


# Bioinspired in-sensor spectral adaptation for perceiving spectrally distinctive features

Received: 25 October 2023

Accepted: 18 June 2024

Published online: 15 July 2024

 Check for updates

Bangsen Ouyang<sup>1,2,3</sup>, Jialiang Wang<sup>1,2,3</sup>, Guang Zeng<sup>1,2</sup>, Jianmin Yan<sup>1,2</sup>, Yue Zhou<sup>1,2</sup>, Xixi Jiang<sup>1,2</sup>, Bangjie Shao<sup>1,2</sup> & Yang Chai<sup>1,2</sup>  

In challenging lighting conditions, machine vision often yields low-quality results. In situations where particular spectral signatures carry critical information, adapting the spectral sensitivity of vision systems to match the predominant spectra of the surrounding environment can improve light capture and image quality. Here we report spectra-adapted vision sensors based on arrays of back-to-back photodiodes. The spectral sensitivity of these bioinspired sensors can be tuned to match either the broadband visible spectrum or a narrow band within the near-infrared spectrum by applying different bias voltages. The process of spectral adaptation takes tens of microseconds, which is comparable with the frame rate (around 100 kHz) of state-of-the-art high-speed cameras. The spectral adaptation increases the Weber contrast of the scene by over ten times, resulting in increased recognition accuracy (from 33% to 90%) of features when exposed to intense visible-light glare.

Natural light spectra span a wide range of wavelengths. The visual information carried by different spectra dynamically varies with the environmental factors, including the spectral distribution of lighting, the spectral transmittance of media and the spectral reflectance of objects<sup>1,2</sup>. However, state-of-the-art complementary metal–oxide–semiconductor (CMOS) and charge-coupled device (CCD) image sensors have fixed spectral responses, and lack adaptability to other spectral regions for imaging under challenging lighting conditions, such as colour-cast lighting, background-glare interference and strong attenuation effects (Supplementary Fig. 1)<sup>3–5</sup>. This mismatch between the spectral response of image sensors and the spectra of the surrounding environment leads to low imaging quality in conventional CMOS-/CCD-based machine vision systems, leading to the ineffective extraction of visual features<sup>6–8</sup>.

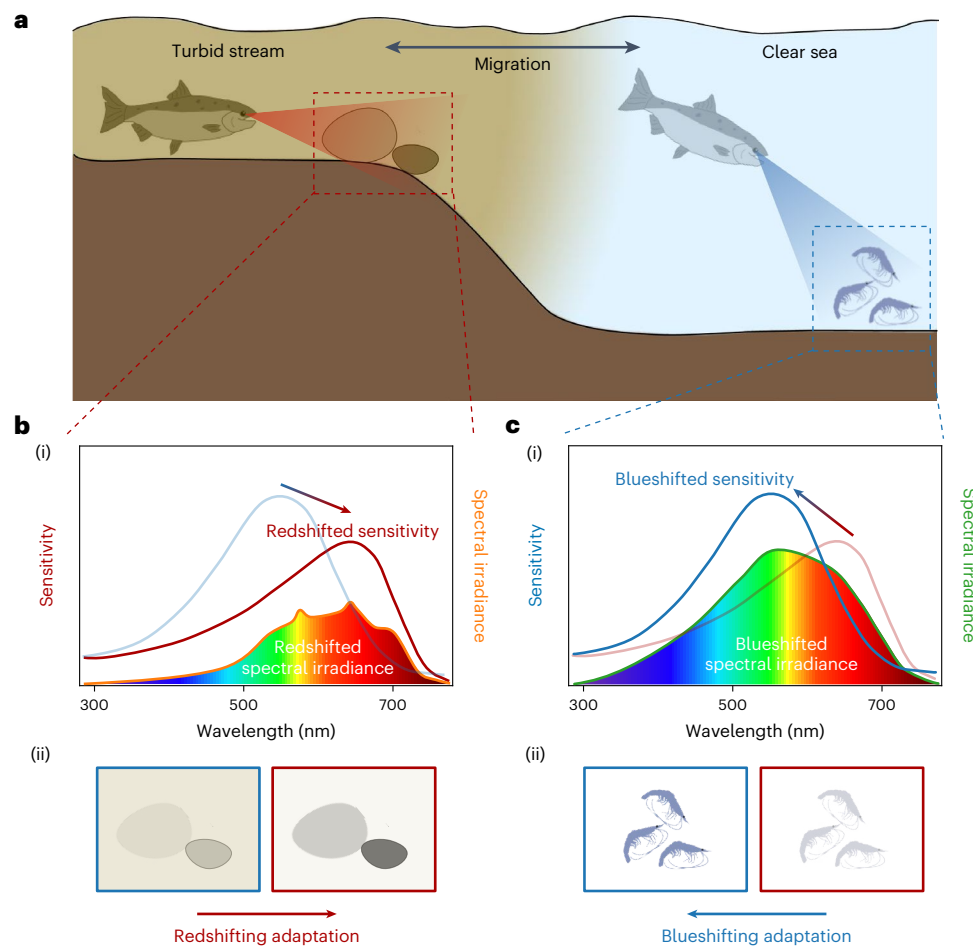
Post-imaging algorithms (such as colour correction, noise cancellation and interference suppression)<sup>9–11</sup> can be used for visual information restoration from low-quality images<sup>12</sup>, but they require additional power<sup>13–16</sup>. Recently reported approaches for light intensity adaptation offer a solution for capturing accurate images under high-dynamic-range light illumination<sup>17</sup>. By adapting the vision sensors to the dynamic environment spectra, machine vision can accurately

perceive the visual information associated with specific spectral signatures, and this could be used in various applications including autonomous vehicles<sup>18</sup>, medical diagnoses<sup>19</sup>, industrial manufacturing<sup>20</sup> and surveillance systems<sup>21</sup>.

Organisms in the natural world have evolved adaptive vision mechanisms that go beyond light intensity adaptation, enabling a more precise perception of visual information. Certain amphibious and migratory species, like Pacific salmon (*Oncorhynchus keta*), have developed spectrally adaptive vision to inhabit environments with variable spectral conditions<sup>22</sup>. This spectral adaptation function originates from two kinds of photosensitive visual pigment—vitamin-A<sub>1</sub>-based visual pigments (VP<sub>1</sub>) and vitamin-A<sub>2</sub>-based visual pigments (VP<sub>2</sub>)—with tunable ratios (Supplementary Fig. 2)<sup>23</sup>. VP<sub>1</sub> and VP<sub>2</sub> are more sensitive to shorter and longer wavelengths, respectively. By converting vitamin A<sub>1</sub> into vitamin A<sub>2</sub> through an enzyme mechanism (Supplementary Note I), the retina can dynamically shift its spectral sensitivity to match the spectra in the surrounding environment.

In this Article, we report a spectra-adapted vision sensor based on an array of back-to-back photodiodes. Inspired by the vision systems of Pacific salmon, we design switchable junctions—between surface shallow and bottom deep junctions by changing the bias

<sup>1</sup>Department of Applied Physics, The Hong Kong Polytechnic University, Hong Kong, China. <sup>2</sup>Joint Research Center of Microelectronics, The Hong Kong Polytechnic University, Hong Kong, China. <sup>3</sup>These authors contributed equally: Bangsen Ouyang, Jialiang Wang. ✉e-mail: [ychai@polyu.edu.hk](mailto:ychai@polyu.edu.hk)



**Fig. 1 | Spectral adaptation behaviour of Pacific salmon.** **a**, As a kind of migratory fish, Pacific salmon inhabit spectrally variable environments, from turbid inland streams to clear open seas, in different stages of their life cycles. **b**, In turbid streams, Pacific salmon adapt to the spectral sensitivity (red curve) to match the redshifted spectral irradiance (orange curve) (i). The redshifting

adaptation increases the clarity of features (ii). **c**, In clear seas, the reverse blueshifting adaptation (i) enables Pacific salmon to also use short-wavelength light for improving the clarity of the features (ii) (green curve, blueshifted spectral irradiance; blue curve, blueshifted spectral sensitivity).

voltages—with different spectral sensitivities that can be adjusted to match either the broadband visible spectrum or narrowband near-infrared spectrum. This approach allows in-sensor spectral adaptation with an operational frequency of 100 kHz. We experimentally demonstrate the bioinspired in-sensor spectral adaptation with an  $8 \times 8$  vision sensor array, where the Weber contrast exhibits a ten times enhancement through the spectral adaptation process. Our spectra-adapted vision sensors can collect information in broad spectra and resist the interference of visible light, which helps machine vision systems improve recognition accuracy in high-glare, smoky and foggy environments.

### Spectral adaptation of biological visual systems

Pacific salmon possess spectrally adaptive vision systems that allow them to inhabit spectrally variable environments, from inland streams to open seas (Fig. 1a)<sup>24</sup>. In turbid inland streams, the spectral distribution of the light under water (Fig. 1b(i), orange curve) shifts to long wavelengths due to the preferential absorption and scattering at short wavelengths<sup>1,25</sup>. Pacific salmon can tune the spectral sensitivity of their visual systems (Fig. 1b(ii), red curve) to match the redshifted aquatic environment. In clear open seas, the spectral distribution of light (Fig. 1c(i), green curve) has higher intensity and relatively blueshifted spectral distribution because of the absence of absorption and scattering effects. Visual systems of the Pacific salmon reversely adapt the

spectral sensitivity (Fig. 1c(i), green curve) to match the blueshifted spectral distribution in the environment.

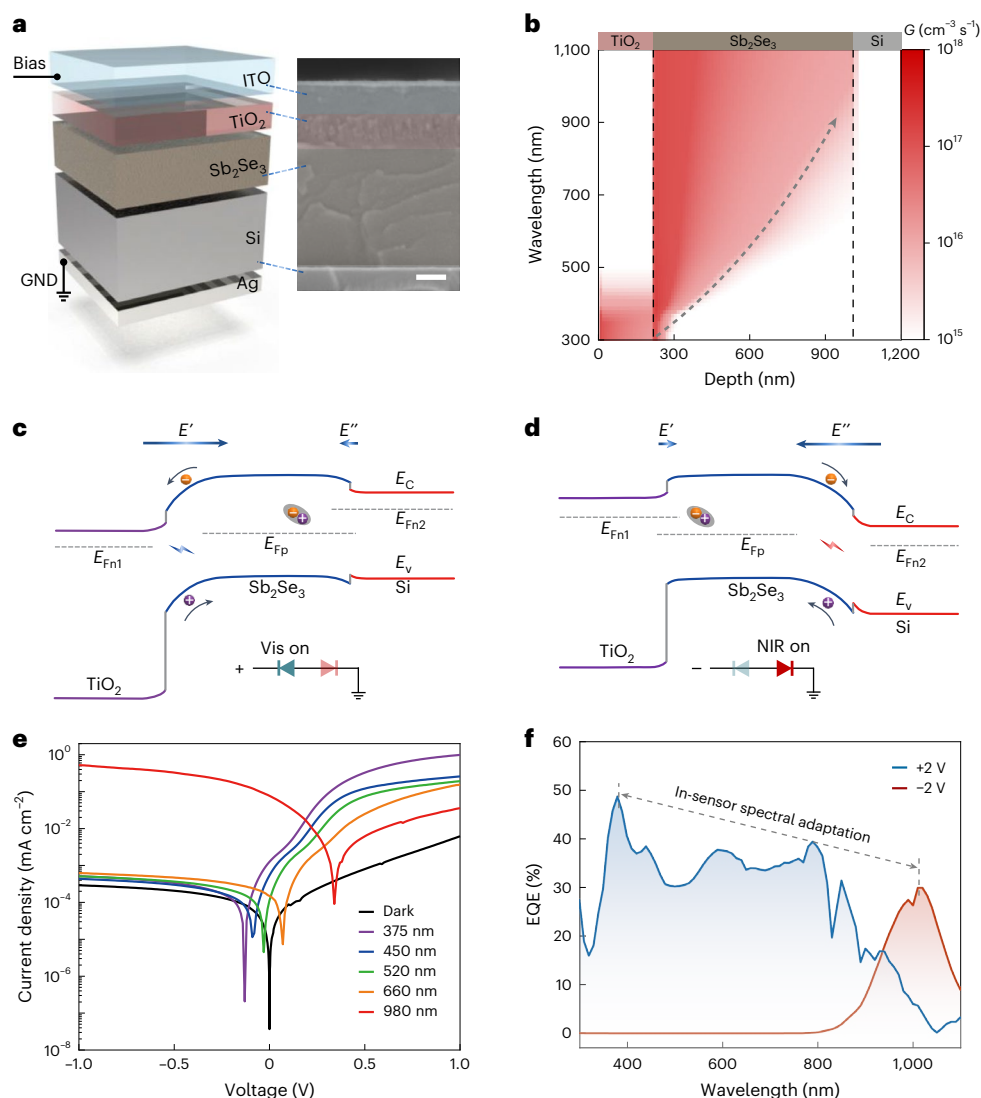
The alignments between the spectral sensitivity and predominant wavelengths in the environment enable the visual system to capture more light and perceive more intense visual features, which can be described according to equations (1) and (2):

$$L_{v,f} = \int_{\lambda} K_m V(\lambda) L_{e,f}(\lambda) d\lambda, \quad (1)$$

$$L_{v,b} = \int_{\lambda} K_m V(\lambda) L_{e,b}(\lambda) d\lambda, \quad (2)$$

where  $K_m$  is the peak luminous efficacy and  $V(\lambda)$  (luminous efficiency function) is the spectral sensitivity of the vision system.  $L_{e,f}(\lambda)$  and  $L_{e,b}(\lambda)$  are the spectral radiance of the targeting features and non-feature background in the environment, and  $L_{v,f}$  and  $L_{v,b}$  are the corresponding luminance perceived by the vision system, respectively. Weber contrast ( $C_W$ ) is defined according to equation (3)<sup>26</sup>:

$$C_W = \frac{L_{v,f} - L_{v,b}}{L_{v,b}}. \quad (3)$$



**Fig. 2 | Bias-switchable in-sensor spectral adaptation.** **a**, Cross-sectional schematic and SEM image of the device structure. Scale bar, 200 nm. **b**, Depth profile of charge generation rate across the device under light illuminations with different wavelengths. **c,d**, Energy band diagrams of the device when positive (**c**) and negative (**d**) voltages are applied to the ITO electrode. **e**,  $I$ - $V$  curves of the back-to-back photodiode under the dark condition and light illumination at

different wavelengths. Under UV–visible (near-infrared) light illumination, the device shows a higher photocurrent-to-dark-current ratio at a positive (negative) bias voltage. **f**, Spectral response of the back-to-back photodiode under different bias voltages. The spectral response fully covers the visible (Vis) spectrum at a bias voltage of 2 V and narrowly distributes in the NIR spectrum at a bias voltage of  $\sim 2$  V.

Consequently, the redshifting and blueshifting adaptation can increase the  $C_w$  value of the perceived features, and allow Pacific salmon to clearly see the surroundings in spectrally distinctive environments (Fig. 1b(ii),c(ii)).

### In-sensor spectral adaptation of back-to-back photodiodes

Conventional semiconductor vision sensors usually have fixed spectral responses because of their sole sensing semiconductor. To emulate the spectral adaptation behaviour of Pacific salmon, we develop vision sensors with distinctive spectral responses at different bias voltages. The designed device has a stacked structure of indium tin oxide (ITO)/n-type titanium dioxide (n-TiO<sub>2</sub>)/p-type antimony selenide (p-Sb<sub>2</sub>Se<sub>3</sub>)/n-type silicon (n-Si)/silver (Ag) (Fig. 2a). The detailed fabrication process of the device can be found in Methods and Supplementary Fig. 3. The thickness of ITO, TiO<sub>2</sub>, Sb<sub>2</sub>Se<sub>3</sub> and Si are 0.22, 0.22, 0.80 and 400.00  $\mu\text{m}$ , respectively (Fig. 2a and Supplementary Fig. 4a). The energy-dispersive X-ray spectroscopy and mapping data further

confirm the composition of the device structure (Supplementary Fig. 4b). The Raman and X-ray diffraction spectra (Supplementary Fig. 4d–f) indicate the successful deposition of pure-phase TiO<sub>2</sub> and Sb<sub>2</sub>Se<sub>3</sub>, good crystallinity of Sb<sub>2</sub>Se<sub>3</sub> and amorphous structure of TiO<sub>2</sub>.

Light with a shorter (longer) wavelength usually has a larger (smaller) absorption coefficient ( $\alpha$ ) and decays exponentially in the light-absorbing material (Supplementary Fig. 5b). Therefore, the electron–hole pairs generation rate ( $G$ ) has the characteristic of  $G \propto \alpha P e^{-\alpha x}/hc$ , where  $P$ ,  $\lambda$  and  $c$  denote the power density, wavelength and speed of light, respectively;  $x$  represents the propagation depth in the light-absorbing material; and  $h$  is the Planck constant. This characteristic contributes to the wavelength-dependent electron–hole pair generation depth across the vertical profile of the light-absorbing material. Under light illumination with shorter (longer) wavelength, electron–hole pairs are generated in a shallower (deeper) region within the light-absorbing material (Fig. 2b). Additionally, the bandgaps of the functional layers decrease sequentially ( $E_{g,\text{TiO}_2} > E_{g,\text{Sb}_2\text{Se}_3} > E_{g,\text{Si}}$ ), further contributing to the wavelength-dependent electron–hole pair

generation depth across the vertical profile of the whole device. When the wavelength of incident light is shorter than 376 nm (photon energy,  $h\nu \geq 3.3$  eV), electron–hole pairs are mainly generated in the top junction with 220-nm-thick  $\text{TiO}_2$ , because the photon energy is higher than the bandgap of  $\text{TiO}_2$  ( $E_g = 3.3$  eV; Supplementary Fig. 5c). The light illumination with a wavelength between 376 and 954 nm ( $3.3 \geq h\nu \geq 1.3$  eV) can excite electron–hole pairs in the 800-nm-thick  $\text{Sb}_2\text{Se}_3$  layer, because the photon energy is larger than the bandgap of  $\text{Sb}_2\text{Se}_3$  ( $E_g = 1.3$  eV; Supplementary Fig. 5d). When the wavelength of incident light is between 954 and 1,107 nm ( $1.30 \geq h\nu \geq 1.12$  eV), electron–hole pairs can be generated in silicon, because the photon energy is larger than the bandgap of Si ( $E_g = 1.12$  eV) (ref. 27), but smaller than the bandgaps of  $\text{TiO}_2$  and  $\text{Sb}_2\text{Se}_3$ . The wavelength-dependent electron–hole pair generation depth across the vertical profile of the device facilitates the subsequent selective separation of short- and long-wavelength excited electron–hole pairs.

The working mechanism of the bias-switchable spectral responses can be further elucidated through the energy band diagram. Ultraviolet photoelectron spectroscopy characterization allows to identify the location of the Fermi level ( $E_f$ ) and valence band maximum ( $E_v$ ) of the functional layers (Supplementary Fig. 6a–c). Supplementary Fig. 6d,e further shows the energy band alignment before and after contact, indicating the formation of back-to-back photodiodes with an n–p–n structure. When a positive voltage is applied to the ITO electrode, the shallow  $\text{TiO}_2/\text{Sb}_2\text{Se}_3$  junction is under the reverse-bias condition and the deep  $\text{Sb}_2\text{Se}_3/\text{Si}$  junction is forward biased (Fig. 2c and Supplementary Fig. 7a). Thus, the separation efficiency of photogenerated electron–hole pairs within the shallow (deep) junction improves (decreases) due to the enhanced (weakened) electric field  $E'$  ( $E''$ ) and broadened (narrowed) space-charge region<sup>28</sup>.  $E'$  and  $E''$  are the electric fields within the depletion region of the shallow and deep junctions, respectively. Under short-wavelength light illumination, the electron–hole pairs are mainly generated in the shallow region of the device. Therefore, electron–hole pairs can be effectively separated by the shallow  $\text{TiO}_2/\text{Sb}_2\text{Se}_3$  junction, and electrons (holes) drift towards the  $\text{TiO}_2$  ( $\text{Sb}_2\text{Se}_3$ ) side to contribute to the photocurrent. In the case of long-wavelength light illuminations, electron–hole pairs generated in the deep region of the device recombine and hardly contribute to the photocurrent, because of the weakened separation capability of the deep junction and the long diffusion distance to the shallow junction. Thus, the device is more sensitive to short-wavelength light when a positive voltage is applied to the ITO electrode (visible-spectrum adaptation).

When a negative voltage is applied to the ITO electrode, the shallow  $\text{TiO}_2/\text{Sb}_2\text{Se}_3$  junction is under the forward-bias condition and the deep  $\text{Sb}_2\text{Se}_3/\text{Si}$  junction is reversely biased (Fig. 2d and Supplementary Fig. 7b). Thus, the separation efficiency of electron–hole pairs within the shallow (deep) junction decreases (improves) due to the weakened (enhanced) electric field  $E'$  ( $E''$ ) and narrowed (broadened) space-charge region<sup>28</sup>. Under light illuminations with short wavelengths, the electron–hole pairs are mainly generated in the shallow region of the devices. The electron–hole pairs recombine and hardly contribute to the photocurrent due to the weakened separation capability of the shallow junction and the long diffusion distance to the deep junction. Thus, the  $\text{TiO}_2$  and  $\text{Sb}_2\text{Se}_3$  regions also work as light filters, which absorb most of the short-wavelength light<sup>29,30</sup>. In the case of long-wavelength light illumination, electron–hole pairs can be generated in the  $\text{Sb}_2\text{Se}_3/\text{Si}$  interface region of the device and effectively separated by the deep  $\text{Sb}_2\text{Se}_3/\text{Si}$  junction, and electrons (holes) drift towards the Si ( $\text{Sb}_2\text{Se}_3$ ) side to contribute to the photocurrent. Because much of the short-wavelength light is absorbed and the cut-off wavelength of Si is 1,107 nm, the device possesses a narrow spectral response range in the near-infrared (NIR) region (NIR-spectrum adaptation).

The control devices ( $\text{TiO}_2/\text{Sb}_2\text{Se}_3$  and  $\text{Sb}_2\text{Se}_3/\text{n-Si}$ ) exhibit expected rectification characteristics (Supplementary Fig. 8a),

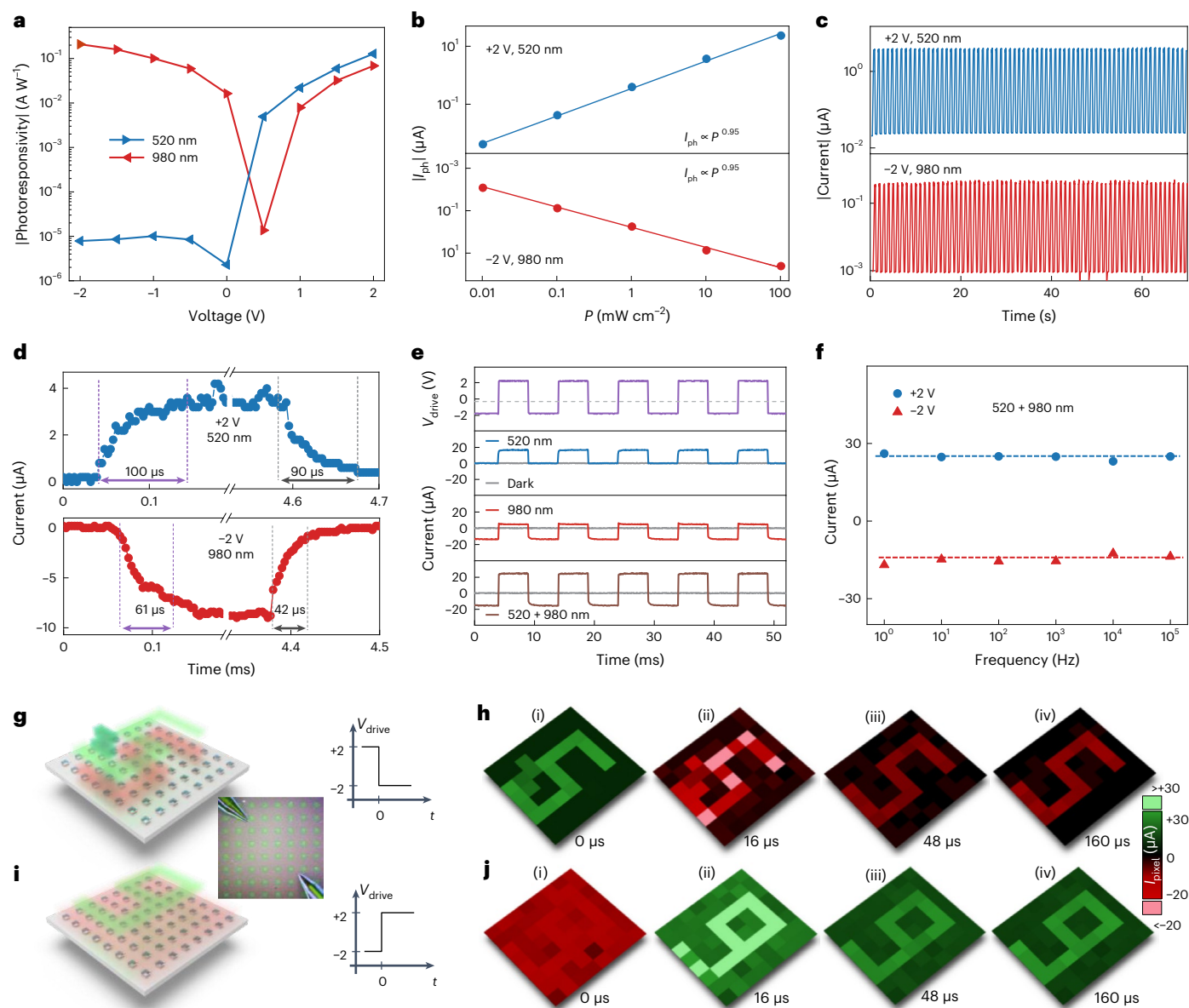
which indicate the formation of the designed heterojunctions. The  $I$ – $V$  curves of the back-to-back photodiodes (Fig. 2e) display a much balanced current under different voltage polarities compared with the single-junction devices. This behaviour indicates the successful formation of back-to-back heterojunctions. The device exhibits expected wavelength-dependent behaviours under light illuminations. Under short-wavelength illuminations (375, 450, 520 and 660 nm) at positive voltages, the device shows prominent photocurrents with sequential photocurrent-to-dark current ratios of 160, 42, 31 and 25. At negative voltages, the ratios are distinctly lower (1.8, 1.5, 1.7 and 2.1). Under 980 nm light, the device shows a prominent photocurrent at a negative voltage with a photocurrent-to-dark-current ratio of  $1.8 \times 10^3$ . However, under a positive bias, the photocurrent is considerably smaller, with a low ratio of 6. These results closely align with the designed bias-switchable spectral responses, where positive voltages correspond to a short-wavelength response and negative voltages correspond to a long-wavelength response. The extracted open-circuit voltages and short-circuit currents of the device exhibit opposite polarity under short- and long-wavelength illuminations (Supplementary Fig. 8c). These behaviours suggest the presence of two opposite built-in electric fields within the device, which is consistent with the proposed working mechanism.

Furthermore, the back-to-back photodiodes show minimal hysteresis during cycled dual-sweep  $I$ – $V$  scans (Supplementary Fig. 8d). This characteristic enables the rapid response of the device when the bias voltage transitions between positive and negative polarities. Supplementary Fig. 8e shows the temperature-dependent  $I$ – $V$  curves of the device. The consistent photocurrents within a temperature range of 25–95 °C enable the device to reliably operate at high temperatures.

Figure 2f shows the spectral external quantum efficiency (EQE) of the device as a function of bias voltage. The spectral response of the device adapts from short-wavelength light to long-wavelength light, when the bias voltage changes from +2 to –2 V. The shift in the spectral response can emulate the adaptation process when Pacific salmon migrate between clear seas and turbid streams. The role of the bias voltage applied to the device is similar to the P450 enzymes in the biological vision system. Under a positive bias voltage, the full-width at half-maximum of the EQE–wavelength curve is 540 nm (between 340 and 880 nm), which fully covers the visible spectrum (380 to 750 nm)<sup>31</sup>. Under negative voltages, the device exhibits a narrowband response in the NIR region (750–1,500 nm). The single-peak EQE–wavelength curve has a full-width at half-maximum of 145 nm (between 930 and 1,075 nm) within the NIR region (750–1,500 nm). The sub-200-nm narrowband response arises from the ‘self-filtering’ structure implemented in the back-to-back photodiode configuration<sup>32</sup>. The top junction absorbs short-wavelength components in the incident light, effectively suppressing the EQE of the bottom junction at these wavelengths. The narrowband NIR response enables the device to selectively capture and distinguish information in the NIR band without being interfered by information in the visible band.

The bias-switchable spectral responses allow the device to selectively operate either in the broadband visible spectrum or in the narrowband near-infrared spectrum without requiring extra footprint and wiring for multiple subpixels or relying on additional optical filters<sup>33,34</sup>. This spectral adaptation function enables the vision sensor to resolve spectrally distinctive features in the visible or NIR band within the sensor. The device shows considerable stability over time, retaining 80% of its EQE after being stored in the ambient condition for 1 year (Supplementary Fig. 9a). The EQE curves of the device also exhibit sub-spectrum differences at different voltages within the same polarity (Supplementary Fig. 9b–d and Supplementary Note II). This behaviour further enables the sensor to work for different sub-spectrum bands in a more versatile manner<sup>33,35</sup>.





**Fig. 3 | Dynamic in-sensor spectral adaptation characteristics of the bias-switchable vision sensor array.** **a**, Photoresponsivity as a function of bias voltages under light illuminations with different wavelengths (520 and 980 nm). **b**, Photocurrent as a function of light intensity, showing the linear dependence of photocurrent on light intensity. **c**,  $I$ - $t$  curves of the device when subjected to cyclic illuminations of 520 nm light (with bias voltages of +2 V) and 980 nm light (with bias voltages of -2 V). **d**, Response time and recovery time to 520 and

980 nm light stimuli. **e**,  $I$ - $t$  curves under constant 520 nm (top), constant 980 nm (middle) or 520 + 980 nm (bottom) light conditions when a periodical bias voltage is applied to the device. **f**, Output current as a function of the frequency of bias voltage. **g**, **h**, Illustration of an  $8 \times 8$  spectra-adapted vision sensor array in the NIR-spectrum adaptation test (**g**) and the time course of the imaging result (**h**). **i**, **j**, Illustration of an  $8 \times 8$  spectra-adapted vision sensor array in the visible-spectrum adaptation test (**i**) and the time course of the imaging result (**j**).

## Dynamic in-sensor adaptation of the vision sensor array

Machine vision systems on autonomous vehicles or traffic surveillance systems need to identify the target over a very short temporal scale<sup>36</sup>. In such high-stakes scenarios that demand real-time responses to rapidly changing environmental conditions, even a minuscule delay or misjudgement can lead to serious consequences, which urgently demands high-speed and highly accurate processing of data-intensive vision information at sensory terminals. Thus, it requires the vision sensors to rapidly adapt between the visible and NIR spectra.

To further investigate the dynamic adaptation characteristics, we adopt illumination with the wavelengths of 520 and 980 nm to represent the visible and NIR light environments, respectively. Figure 3a

(extracted from Supplementary Fig. 9c) shows the photoresponsivity ( $R$ ) of the device under 520 and 980 nm light stimuli as a function of the bias voltage. The photoresponsivity increases with the absolute value of bias voltage. The photoresponsivity under 2 and -2 V of the device reaches  $-0.2 \text{ A W}^{-1}$ , which is comparable with that in previously reported devices<sup>29,37</sup>. Suppressing the noise signal carried by visible light under strongly scattered and high-glare conditions can provide better image quality in challenging lighting conditions. Thus, a high spectral rejection ratio ( $R_{980}/R_{520}$ ) is highly desired. The  $R_{980}/R_{520}$  value reaches  $-10^4$  at -2 V, which proves the device works in the NIR region with negligible interference from visible light.

Figure 3b shows the  $I_{\text{ph}}-P$  curves under 520 nm light (with bias voltages of +2 V) and 980 nm light (with bias voltages of -2 V), where

$I_{ph} = I_{light} - I_{dark}$  represents the photocurrent under a specific power density ( $P$ ) of light stimuli. These curves were fitted according to the formula  $I_{ph} \propto P^\alpha$ . The fitted  $\alpha$  value of 0.95, which closely approaches 1.00, indicates the good linearity of the  $I_{ph}$ - $P$  curves. This linear response characteristic facilitates the faithful reconstruction of the feature with the vision sensors<sup>38</sup>. The linear dynamic range (LDR) is subsequently calculated as  $LDR = 20\log(P_{max}/P_{min})$ , where  $P_{max}$  and  $P_{min}$  are the maximum and minimum power density of light stimuli, respectively. The dynamic range under the wavelengths of 520 and 980 nm light stimuli exceeds 80 dB. These results are larger than that of the state-of-the-art CMOS image sensors (70 dB) (ref. 17).

Figure 3c demonstrates the  $I$ - $t$  curves of the device when subjected to cyclic illuminations (1 Hz) of 520 nm light (bias voltage, +2 V) and 980 nm light (bias voltage, -2 V). The photocurrent shows no observable deviation after 70 cycles, indicating the excellent repeatability of the device. Figure 3d displays the  $I$ - $t$  curves of the device under 520 nm (bias voltage, +2 V) and 980 nm light (bias voltages, -2 V), at an on-off frequency of approximately 100 Hz. The result indicates that the response and recovery times are 100 and 90  $\mu$ s for the 520 nm light and 60 and 42  $\mu$ s for the 980 nm light, respectively. The response time is defined as the duration when the photocurrent increases from 10% to 90% of the stable maximum value, whereas the recovery time is defined as the duration when the photocurrent drops from 90% to 10% of the stable maximum value. This outcome indicates that the device is suitable for most scenarios that require real-time responses to rapidly changing environments, as it can respond faster than human eyes (~20 ms).

Quick adaptation between the visible and NIR regimes enables the vision sensors to alternatively capture the images of the target in the visible and NIR bands within a short time interval. Figure 3e shows the  $I$ - $t$  curves under constant 520 nm, constant 980 nm or 520 + 980 nm light conditions when the bias voltage periodically switches between 2 and -2 V. The device can rapidly adapt between the visible and NIR ranges with good repeatability. The current shows negligible attenuation even at a frequency of 100 kHz, which proves that the device can quickly adapt between light illuminations with wavelengths of 520 and 980 nm (Fig. 3f). This frequency is comparable with the frame rate of state-of-the-art commercial image sensors (frame rate, ~100 kHz)<sup>39</sup>.

Based on the bias-switchable spectral responses of the device, we carried out a bidirectional spectral adaptation test with an  $8 \times 8$  vision sensor array (Fig. 3g,i and Supplementary Fig. 10) to emulate the spectral adaptation behaviours. We adopt a 1 kHz square wave with  $\pm 2$  V amplitude and 10  $\mu$ s rise/fall edge ( $V_{drive}$ ) to drive the devices, which enables the switching between broadband response (+2 V) and narrowband NIR response (-2 V) of the devices, every time the polarity of  $V_{drive}$  flips.

To demonstrate the NIR-spectrum adaptation, we adopt an illumination pattern to resemble the digit '5' as the targeting feature in the imaging task (Fig. 3h). The pattern of 5 contains both NIR and visible illuminations. An additional visible pattern (Fig. 3g, dark-green-coloured zigzag) was simultaneously projected onto the array to mimic the interference in the environment. Supplementary Fig. 11 provides the details of the test setup and representative raw data. Figure 3h shows the photocurrent mapping at different adaptation times. Initially, the output cannot correctly represent the target 5 pattern, but exhibits a pattern like '6' (Fig. 3h(i)). Following the  $V_{drive}$  flipping from +2 to -2 V, the array switches to the narrowband NIR response. Consequently, the photocurrent mapping represents the correct 5 pattern (Fig. 3h(ii)-(iv)). To quantify the quality of the imaging results, we calculate the Weber contrast ( $C_w$ ) of the pattern 5 against the remaining part of the image (Supplementary Fig. 12)<sup>17,40</sup>. The detailed calculation of  $C_w$  is described in Supplementary Notes III and IV. The  $C_w$  value is 14.7 at the beginning. After the voltage flips for 16  $\mu$ s, the sensor is still in a transition period due to the junction capacitance; as the NIR-spectrum adaptation continues, the output image exhibits decreasing interference

from the visible 'zigzag' pattern. The  $C_w$  value increases to 59.4 at 48  $\mu$ s (Fig. 3h(iii)) and 62.7 at 160  $\mu$ s (Fig. 3h(iv)). The NIR-spectrum adaptation of the sensor array is similar to the redshifting adaptation of Pacific salmon. The narrowband NIR response matches well with the spectral intensity of the target feature of pattern 5, avoiding the zigzag interference from the visible band. Therefore, the resulting  $C_w$  value is over four times higher than that before the adaptation.

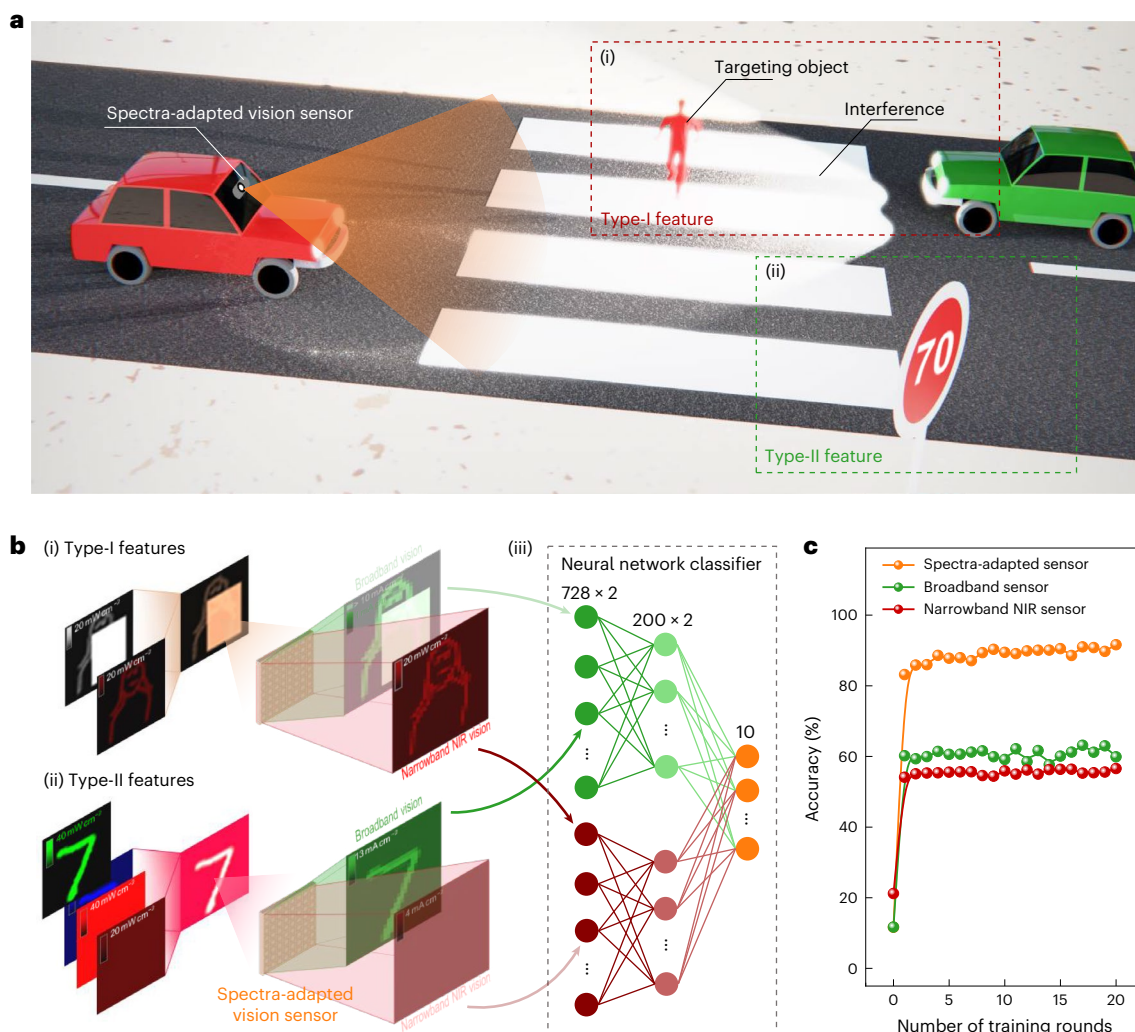
We also carry out the visible-spectrum adaptation test by projecting a pattern of '9' carried by visible illumination and an even background with NIR illumination onto the sensor array (Fig. 3i and Supplementary Fig. 13). Initially, the array with the narrowband NIR response fails to perceive the 9 pattern (Fig. 3i(i)), because the sensor array has no response to visible illumination. Following  $V_{drive}$  switching from -2 to +2 V, the sensor array switches from the narrowband NIR response to the broadband response. Sequentially, the 9 pattern shows up quickly in the imaging results (Fig. 3i(ii)-(iv)). The  $C_w$  value of the 9 pattern in the imaging result increases from 0.2 (0  $\mu$ s) to 1.8 (48  $\mu$ s) and finally to 3.0 at 160  $\mu$ s (Supplementary Fig. 14). The test result of the visible-spectrum adaptation indicates that the vision sensors can also work in scenarios where visible light carries critical information.

The spectra-adapted vision sensors realize two bias-switchable and spectrally distinctive visions, facilitating the capture of high-contrast features in spectrally variable environments. The fast adaptation speed for high  $C_w$  enhancement (400% in the NIR-spectrum adaptation test and 900% in the visible-spectrum adaptation test at 48  $\mu$ s) enables the sensors to capture spectrally distinctive features co-existing in the same scene, which is crucial for delay-sensitive applications such as autonomous vehicles.

## Accurate recognition of spectrally distinctive features

Accurately capturing and recognizing spectrally distinctive features in challenging light environments is difficult for conventional image sensors. Figure 4a illustrates a specific scenario of the situation. The machine vision system in an autonomous vehicle needs to recognize objects on the road exposed to background glares (Fig. 4a(i)). We define them as type-I features. Visible-light glare results in a high-illuminance background, and decreases the contrast of the targeting objects (Supplementary Fig. 15)<sup>41</sup>. To mitigate the interference from the visible-light glare and acquire clear images<sup>5,42</sup>, we can adopt a narrowband spectral response in the NIR region. In the other case, the machine vision system needs to recognize the information from the traffic sign with white digits and a red background (Fig. 4a(ii)). We define them as type-II features. As colour is a visual perception from the combination of visible-light stimuli at multiple wavelengths, a broadband response covering the visible spectrum is a prerequisite for accurately capturing the features represented by the colour differences (Supplementary Fig. 16)<sup>43</sup>. Conventional image sensors possess fixed spectral responses, which makes it difficult to maintain consistent performance in tackling type-I and type-II features. We build a customized dataset, including five classes of type-I features and five classes of type-II features (Supplementary Figs. 17 and 18)<sup>44,45</sup>, and adopt the spectra-adapted vision sensors to recognize spectrally distinctive features.

As shown in Fig. 4b, the spectra-adapted vision sensor captures the features in both broadband vision and narrowband NIR vision, and feeds the two-channel imaging output to the pre-trained neural network for recognizing the features. Figure 4b(i) exhibits the responses of the vision sensor to a pedestrian exposed to background glare, a typical type-I feature. The pedestrian exhibits reduced clarity in broadband vision because of the interference from the background glare. The narrowband NIR vision provides a clear outline of the pedestrian by effectively filtering out the glare. Consequently, the spectra-adapted sensor achieves 90% recognition accuracy in classifying the type-I features (Supplementary Fig. 20). This value is equal to the accuracy in recognizing the objects without the interference from background



**Fig. 4 | Spectra-adapted vision sensor for imaging and classifying spectrally distinctive features.** **a**, Schematic of the application scenario of the spectra-adapted vision sensors. **b**, Operational scheme of the spectra-adapted vision system. Panels (i) and (ii) show examples of the imaging of typical type-I and type-II features by the spectra-adapted vision sensor, and panel (iii) sketches the

architecture of the two-channel artificial neural network. The displayed type-I and type-II features are adapted from ‘The Mona Lisa’ image from the Quick, Draw! dataset and a ‘7’ image from the MNIST test set, respectively. **c**, Recognition accuracies of the ten classes in the two types of spectrally distinctive feature.

glare (Supplementary Note V and Supplementary Figs. 19 and 20). This in-sensor spectral adaptation realizes the anti-glare function without adopting optical accessories or complex algorithm resources. In contrast, the vision system solely based on a conventional broadband sensor fails to recognize the interfered object features with a low recognition accuracy of 33% (Supplementary Fig. 21).

Figure 4b(ii) exhibits the responses of the vision sensor to a white ‘7’ pattern on the red background, resembling a typical type-II feature. Although the narrowband NIR vision does not produce notable contrast for the digit, the broadband visible vision can yield clear contrast for the ‘7’ pattern. The spectra-adapted vision system achieves 90% accuracy in recognizing type-II features (Supplementary Fig. 20), which is comparable with that achieved by a conventional broadband sensor (Supplementary Fig. 21). In contrast, another reference vision system solely based on the narrowband NIR vision fails to recognize the type-II features (Supplementary Fig. 21).

Figure 4c summarizes the accuracy in simultaneously classifying all the ten classes in the two types of spectrally distinctive feature. The spectra-adapted vision sensors can capture two types of spectrally distinctive feature, resulting in a recognition accuracy of 90%. In contrast, the conventional vision sensors, broadband or narrowband NIR, show

poor performance on either type-I or type-II features, thereby yielding low recognition accuracies of 60% and 57%, respectively. The simultaneous recognition of spectrally distinctive features contributes to the comprehensive understanding of the surrounding environment. The spectra-adapted vision sensors provide a way to enhance the reliability of machine vision in challenging lighting conditions.

## Conclusions

We have reported a vision sensor array that is based on back-to-back photodiodes with an ITO/TiO<sub>2</sub>/Sb<sub>2</sub>Se<sub>3</sub>/Si/Ag structure and can emulate the spectral adaptation functions of the retina of migratory fish. The vision sensor includes a surface shallow TiO<sub>2</sub>/Sb<sub>2</sub>Se<sub>3</sub> junction for a broad spectral response range in the visible spectrum and a bottom deep Sb<sub>2</sub>Se<sub>3</sub>/Si junction for a narrow spectral response range within the NIR spectrum. The bias voltage applied to the device can be used to switch between the shallow and deep junctions. Consequently, the central photoresponse can be shifted between the visible and NIR spectra, enabling in-sensor spectral adaptation. The spectral adaptation can improve the Weber contrast of the captured images by adapting the spectral responsivity to match the predominant wavelengths of the features of interest, thereby facilitating the recognition of spectrally



distinctive features in challenging lighting conditions. Our in-sensor spectral adaptation approach could be used to enhance the reliability of machine vision without consuming additional computational resources.

## Methods

### Fabrication of back-to-back photodiodes array

First, n-type Si wafers with a thickness of 400  $\mu\text{m}$  (phosphorus-doped (100) wafer, 0.05–0.20  $\Omega\text{ cm}$ ) were ultrasonically cleaned in turn by deionized water, isopropyl alcohol, ethyl alcohol and acetone, with each step lasting 5 min. The natural oxide layer of Si wafers was removed by buffered oxide etch. After the chamber pressure of the thermal evaporator reached below  $3 \times 10^{-6}$  torr, the 0.8  $\mu\text{m}$   $\text{Sb}_2\text{Se}_3$  thin film was evaporated onto the cleaned Si wafers at a substrate temperature of 300  $^\circ\text{C}$ , at a deposition rate of  $\sim 3 \text{ \AA s}^{-1}$ . To improve the crystallinity of the  $\text{Sb}_2\text{Se}_3$  thin film,  $\text{Sb}_2\text{Se}_3/\text{Si}$  was annealed at 300  $^\circ\text{C}$  in a tube furnace with the protection of the argon atmosphere for an hour. Then, n-type  $\text{TiO}_2$  was sputtered on the annealed substrate at room temperature with a thickness of 220 nm to form a  $\text{TiO}_2/\text{Sb}_2\text{Se}_3/\text{Si}$  structure. Subsequently, a hard mask was placed on the surface of  $\text{TiO}_2$ , after which a 0.22- $\mu\text{m}$ -thick patterned ITO top electrode was deposited onto the  $\text{TiO}_2$  layer using magnetron sputtering at room temperature. Finally, a 300 nm Ag bottom electrode was deposited on the back of the Si substrate through magnetron sputtering at room temperature. The top view of the fabricated  $14 \times 16$  sensor array is displayed in Supplementary Fig. 10a.

### Materials and device characterization

The surface morphologies of each material and cross-sectional images of the device were acquired by a scanning electron microscope (Tescan MAIA3). The X-ray diffraction spectra of each material were obtained by an X-ray diffractometer (Rigaku SmartLab). The transmission spectra of these thin films were characterized by a UV–visible–NIR spectrometer (PerkinElmer). The electrical signal was obtained by a Keithley 4200-SCS source meter connected to a probe station (Lakeshore). Lasers equipped with a pulse generator were utilized to generate constant or periodical light illumination. EQE and responsivity–wavelength curves of the device were obtained by a monochromator system connected to a Keithley 2400 source meter.

### Technology computer-aided design simulation

The photoelectron generation rate in the bias-switchable vision sensor was simulated using Silvaco technology computer-aided design software. A two-dimensional model was constructed with the dimensions of  $10 \times 400 \mu\text{m}^2$ , consisting of 0.22- $\mu\text{m}$ -thick ITO, 0.22- $\mu\text{m}$ -thick  $\text{TiO}_2$ , 0.80- $\mu\text{m}$ -thick  $\text{Sb}_2\text{Se}_3$  and 400.00- $\mu\text{m}$ -thick n-type uniformly doped Si substrate ( $\sim 10^{18} \text{ cm}^{-3}$ ). The bandgaps of  $\text{TiO}_2$  and  $\text{Sb}_2\text{Se}_3$  are set to 3.3 and 1.3 eV, respectively. To balance the computing time and precision, the mesh space in the x direction is set to 0.50  $\mu\text{m}$ , whereas the mesh space in the z direction is set to 0.01, 0.01 and 0.02  $\mu\text{m}$  for  $\text{TiO}_2$ ,  $\text{Sb}_2\text{Se}_3$  and Si materials, respectively. This mesh setting was chosen for simulations since finer grid points have been attempted without notable changes observed. An incident beam was introduced onto the upper surface of the structure at an incident angle of 90 $^\circ$ .

### Simulation of neural network classifier

In the spectra-adapted vision system, the neural network classifier has a two-channel input layer with 784 neurons in each channel and a corresponding two-channel hidden layer with 200 neurons in each channel. The output layer has ten neurons corresponding to the ten classes in the recognition tasks. The activation function used between the input layer and hidden layer is sigmoid. The activation function between the hidden layer and output layer is softmax. In the reference vision systems, the neural network has the same three-layer architecture, with a single-channel 784-neuron input layer, single-channel 400-neuron

hidden layer and 10-neuron output layer. The activation function used between the input layer and hidden layer is sigmoid. The activation function used between the hidden layer and output layer is softmax. The detailed training and testing processes of the vision system are noted in Supplementary Note V.

## Data availability

Source data are provided with this paper. Other data related to this study are available from the corresponding author upon reasonable request.

## Code availability

The code used in this study is available via GitHub at [https://github.com/Jialiang-AP-WANG/ADBPD\\_simulation/](https://github.com/Jialiang-AP-WANG/ADBPD_simulation/).

## References

- Härer, A., Meyer, A. & Torres-Dowdall, J. Convergent phenotypic evolution of the visual system via different molecular routes: how Neotropical cichlid fishes adapt to novel light environments. *Evol. Lett.* **2**, 341–354 (2018).
- Al Naboulsi, M., Sizun, H. & de Fornel, F. Fog attenuation prediction for optical and infrared waves. *Opt. Eng.* **43**, 319–329 (2004).
- Zang, S. Z. et al. The impact of adverse weather conditions on autonomous vehicles: how rain, snow, fog, and hail affect the performance of a self-driving car. *IEEE Veh. Technol. Mag.* **14**, 103–111 (2019).
- Panetta, K., Gao, C. & Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **41**, 541–551 (2016).
- Gao, X. C. et al. Removing light interference to improve character recognition rate by using single-pixel imaging. *Opt. Lasers Eng.* **140**, 106517 (2021).
- Pi, L. J. et al. Broadband convolutional processing using band-alignment-tunable heterostructures. *Nat. Electron.* **5**, 248–254 (2022).
- Lee, S., Peng, R., Wu, C. & Li, M. Programmable black phosphorus image sensor for broadband optoelectronic edge computing. *Nat. Commun.* **13**, 1485 (2022).
- Hwang, A. et al. Visible and infrared dual-band imaging via Ge/MoS<sub>2</sub> van der Waals heterostructure. *Sci. Adv.* **7**, eabj2521 (2021).
- Shraddha, C., Chayadevi, M. L. & Anusuya, M. A. Noise cancellation and noise reduction techniques: a review. In *2019 1st International Conference on Advances in Information Technology* 159–166 (IEEE, 2019).
- Almalioglu, Y., Turan, M., Trigon, N. & Markham, A. Deep learning-based robust positioning for all-weather autonomous driving. *Nat. Mach. Intell.* **4**, 749–760 (2022).
- Verma, G. & Kumar, M. Under-water image enhancement algorithms: a review. *AIP Conf. Proc.* **2721**, 040031 (2023).
- Laiho, M., Poikonen, J. & Paasio, A. *Focal-Plane Sensor-Processor Chips* (Springer, 2011).
- Chai, Y. In-sensor computing for machine vision. *Nature* **579**, 32–33 (2020).
- Zhou, F. C. & Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **3**, 664–671 (2020).
- Zhou, F. et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **14**, 776–782 (2019).
- Wan, T., Ma, S., Liao, F., Fan, L. & Chai, Y. Neuromorphic sensory computing. *Sci. China Inf. Sci.* **65**, 141401 (2021).
- Liao, F. Y. et al. Bioinspired in-sensor visual adaptation for accurate perception. *Nat. Electron.* **5**, 84–91 (2022).
- Nascimento, A. M. et al. A systematic literature review about the impact of artificial intelligence on autonomous vehicle safety. *IEEE Trans. Intell. Transp. Syst.* **21**, 4928–4946 (2020).



19. Esteva, A. et al. A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
20. Oike, Y. Expanding human potential through imaging and sensing technologies. In *2022 International Electron Devices Meeting (IEDM) 1.2.1–1.2.5* (IEEE, 2022).
21. Baytamouny, M., Kolandaisamy, R. & AlDharhani, G. S. AI-based home security system with face recognition. In *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)* 1038–1042 (IEEE, 2022).
22. Corbo, J. C. Vitamin A<sub>1</sub>/A<sub>2</sub> chromophore exchange: its role in spectral tuning and visual plasticity. *Dev. Biol.* **475**, 145–155 (2021).
23. Enright, J. M. et al. Cyp27c1 red-shifts the spectral sensitivity of photoreceptors by converting vitamin A<sub>1</sub> into A<sub>2</sub>. *Curr. Biol.* **25**, 3048–3057 (2015).
24. Beatty, D. D. A study of the succession of visual pigments in Pacific salmon (*Oncorhynchus*). *Can. J. Zool.* **44**, 429–455 (1966).
25. Chen, X. W. et al. Turbidity compensation method based on Mie scattering theory for water chemical oxygen demand determination by UV-vis spectrometry. *Anal. Bioanal. Chem.* **413**, 877–883 (2021).
26. Peli, E. Contrast in complex images. *J. Opt. Soc. Am. A* **7**, 2032–2040 (1990).
27. Jang, H. et al. In-sensor optoelectronic computing using electrostatically doped silicon. *Nat. Electron.* **5**, 519–525 (2022).
28. Pan, C., Zhai, J. & Wang, Z. L. Piezotronics and piezo-phototronics of third generation semiconductor nanowires. *Chem. Rev.* **119**, 9303–9359 (2019).
29. Xu, Y. et al. Chalcogenide-based narrowband photodetectors for imaging and light communication. *Adv. Funct. Mater.* **33**, 2212523 (2022).
30. Fang, Y. J., Dong, Q. F., Shao, Y. C., Yuan, Y. B. & Huang, J. S. Highly narrowband perovskite single-crystal photodetectors enabled by surface-charge recombination. *Nat. Photon.* **9**, 679–686 (2015).
31. Tang, X., Ackerman, M. M., Chen, M. L. & Guyot-Sionnest, P. Dual-band infrared imaging using stacked colloidal quantum dot photodiodes. *Nat. Photon.* **13**, 277–282 (2019).
32. Xie, B. et al. Self-filtering narrowband high performance organic photodetectors enabled by manipulating localized Frenkel exciton dissociation. *Nat. Commun.* **11**, 2871 (2020).
33. Blair, S. et al. Hexachromatic bioinspired camera for image-guided cancer surgery. *Sci. Transl. Med.* **13**, eaaw7067 (2021).
34. Blair, S. et al. Decoupling channel count from field of view and spatial resolution in single-sensor imaging systems for fluorescence image-guided surgery. *J. Biomed. Opt.* **27**, 096006 (2022).
35. Yoon, H. H. et al. Miniaturized spectrometers with a tunable van der Waals junction. *Science* **378**, 296–299 (2022).
36. Chen, J. et al. Optoelectronic graded neurons for bioinspired in-sensor motion perception. *Nat. Nanotechnol.* **18**, 882–888 (2023).
37. Li, K. et al. Filter-free self-power CdSe/Sb<sub>2</sub>(S<sub>1-x</sub>Se<sub>x</sub>)<sub>3</sub> nearinfrared narrowband detection and imaging. *InfoMat* **3**, 1145–1153 (2021).
38. Chai, Y. Silicon photodiodes that multiply. *Nat. Electron.* **5**, 483–484 (2022).
39. *IDT MotionPro® Y-Series Digital Cameras* (2023); <https://www.delimaging.com/camera/idt-motionpro-y-series-compact-digital-cameras/>
40. Kwak, D., Polyushkin, D. K. & Mueller, T. In-sensor computing using a MoS<sub>2</sub> photodetector with programmable spectral responsivity. *Nat. Commun.* **14**, 4264 (2023).
41. Bullough, J. D., Van Derlofske, J., Fay, C. R. & Dee, P. Discomfort glare from headlamps: interactions among spectrum, control of gaze and background light level. *SAE Tech. Pap.* 2003-2001-0296 (2003).
42. Hu, J. B., Guo, Y. P., Wang, R. H., Ma, S. & Yu, A. L. Study on the influence of opposing glare from vehicle high-beam headlights based on drivers' visual requirements. *Int. J. Environ. Res. Public Health* **19**, 2766 (2022).
43. Bloj, M. & Hedrich, M. In *Handbook of Visual Display Technology* Ch. 15 (Springer, 2012).
44. Jongejan, J., Rowley, H., Kawashima, T., Kim, J. & Fox-Gieg, N. The Quick, Draw! Dataset (2016); <https://quickdraw.withgoogle.com/>
45. Yann, L., Corinna, C. & Christopher, J. C. B. The MNIST Database (1998); <http://yann.lecun.com/exdb/mnist/>

## Acknowledgements

This work is supported by MOST National Key Technologies R&D Programme (SQ2022YFA1200118-04), Research Grant Council of Hong Kong (CRS\_PolyU502/22) and The Hong Kong Polytechnic University (1-ZE1T, YXBA and WZ4X).

## Author contributions

Y.C. conceived the concept and supervised the project. B.O. fabricated the devices. B.O. and J.W. designed the test protocol and performed the experiments. Y.C., B.O. and J.W. analysed the experimental data. G.Z. performed the technology computer-aided design simulation. J.W. performed the simulation of the artificial neural networks. B.O., J.W. and Y.C. co-wrote the paper. All the authors discussed the results and commented on the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41928-024-01208-x>.

**Correspondence and requests for materials** should be addressed to Yang Chai.

**Peer review information** *Nature Electronics* thanks Weida Hu, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2024