Novel Algorithms for Novel Data:
Machine Learning for Neuromorphic Data from the International Space Station

Stefan Doucette, Nicole Lape, Tom Swinicki, Kevin Dickey, Tim Welsh
The MITRE Corporation

Dr. Geoff McHarg[1], Dr. Greg Cohen[2],
United States Air Force Academy[1], Western Sydney University[2]

CONFERENCE PAPER

MITRE presents processing algorithms developed for the Falcon Neuro Event-Based Sensor (EBS) operating aboard the International Space Station (ISS).

1. INTRODUCTION

Unlike traditional cameras that generate image frames at a fixed rate, neuromorphic sensors generate event streams based on changes of light intensity at each pixel. These novel sensors also have unusual characteristics, including asynchronous pixel operation, high dynamic range exceeding 120 dB, microsecond time resolution, and potentially lower data generation rates compared to traditional focal plane arrays. Research funded by MITRE's independent research and development program has focused on the development of novel algorithms for the processing and exploitation of the unique data produced by these detectors. A collaborative agreement between MITRE and the Space Physics and Atmospheric Research Center (SPARC) at the United States Air Force Academy (USAFA) provides access to a rare data source: the USAFA Falcon Neuro experiment, which provides neuromorphic imagery from cameras aboard the International Space Station (ISS).

The MITRE-developed algorithms described herein broadly aim to correlate neuromorphic events with categorizable semantic information, including sensor noise, geographic features, and in-orbit objects, with a goal of demonstrating the contribution of event-based sensors in space and edge-processing to the space domain.

2. BACKGROUND

Neuromorphic, or event-based, sensors (EBS) were developed to mimic the operating principles of the human eye – itself a highly efficient detector coupled to a neural network. Unlike standard complementary metal oxide semiconductor (CMOS) and charge-coupled device (CCD) detectors, which sample light across all pixels during the exposure time, neuromorphic sensors produce data according to per-pixel brightness changes, individually known as

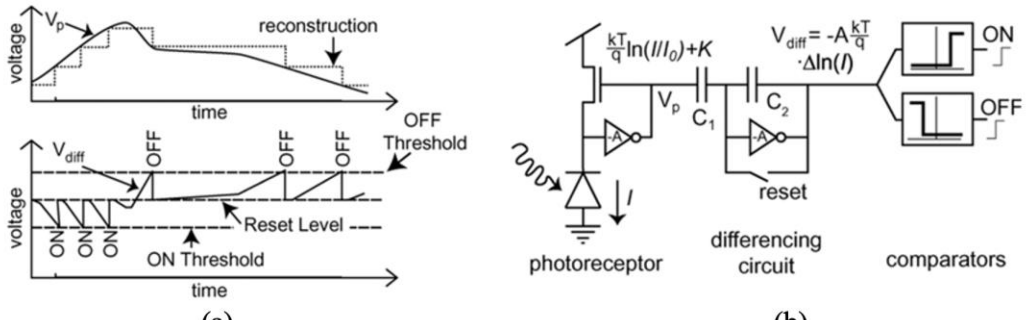


Fig. 1. (a) Pixel response function of an event-based sensor. (b) Simplified event-based sensor pixel diagram. Reprinted from Lichsteiner, et. al [4].

"events." Each pixel converts a continuous intensity signal from the photoreceptor into a discrete step function, as shown in in Fig. 1a. Pixel-level circuitry, shown in a simplified diagram in Fig. 1b, registers an event when the difference between the continuously read photoreceptor and the voltage of the previous event exceeds a user set threshold. As such, there is no fixed exposure time, and the time between logging events at each pixel, or the "refractory period," is generally limited by the transient properties of the circuit. The event is described by the pixel location ("x" and "y" values) and polarity ("0" for an increased illumination change, and "1" for a decreased illumination change). Upon observance, the recorded event is time-stamped at microsecond resolution by the arbiter and sent from the sensor to the host. The pixel then resets the current level in accordance with the polarity direction previously sent to the arbiter, and the process repeats.

Because the primary method of computation is the analog pixel circuitry, the high temporal resolution of the sensor and low latency resemble high-speed camera capabilities; previous evaluation of neuromorphic sensors have demonstrated their ability to capture bullets in flight (work not shown). The asynchronous nature of the pixels theoretically allows for low power consumption and data generation since both are functions of the activity in the scene, where traditional cameras generate a fixed amount of data per image. With no exposure time, EBS data are not subject to motion blur. Finally, the asynchronous logarithmic adjustment of the reset level for each pixel results in high dynamic range (>120 dB).

Event camera data streams are simply chains of events that are generated by the camera as they occur. Each event contains the (x,y) coordinate of the detected change, the polarity of the change (i.e., brighter or darker), and a timestamp. For processing in machine learning applications, it is common to process windows of events, either with fixed numbers of events or with a fixed time window.

3. DATA SOURCE

We developed algorithms using data collected from sensors installed on the International Space Station (ISS) by the United States Air Force Academy (USAFA) Space Atmospheric Research Center (SPARC). The SPARC-produced Falcon Neuro system (Fig. 2) is based on two commercially available iniVation DAVIS 240C neuromorphic cameras, modified for use aboard the ISS. One sensor – the *ram* camera – points towards the direction of the ISS travel path and the other sensor – the *nadir* camera – is pointing towards Earth and 20 degrees starboard to avoid occlusion by ISS structure. We focused most of its processing efforts on the data from nadir camera; results shown here are representative of data from this sensor as it passes over different areas of Earth.

Fig. 2. (a) Rendering of the Falcon Neuro system, showing two neuromorphic cameras (D,E), and data processing boards (A,B,C). (b) Falcon Neuro system (highlighted in red) after installation on the ISS. Both images are courtesy of SPARC.

Built from commercial-off-the-shelf (COTS) components, the Nadir camera is a DAVIS 240C focal plane array with a 25mm f/1.8 lens. With a resolution of 240x180 pixels and 18.5µm pixel pitch, the imager has an instantaneous

field of view (IFOV) of 310m/pixel and a ground swath of roughly 73x55km. The camera is controlled by a field-programmable gate array (FPGA) on an advanced reduced instruction set computer machine (ARM) processor board, developed by members of the Falcon Neuro team. The FPGA is tasked with handling camera event streams, providing timestamps, pre-filtering spurious events, and storing data locally before transmission to a ground station. This controller is also tasked with managing the 20 independent sensor-specific bias settings of the 240C, which tune the event rate, pixel response function, and noise sensitivity of the camera. The data are streamed periodically to USAFA and stored as hierarchical data format 5 (HDF5) files for processing and dissemination.

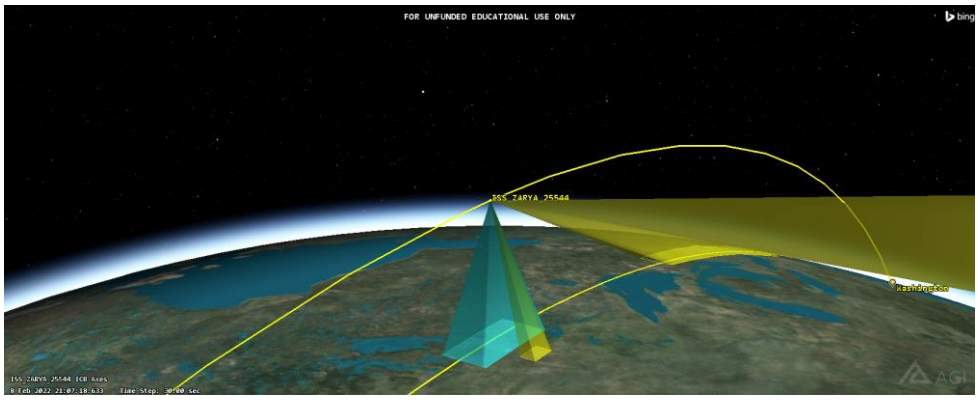


Fig. 3. Field of view of Falcon Neuro (yellow) and the NASA HD camera (blue) on the ISS.

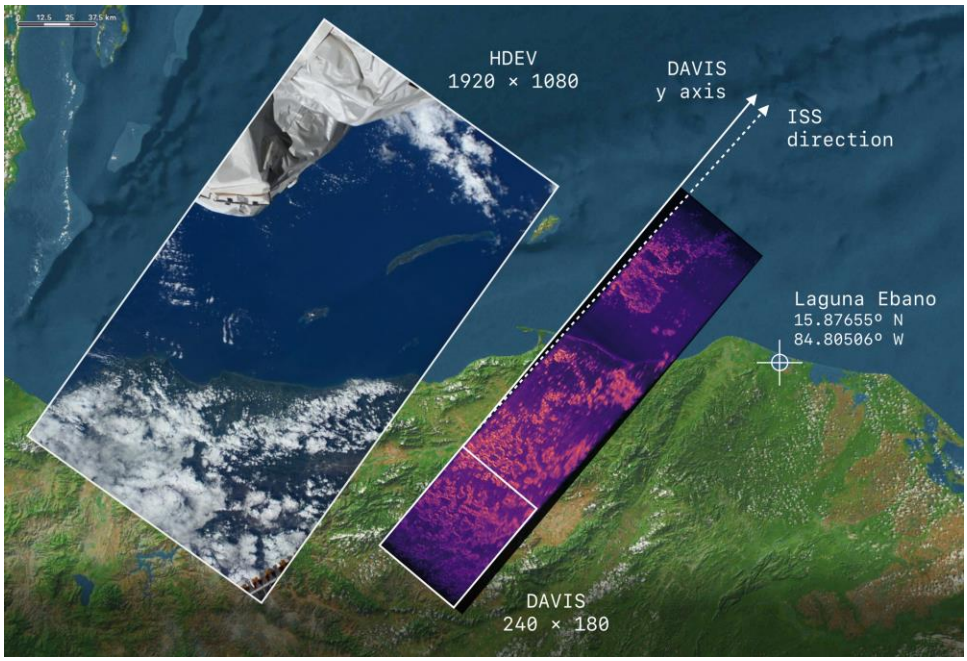
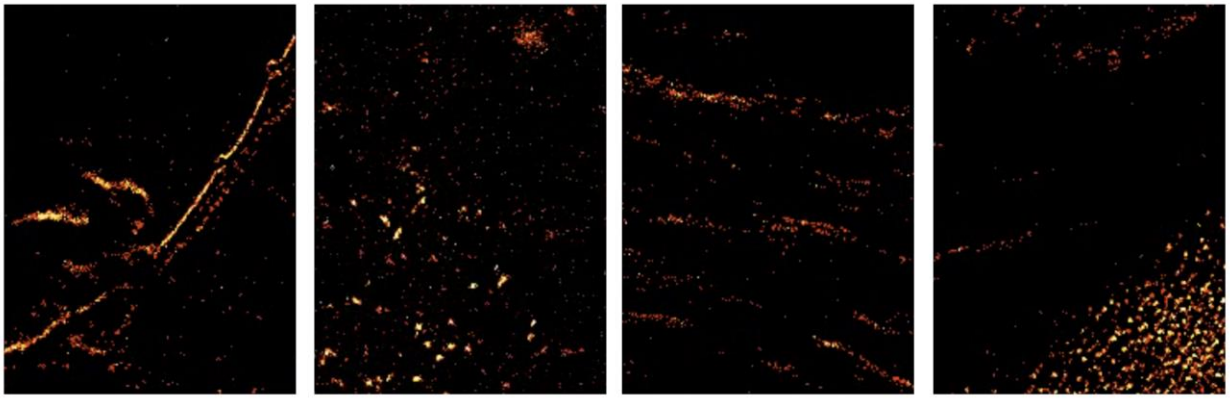


Fig. 4. The above image shows the field of view of the Falcon Neuro Nadir sensor (right inset) and the NASA HD camera (left inset) superimposed on imagery of the Earth. The DAVIS240C sensor in Falcon Neuro has a pixel resolution of 240 x 180 pixels, as shown by the white rectangle within the inset. The continuous data from the Nadir sensor is then motion-compensated to produce continuous image strips of fixed width. The resolution of the NASA HD camera is also provided for reference. Base layer source: Apple Maps (Earth Imaging data).

Geography such as that in Fig. 5 captured in the field of view of Falcon Neuro can be identified both by the known location of the International Space Station and the NASA HD camera as in Fig. 3 and Fig. 4.



(a) (b) (c) (d)

Fig. 5. Falcon Neuro data of (a) the Baja Peninsula, (b) Central Europe at night with city lights, (c) New Zealand, and (d) the Great Lakes region, all visualized as frames where the value of each pixel is the total number of positive polarity events recorded at that location during a 10-millisecond time window.

4. ALGORITHMS OVERVIEW

The algorithms described in this section were developed by MITRE-funded research projects and adapted as needed to be applied to the Falcon Neuro data. These tools include frequency domain source segmentation, graph representation, anomaly detection, and streak detection. Additional work is being explored following research from Zhu, et al. [3] which uses an unsupervised machine learning approach to estimate the optical flow of the events, along with ego-motion and depth perception. Results from this approach are under evaluation and are not yet available.

In general, event data can be processed either as native, polarity-based data or as frames resembling traditional images, generated by "flattening" events within a short time window into a two-dimensional array. While frame-based processing enables the use of traditional image processing algorithms for non-sparse two-dimensional datasets, this approach effectively down-samples data in the temporal domain, negating some of the inherent advantages of event-based imagers.

Frequency Domain Source Segmentation (FDSS)
Some light sources observed from the ISS, such as city lights and lightning strikes, were suspected to be periodic in nature. To find frequencies in neuromorphic data, the algorithm calculates the correlation between neighboring events, forms clusters based on similar correlation values, and then calculates frequencies using a Fast Fourier Transform (FFT).

For FDSS, event data is segmented into chunks of 100ms each. Within each chunk, only polarity events that occurred more than once at each pixel are preserved – a simple but effective approach to reduce noise in the data. After removal of suspected noisy events, the chunk is further subdivided into frames; for our experiments, we separated each chunk into two subframes. The result of this process is a 50-ms time window of events, with noise and other artifacts removed over a longer 100-ms window.

The events in each frame were spatially aggregated into blocks by moving a 4x4 kernel across the entire frame. Of these blocks, only the largest 10% were used for correlations, from which a normalized correlation was taken using the following equation:

$$\frac{CC(y_2, y_1)}{\sqrt{med\big(CC(y_1, y_1)\big) * med(CC(y_2, y_2))}}$$

where *CC* is the discrete linear cross-correlation between two N-dimensional arrays and *med* is the median of the output. Here, y_1 and y_2 are the two neighbor event blocks that are compared. These correlations are tracked for each adjacent neighbor group at each pixel.

Clusters are formed from the neighbors of the blocks that have correlation values greater than 10% of the maximum correlation value. For each cluster, the timestamps of each of the contained events are counted and converted to a time signal, which is then normalized against the mean of the signal and used when taking the FFT. The maximum value from this FFT is the fundamental frequency of the cluster, which is assigned to the cluster for that frame. If the same cluster is temporally consistent with previous frames, the counter will be updated with the values from the previous frame and the FFT will be taken again to produce an improved estimate of fundamental frequency from a longer temporal sequence of events

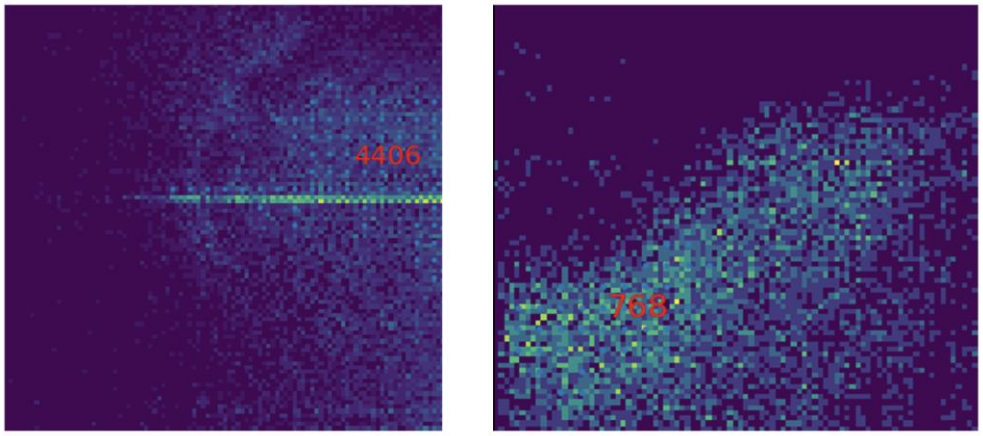


Fig. 6. Visualizations of lightning events from Falcon Neuro nadir camera over (a) the Democratic Republic of the Congo and (b) Sri Lanka.

As seen in Fig. 6, results included initial analysis of lightning events. Lightning events consist of shorter events where the return stroke is the main visible event. An example depiction of a lightning event is shown in Fig. 7, where the duration of return strokes average about 65 microseconds and the time between the return strokes averages 1.5 milliseconds [6,7]. Thus, the expected frequency would be roughly 1KHz. Processing the lightning event data resulted in frequencies as low as 10Hz and as high as 4.4kHz. Not enough samples of these events were collected for a statistically significant assessment of measurement accuracy. The most common frequencies found in the event data were those captured of clouds, which we did not anticipate returning frequencies at all. Ego-motion of the sensor may account for some of these spurious frequency-domain signals. FDSS analysis of the Falcon Neuro data provided inconclusive results for common phenomena, requiring additional investigation as more data are collected.

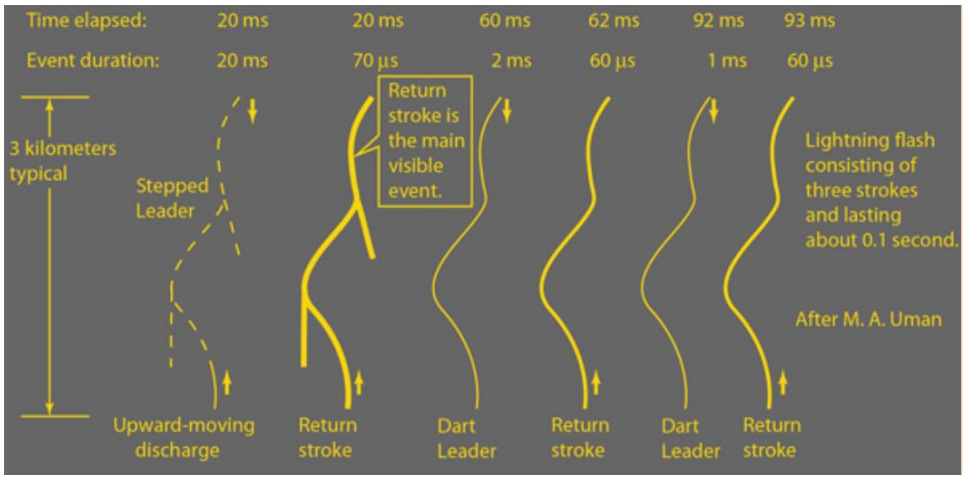


Fig. 7. Lightning event progression including multiple return strokes at an average frequency of ~10Hz. Reprinted from Nave [7].

We believe that the most commonly observed frequency from FDSS resulted from the ego-motion of the sensor on the ISS. The expected rate that an object on earth would traverse pixels is based on the field of view, altitude, and velocity of the ISS, which we calculated as roughly 25 pixels per second along the frame. Several captures that were analyzed using FDSS had frequency outputs around the value of 25 Hz, which implies possible influence from the ego-motion of the sensor. Future work utilizing software correction to remove this frequency where appropriate will be explored as well as with different timeframes and Fourier methods.

Graph Representation

The graph-theory algorithm developed within this study connects events based on spatial and temporal correlations. Each event is represented as a node in a graph, and nodes are only interconnected by edges that indicate high correlation according to a calculated velocity and a threshold set by the user. The velocity between two nodes is calculated as

$$v = \frac{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}{|t_2 - t_1|}$$

where x_1, y_1, and t_1 are the coordinates and timestamp of the first event, and x_2, y_2, and t_2 belong to the second event. The algorithm sets the velocity to be zero if the timing between the two events is too small to be significant; this prevents division-by-zero errors. Typically, this threshold is 1 nanosecond. Using the calculated velocities between nodes, the graph is pruned to remove edges with insufficient velocity or correlation. The algorithm can also be tuned to remove nodes with low connectivity, maintaining only the most critical nodes in the graph. As seen in Fig. 8, the resulting representation can reveal events according to the motion of an object in the scene, like the movement of the rotors of a UAV over time.

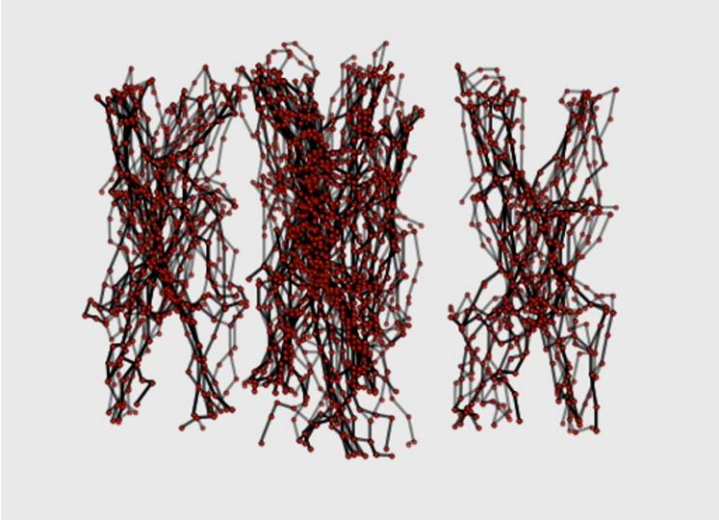


Fig. 8. The graph-based algorithm applied to event data collected from the rotors of a quadcopter in flight.

Serving as both an alternate visualization of event data and as a format for computing spatiotemporal correlation of each event to other events, the graph representation of Falcon Neuro data from the ISS may provide utility for more generalized processing of event-based imagery. In this representation, objects with velocities relative to the surface of the earth, or those higher than the surface of the earth, may be represented as independent graphs, providing trivial object segmentation exceptional time resolution. Additionally, it is hypothesized that the structure of Earth features such as clouds and coastlines will result in categorically differing structures that machine learning models can learn to distinguish regardless of their relative velocities. Further efforts to process and understand graphs produced from the Falcon Neuro data, such as Fig. 9, are planned.

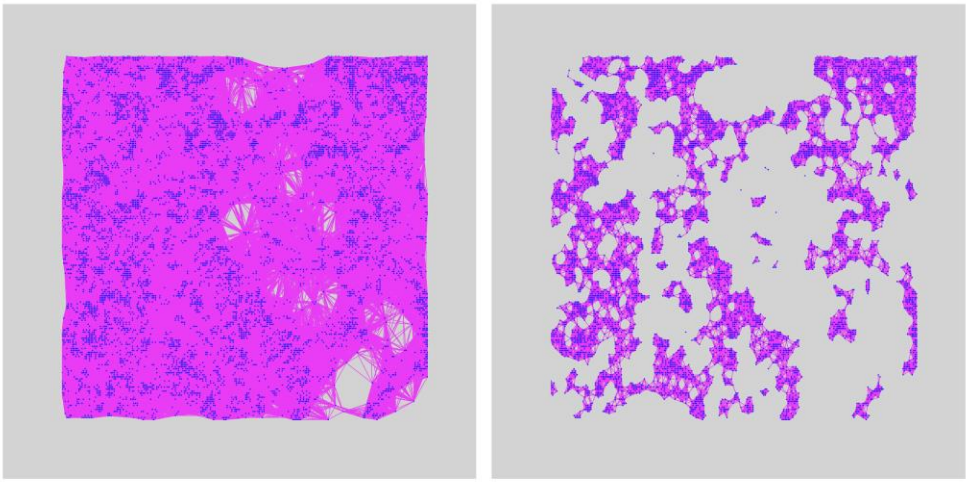


Fig. 9. (a) A top-down view of a graph representation of a slice of event data, where each node is an event that has not been pruned, and each edge represents a correlation in time and space between one event and another. (b) Using the same input data as used in Fig. 9a, a stricter threshold for event correlation is applied.

Anomaly Detection

Machine learning approaches were applied to Falcon Neuro event data to attempt to reduce noise and detect data outliers. We tested approaches for adapting to periods of strong noise produced by the sensor, as well as the detection of high-speed objects such as orbital debris. Each challenge shares the common tasks of data segmentation and windowing, model selection and training, and parameter tuning for optimal results.

Event data can be processed one event at a time, as correlated to surrounding pixels and their events, through computational techniques like time surfaces or in time-based bins of events similar to frames. To create a single frequency frame, a slice of data containing every event detected during a 10,000 microsecond (0.01 second) time frame is created, followed by the allocation of an empty array matching the native resolution of the imager at 240x180 pixels. Each location in the frame is populated by the total count of events detected for that pixel. Only positive polarity events were considered at this stage. Including both positive and negative polarities tended to provide redundant information. Another option for creating frames is a so-called "surface active" event frame, where each pixel in the array is updated with the most recent timestamp of an event that fired at that location rather than a total count of all events that fired at that location. Tools were created for both frame types, as well as other types like time surfaces, but most experimentation and visualization was performed using frequency frames.

Adapted autoencoder models from the Python Outlier Detection library [5] were used to discover anomalies in the Falcon Neuro data. After testing various model layer and learning rate combinations, it became evident that the models did not have to be large or complex to provide significant results. Fig. 10 shows that frames with large amounts of noise, believed to be the result of environmental issues at the sensor, were flagged by the model as highly anomalous compared to the rest of the data. As mentioned at the start of this section, the downside to processing event data as frames is the loss of rich temporal information that can be provided by processing the data in its raw event form. This process could be used for a first pass in identifying windows of data that are highly noisy – potentially too noisy to contain useful information.

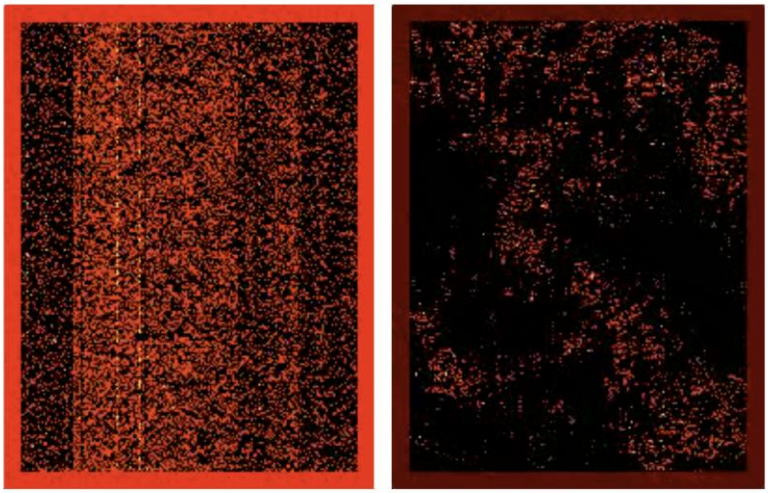


Fig. 10. (a) A frame that has been flagged as highly anomalous by the model (bright red border). (b) A frame exhibiting a lower score for being anomalous (darker red border).

Streak Detection

Internal streak detection algorithms were modified and applied to the data in order to find events that may serve as orbiting objects of interest. The algorithm for this detection is random sample consensus (RANSAC)-based. It samples random points from the input and calculates a linear regression model. It subsequently finds inliers to the model and updates accordingly. If a sufficient number of inliers – based on the Euclidean distance from a three-dimensional line – are found, a streak is considered to be detected.

The Falcon Neuro data may or may not contain events caused by nearby objects in space. If present, such events would be rare given the low altitude of the ISS, the limited field of view of the neuromorphic sensors, and the sparsity of objects in Low Earth Orbit (LEO). Further, objects passing within the field of view of the sensor would do so in a very short amount of time since the potential velocity difference in low earth orbit (LEO) satellites can be up to 14 km/s. Identifying events generated by nearby objects amid abundant false positives requires ground truth, which was not available. Existing modeling and simulation tools could be used to predict whether known objects might pass through the field of view of a sensor, but these analyses were beyond the scope of this research.

To estimate theoretical detection performance on Falcon Neuro data, as well as to create a dataset to test future algorithms, we developed code to insert synthetic streaks into real event data. The inserted streaks have user-defined values for streak radius, speed, and direction. Fig. 11 shows two synthetic streaks in blue, both successfully detected by our streak detection algorithm as shown in red.

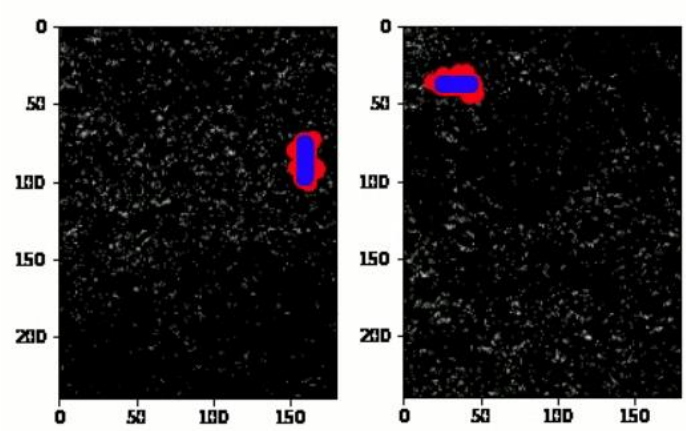


Fig. 11. Synthetic streaks being detected by the algorithm with truth in blue and output detections in red.

We expect our algorithm to be able to detect objects with motion similar to the synthetic streaks traveling through the field of view. While the synthetic streaks simulate a simple case and only provide a first-order model of such an event with no consideration of astrodynamics or spacecraft structures, it provides a promising initial outcome for event-based detection of moving objects aboard Falcon Neuro.

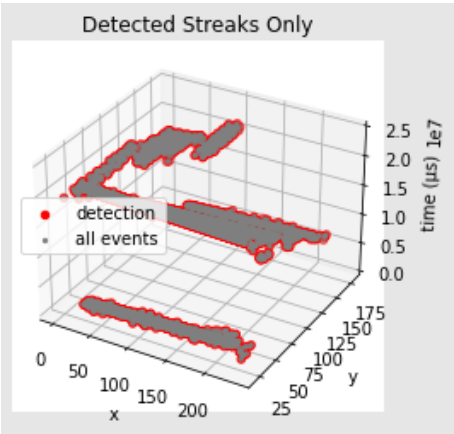


Fig. 12. A three-dimensional representation of streaks detected from synthetic data, with time shown in the z direction.

5. APPLICATIONS

Events that happen in or near space that might be of interest to scientific or commercial observers include the movement of atmospheric or space vehicles, tracking of manmade or natural orbiting objects (e.g., satellites, asteroids, and comets), and space and atmospheric weather. Interestingly, many of the events observed in or from space exhibit traits that appear to be a good match for event-based imaging technology—i.e., events that occur rarely but happen over a very short period of time or have features requiring high temporal resolution. Future applications of neuromorphic sensing technology will be most effective when exploiting the unique strengths of the sensor, which include high dynamic range, high temporal resolution, and potentially lower data generation and power consumption resulting from asynchronous pixel operation. Commercially available event-based sensors have a

significantly lower sensitivity when compared to traditional cameras, so scenarios with very dim targets are likely not appropriate for this technology until design improves.

Space-based neuromorphic sensors will likely benefit from edge processing. Converting data to consumable information onboard a satellite could reduce load on communications systems, especially compared to traditional camera data, and allow for autonomous operations. The ability to process event-based imagery in real time using machine learning aboard a spacecraft may enable fast, automated responses to rare events that would otherwise require the extended operation – and significant size, weight, and power (SWAP) penalties – of a high-speed framing camera.

6. SUMMARY AND CONCLUSION

We have developed several in-house research software packages that enable us to quickly process data, test research algorithms, and newly visualize the unique event data collected by the SPARC team from the ISS. Each of our algorithms can be expanded upon using variations on data processing and model tuning. By continuing collaboration and sharing with the research community, we can further develop useful tools which enhance the space-based observation of Earth and beyond.

7. ACKNOWLEDGEMENTS

We would like to thank Dr. Kris Rosfjord, Guido Zarrella, and the MITRE Independent Research and Development Program for funding and supporting the development of neuromorphic collection and processing within MITRE. We would also like to thank the United States Air Force Academy Space Physics and Atmospheric Research Center and Western Sydney University teams for their ongoing collaboration, support, and sharing of data.

8. REFERENCES

[1] E Grun, HA Zook, H Fechtig, and RH Giese. Collisional Balance of the Meteoritic Complex, *Icarus*, 62(2):244-272, 1985.

[2] NL Johnson and DS McKnight. *Artificial Space Debris*, Orbit Book Company, 1991.

[3] AZ Zhu, L Yuan, K Chaney, and K Daniilidis. *Unsupervised Event-based Learning of Optical Flow, Depth, and Egomotion.* University of Pennsylvania. 2019.

[4] P Lichsteiner, C Posch, and T Delbruck. *A 128x128 120 dB 15 µs latency asynchronous temporal contrast vision sensor*. IEEE Journal of Solid State Circuits. 2008.

[5] Y Zhao, Z Nasrullah, and Z Li. *PyOD: A Python Toolbox for Scalable Outlier Detection.* Journal of machine learning research (JMLR), 20(96):1-7. 2019.

[6] MA Uman. *Everything You Always Wanted to Know About Lightning but Were Afraid to Ask*. The Saturday Review, 36-41. 1972.

[7] CR Nave. *HyperPhysics: Lightning Time Sequence.* Georgia State University. http://hyperphysics.phy-astr.gsu.edu/hbase/electric/lightning2.html. 2016.