Processing event data from a neuromorphic vision sensor

Marcin Kowalczyk

AGH University of Krakow

Department of Automatic Control and Robotics

kowalczyk@agh.edu.pl

Introduction

Event cameras (DVS) are neuromorphic sensors modelled on the human visual system. Unlike traditional cameras, they do not record full frames at fixed intervals, but only send information about pixels where a change in brightness has occurred. Each event contains pixel coordinates, polarisation (brighter/darker) and a timestamp. By recording only changes, DVS offers a number of advantages:

- very high time resolution (up to 1 us) allows tracking of fast movements,
- low latency (less than 100 us) results from the independent operation of each pixel,
- low data redundancy – only active pixels are processed, resulting in very low power consumption (approx. 50 mW),
- high dynamic range (>120 dB) enables operation in difficult lighting conditions thanks to the logarithmic sensitivity of photoreceptors.

A comparison of data from the event camera and the standard camera is shown in Figure 1. The upper graph shows standard video frames, while the lower graph shows the event stream from the DVS camera. Three situations are presented: slow object movement (left), no movement (centre) and fast object movement (right). The event camera does not generate any output when the object is not moving, but is more resistant to blurring during fast object movement.

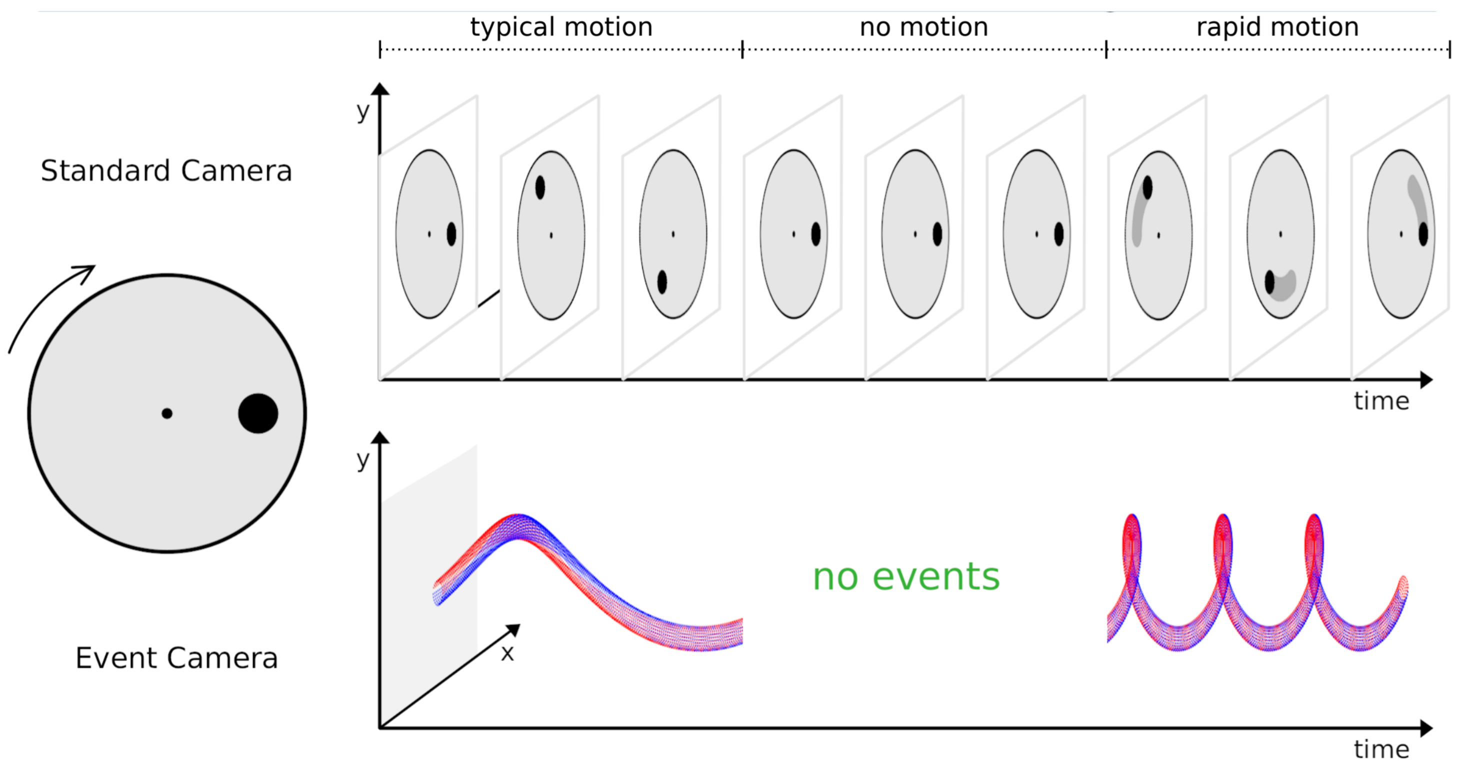

Fig. 1: Comparison of output data from a standard camera (top) and an event camera (bottom) [1].

Event cameras generate a data stream whose intensity depends on the movement in the observed scene. Resolution and bandwidth of these sensors are constantly increasing, placing ever greater demands on data processing systems. Traditional computing units, such as CPUs or GPUs, may have difficulty handling a large number of events in real time, especially with power constraints and low latency requirements. For this reason, FPGAs (Field Programmable Gate Arrays) or other dedicated computing architectures are a particularly attractive platform for processing data from DVS.

Tracking and counting objects

One of the basic applications of event cameras is tracking and counting moving objects. By recording only changes in brightness in the scene, DVS allows you to focus exclusively on dynamic elements, eliminating the need to process static backgrounds.

Tracking objects in DVS cameras requires a different approach than in classic frame cameras. The lack of complete frames means that typical methods based on object detection and re-identification cannot be directly applied. Instead, event segmentation can be used, which involves assigning events to individual objects based on their spatio-temporal consistency. Object counting involves detecting when objects cross a specific detection line (e.g. a virtual gate). Thanks to high temporal resolution, it is possible to count objects even when they are moving at high speeds and in difficult lighting conditions. Object counting usually requires tracking objects and determining the moment when they cross the detection line. The visualisation of tracking with object counting is shown in Figure 2.

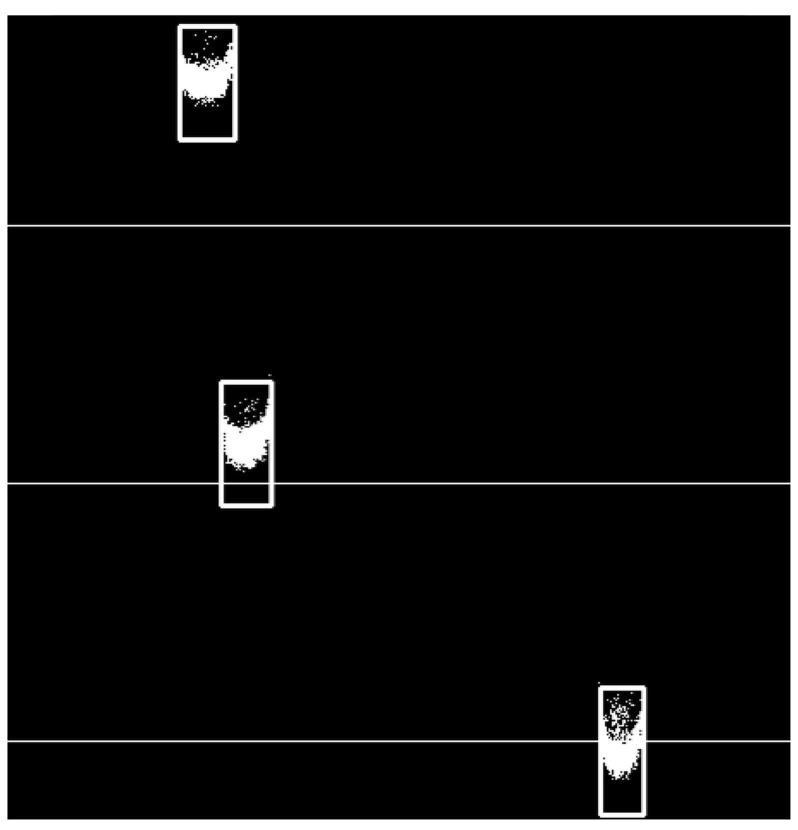

Fig. 2: Visualisation of an algorithm for counting falling corn grains [2].

Filtration

The presence of noise in data recorded by event sensors has a significant impact on the effectiveness of algorithms and the entire system. For example, it can significantly reduce the effectiveness of object classification or detection. For this reason, data filtering is used in virtually all systems that utilise sensor data, such as autonomous vehicles, driver assistance systems, security and monitoring systems, etc. In currently available classic vision sensors, interference usually takes the form of additive Gaussian noise. The situation is different in the case of event sensors, where interference takes the form of impulse noise. These are much more visible and can have a significant impact on the operation of a system.

In tasks where real-time system operation is required, filtering must be able to process a very large number of events in a short time and with as little delay as possible. In the conducted research, it was decided to divide the sensor matrix into square areas and use an independent IIR filter for each of them, which were responsible for determining the timestamp and event interval in a given area. They were then used to decide whether a given event should be rejected or forwarded for further processing. The hardware architecture was implemented and tested on the Xilinx Zynq Ultrascale+ MPSoC system. Several architectures were designed, which differed in terms of filtration efficiency and speed, achieving a throughput of up to 403 million events per second, with low resource utilisation and low power consumption. To the best of the authors' knowledge, the designed solutions provide the highest throughput for filtering data from vision event sensors among the solutions described in the scientific literature. The work on filtering has been described in the articles [3], [4], [5] and [6]. Figure 3 shows a simplified diagram of the event data filtering system.

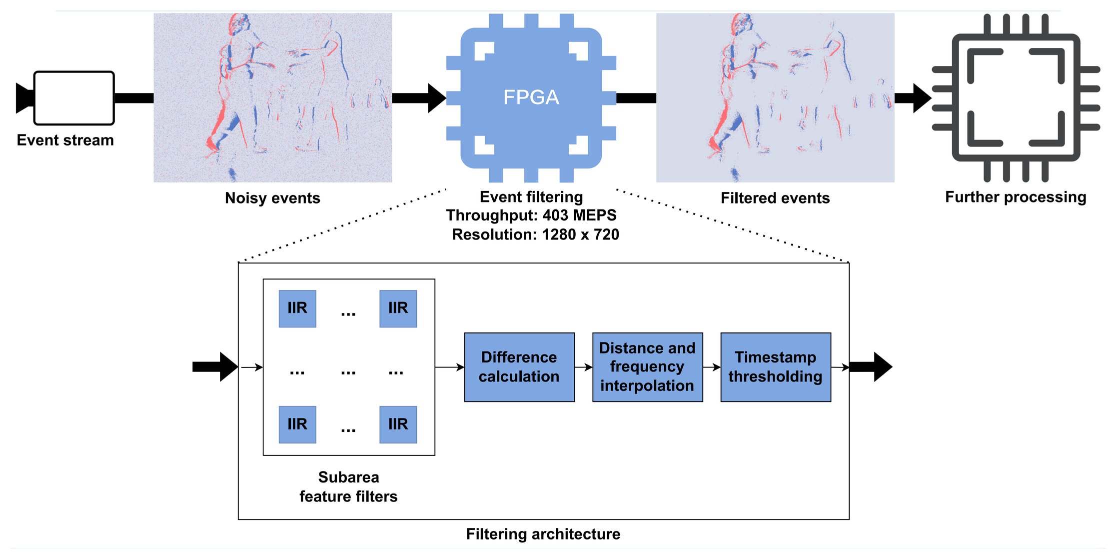

Fig. 3: Schematic diagram of the event data filtration system.

Driver monitoring system

Another area where the use of event cameras has great potential is driver monitoring systems (DMS). Their task is to analyse the driver's mental and physical condition in real time – detecting signs of fatigue, distraction, falling asleep, as well as monitoring the direction of gaze, head position and blink rate. Traditional DMS systems use RGB or infrared (NIR) cameras, but these are limited by their sensitivity to lighting conditions (e.g. sunlight, night-time, reflections). Due to their high resistance to lighting conditions, DVS can improve the effectiveness of these systems. Figure 4 shows an event frame depicting the driver's view in a car.

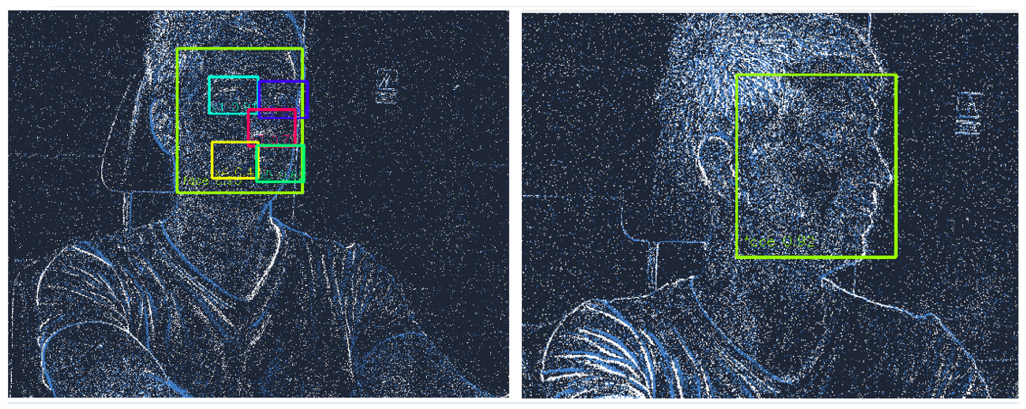

Fig. 4: An example event frame for DVS in the DMS system.

Bibliography

[1] Hanme Kim, Stefan Leutenegger, and Andrew J Davison. "Real-time 3D reconstruction and 6-DoF tracking with an event camera". In: *European Conference on Computer Vision*. Springer. 2016, pp. 349–364.

[2] Kamil Bialik et al. "Fast-moving object counting with an event camera". In: *arXiv preprint arXiv:2212.08384* (2022).

[3] Marcin Kowalczyk and Tomasz Kryjak. "Hardware architecture for high throughput event visual data filtering with matrix of IIR filters algorithm". In: *2022 25th Euromicro Conference on Digital System Design (DSD)*. IEEE. 2022, pp. 284–291.

[4] Marcin Kowalczyk and Tomasz Kryjak. "Interpolation-based event visual data filtering algorithms". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 4056–4064.

[5] Marcin Kowalczyk and Tomasz Kryjak. "High throughput event filtering: The interpolation-based DIF algorithm hardware architecture". In: *Microprocessors and Microsystems* (2025), p. 105171.

[6] Marcin Kowalczyk, Kamil Jeziorek, and Tomasz Kryjak. "Learning from Noise: Enhancing DNNs for Event-Based Vision through Controlled Noise Injection". In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. 2025, pp. 5092–5102.