

Neuromorphic Vision Computing

A Dissertation

presented to

The Faculty of the School of Engineering and Applied Sciences

at the

University of Virginia

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in

Department of Electrical and Computer Engineering

By

Minseong Park

August 2023

APPROVAL SHEET

This Dissertation is submitted in partial fulfillment of the requirements for the degree

of:

Doctor of Philosophy

Authored by:

Minseong Park

This Dissertation has been read and approved by the examining committee:

Kyusang Lee, PhD
Advisor

Joe. C. Campbell, PhD
Committee Chair

Mona Zebarjadi, PhD
Committee Member

Miaomiao Zhang, PhD
Committee Member

Seongkook Heo, PhD
Committee Member

Accepted for the School of Engineering and Applied Science:



Jennifer L. West, School of Engineering and Applied Science
August 2023

ABSTRACT

Neuromorphic computing, referred to as brain-inspired computing for big-data processing and accelerating artificial intelligence (AI) computation, has received a significant boost from the emergence of memristors and associated computing algorithms over the past decade. Recent advancements in memristive systems have enabled the integration of sensing and computing on a chip, known as in-sensor computing, leveraging the memory and dynamic processing capabilities associated with synaptic long-term and short-term plasticity. Among the senses, vision plays a pivotal role in information processing, enabling remote sensing for navigation, learning, and communication. While current neuromorphic systems utilizing advanced memristors have primarily focused on two-dimensional (2D) vision applications, akin to human visual perception, three-dimensional (3D) vision is also vital for machines to tackle more complex tasks by obtaining additional depth information. In this dissertation, we present a comprehensive approach to neuromorphic vision computing that encompasses both 2D and 3D information processing in conjunction with artificial vision dynamics. We demonstrate one III-V photodiode and one nonvolatile memristor (1P1R) array capable of visual sensing, memory, and computing functions. This enables in-sensor computing protocols such as *in-situ* visual classification and encoding, referred to as 2D neuromorphic vision computing. We also introduce a bio-inspired 3D sensing technique utilizing nonvolatile memristors, known as the resistive time-of-flight (RTof) principle, enabling unprecedented 3D neuromorphic vision computing. we lastly achieve dynamic bio-inspired vision by integrating conventional high-electron-mobility transistors (HEMTs) with emerging 2D ferroelectric materials that emulate synaptic plasticity, potentially enabling mixed 2D/3D neuromorphic vision. This multidimensional approach to neuromorphic vision computing paves the way for empowering advanced computer vision and augmented reality applications.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Dr. Kyusang Lee, whose wholehearted instruction and guidance have enabled my PhD career. Thanks to his thorough support, insight, encouragement, and enthusiasm for research, I have been able to accomplish my research during these years and become an academic scientist working towards a better world. I extend my sincere gratitude to all my proposal and dissertation committee members, Dr. Andreas Beling, Dr. Nikhil Shukla, and Dr. Benton Calhoun, Dr. Joe C. Campbell, Dr. Mona Zebarjadi, Dr. Miaomiao Zhang, and Dr. Seongkook Heo. Their substantial feedback has been invaluable in developing and solidifying my research. I am also grateful to our group members: Dr. Sihwan Kim, Dr. Seunghyun Lee, Dr. Hee Sung Lee, Dr. Doeon Lee, Dr. Yongmin Baek, Byungjoon Bae, Taegeon Kim, Jeong Yong Yang, and Inbo Sim. Their collaboration and support have greatly contributed to my journey.

Lastly, I would like to express my heartfelt thanks to my family: my father Yoohyun Park, my mother Myungsook Heo, and my sister Jisun Park. Their unwavering support, understanding, and sacrifices throughout my PhD journey have been immeasurable. With their abundant love and encouragement, they have been a constant source of strength, guiding me through challenges and inspiring me to pursue my dreams. I am forever grateful for their presence in my life, as their unwavering love and belief in me have been the driving force behind my accomplishments.

Minseong Park

University of Virginia

August 2023

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES	viii
CHAPTER 1 Neuromorphic Computing	
1.1 Introduction	14
1.2 Biological Synaptic Plasticity	17
1.3 Memristors	18
1.4 Memristor-Based Hardware Accelerators	23
1.5 Conclusion	26
CHAPTER 2 Three-dimensional (3D) Neuromorphic Vision	
2.1 Time-of-Flight (ToF) Principle.....	27
2.2 GaN high-electron-mobility transistor (HEMT).....	28
2.3 GaN HEMT-based Laser Driver.....	29
2.4 Resistive ToF (RToF) Principle.....	35
2.5 3D RToF Imaging.....	38
2.6 Non-Line-of-Sight Detection Based on RToF Principle	42
2.7 Conclusion	48
CHAPTER 3 Two-dimensional (2D) Neuromorphic Vision	
3.1 Visible and Infrared Dual-Band Imaging	50
3.2 In-Sensor Computing.....	59
3.3 One-Photodiode One-Resistor (1P1R) Crossbar Array	61
3.4 1P1R-Based In-Sensor Computing.....	69
3.5 Conclusion	74

CHAPTER 4	Dynamic Bio-Inspired Vision	
4.1	GaN/ α -In ₂ Se ₃ Ferroelectric HEMTs	77
4.2	GaN/CuInP ₂ S ₆ Ferroelectric HEMTs	85
4.3	Synaptic GaN HEMT-Based Reservoir Computing	91
4.4	Artificial Oculomotor Dynamics	97
4.5	Conclusion	106
CHAPTER 5	Future Works	
5.1	Second-Order Spiking Neuromorphic Vision	108
5.2	Ultrasound RToF	109
REFERENCE	111

LIST OF FIGURES

Figure 1.1 Comparison of conventional Von Neumann architecture and emerging neuromorphic architecture ¹	15
Figure 1.2 Neuromorphic devices forecast by market segment ²	16
Figure 1.3 Emulating biological synaptic plasticity ³	17
Figure 1.4 Various non-volatile resistive switching mechanism and corresponding materials of memristors ⁴	19
Figure 1.5 Fabricated HfO ₂ two-terminal and non-volatile redox memristor. a Photograph of HfO ₂ memristor. b Current-voltage (<i>I-V</i>) characteristics HfO ₂ memristor. The round-trip voltage swing switches the memristor conductance ⁵	20
Figure 1.6 Comparison of non-volatile resistive switching mechanisms in terms of dimension, stochasticity, number of distinguishable states, switching energy, switching speed, endurance, and retention time ⁴	21
Figure 1.7 Realization of thousand states of single redox memristor. a Photograph of wafer-scale production of thousand state redox memristors. b <i>I-V</i> characteristics of 256 × 256 redox memristors, achieving thousand states. c Random telegraph noise (RTN) before and after ‘denoising’ process. 3	23
Figure 1.8 Implementation of fully hardware non-volatile and two-terminal memristor crossbar array. a Embedding of software artificial neural network (ANN) architecture into physical ANN building blocks. b Photograph of single crossbar tile. c Cross-sectional transmission electron microscopy image of the one-transistor one-resistor (1T1R) cell d Block diagram of in-memory computing architecture Detailed structure of e in-memory unit f 1T1R structure ⁶	24
Figure 1.9 Classification accuracies of hardware-implemented artificial intelligence (AI) accelerator for MNIST and CIFAR-10 datasets and other long short-term memory (LSTM) and restricted Boltzmann machine (RBM) machine learning architectures ⁶	25
Figure 2.1 Schematic of time-of-flight (ToF) principle-based three-dimensional (3D) depth acquisition. a Schematic of ToF procedure, calculating the time difference between emitter and receiver. b Detailed implementation of each ToF building block ⁷	27
Figure 2.2 Structures of the heterogeneously integrated devices. Schematic of ToF principle-based 3D depth acquisition. a 3D schematic illustration of the devices. The vertical-cavity surface-emitting laser (VCSEL) is integrated into the drain region of the high-electron-mobility transistor (HEMT). Scale bar: 100 μm. b 3D structure of the device (not scaled). c Circuit diagram of the heterogeneously integrated device. d Light-current-voltage (<i>L-I-V</i>) characteristics of the VCSELs at four different temperatures (0 C°, 25 C°, 40 C°, and 80 C° respectively, starting from the darkest curve). The green lines are voltage (left y-axis), and the red lines are optical power (right y-axis). The optical power is measured at 8 mA, and the turn-on voltage ranges from 1.5 V to 1.9 V.....	30
Figure 2.3 ToF characterization of GaN HEMT laser driver and VCSEL emitter. a Rise and fall time response of the AlGaIn/GaN HEMT and Si metal-oxide-semiconductor field-effect transistor (MOSFET, LND 150). The navy line is gate voltage (left y-axis), and the red and blue lines are drain current (right y-axis). Inset: close-up of the black-dotted area to compare rising responses at the rising interval (3.6 μs~4.0 μs). The two rise times are measured from 90% of pulse-on voltage and 10% of pulse-off voltage. b Linear results of ToF ranging. The minimum distance is 50 mm, and the maximum distance is 210 mm. 300 measurements are shown as orange dots, and the ideal ranging line is shown as a black-dotted line (Actual distance = Measured distance). The R-square linear fitting parameter is 0.97. c 3D ToF imaging of ‘UVA’ letters. The z-axis (depth) is an arbitrary unit; the yellow points signify the highest depth, and blue points signify the lowest depth. 32×32 grids per each letter are adopted.....	32
Figure 2.4 Demonstration of resistive ToF (RToF) principle. a Schematic illustration of biological STDP behavior and RToF. The rod cell and the avalanche photodiode (APD) receive the optical signals and transmit to the synapse and the memristor, respectively. Inset: scanning electron microscopy	

(SEM) image of memristor. Scalebar: 20 μm . b Schematic illustration of RToF system including classification. c Principle of RToF system. The ToF is calculated by subtracting the receiving time (t_r) from the transmitting time (t_t). The ToF ($\Delta t_1 < \Delta t_2 < \Delta t_3$) corresponds to the target distances ($d_1 < d_2 < d_3$). The shorter distance enables larger transitions of the conductance of the memristor, and the conductance is measured to range the distance.....	36
Figure 2.5 Schematic illustration of RToF experimental setup.	38
Figure 2.6 Experimental demonstration of RToF depth ranging. a Schematic illustration of RToF system. The optical signal received by the APD is converted to the electrical Rx signal that couples with the Tx signal to program the memristor. b Spike-timing-dependent plasticity (STDP) waveforms from the RToF ranging (40 cm, 25 cm and 10 cm object distances from left to right).	40
Figure 2.7 Target scene and experimental RToF 3D depth imaging (16 \times 8 pixels).....	41
Figure 2.8 RToF non-line-of-sight (NLOS) detection. a Schematic of neuromorphic NLOS detection system. The HfO ₂ memristor stores the NLOS information as conductance via STDP. b Illustration of experimental neuromorphic NLOS detection setup. The presynaptic spike is generated via pulse generator, branched to both memristor and laser. The postsynaptic spike is converted from the reflected optical signal via the avalanche photodiode and the following transimpedance amplifier (TIA). The two spikes sandwich the memristor and program the memristor only if a target object is scanned. c Schematic waveforms of presynaptic and postsynaptic spikes, reset/read and superpositioned pulse signals to memristor from top to bottom for neuromorphic NLOS detection. The postsynaptic spike 2 is generated from the light scattered by the target object and is coupled with the presynaptic spike. The superposed waveform programs the memristor with a superposition peak higher than V_{th} . Two read pulses are applied to the memristor to achieve G_{final} and G_{init} . The memristor conductance is initialized by the negative reset pulse.	43
Figure 2.9 Schematic illustration of NLOS situation. The laser beam with one angular position (10°, 20°, or 30°) reaches the wall, partially reflected to the laser (Reflection 1). The propagating beam is also reflected from the hidden object (p_1 , p_2 , or p_3) and then from the wall (Reflection 2). Both reflection signals are detected by the ToF sensor, and the laser steers the angle to another angular position for the next scanning.	45
Figure 2.10 Experimental demonstration of RToF-based NLOS detection. a Experimental results of the NLOS detection. The second peak is coupled to the electrical presynaptic spike to generate the STDP programming (top). Results without coupling for objects located at each p_1 , p_2 , and p_3 (bottom). b Experimental comparison of superposed spiking energy between the scanning positions. As the object is located at the non-scanning points, less coupling leads to less conductance switching of the memristor. The error bars indicate the standard deviation.	46
Figure 3.1 Visible (VIS) and infrared (IR) dual-band Ge/MoS ₂ multi-spectral photodetector. a Schematic of the device application. The selective dual-band detection enables vision for a harsh environment (e.g., fog, top) using VIS and infrared IR visions simultaneously and separately (bottom). b Schematic of the device structure. c X-ray photoelectron spectroscopy (XPS) measurement of the Ge/MoS ₂ heterojunction structure.	51
Figure 3.2 Design and operation principle of the Ge/MoS ₂ multi-spectral photodetector. a Cross-sectional transmission electron microscopy (TEM) image. b Energy band diagram in equilibrium.	53
Figure 3.3 Energy band diagrams under a near-equilibrium (near-photovoltaic mode) and b strong reverse bias (photoconductive mode).	54
Figure 3.4 I - V characteristics of Ge/MoS ₂ multi-spectral photodetector with respect to a wavelengths and incident power densities at b 406 nm and c 1550 nm.	55
Figure 3.5 Synaptic STP behavior of Ge/MoS ₂ multi-spectral photodetector. Temporal response of the device at various illumination wavelengths in the a near-photovoltaic and b photoconductive modes. The temporal responses for different wavelengths are offset for clarity. c Calculated rising and falling times under the near-photovoltaic and photoconductive modes at various wavelengths.	56

Figure 3.6 VIS/IR dual-band selective imaging. a Schematic of the dual-band selective imaging experiment. The double-side-patterned silicon target object exhibits angry and smile face images. Results of the dual-band imaging for transmission imaging (IR illumination) operating at the b near-photovoltaic and c photoconductive modes. Results of the dual-band imaging for reflection imaging (VIS illumination) operating at the d photoconductive and e near-photovoltaic modes.	58
Figure 3.7 Sensory-based computing. a Conventional computing architecture. b near-sensor and c in-sensor computing architectures. ⁸²	60
Figure 3.8 In-sensor computing system with an in-pixel direct computing functionality emulating human vision system. a Schematic illustration of human vision system from the retina to the brain, and its visual cognitive processing with in-retina image encoding process. b Schematic illustration of a process flow of the one-photodiode one-resistor (1P1R) in-sensor computing system, which emulates the human vision system.....	62
Figure 3.9 Design of single InGaAs 1P1R integrated device. a Schematic illustration of the fabricated InGaAs 1P1R device. b Optical image of the fabricated 1P1R focal plane array. Scale bar: 7 mm. c A optical microscope image of the fabricated 1P1R array. Dashed red (blue) box in shows a 1P1R pixel in the array. Scale bar: 400 μm . d Enlarged optical microscope image of the dashed blue box area in c . Scale bar: 50 μm	64
Figure 3.10 Working principle of single InGaAs 1P1R integrated device. a Equivalent circuit diagram of the 1P1R structure, and three key operations in the 1P1R device depending on applied bias voltage: i) memorization, ii) computation, and iii) erasing operations. b <i>I-V</i> characteristics of the single 1P1R integrated device under light illumination with an incident power density of 67 mW/cm^2 and wavelength of 532 nm.....	66
Figure 3.11 Characterization and image memorization process of the fabricated 1P1R crossbar array. a Circuit diagram for a pixel in the 16×16 1P1R crossbar array. The pixels are colored in yellow (blue) when the memristors are in a SET (RESET) state in (b-d). b-d Schematic illustration of image memorization b , readout c , and erasing d processes in the 1P1R crossbar array.....	68
Figure 3.12 Schematic illustration of multiply–accumulate operation for an image encoding process in a conventional memristor crossbar array and b 1P1R crossbar array. Corresponding matrix-vector multiplication is depicted with parameters of the input voltage (V_j), conductance of memristor (G_{ij}), and output current (I_j).	69
Figure 3.13 Image memorization, encoding and classification via an in-pixel neuromorphic computing. a Schematic illustration of an example of in-pixel image memorization, encoding, and classification process with a 5×5 1P1R array. At the initial state, conductance of memristors is G^0 . After image memorization, conductance of each memristor is indicated as G_{ij} . Once an image is memorized in the sensor, pre-trained weight voltages ($-1.3 \text{ V} < V_i < -0.5 \text{ V}$) are applied to the rows of the crossbar to perform a multiply–accumulate operation in the sensor for the encoding process. The encoded data are transferred to a post-ANN to classify the image. b Classification results from the memorized ‘4’ and ‘8’ digit images before and after training the ANN.	71
Figure 3.14 Application of in-sensor computing and memorization. a Schematic illustration of in-sensor computing mode. The pre-trained weight matrix is mapped to the memristor array, and $V_{\text{bias,low}}$ is sequentially applied to each row. The output currents of each column are accumulated (I_{out}). b Results of simulated classification accuracy. The device scheme with sequential biasing exhibits classification performance (up to 82%), while the accuracy of the device scheme with conventional simultaneous is up to 43%. Inference results for noised c ‘0’ and d ‘3’ images. Both software and device scheme identified the noised images correctly.	73
Figure 4.1 GaN/ α -In ₂ Se ₃ HEMTs (FeHEMTs). a Scanned piezoelectric force measurement (PFM) phase image after applying a bias of + 8 V and - 8 V, sequentially. b Amplitude and phase graph versus external tip bias measured on α -In ₂ Se ₃ /Si substrate. The atomic structure of α -In ₂ Se ₃ and pulse scheme for c program and d erase states. Band diagrams of fabricated device under e program and f	

erase states, respectively.....	79
Figure 4.2 Design of GaN/ α -In ₂ Se ₃ FeHEMTs. a Cross-section schematic of GaN/ α -In ₂ Se ₃ MOS-HEMT before and after the etching process. Optical microscopy image of the region where Raman spectroscopy analysis was conducted on α -In ₂ Se ₃ b before and c after the etching process (white dashed circle).	81
Figure 4.3 Electrical characterization of GaN/ α -In ₂ Se ₃ FeHEMTs. a Normalized bidirectional I_{DS} vs gate voltage (V_{GS}) transfer curve of GaN/ α -In ₂ Se ₃ MOS-HEMT at drain voltage (V_{DS}) = 0.1, 0.5, 1 V, respectively. An inset shows the linear scale of bidirectional curve of Figure 4.3a. b Normalized I_{DS} vs V_{DS} output curve of GaN/ α -In ₂ Se ₃ MOS-HEMT at V_{GS} range of 0 V ~ -10 V. c and d Minimum subthreshold swing (SS_{min}) and ΔV versus different V_{GS} sweep ranges. Parameters were extracted from the bidirectional I_{DS} vs V_{GS} transfer curve of GaN/ α -In ₂ Se ₃ MOS-HEMT.	82
Figure 4.4 Subthreshold swing (SS) and hysteresis characterizations of GaN/ α -In ₂ Se ₃ FeHEMTs. a Bidirectional I_{DS} vs V_{GS} transfer curve of GaN/ α -In ₂ Se ₃ MOS-HEMT at V_{DS} = 1 V, measured under a slow, normal, and fast sweep speed mode. b SS_{min} and c ΔV extracted from Fig. 4.3a at V_{end} = 0, 2 V under a slow, normal, and fast sweep speed.....	84
Figure 4.5 Material analysis of GaN/CuInP ₂ S ₆ (CIPS) FeHEMT. a Polarization-voltage (P - V) characterization of Ti/Au/CIPS/Ti/Au metal-ferroelectric-metal (MFM) capacitor under 50-kHz triangular pulse. Inset: positive-up negative-down (PUND) waveform. b Polarization switching endurance of CIPS membrane. Red and blue lines: endurance under downward and upward polarization, respectively. c PFM characterization of CIPS membrane. The hysteresis loops of both the amplitude and phase PFM verify the ferroelectricity of the CIPS membrane. d PFM imaging of CIPS membrane on metal (15 μ m \times 10 μ m horizontal scan). The bright region indicates the piezoelectric response of the CIPS membrane. e Cross-sectional transmission electron microscopy (TEM) image of GaN/CIPS FeHEMT. f Raman spectroscopy of GaN/CIPS FeHEMT, including GaN (left) and CIPS (right) peaks. The GaN peak is at 568 cm^{-1} , and the CIPS peaks include the multiple vibration peaks (150-400 cm^{-1}) and ionic peaks (100 cm^{-1} and 315 cm^{-1}). g Energy dispersive X-ray spectroscopy (EDS) characterization of cross-sectional GaN/CIPS FeHEMT. White dash line: CIPS region. Cu, In, S, and P are uniformly distributed in the CIPS region.	86
Figure 4.6 Working principle of GaN/CIPS FeHEMT. a Cross-sectional device structure of CIPS/GaN HEMT. b Band diagrams of program and erase states of GaN/CIPS FeHEMT. The IL in the figure indicates an oxygen-abundant GaN interlayer between Al ₂ O ₃ and AlGaN. The positive bias generates downward polarization, lowering the barrier height and increasing the transconductance of the GaN/CIPS FeHEMT. In contrast, the negative bias generates the upward polarization that increases the barrier height and decreases the transconductance of the GaN/CIPS FeHEMT.	88
Figure 4.7 Electrical characterization of GaN/CIPS FeHEMT. a I - V transfer curve of the fabricated CIPS/GaN HEMT. The double sweep switches the ferroelectric polarization of the CIPS, thus the hysteresis loop forms. b I - V output curve of the GaN/CIPS FeHEMT with an output current of approximately 2 mA. c Comparison of subthreshold swing (SS) between GaN/CIPS FeHEMT and GaN HEMT. With the CIPS membrane, the SS overcomes Boltzmann limitation (60 mV/dec).....	89
Figure 4.8 Short-term plasticity (STP) characterizations of GaN/ α -In ₂ Se ₃ FeHEMTs. a Excitatory postsynaptic current (EPSC) responses for different gate input widths from 1 ms to 500 ms. Each output current of the EPSC is collected from the drain electrode of GaN/ α -In ₂ Se ₃ MOS-HEMT while the input pulses are applied to the gate electrode. Cumulative EPSC responses for sequential pulses with a time interval of b 10 ms and c 100 μ s, and d with incremental amplitudes. e For various time intervals between two sequential pulses, ranging from 10 ms to 500 ms, the pair-pulse-facilitation (PPF) was calculated from 1 st and 2 nd peaks of each EPSC. f Endurance cycle for the program and erase operations. Responses current was measured after applying a program and erase pulses at V_{DS} = 1 V.	92
Figure 4.9 Schematic of reservoir computing (RC) based on GaN/ α -In ₂ Se ₃ FeHEMTs for vowel image classification (example of 'A'). 3 \times 5 'A', 'E', 'I', 'O', and 'U' datasets are prepared, and the pixel	

intensities are converted to voltages ('0' \rightarrow -2 V and '1' \rightarrow +2 V). The accumulated pulses generate corresponding EPSCs at $V_D=2$ V and $V_G=-6$ V. The readout EPSCs are normalized and fed into the pre-trained fully connected layer (3-5-8-5 nodes) for classification. All pulses are 50-ms pulses with 50% duty cycle. 93

Figure 4.10 Experimental EPSC responses of letter **a** 'A', **b** 'E', and **c** 'I'. The EPSC peaks correspond to the '1' pixel intensity in each letter. 95

Figure 4.11 RC-based vowel classification. **a** EPSC responses with respect to each pulse number. The current values corresponding to each pulse pattern are averaged. **b** Results of RC-based classification. For each train and test procedure, up to 97% and 95% accuracies are achieved, respectively. **c** Confusion matrix of vowel image classification. Inference results for noised (modifying one pixel) inputs, **d** 'E', **e** 'I', and **f** 'O'. 96

Figure 4.12 Schematic illustration of mimicking neuromuscular junction (NMJ). **a** Schematic illustration of oculomotor system. The stimulus transmitted through the optic nerve propagates towards the target muscle via the central nervous system (CNS), oculomotor nucleus/nerve, and NMJ. **b** Biological NMJ. The NMJ connects the terminal of the oculomotor nerve with muscle fibers. **c** Emulation of the NMJ for fast stimulus enhancement. The fast response is emulated by the programmable GaN/CIPS FeHEMT and integrate-fire unit (IFU). **d** Emulation of NMJ dynamics. The CNS generates an action potential on the efferent nerve that actuates the target muscle through the NMJ. The trained GaN/CIPS FeHEMT with an IFU allows for enhanced reflexes. **e** Schematic of GaN/CIPS FeHEMT structure. The polarization of the GaN/CIPS FeHEMT mimics synaptic plasticity and generates a programmable two-dimensional electron gas (2DEG) current that serves as an artificial EPSC. 98

Figure 4.13 Emulation of biological enhanced reflexes based on GaN/CIPS FeHEMT. **a** Schematic illustration of integrated GaN/CIPS FeHEMT with an IFU. Output spiking responses **b** before programming and **c** after programming (trained). After the training process, a 14% faster response (output spike) was achieved (24 ms) compared to the response without repeated training (28 ms). 100

Figure 4.14 Integration of microelectromechanical system (MEMS) mirror and GaN/CIPS FeHEMT. **a** Direct amplifier-less operation of actuator via GaN/CIPS FeHEMT. The beam steering emulates the adduction and abduction motions. The two-terminal microelectromechanical system (MEMS) mirror is connected to the drain of the GaN/CIPS FeHEMT. **b** Experimental results of beam steering. The displacement of the laser beam is dependent on the steering angle of the MEMS mirror, which actuation is driven by the I_{DS} of the GaN/CIPS FeHEMT. **c** Voltage-dependent steering angles. The steering angle is proportional to I_{DS} , modulated by V_{GS} 101

Figure 4.15 Schematic of in-situ mechanical object tracking. The tracking system additionally incorporates an ultrasound distance sensor integrated to the gate of the GaN/CIPS FeHEMT. The position of the car is captured by the ultrasound distance sensor and converted to the analog voltage that drives I_{DS} to steer the mirror. 102

Figure 4.16 Emulation of biological oculomotor dynamics. **a** Motion input and tracking output time-series data. The trained GaN/CIPS FeHEMT shows a higher I_{DS} (red line) that is more closely matched with the motion input signal for effective object tracking through beam steering. **b** Symmetric absolute percentage error (sAPE) plots of tracking output (I_{DS}) with respect to motion input. The trained GaN/CIPS FeHEMT drives larger initial I_{DS} , improving the cumulative sAPE of tracking. Inset: Zoom-in plot of cumulative sAPE. **c** Phase plots of motion input and tracking output. The trained system shows closer phase tracking of the motion input. 104

Figure 5.1 Artificial **a** synaptic and **b** neuronal devices for first-, second-, and third-order information processing. In this thesis, nonvolatile memristors have mainly been discussed to realize 2D/3D neuromorphic vision¹⁷⁹. 109

© Copyright by

Minseong Park

All rights reserved

2023

This dissertation is dedicated to my family

Chapter 1

Neuromorphic Computing

1.1 Introduction

The current front-end frameworks used in devices such as PCs, smartphones, and laptops are based on the Von Neumann architecture. This architecture is characterized by separate memory and computing units, which allows for efficient data processing. Recent advancements in artificial intelligence (AI), particularly the development of artificial neural networks (ANNs), have led to breakthroughs in various areas of information processing, including image recognition, language processing, and time-series analysis¹. In order to effectively simulate human brain functions, AI algorithms now require more complex ANN architectures known as deep neural networks (DNNs)^{1,2}. As a result, the dimensionality and volume of data have significantly increased, necessitating faster and more reliable information processing capabilities. To address these challenges, there is a need for more efficient physical chips that can handle the demands of AI algorithms^{3,4}. This requires further scaling down of solid-state layouts and fabrication processes, which has been limited by emerging issues related to degradation and yield in the nanometer-scale processes⁵.

Neuromorphic computing offers an alternative computing paradigm inspired by the biological brain^{6,7}. It specifically focuses on handling complex data structures and processing big data. Unlike the current mainstream Von Neumann architecture, neuromorphic computing utilizes multi-functional and analog devices that can function both as memory and computing units⁶⁻⁸. These devices can have multiple states beyond the traditional binary states of '0' and '1' in digital

computing, enabling a new in-memory computing architecture. By integrating both memory and computing functions within the same physical layer, the need for data transfer and processing is significantly reduced in terms of footprint and time/energy requirements.

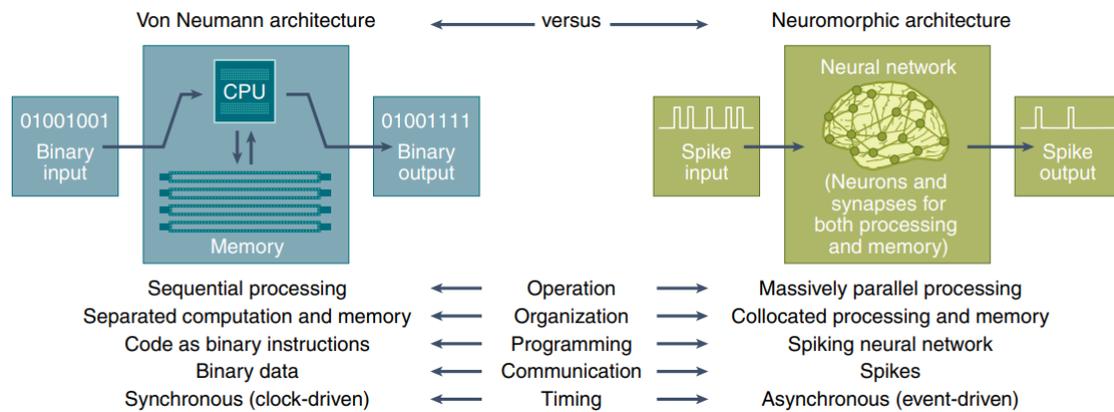


Figure 1.1 Comparison of conventional Von Neumann architecture and emerging neuromorphic architecture⁷.

Figure. 1.1 depicts the comparison of conventional Von Neumann architecture and emerging neuromorphic architecture⁷. Neuromorphic computing directly realizes the computing protocols in ANNs/DNNs, suited to massively parallel processing, memorization, and asynchronous (event-driven) computing similar to neuronal spiking dynamics in human^{9–11}. The market of the neuromorphic regime is expected to skyrocket, toward industrial, mobile, and automotive applications (Fig. 1.2)⁶. In 2034, neuromorphic computing for the automotive, industry, and mobile markets is anticipated to reach \$10B, \$5.0B, and \$2.5B, respectively⁶.

2024, 2029 and 2034 neuromorphic sensing & computing device sales forecasts - By market segment

(Source: Neuromorphic Sensing and Computing 2019 report, Yole Développement, 2019)

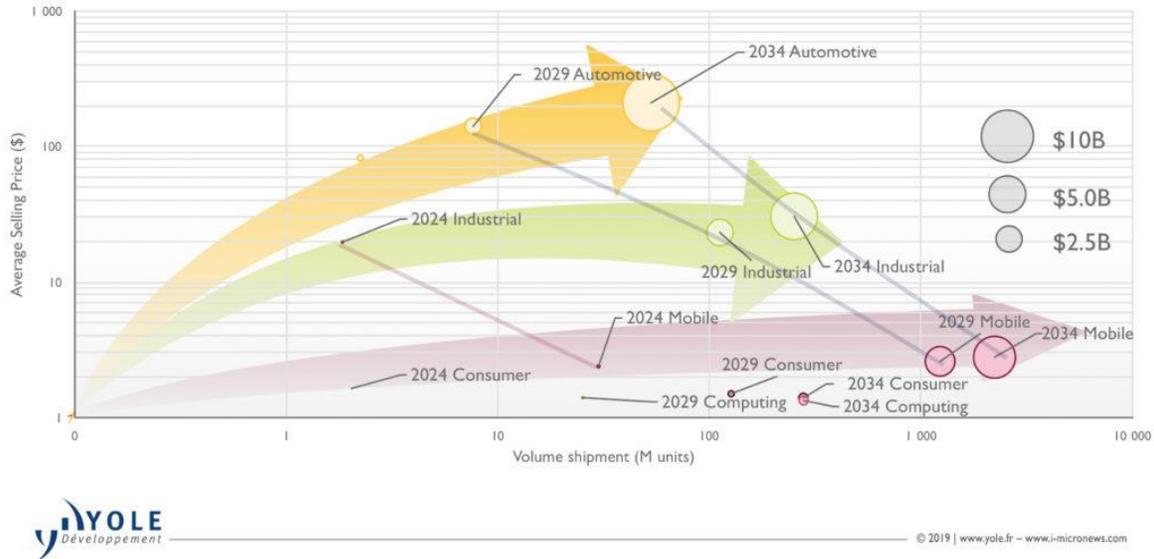


Figure 1.2 Neuromorphic devices forecast by market segment⁶.

While several neuromorphic computing protocols have been demonstrated, primarily utilizing silicon (Si) or resistive-switching materials (memristors), there is still ongoing research to establish a standard procedure for neuromorphic computing that can compete with the current Von Neumann computing chips^{12–15}. The field of neuromorphic computing is still in its early stages but holds promise for further advancements and standardization through the integration of additional bio-inspired computing paradigms and protocols. Achieving this goal requires interdisciplinary research involving various fields such as material engineering, computer engineering, mechanical engineering, electrical engineering, neurology, and biomedical engineering. By combining expertise from these diverse disciplines, the development of neuromorphic computing can be advanced and standardized for practical applications.

1.2 Biological Synaptic Plasticity

One prominent approach in the development of neuromorphic computing protocols involves the emulation of biological neuronal plasticity^{16–18}. Specifically, this entails replicating the in-memory functionality found in biological neurons, which is achieved through synapses—gaps between neurons responsible for signal transmission and memorization. There are three key types of plasticity in synapses that regulate the "weights" of synapses, acting as metrics to determine whether incoming spikes should be transmitted or not.

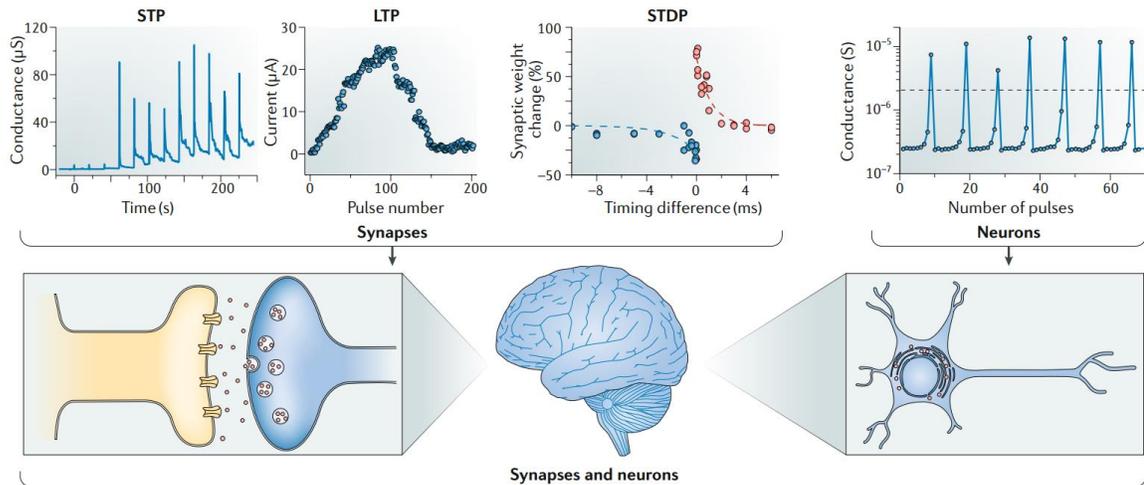


Figure 1.3 Emulating biological synaptic plasticity¹⁸.

Figure 1.3 illustrates the emulation of biological synaptic plasticity, which includes short-term plasticity (STP), long-term plasticity (LTP), and spike-timing-dependent plasticity (STDP). STP exhibits non-linear temporal dynamics without long-term persistence of input stimuli (spikes). It is commonly utilized in dynamic neuromorphic computing systems for processing temporal data, reservoir computing, and bio-plausible computing^{19–21}. On the other hand, LTP preserves the input stimulus for extended periods, making it a memory-building component in neuromorphic

computing architectures²². As shown in Fig. 1.3, the conductance of a synapse increases with the number of input pulses (spikes), enabling multiple states for analog computing¹⁸. STDP, another form of LTP, utilizes two subsequent spikes and relies on the timing difference between them. One type of STDP involves exponentially increasing the weights of synapses as the timing difference decreases. This STDP variant is particularly well-suited for processing the timing between two input signals compared to LTP.

In the case of neurons, the soma collects spikes from multiple synapses and generates a single spike at its terminal (dendrite) when the accumulated spike reaches or exceeds a threshold level²³. This process is referred to as the "integrate" process^{24,25}. Conversely, if the accumulated spike is below the threshold level, the dendrite does not generate a spike but instead relaxes, known as the "leaky" process. These processes are combined into a single plasticity mechanism called "leaky-integrate-and-fire," which emulates the activation functions in artificial neural networks (ANNs) and deep neural networks (DNNs)²⁴⁻²⁸.

The fascinating non-linear and temporal dynamics of STP find extensive application in temporal data processing, including prediction, classification, motion detection, and noise reduction. LTP, on the other hand, is often employed as in-memory computing units, especially in hardware AI accelerators based on memristor crossbars, due to its ability to retain information for longer periods, typically exceeding 1 hour. STDP has also been utilized as memory-building blocks in in-memory computing chips, with various input spike geometries being employed to compute timing differences among them, including square, triangle, and sinusoidal pulses.

1.3 Memristors

Recently, memristors, which are bio-inspired resistive-switching memories, have emerged as promising intelligent materials for interacting with the environment and enabling in-memory computing²⁹. These memristors have been utilized to mimic the synaptic responses of neurons^{12,24,30}. Through the manipulation of atomic placement in the active medium using external electric fields, memristors can be programmed to adjust their conductance²⁹.

Memristors can be categorized based on various metrics, including volatility, geometry, and materials¹⁸. Volatility is associated with memory capability and is divided into non-volatile and volatile memories^{31–33}. Non-volatile memristors emulate the characteristics of biological LTP/STDP, while volatile memristors replicate the behavior of STP¹⁸. The geometry of memristors can be classified as diode-/tunnel junction-type (two-terminal) or transistor-type (three-terminal)^{8,18}. Two-terminal memristors offer simpler fabrication processes and provide more compact platforms. However, three-terminal memristors, which incorporate an additional gate electrode, generally exhibit more reliable LTP, STP, and STDP plasticity control^{16,34–37}.

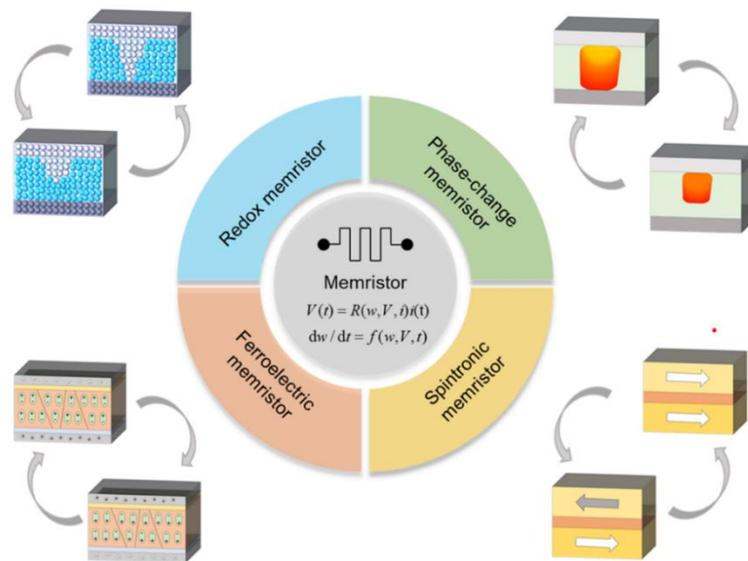


Figure 1.4 Various non-volatile resistive switching mechanisms and corresponding materials of

memristors¹³.

Figure 1.4 shows the four representative non-volatile resistive switching mechanisms in currently invented memristors¹³. Redox memristors are the type of memristor first found in 2008, generally incorporating the transition of atomic vacancies, defects, or ions in the memristor medium induced by electric fields²⁹. Oxygen vacancies and atomic defects typically feature non-volatility, while ionic transitions such as mobile metal cations (copper and silver) and organic compounds mainly exhibit volatility with partial nonvolatility³⁸. The materials for the oxygen vacancies-based transitions include various oxide thin films such as HfO_2 , Ta_2O_3 , SiO_2 , ZrO_2 , and some complex oxides, while the defects-based transitions generally occur at two-dimensional (2D) materials such as graphene, hexagonal BN (hBN), and transition metal dichalcogenides (TMDs)^{39–43}.

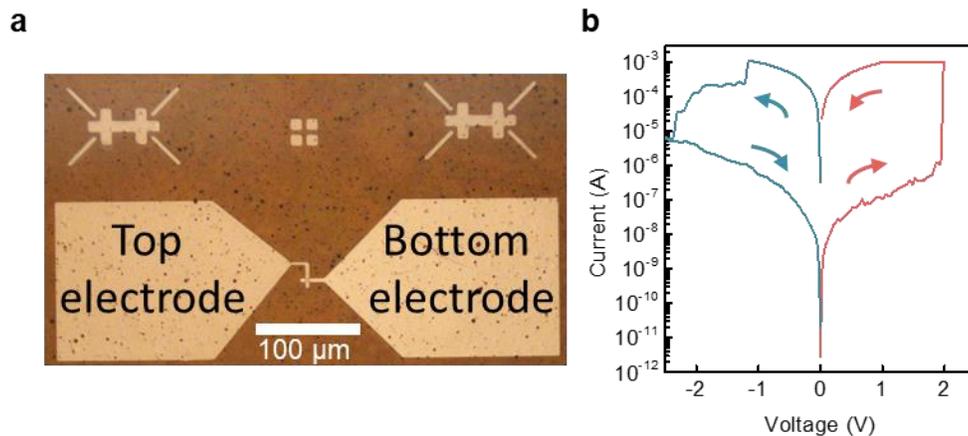


Figure 1.5 Fabricated HfO_2 two-terminal and non-volatile redox memristor. **a** Photograph of HfO_2 memristor. **b** Current-voltage (I - V) characteristics HfO_2 memristor. The round-trip voltage swing switches the memristor conductance⁴⁴.

Figure 1.5a shows a photograph of a fabricated HfO₂ memristor. Current-voltage (*I-V*) characteristics of the fabricated HfO₂ memristor are shown in Figure 1.5b⁴⁴. The set and reset voltages are approximately 2 V and -1.5 V, respectively. The maximum and minimum resistance at the -0.1 V read voltage is approximately 90 MΩ and 1 kΩ, respectively.⁴⁴

Ferroelectric memristors utilize the polarization switching of dipoles in the ferroelectric memristor medium, which can be induced by electric fields⁴⁵. The retention time of the polarization switching depends on the applied voltages, materials, and device geometry, which can serve as either STP, LTP, or hybrid STP/LTP (short-time relaxation with preserved baseline shift). Other mechanisms include phase-change and spintronic memristors induced by Joule heating and magnetic fields, respectively, also capable of non-volatile memories for neuromorphic computing⁴⁶⁻⁴⁸.

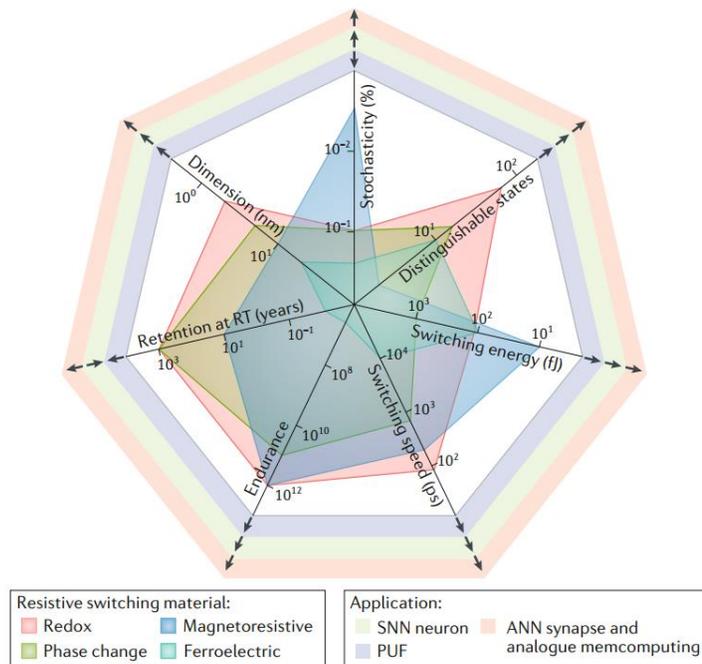


Figure 1.6 Comparison of non-volatile resistive switching mechanisms in terms of dimension,

stochasticity, number of distinguishable states, switching energy, switching speed, endurance, and retention time¹³.

Figure 1.6 compares the performance of memristors based on various mechanisms, considering factors such as dimension, stochasticity, number of distinguishable states, switching energy, switching speed, endurance, and retention time¹³. These mechanisms operate at the nanometer scale, with ferroelectric switching being achievable in ultrathin ferroelectric oxides. On the other hand, redox memristors exhibit the most reliable LTP compared to other mechanisms. Recently, they have achieved wafer-scale production (as shown in Fig. 1.7a) and can support more than a thousand states (as depicted in Fig. 1.7b)⁴⁹. To ensure stability and concentration of the oxygen vacancy distribution, a "denoising" process involving multiple programming and verification steps is employed (illustrated in Figure 1.7c)⁴⁹. Redox and ferroelectric memristors also demonstrate comparable switching speed and endurance when compared to other mechanisms. It is important to note that this thesis primarily focuses on fabricated redox and ferroelectric memristors, including both two-terminal and three-terminal structures, as well as STP/LTP characteristics.

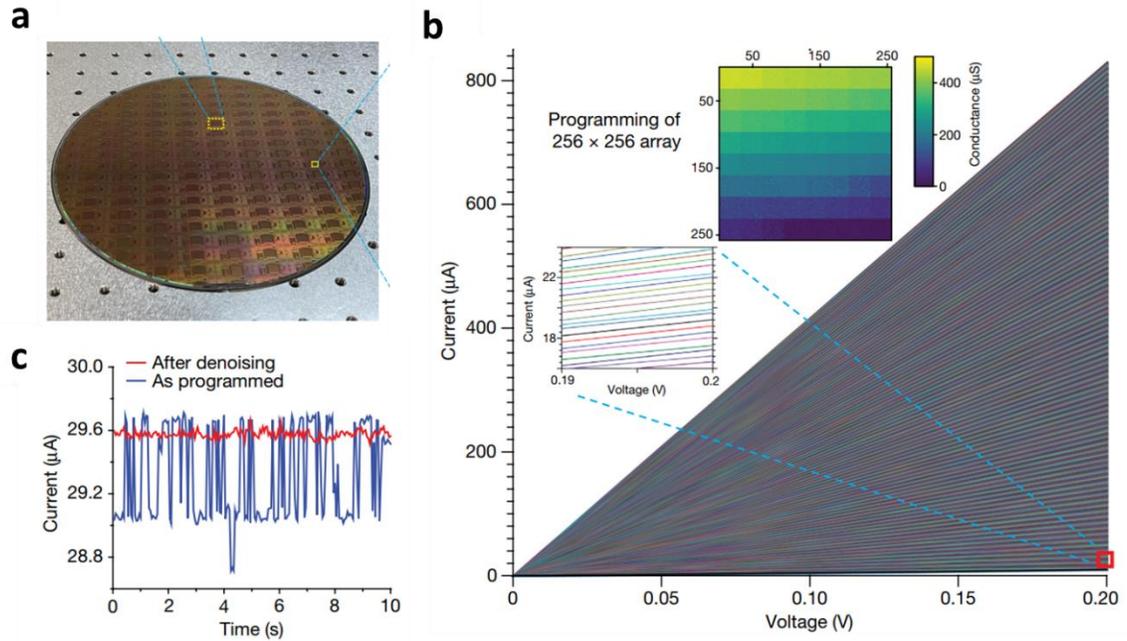


Figure 1.7 Realization of thousand states of single redox memristor. **a** Photograph of wafer-scale production of thousand state redox memristors. **b** I - V characteristics of 256×256 redox memristors, achieving thousand states. **c** Random telegraph noise (RTN) before and after ‘denoising’ process⁴⁹.

1.4 Memristor-Based Hardware Accelerators

Using an array of non-volatile memristors allows for the implementation of multiply-accumulate (MAC) operations by utilizing current inputs through Ohm's law^{11,50,51}. This approach finds wide applicability in numerous ANNs and DNNs applications, including classification, forecasting, and reinforcement learning^{19,22,52}. Among the various neuromorphic architectures based on memristors, synaptic spiking models such as Hodgkin-Huxley, Hebbian, and non-Hebbian learning algorithms offer high bio-fidelity, low-power consumption, and sparse learning capabilities²⁷. These models can be effectively utilized for processing various types of visual

information, including texts, diagrams, and images.

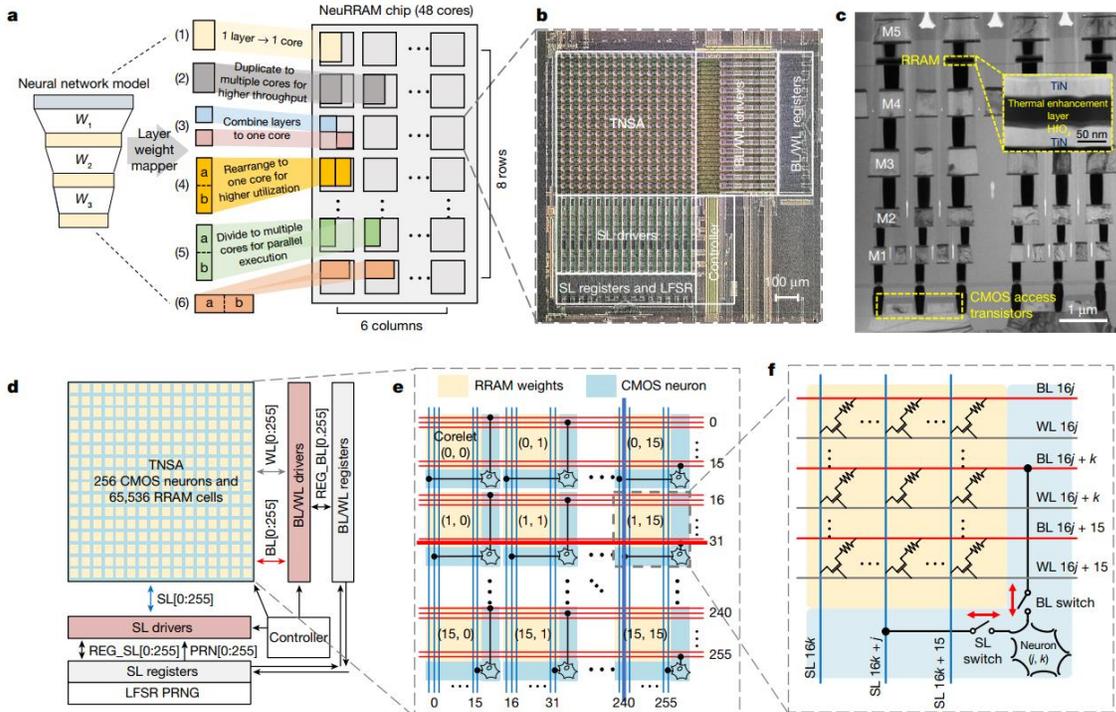


Figure 1.8 Implementation of fully hardware non-volatile and two-terminal memristor crossbar array. **a** Embedding of software artificial neural network (ANN) architecture into physical ANN building blocks. **b** Photograph of single crossbar tile. **c** Cross-sectional transmission electron microscopy image of the one-transistor one-resistor (1T1R) cell **d** Block diagram of in-memory computing architecture Detailed structure of **e** in-memory unit **f** 1T1R structure⁵³.

However, the main drawback of memristor-only crossbar is the leakage current driven by high-resistance-state memristor cells (OFF state), which degrades the computing accuracy and scalability of the crossbar⁵⁴. To alleviate this issue, selecting device is integrated into each memristor cell, referred to as one-transistor one-resistor (1T1R) cell⁵⁵. By integrating such switching transistors into each memristor cell, the transistor minimizes the leakage current of each

cell associated with the ultra-low OFF current of the transistors. Figure. 1.8a shows the recent implementation of a full hardware memristor 1T1R system for AI acceleration⁵³. The software ANN architecture is embedded into the physical tile of the chip. Each tile incorporates the 1T1R crossbar, SL/WL/BL drivers, and registers (Fig. 1.8b). The 1T1R is the integration of memristor and complementary metal-oxide-semiconductor (CMOS), achieving reliable and scalable production in CMOS foundry (Fig. 1.8c). The detailed circuit interconnection and 1T1R configuration (Fig. 1.8d) is shown in Fig. 1.8e, f⁵³.

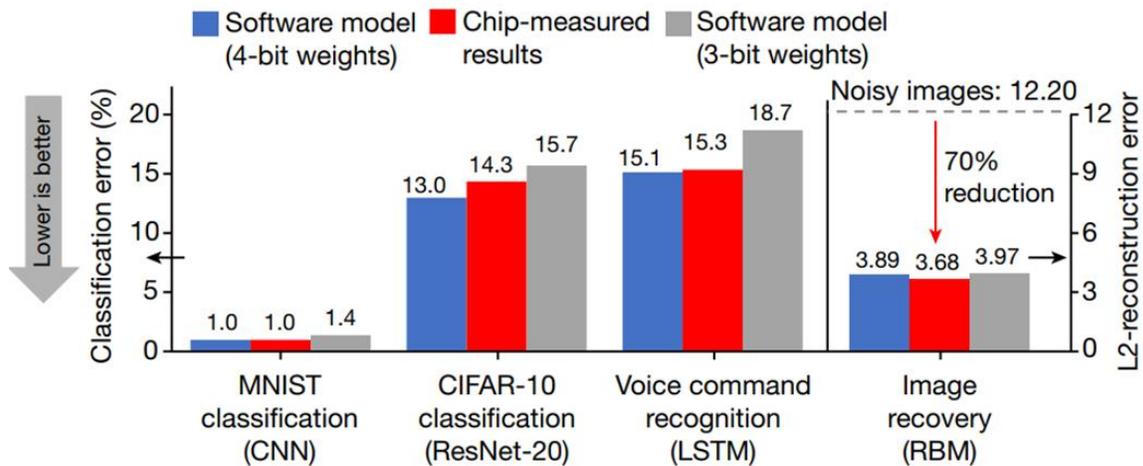


Figure 1.9 Classification accuracies of hardware-implemented artificial intelligence (AI) accelerator for MNIST and CIFAR-10 datasets and other long short-term memory (LSTM) and restricted Boltzmann machine (RBM) machine learning architectures⁵³.

Based on the system, the memristor-based AI accelerator achieved compatible classification accuracies for standard classification datasets such as MNIST and CIFAR-10 and other machine learning architectures, long short-term memory (LSTM), and restricted Boltzmann machine (RBM), as shown in Fig. 1.9⁵³. Unlike the classification via the pure software domain, these hardware-implemented architectures feature ideally one-shot computation time for the

inference process compared to the multiple computation steps required for conventional Von Neumann architecture⁵³.

1.5 Conclusion

Neuromorphic computing is an emerging alternative paradigm that aims to overcome the current limitations of the Von Neumann architecture, particularly in handling big data processing and high-dimensional data. The introduction of memristors in 2008 has significantly advanced the field of neuromorphic computing by enabling the practical implementation of large-scale 1T1R structures for AI accelerators^{29,55}. Memristors can be classified as non-volatile or volatile, as well as two-terminal or three-terminal, based on their material properties and geometries, respectively¹⁸. Extensive research has been conducted on different resistive switching mechanisms for non-volatile memristors, with a particular focus on redox and ferroelectric memristors, which offer reliable analog state generation.

In this thesis, I conducted a benchmark study on memristor crossbars, similar to the 1T1R structure, and fabricated redox and ferroelectric memristors for both two-dimensional (2D) and three-dimensional (3D) visual information processing. This novel approach is referred to as "neuromorphic vision computing," combining the principles of neuromorphic computing with visual data processing. By leveraging these advanced memristor technologies, the goal is to achieve efficient and effective processing of visual information in a neuromorphic computing framework.

Chapter 2

Three-dimensional (3D) Neuromorphic Vision

2.1 Time-of-Flight (ToF) Principle

Vision is the essential sense to obtain information from surroundings in navigation, object recognition, and complex environment exploration^{9,44}. In mammalian visual systems, the stimulus and responses in vision are driven by the STDP process through 10^{11} neurons linked with 10^{14} - 10^{15} synapses with extremely low energy consumption (1-10 fJ per synapse)⁵⁶. The STDP rule is an asymmetrical and temporal form of Hebbian learning⁵⁷ among neurons and allows sparse asynchronous spiking dynamics of synapses with low-power neurotransmission.

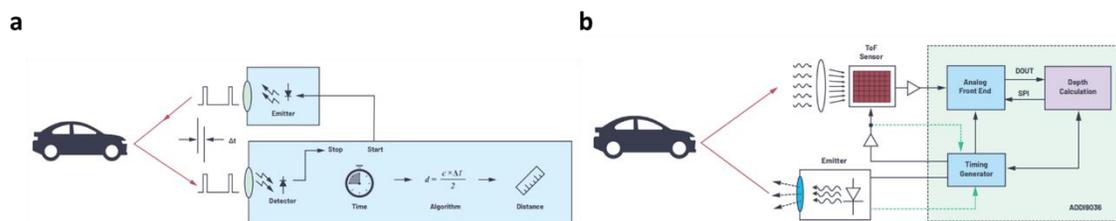


Figure 2.1 Schematic of time-of-flight (ToF) principle-based three-dimensional (3D) depth acquisition. **a** Schematic of ToF procedure, calculating the time difference between the emitter and receiver. **b** Detailed implementation of each ToF building block⁵⁸.

Time-of-flight (ToF) principle determines the 3D depth by calculating the time difference

between the transmitted and reflected flux as $d = \frac{1}{2}c \cdot \tau_{ToF}$, where d is the distance between a ToF sensor and target objects, c is the speed of light, and τ_{ToF} is the time difference between the emitter (Tx) and receiver (Rx) (Fig. 2.1)^{58,59}. In particular, the light-based ToF system is referred to as light-detection and ranging (LiDAR), which has widely been investigated to achieve high spatiotemporal resolution of 3D imaging (Fig. 2.1b)^{58,59}.

Nature includes abundant examples based on the ToF principle^{60–64}. Echolocation in mammals and birds calculates the time or phase difference between transmitting and receiving particles or waves in order to detect objects remotely⁶⁵. The ToF principle is generally engineered via optical flux in the sensor to achieve high accuracy and frame rate.⁶⁶ Therefore, ToF-based sensors have been widely employed for ranging and mapping technology, leveraging many advanced 3D applications in machine vision and the biomedical industry, etc.^{67–74} However, current ToF building blocks, such as emitters, receivers, and drivers, are still necessary to simplify their structural complexity, cost-ineffectiveness, and large form-factors, and to enhance their insufficient optical power, speed, and sensitivity^{74,75}.

2.2 GaN high-electron-mobility transistor (HEMT)

To alleviate these issues, hetero-integration of GaN-based high electron mobility transistors (HEMTs) and GaAs-based vertical cavity surface emitting lasers (VCSELs) on a single platform via a cold-welding process that provides the potential for high-resolution 3D real-world imaging was demonstrated⁵⁹. GaN-based HEMTs have been promising candidates for high power and high frequency applications due to the advantageous intrinsic material properties of GaN for device applications, including low parasitic capacitance, large breakdown endurance, and low on-resistance^{76–79}. AlGaIn/GaN heterostructures in GaN HEMTs allow the

formation of two-dimensional electron gas (2DEG) on the channel with a high density of the electron gas ($> 1 \times 10^{13} \text{ cm}^{-2}$) by piezoelectric polarization effects, resulting in high-speed operation and low node-to-node junction capacitance^{77,80}. Recently, GaN-based HEMTs have been employed to drive high-peak and narrow-pulse optoelectrical device applications, including ToF ranging, 3D imaging, and LiDAR^{75,77}. These applications require sub-nanosecond acquisitions; therefore, the switching performance of devices is critical. Switching performance is affected by gate resistance, capacitance, mobility, and inductance of driving transistors. GaN HEMTs show outstanding performance, with faster switching performance than that of silicon-based metal-oxide semiconductor field effect transistors (MOSFETs).

In addition, high-power emitters are the essential device component for efficient pulse-based optical communication. VCSELs are one of the most reliable and high-performing emitters. This is due to their surface-normal structure, large current densities, directive emission, narrow bandwidth, low divergence angle, and low power consumption.^{81,82} Although VCSEL chips are available with microscale dimensions, additional fabrication or packaging processes are required to be combined with other electronic/optoelectronic components. Monolithic fabrication of modules is advantageous for preventing device degradation related to material incompatibility and process reliability⁸³. Despite this, straightforward and inexpensive processes are still challenging.

2.3 GaN HEMT-based Laser Driver

Heterogeneous integration provides the potential for system-level applications without the issues of incompatibilities stemming from intrinsic material properties, complex spatial layouts, and geometric limitations.⁸⁴⁻⁸⁸ Here, I fabricated a heterogeneously integrated ToF sensor by

combining depletion-mode AlGaIn/GaN HEMTs and VCSELs via the cold-welding process.⁵⁹ The simple cold-welding process enabled quick and low-temperature bonding for microscopic structures⁸⁹. The electrical performance of the heterogeneously integrated devices was characterized by a KEYSIGHT B1500A Semiconductor Device Analyzer equipped with a waveform generator/fast measurement unit, a pulse generator (KEYSIGHT 33600A Series), and Digilent Analog Discovery 2. To measure pulse repetition rate (PRF) and photodiode current, an oscilloscope (KEYSIGHT DSO-X 3024T), and a current amplifier (Edmund 59-179) were employed to minimize external noise. A continuous measurement mode was adopted for the I - V characteristics. An additional power supplier (GS-1325-ND) was used for ToF ranging and imaging.

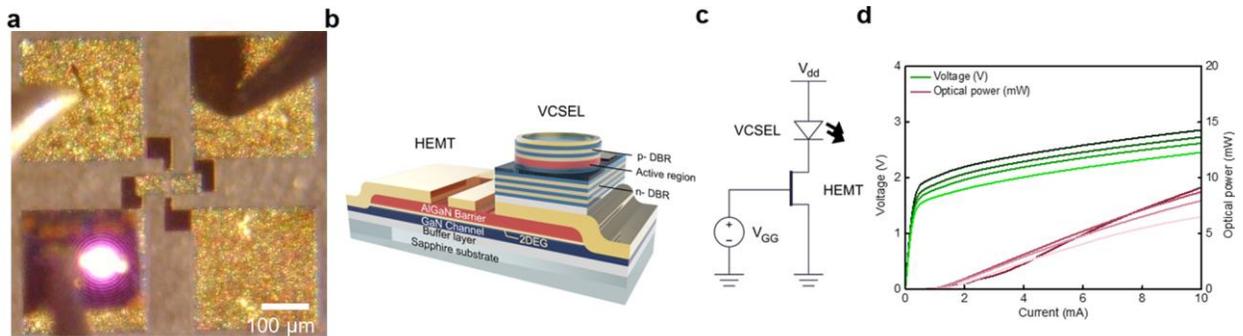


Figure 2.2. Structures of the heterogeneously integrated devices. Schematic of ToF principle-based 3D depth acquisition. **a** Optical microscopy (OM) image of the devices. The vertical-cavity surface-emitting laser (VCSEL) is integrated into the drain region of the high-electron-mobility transistor (HEMT). Scale bar: 100 μm. **b** 3D structure of the device (not scaled). **c** Circuit diagram of the heterogeneously integrated device. **d** Light-current-voltage (L - I - V) characteristics of the VCSELs at four different temperatures (0 C°, 25 C°, 40 C°, and 80 C° respectively, starting from the darkest curve). The green lines are voltage (left y-axis), and the red lines are optical power

(right y-axis). The optical power is measured at 8 mA, and the turn-on voltage ranges from 1.5 V to 1.9 V⁵⁹.

Figure 2.2a shows an optical microscopy (OM) image of the heterogeneously integrated device. A VCSEL chip was integrated on a HEMT using a thermally assisted cold-welding method⁹⁰. 200 MPa force was applied to the pre-cleaned surface between the two gold films at 270 °C. To apply uniform force over the interfaced area, a Teflon film was inserted between the sample and the pressing head. Figure 2.2b depicts the detailed device structure. The heterostructure of the AlGaIn/GaN HEMT was grown on a sapphire substrate by metal-organic chemical vapor deposition (MOCVD). The low-temperature (LT) and high resistance (HR) layers were grown as a buffer layer for the high-quality growth of the epitaxial GaN HEMT structure. For device fabrication, after the mesa etching, ohmic contacts were metalized and improved by annealing using rapid thermal annealing (RTA). Finally, gold pads for the source, drain, and gate were deposited using e-beam evaporation. The VCSEL structure was grown by MOCVD on a GaAs substrate. The epilayer structure is composed of three InGaAs–AlGaAs multiple-quantum-wells (MQWs) sandwiched between a 38-pair n-type and 21-pair p-type Al_{0.90}Ga_{0.10}As–Al_{0.05}Ga_{0.95}As distributed-Bragg-reflector (DBR) layers with a p-type Al_{0.98}Ga_{0.02}As layer (30 nm thickness) above the MQWs for oxidation. An oxidation technique was used to define a circular current-confined area of 10 μm in diameter. The 10-μm aperture size has been employed to maintain the consistency of the fabrication and to achieve the optimum number of cavity modes and gain for our applications. On the backside of the GaAs substrate, the gold film was deposited by e-beam evaporation, which serves as a bottom electrode.

The VCSEL chip was integrated on the drain side of the GaN HEMT to prevent

fluctuations related to turn-on stages from affecting the source voltage of the GaN HEMT, as shown in Fig. 2.2c. The light-current-voltage ($L-I-V$) characteristics of the VCSELs are shown in Fig. 2.2d. The threshold currents were nearly constant at various temperatures. Thermal effects slightly diminished optical power as the operating temperature increases. The fabricated phase-based ToF module can operate reliably up to 85 C° , which is similar to the operating temperature of commercialized ToF modules⁹¹. The current level was determined by device endurance and power conversion efficiency; thus, less than 8 mA current was applied to the heterogeneously integrated device.

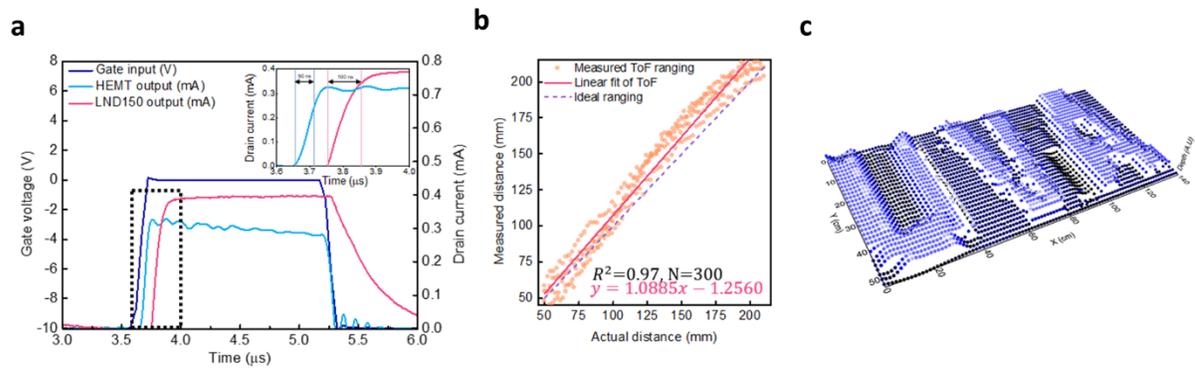


Figure 2.3 ToF characterization of GaN HEMT laser driver and VCSEL emitter. **a** Rise and fall time response of the AlGaIn/GaN HEMT and Si metal-oxide-semiconductor field-effect transistor (MOSFET, LND 150). The navy line is gate voltage (left y-axis), and the red and blue lines are drain current (right y-axis). Inset: close-up of the black-dotted area to compare rising responses at the rising interval ($3.6\ \mu\text{s}\sim 4.0\ \mu\text{s}$). The two rise times are measured from 90% of pulse-on voltage and 10% of pulse-off voltage. **b** Linear results of ToF ranging. The minimum distance is 50 mm, and the maximum distance is 210 mm. 300 measurements are shown as orange dots, and the ideal ranging line is shown as a black-dotted line (Actual distance = Measured distance). The R-square linear fitting parameter is 0.97. **c** 3D ToF imaging of ‘UVA’ letters. The z-axis (depth) is an

arbitrary unit; the yellow points signify the highest depth, and blue points signify the lowest depth. 32×32 grids per each letter are adopted⁵⁹.

Phase shift based ToF was employed for this study. To record the exact phase difference from both transmitter and receiver signal, OPT3101 (Texas Instruments) analog front-end (AFE) was used. The transmitter square-shaped waves are based on continuous-wave modulation, which was then used to calculate the phase difference between transmitted and received signal. The calibration is performed to eliminate the crosstalk caused by the background signal for precise measurement. To calibrate the module, the software development kit (SDK) code was run at various temperatures to measure both crosstalk and phase. Then, the phase at various ambient light brightness was measured; the values gained and extracted from the calibration were used to the correct SDK method which relates these values to register writes.

Figure 2.3a confirms advantageous fast switching of the fabricated GaN HEMTs. The rise time of the HEMTs was 50 ns, which was over twice as fast as that of the commercialized Si-MOSFETs (Microchip Technology, LND 150). Furthermore, the fall time of HEMTs was remarkably faster than that of MOSFETs as well. The superior rise and fall times of the fabricated GaN-based HEMT were comparable to those of commercial GaN-based HEMTs⁹². The difference in turn-on delay times of both HEMTs and Si-MOSFETs originated from the difference in threshold voltages; GaN HEMTs with -7 V and Si-MOSFETs with -1.5 V. The short rise and fall times supported the performance of GaN HEMTs as fast switching devices to drive the current for a ToF sensor.

The characterization of ToF measurement is shown in Fig. 2.3b. The device exhibited consistent trends with the transition of the distance between the object and the VCSEL. The

distribution of the 300 measurements was interpreted as a linear regression model with a reliable R-square value (0.97). Furthermore, the heterogeneously integrated device enabled ToF-based 3D imaging, as shown in Fig. 2.3c. ‘UVA’ letters were used for this 3D depth imaging. The edges of each letter were recognizable, which supports the possibility of using a demonstrated ToF sensor for object and material recognition applications⁹³⁻⁹⁶.

Direct ToF ranging is calculated as $d = \frac{1}{2}c \cdot \tau_{ToF}$, where d is the distance between a sensor and object, c is the speed of light, and τ_{ToF} is the time difference between transmitter and receiver signals. The τ_{ToF} can be derived from time-to-digital converters (TDCs), but can also be calculated by the phase difference between the transmitter and receiver signals as $\varphi = 2\pi f\tau_{ToF}$ and $d = \frac{c}{4\pi f}\varphi$, where f is modulation frequency, and φ is a phase difference between illumination and reflection. The phase difference is calculated from the intensity of received signal at four different points such that $\varphi = \arctan\left(\frac{A_1-A_3}{A_2-A_4}\right)$, where A_1, A_2, A_3 and A_4 are the measurement at four different phases (0 deg, 90 deg, 180 deg and 270 deg each). The phase shift based ToF model was employed in this study due to its high accuracy in the range of millimeters with high surface reflectivity. The noise at the edge of the square waves was mitigated by damping the parasitic circuits to relax the resonance behavior. The mismatch between the linear fit and the ideal line was possibly caused by an inertial fluctuation of the setup and a lack of further precise calibration without consideration of humidity and ambient light conditions. These factors are mostly related to the SNR and elicit measurement deviation as $\sigma_{\text{distance}} = \frac{d}{\text{SNR}}$. Normally, one millimeter accuracy corresponds to 6.6 picoseconds pulses, which is challenging for a conventional silicon-based photodiode⁶⁶. In this experiment, the range was 50 mm to 210 mm in order to maintain the minimum distance between the microprobes and object, and the maximum

distance was limited by the optical power of the emitter and detectivity of the photodetector. Therefore, further ranging is possible by maximizing the SNR using a high optical power emitter with a high sensitivity receiver. Sub-watt optical power of the VCSELs can be easily achieved by employing multiple VCSEL-based emitters⁹⁷⁻⁹⁹. Moreover, the VCSEL beamwidth is small and directional. Thus, combination with additional beam controllers, such as rotary motors⁷⁵, micromirrors^{100,101}, optical phased array (OPA)¹⁰²⁻¹⁰⁴, and liquid crystal (LC) based beam steerers¹⁰⁵, is expected to pave the way for future angle-dependent ranging and high-resolution 3D imaging⁵⁹.

2.4 Resistive ToF (RToF) Principle

In the biological visual system, the rod cell receives optical signals and converts them to electrical spikes towards the synapse that lies between the rod cell (presynaptic neuron) and the bipolar neuron (postsynaptic neuron)^{44,106}. If the timing is small enough, the synapse releases an acetylcholine neurotransmitter that transmits signals to the post neuron. Here, I have mimicked the biological synaptic responses in visual systems by implementing STDP in the memristors. The basic principle of STDP in the biological system is the transition of synaptic weights hinging on the time difference between presynaptic and postsynaptic spikes.

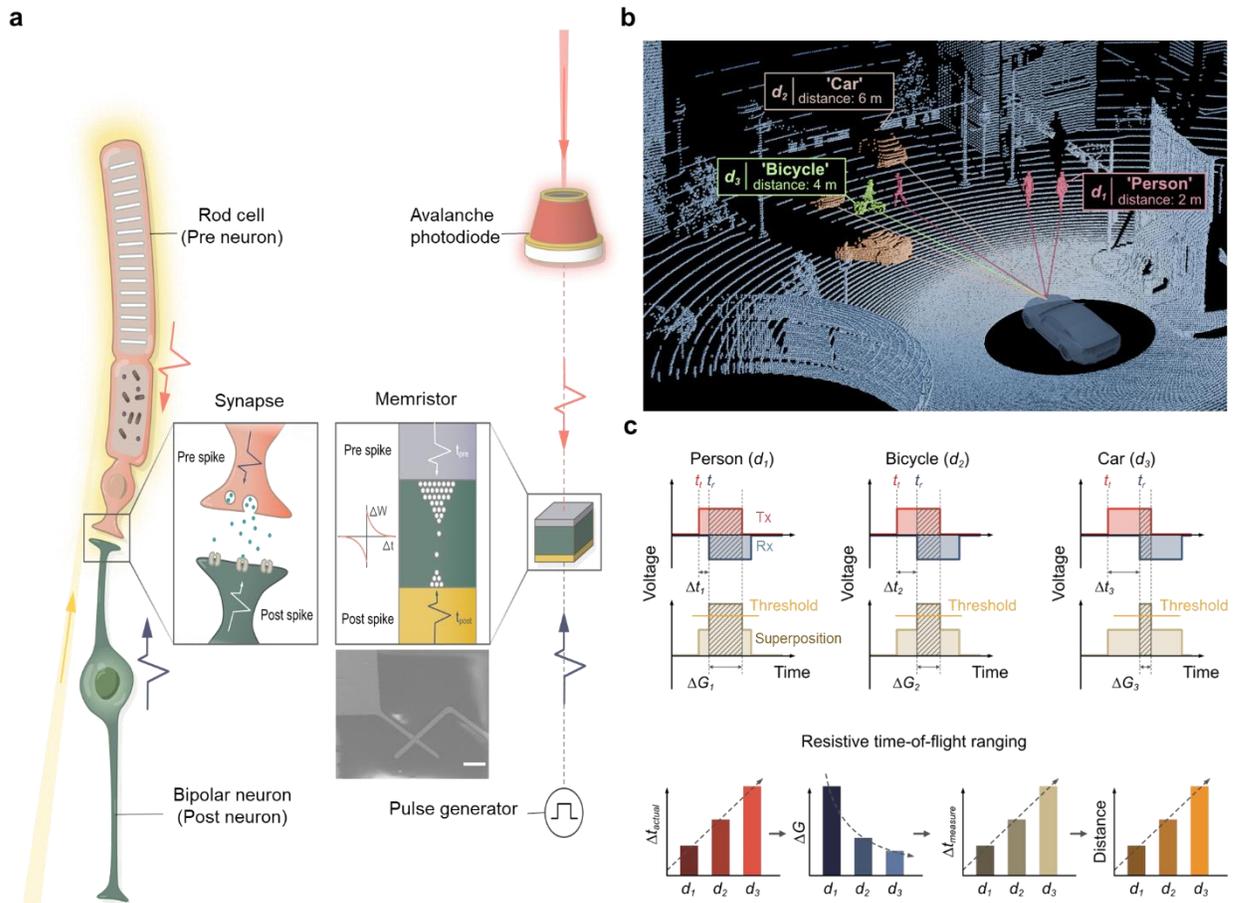


Figure 2.4 Demonstration of resistive ToF (RToF) principle. **a** Schematic illustration of biological STDP behavior and RToF. The rod cell and the avalanche photodiode (APD) receive the optical signals and transmit to the synapse and the memristor, respectively. Inset: scanning electron microscopy (SEM) image of memristor. Scalebar: 20 μm . **b** Schematic illustration of RToF system including classification. **c** Principle of RToF system. The ToF is calculated by subtracting the receiving time (t_r) from the transmitting time (t_t). The ToF ($\Delta t_1 < \Delta t_2 < \Delta t_3$) corresponds to the target distances ($d_1 < d_2 < d_3$). The shorter distance enables larger transitions of the conductance of the memristor, and the conductance is measured to range the distance⁴⁴.

Figure 2.4a shows the structural and functional analogy between the biological and artificial synaptic systems. The STDP learning rule applies to memristors since a conductive bridge by oxygen vacancies in the active medium is formed depending on the overlapped period between the presynaptic (to the top electrode of the memristor) and postsynaptic spikes (to the bottom electrode of the memristor). Therefore, the temporal difference between spikes can be obtained by reading the programmed conductance of the memristors¹⁰⁷. The inset shows the scanning electron microscope (SEM) image of the fabricated memristors. I demonstrated a new principle of RToF that converts memristor conductance to depth information. Figure 2.4b shows a schematic illustration of the RToF principle and classification. Combined with Fig. 2.4c, the signals from the transmitter and receiver exhibit time-difference-based responses similarly to the biological time-variant sandwiching spikes. Therefore, the polarities of the presynaptic and postsynaptic spikes are opposite. Following the ToF principle, the presynaptic spike in the RToF sensor is generated by an electrically transmitted signal, which also operates a picosecond laser. The postsynaptic spike in the system is transmitted by the current pulse of a received signal via APDs. The APDs adopted in the ToF sensing systems^{108,109} as an optical receiver provide high internal multiplication gain, low dark current, and low excess noise¹¹⁰. The larger spike time differences between the transmitted and received signals ($\Delta t_1 < \Delta t_2 < \Delta t_3$) lead to longer programming times and higher peak amplitudes in the memristor by the superposition of the semi-synchronous positive presynaptic peak and the negative postsynaptic peak. The longer and higher peak signals impart larger transitions of conductance ($\Delta G_1 > \Delta G_2 > \Delta G_3$) to the memristor. As a result, the conductance transitions of the memristor are converted to time-difference information following the pre-calibrated exponential-fitted relationship between the conductance and the programming time, which leads to the final depth information ($d_1 < d_2 < d_3$).

2.5 3D RToF Imaging

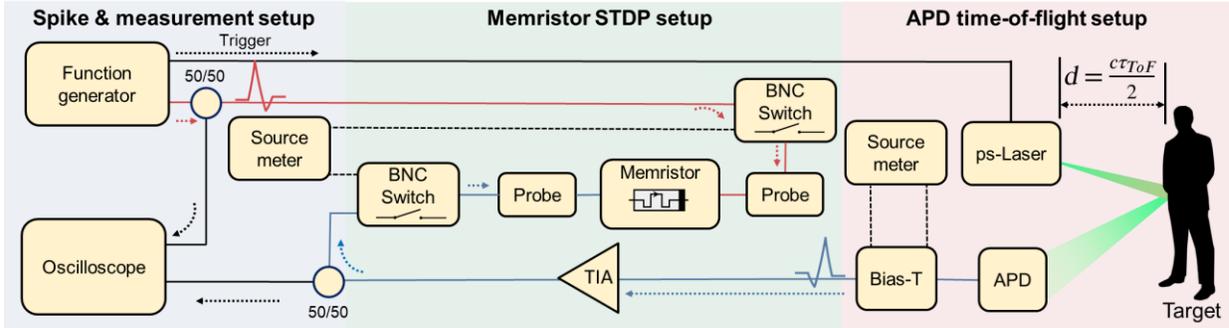


Figure 2.5 Schematic illustration of RToF experimental setup⁴⁴.

The schematic RToF experimental setup is shown in Fig. 2.5. The HfO_2 memristors were fabricated on a 25- μm polyimide substrate. To prevent the expansion of the polyimide during the post process, a pre-annealing of the polyimide substrate was performed at 200 °C for 30 min on a hot plate, followed by deposition of a 100-nm Al_2O_3 buffer layer via atomic layer deposition (ALD) at 200 °C. A 3/25 nm thick Ti/Pt bottom contact was deposited and patterned by using e-beam evaporation and a photolithography process, respectively. A 5-nm HfO_2 switching layer was deposited using ALD at 200 °C. A 50-nm Ta top metal contact was deposited via DC magnetron sputtering, followed by photolithography patterning. The surface area of the fabricated HfO_2 memristor was $5 \times 5 \mu\text{m}^2$. The $\text{Al}_{0.7}\text{InAsSb}$ APDs were grown on a n-type GaSb substrate ($1.9 \times 10^{17} \text{cm}^{-3}$) with the following epitaxial structure: a 300-nm n-type ($2 \times 10^{18} \text{cm}^{-3}$) GaSb buffer layer, a 300-nm n-type ($2 \times 10^{18} \text{cm}^{-3}$) $\text{Al}_{0.7}\text{InAsSb}$ contact layer, a 1500-nm $\text{Al}_{0.7}\text{InAsSb}$ unintentionally doped (UID) layer, and a 100-nm p-type ($2 \times 10^{18} \text{cm}^{-3}$) $\text{Al}_{0.7}\text{InAsSb}$ contact layer. The epitaxial growth was capped with a 50-nm p-type ($1 \times 10^{19} \text{cm}^{-3}$) GaSb layer. Be and Te were used as the p-

type and n-type dopants, respectively. The top-illuminated PIN APDs were defined using standard photolithography techniques. The mesas were etched into the n-contact layer with a citric/phosphoric acid solution prior to a 12-nm Ti and 85-nm Au contact deposition via electron beam evaporation.

Electrical performance of the memristors was characterized using a KEYSIGHT B1500A Semiconductor Device Analyzer equipped with a waveform generator/fast measurement unit and a pulse generator (KEYSIGHT 33600A Series). To measure resistive STDP and APD current, an oscilloscope (KEYSIGHT DSO-X 3024T), and two current amplifiers (Edmund 59-179 and EOC DLPCA-200) were employed. A Hamamatsu S9073 APD and a picosecond laser (Alphas PLDD-50M) were employed for RToF ranging measurement with 100 V reverse bias applied to the APD and low-noise 10^3 gain of the TIA. The physical channel lengths of the presynaptic and postsynaptic spikes were matched to minimize the parasitic time delay. The peak programming voltage was limited to 2 V to protect the device from exceeding the transition ranging of the resistance. The remainder time delay generated by parasitic capacitance and inductance of the external cables and wire connections was compensated empirically by modulating the pulse timing via MATLAB. For the one-dimensional (1D) RToF measurement, a commercially retro-reflective (metal coated, 0.9 reflectivity) object was adopted. In the dark room condition, 150 V reverse bias was applied to the APD to maximize the multiplicatively reflected optical signals. For the 1D RToF measurement, the picosecond laser was set to 203 mW and 750 mW peak power for short and long range, respectively. 0.5 V amplitude and 20-ns pulse width are employed for the presynaptic spikes. Each position was measured 5 times.

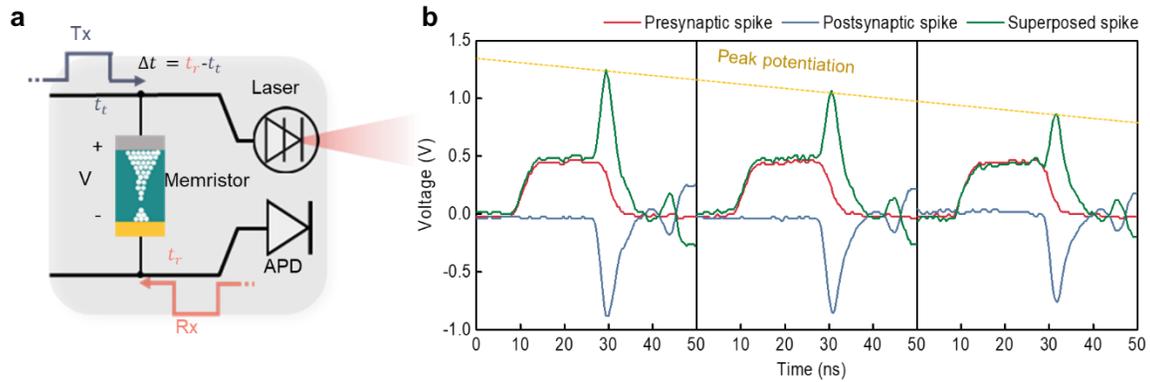


Figure 2.6 Experimental demonstration of RToF depth ranging. **a** Schematic illustration of RToF system. The optical signal received by the APD is converted to the electrical Rx signal that couples with the Tx signal to program the memristor. **b** Spike-timing-dependent plasticity (STDP) waveforms from the RToF ranging (40 cm, 25 cm and 10 cm object distances from left to right)⁴⁴.

Figure 2.6a shows schematic illustration of the RToF device. The postsynaptic spike is originated from the photogenerated current of the APD usually in the range of 1 μ A to 0.1 mA and is converted to the amplified voltage signal via a transimpedance amplifier (TIA) approximately 1 V to 2 V to program the memristors. The RToF waveforms depending on the object locations are shown in Fig. 2.6b. The peak of the superposed programming signal was low (high) when the time difference between the presynaptic spike and the postsynaptic spike was large (small), which leads to weak (strong) coupling. The analog transitions of the memristor conductance follow the analog programming via the pulse train response generated by overlapped pre- and post-synaptic spikes, which verifies the depth memory functionality of the RToF system.

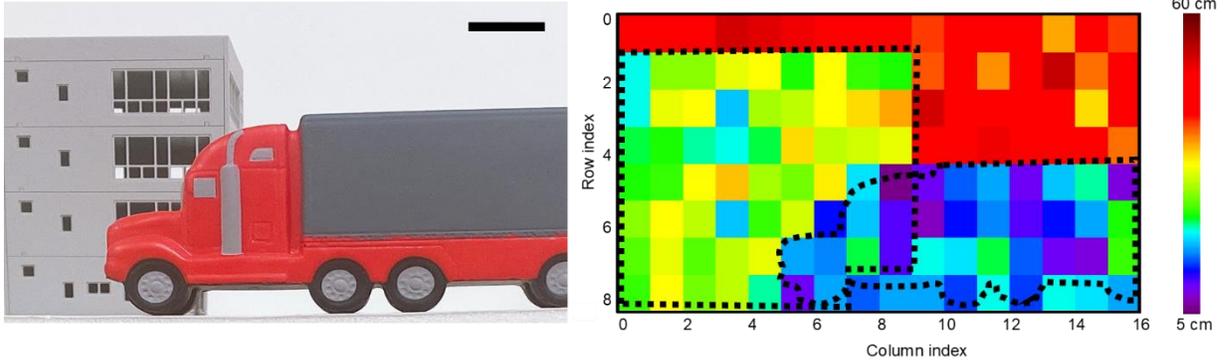


Figure 2.7 Target scene and experimental RToF 3D depth imaging (16×8 pixels)⁴⁴.

Using the RToF system, I demonstrated a RToF 3D LiDAR scan. As shown in Fig. 2.7, I emulated a street scene that include a car, building, and wall miniature. I measured distances of a miniaturized truck at 15 cm, a scaled building (EatingBiting DIY 1:160 Outland Models) at 25 cm, and a wall at 45 cm. The target scene is shifted for the imaging to enable a point-by-point LiDAR scanning method (16×8 pixels), and a scattering film are attached to the ranging spot. Due to the scattering effect that represents the realistic 3D ranging environment, different bias voltages are applied to the APD (15 cm for -140 V, 25cm for -150 V, and 45 cm for -155 V). By employing series of single-point ranging, I observed the RToF-based imaging capability.

The analog-domain depth ranging and storing capabilities of the RToF exhibit comparable power consumption and small footprint compared to the conventional TDC architectures and enable two high degrees of freedom for further designing a 3D sensing system. First, the system can be electrically and optically modulated by tuning the spike schemes and gains from the APD and the TIA. Both memristive STDP and RToF behaviors are dependent on amplitude, decay, and rise times of presynaptic and postsynaptic spikes. Additionally, the gain modulation relaxes a target object’s reflectivity-dependent amplitude transition. In a practical LiDAR application, the

intensity of the received optical signals depends on the distance from the objects as well as the reflectivity of the objects. The detection of low reflectivity objects can be improved through an optimized voltage window via the gain and spike modulation. Furthermore, multiple echoes from multiple objects within a single ranging shot are separable when the distance between the individual objects are far enough. In this case, only one postsynaptic spike (from the closer object) is coupled to the presynaptic spike. Second, a high geometrical degree of freedom can be achieved including small footprint and CMOS compatibility. Previous memristor studies have demonstrated CMOS-compatible architectures, small-footprint fabrication, and low power consumption, since the memristor medium is also widely used for high-k dielectric layers in standard silicon CMOS process^{111,112}. Furthermore, sub-nanoseconds switching speed of memristors has been reported^{113,114}, indicating further improvement in the resolution of the RToF sensor (15 cm in this work) towards a millimeter ranging system. Ferroelectric materials such as Ag/BaTiO₃/Nb:SrTiO₃¹¹⁵ and LiTaO₃¹¹⁶, and scalable magnetoelectric spin-orbit devices⁴⁸ are promising candidates due to their faster switching speed (more than one order) compared to the resistive HfO₂ layer⁴⁴.

2.6 Non-Line-of-Sight (NLOS) Detection Based on RToF

Principle

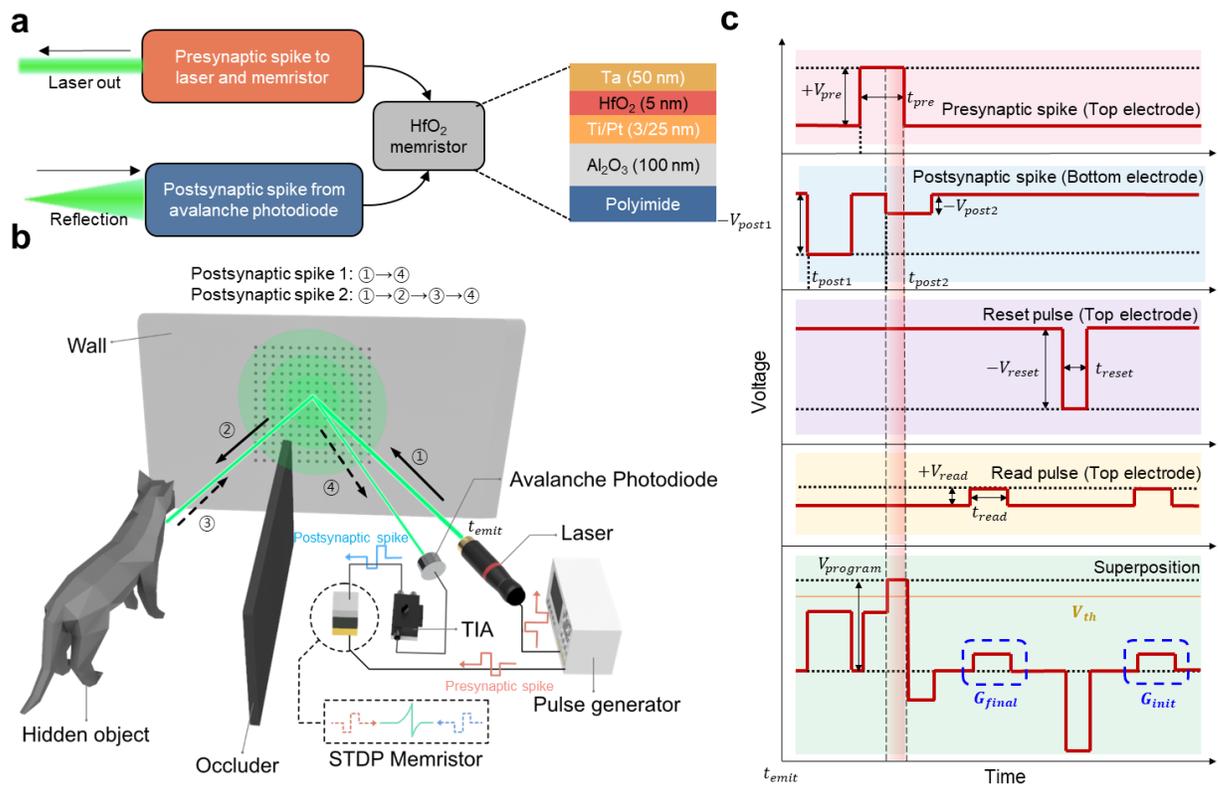


Figure 2.8 RToF non-line-of-sight (NLOS) detection. **a** Schematic of neuromorphic NLOS detection system. The HfO_2 memristor stores the NLOS information as conductance via STDP. **b** Illustration of experimental neuromorphic NLOS detection setup. The presynaptic spike is generated via pulse generator, branched to both memristor and laser. The postsynaptic spike is converted from the reflected optical signal via the avalanche photodiode and the following transimpedance amplifier (TIA). The two spikes sandwich the memristor and program the memristor only if a target object is scanned. **c** Schematic waveforms of presynaptic and postsynaptic spikes, reset/read and super-positioned pulse signals to memristor from top to bottom for neuromorphic NLOS detection. The postsynaptic spike 2 is generated from the light scattered by the target object and is coupled with the presynaptic spike. The superposed waveform programs the memristor with a superposition peak higher than V_{th} . Two read pulses are applied to the memristor to achieve G_{final} and G_{init} . The memristor conductance is initialized

by the negative reset pulse¹¹⁷.

Non-line-of-sight (NLOS) ranging is a promising 3D imaging application based on the ToF principle, which detects indirect optical signals for visualizing hidden objects^{67,117}. To demonstrate a RToF system for NLOS sensing, I integrated the HfO₂ memristors with an APD and a laser light source as shown in Fig. 2.8a. Figure 2.8b illustrates the NLOS detection process using the RToF system. An electrical signal from a pulse generator was applied to both the memristor and a picosecond laser, which served as the presynaptic spike for the system (Alphas PLDD-50M, 750 mW peak power). The emitted laser pulse was reflected off the target object and received by an APD, which converted the signal into an electric current (Hamamatsu S9073). To detect NLOS objects, the emitted laser pulse first hits a reflective wall and partially scatters. The partially scattered rays then reach the hidden object. Subsequently, both the rays reflected from the wall and hidden object reach the APD sequentially with respect to the distances of the ray propagations. The resulting signals are usually weak because of multiple reflections. Thus, the observed current signal is amplified and converted into a voltage signal (postsynaptic spike) using a tunable transimpedance amplifier (Edmund 59-179). In the RToF system, the input signal generated by the superposition of the presynaptic and postsynaptic spikes is used to program the depth information as the resistance of the memristor connected to the avalanche photodiode.

Figure 2.8c shows the waveform details of each spike and pulse in the RToF system. The observed information is converted into a postsynaptic spike that generates a superposed spike (V_{program}) capable of programming the connected memristor upon exceeding a threshold voltage (V_{th}). Presynaptic and postsynaptic spikes were applied to the top and bottom electrodes of the memristor, respectively. Therefore, the programming signal applied to the memristor is the

potential difference between the positive presynaptic and negative postsynaptic spikes. The first postsynaptic spike originated from the rays that were directly reflected from the wall. Thus, the first postsynaptic spike did not contain information from the target object and was not coupled to the presynaptic spike. The first read pulse ($V_{\text{read}} = 0.1 \text{ V}$ and $t_{\text{read}} = 10 \mu\text{s}$), after the neuromorphic ToF learning process, provides a programmed conductance (G_{final}). The programmed conductance is initialized by applying a reset pulse ($V_{\text{reset}} = -1.5 \text{ V}$ and $t_{\text{reset}} = 10 \mu\text{s}$) and then confirmed by the second read pulse (G_{initial}). The conductance transition is calculated as $\Delta G = \frac{(G_{\text{final}} - G_{\text{init}})}{G_{\text{init}}} \times 100(\%)$. All the pulses were applied to the top electrode of the memristor, except for the postsynaptic spike.

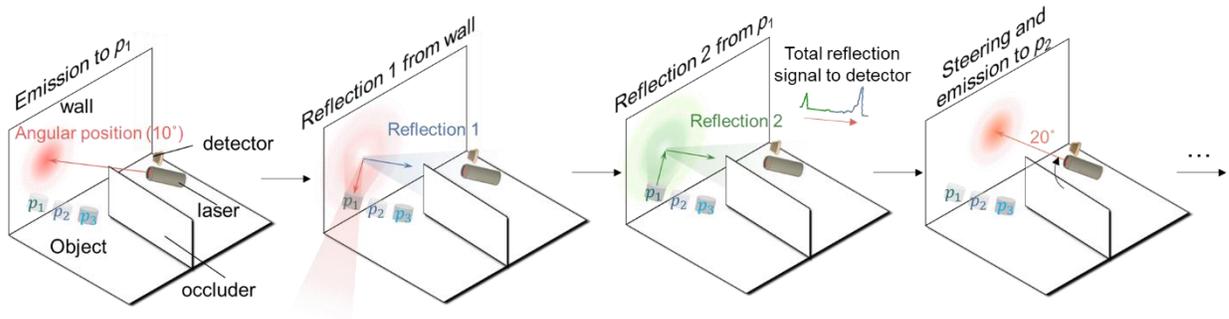


Figure 2.9 Schematic illustration of NLOS situation. The laser beam with one angular position (10° , 20° , or 30°) reaches the wall, partially reflected to the laser (Reflection 1). The propagating beam is also reflected from the hidden object (p_1 , p_2 , or p_3) and then from the wall (Reflection 2). Both reflection signals are detected by the ToF sensor, and the laser steers the angle to another angular position for the next scanning¹¹⁷.

The RToF system is used for NLOS detection using a time gated STDP learning rule. This was achieved by overlapping the reference electrical signal and spike signal received after multiple

reflections from the hidden target object and scattering wall, as shown in Fig. 2.9. The APD received two optical signals (Reflection 1 and Reflection 2), originates from the wall and the target object, respectively. The time gated RToF system allowed only the signal reflected from the object to be coupled with the presynaptic spike to program the memristor. This was achieved by applying a presynaptic spike after the first strong reflection signal was received without forming highly superposed peaks. Thus, the presynaptic spike was coupled only with the signal reflected from the object when the hidden object was in the scanning region. As expected, no second postsynaptic spike was generated unless the object existed in the scanning region.

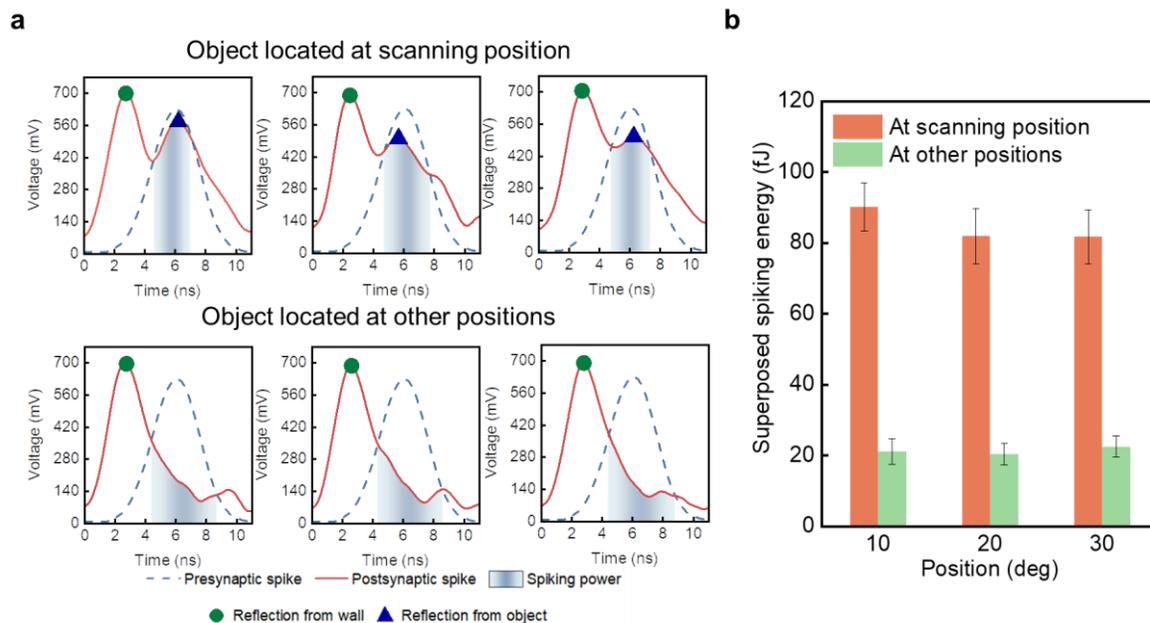


Figure 2.10 Experimental demonstration of RToF-based NLOS detection. **a** Experimental results of the NLOS detection. The second peak is coupled to the electrical presynaptic spike to generate the STDP programming (top). Results without coupling for objects located at each p_1 , p_2 , and p_3 (bottom). **b** Experimental comparison of superposed spiking energy between the scanning positions. As the object is located at the non-scanning points, less coupling leads to less

conductance switching of the memristor. The error bars indicate the standard deviation¹¹⁷.

I conducted an experimental demonstration of the RToF-based NLOS detection, as illustrated in Fig. 2.10a. The target object was positioned 48 cm from the center of the wall and rotated at an angular position of 10° to 30°. The laser source and receiver were co-rotated horizontally to detect the target objects located at different positions (10°, 20°, and 30° from the wall). Electrical input with an amplitude of 0.7 V and a pulse width of 30 ns was utilized as presynaptic spikes. The RToF system rotates to select one positions (10°, 20°, or 30°), and the target object is detected when the scanning direction matches that of the object. The second peak in Fig. 2.10a represents the signal reflected from the NLOS object, which was coupled to the memristor via the STDP learning rule in the neuromorphic ToF system. The first postsynaptic spike was time gated with an asynchronous postsynaptic spike to prevent programming through direct reflection from the wall. Therefore, only the NLOS signal (second peak) is coupled with the presynaptic spike. When the NLOS optical scan is applied to an object located at the target angle, a superimposed spiking signal is applied to the memristor for programming, which represents the NLOS detection, as shown in Fig. 2.10b. The superposed spiking power was calculated by squared voltage signals, multiplied by the average value of high resistance state (HRS) memristor of 7.287 k Ω , and integrated by time.

Combined with this technique, the localization in the two angular bases in the polar coordinate can be carried out by a 2-axis steering system such as double axis MEMS mirrors^{67,118}. Further practical implementation is also viable by using APD array that can scan multiple spots in a single laser shot so that the entire scanning time can significantly be reduced¹¹⁹. Integration with multiple focusing lenses and an intensity-thresholding components including comparator,

amplifier, and filter circuits¹²⁰ can alleviate the target reflectivity and accurate detection timing issues, respectively. Moreover, digital and analog co-designed integrated circuits can offer practical solutions for addressing distance-related attenuation by using automatic gain modulation through digital logics for different distances and employing the signal regulation of the postsynaptic spike through an analog regulation circuit¹²¹.

The distance-dependent attenuation issue in the proposed system, similarly in LiDAR systems, is influenced by the ambient environment, target reflectivity, and distance. One effective way to address the ambient light effect is to use a near-infrared (NIR) light source, which can propagate robustly through the air⁴⁴. Neuromorphic ToF ranging and NLOS detection can be further improved by minimizing the parasitic parameters through system-on-chip structures and by using faster switching artificial synapses, such as ferroelectric^{115,122,123}, spin-torque^{48,124}, and phase-change devices¹²⁵⁻¹²⁷.

2.7 Conclusion

In conclusion, I have invented a neuron-like RToF method for 3D sensing and NLOS detection. A memristor has been employed to store the ranging information programmed by superposed pre- and post-synaptic spikes. The RToF system attained up to 55 cm ranging in a single mode, and further ranging was achieved by controlling the spiking scheme combined with the gain modulation of an APD and a TIA. To prove the feasibility of RToF sensing, I have demonstrated the proof-of concept 3D imaging. In the case of RToF-based NLOSE detection, the coupling between presynaptic and postsynaptic spikes, electrical turn-on signals, and multiple-reflected optical signals enabled the NLOS detection via RToF. The presynaptic and second postsynaptic spikes were coupled when the target object was in the same direction as the scanning

direction, resulting in high programming power for the memristors. The proposed neuromorphic NLOS detection features the processing of detection information in the memristor medium, thereby minimizing the need for additional signal processing algorithms and circuitry required for 3D ranging. I believe that the proposed RToF principle can provide a simplified and miniaturized 3D vision computing platform with potential application in various engineering fields, such as automotive navigation¹²⁸⁻¹³⁰, machine learning¹³¹⁻¹³⁴, and biomedical engineering¹³⁵⁻¹³⁷.

Chapter 3

Two-dimensional (2D) Neuromorphic Vision

3.1 Visible and Infrared Dual-Band Imaging

The combination of visible (VIS) and infrared (IR) photodetectors is a promising technique for use in various applications such as computer vision^{138,139}, biomedical engineering¹⁴⁰, and tactical vision^{141,142}. Such multi-spectral image fusion can provide a more accurate understanding of the surrounding information compared to data acquisition by a single spectral band¹⁴³. Along with the VIS spectrum detection, the imaging of IR spectrum near 1550 nm is advantageous due to the high atmospheric transmission¹⁴⁴ and eye safety concerns. Furthermore, IR spectrum imaging, attracts considerable interest for night vision¹⁴⁵ and 3D ranging¹⁰⁹ etc. However, the integration of individual VIS and IR photodetectors for multi-spectral detection is limited by structural complexity. The conventional side-by-side configuration of independent VIS and IR photodetectors leads to large pixel size and optical crosstalk between photodetectors. Furthermore, related external circuits require large footprint and cost for switching capabilities. To address such challenges, vertical-stacked and multi-spectral photodetectors have been proposed^{146–152}. The multi-spectral functionality has been realized via broadband and selective photodetectors. Broadband photodetectors exhibit photonic responsivity to both VIS and NIR spectra simultaneously and can hardly resolve the VIS and NIR information^{146–148}. Photodetectors based on organic heterojunctions¹⁵⁰, quantum dot nanomaterials¹⁴⁹, and semiconductor multiple junctions¹⁵² obtain selective band information from broadband signal. However, the structures are based on a multi-junction structure including back-to-back-connected diodes with three

terminals that requires dual bias polarity operations for selective detection capabilities, resulting in complex fabrication processes and non-conventional back plane and readout circuit designs.

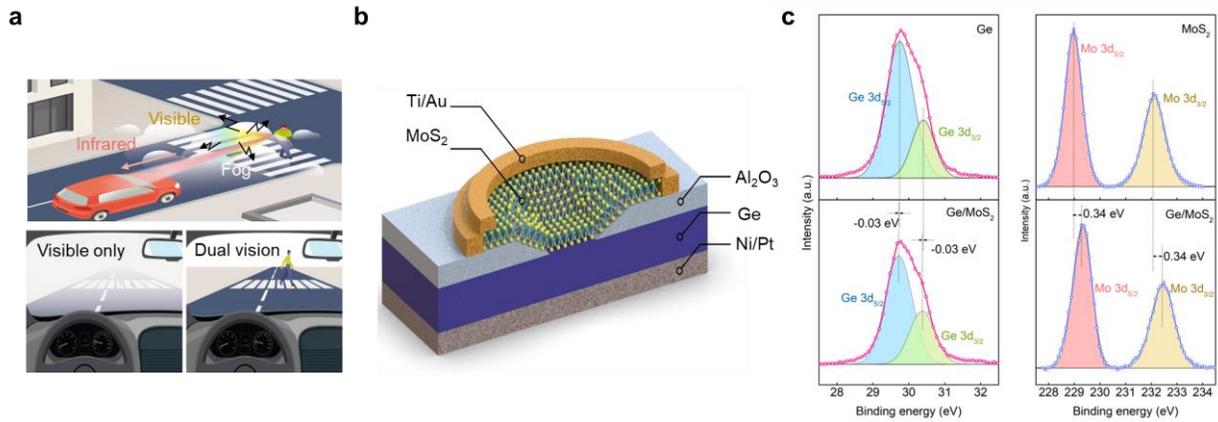


Figure 3.1 Visible (VIS) and infrared (IR) dual-band Ge/MoS₂ multi-spectral photodetector. **a** Schematic of the device application. The selective dual-band detection enables vision for a harsh environment (e.g., fog, top) using VIS and infrared IR visions simultaneously and separately (bottom). **b** Schematic of the device structure. **c** X-ray photoelectron spectroscopy (XPS) measurement of the Ge/MoS₂ heterojunction structure¹⁴².

The selective detection capability can be applied to vision under adverse weather conditions such as night, fog, and dust (Fig. 3.1a). The vision information obtained by the different wavelength spectra achieves advanced ambient information systems. I fabricated a p-Ge/n-MoS₂ photodetector with bias-dependent wavelength selection capability (Fig. 3.1b). A 20 nm oxide (Al₂O₃) layer was deposited on a p-Ge substrate ($5 \times 10^{17} \text{ cm}^{-3}$) at 200 °C by atomic layer deposition, to isolate the metal contact of MoS₂. To control the junction area, a circular shape with a diameter of 6 μm was etched in 1% diluted hydrofluoric acid (HF) in the Al₂O₃ layer until the Ge surface was completely exposed, resulting in a sloped sidewall at the Ge/Al₂O₃ boundary¹⁵³.

To ensure air stability, the Ge substrate was passivated by a solution of HCl:H₂O in a ratio of 1:1 for 1 min. Multi-layer MoS₂ flakes were mechanically exfoliated from bulk MoS₂ crystals (Graphene Supermarket, USA) and then transferred onto the exposed Ge surface. Ti/Au (5/60 nm) was deposited as an Ohmic metal of MoS₂ by electron-beam evaporation using a standard lift-off technique. A 5/60-nm Ni/Pt p-contact was deposited on the back side of the p-Ge substrate by electron-beam evaporation. The device was annealed for 1 h at 150 °C in vacuum using a rapid thermal processor to form a reliable junction at the Ge–MoS₂ interface.

Figure 3.1c shows X-ray photoelectron spectroscopy (XPS) results of the heterostructure device. The high-resolution spectra of the C 1s, Mo 3d_{3/2}, Mo 3d_{5/2}, and Ge 3d peaks were collected in steps of 0.1 eV. The measurement of the band levels before and after the heterojunction reveals the band shift indicating the energy barriers of conduction (ΔE_{Ge}) and valence (ΔE_{MoS_2}) bands of approximately 0.03 and 0.34 eV, respectively¹⁵⁴. The apparent valence band maximum (VBM) locations are 0.14 eV for Ge and 0.89 eV for MoS₂ via ultraviolet photoelectron spectroscopy (UPS), and additional electron energy loss spectroscopy (EELS) results also verify the bandgaps. The heterojunction was also verified by cross-sectional transmission electron microscopy (TEM) images, as shown in Fig. 3.2a. The heterojunction is shown as a white region between the bulk Ge and MoS₂ regions. Based on the apparent VBM and energy shift, the equilibrium band diagram of the p-Ge/n-MoS₂ heterostructure device was calculated (Fig. 3.2b).

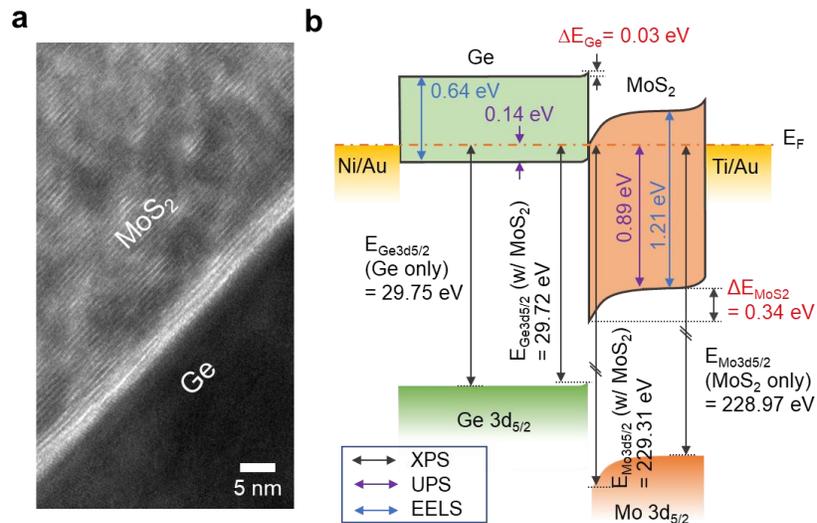


Figure 3.2 Design and operation principle of the Ge/MoS₂ multi-spectral photodetector. **a** Cross-sectional transmission electron microscopy (TEM) image. **b** Energy band diagram in equilibrium¹⁴².

Figures 3.3a, b show energy band diagrams of the p-Ge/n-MoS₂ heterostructure under different biases. A heterojunction with asymmetric transport is designed to enable selective spectrum detection. As shown in Fig. 2A, the transport of the generated holes in the n-MoS₂ region due to the VIS absorption is blocked by the downward band bending, which are eventually recombined with the electrons. Thus, the overall transport of photo-generated carriers under VIS illumination is limited. By contrast, both generated electrons and holes under IR illumination in the p-Ge region easily pass through the heterojunction. Furthermore, the band configuration under a strong bias is harnessed to achieve a completely selective detection functionality. The mechanism of the VIS spectrum detection is illustrated in Fig. 2B. Under IR illumination, the transport from the generated electrons and holes in the p-Ge region is weak and diluted by the increased dark current. However, under VIS illumination, the strong bias alters the band structure bending downward toward the interface of n-MoS₂ and Ti/Au, which elicits a higher photocurrent

compared with the IR illumination. Moreover, the photo-generated holes in the n-MoS₂ region were trapped in the heterojunction interface, resulting in the downward band bending of the n-MoS₂ region¹⁵⁵. As a result, these trapped holes increase the tunneling current included in the total output current.

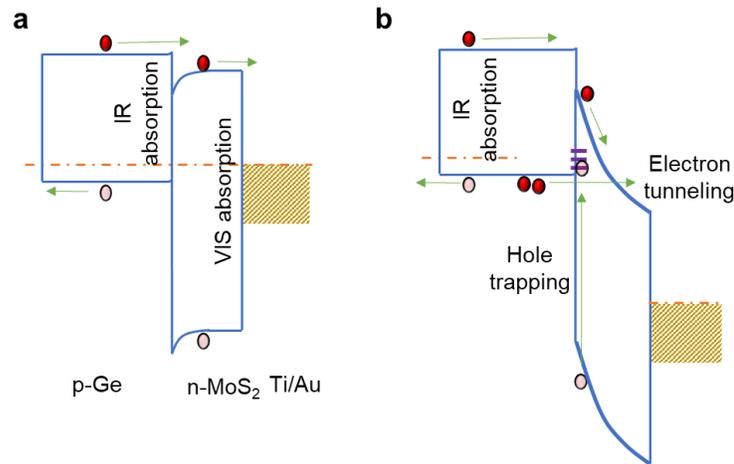


Figure 3.3 Energy band diagrams under **a** near-equilibrium (near-photovoltaic mode) and **b** strong reverse bias (photoconductive mode)¹⁴².

Figure 3.4a shows the $I-V$ characteristics of the p-Ge/n-MoS₂ photodetectors with different wavelengths of illumination. The wavelength is in the range of 406 to 1550 nm. The power density of the incident light is fixed. The longer wavelength exhibits higher responses at -0.5 V (near-photovoltaic mode), whereas the shorter wavelength exhibits higher responses at -3.5 V (photoconductive mode). I calculated the ideality factor ($n=1.31$) based on the experimental dark $I-V$ characteristics. Under a small positive bias (less than 0.5 V), Figure. 3.4b shows the exponential relationship between the current and the voltage.

The selective multi-spectral detection performance of the p-Ge/n-MoS₂ photodetectors is shown in Fig. 3.4b, c. The $I-V$ characteristics of the p-Ge/n-MoS₂ photodetectors exhibit selective

detection at wavelengths of 406 and 1550 nm under various incident power densities. The basic rectifying behavior of the p-Ge/n-MoS₂ heterojunction is also observed in the dark-*I-V* characteristics. Under the 406 nm illumination, a large photocurrent is generated in the photoconductive mode (VIS vision). By contrast, under the 1550 nm illumination, most of the photocurrent is generated at the near-photovoltaic mode (IR vision). The bias-dependent selective phenomena are attributed to the interplay of band energy configurations, such as band offset, trapping, and tunneling electron, and hole currents at different biases. Although the illumination spot is larger than the Ge/MoS₂ junction region, most regions out of the junction are blocked by the metal. Therefore, negligible amount of light generates an undesirable Schottky junction effect by reaching the MoS₂ out of the junction.

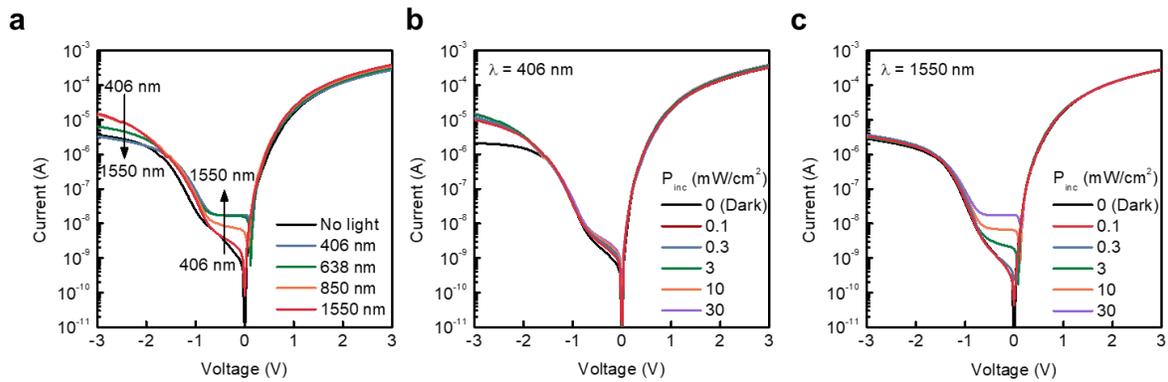


Figure 3.4 *I-V* characteristics of Ge/MoS₂ multi-spectral photodetector with respect to **a** wavelengths and incident power densities at **b** 406 nm and **c** 1550 nm¹⁴².

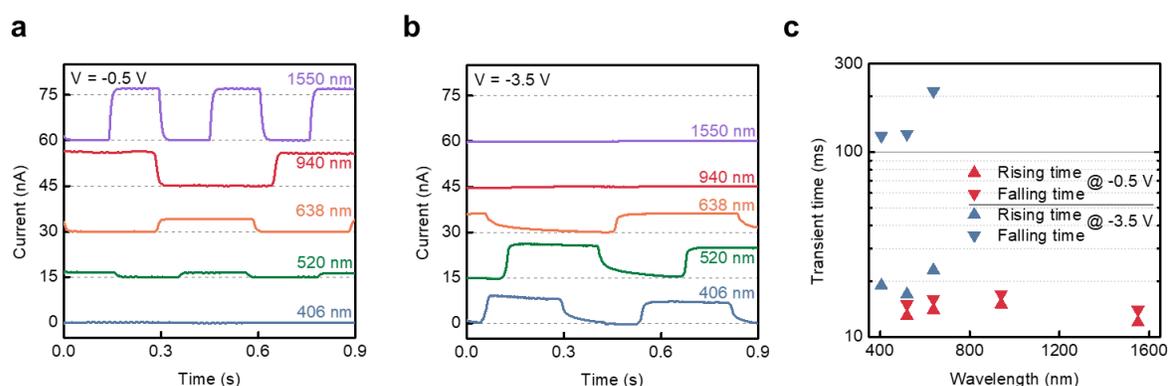


Figure 3.5 Synaptic STP behavior of Ge/MoS₂ multi-spectral photodetector. Temporal response of the device at various illumination wavelengths in the **a** near-photovoltaic and **b** photoconductive modes. The temporal responses for different wavelengths are offset for clarity. **c** Calculated rising and falling times under the near-photovoltaic and photoconductive modes at various wavelengths¹⁴².

Figures 3.5a, b show the transient characteristics of the p-Ge/n-MoS₂ photodetectors, similar to biological STP behavior. The power density is fixed to 30 mW/cm² with bias voltages of -0.5 and -3.5 V. The photo responses at the near-photovoltaic mode show a higher photocurrent when the wavelength becomes longer. Figure 3.5c shows the wavelength-dependent rise and fall times of the device. The two metrics are similar, 16 ms, in the near-photovoltaic mode. In contrast, the fall time is approximately five times larger than the rise time in the photoconductive mode. These bias-dependent variations are attributed to the different absorption aspects of Ge (~0.67 eV) and MoS₂ (~1.23 eV). The bulk Ge layer is covered by the exfoliated 66-nm MoS₂, and the corresponding penetration depths for various VIS wavelengths are comparatively lower than the MoS₂ thickness (66 nm), implying that most of the illuminated VIS are absorbed in the MoS₂ region before reaching the Ge region. Notably, the higher bandgap of MoS₂ with the

heterojunction allows VIS absorption and transport of the generated carriers at a higher bias (-3.5 V). Under weak reverse bias and illumination, the electrons in Ge and holes in the MoS₂ region move toward the barrier. Owing to the modulated heterojunction barrier, the electrons easily move over the barrier, while the transport of the holes is hindered by the downward bending of the MoS₂ valence band. By contrast, under a strong bias, holes in the n-MoS₂ region pass through the heterojunction interface via trap-assisted tunneling. The trapped hole carriers require an additional time to be fully released from the trap states when the light is off. This trap-assisted tunneling mechanism increases the photocurrent decay time, resulting in the slow switching time. Note that the built-in potential of the heterojunction is smaller than the reverse bias of the photoconductive mode and that the response time is dominated by the carrier diffusion in the bulk region at the near photovoltaic mode. Therefore, the temporal effect of the built-in potential is negligible.

Figure 3.6a shows a schematic of the experimental setup of the dual imaging. I prepared a double-sided target object that can transmit and reflect NIR and VIS sources, respectively. The smile and angry images were fabricated on a Si substrate with a 230 nm SiO₂ antireflective coating. At the backside, a 100/300 nm thick Ti/Au array image was deposited and patterned. On the front side, a 300 nm Al layer was deposited. Both metal layers were deposited and patterned by electron-beam evaporation and a standard lift-off technique, respectively. The VIS and NIR spectra were sequentially illuminated on the sample. The NIR light penetrates the backside of the sample, except in the region where the metal is not present (transmission imaging). Similarly, the VIS light toward the front side of the sample is reflected, where the metal is present (reflection imaging). As a result, the transmission imaging exhibits an angry face in the near-photovoltaic mode (Fig. 3.6b), while no visual information is achieved in the photoconductive mode (Fig. 3.6c). However, the reflection mode offers hidden visual information (smile) via the photoconductive

mode (Fig. 3.6d) without achieving an angry image in the near-photovoltaic mode (Fig. 3.6e). This selective imaging capability is attributed to the specific responsivities of the different spectra via switching between the near-photovoltaic and photoconductive modes.

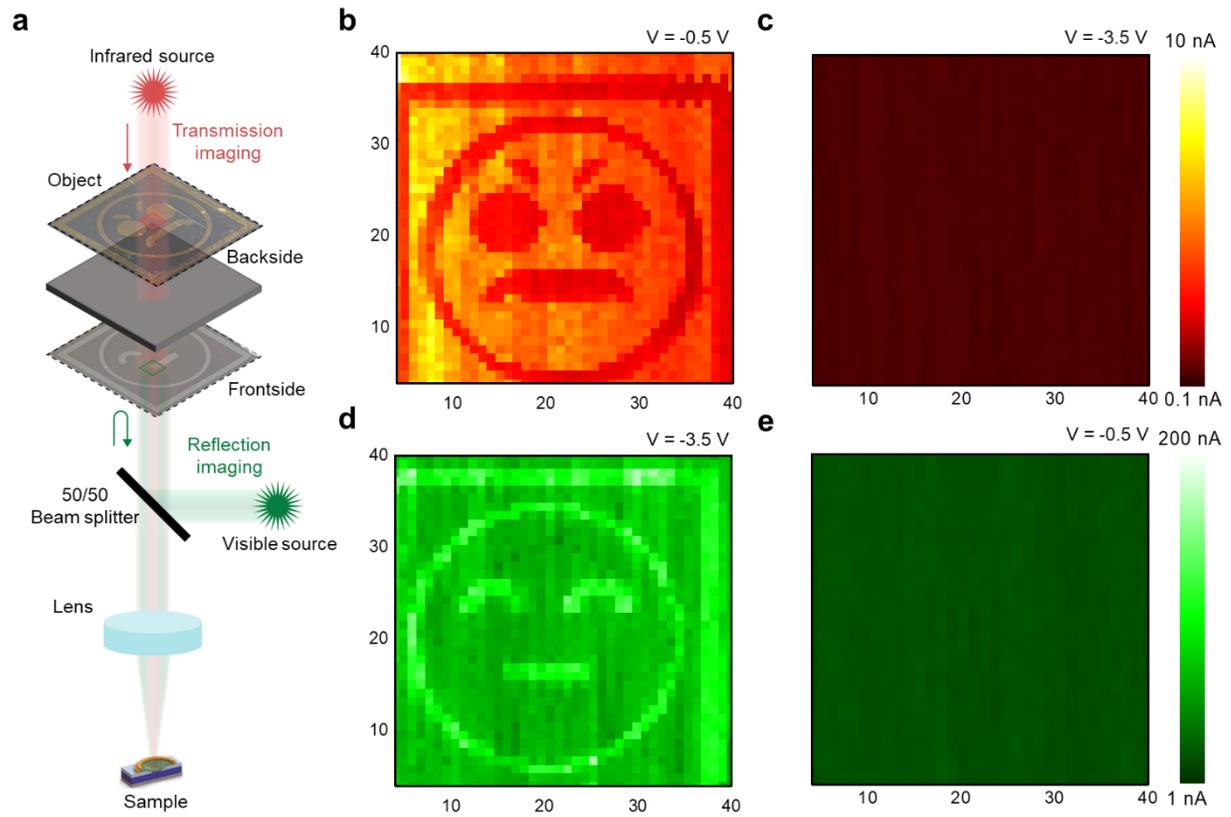


Figure 3.6 VIS/IR dual-band selective imaging. **a** Schematic of the dual-band selective imaging experiment. The double-side-patterned silicon target object exhibits angry and smile face images. Results of the dual-band imaging for transmission imaging (IR illumination) operating at the **b** near-photovoltaic and **c** photoconductive modes. Results of the dual-band imaging for reflection imaging (VIS illumination) operating at the **d** photoconductive and **e** near-photovoltaic modes¹⁴².

Conventional broad-band photodetectors provide inseparable VIS and NIR detection capabilities, whereas the van der Waals heterojunction photodetector offers selective VIS and NIR

vision. The separation of each spectrum is important to extract more synergistic information. For instance, VIS and NIR lights generally propagate simultaneously due to multiple driving forces such as black body radiations (temperature distribution), spectral reflections, and direct illuminations from light sources. Each spectrum is thus originated from the different mechanisms, and by applying a temporal voltage pulse train to our device, separable visual information can be achieved. The separated VIS and NIR vision could be thus harnessed in various applications such as image fusion¹⁴³, red–green–blue–depth imaging¹⁵⁶, and classification via image segmentation¹⁵⁷. As a future direction, the use of two-dimensional materials would provide strong light–matter interactions and ease of integration with conventional silicon readout electronics as well as array-scale implementation^{158–160}. Furthermore, various heterogeneous integration methods would allow the device to be formed in arbitrary 3D geometries, which offers a wide-field-of-view vision along with compact bio-inspired functionalities¹⁴².

3.2 In-Sensor Computing

Analogous to the neural circuit and photoreceptor cells in the retina, photodetectors in image sensors can be directly integrated with artificial synapses (e.g., oxide memristors, phase change memories), constructing an artificial neural circuit^{161–163}. Integrated analog computing units can store image information as resistance states, while also performing computational tasks to implement an ANN for a cognitive algorithm. In this way, so-called “in-sensor computing,” the image information can be processed within the sensors, significantly reducing data movement by processing it at the edge of the system in a similar way to mammalian vision^{164,165}. There have been efforts to demonstrate in-sensor computing systems for analog machine vision using optoelectronic memory devices. However, previously demonstrated in-sensor machine vision systems mostly conducted image memorization and pre-processing (e.g., image enhancement)

within the sensors¹⁶⁶⁻¹⁷⁶. Thus, all the image data still needs to be transferred to a back-end post-processor for high-level image processing, such as feature extraction, image encoding, and image classification. However, minimizing unnecessary data transportation between the sensor, memory, and computation unit is the key feature and goal of in-sensor computing to achieve an energy-efficient and latency-free sensor/processor system. As such, advancements in-sensor computing systems require the capability to memorize images and directly perform high-level image processing within the sensor.

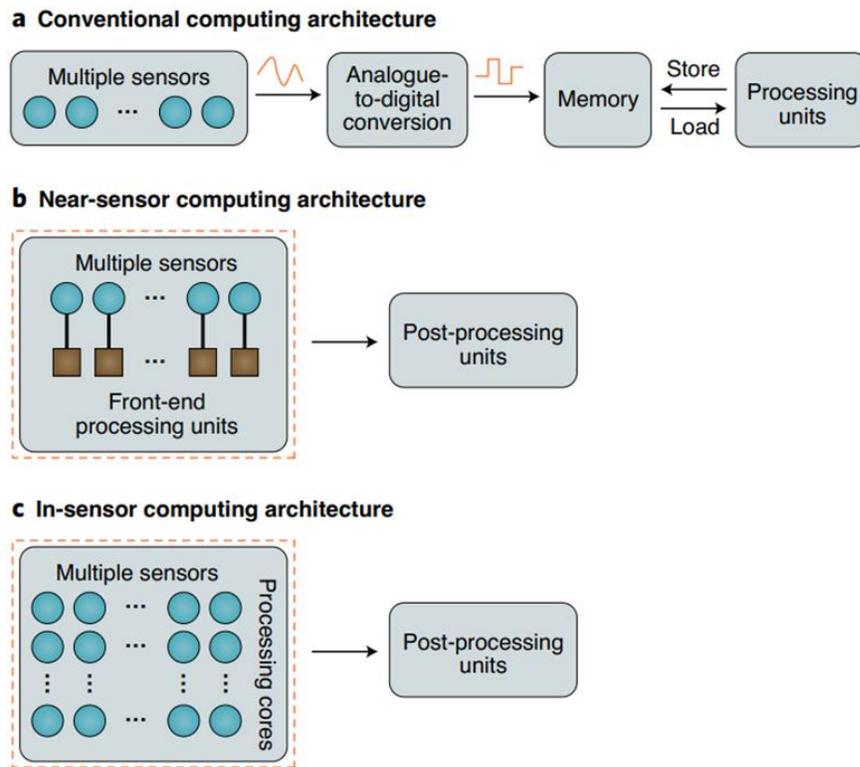


Figure 3.7 Sensory-based computing. **a** Conventional computing architecture. **b** near-sensor and **c** in-sensor computing architectures¹³⁴.

Figure. 3.7a shows the conventional computing architecture for sensory module¹³⁴. The

conventional sensors receive raw analog signals and transmit them to the ADC module for digital processing compatible to high-level CMOS processors, which requires inefficient energy and time to transfer and process the data signals. Unlike the conventional architecture, however, the near-sensor computing architecture encompasses the front-end processing units (ADCs) into the sensory array or module, improving the energy efficiency and reducing the latency issue (Fig. 3.7b). Recent synaptic sensors are even capable of the front-end processing themselves, further enhancing the sensory computing capabilities and minimizing the footprint of the sensory tile (Fig. 3.7c)¹³⁴. In this chapter, I aim to realize the proof-of-concept in-sensor computing protocol associated with the 2D neuromorphic vision computing based on the redox non-volatile memristors and conventional III-V-based photodiodes.

3.3 One-Photodiode One-Resistor (1P1R) Crossbar Array

In this thesis, I developed an advanced in-sensor computing system with neuromorphic image memorization and encoding capabilities within the pixel for visual cognitive processing, emulating the biological visual processing system of the mammalian retina and biological long-term plasticity¹³³. This in-pixel computing system efficiently computes and conveys visual information to minimize data transportation bottlenecks. In a single sensor, a photodiode is directly integrated with a non-volatile redox memristor to construct 1-photodiode 1-resistor (1P1R) pixels, where HfO₂-based memristors are fabricated on InGaAs-based p-i-n photodiodes. First, I fabricated a 1P1R single pixel with an InGaAs photodetector and HfO₂-based memristor and characterized the electrical and optoelectrical properties for the image memorization and data processing of visual information within the sensor. Subsequently, a 16×16 1P1R crossbar array with an InGaAs photodiode and a HfO₂-based memristor was fabricated and characterized.

Subsequently, the imaging of the MNIST handwritten digits was performed, where the visual stimuli of the images were effectively stored in pixels. Subsequently, using the fabricated 1P1R array, the biomimicking image encoding process was performed by in-memory vector-matrix-multiplication without data transfer. In contrast to typical in-memory computing methods, input image data were stored in memristors, and weights were applied to the crossbar array as input voltages, in which 2D-1D vectorization was no longer needed and the size of the crossbar array was significantly reduced. The encoded images were then conveyed to a central processor for image classification. In the in-sensor computing process, the saved images in the pixels were directly computed via in-memory computing without any data transportation between an external memory and/or processor. My process dramatically reduces redundant data movement between the sensor, memory, and processor by performing in-pixel encoding, possibly alleviating data transportation bottlenecks and energy overconsumption.

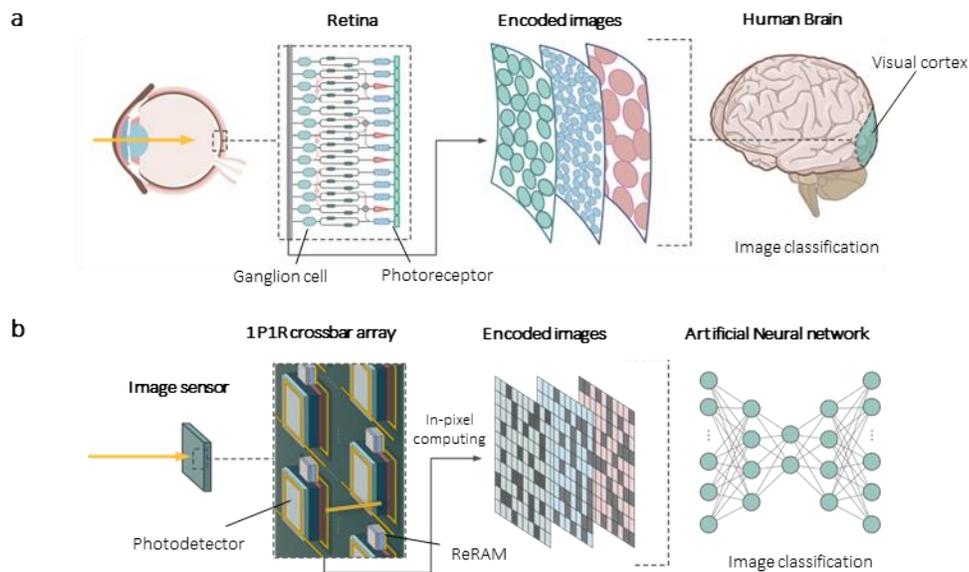


Figure 3.8 In-sensor computing system with an in-pixel direct computing functionality emulating human vision system. **a** Schematic illustration of human vision system from the retina to the brain,

and its visual cognitive processing with in-retina image encoding process. **b** Schematic illustration of a process flow of the one-photodiode one-resistor (1P1R) in-sensor computing system, which emulates the human vision system¹³³.

Figure 3.8a shows a schematic illustration of the image encoding and classification processes in the mammalian retinal and brain system. The outputs of the rod and cone photoreceptors are decomposed into approximately 12 parallel information streams, which are then connected to the retinal ganglion cells. Bipolar and amacrine cell activity are combined in a ganglion cell to create diverse encodings of features extracted from the visual world, such as edges, direction, and color; the retina then transmits these pre-processed data to the brain¹⁷⁷⁻¹⁷⁹. By reducing redundant information, the retina can effectively convey image data to the central processor with a minimal transport delay. In the visual cortex, higher-level visual cognitive processes are conducted using encoded images from the retina¹⁷⁷⁻¹⁷⁹,

In this thesis, I designed and demonstrated an in-sensor neuromorphic machine vision system with functionalities of image memorization and processing, by mimicking the above-mentioned neural circuit and visual classification system in the human eye, as shown in Fig. 3.8b. The image sensor consists of a crossbar array of photodetectors and resistive memory cells, which correspond to photoreceptors and ganglion cells in the retina, respectively. In the retina, the ganglion cells operate as a pre-computing processor unit, whereas the memristor in our system serves both as a memory and computation unit, depending on the polarity and magnitude of the applied bias to each pixel. When a reverse bias with respect to the photodiode is applied to the 1P1R pixels, the sensor operates in a memorization mode, in which incident light stimuli are converted to electrical signals in the photodetectors, and the photocurrents are subsequently stored in the memory cells by changing the conductance of the memory. Under a forward-biased voltage

with respect to the photodiode (lower than the threshold voltage for the erase operation), the sensor operates in the computing mode to process the stored image at the pixels via analog in-memory computing for vector-matrix multiplication. Because vector-matrix multiplication is a key operation in the ANN algorithm, I utilized the 1P1R crossbar array to execute in-sensor image encoding, which extracts critical features from the original image to alleviate the data transfer burden at the sensor and processor interface, paralleling biological processes in the human retina. Finally, image classification was conducted in the post-processing unit with the encoded images delivered through an ANN. While the encoded images possess compressed information compared to the original images, the ANN successfully classifies the object with less computational load.

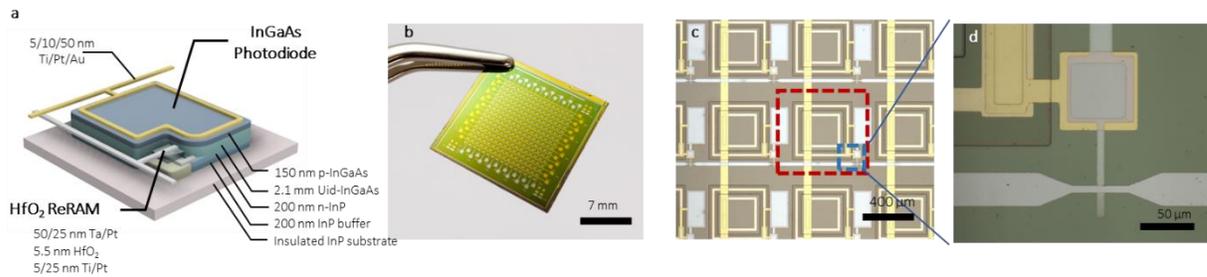


Figure 3.9 Design of single InGaAs 1P1R integrated device. **a** Schematic illustration of the fabricated InGaAs 1P1R device. **b** Optical image of the fabricated 1P1R focal plane array. Scale bar: 7 mm. **c** A optical microscope image of the fabricated 1P1R array. Dashed red (blue) box in shows a 1P1R pixel in the array. Scale bar: 400 μm . **d** Enlarged optical microscope image of the dashed blue box area in **c**. Scale bar: 50 μm ¹³³.

Figure 3.9a shows a schematic illustration of a single 1P1R pixel. InGaAs p-i-n layers were grown on a InP substrate by general molecular beam epitaxy (MBE)¹⁸⁰. The 1P1R crossbar array fabrication starts with a mesa etching of p-InGaAs/Uid-InGaAs layers. The mesa areas were

protected with a bilayer photoresist (PR; LOR3A/AZ5214) by photolithography, and the unprotected InGaAs area was etched using inductively coupled plasma-reactive ion etching (ICP-RIE; BCl₃ 20 sccm, 600 W ICP power, 150 W forward power, 7 mTorr, 20 °C stage temperature for 6 min), followed by wet etching for 1 min in a solution of H₃PO₄:H₂O₂:H₂O = 3:1:25, which stopped at the n-InP layer. The PR mask was then removed in the Remover PG (Kayaku Advanced Materials) at 60 °C. With a single PR (AZ5214) patterning, n-InP mesa for the bottom metal electrodes was defined, followed by a wet etch process for the n-InP/InP buffer layer using a solution of HCl:H₃PO₄ = 3:1 (30 s). Next, a dielectric insulator layer of 150 nm Al₂O₃ was deposited by plasma-enhanced atomic layer deposition (PE-ALD). The via holes were etched with a bilayer PR mask using ICP-RIE (BCl₃ 20 sccm, 50 W ICP power, 200 W forward power, 5 mTorr, 20 °C stage temperature for 6 min). The top and bottom electrodes of the photodiodes were simultaneously deposited by photolithography with bilayer PR and e-beam evaporation of Ti/Pt/Au (5/10/50 nm), which was lifted off in the Remover PG at 60 °C. Another dielectric insulator of 150 nm Al₂O₃ was deposited by PE-ALD, and through holes were opened on the bottom electrodes of the photodiodes. Subsequently, the bottom electrodes of the memristors, which are connected to the bottom electrodes of the photodiodes, were deposited using bilayer photolithography, e-beam evaporation of Ti/Pt (5/25 nm), and a lift-off process with the Remover PG. 5.5 nm of a HfO₂ layer was deposited by PE-ALD, followed by metal deposition of Ta/Pt (50/25 nm) on the top electrodes of the memristors using DC magnetron sputtering of Ta (25 W RF power, 5 mTorr, Ar 20 sccm, room temperature for 18 min) and e-beam evaporation of Pt. Finally, the HfO₂ mesa areas were defined by dry etching with ICP-RIE with a bilayer PR mask. The optical images of the fabricated chip and optical microscope images of the 1P1R crossbar array are shown in Fig. 1b and Figs 1c, d, respectively. Each pixel consisted of an InGaAs photodiode and memristor. The row lines share the top electrodes of the memristors (Ta/Pt

electrodes) and the column lines share the top electrode of the InGaAs photodiodes (p+ electrodes). Prior to operating the 1P1R array, I studied the optical and electrical characteristics of a single 1P1R pixel.

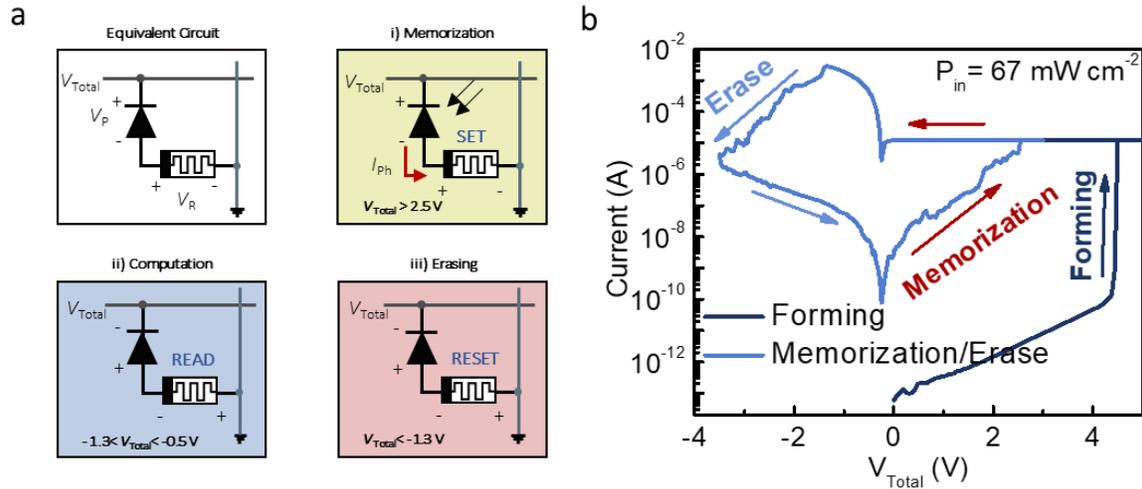


Figure 3.10 Working principle of single InGaAs 1P1R integrated device. **a** Equivalent circuit diagram of the 1P1R structure, and three key operations in the 1P1R device depending on applied bias voltage: i) memorization, ii) computation, and iii) erasing operations. **b** I - V characteristics of the single 1P1R integrated device under light illumination with an incident power density of 67 mW/cm^2 and wavelength of 532 nm ¹³³.

Figure 3.10a depicts its equivalent circuit diagram, demonstrating a working principle of a single 1P1R pixel, where V_P and V_R are the applied voltages of the photodiode and memristor, respectively, and $V_{total}=V_P + V_R$. The 1P1R pixel is composed of an InGaAs p-i-n photodiode and HfO_2 -based memristor, which converts incoming optical signals to electrical signals and memorizes the optical information as its resistance. Additionally, the memristor is utilized as an in-memory computation unit when it operates in the computation mode. The three primary

operations of the 1P1R device, depending on the applied bias voltage (V_{Total}), are depicted in Fig. 3.10a: i) memorization, ii) computation, and iii) erasing operations. When $V_{\text{Total}} > 2.5$ V under light illumination, the photodiode is reverse-biased such that an incident optical signal generates a photocurrent to modulate the resistance of the memristor by forming a conductive filament (memorization operation). Thus, the optical signal can be stored in the form of resistance in a synaptic device. The stored image data in the memristor can then be directly used for high-level in-sensor processing (computation operation). When -1.3 V $< V_{\text{Total}} < -0.5$ V, the photodiode operates under the ohmic regime with relatively low resistance (< 50 Ω) compared to a resistance range of the memristor ($> 1 \cdot 10^3$ Ω); in this way, the 1P1R circuit can be approximated to a single memristor circuit. Thus, the 1P1R crossbar array can be used for synaptic in-memory computing based on Ohm's and Kirchhoff's laws.^{164,181,182} Therefore, memristor serves as a cross-functional device for both the memory unit and processing unit for high-level in-sensory image processing. When a high negative bias voltage over the RESET threshold voltage ($V_{\text{Total}} < -1.3$ V) is applied across the 1P1R device, the memorized data in the memristor are erased (Erase operation). These three operations are key functions for realizing neuromorphic in-pixel image processing with a 1P1R crossbar array.

All electrical measurements were performed using a semiconductor analyzer (B1500A, Keysight). The devices were illuminated using a diode laser (DJ532, Thorlabs) with a wavelength of 532 nm, where the incident power was controlled by a neutral density filter. The conductive filament was formed by the first positive voltage sweep under light illumination, corresponding to the forming loop shown in Fig. 3.10b. After the forming process, the memristor is switched to the OFF state by applying a negative voltage sweep, where the light illumination has no effect on the erase operation (Fig. S1). Subsequently, a positive voltage sweep was conducted on the 1P1R device under light illumination, switching the memristor to the ON state and memorizing the light

information (Fig. 3.10b). Under dark conditions, the memristor in the 1P1R device cannot be switched ON via a positive voltage sweep owing to the lack of sufficient driving current to build the conductive filament, because the current flow is limited by the reverse-biased dark current of the photodiode. This result clearly demonstrates the capability of the 1P1R device as binary optoelectronic memory.

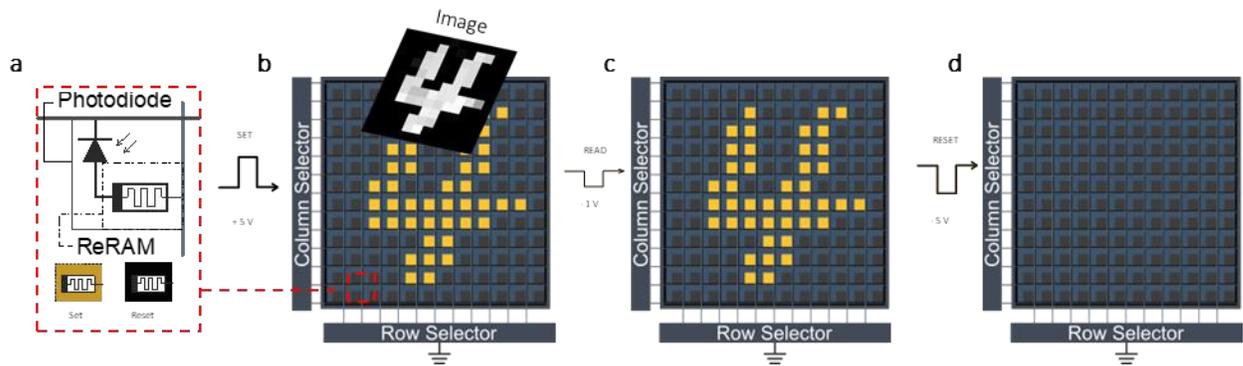


Figure 3.11 Characterization and image memorization process of the fabricated 1P1R crossbar array. **a** Circuit diagram for a pixel in the 16×16 1P1R crossbar array. The pixels are colored in yellow (blue) when the memristors are in a SET (RESET) state in **(b-d)**. **b-d** Schematic illustration of image memorization **b**, readout **c**, and erasing **d** processes in the 1P1R crossbar array¹³³.

Figures 3.11 shows the schematic illustrations of the image memorization, read, and erasing processes, respectively, with the 16×16 1P1R focal plane array. MNIST handwritten digit images of ‘4’ and ‘8’ are printed out on photomasks using a direct write lithography tool (MicroWriter ML3, Durham Magneto Optics Ltd), and the diode laser light with a wavelength of 532 nm ($P = 50 \text{ mJ/cm}^2$) is illuminated on the 1P1R device by passing through the printed digit images. The projected image (‘4’ or ‘8’) on the sensor is memorized by applying a +5 V voltage pulse to each pixel using a semiconductor analyzer. After the memorization process, -1 V pulse

is applied to each pixel to readout the saved image. Then, the stored image is erased by applying -5 V pulse to each pixel. A circuit diagram of the pixels is depicted in Fig. 3a, where a memristor in a yellow (dark blue)-colored pixel is in the LRS (HRS) state. To control the 1P1R array, memorization, read, and erase operations were utilized. For the image memorization process, voltage pulses of +5 V (100 μ m pulse width) were applied across the individual 1P1R pixels, where the photodiodes were reverse-biased, to store incident image information in the memristors, as shown in Fig. 3.11b. The stored image is then read by applying voltage pulses of -1 V to each 1P1R pixel, where the photodiode is forward-biased, to read the resistance states of the memristors (Fig. 3.11c). To erase the saved image in the sensor, voltage pulses of -5 V (100 μ m pulse width) were applied to each 1P1R pixel to switch all pixels to the HRS state, enabling the sensor to be ready to capture the next images (Fig. 3.11d).

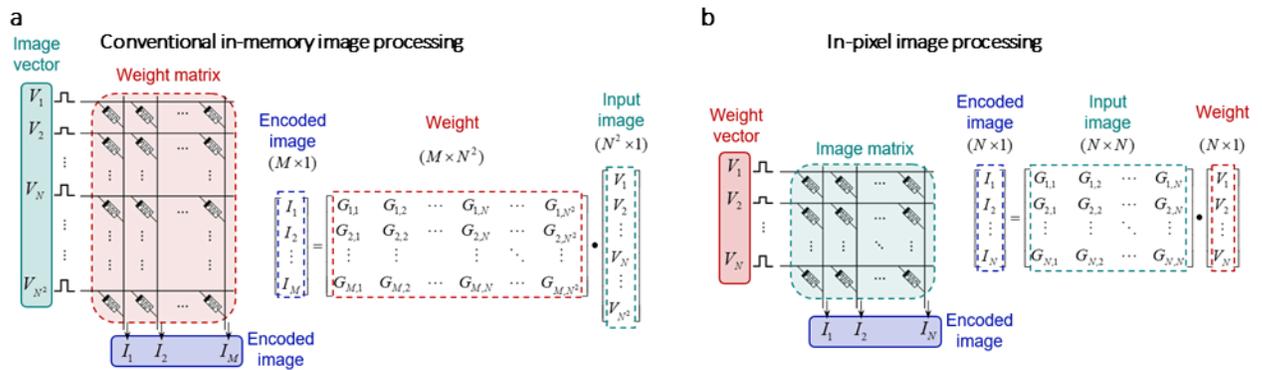


Figure 3.12 Schematic illustration of multiply-accumulate operation for an image encoding process in **a** conventional memristor crossbar array and **b** 1P1R crossbar array. Corresponding matrix-vector multiplication is depicted with parameters of the input voltage (V_j), conductance of memristor (G_{ij}), and output current (I_j)¹³³.

3.3 1P1R-Based In-Sensor Computing

The fabricated 1P1R crossbar focal plane array fuses sensing, learning, and computing capabilities similar to those of biological retinas. To realize a neuromorphic vision system, I stored the vision information in each 1P1R cell as a matrix geometry and simultaneously harnessed the data using emulated vision encoding. Previously demonstrated conventional crossbar geometries of neuromorphic in-memory computing systems for image processing are associated with pre-trained weight values in the ANN matrices, and input image data are applied to the crossbar column as a vectorized electrical signal (Fig. 3.12a)^{161,164,181}. Because the format of image data is usually a 2D $N \times N$ array, 2D-to-1D conversion (vectorization) must be applied as a vector input, which is an $N^2 \times 1$ vector, to the column of the memristor crossbar array. In this case, extra complex circuit components (e.g., ADCs, DACs, and multiplexers) must be added to a peripheral circuit to control a large number of input signals, increasing energy consumption and operational complexity^{164,165,181}. However, our in-pixel image processing system transposes image data to the weights of the ANN, in which the input image is applied and stored in the crossbar array in a weight vector matrix form, as shown in Fig. 3.12b. Therefore, the 2D-to-1D conversion of the image data is no longer necessary for this configuration, significantly reducing the circuit complexity and improving the operational efficiency. Moreover, data transportation from image memorization to the image encoding process is significantly diminished because the image information is directly processed in the pixels without any data transfer.

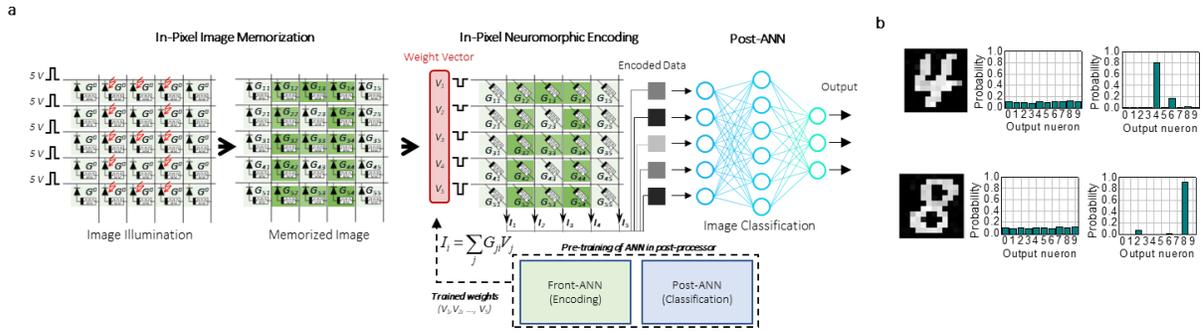


Figure 3.13 Image memorization, encoding and classification via an in-pixel neuromorphic computing. **a** Schematic illustration of an example of in-pixel image memorization, encoding, and classification process with a 5×5 1P1R array. At the initial state, conductance of memristors is G^0 . After image memorization, conductance of each memristor is indicated as G_{ij} . Once an image is memorized in the sensor, pre-trained weight voltages ($-1.3 \text{ V} < V_i < -0.5 \text{ V}$) are applied to the rows of the crossbar to perform a multiply–accumulate operation in the sensor for the encoding process. The encoded data are transferred to a post-ANN to classify the image. **b** Classification results from the memorized ‘4’ and ‘8’ digit images before and after training the ANN¹³³.

Figure 3.13a shows the in-pixel computing process using the fabricated 1P1R array. The 12×12 image of ‘8’ is optically mapped onto the 1P1R array (sensing) and preserved as the conductance of the memristors (learning). Meanwhile, the front-ANN and post-ANN is pre-trained with 10,000 datasets of the MNIST handwritten numbers in the post-processor to extract the optimum weight vector.¹⁸³ The pre-trained 1D weight vector is then converted to electrical signals and applied to the 1P1R array, enabling the physical matrix multiplication for the in-pixel ANN computation via Ohm’s and Kirchhoff’s law (computing). The output current signals from the voltage-conductance multiplication thus represent the encoded vector of the image ‘8,’ achieved without data transportation by emulating the biological encoding capability.

The encoder and classifier models were implemented using Python. I combined a matrix-to-vector encoder (12×12-12) and a fully connected layer classifier with two hidden layers (12-20-16-10) on the Modified National Institute of Standards and Technology (MNIST) dataset. Each original MNIST image was resized to 12×12 pixels, and I trained and tested 10,000 images (with 64 batch sizes for 100 epochs) and 100 MNIST images. For the backpropagation learning process, I employed an RMSprop optimizer, rectifier (softmax for the last output) nonlinearity activations, and an initial time-decaying learning rate (0.001). Figures 3.13b shows the classification results from the measured and memorized ‘4’ and ‘8’ digit images. Before training, the activation level of each digit is randomly distributed. However, the activation level of the ANN output neurons of the ANN is concentrated on a single digit after training. A digit with the highest activation level was adopted as the classified ‘answer’.

Compared to the general 28×28 MNIST handwritten digit dataset, the size of the MNIST dataset in this work is much smaller (12×12), increasing the difficulty of the image classification task. Thus, if I design and fabricate a 1P1R array with a larger number of pixels, I can possibly improve the accuracy of the image classification based on the proposed in-sensor encoding concept. Further accuracy improvement could be achieved by employing a dual encoding neural layer in the ANN, conserving both row- and column-wise features of images, which can be realized by employing bi-directional peripheral circuitry. I believe that more practical in-sensor image processing can be realized by increasing the number of pixels in the 1P1R array and modifying the peripheral circuitry of the system.

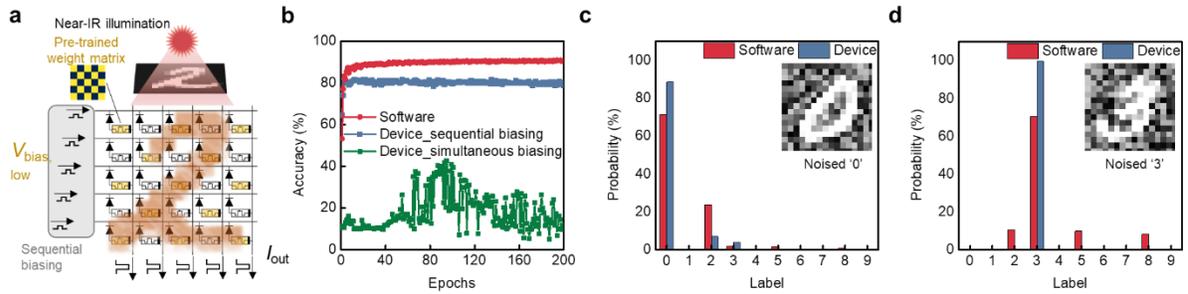


Figure 3.14 Application of in-sensor computing and memorization. **a** Schematic illustration of in-sensor computing mode. The pre-trained weight matrix is mapped to the memristor array, and $V_{bias,low}$ is sequentially applied to each row. The output currents of each column are accumulated (I_{out}). **b** Results of simulated classification accuracy. The device scheme with sequential biasing exhibits classification performance (up to 82%), while the accuracy of the device scheme with conventional simultaneous is up to 43%. Inference results for noised **c** ‘0’ and **d** ‘3’ images. Both software and device scheme identified the noised images correctly¹⁸⁴.

I also conducted simulation with the fabricated 1P1R artificial retina which features two in-sensor modes: the computing mode and the memorization mode¹⁸⁴. Figure. 3.14a shows the schematic illustration of the in-sensor computing mode. I have employed the MNIST digit dataset for the image classification task¹⁸⁵. The total number of labels is 10, from ‘0’ to ‘9’ digits. I first simulated the ANN that includes the in-sensor computing layer implemented via 1P1R pixel configuration and the following fully connected layer (FCL). The pre-trained weight matrix of the in-sensor computing layer is mapped as the resistance of the memristors. To activate the photodiodes, 3 V is applied to the column of the 1P1R artificial retina. Note that 3 V is a low enough voltage level so that the memristors are not programmed.

Unlike a conventional approach that uses the simultaneous biasing of 3 V, I improved the classification performance by employing the sequential biasing of 3 V to each row, which

increases the dimensionality of the output (Fig. 3.14b) by implementing elementwise multiplication. The sequential biasing scheme exhibits up to 82% classification accuracy, while the simultaneous biasing approach shows the accuracy only up to 43% during 200 epochs. The pre-trained 1P1R artificial retina is also robust enough to perform the classification tasks using noised images, as shown in Fig. 3.14c and 3.14d. Both the software-based and device-based simulation results identified the noised images correctly. I simulated the MNIST classification task based on the NeuroSim simulator with experimentally extracted memristor parameters.¹⁸⁶ The number of training and test images is 10,000 and 2,500, respectively. To match the pixel dimensions with our fabricated 1P1R artificial retina, the original 28×28 MNIST images were resized to 16×16. The ANN architecture includes the in-sensor computing layer and 256-64-32-10 fully connected layers. The initial learning rate was 0.12, and the momentum was 0.7. All the weight values in each ANN layer were initialized with a Gaussian distribution, and the feedforwarding batch size was 32 with 200 epochs.

The energy consumption of in-sensor operation (1 cycle of write-read-erase-read) is calculated to be 187.5 μ W, when the pulse 500 μ s with a 50% duty cycle is applied. With respect to computing time, the required time to read the result of in-sensor computing is: $T_{read} = (500 \mu\text{s}) / 0.5 \times 16 = 16 \text{ ms}$. As it can be seen from the calculation, the consumed energy and the processing time are dependent on the employed pulse width, thus it can further be minimized by employing the short pulse for the set/read/erase operations.

3.4 Conclusion

I demonstrated a spectrally selective photodetector based on the heterojunction of the

fabricated p-Ge/n-MoS₂ photodetectors. The two-terminal heterojunction device exhibited both VIS and IR responsivities, depending on the reverse bias level. The selective detection capability is attributed to the heterojunction band modulation and reverse bias level, which promotes or hinders the carrier transport. The vertical heterostructure and single polarity of the bias voltage offer miniaturization of the system and dual-vision imaging capability. I believe that the multi-spectral devices will be useful in the realization of neuromorphic vision systems for various advanced applications including light detection and ranging, healthcare, computer vision, and *in vivo* biomedical imaging¹³³.

I also demonstrated a neuromorphic machine vision system with an in-sensor encoding process inspired by mammalian vision. The focal plane array is based on an InGaAs photodiode directly integrated with HfO₂ memristor, constructing the 1P1R optoelectronic memory and computing pixels. The optoelectronic and memory functionality of the fabricated 1P1R pixel under light illumination showed reliable digital and multibit memory operation and endurance performance. Furthermore, a 16×16 1P1R crossbar array with an InGaAs photodiode and HfO₂-based memristor was used to perform edge computing of the handwritten numbers. Finally, we demonstrated biological image encoding with the developed 1P1R crossbar array, utilizing direct image memorization and in-memory vector matrix multiplication. The encoded images were conveyed to the ANN for image classification, which revealed an accuracy of 82% with 100 training epochs. This slightly low classification accuracy is attributed to the structure of the encoding neural network, which consists of twelve 12×1 fully connected layer. The architecture of the neural network is inevitably determined by the hardware circuit structure of the 1P1R crossbar array. The classification accuracy of our sensor system can be further improved using a dual-encoding neural layer in the ANN. The in-sensor computing concept introduced in this study is a novel method for storing and processing image information directly within pixels without any

data transportation between external computing components and is seamlessly scalable with conventional semiconductor fabrication technology¹⁸⁴.

Chapter 4

Dynamic Bio-Inspired Vision

4.1 GaN/ α -In₂Se₃ Ferroelectric HEMTs

HEMTs are attractive for high-frequency applications owing to their high electron mobility ($>2000 \text{ cm}^2/\text{Vs}$)^{20,187-189}. Recently, ferroelectric-based GaN-MOS HEMTs using HZO, PZT, and HfO₂ dielectrics have been proposed to reduce the device dimensions and achieve lower power consumption via a negative capacitance effect¹⁹⁰⁻¹⁹⁸. However, it has been still challenging for ferroelectric-GaN MOS-HEMTs to achieve ideal performance of a large hysteresis (ΔV) and low subthreshold slope (SS) owing to the depolarization effect caused by the interface charge trapping and polarization of AlGaN barrier layer. To maximize the ΔV and steep SS , the integration of a ferroelectric material with strong ferroelectric polarization and without dangling bonds is ideal for ferroelectric-GaN MOS-HEMTs¹⁹⁹.

In this thesis, I employed a highly crystalline two-dimensional (2D) ferroelectric material, α -In₂Se₃, as a ferroelectric gate heterostructure in GaN MOS-HEMTs because it is free from the charge trapping effect in dangling bonds and has a high carrier concentration. The AlGaN/GaN heterojunction was grown by metal-organic chemical vapor deposition on a 2-inch silicon substrate. It consists of a 3 nm-GaN cap layer, 25 nm-Al_{0.26}Ga_{0.74}N barriers, 1 nm-AlN interlayer, 2 μm -GaN channel layer, and 1 μm -GaN buffer layer. The measured room-temperature Hall mobility and extracted sheet carrier concentrations were $> 1300 \text{ cm}^2/\text{Vs}$ and $\sim 10^{13} \text{ cm}^{-2}$, respectively. Device fabrication was initiated by mesa isolation using a BCl₃/Cl₂-based

inductively coupled plasma reactive ion etching system. A Ti/Au/Al/Ni/Au metal stack was deposited through e-beam evaporation, followed by rapid thermal annealing at 830 °C for 30 s in N₂ environment for ohmic contact formation. Subsequently, 8-nm thick Al₂O₃ for the gate dielectric was deposited using an atomic layer deposition (ALD) system with trimethylaluminum and ozone precursors at a temperature of 450 °C. The α -In₂Se₃ flakes exfoliated from the α -In₂Se₃ bulk (HQ graphene) were then transferred as ferroelectric layer. Subsequently, a Ni/Au (20 nm/50 nm) metal stack was deposited as gate electrode through e-beam evaporation. Finally, the metal-gate aligned α -In₂Se₃ wet etching process was carried out for 60 s in an HCl: H₂O₂ = 1:5 mixture solution.

The α -In₂Se₃ exhibits both ferroelectric and semiconducting characteristics^{200,201}. Thus, it provides a fast speed barrier height modulation along with a relatively robust against depolarization field²⁰². Moreover, the enhancement of the vertical polarization in the ferroelectric layer was achieved by developing a self-aligned structure between the gate electrode and α -In₂Se₃, which suppresses the polarization in the lateral directions. The effectiveness of the aligned structure was also verified using physics-based technology computer aided design (TCAD) simulations. To prove the ferroelectric polarization effect in the GaN/ α -In₂Se₃ MOS-HEMT structure, I demonstrated modulation of the effective schottky barrier height ($\Delta\Phi_{\text{eff}}$) in the GaN/ α -In₂Se₃ metal-semiconductor-HEMT (MES-HEMT) structure. Hence, superior *SS* characteristics of ~10 mV/dec and a ΔV of ~2 V were achieved using the GaN/ α -In₂Se₃ MOS HEMT compared to conventional oxide-based ferroelectric GaN MOS-HEMTs¹²².

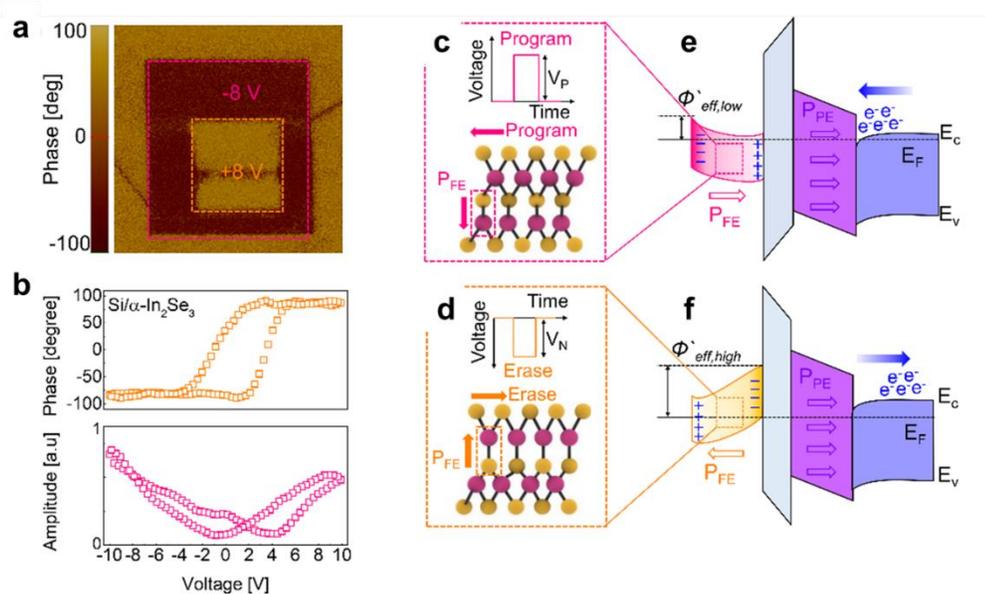


Figure 4.1 GaN/ α -In₂Se₃ HEMTs (FeHEMTs). **a** Scanned piezoelectric force measurement (PFM) phase image after applying a bias of + 8 V and - 8 V, sequentially. **b** Amplitude and phase graph versus external tip bias measured on α -In₂Se₃/Si substrate. The atomic structure of α -In₂Se₃ and pulse scheme for **c** program and **d** erase states. Band diagrams of fabricated device under **e** program and **f** erase states, respectively²⁰.

Figure 4.1a shows the piezo-response force microscopy (PFM) mapping for out-of-plane (OOP, vertical) polarizations of α -In₂Se₃ by applying a sequential bidirectional DC bias of +8 V and -8V. The OOP and in-plane (IP, lateral) ferroelectric polarizations of α -In₂Se₃ are switched simultaneously via an external electric field, resulting in a dipole-locking effect^{132,203}. However, the OOP ferroelectric polarization of α -In₂Se₃ dominantly affects the channel conductance of the GaN/ α -In₂Se₃ MOS-HEMT as a gate structure. Figure 4.1b shows the amplitude and phase responses to the application of external voltages, confirming the clear transition of the polarization

above the coercive field. The ferroelectricity of layered α -In₂Se₃ originates from the shift of the Se atom by an external electric field, as shown in Fig. 4.1c, d. Using the ferroelectric property of α -In₂Se₃, I have demonstrated a GaN/ α -In₂Se₃ MOS-HEMT that exhibits output current states. The application of positive (for program, V_P) and negative (for erase, V_N) pulses to the gate electrode tunes the ferroelectric polarization of the α -In₂Se₃ layer coupled with the V_P channel in the GaN MOS-HEMT.

The multilayer ferroelectric polarization switching characteristic of α -In₂Se₃ allows the emulation of biological STP. Specifically, STP behavior can be achieved by the local polarization of the α -In₂Se₃ membrane, offering reservoir computing (RC) capability. In conjunction with high-frequency operation, the heterogeneously integrated GaN/ α -In₂Se₃ MOS-HEMT serves as a single synapse. Figure 4.1e shows the band diagram of the GaN/ α -In₂Se₃ MOS-HEMT heterostructure with the application of a programming pulse, where Φ_{eff} indicates the effective barrier height through the α -In₂Se₃ layer. The V_P leads to the alignment of the ferroelectric polarization to in the same direction as the polarization in the AlGaN barrier (P_{PE}). Therefore, carriers in the AlGaN/GaN 2DEG channel accumulate when the conduction band is lowered below the Fermi level. With the application of V_N , the ferroelectric polarization is aligned in the opposite direction of the P_{PE} , which counters the polarization in the AlGaN barrier. Thus, the conduction band in the heterostructure rises above the Fermi level and the carriers in the 2DEG channel are fully depleted, as shown in Figure 4.1f. Furthermore, the accumulation and depletion of carriers within semiconducting α -In₂Se₃ layer depending on applied gate bias modulate the Φ_{eff} , inducing the depletion capacitance (C_{dep}) at the interface of α -In₂Se₃/Al₂O₃.

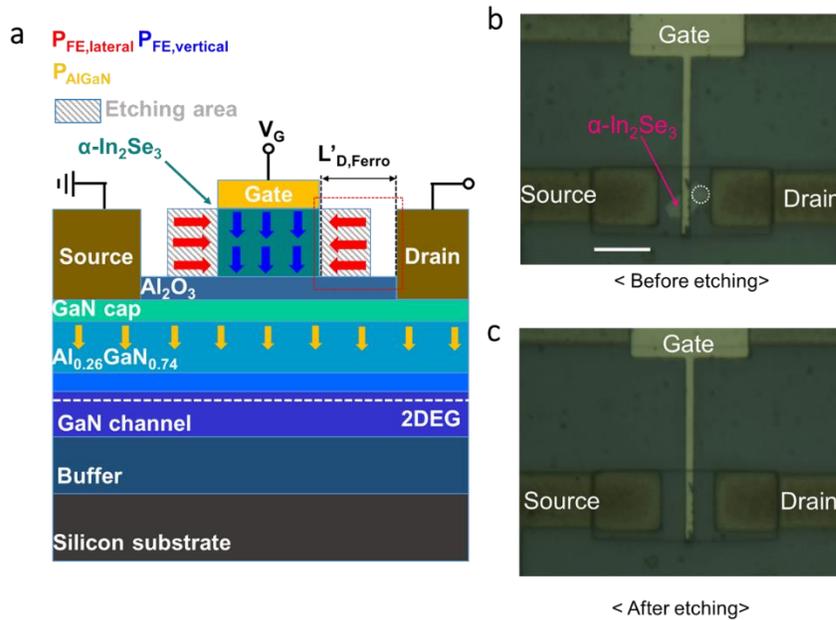


Figure 4.2 Design of GaN/ α -In₂Se₃ FeHEMTs. **a** Cross-section schematic of GaN/ α -In₂Se₃ MOS-HEMT before and after the etching process. Optical microscopy image of the region where Raman spectroscopy analysis was conducted on α -In₂Se₃ **b** before and **c** after the etching process (white dashed circle)²⁰.

To maximize the ferroelectric effect on the coupled 2DEG channel, I confined the α -In₂Se₃ area by selectively removing the film outside the gate metal region through a wet-etching process. For the GaN/ α -In₂Se₃ MOS-HEMT structure without a confined α -In₂Se₃ layer, the area of α -In₂Se₃ adjacent to the drain electrode was unintentionally polarized owing to the drain bias. Thus, when an V_N and V_P are applied to the gate electrode to align the polarization, the induced parasitic polarization from the drain bias offsets the polarization derived via the gate pulse. Therefore, this limits the ferroelectric polarization of the GaN/ α -In₂Se₃ MOS-HEMT in the OOP direction. Herein, I constrained the unwanted parasitic IP polarization by self-defining the area of α -In₂Se₃ matching with the gate electrode. Figure 4.2a shows a schematic of the GaN/ α -In₂Se₃ MOS-

HEMT before and after confining the α -In₂Se₃ layer, where the grey dashed box represents the etched area. The red and blue arrows indicate IP and OOP polarizations in the α -In₂Se₃ layer, respectively. The $L'_{D,Ferro}$ (4.5 μ m) is the length between the α -In₂Se₃ and drain metal after the self-defined wet etching process. Optical microscopy images of the device before and after the self-defined etching process are shown in Fig. 4.2b, c.

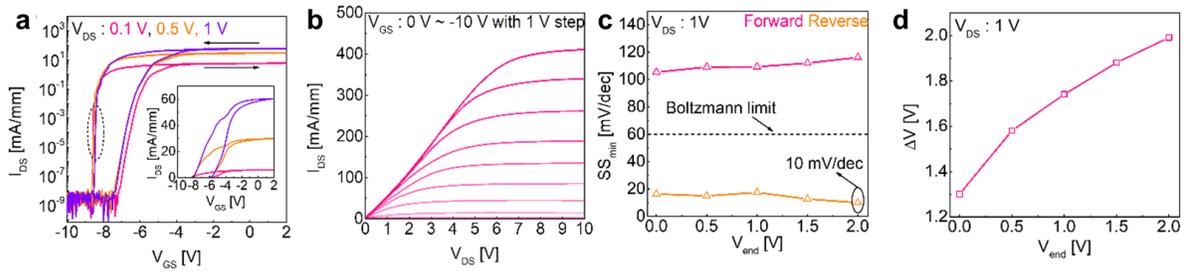


Figure 4.3 Electrical characterization of GaN/ α -In₂Se₃ FeHEMTs. **a** Normalized bidirectional I_{DS} vs gate voltage (V_{GS}) transfer curve of GaN/ α -In₂Se₃ MOS-HEMT at drain voltage (V_{DS}) = 0.1, 0.5, 1 V, respectively. An inset shows the linear scale of bidirectional curve of Figure 4.3a. **b** Normalized I_{DS} vs V_{DS} output curve of GaN/ α -In₂Se₃ MOS-HEMT at V_{GS} range of 0 V ~ -10 V. **c** and **d** Minimum subthreshold swing (SS_{min}) and ΔV versus different V_{GS} sweep ranges. Parameters were extracted from the bidirectional I_{DS} vs V_{GS} transfer curve of GaN/ α -In₂Se₃ MOS-HEMT²⁰.

Figure 4.3a shows the bi-directional transfer curves at drain voltages (V_{DS}) of 0.1, 0.5, and 1 V. Counterclockwise ΔV is referred to as switching of the ferroelectric polarization of α -In₂Se₃ in the GaN/ α -In₂Se₃ MOS-HEMT structure. To confirm the output characteristics of the GaN/ α -In₂Se₃ MOS-HEMT, I measured the I_{DS} vs. V_{DS} output curve for the V_{GS} range from 0 V to -10 V with a 1 V step, as shown in Fig. 4.3b. Furthermore, the SS_{min} and ΔV under various V_{GS} sweep

ranges (V_{end} indicate the end point of V_{GS} sweep range), -10 V ~ 0 V and -10 V ~ 2V, were extracted from the bi-directional transfer curve at $V_{\text{DS}} = 1$ V. Steep SS in ferroelectric-FET (FeFET) is advantageous for neuromorphic computing applications because it minimizes the power consumption by preventing unwanted switching behavior until it reaches the threshold voltage. Specifically, the low subthreshold current suppresses the leakage current generated by a sequential input pulse train²⁰⁴. Figure 4.3c shows SS_{min} vs. V_{end} for the forward and reverse sweeps. The SS_{min} is maintained below 20 mV/dec at V_{end} range from 0 V to 2 V, and it is decreased further to 10 mV/dec when $V_{\text{end}} = 2$ V. Owing to the coercive field (for both negative and positive amplitude) in α - In_2Se_3 , as confirmed in PFM data, the direction of ferroelectric polarization can be flipped by an external electric field. Figure 4.3d exhibits the ΔV graph extracted from the transfer curve for V_{end} from 0 V to 2 V, indicating a gate sweep range (from -10 V to 2 V) sufficiently induces ferroelectricity of α - In_2Se_3 .

The low SS_{min} achieved in the GaN/ α - In_2Se_3 MOS-HEMT structure can be explained by the emergence of the additional C_{dep} induced by α - In_2Se_3 layer owing to its ferroelectric and semiconductor properties. Introduction of semiconducting α - In_2Se_3 layer causes the depletion at the interface of α - $\text{In}_2\text{Se}_3/\text{Al}_2\text{O}_3$, and thus induces additional C_{dep} . The equivalent circuit of a classical ferroelectric MOS-HEMT is a series combination of $C_{\text{pol}} - C_{\text{ins}} - C_{\text{AlGaN}} - C_{\text{GaN}}$ without C_{dep} , where C_{pol} and C_{ins} are the capacitances of the ferroelectric polarization and dielectric layer, respectively¹⁹². The SS extracted from the ferroelectric MOS-HEMT can be expressed as below:

$$SS = 60 \cdot \left(1 + \frac{C_{\text{HEMT}}}{C_{\text{ins}}}\right) \cdot \left(1 - \frac{C_{\text{MOS}}}{|C_{\text{pol}}|}\right), C_{\text{MOS}} = \frac{C_{\text{HEMT}} \cdot C_{\text{ins}}}{C_{\text{HEMT}} + C_{\text{ins}}} \quad (1)$$

$$\frac{1}{A_v} = \left(1 - \frac{C_{\text{MOS}}}{|C_{\text{pol}}|}\right) \quad (2)$$

The term $\left(1 - \frac{C_{MOS}}{|C_{pol}|}\right)$ in Eq. (1) indicates that ferroelectric polarization can further decrease SS below 60 mV/dec. Specifically, the term $\left(1 - \frac{C_{MOS}}{|C_{pol}|}\right)$ is associated with voltage amplification (A_V), and the term $1/A_V$ can be described as shown in Eq. (2). Under the specific conditions ($C_{ins} > |C_{pol}|$), $1/A_V$ has a value of less than 1 ($1/A_V < 1$), and it is mainly attributed to the polarization switching in α -In₂Se₃¹⁹². C_{ins} of 1.593×10^{-6} F/cm² and C_{pol} of $\sim 0.2 \times 10^{-6}$ F/cm² are extracted from C-V measurement using the fabricated MOSCAP and through the TCAD simulation, respectively. By comparing C_{ins} and C_{pol} , I confirmed that the condition of $C_{ins} > |C_{pol}|$ for proving $1/A_V < 1$ is satisfied, and the SS below 60 mV/dec at room temperature can be explained using Eq. (1), (2)¹⁹². In addition to the ferroelectric polarization effect, the reduction in C_{pol} by adding a series capacitance results in a further decrease in the SS . Hence, C_{dep} can even further enhance the subthreshold performance of the device. Thus, the semiconducting property of α -In₂Se₃ uniquely contributed to the superior performance of the GaN/ α -In₂Se₃ MOS-HEMT.

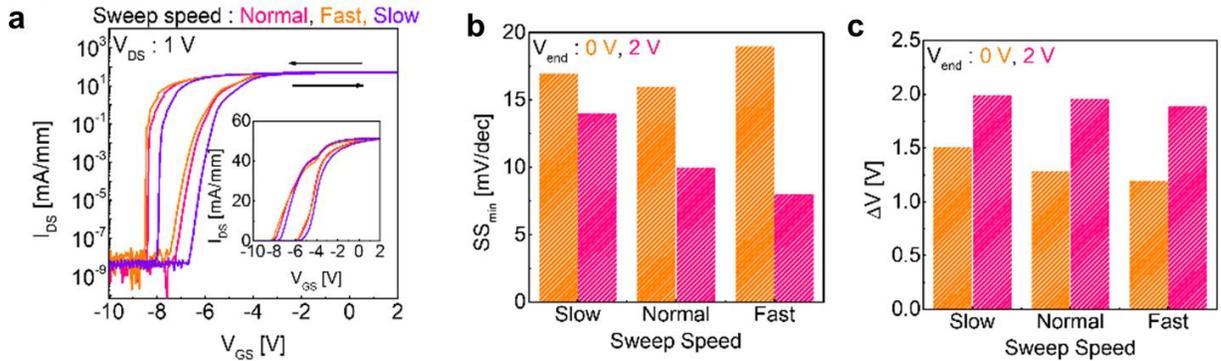


Figure 4.4 Subthreshold swing (SS) and hysteresis characterizations of GaN/ α -In₂Se₃ FeHEMTs.

a Bi-directional I_{DS} vs V_{GS} transfer curve of GaN/ α -In₂Se₃ MOS-HEMT at $V_{DS} = 1$ V, measured under a slow, normal, and fast sweep speed mode. **b** SS_{min} and **c** ΔV extracted from Fig. 4.3a at $V_{end} = 0, 2$ V under a slow, normal, and fast sweep speed²⁰.

Figure 4.4a shows the bi-directional transfer curves for the slow (0.3 V/s), normal (0.9 V/s), and fast (2.9 V/s) sweep speed modes. It coherently exhibits counterclockwise ΔV , and threshold voltage shifts to the negative direction as the sweep speed decreases because the gate bias polling time is reduced. The SS_{\min} depending on the sweep speed for $V_{\text{end}} = 0$ and 2 V extracted from the bi-directional curve is shown in Fig. 4.4b. The SS_{\min} for each case was lower than 20 mV/dec; in particular, the SS_{\min} decreased to 8 mV/dec at $V_{\text{end}} = 2$ V under a fast sweep speed mode. In addition, ΔV is dependent on the sweep speed that supports the polarization matching effect. The ΔV has been weakened in the fast sweep speed mode, as shown in Fig. 4.4c²⁰.

4.2 GaN/CuInP₂S₆ Ferroelectric HEMTs

I employed a CIPS/GaN ferroelectric HEMT (FeHEMT) to achieve high-power driving capability. GaN-based HEMTs are widely employed in radio frequency (RF) and power electronics applications because of their high output current and fast switching capability by employing an AlGaN/GaN heterostructure, which forms a two-dimensional electron gas (2DEG) transport channel²⁰⁵. Thus, it offers a high electron saturation velocity, a high breakdown electrical field, and a high electron mobility suitable for high-power and high-frequency applications^{188,206}.

Epilayers for HEMT structure were grown by metal-organic chemical vapor deposition (MOCVD) on a 2 inch silicon substrate (MSE supplies, USA). The heterostructure was etched to isolate the mesa structure via a BCl₃/Cl₂-based inductively coupled plasma reactive ion etching (ICP-RIE) system. A Ti/Al (25/140 nm) metal stack was deposited via sputtering, followed by the deposition of Ni/Au (40/50 nm) via an electron-beam (e-beam) evaporator. The sample was annealed using a rapid thermal annealing (RTA) system at 830 °C for 30 s in an N₂ atmosphere

to form ohmic contacts. A 5 nm Al_2O_3 gate dielectric was deposited using an atomic layer deposition (ALD) system with trimethylaluminum and ozone precursors at 450 °C. A CIPS flake (60 nm) was then exfoliated from the CIPS bulk (HQ graphene) and transferred to the HEMT channel interface. The final Ni/Au (40/80 nm) gate electrode was deposited by e-beam evaporation.

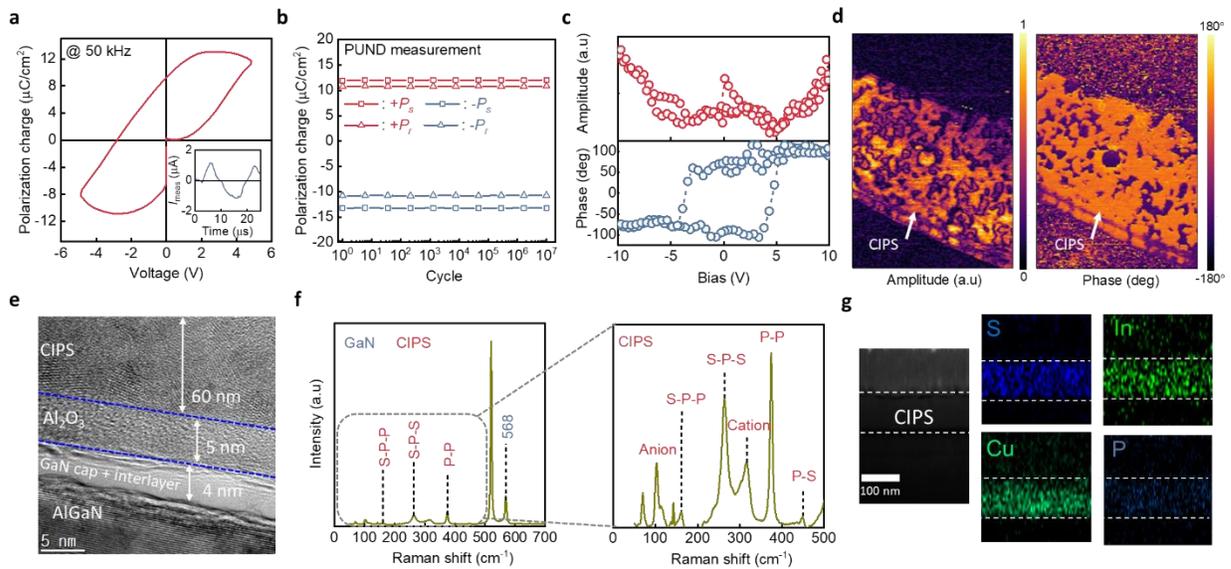


Figure 4.5 Material analysis of GaN/CuInP₂S₆ (CIPS) FeHEMT. **a** Polarization-voltage (P - V) characterization of Ti/Au/CIPS/Ti/Au metal-ferroelectric-metal (MFM) capacitor under 50-kHz triangular pulse. Inset: positive-up negative-down (PUND) waveform. **b** Polarization switching endurance of CIPS membrane. Red and blue lines: endurance under downward and upward polarization, respectively. **c** PFM characterization of CIPS membrane. The hysteresis loops of both the amplitude and phase PFM verify the ferroelectricity of the CIPS membrane. **d** PFM imaging of CIPS membrane on metal ($15\ \mu\text{m} \times 10\ \mu\text{m}$ horizontal scan). The bright region indicates the piezoelectric response of the CIPS membrane. **e** Cross-sectional transmission electron microscopy (TEM) image of GaN/CIPS FeHEMT. **f** Raman spectroscopy of GaN/CIPS FeHEMT, **g**

including GaN (left) and CIPS (right) peaks. The GaN peak is at 568 cm^{-1} , and the CIPS peaks include the multiple vibration peaks ($150\text{-}400\text{ cm}^{-1}$) and ionic peaks (100 cm^{-1} and 315 cm^{-1}). **G** Energy dispersive X-ray spectroscopy (EDS) characterization of cross-sectional GaN/CIPS FeHEMT. White dash line: CIPS region. Cu, In, S, and P are uniformly distributed in the CIPS region.

To analyze the ferroelectric properties of the CIPS membrane, including the polarization window, endurance, and spatial ferroelectricity, I first fabricated a two-terminal metal-ferroelectric-metal (MFM) capacitor. Figure 4.5a shows the bipolar polarization–voltage (P – V) curve of the CIPS without a preset loop. The P – V curve exhibited a hysteresis loop, a remanent polarization of approximately $10\text{ }\mu\text{C}/\text{cm}^2$, and a coercive voltage of $\pm 3\text{ V}$ driven by a 50 kHz triangular pulse. I also characterized the endurance performance using the repeated positive-up and negative-down (PUND) measurement technique (Fig. 4.5b). The polarization switching between saturation (P_s) and remanent polarization charge (P_r) persisted over 10^7 cycles, confirming the stability and programmability of the CIPS membrane. The results of piezoelectric force microscopy (PFM) measurements are shown in Fig. 4.5c, d. The hysteresis loops in both the amplitude and phase of the PFM measurements indicated ferroelectricity in the CIPS membrane (Fig. 4.5c). The spatially resolved amplitude and phase of the piezoelectric potential were mapped in the PFM images (Fig. 4.5d), confirming the piezoelectric response of the CIPS membrane. Figure 4.5e displays a cross-sectional transmission electron microscopy (TEM) image of the CIPS membrane, which was mechanically exfoliated from the bulk crystal. The freestanding membrane was heterogeneously integrated onto an AlGaN/GaN HEMT epilayer passivated with Al_2O_3 as a gate dielectric through van der Waals forces. The thicknesses of the CIPS and Al_2O_3 layers were 60 nm and 5 nm , respectively. Figure 4.5f shows the Raman spectra of the GaN/CIPS FeHEMT.

The peak at 568 cm^{-1} indicates the GaN layer (Fig. 2F), and the peaks at 160 , 263 , and 375 cm^{-1} correspond to the S-P-P, S-P-S, and P-P vibrations of the CIPS membrane, respectively. The ionic responses of the CIPS membrane were also characterized, with $\text{P}_2\text{S}_6^{4-}$ anion and Cu^+ cation peaks observed at approximately 100 and 315 cm^{-1} , respectively²⁰⁷. Figure 4.5g shows the cross-sectional energy-dispersive X-ray spectroscopy (EDS) mapping results of the GaN/CIPS FeHEMT. The EDS mapping clearly confirms that the CIPS layer is composed of uniformly distributed Cu, In, P, and S. A cross-sectional GaN/CIPS FeHEMT structure was analyzed using a focused ion beam (FIB) system and HR-TEM (JEM-2100F(HR), JEOL Ltd.) with EDS elemental mapping. The PFM amplitude and phase were characterized using a PFM tip with a writing bias range of $\pm 10\text{ V}$ in the electrostatic force microscopy (EFM) mode (NX-10/Park Systems). Micro-Raman spectroscopy was performed at room temperature using a Raman imaging microscope (DXR 2xi) with a laser excitation wavelength of 532 nm .

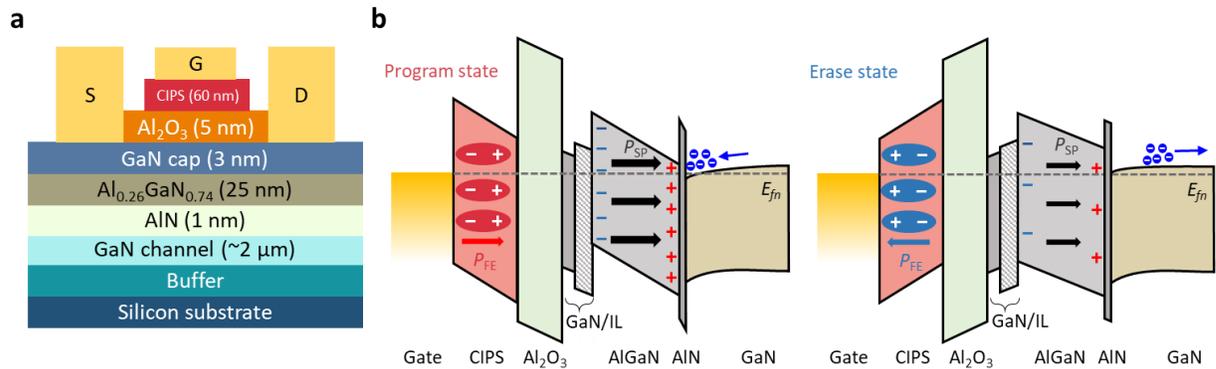


Figure 4.6 Working principle of GaN/CIPS FeHEMT. **a** Cross-sectional device structure of CIPS/GaN HEMT. **b** Band diagrams of program and erase states of GaN/CIPS FeHEMT. The IL in the figure indicates an oxygen-abundant GaN interlayer between Al_2O_3 and AlGaIn . The positive bias generates downward polarization, lowering the barrier height and increasing the transconductance of the GaN/CIPS FeHEMT. In contrast, the negative bias generates the upward

polarization that increases the barrier height and decreases the transconductance of the GaN/CIPS FeHEMT.

Figure 4.6a provides a detailed device structure of the fabricated GaN/CIPS FeHEMT, which includes a Ti/Al/Ni/Au (25/140/40/50 nm) source/drain electrode, a Ni/Au (40/50 nm) gate electrode, a CIPS/Al₂O₃ (60/5 nm) gate dielectric, a GaN capping layer (3 nm), Al_{0.26}Ga_{0.74}N (25 nm) layer, AlN (1 nm), and a GaN channel layer (~2 μm). Figure 4.6b illustrates the band diagrams for the GaN/CIPS FeHEMT under programmed and erased states, estimated by physics-based technology computer-aided design (TCAD) simulations. An external electric field alters the polarization state of the CIPS gate dielectric. A positive bias generates a downward polarization state that reinforces the spontaneous polarization of AlGaN and increases the electron carrier concentration in the 2DEG of the HEMT structure. In contrast, a negative bias switches the polarization state upwards, decreasing the electron carrier density in the HEMT channel by counterbalancing the spontaneous polarization of AlGaN¹²². This ferroelectric switching capability, coupled with the carrier concentration in the 2DEG through external stimuli, enables a programmable output current.

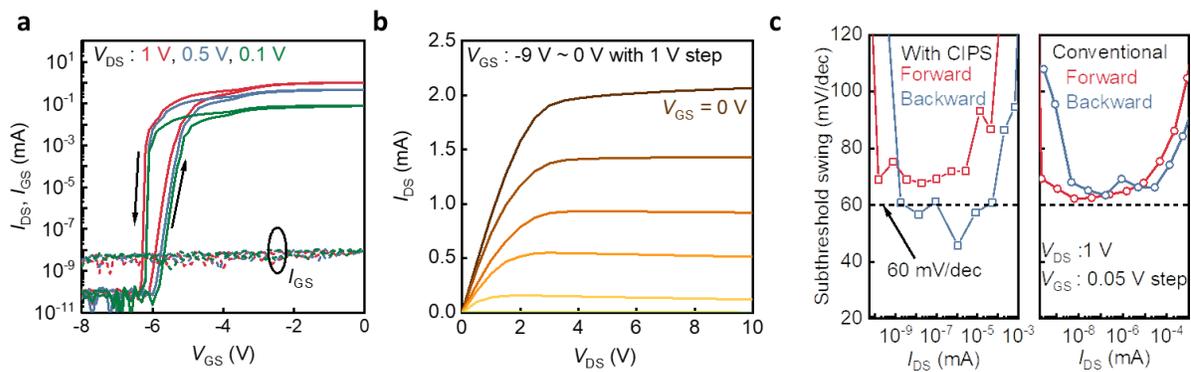


Figure 4.7 Electrical characterization of GaN/CIPS FeHEMT. **a** I - V transfer curve of the

fabricated CIPS/GaN HEMT. The double sweep switches the ferroelectric polarization of the CIPS, thus the hysteresis loop forms. **b** I - V output curve of the GaN/CIPS FeHEMT with an output current of approximately 2 mA. **c** Comparison of subthreshold swing (SS) between GaN/CIPS FeHEMT and GaN HEMT. With the CIPS membrane, the SS overcomes Boltzmann limitation (60 mV/dec).

Figure 4.7a presents the transfer I - V characteristics of the fabricated GaN/CIPS FeHEMT during forward and backward bias sweeps. All electrical measurements were performed using a semiconductor parameter analyzer with a pre-amp (Keithley-4200A-SCS). The piezoelectric properties of α - In_2Se_3 were characterized using piezoelectric force measurements (PFM; NX-10/Park Systems), and the thickness of the α - In_2Se_3 flake was verified using atomic force microscopy (XE-150/Park Systems). The Raman shift was measured using a Raman spectrometer (DXR2xi/Thermo).

The forward sweep led to the switching of out-of-plane polarization of the GaN/CIPS FeHEMT, decreasing a threshold voltage up to 0.5 V with an $\sim 1.5 \times 10^8$ ON/OFF ratio during the backward sweep. The gate leakage current (I_{GS}) of the GaN/CIPS FeHEMT was substantially smaller (~ 5 pA) than the on-current (I_{DS} , 0.95 mA). During a backward sweep, the CIPS polarization changed to downward and enhance the polarization in the AlGaN barrier layer. The CIPS down-polarization induced the accumulation of 2DEG, allowing the transport channel to be more conductive. This results in an increase in I_{DS} at a fixed gate bias, and thus negatively shifts the threshold voltage (V_{th}). During a forward sweep, the depletion of the 2DEG resulted in a positive shift in V_{th} owing to the reduced current at a fixed gate.

Figure 4.7b shows the output I - V characteristics of the GaN/CIPS FeHEMT. The saturation current of the GaN/CIPS FeHEMT after polarized switching was 2 mA at gate voltage

($V_{GS} = 0$ V) and drain voltage ($V_{DS} = 10$ V), which corresponds to a current density of 200 mA/mm. As shown in Fig. 4.7c, the subthreshold swing (SS) was improved by using the CIPS/GaN van der Waals heterostructure due to the negative capacitance effect of the CIPS gate dielectric. The lower SS drives a sufficient current at low drain bias and suppresses a leakage current, reducing the energy consumption for neuronal circuitry such as spiking neural networks and reservoir computing²⁰⁸. After polarization switching, the SS at the backward sweep overcomes the 60 mV/dec Boltzmann limitation (46 mV/dec) at room temperature, whereas the conventional AlGaIn/GaN HEMT without the CIPS membrane only shows a SS of near 60 mV/dec. With a programming gate pulse of 2 ms, the retention time of the out-of-plane polarization persisted for more than 5 hours, which is sufficient for time-varying high-order neuromorphic computing applications.

4.3 Synaptic GaN HEMT-Based Reservoir Computing

RC is well known for being suitable for processing temporal data and data with high dimensionality^{14,19}. RC process, which mimic the biological STP, allows for a non-linear encoding framework in a time domain, reducing the data dimensionality without the need for additional training processes²⁰⁹.

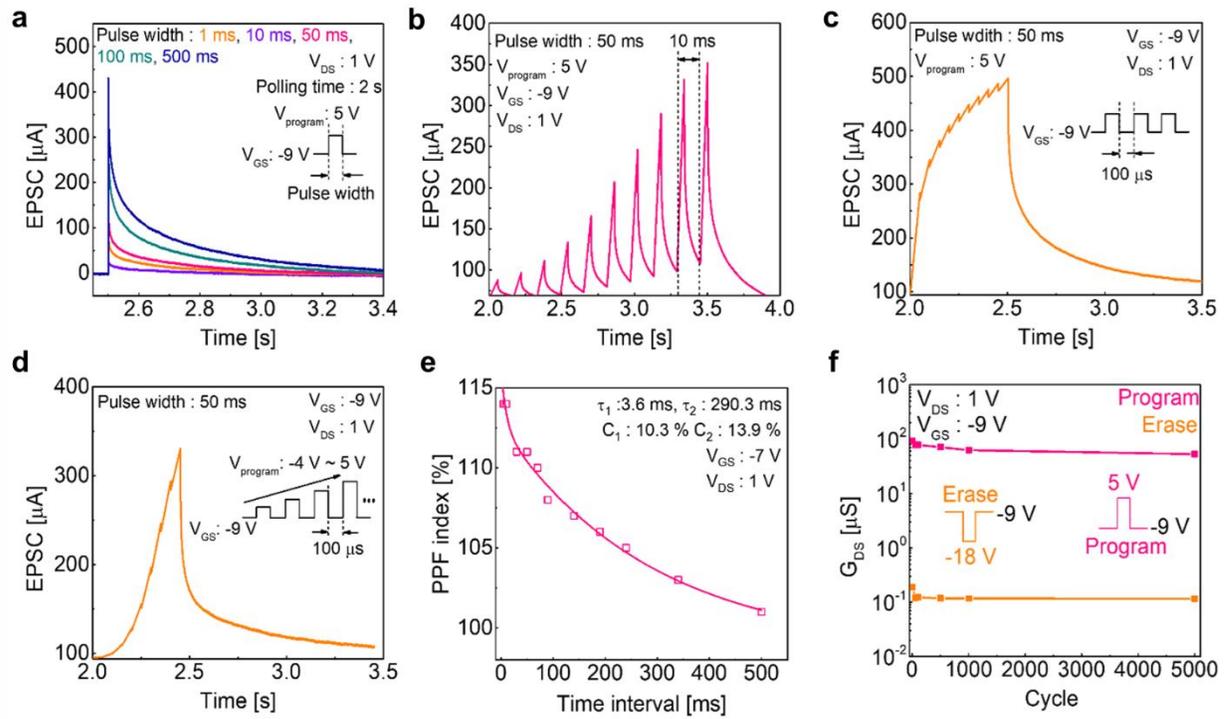


Figure 4.8 Short-term plasticity (STP) characterizations of GaN/ α -In₂Se₃ FeHEMTs. **a** Excitatory postsynaptic current (EPSC) responses for different gate input widths from 1 ms to 500 ms. Each output current of the EPSC is collected from the drain electrode of GaN/ α -In₂Se₃ MOS-HEMT while the input pulses are applied to the gate electrode. Cumulative EPSC responses for sequential pulses with a time interval of **b** 10 ms and **c** 100 μ s, and **d** with incremental amplitudes. **e** For various time intervals between two sequential pulses, ranging from 10 ms to 500 ms, the pair-pulse-facilitation (PPF) was calculated from 1st and 2nd peaks of each EPSC. **f** Endurance cycle for the program and erase operations. Responses current was measured after applying a program and erase pulses at $V_{DS} = 1$ V²⁰.

The unique properties of heterointerface between α -In₂Se₃ and AlGaIn/GaN based 2DEG enable the neuromorphic function GaN/ α -In₂Se₃ MOS-HEMT in STP. Figure 4.8a shows the

excitatory postsynaptic (EPSC) responses for pulse widths of 1–500 ms, mimicking the biological STP. To compensate for the initial polarization caused by the AlGaN barrier layer, an initial set-up pulse has been applied for 2 s. The EPSC showed a dependency on the pulse width, and the postsynaptic signal intensity has been increased as the wider pulse width has been applied. Figs. 4.8b, c show the sequential EPSC for two pulses at a time interval of 10 ms and 100 μ s, respectively. Unlike the 10-ms interval, the shorter time interval (100 μ s) allows the accumulation of input spikes. The input spikes can be accumulated by increasing the pulse amplitude (Fig. 4.8d). Figure. 4.8e shows the corresponding pair-pulse-facilitation (PPF) index plot. The time interval for extracting the parameters of the PPF index increased gradually from 10 to 500 ms. The maximum PPF index value was 114% at a time interval of 10 ms, whereas the minimum value was 100% at a time interval of 500 ms. Figure 4.8f shows the endurance times for the erase and program operations. Both operations were maintained for 5000 cycles without significant variations in the channel conductance (G_{DS}).

Owing to the high electron mobility of AlGaN/GaN HEMT, the extracted cut-off frequency (f_T) of the fabricated device is estimated to be 3.3 GHz based on the transconductance of 156 μ S and the total gate capacitance of 7.32×10^{-15} F extracted from I_{DS} vs V_{GS} and C-V measurement, respectively.

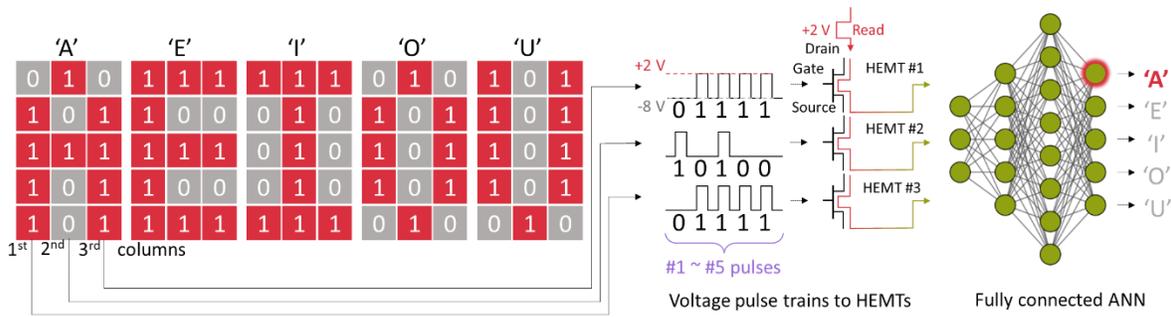


Figure 4.9 Schematic of reservoir computing (RC) based on GaN/ α -In₂Se₃ FeHEMTs for vowel

image classification (example of 'A'). 3×5 'A', 'E', 'I', 'O', and 'U' datasets are prepared, and the pixel intensities are converted to voltages ('0' \rightarrow -2 V and '1' \rightarrow +2 V). The accumulated pulses generate corresponding EPSCs at $V_D=2$ V and $V_G=-6$ V. The readout EPSCs are normalized and fed into the pre-trained fully connected layer (3-5-8-5 nodes) for classification. All pulses are 50-ms pulses with 50% duty cycle²⁰.

RC operation is feasible via multistate dipole polarization of the GaN/ α -In₂Se₃ MOS-HEMT that corresponds to the randomly distributed nodes in the reservoir layer. Furthermore, the coupled 2DEG channel allows for processing high-frequency input signals in RC system. Here, based on the described electrical and neuromorphic characteristics above, I demonstrated the RC capability of the GaN/ α -In₂Se₃ MOS-HEMT. Figure 4.9 shows a schematic of the RC system. I prepared 3×5 'A', 'E', 'I', 'O', and 'U' datasets for classification, and the pixel intensities of the datasets were converted to voltage amplitudes. Then, three 1×5 voltage pulses (corresponding to each letter) have been applied to the gate electrode of the GaN/ α -In₂Se₃ MOS-HEMT. The pixel intensities of the letter images were encoded to voltage amplitudes (0 for -2 V and 1 for 2 V). The erase, read, and program voltages were -20, 1, and 2 V, respectively. All pulse widths were 50 ms. One hundred encoded data were collected (20 per letter) and divided into 80 training and 20 test images. A 3-5-8-5 FCNN was employed with ReLU activation layers (softmax for the last output). An RMSprop optimizer was employed with an initial time-decaying learning rate of 0.001.

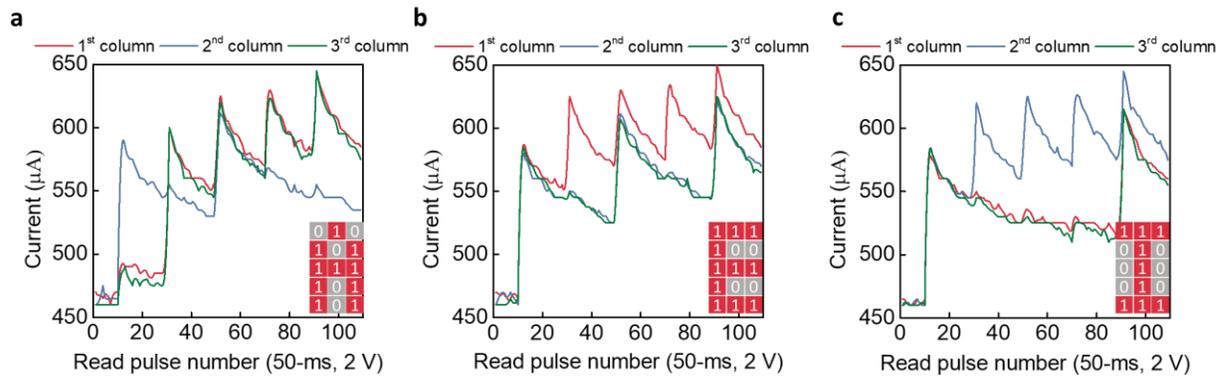


Figure 4.10 Experimental EPSC responses of letter **a** ‘A’, **b** ‘E’, and **c** ‘I’. The EPSC peaks correspond to the ‘1’ pixel intensity in each letter²⁰.

The input pulses in the letter images generated accumulated EPSC in a time domain (Figure 4.10), which is then applied as an input to the FCNN. For example, the ‘1’ pixel intensity generates the EPSC spike, whereas the ‘0’ pixel intensity relaxes the EPSC intensity. As a result, the 3×5 input matrix is nonlinearly embedded into the distinguishable 3×1 vector (at the read pulse) as an input to the FCNN, offering the miniaturization of the FCNN model for the faster classification.

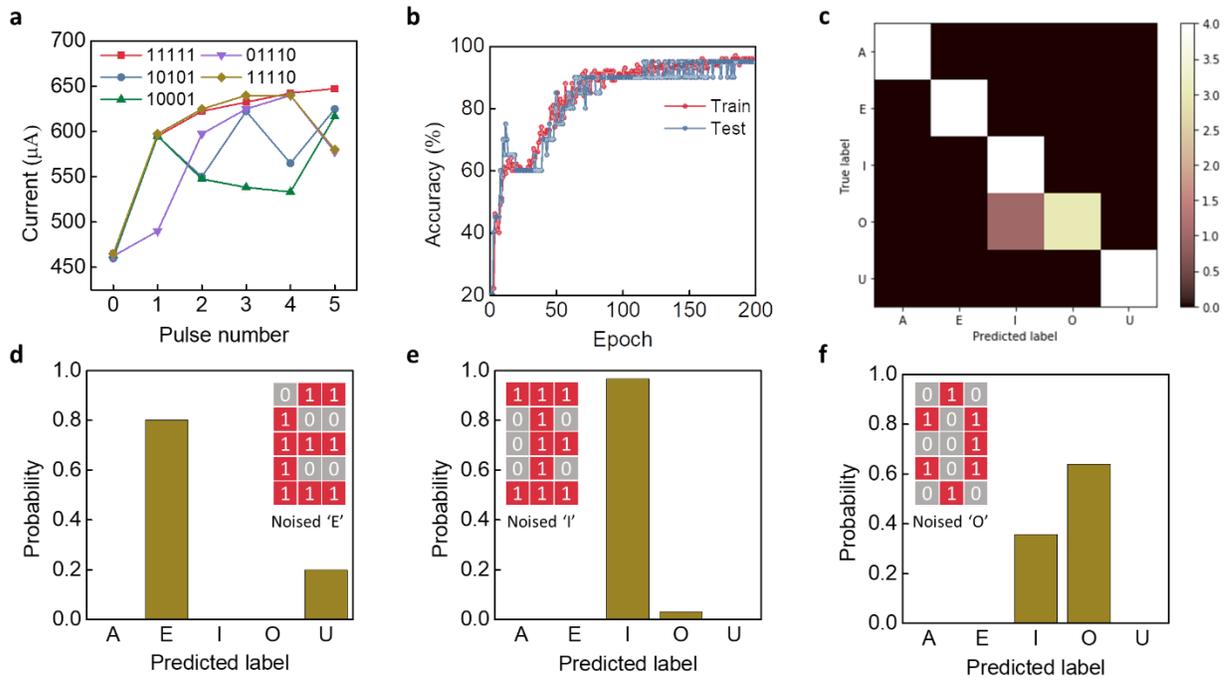


Figure 4.11 RC-based vowel classification. **a** EPSC responses with respect to each pulse number. The current values corresponding to each pulse pattern are averaged. **b** Results of RC-based classification. For each train and test procedure, up to 97% and 95% accuracies are achieved, respectively. **c** Confusion matrix of vowel image classification. Inference results for noised (modifying one pixel) inputs, **d** 'E', **e** 'I', and **f** 'O'²⁰.

The corresponding readout currents for certain patterns are shown in Figure 4.11a. The continuous '1' input signals accumulated as increased EPSC intensity, whereas '0' signals relaxed the EPSC from GaN/ α -In₂Se₃ MOS-HEMT due to its STP. I have collected the encoded data and pretrained the FCNN. Figure 4.11b shows the training and test accuracies of the RC-based neural network. The accuracy converged to approximately 96% after the 50th epoch via the continuous optimization of the learning parameters, with a slight fluctuation during the first 50 epochs. Consequently, the training and test processes achieved accuracies of up to 97% and 95%,

respectively. Figure 4.11c presents the confusion matrix for the classification. Although the classification accuracy of the 'O' is slightly degraded, the FCNN is robustly trained to infer noised images. For instance, noised 'E', 'I', and 'O' images are also classifiable via the FCNN, exhibiting the highest probability for the corresponding predictions, as shown in Figure 4.11d-f. The spontaneous STP behavior of the GaN/ α -In₂Se₃ MOS-HEMT can perform RC, encoding the input matrix as a vector for an effective classification task.

4.4 Artificial Oculomotor Dynamics

The somatosensory system in biological organisms is responsible for detecting and responding to various external stimuli, including vision, sound, odor, pressure, and temperature²¹⁰⁻²¹². To achieve such reactions, external stimuli detected by the afferent nerve (sensory neurons) are first transferred to the central nervous system (CNS)²¹³. The CNS then generates an action potential for the efferent nerve (motor nerve) that actuates the target muscle through the neuromuscular junction (NMJ). The NMJ is a unique and essential synaptic connection between the efferent nerve and muscle fibers that triggers motion via the transmission of action potentials through it²¹⁴. Consequently, the stimulated muscle fibers contract and relax, becoming capable of triggering macro-motions. Macro-motions generally require substantially higher energy than that required for computation. Therefore, it has been challenging to emulate NMJs to fulfill both synaptic plasticity and the capability to drive large amounts of energy²¹⁵⁻²¹⁸.

To address these challenges, here I demonstrate synaptic transistors by heterogeneously integrating a CIPS ferroelectric membrane as a gate dielectric material with an AlGaN/GaN HEMT. The programmable transconductance of ferroelectric transistors allows for artificial synaptic plasticity, which is attributed to the polarization of the ferroelectric gate dielectric layer.

Among many ferroelectric materials, including Zr-doped $\text{Hf}_{1-x}\text{Zr}_x\text{O}^{123}$, $\text{PbZr}_{0.52}\text{Ti}_{0.48}\text{O}_3^{219}$, BaTiO_3^{220} , and BiFeO_3^{221} , CIPS is a unique two-dimensional (2D) van der Waals (vdW) material with out-of-plane ferroelectricity that offers high integrability, flexibility, and responsivity to electrical signals^{45,222–226}. The out-of-plane ferroelectricity in the CIPS is attributed to the ionic dynamics of the Cu and In cations, which are vertically displaced in the sulfur framework²²⁷.

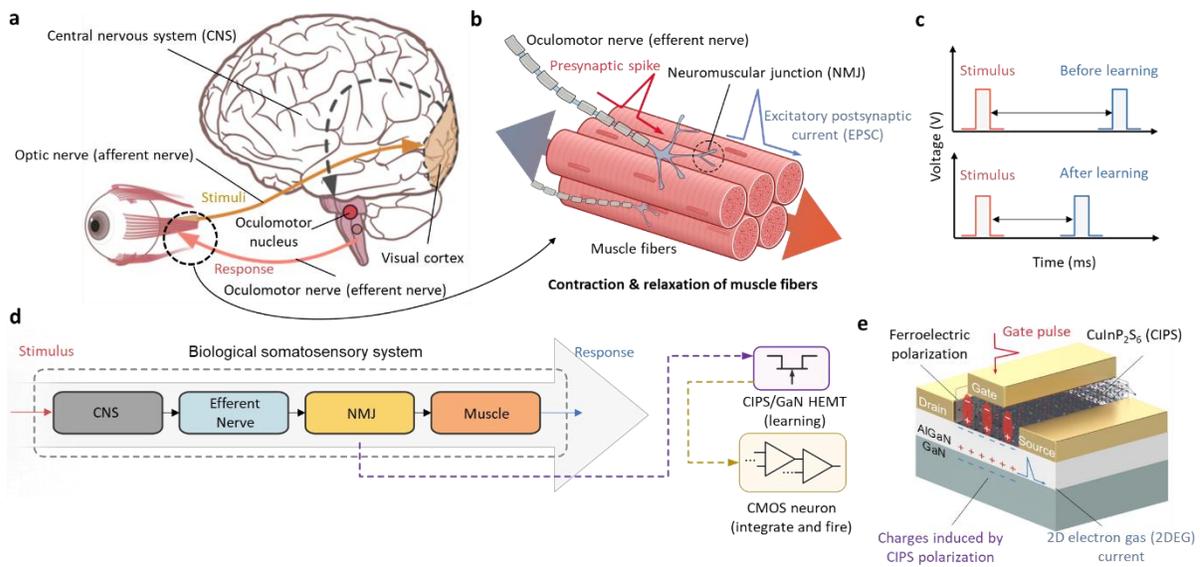


Figure 4.12 Schematic illustration of mimicking neuromuscular junction (NMJ). **a** Schematic illustration of oculomotor system. The stimulus transmitted through the optic nerve propagates towards the target muscle via the central nervous system (CNS), oculomotor nucleus/nerve, and NMJ. **b** Biological NMJ. The NMJ connects the terminal of the oculomotor nerve with muscle fibers. **c** Emulation of the NMJ for fast stimulus enhancement. The fast response is emulated by the programmable GaN/CIPS FeHEMT and integrate-fire unit (IFU). **d** Emulation of NMJ dynamics. The CNS generates an action potential on the efferent nerve that actuates the target muscle through the NMJ. The trained GaN/CIPS FeHEMT with an IFU allows for enhanced reflexes. **e** Schematic of GaN/CIPS FeHEMT structure. The polarization of the GaN/CIPS

FeHEMT mimics synaptic plasticity and generates a programmable two-dimensional electron gas (2DEG) current that serves as an artificial EPSC.

Figure 4.12a shows a schematic of the biological oculomotor system. When visual stimuli via the optic nerve—categorized as an afferent nerve—trigger the CNS²¹³, the CNS generates an action potential via the oculomotor nucleus in the midbrain. The potential is then transmitted to the NMJ via the oculomotor nerve—categorized as an efferent nerve—and finally actuates the target muscle²²⁸. Figure 4.12b illustrates the neuromuscular dynamics driven by stimuli from the CNS. The NMJ connects the oculomotor nerve to the muscle fibers, generating an excitatory postsynaptic current (EPSC) that triggers contraction and relaxation of the muscle fibers²²⁹. Repeated stimuli reinforce synaptic connectivity for signal transmission between NMJs and muscle fibers in the somatosensory system, resulting in enhanced reflexes²¹⁵. Trained athletes' quick reactions, such as sprint starts, swim starts, and dribbling, are good examples of enhanced conscious responses. This study demonstrated the enhancement using artificial synaptic plasticity of the ferroelectric field-effect transistor (FeFET) through a training process (Fig. 4.12c).

I used a ferroelectric CIPS membrane integrated into a GaN HEMT as a gate dielectric layer (Fig. 4.12d) to emulate neuromuscular synaptic plasticity and motion driving capability. The ferroelectric polarization of CIPS enables the GaN HEMT to be programmable, which is a critical function of biological STP and LTP. I also emulated the enhanced response by integrating artificial synapses based on the GaN/CIPS FeHEMT as an NMJ with an IFU that accumulated the incoming pulse train and generated a single firing spike.

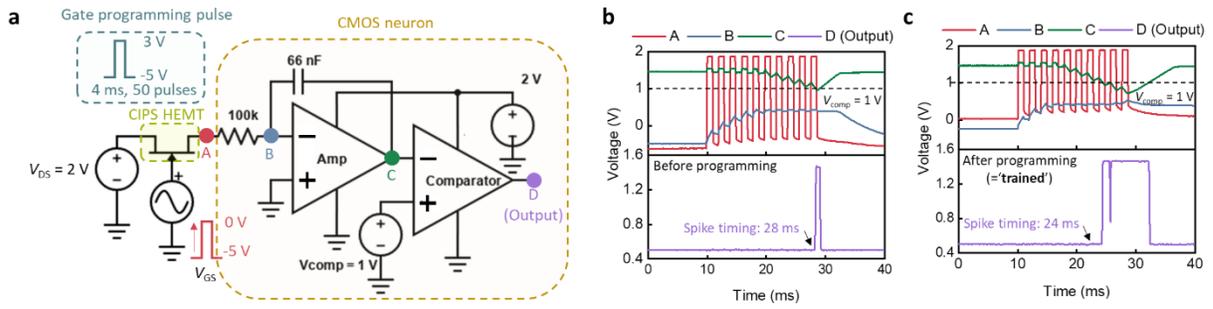


Figure 4.13 Emulation of biological enhanced reflexes based on GaN/CIPS FeHEMT. **a** Schematic illustration of integrated GaN/CIPS FeHEMT with an IFU. Output spiking responses **b** before programming and **c** after programming (trained). After the training process, a 14% faster response (output spike) was achieved (24 ms) compared to the response without repeated training (28 ms).

The combination of artificial LTP and STP enables spatiotemporal processing, which provides a pathway for emulating the enhanced reflexes. Figure 4.13a shows the GaN/CIPS FeHEMT connected to the IFU that accumulates incoming spikes and generates different spike timings with respect to the programmed states of the GaN/CIPS FeHEMT. The IFU includes integrating and firing building blocks. The integrating building block includes an amplifier, resistor, and capacitor that accumulate the input voltage signals. The following comparator serves as the firing building block, generating an output spike when the accumulated input signal exceeds the predefined threshold voltage (V_{comp} , 1 V). I applied repeated external stimulation (50 gate pulses), referred to as the training process, to strengthen the synaptic connections toward the target muscle. Each pulse features a 50% duty cycle, a baseline of -5 V, an amplitude of 8 V, and a pulse width of 4 ms. The transconductance of the trained GaN/CIPS FeHEMT increased, generating larger spikes and triggering a faster corresponding spike from the comparator. After repeated training procedures, a smaller number of external stimulations (10 gate pulses) generated a higher

EPSC that mimicked the increased connectivity between the biological muscle and the NMJ. The EPSC generated by the GaN/CIPS FeHEMT is accumulated by the IFU; then, by firing a final output spike upon exceeding the threshold V_{comp} (1 V), as shown in Fig. 4.13b and c, a 14% faster response (output spike) was achieved (24 ms) compared to the response without repeated training (28 ms). The redundant floating effect at node A after the discharging process could also increase the baseline of the voltage at point A to the positive voltage, but most of the upshift was attributed to the training process.

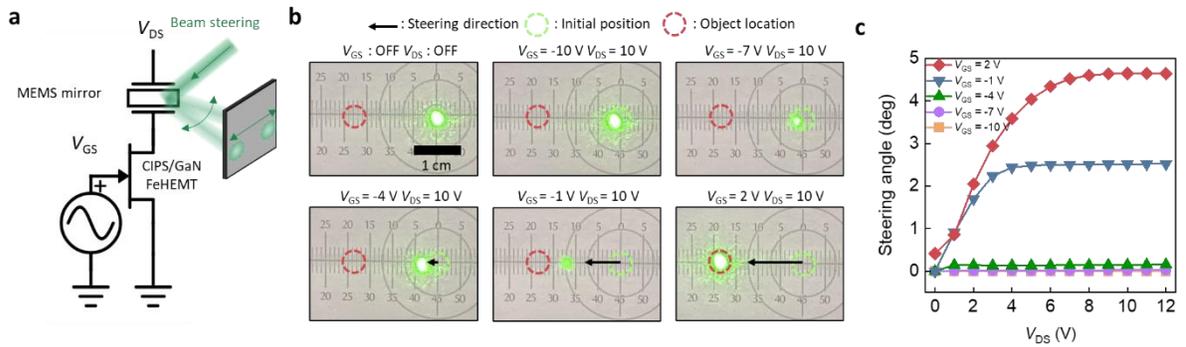


Figure 4.14 Integration of microelectromechanical system (MEMS) mirror and GaN/CIPS FeHEMT. **a** Direct amplifier-less operation of actuator via GaN/CIPS FeHEMT. The beam steering emulates the adduction and abduction motions. The two-terminal microelectromechanical system (MEMS) mirror is connected to the drain of the GaN/CIPS FeHEMT. **b** Experimental results of beam steering. The displacement of the laser beam is dependent on the steering angle of the MEMS mirror, which actuation is driven by the I_{DS} of the GaN/CIPS FeHEMT. **c** Voltage-dependent steering angles. The steering angle is proportional to I_{DS} , modulated by V_{GS} .

Based on the biological plausibility of LTP and STP, the GaN/CIPS FeHEMT exhibits programmable temporal dynamics with a high enough output current to drive external actuators. To demonstrate the feasibility of this direct actuation, I performed experiments of oculomotor dynamics using the GaN/CIPS FeHEMT-based artificial NMJ for in-situ mechanical object tracking. The oculomotor dynamics are mediated by multiple extraocular muscles, including the obliques and recti. The lateral and medial recti govern the lateral rotation of the eyeballs via adduction and abduction²³⁰. I integrated a MEMS mirror (S12237-03P, Hamamatsu) directly with the GaN/CIPS FeHEMT to emulate these extraocular movements (Fig. 4.14a). The movements were visualized by exhibiting displacements of the reflected laser beam on the MEMS mirror towards the target location, depending on the applied V_{GS} to the GaN/CIPS FeHEMT (Fig. 4.14b). I confirmed that the high driving output current of the GaN/CIPS FeHEMT was sufficient to operate the MEMS mirror without additional amplifying circuitry (Fig. 4.14c).

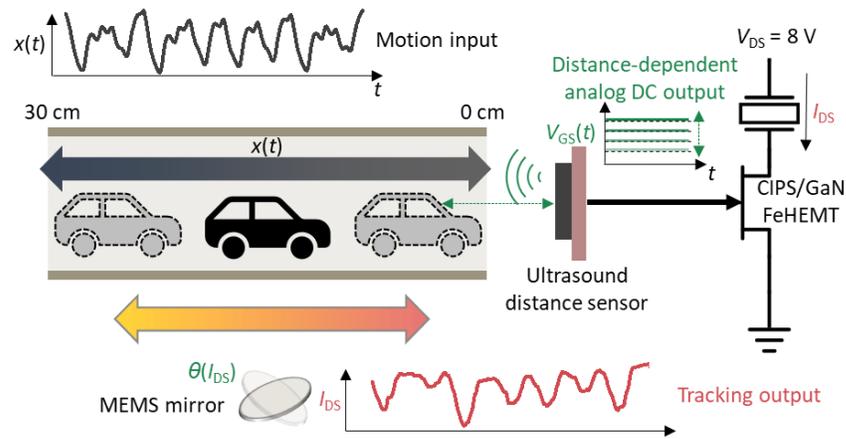


Figure 4.15 Schematic of in-situ mechanical object tracking. The tracking system additionally incorporates an ultrasound distance sensor integrated to the gate of the GaN/CIPS FeHEMT. The

position of the car is captured by the ultrasound distance sensor and converted to the analog voltage that drives I_{DS} to steer the mirror.

By directly integrating a sensor into the system, the synaptic plasticity of the GaN/CIPS FeHEMT can be utilized to improve the in-situ object-tracking functionality (Fig. 4.15). I employed the programmable robot car (Zumo Robot v1.2, Pololu) as a tracked object that moves back and forth based on the chaotic time series input. The gate of the GaN/CIPS FeHEMT was connected externally with the ultrasound sensor (URM09, DFRobot) that detects the distance between the sensor and object ($x(t)$). $x(t)$ ranges from 0 to 30 cm corresponding to the chaotic oscillation range of the robot car. The distance-dependent analog voltage from the ultrasound sensor was applied as V_{GS} to the GaN/CIPS FeHEMT. The output drain current (I_{DS}) of the GaN/CIPS FeHEMT was dependent on the movement of the target object and modulated the steering angle ($\theta(I_{DS})$) of the MEMS mirror for in-situ object tracking.

I employed the chaotic Mackey-Glass time-series data for a moving signal, $x(t)$, generated by the discrete equation as shown in Eq. (3)²³¹⁻²³⁴:

$$x(t + 1) = c \cdot x(t) + \left(\frac{a \cdot x(t-d)}{b \cdot x(t-d)^e} \right) \quad (3)$$

where a (0.05), b (0.2), c (0.9), d (17), and e (10) are constants. The first 100 data points were removed to achieve a stable oscillating chaotic signal. The remaining 400 data points were then normalized between 0 and 1. In this experiment, the output of the ultrasound distance sensor was between 0 V and 0.8 V, but further distance ranging is possible when the larger space is available, reaching the operating voltage of the ultrasound distance sensor (5.5 V). Thus, within

the maximum possible voltage range of the sensor (5.5 V), I additionally amplified the output signal (59-179, Edmund) that reaches up to 4 V only for the visualization purpose. The I_{DS} output signals were interpolated from 200 to 400 data points and normalized, to match the dimensions with that of $x(t)$. The tracking spot was illuminated by a 532-nm diode laser (DJ532, Thorlabs).

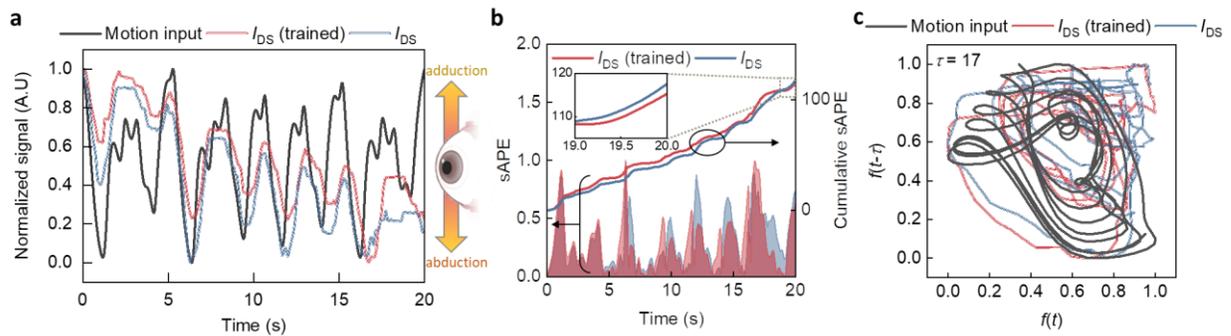


Figure 4.16 Emulation of biological oculomotor dynamics. **a** Motion input and tracking output time-series data. The trained GaN/CIPS FeHEMT shows a higher I_{DS} (red line) that is more closely matched with the motion input signal for effective object tracking through beam steering. **b** Symmetric absolute percentage error (sAPE) plots of tracking output (I_{DS}) with respect to motion input. The trained GaN/CIPS FeHEMT drives larger initial I_{DS} , improving the cumulative sAPE of tracking. Inset: Zoom-in plot of cumulative sAPE. **c** Phase plots of motion input and tracking output. The trained system shows closer phase tracking of the motion input.

Figure 4.16a confirms the in-situ object tracking capability using the assembled system consisting of the ultrasound sensor, GaN/CIPS FeHEMT, and MEMS mirror. Here, I showed that the motion of the robot car stimulated the dynamics of adduction and abduction using an artificial NMJ, depending on its moving directions. The time-dependent information of the object position

($x(t)$, black line) was converted into an electrical input signal of the GaN/CIPS FeHEMT by the ultrasound sensor. The mechanical steering ($\theta(I_{DS})$) of the MEMS mirror was triggered by the ultrasound sensor and the assembled GaN/CIPS FeHEMT, corresponding to the oculomotor controls by the lateral and medial recti. I examined the GaN/CIPS FeHEMT before and after the training to achieve the enhanced reflexes (blue and red line, respectively), emulating the enhancement of the synaptic connections in biological neural systems. The trained GaN/CIPS FeHEMT exhibited increased transconductance, resulting in enhanced I_{DS} for an efficient object tracking process. Figure. 4.16b shows the error plots of I_{DS} with and without the training process. I employed a symmetric absolute error (sAPE)²³⁵ to avoid undefined output for zero values as shown in Eq. (4).

$$\text{sAPE} = \frac{|x(t) - I_{DS}^*(t)|}{(x(t) + I_{DS}^*(t))/2} \quad (4)$$

where I_{DS}^* is normalized I_{DS} between 0 and 1. As time progresses, the cumulative sAPE decreases in the enhanced state (red line), which is attributed to the programmed transconductance of the GaN/CIPS FeHEMT. Figure 4.16c shows the phase plots of each signal (represented as $f(t)$) after the baseline correction for I_{DS} . I_{DS} with the training process shows closer phase tracking with the motion input.

Analogous to conventional GaN HEMTs, the proposed GaN/CIPS FeHEMT also potentially attains nanosecond operation by etching the AlGaN barrier layer to minimize the surface potential and by using the thinner CIPS membrane to alleviate the voltage drops at CIPS^{205,236}. The direct actuation of the mechanical platform using the artificial NMJ provides a wide range of neuromorphic sensing-to-action applications, including time-of-flight ranging^{44,59}, in-situ object tracking¹²⁸, in-sensor computing^{132,133}, and human-computer interaction (HCI)²³⁷.

In this study, I achieved a normalized output current of 200 mA/mm with the GaN/CIPS FeHEMT, which is notably greater than that of recently reported synaptic transistors^{16,34,36,51,238–243}. Therefore, the GaN/CIPS FeHEMT is potentially deployable as an artificial NMJ in robotic systems to operate mechanical actuators that require a milliampere-scale driving current for macro-motion. Moreover, the 2DEG transport channel of HEMT offers a GHz-range frequency response that can be coupled with integrated ferroelectricity for reconfigurable RF applications^{20,122}.

4.5 Conclusion

I have demonstrated ferroelectric GaN/ α -In₂Se₃ MOS-HEMTs with superior electrical performance by implementing a semiconducting α -In₂Se₃ layer to induce C_{dep} , which maximizes the effect of its ferroelectricity. The lowest SS (8 mV/dec) and large ΔV (2 V) were obtained owing to the polarization matching effect. The ferroelectric polarization and $\Delta\Phi_{eff}$ was confirmed in the GaN/ α -In₂Se₃ MES-HEMT structure, which sufficiently controlled the 2DEG conductivity. In addition, the self-aligned etching process for the α -In₂Se₃ layer confines polarization in the vertical direction, which can further enhance ferroelectricity confirmed both by experiment and simulation using physics-based TCAD. Finally, I demonstrated the GaN/ α -In₂Se₃ MOS-HEMT for neuromorphic computing applications by using its STP. RC was conducted using the measured and extracted characteristics of GaN/ α -In₂Se₃ MOS-HEMT. The recognition accuracy for RC reached 95% during the test process. My study suggests that the GaN/ α -In₂Se₃ MOS-HEMT is beneficial for demonstrating fast switching neuromorphic applications owing to the combination of semiconducting property of α -In₂Se₃ and the high carrier mobility of GaN HEMT²⁰.

I also applied a CIPS ferroelectric membrane integrated with a GaN FeHEMT for artificial oculomotor dynamics. The high output current achieved by the AlGaIn/GaN 2DEG enabled amplifier-less actuation for the adduction and abduction motions. The polarization of the CIPS tuned the output current of the GaN/CIPS FeHEMT using the training pulse at its gate node. This non-volatile artificial synaptic device was connected to the CMOS-based efferent system that integrates and fires the spike to actuate the mechanical platform, such as a mechanical object tracker. The temporal dynamics of the GaN/CIPS FeHEMT were enhanced by the training process, analogous to the biological stimulus response. I believe that the GaN/ α -In₂Se₃ and GaN/CIPS FeHEMTs have immense potential to process high-frequency input signals and to realize functional bio-inspired elements for artificial muscles and smart robotic applications.

Chapter 5

Future Works

5.1 Second-Order Spiking Neuromorphic Vision

In this thesis, the focus has primarily been on non-volatile memristors to achieve 2D/3D neuromorphic vision, inspired by the biological vision system. However, volatile memristors, such as ionic memristors utilizing copper, silver, and other organic compounds, as well as Mott memristors, have also been extensively used for temporal dynamic processing^{33,35,209,244}. Mott memristors utilize a Mott insulator, characterized by strong electron-electron interactions and Joule heating. It is interesting to note that ionic volatile memristors exhibit STP due to their capacitive behavior, while Mott memristors demonstrate an abrupt transition from insulator to metal without temporal relaxation, making them suitable for ultrafast computing paradigms and activation functions in ANNs⁸.

Figure 5.1 illustrates the first-, second-, and third-order information processing capabilities of various memristors, including Mott memristors⁸. Non-volatile memristors such as redox and ferroelectric memristors enable the realization of simple first-order synaptic devices. Second-order synaptic devices incorporate both LTP and STP, allowing for non-volatile but dynamic processing capabilities. In the case of neuronal devices, a single Mott memristor exhibits an abrupt transition from HRS to LRS, mimicking the firing behavior of neurons and the activation function in ANNs. When combined with internal capacitance and a load resistor, the neuronal Mott insulator exhibits higher-order behavior, including self-oscillations. Unlike the Mott memristor's

first-order behavior, the internal capacitance and the Mott transition introduce two state variables: charge distribution and temperature (Joule heating), respectively. Further advancements in nanoscale fabrication processes allow for additional state variables, leveraging the ultrafast Mott transition in nanocircuit components for neuromorphic computing purposes⁸.

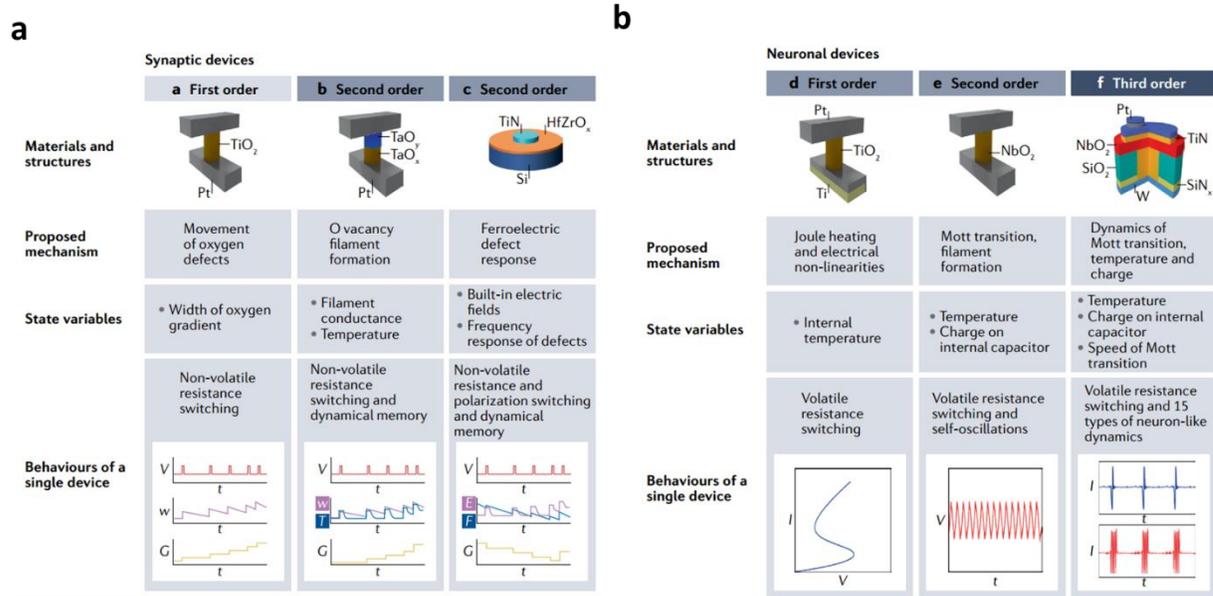


Figure 5.1 Artificial **a** synaptic and **b** neuronal devices for first-, second-, and third-order information processing. In this thesis, nonvolatile memristors have mainly been discussed to realize 2D/3D neuromorphic vision⁸.

5.2 Ultrasound RToF

I have successfully demonstrated an unprecedented memristor-based 3D sensing approach, known as the RToF principle, which is associated with the biological STDP behavior. While light-based RToF systems using ToF with light as the source flux provide high-resolution 3D imaging platforms like LiDAR, the challenge lies in minimizing the parasitic capacitance and inductance induced by metal wires and interconnections in the fabricated RToF chips for compact circuit

design²⁰⁵. In my proof-of-concept implementation of the light-based RToF, the depth resolution was compromised (> 15 cm) due to neglecting these parasitic issues⁴⁴.

In contrast, ultrasound can also serve as a source flux for RToF, especially in sonar navigation and distance sensors for automotive applications^{245–247}. This is because sound propagates at a much lower speed compared to light. Therefore, an RToF system based on ultrasound as the source flux can be employed as an alternative, avoiding the need for compact circuit design, and alleviating the parasitic issues encountered in light-based RToF systems.

Another challenge is the integration of signal pre-processing modules that effectively handle noisy and oscillating raw ultrasound sensory signals²⁴⁸. This integration includes components such as amplifiers, low-pass filters, ADCs, and comparators²⁴⁹. In future work, the goal is to integrate these circuit components with fabricated nonvolatile memristors and thin-film ultrasound transducers/receivers. By utilizing the ultrasound-based RToF approach, in contrast to the current light-based RToF, stable and reliable depth ranging can be achieved, even when implemented on flexible substrates, enabling in-sensor 3D depth ranging and imaging based on the neuromorphic domain.

REFERENCE

1. Krizhevsky, A. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *In Advances in neural information processing systems* 1097–1105 (2012).
2. Silver, D. *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* (2016) doi:10.1038/nature16961.
3. Shafiee, A. *et al.* ISAAC: A Convolutional Neural Network Accelerator with in-Situ Analog Arithmetic in Crossbars. *SIGARCH Comput. Archit. News* **44**, 14–26 (2016).
4. Giordano, M. *et al.* CHIMERA: A 0.92 TOPS, 2.2 TOPS/W edge AI accelerator with 2 MByte on-chip foundry resistive RAM for efficient training and inference. in *2021 Symposium on VLSI Circuits* 1–2 (IEEE, 2021).
5. Waldrop, M. M. The chips are down for Moore’s law. *Nature News* **530**, 144 (2016).
6. Yole Intelligence. *Neuromorphic Computing and Sensing 2021. Optics and Photonics News* vol. 33 <https://www.yolegroup.com/product/report/neuromorphic-computing-and-sensing-2021/> (2021).
7. Schuman, C. D. *et al.* Opportunities for neuromorphic computing algorithms and applications. *Nat Comput Sci* **2**, 10–19 (2022).
8. Kumar, S., Wang, X., Strachan, J. P., Yang, Y. & Lu, W. D. Dynamical memristors for higher-complexity neuromorphic computing. *Nat Rev Mater* 1–17 (2022).
9. Medathati, N. V. K., Neumann, H., Masson, G. S. & Kornprobst, P. Bio-inspired computer vision : Towards a synergistic approach of artificial and biological vision. **150**, 1–30 (2016).
10. Taeckens, E., Dong, R. & Shah, S. A Biologically Plausible Spiking Neural Network for Decoding Kinematics in the Hippocampus and Premotor Cortex. *bioRxiv* 2011–2022 (2022).
11. Woźniak, S., Pantazi, A., Bohnstingl, T. & Eleftheriou, E. Deep learning incorporating biologically inspired neural dynamics and in-memory computing. *Nat Mach Intell* **2**, 325–336 (2020).
12. Cai, F. *et al.* A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations. *Nat Electron* **2**, 290–299 (2018).
13. Yang, K., Joshua Yang, J., Huang, R. & Yang, Y. Nonlinearity in memristors for neuromorphic dynamic systems. *Small Science* **2**, 2100049 (2022).
14. Cao, J. *et al.* Emerging dynamic memristors for neuromorphic reservoir computing. *Nanoscale* **14**, 289–298 (2022).
15. Wang, R. *et al.* Bipolar Analog Memristors as artificial synapses for neuromorphic computing. *Materials* (2018) doi:10.3390/ma11112102.
16. Ren, Y. *et al.* Gate-tunable synaptic plasticity through controlled polarity of charge trapping in fullerene composites. *Adv Funct Mater* **28**, 1805599 (2018).
17. Prezioso, M., Merrih Bayat, F., Hoskins, B., Likharev, K. & Strukov, D. Self-Adaptive Spike-Time-Dependent Plasticity of Metal-Oxide Memristors. *Sci Rep* **6**, 2–7 (2016).
18. Wang, Z. *et al.* Resistive switching materials for information processing. *Nat Rev Mater* **5**, 173–195 (2020).
19. Moon, J. *et al.* Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat Electron* **2**, 480–487 (2019).
20. Yang, J. Y. *et al.* Reconfigurable Physical Reservoir in GaN/ α -In₂Se₃ HEMTs Enabled by Out-of-Plane Local Polarization of Ferroelectric 2D Layer. *ACS Nano* **17**, 7695–7704 (2023).
21. Yi, W. *et al.* Biological plausibility and stochasticity in scalable VO₂ active memristor neurons. *Nat Commun* **9**, 4661 (2018).
22. Peng, X., Huang, S., Jiang, H., Lu, A. & Yu, S. DNN+ NeuroSim V2. 0: An end-to-end benchmarking framework for compute-in-memory accelerators for on-chip training. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **40**, 2306–2319

- (2020).
23. Li, X. *et al.* Power-efficient neural network with artificial dendrites. *Nat Nanotechnol* **15**, 776–782 (2020).
 24. Lin, J. *et al.* A memristor-based leaky integrate-and-fire artificial neuron with tunable performance. *IEEE Electron Device Letters* **43**, 1231–1234 (2022).
 25. Yang, J.-Q. *et al.* Leaky integrate-and-fire neurons based on perovskite memristor for spiking neural networks. *Nano Energy* **74**, 104828 (2020).
 26. Merolla, P. A. *et al.* A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science (1979)* **345**, 668–673 (2014).
 27. Izhikevich, E. M. Simple Model of Spiking Neurons. **14**, 1569–1572 (2003).
 28. Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T. & Maida, A. Deep learning in spiking neural networks. *Neural Networks* **111**, 47–63 (2019).
 29. Strukov, D. B., Snider, G. S., Stewart, D. R. & Williams, R. S. The missing memristor found. *Nature* **453**, 80–83 (2008).
 30. Kumar, S. *et al.* Oxygen migration during resistance switching and failure of hafnium oxide memristors. *Appl Phys Lett* **110**, 1–5 (2017).
 31. Vaidya, J. *et al.* A three-terminal non-volatile ferroelectric switch with an insulator–metal transition channel. *Sci Rep* **12**, 1–7 (2022).
 32. Chen, W. H. *et al.* CMOS-integrated memristive non-volatile computing-in-memory for AI edge processors. *Nat Electron* **2**, 420–428 (2019).
 33. Park, S. O., Jeong, H., Park, J., Bae, J. & Choi, S. Experimental demonstration of highly reliable dynamic memristor for artificial neuron and neuromorphic computing. *Nat Commun* **13**, (2022).
 34. Xie, Z. *et al.* All-Solid-State Vertical Three-Terminal N-Type Organic Synaptic Devices for Neuromorphic Computing. *Adv Funct Mater* **32**, 2107314 (2022).
 35. Deng, X. *et al.* A flexible mott synaptic transistor for nociceptor simulation and neuromorphic computing. *Adv Funct Mater* **31**, 2101099 (2021).
 36. Kim, S. *et al.* Pattern recognition using carbon nanotube synaptic transistors with an adjustable weight update protocol. *ACS Nano* **11**, 2814–2822 (2017).
 37. Dai, S. *et al.* Recent advances in transistor-based artificial synapses. *Adv Funct Mater* **29**, 1903700 (2019).
 38. Yeon, H. *et al.* Alloying conducting channels for reliable neuromorphic computing. *Nat Nanotechnol* **15**, 574–579 (2020).
 39. Hus, S. M. *et al.* Observation of single-defect memristor in an MoS₂ atomic sheet. *Nat Nanotechnol* **16**, 58–62 (2021).
 40. Völkel, L. *et al.* Resistive Switching and Current Conduction Mechanisms in Hexagonal Boron Nitride Threshold Memristors with Nickel Electrodes. *Adv Funct Mater* 2300428 (2023).
 41. Ismail, M., Abbas, H., Choi, C. & Kim, S. Stabilized and RESET-voltage controlled multi-level switching characteristics in ZrO₂-based memristors by inserting a-ZTO interface layer. *J Alloys Compd* **835**, 155256 (2020).
 42. Bousoulas, P., Sakellaropoulos, D. & Tsoukalas, D. Tuning the analog synaptic properties of forming free SiO₂ memristors by material engineering. *Appl Phys Lett* **118**, (2021).
 43. Abbas, Y. *et al.* Compliance-free, digital SET and analog RESET synaptic characteristics of sub-tantalum oxide based neuromorphic device. *Sci Rep* **8**, 1228 (2018).
 44. Park, M. *et al.* Neuron-Inspired Time-of-Flight Sensing via Spike-Timing-Dependent Plasticity of Artificial Synapses. *Advanced Intelligent Systems* **n/a**, 2100159 (2021).
 45. Ryu, H., Xu, K., Li, D., Hong, X. & Zhu, W. Empowering 2D nanoelectronics via ferroelectricity. *Appl Phys Lett* **117**, 80503 (2020).
 46. Lencer, D., Salinga, M. & Grabowski, B. A map for phase-change materials. *Nat Mater* (2008) doi:10.1038/nmat2330.
 47. Lee, S. H., Jung, Y. & Agarwal, R. Highly scalable non-volatile and ultra-low-power phase-change nanowire memory. *Nat Nanotechnol* (2007) doi:10.1038/nnano.2007.291.

48. Manipatruni, S. *et al.* Scalable energy-efficient magnetoelectric spin-orbit logic. *Nature* **565**, 35–42 (2019).
49. Rao, M. *et al.* Thousands of conductance levels in memristors integrated on CMOS. *Nature* **615**, 823–829 (2023).
50. Sebastian, A., Gallo, M. Le, Khaddam-aljameh, R. & Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat Nanotechnol* **15**, 529–544 (2020).
51. Ning, H. *et al.* An in-memory computing architecture based on a duplex two-dimensional material structure for in situ machine learning. *Nat Nanotechnol* 1–8 (2023).
52. Wang, Z. *et al.* Reinforcement learning with analogue memristor arrays. *Nat Electron* **2**, 115–124 (2019).
53. Wan, W. *et al.* A compute-in-memory chip based on resistive random-access memory. *Nature* **608**, 504–512 (2022).
54. Lee, H. S. *et al.* Efficient Defect Identification via Oxide Memristive Crossbar Array Based Morphological Image Processing. *Advanced Intelligent Systems* **2000202**, 2000202 (2020).
55. Merced-grafals, E. J., Dávila, N., Ge, N., Williams, R. S. & Strachan, J. P. Repeatable , accurate , and high speed multi- level programming of memristor 1T1R arrays for power ef fi cient analog computing applications. *Nanotechnology* **27**, (2016).
56. Rajendran, B., Sebastian, A., Schmuker, M., Srinivasa, N. & Eleftheriou, E. Low-Power Neuromorphic Hardware for Signal Processing Applications: A review of architectural and system-level design approaches. *IEEE Signal Process Mag* **36**, 97–110 (2019).
57. Hebb, D. O. The Organization of Behavior : A Neuropsychological Theory. **50**, 0–1 (2018).
58. Brand, T. Time of Flight System for Distance Measurement and Object Detection. *Çevrimiçi*. Available: <https://www.analog.com/en/technicalarticles/tof-system-for-distance-measurement-and-object-detection.html>. [Erişildi: 10 05 2021] (2021).
59. Park, M. *et al.* Hetero-integration enables fast switching time-of-flight sensors for light detection and ranging. *Sci Rep* **10**, 2764 (2020).
60. Brinkløv, S., Fenton, M. B. & Ratcliffe, J. M. Echolocation in oilbirds and swiftlets. *Front Physiol* **4 MAY**, 1–12 (2013).
61. Novacek, M. J. Oldest Known Bats. 22–23 (1985).
62. Benoit-Bird, K. J. & Au, W. W. L. Automatic gain control in the echolocation system of dolphins. *Nature* **423**, 861–863 (2003).
63. Zhou, X. *et al.* Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nat Commun* **4**, 1–6 (2013).
64. Bhandari, A. & Raskar, R. Signal Processing for Time-of-Flight Imaging Sensors: An introduction to inverse problems in computational 3-D imaging. *IEEE Signal Process Mag* **33**, 45–58 (2016).
65. Horaud, R., Hansard, M., Evangelidis, G. & Ménier, C. An overview of depth cameras and range scanners based on time-of-flight technologies. *Mach Vis Appl* **27**, 1005–1020 (2016).
66. Li, L. Time-of-Flight Camera – An Introduction. (2014).
67. Otoole, M., Lindell, D. B. & Wetzstein, G. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature* **555**, 338–341 (2018).
68. Kirmani, A. *et al.* First-Photon Imaging. **343**, 58–62 (2014).
69. McCarthy, A. *et al.* Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detection. *Opt Express* **21**, 8904 (2013).
70. Schwarz, B. Lidar: Mapping the world in 3D. *Nat Photonics* **4**, 429–430 (2010).
71. Van Tricht, K. *et al.* Clouds enhance Greenland ice sheet meltwater runoff. *Nat Commun* **7**, (2016).
72. Pedreros Bustos, F. *et al.* Remote sensing of geomagnetic fields and atomic collisions in the mesosphere. *Nat Commun* **9**, (2018).
73. Sutfin, N. A. & Wohl, E. Elevational differences in hydrogeomorphic disturbance regime influence sediment residence times within mountain river corridors. *Nat Commun* **10**, 1–14

- (2019).
74. McManamon, P. F. Errata: Review of lidar: a historic, yet emerging, sensor technology with rich phenomenology. *Optical Engineering* **51**, 089801–1 (2012).
 75. Glaser, J. How GaN Power Transistors Drive High-Performance Lidar. *IEEE Power Electronics Magazine* **4**, 25–35 (2017).
 76. High breakdown voltage achieved on AlGaIn/GaN HEMTs with integrated slant field plates. *IEEE Electron Device Letters* **27**, 713–715 (2006).
 77. Zeng, F. *et al.* A Comprehensive Review of Recent Progress on GaN High Electron Mobility Transistors: Devices, Fabrication and Reliability. *Electronics (Basel)* **7**, 377 (2018).
 78. Russo, S. & Di Carlo, A. Influence of the source-gate distance on the AlGaIn/GaN HEMT performance. *IEEE Trans Electron Devices* **54**, 1071–1075 (2007).
 79. Oka, T., Ina, T., Ueno, Y. & Nishii, J. $1.8\text{m}\Omega\cdot\text{cm}^2$ vertical GaN-based trench metal-oxide-semiconductor field-effect transistors on a free-standing GaN substrate for 1.2-kV-class operation. *Applied Physics Express* **8**, 2–4 (2015).
 80. Heikman, S., Keller, S., Green, D. S., DenBaars, S. P. & Mishra, U. K. High conductivity modulation doped AlGaIn/GaN multiple channel heterostructures. *J Appl Phys* **94**, 5321–5325 (2003).
 81. Mutig, A. & Bimberg, D. Progress on high-speed 980nm VCSELs for short-reach optical interconnects. *Advances in Optical Technologies* **2011**, (2011).
 82. Al-Omari, A. N., Alias, M. S., Ababneh, A. & Lear, K. L. Improved performance of top-emitting oxide-confined polyimide-planarized 980 nm VCSELs with copper-plated heat sinks. *J Phys D Appl Phys* **45**, (2012).
 83. Liu, C., Cai, Y., Jiang, H. & Lau, K. M. Monolithic integration of III-nitride voltage-controlled light emitters with dual-wavelength photodiodes by selective-area epitaxy. *Opt Lett* **43**, 3401 (2018).
 84. Soh, M. *et al.* Heterogeneous Integration of GaN and BCD Technologies. *Electronics (Basel)* **8**, 351 (2019).
 85. Lapis, M., Stemme, G. & Niklaus, F. Wafer-level heterogeneous integration for MOEMS, MEMS, and NEMS. *IEEE Journal on Selected Topics in Quantum Electronics* **17**, 629–644 (2011).
 86. Xie, L. *et al.* Heterogeneous integration of bio-sensing system-on-chip and printed electronics. *IEEE J Emerg Sel Top Circuits Syst* **2**, 672–682 (2012).
 87. Hwang, S. W. *et al.* 25th anniversary article: Materials for high-performance biodegradable semiconductor devices. *Advanced Materials* **26**, 1992–2000 (2014).
 88. Kang, D. *et al.* Flexible Opto-Fluidic Fluorescence Sensors Based on Heterogeneously Integrated Micro-VCSELs and Silicon Photodiodes. *ACS Photonics* **3**, 912–918 (2016).
 89. Lu, Y., Huang, J. Y., Wang, C., Sun, S. & Lou, J. Cold welding of ultrathin gold nanowires. *Nat Nanotechnol* **5**, 218–224 (2010).
 90. Lee, K., Zimmerman, J. D., Hughes, T. W. & Forrest, S. R. Non-destructive wafer recycling for low-cost thin-film flexible optoelectronics. *Adv Funct Mater* **24**, 4284–4291 (2014).
 91. Zaffar, M., Ehsan, S., Stolkin, R. & Maier, K. M. D. Sensors, SLAM and Long-term Autonomy: A Review. *2018 NASA/ESA Conference on Adaptive Hardware and Systems, AHS 2018* 285–290 (2018) doi:10.1109/AHS.2018.8541483.
 92. Fornetti, F., Morris, K. A. & Beach, M. A. Pulsed operation and performance of commercial GaN HEMTs. *Proceedings of the 2009 European Microwave Integrated Circuits Conference (EuMIC)* 226–9|56 (2009).
 93. Maturana, D. & Scherer, S. VoxNet: A 3D Convolutional Neural Network for real-time object recognition. *IEEE International Conference on Intelligent Robots and Systems 2015-Decem*, 922–928 (2015).
 94. Jing, H. & You, S. Point cloud labeling using 3D Convolutional Neural Network. *Proceedings - International Conference on Pattern Recognition* 2670–2675 (2017)

- doi:10.1109/ICPR.2016.7900038.
95. Qi, C. R., Liu, W., Wu, C., Su, H. & Guibas, L. J. Supplementary Material for Frustum PointNets for 3D Object Detection from RGB-D Data. *Cvpr* 1–6 (2017) doi:10.1109/CVPR.2018.00102.
 96. Su, S. *et al.* Material Classification Using Raw Time-of-Flight Measurements. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem*, 3503–3511 (2016).
 97. Moench, H. *et al.* High-power VCSEL systems and applications. *High-Power Diode Laser Technology and Applications XIII* **9348**, 93480W (2015).
 98. Seurin, J.-F. *et al.* High-brightness pump sources using 2D VCSEL arrays. *Vertical-Cavity Surface-Emitting Lasers XIV* **7615**, 76150F (2010).
 99. Seurin, J.-F. *et al.* High-power red VCSEL arrays. *Vertical-Cavity Surface-Emitting Lasers XVII* **8639**, 86390O (2013).
 100. Yoo, H. W. *et al.* MEMS-based lidar for autonomous driving. *Elektrotechnik und Informationstechnik* **135**, 408–415 (2018).
 101. Wang, Y. & Wu, M. C. Micromirror based optical phased array for wide-angle beamsteering. *Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS)* 897–900 (2017) doi:10.1109/MEMSYS.2017.7863553.
 102. Kwong, D. *et al.* On-chip silicon optical phased array for two-dimensional beam steering. *Opt Lett* **39**, 941 (2014).
 103. Poulton, C. V. *et al.* Large-scale silicon nitride nanophotonic phased arrays at infrared and visible wavelengths. *Opt Lett* **42**, 21 (2017).
 104. Komljenovic, T., Helkey, R., Coldren, L. & Bowers, J. E. Sparse aperiodic arrays for optical beam forming and LIDAR. *Opt Express* **25**, 2511 (2017).
 105. Davis, S. R., Farca, G., Rommel, S. D., Martin, A. W. & Anderson, M. H. Analog, non-mechanical beam-steerer with 80 degree field of regard. *Acquisition, Tracking, Pointing, and Laser Systems Technologies XXII* **6971**, 69710G (2008).
 106. Ghezzi, D. *et al.* sensitivity in blind rat retinas. **7**, (2013).
 107. Panwar, N., Rajendran, B. & Ganguly, U. Arbitrary Spike Time Dependent Plasticity (STDP) in memristor by analog waveform engineering. *IEEE Electron Device Letters* **38**, 740–743 (2017).
 108. Makino, K. *et al.* Development of an InGaAs SPAD 2D array for flash LIDAR. 19 (2018) doi:10.1117/12.2289270.
 109. Jiang, X. *et al.* InGaAsP/InP Geiger-mode APD-based LiDAR. in *Optical Sensing, Imaging, and Photon Counting: From X-Rays to THz* (eds. Mitrofanov, O., Tan, C. H., Vizcaíno, J. L. P. & Razeghi, M.) vol. 10729 33–44 (SPIE, 2018).
 110. Jones, A. H., March, S. D., Bank, S. R. & Campbell, J. C. Low-noise high-temperature AlInAsSb/GaSb avalanche photodiodes for 2- μ m applications. *Nat Photonics* (2020) doi:10.1038/s41566-020-0637-6.
 111. Pi, S. *et al.* Memristor crossbar arrays with 6-nm half-pitch and 2-nm critical dimension. *Nat Nanotechnol* **14**, 35–39 (2019).
 112. Xue, C. X. *et al.* A CMOS-integrated compute-in-memory macro based on resistive random-access memory for AI edge devices. *Nat Electron* **4**, 81–90 (2021).
 113. Lee, H. Y. *et al.* Evidence and solution of Over-RESET Problem for HfO X Based Resistive Memory with Sub-ns Switching Speed and High Endurance. 19–22 (2010).
 114. Choi, B. J. *et al.* High-Speed and Low-Energy Nitride Memristors. 5290–5296 (2016) doi:10.1002/adfm.201600680.
 115. Ma, C. *et al.* Sub-nanosecond memristor based on ferroelectric tunnel junction. *Nat Commun* **11**, 1–9 (2020).
 116. Cho, Y., Hashimoto, S., Odagawa, N., Tanaka, K. & Hiranaga, Y. Realization of 10 Tbit in.2 memory density and subnanosecond domain switching time in ferroelectric data storage. *Appl Phys Lett* **87**, 1–3 (2005).
 117. Park, M. *et al.* Non-Line-of-Sight Detection Based on Neuromorphic Time-of-Flight Sensing.

- ACS Photonics* (2023).
118. Yamashita, H. *et al.* Dual-axis hybrid-MEMS mirror having Si and polymer hinges for large-area 2D raster scan. in *2014 International Conference on Optical MEMS and Nanophotonics* 167–168 (IEEE, 2014).
 119. Tan, C. *et al.* Long-Range Daytime 3D Imaging Lidar With Short Acquisition Time Based on 64×64 Gm-APD Array. *IEEE Photonics J* **14**, 1–7 (2022).
 120. Chevrier, M. & Campanella, G. How to build a LIDAR system with a time-to-digital converter. *Analog Applications Journal* 1–6 (2017).
 121. Zhang, Y. & An, Y. Computational three-dimensional imaging method of compressive LIDAR system with gain modulation. in *LIDAR Imaging Detection and Target Recognition 2017* (eds. Lv, Y. *et al.*) vol. 10605 44–48 (SPIE, 2017).
 122. Yang, J. Y. *et al.* Reconfigurable Radio-Frequency High-Electron Mobility Transistors via Ferroelectric-Based Gallium Nitride Heterostructure. *Adv Electron Mater* 2101406 (2022).
 123. Ryu, H., Wu, H., Rao, F. & Zhu, W. Ferroelectric Tunneling Junctions Based on Aluminum Oxide/ Zirconium-Doped Hafnium Oxide for Neuromorphic Computing. *Sci Rep* **9**, 1–8 (2019).
 124. Tokura, Y., Seki, S. & Nagaosa, N. Multiferroics of spin origin. *Reports on Progress in Physics* **77**, 76501 (2014).
 125. Wright, C. D., Liu, Y., Kohary, K. I., Aziz, M. M. & Hicken, R. J. Arithmetic and biologically-inspired computing using phase-change materials. *Advanced Materials* **23**, 3408–3413 (2011).
 126. Burr, G. W. *et al.* Experimental demonstration and tolerancing of a large-scale neural network (165,000 synapses), using phase-change memory as the synaptic weight element. *2014 IEEE International Electron Devices Meeting* **XX**, 29.5.1–29.5.4 (2014).
 127. Kim, S. *et al.* NVM neuromorphic core with 64k-cell (256-by-256) phase change memory synaptic array with on-chip neuron circuits for continuous in-situ learning. *Technical Digest - International Electron Devices Meeting, IEDM 2016-Febru*, 17.1.1–17.1.4 (2015).
 128. Hinz, G. *et al.* Online multi-object tracking-by-clustering for intelligent transportation system with neuromorphic vision sensor. in *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)* 142–154 (Springer, 2017).
 129. Lamoureux, A., Lee, K., Shlian, M., Forrest, S. R. & Shtein, M. Dynamic kirigami structures for integrated solar tracking. *Nat Commun* **6**, 1–6 (2015).
 130. Lee, K., Lee, J., Mazor, B. A. & Forrest, S. R. Transforming the cost of solar-to-electrical energy conversion: Integrating thin-film GaAs solar cells with non-tracking mini-concentrators. *Light Sci Appl* **4**, e288–e288 (2015).
 131. Bae, B., Park, M., Lee, D., Sim, I. & Lee, K. Hetero-Integrated InGaAs Photodiode and Oxide Memristor-Based Artificial Optical Nerve for In-Sensor NIR Image Processing. *Adv Opt Mater* **11**, 2201905 (2023).
 132. Wang, T. *et al.* Reconfigurable optoelectronic memristor for in-sensor computing applications. *Nano Energy* **89**, (2021).
 133. Lee, D. *et al.* In-sensor image memorization and encoding via optical neurons for bio-stimulus domain reduction toward visual cognitive processing. *Nat Commun* **13**, 5223 (2022).
 134. Zhou, F. & Chai, Y. Near-sensor and in-sensor computing. *Nat Electron* **3**, 664–671 (2020).
 135. Ciesla, C. M. *et al.* Biomedical applications of terahertz pulse imaging. *Commercial and Biomedical Applications of Ultrafast Lasers II* **3934**, 73 (2000).
 136. El-Hajj, C. & Kyriacou, P. A. A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. *Biomed Signal Process Control* **58**, 101870 (2020).
 137. Tanveer, Md. S. & Hasan, Md. K. Cuffless blood pressure estimation from electrocardiogram and photoplethysmogram using waveform based ANN-LSTM network. *Biomed Signal Process Control* **51**, 382–392 (2019).
 138. Ma, J., Chen, C., Li, C. & Huang, J. Infrared and visible image fusion via gradient transfer and total variation minimization. *Information Fusion* **31**, 100–109 (2016).

139. Ma, J., Yu, W., Liang, P., Li, C. & Jiang, J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Information Fusion* **48**, 11–26 (2019).
140. Kamimura, M., Kanayama, N., Tokuzen, K., Soga, K. & Nagasaki, Y. Near-infrared (1550 nm) in vivo bioimaging based on rare-earth doped ceramic nanophosphors modified with PEG-b-poly(4-vinylbenzylphosphonate). *Nanoscale* **3**, 3705–3713 (2011).
141. McDaniel, R. V *et al.* Image fusion for tactical applications. in *Infrared Technology and Applications XXIV* (eds. Andresen, B. F. & Strojnik, M.) vol. 3436 685–695 (SPIE, 1998).
142. Hwang, A. *et al.* Visible and infrared dual-band imaging via Ge/MoS₂ van der Waals heterostructure. *Sci Adv* **7**, eabj2521 (2021).
143. Zou, Y. *et al.* Infrared visible color night vision image fusion based on deep learning. in *AI and Optical Data Sciences II* vol. 11703 62 (SPIE, 2021).
144. Naboulsi, M. C. Al, Sizun, H. & de Fornel, F. Fog attenuation prediction for optical and infrared waves. *Optical Engineering* **43**, 319–329 (2004).
145. Xia, Z. & Wei, Z. Z. Fabrication and performance of a NiMn₂O₄/LaNiO₃ composite film detector with multi electrodes for night vision imaging application. *Sens Actuators A Phys* 112713 (2021) doi:<https://doi.org/10.1016/j.sna.2021.112713>.
146. Cao, Y. *et al.* Ultra-Broadband Photodetector for the Visible to Terahertz Range by Self-Assembling Reduced Graphene Oxide-Silicon Nanowire Array Heterojunctions. *Small* **10**, 2345–2351 (2014).
147. Lee, J. W., Kim, D. Y., Baek, S., Yu, H. & So, F. Inorganic UV–Visible–SWIR Broadband Photodetector Based on Monodisperse PbS Nanocrystals. *Small* **12**, 1328–1333 (2016).
148. Hu, W. *et al.* Germanium/perovskite heterostructure for high-performance and broadband photodetector from visible to infrared telecommunication band. *Light Sci Appl* **8**, 106 (2019).
149. Tang, X., Ackerman, M. M., Chen, M. & Guyot-Sionnest, P. Dual-band infrared imaging using stacked colloidal quantum dot photodiodes. *Nat Photonics* **13**, 277–282 (2019).
150. Lan, Z. *et al.* Near-infrared and visible light dual-mode organic photodetectors. *Sci Adv* **6**, (2020).
151. Wang, Y. *et al.* Stacked Dual-Wavelength Near-Infrared Organic Photodetectors. *Adv Opt Mater* **9**, 2001784 (2021).
152. Simola, E. T. *et al.* Voltage-tunable dual-band Ge/Si photodetector operating in VIS and NIR spectral range. *Opt. Express* **27**, 8529–8539 (2019).
153. Song, D. Y., Chu, D., Lee, S. K., Pak, S. W. & Kim, E. K. High photoresponsivity from multilayer MoS₂/Si heterojunction diodes formed by vertically stacking. *J Appl Phys* **122**, 124505 (2017).
154. Chiu, M. H. *et al.* Determination of band alignment in the single-layer MoS₂ WSe₂ heterojunction. *Nat Commun* **6**, 1–6 (2015).
155. Hu, L. *et al.* All-Optically Controlled Memristor for Optoelectronic Neuromorphic Computing. *Adv Funct Mater* **31**, 2005582 (2021).
156. Lachat, E., Macher, H., Landes, T. & Grussenmeyer, P. Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling. *Remote Sens (Basel)* **7**, 13070–13097 (2015).
157. Cui, B., Ma, X., Xie, X., Ren, G. & Ma, Y. Classification of visible and infrared hyperspectral images based on image segmentation and edge-preserving filtering. *Infrared Phys Technol* **81**, 79–88 (2017).
158. Manzeli, S., Ovchinnikov, D., Pasquier, D., Yazyev, O. V & Kis, A. 2D transition metal dichalcogenides. *Nat Rev Mater* **2**, 17033 (2017).
159. Goossens, S. *et al.* Broadband image sensor array based on graphene–CMOS integration. *Nat Photonics* **11**, 366–371 (2017).
160. Mueller, T. & Malic, E. Exciton physics and device application of two-dimensional transition metal dichalcogenide semiconductors. *NPJ 2D Mater Appl* **2**, 29 (2018).
161. Prezioso, M., Hoskins, B. D., Adam, G. C., Likharev, K. K. & Strukov, D. B. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **521**,

- 61–64 (2015).
162. Wang, Z. *et al.* Fully memristive neural networks for pattern classification with unsupervised learning. *Nat Electron* **1**, 137–145 (2018).
 163. Yao, P. *et al.* Fully hardware-implemented memristor convolutional neural network. *Nature* **577**, 641–647 (2020).
 164. Zhou, F. & Chai, Y. Near-sensor and in-sensor computing. *Nat Electron* **3**, 664–671 (2020).
 165. Chai, Y. In-sensor computing for machine vision. *Nature* **579**, 32–33 (2020).
 166. Wang, H. *et al.* A Ferroelectric/Electrochemical Modulated Organic Synapse for Ultraflexible, Artificial Visual-Perception System. *Advanced Materials* **30**, (2018).
 167. Chen, S., Lou, Z., Chen, D. & Shen, G. An Artificial Flexible Visual Memory System Based on an UV-Motivated Memristor. *Advanced Materials* **30**, (2018).
 168. Jang, H. *et al.* An Atomically Thin Optoelectronic Machine Vision Processor. *Advanced Materials* **32**, (2020).
 169. Gao, S. *et al.* An Oxide Schottky Junction Artificial Optoelectronic Synapse. *ACS Nano* **13**, 2634–2642 (2019).
 170. Choi, C. *et al.* Curved neuromorphic image sensor array using a MoS₂-organic heterostructure inspired by the human visual recognition system. *Nat Commun* **11**, 1–9 (2020).
 171. Xue, J. *et al.* Narrowband Perovskite Photodetector-Based Image Array for Potential Application in Artificial Vision. *Nano Lett* **18**, (2018).
 172. Zhou, F. *et al.* Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat Nanotechnol* **14**, 776–782 (2019).
 173. Wang, T. *et al.* Reconfigurable optoelectronic memristor for in-sensor computing applications. *Nano Energy* **89**, (2021).
 174. Hong, S. *et al.* Neuromorphic Active Pixel Image Sensor Array for Visual Memory. *ACS Nano* **15**, 15362–15370 (2021).
 175. Molina-mendoza, A. J. & Mueller, T. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–65 (2020).
 176. Zhu, Q. *et al.* A flexible ultrasensitive optoelectronic sensor array for neuromorphic vision systems. *Nat Commun* **12**, 1–7 (2021).
 177. Turner, M. H., Gonzalo, L., Giraldo, S., Schwartz, O. & Rieke, F. Stimulus- and goal-oriented frameworks for understanding natural vision. *Nat Neurosci* **22**, 15–24 (2019).
 178. Liu, B., Hong, A., Rieke, F. & Manookin, M. B. Predictive encoding of motion begins in the primate retina. *Nat Neurosci* **24**, 1280–1291 (2021).
 179. Masland, R. H. The Neuronal Organization of the Retina. *Neuron* **76**, 266–280 (2012).
 180. Fan, D., Lee, K. & Forrest, S. R. Flexible Thin-Film InGaAs Photodiode Focal Plane Array. *ACS Photonics* **3**, 670–676 (2016).
 181. Sebastian, A., Gallo, M. Le, Khaddam-aljameh, R. & Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat Nanotechnol* **15**, 529–544 (2020).
 182. Li, C. *et al.* Analogue signal and image processing with large memristor crossbars. *Nat Electron* **1**, 52–59 (2018).
 183. Deng, L. The MNIST Database of Handwritten Digit Images for Machine Learning Research. *IEEE Signal Process Mag* **29**, 141–142 (2012).
 184. Bae, B., Park, M., Lee, D., Sim, I. & Lee, K. Hetero-Integrated InGaAs Photodiode and Oxide Memristor-Based Artificial Optical Nerve for In-Sensor NIR Image Processing. *Adv Opt Mater* **11**, 2201905 (2023).
 185. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324 (1998).
 186. Chen, P., Peng, X. & Yu, S. NeuroSim: A Circuit-Level Macro Model for Benchmarking Neuro-Inspired Architectures in Online Learning. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **37**, 3067–3080 (2018).
 187. Thyayne, I. Advanced III-V HEMTs. *III-Vs Review* **16**, 48–51 (2003).

188. Hamza, K. H. & Nirmal, D. A review of GaN HEMT broadband power amplifiers. *AEU-International Journal of Electronics and Communications* **116**, 153040 (2020).
189. Keshmiri, N., Wang, D., Agrawal, B., Hou, R. & Emadi, A. Current status and future trends of GaN HEMTs in electrified transportation. *IEEE Access* **8**, 70553–70571 (2020).
190. Zhu, J. *et al.* Ferroelectric gate AlGaIn/GaN E-mode HEMTs with high transport and sub-threshold performance. *IEEE Electron Device Letters* **39**, 79–82 (2017).
191. Wu, C.-H. *et al.* High-performance normally-OFF GaN MIS-HEMTs using hybrid ferroelectric charge trap gate stack (FEG-HEMT) for power device applications. *IEEE Electron Device Letters* **39**, 991–994 (2018).
192. Chandrasekar, H. *et al.* Demonstration of Wide Bandgap AlGaIn/GaN Negative-Capacitance High-Electron-Mobility Transistors (NC-HEMTs) Using Barium Titanate Ferroelectric Gates. *Adv Electron Mater* **6**, 2000074 (2020).
193. Li, G. *et al.* Heteroepitaxy of Hf_{0.5}Zr_{0.5}O₂ ferroelectric gate layer on AlGaIn/GaN towards normally-off HEMTs. *Appl Surf Sci* **597**, 153709 (2022).
194. Wu, C. *et al.* Hf_{0.5}Zr_{0.5}O₂-based ferroelectric gate HEMTs with large threshold voltage tuning range. *IEEE Electron Device Letters* **41**, 337–340 (2020).
195. Chen, L. *et al.* Hetero-integration of quasi two-dimensional PbZr_{0.2}Ti_{0.8}O₃ on AlGaIn/GaN HEMT and non-volatile modulation of two-dimensional electron gas. *Appl Phys Lett* **115**, (2019).
196. Song, W. *et al.* Steep subthreshold swing in GaN negative capacitance field-effect transistors. *IEEE Trans Electron Devices* **66**, 4148–4150 (2019).
197. Ye, H. *et al.* Ferroelectric-Gated GaN HEMTs for RF and mm-Wave Switch Applications. in *2022 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)* 1–2 (IEEE, 2022).
198. Chen, P.-G., Wei, Y.-T., Tang, M. & Lee, M.-H. Experimental demonstration of ferroelectric gate-stack AlGaIn/GaN-on-Si MOS-HEMTs with voltage amplification for power applications. *IEEE Trans Electron Devices* **61**, 3014–3017 (2014).
199. Zhao, D. *et al.* Depolarization of multidomain ferroelectric materials. *Nat Commun* **10**, 2547 (2019).
200. Xie, S. *et al.* Ferroelectric semiconductor junctions based on graphene/In₂Se₃/graphene van der Waals heterostructures. *2d Mater* **8**, 045020 (2021).
201. Wang, L. *et al.* Exploring ferroelectric switching in α -In₂Se₃ for neuromorphic computing. *Adv Funct Mater* **30**, 2004609 (2020).
202. Huang, Y. *et al.* Two-dimensional In₂Se₃: A rising advanced material for ferroelectric data storage. *InfoMat* **4**, e12341 (2022).
203. Li, Y. *et al.* Orthogonal electric control of the out-of-plane field-effect in 2D ferroelectric α -In₂Se₃. *Adv Electron Mater* **6**, 2000061 (2020).
204. Loyez, C., Carpentier, K., Sourikopoulos, I. & Danneville, F. Subthreshold neuromorphic devices for spiking neural networks applied to embedded ai. in *2021 19th IEEE International New Circuits and Systems Conference (NEWCAS)* 1–4 (IEEE, 2021).
205. Glaser, J. How GaN power transistors drive high-performance Lidar: Generating ultrafast pulsed power with GaN FETs. *IEEE Power Electronics Magazine* **4**, 25–35 (2017).
206. Ma, C.-T. & Gu, Z.-H. Review of GaN HEMT applications in power converters over 500 W. *Electronics (Basel)* **8**, 1401 (2019).
207. Si, M., Liao, P.-Y., Qiu, G., Duan, Y. & Ye, P. D. Ferroelectric field-effect transistors based on MoS₂ and CuInP₂S₆ two-dimensional van der Waals heterostructure. *ACS Nano* **12**, 6700–6705 (2018).
208. Choi, K.-B. *et al.* A split-gate positive feedback device with an integrate-and-fire capability for a high-density low-power neuron circuit. *Front Neurosci* **12**, 704 (2018).
209. Zhong, Y. *et al.* Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing. *Nat Commun* **12**, 408 (2021).
210. Rolls, E. T. & Baylis, L. L. Gustatory, olfactory, and visual convergence within the primate

- orbitofrontal cortex. *Journal of Neuroscience* **14**, 5437–5452 (1994).
211. Abaira, V. E. & Ginty, D. D. The sensory neurons of touch. *Neuron* **79**, 618–639 (2013).
 212. Spray, D. C. Cutaneous temperature receptors. *Annu Rev Physiol* **48**, 625–638 (1986).
 213. Lynn, B. Somatosensory receptors and their CNS connections. *Annu Rev Physiol* **37**, 105–127 (1975).
 214. Gerschenfeld, H. M. Chemical transmission in invertebrate central nervous systems and neuromuscular junctions. *Physiol Rev* **53**, 1–119 (1973).
 215. Kim, S. *et al.* Artificial stimulus-response system capable of conscious response. *Sci Adv* **7**, eabe3996 (2021).
 216. He, K. *et al.* Artificial Neural Pathway Based on a Memristor Synapse for Optically Mediated Motion Learning. *ACS Nano* (2022).
 217. Wang, C. *et al.* A Braitenberg vehicle based on memristive neuromorphic circuits. *Advanced Intelligent Systems* **2**, 1900103 (2020).
 218. Roe, D. G. *et al.* Humanlike spontaneous motion coordination of robotic fingers through spatial multi-input spike signal multiplexing. *Nat Commun* **14**, 1–7 (2023).
 219. Shen, X. *et al.* Improved Self-Powered Photodetection of Ferroelectric PbZr_{0.52}Ti_{0.48}O₃ Thin Films via Interfacial Engineering. *The Journal of Physical Chemistry C* **126**, 18617–18622 (2022).
 220. Kamalasanan, M. N., Chandra, S., Joshi, P. C. & Mansingh, A. Structural and optical properties of sol-gel-processed BaTiO₃ ferroelectric thin films. *Appl Phys Lett* **59**, 3547–3549 (1991).
 221. Rana, D. S. *et al.* Understanding the nature of ultrafast polarization dynamics of ferroelectric memory in the multiferroic BiFeO₃. *Advanced Materials* **21**, 2881–2885 (2009).
 222. Chen, J. *et al.* Mimicking Neuroplasticity via Ion Migration in van der Waals Layered Copper Indium Thiophosphate. *Advanced Materials* **34**, (2022).
 223. Wu, J. *et al.* High tunnelling electroresistance in a ferroelectric van der Waals heterojunction via giant barrier height modulation. *Nat Electron* **3**, 466–472 (2020).
 224. Jiang, X. *et al.* Manipulation of current rectification in van der Waals ferroionic CuInP₂S₆. *Nat Commun* **13**, 1–8 (2022).
 225. Wang, X. *et al.* Van der Waals engineering of ferroelectric heterostructures for long-retention memory. *Nat Commun* **12**, 1–8 (2021).
 226. Liu, F. *et al.* Room-temperature ferroelectricity in CuInP₂S₆ ultrathin flakes. *Nat Commun* **7**, 1–6 (2016).
 227. Belianinov, A. *et al.* CuInP₂S₆ room temperature layered ferroelectric. *Nano Lett* **15**, 3808–3814 (2015).
 228. Young, L. R. Pursuit eye tracking movements. *The control of eye movements* 429–443 (1971).
 229. Dudel, J. The effect of reduced calcium on quantal unit current and release at the crayfish neuromuscular junction. *Pflügers Archiv* **391**, 35–40 (1981).
 230. Zhang, N. & He, X. Understanding the extraocular muscles and oculomotor, trochlear, and abducens nerves through a simulation in physical examination training: an innovative approach. *Journal of Chiropractic Education* **24**, 153–158 (2010).
 231. Mackey, M. C. & Glass, L. Oscillation and chaos in physiological control systems. *Science* (1979) **197**, 287–289 (1977).
 232. Farmer, J. D. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D* **4**, 366–393 (1982).
 233. Glass, L. & Mackey, M. Mackey-glass equation. *Scholarpedia* **5**, 6908 (2010).
 234. Amil, P., Cabeza, C. & Marti, A. C. Exact discrete-time implementation of the Mackey–Glass delayed model. *IEEE Transactions on Circuits and Systems II: Express Briefs* **62**, 681–685 (2015).
 235. Makridakis, S. Accuracy measures: theoretical and practical concerns. *Int J Forecast* **9**, 527–529 (1993).
 236. Lin, S. *et al.* A GaN HEMT structure allowing self-terminated, plasma-free etching for high-

- uniformity, high-mobility enhancement-mode devices. *IEEE Electron Device Letters* **37**, 377–380 (2016).
237. Annett, M., Grossman, T., Wigdor, D. & Fitzmaurice, G. Medusa: a proximity-aware multi-touch tabletop. in *Proceedings of the 24th annual ACM symposium on User interface software and technology* 337–346 (2011).
238. Jang, S. *et al.* Ultrathin conformable organic artificial synapse for wearable intelligent device applications. *ACS Appl Mater Interfaces* **11**, 1071–1080 (2018).
239. Seo, M. *et al.* First demonstration of a logic-process compatible junctionless ferroelectric FinFET synapse for neuromorphic applications. *IEEE Electron Device Letters* **39**, 1445–1448 (2018).
240. Yang, J. Y., Yeom, M. J., Park, Y., Heo, J. & Yoo, G. Ferroelectric α -In₂Se₃ Wrapped-Gate β -Ga₂O₃ Field-Effect Transistors for Dynamic Threshold Voltage Control. *Adv Electron Mater* **7**, 2100306 (2021).
241. Ji, X. *et al.* Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor. *Nat Commun* **12**, 2480 (2021).
242. Wang, J. *et al.* Deep-ultraviolet-triggered neuromorphic functions in In-Zn-O phototransistors. *Appl Phys Lett* **113**, 151101 (2018).
243. Wang, Y. *et al.* Photonic synapses based on inorganic perovskite quantum dots for neuromorphic computing. *Advanced materials* **30**, 1802883 (2018).
244. Tesler, F. *et al.* Relaxation of a spiking Mott artificial neuron. *Phys Rev Appl* **10**, 054001 (2018).
245. Khyam, M. O., Ge, S. S., Li, X. & Pickering, M. R. Highly accurate time-of-flight measurement technique based on phase-correlation for ultrasonic ranging. *IEEE Sens J* **17**, 434–443 (2016).
246. Jackson, J. C. *et al.* Time-of-flight measurement techniques for airborne ultrasonic ranging. *IEEE Trans Ultrason Ferroelectr Freq Control* **60**, 343–355 (2013).
247. Paredes, J. A., Álvarez, F. J., Aguilera, T. & Villadangos, J. M. 3D indoor positioning of UAVs with spread spectrum ultrasound and time-of-flight cameras. *Sensors* **18**, 89 (2017).
248. Lenk, C. *et al.* Neuromorphic acoustic sensing using an adaptive microelectromechanical cochlea with integrated feedback. *Nat Electron* 1–11 (2023).
249. Moro, F. *et al.* Neuromorphic object localization using resistive memories and ultrasonic transducers. *Nat Commun* **13**, 3506 (2022).