

阻变存储器：从物理机理到集成应用

刘琦^{1,2,3*}, 陈超^{1,2,3}, 陈沛^{1,2,3}, 张续猛^{1,2,3}

(1. 复旦大学集成芯片与系统全国重点实验室, 上海 200433; 2. 复旦大学集成电路与微纳电子创新学院, 上海 200433; 3. 复旦大学芯片与系统前沿技术研究院, 上海 200433)

摘要:阻变存储器(resistive random-access memory, RRAM)作为最具潜力的新型非易失性存储技术之一,因其结构简单、功耗低、速度快及良好的可扩展性,正成为突破传统“存算分离”架构下数据搬运瓶颈的重要方向。然而,其大规模应用仍受限于开关均匀性、循环耐久性及集成可靠性等挑战。本文系统综述了RRAM在机理解析、性能调控、工艺集成与系统应用方面的研究进展。从阻变机制出发,总结了工艺优化及电学编程策略改善器件一致性和可靠性的关键方法;在集成层面,系统回顾了RRAM在先进工艺节点下与互补金属氧化物半导体工艺兼容性的探索,以及高密度三维集成方面的最新进展;在应用方面,重点分析了RRAM在高能效存算一体、类脑计算、智能感知及安全芯片中的发展趋势。面向未来,本文指出需在机理-材料-架构层面开展跨尺度协同创新,以支撑智能计算与信息技术融合发展中的战略目标。

关键词:阻变存储器;物理机理;性能调控;三维集成;存内计算;神经形态计算

中图分类号:TN43 **文献标志码:**A **文章编号:**1001-2486(2026)02-331-18

Resistive random-access memory: from physical mechanisms to integration and applications

LIU Qi^{1,2,3*}, CHEN Chao^{1,2,3}, CHEN Pei^{1,2,3}, ZHANG Xumeng^{1,2,3}

(1. State Key Laboratory of Integrated Chips and Systems, Fudan University, Shanghai 200433, China;
2. College of Integrated Circuits and Micro-Nano Electronics, Fudan University, Shanghai 200433, China;
3. Frontier Institute of Chip and System, Fudan University, Shanghai 200433, China)

Abstract: RRAM (resistive random-access memory) has emerged as a promising non-volatile memory technology due to its simple device structure, low power consumption, fast switching speed, and excellent scalability, addressing the data movement bottleneck in traditional compute-memory separation architectures. However, challenges in switching uniformity, cycling endurance, and integration reliability hinder its widespread adoption. This review systematically examined recent advances in RRAM, covering mechanism analysis, performance modulation, process integration technologies, and innovative applications. Starting from resistive switching mechanisms, key approaches based on process optimization and electrical programming strategies were summarized to enhance device uniformity and reliability. At the integration level, recent advances in CMOS (complementary metal-oxide-semiconductor) compatibility at advanced technology nodes and high-density 3D (three-dimensional) integration of RRAM were systematically reviewed. In terms of applications, the development trends of RRAM in high-energy-efficiency in-memory computing, neuromorphic computing, intelligent sensing, and secure chips were analyzed in detail. Towards the future, synergetic cross-scale innovation spanning mechanism, material, and architectural levels were emphasized, supporting the strategic goals of the integrated development of intelligent computing and information technologies.

Keywords: memristor; physical mechanism; performance modulation; 3D integration; in-memory computing; neuromorphic computing

在数字经济与信息技术高速迭代的今天,数据已成为核心生产要素,全球数据量正以每三年翻倍的速度爆发式增长^[1],人工智能、物联网、自动驾驶、云计算等新兴领域的崛起,对数据的存

储、访问与处理效率提出了前所未有的要求。存储器作为承载数据的核心硬件载体,是信息技术产业链的基石,其性能(容量、速度、功耗)直接决定了终端设备、数据中心乃至国家算力体系的运

收稿日期:2025-12-24

基金项目:国家自然科学基金资助项目(T2293732, 62374040)

*第一作者:刘琦(1980—),男,安徽宿州人,教授,博士,博士生导师,E-mail:qi_liu@fudan.edu.cn

引用格式:刘琦,陈超,陈沛,等.阻变存储器:从物理机理到集成应用[J].国防科技大学学报,2026,48(2):331-348.

Citation: LIU Q, CHEN C, CHEN P, et al. Resistive random-access memory: from physical mechanisms to integration and applications[J].

Journal of National University of Defense Technology, 2026, 48(2): 331-348.

行效率。从国家战略层面看,高性能存储器技术的自主可控是保障信息安全、突破半导体产业“卡脖子”瓶颈的关键环节,全球主要经济体均将新型存储技术研发列为重点发展方向,推动其成为科技竞争的核心赛道之一^[2-3]。无论是个人消费电子、工业互联网,还是前沿科学计算、国家安全领域,存储器的技术迭代与创新都扮演着不可替代的支撑角色,成为衡量一个国家信息技术核心竞争力的重要标志。

传统计算机体系基于冯·诺依曼架构,其核心特征是“存储单元与运算单元分离”,数据需在内存(如 DRAM)与处理器(CPU/GPU)之间频繁搬运才能完成运算。随着数据量呈指数级增长,这种“存算分离”模式引发了严重的“内存墙”与“功耗墙”问题:一方面,数据搬运过程产生的延迟占比已超过运算本身,导致系统整体效率受限;另一方面,数据迁移消耗的能耗占比过高,与低碳化发展需求相悖。

面向未来算力需求,新型存储技术需满足“高容量、高速度、低功耗、非易失性、长寿命、可扩展、存算一体兼容”的核心诉求:既要突破传统存储的性能瓶颈,又要适配“存算一体”架构的发展趋势,通过将运算功能融入存储单元,减少数据搬运,从根源上解决延迟与功耗问题。这种需求驱动下,阻变存储器(resistive random-access memory, RRAM)、相变存储器(phase-change memory, PCM)、磁随机存储器(magnetoresistive random-access memory, MRAM)等新型非易失性存储技术应运而生,成为突破冯·诺依曼架构局限、构建下一代高性能计算系统的关键候选方案。图 1 总结了这些技术在关键性能指标上的对比^[4-6]。

其中,PCM 具备良好的非易失性保持特性与集成密度,可用于构建大规模长期存储阵列。然而,其基于热致相变的写入机制通常需要较高的写入能量,并伴随热管理与热串扰问题,限制了其在低功耗系统中的应用。MRAM 通过磁隧道结结构实现高速、低能耗写入并具备优异耐久性,在嵌入式存储与缓存场景中具有显著优势。然而,MRAM 器件结构相对复杂,且可区分状态数有限,使得其在高密度大规模阵列集成及多值表达方面存在严峻的挑战。相较而言,RRAM 在单元面积、读写速度与多值存储等方面具有综合优势。特别地,RRAM 制备与后道互补金属氧化物半导体(complementary metal-oxide-semiconductor, CMOS)工艺兼容,是先进节

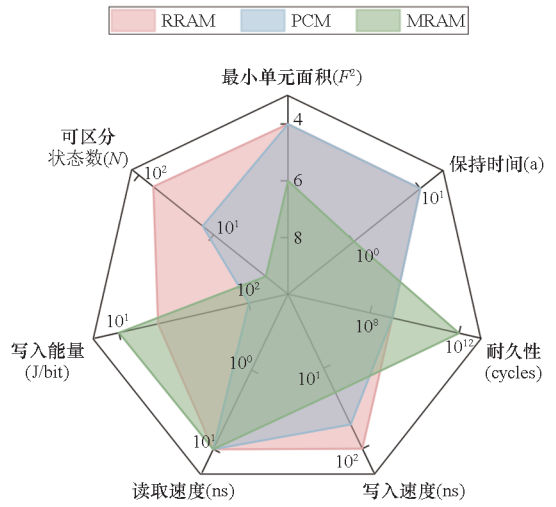


图 1 RRAM、PCM 与 MRAM 的性能对比^[4-6]
Fig. 1 Performance comparison of RRAM, PCM, and MRAM^[4-6]

点下嵌入式非易失性存储器的有利候选者,且其交叉阵列结构天然适配并行的向量矩阵乘加计算,受到产业界的广泛关注并大力推进产业应用。因此,本文将重点围绕 RRAM 展开讨论,系统分析其材料与器件机制、性能优化、阵列集成及其应用潜力。

RRAM 的核心原理基于功能材料在电场作用下的电阻态可逆转变,如高阻态(high-resistance state, HRS)与低阻态(low-resistance state, LRS),以实现数据存储。其物理现象可追溯至 20 世纪 60 年代的早期阻变研究,而理论基石则源于 1971 年蔡少棠教授预言的忆阻器——作为除电阻、电容、电感之外的第四种基本电路元件,其电阻值可随历史电荷量变化,为 RRAM 提供了关键的理论框架^[7]。2008 年,惠普实验室首次在 TiO₂ 薄膜器件中明确实现了与忆阻器理论对等的物理原型^[8],引发全球科研界与产业界的广泛关注。此后 RRAM 技术进入快速发展期,在材料体系(从氧化物到硫化物、有机材料)、器件结构(垂直堆叠、交叉阵列)、性能优化(循环寿命、数据保持性)等方面取得一系列突破。

目前,RRAM 已成为忆阻器的典型代表,二者在广义上常被等同视之,本文统一采用 RRAM 的称谓。尽管其实验研究展现出巨大潜力,实际应用仍面临开关速度、循环耐久性及器件间一致性等多重性能与可靠性挑战。这些挑战导致系统能效下降与计算精度损失,严重制约了 RRAM 在高密度集成与复杂计算任务中的大规模部署。因此,系统性地解决性能与可靠性问题,已成为推动 RRAM 技术从实验室走向商业

化应用的核心。

本文采用从物理机制、性能调控、阵列集成到应用拓展的分类框架,如图 2 所示,系统综述 RRAM 领域的进展,探究其电阻转变与导电通路的微观机制,分析器件性能优化策略、高密度存储

实现方案及相关应用路径。通过整合理论与实验成果,本文进一步讨论了当前在器件变异性、串扰管理和可靠性退化等方面的挑战,提出了跨物理机制-材料设计-系统架构的协同优化路径,为 RRAM 的开发应用提供理论指导。

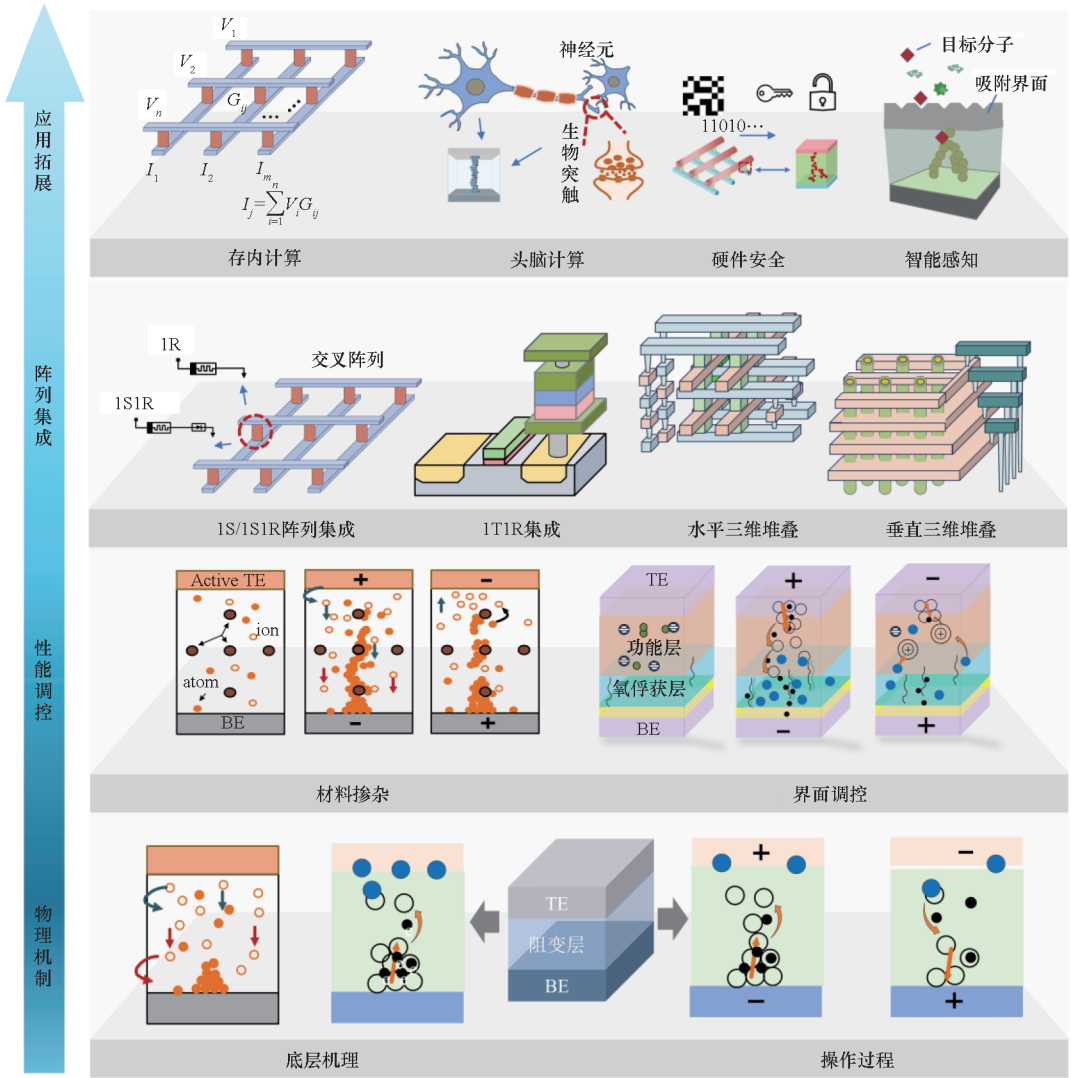


图 2 阻变随机存储器:物理机制、集成与应用

Fig. 2 Schematic of RRAM: physical mechanism, integration, and applications

1 阻变机理与导电通路动力学

本节以导电通路作为器件阻变行为的核心物理主线,围绕阻变机制分类、通路形成与断裂动力学、量子尺度输运行为三个层级展开系统讨论。通过多尺度递进分析,本节阐明了阻变效应的物理本质,并总结了器件在不同尺度下的关键调控规律。

1.1 阻变机制分类及导电通路起源

RRAM 的电阻转变行为主要由三种物理机制驱动:电化学金属化 (electrochemical metallization, ECM)、价变机制 (valence change mechanism,

VCM) 和热化学机制 (thermochemical mechanism, TCM)。这些机制分别基于金属离子迁移、氧空位行为或相变等物理原理。其中,ECM 和 VCM 是当前主流机制。如图 3(a) 所示,在 ECM 机制中,活性电极 (例如 Ag 或 Cu) 发生氧化还原反应,并通过离子迁移形成导电细丝,采用该机制的器件又被称为导电桥 RRAM (conductive-bridge random-access memory, CBRAM)。与之相比,如图 3(b) 所示的 VCM 的核心在于氧离子迁移与氧化物价态变化共同形成氧空位导电通道,这类器件亦称为氧化物阻变存储器 (oxide-based resistive random-access memory, OxRAM)。此外,

TCM 主要借助热诱导相变实现开关,例如 VO₂ 基器件可以实现高速阈值响应,但其固有的易失性不利于数据长期保持^[9]。本文将重点讨论 ECM 和 VCM 这两种主流机制。从动力学角度看,ECM 的速率控制通常由阳离子迁移与阴极还原沉积的耦合决定,而 VCM 的速率控制更常取决于氧空位/氧离子扩散与界面氧交换,因此两类器件在随机性与一致性上的差异本质源于“成核位置稳定性”与“迁移通道可重复性”的不同。ECM 器件的优势在于开关窗口大、功耗低,但一致性和稳定性通常较差;而 VCM 器件则在一致性和耐久性方面表现更优,但开关窗口相对较小。总体而言,不同机制的 RRAM 在开关速度、耐久性、集成潜力等关键性能上各具特点,在实际应用中需根据需求权衡。

持特性相对较差。

综上所述,RRAM 中不同的阻变机制及相应的材料策略共同决定了其在转变速度、循环耐久性、器件间变异性和数据保持能力等关键性能指标的表现。因此,理解主导机制背后的物理原理及其相互作用,是优化器件和指导设计的关键。

借助先进的理论与计算手段,从原子/微观尺度出发,有助于深入理解电阻转变行为背后的物理起源。第一性原理计算、量子模型及多尺度数值模拟,已成为理解 RRAM 动态行为、预测新现象和分析变异性的关键工具。以最常见的 VCM 器件为例,其多尺度建模框架通常整合了从原子尺度(第一性原理计算缺陷性质)到介观尺度(动力学蒙特卡罗模拟导电细丝的演化)再到器件尺度(紧凑模型用于电路仿真)三个层次^[10-12]。这些方法为定量理解氧空位迁移、金属离子还原等过程提供了坚实的理论支撑。

1.2 导电通路的形成、演化与断裂动力学

导电通路的形成与断裂是 RRAM 电阻开关的核心物理过程,主要由氧空位和金属阳离子的迁移驱动。在此过程中,材料微观结构、电场和热效应分别从不同维度对通路演化过程进行调控:微观结构“筛选”迁移通道(如晶界、界面等缺陷),电场驱动离子漂移成核,热效应限定扩散动力学与热失稳边界。三者共同作用,最终决定了导电细丝的位置和状态,并形成器件间与周期非均一性的来源。这些机制受材料微观结构、电场和热效应调控,并直接决定开关均匀性、多值存储能力和器件可靠性。本节将系统分析氧空位与金属阳离子的动态过程。

氧空位迁移作为 OxRAM 中导电细丝形成的主导机制,其动力学受外部电场和材料微观结构(如晶界、缺陷密度)调控,直接决定形成电压、转变均匀性及多值能力等关键电学特性。在 HfO_x、TaO_x 等材料体系中,氧空位浓度和分布的精确控制可实现高开关比和多值行为。因此,深入理解氧空位迁移机制是优化 RRAM 性能的基础。如图 4 所示,RRAM 的形成电导丝(SET)过程为在外加电场驱动下,氧空位定向迁移并聚集,在电极间形成连续导电通路,使器件从 HRS 转为 LRS,实现信息“写入”。其细丝断裂(RESET)过程则是通过反向电场或焦耳热效应,促使氧空位与氧离子重新复合,导致导电细丝断裂/退化,器件从 LRS 恢复至 HRS,完成信息“擦除”。需要强调的是,RESET 操作通常并非意味着导电通路的完全消失,实则表现为局域通路的变细和断裂,这种变

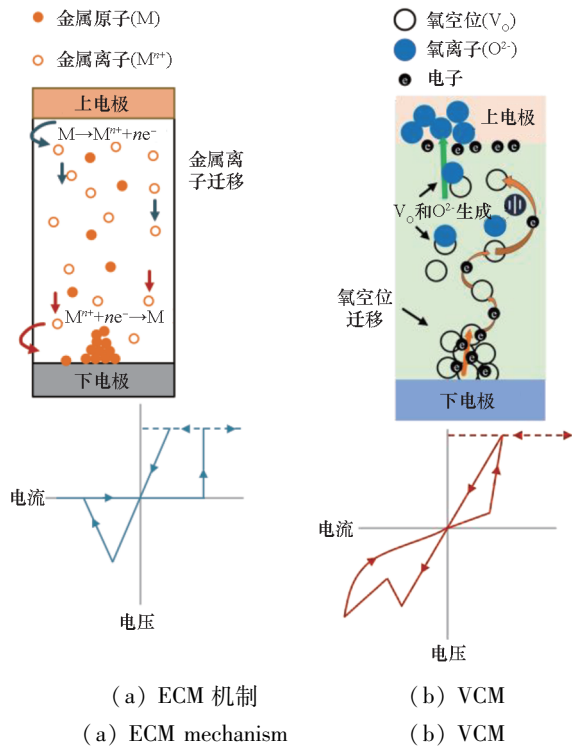


图 3 RRAM 两种典型工作机制

Fig. 3 Two classical working mechanisms of RRAM

上述阻变机制的核心,是介质层中导电细丝的形成或断裂所导致的器件整体电阻变化。此外,界面型阻变机制也至关重要。界面机制的核心是肖特基势垒的调制和界面处电荷俘获/释放过程。具体而言,界面陷阱占据会改变局域电场并造成能带弯曲。当陷阱占据在脉冲激励下呈渐变演化时,器件可自然呈现连续(模拟)电导调节,但在无偏置条件下易发生由陷阱占据导致的保持退化。界面电荷俘获机制可实现连续、模拟的电阻转换,并展现出较小的器件间差异性,但保

化多集中于电流和温度集中的区域。正是这种导电路径从连续向间断/隧穿模式的转变,器件 RESET 后的 HRS 分布会随电流大小以及散热条件的变化而显著改变。

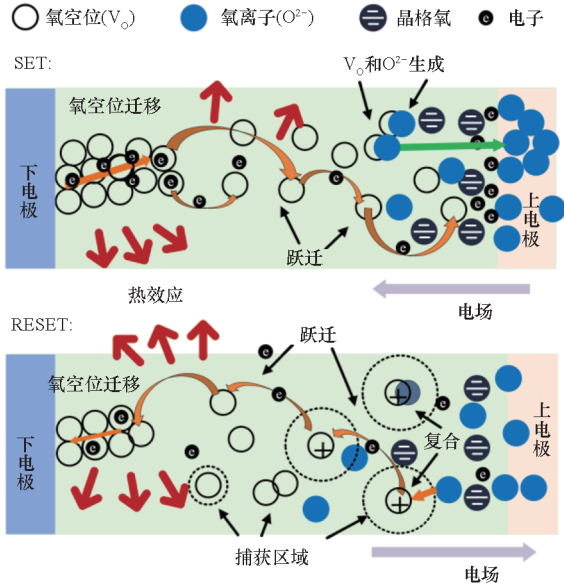


图4 RRAM的SET与RESET阻态切换过程

Fig. 4 RRAM's SET and RESET resistance state switching

金属阳离子(如 Ag^+ 、 Cu^{2+})的迁移是除氧空位机制外的另一核心物理过程,该机制通过电化学方式形成金属导电细丝,实现电阻转变。在外加电场作用下,金属离子经历氧化、迁移、还原与沉积的循环,可逆地形成或断裂金属导电细丝,从而实现器件在低阻态与高阻态的切换^[13]。

此外,两种机制还可以协同工作,形成双通路系统,有助于实现更低操作电压和更高开关比。因此,离子与氧空位的迁移动力学及其所形成的微观结构,是理解与调控器件宏观性能的关键。下文将从宏观动力学与材料结构两个维度展开分析:首先分析电场与温度的宏观动力学影响,继而探讨在不同材料结构下导电通路演化的独特行为及其对应的器件功能。

电场与温度是最关键的两个外部参数,它们通过调控离子与氧空位的迁移动力学,深刻影响RRAM器件的开关特性、可靠性及能耗。例如,高电导率促进氧空位生成从而降低形成电压;而高热导率增强散热会提高形成电压^[14]。这一现象的物理本质在于,形成电压对应着缺陷临界连通所需的能量阈值。高电导率通过提升局域电流与能量聚集,降低了这一阈值;而高热导率则通过消散热量、打断热-缺陷正反馈,提高了这一阈值。因此两者对形成门槛呈相反影响。

电场作为直接驱动力,主导了离子的定向迁

移,可以实现对离子迁移路径的精确调控与能耗优化。温度则通过改变热激活能,影响迁移速率与机制。高温下,增强的热激活过程会加剧离子运动,导致电阻态的弛豫与退化;而在低温下,热激活作用减弱,表现为转变电压升高,转变行为也从连续的渐变式切换转变为突变的数字式切换^[15]。此外,由焦耳热效应引发的局部纳米尺度热分布与电场协同作用,是驱动导电细丝成核与生长的关键因素。该协同的核心是一个自增强的热-电正反馈:通过局部温升降低势垒,从而加速导电通路的形成和器件电导增加。电导提升进一步引发电流在局域区域内集中,从而增加焦耳热并加剧局部热积累,形成持续强化的正反馈循环。因此引入电流限制或热扩散层的本质目的在于切断该正反馈循环,以提升器件操作的可控性与耐久性。电场与温度的交互作用,共同定义了器件稳定工作的功能边界,是RRAM应用中动态性能调控的关键。

除宏观的动力学规律以外,材料的结构特性从根本上主导导电通路的形成与演化,这使得材料工程成为实现器件功能定制化的核心途径。在有序的单晶材料体系中,规整的晶格为导电细丝提供了可预测的迁移路径,易于实现快速、低随机性的多值转变特性,适用于高精度计算。在高离子电导材料中,高效的氧空位迁移与聚集能够模拟生物突触中复杂的动态过程,从而支持丰富的神经形态行为^[16]。而在二维范德瓦尔斯异质结构及相变系统中,界面效应和限域环境为调控导电细丝行为提供了新的研究空间,其迁移可诱导局域结构相变,并增强器件在循环过程中的热稳定性与耐久性^[17-18]。这些基于材料学的探索表明,对导电通路的理解与操控,是实现高性能、多功能计算与存储应用的关键。

1.3 量子尺度下的阻变物理机理

当阻变器件尺寸缩减至原子/量子尺度时,量子效应从次要影响因素转变为主导阻变行为的核心物理根源,其阻变机理直接源于量子力学层面的电荷传输、粒子耦合及低维约束效应。因此,从量子尺度重新审视阻变机理,为理解器件微观随机性,支撑超高密度与新型计算架构发展提供了前沿理论支撑。

首先,导电细丝在原子尺度展现出量子化特征,其电导以量子电导 G_0 的整数或半整数倍变化。量子化台阶意味着电导由少数原子级通道主导,其导电能力主要取决于最狭窄区域的原子排列与键合状态。此时的电导波动往往来源于单原

子附着或脱附等离散事件,而非连续几何尺寸变化。其次,电子-离子耦合(特别是电子反冲力)是驱动细丝演化的关键微观机制,这已得到理论预测和实验观察(如透射电镜原位研究)的共同支持^[19]。在高电流密度下,电子动量传递可对离子/缺陷产生有效驱动力,使细丝中的狭窄区域处更易发生原子迁移与重排;这类非平衡驱动会放大周期间波动,并与热扩散共同决定细丝的形成与断裂路径。在低维材料中,第一性原理计算揭示量子限制效应会诱导高开关比和循环可变性^[20-21]。最后,非平衡态下的热电子传输和界面散射,也是造成器件参数变异和耐久性退化的微观根源之一。这些现象凸显了量子与纳米尺度传输在理解器件性能变异中的核心作用。

综上,量子尺度下的阻变行为本质是量子效应主导的多物理场耦合过程。这些量子机理的揭示为突破 RRAM 器件性能瓶颈提供了全新的理论视角与技术路径。

1.4 小结

本节围绕阻变存储器的物理本质,从电阻转变机制分类、导电通路的形成与演化以及量子尺度输运现象等多个层面,系统梳理了 RRAM 电阻切换行为的微观起源。导电细丝的形成与断裂过程具有显著的动态演化特征与随机性,是器件非均一性与可靠性问题的核心物理根源。基于 ECM 机制与 VCM 在导电通路的构建方式、开关过程稳定性以及工艺集成潜力等方面各具特点,但均不可避免地会受到细丝演化过程的统计波动制约。随着器件尺寸进一步微缩至纳米尺度,量子化输运与电子-离子耦合效应导致了更复杂的阻变过程,并进一步加剧了器件的非均一性。尽管多尺度建模与原位表征技术的发展显著深化了对阻变机理的理解,但在低功耗约束下如何有效抑制导电通路随机性、提升阻变过程的可控性,仍是未来器件工程优化与系统级集成亟须突破的关键科学问题。

2 性能调控与可靠性优化

上一节通过分析 RRAM 的阻变机理,为性能调控提供了理论支撑。针对实际应用中面临的稳定性、均匀性等关键问题,本节从材料掺杂、界面工程及电学编程方案优化三个方面,系统探讨 RRAM 性能调控与可靠性优化策略。

2.1 材料掺杂

材料掺杂通过精准调制缺陷浓度、离子迁移

动力学和界面特性,已成为优化 RRAM 开关均匀性、多值存储能力和神经形态功能的通用手段。本节将依次讨论过渡金属氧化物掺杂、材料缺陷工程、有机-无机杂化材料成分优化以及合金化策略,系统阐述这些方法在多种材料体系中的性能增强效果。

对于过渡金属氧化物 RRAM,金属和氮掺杂通过精确调控氧空位的形成与迁移,已成为优化开关参数、实现可靠多值操作及提升循环耐久性的关键策略。该策略已在多种材料体系中得到验证。如图 5 所示,金属掺杂形成的纳米晶体与扩散的金属原子能引导导电通路生长路径并增强局部电场,从而降低形成难度。该机制的本质是:掺杂相/纳米晶体通过构建介电常数或导电率的不均匀分布,实现局域电场集中;与此同时,掺杂引入的化学键环境可调控氧空位形成能,促使成核过程固定于掺杂富集区,进而抑制随机成核引发的形成电压波动。具体而言,电负性金属(如 Au、Pt 和 Pd 掺杂 Al_2O_3)可促进氧空位成核,从而提升转变重复性与成品率^[22];对 HfO_2 进行氮掺杂,能形成稳定团簇以增强离子电导率与开关比,同时改善操作电压的均匀性^[23];采用 Al、N 等多元素共掺 TaO_x ,则可协同优化界面阻变行为,获得高耐久性与优异的多值保持能力^[24]。尽管掺杂元素存在差异,但上述结果可统一归纳为“缺陷工程”机制:通过调控氧空位的形成与聚集,使缺陷从随机分布状态转变为受控聚集状态,进而同步降低器件的形成电压并优化周期间可重复性。除了传统氧化物,二维材料中的缺陷工程为实现超低变异、高均匀的转变行为提供了新途径。例如,Ag 掺杂的 MXene 能够有效降低操作参数变异,提升器件可靠性^[25]。进一步地,通过精密的分子设计优化有机-无机杂化材料的成分,能获得双模开关、低压操作等特性,这拓展了其在柔性神经形态电子与传感内计算系统中的应用前景。

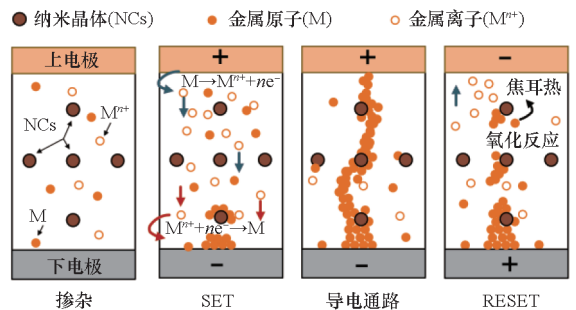


图 5 掺杂优化导电细丝演化的过程
Fig. 5 Process of doping for optimized conductive filament evolution

最后,合金化策略通过稳定导电细丝并调制电化学过程,能显著提升RRAM的转变均匀性、数据保持能力及操作对称性。例如,高熵合金通过高熵效应有效稳定导电通路,实现低操作电压和高循环耐久性^[26-27]。然而,合金化也可能引入杂质变异,因此纯度控制对稳定性至关重要。此外,自组装量子点、有序空位半导体等协同成分调控策略,可通过引导细丝生长等机制进一步降低阈值电压并均匀化转变参数。

2.2 界面工程

界面调控通过优化电极-介质接触改变势垒高度与载流子注入效率,并调控导电细丝的形成与稳定,从而提升RRAM的开关均匀性、耐久性,并实现低功耗操作。界面层的作用机制通常可概括为三条路径:一是调节肖特基势垒高度与陷阱态密度,改变电子注入机制;二是作为储氧层或阻氧层调控界面处氧空位的动力学行为,进而调制氧空位的浓度梯度分布;三是通过几何或化学限域作用,实现导电通路的成核位置由随机分布向特定区域收敛。例如,TiO₂层等能优化界面能带并降低陷阱密度^[28-29];Al₂O₃和SiO₂层则能调控氧离子迁移以提升开关比和耐久性^[30-31]。此外,TiN纳米晶体界面则通过优化导电细丝操作来改善开关均匀性^[32]。总之,这些界面层通过多维度的缺陷工程,为制备高性能、高均匀性的RRAM器件奠定了关键材料基础。

主要的界面优化策略包括电极材料选择、形貌工程与肖特基势垒调制。这些策略旨在通过调控电场分布、限制导电路径来改善开关电压、开关比及可靠性。首先,电极材料选择直接决定界面动力学。例如,在MXene基器件中采用钨电极,可实现低开关电压与高开关比^[33]。其次,形貌工程通过构建粗糙表面或纳米结构来集中电场、引导细丝,从而降低开关电压。最后,肖特基势垒调制利用氧空位/离子迁移动态重构界面势垒,这不仅可模拟电阻态与自整流特性,还能提升循环稳定性。

为了进一步约束导电细丝在介质内部的随机行为,研究者将目光从单一界面转向了构建更为复杂的多层异质结构。多层异质结构利用更复杂的界面工程调控氧空位迁移与导电通路动态演化,从而显著提升RRAM的开关均匀性,更有利于多态存储和计算。如图6所示,氧俘获层促使氧离子从功能区中移出,更容易形成导电细丝。Ta₂O₅/HfO₂双层结构^[34]、超晶格结构^[35]以及二

维材料异质界面^[36]等结构,均通过界面机制实现了低变异的开关行为。

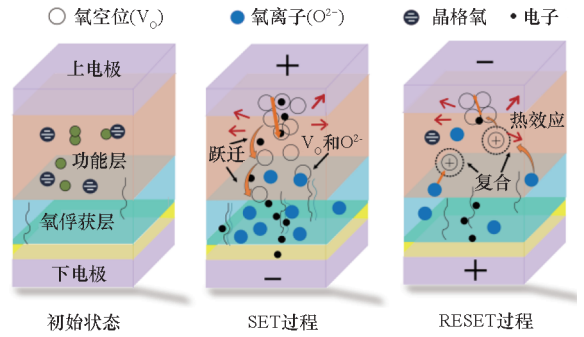


图6 界面调控下的阻变过程

Fig.6 Resistance change process under interface regulation

此外,在电极和主介质层之间插入辅助缓冲层,通过调控细丝路径和氧空位分布,已成为提升RRAM开关一致性与计算可靠性的另一项核心策略。Al₂O₃/AlN叠层^[30]、TiSbTe层^[37]和MoS₂层^[38]等缓冲层,通过稳定细丝路径、产生阻氧效应或空间限制细丝位置等机制降低性能波动。

尽管上述策略共同推动RRAM在速度、功耗与可靠性方面的进展,但其实际应用仍面临制造复杂性、氧空位随机分布等挑战,仍需在性能与工艺之间进行权衡。

2.3 电学编程方案优化

脉冲编程是调控RRAM的关键电学手段,通过精确调控电压/电流脉冲的宽度、幅度、频率和形状等关键参数,直接操纵阻变器件中的离子迁移、电荷捕获与导电细丝动力学^[39]。因此,该技术为实现高性能存算(如多值存储、模拟突触可塑性)提供了关键的实现途径。本节将系统阐述如何通过脉冲参数的协同优化,平衡多级状态精度、操作能效与长期可靠性之间的矛盾。

脉冲编程的首要目标是实现高精度、可控的多级电导状态。脉冲参数直接决定了氧空位迁移和细丝构型的动力学过程,进而影响电导状态的线性度、精度与噪声。例如,超快脉冲策略通过精细控制细丝生长尺寸,可在器件中实现数百个非易失性电导水平,为高密度存储和低功耗加速器奠定基础^[39]。然而,过快的操作会因离子迁移固有的随机性而引入显著的参数涨落,从而损害电导状态的稳定性。因此,为确保状态稳定,需要引入额外的调控机制。例如,施加电流限制可使器件实现多值存储和高精度突触响应;而采用特定波形设计(如三角波脉冲)则能平缓调控缺陷分布,有效抑制状态写入过程中的过冲与不稳定性,

进而改善人工神经网络的识别准确率^[40]。这些策略共同表明,通过协同优化脉冲序列,可以在不大幅牺牲速度的前提下,显著提升多值的可编程性与精度。

在追求高精度与低功耗的同时,确保电导状态的长期稳定性(如抑制弛豫、漂移)是脉冲编程面临的另一关键挑战。研究表明,写入验证、迭代写入等定制化的脉冲方案能抑制细丝过度弛豫,提升多值存储能力。更进一步,当脉冲策略与材料工程(如通过嵌入纳米颗粒以钉扎氧空位)结合时,能协同增强器件的耐久性和数据保持能力^[41]。然而,稳定性提升往往需要付出其他性能代价。例如,采用更为平缓的脉冲策略,通常会导致操作电压升高或开关时间延长,这凸显了开关速度与数据保持能力之间固有的权衡关系。因此,脉冲编程方案的设计需以应用为导向,在速度、能耗与可靠性之间做出针对性的权衡。

综上所述,脉冲编程方案通过精确调控阻变动力学,为实现高性能阻变存储器提供了灵活而强大的技术路径。未来的优化方向将超越单一的脉冲参数调整,迈向脉冲-材料-算法的跨层级协同设计。例如,将界面工程赋予的高开关比和多值能力^[42-43]与自适应脉冲算法结合,有望在系统层面实现动态重构与能效优化,最终突破单一技术路径的性能瓶颈。

2.4 小结

材料掺杂、界面工程与电学编程等策略,虽在实验层面为 RRAM 带来了开关均匀性、循环耐久性与多值稳定性的显著提升。然而,这些性能增益往往以制造工艺复杂化、工艺窗口收窄以及系统集成与实现难度提升为代价。综合来看,RRAM 的性能优化绝非单一指标的突破,而是一个涉及多维度权衡的系统工程。其根本挑战在于如何平衡器件性能提升与规模制造可行性的固有矛盾。

综上,RRAM 的性能优化本质上涉及多维指标之间的权衡,不同调控策略在速度、功耗、可靠性和制造复杂度方面各具优势,其适用性高度依赖于目标应用需求。例如,对于高精度存内计算和神经形态计算,通过适度增加工艺复杂度来换取稳定性与可编程性,往往是可接受的;但在大容量嵌入式存储及规模化集成场景中,工艺简洁性和良率应优先于单器件性能提升。因此,未来 RRAM 的发展重点将不再是孤立追求单一性能指标,而是通过材料、器件、编程与系统架构的协同

设计,在性能与制造可行性之间构建可持续的工程平衡。

3 阵列集成技术

RRAM 技术的实际落地,主要受限于阵列集成密度、互连效率、工艺兼容性及系统扩展性等核心问题。因此,阵列集成技术已成为学术界与产业界的共同焦点。本节将系统梳理 RRAM 阵列集成技术的最新研究进展,并剖析其面临的关键挑战。

3.1 交叉阵列设计

交叉阵列是 RRAM 阵列集成中最具代表性的核心架构。其拓扑简单、集成密度高,且与存算一体系统天然兼容,因而成为突破集成瓶颈的关键研究方向。典型的交叉阵列设计如图 7 所示,RRAM 被集成于行/列电极的交叉点,形成简单的“1R”结构。尽管该结构有利于高密度集成,但其存在严重的旁路电流问题,如图 8 所示。具体而言,读写操作时电流会通过同一行/列的非目标单元形成漏电路径,导致误操作、能耗增加与可靠性劣化。此旁路电流问题是阻碍 1R 阵列实用化的关键瓶颈。为抑制旁路电流、提升阵列性能与存算适用性,研究者们围绕器件结构提出并发展多种抑制方案,见表 1,包括通过器件非线性优化的 1R 自整流结构、集成选择器的 1S1R 结构以及晶体管选通的 1T1R 单元结构等^[44-45]。

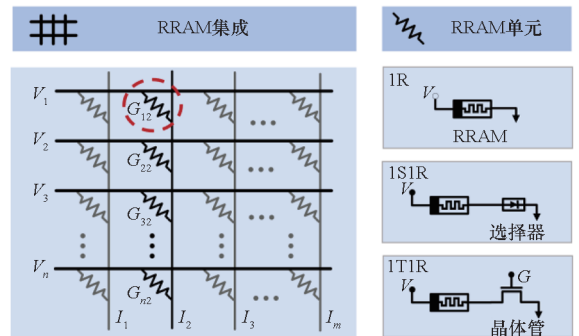


图 7 RRAM 单元结构和阵列

Fig. 7 RRAM cell structure and array

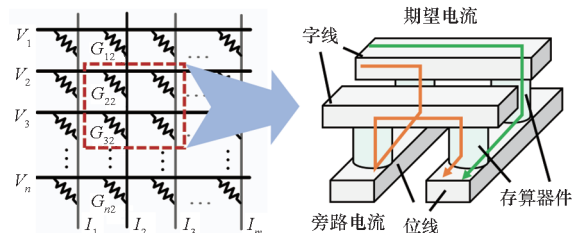


图 8 RRAM 阵列中的旁路电流

Fig. 8 Sneak current in RRAM arrays

表1 不同 RRAM 阵列集成方案的技术对比^[44-45]Tab.1 Technical comparison of different RRAM array integration schemes^[44-45]

方案	密度	串扰抑制	三维集成能力
1R 自整流结构	高($4F^2/n$)	中	平面、垂直
1S1R	高($4F^2/n$)	中	平面
1T1R	中($6F^2$)	高	不支持

注: n 表示三维堆叠的层数, F 为特征尺寸。

1R 方案的关键在于通过 RRAM 器件自身的非线性和整流特性实现选通功能,在实现高密度存储方面展现出巨大潜力。这类器件的物理实现通常依赖势垒工程、界面调控与材料缺陷调节等机制。具体而言,通过势垒工程可在金属/氧化物界面引入整流行为,从而在实现高整流比的同时降低器件操作电流^[31,46];通过基于材料掺杂优化氧空位浓度分布与界面处的肖特基势垒,从而提升导电通路形成/断裂过程的可控性^[24,47]。在阵列应用中,1R 方案可在无须额外选通器件的前提下有效降低旁路电流,提升阵列能效与可靠性。然而,该方案面临器件非线性能力不足及均一性问题,在大规模阵列下难以满足高精度存算对稳定性与可控性的要求。

与之相比,1S1R 在有效抑制旁路电流的同时仍可保持较高的阵列集成密度。通过串联阈值型选择器,1S1R 结构能够更有效地降低阵列中的旁路漏电,从而提升读写窗口与可靠性,并增强存算一体计算的可扩展性。因此,在 1S1R 结构中,优化选择器性能成为更有效抑制旁路电流、提升存储密度与能效的关键。通过材料掺杂、界面工程等策略,可系统调控选择器的阈值电压、非线性度、开关速度及循环寿命。这对增强阵列均匀性与能效至关重要。例如,材料掺杂(Si、Te)可优化开关特性^[48-49];界面工程(缺陷石墨烯或等离子体处理)能增强细丝控制与耐久性^[50-51];动力学调制(Ag 细丝调控)可调节阈值与寿命^[52-53]。这些策略不仅提升了选择器性能,还有助于抑制旁路电流,为 RRAM 大规模集成提供了可靠支持。然而,1S1R 单元中的选择器需同时满足低阈值波动、高开关速度、高耐久与低功耗等多目标要求,在系统级扩展时仍面临挑战。

与 1S1R 相比,1T1R 结构利用晶体管提供强选通能力,可实现更优的单元隔离与写入控制,通常具有更高的开关比、更强的均一性以及更好的写入可重复性,在高可靠存储阵列与工艺可控的嵌入式应用中具有明显优势。然而,1T1R 的主要

限制在于晶体管引入显著面积开销,使单元尺寸难以达到交叉阵列的极限密度,且外围选通电路会提高系统复杂度与功耗。其结构层面的约束决定了其更适合对一致性和稳定性要求更高、而对密度要求相对次要的应用场景。

3.2 三维高密度集成

追求更高的面积利用率,驱动着研究重心从平面交叉阵列转向三维集成架构。三维集成不仅通过立体堆叠显著提升了存储密度,也通过增加数据并行处理通道增强了处理效率。该架构采用多层交叉阵列与垂直单元,在芯片面积和处理延迟两方面相比二维方法均有显著改善^[54-55]。目前主流的三维集成技术可分为水平堆叠与垂直堆叠两类,如图 9 所示。水平堆叠是在垂直方向层叠多个二维阵列;垂直堆叠则是先构建一个贯穿多层的立体结构(如柱状电极),再围绕其形成各层忆阻单元。这两种堆叠方式各有优势,但高密度集成也带来了串扰电流和制造复杂化等新挑战。对此,需进一步优化 RRAM 器件以抑制串扰,并优化可堆叠开关以保障可靠性。总体而言,三维堆叠是通向高密度存算的必经之路,但其成败关键在于材料与集成工艺的突破。

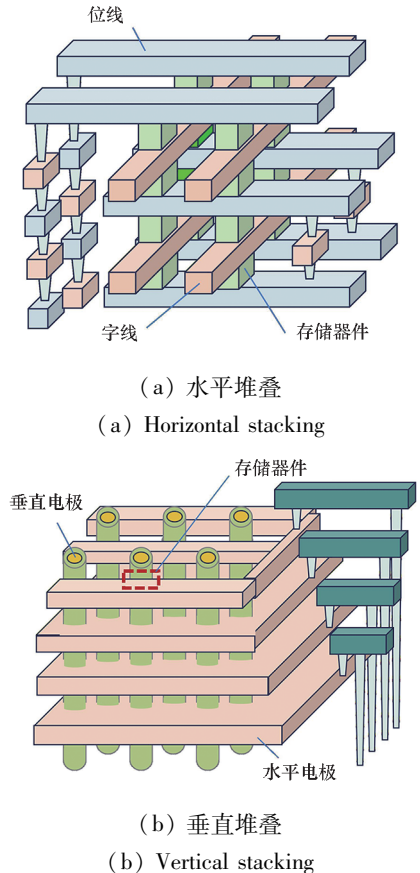


图9 三维堆叠 RRAM 结构

Fig.9 Structure of the 3D stacking RRAM

除了可扩展性,功耗与热管理也是三维集成存内计算系统面临的核心问题。在三维集成存内计算系统中,功耗优化和热管理对于抑制热积累、保障系统稳定运行至关重要。模拟存内计算集成本身的设计复杂性使得三维集成系统的功耗与热管理问题尤为突出。功耗优化旨在提升能效与加强热控制来保障系统性能与可靠性。但这往往需要在精度、制造可行性及架构灵活性之间进行权衡。

3.3 CMOS 后道集成技术

RRAM 以其结构简单、制程温度低的特点,可在 CMOS 后段 (back-end-of-line, BEOL) 完成沉积与互连,如图 10 所示,被视为先进工艺节点下高密度非易失性存储的优选方案。近年来,台积电、中芯国际、中国科学院微电子所等企业/高校陆续展示了 28 nm/14 nm 等先进节点下的嵌入式 RRAM 集成工艺^[56],并向 7 nm/3 nm 等更先进节点推进。通过在 M1 - M3 互连层之间直接堆叠 RRAM 单元,无须改动前端逻辑电路即可实现大容量片上存储;若进一步结合 3D VRRAM、1T2R 等新型单元结构,阵列密度仍有持续提升的空间^[57-58]。

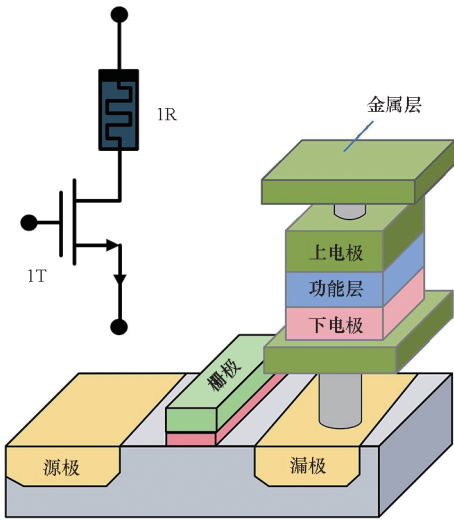


图 10 1T1R 型 RRAM 与 CMOS 的集成结构

Fig. 10 1T1R RRAM integrated with CMOS structure

在具体工艺方面,CMOS 兼容性主要取决于材料选择与温度窗口^[59-60]:常用的 HfO_x/TiN 、 TaO_x/TiN 等器件堆叠可在低于 400 °C 的后段工艺中通过原子层沉积 (atomic layer deposition, ALD) 或磁控溅射制备,既能保证薄膜均匀性与导电细丝形成的可控性,又可避免对低 k 介质和铜互连造成热损伤^[61-62]。面向 28 nm/14 nm 量产平台,业界已形成一套成熟的“晶体管上叠层

RRAM”工艺流程,涵盖下电极平坦化、阻变层沉积与刻蚀、顶部电极填充等关键步骤^[56];结合位线转置、选通管共享等电路设计,可将器件操作电压降至 1 V 以下,兼顾可靠性与能效^[63]。

目前仍存在三方面主要挑战:其一,电极平坦度与薄膜缺陷控制直接影响阵列均匀性,需要依靠高精度化学机械抛光 (CMP)、刻蚀工艺及应力管理来实现^[64];其二,1T1R 架构受限于选通管尺寸,需通过材料与器件协同优化来降低操作电压与电流,或引入 1T2R、1S1R 等组合结构以提升集成密度^[65];其三,阵列级可靠性指标 (如写入窗口、循环寿命、数据保持特性与软错误率) 仍需借助冗余列、错误校正与在线校准等策略进行系统优化^[66]。总体而言,随着 BEOL 集成工艺模块在 30.48 cm 平台上完成验证,RRAM 的 CMOS 兼容路线已逐步从研发阶段走向量产,预计其在嵌入式高密度存储、存算一体及安全芯片等领域中的应用将进一步加速。

3.4 小结

本节围绕 RRAM 的阵列级集成问题,系统讨论了交叉阵列架构、三维高密度集成方案以及 CMOS 后道集成技术的最新进展。针对大规模交叉阵列中旁路电流与串扰效应对阵列可靠性和存算精度的制约,本节从单元密度、抑制能力与集成可扩展性等维度,对 1R、1S1R 以及 1T1R 单元结构进行了对比分析。进一步地,本节指出三维堆叠与 CMOS 兼容工艺为突破存储密度与系统能效瓶颈提供了关键路径,但同时引入制造复杂性、热管理以及阵列级可靠性等工程挑战。总体而言,阵列层面的性能优化高度依赖器件与工艺层的协同创新,其成熟度将直接决定 RRAM 技术向大规模系统应用推进的可行性。

4 RRAM 的应用方向

RRAM 阵列集成技术的不断突破,正使其成为信息存储与计算领域跨界融合的关键使能技术。凭借其非易失性、可并行调控特性及与 CMOS 工艺的良好兼容优势,RRAM 可以直接满足不同应用场景的核心硬件需求。本节将围绕存算一体技术、神经形态计算、智能感知和硬件安全四大核心应用方向,系统综述 RRAM 的实用化进展。

4.1 存算一体技术

基于 RRAM 的存算一体技术可以在交叉阵列中直接执行模拟矩阵 - 向量乘,如图 11 所示,

其物理基础在于利用欧姆定律与基尔霍夫定律。该技术通过存储器件的电导状态实现模拟计算,从而规避冯·诺依曼瓶颈,可以显著提升 AI 加速器的速度、能效及吞吐量。代表性研究已证实了其高效矩阵运算能力,例如支持基本的向量运算,以及通过定制宏电路实现低延迟点积计算。这些硬件创新结合电路与算法的协同设计,正为可扩展的深度学习硬件系统奠定基础。

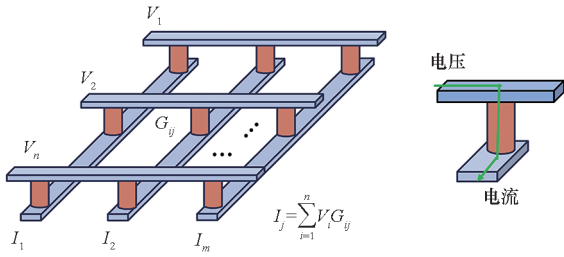


图 11 RRAM 原位存内计算示意图

Fig. 11 Schematic diagram of RRAM-based in-situ computing-in-memory architecture

上述进展主要集中于基础算子的硬件实现。基于矩阵运算机制,存算一体架构通过对 RRAM 等非易失性存储器的硬件优化,在深度学习任务中实现了能效、精度和吞吐量的同步提升。图 12 对比了 2020—2025 年不同存储技术的存算一体芯片能效,以 TOPS/W (tera operations per second per watt) 作为衡量指标,包括 RRAM、SRAM、DRAM、Flash、PCM 和 MRAM^[67-90],其中 RRAM 在现有技术中展现出能效领先优势。

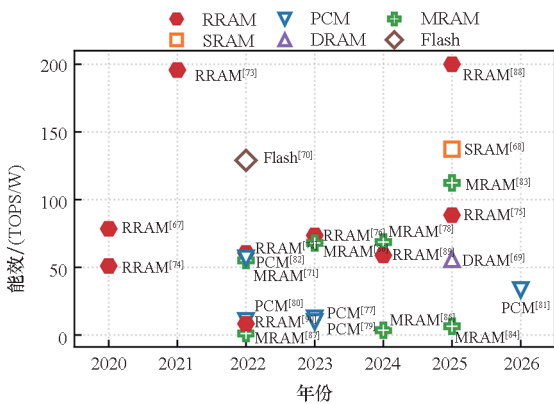


图 12 不同存储器类型的存内计算能效对比

Fig. 12 Comparison of energy efficiency of compute-in-memory based on different memory technologies

在卷积神经网络 (convolutional neural network, CNN) 中,存内计算已能够实现高精度推理。例如,在 ImageNet 和 MNIST 等标准数据集上,其推理准确率已达到与软件相当的水平^[91-92]。此外,基于 RRAM 的存算一体芯片在

多种任务中均展现出高能效特性^[93]。对于循环神经网络 (recurrent neural network, RNN) 及实时处理任务,集成了非线性激活函数的忆阻阵列展现出独特优势,基于 RRAM 的存算架构也适用于音频分类与视频实时处理等场景。综上,通过算法-硬件的协同设计,存内计算技术能够在多种网络架构下兼顾精度与能效,具备了向边缘端部署的潜力。为适应边缘人工智能 (edge artificial intelligence, edge AI) 场景,存内计算架构需进行针对性设计,以满足其对低延迟、高能效及小型化的严苛要求。通过采用近存储计算、传感器融合与先进热管理等一系列策略,存内计算能够有效克服传统冯·诺依曼架构的数据移动瓶颈,以满足物联网及各类实时应用对超低功耗、快速响应和高鲁棒性的需求。例如,采用 RRAM-SRAM 混合精度的处理器,能在深度学习任务中同时实现高能效计算和快速唤醒^[94]。

然而,要实现存内计算的高性能,必须克服硬件非理想特性导致的精度损失与器件变异等挑战。为应对这些挑战,领域内正发展材料、电路与算法多层次协同的创新路径,以系统化地管理精度损失与器件变异,在保障可靠性的同时维持高能效。具体而言,在材料与器件层面,可通过掺杂、异质结设计及界面工程等策略提升器件均匀性。在电路架构层面,则可通过选择器集成与脉冲编程等技术抑制器件变异性等不利影响。而在算法与系统层面,则可利用容错训练、混合精度计算等策略,从计算本质上补偿硬件的非理想性。上述跨层级协同策略,是保障当前精度与能效的关键,将最终推动下一代 AI 硬件的发展。

4.2 神经形态计算

神经形态计算通过模拟生物神经网络的信息处理机制,致力于实现高能效和自适应的智能硬件。RRAM 作为实现人工突触的核心组件,是模拟突触可塑性、构建脉冲神经网络 (spiking neural network, SNN)、探索储备池计算及集成神经元-突触系统的硬件基石。本节将系统综述 RRAM 在这些应用场景中的架构设计与最新进展。

类脑计算的核心之一是模拟生物突触的信息加权功能,而 RRAM 是实现该功能的典型硬件突触载体,两者的对应关系可通过图 13 直观体现。作为人工突触,RRAM 通过模拟 STDP、LTP/LTD 等可塑性规则及记忆的动态转换,支持了高能效神经形态计算^[95]。例如,氧化物基和金属离子基 RRAM 分别利用氧空位迁移和电化学机制模拟了 STDP、LTP/LTD 等突触可塑性行为及突触权重

变化。

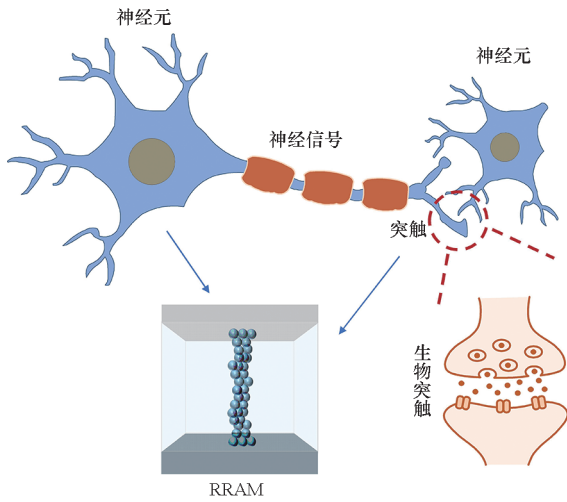


图 13 类脑计算中生物突触与 RRAM 突触的对应关系

Fig. 13 Correspondence between biological and RRAM synapses in neuromorphic computing

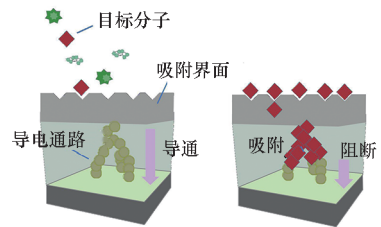
在 RRAM 人工突触的基础上构建的 SNN,能够利用其硬件特性高效编码和处理时空信息,从而在动态环境中实现类脑的模式识别。通过将脉冲时序机制与 RRAM 结合,SNN 显著提升了能效和鲁棒性,并在时序信号处理任务中实现了高分类精度^[96]。其优异的时序处理能力源于仿生的架构设计^[97]。

为追求更高的系统集成度与生物相似性,神经元-突触集成系统应运而生。它在单一器件或阵列中融合了突触可塑性与神经元动态行为,实现了功能可重构和高能效的处理架构。典型如 NbO_x 基等体系,可在单器件中实现突触与神经元功能的协同模拟,且通过泄漏积分发放神经元模型(leaky integrate-and-fire, LIF)等动力学仿生设计以增强适应性^[98-99]。通过器件结构优化和系统级协同创新(如三维交叉阵列),此类系统在图像分类等任务中实现了极高计算密度和极低能耗。

4.3 智能感知

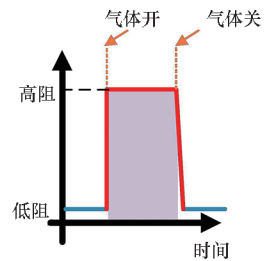
除了在神经形态计算领域的探索,RRAM 也正向智能感知系统扩展,用于构建集多模态感知与前端处理能力于一体的硬件。光电 RRAM 通过将图像感知、非易失性存储和预处理功能集成于单一器件,有效克服了传统视觉系统因功能分离导致的高能耗和数据传输瓶颈问题,显著提升了人工视觉系统的能效与处理准确性。在材料层面,研究已证实铝掺杂氧化锌(AZO)柔性 RRAM 阵列和无铅钙钛矿器件等体系展示了宽带光学响应与长时图像存储能力^[100-101]。在系统层面,全

集成 CMOS 与光电 RRAM 阵列通过多模式操作进一步提高了图像识别准确率并降低能耗^[102]。这些进展凸显了光电 RRAM 在实现传感-存储-计算一体化架构中的核心硬件地位,为高效静态图像处理奠定了硬件基础。此外,RRAM 技术的感知能力并不限于视觉领域。基于电阻切换机制,RRAM 在气体、化学与生物传感领域展现了高灵敏度检测与快速响应优势,可被应用于环境监测、医疗诊断及生物集成系统。在气体传感中如图 14 所示,诸如 Ti/TiO₂/Pt 等 RRAM 器件能在室温下对氧化/还原性气体实现选择性检测;若与人工神经网络结合,则可进一步实现高精度气体识别^[103]。在化学传感方面,RRAM 可用于检测特定离子浓度,并展现出线性响应和多态记忆特性^[104]。而对于生物传感,RRAM 则能实现对生物分子的高灵敏度检测^[105]。这些 RRAM 传感器不仅提升了单一模态的检测性能,也为构建多模态集成感知系统奠定了基础。然而,要走向实际应用,仍需解决材料稳定性与环境适应性等挑战。



(a) 吸附示意图

(a) Adsorption diagram



(b) 阻态变化

(b) Resistance change

图 14 基于 RRAM 的气体目标物检测及阻态调控

Fig. 14 Schematic diagram of gas target detection process and resistance state regulation based on RRAM

进一步地,RRAM 在系统集成与边缘部署方面也展现出重要价值。通过集成存储与模拟计算功能,RRAM 能为物联网及边缘设备提供低功耗、高效率的传感器系统解决方案,适用于健康监测、实时通信与能源收集等多种应用场景。例如在可穿戴健康监测中,基于 RRAM 的系统已实现生理

信号采集及人机交互功能^[106-107];同时,柔性可与可拉伸 RRAM 通过材料与结构创新,凭借其可拉伸、自愈合特性,显著提升了设备的可靠性与兼容性。在实时通信^[108]以及结合自供电与神经网络近传感器计算系统中^[109-110],RRAM 均被证明能有效降低延迟与能耗,提升系统的可持续性。因此,RRAM 的一体化架构是解决边缘计算资源瓶颈、赋能物联网应用的关键技术之一。

4.4 硬件安全

随着 RRAM 在智能感知领域走向系统集成,其在安全敏感场景下的独特价值也日益凸显。硬件安全的核心在于利用器件固有特性构建安全机制。RRAM 凭借其固有变异性和可控开关特性,为物理不可克隆函数(physical unclonable function, PUF)和真随机数生成器(true random number generator, TRNG)提供了高效的物理熵源。图 15 展示了其典型的密钥生成形式,同时区分了弱 PUF 与强 PUF 的结构及功能差异。基于电阻开关机制,RRAM 能够实现高可靠且独特的 PUF 响应。尽管多项研究基于 RRAM 实现了高可靠的 PUF 设计^[111-112],但高性能 PUF 往往需要对导电细丝进行精细调控,这增加了工艺复杂性。在 TRNG 方面,RRAM 利用其随机开关行为生成高熵输出,例如有研究利用概率延迟机制实现了高比特率并成功通过随机性测试^[113]。总体而言,RRAM 基 PUF 和 TRNG 凭借其物理随机性为安全硬件奠定了坚实基础,但仍面临随机性质量与可靠性的权衡问题,且方案的通用性有待提升。不过以此为基础,RRAM 可进一步支撑逻辑锁定等更为复杂的安全应用。

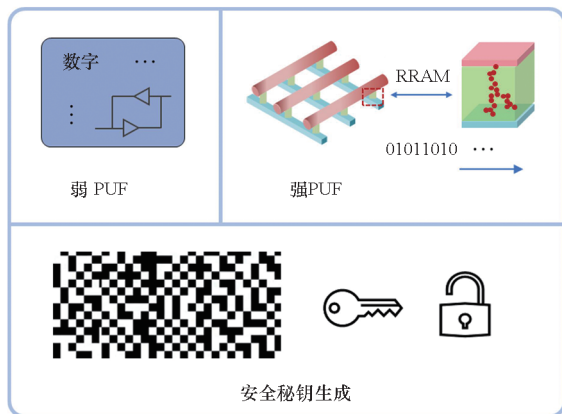


图 15 基于 RRAM 的物理不可克隆函数示意图

Fig. 15 Schematic diagram of RRAM-based PUF

为防御逆向工程与未授权访问,逻辑锁定及可证明密钥销毁等方案可利用 RRAM 的物理随机性作为硬件基石。例如,基于 HfO_2 的忆阻交叉

阵列已成功实现了可证明的密钥销毁机制^[114]。此外,RRAM 的电阻开关特性也使其能够实现可重构的安全逻辑功能。例如, HfO_2 基 RRAM 利用比例逻辑来执行逻辑操作,但其可靠性仍是实际应用的挑战。尽管 RRAM 在安全硬件系统中潜力显著,但为充分发挥这一潜力,未来仍需通过材料优化与可靠性提升来克服现有局限。

更进一步,RRAM 技术通过将加密操作与安全数据传输集成于单一硬件平台,构建从边缘设备到数据中心的高能效安全计算系统。在加密与通信硬件领域,RRAM 可结合 PUF、同态加密技术优化协议^[115-116]以及并行存内计算架构增强数据传输性能^[117-118]。然而,此类基于 RRAM 的加密系统在实际应用中的精度损失等问题仍有待解决。

总体而言,RRAM 为安全关键应用提供了一种可扩展、高能效的硬件架构选择。未来需解决集成通用性等挑战,以促进其广泛落地应用。

4.5 小结

本节综述了 RRAM 在存算一体、神经形态计算、智能感知及硬件安全等关键应用领域的研究进展。得益于其非易失性、可并行调控及与 CMOS 工艺的良好兼容性,RRAM 在突破“存算分离”瓶颈、实现高能效智能计算方面展现出独特优势。然而,不同应用场景对精度、可靠性与能效的侧重点存在显著差异,使得器件变异性、非理想效应及系统级稳定性成为制约其实际落地的主要因素。未来,RRAM 的应用发展将更加依赖材料、器件、电路与算法之间的跨层级协同优化,以实现从单器件演示向系统集成的转变。

5 总结与展望

本文系统梳理了 RRAM 在机理解析、性能调控、工艺集成及系统应用层面的研究进展,并总结了其在存算一体、神经形态计算、智能感知与硬件安全等方向中的应用潜力。然而,从实验室研究走向规模化应用,RRAM 仍面临多方面的关键挑战,亟须在未来研究中加以系统突破。

5.1 器件优化与协同调控

器件间及循环间一致性不足以及长期可靠性受限,是当前 RRAM 技术发展面临的核心瓶颈。未来研究需要在物理机理与材料设计层面实现更强的可控性。

器件优化方面,可通过缺陷工程、界面调控等策略,引导导电通路在空间分布和能量势垒上的

可预测演化,从源头降低器件参数离散性。此外,当前发展趋势将由单一材料或结构优化,转向材料设计与电学编程策略的协同调控。例如,通过稳定缺陷构型、抑制局域过热效应,并引入自适应脉冲编程、写入验证及在线校准机制,在提升循环耐久性与数据保持特性的同时兼顾多值精度。

5.2 阵列集成与跨层设计

在高密度交叉阵列与三维集成架构中,旁路电流、热串扰、电串扰及器件非理想性等问题影响了 RRAM 的存算精度、能效与系统稳定性。研究的重点将从单纯抑制非理想效应转向器件 - 电路 - 算法的跨层协同设计,以系统化方式应对阵列级复杂问题。

在器件与电路层面,通过发展高性能选择器、自整流 RRAM 以及内置非线性结构,可有效缓解旁路电流与阵列串扰问题;在集成层面,三维堆叠与 CMOS 后道集成仍是提升存储密度与并行计算能力的必由之路。未来需关注低温工艺窗口内材料体系的优化、堆叠结构的简化以及层间热管理设计,从集成验证走向稳定量产。在系统与算法层面引入容错训练、混合精度计算及误差补偿机制,在一定程度上消化器件和阵列层的非理想性,实现精度与能效的协同提升。

5.3 应用驱动的优化趋势

随着 RRAM 在人工智能推理、边缘计算、智能感知及硬件安全等应用场景中的需求逐渐清晰。不同于以往以器件性能指标为主导的研究模式,未来研究将更多由应用需求反向约束架构设计、器件参数及材料选择,推动形成“应用 - 架构 - 器件 - 材料”协同优化的闭环。

例如,在边缘 AI 与存算一体应用中,对低功耗、高效能与高鲁棒性的需求,将引导器件设计更加关注一致性与能效而非极限开关比;在神经形态计算与智能感知系统中,对动态可塑性与多模态响应能力的需求,则促使器件与材料向功能可重构方向发展;而在硬件安全领域,对随机性质量与长期稳定性的双重要求,也将推动对阻变随机性“可控利用”的深入研究。应用驱动的研究范式,将避免性能指标的无序堆叠,显著提升 RRAM 技术路线的工程可行性与产业价值。

5.4 小结

总体而言,RRAM 的未来突破将不再依赖单一层面的渐进式改进,而是需要在物理机制理解、材料与器件设计、阵列集成工艺以及系统与算法架构层面实现深度协同创新。随着相关理论模型

的不断完善、制造工艺的逐步成熟以及标准化体系的建立,RRAM 有望在高密度非易失性存储、存算一体计算及智能硬件系统中发挥关键支撑作用,为构建新一代低功耗、高效能的信息处理体系提供重要技术基础。

参考文献 (References)

- [1] REINSEL D, GANTZ J, RYDNING J. The digitization of the world from edge to core [R]. Framingham, USA: IDC, 2025.
- [2] 中商产业研究院. 2025 年中国半导体存储器行业最新政策汇总一览 [EB/OL]. (2025 - 07 - 24) [2025 - 12 - 03]. [https://www. askci. com/news/chanye/20250724/094336275332141517524703. shtml](https://www.askci.com/news/chanye/20250724/094336275332141517524703.shtml). China Business Industry Research Institute. 2025 China semiconductor memory industry latest policy summary overview [EB/OL]. (2025 - 07 - 24) [2025 - 12 - 03]. [https://www. askci. com/news/chanye/20250724/094336275332141517524703. shtml](https://www. askci. com/news/chanye/20250724/094336275332141517524703.shtml). (in Chinese)
- [3] CHIPS for America. The CHIPS program office vision for success: two years later [R]. USA: CHIPS for America, 2025.
- [4] WANG Z R, WU H Q, BURR G W, et al. Resistive switching materials for information processing [J]. *Nature Reviews Materials*, 2020, 5(3): 173 - 195.
- [5] COVI E, DONATI E, LIANG X P, et al. Adaptive extreme edge computing for wearable devices [J]. *Frontiers in Neuroscience*, 2021, 15: 611300.
- [6] YU S M. Semiconductor memory devices and circuits [M]. Boca Raton: CRC Press, 2022.
- [7] CHUA L. Memristor-the missing circuit element [J]. *IEEE Transactions on Circuit Theory*, 1971, 18(5): 507 - 519.
- [8] STRUKOV D B, SNIDER G S, STEWART D R, et al. The missing memristor found [J]. *Nature*, 2008, 453(7191): 80 - 83.
- [9] HUANG W, WAN P, ZHANG Z C, et al. Memristors enabling biomimetic pain perception based on pure VO₂ nanoplates synthesized via annealing of VSe₂ nanoplates [J]. *ACS Applied Electronic Materials*, 2025, 7(20): 9534 - 9541.
- [10] ZHANG Y, MAO G Q, ZHAO X L, et al. Evolution of the conductive filament system in HfO₂-based memristors observed by direct atomic-scale imaging [J]. *Nature Communications*, 2021, 12: 7232.
- [11] DAI L J, WU H Q, GAO B, et al. A compact model for the SET parameter variations of oxide RRAM array [C]// *Proceedings of 2016 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*, 2016: 1 - 2.
- [12] ROHAN J N, ZHUANG P P, NIBHANUPUDI S T, et al. Neural network assisted compact model for accurate characterization of cycle-to-cycle variations in 2-D h-BN based RRAM devices [C]// *Proceedings of 2019 Device Research Conference (DRC)*, 2019: 103 - 104.
- [13] LIU Q, SUN J, LV H B, et al. Real-time observation on dynamic growth/dissolution of conductive filaments in oxide-electrolyte-based ReRAM [J]. *Advanced Materials*, 2012, 24(14): 1844 - 1849.

- [14] GOORAN-SHOORAKCHALY A, SHARIF S S, BANAD Y M. Investigating the effect of electrical and thermal transport properties on oxide-based memristors performance and reliability[J]. *Scientific Reports*, 2025, 15: 18646.
- [15] GARCÍA H, BOO J, VINESA G, et al. Influences of the temperature on the electrical properties of HfO₂-based resistive switching devices[J]. *Electronics*, 2021, 10(22): 2816.
- [16] CHENG C D, WANG Y H, XU L Y, et al. Artificial astrocyte memristor with recoverable linearity for neuromorphic computing[J]. *Advanced Electronic Materials*, 2022, 8(8): 2100669.
- [17] YAO L D, INKINEN S, VAN DIJKEN S. Direct observation of oxygen vacancy-driven structural and resistive phase transitions in La_{2/3}Sr_{1/3}MnO₃[J]. *Nature Communications*, 2017, 8: 14544.
- [18] WANG M, CAI S H, PAN C, et al. Robust memristors based on layered two-dimensional materials[J]. *Nature Electronics*, 2018, 1(2): 130–136.
- [19] KHARLANOV O G. Effects of quantum recoil forces in resistive switching in memristors[J]. *JETP Letters*, 2023, 117(5): 384–391.
- [20] SANGWAN V K, JARIWALA D, KIM I S, et al. Gate-tunable memristive phenomena mediated by grain boundaries in single-layer MoS₂[J]. *Nature Nanotechnology*, 2015, 10(5): 403–406.
- [21] LEHENKARI T, HUANG S D, KORDAS K, et al. First-principles study of Ag, Au, Cu, and Li defects in MoS₂ and their application to memristors[J]. *Physical Review Materials*, 2025, 9: 014001.
- [22] TAN Z J, SOMJIT V, TOPARLÍ C, et al. Electronegative metal dopants improve switching variability in Al₂O₃ resistive switching devices[J]. *Physical Review Materials*, 2022, 6(10): 105002.
- [23] AZIZ J, KHAN M F, NEUMAIER D, et al. Improved memory performance of ALD grown HfO₂ films by nitrogen doping[J]. *Materials Science and Engineering: B*, 2023, 297: 116755.
- [24] KIM D E, CHABUNGBAM A S, KIM G, et al. Doping engineering for optimized self-rectifying TaO_x memristor for crossbar array neuromorphic applications[J]. *Advanced Functional Materials*, 2025, 35(45): 2503883.
- [25] CHAI X J, YAN Z Y, WANG M R, et al. High-performance multi-level memristor based on Ag NPs-doped MXene for reliable artificial synapse[J]. *Chemical Engineering Journal*, 2025, 522: 168181.
- [26] LI S Y, ZENG Y L, JIANG J Y, et al. Towards outstanding resistive switching and synaptic behaviors using high-entropy alloys-based memristors for neuromorphic computing[J]. *Applied Surface Science*, 2025, 713: 164292.
- [27] PASSERINI E, LEWERENZ M, CSANTOS M, et al. Controlling volatility and nonvolatility of memristive devices by Sn alloying[J]. *ACS Applied Electronic Materials*, 2023, 5(12): 6842–6849.
- [28] LEE M J, KIM T H, LEE S H, et al. Intermediate layer-assisted trap density reduction in low-power optoelectronic memristors for multifunctional systems[J]. *Advanced Functional Materials*, 2025, 35(22): 2421080.
- [29] LEE J W, CAI L, NAM J S, et al. Bifacially engineered perovskite-based synaptic memristors achieve high linearity and symmetry for accurate and robust neuromorphic computing[J]. *Advanced Science*, 2025, 12(42): e11489.
- [30] NOH M, KIM S. Material-engineered self-compliant memristor enabling multibit synaptic learning and in-memory computing[J]. *Journal of Materials Chemistry C*, 2025, 13(40): 20675–20689.
- [31] KIM M, KIM S, HWANG S. Self-compliance and forming-free memristor arrays with a SiO₂ scavenging barrier for energy-efficient neuromorphic computing[J]. *Journal of Physics D: Applied Physics*, 2025, 58(37): 375102.
- [32] YANG S, KIM T, KIM S, et al. Demonstration of synaptic characteristics in VRRAM with TiN nanocrystals for neuromorphic system[J]. *Advanced Materials Interfaces*, 2023, 10(21): 2300290.
- [33] GRACIO M, TEIXEIRA H, DIAS C, et al. Effect of electrode tuning on the resistive switching behaviour of MXene-based composites[J]. *Polymers*, 2025, 17(10): 1309.
- [34] LIU Y L, CHEN Q L, GUO Y B, et al. Enhancing the uniformity of a memristor using a bilayer dielectric structure[J]. *Micromachines*, 2024, 15(5): 605.
- [35] WANG C X, MAO G Q, HUANG M H, et al. HfO_x/AlO_y superlattice-like memristive synapse[J]. *Advanced Science*, 2022, 9(21): 2201446.
- [36] LI Y S, CHEN S, YU Z G, et al. In-memory computing using memristor arrays with ultrathin 2D PdSeO_x/PdSe₂ heterostructure[J]. *Advanced Materials*, 2022, 34(26): 2201488.
- [37] CHEN L L, MA Z Y, LENG K M, et al. Artificial synapse consisted of TiSbTe/SiC_x: H memristor with ultra-high uniformity for neuromorphic computing[J]. *Nanomaterials*, 2022, 12(12): 2110.
- [38] JIN S, KWON J D, KIM Y. Statistical analysis of uniform switching characteristics of Ta₂O₅-based memristors by embedding in-situ grown 2D-MoS₂ buffer layers[J]. *Materials*, 2021, 14(21): 6275.
- [39] XIAO M, HELLENBRAND M, STRKALJ N, et al. Ultra-fast non-volatile resistive switching devices with over 512 distinct and stable levels for memory and neuromorphic computing[J]. *Advanced Functional Materials*, 2025, 35(29): 2418980.
- [40] FENG Y L, HUANG P, ZHAO Y D, et al. Improvement of state stability in multi-level resistive random-access memory (RRAM) array for neuromorphic computing[J]. *IEEE Electron Device Letters*, 2021, 42(8): 1168–1171.
- [41] MAHATA C, ALGADI H, ISMAIL M, et al. Controlled multilevel switching and artificial synapse characteristics in transparent HfAlO-alloy based memristor with embedded TaN nanoparticles[J]. *Journal of Materials Science & Technology*, 2021, 95: 203–212.
- [42] KIM M, JU D, KANG M, et al. Improved resistive and synaptic characteristics in neuromorphic systems achieved using the double-forming process[J]. *Nanomaterials*, 2023, 13(21): 2859.
- [43] RAHMANI M K, KHAN S A, PARK J H, et al. Modulation of the electrical characteristics on poly(3,4-ethylenedioxythiophene)-poly(styrenesulfonate)-based resistive random access memory device by the impact of top electrode materials[J]. *Thin Solid Films*, 2022, 748: 139168.
- [44] LIN C Y, TSENG Y T, CHEN P H, et al. A high-speed MIM resistive memory cell with an inherent vanadium

- selector[J]. *Applied Materials Today*, 2020, 21: 100848.
- [45] SHARMA A A, SKOWRONSKI M, BAIN J A, et al. Novel CMOS-compatible a-Si based oscillator and threshold switch[C]//*Proceedings of the 45th European Solid State Device Research Conference (ESSDERC)*, 2015: 93–96.
- [46] RANJAN H, SINGH C P, SINGH V P, et al. Self-rectifying and forming-free resistive switching with Cu/BN/SiO₂/Pt bilayer device [J]. *Materials Science in Semiconductor Processing*, 2024, 183: 108744.
- [47] KIM M, LEE S, KIM S J, et al. Study on the sodium-doped titania interface-type memristor[J]. *ACS Applied Materials & Interfaces*, 2024, 16(13): 16453–16461.
- [48] PAL P, SINGH A, WANG Y H. Controllable coexistence of threshold and non-volatile crosspoint memory for highly linear synaptic device applications [J]. *Journal of Physics D: Applied Physics*, 2023, 56(28): 285102.
- [49] YANG Y F, XU M K, JIA S J, et al. A new opportunity for the emerging tellurium semiconductor; making resistive switching devices[J]. *Nature Communications*, 2021, 12: 6081.
- [50] SONG M, LEE S, NIBHANUPUDI S S T, et al. Self-compliant threshold switching devices with high on/off ratio by control of quantized conductance in Ag filaments[J]. *Nano Letters*, 2023, 23(7): 2952–2957.
- [51] LEE M J, LEE D, CHO S H, et al. A plasma-treated chalcogenide switch device for stackable scalable 3D nanoscale memory[J]. *Nature Communications*, 2013, 4: 2629.
- [52] ZHAO X L, NIU J B, YANG Y, et al. Modulating the filament rupture degree of threshold switching device for self-selective and low-current nonvolatile memory application[J]. *Nanotechnology*, 2020, 31(14): 144002.
- [53] WANG W, WANG M, AMBROSI E, et al. Surface diffusion-limited lifetime of silver and copper nanofilaments in resistive switching devices[J]. *Nature Communications*, 2019, 10: 81.
- [54] DING Y X, YANG J G, LIU Y, et al. 16-layer 3D vertical RRAM with low read latency (18 ns), high nonlinearity (>5 000) and ultra-low leakage current (~pA) self-selective cells[C]//*Proceedings of 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, 2023: 1–2.
- [55] CHOI S, SHIN J, PARK G, et al. 3D-integrated multilayered physical reservoir array for learning and forecasting time-series information[J]. *Nature Communications*, 2024, 15: 2044.
- [56] XU X X, YU J, GONG T C, et al. First demonstration of OxRRAM integration on 14 nm FinFet platform and scaling potential analysis towards sub-10 nm node[C]//*Proceedings of 2020 IEEE International Electron Devices Meeting (IEDM)*, 2020: 24.3.1–24.3.4.
- [57] CAO Y, YANG H H, YANG J G, et al. A 67F² reconfigurable PUF using 1T2R RRAM switching competition in 28 nm CMOS with 5e-9 bit error rate[C]//*Proceedings of 2024 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, 2024: 1–2.
- [58] YANG J G, XUE X Y, XU X X, et al. A 2.8 nm 1.5 Mb embedded 1T2R RRAM with 14.8 Mb/mm² using sneaking current suppression and compensation techniques [C]//*Proceedings of 2020 IEEE Symposium on VLSI Circuits*, 2020: 1–2.
- [59] WALCZYK C, SOWINSKA M, WALCZYK D, et al. (Invited) resistive switching and current status of HfO₂-based RRAM[J]. *ECS Transactions*, 2014, 61(2): 315.
- [60] SHAHRABI E, LAGRANGE T, DEMIRCI T, et al. Performance improvement of chip-level CMOS-integrated ReRAM cells through material optimization [J]. *Microelectronic Engineering*, 2019, 214: 74–80.
- [61] CHAKRABARTI S, ANG J M, THONG J R, et al. Effects of thermal annealing on Ta₂O₅ based CMOS compatible RRAM[C]//*Proceedings of 2020 IEEE Silicon Nanoelectronics Workshop (SNW)*, 2020: 79–80.
- [62] MOLINA J, TORRES R, RANJAN A, et al. Resistive switching characteristics of MIM structures based on oxygen-variable ultra-thin HfO₂ and fabricated at low temperature[J]. *Materials Science in Semiconductor Processing*, 2017, 66: 191–199.
- [63] YANG J G, XUE X Y, XU X X, et al. 24.2 A 14 nm FinFET 1 Mb embedded 1T1R RRAM with a 0.022 μm² cell size using self-adaptive delayed termination and multi-cell reference[C]//*Proceedings of 2021 IEEE International Solid-State Circuits Conference (ISSCC)*, 2021: 336–338.
- [64] FAN Y, KING S W, BIELEFELD J, et al. Characterization of porous BEOL dielectrics for resistive switching[J]. *ECS Transactions*, 2016, 72(2): 35–50.
- [65] UHLMANN M, PÉREZ-BOSCH QUESADA E, FRITSCHER M, et al. One-transistor-multiple-RRAM cells for energy-efficient in-memory computing[C]//*Proceedings of the 21st IEEE Interregional NEWCAS Conference (NEWCAS)*, 2023: 1–5.
- [66] LING Y T, WANG Z W, YU Z Z, et al. Temperature-dependent accuracy analysis and resistance temperature correction in RRAM-based in-memory computing[J]. *IEEE Transactions on Electron Devices*, 2024, 71(1): 294–300.
- [67] LIU Q, GAO B, YAO P, et al. 33.2 A fully integrated analog ReRAM based 78.4 TOPS/W compute-in-memory chip with fully parallel MAC computing[C]//*Proceedings of 2020 IEEE International Solid-State Circuits Conference (ISSCC)*, 2020: 500–502.
- [68] HU X F, MUN H, MENG J, et al. A 28 nm 20.9–137.2 TOPS/W output-stationary SRAM compute-in-memory macro featuring dynamic look-ahead zero weight skipping and runtime partial sum quantization[C]//*Proceedings of 2025 IEEE Custom Integrated Circuits Conference (CICC)*, 2025: 1–3.
- [69] ZHANG X Y, JO Y J, KIM T T H. A 65-nm 55.8-TOPS/W compact 2T eDRAM-based compute-in-memory macro with linear calibration[J]. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2025, 33(5): 1477–1481.
- [70] KIM M, LIU M Q, EVERSON L R, et al. An embedded nand flash-based compute-in-memory array demonstrated in a standard logic process [J]. *IEEE Journal of Solid-State Circuits*, 2022, 57(2): 625–638.
- [71] CHIU Y C, YANG C S, TENG S H, et al. A 22 nm 4 Mb STT-MRAM data-encrypted near-memory computation macro with a 192 GB/s read-and-decryption bandwidth and 25.1–55.1 TOPS/W 8 b MAC for AI operations[C]//*Proceedings of 2022 IEEE International Solid-State Circuits Conference (ISSCC)*, 2022: 178–180.
- [72] CHANG M Y, SPETALNICK S D, CRAFTON B, et al. A 40 nm 60.64 TOPS/W ECC-capable compute-in-memory/

- digital 2.25 MB/768 KB RRAM/SRAM system with embedded cortex M3 microprocessor for edge recommendation systems[C]//Proceedings of 2022 IEEE International Solid-State Circuits Conference (ISSCC), 2022: 1–3.
- [73] XUE C X, HUNG J M, KAO H Y, et al. 16.1 A 22 nm 4 Mb 8b-precision ReRAM computing-in-memory macro with 11.91 to 195.7 TOPS/W for tiny AI edge devices [C]//Proceedings of 2021 IEEE International Solid-State Circuits Conference (ISSCC), 2021: 245–247.
- [74] HE W X, YIN S H, KIM Y, et al. 2-bit-per-cell RRAM-based in-memory computing for area-/energy-efficient deep learning [J]. *IEEE Solid-State Circuits Letters*, 2020, 3: 194–197.
- [75] WANG L F, LI W Z, ZHOU Z D, et al. A near-threshold memristive computing-in-memory engine for edge intelligence[J]. *Nature Communications*, 2025, 16: 5897.
- [76] CHANG M Y, LELE A S, SPETALNICK S D, et al. A 73.53 TOPS/W 14.74 TOPS heterogeneous RRAM in-memory and SRAM near-memory SoC for hybrid frame and event-based target tracking [C]//Proceedings of 2023 IEEE International Solid-State Circuits Conference (ISSCC), 2023: 426–428.
- [77] AMBROGIO S, NARAYANAN P, OKAZAKI A, et al. An analog-AI chip for energy-efficient speech recognition and transcription[J]. *Nature*, 2023, 620(7975): 768–775.
- [78] DEAVILLE P, ZHANG B N, VERMA N. A fully row/column-parallel in-memory computing macro in foundry MRAM with differential readout for noise rejection[J]. *IEEE Journal of Solid-State Circuits*, 2024, 59(7): 2070–2080.
- [79] LE GALLO M, KHADDAM-ALJAMEH R, STANISAVLJEVIC M, et al. A 64-core mixed-signal in-memory compute chip based on phase-change memory for deep neural network inference[J]. *Nature Electronics*, 2023, 6(9): 680–693.
- [80] KHADDAM-ALJAMEH R, STANISAVLJEVIC M, FORNT MAS J, et al. HERMES-core: a 1.59-TOPS/mm² PCM on 14-nm CMOS in-memory compute core using 300-ps/LSB linearized CCO-based ADCs[J]. *IEEE Journal of Solid-State Circuits*, 2022, 57(4): 1027–1038.
- [81] YANG X Y, LEI Y, TAN J Z, et al. Bayesian neural network realization by exploiting random noise and mitigating the impact of conductance drift of PCM[J]. *Advanced Materials Technologies*, 2026, 11(4): e01278.
- [82] ZHOU C T, REDONDO F G, BÜCHEL J, et al. ML-HW co-design of noise-robust TinyML models and always-on analog compute-in-memory edge accelerator[J]. *IEEE Micro*, 2022, 42(6): 76–87.
- [83] LI H M, CHAI Z, DONG W R, et al. A lossless and fully parallel spintronic compute-in-memory macro for artificial intelligence chips [J]. *Nature Electronics*, 2025, 8(11): 1046–1058.
- [84] MENG H R, WU Y J, ZHOU S C, et al. A 40 nm STT-MRAM near-memory computing macro for memory-augmented neural network acceleration [C]//Proceedings of 2025 IEEE International Symposium on Circuits and Systems (ISCAS), 2025: 1–5.
- [85] ZHANG Q R, FAN Z C, AN H, et al. RoboVisio: a micro-robot vision domain-specific SoC for autonomous navigation enabling fully-on-chip intelligence via 2-MB eMRAM [J]. *IEEE Journal of Solid-State Circuits*, 2024, 59(8): 2644–2658.
- [86] CHIU Y C, KHWA W S, YANG C S, et al. A CMOS-integrated spintronic compute-in-memory macro for secure AI edge devices [J]. *Nature Electronics*, 2023, 6(7): 534–543.
- [87] KIM T, JANG Y, KANG M G, et al. SOT-MRAM digital PIM architecture with extended parallelism in matrix multiplication[J]. *IEEE Transactions on Computers*, 2022, 71(11): 2816–2828.
- [88] HUMOOD K, FOSTER P, WANG S W, et al. SPIKA: 200-TOPS/W RRAM-based neural network accelerator chip [C]//Proceedings of 2025 IEEE International Symposium on Circuits and Systems (ISCAS), 2025: 1–5.
- [89] SPETALNICK S D, CHANG M Y, KONNO S, et al. A 40-nm compute-in-memory macro with RRAM addressing IR drop and off-state current [J]. *IEEE Solid-State Circuits Letters*, 2024, 7: 10–13.
- [90] HUO Q, YANG Y M, WANG Y M, et al. A computing-in-memory macro based on three-dimensional resistive random-access memory[J]. *Nature Electronics*, 2022, 5(7): 469–477.
- [91] YAO P, WU H Q, GAO B, et al. Fully hardware-implemented memristor convolutional neural network [J]. *Nature*, 2020, 577(7792): 641–646.
- [92] JOSHI V, LE GALLO M, HAEFELI S, et al. Accurate deep neural network inference using computational phase-change memory[J]. *Nature Communications*, 2020, 11(1): 2473.
- [93] WANG R, SHI T, ZHANG X M, et al. Implementing in-situ self-organizing maps with memristor crossbar arrays for data mining and optimization [J]. *Nature Communications*, 2022, 13: 2289.
- [94] KHWA W S, WEN T H, HSU H H, et al. A mixed-precision memristor and SRAM compute-in-memory AI processor [J]. *Nature*, 2025, 639(8055): 617–623.
- [95] LI C, ZHANG X M, CHEN P, et al. Short-term synaptic plasticity in emerging devices for neuromorphic computing [J]. *iScience*, 2023, 26(4): 106315.
- [96] LI C, YU J, ZHANG X M, et al. Experimental demonstration of a CT-FeFET array with intrinsic long-short-term plasticity for low-cost trajectory prediction [C]//Proceedings of 2024 IEEE International Electron Devices Meeting (IEDM), 2024: 1–4.
- [97] ZHANG X M, WU Z H, LU J K, et al. Fully memristive SNNs with temporal coding for fast and low-power edge computing [C]//Proceedings of 2020 IEEE International Electron Devices Meeting (IEDM), 2020: 29.6.1–29.6.4.
- [98] ZHANG X M, ZHUO Y, LUO Q, et al. An artificial spiking afferent nerve based on Mott memristors for neurorobotics [J]. *Nature Communications*, 2020, 11: 51.
- [99] YANG Y, ZHU F D, ZHANG X M, et al. Firing feature-driven neural circuits with scalable memristive neurons for robotic obstacle avoidance [J]. *Nature Communications*, 2024, 15: 4318.
- [100] GAO Q Q, MENG H T, WU E X, et al. A flexible, broadband optoelectronic memristor for neuromorphic visual perception and intelligent image processing [J]. *Advanced Functional Materials*, 2026, 36(19): e20981.
- [101] YE H B, LIU Z Y, SUN B, et al. Optoelectronic resistive memory based on lead-free Cs₂AgBiBr₆ double perovskite for artificial self-storage visual sensors [J]. *Advanced Electronic*

- Materials, 2023, 9(2): 2200657.
- [102] HUANG H Y, LIANG X P, WANG Y Y, et al. Fully integrated multi-mode optoelectronic memristor array for diversified in-sensor computing[J]. Nature Nanotechnology, 2025, 20(1): 93–103.
- [103] KIM J S, NAM Y, KIM D Y, et al. Forming-free chemi-memristive gas sensing for artificial olfactory system[J]. Sensors and Actuators B: Chemical, 2025, 444: 138315.
- [104] ZRINSKI I, KNAPIC D, HASSEL A W, et al. Anodic HfO_2 crossbar arrays for hydroxide-based memristive sensing in liquids[J]. Journal of Electrochemical Science and Engineering, 2023, 13(5): 805–815.
- [105] DUTTA M, MAIKAP S, QIU J T. Controlling conductive filament and tributyrin sensing using an optimized porous iridium interfacial layer in $\text{Cu}/\text{Ir}/\text{TiN}_x\text{O}_y/\text{TiN}$ [J]. Advanced Electronic Materials, 2019, 5(2): 1800288.
- [106] KIM H, IM I H, HYUN D, et al. Stretchable resistive switching memory devices for wearable systems[J]. Journal of the Korean Ceramic Society, 2025, 62(3): 397–411.
- [107] SIM K, RAO Z L, ZOU Z N, et al. Metal oxide semiconductor nanomembrane-based soft unnoticeable multifunctional electronics for wearable human-machine interfaces[J]. Science Advances, 2019, 5(8): eaav9653.
- [108] HUANG Y, HE C Y, LING Y Z, et al. Radio frequency signal processing with a memristive system-on-a-chip[J]. Nature Electronics, 2025, 8(7): 587–596.
- [109] FU T D, LIU X M, FU S, et al. Self-sustained green neuromorphic interfaces [J]. Nature Communications, 2021, 12: 3351.
- [110] CHOI S, KIM Y, VAN NGUYEN T, et al. Low-power self-rectifying memristive artificial neural network for near Internet-of-Things sensor computing [J]. Advanced Electronic Materials, 2021, 7(6): 2100050.
- [111] MIN K P, LI C Y, WANG T J, et al. Performance improvement of AlN -based RRAMs using Ag layer for hardware security applications [J]. IEEE Transactions on Electron Devices, 2023, 70(8): 4115–4121.
- [112] GAO B, LIN B H, PANG Y C, et al. Concealable physically unclonable function chip with a memristor array[J]. Science Advances, 2022, 8(24): eabn7753.
- [113] QIN Y B, WANG Z W, YANG Y F, et al. A high-speed true random number generator based on unified selector-RRAM[J]. IEEE Electron Device Letters, 2023, 44(12): 1967–1970.
- [114] JIANG H, LI C, ZHANG R, et al. A provable key destruction scheme based on memristive crossbar arrays[J]. Nature Electronics, 2018, 1(10): 548–554.
- [115] LI X Q, GAO B, QIN Q, et al. Federated learning using a memristor compute-in-memory chip with in situ physical unclonable function and true random number generator[J]. Nature Electronics, 2025, 8(6): 518–528.
- [116] CHEONG W H, IN J H, JEON J B, et al. Stochastic switching and analog-state programmable memristor and its utilization for homomorphic encryption hardware[J]. Nature Communications, 2024, 15: 6318.
- [117] WANG C, RUAN G J, YANG Z Z, et al. Parallel in-memory wireless computing[J]. Nature Electronics, 2023, 6(5): 381–389.
- [118] PI S, GHADIRI-SADRABADI M, BARDIN J C, et al. Nanoscale memristive radio frequency switches[J]. Nature Communications, 2015, 6: 7519.