Preview

# Artificial perception system for dynamic recognition and trajectory extraction

**Pengshan Xie[1,2] and Johnny C. Ho[1,2,3,4,*]**
[1]Department of Materials Science and Engineering, City University of Hong Kong, Hong Kong SAR 999077, China
[2]Shenzhen Research Institute, City University of Hong Kong, Shenzhen 518057, China
[3]State Key Laboratory of Terahertz and Millimeter Waves, City University of Hong Kong, Hong Kong SAR 999077, China
[4]Institute for Materials Chemistry and Engineering, Kyushu University, Fukuoka 816-8580, Japan
*Correspondence: johnnyho@cityu.edu.hk
https://doi.org/10.1016/j.matt.2025.102060

**Multi-modal in-sensor computing offers a promising approach to enhancing data processing efficiency and achieving breakthroughs in bandwidth and energy consumption. In a recent study published in *Device*, Chang et al. introduced a porous-Si/Si artificial visual system created through electrochemical etching. The voltage regulation performance, driven by the photovoltaic effect, has the potential to inspire the development of next-generation neuromorphic devices.**

The visual system is essential for the survival of living organisms, serving as a primary means of gathering environmental information. It allows organisms to perceive various visual cues, such as light, color, shape, motion, and depth, all crucial for survival and reproduction. Over millions of years, biological vision has evolved to meet the specific survival needs of different species. To overcome the limitations of natural vision, machine vision, inspired by biomimicry, has been developed and is advancing rapidly. Combined with deep learning and artificial intelligence, machine vision systems have great potential for integrating spatial and temporal data streams.[1] For machine vision that draws inspiration from biology and aims to surpass it, the primary applications are autonomous driving, industrial automation, robotics, and medical imaging. Consequently, broadband in-sensor processing and dynamic sensing are critical.

In a recent issue of *Device*, Chang et al. introduced an innovative approach to integrating dynamic recognition and trajectory extraction.[2] Their use of a porous-p-Si/p-Si heterojunction, created through electrochemical etching, features a strong built-in electric field that efficiently separates photo-generated electron-hole pairs, resulting in a rapid photoelectric response within microseconds (Figure 1A). Notably, the device's differential photoelectric response was leveraged to extract the trajectory of light sources. The source-drain voltage difference, caused by the irradiation of a visible spot at various locations in the device channel, allows for capturing the spot's movement path. This position-dependent differential voltage naturally oscillates between positive and negative, paving the way for further integration of edge enhancement algorithms.

Some pioneering efforts have focused on the neuromorphic response and in-sensor computing of devices by incorporating various functional layers, such as ferroelectric, phase change, and thermoelectric materials. Among these, constructing heterojunctions has been a complex choice due to challenges like large-area integration, device uniformity, and compatibility with current complementary metal-oxide semiconductor (CMOS) integration processes.[3] Introducing defect energy levels for carrier trapping and release offers another pathway to realizing a bionic vision system, especially for large-area channel preparation.[4] In Chang's study, electrochemical etching created a porous-p-Si layer on the p-Si surface. The etching process can be precisely controlled by adjusting the current level and etching time, while large-area etching ensures device uniformity. The entire device, utilizing silicon as the channel material, is compatible with current chip fabrication processes, offering the potential for large-scale integration.

Traditionally, sensing systems are physically separate from computation units due to differing usage requirements and practical process limitations. As a result, sensor units capture large amounts of raw data locally, which must then be transferred to a local computing unit or a cloud-based system.[5] Meanwhile, computations are typically conducted digitally using traditional von Neumann computing architectures, which lack parallel computation capabilities. This process poses significant energy consumption, response time, data redundancy, communication bandwidth, and security challenges. In an in-sensor computing architecture, individual sensors or integrated sensor arrays can directly process environmental information, eliminating the need for constant data exchange between the sensor and processor, thereby combining sensing and computational functionality.[6] Thus, in Chang's work, energy band bending due to the porous-p-Si/p-Si heterojunction enabled a real-time optical signal response, ensuring effective front-end signal collection. Unlike other approaches, this work achieved real-time localization of the light source signal position. With a nearly closed four-terminal electrode structure, the trajectory of the light source can be effectively pinpointed using the voltage difference between opposing electrodes. The inherent directionality of the voltage difference even allows for identifying the light source's movement path, including its direction and pattern. Figures 1B and 1C illustrate the change in voltage difference between two pairs of electrodes as the laser spot moves

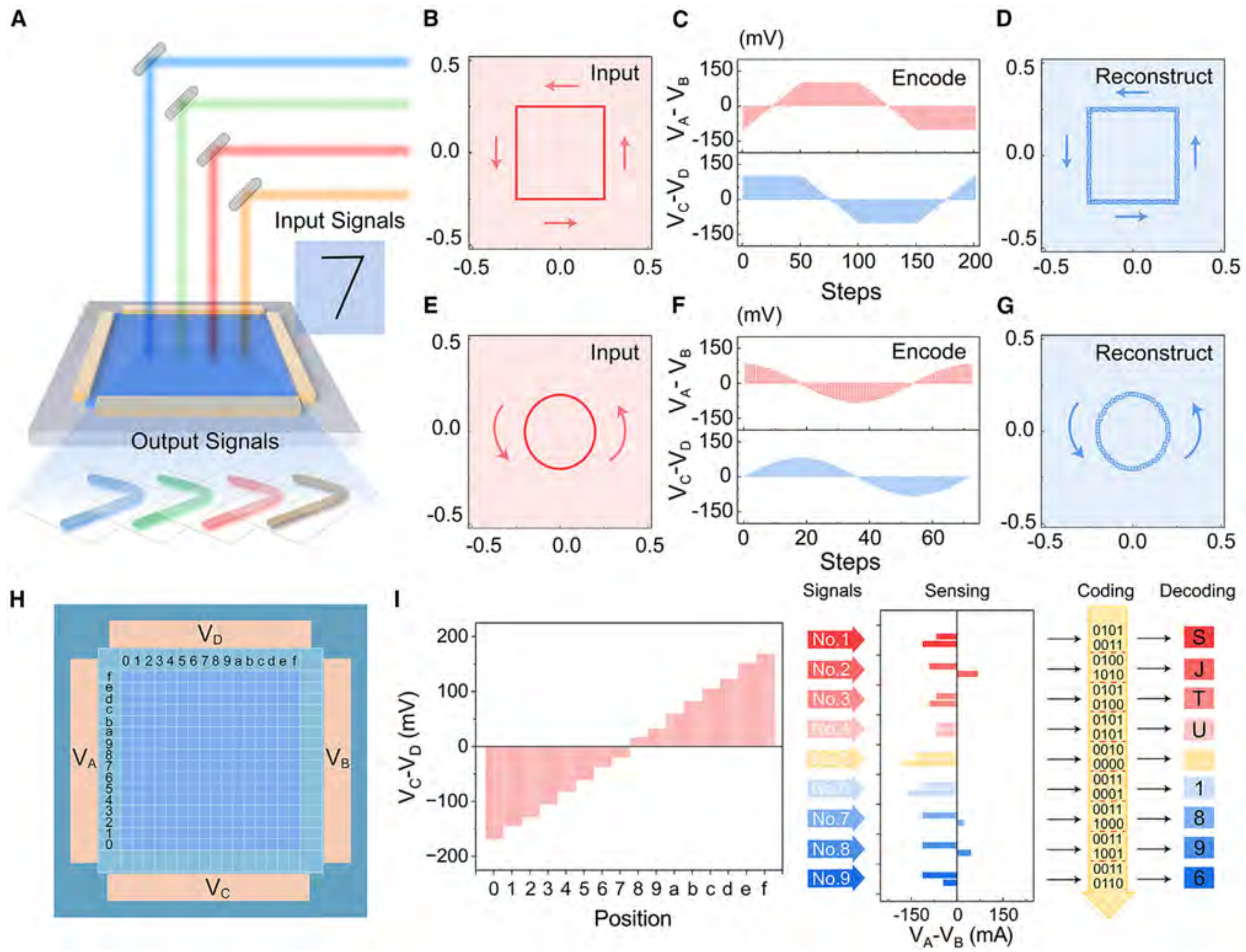


**Figure 1. The structure and optoelectronic working processes of the device**
(A) Schematic diagram of device structure and broadband sensing.
(B) Square moving trajectory at optical signal input.
(C) Voltage difference between opposite electrodes with the square moving trajectory.
(D) Reconstructing the square trajectory of the moving light source.
(E) Circular moving trajectory at optical signal input.
(F) Voltage difference between opposite electrodes with the circular moving trajectory.
(G) Reconstructing the circular trajectory of the moving light source.
(H) Schematic diagram of device sensing area division with 8-bit information coding.[2]
(I) The relationship between the vertical position of the laser spot and the voltage difference and the schematic diagram of the sensing and encoding process. Adapted from Chang et al.[2]

in a counterclockwise square. The optical signal's trajectory can then be reconstructed (Figure 1D). This perception and reconstruction process is also effective for signals from circular moving trajectories (Figures 1E–1G).

To further advance the in-sensor computing process, devices continually evolve toward multi-dimensional sensing, finer front-end processing, easier integration and fabrication, faster response times, and lower power consumption. At the front-end perception level, multi-modal perception is a priority for devices to capture external information. Currently, biomimetic devices based on biological perception primarily imitate a few key functions. However, for example, in visual bionics, the visual sensing wavelength of certain organisms is often broad. Simultaneously, the visual system can sense multi-dimensional information such as temperature, humidity, gas, and pressure. Therefore, broadband sensing is a vital feature of visual sensing devices. Compared to materials with relatively poor environmental stability, such as perovskite, black phosphorus, and organic semiconductor films, the porous-p-Si/p-Si heterojunction demonstrated stable detection from ultraviolet to near-infrared in the air atmosphere.[7] Beyond information perception, improving signal-to-noise ratio, suppressing background noise, and enhancing feature signals are also critical. Effective signal extraction can significantly reduce computational load and increase the efficiency of high-level processing tasks. In Chang's work, the device was divided into 20 regions based on the laser spot size (Figure 1H). In this quadrilateral electrode structure, 8-bit binary signals can be encoded by combining the respective encoding capabilities of the voltage difference between the transverse and longitudinal electrodes (Figure 1I). Achieving this position-dependent multi-bit writing capability hinges on the apparent signal difference between neighboring electrodes. In this study, the potential difference between adjacent intervals exceeded 20 mV, ensuring accurate signal recognition.

High integration is the future direction for device development, encompassing materials, devices, architectures, and algorithm co-design. The memristor offers a viable concept, with classical crossbar array structures easily integrated at high density into CMOS-compatible processes. This two-dimensional transverse and longitudinal structure can naturally perform vector-matrix multiplication, one of the most important and frequent operations in machine learning. As data are stored and transmitted by the array, conductance can be multiplied and summed under Ohm's law and Kirchhoff's current law.[8] Another emerging strategy is integrating resistive switching devices with three-terminal transistor devices. Integrating one transistor and one memristor (1T1R) or multiple transistors and one memristor (MT1R) can enhance the device's computational dimension. When equipped with threshold switches, devices of this structure offer a broader range of neural network compatibility.[9] Arrays of in-sensor computing devices based on dynamic recognition and trajectory extraction also hold the potential for further integration and versatility. Smaller device partitions and the integration of resistive switches provide strategies for subsequent development.

After decades of reliance on von Neumann architectures, we are transitioning into the next era. With the rapid development of neuromorphic computing, we need new approaches to handle the burgeoning data processing demands and computing paradigms ahead. In-sensor computing is currently a crucial development direction and a key breakthrough from existing architectures. Combining large-area, multi-component integration and multi-modal sensing allows for more sophisticated adaptive recognition capabilities under low-energy conditions. In this work, the combination of in-sensor processing and real-time trajectory recognition demonstrates the ability for position-dependent multi-bit coding with broadband detection, showcasing potential applications in smart driving, robotics, and intelligent identification.

## ACKNOWLEDGMENTS

We acknowledge the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. CRS_CityU101/24), the Innovation and Technology Fund (MHP/126/21) from the Innovation and Technology Commission of the Hong Kong Special Administrative Region, China, the Science Technology and Innovation Committee of Shenzhen Municipality (project no. JCYJ20230807114910021), and the Guangdong Basic and Applied Basic Research Fund (project no. 2024A1515011922).

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Chen, C., Chen, Z., Liu, D., Zhang, X., Gao, C., Shan, L., Liu, L., Chen, T., Guo, T., and Chen, H. (2024). Three-terminal quantum dot light-emitting synapse with active adaptive photoelectric outputs for complex image processing/parallel computing. Matter *7*, 3891–3906. https://doi.org/10.1016/j.matt.2024.06.050.

2. Chang, K., Yu, X., Ren, W., He, Y., Gan, Z., Ma, S., Huang, H., Jin, X., Zhang, Y., and Wang, H. (2024). Artificial vision system based on porosity heterojunction for in-sensor processing and dynamic trace extraction. Device *3*, 100598. https://doi.org/10.1016/j.device.2024.100598.

3. Chen, J., Zhou, Z., Kim, B.J., Zhou, Y., Wang, Z., Wan, T., Yan, J., Kang, J., Ahn, J.-H., and Chai, Y. (2023). Optoelectronic graded neurons for bioinspired in-sensor motion perception. Nat. Nanotechnol. *18*, 882–888. https://doi.org/10.1038/s41565-023-01379-2.

4. Dang, B., Zhang, T., Wu, X., Liu, K., Huang, R., and Yang, Y. (2024). Reconfigurable in-sensor processing based on a multi-phototransistor–one-memristor array. Nat. Electron. *7*, 991–1003. https://doi.org/10.1038/s41928-024-01280-3.

5. Liao, F., Zhou, Z., Kim, B.J., Chen, J., Wang, J., Wan, T., Zhou, Y., Hoang, A.T., Wang, C., Kang, J., et al. (2022). Bioinspired in-sensor visual adaptation for accurate perception. Nat. Electron. *5*, 84–91. https://doi.org/10.1038/s41928-022-00713-1.

6. Shao, H., Wang, W., Zhang, Y., Gao, B., Jiang, C., Li, Y., Xie, P., Yan, Y., Shen, Y., Wu, Z., et al. (2025). Adaptive In-Sensor computing for enhanced feature perception and broadband image restoration. Adv. Mater. *37*, 2414261. https://doi.org/10.1002/adma.202414261.

7. Wang, Z., Wu, H., Burr, G.W., Hwang, C.S., Wang, K.L., Xia, Q., and Yang, J.J. (2020). Resistive switching materials for information processing. Nat. Rev. Mater. *5*, 173–195. https://doi.org/10.1038/s41578-019-0159-3.

8. Yang, Y., Pan, C., Li, Y., Yangdong, X., Wang, P., Li, Z.-A., Wang, S., Yu, W., Liu, G., Cheng, B., et al. (2024). In-sensor dynamic computing for intelligent machine vision. Nat. Electron. *7*, 225–233. https://doi.org/10.1038/s41928-024-01124-0.

9. Zhou, F., and Chai, Y. (2020). Near-sensor and in-sensor computing. Nat. Electron. *3*, 664–671. https://doi.org/10.1038/s41928-020-00501-9.