

基于磁性隧道结的群体编码实现无监督聚类

张亚君 蔡佳林 乔亚 曾中明 袁喆 夏钊

Implementation of unsupervised clustering based on population coding of magnetic tunnel junctions

Zhang Ya-Jun Cai Jia-Lin Qiao Ya Zeng Zhong-Ming Yuan Zhe Xia Ke

引用信息 Citation: *Acta Physica Sinica*, 71, 148506 (2022) DOI: 10.7498/aps.71.20220252

在线阅读 View online: <https://doi.org/10.7498/aps.71.20220252>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

MgO基磁性隧道结温度-偏压相图的理论研究

Theoretical study on temperature-bias phase diagram of MgO-based magnetic tunnel junctions

物理学报. 2021, 70(10): 107302 <https://doi.org/10.7498/aps.70.20201905>

NbO_x 忆阻神经元的设计及其在尖峰神经网络中的应用

Design of NbO_x memristive neuron and its application in spiking neural networks

物理学报. 2022, 71(11): 110501 <https://doi.org/10.7498/aps.71.20220141>

基于工艺偏差的电压调控磁各向异性磁隧道结电学模型及其在读写电路中的应用

Process deviation based electrical model of voltage controlled magnetic anisotropy magnetic tunnel junction and its application in read/write circuits

物理学报. 2020, 69(19): 198502 <https://doi.org/10.7498/aps.69.20200228>

基于工艺偏差的自旋转移矩辅助压控磁各向异性磁隧道结电学模型及其应用研究

Process deviation based electrical model of spin transfer torque assisted voltage controlled magnetic anisotropy magnetic tunnel junction and its application

物理学报. 2022, 71(10): 107501 <https://doi.org/10.7498/aps.71.20211700>

基于人工神经网络在线学习方法优化磁屏蔽特性参数

Online learning method based on artificial neural network to optimize magnetic shielding characteristic parameters

物理学报. 2019, 68(13): 130701 <https://doi.org/10.7498/aps.68.20190234>

基于深度神经网络的时空编码磁共振成像超分辨率重建方法

Super-resolved reconstruction method for spatiotemporally encoded magnetic resonance imaging based on deep neural network

物理学报. 2022, 71(5): 058702 <https://doi.org/10.7498/aps.71.20211754>

专题: 面向类脑计算的物理电子学

基于磁性隧道结的群体编码实现无监督聚类*

张亚君¹⁾ 蔡佳林²⁾ 乔亚¹⁾ 曾中明^{2)†} 袁喆^{1)‡} 夏钊³⁾

1) (北京师范大学物理学系, 高等量子研究中心, 北京 100875)

2) (中国科学院苏州纳米技术与纳米仿生研究所, 苏州 215123)

3) (北京计算科学研究中心, 北京 100193)

(2022 年 2 月 9 日收到; 2022 年 4 月 13 日收到修改稿)

利用新型材料器件发展类脑计算硬件研究的关键问题是发展出合适的算法, 能够发挥新器件的特点和优势. 群体编码是生物神经系统常见的编码方式, 能够有效去除噪音, 实现短时程记忆及复杂的非线性映射功能. 本文选择自旋电子学器件中研究较多、工艺较成熟的磁性隧道结, 应用其可调控的随机动力学实现群体编码. 作为一个应用的例子, 超顺磁隧道结构建的二层脉冲神经网络成功完成了鸢尾花数据集的无监督聚类. 数值仿真表明基于磁性隧道结的群体编码可以有效对抗器件的非均匀性, 为类脑计算硬件研究提供重要的参考.

关键词: 磁性隧道结, 群体编码, 脉冲神经网络, 无监督学习**PACS:** 85.70.-w, 07.05.Mh**DOI:** 10.7498/aps.71.20220252

1 引言

得益于神经网络算法的进步和新型计算硬件的发展, 人工智能在近十多年取得了巨大的进步. 其中在图像识别、自动驾驶和自然语言处理等方面的应用已经进入并逐步改变了人们的日常生活^[1]. 然而当前人工智能的发展仍然面临诸多问题. 一方面, 随着摩尔定律的终结, 处理器性能的增长开始放缓; 另一方面, 传统冯·诺依曼架构采用的存算分离系统, 使数据在中央处理器和内存之间搬运消耗了大量的时间和能量. 为了解决这些问题, 科学家试图从大脑等生物神经系统寻求解决方案. 大脑的神经系统具有存算一体、高度并行、事件驱动等特点^[2]. 目前已经有多款基于传统互补金属氧化物半导体 (complementary metal oxide semiconductor, CMOS) 工艺类脑芯片问世^[3,4], 展示了类脑芯片

在低能耗方面的巨大优势, 但实现单个的脉冲神经元和突触需要大量晶体管. 通常类脑神经元之间所需要的高连接度也限制了类脑芯片的规模和可扩展性. 因此, 利用基于阻变材料、相变材料等新材料器件实现类脑计算的基本单元成为研究的重点, 有望从根本上改进类脑芯片的计算架构和性能^[5–9]. 自旋电子学器件具有低能耗、高速和非线性磁动力学及非易失性等特点, 已经在硬盘、磁性随机存储器等存储设备上实现了商业化应用. 近年来, 基于自旋电子学的存内计算^[10]和类脑计算^[11–21]也获得了广泛的关注. 如何利用自旋电子学器件的天然优势, 结合新型类脑算法, 实现软硬件的协同设计, 是基于自旋器件类脑计算的核心问题.

大脑中存在大量的神经元来执行日常的感知、情感和记忆等高级功能. 不同于计算机中精确的数值表示, 大脑中信息传递依赖于神经元之间的脉冲发放, 依靠神经元之间的连接形成的复杂网络执行

* 国家自然科学基金 (批准号: 11734004, 12174028) 资助的课题.

† 通信作者. E-mail: zmzeng2012@sinano.ac.cn‡ 通信作者. E-mail: zyuan@bnu.edu.cn

相关的任务. 尽管每个神经元细胞的信息传递和膜电位变化过程中伴随大量噪音和随机动力学, 但大脑可以通过群体神经元共同编码信号来实现复杂和精确的认知功能. 群体编码是生物系统在方向和空间感知^[22]、颜色分辨、嗅觉识别和记忆^[23]等方面的基础模式^[24], 并具有很强的抗噪能力, 对神经细胞的异质性也具有很好的容忍度. 群体编码的优势可以解决制造工艺的局限导致新材料类脑器件的非均一性问题, 大幅降低新型类脑器件的制造要求. 基于传统 CMOS 技术^[25]、相变器件^[26]和自旋器件^[27]的群体编码系统都表现了对器件非均一性的较强耐受能力, 基于相变器件^[26]的群体编码利用了脉冲发放的随机特性, 其发放频率与输入信号的脉宽呈线性关系; 而随机磁性隧道结的发放频率与输入电压呈现单峰的非单调关系^[27], 更符合生物系统的调谐曲线.

本文利用超顺磁磁性隧道结的随机翻转特性, 结合群体编码算法, 提出一种基于自旋电子学器件的脉冲神经网络方案. 通过对鸢尾花数据集的无监督分类研究, 展示了该网络的鲁棒性和低能耗. 首先给出了在实验上观测到的磁性隧道结的随机翻转特性及其电流调控, 建立了基于超顺磁隧道结的群体编码基础, 同时验证了奈尔-布朗 (Néel-Brown) 理论模型对器件的描述能力. 研究了基于磁性隧道结的群体编码, 并结合延时脉冲发放和脉冲时序依赖突触可塑性设计了基于自旋电子学器件的神经网络. 探究了磁性隧道结数目和器件参数对网络性能的影响, 并分析了该网络在能耗方面的优势.

2 利用超顺磁隧道结的群体编码

2.1 群体编码

在线性回归的数值计算中, 我们的目标是由离散的实验数据点得到直线的斜率和截距. 由于实验数据存在误差, 我们期望通过更多的数据来降低数据误差带来的影响. 生物系统同样面临这样的问题, 单个神经元的信号噪声很大, 为了获取更加精准的感知信号, 生物系统总是通过一组神经元的共同响应来感知外界的刺激. 这种利用一群神经元对刺激的共同响应来编码信息的方式被称为群体编码, 其对于连续变量的编码尤为有效, 例如颜色、空间位置等. 在群体编码中每个神经元对不同的外界刺激输入信号具有一定的响应分布, 称为该神经

元的调谐曲线; 每个神经元对某个特定的外界输入刺激会比较敏感, 称为该神经元的偏好刺激. 因此通常可以用类似高斯函数的形式来描述调谐曲线, 峰值即对应偏好刺激. 当神经元足够多, 所有神经元的偏好刺激可以覆盖全部输入刺激的取值范围 (感受野) 时, 群体编码即相当于把一个输入信号用一组高斯型核函数展开, 因此即使若干个神经元上信息缺失也不影响整体工作效果. 相比于用单个神经元来处理精确的感知或运动控制, 群体编码能够显著提高准确性, 并且具有去除噪音、实现短时记忆和复杂非线性响应等功能^[24].

2.2 磁性隧道结的动力学

磁性隧道结是常见的自旋电子学器件, 由两个铁磁层夹着一个较薄的绝缘层组成, 如图 1(a) 所示, 其中上层的磁化方向固定, 中间是隧穿层, 下层的磁化方向可以通过施加外界磁场或激励电流控制翻转. 磁性隧道结有两个稳定的状态 (能量极小值), 当自由层和固定层磁矩平行时称为平行态, 此时电阻较低, 反之称为反平行态, 具有高电阻. 当两个状态之间势垒高度 ΔE 远大于热扰动 $k_B T$ 时, 磁性隧道结的状态是非常稳定的, 适合作为磁随机存储器中的非易失二值存储单元. 而当势垒高度与 $k_B T$ 相当时, 磁性隧道结的状态会受外界的热噪声影响而改变^[28]. 偏压或电流可以给自由层施加自旋转移力矩^[29], 从而调节隧道结在两个状态上的概率分布, 进而调控其作为神经元的调谐曲线, 实现基于磁性隧道结的群体编码系统.

本文使用的磁性隧道结多层薄膜通过磁控溅射沉积, 其核心结构自下而上为 PtMn(15 nm)/Co₇₀Fe₃₀(2.5 nm)/Ru(0.85 nm)/Co₄₀Fe₄₀B₂₀(2.4 nm)/MgO(0.80 nm)/Co₆₀Fe₂₀B₂₀(2.1 nm). 薄膜沉积后在 1 T 磁场下, 300 °C 退火 2 h. 随后通过光刻、蚀刻制备成 50 nm × 130 nm 的椭圆形器件. 沿与磁性隧道结器件难轴夹角 20° 方向施加适当强度的磁场, 然后向器件注入直流偏置电流并通过示波器测量器件两端电压. 电流产生的自旋转移力矩与易轴方向的磁场分量竞争使自由层磁矩在与钉扎层磁矩平行和反平行方向翻转, 通过示波器可以观察到电压信号随机跳变, 且随机跳变的频率随电流强度变化. 随机翻转测试在室温下进行. 图 1(b) 展示了实验上用电流调节磁性隧道结翻转频率的测量结果. 当电流为 -60 μ A 时, 磁性隧道结处于高阻态

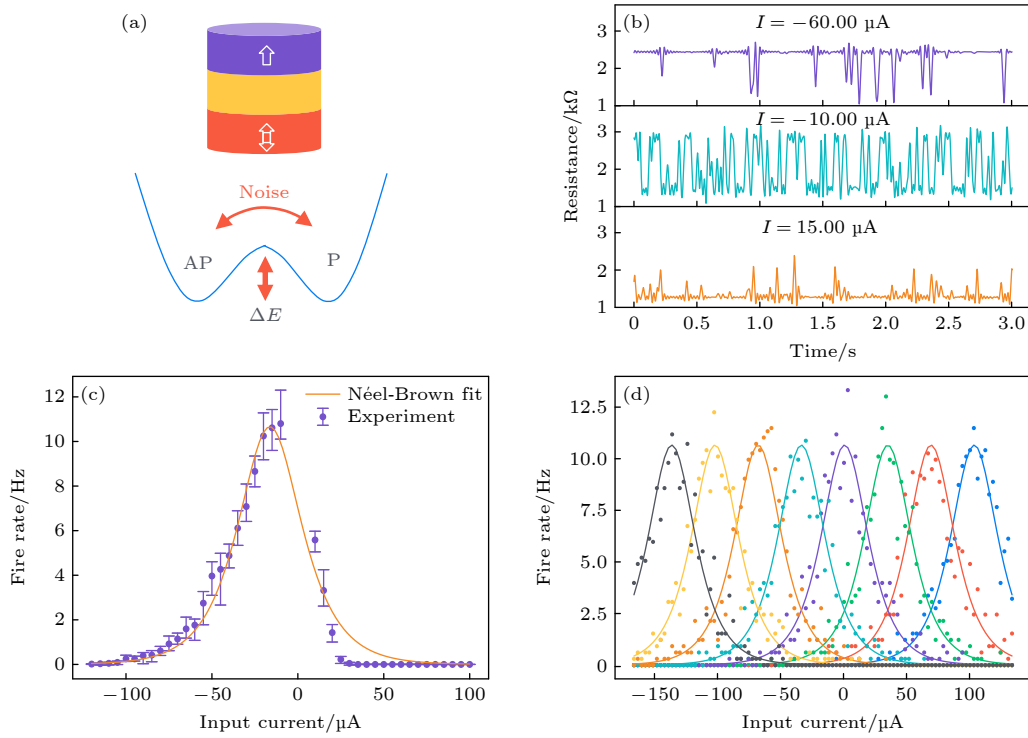


图 1 (a) 磁性隧道结示意图; (b) 不同电流下磁性隧道结电阻随时间的变化; (c) 磁性隧道结翻转频率与电流的函数关系; (d) 一组存在不同偏置电流的隧道结实现群体编码 (散点为 Néel-Brown 理论仿真结果, 实线对应隧道结的调谐曲线)

Fig. 1. (a) Schematic of a magnetic tunnel junction; (b) measured electrical resistance of a magnetic tunnel junction as a function of time under different electrical currents; (c) switching frequency of a magnetic tunnel junction as a function of electrical current; (d) population coding using a group of magnetic tunnel junctions with different bias currents (The dots are simulated data using the Néel-Brown theory and the solid lines are the corresponding tuning curves).

的概率较大; 而当电流为 $-10 \mu\text{A}$ 时, 磁性隧道结在高阻态和低阻态之间随机翻转; 当电流为 $15 \mu\text{A}$ 时, 磁性隧道结处于低阻态的概率较大. 注意到在不同电流下磁性隧道结的高阻态不同, 这是由于磁性隧道结的电流-电压曲线不是线性的, 在高阻态下呈抛物线型. 如果将磁性隧道结在高低阻态之间的一次翻转当作神经元的一次脉冲发放, 就可以通过调节输入电流来改变磁性隧道结的发放率. 随着输入电流的增大, 磁性隧道结的发放率先增大后减小, 如图 1(c) 所示. 图中的每个数据点为 10 次独立采样结果的平均值和标准差, 每次独立采样 97300 次, 大约 31 s.

2.3 Néel-Brown 理论

Néel-Brown 理论可以很好地描述超顺磁隧道结的随机动力学翻转 [30]. 处于平行态或反平行态上的磁矩翻转的弛豫时间 $\tau_{\text{P/AP}}$ 由势垒高度 ΔE 和经过隧道结的电流 I 共同决定, 即

$$\frac{1}{\tau_{\text{P/AP}}} = \phi_0 \exp \left[-\frac{\Delta E}{k_B T} \left(1 \pm \frac{I}{I_c} \right) \right], \quad (1)$$

式中 ϕ_0 为尝试频率, 与器件的翻转机制有关, 本文中取为 10^9 Hz . k_B 为玻尔兹曼常数, T 为环境温度, I_c 为磁性隧道结翻转的临界电流. 在群体编码中用输入电流的数值大小 I 来编码外界刺激信号. 磁性隧道结处于平行态或反平行态的概率表示为

$$P_{\text{P/AP}} = 1 - \exp \left(-\frac{\Delta t}{\tau_{\text{P/AP}}} \right), \quad (2)$$

其中 Δt 为采样时间, $\Delta t = 326.5 \mu\text{s}$. 如果将连续两次采样中磁矩发生改变定义为该神经元发放一次, 则 Néel-Brown 理论给出了超顺磁隧道结神经元的发放频率为

$$r(I) = \frac{\phi_0 \exp \left(-\frac{\Delta E}{k_B T} \right)}{2 \cosh \left(\frac{\Delta E}{k_B T} \cdot \frac{I}{I_c} \right)}. \quad (3)$$

可以用 (3) 式拟合实验数据, 结果见图 1(c) 实线, 可得 $\frac{\Delta E}{k_B T} = 17.7$, $I_c = 293.15 \mu\text{A}$. 拟合得到的能量势垒与临界电流与文献 [27] 中类似器件的参数可比. 可以看出当电流 $I = -16.27 \mu\text{A}$ 时发放率最大,

这是因为此时电流施加的自旋转移力矩刚好抵消了隧道结自身的偏置场。

为了使各神经元的偏好刺激均匀分布在输入刺激 I 的变化范围内, 可对各个磁性隧道结施加等间隔电流偏置 I_{bias} , 最终磁性隧道结的实际输入电流为 $I = I_{\text{input}} + I_{\text{bias}}$. 此时这一组神经元的调谐曲线均匀排列, 如图 1(d) 所示, 实线为该组磁性隧道结的调谐曲线, 散点代表利用 Néel-Brown 方程模拟的结果. 由于磁性隧道结的随机发放特性, 模拟结果与调谐曲线存在一定偏差。

3 群体编码脉冲神经网络和无监督聚类

3.1 脉冲神经网络模型

为展示基于超顺磁隧道结的群体编码方案, 设计了一个两层的脉冲神经网络, 如图 2(a) 所示. 网络前端是包含若干神经元的输入层, 每个神经元将一个输入信号编码为电流通入到下一层; 第一层是群体编码层, 每个输入层神经元的输入信号传递到一组超顺磁隧道结进行群体编码; 第二层输出层包含若干有适应性阈值的泄漏累积发放神经元 [31]. 群体编码层和输出层之间为全连接. 输出层的神经

元具有横向抑制连接, 保证该层神经元最多仅有一个发放。

图 2(a) 的神经网络用来实现鸢尾花数据集的无监督聚类. 鸢尾花数据集 [32] 中收集了 3 个类别的样本, 分别是 Setosa 鸢尾花, Versicolour 鸢尾花和 Virginica 鸢尾花, 每个类别有 50 个样本, 共计 150 条数据. 每条数据包含 4 个属性, 分别代表该朵鸢尾花花瓣的长度和宽度, 花萼的长度和宽度. 这 4 个数据通过 4 个输入神经元 x_i 分别把对应的电流输入给中间层的群体编码神经元, 即若干超顺磁隧道结中. 输入电流大小决定了磁性隧道结的翻转频率, 而磁性隧道结发放的脉冲会引起输出层输出神经元 y_i 膜电位的累积, 当输出神经元的膜电位达到该神经元阈值后将发放脉冲, 代表该鸢尾花样本属于发放的输出神经元所对应的类别. 由于输出层神经元之间存在抑制型突触 (紫色连线), 某个神经元发放后会抑制其他神经元的发放. 该抑制型突触可以设置为 $N \times N$ 的连接矩阵 (输出层神经元个数为 N), 其对角元为零, 非对角元为负值, 这里我们取值为 -17.5. 中间层与输出层之间的突触连接强度是可学习的, 应用依赖于当前权重的脉冲时序依赖可塑性 (spike-timing-dependent

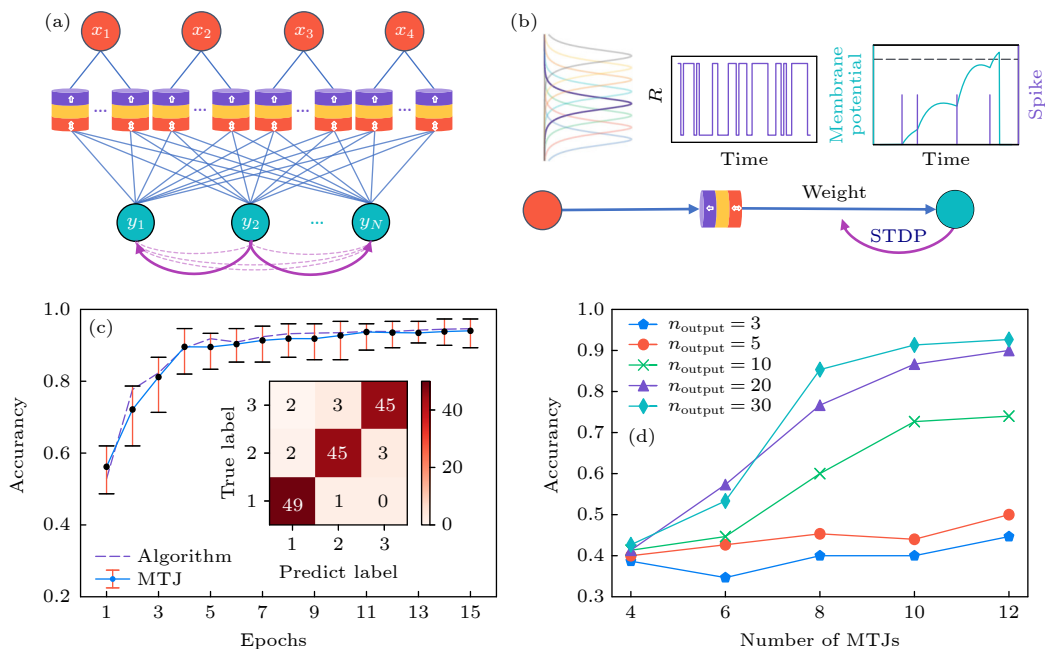


图 2 (a) 群体编码脉冲神经网络示意图; (b) 网络训练过程示意图; (c) 鸢尾花数据集无监督聚类测试结果; (d) 用于编码一个数据的隧道结数目与输出神经元数量对网络聚类正确率的影响

Fig. 2. (a) Schematic of spiking neural network; (b) schematic illustration of the network learning process; (c) test results of the unsupervised classification of the iris data set; (d) the influence of number of magnetic tunnel junctions used in the population coding and number of output neurons.

plasticity, STDP) 算法^[31]训练. 训练规则如下: 当突触前神经元发放时,

$$\Delta w = -\eta_{\text{pre}} x_{\text{post}}(w - w_{\text{min}}); \quad (4)$$

当突触后神经元发放时,

$$\Delta w = \eta_{\text{post}} x_{\text{pre}}(w_{\text{max}} - w). \quad (5)$$

这里, Δw 为权重变化量; 前突触学习率 η_{pre} 和后突触学习率 η_{post} 分别为 0.001 和 0.01; 权重 w 最大值 $w_{\text{max}} = 1$, 最小值 $w_{\text{min}} = 0$; x_{pre} 和 x_{post} 分别为突触前后神经元的膜电位. 图 2(b) 为网络运行的流程图, 从左到右分别为群体编码的调谐曲线, 磁性隧道结的电阻变化示意, 泄漏型积分发放神经元的动力学示意.

3.2 鸢尾花样本的无监督聚类

首先用理想调谐曲线的算法进行数值仿真, 每个输入变量用 12 个磁性隧道结编码, 输出神经元的数量为 30. 在每一轮训练中随机通入 100 条数据, 注意这里的数据是随机挑选, 并且由于是无监督学习, 所以我们并没有区分训练集和测试集. 输出神经元的标签是动态调整的, 每完成一轮训练后, 将对输出神经元刺激最强的鸢尾花类别作为该输出神经元的标签. 在每轮训练结束后, 将 150 个样本随机输入网络进行测试, 测试得到的聚类正确率如图 2(c) 中的虚线所示. 经过 15 轮训练后网络的聚类正确率达到了饱和值的 92.6%.

生物系统的神经元在不同时刻收到相同的刺激时, 发放模式并不是完全重复调谐曲线, 而是具有一定的随机性, 但并不影响大脑精确地执行各项复杂的动作. 同样, 外界热噪声扰动下的磁性隧道结的动力学具有一定的随机性, 因此超顺磁隧道结的群体编码可以很好地模拟生物系统中的神经元随机动力学. 为了检验网络性能, 我们分别进行了 10 次独立的试验, 并把测试正确率的平均值和标准差画在图 2(c) 中. 可以看出, 随着训练轮次的增加, 基于磁性隧道结的群体编码网络正确率最终也达到了使用理想调谐曲线的算法仿真极限值 (图 2(c) 虚线). 图 2(c) 中插图 15 轮训练后测试结果的融合矩阵, 其中的标签 1, 2, 3 依次对应 Setosa 鸢尾花、Versicolour 鸢尾花和 Virginica 鸢尾花, 红色的深度与正确预测的标签数目相对应.

随着群体编码使用隧道结 (神经元) 数量的增加, 偏好刺激在感受野中的分布更加密集, 可以使

群体编码包含的信息量更大, 鲁棒性更强^[33]. 同时, 随着磁性隧道结数目的增加, 输入数据将被投影到更高维的空间, 该组磁性隧道结的发放脉冲包含的信息随之增加, 输出层神经元能够解码出更多的信息, 包含了数据的更多细节, 因此磁性隧道结数目的增加有助于提高网络的性能. 图 2(d) 展示了随着用于编码每个变量的磁性隧道结的增加, 网络性能逐渐增强, 当磁性隧道结数目达到 12 个时, 无监督聚类的正确率也达到饱和.

本文还探究了输出层神经元对网络性能的影响. 执行无监督聚类的网络事先并不知道样本数据种类数量, 因此输出层神经元的数量应当大于或等于数据种类的数量. 输出层神经元的标签由引起它发放次数最多的鸢尾花类别决定, 因此, 每一类鸢尾花由一组输出层神经元共同标注. 当输出层的神经元达到 20 个时, 正确率可以达到 90%; 当输出层的神经元达到 30 个, 正确率就可以达到 92.6%, 达到了网络性能的极限. 需要强调, 我们的无监督学习中使用了 STDP 的学习规则, 在实际训练过程中表现出一定的不稳定性, 对权重初值较为敏感, 而更多的输出神经元代表网络有更大的初始参数空间, 这给训练过程带来了更多的可能性, 更加有利于 STDP 的学习, 从而避免参数落入局域最优解导致训练失败, 这也解释了为什么更多的输出神经元会提高网络的性能.

尽管我们仅测试了样本个数较少的鸢尾花数据集, 但该网络在软件层面已经成功应用于大规模数据集的无监督分类^[31,34]. 因此, 基于磁性隧道结的群体编码硬件网络也同样具有可扩展性, 能够推广到更大规模的数据分类任务. 此外, 本文探究了权重的精度对网络性能的影响, 模拟结果表明在鸢尾花聚类的任务中, 3 比特权重下网络的性能就可以达到 90% 以上, 这表明该网络的权重可以映射到磁畴壁等自旋多态器件, 有望实现全自旋脉冲神经网络.

4 器件的非均一性与能耗

即使同一批生长的材料和相同的刻蚀工艺制造的磁性隧道结器件通常也会有所差异, 这种器件的非均一性是类脑计算硬件不可避免的问题. 具体到超顺磁隧道结, 其非均一性表现在器件两个参数的差别, 即势垒高度 ΔE 和临界电流 I_c . 这两个参

数的差别导致隧道结调谐曲线的高度和宽度都会发生改变^[27]. 为了探究器件的不均一性对网络性能的影响, 我们测试了磁性隧道结的势垒 ΔE 存在0—10% 涨落范围下无监督聚类的准确性. 测试中的群体编码均采用每组8个隧道结, 输出层包含50个神经元. ΔE 增大(减小)会降低(升高)磁性隧道结的翻转频率, 因此 ΔE 的涨落影响器件调谐曲线的高度. 例如势垒过高的器件翻转1次所需的时间很长, 因此平均发放速率相对要低很多. 图3(a)插图展示了 ΔE 的涨落为6%时一组磁性隧道结的调谐曲线. 将这组器件放入网络中进行群体编码, 最终实现无监督分类的正确率达到80%. 整体上随着器件势垒高度涨落增大, 网络的性能略有降低^[27], 如图3(a)所示.

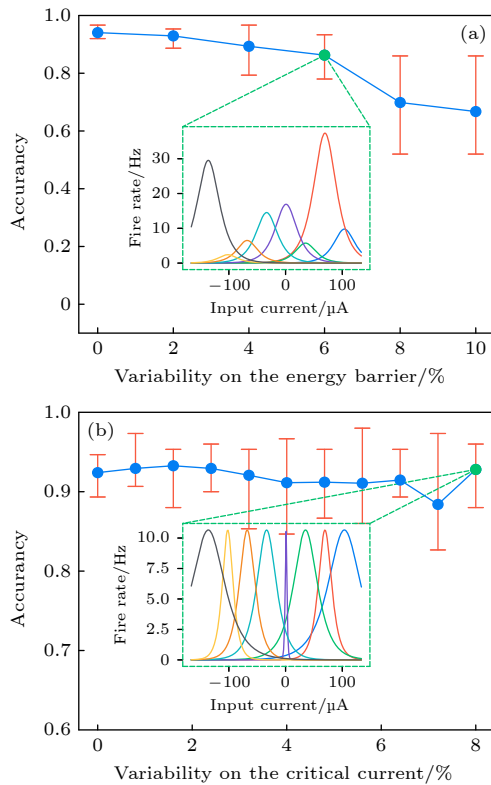


图3 磁性隧道结 (a) 势垒和 (b) 翻转临界电流的不均一性对网络性能的影响

Fig. 3. Population coding using the magnetic tunnel junctions with a fluctuation in their (a) energy barriers and (b) critical currents for switching.

我们进一步测试了磁性隧道结的临界电流 I_c 存在0—80% 涨落时网络的性能, 结果如图3(b)所示. I_c 的变化会影响调谐曲线的宽度, 图3(b)的插图展示了 I_c 存在80% 涨落时的一组器件的调谐曲线, 此时网络的正确率依然能够达到90%. 因此

不同器件临界电流并不会影响网络的鲁棒性, 势垒高度的变化对网络性能的影响更加显著.

上述测试结果与生物系统中群体编码的特征一致. 在生物系统中, 不同神经元在收到相同的刺激时响应也不是完全相同的, 这种异质性神经元的群体编码并不影响生物神经系统的正常工作.

磁性隧道结的发放每个采样时间窗口内是随机的, 我们可以改变观测时间, 在1个观测时间内对磁性隧道结进行多次采样, 然后将多次采样的结果转换为1次脉冲, 这样可以在统计意义上消除一定的随机性, 从而提高网络的性能. 但随着观测时间的增加, 每次运行时器件能耗也会随之增加. 为了探究器件耗能和网络性能的关系, 我们使用每组群体编码包含了16个磁性隧道结, 模拟观测时间与采样时间的比值对网络性能的影响. 在一个采样时间内磁性隧道结的最大输入电流为 $100 \mu A$, 电阻取高阻态与低阻态的均值 $2.016 k\Omega$, 因此最大能耗为 $6.58 \times 10^{-9} J$. 将其乘以观测时间与采样时间的比值, 就可得到观测时间对应的最大能耗. 图4给出了不同观测时间的聚类正确率, 其中上轴表示每个观测时间对应的能耗. 当观测时间为10次采样时间时网络达到正确率的饱和值, 对应的最大能耗为 $6.58 \times 10^{-8} J$. 随着实验工艺的进步, 降低采样时间可以有效地降低器件能耗, 目前实验上报道的隧道结采样时间可以低至 $8 ns$ ^[35],

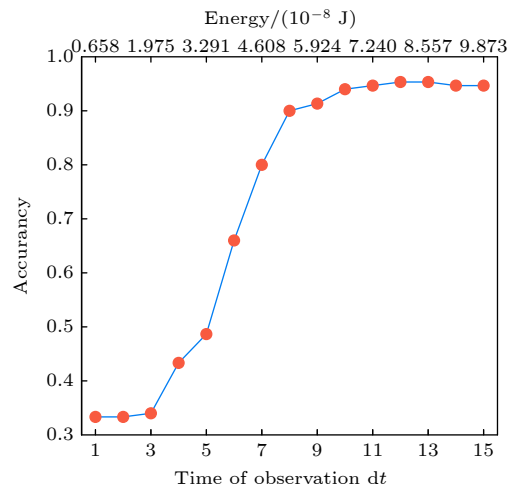


图4 不同观测时间 (以采样时间 dt 为单位) 对网络执行无监督聚类正确率的影响 (上轴表示每个观测时间对应的能耗)

Fig. 4. Classification accuracy of unsupervised clustering performed by the network as a function of the different observation time (The upper axis shows the energy consumption corresponding to each observation time).

相比于 CMOS 的群体编码系统^[25], 采样速度提高了 3 个量级, 采用该器件实现群体编码有望将能耗降到 4×10^{-12} J. 在系统集成方面, 磁性隧道结将连续信号直接转换为二值信号, 有利于后续数字信号的处理, 而基于 CMOS 的群体编码需要设计额外的数模转换电路. 因此, 基于自旋电子器件的群体编码在面积、能耗和速度方面都有突出的优势.

5 结 论

本文提出了一种使用超顺磁隧道结的群体编码方案, 利用热扰动下超顺磁隧道结的随机翻转过程及其自旋转移力矩的调控, 可以用带有偏置力矩的若干隧道结的偏好刺激编码整个神经元的感受野. 作为一个简单且典型的例子, 我们把该群体编码方案应用到一个三层的神经网络中, 实现了对鸚尾花数据集的无监督聚类, 探究了用于群体编码的隧道结个数等网络参数对聚类正确率的影响. 群体编码方案在应用中可以很好地克服器件非均一性等技术和工艺问题, 并且能够有效地降低神经网络的能耗.

参考文献

- [1] LeCun Y, Bengio Y, Hinton G 2015 *Nature* **521** 436
- [2] Roy K, Jaiswal A, Panda P 2019 *Nature* **575** 607
- [3] Davies M, Srinivasa N, Lin T H, et al. 2018 *IEEE Micro.* **38** 82
- [4] Pei J, Deng L, Song S, et al. 2019 *Nature* **572** 106
- [5] Ambrogio S, Narayanan P, Tsai H, Shelby R M, Boybat I, di Nolfo C, Sidler S, Giordano M, Bodini M, Farinha N C P, Killeen B, Cheng C, Jaoudi Y, Burr G W 2018 *Nature* **558** 60
- [6] Torrejon J, Riou M, Araujo F A, et al. 2017 *Nature* **547** 428
- [7] Yao P, Wu H Q, Gao B, Tang J S, Zhang Q T, Zhang W Q, Yang J J, Qian H 2020 *Nature* **577** 641
- [8] Zhang X M, Zhuo Y, Luo Q, et al. 2020 *Nat. Commun.* **11** 51
- [9] Zhang Y, Wang Z R, Zhu J D, Yang Y C, Rao M Y, Song W H, Zhuo Y, Zhang X M, Cui M L, Shen L L, Huang R, Yang J J 2020 *Appl. Phys. Rev.* **7** 011308
- [10] Jung S, Lee H, Myung S, et al. 2022 *Nature* **601** 211
- [11] Grollier J, Querlioz D, Camsari K Y, Everschor-Sitte K, Fukami S, Stiles M D 2020 *Nat. Electron.* **3** 360
- [12] Lan X K, Cao Y, Liu X Y, Xu K J, Liu C, Zheng H Z, Wang K Y 2021 *Adv. Intell. Syst.* **3** 2000182
- [13] Jiang W C, Chen L N, Zhou K Y, Li L Y, Fu Q W, Du Y W, Liu R H 2019 *Appl. Phys. Lett.* **115** 192403
- [14] Zhang Y J, Zheng Q, Zhu X R, Yuan Z, Xia K 2020 *Sci. China Phys. Mech. Astron.* **63** 277531
- [15] Zheng Q, Mi Y Y, Zhu X R, Yuan Z, Xia K 2020 *Phys. Rev. Appl.* **14** 044060
- [16] Zheng Q, Zhu X R, Mi Y Y, Yuan Z, Xia K 2020 *AIP Adv.* **10** 025116
- [17] Sengupta A, Roy K 2016 *Phys. Rev. Appl.* **5** 024012
- [18] Yu W C, Xiao J, Bauer G E W 2021 *Phys. Rev. B* **104** L180405
- [19] Song K M, Jeong J S, Pan B, et al. 2020 *Nat. Electron.* **3** 148
- [20] Kurenkov A, DuttaGupta S, Zhang C, Fukami S, Horio Y, Ohno H 2019 *Adv. Mater.* **31** e1900636
- [21] Romera M, Talatchian P, Tsunegi S, et al. 2018 *Nature* **563** 230
- [22] Banino A, Barry C, Uria B, et al. 2018 *Nature* **557** 429
- [23] Mi Y, Katkov M, Tsodyks M 2017 *Neuron* **93** 323
- [24] Pouget A, Dayan P, Zemel R 2000 *Nat. Rev. Neurosci.* **1** 125
- [25] Thakur C S, Wang R, Hamilton T J, Tapson J, van Schaik A 2016 *IEEE Transactions on Circuits and Systems I: Regular Papers* **63** 211
- [26] Tuma T, Pantazi A, Le Gallo M, Sebastian A, Eleftheriou E 2016 *Nat. Nanotechnol.* **11** 693
- [27] Mizrahi A, Hirtzlin T, Fukushima A, Kubota H, Yuasa S, Grollier J, Querlioz D 2018 *Nat. Commun.* **9** 1533
- [28] Cai J L, Fang B, Zhang L K, Lv W X, Zhang B S, Zhou T J, Finocchio G, Zeng Z M 2019 *Phys. Rev. Appl.* **11** 034015
- [29] Cai K M, Yang M Y, Ju H L, et al. 2017 *Nat. Mater.* **16** 712
- [30] Li Z, Zhang S 2004 *Phys. Rev. B* **69** 134416
- [31] Diehl P U, Cook M 2015 *Front. Comput. Neurosci.* **9** 99
- [32] Fisher R A 1936 *Annals of Eugenics* **7** 179
- [33] Dayan P, Abbott L F 2001 *Theoretical Neuroscience* (Cambridge, MA: MIT Press) pp108–112
- [34] Biswas A, Prasad S, Lashkare S, Ganguly U 2016 arXiv: 1612.02233
- [35] Hayakawa K, Kanai S, Funatsu T, Igarashi J, Jinnai B, Borders W A, Ohno H, Fukami S 2021 *Phys. Rev. Lett.* **126** 117202

SPECIAL TOPIC—Physical electronics for brain-inspired computing

Implementation of unsupervised clustering based on population coding of magnetic tunnel junctions^{*}

Zhang Ya-Jun¹⁾ Cai Jia-Lin²⁾ Qiao Ya¹⁾ Zeng Zhong-Ming^{2)†}
Yuan Zhe^{1)‡} Xia Ke³⁾

1) (*Center for Advanced Quantum Studies, Department of Physics, Beijing Normal University, Beijing 100875, China*)

2) (*Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou 215123, China*)

3) (*Beijing Computational Science Research Center, Beijing 100193, China*)

(Received 9 February 2022; revised manuscript received 13 April 2022)

Abstract

Developing suitable algorithms that utilize the natural advantages of the corresponding devices is a key issue in the hardware research of brain-inspired computing. Population coding is one of the computational schemes in biological neural systems and it contains the mechanisms for noise reduction, short-term memory and implementation of complex nonlinear functions. Here we show the controllable stochastic dynamical behaviors for the technically mature spintronic device, magnetic tunnel junctions, which can be used as the basis of population coding. As an example, we construct a two-layer spiking neural network, in which groups of magnetic tunnel junctions are used to code input data. After unsupervised learning, this spiking neural network successfully classifies the iris data set. Numerical simulation demonstrates that the population coding is robust enough against the nonuniform dispersion in devices, which is inevitable in fabrication and integration of hardware devices.

Keywords: magnetic tunnel junction, population coding, spiking neural network, unsupervised learning

PACS: 85.70.-w, 07.05.Mh

DOI: 10.7498/aps.71.20220252

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 11734004, 12174028).

[†] Corresponding author. E-mail: zmzeng2012@sinano.ac.cn

[‡] Corresponding author. E-mail: zyuan@bnu.edu.cn