Emergence of Gabor-like Receptive Fields in a Recurrent Network of Mixed-Signal Silicon Neurons

Valentina Baruzzi
Dept. of Computer Science,
Bioengineering, Robotics and Systems Eng.
University of Genoa
Genoa, Italy
Email: baruzzivalentina@gmail.com

Giacomo Indiveri
Institute of Neuroinformatics
University of Zürich and
ETH Zürich
Zurich, Switzerland
Email: giacomo@ini.uzh.ch

Silvio P. Sabatini
Dept. of Computer Science,
Bioengineering, Robotics and Systems Eng.
University of Genoa
Genoa, Italy
Email: silvio.sabatini@unige.it

***Abstract*—Mixed signal analog/digital neuromorphic circuits offer an ideal computational substrate for testing and validating hypotheses about models of sensory processing, as they are affected by low resolution, variability, and other limitations that affect in a similar way real neural circuits. In addition, their real-time response properties allow to test these models in closed-loop sensory-processing hardware setups and to get an immediate feedback on the effect of different parameter settings. Within this context we developed a recurrent neural network architecture based on a model of the retinocortical visual pathway to obtain neurons highly tuned to oriented visual stimuli along a specific direction and with a specific spatial frequency, with Gabor-like receptive fields. The computation performed by the retina is emulated by a Dynamic Vision Sensor (DVS) while the following feed-forward and recurrent processing stages are implemented by a Dynamic Neuromorphic Asynchronous Processor (DYNAP) chip that comprises adaptive integrate-and fire neurons and dynamic synapses. We show how the network implemented on this device gives rise to neurons tuned to specific orientations and spatial frequencies, independent of the temporal frequency of the visual stimulus. Compared to alternative feed-forward schemes, the model proposed produces highly structured receptive fields with a limited number of synaptic connections, thus optimizing hardware resources. We validate the model and approach proposed with experimental results using both synthetic and natural images.**

I. Introduction

Todays's neuromorphic systems represent a promising alternative to von Neumann architectures in terms of power efficiency, computational flexibility, and robustness. In particular, reproducing the dynamics of biological neural systems using sub-threshold analog circuits and asynchronous digital ones make these systems ideal platforms to implement low-power bio-inspired devices for a wide range of application domains [1] [2]. Despite these principled assets, neuromorphic system design has to cope with the limited resources presently available on hardware. In this paper, we propose an economic way to implement spike-based early-vision detectors of oriented features in given spatial frequency bandwidths that mimic the known properties of Gabor-like simple cells receptive fields (RFs) in the primary visual cortex (V1) [3]. The computation performed by the biological retina is emulated by an asynchronous event-driven Dynamic Vision Sensor (DVS) [4], which only indicates luminance temporal changes in the image impinging on the photodiode array. Its output feeds a neuromorphic processor (DYNAP) [5] with re-configurable silicon neurons that can be interconnected to design real-time spiking neural networks (SNNs). Specifically, a recurrent clustered inhibition scheme is used to obtain the desired receptive field functionality. Different architectures for cortical-like orientation selective image filtering have been proposed (e.g., [6] [7] [8]) as potential building blocks for biologically inspired artificial perception systems (e.g., [9] [10] [11]). Though, most of them followed a highly custom design process, which typically relies on dedicated continuous time analog processing circuits that eventually pass the output to a neuron circuit for ultimate spike-train encoding. The work here presented is a step further in that direction, as we directly process event-based input stream and the architecture is implemented with high efficiency on a general-purpose reconfigurable neuromorphic processor.

II. Materials and Methods

Feedback is often proposed as a mechanism for refining the orientation and the spatial-frequency tuning of simple cells. In particular, in a previous work [12], we showed how the (linear) superposition of a retinocortical (i.e., feed-forward) oriented bias and a recurrent (i.e., feed-back) cross-orientation inhibition can give rise to highly structured Gabor-like receptive profiles when inhibition arises from laterally distributed clusters.

The network is conceptually composed of two layers that represent two homogeneous populations of retinal and cortical (i.e., V1) neurons, respectively. Accordingly, the excitation e of a V1 neuron with orientation preference θ in the spatial position $\mathbf{x} = (x, y)$ can be modeled as the solution of:

$$e(\mathbf{x}) = a(h_0 * s)(\mathbf{x}) - b(w * e)(\mathbf{x}) \tag{1}$$

978-1-7281-3320-1/20/$31.00 ©2020 IEEE

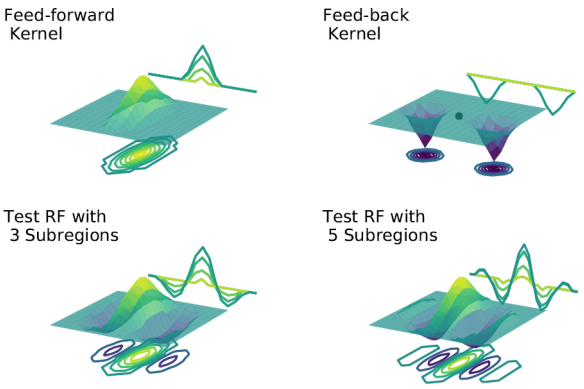


Fig. 1. (Top) Profiles of the feed-forward (h_0) and the feed-back (w) kernels; the black dot represents the target neuron of the V1 layer that receives the feed-forward excitation from the retina layer and the recurrent clustered inhibition from other neurons of the V1 layer. The connection scheme replicates identical for every neuron of the V1 layer. (Bottom) The profiles of the receptive fields obtained directly by feed-forward connections, only.

where $*$ denotes the spatial convolution operator, $s(\mathbf{x})$ is the visual stimulus, a is the strength of the feed-forward contribution, and b the strength of inhibition. The feed-forward kernel is modeled as an elongated Gaussian:

$$h_0(x_\theta, y_\theta) = \frac{1}{2\pi p \sigma_h^2} \cdot \mathrm{e}^{-\frac{x_\theta^2/p^2 + y_\theta^2}{2\sigma_h^2}}, \tag{2}$$

and the inhibitory interaction kernel is modeled by two displaced Gaussians along the direction orthogonal to the major axis of the feed-forward kernel:

$$w(x_\theta, y_\theta) = \frac{1}{2\pi\sigma_k^2} \left(\mathrm{e}^{-\frac{(x_\theta+d)^2+y_\theta^2}{2\sigma_k^2}} + \mathrm{e}^{-\frac{(x_\theta-d)^2+y_\theta^2}{2\sigma_k^2}} \right) \tag{3}$$

where σ_h and p set the size and elongation of the feed-forward kernel, σ_k and d set the size and distance of the clustered inhibition, and (x_θ, y_θ) are the rotated spatial coordinates. The resulting receptive field $h(x_\theta, y_\theta)$ can be approximated by a Gabor function, characterized by a radial peak frequency k_0 and spatial extension σ:

$$h(x_\theta, y_\theta) = \mathrm{e}^{-\frac{x_\theta^2+y_\theta^2}{2\sigma^2}} \cos k_0 x_\theta. \tag{4}$$

The shapes of the kernels are shown in Fig. 1.

A. The simulated spiking neural network

The spiking network has been designed following the architectural principles described above, yet taking into account some major differences, such as the problem of discrete quantities and finite neurons' populations, and the separate presence of ON and OFF channels. The neural network functionality has been first tested using `Brian2` [13], a simulator for SNNs written in the Python programming language, and its toolbox `teili` [14] that implements the DPI neuron model, a variation of the adaptive exponential I&F model, which is also physically implemented on the DYNAP board.

The simulated network behavior confirmed that the inhibitory parameters d, σ_k and b play a key role for the tuning of the neurons on a specific orientation and spatial frequency. The activity of the retina layer retraced the data recorded with the event camera DVS128, so that to each pixel corresponds a neuron in the retina population and to each ON-event from the pixel corresponds a spike of the neuron. Only the (19×19) central region of the DVS128's array of pixels, corresponding to the dimension of the retina layer, has been considered and ON and OFF events have been handled separately.

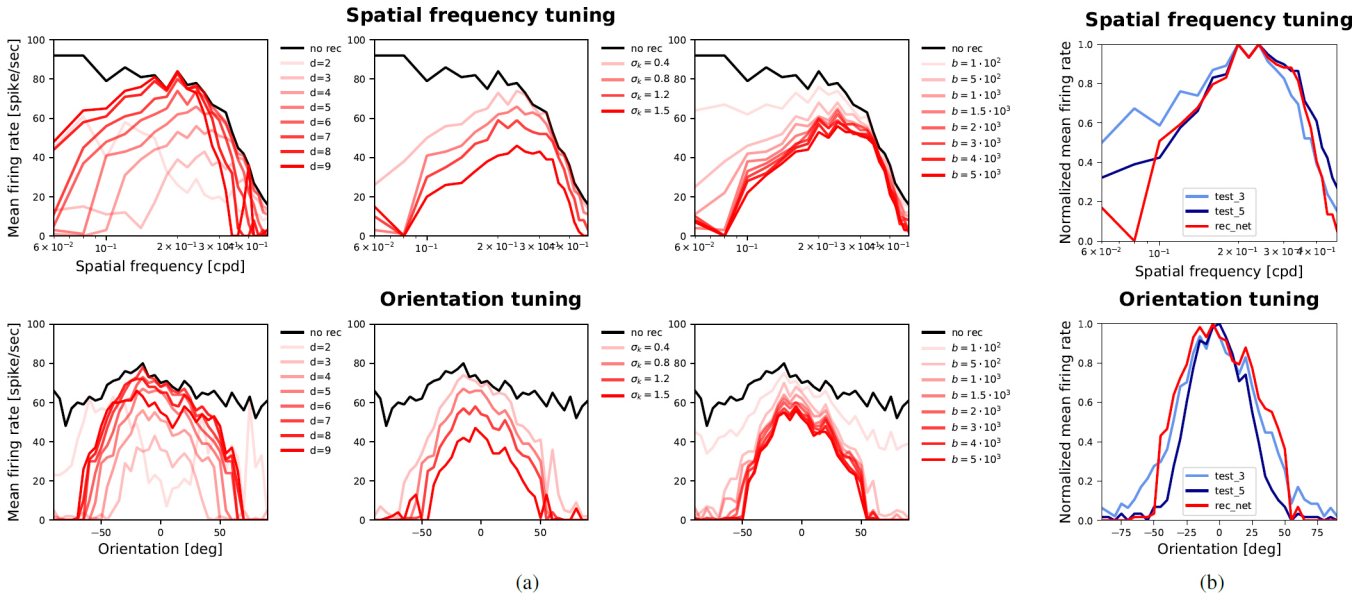
To have a benchmark for comparing the results achieved by the proposed method, we considered two test networks that implement the RFs of the neurons of interest (see Fig. 1 bottom) by feed-forward excitatory and inhibitory connections, directly obtained by sampling the desired Gabor function in Eq. 4. Specifically, in order to have the excitatory central region 5 pixel wide (i.e., of the same size as the feed-forward kernel), we have fixed $k_0 = 0.7$ and set $\sigma = 3.5$ and $\sigma = 4.7$ to obtain a 3-subregion and a 5-subregion receptive field, respectively. The `Brian2` simulation allowed us to study the effect of recurrent inhibition in a SNN and to determine the combination of parameters that yields the best tuning curves, without dealing with the restrictions posed by DYNAP in terms of the maximum number of synapses per neuron, and of the quantization of synaptic weights.

B. Implementation of the network on the DYNAP board

On the basis of the simulations, we have fixed the values of d and σ_k that yield the best results and then implemented the network on the neuromorphic processor DYNAP. The DYNAP board was interfaced with a PC through the software CTXCTL Primer [15]. The network structure was then slightly modified to overcome the restrictions posed by the DYNAP board:

- R1: each neuron can have at most 64 afferent connections;
- R2: each neuron in a core shares the same biases, including the synaptic weight of the afferent connections;
- R3: each neuron has two types of excitatory synapses and two types of inhibitory synapses, thus limiting to two the maximum number of different excitatory and inhibitory weights for each core;
- R4: only the shunting-type inhibitory synapse could be used since the other type was not effective in lowering the "membrane voltage" of the target neuron.

To overcome restriction R1, an extra layer of neurons, the relay layer, was added. The relay layer receives excitation from the retina layer through the feed-forward kernels, and projects one-to-one connections to the V1 layer, where inhibition takes place. For this network to behave as the one with the previous structure, the activity of the V1 layer should mimic the activity of the relay layer as closely as possible if the recurrent inhibition is excluded. The weights of the one-to-one connections between the relay layer and the V1 layer need to be adjusted accordingly. Due to restrictions R2, R3, and R4, the connection weights that define the kernels cannot be assigned by sampling the Gaussian profiles, as in the simulations, but have to be set to a single value, in the case



(a) (b)

Fig. 2. (a) In red the tuning curves for the central neuron of the V1 layer for different values of d, σ_k and $b = 3 \cdot 10^3$. In black the curves when the recurrent inhibition is removed. (b) Comparison between the tuning curves for the central neuron of the V1 layer of the recurrent network (obtained for $d = 5$, $\sigma_k = 1.2$, and $b = 3 \cdot 10^3$) and the neurons of the test networks with Gabor-like feed-forward receptive fields with 3 and 5 subregions.

of the recurrent kernel, or quantized by two levels in the case of the feed-forward kernel. The relay layer and the V1 layer were placed on different cores of the DYNAPS chip. The retina layer was assigned to 'virtual neurons', which are implemented by a module that acts as a spike generator, providing input spikes to the physical neurons on the target chip. The spiking activity of the silicon neurons can be recorded and further processed off chip.

C. Dataset of visual stimuli and experimental set-up

Sinusoidal gratings are widely used as visual stimuli to study the response of cells in the primary visual cortex; they allow us to obtain the *tuning curves* that characterize a neuron's behavior. PsychoPy [16] was used to generate moving sinusoidal gratings with specific orientation, spatial frequency and temporal frequency. The moving gratings were displayed on a screen with maximum brightness and recorded with the DVS128 event camera in a semi-dark room to reduce the refraction of the screen. The DVS128 was connected to a PC through a USB port, and the software jAER [17] enabled the real-time visualization and recording of streams of events into AEDAT files and allowed us to set the camera biases and apply filters on the data. Since the DVS128 is sensitive to local contrast changes, bands of ON and OFF events occur where the sinusoidal contrast profile is steep enough. Where the profile is almost flat, contrast differences are too small to be sensed and no events are generated, thus resulting in bands without events that are larger or narrower, according to the contrast sensitivity threshold. The spatial frequency information in preserved, encoded in the distance between the bands of events, not in their width, but the phase is shifted by $\pi/2$. The AEDAT files were converted into matrices containing the x and y pixel coordinates, the timestamp and the polarity of each recorded event. The matrices were then divided into segments, each one representing 1 s of recording, containing the data corresponding to a moving grating with fixed orientation, spatial frequency and temporal frequency.

III. Results

A. Simulated network

First, we tested the linearity assumption. If the network behaves in a linear way, the firing rate of the neuron of interest should oscillate with the same temporal frequency of the grating used as input. This condition is fulfilled for different temporal and spatial frequencies of the stimulus, both when the inhibitory recursion is switched off and on. Considering fixed parameters for the feed-forward kernel (in particular $\sigma_h = 3.5$, $p = 1/3$, and multiplicative weight of the feed-forward excitatory connections $a = 10^3$), the parameters that influence the effect of the recurrent clustered inhibition are the distance d between the target neuron and each of the inhibitory clusters, the dimension σ_k of the clusters and the multiplicative weight of the recurrent inhibitory connections b. The parameter that most influences the efficacy of the recursion in inducing a Gabor-like receptive field is d, since it is crucial that the clusters are correctly displaced. The first column of Fig. 2a shows how the tuning curves change according to d. Keeping σ_k fixed at 0.8 and b at $3 \cdot 10^3$, $d = 5$ yielded the best tuning curves both for spatial frequency and orientation. For lower values of d the neuron is not properly tuned, and the inhibition reduces the firing rate significantly. For higher values of d the tuning curves broaden and the peak spatial frequency shifts to the left. By keeping d and b constant, the tuning curves are broader for low values of

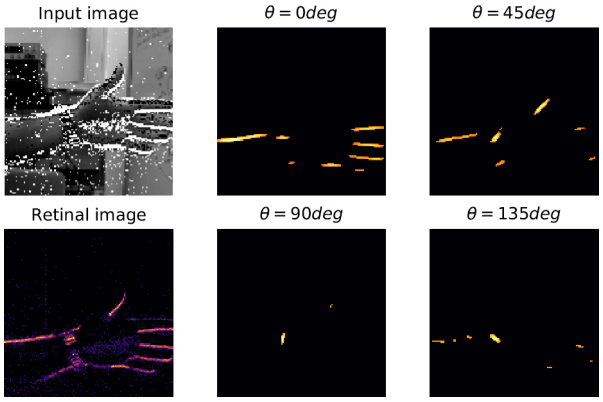


Fig. 3. A snapshot of the activity of the retina layer and of the V1 layer for different preferred orientations in response to the stream of events generated by the DVS, represented in the input image, for a real scene. Cropped DVS240 recording from DVSFLOW16 dataset [18].

σ_k and for high values the inhibition lowers the firing rate, as shown in the central column of Fig. 2a. If we compare the normalized curves, $\sigma_k = 1.2$ appears to be the best choice, yielding the sharpest tuning. Changing the value of b does not significantly affect the tuning of the neuron, given that the strength of the inhibition is at least sufficient to induce the tuning, as it can be seen in the rightmost column of Fig. 2a. Figure 2b shows the comparison between the tuning curves obtained by the recurrent inhibition and those obtained by the two test feed-forward networks. In terms of the spatial frequency tuning curve, the recurrent network yields the best results, while the orientation tuning curves are comparable. Anyhow, the recurrent inhibition method requires far less synaptic connections, as shown in Table I, which is a relevant feature if one considers the limitations posed by neuromorphic processors like the DYNAP. Figure 3 shows an example of network response, for differently oriented Gabor-like filters, when the input is a real scene recorded with the DVS.

	Number of afferent connections per neuron
Recurrent network with $\sigma_h = 3.5$ and $\sigma_k = 1.2$	101
Equivalent test network with 3 subregions	127
Equivalent test network with 5 subregions	241

TABLE I
COMPARISON BETWEEN RECURSIVE AND FEED-FORWARD SCHEME.

B. Network implemented on the DYNAP board

The tuning curves of the implemented filters for spatial frequency and orientation were obtained by reproducing the DVS128 recordings through the activity of the retina layer, and by recording the spikes of the central neuron of the V1 layer. Since neurons and synapses behavior on the DYNAP is not deterministic due to device mismatches, the curves were mediated over 10 sessions. The results are shown in Fig. 4. A neuron that receives only the feed-forward input is not tuned to any specific value of the features and its firing rate changes

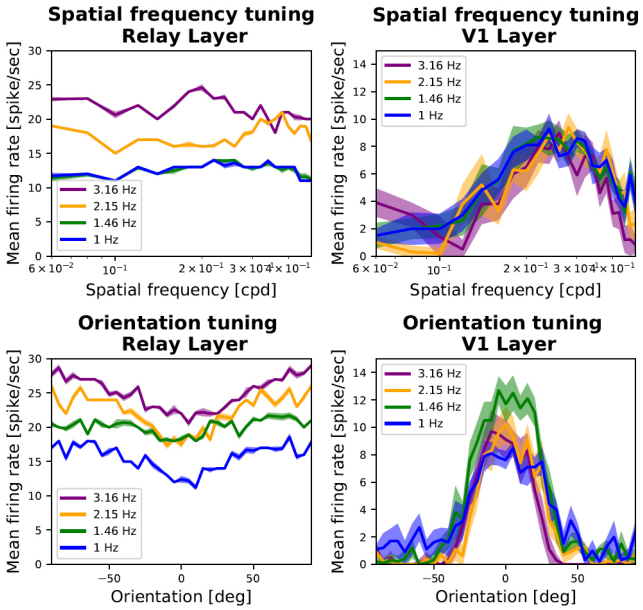


Fig. 4. Tuning curves for spatial frequency and orientation. The relay layer receives only feed-forward input, while the V1 layer receives also the recurrent inhibition contribution. Different colors indicate different temporal frequencies of the gratings used as visual stimuli.

according to the temporal frequency of the grating (since faster gratings elicit more events on the DVS128, and thus higher firing rate of the retina neurons that project to the relay layer and then to the V1 layer). When the recurrent inhibition is switched on, the neuron is clearly tuned to a specific spatial frequency and to a specific orientation. The curves obtained for different temporal frequencies overlay: this is evidence of the fact that the emergence of ON and OFF subregions in the receptive field induced by recurrent inhibition successfully normalizes the firing rate in input.

IV. CONCLUSION

We proposed a step forward in the design of spike-based, brain-inspired artificial vision processing on neuromorphic hardware. Different spiking network structures were designed and tested in simulation and on the neuromorphic processor DYNAP, to obtain silicon neurons that are tuned to visual stimuli oriented at specific angles and with specific spatial frequencies, provided by an event camera. Recurrent inhibition was successfully tested on SNNs, both in simulation and on the DYNAP board, to obtain neurons with highly structured Gabor-like receptive fields; these neurons are characterized by tuning curves that are sharper or at least comparable to the ones obtained using only feed-forward connections, but need a significantly lower number of synapses.

ACKNOWLEDGMENT

This project has received funding from the European Research Council under the Grant Agreement No. 724295 (NeuroAgents).

REFERENCES

[1] C. Mead, "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990.

[2] G. Indiveri and S. Liu, "Memory and information processing in neuromorphic systems," *Proceedings of the IEEE*, vol. 103, no. 8, pp. 1379–1397, 2015.

[3] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," *Journal of neurophysiology*, vol. 58, no. 6, pp. 1233–1258, 1987.

[4] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128x128 120 db 15us latency asynchronous temporal contrast vision sensor," *IEEE journal of solid-state circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[5] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, "A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (dynaps)," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 1, pp. 106–122, 2018.

[6] L. Raffo, S. P. Sabatini, G. M. Bo, and G. M. Bisio, "Analog VLSI circuits as physical structures for perception in early visual tasks," *IEEE Transactions on Neural Networks*, vol. 9, no. 6, pp. 1483–1494, 1998.

[7] T. Y. W. Choi, P. A. Merolla, J. V. Arthur, K. A. Boahen, and B. E. Shi, "Neuromorphic implementation of orientation hypercolumns," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 6, pp. 1049–1060, 2005.

[8] T. Y. W. Choi, B. E. Shi, and K. A. Boahen, "An on-off orientation selective address event representation image transceiver chip," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 2, pp. 342–353, 2004.

[9] L. Raffo, S. P. Sabatini, M. Mantelli, A. De Gloria, and G. M. Bisio, "Design of an asip architecture for low-level visual elaborations," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 5, no. 1, pp. 145–153, 1997.

[10] F. Folowosele, R. J. Vogelstein, and R. Etienne-Cummings, "Towards a cortical prosthesis: Implementing a spike-based hmax model of visual object recognition in silico," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 516–525, 2011.

[11] G. Orchard, X. Lagorce, C. Posch, S. B. Furber, R. Benosman, and F. Galluppi, "Real-time event-driven spiking neural network object recognition on the spinnaker platform," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2015, pp. 2413–2416.

[12] S. P. Sabatini, "Recurrent inhibition and clustered connectivity as a basis for gabor-like receptive fields in the visual cortex," *Biological cybernetics*, vol. 74, no. 3, pp. 189–202, 1996.

[13] M. Stimberg, D. F. Goodman, V. Benichoux, and R. Brette, "Equation-oriented specification of neural models for simulations," *Frontiers in neuroinformatics*, vol. 8, p. 6, 2014.

[14] M. Milde, A. Renner, R. Krause, A. M. Whatley, S. Solinas, D. Zendrikov, N. Risi, M. Rasetto, K. Burelo, and V. R. C. Leite, "teili: A toolbox for building and testing neural algorithms and computational primitives using spiking neurons," 2018, unreleased software, Institute of Neuroinformatics, University of Zurich and ETH Zurich.

[15] aiCTX. (2018) CTXCTL primer. Last checked on Nov 20, 2019. [Online]. Available: https://ai-ctx.gitlab.io/ctxctl/primer.html

[16] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv, "PsychoPy2: Experiments in behavior made easy," *Behavior Research Methods*, vol. 51, no. 1, pp. 195–203, 2019.

[17] Delbruck. (2007) jaer open source project. [Online]. Available: http://jaer.wiki.sourceforge.net.

[18] B. Rueckauer and T. Delbruck, "Evaluation of event-based algorithms for optical flow with ground-truth from inertial measurement sensor," *Frontiers in neuroscience*, vol. 10, p. 176, 2016.