

Нейроморфные системы машинного зрения

А. В. Теплюк¹

Нейроморфные системы компьютерного зрения, также называемые событийными камерами, представляют собой сенсоры, основной принцип работы которых был заимствован из физиологического принципа работы человеческого зрения. Они отличаются от традиционных камер тем, что вместо того, чтобы выполнять снимки через фиксированные интервалы времени, они асинхронно детектируют изменения яркости каждой точки пространства, попадающего в объектив камеры. На выходе эти сенсоры генерируют поток пакетов с данными, где каждому пакету соответствует событие изменения яркости, а сам пакет содержит информацию о времени возникновения события, координатах пикселя и полярности изменения яркости (увеличение либо уменьшение). В статье приводится описание принципов работы событийных камер. Дается обзор их основных преимуществ, а также наиболее перспективных направлений для применения данной технологии в различных областях деятельности человека.

Ключевые слова: Событийные камеры, машинное зрение, нейроморфные системы, компьютерное зрение, асинхронные сенсоры, широкий динамический диапазон, робототехника.

1. Введение

Одна из сложнейших задач, над которой работают многие ученые, заключается в исследовании понимания механизмов работы мозга и воспроизведения этих принципов в вычислительных устройствах. Зрение же является основным органом чувств, с помощью которого человек, используя свой мозг, познает окружающее его пространство. Технология событийных камер изначально задумывалась, как попытка воспроизвести «в железе» биологические принципы работы человеческого зрения, с целью её дальнейшего практического использования в широком спектре различных задач, связанных с компьютерной обработкой визуальной информации.

Событийные камеры являются асинхронными сенсорами, позволяющими детектировать непрерывное динамическое изменение состояния

¹ Теплюк Александр Викторович — Аспирант, Сколковский институт науки и технологий, e-mail: Alexander.Teplyuk@skoltech.ru.

Teplyuk Aleksandr Viktorovich — PhD student, Skolkovo Institute of Science and Technology.

наблюдаемой сцены. Это кардинально отличает их от повсеместно используемых видео камер, которые способны фиксировать визуальное состояние сцены только в определенные равноудаленные промежутки времени.

Основными преимуществами событийных камер являются их высокая скорость работы, широкий динамический диапазон, позволяющий вести съемку в условиях крайне низкой или наоборот крайне высокой освещенности и низкое энергопотребление. Эти особенности данных камер привели в последнее время к повышенному интересу к ним со стороны компаний, занимающихся робототехникой и автономными устройствами.

2. Принцип работы событийной камеры

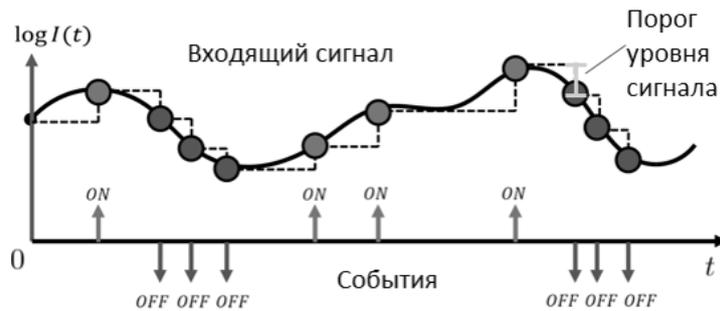


Рис. 1. Моменты генерации событий

Основной принцип работы событийной камеры (рис. 1) состоит в непрерывном отслеживании изменений яркости каждой точки наблюдаемой сцены и генерации события в тот момент, когда величина данного изменения превысит пороговое значение. Условие, при котором возникает генерация события описывается по формуле 1.

$$e = \{x, y, \check{p}, \check{t}\} : \exists e_{x,y}[j] \left(\check{p} = \begin{cases} 1 & \text{for } \int_{T[j-1]}^{\check{t}} C_{x,y}^t(t) + \dot{N}_{x,y}(t) dt \geq \theta^{ON} \\ 0 & \text{for } \int_{T[j-1]}^{\check{t}} C_{x,y}^t(t) + \dot{N}_{x,y}(t) dt \leq \theta^{OFF} \end{cases} \right)$$

Формула 1. Условие возникновения события. e – событие, N – шум, θ – пороговое значение, x, y – координаты пикселя, p – полярность, t – время, C – контрастность

Контрастность C в момент времени t рассчитывается по формуле 2.

$$C_{x,y}^t(t) = \frac{1}{I_{x,y}(t)} \frac{dI_{x,y}(t)}{dt} = \frac{d(\ln(I_{x,y}(t)))}{dt}$$

Формула 2. Контрастность. t – время, I – яркость

3. Способы обработки данных, сгенерированных событийной камерой

Одним из ключевых моментов, связанных с практическим использованием событийных камер, является способ извлечения полезной информации из потока данных, генерируемого камерой. Конкретный способ обработки потока выходных данных, а также выбор алгоритмов, напрямую зависит от контекста решаемой задачи. В зависимости от количества одновременно обрабатываемых событий, можно выделить две категории алгоритмов: а) основанные на последовательной обработке каждого события с последующим изменением системы (рис.2), что позволяет достичь максимально быстрой реакции. б) основанные на одновременной обработке целой группы событий, выделенных на определенном временном интервале (рис. 3). Различие в этих подходах достаточно условное, ввиду того, что одно событие не содержит достаточно информации для оценки характера изменения наблюдаемой сцены, и необходимо использовать данные из предыдущих событий.

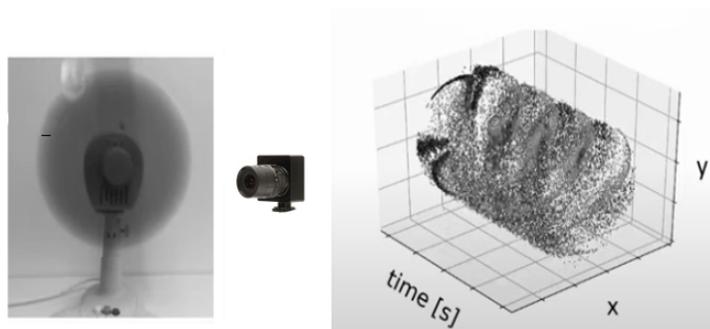


Рис. 2. Поток событий (справа), зафиксированных событийной камерой, направленной на вращающийся вентилятор (слева)



Рис. 3. Агрегация событий на определенном временном интервале (справа), зафиксированных событийной камерой, направленной на движущуюся руку (слева)

4. Практическое применение технологии

На сегодняшний день существует множество прикладных задач, для решения которых применяются событийные камеры. Это трекинг объектов [1], [2], распознавание жестов [3], визуальная вибродиагностика, подсчет быстродвижущихся мелких объектов на конвейерной ленте предприятия и др.

Список литературы

- [1] T. Delbruck and M. Lang, “Robotic goalie with 3ms reaction time at 4% CPU load using event-based dynamic vision sensor”, *Frontiers in Neuroscience*, **7** (2013), 223.
- [2] A. Glover and C. Bartolozzi, “Event-driven ball detection and gaze fixation in clutter”, *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [3] G. Orchard, C. Meyer, R. Etienne-Cummings, C. Posch, N. Thakor, and R. Benosman, “HFirst: A temporal approach to object recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**:10 (2015), 2028–2040.

Neuromorphic machine vision systems

Теплюк А.В.

Neuromorphic computer vision systems, also called event cameras, are sensors whose basic principle of operation was borrowed from the physiological principle of how human vision works. They differ from traditional cameras in that instead of taking pictures at fixed intervals, they asynchronously detects changes in the brightness of each point in the observed space. At the output, these sensors generate a stream

of data packets, where each packet corresponds to a brightness change event, and the packet itself contains information about time the event occurs, pixel coordinates and the nature of the brightness change (increase or decrease). The article describes principles of event cameras operations. Provides an overview of main advantages of event cameras, as well as the most promising areas for using this technology in various fields of human activity.

Keywords: Event cameras, machine vision, neuromorphic systems, computer vision, asynchronous sensors, high dynamic range, robots.

References

- [1] T. Delbruck and M. Lang, “Robotic goalie with 3ms reaction time at 4% CPU load using event-based dynamic vision sensor”, *Frontiers in Neuroscience*, **7** (2013), 223.
- [2] A. Glover and C. Bartolozzi, “Event-driven ball detection and gaze fixation in clutter”, *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [3] G. Orchard, C. Meyer, R. Etienne-Cummings, C. Posch, N. Thakor, and R. Benosman, “HFirst: A temporal approach to object recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**:10 (2015), 2028–2040.