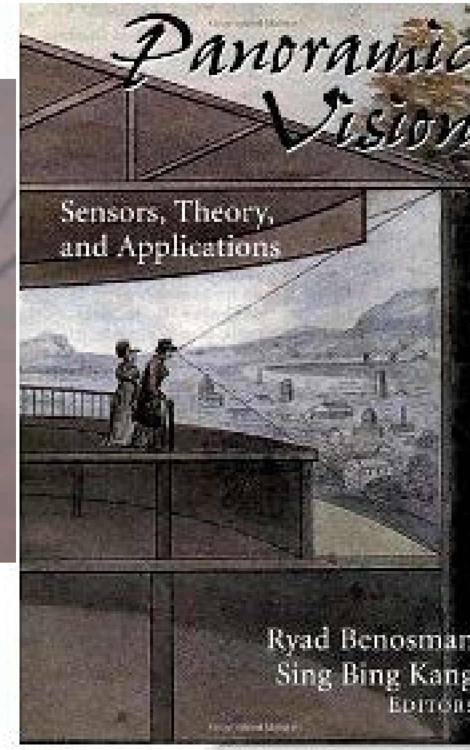
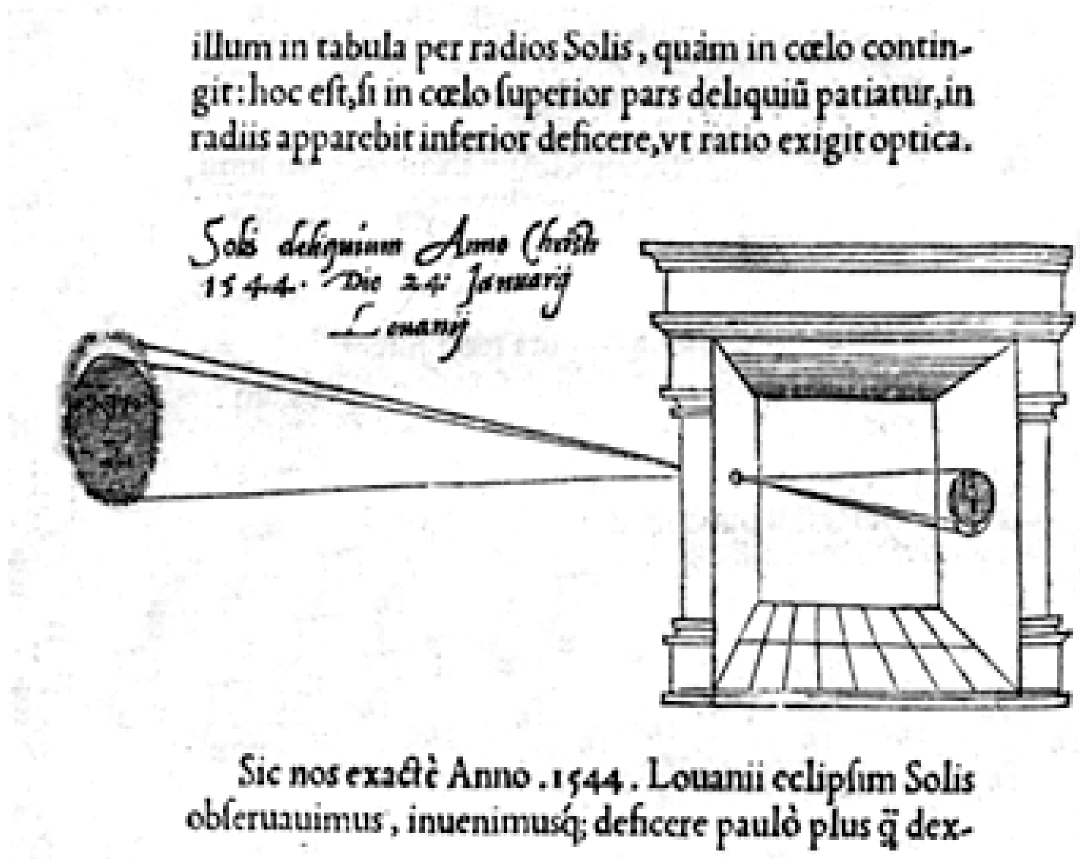
Neuromorphic General Purpose Computation Using Precise Timing

R.B. Benosman,
Eye and Ear Institute,
BST-3, Rm 2046, 3501 Fifth Avenue
Pittsburgh, PA 15213
benosman@pitt.edu

Omnidirectional Vision



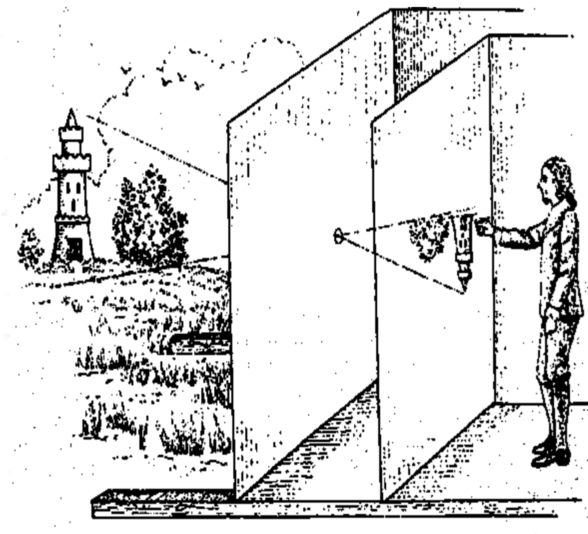
Origins of Imaging



illum in tabula per radios Solis, quàm in cœlo contin-
git: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in
radiis apparebit inferior deficere, vt ratio exigit optica.

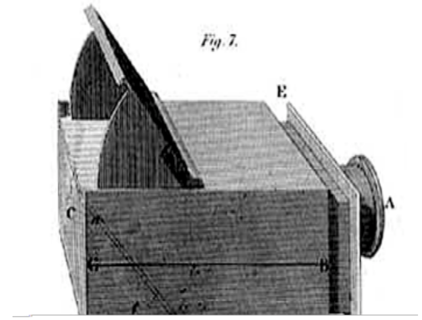
Solis deliquium Anno Christi
1544. Die 24 Januarij
Louanij

Sic nos exactè Anno .1544. Louanii eclipſim Solis
obſeruauimus, inuenimusq; deficere paulò plus q̃ dex-



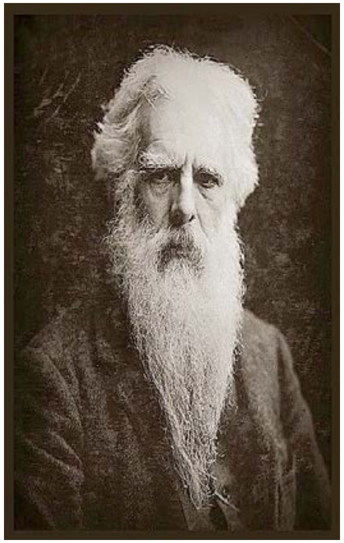
- Invention of the camera obscura in 1544 (L. Da Vinci)
- The mother of all cameras
- A more realistic and fast depiction of reality

Origins of Imaging

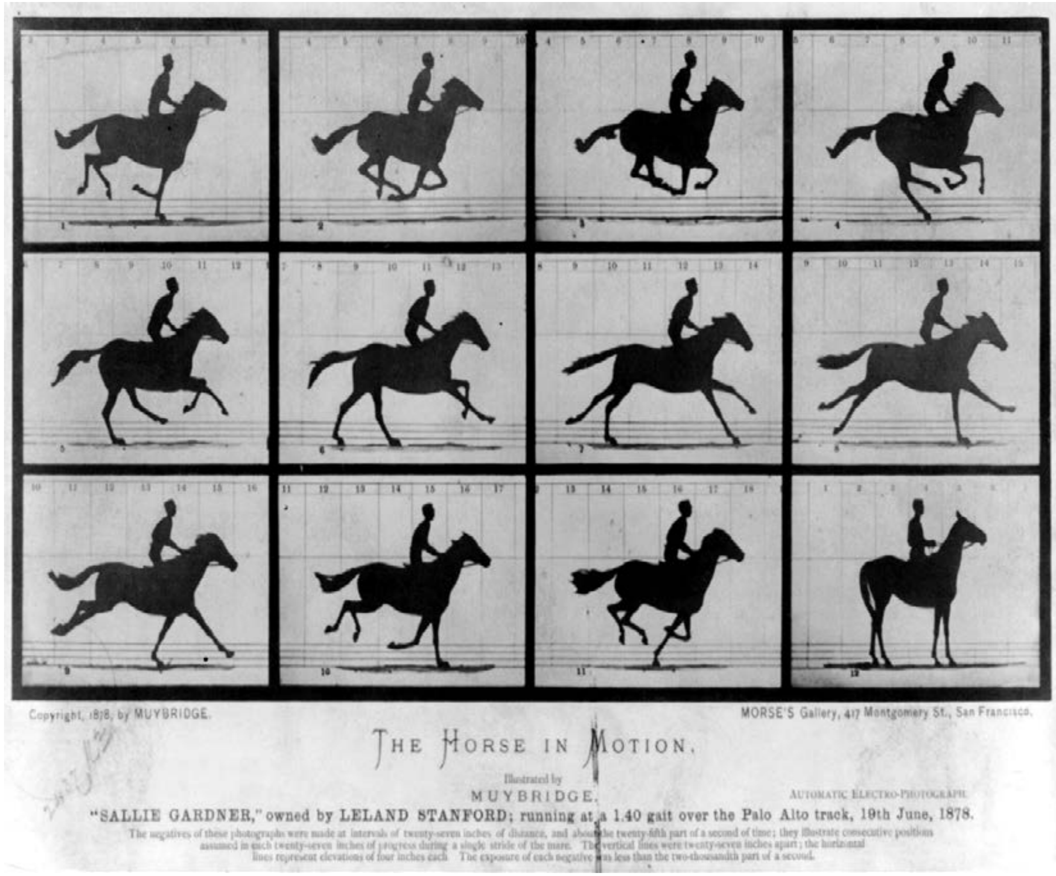


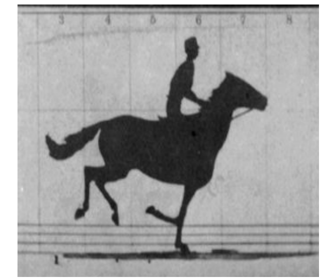
- Increasing painters profits: painting faster
- Evolution from portable models for travellers to current digital cameras
- Evolving from canvas, to paper, to glass, to celluloid, to pixels

Motion Picture: origins of video



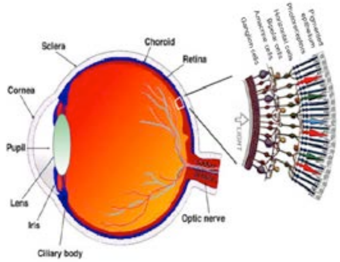
Eadweard Muybridge
(1830-1904)

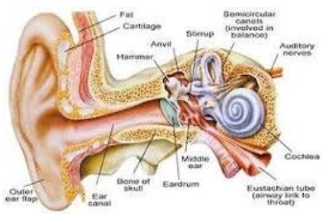


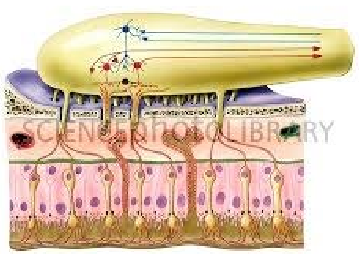


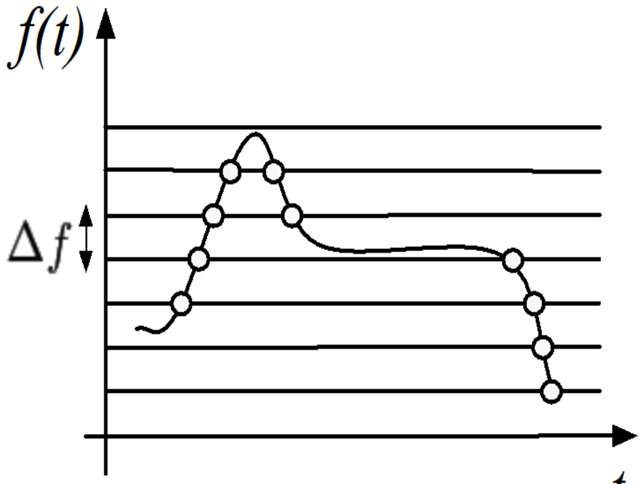
- Early work in motion-picture projection
- known for his pioneering work on animal locomotion in 1877 and 1878, which used multiple cameras to capture motion in stop-motion photographs

Neural acquisition

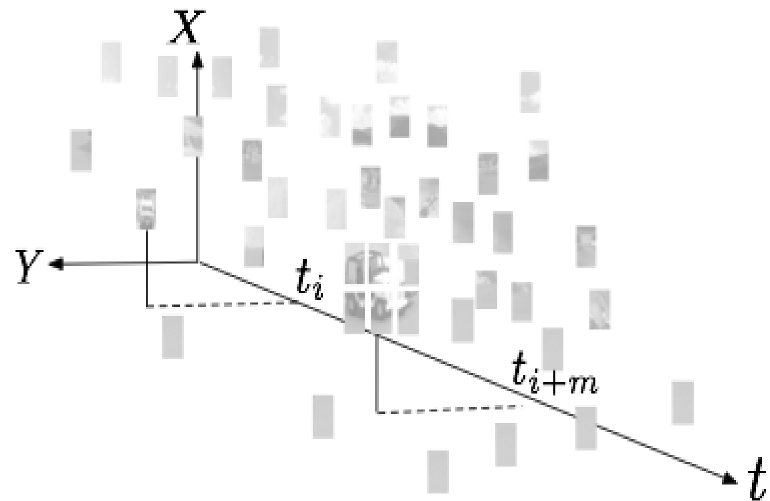






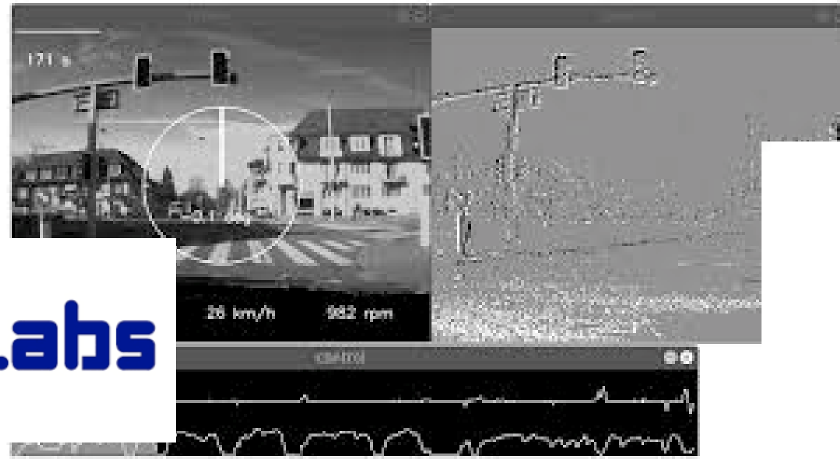


(b)

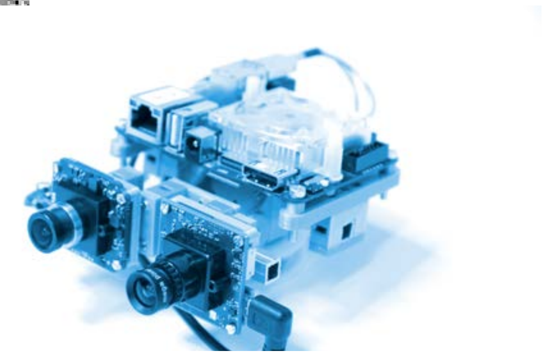


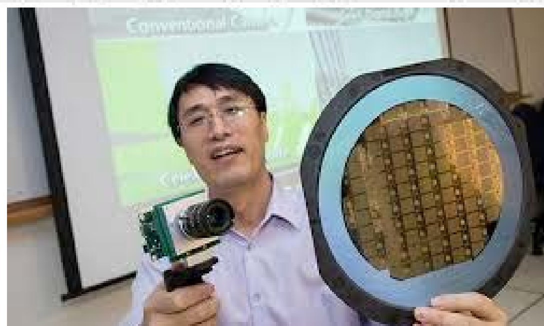
- Amplitude sampling
- Information is sent when it happens
- When nothing happens, nothing is sent or processed
- Sparse information coding
- Time is the most valuable information

Event-based Cameras



iniLabs





insightness

SAMSUNG

HILLHOUSE

SEE DIFFERENT, SEE FURTHER

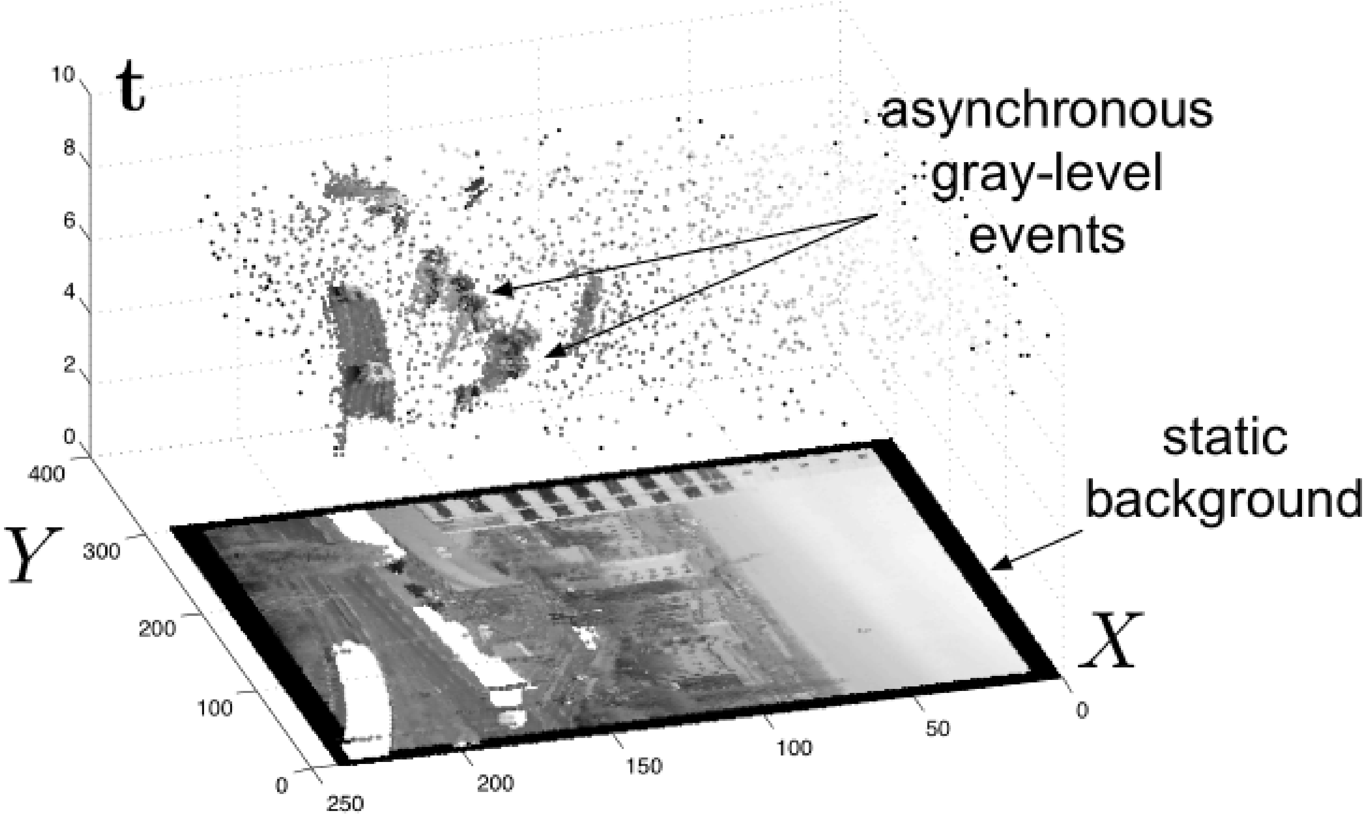
Insightness

PROPHESEE

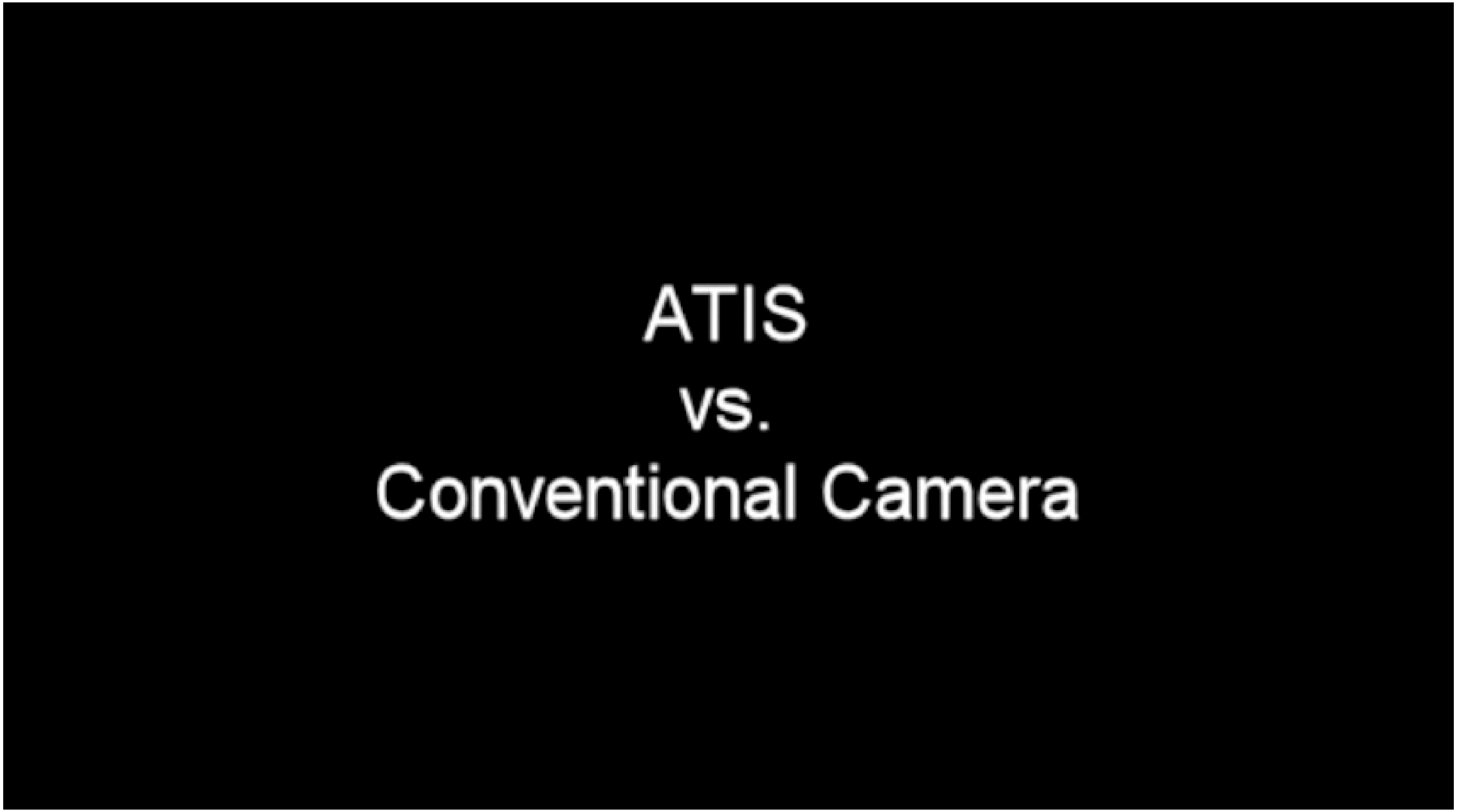
METAVISION FOR MACHINES

- Event-based cameras have become **a commodity**

Data Space of Events



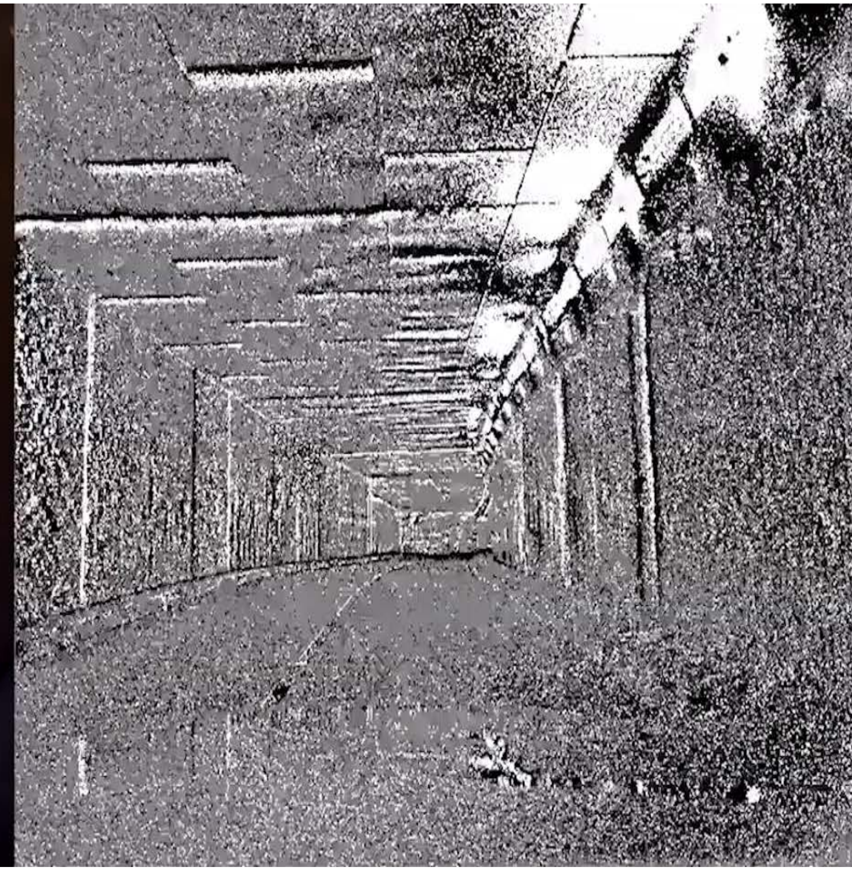
Why Event Based sensors?



Why Event Based sensors?

Conventional Camera

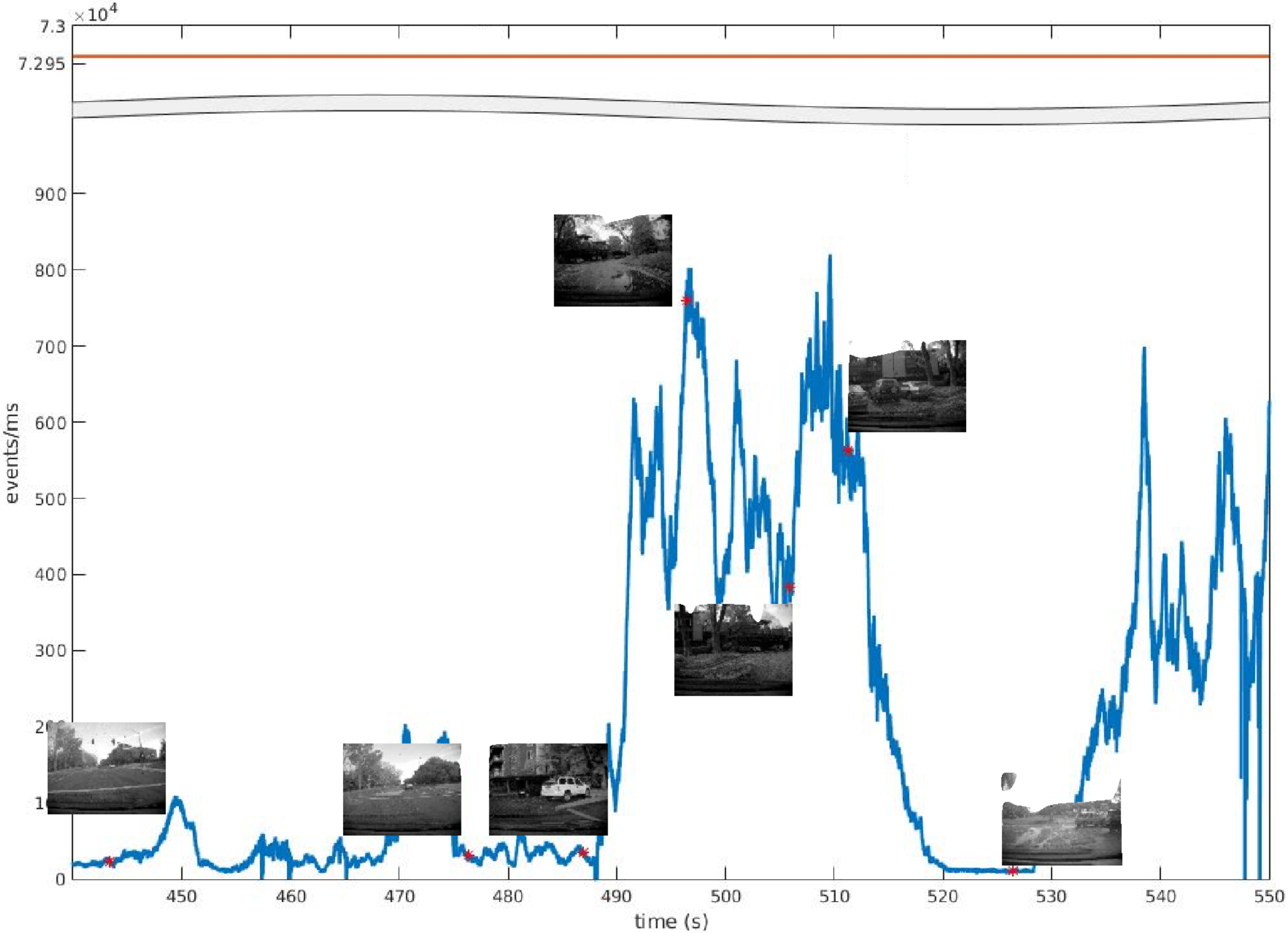
Event based Camera



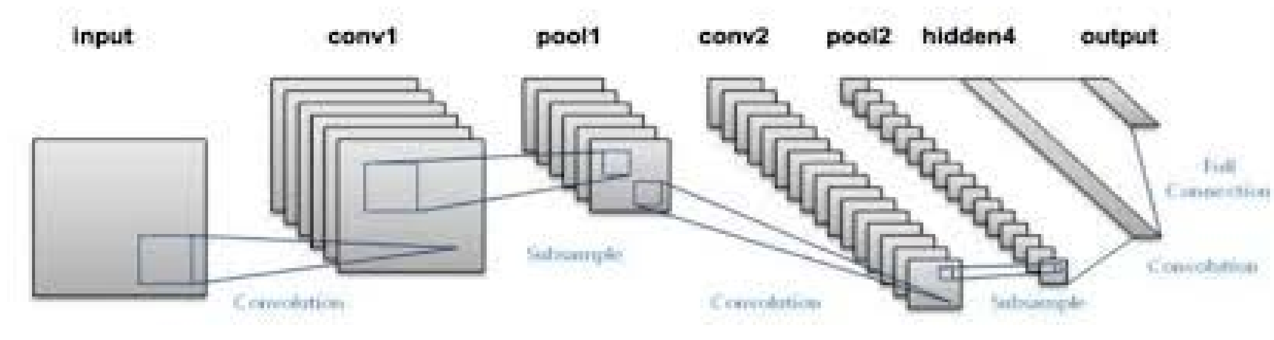
- Data driven: only moving edges produce data
- Temporal edges, precisely timed

Why Event Based sensors?

Why Event Based sensors?



What can not be Event based computation?

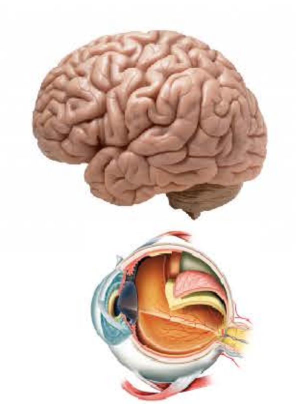


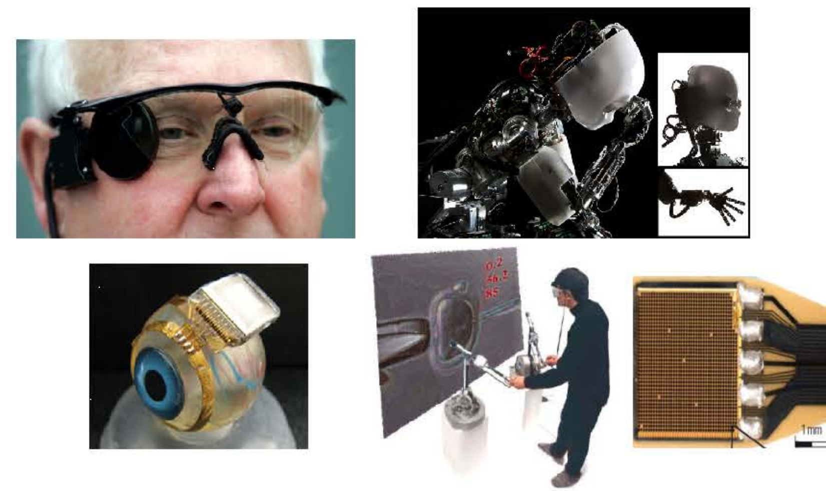
- Creating frames from events at the cost of heavy computation costs
- Using CNN and artificial binary frames

What can not be Event based computation?

- Creating frames from events at the cost of heavy computation costs
- Using CNN and artificial binary frames

Neuromorphic engineering





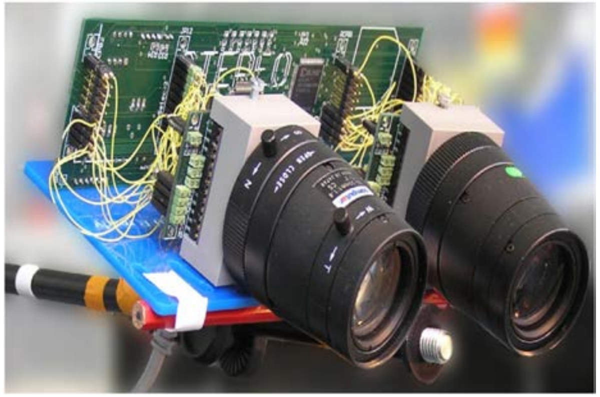
Physiology

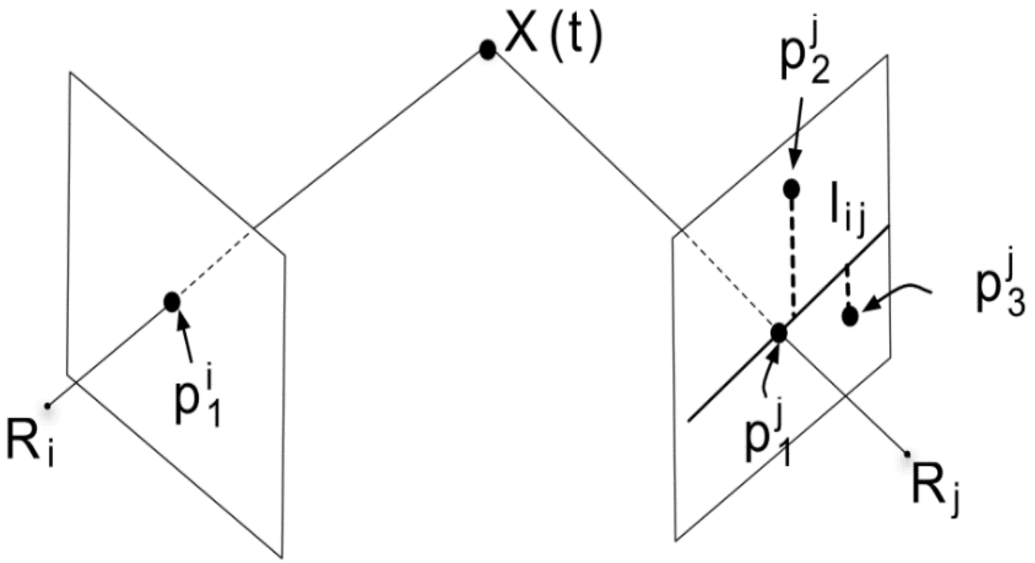
- Models
- Hardware

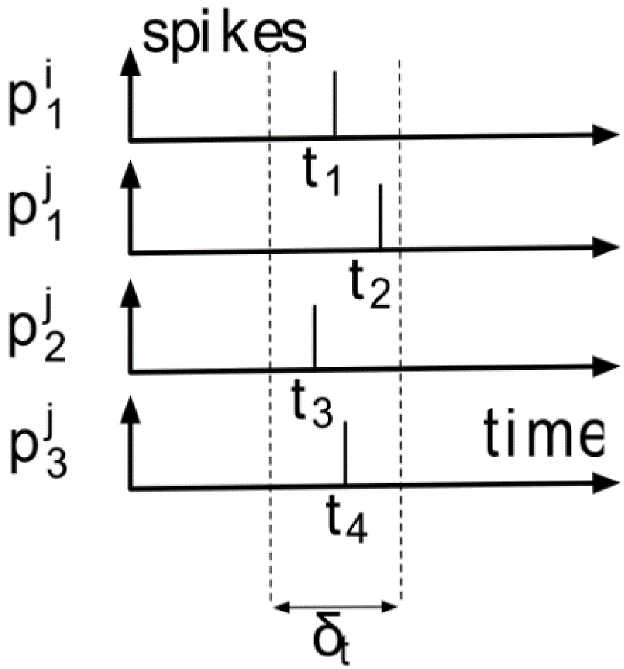
- Prosthetics
- Robotics
- Computation

- Makes of machine vision a science!
- Develop new bidirectional methodology to understand the brain
- Merging Computational and Biological Vision

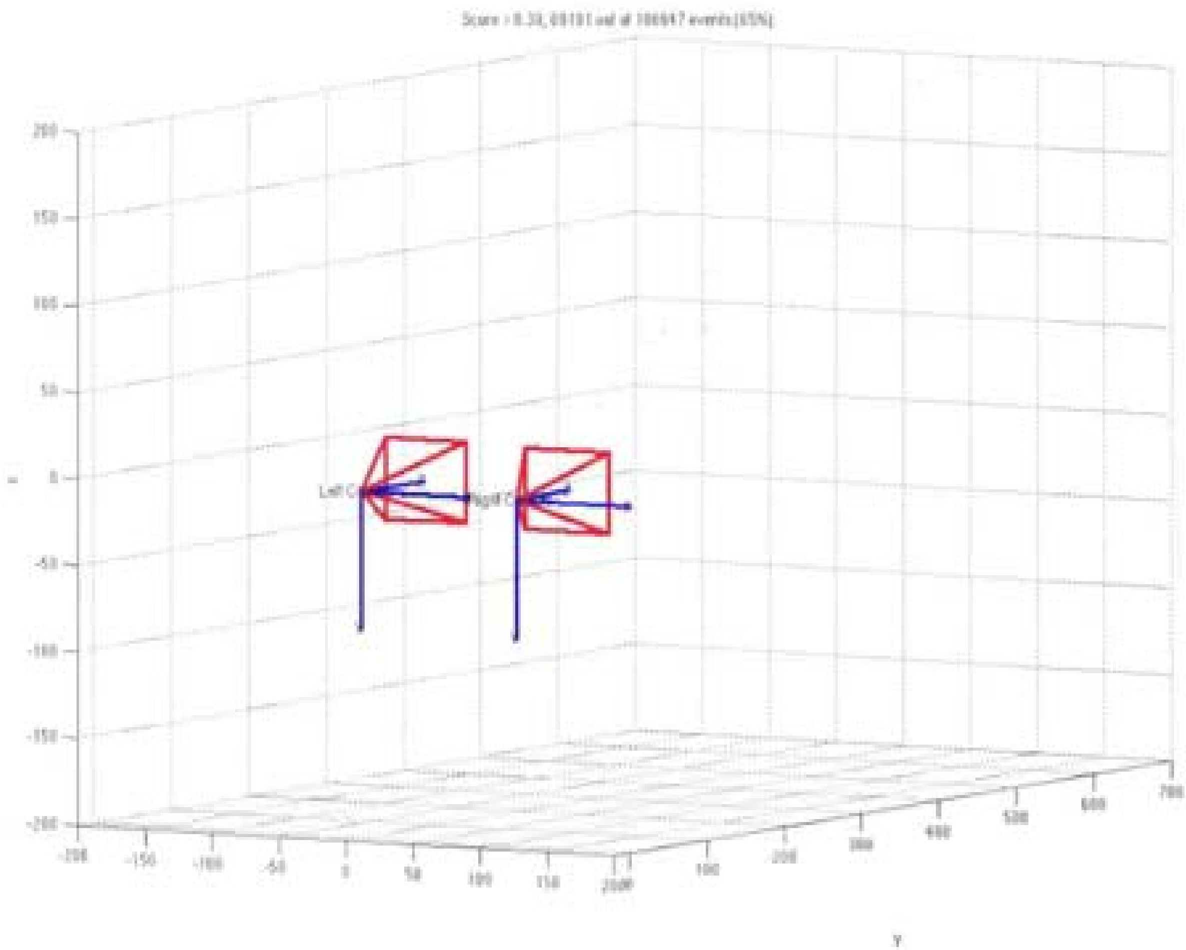
Applications: Stereovision

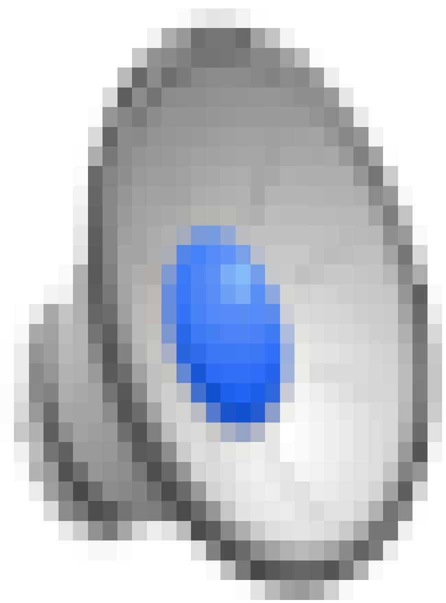




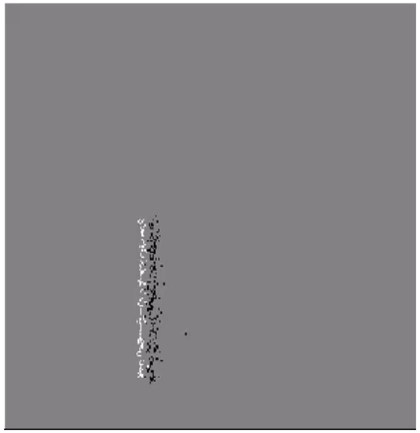


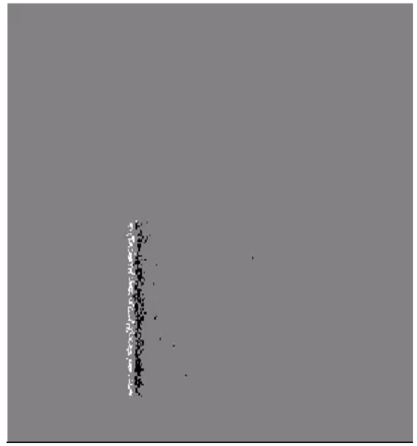
- Matching binocular events only using the time of events
- Two events arriving at the same time and fulfilling geometric constraints are matched

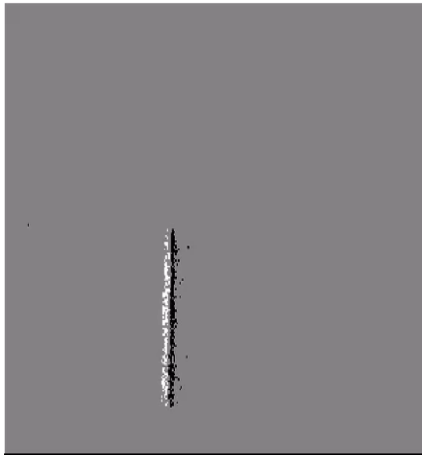

200
150
100
50
0
-50
-100
-150
-200
-200
-150
-100
-50
0
50
100
150
200
100
200
300
400
500
600
700
y

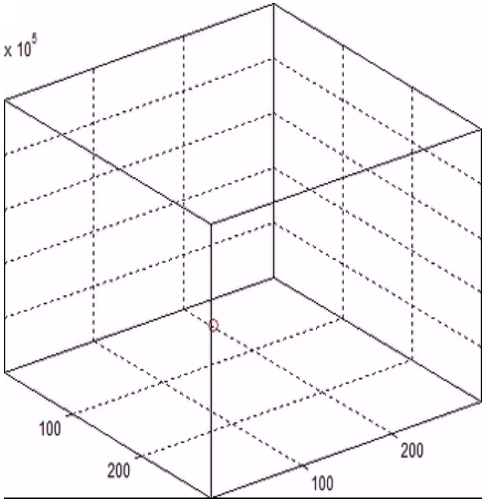


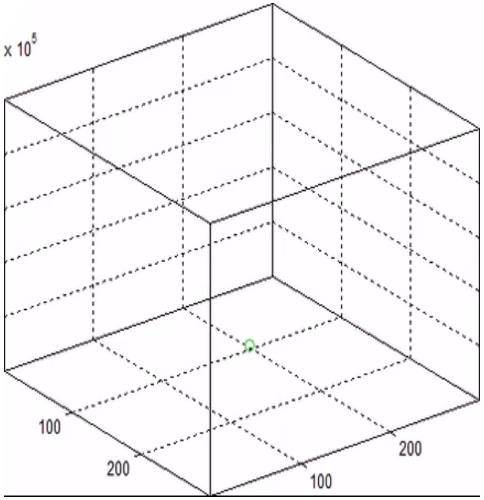
Motion estimation: optical flow

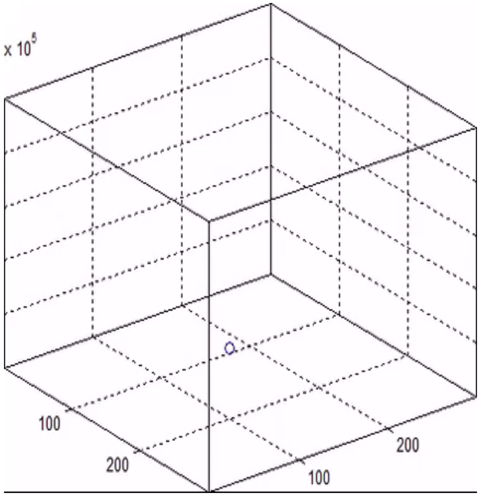






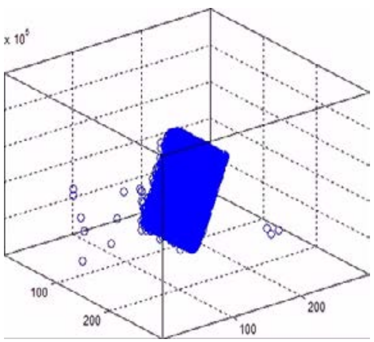






Optical flow

time

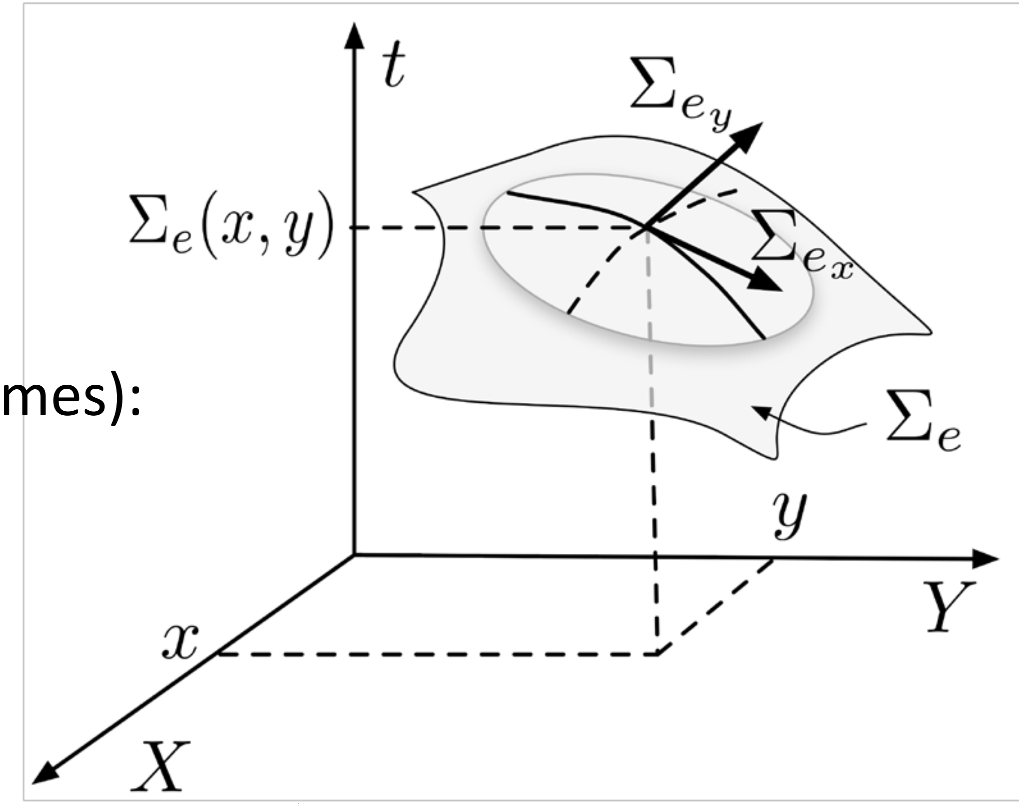


x

y

- High temporal resolution allows to generate smooth space-time surface
- The slope of the local surface contains the orientation and amplitude of the optical flow

Visual Motion flow:



For an incoming event :

$$e(\mathrm{p}, t) = (\mathrm{p}, t)^T$$

Form the surface (image of times):

$$\Sigma_e : \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

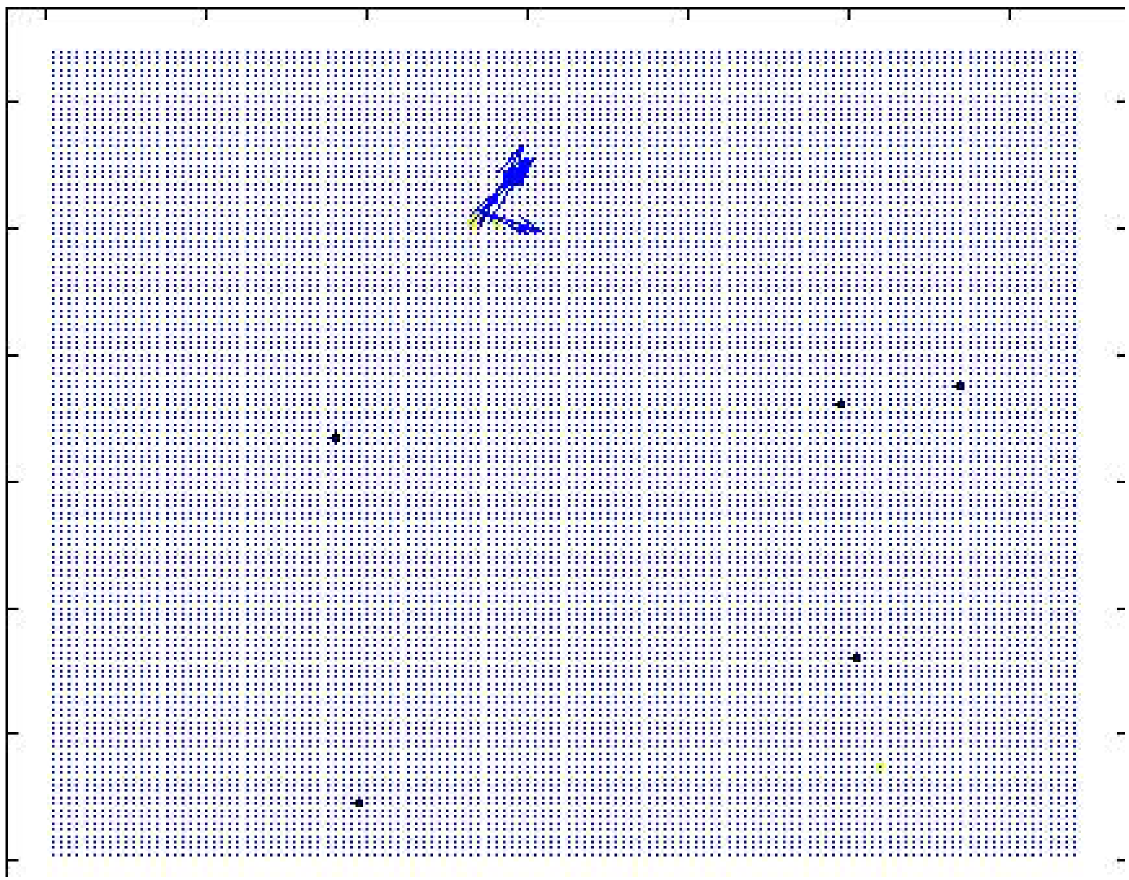
$$\mathrm{p} \mapsto t = \Sigma_e.$$

We then have:

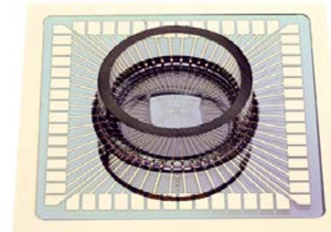
$$\frac{\partial \Sigma_e}{\partial x}(x, y_0) = \frac{\mathrm{d}\Sigma_e|_{y=y_0}}{\mathrm{d}x}(x) = \frac{1}{v_x(x, y_0)},$$

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

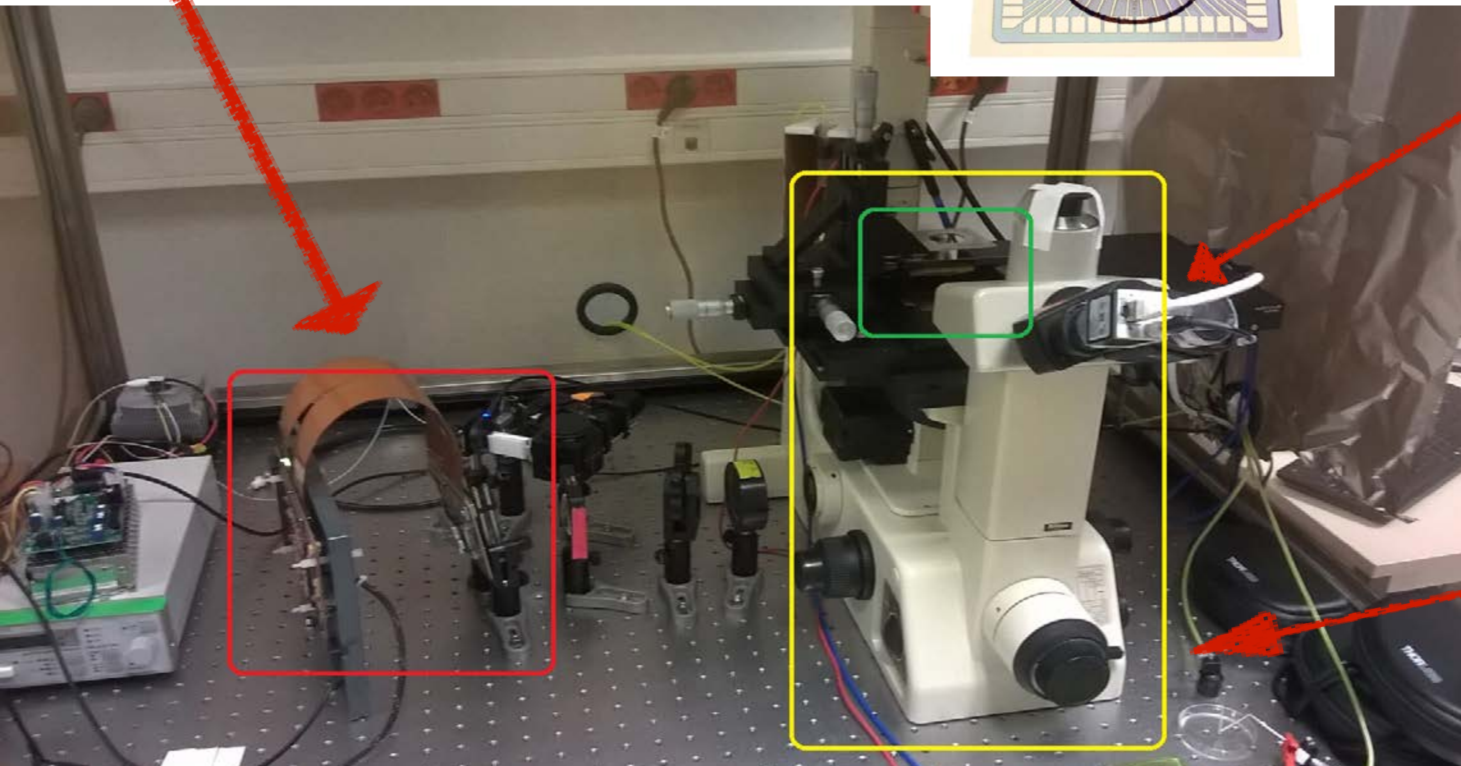
$$\Rightarrow \quad \nabla \Sigma_e = \left(\frac{1}{v_x}, \frac{1}{v_y}\right)^T,$$



projection system

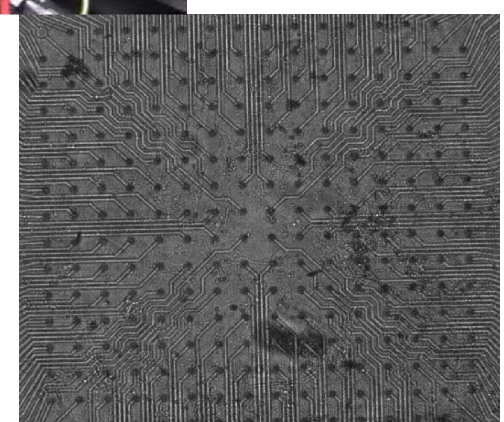


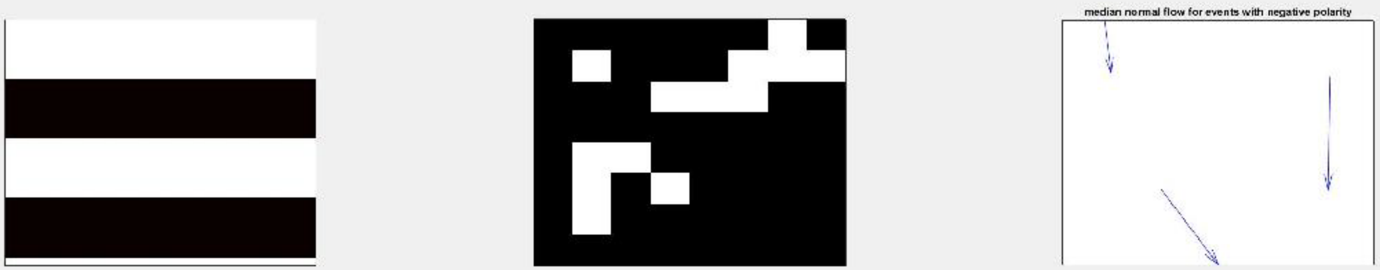
electrodes



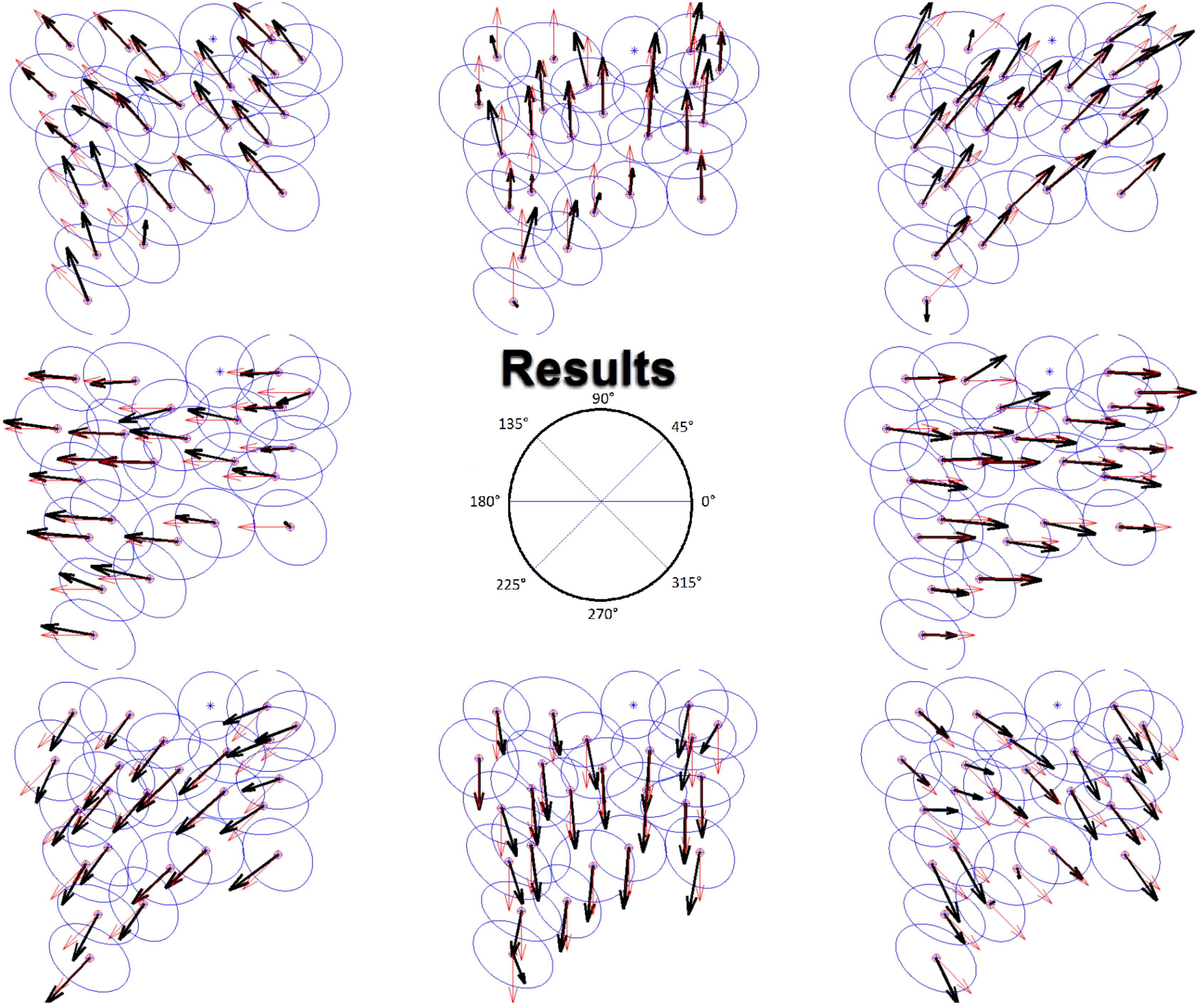
microscope

- Multichannel system 16*16 electrodes
- Visual stimulation frequency up to 1ms
- 20kHz recording precision of neural activities

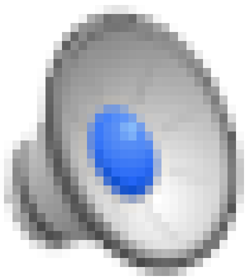


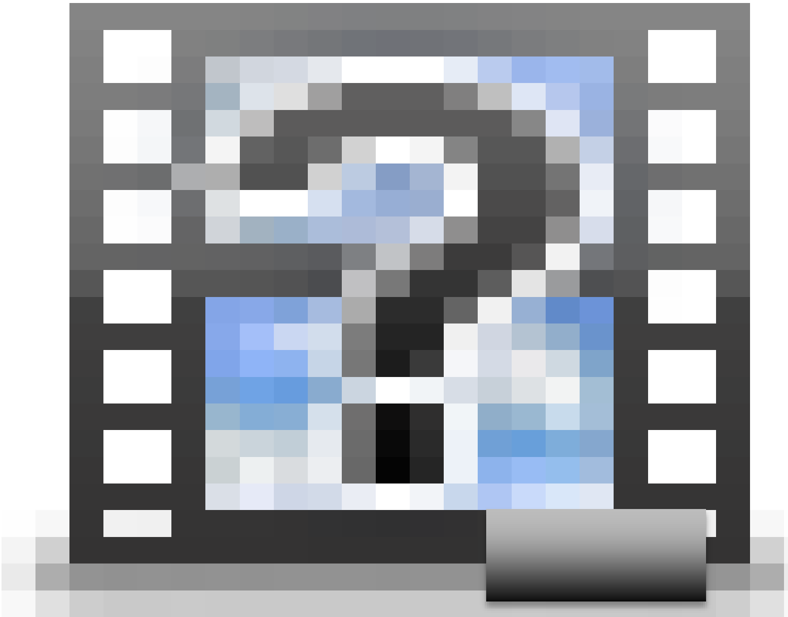

median normal flow for events with negative polarity

Results

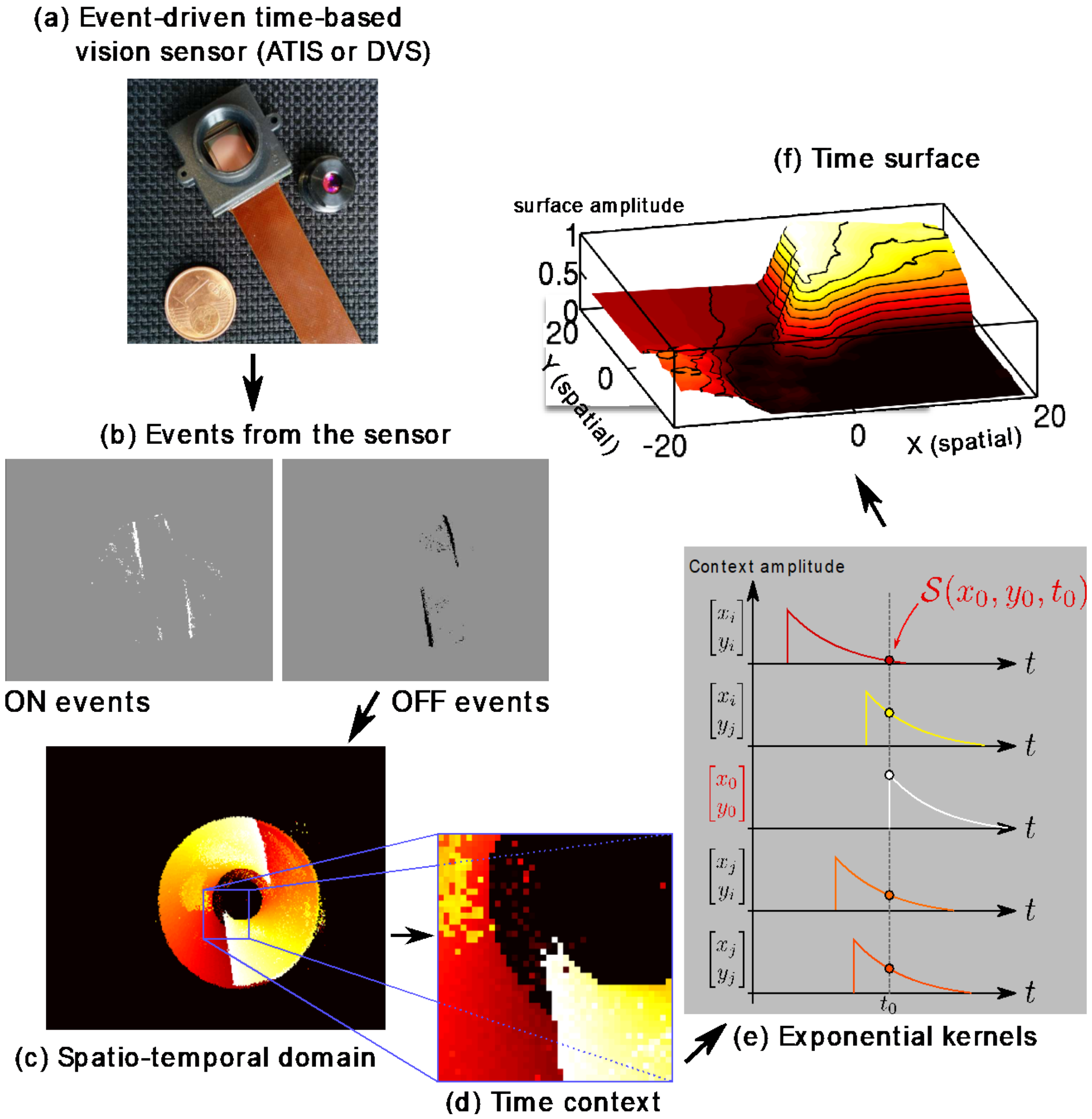


Rewriting the whole computer vision

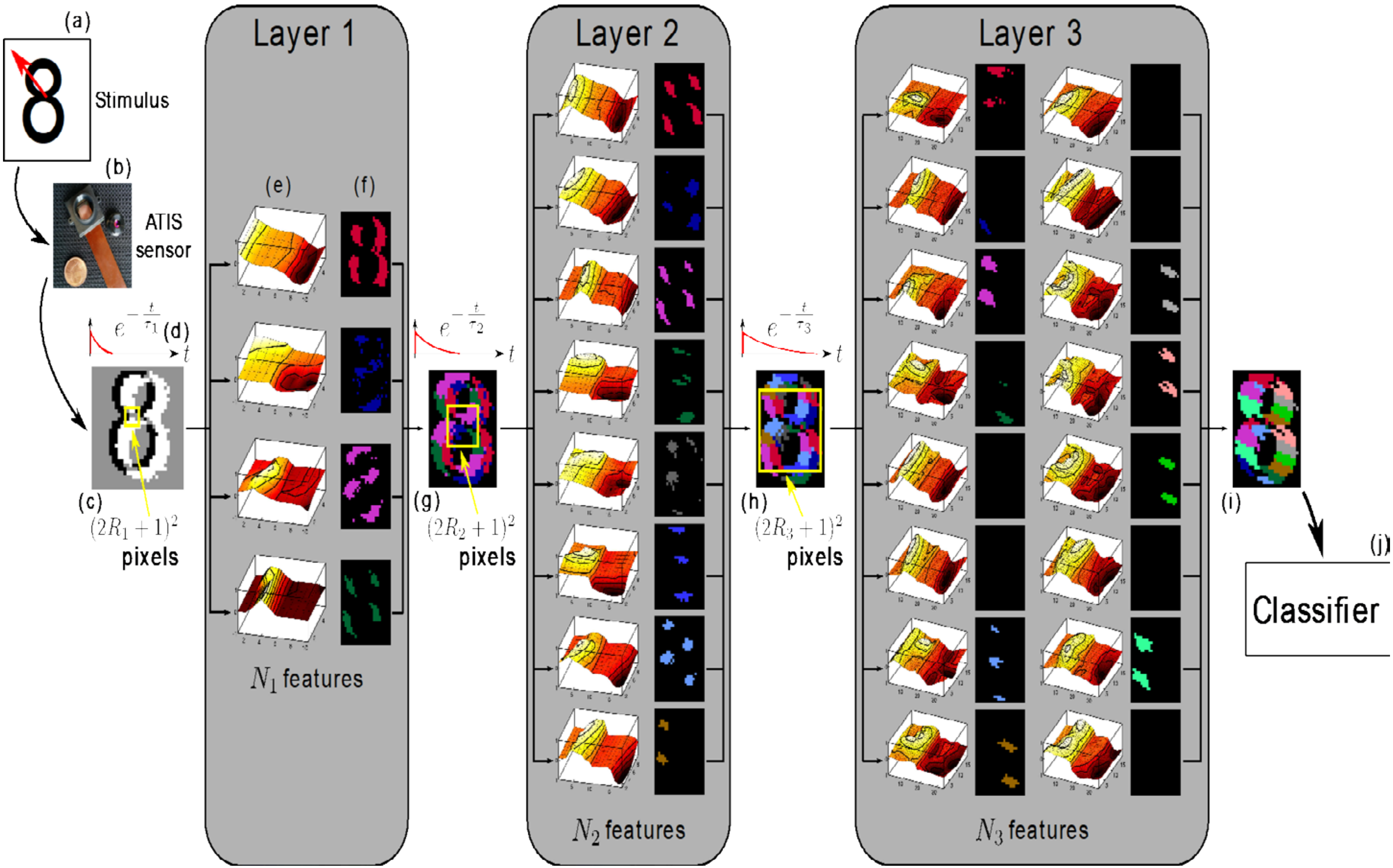




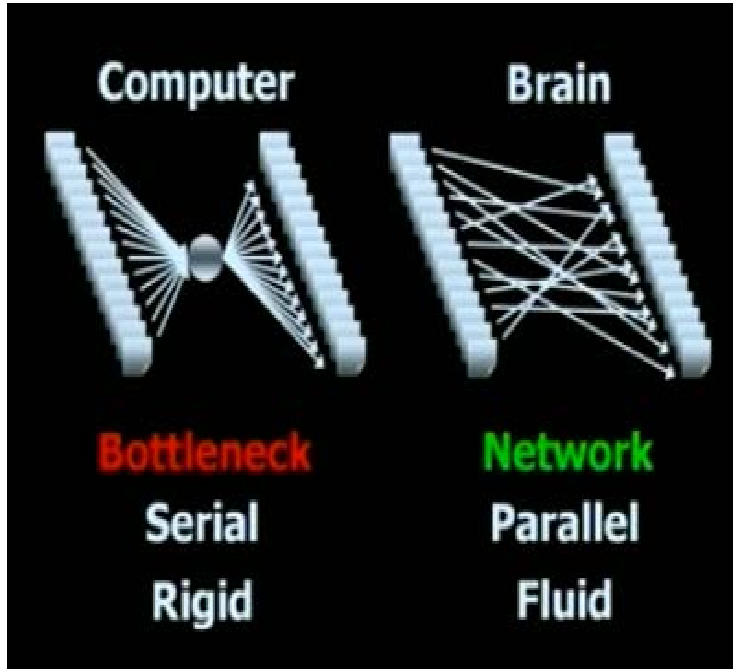
Dynamic Machine Learning: time surfaces

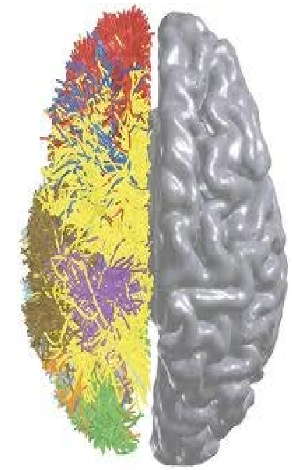


HOTS: A Hierarchy Of event-based Time-Surfaces

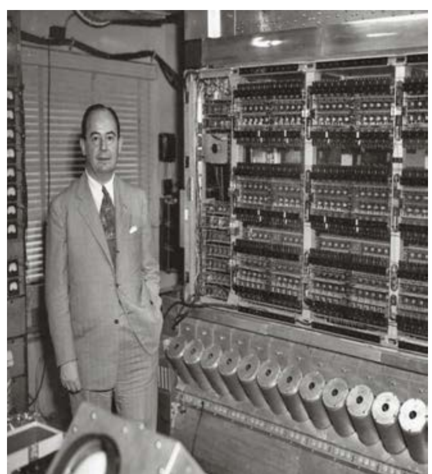


Renaissance of Event-based computing



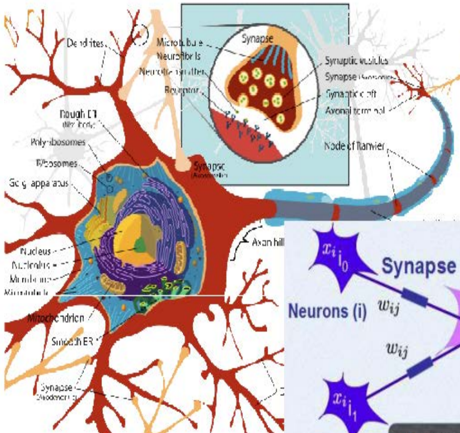


Von Neumann Architecture

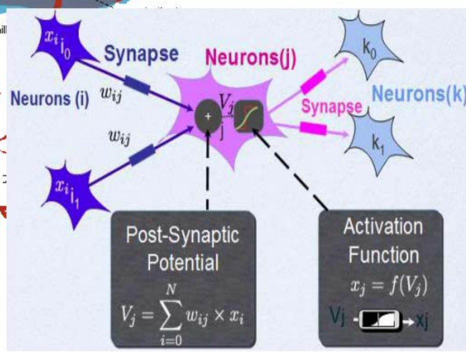


Artificial Neural processors

Neuromorphic Computing, an old story!



Biological neural network

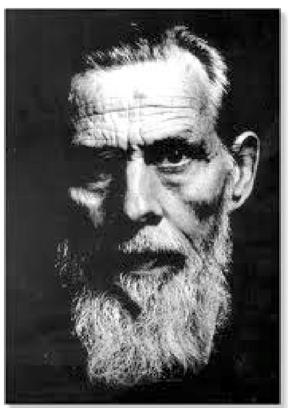


Artificial neural network

Synapse

- Historical memory
- Analog levels

[1] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophysics, no. 5, pp. 115-133, **1943**.



Warren McCulloch

Walter Pitts

A Logical Calculus of Ideas Immanent in Nervous Activity

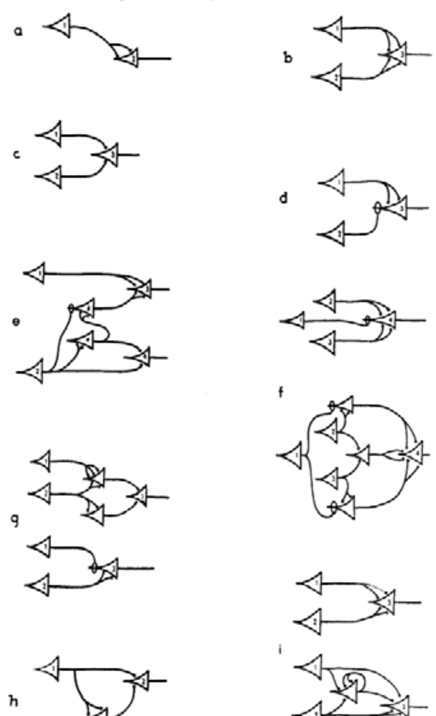


FIGURE 1

A Logical Calculus of Ideas Immanent in Nervous Activity

observations and of these to the facts is all too clear, for it is apparent that every idea and every sensation is realized by activity within that net, and by no such activity are the actual afferents fully determined.

There is no theory we may hold and no observation we can make that will retain so much as its old defective reference to the facts if the net be altered. Tinnitus, paraesthesias, hallucinations, delusions, confusions and disorientations intervene. Thus empiry confirms that if our nets are undefined, our facts are undefined, and to the "real" we can attribute not so much as one quality or "form." With determination of the net, the unknowable object of knowledge, the "thing in itself," ceases to be unknowable.

To psychology, however defined, specification of the net would contribute all that could be achieved in that field—even if the analysis were pushed to ultimate psychic units or "psychons," for a psychon can be no less than the activity of a single neuron. Since that activity is inherently propositional, all psychic events have an intentional, or "semiotic," character. The "all-or-none" law of these activities, and the conformity of their relations to those of the logic of propositions, insure that the relations of

EXPRESSION FOR THE FIGURES

In the figure the neuron c_i is always marked with the numeral i upon the body of the cell, and the corresponding action is denoted by 'N' with i as subscript, as in the text.

Figure 1a $N_2(t) . \equiv . N_1(t-1)$

Figure 1b $N_3(t) . \equiv . N_1(t-1) \vee N_2(t-1)$

Figure 1c $N_3(t) . \equiv . N_1(t-1) . N_2(t-1)$

Figure 1d $N_3(t) . \equiv . N_1(t-1) . \sim N_2(t-1)$

Figure 1e $N_3(t) : \equiv : N_1(t-1) . \vee . N_2(t-3) . \sim N_2(t-2)$
$N_4(t) . \equiv . N_2(t-2) . N_2(t-1)$

Figure 1f $N_4(t) : \equiv : \sim N_1(t-1) . N_2(t-1) \vee N_3(t-1) . \vee . N_1(t-1) . N_2(t-1) . N_3(t-1)$
$N_4(t) : \equiv : \sim N_1(t-2) . N_2(t-2) \vee N_3(t-2) . \vee . N_1(t-2) . N_2(t-2) . N_3(t-2)$

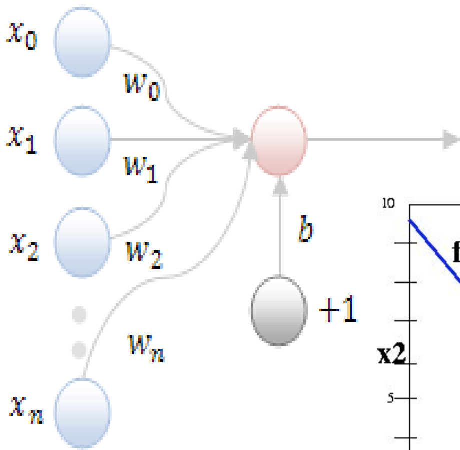
Figure 1g $N_3(t) . \equiv . N_2(t-2) . \sim N_1(t-3)$

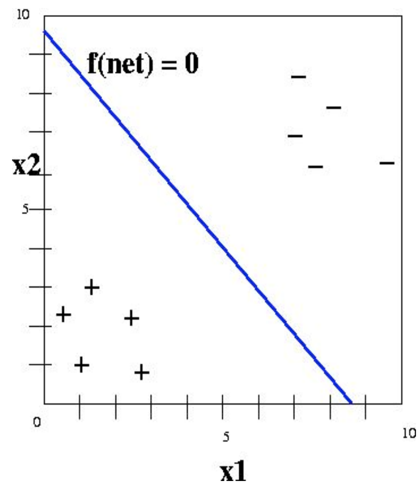
Figure 1h $N_2(t) . \equiv . N_1(t-1) . N_1(t-2)$

Figure 1i $N_3(t) : \equiv : N_2(t-1) . \vee . N_1(t-1) . (Ex)t-1 . N_1(x) . N_2(x)$

(Slide adapted from Christian Gamrat's ESSDERC 2012 presentation)

Perceptron: first neuromorphic engine

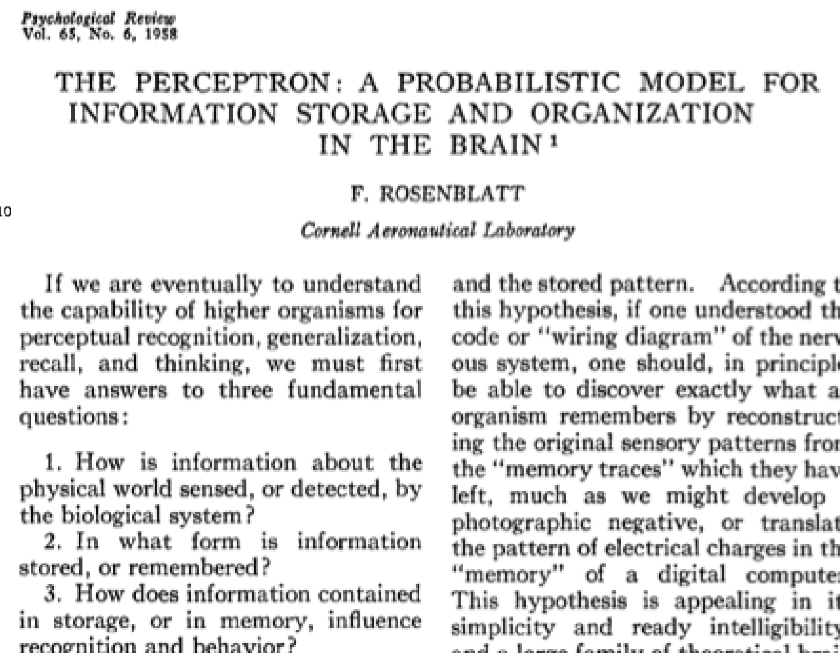




[1] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," Psychological Review, vol. 65, no. 6, pp. 386-408, **1958**.

(Robert Hecht-Nilsen: Neurocomputing, Addison-Wesley, 1990)

Frank Rosenblatt



Psychological Review
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN [1]

F. ROSENBLATT

Cornell Aeronautical Laboratory

If we are eventually to understand the capability of higher organisms for perceptual recognition, generalization, recall, and thinking, we must first have answers to three fundamental questions:

1. How is information about the physical world sensed, or detected, by the biological system?
2. In what form is information stored, or remembered?
3. How does information contained in storage, or in memory, influence recognition and behavior?

and the stored pattern. According to this hypothesis, if one understood the code or "wiring diagram" of the nervous system, one should, in principle, be able to discover exactly what an organism remembers by reconstructing the original sensory patterns from the "memory traces" which they have left, much as we might develop a photographic negative, or translate the pattern of electrical charges in the "memory" of a digital computer. This hypothesis is appealing in its simplicity and ready intelligibility, and a large family of theoretical brain

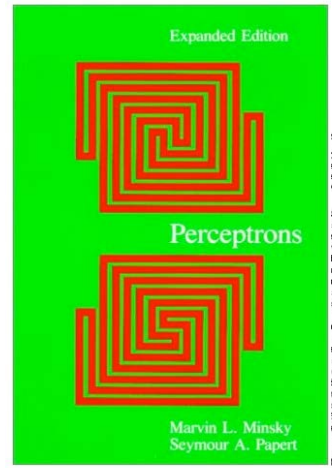
(Slide adapted from Christian Gamrat's ESSDERC 2012 presentation)

The big depression of the 1970's

- Minsky an Papert's book on Perceptrons is seen by many as the cause of the drop in ANN research (the XOR problem)

[1] M. L. Minsky and S. A. Papert, Perceptrons: An Introduction to Computational Geometry. The MIT Press, **1970**.

Marvin Minsky & Seymour Papert



Book Reviews

standing of Information Processes

figure. Let there be a set of predicates, call them ϕ, which one can think of as elementary measurements on the space R. Then a perceptron is a predicate which can be represented in the form:

$$\psi(X) \text{ is true if } \sum_{\text{all } \phi} \alpha(\phi)\phi(X) > \theta$$

$$\psi(X) \text{ is false if } \sum_{\text{all } \phi} \alpha(\phi)\phi(X) \leq \theta$$

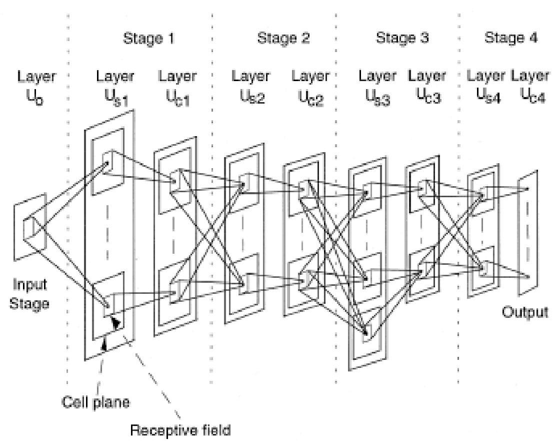
where the coefficients, α, and the threshold, θ, are real numbers and the values

in some fixed way, the α, and ask if the evidence adds up to enough, θ, to warrant saying that X is an instance of the pattern (equivalently, deciding yes). Although this corresponds to the oft-expressed intuitive notion that judgments are made by "weighing the evidence," it must be made clear that perceptrons are an extremely restricted class of decision devices. In most real decisions there is much exploring of consequences, returning for new information, redefinition of the situation, and so on. None of these processes find expression in the perceptron, as formulated. Nevertheless, perceptrons still constitute a nontrivial type of decision element, and—as Minsky and Papert note—if we cannot understand the behavior of perceptrons we have little chance with the more complex decision processes.

The book states and proves a large number of theorems about perceptrons. For any interesting theory, one must restrict the elementary measurements (the ϕ), since otherwise the whole burden of the decision could be put on them, the combinational aspect that is the essence of the definition thus being bypassed entirely. Two restrictions are proposed: *diameter-limited* perceptrons, in which the points on which a ϕ depends must all lie within a circle of given diameter (though the whole collection of ϕ can cover R many times over); and *order-limited* perceptrons, in which the number of points on which a ϕ depends must be less than a given number (though the points can be located anywhere on the retina). Both restrictions fit an intuitive notion that the ϕ are somehow simple, limited and local predicates, so that the act of

ceptron that can recognize when a figure is connected, as opposed to being disconnected. This holds for both diameter-limited and order-limited perceptrons, though the proof for the first is direct and for the latter quite complex. In general the results are of this negative character. For instance, it is possible for there to be perceptrons of order 1 for two predicates, yet no perceptron of finite order that will recognize the disjunction (or, similarly, the conjunction) of the two predicates. In the development of the theory some powerful tools are constructed. Perhaps the most central is the group-invariance theorem, which states that if a perceptron is to be invariant over a (finite) group of transformations on the retina, then there must exist a particularly simple form of the weighted sum (namely, where all coefficients of those ϕ which are equivalent under the group are the same). The power of this theorem arises from the close connection between notions of what is interesting geometrically and properties that are invariant under groups of transformation. Thus the theorem reflects something of the geometry of the retina in the algebraic structure of the perceptron.

Still other results concern the fact that though order-limited perceptrons exist for some classes of patterns, their coefficients (more precisely, the ratio between the smallest and largest coefficient) may be exceedingly large—so large, indeed, that one might as well store the instances directly, since that would require fewer bits than storing the coefficients. There is a chapter on learning in perceptrons in which one considers the ϕ fixed and asks what procedures might discover appropriate weights to do a particular pattern-recognition task. The information from which the weights are inferred is a sequence of instances of the patterns. There is a perceptron convergence theorem which states that a particularly simple form of feedback modification of the weights under the impact of the sequence will indeed find a workable set of weights if such exists. Finally, there is a comparison of the perceptron with various highly serial algorithms



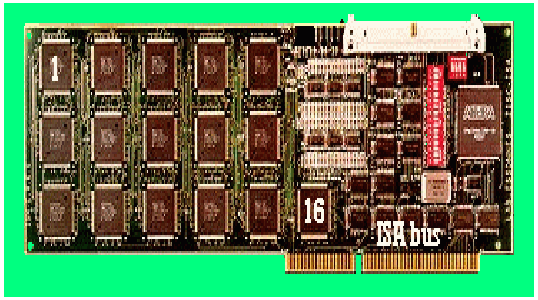
Kunihiko Fukushima

[1] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biological Cybernetics, vol. 36, no. 4, pp. 193-202, 1980.

(Slide adapted from Christian Gamrat's ESSDERC 2012 presentation)

1980's Neurocomputers...

- Siemens : MA-16 Chips (SYNAPSE-1 Machine)
  - Synapse-1, neurocomputer with 8xM-A16 chips
  - Synapse3-PC, PCI board with 2xMA-16 (1.28 Gpcs)
- Adaptive Solutions : CNAPS
  - SIMD // machine based on a 64 PE chip.
- IBM : ZISC
  - Vector classifier engine
- Philips : L-Neuro (M. Duranton)
  - 1st Gen 16PEs 26 MCps
  - 2nd Gen 12 PEs 720 MCps
- + Intel (ETANN), AT&T (Anna), Hitachi (WSI), NEC, Thomson (now THALES), etc…



(Slide adapted from Christian Gamrat's ESSDERC 2012 presentation)

How to encode numbers with neurons?

Necessity to find an alternative to binary

Development of Elementary Numerical Abilities: A Neuronal Model

Stanislas Dehaene
INSERM and CNRS, Paris

Jean-Pierre Changeux
Institut Pasteur, Paris

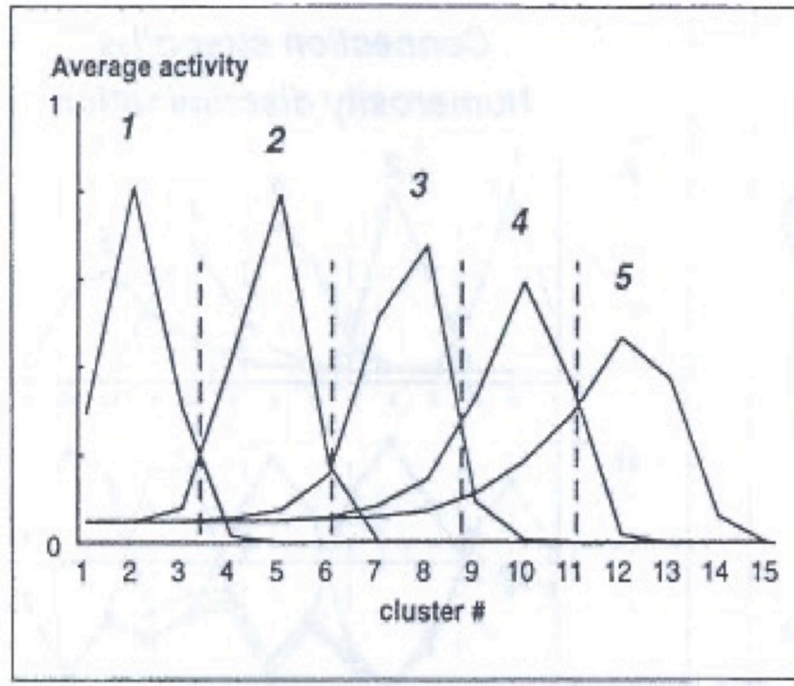
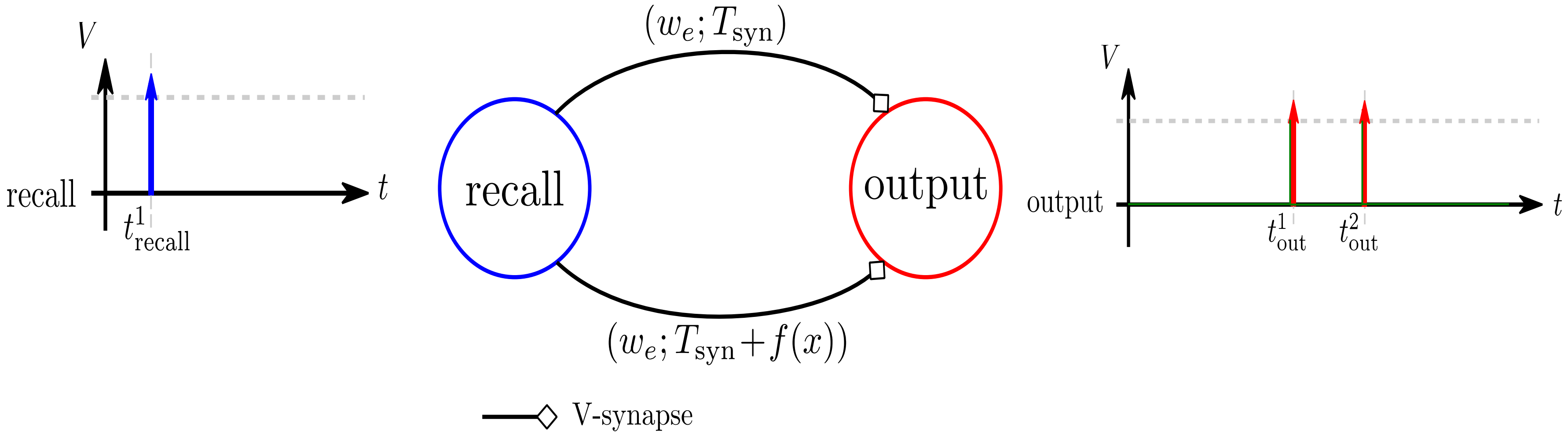


Figure 4. Average activity of numerosity clusters when random sets of 1, 2, 3, 4, or 5 objects were presented for input. For each input numerosity, only a small number of clusters were selectively activated (e.g., clusters 1, 2, and 3 responded only when a single object was presented). The activity peaks were lower and wider for larger numerosities, implying a decrease in discriminability with increasing numerosity (Fechner's law).

How to encode numbers ?

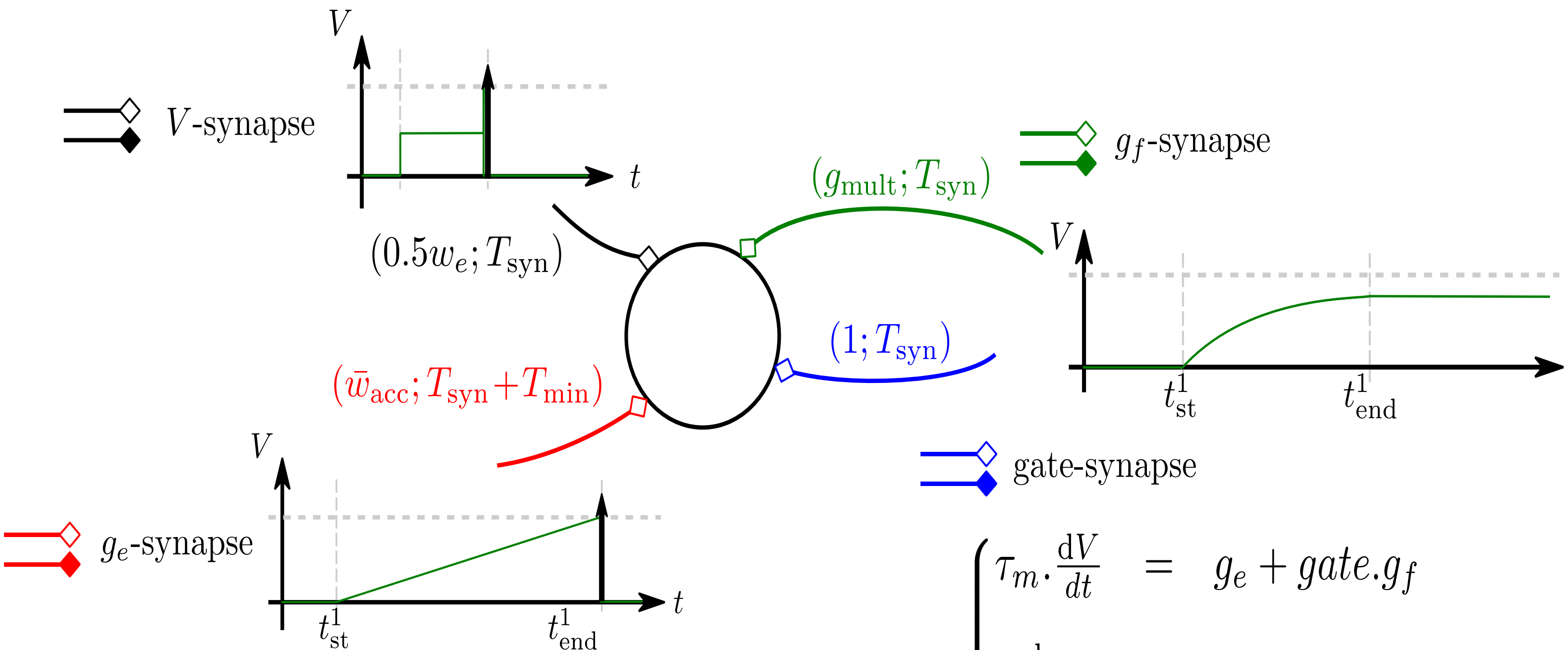


$$\Delta t = f(x) = T_{\text{min}} + x.T_{\text{cod}}$$

STICK: Spike Time Interval Computational Kernel, a Framework for General Purpose Computation Using Neurons, Precise Timing, Delays, and Synchrony

Xavier Lagorce
xavier.lagorce@upmc.fr
Ryad Benosman
ryad.benosman@upmc.fr

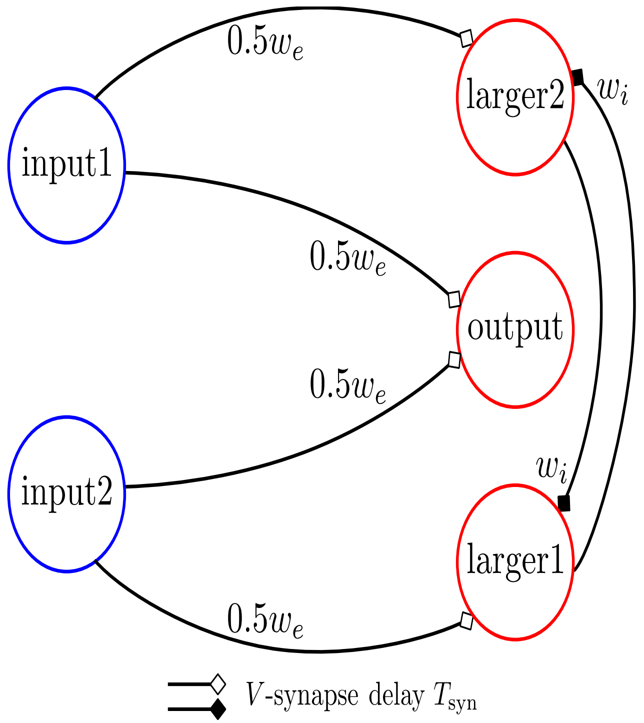
elementary units

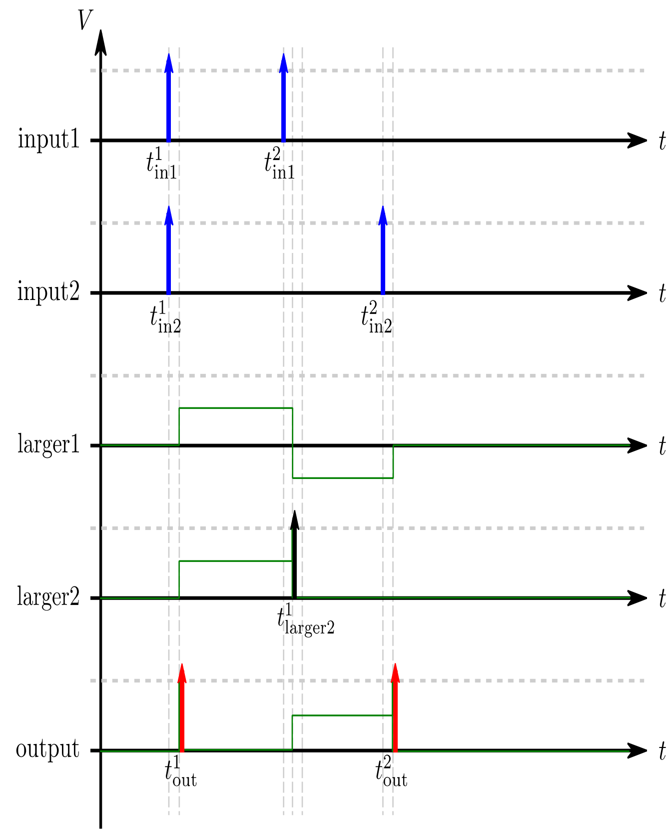


State evolution:

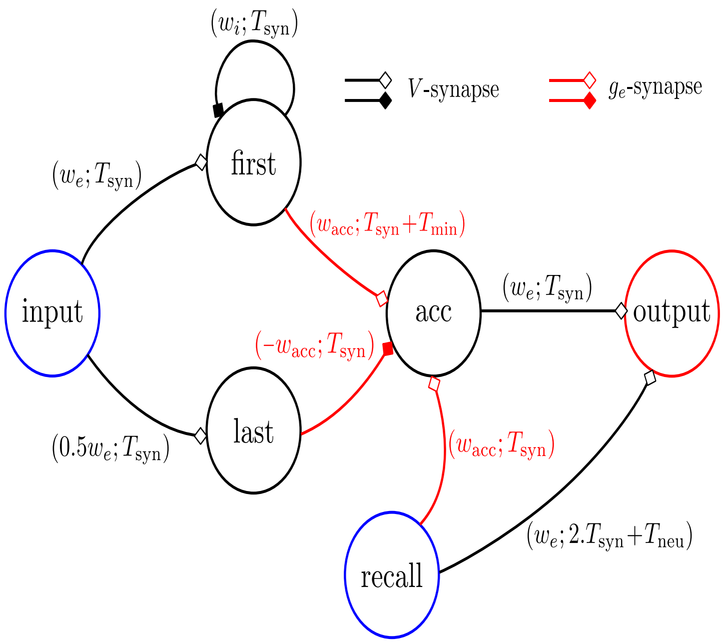
$$\begin{cases} \tau_m . \dfrac{\mathrm{d}V}{dt} & = & g_e + gate.g_f \\ \dfrac{\mathrm{d}g_e}{dt} & = & 0 \\ \tau_f . \dfrac{\mathrm{d}g_f}{dt} & = & -g_f \end{cases}$$

Compute Maximum

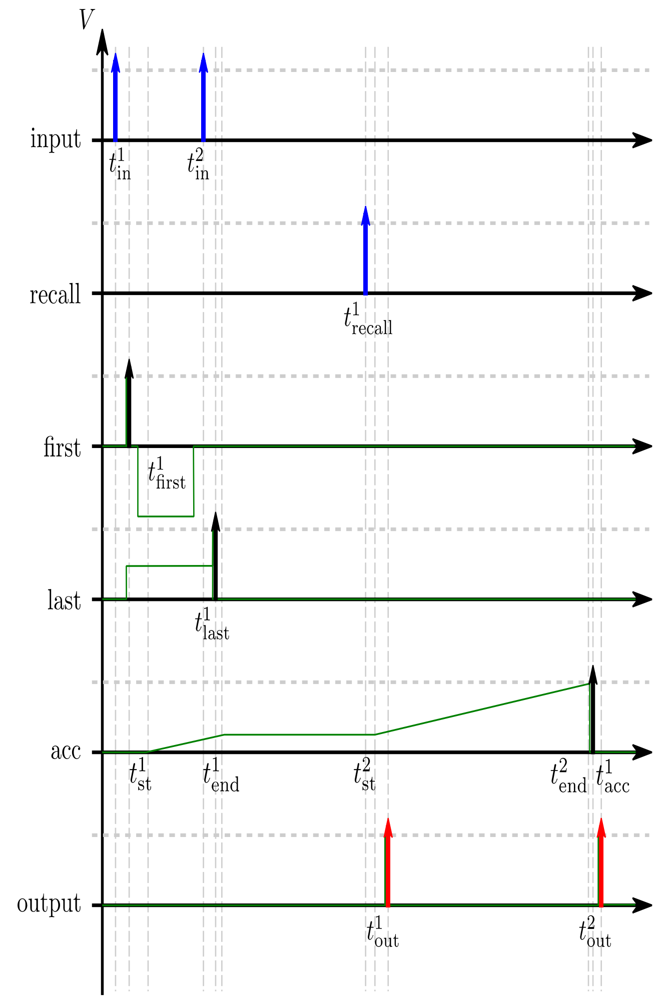




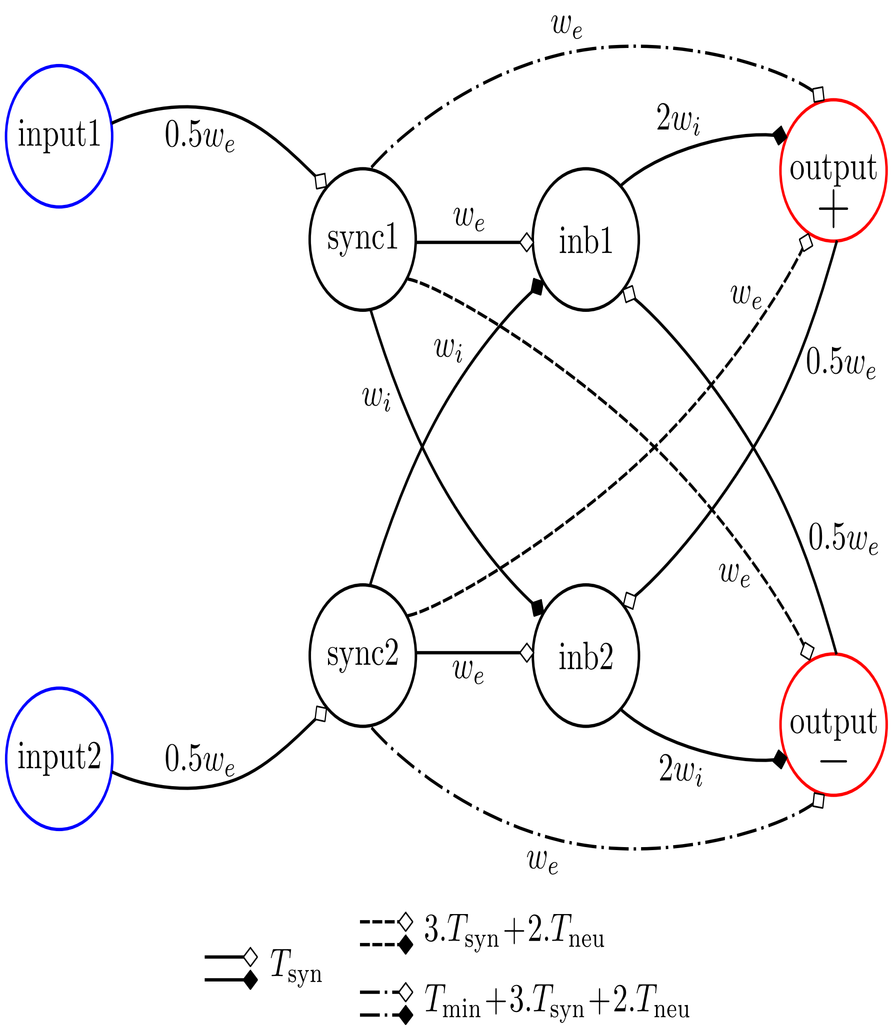
Storing information: an inverting Memory network

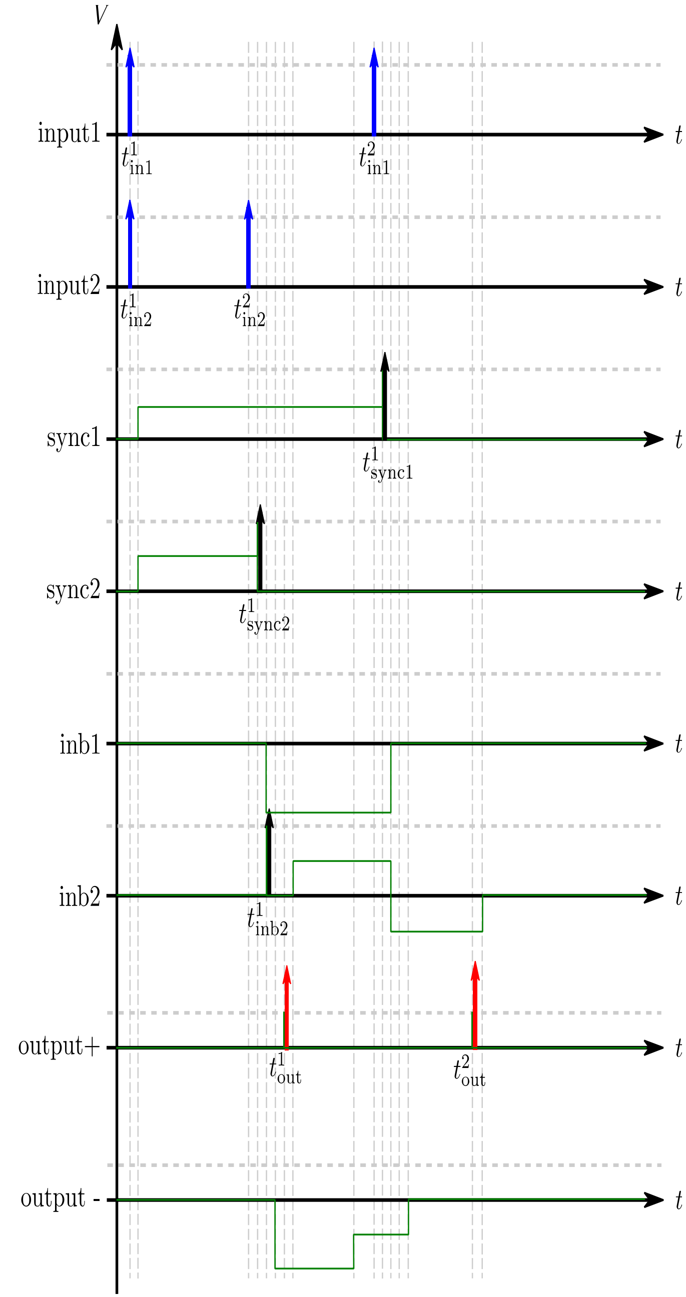


$$\begin{cases} \tau_m.\dfrac{\mathrm{d}V}{dt} = g_e + gate.g_f \\ \dfrac{\mathrm{d}g_e}{dt} = 0 \\ \tau_f.\dfrac{\mathrm{d}g_f}{dt} = -g_f \end{cases}$$

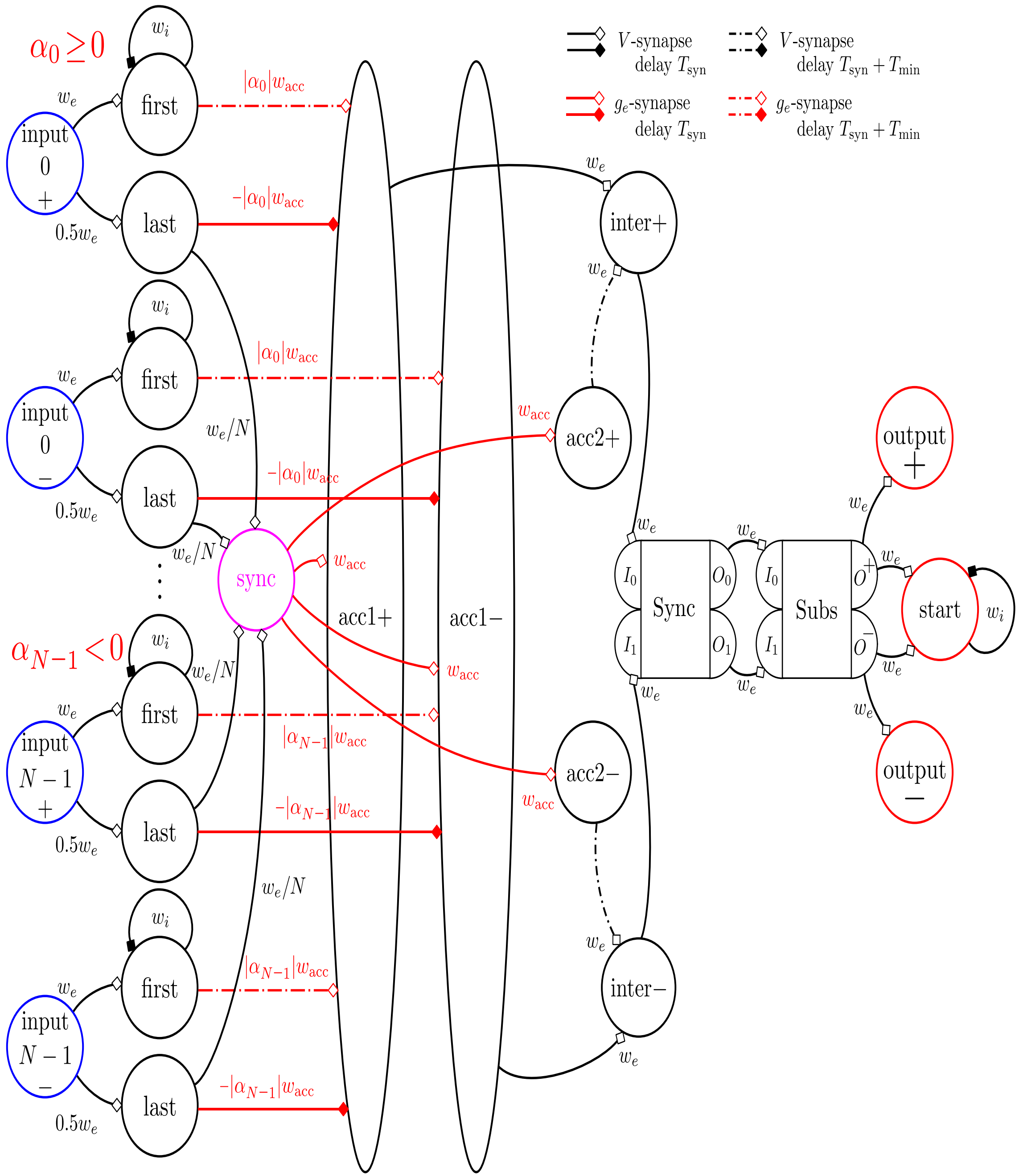


Subtractor network



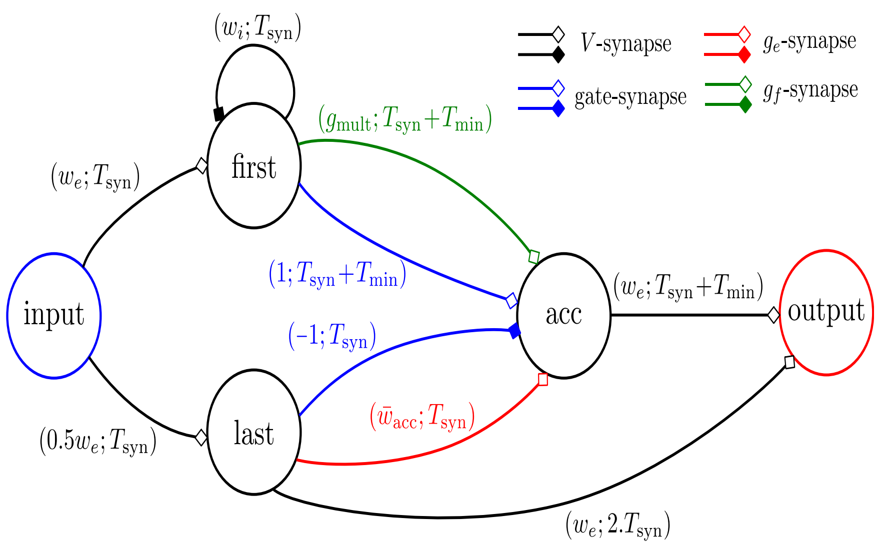


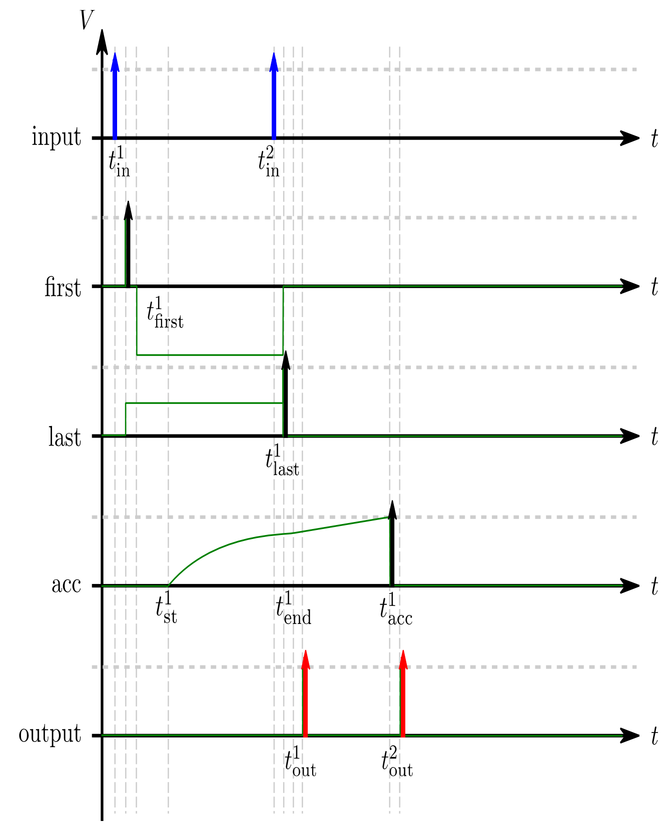
Linear Combination network



Non-linearities: exponential

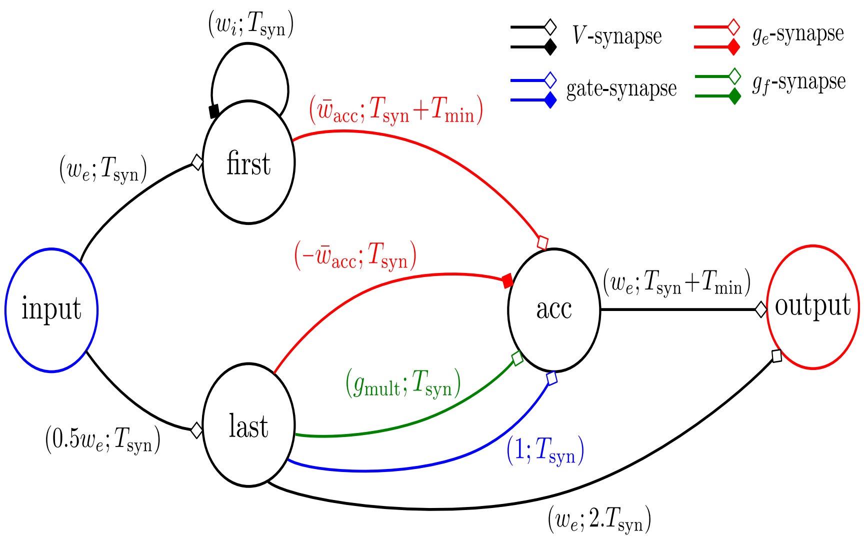
$$\begin{cases} \tau_m . \frac{\mathrm{d}V}{dt} & = & g_e + gate.g_f \\ \frac{\mathrm{d}g_e}{dt} & = & 0 \\ \tau_f . \frac{\mathrm{d}g_f}{dt} & = & -g_f \end{cases}$$

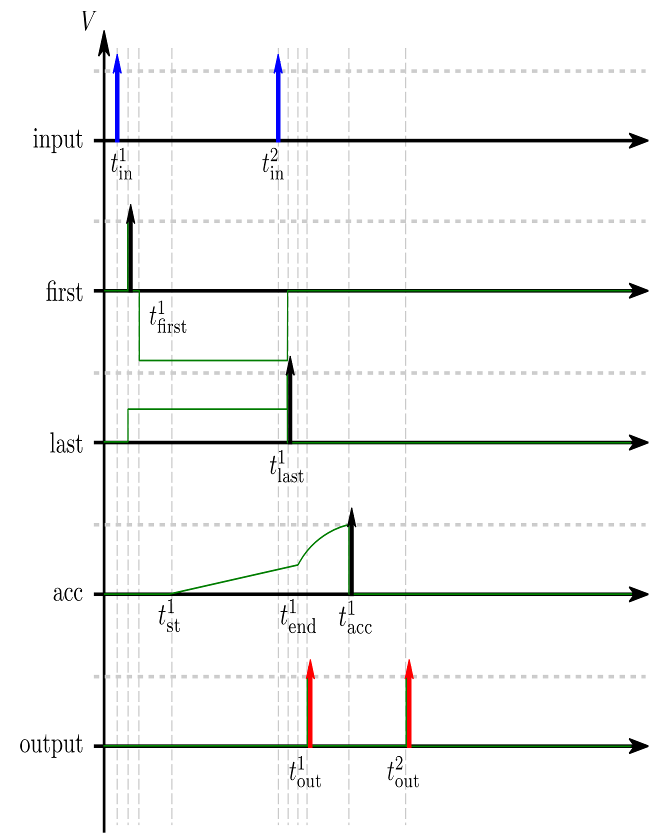




Non-linearities: logarithm

$$\begin{cases} \tau_m.\dfrac{\mathrm{d}V}{dt} &= g_e + gate.g_f \\ \dfrac{\mathrm{d}g_e}{dt} &= 0 \\ \tau_f.\dfrac{\mathrm{d}g_f}{dt} &= -g_f \end{cases}$$

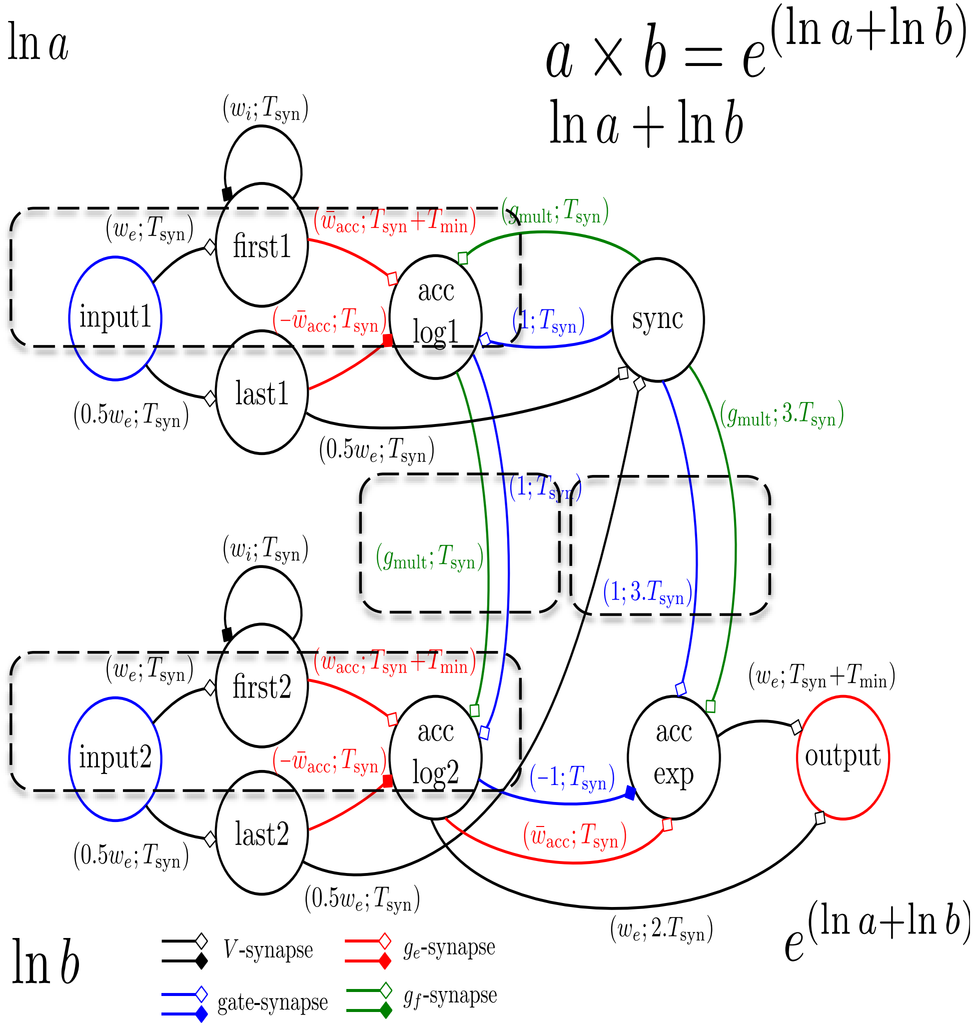


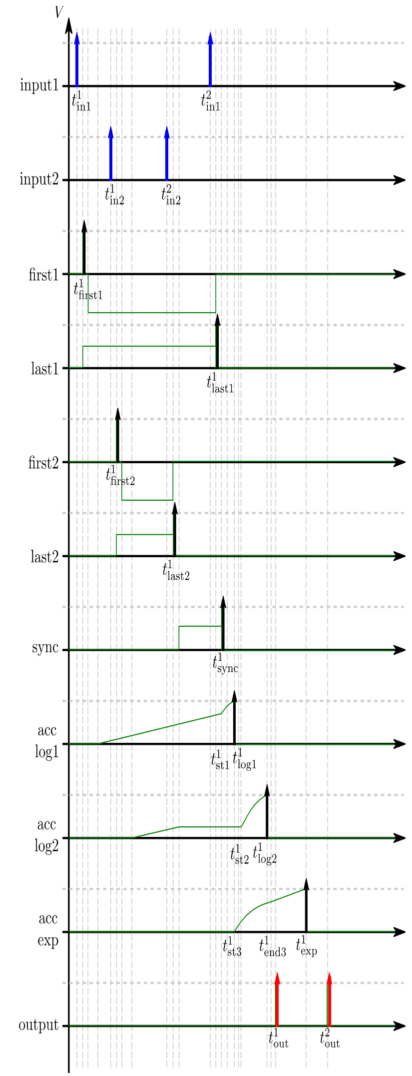


Building a Multiplier network

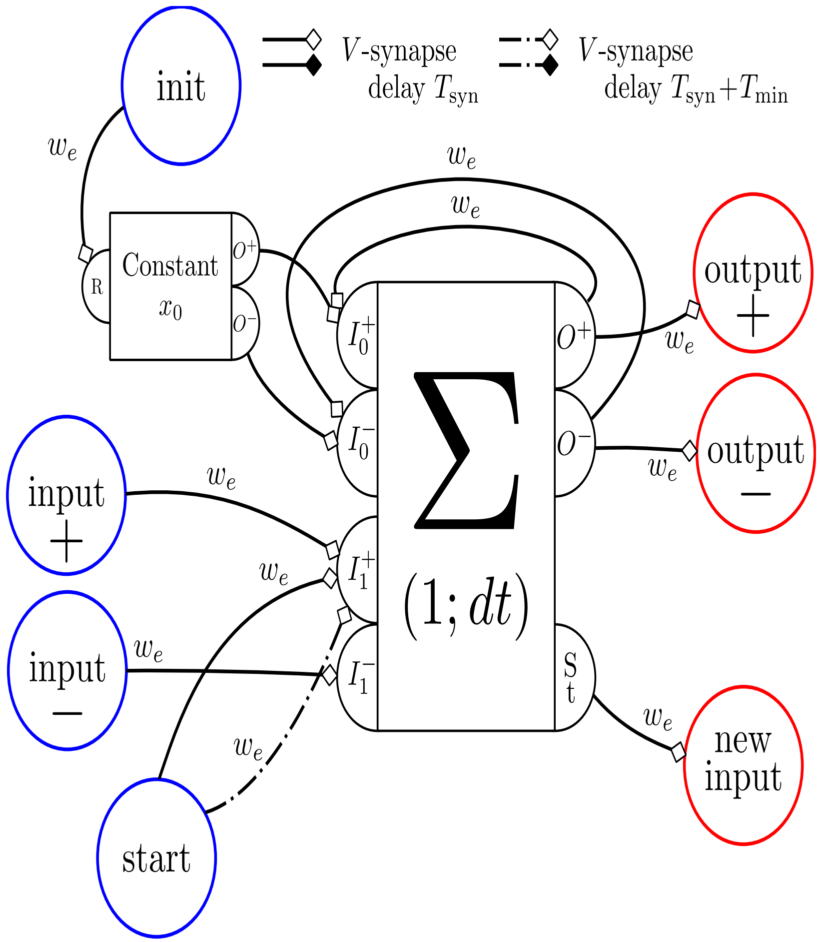
$$a \times b = e^{(\ln a + \ln b)}$$

$$\ln a + \ln b$$



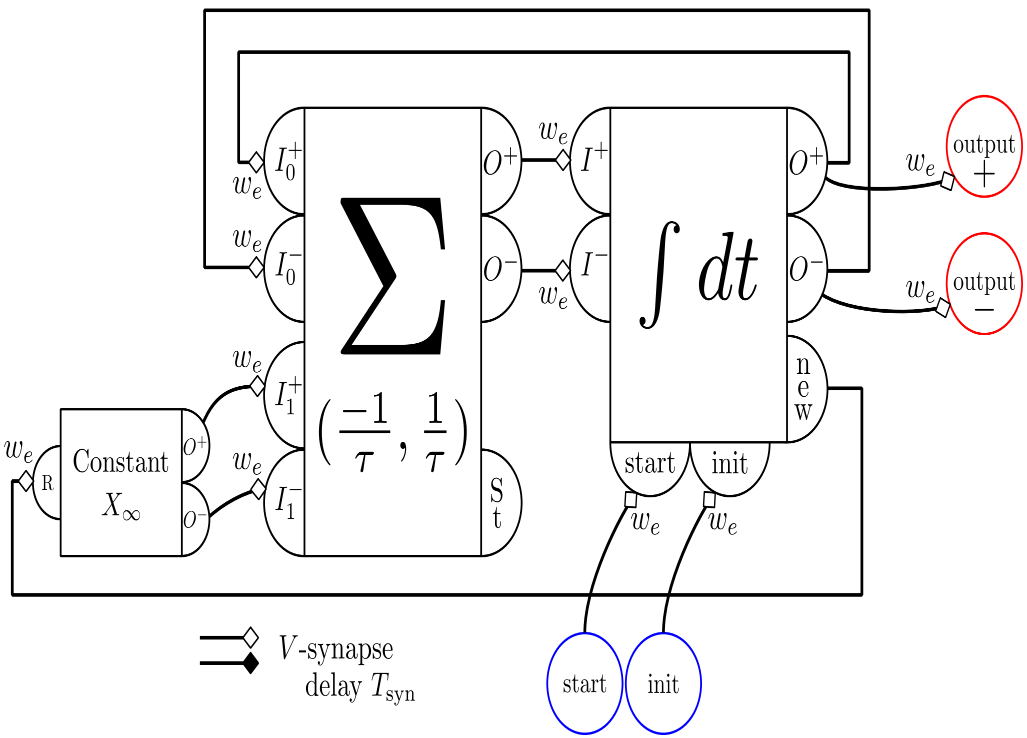


Integrating signals

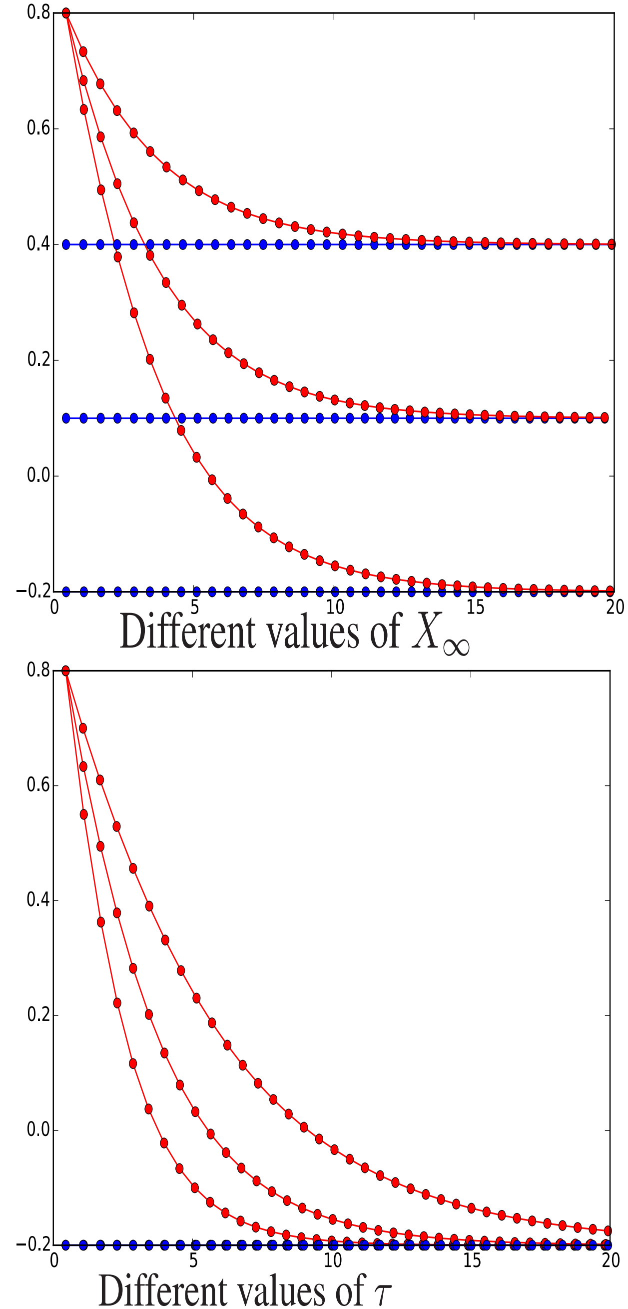


Computing a first-order ODE

$$\tau.\frac{\mathrm{d}X}{dt} + X(t) = X_\infty$$

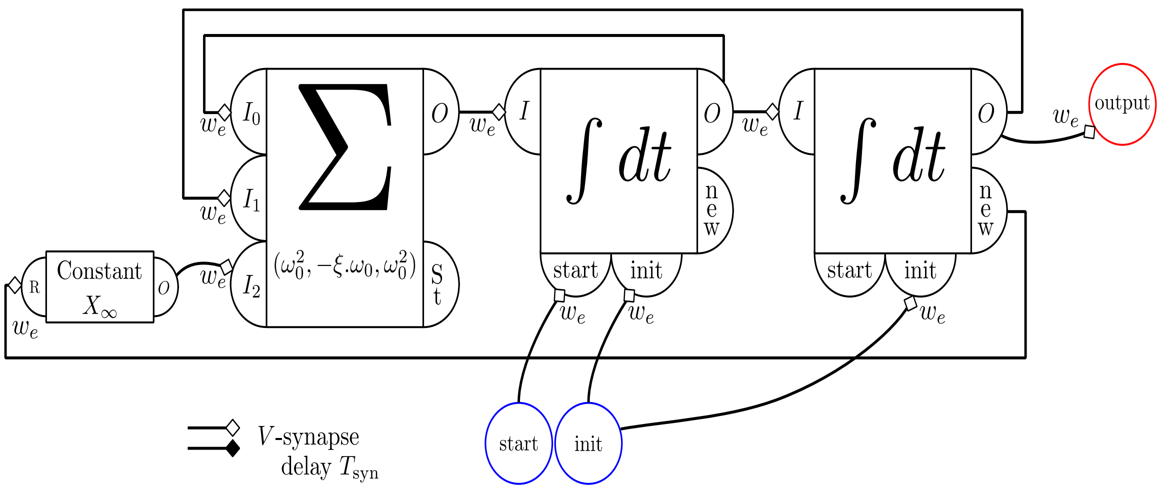


118 neurons

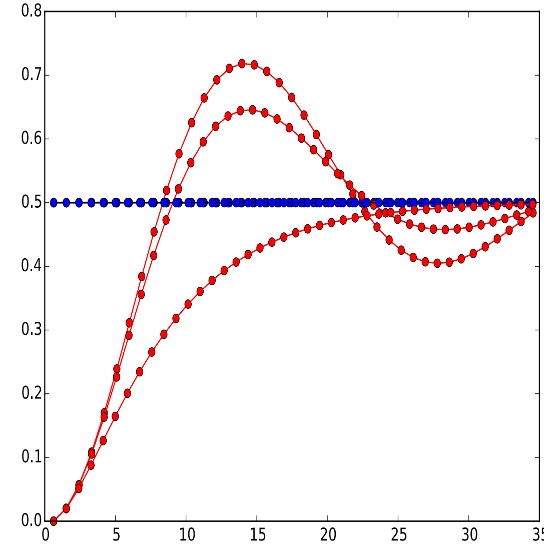


Computing a second-order ODE

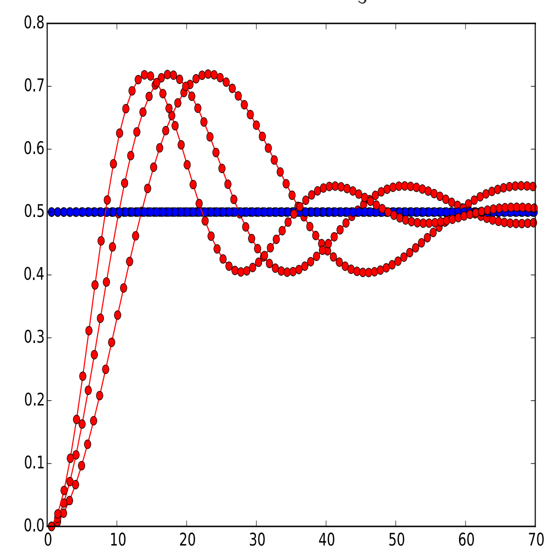
$$\frac{1}{\omega_0^2}\cdot\frac{\mathrm{d}^2X}{dt^2}+\frac{\xi}{\omega_0}\cdot\frac{\mathrm{d}X}{dt}+X(t)=X_\infty$$



187 neurons



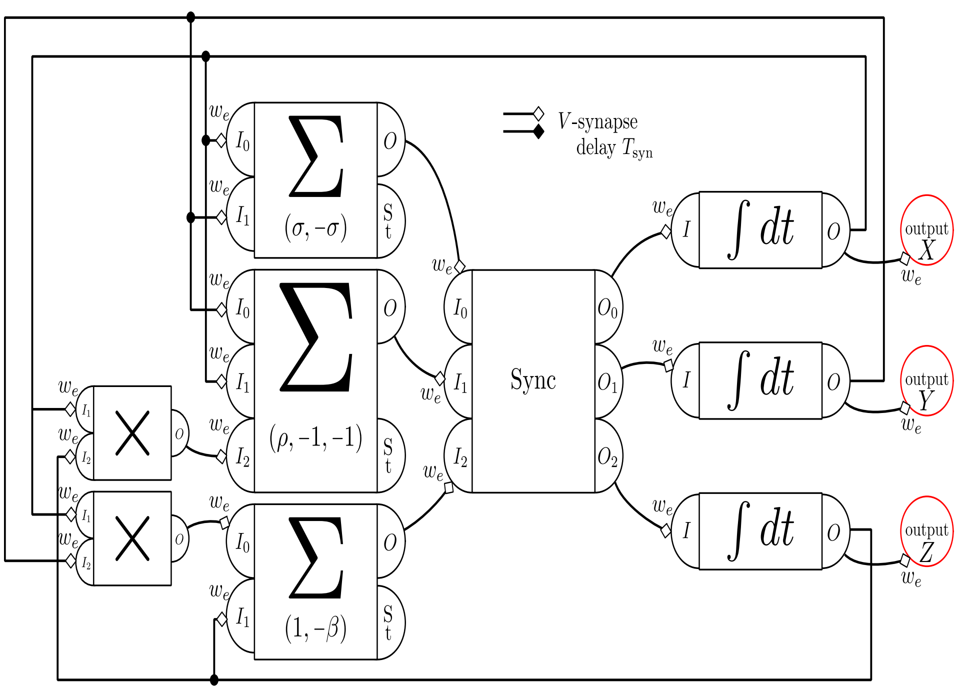
Different values of ξ

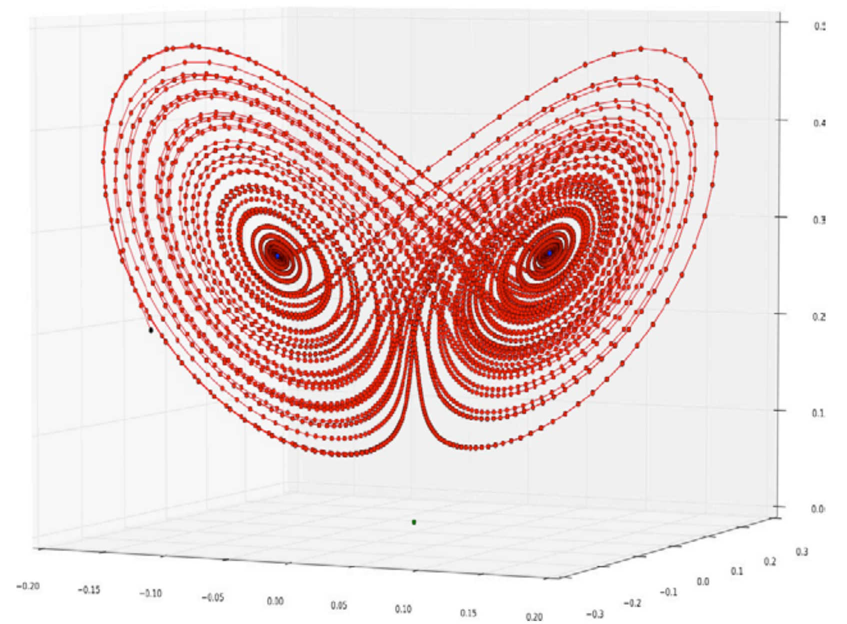


Different values of ω_0

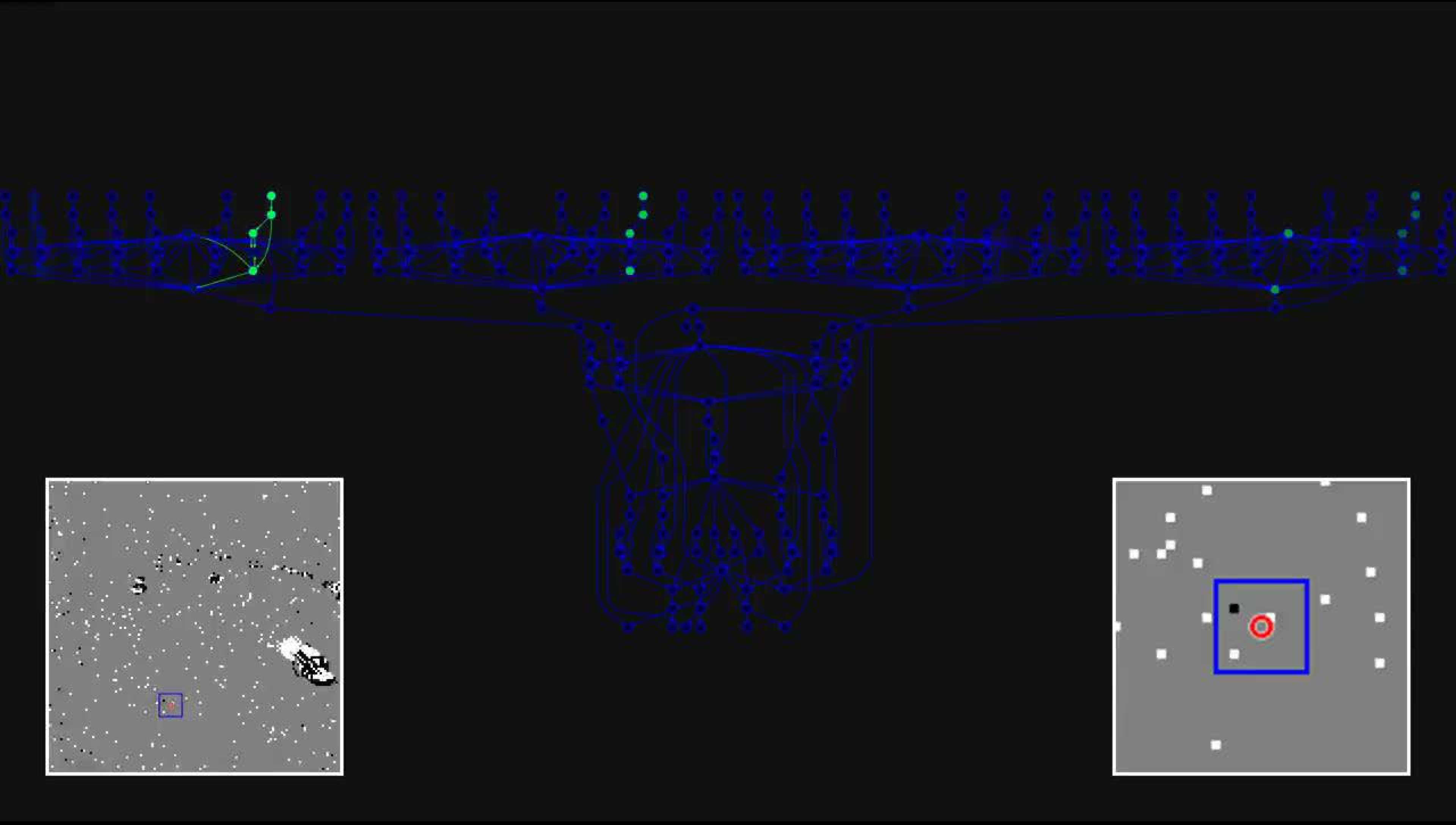
Simulating a Lorenz attractor

$$\frac{\mathrm{d}X}{dt} = \sigma(Y(t) - X(t))$$
$$\frac{\mathrm{d}Y}{dt} = \rho X(t) - Y(t) - X(t).Z(t)$$
$$\frac{\mathrm{d}Z}{dt} = X(t).Y(t) - \beta Z(t)$$



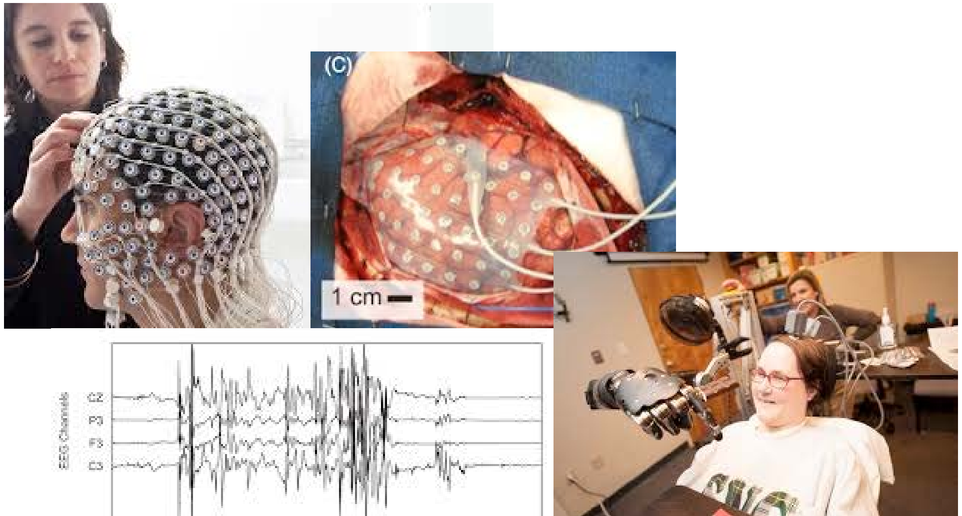


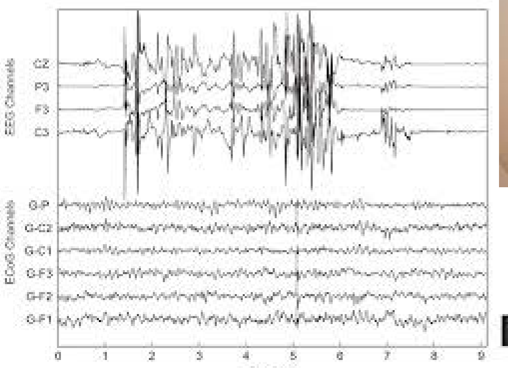
280 neurons




FUTURE MAG

& much more…



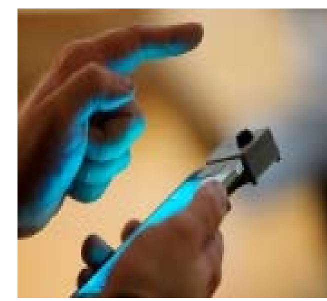


Low power Online
decoding and classification

Robotics

Decision making: game theory
stock Market

Autonomous driving



Always on sensing

Conclusions

- A paradigm shift in AI
- Operate on time rather than luminance information
- Several possible sensors
- Adapted to IOT and low power computation
- Low data bandwidth
- Outperforms conventional image based acquisition