2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Neuromorphic Event Alarm Time-Series Suppression

Shane Harrigan
SCEIS-CRL
Ulster University
Derry, UK
sp.harrigan@ulster.ac.uk

Sonya Coleman
SCEIS-CRL
Ulster University
Derry, UK
sa.coleman@ulster.ac.uk

Dermot Kerr
SCEIS-CRL
Ulster University
Derry, UK
d.kerr@ulster.ac.uk

Justin Quinn
SCEIS-CRL
Ulster University
Derry, UK
jp.quinn@ulster.ac.uk

Kyle Madden
SCEIS-CRL
Ulster University
Derry, UK
k.madden@ulster.ac.uk

Leeanne Lindsay
SCEIS-CRL
Ulster University
Derry, UK
l.lindsay@ulster.ac.uk

Benn Henderson
SCEIS-CRL
Ulster University
Derry, UK
b.henderson@ulster.ac.uk

Shammi Rahman
SCEIS-CRL
Ulster University
Derry, UK
s.rahman@ulster.ac.uk

***Abstract*—The field of neuromorphic vision systems aims to replicate the functionality of biological visual systems by mimicking their physical structure and electrical behaviour. Unlike traditional full-frame sensors, neuromorphic systems process data asynchronously and at the pixel level, modelling biological signalling processes. This allows for high-speed operations with lower energy consumption, making them suitable for applications like autonomous vehicles and embedded robotics. This work introduces the Neuromorphic Event Alarm Time-Series Suppression (NEATS) framework, designed to filter noise and detect outlier behaviours in event data without the need for 2-D transformations. NEATS employs rolling statistics and advanced neuromorphic data structures to minimise noise while identifying changes in scene dynamics. This framework injects attention into scene processing, similar to summarisation frameworks in traditional image processing. A novel event-vision alarm change collection (EACC) database is presented, containing controlled stimuli pattern changes captured using leading neuromorphic imaging devices. This database facilitates future benchmarking of neuromorphic attention frameworks, advancing the development of efficient and accurate artificial vision systems.**

***Index Terms*—Bio-inspired, Signal Processing, Pattern Recognition**

I. Introduction

Biologically-inspired artificial vision systems, or neuromorphic vision systems [3], [6], [12], are designed to emulate the function of modelled biological visual systems by mimicking the their physical structure and electrical behaviour. Classic full-frame sensors, or active-pixel sensors (APS), work by polling a 2-D array of pixels at a certain time interval (static or dynamic) and digitising the data values of all pixels during the polling to produce a numerical representation of the light levels detected at the polling moment. When displayed on a 2-D surface, these data values will contain spatial information at the moment of capture which can be processed further by classical image processing techniques. Taking a series of 2-D matrices formed from the integration of events time (with events behaving as a Dirac delta singleton), and represented as a sequence from oldest captured to newest, we can develop a representation of change (motion image) within an observed scene.

The classical full-frame approach is well understood and is the foundation of most imaging research to date but it has disadvantages in terms of time and energy since its formalisation. The power-to-speed ratio is directly linear such that in order to poll and digitise scene information at high-speeds, we always need to increase the power consumption in the vision system; conversely to decrease the power consumption we need to decrease the polling and digitisation slowing the capturing process. This ratio is acceptable in a number of existing areas but some key areas of research, such as autonomous vehicles and general embedded robotics, require more responsible energy usage while maintaining high-speeds.

Consider a time-series E of N event data such that $E_i = \langle t_i, \langle x_i, y_i \rangle, p_i \rangle$ where $i = [0, N)$, t represents relative time, $\langle x, y \rangle$ is a spatial identifier, and $p = \pm 1$ is a polarity value indicating the luminance directional change. Unlike classical full-frames which are depending on an integration or synchronisation time, neuromorphic data are asynchronous as all pixels are independent of each other and model biological signalling processes. This architecture allows neuromorphic sensors to operate at microsecond domains for each individual pixel with the speed impeded by the data transmission bus. It is popular at present within neuromorphic research to translate the time-series E onto a 2-D surface to allow for classical image processing techniques to be utilised [2], [8], [9], [12]. This is achieved by reintroducing integration over time; given enough time a similar 2-D matrix like a full-frame can be produced. Integrating neuromorphic signals overtime can still be faster than the full-frame methods but they often consist of higher levels of noise (blurring, system level etc.) and

Funded by Innovate UK through the Smart Manufacturing Data Hub (SMDH) project.

an obvious decrease in speed. Some research [14], [15] has shown that it is possible to evaluate spatial information as it is initially processed rather than transforming all event data simultaneously, and other methods [13] have been introduced which do not rely on transformation for rich feature extraction.

Event data are known to be advantageously sparse when compared with frame-based representations but being sparse also leads to higher susceptibility to noise, radiating from the sensing array, for event data leveraging pipelines. In this paper we introduce the neuromorphic event alarm time-series suppression (NEATS) framework. NEATS is designed to act both as a noise filtering process and as a neuromorphic data alarm system for identifying outlier behaviours in event data without the requirement of 2-D transformations. The NEATS framework is designed to minimise system and environmental noise, using a rolling statistics model and a state-of-the-art neuromorphic data structure [14], while acting as a pattern alarm system which can be used to indicate when the underlying scene dynamics change. In essence the NEATS framework is designed to inject attention as a quantitative element of scene processing. Attention frameworks in neuromorphic vision are closely related to summarisation frameworks in frame-based approaches [10] with a commonality being the use of 3-D structures [7], [11], [17] to process 2-D (spatial) information over time sequences; sequencing naturally leads to integration-over-time slowing the speed of neuromorphic systems down.

Since NEATS is the first framework of its kind, we present an event vision database called the event-vision alarm change collection (EACC) database which is a novel collection of controlled stimuli pattern change captured using leading neuromorphic imaging devices. The database is designed to allow for future benchmarking of neuromorphic attention frameworks and will be open source.

II. Methodology

In this section we outline the NEATS algorithm in terms of theory and pseudo-implementation (Section II-A). Additionally we elaborate on the framework for using NEATS (Section II-B) as utilised during the experiments outlined in Section IV.

A. NEATS Algorithm

The neuromorphic event alarm time-series suppression is designed to minimise, or filter, the noise within a scene while allowing actual scene activity (called actions) to be retained and key pattern changing events to be flagged for evaluation. The NEATS filter works by first expressing event data as a series of energy ψ such that

$$\psi = \|\mathbb{E}\| \tag{1}$$

for $\mathbb{E} \subset E$ such that $0 \leq t \leq T$ where T represents a specific point in time. The energy ψ is also expressible at pixel level $\psi(\hat{x}, \hat{y}, T)$ such that

$$\psi(\hat{x}, \hat{y}, T) = \|\mathbb{E}\| \mid \forall \mathbb{E} \text{ where } \hat{x} = x, \hat{y} = y \tag{2}$$

It is common practice to convert $\mathbb{E}$ into a frame with integration-over-time, which is a popular technique, such that an event-frame $eF \equiv \mathbb{R}^3$ is expressible as

$$eF(\hat{x}, \hat{y}, T) = \int_a^T \psi(\hat{x}, \hat{y}, t)\ dt \tag{3}$$

where each pixel $eF(\hat{x}, \hat{y})$ is equal to the behaviour of $\psi(\hat{x}, \hat{y}, T)$ at pixel level for all events not exceeding T and $a = 0$. For NEATS we are interested in the delta energy $\Delta\mathbb{F}$ of a scene, that is the difference between the current energy and the previously observed energy, expressed as

$$\Delta\mathbb{F}_i(\psi) = \int_{i-1}^{i} \psi(\hat{x}, \hat{y}, a)\ dt \tag{4}$$

or more succinctly as

$$\Delta\mathbb{F}_i(\psi) = eF(\hat{x}, \hat{y}, i) - eF(\hat{x}, \hat{y}, i-1) \tag{5}$$

Within this manuscript, we are interested in the delta values of our scene energy and use these to form a rolling time-series D such that

$$D = \langle \Delta\mathbb{F}_1(\psi), \Delta\mathbb{F}_2(\psi), \ldots, \Delta\mathbb{F}_{\|F\|}(\psi) \rangle \tag{6}$$

where F is the number of samples possible within our rolling window. The produced D time-series is an unlabelled data series containing sequential representative energy values of a scene.

In NEATS, the time-series D is operated over by a calculated p-value to produce a power-Martingale value (an assumption-based betting strategy) [4] and defined as

$$M_n^{(e)} = \prod_{i=1}^{n} (\epsilon p_i^{\epsilon - 1}) \tag{7}$$

where p_i is computed as

$$p_i = \frac{\#\{j : D_j > D_i + \theta_i \#\{j : D_j = D_i\}}{i} \tag{8}$$

with $\theta_i = [0, 1]$, $j = [1, 2, 3, \ldots, i]$ and $\#$ represents the cardinality of the data set. In [4] the current power-Martingale can be expressed as the current operating p-value factored by the previous power-Martingale value of the last p-value such that

$$M_n^{(e)} = \epsilon p_i^{\epsilon - 1} M_{n-1}^{(e)} \tag{9}$$

and we establish a special rule where $M_1 = 1$ as a cost of preventing time-series vanishing propagation. Taking our new time-series as M we compute the z-score equivalent over M, resulting in a normalised M such that $\|M\| = \mathbb{M}$ in z-space. The advantage of using z-space is that a value compared within a z distribution of data can be considered to be normal provided $-3 \leq z \leq 3$; in NEATS the compliment $\hat{\mathbb{M}}$ of normal $(-3 > zandz > 3)$ can be used to identify $\mathbb{M}$ values that are candidates for abnormal behaviour, and thus candidates for actions within a scene.

To summarise, the NEATS algorithm accepts a collection of event data $\mathbb{E}$ and computes time regions within the collection in which the form of highlighted abnormal data $\hat{\mathbb{M}} > 3$ or

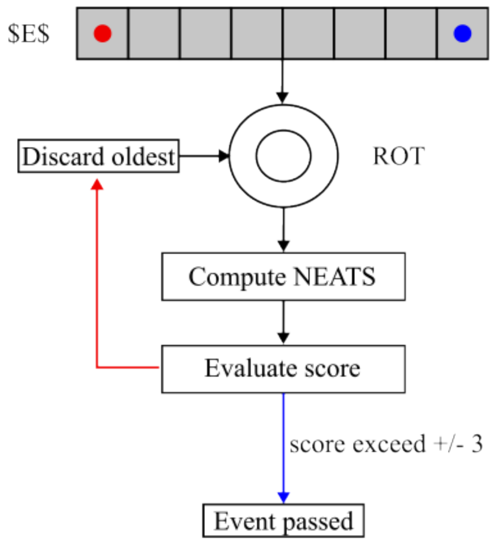


Fig. 1: An illustration of the NEATS framework showing the usage of an ROT tree as the rolling engine of the framework. An example of how the oldest (red) and newest (blue) event data is treated by the algorithm is also shown.

$\hat{\mathbb{M}} < -3$. In practice we compute $|\hat{\mathbb{M}}|$ such that thresholding follows $|\hat{\mathbb{M}}| > 3$. However, we could extrapolate the z-space polarity of energy over time, for use in more advanced photo-event algorithms.

B. NEATS Framework

The NEATS framework is based on a linearly devolved Reducing-Over-Time (ROT) tree [15] (similar to a ring buffer) R. A ROT tree is a data structure engineered for event-data and has been shown to represent event data in a rapidly accessible and processor-friendly manner. ROT trees operate by automatically keeping event data balanced within a binary search tree structure while automatically identifying, selecting and pruning nodes which fall outside of "best by" timestamp threshold τ range which can be static or dynamic in nature. In the case of NEATS, if an event E_i is added to R such that it causes $|M_i| > 3$ we consider this event to be abnormal from the background data and therefore meets our definition of an action event. As NEATS evolves R can be thought of as a revolving attention window which retains recent past events. As R fills, the oldest events are overwritten with newer events. Algorithm 1 outlines the NEATS framework procedurally.

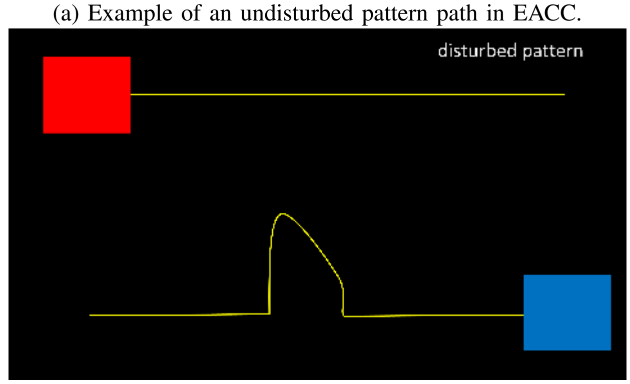
III. Database

In this section we discuss the novel event-vision alarm change collection (EACC) database capture and utilised in this paper. The EACC (version 1) database is a collection of over 10000 recordings of various stimuli captured using a DAVIS346 housed in a darkroom with true-colour calibration on an LCD monitor. Each recording contains a controlled stimuli projected onto the screen with a predetermined motion pattern. At documented intervals of time the pattern is temporarily randomised to induce Brownian behaviours within the event data. The database can be split into two groups based on motion pattern drawing: the first half consists of EACC recordings where stimuli motion was projected at the full screen-rate capacity of the LCD display. The second half of the EACC recordings are captured using the popular strobing methodology in which the projected motion pattern is refreshed at

Algorithm 1 NEATS Algorithm

Require: R, $E \neq \varnothing$
Ensure: $M_1 = 1$
$i \leftarrow 1$
while $i <$ some condition **do**
  add E_i to R {overwrite oldest}
  compute $\hat{M}_i$ using Eq. (9) over R
  $j \leftarrow |\hat{M}_i|$
  if $j > 3$ **then**
    release E_i
  end if
  $i \leftarrow i + 1$
end while=0

(a) Example of an undisturbed pattern path in EACC.



(b) Example of an disturbed pattern path in EACC.

Fig. 2: (2a) shows the pattern path of an undisturbed scenario. (2b) shows the pattern path of a disturbed scenario. Pattern paths are denoted as a yellow line.

Fig. 3: The setup used to collect EACC data (this does not reflect the true capture conditions which were automated, in a darkroom with no human attendance).

a lower pace causing a strobing effect to appear stimulating the camera sensor. The stimuli itself involves a collection of basic shapes (rectangles, triangles, circles etc.) which are given the appearance of motion using the psychtoolbox [1] which is a reliable stimuli generation tool used across multiple fields. Each recording is accompanied with metadata including extracted frames, events, and timestamp files reflecting when pattern was disturbed relative to the camera timestamps. Figure 2 shows two example scenarios recorded in EACC, Figure 2a shows an example of an undisturbed pattern path in EACC. Figure 2b shows an example of a disturbed pattern path in EACC. In both figures the pathing is denoted as a yellow line.

Figure 3 shows the capturing setup used to build EACC using a DAVIS346; it should be noted that actual recording sessions were automated and took place in a darkroom. Utilising EACC will enable alarm frameworks based on event data to be benchmarked; by noting the time of pattern disturbance and the resumption of normal patterns we are able to designed specific temporal segments in which an alarm framework needs to trigger. Figure 4 shows a sample from the NEATS framework over a splice of EACC data. The M value is represented by the blue line, pattern disturbances are marked with dashed vertical black lines, and alarms generated by NEATS are shown as red 'x' symbols.

IV. Experiment and Results

The NEATS framework is a first-of-its-kind event-based alarm framework which leverages the neuromorphic ROT data structure for data handling and operates over the EACC database. As both NEATS and EACC are new and state-of-the-art, it is difficult to evaluate the NEATS framework; in this paper we offer the first usage of the EACC database as a benchmarking tool of neuromorphic alarm frameworks by reporting the following statistics gained from NEATS operating over EACC:

1) Percentage of events removed in total as noise/old information.
2) Percentage of true alarms triggered.
3) Percentage of false no-alarms.
4) Average time to process a single unit of data.

We are particularly interested in statistic measures 2 and 3 which report the percentage of true and false no-alarms generated by the NEATS framework as a true positive rate (tPR) and a false negative rate (fNR). A tPR is when the NEATS framework positively signals an alarm during a pattern disturbance period; an fNR is a failure of the NEATS framework to successfully signal an alarm during a pattern disturbance period. Given the documented disturbance times, which record both commencement and cessation of abnormal pattern behaviour, we can treat the start and end of the disturbance times as a indicator of when an alarm should be raised; we classify tPR as any alarm which is reported between the start and end of a disturbance. The fNR reports the number of disturbance areas which saw no alarm reported by the framework (the framework was unable to recognise the need for an alarm to be triggered).

TABLE I: Evaluation Results for NEATS Framework

Statistic (Average)	**Value**
1) Events Removed (Noise/Old Info)	73.82%
2) True Alarms (tPR)	83.11%
3) False No-Alarms (fNR)	16.89%
4) Time per Data Unit (nSec)	247nS

Table I reports the results of the NEATS framework collected while operating over the EACC database. The events removed metric indicates that the NEATS framework filtered out approximately 74% of event data (which is a large percentage of a "sparse" dataset) while operating over data recordings, this is inline with pairing filtering capabilities of the ROT tree [14], which is the core data structure of the framework, with the selective properties that the NEATS probabilistic evaluation of neuromorphic data provides; despite removing $\frac{3}{4}$ of the neuromorphic data the NEATS framework was capable of an approximately 83% positive identification of pattern change while operating over the database. Inversely the NEATS algorithm produces 17% false no-alarm (tied to the tPR) and took only 247 nanoseconds to achieve decision from data input to labelling. These results indicate that the NEATS framework, while filtering over 70% of the database on average, is capable of correctly identifying when the motion pattern of a scene changes by approximately $\frac{4}{5}$ of the time while still operating in the nanosecond domain.

V. Conclusion

We present the NEATS, or neuromorphic event alarm time-series suppression, algorithm and accompanying framework. NEATS is a filtering algorithm and accompanying pattern observation framework for determining if neuromorphic data belong to an action/pattern or system level noise without transformation to 2-D. Unlike other popular neuromorphic data

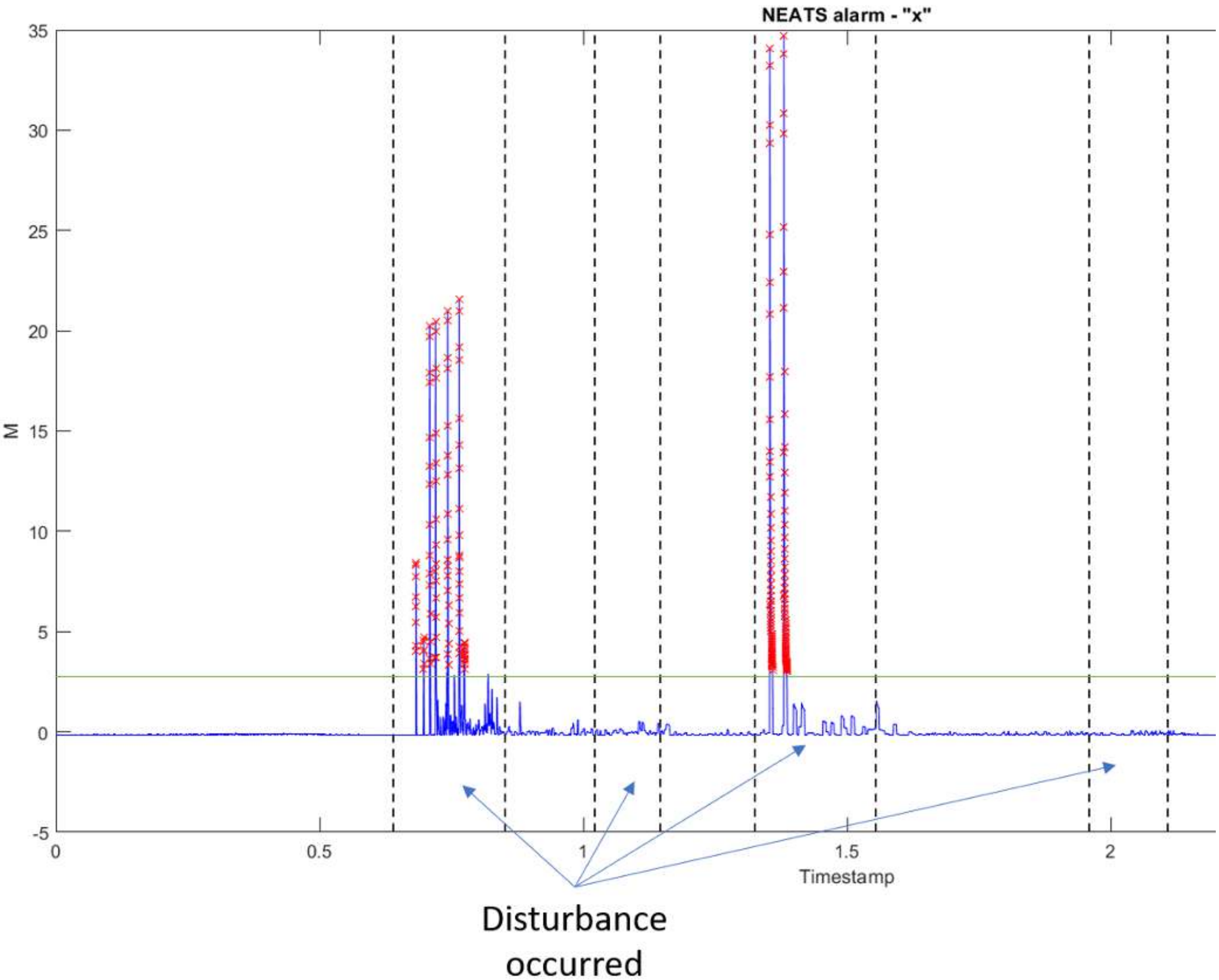


Fig. 4: A sample output of the NEATS framework operating over EACC data in nanosecond time showing triggered alarms (red 'x'), disturbed pattern occurrences (black dashed lines) and the output M values (blue line). The green line indicates the 3 standard deviation z-space threshold.

focussed algorithms and models, which handle noise filtering over a 2-D visual surface, NEAT is capable of working in 1-D requiring no spatial transformation as is common in leading techniques. It achieves this by introducing the concept of attention into the data stream. The NEATS algorithm is also sensor invariant; the scope of this paper is vision sensors but all neuromorphic data streams (for example the popular silicon cochlea [5]) produce data which NEAT can operate over allowing for a wide range of applications.

Furthermore we hypothesis that the NEATS algorithm and framework can be summarised using machine learning practices reducing the stages into a $1:1$ relationship between input and output. In future works we will explore this neural-version of NEATS (styled as N-NEATS), it is likely that N-NEATS will result in a speed-up of the NEATS framework filtering capabilities. Additionally we hypothesis that the ROT-Harris algorithm [14] can be modified to leverage the high filtering capabilities of NEATS to compute corner points. An interesting topic of further research is the time-to-first-label (similar to the time-to-first-spike [16] of neuroscience), as observable in Figure 4 where pattern changes are detected at varying times and with distinctive patterning, it maybe be possible to further explore these discrepancies in the future to extract richer features from NEATS operated data sets.

References

[1] M. Kleiner, D. Brainard, and D. Pelli, "What's new in psychtoolbox-3?," 2007.

[2] T. Delbruck, "Frame-free dynamic digital vision," in *Proceedings of Intl. Symp. on Secure-Life Electronics, Advanced Electronics for Quality Life and Society*, Citeseer, vol. 1, 2008, pp. 21–26.

[3] P. Lichtsteiner, C. Posch, and T. Delbruck, "A $128\times$times 128×120 dB $15mu$s latency asynchronous temporal contrast vision sensor," *IEEE journal of solid-state circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[4] S.-S. Ho and H. Wechsler, "A martingale framework for detecting changes in data streams by testing exchangeability," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2113–2127, 2010.

[5] S.-C. Liu, A. Van Schaik, B. A. Minch, and T. Delbruck, "Event-based 64-channel binaural silicon cochlea with q enhancement mechanisms," in *2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, 2010, pp. 2027–2030.

[6] T. Serrano-Gotarredona and B. Linares-Barranco, "A 128×128 1.5% contrast sensitivity 0.9% fpn 3 μs latency 4 mw asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 3, pp. 827–838, 2013.

[7] P. Bardow, A. J. Davison, and S. Leutenegger, "Simultaneous optical flow and intensity estimation from an event camera," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 884–892.

[8] X. Lagorce, G. Orchard, F. Galluppi, B. E. Shi, and R. B. Benosman, "Hots: A hierarchy of event-based time-surfaces for pattern recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 7, pp. 1346–1359, 2016.

[9] I. Alzugaray and M. Chli, "Asynchronous corner detection and tracking for event cameras in real time," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3177–3184, 2018.

[10] L. Annamalai, A. Chakraborty, and C. S. Thakur, "Evan: Neuromorphic event-based anomaly detection," *arXiv preprint arXiv:1911.09722*, 2019.

[11] L. Wang, Y.-S. Ho, K.-J. Yoon, *et al.*, "Event-based high dynamic range image and very high frame rate video generation using conditional generative adversarial networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 081–10 090.

[12] G. Gallego, T. Delbrück, G. Orchard, *et al.*, "Event-based vision: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 1, pp. 154–180, 2020.

[13] S. Harrigan, S. Coleman, D. Kerr, P. Yogarajah, Z. Fang, and C. Wu, "Neural coding strategies for event-based vision data," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020, pp. 2468–2472.

[14] S. Harrigan, S. Coleman, D. Ker, P. Yogarajah, Z. Fang, and C. Wu, "Rot-harris: A dynamic approach to asynchronous interest point detection," in *2021 17th International Conference on Machine Vision and Applications (MVA)*, IEEE, 2021, pp. 1–6.

[15] S. Harrigan, S. Coleman, D. Kerr, P. Yogarajah, Z. Fang, and C. Wu, "Reducing-over-time tree for event-based data," in *2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, 2021, pp. 1765–1772.

[16] L. Bonilla, J. Gautrais, S. Thorpe, and T. Masquelier, "Analyzing time-to-first-spike coding schemes: A theoretical approach," *Frontiers in Neuroscience*, vol. 16, p. 971 937, 2022.

[17] J. Cuadrado, U. Rançon, B. Cottereau, F. Barranco, and T. Masquelier, "Optical flow estimation with event-based cameras and spiking neural networks," *arXiv preprint arXiv:2302.06492*, 2023.