VISUAL SLAM WITH AN EVENT CAMERA

Kim H., Davison A. - Imperial College London
{hanme.kim, a.davison}@imperial.ac.uk

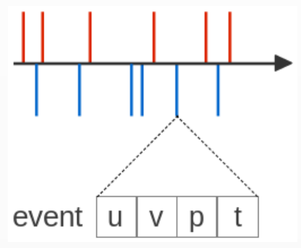
Imperial College London

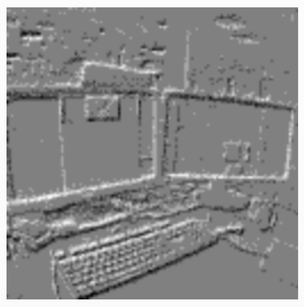
Abstract

An event camera is a silicon retina which outputs not video frames, but a stream of events indicating when individual pixels observe a log intensity change. We aim to develop a visual SLAM system with a single event-driven sensor which is to be no longer restricted by the same limitations imposed to standard cameras and inherently more efficient as it exploits compressed visual data, high temporal resolution, low latency and wide dynamic range provided by this type of sensor.

Event Camera

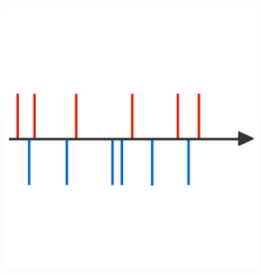
An event camera, such as the Dynamic Vision Sensor (DVS) [1], has no global shutter and does not capture image frames in the traditional sense. Instead, each pixel responds independently to changes in log intensity by generating asynchronous events, each with pixel location, sign and microsecond-precise timing. By encoding only image change, it offers the potential to transmit the information in a standard video but at vastly reduced bitrate, and with huge added advantages of very high dynamic range and temporal resolution.

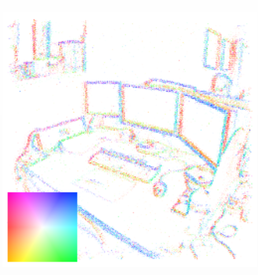




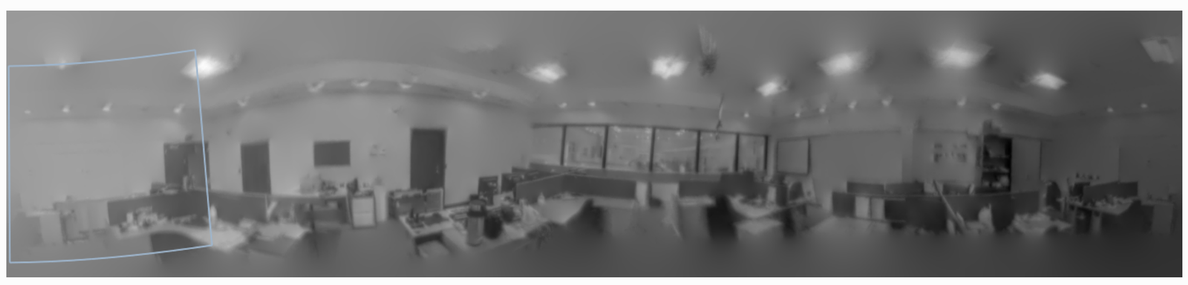
Scene Reconstruction and Tracking

We showed for the first time that an event stream, with no additional sensing, can be used to track camera rotation while building a persistent and high quality mosaic of a scene. Our method involves parallel camera rotation tracking and template reconstruction from estimated gradients, both operating on an event-by-event basis and based on probabilistic filtering.

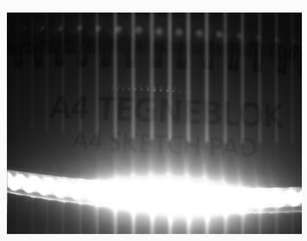




∗ Result 1: Spherical mosaic scene reconstruction.



∗ Result 2,3: High resolution and dynamic range reconstruction.



Towards 3D SLAM

The ultimate goal of this research is to develop a visual SLAM system following a similar approach in the style of [2] using a single event camera.

This system is to be no longer restricted by the limitations imposed to conventional cameras and inherently more efficient because it exploits compressed visual data, high temporal resolution, low latency and wide dynamic range provided by event-driven vision sensors. Additionally, its efficient strategy for tracking and reconstruction both operating on an event-by-event basis will allow the system to be able to track extremely fast motion and build a 3D dense map while it keeps its hardware complexity and power consumption low.

References

[1] P., Lichtsteiner, C., Posch, T., Delbruck, A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor, in *IEEE Journal of Solid-State Circuits*, 2008

[2] R. A., Newcombe, S., Lovegrove, A. J., Davison, DTAM: Dense Tracking and Mapping in Real-Time, in *Proceedings of the International Conference on Computer Vision (ICCV)*, 2011

Acknowledge

This research is supported by an EPSRC DTA studentship and Qualcomm Innovation Fellowship 2014 to Hanme Kim.