Technische Universität Berlin

EINSTEIN CENTER Digital Future

Event-based Robot Vision

Prof. Dr. Guillermo Gallego
Chair: Robotic Interactive Perception

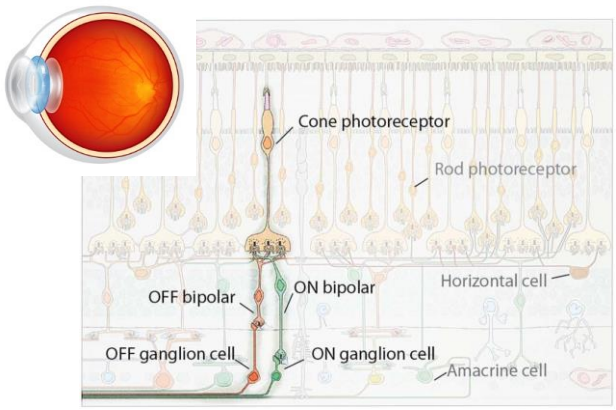
guillermo.gallego@tu-berlin.de

http://www.guillermogallego.es

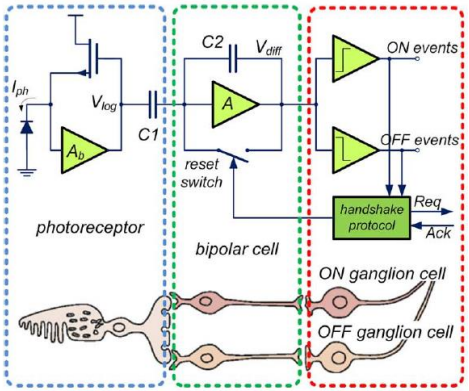
Types of Event Cameras

Dynamic Vision Sensor – DVS

- **Output**: events $e = (x, y, t, p)$ representing brightness changes (i.e., temporal contrast)
- Models the **transient** visual pathway ("where" system)
- Spatial resolution: from 128 x 128 pixels (DVS128 in 2008, **1st commercially available**) to recent 1 Mpixel versions (2020)
- Manufactured by iniVation, Samsung



Highlighted cells:
simplified biological functions



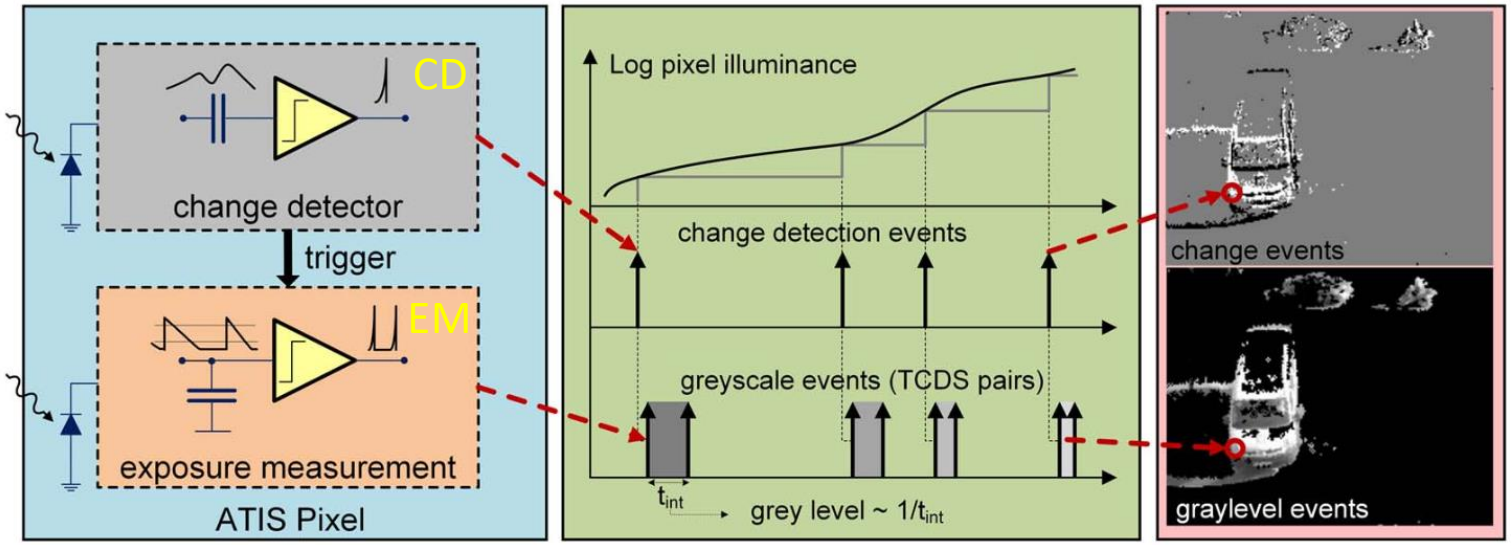
Pixel electrical design

VLSI → Vision sensor

Lichtsteiner, Posch and Delbruck, *A 128x128 120dB 15μs latency asynchronous temporal contrast vision sensor*, IEEE J. Solid-State Circuits, 2008

Asynchronous Time-based Image Sensor - ATIS

- **Output**:
  - Change detection (**CD**) events (i.e., like DVS) that model the transient visual pathway ("where" system)
  - Exposure measurements (**EM**) events (grayscale events) that model the sustained visual pathway ("what" system). Intensity-encoding is time-based
- No frames: both CD and EM events are asynchronous
- Manufactured by Prophesee (Paris)



Posch et al., *A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression …*, IEEE JSSC 2011.

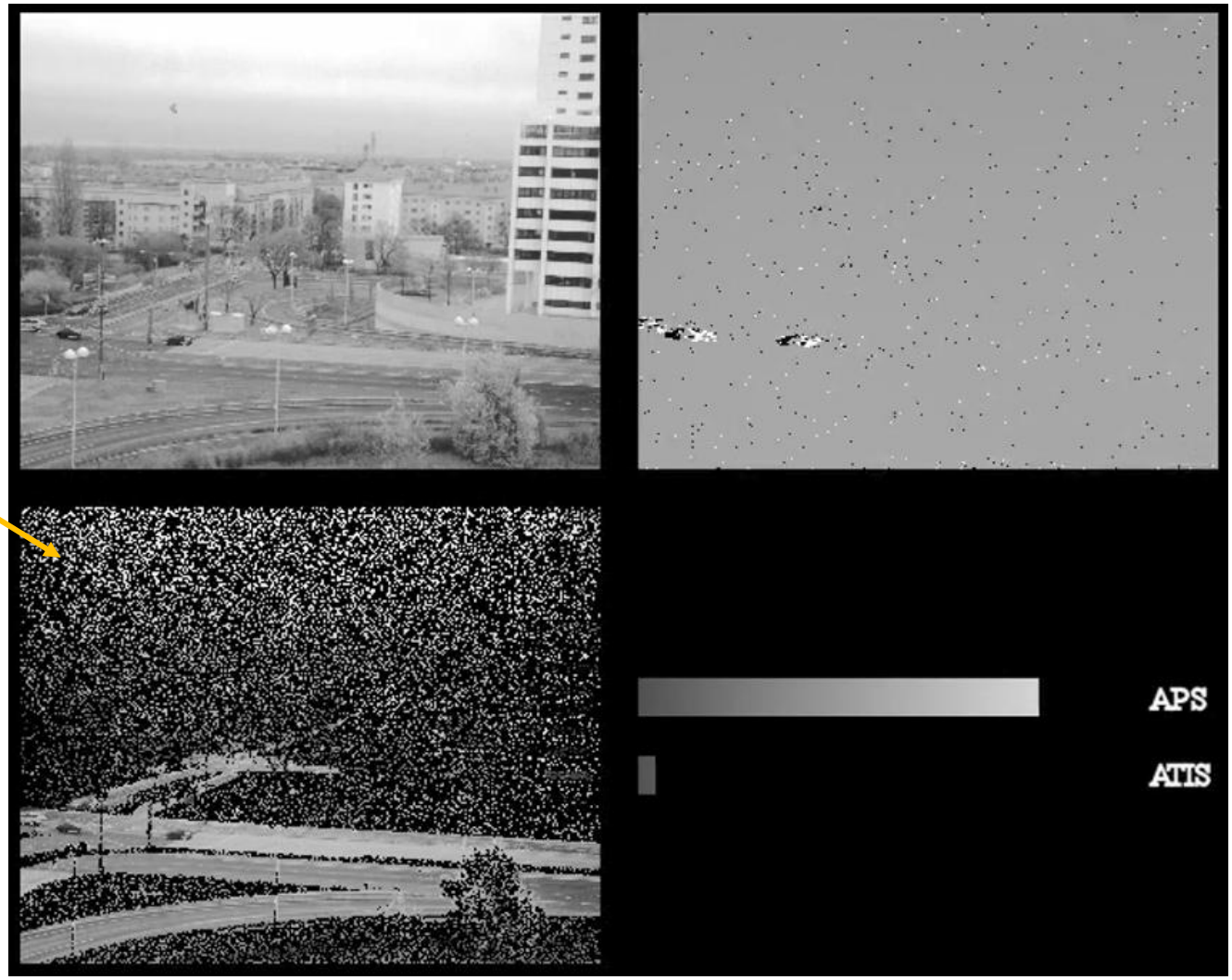
Asynchronous Time-based Image Sensor - ATIS

Standard camera Frame (APS)

(Event "noise")

EM grayscale events

Displaying the last grayscale value per pixel



CD events (like DVS)

Amount of transmitted data

Sample application: Pixel-level video compression

Posch et al., *A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression …*, IEEE JSSC 2011.

Asynchronous **T**ime-based **I**mage **S**ensor - ATIS

Pixel-level video compression by transmitting only EM grayscale events

Transmitter ATIS camera Receiver

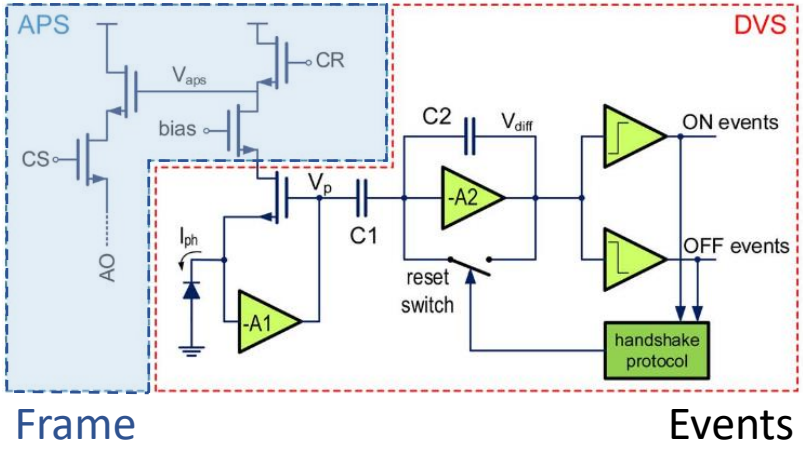
Video by X. Lagorce https://youtu.be/3Wiw8LA8hLs

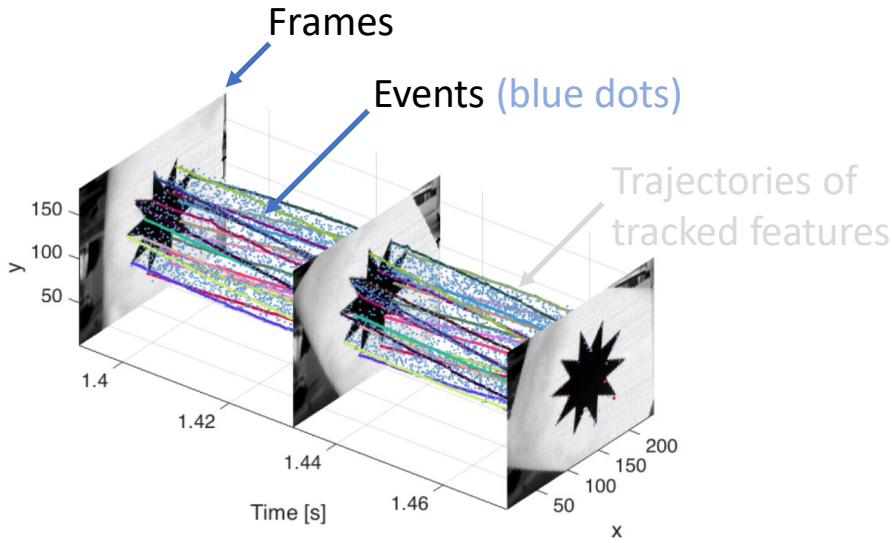
Posch et al., *A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression …*, IEEE JSSC 2011.

Dynamic Pixel and Active Vision Sensor - DAVIS

- **Output**: DVS **events**, standard **frames** and **IMU** data
  - Combines a DVS and a standard camera **in the same pixel array**
  - Frames (rolling or global shutter, grayscale or color) are not HDR (~55 dB)
  - Frames resemble the information in the "what" visual pathway
- Spatial resolution: from 240 x 180 pixels (DAVIS240C) to 640 x 480 pixels (VGA)
- Manufactured by iniVation and Insightness

Add a few transistors to each DVS pixel…

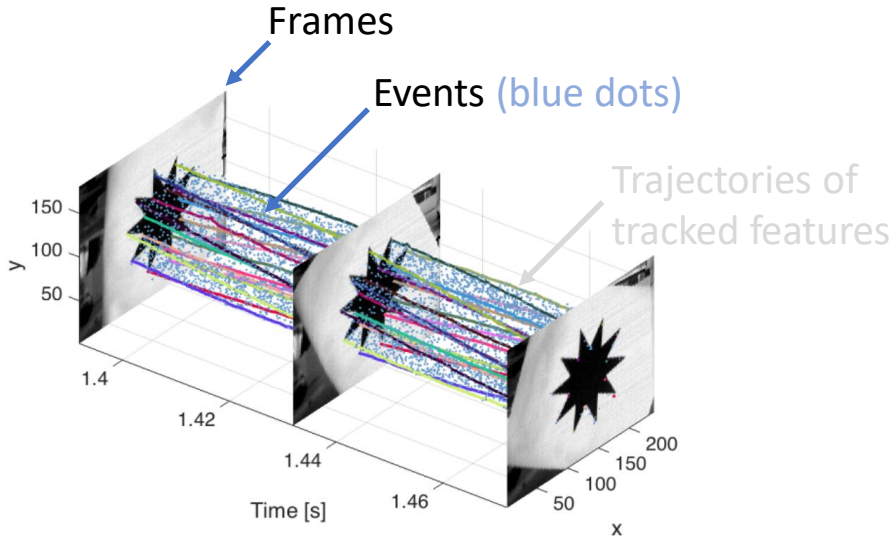


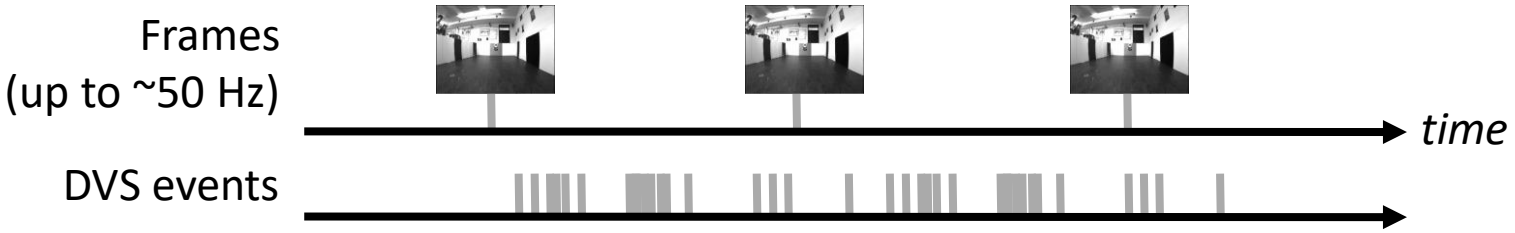


Brandli et al., *A 240x180 130 dB 3 µs Latency Global Shutter Spatiotemporal Vision Sensor*, IEEE J. Solid-State Circuits, 2014

Dynamic Pixel and Active Vision Sensor - DAVIS

- **Output**: DVS **events**, standard **frames** and **IMU** data
  - Combines a DVS and a standard camera **in the same pixel array**
  - Frames (rolling or global shutter, grayscale or color) are not HDR (~55 dB)
  - Frames resemble the information in the "what" visual pathway





Brandli et al., *A 240x180 130 dB 3 µs Latency Global Shutter Spatiotemporal Vision Sensor*, IEEE J. Solid-State Circuits, 2014

Very Short Summary

	DVS (2008)	DAVIS (2014)	ATIS (2011)
Events (change detection)	Yes	Yes	Yes
Grayscale	No	Frames (~ 55dB)	EM events (HDR) Time-based encoding
HDR, Latency, Power	All similar		
Manufacturer	iniVation, Samsung	iniVation, Insightness	Prophesee

Types of Events

Types of Events

- **DVS** or "change detection (**CD**)" event $e = (x, y, t, p)$
- ATIS **grayscale events** (Exposure Measurement) $e = (x, y, t, L)$
- Event from a **stereo** camera
  - Include camera index (left/right): $e = (x, y, t, p, l/r)$
- **Color** events
  - Color DVS (include channel index) $e = (x, y, t, p, channel)$
  - Color ATIS (exposure measurement) $e = (x, y, t, r, g, b)$

- **Augmented** events $e = (x, y, t, p, extra)$
  - Additional information given by the output of some algorithm
  - Lifetime estimation: normal optical flow, lifetime
  - Sensor fusion (event camera + depth): 3D depth of each event

Mueggler et al., *Lifetime Estimation of Events from Dynamic Vision Sensors*, ICRA 2015.
Weikersdorfer et al., *Event-based 3D SLAM with a depth-augmented dynamic vision sensor*, ICRA 2014

The meaning of “Event”

- Events can signal **any kind of information with associated place and time** (x, y, t): intensity, local spatial contrast, etc.
- Ideally, events represent **meaningful information** to reduce data rate and therefore decrease demands on bandwidth, memory, and computer power for transmission, storage and processing.

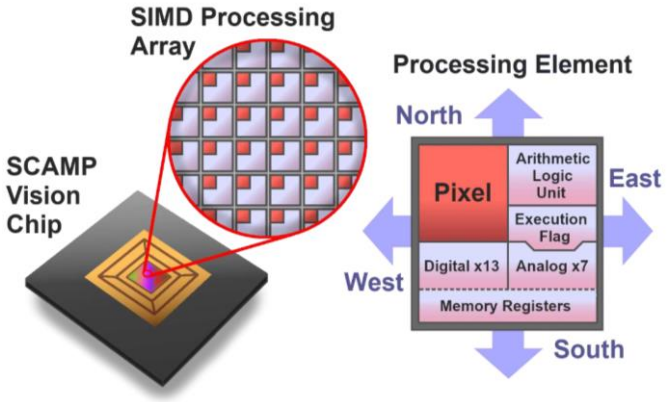
- The term has evolved from address-event (AER) representation. Today “event” mostly refers to “brightness changes” output by DVSs.

Image by Walterio Mayol

Where could this be going?

- **Perform early visual processing on chip, near the image plane** (in analog and transmit only “meaningful information” off the chip)

- Increasing **pixel complexity**:
  - Standard Pixel → DVS pixel → DAVIS pixel → Pixel Processor Arrays (PPAs)



- PPA
  - One mini-processor per pixel (SCAMP-5)
  - Dedicates more area to processing, less for transducing light (photodiode). Light conversion efficiency could be improved with a stacked approach.
  - It is a prototype sensor (academic), not mass-produced.
  - We are entering the realm of computational imaging/photography, co-designing the visual pipeline end-to-end.

References

Reading:

- Journal papers on main sensor designs:
  - **DVS**: Lichtsteiner et al., *A 128x128 120dB 15μs latency asynchronous temporal contrast vision sensor*, IEEE J. Solid-State Circuits, 2008.
  - **ATIS**: Posch et al., *A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS*, IEEE J. Solid-State Circuits, 2011.
  - **DAVIS**: Brandli et al., *A 240x180 130 dB 3 μs Latency Global Shutter Spatiotemporal Vision Sensor*, IEEE J. Solid-State Circuits, 2014.
- Papers comparing the sensors:
  - Posch et al., *Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output*, Proc. IEEE, 2014. PDF
  - T. Delbruck, *The Slow but Steady Rise of the Event Camera*, EE Times 2020.
  - **Section 2.1** of Gallego et al., Event-based Vision: A Survey, IEEE TPAMI 2020.