



HAL
open science

Neuromorphic Event-Based Spatio-temporal Attention using Adaptive Mechanisms

Amélie Gruel, Antonio Vitale, Jean Martinet, Michele Magno

► **To cite this version:**

Amélie Gruel, Antonio Vitale, Jean Martinet, Michele Magno. Neuromorphic Event-Based Spatio-temporal Attention using Adaptive Mechanisms. International Conference on Artificial Intelligence Circuits and Systems (AICAS), Jun 2022, Incheon, South Korea. . hal-03760704

HAL Id: hal-03760704

<https://hal.science/hal-03760704>

Submitted on 25 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Neuromorphic Event-Based Spatio-temporal Attention using Adaptive Mechanisms

Amélie Gruel¹

Antonio Vitale²

Jean Martinet¹

Michele Magno²

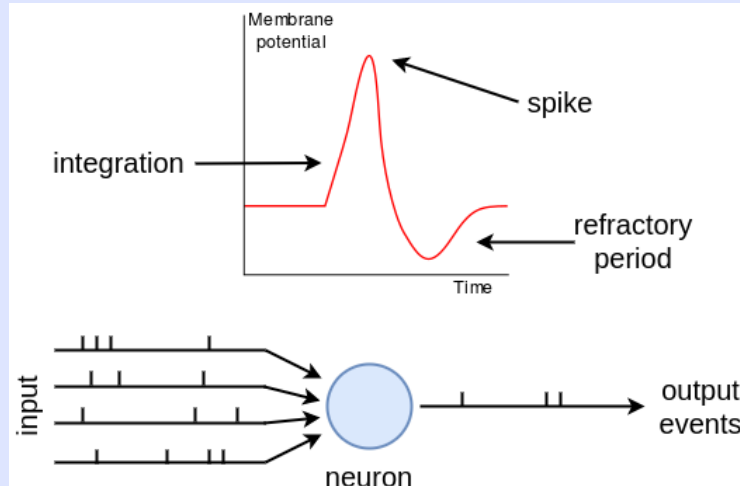
¹Université Côte d'Azur, CNRS, I3S - France

²Center for Project-Based Learning, ETH Zürich - Switzerland

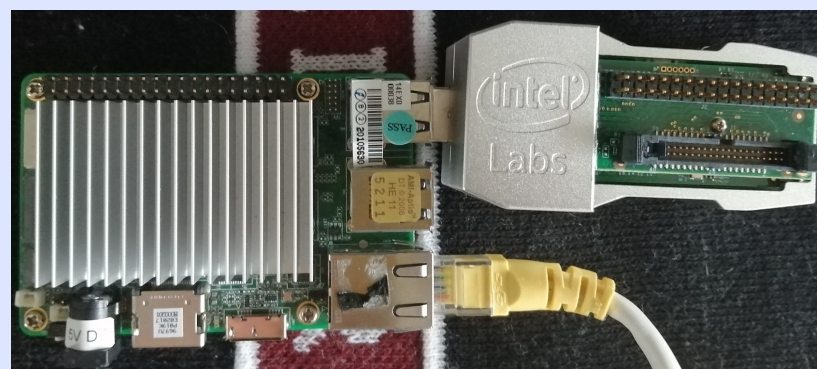
This work was supported by the European Union's ERA-NET CHIST-ERA 2018 research and innovation programme under grant agreement ANR-19-CHR3-0008.

Visual attention: behavioral and cognitive process of selectively focusing on a discrete aspect of sensory cues

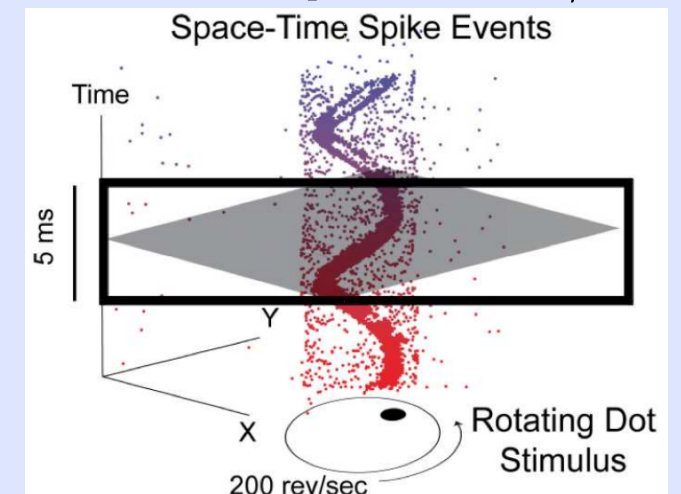
Spiking neural network [Maass, 1997]



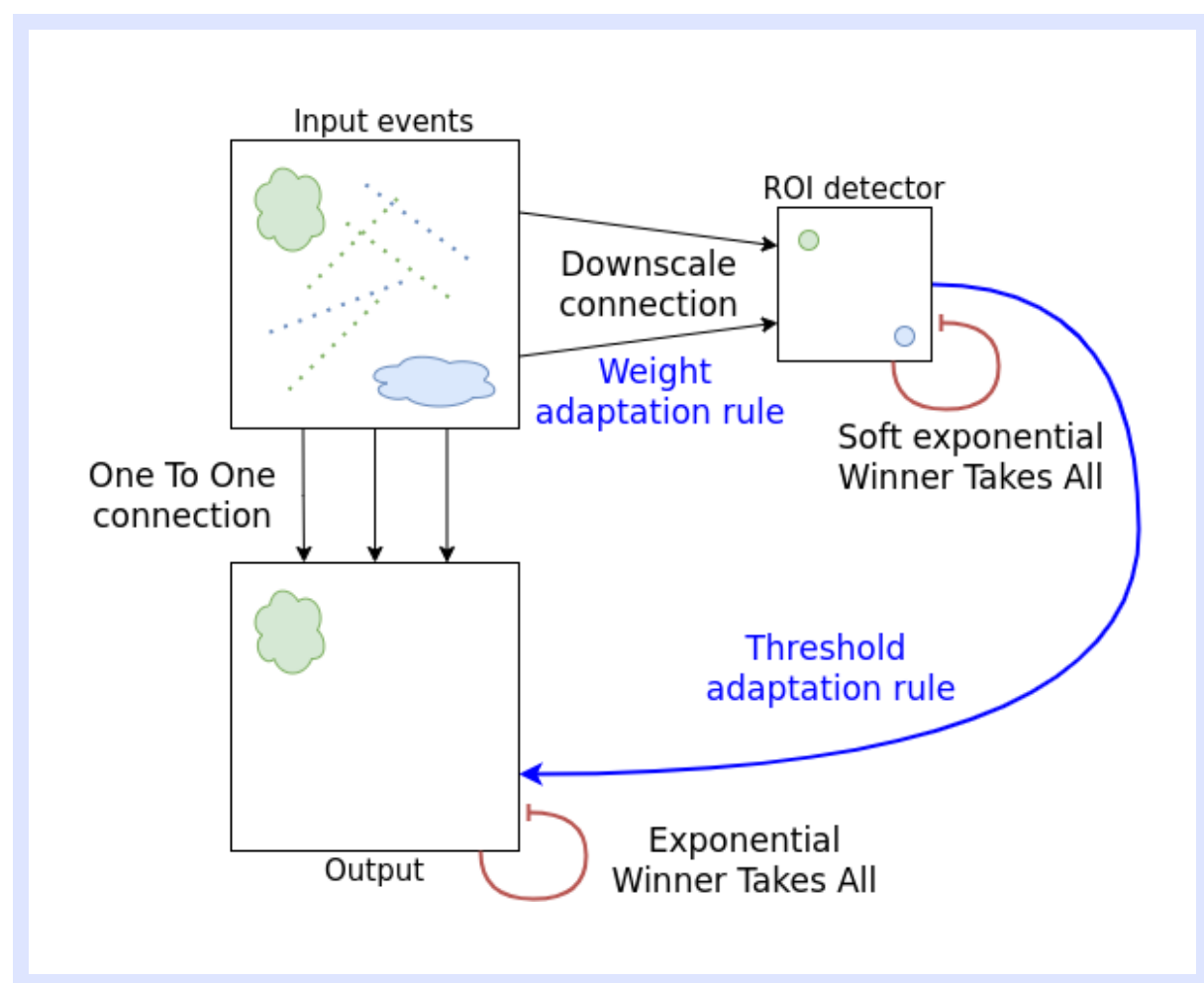
Intel's Loihi [Davies, 2021]



Event camera [Lichtsteiner, 2008]



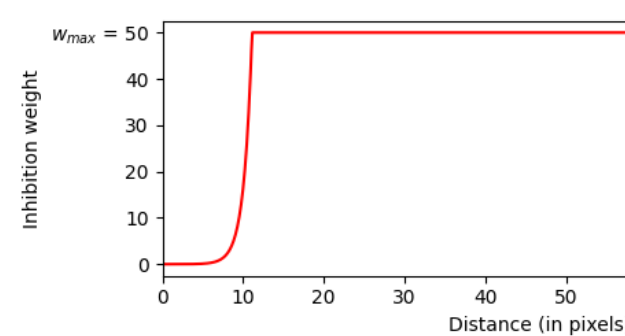
Neuromorphic model



2 adaptive mechanisms:

- dynamic weight adaptation rule inspired by [Hebb, 1949]
- dynamic threshold adaptation rule

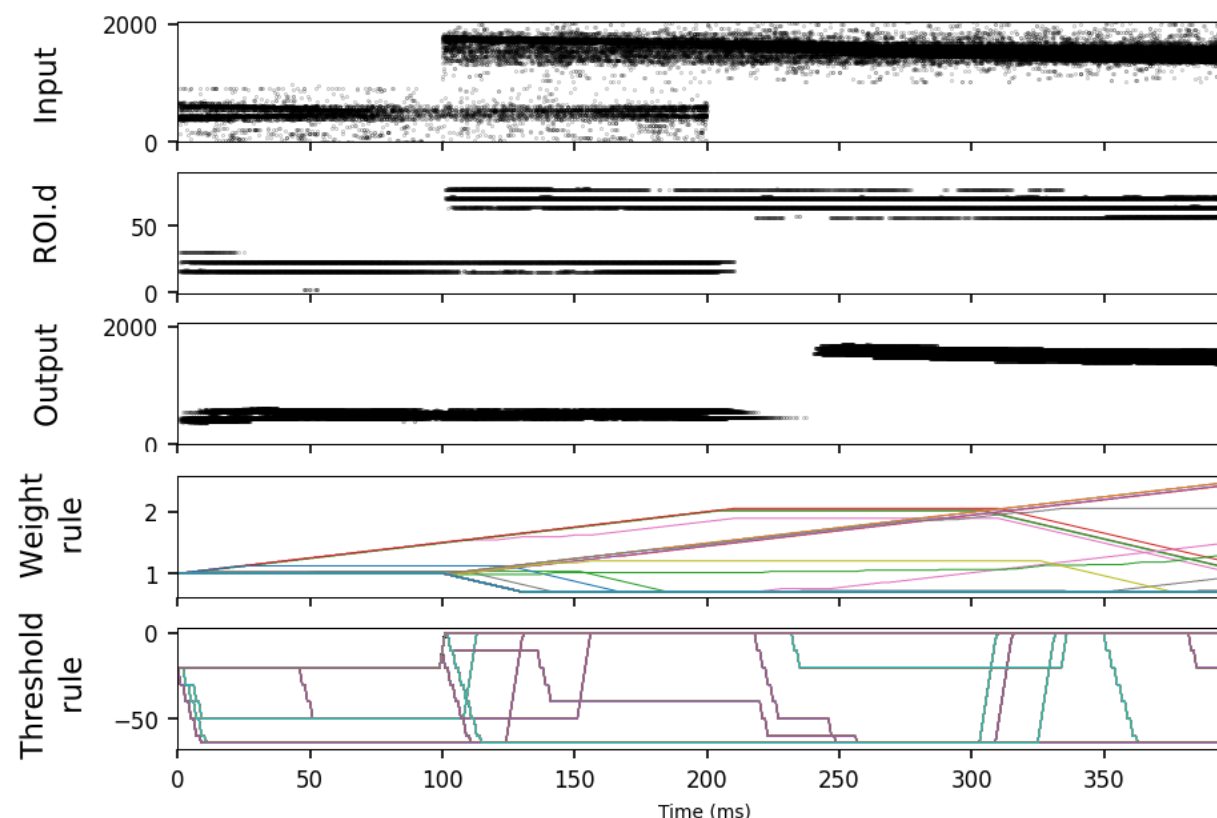
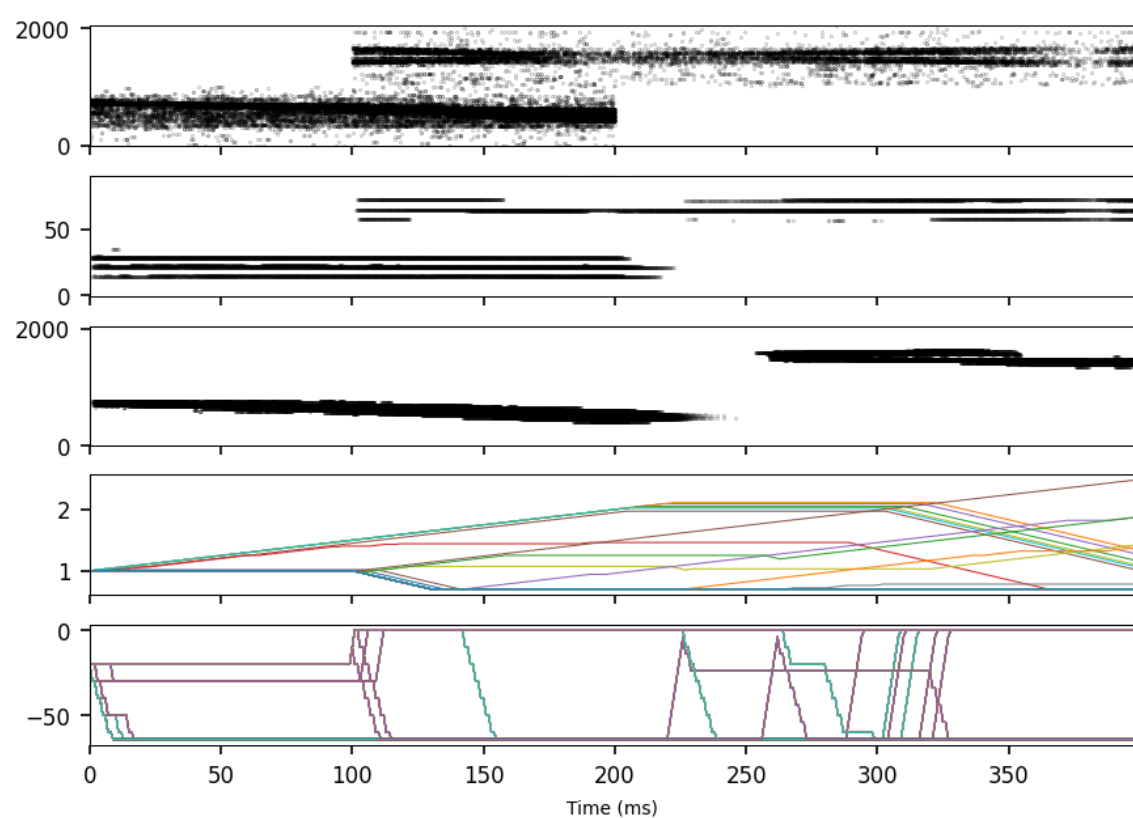
Exponential Winner-Takes-All



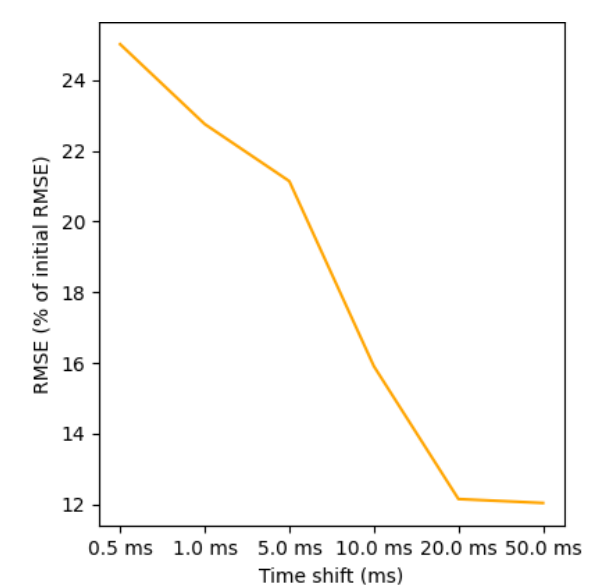
$$W_{WTA} = \max\left(\frac{e^d}{w \times h}, w_{max}\right)$$

- soft WTA on ROI detector to aggregate active regions in distinct segments
- strong WTA to filter out one segment at the time

Experimental results



Simulation of the attention model on PyNN & NEST, applied to two DVS 128 Gesture [Amir, 2017] samples with a time shift of 100 ms. Input data on left corresponds to large gesture first, small gesture second; and vice-versa on the left.



Evolution of the RMSE error as the time shift between two samples increases.

References

- [Maass, 1997] W. Maass (1997). Networks of Spiking Neurons: The third generation of neural network models. *Neural Networks*.
- [Davies, 2021] M. Davies et al. (2021). Advancing neuromorphic computing with loihi: A survey of results and outlook. *Pr. of the IEEE*, vol. 109.
- [Lichtsteiner, 2008] P. Lichtsteiner, C. Posch and T. Delbruck (2008). A 128x128 120 dB 15 us Latency Asynchronous Temporal Contrast Vision Sensor. *IEEE Journal of Solid-State Circuits*.
- [Hebb, 1949] D. Hebb (1949). The organization of behavior: A neuropsychological theory. *Journal of the American Medical Association*, vol. 143.
- [Amir, 2017] A. Amir et al. (2017). A Low Power, Fully Event-Based Gesture Recognition System. *CVPR*.

Simulator	Intel's Loihi	PyNN & NEST
Simulation time	6ms to process 1ms → Ratio: 1:6	3h26min58s to process 400ms → Ratio: 1:12,505
Adaptation rules ?	None	2 adaptation rules
Large gesture first, small gesture second	Avg RMSE = 12.92	Avg RMSE = 0.03
Small gesture first, large gesture second	Avg RMSE = 45.88	Avg RMSE = 0.16