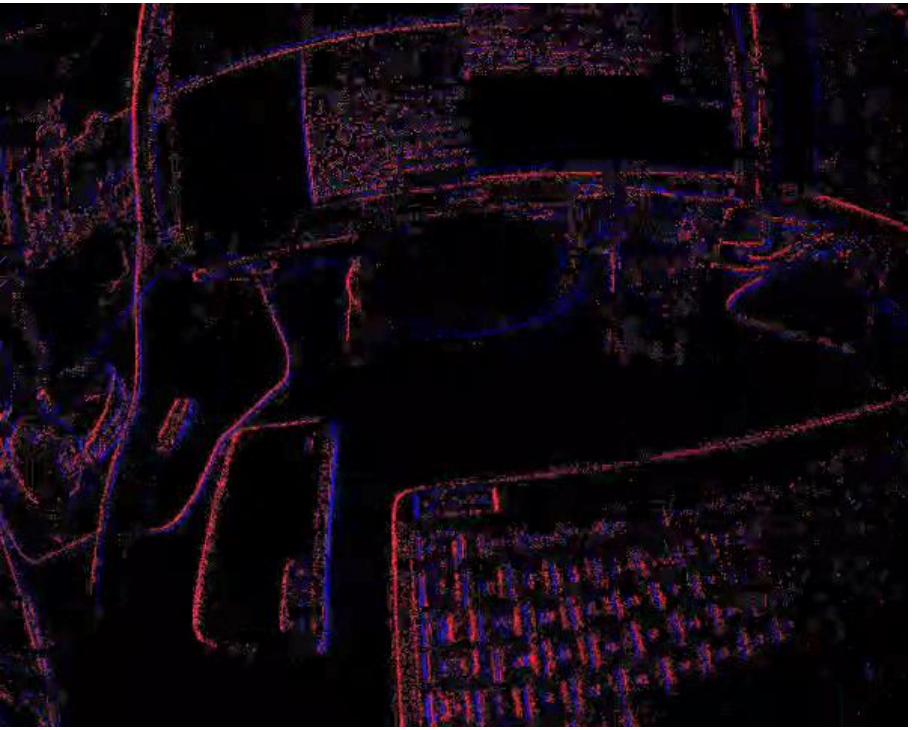


Gallego et al., Focus Is All You Need: Loss Functions for Event-based Vision, CVPR19, PDF.

Image Reconstruction

Image Reconstruction from Events

Events



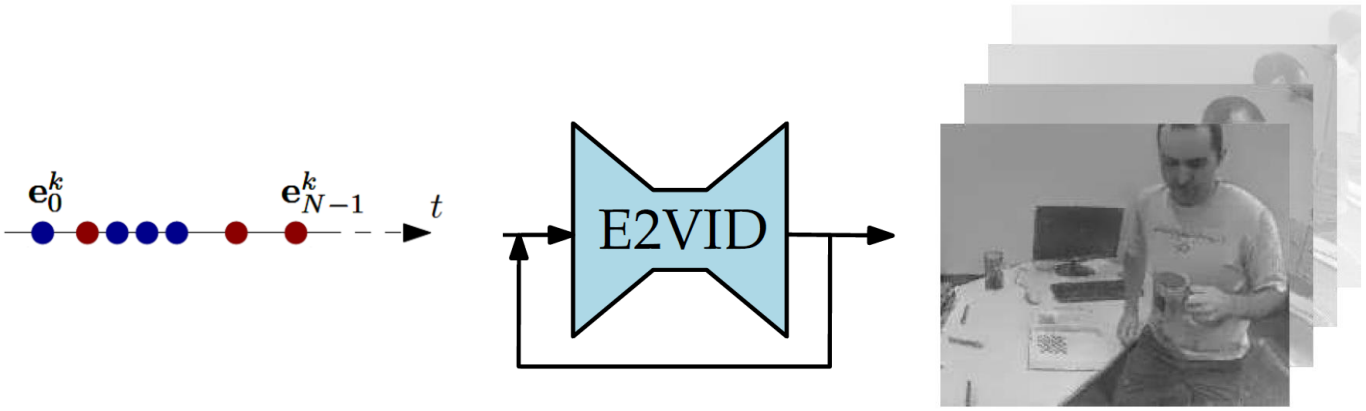
Reconstructed image from events

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., “High Speed and High Dynamic Range Video with an Event Camera”, **T-PAMI’19**. PDF Video Code

Overview

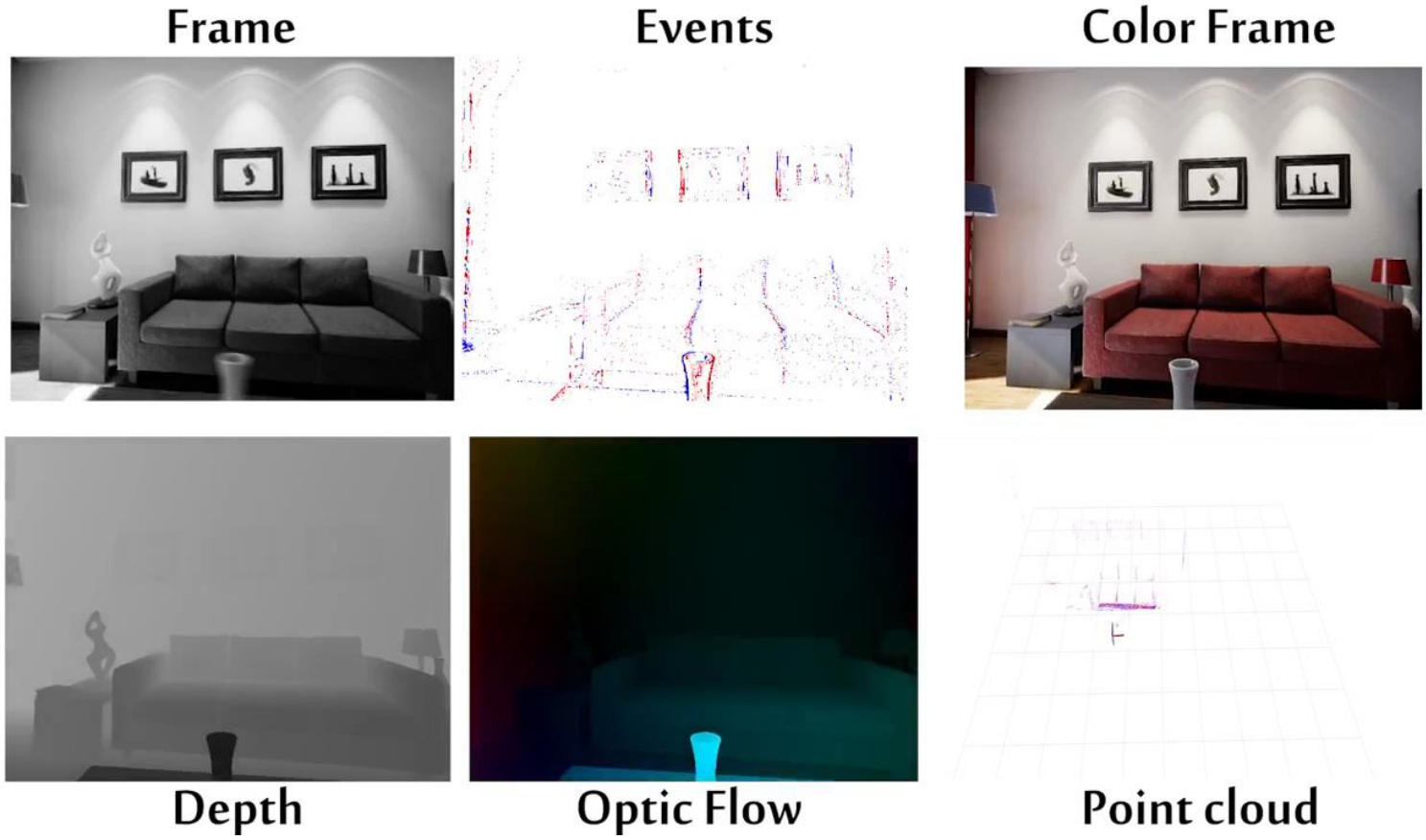
- **Recurrent neural network** (main module: Unet)
- Input: sequences of *event tensors* (spatio-temporal volumes of events[3])
- **Trained in simulation only**, without seeing a single real image



Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", **T-PAMI'19**. PDF Video Code

Event Camera Simulator



Open Source: http://rpg.ifi.uzh.ch/esim.html

Rebecq, ESIM: an Open Event Camera Simulator, CORL'18. PDF, YouTube, Project page

High Speed Video Reconstruction Results

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., “High Speed and High Dynamic Range Video with an Event Camera”, **T-PAMI’19**. PDF Video Code

Bullet shot by a gun (376m/s =1,354km/h)

Recall: trained in simulation only!

Huawei P20 Pro (240 FPS)

Our reconstruction (5400 FPS)

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", **T-PAMI'19**. PDF Video Code

Bullet shot by a gun (376m/s =1,354km/h)

Recall: trained in simulation only!

Huawei P20 Pro (240 FPS)

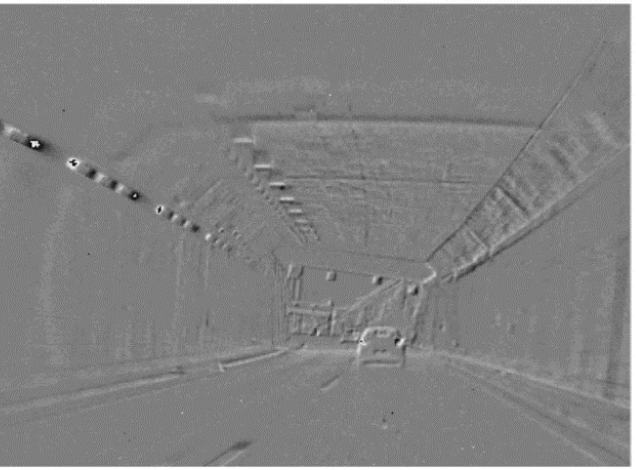
Our reconstruction (5400 FPS)

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

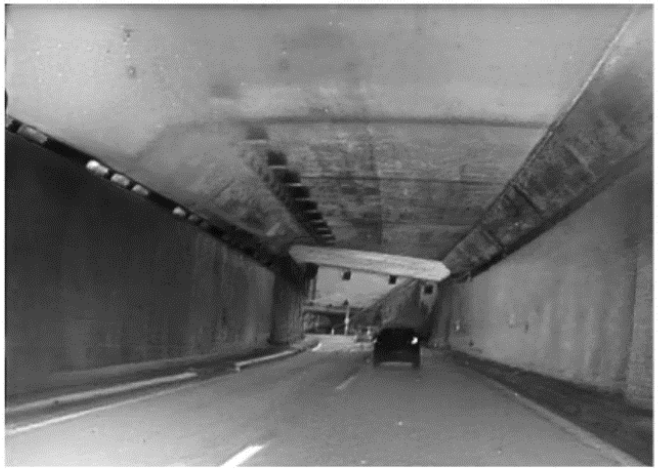
Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", **T-PAMI'19**. PDF Video Code

HDR Video: Driving out of a tunnel

Driving out of a tunnel



Events



Our reconstruction

Phone camera

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., “High Speed and High Dynamic Range Video with an Event Camera”, **T-PAMI’19**. PDF Video Code

HDR Video: Night Drive

Recall: trained in simulation only!

Video courtesy of Prophesee

Our reconstruction from events
(we used a Prophesee sensor)

GoPro Hero 6

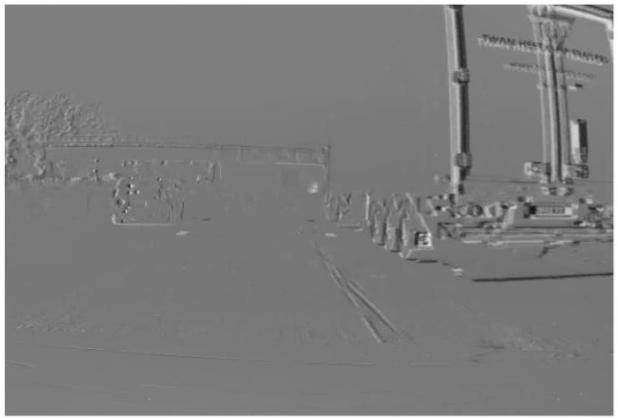
Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", **T-PAMI'19**. PDF Video Code

Downstream Applications

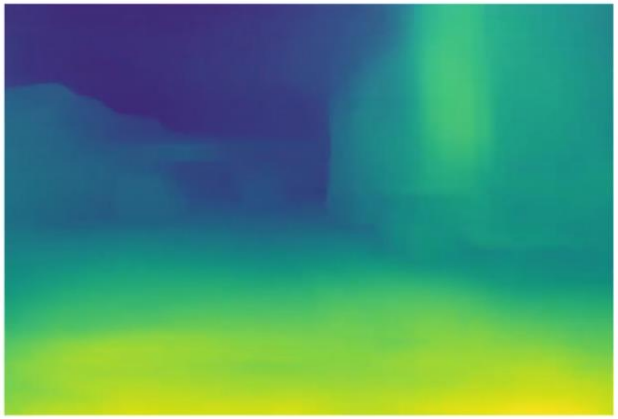
Monocular Depth Estimation

https://youtu.be/eomALySSGVU



Events

Our reconstruction



Monocular depth

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

Rebecq et al., “High Speed and High Dynamic Range Video with an Event Camera”, **T-PAMI’19**. PDF Video Code

Downstream applications (object detection)

YouTube https://youtu.be/eomALySSGVU

Events

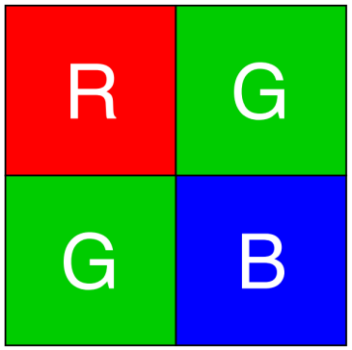
Our reconstruction + object detections (YOLOv3)

Source code & Datasets: https://github.com/uzh-rpg/rpg_e2vid

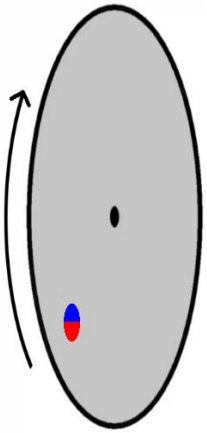
Rebecq et al., "High Speed and High Dynamic Range Video with an Event Camera", **T-PAMI'19**. PDF Video Code

Color Event Camera

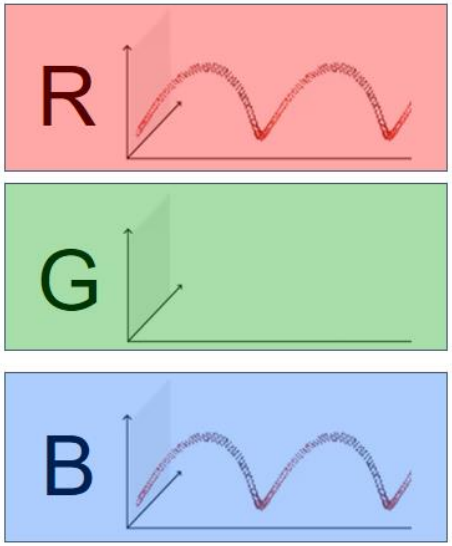
DAVIS346 Red Color



Bayer pattern



Input

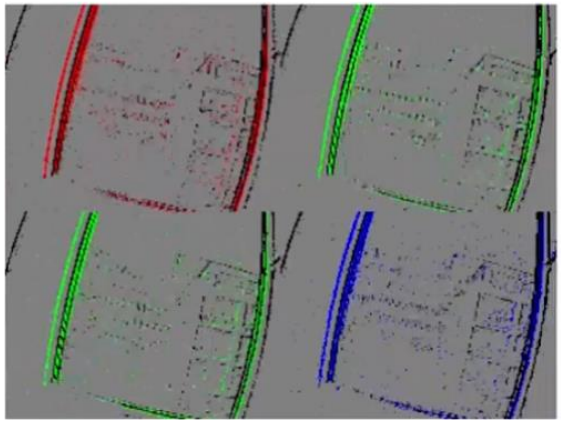


Output

- Each pixel is sensitive to **red, green or blue** light.
- Transmits **brightness changes** in each color channel

Taverni et al., Front and back illuminated Dynamic and Active Pixel Vision Sensors comparison, TCS'18

Color Event Camera Reconstruction (HDR)



Color events

Our reconstruction

Color frame

Color Event Camera Datasets: http://rpg.ifi.uzh.ch/CED.html

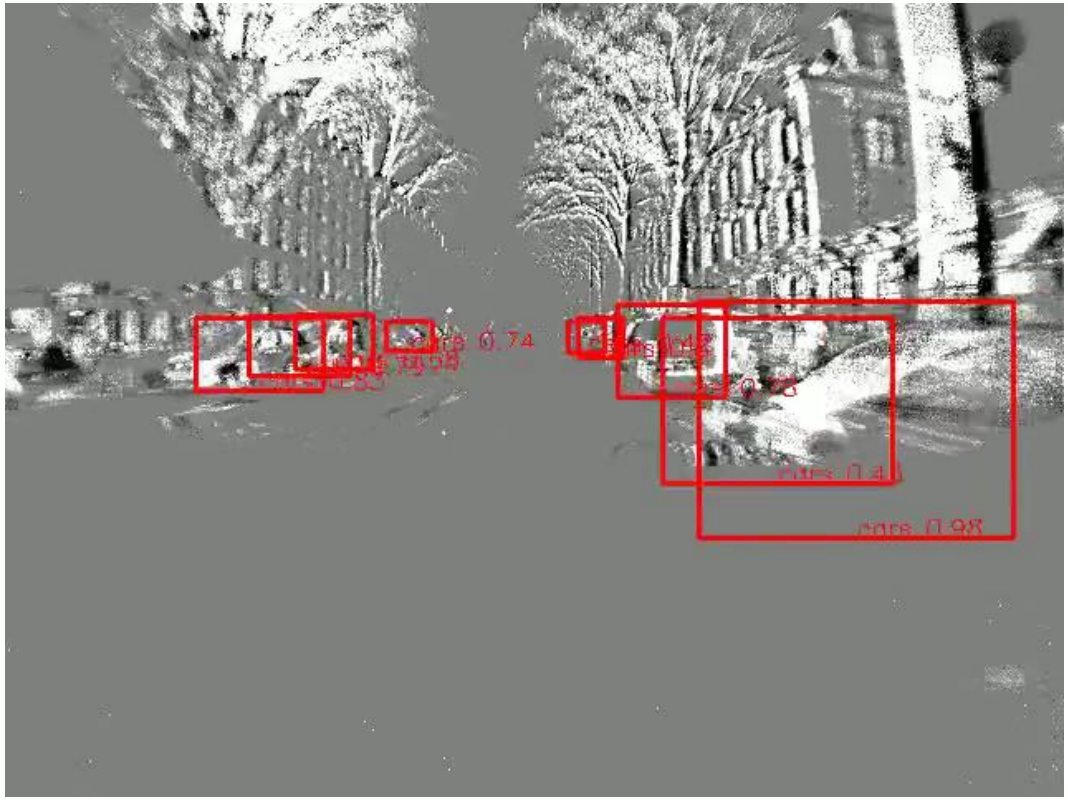
Scheerlinck, Rebecq, Stoffregen, Barnes, Mahony, Scaramuzza
CED: Color Event Camera Dataset. CVPRW, 2019. PDF YouTube Dataset

Does it mean that in order to use event cameras we must first reconstruct an image?

NO!

These results were only to show that it should be possible to design algorithms that process events **end-to-end without passing through image reconstruction**!

Example: End-to-End Object Classification



- Dataset: https://www.prophesee.ai/dataset-n-cars/
- Collected by PROPHESEE (largest event-camera company)
- Contains: Event, Images, car and pedestrian annotations

Sironi et al., “HATS: Histograms of Averaged Time Surfaces for Robust Event-based Object Classification”. CVPR’18

Thanks!

- Code, Dataset, Simulator, tutorials, resources on event cameras: http://rpg.ifi.uzh.ch/research_dvs.html

- Survey paper on event cameras: http://rpg.ifi.uzh.ch/docs/EventVisionSurvey.pdf

- Code, datasets, videos, and publications: http://rpg.ifi.uzh.ch/

 @davsca1

 @davidescaramuzza