

文章编号:1007-2780(2021)12-1664-10

事件视觉传感器发展现状与趋势

方应红^{1,2}, 徐伟^{1*}, 朴永杰¹, 冯汝鹏¹, 郑亮亮¹

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049)

摘要:基于事件视觉传感器是一种新型的仿生型视觉传感器,它更类似于人眼的工作机制使其广受关注。与基于帧的传统相机的工作机制和输出方式不同,基于事件视觉传感器的像素可以单独检测光照强度对数的变化,并在变化量超过一定阈值时输出包含位置、时间、极性的事件信息,拥有低延迟、高动态范围、低功耗的优点,其独特的输出方式和工作特性使其特别适应于有高速运动、光照条件变化较大或小能耗的场合。本文介绍了事件相机的发展历程、分类和工作原理、优缺点,以及事件相机在快速运动的跟踪与监测、目标识别、即时定位与地图构建(SLAM)等领域近些年的应用情况。最后总结了事件相机在不同应用领域仍存在的挑战,并展望其未来的发展。事件相机的广泛应用将为目前传统相机仍很棘手的高速运动和高动态范围场合提供新的解决方案,在未来的不断更新和发展下,它将能够在更多复杂的应用场景发挥作用。

关键词:事件相机;视觉传感器;基于事件的视觉传感器;原理和应用

中图分类号:TP391.4;TH691.9 **文献标识码:**A **doi:**10.37188/CJLCD.2021-0149

Development status and trend of event-based vision sensor

FANG Ying-hong^{1,2}, XU Wei^{1*}, PIAO Yong-jie¹, FENG Ru-peng¹, ZHENG Liang-liang¹(1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Event-based vision sensor is a new type of bionic vision sensor, which is more similar to the working mechanism of human eye, and makes it widely concerned. Different from the working mechanism and output mode of traditional frame-based cameras, the pixels of event-based vision sensors can individually detect logarithmic change of light intensity, and output event information including location, time, and polarity when the amount of change exceeds a certain threshold. The event-based vision sensor has the advantages of low latency, high dynamic range, and low power consumption. Its unique output mode and working characteristics make it especially suitable for occasions with high-speed movement, large changes in lighting conditions, or low energy consumption. This article introduces the development history, classification, working principle, advantages and disadvantages of event-based cameras, as well as their applications in the fields of rapid motion tracking and monitor-

收稿日期:2021-05-31;修订日期:2021-06-29.

基金项目:钱学森空间技术实验室创新工作站开放基金(No. GZZKFJJ2020003)

Supported by Qian Xuesen Laboratory of Space Technology, GAST(No. GZZKFJJ2020003)

*通信联系人, E-mail: xwciomp@126.com

ring, target recognition, simultaneous localization and mapping (SLAM) in recent years. Finally, it summarizes the challenges that the event-based camera still exists in different application fields, and looks forward to its future development. The wide application of event-based cameras can provide new solutions for high-speed motion and high dynamic range occasions where traditional cameras are still tricky. With continuous updates and development in the future, it will be able to play a role in more complex application scenarios.

Key words: event-based camera; vision sensor; event-based vision sensor; principle and application

1 引言

如今相机已经广泛地应用在工业、商业、军事、科研试验、太空探索、日常生活等领域。传统相机使用一种基于帧的视觉传感器,它通过预置一定的曝光时间,来逐帧获取场景信息。这种基于帧的视觉传感器有几个不足:第一,无法得到曝光时间内场景的变化信息,使得其在有快速运动的场景中应用受限;第二,由于每次曝光得到的图像包含场景所有内容,在研究运动物体时,那些没有变化的像素会产生较大的信息冗余,这会增加传输带宽、功耗、存储容量和信息处理的负担;第三,由于传统相机需要获取场景的绝对灰度值,当场景的光照强度过高或者过低时,传统相机会出现过曝光或者欠曝光的情况^[1]。

早在 20 世纪初,研究者们就致力于生物学视觉传感器的研究。研究发现,人眼获取的场景信息是异步地输出给大脑神经网络的,而且人眼对发生运动的物体有着更强的关注度。基于这些生物学现象,一种基于事件视觉传感器逐渐发展起来^[2-3]。这种基于事件视觉传感器(将使用这种传感器的相机称为事件相机)与基于帧的视觉传感器(将使用这种传感器的相机称为传统相机)有着本质的区别:事件相机的每个像素独立检测光照强度对数的变化,并当变化量达到一定阈值时,输出一个包含时间、位置、极性信息的事件。传感器中所有像素同时异步地检测光强对数变化情况并输出事件信息,形成了异步事件流。事件相机没有帧率的限制,而是主动检测场景中发生光强变化的位置并输出变化信息,减少了数据量,并拥有很高的动态范围,这很好地弥补了传统相机的不足。

自从基于事件的视觉传感器问世之后,它低延迟、高时间分辨率、高动态范围、低功耗的优点

吸引了许多研究学者的兴趣,事件相机的功能和性能也被一步步地改进和完善。2008 年,事件相机实现商业化,而鉴于可视化的需求,研究者们又开发了既能够输出事件又能够输出灰度信息的事件相机,这使得事件相机在一些应用场景,已经能够比肩传统相机。

2 发展历程和现状

基于事件的传感器的发展是受到生物学的启发,最初是以视网膜^[3]的形式被少量研究者采用。在 20 世纪 80 年代末,加州理工学院的 Carver Mead 引入了“神经形态”的概念来描述包含模拟和异步数字电子电路的系统,这一概念彻底改变了计算和神经生物学的前沿,引出了一门新的工程学科,其目标是设计和构建人工神经系统,如视觉系统、听觉处理器或自主漫游机器人,该领域被称为神经形态工程学^[4-7]。神经形态工程学的发展促进了一种新型的生物启发的视觉传感器的开发和实现。这种传感器中像素自主工作,通过产生异步数字脉冲(“尖峰”事件)响应光照强度的变化。这些脉冲通过异步仲裁数据总线(地址事件表示, AER^[8-10])触发包含相应像素阵列地址的数据包传输。2006 年, Lichtsteiner 等人提出了一个 128×128 的异步时间对比视觉传感器^[11]。2008 年, Posch 等人在时间对比视觉传感器的基础上,提出了时间对比度红外传感器^[12-13],这意味着动态视觉传感器(DVS)的问世,标志着基于事件的视觉传感器的研究进入到一个新的阶段,事件相机开始投入商用。

事件相机在各领域应用中的出色表现,吸引了一大批研究者,他们的研究又促进了事件相机的快速发展。由于 DVS 能够输出的场景数据太少,很难做到可视化,于是对能够输出灰度信息的事件相机的需求激励了新一代事件相机的问世。

Posch 等在 2011 年提出了 ATIS,它在 DVS 的结构上增加了一个曝光结构,让两个光感受器同时工作,使得在输出事件信息时能够根据光强变化同步地曝光,得到一些灰度信息。然而 ATIS 的曝光方式存在一些缺点,当光强变化较慢时,容易出现曝光异常的情况^[14]。Delbruck 等在 2013 年提出了 DAVIS,这是一种将传统 APS 相机与事件相机结合的产品,通过利用事件相机的光感受器来进行同步曝光,可以同时获得事件信息和场景灰度信息。这种相机虽然容易做到可视化,但它的 APS 部分保留了传统相机的缺陷,也没有充分利用事件相机的优点^[15-16]。Cho 等在 2014 年对生物视觉传感器的发展与应用的综述中,认为在需要快速视觉反馈的视觉跟踪和控制领域,考虑到速度和计算负荷,DVS 相机几乎可以替代传统的基于帧的相机。而在与识别相关的领域中,DAVIS 有着很好的应用前景^[17]。

硬件的研发到 DAVIS 已经发展到了比较成熟的阶段,而对于事件相机的应用算法和数据库的创建和更新,成了开发事件相机的主流。目前,比较成熟的算法主要应用在跟踪与监测、模式识别、SLAM 和光流估计等方面。

截止目前,事件相机的巨大潜力使之成为一个热门的技术和话题,许多机构都在致力于事件相机的开发和研究,较为出色的有法国 Prophesee、瑞士 iniVation、新加坡 CelePixel、韩国 Samsung 等。2017 年 6 月,首届基于事件的视觉国际研讨会在新加坡举行,这是事件相机诞生以来的第一次国际会议;2019 年 6 月,在美国加州长滩举行了第二届国际研讨会,这次研讨会由华为公司赞助,这表明华为也在从事事件相机领域的研究;2021 年 6 月,举行了第三届研讨会。通过这些会议,募集到了许多关于事件相机的资源,包括在各个应用领域的各种计算机算法以及相应的数据集和模拟器、处理器和硬件平台、各国研究人员的论文等,极大地促进了事件相机的发展。

3 硬件分类和原理

目前被广泛应用的事件相机可大致分成 3 类:动态视觉传感器(Dynamic Vision Sensor, DVS),基于异步时间的图像传感器(Asynchronous Time-

based Image Sensor, ATIS),动态主动像素视觉传感器(Dynamic and Active Pixel Vision Sensor, DAVIS)^[2,18-21]。

3.1 DVS

DVS 是最基本的也是最先发展的一种事件相机。它的每个像素单独检测环境中的光照强度,并以光强对数的形式进行记录和比较。当检测到的光照强度的变化达到一定阈值时,便会输出一个包含时间、像素所在地址、亮度增加(ON)或减少(OFF)的事件。所有的像素同时异步地进行光强检测,并异步地持续输出这样的事件,形成了事件流。其电路结构图如图 1 所示。电路由对数光感受器(Photoreceptor)、差分电路(Differencing)和两个比较器(Comparators)组成。

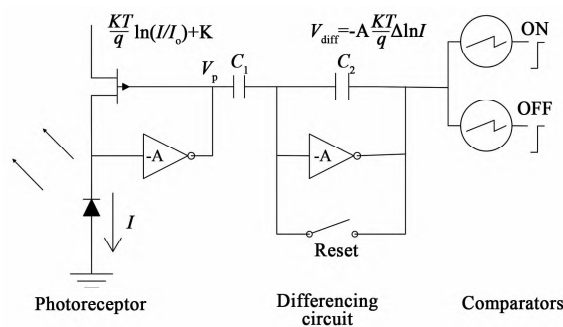


图 1 DVS 像素电路结构^[17]

Fig.1 Circuit structure of DVS pixel^[17]

通常,DVS 输出的事件 e_k 包含该事件发生的像素地址、时间和事件的类型(极性)。即 $e_k = (x_k, t_k, p_k)$,其中 $x_k = (x_k, y_k)^T$,表示像素地址; t_k 表示该事件发生的时间; $p_k \in \{+1, -1\}$ 表示事件的极性, $p_k = +1$ 为变亮事件, $p_k = -1$ 为变暗事件。规定像素 x_k 在 t_k 时刻的光强对数表示为:

$$L(x_k, t_k) = \log(I(x_k, t_k)) \quad (1)$$

其中 I 表示光照强度。事件 e_k 被激发,意味着像素 x_k 上的光强对数与该像素上一次激发事件时的光强对数之差,达到了预先设置的阈值 $\pm C$,即:

$$\Delta L(x_k, t_k) = p_k C \quad (2)$$

其中:

$$\Delta L(x_k, t_k) = L(x_k, t_k) - L(x_k, t_k - \Delta t_k) \quad (3)$$

Δt_k 表示从像素 x_k 上次激发事件开始经过的时间。

上述的事件生成模型,是在没有噪声的理想情况下建立的。实际上,任何相机都会受到来自感光器的外部及内部电路引起的噪声的影响,DVS 也不例外。通过设置阈值 C 的大小,可以调节 DVS 对噪声的敏感度:阈值 C 越大,DVS 对噪声越不敏感,但是 DVS 捕捉到的事件也会相应地减少。阈值 C 越小,DVS 捕捉到的事件越多,但信噪比会严重降低。

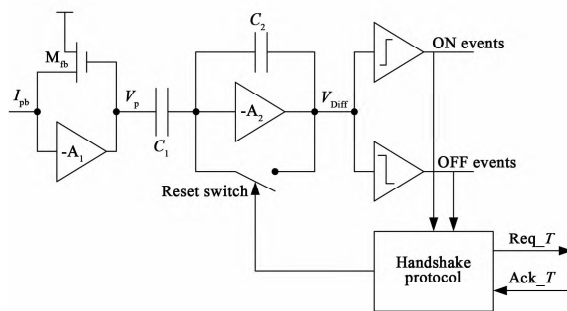
像素上亮度的变化可以由场景亮度的改变、场景中的物体或相机自身发生了移动造成。并且,当场景亮度变化越快、场景中的物体或相机自身移动越快时,单位时间内 DVS 产生的事件就越多。根据硬件和设计的不同,DVS 输出速率可高达 300 Meps(events per second,eps),延迟可低至 $3\ \mu\text{s}$,动态范围可达 130 dB。这样的特性让 DVS 在捕捉高速运动的物体、应对高动态范围的光照环境、低延迟控制等应用上有着巨大的优势。

3.2 ATIS

单纯的 DVS 输出的事件,只提供了事件像素地址、时间和极性,然而这些信息无法满足对可视化的需求,给后期数据处理带来了很大的困扰。能够在提供事件信息的同时还能提供一定灰度信息的需求,促使了 ATIS 的诞生。ATIS 像素结构^[14]分成两个部分(A 和 B),包含两个感光器,如图 2 所示。

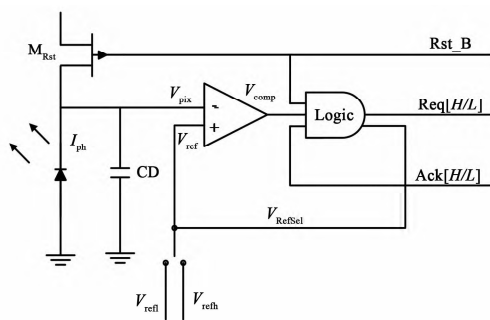
A 部分包含完整的 DVS 像素结构,它可以检测光强的变化并激发相应的事件;与此同时,B 部分包含的感光器检测光照强度的变化并进行曝光。ATIS 像素中 B 部分的曝光方式与传统 APS 相机的曝光方式不同。传统相机的曝光方式是预置曝光时间,通过测定曝光后电容两端电压的大小来确定像素的灰度值。因此,光照强度越大、电压差值越大,像素的灰度值越大。而 ATIS 像素中的 B 部分采用一种基于时间的曝光方式:在 A 子像素激发事件的同时,B 子像素内的一个电容被重置为高电平,随着 B 子像素持续受到光照,电容两端的电压下降,其两端电压从高电平下降到低电平所需的时间决定了该像素的灰度值。电压下降时间越短,说明该像素上光照强度越强,灰度值越高;下降时间越长,则说明灰度值越低。像素的灰度信息是以独立的事件流的形式输出的。

这种由事件触发的基于时间的曝光方式,使得 ATIS 相机在高动态范围的场景下也能获得较



(a) A 部分

(a) Part A



(b) B 部分

(b) Part B

图 2 ATIS 像素电路结构^[14]

Fig.2 Circuit structure of ATIS pixel^[14]

好的灰度信息(传统 APS 相机容易过曝或欠曝),并且只有产生事件的像素会输出灰度信息,主要的输出是场景中物体的边缘像素的灰度信息,减少了信息的时间冗余和空间冗余。然而 ATIS 相机也有一定的缺陷:在亮度过低的情况下曝光时间过长,当曝光时间超过下一次事件触发的时间,电容会被强行重置导致曝光异常终止,从而导致信息丢失。

3.3 DAVIS

DAVIS 相机将 DVS 相机和传统的 APS(有源像素传感器)相机结合起来,能够同时输出场景事件和灰度信息。图 3 为其电路结构图。

DAVIS 的像素结构^[17]也分成两个部分:APS 和 DVS。其中 DVS 检测光强的变化并输出异步的事件流信息,与此同时,APS 进行同步曝光获得场景的灰度信息。与 ATIS 不同的是,DAVIS 的两个部分共用一个感光器,这使得在体积上比 ATIS 小很多。通过 DAVIS 获取的灰度图像,与通过传统 APS 相机获取的灰度图像一样,包含那些光强变化较小或没有变化的像素,具

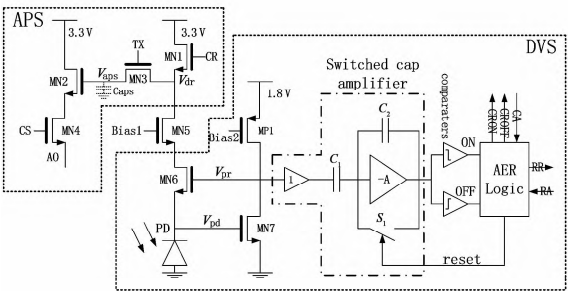


图 3 DAVIS 像素电路结构^[17]

Fig.3 Circuit structure of DAVIS pixel^[17]

有时间冗余性和空间冗余性,并且 APS 部分无法承受高动态范围的环境。

根据原理上的差异,表 1 中列出了这 3 种事件相机的特点和它们的一般用途。

表 1 3 种事件相机特点和用途的比较

Tab.1 Comparison of the characteristics and uses of these three event-based cameras

种类	特点	用途
DVS	仅输出事件信息,能够保持事件相机低延迟、低功耗、高动态范围的优点,但无法输出任何灰度信息,可视性差。	一般用于对可视化要求低、对延迟要求高的场合,如高速运动物体的跟踪、计数或运动监视、即时定位等。
ATIS	能够输出事件信息和发生光强变化像素的灰度信息,保持事件相机低延迟、较低功耗、高动态范围的优点,有一定可视性。但在环境亮度变化缓慢时容易出现曝光异常、信息丢失的问题。	一般用于环境亮度变化频繁且存在高速运动的场合,主要用于工业制造中产品的实时监测。
DAVIS	能够输出事件信息和环境灰度信息,将事件信息和灰度信息结合起来,时间分辨率高、可视性好,能够获得运动的细节信息,但它会受到 APS 相机缺陷的影响。	一般用于对可视化要求高、动态范围小的场合,如目标识别、目标检测、目标的跟踪与定位,特别是在机器人和无人驾驶领域应用较多。

3.4 事件相机的优缺点

从上述事件相机的原理能够发现,相比于传

统摄像机,事件相机有着高动态范围、低延迟、高时间分辨率、能耗低的优点,非常适合应用于高速运动的追踪和高动态范围的目标识别等场景。然而,作为一种新型的相机,事件相机有着和传统相机不同的输出信息,这就带来了算法和数据处理等方面的挑战^[22]。

(1)数据的处理:事件相机输出的事件仅包含少量与场景有关的信息(亮度的变化和位置),目前仅利用这些信息去还原场景或者提取场景中感兴趣的对象仍很困难,仍然需要获得足够多的灰度信息来支撑,但在未来的应用研究中,这些事件信息是否能够满足完成视觉任务仍有待研究。

(2)算法的更新:当前计算机视觉算法大多是基于图像帧,这些算法并不适用于事件相机,因此需要开发出新的适用于事件相机的各种应用算法,创建适用于事件相机的数据集和评价体系。

(3)噪声:由于事件相机是从像素的层面上进行事件输出,加上电路中差分电路的影响,使得噪声对事件相机的输出有较大的影响,如何设计算法能够从输出信息中滤除噪声而保留有用信息是一个亟需解决的问题。

4 主要应用

相比于事件相机,人们对传统相机的应用已经相当充分,传统相机能够提供足够多的目标信息特征,这使得其在视觉传感器的地位很难被动摇。然而传统相机存在的缺陷限制了一些特殊场合的应用,事件相机的独特性能可以为这些难题提供新的解决方案。针对事件相机的特

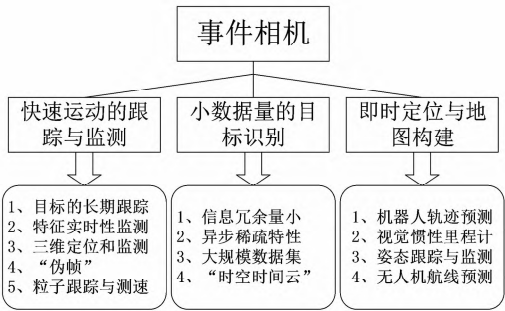


图 4 事件相机的应用

Fig.4 Application of event-based camera

点,在一些与之契合的视觉应用领域,比如快速运动的跟踪与监测、小数据量的目标识别、即时定位与地图构建(SLAM),已经有了较大的突破。下面介绍一些应用事件相机的研究,这些研究在应用事件相机的同时也在不断为事件相机的发展提供助力。图 4 对事件相机的应用进行了罗列。

4.1 快速运动的跟踪与监测

跟踪与监测是指使用视觉传感器来监测和定位场景中我们感兴趣的目标,我们希望这种监测能够不间断,定位能够足够精准。相比于传统相机,事件相机低延迟、高动态范围的优势使得它在目标跟踪领域有着很好的应用前景。

早期用事件相机进行跟踪只是单纯地跟踪对象,而且其背景还需要相对简单。Ramesh 等提出了一种用于事件摄像机的长期目标跟踪算法,跟踪框架通过在线学习来对对象使用区别性表示,并在对象返回视野时检测和重新跟踪对象。其中一个关键的创新是使用了基于事件的本地滑动窗口技术,该技术在背景混乱和有纹理的场景中表现可靠^[22]。

Tedaldi 等提出利用 DAVIS 提供的帧和事件数据来检测和跟踪视觉特征。首先在灰度帧中检测特征,然后使用事件流在帧之间的盲时间中异步跟踪特征。他们还提出了一种基于事件的算法,使用迭代的几何配准方法来跟踪特征,并且用 DAVIS 获得的真实数据评估了该方法的有效性^[23]。

Mitrokhin 等认为现代基于事件的视觉传感器空间分辨率极低,会产生大量噪声。针对这个问题,他们提出了一种新颖的事件流表示方式,它能够利用事件流的动态(时间)成分的信息。他们还基于此方式,提出了一种异步摄像机目标跟踪方法,用参数模型来近似事件流的 3D 几何形状,从而对摄像机进行运动补偿,然后在迭代过程中检测不符合模型的运动对象,来获得目标的运动从而实现跟踪。在实验中,根据可能出现的场景的性质,将记录分成 5 个序列,然后以成功率的形式定义评估指标,对来自 DAVIS 传感器的 RGB 帧对应的每个时间片计算单独的成功率,作为与 RGB 帧中可见的物体重叠至少 50% 的被检测物体的百分比。序列得分的平均值见表 2,可见该方法有一定的可行性^[24]。

表 2 数据集上传递途径的评估^[24]

Tab.2 Evaluation of the pipeline on the proposed dataset

序列	成功率/%
快速移动的无人机	92.78
多个物体	87.32
亮度变化	84.52
什么是背景	89.21
封闭序列	90.83

Boettiger 等将事件相机与传统相机就高动态范围检测和快速跟踪能力进行了实验对比,得出事件相机虽然存在很大优势,但仍然被长时间的数据处理所限制。他们提出了一种将基于帧和基于事件的传感器相结合的处理方法——“伪帧(Pseudo-frames)”。他们通过实验证明“伪帧”的处理效果是最好的,但“伪帧”没有真正充分利用事件相机低功耗、低数据、高速和高动态范围的特性,这表明在检测和跟踪领域,仍存在较大的研究空间^[25]。

事件相机还能用于微观粒子跟踪与测速方面。Wang 等提出了一种新的算法框架(图 5),

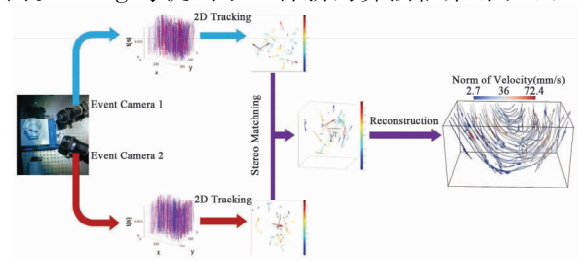


图 5 立体事件 PTV 框架的架构概述。两个事件摄像机捕捉流体内部粒子的运动。它们产生两个事件序列,这里用 $x-y-t$ 空间表示。2D 跟踪步骤为每个序列提供捕获粒子的 2D 速度。然后使用立体匹配步骤建立一个稀疏的三维速度场,用来估计密集的三维流体流动^[26]。

Fig.5 Overview of the architecture of our stereo-event PTV framework. The two event cameras capture the motion of the particles inside the fluid. They generate two sequence of events, represented here in the $x-y-t$ space. A 2D tracking step provide the 2D velocity of the captured particles for each sequence. Then, using a stereo matching step we build a sparse 3D velocity field in order to estimate the dense 3D fluid flow.

它是一种使用一对事件相机来检索流体速度场的密集 3D 测量。首先,跟踪两个事件序列内的粒子以估计它们在两个图像序列中的 2D 速度;然后,执行立体匹配步骤来检索它们的 3D 位置。这些中间输出被合并到一个优化框架中,该框架还包括物理上的正则化,以便检索 3D 速度场^[26]。

4.2 小数据量的目标识别

目标识别是指根据传感器收集到的视觉信息,从中提取感兴趣的特征信息,根据这些特征信息来判断该信息的种类或内容。目前比较流行的机器学习和深度学习的目标识别都是以基于帧的图像信息为输入内容。然而,图像信息中真正有用的信息占比很小,也就是说信息冗余量较大,这减慢了数据处理过程。事件相机输出的事件流信息和一些变化的边缘灰度信息(ATIS)已经足够用来进行一些简单的目标识别,开发出一种基于事件的目标识别的算法,可以提高在一些简单的场合的识别速率。

Damien 等将深度学习算法与基于事件的数据集成到人工框架中的研究中,通过实验发现,灰度图像中可能只有 35% 的信息可能会在训练中被用到,而通过事件相机得到的数据信息所达到的效果和这 35% 的信息几乎一致,这表明传统相机存在大量的数据冗余,而事件相机拥有巨大优势^[27]。

在目标识别和检测过程中,为了追求精度,很多基于事件的视觉算法在很大程度上依赖于使用深度神经网络同步批量处理事件,然而,这种方式丢弃了事件数据的稀疏性和异步性。Messikommer 等提出了一个通用框架,用于将基于同步图像类事件表示的模型转换为具有相同输出的异步模型,从而直接利用事件数据的内在异步和稀疏特性,大幅降低了高容量同步神经网络的计算复杂性和延迟,而不牺牲精度^[28]。

针对当时事件相机空间分辨率较低、缺乏大规模的训练数据集、缺乏完善的基于事件处理的深度学习架构等问题,Perot 等公开发布了第一个用于对象检测的高分辨率大规模数据集。另外,他们还引入了一种新的基于事件检测的递归体系结构和一种训练性能更好的时间一致性损失。值得注意的是,他们提出的方法不需要从事件重建任何强度图像,可以直接从原始事件信息来训练,这比通过中间强度图像更有效、更准确。

从表 3 的数据中可以看到,他们提出的方法在 1 M 像素数据集上达到了与 Gray-Retainet 相同的平均精度均值(mAP),这说明他们的事件相机检测器具有与常用帧相机检测器相当的精度。然而,这种方法仍然需要密集的事件表示,这意味着丢弃了事件信息的稀疏性^[29]。

表 3 两个汽车检测数据集的评价^[29]

Tab.3 Evaluation on the two automotive detection datasets

	1 M 像素检测数据集			第一代检测数据集	
	mAP	Runtime/ ms	Params/ M	mAP	Runtime/ ms
Events- RetinaNet	0.18	44.05	32.8	0.34	18.29
RED (ours)	0.43	39.33	24.1	0.40	16.70
Gray- RetinaNet	0.43	41.43	32.8	0.44	17.35

Wang 等为了处理事件摄像机的独特输出,将事件流视为一组三维时空点,即时空事件云。他们使用 PointNet 神经网络架构来分析事件云和进行手势识别的训练,通过实验证明该方法的识别率高达 97.08%,从而证明了他们提出的事件云概念是一种对事件流的有效表示^[30]。

4.3 即时定位与地图构建

即时定位与地图构建(SLAM)是指通过移动来获取位置信息进行自身定位并据此来构建增量式地图,从而实现自主定位和导航,主要用于机器人、无人机等领域。

Monforte 等在研究机器人的轨迹预测时,使用事件摄像机来检测快速移动的目标,他们使用基于长短期记忆(LSTM)架构的机器学习的方法来获取高度非线性的轨迹,研究了 LSTM 模型如何适应事件相机数据,并通过实验分析证明,异步空间采样优于固定速率采样,LSTM 与事件摄像机相结合,是一种有前途的学习方法^[31]。

Zhu 等提出了一种基于事件的视觉惯性里程计算法,根据从里程计滤波器获得的流量和旋转估计,产生鲁棒的特征轨迹,再将特征轨迹与惯性测量单元的输出融合,来获得全 6 自由度姿态的

精确度量跟踪。通过数据集测试,他们得出,与基于图像的技术相比,基于事件的特征跟踪通常能够在更长的时间内跟踪更多的特征^[32]。

为了获得更敏捷的机器人,Censi 等人提出了一种使用 DVS 和能主动高频(大于 1 000 Hz)闪烁的发光二极管标记器(ALMs)进行低延迟姿态跟踪的方法。DVS 可以检测不同频率的闪烁进行身份标记,然后将不同的标记放到一个“身份地图”中获得不同的姿态信息。他们的算法丢失追踪时间平均为 (0.35 ± 0.10) s,而 PTAM 丢失追踪的时间平均为 (0.80 ± 0.33) s,比翻转一次的时间 (0.56 ± 0.15) s 还要长。实验表明,一旦发光二极管再次闪烁,DVS 能够以可忽略的延迟重新获得稳定的跟踪,而不会遭受运动模糊的影响,DVS 在恢复轨迹丢失方面比 PTAM 方法更快。他们预测这项技术是实现飞行机器人高度自主机动的最有前途的候选技术^[33]。

除此之外,还有许多研究人员在原来的基础上开发新的效率更高的传感器和算法。其中 Jing 等提出了一个具有像素渲染机制和评估异步绝对强度能力的光流运动传感器和一种事件驱动跟踪算法,以实现高精度和实时性的跟踪和光流应用。与 ATIS 和 DAVIS 相比,其设计的传感器没有曝光时间,有助于高速物体的快速响应成像。除了检测照明变化之外,他所提出的运动传感器中的像素是互连的,并且通过像素渲染模块相互传递事件状态。每个活动像素及其 4 个相邻像素报告灰度事件,以在梯度提取中提供足够的数据,可以从根本上解决现有事件驱动的由事件稀疏性和事件强度缺乏引起的光流估计的准确性问题。其提出的算法利用在线自适应搜索区域来实现更精确的定位,运动传感器生成的事件包的灰度级强度被用来重建附加样本并随时间更新跟踪器模型。这种跟踪系统在跟踪静态背景下的单个或多个运动目标时具有优越的性能。然而,当背景和前景目标同时移动时,这种用于自适应搜索区域的事件引导方法将不适用^[34]。

5 结论与展望

事件相机有着非常突出的优点:低延迟、高时间分辨率、高动态范围、低功耗。然而在不同的应用领域,又有不同的缺陷阻碍着它的发展:在目标跟踪与监测领域,相机本身的运动和场景迅速变化,使得事件生成速率大大提高,特别是高速运动的物体,这给数据处理带来了巨大的压力,容易造成数据混乱和数据缺失;在目标识别领域,获取特征往往需要目标的灰度信息甚至彩色信息,对可视化的要求较高,而事件相机输出的事件包能保留的信息太少,无法做到可视化,尽管 DAVIS 能够获得一定的灰度信息,但它还是会受到传统相机缺点的影响,设计出一款能够将事件和灰度信息结合而不受到传统相机限制的传感器仍是一个困难的工作;在机器人、无人驾驶等领域,需要很好的可靠性,还需要强大的控制和处理能力,事件相机的输出数据的形式使之应用很受限制,复杂的场景的快速变化会输出大量事件,如何从这些事件中提取出有用的信息并加以控制和决策,是这个领域亟需解决的问题;噪声的处理是事件相机非常重要的一个问题,由于事件的输出形式的特性和差分器的作用,使事件相机的噪声的影响很大,这是限制事件相机发展的一个重要因素。

在诸多应用中,事件相机低带宽通信、低重量、低功率和高速的特点,使其充分满足航空航天领域对传感器的苛刻要求。目前,事件相机在航空航天领域的研究仍较少,显然,在这方面事件相机还有很大的发展潜力。如果将事件相机应用在导弹、航天飞机、超快速战斗机、海上军舰等军事目标的追踪上,由于这些目标有着显著的外观特征,而且它们的运动速度一般很快,或许可以获得比传统遥感相机更好的效果。另外,将事件相机应用于天基空间态势感知,也为监测卫星和星体提供了新的解决方案。未来,在航空航天和军事领域上,事件相机也将会有更加迅猛的发展。

参 考 文 献:

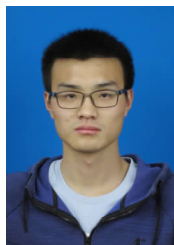
[1] 孔德磊,方正.基于事件的视觉传感器及其应用综述[J].信息与控制,2021,50(1):1-19.

KONG D L, FANG Z. A review of event-based vision sensors and their applications [J]. *Information and Control*,

- 2021, 50(1): 1-19. (in Chinese)
- [2] POSCH C, SERRANO-GOTARREDONA T, LINARES-BARRANCO B, *et al.* Retinomorphic event-based vision sensors: bioinspired cameras with spiking output [J]. *Proceedings of the IEEE*, 2014, 102(10): 1470-1484.
 - [3] CRISTÓBAL G, PERRINET L, KEIL M S. *Biologically Inspired Computer Vision* [M]. Berlin: Wiley-VCH, 2015.
 - [4] LICHTSTEINER P, DELBRUCK T. A 64×64 AER logarithmic temporal derivative silicon retina [C]//*Research in Microelectronics and Electronics*, 2005 *PhD*. Lausanne: IEEE, 2005: 202-205.
 - [5] MEAD C. *Analog VLSI and Neural Systems* [M]. Reading, Mass.: Addison-Wesley, 1989.
 - [6] MEAD C. Neuromorphic electronic systems [J]. *Proceedings of the IEEE*, 1990, 78(10): 1629-1636.
 - [7] MAHER M A C, DEWEERTH S P, MAHOWALD M A, *et al.* Implementing neural architectures using analog VLSI circuits [J]. *IEEE Transactions on Circuits and Systems*, 1989, 36(5): 643-652.
 - [8] POSCH C. Bio-inspired vision [J]. *Journal of Instrumentation*, 2012, 7: C01054.
 - [9] BOAHEN K A. Communicating neuronal ensembles between neuromorphic chips [M]//LANDE T S. *Neuromorphic Systems Engineering*. Boston: Springer, 1998: 229-259.
 - [10] BOAHEN K A. Point-to-point connectivity between neuromorphic chips using address events [J]. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 2000, 47(5): 416-434.
 - [11] LICHTSTEINER P, POSCH C, DELBRUCK T. A 128×128 120 dB 30 mW asynchronous vision sensor that responds to relative intensity change [C]//2006 *IEEE International Solid State Circuits Conference -Digest of Technical Papers*. San Francisco: IEEE, 2006: 2060-2069.
 - [12] LICHTSTEINER P, POSCH C, DELBRUCK T. A 128×128 120 dB 15 μ s latency asynchronous temporal contrast vision sensor [J]. *IEEE Journal of Solid-State Circuits*, 2008, 43(2): 566-576.
 - [13] POSCH C, WOHLGENANNT R, MATOLIN D, *et al.* A temporal contrast IR vision sensor [C]//*Proceedings of SPIE 7100 Optical Design and Engineering III*. Glasgow: SPIE, 2008: 71002A.
 - [14] 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 [J]. *IEEE Journal of Solid-State Circuits*, 2011, 46(1): 259-275.
 - [15] BERNER R, BRANDLI C, YANG M H, *et al.* A 240×180 10 mW 12 μ s latency sparse-output vision sensor for mobile applications [C]//2013 *Symposium on VLSI Circuits*. Kyoto: IEEE, 2013: C186-C187.
 - [16] BRANDLI C, BERNER R, YANG M H, *et al.* A 240×180 130 dB 3 s latency global shutter spatiotemporal vision sensor [J]. *IEEE Journal of Solid-State Circuits*, 2014, 49(10): 2333-2341.
 - [17] CHO D D, LEE T J. A review of bioinspired vision sensors and their applications [J]. *Sensors and Materials*, 2015, 27(6): 447-463.
 - [18] 马艳阳, 叶梓豪, 刘坤华, 等. 基于事件相机的定位与建图算法: 综述[J]. 自动化学报, 2021, 47(7): 1484-1494.
MA Y Y, YE Z H, LIU K H, *et al.* Event-based visual localization and mapping algorithms: a survey [J]. *Acta Automatica Sinica*, 2021, 47(7): 1484-1494. (in Chinese)
 - [19] 闫昌达, 王霞, 左一凡, 等. 基于事件相机的可视化及降噪算法[J]. 北京航空航天大学学报, 2021, 47(2): 342-350.
YAN C D, WANG X, ZUO Y F, *et al.* Visualization and noise reduction algorithm based on event camera [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(2): 342-350. (in Chinese)
 - [20] 周小龙, 刘倩倩, 产思贤, 等. 基于事件相机的视觉跟踪算法综述[J]. 小型微型计算机系统, 2020, 41(11): 2325-2332.
ZHOU X L, LIU Q Q, CHAN S X, *et al.* Event camera-based visual tracking algorithms: a survey [J]. *Journal of Chinese Computer Systems*, 2020, 41(11): 2325-2332. (in Chinese)
 - [21] 桑永胜, 李仁昊, 李耀仟, 等. 神经形态视觉传感器及其应用研究[J]. 物联网学报, 2019, 3(4): 63-71.
SANG Y S, LI R H, LI Y Q, *et al.* Research on neuromorphic vision sensor and its applications [J]. *Chinese Journal on Internet of Things*, 2019, 3(4): 63-71. (in Chinese)
 - [22] RAMESH B, ZHANG S H, LEE Z W, *et al.* Long-term object tracking with a moving event camera [C]//*BMVC*. Newcastle, 2018: 241.

- [23] TEDALDI D, GALLEGRO G, MUEGGLER E, *et al.* Feature detection and tracking with the dynamic and active-pixel vision sensor (DAVIS) [C]//2016 *Second International Conference on Event-based Control, Communication, and Signal Processing (EBCCSP)*. Krakow: IEEE, 2016: 1-7.
- [24] MITROKHIN A, FERMÜLLER C, PARAMESHWARA C, *et al.* Event-based moving object detection and tracking [C]//2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Madrid: IEEE, 2018: 1-9.
- [25] BOETTIGER J P. A comparative evaluation of the detection and tracking capability between novel event-based and conventional frame-based sensors [D]. Ohio: The Air Force Air University, 2020.
- [26] WANG Y H, IDOUGHI R, HEIDRICH W. Stereo event-based particle tracking velocimetry for 3D fluid flow reconstruction [C]//*Proceedings of 16th European Conference on Computer Vision*. Glasgow: Springer, 2020: 36-53.
- [27] DAMIEN J, HUBERT K, FREDERIC C. Convolutional neural network for detection and classification with event-based data [C]//*Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. Prague: VISAPP, 2019: 200-208.
- [28] MESSIKOMMER N, GEHRIG D, LOQUERCIO A, *et al.* Event-based asynchronous sparse convolutional networks [C]//*Proceedings of 16th European Conference on Computer Vision*. Glasgow: Springer, 2020: 415-431.
- [29] PEROT E, DE TOURNEMIRE P, NITTI D, *et al.* Learning to detect objects with a 1 Megapixel event camera [J]. *arXiv preprint arXiv:2009.13436*, 2020.
- [30] WANG Q Y, ZHANG Y X, YUAN J S, *et al.* Space-time event clouds for gesture recognition: from RGB cameras to event cameras [C]//2019 *IEEE Winter Conference on Applications of Computer Vision (WACV)*. Waikoloa: IEEE, 2019: 1826-1835.
- [31] MONFORTE M, ARRIANDIAGA A, GLOVER A, *et al.* Exploiting event cameras for spatio-temporal prediction of fast-changing trajectories [C]//2020 *2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*. Genova: IEEE, 2020: 108-112.
- [32] ZHU A Z, ATANASOV N, DANIILIDIS K. Event-based visual inertial odometry [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017: 5816-5824.
- [33] CENSI A, STRUBEL J, BRANDLI C, *et al.* Low-latency localization by active LED markers tracking using a dynamic vision sensor [C]//2013 *IEEE/RSJ International Conference on Intelligent Robots and Systems*. Tokyo: IEEE, 2013: 891-898.
- [34] HUANG J. Asynchronous high-speed feature extraction image sensor [D]. Singapore: Nanyang Technological University, 2018.

作者简介:



方应红(1997—),男,河南信阳人,硕士研究生,2020年于山东大学获得学士学位,主要从事智能成像方向的研究。
E-mail:941869192@qq.com



徐 伟(1981—),男,黑龙江大庆人,博士,研究员,2008年于中国科学院长春光学精密机械与物理研究所获得博士学位,主要从事新型高分辨率空间有效载荷总体设计等方面的研究。E-mail:xwciomp@126.com