

# 基于 *spiking* 神经网络的自主强化学习模型及其在决策中的应用

A *spiking* Neural Network Based Autonomous Reinforcement Learning Model and Its Application in Decision Making

Guixiang Wang<sup>1</sup>, Yi Zeng<sup>1,2</sup>, and Bo Xu<sup>1,2</sup>

<sup>1</sup> Institute of Automation, Chinese Academy of Science, Beijing, 100190, China

<sup>2</sup> Center for Excellence in Brain Science and Intelligence Technology Chinese Academy of Sciences, Shanghai, 200031, China

yi.zeng@ia.ac.cn, guixiang.wang@ia.ac.cn

(Song Jian, translate)

**摘要:** 本文提出了一种用于决策的自主脉冲神经网络 (SNN, *spiking* Neuron Networks) 模型。该模型是对具有自动环境感知功能的基底神经节电路的扩展, 从图像输入中自动构造环境状态。本文的工作有以下贡献: (1) 在我们的模型中, 建立了简化的 *Hodgkin-Huxley* 计算模型, 以达到接近漏整合发放 (Leaky Integrate-and-Fire, *LIF*) 模型的计算效率, 并用于获取和测试认知中的离子水平特性。(2) 提出了一种基于尖峰的运动感知机制, 在不进行大量训练的情况下, 从原始像素中提取学习过程的关键元素。我们将我们的模型应用到“飞鸟”游戏中, 经过几十次训练后, 它表现得很好。该模型在训练开始时与人的学习性能相似。此外, 我们的模型模拟了 *Hodgkin-Huxley* 模型中阻断钠离子或钾离子通道时的认知缺陷, 这是对深入到离子水平的认知的探索。

**关键词:** *spiking* 神经网络; *Hodgkin-Huxley*; 基底神经节; 运动知觉

## 1 引言

人脑可以很好地处理复杂的任务, 从观察情况到给出可用的计划。当在某些设备上玩游戏时, 大脑不仅会玩游戏, 还会分析它所感知到的东西, 获取关键元素, 并根据反馈给出下一个动作。大脑理解情况, 而传统的人工智能不能“理解”情况。受大脑结构和信息处理方式启发的算法可能会为理解和解决问题提供更好的解决方案。

以大脑为灵感的神经网络为实现一般智能提供了新的机会, 并在最近几年变得流行起来。在本文中, 我们通过分析每一次出现的配置, 建立了一个可以玩战略游戏的 *spiking* 神经网络。智能体 (Agent) 一开始对比赛一无所知。经过几次检测, 它能感觉到运动和可用的运动。了解移动限制和环境后, 智能体将学习如何移动以获得更好的游戏分数。策略学习是一个强化培训过程。我们的 *spiking* 神经网络模型由两个模块组成, 一个是环境感知, 另一个是自主学习。

我们的动机是建立一个大脑激发的认知模型来强化学习任务, 这样有了这个模型的人就可以在不知道环境是什么的情况下拥有自主学习能力。智能体可以用眼睛感知环境, 对任务的关键要素进行抽象, 但先验知识很少。为了完成强化学习等战略任务, 我们的自主认知模型从结构和功能的角度借用了大脑中的想法。我们希望我们的工作能给我们提供一种接近人类智能的可能方法。

本文的其余部分组织如下。第 2 节描述了基底神经节模型和 *Hodgkin-Huxley* 模型。第 3 节是我们的方法, 包括我们的自主学习模型的简化 *Hodgkin-Huxley* 计算方程; 我们提出的运动感知算法和整个自主学习模型。第 4 节是我们模型的实验和应用。最后, 第 5 部分对本文进行了总结, 并讨论了模型的局限性。

## 2 前期工作

## 2.1 基底神经节模型

生物大脑给我们很多启发。基底神经节是大脑重要的结构之一。它们在许多认知行为中都很重要，包括决策和强化学习。基底神经节由纹状体、STN（丘脑下核）、GPe（苍白球外段）和两个输出核、SNr（黑质网状部）和GPi（苍白球内段）组成[3,5-8]。其他核，包括SNc（黑质致密部）和VTA（腹侧被盖区），也被视为基底神经节的一部分。每个核在大脑的学习过程中有不同的功能。

有一些重要的核需要引入。纹状体是基底神经节的输入核，它接受来自皮质区域的直接输入，并调节来自SNc的传入（DA）。它有两种类型的DA受体，分别是D1和D2。StrD1受体增强输入的反应，而StrD2则有相反的作用[5,25]。STN接收来自皮质区域的直接输入。与基底神经节其他核的抑制性输出相比，STN的输出具有兴奋性。GPi是基底神经节的输出核。SNc和VTA释放多巴胺（DA）作为一种对认知行为至关重要的调节信号[4,5,7]。

在认知过程中，基底神经节需要与其他相关脑区协同形成基底神经节回路[4-6,8,25]。基底神经节接收来自前额皮质（PFC）的输入，并将输出发送到丘脑。丘脑将基底神经节的输出送回PFC[25]。对基底神经节的认知研究产生了大量的模型。从基底神经节的“盒箭”模型出发[14,15]根据大脑解剖数据，提出了许多计算模型[5,7-10,13,15,16]。这些模型试图根据基底神经节及其相关脑区的结构和功能来改善原始基底神经节电路。大多数的基底神经节模型是用人工神经网络[5,9,15]建立的，不能很好地描述神经活动。离子水平神经元模型，如Hodgkin-Huxley模型，在这些基底神经节模型中没有使用。此外，这些模型被用来做简单的动作选择实验，而不是解决实际和复杂的认知任务。我们将建立基于基底神经节电路的自主学习模型。在描述我们的模型之前，我们需要介绍给我们灵感的基底神经节的数学模型。

## 2.2 基底神经节的尖峰编码模型

我们的工作与动作选择的数学基底神经节模型[10]及其扩展的spiking编码模型[1,8]有关。

由于两个原因，我们开发了基于spiking编码模型的自主学习模型。（1）该模型计算复杂度低，保持了基底神经节的主要生物学特性。（2）该模型可以处理数百个维度的大范围输入[1]。

尖峰编码模型的选择机制可以描述为一个线性方程[10]：

$$f(x_i) = \begin{cases} 0 & x_i < e_i \\ m(x_i - e_i)e_i & e_i \leq x_i < 1/m + e_i \\ 1 & x_i > 1/m + e_i \end{cases} \quad (1)$$

其中 $f(x_i)$ 是核的输出，而 $x_i$ 是以 $0 \leq x_i \leq 1$ 作为值区间的输入。

图1显示了核之间的联系以及基底神经节中每个核的精确方程式。前额叶皮质向基底神经节投射输入向量（代表一组动作）。后者根据其值选择一个动作，并通过GPi将其输出到丘脑（未绘制）。注意基底神经节的输出是抑制性的，因此，所选择的动作将具有接近零的最小值。

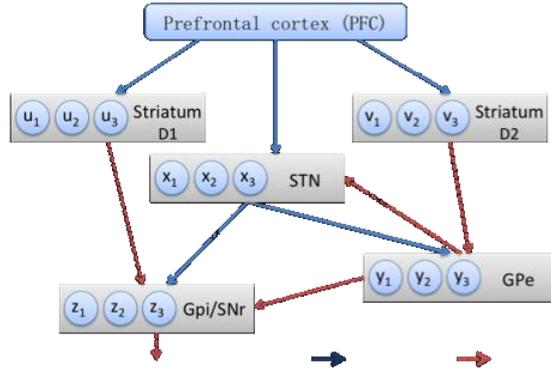


图 1: 基底神经节[1]的数学模型。基底神经节的输入向量来自前额皮质。x、y、z、u 和 v 分别表示每个核的向量输出。方程中的值，如 0.25，来自基底神经节[1]的数学模型。

基底神经节中的每个核由一组生物神经元表示，如图 1 所示，以圆圈表示。一般认为，生物神经元通过产生复杂的尖峰序列来传递信息[1,11,18]。这些峰值将输入刺激编码为频率[17,18]。频率随输入刺激的增加而增加。神经元  $i$  的动作电位可以写成：

$$v_i(t) = G(J_i(x(t))) \quad (2)$$

其中  $v_i(t)$  是神经元的动作电位， $G(x)$  是神经元模型。 $J(x)$  是电流输入，通常是  $x(t)$  的线性函数。当动作电位达到阈值时，释放峰值。使用神经元模型可以获得一个时间段内的峰发放频率  $r_i(x)$ ，因此，输入  $x$  将被编码为一个发放率（firing rate）。

在编码过程之后，我们需要解码操作来恢复估计的输入。输入通常是写为  $f(x(t))$  的函数，其估计结果可以表示为：

$$\hat{f}(x) = \sum_i^n r_i(x) d_i \quad (3)$$

其中  $d_i$  表示线性解码器， $n$  表示用于编码输入刺激的神经元数量。为了估计输入函数，我们可以使用最小二乘法（least-squared-error method）[1,8,12]。估计值与输入值的偏差可写为：

$$E = \int [f(x) - \hat{f}(x)]^2 dx \quad (4)$$

通过最小化偏差  $E$ ，得到译码器  $d_i$ [1,8] 的最小二乘版本估计：

$$d = \Gamma^{-1} \mathfrak{R} \quad \Gamma_{ij} = \int r_i r_j dx \quad \mathfrak{R}_i = \int r_i f(x) dx \quad (5)$$

在我们的自主学习神经网络中，利用 *Hodgkin-Huxley* 模型建立了基底神经节模型。我们将运行 *H-H* 方程来确定发放率和电流输入之间的关系。

### 2.3 Hodgkin-Huxley 模型

*Hodgkin-Huxley* 模型是由 Alan Lloyd Hodgkin 和 Andrew Fielding Huxley 在 1952 年发现的[20]。他们是从鱿鱼巨大的轴突中得到的。利用实验数据确定了模型中的许多参数。在离子水平上，*H-H* 模型是描述动作电位动态特性的最佳模型。该模型被写成四个方程[20]：

$$\begin{aligned} C\dot{V} &= I - \bar{g}_k n^4 (V - E_k) - \bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_L (V - E_L) \\ \dot{n} &= \alpha_n(V)(1-n) - \beta_n(V)n \\ \dot{m} &= \alpha_m(V)(1-m) - \beta_m(V)m \\ \dot{h} &= \alpha_h(V)(1-h) - \beta_h(V)h \end{aligned} \quad (6)$$

其中  $V$  是膜电位， $C$  是膜电容，且  $C=1 \mu F/cm^2$ ， $i$  是外部电流输入。

离子电流由三种成分组成：三个激活门和一个灭活门的钠 ( $Na^+$ ) 电流、四个激活门的

钾 ( $K^+$ ) 电流和主要由氯 ( $Cl^-$ ) 携带的泄漏电流[19,21]。方程式中的  $n$ 、 $m$  和  $h$  分别代表不同离子门的开启概率[21]，其中  $n$  代表钾 ( $K^+$ ) 的， $m$  和  $h$  代表钠 ( $Na^+$ ) 的。

门在开启和关闭状态之间的转换速率用  $\alpha(V)$  和  $\beta(V)$  表示， $\alpha(V)$  和  $\beta(V)$  与电压有关[19,21] 且命名为速率常数。电压钳实验确定了过渡速率函数[20,21]。了解具体方程式请参考[20,21]。

本文采用上述的尖峰编码方法，利用  $H-H$  模型建立基底神经节模型。它可以模拟更多与认知相关的生物效应。我们将该模型应用于实际的认知任务，以测试模型的离子性质，并模拟其对认知的影响。

## 3 方法

### 3.1 简化计算的 *Hodgkin-Huxley* 模型

我们利用 *Hodgkin-Huxley* 方程，在第二节关于基底神经节的先前工作的基础上，建立了强化学习的 *spiking* 神经网络模型。我们在这里使用  $H-H$  模型有两个原因。(1)  $H-H$  模型能准确解释实验结果，并能对神经细胞进行定量电压分析。(2) 钠和钾与人类认知有关。钠或钾通道中离子电导的变化可能会影响决策或学习过程，从而给我们一个在离子水平深入研究“理解力”的机会。

基底神经节模型通常是用漏整合发放神经元 (*LIF*) 建立的。其原因是 *LIF* 模型计算复杂度低，易于控制。然而 *LIF* 神经元对动态动作电位和离子水平活性的描述较差。我们不能仅仅在 *LIF* 神经元水平上很好地模拟认知失败 (*cognitive defeats*)。 $H-H$  模型可以很好地深入到离子水平，这为我们分析和预测生物特性提供了更多的信息。

与 *LIF* 神经元不同， $H-H$  模型更为复杂。为了达到动作电位阈值， $H-H$  模型需要模拟数百个步长，而不是像 *LIF* 一样只有一个步长。如果 *LIF* 只用一步长在 1 毫秒内达到峰值，由于离子通道的动态特性， $H-H$  模型将花费数百个步骤在同一时间段内达到阈值。这种现象由  $H-H$  方程决定。

当电流输入大于某一特定值时， $H-H$  模型会自动释放某一频率的峰值。如果 *LIF* 模型以毫秒为单位运行，则  $H-H$  模型以纳秒为单位运行。这意味着在使用  $H-H$  模型进行模拟时，需要两个回路，外部回路是模拟本身，内部回路是离子水平模拟，以积累动作电位。这需要更多的模拟和成本更高的计算复杂性，因此，有必要为  $H-H$  模型提出简化的计算方程。

为了达到接近 *LIF* 模型的计算效率，本文对  $H-H$  方程做了一些小的改动。(1) 膜电容  $C$  小于初始值，设置为  $0.01-0.03\mu F/cm^2$ 。这将在一步长内大大增加动作电位。在达到临界点之前，需要更少的步长。(2) 为了确保足够大的过滤速率，电流输入应大于  $6-7\mu A/cm^2$ ，而较大的电流输入可确保更好的精度。但是，电流输入不能太大，否则会降低  $H-H$  模型的离子性质。根据我们的实验，30 是电流输入的一个更好的选择。(3)  $H-H$  方程模拟中没有动作电位积累的内环。前两个变化在一步长内获得足够的电压增量。一步长模拟的精度和以前一样。(4)  $H-H$  方程以一定频率自动恢复峰值，因此，为了控制  $H-H$  神经元的活动来编码不同的功能，进行了干预机制。我们需要在尖峰后将电压重置为零，并根据所表示的函数和编码过程为方程分配一个不稳定的时间段。

我们将  $H-H$  模型的电压方程进行了一些改进，改写如下：

$$\begin{aligned}
C\dot{V} &= J(x) - \overline{g}_k n^4 (V - E_k) - \overline{g}_{Na} m^3 h (V - E_{Na}) - \overline{g}_L (V - E_L) \\
V &= 0, \quad \text{if } V > 60mV \\
J(x) &= 30(ax + b), \quad -1 < a, b < 1 \\
C &= 0.02\mu F/cm^2
\end{aligned} \tag{7}$$

其中  $J(x)$  是电流输入，是原始输入  $x$  的线性函数， $a$  和  $b$  是尖峰编码模型中不同神经元的随机值。30 是电流输入的一个标度，以保证  $H-H$  模型的充分发放率。峰值电压约为  $80-100mV$ ，因此，如果电位超过  $60mV$ ，我们会计算一个峰值，并将电压设置为零，让它立即重新启动，或在难以控制结束前等待一段时间。

为了便于计算，通常将静息动作电位设为  $V = V_{rest} = 0$ 。同时，其他电位和电导参数应给出以下值[19,22,24]:

$$\begin{aligned}
E_K &= -12mV & E_K &= 120mV & E_K &= 10.6mV \\
\overline{g}_K &= 36mS/cm^2 & \overline{g}_K &= 120mS/cm^2 & \overline{g}_K &= 0.3mS/cm^2
\end{aligned}$$

当输入电流大于特定值时，该特定值在我们的模拟中约为  $6-7\mu A/cm^2$ ，可以观察到有规律的峰值活动[22]。如果峰值间隔为  $T$ ，则平均拟合率可表示为  $f=1/T$ ，并随着刺激的增强而增加[22,23,24]，如图 2 所示。

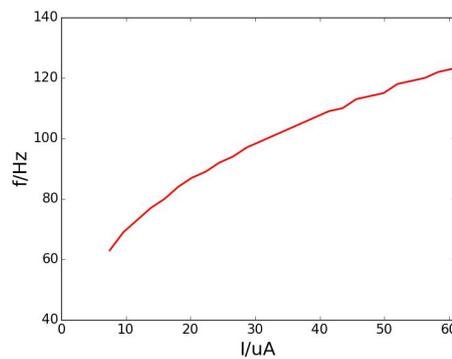


图 2: 在不同电流输入下，*Hodgkin-Huxley* 模型的发放率曲线。在我们的实验中，当电流输入大于  $7.5\mu A/cm^2$  时，神经元开始有稳定的规则尖峰。静息电位设为 0。时间常数是  $0.01ms$ 。仿真周期为  $1s$ 。

### 3.2 基于尖峰的运动感知

眼睛可以检测环境中的运动物体，并将结果传回大脑，作为认知的重要参考。本文建立了一个以智能体为对象的神经网络模型。使用该模型的智能体可以在一定程度上感知和理解环境。以基底神经节为强化学习模型，智能体将通过图像输入获得自主学习能力，而不是简单的动作选择来完成一些复杂的认知任务。

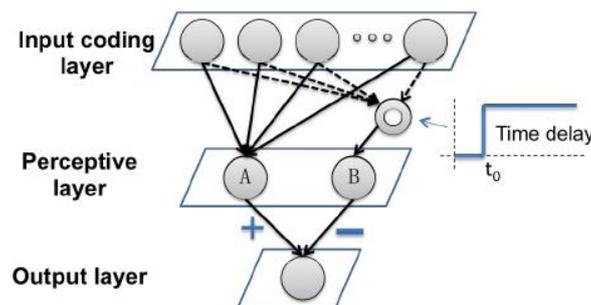


图 3: 多层结构的运动感知模型。这里的时间延迟用于保持上次的 *spiking* 序列，因此，输出层将检测随时间的运动。

图像的感知是多层的，包括输入层、感知层和输出层。对于图像输入，每个像素都有一组神经元来表示其在输入层中的值。利用群编码算法（Population coding algorithm）将像素序列转换为一段时间内的峰值序列。这些尖峰序列是感知层中两个特殊神经元的输入。在这一层中，两个神经元代表一个像素。神经元 A 在当前时间工作，神经元 B 在过去以恒定的时间延迟 $\tau_0$ 工作。这意味着神经元 A 接收电流输入，B 接收最后一个输入。如果两个神经元的位置有移动，它们会有不同的尖峰输出。在输出层中，每个像素有一个神经元。神经元的输入是神经元 A 和 B。神经元 A 被激发，而 B 则有相反的作用。如果没有移动，神经元就没有尖峰，否则会释放尖峰。

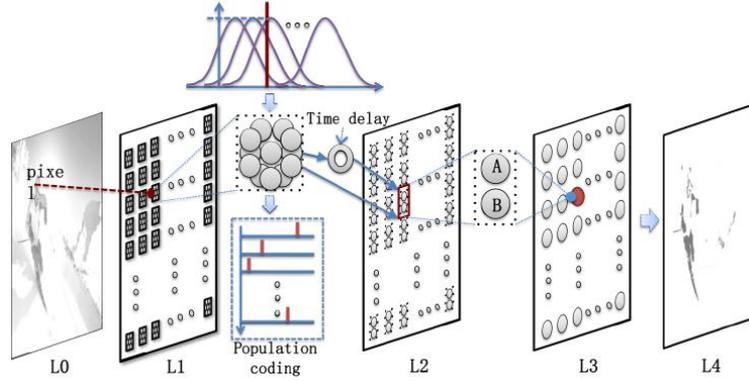


图4: 原始图像输入的运动感知过程。L0为原始图像，L1为输入编码层，L2为感知层，L3为脉冲输出层，L4为原始图像的脉冲活动，表示运动区域。总体编码过程在L1语言中进行。在L1中，每个像素被转换为一组神经元的尖峰序列。

整个运动感知过程如图4所示。图像中的每个像素都有其从输入编码层到输出层的代表神经元。在输入编码层中，每个像素由一组神经元表示，并使用群编码将其转换为峰值序列[26]。这些神经元具有重叠的高斯接受域（gaussian receptive fields），如图4所示。神经元数目和编码间隔可以可变，以获得良好的编码结果。有了良好的参数，群编码可以通过不同的尖峰模式区分两个相似的值。下一节的实验表明它在像素编码方面具有良好的性能。

在群编码中，假设 $[I_{min}, I_{max}]$ 是编码间隔，M是神经元数，每个神经元的高斯接受域中心可以计算为[26]:

$$C_i = I_{min} + (i - 1) \cdot \frac{I_{max} - I_{min}}{m - 2} \quad (8)$$

每个神经元的方差方程写为[26]:

$$\sigma_i = \beta \cdot \frac{I_{max} - I_{min}}{m - 2} \quad (9)$$

其中 $\beta$ 值一般在 $[0,2]$ 中。编码过程将计算一个值与每个高斯接受域的交点，并获得这些高斯神经元的峰值时间。

运动感知的输出为运动区域由尖峰高亮显示的图像，如图4中的L4所示。在背景变化较小的环境中，很容易得到物体的运动方向，并将运动区域与其他背景分开。这对于移动障碍物避免也很有用，将在下一节中介绍。

### 3.3 基于自主强化学习模型的spiking神经网络

在本节中，我们将描述我们的整个自主强化学习模型。该模型由运动知觉模型（motion perception model）和基底神经节模型两个重要的spiking模型组成。我们的自主模型设计用于与环境交互并执行强化学习，以避免在不给出环境状态的情况下移动障碍。本文中所有重

要的模型都是用 *Hodgkin-Huxley* 模型建立的。

我们的自主学习模型如图 5 所示。该模型形成一个 Q 学习循环，学习如何与环境交互。运动感知模型用于从图像输入中检测运动物体，并获得新的环境状态函数（environment state function）。基底神经节模型以状态函数为输入，在每个时间步长的所有可用动作中进行动作选择。

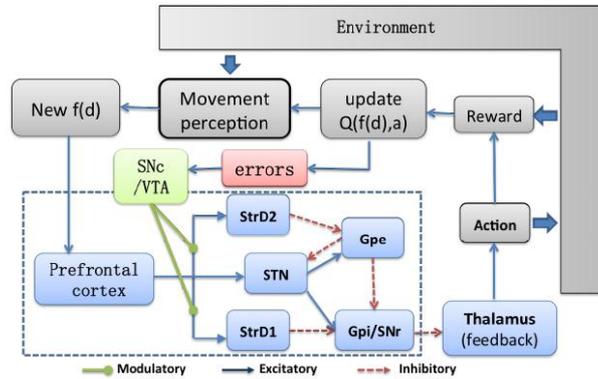


图 5: 利用 *spiking* 神经网络建立了自主强化学习模型。基底神经节结构及其作用选择函数(矩形内)参照[1]和[2]。 $f(d)$ 是距离  $d$  的函数，称为状态函数。Q 值是状态函数和动作  $a$  的函数，环境状态由状态函数计算而不是人为创建。除了一些基本的生存法则，这个智能体什么都不知道。

如第 2 节所述，使用参考文献[1,2]中的方法建立基底神经节模型。基底神经节电路的学习是通过更新基底神经节和前额叶皮质之间的突触连接来实现的。直到基底神经节获得良好的动作选择，并且智能体与环境很好地相互作用，学习才会结束。

智能体要想在环境中生存，必须遵循两条规则：（1）智能体的前进方向不应有障碍。（2）如果没有其他目标，智能体应尽可能远离障碍物。状态函数和奖励根据这两条规则计算。

## 4 实验与应用

在这一部分中，我们将我们的模型应用到名为“飞鸟”的游戏中，以测试我们的方法的性能。如上所述，通过我们的模型使用简化计算的 *Hodgkin-Huxley* 方程，还需要实验。下面，我们将首先对 *H-H* 模型进行实验。接下来，我们使用我们的模型在“飞鸟”游戏中进行了检验。

### 4.1 简化的 *Hodgkin-Huxley* 模型

如上所述，*Hodgkin-Huxley* 模型具有许多 *LIF* 模型没有的离子性质，我们提出了简化的 *H-H* 方程，以降低计算复杂性，同时保持其离子性质。图 6 中的实验显示了我们简化的 *H-H* 模型的性能。实验结果表明，简化后的 *H-H* 模型具有较低的计算复杂度和与 *LIF* 模型相似的性能。

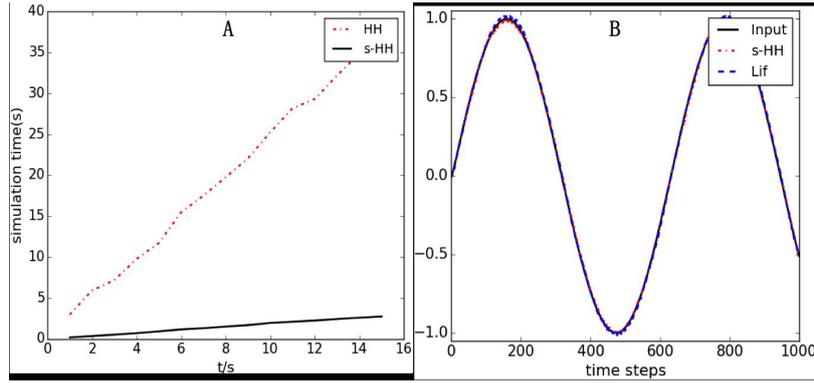


图 6: 简化  $HH$  方程的实验研究。A:原  $HH$  模型与简化模型的仿真时间比较。B:简化  $HH$  模型和  $LIF$  模型的函数跟踪实验。结果表明, 简化后的  $HH$  模型具有与  $LIF$  模型相似的恢复性能。

## 4.2 游戏中的自主强化学习模型

在玩“飞鸟”游戏之前, 我们将解释如何计算环境状态函数和奖励。

在“飞鸟”游戏中, 管道朝着鸟移动, 鸟上下移动以调整其位置, 试图通过管道间隙而不发生碰撞。根据上述两条幸存的规则, 鸟类应避免沿其移动方向与管道发生碰撞, 并尽量远离每个管道的终端。因为上面和下面都有管子, 鸟应该尽量靠近管子缝隙的中心。

环境状态是描述环境的关键元素。如果没有事先指定, 它可以是管道间隙中心距离的函数, 写为  $f(d)$ 。当一对管道出现时, 鸟用运动感知模型来获取管道间隙的中心。环境状态函数  $f(d)$  根据鸟的位置计算每一时间步长。对于状态函数的每个值, 鸟可以采取两个动作, 即上和下。 $Q$  学习过程的  $Q$ -值函数由  $f(d)$  和动作  $a$  的函数表示, 写为  $Q(f(d), a)$ 。每一时间步长的  $Q$  值更新时, 该误差用于自主模型的学习。

奖励对学习过程也很重要。它可以根据运动感知结果来计算。如果鸟的前进方向没有管道, 奖励是正的, 如果不是, 奖励应该是负的。如果考虑到管道间隙的中心, 奖励也可以是距离  $d$  的函数, 并且随着  $d$  的变化, 奖励值也不同。

### A. 运动感知

我们的模型不是告诉模型环境的关键元素, 而是处理原始图像输入。提出的运动感知算法将有助于检测障碍物的运动, 并根据尖峰结果计算出障碍物的状态和奖励等关键信息。运动感知结果如图 7 所示。

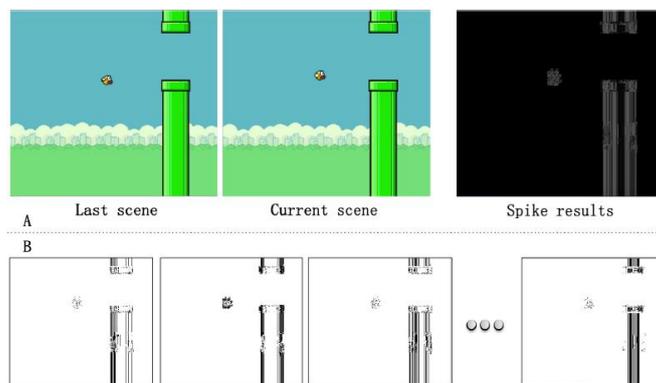


图 7: 飞鸟游戏的运动感知结果。A:运动感知模型将游戏场景作为当前时间步长的输入, 在下一个时间步给出峰值结果。B:每个群编码步骤的尖峰

结果。

实验结果表明，该算法具有良好的输入编码性能。该编码过程将一个像素在一定的时间间隔内转换成脉冲序列。在这个时间间隔中有几十个编码步骤。从图 7B 所示，我们可以看到，运动感知过程可以在单一的时间步长内检测出主要的运动区域。图 7A 中出现了尖峰现象，是所有编码时间步长的结果集合。

## B.玩游戏

我们在“飞鸟”游戏中训练我们的自主学习模式。实验结果如图 8 所示。在 *Hodgkin-Huxley* 模型下，我们的模型可以在 50 次尝试后玩得很好。从图 8B 可以看出，在开始的时候，鸟在碰撞管子之前通常会得低分。经过几次尝试，这只鸟学会了如何避开管道，并在最后一次尝试中获得 30 多分，如果游戏继续进行，甚至会获得更高的分数。这一成就可以与一个优秀的人类玩家相比。

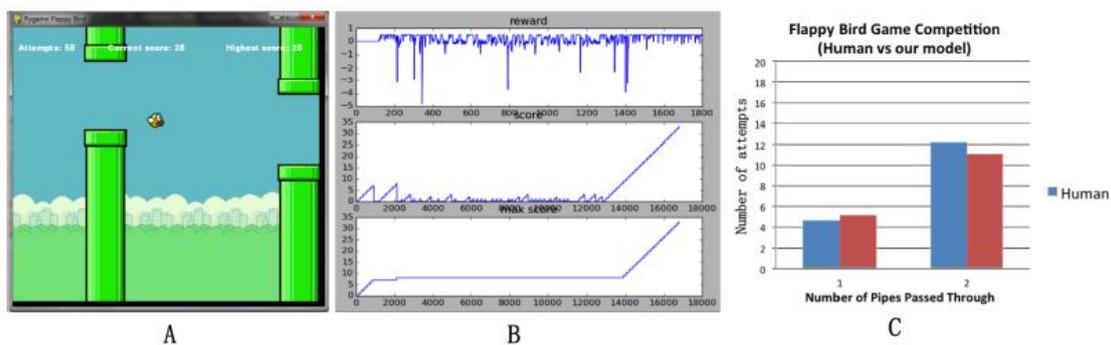


图 8: 使用我们的模型进行游戏的结果。如果鸟撞到管子，就得付出一次尝试的代价。如果鸟通过一根管子，它的分数加 1。A: 游戏场景。B: 学习过程中的分数记录。C: 人和我们的模型的性能比较。

在图 8C 中，我们比较了人类玩家和我们的模型。结果表明，在学习初期，我们的模型具有与一般人类玩家相似的性能。这意味着我们的自主学习模型能够模拟简单的人类决策过程，并得到类似的结果。这对于使用 *spiking* 神经网络建立模型的应用非常重要。

## C. H-H 模型的离子性能测试

*H-H* 方程中的钠和钾离子对生物大脑的认知过程非常重要。图 9 给出了 *H-H* 模型中这两种离子不足的仿真结果。实验表明，缺乏钠或钾离子会导致认知障碍。图 9C 所示，这只鸟只能通过几根管道，即使在较长时间的训练中也不能得到更高的分数。

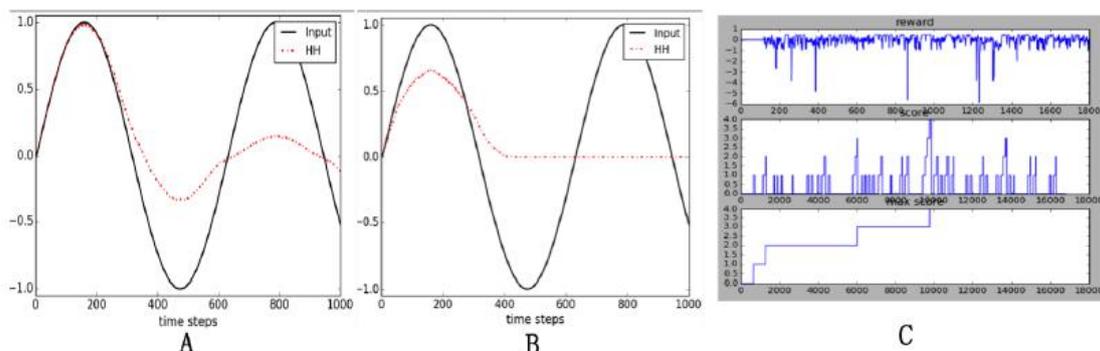


图 9: 钠、钾不足实验。A: 钠离子缺乏。B: 缺钾离子。C: 缺乏钠离子时，玩飞鸟游戏。

## 5 结论

本文提出了一种自主决策模型。该模型是利用 *spiking* 神经网络建立的，可以在不透露大量环境信息的情况下进行自主学习。

在本文中，我们提出了简化的 *Hodgkin-Huxley* 计算模型，以达到接近 *LIF* 模型的计算效率，并保持离子水平特性的认知。为了检测环境中的运动物体，开发了一种基于尖峰的运动感知机制，在不需要大量训练的情况下，从原始像素中提取关键元素。在实验部分，我们将我们的模型应用到飞鸟游戏中，经过不到 60 次的训练后，它表现得很好。此外，我们的模型模拟了 *Hodgkin-Huxley* 模型中阻断钠离子或钾离子通道时的认知缺陷，这是对深入到离子水平的认知的探索。

然而，我们的模型并不完美，它有几个局限性。当背景随时间变化时，不能正确检测运动物体。与大脑相比，它的基底神经节模型还不够真实。我们正朝着这些限制前进。我们将改进我们的模型以获得更好的认知性能。

## 感谢

本研究由中国科学院战略优先研究计划 (*XDB0206007*) 和北京市科学技术委员会 (*Z15110000915070*、*Z161100000216124*) 资助。

## 参考文献

1. Stewart, T. C., Bekolay, T. and Eliasmith, C.: Learning to select actions with spiking neurons in the basal ganglia. *Frontiers in Neuroscience*, vol. 6, no. 2, 1-14 (2012)
2. Bekolay, T. and Eliasmith, C.: A general error-modulated STDP learning rule applied to reinforcement learning in the basal ganglia. *Computational and Systems Neuroscience conference*, Salt Lake City, Utah, 24-27 (2011)
3. Eliasmith, C.: *How to build a brain*. Reprint edition. New York: Oxford, 121-171 (2013)
4. Chakravarthy, V. S., Joseph, D. and Bapi, R. S.: What do the basal ganglia do? A modeling perspective. *Biological Cybernetic*, vol. 103, no. 3, 237-253 (2010)
5. Frank, M. J.: Dynamic dopamine modulation in the Basal Ganglia: a neuro computational account of cognitive deficits in medicated and nonmedicated parkinsonism. *Journal of Cognitive Neuroscience*, vol. 17, no. 1, 51-72 (2005)
6. Uter, A. A. and Basso, M. A.: The basal ganglia: an overview of circuits and function. *Neuroscience and Biobehavioral Reviews*, vol. 32, no. 3, 333-342 (2008)
7. Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., et al: Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nature Reviews Neuroscience*, vol. 11, 760-772 (2011)
8. Stewart, T. C., Choo, X. and Eliasmith, C.: Dynamic behavior of a spiking model of action selection in the basal ganglia. *Proceedings of the 10th International Conference on Cognitive Modeling*, 5-8 (2010)
9. Frank, M. J.: Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks*, vol. 19, no. 8, 1120-1136 (2006)
10. Gurney, K., Prescott, T. J. and Redgrave, P.: A computational model of action selection in the basal ganglia. *Biological Cybernetics*, vol. 84, no. 6, 401-410 (2001)
11. Stewart, T. C. and Eliasmith, C.: Large-scale synthesis of functional spiking neural circuits. *Proceedings of IEEE*, vol. 102, no. 5, 881-898 (2014)
12. MacNeil, D. and Eliasmith, C.: Fine-tuning and the stability of recurrent neural networks. *Public Library of Science (PLoS One)*, vol. 6, no. 9, 1-16 (2011)
13. Gurney, K., Prescott, T. J., Wickens, J. R. and Redgrave, P.: Computational models of the basal ganglia: from robots to membranes. *Trends in Neuroscience*, vol. 27, no. 8, 453-459 (2004)
14. Albin, R. L., Young, A. B. and Penney, J. B.: The functional anatomy of basal ganglia disorders. *Trends in Neuroscience*, vol. 12, no. 10, 366-375 (1989)
15. Bar-Gad, I. and Bergman, H.: Stepping out of the box: information processing in the neural networks of the basal ganglia. *Current Opinion in Neurobiology*, vol. 11, no. 6, 689-695 (2011)
16. Igarashi, J., Shouno, O., Fukai, T. and Tsujino, H.: Real-time simulation of a spiking neural network model of the basal ganglia circuitry using general purpose computing on graphics processing units. *Neural Networks*, vol. 24, no. 9, 950-960 (2011)
17. Cessac, B., Paugam-Moisy, H. and Vi'eville, T.: Overview of facts and issues about neural coding by spikes. *Journal of Physiology-Paris*, vol. 104, no. 1, 5-18 (2010)
18. Dayan, P. and Abbott, L. F.: *Computational and mathematical modeling of neural systems: Model Neurons I: Neuroelectronic*. Cambridge: MIT Press, (2003)

19. Izhikevich, E. M.: *Dynamical systems in neuroscience: the geometry of excitability and bursting*. Cambridge, MA: MIT Press, (2004)
20. Hodgkin, A. L. and Huxley, A. F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, vol. 117, no. 4, 500-544 (1952)
21. Nelson, M. E.: *Electrophysiological Models In: databasing the brain: from Data to Knowledge*. Wiley, New York, (2004)
22. Gerstner, W. and Kistler, W. M.: *Spiking neuron models. Single Neurons, Populations, Plasticity*. Cambridge: Cambridge University Press, (2002)
23. Wells, R. B.: *Introduction to Biological signal processing and computational Neuroscience*. Moscow, ID, USA, (2010)
24. Long, L. N. and Fang, G. L.: A review of biologically plausible neuron models for spiking neural networks. *AIAA InfoTech Aerospace Conference*, Atlanta, GA, 20-22 (2010)
25. Weber, C., Elshaw, M., Wermter, S., Triesch, J. and Willmot, C.: *Reinforcement Learning: Theory and Applications: Reinforcement learning embedded in brains and robots*. Vienna, Austria, (2008)
26. Bohte, S. M., Poutre, H. L. and Kok, J. N.: Unsupervised clustering with spiking neurons by sparse temporal coding and multilayer RBF networks. *IEEE Transactions on Neural Networks*. vol. 13, no. 2, 426-435 (2002)