This CVPR Workshop paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Live Demonstration: Joint Estimation of Optical Flow and Intensity Image from Event Sensors

Prasan Shedligeri Kaushik Mitra
Department of Electrical Engineering
Indian Institute of Technology Madras
{ee16d409, kmitra}@ee.iitm.ac.in

Abstract

This demonstration will show a deep learning based method to predict intensity images and optical flow from an event based sensor. The deep neural network is trained to take in as input a sequence of event frames and output a sequence of intensity images corresponding to the event frames and the optical flow between successive event frames. This setup will allow visitors to see that neural network based methods can be utilized for processing event sequence efficiently.

1. Introduction

Event based sensors asynchronously sense only the pixel-level brightness changes with microsecond resolution. These sensors provide promising advantages over traditional frame-based image sensors in challenging conditions such as high-speed motions, difficult-lighting and so on. Event sensors have shown promising results in the field of computer vision tasks like visual odometry/SLAM [5, 2, 1], 3D reconstruction [1] and so on. Unlike traditional image sensors, event sensors do not sample data at all the pixels at uniform time intervals and hence saves on sensor bandwidth requirement. However, this also poses a challenge while designing computer vision algorithms where one usually needs a grid-like representation.

In this demo we present a deep learning based algorithm to process events as a sequence of event frames. We show how standard deep learning algorithms can be made used to simultaneously reconstruct high frame rate intensity images and optical flow from event sensor data.

2. Demonstration Setup

The demonstration is fully functional and it consists of a Dynamic Vision Sensor (DVS) camera with a resolution of 128x128 [3] and a laptop. The DVS will be connected to the laptop through an USB cable. The deep learning based method will be run locally on the laptop itself and the data to be fed into the algorithm will be collected live from the venue. The visitors can themselves move the camera in various motions to generate the events. The neural network can predict intensity frames and optical flow for upto 150 event frames per second at the DVS resolution of 128×128

3. Visitor Experience

The visitors to the demonstration will see how the events generated from the DVS can be fed into a neural network based method and simultaneously predict the intensity image and the optical flow. Events which are to be fed into the algorithm can be generated by making various actions in front of the DVS. The visitors can also move the DVS to generate events and they will get to see simultaneously in real-time the generated events as frames, the predicted intensity images and the optical flow. In offline, they can also view the high frame rate reconstruction of the scene captured using DVS. The predicted intensity images will have a similar dynamic range as that of the DVS.

4. Prior Publications

At the time of writing, the manuscript with algorithmic details is being prepared for submission to BMVC 2019.

References

[1] Guillermo Gallego, Henri Rebecq, and Davide Scaramuzza. A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3867–3876, 2018.

[2] 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*, pages 349–364. Springer, 2016.

[3] Patrick Lichtsteiner, Christoph Posch, and Tobi Delbruck. A 128 $\times$128 120 db 15μ s latency asynchronous temporal

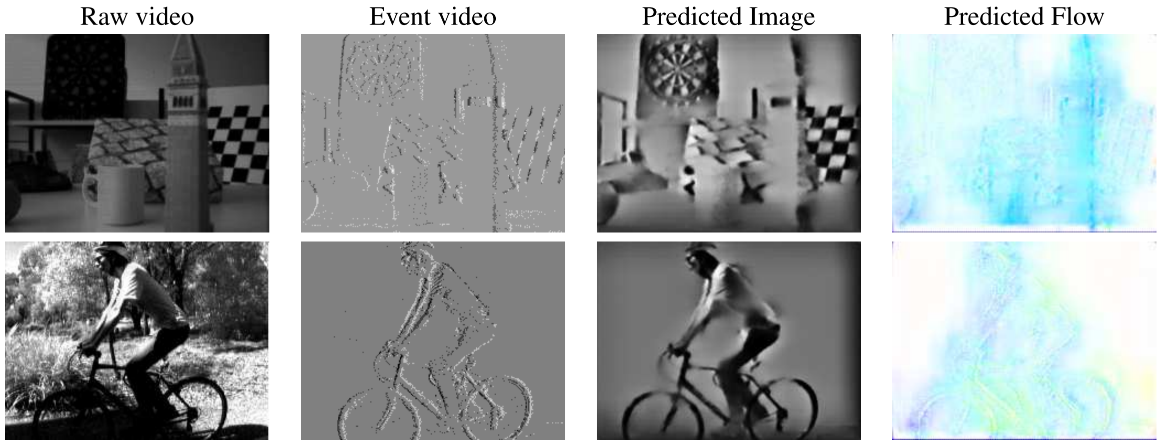

Figure 1. Qualitative reconstruction results from our method [7] shown on sequences from [4, 6]. We see that the reconstructed intensity images from our method are noise and artifact free. The predicted intensity images also retain the high-dynamic range property of the event sensors. *(If document is opened in Adobe Reader, videos can be viewed by clicking on the images).*

contrast vision sensor. *IEEE journal of solid-state circuits*, 43(2):566–576, 2008.

[4] Elias Mueggler, Henri Rebecq, Guillermo Gallego, Tobi Delbruck, and Davide Scaramuzza. The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and slam. *The International Journal of Robotics Research*, 36(2):142–149, 2017.

[5] Henri Rebecq, Timo Horstschäfer, Guillermo Gallego, and Davide Scaramuzza. Evo: A geometric approach to event-based 6-dof parallel tracking and mapping in real time. *IEEE Robotics and Automation Letters*, 2(2):593–600, 2017.

[6] Cedric Scheerlinck, Nick Barnes, and Robert Mahony. Continuous-time intensity estimation using event cameras. *arXiv preprint arXiv:1811.00386*, 2018.

[7] Prasan Shedligeri and Kaushik Mitra. Joint estimation of optical flow and intensity image for event sensors. *Manuscript in preparation for submission to BMVC 2019*.