– Supplementary Notes –

Recurrent models of orientation selectivity enable robust early-vision processing in mixed-signal neuromorphic hardware

December 22, 2024

Valentina Baruzzi
Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Via Opera Pia 13, I-16145, Genoa, Italy

Giacomo Indiveri
Institute of Neuroinformatics, University of Zurich and ETH Zurich, Switzerland

Silvio P. Sabatini[1]
Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Via Opera Pia 13, I-16145, Genoa, Italy

1 Effect of additional recursive excitatory connections

The addition of recursive excitatory connections was tested as a means to improve the tuning of the central neuron of the simulated V1 layer to a specific orientation and spatial frequency. The recurrent excitatory kernel, analogous to the inhibitory one, was defined as two Gaussian functions equidistant from the target neuron, but in the orthogonal direction (i.e., aligned with the initial orientation of the feed-forward contribution from the retina layer):

$$w^E(x_\theta, y_\theta) = \frac{1}{2\pi\sigma_k^2}\left[\exp\left(-\frac{x_\theta^2 + (y_\theta + d)^2}{2\sigma_k^2}\right) + \exp\left(-\frac{x_\theta^2 + (y_\theta - d)^2}{2\sigma_w^2}\right)\right] \quad (1)$$

Accordingly, the excitation e of the central neuron of the V1 layer becomes the solution of:

$$e(\mathbf{x}) = a(h_0 * s)(\mathbf{x}) - b(w * e)(\mathbf{x}) + b^E(w^E * e)(\mathbf{x}) \quad (2)$$

where b^E is the strength of the excitation. The results are shown in Supplementary Fig. S1: the addition of recursive excitation does indeed result in narrower tuning curves. However, if the strength of the excitation exceeds the maximum value shown in Supplementary Fig. S1, the network becomes unstable and the neuron is no longer tuned to specific features.

2 Visual stimuli definition and recording

Moving gratings with a sinusoidal contrast profile were generated using PsychoPy, an open-source package for running neuroscience experiments in Python, employing the built-in class *GratingStim*. Each

[1]Corresponding author, e-mail: silvio.sabatini@unige.it

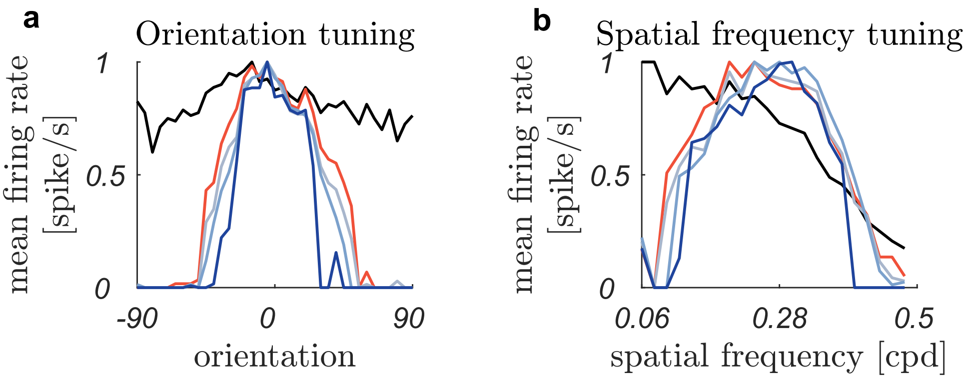

Figure S1: **Tuning curves obtained by adding recursive excitation.** Comparison between the tuning curves for the central neuron of the simulated V1 layer when only recurrent inhibition is present (red curves, parameters set to $d = 5$, $\sigma_k = 1.2$, and $b = 3 \cdot 10^3$) and when recurrent excitation is also added (blue curves, parameters set to $d = 5$, $\sigma_k = 1.2$, $b = 3 \cdot 10^3$, and $b^E \in \{50, 100, 200\}$, increasing b^E values represented by increasing color saturation). **a** Orientation tuning curves. **b** Spatial frequency tuning curves. Black lines represent the curves obtained when both recurrent inhibition and excitation are removed.

grating is described in a polar coordinate system in terms of its orientation θ and radial spatial frequency k_s, with units of degrees and cycles/deg (cpd), respectively. The unit of measure cpd is commonly used when describing visual stimuli in psychophysics experiments and indicates the number of cycles per degree of visual field (in this case, a cycle is a spatial period of the sinusoid). As such, it is dependent on the distance of the observer from the visual stimuli, or, in this case, the recording device. PsychoPy automatically adjusts the size of the gratings reproduced on the screen based on the distance from the recording device and the monitor's resolution in pixels, to match the desired spatial frequency value. Since only moving stimuli generate a response from the DVS, drifting gratings were used instead of static ones. Each grating is thus also described by its temporal frequency, i.e., the number of grating cycles that pass a point in the image plane per unit time. Constant temporal frequency results in the phase of the sinusoidal grating evolving linearly with time. Supplementary Fig. S2a shows an example of consecutive snapshots of a moving sinusoidal grating as reproduced on screen.

The visual stimuli were displayed on a screen with maximum brightness and acquired by the DVS event camera (at a fixed distance of 40 cm) in a semi-dark room to reduce as much as possible the refraction of the screen. Supplementary Fig. S2b shows examples of the DVS response to the moving sinusoidal gratings as visualized through the jAER interface. Each pixel of the DVS is represented as a square that appears grey if the pixel is silent and becomes white or black if ON and OFF events, respectively, are generated. This figure illustrates in practice the DVS response to local contrast changes, explained theoretically in Supplementary Fig. S3.

It is worth noting that, due to the nature of the DVS, fast-moving stimuli elicit a higher number of events than slow ones. As such, for moving gratings with the same spatial frequency, the higher the radial spatial frequency k_s the lower the velocity of the grating, resulting in a variable average number of events recorded. Considering a constant drifting velocity instead of constant spatial frequency does not produce a constant average number of recorded events either, because a higher number of contrast differences for unit of space (i.e., higher stimulus spatial frequency k_s) elicit more events. This phenomenon is shown in Supplementary Fig. S4.

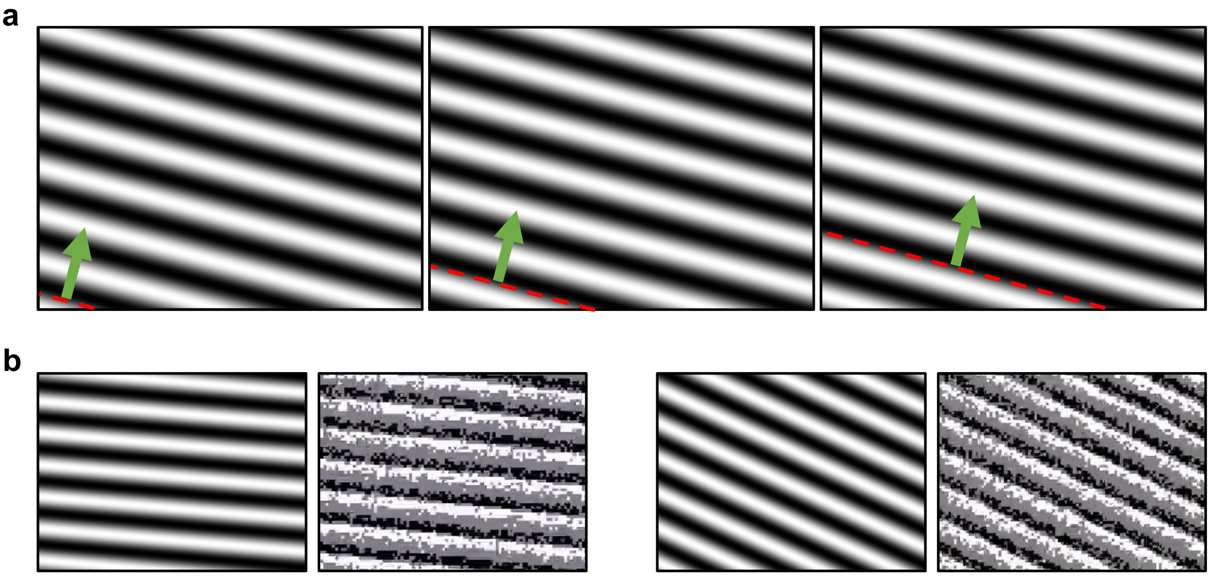

Figure S2: **Moving sinusoidal gratings. a** Consecutive snapshots of a moving sinusoidal grating with a temporal frequency of 3 Hz captured at intervals of 0.2 s. The red dashed line highlights the same wavefront in all snapshots, whereas the green arrow indicates the direction of movement, perpendicular to the wavefront. **b** Pairs of snapshots of moving sinusoidal gratings as reproduced on the screen and of the corresponding DVS recordings as visualized through the jAER interface.

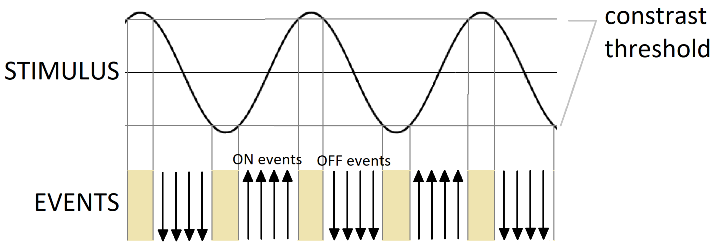

Figure S3: **How the sinusoidal grating is converted in events by the DVS.** Upward arrows represent ON events, whereas downward arrows represent OFF events. Yellow regions denote temporal intervals with no events.

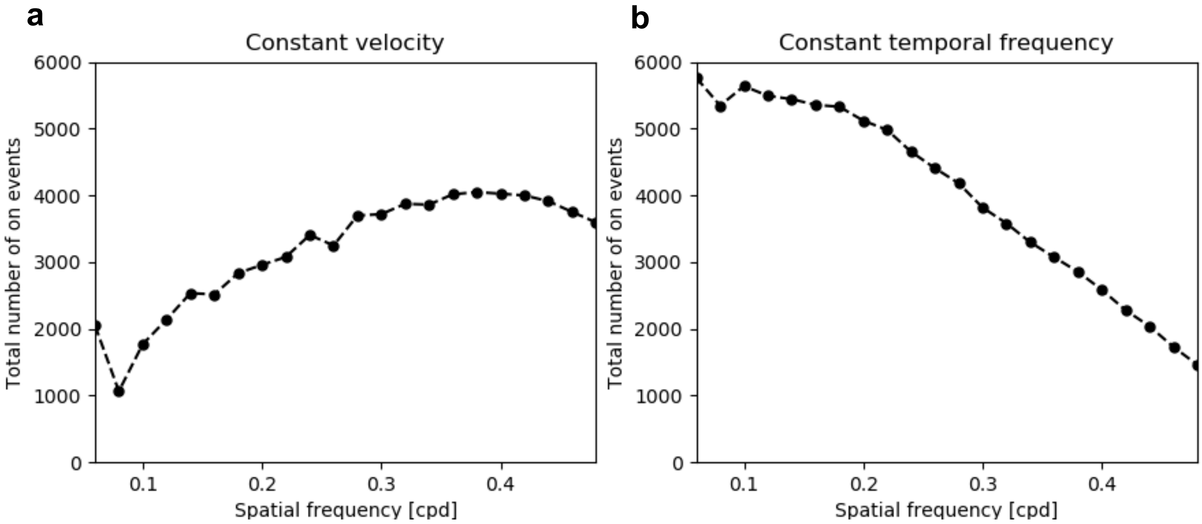

Figure S4: **Variation of the total number of events as a function of spatial frequency.** The total number of events registered by the DVS in recording slots of 1 s each varies according to the spatial frequency k_s of the visual stimulus. Panel **a** displays the case of constant velocity (10 Hz/cpd), whereas panel **b** displays the case of constant temporal frequency (3 Hz). The total number of events is obtained considering ON events in a central patch of 19×19 pixels.