Modernised Reduction: Adapting the ROT tree

Modernised Reduction: Adapting the ROT tree

Shane Harrigan, Sonya Coleman, Dermot Kerr, and Justin Quinn

Ulster University, Northern Ireland, BT487JL, UK
{sp.harrigan,sa.coleman,d.kerr,jp.quinn}@ulster.ac.uk

Abstract. Neuromorphic vision data offers a new means of evaluating digitise spatial and temporal representations of evolving scene dynamics. Reduction-Over-Time (ROT) trees have seen growing popularity as a medium for storing and operating over the temporally asynchronous data produced from neuromorphic sensors given their 1-D nature, spatial preservation, and speed. In this paper we propose a variation of the ROT tree called R-ROT which allows for greater adaptability within structure when compared to the originally proposed ROT tree using adaptive self-pruning. The R-ROT structure is evaluated against the original ROT model and is shown to achieve high accuracy results in shorter time across a widely popular benchmark database.

Keywords: smart sensors and actuators, smart factory, neuromorphic data

1 Introduction

The Reduction-Over-Time (ROT) tree, [1, 2], is an abstract data structure popularly applied to neuromorphic data, particularly the output of models based on silicon retinas. Neuromorphic vision sensors, and their systems [3–5], emulate the function of modelled biological visual systems by mimicking the structures and electrical behaviour of these systems in silicon. Neuromorphic sensing is a paradigm shift enabling a completely new avenue of approach to understand observable scenes; sensors such as the Dynamic Vision Sensor (DVS) model family [3] are growing increasingly popular in embedded, high-speed, and/or low-power systems for their compact size, speed increases, and power demand reduction. We contrast sensors such as those of DVS with the classical active-pixel sensor (APS) based approaches.

The classical APS approach is well understood and is the foundation of most imaging research to date but it has disadvantages in terms of time and energy since its formalisation. Classic APS sensors, or active-pixel sensors (APS), work by polling a 2-D array of pixels at a certain time interval (static or dynamic) and digitising the data values of all pixels during the polling to produce a numerical representation of the light levels detected at the polling moment. When displayed on a 2-D surface, these data values will contain spatial information at

Funded by Innovate UK through the Smart Manufacturing Data Hub (SMDH) project

the moment of capture which can be processed further by classical image processing techniques.
Taking a series of 2-D matrices formed from the event time at an instance in time, and representing in a sequence from oldest captured to newest, we can develop a representation of change within an observed scene. The power-to-speed ratio is directly linear such that in order to poll and digitise scene information at high-speeds we always need to increase the power consumption in the vision system; conversely to decrease the power consumption we need to decrease the polling and digitisation slowing the capturing process down. This ratio is acceptable in a number of existing areas but some key areas of research, such as autonomous vehicles, power and/or data conscious [6] systems, general embedded robotics, and energy-focussed manufacturing require more responsible energy usage while maintaining high-speeds.

Sensors, such as the DVS (including the hybrid DAVIS [7]) family, are emulations of retinal neural behaviour [**?**,7] which asynchronously release information on a pixel-by-pixel basis unlike APS sensors which operate on fixed cyclic emission rates . The core datatype of neuromorphic vision sensors is the event $e_i = \langle t_i, x_i, y_i, p_i \rangle$ where t is the event emission time, x and y are identifiers of the emitting pixel, $p = \pm 1$ indicates the polarity of light intensity change and i is the index of the event within $E = \{e_1, \dots, e_i\}$.
The ROT tree offers a fast means of processing neuromorphic data with observed enhanced noise reduction. The ROT tree is a paradigm shifting structure in terms of neuromorphic data processing which can operate in 1-D space as opposed to the 2-D approaches common currently in the field [8, 9, 5]. An ROT tree uses a event-time differences reduction model, which models information decay over time, to perform self-pruning) while maintaining the self-balancing nature of the tree for read operations. The ROT tree has been applied to neuromorphic data because:

1. ROT trees are well-suited for handling high-dimensional, continuous data because they can perform data reduction, as the data streams in, without requiring large amounts of memory or computing power.
2. ROT trees can efficiently [2] extract relevant information from the neuromorphic data while preserving its temporal and spatial structure. This is important because neuromorphic data often has complex temporal and spatial patterns that traditional data reduction techniques may not be able to capture.
3. ROT trees can be used to detect and track changes in the neuromorphic data over time. This is important because neuromorphic data often contains dynamic, evolving patterns that may be difficult to detect using other methods.

In this paper, we will show that ROT trees can be easily modified and adapted to suit specific types of neuromorphic data. We achieve this by comparing the ROT tree [2] which is based on a fixed data reduction model with an ROT variation which uses an adaptive data reduction model based on data temporality. We will refer to the ROT variation within this paper as R-ROT.

2 Methodology

The ROT tree is a self-pruning temporal device for storing neuromorphic data based on relevance over time, historically the data structure has made use of a forgetting curve designed to model the information retention observed across many psychological studies. This approach has proved to be efficient in terms of time and memory but recent implementations of ROT have favoured the use of fixed-thresholding to evaluate temporal information as it develops over time. Spatial ROT trees [1] have shown little consequences of storing spatial information over time but temporal information is based on an all-or-nothing decay model. Instead of organising ROT trees purely based on data reduction through temporal delta (the difference in time against events) we adjust for the natural variance of data. Consider the original data reduction model $P(k) \longrightarrow 0.184/\log_{10} k^{1.25} + 0.184$ where k is the time difference between the current event and previous events and the rate of reduction can be considered at local maxima to be 25% , we adjust such that $P(k) \longrightarrow 0.184/\log_{10} k^{1+\varkappa} + 0.184$ such that $\varkappa$ is a inverse normalised percentage value of k value variance ϱ, such that $\varkappa = 1 - \varrho$, as the neuromorphic data evolves over time.

3 Experiment and Results

We compare T--ROT against an original [2] this includes the same experimental setup using a popular and publicly available database [10]. We will compare ROT and R-ROT using the the shapes, boxes, walking and run datasets. The datasets are captured using a 240×180 resolution event-based and frame-based hybrid camera known as the DAVIS240-C [7]. Fig. 1 shows the output sample from the experiment, the ROT-Harris (proposed in [2] operating a Harris response model in 1-D) response is captured as computed from neuromorphic data held in ROT (green) and R-ROT (red) tree respectively. The original scene image is also provided to allow for contrasting. An assessment of ROT and R-ROT contents revealed a near $1:1$ ratio indicating that, at sampling time, the ROT and R-ROT contained the same events at that stage.

To evaluate ROT and R-ROT we compute accuracy, similarly to the original ROT paper, using ground-truth of the datasets to form true positive (tP) and false positive (fP) statistics as showing in (1).

$$\text{Accuracy} = \frac{\text{tP}}{\text{tP} + \text{fP}} \tag{1}$$

As in [2] we use the spatial identifiers (x and y) of the neuromorphic data as purely identifiers, we do not consider event data within this paper to exceed 1-D in operational space, this architectural decision has been shown to produce faster and more memory conservative means frameworks for operating over neuromorphic data at the cost of some thread-safe management. The results of [2] highlighted the benefits of using an ROT as an underlying engine for corner point extraction using an adapted Harris algorithm [11]. In terms of accuracy

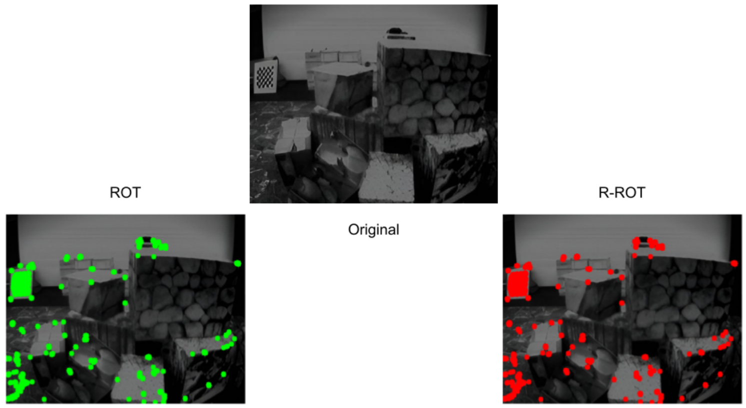

Fig. 1. A sample output of operating the 1-D ROT-Harris corner detector over data retained in the ROT (green) and R-ROT (red) trees respectively. The original underlaying image is provided for clarity

the original ROT algorithm achieved an average result of 68.45%, the R-ROT achieved an averaged result of 73.14%. It is important to note that on average we saw a factor of 4 decrease in overall memory usage while utilising R-ROT. Comparing ROT and R-ROT in time we also notice a significant change, the ROT reported an average of 83.25 nanoseconds per execution instruction (that is the amount of time taken to determine if an event is a corner or not) while R-ROT obtained an average of 64.33(rec.); this speed-up is likely the result of the R-ROT retaining fewer events when gaps (temporal voids) appear within the data while the data stream progressing.

4 Conclusion

This paper presents a variation of the classical ROT tree applied to neuromorphic vision data which we denote as R-ROT. The R-ROT tree utilises an adaptive data reduction approach to allow the self-pruning mechanism residing with ROT tree design to adapt based on the pattern of a data stream. The R-ROT structure is compared against the original ROT structure in the area of corner detection using the 1-D neuromorphic ROT-Harris algorithm to identify corners as interest points within a widely used database. The results show that the R-ROT structure is able to outperform the original ROT structure in accuracy (as calculated using accepted ground truth) and time (as calculated in decision time). For future work we will explore introducing a deep-learning model to replace or enhance the ROT tree features with the goals of deploy-ability and generalisation, alongside accuracy and speed increases.

References

1. Harrigan, S., Coleman, S., Kerr, D., Yogarajah, P., Fang, Z., Wu, C.: Reducing-over-time tree for event-based data. In: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE (2021) 1765–1772
2. Harrigan, S., Coleman, S., Ker, D., Yogarajah, P., Fang, Z., Wu, C.: Rot-harris: A dynamic approach to asynchronous interest point detection. In: 2021 17th International Conference on Machine Vision and Applications (MVA), IEEE (2021) 1–6
3. Lichtsteiner, P., Posch, C., Delbruck, T.: A 128×times 128×120 db $15mu$s latency asynchronous temporal contrast vision sensor. IEEE journal of solid-state circuits **43**(2) (2008) 566–576
4. Serrano-Gotarredona, T., Linares-Barranco, B.: A 128×128 1.5% contrast sensitivity 0.9% fpn 3 μs latency 4 mw asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers. IEEE Journal of Solid-State Circuits **48**(3) (2013) 827–838
5. Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A.J., Conradt, J., Daniilidis, K., et al.: Event-based vision: A survey. IEEE transactions on pattern analysis and machine intelligence **44**(1) (2020) 154–180
6. Yu, H., Cong, Y., Sun, G., Hou, D., Liu, Y., Dong, J.: Open-ended online learning for autonomous visual perception. IEEE Transactions on Neural Networks and Learning Systems (2023)
7. Brandli, C., Berner, R., Yang, M., Liu, S.C., Delbruck, T.: A 240 x 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. IEEE Journal of Solid-State Circuits **49**(10) (2014) 2333–2341
8. Alzugaray, I., Chli, M.: Asynchronous corner detection and tracking for event cameras in real time. IEEE Robotics and Automation Letters **3**(4) (2018) 3177–3184
9. Cuadrado, J., Rançon, U., Cottereau, B., Barranco, F., Masquelier, T.: Optical flow estimation with event-based cameras and spiking neural networks. arXiv preprint arXiv:2302.06492 (2023)
10. Mueggler, E., Rebecq, H., Gallego, G., Delbruck, T., Scaramuzza, D.: The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and SLAM. The International Journal of Robotics Research **36**(2) (2017) 142–149
11. Harris, C., Stephens, M.: A Combined Corner and Edge Detector. Procedings of the Alvey Vision Conference 1988 (1988) 23.1–23.6