SCIENCE CHINA
Information Sciences

• REVIEW •

April 2022, Vol. 65 141401:1–141401:14
https://doi.org/10.1007/s11432-021-3336-8

Neuromorphic sensory computing

Tianqing WAN[1], Sijie MA[1], Fuyou LIAO[1,2], Lingwei FAN[1] & Yang CHAI[1,2*]

[1]*Department of Applied Physics, The Hong Kong Polytechnic University, Hong Kong, China;*
[2]*The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 515100, China*

Received 31 July 2021/Revised 14 September 2021/Accepted 22 September 2021/Published online 16 December 2021

Abstract The number of sensory nodes in the Internet of everything continues to increase rapidly and generate massive data. The generated information from sensory nodes is much larger than the total collective human sensory throughput. It is quite challenging to send all of the data produced at sensory terminals to the "Cloud" computation center, especially for those time-delay sensitive applications. This situation demands a dramatic increase in the computation near or inside sensory networks. Inspired by biological sensory systems with a high data compression ratio, neuromorphic sensory computing provides a way to efficiently acquire and process a large volume of data from complex environments. Researchers have been investigating emerging materials, devices, circuits, and computing architectures to implement an artificial sensory system with high energy efficiency, speed, and density. Here we summarize the important features of biological systems and their hardware implementations. Electrons and photons are two representative information carriers, in which electron carrier allows high integration density for complex computing and photon carrier has high connectivity, high speed, wide bandwidth, and low power consumption. We overview the electronic and optical neuromorphic sensory computing and hybrid opto-electronic sensory computing, and present advances on multimodal sensory computing and their potential challenges.

Keywords neuromorphic, sensory computing, multimodal sensory computing, electronic sensory computing, optical sensory computing

Citation Wan T Q, Ma S J, Liao F Y, et al. Neuromorphic sensory computing. Sci China Inf Sci, 2022, 65(4): 141401, https://doi.org/10.1007/s11432-021-3336-8

1 Introduction

Sensory systems are indispensable parts for acquiring information from complex environments. The collected raw data are perceived and computed for high-level intelligence, e.g., memorizing, learning, making decisions, which constructs the base of our interactions with the environments. The sensory system has a general architecture composed of receptors (data acquisition), pathway (data transmission), and processing units (computing and memory) [1, 2], as illustrated in Figure 1. The performance of receptors is under steady progress in terms of feature size, speed, power consumption, sensitivity, detecting range, and cost, benefiting from the development of new materials and processing technology. Nowadays, various high-performance receptors (e.g., vision, audio, pressure, chemicals, etc.) keep producing a large volume of data, which are further encoded, transferred, and processed. As the number of sensory nodes increases, the huge amount of raw data (containing lots of redundant data) become a heavy burden to the transmission network and processing units. In addition, the computing units and memory in conventional von Neumann architecture are separated in physical space. The physical separation between receptors, computing units, and memories gives rise to frequent data movement along the interconnects, degrading the overall performance in terms of speed and energy efficiency [3, 4].

In sharp contrast, biological counterparts can efficiently process the sensory information from complex environments with high energy efficiency, owing to the well-developed hierarchy architecture, co-localization of computing and memory, and complex neural network. For example, the human retinas consist of various functional cells arranged layer-by-layer, in which the raw data from the photo-receptors

* Corresponding author (email: ychai@polyu.edu.hk)

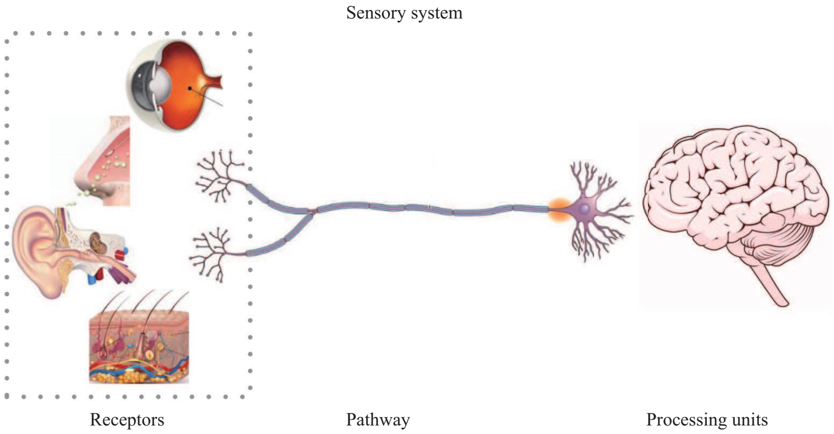

Figure 1 (Color online) The schematic of the sensory system. The system acquires information from environments through various types of receptors (e.g., visual, olfactory, auditory, tactile sensation, etc.). Then the outputs of receptors are transmitted to the complex processing units through the pathway.

are pre-processed before they are transferred to the complex central nervous system [5]. By conducting effective feature extraction close to the receptors, the biological sensory system effectively reduces the data that requires a long-distance transfer, greatly improving the energy efficiency and speed. This strategy has been implemented in artificial sensory systems by analog very large scale integration (VLSI) [6–9]. As the feature size of Si-based transistors shrinks, more complex CMOS-based elements can be integrated very close to the receptors, making the receptors output a more efficient representation of the external stimuli. However, there are inherent distinctions between CMOS-based processing units (logical gates, SRAM, DRAM) and biological counterparts (neurons and synapses). For example, a single neuron or synapse has complex dynamics for computing and memory [10–12]. To emulate a corresponding function, it requires hundreds of transistors, which restricts the integration density, energy efficiency, and speed.

A biological neural network relies on ion dynamics to transmit information and works at a relatively slow speed (hundreds of Hz). Comparatively, modern computers based on Si-based transistors use electrons to carry information, exhibiting much higher speed. They are more suited to running sequential, digital, procedure-based programs [13]. However, neuromorphic computing requires distributed and massively parallel architecture, which is hard to realize in conventional electronic computers with low connectivity. Although time-division multiplexed communication bus sacrifices the bandwidth for connectivity, this trade-off in speed and interconnectivity restricts the overall performance. Apart from electrons, light is another information carrier for high-speed communication. The optical interconnects have intrinsic high connectivity, which makes photonic neuromorphic computing potential candidates for realizing low latency, high bandwidth, and low energy. However, optical modules are quite bulky and have limited computing complexity, which restricts the integration density. To achieve high efficiency, it is appealing to design a hybrid optical-electronic computing architecture, exploiting their respective advantages.

In this review, we will first introduce the important functions that contribute to the high efficiency of the biological sensory system and corresponding electronic hardware. In the third part, we summarize the optical implementations of neuromorphic computing. Then, we review the essential optic-electronic devices that conduct signal transduction. Next, we discuss multimodal sensory computing, in terms of the fundamental mechanism and hardware implementations. At last, we propose some figure-of-merits that are important for evaluating the performance of a sensory system.

2 Electronic sensory computing

Electronic devices with diverse and complex functions at the back-end of sensory systems can process the output signals from various receptors. Inspired by the high efficiency of biological sensory systems, researchers have adopted various electronic devices and circuits to emulate the basic modules and architecture of biological counterparts. In this section, we summarize a few key features of neuromorphic sensory computing with electronic devices.

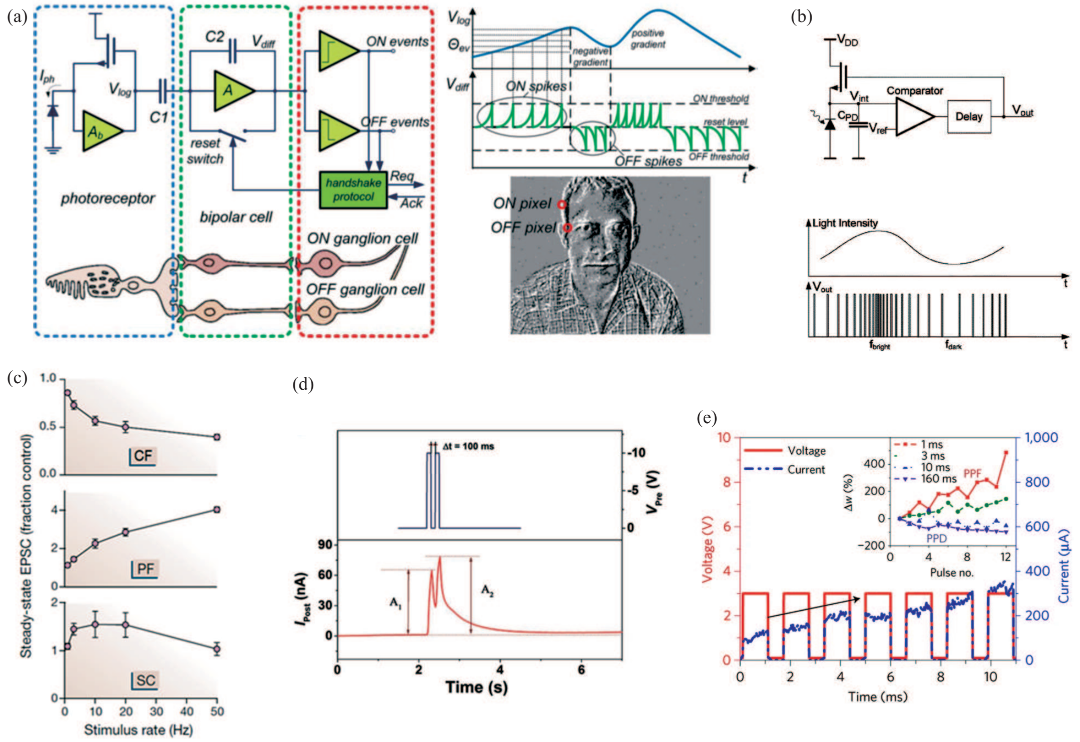

Figure 2 (Color online) (a) Three-layer model of a human retina, corresponding dynamic vision sensor (DVS) pixel circuitry and waveforms [15] @Copyright 2014 IEEE. (b) Schematic of the frequency-based visual sensor [14]@Copyright 2014 IEEE. (c) Short-term memory can be used to construct filters. The initial states influence the types of filters (low-pass, high-pass, and band-pass). CF, PF, and SC are climbing fibre, parallel fibre, Schaffer collateral synapses, respectively [10] @Copyright 2000 JNeurosci. (d) Pair-pulse facilitation [27] @Copyright 2017 Wiley. (e) Pair-pulse depression [25] @Copyright 2017 Springer Nature.

2.1 Event-driven sensory outputs

Biological sensory systems work in an event-driven way. Only when the stimuli exceed a threshold, an event is triggered and transmitted to the following units asynchronously, which allows to efficiently extract useful information. For example, conventional frame-based image sensors can produce a large volume of redundant data, imposing a strict requirement on the bandwidth, energy, speed, and storage capacity of the system. Instead of depending on external control signals, the bio-inspired visual sensors (e.g., dynamic vision sensor, asynchronous time-based image sensor, etc.) capture the information depending on the individual pixel intensity [14, 15]. As shown in Figure 2(a) [15], when the change of pixel light intensity exceeds a threshold (positive-ON event, negative-OFF event), the comparator outputs an event spike, in which the timing of spike encodes the light intensity. After that, a reset signal pulls the $V_{\rm diff}$ back to the reset level, starting another cycle. This time-domain spike encoding scheme has the advantage of high efficiency and immunity to voltage degradation, because the spike timing encodes information. However, it is sensitive to noise, caused by the variation of devices and electrical disturbance during the transmission. To improve the robustness, an alternative way is to use the number of spikes in a period of time (rate-encoding), yet at the cost of energy efficiency (Figure 2(b) [14]). After the event-driven sensors, the output sparse spike trains transmit according to address event representation (AER) protocol and are computed by corresponding event-driven processors (e.g., spiking neural network), exhibiting high energy efficiency and speed.

2.2 Low-level processing

Usually, there are a few pre-processing steps before the data are transmitted to the central nervous systems for high-level perception. In this hierarchical architecture, the raw data from the sensors are pre-processed to extract useful information, greatly reducing the volume of data that requires long-distance transmission. Researchers have adopted the so-called near-sensor computing scheme to perform information pre-processing with the advancement of analog VLSI, especially in visual and auditory sensors [16–24]. Recently, various emerging devices (e.g., resistive random access memory (RRAM), electrolyte-based transistors, ionic cable, etc.) show diverse and intriguing dynamics, which can be exploited to pre-process the data in a small footprint [25, 26]. In addition, they can be integrated with sensors and directly process the analog outputs, eliminating the power-consuming ADC/DAC modules.

Filtering is a common method in the data processing. The data in a specific frequency range can be suppressed through the design of filters. Some devices (e.g., RRAM, electrolyte-based transistors, etc.) exhibit short retention of their conductance change when excited by a stimulus. Once the stimulus is removed, its conductance gradually relaxes to the initial value. This short-term memory can be exploited to design different kinds of filters. As shown in Figure 2(c) [10], when the device is initially at a high conductance state, the stimulus induces conductance decrease. And the high-frequency stimulus leads to more decrease. This relationship between conductance change and frequency can realize a low-pass filter. On the contrary, when the device is initially at a low conductance state, the excitatory stimulus can emulate the function of a high-pass filter [27]. Since the effect of stimulus has a close relationship with the initial state of devices, a band-pass filter can be constructed by setting the device at an intermediate state. Many devices show the high-pass filtering feature in the form of pair pulse facilitation (PPF). Due to the residual conductance change induced by the previous stimulus, the following stimulus can give rise to a larger response ($A_2 > A_1$), as shown in Figure 2(d). When the sensors have a transient response to the sensory stimulus, with this short-term memory, the duration of the sensory stimulus can also be recorded.

Some diffusive memristors exhibit short-term depression in the form of pair pulse depression (PPD) [25], as shown in Figure 2(e) [25]. During the longer interval, the device relaxes to a lower conductance state, which means the decay effect is stronger than the stimulus. Therefore, this device shows PPD under a low-frequency stimulus, realizing a high-pass filter. This feature has a significant role in the data processing. By decreasing the response to high-frequency signals, the original signal is scaled, the degree to which is defined by the signal itself [10, 11]. Through this dynamic gain control, the output signal represents the relative values instead of the absolute values. Therefore, the noise, which is relatively small compared to signals, can be precisely differentiated. Notably, the state of the short-term memory reflects the scale degree, which means the original signal can be restored. Taking inspirations from the above-mentioned dynamic adaptation, we can realize sensory computing unit with high sensitivity and wide dynamic range.

Feature extraction is important for information pre-processing. By extracting useful information and building a more concise representation, the data transfer can be greatly reduced. Convolution is a common method widely used in spatial feature extraction, in which the multiplication and accumulation (MAC) operation is the core. The convolution kernel can be mapped into various hardware devices, among which the RRAM crossbar array has been widely used to accelerate the MAC operation in convolutional neural network (CNN) for image pattern classification (Figure 3(a)) [28]. The conductance of the RRAM multiplies with the voltage input, outputting current based on Ohm's law. Then the current is summarized naturally according to Kirchhoff's current law. Therefore, the RRAM array can be programmed to produce different convolution kernels. Apart from that, the axon, or the pathway between receptors and processing units, can execute convolution, leveraging the signal degradation along with the transmission. Due to the decay of excitatory post-synaptic current (EPSC), the signals travelling through longer ionic cable have smaller amplitude when they reach the synapse (Figure 3(b)) [29]. In the multi-gate electrolyte-based transistor, the gates with different distances to the channel have different impacts on the channel conductance [30]. The spatial distribution of gates defines a convolution kernel. However, these pathways are fixed after the fabrication, which means the kernel cannot be trained. An alternative method is to replace the fixed resistor or capacitor with programmable devices, like transistors [31].

2.3 High-level processing

Through the above-mentioned design, the raw data from the sensor outputs can be processed into a more concise form. To process these data for high-level perception, complicated processors are essential. The artificial neural network (ANN) is especially suitable for data-intensive tasks. To build a hardware neural network and reproduce the high efficiency of biological systems, it requires devices with intrinsic similarity to the synapses. RRAM is a promising candidate in neuromorphic computing, which can dynamically change its conductance under external stimuli. Meanwhile, the conductance change of RRAM can retain over a long time, mimicking the in-memory computing nature of synapses (Figure 3(c) [32]). With the compact design of hardware, RRAM-based neural network has shown great performance in

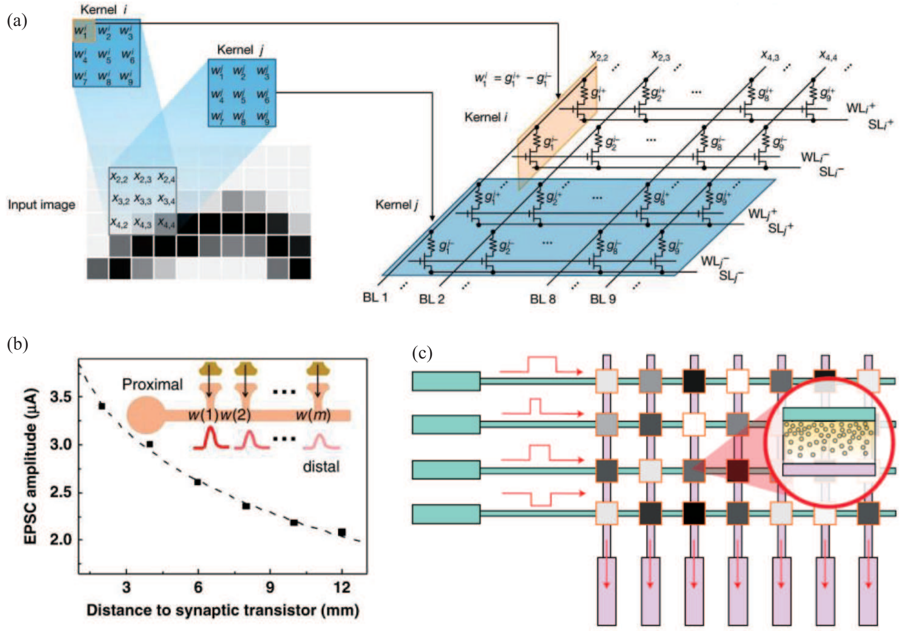

Figure 3 (Color online) (a) The schematic of convolution and its mapping to the RRAM array [28] @Copyright 2020 Springer Nature. (b) The EPSC amplitude decreases as the distance between inputs and synapses [29] @Copyright 2020 Springer Nature. (c) RRAM with in-memory computing characteristic can emulate the synapses [32] @Copyright 2017 Springer Nature.

various fields, e.g., pattern classification [28, 32–36], associative memory [37–39], temporal sequence processing [40], and forecasting [40–42]. Based on the outputs of the neural network, one can make reactions to the environments, such as muscle control.

3 Optical sensory computing

Photon is another information carrier that exhibits high speed and low power consumption, which is suitable for constructing a highly efficient neural network. It has different working principles compared to electronic processing units. In this section, we summarize the optical implementations of important neuromorphic modules, such as synapses and neurons.

3.1 Matrix multiplication

Matrix multiplication is of most importance in the artificial neural network, consisting of basic multiplication and accumulation operations. There are several methods to implement matrix multiplication in optical elements, which are based on different optical effects, such as transmission, diffraction, interference, coupling, and scattering. With the image patterns as the input, the light intensity is used for information representation. The transmission-based method executes multiplication through modulating light intensity via space light modulators (SLM), which represents an interconnection weight. After the SLM, the light signals are collected by photodetectors for accumulation, in which light intensity is transformed to photo-current values. The details of other methods are referred to more coherent papers [13, 43].

Although matrix multiplication is a core operation in convolution, there are some optical elements that are customized for convolution to reduce the cost of sliding convolution kernels. Passive diffraction layer by angle sensitive pixel (ASP) can perform optical edge filtering for the CNN first layer. Chen et al. [44] proposed the concept of ASP-vision in CNN for the first time in 2016. The ASP is a CMOS sensor that is typically composed of two gratings (a diffraction grating and an analyzer grating) above a single photodiode and has a sinusoidal light sensitivity on the basis of the Talbot effect. The Talbot effect refers to that when a plane wave is an incident upon a periodic diffraction grating, the image of the grating is repeated at regular distances away from the grating plane, and the image generated by the Talbot effect shifts horizontally with the angle of the incident light. If the second grating is placed one Taber length after the first one, the intensity of incident light can be adjusted periodically. For a two-dimensional

ASP, the sinusoidal angular response induced by the Talbot effect can be described according to [45]

$$i^{(\alpha,\beta,\gamma)}(\theta) = 1 + m\cos(\beta(\cos(\gamma)\theta_x + \sin(\gamma)\theta_y) + \alpha), \tag{1}$$

where θ_x and θ_y are incident angles, α is the phase shift of the two gratings, β is frequency, and γ is grating orientation. They designed two differential pixels of phase α and phase $\alpha + \pi$ (pixels A and B of Figure 4(c)), and subtracted their responses, to obtain the sinusoidal term of (1). During the capture process, impulse responses are convolved optically with objects in the scene, achieving edge filtering, as shown in Figure 4(d) [44,46]. ASPs can receive and filter the image at the same time. This function layer significantly saves the power consumption of the system and reduces the data bandwidth. Whereas, ASPs will greatly reduce the image resolution and have a great loss of light after the diffraction processing.

Another method of image convolution is using an ordinary lens to produce the Fourier transform (FT) of the complex amplitude of an incident electro-magnetic field. The convolution theorem points out that a convolution in one domain corresponds to a product in another domain, and the FT can convert a signal between time domain and frequency domain. Weaver et al. [47] demonstrated the application of the joint-transform correlator in 1966. Chang et al. [48] reported optical convolutional processing based on diffracted optical elements. In linear optical systems, the $4f$ system of two convex lenses with the focal length of f can realize the Fourier and inverse Fourier transforms in a cascade. By placing a phase plate in the Fourier plane in the middle of the $4f$ system, they modulated the amplitude and phase of the incident light. In order to realize the multiple convolution kernels, the phase plate can be divided into several tiled convolution cores. However, the image resolution of this system also decreases correspondingly, and only non-negative convolution can be obtained. LeCun et al. [49] presented a silicon photonics-based architecture with Mach-Zehnder interferometer (MZI) array in 1995. It similarly exploited the optical Fourier transform and allowed complex data to be encoded in a 2-dimensional grid. Shen et al. [50] demonstrated a programmable nanophotonic processor featuring a cascaded array of 56 programmable MZIs in a silicon photonic integrated circuit in 2017. However, the signals need to be first preprocessed to a high-dimensional vector on a computer and then encoded in the amplitude of optical pulses propagating in the photonic integrated circuit.

3.2 Nonlinear activation functions

Neural networks require continuous differentiable nonlinear activation functions to fit arbitrary complex functions. The types of nonlinear activation functions mainly include a step function, sigmoid function and Tanh (hyperbolic tangent function), threshold linear function, or others [51]. For single nonlinear optical devices, the output power is much lower than the input power and cannot drive even a single neuron, because the nonlinear output comes from the high order terms of the electrical susceptibility. In addition, the output of optical interconnections from one to N computing elements is accompanied by an N-fold loss of light power for each connection [52]. Therefore, nonlinear functions are usually implemented electronically in most hybrid optical neural networks (ONNs). To solve this problem, researchers adopt carrier regeneration approaches for optical amplification in ONNs [53], and propose different nonlinear optical components with the development of optics and materials science.

Since the steady-state power transfer curve of semiconductor optical amplifier (SOA) resembles the upper part of the tanh-curve, SOA is suitable for implementing all-optical nonlinear activation [54]. Hill et al. [53] demonstrated the system of two coupled lasers with SOA, which can provide a useful sigmoid or thresholding function in the optical domain. Kerr effect (a nonlinear variation in the refractive index of a material in response to an applied electric field) in microring resonators [55] and nonlinear fibers [56] can also provide optical power-dependent nonlinear responses. Photorefractive effect is a special phenomenon of photoinduced refractive index change. Photorefractive crystals can produce a great variance in refractive index with a much lower light intensity due to the photoinduced electric field, which is considered as nonlinear activators, but its response time is much longer than the Kerr effect [57]. In addition, Zuo et al. [58] implemented nonlinear optical activation functions with laser-cooled 85Rb atoms in a dark-line two-dimensional magneto-optical trap (MOT) on the basis of electromagnetically induced transparency (EIT), a light-induced quantum interference effect among atomic transitions, in which the atomic medium is opaque to the resonant probe beam without a coupled beam.

Phase change materials (PCM) can be switched between crystalline and amorphous by the input light transmitted in waveguides, which has been adopted in emulating the behavior of spiking neurons.

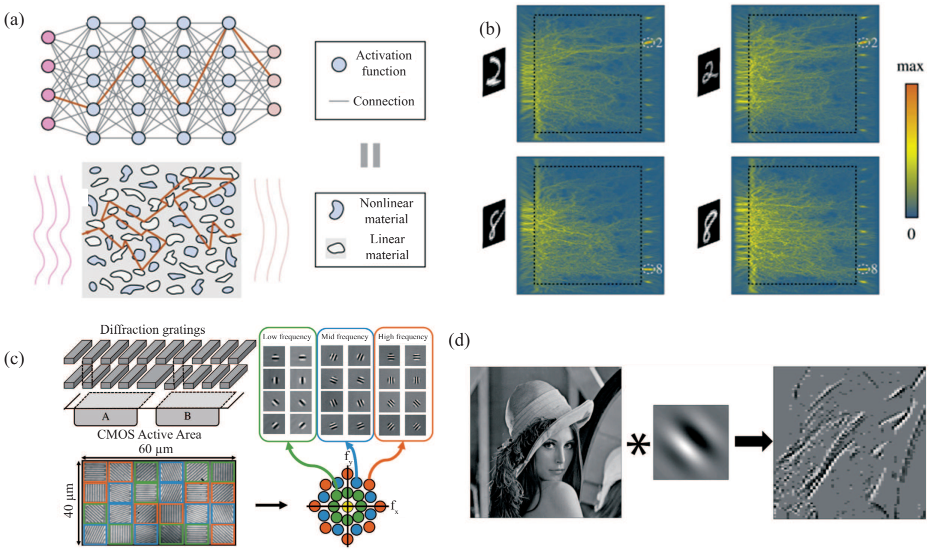

Figure 4 (Color online) (a) Schematic diagrams of the conventional ANN architecture and the passive neural computing through a nanophotonic medium. (b) Schematic illustration of nanophotonic neural medium (NNM) trained to recognize handwritten digits [61] @Copyright 2019 OSA. (c) Differential pixel's impulse responses across an ASP tile. (d) An edge filtered image of a scene after optical convolution [44, 46] @Copyright 2016 IEEE.

Chakraborty et al. [59] proposed an integrate-fire spiking neuron utilizing a $Ge_2Sb_2Te_5$(GST)-ring resonator system to replace rectified linear unit (ReLU) activation function in spiking neural network (SNN). They set the GST elements initially in a crystalline state. The intensity of a single pulse is insufficient to amorphize the GST. When the membrane potential accumulates a few write pulses over a period of time and crosses its threshold, leading the GST to full amorphization. Then the photonic circuit generates a spike. After the neuron fires, a RESET pulse will be passed to reset the states of the devices to their initial states. Feldmann et al. [60] also realized all-optical spiking neurosynaptic networks by PCM and achieved nonlinear activation function by microring resonators integrated with PCM cells. When the PCM element is in the crystalline state, no output can be observed. Conversely, if the instantaneous summed power of the weighted input pulses is high enough to cause the PCM cell to switch to an amorphous state, the light in the output waveguide no longer couple into the ring resonator and the output can be generated satisfying the ReLU function.

In addition, Khoram et al. [61] demonstrated artificial neural computing through a continuous and layer-free nanophotonic medium by leveraging optical reflection and scattering, which enables ultra-high computing density. Light signals come from the left side and interfere strongly by nanostructures (air holes or inclusions) in the SiO_2 host medium, as shown in Figure 4(a). Then they would be guided toward one of several light receivers. The different output positions represent the different numbers in input images. According to the position of the highest share of energy intensity on the right side of the medium, they realized image digital recognition (Figure 4(b) [61]). And the spatial scattering implements the nonlinear operation (ReLUs function) via dye semiconductors or graphene saturable absorbers.

4 Optic-electronic sensory computing

Electrons and photons are two representative information carriers with different properties. Electron carrier allows high integration density for complex computing, based on which diverse processing units can be constructed, such as nonlinear activation functions. However, electronic units suffer from low connectivity and relatively high energy consumption, which restrict the network dimension and computing capability. On the other hand, optical computing exhibits superior performance in connectivity, speed, bandwidth, and power consumption, which is a promising candidate in the artificial neural network. By combining the advantages of electronics and optics, a highly efficient and complex neural network can be obtained. To construct a hybrid optic-electronic hardware platform, optic-electronic devices which transduce light to electronic signals or vice versa are essential.

Optoelectronic devices transduce light to electronic signals, for example, photodiode, optoelectronic resistive random-access memory (ORRAM), and phototransistor. The photodiode is widely used in various fields of photo-detection, due to its simple device structure and fast response. However, the photodiode has a linear relationship between optical input and electrical output, which exhibits optical-to-electronic conversion without any nonlinear computing function. The processing of sensory information requires the complex peripheral circuits or processing unit, which increases footprint, decreases speed, and results in high-power consumption.

As emerging optic-electronic devices, ORRAM and phototransistors based on new materials (e.g., 2D transition-metal dichalcogenides (TMDCs) materials and transition metal oxide) can output the electronic signals with a non-linear relationship to the input light, which enables simple computing functions. The simple two-terminal ORRAM synaptic devices exhibit light-tunable synaptic behaviors. An output image can realize image contrast enhancement through the ORRAM devices array for in-sensor non-linear computing [62]. The three-terminal phototransistor enables to modulate the conductance of the semiconductor channel by light and electrical stimuli. In phototransistors, light stimuli are regarded as pre-synaptic spikes to trigger the synaptic responses. Researchers have mimicked basic synaptic response and unique neural functions using phototransistors [63–65], including neuromorphic reinforcement learning [64], image contrast enhancement [66], colored and color-mixed pattern recognition [67]. The in-sensor computing optic-electronic device provides the potential to simplify the circuitry of a neuromorphic visual system.

5 Multimodal sensory computing

The processes of synthesizing and organizing various huge amounts of inputs are fundamental to effective perception and cognitive functioning. Human can detect and interpret the events from a complex and dynamic environment, and then react depending on the perception results. A single sensory input typically leads to inevitable uncertainties, including the randomness of the signal itself and the noise of the signal processing. From a statistical point of view, a straightforward way of reducing uncertainty and increasing perceptual sensitivity is to combine the information from multiple and independent measurements. Human brains can integrate these inputs from multiple sensory systems, including vision, audition, and somatosensation. This multisensory processing enables appropriate response under complicated circumstances and is helpful for more rapid and accurate information acquisition, in which one sensing channel is inadequate. For example, the integration of information enhances the recognition of objects in the direction of the sound source and shortens the scan time.

5.1 Mathematical models of multisensory integration

The combined outputs of different stimuli rely on not simply linear superposition but the nonlinearity (NL) of weight update. According to psychophysics, the mechanisms for integrating multimodal signals in the brain generally follow two laws. (1) The principle of inverse effectiveness (PoIE), integrated multisensory stimuli are inversely proportional to the effectiveness of the best unisensory response. As the responsiveness to individual sensory stimuli decreases, the strength of multisensory integration increases. Highly prominent individual cues are easily detected, and their combination should have a moderate effect on neural activity. While weaker individual cues elicit relatively few nerve impulses, and the result of integration authentically enhances neuronal stimulation. This effect has a significant positive impact on improving the speed and likelihood of detecting and locating events [68]. (2) The spatial/temporal principle, multisensory integration is more likely or stronger when the constituent unisensory stimuli arise at approximately the same time or the same location, while the combined response becomes suppressed when two stimuli are far enough apart in time or space. This phenomenon can help organisms combine and categorize clues for informed judgment [69]. The output of multisensory integration can be described by additivity index, which is defined as

$$A = \frac{I_{\rm combined}}{\sum_{i=1}^{n} I_{\rm inputi}}, \tag{2}$$

where $I_{\rm combined}$ represents the output intensity of integration of all input stimuli, $I_{\rm inputi}$ represents the output intensity of single stimuli individually. When the additivity index > 1, it means the combined response is greater than the sum of the unisensory responses (super-additivity); when the additivity index < 1, it means the combined response is smaller than the sum of the unisensory responses (sub-additivity).

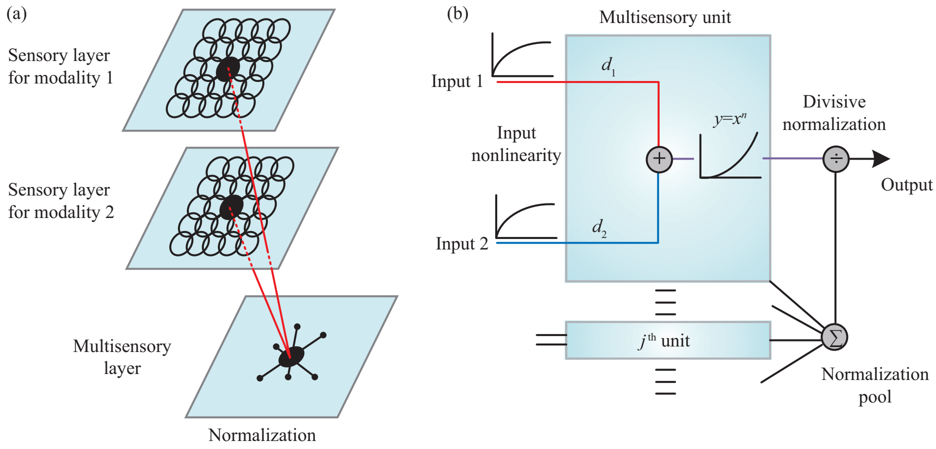

Figure 5 (Color online) Schematic illustration of the normalization model of multisensory integration [72] @Copyright 2011 Springer Nature. (a) Unisensory neurons from separate populations send inputs to a topographically aligned multisensory neuron; (b) the driving input to a particular multisensory neuron is generated by multiplying unisensory inputs by their corresponding synaptic weights and then summing up.

Colonius et al. [70] proposed a time-window-of-integration model related to the spatial/temporal principle from a saccadic experiment early. They found the reaction time to visual targets tends to be fast when the stimuli from another modality are presented in close temporal or spatial proximity. From the viewpoint of mathematics, Ohshiro [71] simulated both the principle of inverse effectiveness and the spatial/temporal principle by a divisive normalization model. Each unimodal input to the spatial integration model is specified by its intensity c and its spatial position in Cartesian coordinates, $\theta = (x_\theta, y_\theta)$, as shown in Figure 5 [72]. The spatial receptive field of each primary sensory neuron is modeled as a two-dimensional Gaussian function. The stimulus intensities also follow the time-dependent Gaussian function. The activity of each neuron is divided by the net activity of all multisensory neurons to produce the final response, which can be represented according to

$$E = d_1 \cdot I_1(x_0, y_0) + d_2 \cdot I_2(x_0, y_0), \tag{3}$$

$$R = \frac{E^n}{\alpha^n + \frac{1}{N}\sum_{j=1}^{N} E_j^N}, \tag{4}$$

where E is the weighted linear sum of its unisensory inputs $I_1(x_0, y_0)$ and $I_2(x_0, y_0)$. The parameters that govern the response of each multisensory neuron include the modality dominance weights (d_1 and d_2), the exponent (n) of the output nonlinearity, the semi-saturation constant (α), and the locations of the receptive fields (x_0, y_0). The semi-saturation constant determines the overall neuron sensitivity to stimulus intensity. A larger value will yield greater super-additivity for a fixed stimulus.

5.2 Multisensor and their integration

The multisensory integration mechanism has inspired the investigation of the devices with the capabilities of sensing two or more physical parameters. In addition to vision, researchers also developed other neural sensing networks, such as biomimetic audiomorphic devices for sound localization, biological semicircular canals for detecting rotational motion, and miniature artificial electronic nose for the detection of mixed harmful gases [73].

Researchers integrated different types of sensor receptors (e.g., mechanoreceptors, thermoreceptors, nociceptors, etc.) by multilayered design layout and connected sensory nodes by meandering wires to achieve multifunctional sensing performance [74,75]. In this way, they realized highly sensitive detections for seven different stimuli, including temperature, strain, humidity, light, magnetic, pressure, and proximity [74]. Lu et al. [75] extended the same strategy to multimodal plant healthcare flexible sensor system based on stacked $ZnIn_2S_4$(ZIS) nanosheets, which can detect both light and humidity. Bao's research team designed and fabricated a multimodal receptor based on the ion relaxation dynamics of a deformable ion conductor, decoupling temperature, and strain sensing in a single unit through different kinds of output signals. The relaxation time and capacitance can be respectively used as a strain-insensitive and temperature-insensitive intrinsic variable for detecting temperature and strain [76].

Yu et al. [77] and Wu et al. [78] demonstrated the integration of vision and touch, which greatly improves the accuracy of handwriting recognition. They integrated flexible triboelectric nanogenerator (TENG) in contact-separation mode on the phototransistor. The phototransistor could be prepared with graphene/MoS_2 heterostructure or lead-free perovskite ($Cs_2AgBiBr_6$). TENG can supply an equivalent gate voltage to drive the synaptic transistor. The output integrated signal corresponding to the EPSC was represented by the channel current of the phototransistor. This synaptic transistor with multi-modal sensory functions realizes the long-term memory and consecutive neural facilitation. Especially, Wu et al. [78] reported the super/sub-additivity under light and touch pulses with the same interval time, which matches the inverse principle. They also observed enhanced EPSC when the temporal interval of these two stimuli decreases, which obeys the temporal principle. This structure is expected to further mimic the behavior of multisensory integration. Also, Wan et al. [29] developed a bimodal artificial sensory neuron that collects optic and pressure information from the photodetector and pressure sensors, respectively. The information is transmitted through two artificial sensory channels (ionic cables) to the electrolyte gated synaptic transistors for further integration and processing. The weight of the input is dependent on the distance between the transistor and the sensor unit. Integrating different stimulation via ion cables is more suitable for universal multimodal signal integration when sensor systems involve more than two different kinds of signals.

There are still challenges for multimodal sensory computing, including the underlying logic of multimodal integration computing and the integration techniques of artificial neurons. It is quite important to control the reception and processing time hysteresis of multiple signals and coordinate the range and precision of different sensory stimuli. Therefore, the raw multimodal data with great heterogeneity should be firstly normalized before subsequent processing. In addition, it is also a grand challenge to identify the direct relationships between the elements from two or more different.

6 Perspective

6.1 Benchmark

The sensory system plays an essential role in our interaction with the environments, in which we acquire information, analyze, and make reactions. With the advances in microelectronics, algorithm, and architecture [28, 79, 80], the performance of the sensory system is under stable progress. Here, we propose some benchmarks to evaluate a sensory system.

The most common task for a sensory system is pattern recognition, e.g., image patterns, speech, odor signal. The training stage consumes substantial power and time, which is usually conducted in the cloud. After the optimization of synaptic weights, the network models are transferred to edge devices, which can distinguish learned patterns. The practical hardware resources determine the size and complexity of the network models, which in turn affect the capability to process high-dimensional patterns. The recognition accuracy of a two-layer perceptron can severely decay as the number of input neurons increase. To effectively process the sensory stimuli from complex environments, it is essential to increase the network size and deepness in edge devices. However, the power consumption and area are two important merits for practical use. The scaling down of feature size in the semiconductor industry enables a more compact and powerful system in edge devices. To evaluate the energy efficiency of a sensory system, we can calculate the average power consumption that the sensory system consumes to recognize one pattern. Since the sensory system is a multi-discipline platform that relates to various signal types, the number of operations varies greatly. In addition, this value can be normalized by the number of pixels or areas for a more effective representation of energy efficiency. The energy efficiency of the sensory system depends on several factors, e.g., the power consumption of the receptors, transmission line, and the architecture of the processing units. In von Neumann computers, the frequent data transfer between computing units and memory can greatly deteriorate the overall energy efficiency, especially for data-intensive tasks. To overcome the bottleneck, optical computing and in-memory computing empowered by memristors are potential candidates. The power-hungry matrix multiplication operations can be executed in a more efficient and parallel way, compared to conventional digital and sequential processing units.

The working speed is of most importance for some interactive systems, which requires fast response in some application scenarios, e.g., auto-driving. We define the working speed as the number of patterns that the system can process in one second, similar to the frame rate in a vision sensor. It depends on the response time of receptors, transmission delay, and the working speed of processing units. The photo-detectors based on conventional semiconductors and emerging 2D materials have a very fast light response (in the timescale of picosecond) [81–88], which allows the realization of a fast vision sensor. However, many chemical and haptic sensors work at a relatively slow speed, compared to the high-speed transmission bus and processing units. As for the transmission, the outputs of slow receptors can be transmitted in a time-multiplexing way, replacing dedicated one-to-one connections by a few metal wires and switches [15]. The speed discrepancy between sensors and processing units allows multiple iteration cycles of processing in one sampling period, which benefits the working speed of the whole system [79]. As for the processing units, the data transfer in von Neumann architecture deteriorates the overall working speed. Besides, the peripheral circuits for analog/digital signal conversion and control logic also play an important role. For parallel computing, the working speed in our definition has a weak relationship with the pixel array size. But the increase in pixel density indirectly influences the working speed due to the subsequent data transmission and processing.

Apart from those basic merits to evaluate a hardware processor, there are some specific merits for a sensory system. Since it interacts with a complex, noisy, and ever-changing environment, high robustness is important for the correct and precise interpretation of sensory information. It depends on the sensitivity of receptors and the tolerance to noise signals. Researchers have been developing emerging materials and device structures for higher response to the sensory stimuli [89–91]. With higher sensitivity, the smaller change in sensory stimuli can be detected and transformed into effective signals for further process. Otherwise, those sensory stimuli will be mixed up with the noise signals, degrading the recognition accuracy of the sensory system. The reluctance to noise signals depends on not only hardware devices but also algorithms. The emerging optic-electronic devices can suppress the noise signals through short-term memory, realizing in-sensor computing [62]. Besides, some pre-processing circuits can be placed near the sensor to suppress the random and low-intensity noise signals. Those hardware implementations greatly enhance the contrast of input patterns and reduce the pattern dimension, contributing to high performance and energy efficiency. As for algorithm, although artificial neural network has relatively high robustness to various noise signals, the performance can degrade quickly at high noise level. The bio-inspired spiking neural network has better performance in recognizing noise-interfered patterns [79]. It can efficiently discriminate between random noise signals and coherent pattern signals, which is a powerful candidate for processing complex sensory information.

Another important merit is the dynamic range, which indicates the intensity range that the sensory system can effectively process. The dynamic range of the single receptor is determined by the minimal and maximum intensity with an effective response. The dynamic range of a single receptor is usually limited. Through circuit and algorithm design for wide-range adaptation, e.g., dynamic gain control, the dynamic range of the whole sensory system can be effectively enlarged [10, 11, 92, 93].

6.2 Outlook

Apart from the improvement in recognition accuracy, the sensory system is expected to deliver high-level intelligence like human, such as the online learning capability. The sensory system should continuously adjust the network parameters to adapt to the ever-changing environments. The practical circumstance can be very different from the training patterns, which possibly leads to serious failure. First, the basic devices should be programmable so that their states can be modulated by external stimuli. Second, the architecture of the processing units should be carefully designed to execute training algorithms, e.g., the loss functions and error backpropagation. However, the training stage can consume lots of energy and time, which poses a trade-off between online learning capability and hardware constraints. In this sense, a spiking neural network that exhibits event-driven capability and local learning rules is more suited for sensory system [79].

Many existing studies focused on performance improvement in the single-modal sensory system. However, it is more reliable to rely on multiple sensory channels. On the one hand, the performance of the artificial sensory system is much weaker than that of humans. Combining multiple sensory stimuli can greatly enhance the overall sensitivity and accuracy. On the other hand, a single sensory input leads to inevitable uncertainties, which can induce serious accidents when applied in practice. Therefore, multi-modal sensory computing that effectively processes different types of sensory stimuli is essential to react to complicated circumstances.

Bio-inspired neuromorphic computing is a powerful tool for processing enormous data, especially for sensory stimuli that are unstructured, contain enormous random noise signals, and change continuously. However, the performance of artificial sensory system still requires substantial improvement in accuracy, speed, energy efficiency, footprint, and high-level intelligence. This field calls for collaboration between researchers from different fields (e.g., materials, chemistry, biology, mechanics, engineering, neuroscience, etc.), which is crucial for the practical application of the sensory system.

Acknowledgements This work was supported by Research Grant Council of Hong Kong (Grant No. PolyU 15205619), Science, Technology and Innovation Commission of Shenzhen (Grant No. JCYJ20180507183424383), and Hong Kong Polytechnic University (Grant No. 1-ZE1T).

References

1 Ji X L, Zhao X Y, Tan M C, et al. Artificial perception built on memristive system: visual, auditory, and tactile sensations. Adv Intell Syst, 2020, 2: 1900118
2 Wan C L, Cai P Q, Wang M, et al. Artificial sensory memory. Adv Mater, 2020, 32: 1902434
3 Chai Y. In-sensor computing for machine vision. Nature, 2020, 579: 32–33
4 Zhou F C, Chai Y. Near-sensor and in-sensor computing. Nat Electron, 2020, 3: 664–671
5 Liao F Y, Zhou F C, Chai Y. Neuromorphic vision sensors: principle, progress and perspectives. J Semicond, 2021, 42: 013105
6 Cottini N, Gottardi M, Massari N, et al. A 33 μW 64×64 pixel vision sensor embedding robust dynamic background subtraction for event detection and scene interpretation. IEEE J Solid-State Circ, 2013, 48: 850–863
7 Lichtsteiner P, Posch C, Delbruck T. A 128×128 120 dB 15 μs latency asynchronous temporal contrast vision sensor. IEEE J Solid-State Circ, 2008, 43: 566–576
8 Moini A. Vision Chips2000. Berlin: Springer, 2000
9 Ruedi P F, Heim P, Kaess F, et al. A 128×128 pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction. IEEE J Solid-State Circ, 2003, 38: 2325–2333
10 Dittman J S, Kreitzer A C, Regehr W G. Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. J Neurosci, 2000, 20: 1374–1385
11 Abbott L F, Varela J A, Sen K, et al. Synaptic depression and cortical gain control. Science, 1997, 275: 221–224
12 Rothman J S, Cathala L, Steuber V, et al. Synaptic depression enables neuronal gain control. Nature, 2009, 457: 1015–1018
13 Shastri B J, Tait A N, de Lima T F, et al. Photonics for artificial intelligence and neuromorphic computing. Nat Photonics, 2021, 15: 102–114
14 Chen D G, Matolin D, Bermak A, et al. Pulse-modulation imaging-review and performance analysis. IEEE Trans Biomed Circ Syst, 2011, 5: 64–82
15 Posch C, Serrano-Gotarredona T, Linares-Barranco B, et al. Retinomorphic event-based vision sensors: bioinspired cameras with spiking output. Proc IEEE, 2014, 102: 1470–1484
16 Finateu T, Niwa A, Matolin D, et al. A 1280×720 back-illuminated stacked temporal contrast event-based vision sensor with 4.86 μm pixels, 1.066GEPS readout, programmable event-rate controller and compressive data-formatting pipeline. In: Proceedings of IEEE International Solid-State Circuits Conference, San Francisco, 2020
17 Hasler P, Smith P D, Graham D, et al. Analog floating-gate, on-chip auditory sensing system interfaces. IEEE Sens J, 2005, 5: 1027–1034
18 Hsu T H, Chen Y K, Wu J S, et al. A 0.8 V multimode vision sensor for motion and saliency detection with ping-pong PWM pixel. In: Proceedings of IEEE International Solid-State Circuits Conference, San Francisco, 2020
19 Jimenez-Fernandez A, Cerezuela-Escudero E, Miro-Amarante L, et al. A binaural neuromorphic auditory sensor for FPGA: a spike signal processing approach. IEEE Trans Neural Netw Learn Syst, 2017, 28: 804–818
20 Kyuma K, Lange E, Ohta J, et al. Artificial retinas — fast, versatile image processors. Nature, 1994, 372: 197–198
21 Lichtsteiner P, Posch C, Delbruck T. A 128×128 120 dB 30 mW asynchronous vision sensor that responds to relative intensity change. In: Proceedings of IEEE International Solid-State Circuits Conference, San Francisco, 2006
22 Lyon R F, Mead C. An analog electronic cochlea. IEEE Trans Acoust Speech Signal Process, 1988, 36: 1119–1134
23 Song Y M, Xie Y, Malyarchuk V, et al. Digital cameras with designs inspired by the arthropod eye. Nature, 2013, 497: 95–99
24 Wen B, Boahen K. A 360-channel speech preprocessor that emulates the cochlear amplifier. In: Proceedings of IEEE International Solid State Circuits Conference, San Francisco, 2006
25 Wang Z R, Joshi S, Savel'ev S E, et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. Nat Mater, 2017, 16: 101–108
26 Ling H F, Koutsouras D A, Kazemzadeh S, et al. Electrolyte-gated transistors for synaptic electronics, neuromorphic computing, and adaptable biointerfacing. Appl Phys Rev, 2020, 7: 011307
27 Zang Y P, Shen H G, Huang D Z, et al. A dual-organic-transistor-based tactile-perception system with signal-processing functionality. Adv Mater, 2017, 29: 1606088
28 Yao P, Wu H Q, Gao B, et al. Fully hardware-implemented memristor convolutional neural network. Nature, 2020, 577: 641–646
29 Wan C J, Cai P Q, Guo X T, et al. An artificial sensory neuron with visual-haptic fusion. Nat Commun, 2020, 11: 4602
30 He Y L, Nie S, Liu R, et al. Spatiotemporal information processing emulated by multiterminal neuro-transistor networks. Adv Mater, 2019, 31: 1900903
31 Das S, Dodda A, Das S. A biomimetic 2D transistor for audiomorphic computing. Nat Commun, 2019, 10: 3450
32 Sheridan P M, Cai F X, Du C, et al. Sparse coding with memristor networks. Nat Nanotech, 2017, 12: 784–789
33 Li C, Belkin D, Li Y N, et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. Nat Commun, 2018, 9: 2385
34 Prezioso M, Merrikh-Bayat F, Hoskins B D, et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature, 2015, 521: 61–64
35 Wang Z R, Joshi S, Savel'ev S, et al. Fully memristive neural networks for pattern classification with unsupervised learning. Nat Electron, 2018, 1: 137–145
36 Serb A, Bill J, Khiat A, et al. Unsupervised learning in probabilistic neural networks with multi-state metal-oxide memristive synapses. Nat Commun, 2016, 7: 12611
37 Eryilmaz S B, Kuzum D, Jeyasingh R, et al. Brain-like associative learning using a nanoscale non-volatile phase change synaptic device array. Front Neurosci, 2014, 8: 205

38 Hu S G, Liu Y, Liu Z, et al. Associative memory realized by a reconfigurable memristive Hopfield neural network. Nat Commun, 2015, 6: 7522

39 Pershin Y V, Di Ventra M. Experimental demonstration of associative memory with memristive neural networks. Neural Netw, 2010, 23: 881–886

40 Moon J, Ma W, Shin J H, et al. Temporal data classification and forecasting using a memristor-based reservoir computing system. Nat Electron, 2019, 2: 480–487

41 Marinella M J, Agarwal S. Efficient reservoir computing with memristors. Nat Electron, 2019, 2: 437–438

42 Zhong Y N, Tang J S, Li X Y, et al. Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing. Nat Commun, 2021, 12: 408

43 Xu R Q, Lv P, Xu F J, et al. A survey of approaches for implementing optical neural networks. Opt Laser Tech, 2021, 136: 106787

44 Chen H G, Jayasuriya S, Yang J Y, et al. ASP vision: optically computing the first layer of convolutional neural networks using angle sensitive pixels. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. 27–30

45 Hirsch M, Sivaramakrishnan S, Jayasuriya S, et al. A switchable light field camera architecture with angle sensitive pixels and dictionary-based sparse coding. In: Proceedings of IEEE International Conference on Computational Photography (ICCP), 2014

46 Wang A, Sivaramakrishnan S, Molnar A. A 180 nm CMOS image sensor with on-chip optoelectronic image compression. In: Proceedings of IEEE Custom Integrated Circuits Conference, 2012

47 Weaver C S, Goodman J W. A technique for optically convolving two functions. Appl Opt, 1966, 5: 1248

48 Chang J L, Sitzmann V, Dun X, et al. Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. Sci Rep, 2018, 8: 12324

49 LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: Handbook of Brain Theory and Neural Networks. Cambridge: MIT Press, 1995

50 Shen Y C, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits. Nat Photon, 2017, 11: 441–446

51 Bueno S, Salmeron J L. Benchmarking main activation functions in fuzzy cognitive maps. Expert Syst Appl, 2009, 36: 5221–5229

52 Goodman J W. Fan-in and fan-out with optical interconnections. Opt Acta-Int J Opt, 1985, 32: 1489–1496

53 Hill M T, Frietman E E E, de Waardt H, et al. All fiber-optic neural network using coupled SOA based ring lasers. IEEE Trans Neural Netw, 2002, 13: 1504–1513

54 Vandoorne K, Dierckx W, Schrauwen B, et al. Toward optical signal processing using photonic reservoir computing. Opt Express, 2008, 16: 11182

55 Mesaritakis C, Papataxiarhis V, Syvridis D. Micro ring resonators as building blocks for an all-optical high-speed reservoir-computing bit-pattern-recognition system. J Opt Soc Am B, 2013, 30: 3048

56 Rosenbluth D, Kravtsov K, Fok M P, et al. A high performance photonic pulse processing device. Opt Express, 2009, 17: 22767

57 Yan T, Wu J M, Zhou T K, et al. Fourier-space diffractive deep neural network. Phys Rev Lett, 2019, 123: 023901

58 Zuo Y, Li B H, Zhao Y J, et al. All-optical neural network with nonlinear activation functions. Optica, 2019, 6: 1132

59 Chakraborty I, Saha G, Sengupta A, et al. Toward fast neural computing using all-photonic phase change spiking neurons. Sci Rep, 2018, 8: 12980

60 Feldmann J, Youngblood N, Wright C D, et al. All-optical spiking neurosynaptic networks with self-learning capabilities. Nature, 2019, 569: 208–214

61 Khoram E, Chen A, Liu D J, et al. Nanophotonic media for artificial neural inference. Photon Res, 2019, 7: 823

62 Zhou F C, Zhou Z, Chen J W, et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. Nat Nanotechnol, 2019, 14: 776–782

63 Guo Y B, Zhu L Q. Recent progress in optoelectronic neuromorphic devices. Chin Phys B, 2020, 29: 078502

64 Zhu Q B, Li B, Yang D D, et al. A flexible ultrasensitive optoelectronic sensor array for neuromorphic vision systems. Nat Commun, 2021, 12: 1798

65 Islam M M, Dev D, Krishnaprasad A, et al. Optoelectronic synapse using monolayer MoS_2 field effect transistors. Sci Rep, 2020, 10: 21870

66 Choi C, Leem J, Kim M S, et al. Curved neuromorphic image sensor array using a MoS_2-organic heterostructure inspired by the human visual recognition system. Nat Commun, 2020, 11: 5934

67 Seo S, Jo S H, Kim S, et al. Artificial optic-neural synapse for colored and color-mixed pattern recognition. Nat Commun, 2018, 9: 5106

68 Stein B E, Stanford T R. Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci, 2008, 9: 255–266

69 Holmes N P. The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. Neuropsychologia, 2007, 45: 3340–3345

70 Colonius H, Diederich A. Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J Cogn Neurosci, 2004, 16: 1000–1009

71 Ohshiro T, Angelaki D E, DeAngelis G C. A normalization model of multisensory integration. Nat Neurosci, 2011, 14: 775–782

72 Fetsch C R, DeAngelis G C, Angelaki D E. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. Nat Rev Neurosci, 2013, 14: 429–442

73 Zhang J Y, Xue Y Y, Sun Q Y, et al. A miniaturized electronic nose with artificial neural network for anti-interference detection of mixed indoor hazardous gases. Sens Actuat B-Chem, 2021, 326: 128822

74 Hua Q L, Sun J L, Liu H T, et al. Skin-inspired highly stretchable and conformable matrix networks for multifunctional sensing. Nat Commun, 2018, 9: 244

75 Lu Y Y, Xu K C, Zhang L S, et al. Multimodal plant healthcare flexible sensor system. ACS Nano, 2020, 14: 10966–10975

76 You I, Mackanic D G, Matsuhisa N, et al. Artificial multimodal receptors based on ion relaxation dynamics. Science, 2020, 370: 961–965

77 Yu J R, Yang X X, Gao G Y, et al. Bioinspired mechano-photonic artificial synapse based on graphene/MoS_2 heterostructure. Sci Adv, 2021, 7: 9117

78 Wu X M, Li E L, Liu Y Q, et al. Artificial multisensory integration nervous system with haptic and iconic perception behaviors. Nano Energy, 2021, 85: 106000

79 Imam N, Cleland T A. Rapid online learning and robust recall in a neuromorphic olfactory circuit. Nat Mach Intell, 2020, 2: 181–191
80 Pei J, Deng L, Song S, et al. Towards artificial general intelligence with hybrid Tianjic chip architecture. Nature, 2019, 572: 106–111
81 Koscielniak W C, Pelouard J L, Littlejohn M A. Dynamic behavior of photocarriers in a GaAs metal-semiconductor-metal photodetector with sub-half-micron electrode pattern. Appl Phys Lett, 1989, 54: 567–569
82 Beling A, Campbell J C. InP-based high-speed photodetectors. J Lightwave Technol, 2009, 27: 343–355
83 Mueller T, Xia F N, Avouris P. Graphene photodetectors for high-speed optical communications. Nat Photon, 2010, 4: 297–301
84 Urich A, Unterrainer K, Mueller T. Intrinsic response time of graphene photodetectors. Nano Lett, 2011, 11: 2804–2808
85 Xia F N, Mueller T, Lin Y M, et al. Ultrafast graphene photodetector. Nat Nanotech, 2009, 4: 839–843
86 Youngblood N, Li M. Ultrafast photocurrent measurements of a black phosphorus photodetector. Appl Phys Lett, 2017, 110: 051102
87 Park H L, Kim H, Lim D, et al. Retina-inspired carbon nitride-based photonic synapses for selective detection of UV light. Adv Mater, 2020, 32: 1906899
88 Kim S G, Kim S H, Park J, et al. Infrared detectable MoS_2 phototransistor and its application to artificial multilevel optic-neural synapse. ACS Nano, 2019, 13: 10294–10300
89 Qian C, Choi Y, Choi Y J, et al. Oxygen-detecting synaptic device for realization of artificial autonomic nervous system for maintaining oxygen homeostasis. Adv Mater, 2020, 32: 2002653
90 Yang J, Chen J, Su Y J, et al. Eardrum-inspired active sensors for self-powered cardiovascular system characterization and throat-attached anti-interference voice recognition. Adv Mater, 2015, 27: 1316–1326
91 Zhu B W, Wang H, Liu Y Q, et al. Skin-inspired haptic memory arrays with an electrically reconfigurable architecture. Adv Mater, 2016, 28: 1559–1566
92 Ledda P, Santos L P, Chalmers A. A local model of eye adaptation for high dynamic range images. In: Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, 2004. 151–160
93 Pattanaik S N, Ferwerda J A, Fairchild M D, et al. A multiscale model of adaptation and spatial vision for realistic image display. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, 1998. 287–298