Event-Based Object Detection and Tracking for Space Situational Awareness

Saeed Afshar[ID], *Member, IEEE*, Andrew Peter Nicholson, Andre van Schaik[ID], *Fellow, IEEE*, and Gregory Cohen[ID], *Member, IEEE*

Abstract—In this work, we present an optical space imaging dataset using a range of event-based neuromorphic vision sensors. The unique method of operation of event-based sensors makes them ideal for space situational awareness (SSA) applications due to the sparseness inherent in space imaging data. These sensors offer significantly lower bandwidth and power requirements making them particularly well suited for use in remote locations and space-based platforms. We present the first publicly-accessible event-based space imaging dataset including recordings using sensors from multiple providers, greatly lowering the barrier to entry for other researchers given the scarcity of such sensors and the expertise required to operate them for SSA applications. The dataset contains both day time and night time recordings, including simultaneous co-collections from different event-based sensors. Recorded at a remote site, and containing 572 labeled targets with a wide range of sizes, trajectories, and signal-to-noise ratios, this real-world event-based dataset represents a challenging detection and tracking task that is not readily solved using previously proposed methods. We propose a highly optimized and robust feature-based detection and tracking method, designed specifically for SSA applications, and implemented via a cascade of increasingly selective event filters. These filters rapidly isolate events associated with space objects, maintaining the high temporal resolution of the sensors. The results from this simple yet highly optimized algorithm on the space imaging dataset demonstrate robust high-speed event-based detection and tracking which can readily be implemented on sensor platforms in space as well as terrestrial environments.

Index Terms—Space situational awareness, event-based detection, event-based features, event-based tracking, event-based processors.
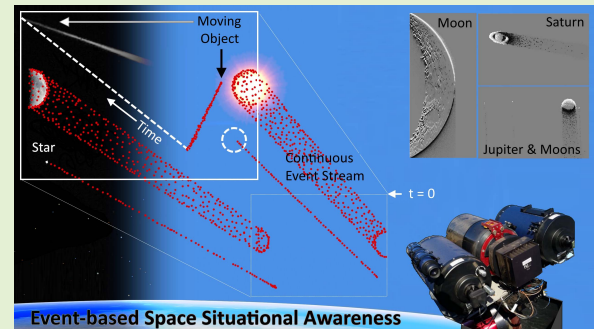
I. INTRODUCTION

A. Event-Based Space Situational Awareness

OUR increasing reliance on space-based technologies for communication, navigation and security tasks as well as the recent dramatic drop in the cost of space launches has created an immediate need for better methods for detecting and tracking objects in orbit around the earth [1]. The cost of collisions in space poses a significant risk to both our space infrastructure and future space missions.

Space Situational Awareness (SSA), and Space Traffic Management (STM) — its civilian counterpart — are therefore critical tasks for regulation and enforcement of the use of space, and to prevent a future catastrophic

space event, such as described by the Kessler effect [2]. Space Situational Awareness is defined by the European Space Agency (ESA) as comprising three segments: Space Surveillance and Tracking (SST), Space Weather and Near Earth Objects (NEO) [3]. This work contributes primarily to the task of space surveillance and tracking, specifically applied to satellites in orbit around the earth. Currently, over 80 countries have a presence in space and this is likely to increase, driven by both national space efforts and private industry [4].

Currently, the majority of SSA data originates from dedicated radar installations operated by the United States Air Force [5]. However, radars are an expensive technology to install and operate and there is an increased focus on looking toward optical telescopes to provide a more flexible, cost-effective and responsive means of obtaining accurate SSA data [4]. In our previous work, we have demonstrated that event-based neuromorphic cameras offer a novel means of performing SSA tasks and provide capabilities that cannot be achieved using conventional astronomy cameras [6].

Event-based cameras operate in a different imaging paradigm, emitting data as a spatio-temporal pattern rather
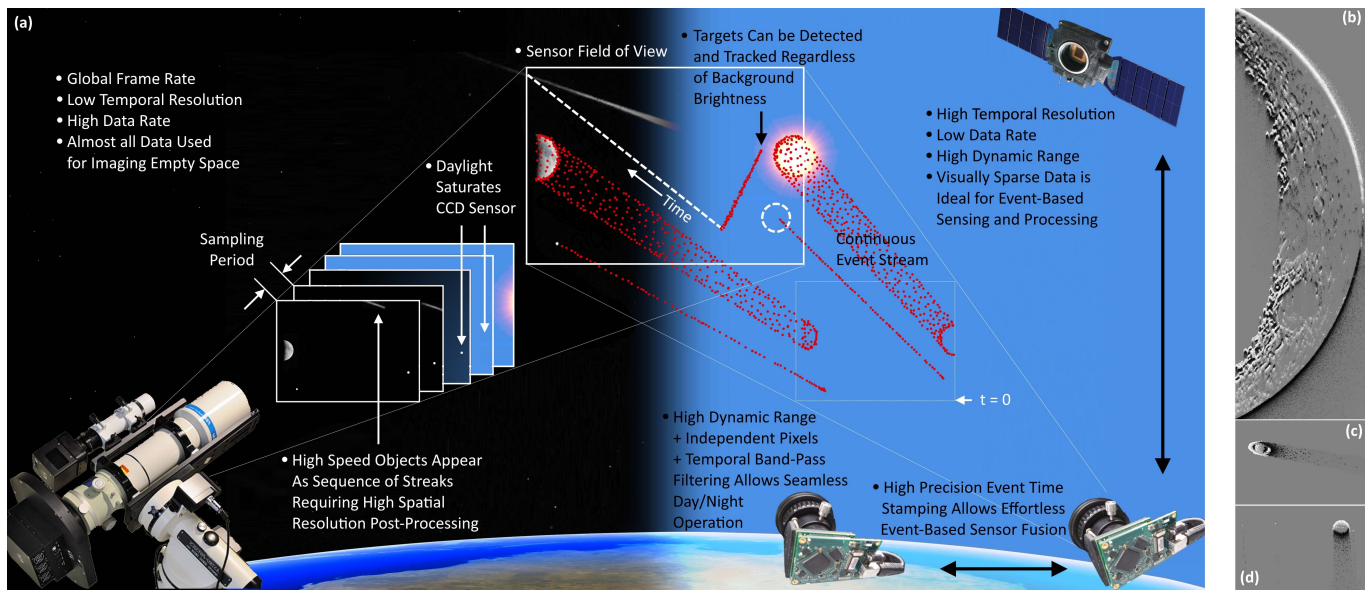
Fig. 1.   **Event-based Space Situational Awareness (EBSSA) compared to the standard CCD sensor approach.** (a) Event-based sensors provide high temporal resolution imaging data of the sparse space environment allowing rapid sensor fusion, low bandwidth communication and operation during continuous operation during day and night time. (b) Examples of event-based imaging of the Moon, (c) Saturn, (d) Jupiter and moons.

than using conventional frames [7]. The pixels are also independent and asynchronous, giving the device a high temporal resolution and a very high dynamic range [8]. The characteristics of these devices enable unique opportunities for space imaging and novel approaches to satellite detection and tracking [9], high-speed adaptive optics [10], satellite identification [11] and real-time in-frame astrometry [12].

This work builds upon those findings and presents two methods for tracking objects in the spatio-temporal output of an event-based camera. There exist many event-based trackers, such as those for long-term object tracking [13], real-time particle tracking [14], micro-particle tracking [15], corner detectors [16] and more complex kernel tracking algorithms [17]. These methods are all very specific to both their specific application and data, but do not generalize well and are not easily applicable to event-based space imaging (EBSI) data. Figure 1 shows two-dimensional renderings of the three dimensional spatio-temporal event-based data recordings.

Event-based Space Situational Awareness (EBSSA) is a new and emerging field of study [6], [12], [18], [19]. The most relevant work to that presented here is the frame-based star tracking method proposed in [19]. Here, an event-based camera was used to capture simulated star imaging data displayed on a monitor. A frame-based rotation averaging and bundle adjustment method was used to determine the dominant motion in the field of view. However, this system was designed and tested on simulated ideal data which did not exhibit the noise and dynamics of real-world event-based space imaging environment. Furthermore, the method was aimed at extracting a single velocity vector from a high Signal to Noise Ratio (SNR) star field as opposed to the detection of independent moving targets with a wide range of SNRs as is presented in this work.

The application of event-based cameras to real-world space imaging leverages the unique nature of the independent and

asynchronous pixels in the sensor. This is accomplished with dedicated in-pixel circuitry which cannot be replicated or generated by a conventional sensor. The sensors can continuously image whilst in motion, allowing the telescope to move arbitrarily whilst still capturing valid data. This is in contrast to conventional telescope tracking, in which the motion is induced to keep objects centered in the field of view. Event-based sensors can therefore capture both the object being centered, and the background stars without generating the streaks commonly found in conventional space imaging. Furthermore, the in-pixel circuitry also generates events in response to changes in log illumination at each pixel, providing a very high dynamic range for each pixel allowing for imaging across both day and night lighting conditions. Whilst daytime imaging can be performed with a standard camera [20], the change-based detection provided by event-based sensors alleviates many of the problems associated with saturation and noise in this environment.

It therefore allows for different and novel approaches to space imaging which can overcome many of the current limitations in space situational awareness systems. In our previous work, we demonstrated the ability to detect a resident space object in orbits ranging from low-earth orbit (LEO) to geosynchronous orbits (GEO) [12]. We also demonstrated the ability to observe objects during the day with an event-based camera, and without any modifications to the optics.

Figure 1 provides an overview of the benefits of a neuromorphic approach to space imaging. The low-power and low-bandwidth operation of event-based sensors makes them highly suitable for use on orbital platforms, and the ability to synchronize cameras in a highly efficient manner also creates the potential for large distributed SSA observation networks.

The continuous nature of the imaging provided by event-based sensors allows for the camera to image whilst moving, and as a result, allows the device to operate in less
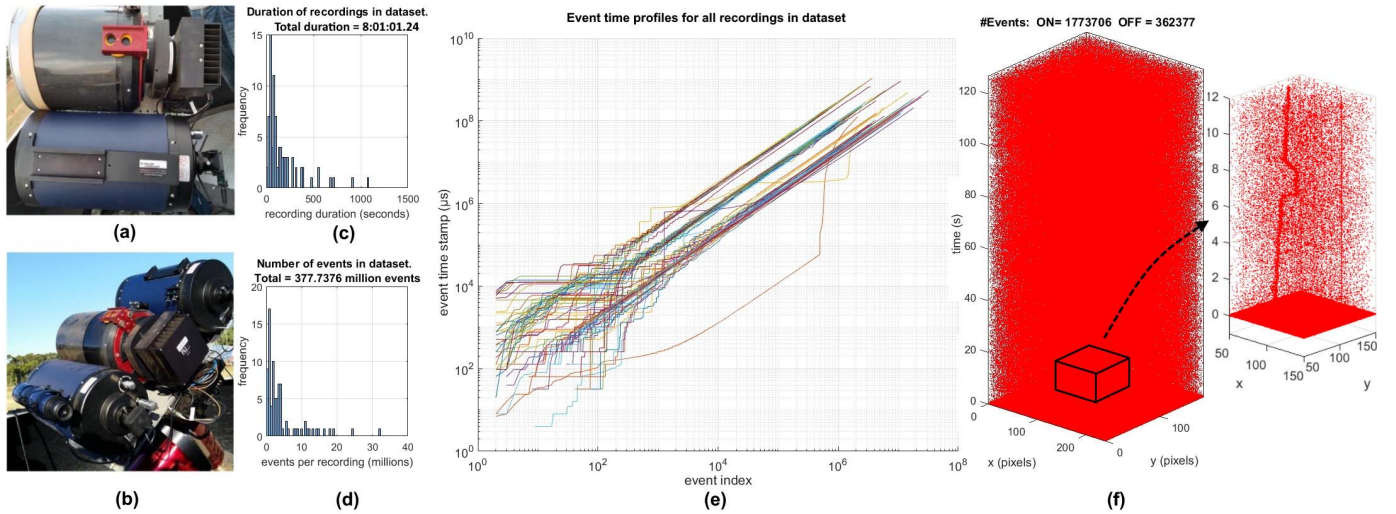
Fig. 2. **Space imaging set up and resulting dataset.** Panels (a) and (b) show photographs of the equipment used in the recording of the space imaging data. Two identical telescopes were used for the ATIS and DAVIS event-based sensors alongside a conventional astronomy CCD camera (FLI Proline PL4710). (a) The ATIS camera is attached to the base of the lower telescope with the CCD camera shown at the top. (b) Shows the set up used in the simultaneous co-collects from both the ATIS and DAVIS cameras. Note that the optics for the telescopes for the event-based cameras were not altered between daytime and nighttime operations. Panels (c) and (d) show the distribution of the recordings in terms of duration and number of events respectively. (e) Plots the timestamp of all recordings in the dataset as a function of their index. Each line indicates one recording. (f) The Dimetric projection of the event stream from a two-minute recording of the rocket body SL-8 R/B [22] (NORAD ID:21938) with time as the vertical axis. A rendered video of this recording is provided with the dataset. The inset in the panel focuses on a small sub-region of the spatio-temporal event stream centered around the object with the highest SNR which is the rocket body. The plots of the full recording and the small sub-region show the high noise rate of a typical EBSI recording as well data artifacts such as noisy hot pixels (unbroken vertical red lines in the inset), event timing jumps (at approximately t = 93 s), and event dumps (at approximately t = 0 s). Such non-ideal behavior was observed with both the DAVIS and the ATIS sensors under different conditions. This panel serves to highlight the significant difference in SNR between real-world EBSI data and typically high SNR terrestrial event-based datasets.

stable environments than conventional astronomy cameras. The field of EBSSA requires the development of robust real-time space object detectors and trackers that can operate reliably in the presence of unexpected and random saccade-like ego-motion and in the presence of a wide range of noise conditions. This makes the task significantly different from conventional detection and tracking problems.

This work introduces a number of novel contributions to the literature:

- A large event-based space imaging dataset.
- A novel volume-based method for evaluating the information content of an event stream enabling an unbiased comparison of the output of event-based systems with each other and to their raw input.
- A novel high-speed event-based feature detector optimized for the detection of low SNR space objects and one that is robust to target size, orientation, and speed.
- A novel sequential least-squares-based tracker that dynamically weighs target orientation information as a function of target velocity.
- A cascaded filter design that applies progressively more refined and computationally expensive filters on the event data which applies complex, highly selective filters the data without slowing the overall system.

II. METHODOLOGY

This section describes the structure and nature of the events generated by the event-based cameras, the method used to generate the event-based space imaging dataset, the methodology used when labeling the dataset, and the metrics used to report sensitivity, specificity and informedness

from the event streams. The section further details a complete event-based detection and tracking system, as well as discussion of alternatives methods for benchmarking performance.

A. Generation of the Space Imaging Dataset

The space imaging dataset was captured using both ATIS sensors [8] and DAVIS sensors [21] and was undertaken at the Australian Defence Science and Technology Group's research facility in Edinburgh, South Australia. The experiments made use of their robotic electro-optic telescope facility, which was modified to support the event-based sensors and the existing astronomy equipment simultaneously. The DAVIS camera used to generate the dataset has a 240×180 pixel resolution at 18 μm with a 2000mm focal length $= 7.44 \times 5.58$ arc-minutes $= 0.124 \times 0.093$ degrees. The ATIS camera used has a 304×240 pixel resolution at 30 μm with a 2000mm focal length $= 15.66 \times 12.36$ arc-minutes $= 0.261 \times 0.206$ degrees.

The conventional telescope configuration comprised an Officina Stellare RH200 telescope and an FLI Proline PL4710 camera. This telescope and camera were used to provide ground truth and to build an accurate mount and pointing model, allowing the event-based cameras to track and to be accurately pointed at objects. The telescopes were both mounted on a Software Bisque Paramount MEII robotic mount, as shown in Figure 2. The system is housed in a 7 ft Aphelion Dome which contains a PC that controls the robotic telescope and controls the event-based cameras.

The event-based cameras were attached to an 8" Meade LX200 telescope, as shown in Figure 2. When performing

co-collects with both event-based sensors, a second Meade LX200 was attached on the other side of the primary telescope as shown in (c).

With over 8 hours and 377 million events, the presented dataset is the first publicly available event-based space imaging dataset in the literature. The dataset consists of 84 separate labeled recordings, 45 using the DAVIS sensor and 39 using the ATIS. The dataset contains event-based data and manually-labeled ground truth describing the motion of objects through the field of view. Due to the novelty of using event-based cameras for space imaging, it was not possible to link these tracked objects with verifiable resident space objects in every case. Therefore, the labeled data refers to the position of the objects as observed by the camera, rather than the actual resident space object. The techniques presented in this work provide the fundamental techniques for developing the ability to generate such a labeled dataset, and this is the focus of our current and future work. The full dataset, supporting material presented in the work can be accessed at [23]. The timestamps of the recordings are relative to the start of the recording. The recordings contain portions where the satellites are actively being tracked, where the stars are siderally tracked, and where the telescope is static. Due to the continuous nature of the event-based recordings, some recordings contain portions of all three types of operation. It is important for event-based systems to reliably track across all modes of movement, as this is what will usually be present when operating these systems.

In addition, a further 149 unlabeled data streams containing 5 hours of recording and containing 2513 million events are provided. These include 15 recordings from the 180×240 DAVIS sensor, and 27, 100 and 7 using, a 304×240 pixel ATIS camera, a larger format 640×480 pixel ATIS prototype camera, and the BSI variant of the DAVIS sensor described in [24] respectively.

The larger unlabeled dataset enables further exploration of almost all currently available EBSI data by the research community. As shown in Figure 2(e), the time-stamp profiles of all recordings in the dataset show the heterogeneity and non-idealities in the dataset. The discontinuous staircase features in the time-stamp profiles represent event stream timing artifacts. These event stream 'jumps' and 'dumps' occur when multiple events are erroneously assigned simultaneous time-stamps often at periodic intervals. This effect is likely due to USB communication delays in the cameras.

Presented on a log-log scale, these discontinuities in time and event index can be observed more frequently at the lower scale at the lower-left corner but are present with decreasing frequency at the higher scales, as the plots move to the top-right corner where discontinuities represent more severe artifacts. The effect of these artifacts on the data stream is also illustrated in Figure 2(f). This particular recording of the rocket body SL-8 R/B is an especially instructive data stream in that it contains nearly all the sensor non-idealities, scene complexities and processing challenges that can be found in the dataset as a whole. It will therefore be used repeatedly in this work to illustrate many of the event-based processing problems and solutions presented in this work.

B. Labeling the Dataset

Generating ground truth labeling for real-world event-based space imaging data is a non-trivial task. Even when the true position, velocity, size and luminance of all targets in the field of view of the sensor are known, their detection by the event-based sensor is far from guaranteed. The clearest demonstration of this problem is in cases where within the same recording, the biases and circuitry of the camera are configured optimally for one event polarity such that space objects are clearly visible in one polarity but produce zero events in the other polarity. In these and analogous situations, the use of any ground truth labels from external information sources such as the co-collection recordings from a conventional camera (or a sky catalog or database as used in [19]), would likely result in an incorrect evaluation of any event-based algorithm operating on the actual observed real-world event stream.

Such a comparison was recently performed in [18] where the DAVIS event-based sensor was estimated to have lower sensitivity relative to the global shutter CMOS sensor. This difference in relative sensitivity was measured via the limiting magnitude, defined as the faintest magnitude of a celestial body that is detectable. The event-based sensor was estimated to be less sensitive than the CMOS sensor by a magnitude between 1.32 and 1.78. While these results do not necessarily generalize to other event-based sensor configurations and recording environments, they do highlight the general problem of using external labels to evaluate event-based data. This work evaluates the tracking algorithms and not the performance of the sensor. Hence, ground truth from a different sensor type (such as a conventional CCD, global shutter CMOS, etc) does not directly allow us to predict the accuracy of the tracker. Thus we require a ground truth related to the events generated from the camera, and not from external label sets.

For this reason, to generate a more appropriate label set for the observed event streams, hand-labeling of the data was performed and a committee-of-experts approach was used to determine the ground truth labels. Thus expert human labeling of the highly noisy event stream is here set as a benchmark against which proposed event-based algorithms are tested. In the Supplementary Material section, the labeling procedure is described in and the quality of the expert human labeling procedure is verified using an artificially generated space imaging dataset in which ground truth labels are analytically defined.

C. Measuring Sensitivity, Specificity and Informedness

The data from the event-based sensors have a high temporal resolution, with the event rate varying for each pixel and dependent on the activity in the scene. A robust method is required for measuring how well a given event stream sampled at 1 MHz matches the frame-based expert labeled dataset which is sampled at a much slower 1 kHz. This accuracy measure must also be invariant to the extreme differences in event rates produced by different recording conditions. The measure must also assess the highly noisy raw events of the sensor in the same manner as the extremely sparse detection and tracking output event streams. To achieve this, we propose
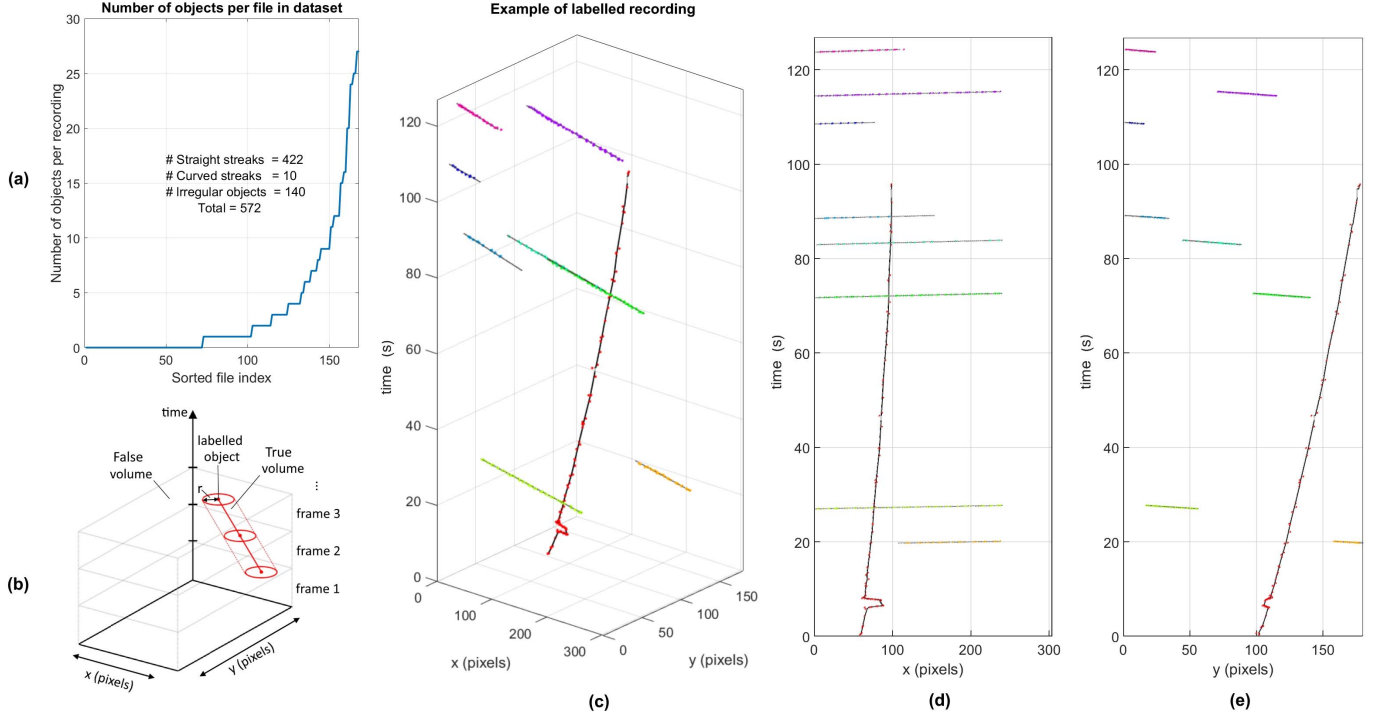
Fig. 3. Dataset labeling. (a) Total number of sub-types and the number of objects per recording in the dataset shown as a sorted list. The 422 straight streaks represent objects that exhibit zero acceleration and move in a straight line in space-time. The 10 curved streaks were objects observed to exhibit uniform acceleration and 140 irregular objects exhibited non-uniform acceleration while in the field of view. (b) Illustrates the method used for calculating sensitivity and specificity of event volumes around labeled data points. A volume of radius r around a line connecting the labeled points marks the boundary between true and false volumes. The volumes are sliced at 10ms intervals. The event density of each sub-region designates its volume as a positive or negative volume depending on whether it is above or below the mean density of the recording as a whole. Panels (c), (d) and (e) show the expert labeled objects in the SL-8 R/B recording in a dimetric projection and across the x and y-axis respectively. The slow-moving object is the SL-8 R/B as it is being tracked by the telescope. The curve in the trajectory of SL-8 R/B is an artifact of the open-loop tracking of the mount. The rapid jerk-like movement around 8 seconds is due to mechanical movements caused by the human operator. The high-speed streaks are background stars.

a metric based on relative event density in the event stream. This method assigns spatio-temporal volume slices to either a positive or a negative state. These states are then compared to the labeled dataset which indicates whether the corresponding volume contains target objects (True), or not (False).

As shown in Figure 3(b), for each frame, the spatio-temporal volume slice surrounding the trajectory of a labeled object by radius r is designated as True and the spatio-temporal volume outside this region and in frames with no labeled object is designated as False. If for any spatio-temporal volume, the event density is above the global event density of the full recording, then the volume is activated as positive. Conversely, the volume is designated as negative if the event density in the volume falls below the global event density of the full recording (i.e. if there are relatively fewer events per $pixel^2$/second in the local volume slice than the total number of events divided by the total recording duration multiplied by the sensor area).

In this way, event streams with drastically different noise profiles and event densities can be directly compared and evaluated by calculating the mean True Positive ($\overline{TP}$), True Negative ($\overline{TN}$), False Positive ($\overline{FP}$) and False Negative ($\overline{FN}$) volumes of each recording. Using these volume-based measures, the event-based sensitivity and specificity of a

particular event stream can be calculated using:

$$Sensitivity = \overline{TP}/(\overline{TP} + \overline{FN}) \quad (1)$$
$$Specificity = \overline{TN}/(\overline{TN} + \overline{FP}) \quad (2)$$

Using these measures, the informedness, or the Bookmaker Informedness of an event stream can be calculated using (3). Informedness, which is a generalization of the Youden's J statistic, provides a single statistic that captures the performance of a binary diagnostic test [25], and "quantifies how informed a predictor is for the specified condition, and specifies the probability that a prediction is informed in relation to the condition (versus chance)" [26].

$$Informedness = Sensitivity + Specificity - 1 \quad (3)$$

Informedness seeks to avoid biases of other common statistics, such as accuracy and precision, which are susceptible to population prevalence and label bias. This makes informedness an accuracy measure suitable for the highly imbalanced EBSI datasets in which the vast majority of the spatio-temporal volumes are labeled as False regions.

As an example, the event density activated volume statistics for the SL-8 R/B recording are detailed in Table I showing clear differences between the raw ON and OFF event streams.

Polarity	Sensitivity	Specificity	Informedness	# Events (ke)
ON Events	0.69	0.68	0.37	1770
OFF Events	0.65	0.79	0.43	360

D. Event Pre-Processing

The algorithms presented in this work are entirely event-based with all components from the sensors to the detectors and trackers operating entirely in the event-based domain. The microsecond time resolution of the sensor is therefore maintained throughout the processing chain. A brief explanation of event-based processing is provided below.

Following the notation in [17], events generated at the sensor, e_i can be described mathematically as:

$$e_i = [\boldsymbol{x}_i, t_i, p_i]^T \tag{4}$$

where i is the index of the event, t_i is the time at which the event occurred and, $\boldsymbol{x}_i$ is the spatial address of the source pixel as defined by (5).

$$\boldsymbol{x}_i = [x_i, y_i]^T \tag{5}$$

where x_i and y_i are the column and row address of the source pixel corresponding its physical location on the sensor.

In this work, the ON and OFF polarity of the events, $p_i \in \{-1, 1\}$ which denote whether the log intensity has increased or decreased is not used. Instead, the ON and OFF recordings are treated as independent recordings. This is due to the significant difference in the SNR characteristic of the ON and OFF events in the recordings. The timestamp t_i has a temporal resolution of $1\mu s$ and is applied to the event in hardware within the event-based sensor.

Event-based algorithms require memory of recent events and can be generated via a range of methods that were investigated in [27]. The method used in this work and one which typically outperforms other approaches is the exponentially decaying event time surface.

In this approach, the matrix $\boldsymbol{T}$ contains the time-stamp of the most recent event at each pixel. At each event e_i, the entry in $\boldsymbol{T}$ at location $\boldsymbol{x}_i$ is updated by the event time stamp t_i with all other entries in $\boldsymbol{T}$ remaining unchanged as described by (6).

$$\boldsymbol{T} : \boldsymbol{x}_i \to \boldsymbol{T}(\boldsymbol{x}_i) = t_i \tag{6}$$

Using $\boldsymbol{T}$, an exponentially decaying time surface matrix $\boldsymbol{S}$ can be calculated for any sub-region of the sensor via (7).

$$\forall \boldsymbol{x_k} \in X : S(\boldsymbol{x}_k) = e^{(T(x_k)-t_i)/\tau} \tag{7}$$

where X is the matrix containing all pixel addresses on the sensor, k is the index of the sensor pixel, and τ is the decay constant of the time surface. All pixels in $\boldsymbol{T}$ and $\boldsymbol{S}$ are initialized to $-\infty$ and 0 respectively. In this work, a value of

$\tau = 0.4$ seconds was chosen as the decay constant. The value of τ was chosen heuristically and found to maximize accuracy of the system on the dataset. A more detailed investigation of this parameter is provided in the Supplementary Material Section.

As shown in Figure 4(a), after updating the time surface, a Region Of Interest (ROI) of size $D \times D$ around the event $e_i = [x_i, y_i, t_i, p_i]^T$ is selected for processing. The $\boldsymbol{ROI}_i$ associated with event e_i is defined as:

$$\boldsymbol{ROI}_i = S(x_i + \boldsymbol{u}_x, y_i + \boldsymbol{u}_y) \tag{8}$$

where $\boldsymbol{u}_x = [-R : R]$ and $\boldsymbol{u}_y = [-R : R]$ subject to the constraint:

$$\sqrt{x^2 + y^2} \le R, \quad \forall x \in \boldsymbol{u}_x, \ \forall y \in \boldsymbol{u}_y \tag{9}$$

Thus the $\boldsymbol{ROI}_i$ generated by event e_i is defined as a disc containing the time surface values S at time t_i from all pixels within a distance R to the location of the current event e_i. This $\boldsymbol{ROI}_i$ is then processed by a surface activation test which is defined as:

$$L < \sum_{x=x_i-R}^{x_i+R} \sum_{y=y_i-R}^{y_i+R} \left((x, y) > \Phi \right) \tag{10}$$

where Φ is the event activation time interval, L is the number of activated pixels required and x and y are subject to the distance constraint given in (9). Thus, if the number of recently activated pixels on the time surface within a disc of radius R around the current event e_i is above L, then the ROI is accepted. Here recency is defined as a pixel that has received an event within Φ seconds. This surface activation test effectively acts as a noise filter and is a generalization of the nearest neighbor filter described in [28] where R was selected as 1. Unlike in terrestrial applications, event filtering of event-based space imaging data requires large ROIs and long time windows Φ for the surface activation test. This is due to the significantly lower SNR in space imaging data and the similarity of the most challenging targets (small dim geostationary satellites) to background noise.

E. Feature Detection

In the next stage of processing, a valid ROI is converted to an angular activation vector $\boldsymbol{\Lambda}$, generated by multiplying the ROI with each of N rotated half-bar templates shown in Figure 4(d).

The half-bar templates are designed to be triggered by events at the tip of a moving streak. By matching the ROI of each new event with rotated versions of this half bar template, a highly selective feature detector can be designed that most strongly activates for events at the tip of a streak generated by a target moving in a particular direction. The template consists of a bar of length $R+1$ and three pixels wide with a magnitude of one. While setting the bar width parameter at three pixels appears an arbitrary choice, this pattern was found heuristically to produce the best performance across a wide range of object sizes and ROI sizes. This is likely due to the three-pixels bar being close to the size of the smallest resolvable streak in the space imaging dataset. In addition to resembling the smallest
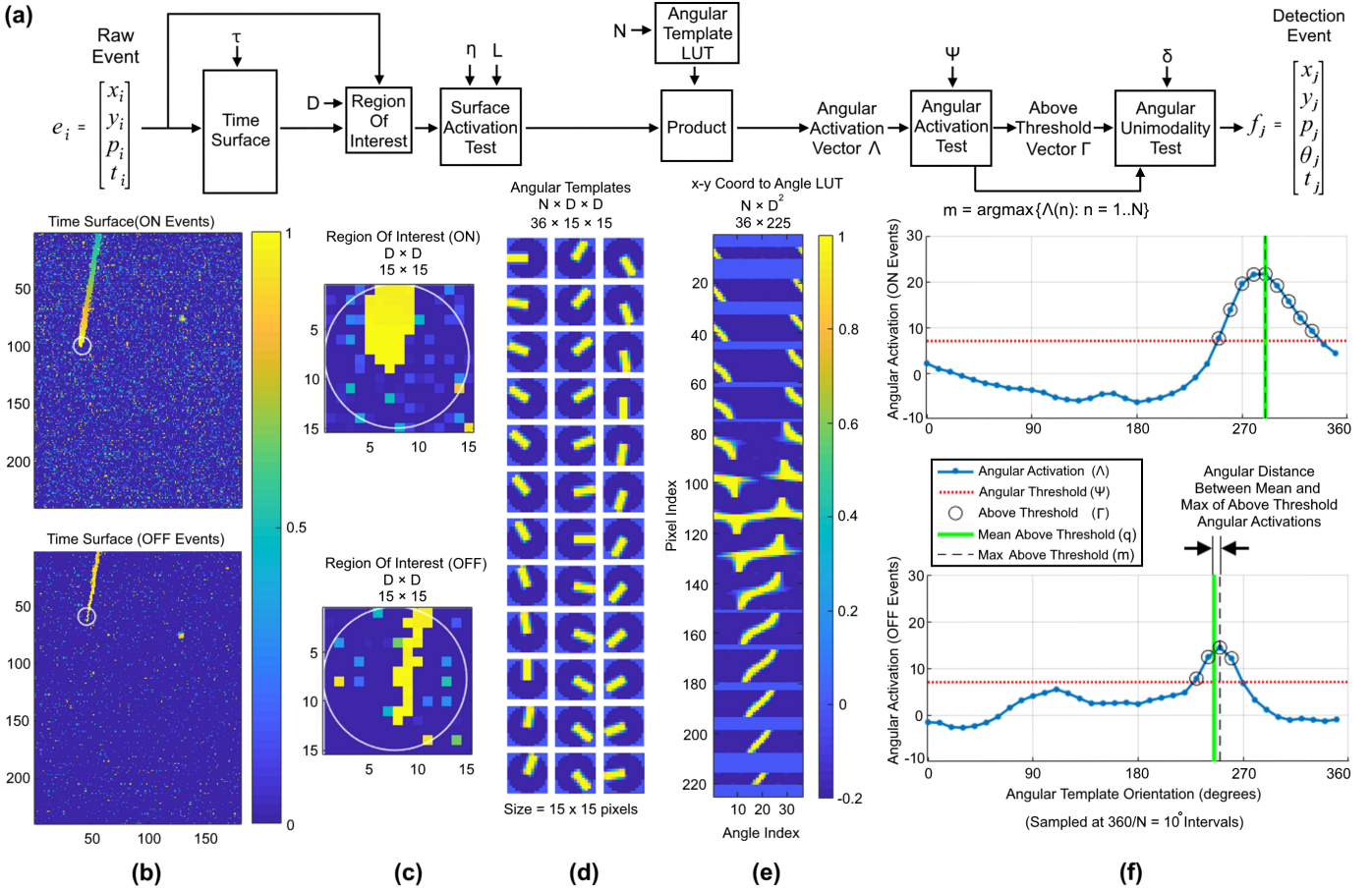
Fig. 4. **Orientation invariant space object detection algorithm and signals at each stage of processing.** Panel (a) shows the block diagram of the algorithm whereby a sequence of increasingly refined tests operate on an event e_i. If the event passes all test a detection event f_j is generated. (b) Shows an instance of the ON and OFF time surface for the SL-8 R/B recording. Note the different noise levels and target sensitivity of the two polarities. (c) Shows the local 15×15 Region Of Interest (ROI_i) around the current event e_i for each polarity. (d) $N = 36$ Streak templates rotated at 10-degree intervals to calculate the angular activation of the ROI. (e) a stored Look Up Table (LUT) converts the ROI values to an angular activation vector Λ through a single vector-matrix multiplication operation. (f) the resultant angular activation is shown for each of the ON and OFF ROIs. If Λ exceeds the angular threshold Ψ, it passes the angular activation test after which the circular mean index q of all angles above the angular threshold Γ, is calculated. If the distance ζ between q to the maximally activated angle m is below the threshold δ the event passes the angular unimodality test resulting in a detection event f_j. Note that for visual simplicity, both the static and the dynamic angular activation thresholds are made static and equal with $\Psi = \Psi_i = 7$.

resolvable streak, the half bar template also acts as a more general orientation-selective feature detector that can detect the tips of streaks which are much thicker than the three-pixel half bar template. Furthermore, when the relative activation of the rotated templates is used as a feature, the pattern of template activations also selectively detects stationary objects of arbitrary size. The response of the feature detector to a wide range of stimuli is demonstrated in the Supplementary Material Section.

Outside of the bar, the rest of the template has a negative magnitude of $s = -0.2$ to penalize activation from noise events. In practice the N, $D \times D$ templates are re-arranged into an $N \times D^2$ Look Up Table (LUT) and the $D \times D$ ROI vector is rearranged to a $D^2 \times 1$ vector. This vectorization operation is here denoted as the vec() function as described in Equation 11. The multiplication of the ROI vector and the LUT results in an $N \times 1$ Λ vector as illustrated in Figure 4(f).

$$\Lambda_i = LUT \cdot \text{vec}(ROI_i) \qquad (11)$$

Optimization of the subsystem that converts the ROI event timestamps to the angular activation vector Λ is critical in the performance of the proposed algorithm. The calculation of the angular activation vector is, regardless of the implementation environment, significantly more computationally expensive than that of the previous surface activation test, but unlike subsequent stages which are also computationally intensive, this operation is performed on the majority of the events from the camera. This makes the calculation of the angular activation vector the most computationally expensive step relative to the number of events processed, making it a computational bottleneck of the algorithm. This aspect of the algorithm is investigated in more detail in Section III-B.

The resulting angular activation vector Λ is compared to an angular activation threshold of Ψ and if no element of Λ exceeds Ψ the angular activation test fails, otherwise the variable m, which is defined as the index of the highest activated element of Λ, is passed to the next stage of processing along with the vector Γ which contains the index of all elements in Λ above threshold Ψ.

The threshold used for calculating $\boldsymbol{\Gamma}$ can be chosen as a static parameter Ψ, or as a dynamic threshold Ψ_i which is defined as a scalar factor W of the difference between the minimum and maximum of elements of $\boldsymbol{\Lambda}_i$ as described in (12). The use of a static threshold Ψ, simplifies the algorithm implementation whereas the use of a dynamic threshold can provide greater robustness to noise events. Except where stated, in this work, the dynamic method is used with $W = 0.5$.

$$\Psi_i = W(\max(\boldsymbol{\Lambda}_i) + \min(\boldsymbol{\Lambda}_i)) \qquad (12)$$

The angular activation test serves as a filter that removes all ROIs with uniform activation in polar coordinates. This filter is useful in removing events not associated with a streak on the time surface. However, this filter does not distinguish between events that are on or near a streak and those at the streak's tip. To further extract these later events, a statistical unimodality test must be applied to the angular activation vector $\boldsymbol{\Lambda}$. Previously proposed unimodality tests include fitting of parametric mixture models [29], as well as non-parametric tests such as the commonly used Dip Test [30], use of kernel density estimates [31] and recursive methods based on unimodal template transformations [32]. While these methods can provide robust solutions to the unimodality test, they are too computationally expensive for the streak tip detection application where thousand of events must potentially be processed per second possibly by a low power processor on a space-based platform. We therefore propose a highly simplified hardware amenable circular unimodality test for our event-based application. This unimodality test we simply measure the angular distance between the maximum value in $\boldsymbol{\Lambda}$ and the angular mean of all elements above a threshold Ψ_i.

As shown in Figure 4(a), the unimodality block takes as input m_i which is the index of the largest element of $\boldsymbol{\Lambda}_i$. It also takes as input a vector $\boldsymbol{\Gamma}_i$ containing the index of all elements in $\boldsymbol{\Lambda}_i$ higher than Ψ_i: $\boldsymbol{\Gamma}_i = \{n\}$ s.t. $\boldsymbol{\Lambda}_i(n) > \Psi_i$. The unimodality block then outputs a stream of filtered events f_j which have passed the unimodality test. As plotted in Figure 4(f), the unimodality block performs its test by calculating q_i which is the circular mean of $\boldsymbol{\Gamma}_i$ and testing whether the angular distance ζ_i between q_i and m_i is below a threshold δ. In this work, a value of $\delta = 2$ was used equivalent to a tolerance of 20° in the angular distance. In the Supplementary Material Section we detail the behavior of the angular unimodality detection in response to common EBSI input as well as alternative methods for calculating q_i. The distance ζ_i represents how far the peak angular activation is from the mean. This makes ζ_i a simplified yet robust measure of the unimodality of the angular activation vector $\boldsymbol{\Lambda}_i$.

Despite its simplicity, this unimodality test is highly selective and performs remarkably well at extracting space targets from the observed event-based space event streams while being robust to noise, object velocity, object size and orientation. If the event e_i passes this angular unimodality test, it is augmented with the mean orientation variable $\theta_i = q_i$ and results in an output detection event f_j as described by Algorithm 1, and illustrated in Figure 4(a). Note that

---

Algorithm 1 Feature Detection

Require: $\boldsymbol{e}_i = [x_i, y_i, p_i, t_i]^T, i \in \mathbb{N}$
Ensure: $\boldsymbol{f}_j = [x_j, y_j, p_j, \theta_j, t_j]^T, j \in \mathbb{N}$
  $j \Leftarrow 0$
  for every event $\mathbf{e_i}$ **do**
    $\boldsymbol{T}_i(x_i, y_i) \Leftarrow t_i$
    $\boldsymbol{ROI_i} \Leftarrow \boldsymbol{S}_i(x_i + \mathbf{u_x}, y_i + \mathbf{u_y})$ via (8)
    if $\boldsymbol{ROI}_i$ contains recent events verifying (10) **then**
      $\boldsymbol{\Lambda}_i \Leftarrow \boldsymbol{LUT} * vec(\boldsymbol{ROI}_i)$
      if $\max(\boldsymbol{\Lambda}_i) > \Psi$ **then**
        $\boldsymbol{\breve{\Lambda}}_i \Leftarrow \left[\boldsymbol{\Lambda}_i\big((N/2+1):N\big), \boldsymbol{\Lambda}_i\big(1:N/2\big) \right]$
        $\Psi_i \Leftarrow W(\max(\boldsymbol{\Lambda}_i)+\min(\boldsymbol{\Lambda}_i))$
        $\boldsymbol{\Gamma}_i \Leftarrow \{n\}$ s.t. $\boldsymbol{\Lambda}_i(n) > \Psi_i$
        $\boldsymbol{\breve{\Gamma}}_i \Leftarrow \{n\}$ s.t. $\boldsymbol{\breve{\Lambda}}_i(n) > \Psi_i$
        $q_i \Leftarrow 1/G \sum_{h=1}^{G} \boldsymbol{\Gamma}_i(h)$
        $\breve{q}_i \Leftarrow 1/G \sum_{h=1}^{G} \boldsymbol{\breve{\Gamma}}_i(h)$
        $m_i \Leftarrow \text{argmax}_n \boldsymbol{\Lambda}_i(n)$
        $\breve{m}_i \Leftarrow \text{argmax}_n \boldsymbol{\breve{\Lambda}}_i(n)$
        if $\min(|m_i - q_i|, |\breve{m}_i - \breve{q}_i|) < \delta$ **then**
          $j \Leftarrow j + 1$
          $\theta_j \Leftarrow q_i$
          $\boldsymbol{f}_j \Leftarrow [x_i, y_i, p_i, \theta_i, t_i]^T$
        end if
      end if
    end if
  end for

TABLE II
EVENT DENSITY ACTIVATED VOLUME BASED STATISTICS FOR MEASURING THE PERFORMANCE OF THE FEATURE DETECTION EVENTS f_j. THE STATISTICS CALCULATED ARE FROM THE DETECTION EVENTS GENERATED USING THE SL-8 R/B RECORDING WHOSE DATA STREAM IS ILLUSTRATED IN FIGURE 5

Polarity	Sensitivity	Specificity	Informedness	# Events (ke)
ON Events	0.66	0.99	0.65	22.20
OFF Events	0.60	0.98	0.58	15.98

in Algorithm 1, G denotes the number of above-threshold elements in $\boldsymbol{\Lambda}_i$ and is therefore the length of the vector Γ_i.

To illustrate in detail the behavior of the feature detector on a real space imaging data stream, the detection event stream generated from the SL-8 R/B recording is shown in Figure 5 with associated event-based statistics shown in Table II. This shows that whilst the sensitivity of the event stream is slightly lower than in Table I, the much higher specificity results in significantly greater informedness than the raw events.

As shown in Figure 5, due to different noise characteristics and sensitivity of the ON and OFF polarities, significantly different detection event streams are generated from each of the polarities. Note that the high-velocity streaks exhibit discrete orientation distributions whereas the slow-moving object being tracked in the field of view generates a uniform distribution $\theta \in [0, 2\pi)$ since the latter generates a circular image on the time surface triggering detection events that are approximately equally in all directions.

As SL-8 R/B leaves the field of view, the uniform distribution also fades away, leaving only the orientation traces from the high-speed streaks. Also note a 180 degree shifted angular 'shadow' generated by high-speed targets especially
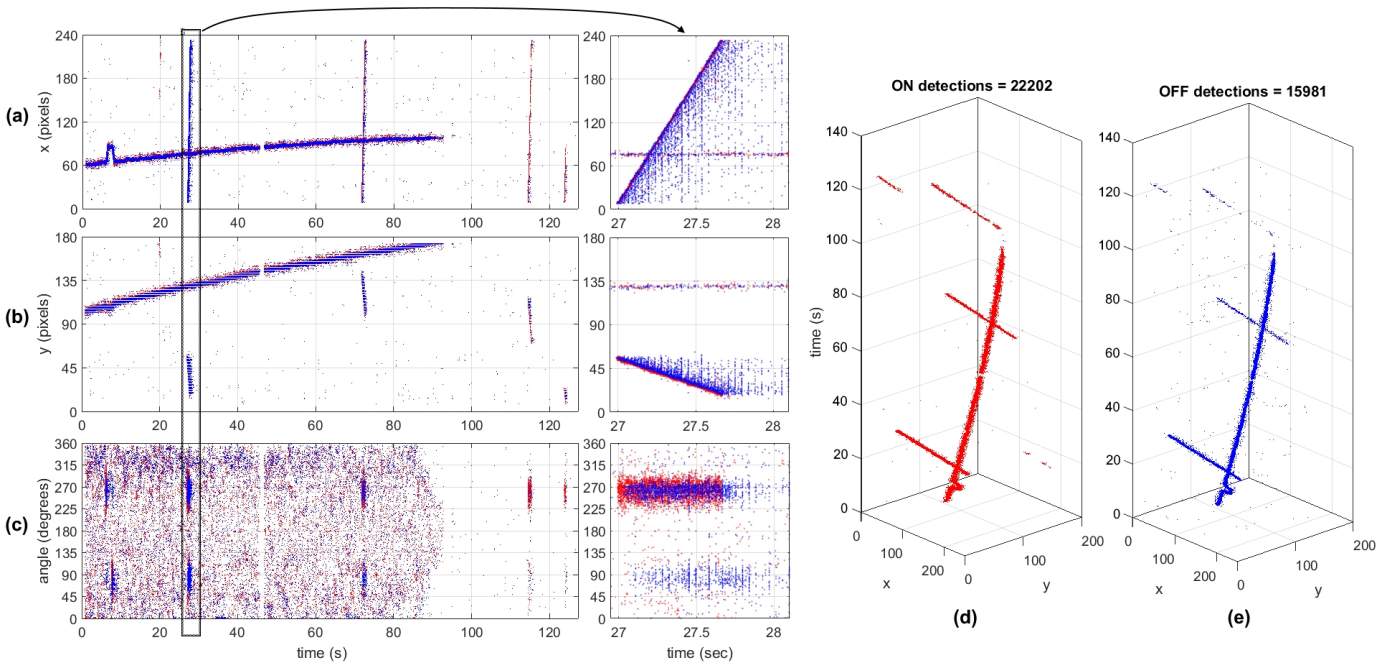
Fig. 5. **Feature detection events from the SL-8 R/B recording.** The panels on the left in (a), (b) and (c) show the x and y location and orientation of the detection events respectively over time for the ON (red) and OFF (blue) detection events. The dashed rectangle marks the time interval around the detection of a high-speed object shown in the close-up right-hand side panels. The panels in (c) show the dominant orientation of the detection events, based on the mean index of above-threshold templates $\theta_i = q_i$. Here, SL-8 R/B forms the long, slightly curved line. Note that the temporal event bands in the close-up panels are artifacts caused by the data interface. Panels (d) and (e) show the dimetric projections of the ON and OFF detection events respectively.

for the noisier OFF events. These false detections, which are pointed in the opposite direction to the true angle of the object's trajectory, are due to late-triggered events along the tail of the streak. These detections have an equal likelihood of being oriented forward or backward. As shown in the close-up panels in Figure 5(a), (b) and (c), even whilst using high sensitivity parameter settings, these false detection events are significantly less frequent and more dispersed in space and time than the detections made at the tip of the streak generated by the fast-moving object and as such can be readily filtered by an event-based tracker.

The proposed feature detector can be viewed as a highly refined filter designed to remove noise events passed to it from the upstream surface activation filter. The far sparser output event stream of the streak detection events can then be passed to a more computationally intensive event-based tracker. The event-based tracker in turn can be viewed as an even more restrictive filter capable of removing spurious detection events not associated with other nearby detection events of the same orientation and velocity. When viewed in this way, as a series of increasingly refined event-filters, the requirement of preserving true events generated by true targets outweighs the value of removing noise events at earlier filtering stages as long as the noise events will eventually be removed by a downstream filter. Thus, as long as the final filter can remove all remaining noise events, the only penalty to permissive parameter settings at the upstream stages is in the increased processing time of the latter filters. This motivates a conservative parameter selection regime which at the feature detection stage involves the selection of parameters that generate a significant level of false positive detection events.

F. Event-Based Tracking

The event-based tracking method used in this work continually generates, updates, and removes hypotheses in an event-based manner. The state of the hypotheses is modeled as a population of leaky integrate and fire neurons whilst the hypotheses trajectories are updated using a sequential least-squares fitting algorithm operating on incoming detection events.

Each active tracked object is modeled as a neuron containing a membrane potential which decays over time and is incremented via detection events $\boldsymbol{f}_j$ assigned to it as detailed later in this section. The membrane potential represents the level of recent observations of the object. If the membrane potential reaches the activation potential M_A, the object is activated. Alternatively, if the membrane potential reaches zero the object is deleted.

$$M_k^{(o)} = \begin{cases} M_{k-1}^{(o)} - \gamma\,(t_j - t_{k-1}) + 1, & \text{if } \boldsymbol{f}_j \text{ is assigned to } \boldsymbol{H}_k^{(o)}. \\ 0, & \text{if } M_{k-1}^{(o)} = 0. \\ M_{k-1}^{(o)} - \gamma\,(t_j - t_{k-1}), & \text{otherwise.} \end{cases}$$
$$(13)$$

where $\boldsymbol{H}_k^{(o)}$ is the kth observation of the oth object, $M_k^{(o)}$ is the membrane potential of $\boldsymbol{H}_k^{(o)}$ at t_k and γ is the decay constant for the membrane potential. If the object activation variable $M_k^{(o)}$ reaches the activation threshold M_A, the object $\boldsymbol{H}_k^{(o)}$ is deemed a true tracked object. A variable $A_k^{(o)}$ tracks the activation level of the object until it reaches M_A and $A_k^{(o)}$ reaches 1. Thereafter $A_k^{(o)}$ remains at 1, permanently indicating the activation of the object regardless of the value of the membrane potential $M_k^{(o)}$.

This behavior is described by (14). In addition to indicating the activation of the object, $A_k^{(o)}$ will be used to weigh the angular distance relative to the spatial coordinates and as such plays an important role in reducing the weight of the angular distance in the earlier stages of tracking where the object's estimated angle tends to be unreliable.

$$A_k^{(o)} = \begin{cases} M_k^{(o)}/M_A, & \text{if } M_k^{(o)} < M_A \text{ and } A_k^{(o)} < 1. \\ 1, & \text{otherwise.} \end{cases} \quad (14)$$

where k denotes the number of previous observations assigned to the nth object.

Given the jth detection event $\boldsymbol{f}_j = [x_j, y_j, p_j, \theta_j, t_j]^T$, $\boldsymbol{z}_j$ is defined as the vector containing the position and angular information excluding the polarity and timestamp entries:

$$\boldsymbol{z}_j = [x_j, y_j, \theta_j]^T \quad (15)$$

The position and velocity of each active object n in space and time, at the kth observation, is defined as:

$$\boldsymbol{H}_k^{(o)} = [\hat{z}_k, \boldsymbol{b}_k, p_k, t_k]^T, \quad o \in \mathbb{N}, \ k \in \mathbb{N} \quad (16)$$

where $\hat{z}_k = [\hat{x}_k, \hat{y}_k, \hat{\theta}_k]^T$ and $\boldsymbol{b}_k = [d\hat{x}/dt_k, d\hat{y}/dt_k, d\hat{\theta}/dt_k]^T$ as estimated via Algorithm 3.

The predicted object position at time t_j is determined using:

$$[\hat{x}_k, \hat{y}_k, \hat{\theta}_k]^T = [\hat{x}_{k-1}, \hat{y}_{k-1}, \hat{\theta}_{k-1}]^T + \boldsymbol{b}_{k-1}(t_j - t_{k-1}) \quad (17)$$

where $\boldsymbol{b}_{k-1} = [dx/dt_{k-1}, dy/dt_{k-1}, d\theta/dt_{k-1}]^T$ as estimated via Algorithm 3.

When estimating the distance of a new detection event to each object $\boldsymbol{H}_k^{(o)}$, the weight of the angular distance $\boldsymbol{w}_k^{(o)}$ in θ, relative to the distance in x and y is proportional to each object's previous speed and the activation measure $A_k^{(o)}$. Thus, the faster the velocity of an object, the higher the weight of the angular distance is with respect to the positional distance. Objects moving at close to zero velocity are assigned near-zero weight since the detection will be oriented at random, whereas objects moving at high speed have sharp clearly distinguishable angles:

$$\boldsymbol{w}_k^{(o)} = V\left(\sqrt{(dx/dt_k^{(o)})^2 + (dy/dt_k^{(o)})^2}\right) A_k^{(o)} \quad (18)$$

where V is a scaling factor which in this work was selected as $V = 0.1$ pixel/degree.

The distance between a new detected event $\boldsymbol{f}_j$ and the predicted position of each active object $\boldsymbol{H}_k^{(o)}$ at t_j is defined as:

$$d_k^{(o)} = \sqrt{(x_j - \hat{x}_k^{(o)})^2 + (y_j - \hat{y}_k^{(o)})^2 + \boldsymbol{w}_k^{(o)}(\theta_j - \theta_k^{(o)})^2} < d_{max} \quad (19)$$

where d_{max} is the threshold acceptable distance to the detected event.

In summary, at each detection event $\boldsymbol{f}_j$, the weighted Euclidean distance between the event and the projected x, y, θ position of every object $\boldsymbol{H}_k^{(o)}$ with an active membrane potential $M_k^{(o)} > 0$ at time t_i, is measured. This distance is

---

Algorithm 2 Detection Assignment

Require: $\boldsymbol{f}_j = [x_j, y_j, p_j, \theta_j, t_j]^T, i \in \mathbb{N}$
Ensure: $\boldsymbol{H}_k^{(o)}$, $\boldsymbol{M}_k^{(o)}$ and $A_k^{(o)}$ as defined by (16), (13) and (14)
  for every feature event $\boldsymbol{f}_j$ **do**
    for every object $\boldsymbol{H}_k^{(o)}$ where $\boldsymbol{M}_k^{(o)} > 0$ **do**
      compute predicted postition $[x_k^{(o)}, y_k^{(o)}, \theta_k^{(o)}]$ using (17)
    end for
    if $\exists$ any object satisfying (19) **then**
      $n \Leftarrow \arg\min_m d_k^{(m)}$ s.t. $\boldsymbol{H}_k^{(m)}$ verifies (19)
      $\boldsymbol{M}_k^{(o)} \Leftarrow \boldsymbol{M}_{k-1}^{(o)} + 1$
      Update $\boldsymbol{H}_k^{(o)}$ with $[x_{k+1}^{(o)}, y_{k+1}^{(o)}, \theta_{k+1}^{(o)}]^T$ using Algorithm 3.
    end if
    for every object $\boldsymbol{H}_k^{(o)}$ with $\boldsymbol{M}_k^{(o)} > 0$ **do**
      Update $\boldsymbol{M}_k^{(o)}$ using (13)
    end for
  end for

Algorithm 3 Object Velocity Estimation Algorithm

Require: $\boldsymbol{z}_j$ from $\boldsymbol{f}_j$, $j \in \mathbb{N}$ using (15) and $\boldsymbol{z}_k$ from $\boldsymbol{H}_k^{(o)}$, $k \in \mathbb{N}, o \in \mathbb{N}$ verifying (19)
Ensure: $\boldsymbol{H}_k^{(o)}$ as defined by (16)
  for every feature event $\boldsymbol{f}_j$ **do**
    $t_k \Leftarrow t_j, \boldsymbol{z}_k \Leftarrow \boldsymbol{z}_j$
    if $k = 1$ **then**
      $\bar{t}_k \Leftarrow t_k, \bar{z}_k \Leftarrow \boldsymbol{z}_k$
      $\sigma_k^2 \Leftarrow 0, \Sigma_k^2 \Leftarrow 0$
      $c_k^2 \Leftarrow 0, C_k^2 \Leftarrow 0$
      $\boldsymbol{b}_k \Leftarrow 0, \hat{t}_k \Leftarrow t_k, \hat{z}_k \Leftarrow \boldsymbol{z}_k$
    else if $k > K$ **then**
      $\bar{t}_k \Leftarrow \bar{t}_{k-1} + (t_k - t_{k-K})/K$
      $\bar{z}_k \Leftarrow \bar{z}_{k-1} + (\boldsymbol{z}_k - \boldsymbol{z}_{k-K})/K$
      $\sigma_k^2 \Leftarrow \sigma_{k-1}^2 + (t_k - \bar{t}_{k-1})(t_k - \bar{t}_k) - (t_{k-K} - \bar{t}_{k-1})(t_{k-K} - \bar{t}_k)$
      $c_k^2 \Leftarrow c_{k-1}^2 + (\boldsymbol{z}_k - \bar{z}_{k-1})(t_k - \bar{t}_k) - (\boldsymbol{z}_{k-K} - \bar{z}_{k-1})(t_{k-K} - \bar{t}_k)$
      $\Sigma_k^2 \Leftarrow \sigma_k^2/(K-1)$
      $C_k^2 \Leftarrow c_k^2/K$
      $\boldsymbol{b}_k \Leftarrow C_k^2/\Sigma_k^2$
      $\hat{t}_k \Leftarrow \bar{t}_k + (t_k - \bar{t}_{k+K})/2$
      $\hat{z}_k \Leftarrow \bar{z}_k + \boldsymbol{b}_k(t_k - \bar{t}_{k+K})/2$
    else
      $\bar{t}_k \Leftarrow \bar{t}_{k-1} + (t_k - \bar{t}_{k-1})/k$
      $\bar{z}_k \Leftarrow \bar{z}_{k-1} + (\boldsymbol{z}_k - \bar{z}_{k-1})/k$
      $\sigma_k^2 \Leftarrow \sigma_{k-1}^2 + (t_k - \bar{t}_{k-1})(t_k - \bar{t}_k)$
      $c_k^2 \Leftarrow c_{k-1}^2 + (\boldsymbol{z}_k - \bar{z}_k)(t_k - \bar{t}_k)$
      $\Sigma_k \Leftarrow \sigma_k^2/(k-1)$
      $C_k^2 \Leftarrow c_k^2/k$
      $\boldsymbol{b}_k \Leftarrow C_k^2/\Sigma_k^2$
      $\hat{t}_k \Leftarrow \bar{t}_k + k(t_k - \bar{t}_1)/2(k-1)$
      $\hat{z}_k \Leftarrow \bar{z}_k + \boldsymbol{b}_k k(t_k - \bar{t}_1)/2(k-1)$
    end if
    Update $\boldsymbol{H}_k^{(o)}$ with $t_k, \boldsymbol{z}_k, \hat{t}_k$ and $\hat{z}_k$
  end for

---

then compared to the threshold d_{max}. The detection event is assigned to the closest object with a distance below d_{max}. If no object is within d_{max} of the current detection event, a new object $\boldsymbol{H}_1^{(o+1)})$ is created. This algorithm is described by Algorithm 2.

To estimate the position of each hypothesis $\boldsymbol{H}_k^{(o)}$ at the time of each detection event $\boldsymbol{f}_j$, a sequential least-squares method is implemented involving the sequential calculation of the ratio of the covariance of the position and timing of the object over the variance of the timing. In this event-based approach, the covariance and variance measures are calculated online in a sequential manner. Each measure is calculated using a fixed rolling window of length K. This online approach
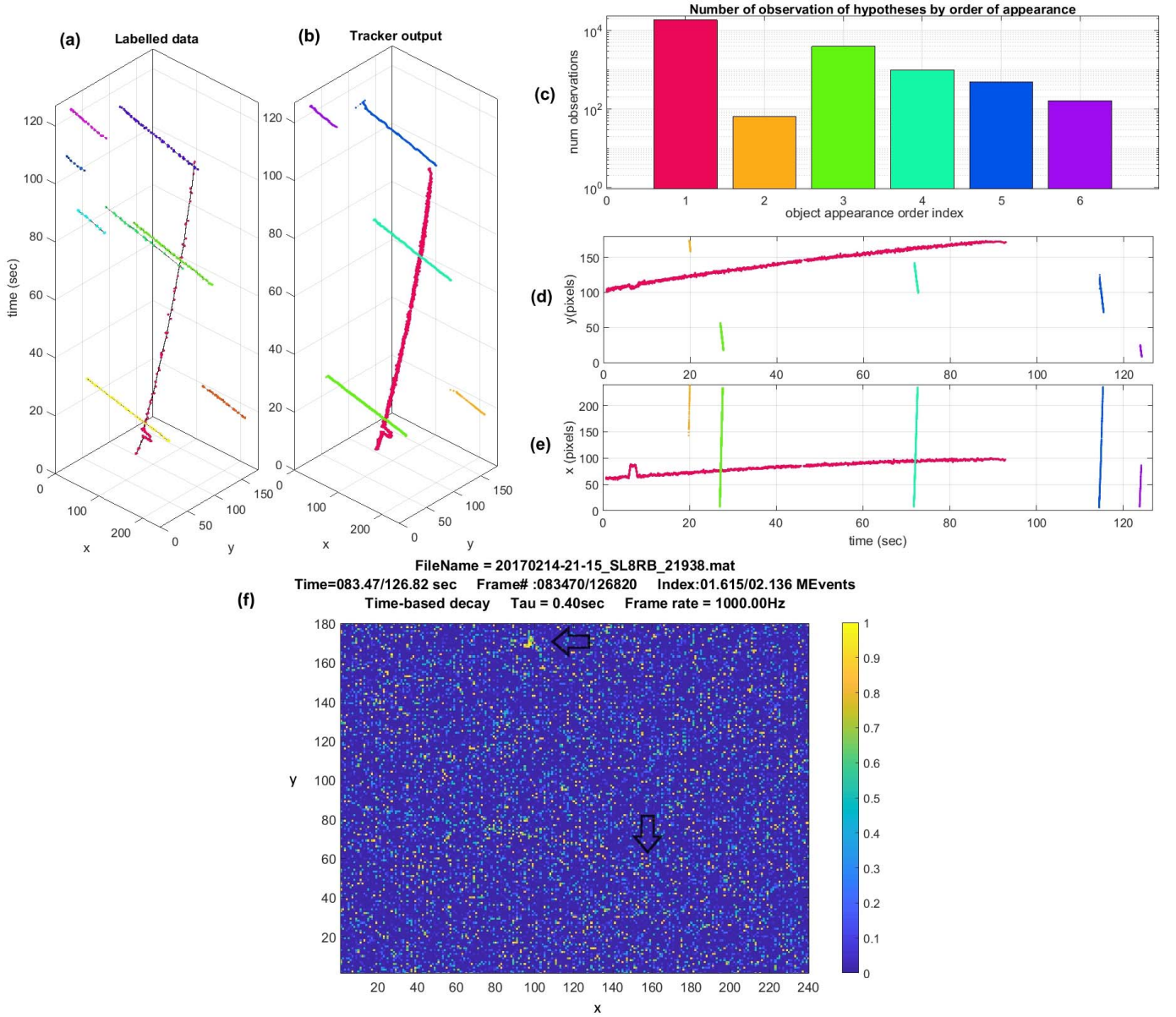
Fig. 6. **Output of tracking algorithm.** Panels (a) and (b) show the dimetric projection of the labeled data and the output of the event-based tracker respectively for the SL-8 R/B recording. (c) Shows the number of tracking events per object. Panels (d) and (e) show the tracker event position in x and y respectively over time. (f) Example of an expert labeled object not detected by the algorithm showing the difficulty level at which the algorithm fails. The missed object is at the tip of a streak that enters the field of view at (1,90) and continues to its location at (158, 56). SL-8 R/B is located at (97, 170). A rendered video of this recording is provided with the dataset.

allows the rapid calculation of the velocity of each object in $\{x, y, \theta\}$ space without the need to perform least-squares on previous observations. The event-based tracker update method is described using Algorithm 3.

As shown in Figure 6, the event-based line fitting tracker algorithm removes virtually all false detection events remaining after the feature detection stage while correctly clustering events from each object. The output events g_l from the tracker can be represented in the form of an event stream defined as:

$$g_l = [x_l, y_l, p_l, \theta_l, o_l, t_l]^T \qquad (20)$$

where o_l is the object index of the lth event generated by the tracker. Figure 6 compares the output event stream of the

tracking algorithm to the labeled data. Figure 6(f) shows an example of a labeled object missed by the end-to-end system. In the example SL-8R/B recording, three such faint high-speed objects are missed, demonstrating the superior performance of human experts over the proposed algorithm.

Table III details the statistics generated from the output of the event-based tracker demonstrating improved sensitivity, specificity and informedness with respect to the raw and detection event streams detailed in Tables I and II respectively.

G. Alternative Algorithms

To further evaluate and benchmark the performance of the feature detection algorithm, three additional high-speed event-based methods were implemented and tested on the
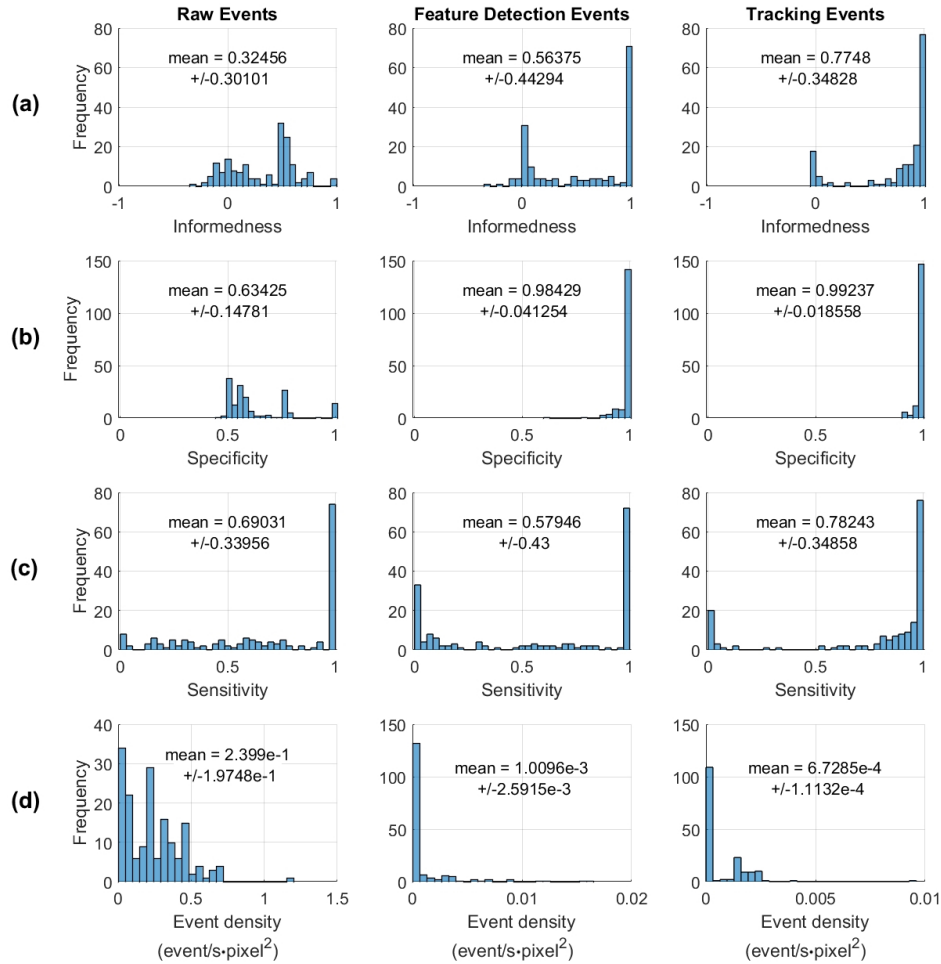
Fig. 7. **Per Recording Histogram Results on the Space Imaging Dataset.** From left to right, the panels show results from the raw events, the detection events and the tracking events. From top to bottom the panels show (a) informedness, (b) specificity, (c) sensitivity and (d) the event density of each of the event streams.

TABLE III
EVENT DENSITY ACTIVATED VOLUME STATISTICS FOR MEASURING THE PERFORMANCE OF THE TRACKING EVENT STREAM g_l **AGAINST LABELS.** THE STATISTICS CALCULATED GENERATED USING THE SL8R/B RECORDING WHOSE DATA STREAM IS ILLUSTRATED IN FIGURE 6

Polarity	Sensitivity	Specificity	Informedness	# Events (ke)
ON Events	0.90	0.99	0.89	20.21
OFF Events	0.87	0.97	0.84	15.13

space imaging dataset. The first method used is the event-based Global Maximum Detection (GMD) algorithm. The GMD sequentially finds the most prominent region of activity in the scene and assigns a detection to this point. This method can be viewed as a highly optimized event-based blob detector that is particularly suitable to the star tracking environment due to the sparsity in most recordings. The second, an event-based Hough transform algorithm and the third method, combined the best performance of the previous two methods via access to ground truth labeling. In the Supplementary Material Section, we describe these alternative methods in detail and discuss their performance in relation to the proposed algorithm. In all three approaches the events are first processed through the

same time surface generation and surface activation filter described in Section II-E. Following each of the alternative feature detection algorithms, the same tracking algorithm described here was used on the detection event stream providing an unbiased comparison between the methods.

III. RESULTS

A. Performance on Real World Space Imaging Dataset

The detailed results for all recordings in the dataset are summarized in Figure 7. The first three rows of results (a), (b) and (c) plot informedness, specificity and sensitivity respectively. The results demonstrate how each stage of processing shifts the distribution toward 1 resulting in a more informative event stream. The bottom row (d) shows how, at each stage of processing, the event density of the recordings is reduced into an ever more efficient representation of the data. Together these results demonstrate that over the wide range of heterogeneous input event streams, the proposed algorithm generates a sparse yet informative output event stream. (b) Shows the per recording specificity distribution is shifted from a mean of 0.63 for the raw events to 0.98 and 0.99 for the detection and tracking events with most results at 1. Similarly (c) shows the per recording sensitivity
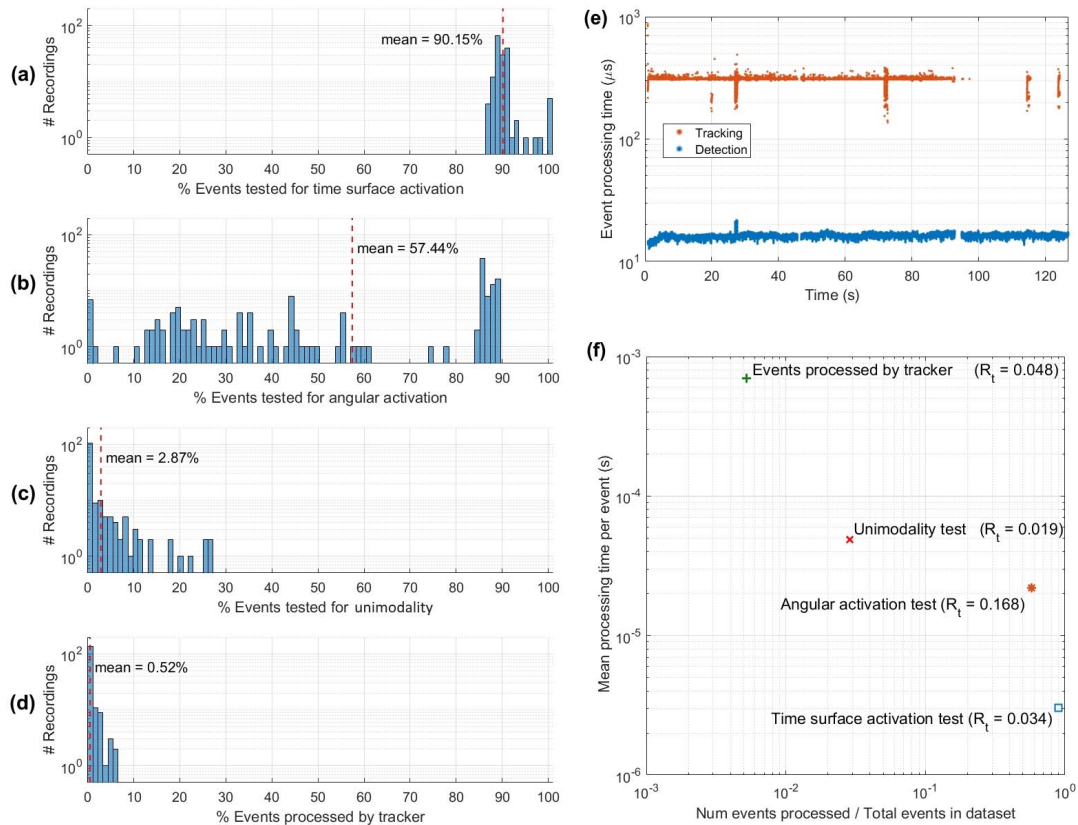
Fig. 8. Reduction in event numbers and associated processing time at each stage of the algorithm. Panel (a) shows the distribution of the number of events processed in each recording by the initial time surface activation test. (b) Shows the angular activation test, (c) the angular unimodality test and (d) the tracker. (e) Shows processing time per event of the detection and tracking algorithm for the SL-8 R/B recording. (f) The number of events at each stage of processing against the mean processing time per event at that stage. The processing time ratio R_t is the total processing time of each stage divided by the duration of the recording being processed.

distributions for the raw, detection and tracking event streams. Here, the sensitivity distribution is actually reduced in the detection stream in comparison to the raw events. This is primarily due to the relative sparseness of the detection stream. When the sparser detection event stream is interpolated via the tracker, the sensitivity rises above the raw events. Together the higher sensitivity and specificity result in a significantly higher informedness distribution as shown in (a). These results demonstrate the effectiveness of the end-to-end system in transforming noisy raw input events from space imaging data, into sparse highly informative noise-free event streams using a series of simple hardware implementable filters.

B. Processing Time Results

In this section the processing time and filtering operation of the algorithm is detailed. The processing time was evaluated in the MATLAB 2017a environment on a laptop with a 64-bit 4.00 GHz i7-6700 CPU processor and 64GB of RAM. Figure 8 shows the cascaded event filter design of the proposed system, where at each of the increasingly refined processing stages, the increased computation time is accompanied by a corresponding reduction in events.

As the distributions shown in panels (a) to (d) of Figure 8 demonstrate, the event rates at each stage of processing of the space imaging dataset becomes reduced requiring an ever-smaller number of events to be processed by the subsequent stage. Furthermore, as panel (e) shows for the example SL-8 R/B recording, due to the sparseness of activation in space imaging event streams, the processing speed of the algorithms is remarkably stable over time within a recording. In other words, given the small size of the area occupied by space objects relative to the entire field of view, the presence or absence of even bright target objects in the field of view makes little difference in the global event rate of the raw events. This is in contrast to terrestrial applications where, due to the complexity and the relative size of the objects in the visual environment, the event rate can vary by many orders of magnitude depending on the relative velocity of the visual scene. The relative stability of event rates within EBSI recordings can be exploited at every stage of processing. This property of the data provides yet another important distinction between EBSI processing algorithms and more general event-based systems. Panel (f) shows the timing response of the entire system for each processing stage. Here we can observe that as envisioned, at each stage of processing, the increase in complexity of the following stage is accompanied by an approximately commensurate reduction in input event rate such that the entire end-to-end system can process all events at slightly faster than an eighth of the speed of the first simple surface activation test. This is despite the fact that the last processing stage, the tracker, processes events at a rate that is more than 230 times slower than the first stage.

TABLE IV
SUMMARY OF RESULTS OF TESTED ALGORITHMS ON THE SPACE IMAGING DATASET

Algorithm	Informedness mean	std	Sensitivity mean	std	Specificity mean	std	Processing time ratio
Raw Events	0.324	0.301	0.690	0.340	0.634	0.148	1
Hough Detector	0.244	0.343	0.552	0.408	0.692	0.224	0.632
Hough Detector + Tracker	0.417	0.478	0.442	0.488	0.975	0.073	0.781
GMD Detector	0.609	0.323	0.756	0.284	0.853	0.117	**0.113**
GMD Detector + Tracker	0.664	0.374	**0.813**	0.314	0.851	0.223	0.671
max(GMD,Hough) Detector	0.617	0.309	0.754	0.286	0.863	0.103	0.755
max(GMD,Hough)Detector+Tracker	0.753	0.344	0.804	0.331	0.950	0.096	1.174
Feature Detector	0.564	0.443	0.580	0.430	0.984	0.041	0.222
Feature Detector + Tracker	**0.775**	0.348	0.782	0.349	**0.992**	0.019	0.270

Finally, note the position of the angular activation test above and to the right of the diagonal formed by the other tests. This position identifies this stage as the bottleneck in the system as discussed in Section II-E.

C. Comparison of Proposed Method With Alternative Algorithms

To provide a benchmark for comparison Table IV details the results of the feature-based detection and tracking algorithm against alternative event-based high-speed algorithms that could be used on the space imaging dataset against expert labeling. For a more detailed discussion of the behavior of the alternative algorithms see the Supplementary Material Section. All algorithms detailed in the table operated on the same event stream which was pre-processed with the same initial local surface activation filter and were paired with an identical event-based tracker for the tracking results. The first row in the table sets the raw events as a baseline showing low informedness primarily due to the low mean specificity of the event streams. Given the high noise rate, the Hough line detection algorithm is the worst performing algorithm in this context with informedness lower than the raw events. This however is primarily due to the poor sensitivity of the Hough detector to a great number of the observed objects in the dataset that are extremely slow-moving. These slow objects generate faint point source-like activation patterns which the Hough detector can not detect. When augmented with the event-based tracker, the sensitivity of the system is slightly reduced but specificity rises to close to 1 resulting in a near doubling of the informedness. In contrast to the Hough detector, the GMD detector performs best on the more common slower-moving targets thus resulting in significantly higher sensitivity and thus informedness. The GMD detector however performs poorly in noise filtering. This is especially the case for neighboring clusters of noise events from overactive 'hot pixels' on the sensor which are a challenging feature of the dataset and which the GMD fails to remove. Furthermore, these localized stationary clusters of noise activation are also difficult for the tracker to remove. For this reason, the specificity of the GMD system is about the same with or without the tracker. However, the tracker does slightly improve sensitivity mainly through interpolating between periods of higher activity of slow-moving objects. Next, when the performance of the GMD and the Hough detector are combined in a post hoc manner the highest informedness is achieved. When the output of this detection system is processed by the tracker, a result of 0.804 sensitivity, 0.95 specificity and

0.753 informedness is achieved. The performance of this combined system with access to the ground truth labeling serves as a benchmark for comparison to the feature-based detection algorithm. When the feature-based detection event stream is evaluated alone we observe a low sensitivity value of 0.58 but the highest specificity so far at 0.984. However, when combined with the tracker the sensitivity jumps to 0.782, the specificity to 0.992 and the informedness to 0.775 with the latter two being the highest achieved measures on the dataset, exceeding even the combined GMD-Hough system. Together these results show that after the tracking stage is completed, the proposed feature-based detection approach outperforms all other methods tested including the post hoc combined Hough-GMD detector with unrealistic access to ground truth demonstrating the performance of the proposed approach on this challenging space imaging dataset. Finally, as detailed in the last column of Table IV, the processing time of the feature-based detector at 0.222 real-time duration, is approximately double the much simpler and lower performing GMD detector. When augmented with the tracker the feature-based detection and tracking system process events faster than all other approaches at only 0.27 times real-time duration. This is less than half the processing time of the GMD detection and tracking system which passes through a significant number of noise detection events to the more computationally expensive tracker as is evidenced by lower specificity of the GMD detector relative to the feature-based detector. This best-of-both-worlds performance, of high processing speed and high algorithm complexity resulting in high accuracy, is only possible due to the highly optimized cascaded event-based filtering design described.

IV. DISCUSSION

In terrestrial event-based recording conditions a typically complex, feature-rich scenery is observed at a relatively high SNR, generating event streams with high variance in event rates. In contrast, event-based space imaging typically contains sparse simple featured scenes with low SNR and very stable event rates. In this context, the primary challenge is not the processing of a complex environment, but the extraction of simple faint detections from a highly noisy random event stream. Here, even the most simple event-based algorithms such as hot pixel filters can become problematic given the similarity of noisy pixels to the stationary point sources targeted in EBSSA. Thus EBSSA is to a significant degree an exercise in SNR enhancement. However, it should be noted that it is always desirable to operate the system as close to the

noise limit as possible. If the SNR is improved for a specific sensitivity, then it would potentially be desirable to make the sensor more sensitive and detect fainter objects at the highest tolerable noise level.

Two entirely independent solutions to this problem of low SNR are of course the design of specialized event-based space imaging sensors and more immediately online automated optimization of current event-based sensor biases to recording conditions. The development of more sensitive dedicated EBSI sensors is the subject of our future work. Among the recordings in the dataset are instances where due to the incidental alignment of sensor biases to the recording conditions extremely faint low earth orbits objects exhibiting random trajectories are observed. In theory, such LEO observations should populate all recordings in the dataset, yet they are present in only a few. On the other hand, regardless of future improvements in sensor technologies, improved observing conditions and future implementations of online sensor bias optimization systems, there will always remain fainter space objects to be observed and extracted from the event stream. This perpetual requirement for higher sensitivity will continue to motivate the configuration of sensor biases for higher sensitivity (and higher noise) in the space imaging context. This ensures that such event-based datasets will continue to be noisy and in need of robust detection and tracking algorithms like those described in this work.

One important hyperparameter in the algorithms presented and in all low SNR event-based applications is the size of the ROI patch used. While small ROIs with faster decaying memory suffice in high SNR contexts, in low SNR applications such as EBSI, larger-sized ROIs with slower decaying memory collect more information from a larger spatio-temporal volume which typically results in better performance. On the other hand, increasing a system's ROI size reduces its speed. Through heuristic testing of the space imaging dataset and the algorithms presented in this work, an ROI of fifteen by fifteen pixels was found to provide a reasonable trade-off between performance and speed.

Another important hyperparameter that was investigated in detail was the shape and weights of the LUT templates used to generate the angular activation vector Λ. Initially, it was assumed the precise image used for the template and its fidelity to observed space object shapes would significantly impact the accuracy of the overall system and be highly specific for each particular class and size of the objects observed. In practice, it was found through experimentation with a range of different bar shapes, lengths, widths and template values, that as long as the template was strongly uni-directional, the precise shape of the template did not significantly impact performance.

In this work the proxy signal ζ estimating unimodality of the angular activation Λ was used for scale, speed and rotation invariant detection of point sources. In typical terrestrial event-based contexts with their higher SNRs and more complex features a more local plane fitting optical flow algorithm is used as the first step in detecting events on moving edges [33]. These events are then augmented with orientation information that is analogous to θ_i in this work. In future work, we will apply the optimized hardware-implemented feature-based detection algorithm presented here to extremely low SNR terrestrial contexts where the larger ROI are likely to provide improved performance over more localized optical flow detection algorithms.

By performing all stages of the detection and tracking operation in the event-based domain, the presented system maintains the high temporal precision of the sensor allowing it to process high-speed targets even at close range extending the potential application of the system to a wider range of challenging high-speed sensing environments.

V. CONCLUSION

In this work, the first event-based space imaging dataset was presented. The labeled dataset, augmented with a larger unlabeled dataset, provides a test bench for investigation of event-based algorithms for the unique and challenging space imaging environment. Statistical measures were introduced where event density activated spatio-temporal volume slices can be used to compare the sensitivity, specificity and informedness of extremely heterogeneous event streams. In this way, the output of the proposed detection and tracking systems can be directly compared to the raw input events quantifying improvements at each stage and providing insights into properties of the dataset as well as the operation of the algorithm. The expert labeling procedure used was validated using an artificial dataset with analytically defined ground truth. The expert labeling procedure was shown to provide a highly accurate label set across a wide range of SNR environments. Several high-speed event-based algorithms with different levels of complexity were tested on the dataset with the feature-based detection and tracking method outperforming the other methods combined, both in terms of accuracy as well as in speed of operation. By measuring an optimized proxy measure for the unimodality of angular activation over a fairly large, slow decaying local time surface region, the feature-based method was shown to provide a scale, rotation and speed invariant target detection capability that is ideal for the event-based space imaging context. In terms of speed of operation, the cascaded event-filter design of the detection and tracking system provides a high-speed event processor.

REFERENCES

[1] F. Schiemenz, J. Utzmann, and H. Kayal, "Survey of the operational state of the art in conjunction analysis," *CEAS Space J.*, vol. 11, no. 3, pp. 255–268, Sep. 2019.

[2] D. J. Kessler and B. G. Cour-Palais, "Collision frequency of artificial satellites: The creation of a debris belt," *J. Geophys. Res.*, vol. 83, no. A6, p. 2637, 1978, doi: 10.1029/ja083ia06p02637.

[3] N. Bobrinsky and L. Del Monte, "The space situational awareness program of the European Space Agency," *Cosmic Res.*, vol. 48, no. 5, p. 392, Oct. 2010. [Online]. Available: http://link.springer.com/10.1134/S0010952510050035

[4] B. Lal, E. J. Sylak-Glassman, M. Mineiro, N. Gupta, L. M. Pratt, and A. Azari, "Global trends in space volume 1: Background and overall findings," IDA Sci. Technol. Policy Inst., Washington DC, USA, Tech. Rep. P-5242, 2015.

[5] *Continuing Kepler's Quest*, National Academies Press, Washington, DC, USA, Sep. 2012.

[6] G. Cohen *et al.*, "Event-based sensing for space situational awareness," *The J. Astron. Sci.*, vol. 66, pp. 125–141, Jun. 2019.

[7] C. Posch, T. Serrano-Gotarredona, B. Linares-Barranco, and T. Delbruck, "Retinomorphic event-based vision sensors: Bioinspired cameras with spiking output," *Proc. IEEE*, vol. 102, no. 10, p. 1470, Oct. 2014.

[8] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, p. 259, Jan. 2011.

[9] B. Cheung, M. Rutten, S. Davey, and G. Cohen, "Probabilistic multi hypothesis tracker for an event based sensor," in *Proc. 21st Int. Conf. Inf. Fusion*, Jul. 2018, p. 933.

[10] A. Lambert and G. Cohen, "Study of a spatio-temporal sensor for turbulence characterisation and wavefront sensing (conference presentation)," *Proc. SPIE Environ. Effects Light Propag. Adapt. Syst.*, vol. 10787, Art. no. 1078702, Oct. 2018.

[11] D. M. Palmer and R. M. Holmes, "Extremely low resource optical identifier: a license plate for your satellite," *J. Spacecraft Rockets*, vol. 55, no. 4, p. 1014, 2018.

[12] G. Cohen, S. Afshar, and A. van Schaik, "Approaches for astrometry using event-based sensors," in *Proc. Adv. Maui Opt. Space Surveill. Technol. Conf.*, 2018, p. 1.

[13] P. Burt *et al.*, "Object tracking with a moving event camera," *Proc. Workshop Vis. Motion*, 2016, p. 2.

[14] D. Drazen, P. Lichtsteiner, P. Häfliger, T. Delbrück, and A. Jensen, "Toward real-time particle tracking using an event-based dynamic vision sensor," *Experim. Fluids*, vol. 51, no. 5, p. 1465, 2011.

[15] Z. Ni, S. Ieng, C. Posch, S. Régnier, and R. Benosman, "Visual tracking using neuromorphic asynchronous event-based cameras," *Neural Comput.*, vol. 27, no. 4, p. 925, 2015.

[16] D. Tedaldi, G. Gallego, E. Mueggler, and D. Scaramuzza, "Feature detection and tracking with the dynamic and active-pixel vision sensor (DAVIS)," in *Proc. 2nd Int. Conf. Event-Based Control, Commun., Signal Process. (EBCCSP)*, 2016, p. 1. [Online]. Available: http://ieeexplore.ieee.org/document/7605086/

[17] X. Lagorce, C. Meyer, S.-H. Ieng, D. Filliat, and R. Benosman, "Asynchronous event-based multikernel algorithm for high-speed visual features tracking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1710–1720, Sep. 2014.

[18] M. Z. O. L. Nowski, R. Reszelewski, D. P. Moeys, T. Delbruck, and K. Kamiński, "Observational evaluation of event cameras performance in optical space surveillance," in *Proc. 1st ESA NEO Debris Detection Conf.*, 2019, pp. 1–7.

[19] T.-J. Chin, S. Bagchi, A. Eriksson, and A. van Schaik, "Star tracking using an event camera," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 1–10.

[20] N. Estell, D. Ma, and P. Seitzer, "Daylight imaging of leo satellites using cots hardware," in *Proc. AMOS*, 2019. [Online]. Available: https://scholar.google.com.au/scholar?cluster=7475790139432626288&hl=en&as_sdt=0,5

[21] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[22] A. Aggarwal. (2019). *Sl-8 R/B Satellite Details 1968-040b Norad 3230*. [Online]. Available: https://www.n2yo.com/satellite/?s=3230

[23] S. Afshar, A. van Schaik, and G. Cohen. (2019). *Space Imaging Dataset*. https://www.westernsydney.edu.au/icns/reproducible_research/publication%_support_materials/space_imaging

[24] G. Taverni *et al.*, "Front and back illuminated dynamic and active pixel vision sensors comparison," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 5, p. 677, Oct. 2018.

[25] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, p. 32, 1950.

[26] D. M. W. Powers, "Evaluation: From precision, recall and F-factor to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, pp. 1–24, 2007. [Online]. Available: https://scholar.google.com.au/scholar?cluster=2000221564118550637&hl=en&as_sdt=0,5

[27] S. Afshar, T. J. Hamilton, J. Tapson, A. van Schaik, and G. Cohen, "Investigation of event-based surfaces for high-speed detection, unsupervised feature extraction, and object recognition," *Frontiers Neurosci.*, vol. 12, p. 1047, Jan. 2019.

[28] D. Czech and G. Orchard, "Evaluating noise filtering for event-based asynchronous change detection image sensors," in *Proc. 6th IEEE Int. Conf. Biomed. Robot. Mechatronics (BioRob)*, Jun. 2016, p. 19.

[29] D. Cox, "Notes on the analysis of mixed frequency 1 distributions," *Brit. J. Math. Stat. Psychol.*, vol. 19, no. 1, p. 39, 1966.

[30] J. A. Hartigan *et al.*, "The dip test of unimodality," *Ann. Statist.*, vol. 13, no. 1, p. 70, 1985.

[31] B. W. Silverman, "Using kernel density estimates to investigate multimodality," *J. Roy. Stat. Soc., Ser. B Methodol.*, vol. 43, no. 1, p. 97, 1981.

[32] M.-Y. Cheng, T. Gasser, and P. Hall, "Nonparametric density estimation under unimodality and monotonicity constraints," *J. Comput. Graph. Statist.*, vol. 8, no. 1, p. 1, Mar. 1999.

[33] R. Benosman, S.-H. Ieng, C. Clercq, C. Bartolozzi, and M. Srinivasan, "Asynchronous frameless event-based optical flow," *Neural Netw.*, vol. 27, p. 32, Oct. 2012.

Saeed Afshar (Member, IEEE) received the B.Eng. degree in electrical engineering from the University of New South Wales, Sydney, Australia, the M.Eng. degree in electrical engineering from the University of Western Sydney, Sydney, Australia. He is a Postdoctoral Research Fellow with the International Centre for Neuromorphic Systems, Western Sydney University. His research interests include event-based vision and audio processing for neuromorphic hardware.

Andrew Peter Nicholson received the B.Eng. (Hons.) degree in computer engineering from the University of Newcastle, Callaghan, NSW, Australia, the M.Eng. degree in electrical engineering from the University of Sydney, specializing in wireless, and the Ph.D. degree from the School of Electrical Engineering and Telecommunications, University of New South Wales, Australia. He is a Postdoctoral Research Fellow with the International Centre for Neuromorphic Systems.

Andre van Schaik (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the University of Twente, Enschede, The Netherlands, in 1990, and the Ph.D. degree in neuromorphic engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1998. He is a pioneer of Neuromorphic Engineering and the Director of the International Centre for Neuromorphic Systems, Western Sydney University. From 1991 to 1994, he was a Researcher with the Swiss Centre for Electronics and Microtechnology (CSEM), where he developed the first commercial neuromorphic chip–the optical motion detector used in Logitech trackballs since 1994. In 1998, he was a Postdoctoral Research Fellow with the Department of Physiology, University of Sydney. In1999, he became a Senior Lecturer with the School of Electrical and Information Engineering and a Reader in 2004. In 2011, he became a Full Professor with Western Sydney University. His research interests include neuromorphic engineering, encompassing neurophysiology, computational neuroscience, software and algorithm development, and electronic hardware design. He has authored more than 200 articles and is an inventor of more than 35 patents. He has founded three technology start-ups. He is a Fellow of the IEEE for contributions to Neuromorphic Circuits and Systems.

Gregory Cohen (Member, IEEE) is an Associate Professor of Neuromorphic Systems with the International Centre for Neuromorphic Systems (ICNS), Western Sydney University, and a program lead for neuromorphic algorithms and space applications. Prior to returning to research from industry, he worked in several start-ups and established engineering and consulting firms including working a consulting engineer in the field of large-scale HVAC from 2007 to 2012. He is a pioneer of event-based and neuromorphic sensing for space imaging applications, unsupervised feature extraction, bio-inspired machine learning, and neuromorphic computation systems.