DOI 10.5162/SMSI2025/B6.3

Motion Compensation of Event-Streams for Multitarget Tracking in Sensor-Based Sorting

Paul Bäcker [1] and Georg Maier[1]
[1] Fraunhofer IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany paul.baecker@iosb.fraunhofer.de

Summary: Event-based cameras excel in representing dynamic changes efficiently but can be overwhelmed in scenarios with strong global motion, such as conveyor belt systems. We propose a motion compensation method that filters events to focus on relative motion deviations. Our method significantly reduces the event-stream size while maintaining or even improving tracking accuracy. Our approach incorporates prior knowledge of global motion and demonstrates its efficacy in sensor-based sorting tasks.

Keywords: event-based vision, motion compensation, multitarget tracking, sensor-based sorting

Introduction

Event-based cameras efficiently capture dynamic scene changes by generating data only during motion or intensity changes, making them ideal for applications like multitarget tracking. However, in scenarios like sensor-based sorting on a conveyor belt with uniform motion, the constant motion generates a high volume of events, potentially overwhelming the system. To address this, we propose a method to filter event streams with a strong global motion prior, retaining only events corresponding to deviations from the expected motion.

Background, Motivation and Objective

Several works use contrast maximization techniques for compensation of scene motion in event-based vision [1]. Generally, these methods do not utilize a priori motion knowledge and therefore use computationally costly methods for estimating local scene motion from an event stream. [2] utilize a spiking neural networks for motion estimation and filtering, fully utilizing the asynchronous nature of event-based cameras. Although promising, spiking neural networks are still in their infancy, limiting their feasibility for widespread industrial application. [3] use on-sensor circuitry to extract object motion, by comparing local change-rates. However, some of their scene assumptions do not apply to our application. Further, to the best of our knowledge, the sensors are not commercially available. We propose a method, that efficiently filters events from commonly used event-based cameras by incorporating strong priors on global scene motion. The method is designed to be easily incorporated into tracking methods. We demonstrate our method on a multitarget tracking problem in sensor-based sorting, where we achieve equal tracking accuracy compared to a method using the full event-stream, while reducing the number of events by 70$\%$, reducing computational cost.

Method

We assume, that for a pure planar motion, every pixel of the event-based camera is expected to trigger the same amount of events. A property of the event-based sensor is, that on-events suffer from lower temporal jitter, than off-events. We therefore only consider on-events.

Consider $L(x,t)$ the radiance of a planar scene, we introduces $f(x,t)$ the motion compensated scene radiance.

$$L(x,t) = f(x + \boldsymbol{v}t, t) \tag{1}$$

relative to a planar transformation velocity $\boldsymbol{v}$ with $[v] = px/s$. A camera pointing towards a conveyor belt would capture radiance L, whereas a hypothetical camera moving along with the conveyor belt captures f. For a small temporal window δ_t we can assume a static scene $f(x)$. Using the event-generation equation from [1] we derive

$$\begin{aligned} pC &= L(x,t) - L(x,t_0) \\ &= f(x + \boldsymbol{v}t) - f(x + \boldsymbol{v}(t_0)). \end{aligned} \tag{2}$$

With p the polarity of the triggered event, $t_0 = t - \delta t$ the timestamp of the previously triggered event and C the contrast threshold of the event camera. Interpreting this, each event gives a measurement of the spatial change of a point in the scene, relative to the scene motion (e.g. the edge of a particle laying on top of a conveyor belt). However, since events suffer temporal jitter and only give binary information, we consider a spatio-temporal window around x, t for estimating the spatial gradient in ∇f relative to the speed v.

$$\nabla f(x,t)\boldsymbol{v} \propto \sum_{(x_i,y_i,t_i,p_i)\in\mathcal{E}} p_i\, \chi_x(x_i)\, \chi_y(y_i)\, \chi_t(t_i) \tag{3}$$

DOI 10.5162/SMSI2025/B6.3

With $\mathcal{E}$ the event-stream and $\chi_x(x_i)$, $\chi_y(y_i)$, $\chi_t(x_i)$ indicator functions for whether event i is within the spatio-temporal window around x, t.

For every timestep Δt, we update all estimators for $\nabla f \boldsymbol{v}$. We choose Δt to be $1px/|\boldsymbol{v}| \approx 0.25ms$. The temporal window for the estimation of $\nabla f \boldsymbol{v}$ is chosen as $\Delta T = 5 \cdot \Delta t$, the spatial window a grid of 5 pixels around x. After updating the estimators, we compare $\nabla f(x,t)\boldsymbol{v}$ to $\nabla f(x,t_0)\boldsymbol{v}$, with t_0 the timestamp of the last motion compensated event, we triggered. Our event-trigger function $g(t)$ is defined as

$$g(t) = \begin{cases} 1 & \text{if } \nabla f(x,t)\boldsymbol{v} - \nabla f(x,t_0)\boldsymbol{v} > C_\nabla, \\ -1 & \text{if } \nabla f(x,t)\boldsymbol{v} - \nabla f(x,t_0)\boldsymbol{v} < C_\nabla, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Where C_∇ is the threshold for triggering an event. With 0 no event triggered, else $g(t)$ the polarity of the event. Figure 1 shows an example output of the motion filter algorithm.

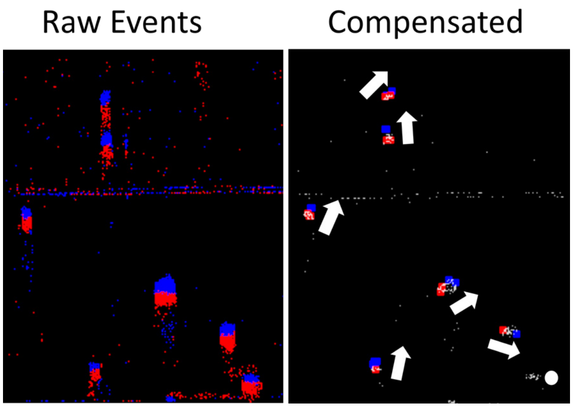

Fig. 1: Illustration of motion compensation. With index increasing proportional to the object speed, objects moving at the expected speed are projected to the same point in the compensated image (right). Using the relative position of the positive- and negative events, cues about object motion may be extracted, as indicated by the arrows.

Results

We apply the motion-filter to an existing event-stream, using different thresholds, resulting in different ratios of filtered events. In this work, we used an existing tracking approach using mean-shift tracking [4], with hyper parameters optimized using the full event-stream. We then applied different, motion-filtered event-streams without re-adjusting the tracking parameters. As shown in figure 2, we achieve a slight improvement of tracking accuracy, when reducing the number of events significantly, as only at very high filter ratios, the tracking accuracy is reduced. We further show that the filtered event-streams carry much more dense additional information, compared to the raw event-stream 1. This information could be used to further improve tracking accuracy.

Discussion and Future Work

Current method induces latency, since changes in event-frequency are only detected over time. Lower latency may be achieved by incorporating event-detection probabilities [5] or using more sophisticated changepoint detectors such as in [6]. The current implementation not optimized for realtime processing, but can be easily parallelized, since it applies operations globally.

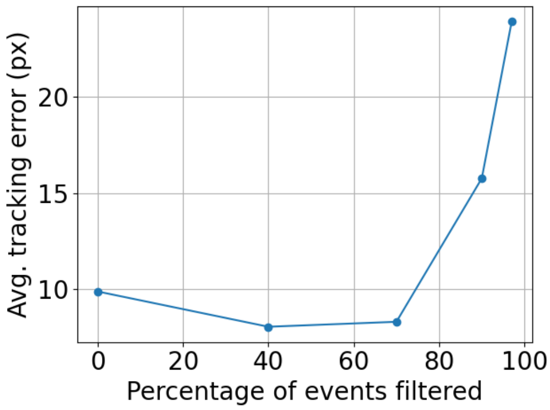

Fig. 2: Average tracking error, as given by distance to the ground-truth trajectory in pixels.

References

[1] T. Stoffregen, G. Gallego, T. Drummond, L. Kleeman, D. Scaramuzza, Event-Based Motion Segmentation by motion compensation, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 7243-7252 (2019); doi: 10.1109/ICCV.2019.00734.

[2] C. P. Rizzo, C. D. Schuman, J. S. Plank, Speed-based Filtration and DBSCAN of Event-based Camera Data with Neuromorphic Computing, *arXiv* (2024); doi: 10.48550/arXiv.2401.15212.

[3] S. Snyder, H. Thompson, M. A.-A. Kaiser, G. Schwartz, A. Jaiswal, M. Parsa, Object Motion Sensitivity: A Bio-inspired Solution to the Ego-motion Problem for Event-based Cameras, *arXiv* (2023); doi: 10.48550/arXiv.2303.14114.

[4] P. Bäcker, N. Terzer, G. Maier, U. D. Hanebeck, T. Längle, J. Beyerer, Fusion between Event-Based and Line-Scan Camera for Sensor Based Sorting, *10th Sensor-Based Sorting & Control 2024*, 89 (2024).

[5] A. Manakov, H. Herrman, B. Jähne, Performance comparison of area-scan and event-based camera, *Forum Bildverarbeitung 2024* (2024); doi: 10.5445/KSP/1000174496.

[6] V. Sundar, M. Dutson, A. Ardelean, C. Bruschini, E. Charbon, M. Gupta, Generalized Event Cameras, *arXiv* (2024); doi: 10.48550/arXiv.2407.02683.