

# A reconfigurable heterostructure transistor array for monocular 3D parallax reconstruction

Received: 14 November 2023

Accepted: 16 September 2024

Published online: 10 January 2025



Zhexin Li<sup>1,2</sup>, Hao Xu<sup>1</sup>, Yiqiang Zheng<sup>1</sup>, Lingchen Liu<sup>1,2</sup>, Linlin Li<sup>1,2</sup>, Zheng Lou<sup>1,2</sup>✉ & Lili Wang<sup>1,2</sup>✉

Sensors that are capable of three-dimensional detection of depth field information in the spatial domain are of potential use in applications such as robotics, satellite imaging and medical assistance. However, current techniques require a precise light source for complex phase detection and diffraction, or involve static multidirectional reflection imaging. Here we report a reconfigurable heterostructure transistor array for monocular three-dimensional parallax reconstruction. The phototransistors are based on heterostructures of indium gallium zinc oxide and tungsten diselenide, and can operate as n-type, p-type or ambipolar transistors depending on electrostatic modulation. The arrays can be switched between two modes: a real-time constant perception mode for static imaging and a spatiotemporal planar configuration mode with memory for dynamic imaging. To switch between the modes, the dominant carrier polarity is changed via a complementary metal–oxide–semiconductor-compatible multiterminal addressing architecture. We show that the system can be used for three-dimensional morphology reconstruction, two-dimensional depth field mapping and multi-view coupling.

Light-field detection can be used to collect multidimensional information that goes beyond the capabilities of conventional intensity-based two-dimensional (2D) photography<sup>1–3</sup>. Depth field information in the spatial domain can also be obtained by using the phase information of an object<sup>4–6</sup>, diffracted light waves<sup>7–9</sup> or reflection azimuth in multi-dimensional detection<sup>10</sup>. However, with current light-field detection, the phase or diffraction field requires specific light sources to induce light–matter interactions for optical resonances<sup>11,12</sup>, and comprehensive acquisition of information with the reflection field requires multi-aperture photography<sup>13</sup>.

As a result, integrating advanced light-field detection into complementary metal–oxide–semiconductor (CMOS) architectures are complex and costly<sup>14–16</sup>. Reference calibration in multi-camera systems can also be intricate<sup>17</sup>. **Several sensors have been created that are based on structured light and binocular images, and can be used to measure**

**depth fields<sup>18–21</sup>. However, developing advanced imaging hardware for monocular three-dimensional (3D) information reconstruction remains challenging.**

In this Article, we report a reconfigurable polarity conversion heterostructure transistor (PCHT) array for monocular 3D information reconstruction. **Our approach is inspired by the correlated method of structure from motion, which can reconstruct 3D information from 2D image sequences<sup>22,23</sup>.** In particular, we create a monocular imaging system that can record both optical intensity and the spatiotemporal parallax relationship at the sensor level. The system can reconstruct 3D objects from stored, dynamic image sequences using the binocular parallax principle.

We use pixelated phototransistors based on heterostructures of indium gallium zinc oxide (IGZO) and tungsten diselenide (WSe<sub>2</sub>) that can be operated as n-type, p-type or ambipolar transistors depending

<sup>1</sup>State Key Laboratory for Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China. <sup>2</sup>Center of Materials Science and Optoelectronic Engineering, University of Chinese Academy of Sciences, Beijing, China. ✉e-mail: [zlou@semi.ac.cn](mailto:zlou@semi.ac.cn); [liliwang@semi.ac.cn](mailto:liliwang@semi.ac.cn)

on electrostatic modulation (polarity conversion). These pixelated units can modulate both the phase and amplitude of the electric input, and can transfer between real-time constant perception for static imaging and spatiotemporal parallax configuration storage for dynamic imaging. Our PCHT array uses a CMOS-compatible multiterminal addressing architecture that is based on silicon selector transistors. Our system can reconstruct the spatial morphology of the free azimuth, and can compress the depth field into a planar map. We show that the approach can be used for 3D morphology reconstruction, 2D depth field mapping and multi-view coupling.

## Reconfigurable PCHT architecture

Figure 1a shows a schematic of our PCHT array architecture. Our reconfigurable PCHT devices are based on the inorganic covalent compound IGZO, which forms an n-type induction layer and the van der Waals crystal WSe<sub>2</sub>, which forms a p-type induction layer. We chose these materials because they have similar crystal configurations and lattice constants, which reduces the interfacial barrier that forms if the lattices are mismatched and reduces stress concentration (Supplementary Fig. 1a,b). These semiconductors also show efficient electric induction and tunable electrostatic doping<sup>24–26</sup>. Furthermore, WSe<sub>2</sub> has a layer dependent tunable optical bandgap (Supplementary Fig. 2b,d). We fabricated a monolayer that shows typical p-type semiconductor behaviour<sup>27–31</sup>.

Polarization Raman characterization shows that the origin of the anisotropy is mostly from the WSe<sub>2</sub> and that the structural symmetry is not disrupted by the heterogeneous integration with IGZO (Supplementary Fig. 3). Detailed characterization information can be found in the Methods and Supplementary Note 1. Phototransistor arrays with a back-gated configuration can decrease photon absorption of metal and enhance light–matter interactions in semiconductor<sup>32,33</sup>, and thus it is the device design we adopt here for our PCHTs. The output terminal is located above the depletion area of heterojunction and can output the intermediate potential and photo-induced carriers from the depletion area. The source and drain electrodes select the carrier type of our heterostructures through electrostatic doping, and the transport characteristics of photogenerated carriers is tuned by the electric induction of the gate electrode. Figure 1b and Supplementary Fig. 1c show interfacial spectrum mapping and cross-sectional element distribution of our heterostructure transistors.

We use a pseudo-crossbar array and combine it with a packaged read-out circuit. We reconstruct depth field information in the spatial domain through mapping the parallax relationship between sequential, stored configurations from a temporal series<sup>34</sup>. The relative location changes between sequential images (planar intensity domain) of a three-dimensional object in motion (spatial domain) dependent on the distance between the object and the photographic array. Objects further away from the imaging array have smaller displacements after planar mapping. Furthermore, in theory, an object at the infinity point would have approximately zero planar mapping displacement. Therefore, spatial depth field information can be reconstructed using the stored planar configuration according to the parallax principle (Fig. 1c).

Tunable macroscopical electric output characteristic of our heterostructures (polarity conversion) originates from selective, microscopic carrier activation (the dominant carrier) through electrostatic doping by different electrodes (Fig. 1d,e). When only the IGZO layer of the heterostructures is activated through electrostatic doping ( $V_1 = 5$  V,  $V_2 = 0$  V), we observe that the dominant carriers are electrons and our monolithic heterostructure shows a typical n-type transfer characteristic (Fig. 1e). If, in addition, the WSe<sub>2</sub> portion of heterostructure is activated under increasing bias with a positive gate voltage ( $V_1 = 5$  V,  $V_2 > 0$  V) an output enhancement is observed. An ambipolar transfer characteristic is observed when both semiconductors are activated by a similar voltage ( $V_1 \approx V_2 = 5$  V). However, when only the WSe<sub>2</sub> layer of

heterostructure is activated ( $V_1 = 0$  V,  $V_2 = 5$  V), the dominant carriers are holes and our monolithic heterostructure shows a typical p-type transfer characteristic (Supplementary Fig. 4a). This competition between the dominant carriers in the device leads to a tunable carrier polarity conversion and therefore the heterostructure transistor can modulate the amplitude and phase of a periodic triangular wave input voltage.

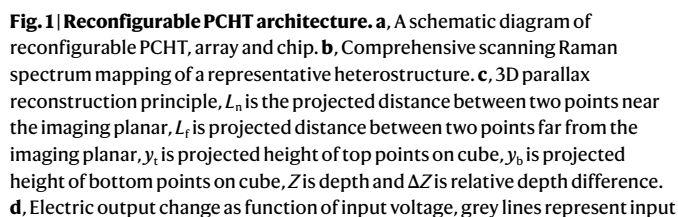
The heterostructure act as an amplitude rectifier when electrons are the dominant carriers. When holes are the dominant carriers the heterostructure operates as a phase inverter. Furthermore, we show that the output response is proportional to activation bias (Supplementary Fig. 4b–d). Notably, our reconfigurable PCHT shows ultralow power consumption ( $P = V_{dd} \times I_d$ , where  $I_d$  is channel current and  $V_{dd}$  is source-drain bias) according to the calculation in ref. 35, with a peak power consumption of 142 pW when  $V_{dd}$  is 5 V and the device operates as a phase inverter (Supplementary Fig. 4e). Supplementary Figs. 5 and 6 show leakage current, noise power density and the transfer behaviour of PCHTs in the array. We observe uniform behaviour across the devices, which is needed for robust imaging.

Supplementary Fig. 4f–h shows the distribution mapping of the current output and further supports our claims of polarity conversion of the dominant carriers. The gradient vectors of the current distribution depend on the gate voltage, and electrostatic doping from the source or drain have different effects on the PCHT, which further supports that the output distribution correlates with the domination carrier type. Finite element modelling supports our claims of a change in carrier polarity (Supplementary Fig. 7a–d). Our simulation results show that hole concentration dominates in the whole heterostructure when the WSe<sub>2</sub> layer of heterostructure is activated ( $V_1 = 0$  V,  $V_2 = 5$  V). On the other hand, the electrons concentration dominates when IGZO layer is activated ( $V_1 = 5$  V,  $V_2 = 0$  V). The n-type semiconductor shows a greater sensitivity to the gate control; an observation that is consistent with our experimental results.

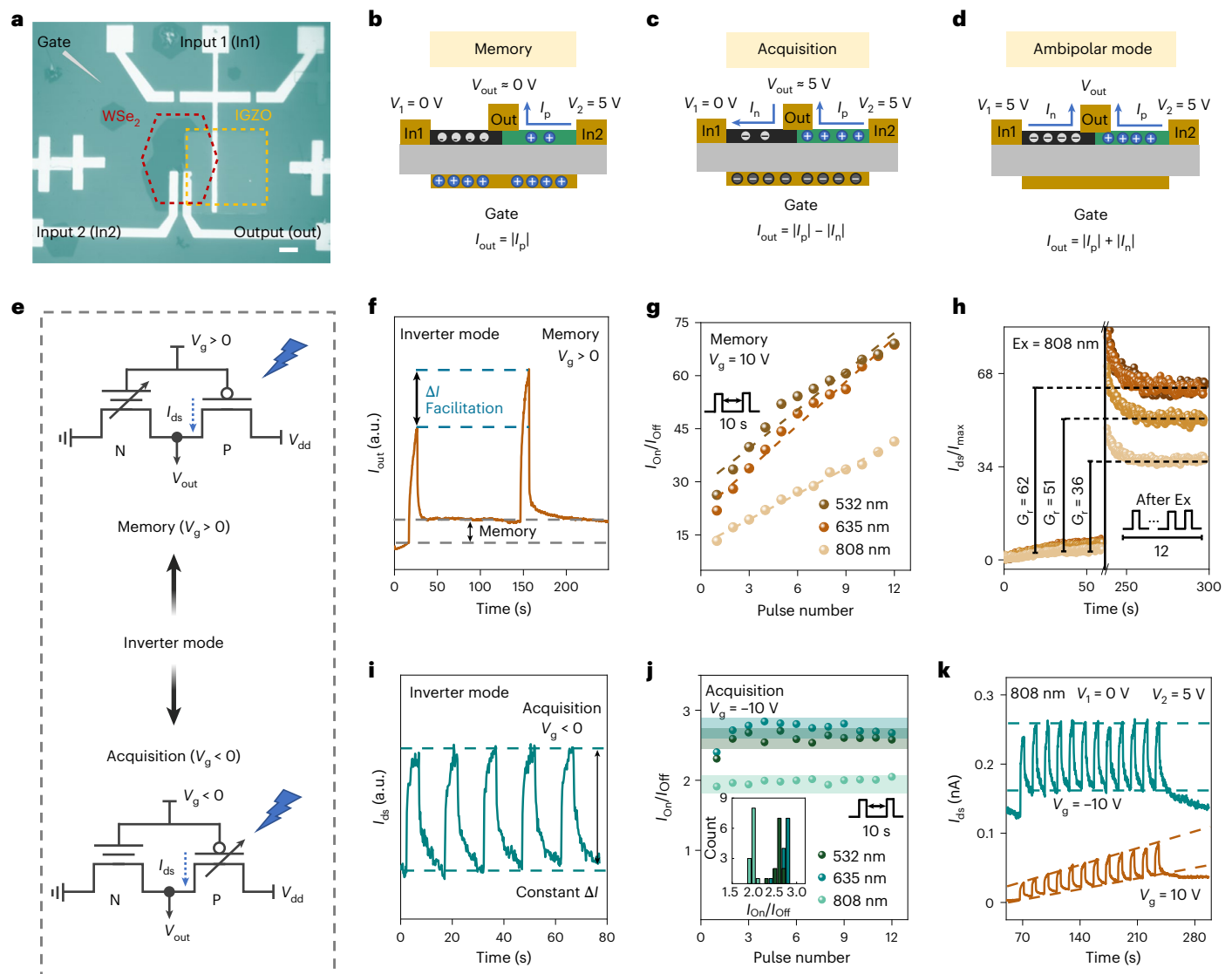
The polarity conversion results in a tunable macroscopical optoelectronic output characteristic (Fig. 1f). This optoelectronic tunability originates from selective carrier activation (via electrostatic doping and electric induction) and the competition between the dominant carriers. For optimum optoelectronic control, we operate the device in the phase inverter mode in which only the WSe<sub>2</sub> is activated (Fig. 1d, bottom). This ensures the preferable gate sensitivity of the IGZO does not totally dominate the output behaviour. The current output shows temporal-dependent enhancement through photo-induced relaxation when a positive voltage is applied on the gate. Interfacial defect states lead to in situ recombination relaxation of carriers for pixel information storage. On the other hand, the current output shows constant perception when a negative voltage is applied on the gate. The hole carriers activated by a small potential different and the potential-driven electron carriers from a large potential difference compete with each other until reaching a dynamic equilibrium.

## Optoelectronic tunability

We fabricated an optoelectronic, reconfigurable PCHT for optoelectronic conversion using lithographic patterning (Fig. 2a). Optoelectronic conversion originates from the tunable band structure and migrating carrier polarity. The selective carrier activation through electrostatic doping determines the barrier structure, and the electric induction modulates the Fermi level of the heterostructure. To accurately calibrate the band structure of the heterostructure, we theoretically model the flat band structure using CASTEP software (the Cambridge Sequential Total Energy Package). We experimentally measured the work function and bandgap using Kelvin probe force microscopy, which was calibrated via spectrum characterization (Supplementary Fig. 2c–f). The work function of the p-type and n-type semiconductors was measured to be 5.58 and 5.24 eV. The Fermi-level difference in the heterostructure, before contact, was calculated to be 0.34 eV.



We establish a macroscopic photocurrent flow model of the optoelectronic conversion according to the microscopic tunability



**Fig. 2 | PCHT mode-dependent optoelectronic performance.** **a**, Optical microscope image of the patterned reconfigurable PCHT device. Scale bar, 40  $\mu\text{m}$ . **b**, Model of the operation mechanism of the dynamic imaging mode with temporal-dependent storage.  $I_p$  is current flow in p-type semiconductor. **c**, Model of the operating mechanism of the constant perception mode for static imaging.  $I_n$  is current flow in n-type semiconductor. **d**, Model of the operation mechanism of the ambipolar mode. **e**, Equivalent functional circuit of the reconfigurable PCHT,  $V_g$  is gate voltage,  $V_{out}$  is voltage on output terminal, P represents a p-type semiconductor, N represents an n-type semiconductor and  $I_{ds}$  is the driven current in the channel. **f**, Temporal-varying current output of the PCHT in response to an optical excitation (532 nm, 0.38  $\text{mW cm}^{-2}$ ) when the device is in the temporal-dependent memory mode. **g**, Wavelength and excitation pulse number dependent optical response when the device is in the temporal-dependent memory mode. The pulse width is 5 s and the excitation intensities

of 532, 635 and 808 nm reach 0.38, 1.94 and 3.00  $\text{mW cm}^{-2}$ , respectively. **h**, Retention storage after optical excitation,  $G_i$  is stored grey level,  $E_i$  is excitation wavelength,  $I_{ds}$  is photoresponse and  $I_{max}$  is saturation response. **i**, Temporal-varying current output response to an optical excitation (532 nm, 0.38  $\text{mW cm}^{-2}$ ) while the PCHT operates in the constant perception mode,  $\Delta I$  is the variation of the photocurrent. **j**, Wavelength and excitation number dependent optical response of the device operated in the constant perception mode. The pulse width is 5 s and the excitation intensities of 532, 635 and 808 nm reach 0.38, 1.94 and 3.00  $\text{mW cm}^{-2}$ , respectively. The shaded regions indicate the on/off ratio window in which the device operates in the constant perception mode. **k**, Optoelectronic current,  $I_{ds}$ , response of the PCHT after optical exposure with wavelength of 808 nm. The pulse width is 5 s and excitation intensity of 808 nm reaches 3.00  $\text{mW cm}^{-2}$  in the static ( $V_g = -10\text{ V}$ ) and dynamic ( $V_g = 10\text{ V}$ ) mode.

discussed above. We operated the heterostructure transistor in the phase inverter mode to maximize the competition between carriers. When a positive voltage is applied on the gate, the photocurrent flows from the activation bias terminal to the output terminal of the p-type semiconductor according to the partial potential. Any positive electric induction inhibits the magnitude of this current. The n-type semiconductor is inactivated and contributes little to the carrier migration. The total output can be approximated by the intrinsic temporal-dependent photocurrent enhancement of the p-type semiconductor (Supplementary Fig. 8), which is described by  $I_{out} = |I_p(t)|$  (Fig. 2b). When a negative

voltage is applied to the gate, electric induction leads to an increase in the of hole concentration in the p-type semiconductor but this contributes little to the potential field and towards driving the carriers. The electrons in the n-type region reverse their flow direction due to the large partial voltage that leads to a reversed photocurrent. The total output can be described by  $I_{out} = |I_p(t)| - |I_n(t)|$  and includes the photocurrent competition affecting the dynamic equilibrium (Fig. 2c). Notably, the output of the heterostructure operated in the ambipolar mode is an improvement when compared with operating the device in the phase inverter mode. The bilateral biases lead to codirectional



photocurrent with synergetic enhancement and the total output can be described by  $I_{\text{out}} = |I_p(t)| + |I_n(t)|$  (Fig. 2d).

The PCHT circuit function is equivalent to a CMOS structure with two antitypic transistors (Fig. 2e and Supplementary Fig. 10a,b). Reconfigurable modulation affects the output from a serial or parallel architecture that leads to a tunable device that can be used for constant perception in form of static imaging and for spatiotemporal parallax storage used for dynamic imaging. The dynamic imaging mode depends on the recombination relaxation, which is similar to neuromorphic double pulse facilitation<sup>37,38</sup>. The dark current increases and is maintained at a relatively high level to form a memory window after 532 nm excitation (Fig. 2f). This temporal-dependent enhancement is linearly dependent on the optical pulse number. After 12 optical exposures, the on/off ratio improved from 13.32 to 41.41 under the illumination of an optical source with a wavelength of 808 nm  $s^{-1}$  ( $P = 3.00 \text{ mW cm}^{-2}$ ), from 21.86 to 68.73 under the illumination of an optical source with a wavelength of 635 nm ( $P = 1.94 \text{ mW cm}^{-2}$ ) and from 26.27 to 69.04 under the illumination of an optical source with a wavelength of 532 nm ( $P = 0.38 \text{ mW cm}^{-2}$ ), respectively (Fig. 2g).

To mimic the shuttered exposure of a conventional camera we operate our PCHT under a pulsed light exposure (rather than under a continuous illumination). The dark (resting) state level increase acts as a memory window and shows that photosensitive information is stored. In a conventional photo array, the dark state returns to its original resting state. The size of the memory window is correlated with the reflection intensity of captured features. A higher intensity leads to a larger memory window.

We divide memory window between dark state ( $I_{\text{dark}} \approx 0$ ) and saturation response ( $I_{\text{max}}$ ) into  $2^8$  levels (each level refers to one degree). The normalized  $I_{\text{ds}}/I_{\text{max}}$  level ( $G_r$ ) of the PCHT remains at  $62^\circ$  after the application of 12 illuminations at an intensity of  $4.55 \text{ mW cm}^{-2}$ , and shows a retention of  $51^\circ$  and  $36^\circ$  after illumination with an intensity of  $2.96$  and  $1.32 \text{ mW cm}^{-2}$ , respectively (Fig. 2h).

Static imaging demonstrates robustness to the optical excitation pulse number. The device response to an optical excitation (532 nm excitation,  $0.38 \text{ mW cm}^{-2}$ , 5 s excitation and 10 s resting) shows little change with temporal variation in the constant perception mode (Fig. 2i). The on/off ratio is relatively insensitive to the number of pulses irrespective of the exposure wavelength (the average on/off ratio is 1.98 at 808 nm, 2.73 at 635 nm and 2.59 at 532 nm, respectively) (Fig. 2j). The photoresponse at a wavelength of 808 nm shows optoelectronic conversion (Fig. 2k). The peak response reaches 0.26 nA at constant perception when a voltage of  $-10 \text{ V}$  is applied. The output response is larger than the device response in the temporal-dependent storage mode. A similar device output was observed in response to illumination at different wavelengths (Supplementary Fig. 10d–f). The difference in the response of the two imaging modes originates from the dominant carrier concentration due to electric induction. Under the application of a negative voltage at the gate, the activated p-type semiconductor has a larger hole concentration than when a positive voltage is applied at the gate. The photoresponse in the ambipolar mode is also consistent with the effect of carrier concentration modulation. Both bilateral semiconductors are activated and the synergetic current results in a larger response compared with the phase inverter imaging mode (Supplementary Fig. 10g–i). Moreover, the signal output is positively correlated to the exposure length of each activation. A longer exposure length results in a higher output that is similar to how CMOS imagers operate (Supplementary Fig. 11).

Notably, the PCHT shows temporal-dependent suppression under light excitation with a wavelength of 450 nm. Photoluminescence spectrum measurements show band edge excitation and a high sensitivity of IGZO to light at the ultraviolet wavelength. At this wavelength, the generated photoelectrons effect the dominant carrier distribution and with this the competition that breaks the dynamic equilibrium. This results in suppression of the device output and is

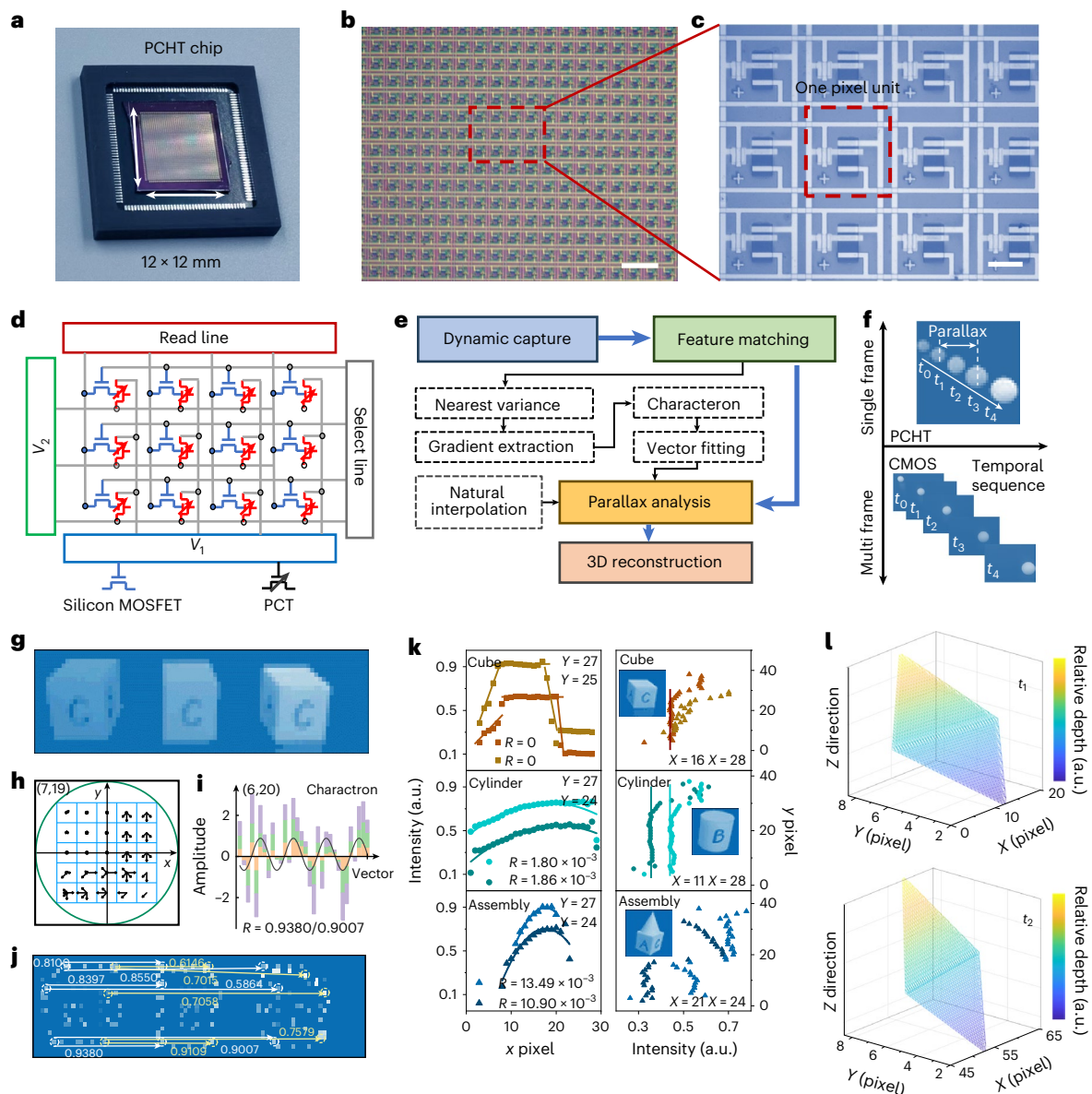
consistent with macroscopic photocurrent flow model in the constant perception mode (Fig. 2c). We propose that a colour balanced device could be achieved using three parallel PCHTs operated with different gate voltages. These PCHTs could modulate red (R), green (G) and blue (B) colour channels during pulse exposures, and could be used to transform images from cold to warm tones. Supplementary Fig. 12 shows the modulation range in the colour space on the basis of our response results. The histogram distribution of each RGB image after each exposure period shifts demonstrating balance modulation (Supplementary Fig. 12).

## Monolithic integration of the PCHT array

We monolithically integrate a PCHT array using lithographically patterning of heterostructure transistor integrated with addressing selective silicon metal–oxide–semiconductor field-effect transistors (MOSFETs). We fabricated a  $40 \times 40$  reconfigurable PCHTs array packaged with a read-out circuit. The array shows good static consumption ( $\sim 0.14 \text{ nW}$  per pixel) and high integration as its fabrication is CMOS-compatible (Supplementary Fig. 13). Other performance characteristics and comparisons are shown in Supplementary Table 2. An optical image of our PCHT array is shown in Fig. 3a. A high-density pixel cluster is aligned to a pseudo-crossbar using on-chip integration (Fig. 3b). The PCHT array has a total photosensitive area of  $12 \times 12 \text{ mm}$  and an integrated pixel size of  $300 \times 300 \mu\text{m}$ . The pixel photosensitive area is  $60 \times 120 \mu\text{m}$  (Fig. 3c). Silicon transistors act as selectors to avoid read-disturbance and are wired in series with the PCHT to form a pixelate cell. The detailed monolithic integration fabrication process of our PCHT array is shown in Supplementary Fig. 14. We design an addressing architecture for progressive scanning as read-out strategy. Such an addressing architecture is good for large-scale integration and optimizes the control logic burden. The selection terminals scan row-by-row, and the photoelectronic signals is read out from every column via read-out terminals. This architecture ensures just one PCHT output is performed in every column but 40 column groups are read out simultaneously in parallel for read-out efficiency and disturbance protection (Fig. 3d).

Our hardware architecture supports the 3D parallax reconstruction algorithmic methodology. We demonstrate rapid detection suitable for invariant feature matching and parallax reconstruction. **Our methodology is inspired by the scale invariant feature transform (SIFT)<sup>39</sup> and the Harris corner and edge detection algorithm<sup>40</sup>.** The methodological architecture and pipeline are illustrated in Fig. 3e. First, the integrated reconfigurable PCHT array captures motional features in different locations. Then optical flow gradient (OFG) correlated with feature and environmental optical interaction is constructed using the variance and gradient. The extracted OFG are encoded to form a hyperparameter vector describing their characteristics, which we call characterons. We compared these key feature characterons from different locations using vector fitting and match key feature pairs with high Euclidean correlation. These matched feature pairs only have planar coordinates, and the parallax analysis provides a depth field coordinate for these matched feature pairs. The complete spatial information is reconstructed using parallax-based spatial interpolation where the interpolation planar mapping coordinates are derived from the planar static images. The full spatial field can be reconstructed with our integrated PCHT array-based 3D parallax methodology. 3D parallax reconstruction is possible using our PCHT as both intensity and spatiotemporal parallax at the sensor level is recoded. On the other hand, conventional CMOS only records intensity and reconstructs parallax through a multi-frame approach (Fig. 3f).

The integrated reconfigurable PCHT array in the dynamic imaging mode captures a cube in motion (Fig. 3g). We use dynamic imaging to record three planar configurations of the cube. We record the object throughout its motion and read out the stored information in the PCHT array before each subsequent exposure. For each recorded



**Fig. 3 | 3D parallax reconstruction methodology.** **a**, Optical image of our monolithic integrated reconfigurable PCHT array. **b**, Low resolution optical microscope image of the integrated PCHT array. Scale bar, 600  $\mu\text{m}$ . **c**, High-resolution optical microscope image of a single pixel unit. The dashed black box encloses a single pixel unit. Scale bar, 100  $\mu\text{m}$ . **d**, Schematic of the reconfigurable PCHT array hardware architecture. **e**, Algorithmic methodology of the 3D parallax reconstruction. **f**, Schematic showing the conceptual difference between conventional CMOS imaging and PCHT array imaging of an object in

motion at different times,  $t_0$  to  $t_4$ . **g**, Dynamic imaging of a cube in motion at three moments in time. **h**, Distribution of the OFG vectors. **i**, Hyperparameter structure of the characterons associated with feature key points. **j**, Distribution and correlation of feature key points for feature matching. **k**, Surface reconstruction through the 2D optical intensity,  $R$  is radius of curvature,  $X$  is pixel column on imaging planar and  $Y$  is pixel row on imaging planar. **l**, 3D reconstruction of different configurations.

grayscale matrix at a certain moment in time, the variance of every pixel with respect to each of the adjacent 8 pixels is calculated to form a new variance matrix. The matrix mapping extracts some edges and corners due to large optical changes at these features (Supplementary Fig. 15). To filter subvariational points and improve the calculation efficiency, maximum scanning with a similar convolution kernel is used for down-sampling to find feature key points. These key points are used to calculate their direction dependent OFG vectors using the following method:  $G = (I(x, y) - I(x', y')) / \sqrt{(x - x')^2 + (y - y')^2}$ , where  $G$  is the variation gradient,  $I$  is the grey level of pixel,  $(x, y)$  and  $(x', y')$  are pixel coordinates. A bit space of global pixels in array with nine orientations

was established (Supplementary Fig. 16). Defining the feature key points as the coordinate origin, a new coordinate system is established in the range of a  $5 \times 5$  kernel. The vectors along the same direction in each quadrant are summarized to form four supervectors that include the information from the nine directions (Fig. 3h). The characterons with 36 hyperparameter vectors are constructed by these four supervectors arrangements. Figure 3i shows that features with high correlation have similar hyperparameter envelopes. For example, a sinusoidal distribution of key point (6, 20) is similar to its matching point (34, 20), (63, 20), which has a high coefficient of correlation in 0.9380 and 0.9007, respectively. Total feature key points, their matching pairs and coefficient of correlation were calibrated and are shown in Fig. 3j.

Surface reconstruction plays an important role in 3D parallax reconstruction after feature matching. Although some feature key points were calibrated, the spatial surface configuration among these points needs to be reconstructed. Our PCHT array also has the ability to store 2D information together with their spatial location relationship and optical intensity distribution, and can be used to extract the spatial morphology of the surface configuration. A plane has a linear optical intensity distribution while a curved surface has a quadratic or curved distribution. A frontal facing cube has a relatively constant distribution of optical intensities associated with its front facing side while it shows a linearly reduced distribution from its sides. A cylinder shows a smooth curved distribution with the same radius of curvature on its side and a linear distribution along its axis. A cone is represented as a smooth curve with a different radius of curvature. Figure 3k shows the optical distributions correlated with different spatial surface configurations. The detailed mathematical methodology of our parallax reconstruction and dynamic reconstruction of a plane, curved, circular and elliptic surface is shown in Supplementary Notes 2 and 3 and Supplementary Figs. 17–19. The results show that the spatial configuration of a cube in motion viewed from the top at two moments in time is well reconstructed. The reconstructed surfaces show different orientations due to their different locations (Fig. 3l).

We show that we can perform 3D parallax reconstruction when operating our monolithic PCHT array system in the memory mode at a frame rate of 33 Fps, which is comparable to the frame rate of a standard camera. The frame rate using the acquisition mode can also reach 30 Fps (Supplementary Fig. 20a). The average computation time for feature matching and reconstruction is around 4.648 s during a whole calculation period (Supplementary Fig. 20b). We note that the reconstruction speed for different shapes is slightly different (Supplementary Fig. 20c–e). Compared with a standard camera, the monolithic PCHT is optimized in terms of memory and computation as both intensity and spatiotemporal parallax are recorded at the sensor level. The memory burden is reduced by an order of magnitude according to our experiment and simulation results, and the read speed is improved (Supplementary Fig. 21). A detailed comparison of the PCHT array performance with other methods for 3D reconstruction is shown in Supplementary Table 2.

### 3D parallax reconstruction

Three-dimensional parallax reconstruction using our monolithic integrated PCHT array is feasible for 3D morphology reconstruction. In standard imaging, a planar configuration is captured of an object at a certain moment in time and it is not possible to view its spatial structure in different orientations. To demonstrate the advantage of the PCHT in overcoming this limitation, we reconstructed the stereo morphology of a complex object assembly and show its spatial structure from different points of view. Camera calibration and movement type are determined for the reconstruction (Supplementary Note 4 and Supplementary Fig. 22). Figure 4a shows the results of this morphology reconstruction. The optical grayscale imaging of the static planar configuration, ground truth of the assembly, is shown on the upper left. The morphology of the cone and cater-cornered cube are relatively undistorted. The uncovered top of the cube is clearly separation from the cone. The front view is almost identical to the ground truth. The left and right view shows the different rotation angles and also shows good reconstruction of the spatial configurations. Figure 4b shows 2D depth field mapping of the object assemblies. The spatial configuration of the complex assemblies is projected onto the 2D plane and the third-dimensional depth field information is represented by the colour mapping. Warm colours show that the points are closer to the reconfigurable PCHT array and cold tones represent locations that are further away. The 2D colour mapping accurately reflects depth field information of the spatial domain. The spatial morphology and 2D depth field mapping of other objects such as a frontal cube, cater-cornered cube, cone, cylinder and simple

assembly, are shown in Supplementary Figs. 23 and 24. Moreover, the parallax reconstruction of movement perpendicular to the camera plane further demonstrates the feasibility of parallax reconstruction for general motion scenarios (Supplementary Fig. 25).

As a proof-of-concept demonstration, we show multi-view coupling of different configurations for multidimensional reconstruction. Multi-viewing coupling reconstructs 2D depth field mapping and can uncover some unreflective features that are hidden by the objects structure. It comprehensively maps all the features from every planar configuration capture in the time series to a single 2D map. We demonstrated this for the frontal cube shown in Supplementary Fig. 23a. The first exposure only images the surface with the letters 'B' and 'C'. The second exposure only maps the surface with the letters 'A' and 'C'. To operate transformation of coordinates for multi-view coupling, the statistically distributed histogram of Euclidean distance and Euclidean angle shows an average horizontal pixel displacement of 46 and 0.6° of vertical shift (Fig. 4c), and from this the coordinates without offset can be calculated:  $\begin{vmatrix} x'_1 \\ y'_1 \end{vmatrix} = \begin{vmatrix} x_2 \\ y_2 \end{vmatrix} - X_{ed} \cdot \begin{vmatrix} \cos \theta_{ea} \\ \sin \theta_{ea} \end{vmatrix}$ , where  $\theta_{ea}$  is Euclidean angle,

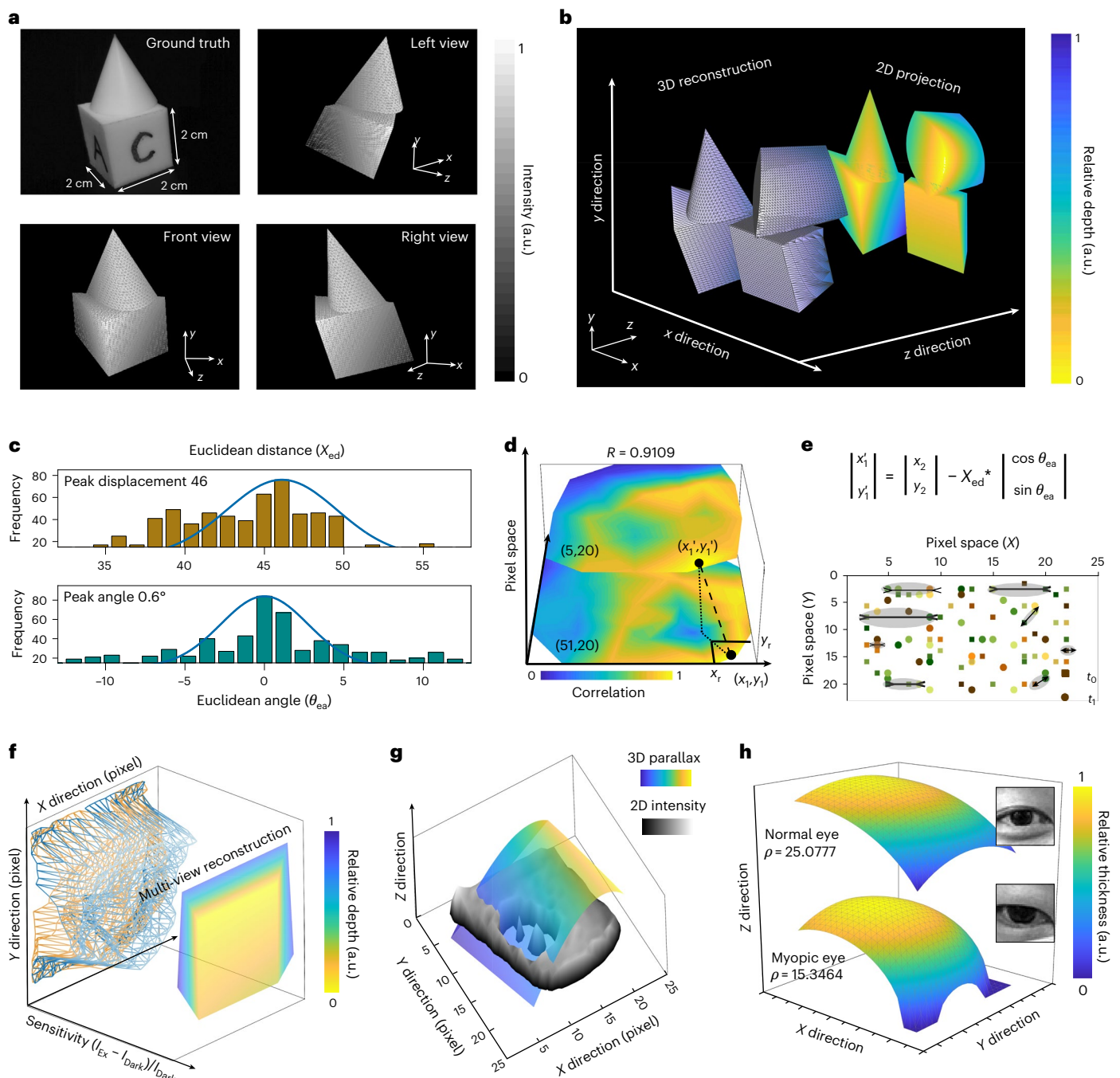
$X_{ed}$  is Euclidean distance ( $x_2, y_2$ ) is original coordinate and ( $x'_1, y'_1$ ) is coordinate after transformation. After the transformation of the coordinates, each key point from the different configurations is matched and paired and forms unified planar coordinates. The average  $x$  and  $y$  coordinates of the point pair with maximal correlation forms the updated key point coordinates (Fig. 4d). The key point distribution after the coordinate transformation is shown in Fig. 4e. The parallax analysis is used to reconstruct new spatial configurations on the basis of these new correlation key point pairs. The experimental captured dynamic data of the optical intensity in the two configurations, and the coupled multidimensional reconstruction are shown in Fig. 4f. Lateral opposite sides that cannot be displayed at the same time are simultaneously reconstructed. The depth field is also reconstructed via colour mapping. The relationship between the depth and structure of each object in the real spatial domain is considered and we show the reconstructed planar configuration after multi-view coupling and the distribution of the depth field.

Our integrated PCHT array can dynamically capture the spatiotemporal parallax configuration of objects in motion and reconstruct their spatial domain depth field using 3D parallax imaging. Compared with intensity-contrast 2D imaging, 3D parallax imaging can construct the third-dimensional depth field and map surface morphology of the spatial configuration. The depth field constructed by parallax is more robust to interference of surface texture than intensity-based reconstruction (Fig. 4g). Furthermore, the memory function of our spatiotemporal parallax configuration enables monocular dynamic detection. It captures multiple images of an object in motion at a fixed frame with fixed camera parameters in a single coordinate system and thus does not require calibration of each coordinate parameters, which is the case when using multiple camera imaging. These advantages mean our integrated reconfigurable PCHT array hardware could be used for advanced light-field detection applications. Figure 4h shows a proof-of-concept scenario that uses our hardware to assist the detection of eye diseases. 3D reconstruction of bulbus oculi using our PCHT based hardware could be used to analyse keratoconus and myopia in natural light conditions without special reflected ray tracing. Our results show that a myopic eye is more curved and has a smaller radius of curvature compared to a normal eye, which confirms myopia has a correlation with a larger curved and convex cornea (Supplementary Fig. 26).

### Conclusions

We have reported a reconfigurable PCHT array for monocular 3D parallax reconstruction. Our pixelated PCHTs exhibit polarity conversion between n-type, ambipolar and p-type via selective carrier activation through electrostatic doping and electric induction. These pixelated





**Fig. 4 | 3D parallax reconstruction demonstration.** **a**, Stereo morphology reconstruction of a complex object assembly. Scale bars, 10, 5 and 5 pixels in  $x$ ,  $y$ ,  $z$ , respectively. **b**, 2D depth field mapping of two spatial configurations. **c**, Statistical distribution of the Euclidean parameters. **d**, Correlation space of matched key point. **e**, Coordinate transformation methodology (top) and the

distribution of key points after coordinate transformation (bottom).

**f**, Demonstration of multi-view coupling,  $I_{ex}$  is excitation photocurrent and  $I_{dark}$  is dark current. **g**, Comparison between 2D intensity-contrast imaging and 3D parallax imaging. **h**, Surface reconstruction of the bulbus oculi of a normal (top) and myopic eye (bottom).  $\rho$  is the radius of curvature of the reconstructed eyes.

units can record both the intensity and spatiotemporal parallax configuration at the sensor level. The transistor can be toggled between a real-time constant perception mode for static imaging and a spatiotemporal planar configuration mode with memory for dynamic imaging. We demonstrated CMOS-compatible integration of our reconfigurable PCHT array with a multiterminal addressing hardware architecture, as well an algorithmic methodology for 3D parallax reconstruction. To illustrate the capabilities of this system, we used it for 3D morphology reconstruction, 2D depth field mapping and multi-view coupling.

## Methods

### Synthesis of polarity conversion heterojunction

To fabricate the polarity conversion heterostructure, p-type tungsten diselenide ( $\text{WSe}_2$ ) was synthesized via chemical vapour deposition. A mixture containing 0.010 g of  $\text{WO}_3$  powder and 0.002 g of NaCl powder was poured into a small oxide boat, the mixture was loaded at the centre of furnace and a piece of cleaned Si substrate covered with a 300 nm  $\text{SiO}_2$  was placed on the mixture with the polished surface down. The excess selenium powder was loaded at the upstream side. After continuous Ar flow cleaning in 30 min, the chemical vapour deposition



furnace was ramped up to 750 °C in 15 min, the vapour transportation was maintained for 5 min. During the preparation, 60 sccm gaseous mixture (10% H<sub>2</sub> + 90% Ar) was continuously supplied, all the reactions were carried out at atmospheric pressure. The temperature was naturally cooled down to room temperature after the preparation process. A standard photolithography process was used to define the area of the n-type semiconductor. The n-type IGZO was synthesized via facile magnetron sputtering. The n-type layer was deposited using a 70 W radio-frequency power source with 20 sccm argon. Then lithography and lift-off were used to fabricate the heterojunction devices.

### Transistor and array fabrication

Multi-mask photolithography was used to fabricate PCHTs and the pixel array. A schematic diagram of the fabrication process is shown in Supplementary Fig. 9. First, the Si window was defined and prepared by pattern-assisted wet etching. Then the heterojunction semiconductor was synthesized by magnetron sputtering and chemical vapour deposition. The silicon MOSFET source and drain and polarity conversion heterostructure were defined. Thermal evaporation was used to deposit Cr/Au (10/50 nm) for the electrodes. The source and drain contacts were revealed by lift-off in acetone. The dielectric layer was deposited and patterned by spin-coating SU-8 photoresist solution. The top gate of the silicon MOSFET was patterned by photolithography and Cr/Au (10/50 nm) was deposited by thermal evaporation.

### Material characterization

The lattice structure and element distribution of semiconductors in the heterojunction was characterized using a high-resolution transmission electron microscope with an EDS analyser (JEM-2010F) and a scanning electron microscope (Zeiss Supra55(VP)). The morphologies and work function of the semiconductors used in the heterojunction were characterized by atomic force microscopy and Kelvin probe force microscopy (Bruker). The spectral characteristic of materials was characterized by a laser Raman confocal microscope (Renishaw in Via-Reflex) with an excitation laser of 532 nm, and a photoluminescence with two excitation lasers of 405 and 635 nm. A probe station supporting a semiconductor characterization system (Keithley 1500) was used to measure the photoelectric characteristics of the materials. All measurements were performed at room temperature.

### Correlation calculation

The Pearson correlation coefficient,  $r$ , was used to evaluate the correlation between two different characters ( $A$  and  $B$ ) using the following equation:

$$r = \frac{\mathbf{A} \cdot \mathbf{B}}{\sqrt{\sum_{i=1}^n \mathbf{A}_i^2} \sqrt{\sum_{i=1}^n \mathbf{B}_i^2}}$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are two hyperparameter vectors and  $\mathbf{A}_i$  and  $\mathbf{B}_i$  are parameter-in-hyperparameter vectors. Two different characters have a strong correlation if  $|r| > 0.7$ , and the feature points represented by these two characters are matched.

### Density functional theory calculation

Ab initio calculations were performed using CASTEP on the basis of density functional theory. The generalized gradient Perdew–Burke–Ernzerhof approximation was adopted to handle the electron exchange–correlation functional. The kinetic energy cut-off for WSe<sub>2</sub> was set to 650 eV and the kinetic energy cut-off for IGZO was set to 780 eV.

### Finite element analysis

The finite element analysis of device model was calculated using COMSOL Multiphysics. The steady study of semiconductor module

was used to analysis carrier distribution. The 2D semiconductors were simplified into a block of two pieces. The surface area of the p-type semiconductor was set at 20 × 50 μm while the surface area of n-type was 50 × 50 μm. The covered area of electrode on each semiconductor was 5 × 50 μm. The effective state density of valence band and conduction band in p-type semiconductor was set 7.57 × 10<sup>18</sup> and 4.75 × 10<sup>18</sup> 1 per m<sup>3</sup>, respectively. The effective state density of valence band and conduction band in n-type semiconductor was set 5 × 10<sup>18</sup> and 5 × 10<sup>18</sup> 1 per m<sup>3</sup>, respectively.

### Data availability

The data that support the findings of this study are available via Figshare at <https://doi.org/10.6084/m9.figshare.26764015> (ref. 41). Source data are provided with this paper.

### Code availability

The code used for image reconstruction is available from the corresponding author on reasonable request.

### References

- Wu, J. et al. An integrated imaging sensor for aberration-corrected 3D photography. *Nature* **612**, 62–71 (2022).
- Lien, M. et al. Ranging and light field imaging with transparent photodetectors. *Nat. Photonics* **14**, 143–148 (2020).
- Feng, X., Ma, Y. & Gao, L. Compact light field photography towards versatile three-dimensional vision. *Nat. Commun.* **13**, 3333 (2022).
- Jing, X. et al. Single-shot 3D imaging with point cloud projection based on metadvice. *Nat. Commun.* **13**, 7842 (2022).
- Huang, L. et al. Three-dimensional optical holography using a plasmonic metasurface. *Nat. Commun.* **4**, 2808 (2013).
- Park, J., Lee, K. & Park, Y. Ultrathin wide-angle large-area digital 3D holographic display using a non-periodic photon sieve. *Nat. Commun.* **10**, 1304 (2019).
- Vamvakeros, A. et al. 5D operando tomographic diffraction imaging of a catalyst bed. *Nat. Commun.* **9**, 4751 (2018).
- Li, J. et al. Rapid sensing of hidden objects and defects using a single-pixel diffractive terahertz sensor. *Nat. Commun.* **14**, 6791 (2023).
- Cherukara, M. J. et al. Three-dimensional X-ray diffraction imaging of dislocations in polycrystalline metals under tensile loading. *Nat. Commun.* **9**, 3776 (2018).
- Yi, L., Hou, B., Zhao, H. & Liu, X. X-ray-to-visible light-field detection through pixelated colour conversion. *Nature* **618**, 281–286 (2023).
- Shen, Z. et al. Monocular metasurface camera for passive single-shot 4D imaging. *Nat. Commun.* **14**, 1035 (2023).
- Huang, X. et al. Polarization structured light 3D depth image sensor for scenes with reflective surfaces. *Nat. Commun.* **14**, 6855 (2023).
- Fang, J., Huang, K., Wu, E., Yan, M. & Zeng, H. Mid-infrared single-photon 3D imaging. *Light Sci. Appl.* **12**, 144 (2023).
- Zhang, X., Kwon, K., Henriksson, J., Luo, J. & Wu, M. C. A large-scale microelectromechanical systems-based silicon photonics LiDAR. *Nature* **603**, 253–258 (2022).
- Rogers, C. et al. A universal 3D imaging sensor on a silicon photonics platform. *Nature* **590**, 256–261 (2021).
- Goossens, S. et al. Broadband image sensor array based on graphene-CMOS integration. *Nat. Photonics* **11**, 366–371 (2017).
- Zhang, Z. A flexible new technique for camera calibration. *IEEE T. Pattern. Anal.* **22**, 1330–1334 (2000).
- Tachella, J. et al. Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers. *Nat. Commun.* **10**, 4984 (2019).
- Zhou, K. et al. Parallelized computational 3D video microscopy of freely moving organisms at multiple gigapixels per second. *Nat. Photon.* **17**, 442–450 (2023).

20. Siavashani, M. et al. 3D imaging using scanning diffractometry. *Sci. Rep.* **11**, 482 (2021).
21. Scholler, J. et al. Dynamic full-field optical coherence tomography: 3D live-imaging of retinal organoids. *Light Sci. Appl.* **9**, 140 (2020).
22. Ozyesil, O., Voroninski, V., Basri, R. & Singer, A. A survey of structure from motion. *Acta Numer.* **26**, 305–364 (2017).
23. Schonberger, J. L. & Frahm, J. Structure-from-motion revisited. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 4104–4113 (IEEE, 2016).
24. Mennel, L. et al. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–66 (2020).
25. Lee, J. et al. Thin-film image sensors with a pinned photodiode structure. *Nat. Electron.* **6**, 590–598 (2023).
26. Pan, C. et al. Reconfigurable logic and neuromorphic circuits based on electrically tunable two-dimensional homojunctions. *Nat. Electron.* **3**, 383–390 (2020).
27. Choi, H. et al. A steep switching WSe<sub>2</sub> impact ionization field-effect transistor. *Nat. Commun.* **13**, 6076 (2022).
28. Seyler, K. L. et al. Electrical control of second-harmonic generation in a WSe<sub>2</sub> monolayer transistor. *Nat. Nanotechnol.* **10**, 407–411 (2015).
29. Pudasaini, P. R. et al. High-performance multilayer WSe<sub>2</sub> field-effect transistors with carrier type control. *Nano Res.* **11**, 722–730 (2018).
30. Cong, X. et al. Interplay of valley polarized dark trion and dark exciton-polaron in monolayer WSe<sub>2</sub>. *Nat. Commun.* **14**, 5657 (2023).
31. Ren, Y. et al. Recent advances in ambipolar transistors for functional applications. *Adv. Funct. Mater.* **29**, 1902105 (2019).
32. Liao, F. et al. Bioinspired in-sensor visual adaptation for accurate perception. *Nat. Electron.* **5**, 84–91 (2022).
33. Li, Z. et al. Parallel photoelectron storage and visual preprocessing based on nanowire defect engineering for image degradation. *Adv. Funct. Mater.* **34**, 2304119 (2024).
34. Kellnhofer, P. et al. Motion parallax in stereo 3D: model and applications. *ACM T. Graphic.* **35**, 176 (2016).
35. Sun, X. et al. Reconfigurable logic-in-memory architectures based on a two-dimensional van der Waals heterostructure device. *Nat. Electron.* **5**, 752–760 (2022).
36. Chen, Y. et al. Ferroelectric-tuned van der Waals heterojunction with band alignment evolution. *Nat. Commun.* **12**, 4030 (2021).
37. Liu, K. et al. An optoelectronic synapse based on  $\alpha$ -In<sub>2</sub>Se<sub>3</sub> with controllable temporal dynamics for multimode and multiscale reservoir computing. *Nat. Electron.* **5**, 761–773 (2022).
38. Li, P. et al. Reconfigurable optoelectronic transistors for multimodal recognition. *Nat. Commun.* **15**, 3257 (2024).
39. Lowe, D. G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**, 91–110 (2004).
40. Harris, C. & Stephens, M. A combined corner and edge detector. In *Proc. Alvey Vision Conference* (ed. Taylor, C. J.) 147–151 (Alvey Vision Club, 1988).
41. Li, Z. Source data. *Figshare* <https://doi.org/10.6084/m9.figshare.26764015> (2024).

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant nos. 62422409, 62174152 and 62374159). The authors acknowledge financial support from Open Fund of State Key Laboratory of Infrared Physics (grant no. SITP-NLIST-YB-2024-04) and the Youth Innovation Promotion Association of the Chinese Academy of Sciences (grant no. 2020115). Thanks to J. Chang (Institute of Semiconductors, CAS) for technical guidance of etching process and thanks to J. Chen (Institute of Semiconductors, CAS) for technical guidance of spectrum characterization.

## Author contributions

Z. Li, Z. Lou and L.W. designed the research, Z. Li and L.W. wrote the paper, Z. Li, H.X., Y.Z., L. Liu and L. Li performed the experiments, Z. Li and L. Li performed the first-principles calculations and simulation. Z. Li, L. Liu and Z. Lou analysed the data. Z. Lou and L.W. revised the paper, Z. Lou and L.W. supervised the project. All authors contributed to research and reviewed 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/s41928-024-01261-6>.

**Correspondence and requests for materials** should be addressed to Zheng Lou or Lili Wang.

**Peer review information** *Nature Electronics* thanks Jongchan Park 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 2025