Why is Neuromorphic Event-based Engineering the future of AI?

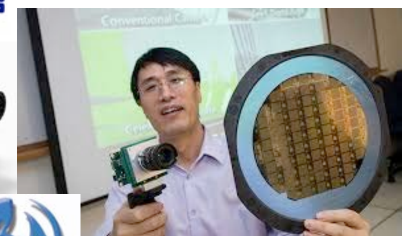
R.B. Benosman,
Professor

Neuromorphic Started in a still **very active group** that meets every year in **Telluride** for **hands on projects** and **talks**…

Sensors

OmniVision

iniLabs



insightness

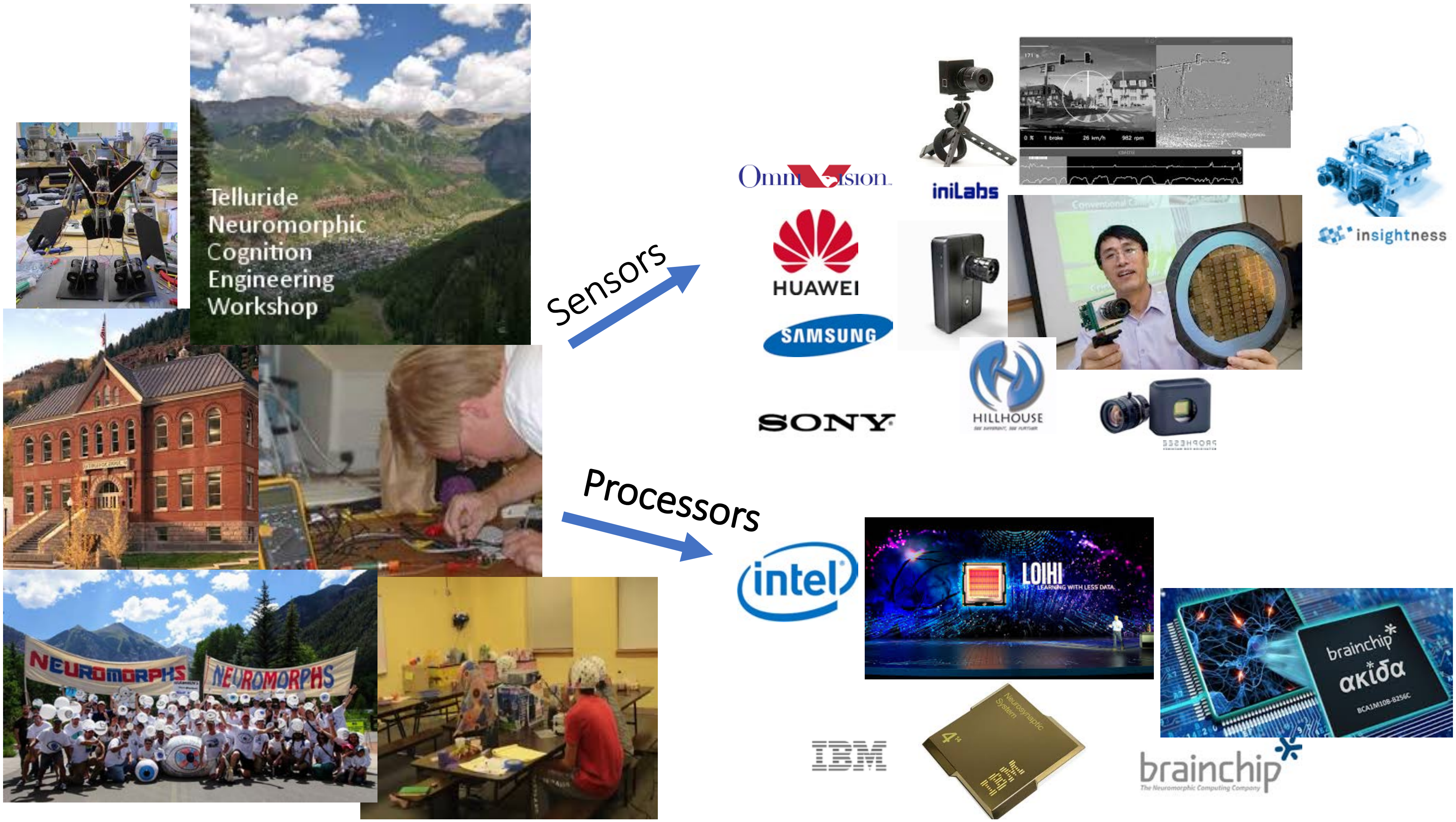
HUAWEI

SAMSUNG

SONY

HILLHOUSE

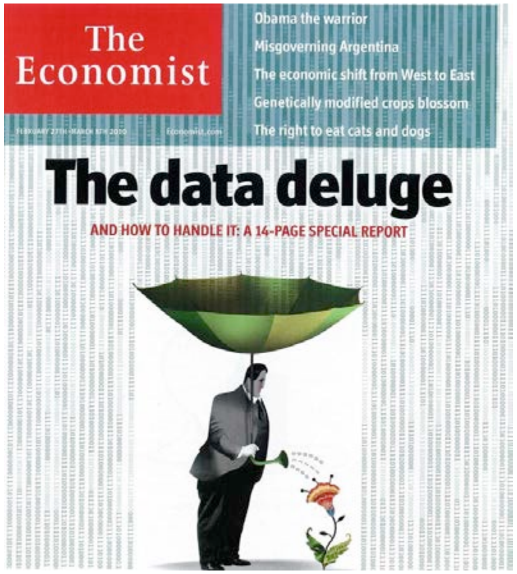
PROPHESEE

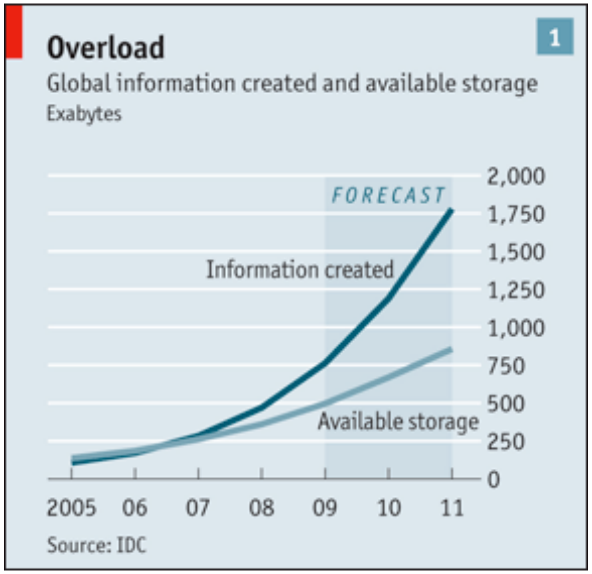

Telluride
Neuromorphic
Cognition
Engineering
Workshop
NEUROMORPHS
NEUROMORPHS
Sensors
Processors
OmniVision
iniLabs
HUAWEI
SAMSUNG
SONY
HILLHOUSE
insightness
intel
LOIHI
LEARNING WITH LESS DATA
IBM
brainchip
ακίδα
brainchip
The Neuromorphic Computing Company

Why Neuromorphic?

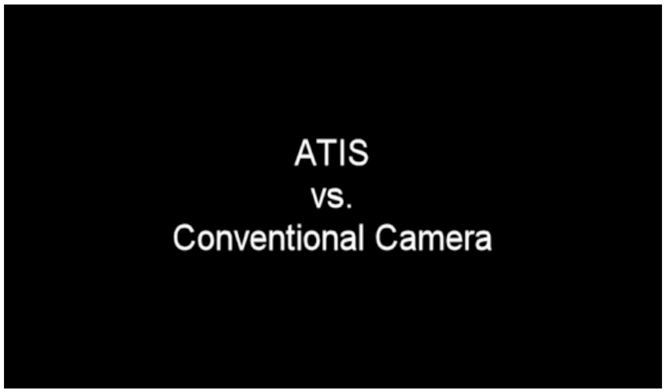
Data deluge

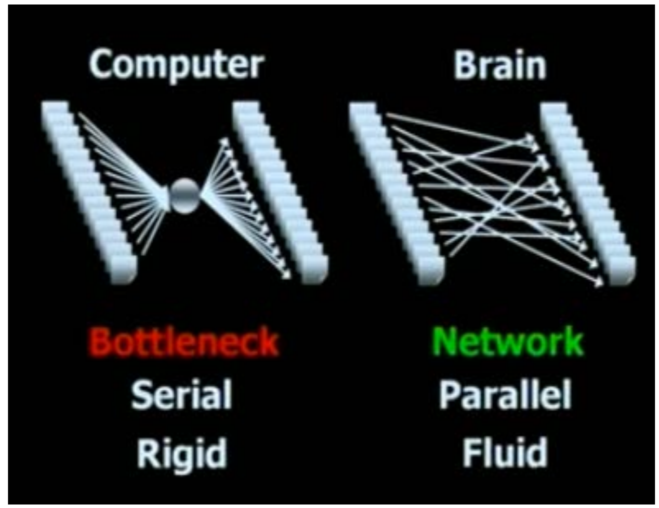
- In 2010 the world generated more than 1.2 Zetta bytes (10^21) of new data
- Equivalent of 300000km of DVD stack (<u>distance between the earth and the moon</u>)
- The amount of data increases faster than the computing power
- <u>Quest for Local computation</u>

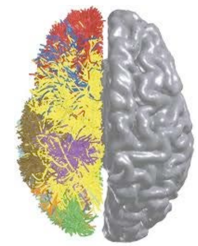




Why Neuromorphic?



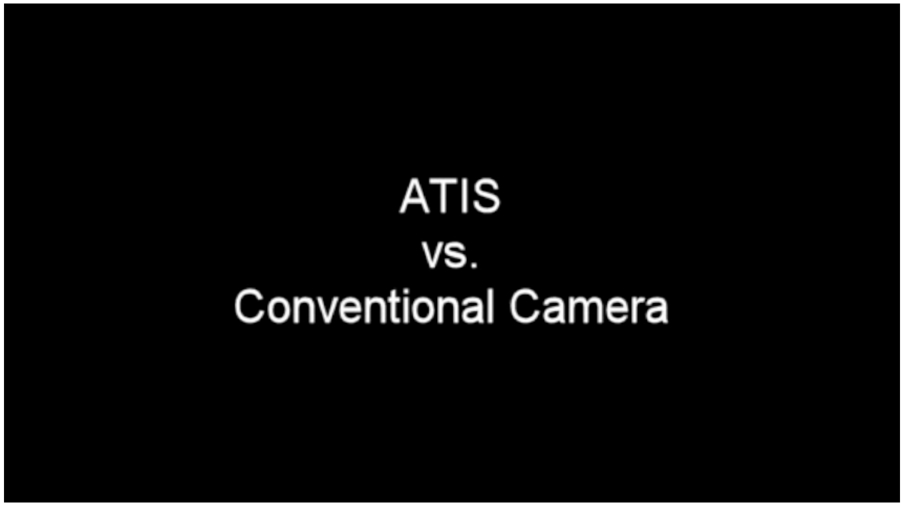


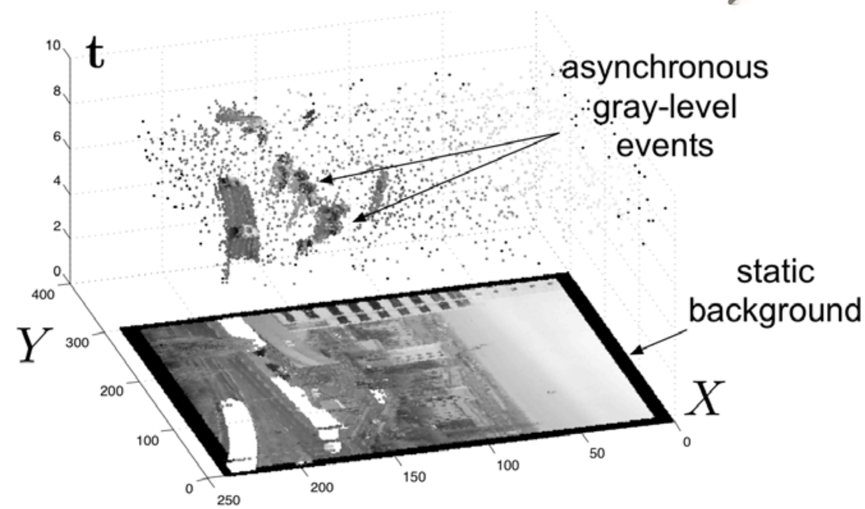


- Low Power & Low Latency
- High Temporal resolution
- Light independent
- Real time processing beyond 1KHz

- Low Power Inference
- Go beyond processor-memory bottleneck
- Real time processing beyond 1KHz

Data Space of Events

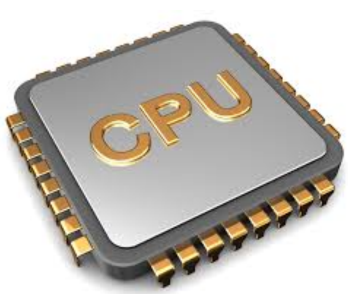




- Amplitude sampling
- Information is sent when it happens
- When nothing happens, nothing is sent or processed
- Sparse information coding
- Time is the most valuable information

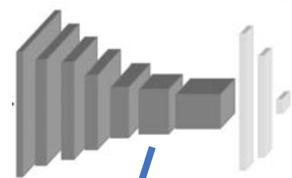
Current Alternatives

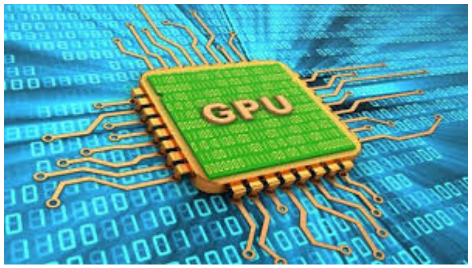
Generate frames (binary, grey level,…)



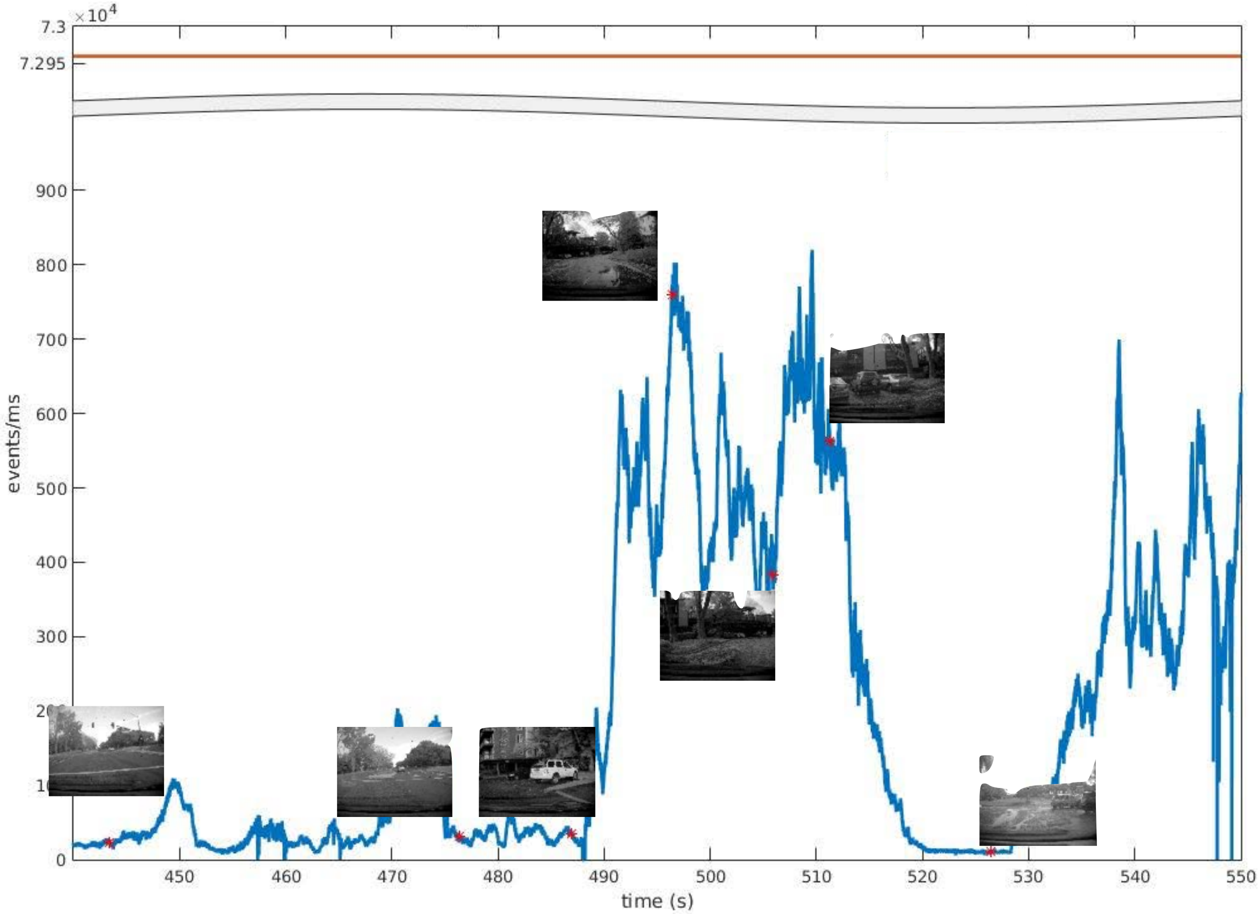
$\sum_t$

Integrate
over time

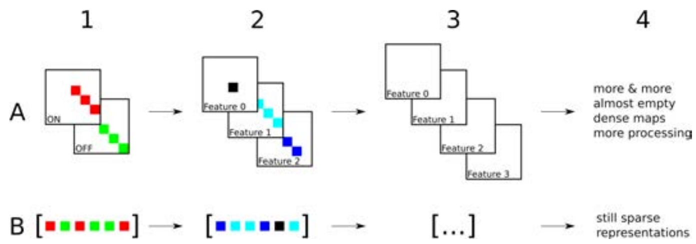


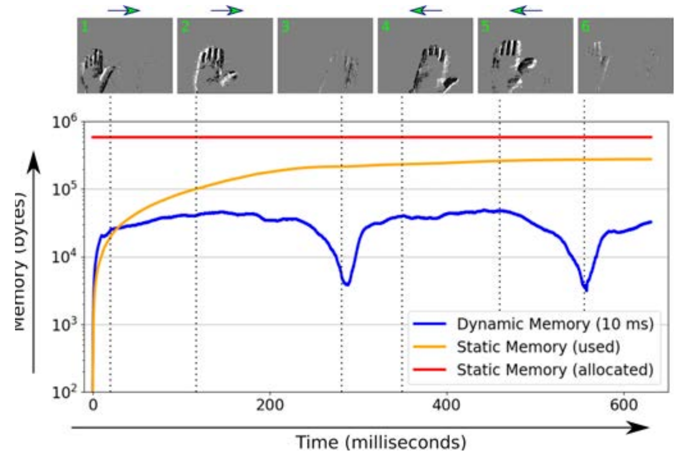


Why Event Based sensors?



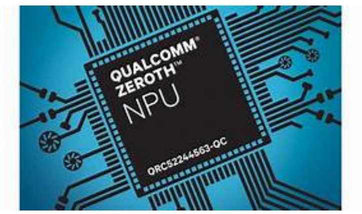
Generating Frames is inadequate





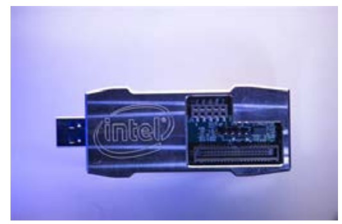
Dataset (Mean Event Rate & Sensor Size)	**PokerDVS** 170.4 ev/ms 35x35			**N-MNIST** 13.6 ev/ms 28x28			**DvsGesture** 56.9 ev/ms 128x128			**NavGesture-walk** 188.6 ev/ms 304x240		
1. Time Window (ms)	**1**	**10**	**100**	**1**	**10**	**100**	**1**	**10**	**100**	**1**	**10**	**100**
2. Mean Number 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%

Existing Neuromorphic Processing Hardware is based on silicon neurons



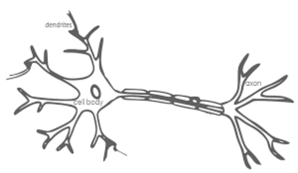
Qualcomm Zeroth (2013)

IBM TrueNorth (2014)

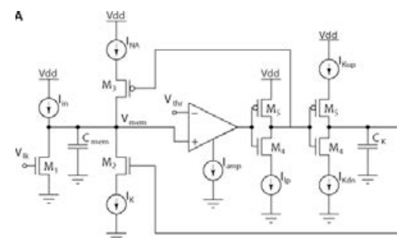


Intel Loihi (2017)

BrainChip (2019)

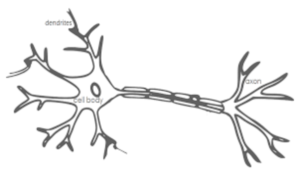


replicate

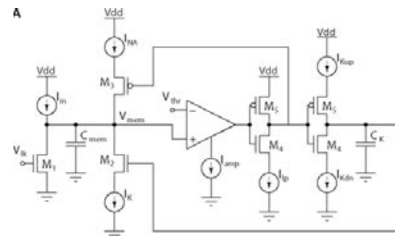


Existing hardware is based on the concept of replicating biological neurons into silicon

This approach is limited: No real theory available, wastefull in silicon area, general computation limited (materials, theories, …). We still know so little of the Brain.



replicate

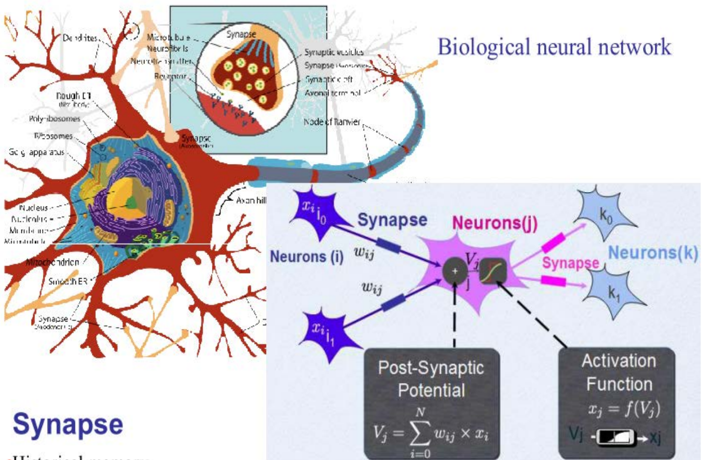


Replicating nature's solutions is not always the optimal path to solve an engineering problem.

understand

Understanding rather than replicating

Neuromorphic Computing, an old story!



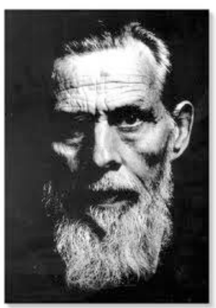
Biological neural network

Artificial neural network

Synapse

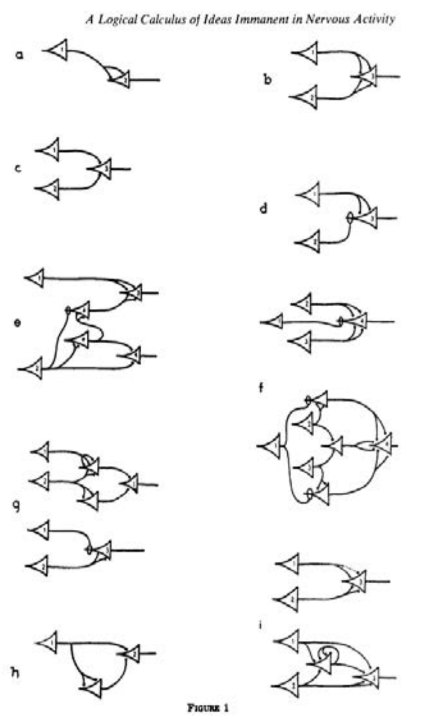
- Historical memory
- Analog levels

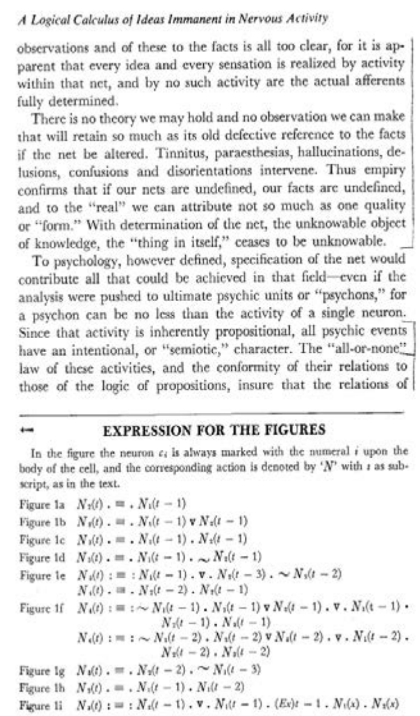
[1] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” Bull. Math. Biophysics, no. 5, pp. 115-133, **1943**.



Warren McCulloch

Walter Pitts





A Logical Calculus of Ideas Immanent in Nervous Activity

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 $N_3(t) :\equiv: N_1(t-1) \,.\vee.\, N_2(t-3) \,.\, {\sim} N_2(t-2)$
$N_4(t) \,.\!\equiv.\, N_2(t-2) \,.\, N_2(t-1)$

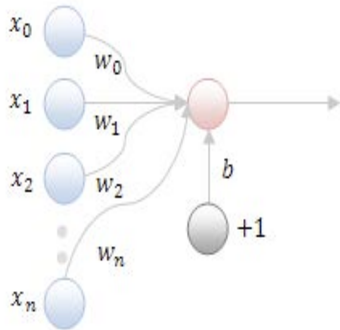
Figure 1f $N_4(t) :\equiv: {\sim} N_1(t-1) \,.\, N_2(t-1) \vee N_3(t-1) \,.\vee.\, N_1(t-1) \,.\, N_2(t-1) \,.\, N_3(t-1)$
$N_4(t) :\equiv: {\sim} N_1(t-2) \,.\, N_2(t-2) \vee N_3(t-2) \,.\vee.\, N_1(t-2) \,.\, N_2(t-2) \,.\, N_3(t-2)$

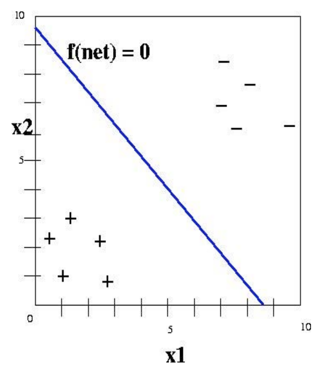
Figure 1g $N_3(t) \,.\!\equiv.\, N_2(t-2) \,.\, {\sim} N_1(t-3)$

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

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

Perceptron: first neuromorphic engine

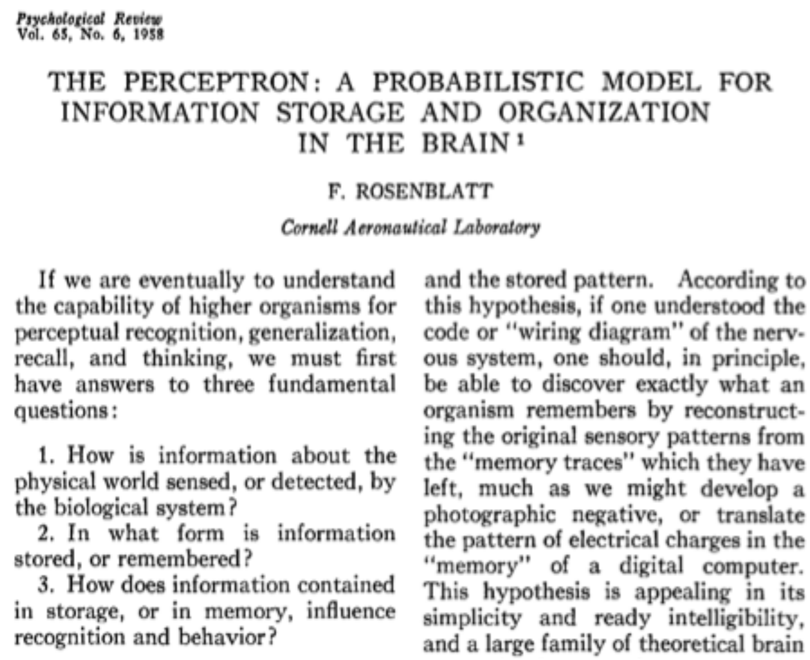




[1] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," Psychological Review, vol. 65, no. 6, pp. 386-408, **1958**.

(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

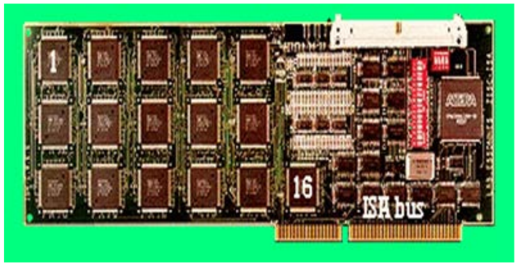
If we are eventually to understand the capability of higher organisms for perceptual recognition, generalization, recall, and thinking, we must first have answers to three fundamental questions:

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

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...



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

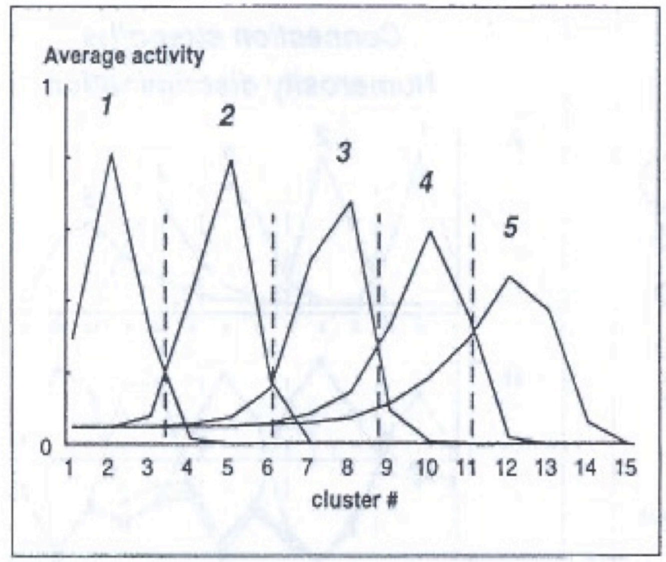
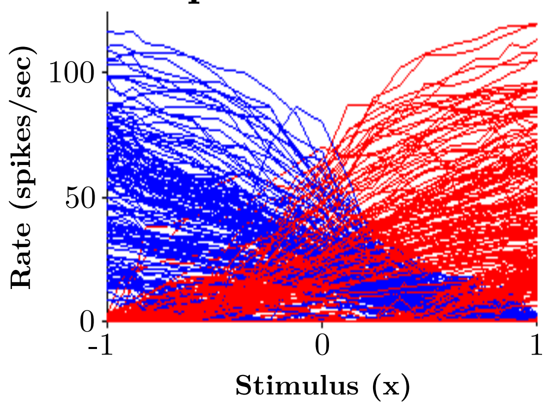


Figure 4. Average activity of numerosity clusters when random sets of 1, 2, 3, 4, or 5 objects were presented for input. For each input numerosity, only a small number of clusters were selectively activated (e.g., clusters 1, 2, and 3 responded only when a single object was presented). The activity peaks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

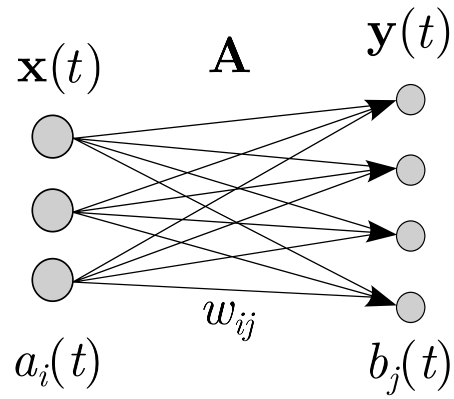
The Neural Engineering Framework

Chris Eliasmith and Charles H. Anderson

Representation

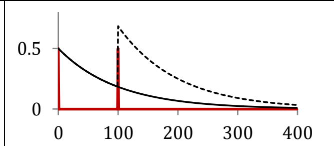
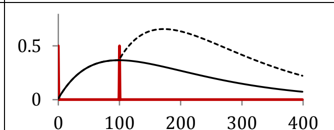
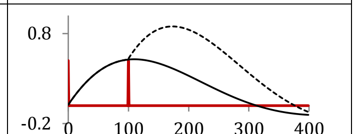
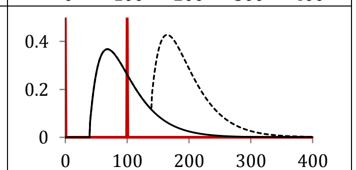
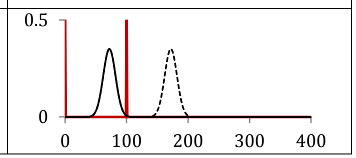


Transformation



Representation: A multi-dimensional stimulus $\mathbf{x}(t)$ is nonlinearly encoded as a spike rate $a_i(\mathbf{x}(t))$—represented by the *neuron tuning curve*—that is linearly decoded to recover an estimate of $\mathbf{x}(t)$, $\hat{\mathbf{x}}(\mathbf{t}) = \sum_i a_i(\mathbf{x}(t))\phi_i^{\mathbf{x}}$, where $\phi_i^{\mathbf{x}}$ are the decoding weights.

Variety of synaptic responses

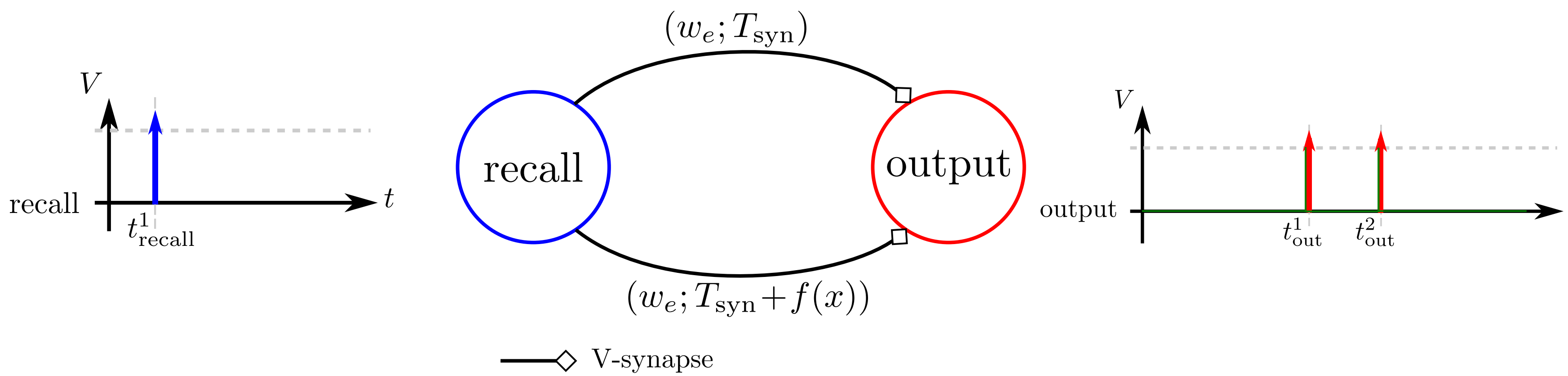
Kernel Type	Expression for Impulse Response	Typical Function (Spikes at $t = 0$, $t = 100$)
Stable recurrent connection (leaky integration) with nonlinear leak	$g(t) = \frac{1}{1+(g(t-\tau))^2}\int_{t_0}^{t}\sum_{i=1}^{L} w_{ji}^{(1)}\, x_{i,t}\ dt$	
Alpha function	$g(t) = \left[\sum_{i=1}^{L} w_{ji}^{(1)}\, x_{i,t}\right]\frac{t}{\tau}e^{-\frac{t}{\tau}}$	
Damped resonant synapse	$g(t) = \left[\sum_{i=1}^{L} w_{ji}^{(1)}\, x_{i,t}\right]e^{-\frac{t}{\tau}}\sin(\omega t)$	
Synaptic or dendritic delay with alpha function	$for\ t \geq \Delta t$: $g(t) = \left[\sum_{i=1}^{L} w_{ji}^{(1)}\, x_{i,t}\right]\frac{t-\Delta T}{\tau}e^{-\frac{t-\Delta T}{\tau}}$ $t < \Delta t$: $g(t) = 0$	
Synaptic or dendritic delay with Gaussian function	$g(t) = \left[\sum_{i=1}^{L} w_{ji}^{(1)}\, x_{i,t}\right]\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-\Delta T)^2}{2\sigma^2}}$	

Time and delays plays an important role

- Progresses in Neuroscience demonstrated the weaknesses of the perceptron approach and introduced LTP/LTD and STDP

from Markram et al. "A history of spike-timing-dependent plasticity," in *Frontiers in Synaptic neuroscience*, Vol 3, August 2011

How to encode numbers ?

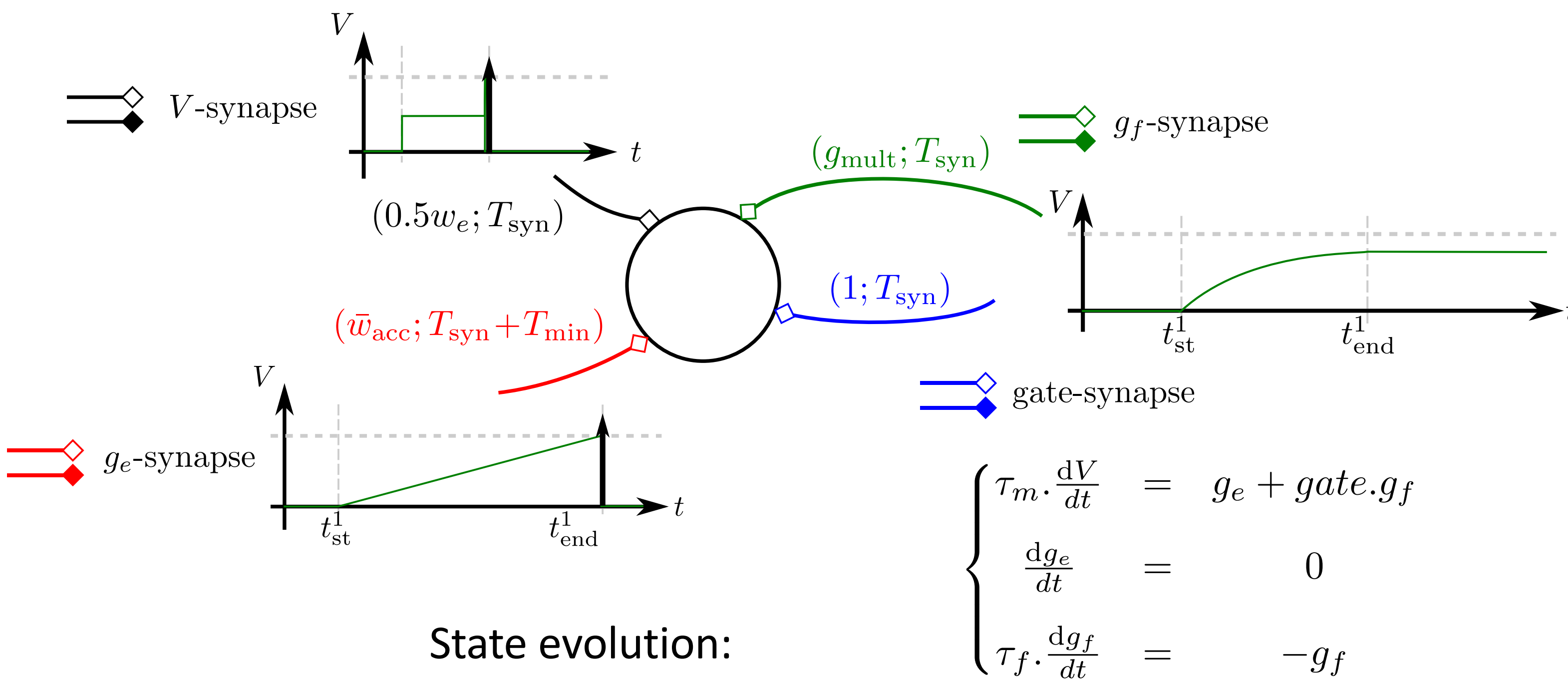


$$\Delta t = f(x) = T_{\min} + x.T_{\text{cod}}$$

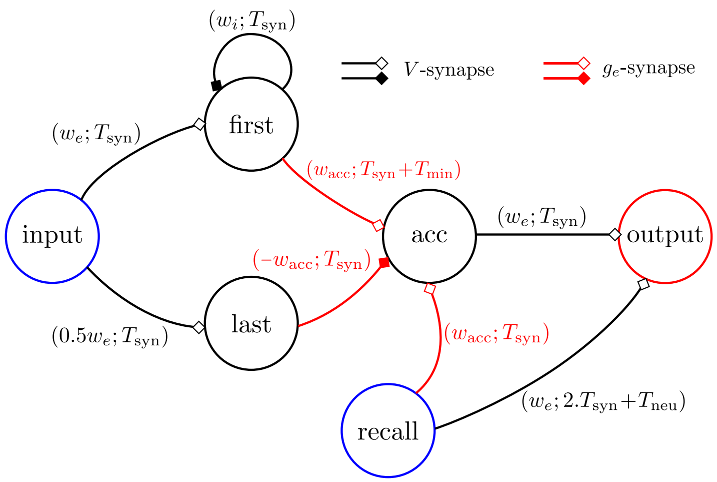
ARTICLE **Communicated by Terrence Sejnowski**

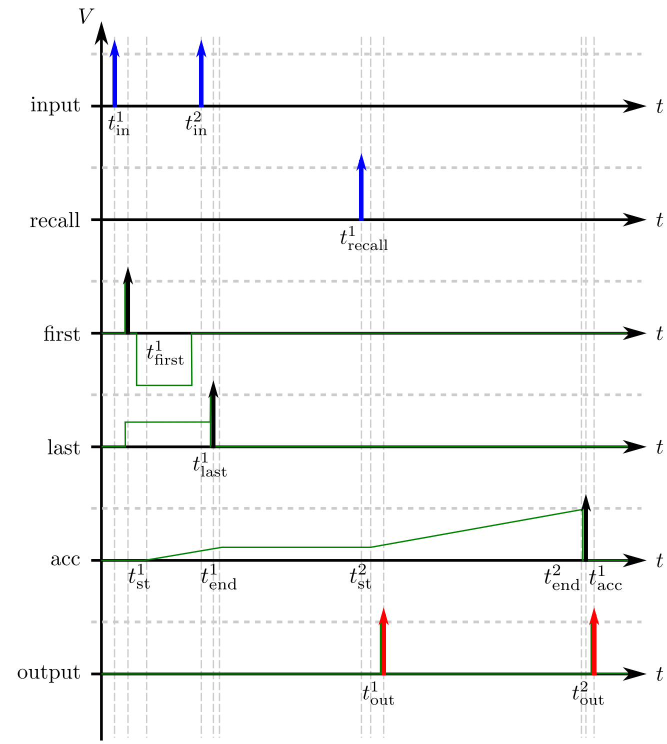
STICK: Spike Time Interval Computational Kernel, a Framework for General Purpose Computation Using Neurons, Precise Timing, Delays, and Synchrony

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



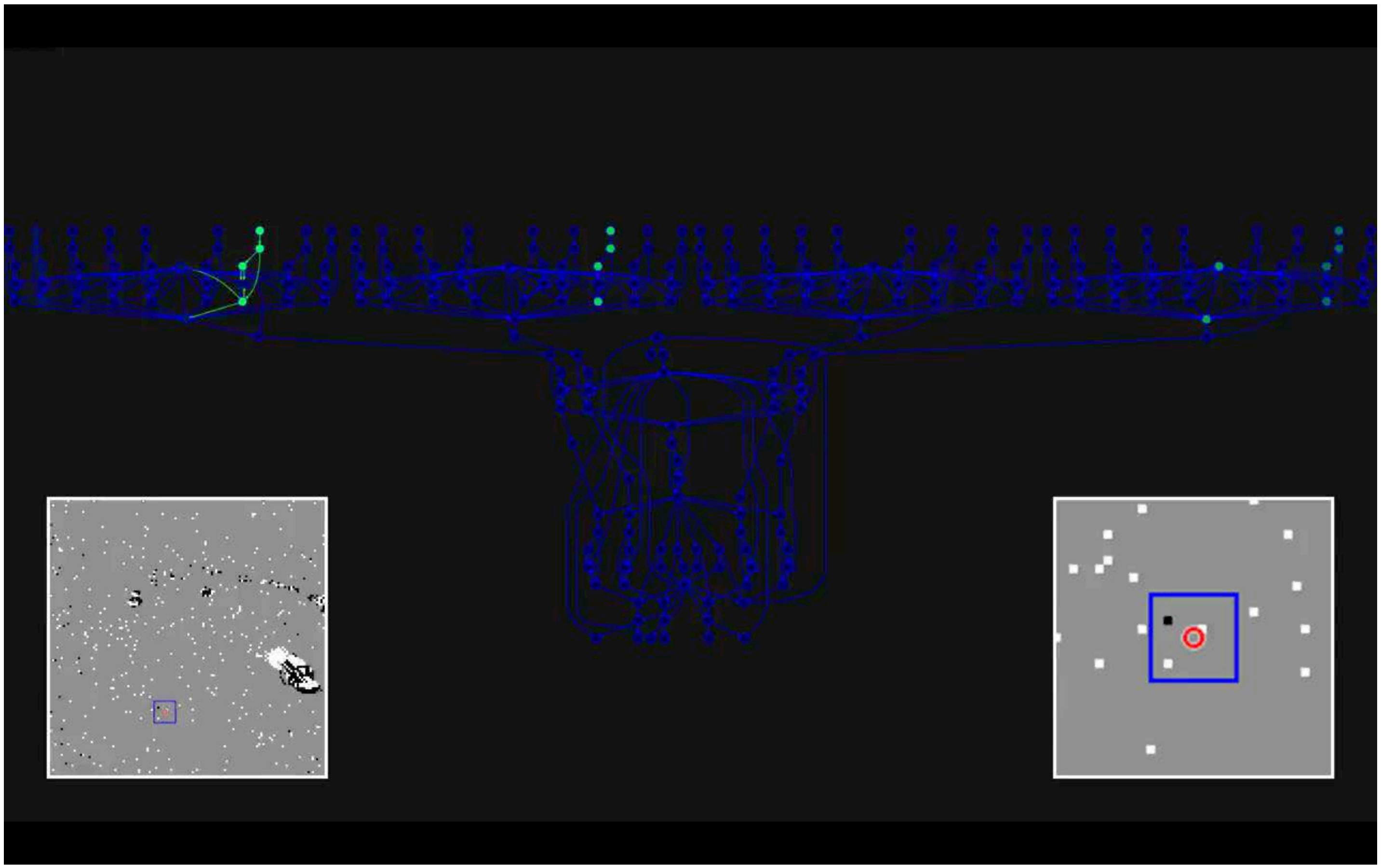
Storing information: an inverting Memory network



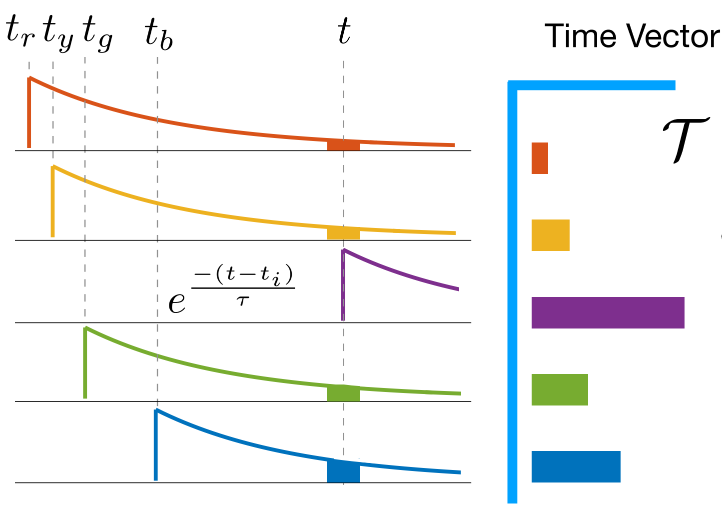


$$
\begin{cases}
\tau_m.\frac{\mathrm{d}V}{dt} = g_e + gate.g_f \\
\frac{\mathrm{d}g_e}{dt} = 0 \\
\tau_f.\frac{\mathrm{d}g_f}{dt} = -g_f
\end{cases}
$$

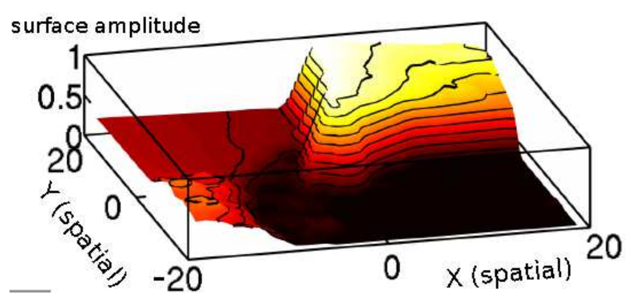
Computation of the Optical with STICK

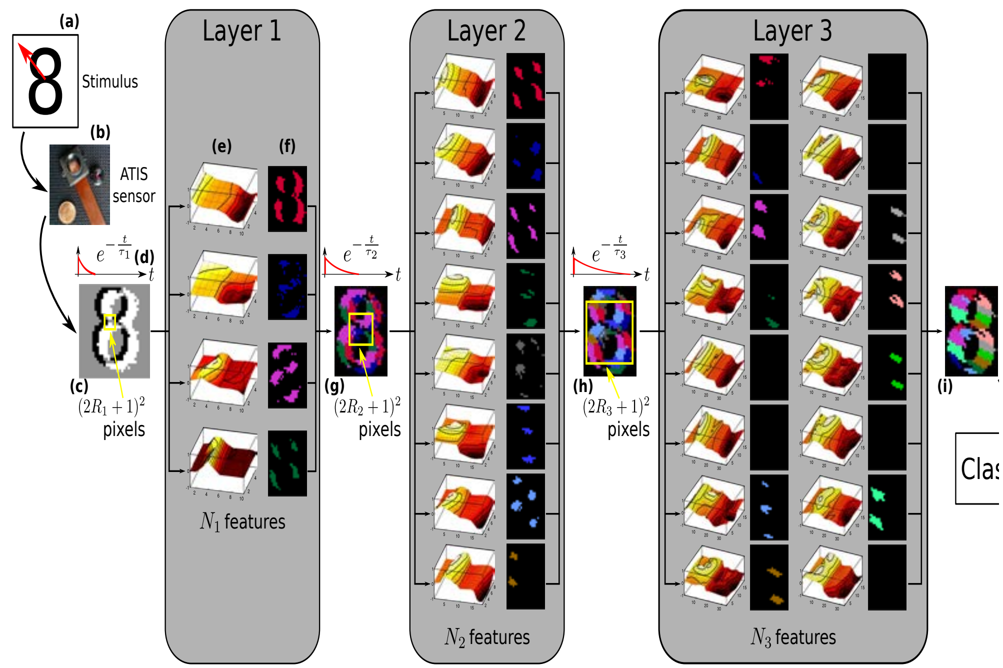


HOTS: A Hierarchy Of event-based Time-Surfaces



Time Surface or vectors





Most Work in Neuromorphic Vision relies on computing Equations

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. ?, NO. ?, 2017

Event-based, 6-DOF Camera Tracking from Photometric Depth Maps

Guillermo Gallego, Jon E.A. Lund, Elias Mueggler, Henri Rebecq, Tobi Delbruck, Davide Scaramuzza

Abstract—Event cameras are bio-inspired vision sensors that output pixel-level brightness changes instead of standard intensity frames. These cameras do not suffer from motion blur and have a very high dynamic range, which enables them to provide reliable visual information during high-speed motions or in scenes characterized by high dynamic range. These features, along with a very low power consumption, make event cameras an ideal complement to standard cameras for VR/AR and video game applications. With
accurate, low-latency tracking of an event camera from an existing
built via classic dense reconstruction pipelines. Our approach tracks the
nt, thus virtually eliminating latency. We successfully evaluate the method
of the technological advantages of the event camera—our pipeline works
unaccessible to standard cameras.

ision Sensor, Bayes filter, Asynchronous processing, Conjugate priors,

Fig. 1: Sample application: 6-DOF tracking in AR/VR (Aug-

Home / Journals / IEEE Transactions on Pattern Analysis and Machine Intelligence / Preprints

Real-time high speed motion prediction using fast aperture-robust event-driven visual flow

PrePrints pp. 1-1,
DOI Bookmark: 10.1109/TPAMI.2020.3010468

Keywords

Visualization, Cameras, Heuristic Algorithms, Vehicle Dynamics, Apertures, Robot Sensing Systems, Event Driven, Neuromorphic, Optical Flow, Motion Prediction

Authors

Himanshu Akolkar, Department of Ophthalmology, University of Pittsburgh, 6614 Pittsburgh, Pennsylvania United States (e-mail: akolkar@pitt.edu)
Sio Hoi Ieng, ISIR, University Pierre et Marie Curie, Paris, Paris France 75252 (e-mail: sio-hoi.ieng@upmc.fr)
Ryad Benosman, ISIR, University Pierre et Marie Curie, Paris, Paris France 75252 (e-mail: ryad.benosman@upmc.fr)

Abstract

Optical flow is a crucial component of the feature space for early visual processing of dynamic scenes especially in new applications such as self-driving vehicles, drones and autonomous robots. The dynamic vision sensors are well suited for such applications because of their asynchronous, sparse and temporally precise representation of the visual dynamics. Many algorithms proposed for computing visual flow for these sensors suffer from the aperture problem as the direction of the estimated flow is governed by the curvature of the object rather than the true motion direction. Some methods that do overcome this problem by temporal windowing under-utilize the true precise temporal nature of the dynamic sensors. In this paper, we propose a novel multi-scale plane fitting based visual flow algorithm that is robust to the aperture problem and also computationally fast and efficient. Our algorithm performs well in many scenarios ranging from fixed camera recording simple geometric shapes to real world scenarios such as camera mounted on a moving car and can successfully perform event-by-event motion estimation of objects in the scene to allow for predictions of up to 500 ms i.e. equivalent to 10 to 25 frames with traditional cameras.

Motion Equivariant Networks for Event Cameras with the Temporal Normalization Transform

Alex Zihao Zhu [1] Ziyun Wang [1] Kostas Daniilidis [1]

Abstract

In this work, we propose a novel transformation for events from an event camera that is equivariant to optical flow under convolutions in the 3-D spatiotemporal domain. Events are generated by changes in the image, which are typically due to motion, either of the camera or the scene. As a result, different motions result in a different set of events. For learning based tasks based on a static scene such as classification which directly use the events, we must either rely on the learning method to learn the underlying object distinct from the motion, or to memorize all possible motions for each object with extensive data augmentation. Instead, we propose a novel transformation of the input event data which normalizes the x and y positions by the timestamp of each event. We show that this transformation generates a representation of the events that is equivariant to this motion when the optical flow is constant, allowing a deep neural network to learn the classification task without the need for expensive data augmentation. We test our method on the event based N-MNIST dataset, as well as a novel dataset N-MOVING-MNIST, with significantly more variety in motion compared to the standard N-MNIST dataset. In all

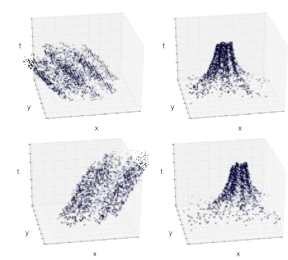


Figure 1. Classical convolution layers would not be equivariant to event motions on the left, since they are shear deformations of the event volume. After transforming to canonical coordinates on the right, the volume translates uniformly, resulting in equivariance to the motion. Left: Raw input events. Right: Corresponding transformed events.

latency, high dynamic range, and low power consumption. These benefits provide a compelling reason to utilize these cameras in traditional vision tasks such as image classifi-

The future is ours … but, we need:

- Allow to execute Learning and General Computation (equations)

- Allow for incremental processing ensuring a fast access for each incoming event to local resources

- The retrieval of relevant local information around incoming events (access times) to match the high temporal properties of event-based cameras and ensure computation can be carried out at the native elementary temporal step of event-based cameras (1μs)

- Allow for sparse memory use following the scene-driven properties of event-based cameras and temporal requirements of the used incremental algorithms.

- Meet the current urgent need to handle >5 Giga Events/second at few mW