This is a repository copy of *Is structural awareness the key to event camera data cleansing for enhancing veracity?*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/235941/

Version: Published Version

Proceedings Paper:
Li, H. and Abhayaratne, C. orcid.org/0000-0002-2799-7395 (2025) Is structural awareness the key to event camera data cleansing for enhancing veracity? In: 36th British Machine Vision Conference 2025, BMVC 2025, Sheffield, UK, November 24-27, 2025. 36th British Machine Vision Conference, 24-27 Nov 2025, Sheffield, UK. BMVA. Article no: 1007.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Is Structural Awareness the Key to Event Camera Data Cleansing for Enhancing Veracity?

Haiyu Li[1]
hli108@sheffield.ac.uk
Charith Abhayaratne[1,2]
c.abhayaratne@sheffield.ac.uk

[1] School of Electrical and Electronic Engineering
The University of Sheffield
Sheffield, United Kingdom

[2] Centre for Machine Intelligence
The University of Sheffield
Sheffield, United Kingdom

Abstract

Neuromorphic vision sensors, also known as event cameras, offer significant advantages over conventional frame-based cameras, in terms of high dynamic range, low latency, and low power consumption. However, their high sensitivity to illumination changes and asynchronous operation introduces substantial data (typically in the form of false or structure-irrelevant event data) posing challenges to the veracity of the acquired information in downstream vision tasks, such as, object recognition, feature tracking and human action recognition. Traditional cleansing methods for neuromorphic vision sensor event data typically rely on denoising techniques guided by reference-based metrics, which require auxiliary modalities such as Active Pixel Sensor (APS) frames or manual annotations. These references are often unavailable in real-world scenarios. Moreover, existing reference-free metrics generally overlook structural integrity, leading to deceptively high scores when aggressive noise removal results in the loss of meaningful structure. In this paper, we propose the *Temporal Structural Event Index* (TSEI), a novel, reference-free, structure-aware metric designed to assess the veracity of cleansed neuromorphic vision sensor event data. TSEI integrates temporal structural similarity (TSSM) and contrast normalization within an adaptive segmentation framework to jointly evaluate signal preservation and noise suppression. Experiments on both synthetic and real-world datasets demonstrate that TSEI strongly correlates with structural fidelity and recognition accuracy, outperforming existing metrics in detecting over-cleansing and structural degradation. These findings highlight that structural awareness is a critical factor in enhancing the veracity of neuromorphic event data and ensuring reliable performance in visual recognition tasks.

1 Introduction

Neuromorphic vision sensors, particularly event cameras, represent a novel class of biologically inspired vision device [20, 22]. Unlike conventional cameras that capture synchronized

© 2025. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

two-dimensional frames at fixed intervals, event cameras asynchronously record per-pixel brightness changes with microsecond-level temporal resolution [8]. Each recorded event encodes a polarity-based change in illumination, along with precise spatial and temporal information [15]. These distinctive characteristics endow event cameras with exceptionally high dynamic range, ultra-low latency, and low power consumption, making them highly suitable for real-time vision tasks such as optical flow estimation [18, 29, 30], high-speed video reconstruction [21, 26], feature tracking [6, 9, 11], and human action recognition [1] and simultaneous localization and mapping (SLAM) [23, 27, 28].

However, the high temporal sensitivity of neuromorphic vision sensors also makes them vulnerable to noise, typically in the form of spurious events or structure-irrelevant activity induced by background motion or illumination changes, especially under dynamic lighting conditions [17]. Even subtle variations in background or ambient lighting can generate substantial amounts of spurious activity, commonly referred to as background noise [25]. Consequently, the event streams captured by these sensors often contain a significant proportion of non-informative or misleading data, posing major challenges for downstream visual processing tasks. To address this, effective denoising that removes noise while preserving structurally meaningful events has become a critical pre-processing step in neuromorphic vision systems.

To evaluate the effectiveness of event data cleansing algorithms, a variety of metrics have been proposed. Most existing approaches are reference-based, relying on auxiliary modalities such as Active Pixel Sensor (APS) frames, Inertial Measurement Unit (IMU) signals, or manually annotated labels [2]. Although effective in controlled settings, these references are often unavailable or unreliable in real-world scenarios. Moreover, the sheer volume of neuromorphic vision sensor event data (often reaching millions of events per second) makes manual annotation impractical [5, 17].

In contrast, reference-free metrics such as the Total Sum of Squares (TSS) rely solely on contrast-based evaluations. While more practical, they are highly sensitive to event density and tend to overestimate performance in high-density or noisy regions, without adequately penalizing structural distortions.

In response to the limitations of existing metrics, several reference-free alternatives have been proposed. However, many of these overlook a critical aspect of neuromorphic event data quality: structural fidelity. That is, a cleansing algorithm may aggressively suppress noise yet inadvertently discard structurally meaningful events, still receiving high scores from contrast or density-based metrics. This shortcoming raises a central question: *Is structural awareness the key to improving the veracity of neuromorphic vision sensor event data and enabling reliable downstream recognition?*

To address this, we propose the **Temporal Structural Event Index (TSEI)**—a novel, reference-free, structure-aware metric that explicitly captures structural consistency within neuromorphic event data. TSEI integrates contrast normalization and temporal structural similarity (TSSM) within an adaptive segmentation framework. Our contributions are summarized as follows:

- We introduce a **density-aware normalized TSS formulation**, which employs adaptive sigmoid weighting to emphasize informative event clusters while suppressing sparse and noisy regions, enabling fair evaluation across scenes with varying densities.

- We incorporate **temporal structural similarity (TSSM)** between adjacent event win-

dows, capturing consistent spatial patterns over time and providing the first structure-aware evaluation metric specifically designed for neuromorphic event data cleansing.

- We design a **flexible scoring function** that fuses contrast and structural similarity, allowing tunable trade-offs between signal preservation and noise suppression.

Extensive experiments on both synthetic and real-world datasets demonstrate that TSEI more accurately reflects structural degradation and correlates strongly with recognition performance. These findings underscore the importance of structural awareness in estimating the veracity of neuromorphic event data and ensuring the reliability of vision systems built upon it.

2 Related Work

Evaluating denoised event data quality is a core challenge in neuromorphic vision. Early metrics are typically reference-based, requiring auxiliary data or manual annotations. For example, the Percentage of Signal Remaining (PSR) is a metric that uses bounding boxes to quantify the amount of signal relative to noise. [17], a recent approach utilizes APS and IMU signals to estimate the validity of events [2]. Although effective in controlled settings, such methods are limited in real-world use due to the lack of reliable references.

To address this, several reference-free metrics have been proposed. TSS [7] and similar contrast-based methods are efficient but often biased by event density, rewarding noise-heavy outputs and overlooking structural degradation. More recent work applies normalized contrast with spatial penalties, but remains static and lacks temporal modeling [10, 24].

In contrast, our proposed Temporal Structural Event Index (TSEI) is the first structure-aware, reference-free metric that jointly considers temporal structural similarity and density-normalized contrast. By adaptively segmenting event streams and capturing inter-window coherence, TSEI enables robust assessment of structural fidelity and avoids over-cleansing.

3 Methodology

TSEI combines three key components—adaptive temporal segmentation, normalized contrast measurement, and temporal structural similarity—into a unified framework for robust, structure-aware evaluation of cleansing performance across varying event densities and noise conditions. The resulting evaluation score, referred to as the *Temporal Structural Event Index (TSEI)* and denoted by E, reflects both temporal and spatial quality. These components respectively balance event density over time, enable density-invariant saliency measurement, and capture spatial consistency across adjacent temporal windows.

3.1 Adaptive Event Segmentation (AES)

To ensure fair comparison across segments with varying event densities, we partition the input event stream into segments containing approximately uniform numbers of events. The number of events per segment, denoted by n, is computed as:

$$n = \max\left(n_{\min}, \rho \log_{10}(N)\right), \tag{1}$$

where N is the total number of events, and ρ is the normalized spatiotemporal event density defined as:

$$\rho = \frac{N}{XYT}, \tag{2}$$

with X, Y, and T representing the spatial and temporal extents of the event stream. The total number of segments S is then determined by:

$$S = \left\lceil \frac{N}{n} \right\rceil. \tag{3}$$

This adaptive segmentation balances the number of events in each segment, thereby mitigating the effects of data imbalance and improving the robustness of subsequent evaluations.

3.2 Normalized Contrast Measure

Although contrast metrics such as the total sum of squares are widely employed in event data analysis, they suffer from two key limitations:

(1) they scale directly with the number of events, rendering them overly sensitive to density variations rather than the actual signal quality;

(2) they disregard spatial structure, often overemphasizing noise in dense regions while neglecting meaningful patterns in sparse areas.

To overcome these shortcomings, we propose a density-normalized contrast formulation incorporating two key modifications:

- **Z-score normalization**: Contrast values across temporal windows are standardized to remove biases introduced by raw event counts.
- **Density-aware weighting**: A sigmoid-based weighting factor is introduced to suppress the influence of overrepresented dense regions and promote fairness across different event distributions.

Specifically, the normalized contrast score for window i, denoted as D_i, is computed as:

$$D_i = \frac{D_i' - \mu_D}{\sigma_D + \varepsilon}, \tag{4}$$

where D_i' the unnormalized contrast value, which is defined as:

$$D_i' = w_i \sum_x I^2(x), \tag{5}$$

with $I(x)$ representing the warped event intensity at pixel x, calculated as:

$$I(x) = \sum_{e_k \in W_i} p_k \delta(x - x_k). \tag{6}$$

Here, p_k denotes the polarity of the k-th event, x_k denotes its spatial coordinate and $\delta(\cdot)$ denotes the Dirac delta function, approximated via discrete binning.

To mitigate contrast inflation in extreme density scenarios, a weighting factor w_i is introduced as follows:

$$w_i = \frac{1}{1 + e^{-\alpha_w(\rho_i - \beta_w)}}, \tag{7}$$

where ρ_i is the event density in window W_i and α_w, β_w are the hyper-parameters controlling the steepness and midpoint of the sigmoid function, respectively. This formulation ensures that the contrast measure captures salient structural patterns in the event data, independent of raw event density.

3.3 The Proposed Temporal Structural Similarity

To assess the preservation of temporal structural coherence in the event stream, we compute the structural similarity between each pair of consecutive temporal windows. Let $S(W_i, W_{i+1})$ denote the similarity score between window W_i and its successor W_{i+1}, defined as:

$$S(W_i, W_{i+1}) = \frac{(2\mu_i\mu_{i+1} + C_1)(2\sigma_{i,i+1} + C_2)}{(\mu_i^2 + \mu_{i+1}^2 + C_1)(\sigma_i^2 + \sigma_{i+1}^2 + C_2)}. \tag{8}$$

In this formulation, μ_i and μ_{i+1} represent the mean event intensities of windows W_i and W_{i+1}, respectively, while σ_i and σ_{i+1} denote their corresponding variances. The term $\sigma_{i,i+1}$ indicates the covariance between the two windows. The constants C_1 and C_2 are small values introduced to stabilize the computation and avoid division by zero.

This similarity measure captures the degree of structural continuity across adjacent temporal segments. It penalizes excessive filtering or cleansing operations that disrupt consistent spatial patterns over time, ensuring that meaningful temporal dynamics are preserved.

3.4 Temporal Structural Event Index (TSEI)

The final structural event score for each temporal window is computed as a weighted combination of the normalized contrast and the temporal structural similarity:

$$E_i = \alpha D_i + \beta S(W_i, W_{i+1}), \tag{9}$$

where α and β are tunable parameters that control the trade-off between contrast preservation and structural consistency.

The overall score for the entire event sequence is obtained by averaging the per-window scores:

$$E = \frac{1}{N}\sum_{i=1}^{N} E_i. \tag{10}$$

This unified formulation provides a robust, interpretable, and structure-aware evaluation metric for assessing the integrity of cleansed event data, without requiring any ground-truth or external references.

4 Experimental Results

4.1 Quantitative Comparison Under Synthetic Noise

We begin by evaluating the proposed TSEI metric alongside two classical metrics—TSS and MESR [4]—on synthetic datasets with controlled noise levels. Event streams are synthetically generated by superimposing background noise points onto known spatial structures,

Table 1: Performance comparison under 10% and 30% noise levels in different thresholds.

Cleansing Method	10% Noise			30% Noise		
	TSEI (Ours)	MESR [4]	TSS [7]	TSEI (Ours)	MESR	TSS
Pure_Event	0.1515	0.4322	0.1914	0.1515	0.3879	0.1556
Event_Noise	0.1565	0.4033	0.1976	0.1152	0.3281	0.1743
0.01	0.2998	0.5666	0.0314	0.3052	0.5093	0.0226
0.1	0.2837	0.5706	0.0397	0.2922	0.5044	0.0284
0.5	0.2407	0.5857	0.0758	0.2421	0.5179	0.0526
1	0.2243	0.6425	0.0929	0.2232	0.7867	0.1325

Table 2: Performance comparison under 50% and 70% noise levels in different thresholds.

Cleansing Method	50% Noise			70% Noise		
	TSEI (Ours)	MESR	TSS	TSEI (Ours)	MESR	TSS
Pure_Event	0.1503	0.4643	0.2143	0.1471	0.5125	0.2662
Event_Noise	0.1086	0.3532	0.2454	0.0959	0.3554	0.3102
0.01	0.2895	0.6297	0.0363	0.2776	0.7694	0.0491
0.1	0.2680	0.6314	0.0469	0.0959	0.7683	0.0630
0.5	0.2243	0.6425	0.0929	0.2232	0.7867	0.1325
1	0.2134	0.6366	0.1177	0.2110	0.8082	0.1711

with noise ratios set at 10%, 30%, 50%, and 70%. A temporal coherence-based cleansing algorithm with a tunable threshold is employed to filter out noise events. This algorithm retains events exhibiting strong local temporal consistency and discards temporally inconsistent events, thereby enabling controlled noise suppression at different aggressiveness levels.

As shown in Tables 1 and 2, TSS and MESR scores increase consistently with noise, often overstating performance despite structural degradation. In contrast, TSEI reflects true data quality remaining stable or decreasing under heavy noise by capturing structural loss, thus offering a more reliable assessment of event data veracity.

To further illustrate this behavior, Figure 1 plots the values of TSEI, MESR, and TSS across a range of cleansing thresholds under four noise levels. While MESR and TSS increase monotonically with more aggressive cleansing, TSEI exhibits a peak followed by a decline—initially rising as noise is removed, but dropping as structurally important events are lost. This turning-point pattern, uniquely captured by TSEI, demonstrates its effectiveness in balancing noise suppression with structural preservation—an essential requirement for reliable vision systems.

4.2 Structure-Preserving Evaluation via Hausdorff

To assess structural preservation during cleansing, we use the Hausdorff distance (HD)—a geometric metric that measures the maximum spatial deviation between the cleansed and ground-truth event sets. Applied to our synthetic datasets, it quantifies structural misalignment or loss. A rising Hausdorff distance with increasing threshold indicates over-cleansing and structural degradation, aligning well with TSEI trends.

Figure 2 illustrates the trends of TSEI, MESR, and TSS alongside the Hausdorff distance across varying noise levels and cleansing thresholds. As the threshold increases, the Hausdorff distance rises sharp, particularly under high-noise conditions, indicating progressive structural degradation. Among the three metrics, only TSEI reliably reflects this degradation, exhibiting a consistent downward trend that correlates strongly with the increase in Hausdorff distance. By contrast, MESR and TSS either remain flat or increase monotonically, failing to capture the underlying structural disruption.

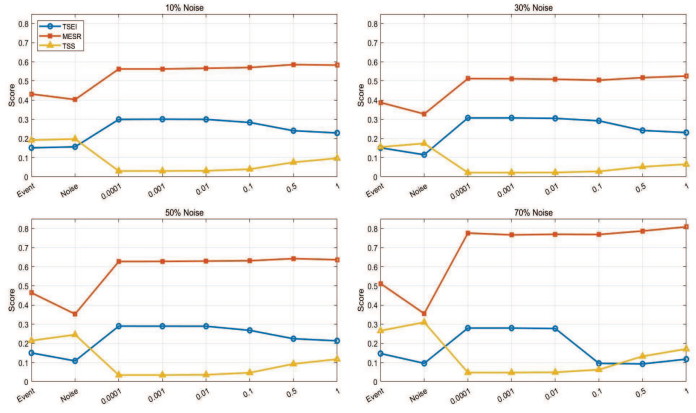

Figure 1: TSEI rises with noise removal but drops once structure degrades, while MESR and TSS keep increasing, unaffected by structural loss.

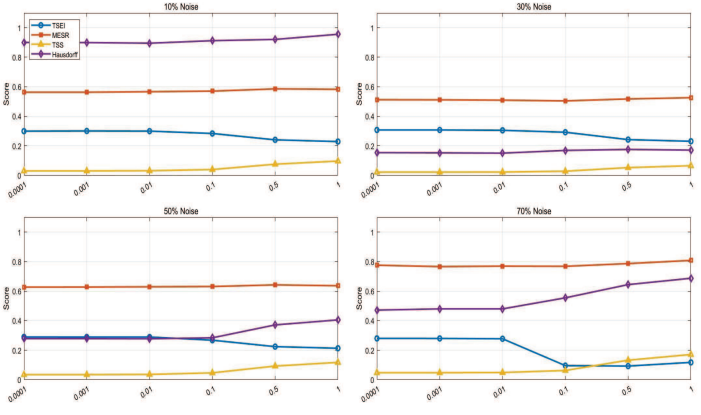

Figure 2: TSEI aligns closely with Hausdorff distance, while MESR and TSS remain insensitive to structural degradation.

To further visualize this behavior, Figure 3 presents heatmaps of TSEI, MESR, and TSS across all threshold-noise combinations. The TSEI heatmap distinctly highlights a "structural collapse zone" in high-noise, high-threshold regions, indicated by a sudden drop in scores. In contrast, the MESR and TSS heatmaps appear less discriminative, showing smooth or monotonic gradients that do not correlate with structural fidelity loss.

These results validate TSEIs effectiveness as a structure-aware quality metric. Its alignment with the Hausdorff distance—a direct geometric measure demonstrates its ability to penalize over-cleansing and detect structural damage. This sensitivity is crucial for maintaining the veracity of event data and supporting downstream vision tasks that rely on fine-grained spatial structure.

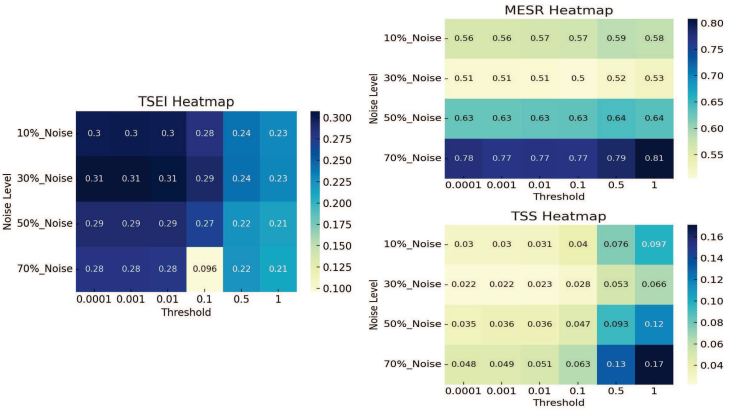

Figure 3: TSEI exhibits clear structural sensitivity, with low values concentrated in high-noise, high-threshold regions.

4.3 Comparison Across Cleansing Methods with Real Data

We further evaluate the proposed TSEI metric by comparing four representative event data cleansing methods—AW-GATCN [13], BAF [3], DWF [10], and TS [12]—across five diverse event datasets: Caltech101 [16], Cifar10 [14], D-END [4], N-CARS [19], and N-END [4]. For each method-dataset pair, we compute the average TSEI and MESR score before and after cleansing, and report the improvement (ΔTSEI and ΔMESR) to assess structural enhancement in Table 3. TSS is not used for comparison because it decreases with fewer event points, making it unreliable after noise reduction.

Background Activity Filter (BAF) and **Dual Window Filter (DWF)** are classical statistical filters. BAF estimates event activity within an 8-connected neighborhood and removes events below a threshold. DWF uses a FIFO queue to retain only recent events, filtering out those with low temporal correlation.

Time Surface (TS) converts events into a continuous temporal surface using logarithmic decay. Events disrupting surface smoothness are removed, enhancing robustness in structured scenes.

Adaptive Weighted Graph Attention Convolutional Network (AW-GATCN) treats events as spatiotemporal graphs with edges weighted by motion and polarity cues. A variance-based thresholding scheme and attention mechanism help identify and preserve structurally significant events.

As shown in Table 3, ΔMESR and ΔTSEI exhibit varying trends across methods and datasets. On *Caltech101*, TS yields the highest ΔMESR but a negative ΔTSEI, while AW-GATCN achieves the highest ΔTSEI. For *Cifar10*, TS shows the largest ΔTSEI, though most methods have negative ΔMESR. On *D-END*, BAF reports the highest ΔMESR, whereas AW-GATCN leads in ΔTSEI. In *N-CARS*, AW-GATCN achieves the highest improvements in both metrics, and on *N-END*, it shows the largest ΔMESR, while TS yields the highest ΔTSEI.

These variations indicate that ΔMESR and ΔTSEI do not always align. Since the two metrics operate on different numerical scales, their absolute values are not directly comparable; instead, the trend of change offers more insight into whether a method may be over-cleansing the data.

Table 3: Comparison of metrics improvements across different cleansing methods and datasets. Bold indicates the best performance per column.

Cleansing Method	Caltech101		Cifar10		D-END	
	ΔMESR	ΔTSEI (Ours)	ΔMESR	ΔTSEI (Ours)	ΔMESR	ΔTSEI (Ours)
AW-GATCN [13]	0.1886	**0.0190**	**0.0072**	0.0593	0.024	**0.0514**
BAF [3]	0.1056	-0.0275	-0.0618	0.0755	**0.1343**	0.0444
DWF [10]	0.1421	-0.0218	-0.0651	0.0427	0.1293	0.0447
TS [12]	**0.4769**	-0.1371	-0.3282	**0.1025**	0.1183	0.0330

Cleansing Method	N-CARS		N-END	
	ΔMESR	ΔTSEI (Ours)	ΔMESR	ΔTSEI (Ours)
AW-GATCN [13]	**0.2725**	**0.1866**	0.7729	0.0273
BAF [3]	0.0096	0.0116	**0.1872**	0.0053
DWF [10]	0.0358	0.0516	0.1863	0.0074
TS [12]	0.0235	0.0023	0.0458	**0.0326**

Table 4: Classification accuracy (%) across datasets and cleansing methods. Bold values indicate the highest accuracy per dataset.

Cleansing Method	Caltech101	Cifar10	D-END	N-CARS	N-END
AW-GATCN [13]	**74.07**	60.53	**85.73**	**90.18**	**75.47**
BAF [3]	50.87	65.70	80.43	66.67	63.33
DWF [10]	50.33	55.56	82.61	71.88	64.76
TS [12]	36.49	**71.69**	75.06	49.79	73.06

4.4 Correlation with Downstream Tasks

Building on the metric variations discussed in Section 4.3, we further evaluate the practical relevance of these metrics by examining their relationship with downstream classification accuracy. Table 4 presents classification results across five benchmark datasets and four representative cleansing methods, consistent with the configurations in Section 3.3.

A general alignment is observed between higher TSEI scores and improved classification performance. For instance, AW-GATCN consistently achieves top or near-top TSEI values (Table 3) and delivers strong accuracy across datasets. Notably, some exceptions exist, for example, TS shows relatively high accuracy on *N-END* despite moderate TSEI scores, suggesting that structural integrity is not the sole factor influencing task performance.

These findings empirically support the trends observed in Section 4.3: while traditional metrics such as MESR may fluctuate due to variations in event density, they do not consistently reflect structural changes introduced by different cleansing methods. In contrast, TSEI is specifically designed to account for structural fidelity and temporal coherence, making it more sensitive to meaningful improvements in data quality. The alignment between high TSEI scores and strong classification performance across datasets suggests that structure-aware metrics provide more reliable and task-relevant evaluations. This connection reinforces the utility of TSEI not only as a diagnostic tool for cleansing effectiveness but also as a practical indicator for selecting structure-preserving strategies in event-based vision applications.

5 Conclusions

We presented *TSEI*, a novel structure-aware, reference-free metric for evaluating event data cleansing quality. By integrating temporal structural similarity and density-aware contrast

normalization, TSEI effectively captures structural fidelity under diverse noise conditions. Experiments demonstrate that TSEI not only detects over-cleansing but also correlates strongly with geometric integrity and downstream classification performance. In addition to an evaluation, TSEI serves as a practical indication for guiding structure-preserving cleansing strategies, especially in real-world scenarios where reference data is unavailable. Its lightweight, unsupervised nature makes it applicable to embedded real-time vision systems, such as, robotics or autonomous driving.

References

[1] Salah Al-Obaidi, Hiba Al-Khafaji, and Charith Abhayaratne. Making sense of neuromorphic event data for human action recognition. *IEEE Access*, 9:82686–82700, 2021. doi: 10.1109/ACCESS.2021.3085708.

[2] R. Wes Baldwin, Mohammed Almatrafi, Vijayan Asari, and Keigo Hirakawa. Event probability mask (EPM) and event denoising convolutional neural network (EDnCNN) for neuromorphic cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1698–1707. IEEE, 2020. doi: 10.1109/CVPR42600.2020.00177.

[3] Tobi Delbrück. Frame-free dynamic digital vision. *Secure-Life Electronics: Advanced Electronics for Quality Life and Society*, 1(1):21–26, 2008. doi: 10.5167/UZH-17620.

[4] Saizhe Ding, Jinze Chen, Yang Wang, Yu Kang, Weiguo Song, Jie Cheng, and Yang Cao. E-MLB: multilevel benchmark for event-based camera denoising. *IEEE Transactions on Multimedia*, 26:1–12, 2024. doi: 10.1109/TMM.2023.3260638.

[5] Yifan Feng, Yixuan Liu, Yang Wang, Tianxiang Wang, and Zheng-Jun Zha. Event density based denoising method for dynamic vision sensor. *Applied Sciences*, 10(6): 2024, 2020. doi: 10.3390/app10062024.

[6] Guillermo Gallego, Henri Rebecq, and Davide Scaramuzza. Event-based, 6-DOF camera tracking from photometric depth maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(10):2402–2412, 2018. doi: 10.1109/TPAMI.2017.2769655.

[7] Guillermo Gallego, Daniel Gehrig, and Davide Scaramuzza. Focus is all you need: Loss functions for event-based vision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12272–12281. IEEE, 2019. doi: 10.1109/CVPR.2019.01256.

[8] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jorg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):154–180, 2022. doi: 10.1109/TPAMI.2020.3008413.

[9] Daniel Gehrig, Henri Rebecq, Guillermo Gallego, and Davide Scaramuzza. EKLT: asynchronous photometric feature tracking using events and frames. *International Journal of Computer Vision*, 128(3):601–618, 2020. doi: 10.1007/s11263-019-01209-w.

[10] Shizhe Guo and Tobi Delbruck. Low cost and latency event camera background activity denoising. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1): 785–795, 2023. doi: 10.1109/TPAMI.2022.3152999.

[11] Benjamin Kueng, Elias Mueggler, Guillermo Gallego, and Davide Scaramuzza. Low-latency visual odometry using event-based feature tracks. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 16–23. IEEE, 2016. doi: 10.1109/IROS.2016.7758089.

[12] Xavier Lagorce, Garrick Orchard, Francesco Galluppi, Shi, and Benosman. HOTS: a hierarchy of event-based time-surfaces for pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(7):1346–1359, July 2017. doi: 10.1109/TPAMI.2016.2574707.

[13] Haiyu Li and Charith Abhayaratne. AW-GATCN: adaptive weighted graph attention convolutional network for event camera data joint denoising and object recognition. In *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, 2025. arXiv preprint arXiv:2505.11232.

[14] Hongmin Li, Guangzhi Liu, Xiangyang Ji, Guoqi Li, and Luping Shi. CIFAR10-DVS: an event-stream dataset for object classification. *Frontiers in Neuroscience*, 11:309, May 2017. doi: 10.3389/fnins.2017.00309.

[15] Patrick. Lichtsteiner, Christoph. Posch, and Tobi. Delbrück. A 128 $\times$ 128 120 db 15 μs latency asynchronous temporal contrast vision sensor. *IEEE Journal of Solid-State Circuits*, 43(2):566–576, 2008. doi: 10.1109/JSSC.2007.914337.

[16] Garrick Orchard, Ajinkya Jayawant, Gregory K. Cohen, and Nitish V. Thakor. Converting static image datasets to spiking neuromorphic datasets using saccades. *Frontiers in Neuroscience*, 9:437, 2015. doi: 10.3389/fnins.2015.00437.

[17] Venkatesh Padala, Arindam Basu, and Garrick Orchard. A noise filtering algorithm for event-based asynchronous change detection image sensors on truenorth and its implementation on truenorth. *Frontiers in Neuroscience*, 12:118, 2018. doi: 10.3389/fnins.2018.00118.

[18] Lei Pan, Ming Liu, and Richard Hartley. Single image optical flow estimation with an event camera. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1669–1678. IEEE, 2020. doi: 10.1109/CVPR42600.2020.00174.

[19] Amos Sironi, Marc Brambilla, Nicolas Bourdis, Xavier Lagorce, and Ryad Benosman. HATS: histograms of averaged time surfaces for robust event-based object classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1731–1740. IEEE, 2018. doi: 10.1109/CVPR.2018.00186.

[20] Gemma Taverni, Diederik Paul Moeys, Chenghan Li, Celso Cavaco, Vasyl Motsnyi, David San Segundo Bello, and Tobi Delbrück. Front and back illuminated dynamic and active pixel vision sensors comparison. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(5):677–681, 2018. doi: 10.1109/TCSII.2018.2824899.

[21] Sergey Tulyakov, Zhixin Liu, Xueying Wang, Ming-Yu Liu, and Jan Kautz. Time lens: Event-based video frame interpolation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16150–16159. IEEE, 2021. doi: 10.1109/CVPR46437.2021.01589.

[22] Lin Wang, I. S. Mohammad Mostafavi, Yo-Sung Ho, and Kuk-Jin Yoon. 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 (CVPR)*, pages 10073–10082. IEEE, 2019. doi: 10.1109/CVPR.2019.01032.

[23] David Weikersdorfer, Darius B. Adrian, Daniel Cremers, and Jorg Conradt. Event-based 3D SLAM with a depth-augmented dynamic vision sensor. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 359–364. IEEE, 2014. doi: 10.1109/ICRA.2014.6906882.

[24] Jing Wu, Chao Ma, Lei Li, Weisheng Dong, and Guangming Shi. Probabilistic undirected graph based denoising method for dynamic vision sensor. *IEEE Transactions on Multimedia*, 23:1148–1159, 2021. doi: 10.1109/TMM.2020.2993957.

[25] Jun Xu, Jing Zou, and Zhen Gao. Comment on "Temperature and Parasitic Photocurrent Effects in Dynamic Vision Sensors". *IEEE Transactions on Electron Devices*, 65 (7):3081–3082, 2018. doi: 10.1109/TED.2018.2833106.

[26] Zhe Yu, Boyu Jiang, Yizhou Liu, Yitong Wang, Yinqiang Wang, and Qifeng Zhang. Training weakly supervised video frame interpolation with events. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 14569–14578. IEEE, 2021. doi: 10.1109/ICCV48922.2021.01432.

[27] Xinyu Zhang and Li Yu. Unifying motion deblurring and frame interpolation with events. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17744–17753. IEEE, 2022. doi: 10.1109/CVPR52688.2022.01724.

[28] Yizhou Zhou, Guillermo Gallego, and Shaojie Shen. Event-based stereo visual odometry. *IEEE Transactions on Robotics*, 37(5):1433–1450, 2021. doi: 10.1109/TRO.2021.3062252.

[29] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. EV-FlowNet: self-supervised optical flow estimation for event-based cameras. In *Proceedings of Robotics: Science and Systems (RSS)*, 2018. doi: 10.15607/RSS.2018.XIV.062.

[30] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Unsupervised event-based learning of optical flow, depth, and egomotion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 989–997, 2019. doi: 10.1109/CVPR.2019.00108.