SCIENCE CHINA
Information Sciences

• LETTER •

February 2026, Vol. 69, Iss. 2, 129405:1–129405:2
https://doi.org/10.1007/s11432-025-4641-7

EE-Extractor: a near-sensor real-time effective event extractor for dynamic vision sensor

Feiqiang LI†, Yujie HUANG*†, Mingyu WANG, Wenhong LI, Ming'e JING* & Xiaoyang ZENG

State Key Laboratory of Integrated Chips and Systems, Fudan University, Shanghai 200000, China

Received 3 June 2025/Revised 4 August 2025/Accepted 17 October 2025/Published online 4 January 2026

Citation Li F Q, Huang Y J, Wang M Y, et al. EE-Extractor: a near-sensor real-time effective event extractor for dynamic vision sensor. Sci China Inf Sci, 2026, 69(2): 129405, https://doi.org/10.1007/s11432-025-4641-7

Unlike traditional frame-based cameras, the dynamic vision sensor (DVS)'s output (i.e., events) comes from changes in pixel brightness [1]. It requires no exposure time and generates an event in only microseconds from a change in pixel brightness, contributing to less redundancy, latency, and power consumption. Hence, DVS has potential for machine vision applications such as video surveillance, industrial inspection, and autonomous driving, especially in scenarios involving high-speed motion and extreme lighting. However, due to the sparsity of real events, DVS is sensitive to background active (BA) noise. BA noise confuses events, wasting bandwidth and affecting subsequent processing. Therefore, extracting effective events is a significant task for DVS.

To extract effective events, spatiotemporal correlation filters [2] are widely used. However, it requires array-sized memory cells to store the timestamp of the most recent event, which is unfriendly to hardware implementation. Some studies [3, 4] try to reduce storage overhead by sharing memory cells, but their approaches destroy the complete spatiotemporal correlation, reducing the performance of denoising. In addition, when there are a large number of noise events, classical spatiotemporal filters cannot effectively remove them. Some studies use more complex methods for denoising, such as EDformer [5]. However, these methods require a large storage and computing resource overhead solely for denoising, which is not cost-effective.

In this work, an EE-Extractor near the DVS pixel array is proposed, including a Bit Queue spatiotemporal filter and a boundary-box-based clusterer. For the proposed filter, according to the output characteristics of the DVS pixel array, the Bit Queue memory structure and the Bit Mask Computer (BMC) computation module are designed, through which the memory overhead and computational complexity are reduced. For the clusterer, a boundary-box representation is proposed, achieving a balance between effectiveness and computational complexity. Then, an ending mechanism and a category coverage mechanism are presented, which ensure the integrity of information after clustering. Experiment results demonstrate that the proposed filter achieves the best performance with the least memory among the recent spatiotemporal filters. The clusterer further removes noise while extracting targets, and it outperforms other event-based clustering methods on ROI extraction.

Proposed algorithm. Figure 1(b) is the algorithm block diagram of the proposed filter. Leveraging the row-wise arbitration characteristics of the DVS pixel array output, a Bit Queue is designed for the filter. Rather than storing event coordinates (x, y), Bit Queue directly stores one-row events from the DVS array. It maintains a sliding window of the most recent R rows of binary event data, where each bit represents the presence of the event. For the incoming event (x_i, y_i), the algorithm extracts an adjacent event matrix E that covers the spatial neighborhood defined by $x \in [x_i - D, x_i + D]$ and $y \in [y_i - D, y_i + D]$ from the Bit Queue. The event is validated as real if any '1' exists in E. Figure 1(c) is the algorithm block diagram of the proposed clustering. The category is represented by the boundary box and its event count, which also serves as the fundamental unit for event-by-event clustering. An event is assigned to a category when its distance to the boundary box falls below the clustering threshold G. Following this assignment, the system dynamically updates both the boundary box and the event count for the category. Furthermore, based on the output characteristics of the DVS pixel array, a category coverage mechanism is presented, which reclaims inactive categories, ensuring efficient resource utilization. The clustering ending mechanism terminates processing based on the completion of a full frame scan and timeout thresholds, which ensure the integrity of information after clustering. Details about the algorithm are presented in Appendix A.

Hardware design. The hardware design of the EE-Extractor is illustrated in Figure 1(a). As shown in Figure 1(d), for the filter, the Bit Queue employs eight register groups to store the past eight rows of 512-bit event data, while Row Num stores corresponding row numbers. The Read Ctrl module extracts eight *Read_data* segments centered around the current event's y-coordinate from the Bit Queue. The BMC module then performs bit-mask operations to determine spatiotemporal correlation, achieving real-time denoising. As shown in Figure 1(e), for the clusterer, the Category information memory stores boundary coordinates and event counts for eight categories. The Updateboundary module

* Corresponding author (email: yujiehuang@fudan.edu.cn, mejing@fudan.edu.cn)
† These authors contributed equally to this work.

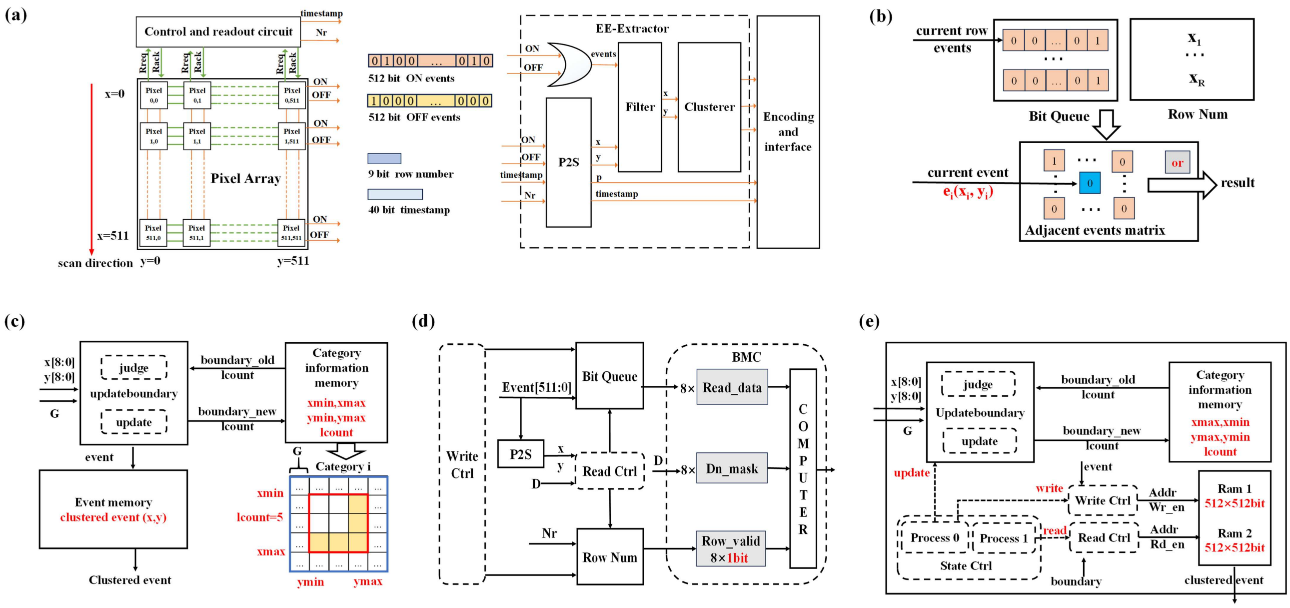

Figure 1 (Color online) (a) Overview of the proposed EE-Extractor, which is positioned near the DVS pixel array; (b) algorithm of the Bit Queue spatiotemporal filter; (c) algorithm of the boundary-box-based clustering; (d) hardware architecture of the filter; (e) hardware architecture of the clusterer.

processes each incoming event to determine category assignment and updates boundaries accordingly. Two Rams in a ping-pong structure alternate between the clustering and output phases, ensuring continuous processing. The State Ctrl module manages state transitions between data reception and output phases, while the Read Ctrl module supports both coordinate-based and binary-frame output formats for subsequent processing. Details about the design are presented in Appendix B.

Performance evaluation. A series of experiments was conducted to evaluate the proposed EE-Extractor in comparison with state-of-the-art methods. Quantitative experiments demonstrate that the proposed filter outperforms recent spatiotemporal filters under different lighting conditions, improving performance by at least 4.1% while reducing memory by at least 77%. On the public dataset, visual comparisons confirm that the proposed filter effectively preserves real events while removing noise. The proposed clusterer further removes residual noise that passes through the filter. For ROI extraction on the public dataset, the proposed clusterer improves the weighted F1-score by an average of 15.7% compared to recent event-based clustering methods. Details about the experiments are presented in Appendix C.

Hardware implementation. The EE-Extractor is implemented on an FPGA board equipped with an Xilinx xc7k325tffg900-3 chip, which is operated at 100 MHz. It takes only 110 ns from the generation of the first event to the output of the denoising result and 538.2 μs to complete the output of clustering. Finally, the EE-Extractor is logically synthesized at 250 MHz using the HLMC 55 nm. The result shows that the entire EE-Extractor has an area of 1.04 mm^2 and a power consumption of 58.77 mW. At 1 Meps, the system processed an average of 513 events per clustering, and the throughput is 1949 fps (513 μs for the accumulation of events, 0.08 μs for processing delay).

Conclusion. A real-time EE-Extractor is proposed in this study, including a novel Bit Queue spatiotemporal filter and a boundary-box-based clusterer. Experiment results demonstrate that the performance of the proposed filter has improved by at least 4.1% compared to recent spatiotemporal filters, and memory is reduced by at least 77%. The clusterer further removes noise while extracting targets, and it outperforms other event-based clustering methods on ROI extraction. Finally, the EE-Extractor is implemented on both the FPGA and the HLMC 55 nm process. It achieves a throughput of 1949 fps, has an area of 1.04 mm^2 and a power consumption of 58.77 mW.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant No. 62074041), Key R&D Program of Shandong Province (Grant No. 2022CXGC010504), State Key Laboratory of ASIC and System (Grant No. 2021KF009), and Zhuhai Fudan Innovation Institute.

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Posch C, Serrano-Gotarredona T, Linares-Barranco B, et al. Retinomorphic event-based vision sensors: bioinspired cameras with spiking output. Proc IEEE, 2014, 102: 1470–1484
2 Delbruck T. Frame-free dynamic digital vision. In: Proceedings of International Symposium on Secure-Life Electronics, 2008
3 Liu H, Brandli C, Li C, et al. Design of a spatiotemporal correlation filter for event-based sensors. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), 2015. 722–725
4 Khodamoradi A, Kastner R. O(N)-space spatiotemporal filter for reducing noise in neuromorphic vision sensors. IEEE Trans Emerg Top Comput, 2021, 9: 15–23
5 Jiang B, Xiong B, Qu B, et al. EDformer: transformer-based event denoising across varied noise levels. In: Proceedings of European Conference on Computer Vision, 2024