SCIENCE ROBOTICS | RESEARCH ARTICLE

BIOMIMETICS

Stereoscopic artificial compound eyes for spatiotemporal perception in three-dimensional space

Byungjoon Bae[1]†, Doeon Lee[1]†, Minseong Park[1]†, Yujia Mu[1], Yongmin Baek[1], Inbo Sim[1], Cong Shen[1], Kyusang Lee[1,2]*

Arthropods' eyes are effective biological vision systems for object tracking and wide field of view because of their structural uniqueness; however, unlike mammalian eyes, they can hardly acquire the depth information of a static object because of their monocular cues. Therefore, most arthropods rely on motion parallax to track the object in three-dimensional (3D) space. Uniquely, the praying mantis (Mantodea) uses both compound structured eyes and a form of stereopsis and is capable of achieving object recognition in 3D space. Here, by mimicking the vision system of the praying mantis using stereoscopically coupled artificial compound eyes, we demonstrated spatiotemporal object sensing and tracking in 3D space with a wide field of view. Furthermore, to achieve a fast response with minimal latency, data storage/transportation, and power consumption, we processed the visual information at the edge of the system using a synaptic device and a federated split learning algorithm. The designed and fabricated stereoscopic artificial compound eye provides energy-efficient and accurate spatiotemporal object sensing and optical flow tracking. It exhibits a root mean square error of 0.3 centimeter, consuming only approximately 4 millijoules for sensing and tracking. These results are more than 400 times lower than conventional complementary metal-oxide semiconductor–based imaging systems. Our biomimetic imager shows the potential of integrating nature's unique design using hardware and software codesigned technology toward capabilities of edge computing and sensing.

INTRODUCTION

Arthropod eyes are sophisticated and exquisite convex visual organs prevalent among insects (*1*, *2*). These eyes are composed of numerous ommatidia, with each ommatidium containing a photoreceptor cell capable of detecting light from a specific segment of the visual field. Consequently, arthropod eyes exhibit distinctive attributes, including a wide field of view (FoV), high motion sensitivity, and infinite depth of field, rendering them highly effective for optical flow sensing (Fig. 1A) (*3*–*7*). However, in determining the distance to an object, most insects and crustaceans rely on the relative movement of either the object or themselves. This is because they use monocular vision to perceive the three-dimensional (3D) world, which limits depth perception for stationary objects (Fig. 1B). Notably, the visual system of praying mantises provides considerable overlap between the visual fields of their left and right eyes, unlike in many insects, where there is a distinct bifurcation (Fig. 1C) (*8*–*11*). This overlapping FoV enables binocular disparity, facilitating depth perception through stereopsis, even in the absence of motion parallax. Inspired by the praying mantis stereopsis, we have demonstrated optical flow sensing in 3D space with stereoscopic artificial compound eyes.

Obtaining 3D information through the imager in traditional signal processing techniques, such as the time-of-flight and triangulation (*12*) methods, typically necessitates complex peripheral circuitry, extensive memory storage, and considerable communication bandwidth between memory and computing units. In conventional systems, optical flow information is often derived from consecutive images (*13*, *14*) via dedicated algorithms such as Lucas-Kanade (*15*, *16*) and Horn-Schunck (*17*) methods. Although these algorithms can generate a vector field that represents motion, they demand substantial computational resources, large memory units, extended processing times, and considerable energy consumption. In contrast, machine learning (ML) hardware and algorithms offer a more efficient alternative for processing complicated spatiotemporal data in 3D space, bypassing the limitation inherent to geometry-based calculation methods. However, deploying sophisticated and large-scale ML algorithms, such as deep neural networks (DNNs), still requires high computing power (*18*–*20*). Thus, big data processing with DNN often relies on cloud computing to achieve the desired classification and regression outcomes, which poses challenges for on-site data processing (*21*, *22*).

To alleviate the issues, we have adopted both in-sensor and near-sensor data processing approaches, using artificial synapses integrated into each pixel and a back-end artificial neural network (ANN) implemented on a local processer, respectively. Drawing inspiration from the vision systems of insects, which rely on nonspiking graded neurons (*23*), artificial synapses are integrated at the pixel level for rapid motion perception, mirroring the functionality of graded neurons. This integration of artificial synapses with a photodetector array enables direct encoding of spatiotemporal information at the sensor level through optical programing and one-shot readout processes (*24*–*26*), substantially reducing power consumption by decreasing memory usage and frequency of communication between memory and computing units. The processed output from the sensor was immediately available for use by computing units, facilitating swift motion detection and spatiotemporal object tracking while minimizing data transfer and energy consumption. However, the limited access to powerful computing hardware limits the implementation of DNN for edge computing. To overcome these challenges, we used compact split learning (SL) and federated learning (FL) (*27*) to achieve high accuracy and low latency with minimal data storage on limited computing resources at local system (*28*–*30*). Federated split learning (FSL) combines the benefits of FL, which enables parallel processing across

---

[1]Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904, USA. [2]Department of Material Science and Engineering, University of Virginia, Charlottesville, VA 22904, USA.
*Corresponding author. Email: kl6ut@virginia.edu
†These authors contributed equally to this work.

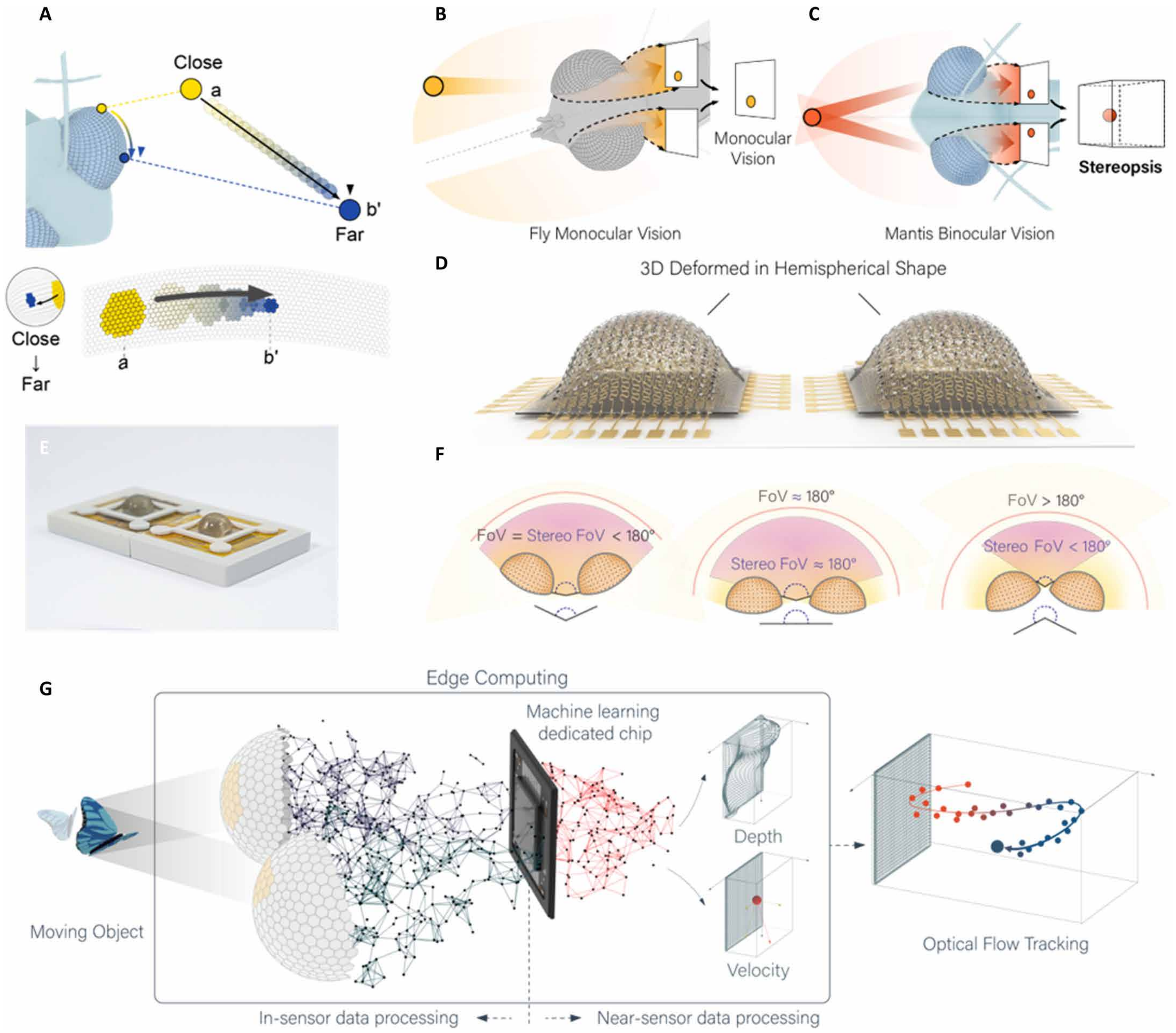

Fig. 1. Illustration of stereoscopic artificial compound eye system. (**A**) Schematic illustration of 3D object and motion perception using compound eyes via motion parallax. The compound eye is sensitive to motion, but it is challenging for insects to estimate the distance to the static object. The figure illustrates that the object in the "a" position was detected by more ommatidia than the identical object in the position "b." (**B** and **C**) Comparison between vision systems of fly and mantis. (B) There is minimal visual field overlap between two eyes in the vision of the fly. (C) Mantis uses stereopsis that utilizes two eyes to collect visual information of a single object. (**D**) Schematic illustration of stereoscopic artificial compound eyes. A pair of hemispherical-shaped 16-by-16 FPAs are arranged in parallel. (**E**) Photograph of fabricated and stereoscopic 16-by-16 artificial compound eyes. (**F**) Stereoscopic FoV of paired compound eyes depending on their arrangement. The visual system with two eyes placed at an angle of 180° exhibits the maximum stereoscopic FoV close to 180°. (**G**) Process flow of the object tracking in 3D space using stereoscopic compound eyes and ANNs. Crossbar array of pixel encodes the optical stimulus via in-sensor processing, and then the generated data from the image sensor are fed into the FSNN loaded in the local processor to calculate the velocity of object, direction of movement, and position in 3D space.

distributed clients, with SL, which partitions the model between clients and a central server during training. This approach not only offers a shorter training time compared with SL alone but also alleviates the storage and processing demands on resource-constrained devices compared with FL. SL divides DNN layers into segments, each managed and stored by different agents, whereas FL enables multiple agents to cooperatively train intricate ML models locally without transmitting individual data to the central processor. The integration of SL and FL principles in FSL facilitates efficient computing in terms of communication, storage, and computation. In our demonstration, the agents in FL are represented by the left and right artificial eyes, and FSL has been effectively used for object tracking, leveraging its ability to extract complex features from multiple sensors.

In this study, we implemented the previously discussed FSL as a near-sensor computing algorithm for the accurate decision and in situ response. Our FSL design involves processing encoded outputs from both the left and right artificial eyes through two separate convolutional neural networks (CNNs) that share weight values. This setup is followed by a single fully connected layer (FCL) that merges the outputs. The result from this neural network configuration indicates the position of the detected object in 3D coordinates and its velocity, providing comprehensive tracking information in real time. The efficacy of our system for object tracking in 3D space was empirically evaluated by calculating mean square error (*31*). Our experimental results confirmed that the codesigned software and hardware system achieved precise in situ object movement tracking. Notably, the system maintained an error margin of less than 0.3 cm and demonstrated a rapid temporal processing rate of 1.8 ms. This level of performance was attained although we used a compact, low-performance microprocessor for computing.

RESULTS

Overall system

Using the codesigned hardware and software platform previously outlined, we demonstrated the capabilities of stereoscopic artificial compound eyes for object detection and optical flow sensing in 3D space. Initially, we used epitaxial liftoff technology to fabricate thin-film indium gallium arsenide (InGaAs) photodiodes on a flexible substrate. These photodiodes were then paired with hafnium oxide (HfO_2)–resistive random-access memory (ReRAM) units to create a one photodiode–one resistor (1P-1R)–based focal plane array (FPA) on a flexible Kapton substrate. This configuration allows the optical signals received by the photodiodes to be modulated, effectively emulating pigment cells and receptors found in the arthropod visual system (*32*). Subsequently, each pixel was combined with a polymethyl methacrylate–based microlens array to improve focusing capabilities, mimicking the structure and function of an ommatidium. Then, we shaped the fabricated FPAs into a hemispherical form with a 20-mm radius, exploiting the mechanical flexibility of the membrane-based system to modulate both the shape and radius of the structure. To facilitate stereoscopy, a pair of artificial compound eyes were positioned 2 cm apart on a planar surface, achieving a 160° FoV (Fig. 1D and fig. S1). The artificial compound eyes were then mounted on 3D printed structures (Fig. 1E) and aligned to maximize the stereo FoV area, which was the overlapped FoV area between the two eyes (Fig. 1F). Last, the compound eyes were installed onto a customized printed circuit board for the inference readout process and electrical control (fig. S2), integrating the entire system into a cohesive unit capable of advanced optical sensing and object detection in 3D space.

Figure 1G illustrates the process flow for object sensing and signal processing using our hardware and software codesigned stereoscopic compound eyes. Conventional imaging and processing systems necessitate a range of components, including a backplane circuit, readout circuit, analog-to-digital converter, memory, and computing units, to capture data at the sensor level and to process signals at the processor level. In contrast, our approach integrates ReRAMs directly with the sensor and a compact microprocessor connected to the sensor for edge computing capabilities. The ReRAMs, programmable through photogenerated currents from the photodetectors, enable in-sensor processing that generates 1D encoded data from consecutive images through the one-shot inference process, bypassing the need for a sequential readout process. By processing signals directly at the pixel level using ReRAMs, we reduced the dimension/size of the generated data; the need for data transportation between sensors, memory, and computing units; and the reliance on software-based calculations and complex peripheral circuitry (described in the following section in detail). With the described scheme, the processed data at each pixel were fed into the local neural network to detect objects in the 3D field using the compact and dedicated ML system (Sony Spresense board).

Fabrication and characterization of stereoscopic compound eyes

To demonstrate the optical flow sensing in 3D space, we designed and fabricated artificial compound eyes that consist of 16-by-16 InGaAs photodiode and HfO_2 ReRAM (1P-1R) pixels on a flexible Kapton substrate (Fig. 2A). Here, a square pixel array was used instead of a symmetrical structure to simplify row-by-row and column-by-column current summation process. Figure 2B depicts magnified view of the compound eye, consisting of thin-film InGaAs-based 1P-1R and the serpentine-shaped interconnections between the 1P-1R FPAs, featuring robustness to external strain and mechanical deformation (*33*–*35*). The detailed fabrication process of the 1P-1R array is described in the “Fabrication of 1P-1R array on flexible platform” section in Materials and Methods and figs. S3 to S5. After the fabrication of the 1P-1R array, a microlens array was integrated (see the “Integration of microlens array on FPA” section in the Supplementary Materials and fig. S6) onto the photodiode array, and FPA was deformed into a hemispherical shape (Fig. 2C and fig. S7). Figure 2D shows the schematic description of the microlens and 1P-1R array. The microlens was integrated on an individual photodiode as a corneal lens in compound eye structure (*36*), which samples a limited visual field only within its acceptance angle. This configuration emulates the ommatidium in arthropod eyes, which enhances their sensitivity for the motion detection (*37*, *38*). We designed the microlens on the basis of the optical simulation conducted using COMSOL software (see the “Microlens simulation” section in Supplementary Materials and fig. S8). The detailed fabrication procedure of microlens array is described in the “Fabrication of microlens” section in Materials and Methods. With the fabricated photodiode integrated with the microlens, we characterized the FoV by measuring the photogenerated current under a −3-V bias applied to the photodiode, across various light incident angles ranging from −35° to 35°. The fabricated pixel integrated with microlens exhibited 8.5° of light acceptance angle (Fig. 2E).

After the fabrication of compound eyes, we first examined the characteristics of individual photodiode and ReRAM. To confirm the dynamic range of photodiode, we measured the current-voltage (*I-V*) characteristics of the fabricated InGaAs photodiode under the illumination of light at a wavelength of $\lambda = 530$ nm with varying light intensities from dark to 1.74 W/m^2 (Fig. 2F). Figure 2G shows the *I-V* characteristics of the ReRAM, featuring the nonvolatile resistive switching behavior after the initial forming process. Statistical distribution of set and reset current measured using 100 fabricated ReRAMs is shown in fig. S9. The transition between the high-resistance state (HRS) and the low-resistance state (LRS) is associated with the drift of oxygen vacancies in the HfO_2 medium (*39*), driven by external electric fields. For the initial forming process, voltage sweeping was carried out from 0 to 8 V, with a 1-mA current compliance. The formation of a conductive filament became evident upon observing abrupt conductance change (indicated by the gray line). The ReRAM was subsequently reset to HRS using a sweeping voltage ranging from

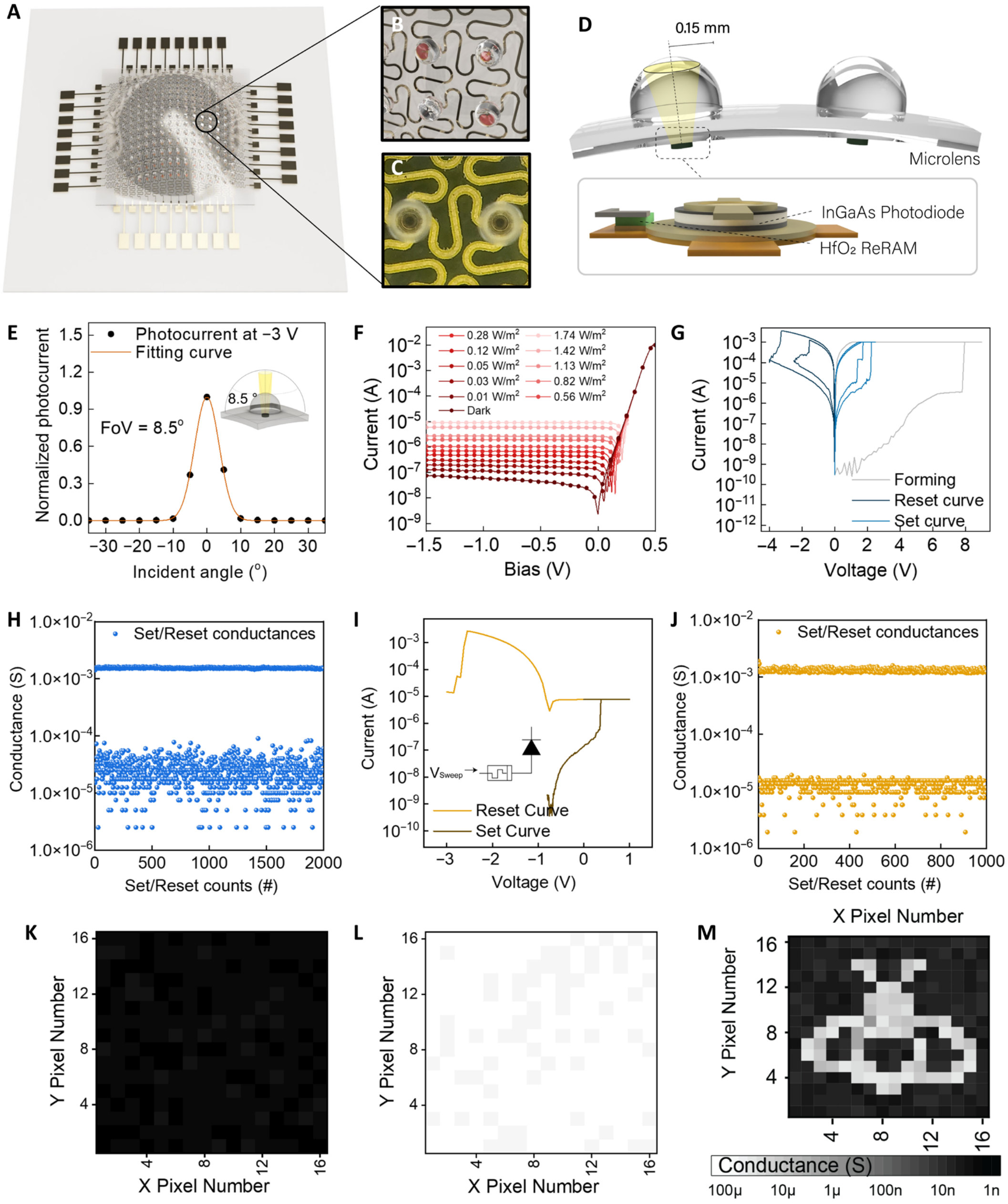

Fig. 2. Device characterizations. (**A**) Illustration of hemispherically formed FPA integrated with microlens array. (**B**) The image displays serpentine metal pattern connection between the pixels and integrated microlens. (**C**) Photograph of the fabricated pixels integrated with microlens. (**D**) Illustration of the fabricated microlens and 1P-1R ReRAM array structure. InGaAs photodiode and HfO_2-based ReRAM were integrated to implement 1P-1R structure. To focus the light within the limited acceptance angle, the microlens with 150-μm radius and 100-μm pilar height was fabricated and integrated on the top of the photodiode. (**E**) FoV of each pixel integrated with microlens. The experimental data indicate an FoV of 8.5°. Inset: Schematic illustration describes the light acceptance range of the pixel. (**F**) *I-V* characteristics of InGaAs photodiode under various light illumination intensities from dark to 1.74 W/m^2. (**G**) *I-V* characteristics of ReRAM. The gray line shows the initial forming process. The dark blue line indicates the resetting process, whereas the light blue line indicates the setting process. (**H**) Endurance characterization of ReRAM during the application of 2000 cycles of set and reset pulses (1.2 and −2 V for 0.5 ms, respectively). (**I**) Optical programming of 1P-1R pixel. *I-V* characteristic of 1P-1R. Brown line indicates the optical programming (set ReRAM by the photocurrent under 1.6 W/m^2 of incident light) process, and yellow line shows the resetting process. Inset shows the direction of sweeping voltage during the programming process. (**J**) Endurance characterization of 1P-1R during the application of 1000 cycles of 1.5-V set and −2.5-V reset pulses with 0.05-ms pulse width under 1.6 W/m^2 of incident light. Conductance maps (**K**) before and (**L**) after a forming process of the fabricated 1P-1R array. (**M**) Readout signal from 1P-1R FPA after optical programing of each pixel using light illumination of bee shape.

0 to −4 V (depicted by dark blue lines), followed by the setting process using a sweeping voltage from 0 to 2.5 V (light blue lines). To evaluate the endurance of the ReRAM, we applied consecutive set (1.2 V/500 μs) and reset voltage (−2 V/500 μs) pulses in turn, toggling the state of ReRAM a total of 2000 times. Subsequently, we applied readout voltage pulses (0.2 V/500 μs) and acquired a series of set and reset currents (Fig. 2H). Statistical analysis of set and reset conductance showed small standard deviations (SDs) of $\sigma = 9.273 \times 10^{-6}$ and 3.583×10^{-6} S, respectively, which confirmed the superior endurance performance of the fabricated ReRAM on Kapton substrate.

On the basis of characteristics of individual devices, we designed and characterized the 1P-1R pixel to be used for in-sensor computing. Figure 2I shows the *I-V* characteristics of the 1P-1R pixel under illumination (1.6 W/m^2). The notable conductance shift of the ReRAM was observed after the application of sweep voltage from 0 to 1 V (brown lines) to the pixel under the light illumination with intensity of 1.6 W/m^2. The ReRAM was reset to HRS by sweeping the input voltage from 0 to −3 V (yellow lines) to the pixel. These results confirm that the photogenerated current can be used to program the 1P-1R pixel. Then, the endurance of 1P-1R pixel was examined by applying the alternating set (1.2 V/500 μs) and reset (−2.5 V/500 μs) pulses for a total of 1000 cycles under the light illumination at an intensity of 1.6 W/m^2 (Fig. 2J). This result verifies the durability of 1P-1R to multiple pulse-based operations that enable continuous detection of the object. Furthermore, we evaluated the programmability of the entire 16-by-16 1P-1R array. Figure 2K shows the conductance of the 256 ReRAMs within the array before the forming process. After the forming process, all pixels demonstrated a successful transition from the HRS to the LRS, as represented in Fig. 2L. After the resetting process, we were able to selectively program the ReRAMs to a desired pattern (here, the shape of a bee) when the array was exposed to the patterned light illumination with an intensity of 0.6 W/m^2, as described in Fig. 2M. The signal variation analysis result of the fabricated hemispherical 1P-1R array along the end-to-end lines is illustrated in fig. S10.

In-sensor computing system

After the characterization of individual pixels and arrays, we examined the functionality of fabricated 16-by-16 1P-1R to directly encode optical stimuli from a wide FoV at the sensor for the reduction of data dimension and size. Figure 3A describes the schematic illustration of the projected area on the hemispherical surface when the target object moves in 3D space during the time τ. In this schematic illustration, the object moves away from the sensor; thus, the projected area on the hemispherical surface became smaller. To achieve energy-efficient optical flow sensing, we first encoded the sequential data detected from moving objects using in-sensor memory-based processing. The initial position (at time t_1) of the object was stored in ReRAMs via optical programming using the photogenerated current from a serially connected photodiode when it was exposed to light. If there was no optical signal to photodiode at the time t_1, the ReRAM maintained its HRS. Once the ReRAM was programmed by optical signal, ReRAM functioned as a load series resistance (R_L) to the integrated photodiode at each pixel. Thus, under the forward bias, the current output from 1P-1R was influenced by total resistance ($R_L + R_{\text{photodiode}}$) of the pixel, whereas the effect of photogenerated current was dominant under the reverse bias (see the "In-sensor computing principles" section in the Supplementary Materials). Therefore, the resistance stored in ReRAM and photogenerated current from the pixel can be individually read by measuring the output current of the pixel in forward and reverse bias regimes, respectively (fig. S11). Readout of the output current of the pixel in both reverse and forward bias regimes allowed 1P-1R to obtain consecutive scene data at the same time (at time t_2). After the readout process, all the ReRAMs in the pixel were reset to HRS to encode the next consecutive images. Figure 3B shows the representative *I-V* characteristic of 1P-1R that emphasizes the effect of programmed ReRAM to the pixel with and without light illumination in both photoconduction and photodetection modes.

Figure 3C shows the designed and applied pulse scheme for the sequential set/read/reset process. Each pulse cycle consisted of a set (memorization) pulse for ReRAM at time t_1, two readout pulses (photoconduction and photodetection modes) at time t_2, and a reset pulse (erasing) right after the readout. Set and reset of each pixel were achieved through the application of 1.0- and −2.5-V voltage pulses with a duration of 100 μs. Readout was performed 1 ms after the set process by applying −0.4- and 0.8-V voltage pulses with a duration of 50 μs. We applied −0.2 V to the pixel to prevent the substantial current from flowing to ReRAM when there was no input pulse applied. This readout scheme was advantageous for in-sensor computing by simplifying the in and out (I/O) operation of obtained image data to external memory (see the "I/O operation in memories" section in the Supplementary Materials).

Figure 3D describes experimental readout results of the 1P-1R pixel with respect to all possible consecutive optical detection scenarios: An object remained within the detection range of the pixel (Fig. 3D-1); an object entered the detection range at time t_2 (Fig. 3D-2); an object exited the detection range at time t_2 (Fig. 3D-3); and the object was not detected to the pixel (Fig. 3D-4). We first applied voltage pulse (−1.0 V) under the light illumination of 1.3 W/m^2 to the pixel to store the light information at t_1. Then, sequential readout pulses for both photoconduction and photodetection modes (0.8 and −0.4 V, respectively) were applied to each pixel at time t_2. When the object was placed on the original position at t_1 and moved out to the position at t_2, a distinguishable combination of readout currents via the application of both forward and reverse bias pulses was achieved for the corresponding combination of output signals. The experimentally measured current outputs from individual pixels using the designed pulse scheme under four different scenarios are shown in fig. S11.

To demonstrate the in-sensor computing process, we showed output from the inference of 16-by-16 artificial compound eyes. Figure 3E illustrates the geometry of setup for the characterization of the array. The left and right eyes were placed 2 cm apart from each other on the flat surface, and the target was placed 5.5 cm above the center of the two eyes at time t_1. Then, we moved the target 10 mm to the right from the initial position at time t_2. The overall experimental process is described in movie S1. To perform the in-sensor computing, instead of procuring pixel-by-pixel image data by applying sequential pulses to the array that the conventional image readout process uses, we encoded the image data using the single inference read pulse that can reduce the dimension of output from 2D to 1D (Fig. 3F). To achieve one-shot inference of the crossbar pixel array of the compound eyes without losing notable optical information, we applied a readout pulse scheme, −0.4- and 0.8-V pulses as described in the above section, to the array both row by row and column by column. Thus, we acquired two pairs of 1D encoded datasets from column-wise and row-wise summation, respectively. Using this readout process, we were able to encode the optical flow cues from 16-by-16

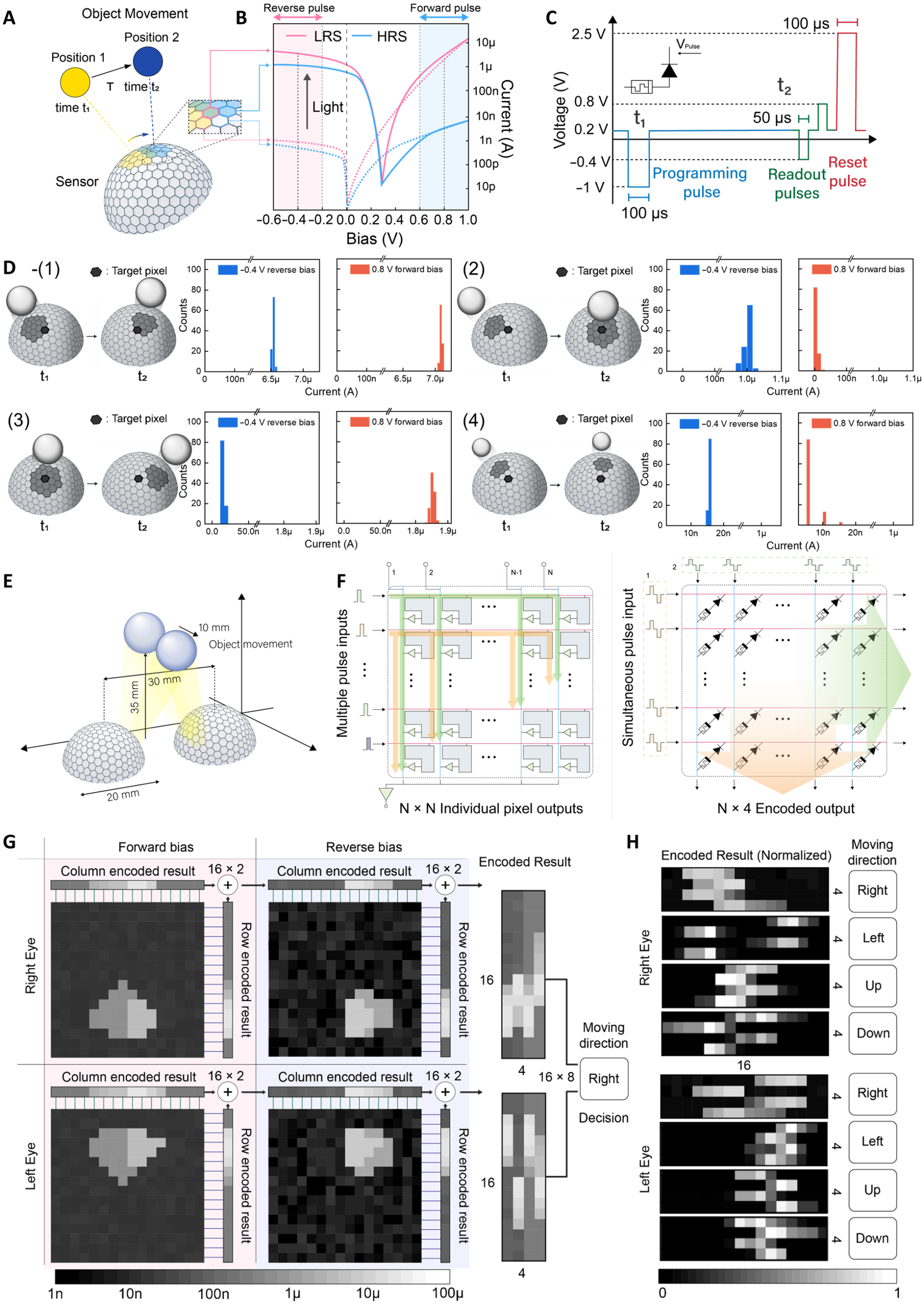

Fig. 3. FPA characterization for in-sensor computing. (**A**) Illustration of activated region on hemispherical compound eye surface during the movement of the target object positions in 3D space. (**B**) *I-V* curves of single pixel for four cases: the HRS of ReRAM under illumination, HRS under dark conditions, LRS under illumination, and LRS under dark conditions. Dashed and solid lines indicate the *I-V* curve without and with light, respectively. Red and blue lines refer to the result when the ReRAM in the 1P-1R structure was in the LRS and HRS, respectively. The readout was performed in both forward and reverse bias regimes (0.8 and −0.4 V, respectively) (**C**) Applied pulse scheme to each pixel for programing/read/reset. Programing and resetting occur through the application of 1.5- and −2.5-V voltage pulse with a duration of 100 μs. Readout is performed 1 ms after the setting pulse by applying 0.8- and −0.4-V voltage pulses with a duration of 50 μs. We applied negative voltage (−0.2 V) bias during the resting time to prevent the notable current flowing to ReRAM. (**D**) Schematic illustration of motion detected on the hemispherical surface and its readout results depending on four cases. (i) The object was detected at time t_1 and still in the range of detection despite the movement at time t_2, (ii) the object was detected at time t_1 but moved out of detection range at time t_2, (iii) the object was not detected at time t_1 but moved into the detection range at time t_2, and (iv) the object was out of detection range at both time t_1 and t_2. Each case was distinguishable according to the measurement result. (**E**) Illustration of experimental geometry for object tracking. (**F**) Readout scheme of the array to obtain image data from the sensor by using conventional CMOS-based readout system (left) and our inference readout system (right). Instead of procuring image data pixel by pixel by applying sequential pulses, we retrieved the encoded result from the image sensor by applying one-shot inference pulses. (**G**) Readout map of individual pixels when measured in the photoconductance regime (0.8 V, 50 μs) (left) and photodetection regime (−0.4 V, 50 μs) (right) at time t_2. Accumulation of row-wise and column-wise output current is also shown in the bottom and left side of figures from the inference operation. (**H**) Encoded data (16-by-4-by-2) were obtained for various object movements. The figure illustrates the examples of the encoded result from both right and left compound eyes in conjunction with the object moving direction.

hemispherical FPA into four channels of a 16-by-1 1D dataset. This process diminished latency and power consumption for the in situ operation of both the front- and back-end systems by alleviating data storage and transmission demands because of the reduced data dimensions from 2D to 1D. Figure 3G describes the readout result of individual pixels in a 2D map to confirm the programmability of the array and the 1D encoded readout result from FPA by applying forward and reverse bias pulses in all rows and columns. Figure 3I illustrates the examples of the encoded images, showing the distinctive result in conjunction with the moving direction of the object. The encoded data were fed into the ANN and loaded in the connected local ML system for the following near-sensor computing.

Back-end near-sensor computing system

Figure 4A describes the edge computing process including in-sensor and near-sensor processing steps for object perception and recognition. For near-sensor computing, an FSL-based neural network (FSNN) was used for object tracking in 3D space with the local microprocessor (Fig. 4B). The encoded data by each compound eye were processed independently by two separate CNNs that share the weight updates. A 3-by-3 kernel was applied to each CNN layer with identical padding. Then, we applied rectified linear unit (ReLU) function to activate the layers, followed by a 2-to-1 average pooling layer to compress the features. The outputs from each CNN were concatenated into an FCL to combine all the learned features for the regression. Eventually, the network generates a 6-by-1 output vector, comprising two distinct 3-by-1 components. The first 3-by-1 vector represents the moving velocity of the object in 3D space, whereas the second 3-by-1 vector corresponds to the position of the object before the movement. To achieve accurate results, the neural network requires a substantial amount of data for the training process. Therefore, we generated 100,000 training and 20,000 test datasets using a ray-tracing 3D simulation (details in the "Generating training and test dataset" section in Materials and Methods and fig. S12). In the simulation, we designed the identical stereoscopic compound eye used in the experiment, including geometry/arrangement of the sensors and the light acceptance angle through the microlens at each pixel. With the simulated dataset, we trained FSNN using the root mean square error (RMSE) as the loss function to gauge the discrepancy between the truth and the tracked values (Fig. 4C). This process yielded training and validation losses of 0.015 and 0.022 during the training process, respectively.

The trained FSNN model was converted to hex data and was uploaded to a local processor using Arduino integrated development environment. The local portable processor used in this study operates at 5 V with a clock frequency of 156 MHz, equipped with a 1.5-megabyte on-chip static random-access memory. Reduced data size and dimensionality through in-sensor computing enable processing of the information using this compact system, negating the necessity for high-powered computing hardware. To confirm the effectiveness of the system for the 3D perception, we used the stereoscopic compound eye system to track objects, detecting both the velocity and the position of an object in the 3D space by simulation based on experimental data (Fig. 4, D and E). Figure 4D shows the tracking outcome, illustrating that the truth value of optical flow sensing (yellow line) closely aligns with the tracked value (blue line; movie S2). Figure 4E presents the recognition results of the object position in 3D space, where the tracked x, y, and z coordinates closely match the truth values (fig. S13). Object tracking conducted under different circumstances (object shapes and light conditions) also shows nearly identical performance (described in figs. S14 and S15). These in situ object tracking computations were achieved within 1.8 ms per frame without transporting the data to the external computation units and memories ("Prediction on edge system" section in Materials and Methods). As shown in Fig. 4 (F and G), notably low RMSE (lower than 0.3 cm) was achieved for both depth and 2D position detection using the designed stereoscopic compound eyes.

Figure 4 (I and J) illustrates the notably simplified object tracking process in 3D space using stereoscopic compound eyes compared with conventional complementary metal-oxide semiconductor (CMOS) image sensor and computing units. The CMOS image sensor (CIS)–based approach requires the process through the image signal processor to read out the signal as a 2D image, through an amplifier and analog-to-digital converter (ADC). Furthermore, optical flow computation and 3D image perception using acquired image data mandate frequent data transportation between memory (e.g., static random access memory, SRAM, and dynamic random access memory, DRAM) and computing units (e.g., arithmetic logic units, ALU, and multiply-accumulate operation unit, MAC). In contrast, our edge computing system does not require frequent memory access to store each scene data or to compute the object tracking in 3D space because the optical flow cues from the scene were processed within the sensor (details in the "Comparison of conventional object tracking with compound eye sensor system" section in the Supplementary Materials). On the basis of the described process flow, we estimated and compared the power consumption of both systems (Fig. 4H). The calculation showed that our compound eye system requires ~75 times less energy than a conventional system for data acquisition and processing to generate the optical flow cues. Furthermore, postsignal processing for optical flow sensing by using FSNN consumed approximately 7000 times less energy than using conventional signal processing hardware and algorithm. The estimation details are shown in the "Power consumption calculations during the preneural network processing" section in the Supplementary Materials. Moreover, the latency for the signal processing was notably reduced (~100 and ~300 times for pre- and postprocessing, respectively) when we used the compound eye system with edge computing (details in the "Power consumption calculations during the neural network processing" section in the Supplementary Materials). This result confirms the unique feature of our system to harness computing at sensor with minimal energy consumption and latency and without using costly and bulky hardware and/or cloud computing, which make it suitable as an in situ object tracking and navigation system.

DISCUSSION

Here, we demonstrated stereoscopic artificial compound eyes integrated with an edge computing system that can track object movement in 3D space. Instead of using a disparity calculation that requires frequent communication between processor and memory, our sensor captures sequential changes in the observed scene using the integrated ReRAM. Furthermore, without using a conventional pixel-by-pixel readout process, a one-shot inference–based readout scheme was used with the 1P-1R crossbar array to encode the data at the sensor. The encoded data were subsequently relayed to a neural network implemented by a compact computing unit adjacent to the sensor, which tracks the position of object in 3D space and its moving direction

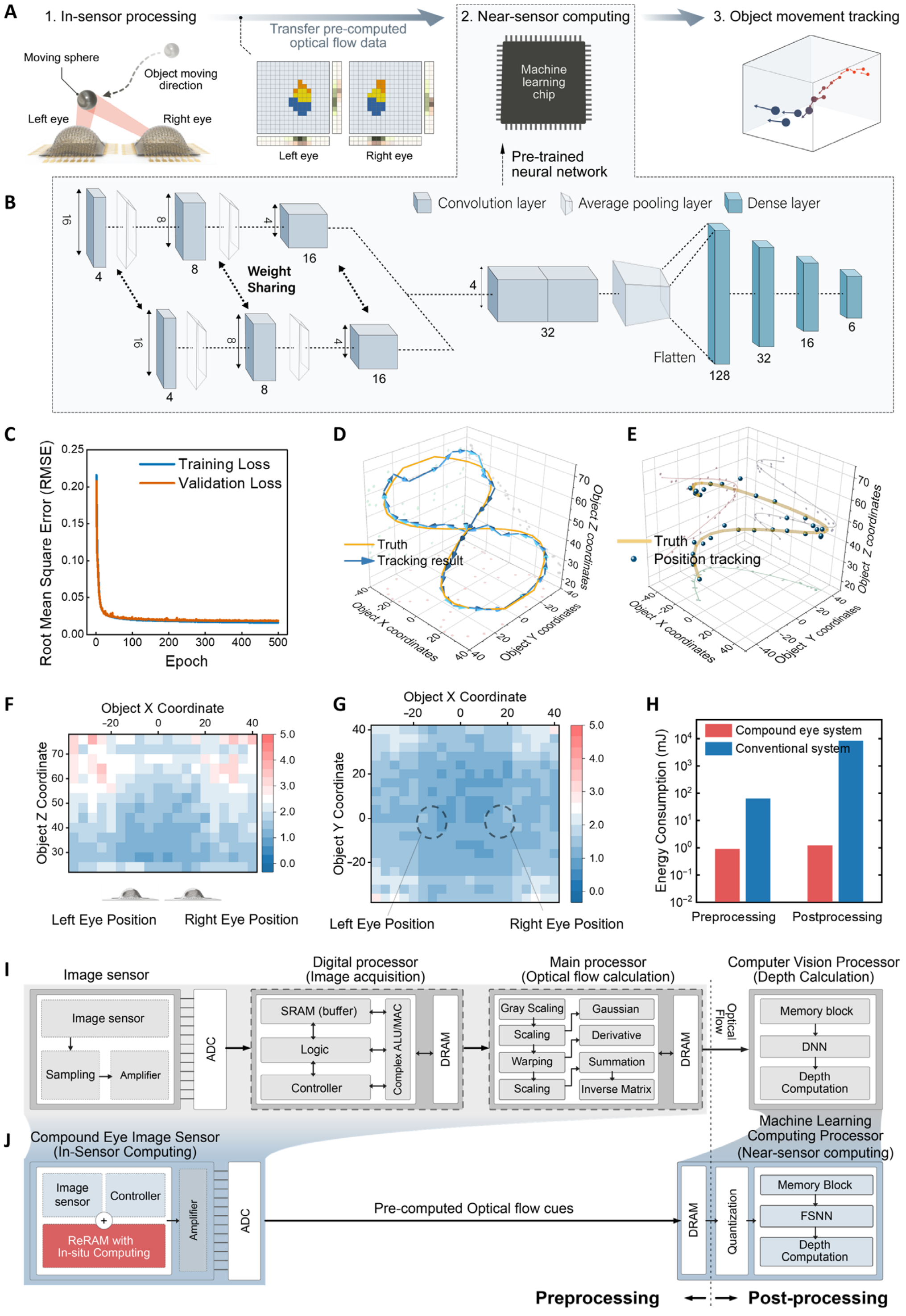

Fig. 4. FSNN-based tracking via near-sensor computing. (**A**) Illustration of the perception and cognition process used in this work. Consecutive scenes encoded by stereoscopic compound eye were processed through a local processor using an FSNN. The local processor provides the result of the object position and optical flow. (**B**) Illustration of the FSNN used in the system. Encoded data from each compound eye were processed individually by two distinct CNN networks, which share weight updates with each other. The obtained features from each CNN network were then merged and fed into an FCL. Six outputs of the FCL indicate the tracked value of the object coordinates and moving direction in Euclidean coordinates. (**C**) Training and validation result for the object position and optical flow detection using neural network with 100,000 datasets. (**D**) 3D plot of detected object moving directions and its truth values. The tracked result and truth movement of the object were well aligned without significant error. (**E**) Detected object position in 3D space and its truth value. 2D contour map of the average RMSE between detected position and truth position for (**F**) the depth ranging of object and (**G**) *x*-*y* position of the object. (**H**) Power consumption comparison between our edge computing–based and conventional CIS–based systems. The calculated consumed energy required for signal preprocessing using conventional system and compound eye system was 65.0 mJ and 895 μJ, respectively. For the postprocessing, the energy consumption using conventional system was 8.63 J per inference, whereas the compound eye system only consumed 1.209 mJ. The result confirmed the effectiveness of compound eyes for energy-efficient signal processing at the sensor. Signal processing flow diagram is shown for the 3D perception using (**I**) the conventional CIS and (**J**) the demonstrated artificial compound eye system.

and velocity. The depth detection accuracy of the system was further enhanced by incorporating stereoscopic sensors. This hardware and software codesigned edge computing system simplifies the complex front- and back-end system to overcome the challenges in in situ perception and cognition toward offering a versatile computing-in-sensor platform.

MATERIALS AND METHODS

Fabrication of 1P-1R array on flexible platform

The inverted InGaAs p-i-n photodiode active device with an AlAs sacrificial layer was grown on InP substrate using molecular beam epitaxy (fig. S3A) (*33*). Ti/Au (10/150 nm) was deposited on top of both the InGaAs wafer and the Kapton substrate using an e-beam evaporator and cold-welded under 2-kN pressure and 205°C conditions for 5 min (*40*, *41*). Then, the bonded substrate was immersed into a mixture solution of $HCl:H_3PO_4 = 4:1$ for 2 hours to etch out the InP wafer. The etched p-InGaAs/Uid-InGaAs films were then etched using inductively coupled plasma–reactive ion etching (ICP-RIE) to form a mesa for 6 min under the conditions of BCl_3 of 20 standard cubic centimeter per minute (sccm), 600-W ICP power, 150-W forward power, 7 mtorr, and 20°C stage temperature, followed by a wet-etching process. The n-InP mesa for the bottom metal electrodes was patterned by 30 s of wet etching using a solution of $HCl:H_2PO_4 = 3:1$. Then, a 150-nm Al_2O_3 dielectric insulator layer was formed via plasma-enhanced atomic layer deposition (PEALD). The via holes were etched using ICP-RIE under the conditions of BCl_3 of 20 sccm, 50-W ICP power, 200-W forward power, 5 mtorr, and 20°C stage temperature for 6 min. The Ti/Pt/Au (5/10/50 nm) top and bottom electrodes of the photodiodes were deposited simultaneously using e-beam evaporation and were patterned by photolithography. Then, PEALD was used to deposit an additional dielectric insulator of 150-nm Al_2O_3, and through holes were patterned on the bottom electrodes of the photodiodes. Subsequently, the bottom electrodes of the ReRAMs were connected to the bottom electrodes of the photodiodes by depositing Ti/Pt (5/25 nm). Then, 5.5-nm HfO_2 was deposited to form the active medium of ReRAM using PEALD, followed by Ta/Pt (50/25 nm) metal deposition through dc magnetron sputtering of Ta (25-W radio frequency, RF power of 5 mtorr, Ar of 20 sccm, and room temperature for 18 min) and e-beam evaporation of Pt. Last, the flexible Kapton substrate was etched out with a serpentine pattern (*42*) to be deformed into the hemispherical shape, using ICP-RIE under the conditions of O_2 of 10 sccm, SF_6 of 5 sccm, 600-W ICP power, 150-W forward power, 7 mtorr, and 20°C stage temperature. The specific processes to fabricate the artificial compound eye are described in fig. S3B.

Fabrication of microlens

To fabricate microlenses, we created a mold with a 16-by-16 pattern of microlens (each with a 150-μm radius) on a silicon substrate. To eliminate the native oxide layer on the prepared silicon substrate, we subjected it to plasma treatment using 950 W of ICP power and 15 W of bias power for 20 s. The treatment used a gas mixture comprising $C_4F_8:SF_6:CF_4$ in a ratio of 60:24:27, with a total flow rate of 110 sccm, whereas the gas pressure was maintained at 20 mtorr. Then, the silicon substrate was isotropically dry-etched (*43*, *44*) under the conditions of SF_6 of 25 sccm, O_2 of 2 sccm, 1800-W ICP power, 30 mtorr, and 20°C stage temperature for 10 min to form hemispherical holes. The holes were deepened using the Bosch process (*45*) consisting of two alternating steps. The first process was to deposit a passivation film on the hole, under the conditions of SF_6 of 5 sccm, C_4F_8 of 50 sccm, 650-W ICP power, 5-W RF power, 7-W dc power, 40 mtorr, and 20°C stage temperature. The next process was to etch out the hole by applying high dc power, under the conditions of SF_6 of 60 sccm, C_4F_8 of 25 sccm, 650-W ICP power, 20-W RF power, 170-W dc power, 30 mtorr, and 20°C stage temperature. We repeated these two processes 100 times to anisotropically deepen the hole. After the Bosch process, we cleaned the silicon mold with oxygen plasma treatment and removed the photoresist layer and residues formed on the substrate. Then, the silicon mold was coated with a hydrophobic layer using ICP-RIE under the conditions of O_2 of 10 sccm, C_4F_8 of 50 sccm, 600-W ICP power, 5-W RF power, 10 mtorr, and 20°C stage temperature for 1 min (*46*). After the mold preparation, the polymethyl methacrylate was prepared with a mixture of toluene:polymethyl methacrylate = 8:1 (weight %) and spin-coated on the fabricated silicon mold with 200 rounds per minute. The silicon mold was left in the vacuum oven for 10 min to remove bubbles and annealed under 80°C for 20 min. After the annealing process, the microlens array was demolded from the silicon mold.

Generating training and test dataset

To generate the training and test dataset, we used UNITY software to emulate the experimental conditions. The virtual space was crafted on the basis of a Euclidean coordinate system, converting real-world measurements such that 1 mm equated to 1 U in the virtual space (fig. S11A). Within this virtual space, we positioned two compound eye sensors, each composed of a 256-pixel array in a 16-by-16 configuration. Each pixel of the compound eye was represented by an individual camera providing an 8.5° FoV within a 40-by-40 square dimension. To replicate the hemispherical structure of the actual compound eye sensor, the cameras were repositioned within a 20-by-20 dimension (equivalent to 20 mm by 20 mm in real space; fig. S11B). The centers of two sensors were positioned at the coordinates (−15, 0, 0) and (15, 0, 0), respectively.

To emulate the object tracking situation, we used a moving sphere with a 6-U radius as the detection target in a virtual environment. This object was coded to move randomly within set boundaries: x coordinates between −40 and 40, y coordinates between −40 and 40, and $z$coordinates between 25 and 120. Each camera reported on the presence of the moving object. If the moving sphere was within a camera's FoV, the camera transmitted a "1" signal; otherwise, it sent "0." In two consecutive time periods (before and after the movement of the object), all 512 cameras in the field transmitted signals to the system twice. To simulate the readout procedure, a signal dimension of 512 by 2 was dispatched to the system. In addition, the object's position in Euclidean coordinates, both before and after its movement (comprising x, y, and z coordinates, 3-by-2 matrix), was also relayed to the system. As a result, a dataset with a length of 120,000 comprising 1030 data points was acquired. The dataset examples are depicted in fig. S11 (C and D).

Prediction on edge system

Most microprocessors, especially those used in embedded systems, are limited in memory and computational power. Consequently, this necessitates the use of compact neural network models. Initially, we designed and tested a streamlined yet highly functional neural network intended for use in portable microcontroller units. The architecture for near-sensor object tracking consists of three 1D CNN layers, followed by an FCL. Each CNN layer uses a 3-by-1 kernel with same padding, activated by the ReLU function, and includes a 3-to-1 average pooling layer to condense features. The left and right feature maps from each layer are then input into the subsequent CNN layer for regression and further compression. These inputs are processed to compute feature maps and subsequently concatenated across the channel dimension to form a cost volume. The merged features are input into an FCL network with three hidden layers (128-32-16-6), which track object movement. The first two layers of this multilayer perceptron use ReLU activation, whereas the final layer uses a sigmoid function to scale the output between 0 and 1. RMSE loss is used to calculate the regression loss between the initial and final positions of the moving object. We evaluated our model using datasets comprising 80,000 training, 20,000 validation, and 20,000 testing samples. Data labels were normalized to a 0-to-1 range using a minimum-maximum scaler during preprocessing. The model was trained over 5000 epochs with a batch size of 128, using stochastic gradient descent with a momentum optimizer in the Keras framework. Model verification was conducted using the mean absolute error as the evaluation metric.

For deployment on microcontrollers, the model underwent modifications to ensure compatibility with microprocessor constraints. These adjustments included techniques such as compression, conversion, quantization, and pruning (*47*). Compression reduced the model's size to fit the microprocessor's limited memory, and the quantization and pruning simplified the model by converting the 32-bit floating-point model to an 8-bit signed integer format. The model was then compiled into a header (.h) file for interpretation by the microprocessor and subsequently uploaded. The uploaded model size was 16.54 kilobytes.

Supplementary Materials

This PDF file includes:
Supplementary Methods
Figs. S1 to S15
References (*48*–*54*)

Other Supplementary Material for this manuscript includes the following:
Movies S1 and S2
MDAR Reproducibility Checklist

REFERENCES AND NOTES

1. Marshall Cavendish Corporation, *Insects and Spiders of the World* (Marshall Cavendish, 2003).
2. E. Gaten, Optics and phylogeny: Is there an insight? The evolution of superposition eyes in the Decapoda (Crustacea). *Contrib. Zool.* **67**, 223–235 (1998).
3. J.-J. Kim, H. Liu, A. O. Ashtiani, H. Jiang, Biologically inspired artificial eyes and photonics. *Reports Prog. Phys.* **83**, 047101 (2020).
4. M. F. Land, D.-E. Nilsson, *Animal Eyes* (Oxford Univ. Press, ed. 2, 2012).
5. N. Franceschini, F. Ruffier, J. Serres, A bio-inspired flying robot sheds light on insect piloting abilities. *Curr. Biol.* **17**, 329–335 (2007).
6. C. L. Choong, M. B. Shim, B. S. Lee, S. Jeon, D. S. Ko, T. H. Kang, J. Bae, S. H. Lee, K. E. Byun, J. Im, Y. J. Jeong, C. E. Park, J. J. Park, U. I. Chung, Highly stretchable resistive pressure sensors using a conductive elastomeric composite on a micropyramid array. *Adv. Mater.* **26**, 3451–3458 (2014).
7. D. Floreano, M. V. Srinivasan, J. C. Zufferey, C. Ellington, *Flying Insects and Robots* (Springer, 2010).
8. M. V. Srinivasan, M. Poteser, K. Kral, Motion detection in insect orientation and navigation. *Vision Res.* **39**, 2749–2766 (1999).
9. V. Nityananda, G. Tarawneh, R. Rosner, J. Nicolas, S. Crichton, J. Read, Insect stereopsis demonstrated using a 3D insect cinema. *Sci. Rep.* **6**, 1–9 (2016).
10. C. Ruiz, J. Theobald, Insect vision: Judging distance with binocular motion disparities. *Curr. Biol.* **28**, R148–R150 (2018).
11. R. Rosner, J. von Hadeln, G. Tarawneh, J. C. A. Read, A neuronal correlate of insect stereopsis. *Nat. Commun.* **10**, 2845 (2019).
12. A. Saxena, S. H. Chung, A. Y. Ng, 3-D depth reconstruction from a single still image. *Int. J. Comput. Vis.* **76**, 53–69 (2008).
13. D. H. Warren, E. R. Strelow, *Electronic Spatial Sensing for the Blind: Contributions from Perception, Rehabilitation, and Computer Vision*, D. H. Warren, E. R. Strelow, Eds., vol. 99 of *NATO Science Series E:* (Springer Netherlands, 1985).
14. A. Burton, J. Radford, *Thinking in Perspective: Critical Essays in the Study of Thought Processes*, vol. 66 of *Psychology in Progress*, A. Burton, J. Radford, Eds. (Methuen, 1978).
15. K. Blachut, T. Kryjak, M. Gorgon, Hardware implementation of multi-scale Lucas-Kanade optical flow computation algorithm – A demo, in *Proceedings of the 2018 Conference on Design and Architectures for Signal and Image Processing (DASIP)* (IEEE, 2018), pp. 60–61.
16. C. H. Chang, C. N. Chou, E. Y. Chang, CLKN: Cascaded Lucas-Kanade networks for image alignment, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 3777–3785.
17. G. Zhang, H. Chanson, Application of local optical flow methods to high-velocity free-surface flows: Validation and application to stepped chutes. *Exp. Therm. Fluid Sci.* **90**, 186–199 (2018).
18. S. Yi, J. D. Kendall, R. S. Williams, S. Kumar, Activity-difference training of deep neural networks using memristor crossbars. *Nat. Electron.* **6**, 45–51 (2023).
19. Z. Lu, S. Rallapalli, K. Chan, S. Pu, T. La Porta, Augur: Modeling the resource requirements of convnets on mobile devices. *IEEE Trans. Mob. Comput.* **20**, 352–365 (2021).
20. D. Li, X. Chen, M. Becchi, Z. Zong, Evaluating the energy efficiency of deep convolutional neural networks on CPUs and GPUs, in *2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BdCloud-SocialCom-SustainCom)* (IEEE, 2016), pp. 477–484.
21. H. Chiroma, U. A. Abdullahi, S. M. Abdulhamid, A. Abdulsalam Alarood, L. A. Gabralla, N. Rana, L. Shuib, I. A. Targio Hashem, D. E. Gbenga, A. I. Abubakar, A. M. Zeki, T. Herawan, Progress on artificial neural networks for big data analytics: A survey. *IEEE Access* **7**, 70535–70551 (2019).
22. I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, S. U. Khan, The rise of "big data" on cloud computing: Review and open research issues. *Inf. Syst.* **47**, 98–115 (2015).
23. J. Chen, Z. Zhou, B. J. Kim, Y. Zhou, Z. Wang, T. Wan, J. Yan, J. Kang, J.-H. Ahn, Y. Chai, Optoelectronic graded neurons for bioinspired in-sensor motion perception. *Nat. Nanotechnol.* **18**, 882–888 (2023).
24. B. Bae, M. Park, D. Lee, I. Sim, K. Lee, Hetero-integrated ingaas photodiode and oxide memristor-based artificial optical nerve for in-sensor nir image processing. *Adv. Opt. Mater.* **11**, 202201905 (2022).
25. D. Lee, M. Park, Y. Baek, B. Bae, J. Heo, K. Lee, In-sensor image memorization and encoding via optical neurons for bio-stimulus domain reduction toward visual cognitive processing. *Nat. Commun.* **13**, (2022).
26. M. Park, Y. Yuan, Y. Baek, A. H. Jones, N. Lin, D. Lee, H. S. Lee, S. Kim, J. C. Campbell, K. Lee, Neuron-inspired time-of-flight sensing via spike-timing-dependent plasticity of artificial synapses. *Adv. Intell. Syst.* **4**, 2100159 (2021).
27. C. Thapa, M. A. P. Chamikara, S. Camtepe, SplitFed: When federated learning meets split learning. *Proc. AAAI Conf. Artif. Intell.* **36**, 8485–8493 (2020).
28. K. Yang, S. Chen, C. Shen, On the convergence of hybrid server-clients collaborative training. *IEEE J. Sel. Areas Commun.* **41**, 802–819 (2023).
29. X. Wei, C. Shen, Federated learning over noisy channels: Convergence analysis and design examples. *IEEE Trans. Cogn. Commun. Netw.* **8**, 1253–1268 (2022).
30. Y. Li, Y. Guo, M. Alazab, S. Chen, C. Shen, K. Yu, Joint optimal quantization and aggregation of federated learning scheme in VANETs. *IEEE Trans. Intell. Transp. Syst.* **23**, 19852–19863 (2022).
31. D. Eigen, C. Puhrsch, R. Fergus, Depth map prediction from a single image using a multi-scale deep network. *Adv. Neural Inf. Process. Syst.* **3**, 2366–2374 (2014).
32. R. H. Masland, The neuronal organization of the retina. *Neuron* **76**, 266–280 (2012).
33. D. Fan, K. Lee, S. R. Forrest, Flexible thin-film InGaAs photodiode focal plane array. *ACS Photonics* **3**, 670–676 (2016).
34. S. Xu, Y. Zhang, J. Cho, J. Lee, X. Huang, L. Jia, J. A. Fan, Y. Su, J. Su, H. Zhang, H. Cheng, B. Lu, C. Yu, C. Chuang, T. Il Kim, T. Song, K. Shigeta, S. Kang, C. Dagdeviren, I. Petrov, P. V. Braun, Y. Huang, U. Paik, J. A. Rogers, Stretchable batteries with self-similar serpentine interconnects and integrated wireless recharging systems. *Nat. Commun.* **4**, 1543–1548 (2013).
35. J. A. Rogers, T. Someya, Y. Huang, Materials and mechanics for stretchable electronics. *Science* **327**, 1603–1607 (2010).
36. D.-E. Nilsson, Vision optics and evolution. *Bioscience* **39**, 298–307 (1989).
37. M. V. Srinivasan, G. D. Bernard, The effect of motion on visual acuity of the compound eye: A theoretical analysis. *Vision Res.* **15**, 515–525 (1975).
38. D. Floreano, R. Pericet-Camara, S. Viollet, F. Ruffier, A. Brückner, R. Leitel, W. Buss, M. Menouni, F. Expert, R. Juston, M. K. Dobrzynski, G. L'Eplattenier, F. Recktenwald, H. A. Mallot, N. Franceschini, Miniature curved artificial compound eyes. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9267–9272 (2013).
39. Y. Zhang, G. Q. Mao, X. Zhao, Y. Li, M. Zhang, Z. Wu, W. Wu, H. Sun, Y. Guo, L. Wang, X. Zhang, Q. Liu, H. Lv, K. H. Xue, G. Xu, X. Miao, S. Long, M. Liu, Evolution of the conductive filament system in HfO_2-based memristors observed by direct atomic-scale imaging. *Nat. Commun.* **12**, 1–10 (2021).
40. K. Lee, J. D. Zimmerman, T. W. Hughes, S. R. Forrest, Non-destructive wafer recycling for low-cost thin-film flexible optoelectronics. *Adv. Funct. Mater.* **24**, 4284–4291 (2014).
41. G. S. Ferguson, M. K. Chaudhury, G. B. Sigal, G. M. Whitesides, Contact adhesion of thin gold films on elastomeric supports: Cold welding under ambient conditions. *Science* **253**, 776–778 (1991).
42. Y. S. Choi, Y. Y. Hsueh, J. Koo, Q. Yang, R. Avila, B. Hu, Z. Xie, G. Lee, Z. Ning, C. Liu, Y. Xu, Y. J. Lee, W. Zhao, J. Fang, Y. Deng, S. M. Lee, A. Vázquez-Guardado, I. Stepien, Y. Yan, J. W. Song, C. Haney, Y. S. Oh, W. Liu, H. J. Yun, A. Banks, M. R. MacEwan, G. A. Ameer, W. Z. Ray, Y. Huang, T. Xie, C. K. Franz, S. Li, J. A. Rogers, Stretchable, dynamic covalent polymers for soft, long-lived bioresorbable electronic stimulators designed to facilitate neuromuscular regeneration. *Nat. Commun.* **11**, 1–14 (2020).
43. P. Panduranga, A. Abdou, Z. Ren, R. H. Pedersen, M. P. Nezhad, Isotropic silicon etch characteristics in a purely inductively coupled SF6 plasma. *J. Vac. Sci. Technol. B* **37**, 061206 (2019).
44. J. Albero, L. Nieradko, C. Gorecki, H. Ottevaere, V. Gomez, H. Thienpont, J. Pietarinen, B. Päivänranta, N. Passilly, Fabrication of spherical microlenses by a combination of isotropic wet etching of silicon and molding techniques. *Opt. Express* **17**, 6283–6292 (2009).
45. A. M. Baracu, C. A. Dirdal, A. M. Avram, A. Dinescu, R. Muller, G. U. Jensen, P. C. V. Thrane, H. Angelskår, Metasurface fabrication by cryogenic and bosch deep reactive ion etching. *Micromachines* **12**, 501 (2021).

46. J. Lee, J. Lee, H. W. Lee, K.-H. Kwon, Anti-adhesive characteristics of CHF3/O2 and C4F8/O2 plasma-modified silicon molds for nanoimprint lithography. *Mater. Res. Bull.* **69**, 120–125 (2015).
47. P. Novac, G. B. Hacene, A. Pegatoquet, B. Miramond, V. Gripon, Quantization and deployment of deep neural networks on microcontrollers. *Sensors* **21**, 2984 (2021).
48. B. Matas, C. DeSubercasau, *Memory 1997: Complete Coverage of DRAM, SRAM, EPROM, and Flash Memory ICs* (Integrated Circuit Engineering Corporation, 1997).
49. K. Kagawa, S. Shishido, M. Nunoshita, J. Ohta, A 3.6pW/frame-pixel 1.35V PWM CMOS imager with dynamic pixel readout and no static bias current, in *2008 IEEE International Solid-State Circuits Conference-Digest of Technical Papers* (IEEE, 2008), pp. 54–56.
50. M. Y. Ng, 0.18um low voltage 12-bit successive-approximation-register analog-to-digital converter (SAR ADC), in *2011 3rd Asia Symposium on Quality Electronic Design (ASQED)* (IEEE, 2011), pp. 277–281.
51. Z. Liu, T. Park, H. S. Park, N. S. Kim, Ultra-low-power image signal processor for smart camera applications. *Electron. Lett.* **51**, 1778–1780 (2015).
52. Micron Technology Inc., Technical note: DDR3 point-to-point design support (2013).
53. A. Petreto, A. Hennequin, T. Koehler, T. Romera, Y. Fargeix, B. Gaillard, M. Bouyer, Q. L. Meunier, L. Lacassagne, Energy and execution time comparison of optical flow algorithms on SIMD and GPU architectures, in *2018 Conference on Design and Architectures for Signal and Image Processing (DASIP)* (IEEE, 2018), pp. 25–30.
54. G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 4700–4708.

Acknowledgments

Funding: This work was supported by National Science Foundation (NSF) grant no. 1942868 (B.B., D.L., M.P., Y.B., and K.L.) and U.S. Air Force Office of Scientific Research (AFOSR) Young Investigator Program (YIP) grant no. FA9550-23-1-0159 (B.B., Y.B., I.S., and K.L.). The work of Y.M. and C.S. is partially supported by the National Science Foundation under grants CNS-2002902, ECCS-2033671, and ECCS-2143559. **Author contributions:** Conceptualization: K.L., B.B., D.L., and M.P. Methodology: B.B., D.L., and Y.M. Investigation: K.L., C.S., B.B., D.L., and I.S. Visualization: B.B. and Y.B. Funding acquisition: K.L.. Project administration: K.L. and C.S. Supervision: K.L. Writing: B.B., M.P., and Y.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the Supplementary Materials. The code and data used for our quantitative analysis have been made available for download at https://doi.org/10.5061/dryad.6m905qg7k and https://doi.org/10.5281/zenodo.10980343.

Submitted 13 October 2023
Accepted 17 April 2024
Published 15 May 2024
10.1126/scirobotics.adl3606

Science Robotics

Stereoscopic artificial compound eyes for spatiotemporal perception in three-dimensional space

Byungjoon Bae, Doeon Lee, Minseong Park, Yujia Mu, Yongmin Baek, Inbo Sim, Cong Shen, and Kyusang Lee

Sci. Robot. **9** (90), eadl3606. DOI: 10.1126/scirobotics.adl3606

Editor's summary
The praying mantis has an overlap in its field of view between its left and right eyes that generates binocular vision and is capable of depth perception through stereopsis. Inspired by the praying mantis, Bae *et al.* have developed artificial compound eyes with stereoscopic vision for spatiotemporal perception in three-dimensional space using edge computing. They adopted federated split learning to process information from both the left and right artificial eyes and showed that the system was capable of tracking object movement at a rapid temporal processing rate. —Amos Matsiko

---

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.