

基于运动视觉与视差协同计算的迫近感知神经网络

周浩艇¹, 房川凯¹, 刘稔远¹, 彭济根^{1,2}, 傅沁冰^{1,2*}

(1. 广州大学数学与信息科学学院, 广东广州 510006; 2. 广州大学机器生命与智能研究中心, 广东广州 510006)

摘要: 从自然界中动物的正常生存到工业中机器的安全运作, 碰撞感知能力始终都是至关重要的. 受蝗虫视觉神经元 LGMD(Lobula Giant Movement Detector)的启发, 许多仿生的计算模型已经被用于实时可靠的碰撞感知. 然而, 受限于二维单目的输入信号, 目前的方法难以捕捉运动目标的深度特征, 进而无法满足在复杂的真实动态场景下进行迫近感知的需求. 因此, 本研究提出了一种融合生物似然性运动通路和视差通路的三维迫近感知模型. 在突触前神经网络, 通过对两种视觉通路从时空维度上进行实时的神经信号整合, 所提出的模型不仅能够有效排除大范围的背景杂波干扰, 而且可以明显抑制前景非迫近运动所产生的视觉刺激, 同时还降低了对突然出现在感受野的目标的关注度, 进一步提高在未知现实环境中对迫近运动的选择性. 真实场景数据集的离线测试, 以及在线的机器人测试的实验结果显示, 与目前最先进的方法相比, 我们的模型在时间复杂度降低了一个数量级的前提下, 准确率提升至 96.09%, 且能够协助移动机器人在自主导航时实时稳健地检测和避免潜在的碰撞威胁. 本研究综合揭示出, 该迫近感知神经网络对于运动通路的高效性以及视差通路的可靠性, 具备显著协同的兼顾能力.

关键词: 迫近感知; 视差; 蝗虫视觉神经系统; 生物似然; 神经信号融合

基金项目: 国家自然科学基金 (No.62376063)

中图分类号: TP183; TP242

文献标识码: A

文章编号: 0372-2112(2025)xx-xxxx-xx

网址:

DOI:10.12263/DZXB.20250337

A Bio-plausible Neural Network Integrating Motion and Disparity Pathways for Looming Perception

ZHOU Hao-ting¹, Fang Chuan-kai¹, Liu Ren-yuan¹, Peng Ji-gen^{1,2}, Fu Qin-bing^{1,2*}

(1. School of Mathematics and Information Science, Guangzhou University, Guangzhou, Guangdong 510006, China;

2. Machine Life and Intelligence Research Center, Guangzhou University, Guangzhou, Guangdong 510006, China)

Abstract: The capacity to perceive collisions is essential, from the survival of animals in nature to the safe operation of machines in industrial environments. Inspired by the locust visual neuron LGMD (Lobula Giant Movement Detector), numerous biomimetic computational models have been developed for real-time and reliable collision detection. However, constrained by two-dimensional monocular visual input, existing methods struggle to capture the depth features of moving objects, thus failing to meet the demands of looming perception in complex dynamic scenarios. To address this, this study proposes a 3D looming perception model that integrates bio-plausible motion and disparity pathways. In the presynaptic neural network, the proposed model achieves spatiotemporal integration of neural signals from both visual pathways. This not only effectively eliminates background clutter interference but also significantly suppresses visual stimuli caused by non-looming foreground motion, while reducing attention to targets suddenly appearing within the field of vision. Consequently, the model enhances selectivity for approaching objects in unknown realistic environments. The experimental results of offline tests on real scene datasets and online tests on robot validate that our model attains an accuracy of 96.09% while reducing time complexity by an order of magnitude compared with the state-of-art method. Furthermore, it enables mobile robots to detect and avoid potential collisions in real-time during autonomous navigation. The study demonstrates a significantly synergistic fusion of the motion pathway's efficiency and the disparity pathway's reliability accomplished by the proposed neural network.

Key words: looming perception; disparity; locust visual system; bio-plausible; neural signal integration

Foundation Item(s): National Natural Science Foundation of China (No.62376063)

1 引言

近年来,随着智能技术在工业生产中的渗透,无人机、无人驾驶汽车等智能设备的应用,已经在很多实际场景中得到了广泛的推广.作为自主移动机器人的安全保障关键,实时稳健的碰撞检测系统已成为其实际运行中的核心环节[25].然而,在复杂未知的自然环境中,即便是搭载了当前最先进的传感器的智能机器人[26–28],也难以确保能够时刻进行准确无误的碰撞探测.因此,如何有效地减轻外界动态复杂环境对硬件设备中碰撞检测系统的影响,以进一步提升其在实际应用中的稳健性,是当前的一大研究热点.

所谓碰撞检测,就是对即将与观察者在物理上发生碰撞的这一事件,进行的提前预判过程.而作为一种运动模态,迫近则指的是运动物体在深度上逐渐靠近观察者,并且当其迫近的显著性达到一定程度(超过某个阈值)时,即可认为碰撞即将发生.因此,迫近感知是碰撞检测的前提,准确感知迫近是提前探测到潜在碰撞的关键.

截至目前,一些基于深度信息进行场景重建的碰撞感知方法[1–3,61]已经被提出用于优化该系统在实际应用中的表现.然而,尽管立体视觉匹配算法已在硬件层面历经数十年的优化[18,62],但无论是基于 Euclid 距离场[1]还是三维点云[30]的场景重建方法,均需维护大量数据信息,因而强大的计算系统在这类方法下是不可或缺的[29].与此同时,为了进一步提升碰撞检测系统的稳健性,一些研究者也提出了多种基于深度学习的方法[4,5].这类方法通常能够在感知碰撞时展现出相对更高的准确性,但是这往往需要以提高计算成本作为代价.鉴于此,为了在实际应用中更好地服务于硬件系统,我们需要一种更加高效、更低能耗的方法来实现准确的碰撞探测[6].

历经了数百万年的进化过程,自然界中的生物体已经具备了高效躲避障碍物的能力.尤其是只有十万数量级的神经元的昆虫,它们能够仅仅依赖自身的视觉神经系统,来进行实时的碰撞检测等行为.其中最为典型的就蝗虫的视觉神经系统,它们甚至能够在视觉处理能力有限的情况下,实现协调的群体飞行运动[31,32].早在上个世纪,生物学家在对蝗虫大脑的长期解剖学研究中,发现了一组对迫近运动具有选择性的视觉神经元,即小叶巨型运动检

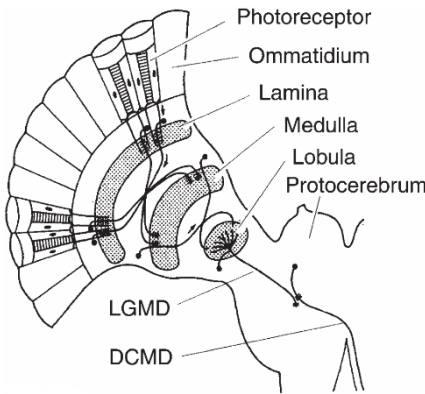


图 1 LGMD 神经元的突触前与突触后的神经连接示意图[53]

测器(lobula giant movement detectors, LGMDs)[24,33],其突触前和突触后结构如图 1 所示.在进一步的研究中,Rind 等人又首次将这类神经元区分成了在形态学上相邻的 LGMD1 神经元和 LGMD2 神经元[34],并且发现了与 LGMD1 神经元不同的是, LGMD2 神经元只对相对于背景来说的暗物体的迫近运动敏感[20].另一方面,在果蝇的视觉神经系统的研究中,生物学家同样发现了一种功能类似的神经元,称为小叶板柱型-II(lobula plate lobula column type-II, LPLC2)神经元,其对来自感受野正中央的迫近物体展现出了极其独特的选择性[9].这些昆虫中简单而可靠的碰撞感知神经机制,为开发兼具准确性与计算效率的碰撞检测方法提供了重要启示.

受生物神经元结构的启发,目前相关学者已经提出了多种实时、高效、低能耗的碰撞感知计算模型[6–8].上世纪末,Rind 等人首次构建了一种神经网络模型,旨在通过兴奋信号和抑制信号之间的相互竞争来刻画 LGMD1 神经元的特性[17].而 Gabbiani 等人[36]则通过电生理学实验的结果,揭示出 LGMD 神经元的放电频率与运动物体的角速度和角大小之间的非线性关系,并在所提出的计算模型中利用对数-指数变换实现了这种乘法运算机制.上述这些早期的研究和猜想,都为后续生物启发的碰撞探测方法的提出和改进,提供了重要支持.然而,目前这些模型大多只基于单目的二维平面信息来构建,在真实动态场景中容易受到感受野中图案花纹和其他非迫近运动模式的干扰.准确地区分出真正的迫近运动与这些视觉噪声之间的不同,能够帮助我们在现实应用中显著提升这类碰撞检测系统的稳健性.因此,本文的研究将尝试突破二维平

面空间的局限,从三维立体空间信息的角度出发,重新刻画蝗虫 LGMD 神经元的特性,以在保证计算效率的前提下,增强仿生计算模型对复杂现实环境的适应能力。

在自然界中,深度感知是动物获取三维立体空间中的距离信息的一种能力,而视差则是进行深度感知的一种重要线索。所谓视差,就是两个不同位置的观测者观察同一目标时,在视觉上所产生的细微差异,这种差异通常也被认为是生物用于感知三维空间结构的途径[41]。从几何学上来看,视差信息和深度信息实际上是等价的[39],并且二者具备一种反比映射关系。一般来说,视差可以分为单目运动视差和双目视差,它们可以单独或者同时作为距离估计的线索[38]。传统生物学观点认为,昆虫复眼的固定透镜结构使其仅能处理像素化的图像,而难以获得具体的深度信息[22,23,37]。然而,随着神经科学研究的深入,研究者发现,蝗虫可以通过反复调整每个小眼单元观察角度(即凝视过程)来获取运动视差,然后进一步来对目标物体进行距离的估计[12,13]。不仅如此,Kral 等人还证实了这一机制能够协助蝗虫在动态环境中,也可以准确地区分出前景的运动目标以及背景的运动目标[13]。另外,在近期的研究中,Kemppainen 等[22]还发现了,果蝇左右眼镜像对称的轻微跳动这一过程,会在时间上产生相位差的信号,其视觉神经系统能够同步对这些信号进行神经编码,进而产生立体视觉,帮助果蝇的大脑在三维空间中精确地计算出目标物体的距离。这些神经科学发现揭示了,昆虫视觉神经系统不仅依赖于二维平面中的亮度变化信息来进行运动感知,三维立体空间中的视差信息也是信号处理过程中的一部分。

受上述生物视差研究文献的启发,本文首次提出了一种融合运动通路和视差通路的计算建模框架,建立了一个全新的生物似然性逼近感知的神经网络,即 LPNN 模型。该模型基于三维空间中的视觉信号,进一步解决了目前基于二维平面空间信号的单目模型在现实动态复杂场景中稳健性的欠佳。在实时的视觉信号处理过程中,所提出的模型需要一张灰度图像和一张对应的视差图像,分别作为运动通路以及视差通路的输入信号。随后,为了进一步增强模型对逼近物体的选择性,来自两个神经通路的视觉信号将在 LGMD 神经元的突触前区域,实时地进行三处重要的信息融合:

- 1) 对于运动通路所捕获的初始运动信息,借助于视差通路中的实时信号,我们将更为精确地过滤掉对应深度值较大的背景部分,只提取更为重要的前景变化信息,作为后续神经元的输入信号,从而在不失去重要运动信息的基础上最大化地去除背景噪声杂波的干扰。

- 2) 基于视差的变化信号,我们将对视野内所提取到的运动边缘信息进行重新整合,建立一种运动边缘信号的划分机制和空间上的相互竞争机制,使得最终只保留不断扩张的边缘信号进入到 LGMD 神经元,以应对在绝大部分现实场景中的复杂运动下的碰撞感知。

- 3) 在前馈抑制神经元的编码中,我们的模型首次引入了视差的变化信息,从亮度的变化和视差的变化两个角度,来刻画感受野内突然发生的大范围变化,进一步加强对感受野的边缘区域瞬间产生的运动刺激的抑制作用,降低神经元对突然出现的目标的关注度。

系统性的离线实验验证了,基于这种计算建模框架的 LPNN 模型,即便是在复杂动态场景(如具有抖动光流、光影噪声等视觉杂波)中也具备很强的鲁棒性,只有当移动物体处于不断逼近的过程中才会被明显地激活。不仅如此,所提出的模型在机器人的在线闭环测试,也证实了能够为移动机器人在自主导航时提供一套稳健而又实时高效的嵌入式碰撞感知系统。值得一提的是,与仅依赖于视差深度的方法相比,我们的模型在时空上进行了协同计算,能够显著区分出不同类型的运动模态(如逼近、平移、远离),进而实现更为可靠的碰撞检测。一个典型的例子是,对于视差较大(即距离较近)的平移物体,基于距离估计和简单阈值分割的方法容易将其误判为潜在的碰撞威胁,而忽略了其并非处在逼近过程中。进一步地,这种时空关联运算有助于准确预测即将发生的碰撞,为感受野内快速靠近的物体提供可靠的早期预警[40]。

在本文接下来的章节中,我们将回顾与本研究相关领域的发展。第三节将详细介绍所提出的模型的计算方法以及具体的参数设置。然后,我们将给出具体的实验设置、模型的对比标准,以及实验结果的分析。最后,第五个章节总结了本研究的结论,并探讨了未来的研究方向。

2 相关工作

2.1 生物似然的碰撞感知方法

随着生物似然的碰撞感知模型的发展,目前在这方面的生物计算建模,主要是为了探寻一些实时、可靠、低能耗的检测方法[6],因此这方面的建模工作需要平衡算法的准确度与复杂度。在此背景下,Silva 等人[59]基于蝗虫复眼的像素重映射,曾提出了一种改进的碰撞感知计算模型,有效提升了其的抗噪能力以及处理速度。随后,受生物上 ON/OFF 机制[14]的启发,Fu 等人从不同极性的亮度变化的角度,重新构建了一种基于 LGMD1 神经元结构的网络框架[10]用于感知碰撞,并且首次提出了一种 LGMD2

碰撞感知模型[11], 实现了幼年期的蝗虫大脑中 LGMD2 神经元的独特选择性[20]. 同样, 受到了果蝇视觉神经系统的启发, Hua 等人的研究[48]通过融合视野坐标域中四个不同方向的基本运动检测器 EMD[49]的输出, 成功实现了果蝇 LPLC2 神经元的特殊选择性[9].

为了进一步提升这类碰撞检测方法的鲁棒性, 近年来, 许多实际应用效果更好的计算模型被陆续提出. 其中, Lei 等人的研究[50]提出了一种基于 LGMD 视觉神经模型, 旨在通过 ON 通道的信号和 OFF 通道的信号之间的相互竞争机制, 来排除平移运动的干扰. 另外, Li 等人在近期工作[35]中, 构建了一种果蝇视觉联合的人工感知神经网络, 根据抛物型曲线对视野域的区域划分, 该模型能够同时检测多个区域的碰撞威胁. 然而, 上述这类基于二维平面信息来感知运动的碰撞检测方法, 在实际复杂多变的自然环境中的应用, 仍然是难以满足需求的. 因此, 一些研究者开始尝试从生物的立体视觉的角度, 探寻一些稳健性更高的碰撞感知方法. 作为一种节肢动物, 螳螂是目前为止唯一一种被证实能够借助立体视觉来进行深度感知的昆虫[42]. 最近, Rosner 等的研究[43,44]也首次在螳螂的大脑中发现了对双目视差特别敏感的神经元的存在, 这些神经元能够对视差和偏心率进行特定的编码, 从而估计目标物体在三维世界中的位置. 虽然其处理视差信息的完整神经回路尚不明了[43], 但是受此启发, O’Keefe 等人[46]基于一定的猜想也构建出了一个同时对目标物体的视差和大小敏感的双目神经元模型, 并且在一定程度上模拟了螳螂在捕食中对猎物的距离和大小的感知过程. 除此之外, Zheng 等人在最新的研究[16]中, 提出了一种双目结构的 LGMD 碰撞检测模型, 旨在通过对左右图像中运动边缘信息的自相关运算, 来获取运动物体的视差, 进而估计出对应的深度值以用于更为可靠的碰撞感知.

总而言之, 因为视差与深度距离之间存在对应关系, 所以其能够更为直观地体现运动物体准确的立体信息而不受无关运动的影响. 因此, 在我们的研究工作中, 运动信息与视差信息的相互结合, 能够在实际复杂多变的自然环境中, 协助基于单目视觉的运动感知神经, 更为准确地从杂乱环境中提取出正在逼近的目标物体.

2.2 机器人中的生物启发式碰撞检测

相比于其他类型的方法, 由于仿生碰撞检测方法具备更高的计算效率, 因此其能够为计算资源有限的移动机器人提供一套实时的避障闭环系统. 截至目前, 许多受生物启发的方法, 已经在地面移动机器人中进行了嵌入式开发与应用[10,11,45,60]. 值得一提的是, 基于前期的建模工作[51], Fu 等人[52]曾将蝗虫的 LGMD1 和 LGMD2 两种神经元与果蝇的两个方向选择性神经元进行结合, 提出了一种不易受到平移运动和远离运动影响的嵌入式碰撞探

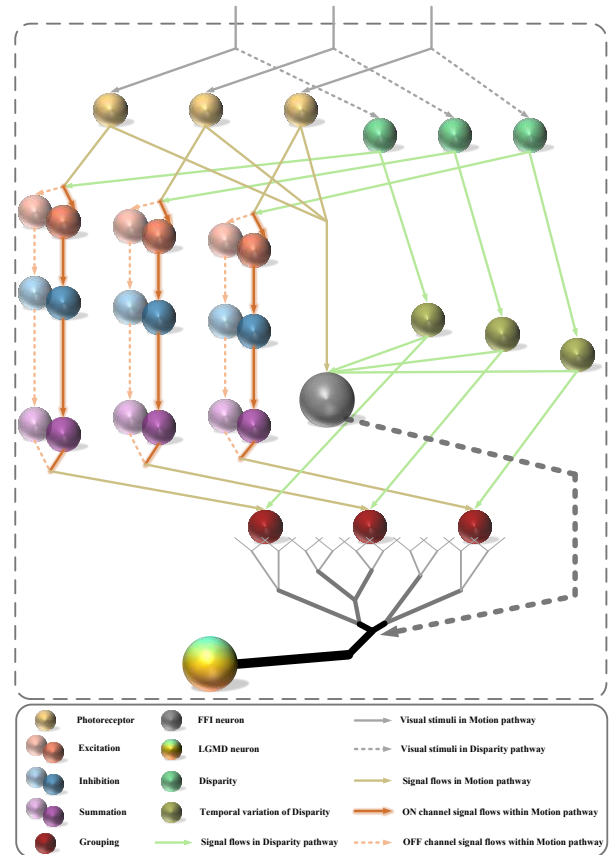


图 2 所提出的 LPNN 神经网络模型的架构图. 该神经网络包含两个视觉神经通路, 黄色实线代表的是其中的运动通路所处理的信号流, 对应的是二维空间中的亮度信号; 而绿色实线则表示视差通路所处理的信号流, 对应的是三维空间中的视差信号. 在该计算框架下, 运动通路提取的是感受野内物体的运动边缘特征, 视差通路提取的是运动物体动态距离特征, 二者的神经信号在突触前的前馈神经网络中被重新整合. 随后, 融合了两种视觉特征的神经信号, 再传递给 LGMD 细胞进行突触后的神经编码, 并最终用电脉冲的形式作为输出信号.

测系统, 并且在微型机器人中验证了其的实时可靠性. 而 Čížek 等人的工作[58], 则提出一种基于 LGMD 神经元和自监督学习的视觉避障系统, 实现了六足机器人在真实复杂场景下的自主避碰. 不仅如此, 一些生物启发式的碰撞检测方法[54–56]还能够协助无人机, 在飞行导航过程中进行实时有效的自主避障; 甚至在无需建立相互通信的情况下, 也能够实现多无人机的自主群体飞行[57].

然而, 目前这些仅仅依靠单目二维图像的嵌入式系统, 对于自然环境中复杂视觉噪声的影响, 仍然是难以避免的. 因此, 考虑到基于深度信息的避障策略已经能够在机器人的自主路径规划中得到很好的应用[30,47], 我们的工作聚焦于运动视觉信息和与视差深度信息的相互融合, 以同时兼顾前者的高效性和后者的可靠性, 并且在实验过程中利用 TurtleBot 机器人验证了我们方法的可行性, 解决了目前生物启发式碰撞探测系统在机器人自主导航时,

难以适应复杂多变的真实场景的问题。

3 模型描述

在本章节中，我们将详细介绍所提出的神经网络模型 LPNN 的整体架构，以及具体的计算方法。如图 2 所示，由于视差是生物进行深度估计最为重要的线索[15]，因此所提出的计算模型将在 LGMD 细胞的突触前区域融合视差信息，通过整合运动通路和视差通路两个视觉神经通路的信息，达到减弱无关噪声的干扰的效果，同时提高对不断靠近的物体的选择性。

以视频序列作为输入的视觉信号，所提出的 LPNN 模型涵盖了在视网膜层、椎板层、髓质层、前馈抑制层(Feed-Forward Inhibition, FFI)以及小叶层上的仿生计算，并且最终以电脉冲的形式作为输出信号。值得注意的是，神经网络中的第二、三、四层结构，是对来自于运动通路和视差通路的神经信号进行融合的核心过程。

3.1 视网膜层仿生计算

以实时的视差图像作为输入信号，视网膜中的视差通路对应的是中视差信号的计算处理过程，其神经元细胞按二维矩阵的形式而排列。由于视觉传感器所提供的视差信息存在由外界环境所引起的一定误差，因此对于输入的视差信息，首先需要进行高斯滤波的预处理，即：

$$\widetilde{DP}(x, y, t) = \iint_{\Lambda} DP(u, v, t) G_{\sigma}(x - u, y - v) du dv \quad (1)$$

其中， G_{σ} 是一个半径为 5、标准差为 σ 的高斯方程， Λ 是局部的 5×5 邻域。

随后，计算连续帧之间的差分信号，以刻画视差的变化信号。计算公式如下定义：

$$\Delta DP(x, y, t) = \widetilde{DP}(x, y, t) - \widetilde{DP}(x, y, t - 1) \quad (2)$$

预处理后的视差信号 $\widetilde{DP}(x, y, t)$ 和视差的变化信号 $\Delta DP(x, y, t)$ ，都将用于后续神经层中运动通路和视差通路的信息融合。

而在视网膜层的运动通路中，首先需要对感受野内发生的运动进行初步的感知，该过程由感光细胞完成。感光细胞是以二维矩阵形式排列而形成的，其捕获的是输入图像的亮度信息(即相机所提供的灰度信息)，然后对其进行连续两帧的差分运算，便可得到初步的运动信息，如下公式所示：

$$P(x, y, t) = L(x, y, t) - \int L(x, y, \tau) \delta(t - \tau - 1) d\tau \quad (3)$$

其中 δ 表示单位脉冲函数， $L(x, y, t)$ 为空间上的像素位置

(x, y) 处在 t 时刻的亮度信号，而 $P(x, y, t)$ 则代表感光细胞所捕获的亮度变化信息。

由于动态复杂场景下的背景运动信息将带来不可估量的噪声影响，因此对于感光细胞的输出信号，我们将提取出其中更值得注意的前景运动信息来进行后续的处理，即本文所提出的前景提取机制(Foreground Extraction Mechanism, FEM)。由于早期的研究发现，蝗虫善于利用运动视差来进行对目标物体的深度距离感知[12]，并且在动态场景下仍能够借助于运动视差来准确区分前景和背景的运动信息[13]，因此我们将借助于视差通路中预处理后的视差信息(公式 1)，在此处准确地提取出运动通路中的前景运动信息(如图 3(a)(c)所示)。所定义的前景信号提取方式为：

$$\tilde{P}(x, y, t) = \rho \cdot P(x, y, t) \quad (4)$$

根据分离阈值 T_1 ，公式 4 的系数 ρ 可以定义为：

$$\rho = \min(1, e^{\beta(\widetilde{DP}(x, y, t) - T_1)}) \quad (5)$$

3.2 椎板层极性视觉仿生计算

之后，在椎板层中，对于分离出的前景运动信号，根据半波整流的机制，我们将负(OFF)极性信号转换为正极性，而正(ON)极性信号则保持不变。对应的处理过程由下式给出：

$$\widetilde{P}_{on}(x, y, t) = [\tilde{P}(x, y, t)]^+ \quad (6)$$

$$\widetilde{P}_{off}(x, y, t) = [\tilde{P}(x, y, t)]^- \quad (7)$$

其中 $[x]^+ = \max(0, x)$ ， $[x]^- = \max(0, -x)$ 。分离后的信号分别流入 ON 通道(亮度增加)和 OFF 通道(亮度减少)，并在后续的神经层中并行计算[14]。

3.3 髓质层兴奋抑制时空交互仿生计算

在髓质层的运动通路中，对于其 ON 通道和 OFF 通道中的信号，兴奋信号和抑制信号的计算方式类似，都是用于提取物体运动的边缘信息。如图 3 (a)所示，以 ON 通道为例，兴奋信息接收的是前景运动信息经过半波整流后的信号：

$$E_{on}(x, y, t) = \widetilde{P}_{on}(x, y, t) \quad (8)$$

而侧向的抑制信息则是通过对时间上延迟的兴奋信号进行局部的卷积操作来刻画：

$$I_{on}(x, y, t) = \iiint_{\Pi} W_1(x - u, y - v, t - \tau) \times \widetilde{P}_{on}(u, v, \tau) du dv d\tau \quad (9)$$

此处的 Π 为局部的 3×3 邻域， W_1 是从时空维度上刻画的一个卷积核，可表示成：

$$W_1(x, y, t) = \psi(t) \cdot \Gamma(x, y) \quad (10)$$

其中， $\psi(t)$ 表示的是一个基于一节低通滤波的时延函数[10]，其延迟时间（单位：毫秒）在空间上的分布为：

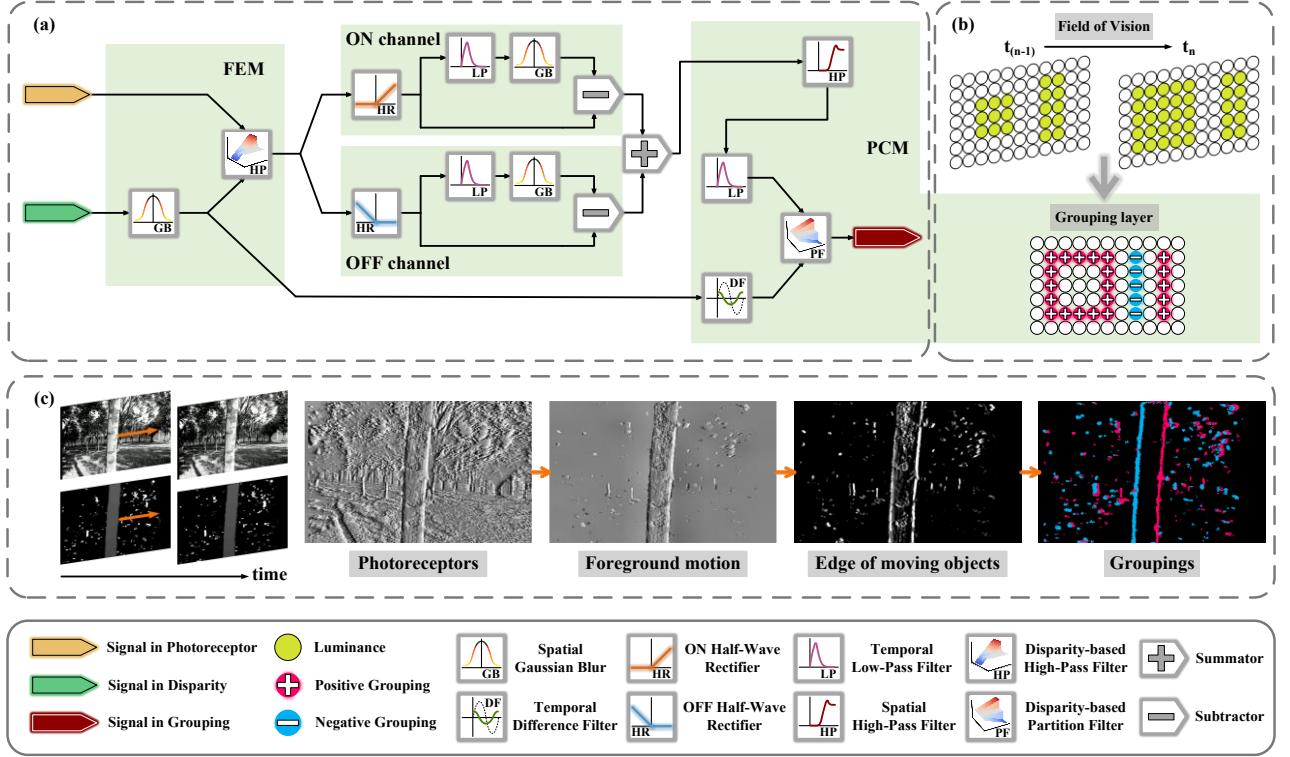


图 3 所提出的模型对视觉信号的实时处理流程图. 子图(a)展示的是突触前区域的每个视觉神经单元中神经信号的传输与编码过程, 包含了运动通路中的 ON/OFF 极性通道以及两处重要的神经通路整合过程. 其中, 基于视差的前景提取机制(FEM)被用于准确过滤感光细胞中无关的背景运动信号; 而划分-抑制机制(PCM)则对信号增强单元中的运动边缘信号进行了极性的划分, 以便它们在空间上的相互竞争来抑制非迫近的前景运动所带来的视觉刺激. 子图(b)以两种基本运动模式(迫近与平移)为例, 展示了在 PCM 机制中, t_n 时刻正向的 Grouping 单元与负向的 Grouping 单元在视觉空间上的相互竞争关系. 子图(c)可视化了基于灰度图像和视差图像作为输入信号的 LPNN 模型中间神经层的输出结果.

$$\begin{bmatrix} 60 & 30 & 60 \\ 30 & 0 & 30 \\ 60 & 30 & 60 \end{bmatrix} \quad (11)$$

空间卷积核 $\Gamma(x,y)$ 的权重分布则定义为:

$$\Gamma = \begin{bmatrix} 1/8 & 1/4 & 1/8 \\ 1/4 & 0 & 1/4 \\ 1/8 & 1/4 & 1/8 \end{bmatrix} \quad (12)$$

随后, ON/OFF 通道中的兴奋信号和抑制信号通过线性求和的方式, 来体现出二者之间的一个竞争机制:

$$S_{on}(x, y, t) = E_{on}(x, y, t) - \omega_1 \cdot I_{on}(x, y, t) \quad (13)$$

$$S_{off}(x, y, t) = E_{off}(x, y, t) - \omega_2 \cdot I_{off}(x, y, t) \quad (14)$$

其中的 ω_1, ω_2 均为抑制信号的权值.

同时, 重新整合 ON 通道和 OFF 通道中的求和信号, 即可得到运动通路中所输出的求和信号值. 具体的数学公式为:

$$S(x, y, t) = \theta_1 \cdot S_{on}(x, y, t) + \theta_2 \cdot S_{off}(x, y, t) \quad (15)$$

上式中的 $\{\theta_1, \theta_2\}$, 代表这个整合过程中每一项的系数.

为了进一步增强上述信号(公式 15)中有效的运动信号, 我们首先需要对其进行卷积操作:

$$G(x, y, t) = \iint_{\Pi} S(u, v, t) W_2(x - u, y - v) du dv \quad (16)$$

其中,

$$W_2 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (17)$$

为一个等权重的卷积核[7,10,11]. 然后, 通过设置一个过滤阈值 T_2 , 去除掉其中孤立的兴奋信号, 同时增强了相对而言更为聚集的兴奋信号. 具体公式为:

$$\hat{G}(x, y, t) = \begin{cases} G(x, y, t), & \text{if } G(x, y, t) \geq T_2 \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

至此, 我们已经将前景运动物体的运动边缘信息完整地提取出来了. 然而, 这些边缘信号中仍然保留着由于快速平移运动、左右两侧的经过物体、抖动光流等视觉杂波所引起的兴奋信号, 而这些噪声信号使得 LGMD 神经元被提前激活的问题, 在目前基于单目二维平面信息的 LGMD 模型中还是难以解决的. 因此, 如何利用来自视差通路中的信息来去除这些噪声信号的干扰, 是我们的工作的关键.

如图 3 (a)(b) 所示, 我们提出一种对于运动边缘信

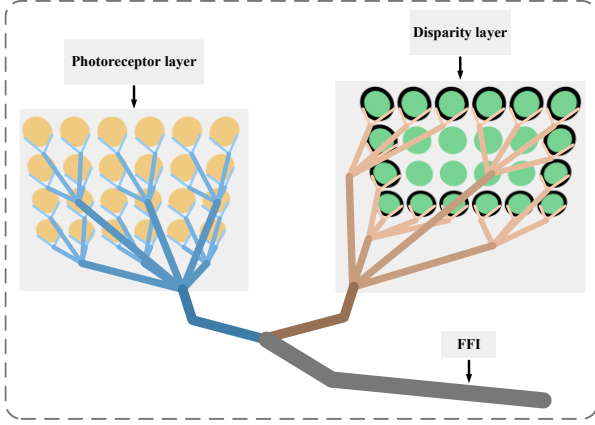


图 4 LPNN 模型中 FFI 神经元对来自运动通路和视差通路的视觉神经信号的融合过程。其中，在运动通路中整合的是感受野内所有感光细胞的信号，而在视差通路中则仅整合来自感受野边缘区域(本文所定义的边缘区域为外围占整个感受野 20%的区域)内的信号。随后，二者再进行线性的求和以得到最终 FFI 神经元的输出信号。

息的划分-竞争的机制(Partition-Competition Mechanism, PCM). 采用一阶低通滤波的方式, 我们首先对增强后的运动信号(公式 18)进行一定时间 τ_1 的延迟:

$$\frac{dG'(x, y, t)}{dt} = \frac{1}{\tau_1} (\hat{G}(x, y, t) - G'(x, y, t)) \quad (19)$$

随后, 融合视差通路中输出的视差变化信号(公式 2), 便可对运动通路中的延迟信号 $G'(x, y, t)$ 根据阈值 T_3 ($T_3 > 0$) 进行边缘信息的划分。具体操作为, 对于视差变化信号为小于 $-T_3$ 的, 将其对应的运动边缘信号也转换为负信号; 而若视差变化信号在 $-T_3$ 到 T_3 之间, 则认为其对应的运动边缘无效, 并将其信号值设为 0; 最终只保留 $\Delta DP(x, y, t) \geq T_3$ 所对应的运动边缘信号不变。该处理过程的数学公式定义如下所示:

$$\tilde{G}(x, y, t) = \left\lfloor \frac{\text{sgn}(\Delta + T_3) + \text{sgn}(\Delta - T_3)}{2} \right\rfloor G'(x, y, t) \quad (20)$$

其中, Δ 表示的是视差的变化信息 $\Delta DP(x, y, t)$ 。这样, 我们将所提取出来的运动边缘(公式 18)划分为靠近边缘(正信号)、远离边缘(负信号)以及无效边缘。

由于运动边缘信号中负信号的存在, 因此在后续的池化过程中, 经过不同类型的运动边缘信号之间的相互竞争, 大部分的噪声信号将会在空间上被这些负信号所抑制, 从而最终只剩下了不断靠近的物体的运动边缘信号, 如图 3(b)(c)所示。

3.4 前馈抑制仿生计算

在前馈抑制(FFI)机制中, 如果感受野内突然发生了大面积的变化, 这种神经元可以直接抑制 LGMD 神经元被迅速激活。然而, 与之前的建模方法[7, 10]略有不同, 我

们的模型是从亮度变化和视差变化两个方面来同时更为全面地刻画这种“视野内的大面积变化”。如图 4 所示, 通过整合运动通路和视差通路的信号, 所提出的模型可以进一步增强特别是对感受野的边缘区域突然产生的运动刺激的抑制作用。这一过程在数学上的定义是计算所有感光细胞所捕获的绝对亮度变化的平均值, 以及感受野边缘区的平均视差变化的绝对值:

$$\begin{aligned} F_1(t) &= \frac{1}{C \cdot R} \int_1^C \int_1^R |P(x, y, t)| dx dy \\ F_2(t) &= \left| \frac{1}{N_{cell}} \iint_{\Omega} \Delta DP(x, y, t) dx dy \right| \end{aligned} \quad (21)$$

其中, Ω 代表的是感受野的边缘区域, N_{cell} 对应的是该边缘区域的神经细胞数量, 而 C 和 R 则分别表示输入的二维图像矩阵的列数和行数。

然后, 分别对其进行一定时间 τ_2 的延迟:

$$\frac{dF'_i(t)}{dt} = \frac{1}{\tau_2} (F_i(t) - F'_i(t)) \quad (i = 1, 2) \quad (22)$$

并且根据初始设定的阈值, 重新整合来自两个神经通路的信号, 得到前馈抑制神经元的输出信号:

$$FFI(t) = [F'_1(t) - T_4]^+ + [F'_2(t) - T_5]^+ \quad (23)$$

其中, $[x]^+ = \max(0, x)$ 。

一旦前馈抑制神经元在当前时刻被激活, 那么最终输出的膜电位 $K(t)$ 将会被设置到静息电位:

$$K(t) = 0.5, \text{ if } FFI(t) > 0 \quad (24)$$

3.5 小叶层仿生计算

输入信号在经过上述的突触前部分的处理后, 将到达 LGMD 神经元。小叶层的神经元细胞将对突触前区域的输出信号进行重新的整合, 然后得到对应的膜电位:

$$k(t) = \left[\int_1^C \int_1^R \tilde{G}(x, y, t) dx dy \right]^+ \quad (25)$$

其中, $[x]^+ = \max(0, x)$, C 和 R 分别代表输入的二维图像矩阵的列数和行数。随后, 将该膜电位映射到 $[0.5, 1]$, 以得到最终在 t 时刻的输出 $K(t)$ 。具体的映射公式为:

$$K(t) = (1 + e^{-k(t) \cdot (C \cdot R \cdot \gamma)^{-1}})^{-1} \quad (26)$$

公式中的 γ 为一个缩放系数。

最后, 一旦膜电位 $K(t)$ 超过了初始设定的阈值 T_6 , 那么当前 t 时刻将会产生一个脉冲信号:

$$Spike(t) = \begin{cases} 1, & \text{if } K(t) \geq T_6 \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

与此同时, 为了更加方便有效地判断当前是否处于碰撞危险时刻, 我们定义如下公式:

$$Col(t) = \begin{cases} 1, & \text{if } \int_{t-t_n}^t Spike(\tau) d\tau \geq N_{sp} \\ 0, & \text{otherwise} \end{cases} \quad (28)$$

当产生的脉冲能够持续至少 t_n 个时刻, 那么就认为当前 t

时刻极有可能会发生碰撞，并且该神经层将发出一个碰撞预警信号，然后传递到机器人的运动控制系统中，以便机器人对靠近的障碍物进行及时的避碰。

3.6 模型的参数设置

所提出的 LPNN 模型，在上述计算过程的具体分析中用到的参数设置，均如表 1 所示。值得注意的是，在其中的前景提取机制中，过滤阈值 T_1 的选择很重要，过小的阈值可能会导致较多的背景噪声信号流入后续的神经元，而过大的阈值反而会使得神经元无法提前预判潜在的碰撞危险。另外，在真实动态场景尤其是抖动环境下测试或应用的时候，应当在给定区间内适当地提高模型中的阈值 T_4 和 T_5 ，以避免过度激活前馈抑制神经元。

表 1 典型结构参数

Parameter	Value	Parameter	Value
τ_1, τ_2	30~60ms	ω_1	0.3
σ	1	ω_2	0.6
T_1	6~10px	β	2
T_2	10	γ	0.7
T_3	1~2px	θ_1	0~1
T_4	15~40	θ_2	0~1
T_5	0.8~4px	t_n	3
T_6	0.7	N_{sp}	3

4 实验结果与分析

为了验证所提出的模型在不同场景下的有效性，本节将在多种场景下系统性地设计一系列的实验，并且对实验结果进行详细地分析和总结。所设计的实验可以分为：(1)离线实验，(2)在线实验。其中，离线实验还包括了真实静态场景下的实验，以及真实动态场景下的实验。

4.1 实验设置方法

在第一类的离线实验中，进行测试的所有视觉刺激，均由 OAK-D Lite 深度相机所拍摄捕获的视频序列而生成。其中，该相机的焦距参数为 451.91 像素，基线长度为 0.075 米。对于这些视频序列，每一帧图像的分辨率为 640×400 像素，采样频率为 30 Hz。需要注意的是，相机所捕获的灰度图像序列和对应生成的视差图像序列，将共同作为 LPNN 模型的输入信号。为了更加直观地凸显所提出的模型的稳健性，团队前期的研究中所提出的 LGMD1 模型[10]、LGMD2 模型[11]、Bi-LGMD 模型[16]，将被用于与我们的模型在离线实验中进行性能的对比如分析。

具体而言，在我们设计的离线实验中，静态场景下的视觉刺激由小球滚动而形成，包括滚球的靠近运动、近距离平移运动以及远离运动。另一方面，对于动态场景下的视觉刺激，我们以 Turtle-Bot4 机器人(图 5)的第一

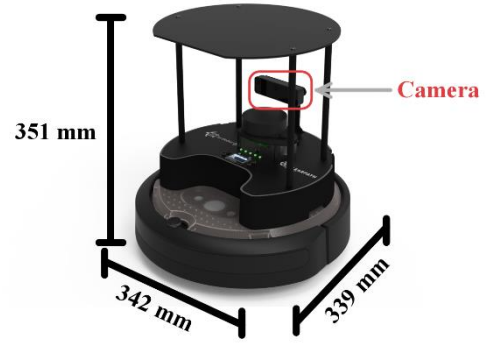


图 5 实验中所使用的 TurtleBot 4 机器人的示意图

视角，录制其在室内移动的整个过程，来模拟现实中汽车行驶的真实场景；同时，我们还在室外移动录制了更为复杂且极具挑战性的视频，来作为模型测试的另一类数据集。值得一提的是，为了尽可能地接近现实世界可能发生的各种复杂情景，相机均是处于一定的晃动状态下来进行移动拍摄的。

而在另一类的在线实验中，我们将所提出的 LPNN 模型在移动机器人 TurtleBot 4(图 5)中进行嵌入式开发，并且进行了在线的机器人闭环测试，以评估所提出的算法在真实物理环境中碰撞感知的可行性与稳健性。需要注意的是，机器人在实验中实时处理的图像分辨率为 320×200 像素。

4.2 模型对比标准

在对比分析中，为了量化所提出的模型与另外三个比较模型在进行碰撞感知时的有效性，我们根据总共 34 组离线测试的数据结果，引入两个具体的衡量标准，即准确率(Accuracy)和召回率(Recall)，来综合评估模型在测试数据集集中的表现。其中，准确率的数学计算公式被定义为：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (29)$$

召回率的数学计算公式则被定义为：

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (30)$$

具体来说，在已被测试的数据中，TP(true positives)代表的是模型所生成的正确碰撞警报的数量，FP(false positives)表示的是在非碰撞场景下模型生成的错误警报数量。另一方面，TN(true negatives)表示模型所能识别出的非碰撞时刻的数量，而 FN(false negatives)则代表在碰撞场景下模型无法正确检测出碰撞的数量。

4.3 真实静态场景实验

在真实静态场景的实验中，我们采用滚球的运动来产生一定的视觉刺激，这些运动可以分为三类：靠近运动、远离运动、近距离平移运动，如图 6 所示。实验结果表明，

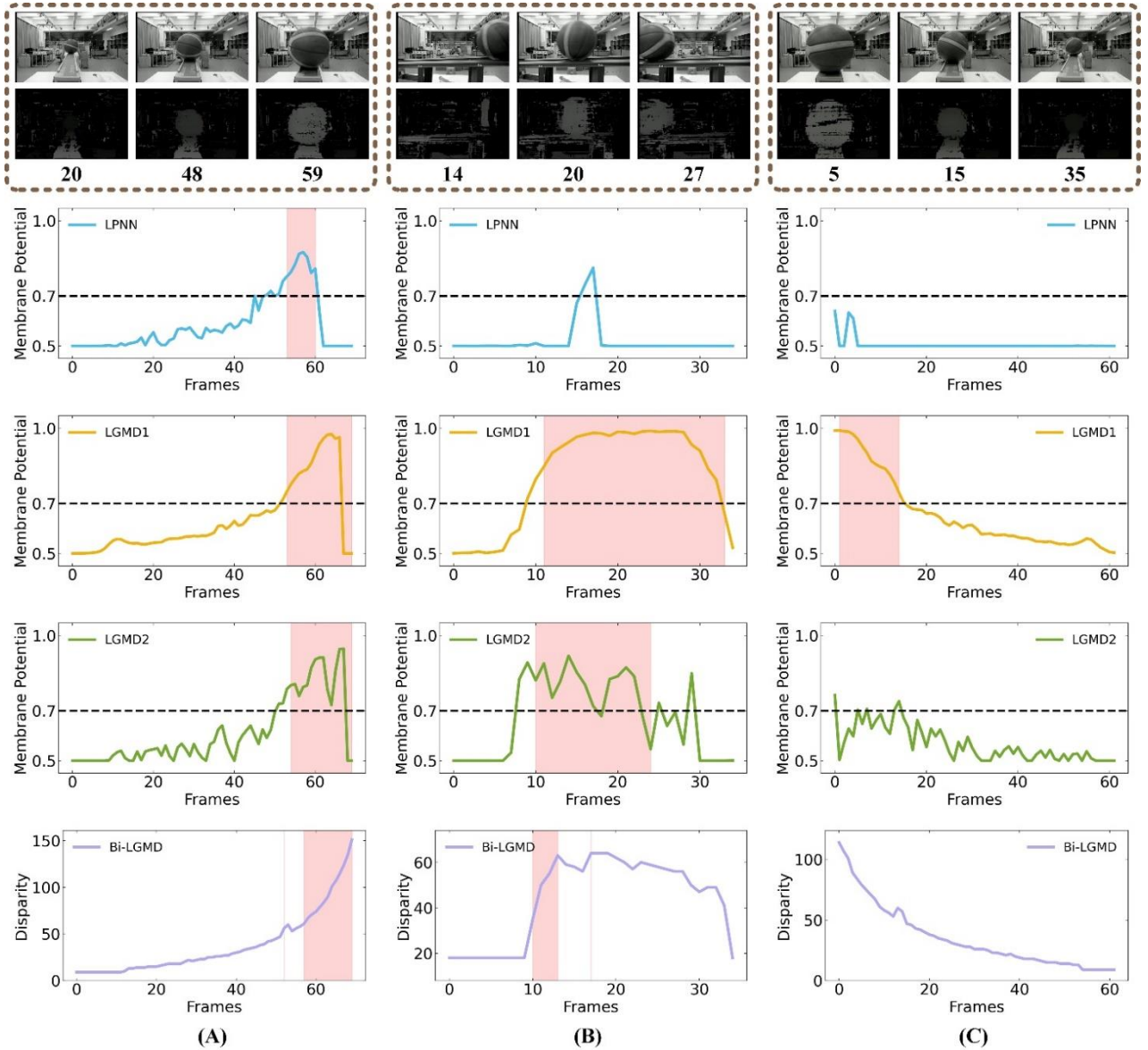


图 6 所提出的模型与另外三个对比模型对于真实静态场景下的视觉刺激的反应结果，红色区域代表的是测试过程中模型产生碰撞警报的时间窗口。每个子图的第一行均为所输入的视觉刺激中部分帧的展示。其中，子图(A)实验、子图(B)实验、子图(C)实验的视觉刺激，分别来源于滚球的迫近运动、近距离平移运动、远离运动。在这三种基本运动的刺激下，LPNN 模型展现出对迫近运动相对更高的选择性。

所提出的模型保留了 LGMD 模型对靠近物体的选择性，并且与另外三个模型相比都能在接近同一时间开始产生碰撞警报。不过，由于在近距离双目视觉的立体匹配的失效，即 OAK-D Lite 深度相机无法正确提供物体在一定距离范围内的视差信息，因此，相比于其他三个比较模型，当正在靠近的滚球运动到距离相机一定距离之内，LPNN 模型输出的膜电位曲线将提前出现拐点，并且开始迅速下降到达静息电位。

另一方面，对于滚球的近距离平移运动，相比于比较的模型，我们的模型由于划分-竞争机制的存在而显得更为平静，结果也更加接近于早期生物学上对 LGMD 神经元的研究发现[20]。值得一提的是，其中的 Bi-LGMD 模型

也仅在平移的滚球刚出现在感受野的期间产生少量的脉冲，而在后续的运动过程中，由于其对深度距离不变的平移物体计算所得到的视差基本保持不变[16]，因此也并未产生任何的警报响应。总的来说，这两个模型都基本上解决了所对比的 LGMD1 模型和 LGMD2 模型无法区分出近距离平移运动的这一问题。最后，在图 6 的滚球远离实验中，仅有 LGMD1 模型在运动的初期产生了较多的碰撞警报，显然这些都是由于错误激活 LGMD1 神经元而生成的假预警。相反，LGMD2 模型对这种暗物体的远离运动并不敏感；而所提出的模型与 Bi-LGMD 模型则由于结合了视差计算的理念，均没有产生任何的脉冲信号。

4.4 真实动态场景实验

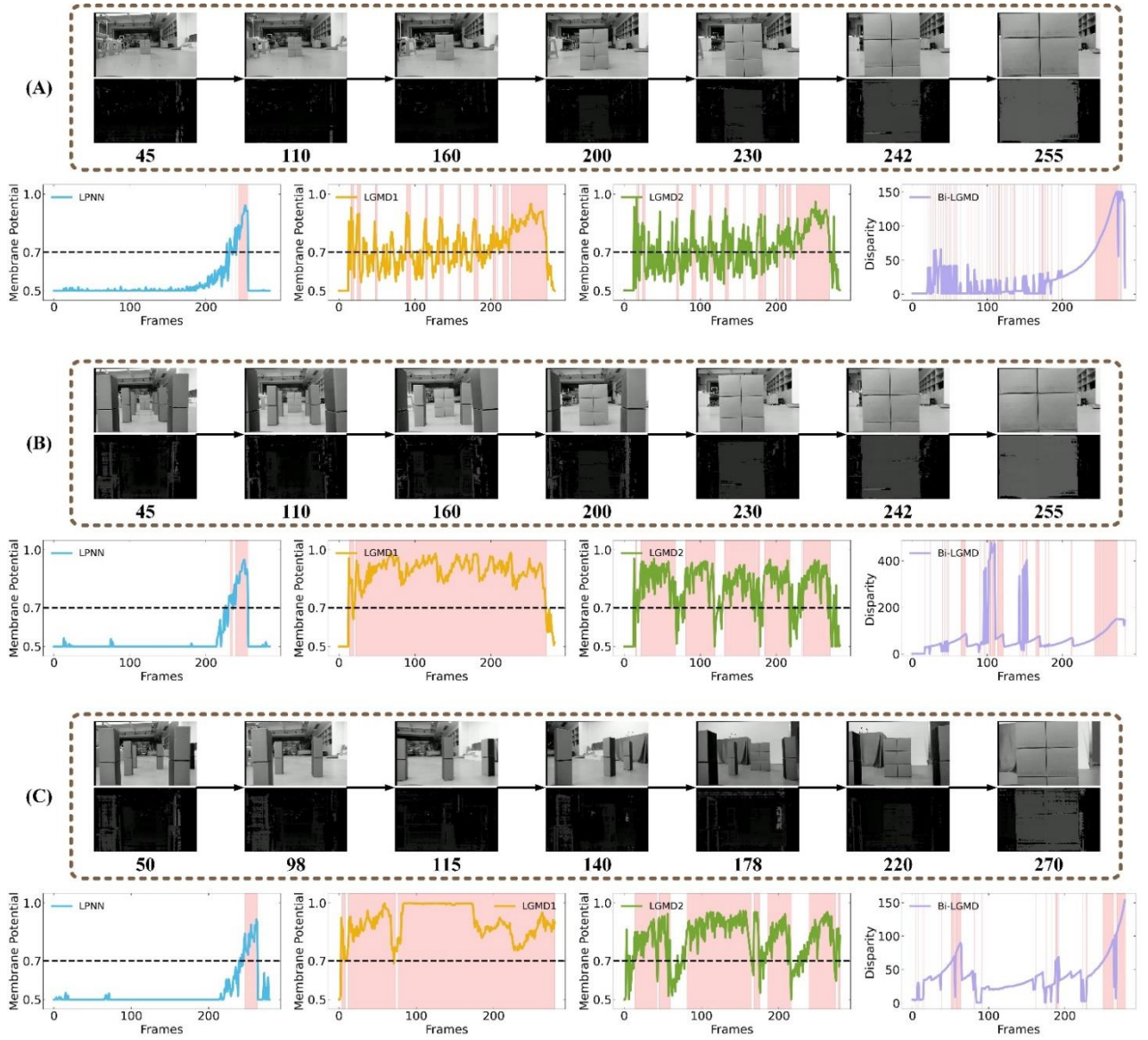


图 7 在机器人的自主移动过程中所捕获的视觉刺激下，所提出的模型与另外三个对比模型的输出结果，红色区域代表的是测试过程中模型产生碰撞警报的时间窗口。这些室内的动态场景视觉刺激来源于以下场景：(A)机器人在行驶路径上无干扰的情况下驶向一个障碍物，(B)机器人在左右两侧存在视觉杂波的干扰下驶向一个障碍物，(C)机器人在左右两侧存在视觉杂波的干扰下进行不减速的转向，随后再驶向一个障碍物。图中每个子图的第一行展示的均为这些场景中的关键帧。

首先，为了模拟汽车在现实行驶中可能遇到的不同复杂情况，我们操控移动机器人 TurtleBot 4 在实验室内行驶来收集各种动态复杂背景下的实验素材。其中，我们还在行驶的机器人的两侧放置较为密集的箱子，来模拟汽车行驶过程中道路两侧的树木、人流、车流等光流所引起的杂乱噪声，如图 7(B)(C) 实验所示。(B) 对照实验的结果在一方面显示出，在这种杂乱的视觉噪声的影响下，所提出的 LPNN 模型的输出膜电位曲线，仍然能够接近拟合出在不含有此类型噪声下模型输出的膜电位曲线；而与之相反的是，三个比较的模型均会受到这些噪声杂波很严重的影响，从而全过程中产生了很多假的碰撞警报。另一方面，这组实验的结果也揭示了，在抖动光流的影响下，相

比于其他模型在输出上有很大的波动，我们的模型并未受到明显的干扰，同时还能更为精确地提取出运动过程中存在的碰撞威胁。

其次，在上述实验所设计的两种特殊噪声杂波的基础上，我们进一步增加了动态环境的复杂性，如图 7(C) 实验所示，我们利用机器人模拟出汽车在行驶中的不减速转向过程。结果展现出，只有所提出的模型能够很有效地忽略掉这些不相关的刺激。同时，相比于 LGMD1 模型和 LGMD2 模型，Bi-LGMD 模型也能够较少地受到这类刺激的干扰。

随后，我们还使用了极具挑战性的室外动态场景数据集，进行了 4 组对比实验。值得一提的是，图 8(B)实

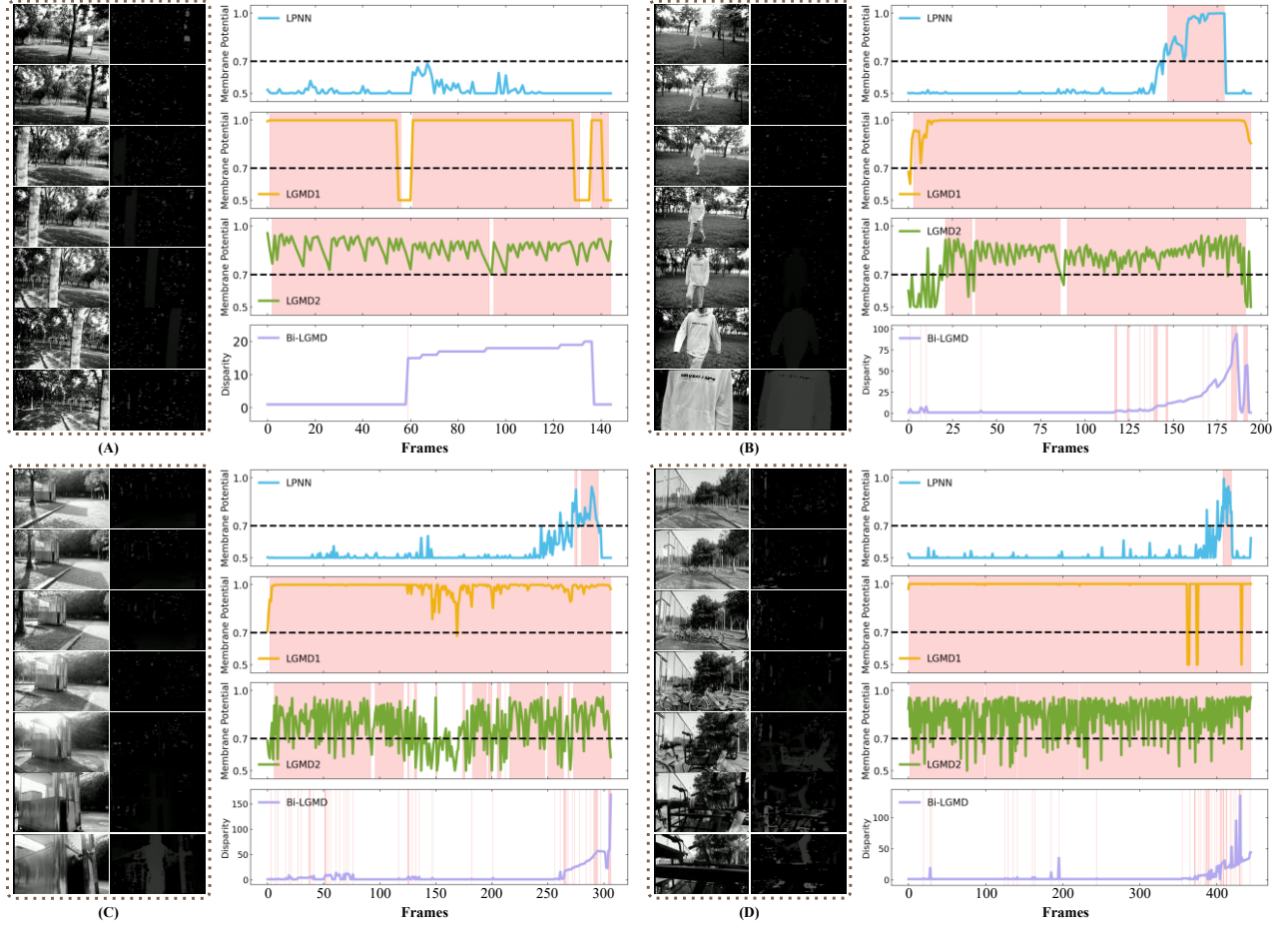


图 8 所提出的模型与另外三个对比模型在室外拍摄的动态场景中，对于复杂视觉刺激的反应结果，红色区域代表的是测试过程中模型产生碰撞警报的时间窗口。所展示的室外动态场景包括了：(A)在树林中进行缓慢旋转的非碰撞过程，(B)注释跟随过程中所发生的碰撞，以及(C)&(D)在树林中光影噪声的干扰下的碰撞场景。其中，每个子图的第一列均为模型所要处理的视频序列中的部分帧。

表 2 四个模型在 34 组真实数据集的测试下的有效性评估结果

Criteria	LPNN	LGMD1	LGMD2	Bi-LGMD
Accuracy in Static Scene	95.91%	66.06%	79.82%	90.17%
Accuracy in Dynamic Scene	96.29%	16.97%	36.23%	83.79%
Overall Accuracy	96.09%	42.96%	59.31%	87.16%
Overall Recall	65.93%	95.23%	90.03%	58.51%

验的测试视频序列，是我们在室外模拟果蝇的跟踪注视这一过程[19]，并且被注视的移动物体在最后发生了碰撞行为。图 8(A)(B)的实验结果展示了，与所比较的模型不同的是，我们的模型并不会受到在杂乱环境下的旋转过程的干扰，而是能够从这些复杂的噪声杂波中准确地提取出正在逼近的物体。而图 8(C)(D)实验则显示，现实中杂乱的光影噪声，都使得 LGMD1 模型和 LGMD2 模型的膜电位迅速上升且接近于饱和状态；而对于所提出的模型，由于只有当物体处于不断靠近的过程前景的所有运动边缘信号才

能被完整地保留，因此其能够只在探测到碰撞威胁的阶段才被迅速激活。

4.5 实验结果的统计分析

为了排除偶然因素的影响，以确保图 6、图 7、图 8 中所有实验结果的可靠性，本文总共挑选了 34 组具有代表性的室内外场景视频序列(包括 18 组静态场景刺激和 16 组动态场景刺激)来进行测试，根据公式 29 所计算的准确率以及公式 30 所计算的召回率，来综合比较所提出的模型与另外三个模型之间的性能。基于不同标准的评估结

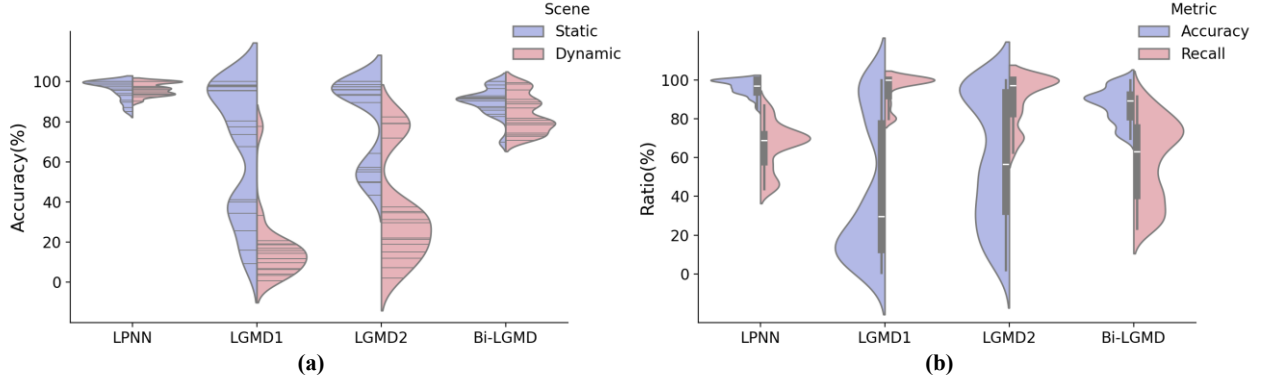


图 9 四个模型在 34 组真实数据集上的统计对比展示. 子图(a)展示的是四个模型分别在静态场景数据集和动态场景数据集上的碰撞感知准确率分布图, 子图(b)则可可视化了它们在所有测试数据集上的准确率和召回率分布图.

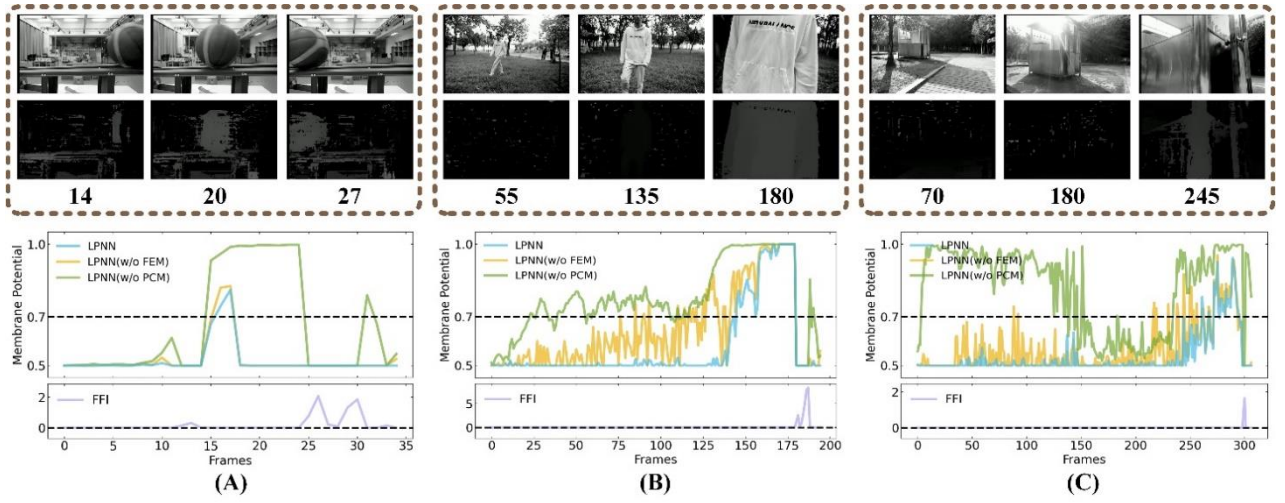


图 10 所提出的模型(LPNN)与没有 FEM 机制的 LPNN 模型(LPNN w/o FEM)、没有 PCM 机制的 LPNN 模型(LPNN w/o PCM), 在 3 组真实场景的视觉刺激下的反应结果对比. 每个子图的第一行展示的是对于场景中选取的关键帧, 随后展示的是这 3 个模型分别输出的膜电位曲线. 值得一提的是, 我们还可可视化了对应过程中 FFI 神经元的输出 (公式 23), 揭示其对感受野边缘区域突然产生的视觉兴奋具有明显的抑制作用. 与另外两个模型相比, 融合了两种机制的 LPNN 模型展现出了更高的稳健性.

果汇总如表 2 所示. 图 9 的具体统计结果一方面在整体上显示出了, 所提出的模型无论在何种场景下都拥有最高的准确率, 验证了其具备相对更高的稳健性. 另一方面, 我们的模型在不同类型场景下的测试结果显得更为稳定; 相比之下, 由于在大多数真实动态场景下的表现要远差于在静态场景下的表现, LGMD1 模型和 LGMD2 模型在性能上的不稳定性要远高于所提出的模型.

除此之外, 图 9(b)中 LPNN 模型相对较低的召回率也揭示了所提出的模型存在的缺陷, 即并不能完全识别出所有的碰撞样本, 这本质上也是由于深度相机难以提供近距离物体的有效视差信息. 例如在图 7(A)与图 8(B)的测试中, 当迫近目标处于即将发生碰撞的最后时间段内, 我们的模型整合了无效的视差信号, 从而输出的膜电位过早地回落到静息电位, 间接导致了对该时间段内碰撞样本的漏

检. 因此, 为了有效减轻深度传感器的这一约束对模型的影响, 未来我们考虑从“有条件整合”的角度来进一步优化, 即在视差计算通路失效的情况下, 不受影响的运动通路将作为感知迫近的唯一神经通路.

4.6 消融实验

在本节中, 我们通过消融实验, 进一步探究了前景提取机制(FEM)与划分-竞争机制(PCM)这两个重要机制的相互融合以及 FFI 机制的协助, 对我们的模型所产生的影响. 如图 10 中 3 组对比实验的结果所示, FEM 机制的缺失会导致在运动发生的初期, 模型明显受到大范围背景视觉杂波的影响, 从而其输出膜电位的波动起伏较大. 另一方面, PCM 机制则在我们的模型中扮演了更加重要的角色, 其能够很大程度地抑制前景中的很多无关运动所产生的视觉兴奋, 从而进一步提高模型对不断靠近物体的选择

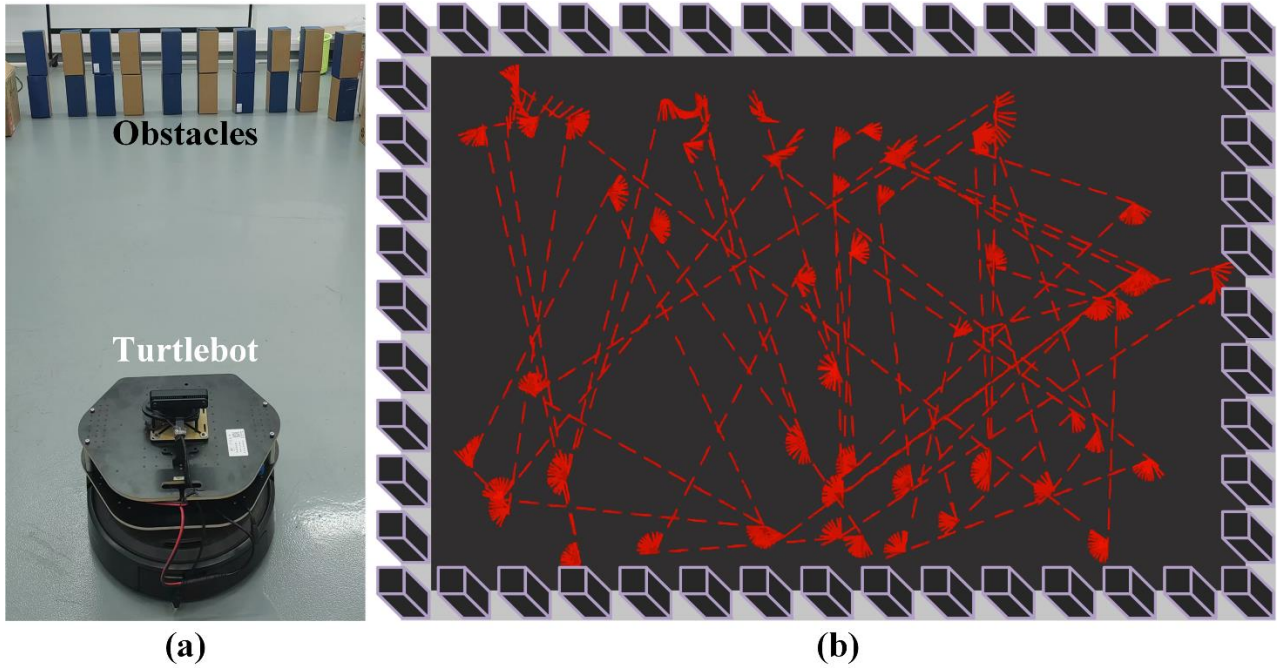


图 11 本实验展示所提出的 LPNN 模型在移动机器人的嵌入式应用以及闭环测试的表现结果. 左侧的子图(a)展示了机器人实际测试的真实场景, 包括了障碍物的设置方式; 子图(b)则是对机器人在 10 分钟的测试过程中实时轨迹的可视化: 其中红色虚线标记的是 TurtleBot 机器人实时的视觉朝向, 红色节点标记了当机器人检测到潜在碰撞威胁时的避碰转向过程.

性, 确保模型能够准确地捕捉到感受野内真正的碰撞威胁.

除此之外, 图 10(A)也揭示了, 所提出的融合运动通路和视差通路的 FFI 机制, 能够在运动物体突然进入或者离开感受野区域的过程中, 对可能产生的视觉兴奋进行有效的抑制, 进而降低了模型对感受野边缘区域突然产生的视觉刺激的关注度.

4.7 在线的机器人实验

在本节中, 我们将验证所提出的 LPNN 模型在移动机器人自主导航中进行实时地检测和避免碰撞的可行性与稳健性. 实验中所使用的 TurtleBot 机器人, 由于其搭载了深度相机、激光雷达等众多传感器, 因此在实际自主导航中已经得到了广泛的应用[21]. 然而, 在我们的实验中, 所提出的基于视觉的嵌入式碰撞感知系统, 将作为机器人进行实时碰撞感知的唯一方法, 而并不需要依赖于其他更加先进的传感器. 在闭环测试中, 我们设置机器人的初始状态为以 25cm/s 的恒定线速度进行直线行驶, 而一旦探测到碰撞威胁的存在, 机器人将先以 50 cm/s 的线速度在当前这一时刻做出后退反应, 随后再随机地选择向左或者向右执行 60°左右的转向指令, 从而避开了潜在的碰撞危险并且进入下一阶段的直线行驶和碰撞感知过程.

如图 11(b)中的实验结果所示, 红色曲线展示了在持续 10 分钟的闭环测试过程中机器人的行驶轨迹, 其揭示了每当实际物理环境中的障碍物成为了机器人在移动过程中的碰撞威胁时, 所提出的方法都能够协助机器人及时

地进行转向躲避的操作. 因此, 这也验证了融合运动通路和视差通路的视觉神经网络, 同时兼顾了视差通路的准确可靠性和运动通路的实时高效性, 能够为移动机器人在真实复杂环境中进行自主导航时, 提供一套实时稳健的碰撞感知嵌入式系统.

5 总结

综上所述, 本研究首次提出了一个实时整合来自运动通路和视差通路的视觉信号的碰撞感知神经网络模型, 为后续的建模研究提供一种全新的计算框架. 该模型取长补短, 分别克服了当前视差深度模型运算复杂度较高和难以区分不同运动模态的问题, 以及 LGMD 模型在动态复杂环境中易受视觉噪声影响的问题, 在保持较高的计算效率下, 进一步提高了对迫近运动的选择性. 具体而言, 通过视差信号对物体运动边缘信号的滤波与划分处理, 以及二者在前馈抑制神经元中的线性融合, 我们的模型充分结合了运动目标的扩张边缘特征以及动态的深度特征, 有效减轻了前后景中视觉噪声杂波的干扰, 并且降低了对感受野突然出现的目标的关注度, 以增强模型在复杂动态场景下进行碰撞感知的稳健性. 同时, 所提出的神经网络模型具备在硬件系统(如自主移动的机器人)上实际运行的可行性, 能够为移动机器人在绝大多数场景的自主导航中提供一套实时的、低能耗的、相对稳健的碰撞检测方法, 在机

机器人的自主导航领域中拥有广阔的应用前景。

通过系统性的实验,可以验证的是,一方面,与目前基于二维平面信息的 LGMD 方法相比,对于具备四个维度信息(即物理空间中的三维信息以及时间维度的信息)的 LPNN 模型而言,深度融合来自运动通路和视差通路的神经信号,使得我们的模型在不损失重要运动信息的前提下,既能排除背景视觉杂波的干扰,又能明显地抑制前景噪声信号的刺激,进而最大程度地提高了该模型在实际应用中的稳健性。另一方面,与基于深度图像进行碰撞检测的传统方法相比,由于我们的模型基于时空关联来协同运算,因此其能够根据障碍物移动的速度与运动模态,提供精确的“碰撞时间”信息和早期预警[40],以应对视野范围内突然出现的快速靠近物体[29]。与此同时,所提出的模型并不需要对真实环境进行场景重建,因此其无需维护三维空间中繁重的数据,这使得其在计算资源有限的情况下,仍然能够进行实时稳定的碰撞检测。

然而,由于双目立体视觉在近距离的匹配限制,我们难以获得物体在一定距离之内的真实视差信息,这将导致所提出的模型在视差计算通路失效的情况下,无法有效感知近距离的碰撞。鉴于此,未来模型的碰撞探测可以基于两个神经通路信号的“有条件融合”,即当运动目标进入到视差计算通路失效的区域内,系统仅通过运动视觉通路来感知逼近物体。除此之外,照明条件差可能会同时引起亮度信息和视差信息的严重失真,我们的模型在这种场景下仍会面临较大的挑战。因此,我们在接下来的研究中将进一步优化当前的模型,通过将该计算框架泛化到红外相机、事件相机等传感器,以探索减轻这种极端环境对模型产生影响的可能性。另外,未来的工作还将着眼于模型在移动机器人中的集群应用,以完成更为复杂的群体协作任务。

参考文献

- [1] LU L, CARRIO A, SAMPEDRO C, et al. A robust and fast collision-avoidance approach for micro aerial vehicles using a depth sensor[J]. Remote Sensing, 2021, 13(9): 1796.
- [2] CARRIO A, TORDESILLAS J, VEMPRALA S, et al. Onboard detection and localization of drones using depth maps[J]. IEEE Access, 2020, 8: 30480-30490.
- [3] LIN J, ZHU H, ALONSO-MORA J. Robust vision-based obstacle avoidance for micro aerial vehicles in dynamic environments[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 2682-2688.
- [4] WANG D, LI W, LIU X, et al. UAV environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution[J]. Computers and Electronics in Agriculture, 2020, 175: 105523.
- [5] SALEH S, MANOHARAN S, NINE J, et al. Towards robust perception depth information for collision avoidance[C]//2020 IEEE Congreso Biental de Argentina (ARGENCON). IEEE, 2020: 1-4.
- [6] FU Q, WANG H, HU C, et al. Towards computational models and applications of insect visual systems for motion perception: A review[J]. Artificial Life, 2019, 25(3): 263-311.
- [7] YUE S, RIND F C. Collision detection in complex dynamic scenes using an LGMD-based visual neural network with feature enhancement[J]. IEEE Transactions on Neural Networks, 2006, 17(3): 705-716.
- [8] GU B, FENG J, SONG Z. Looming detection in complex dynamic visual scenes by interneuronal coordination of motion and feature pathways[J]. Advanced Intelligent Systems, 2024, 6(9): 2400198.
- [9] KLAPOETKE N C, NERN A, PEEK M Y, et al. Ultra-selective looming detection from radial motion opponency[J]. Nature, 2017, 551(7679): 237-241.
- [10] FU Q, HU C, PENG J, et al. Shaping the collision selectivity in a looming sensitive neuron model with parallel on and off pathways and spike frequency adaptation[J]. Neural Networks, 2018, 106: 127-143.
- [11] FU Q, HU C, PENG J, et al. A robust collision perception visual neural network with specific selectivity to darker objects[J]. IEEE Transactions on Cybernetics, 2019, 50(12): 5074-5088.
- [12] SOBEL E C. The locust's use of motion parallax to measure distance[J]. J. Comp. Physiol. A, 1990, 167(5): 579-588.
- [13] KRAL K, POTESER M. Motion parallax as a source of distance information in locusts and mantids[J]. Journal of Insect Behavior, 1997, 10: 145-163.
- [14] FU Q. Motion perception based on ON/OFF channels: a survey[J]. Neural Networks, 2023, 165: 1-18.
- [15] VIENNE C, PLANTIER J, NEVEU P, et al. (Disparity-driven) Accommodation response contributes to perceived depth[J]. Frontiers in Neuroscience, 2018, 12: 973.
- [16] ZHENG Y, WANG Y, WU G, et al. Enhancing LGMD-based model for collision prediction via binocular structure[J]. Frontiers in Neuroscience, 2023, 17: 1247227.
- [17] RIND F C, BRAMWELL D I. Neural network based on the input organization of an identified neuron signaling impending collision[J]. Journal of Neurophysiology, 1996, 75(3): 967-985.
- [18] TIPPETTS B, LEE D J, LILLYWHITE K, et al. Review of stereo vision algorithms and their suitability for resource-limited systems[J]. Journal of Real-Time Image Processing, 2016, 11: 5-25.
- [19] BAHL A, AMMER G, SCHILLING T, et al. Object tracking in motion-blind flies[J]. Nature Neuroscience, 2013, 16(6): 730-738.
- [20] SIMMONS P J, RIND F C. Responses to object approach by a wide

- field visual neurone, the LGMD2 of the locust: characterization and image cues[J]. *Journal of Comparative Physiology A*, 1997, 180: 203-214.
- [21] PIETRZIK S, CHANDRASEKARAN B. Testing autonomous path planning algorithms and setup for robotic vehicle navigation[C]//2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2018: 485-488.
- [22] KEMPPAINEN J, SCALES B, RAZBAN HAGHIGHI K, et al. Binocular mirror-symmetric microsaccadic sampling enables *Drosophila* hyperacute 3D vision[J]. *Proceedings of the National Academy of Sciences*, 2022, 119(12): e2109717119.
- [23] LAUGHLIN S B. The role of sensory adaptation in the retina[J]. *Journal of Experimental Biology*, 1989, 146(1): 39-62.
- [24] O'SHEA M, WILLIAMS J L D. The anatomy and output connection of a locust visual interneurone; the lobular giant movement detector (LGMD) neurone[J]. *Journal of Comparative Physiology*, 1974, 91: 257-266.
- [25] SHARKAWY A N, ALI M M. NARX neural network for safe human-robot collaboration using only joint position sensor[J]. *Logistics*, 2022, 6(4): 75.
- [26] RAMASAMY S, SABATINI R, GARDI A, et al. LIDAR obstacle warning and avoidance system for unmanned aerial vehicle sense-and-avoid[J]. *Aerospace Science and Technology*, 2016, 55: 344-358.
- [27] WESSENDORP N, DINAUX R, DUPEYROUX J, et al. Obstacle Avoidance onboard MAVs using a FMCW RADAR[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 117-122.
- [28] FALANGA D, KLEBER K, SCARAMUZZA D. Dynamic obstacle avoidance for quadrotors with event cameras[J]. *Science Robotics*, 2020, 5(40): eaaz9712.
- [29] LU L, FASANO G, CARRIO A, et al. A comprehensive survey on non-cooperative collision avoidance for micro aerial vehicles: Sensing and Obstacle Detection[J]. *Journal of Field Robotics*, 2023, 40(6): 1697-1720.
- [30] WANG Y, JI J, WANG Q, et al. Autonomous flights in dynamic environments with onboard vision[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 1966-1973.
- [31] BAZAZI S, BUHL C, HALE J J, et al. Collective motion and cannibalism in locust migratory bands[J]. *Current Biology*, 2008, 18(10): 735-739.
- [32] KRONGAUZ D L, AYALI A, KAMINKA G A. Vision-based collective motion: a locust-inspired reductionist model[J]. *PLoS Computational Biology*, 2024, 20(1): e1011796.
- [33] YAKUBOWSKI J M, MCMILLAN G A, GRAY J R. Background visual motion affects responses of an insect motionsensitive neuron to objects deviating from a collision course[J]. *Physiological Reports*, 2016, 4(10): e12801.
- [34] RIND F C. Non-directional, movement sensitive neurones of the locust optic lobe[J]. *Journal of Comparative Physiology A*, 1987, 161: 477-494.
- [35] LI L, ZHANG Z, LU J. Artificial fly visual joint perception neural network inspired by multiple-regional collision detection[J]. *Neural Networks*, 2021, 135: 13-28.
- [36] GABBIANI F, KRAPP H G, KOCH C, et al. Multiplicative computation in a visual neuron sensitive to looming[J]. *Nature*, 2002, 420(6913): 320-324.
- [37] READ J C A. Stereopsis without correspondence[J]. *Philosophical Transactions of the Royal Society B*, 2023, 378(1869): 20210449.
- [38] KELLNHOFFER P, DIDYK P, RITSCHER T, et al. Motion parallax in stereo 3D: Model and applications[J]. *ACM Transactions on Graphics (TOG)*, 2016, 35(6): 1-12.
- [39] SUN X, JIANG Y, JI Y, et al. Distance measurement system based on binocular stereo vision[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2019, 252(5): 052051.
- [40] RIND F C. Recent advances in insect vision in a 3D world: looming stimuli and escape behaviour[J]. *Current Opinion in Insect Science*, 2024, 63: 101180.
- [41] WHEATSTONE C. On some remarkable and hitherto unobserved phenomena of binocular vision[J]. *The Optometric Weekly*, 1962, 53: 2311-2315.
- [42] NITYANANDA V, BISSIANNA G, TARAWNEH G, et al. Small or far away? Size and distance perception in the praying mantis[J]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2016, 371(1697): 20150262.
- [43] ROSNER R, VON HADELN J, TARAWNEH G, et al. A neuronal correlate of insect stereopsis[J]. *Nature Communications*, 2019, 10(1): 2845.
- [44] ROSNER R, TARAWNEH G, LUKYANOVA V, et al. Binocular responsiveness of projection neurons of the praying mantis optic lobe in the frontal visual field[J]. *Journal of Comparative Physiology A*, 2020, 206(2): 165-181.
- [45] ROSNER R, TARAWNEH G, LUKYANOVA V, et al. Binocular responsiveness of projection neurons of the praying mantis optic lobe in the frontal visual field[J]. *Journal of Comparative Physiology A*, 2020, 206(2): 165-181.
- [46] O'KEEFFE J, YAP S H, LLAMAS-CORNEJO I, et al. A computational model of stereoscopic prey capture in praying mantises[J]. *PLoS Computational Biology*, 2022, 18(5): e1009666.

- [47] SANTOS M C P, SANTANA L V, BRANDAO A S, et al. UAV obstacle avoidance using RGB-D system[C]//2015 international conference on unmanned aircraft systems (ICUAS). IEEE, 2015: 312-319.
- [48] HUA M, FU Q, PENG J, et al. Shaping the ultra-selectivity of a looming detection neural network from non-linear correlation of radial motion[C]//2022 International Joint Conference on Neural Networks (IJCNN). IEEE, 2022: 1-8.
- [49] FRANCESCHINI N. Small brains, smart machines: from fly vision to robot vision and back again[J]. *Proceedings of the IEEE*, 2014, 102(5): 751-781.
- [50] LEI F, PENG Z, LIU M, et al. A robust visual system for looming cue detection against translating motion[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 34(11): 8362-8376.
- [51] FU Q, YUE S. Modelling Drosophila motion vision pathways for decoding the direction of translating objects against cluttered moving backgrounds[J]. *Biological Cybernetics*, 2020, 114(4): 443-460.
- [52] FU Q, YUE S. Complementary visual neuronal systems model for collision sensing[C]//2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM). IEEE, 2020: 609-615.
- [53] RIND F C, Simmons P J. Seeing what is coming: building collision-sensitive neurones[J]. *Trends in Neurosciences*, 1999, 22(5): 215-220.
- [54] ZHAO J, WANG H, BELLOTTO N, et al. Enhancing LGMD's looming selectivity for UAV with spatial-temporal distributed presynaptic connections[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 34(5): 2539-2553.
- [55] ZHAO J, XIE Q, SHUANG F, et al. An Angular Acceleration Based Looming Detector for Moving UAVs[J]. *Biomimetics*, 2024, 9(1).
- [56] SALT L, HOWARD D, INDIVERI G, et al. Parameter optimization and learning in a spiking neural network for UAV obstacle avoidance targeting neuromorphic processors[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(9): 3305-3318.
- [57] PETRÁČEK P, WALTER V, BÁČA T, et al. Bio-inspired compact swarms of unmanned aerial vehicles without communication and external localization[J]. *Bioinspiration & Biomimetics*, 2020, 16(2): 026009.
- [58] ČÍŽEK P, FAIGL J. Self-supervised learning of the biologically-inspired obstacle avoidance of hexapod walking robot[J]. *Bioinspiration & Biomimetics*, 2019, 14(4): 046002.
- [59] SILVA A, SANTOS C. Computational model of the LGMD neuron for automatic collision detection[C]//2013 IEEE 3rd Portuguese Meeting in Bioengineering (ENBENG). IEEE, 2013: 1-4.
- [60] ZHAO J, XI S, LI Y, et al. A fly inspired solution to looming detection for collision avoidance[J]. *iScience*, 2023, 26(4).
- [61] 陈昌川, 丁锡烽, 王亦何, 孙霞, 乔飞. 基于光线碰撞检测的轻量化实时三维重建系统[J]. *电子学报*, 2024, 52(12): 4074-4086.
- CHEN Chang-chuan, DING Xi-feng, WANG Yi-he, SUN Xia, QIAO Fei. Lightweight Real-Time 3D Reconstruction System Based on Light Collision Detection[J]. *Acta Electronica Sinica*, 2024, 52(12): 4074-4086.
- [62] 周晓清, 王翔, 郑锦, 百晓. 基于自适应空间稀疏化的高效多视图立体匹配[J]. *电子学报*, 2023, 51(11): 3079-3091.
- ZHOU Xiao-qing, WANG Xiang, ZHENG Jin, BAI Xiao. Adaptive Spatial Sparsification for Efficient Multi-View Stereo Matching[J]. *Acta Electronica Sinica*, 2023, 51(11): 3079-3091.