



Neuromorphic Vision to Control for UAVs

APPLICATION NOTE

October 6th, 2025



Platform Goals

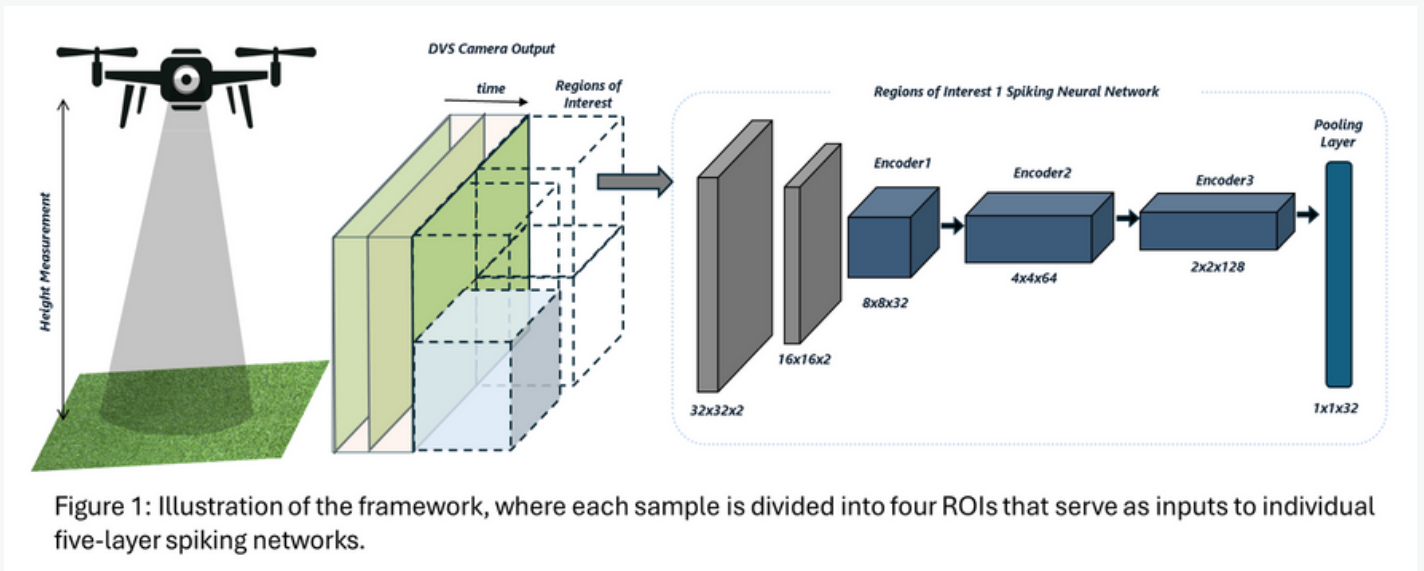
The goal of this research is to present a fully neuromorphic vision to control pipeline for a flying drone, demonstrating autonomous vision based ego motion estimation and control to perform maneuvers like hovering, landing, and sideways movement, while leveraging the low latency and high energy efficiency potential of neuromorphic systems. This pipeline can be crucial for real world autonomous vehicles, particularly small flying drones, because it offers a solution for deploying complex deep neural networks on lighter platforms with highly stringent Size, Weight, and Power (SWaP) constraints.

Use Case

To avoid the "reality gap," the vision Spiking Neural Network (SNN) was trained and evaluated exclusively on a real world dataset of approximately 40 minutes of event data, collected using the same drone and downward facing DVS 240 event camera used in the final flight experiments. The dataset, recorded in an indoor environment with a texture-rich floor, was split into 25 minutes for training and 15 minutes for evaluation. The data distribution includes a wide variety of motions, from slow to fast, including rapid rotations of approximately 4 radians per second, ensuring the network was trained on a representative range of ego motions. The raw input data from the camera, which has a native resolution of 240x180 pixels, undergoes significant preprocessing to meet hardware constraints; it is down sampled and cropped into four independent Regions of Interest (ROI) located in the image corners, resulting in a final input dimension of 16x16 pixels for each of the four replicated SNNs. Furthermore, to manage I/O bandwidth, the number of events processed in each 5-millisecond window is limited to a maximum of 90 per ROI.

Architecture

The system's architecture is a fully neuromorphic vision to control pipeline divided into two distinct networks: a vision network and a control network. The input to the vision network is raw data from an event camera, which is downsampled and cropped into four separate 16x16 pixel ROI located in the image corners. The vision network is a SNN that processes these ROIs independently using four identical copies of a five-layer architecture, totaling 28,800 neurons. Each of these replicated networks incorporates an input layer, three self-recurrent encoder layers, and a final spiking pooling layer, all using a current-based Leaky Integrate and Fire (CUBA LIF) neuron model. This SNN is responsible for estimating optical flow from the event data. The control network is a much simpler, single linear mapping layer that takes the optical flow estimates from the vision network, along with the drone's attitude and a user-provided set point, and maps them to low level motor commands. To create an efficient, end to end neuromorphic pipeline, the linear decoding of spikes from the vision network and the linear control mapping were mathematically merged into a single matrix multiplication, directly transforming vision spikes into control commands.



Technical Performance

The vision SNN, trained with a self-supervised method, produced optical flow estimates that closely captured the drone's ego-motion, matching ground-truth data even during fast rotations of ~ 4 rad/s. The primary metric for vision accuracy was the Average Endpoint Error (EPE), the Euclidean distance between predicted and ground-truth optical flow vectors. **Our neuromorphic accelerator (implemented on a Xilinx/AMD Artix UltraScale+ (XCAU15P) FPGA) delivers 24 GOP/s at ~ 0.92 W (~ 26.1 GOP/s/W).** Under the same workload, Jetson Nano (10 W mode) achieves ~ 36.0 GOP/s throughput and ~ 3.6 GOP/s/W efficiency.

Net: Our FPGA performed $\sim 7\times$ higher energy efficiency than Jetson Nano while sustaining real-time, on-device processing without cloud dependence.

Roadmap

Current priorities focus on characterizing the baseline vision and control model pipeline on the custom neuromorphic hardware, despite the current challenges posed by hardware capabilities and configuration within the HLS framework. We will bring up the fully neuromorphic vision to control stack on our HDK (Artix-7/US+, PCIe Gen2 x4) by quantizing/partitioning the four ROI 16×16 SNN, enforcing the ≤ 90 events/ROI/5 ms budget, and fusing spike decoding with the linear control head to meet ≥ 40 FPS at ≤ 1 W; stand up an end to end profiling harness and run scene-wise studies to quantify latency, energy, and EPE, including ablations for temporal sparsity, ROI count, and encoder depth, targeting $\geq 10\times$ GOP/s/W vs. Jetson Nano. Our design will execute closed-loop flight trials (hover/land/strafe) with IMU/motor telemetry to validate stability and control error under fast rotations (~ 4 rad/s) and publish per-scene tables and CPU/GPU/Jetson baselines; and in parallel, derive compute/SRAM/NoC budgets from HDK traces to advance a 28 nm ASIC (~ 1.2 TOP/s at ~ 0.3 W, ~ 4.0 TOP/s/W) through pre-silicon emulation, MPW shuttle tape out, and post silicon correlation while de-risking multi-chip scaling via locality aware mapping, traffic-shaped NoC, and SRAM banking.

✉ hello@type1compute.com

☎ 503.810.8092

🌐 www.type1compute.com

T1C