

Spike timing-based unsupervised learning of orientation, disparity, and motion representations in a spiking neural network.

Thomas Barbier, Céline Teulière, Jochen Triesch

Université Clermont-Auvergne, CNRS, SIGMA Clermont, Institut Pascal, Frankfurt Institute for Advanced Studies.

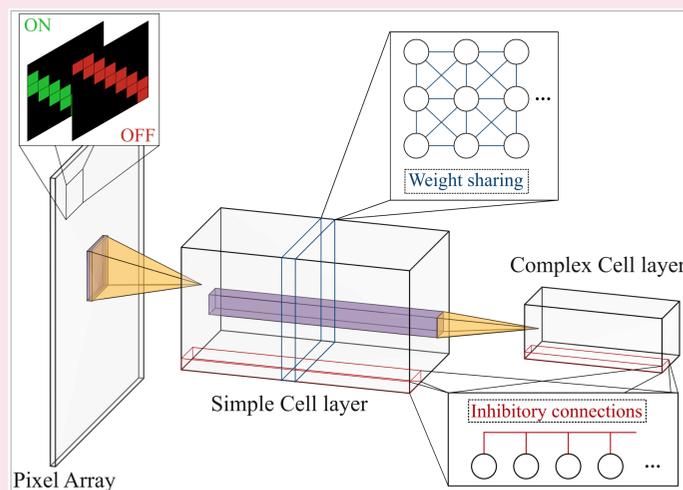


Method

- We present a network of leaky integrate and fire (LIF) neurons that learns representations similar to those of simple and complex cells in the primary visual cortex of mammals.
- Based on a wide range of homeostatic mechanisms, such as refractory periods (RP), spike rate adaptations (SRA) or lateral inhibitions.
- A (LIF) neuron membrane potential $V(t)$ variation in time can be summed up as:

$$\tilde{V}(t + \Delta t) = V(t) e^{-\frac{\Delta t}{\tau_m}} - V_{SRA}(t) e^{-\frac{\Delta t}{\tau_{SRA}}} - \eta_{RP} e^{-\frac{t_s - t - \Delta t}{\tau_{RP}}} + w_i(t)$$

with w_i as the synaptic inputs, and the other terms as membrane potential regulators.

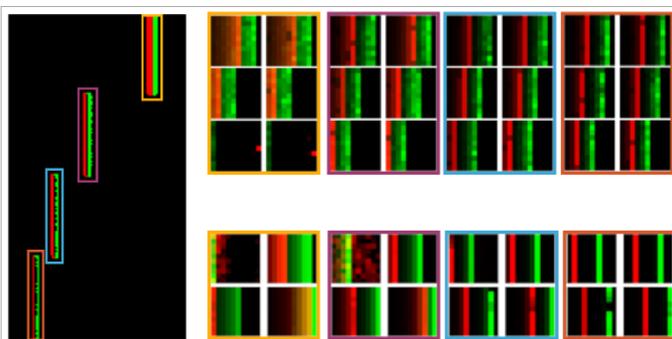


Spiking neural network diagram

- The network learns using the Spike Timing-Dependent Plasticity (STDP) rule.

Learning orientation, motion and disparity

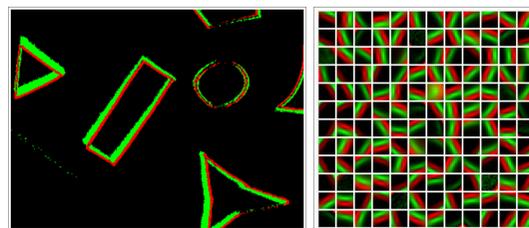
- The network learns visual feature detectors for orientation, disparity, and motion in a fully unsupervised fashion.



Simple cell receptive fields learned on vertical bars (left image) moving at varying speed. top right image are receptive fields learned with 3 multi-synaptic delayed synapses, the bottom right image represents receptive fields learned with a pair of stereoscopic cameras. The shift respectively represents the speed and disparity of the moving bars.

Evaluating the network cell basis

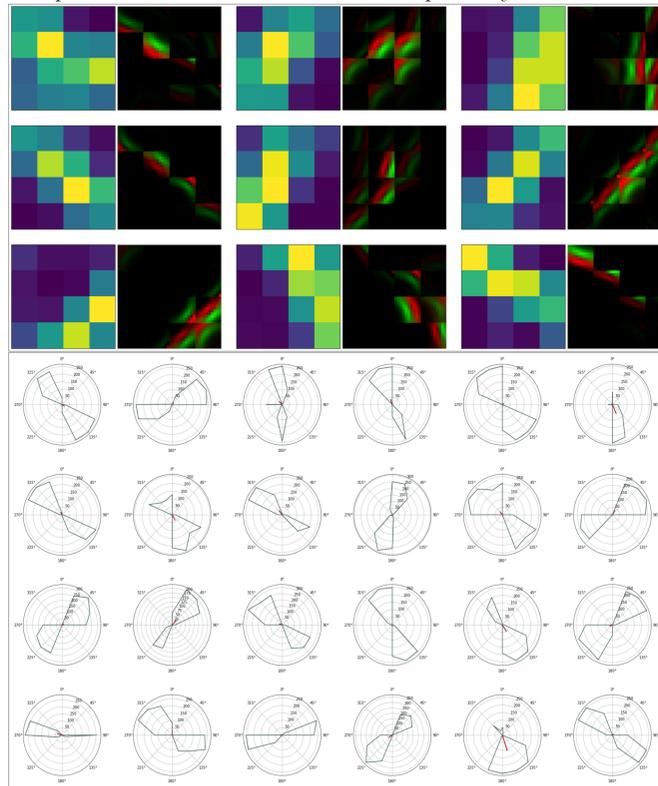
- Using real event videos of various moving shapes, we learn a diverse basis of simple and complex cell receptive fields.
- Learned neuronal representations resemble closely biological characteristics, such as gabor functions for the simple cells.
- By showing diverse inputs, the network is able to learn efficient feature detectors.



Screenshot of an event video of various moving shapes (left). Resulting simple cell learned receptive fields.

Measuring the orientation selectivity

- We measure the complex cell responses to oriented grating stimulus from their average spiking activity.
- We demonstrate that after learning, simple and complex cells become very selective to specific orientations and directions respectively.



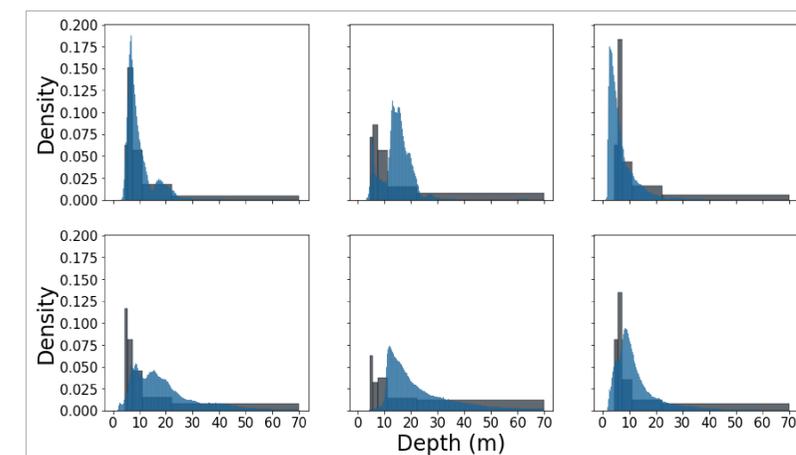
Top: complex cells representations made from an average (left, heatmap) and the maximum (right, receptive fields) of the simple cell connection strengths. Bottom: complex cell average responses in direction space measured with rotating oriented grating stimulus.

Estimating depth

- We performed scene acquisitions from a stereoscopic pair of event cameras mounted on a robotic mobile platform in an urban environment.
- We estimate the average scene depth for specific regions of the visual field from the shift between simple cell stereo receptive fields. We compare it to a traditional frame-based depth analysis.
- This demonstrate the network ability to learn disparity representations in complex stereo environments.



Robotic platform.



Comparison of the depth estimation histograms computed with traditional frames (blue) and the learned receptive fields of the simple cells (grey).

Future work

- Our network is well-suited for implementation on modern neuromorphic spiking network hardware.
- We would like to extend this work to the autonomous self-calibration of spike-based active binocular vision systems.

References

- [1] P. Stefano A. Himanshu and B. Chiara. Spike time based unsupervised learning of receptive fields for event-driven vision. *ICRA*, 2015.
- [2] A. Wendt L. Paulun and N. Kasabov. A retinotopic spiking neural network system for accurate recognition of moving objects using NeuCube and dynamic vision sensors. *Frontiers in Computational Neuroscience*, 2018.
- [3] K. Scheper F. Paredes-Valles and G. De Croon. Unsupervised learning of a hierarchical spiking neural network for optical flow estimation: From events to global motion perception. *PAMI*, 2018.