Event-Based State Estimation of Planetary Robots using Propeller Sensing

Ravi Kumar Thakur*, Luis Granados Segura, Jan Klivan, Radim Špetlík
Tobiáš Vinklárek, Matouš Vrba, Martin Saska

***Abstract*—Robotic missions for planetary exploration are challenging due to difficult or inaccessible terrain and a limited capability to explore scientific interests beyond the landing zone. The use of Unmanned Aerial Vehicles (UAVs) for such cases offers advantages in terms of high-resolution spatial coverage, support science missions, and terrain mapping to enable planning for ground rovers. This potential can be augmented with the use of a swarm of aerial robots. Such multi-robot system could enable transport of scientific payloads on the planetary surface or mapping of areas inaccessible by rovers and relay orbiters. We propose a framework for relative state estimation for such a system using event-based propeller sensing. The propellers in the event stream are detected, tracked, and sampled asynchronously to extract their rotational frequencies per-propeller. These frequency measurements drive a kinematic state estimation module as a thrust input, while camera-derived position measurements provide the update step. Additionally, we use geometric primitives derived from these event streams to estimate the orientation of the quadrotor using perspective-1-ellipsoid correspondence. We evaluate our approach on a set of real-world outdoor flight sequences obtained from quadrotors in a leader-follower formation, providing a method for decentralized relative localization for multi-robot systems using event cameras.**

***Index Terms*—Event Camera, Propeller Sensing, Frequency Estimation, Relative Localization**

I. INTRODUCTION

The exploration of celestial bodies using robots introduces unique challenges that have not been faced on Earth. The lack of accessibility and communication delays forces the robotic mission to have higher safety standards with harsher constraints on size, weight, and power. Over the years, such robotic systems, including space probes and reconnaissance spacecraft, have played a crucial role in furthering our understanding of planetary environments. Recent advances in the capabilities of robotic landers and rovers allowed us to perform in-situ measurements of planetary surfaces and atmosphere. However, challenges remain when it comes to scaling the exploration mission or building infrastructure for future manned flights. These challenges can be addressed using the autonomous formation of multiple aerial robots.

To enable such a multi-robot aerial swarms requires that robots estimate the states of others accurately. In an aerial swarm, the robots can interact with each other to pursue a common goal. Accurate estimation of linear and rotational kinematic states is necessary for multi-robot coordination by performing collision avoidance, infer intent, predicting trajectories and maintaining relative poses with respect to each other. This is commonly achieved by explicit communication between them [1] or by using relative localization through various sensing modalities [2]–[4]. Most formation flights involving quadrotors rely on robust communication systems to exchange state information. However, performance can be degraded in the presence of communication latency or packet loss, and in some cases, the stability of formation may be compromised. To solve this problem, we propose event-based propeller sensing to estimate the relative states of quadrotors, thus enabling decentralized swarm flight in visually challenging conditions

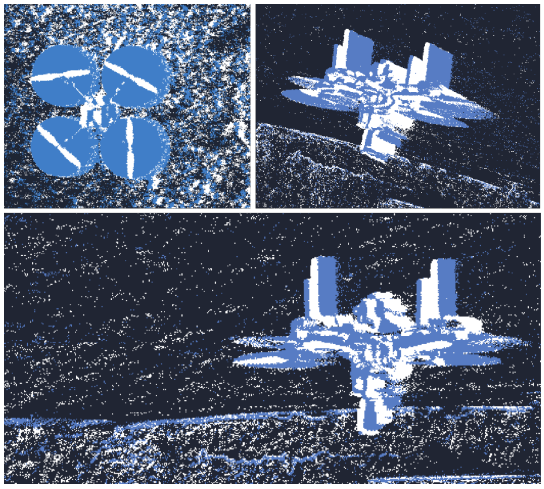


Fig. 1. Visuals of a quadrotor in-flight as seen in pseudo-frame representation of streams accumulated from the event camera.

A resilient autonomous swarm flight involves multiple quadrotors performing tasks with limited communication, typically based on vision-based perception. However, most vision-based methods rely on a frame-based RGB camera which records the light intensity per pixel for every frame. The frame-based camera has low temporal resolution, high latency, and low dynamic range in comparison to event cameras. These limitations make it a less suitable choice where the environment introduces motion blur, extreme illumination changes, and background noise due to clutter. Thus, frame-based cameras introduce limitations for multi-robot formation flight especially if the quadrotors are flying in a perceptually

Authors are with the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Karlovo Namesti 13, 121 35 Prague 2, Czechia. *Corresponding author: ravi.thakur@cvut.cz

This work was funded by the Czech Science Foundation (GAČR) under research project no. 26-22419S

challenging environment.

To address the limitations of frame-based cameras in multi-robot formation flight, our work relies on event cameras. These sensors are low-powered, has low latency and high temporal resolution [5]. The sensor area is activated only when there is a change in the visual appearance of the scene. The event streams are generated by motion in the sensor area. Due to these advantages, event cameras are considered for autonomous navigation tasks in which there are dynamic obstacles and the quadrotors are moving at high speed [6], [7].

Such an autonomous swarm of quadrotors can enable future planetary exploration in multiple ways. The distributed network of quadrotors can act as a synthetic aperture or form a sensor network [8]. This will allow to form a configurable sensor array depending upon the changes in mission requirements. The swarm can also enable high resolution spatial coverage of difficult terrains inaccessible by rovers beyond the safer landing zones. The formation of quadrotors can be part of heterogeneous robot teams for the exploration of canyon-like structures where communication is challenging [9], [10], thus reducing dependency on relay orbiters.

In this paper, we propose the use of a neuromorphic camera as the primary sensor to perform relative state estimation for swarm of aerial robots. Our approach utilizes the event streams from the camera to detect the quadrotor, localize it, and estimate the frequency of its propellers. The frequency of the propeller is used by the state estimation pipeline to get first-order kinematics of the quadrotor. We make use of the shape of the propeller disk detected in the event streams to estimate the orientation of the quadrotor using Perspective-1-Ellipse (P1E) [11]. Our proposed framework for state estimation improves the aerial mobility system, enabling future scientific missions for planetary exploration.

II. RELATED WORK

A. Aerial Robots for Planetary Exploration

The use of aerial robots has attracted interest for planetary exploration due to its potential for high-resolution spatial coverage and ability to access difficult terrains. These include conceptual missions for Mars and Venus using platforms like fixed-wing UAVs, rotorcraft and balloons [12]. A formation of explorer robots consisting of autonomous balloons was proposed to carry out studies at atmospheric altitudes [13]. These balloons were conceptualized to have capabilities for trajectory control and deployable miniature probes. The Ingenuity helicopter was sent as a secondary payload along with the Perseverance rover [14]. It enabled high-resolution surface imaging and supported trajectory planning for the rover. A Mars Science Helicopter mission is being considered to follow with a larger payload capacity and capabilities to study the atmosphere at higher altitudes [15]. The Dragonfly mission is planned to carry out an astrobiological study of Titan using a rotorcraft platform [16]. The Skyfall mission is planned to explore Mars using a formation of three Ingenuity helicopters [17]. It will test the deployment of helicopters without a conventional landing system. The drones will operate autonomously to capture high resolution images, map terrain using ground-penetrating radar, and scout for future landing sites.

B. State Estimation using Event Camera

The estimation of the state of UAV and its relative pose with respect to other robots using an RGB camera is a common approach. However, those approaches cannot be directly applied to the event camera. The potential use of propeller sensing to estimate rotational attitude parameters was proposed in [18]. In this approach, the event streams are treated as the Poisson process to distinguish events generated by a propeller and the background. The pose is estimated by fitting an ellipse over the aggregated propeller events. HelixTrack extends this work by estimating the pose of the rotrating propeller using a batched Gauss-Newton optimization. They test their algorithm on propeller dataset with different levels of ego-motion [19]. Eventpro [20] proposes the estimation of the propeller frequency using hierarchical event processing and known propeller geometry. The flight commands are then predicted using a trained Support Vector Machine (SVM) classifier. In [21], an estimation of ego-motion of kinematic states based on monocular event vision is proposed, relying on event-based optical flow without propeller detection. Unlike these methods, our approach directly links quadrotors propeller frequencies to a physical dynamics model for state estimation.

III. STATE ESTIMATION USING PROPELLER SENSING

A. Event Camera Data Representation

An event camera responds to any change in brightness in the sensor region. They are low-powered device with high dynamic range and microsecond resolution. In contrast to frame-based cameras, they output asynchronous event streams where each event is represented as a tuple

$$e_i = (x_i, y_i, t_i, p_i), \tag{1}$$

where (x_i, y_i) denote the pixel coordinates, t_i is the microsecond-precision timestamp, and $p_i \in \{-1, +1\}$ indicates the polarity (brightness increase or decrease). An event is triggered when the logarithmic intensity change at a pixel exceeds a threshold C:

$$|\log I(x, y, t_i) - \log I(x, y, t_{i-1})| > C. \tag{2}$$

This enables the event cameras to respond to fast changes. They naturally respond to features like edges and corners by design. However, because of their inherent sensitivity to motion, they capture background noise as well. The event streams can be visualized using intermediate pseudo-frame representations like time-surface, event histogram and event frame etc. A comprehensive survey of event-based vision, the type of data representation, and their processing is provided in [5] .

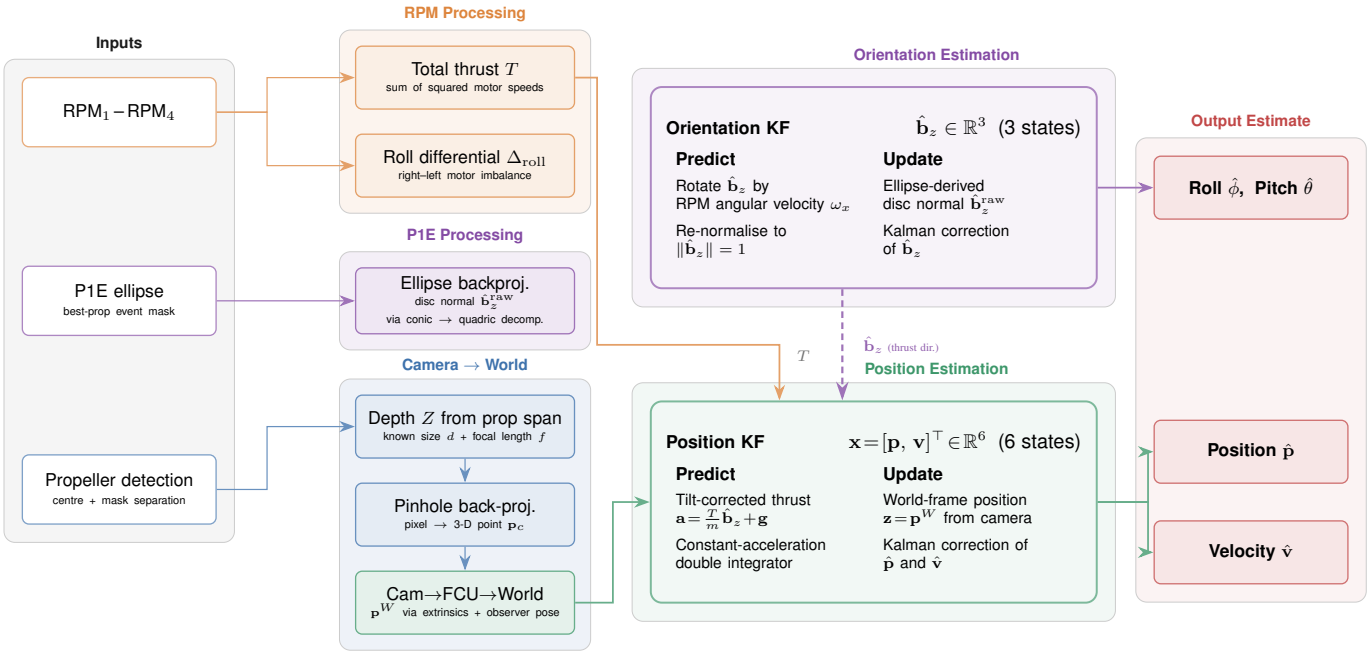


Fig. 2. State estimation pipeline. Two coupled KFs operate on complementary measurements. Orientation KF tracks the body z-axis $\hat{\mathbf{b}}_z$, predicted by RPM-derived angular velocity and corrected by P1E ellipse backprojection. Position KF tracks position and velocity, predicted by tilt-corrected thrust and corrected by camera-derived world-frame position. The orientation estimate provides the thrust direction for the position prediction step. Roll and pitch are extracted from $\hat{\mathbf{b}}_z$ via a yaw-free decomposition since yaw is unobservable from the propeller ellipse.

B. Propeller Detection Pipeline

The proposed approach processes event streams in temporal chunks of $\Delta t = 10\,\text{ms}$. The outdoor data set contains noise generated by the background due to events generated by camera movement. These events are filtered out using spatio-temporal contrast filtering. We implemented two approaches to detect the propeller region. The first uses Connected Components (CC) labeling on the event frame and the other uses Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) clustering [22] on the event streams. The goal of both detection approaches is to identify the regions of interest for estimating frequency. Detected regions are tracked across frames and assigned stable tracking IDs based on spatial position.

C. Tracking of the quadrotor

The detected propeller regions were tracked across temporal chunks using Norfair [23], an open-source multi-object tracking framework. This tracker maintains track continuity by associating detections between frames using Euclidean distance in position space. These detections can come from either CC labeling or density based clustering. Each detection is represented by its centroid (c_x, c_y), which serves as both the tracked point and the re-identification embedding for similarity matching. The four centroids can be used to derive the center of the quadrotor and track its position across time. A general survey of event-based quadrotor detection methods is given in [24]. In this work, the objective of propeller detection and tracking is also to validate the accuracy of frequency estimation by matching the tracked propeller ID with the ground truth.

The tracker employs a distance threshold and maintains tracks for a few frames during temporary occlusions or detection failures. When a propeller is briefly missed (e.g., when detection fails due to CC labeling merging the blobs or occlusion of the propellers), the tracker predicts its position based on motion history, preventing track fragmentation.

D. Propeller Frequency Estimation

The frequency is computed by observing the event counts in the region of interest defined by a sub-quadrant in the bounding box around detected propeller. In these regions, we temporally bin the events and create a sliding window consisting of time bins sufficient to capture multiple blade passages, resulting in an event-count signal. The Fast Fourier Transform (FFT) is applied to the binned event count signal to detect the dominant frequency peak corresponding to the blade passage rate. A KF smooths the raw estimates. The filter is initialized from the first FFT measurement to avoid convergence delays. The filter provides temporal consistency by rejecting outliers, smoothing out measurement noise.

E. Relative State Estimation

The kinematics and dynamics model of the quadrotor is formulated as Eq. (3)–Eq. (6). The equations describing thrust T is generated along the body z-axis proportional to the square of the motor angular velocity Ω_i. The rotation matrix corresponding to the quaternion $\mathbf{q}$ is denoted by $\mathbf{R}(\mathbf{q})$. It converts vectors in the body-fixed frame of the quadrotor to the

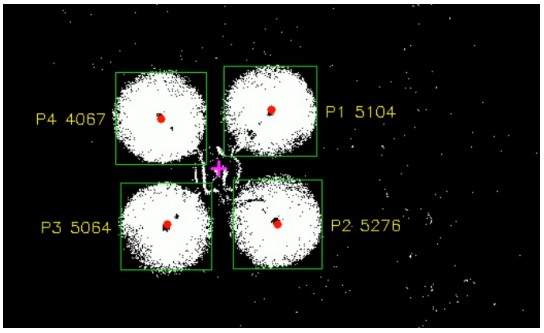


Fig. 3. The event frame showing instantaneous RPM and center of quadrotor. The propellers are numbered 1–4 from top-right.

world frame. The total thrust of the target quadrotor is obtained as $T = c_f \sum_{i=1}^{4} \Omega_i^2$, with Ω_i ($\mathrm{rad\,s^{-1}}$) denoting the angular velocity of motor i obtained from the event-camera propeller frequency estimates and c_f denoting thrust coefficient of the motor. The position and velocity are denoted by $\mathbf{p}$ and $\mathbf{v}$ respectively. A detailed modeling of dynamics and control of a quadrotor system is described in [25].

$$\dot{\mathbf{p}} = \mathbf{v}, \tag{3}$$

$$\dot{\mathbf{v}} = \mathbf{R}(\mathbf{q}) \begin{bmatrix} 0 \\ 0 \\ T/m \end{bmatrix} + \mathbf{g}, \tag{4}$$

$$\dot{\mathbf{q}} = \tfrac{1}{2}\, \mathbf{q} \otimes [0,\, \boldsymbol{\omega}], \tag{5}$$

$$\dot{\boldsymbol{\omega}} = -\beta\, \boldsymbol{\omega}, \tag{6}$$

the relative state estimation pipeline has position and orientation estimation modules as shown in Fig. 2. The position estimation module has a vector $\mathbf{x} = [\mathbf{p}^\top,\ \mathbf{v}^\top]^\top \in \mathbb{R}^6$. The horizontal components do not receive thrust input and are propagated by constant-velocity extrapolation, making the contribution of propeller frequency directly observable in vertical tracking performance. The state is tracked with a standard linear KF.

The orientation is estimated by the tracking z-axis $\hat{\mathbf{b}}_z = [0, 0, 1]^\top \in \mathbb{R}^3$ representing thrust direction in the world frame. The propeller frequency estimates are fused with camera-derived quadrotor position measurements in the world in this filter. The filters are coupled such that orientation estimate provides the thrust direction used by the position estimation module. The total thrust as a function of the motor constant and the propeller frequencies gives the world frame input acceleration. We describe both the modules below.

1) Position Estimation: The world-frame acceleration is computed using $\hat{\mathbf{b}}_z$ from the orientation filter. This provides a transformation to obtain the acceleration in the world frame as

$$\mathbf{a} = \frac{T}{m}\, \hat{\mathbf{b}}_z + \mathbf{g}, \tag{7}$$

where m is the mass of the quadrotor and gravity is denoted by $\mathbf{g} = [0, 0, -9.81]^\top\, \mathrm{m\,s^{-2}}$. When the orientation estimate is not yet available, the filter defaults to the hover assumption $\hat{\mathbf{b}}_z = [0, 0, 1]^\top$, so that thrust contributes only to the vertical component. The state transition matrix $\mathbf{F}$ and input $\mathbf{u}$ gives the prediction model as (9) and (10)

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & \mathbf{I}\Delta t \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \frac{1}{2}\mathbf{a}\Delta t^2 \\ \mathbf{a}\Delta t \end{bmatrix}, \tag{8}$$

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}\hat{\mathbf{x}}_{k|k} + \mathbf{u}, \tag{9}$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}\mathbf{P}_{k|k}\mathbf{F}^\top + \mathbf{Q}\Delta t, \tag{10}$$

where $\mathbf{P}$ is the state error covariance matrix and $\mathbf{Q}$ is the process noise covariance.

The world-frame position measurement $\mathbf{z}_k$ is derived from event-frame detections. A known physical diagonal length between the propellers and the pixel length between the detected propellers is used to estimate depth using the stereo disparity and camera intrinsics. The depth obtained is used to convert the coordinate of the quadrotor center to the world frame. The downward-facing event camera is mounted on the observer quadrotor at a known position and orientation. The quadrotor center in the camera frame is transformed to the observer frame using the rigid transformation. Using the known pose of the observer, this quadrotor center is transformed from the observer frame to the world frame. The position KF observation matrix and gain is

$$\mathbf{H} = [\mathbf{I}_3\ \mathbf{0}_3], \tag{11}$$

$$\mathbf{K} = \mathbf{P}\mathbf{H}^\top(\mathbf{H}\mathbf{P}\mathbf{H}^\top + \mathbf{R})^{-1}, \tag{12}$$

the monocular depth estimation using the event frame has inherent uncertainty. To address that, we use measurement noise covariance $\mathbf{R} = \mathrm{diag}(\sigma_{\mathrm{lat}}^2, \sigma_{\mathrm{lat}}^2, \sigma_{\mathrm{depth}}^2)$, with separate lateral and depth noise terms.

2) Orientation Estimation: The orientation prediction makes use of differential propeller frequencies which produces a roll angular velocity in the body frame [25] as

$$\omega_x^{\mathrm{body}} = k_{\mathrm{roll}} \left(\sum_{i \in \mathrm{right}} \Omega_i^2 - \sum_{i \in \mathrm{left}} \Omega_i^2 \right) \tag{13}$$

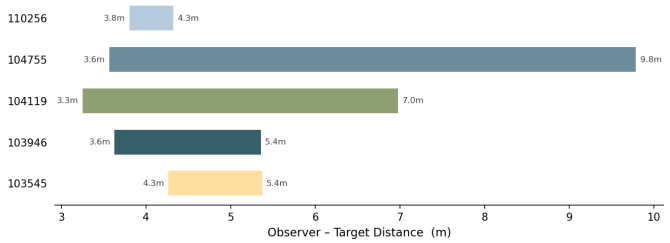
The body-frame angular velocity $\boldsymbol{\omega}^{\mathrm{body}}$ is rotated into the world frame using the current $\hat{\mathbf{b}}_z$ and used to propagate the unit vector using a constant angular velocity model

$$\hat{\mathbf{b}}_{z,k+1|k} = \hat{\mathbf{b}}_{z,k} + (\boldsymbol{\omega}_w \times \hat{\mathbf{b}}_{z,k})\, \Delta t, \quad \hat{\mathbf{b}}_z = \frac{\hat{\mathbf{b}}_z}{\|\hat{\mathbf{b}}_z\|} \tag{14}$$

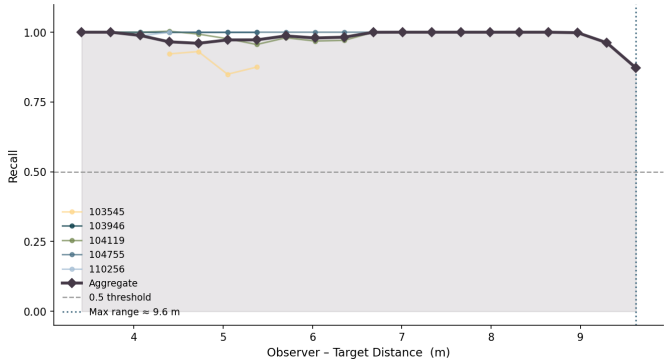
with covariance update given as follows

$$\mathbf{P}_{\mathrm{ori}} = \mathbf{P}_{\mathrm{ori}} + \mathbf{Q}_{\mathrm{ori}}\, \Delta t. \tag{15}$$

The event stream of a propeller disk forms an ellipse on the image plane of the event frame. We find the best fit ellipse [26] on the propeller with highest number of events. The ellipse encodes orientation information [11], [27]. We estimate the normal of the propeller by reverse projection of the best-fit ellipse through the intrinsics of the camera and the



(a) Observer-Target Distance Range per Test Sequence



(b) Detection Recall vs Observer-Target Distance

Fig. 4. The propeller detection results on the test dataset a) The figure shows range of distance between the target and the observer quadrotor in each sequence, b) shows detection recall vs the distance.

P1E formulation [11]. This formulation yields a measurement of the body z-axis $\hat{\mathbf{b}}_z$ in the world frame. The standard Kalman update is applied with $\mathbf{H} = \mathbf{I}_3$. The measurement noise covariance $\mathbf{R}_{\text{ori}} = \sigma_m^2 \mathbf{I}_3$ reflects uncertainty in ellipse backprojection.

Roll and pitch are extracted from the filtered $\hat{\mathbf{b}}_z$ via a yaw-free decomposition:

$$\phi = \arcsin(-\hat{b}_{z,y}), \quad \theta = \arctan 2(\hat{b}_{z,x},\ \hat{b}_{z,z}), \tag{16}$$

the yaw component is not observable from the propeller ellipse. When the disk rotates about its normal, it produces an identical projection, so yaw cannot be recovered from the ellipse geometry alone [11]. Thus, we restrict the orientation estimation to roll and pitch.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

The experiments were conducted outdoors using two quadrotors in a leader-follower formation. The follower acted as an observer with a downward-facing event camera. It carried the Prophesee IMX636 event sensor, such that the target's spinning propellers were in the camera's field of view during flight. The leader was designated as the target. A total of six flight sequences were recorded in the outdoor environment. The observer moved at nearly fixed altitude while the target performed a variety of maneuvers beneath it while varying the vertical separation between the two vehicles. In some trials,

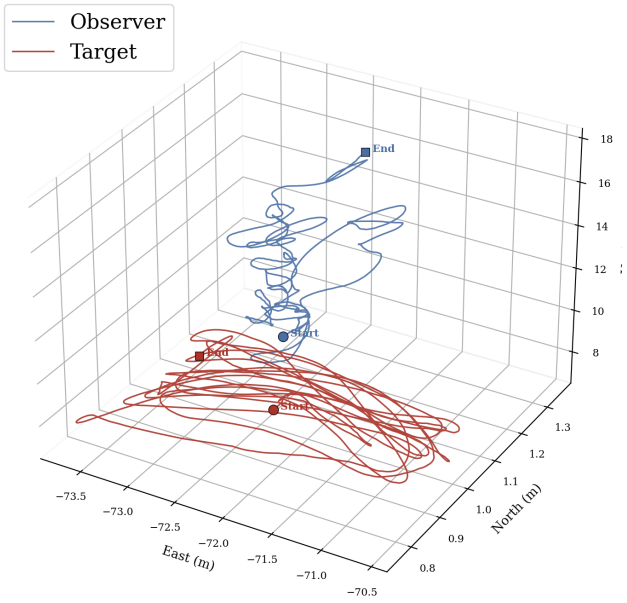


Fig. 5. Target-observer 3-D trajectories in a single validation flight sequence

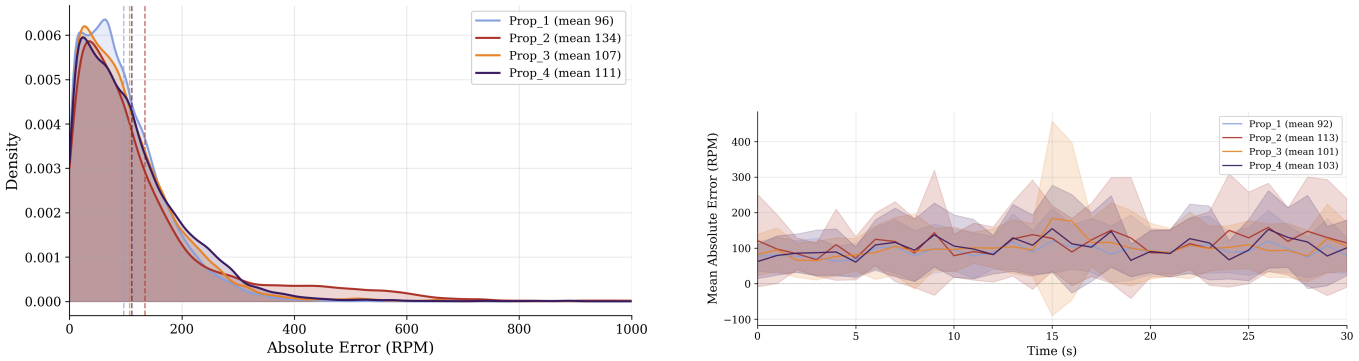
the target performs aggressive lateral motions. In all the flight sequences, the four propellers of target are visible, except a few instances where it went out the frame. The performance of the propeller detection and the range of distances in the test sequences is shown in Fig. 4

Telemetry was performed using Real-Time Kinematic Global Navigation Satellite System (RTK-GNSS) and Inertial Measurement Unit (IMU) to record the position and orientation of the two quadrotors. In addition to states, the propeller RPM ground truth was obtained using an external sensor and logged synchronously with the event stream. For the purpose of development and parameter tuning, we used one of the sequences as validation data. This validation sequence had the target's four propellers operated between $143\,\text{Hz}$ and $197\,\text{Hz}$ with a mean of approximately $166\,\text{Hz}$, RPM across all motors. The trajectory of the target and the observer in the world frame is shown in Fig. 5.

B. Aggregate RPM Results

For each propeller, we computed Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). This was performed for the five flight sequences in the test data using both the propeller detection approaches, the density based clustering method and connect component labeling. Table I summarizes the overall performance in all propellers and datasets.

We found that estimating the RPM using both the detection methods achieved nearly identical accuracy. This is indicative since detection only plays a role in identifying the region of interest for frequency computation. Thus, the choice of the propeller detection method has negligible effect on RPM estimation accuracy. The connected component labeling method was used in the state estimation pipeline, since it was



(a) RPM error distribution

(b) Error over time

Fig. 6. The RPM error analysis for the connected component labeling based propeller detection. a) The figure shows per-propeller absolute error distributions across the five test flight sequences, b) shows propeller frequency estimation error over time. The error was nearly identical for the HDBSCAN based method.

TABLE I
PROPELLER FREQUENCY ESTIMATION PERFORMANCE

Method	MAE	RMSE	MAPE	Corr.
CC	136.2±63.0	360.9±55.0	2.75%	0.442
HDBSCAN	135.3±63.3	360.6±54.5	2.73%	0.445

faster than the density based clustering approach. The detected propeller regions on an event frame and their respective tracking IDs are shown in Fig. 3.

C. Error Analysis

We performed an analysis with the absolute error distributions per-propeller aggregated across all five test flight sequences for both the detection methods. The results were nearly identical for both methods exhibiting similar distribution shapes, suggesting the propeller detection method did not influence RPM estimation. The error histograms for connected component based detection is shown in Fig. 6. The histogram aggregates all individual sample-level errors with a density estimation overlay, and the vertical lines indicate mean and median error. The error for Prop 2 shows the widest spread due to an outlier dataset.

The temporal evolution of mean absolute error with $\pm 1\sigma$ bands (first 30 seconds) is shown in Fig. 6. We define the windowed error for propeller p in time bin $W_k = [k, k+1)$:

$$\bar{e}_p(k) = \frac{1}{|W_k|} \sum_{t \in W_k} |\Omega_{\text{est},p}(t) - \Omega_{\text{gt},p}(t)|. \tag{17}$$

The error bands are computed across the datasets in each time bin. It was found that both detection methods exhibit a stable error throughout the sequence with no significant transients.

D. Per-Propeller Analysis

The redundant events of the background can cause significant noise despite the use of a spatio-temporal clustering filter. The MAE per motor and the correlation for both methods are in Table II. Prop 1 achieves the lowest MAE, while Prop 2 exhibits the highest. This variation is due to partial occlusions during specific flight maneuvers. Prop 4 achieves the highest correlation for both methods.

TABLE II
PER-PROPELLER MAE (RPM) AND CORRELATION

	Prop 1	Prop 2	Prop 3	Prop 4
CC MAE	117.6	171.7	127.2	128.2
HDBSCAN MAE	117.1	172.3	122.8	129.1
CC Corr.	0.372	0.281	0.461	0.655
HDBSCAN Corr.	0.373	0.281	0.471	0.655

E. State Estimation Results

To analyze the state estimation pipeline, we computed all the components of position and velocity. The orientation of the body z-axis was decomposed in roll and pitch angles. The comparison is reported in Table III across the five flight sequences. The ground-truth states were obtained using RTK-GNSS and IMU. Since the RPM KF is initialized from the first measurement to avoid bias from a fixed prior, the first $300\,\text{ms}$ of each sequence is excluded from evaluation to allow the sliding-window buffer to fill.

The z-axis dominates position error due to noise from the monocular depth estimation, while the lateral axes benefit from direct camera measurements. Since the diagonal pixel length between the propellers is estimated using a noisy blob centroid, there is a systematic depth uncertainty. The estimation pipeline infers the velocity via the position-velocity cross-covariance, yielding mean RMSE below $0.5\,\text{m}\,\text{s}^{-1}$ on all axes

In the five test flight sequences, the quadrotor exhibits lateral motion more than forward-backward motion. Thus, the roll range is larger than the pitch. Though the mean RMSE for both roll and pitch are comparable, The orientation KF estimates the roll stronger than the pitch taking advantage of left-right differential in propeller frequencies. The yaw is not estimated since it is not observable from ellipse fitting. The sequence 104119, in Table III, exhibits the highest errors in both position and orientation, attributable to aggressive lateral maneuvers

TABLE III
STATE ESTIMATION RMSE ACROSS FIVE TEST SEQUENCES.

	Position			Velocity			Orientation	
Sequence	x (m)	y (m)	z (m)	x ($\mathrm{m\,s^{-1}}$)	y ($\mathrm{m\,s^{-1}}$)	z ($\mathrm{m\,s^{-1}}$)	roll (°)	pitch (°)
103545	0.363	0.138	0.919	0.675	0.187	0.590	5.12	9.12
103946	0.059	0.097	1.092	0.086	0.134	0.258	5.78	7.47
104119	0.945	0.253	1.147	0.916	0.228	0.697	6.13	13.45
104755	0.044	0.027	0.798	0.096	0.075	0.312	3.16	6.98
110256	0.070	0.102	0.526	0.237	0.330	0.428	8.43	8.85
Mean±Std	0.30±0.35	0.12±0.07	0.90±0.22	0.40±0.34	0.19±0.09	0.46±0.17	5.72±1.70	9.18±2.29

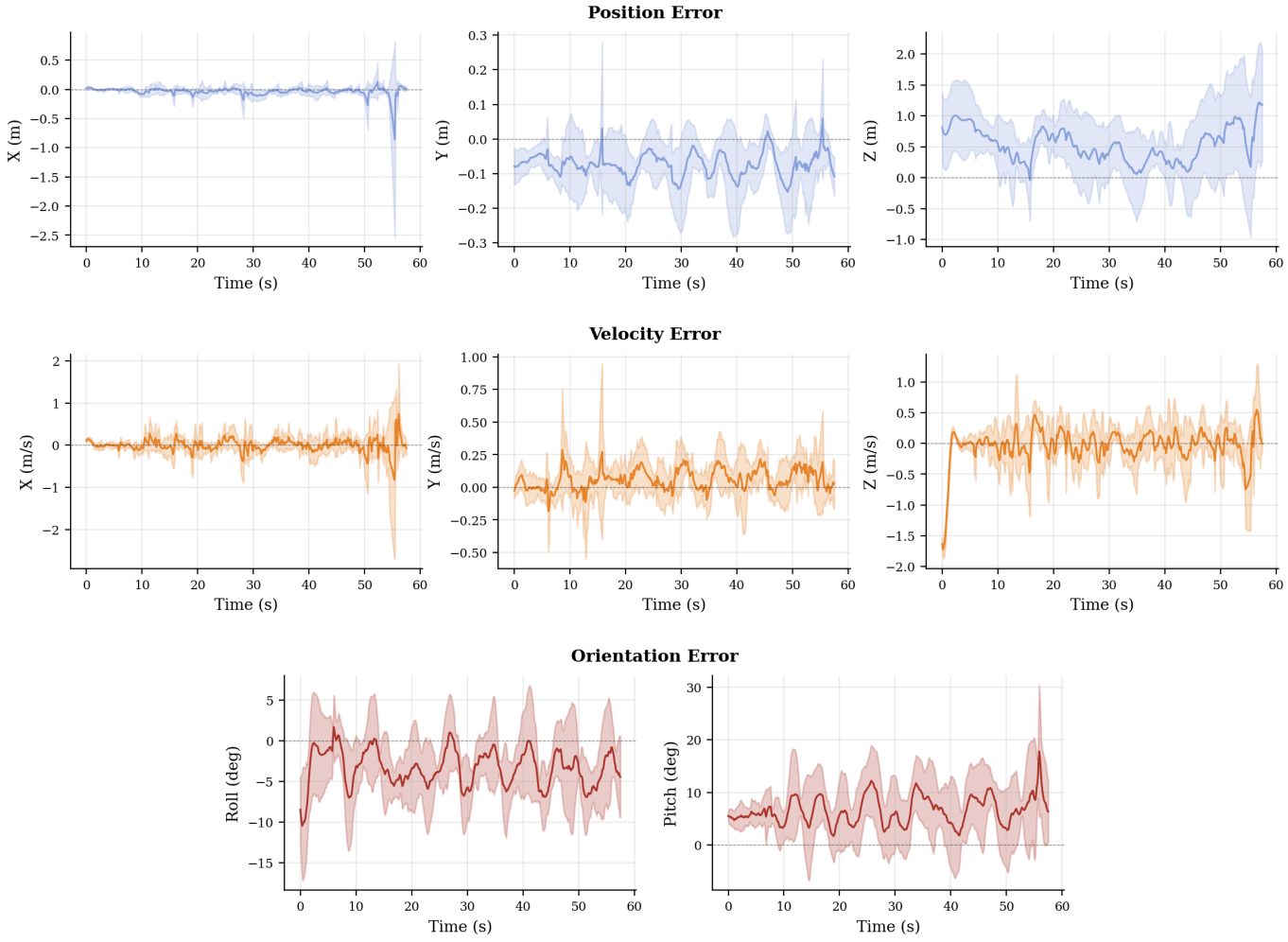


Fig. 7. Estimation error (mean $\pm 1\sigma$ across five sequences) over time. Top row: position error (m). Middle row: velocity error ($\mathrm{m\,s^{-1}}$). Bottom row: orientation error (°). The shaded band represents ± 1 standard deviation across sequences.

during that flight. The temporal evolution of the estimation error for all the element of state vector is shown in Fig. 7.

F. Discussion

We made a few observations from the multi-dataset evaluation. The propeller detection method has a negligible effect on the accuracy of the RPM. CC labeling and HDBSCAN clustering achieve nearly identical overall accuracy. The per-propeller error analysis suggests that the RPM estimation pipeline dominates accuracy, not the spatial detection method, since the propeller region is used only to identify the region of interest. CC processes on average faster per sequence than HDBSCAN. Therefore, the connected component method is preferred for real-time or resource-constrained applications.

The variance of the per-propeller RPM error indicates that propeller-specific factors such as viewing angle, blade geometry, and occlusion affect the quality of the estimation more than

the detection algorithm. The accuracy of the RPM depends on the quality of the event signal from the propeller region. The detection method may merge adjacent blobs, or miss them when the propellers are outside the frame, specifically during fast lateral motion.

From RPM-driven state estimation we found that the z-axis dominates the position error due to the monocular depth uncertainty, while lateral axes benefit from direct camera measurements. The orientation KF recovers roll more accurately than pitch, reflecting the stronger observability of the left–right motor differential compared to the front–rear differential, due to a left-right velocity bias of the target in the test dataset. The estimated body z-axis from the orientation filter determines the thrust direction used in the position prediction step, enabling tilt-corrected acceleration estimates.

V. CONCLUSIONS AND FUTURE WORK

This paper demonstrates an event-based approach to relative state estimation for quadrotors. Propeller frequencies are extracted from the periodic event signal in each propeller's region of interest and used to enhance the orientation estimation of observed UAVs. A system of coupled KFs, one for position and the other for orientation, fuse RPM-derived thrust and angular velocity with camera-derived position and tilt measurements. The position KF estimates the position and velocity of the quadrotor, while the orientation KF recovers roll and pitch. The propeller frequencies provide a direct vertical acceleration signal complementary to the camera-derived position, and the differential motor RPMs allow estimation of the attitude through the geometry of the propeller disk. This proposed relative state estimation method is evaluated across five outdoor flight sequences. This event-based state estimation approach can augment the capabilities of swarms of aerial robots for future planetary explorations. Future work will explore multi-quadrotor scenarios and further leverage the use of event vision to extract features based on periodic events and geometric primitives to estimate relative states.

REFERENCES

[1] S.-J. Chung, A. A. Paranjape, P. Dames, S. Shen, and V. Kumar, "A survey on aerial swarm robotics," *IEEE Transactions on robotics*, vol. 34, no. 4, pp. 837–855, 2018.

[2] V. Walter, N. Staub, A. Franchi, and M. Saska, "Uvdar system for visual relative localization with application to leader–follower formations of multirotor uavs," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2637–2644, 2019.

[3] J. Ulrich, J. Blaha, A. Alsayed, T. Rouček, F. Arvin, and T. Krajník, "Real time fiducial marker localisation system with full 6 dof pose estimation," *ACM SIGAPP Applied Computing Review*, vol. 23, no. 1, pp. 20–35, 2023.

[4] M. Faessler, E. Mueggler, K. Schwabe, and D. Scaramuzza, "A monocular pose estimation system based on infrared LEDs," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2014, pp. 907–913.

[5] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis *et al.*, "Event-based vision: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 1, pp. 154–180, 2020.

[6] A. R. Vidal, H. Rebecq, T. Horstschaefer, and D. Scaramuzza, "Ultimate slam? combining events, images, and imu for robust visual slam in hdr and high-speed scenarios," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 994–1001, 2018.

[7] N. J. Sanket, C. D. Singh, C. M. Parameshwara, C. Fermüller, G. C. De Croon, and Y. Aloimonos, "Evpropnet: Detecting drones by finding propellers for mid-air landing and following," *arXiv preprint arXiv:2106.15045*, 2021.

[8] M. Saska, "Large sensors with adaptive shape realised by self-stabilised compact groups of micro aerial vehicles," in *Robotics Research: The 18th International Symposium ISRR*. Springer, 2019, pp. 101–107.

[9] J. Stadler, H. Kayal, A. Maurer, J. Mutter, and C. Riegler, "Exploring transient phenomena in the martian atmosphere," in *76th International Astronautical Congress (IAC)*, 2025.

[10] C. Riegler, H. Kayal, A. Maurer, J. Mutter, and J. Stadler, "Towards autorotation landers for communication and sensor networks on mars," in *76th International Astronautical Congress (IAC)*, 2025.

[11] V. Gaudillière, G. Simon, and M.-O. Berger, "Perspective-1-ellipsoid: formulation, analysis and solutions of the camera pose estimation problem from one ellipse-ellipsoid correspondence," *International Journal of Computer Vision*, vol. 131, no. 9, pp. 2446–2470, 2023.

[12] D. Rice and A. Abdelkefi, "Evolution of space drones for planetary exploration: A review," *Progress in Aerospace Sciences*, vol. 97, pp. 61–105, 2018.

[13] A. Pankine, K. Aaron, M. Heun, K. Nock, R. Schlaifer, C. Wyszkowski, A. Ingersoll, and R. Lorenz, "Directed aerial robot explorers for planetary exploration," *Advances in Space Research*, vol. 33, no. 10, pp. 1825–1830, 2004.

[14] T. Tzanetos, M. Aung, J. Balaram, H. F. Grip, J. T. Karras, T. K. Canham, G. Kubiak, J. Anderson, G. Merewether, M. Starch *et al.*, "Ingenuity mars helicopter: From technology demonstration to extraterrestrial scout," in *2022 IEEE aerospace conference (AERO)*. IEEE, 2022, pp. 01–19.

[15] T. Tzanetos, J. Bapst, G. Kubiak, L. P. Tosi, S. Sirlin, R. Brockers, J. Delaune, H. F. Grip, L. Matthies, J. Balaram *et al.*, "Future of mars rotorcraft-mars science helicopter," in *2022 IEEE aerospace conference (AERO)*. IEEE, 2022, pp. 1–16.

[16] R. D. Lorenz, E. P. Turtle, J. W. Barnes, M. G. Trainer, D. S. Adams, K. E. Hibbard, C. Z. Sheldon, K. Zacny, P. N. Peplowski, D. J. Lawrence *et al.*, "Dragonfly: A rotorcraft lander concept for scientific exploration at titan."

[17] P. Voosen. (2026, mar) Nasa plans to send a nuclear-powered spacecraft to mars in 2028. [Online]. Available: https://www.science.org/content/article/nasa-plans-send-nuclear-powered-spacecraft-mars-2028

[18] R. Spetlik, T. Uhrova, and J. Matas, "Efficient real-time quadcopter propeller detection and attribute estimation with high-resolution event camera," in *Image Analysis*, 2025.

[19] R. Spetlik, M. Pliska, V. Vrba, and J. Matas, "Helixtrack: Event-based tracking and rpm estimation of propeller-like objects," 2026. [Online]. Available: https://arxiv.org/abs/2603.09235

[20] X. Chen, J. Xu, W. Ding, H. Wang, X. Luo, R. Duan, J. Chen, X. Wang, Y. Liu, and X. Chen, "Count every rotation and every rotation counts: Exploring drone dynamics via propeller sensing," *arXiv preprint arXiv:2511.13100*, 2025.

[21] X. Lu, Y. Zhou, J. Mai, K. Dai, Y. Xu, and S. Shen, "Event-based visual-inertial state estimation for high-speed maneuvers," *IEEE Transactions on Robotics*, 2025.

[22] R. J. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," in *Pacific-Asia conference on knowledge discovery and data mining*. Springer, 2013, pp. 160–172.

[23] J. Alori, A. Descoins, javier, F. Lezama, KotaYuhara, D. Fernández, A. Castro, fatih, David, R. C. Linares, F. Kurucz, B. Ríos, shafu.eth, K. Nar, D. Huh, and Moises, "tryolabs/norfair: v2.2.0," Jan. 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7504727

[24] G. Magrini, L. Berlincioni, F. Becattini, L. Cultrera, and P. Pala, "Drone detection with event cameras," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2025, pp. 4703–4714.

[25] R. Mahony, V. Kumar, and P. Corke, "Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor," *IEEE robotics & automation magazine*, vol. 19, no. 3, pp. 20–32, 2012.

[26] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least squares fitting of ellipses," in *Proceedings of 13th international conference on pattern recognition*, vol. 1. IEEE, 1996, pp. 253–257.

[27] D. Lo´ pez de Ipin a, P. R. Mendonça, A. Hopper, and A. Hopper, "Trip: A low-cost vision-based location system for ubiquitous computing," *Personal and Ubiquitous Computing*, vol. 6, no. 3, pp. 206–219, 2002.