

# A Spiking Visual Neuron for Depth Perceptual Systems

**Chunsheng Chen**

Nanjing University

**Yongli He**

Nanjing University

**Huiwu Mao**

Nanjing University

**Li Zhu**

Nanjing University

**Xiangjing Wang**

Nanjing University

**Ying Zhu**

Nanjing University

**Yixin Zhu**

Nanjing University

**Yi Shi**

Nanjing University <https://orcid.org/0000-0002-1082-0315>

**Changjin Wan**

Nanyang Technological University

**Qing Wan** (✉ [wanqing@nju.edu.cn](mailto:wanqing@nju.edu.cn))

Nanjing University

---

## Article

### Keywords:

**Posted Date:** November 2nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-990982/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# A Spiking Visual Neuron for Depth Perceptual Systems

Chunsheng Chen<sup>1</sup>, Yongli He<sup>1</sup>, Huiwu Mao<sup>1</sup>, Li Zhu<sup>1</sup>, Xiangjing Wang<sup>1</sup>, Ying Zhu<sup>1</sup>, Yixin Zhu<sup>1</sup>, Yi Shi<sup>1</sup>, Changjin Wan<sup>1\*</sup>, and Qing Wan<sup>1\*</sup>

<sup>1</sup>School of Electronic Science & Engineering, and Collaborative Innovation Centre of Advanced Microstructures, Nanjing University, Nanjing 210093, China

\*Corresponding author. Email: [wanqing@nju.edu.cn](mailto:wanqing@nju.edu.cn); [cjwan@nju.edu.cn](mailto:cjwan@nju.edu.cn).

## Abstract

The biological visual system encodes information into spikes and processes them parallelly by the neural network, which enables the perception with high throughput of visual information processing at an energy budget of a few watts. The parallelism and efficiency of bio-visual system motivates electronic implementation of this biological computing paradigm, which is challenged by the lack of bionic devices, such as spiking neurons that can mimic its biological counterpart. Here, we present a highly bio-realistic spiking visual neuron based on an Ag/TaO<sub>x</sub>/ITO memristor. Such spiking visual neuron collects visual information by a photodetector, encodes them into action potentials through the memristive spiking encoder, and interprets them for recognition tasks based on a network of neuromorphic transistors. The firing spikes generated by the memristive spiking encoders have a frequency range of 1-200 Hz and sub-micro watts power consumption, very close to the biological counterparts. Furthermore, a spiking visual system is demonstrated, replicating the distance-dependent response and eye fatigue of biological visual systems. The mimicked depth perception shows a recognition improvement by adapting to sights at different distance. Our design

presents a fundamental building block for energy-efficient and biologically plausible artificial visual systems.

## **Introduction**

Whereas information in a conventional sensing system is generally processed in sequence and encoded by amplitude, biological systems operate parallelly and use the “spike” as its language<sup>1</sup>. As a result, the biological systems show unparalleled energy efficiency compared to electronic counterparts. Hence, there is a trend to develop systems implicated in biologically plausible or bioinspired methodologies, which has boomed many fields including neuromorphic engineering<sup>2, 3</sup>, neurorobotics<sup>4</sup>, and bioinspired sensing<sup>5-7</sup>. Spiking neural networks (SNNs) with higher biological fidelity are thought to be a promising competitor to classical artificial neural networks (ANNs) for further improvement of computing efficiency<sup>8,9</sup>. Therefore, many efforts have been made recently to build a SNN-based artificial visual system<sup>10</sup>. More importantly, visual information constitutes the largest portion of what we perceive in our interaction with the environment. There is ~1 Gigabit of visual information transmitted through the retina per second, and the visual system can organize, identify, and interpret such a high volume of information within an energy budget of a few watts<sup>11</sup>. Specifically, the biological visual systems rely on spiking bipolar cells, as light waves from the object enter the eye through the cornea and reach the retina where the photoreceptors respond to light and influence the membrane potential of the bipolar cells. Then the potential spikes are generated, conveying the visual information through the retinofugal pathway to the central neural systems, which can be well interpreted by the so-called leaky integrate-and-fire (LIF) models<sup>10, 12</sup>.

Hence a spiking visual neuron (sVN) with high biological plausibility is crucial for building an SNN-based artificial visual system. The spiking component can be achieved by classical oscillator circuits such as ring oscillators with multiple transistors. A carbon nanotube (CNT)-based three-stage ring oscillator has been exploited in an artificial tactile neuron with an active/inactive energy consumption of several microwatts<sup>13, 14</sup>. Recent achievements on threshold switching (TS) memory have been demonstrated, featuring good scalability and thus integration density. In addition, they can well mimic the oscillation behaviors of neurons based on the intrinsic ionic dynamics. For example, a hafnium-based TS device that takes the advantage of Ag filamentary rupture/formation successfully achieved self-oscillation behaviors with a low operation voltage of <0.6 V and power consumption of  $\sim 1.8 \mu\text{W}/\text{spike}$ <sup>15</sup>. Although there is a wide spectrum of materials including transition metal oxide (e.g.,  $\text{HfO}_2$  and  $\text{TaO}_x$ )<sup>16-20</sup>, phase change materials (e.g.,  $\text{Ge}_2\text{Sb}_2\text{Te}_5$ )<sup>21, 22</sup>, insulator–metal-transition (e.g.,  $\text{NbO}_2$  and  $\text{VO}_2$ )<sup>23-25</sup>, the oscillation parameters like frequency range and power consumption of current spiking encoder devices are still far from that of their biological counterparts. Furthermore, how to mimic the rich neural behaviors, such as depth perception, and how to leverage such behaviors to benefit neuromorphic visual computing are still lacking.

Here we demonstrate an Ag/TaO<sub>x</sub>/ITO memristor-based spiking visual neuron with high biological plausibility. Such a memristor shows excellent TS characteristics including steep slope (1.37 mV/dec) and fast turn-on/off delay time (10/40 ns). The TS-based spiking encoder exhibits four critical features of action-potential-based

computing: the all-or-nothing spiking of an action potential, threshold-driven spiking, a refractory period, and strength-modulated frequency response. The TS-based visual neuron contains three components: the artificial photoreceptor, the TS-based spiking encoder, and the neural network, which mimics biological behaviors like a sigmoid response to increasing light intensity and eye fatigue. It shows a frequency range of 1-200 Hz and the sub-microwatt power consumption, representing a step forward to the biological comparable performance. A binocular visual system built by such an artificial visual neuron has been verified by simulation, showing a recognition improvement by refocusing on sights at different distances. Our design presents a fundamental building block for energy-efficient and biologically plausible artificial visual systems.

## **Results**

In biological visual systems, photoreceptor cells convert external optical stimuli into spiking potentials which are eventually interpreted and processed by the visual cortex. A biological neuron receives input spikes from other neurons through connected synapses and then integrates all the input spikes. A firing spike of membrane potential is generated when the integrated potentials collected from the dendrites exceed the threshold<sup>26-28</sup>. The leaky integrate-and-fire (LIF) neuron model could well describe such a spiking generation process. In parallel to the biological visual system, a spiking visual neural network (sVNN) is proposed, as shown in the bottom panel of Fig.1a. The sVNN consists of sensory transduction of external stimuli to artificial receptors, the

subsequent LIF neurons functioning as spiking encoders, and neuromorphic computing for deep processing.

Firstly, a spiking encoder is built based on the magnetron sputtering deposited Ag/TaO<sub>x</sub>/ITO memristor as shown in Fig. 1b. The top and bottom electrodes of the memristor implement the functions of the dendrites and axons of a neuron. The active layer acts as the soma of neurons which integrates signals from dendrites with the assistance of a parallel capacitor, representing the plasma membrane. The SEM image of the cross-section of the Ag/TaO<sub>x</sub>/ITO memristor confirms a vertically layered structure, shown in Fig.1b along with a structural schematic in the inset. In the 100 nm-thick oxide layer, the existence of Ta/O is verified by XPS in supplementary materials. This diffusive memristor has an electrochemical metallization (EM) based switching property<sup>29, 30</sup>. The Ag ions move from the top Ag electrode to the bottom electrode by electro-chemical-thermal driving force under a bias voltage. The Ag atoms form localized conductive filaments at a high electric field with the displacement and drift of Ag atoms. When the voltage reaches the threshold voltage ( $V_{TH}$ ) of 0.2~0.4 V and forms a conductive filament, the memristor turns on and switches sharply from a high resistance state (HRS) to a low resistance state (LRS), as shown in Fig.1c. The voltage of the conductive filament formation will increase with thickening the oxide layer. When the voltage drops below the hold voltage ( $V_{Hold}$ ) of 0.1 V, the memristor will return to OFF-state (HRS). A compliance current ( $I_{cc}$ ) of 1  $\mu$ A is necessary during the electroforming process to activate the volatile threshold switching behavior. The threshold switch (TS) characteristics exhibit no obvious degradation during 1,000

consecutive cycles. These characteristics prove the memristor is reliable and durable. The current abruptly jumps from  $2 \times 10^{-10}$  to  $1 \times 10^{-6}$  A within 5 mV during the DC voltage sweep, indicating a steep switching slope of 1.37 mV/dec. The ON-OFF sweeping speed of the memristor is detected utilizing pulse measurement. The memristive switching can be triggered by a voltage pulse with an amplitude of 1 V and a width of 2  $\mu$ s, and the turn-on delay time is only 10 ns, as shown in Fig. 1d. In contrast, the memristor quickly returns to the OFF-state with a 40ns delay under a read voltage of 0.1 V.

A biological neuron generally fires with a rate of around 1-200 spike/s and power consumption of  $\sim 10^{-10}$  W/spike<sup>31-33</sup>. In contrast, an electronic spiking encoder was developed to fire at a much higher rate (e. g. up to 250 MHz<sup>34</sup>) to achieve fast encoding and low energy consumption per spike (the lower pulse width the lower energy per spike), which are required for hardware acceleration of neuromorphic computing. However, this may not be energy efficient compared to the biological counterpart, due to the much higher power consumption and firing rate. Furthermore, the frequency and power consumption of these devices are very different from those of the biological neuron, making them difficult to be deployed in bio-plausible systems and challenging to interface with biological neural networks. Fig.1e shows a comparison of benchmarks between several kinds of LIF spiking encoder and biological neurons<sup>15, 19, 20, 24, 25, 35-38</sup>. Our work features both biological plausibility with a spiking rate of 1-200 spikes/s and low power consumption of 0.25-0.5  $\mu$ W/spike.



The circuit for LIF spiking encoder is shown in Fig.S2. The EM threshold switch memristor is connected in series with a resistor ( $R_0=200\text{ k}\Omega$ ), representing an artificial dendrite that integrates input spikes from the presynaptic neurons, and in parallel with a capacitor ( $C=10\text{ nF}$ ) to work as an artificial axon for spike generation. The input spikes were applied to this component through a resistor ( $R_S=10\text{ M}\Omega$ ), and the output spikes were measured on  $R_0$ . The operational mechanism is divided into the integration process and the firing process, the same as that of a biological neuron. During the integration process, input pulses are applied to the neural circuit to charge the capacitor until the TSM voltage reaches  $V_{TH}$ , and the memristor switches from the HRS to LRS with the formation of Ag filaments. During the neuron firing, the charge accumulated on the capacitor is released through the memristor, and a spike voltage on  $R_0$  is detected. When the TSM voltage decreases below  $V_{Hold}$ , the device reinstates to the HRS due to the rupture of Ag filament. The EM threshold switch operates through the formation and dissolution of the metallic filament in the interlayer. Fig. 2a illustrates the leaky integrate-and-fire behavior of TSM neurons by applying a train of voltage spikes. When the series of voltage spikes (1 kHz, 500  $\mu$ s, 5V) is fed to the node “IN”, the capacitor will accumulate the charges and raise the voltage potential on the capacitor during the integration process. The TSM remains in the HRS ( $R_H=1\text{ G}\Omega$ ) in the charging process, implying a negligible voltage on  $R_0$ . When the voltage on the capacitor goes beyond the memristor reaches the  $V_{TH}$ , the memristor switches from HRS to LRS. In contrast, the RC constant ( $T_C=R_S\times C$ ) in the charging loop is much longer than that ( $T_D=(R_L+R_0)\times C$ ) in the discharging loop as the device is in the LRS ( $R_L\sim 3\text{ k}\Omega$ ). In this case, the

capacitor will discharge, and the neuron will fire an output voltage spike. Along with the discharging process, the voltage on the memristor and  $R_0$  gradually decreases. When the voltage on the memristor falls below  $V_{\text{Hold}}$ , the TSM will spontaneously relax back to the HRS again and the neural circuit will start the next integration cycle. Under fixed voltage pulses, the neuron spikes at a constant frequency. Larger voltage pulses would reduce the charging time to reach the threshold potential, thereby increasing the frequency of the firing spike. In Fig.2b, consecutive input pulses with different amplitudes ranging from 2V~5 V are applied to the TSM neuron. The output spiking frequency increases with the increase of input amplitude, as shown in Fig.2c. For example, the firing rate is as low as 1 spike/s with 2.5 V input pulses. When the input amplitude reaches 5 V, the highest firing rate is 10 spikes/s. This is mainly due to the higher charging speed of a capacitor under higher input voltage amplitude. Fig.2c shows the statistical relation between voltage and spike-frequency of the neuron. During the capacitor charging towards the threshold voltage, a smaller voltage requires more charging cycles and leads to a smaller firing rate. The strength-modulation of spiking frequency in biological neurons are realized by the modulating the amplitude of input voltage spikes<sup>31</sup>. Besides, an increasing RC time constant means a longer integration time. Fig. 2d demonstrates that a larger RC constant decreases the output spiking frequency. The firing rates decreases from 200 spikes/s to 1 spikes/s as the RC constant raises from 10  $M\Omega \cdot nF$  to 1000  $M\Omega \cdot nF$ . Through engineering the circuit parameters, our system can reduce energy consumption and improve efficiency.

The human brain receives perceptual information through sensory organs for cognitive processing<sup>39</sup>. Cone cells in human eyes detect light signals and then integrate the signals before propagating them to the visual cortex of the brain. The ciliary muscle of the eyes is thus adjusted to adapt either near vision or distance vision, as shown in Fig.3a. For near vision, the ciliary muscle shrinks outward and the crystalline lens becomes a flatter surface. For distance vision, the crystalline lens is squeezed by the muscle to become rounder and rounder. As a result, both the near and distance vision can be focused clearly on the retinal plane. To simulate the function of the human visual system, a spiking visual neuron (sVN) was built based on our LIF spiking encoder as shown in Fig.3b. A photoresistor (Light:  $R_P=10\text{ k}\Omega$ , Dark:  $R_P=100\text{ M}\Omega$ ) and a resistor ( $R_1=1\text{ M}\Omega$ ) are connected in series, and the voltage on the resistor is used as the input of the electronic circuit. While the light intensity increases from 0 to  $7.2\text{ mW/cm}^2$ , modulating  $R_P$  from  $100\text{ M}\Omega$  to  $10\text{ k}\Omega$ , the voltage on  $R_1$  rises from 1% to 99% of  $V_{DD}$ . The pulse voltage on  $R_1$  gradually increases and the capacitor starts being charged when the photoresistor is illuminated by pulses. This charging cycle is the same as that in the electric neuron circuit. The voltage on the capacitor rises during the integration process until the memristor reaches the  $V_{TH}$  and switches from HRS to LRS. As a result, the capacitor will discharge, and an output voltage spike will be produced in the firing process, as shown in Fig.3c. After the discharging process, another charging cycle starts subsequently. The light intensity is inversely proportional to the square of the distance between the light source and photoresistor as well as the resistance of the photoresistor. As the distance increases, the light intensity on the photoresistor decreases and  $R_P$

increases in turn. Consequently, the voltage on  $R_1$  decreases, and the firing rate in Fig.3d decreases as well with a sigmoid shape. When biological neurons are overused for a long time, fatigue would occur, resulting in performance degradation with bad recognition accuracy. This is because the neurotransmitter, which triggers the membrane potential of the cone cell, would not be released until a segment of light pulses is continuously superimposed<sup>40, 41</sup>. Our sVN can mimic such behavior upon continuous illumination. After  $10^5$  cycle of light pulses (Intensity=6 mW/cm<sup>2</sup>, duration=5 ms, frequency=100 Hz), fatigue attenuation occurs. The detailed simulation process is illustrated in Fig.S3. Although the distance-firing rate-dependent curve shows a similar sigmoid trend, it becomes flatter than the initial state, indicating the sensitivity decreases in the fatigue state. For example, the firing rate in response to a target with distance of 5 cm decreases from 120Hz to 90 Hz in the initial and fatigue states, respectively.

Our visual system can focus on distant objects through the adjustment of the ciliary muscle, which enables us to acquire depth perception for precise recognition and learning. Such neuromuscular adjustment is processed and computed based on the slight difference of spatial information (i. e., parallax) perceived by each eye. Figure 4a shows such a process. The firing rates of the triggered spikes ( $f_1, f_2$ ) from the two optical neuronal pathways are dependent on the distance ( $L_1, L_2$ ) between an object to the left and right eyes, respectively. The interpupillary distance and the distance between the eye plane and object plane are  $d$  and  $z$ , respectively. The visual angle  $\theta$  is correlated to the relative orientation of the target sight. The visual system couldn't achieve a clear

image of the object due to the lack of depth information, leading to poor recognition. By comparing the firing rates between the two eyes, the ciliary muscle can be commanded to produce a clear vision. Two sVNs were used to obtain the firing rate differences for depth perception. The  $d$  and  $z$  are set to 20 and 100 mm ( $z=5d$ ), respectively. The frequencies of the two sVNs are 110 spikes/s and 40 spikes/s when  $L_1$  and  $L_2$  are 70 mm and 150 mm, respectively. Furthermore, the polar coordinate in Fig.4c shows the relationship between angle  $\theta$  and the firing rate differences between the left and right eyes. When the angle  $\theta$  is  $\arctan(2z/d) = 84.3^\circ$ , the firing rate difference is zero. The red and blue lines show the initial state and fatigue state, respectively. The difference curve is axially symmetric with respect to the angle of  $84.3^\circ$  in polar coordinates, which corresponds to the symmetrical equivalent relationship between the left and the right eyes. When the value of  $\theta$  gradually increases from  $40^\circ$  to  $140^\circ$  (or  $220^\circ$  to  $320^\circ$ ), the firing rate difference decreases to the minimum at  $84.3^\circ$ , and then gradually increases.

As the biological visual system can take advantage of the firing rate difference from the binocular vision, this can be exploited to estimate the distance of the object and focus on it to show a sharp image (Fig. S4). Without this process, blurred images will be formed in the eyes. This is because, when the object refracted by the crystalline lens is exactly at the plane of the retina, it appears as the clearest focused image. When the lens becomes a little thicker, the image projected on the retina is slightly out of focus, making it less clear after refraction. If the lens continues to thicken, severe defocusing will occur, and the image on the retina becomes harder to recognize. Nine groups of

face images are chosen from the Yale Face Database and down-sized to 784 pixels in  $28 \times 28$  size, with each group comprising 585 images. Fig. 4d shows examples of label 1-3 with different clarities. The image is in grey scale where each pixel value ranges from 0 to 255 with a larger value corresponding to a darker pixel. At the first stage, the visual system just receives the distance vision without proper focusing. The face images formed in the retina would be blurry as shown in the first column. At the second stage, the visual system adjusts the crystalline lens to obtain a clearer vision as shown in the second column. At last, the eye can focus on the object by comparing the firing rate differences as shown in the third column. The image datasets comprising of the Yale faces with three clarities ( $r=5, 1, 0$ , respectively), were built for training the network. The value of  $r$  represents the ratio value between the actual focal plane of the lens and the retinal plane, which is used for simulating the defocused images. A three-layer perceptron, with 784, 28, and 9 nodes in each layer, is adopted by the neural network for the recognition task using the back-propagation algorithm. Each synaptic connection in the network is simulated by extracting the parameter from an IGZO-based floating-gate synaptic transistors with 5-bits precision (Fig. S7)<sup>42</sup>. The neural network is trained to distinguish one person's face from others. The recognition consists of two phases: training and testing. The learning rate is 5% in the training procedure. The training phase includes two sub-procedures: forward-pass and weight update. During the training process, the weights are optimized by minimizing the difference between the predicted classification and the ground truth. During the testing process, the network was fed up with images of different clarity to identify the category of corresponding

faces, and the recognition accuracy is thus obtained. Fig.4e shows the relationship between the recognition accuracy and training epoch in three different focusing states. As the training epoch increases, the recognition accuracy is improved. For clearly focused images, the recognition accuracy is able to reach 90% after 20 training epochs. In contrast, for the severely defocused control group, the recognition accuracy can only reach 68% under the same situation. This simulation successfully demonstrates that depth perception has a great positive impact on the recognition accuracy of an artificial visual neural network. The function of binocular positioning based on the spiking visual neuron proposed in this work has a great potential for future neuromorphic computing, image acquisition, and recognition.

## **Discussion**

In conclusion, a memristor-based spiking visual neuron was built with bio-realistic properties capable of encoding light signals into spike trains. The firing rate of such a spiking visual neuron is distance-dependent and shifts with the stimulation time, analogous to the depth perception and eye fatigue in biological visual systems. An artificial binocular visual system with two sVNs can generate firing rate difference due to the distance difference from the target point to the two sVNs, which could be used to infer the depth of the vision. Such a binocular visual system can take advantage of the depth perception in achieving high accuracy pattern recognition. Face visions with three defocusing levels corresponding to different depth perception stages were simulated. Assisted by such an artificial binocular visual system, the accuracy of a

hardware neural network's image recognition increases from 68% to 90% after 20 training epochs. Such a spiking visual neuron is a promising candidate for hardware visual systems with high biological plausibility, energy efficiency and perceptual ability, which might open up a new frontier for neuromorphic engineering and bionic robotics.

## **Materials and Methods**

The threshold switching memristor was fabricated on the silicon substrate. Firstly, ITO (Indium Tin Oxide) bottom electrode was deposited on silicon with a thickness of 300 nm by radio frequency magnetron sputtering using an ITO target (90 wt%  $\text{In}_2\text{O}_3$  and 10 wt%  $\text{SnO}_2$  mass ratio) at room temperature. Secondly, the  $\text{TaO}_x$  buffer layer was deposited by magnetron sputtering  $\text{Ta}_2\text{O}_5$  target (100 wt%  $\text{Ta}_2\text{O}_5$ ) utilizing an Ar:  $\text{O}_2$  ratio of 30:2 and at room temperature with a radio-frequency power of 100 W. Finally, Ag top electrodes (80  $\mu\text{m}$ ×80  $\mu\text{m}$ ) were deposited on the  $\text{TaO}_x$  buffer layer by thermal evaporation with a nickel shadow mask. In the electrical neuron function measurement, a Tektronix AFG3102 pulse generator acts as the input source. In the optical measurement, the wavelength of laser is 532 nm. A Tektronix DPO3032 oscilloscope was used to measure the input/output voltages. The electrical characteristics of the TSM device were tested by Fs-Pro PX500. The exact stoichiometric ratios of the  $\text{TaO}_x$  buffer layer were confirmed by X-ray photoelectron spectroscopy (PHI5000 Versa Probe).



## References

1. Abbott, L. F. and Regehr, W. G. Synaptic computation. *Nature* **431**, 796-803 (2004).
2. Mead, C. How we created neuromorphic engineering. *Nat. Electron.* **3**, 434-435 (2020).
3. Liu, Z. *et al.* Multichannel parallel processing of neural signals in memristor arrays. *Sci. Adv.* **6**, abc4797 (2020).
4. Akbari, M. K. and Zhuiykov, S. A bioinspired optoelectronically engineered artificial neurorobotics device with sensorimotor functionalities. *Nat. Commun.* **10**, 3873 (2019).
5. Zhang, C. *et al.* Oxide synaptic transistors coupled with triboelectric nanogenerators for bio-inspired tactile sensing application. *IEEE Electron Device Lett.* **41**, 617-620 (2020).
6. Wan, C. *et al.* An artificial sensory neuron with tactile perceptual learning. *Adv. Mater.* **30**, 26258-26266 (2018).
7. Yu, J. *et al.* Bioinspired mechano-photonic artificial synapse based on graphene/MoS<sub>2</sub> heterostructure. *Sci. Adv.* **7**, abd9117 (2021).
8. Ghosh-Dastidar, S. and Adeli, H. Spiking neural networks. *Int. J. Neural Syst.* **19**, 295-308 (2009).
9. Ponulak, F. and Kasinski, A. Introduction to spiking neural networks: Information processing, learning and applications. *Acta. Neurobiol. Exp.* **71**, 409-433 (2011).
10. Wu, Q. *et al.* Spike encoding with optic sensory neurons enable a pulse coupled neural network for ultraviolet image segmentation. *Nano Lett.* **20**, 8015-8023 (2020).
11. Li, Z. A new framework for understanding vision from the perspective of the primary visual cortex. *Curr. Opin. Neurobiol.* **58**, 1-10 (2019).
12. Radhakrishnan, S. S., Sebastian, A., Oberoi, A., Das, S., and Das, S. A biomimetic neural encoder for spiking neural network. *Nat. Commun.* **12**, 2143 (2021).
13. Tee, B. C. K. *et al.* A skin-inspired organic digital mechanoreceptor. *Science* **350**, 313-316 (2015).
14. Kim, Y. *et al.* A bioinspired flexible organic artificial afferent nerve. *Science* **360**, 998-1003 (2018).
15. Hua, Q. *et al.* Low-voltage oscillatory neurons for memristor-based neuromorphic systems. *Glob. Chall.* **3**, 1900015 (2019).
16. Yin, J. *et al.* Self-modulating interfacial cation migration induced threshold switching in bilayer oxide memristive device. *J. Phys. Chem. C* **123**, 878-885 (2019).

17. Huang, H.-M. *et al.* Implementation of dropout neuronal units based on stochastic memristive devices in neural networks with high classification accuracy. *Adv. Sci.* **7**, 2001842 (2020).
18. Sun, Y. *et al.* Performance-enhancing selector via symmetrical multilayer design. *Adv. Funct. Mater.* **29**, 1808376 (2019).
19. Zhao, Y. *et al.* A compact model for drift and diffusion memristor applied in neuron circuits design. *IEEE T. Electron. Dev.* **65**, 4290-4296 (2018).
20. Zhang, X. *et al.* An artificial neuron based on a threshold switching memristor. *IEEE Electron Device Lett.* **39**, 308-311 (2018).
21. Tuma, T., Pantazi, A., Le Gallo, M., Sebastian, A., and Eleftheriou, E. Stochastic phase-change neurons. *Nat. Nanotechnol.* **11**, 693-699 (2016).
22. Delaney, M. *et al.* Nonvolatile programmable silicon photonics using an ultralow-loss Sb<sub>2</sub>Se<sub>3</sub> phase change material. *Sci. Adv.* **7**, eabg3500 (2021).
23. Gao, L., Chen, P.-Y., and Yu, S. NbOx based oscillation neuron for neuromorphic computing. *Appl. Phys. Lett.* **111**, 103503 (2017).
24. Yi, W., Tsang, K. K., Lam, S. K., Bai, X., Crowell, J. A., and Flores, E. A. Biological plausibility and stochasticity in scalable VO<sub>2</sub> active memristor neurons. *Nat. Commun.* **9**, 4661 (2018).
25. Pickett, M. D., Medeiros-Ribeiro, G., and Williams, R. S. A scalable neuristor built with Mott memristors. *Nat. Mater.* **12**, 114-117 (2013).
26. Choi, S., Yang, J., and Wang, G. Emerging memristive artificial synapses and neurons for energy-efficient neuromorphic computing. *Adv. Mater.* **32**, 2004659 (2020).
27. Wang, Z. *et al.* Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* **1**, 137-145 (2018).
28. Burkitt, A. N. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biol. Cybern.* **95**, 1-19 (2006).
29. Kim, T. *et al.* An artificial neuron using a bipolar electrochemical metallization switch and its enhanced spiking properties through filament confinement. *Adv. Electron. Mater.* **7**, 2000410 (2021).
30. Yang, Y. *et al.* Electrochemical dynamics of nanoscale metallic inclusions in dielectrics. *Nat. Commun.* **5**, 4232 (2014).

31. Wang, Z., Feng, Z., and Wei, X. Axonal stimulations with a higher frequency generate more randomness in neuronal firing rather than increase firing rates in rat hippocampus. *Front. Neurosci.* **12**, 783 (2018).
32. Silver, R. A. Neuronal arithmetic. *Nat Rev Neurosci* **11**, 474-489 (2010).
33. Merolla, P. A. *et al.* A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**, 668-673 (2014).
34. Sharma, A. A., Li, Y., Skowronski, M., Bain, J. A., and Weldon, J. A. High-frequency TaO<sub>x</sub>-based compact oscillators. *IEEE Electron Device Lett.* **62**, 3857-3862 (2015).
35. Midya, R. *et al.* Artificial neural network (ANN) to spiking neural network (SNN) converters based on diffusive memristors. *Adv. Electron. Mater.* **5**, 1900060 (2019).
36. Han, J.-K., Seo, M., Yu, J.-M., Suh, Y.-J., and Choi, Y.-K. A single transistor neuron with independently accessed double-gate for excitatory-inhibitory function and tunable firing threshold voltage. *IEEE Electron Device Lett.* **41**, 1157-1160 (2020).
37. Hua, Q., Jiang, C., and Hu, W., Ag/HfO<sub>2</sub>-based threshold switching memristor as an oscillatory neuron, in *5th IEEE Electron Devices Technology and Manufacturing Conference (EDTM)*, (2021).
38. Lim, H. *et al.* Relaxation oscillator-realized artificial electronic neurons, their responses, and noise. *Nanoscale* **8**, 9629-9640 (2016).
39. Zhang, J. *et al.* Retina-inspired organic heterojunction-based optoelectronic synapses for artificial visual systems. *Research* **2021**, 7131895 (2021).
40. Bao, L. *et al.* Artificial shape perception retina network based on tunable memristive neurons. *Sci. Rep.* **8**, 13727 (2018).
41. Chen, S., Lou, Z., Chen, D., and Shen, G. An artificial flexible visual memory system based on an UV-motivated memristor. *Adv. Mater.* **30**, 1705400 (2018).
42. He, Y. *et al.* IGZO-based floating-gate synaptic transistors for neuromorphic computing. *J. Phys. D. Appl. Phys.* **53**, 215106 (2020).

**Acknowledgments:** We appreciate X. J. Liu for discussion and comments on organizational structure of the article. This work is supported by National key R&D Program of China (Grant No. 2019YFB2205400) and National Natural Science Foundation of China (Grant No. 62074075 and No. 62174082).

**Author contributions:**

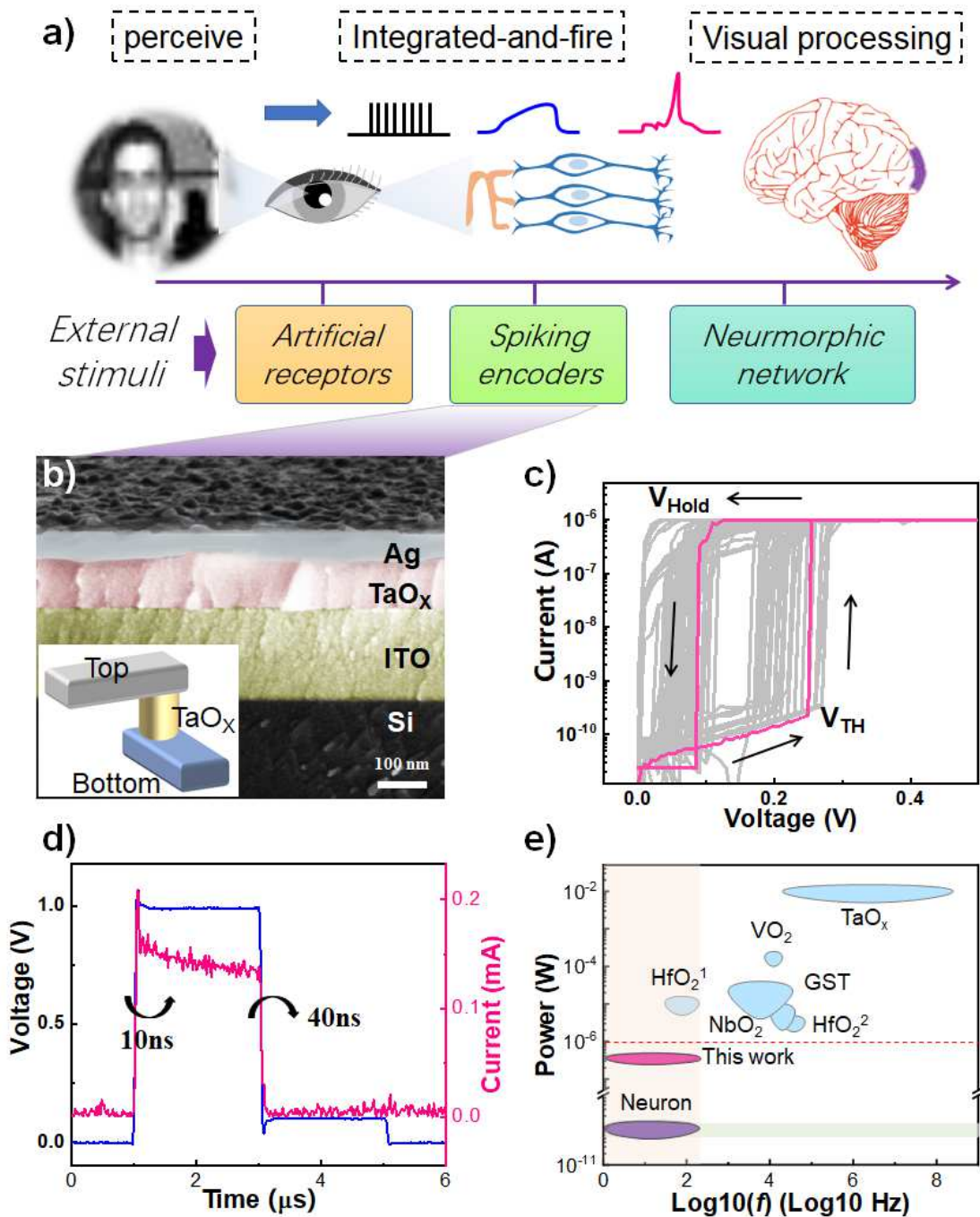
C. J. W. and Q. W. conceived and supervised the project. C. S. C, C. J. W, Q. W, Y. S designed the device and experiment. C. S. C., Y. L. H., H. W. M., L. Z., X. J. W, Y. Z, Y. X. Z, and C. J. W. conducted most of the experiments, including the device fabrication and characterization. Y. L. H. fabricated and tested the floating-gate synaptic transistors and extracted the parameters of such devices for simulation. C. S. C, Y. L. H. and C. J. W. designed and conducted the simulation for face recognition tasks. C. S. C., C. J. W., and Q. W. wrote the manuscript.

**Competing interests:**

The authors declare that they have no competing interests.

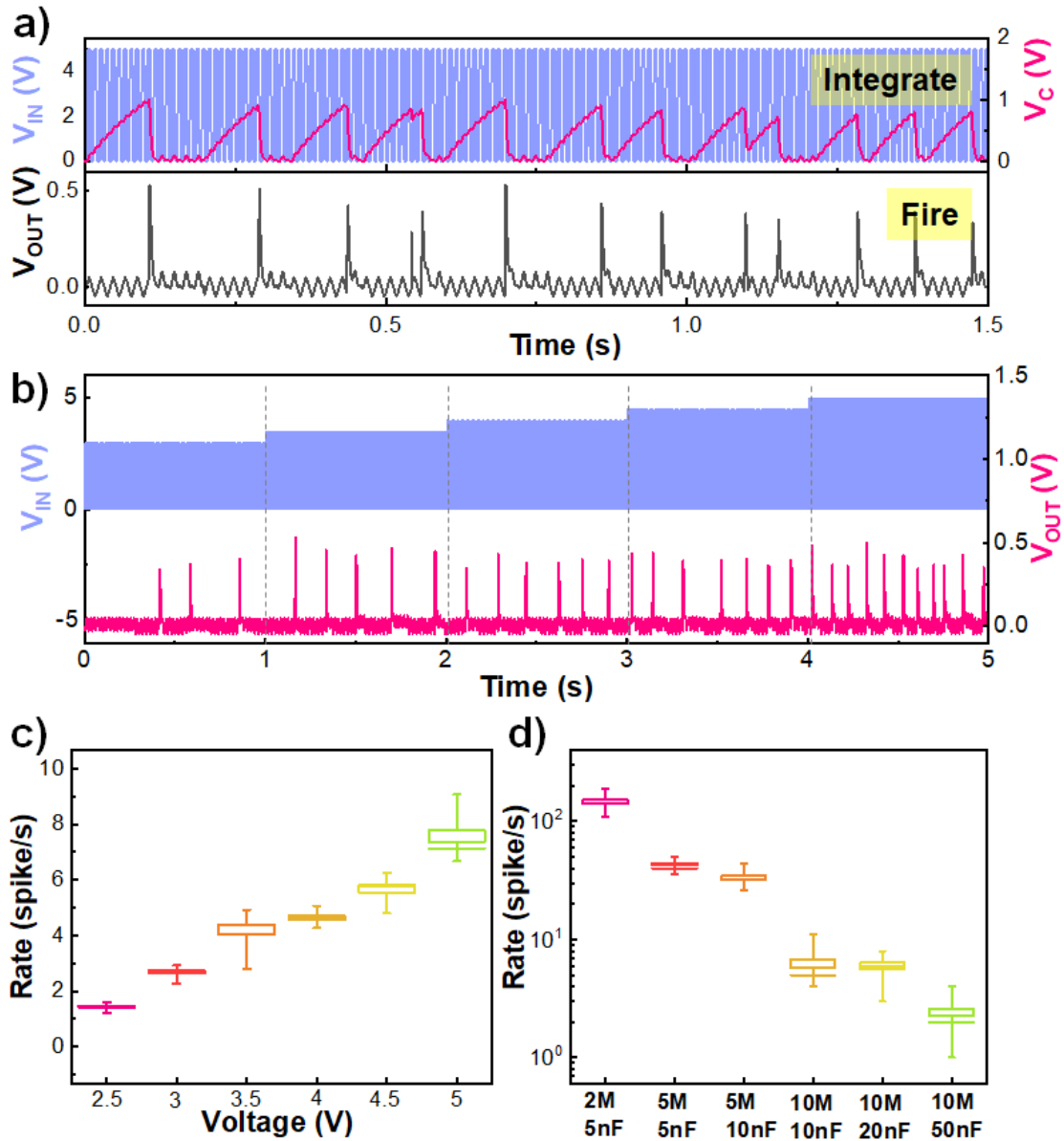
**Data and materials availability:**

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

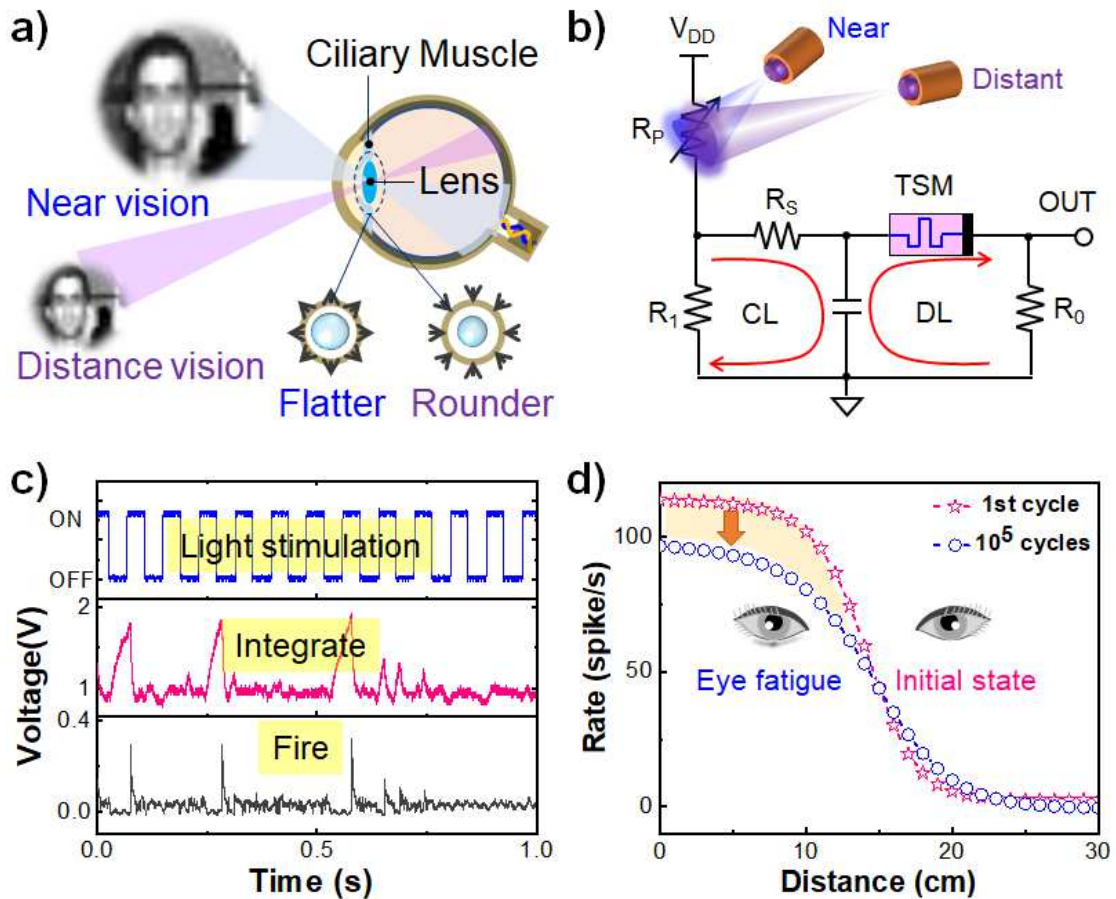


**Fig.1 A spiking visual neural network based on Ag/TaO<sub>x</sub>/ITO memristor in comparison with a biological one.** (a) Photoreceptors convert external stimuli into spikes which are encoded according to leaky integrate-and-fire model. The encoded information feeds the central nervous system for further processing. (b) SEM image of the cross-section of the memristor. Inset: a structural diagram of the two-terminal TaO<sub>x</sub> memristor. (c) Current-voltage curves of the device in 1000 sweep loops. (d) The I-V characteristics of the TSM device and the corresponding switching speed of the resistance. (e) The

summary of some recent memristor-based spiking neurons vs. biological neuron in terms of firing rate and power consumption per spike.

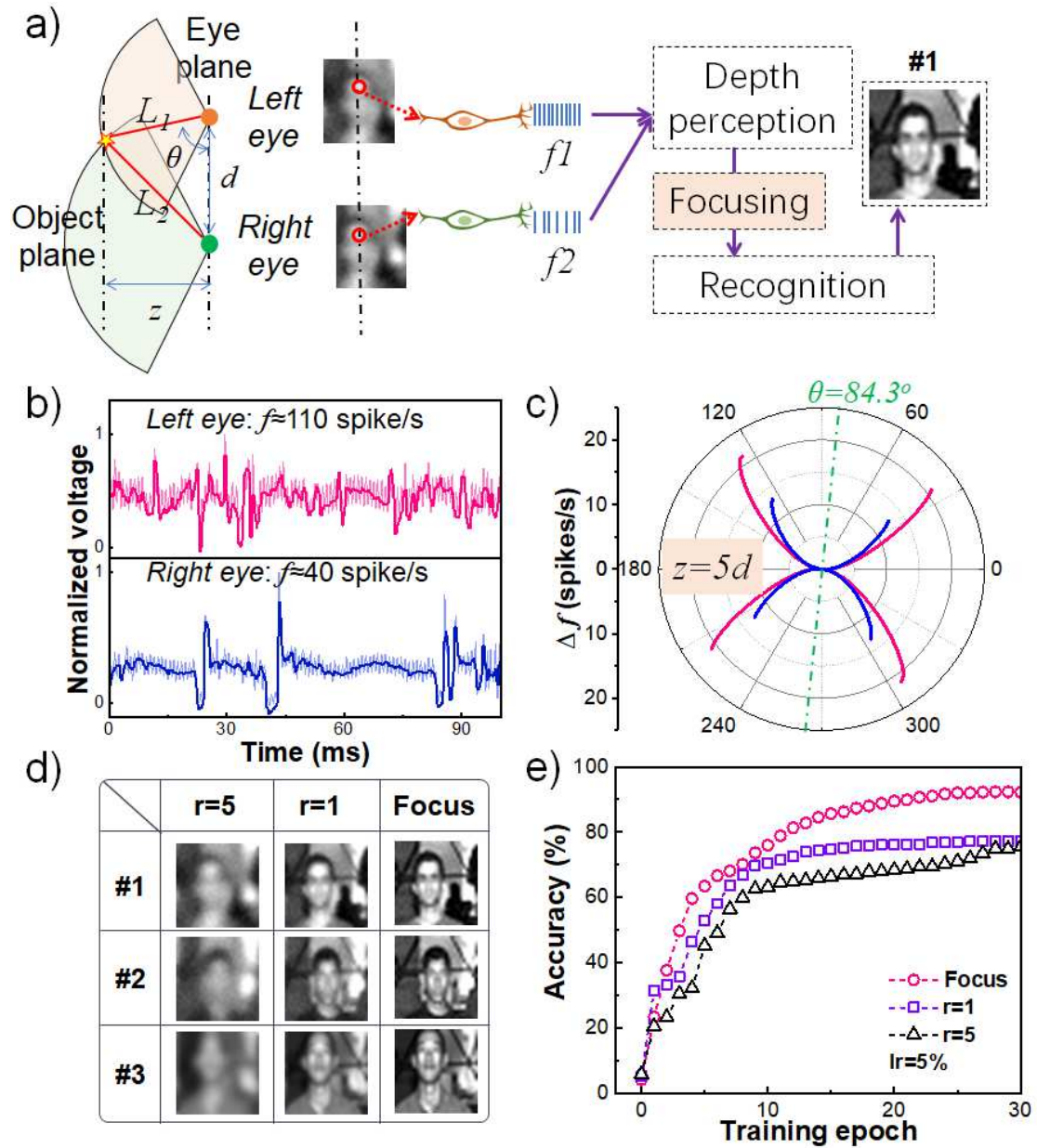


**Fig.2** The integrate and fire dynamics of the TS memristor-based spiking encoder. (a) The voltage across the capacitance ( $V_C$ : pink curve) and output fire (black curve) under a train of input voltage pluses ( $V_{IN}$ : blue curve). (b) The output spikes of the spiking encoder in response to different input voltage amplitudes. (c) Statistical voltage and spiking rate relationship of the neuron. Increasing voltage amplitude leads to increasing spiking rate. (d) A larger RC time constant decreases the output spiking rate.



**Fig.3 A spiking visual neuron based on the TS memristor.** (A) Schematic illustration of the eye adjusted to adapt to near vision and distance vision through the ciliary muscle. (B) The circuit scheme for the spiking visual neuron. (C) The proposed spiking visual neuron encodes the light pulse based on the leaky integrate-and-fire model. (D) The relationship between firing rate and distance for 1<sup>st</sup> (pink) and 10<sup>5</sup> (blue) cycles, corresponding to the fresh and fatigue states of eye, respectively.





**Fig.4 Emulation of binocular vision for high accuracy recognition based on the TS memristor-based spiking visual neuron.** (a) Schematic image of binocular positioning process in the human brain. The visual system can take advantage of the parallax between the two eyes for image focusing. (b) The firing rate of the left eye and the right eye at a certain distance, respectively. (c) The firing rate difference between left and right eyes as a function of visual angle. (d) The three columns of face images are corresponded to severely defocus ( $r=5$ ), slightly defocus ( $r=1$ ), and clearly focus face images from Yale Face Database. (e) When the learning rate is 5%, evolution classification accuracy as a function of

training epochs for clearly focus, slightly defocus ( $r=1$ ), and severely defocus ( $r=5$ ) face image datasets, respectively.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.pdf](#)