

基于事件相机的机器人感知与控制综述

粟 僂¹ 杨 帆¹ 王向禹¹ 郭川东² 童良乐¹ 胡 权²

摘 要 事件相机作为一种新型动态视觉传感器, 通过各个像素点独立检测光照强度变化并异步输出“事件流”信号, 具有数据量小、延迟低、动态范围高等优秀特性, 给机器人控制带来新的可能. 本文主要介绍了近年来涌现的一系列事件相机与无人机、机械臂和人形机器人等机器人感知与运动控制结合的研究成果, 同时聚焦基于事件相机的控制新方法、新原理以及控制效果, 并指出基于事件相机的机器人控制的应用前景和发展趋势.

关键词 事件相机, 延迟低, 机器人控制, 感知与运动, 新型控制算法

引用格式 粟僂, 杨帆, 王向禹, 郭川东, 童良乐, 胡权. 基于事件相机的机器人感知与控制综述. 自动化学报, 2022, 48(8): 1869–1889

DOI 10.16383/j.aas.c210263

A Survey of Robot Perception and Control Based on Event Camera

SU Li¹ YANG Fan¹ WANG Xiang-Yu¹ GUO Chuan-Dong² TONG Liang-Le¹ HU Quan²

Abstract As a new type of dynamic vision sensor, the event camera detects the change of illumination intensity independently through each pixel and outputs the “event stream” signal asynchronously. It has excellent characteristics such as small amount of data, low latency and high dynamic range, which brings new possibilities for robot control. In this survey, a series of the latest research results on the combination of event camera and robot perception and motion control, including unmanned aerial vehicles, manipulators and humanoid robots, are introduced. At the same time, the new control methods, new principles and control effects based on event camera are focused, and the application prospect and development trend of robot control based on event camera are pointed out.

Key words Event camera, low latency, robot control, perception and motion, new control algorithm

Citation Su Li, Yang Fan, Wang Xiang-Yu, Guo Chuan-Dong, Tong Liang-Le, Hu Quan. A survey of robot perception and control based on event camera. *Acta Automatica Sinica*, 2022, 48(8): 1869–1889

近年来, 类似无人机、机械臂等机器人系统在各个领域 (如消防安防、植保农业、工厂制造等) 得到日益广泛的应用, 四足机器人、人形机器人等系统也成为机器人领域的研究热点; 可以预见机器人系统将在未来的智能制造、工业 4.0 革命中发挥愈发突出的作用. 上述机器人系统均为结构复杂、高度集成的机电系统, 系统的运动规划和控制常依赖对所在环境和目标的感知测量. 目前大部分的无人机、机械臂均采用传统的帧相机作为感知器件, 但是, 帧相机具有固有的高数据量、低时间分辨率、高

延迟等特点, 对快速运动的物体感知能力较弱, 极大地限制了机器人的操控能力.

在此背景下, 一种基于生物视觉成像原理的神经形态传感器——事件相机受到机器人领域学者的关注, 将其与机器人感知和控制结合, 涌现出一批突破传统帧相机限制的机器人控制成果. 本文从事件相机的基本原理开始, 依次综述事件相机与无人机、机械臂、人形机器人结合的最新成果, 并介绍在控制方法上, 结合事件相机特性的控制技术最新发展, 以期对这一快速发展的领域提供参考. 本文结构图如图 1 所示.

1 事件相机简介

1.1 事件相机的基本原理

事件相机 (Event camera) 是受生物视觉系统启发的一类新型神经形态视觉传感器, 与基于帧的视觉传感器在工作原理上存在着不同. 帧相机 (Frame camera) 的曝光时间是固定的, 即便某个像素上的光照没有发生变化, 它也会重复曝光. 相比

收稿日期 2021-03-31 录用日期 2021-11-26
Manuscript received March 31, 2021; accepted November 26, 2021

国家自然科学基金 (11872011), 国防科工局稳定支持科研项目基金 (HTKJ2020KL502013) 资助

Supported by National Natural Science Foundation of China (11872011) and Stable Support of Scientific Research Projects by the State Administration of Science, Technology and Industry for National Defense (HTKJ2020KL502013)

本文责任编辑 陈胜勇

Recommended by Associate Editor CHEN Sheng-Yong

1. 首都师范大学信息工程学院 北京 100048 2. 北京理工大学宇航学院 北京 100081

1. School of Information Engineering, Capital Normal University, Beijing 100048 2. School of Astronautics, Beijing Institute of Technology, Beijing 100081

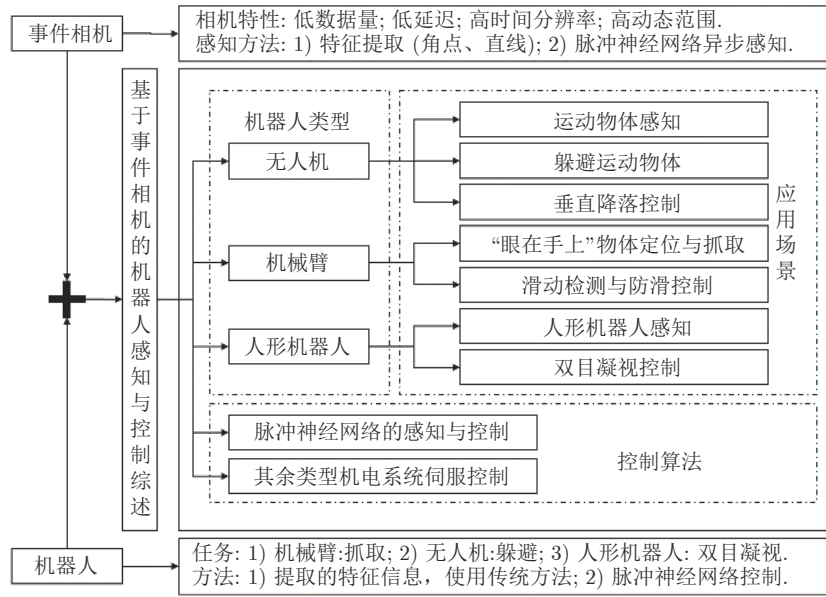


图 1 本文结构图

Fig.1 Structure diagram of this paper

之下, 事件相机的每个像素点独立的检测其上的亮度变化, 生成异步的事件流数据, 该数据包括时间戳、像素地址和亮度变化的极性.

事件相机像素点上的亮度变化要由目标对象或传感器运动引起^[1]. 当多个像素点同时请求事件输出时, 这些事件将以亚微秒级时延异步输出并构成事件流. 旋转圆盘场景下帧相机与事件相机输出对比如图 2 所示, 展示了传统帧相机与事件相机对高速转动圆盘上固定点进行一段时间跟踪拍摄得到的结果. 图 2(a) 显示传统帧相机以固定频率记录整个圆盘信息, 信息冗余度高, 帧间丢失了跟踪点的信息. 而图 2(b) 的事件相机则连续记录跟踪点对应的运动事件, 以异步的方式记录有用的运动信息, 具有低时延、低带宽需求的特点.

事件被触发以像素地址-事件流的形式输出, 地址-事件流数据包含事件的像素坐标、触发时间和极性 (亮度变化的信号) 三类信息. 触发的事件表示为

$$e = (x, y, t, \rho)^T \quad (1)$$

事件 e 表示在事件相机上位于 $(x, y)^T$ 处的像素点在 t 时刻因亮度变化而被触发的事件. $\rho(p, c)$ 为一个截断函数.

$$\rho(p, c) = \begin{cases} +1, & p \geq c \\ -1, & p \leq -c \end{cases} \quad (2)$$

c 为事件点的激发阈值, p 为亮度变化值. 当亮度增量大于 c 时, 激发正极性事件点, 当亮度增量小于 $-c$ 时, 激发负极性事件点, 当亮度增量的绝对值小于 c 时, 事件相机无输出.

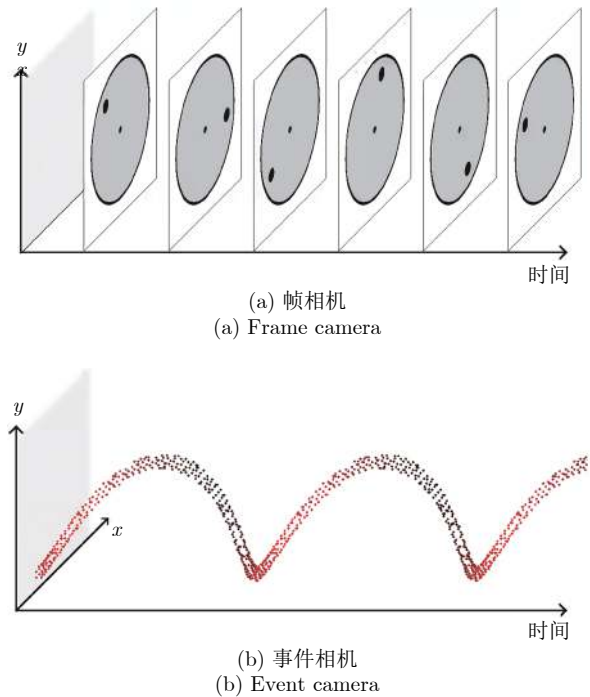


图 2 旋转圆盘场景下帧相机与事件相机输出对比^[2-3]. 帧相机在每一帧图像中记录全部像素点上的数据, 即使圆盘中大部分区域的信息是无用的; 事件相机仅记录圆盘中的黑点位置, 因此仅对运动的有效信息输出事件

Fig.2 Comparison of outputs between frame camera and event camera in rotating disc scene^[2-3]. The frame camera records the data on all pixels in each frame image, even if the information in most areas of the disc is useless;

The event camera only records the position of black spots in the disc, so it only outputs events for the effective information of motion

1.2 事件相机的发展历程

事件相机的出现, 可以追溯到 1992 年, Mahowald^[4] 在博士论文中提出了一种“硅视网膜”视觉传感器. 这是首个输出地址-事件流的视觉传感器, 但它只是事件相机的雏形, 像素面积过大导致它无法得到实际使用. Boahen^[5] 在 2002 年开发了编码增量视觉信号的视网膜启发芯片, 模拟生物视网膜的空间视觉通路 (Parallel pathway).

Kramer^[6] 在 2002 年提出了一种积分光瞬态传感器, 缺点是传感器的对比灵敏度较低. Ruedi 等^[7] 在 2003 年提出了空间对比度和局部定向视觉传感器, 其输出编码的是空间对比度而不是时间对比度. 在一个全局帧积分周期之后, 该设备按从高到低的空间对比顺序传输事件, 每个对比度事件后面都跟着一个编码梯度方向的事件. Luo 等^[8] 在 2006 年提出了一种基于首次时间脉冲 (Time-to-first-spike, TTFS) 技术的 CMOS (Complementary metal oxide semiconductor) 相机, 该相机利用每个像素的单个脉冲时序来编码每个像素的光度. 这种光度的时间表示可以将图像传感器的动态范围扩大到 100 dB 以上, 并引入了异步地址事件读取技术以降低功耗. Chi 等^[9] 在 2007 年提出了一种时间变化阈值检测相机, 它对传统源像素传感器 (Active pixel sensor, APS) 的像素进行了改进, 使其能够检测绝对光度变化, 该同步装置存储信号变化的像素地址, 构成一种同步地址事件表示 (Address event representation, AER) 的图像传感器. Lichtsteine 等^[10] 在 2008 年提出了第一台商用的事件相机, 称为基于异步事件的动态视觉传感器 (Dynamic vision sensor, DVS). 在 DVS 的基础上, 目前已经开发了几种具有附加功能的事件相机. 如 DAVIS、ATIS 等, 详见第 1.4 节.

1.3 事件相机的主要特点

事件相机异步地测量每个像素的亮度变化, 而非以固定速率捕获图像帧, 因此能克服传统帧相机的局限性并且具有特有的优异属性, 如高时间分辨率、低延迟、低功耗、高动态范围等, 在高速和高动态范围场景中有着广阔的应用空间.

1) 高时间分辨率. 事件相机是通过检测像素点处的亮度变化产生事件, 其对亮度变化产生的响应速度极快, 输出频率可达 1 MHz, 即事件时间戳分辨率为微秒级. 因此, 事件相机可以捕捉到高速运动, 不会受到运动模糊的影响.

2) 低延迟. 事件相机可近乎实时地输出亮度变化, 平均延迟在 μs 数量级, 而传统相机受帧率限制,

采样等待时间是 0 和 $1/f$ 之间的均匀分布, 其中 f 是帧速率. 因此, 事件相机对输入变化具有超低延迟, 实验测试中时延约为 $10 \mu\text{s}$, 在真实应用中可达亚毫秒级^[11].

3) 低数据量和低功耗. 事件相机以稀疏事件流的形式输出视觉信号, 事件流中不包含静态背景的任何信息, 因此过滤了大量冗余数据. 因为仅传输像素点的亮度变化, 不需要用于像素读取的模数转换器, 从而避免了大量冗余数据的传输, 能耗仅用于处理变化的像素. 大多数事件相机功耗约在 10 mW 级, 甚至有些相机原型的功耗低于 $10 \mu\text{W}$. 基于事件相机构建的嵌入式系统, 其系统整体的功耗一般大约为 100 mW, 甚至更低^[12-14].

4) 高动态范围. 事件相机不受白平衡、感光度等统一成像参数影响, 在图像过暗、曝光过度、光线突变等情况下, 依然可以通过触发事件来获取视觉信息. 因此, 具有极高的动态范围 ($>120 \text{ dB}$), 甚至可高达 143 dB ^[15]. 而帧相机通常只能达到 60 dB, 所以, 事件相机在光照条件良好的白天或者光线较暗的夜晚均能够有效地工作.

上述优点与各类机器人控制结合, 能够突破传统帧相机的限制, 实现机器人系统对快速运动目标的感知和响应、大大降低感知系统的功耗、提高在复杂光照 (强光、暗光) 条件中的感知能力. 尽管如此, 目前在应用上也存在一些挑战, 主要包括硬件缺陷造成的挑战以及处理范式转变带来的挑战^[11]:

1) 噪声. 光子中存在固有的散粒噪声以及晶体管电路中存在的噪声, 使得视觉传感器本身就携带大量的噪声事件, 当物体或相机运动时也会产生一系列噪声事件, 并且它们是非理想性的. 因此如何从噪声中提取有效事件信息是研究者所要解决的问题.

2) 信息的处理. 与帧相机提供的灰度信息相比, 事件相机产生的每个事件仅提供二进制亮度变化信息 (亮度增加或减少, 以事件的极性表示). 并且亮度变化不仅取决于场景亮度, 还取决于场景和相机之间的过去和当前的运动变化, 处理这种二进制信息具有挑战性.

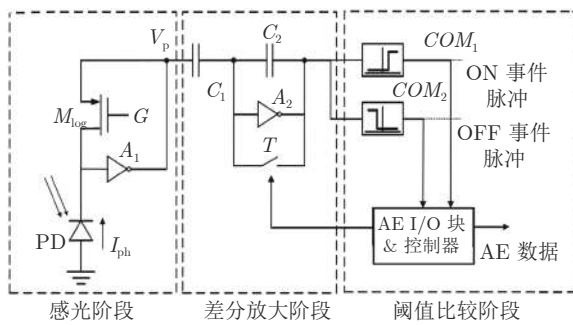
3) 新的算法. 事件相机的输出与帧相机有着根本的不同, 事件是异步的、空间稀疏的, 而图像是同步的、密集的. 因此, 现有的大多数计算机视觉算法在事件相机上无法直接使用, 因此开发特有的新算法是充分应用事件相机的前提.

1.4 事件相机的代表性产品

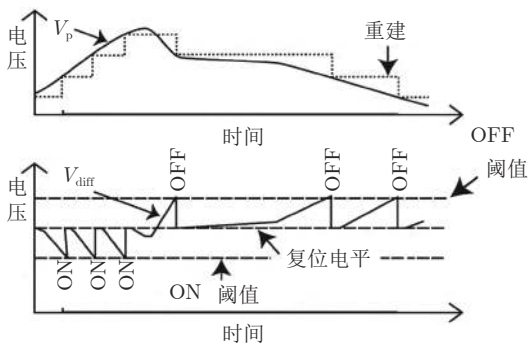
1.4.1 DVS

DVS 为事件相机的一种, 是众多类视觉传感器的基础. 它是一种基于时间的异步动态视觉传感器,

模拟生物视网膜空间视觉通路的功能, 试图感知场景的动态信息. DVS 像素由快速对数光感受器 (Photoreceptor)、差分电路 (Differencing) 和两个比较器 (Comparators) 组成, 以微秒级时间分辨率响应亮度相对变化. 它的单个像素结构和工作原理简化图^[10, 16] 分别如图 3(a) 和图 3(b) 所示. 感光电路主要由一个光电二极管、一个串联的晶体管和一个负反馈回路构成, 它能感知光线强度变化并及时给出响应. 差分电路可以将感光电路的输出进行放大, 并减少晶体管因工艺制作带来的误差. 阈值比较电路主要是通过比较器比较电压的变化实现 ON/OFF 事件脉冲的输出. 若光线变亮, 电压变化量大于 ON 事件的阈值, 则会输出一个脉冲表示 ON 事件; 若光线变暗, 电压变化量小于 OFF 事件的阈值, 则会输出一个脉冲表示 OFF 事件. 如果电压达到饱和值, 复位信号会将电压进行复位, 不产生 ON 或 OFF 信号. 2020 年 Samsung 开发的最新款 DVS 具有 1280×960 的分辨率, 每像素功耗低至 122 nW 与 $4.95 \text{ }\mu\text{m}$ 的像素间距^[17].



(a) 像素电路结构图^[16]
(a) Circuit diagram of pixel^[16]



(b) 工作原理^[10]
(b) Working principle^[10]

图 3 DVS 原理图

Fig. 3 Schematic diagram of DVS

1.4.2 ATIS

Posch 等^[15, 18] 在 2011 年提出基于异步时间的图像传感器 (Asynchronous time-based image

sensor, ATIS). ATIS 在 DVS 上进行改进, 加入了脉冲宽度调制环节, 在其受到光强变化产生事件信号时激发另一子像素进行曝光成像, 可以得到亮度变化超过阈值部分像素的强度图像, 也就是可以在输出光线强度变化信息的同时输出对应点的光线强度信息. Chronocam 供应的系列最新款 ATIS 视觉传感器拥有 304×240 的分辨率, 最高达 143 dB 的动态范围与 $30 \text{ }\mu\text{m} \times 30 \text{ }\mu\text{m}$ 的像素面积.

1.4.3 DAVIS

为了解决 DVS 难以直接应用现有算法的问题, Brandli 等^[19] 和 Taverni 等^[20] 开发了动态和有源像素视觉传感器 (Dynamic and active pixel vision sensor, DAVIS). DAVIS 是 DVS 相机和传统的 APS 相机的结合体, 将 APS 电路和 DVS 电路共用一个光电二极管, 实现异步事件和同步帧级图像的同时输出. DAVIS 像素也分为两个子结构, 其中一个子结构用于监测光照变化的事件, 而另一个子结构像传统的 APS 相机一样进行同步的曝光. 不同于 ATIS 的是, DAVIS 像素的两个子结构是共用同一个感光器的, 而 ATIS 像素的两个子像素有各自的感光器. 因此, DAVIS 相机的像素面积较 ATIS 相机的像素面积更小, 前者的填充因数较后者的更大. 通过 DAVIS 获取的灰度图和通过传统 APS 相机获取的灰度图一样, 具有时间冗余性和空间冗余性, 并且无法承受高动态范围的环境. 由 iniVation 公司供应的系列最新款 DAVIS 346 视觉传感器拥有 346×260 的分辨率, 最高达 120 dB 的动态范围与 $18.5 \text{ }\mu\text{m} \times 18.5 \text{ }\mu\text{m}$ 的像素面积.

1.4.4 CeleX

2019 年上海芯仑科技发布 CeleX-V 事件相机^[21], 具有 1280×800 的分辨率, 同时时域最大输出采样频率为 160 MHz , 动态范围为 120 dB , 是第一款达到百万像素级别的相机, 该传感器采用 65 nm CIS (CMOS image sensor) 工艺实现, 像素尺寸为 $9.8 \text{ }\mu\text{m} \times 9.8 \text{ }\mu\text{m}$, 具有高空间分辨率、高时域分辨率和高动态范围的优势, 引起了当前事件相机领域的关注. CeleX-V 将多个视觉功能集成到一个图像传感器中, 实现全阵列并行运动检测和全帧图像输出, 它还可以实现片上光流提取.

表 1 总结了目前主流事件相机的性能.

2 基于事件相机的无人机感知和运动控制

无人机 (Unmanned aerial vehicle, UAV) 具有体积小、造价低、使用方便、对环境要求低、生存能力较强等优点, 广泛应用于航拍、植保、测绘、

表 1 几种事件相机的性能比较表
Table 1 Performance comparison table of several event cameras

| 文献 | 事件相机型号 | 产品年份 | 分辨率 (pixel) | 最高动态范围 (dB) | 功耗 (mW) | 延迟 (μs) | 像素大小 (μm^2) |
|------|----------|------|-------------|-------------|-----------|----------------------|--------------------------|
| [10] | DVS128 | 2008 | 128×128 | 120 | 132 ~ 231 | 12 | 40 × 40 |
| [17] | DVS1280 | 2020 | 1280×960 | 120 | 150 | — | 4.95 × 4.95 |
| [15] | ATIS | 2011 | 304×240 | 143 | 50 ~ 175 | 3 | 30 × 30 |
| [19] | DAVIS240 | 2014 | 240×180 | 130 | 5 ~ 14 | 12 | 18.5 × 18.5 |
| [20] | DAVIS346 | 2018 | 346×260 | 120 | 10 ~ 170 | 20 | 18.5 × 18.5 |
| [21] | CeleX-V | 2019 | 1280×800 | 120 | 390 ~ 470 | 8 | 9.8 × 9.8 |

巡检等领域. 为了进一步提升无人机对环境的感知能力, 已有众多学者结合事件相机的低数据量、低延迟、高动态范围等特性, 对无人机运动控制问题进行了研究, 实现了在基于传统帧相机为感知器件时, 较难实现的快速感知、避障和垂直降落动作.

在无人机的实际应用中, 会受到如鸟类、敌方目标的干扰, 因此需要在较短的时间内进行障碍物的躲避. 无人机受到重量、功耗等限制, 对数据处理、计算复杂性、功耗、延迟有着极大的限制, 因事件相机低功耗、低数据量、低延迟的特性, 很适合装载于无人机上, 完成相应任务. 从理论上来说, 传统相机对环境的感知延迟约 20 ~ 40 ms, 事件相机的感知延迟仅为 2.2 ± 2 ms, 能够为无人机预留更长的机动时间. 下面分别从基于事件相机的无人机感知、避障、垂直降落等方面, 介绍该方向的工作.

2.1 无人机基于事件相机感知运动物体

因为无人机较快的运动速度, 对数据处理实时性提出了较高的要求, 为了研究事件相机的极限跟踪效果, Chamorro 等^[22]提出了一种超快速跟踪的方法. 使用如图 4 所示的实验环境, 事件相机静止放置于四杆装置前 20 cm, 由直流电机驱动四杆装置进行快速晃动, 该装置上搭载画有几何图形的平面, 用以产生事件, 并通过事件相机进行观测, 用以求解物体的位姿.



(a) 跟踪器的一次视觉快照 (b) 四杆运动装置
(a) A visual snapshot of tracker (b) Four-bar motion device

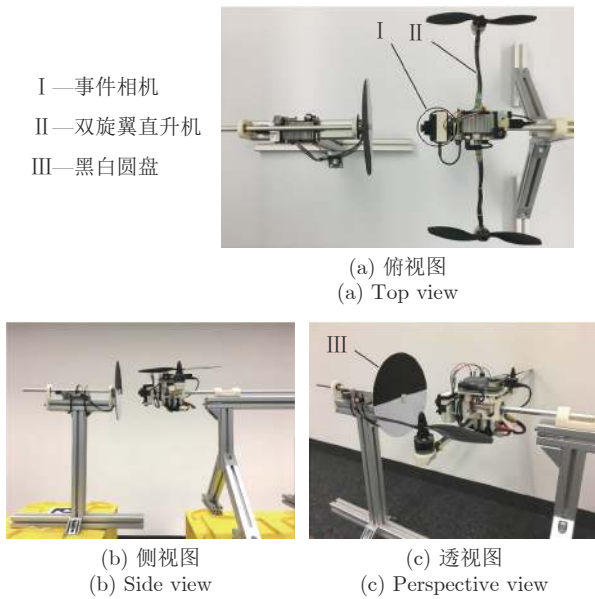
图 4 实验环境^[22]

Fig. 4 Experimental environment^[22]

图 4(a) 模拟了事件相机成像平面. 这项实验具有 Lie 参数化功能^[23], 提出了一种新的 Lie-EKF 公式用于跟踪物体的 6 自由度状态. 为了测试该方法的追踪极限, 直流电机的速度逐渐增加, 直到大约 950 r/min (15.8 Hz), 跟踪性能开始下降. 在这样的角速度下, 装置的速度分析报告的最大目标速度为 2.59 m/s. 线性目标加速度达到 253.23 m/s^2 或 25.81 g , 远远高于最苛刻的机器人应用的预期范围. 根据所使用的运动模型和状态参数, 跟踪器能够处理 89.1% 到 97.7% 的输入数据, 达到实时性能, 并以 10 kHz 的速率产生姿态更新, 并且即使在亮度骤然变化时 (开、关灯), 仍然能保证较好的性能.

Mueggler 等^[2]提出了带有事件相机的四旋翼高速机动的全 3D 轨迹跟踪, 通过搭载在无人机上的事件相机记录事件流, 事件数据由无线网络传输到笔记本电脑上实时可视化, 同时记录了前置标准 CMOS 相机的视频. 实验过程成功演示了在四旋翼翻转期间转速高达 $1200^\circ/\text{s}$ 时的位姿跟踪. 但是该方法适用环境有限, 也不能用于闭环控制.

Dimitrova 等^[24]在装有事件相机的无人机平台上解决了一维姿态跟踪的问题, 并实现了闭环控制. 在双旋翼直升机上搭载事件相机作为感知设备. 实验装置如图 5 所示, 其中黑白圆盘为参考图, 位于事件相机前方 10 cm 的位置, 黑白两部分之间的分界线为要跟踪的参考基准线, 在该实验中, 将卡尔曼滤波器与改进的 Hough 变换算法相结合, 构成基于事件的状态估计器, 计算双旋翼直升机相对于基准线的滚转角和角速度. 并通过旋转编码器提供转盘和无人机的真实测量角度, 用于评估基于事件的状态估计器和控制器的性能. 实验中双旋翼无人机已最小化地面和周围结构的空气动力学影响, 状态估计器以 1 kHz 的更新速率和 12 ms 的延迟执行, 能够以 $1600^\circ/\text{s}$ 的速度跟踪基准线, 并同时控制无人机以 $1600^\circ/\text{s}$ 的高速跟踪基准线. 该方法呈现了良好的事件相机驱动闭环控制的结果. 但是实验中的状态估计步骤繁琐, 适用范围也不广. 该文作者后续也将继续研究如何使用事件直接编码控制信

图 5 实验装置^[24]Fig. 5 Experimental platform^[24]

号, 完全消除状态估计步骤.

2.2 无人机躲避高速运动物体

文献 [25] 是第一个使用事件相机进行躲避高速运动物体的工作. 基于无人机携带的双目事件相机, 设计了一种球形物体追踪器, 使用扩展卡尔曼滤波进行数据处理, 可以在 250 ms 内预测碰撞, 并可以躲避 6 m 远处以 10 m/s 的速度飞来的球. 但因实验平台的限制, 仅从理论上验证了在无人机这样的小型嵌入式平台上, 检测并躲避障碍物的可行性.

为了进一步探究事件相机的优势, Falanga 等^[26]对搭载单目、双目以及事件相机 3 种传感器的无人机进行了实验. 实验结果表明了在 2 m 的感知范围下, 事件相机的延迟 (2 ~ 4 ms) (指从视觉传感器被触发到将数据传输到处理模块的时间) 显著低于普通单目相机 (26 ~ 40 ms) 和双目相机 (17 ~ 70 ms). 对于已知大小的障碍物, 使用聚类 and 光流算法对事件流数据进行处理来跟踪物体, 仅考虑水平移动, 搭载事件相机的无人机可在 2 m 的感知范围内躲避相对速度高达 9 m/s 的障碍物. 实验平台如图 6(a) 所示.

文献 [27] 使用了浅层神经网络从事件流中分割出独立运动对象, 并对其三维运动进行推理, 以执行闪避任务, 闪避场景如图 6(b) 所示, 四旋翼无人机使用 EVDodgeNet 进行障碍物躲避. 然后将其扩展到追踪任务, 这是第一个基于深度学习的解决方案, 网络如图 6(c) 所示, 由 3 个网络组成, 去模糊网络 EVDeblurNet, 用于将 δt 内产生的事件对

齐; 单源性估计网络 EVHomographyNet, 用于计算相机自我运动; 以及用以分割运动物体并进行光流追踪的 EVSegFlowNet, 共包含 130 万个参数, 相比于文献 [26] 实现了飞行器的三维避障. 并且对于没有先验信息的障碍物, 仍有 76% 的躲避成功率, 无需重新训练即可泛化到实际场景中. 但最终的响应延迟长达 60 ms, 极大地限制了无人机的运动.

在文献 [26–27] 的理论基础上, 文献 [28] 在相同的实验环境下, 仅使用事件相机进行数据捕捉, 该方法使用精度换时间的策略, 从收到待处理事件到发出第一条避障指令, 时延仅为 3.5 ms, 且成功率超过 90%. 在感知环节, 为降低计算量, 对文献 [29] 中提出的方法进行改进, 仅使用 IMU (Inertial measurement unit) 进行自我运动补偿, 将事件数据区分为静态场景与动态对象. 再通过聚类区分不同物体并结合双目深度信息计算三维坐标, 并将该位置作为卡尔曼滤波的测量输入计算其速度信息. 而在避障环节, 为降低计算量, 使用了基于反应式的人工势场法回避方案^[30], 避免了由于数值优化而引入延迟. 同时为方便处理, 利用二维平面中的数据将障碍物构造成椭圆体建立排斥场, 目标运动位置构造引力场, 最终向控制器输出运动速度. 算法的总延迟从文献 [27] 中的 60 ms 提升到了 3.5 ms, 且能够在板载处理器上实时运行, 极大地提升了响应速度. 在室内及室外的实验中, 均可躲避相对速度 10 m/s 的障碍物, 可以连续躲避多个障碍物, 并且躲避成功率超过 90%.

而在基于传统相机的避障工作中, Lin 等^[31]的单目方案延迟可优化到 150 ms, 可以在 2.2 m/s 的移动速度下躲避障碍物. Oleynikova 等^[32]的双目方案, 从图像曝光到路点规划的整体时延为 14.1 ms (曝光时间 3 ms), 能够在 5 m/s 的飞行速度下, 对 0.07 m 远处的障碍物进行响应. Barry 等^[33]在无人机上搭载双目 120 Hz 相机, 能在 14 m/s 的速度下实现避障. Huang 等^[34]使用 30 Hz 的 RGB-D 相机在机载芯片上数据处理时间为 25 ms. 因为传统图像数据量较大, 目前很多方案会将数据传回地面站进行计算, 更进一步加大了延迟, 可见因事件相机的独特优势, 已经逐渐展现出了应用潜力.

当前的这些工作中, 导致躲避失败的主要因素是检测到障碍物的时间太晚, 导致无人机无法及时地完成规避动作, 主要原因是: 1) 当障碍物进入到相机视野时, 距离已经不足以使飞行器完成规避动作; 2) 障碍物的运动没有产生足够的事件 (如物体垂直于像平面运动, 仅物体边缘产生少量事件), 算法无法及时检测到障碍物.

文献 [26] 的研究中也发现事件相机的延迟并

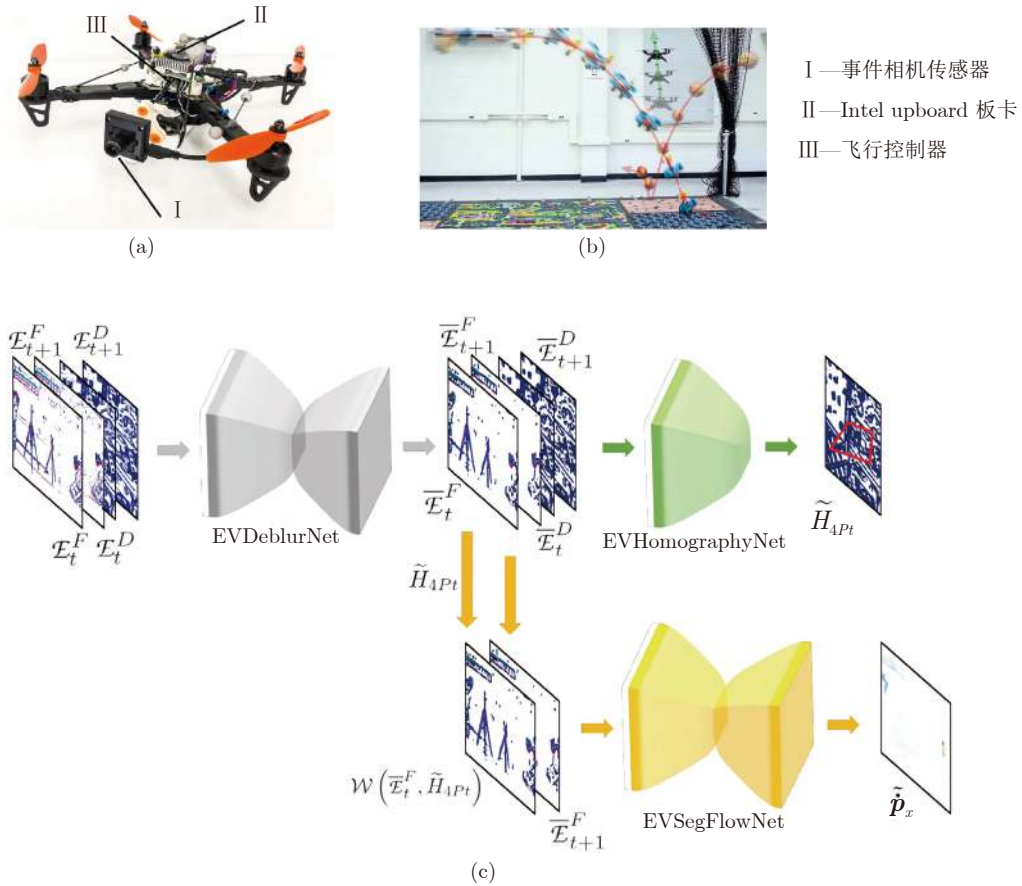


图 6 无人机躲避高速运动物体的相关实验 ((a) 无人机实验平台^[26]; (b) 真实无人机躲避两个飞行障碍物, 四旋翼无人机使用 EVDodgeNet 进行障碍物躲避, 其中无人机左右两侧的弧线轨迹为同时掷向无人机的两个障碍物, 中间的轨迹为无人机闪躲轨迹^[27]; (c) 神经网络架构图^[27], 其中 EVDeblurNet: 事件帧去噪网络, 在模拟数据集上进行训练就可推广到真实场景, 无需重新训练或者微调; EVHomographyNet: 单应性估计网络, 用于自我运动估计, 计算相机的自我运动. 第一个使用事件相机的单应性估计方案; EVSegFlowNet: 物体分割与光流计算网络, 该网络可分割场景中的运动物体并获取其光流信息)

Fig.6 Experiments of UAV avoiding high-speed moving obstacles ((a) UAV experimental platform^[26]; (b) The real UAV avoids two flying obstacles, and the quadrotor UAV uses EVDodgeNet to avoid obstacles, in which the arc trajectory on left and right sides of the UAV is two obstacles thrown at the same time, and the middle trajectory is the UAV dodge trajectory^[27]; (c) Neural network architecture diagram^[27], in which EVDeblurNet: event frame denoising network, which can be extended to real scenes by training on simulated datasets without retraining or fine tuning; EVHomographyNet: homography estimation network, which is used to estimate self-motion and calculate self-motion of camera. The first homography estimation scheme using event camera; EVSegFlowNet: an object segmentation and optical flow computing network, which can segment moving objects in the scene and obtain their optical flow information)

不是恒定的, 而是取决于相机与障碍物之间的距离与速度, 障碍物速度越大、距离越近, 响应时间越短. 因此该文献认为, 在当前, 双目相机在性能、成本上能够有很好的折衷, 仍然是无人机的最佳选择之一, 而随着技术的发展, 事件相机能够在未来提供更好的解决方案.

2.3 无人机垂直降落控制

为了模仿鸟类着陆, 现有的无人机垂直着陆解决方案均在无人机底部安装事件相机, 如图 7 所示, 以观察地面信息.

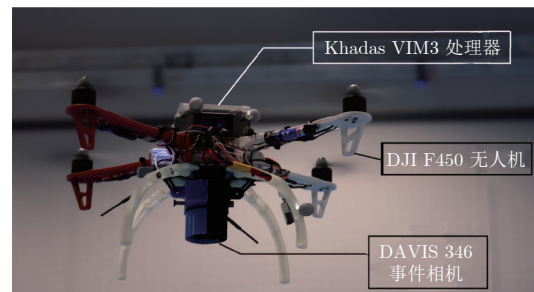


图 7 无人机垂直降落实验装置^[35]
Fig.7 Experimental device for vertical landing of UAV^[35]

文献 [35] 使用如图 7 所示的实验装置. 同时在事件流和基于事件帧的图像上进行直线追踪, 使用扩展卡尔曼滤波将二者结合, 利用前者的快速响应和后者的鲁棒性, 可以高速、有效地跟踪直线, 进而在时空中生成直线可能所在的平面, 整体频率高达 339 Hz. 再利用 Tau 理论^[36] 进行位置调整可以有效地引导无人机, 为着陆、悬停提供空间信息.

Pijnacker Hordijk 等^[37] 最先将基于事件的光流集成到无人机应用中, 提出了一种高效的平面几何光流估计技术, 尽管光流本身并不能提供运动的度量标度, 但是来自光流场的信息对于一些导航任务(包括着陆)十分有用. 实现了无人机的快速平稳着陆.

在文献 [37] 的理论基础上, 文献 [38] 使用神经进化 (Neuro-evolution) 优化人工神经网络, 设计了神经控制器, 学习现实世界中基于事件的光流控制. 对于使用事件相机与普通相机同时采集的数据, 该控制器输出了相近的控制策略, 证明了该控制器的鲁棒性.

文献 [39] 将文献 [38] 中的方法扩展到脉冲神经网络, 脉冲神经网络 (Spiking neural network, SNN) 是一种生物神经网络, 也是第三代神经网络模型^[40]. 它采用脉冲神经元为基本单位, 脉冲神经元通过接收脉冲改变内部状态, 当状态超过阈值才会输出脉冲, 当未接收脉冲时脉冲神经元不工作, 因此 SNN 可以在极低的功耗下工作. 使用 SNN 可以更加充分地利用事件流信息, 以获得更有效的方案. 飞行昆虫能够在混乱的环境中快速飞行, 可以灵活地避开障碍物. 而自主微型飞行器 (Micro aer-

ial vehicle, MAV) 远远落后于生物飞行器, 消耗的能量非常多. SNN 不仅可以用来模拟人脑的神经网络, 文献 [39] 提出该网络还可以利用镜头向下的事件相机产生的光流散度控制 MAV 的飞行和着陆. 这项工作是第一次将 SNN 集成到真实飞行机器人的控制回路中. 不仅研究了如何大幅降低 SNN 控制器的脉冲率, 而且该方案可以节省大量的能量. 但是该控制器只能实现在常规芯片上, 不能在神经形态芯片上实现.

以上基于事件相机的无人机运动控制研究与结果描述如表 2 所示.

当前研究的主要目标为实现闭环控制以及实现低延迟高带宽的飞行控制. 但是当前研究中大部分仍然是基于传统的视觉算法, 并未充分利用事件相机异步更新的特性.

3 基于事件相机的机械臂感知和运动控制

机械臂是由多个驱动关节串联组成的多自由度机电系统, 其通过安装于末端的执行器完成各类操作. 但是机械臂的运动规划和控制均依赖其对周围环境的感知, 确定被操作对象的位姿、状态、是否有障碍等等. 现有的感知手段包括视觉、触觉、力觉等. 其中基于视觉的感知具有低成本、易用性强等特点, 已成为机械臂感知的主流方法. 目前已有学者开始将事件相机作为机械臂的感知器件, 完成基于事件数据的静态物体抓取以及被夹持物体的滑动检测. 与传统的基于图像的方法相比, 事件相机在

表 2 无人机部分实验场景
Table 2 Experimental scenarios of UAV

| 文献 | 无人机型号 | 事件相机 | 处理器或控制器 | 事件流处理方法 | 备注 |
|-------|-------------------------------|--------------------------|---|------------------------|--|
| 基于事件帧 | | | | | |
| [2] | Parrot AR.Drone 2.0 | DVS | Odroid U2 onboard computer | Hough 变换 | 跟踪速度高达 1200°/s |
| [22] | — | DVS | Standard PC | 卡尔曼滤波, Lie-EKF 公式 | 跟踪速度超过 2.5 m/s, 加速度超过 25.8 g |
| [24] | 组装 | DAVIS 240C | Intel Upboard Lumenier F4 AIO flight controller | Hough 变换, 卡尔曼滤波 | 跟踪速度高达 1600°/s |
| [27] | Intel Aero Ready to Fly Drone | DAVIS 240C DAVIS 240B | NVIDIA TX2 CPU+GPU | 浅层神经网络 | 碰撞预测时间 60 ms |
| [37] | 定制 MavTec | DVS128 | Lisa/MX Odroid XU4 | 聚类, 光流算法 | — |
| 基于事件流 | | | | | |
| [25] | 组装 | 双目 DVS 128 | Odroid U3 quad-core computer | 高斯核函数, 卡尔曼滤波 | 碰撞预测时间 250 ms |
| [26] | 基于 DJI F330 定制 | Insightness SEEM1 | Intel Upboard Lumenier F4 AIO flight controller | DBSCAN 聚类, 光流算法 | 碰撞预测时间 68 ms (2 m 处) 183 ms (3 m 处) |
| [28] | 定制 | Insightness SEEM1 | Qualcomm Snapdragon Flight NVIDIA Jetson TX2 | DBSCAN 聚类, 光流算法, 卡尔曼滤波 | 碰撞预测时间 3.5 ms |
| [35] | 基于 DJI F450 | DAVIS 346 | PixRacer Kadas VIM3 | Hough 变换, 卡尔曼滤波 | 规划时间 2.95 ms |
| [39] | Parrot Bebop2 | — | Dual-core ARM cortex A9 processor. | 脉冲神经网络, 光流算法 | 神经形态控制器效果良好 |

此场景中提供了更高的灵敏度、更低的延迟以及更少的功耗。

3.1 “眼在手上”的物体定位感知和抓取

抓取物体是机械臂的基本能力。眼在手上是指视觉传感设备安装于机械臂的末端, 伴随机械臂运动。Muthusamy 等^[41]首先在该问题上做出了尝试, 提出一种眼在手上的基于事件相机的机械臂抓取方案。该方案初始化一个随机目标点, 控制机械臂使相机向该点运动以产生事件, 利用基于事件的视觉伺服 (Event-based visual servoing, EBVS), 控制机械臂使目标中心与相机中心重合, 最后调整抓取角度以实现稳定抓取。其中, EBVS 采用活跃事件表面 (Surfaces of active events, SAE) 异步执行角点检测。得到的角点储存在活跃角点事件表面 (Surfaces of active corner events, SACE) 并用于进一步跟踪和目标中心点计算。

目标中心点储存在活跃虚拟事件表面 (Surfaces of active virtual events, SAVE), 用于二自由度视觉伺服和抓取角度计算。3 个活跃表面如图 8 (a) 所示。该抓取方案采用如图 8 (b) 所示的结构, 机械臂末端安装两个真空吸盘, 在 $1.2\text{ m} \times 1.0\text{ m}$ 的工作台上对三棱柱、立方体、五棱柱三种物体进行了抓取实验。在不同速度和不同光照条件下, 机械臂均能实现稳定抓取, 平均抓取误差 (实际中心与抓取中心的误差距离) 为 16.1 mm 。但是这种抓取方式只适用于规则物体, 且要求像平面平行于工作台。

3.2 机械臂的滑动检测感知与防滑控制

机械臂夹持物体后, 物体有可能产生滑动, 如

果单纯地使用最大加持力, 则有可能损坏物体, 因此根据物体的滑动状态确定夹持控制方案, 能够恰好好处的完成机械臂夹取物体。但是滑动是一瞬间发生的运动, 事件相机的低延迟特性, 恰好能够解决该场景对感知系统的需求。

文献 [42] 首次使用事件相机作为“触觉传感器”, 仅通过事件相机感知透明硅材料与不同物体之间的接触面积, 并通过传统的图像处理方法, 分析重建帧上的事件分布以进行滑动检测, 无需事先了解物体的属性或摩擦系数也可以成功检测到位移。实验环境如图 9 所示, 在复杂光照下, 使用不同材质、形状、重量的物体上进行了实验, 滑动检测的平均精确率为 0.85, 平均延迟为 44.1 ms 。并同时使用高速相机进行了验证, 证明了方法的精确性。但是不同的材质会对结果产生较大影响, 金属物体由于摩擦和物理粘性, 具有更好的效果。并且会受到接触面积的影响, 接触面积较小的物体会产生更高的延迟。

相比于文献 [42] 中使用纯事件相机的方式, Taunayzov 等将触觉传感器与事件相机结合, 提出了视觉-触觉脉冲神经网络 (Visual-tactile spiking neural network, VT-SNN)^[43], 将视觉和触觉两种感知方式用于监督学习中。而且通过二者相结合的方式 (Prophesee 事件相机与 NeuTouch 触觉传感器), 实现了容器分类和物体的旋转滑动分类, 取得了比传统深度学习方法更好的效果。最后在神经形态处理器 Intel Loihi^[44] 上进行实验, 相比于使用 GPU, 推理效率更高, 功耗表现更优。

但是文献 [42] 中的实验场景显然不具备机动性, 因此 Baghaei Naeini 等^[45]提出了一种新的基于视觉测量 (Vision-based measurement, VBM) 的方

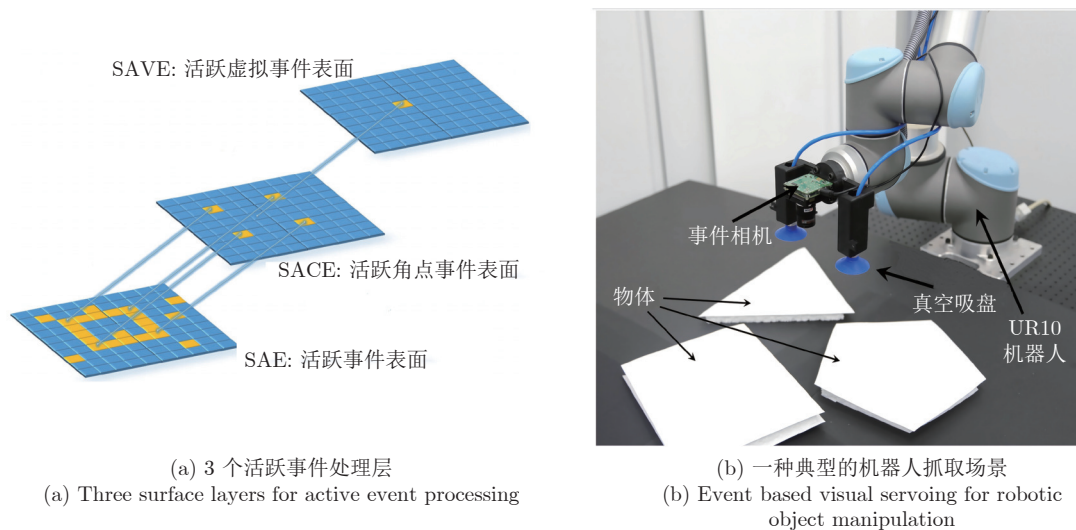


图 8 处理步骤以及抓取场景^[41]

Fig.8 Processing steps and capture scene of reference^[41]

法, 夹持器示意图如图 10 (a) 所示, 直接采用透明硅介质作为夹持器. 这是第一个基于事件相机来测量接触力并在单次抓握中进行材料分类的方法, 提供了更高的灵敏度, 更低的时延以及更少的功耗. 实验通过图 10 (b) 中的时延神经网络 (Time delay neural network, TDNN) 和高斯过程 (Gaussian process) 估计相同形状、尺寸物体在抓取和释放两个阶段的接触力, 并使用深度神经网络仅通过接触力进行材料分类. 其中, TDNN 方法的平均准确率为 79.17%, 时延为 21 ms, 均方误差 (Mean square error, MSE) 为 0.16 N, 证明了基于事件的传感器对于机器人抓取应用的适用性.

但是因为不同尺寸的物体与透明硅材料的接触面积不同, 因此对于尺寸不同的物体, 需要复杂的动态方法来关联每个时间戳上的事件与力的测量

值. 为了克服无法有效检测不同尺寸物体的问题, 因此在文献 [45] 的基础上, 提出了使用不同的长短记忆网络 (Long short-term memory, LSTM) 结构为传感器提供记忆, 进而动态估计接触力的方法^[46]. 传感器在抓取的早期阶段识别物体的尺寸, 使用卷积网络与循环层结合的方式, 使传感器能够根据物体的大小估计相应的接触力. 在与文献 [45] 相同的实验条件下, 时延提升到 10 ms, 误差从文献 [45] 中的 0.16 N 降低到 0.064 N (MSE), 有着更优良的表现, 可以应用于实时抓取应用. 但是对于未知物体的泛化能力较弱.

Muthusamy 等^[47] 使用事件相机作为滑动感知器件, 同样使用透明塑料板作为夹持器, 通过在物体上添加负载的方式使物体产生滑动; 能够以 2 kHz 的采样率实现夹持物体的滑动检测; 并根据检测到



图 9 实验装置的俯视图和侧视图^[42]

Fig.9 Top view and side view of the experimental device^[42]

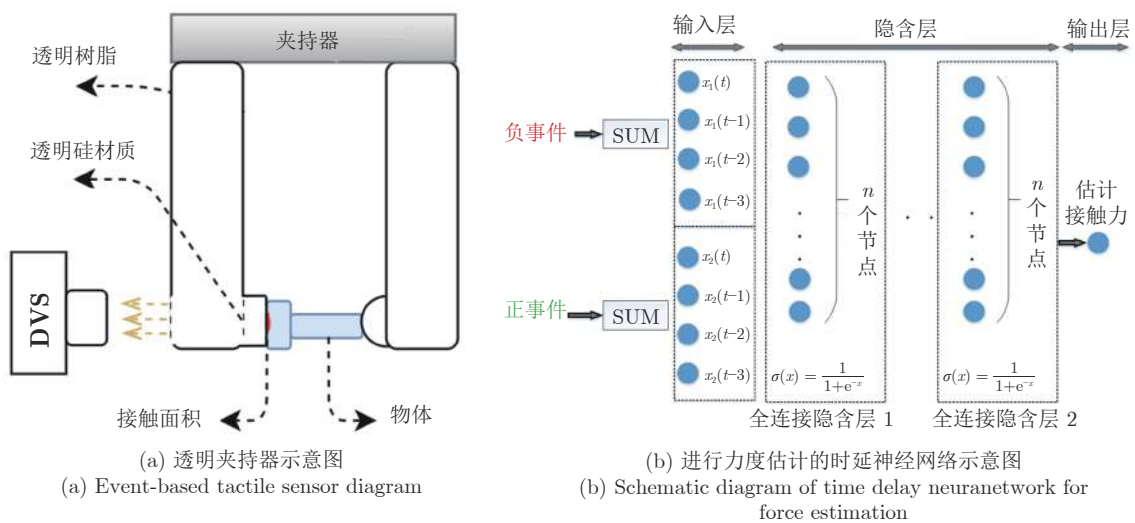


图 10 夹持器及网络示意图^[45]

Fig.10 Schematic diagram of gripper and network^[45]

的滑动幅度, 使用 Mamdani 型模糊逻辑控制器来调节握力. 另外, 文献中提出了基于事件数量阈值与基于特征两种滑动检测方法, 能够对环境进行噪声采样并自动校准. 其中对基于特征的滑动检测, 评估了 3 种特征检测算法、两种采样速率下以及复杂光照、震动环境下的性能. 实验证明, 使用 e-Harries^[48] 特征的滑动检测具有更强的鲁棒性, 成功率超过 90%. 但是需要被抓取对象有较为明显的纹理信息, 如边缘或角点.

文献 [49] 提出了一种基于事件的触觉图像传感器. 在黑色半球形的弹性体内表面嵌入白色标记点, 通过事件相机观测标记点的位移变化, 即可检测到指尖位置因受力而发生的形变, 能够以 0.5 ms 的响应速度进行检测, 处理之后可以检测到接触、滑动、位置、方向等信息. 但是这样的处理方式对于震动较为敏感, 容易出现误检测.

可以看出, 在当前滑动检测的应用中, 大部分论文都选择了将物体与透明板接触, 然后使用事件相机观测面积变化的方式. 但这种方法的问题就在于仍然要将事件流恢复成事件帧, 然后才能进行后续操作. 本节所述算法的总结如表 3 所示.

本节主要总结了基于事件相机的机械臂伺服控制研究. 尽管针对事件相机的机械臂伺服控制虽然已经有了一定的研究, 但很多方面仍处于起步阶段, 许多文献仅对于静态物体的抓取以及滑动检测有了较深入的研究, 而对于一些典型的视觉伺服场景, 同时也是事件相机最擅长的场景——运动物体的抓取, 尚未有深入研究, 这也是最能挖掘事件相机潜力的应用场景. 同时当前的研究也存在一定的缺陷, 如文献 [41] 中虽然实现了物体的抓取, 但仅能抓取静态的规则物体; 大部分文献中的方法仍然需要将事件流恢复为事件帧处理, 无法快速响应; 文献 [45–46] 使用的深度神经网络方法无法充分发挥事件流异步更新的优势.

总的来说, 当前基于事件相机的机械臂伺服面临的问题主要在于, 抓取应用仅适用于简单场景以

及规则物体; 并且无法充分利用事件异步更新的优势. 前者使得事件相机的应用场景受到了极大的限制; 后者在很大程度上削弱了事件相机高时间分辨率的优势. 因此对基于事件相机的机械臂伺服的后续研究, 工作重点应尽可能减少复杂环境下, 静态场景的影响, 以及寻找异步处理事件的方法.

4 基于事件相机的人形机器人感知和控制

人形机器人是研制通用机器人的重要解决方案. 人形机器人能够自然地适应人类环境, 但是目前从各个层面, 如感知、规划、驱动、控制等, 现有能力均与人形机器人的预期有极大的差距. 事件相机基于生物视觉成像原理, 呈现出神经形态特性, 其特点符合对人形机器人的感知能力预期, 因此已有学者将其应用于人形机器人的感知和控制中. 本节系统介绍在此方向上, 事件相机的应用和发展.

目前, 事件相机与人形机器人结合的研究, 大都是依托人形机器人 iCub 展开的. iCub 是由意大利技术研究院 (Italian Institute of Technology, IIT) 建造的一个通用机器人开发平台, 无需任何特殊设备即可在任何实验室中工作, 可用于研究小型人形机器人的爬行、行走、视觉、触摸、人工智能、认知、操纵、学习等课题. iCub 共具有 53 个自由度, 囊括了摄像头、麦克风、力/扭矩传感器、全身皮肤、陀螺仪和加速度计以及每个关节中的编码器等传感器; 在软件层面, iCub 设有 ROS (Robot operating system) 接口, 并使用 YARP (Yet another robot platform) 作为中间件. 传统帧相机在数据的传输、储存和处理过程中需要消耗大量的计算资源, 难以满足 iCub 自主控制的低功耗、低延迟需求; 而事件相机由于其高时间分辨率、低延迟、低功耗、高动态范围等特点, 有助于 iCub 实现自主计算, 进而实现 iCub 自主控制.

4.1 基于事件相机的人形机器人感知

为了实现事件相机与 iCub 的结合, IIT 机器人

表 3 部分基于事件相机的实验设备
Table 3 Experimental equipment based on event camera

| 文献 | 使用传感器 | 机械臂设备 | 方法 |
|------|------------------------------------|------------------------------|-----------|
| [41] | DAVIS240C | UR10 | 事件流 |
| [43] | Prophesee Gen3, NeuTouch | 7 DOF Franka Emika Panda arm | 事件流 + SNN |
| [42] | DAVIS240C | Baxter robot arm | 事件帧 |
| [45] | DAVIS240C | AX-12A Dynamixel servo motor | 事件帧 + DNN |
| [46] | DAVIS240C, ATI F/T sensor (Nano17) | Baxter robot arm | 事件帧 + CNN |
| [47] | DAVIS240C, ATI F/T sensor (Nano17) | Baxter robot daul-arm | 事件帧 |
| [49] | DVS128 | — | 事件帧 |

事件驱动感知团队 (Event-driven Perception for Robotics, EDPR) 在 iCub 的眼球上嵌入了一套事件视觉系统^[50-52], 同时将帧相机重新布置在 iCub 的头顶作为色彩等视觉信息的补充. 如图 11 所示, 事件视觉系统由两个 DVS 事件相机或 ATIS 事件相机、一个通用地址事件处理器以及一块 FPGA 芯片组成. 随着研究的逐步深入, 他们在 YARP 框架下开发了基于事件驱动的软件库^[53], 这使得事件相机可以与包括 iCub 在内的基于 YARP 的机器人一起使用^[54]. 该软件库可以不依赖 iCub 单独使用, 库中包含了光流、聚类追踪、角点检测、圆检测、粒子滤波、相机标定、预处理、可视化等功能模块, 并提供了示例程序及数据集.

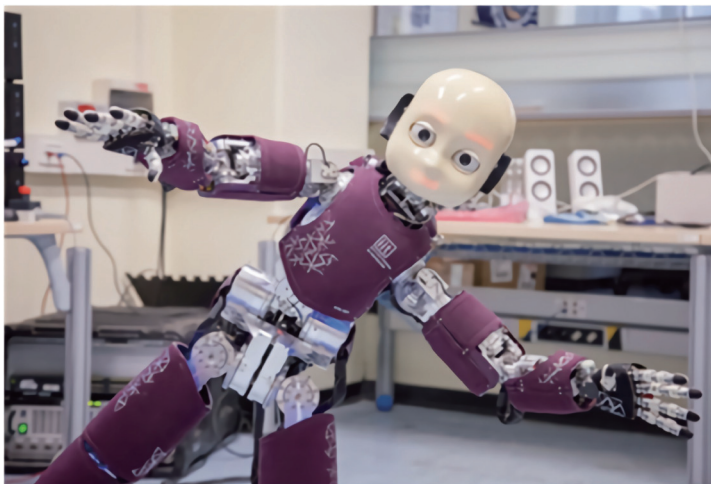
基于在 iCub 上嵌入的事件视觉系统, 学者们在圆检测、角点检测、机器视觉注意力机制等方面开展了大量研究. 在圆检测方面, Wiesmann 等^[55]在桌面静置小球, 通过转动 iCub 的头部或眼球, 使小球与 iCub 产生相对运动, 进而产生边缘事件, 然后对事件聚类, 实现图像分割; 接着采用 Hough 圆变换检测出圆形, 精确识别出小球. 针对复杂背景下的圆检测问题, Glover 等^[56]提出了一种基于光流的定向 Hough 圆变换算法, 即利用光流信息估计圆心的运动方向, 只在该方向上开展 Hough 圆检测.

在角点检测方面, Vasco 等^[48]提出了事件驱动的 e-Harris 算法. e-Harris 算法为每个异步事件创建一个局部检测窗口, 充分利用累计的事件在局部检测窗口应用 Harris 算法计算角点响应值. Vasco 等通过转动 iCub 的头部及眼球扫描静置物体, 检测到了对应的角点事件, 验证了算法的有效性. 但

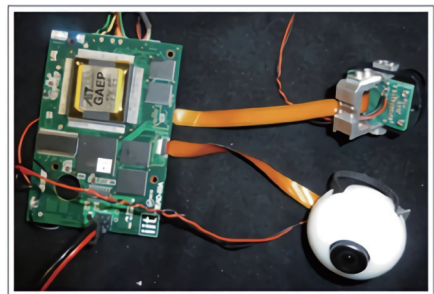
是当事件相机与目标均在运动时, 会产生大量背景事件, 这使得跟踪或识别目标物体的运动变得更加困难. 因此需要将由目标自身运动所产生的目标角点事件, 与由事件相机运动所产生的背景角点事件进行区分. 在文献^[57]中, Vasco 等通过聚类追踪 e-Harris 检测到的角点, 并估计角点的运动速度; 同时根据 iCub 的关节运动, 应用支持向量机和径向基函数核开展有监督学习. 最终实现了背景角点事件与目标角点事件的区分, 且精度高达 90% 以上.

计算机视觉中的注意力机制具有广泛的应用, 这种机制能够忽略无关信息而关注重点信息, 进而节约有限的计算资源. Rea 等^[58]利用事件视觉系统, 为 iCub 开发了一套低延迟的人工注意力系统, 可以快速计算出需要聚焦的位置. 实验结果表明, 相比于基于帧相机的人工注意力系统, 事件驱动人工注意力系统的延迟要低两个数量级, 所占用 CPU 资源也要低近乎一个数量级. 在高动态场景中, 事件驱动的人工注意力系统可以准确识别出视野中需要聚焦的点, 而基于帧相机的人工注意力系统完全无法进行识别. Iacono 等^[59]将原型对象注意力模型与嵌入 iCub 人形平台上的神经形态事件驱动相机配合使用, 为机器人提供了低延迟、计算效率高的注意力系统.

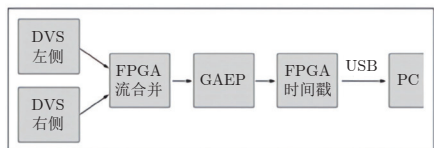
此外, 基于 iCub 及其嵌入事件视觉系统, 结合机器学习技术, 还有许多有意义的研究项目. 例如 Iacono 等^[51]将深度学习应用于 iCub 上的事件相机开展目标检测; Monforte 等^[50]采用长短期记忆人工神经网络开展轨迹预测等.



(a) iCub 机器人^[54]
(a) iCub robot^[54]



(b) iCub 嵌入的事件视觉系统^[52]
(b) The event vision system embedded in iCub^[52]



(c) 数据流^[52]
(c) Data flow^[52]

图 11 人形机器人事件视觉系统
Fig.11 Event vision system of humanoid robot

4.2 人形机器人双目凝视控制

基于在 iCub 上嵌入的事件视觉系统, 可以实现驱动 iCub 的头部及眼球关节实现凝视追踪、深度估计等复杂操作. Glover 等^[66] 利用基于光流的定向 Hough 圆变换检测算法识别小球的空间位置, 并驱动 iCub 转动眼球及头部, 实现对小球的实时凝视跟踪. 在大多数情况下, 定向 Hough 圆变换都能取得很好的检测效果. 但是由于 Hough 圆变换需要固定检测窗口, 当场景中有多个以不同速度移动的物体时, 事件发生率就会有很大波动, 最佳检测阈值将针对不同的对象而变化很大, 并且还会随时间而变化; 当相机追踪球体时, 若相对运动很小, 则可能丢失追踪目标. 为了解决上述问题, Glover 等^[66] 又提出了一种粒子滤波算法, 展现出了更卓越的鲁棒性.

在深度估计方面, iCub 可以通过驱动双目运动, 使其视线聚焦, 进而通过双目之间的相对姿态获得深度信息^[61]. 但帧相机存在延迟高、鲁棒性差等问题, 目标快速运动时算法甚至会失效. 针对该问题, Vasco 等^[62] 用基于事件驱动的视差调谐双目 Gabor 滤波器得到双目视差, 然后驱动眼动使视差收敛, 进而估计出深度信息. 这一改进使得延迟从帧相机的 1 s 降低至事件相机的 200 ms, 且降低了对光照等环境条件的要求, 实现了更好的鲁棒性.

在 iCub 上同时安装麦克风与事件视觉系统, 还可实现多信息融合. Akolkar 等^[63] 在 iCub 的头部安装麦克风, 首先利用 iCub 上的事件视觉系统检测视觉空间中的物体“碰撞”信息. 基于视觉检测到的“碰撞”事件, 可能是真实碰撞, 也可能是假碰撞, 如两个物体在视觉空间中相互遮挡. 真实碰撞通常会产生声音, 因此发生真实碰撞时麦克风可以检测到声音事件. 当通过视觉与听觉同时检测到碰撞事件时, 即可认为视觉空间中发生了真实碰撞. 该基于视听信息融合的碰撞检测算法在 iCub 检测人拍掌的场景中得到了验证.

综上, 相比于传统帧相机, 事件相机在 iCub 上的应用, 以耗费更少的计算资源降低了系统的延迟, 提高了 iCub 在不同光照强度、复杂背景条件、高动态场景下的自主控制能力, 但目前的研究仍较为基础, 主要是驱动 iCub 的头部和眼球进行凝视追踪, 更进一步的应用还有待进一步挖掘, 如抓取、接球等更为复杂的操作.

5 结合事件相机与脉冲神经网络的感知和控制方法

事件相机的离散、异步的数据特点与脉冲神经

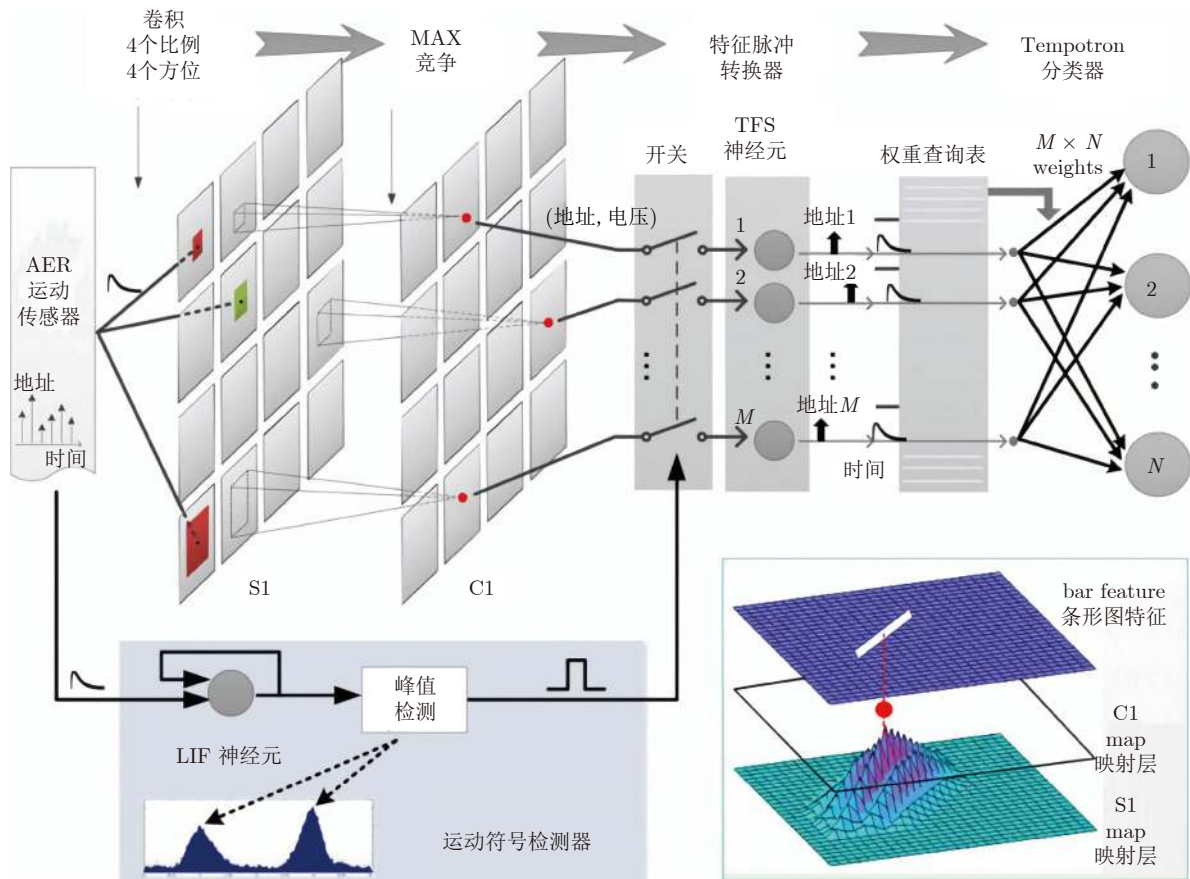
网络的输入需求自然的契合, 因此将它们结合, 产生了一系列感知和控制的新方法. 本节重点介绍事件相机与脉冲神经网络结合的感知和控制工作, 侧重算法和网络的设计.

第 2.3 节中提到, 脉冲神经网络 (SNN) 采用脉冲神经元为基本单位, 由前馈型、递归型以及混合型 3 种拓扑结构, 而且可以在极低的功耗下工作. 脉冲神经元从神经科学角度出发进行建模, 现有的模型包括 HH (Hodgkin-huxley) 模型、LIF (Leaky integrate and fire) 模型、SRM (Spike response model) 模型等. 脉冲神经网络的训练方式包括有监督学习 (如 SpikeProp、Multi-SpikeProp、Hebbian、ReSuMe、Chronotron、SPAN、SWAT、Tempotron 等)、基于 STDP (Spike-timing-dependent plasticity) 的无监督学习、强化学习等. 事件相机以事件流的形式输出, 这与脉冲神经网络的输入十分契合, 因此选择脉冲神经网络搭配事件相机工作具有相当优势.

结合事件相机和 SNN 的识别研究工作较多. Zhao 等^[64-65] 开发了一种基于事件的前馈分类系统, 该系统采用基于地址事件的 Tempotron 分类器实现分类. 如图 12 所示, 系统由卷积、竞争、特征脉冲转换器、运动符号检测器和 Tempotron 分类器构成. 事件相机所产生的每个地址事件都会先由 Gabor 滤波器滤波, 然后与领域事件竞争, 获胜者才能进入图 12 所示的 C1 特征层. 同时, 地址事件会传递到运动符号检测器 (由 LIF 脉冲神经元和峰值检测单元组成) 当中, 当 LIF 脉冲神经元的输出超过峰值时, 将产生脉冲, 进而导通图 12 中的开关, 此时可以将 C1 特征层中胜出的单元导入到特征脉冲转换器中, 由特征脉冲转换器将每个特征编码为时域脉冲. 最终由 Tempotron 分类器实现分类. Tempotron 分类器由 LIF 脉冲神经元所构成, 是一类有监督学习算法. 上述分类器在基于事件相机的人体姿态识别、卡牌识别、数字识别等方面能取得较好效果.

Zhao 等在文献 [66] 中, 对前述前馈分类系统结构进行简化, 取消了特征脉冲转换器和运动符号检测器, 同时由 LIF 脉冲神经元构建 S1 层. 对地址事件, 在采用 Gabor 滤波器滤波后, 经 S1 层、竞争之后, 直接由 Tempotron 分类器进行分类. 对该系统在 MNIST-DVS 数据集上进行测试, 准确性达到 88.14%. Nan 等^[67] 提出了一个基于事件的层次结构模型, 该模型由特征提取层和 SNN 组成. 特征提取层包括 Gabor 滤波、池化、事件时间表面特征提取等操作. SNN 采用 Tempotron 算法训练, 人脸识别的精度可以达到 92.5%.

Shen 等^[68] 应用 DVS128 事件相机和 SNN 的结合,

图 12 前馈分类系统的结构^[65]Fig. 12 Structure of feedforward classification system^[65]

开发了篮球比赛中的自动得分检测系统, 识别准确率高达 91%. 地址事件处理框架如图 13 所示, 主要由特征提取器、峰值检测器和不平衡 Tempotron 分类器组成. 事件相机输出的地址事件先经 Gabor 滤波, 然后在 S1 层和 C1 层分别完成卷积和竞争; 峰值检测器由 LIF 脉冲神经元构成, 当神经元达到阈值后即输出脉冲, 此时将 C1 层提取的数据传输到 Tempotron 分类器处理.

值得注意的是, 上述结合事件相机和 SNN 的识别算法结构具有很大的相似性, 决策端均是采用仅包含输入层与输出层的两层脉冲神经网络. Tempotron 是一个二层网络学习算法, 对这样的两层脉冲神经网络的训练是有效的, 但是无法扩展到多层脉冲神经网络. Massa 等^[69] 使用 DvsGesture 数据集训练深度神经网络, 通过将训练后的深度神经网络转换为脉冲神经网络, 并在 Loihi 上实现, 可以实现实时手势识别, 精度高达 89.65%. Camunas-Mesa 等^[70-71] 在 Spartan6 FPGA 上实现了一个 4 层卷积脉冲神经网络, 可以实现扑克牌花色识别. 该卷积脉冲神经网络中共包含 4 个卷积层、两个采样层以及一个决策层, 卷积模块基于 LIF 模型实现, 模型

参数由图像驱动的卷积神经网络映射而来^[72]. 在测试中, 他们使用 DVS 相机在 1 s 内拍摄了 40 张扑克牌, 可以达到 97.5% 的识别率.

在 SNN 无监督学习方面, STDP 是一种常用方法. Diehl 等^[73] 提出了一种两层架构的脉冲神经网络结构, 第 1 层为 28×28 的输入层, 维度与输入数据相同; 第 2 层为处理层, 又可细分为兴奋层和抑制层, 两层由同等数量的兴奋性神经元和抑制性神经元组成. 输入层与兴奋层之间全连接; 每个兴奋性神经元一对一地连接到抑制性神经元 (如图 14 中兴奋层左侧的射线), 同时抑制性神经元会投射到所有兴奋性神经元 (如图 14 中抑制层左侧的射线). 对该网络采用 STDP 算法进行训练, 在 MNIST 数据集上可以达到 95% 的精度. 但是 Diehl 等^[73] 在预处理 MNIST 数据集时没有采用事件相机的数据格式. Iyer 等^[74] 采用 N-MNIST 数据集 (N-MNIST 数据集为采用 DVS 事件相机记录的 MNIST 数据集) 重新测试了 Diehl 等^[73] 提出的网络, 精度可以达到 80.63%.

SNN 与事件相机的结合, 不仅可用于识别工作, 还能进一步助力机电伺服、运动检测与追踪等.

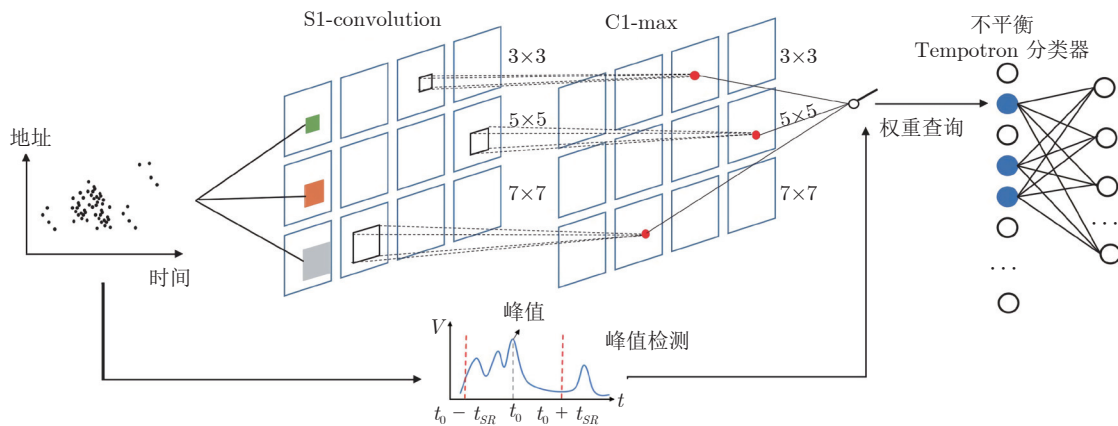


图 13 篮球得分检测框架^[68]
 Fig. 13 Basketball score detection framework^[68]

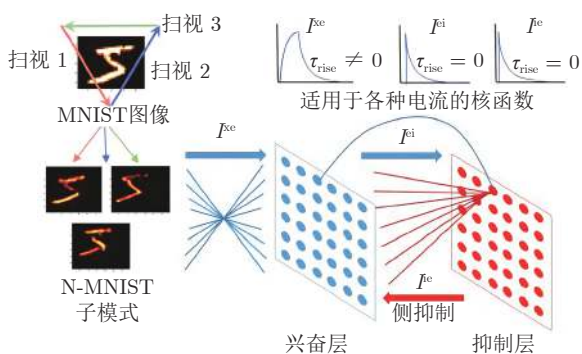


图 14 两层脉冲神经网络结构示意图^[74]
 Fig. 14 Schematic diagram of two-layer spiking neural network structure^[74]

Cheng 等^[75] 将 DVS 事件相机、SpiNNaker 平台以及机电伺服模块集成到由脉冲神经网络控制的自主机器人当中, 实现守门任务. 机电伺服模块由伺服电机、长臂以及长臂末端的“守门员”组成. SpiNNaker 平台上运行的脉冲神经网络仅有输入层和输出层, 不包含中间层, 共有 8 个输出神经元, 对应“守门员”8 个位置. 当球来临时, DVS 相机捕捉到球的运动, 生成地址事件传递给单片机, 单片机将数据打包发送到 SpiNNaker 平台. 当某个输出神经元输出脉冲时, 由 SpiNNaker 平台向单片机发送数据包, 进而由单片机驱动机电伺服模块, 使“守门员”运动到指定位置拦截小球.

Ting 等^[76] 提出了一种六足机器人的步态模仿解决方案, 该方案采用基于 CeleX5 事件相机的前馈脉冲神经网络实现. 如图 15 所示, 在该方案中, Ting 等^[76] 采用 CeleX5 事件相机观察“专家”的步态, 所产生的事件采用 Andpool 进行滤波, 然后采用两层脉冲神经网络处理, 实现“学生”步态与“专家”步态的同步控制. 脉冲神经网络输出层有 6 个

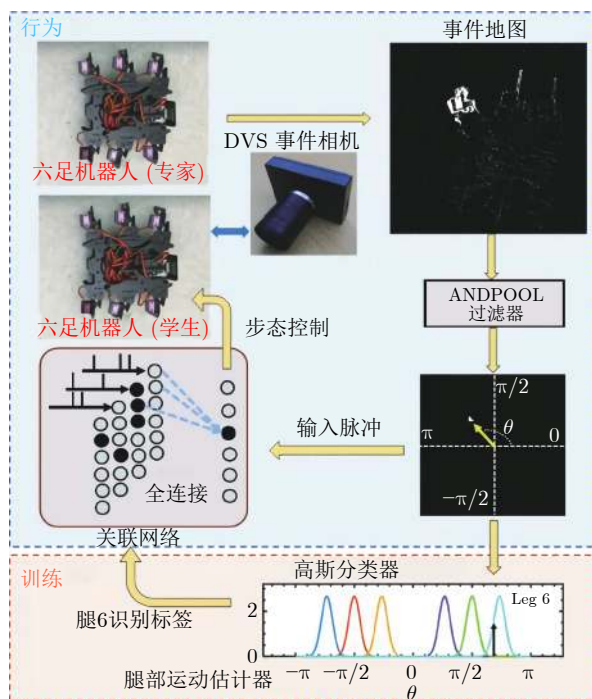


图 15 步态模仿系统结构示意图^[76]
 Fig. 15 Schematic diagram of gait simulation system^[76]

神经元, 对应“学生”的 6 条腿; 输入层与输出层全连接. Youssef 等^[77] 在鳗鱼型水下机器人 Envirobot 头部安装双目 DVS 事件相机, 采用 LIF 脉冲神经网络处理事件数据, 产生 Envirobot 躯干关节控制信号, 驱动 Envirobot 向探测到的目标游动.

Blum 等^[78] 在神经形态芯片 ROLLS 上部署 SNN, 采用 DVS 事件相机作 (像素为 128×128) 为传感器, 采用 Parallella 并行计算平台进行实时计算, 实现了 PushBot 无人小车的避撞控制. 神经网络的输出分为 3 个簇, 每簇 16 个神经元, 分别输出信号控制 PushBot 左转、右转以及前进, 3 个方向的速度

与被激活的输出神经元的数目相关. 在进行避撞控制时, 将 DVS 视图的下半部分用于障碍探测, 每个 4×64 的像素阵列连接一个输入神经元, 共 32 个神经元构成输入层, 输入层与前进簇抑制连接; 像素阵列左右两侧连接的输入神经元分别与左转、右转簇神经元采用激活连接, 左转与右转簇之间有抑制连接. 同时为了在没有障碍物的情形下驱动 PushBot 持续前进, 在前进簇神经元旁边设置了 8 个神经元, 与前进簇激活连接. PushBot 通过 DVS 视图的上半部分探测目标, 每个 2×64 的像素阵列连接一个输入神经元, 共 64 个神经元构成输入层, 用于指引 PushBot 的前进目标方向.

此外, Renner 等^[79] 基于 DAVIS 240C 事件相机和在 Loihi 上实现的递归脉冲神经网络, 实现了用于选择性关注和跟踪的事件驱动视觉和处理系统. Gehrig 等^[80] 使用卷积脉冲神经网络估计了一个旋转运动的 DAVIS240C 事件相机三自由度角速度. 该卷积脉冲神经网络由 5 个卷积层、一个全局平均池化层以及一个全连接层组成, 其神经元采用 SRM 模型, 采用一阶优化方法训练.

6 基于事件相机的其余类型机电系统感知和控制

除了无人机、机械臂和人形机器人, 事件相机也在其余类型的机电系统中有典型的应用. 如铅笔倒立控制等小车的追踪控制等. Conrardt 等^[13, 81] 采用由两部 DVS 事件相机和一个由两台伺服电机驱动的工作台, 实现了铅笔倒立平衡控制, 如图 16 所示. 两部 DVS 相机分别安装在平台 X 、 Y 两个方向上, 当铅笔倾斜时会可以分别获得两个方向上的地址事件, 进而分别实现平面线检测; 基于两个相机平面线检测的结果可以对铅笔进行三维线估计, 进而采用 PD 控制律驱动两部伺服电机, 最终实现了铅笔倒立平衡控制.

在机器人控制方面, Moeys 等^[82] 在 Summit XL 小车上安装 DAVIS 相机, 结合卷积神经网络, 实现对另一部 Summit XL 机器人的追踪. DAVIS 相机可以同时输出地址事件和帧图像, 在预处理数据时每记录 5 000 个地址事件就合成一张直方图, 合成直方图和帧图像作为卷积神经网络的输入. 网络结构为 4C5-R-2S-4C5-R-2S-40F-R-4F, 4 个输出分别是左转、右转、前进、消失, 追踪的准确度可以达到 80%, 如图 17 所示. 在守门机器人驱动方面, 除了 Cheng 等^[75] 采用的将事件相机与 SNN 结合的算法, Delbruck 等^[83-84] 还采用聚类算法识别和追踪小球的运动, 应用小球的速度和位置信息驱动伺服



图 16 铅笔倒立平衡控制系统: 两台 DVS 事件相机进行线检测 (右上角及左下角), 左上角运动平台由中间两部伺服电机驱动^[81]

Fig. 16 Photo of balancer hardware: 2 DVS (right top and bottom left), the motion table (top left) actuated by two servos (center)^[81]

系统到达指定位置实现拦截.

Mueller 等^[85-86] 提出了一种基于事件反馈的控制方法, 直接利用事件相机高时间分辨率、低延迟、低功耗、高动态范围的特点. 在实验中, 基于 PD 控制律和控制实验 LQR (Linear quadratic regulator) 调节器控制伺服电机驱动滚筒装置, 取得了较好的效果. Delbruck 等^[87] 将 DAVIS 安装在槽车轨道的上帝视角, 当小车行驶时会产生事件, 这些事件可用于计算小车的加减速指令. 与人才操控的小车比赛, 基于 DAVIS 控制的小车能赢超过 80% 的比赛. Censi^[88] 在将原始事件数据采用线性滤波器滤波后, 研究了在事件相机反馈下的图像空间中的航向调节问题. 在线性系统整定方面, Singh 等^[89-91] 应用事件相机的成像机制, 研究了连续线性时不变系统的二次稳定问题, 并在考虑测量噪声和测量信号离散的基础上设计了一种 H_∞ 控制器, 能实现不稳定系统的稳定控制.

事件相机在机电系统控制、环境识别与感知等方面具有极其广泛的应用, 相比于传统帧相机具有显著优势. 但将事件相机应用于识别与控制仍具有相当的挑战性, 譬如事件相机与脉冲神经网络的结合, 结构简单的网络易于训练但性能有限、而复杂网络的训练仍面临许多挑战; 结合事件相机的机电系统控制能显著提高系统的响应时间, 但新算法、新应用的开发仍需要深入研究.

7 总结与展望

事件相机与传统帧相机具有完全不同的工作机

理, 具有数据量小、延迟低、动态范围高的特点, 因此将其置于机器人控制回路的闭环之中, 降低了测量、感知环节的延迟和数据量. 本文介绍了事件相机在无人机、机械臂、人形机器人等机器人系统中

感知与控制的最新工作以及结合事件相机发展的新型控制算法, 技术发展脉络如图 18 所示, 图中横轴对应感知部分技术, 主要分为两大类: 一类 (负横轴) 先将事件相机采集的事件流数据对应成事件帧,

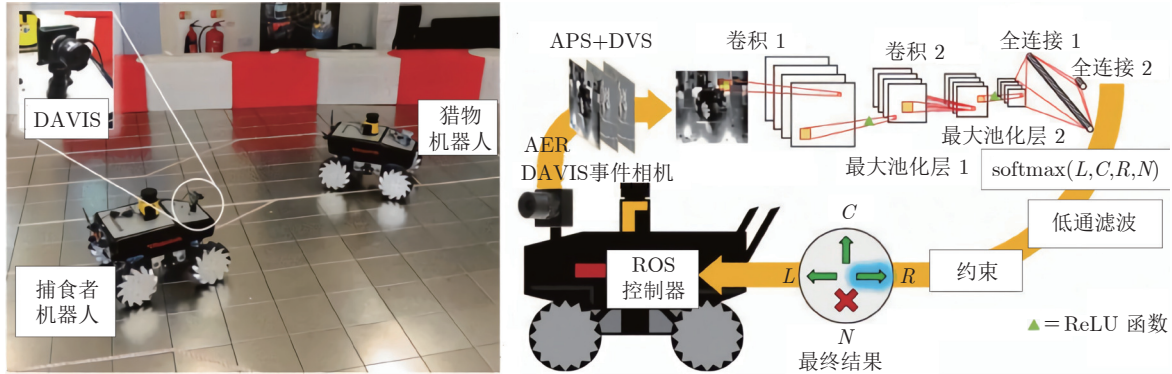


图 17 Summit XL 机器人追踪实验^[82]

Fig.17 Summit XL robot tracking experiment^[82]

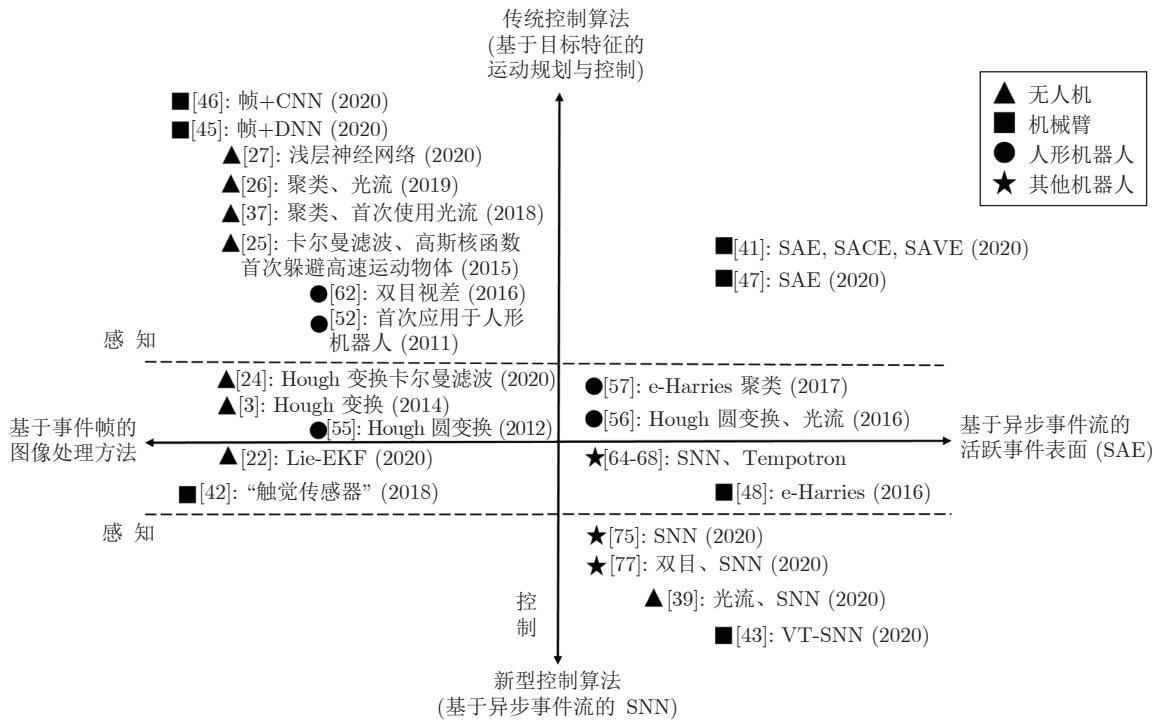


图 18 技术发展脉络图 (横轴为感知技术分类, 负横轴为基于事件帧的图像感知算法, 正横轴为基于异步事件流的感知算法; 纵轴为控制技术分类, 正纵轴为基于目标特征的运动规划与控制的传统方法, 负纵轴为基于异步事件流的新型控制算法, 其中使用相似技术路线的文献处于同一象限, 标注是根据机器人类型进行分类并按时间顺序排列, 虚线内文献只涉及基于事件相机的感知技术)

Fig.18 Development venation map of technology (The horizontal axis is the classification of perception technology, the negative horizontal axis is the image perception algorithm based on event frame, and the positive horizontal axis is the perception algorithm based on asynchronous event stream. The vertical axis is the classification of control technology, the positive vertical axis is the traditional method of motion planning and control based on object features, and the negative vertical axis is a new control algorithm based on asynchronous event stream. The literatures using similar technical routes are in the same quadrant, and the labels are classified according to robot types and arranged in chronological order. The literatures in the dotted line only involve the sensing technology based on event cameras)

可尝试利用图像处理方法解决高速、高动态的应用场景问题。另一类(正横轴)则根据事件相机采集数据自身特点直接设计新的事件流处理算法,如基于异步事件流的活跃事件表面;纵轴对应控制部分技术也分为两大类,包括基于目标特征的运动规划与控制的传统算法(正纵轴),以及充分考虑事件相机异步事件流特性的新型控制算法(负纵轴),代表如SNN。目前的研究根据上述技术分类组合可得到四条主流技术路线,但是仍处于起步阶段,主要存在的问题和发展方向有:

1) 视觉系统是机器人完成跟踪、抓取以及操作的关键基础,而目前基于事件的视觉算法难以满足复杂环境下物体检测、跟踪与识别任务的要求。例如当事件相机随机器人(如无人机、机械臂末端等)运动时,输出的数据包括静止物体、运动物体产生的事件以及噪声事件,从这些杂乱的事件数据中提取可用的目标信息是一大难点。因此,有必要进一步研究基于事件的运动分割技术以及基于机器/深度学习的目标检测与识别技术。

2) 目前在结合事件相机的机器人控制系统中,视觉模块的延迟占比较高,为进一步提升机器人操控的灵活性,有必要根据事件相机特性设计更加快速高效的视觉模块。例如,可深挖事件流的异步特性并开发异步、超低延迟的特征检测与跟踪算法。另外,随着近年来事件相机分辨率的不断提升,在计算效率方面视觉算法将面临更多挑战。

3) 目前在结合事件相机的机器人控制系统中,控制器设计较为简单。由于事件相机对光照变化较为敏感,可能产生包含大量噪声事件的视觉信号,而这些噪声的不确定性将很容易影响控制系统的性能,因此有必要进一步研究新型的控制算法,以实现基于事件相机的机器人稳定控制。同时,也需要研究噪声事件的产生机理,这将有助于去噪算法的设计以及控制器分析。

4) 目前结合事件相机的机器人控制系统仍依赖传统机器人的控制流程,如首先进行视觉特征的检测跟踪,再进行期望轨迹规划与轨迹跟踪等,并没有发挥出事件相机神经形态与高时间分辨率的特点。理论上,事件相机与脉冲神经网络具有自然的适配性,研究有效的脉冲神经网络训练方法,以实现端到端的控制指令生成,具有一定的研究意义。

综上,在未来的研究中,建议充分利用事件相机输出的异步事件流的特性,设计低延迟、低计算量的视觉处理算法,研究新型的控制算法以及研究事件相机与脉冲神经网络结合的方法以实现机器人的灵活与智能操控。

References

- 1 Posch C, Serrano-Gotarredona T, Linares-Barranco B, Delbruck T. Retinomorph event-based vision sensors: Bioinspired cameras with spiking output. *Proceedings of the IEEE*, 2014, **102**(10): 1470–1484
- 2 Mueggler E, Huber B, Scaramuzza D. Event-based, 6-DOF pose tracking for high-speed maneuvers. In: *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Chicago, USA: IEEE, 2014. 2761–2768
- 3 Cho D I, Lee T J. A review of bioinspired vision sensors and their applications. *Sensors and Materials*, 2015, **27**(6): 447–463
- 4 Mahowald M. VLSI Analogs of Neuronal Visual Processing: A Synthesis of form and Function [Ph.D. dissertation], California Institute of Technology, USA, 1992.
- 5 Boahen K. A retinomorph chip with parallel pathways: Encoding increasing, on, decreasing, and off visual signals. *Analog Integrated Circuits and Signal Processing*, 2002, **30**(2): 121–135
- 6 Kramer J. An integrated optical transient sensor. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 2002, **49**(9): 612–628
- 7 Rüedi P F, Heim P, Kaess F, Grenet E, Heitger F, Burgi P Y, et al. A 128×128 Pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction. *IEEE Journal of Solid-State Circuits*, 2003, **38**(12): 2325–2333
- 8 Luo Q, Harris J G, Chen Z J. A time-to-first spike CMOS image sensor with coarse temporal sampling. *Analog Integrated Circuits and Signal Processing*, 2006, **47**(3): 303–313
- 9 Chi Y M, Mallik U, Clapp M A, Choi E, Cauwenberghs G, Etienne-Cummings R. CMOS camera with in-pixel temporal change detection and ADC. *IEEE Journal of Solid-State Circuits*, 2007, **42**(10): 2187–2196
- 10 Lichtsteiner P, Posch C, Delbruck T. A 128×128 120 dB 15 μ s latency asynchronous temporal contrast vision sensor. *IEEE Journal of Solid-State Circuits*, 2008, **43**(2): 566–576
- 11 Gallego G, Delbrück T, Orchard G, Bartolozzi C, Taba B, Censi A, et al. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(1): 154–180
- 12 Amir A, Taba B, Berg D, Melano T, McKinstry J, Di Nolfo C, et al. A low power, fully event-based gesture recognition system. In: *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE, 2017. 7388–7397
- 13 Conradt J, Cook M, Berner R, Lichtsteiner P, Douglas R J, Delbruck T. A pencil balancing robot using a pair of AER dynamic vision sensors. In: *Proceedings of the 2009 IEEE International Symposium on Circuits and Systems*. Taipei, China: IEEE, 2009. 781–784
- 14 Xu H Z, Gao Y, Yu F, Darrell T. End-to-end learning of driving models from large-scale video datasets. In: *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE, 2017. 3530–3538
- 15 Posch C, Matolin D, Wohlgenannt R. A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS. *IEEE Journal of Solid-State Circuits*, 2011, **46**(1): 259–275
- 16 Hu Sheng-Long. Slam for Autonomous Vehicles Based on Stereo Vision in the City Scene with Dynamic Objects [Master thesis], Harbin Institute of Technology, China, 2020. (胡生龙. 城市内动态场景下基于双目视觉的无人车同步定位与建图 [硕士学位论文], 哈尔滨工业大学, 中国, 2020.)
- 17 Suh Y, Choi S, Ito M, Kim J, Lee Y, Seo J, et al. A 1280×960 dynamic vision sensor with a $4.95\text{-}\mu\text{m}$ pixel pitch and motion artifact minimization. In: *Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. Seville, Spain: IEEE, 2020. 1–5

- 18 Posch C, Matolin D, Wohlgenannt R. An asynchronous time-based image sensor. In: Proceedings of the 2008 IEEE International Symposium on Circuits and Systems (ISCAS). Seattle, USA: IEEE, 2008. 2130–2133
- 19 Brandli C, Berner R, Yang M H, Liu S C, Delbruck T. A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-State Circuits*, 2014, **49**(10): 2333–2341
- 20 Taverni G, Moeys D P, Li C H, Cavaco C, Motsnyi V, Bello D S S, Delbruck T. Front and back illuminated dynamic and active pixel vision sensors comparison. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018, **65**(5): 677–681
- 21 Chen S S, Guo M H. Live demonstration: CeleX-V: A 1 M pixel multi-mode event-based sensor. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach, USA: IEEE, 2019. 1682–1683
- 22 Chamorro Hernández W O, Andrade-Cetto J, Solà Ortega J. High-speed event camera tracking. In: Proceedings of the 31st British Machine Vision Virtual Conference (BMVC). Virtual Event, UK, 2020. 1–12
- 23 Deray J, Solà J. Manif: A micro Lie theory library for state estimation in robotics applications. *The Journal of Open Source Software*. 2020, **5**(46): 1371
- 24 Dimitrova R S, Gehrig M, Brescianini D, Scaramuzza D. Towards low-latency high-bandwidth control of quadrotors using event cameras. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation. Paris, France: IEEE, 2020. 4294–4300
- 25 Mueggler E, Baumli N, Fontana F, Scaramuzza D. Towards evasive maneuvers with quadrotors using dynamic vision sensors. In: Proceedings of the 2015 European Conference on Mobile Robots (ECMR). Lincoln, UK: IEEE, 2015. 1–8
- 26 Falanga D, Kim S, Scaramuzza D. How fast is too fast? The role of perception latency in high-speed sense and avoid. *IEEE Robotics and Automation Letters*, 2019, **4**(2): 1884–1891
- 27 Sanket N J, Parameshwara C M, Singh C D, Kuruttukulam A V, Fermüller C, Scaramuzza D, et al. EVDodgeNet: Deep dynamic obstacle dodging with event cameras. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020. 10651–10657
- 28 Falanga D, Kleber K, Scaramuzza D. Dynamic obstacle avoidance for quadrotors with event cameras. *Science Robotics*, 2020, **5**(40): Article No. eaaz9712
- 29 Mitrokhin A, Fermüller C, Parameshwara C, Aloimonos Y. Event-based moving object detection and tracking. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 1–9
- 30 Khatib O. Real-time obstacle avoidance for manipulators and mobile robots. *Autonomous Robot Vehicles*. New York: Springer, 1986. 396–404
- 31 Lin Y, Gao F, Qin T, Gao W L, Liu T B, Wu W, et al. Autonomous aerial navigation using monocular visual-inertial fusion. *Journal of Field Robotics*, 2018, **35**(1): 23–51
- 32 Oleynikova H, Honegger D, Pollefeys M. Reactive avoidance using embedded stereo vision for MAV flight. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA). Seattle, USA: IEEE, 2015. 50–56
- 33 Barry A J, Florence P R, Tedrake R. High-speed autonomous obstacle avoidance with pushbroom stereo. *Journal of Field Robotics*, 2018, **35**(1): 52–68
- 34 Huang A S, Bachrach A, Henry P, Krainin M, Maturana D, Fox D, et al. Visual odometry and mapping for autonomous flight using an RGB-D camera. *Robotics Research*. Cham: Springer, 2016. 235–252
- 35 Eguñuz A G, Rodríguez-Gómez J P, Martínez-de Dios J R, Ollero A. Asynchronous event-based line tracking for time-to-contact maneuvers in UAS. In: Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Las Vegas, USA: IEEE, 2020. 5978–5985
- 36 Lee D N, Bootsma R J, Land M, Regan D, Gray R. Lee's 1976 paper. *Perception*, 2009, **38**(6): 837–858
- 37 Pijnacker Hordijk B J, Scheper K Y W, De Croon G C H E. Vertical landing for micro air vehicles using event-based optical flow. *Journal of Field Robotics*, 2018, **35**(1): 69–90
- 38 Scheper K Y W, De Croon G C H E. Evolution of robust high speed optical-flow-based landing for autonomous MAVs. *Robotics and Autonomous Systems*, 2020, **124**: Article No. 103380
- 39 Hagenaaers J J, Paredes-Vallés F, Bohté S M, de Croon G C H E. Evolved neuromorphic control for high speed divergence-based landings of MAVs. *IEEE Robotics and Automation Letters*, 2020, **5**(4): 6239–6246
- 40 Tavanaei A, Ghodrati M, Kheradpisheh S R, Masquelier T, Maida A. Deep learning in spiking neural networks. *Neural Networks*, 2019, **111**: 47–63
- 41 Muthusamy R, Ayyad A, Halwani M, Swart D, Gan D M, Seneviratne L, et al. Neuromorphic eye-in-hand visual servoing. *IEEE Access*, 2021, **9**: 55853–55870
- 42 Rigi A, Baghaei Naeini F, Makris D, Zweiri Y. A novel event-based incipient slip detection using dynamic active-pixel vision sensor (DAVIS). *Sensors*, 2018, **18**(2): Article No. 333
- 43 Taunyazov T, Sng W, Lim B, See H H, Kuan J, Ansari A F, et al. Event-driven visual-tactile sensing and learning for robots. In: Proceedings of the 16th Robotics: Science and Systems 2020. Virtual Event/Corvallis, Oregon, USA, 2020.
- 44 Davies M, Srinivasa N, Lin T H, Chinya G, Cao Y Q, Choday S H, et al. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 2018, **38**(1): 82–99
- 45 Baghaei Naeini F, AlAli A M, Al-Husari R, Rigi A, Al-Sharman M K, Makris D, et al. A novel dynamic-vision-based approach for tactile sensing applications. *IEEE Transactions on Instrumentation and Measurement*, 2020, **69**(5): 1881–1893
- 46 Baghaei Naeini F, Makris D, Gan D M, Zweiri Y. Dynamic-vision-based force measurements using convolutional recurrent neural networks. *Sensors*, 2020, **20**(16): Article No. 4469
- 47 Muthusamy R, Huang X Q, Zweiri Y, Seneviratne L, Gan D M. Neuromorphic event-based slip detection and suppression in robotic grasping and manipulation. *IEEE Access*, 2020, **8**: 153364–153384
- 48 Vasco V, Glover A, Bartolozzi C. Fast event-based Harris corner detection exploiting the advantages of event-driven cameras. In: Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Daejeon, Korea (South): IEEE, 2016. 4144–4149
- 49 Kumagai K, Shimonomura K. Event-based tactile image sensor for detecting spatio-temporal fast phenomena in contacts. In: Proceedings of the 2019 IEEE World Haptics Conference (WHC). Tokyo, Japan: IEEE, 2019. 343–348
- 50 Monforte M, Arriandiaga A, Glover A, Bartolozzi C. Where and when: Event-based spatiotemporal trajectory prediction from the iCub's point-of-view. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020. 9521–9527
- 51 Iacono M, Weber S, Glover A, Bartolozzi C. Towards event-driven object detection with off-the-shelf deep learning. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 1–9
- 52 Bartolozzi C, Rea F, Clercq C, Fasnacht D B, Indiveri G, Hofstätter M, et al. Embedded neuromorphic vision for humanoid

- robots. In: Proceedings of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Colorado Springs, USA: IEEE, 2011. 129–135
- 53 Glover A, Vasco V, Iacono M, Bartolozzi C. The event-driven software library for YARP-with algorithms and iCub applications. *Frontiers in Robotics and AI*, 2018, **4**: Article No. 73
- 54 IIT. ICub robot [Online], available: <https://icub.iit.it/products/icub-robot>, July 20, 2021
- 55 Wiesmann G, Schraml S, Litzenberger M, Belbachir A N, Hofstätter M, Bartolozzi C. Event-driven embodied system for feature extraction and object recognition in robotic applications. In: Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, USA: IEEE, 2012. 76–82
- 56 Glover A, Bartolozzi C. Event-driven ball detection and gaze fixation in clutter. In: Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Daejeon, Korea (South): IEEE, 2016. 2203–2208
- 57 Vasco V, Glover A, Mueggler E, Scaramuzza D, Natale L, Bartolozzi C. Independent motion detection with event-driven cameras. In: Proceedings of the 18th International Conference on Advanced Robotics (ICAR). Hong Kong, China: IEEE, 2017. 530–536
- 58 Rea F, Metta G, Bartolozzi C. Event-driven visual attention for the humanoid robot iCub. *Frontiers in Neuroscience*, 2013, **7**: Article No. 234
- 59 Iacono M, D'Angelo G, Glover A, Tikhonoff V, Niebur E, Bartolozzi C. Proto-object based saliency for event-driven cameras. In: Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, China: IEEE, 2019. 805–812
- 60 Glover A, Bartolozzi C. Robust visual tracking with a freely-moving event camera. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Vancouver, Canada: IEEE, 2017. 3769–3776
- 61 Gibaldi A, Canessa A, Chessa M, Sabatini S P, Solari F. A neuromorphic control module for real-time vergence eye movements on the iCub robot head. In: Proceedings of the 11th IEEE-RAS International Conference on Humanoid Robots. Bled, Slovenia: IEEE, 2011. 543–550
- 62 Vasco V, Glover A, Tirupachuri Y, Solari F, Chessa M, Bartolozzi C. Vergence control with a neuromorphic iCub. In: Proceedings of the 16th IEEE-RAS International Conference on Humanoid Robots (Humanoids). Cancun, Mexico: IEEE, 2016. 732–738
- 63 Akolkar H, Valeiras D R, Benosman R, Bartolozzi C. Visual-Auditory saliency detection using event-driven visual sensors. In: Proceedings of the 2015 International Conference on Event-Based Control, Communication, and Signal Processing (EBCC-SP). Krakow, Poland: IEEE, 2015. 1–6
- 64 Zhao B, Ding R X, Chen S S, Linares-Barranco B, Tang H J. Feedforward categorization on AER motion events using cortex-like features in a spiking neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, **26**(9): 1963–1978
- 65 Zhao B, Yu Q, Yu H, Chen S S, Tang H J. A bio-inspired feed-forward system for categorization of AER motion events. In: Proceedings of the 2013 IEEE Biomedical Circuits and Systems Conference (BioCAS). Rotterdam, the Netherlands: IEEE, 2013. 9–12
- 66 Zhao B, Chen S S, Tang H J. Bio-inspired categorization using event-driven feature extraction and spike-based learning. In: Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN). Beijing, China: IEEE, 2014. 3845–3852
- 67 Nan Y, Xiao R, Gao S B, Yan R. An event-based hierarchy model for object recognition. In: Proceedings of the 2019 IEEE Symposium Series on Computational Intelligence (SSCI). Xiamen, China: IEEE, 2019. 2342–2347
- 68 Shen J R, Zhao Y, Liu J K, Wang Y M. Recognizing scoring in basketball game from AER sequence by spiking neural networks. In: Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN). Glasgow, UK: IEEE, 2020. 1–8
- 69 Massa R, Marchisio A, Martina M, Shafique M. An efficient spiking neural network for recognizing gestures with a DVS camera on the loihi neuromorphic processor. In: Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN). Glasgow, UK: IEEE, 2020. 1–9
- 70 Camuñas-Mesa L A, Domínguez-Cordero Y L, Linares-Barranco A, Serrano-Gotarredona T, Linares-Barranco B. A configurable event-driven convolutional node with rate saturation mechanism for modular ConvNet systems implementation. *Frontiers in Neuroscience*, 2018, **12**: Article No. 63
- 71 Camuñas-Mesa L A, Linares-Barranco B, Serrano-Gotarredona T. Low-power hardware implementation of SNN with decision block for recognition tasks. In: Proceedings of the 26th IEEE International Conference on Electronics, Circuits and Systems (ICECS). Genoa, Italy: IEEE, 2019. 73–76
- 72 Pérez-Carrasco J A, Zhao B, Serrano C, Acha B, Serrano-Gotarredona T, Chen S C, et al. Mapping from frame-driven to frame-free event-driven vision systems by low-rate rate coding and coincidence processing — Application to feedforward convnets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(11): 2706–2719
- 73 Diehl P U, Cook M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience*, 2015, **9**: Article No. 99
- 74 Iyer L R, Basu A. Unsupervised learning of event-based image recordings using spike-timing-dependent plasticity. In: Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN). Anchorage, USA: IEEE, 2017. 1840–1846
- 75 Cheng R, Mirza K B, Nikolic K. Neuromorphic robotic platform with visual input, processor and actuator, based on spiking neural networks. *Applied System Innovation*, 2020, **3**(2): Article No. 28
- 76 Ting J, Fang Y, Lele A, Raychowdhury A. Bio-inspired gait imitation of hexapod robot using event-based vision sensor and spiking neural network. In: Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN). Glasgow, UK: IEEE, 2020. 1–7
- 77 Youssef I, Mutlu M, Bayat B, Crespi A, Hauser S, Conradt J, et al. A neuro-inspired computational model for a visually guided robotic lamprey using frame and event based cameras. *IEEE Robotics and Automation Letters*, 2020, **5**(2): 2395–2402
- 78 Blum H, Dietmüller A, Milde M B, Conradt J, Indiveri G, Sandamirskaya Y. A neuromorphic controller for a robotic vehicle equipped with a dynamic vision sensor. In: Proceedings of the 13th Robotics: Science and Systems 2017. Cambridge, MA, USA, 2017. 1–9
- 79 Renner A, Evanusa M, Sandamirskaya Y. Event-based attention and tracking on neuromorphic hardware. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach, USA: IEEE, 2019. 1709–1716
- 80 Gehrig M, Shrestha S B, Mouritzen D, Scaramuzza D. Event-based angular velocity regression with spiking networks. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020. 4195–4202
- 81 Conradt J, Berner R, Cook M, Delbruck T. An embedded AER dynamic vision sensor for low-latency pole balancing. In: Proceedings of the 12th International Conference on Computer Vision Workshops. Kyoto, Japan: IEEE, 2009. 780–785
- 82 Moeyns D P, Corradi F, Kerr E, Vance P, Das G, Neil D, et al.

Steering a predator robot using a mixed frame/event-driven convolutional neural network. In: Proceedings of the 2016 Second International Conference on Event-Based Control, Communication, and Signal Processing (EBCCSP). Krakow, Poland: IEEE, 2016. 1–8

- 83 Delbruck T, Lang M. Robotic goalie with 3 ms reaction time at 4% CPU load using event-based dynamic vision sensor. *Frontiers in Neuroscience*, 2013, 7: Article No. 223
- 84 Delbruck T, Lichtsteiner P. Fast sensory motor control based on event-based hybrid neuromorphic-procedural system. In: Proceedings of the 2007 IEEE International Symposium on Circuits and Systems. New Orleans, USA: IEEE, 2007. 845–848
- 85 Mueller E, Censi A, Frazzoli E. Low-latency heading feedback control with neuromorphic vision sensors using efficient approximated incremental inference. In: Proceedings of the 54th IEEE Conference on Decision and Control (CDC). Osaka, Japan: IEEE, 2015. 992–999
- 86 Mueller E, Censi A, Frazzoli E. Efficient high speed signal estimation with neuromorphic vision sensors. In: Proceedings of the 2015 International Conference on Event-Based Control, Communication, and Signal Processing (EBCCSP). Krakow, Poland: IEEE, 2015. 1–8
- 87 Delbruck T, Pfeiffer M, Juston R, Orchard G, Müggler E, Linares-Barranco A, et al. Human vs. computer slot car racing using an event and frame-based DAVIS vision sensor. In: Proceedings of the 2015 IEEE International Symposium on Circuits and Systems (ISCAS). Lisbon, Portugal: IEEE, 2015. 2409–2412
- 88 Censi A. Efficient neuromorphic optomotor heading regulation. In: Proceedings of the 2015 American Control Conference (ACC). Chicago, USA: IEEE, 2015. 3854–3861
- 89 Singh P, Yong S Z, Gregoire J, Censi A, Frazzoli E. Stabilization of linear continuous-time systems using neuromorphic vision sensors. In: Proceedings of the 55th IEEE Conference on Decision and Control (CDC). Las Vegas, USA: IEEE, 2016. 3030–3036
- 90 Singh P, Yong S Z, Frazzoli E. Stabilization of stochastic linear continuous-time systems using noisy neuromorphic vision sensors. In: Proceedings of the 2017 American Control Conference (ACC). Seattle, USA: IEEE, 2017. 535–541
- 91 Singh P, Yong S Z, Frazzoli E. Regulation of linear systems using event-based detection sensors. *IEEE Transactions on Automatic Control*, 2019, 64(1): 373–380



栗 傑 首都师范大学信息工程学院讲师。分别于 2011 年和 2017 年获得北京航空航天大学学士和博士学位。2013 年 ~ 2015 年为在法国南巴黎大学联合培养的博士研究生。主要研究方向为类脑计算和机器人感知与控制。本文通信作者。

E-mail: li.su@cnu.edu.cn

(**SU Li** Lecturer at the College of Information Engineering, Capital Normal University. She received her bachelor degree and Ph.D. degree from Beihang University, in 2011 and 2017, respectively. From 2013 to 2015, she was a visiting Ph.D. student with the University of Paris-Sud, Orsay, France. Her research interest covers neuromorphic computing, perception and control of robot. Corresponding author of this paper.)



杨 帆 首都师范大学信息工程学院硕士研究生。主要研究方向为基于事件相机的目标检测。

E-mail: yangfannie@cnu.edu.cn

(**YANG Fan** Master student at the College of Information Engineering, Capital Normal University. Her

main research interest is object detection based on event camera.)



王向禹 首都师范大学信息工程学院硕士研究生。主要研究方向为基于事件相机的目标识别与跟踪。

E-mail: wangxiangyu@cnu.edu.cn

(**WANG Xiang-Yu** Master student at the College of Information Engineering, Capital Normal Uni-

versity. His research interest covers object recognition and tracking based on event camera.)

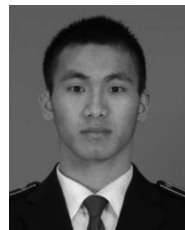


郭川东 北京理工大学宇航学院博士研究生。主要研究方向为复杂柔性航天器动力学、测量与控制。

E-mail: gnochuandong@bit.edu.cn

(**GUO Chuan-Dong** Ph.D. candidate at the College of Aerospace Engineering, Beijing Institute of Tech-

nology. His research interest covers attitude dynamics, measurement and control of complex flexible spacecraft.)

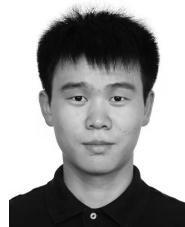


童良乐 首都师范大学信息工程学院硕士研究生。主要研究方向为类脑器件与电路设计。

E-mail: kyttl2021@163.com

(**TONG Liang-Le** Master student at the College of Information Engineering, Capital Normal University.

His research interest covers neuromorphic device and circuit design.)



胡 权 北京理工大学宇航学院副教授。分别于 2009 年和 2014 年获北京航空航天大学学士和博士学位。主要研究方向为多体系统动力学, 空间机器人感知与控制。

E-mail: huquan@bit.edu.cn

(**HU Quan** Associate professor at the College of Aerospace Engineering, Beijing Institute of Technology. He received his bachelor degree and Ph.D. degree from Beihang University in 2009, 2014, respectively. His research interest covers multibody dynamics, perception and control of space robot.)