Check for updates

OPEN ACCESS

EDITED BY Bo Wang, Singapore University of Technology and Design, Singapore

REVIEWED BY

Mikhail A. Mishchenko, Lobachevsky State University of Nizhny Novgorod, Russia Desmond Loke, Singapore University of Technology and Design, Singapore

*CORRESPONDENCE Junjie Wang ⊠ wangjunjie@uestc.edu.cn

RECEIVED 24 March 2023 ACCEPTED 16 May 2023 PUBLISHED 07 June 2023

CITATION

Wang H, Pan X, Wang J, Sun M, Wu C, Yu Q, Liu Z, Chen T and Liu Y (2023) Dual functional states of working memory realized by memristor-based neural network. *Front. Neurosci.* 17:1192993. doi: 10.3389/fnins.2023.1192993

COPYRIGHT

© 2023 Wang, Pan, Wang, Sun, Wu, Yu, Liu, Chen and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Dual functional states of working memory realized by memristor-based neural network

Hongzhe Wang¹, Xinqiang Pan¹, Junjie Wang^{1*}, Mingyuan Sun², Chuangui Wu¹, Qi Yu¹, Zhen Liu³, Tupei Chen⁴ and Yang Liu^{1,5}

¹State Key Laboratory of Electronic Thin Films and Integrated Devices, University of Electronic Science and Technology of China, Chengdu, China, ²Beijing China Changfeng Electromechanical Technology Research and Design Institute, Beijing, China, ³School of Materials and Energy, Guangdong University of Technology, Guangzhou, China, ⁴School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore, ⁵Deepcreatic Technologies Ltd., Chengdu, China

Working memory refers to the brain's ability to store and manipulate information for a short period. It is disputably considered to rely on two mechanisms: sustained neuronal firing, and "activity-silent" working memory. To develop a highly biologically plausible neuromorphic computing system, it is anticipated to physically realize working memory that corresponds to both of these mechanisms. In this study, we propose a memristor-based neural network to realize the sustained neural firing and activity-silent working memory, which are reflected as dual functional states within memory. Memristor-based synapses and two types of artificial neurons are designed for the Winner-Takes-All learning rule. During the cognitive task, state transformation between the "focused" state and the "unfocused" state of working memory is demonstrated. This work paves the way for further emulating the complex working memory functions with distinct neural activities in our brains.

KEYWORDS

memristor, working memory, neural networks, bio-inspired computing, Hebbian learning

1. Introduction

Working memory is an essential brain function that allows for the temporary storage and manipulation of information required for cognitive tasks (Baddeley and Hitch, 1974; Morris, 1986; Baddeley, 1992, 2010). For a long time, it was thought to be presented in the form of persistent neuronal firing during the delay period (Funahashi, 2017). However, recent studies have suggested that synaptic weight can also store information during the delay period, even if persistent neuronal firing has ceased (Mongillo et al., 2008; Stokes, 2015; Silvanto, 2017). This phenomenon is referred to as "activity-silent" working memory. In most previous studies, the sustained neuronal firing and "activity-silent" working memory have been modeled independently, and these mechanisms appear to be fundamentally opposed in principle. On the other hand, in recent years, several studies have provided insights into the interaction between sustained neuronal firing and "activity-silent" working memory. Manohar et al. have proposed a memory model that unites both persistent activity attractors and silent synaptic memory, which is applicable to many empirical phenomena (Manohar et al., 2019). Barbosa et al. have investigated the interplay between persistent activity and activity-silent dynamics in the prefrontal cortex using monkey and human electrophysiology data (Barbosa et al., 2020).

In the past decades, a great number of efforts had been made for the hardware implementation of a wide variety of artificial neural networks (Misra and Saha, 2010; Capra et al., 2020; Nguyen et al., 2021; Ghimire et al., 2022). Recent works on silent synapses and artificial synapses have highlighted their potential for advancing the understanding of the nervous system and developing neuromorphic computing technologies (Loke et al., 2016; Go et al., 2021; Hao et al., 2021). Following this research trend, there is a growing anticipation to realize a highly bio-plausible neuromorphic computing system. Although working memory plays a vital role in biological neurocomputing (Wang et al., 2020), there have been only a handful of studies on the hardware implementation of working memory, and the existing research has mainly focused on the independent neural mechanism of working memory (Brown and Aggleton, 2001; Ji et al., 2022). Proposing a hardware design for working memory that is compatible with both sustained neuronal firing and activity-silent working memory can improve its biological plausibility and expand the breadth of its application.

To achieve the aforementioned dual functional states of working memory, this paper proposes a hardware design for working memory based on memristors. Memristor is a nonlinear two-terminal electrical device that has been extensively studied in the past decade, which is a key element used in artificial neural networks for synapses and neurons due to similarities in electrical behavior (Chua, 1971; Strukov et al., 2008; Thomas, 2013; Li et al., 2018; Camuñas-Mesa et al., 2019; Xia and Yang, 2019). In this work, we propose a memristorbased neural network to realize the dual functional states of working memory. To achieve this, the electrical characteristics of an Au/LNO/Pt memristor based on Single-Crystalline LiNbO3 (SC-LNO) thin films is utilized. The use of the high-quality SC-LNO thin film results in several advantageous properties, including high switching uniformity, long retention time, stable endurance performance, and reproducible multilevel resistance states (Wang et al., 2022). An artificial synapse circuit with simplified Hebbian learning rule is implemented with Au/LNO/Pt memristor. A spiking neural network capable of realizing the winner-takesall (WTA) functionality is constructed, which is utilized to achieve working memory working memory. State transformation between the "focused" state (sustained neuronal firing) and the "unfocused" state (activity-silent working memory) of memristorbased working memory is demonstrated. This hardware solution for bio-plausible working memory with dual functional states, leveraging the intrinsic electrical properties of memristors, has promising implications for the development of advanced bioplausible neuromorphic computing systems.

2. Materials and methods

The memristor utilized in this study is an Au/LNO/Pt memristor, which is based on 30 nm Single-Crystalline $LiNbO_3$ (SC-LNO) thin films. The detailed fabrication method was presented in our prior work (Wang et al., 2022). The electrical characteristics of the memristor were obtained at room temperature using a Keithley 4200-SCS Semiconductor Characterization System. A modified Yakopcic generalized

memristor model (Yakopcic et al., 2011) is employed to fit the experimental data of the LNO memristor, taking into account its inherent instability. The dual funtiaonal states of working memory was validated using the Brian spiking neural network simulator (Brian 2). The memristor-based working memory circuit is designed and verified through SPICE simulations.

2.1. Dual functional states of working memory

Figure 1 depicts the working memory network model that supports "focused" state (sustained neuronal firing) and the "unfocused" state (activity-silent working memory), which builds upon Manohar's working memory model (Manohar et al., 2019) by adjusting it to the form of Spiking Neural Networks (SNNs), thus endowing it with greater biological plausibility. The network comprises two distinct types of neurons: feature-selective neurons and freely-conjunctive neurons. Feature-selective neurons receive unique types of feature information such as colors, orientations, and locations. On the other hand, freely-conjunctive neurons encode a combination of simultaneously active features and establish an associative mapping to feature-selective neurons. Upon arrival of feature information stimulus, the membrane potential of the corresponding feature-selective neuron increases. As the potential of a feature-selective neuron approaches its threshold, the neuron fires a spike. The spike train generated by featureselective neurons can be interpreted as sensory activation of feature information, and it is perceived by freely-conjunctive neurons. Initially, the synaptic weights between the two types of neurons are randomly assigned, reflecting the connection strength between neurons. As spikes arrive from feature-selective neurons, the membrane potential of freely-conjunctive neurons increases, eventually leading to firing. Synaptic connection strength changes according to the temporal relationship between pre- and postsynaptic spikes, in accordance with the Hebbian plasticity rule. Freely-conjunctive neurons also compete with each other through lateral inhibition and self-excitation, following the winner-takes-all (WTA) rule. Through this competition, only one neuron remains active in each feature dimension. The working memory network model includes two independent weight vectors connecting feature-selective neurons and freely-conjunctive neurons, which depend on the direction of spike propagation. In Figure 1, W_{fc} indicates the synaptic weight of the forward direction (featureselective neurons to freely-conjunctive neurons), and W_{cf} indicates the synaptic weight of the backward direction (freely-conjunctive neurons to feature-selective neurons). During Hebbian learning, both feature-to-conjunctive synapses and conjunctive-to-feature synapses are strengthened for the winner among freely-conjunctive neurons, while they are weakened for other freely-conjunctive neurons. This process creates an synaptic mapping between the freely-conjunctive neurons and the feature information stimulus.

Figure 2 shows the sequence of neuronal events in working memory: *encoding*, *attention*, and *retrieval*. Step 1: sensory inputs (color: blue, orientation: +45 degrees, location: top right) are received by feature-selective neurons. Then, freely-conjunctive



neurons perceive spike trains from feature-selective neurons. The freely-conjunctive neurons filled with light red indicate relatively lower and erratic firing rates. During this process, the freely-conjunctive neurons compete and encode the active features. Following this, each feature dimension has a single neuron that wins a competition, which can be seen as neuronal events of encoding in working memory. Step 2: the neurons filled with deep red indicate higher and more stable firing rates. Under the WTA rule, firing from the winner among the freely-conjunctive neurons is stabilized. Step 3: when the sensory input to the feature-selective neuron is withdrawn, the freelyconjunctive neuron that has been firing stimulates the featureselective neuron through the conjunctive-to-feature synapses, this results in the firing of the feature-selective neuron and, in turn causes the freely-conjunctive neuron to fire through the featureto-conjunctive synapses. The self-excitation synapse of the freelyconjunctive neuron enhances its firing through self-stimulation. This sustained neuronal firing can be considered as neuronal events of attention in working memory, which indicate "focused" state of working memory. Step 4: a new round of encoding and competition in the working memory network, and different freely-conjunctive neuron from the previous round may enter the "focused" state. Although the feature information from the previous round is not presented in the working memory network in the "focused" state, it remains encoded in synaptic weights as an activity-silent working memory in the "unfocused" state. Step 5: The providing partial feature information to feature-selective neurons, which reactivates the "focused" state of corresponding freely-conjunctive neurons. This can be considered as neuronal events of retrieval in working memory. Step 6: freely-conjunctive neuron reactivates original features, completing the process of associative recall. From a certain perspective, the activity-silent working memory in this model can be conceptualized as an associative memory with long-term information storage. Although working memory is typically regarded as a form of short-term memory, several studies suggest that associative memory with long-term information storage also contribute to working memory (Burgess and Hitch, 2005; Olson et al., 2005; van Geldorp et al., 2012). During working memory tasks, as many rounds of sensory inputs are applied, the synaptic mapping of activity-silence working memory may be interrupted by new feature information. This is attributed to the limited capacity of working memory, which is consistent with the concept of forgetting in working memory. Although sustained neuronal firing weakens and ceases during the bi-directional feedback between two types of neurons after a relatively short period of time, this process indicates a transition from the "focused" state to the "unfocused" state, rather than forgetting.



2.2. Synaptic behavior of the LNO memristor

The cross-sectional structure of the Au/LNO/Pt memristor used in this work is shown in the inset part of Figure 3A. The insulating layer of the memristor is a Single-Crystalline LiNbO₃ (SC-LNO) thin film with the thickness of 30 nm. The cross mark in Figure 3B display the electronic behavior of LNO memristive conductance under appropriate spike trains applied to the memristor. The blue cross mark depict the LTP characteristic of the memristor. Positive pulse trains comprising 100 sequential 1.5 V pulses, each lasting 5ms, are applied to the top electrode (i.e., the Au electrode) of the memristor while the bottom electrode (i.e., the Pt electrode) is grounded, and the memristive conductance gradually increases, indicating that the memristor undergoes a SET process. The red cross mark depict the LTD characteristic of the memristor. Negative pulse trains comprising 100 sequential -1.5 V pulses, each lasting 5 ms, are applied to the bottom electrode (i.e., the Pt electrode) of the memristor while the top electrode (i.e., the Au electrode) is grounded, and the memristive conductance gradually decreases, indicating that the memristor experiences a RESET process.

In this work, a modified Yakopcic's generalized memristor model (Yakopcic et al., 2011) is used to fit the experimental data of the LNO memristor. The equation of the modified memristor model is as follows:



$$I(x) = \begin{cases} a_1(\frac{1}{r} + \frac{(r-1)x(t)}{r})\sinh(bV(t)), V(t) \ge 0\\ a_2(\frac{1}{r} + \frac{(r-1)x(t)}{r})\sinh(bV(t)), V(t) < 0 \end{cases}$$
(1)

$$g(V(t)) = \begin{cases} A_p(e^{V(t)} - e^{V_p}), V(t) > V_p \\ -A_n(e^{-V(t)} - e^{V_n}), V(t) < -V_n \\ 0, -V_n < V(t) < V_p \end{cases}$$
(2)

$$f(x(t)) = \begin{cases} e^{-\alpha_p(x(t) - x_p)} w_p(x(t), x_p), x(t) \ge x_p \\ 1, x(t) < x_p \end{cases}$$
(3)

$$f(x(t)) = \begin{cases} e^{\alpha_n(x(t)+x_n-1)} w_n(x(t), x_n), x(t) \le 1 - x_n \\ 1, x(t) > 1 - x_n \end{cases}$$
(4)

$$w_p(x, x_p) = \frac{x_p - x}{1 - x_p} + 1$$
(5)

$$w_n(x,x_n) = \frac{x}{1-x_n} \tag{6}$$

$$\frac{dx}{dt} = \eta g(V(t)) f(x(t)) \tag{7}$$

In Equation (1), the hyperbolic sinusoid function is used to fit the I-V relationship of the memristor, along with parameters a_1 , a_2 , and b. The state variable of the memristor is represented by x(t), which varies between 0 and 1. At x(t) = 0, the memristor's resistance reaches its maximum value, while x(t) = 1 corresponds to the minimum resistance. In contrast to the original Yakopcic memristor model, a parameter r is introduced in Equation (1) to represent the ratio of the highest resistance value

to the lowest resistance value. The purpose of this modification is to adjust the resistance range of the memristor to fall within the range of experimental data and to prevent the occurrence of infinitely large or small resistance values. Equation (2) presents the function g(V(t)), which imposes a programming threshold on the memristor. The positive and negative thresholds are denoted by V_p and V_n , respectively, with adjustable parameters for the magnitude of the exponentials represented by A_p and A_n . The state variable motion is modeled by the function f(x(t)), which is expressed in Equations (3) and (4). The motion remains constant until the point x_p or x_n , where the rate of exponential decay is determined by α_p or α_n , respectively. The window function $w_n(x, x_n)$ is utilized to ensure the boundary of the state variable motion, as shown in Equations (5) and (6). Additionally, Equation (7) models the state variable motion, with the direction of the motion represented by η in terms of the voltage polarity. Figure 3A shows the current-voltage (I-V) relationship of the LNO memristor model under a voltage sweep from -2 to 2 V for both experiment and simulation. The LTP and LTD characteristics of the modified Yakopcic's generalized memristor model are illustrated by the green and orange lines in Figure 3B, with the parameters used in this work listed in Table 1. Considering the inherent instability of Cycle-to-cycle and Deviceto-device variations in practical memristors, 40, 10, 25, and 20% noise are introduced to α_p , α_n , A_p , and A_n , respectively. The errorbars on the green and orange lines represent the maximum conductance variations of the memristor after 100 sequential LTD and LTP pulses.

2.3. Memristor-based working memory circuit

In this study, a simplified Hebbian learning rule is introduced as the synaptic plasticity mechanism for feature-to-conjunctive

TABLE 1 The parameters of the LNO memristor model.

<i>a</i> ₁	<i>a</i> ₂			Vp	V _n	Ap	A _n	x _p	x _n	α_p	α_n	x _o	
$7 \times e^{-7}$	$7 \times e^{-7}$	1.5	9	0.5	0.5	0.95	35	0.05	0.3	1.5	5	0.01	1



synapses and conjunctive-to-feature synapses, as illustrated in Figure 4. Long-term potentiation (LTP) and long-term depression (LTD) of synaptic weights are determined by the temporal relation between pre- and post-synaptic spikes. When the pre-synaptic spike is fired first and the post-synaptic spike is fired immediately within a time window of 2ms, the synapse exhibits LTP. When the post-synaptic spike is fired outside of this time window, either earlier or later, the synapse exhibits LTD. The strength of LTP and LTD is determined by the SET/RESET pulse width, denoted by time scalar T_p . The pulse width for SET is three times wider than that for RESET. $\hat{\nu}$ represents the direction of the pulse voltage, with the direction of SET being positive and the direction of RESET being negative.

To harness the LTP/LTD behavior of the LNO memristor, a dedicated synaptic weight update circuit has been designed and optimized based on our previous research (Hu et al., 2013). The synaptic weight update circuit is comprised of a single LTP module, a single LTD module, one memristor, and the peripheral circuit. Schematic representations of the LTP and LTD modules are shown in Figures 5A, B, respectively. The "PRE" and "POST" nodes are connected to the pre-neuron and post-neuron, respectively, while the LTP and LTD nodes represent the output generated by the temporal relation between pre-synaptic and post-synaptic spikes, as specified by the simplified Hebbian learning rule mentioned earlier. At the initial state, both "SP" and "SD" nodes are set to a low level. I₁ and I₃ function as inverters, while I₂ and I₄ function as NAND gates. When the pre-synaptic neuron fires, transistor MP₁ turns on, causing capacitor C₁ to charge up to V_{dd}, and as a result, the "SP"

node rises to V_{dd}. The output of I₄ in the LTD module is held at a low level, which means that the LTD module is inactive while the LTP module is operating, and vice versa. When the LTP module is in operation, C1 begins to discharge through transistors MN1 and MN₂, with the discharge current being regulated by V_{bp}. The maximum duration for the discharge of C1 is denoted as tLTP, which represents the time window of the LTP. When the post-synaptic neuron fires within tLTP, transistor MN3 will be turned on, causing the charge to be redirected to capacitor C2, and setting the "S" input of the S-R latch to a high level. As a result, the LTP output of the S-R latch will be in a "HIGH" state. When the LTD module is in operation, the "S" input of the S-R latch is immediately set to a high level, causing the LTD output of the S-R latch to be in a "HIGH" state. The only purpose of the discharge of C₃ is to deactivate the LTP module while the LTD module is in operation. The LTP/LTD output will remain in a "HIGH" state until a "Ctrl" signal arrives at the S-R latch, which in turn determines the duration of the LTP/LTD output.

The LTP and LTD outputs described above are utilized to implement a Hebbian learning rule in the synapse weight update circuit, as depicted in Figure 6. When the LTP output is in a "HIGH" state, transistors MN7 and M10 turn on, applying a positive voltage Vr to the memristor with the top electrode at high potential and leading to a SET process in the memristor. Similarly, when the LTD output is in a "HIGH" state, transistors MN8 and MN9 turn on, applying V_r to the memristor with the bottom electrode at low potential, leading to a RESET process in the memristor. The duration of the SET/RESET process is equal to the duration of the LTP/LTD output. When the LTP and LTD outputs are in a low state, their inverted outputs LTP and LTD are in a high state. As a result, transistors MN₁₁-MN₁₄ in the path between the presynaptic neuron and the post-synaptic neuron remain on, allowing pre-synaptic spikes to transmit to the post-synaptic neuron through the memristor. However, when the LTP or LTD module is activated, two of the transistors among MN₁₁-MN₁₄ turn off, and the path between the pre-synaptic neuron and the post-synaptic neuron is closed.

To achieve working memory, we utilized a spiking neural network with the winner-takes-all (WTA) functionality previously developed in our research (Wang et al., 2019). All synapses involved in working memory, including feature-to-conjunctive synapses, conjunctive-to-feature synapses, self-exciting synapses, and lateralinhibition synapses, share the same synaptic architecture. The synapse weight of feature-to-conjunctive and conjunctive-tofeature synapses can be modified by the aforementioned synapse weight update circuit that utilizes a simplified Hebbian learning rule. On the other hand, the synaptic weight of self-exciting synapses and lateral-inhibition synapses remain fixed. Both featureselective neurons and freely-conjunctive neurons in our network utilize the leaky integrate-and-fire (LIF) neuron model. The capacitor within the neuron integrates the input current from the







synapses, causing the neuron's membrane potential to increase. Once the neuron potential reaches the threshold, the neuron fires a spike, and its potential returns to its resting state. The network topology of our working memory design, consisting of interconnected artificial neurons and synapses as illustrated in Figure 1.

3. Results and discussion

The functionality of the memristor-based working memory was evaluated using SPICE simulation, leveraging the electrical characteristics of Au/LNO/Pt memristors derived from experimental data. The memristive conductance was normalized to serve as synaptic weights. The working memory employed a total of 9 feature-selective neurons, with each group of 3 neurons corresponding to a distinct feature dimension, including color, orientation, and location. Each neuron was responsible for encoding different feature information within its respective dimension. Additionally, 3 freely-conjunctive neurons were fullyconnected to the feature-selective neurons, resulting in a total of 27 feature-to-conjunctive and 27 conjunctive-to-feature synapses. To enable lateral inhibition, 6 synapses were established between freely-conjunctive neurons, with self-connections excluded. Lastly, 3 self-exciting synapses were connected to the freely-conjunctive network in a self-connected manner, for the purpose of inducing self-excitation.

Figure 7 shows a typical sensory input for working memory. Three input features were selected: Obj_1 (color: red, orientation: -45 degrees, location: bottom-left), Obj_2 (color: yellow, orientation: 0 degrees, location: top-left), and Obj_3 (color: blue, orientation: +45 degrees, location: top-right). The sensory current was applied to corresponding feature-selective neurons, and the magnitude of the sensory current was adjusted to the same value for each of the 3 feature dimensions. The firing rate of the feature-selective neuron was proportional to the



magnitude of the input current. An initialization time of 50 ms was allocated for stabilization at the beginning of working memory. Each feature input lasted for 100 ms, and there was a 50 ms resting time between each two features. The time scalar T_p for the simplified Hebbian learning rule is defined as 1 ms.

Figure 8 shows the activity of each freely-conjunctive neuron. When the first feature input Obj_1 is activated, feature-toconjunctive synapses modify their connectivity based on a simplified Hebbian learning rule, and freely-conjunctive neurons compete with each other. This process is known as *encoding* in working memory. Once the competition is completed, the winning neuron remains active even without the sensory input. The feature information is encoded into working memory and is presented in the form of persistent neuronal firing, which can be regarded as the "focus" state of working memory. Additionally, feature information is silently encoded into the synaptic weights of the feature-to-conjunctive synapses. When the second feature Obj₂ is activated, the previous attention is disturbed by the new input, and a new round of encoding occurs. Then, the neurons that won the competition in this round persistently fire, forming a new attention. Similarly, when the third feature input Obj3 is activated, another "focus" state of working memory is formed as the persistent firing freely-conjunctive neurons map to Obj3. However, the "unfocused" state of working memory exists in the form of synaptic mappings and persists as long as it is not overwritten by new feature information during the working memory task. After a resting time of 50 ms, a sensory input composed of partial feature information of Obj1 (colors: red) is applied. Despite using only a small fraction of the feature information, persistent neuronal firing occurs again, which corresponds to the retrieval of Obj1. This result demonstrates that the proposed memristor-based working memory system can successfully achieve the dual functional states of working memory and accomplish the working memory task.

4. Conclusion

In this paper, a memristor-based working memory that is capable of exhibiting dual functional states is presented. To achieve this, an artificial synapse with a simplified Hebbian learning rule was designed based on the LTP/LTD properties of the Au/LNO/Pt memristor, which uses a single-crystalline LiNbO3 (SC-LNO) thin film as its insulating layer. Two types of artificial LIF neurons were implemented in the network to encode feature information to working memory through the WTA rule and produce persistent neuronal firing patterns. The results show that the proposed system can realize various neuronal events in working memory, including encoding, attention, and retrieval. This study demonstrates that the memristor-based working memory can exist in the dual functional states: the sustained neuronal firing and activity-silent working memory. This study paves the way for the development of advanced bio-plausible neuromorphic computing systems based on memristive neural networks. This research represents a significant step toward the development of advanced bio-plausible neuromorphic computing systems based on memristive neural networks. It is hoped that this work will inspire further research in this exciting and rapidly evolving field.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

Baddeley, A. (1992). Working memory. Science 255, 556-559.

Baddeley, A. (2010). Working memory. Curr. Biol. 20, R136-R140. doi: 10.1016/j.cub.2009.12.014

Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *Psychology of Learning and Motivation*, Vol. 8, ed G. A. Bower (New York, NY: Academic Press), 47-89.

Barbosa, J., Stein, H., Martinez, R. L., Galan-Gadea, A., Li, S., Dalmau, J., et al. (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat. Neurosci.* 23, 1016–1024. doi: 10.1038/s41593-020-0644-4

Brown, M. W., and Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nat. Rev. Neurosci.* 2, 51–61. doi: 10.1038/35049064

Burgess, N., and Hitch, G. (2005). Computational models of working memory: putting long-term memory into context. *Trends Cogn. Sci.* 9, 535–541. doi: 10.1016/j.tics.2005.09.011

Camuñas-Mesa, L. A., Linares-Barranco, B., and Serrano-Gotarredona, T. (2019). Neuromorphic spiking neural networks and their memristorcmos hardware implementations. *Materials* 12, 2745. doi: 10.3390/ma121 72745

Capra, M., Bussolino, B., Marchisio, A., Shafique, M., Masera, G., and Martina, M. (2020). An updated survey of efficient hardware architectures for accelerating deep convolutional neural networks. *Fut. Internet* 12, 113.

Author contributions

HW and JW conceived of the presented idea, designed the circuits and performed the computations and simulations, and prepared the manuscript with contributions from all authors. HW, XP, and JW carried out the experiment. XP and CW fabricated the device. HW and XP analyzed the data. MS, QY, and ZL verified the analytical methods. YL, CW, and TC were in charge of overall direction and planning. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by NSFC under project No. 92064004 and Chengdu Technological Fund under project No. 2019-YF08-00256-GX.

Conflict of interest

YL is employed by Deepcreatic Technologies Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Chua, L. (1971). Memristor-the missing circuit element. IEEE Trans. Circ. Theory 18, 507–519.

Funahashi, S. (2017). Working memory in the prefrontal cortex. Brain Sci. 7, 49. doi: 10.3390/brainsci7050049

Ghimire, D., Kil, D., and Kim, S.-H. (2022). A survey on efficient convolutional neural networks and hardware acceleration. *Electronics* 11, 945. doi: 10.3390/electronics11060945

Go, S. X., Lee, T. H., Elliott, S. R., Bajalovic, N., and Loke, D. K. (2021). A fast, low-energy multi-state phase-change artificial synapse based on uniform partial-state transitions. *APL Mater.* 9, 091103. doi: 10.1063/5.0056656

Hao, S., Zhong, S., Ji, X., Pang, K. Y., Wang, N., Li, H., et al. (2021). Activating silent synapses in sulfurized indium selenide for neuromorphic computing. *ACS Appl. Mater. Interfaces* 13, 60209–60215. doi: 10.1021/acsami.1c19062

Hu, S., Wu, H., Liu, Y., Chen, T., Liu, Z., Yu, Q., et al. (2013). Design of an electronic synapse with spike time dependent plasticity based on resistive memory device. *J. Appl. Phys.* 113, 114502. doi: 10.1063/1.4795280

Ji, X., Hao, S., Lim, K. G., Zhong, S., and Zhao, R. (2022). Artificial working memory constructed by planar 2D channel memristors enabling brain-inspired hierarchical memory systems. *Adv. Intell. Syst.* 4, 2100119. doi: 10.1002/aisy.202100119

Li, Y., Wang, Z., Midya, R., Xia, Q., and Yang, J. J. (2018). Review of memristor devices in neuromorphic computing: materials sciences and device challenges. *J. Phys* D Appl. Phys. 51, 503002. doi: 10.1088/1361-6463/aade3f

Loke, D., Skelton, J. M., Chong, T.-C., and Elliott, S. R. (2016). Design of a nanoscale, CMOS-integrable, thermal-guiding structure for boolean-logic and neuromorphic computation. *ACS Appl. Mater. Interfaces* 8, 34530–34536. doi: 10.1021/acsami.6b10667

Manohar, S. G., Zokaei, N., Fallon, S. J., Vogels, T. P., and Husain, M. (2019). Neural mechanisms of attending to items in working memory. *Neurosci. Biobehav. Rev.* 101, 1–12. doi: 10.1016/j.neubiorev.2019.03.017

Misra, J., and Saha, I. (2010). Artificial neural networks in hardware: a survey of two decades of progress. *Neurocomputing* 74, 239–255. doi: 10.1016/j.neucom.2010.03.021

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769

Morris, N. (1986). Working memory, 1974–1984: a review of a decade of research. *Curr. Psychol. Res. Rev.* 5, 281–295.

Nguyen, D.-A., Tran, X.-T., and Iacopi, F. (2021). A review of algorithms and hardware implementations for spiking neural networks. *J. Low Power Electron. Appl.* 11, 23. doi: 10.3390/jlpea11020023

Olson, I. R., Jiang, Y., and Moore, K. S. (2005). Associative learning improves visual working memory performance. J. Exp. Psychol. Hum. Percept. Perform. 31, 889. doi: 10.1037/0096-1523.31.5.889

Silvanto, J. (2017). Working memory maintenance: sustained firing or synaptic mechanisms? *Trends Cogn. Sci.* 21, 152–154. doi: 10.1016/j.tics.2017.01.009

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405. doi: 10.1016/j.tics.2015.05.004

Strukov, D. B., Snider, G. S., Stewart, D. R., and Williams, R. S. (2008). The missing memristor found. *Nature* 453, 80–83. doi: 10.1038/nature06932

Thomas, A. (2013). Memristor-based neural networks. J. Phys. D Appl. Phys. 46, 093001. doi: 10.1088/0022-3727/46/9/093001

van Geldorp, B., Konings, E. P., van Tilborg, I. A., and Kessels, R. P. (2012). Associative working memory and subsequent episodic memory in Alzheimer's disease. *Neuroreport* 23, 119–123. doi: 10.1097/WNR.0b013e32834ee461

Wang, J., Pan, X., Wang, Q., Luo, W., Shuai, Y., Xie, Q., et al. (2022). Reliable resistive switching and synaptic plasticity in Ar+-irradiated single-crystalline LiNbO3 memristor. *Appl. Surface Sci.* 596, 153653. doi: 10.1016/j.apsusc.2022. 153653

Wang, J., Yu, Q., Hu, S., Liu, Y., Guo, R., Chen, T., et al. (2019). Winner-takesall mechanism realized by memristive neural network. *Appl. Phys. Lett.* 115, 243701. doi: 10.1063/1.5120973

Wang, Y., Yu, L., Wu, S., Huang, R., and Yang, Y. (2020). Memristor-based biologically plausible memory based on discrete and continuous attractor networks for neuromorphic systems. *Adv. Intell. Syst.* 2, 2000001. doi: 10.1002/aisy.2020 00001

Xia, Q., and Yang, J. J. (2019). Memristive crossbar arrays for braininspired computing. *Nat. Mater.* 18, 309–323. doi: 10.1038/s41563-019-0291-x

Yakopcic, C., Taha, T. M., Subramanyam, G., Pino, R. E., and Rogers, S. (2011). A memristor device model. *IEEE Electron Device Lett.* 32, 1436–1438. doi: 10.1109/LED.2011.2163292