A Controller for Adaptive Neuromorphic Sensing using Active Efficient Coding

Aheli Saha Didier Stricker René Schuster
RPTU – University of Kaiserslautern-Landau
DFKI – German Research Center for Artificial Intelligence
firstname.lastname@dfki.de

1. Background

Event cameras have gained significant attention in recent years, due to their asynchronous, energy-efficient operational technique- traditional cameras process intensity values at all pixel positions at fixed time intervals. In contrast, event cameras asynchronously capture data only at positions where the log change in intensity crosses a threshold, mimicking biological vision systems. Such encoding approaches are immediately influenced by signal dynamics, thereby reducing redundant information. This makes them particularly energy-efficient and low-latency alternatives to traditional systems. In neuromorphic systems, event cameras supplement the parallel processing capabilities of neuromorphic processors, enabling effective handling of asynchronous data streams.

Analogous to computational systems, biological systems, too, are faced with the constraint of limited resources. It has been studied and observed that to meet these energy constraints, the brain has evolved to not only optimize neural representations but also use motor actions to vary and adapt the statistics of the input signals. The Active Efficient Coding (AEC) theory [2] postulates that the organism's behaviour that shapes these input sensory signals is tuned in conjunction with the neural representations to arrive at an optimal performance. The authors of [4] provide an excellent illustration of this theory in binocular vision, in which the variance in neural representations of binocular disparities resembles the statistics of disparities that humans experience when engaging with natural scenes. This identifies a fundamental link between brain representations and active eye movement control and suggests an effective strategy to maximize resource allocations while optimizing for performance.

We argue that this concept can be extended to artificial vision systems, specifically event cameras, in our case, to further improve their energy utilization. The conditions for triggering an event are generally pre-determined through calibration procedures and applied once using programmable bias circuits. This indicates that sensor transduction is constant, and there is no dynamic adjustment for varying environmental stimuli or the current task requirements. To overcome this limitation while concurrently promoting enhanced efficiency, we propose jointly optimizing the input statistics and performance of a task-specific neural network. The aim is to develop an intermediate controller that dynamically modulates the internal characteristics of the Dynamic Vision Sensor (DVS).

2. Current Setup

To illustrate and validate the idea, we propose to simulate a dynamic controller for event-based neuromorphic vision. Our study uses the autonomous driving simulator CARLA [1]. CARLA has an integrated DVS camera and allows for accessing the sensor parameters via APIs, providing a controlled setting to study sensor and performance optimization systematically. An event is triggered when

$$\phi(x,t) - \phi(x,t-\delta t) > \tau \mid t - \delta t > \eta \tag{1}$$

where τ indicates the threshold for change in intensity ϕ and η represents the refractory period. It is with these parameters that we seek to modulate the sensory behaviour.

Initial Experiments. We begin by identifying the effects of each of these parameters on the efficiency of the system. Efficiency is quantitatively approximated using two simplistic metrics: the spike count, where a lower count denotes greater efficiency, and the average inter-spike interval, with higher intervals signifying enhanced efficiency. These metrics are computed within temporal windows of 50 ms and subsequently averaged across a simulated dataset for each experimental condition. Our experiments back the intuitive effect that can be theorized, and the variations in Spike Count are summarily presented in Fig. 1.

Neuromorphic Task. We focus on object detection as the downstream task to exemplify the system's objective. Therefore, within the simulation framework, we employ an off-the-shelf event object detector [5], retrained for our purposes, to predict object bounding boxes and produce a set of performance metrics.

Closed-Loop Control. Subsequently, we establish a closed-loop simulation framework within the CARLA environment to investigate and implement the principles of the AEC strategy. Beginning with an initial parameter configuration for the sensor, we systematically collect triggered events and the corresponding ground truth bounding boxes at each simulation step. We compute the aforementioned quantitative event statistics from the triggered events, and simultaneously process them into a dense Stacked Histogram [3] representation. This step yields a three-dimensional tensor representation compatible with standard convolutional operations.

The metric scores of the detector, along with the precomputed spike statistics that denote efficiency, are provided as input to a basic multilayer perceptron (MLP)-based controller. This controller is aimed at predicting prescribed digressions to the current sensor configuration and is trained using a min-max optimization criterion, with the objective of simultaneously enhancing system efficiency and task-specific performance. Intuitively and based on our previous experiments, these objectives are contradictory, as more data (increased spike count) leads to better performance but lower efficiency, and vice versa. A visual representation of the control loop is shown in Fig. 2.

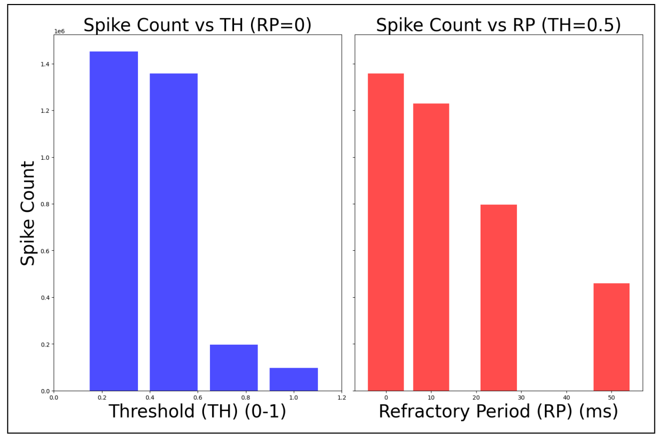

Figure 1. Effect of Threshold (τ) and Refractory Period (η) variations on Spike Count.

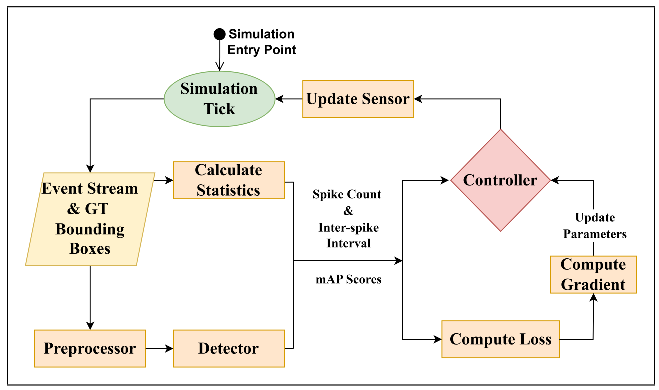

Figure 2. The control loop updates the controller parameters using the contrastive loss from the current timestep before recording data for the next simulation step.

3. Challenges, Limitations, and Future Work

There remain unresolved challenges which are being actively addressed before we can demonstrate the full potential of our proposal. The contrastive nature of the loss makes the optimization process for the controller particularly unstable, impeding convergence on an optimal solution. Further supplementing our dynamic parameter controller, we aim to formulate an adaptive Region of Interest (ROI) filter to constrain inessential computations. With the preliminary proof-of-concept established, we seek to systematically explore and evaluate multiple significant statistical metrics that can be integrated with Spiking Neural Networks to fully leverage the asynchronism and sparsity of the event data.

In a broader view, we aim to investigate the computational models to form a general framework for the efficient coding of sensory systems by unifying cross-modal sensory signals. The adaptation and encoding process of the system are dynamic and adhere to (i) sensory efficiency, via a sensor-specific controller, analogous to our current model, and (ii) a feedback signal from a multi-sensory task-specific controller. We focus on the vision, audio and tactile perceptory modalities, but an objective task that consequentially combines these tracks remains an open question.

References

[1] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017.

[2] Samuel Eckmann, Lukas Klimmasch, Bertram E Shi, and Jochen Triesch. Active efficient coding explains the development of binocular vision and its failure in amblyopia. *Proceedings of the National Academy of Sciences*, 117(11):6156–6162, 2020.

[3] Mathias Gehrig and Davide Scaramuzza. Recurrent vision transformers for object detection with event cameras. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 13884–13893, 2023.

[4] William W Sprague, Emily A Cooper, Ivana Tošić, and Martin S Banks. Stereopsis is adaptive for the natural environment. *Science advances*, 1(4):e1400254, 2015.

[5] Nikola Zubic, Mathias Gehrig, and Davide Scaramuzza. State space models for event cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5819–5828, 2024.