Supplementary Information

Title

Threshold-tunable event-driven vision sensor for adaptive visual processing

Authors

Guanlei Zhao[1]†, Siyu Zhang[1]†, Zhongfang Zhang[1]†, Yingjie Tang[1], Guolei Liu[1], Fanfan Li[1], Kaixin Ge[1], Shuhui Ren[1], Jiayi Mao[1], Weilong Feng[1], Saisai Wang[2], Qi Huang[2], Hong Wang[4*], Bowen Zhu[1,2,3*]

Affiliations

†These authors contributed equally to this work.

[1] Zhejiang Key Laboratory of 3D Micro/Nano Fabrication and Characterization, Department of Electronic and Information Engineering, School of Engineering, Westlake University, Hangzhou, Zhejiang 310030, China.

[2] Westlake Institute for Optoelectronics, Hangzhou 311421, China.

[3] Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou 310024, China

[4] Key Laboratory of Wide Band Gap Semiconductor Technology School of Microelectronics Xidian University, Xi'an 710071, China.

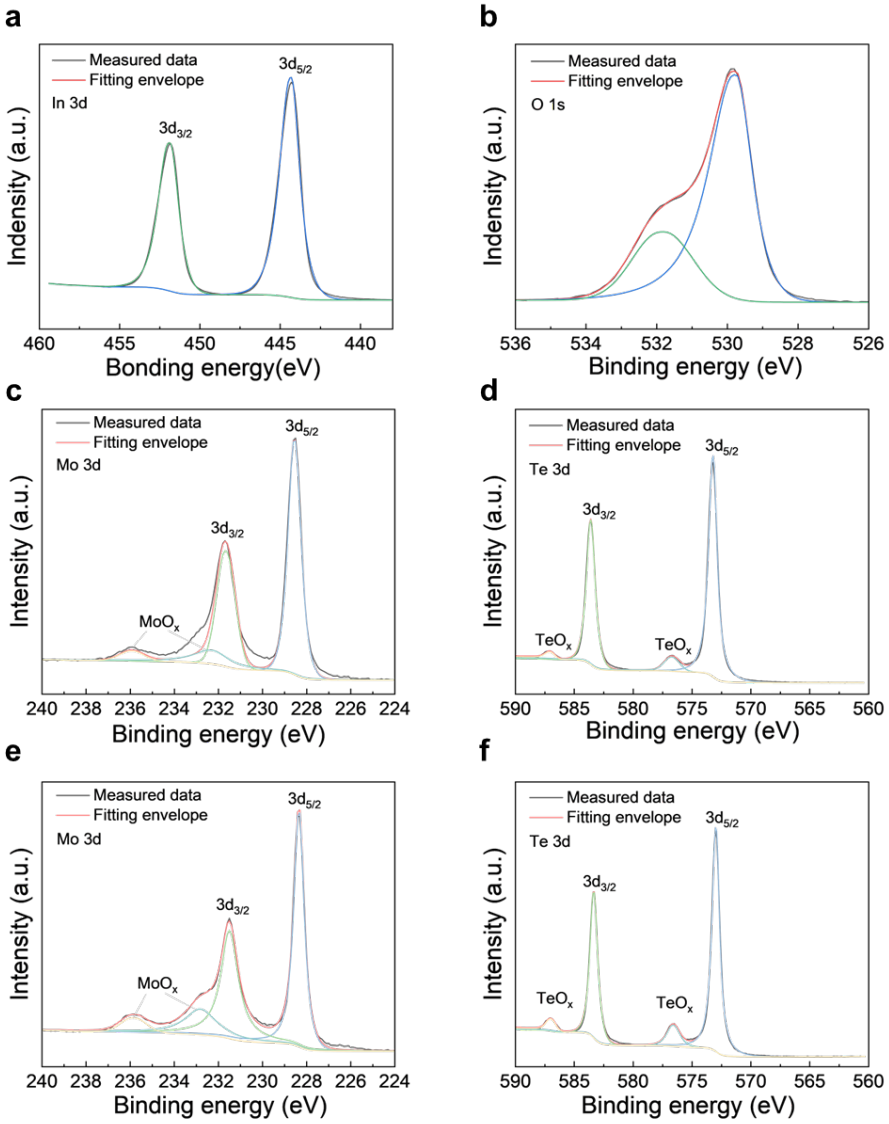
*Corresponding authors. Email: hongwang@xidian.edu.cn (H.W.); zhubowen@westlake.edu.cn (B.Z.)

Supplementary Note 1. Physical significance and calculation method of event entropy. Event entropy is derived from the concept of image entropy but is extended to dynamic scenes to characterize the aggregation properties of event streams generated by event-driven vision sensors. Unlike image entropy, which is defined for static grayscale distributions, event entropy quantifies the spatial and polarity-wise distribution of asynchronous events over time. Event-stream processing modulates event entropy: noise suppression leads to a more concentrated event distribution and reduced entropy, whereas detail enhancement increases entropy by promoting richer event activity. The calculation of event entropy consists of three steps. First, the event stream is spatially partitioned into non-overlapping blocks. The block size is selected to balance computational cost and spatial fidelity. For the 6 × 6 pixels array, a block size of 2×2 pixels are used, while for large-scale arrays (e.g., 1280×720 pixels), a block size of 16×16 pixels are adopted to ensure efficient entropy estimation without excessive computational overhead. Second, at each time instant t, the numbers of positive and negative events within each block $i$are counted, and the corresponding probabilities $p(x, s)$ are calculated.

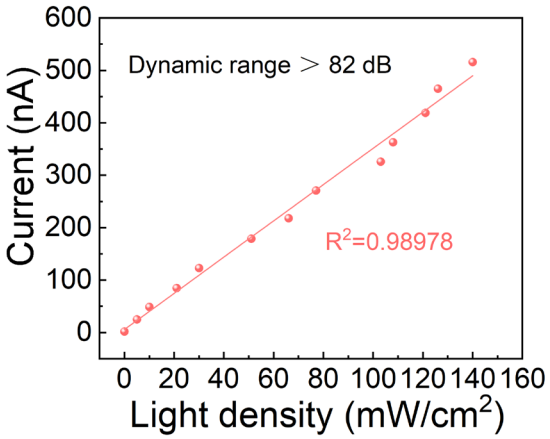
Finally, the event entropy $H(X, S)$is computed as

$$H(X,S) = -\sum_{x=1}^{N} \sum_{s\in\{+1,-1\}} p(x,s) \log_2 p(x,s) \quad (1)$$

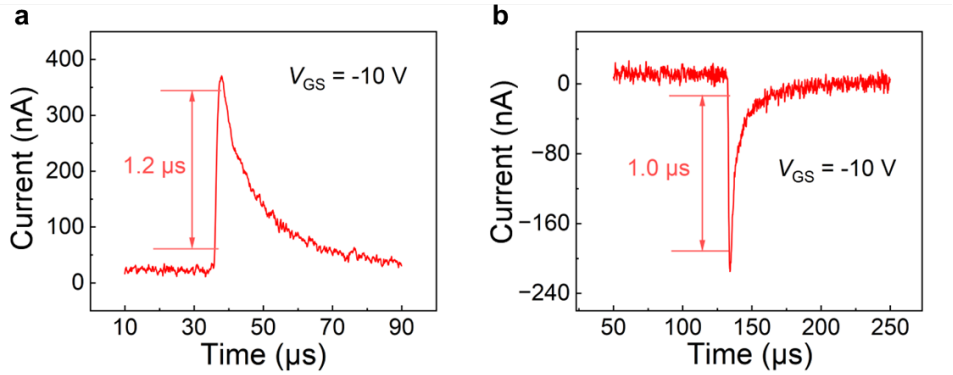
where x indexes the spatial blocks, s denotes event polarity, and $p(x, s)$ is the probability of observing an event of polarity s in block x. Event entropy quantifies the complexity and richness of the event stream.



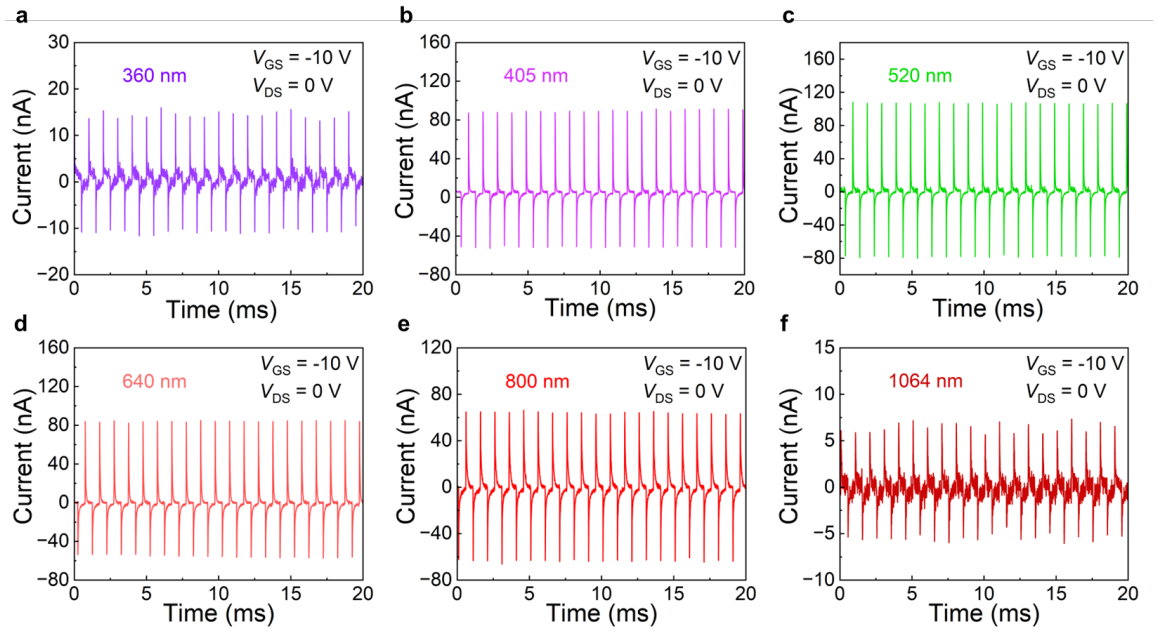
Supplementary Fig. 1. X-ray Photoelectron Spectroscopy (XPS) data. a, **b**, In_2O_3. **c**, **d**, $MoTe_2$ before EBE. **e**, **f**, $MoTe_2$ after EBE.



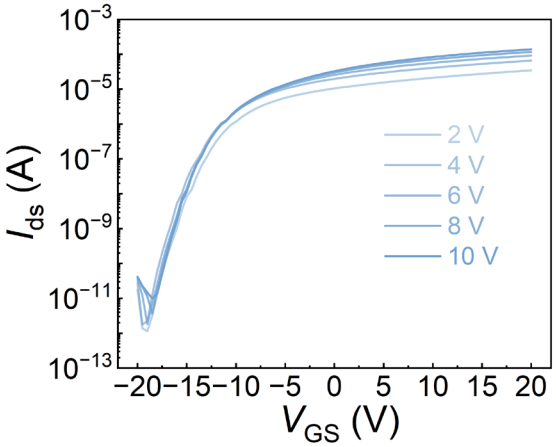
Supplementary Fig. 2. Dynamic range under 800 nm illumination.



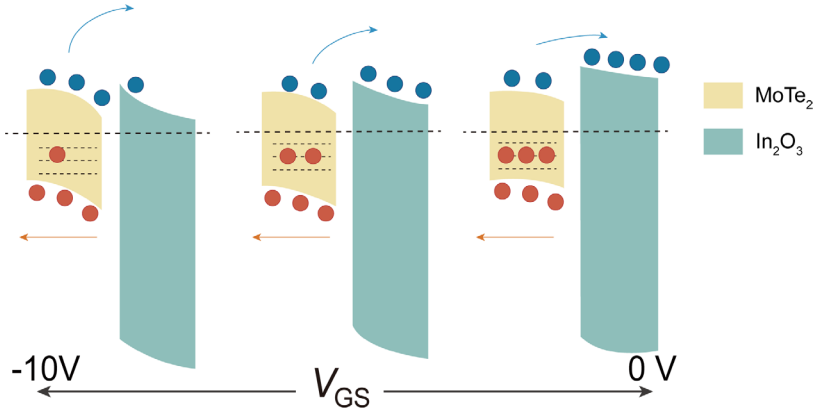
Supplementary Fig. 3. The response time of positive and negative spikes. The device exhibits rapid photoelectric pulse responses, with rise/fall times of 1.2 μs and 1.0 μs for positive and negative pulses, respectively—a performance that is comparable to commercial event-based cameras.



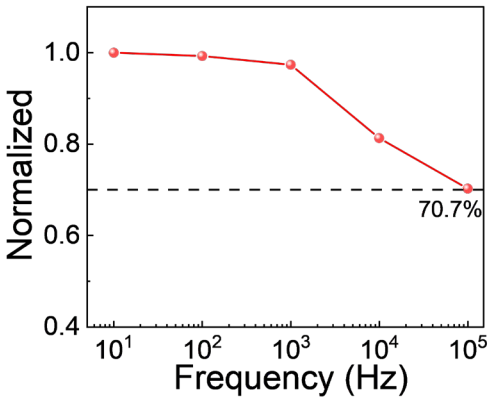
Supplementary Fig. 4. Broad-spectrum response of phototransistors from 360 to 1064 nm. Due to its narrow bandgap, $MoTe_2$ exhibits a broad spectral response ranging from 360 to 1064 nm, making it highly suitable for event-driven vision sensors. This wide responsivity enables reliable operation across diverse lighting conditions, from visible to infrared light, thereby enhancing the sensor's versatility and robustness in real-world applications.



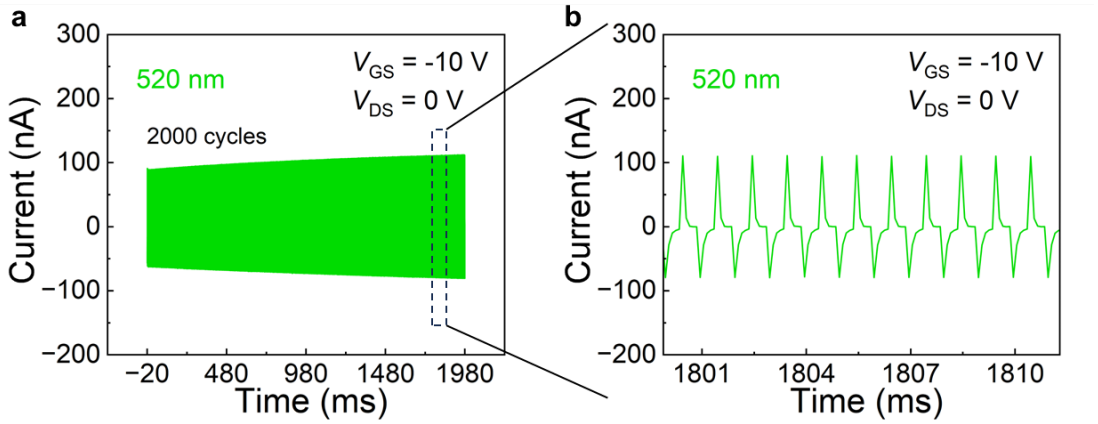
Supplementary Fig. 5. Logarithmic transfer characteristics curve of the $MoTe_2/In_2O_3$ phototransistor.



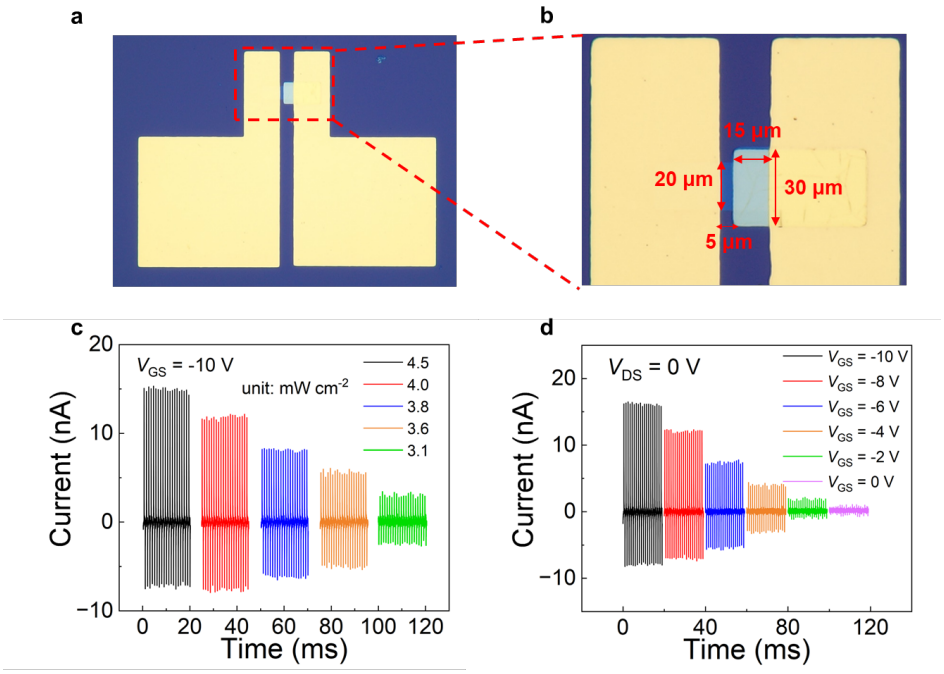
Supplementary Fig. 6. Gate-controlled band diagram. Under negative gate biases (V_{GS} = −10 V), the reduced carrier density in In_2O_3 alters the interfacial band alignment while the electrostatic potential of $MoTe_2$ remains largely unchanged, thereby enhancing the effective junction electric field and promoting more efficient charge separation. As V_{GS} approaches 0 V, the strength of the junction field decreases.



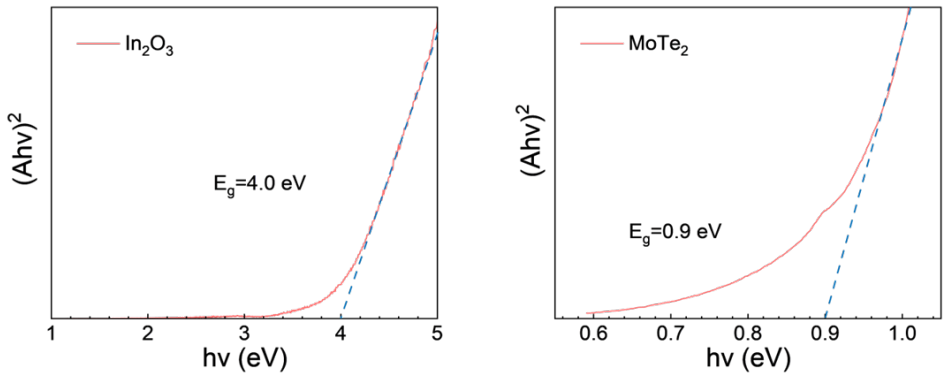
Supplementary Fig. 7. The photoresponse exhibits frequency dependence, modulated by the laser frequency controlled via a signal generator. The 3dB bandwidth is determined to be 10^5 Hz, corresponding to the frequency at which the corresponding photoresponse attenuate by 70.7%



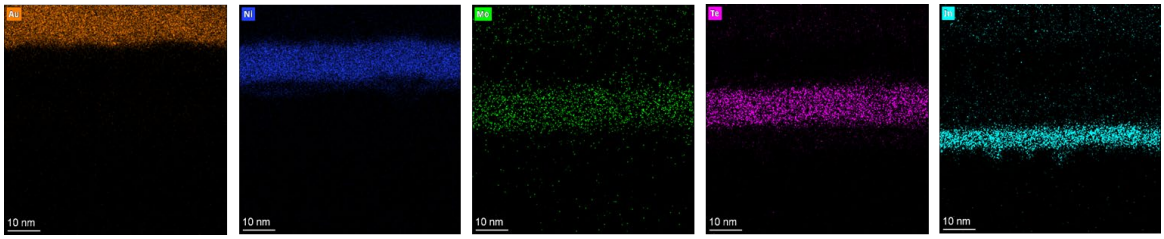
Supplementary Fig. 8. Device cycling stability. The device exhibits excellent stability, as the photoresponse remains nearly unchanged after 2000 cycles of testing.



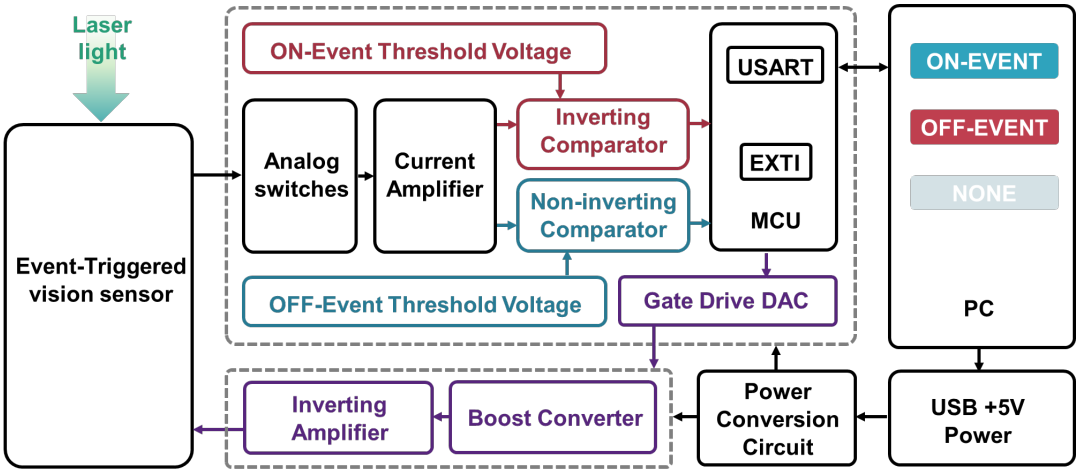
Supplementary Fig. 9. Photoelectronic performance and gate-tunable characteristics of 15×30 μm^2 small-area devices. a, Photograph of the small-area device structure. **b**, Magnified view of the photosensitive region. **c**, Influence of illumination on device response. **d**, Gate-tunable performance of the photoresponse.



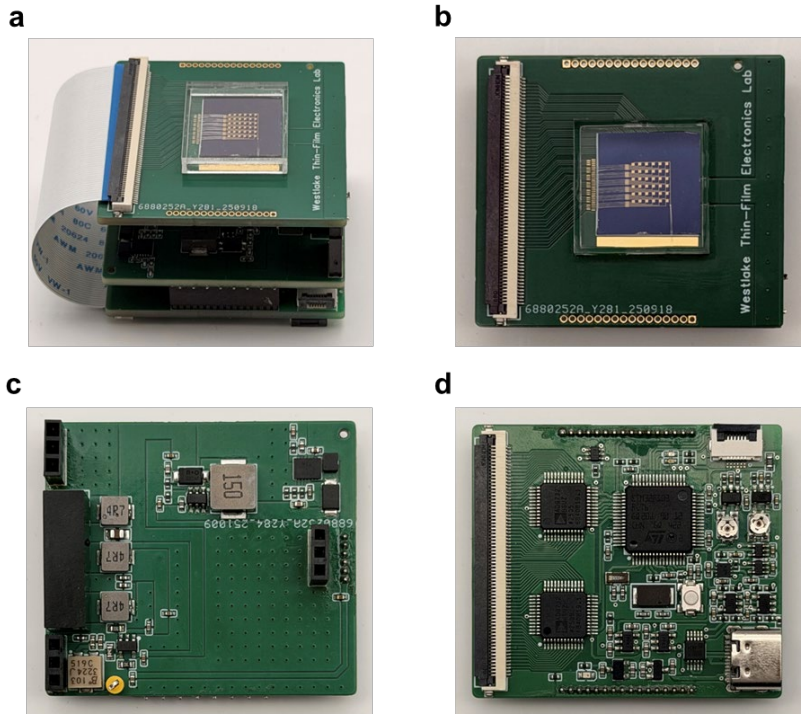
Supplementary Fig. 10. Material absorption spectrum. a, In_2O_3. **b**, $MoTe_2$. The extended lines are used to estimate the band gap of the material.



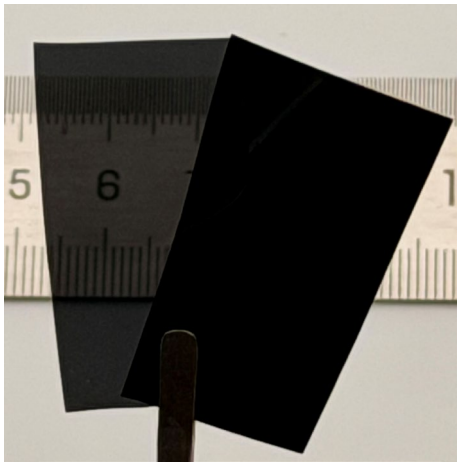
Supplementary Fig. 11. Top-down Energy Dispersive Spectrometer (EDS) data for Au, Ni, Mo, Te, and In.



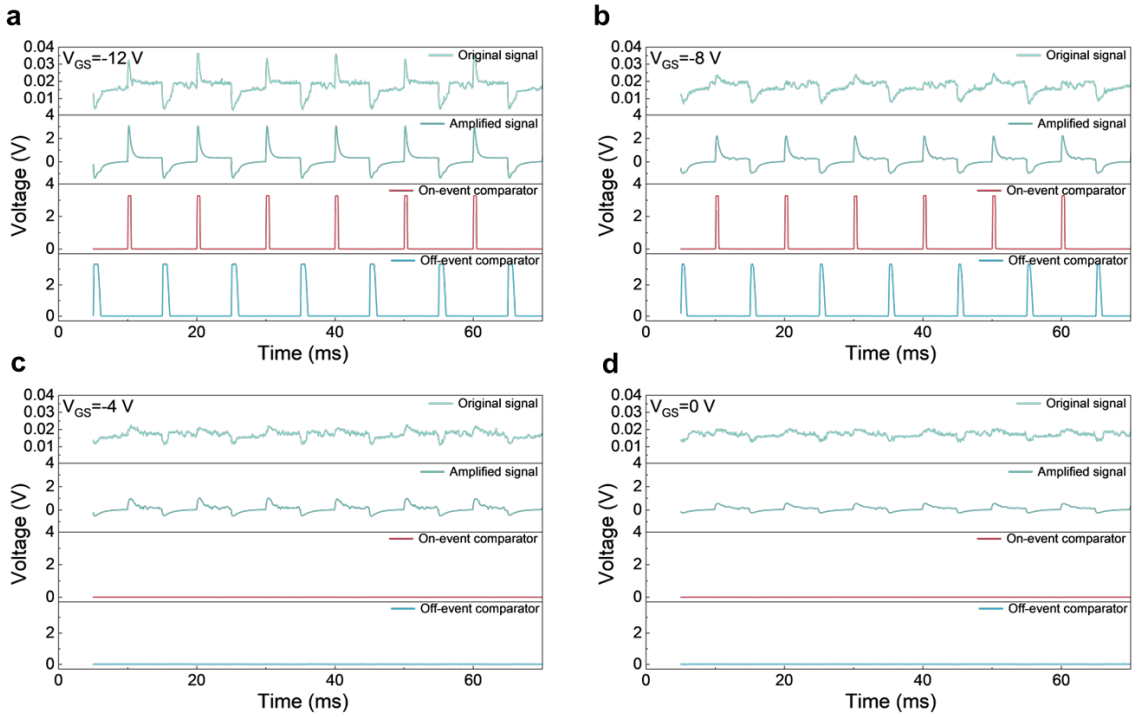
Supplementary Fig. 12. PCB circuit sampling logic. When a laser irradiates the array surface, it generates current signals. Due to the large number of sensors in the array, analog switches are used to sequentially select the source of each transistor in order to reduce circuit size and improve overall integration, thereby enabling the reuse of subsequent circuits. The current signal from the analog switches is converted into a voltage signal via a transimpedance amplifier circuit. Since the pulses generated by the sensors have extremely short pulse widths and contain certain interference signals, a pair of inverting and non-inverting comparators is used to separately capture the positive and negative responses of the pulses. This process yields two wider square waves while filtering out interference signals. In the microcontroller unit (MCU), two external interrupts (EXTI) are employed to separately acquire the square wave signals, and the data is transmitted to the computer host via universal asynchronous receiver/transmitter (UART) in the form of 0 and 1. This solution generates less data volume and maintains the line idle when no event is triggered, enabling the MCU to successfully transmit relatively high-frequency data. While acquiring sensor data, the host computer can also send control signals via UART to adjust the gate voltage. Upon receiving the control signals, the MCU generates an adjustable voltage via a digital-to-analog converter (DAC), which is then converted into the negative gate voltage required through an inverting amplifier. The entire circuit is powered by the computer's +5V USB port, and the different voltages required by various circuits are converted via a power conversion circuit. This data cable provides power while also enabling UART communication, ensuring high integration of the overall system.



Supplementary Fig. 13. Components of the PCB acquisition circuit. a, Overall diagram of the PCB board. **b**, Wire bonding PCB. **c**, Gate power supply PCB. **d**, Acquisition circuit PCB.



Supplementary Fig. 14. Optical thin films for masks with 30% and 0% transmittance.



Supplementary Fig. 15. Modulation of V_{GS} on event thresholds. Using an oscilloscope to measure the original voltage of event-triggered signals on the PCB, as well as the amplified voltage and the signals after passing through the positive and negative comparators, it is clear that the gate modulates the threshold for event generation. When (a) V_{GS}=-12 V and (b) V_{GS}=-8 V, the event pulses exceed the threshold, resulting in output from the subsequent comparators. When (c) V_{GS}=-4 V and (d) V_{GS}= 0 V, the event pulses do not exceed the threshold, resulting in no output from the subsequent comparators.

Supplementary Table 1. Comparison of commercial or prototype event cameras. Shown here are several representative circuit-based event cameras, whose largest array size reaches 1280 × 960 pixels, still far smaller than that of common CMOS image sensors. Owing to circuit complexity, most event cameras exhibit a fill factor below 30%. In addition, complex circuit architectures introduce static noise even in the absence of events, leading to spurious event triggering and additional power consumption, which are common limitations of conventional circuit-based designs.

Supplier	**iniVation**	**Prophesee**		**Samsung**	
Camera Model	DVS128 (2008)[1]	Gen CD (2017)[2]	Gen4 CD (2020)[3]	DVS-Gen3 (2018)[4]	DVS-Gen4 (2020)[5]
Pixels	128×128	480×360	1280×720	640×480	1280×960
Pixel size (μm^2)	40×40	20×20	4.86×4.86	9×9	4.95×4.95
Fill factor (%)	8.1	25	>77	12	22
Latency (μs)	12μs@ 1 klus	40-200	20-150	50	150
Min. contrast sensitivity (%)	17	12	11	15	20
Power consumption (mW)	23	36-95	32-84	40	130
Stationary noise (ev/pix/s) @ 25°C	0.05	0.1	0.1	0.03	NA

Values are approximate since there is no standard measurement tested.

Supplementary Table 2. Recent research on structurally simplified event vision sensors.

Device structure	λ (nm)	Array scale	Dynamic range (dB)	Response time (μs)	DETR (dB)	System-level verification	Ref.
$MoTe_2$/ In_2O_3	360-1064	6×6	82	1.2	>60	Yes	This work
ITO/ZnO/ N2200/ Ion reservoir/ PBDB-T/ MoO_3/Ag	300-800	4×4	200	75	—	—	6
2T2R1C	450-660	3×3	—	5	—	—	7
Au/Ti/Gr/SiO_2 /i-Si/TIA/Au /ZnO/p^+-Si	550-1300	6×6	96	≈20000	—	—	8
Ag/AF teflon/ MoO_3/ BHJ/SnO_2/ITO	600-1300	3 (line array)	69	86	—	—	9
Ga_2O_3/ p-Si	940	Single	60	0.42	—	—	10

Values are approximate since there is no standard measurement tested.

References

1. Lichtsteiner P, Posch C &Delbruck T. A 128×128 120 dB 15 μs latency asynchronous temporal contrast vision sensor. *IEEE J. Solid-State Circuits* **43**, 566–576 (2008).
2. Suh Y, et al. A 1280×960 Dynamic Vision Sensor with a 4.95-μm Pixel Pitch and Motion Artifact Minimization. in *Int. Symp. Circuits Syst. Proc.* 1–5 (IEEE 2020).
3. Prophesee Evaluation Kits, 2020. Available: https://www.prophesee.ai/event-based-evk/.
4. T. Finateu *et al*. 5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86μm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline in *International Solid-State Circuits Conference* 112-114 (IEEE 2020).
5. Y. Suh et al., "A 1280×960 Dynamic Vision Sensor with a 4.95-μm Pixel Pitch and Motion Artifact Minimization," in *International Symposium on Circuits and Systems* 1-5 (IEEE 2020)
6. Lin Q, et al. Event-driven retinomorphic photodiode with bio-plausible temporal dynamics. *Nat. Nanotechnol.* **20**, 1213–1220 (2025).
7. Zhou Y, et al. Computational event-driven vision sensors for in-sensor spiking neural networks. *Nat. Electron.* **6**, 870–878 (2023).
8. Wu, Y. et al. A Spiking Artificial Vision Architecture Based on Fully Emulating the Human Vision. *Adv. Mater.* **36**, 2312094 (2024).
9. Wu, S.-E. et al. Retinomorphic Motion Detector Fabricated with Organic Infrared Semiconductors. *Adv. Sci.* **10**, 2304688 (2023).
10. Kumar M, Park H &Seo H. A Single-Pixel Event Photoactive Device for Real-Time, In-Sensor Spatiotemporal Optical Information Processing. *Adv. Mater.* **37**, 2406607 (2025).