

· 科学论坛 ·

面向神经形态计算的智能芯片与器件技术*

王宗巍^{1†} 杨玉超^{1,2†} 蔡一茂^{1,2**} 朱涛³
丛杨⁴ 王志衡⁵ 黄如^{1,2**}

1. 北京大学微纳电子学系, 北京 100871; 2. 北京大学人工智能研究院, 北京 100871;
3. 重庆大学光电工程学院, 重庆 400044; 4. 中国科学院沈阳自动化研究所, 沈阳 110016;
5. 国家自然科学基金委员会信息科学部, 北京 100085

[摘要] 基于第216期“双清论坛”关于人工智能芯片发展的主题报告和分组讨论, 本文着重介绍了面向神经形态计算的智能芯片与器件技术的国内外主要研究进展及存在的挑战, 包括神经网络加速器、基于传统CMOS的神经形态智能芯片、新型神经形态器件技术及基于新器件的神经形态芯片等内容, 探讨了神经形态智能芯片面临的关键问题和未来发展趋势。

[关键词] 人工智能; 类脑计算; 神经形态芯片; 神经形态器件



王宗巍 2018年获得北京大学博士学位。北京大学微纳电子学研究院博士后/讲师, 电气和电子工程师协会会员。研究领域包括新型存储器、神经形态器件及类脑芯片技术。在IEDM、IEEE-EDL等微电子领域顶级会议和期刊发表学术论文30余篇, 获得中国发明专利授权10余项。



杨玉超 北京大学微纳电子学系研究员。长期从事类脑计算研究工作, 主持国家杰青、重点研发计划等研究项目, 共发表 *Nature Electronics*、*Nature Communications* 等论文70余篇, 撰写中英文专著5章。现任中国电子学会青工委委员、《中国科学: 信息科学》青年编委、*Scientific Reports* 编委。曾获求是杰出青年学者奖、Wiley 青年研究者奖、《麻省理工科技评论》中国区35岁以下科技创新35人、科学探索奖等奖项。



蔡一茂 北京大学微纳电子学系教授, 博士生导师。北京市人工智能研究院智源学者, 北京大学人工智能研究院骨干成员。曾获教育部“长江学者奖励计划”青年学者资助。长期从事神经形态器件及类脑芯片技术研究, 主持包括国家自然科学基金重点项目在内的多项国家级项目, 发表学术论文百余篇。



黄如 北京大学微纳电子学系教授, 中国科学院院士, IEEE Fellow, 国家自然科学基金委创新群体带头人。长期从事半导体新器件及其应用研究。已合作出版著作5本, 发表学术论文300余篇。曾获国家技术发明二等奖、国家科技进步二等奖、北京市科学技术一等奖、教育部自然科学一等奖、教育部科技进步一等奖、中国青年科技奖等多项国家和部委级奖励。

自20世纪50年代末杰克·基尔比(Jack Kilby)和罗伯特·诺伊斯(Robert Noyce)发明集成电路以来, 集成电路的集成度随着摩尔定律持续提高, 推动着传统冯·诺依曼架构计算机的性能快速提升。冯·诺依曼体系架构的基本特征为: 处理器单元与存储器模块分离, 通过预先存储的不同软件程序依指令顺序执行。这一特点使存储和计算单元间的数据频繁交换成为掣肘, 产生“存储墙”问题, 并且无法针对应用场景和需求的变化进行实时、自适应的自我学习演化, 难以实现智能化的数据和信息处理。

与此同时, 随着现代社会迈入信息化、智能化时代, 计算芯片的核心负载呈现多样化和智能化趋势,

收稿日期: 2019-06-04; 修回日期: 2019-10-18

* 本文根据第216期“双清论坛”讨论内容整理。

** 通信作者, Email: caiyimao@pku.edu.cn; ruhuang@pku.edu.cn

† 贡献相同

以自动驾驶、智能机器人、图像/语音识别、自然语言翻译和数据挖掘为代表的一系列智能应用逐渐转变为主流负载。这一系列新的应用场景进一步使冯·诺依曼架构的“存储墙”问题和自适应性差的问题凸显。特别是近年来随着摩尔定律面临尺寸缩小的物理极限,集成电路产业工艺节点推进放缓,单核处理器主频快速提高的进程基本停滞,难以满足未来智能终端和智能计算平台面向移动端的应用场景,学术界和业界开始寻求替代冯·诺依曼体系架构的新型计算范式。

早在 1948 年,图灵就提出“类脑计算”的构想,即设计一种可以通过训练实现自我学习和演化的计算硬件。早期人工神经网络研究借鉴了生物神经网络中的神经元、突触连接性、网络结构等基本概念,在此基础上进行简化,将生物神经网络抽象成由大规模“突触”连接的“神经元”构成的类似于生物大脑的神经网络,进而在网络拓扑和算法层面上实现对生物神经网络的模拟,其工作原理与生物神经网络存在较大的差异^[1]。

与传统的冯·诺依曼体系架构不同,人工智能芯片往往采用模拟人脑神经网络对信息感知和决策的方式进行信息的收集、传输、处理和存储,更适于在硬件平台上实现信息分布式计算与存储一体化,特别是在多感官跨模态等非结构化数据与智能任务处理应用场景下(图像、语音识别和自动驾驶),其在功耗、能效、硬件开销方面的优势显著,有望成为大数据及人工智能时代应对海量实时数据的颠覆性计算范式。因此,人工智能芯片技术受到了学术界和产业界的广泛关注。

1 人工神经网络算法加速芯片

人工智能正引领新的时代变革,近年来深度学习在图像和语音识别等领域取得突破性进展。基于人工神经网络的深度学习迅速崛起,在学术界与产业界的共同努力下,取得了诸多里程碑式的突破,从事深度学习研究的三位先驱也因此获得了 2018 年的图灵奖。

在深度学习基础上,谷歌 2014 年收购的 DeepMind 团队开发了阿尔法围棋(AlphaGo),于 2016 年 3 月和 2017 年 9 月分别战胜了围棋职业九段选手和围棋排名世界第一的选手。然而,AlphaGo 本质上还是运行在传统计算平台上的深度强化学习软件算法,其在对弈时使用了超过 1 000 个通用处理器(CPU)和近 200 个图形处理器(GPU)

来执行围棋对弈这一单一任务,相对于人脑即使在执行多样任务时的功耗水平,AlphaGo 还有巨大差距。由此可见,传统集成电路芯片及计算机采用的信息处理及交换方式是制约人工智能技术发展的重要技术瓶颈之一。

在传统架构基础上衍生出的人工神经网络算法加速芯片应需出现,主要针对现有深度学习等较为成熟的人工神经网络算法,通过 FPGA 或者 ASIC 方式设计芯片模块或者采用分布式存储的整体架构提高能效比和速度。代表性成果有 Google 研制的 TPU、麻省理工学院研制的卷积神经网络加速芯片 Eyeriss 以及我国中科院计算所研制的支持深度神经网络的“寒武纪”芯片等。这类智能芯片的特点是针对人工神经网络算法的关键操作进行硬件固化或者加速,相比传统 CPU 和 GPU,具有速度更快和功耗更低的优势,在近期取得了较为广泛的实际应用进展。但是这类芯片通常面向特定领域的专门应用,并且需要大量的数据训练神经网络,其较长的设计迭代周期也增加了芯片的开发应用成本。

2 神经形态智能芯片

相比传统计算机,人脑无疑是更强大的智能平台,具有在复杂甚至陌生场景条件下的自适应、获取新信息和新技能的能力,以及进行交互、推理从而做出决策的能力。同时,人脑也更高效、可靠,可以在完成上述功能的同时保持低功耗和鲁棒性。总体来说,人脑的生物神经网络具备以下优势:强自主学习能力、高并行性、高容错性、高连接性,以及存算融合和低功耗特性。其中,存算融合特性从根本上解决了冯·诺依曼体系架构的“存储墙”问题。

“类脑芯片”或者“神经形态芯片”是以模拟人脑神经网络计算为基础的一种新型芯片。区别于传统软件算法层面的人工神经网络,神经形态计算芯片是以神经科学理论和生物学实验结果为依据,综合认知科学和信息科学,参考生物神经网络模型和架构,采用现有 CMOS 器件/电路或者新型神经形态器件模拟生物神经元和突触的信息处理特性,构建以人脑神经网络为蓝本的具备信息感知、处理和學習等功能为一体的智能化计算平台。神经形态计算芯片根据实现技术路径主要可分为:基于传统 CMOS 的神经形态芯片和基于新型神经形态器件的神经形态芯片两种类型。其中,基于新型器件的神经形态芯片具有更低功耗、更小硬件代价、自适应、自学习、自演化、高容错等显著优势,但目前尚处于

探索性应用阶段。

2.1 基于传统 CMOS 器件和电路的神经形态计算芯片

自 2004 年以来,各国政府及研究机构投入巨资开展“类脑计算”研究。国内外一系列“脑计划”的开展成功推动了基于传统成熟 CMOS 工艺技术的神经形态芯片研究(表 1)。这类神经形态芯片虽然区别于传统集成电路芯片和人工神经网络加速芯片,但构建芯片的基本结构——人工神经元及其连接人工突触仍然采用基于 CMOS 的数字电路或者数模混合电路来搭建,要模拟单个神经元或者突触的行为往往要靠多个 CMOS 器件组成电路模块来实现,集成密度、功耗和功能模拟准确度都受到限制。尤其是在 CMOS 尺寸已经缩小到接近物理极限的情况下,依赖先进工艺的芯片能够构建的类脑神经网络中神经元和突触数量还是远远小于人脑的规模。

表 1 国内外神经形态芯片主要进展

芯片	核心器件/电路	工艺 (nm)	神经元个数	突触个数
TrueNorth ^[2]	数字电路	28	100 万	2.56 亿
SpiNNaker ^[2]	18 核 ARM 芯片,片上网络互联	130	1.6 万	约 1600 万
BrainScaleS ^[3]	数模混合电路,晶片级集成	180	每片晶圆上约 20 万	每片晶圆上约 4 900 万
Loihi ^[4]	数字电路	14	13.1 万	13.1 亿
天机 2 代 ^[5]	晶体管	28	约 40 000	约 1 000 万
达尔文 ^[6]	晶体管	180	2 048 个神经元	400 万个突触

这些成为限制此类神经形态芯片实现更复杂、更高水平人工智能技术的重要障碍。

2.2 新型神经形态器件及神经形态智能芯片

不同于基于传统 CMOS 器件的神经形态计算芯片,基于新型神经形态器件的神经形态芯片从底层器件仿生的角度出发,在器件层面即开始模拟生物的基本信息处理单元——神经元和突触,在功耗、硬件代价等方面具有显著优势,但目前尚处在探索阶段。现阶段,国内外研究机构已经基于忆阻器(Memristor)、相变存储器(Phase Change Memory)、铁电器件(Ferroelectric Device)、磁隧道结(Magnetic Tunnel Junction)、离子栅控晶体管等新型器件实现了模拟突触和神经元的功能并取得了重要进展^[7-34]。

在诸多新型人工突触器件中,忆阻器凭借其结

构简单、集成密度高和非易失性等一系列优势成为神经形态智能芯片中突触器件的重要候选技术,得到了众多研究者的重点关注。2008 年惠普实验室在《自然》杂志上报道首次实验实现了由华裔科学家蔡少棠(Leon O. Chua)提出的忆阻器^[7]。美国密歇根大学研究人员在 2010 年率先基于银掺杂非晶硅的忆阻器实现了突触可塑性的模拟,展示了忆阻器在神经形态芯片中的应用前景^[8]。然而如何获得突触权值良好的线性可调特性以及长/短时程可塑性、脉冲时间依赖可塑性(STDP)等突触重要功能是基于忆阻人工突触实现神经形态芯片的核心挑战。北京大学研究团队提出了基于不同离子扩散速率的叠层结构过渡金属氧化物忆阻器,利用离子扩散速率的差异调控导电通道生长过程,成功实现了高线性度人工突触^[9],能够高效支持在线学习功能,同时有助于降低电路设计的难度和复杂度。

研究表明,生物神经网络中实现记忆和学习的基本机制与突触的长/短时程可塑性及脉冲时间依赖可塑性密切相关。北京大学研究人员利用忆阻器的两端电极模拟突触前膜和突触后膜,通过在忆阻器两端施加特定脉冲波形使得叠加后的信号与脉冲间隔时间相关,从而在实验上实现了突触的 STDP 特性^[9]。美国密歇根大学研究人员利用电场与热作用的耦合作用,将器件内部的局域温度弛豫作用作为时间信息编码来模拟 STDP 行为^[10]。华中科技大学课题组基于钙钛矿材料的二价忆阻器实现了生物突触中的三相 STDP 规则,可以用于更加复杂的模式识别和轨迹追踪^[11]。国内外多个课题组相继报道了基于忆阻器人工突触实现 STDP 等突触可塑性,这些特性研究成功使神经形态芯片的实现成为可能。

人脑的记忆并非一成不变,既有通过频繁重复使用以增强的长时记忆,也有快速遗忘的短时记忆,两者间的互相转换表现在突触上就是长时程可塑性与短时程可塑性的可逆转变,对于生物体适应环境有着重要意义。日本国立材料科学研究所(NIMS)采用基于电化学原理的忆阻器,利用脉冲频率的变化成功实现了人工突触的长时程可塑性与短时程可塑性的可逆转变^[12]。马萨诸塞大学的研究人员利用扩散型忆阻器本身的“易失性”成功实现了短时程的突触可塑性,这一“易失性”更加符合生物短时程可塑性的机制,进一步丰富了电子突触的功能^[13]。

基于忆阻器人工突触器件及其网络实现生物神经网络中的非联想及联想学习行为是神经形态芯片

实现类脑功能的重要探索。北京大学研究人员基于单个多端忆阻器件成功模拟了在生物神经网络中多个突触和神经元共同作用才能实现的习惯化及敏化等非联想学习功能^[14]，并针对神经形态器件中普遍存在的突触权值改变速率难调问题实现了具有自适应学习速率的三端异源性人工突触，能够模拟对生物系统认知记忆等功能起重要作用的异源突触可塑性^[15]，为实现复杂仿生功能器件和神经形态计算奠定了基础。华中科技大学研究人员利用基于忆阻器的电路模拟巴甫洛夫条件反射实验，成功实现了学习与遗忘功能^[16]。除此之外，研究人员还基于忆阻器人工突触模拟了兴奋性突触后电流(EPSC)、双脉冲易化(PPF)等典型的突触特性，并实现了联想学习等功能。

人工突触还可以基于多种其他新型器件技术实现。研究表明，相变存储器依赖材料相变实现电导变化的物理机制使其模拟突触权值连续增加的变化过程具有显著优势。美国斯坦福大学已基于相变器件实现了多种形式的 STDP 学习规则^[17]，但相变材料非晶化过程趋于突变，难以较好地实现突触权值连续减小的过程。法国研究机构 CEA-LETI 通过将两个相变存储器组成的差分单元整体等效为一个突触的方案缓解了相变材料非晶化过程导致的突触权值陡变问题^[18]。

基于磁隧道结也可以模拟突触可塑性。磁隧道结的自由层磁化方向可以通过外加电流产生的磁场或者自旋转移矩进行调控，当磁化方向与参考层磁化方向同向时表现为低阻态，反之，则表现出高阻态。美国普渡大学研究人员发现可以通过调控磁隧道结的外加激励时间、幅度和频率实现突触长/短时程可塑性的高精度模拟^[19]。德国比勒费尔德大学研究人员基于磁隧道结实现了突触的 STDP 特性^[20]。

此外，传统晶体管结构与新材料的结合也可用来模拟突触特性，实现新原理人工突触。南京大学研究人员基于氧化物双电层材料研制了双电层晶体管，并利用双电层晶体管人工突触成功实现了长/短时程可塑性和 STDP 等学习法则^[21]。北京大学研究人员设计实现了兼具高仿生精度、高线性度、高对称性的离子栅控型人工突触，可以在单个器件中实现多种短时程、长时程突触可塑性及二者之间的可控转变，并实现了与生物突触相当水平的超低能耗(单脉冲 30fJ)，在线性度、对称性、能耗等综合性能方面达到了国际先进水平^[22]。

总之，大脑神经网络中突触可塑性是实现感知、

学习及记忆等高等智能的关键环节。除了上述提到的线性度、对称性、权值精度等权值量化特性和长/短时程可塑性之外，为了可以利用人工突触构建大规模神经网络，衡量人工突触器件性能的指标还包括：可微缩性、一致性、能耗、开关比、保持特性、耐久特性等。基于不同类型器件的特点可以有针对性地优化特定的突触性能，但目前尚无一种器件可以很好地达到所有的指标，因此研究可以低功耗、高精度模拟上述可塑性的电子突触器件对于构建新型神经计算芯片具有重大意义。

实现新型神经形态芯片的另一重要基本元件是可以模拟神经元功能的神经形态器件。相比于突触神经形态器件的研究，神经元神经形态器件的研究进展则要缓慢得多，其中一个主要原因是生物神经元的多样化发放行为难以通过现有器件以低功耗、低硬件成本的方式精准模拟。2012 年，惠普实验室的研究人员在《自然·材料》杂志上报道了基于莫特材料的忆阻器和传统基本电路元件模拟 Hodgkin-Huxley(HH)神经元模型中的钠离子和钾离子通道调节膜电位的功能，从而实现神经元的发放功能，更易于降低神经元的硬件开销，为神经元器件的实现提供了新思路^[23]。IBM 苏黎世研究中心于 2016 年展示了随机相变神经元电路，用相变材料的结晶化过程代替传统 CMOS 神经元的电荷积累过程模拟神经元的积累发放特性，并能有效模拟生物神经元带来的固有随机性^[24]。北京大学在 2018 年基于忆阻器和晶体管级联结构，用很少的硬件代价实现了视网膜神经元的积累发放功能，并搭建了基于此神经元的视网膜边缘检测网络，成功实现了图像处理功能^[25]。

神经形态器件的引入为构建低功耗、高效的神经网络奠定了基础。基于新型神经形态器件的神经形态计算芯片也已经开始崭露头角。加州大学圣芭芭拉分校于 2015 年在《自然》杂志上报道了以金属氧化物忆阻器作为权值阵列的在线学习神经网络，实现了单层感知机算法，能够识别 3×3 像素黑白图像中的图案，可以实现高度并行从而提高计算速度，且利用忆阻器的非易失性存储大大降低了功耗^[26]。美国密歇根大学于 2017 年研制了利用忆阻器神经形态器件与 CMOS 混合集成的原型芯片，可实现稀疏编码，进而实现复杂图像的编码与重构^[27]，并基于忆阻型突触的短时程可塑性结合库网络算法，将时序输入信息映射为高维空间特征信息，实现手写体数字识别^[28]。针对忆阻突触器件涨落性影响神

神经网络功能的共性难题,北京大学提出并实现了具有高容错能力的模糊受限玻尔兹曼机神经网络,提高了网络对于器件涨落性的容忍度,从而提供了通过器件—网络协同设计实现高性能神经网络的新思路和新方法^[29]。马萨诸塞大学在2017年基于1T1R忆阻器阵列成功实现了模拟信号处理和图像压缩处理等功能^[30],并在2018年利用之前提出的扩散性忆阻器实现了神经元积累释放功能,与基于忆阻器的突触阵列相结合实现了全忆阻神经网络,成功实现了无监督学习和图像识别功能^[31]。

3 总结与展望

人工智能是引领新一轮技术革命和产业变革的核心驱动力。随着社会对各个行业智能化产业升级的需求不断增加,未来神经形态计算有着广阔的应用前景。

当前类脑神经形态计算的发展取得了可喜进展和突破,但受限于现阶段对人脑及其复杂神经网络系统的认识,从信息学和工程学角度的仿脑研究仍处于初级阶段,必须通过神经科学、认知科学、计算机科学、微电子学等多学科领域更加紧密的交流与合作才能更进一步地推动对人脑和其神经网络系统的工作原理与模式的深入解析与模仿。需要指出的是,目前基于全神经形态器件的神经形态计算芯片尚处在探索阶段,在器件、模型、架构和算法层面还面临诸多挑战。

(1) **神经形态器件层面**:材料、机理和器件设计的一体化研究体系。一方面,材料的选择将直接影响神经形态器件的性能,需要筛选能够支撑高性能神经形态器件的材料;另一方面,机理的选择和澄清有利于对神经形态器件进行针对性地设计和优化;此外,针对不同应用场景设计神经形态器件是未来可能的发展趋势和研究方向。

(2) **模型层面**:基于生物模型的精确仿生多样化人工神经元和突触模型。当前神经形态芯片中的神经元和突触模型通常较为单一,其中神经元模块只实现了简化神经元模型,尚无法模拟及描述生物神经网络中的区块协同功能。同时,当前神经形态芯片中突触局限于恒定突触强度或单一时程特性的模型,难以协调不同时空位置神经元的连接性,无法实现不同模式输入—输出的非线性整合。

(3) **系统架构层面**:具备自学习能力、鲁棒性和低功耗的可重构系统架构。首先,在未来智能应用场景下,硬件需能够根据环境变化实时调整自身参

数,实现硬件的可编程性;其次,容错能力是智能芯片适应极端环境的保障,在一些复杂环境或存在人为干扰的应用场景下,要求在架构层面提升抗干扰、噪声等的的能力,实现硬件的鲁棒性;同时,功耗约束是智能芯片适应低功耗场景的基本要求。另外,现有架构仍存在信息表达低效和冗余计算较多等问题,需要架构创新来支持更加高效的编码和处理方式、模型可扩展性。

(4) **算法层面**:可映射到硬件神经网络的高效信息处理算法。当前脉冲神经网络训练算法的理论发展还不够成熟,适合硬件实现的算法有待开发,为实现神经元群编码、工作记忆、空间导航、预测跟踪等功能提供有效的计算模型和理论基础。同时,数据约束是智能芯片面临的新难题,特定场景的智能任务不具有大样本数据集,存在数据限制和缺失的问题,因此算法要具有小样本和模糊学习能力。

(5) **脑机混合智能层面**:神经形态智能芯片与生物体的有机结合和智能交互。基于神经形态计算芯片的脑机融合混合智能系统有望实现生物感知、认知基本单元与智能芯片计算功能单元之间的多维深度融合。目前针对脑机融合的混合智能研究在基础理论框架、作用机制等方面仍面临很多挑战,需要进一步探索。

参 考 文 献

- [1] Mitchell TM. Machine learning. New York, USA: the McGraw-Hill companies, 1997.
- [2] Nawrocki RA, Voyles RM, Shaheen SE. A mini review of neuromorphic architectures and implementations. *IEEE Transactions on Electron Devices*, 2016, 63 (10): 3819—3829.
- [3] Meier K. A mixed-signal universal neuromorphic computing system. In 2015 IEEE International Electron Devices Meeting (IEDM) 2015 Dec 7, pp. 4—6.
- [4] Davies M, Srinivasa N, Lin TH, et al. Loihi: a neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 2018, 38(1):82—99.
- [5] Shi LP, Pei J, Deng N, et al. Development of a neuromorphic computing system. In 2015 IEEE International Electron Devices Meeting (IEDM) 2015 Dec 7, pp. 4—3.
- [6] Shen JC, Ma D, Gu ZH, et al. Darwin: a neuromorphic hardware co-processor based on spiking neural networks. *Science China-Information Sciences*, 2016, 59(2): 1—5.
- [7] Strukov DB, Snider GS, Stewart DR, et al. The missing memristor found. *Nature*, 2008, 453(7191): 80.

- [8] Jo SH, Chang T, Ebong I, et al. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Letters*, 2010, 10(4): 1297—1301.
- [9] Wang Z, Yin M, Zhang T, et al. Engineering incremental resistive switching in TaOx based memristors for brain-inspired computing. *Nanoscale*, 2016, 8 (29): 14015—14022.
- [10] Chang T, Jo SH, Lu W. Short-term memory to long-term memory transition in a nanoscale memristor. *ACS Nano*, 2011, 5(9): 7669—7676.
- [11] Yang R, Huang HM, Hong QH, et al. Synaptic suppression triplet-STDP learning rule realized in second-order memristors. *Advanced Functional Materials*, 2018, 28 (5): 1704455.
- [12] Ohno T, Hasegawa T, Tsuruoka T, et al. Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nature Materials*, 2011, 10(8): 591.
- [13] Wang ZR, Joshi S, Savel'ev SE, et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nature Materials*, 2017, 16(1): 101.
- [14] Yang X, Fang YC, Yu ZZ, et al. Nonassociative learning implementation by a single memristor-based multi-terminal synaptic device. *Nanoscale*, 2016, 8(45): 18897—189904.
- [15] Yang YC, Yin MH, Yu ZZ, et al. Multifunctional nanoionic devices enabling simultaneous heterosynaptic plasticity and efficient in-memory boolean logic. *Advanced Electronic Materials*, 2017, 3(7): 1700032.
- [16] Wang ZL, Wang XP. A novel memristor-based circuit implementation of full-function Pavlov associative memory accorded with biological feature. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2018, 65 (7): 2210—2220.
- [17] Kuzum D, Jeyasingh RG, Lee B, et al. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Letters*, 2011, 12 (5): 2179—2186.
- [18] Suri M, Bichler O, Querlioz D, et al. Phase change memory as synapse for ultra-dense neuromorphic systems; Application to complex visual pattern extraction. In 2011 International Electron Devices Meeting 2011 Dec 5, pp. 4—4.
- [19] Sengupta A, Roy K. Short-term plasticity and long-term potentiation in magnetic tunnel junctions; Towards volatile synapses. *Physical Review Applied*, 2016, 5(2): 024012.
- [20] Krzysteczko P, Münchenberger J, Schäfers M, et al. The memristive magnetic tunnel junction as a nanoscopic Synapse-Neuron system. *Advanced Materials*, 2012, 24(6): 762—766.
- [21] Zhou JM, Liu N, Zhu LQ, et al. Energy-efficient artificial synapses based on flexible IGZO electric-double-layer transistors. *IEEE Electron Device Letters*, 2015, 36(2): 198—200.
- [22] Zhu JD, Yang YC, Jia RD, et al. Ion gated synaptic transistors based on 2D van der Waals crystals with tunable diffusive dynamics. *Advanced Materials*, 2018, 30 (21): 1800195.
- [23] Pickett MD, Medeiros-Ribeiro G, Williams RS. A scalable neuristor built with Mott memristors. *Nature Materials*, 2013, 12(2): 114—117.
- [24] Teeter C, Iyer R, Menon V, et al. Generalized leaky integrate-and-fire models classify multiple neuron types. *Nature Communications*, 2018, 9(1): 709.
- [25] Bao L, Kang J, Fang Y, et al. Artificial shape perception retina network based on tunable memristive neurons. *Scientific Reports*, 2018, 8(1): 13727.
- [26] Prezioso M, Merrih-Bayat F, Hoskins BD, et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature*, 2015, 521(7550): 61—64.
- [27] Sheridan PM, Cai FX, Du C, et al. Sparse coding with memristor networks. *Nature Nanotechnology*, 2017, 12(8): 784—789.
- [28] Du C, Cai FX, Zidan MA, et al. Reservoir computing using dynamic memristors for temporal information processing. *Nature Communications*, 2017, 8(1): 2204.
- [29] Zhang T, Yin M, Lu X, et al. Tolerance of intrinsic device variation in fuzzy restricted Boltzmann machine network based on memristive nano-synapses. *Nano Futures*, 2017, 1 (1): 015003.
- [30] Li C, Hu M, Li YN, et al. Analogue signal and image processing with large memristor crossbars. *Nature Electronics*, 2018, 1(1): 52.
- [31] Wang ZR, Joshi S, Savel'ev S, et al. Fully memristive neural networks for pattern classification with unsupervised learning. *Nature Electronics*, 2018, 1(2): 137—145.
- [32] Wang HM, Duan SK, Huang TW, et al. Exponential stability of complex-valued memristive recurrent neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 28(3): 766—771.
- [33] Duan SK, Wang HM, Wang LD, et al. Impulsive effects and stability analysis on memristive neural networks with variable delays. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 28(2): 476—481.
- [34] Duan SK, Hu XF, Dong ZK, et al. Memristor-based cellular nonlinear/neural network; design, analysis, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 26(6): 1202—1213.

Intelligent chip and device technology for neuromorphic computing

Wang Zongwei¹ Yang Yuchao^{1,2} Cai Yimao^{1,2} Zhu Tao³
Cong Yang⁴ Wang Zhiheng⁵ Huang Ru^{1,2}

(1. Institute of Microelectronics, Peking University, Beijing 100871;

2. Academy for Artificial Intelligence, Peking University, Beijing 100871;

3. College of Optoelectronic Engineering, Chongqing University, Chongqing 400044;

4. Shenyang Institute of Automation, Chinese Academy of Science, Shenyang 110016;

5. Department of Information Sciences, National Natural Science Foundation of China, Beijing 100085)

Abstract Recapitulating from the discussion upon development of artificial intelligent chip in the 216th Shuangqing Forum, the recent advances and scientific challenges in intelligent chip and device technology for neuromorphic computing are analyzed in this paper, including artificial neural network accelerator, neuromorphic chip based on conventional CMOS technology, emerging neuromorphic devices and neuromorphic chips. The key challenges and future development trends for neuromorphic chip are outlined.

Key words artificial intelligence; brain-inspired computing; neuromorphic chip; neuromorphic device

《中国科学基金》征稿简则

《中国科学基金》是由国家自然科学基金委员会主管、主办的综合性学术期刊,已被北大核心、CSCI、CSCD及日本《科学技术文献速报》等国内外重要检索系统收录。

欢迎学术思想新颖、观点明确、有学术水平和对科学基金工作者有指导意义的论文和评述文章,尤其是有关科学基金项目的研究进展和基金管理的研讨性论文。

本刊常设栏目:

管理纵横:探讨和交流科学基金项目申请、评审、管理和改革举措等方面的政策研究与深度解读。

研究进展:刊登有关学科的具有战略性、全局性、前瞻性的综述性和评论性文章,以促进学科间的了解、交叉与融合。

科学论坛:围绕科技界普遍关注的热点与焦点问题,如基础研究、研究评价、同行评议、科研诚信、科研管理、科学精神等展开讨论。

成果快报:报道和选登重要的、有影响的、具有代表性的科学基金项目研究成果以及优秀人才和优秀群体介绍。

稿件要求:

稿件应以 Word 文档提交,“研究进展”类综述性文章(8 000 字左右)或“成果快报”类文章(3 000 字左右),投稿时请提交相关的高分辨率彩色图片一张,以供我们遴选作为《中国科学基金》的封面文章使用。

不同栏目稿件的体例格式应严格遵循本刊的相关要求,文章采用通栏排版,图表清晰并能够达到排版印刷的要求,简化字以正式公布的为准,公式符号规范,力求内容充实精炼。文章正文前附中英文题目、作者、单位、摘要(300 字以内)及关键词。稿件应附通信作者的真实姓名、工作单位、职务、通讯地址、邮政编码、电话和电子邮件地址,以及第一作者和通讯作者的 1 寸照片和个人简介(不超过 200 字)。

文章请勿一稿两投,本刊自收到稿件之日起,一个月内将处理结果通知作者,如逾期未予答复,作者有权另行处理,但需及时通知本编辑部。

投稿网址: <http://zkj.cbpt.cnki.net/>

期刊网址: <http://www.nsf.gov.cn/csc/20345/20348/index.html>

联系电话: 010-62327089

编辑部地址: 北京市海淀区双清路 83 号(邮政编码: 100085)