Event-Driven Sensing and Processing for Humanoid Robots

iit iCub
ISTITUTO ITALIANO
DI TECNOLOGIA

Smart Autonomous Systems

Neural vs Digital

Digital outperforms Neural

- High precision, fast, numerical computation

Neural outperforms Digital

- Processing of ambiguous data and generation of an appropriate output behaviour
- Interaction with Real World in Real Time
- Example: Perception (size, position and orientation invariant object recognition)

Digital vs Neural

- largely serial computation
- variables: “0”, “1”
- fast, precise computing elements

- largely parallel computation
- variables: analog
- slow, imprecise computing elements (neurons, synapses)
- adaptive (plasticity), context dependent computation

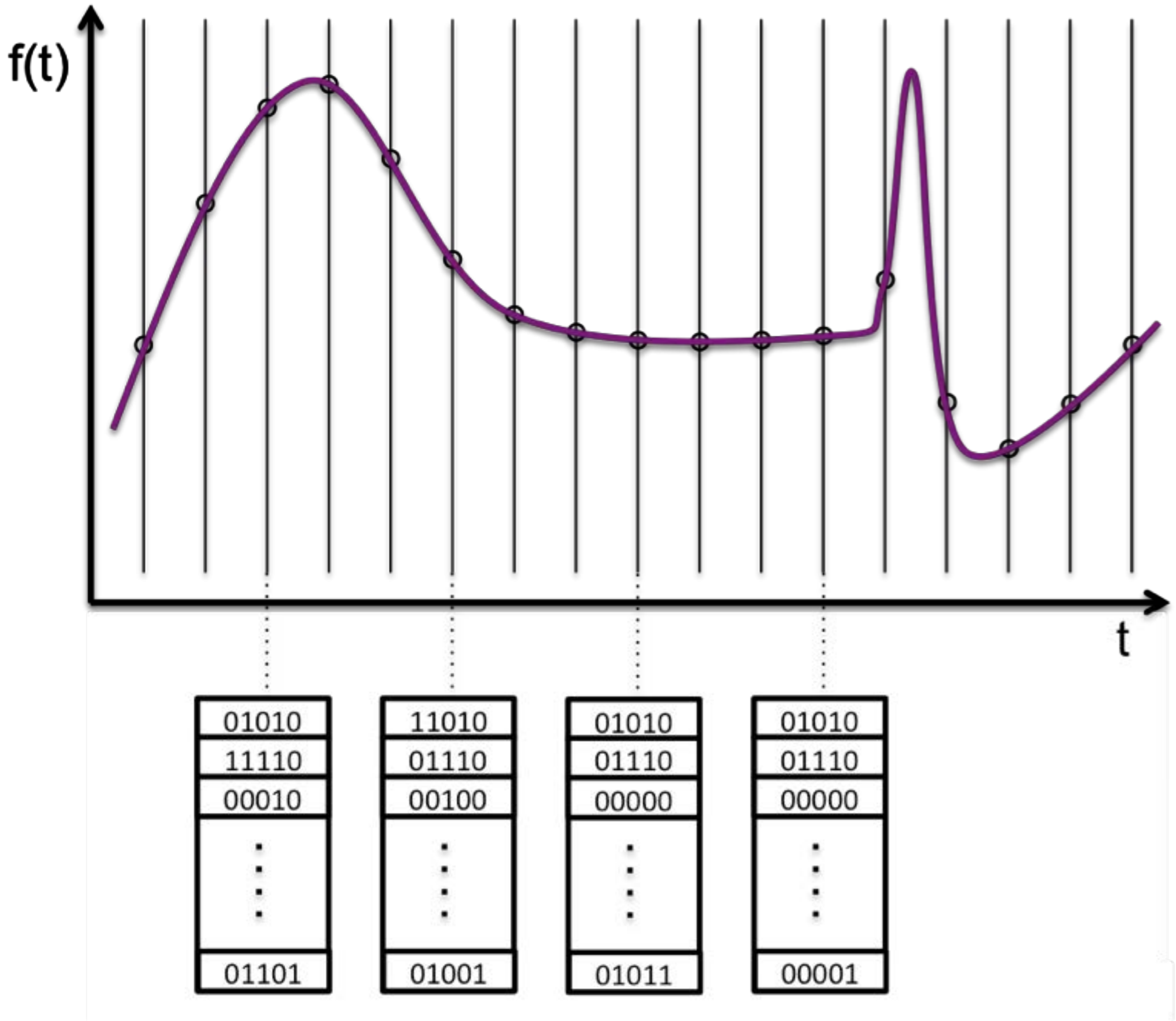
Sensory Acquisition and Processing

- Clock-Driven

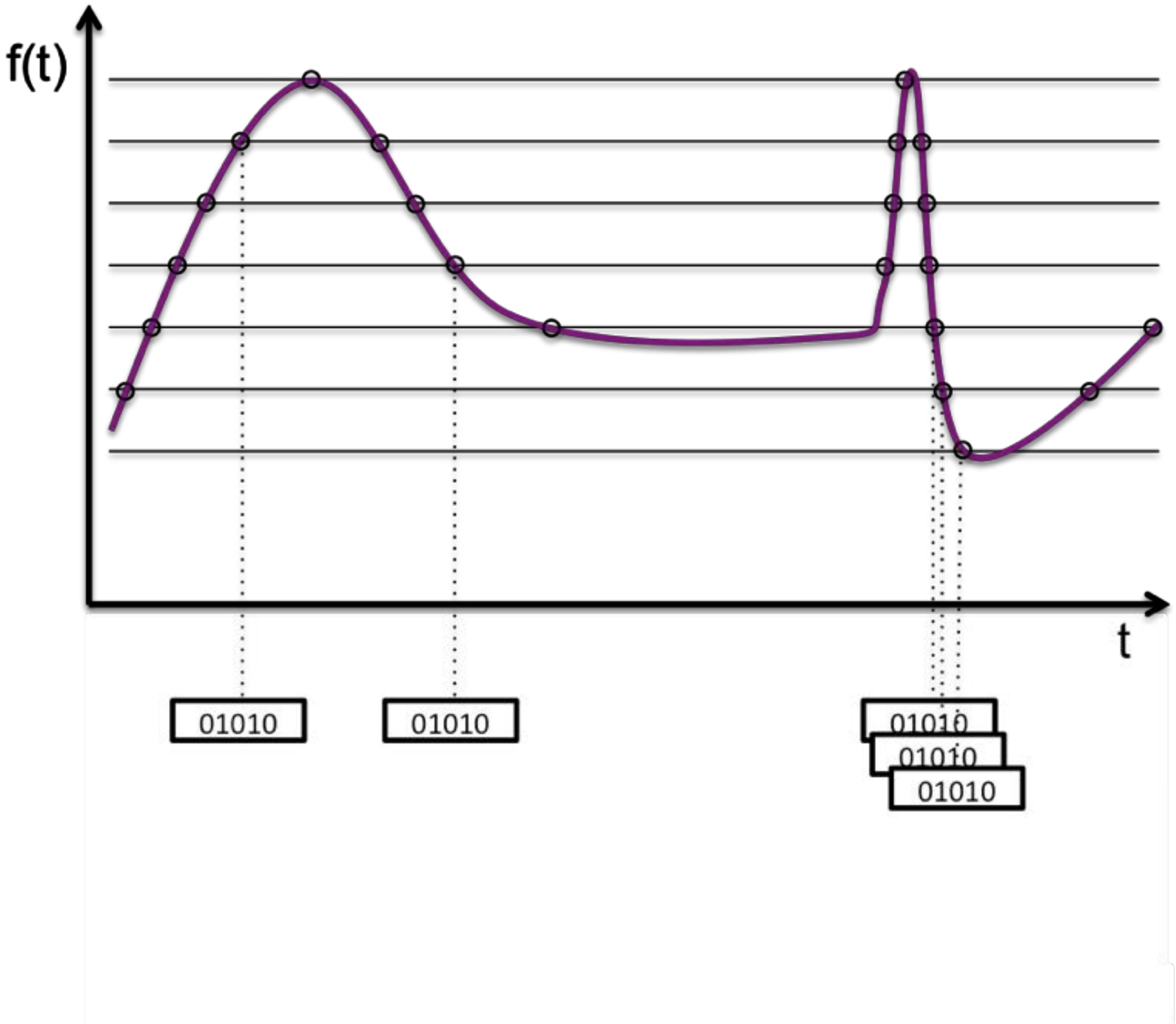
- Event-Driven

Event-Driven Sensing and Processing

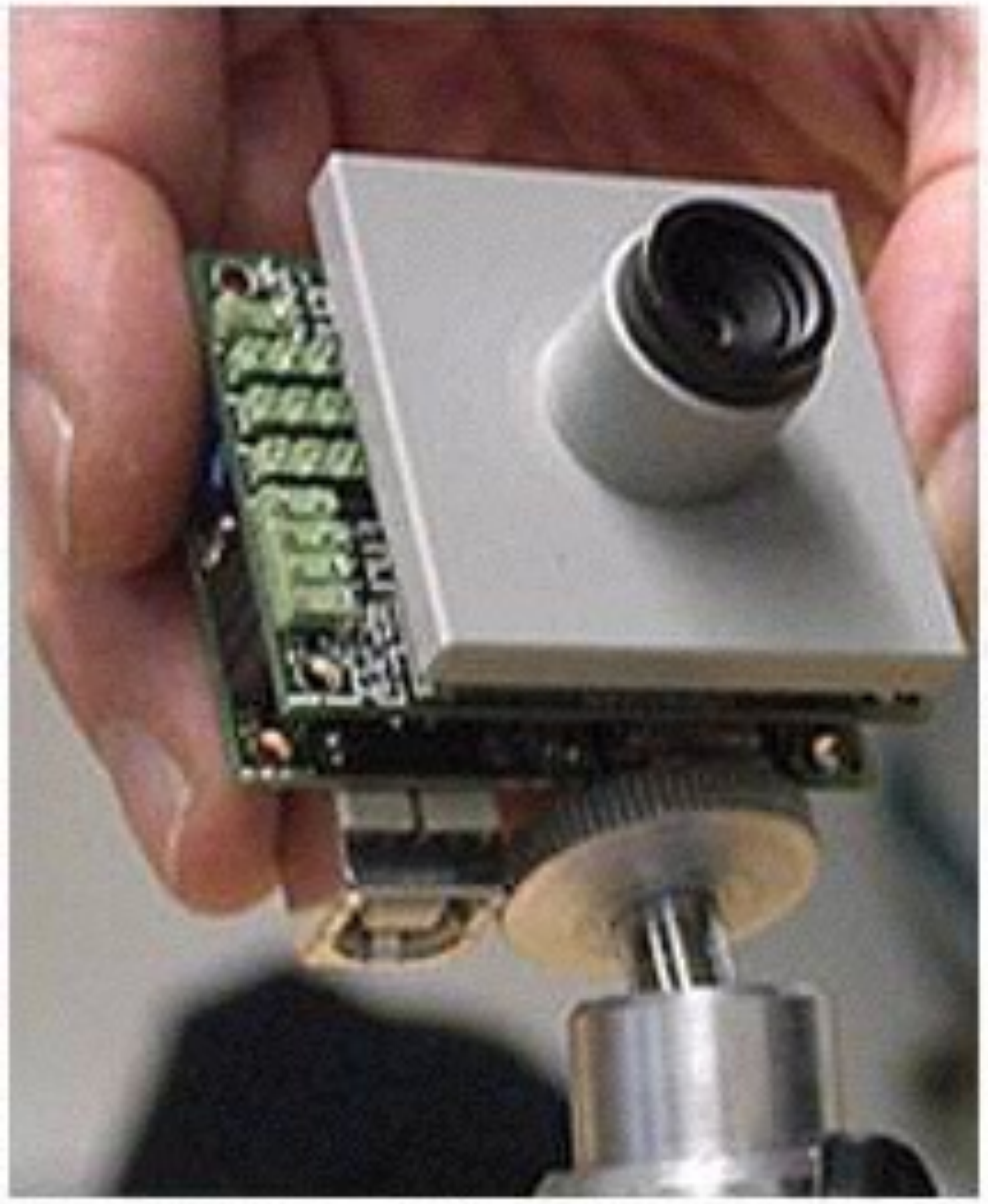
Clock-Based Sampling — fixed Δt

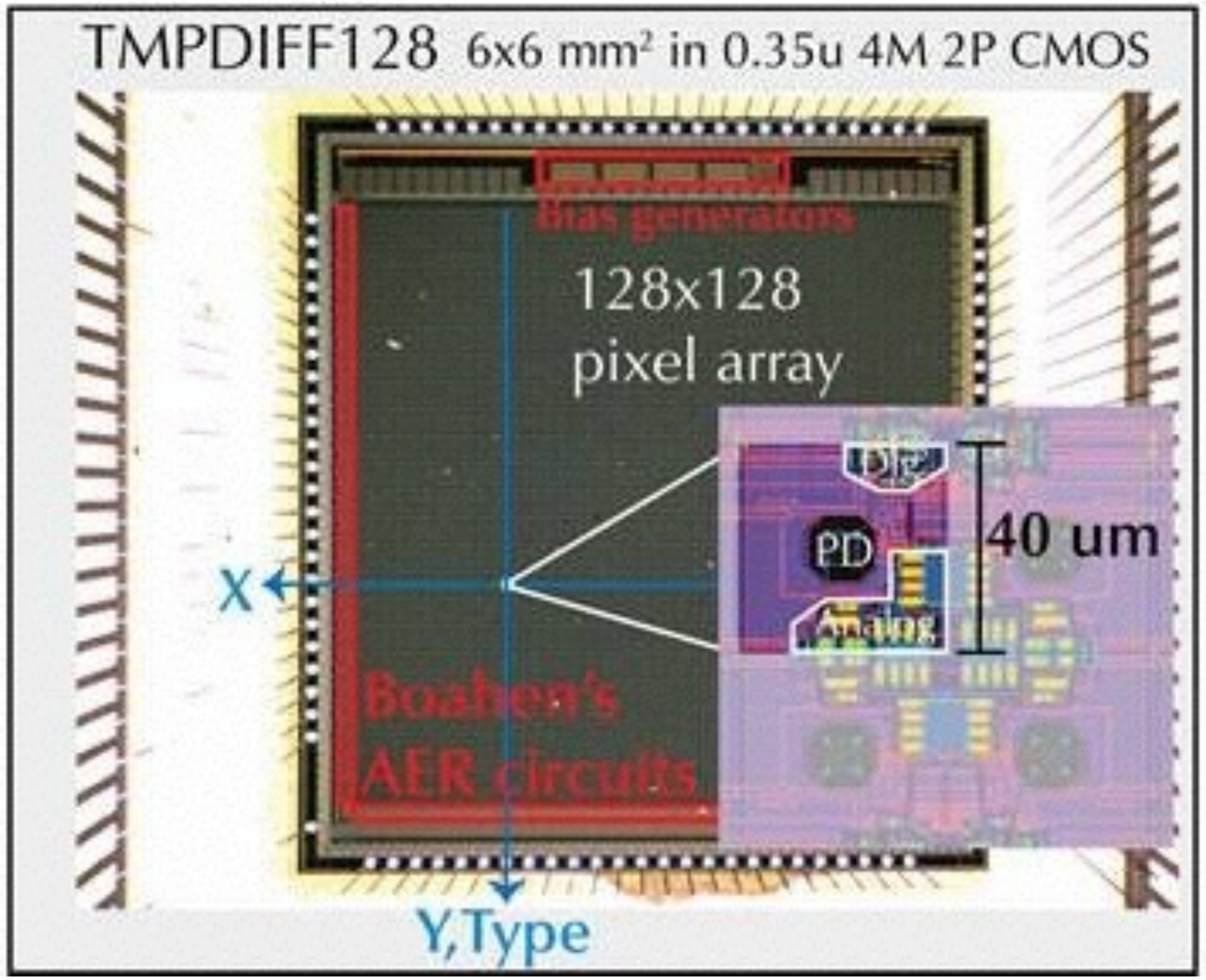


Data-Driven Sampling — fixed Δf



Event-Driven Sensing — Dynamic Vision Sensor





120dB dynamic range

15 µs latency

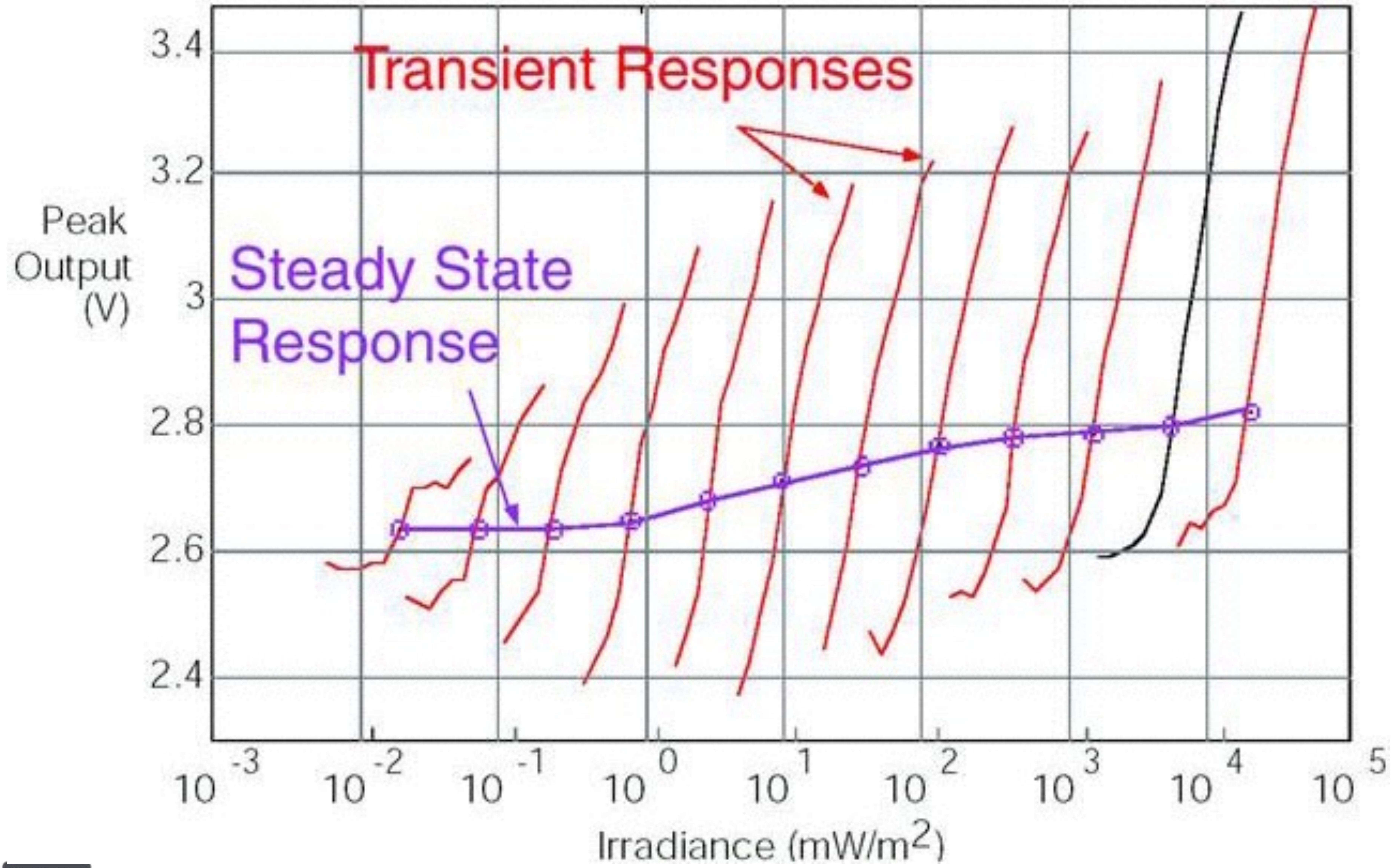
25 ns temporal resolution

23mW power consumption

Event-Driven Sensing — Dynamic Vision Sensor

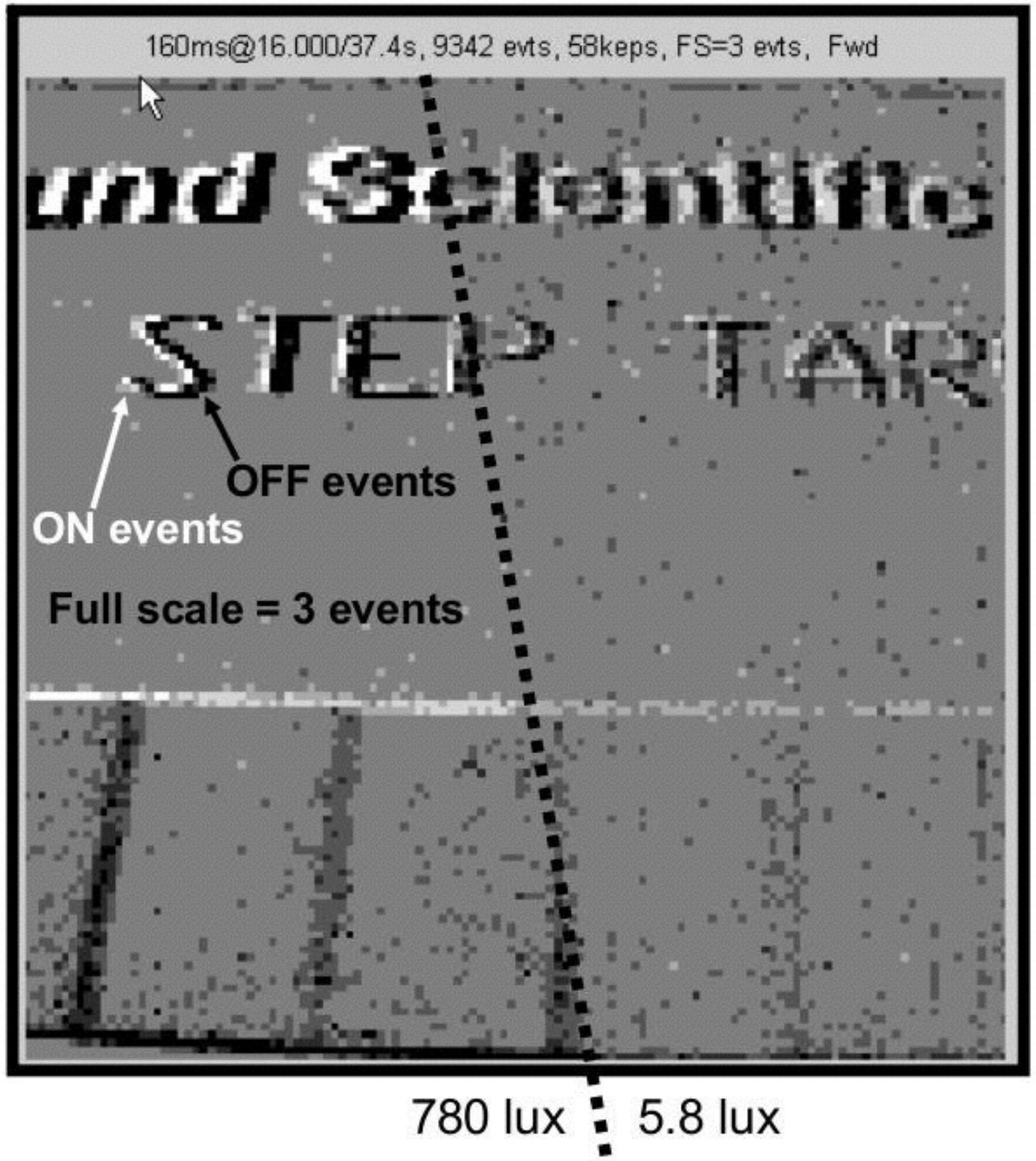
Dynamic Range

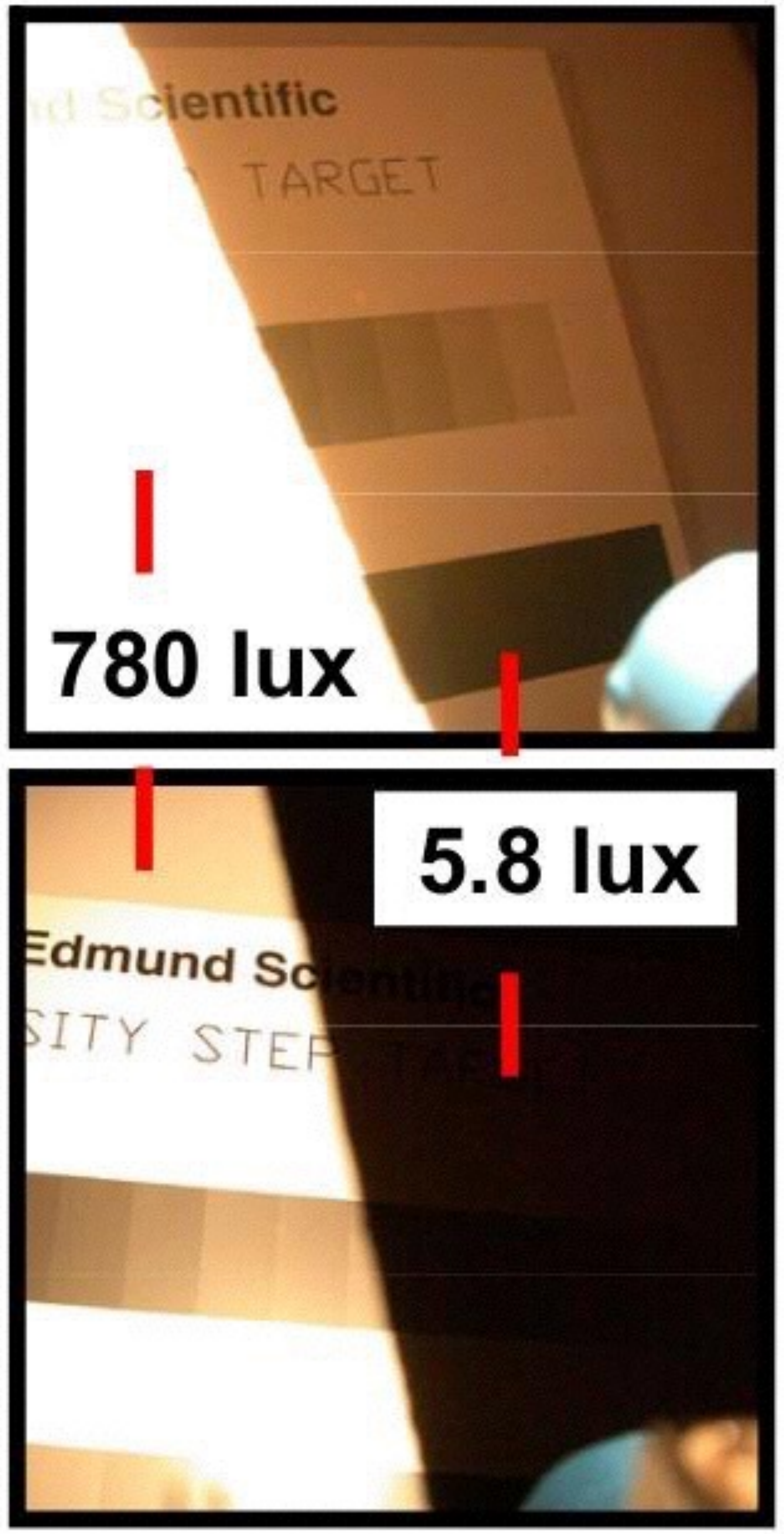
Event-Driven Sensing — Dynamic Vision Sensor



$$\frac{d\log(I)}{dt} = \frac{1}{I}\frac{dI}{dt}$$

Event-Driven Sensing — Dynamic Vision Sensor





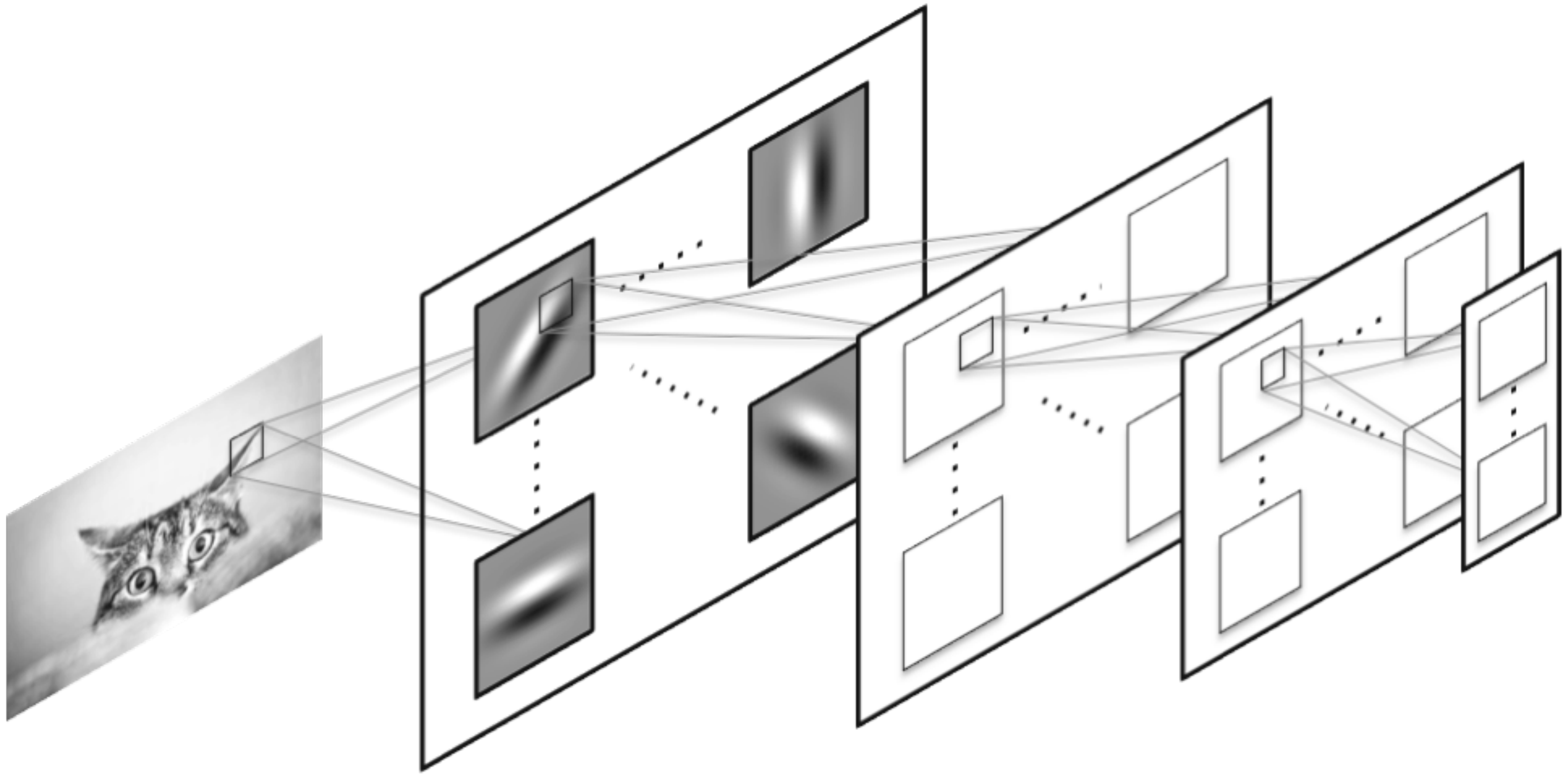
Edmund 0.1 density chart
Illumination ratio=135:1

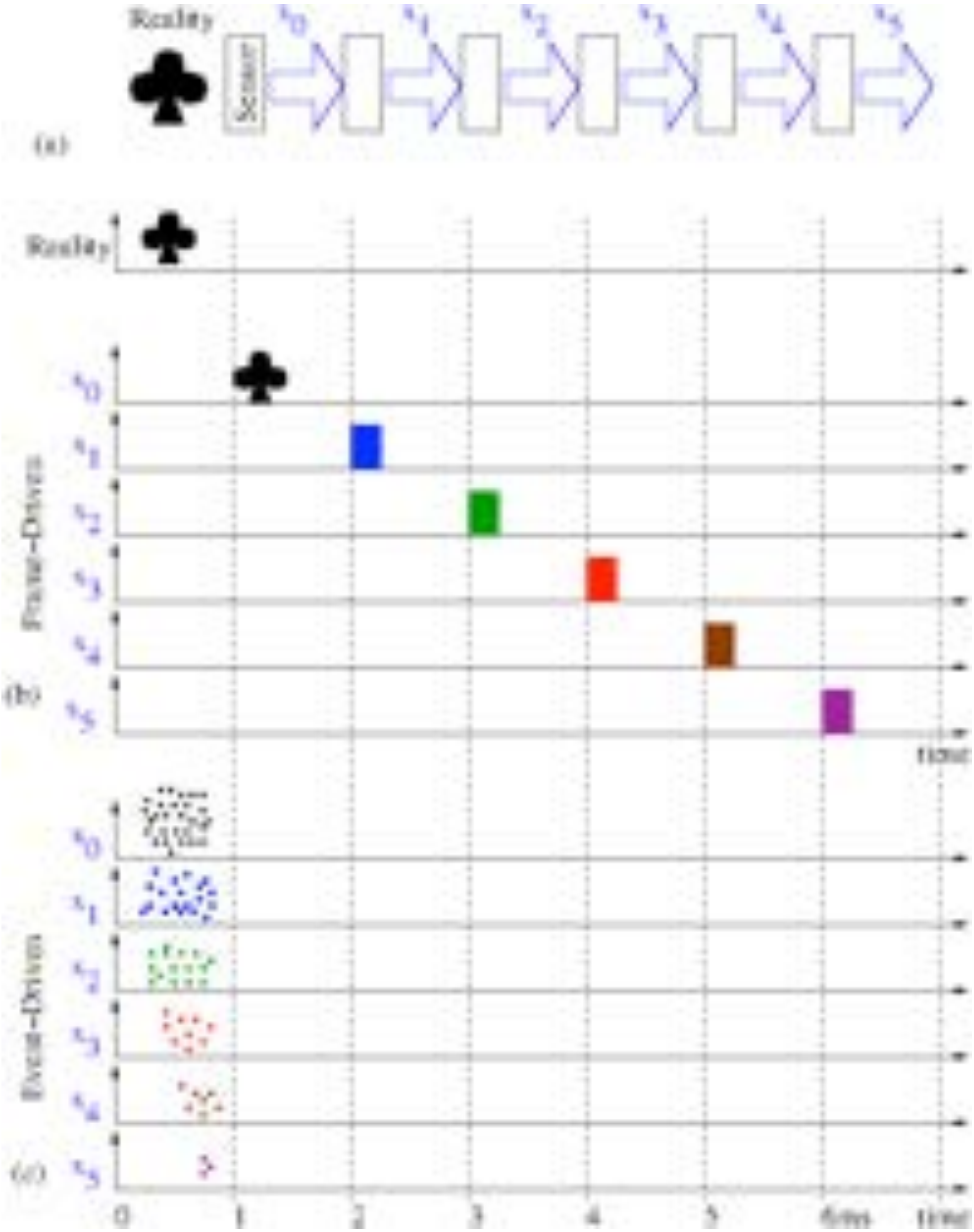
120dB dynamic range

Clocked vs Event-Driven Vision

Conventional high-speed vision systems	Event-driven vision	Benefits
Requires fast PC/GPU/...	Works with any laptop, better with spike-based hardware	Lower costs, Lower power consumption
Highly redundant data, Extremely large data storage (often several TB)	No redundant data, Low storage requirements	Lower costs, More portable, Easier and faster data management
Batch-mode acquisition Off-line post-processing	Real-time acquisition Extremely low latency	Continuous processing No downtime, lower costs
Low dynamic range ordinary sensitivity	High sensitivity	Lower costs
Needs special bright lighting (lasers, strobes, etc.) for short exposure times	No special lighting needed	Simpler data acquisition
Limited dynamic range, typically 50 dB	Very high dynamic range 120 dB	Usable in more real-world situations
Latency - need to wait for one frame to be acquired	Very short latency — 15μs	Enables extremely short reaction times

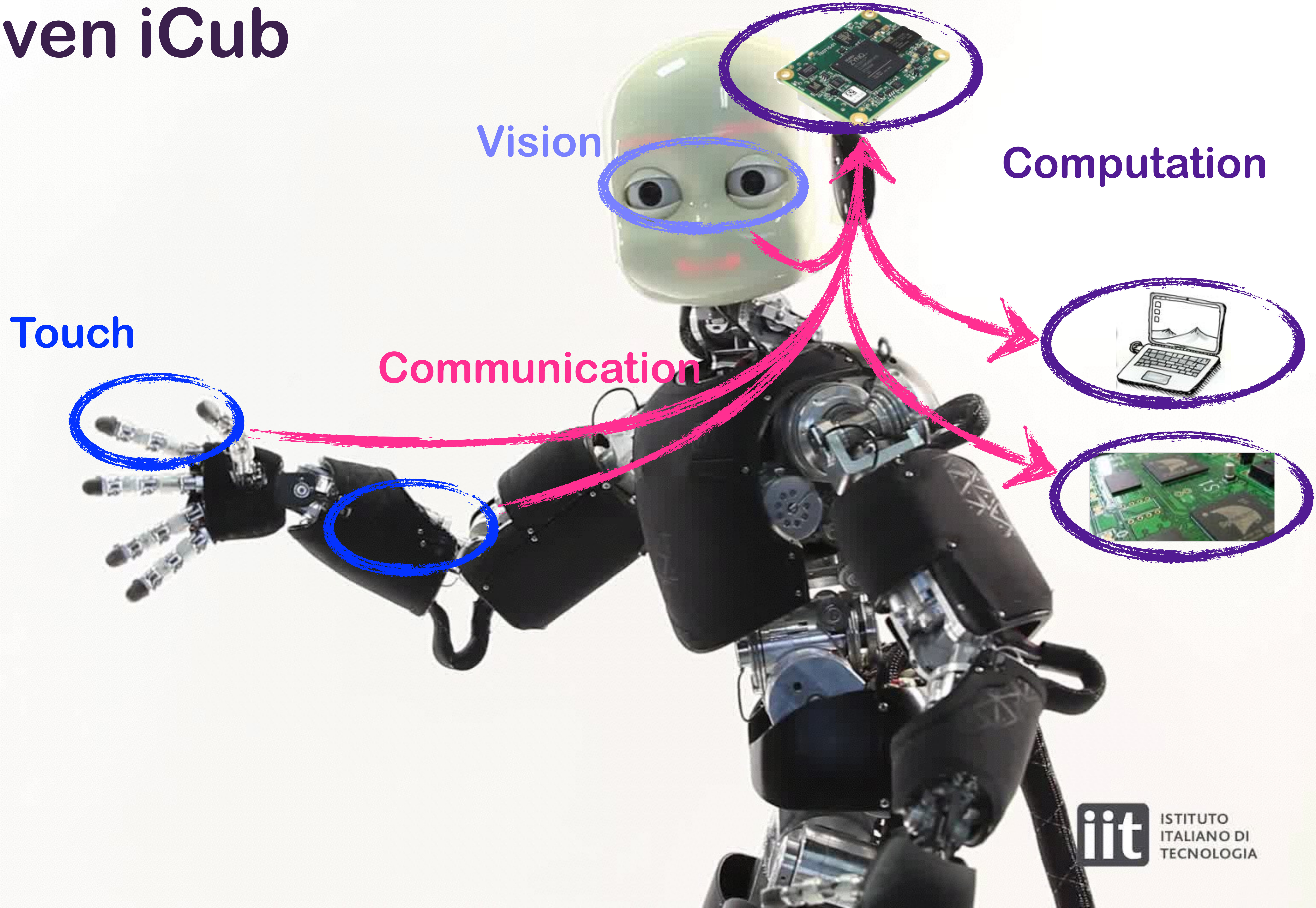
ED Vision – PseudoSimultaneous



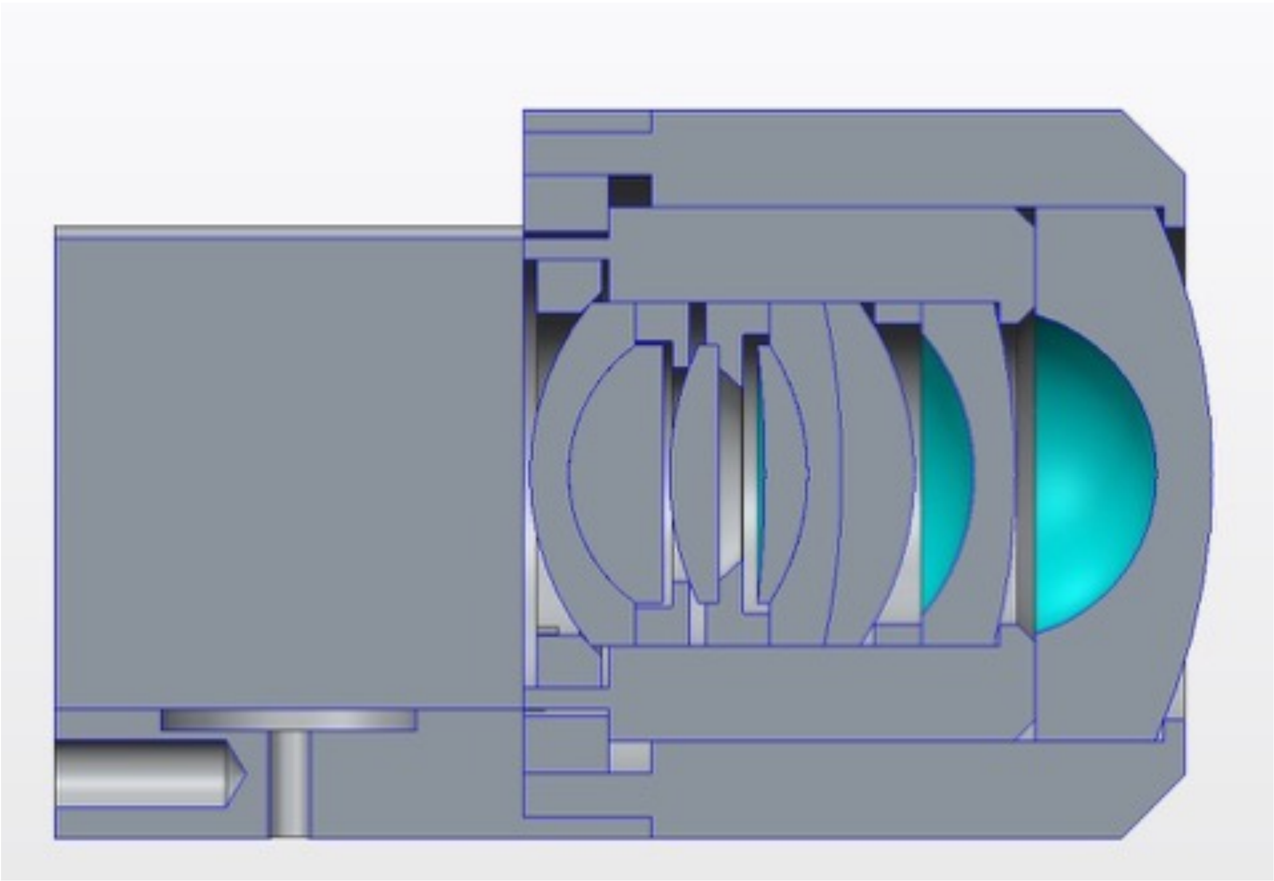


Courtesy of
B. Linares-Barranco

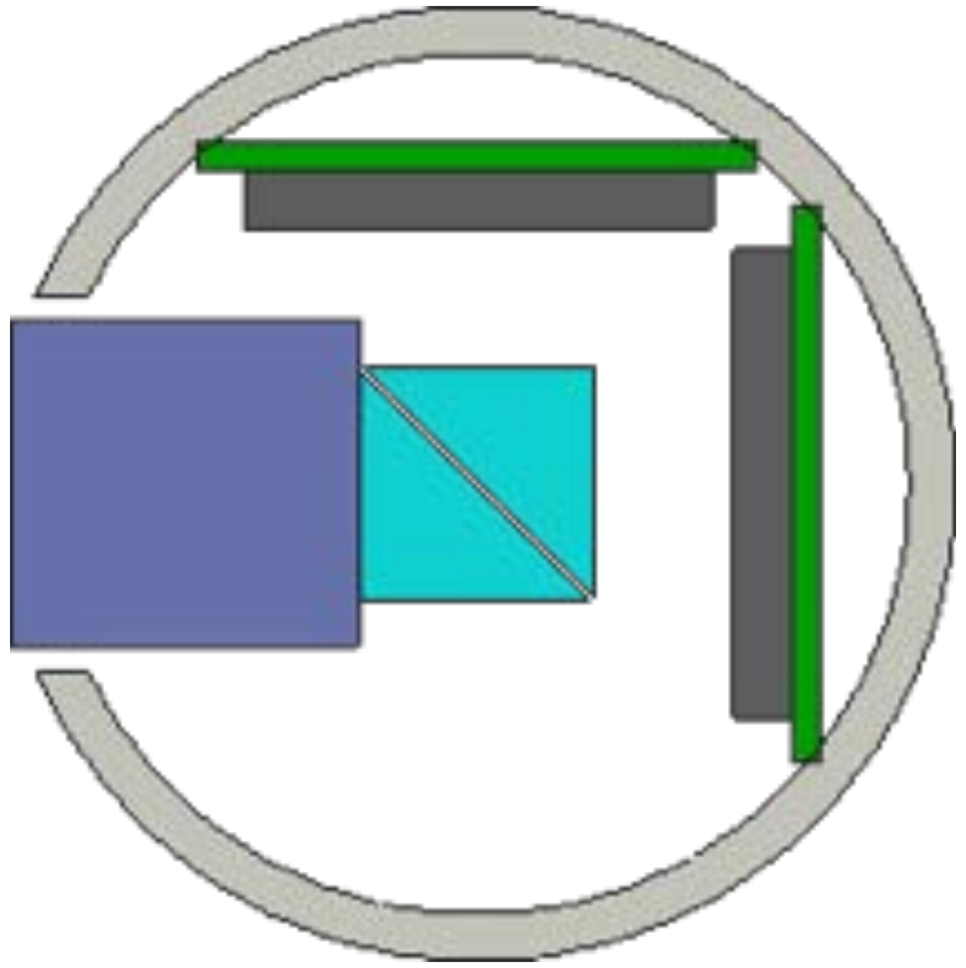
Event-Driven iCub

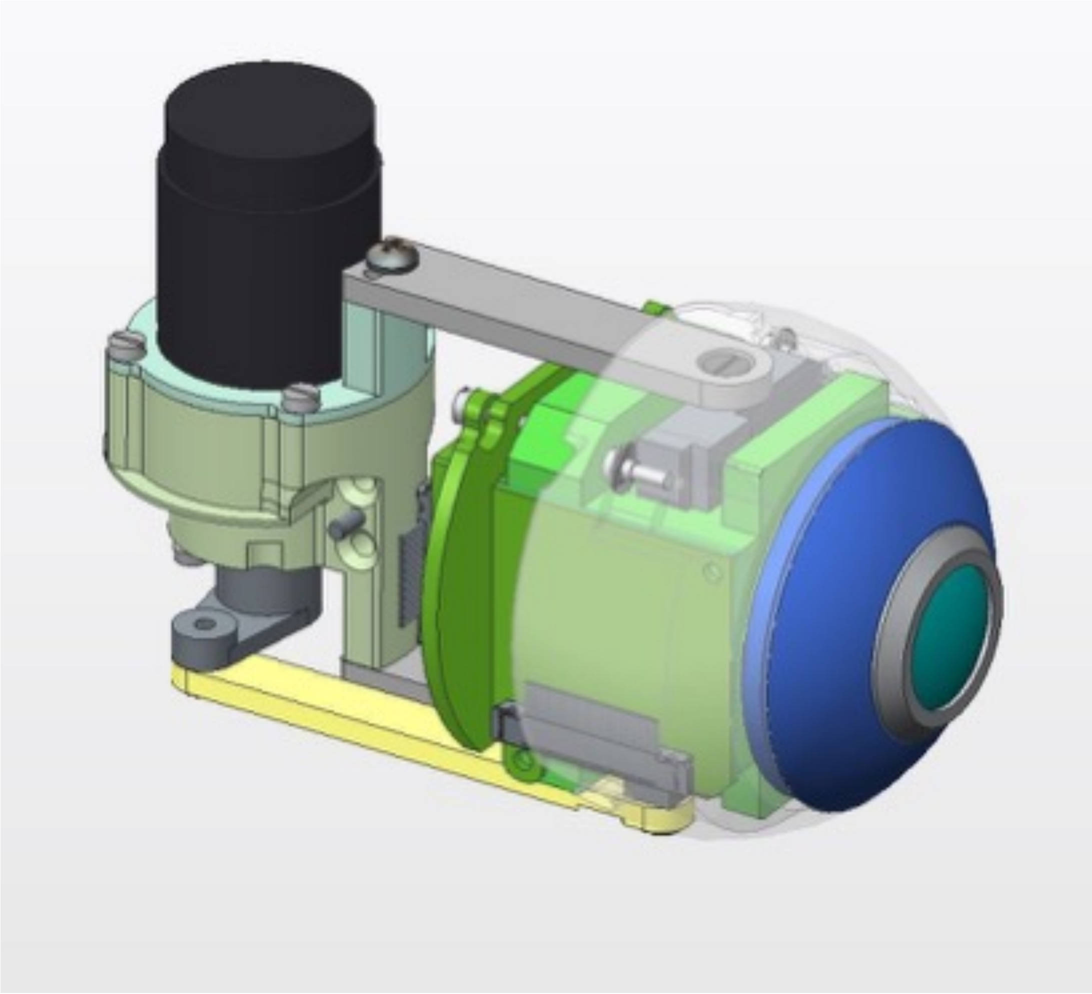


ED Sensors — Vision



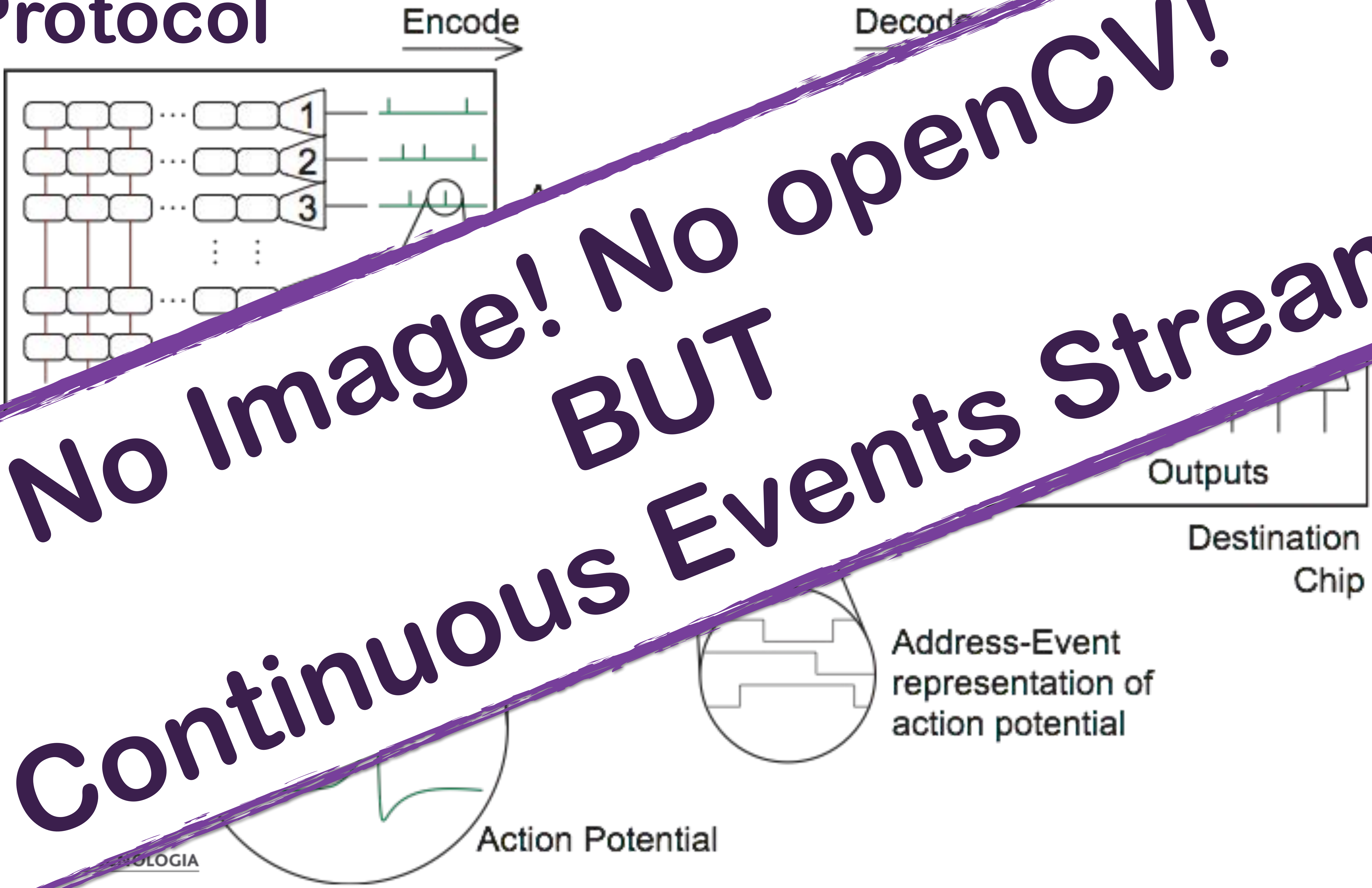
ATIS	CMOS
Spatial Res	240x304 2/3”
Temporal Res	1 µs
AFOV	61°





Python	CMOS
Spatial Res	1,3Mpxl 1/2”
Temporal Res	42 fps (24ms)
AFOV	42°

Event-Driven Sensing and Processing
AER Protocol



No Image! No openCV!
BUT
Continuous Events Stream

ED Vision

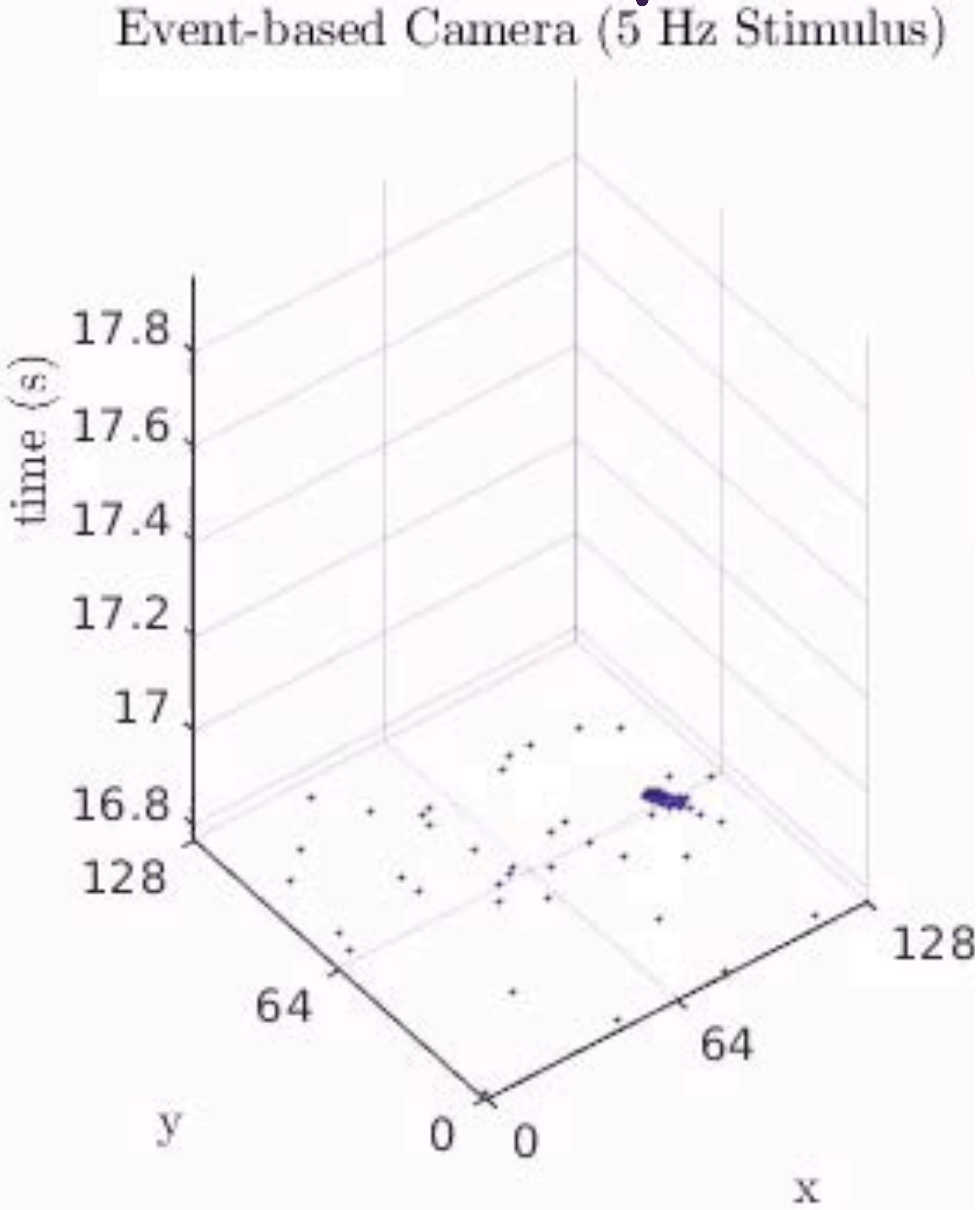
Integrate events and reconstruct a frame -> use openCV

Adapt Computer Vision algorithms to events -> ED Vision → Optical Flow
Corner detection

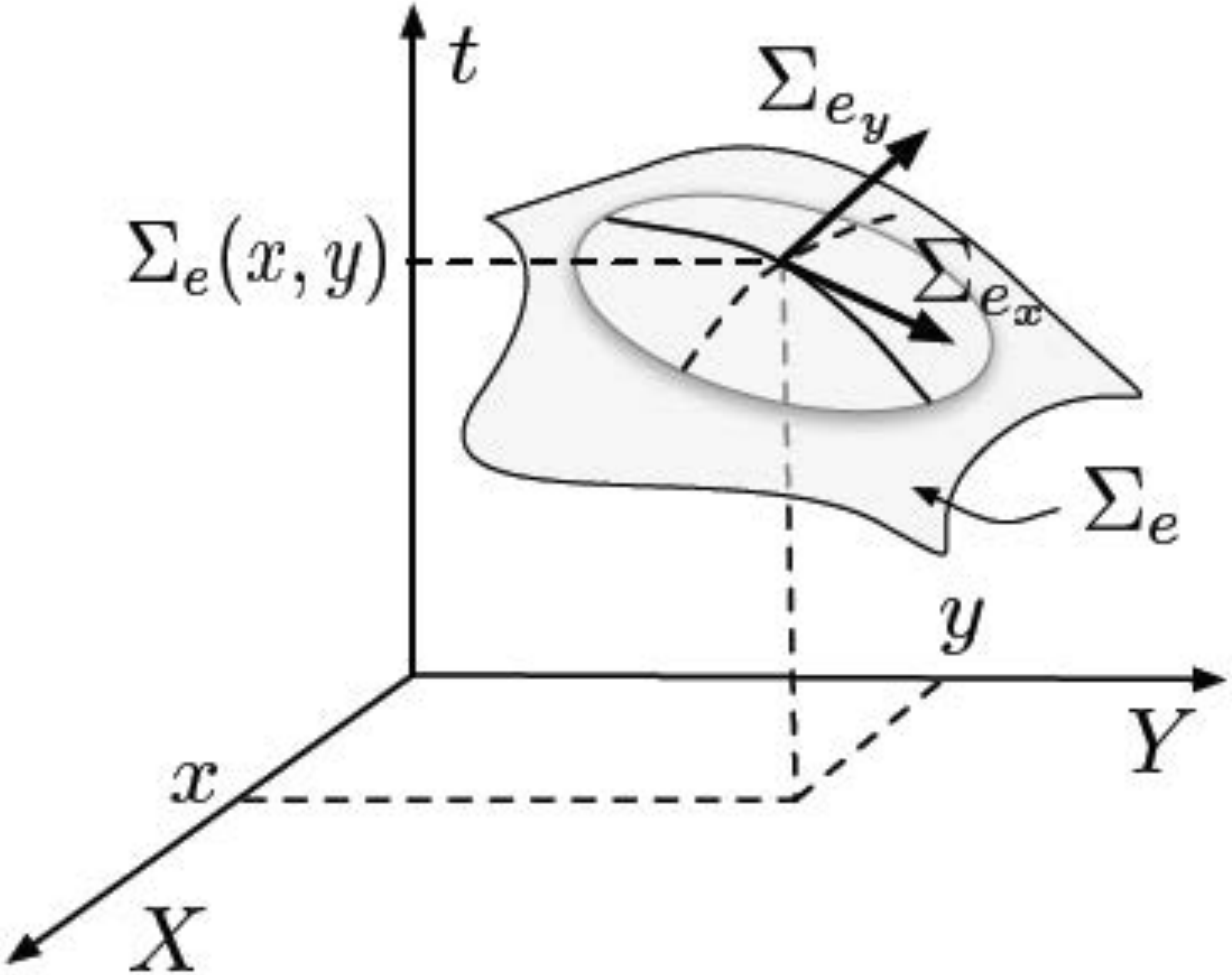
State Estimation -> Probabilistic Filtering → Object Tracking

Implement neurally-inspired models of visual processing → Disparity and vergence

ED Vision — Optical Flow



$$e(\mathbf{p}, \mathbf{t}) :$$
$$\Sigma_e : \mathbb{N} \to \mathbb{R}$$
$$p \longmapsto \Sigma_e(\mathbf{p} = \mathbf{t})$$

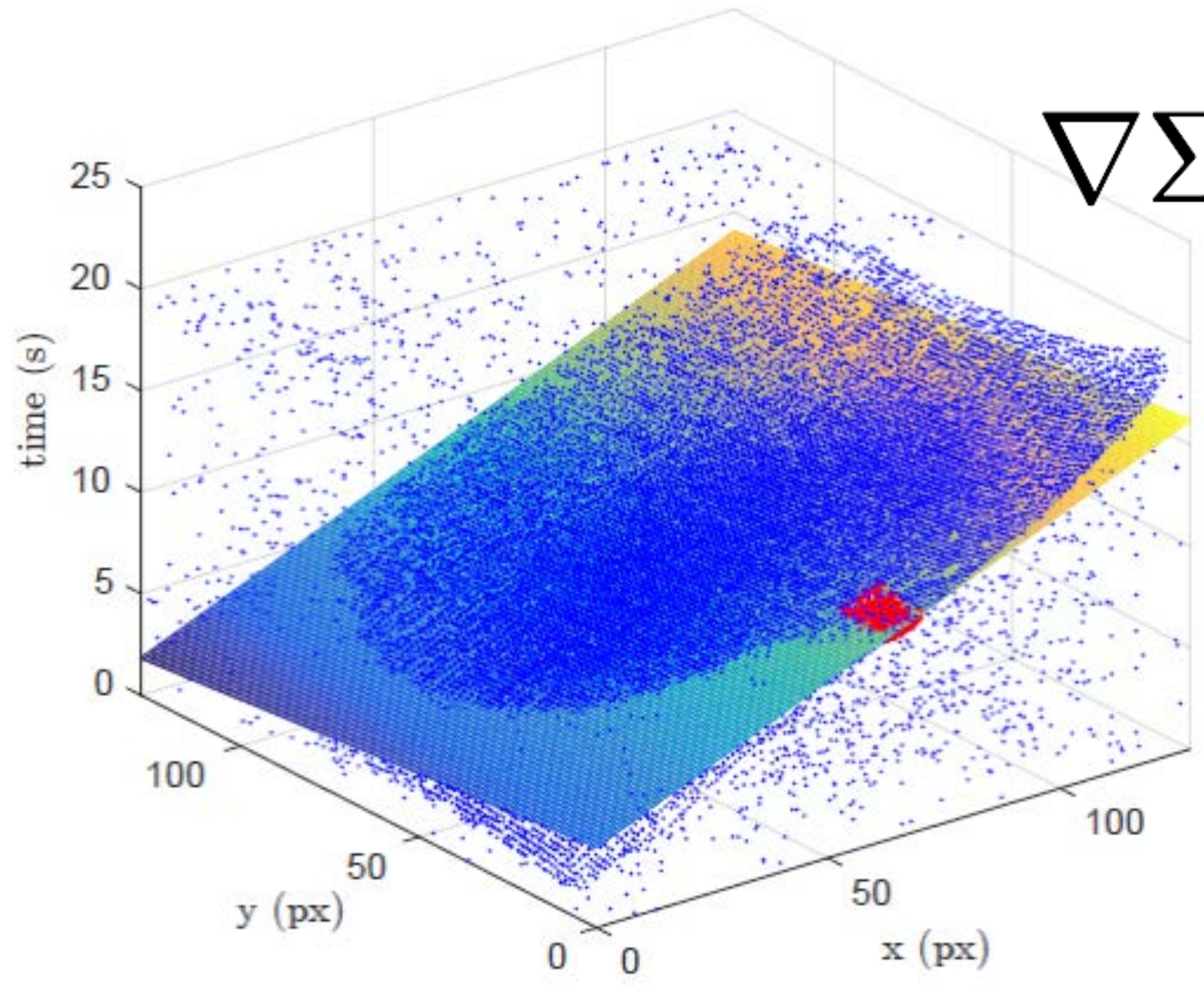


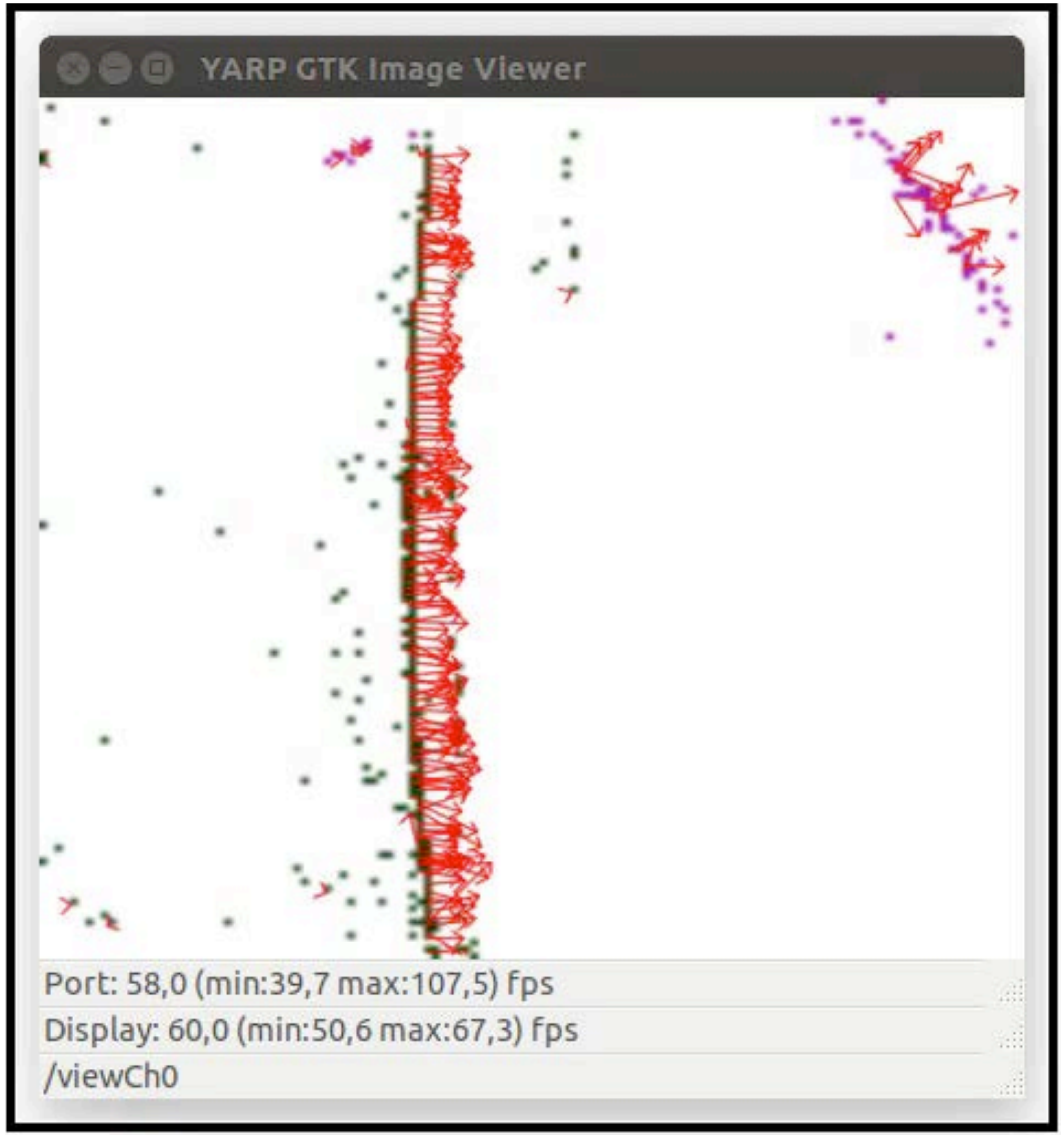
ED Vision — Optical Flow

$$\frac{\partial \Sigma_e}{\partial x}(x, y_0) = \frac{d\Sigma_e|_{y=y_0}}{dx} = \frac{1}{v_x(x, y_0)}$$

$$\frac{\partial \Sigma_e}{\partial y}(x_0, y) = \frac{d\Sigma_e|_{x=x_0}}{dx} = \frac{1}{v_y(x_0, y)}$$

$$\nabla \Sigma_e = \left(\frac{1}{v_x}, \frac{1}{v_y}\right)$$





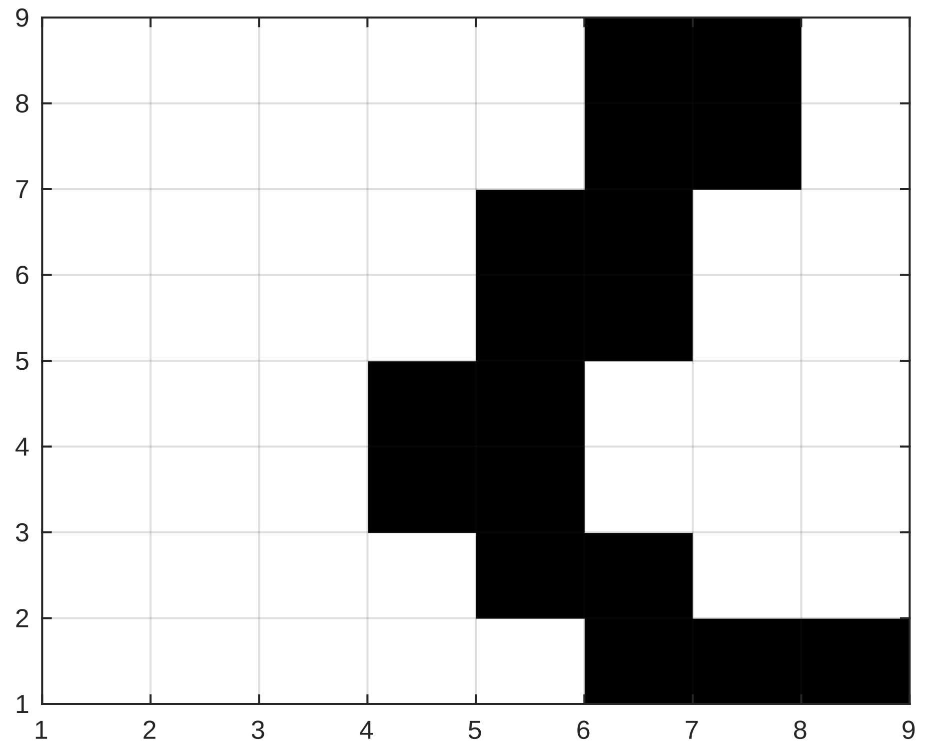
ED Vision — Corner Detection

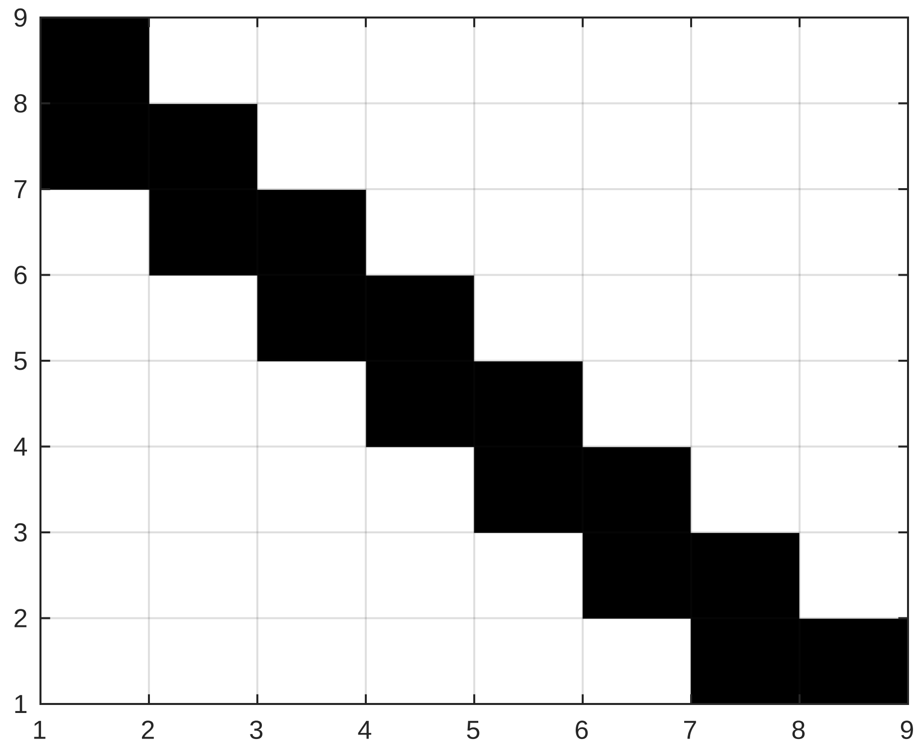
Binary Image Patch

- spatial derivatives

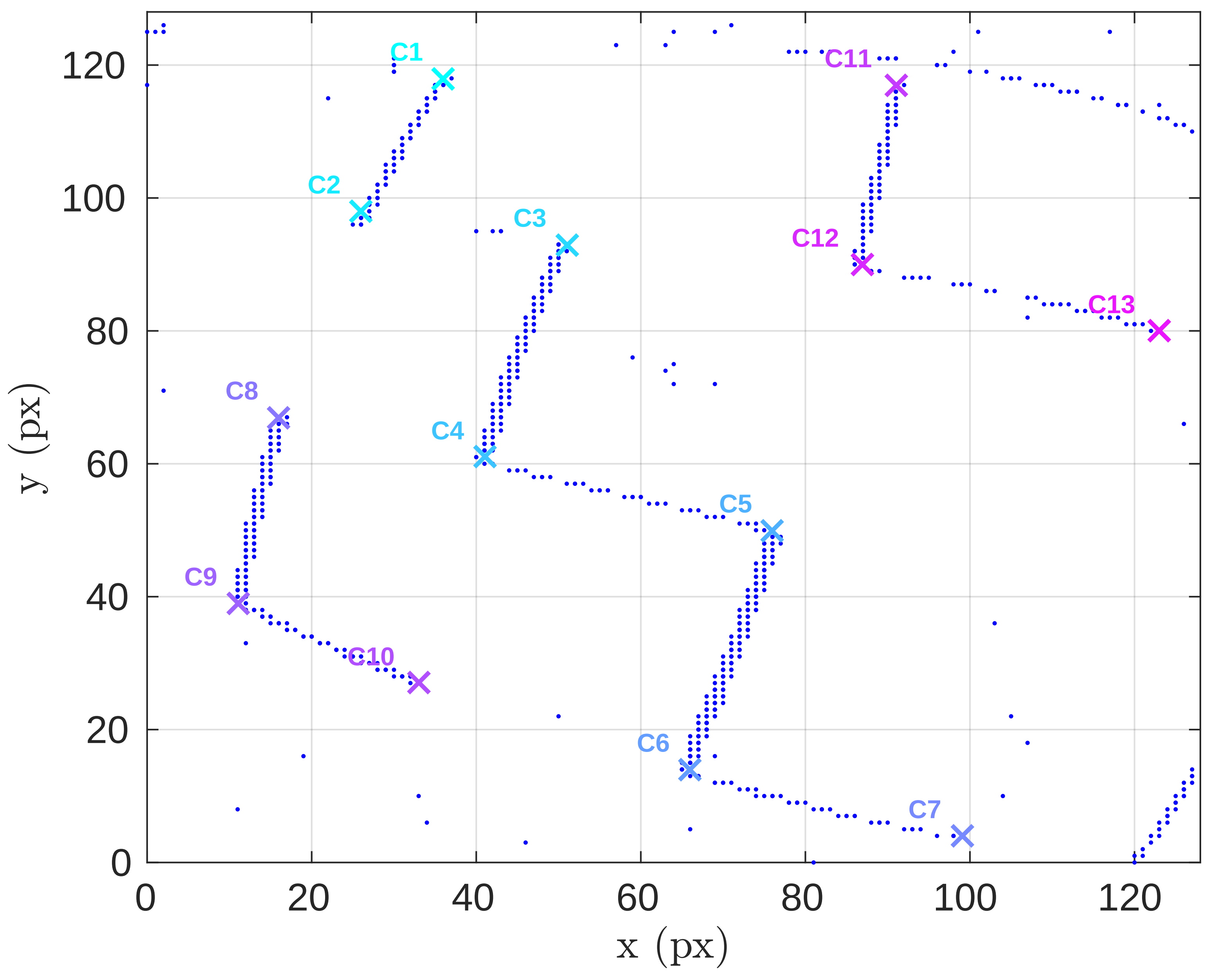
ED Harris Score — R

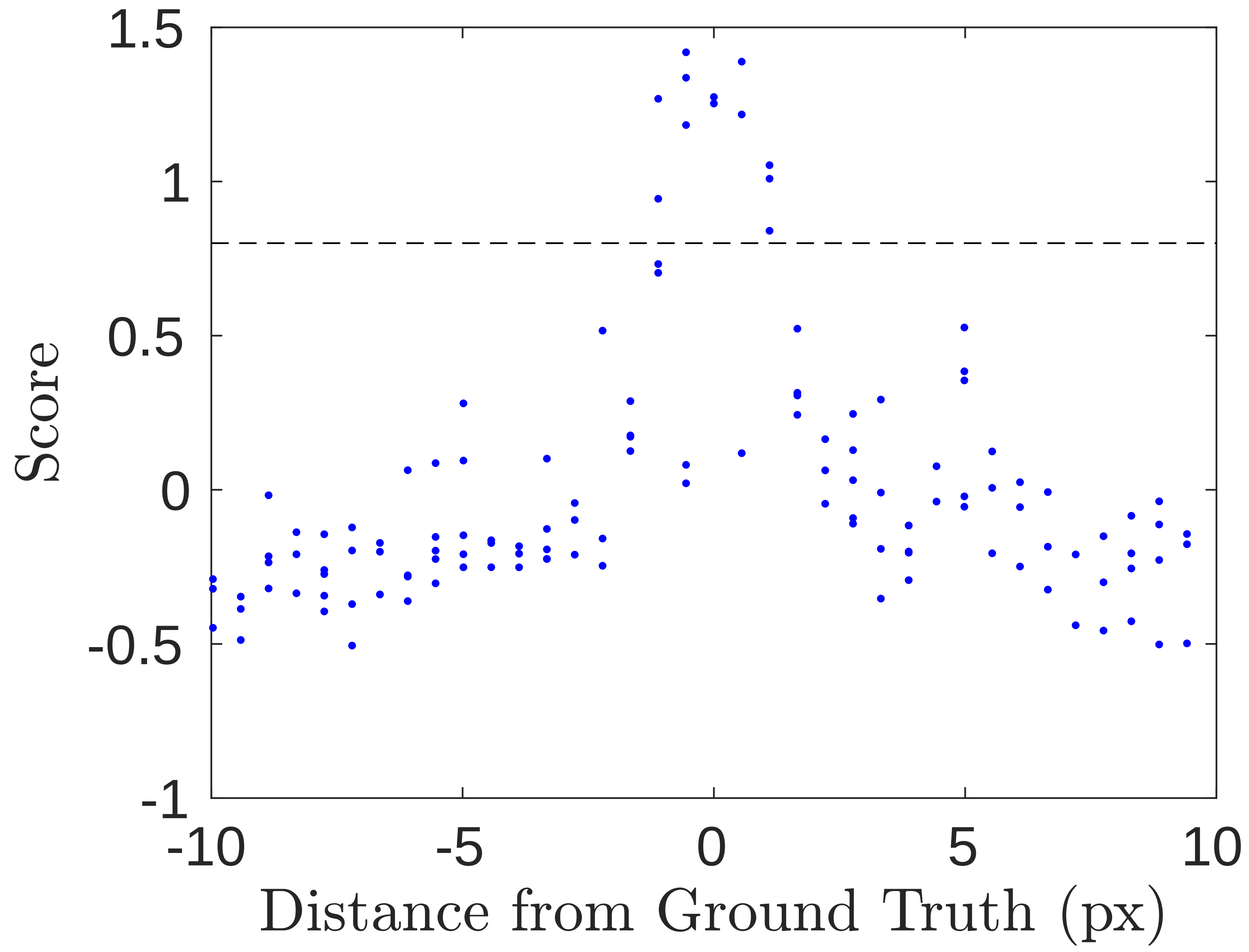
- flat: two small eigenvalues
  R small
- edge: one small and one big
  R negative
- corner: two large eigenvalues
  R big

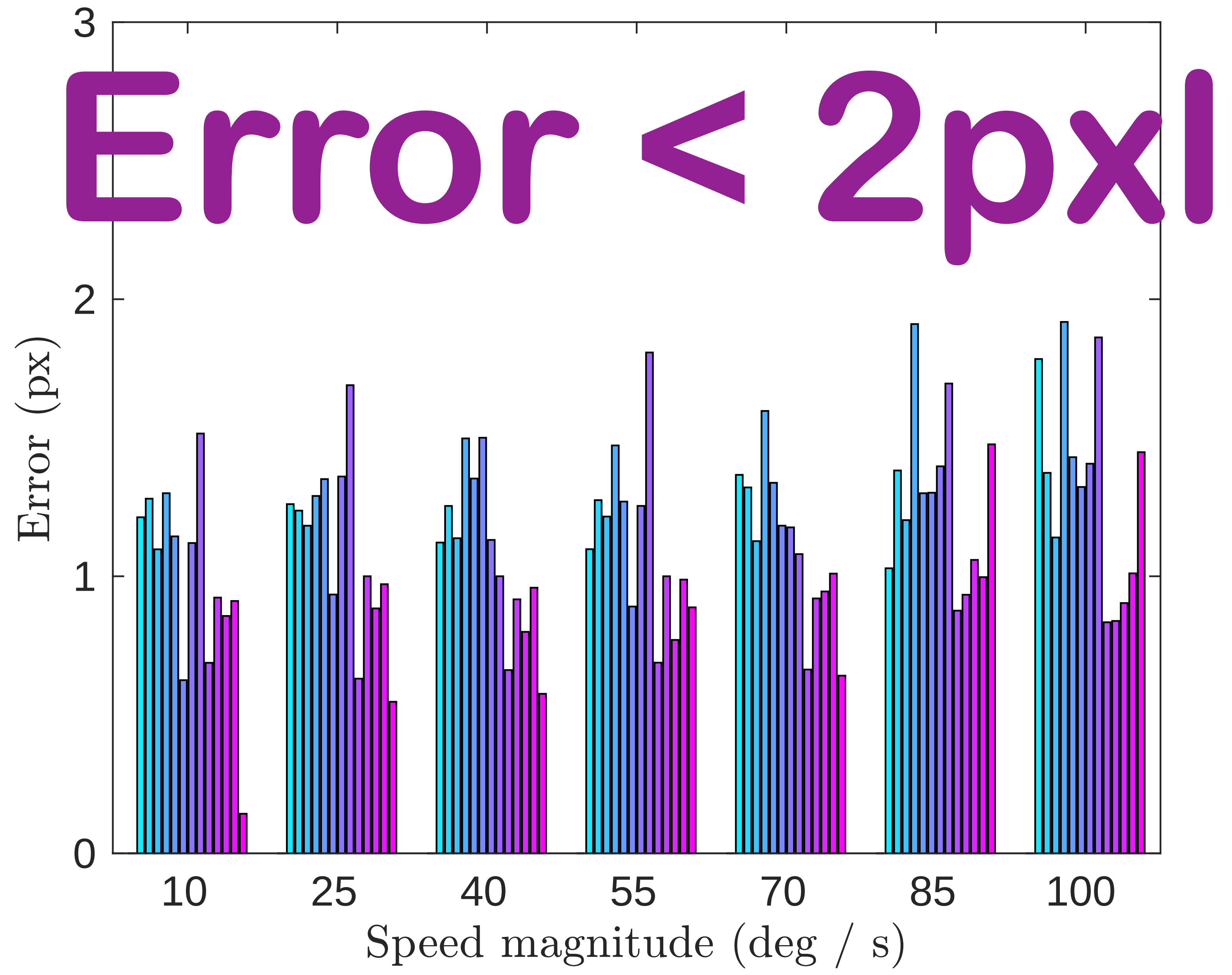


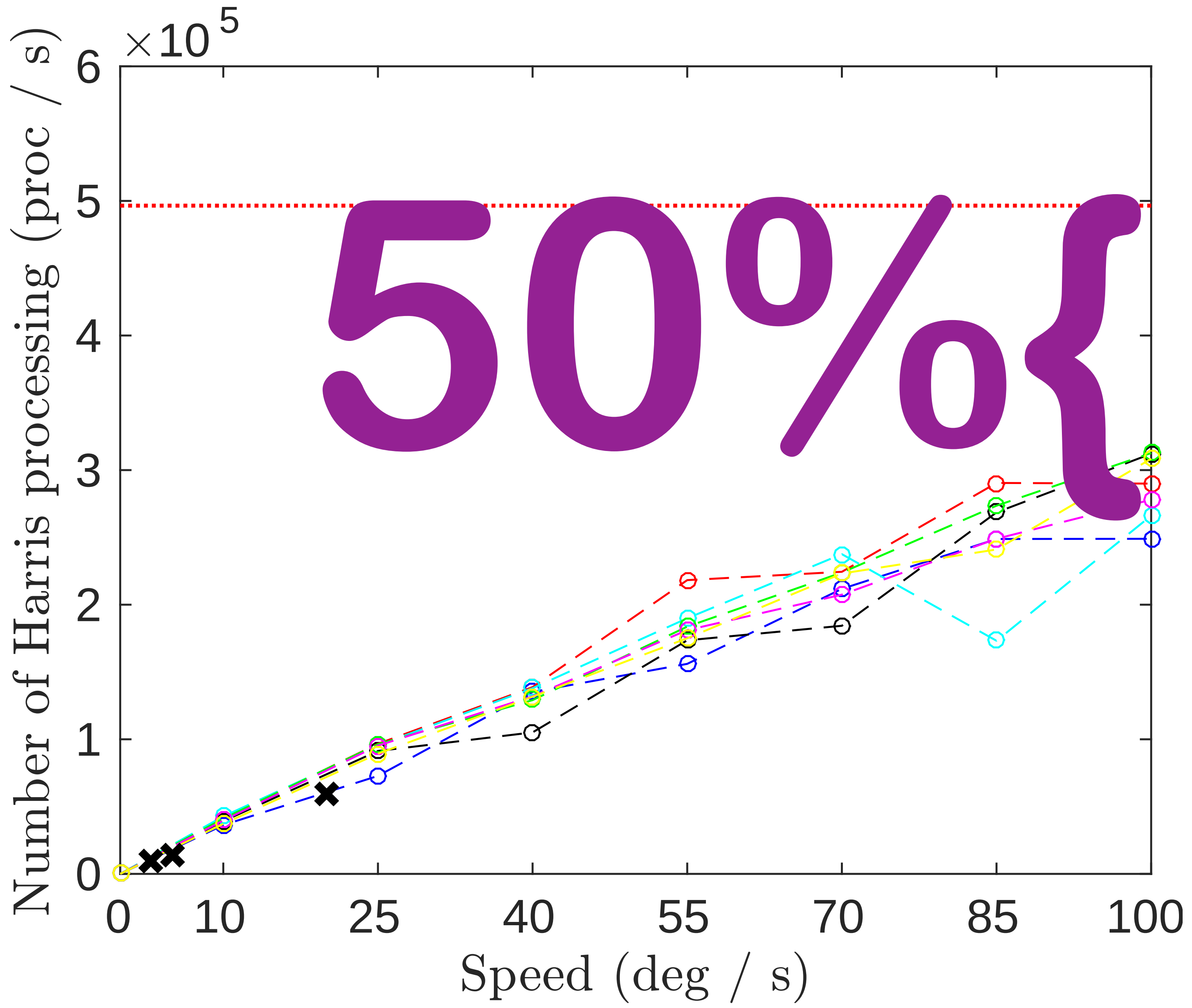


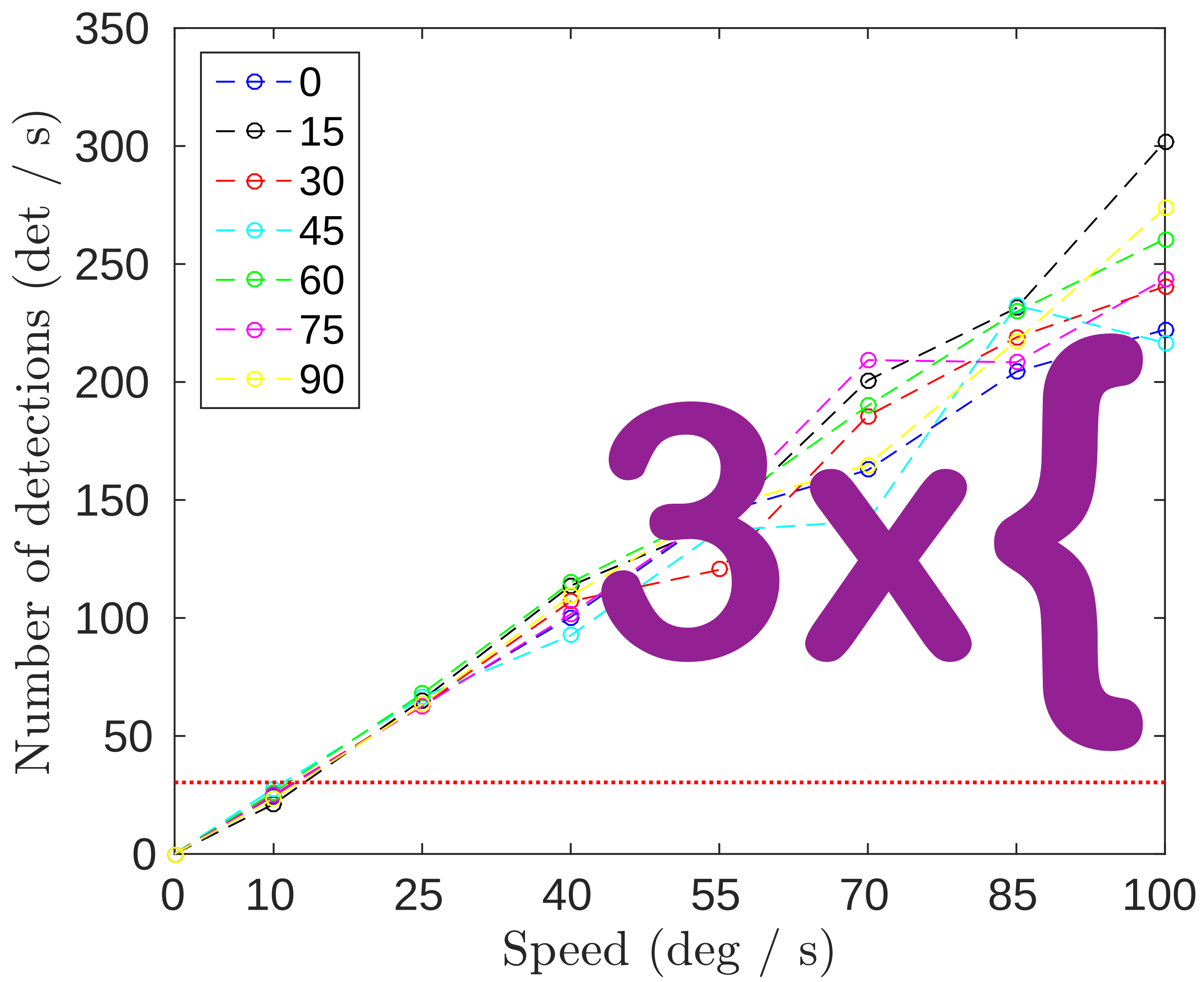
ED Vision — Corner Detection



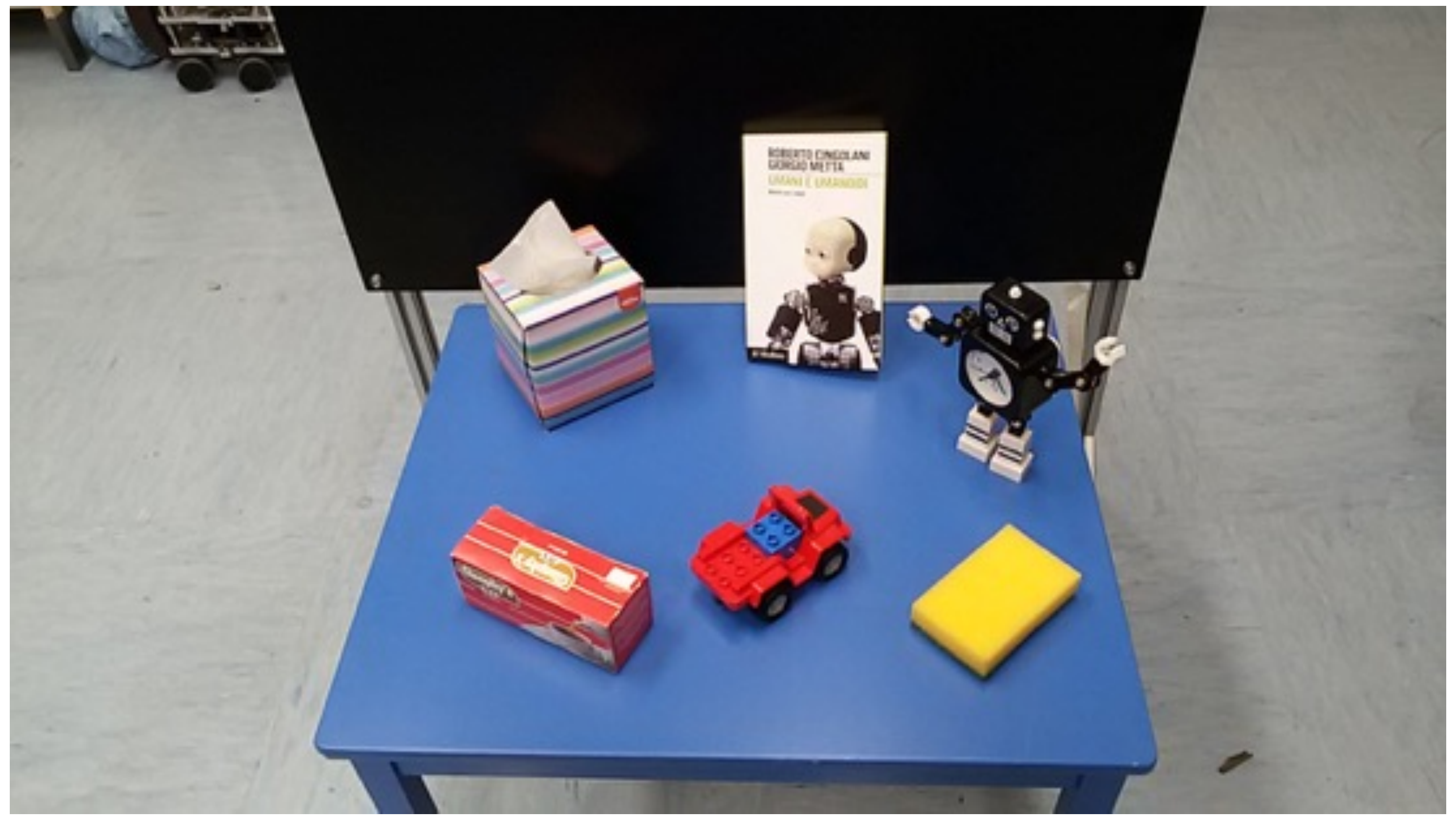


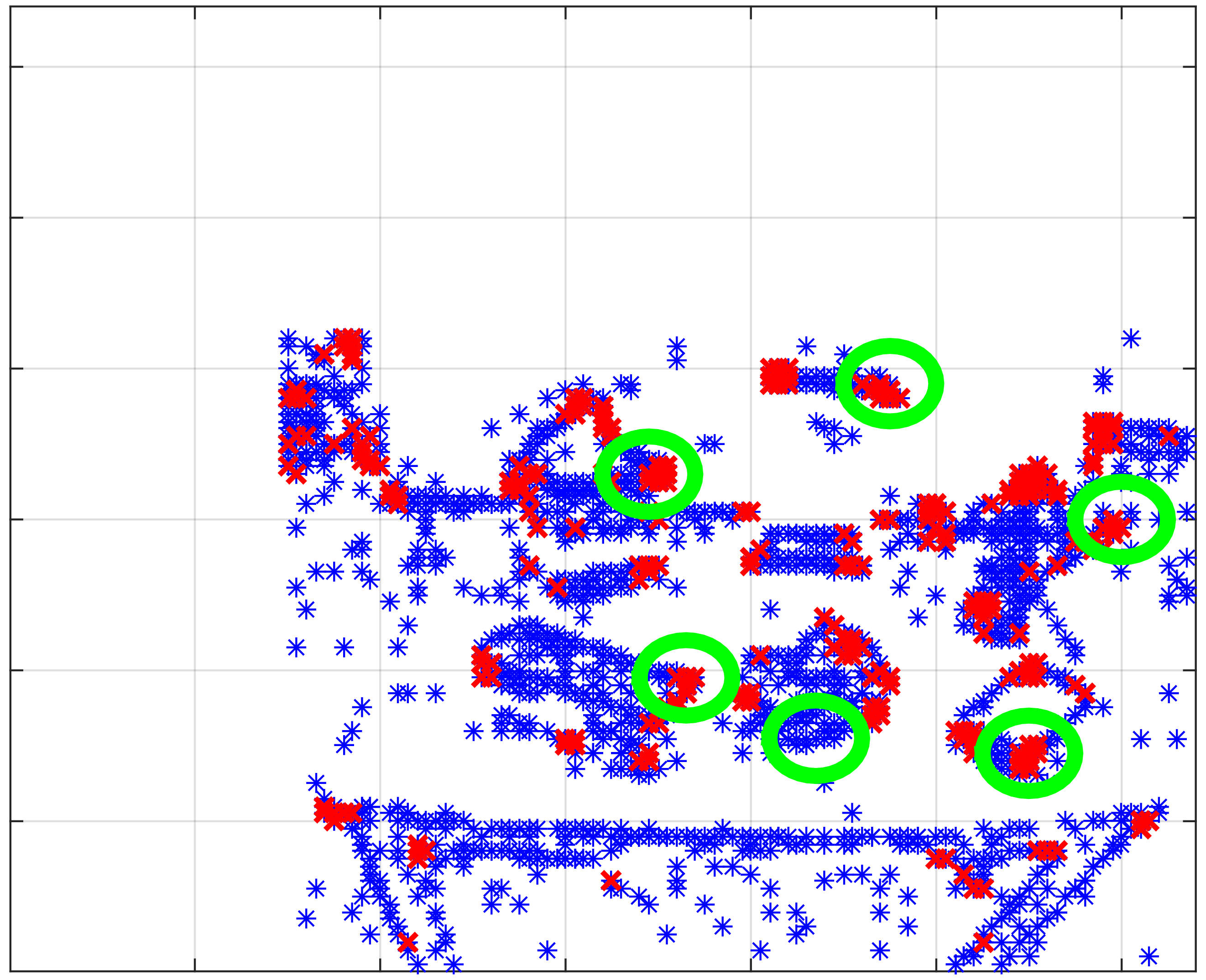


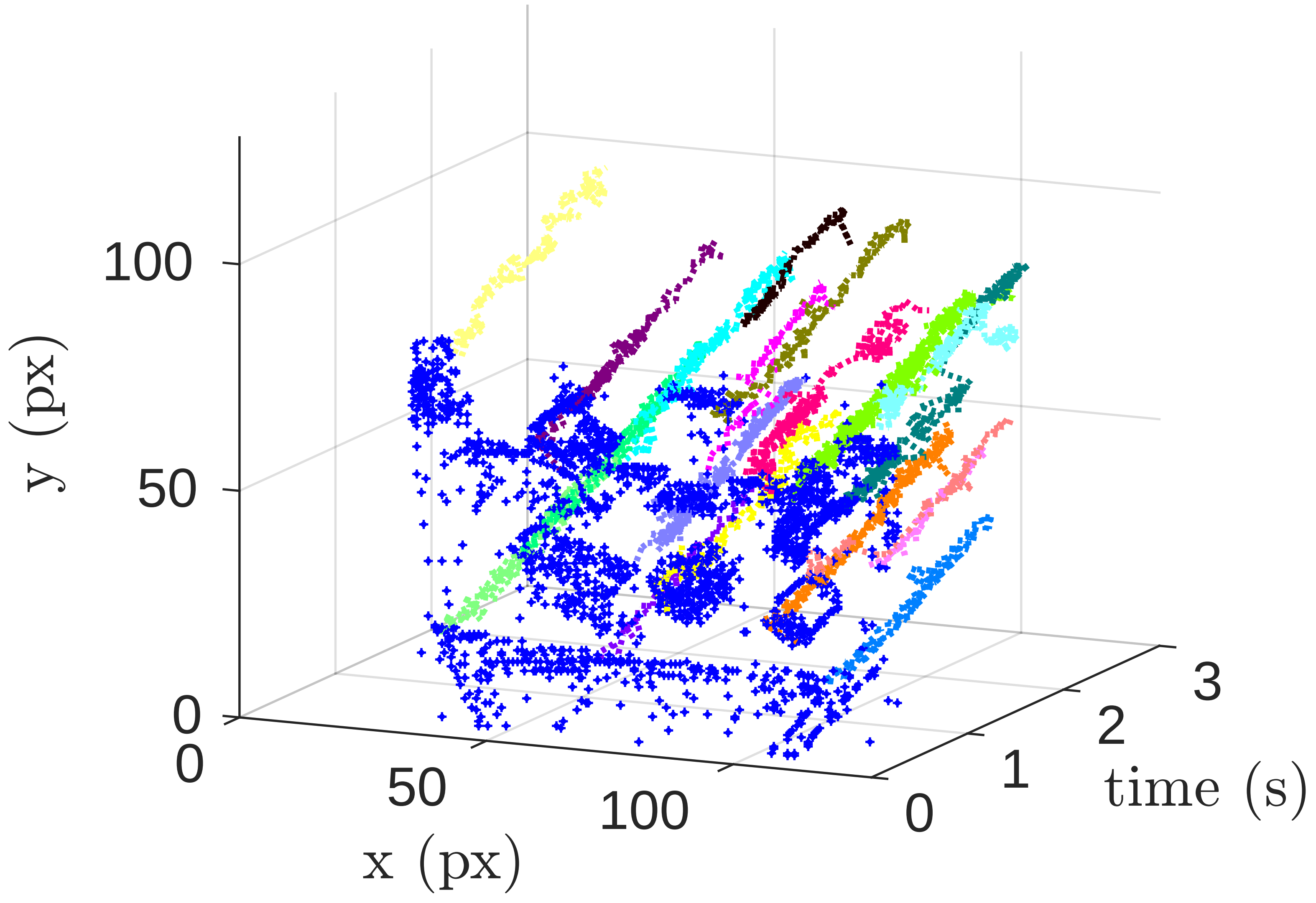


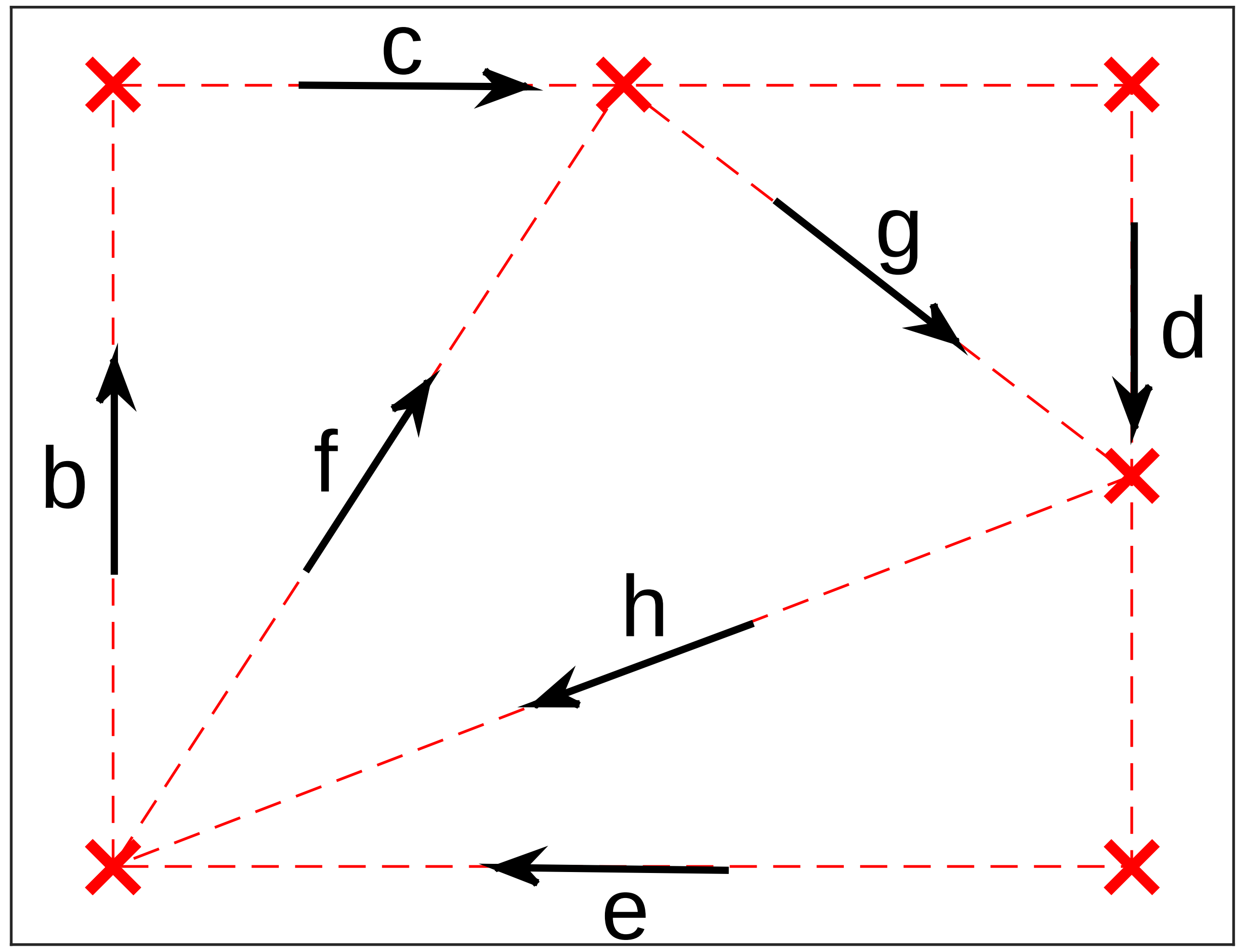


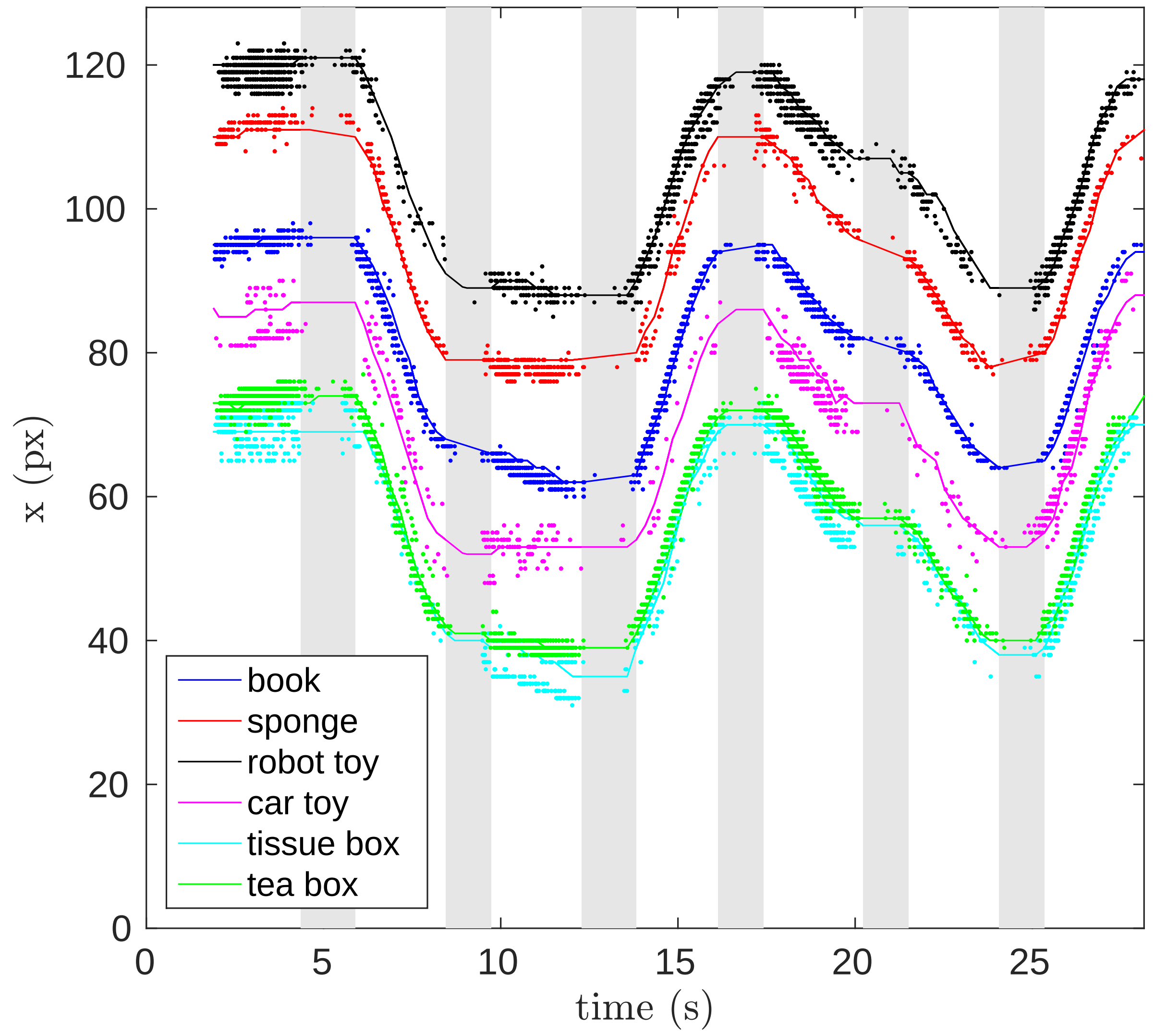
ED Vision — Corner Detection

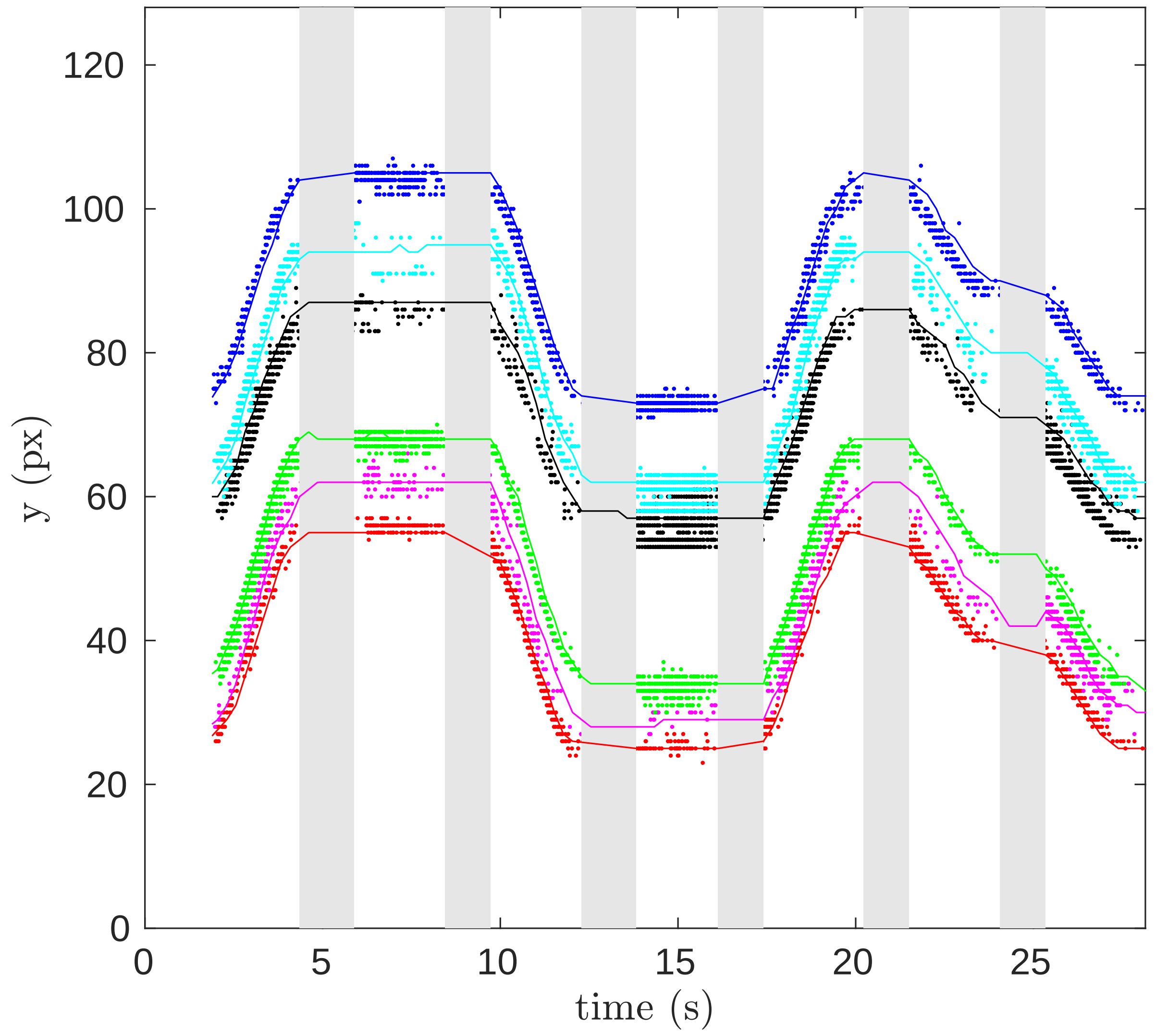


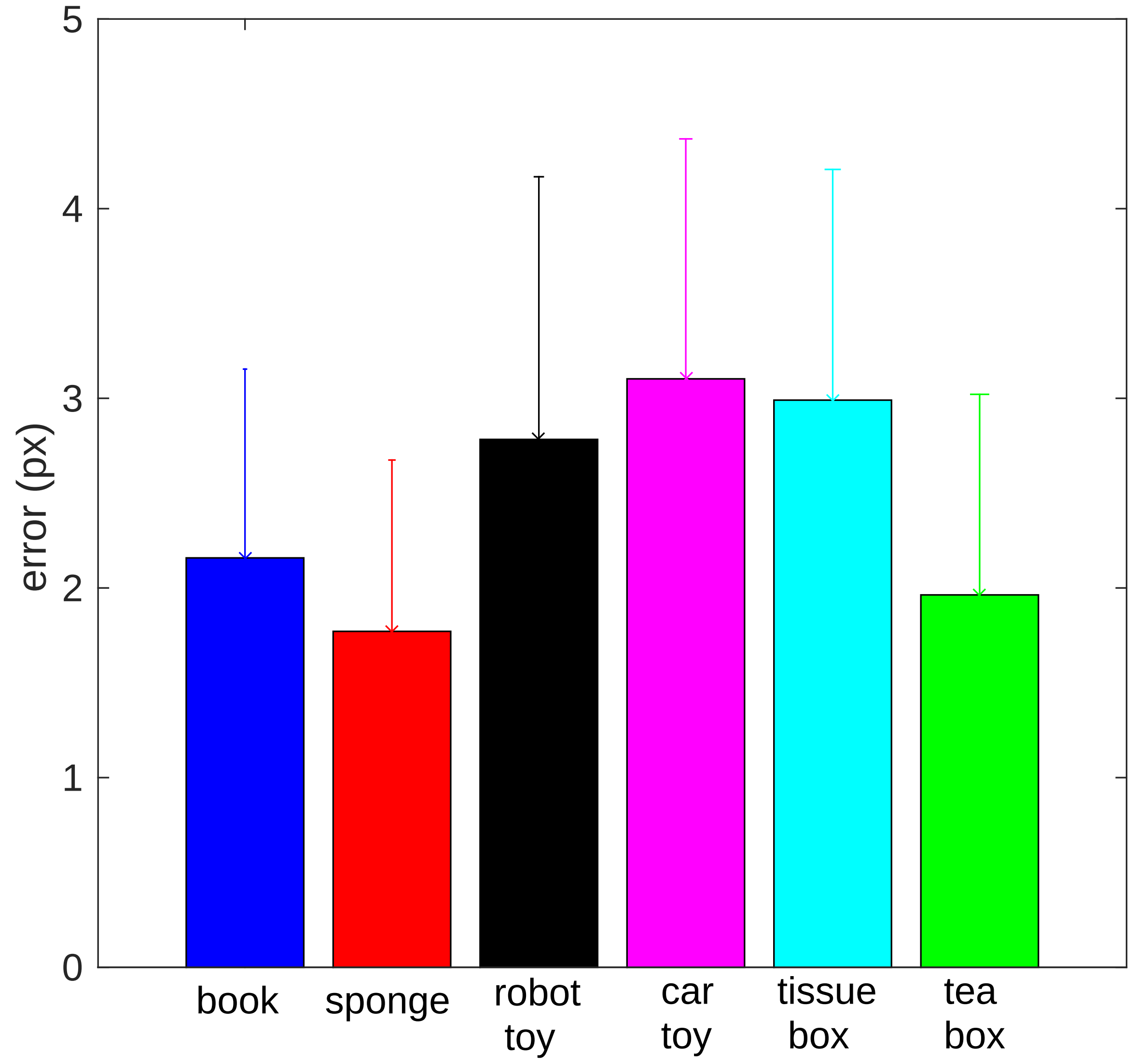




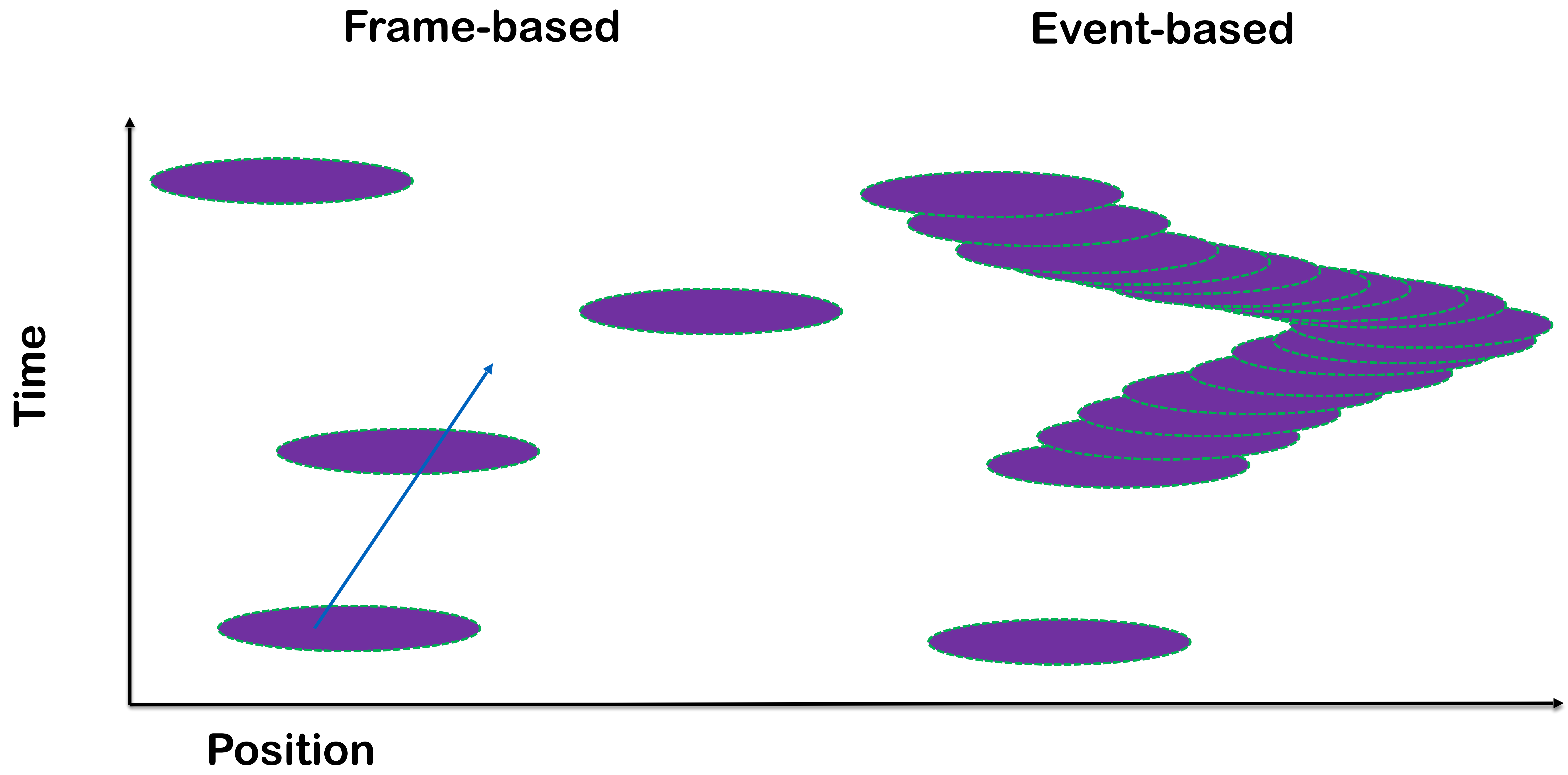




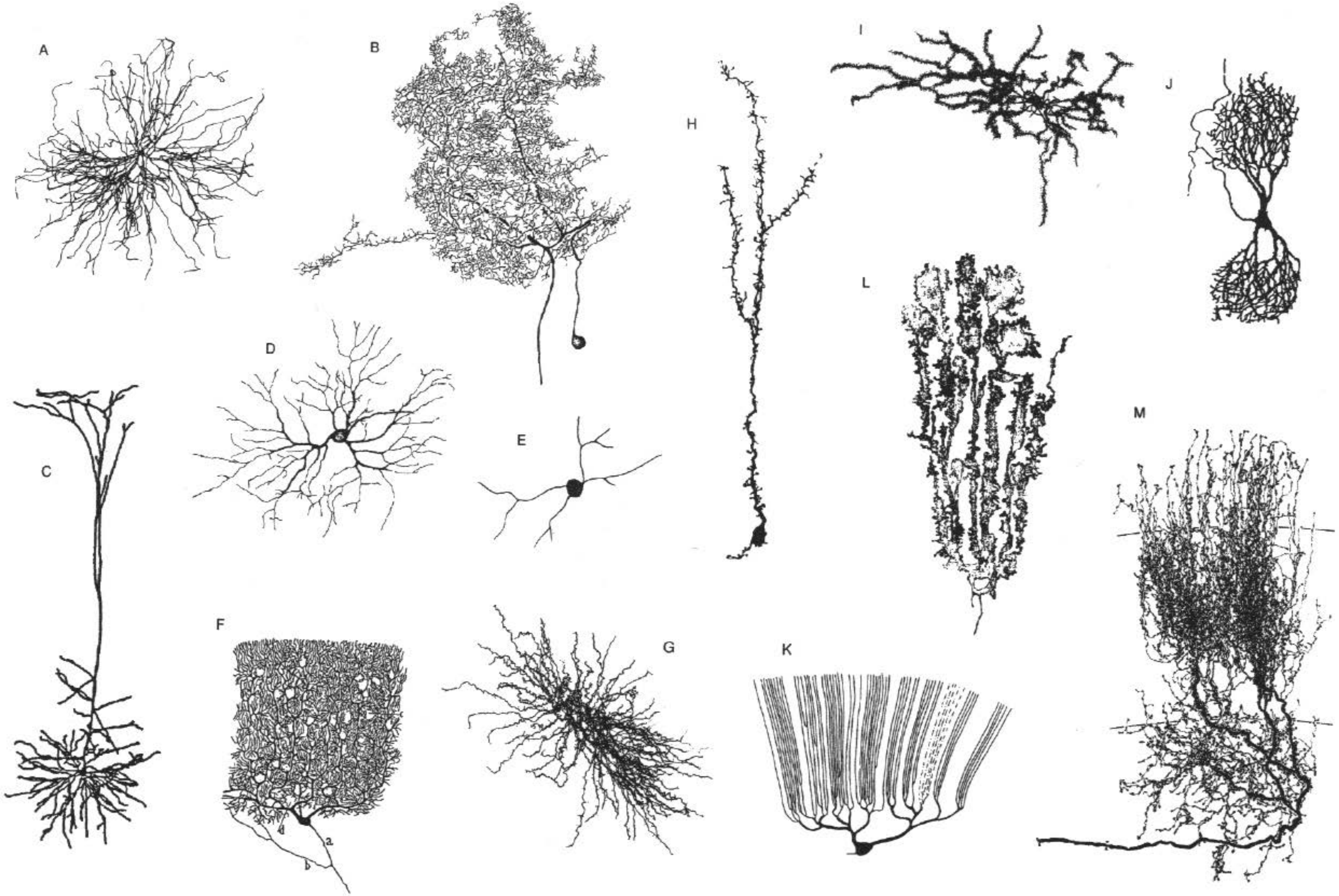




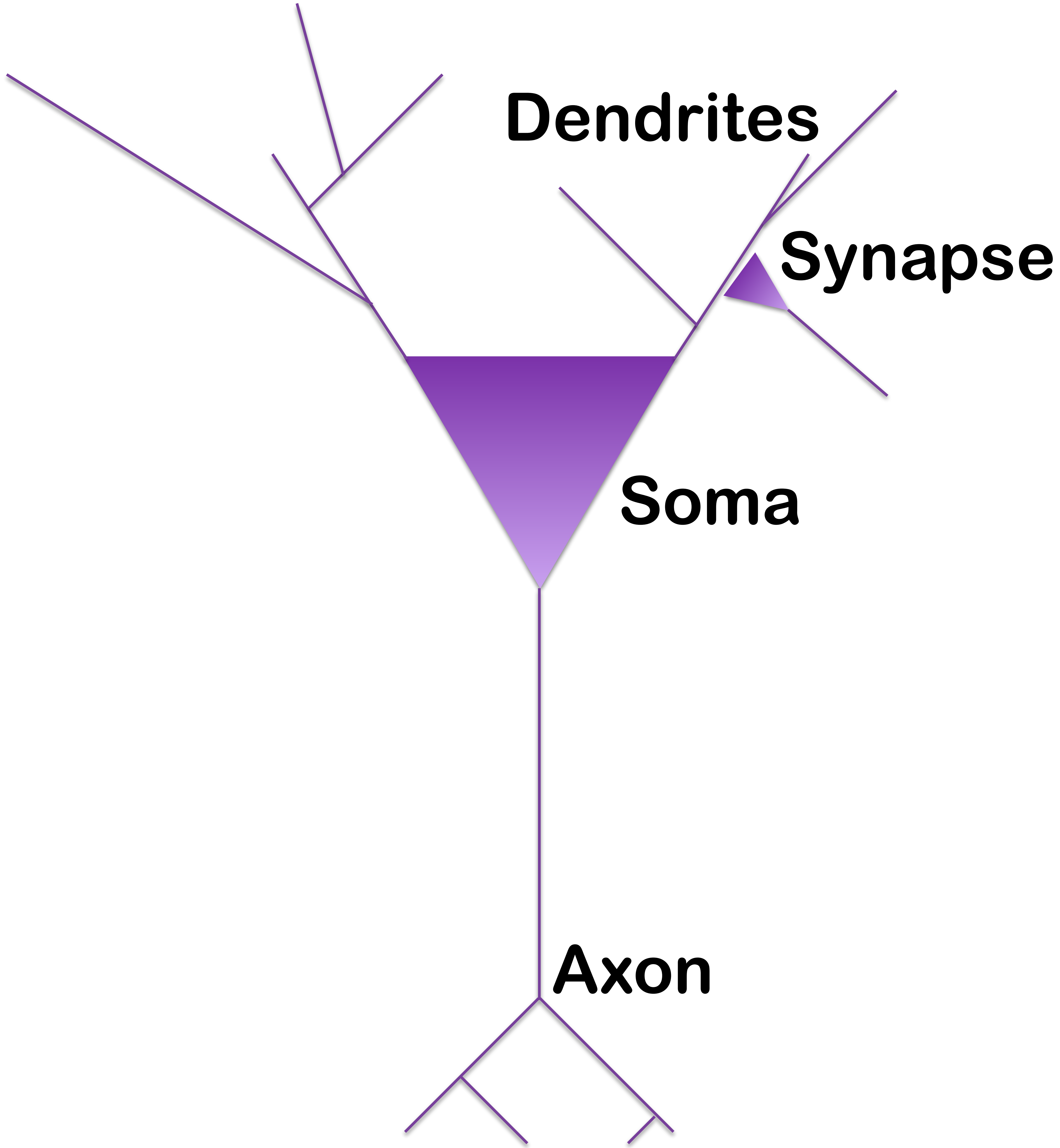
ED Tracking

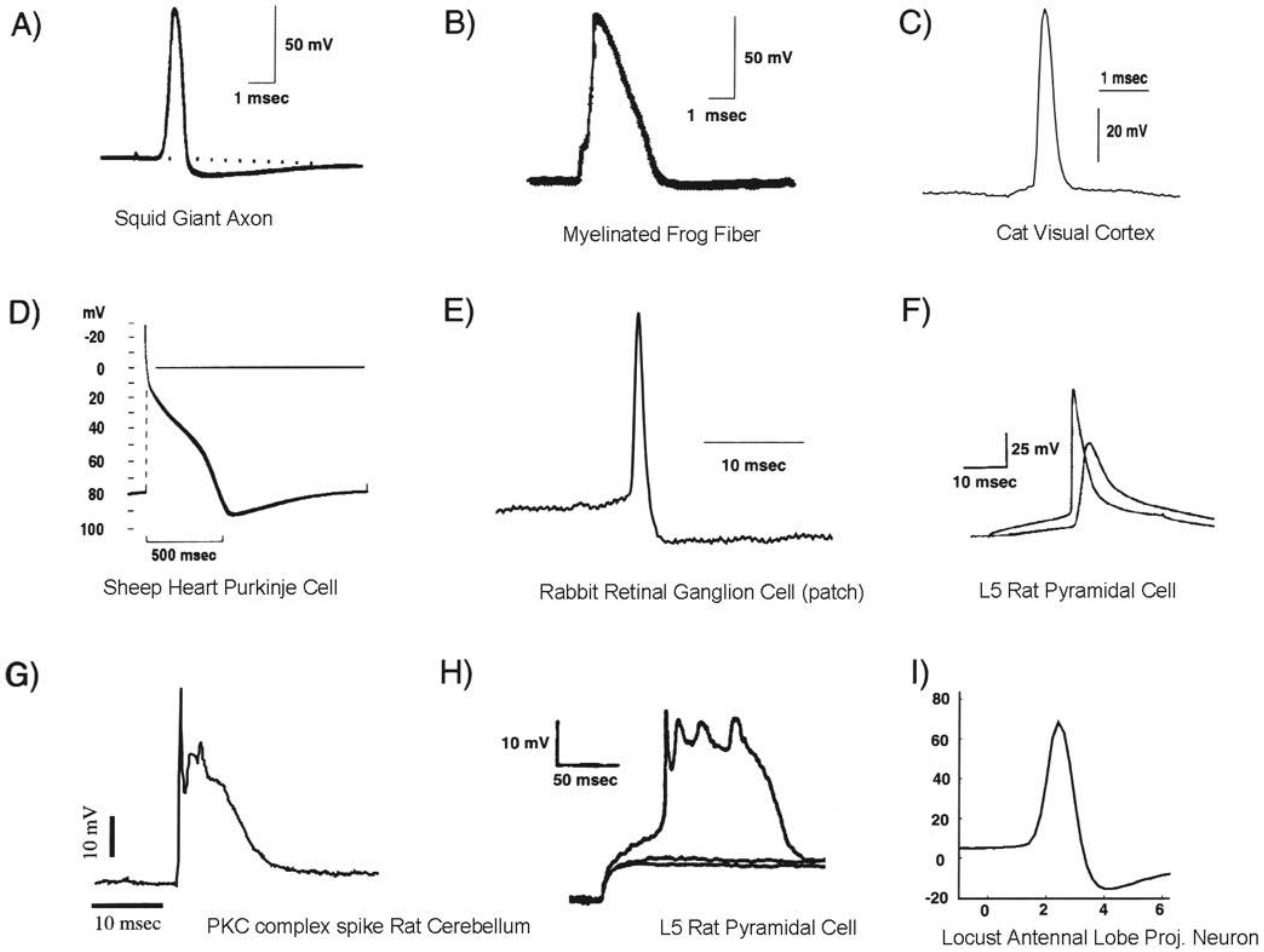


Neurons

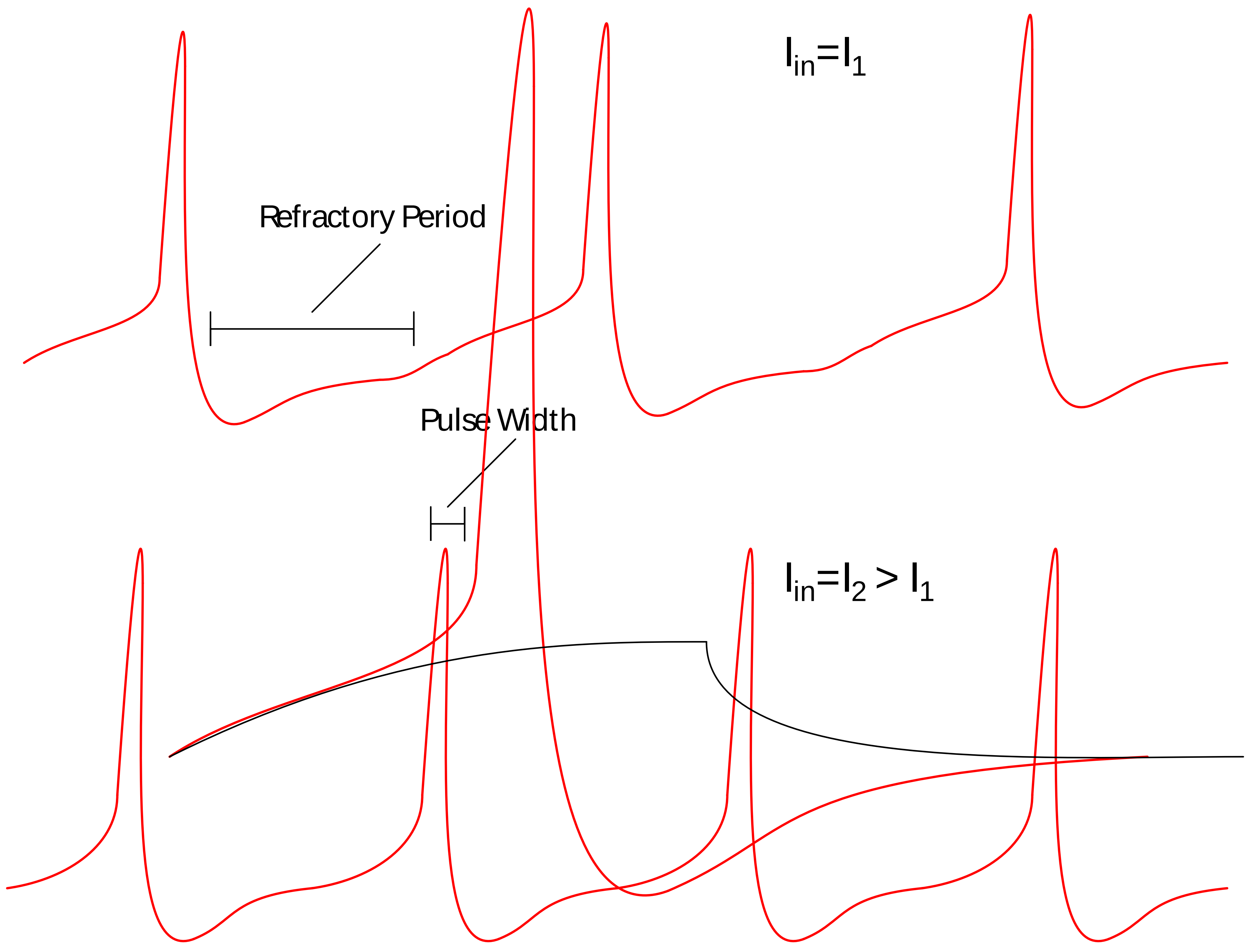


Neurons


Dendrites
Synapse
Soma
Axon


A)
50 mV
1 msec
Squid Giant Axon
B)
50 mV
1 msec
Myelinated Frog Fiber
C)
1 msec
20 mV
Cat Visual Cortex
D)
mV
-20
0
20
40
60
80
100
500 msec
Sheep Heart Purkinje Cell
E)
10 msec
Rabbit Retinal Ganglion Cell (patch)
F)
25 mV
10 msec
L5 Rat Pyramidal Cell
G)
10 mV
10 msec
PKC complex spike Rat Cerebellum
H)
10 mV
50 msec
L5 Rat Pyramidal Cell
I)
80
60
40
20
0
-20
0
2
4
6
Locust Antennal Lobe Proj. Neuron

Neurons — Leaky Integrate & Fire



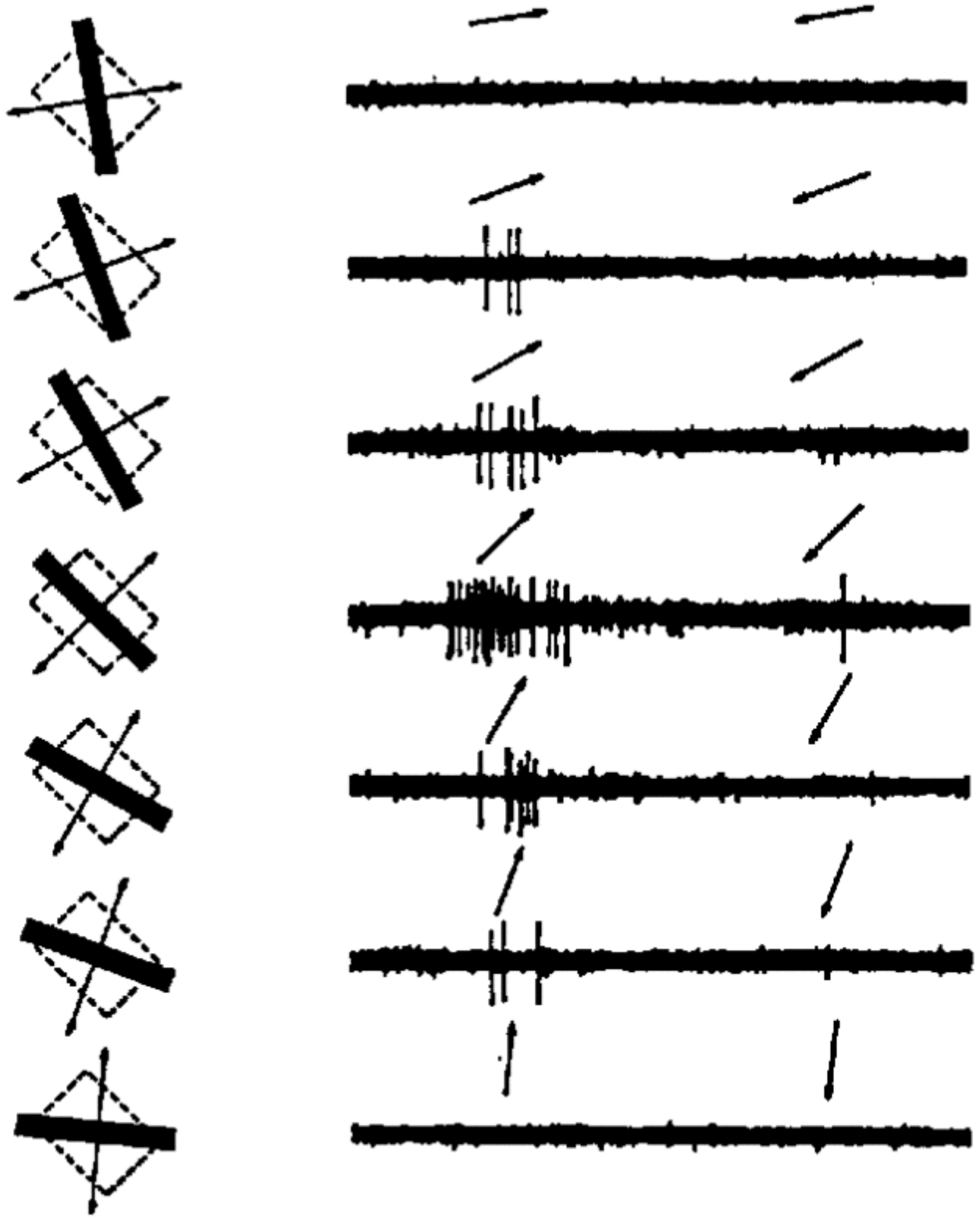
A simplified model is the Leaky Integrate & Fire neuron:

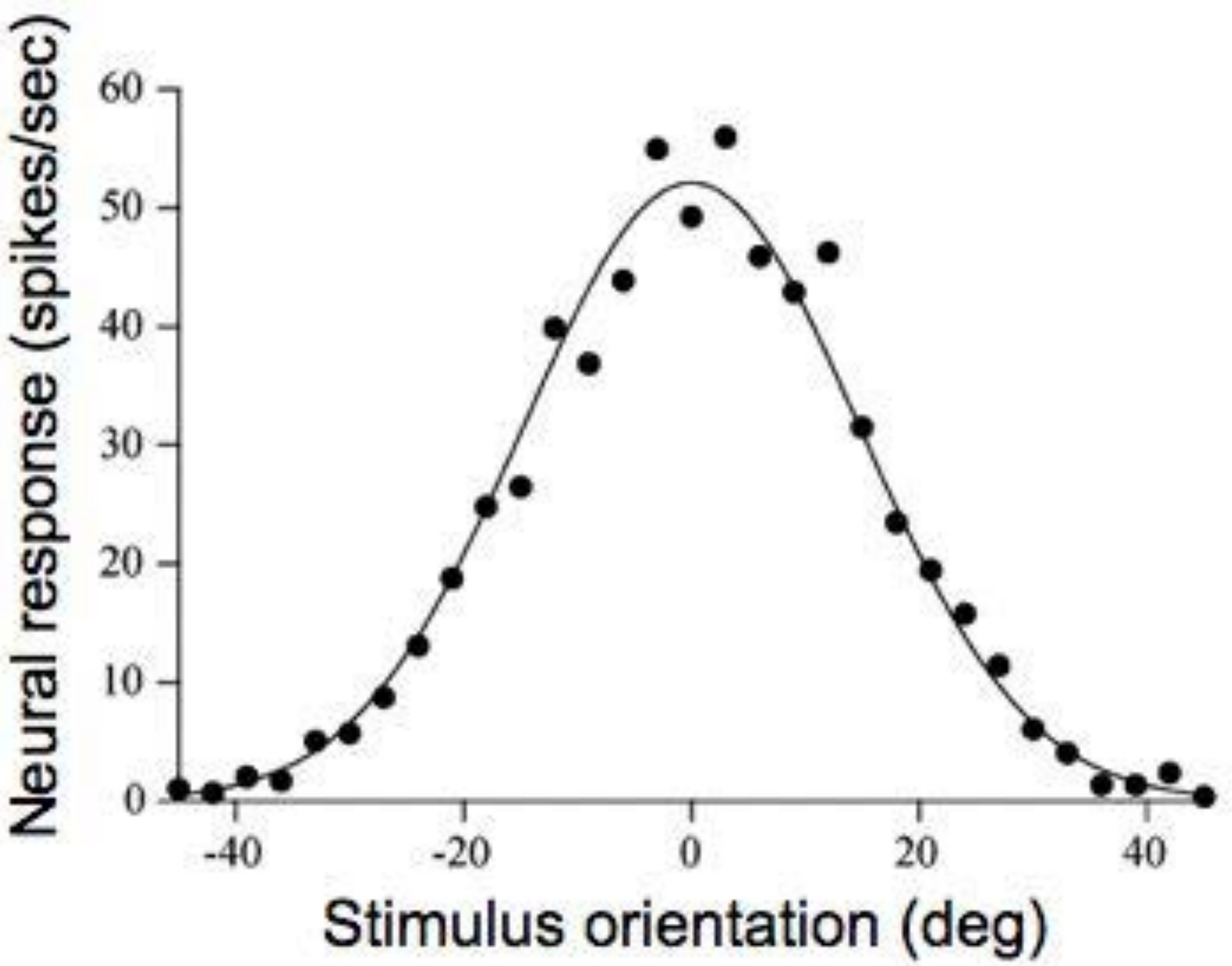
$$\tau_m \frac{dv}{dt} = -v(t) + Ri(t)$$

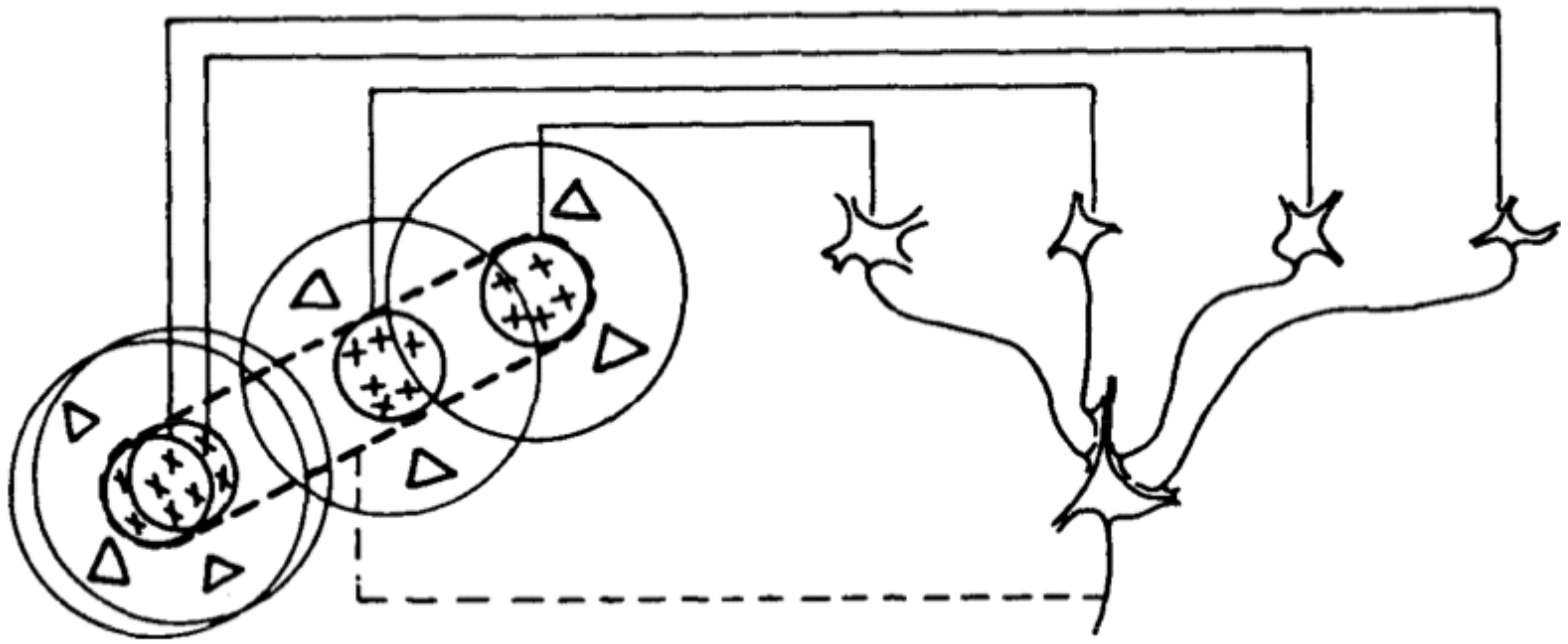
$$\text{if } v > v_{thr} \text{ fire and reset!}$$

Neurons — Receptive Field

Hubel&Wiesel 1968





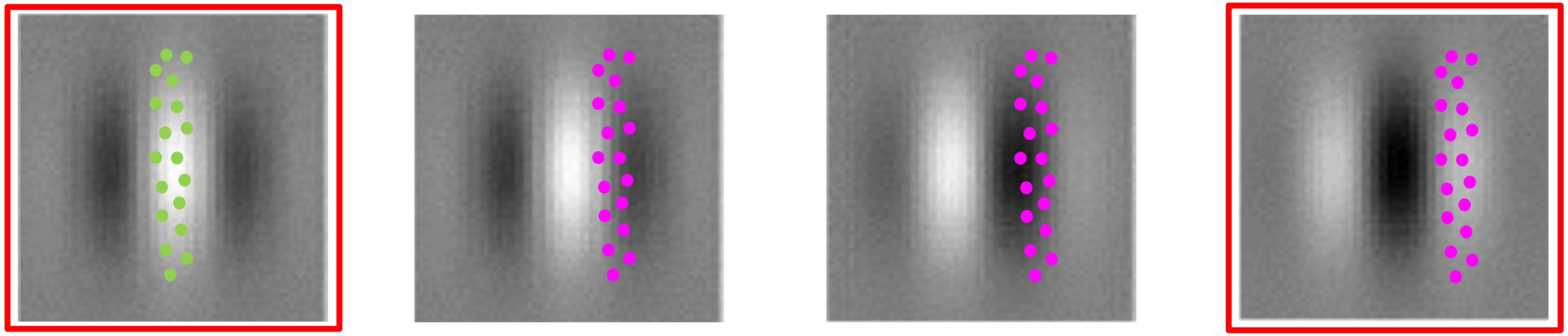


Neurons — Receptive Field

2D Gabor Filter

ED Vision — Disparity and Vergence

LEFT RIGHT



$\psi_L = 0$ ψ_{R1} ψ_{R2} ψ_{R3}

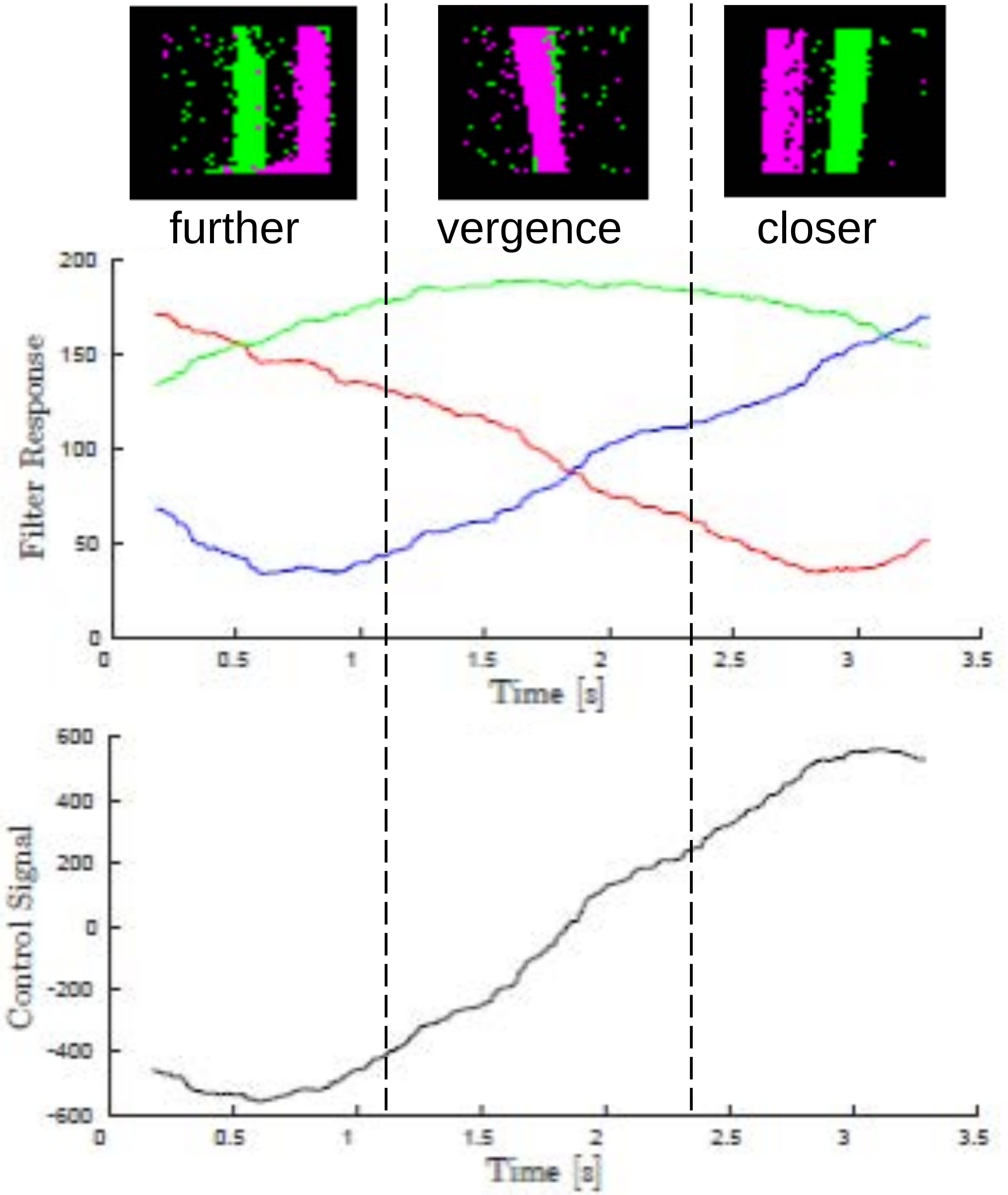
Disparity: Phase-Shift Model

$$g(x, y, \theta, \psi) = e^{-\frac{x^2+y^2}{2\sigma^2}} e^{j(2\pi f_s x_\theta + \psi)}$$

Vergence Control: Proportional Velocity

$$\dot{v} = k_p \frac{\sum_{i=1}^{N_\theta} \sum_{j=1}^{N_\psi} w_{ij} r_{ij}}{N_\theta N_\psi \sum_{i=1}^{N_\theta} \sum_{j=1}^{N_\psi} r_{ij}}$$

ED Vision — Vergence Control



filter tuned to negative disparity

filter tuned to zero disparity

filter tuned to positive disparity

Proportional velocity control, with control signal:

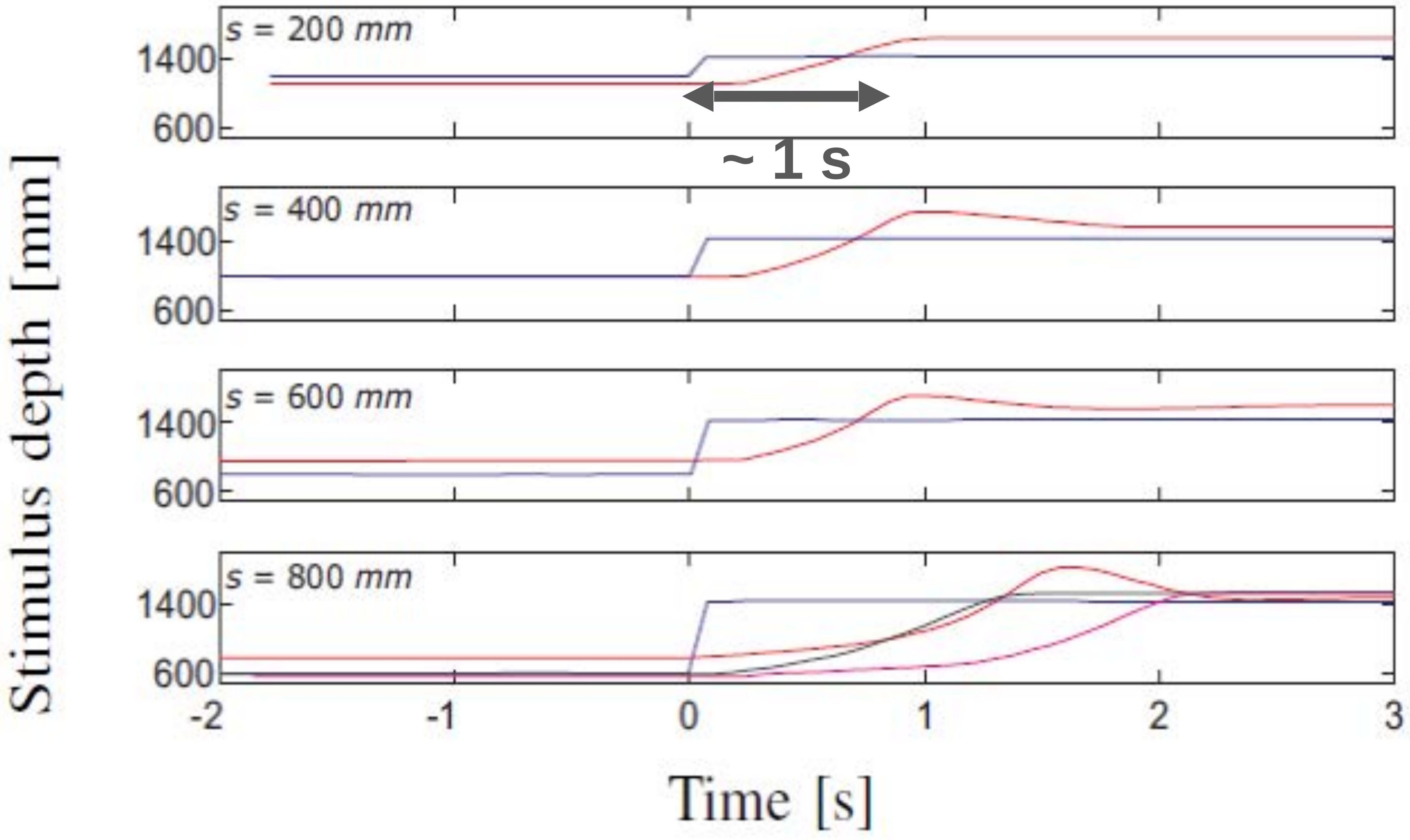
$$\dot{v} = k_p \frac{\sum_{i=1}^{N_\theta} \sum_{j=1}^{N_\psi} w_{ij} r_{ij}}{N_\theta N_\psi \sum_{i=1}^{N_\theta} \sum_{j=1}^{N_\psi} r_{ij}}$$

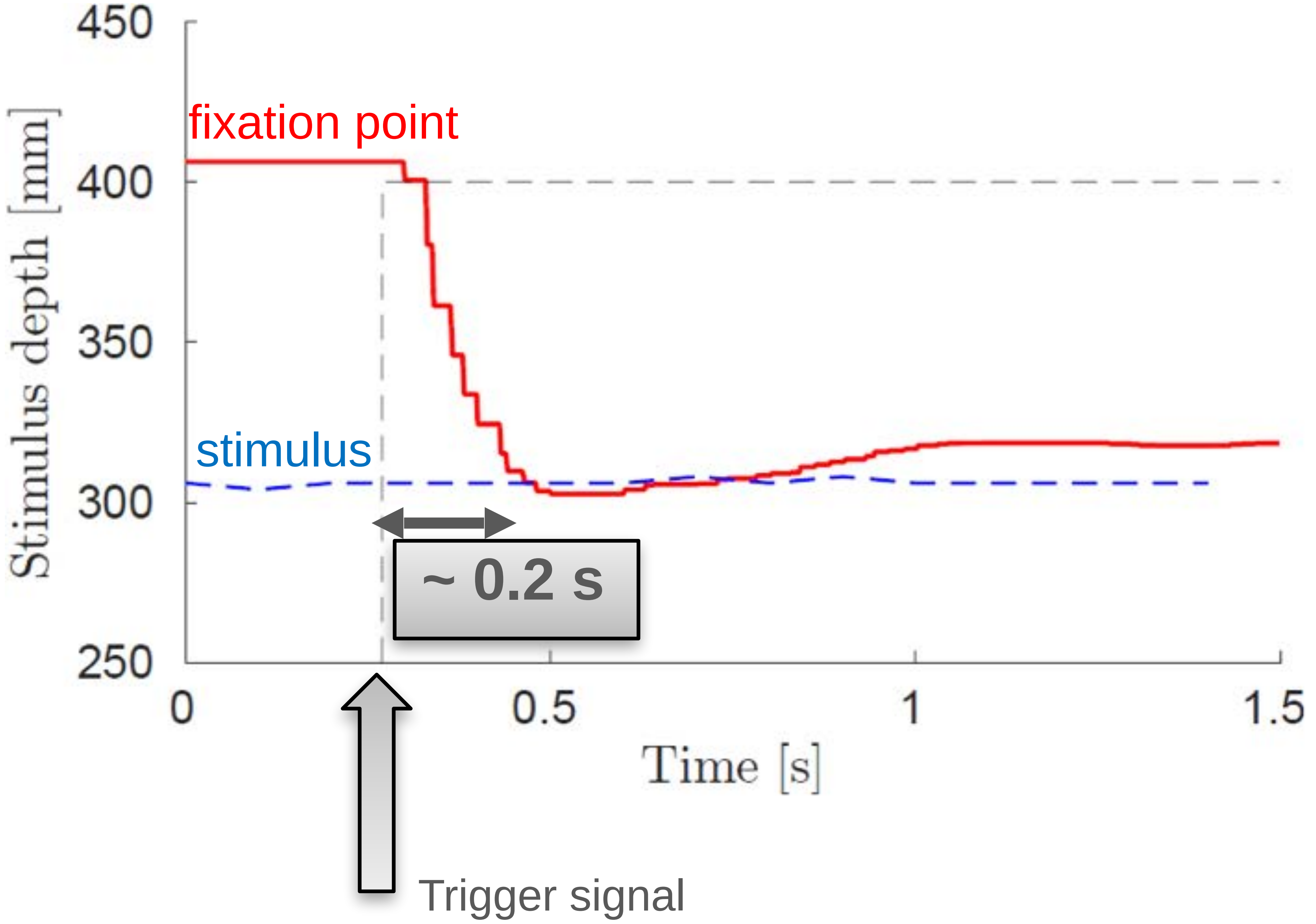
with:

- w_{ij} = weight of the *ij –th* filter
- r_{ij} = response of the *ij –th* filter
- k_p = proportional control gain

ED Vision — Vergence Control

FRAME-BASED Vs. EVENT-BASED





From: A. Gibaldi, A. Canessa, M. Chessa, S.P. Sabatini and F. Solari, **A neuromorphic control module for real-time vergence eye movements on the iCub robot head**, *IEEE-RAS International Conference on Humanoid Robots*

ED Vision — Vergence Control

FRAME-BASED

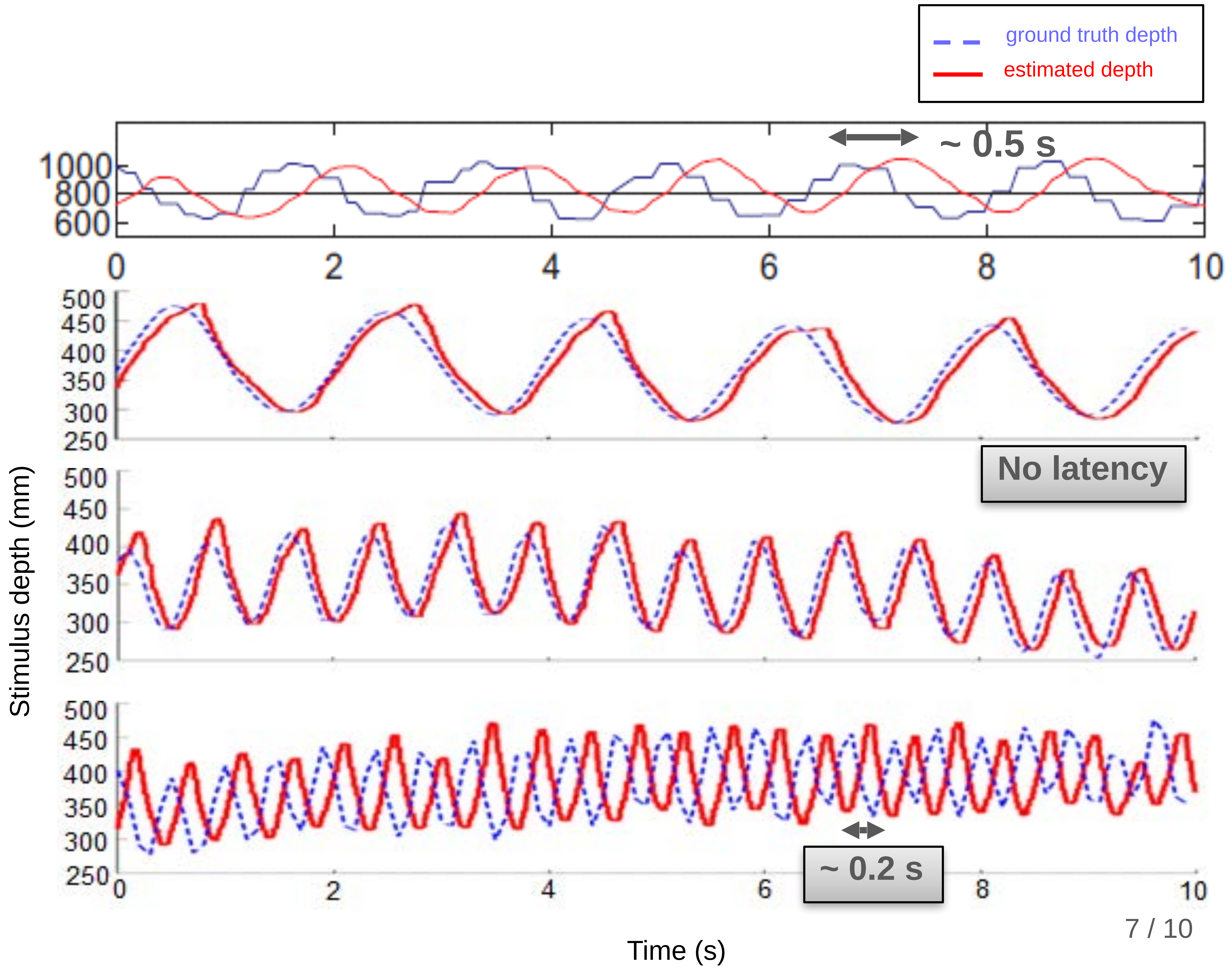
f = 0.5 Hz

EVENT-BASED

f = 0.5 Hz

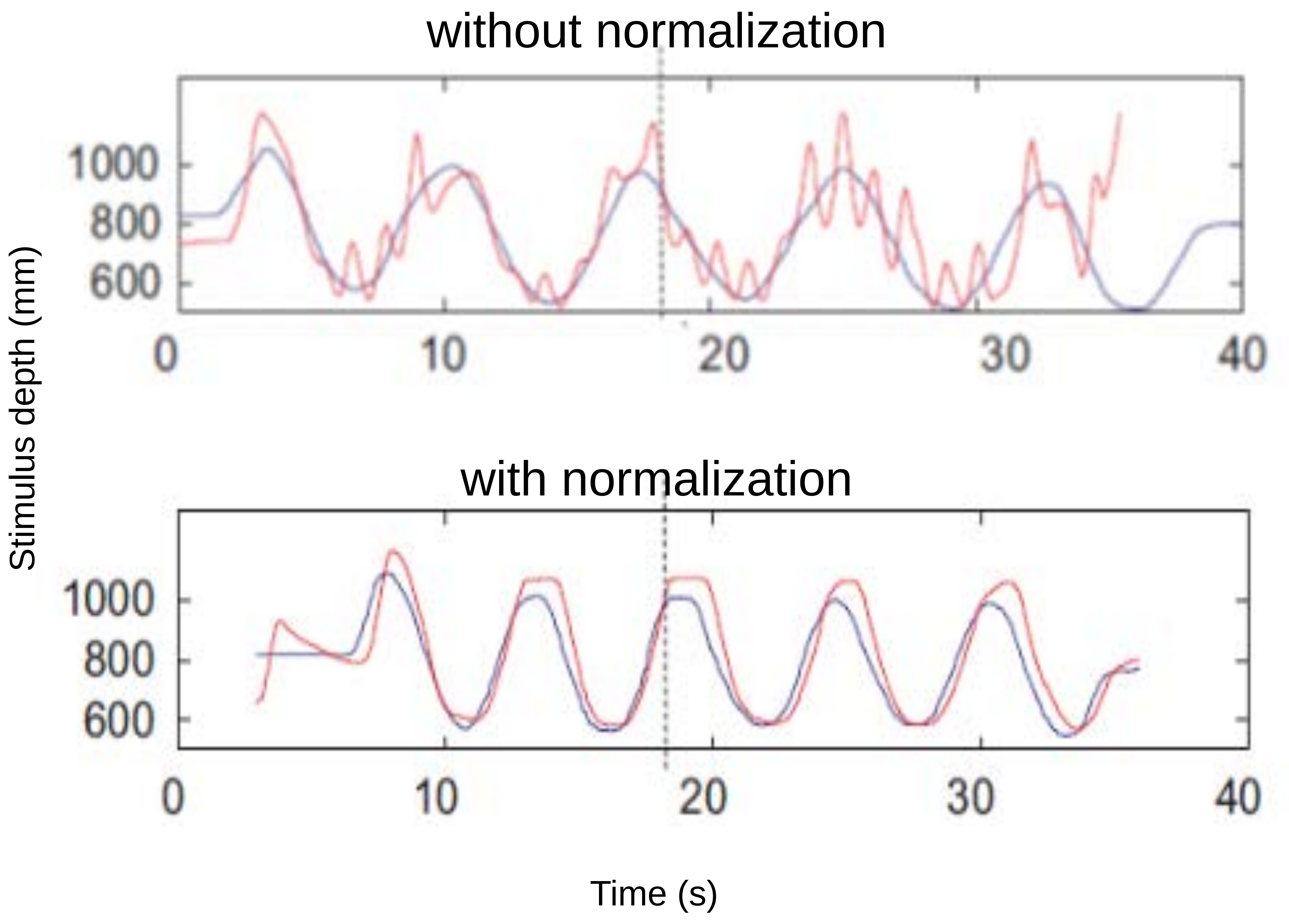
f = 1.25 Hz

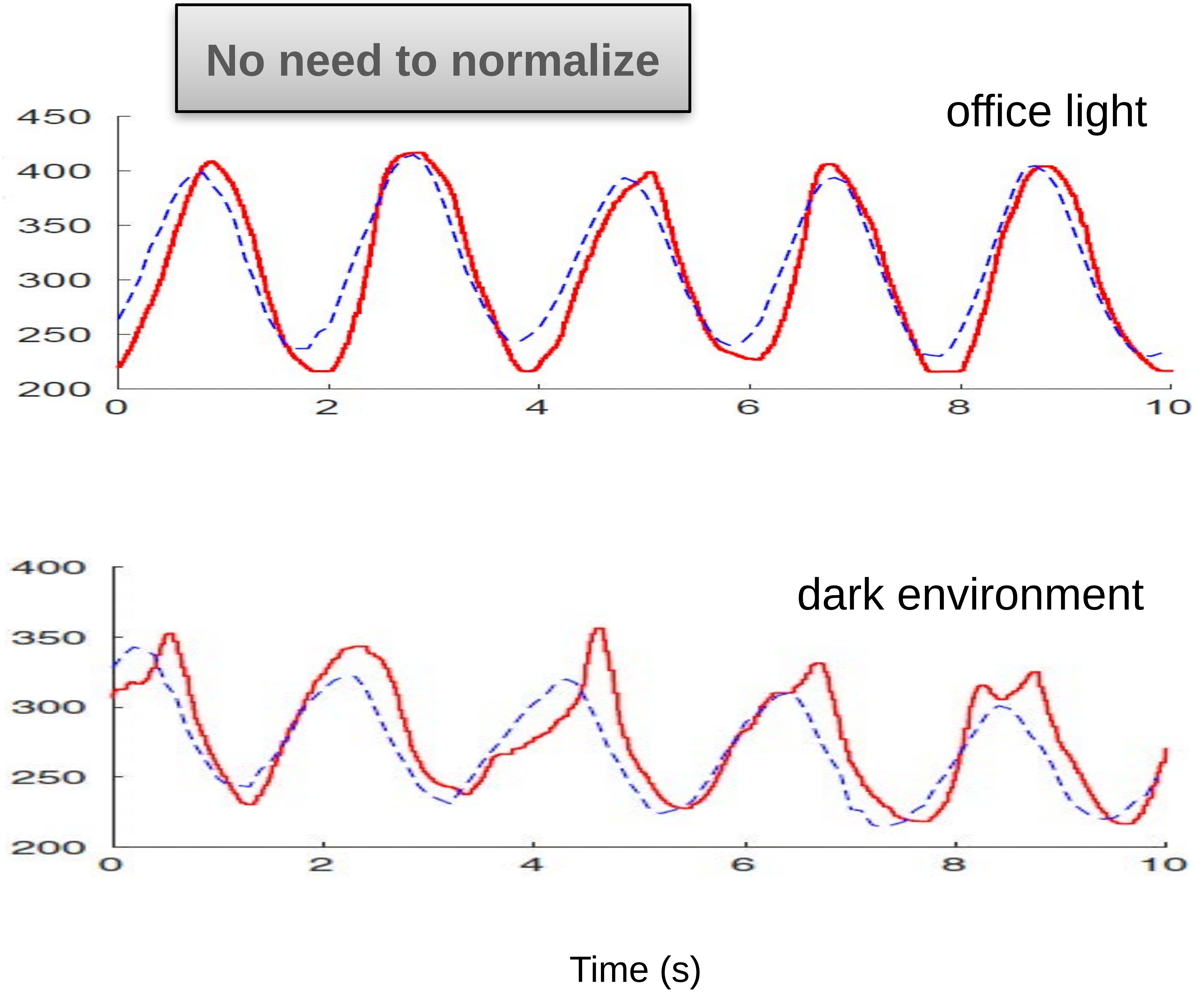
f = 2 Hz



ED Vision — Vergence Control

FRAME-BASED Vs. EVENT-BASED





Vasco et al. Humanoids 2016