Towards Adaptive Neuromorphic Sensing based on Active Efficient Coding

Quentin Halbach[1,2], Vibhu Kaushik[3], Aheli Saha[4], Nguyen H. P. Nguyen[5], Arne Roennau[5], Chiara Bartolozzi[1], Alessio Franci[2], Elisabetta Chicca[6], René Schuster[4,7], and Claudia Lenk[3]

[1] IIT – Italian Institute of Technology
[2] University of Liège
[3] University of Ulm
[4] RPTU – University Kaiserslautern-Landau
[5] KIT – Karlsruhe Institute of Technology
[6] University of Groningen
[7] DFKI – German Research Center for Artificial Intelligence

Abstract. Autonomous systems need a stable, efficient, and fast perception of their environments to react appropriately to dangers, navigate, and interact with others. Adaptation of sensing and processing is one approach to achieve this. We develop a neuromorphic system for perception, which uses three modalities, vision, hearing, and tactile, and implements active efficient coding of the sensory input. Thereby, the sensor properties, including encoding of the signal and the sampling procedure, e.g., movement of the tactile sensor over a surface, are adapted to yield efficient encoding and optimal performance in perception tasks like object detection. Here, we present results regarding the design of adaptive neuromorphic sensors, the identification of sensor parameters, which are most important for performance and efficiency metrics, the development of a controller to tune the sensor properties, and the system setup to merge the signals and processing of the different modalities.

Keywords: active efficient coding · adaptive neuromorphic sensors · control · multi-modality systems.

1 Introduction

Living beings and other autonomous systems like robots need to perceive their environments reliably, efficiently, and fast to react appropriately to dangers, navigate, and interact with others. Stable perception is thereby challenged by the continuously changing inputs and surroundings, which strongly influence the stimuli to the senses. In this context, adaptation of the senses is discussed as a possible solution to overcome these challenges [6,24,30], providing sensory inputs to processing, which are less influenced by environment and stimulus conditions, thus supporting stable and reliable perception. Besides adaptation, robots and humans have to work with restricted power/energy budgets and, thus, be as efficient as possible [19].

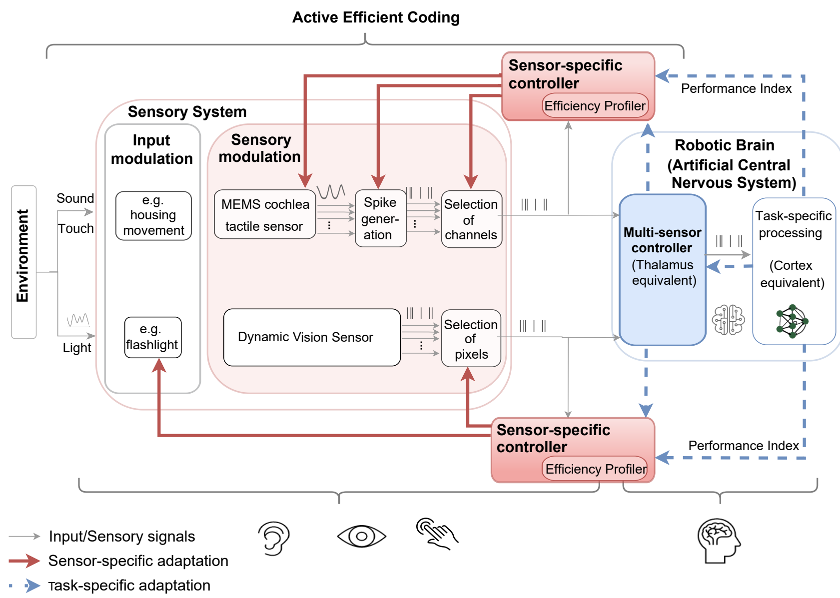


Fig. 1: Overview of the neuromorphic system for implementing active efficient coding using vision, hearing, and touch as sensory modalities.

Recently, active efficient coding (AEC) [8, 26] was proposed to solve both challenges, i.e., adaptation to improve perception while simultaneously trying to be as efficient as possible. AEC is a combination of efficient coding and active sensing. Efficient coding [3] refers to the reduction of redundant data, i.e., minimizing the mutual information between stimulus and sensor output or sensory signal and cortex representation, and reducing energy/power consumption, e.g., by reducing the number of necessary spikes to convey the sensory information. Active sensing refers to actively changing how the input is sampled and to tuning properties to optimally detect the desired stimulus [11]. While simulations of the active efficient coding model demonstrate that its advantages [8], no hardware realization, particularly for mulitmodal systems, has yet been presented.

We work on developing a neuromorphic system to implement AEC for different sensory modalities separately as well as for their combination. A schematic of the intended system is shown in Fig. 1. Particularly, we use event-based cameras like DVS [22] for vision, the MEMS cochlea [15] for hearing and neuromorphic MEMS-based pressure sensors for touch modality. While all these neuromorphic sensors are adaptable, the amount of tunable properties and the range of tuning are quite different, thus requiring adjusted control algorithms. In contrast to other adaptive neuromorphic systems, which base adaptation on signal statistics and stimulus-specifics [6, 27], we will combine signal-based adaptation (using signal statistics and stimulus-specifics) with the simultaneous adaptation based on the task performance, which are, e.g., accuracy or confidence in object identification or sound recognition tasks. Our system is thereby a quadruped ANYmal robot with neuromorphic senses, which can recognize different terrains and select an appropriate walking mode, a task that is difficult to accomplish using only a single sensing modality. To realize this system, we (i) develop adaptive neuromorphic sensors, whose sensing and encoding properties can be tuned dynamically, (ii) identify the important parameters to tune, (iii) develop spike-based, neuromorphic processing using the input from the different modalities, and (iii) develop the controller to adapt sensor properties depending on efficiency and performance. In the following, we will describe our approach and results for these different steps.

2 System Description and Results

2.1 Sensor Design

Neuromorphic sensors can both sense the environment and encode the acquired information into spikes for further processing, with the advantage of being low power and of transmitting information less frequently than conventional sensors sampled at a fixed frequency. Fixed sensor and encoding parameters define an operating point at which encoding properties perform optimally only for a limited set of environmental conditions, not necessarily suited for real world, highly non-stationary sensory inputs. As an example, an event-based camera could be calibrated for a static setting, and perform poorly in a dynamic setting where the camera suffer from motion or vibrations of its support, thus generating many noisy events [28] or missing weak signals that may be relevant for the task. Downstream filtering can (partially) mitigate this issue, but these additional processing steps still incur energy and computational costs that could be avoided if the sensor operated closer to an optimal encoding regime. The optimal encoding strategy may itself be task dependent. Neural responses in early sensory areas are modulated according to task demands, with task-informative neurons showing stronger gain modulation, supporting adaptive representation of stimulus features relevant for the current task [12], which cannot incorporated for fixed encoding parameters.

To address these challenges, adaptive mechanisms that dynamically modify sensor or encoding properties can increase system flexibility and enable more stable performance across diverse environments. At the sensor level, adaptation directly modifies the physical transduction mechanism by changing the transfer characteristics of the sensing element, for example, through feedback, thus altering mechanical or electrical properties [15]. In this approach, the sensing process is actively shaped by feedback signals before spike generation occurs. Alternatively, adaptation can be applied at the level of the edge encoder, tuning the analog-to-spike conversion by varying parameters such as event thresholds, refractory period, or bandwidth.This can significantly reduce the load on downstream processing stages, thereby improving computational efficiency and

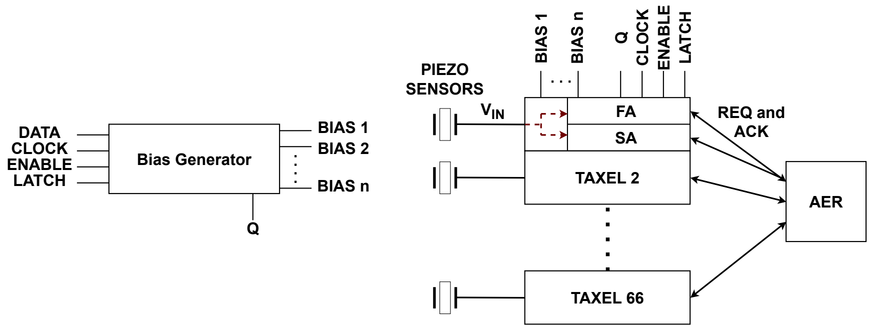


Fig. 2: Architecture of the tactile encoding block implemented in CMOS. Piezoelectric tactile sensors are directly connected to the input pads of the chip. Each taxel independently integrates fast-adapting (FA) and slow-adapting (SA) mechanoreceptor-inspired circuits that encode the sensor's analog output into spike trains. The encoding behavior is controlled by digitally programmable biases.

reducing communication and energy costs. For example, in the work by Delbruck et al. [6], event rate and noise-related statistics are used to dynamically adjust sensor parameters to keep the event rate in specified bounds and improve SNR. In another approach [18], a two time-scale adaptation strategy is used, tuning first refractory period and secondly bandwidth to keep the event rate in specified bounds.

While for hearing (MEMS cochlea) and vision (DVS), adaptable sensors are available, we needed to design an adaptive sensor for touch. To realize an adaptive tactile sensor, either piezoelectric or capacitive, can be combined with an adaptive edge encoding block that replicates two biological tactile mechanoreceptors in CMOS technology: fast adapting (FA) and slow adapting (SA) mechanoreceptors. The FA mechanoreceptor encodes temporal changes in skin deformation, while the SA mechanoreceptor encodes sustained pressure and spatial deformations [4]. Fig. 2 illustrates the developed tactile encoding block. Either piezoelectric or capacitive sensors can be directly connected to the input pads of the CMOS chip. Each taxel integrates one FA and one SA circuit and operates independently of the other taxels. Spike trains are generated and transmitted off chip by FA and SA inspired circuits, both of which process the analog voltage produced by the tactile sensor. The biases of the encoding block are configured through a digital chain of programmable bits, enabling dynamic reconfiguration of the encoding parameters. Initial bias values were obtained through SPICE simulations and optimized for a representative range of input voltages; however, these parameters can be reprogrammed to accommodate different operating conditions. The FA mechanoreceptor is implemented using a circuit derived from the event-based pixel architecture [16], while the SA mechanoreceptor is implemented using a circuit derived from the DPI neuron [17]. This architecture enables the implementation of closed-loop control strategies, either by regulating the output event rate directly or by incorporating feedback from downstream processing stages to update the encoding parameters. Tunable parameters include those associated with the event-based pixel, such as threshold and refractory period, as well as parameters of the DPI neuron, including V_{thr}, V_{tau} for input amplification, V_{refr} for the neuronal refractory period, and parameters governing spike frequency adaptation mechanisms. A controller for active efficient coding relies on adjusting sensor parameters to make the encoding efficient and shape how information is encoded over time. Thus, it is essential to know in advance which parameters are most effective in tuning the sensor's response, which will be described in the next section.

2.2 Sensor Characterization

For controller design, there is a need to understand how its internal parameters affect the way signals are processed. Thereby, sensor characterization focuses not only on basic properties such as sensitivity, linearity, signal-to-noise ratio, but also on the influence on efficient encoding, i.e., the information content in the signal given by entropy measures, and the signal processing itself, analyzed by parameters like accuracy or confidence of sound/object recognition. Depending on the modality and the chosen sensor, several internal parameters might be adaptable to control its behavior. While this flexibility is useful, it is not practical to consider all parameters as controllable. Using many tunable parameters later on can increase system complexity, power consumption, and design effort, and can make it difficult to predict or explain system behavior. For this reason, relevant parameters have to be identified.

As an example how to identify relevant parameters, the strategy is discussed for the MEMS cochlea for hearing modality. The MEMS cochlea is composed of an active cantilever, which acts as a transducer with frequency filtering effect and whose properties can be controlled with electronic feedback. The cantilever voltage output is high-pass filtered and amplified, and then fed into an envelope detection circuit and a spike generation circuit based on a leaky-integrate-and-fire neuron. In each stage, several parameters can be tuned, thus affecting the sensor output. To cover a larger frequency range, an array of sensors with different geometries is required, each with its own controller.

For characterisation, a pure-tone stimulus at the sensors resonance frequency and different amplitudes was applied, while changing one parameter step-by-step and keeping all other parameters at their nominal values. Each recording included a pre-stimulus silence period, which was used to estimate the noise floor for subsequent signal-to-noise ratio calculations. Thus, a total of seven parameters were characterized under identical acoustic conditions: envelope gain and low-pass filter (LPF) cutoff frequency f_c (envelope generation circuit), high-pass filter (HPF) f_c (second-stage amplification), integration time constant and threshold voltage (spike detection circuit), and self-feedback gain a_f and DC bias u_{dc} (feedback circuit). Their effect on sensor performance is studied by analyzing signal-to-noise ratio, sensitivity, time constants (envelope signal), spike rate and offset, and mutual information between sensor signal and stimulus.

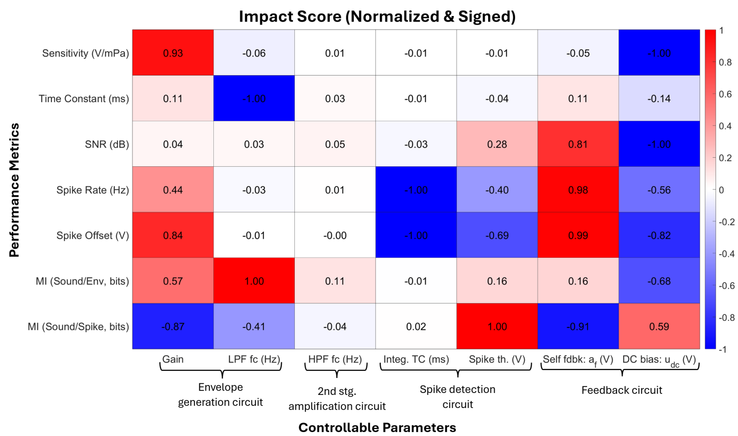


Fig. 3: Impact score matrix demonstrating the influence of each parameter (column label) on different sensor output properties (row label). LPF and HPF denote low-pass and high-pass filters with cut-off frequency f_c, TC denotes time constant, th. denotes spike threshold, fdbk denotes feedback, and MI denotes mutual information.

Based on the influence of each parameter on these output metrics, an impact score is computed as the signed, normalized end-to-end change across the tested range:

$$S_{i,j} = \frac{\Delta m_i(p_j)}{\max_k \left|\Delta m_i(p_k)\right|} \tag{1}$$

where $\Delta m_i(p_j) = m_i\left(p_j^{\max}\right) - m_i\left(p_j^{\min}\right)$ is the signed change in metric m_i when parameter p_j is swept from minimum to maximum value, normalized by the largest absolute change observed for metric m_i across all parameters. A score of ± 1.0 indicates that a parameter has the strongest influence on a given metric, scores close to zero indicate little or no effect.

The resulting impact score is shown in Fig. 3. It reveals, that the parameters fall into two groups. The feedback parameters a_f and u_{dc} show high impact scores across multiple metrics simultaneously, while parameters such as threshold and integration time show moderate impact on spike-domain metrics, but relatively small effects on metrics such as sensitivity or SNR. This suggests that the feedback parameters set the overall operating point of the sensor, while the spike circuit parameters offer more selective control over the encoding stage. However, in terms of spike output and mutual information, other parameters are important as well and can even have a negative impact. In particular, the envelope LPF cutoff frequency stands out as having negligible influence on spectral amplitude metrics (sensitivity, spike rate) but strong impact on time constant and mutual information. This suggests that it primarily shapes the temporal dynamics rather than steady-state levels. Whether this is relevant depends strongly on the specific application: for signals whose discriminative features are mainly amplitude-based (e.g. speech recognition), this parameter may be of secondary importance, whereas for tasks where temporal dynamics play a greater role, e.g. sound localization or sound event detection, it should be included as controllable parameter.

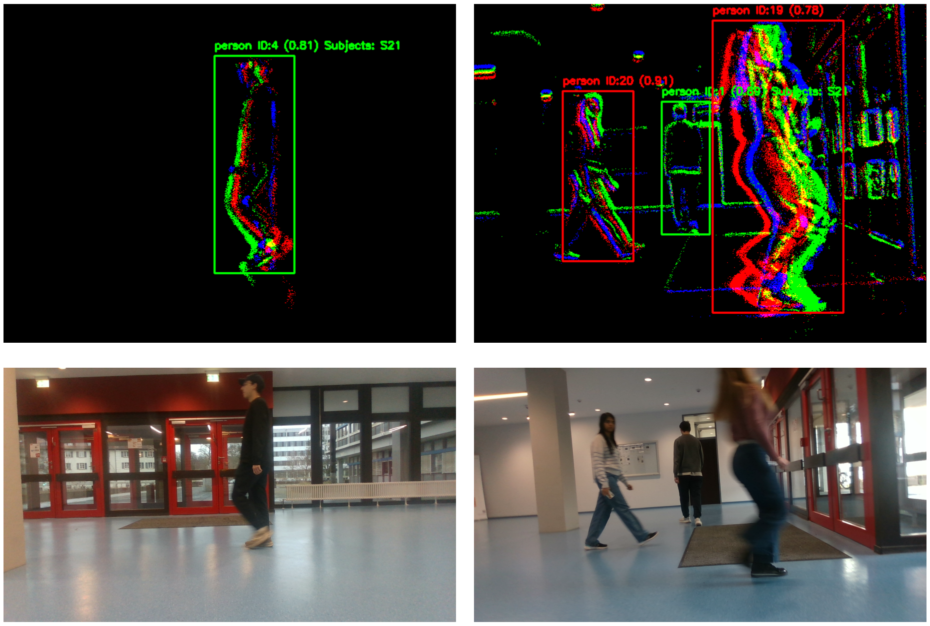
This characterization was carried out using a single pure-tone stimulus, which is a controlled but simplified input condition, while real-world acoustic environments involve broadband and non-stationary signals. Thus, the influence of a subset of parameters onto processing performance was studied for spoken digit recognition, with and without added noise, using a reservoir computing scheme for recognition [14]. Results show that the number of sensors and the feedback factor have the strongest impact on accuracy (up to 30%) while other parameters have only a negligible effect ($< 3\%$ in accuracy). However, for some speakers, changing gain is important, while under noisy conditions, a frequency-dependent bandwidth distribution can increase accuracy by 10%. Although both studies reflect only single-parameter effects without accounting for interactions between parameters, they provide a useful ranking of parameter importance and a structured basis for deciding which parameters to treat as controllable in the AEC controller, and which can be fixed at a default hardware setting. The example from the MEMS cochlea shows how to distinguish parameters which mainly affect output scaling from parameters that fundamentally change response timing, dynamics, and processing performance. The latter are particularly important for developing a controller to implement Active Efficient Coding.

3 SNN based processing

Besides the simple signal-based statistics, processing performance will be included in the controller as well. While for characterization of MEMS cochlea influence on performance, analog signal processing and artificial neural networks were applied, the final systems should use spiking neural networks (SNN) for processing. Due to its maturity, we first developed SNNs-based processing for an event-based camera mounted on the ANYmal D robot from ANYbiotics [13]. The tested task that the robot follows a target person based on walking-gait-based identification.

The proposed pipeline consists of two neural networks. The first network performs pedestrian detection and tracking using YOLOv8x. It detects individual pedestrians over time and accumulates their output bounding boxes. The second network is a three-layer spiking convolutional neural network (S-CNN), which takes temporally stacked cropped bounding boxes from the first network as input and performs person identification in an open-set scenario. In this setting, the individuals appearing during the test experiments are not present in the training set.

After training the S-CNN, we use its penultimate layer as an embedding function. From this layer, an L_2-normalized feature vector is extracted for each detected pedestrian. Before the live experiments, feature vectors of the target individuals are extracted and stored in a gallery. During deployment, the fea-



(a) (b)

Fig. 4: Live identification and following results for subject S21 in (a) with only one pedestrian walking and (b) with multiple pedestrians. Green bounding boxes indicate the tracked preselected target detected and identified by our system; red boxes denote other detected pedestrians. The first row are the event images used by our system, while the second row are the RGB images, which are shown only for visualization and are not used by the proposed algorithms.

ture vector of each detected pedestrian is compared against the gallery, and the identity is assigned according to the closest match. Qualitative results from the experiments are shown in Fig. 4.

3.1 Controller Design

The controller aims to establish a learnable mapping from the current sensor parameters (and therefore statistics) and task performance to the next possible sensor state. Its primary objective is to balance resource consumption for size- or power-sensitive applications with the desired performance on the objective task, conditioned on the current environmental stimuli [26]. Feedback loops influence the sensor's characteristics to provide optimal information for a given task in a specific environment, in contrast to fixed sensor transduction.

To exemplify the objectivity of our neuromorphic controller, object detection is considered as task. For this purpose, we use an off-the-shelf event-based object detector [10, 31]. However, a major hindrance in this approach is the changing

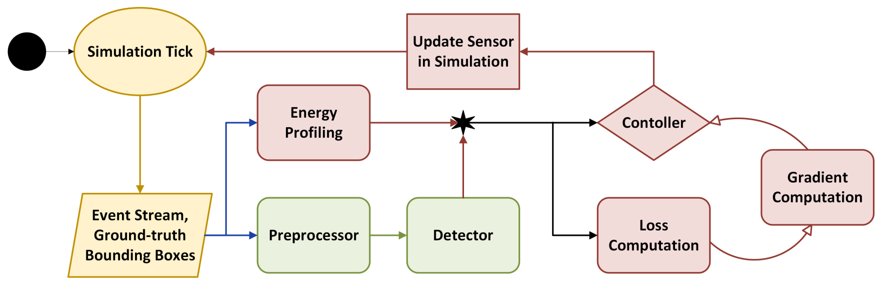
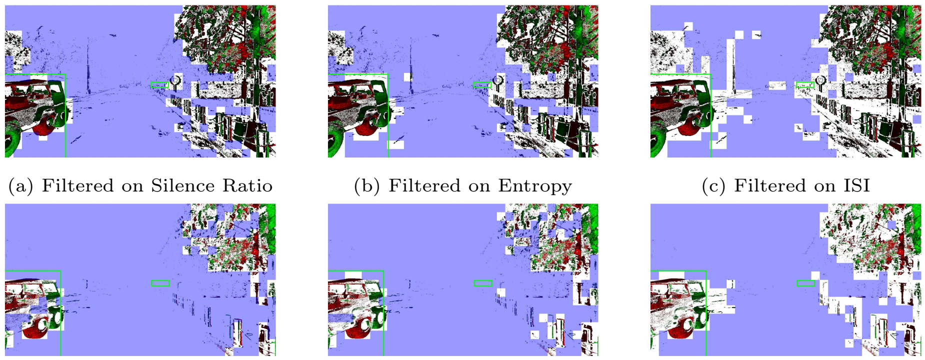


Fig. 5: An overview of the closed-loop simulation environment in CARLA. The preprocessor and detector are off-the-shelf models used for event-based task evaluation, while the controller is a policy-action-based RL model that is trained to optimize energy efficiency and task performance.

sensor input characteristics when the sensor parameters are adapted dynamically. If not handled effectively, this results in constant poor performance by the detector, which introduces a bias in our feedback loop. To ensure that the model is robust to a wide range of input characteristics, and thereby that the score is reflective of the environmental and sensory-adaptation requirements, we retrain and test the model with an expansive set of sensor configurations. This training across diverse sensor parameters enhances the models' robustness by leveraging multi-source domain generalisation and effectively mitigates performance degradation due to fluctuating sensor characteristics.

To validate the idea of adaptive control for a neuromorphic sensors, we consider here neuromorphic vision for vehicular object detection. Therefore, we set up a closed-loop simulation environment in CARLA [7]. At each simulation timestep, sensor readings are retrieved along with ground-truth bounding boxes. The sensor readings serve as input to the detector, whose results are compared against the ground-truth to generate a performance score. Additionally, spike statistics like the firing rate (FR), interspike intervals (ISI) [5] are computed from sensor readings, to characterize the efficiency of sensory transduction. These, alongside the performance metric, are forwarded to our sensory controller as inputs. The controller, in turn, provides the sensor parameter states for the next simulation step. An overview of the closed-loop control is provided in Fig. 5.

While the described sensory controller operates at the level of DVS sensor parameters, globally modulating event generation across the entire focal plane, a complementary and parallel spatial controller operates at the level of image patches. The two controllers act at different levels of the perception hierarchy: the sensory controller governs how the sensor generates events, while the patch controller governs where computational resources are directed within the scene. Together, they constitute a two-level closed-loop active perception system [1,2]. The motivation for patch-level control arises from a fundamental asymmetry in event camera scenes: at any given timestep, the spatial distribution of informa-



(a) Filtered on Silence Ratio (b) Filtered on Entropy (c) Filtered on ISI

Fig. 6: The same threshold, tuned manually on one sensor characteristic (top row), is applied to another (second row). Shaded regions denote patches that are masked. Note how a single spike modality is not sufficient to filter out the objects of interest while discarding redundant background data, even when tuned.

tive activity is highly non-uniform. Regions containing moving vehicles generate dense, structured spike trains, while static background regions produce sparse or silent responses. Forwarding all patches uniformly to the detector is therefore computationally wasteful and, in principle, detrimental; background patches contribute redundant data without an informative signal. Selective spatial allocation, informed by the statistical structure of the event stream, offers a principled route to improving the performance-efficiency tradeoff. Rather than operating on raw events, which are high-dimensional, simulator-specific, and difficult to transfer across hardware, the patch controller is driven by neuro-scientifically grounded spike statistics aggregated at the patch level. Concretely, for each spatial patch at each timestep, spike statistics including the silence ratio (ratio of pixels with a firing rate of 0), ISI, and entropy can be used to characterize the local spiking regime, distinguishing active foreground and silent background in a hardware-invariant manner [5].

Analogously to the sensory controller, the patch controller operates in a closed loop: spike statistics from the current timestep and detector confidence from the previous timestep jointly constitute the controller state. Detector feedback is used as an online observation that informs the controller of whether the current patch selection is sufficient for the task, enabling the system to tighten or relax spatial selectivity dynamically. This closed-loop structure distinguishes the proposed framework from prior work on adaptive or dynamic inference [21, 29], which typically operate feedforward without explicit task-performance feedback.

As a precursor to the full closed-loop patch controller, we conduct a series of controlled experiments in CARLA. Fig. 6 demonstrates that static, manually tuned thresholds, while partially effective, are brittle across scene conditions, thereby motivating the need for adaptive control. The poor performance, even on the tuned settings, demonstrates the need for a weighted combination of different statistics to fully utilize their potential in reducing information redundancy, while ensuring stability in performance on the objective.

3.2 System Integration

Neuromorphic sensors are increasingly being used in robotic applications. For example, event-based cameras enable fully neuromorphic vision and control for autonomous drone flight [20], neuromorphic auditory sensors support sound source localization [23], and event-based tactile sensors enhance robotic manipulation by detecting slip [9]. However, like humans, relying on a single sensory modality is rarely sufficient for robots. Vision alone cannot capture texture or surface compliance; audio cannot reveal shape; and touch is limited to local interactions. Therefore, multimodal sensing systems are necessary.

Some recent work has begun to explore multimodal neuromorphic sensing. For instance, [25] combined an event-based camera with two biologically inspired tactile sensors on a robotic manipulator to classify liquid containers with different fill levels. In their approach, two spiking neural networks (SNNs) independently encoded visual and tactile information into spike trains, which were subsequently concatenated and processed by a task SNN to infer both container type and weight. The experimental results showed that the multimodal system achieved better performance compared to systems using individual sensory modalities in isolation. Despite such promising results, existing systems that typically integrate different neuromorphic modalities are still unexplored. By combining event-driven vision, audition, and tactile sensing, multimodal neuromorphic perception becomes possible, enabling robots to perceive the world in a more biologically inspired manner by integrating complementary streams of sensory events to build a richer and more robust understanding of their surroundings.

To demonstrate the power of such sensor integration, this project proposes a multimodal event-based sensory system comprising an event-based camera, a MEMS cochlea, and event-driven tactile sensors, integrated into an ANYmal quadruped robot [13], which can recognize different terrains and select an appropriate walking mode. The tactile sensors are mounted on a stick, allowing the robot to interact with the environment while simultaneously acquiring tactile feedback. During interaction, the robot applies controlled forces to the stick, inducing changes in the terrain such as deformation, displacement, motion-induced sounds, or texture variations. The event-based sensors continuously capture sensory events, providing rich information about the properties of the terrain. However, not all captured events are useful. Irrelevant signals, such as background noise, external sounds, bouncing phenomena, and especially ego-events resulting from the robot's own motion, can obscure meaningful data. To solve this problem, a preprocessing pipeline is employed, including filtering and debouncing, to suppress noise and ego-events while retaining task-relevant information.

After preprocessing, the refined event streams are directly fed into SNNs. Each sensory modality is processed by a dedicated SNN that learns to extract meaningful spatiotemporal patterns from the raw event data. These networks

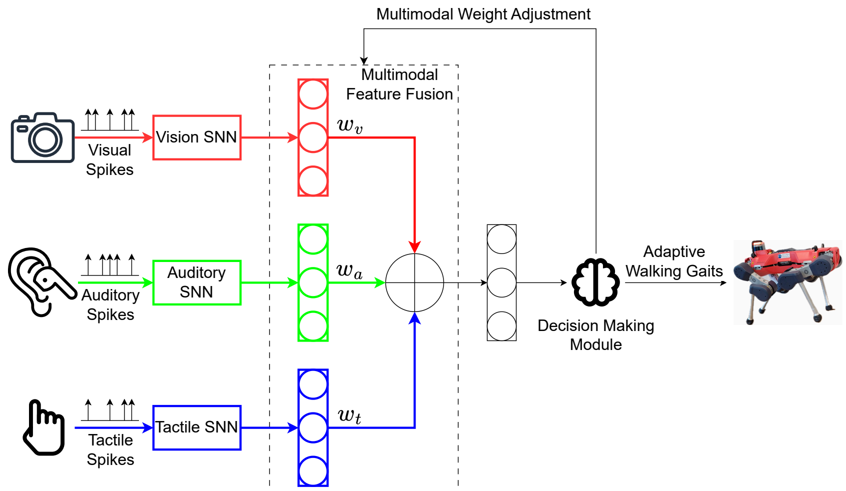


Fig. 7: Overview of the multimodal neuromorphic sensory system integrated into a quadruped robot for terrain recognition and adaptive locomotion selection.

generate spikes encoding features such as motion dynamics, surface textures, and acoustic characteristics derived from visual, tactile, and auditory inputs, respectively. The resulting spike representations capture key properties of the interacted terrain. The modality-specific spikes are then forwarded to a fusion stage, where multimodal information is integrated and dynamically modulated. Finally, the fused representation is processed by a decision-making SNN that determines the terrain type and selects an appropriate locomotion strategy. During operation, the system automatically adjusts the relative weighting of each sensory modality through an outer loop controller based on environmental conditions and the current reliability and performance of each sensor. The processing pipeline is illustrated in Fig. 7.

4 Conclusion

In this work, we presented a system design for building a neuromorphic perception based on multiple modalities and implementing active, efficient coding. In particular, we demonstrated the design of adaptive tactile sensors, the analysis of sensor parameters most important for efficiency and processing for the MEMS cochlea to support controller design, the development of a neuromorphic controller to implement active, efficient coding using global or path-based sensor adaptation, the processing system based on SNNs and the system setup to integrate the input from different modalities for the intended use case.

Acknowledgments. This work is partially funded by the project *SNACE* of the Volkswagen Foundation under the profile area *Exploration* and the funding framework *NEXT: Neuromorphic Computing*.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aloimonos, J., Weiss, I., Bandyopadhyay, A.: Active vision. International journal of computer vision **1**(4), 333–356 (1988) 9
2. Bajcsy, R.: Active perception. Proceedings of the IEEE **76**(8), 966–1005 (1988) 9
3. Barlow, H.B., et al.: Possible principles underlying the transformation of sensory messages. Sensory communication **1**(01), 217–233 (1961) 2
4. Dahiya, R.S., Metta, G., Valle, M., Sandini, G.: Tactile Sensing—From Humans to Humanoids. IEEE Transactions on Robotics **26**(1), 1–20 (Feb 2010). https://doi.org/10.1109/TRO.2009.2033627 4
5. Dayan, P., Abbott, L.F., et al.: Theoretical neuroscience, vol. 806 (2001) 9, 10
6. Delbruck, T., Graca, R., Paluch, M.: Feedback control of event cameras. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 1324–1332. IEEE, Nashville, TN, USA (Jun 2021). https://doi.org/10.1109/CVPRW53098.2021.00146 1, 2, 4
7. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V.: Carla: An open urban driving simulator. In: Conference on robot learning. pp. 1–16. PMLR (2017) 9
8. Eckmann, S., Klimmasch, L., Shi, B.E., Triesch, J.: 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). https://doi.org/10.1073/pnas.1908100117 2
9. Funk, N., Helmut, E., Chalvatzaki, G., Calandra, R., Peters, J.: Evetac: An event-based optical tactile sensor for robotic manipulation. IEEE Transactions on Robotics **40**, 3812–3832 (2024). https://doi.org/10.1109/TRO.2024.3428430 11
10. Gehrig, M., Scaramuzza, D.: Recurrent vision transformers for object detection with event cameras. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 13884–13893 (2023) 8
11. Gibson, J.J.: The senses considered as perceptual systems. (1966) 2
12. Haimerl, C., Ruff, D.A., Cohen, M.R., Savin, C., Simoncelli, E.P.: Targeted V1 comodulation supports task-adaptive sensory decisions. Nature Communications **14**(1), 7879 (Nov 2023). https://doi.org/10.1038/s41467-023-43432-7 3
13. Hutter, M., Gehring, C., Jud, D., Lauber, A., Bellicoso, C.D., Tsounis, V., Hwangbo, J., Bodie, K., Fankhauser, P., Bloesch, M., Diethelm, R., Bachmann, S., Melzer, A., Hoepflinger, M.: ANYmal - a highly mobile and dynamic quadrupedal robot. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 38–44. IEEE, Daejeon, South Korea (Oct 2016). https://doi.org/10.1109/IROS.2016.7758092, http://ieeexplore.ieee.org/document/7758092/ 7, 11
14. Johny, S., Ved, K., Durstewitz, S., Ivanov, T., Ziegler, M., Lenk, C.: Bio-inspired sensor-based sound pre-processing for speech recognition in noisy conditions. In: 2024 IEEE Biomedical Circuits and Systems Conference (BioCAS). pp. 1–5 (2024). https://doi.org/10.1109/BioCAS61083.2024.10798207 7

15. Lenk, C., Hövel, P., Ved, K., Durstewitz, S., Meurer, T., Fritsch, T., Männchen, A., Küller, J., Beer, D., Ivanov, T., Ziegler, M.: Neuromorphic acoustic sensing using an adaptive microelectromechanical cochlea with integrated feedback. Nature Electronics **6**, 370–380 (2023) 2, 3
16. Lichtsteiner, P., Posch, C., Delbruck, T.: A 128x128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor. IEEE Journal of Solid-State Circuits **43**(2), 566–576 (Feb 2008). https://doi.org/10.1109/JSSC.2007.914337 4
17. Livi, P., Indiveri, G.: A current-mode conductance-based silicon neuron for address-event neuromorphic systems. In: 2009 IEEE International Symposium on Circuits and Systems. pp. 2898–2901 (May 2009). https://doi.org/10.1109/ISCAS.2009.5118408 4
18. Nair, G.B., Milford, M., Fischer, T.: Enhancing Visual Place Recognition via Fast and Slow Adaptive Biasing in Event Cameras. In: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 3356–3363 (Oct 2024). https://doi.org/10.1109/IROS58592.2024.10802384 4
19. Niven, J.E.: Neural energetics: Hungry flies turn down the visual gain. Current Biology **24**(8), R313–R315 (2014). https://doi.org/https://doi.org/10.1016/j.cub.2014.03.013, https://www.sciencedirect.com/science/article/pii/S0960982214002723 1
20. Paredes-Vallés, F., Hagenaars, J.J., Dupeyroux, J., Stroobants, S., Xu, Y., De Croon, G.C.H.E.: Fully neuromorphic vision and control for autonomous drone flight. Science Robotics **9**(90), eadi0591 (2024). https://doi.org/10.1126/scirobotics.adi0591, https://www.science.org/doi/10.1126/scirobotics.adi0591 11
21. Peng, Y., Li, H., Zhang, Y., Sun, X., Wu, F.: Scene adaptive sparse transformer for event-based object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16794–16804 (2024) 10
22. Posch, C., Serrano-Gotarredona, T., Linares-Barranco, B., Delbruck, T.: Retinomorphic event-based vision sensors: bioinspired cameras with spiking output. Proceedings of the IEEE **102**(10), 1470–1484 (2014) 2
23. Schoepe, T., Gutierrez-galan, D., Dominguez-Morales, J.P., Greatorex, H., Chicca, E., Linares-Barranco, A.: Event-based sound source localization in neuromorphic systems. https://doi.org/10.36227/techrxiv.21493290.v1, https://www.techrxiv.org/doi/full/10.36227/techrxiv.21493290.v1 11
24. Shamma, S., Fritz, J.: Adaptive auditory computations. Current Opinion in Neurobiology **25**, 164–168 (2014). https://doi.org/https://doi.org/10.1016/j.conb.2014.01.011 1
25. Taunyazoz, T., Sng, W., See, H.H., Lim, B., Kuan, J., Ansari, A.F., Tee, B., Soh, H.: Event-driven visual-tactile sensing and learning for robots. In: Proceedings of Robotics: Science and Systems (July 2020) 11
26. Teulière, C., Forestier, S., Lonini, L., Zhang, C., Zhao, Y., Shi, B., Triesch, J.: Self-calibrating smooth pursuit through active efficient coding. Robotics and Autonomous Systems **71**, 3–12 (2015) 2, 8
27. Vanattou-Saïfoudine, N., Han, C., Krause, R., Vasilaki, E., von der Behrens, W., Indiveri, G.: A robust model of stimulus-specific adaptation validated on neuromorphic hardware. Scientific Reports **11**, 17904 (2021). https://doi.org/https://doi.org/10.1038/s41598-021-97217-3 2
28. Vasco, V., Glover, A., Mueggler, E., Scaramuzza, D., Et., A.: Independent motion detection with event-driven cameras. 2017 18th International Conference on Advanced Robotics (ICAR) (2017). https://doi.org/10.1109/ICAR.2017.8023661 3

29. Wang, Y., Huang, R., Song, S., Huang, Z., Huang, G.: Not all images are worth 16x16 words: Dynamic transformers for efficient image recognition. Advances in neural information processing systems **34**, 11960–11973 (2021) 10
30. Willmore, B.D.B., King, A.J.: Adaptation in auditory processing. Physiological Reviews **103**(2), 1025–1058 (2023). https://doi.org/10.1152/physrev.00011.2022 1
31. Zubic, N., Gehrig, M., Scaramuzza, D.: State space models for event cameras. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5819–5828 (2024) 8