Brain Inspired Color Feature Selection Chip

Sheng Xie, Lihua Wang, Xurui Mao, Jia Cong, Yuhao Zhao, and Hongda Chen

***Abstract*—Inspired by the mechanism of visual attentional selection, a color feature selection unit consisting of photoreceptors and an attentional selection circuit (ASC) is presented. The detection and wavelength recognition of trichromatic light are accomplished by the photoreceptors with RGB filters. The ASC, fabricated in a standard CMOS technology, activates the competition among photoreceptors to select the principal color-component feature defined as the attentional focus. The number of focuses and resolution can be changed with a top-down stimulus. Electrical test results show that the ASC can select and interpret the most important color-component feature. In optical tests, the ASC chip-based unit can extract the principal color-component of single-color and dual-color light in real-time to complete color recognition. The percentage ratio between the intensity of stray light relative to the light with maximum intensity is up to 93.58%. And the maximum achievable recognition frame rate is 1 kHz, which is more than one order of magnitude higher than the frequency of the biological vision systems. Besides, the unit's outputs can be transformed into binary codes as labels for color recognition, and the color recognition process does not require the participation of ADC, CPU, and memory, consequently avoiding the von Neumann bottleneck.**

***Index Terms*—Attentional selection, color feature selection, color recognition, standard CMOS technology.**

I. Introduction

VISION is the most dominant and important sense of all primates including humans, with 80% or more of all sensory information being perceived through sigh [1]. Color is one of the most potent and apparent elements of visual attributes [2]. Therefore, processing the color information perceived by three types of cones in the retina is significant for human beings to sense changes in their surroundings [3]. Mimicking the human eye's spectral response, the artificial image sensor system decodes visible light into three current values through photodetectors covered with wavelength selective filters. Data acquisition is completed with analog-to-digital conversion circuits, and the collected data is amplified, compared, and integrated to realize color recognition, which is completely dependent on CPU and memory [4]. These systems are based on the traditional von Neumann architecture, which consumes a lot of energy in data transmission and processing and may not be able to meet real-time recognition requirements [5]. In particular, these color recognition systems have to deal with the overloaded information at the cost of system complexity or power consumption.

In the human visual system, a tremendous amount of visual information falls on the retina at any given moment, far more than the brain's capacity to process. Surprisingly, the brain can deal successfully and meaningfully with the information in a low-power, efficient and rapid way [6]. This process owes much to the mechanism of attentional selection [7]. Visual attentional selection is the most powerful strategy used by the human visual system, which can select, modulate, and sustain focus on the most behaviorally-relevant information while discarding insignificant ones [8], [9]. Most recently, the mechanism of visual attentional selection has been introduced into visual tasks such as machine translation [10], image recognition [11], and semantic scene segmentation [12]. However, most of these applications are based on very large scale integration (VLSI) circuits, which exist bottlenecks in terms of speed and power consumption [13]. With the boom of deep learning, the artificial neural network is being used in the visual system to simulate attentional selection [14], but in practice it may lead to an inaccurate model due to insufficient data and training time [15]. Therefore, it is of great significance for promoting the development of artificial vision to design a hardware implementation of the attentional selection with simple structure, real-time recognition, low power consumption, and miniaturization.

The color recognition process of the human brain is shown in Fig. 1. The photoreceptor cells located in the retina transform visible light into nerve impulses. The neural networks improve the signal selectivity and contrast by weakening the response of nerve impulses near significant targets to select the significantly associated features, completing the color recognition [16]. Attentional selection is attributed to the regulation of neural activity in the primary visual cortex, including the first visual region (V1) and the frontal lobe eye area [17]. Inspired by the visual attentional selection, a color feature selection scheme was proposed, which is also shown in Fig. 1. The color feature

Manuscript received 17 July 2022; revised 14 August 2022; accepted 6 September 2022. Date of publication 12 September 2022; date of current version 26 September 2022. This work was supported in part by the National Natural Science Foundation of China under Grants 61874104 and 11673019, in part by Jiangsu Provincial Key Research and Development Program under Grant BE2021008-3, and in part by Enterprise Science and Technology Commissioner Project under Grant 21YDTPJC00780. *(Corresponding authors: Xurui Mao; Jia Cong.)*

Sheng Xie is with the Tianjin Key Laboratory of Imaging and Sensing, School of Microelectronics, Tianjin University, Tianjin 300072, China (e-mail: xie_sheng06@tju.edu.cn).

Lihua Wang is with the Tianjin Key Laboratory of Imaging and Sensing, School of Microelectronics, Tianjin University, Tianjin 300072, China, and also with the OmniVision North China IC Ltd., Tianjin 300040, China (e-mail: wanglhua_1995@tju.edu.cn).

Xurui Mao and Hongda Chen are with the State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China (e-mail: maoxurui@semi.ac.cn; hdchen@semi.ac.cn).

Jia Cong is with the School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China (e-mail: congj2015@tju.edu.cn).

Yuhao Zhao is with the Tianjin Key Laboratory of Imaging and Sensing, School of Microelectronics, Tianjin University, Tianjin 300072, China, and also with the Jiangsu Kexu Microelectronics Company, Ltd., Suzhou 215100, China (e-mail: zhaoyuhao@tju.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2022.3205585

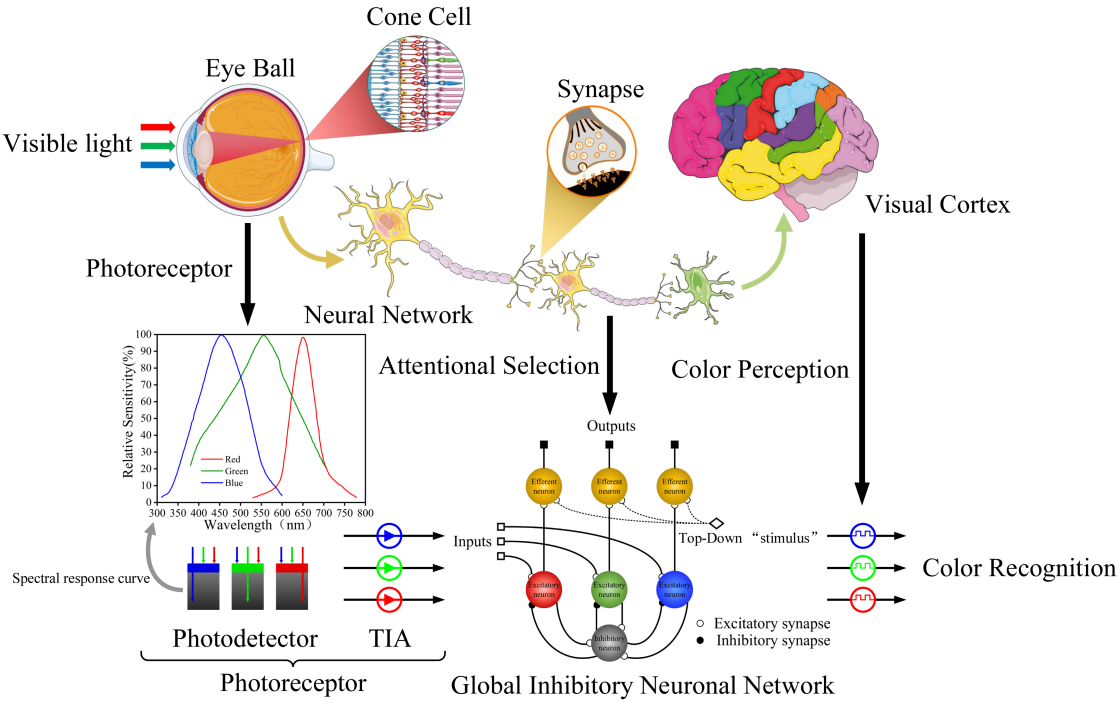

Fig. 1. The functional diagram of the color feature selection unit inspired by visual attentional selection. The photoreceptors are like human cone cells located in the retina that transform visible light into nerve impulses. As one of the most classic global inhibitory neuronal networks, the interaction between excitatory neurons and inhibitory neurons in the ASC can simulate the neural activities related to visual attentional selection.

selection unit can select the most significant information from visible light, and all decisions and information processing are carried out at a low power level. Thus, the unit overcomes the limitations of the Von Neumann architecture [18]. The unit can be devided into two main functional blocks. The first block is photoreceptors, which contain photosensing devices and trans-impedance amplifiers (TIAs), generating analog signals whose amplitude is proportional to the light intensity. The second block is the signal processing module, where the attentional selection circuit (ASC) emulates the visual attentional selection and shifts its attention to the most behaviorally-relevant information in real-time. The feature selection unit has the initiative and selectivity [19] like humans by adaptive adjusting the focus of attention to select the principal color-component feature, and it is a miniaturized hardware implementation of the visual attentional selection, which increases performance over existing computer vision feature extraction approaches.

II. Attentional Selection and Circuit Implementation

In all primates, attention is a critical factor in the regulation of behavior and neurocognitive functioning [20]. The attentional selection works throughout the brain and participates in almost every stage from sensory processing to decision-making and cognition. The research showed that there are two types of attentional selection [21]. The first one driven by environmental stimuli is usually named as "bottom-up" attention. The other is the voluntary attention to external and internal stimuli controlled by the brain, called "top-down" attention. Attentional selection promotes the expression of the most important information in the human brain by improving recognition speed and contrast, enabling organisms to capture prominent targets [22]. The ASC, the hardware implementation of the attention selection mechanism, selects the most significant signal by activating the competition among inputs and actively shifts attention to the position with the salient feature.

The bottom-up attentional shift: Since the silicon-based MOS transistors operating in the subthreshold state have the work function to model biological processes [23], they can achieve a physical process similar to cell membrane current regulation of biological neurons with low power consumption [24]. Based on these fundamental properties, a three-branch attentional selection circuit for color recognition was proposed, which is shown in Fig. 2(a). The proposed circuit responds to the branch with the highest input voltage while suppressing the responses of the other branches, concentrating processing power on the most important information. As a result, its electrical behavior resembles the neurological behavior of specific neural networks in the visual cortex. Fig. 2(b) depicts the I_D-V_{GS} relationship of input transistors M_B, M_G and M_R as excitatory neurons. When the transistors operate in the subthreshold region, the I_D increase exponentially with increasing V_{GS} (gate-source voltage), thus making the circuit extremely sensitive to the variation of input voltage. The common node voltage V_C changes with the input voltage as shown in Fig. 2(c). Similar to the mechanism of inhibitory interneurons in the human brain, the transistor $M_{inhibition}$, as an intermediate inhibitory neuron, mediates global inhibition through the common node. When the input

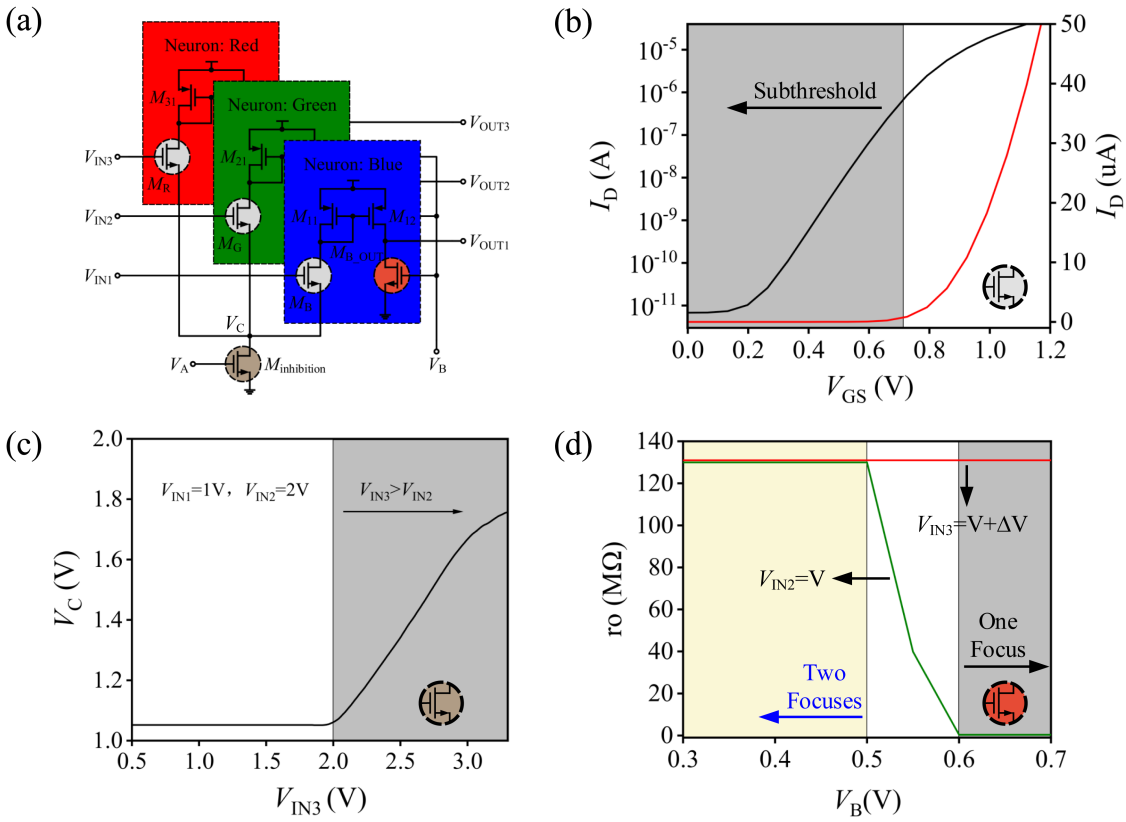

Fig. 2. (a) The attentional selection circuit; (b) The I_D-V_{GS} characteristic of "excitatory neuron". As excitatory neurons, transistors M_B, M_G, and M_R can sense the difference in input voltage; (c) The inhibition curve of "inhibitory neuron". The transistor $M_{inhibition}$ provides global inhibition for the whole circuit as an intermediate inhibition neuron. The electrical behavior of $M_{inhibition}$ is close to the neuroelectrical behavior of middle inhibitory neurons in the human brain [26]; (d) The variation of internal resistance of efferent neurons with top-down stimulus. As efferent neurons, M_{B_OUT}, M_{G_OUT}, and M_{R_OUT} transform the current into spike voltage signals, which indicate this branch becomes the attentional focus. The number of attentional focuses and the voltage resolution can be changed by applying a top-down stimulus, which adjusts the internal resistance of efferent transistors.

voltage of a branch is stronger than that of the other two, this branch gains most of the current from $M_{inhibition}$. The source voltage of the input transistor follows the change of the gate voltage, increasing the source voltage, which is V_C. Since all input transistors share a common source, the responses of other branches are suppressed until they are turned off. Finally, the currents flow from the current mirror formed by M_{i1} and M_{i2} (i = 1, 2, 3) to the efferent transistors M_{B_OUT}, M_{G_OUT}, and M_{R_OUT} as efferent neurons and then are converted into spike voltage signals. Based on the global inhibitory feedback mechanism [25], the proposed circuit quickly switch its attention to the branch having the strongest input and eventually tends to a stable state with only one attentional focus.

The top-down attentional distraction: To be more sensitive and accurate in perceiving changes in the external environment, the brain sometimes actively controls the visual system to distract attention toward the information outside the attentional focus. To achieve the attentional distraction, a top-down stimulus is applied to the efferent transistors. Supposing $V_{IN1} = V - \Delta V$, $V_{IN2} = V$, $V_{IN3} = V + \Delta V$, when no significant stimulus is applied, only the V_{IN3} branch gets the full attention of the circuit. The internal resistance can be iteratively adjusted by changing V_B. The variation of internal resistance of efferent transistors M_{G_OUT} and M_{R_OUT} with V_B is shown in Fig. 2(d). As V_B decreases, the internal resistance, as a reward, gradually increases, and the attention of the V_{IN2} branch is therefore strengthened. When the reward is strong enough, the V_{IN2} branch transforms into an additional focus of attention. As a reward-driven attentional distraction, the reward can also adjust the degree of attentional distraction. When compared with the traditional winner-take-all circuit [27], except for the maximum signal selection, the ASC can change the number of attentional focuses and resolution by applying a top-down stimulus. When different stimuli are applied to the various branches, the circuit's attention only concentrates on the information outside the attentional focus to meet higher-level recognition requirements.

The DC simulation results of the ASC are shown in Fig. 3(a). Where $V_{IN1} = 1$ V, $V_{IN2} = 2$ V, and V_{IN3} gradually increases from 0 V to 3.3 V. When V_{IN3} is far less than 2 V, the circuit only pays attention to the V_{IN2} branch and thus, $V_{OUT2} = 3.3$ V. Similarly, when V_{IN3} exceeds the other two values, the attention switches to its branch. As V_{IN3} approaches V_{IN2}, the ASC adjusts its attention in real-time according to the inputs. The number of attentional focuses and the resolution of the ASC versus V_B are shown in Fig. 3(b). The resolution of ASC with double attentional focuses is defined as the maximum difference between the input voltages of the corresponding branches of focus. When the input difference is within the resolution, the

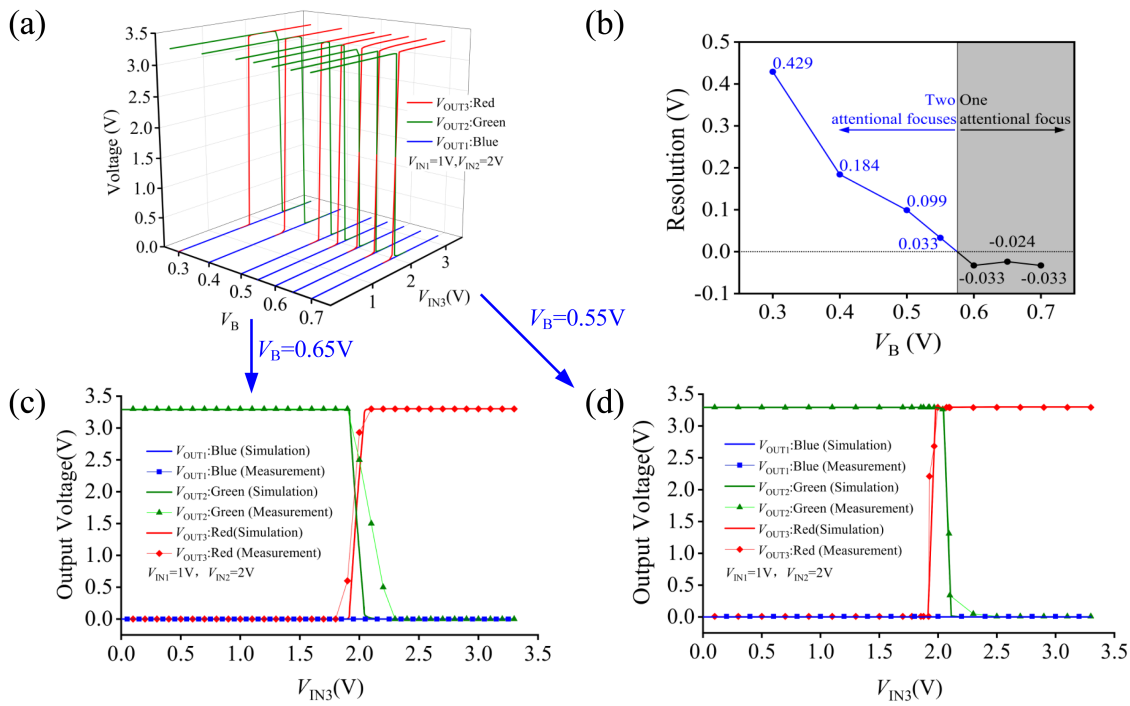

Fig. 3. (a) DC response of the ASC (simulation). The DC curves of the ASC with different V_B voltages; (b) The voltage resolution of the ASC. The evolution of the resolution is shown in Fig. 3(b). The number of attentional focuses and the level of attentional distraction can be adjusted by changing V_B; (c) DC response of the ASC when $V_B = 0.65$ V; (d) DC response of the ASC when $V_B = 0.55$ V. Fig. 3(c) and (d) are the comparisons between experiment and simulation results. It can be seen that ASC can achieve attentional concentration and distraction, and excellent agreement between experiment and simulation results was realized.

circuit can distract its attention to the additional focus. The resolution of ASC with a single attentional focus is the minimum voltage difference that the circuit can recognize when the inhibition weight of $M_{inhibition}$ to each branch and the ASC's attention is constantly adjusted according to the input voltage variation. The circuit switches from the adjustment state to the steady state with only one focus once the voltage difference is outside the resolution. It can be seen that by changing V_B, the voltage resolution and the number of attentional focuses can be adjusted. When V_B is 0.65 V and 0.55 V, the circuit has the optimal voltage resolutions of 24 mV and 33 mV in the corresponding state, respectively.

As a whole, The ASC can simulate the dynamic process of attentional selection, and its simple structure paves the way for low power, efficient hardware designs.

III. Experiments and Test Results

A. Electrical System Measurements: the ASC Chip

The ASC was fabricated in the Taiwan Semiconductor Manufacturing Company (TSMC) 0.18 μm standard complementary metal-oxide-semiconductor (CMOS) technology, and the supply voltage is 3.3 V. To make the circuit more stable and easily accommodate to different conditions, V_A can be adjusted externally. During the test process, the value of V_A is about 0.5 V. The micrograph of the ASC is shown in Fig. 4(a). The ASC chip, including pads, occupies an area of 850 μm $\times$ 800 μm, whereas the core area of the ASC chip is only 0.0285 mm^2 (150 μm $\times$ 190 μm). All the pads of the chip are marked in the figure, and the ASC chip was mounted on a test printed circuit board (PCB) using wires bonding.

The test condition of the ASC is consistent with the simulation. The test results at V_B of 0.65 V are shown in Fig. 3(c). The resolution with single attentional focus is about 80 mV. The test results show that the ASC can focus on the branch with the maximum input voltage in real-time. Note that there is a slight delay in the response of V_{OUT3} compared to the simulation results when V_{IN3} slightly exceeds V_{IN2}, which may be caused by process variations and layout mismatches [28]. The results at $V_B = 0.55$ V are shown in Fig. 3(d). The circuit can select double attentional focuses simultaneously when the inputs are close, and excellent agreement between experiments and simulations is achieved.

B. Optical System Measurements: The Color Feature Selection Unit

Based on the fabricated ASC chip, the proposed color feature selection unit was assembled on a specially designed PCB, as shown in Fig. 4(b). The photoreceptor is located in the blue frame, and the photo-sensing devices (photoreceptors) are in the white frame. It can be seen from in Fig. 1 that the central wavelengths of the three photoreceptors are 450 nm, 550 nm, and 650 nm, respectively. The TIA is located in the yellow frame, and its circuit diagram and frequency response is shown in Fig. 4(c). Operational amplifiers (OPA847, Texas Instruments) are used to

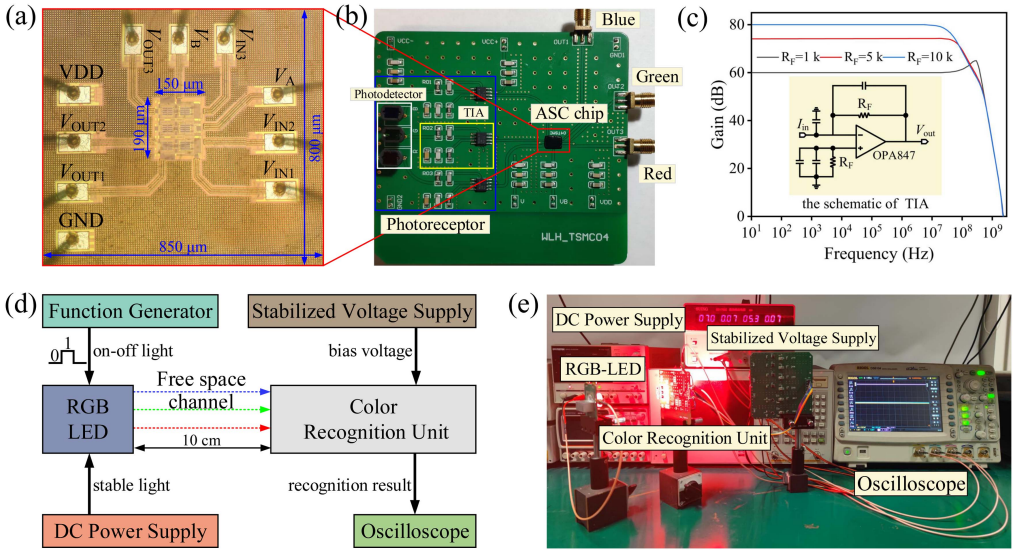

Fig. 4. (a) The micrograph of the ASC chip; (b) The test board of the color feature selection unit; (c) The circuit diagram and frequency response of the TIA; (d) The experimental setup to measure the optical characteristics of the color feature selection unit. The stabilized voltage supply ensures the proper functioning of the ASC; (e) The optical test link. The ratio of RGB colors is achieved by varying the electrical power injected into the LED.

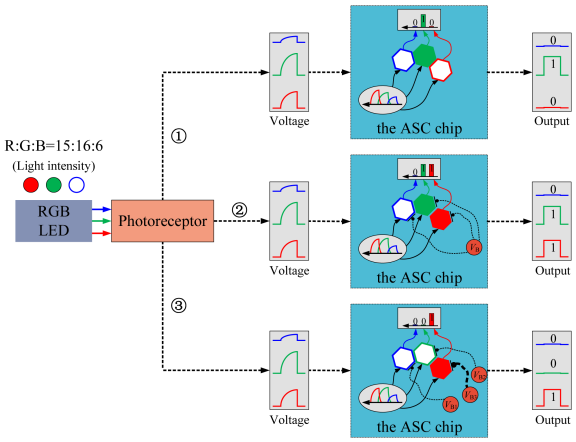

Fig. 5. The schematic diagram of the steady-state color feature selection. The upper bound on the number of attentional focuses is set by V_A by controlling the current flowing through the ASC, while V_B can determine whether the branch outside the attentional focus can become an additional focus by giving it enough "attention". In this paper, the ASC possesses multiple attentional focuses simultaneously. The recognition of single-color light and dual-color light can be achieved by changing the number of attentional focuses. In particular, the special attentional target can be set by adjusting V_B. For example, if the red light is set as the attentional target, the unit's attention can be guided by the target and shifts to the red light while ignoring the green light with the maximum intensity. When different V_B (V_{B1}, V_{B2}, V_{B3}) are applied to the branches, the gain of the red branch is greater than those of the blue and green branches, thus forcing the red branch to receive extra attention. In this case, even though the green light is stronger than the red light, the chip still focuses on the red light unless the green light is strong enough to offset the gain of V_{B3} to the red light.

form TIAs, and the gain of the TIA is adjusted by changing the feedback resistance R_F. To assure the color feature selection chip work properly and meet the input voltage requirements of ASC, whose input common mode range (ICMR) is about 3 V, the output of the TIAs ranges from 0V to 3V by varying the R_F. The experimental setup to measure the optical characteristics is illustrated in Fig. 4(d). An RGB-LED is used as the experimental light source (peak wavelength 455 nm, 550 nm and 635 nm), and its color and light intensity can be adjusted by changing LED's drive voltage. For example, a dual-color light can be obtained by applying DC voltage to the two control ports of the LED, and the periodical on-off light can be realized by using a square-wave signal as a drive voltage. The optical test link is shown in Fig. 4(e). The color feature selection unit is placed on the right side of the LED, and an optimal distance between LED and photoreceptors of about 10 cm is selected. The light intensity on the surface of photoreceptors is measured by the optical power meter (PM100D, Thorlabs), and the results of the color feature selection unit are displayed on the Digital oscilloscope (DS6104, Rigol).

Steady-state optical recognition measurements: Fig. 5 is the steady-state recognition diagram of the color feature selection unit. The intensities measured on the surface of photoreceptors are shown in Fig. 6(a), and the corresponding outputs are shown in Fig. 6(b). Firstly, the RGB light is set as the strongest incident light at test point 1-3, and V_B is fixed at 0.65 V. It can be seen that the unit can recognize the single-color light with maximum intensity and suppress the response of the branches that correspond to the stray light. Then, the dual-color light intensity ratio is set about 1:1 at test point 4-6 and V_B is adjusted to 0.55 V. Note that the unit senses the existence of dual-color light and distracts attention to their corresponding branches in real-time. The corresponding values of the output voltage (V_{OUT1}, V_{OUT2}, V_{OUT3}) can be transformed into the binary format as the color labels, and the obtained truth table is reported in Fig. 6(c). Since the main purpose of this design is to implement the selection of principal color-component features, so the special situation (R = 1, G = 1, and B = 1) is not considered. Moreover, the color feature selection unit can suppress the redundant signals and interference caused by stray light, thus reducing the negative

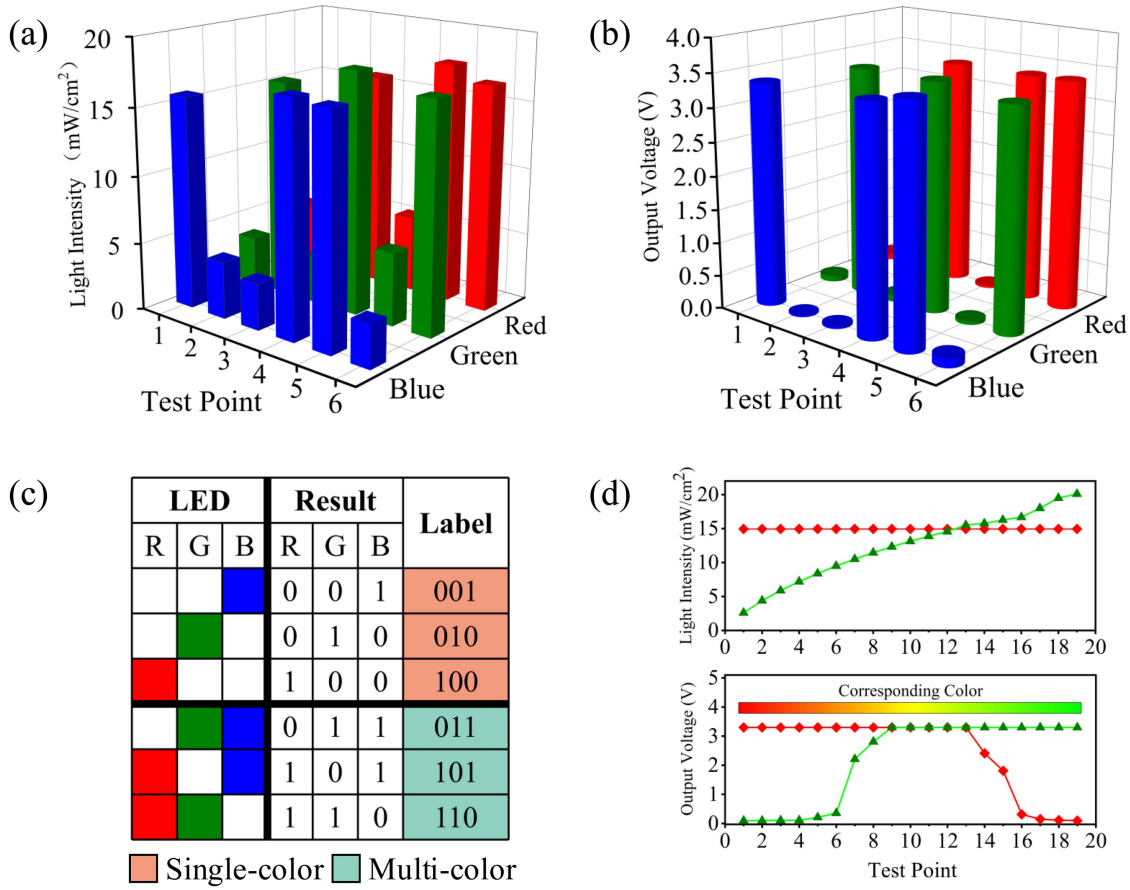

Fig. 6. (a) The light intensity measured on the photoreceptor surface; (b) The recognition results; (c) Measured truth table. The outputs of the unit are defined as digital labels for color recognition. The logic "1" is the attentional focus of the unit, indicating the branch has the maximum light intensity; (d) The light intensity of green and red lights and the corresponding recognition results.

impact on color feature selection caused by the spectral overlap of the photoreceptors [29] and the difficulties of subsequent signal processing.

To further analyze the shifting-attention of the color feature selection unit when the light intensity changes, the red light with fixed intensity and on-off green light are selected for further experiments. The blue light as the background light, whose intensity is relatively weak, would not influence the results. The light intensities and the corresponding recognition results are plotted in Fig. 6(d). By increasing the intensity of green light, the ratio of green-to-red light intensity gradually rises from 2:15 to 20:15. It can be seen that the color feature selection unit can adjust its attention to cope with the change of light intensity, and the recognition results consist with that of the ASC (Fig. 3(d)). At test point 16, the unit can recognize the percentage ratio between the intensity of stray light relative to the attentional focus is up to 93.58% (the output voltage of stray light is 10% lower than that of the strongest light), indicating the unit has excellent light intensity sensitivity and anti-interference ability. That is, the minimum recognition resolution is equivalent to 4-bit ADC. In particular, the sensitivity can be adjusted by V_B to satisfy different requirements.

Color pattern recognition measurements: An experiment was carried out to demonstrate color pattern recognition in image processing. As shown in Fig. 7(a), the input map is a 12×12 pixel array. Each pixel represents the LED luminous color and is proportional to the light intensity. The driving voltage of the LED is configured according to the input map, and the color recognition results are recorded in the recognition map. The recognition map with single attentional focus is shown in Fig. 7(b). The unit can recognize the single-color light with maximum intensity in each pixel. The results are in good qualitative agreement with the color information of the input map, and readers will see "18" if they focus on green and red, respectively. The GB and RB recognition maps with double attentional focus are shown in Fig. 7(c) and 7(d). The recognition unit distracts the attention to the red and green light in the pixels with approximate equivalent intensity and evokes the response of both branches simultaneously. It is obvious which one has higher light intensity when compared to the recognition map of a single attentional focus. Combining the GB and RB recognition maps of the double attentional focus, the RGBY recognition map with double attentional focus is demonstrated in Fig. 7(e). The yellow pixel represents that the attention is drawn to the red and green light. The unit can recognize the input map more accurately and comprehensively, and a pattern of "is" can be obtained due to the additional focus. In practical applications, the recognition sensitivity can be adjusted to meet

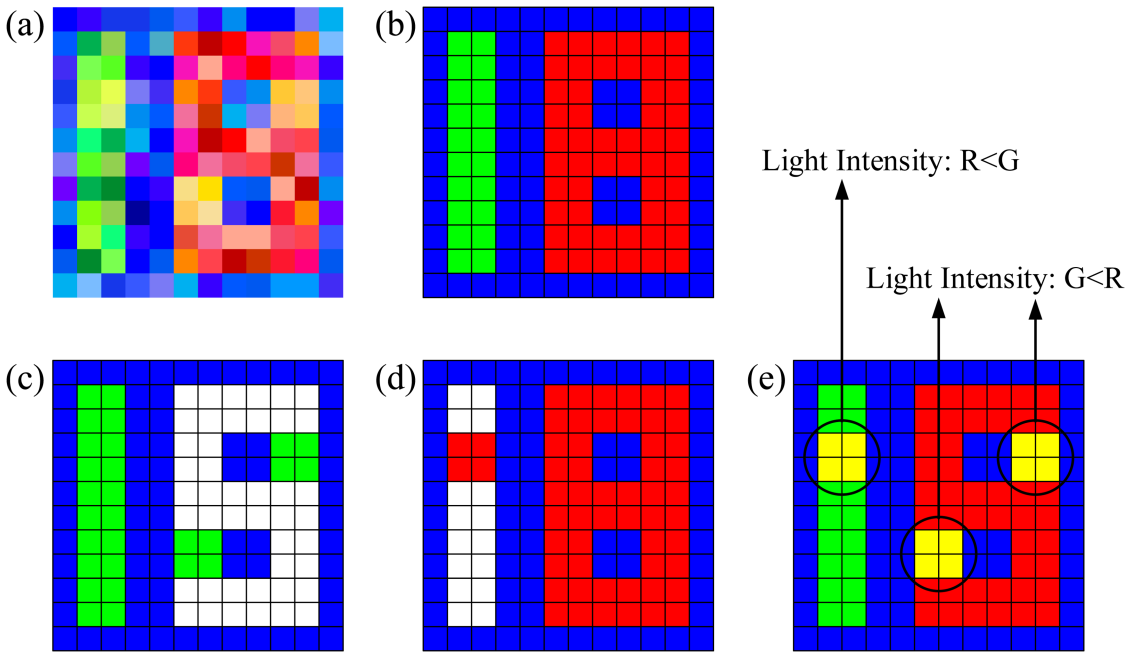

Fig. 7. Color perception experiments. (a) Input map: "is" or "18". Readers will see "18" if they focus on green and red, respectively. Readers will notice "is" if they focus on both green and red; (b) The RGB recognition map with single attentional focus; (c) The GB recognition map with double attentional focuses; (d) The RB recognition map with double attentional focuses; (e) The RGBY recognition map with double attentional focus. The yellow pixel represents that there are two attentional focuses.

different functional requirements by changing V_B. Unlike the traditional pattern recognition systems [30], the proposed color feature selection unit does not require relevant processing such as grayscale and retains the image's color information. More importantly, the unit can directly perform color feature selection at the data position without additional data processing and burdensome training, thereby overcoming the von Neumann bottleneck.

Dynamic optical recognition measurements: The dynamic recognition measurements are carried out to evaluate the dynamic optical characteristics and the sensitivity of the attention. The schematic diagram of dynamic recognition is shown in Fig. 8. The measurement setup is similar to that of the steady-state test. The only difference is that LED is driven by a periodic square wave voltage to generate on-off keying signal. When the driven voltage is high, the RGB light intensity ratio is 15:17:8. Conversely, the intensity ratio is 15:9:8 with low-level driven voltage. The green and red light alternately became the single-color light with maximum intensity, and the blue light as background light does not contribute to the recognition process. V_B is fixed at 0.65 V during this measurement. The driving voltages and the recognition results with the green light at different frequencies are illustrated in Fig. 9(a) and Fig. 9(b). It can be observed from the results that the unit can rapidly switch its attention to the light with the maximum light intensity in any period. The maximum achievable recognition frame rate is 1 kHz. It is confirmed that the color feature selection unit has a strong anti-interference ability in capturing the dynamic change of incident light, even though the stray light has a relatively strong intensity.

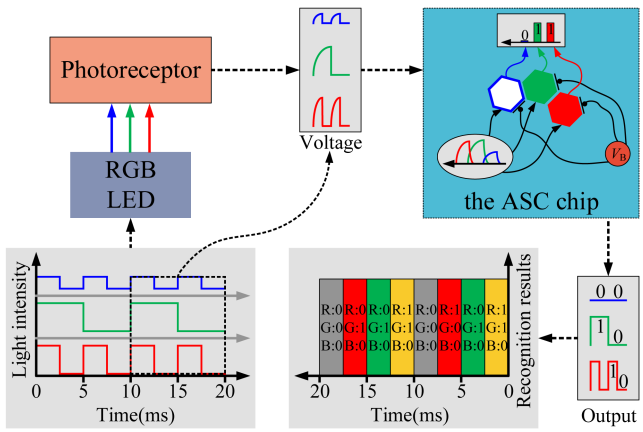

Fig. 8. The schematic diagram of the dynamic color feature selection. In dynamic recognition, the multi-color light in front of photoreceptors flashes on and off freely in different periods, and the final recognition results can show the periodic color change of the color recognition unit.

Furthermore, freely combined optical signals in different frequencies and light intensities (15:15:4) are fed into the recognition unit with a fixed V_B of 0.55 V. Fig. 9(c) demonstrates the driving voltage and recognition results. The unit can respond to periodical on-off light and quickly shift its attention to the branch whose photoreceptor exhibits the most intensive response. And when the red and green light are simultaneously illuminated onto the photoreceptors, the unit can distract the attention to the respective branches. When the results are thoroughly examined, it is not difficult to conclude that the unit's attentional set-shifting and distraction can be implemented in real-time. The

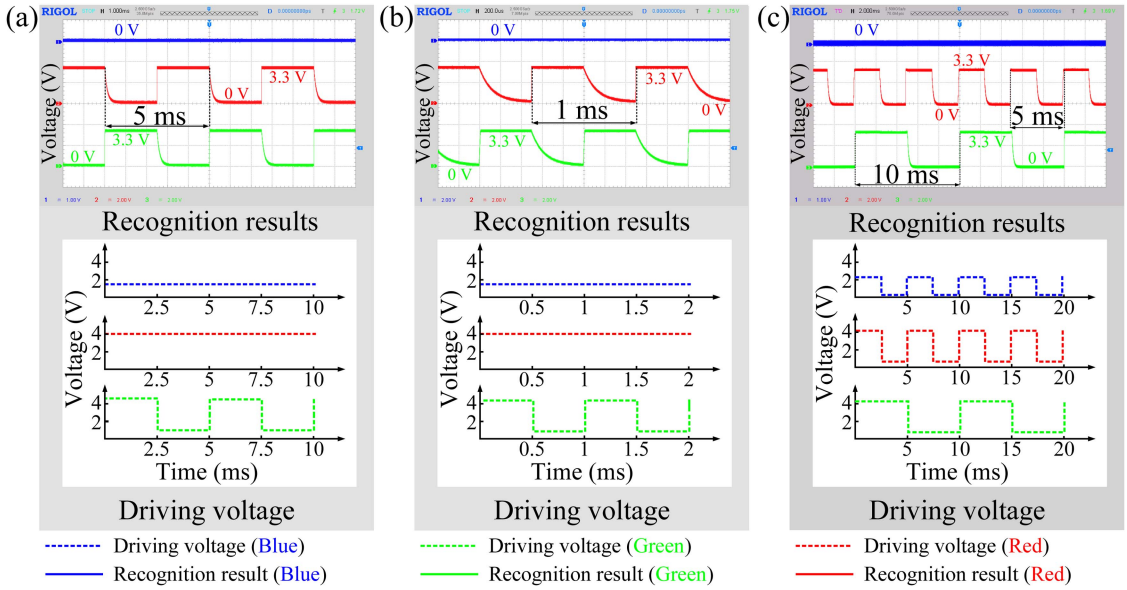

Fig. 9. Dynamic recognition measurements of the color feature selection unit. (a) The recognition results with G-LED driving frequency at 200 Hz; (b) The recognition results with G-LED driving frequency at 1 kHz. It can be seen that the unit can shift its attention according to the change of green light. When the green light is off (driving voltage is low), the red light comes into attentional focus (output voltage is 3.3 V) in real time, and when the green light is on (driving voltage is high), the attentional focus is the green light; (c) The recognition of dynamic trichromatic light. For further confirmation, the unit can recognize the on-off lights in different periods, and the results show the periodic changes of color connected with the incident light.

TABLE I
SUMMARY AND COMPARISON WITH SEVERAL SYSTEMS INSPIRED BY THE BIOLOGICAL VISION SYSTEMS

Inspiration	Fabrication process	Device structure or architecture	Recognition rate (Hz)	
Visual attention mechanism	CMOS VLSI	Neuromorphic hardware	3	[13]
Retinal	Vapour-phase approach	Diode and perovskite nanowire	52-42	[31]
Sensory adaptation	Transition-metal dichalcogenide	Phototransistor	1-3	[32]
Retinal	PAN-coated glasses	NIR organic photosensor	1	[33]
Visual system	Configuration of Pd/MoOx/ITO	Photoelectronic synaptic device	>10	[34]
Visual attention mechanism	Software and algorithms	Deep neural network	12.5	[35]
Visual attention mechanism	Standard CMOS technology	Color feature selection chip	1000	This work

unit can encode the color information into the peak amplitude and frequency of the output signals, enabling real-time optical communication. Moreover, the output signal period follows that of the LED driving voltage without signal cross-talk caused by the spectral overlap of photoreceptors, and the output signals are compatible with subsequent processing circuits and algorithms, thus reducing the design complexity of the whole system.

Note that, the color recognition unit has an inherent delay of 0.5 ms. It may be due to the slow carrier extraction speed of ASC, which results in the inability of the unit to complete the state transition and limits the recognition speeds. Several novel technologies have been proposed to develop the artificial neuromorphic visual system, as listed in Table I. The dynamic recognition speed is an order of magnitude faster than biological vision systems and two orders of magnitude higher than current artificial vision systems. Moreover, the unit based on standard CMOS-compatible processes has significant advantages in terms of low power consumption, anti-interference, and low hardware complexity compared with several visual systems fabricated in advanced manufacturing processes.

IV. CONCLUSION

In this paper, a color feature selection unit consisting of photoreceptors and an ASC was designed. The experiments have shown that the ASC can simulate the brain's visual attentional selection that shifts attention to the most significant color feature in real-time or pay special attention to non-significant color feature by applying a top-down stimulus. The steady-state test results show that the unit can recognize single-color and dual-color light, and the percentage ratio between the intensity

of stray light relative to the strongest light is up to 93.58%. Experiments related to images processing proved that the unit can be used to recognize color patterns and satisfy different color resolution requirements by adjusting the number of attentional focuses and circuit resolution. The dynamic test results show that the unit has a high resolution to periodical on-off visible light, and the maximum achievable frame rate is 1 kHz, which is an order of magnitude faster than biological vision systems. In conclusion, the unit can filter the interference of photoreceptor's spectral overlap and overcomes the von Neumann bottleneck without memory requirements. In addition, it closes the branches outside of the attentional focus, greatly reducing the overall power consumption. Moreover, the recognition unit of RGB light presented in this paper can be easily expanded to more distinct colors and attentional focuses in accordance with the color resolution requirements, which increases performance over existing computer vision feature extraction approaches, providing new insights for the development of artificial neuromorphic visual systems to simulate the initiative, selectivity, and adaptability of biological vision systems.

REFERENCES

[1] R. K. Singh, L. M. Occelli, F. Binette, S. M. Petersen-Jones, and I. O. Nasonkin, "Transplantation of human embryonic stem cell-derived retinal tissue in the subretinal space of the cat eye," *Stem Cells Dev.*, vol. 28, no. 17, pp. 1151–1166, Sep. 2019, doi: 10.1089/scd.2019.0090.

[2] W. S. Tuten, W. M. Harmening, R. Sabesan, A. Roorda, and L. C. Sincich, "Spatiochromatic interactions between individual cone photoreceptors in the human retina," *J. Neurosci.*, vol. 37, no. 39, pp. 9498–9509, Sep. 2017, doi: 10.1523/JNEUROSCI.0529-17.2017.

[3] K. Siuda-Krzywicka and P. Bartolomeo, "What cognitive neurology teaches us about our experience of color," *Neuroscientist*, vol. 26, no. 3, pp. 252–265, Jun. 2020, Art. no. 1073858419882621, doi: 10.1177/1073858419882621.

[4] Z. J. Lan and F. R. Zhu, "Electrically switchable color-selective organic photodetectors for full-color imaging," *ACS Nano*, vol. 15, no. 8, pp. 13674–13682, Aug. 2021, doi: 10.1021/acsnano.1c04908.

[5] D. Joksas et al., "Committee machines-a universal method to deal with non-idealities in memristor-based neural networks," *Nature Commun.*, vol. 11, no. 1, Aug. 2020, Art. no. 4273, doi: 10.1038/s41467-020-18098-0.

[6] C. Gundlach, S. Moratti, N. Forschack, and M. M. Muller, "Spatial attentional selection modulates early visual stimulus processing independently of visual alpha modulations," *Cereb. Cortex*, vol. 30, no. 6, pp. 3686–3703, Jun. 2020, doi: 10.1093/cercor/bhz335.

[7] X. M. Chen, M. Zirnsak, G. M. Vega, E. Govil, S. G. Lomber, and T. Moore, "Parietal cortex regulates visual salience and salience-driven behavior," *Neuron*, vol. 106, no. 1, pp. 177–187, Apr. 2020, doi: 10.1016/j.neuron.2020.01.016.

[8] C. Bartolozzi and G. Indiveri, "Selective attention in multi-chip address-event systems," *Sensors*, vol. 9, no. 7, pp. 5076–5098, Jul. 2009, doi: 10.3390/s90705076.

[9] M. M. Chun, J. D. Golomb, and N. B. Turk-Browne, "A taxonomy of external and internal attention," in *Proc. Annu. Rev. Psychol.*, S. T. Fiske, D. L. Schacter, and S. E. Taylor, Eds., 2011, vol. 62, pp. 73–101.

[10] B. Zhang, D. Y. Xiong, and J. S. Su, "Neural machine translation with deep attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 1, pp. 154–163, Jan. 2020, doi: 10.1109/TPAMI.2018.2876404.

[11] W. Cui et al., "Multi-scale semantic segmentation and spatial relationship recognition of remote sensing images based on an attention model," *Remote Sens.*, vol. 11, no. 9, May 2019, Art. no. 1044, doi: 10.3390/rs11091044.

[12] W. Z. Wang, S. Y. Wang, Y. Li, and Y. S. Jin, "Adaptive multi-scale dual attention network for semantic segmentation," *Neurocomputing*, vol. 460, pp. 39–49, Oct. 2021, doi: 10.1016/j.neucom.2021.06.068.

[13] G. Indiveri, "A neuromorphic VLSI device for implementing 2-D selective attention systems," *IEEE Trans. Neural Netw.*, vol. 12, no. 6, pp. 1455–1463, Nov. 2001, doi: 10.1109/72.963780.

[14] F. Ullah and A. Ben-Hur, "A self-attention model for inferring cooperativity between regulatory features," *Nucleic Acids Res.*, vol. 49, no. 13, Jul. 2021, Art. no. e77, doi: 10.1093/nar/gkab349.

[15] R. C. Zhu et al., "Phase-to-pattern inverse design paradigm for fast realization of functional metasurfaces via transfer learning," *Nature Commun.*, vol. 12, no. 1, May 2021, Art. no. 2974, doi: 10.1038/s41467-021-23087-y.

[16] B. C. Leavell and X. E. Bernal, "The cognitive ecology of stimulus ambiguity: A predator-prey perspective," *Trends Ecol. Evol.*, vol. 34, no. 11, pp. 1048–1060, Nov. 2019, doi: 10.1016/j.tree.2019.07.004.

[17] A. Pooresmaeili, J. Poort, and P. R. Roelfsema, "Simultaneous selection by object-based attention in visual and frontal cortex," *Proc. Nat. Acad. Sci.*, vol. 111, no. 17, pp. 6467–6472, Apr. 2014, doi: 10.1073/pnas.1316181111.

[18] H. Li et al., "A learnable parallel processing architecture towards unity of memory and computing," *Sci. Rep.*, vol. 5, Aug. 2015, Art. no. 13330, doi: 10.1038/srep13330.

[19] M. Mather and M. R. Sutherland, "Arousal-biased competition in perception and memory," *Perspect. Pssychol. Sci.*, vol. 6, no. 2, pp. 114–133, Mar. 2011, doi: 10.1177/1745691611400234.

[20] E. I. Knudsen, "Neural circuits that mediate selective attention: A comparative perspective," *Trends Neurosci.*, vol. 41, no. 11, pp. 789–805, Nov. 2018, doi: 10.1016/j.tins.2018.06.006.

[21] O. Le Meur, P. Le Callet, D. Barba, and D. Thoreau, "A coherent computational approach to model bottom-up visual attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 802–817, May 2006, doi: 10.1109/TPAMI.2006.86.

[22] A. M. Fernandes et al., "Neural circuitry for stimulus selection in the Zebrafish visual system," *Neuron*, vol. 109, no. 5, pp. 805–822, Mar. 2021, doi: 10.1016/j.neuron.2020.12.002.

[23] R. A. John et al., "Ultralow power dual-gated subthreshold oxide neuristors: An enabler for higher order neuronal temporal correlations," *ACS Nano*, vol. 12, no. 11, pp. 11263–11273, Nov. 2018, doi: 10.1021/acsnano.8b05903.

[24] I. Munguira et al., "Glasslike membrane protein diffusion in a crowded membrane," *ACS Nano*, vol. 10, no. 2, pp. 2584–2590, Feb. 2016, doi: 10.1021/acsnano.5b07595.

[25] N. Ness and S. R. Schultz, "A computational grid-to-place-cell transformation model indicates a synaptic driver of place cell impairment in early-stage Alzheimer's Disease," *Plos Comput. Biol.*, vol. 17, no. 6, Jun. 2021, Art. no. e1009115, doi: 10.1371/journal.pcbi.1009115.

[26] T. Li et al., "Action potential initiation in neocortical inhibitory interneurons," *Plos Biol.*, vol. 12, no. 9, Sep. 2014, Art. no. e1001944, doi: 10.1371/journal.pbio.1001944.

[27] R. L. Costea and C. A. Marinov, "A consistent model for Lazzaro winner-take-all circuit with invariant subthreshold behavior," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2375–2385, Nov. 2016, doi: 10.1109/TNNLS.2015.2489862.

[28] M. Shoaran, A. Tajalli, M. Alioto, A. Schmid, and Y. Leblebici, "Analysis and characterization of variability in subthreshold source-coupled logic circuits," *IEEE Trans. Circuits Syst. I. Reg. Papers*, vol. 62, no. 2, pp. 458–467, Feb. 2015, doi: 10.1109/TCSI.2014.2364101.

[29] H. Ren, J. D. Chen, Y. Q. Li, and J. X. Tang, "Recent progress in organic photodetectors and their applications," *Adv. Sci.*, vol. 8, no. 1, Jan. 2021, Art. no. 2002418, doi: 10.1002/advs.202002418.

[30] M. Chu et al., "Neuromorphic hardware system for visual pattern recognition with memristor array and CMOS neuron," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2410–2419, Apr. 2015, doi: 10.1109/TIE.2014.2356439.

[31] L. L. Gu et al., "A biomimetic eye with a hemispherical perovskite nanowire array retina," *Nature*, vol. 581, no. 7808, pp. 278–282, May 2020, doi: 10.1038/s41586-020-2285-x.

[32] S. Hong et al., "Sensory adaptation and neuromorphic phototransistors based on $CsPB(Br_1\text{-}Xix_3)$ perovskite and MoS_2 hybrid structure," *ACS Nano*, vol. 14, no. 8, pp. 9796–9806, Aug. 2020, doi: 10.1021/acsnano.0c01689.

[33] H. L. Wang et al., "A retina-like dual band organic photosensor array for filter-free near-infrared-to-memory operations," *Adv. Mater.*, vol. 29, no. 32, Aug. 2017, Art. no. 1701772, doi: 10.1002/adma.201701772.

[34] F. C. Zhou et al., "Optoelectronic resistive random access memory for neuromorphic vision sensors," *Nature Nanotechnol.*, vol. 14, no. 8, pp. 776–782, Aug. 2019, doi: 10.1038/s41565-019-0501-3.

[35] N. Wagatsuma, A. Hidaka, and H. Tamura, "Correspondence between monkey visual cortices and layers of a saliency map model based on a deep convolutional neural network for representations of natural images," *Eneuro*, vol. 8, no. 1, Jan. 2021, Art. no. 0200-20.2020, doi: 10.1523/ENEURO.0200-20.2020.