Real-Time Intensity-Image Reconstruction for Event Cameras Using Manifold Regularisation

Christian Reinbacher[1]
reinbacher@icg.tugraz.at
Gottfried Graber[1]
graber@icg.tugraz.at
Thomas Pock[1,2]
pock@icg.tugraz.at

[1] Graz University of Technology
Institute for Computer Graphics and Vision
[2] Austrian Institute Of Technology
Vienna

(a) Raw Events (b) Reconstructed Image (c) Event Manifold

Figure 1: Sample results from our method. The image (a) shows the raw events and (b) is the result of our reconstruction. The time since the last event has happened for each pixel is depicted as a surface in (c) with the positive and negative events shown in green and red respectively.

Event cameras or neuromorphic cameras mimic the human perception system as they measure the per-pixel *intensity change* rather than the actual *intensity level*. In contrast to traditional cameras, such cameras capture new information about the scene at MHz frequency in the form of sparse events. The high temporal resolution comes at the cost of losing the familiar per-pixel intensity information.

In this work we aim to bridge the gap between the time-continuous domain of events and frame-based computer vision algorithms. We propose a simple method for simultaneous denoising and intensity reconstruction for neuromorphic cameras in real-time (see Fig. 1 for a sample output of our method). In contrast to very recent work on the same topic by Bardow *et al*. [1], we formulate our algorithm on an event-basis, avoiding the need to simultaneously estimate the optical flow. We cast the intensity reconstruction problem as an energy minimisation, where we model the camera noise in a data term based on the *generalised Kullback-Leibler divergence*. The optimisation problem is defined on a manifold induced by the timestamps of new events (see Fig. 1(c)). We show how to optimise this energy using variational methods and achieve real-time performance by implementing the energy minimisation on a graphics processing unit (GPU). We release software to provide live intensity image reconstruction to all users of DVS cameras[1]. We believe this will be a vital step towards a wider adoption of this kind of cameras.

Image Formation We have given a time sequence of events $(e^n)_{n=1}^N$ from a neuromorphic camera, where $e^n = \{x^n, y^n, \theta^n, t^n\}$ is a single event consisting of the pixel coordinates $(x^n, y^n) \in \Omega \subset \mathbb{R}^2$, the polarity $\theta^n \in \{-1, 1\}$ and a monotonically increasing timestamp t^n.

A positive θ^n indicates that at the corresponding pixel the intensity has increased by a certain threshold $\Delta^+ > 0$ in the log-intensity space. Vice versa, a negative θ^n indicates a drop in intensity by a second threshold $\Delta^- > 0$. We can now reconstruct an intensity image $u^n : \Omega \to \mathbb{R}_+$ by integrating the intensity changes indicated by the events over time.

Taking the $\exp(\cdot)$, the update in intensity space caused by one event e^n can be written as

$$f^n(x^n, y^n) = u^{n-1}(x^n, y^n) \cdot \begin{cases} c_1 & \text{if } \theta^n > 0 \\ c_2 & \text{if } \theta^n < 0 \end{cases}, \tag{1}$$

where $c_1 = \exp(\Delta^+)$, $c_2 = \exp(-\Delta^-)$. Starting from a known u^0 and assuming no noise, this integration procedure will reconstruct a perfect image (up to the radiometric discretisation caused by $\Delta^\pm$).

Image Reconstruction Since the events stem from real camera hardware, there is noise in the events. Also the initial intensity image u^0 is unknown and can not be reconstructed from events alone. Therefore the reconstruction of u^n from f^n can not be solved without imposing some regularity in the solution. We therefore formulate the intensity image reconstruction problem as the solution of the optimisation problem

$$u^n = \underset{u \in C^1(\Omega, \mathbb{R}_+)}{\operatorname{argmin}} \left[E(u) = D(u, f^n) + R(u) \right] , \tag{2}$$

where $D(u, f^n)$ is a *data term* that models the camera noise and $R(u)$ is a *regularisation term* that enforces some smoothness in the solution.

Moving edges in the image cause events once a change in logarithmic intensity is bigger than a threshold. The collection of all events $(e^n)_{n=1}^N$ can be recorded in a spatiotemporal volume $V \subset \Omega \times T$. V is very sparsely populated, which makes it infeasible to directly store it. As in [2], we observe that events lie on a lower-dimensional manifold within V, defined by the most recent timestamp for each pixel $(x, y) \in \Omega$. A visualisation of this manifold for a real-world scene can be seen in Fig. 1(c). Benosman *et al*. [2] fittingly call this manifold the *surface of active events*. We propose to incorporate the surface of active events into our method by formulating the optimisation *directly on the manifold*.

We chose the Total Variation as regularisation term $R(u)$ and the generalized Kullback-Leibler divergence as data term

$$D(u, f^n) := \lambda \int_\Omega (u - f^n \log u) \, \mathrm{d}x , \tag{3}$$

whose minimiser is known to be the correct ML-estimate under the assumption of Poisson-distributed noise between u and f^n [4]. In the paper we show how to formulate both on the *surface of active events* and how to efficiently solve the resulting optimisation problem using the Primal-Dual algorithm of Chambolle and Pock [3].

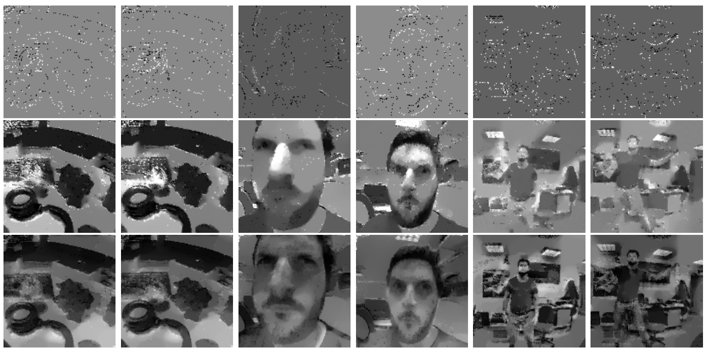

Figure 2: Comparison to the method of [1]. The first row shows the raw input events that have been used for both methods. The second row depicts the results of Bardow *et al*., and the last row shows our result. We can see that out method produces more details (e.g. face, beard) as well as more graceful gray value variations in untextured areas, where [1] tends to produce a single gray value.

[1] Patrick Bardow, Andrew Davison, and Stefan Leutenegger. Simultaneous optical flow and intensity estimation from an event camera. In *CVPR*, 2016.

[2] R. Benosman, C. Clercq, X. Lagorce, S. H. Ieng, and C. Bartolozzi. Event-based visual flow. *IEEE Transactions on Neural Networks and Learning Systems*, 25(2):407–417, 2014.

[3] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1), 2011.

[4] Triet Le, Rick Chartrand, and Thomas J. Asaki. A variational approach to reconstructing images corrupted by poisson noise. *J. Math. Imaging Vision*, 27:257–263, 2007.

[1] https://github.com/VLOGroup/dvs-reconstruction