

Event Computer Vision 10 years Assessment: Where We Came From, Where We Are and Where We Are Heading To

R.B. Benosman

University of Pittsburgh, Carnegie Mellon, Sorbonne Universit  s

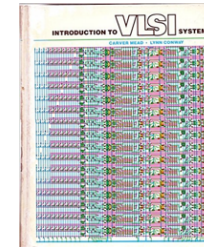
Eye& Ear and McGowan Institute

3025 E Carson st,

Pittsburgh, PA 15203



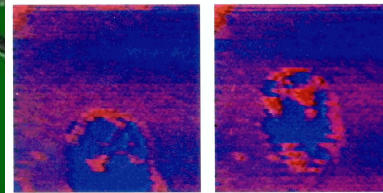
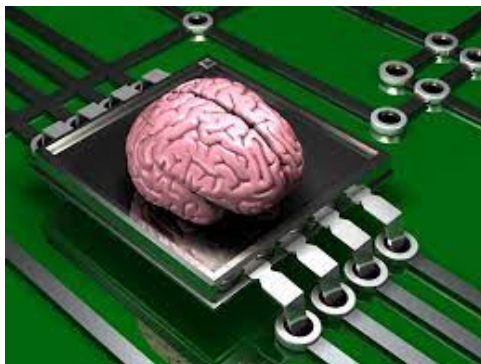
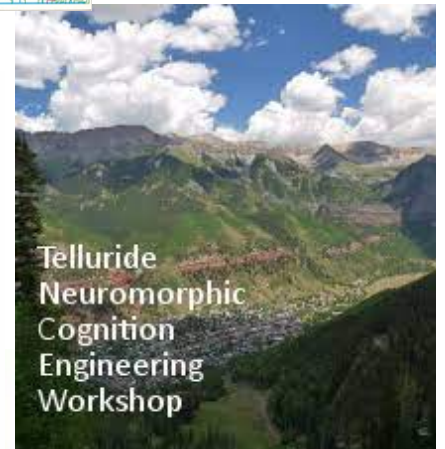
Where we came from...



SCIENTIFIC
AMERICAN

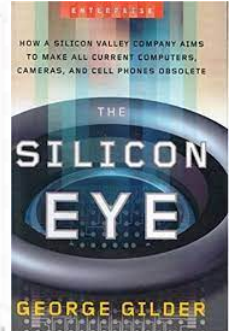
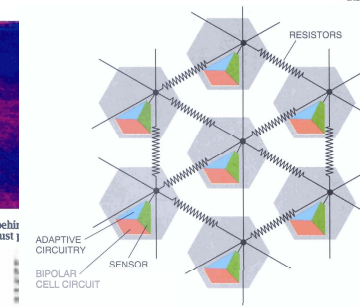
MAY 1991
\$3.95

Exploring the genetic heritage of rhesus monkeys.
Can anyone explain high-temperature superconductivity?
The impact of Hawaii's burning oil wells.



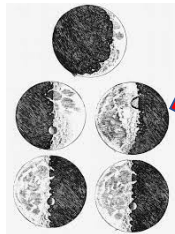
SOCCERBALL in motion shows how the delayed response of the horizontal cell network affects the retina's perception. The ball leaves behind spots have just

Mahowald/Mead retina



Where we came from...

A **Chance** to Move Perception from Engineering to Basic Science!



observe



understand

Relativistic orbital motion

- Light and bodies move on *geodesics* in the spacetime
- The 3D "shadow" of a geodesic in 4D is the *orbit*
- We need an equation for "straight lines" in spacetime!

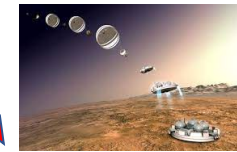
$$\frac{d}{d\tau} \left(\frac{\partial \mathcal{L}}{\partial \dot{\sigma}^i} \right) - \frac{\partial \mathcal{L}}{\partial \sigma^i} = 0$$

Euler-Lagrange equation

$$\mathcal{L} = \frac{1}{2} g_{ij} \dot{\sigma}^i \dot{\sigma}^j$$

The metric tensor g describes spacetime curvature, contains the Schwarzschild metric coefficients

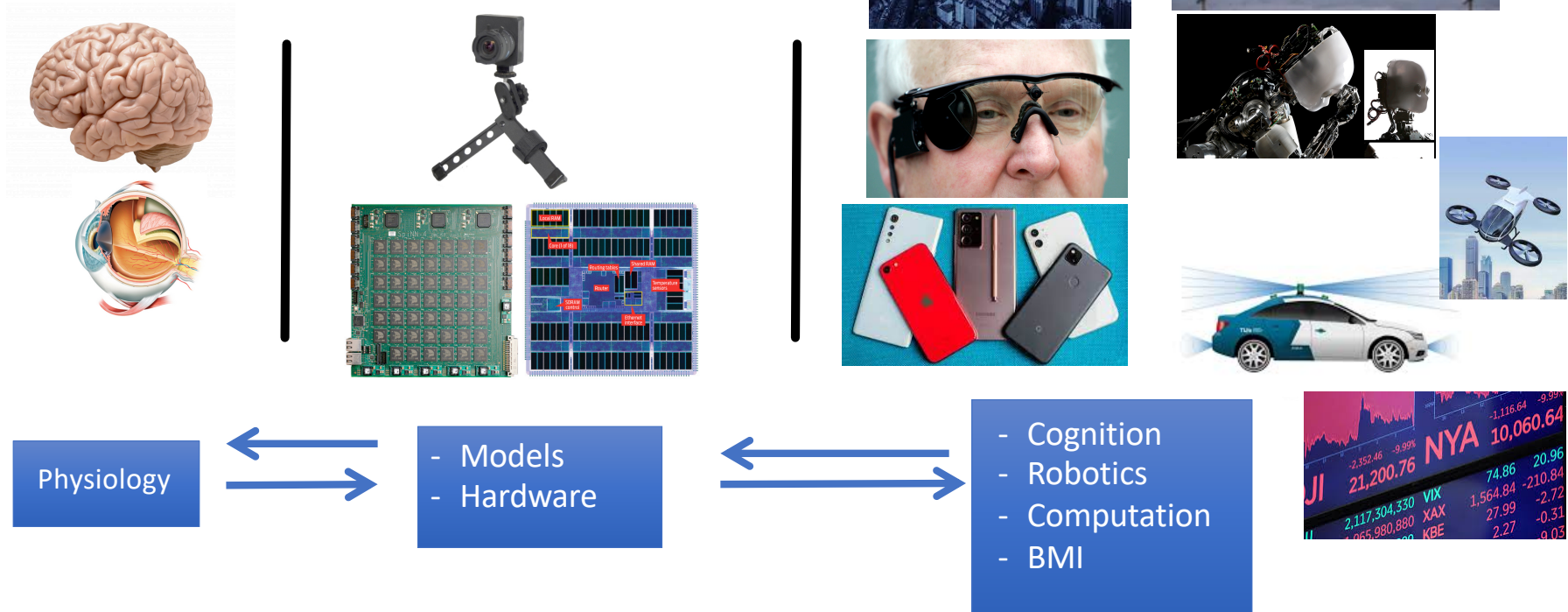
model



application



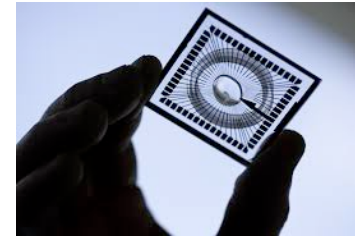
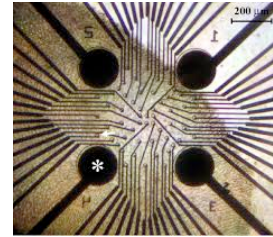
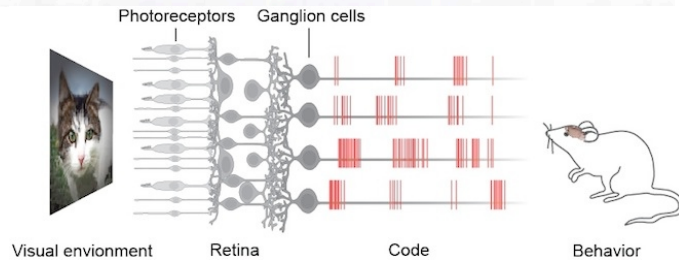
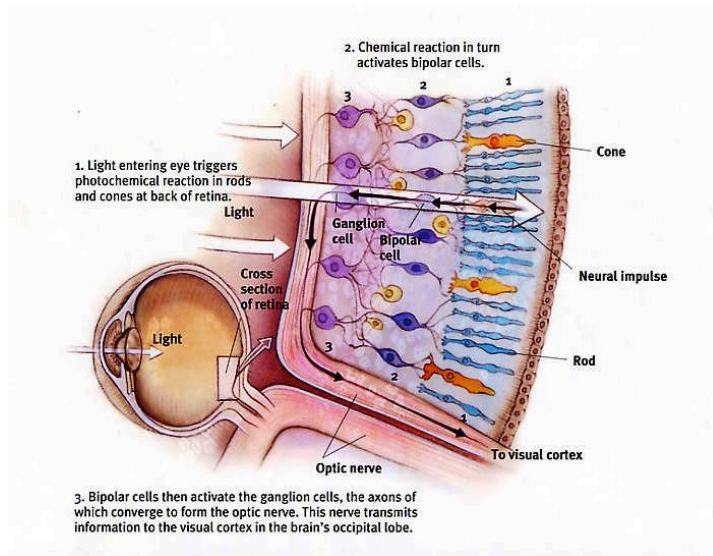
Where we came from...



- **Develop new bidirectional methodology** to understand the brain
- **Merging Computational and Biological Vision**
- **Importance of applications in Brain-Machine Interfaces**

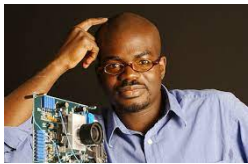
Where we came from...

The retina: “the most approachable portion of the Brain”



Where we came from...

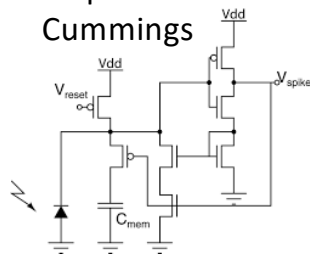
Event-based Cameras: A Long Incremental Quest



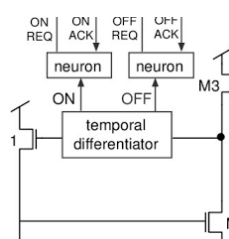
Kwabena Boahen



Ralph Etienne-Cummings



Jorg Kramer



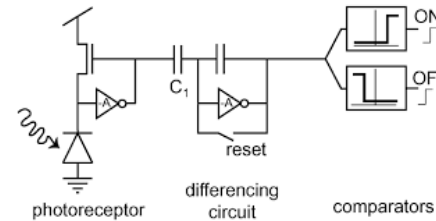
Patrick Lichtsteiner
ETH Zurich



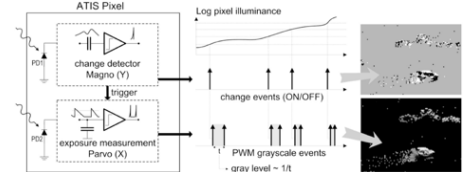
Tobi Delbruck
ETH Zurich



Christoph Posch
Sorbonne Université



Daniel Matolin



2000

2004

2008

Where we are ...

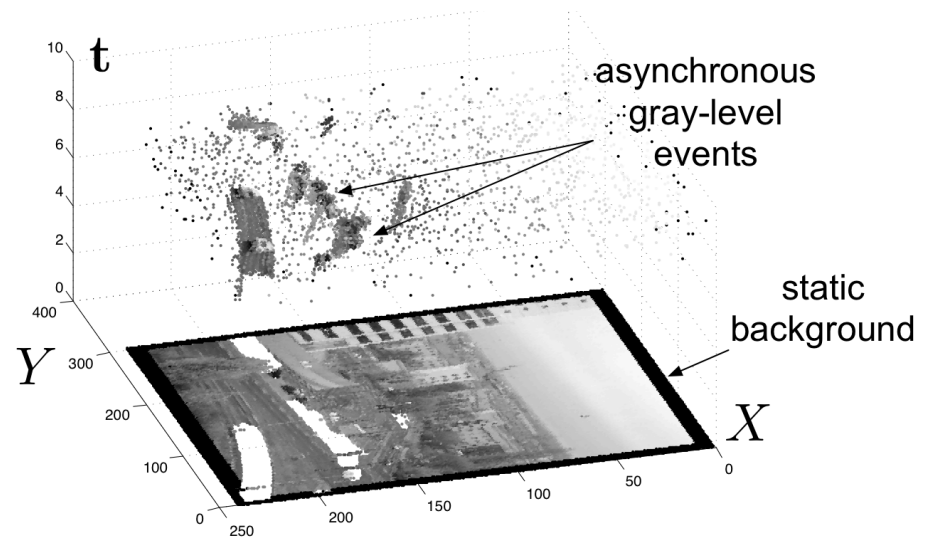
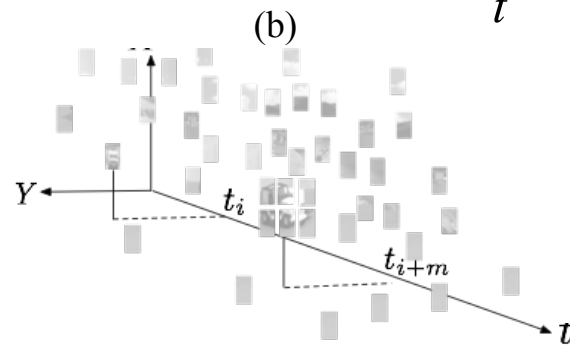
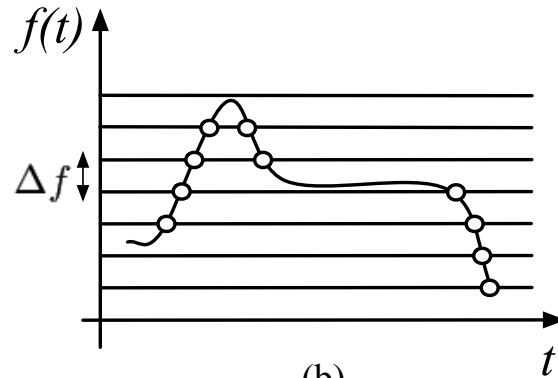
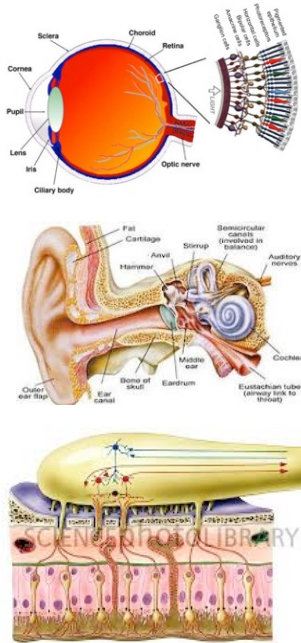
Event-Based Cameras V1.0: Panorama

The collage features the following elements:

- Samsung**: Logo in a blue oval next to a camera module on a tripod.
- Huawei**: Red flower logo.
- iniLabs**: Blue text logo next to a camera module and a comparison image showing a conventional camera view (left) and an event-based camera view (right) of a street scene. The event-based view shows sparse data points. Below the images is a graph showing a signal waveform.
- Sony**: Blue camera module next to the Sony logo.
- insightness**: Blue and white checkered logo next to the insightness text.
- OmniVision**: Logo with a red 'V' and a white eye icon.
- Hillhouse**: Logo with the text 'HILLHOUSE' and 'SEE DIFFERENT, SEE FURTHER'.
- Other images**: A camera module, a person holding a camera module, and a camera module with a lens.

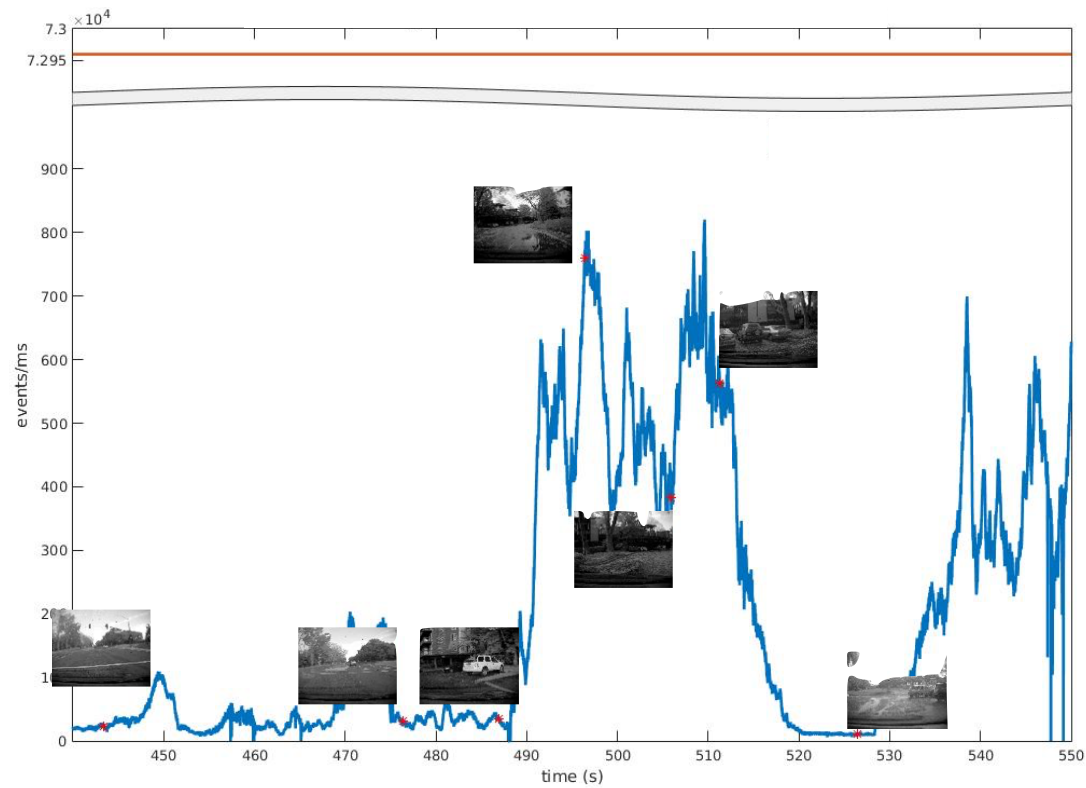
- Event-based cameras have become a commodity

Event Acquisition



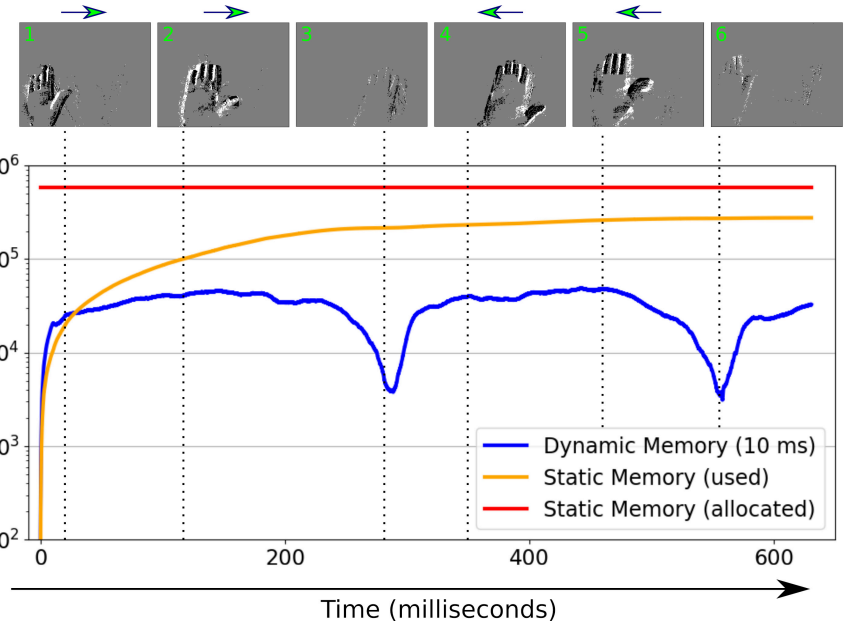
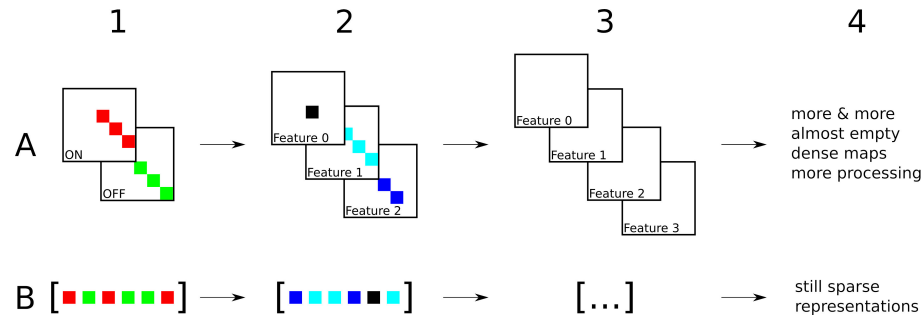
- Information is sent when it happens
- When nothing happens, nothing is sent or processed
- Sparse information coding
- Time is the most valuable information

Properties Event Based sensors?



- Amount of Data generated from a moving car per 10 milliseconds

Generating Frames is non optimal



Dataset (Mean Event Rate & Sensor Size)	PokerDVS 170.4 ev/ms 35x35			N-MNIST 13.6 ev/ms 28x28			DvsGesture 56.9 ev/ms 128x128			NavGesture-walk 188.6 ev/ms 304x240		
1. Time Window (ms)	1	10	100	1	10	100	1	10	100	1	10	100
2. Mean Number of events in TW (percentage of active pixels)	101 (8%)	390 (32%)	486 (40%)	22 (3%)	84 (11%)	229 (29%)	53 (<1%)	340 (2%)	1751 (11%)	285 (<1%)	2818 (4%)	13279 (18%)
3. Max Number of events in TW (percentage of active pixels)	356 (29%)	848 (69%)	1052 (86%)	223 (28%)	312 (40%)	597 (76%)	467 (3%)	2056 (13%)	9191 (56%)	2599 (4%)	18296 (25%)	68128 (93%)
4. Working Memory Size (kB) Dynamic - Average case	0.8	3.1	3.9	0.2	0.7	1.8	0.4	2.7	14.0	2.3	22.5	106.2
5. Working Memory Size (kB) Dynamic - Worst case	2.8	6.8	8.4	1.8	2.5	4.8	3.7	16.4	73.5	20.8	146.3	545.0
6. Allocated Memory Size (kB)	9.8	9.8	9.8	6.3	6.3	6.3	131	131	131	584	584	584
7. Memory ratio dynamic/static (Average Case)	8%	32%	40%	3%	11%	29%	1%	2%	11%	1%	4%	18%
8. Memory ratio dynamic/static (Worst Case)	29%	69%	86%	28%	40%	76%	3%	13%	56%	4%	25%	93%

Computer vision: the impossible trade off of power vs frame rate



Too many data

Too slow

Light-dependent

over-sampling

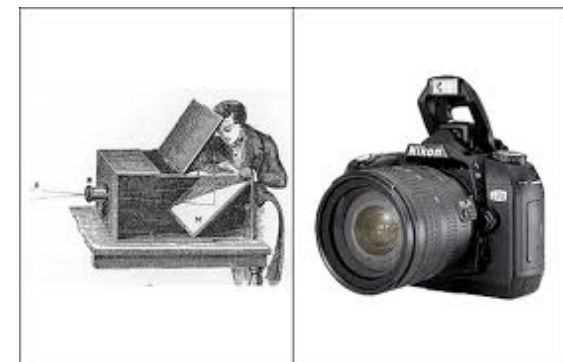
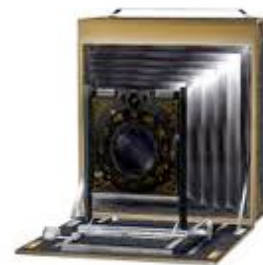
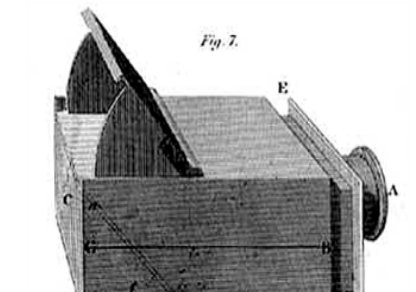
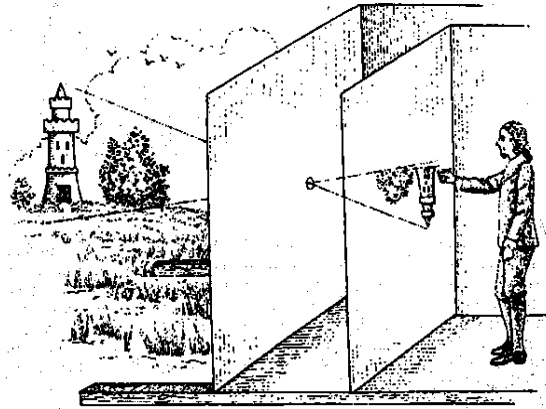
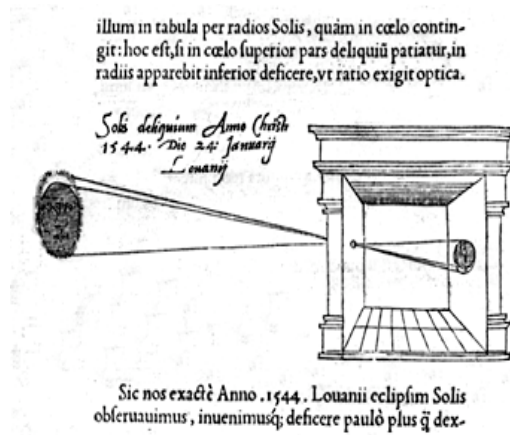
- ▶ redundant useless data
- ▶ power and resource hungry: need to acquire/transmit/store/process

under-sampling

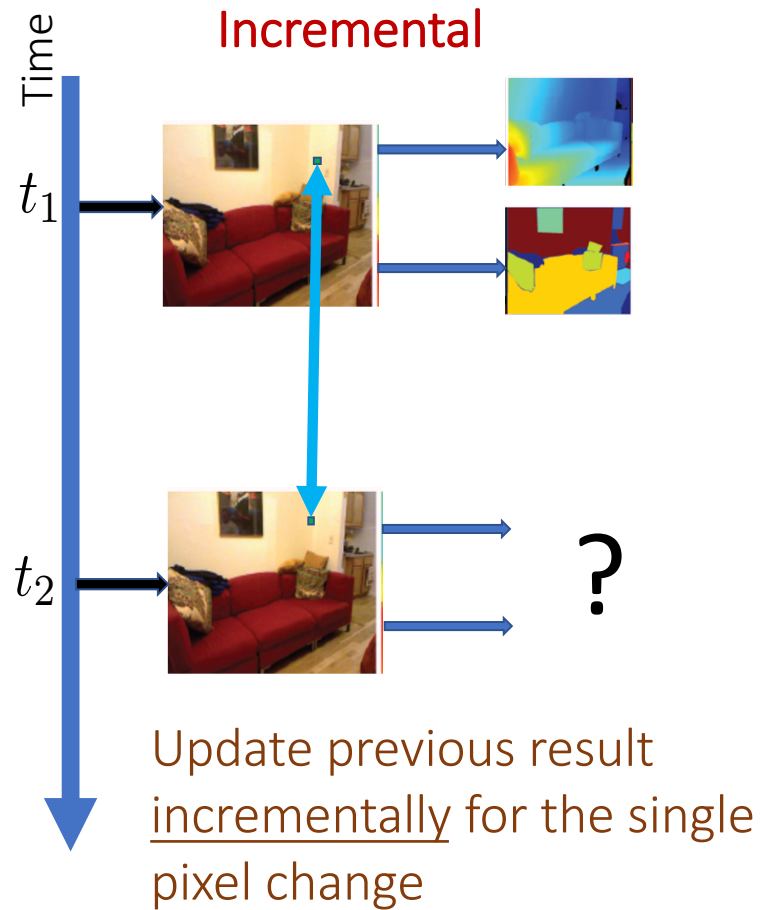
- ▶ motion blur
- ▶ displacement between frames

High Power & High Latency

Origins of Imaging



Event Computation



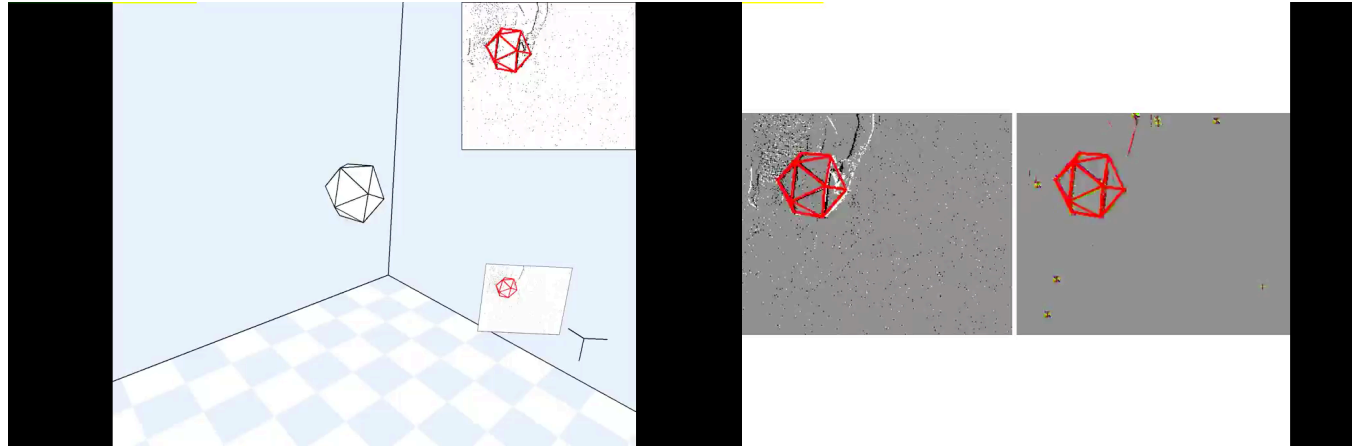
VS

Batch or Frame

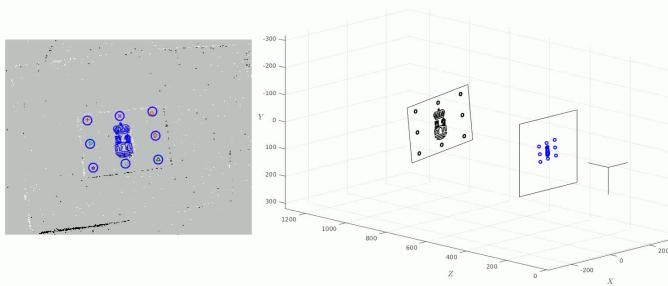


Compute the new 1 pixel-change frame and compute result using the whole frame

Incremental Event Vision: everything can be written as an incremental event process



Event-Based Solution to the PnP Problem

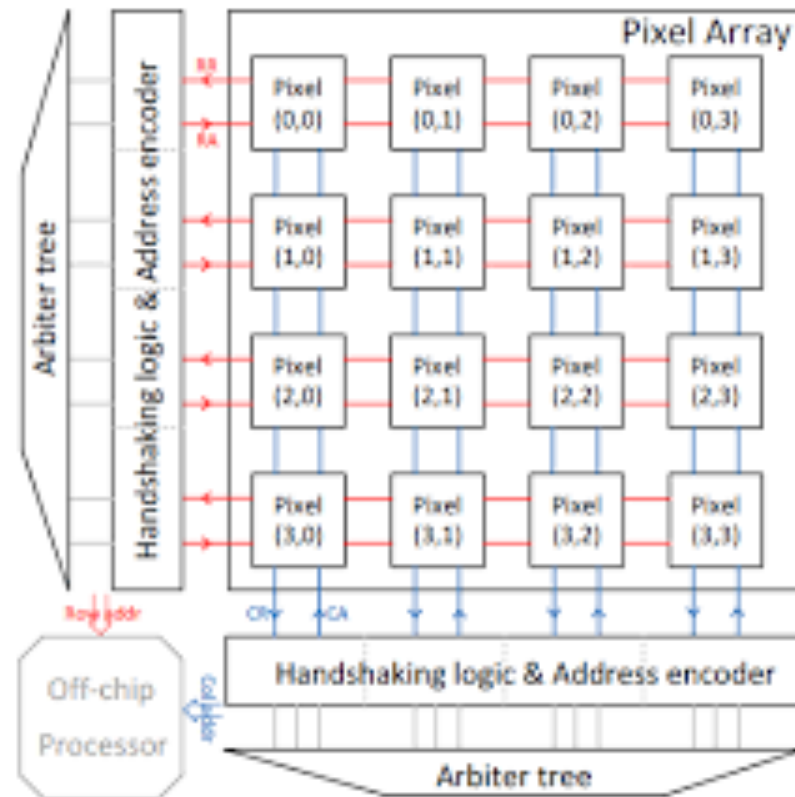


Event-based Face Detection in the Blink of an Eye

Where we are heading to

Replace arbiters

- Current locks are caused by the arbiter that scrambles event times and makes computation difficult
- Event Cameras are becoming compressed Imagers.

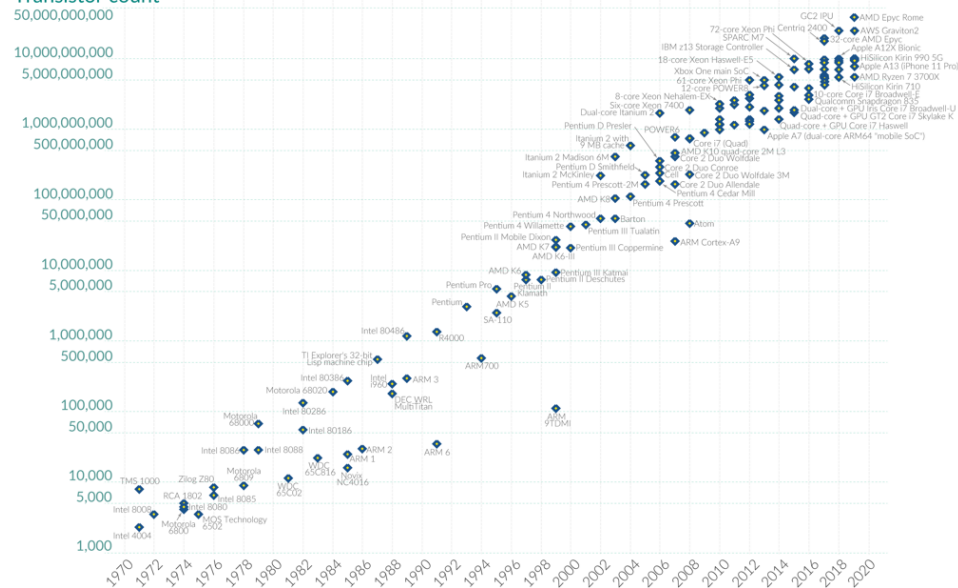


Where we are heading to: the promise

Moore's Law: The number of transistors on microchips doubles every two years

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

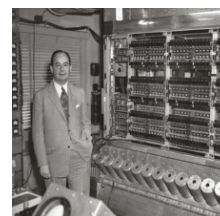
Transistor count



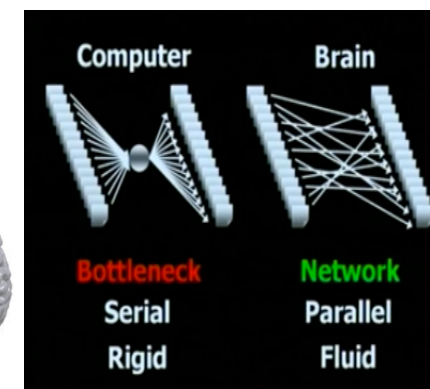
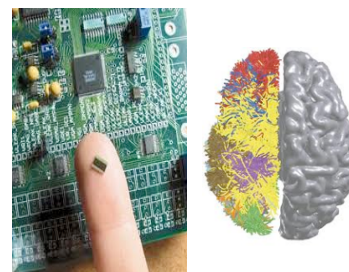
Data source: Wikipedia ([wikipedia.org/wiki/Transistor_count](https://en.wikipedia.org/wiki/Transistor_count))

OurWorldinData.org - Research and data to make progress against the world's largest problems

Licensed under CC-BY by the authors Hannah Ritchie and Max Roser



Von Neumann
bottleneck



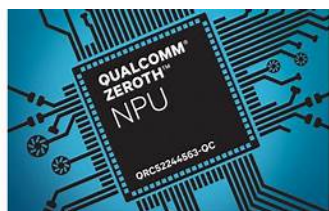
Promise of a new path solving for:

Plateau in current main stream computing technologies

Low Power and Latency

Where we are heading to

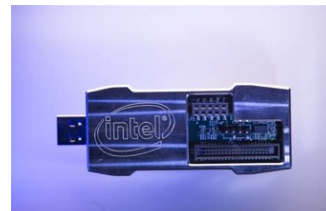
Existing Neuromorphic Processing Hardware is based on what we know of the brain



Qualcomm Zeroth (2013)



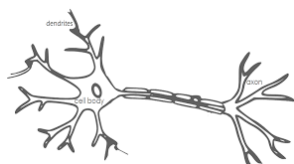
IBM TrueNorth (2014)



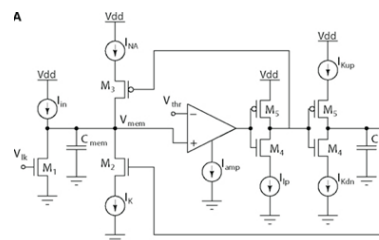
Intel Loihi (2017)



BrainChip (2019)



replicate



Existing hardware is based on the concept of replicating biological neurons into silicon

Limited use cases!

Where we are heading to

Replicating nature's solutions is not always the optimal path to solve an engineering problem.



understand



Understanding rather than replicating

There is a need to find the right level of abstraction

Where We Are Heading To



- We should explore new forms of events' acquisition
- Find a better link between images and events and see how to connect with decades of CV without losing the advantages of events
- We need a dedicated processor adapted to the temporal precision and sparseness of data of events and the amount of generated data
- Event cameras are the future if we explore their temporal properties
- New kind of engineers that understand neurosciences where « biological » events are studied