Spiking motion direction through object motion sensitivity

Riccardo Pignari[1,*], Vittorio Fra[1], Gianvito Urgese[1], James C. Knight[2], Giulia D'Angelo[3]

[1] Politectico di Torino, Turin, Italy
[2] The University of Sussex, UK
[3] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic
[*] Corresponding author: riccardo.pignari@polito.it

Abstract Unlike conventional frame-based imaging, neuromorphic sensing mimics the brain's asynchronous processing. Event-driven cameras[1] capture only changes in light intensity, providing microsecond-scale temporal resolution—crucial for tracking satellites amidst celestial motion and atmospheric interference. This sparse encoding significantly reduces data volume while preserving a high dynamic range, making it well-suited for challenging lighting conditions and enabling motion-blur-free imaging. The combination of event-driven input and spiking neural networks (SNNs), deployed on non-Von Neumann neuromorphic hardware, further supports real-time, adaptive processing with ultra-low power consumption, often in the milliwatt range. To distinguish moving objects while the event-based camera itself is in motion, the agent requires a robust object motion segmentation (OMS) mechanism. Building on the work of D'Angelo et al. [3], this study proposes an SNN architecture for OMS in dynamic scenes, coupled with an additional SNN module for motion direction detection.

Methods The architecture illustrated in Fig. 1 depicts the event-based visual input being processed by the spiking object motion segmentation (OMS) module, which generates an OMS map with enhanced motion contrast. The model employs a Difference of Gaussians (DoG) approach, implemented via two Gaussian kernels: center and surround, whose subtraction accentuates local motion contrast. The resulting OMS map is then passed to the spiking Elementary Motion Detector (sEMD) [2], in which receptive fields encode both the speed and direction of motion within the dynamic scene. The entire architecture is implemented using GPU-enhanced neural networks (GeNN) [1], leveraging Leaky Integrate-and-Fire (LIF) neurons for efficient spiking computation.

Object motion segmentation To isolate relevant motion, three processing steps are applied: (1) subtraction of the surround response from the centre response to enhance contrast, (2) min-max normalisation to scale values between 0 and 1, and (3) a clamping operation to suppress irrelevant background activity based on a threshold.

Motion detection Motion detection is performed downstream of the OMS stage (Fig. 1, stage 2) using four spiking neuron populations, each tuned to one of the cardinal directions: up, down, left, and right. Directional selectivity is achieved through a facilitator-trigger circuit, a binary spiking mechanism that generates spike bursts whose rate and inter-spike intervals encode time-difference information. The direction of motion is inferred from the activity of the neuron population associated with each cardinal direction, which is driven by dedicated facilitator-trigger pairs.

Preliminary results The dataset used in the present study is the Continual Object Recognition 50 (CORe50) dataset, which contains RGB-D recordings of 50 household objects. Data acquisition was conducted across 11 sessions, each designed to introduce variability in background and illumination conditions. For each object and session, a 15-second video was recorded at 20 frames per second using a Kinect 2.0 sensor, yielding 300 temporally aligned RGB-D frames per sequence. Event data were generated from the CORe50 recordings using the ICNS Event-Based Camera Simulator (IEBCS), an open-source tool that converts frame-based sequences into synthetic event streams. To enhance biological realism, microsaccadic motion was introduced during event generation, governed by a probabilistic model derived from empirical studies of human eye movements. The full processing pipeline was implemented using the GPU-enhanced Neuronal Networks (GeNN) framework, allowing efficient simulation of large-scale spiking neural networks.

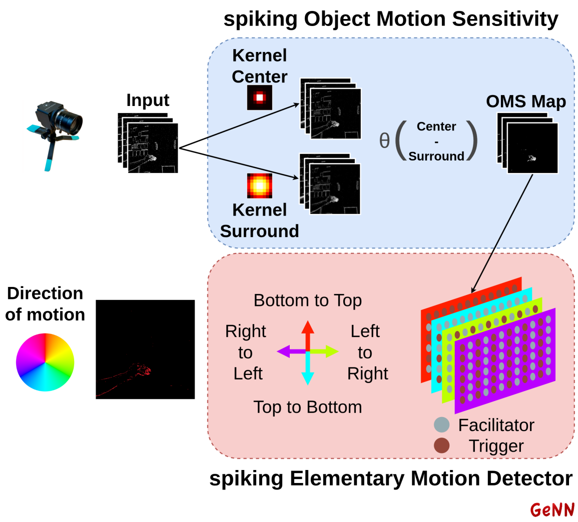

Figure 1: Building blocks of the object segmentation and motion detection pipeline. In the first stage, the segmented object is obtained by kernels. In the second stage, the motion of the OMS map is classified to recover its direction.

Future work will focus on deploying this neuromorphic vision pipeline on an embodied robotic platform. Such an implementation would enable interactive object recognition and motion perception in real-world environments, showcasing the practical potential of the proposed architecture for autonomous systems.

References

[1] Knight, James C., Anton Komissarov, and Thomas Nowotny. "PyGeNN: a Python library for GPU-enhanced neural networks." Frontiers in Neuroinformatics 15 (2021): 659005.

[2] D'Angelo, Giulia, et al. "Event-based eccentric motion detection exploiting time difference encoding." Frontiers in neuroscience 14 (2020): 451.

[3] D'Angelo, Giulia, et al. "Wandering around: A bioinspired approach to visual attention through object motion sensitivity." *Neuromorphic Computing and Engineering* 5.2 (2025): 024019.

[1] Gallego, Guillermo, et al. "Event-based vision: A survey." IEEE Transactions on Pattern Analysis and Machine Intelligence 44.1 (2020): 154–180.