

A Real-time Dynamic Vision Sensor Emulator using Off-the-shelf Hardware

Garibaldi Pineda García

February 28, 2016

Abstract

Vision is one of our most important senses, a vast amount of information is perceived through our eyes. Neuroscientists have performed several studies using vision as input to their experiments. However, computational neuroscience has typically used Poisson-encoded images as spike-based visual sources. Recently neuromorphic Dynamic Vision Sensors have surfaced, while they have excellent capabilities, they remain scarce and difficult to use.

We propose a software-based visual input system, inspired by the behaviour of a DVS, but using a digital camera as a sensor. By using readily-available components, we believe, most scientist would have access to a spiking visual input source. While the primary goal was to use the system as a real-time input, it is also able to transcode well established images and video databases into spike train representations.

1 Introduction

DVS

Carver Mead, Zurich, Sevilla

Cameras are in almost every computing device.

Processing power today allows for these devices to perform calculations on the image stream to convert into a spike representation.

Other models not in real time or require high performance hardware or FPGAs.

In recent years the performance of computer processors has been advancing in smaller increments than it used to a few years ago. This is mainly because manufacturing technologies are reaching their limits. One way to improve performance is to use many processors in parallel, which has been successfully

applied to parallel-friendly applications like computer graphics. Task like pattern recognition are still a hard task for computers even with these technological advances.

Our brains are particularly good at learning and recognizing visual patterns (e.g. letters, dogs, houses, etc.). In order to achieve better performance for similar tasks on computers, scientists have looked into biology for inspiration. This has lead to the rise of brain-like (neuromorphic) hardware, which looks to mimic functional aspects of the nervous system. We can divide neuromorphic hardware into sensors (providing input) and computing devices (make use of information from sensors). Visual input has been traditionally obtained from images that are rate-encoded using Poisson processes, while this might be a biologically-plausible encoding in the first phase of a “visual pipeline” it is unlikely that eyes transmit as much information into later stages. Furthermore, if we think of it in terms of digital networks, having each pixel represented by a Poisson process incurs in high bandwidth requirements.

In 1998, Mead proposed a silicon retina consisting of individual photoreceptors and a resistor mesh that allowed nearby receptors to influence on the output of a pixel[1]. Later, researchers developed frame-free Dynamic Vision Sensors (DVSs) [2], [3]. They feature independent pixels that emit a signal when its intensity value changes above a certain threshold. These sensors have μ -second response time and excellent dynamic range properties, although they are still not as commercially available as regular cameras.

In this work, we propose to emulate the behaviour of a DVS using a conventional digital camera as a sensor. Basing the emulator on widely available hardware, would allow most computational neuroscientists to include video as a spike-based input.

2 Work

Katz, Nikolic, and Delbruck developed a DVS emulator in order to test behaviours for new sensor models. In their work, they transform the image provided a commercial camera into a spike stream at 125 frames per second (fps). They merged the emulator into their jAER framework and can be used there.

RGB images are gamma-encoded, to better utilize bits and be compliant with CRT monitors.

DVS pixels output the logarithm of their input current.

Asynchronous pixel behaviour is approximated via differencing the current and a reference frame.

Rate-based output is possible, though it may overwhelm communication channels.

We propose different time encodings (raw value/binary, threshold multiples/binary, threshold multiples/linear)

Extra behaviours: Local inhibition reduces bandwidth, persistence of stimuli over frames.

Adaptive threshold, slow-charging pixels by reducing the threshold.

Less parameters to configure.

Open source, and available in github.com

3 Results

Real-time DVS emulation on consumer hardware (Intel i5, 8GB ram) sPyN-Naker wrapper for videos and images.

4 Conclusion

GPU computing version would allow for higher resolutions.

Working on ganglion cell/Difference of Gaussian encoding.

References

- [1] C. Mead, *Analog VLSI Implementation of Neural Systems*, C. Mead and M. Ismail, Eds. Boston, MA: Springer US, 1998, ch. Adaptive Retina, pp. 239–246, ISBN: 978-1-4613-1639-8.
- [2] P. Lichtsteiner, C. Posch, and T. Delbruck, “A 128x128 120 db 15 us latency asynchronous temporal contrast vision sensor,” *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008, ISSN: 0018-9200. DOI: 10.1109/JSSC.2007.914337.
- [3] J. A. Lenero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, “A 3.6 us latency asynchronous frame-free event-driven dynamic-vision-sensor,” *IEEE Journal of Solid-State Circuits*, vol. 46, no. 6, pp. 1443–1455, 2011, ISSN: 0018-9200. DOI: 10.1109/JSSC.2011.2118490.
- [4] M. L. Katz, K. Nikolic, and T. Delbruck, “Live demonstration: Behavioural emulation of event-based vision sensors,” in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, 2012, pp. 736–740. DOI: 10.1109/ISCAS.2012.6272143.