2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Direct Angular Rate Estimation Without Event Motion-Compensation At High Angular Rates

Matthew Ng, Xinyu Cai, and Shaohui Foong

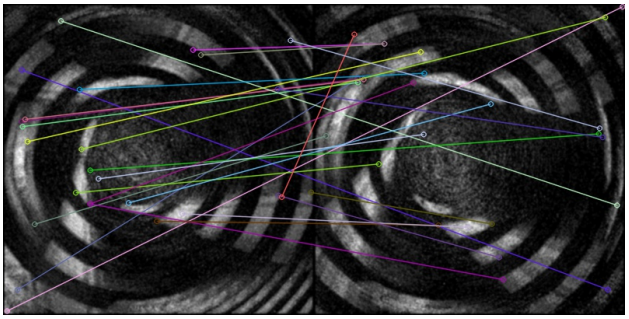
***Abstract*— Feature-based methods are a popular method for camera state estimation using event cameras. Due to the spatiotemporal nature of events, all event images exhibit smearing of events analogous to motion blur for a camera under motion. As such, events must be motion compensated to derive a sharp event image. However, this presents a causality dilemma where motion prior is required to unsmear the events, but a sharp event image is required to estimate motion. While it is possible to use the IMU to develop motion prior, it has been shown that the limited dynamic range of $\pm 2000\,^\circ/\mathrm{s}$ is insufficient for high angular rate rotorcrafts. Furthermore, smoothing of motion-compensated images due to actual event detection time latency in event cameras severely limits the performance of feature-based methods at high angular rates. This paper proposes a Fourier-based angular rate estimator capable of estimating angular rates directly on non-motion compensated event images. This method circumvents the need for external motion priors in camera state estimation and sidesteps problematic smoothing of features in the spatial domain due to motion blur. Lastly, using an NVIDIA Jetson Xavier NX, the algorithm is demonstrated to be real-time performant up to $3960\,^\circ/\mathrm{s}$.**

I. Introduction

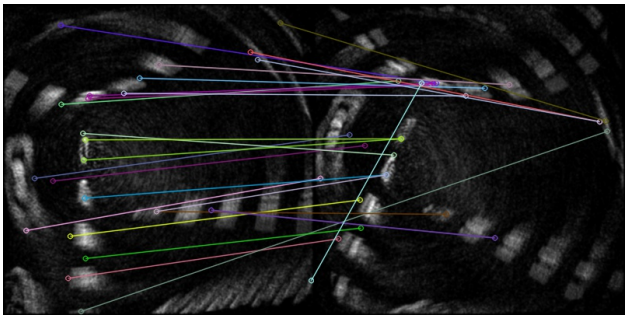
The generation of event images depends on aggregation functions that define the relative spatial influence of events. Due to the spatiotemporal nature of events, prior motion estimates are required to motion-compensate events prior to aggregation. In camera state estimation, this presents a causality dilemma where motion prior is required to estimate motion [1].

As such methods that depend on motion-compensated event images employ IMUs [2] or exploit environmental conditions [3] to bootstrap the algorithm. The Contrast Maximisation framework [4] and Spiking Neural Networks [5] are an exception to this requirement as both are event-only methods capable of directly recovering motion estimates. However, both methods' computation complexity and estimation accuracy tradeoff present a serious obstacle to the tight computation deadline of high dynamic motion.

Consider systems such as the freerotor [6] or damaged quadrotor [7] where rotation about a single axis dominates the state estimates. It has been demonstrated in [1] that as rotational velocity increases, actual event detection time latency among projected pixels describing the same texture becomes more pronounced. Each sub-figure of Fig. 1 shows two consecutive images taken 10ms apart, with a rotational difference of 25° and Fig. 2 the reference image of the IKEA Storabo rug. As seen in Fig. 1b, the normally sharp dashed road lines are smeared by events describing the same texture not being propagated to the correct location due to timing inaccuracies. In event images that use temporal difference to weight motion compensation, this manifests itself as motion blur.



(a) non-motion compensated event image



(b) motion compensated event image

Fig. 1. Two consecutive event images, matched using SIFT feature detector and FLANN based matcher

Feature-based image registration is a popular method for event camera state estimation, with recent works demonstrating performance up to rotational rates of $1150\,^\circ/\mathrm{s}$ [7]. However, above common gyroscope saturation point of $2000\,^\circ/\mathrm{s}$, smoothing of edges due to poor motion compensation degrades the performance of feature-based methods. Note the poor feature matches in Fig. 1 for both motion and non-motion compensated event images. To accurately motion-compensate event images above this saturation point, developing a method capable of estimating angular velocity without prior motion knowledge is necessary.

This paper proposes an angular rate estimator capable of estimating angular rates directly on non-motion compensated event images. Based on the Fourier-Mellin Transform, shown to be performant in frame-based cameras [6], this paper

The authors are with the Faculty of Engineering Product Development, Singapore University of Technology and Design, 8 Somapah Rd, Singapore 487372. (Corresponding Author: foongshaohui@sutd.edu.sg)

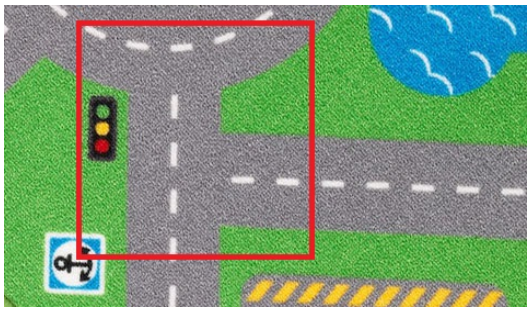


Fig. 2. Visual reference of IKEA Storabo Rug with red box indicating approximate event camera FOV in Fig. 1.

adapts the transform for estimating using event cameras above the gyro saturation limit. The utility of the proposed algorithm is demonstrated by deployment on resource-constrained systems with real-time computation using an NVIDIA Jetson Xavier NX onboard a spinning rotorcraft.

II. Methodology

A. Fourier-Mellin Transform

The Fourier-Mellin Transform (FMT) is a fourier-type similarity transform defined as:

$$\mathcal{M}(u,v) = \frac{1}{2\pi}\int_0^{\infty}\int_0^{2\pi} f(r,\theta)\, r^{-ju} e^{-jv\theta} \mathrm{d}\theta \frac{\mathrm{d}r}{r}. \quad (1)$$

Let $r = e^{\rho}$, the FMT can be shown by the reverse chain rule to be expressed as a Fourier transform of a log-polar transformed function (see (2)).

$$\mathcal{M}(u,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{2\pi} f(e^{\rho},\theta)\, e^{-ju\rho} e^{-jv\theta} \mathrm{d}\theta \mathrm{d}\rho \quad (2)$$

Due to its rotation, scale, and translation invariant properties [8], the FMT has found popular use in image registration [9], [10].

Given two images $I_1(x,y)$ and $I_2(x,y)$, the image registration algorithm is generally broken into 4-steps with the first 3-steps being the FMT:

1) The fourier transform $X(u,v) = \mathcal{F}[I(x,y)]$ is first taken, where $\mathcal{F}[*]$ is the Fourier Transform operation.
2) A high-pass filter is applied in the frequency domain to $\|X(u,v)\|$, and the log-polar transform is taken. This transforms rotation and scale in $I(x,y)$ to translations.
3) A second Fourier transform is applied to the log-polar image.
4) Lastly, phase correlation is used to recover rotation and scale between two transformed images.

B. Proposed Algorithm

This process is extended to tackle issues unique to the event image. Let $\mathcal{E} = \{e_k\}_{k=1}^{N}$ describe an event stream and each event $e_k = (t_k, \mathbf{x}_k)$ where t_k is the timestamp, $\mathbf{x}_k = (x_k, y_k)$ is the x-y position of the event in an event image. The notation for sign polarity is dropped as the method is independent of intensity change. Events are first aggregated using a gaussian aggregation function (see (3)).

$$H(\mathbf{x'}_k) = \sum_{k=1}^{N} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\alpha}{2\sigma^2}\right) \forall \mathbf{x} = \lfloor \mathbf{x'}_k \rceil \pm r\sigma \quad (3)$$

where $\alpha = (x - x_k)^2 + (y - y_k)^2$, $\lfloor \cdot \rceil$ is the rounding function, and $1 \leq r \leq 3$. This produces a non-motion compensated event image. The event image is then apodized using a Tukey window (see (4) and (5)).

$$\omega[n] = \begin{cases} \frac{1}{2}\left\{1 + \cos\left(\frac{2\pi}{r}[n - r/2]\right)\right\} & 0 \leq n < r/2 \\ 1 & \frac{r}{2} \leq n < 1 - \frac{r}{2} \\ \frac{1}{2}\left\{1 + \cos\left(\frac{2\pi}{r}[n - 1 + r/2]\right)\right\} & 1 - \frac{r}{2} \leq n \leq 1 \end{cases} \quad (4)$$

$$H_{apo}(\mathbf{x'}_k) = H(\mathbf{x'}_k) \odot (\omega[x] \otimes \omega[y]) \quad (5)$$

where r is the ratio between the cosine-taper and the entire length of the window, $\odot$ is the Hadamard product, and $\otimes$ is the outer product. Apodization is essential to ensure that the event image is circular continuous at the edges. The Tukey window allows precomputation of the specific contribution of $(\omega[x] \otimes \omega[y])$, while minimizing changes to the overall magnitude spectra.

The first two steps of the FMT are taken, yielding a log-polar transform of the event image $H(\rho,\theta)$. It was shown in [11] that 1-D phase-only correlation is viable for rotation estimation. The ability to estimate scale change in the FMT can be dropped in favour of computing two less Fourier Transform. A similar approach is taken in this paper, excluding the effective line extraction in [11]. A 1-D Discrete Fourier Transform is taken along the log axis of $H(\rho,\theta)$

$$\mathcal{H}(\rho,v) = \sum_{\theta=0}^{\Theta-1} e^{-2\pi j \frac{k\theta}{\Theta}} H(\rho,\theta) \forall \rho \in \{0,...,\mathrm{P}-1\} \quad (6)$$

followed by 1-D normalized cross-correlation taken along the log axis, yielding column-wise peaks. The peaks indicate the estimated rotation, radial from the center of $\mathcal{H}(\rho,v)$. A sum reduction is taken to yield a vector of peaks

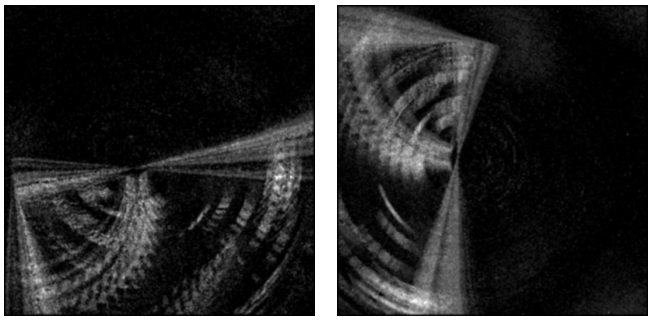
$$h(\theta) = \sum_{\rho=0}^{\mathrm{P}-1} \mathcal{F}^{-1}\left\{\frac{\mathcal{H}_1(\rho,v) \odot \mathcal{H}_2^*(\rho,v)}{|\mathcal{H}_1(\rho,v) \odot \mathcal{H}_2^*(\rho,v)|}\right\}. \quad (7)$$

Lastly, the relative change in rotation between the two frames can be recovered by taking the argmax of $h(\theta)$

$$\Delta\theta = \underset{\theta}{\mathrm{argmax}}\{h(\theta)\}. \quad (8)$$

The algorithm assumes, optimally, that correspondence between the two images exists and rotation is recoverable. However, this may not be true as sparsity in the event image combined with large rotations due to dropped events in $\mathcal{E}$ can cause the algorithm to fail. Fig. 3 illustrates such a case where the time difference between two consecutive

frames is $40\,\mathrm{ms}$ and large sparsity where more than half of the image contains zero events. Note how Fig. 3a contains approximately half the events compared to Fig. 3b, at 66000 and 131000 events respectively, features less motion blur due to time stamping latency. This also shows how motion blur in event images is a function of total events aggregated.



(a) Frame 1630 (b) Frame 1631

Fig. 3. Two consecutive event images of a failed rotation estimate

To improve robustness, the results of the algorithm, $\dot{\theta} = \Delta\theta/\Delta s$ is fed into a perceptron-like median filter [12]. This perceptron tracks the median through a series of perturbations using a learned amplitude. The amplitude depends on the expected magnitude of the tracked state, affecting the track's resolution and the maximum rate it can track (see (9)).

$$\dot{\theta}_{med}(t) = \dot{\theta}_{med}(t-1) + a \cdot Sgn\left(\dot{\theta}(t) - \dot{\theta}_{med}(t-1)\right) \quad (9)$$

where a is the learned amplitude and $Sgn(*)$ is the Signum function. This filter was adopted for its fast computation speed and good long-term median tracking performance on stationary and non-stationary median. It was found empirically that $a = \dot{\theta}_{expected}/100$ is a good starting point for assessing amplitude suitability. Indirectly, this means it will take 100 perturbations to reach $\dot{\theta}_{expected}$. It is a non-issue for the current use case as the filter tracks a moving median from 0. Lastly, if $a = 20$, then the resolution of the filter is ± 20.

III. Experiments

In this section, three experiments are conducted to assess the algorithm. The first experiment validated the ability to estimate angular rate above the gyro saturation point using non-motion compensated event images. The second experiment verifies the algorithm's performance on a dynamic aerial platform. Lastly, the third experiment demonstrates the algorithm's performance on the rig at twice the angular rate of experiment one, demonstrating performance guarantees for smaller, faster-rotating rotorcrafts.

A. Validation of proposed algorithm

For this experiment, the DVXplorer Mini by Inivation is mounted to the rotation rig, as shown in Fig. 4. The rig comprises a T-Motor U8, a US Digital E6 Rotary Encoder, and an Odrive V3.6 that provides closed-loop control of the U8 motor. The DVXplorer Mini is interfaced via USB-C to NVIDIA Jetson Xavier NX. Lastly, the LattePanda V1 provides a wireless interface to the Odrive V3.6.

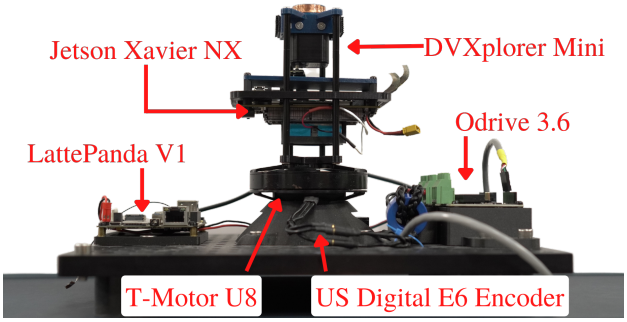


Fig. 4. System setup to accurately rotate the DVXplorer Mini.

As each event corresponds to a pixel-wise intensity change in the event camera, the rate of event throughput is dependent on both the observed texture and the rate at with the texture changes (which can be understood as the rate at which pixels change intensity). This is easily illustrated using two Event-Camera dataset [13] which were generated using a DAVIS240C [14]. Event and IMU Angular Rate of the *'boxes_rotation'* and *'shapes_rotation'* are plotted. Given the *'boxes_rotation'* dataset is more richly textured than the simple *'shapes_rotation'*, Fig. 5 shows the event rate for the former is a magnitude order higher than the latter. This is despite both datasets undergoing similar rotation rates (see Fig. 6).

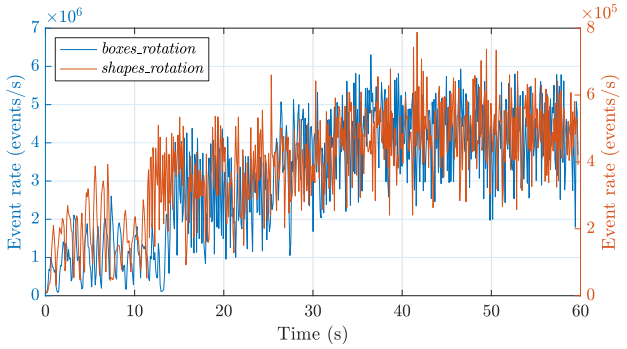
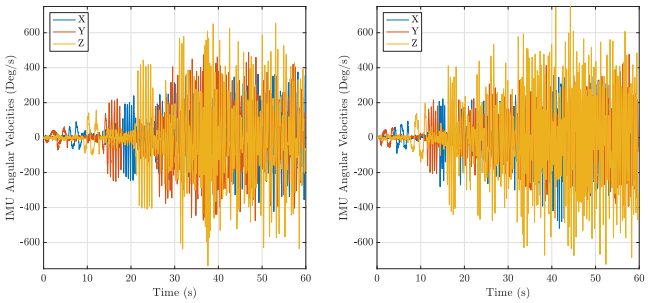


Fig. 5. Event rate of *'boxes_rotation'* (blue) and *'shapes_rotation'* (red)



(a) *'boxes_rotation'* (b) *'shapes_rotation'*

Fig. 6. IMU angular velocity of of *'boxes_rotation'* and *'shapes_rotation'*

The DVXplorer Mini has a maximum event throughput of 450 million events per second. However, typically the DVXplorer Mini transmits an event rate of between 10 million and 20 million events per second. To cope with the large number of events, the algorithm is broken into two nodes on a ROS network (see Fig. 7). The first node reads the event data from the data buffer and aggregates it into an event image. The second node transfers the event image from CPU to GPU memory and computes $\dot{\theta}_{med}$ using the GPU. On the Jetson Xavier NX, the first node takes approximately 10 ms, and the second node takes approximately 5 ms. Resulting in a fixed time offset of 15ms between receipt of event data and rotation estimate. However, the concurrent computation also guarantees a deterministic performance, as the event packet arrives at 100 Hz, one of the goals was to ensure the algorithm is also capable of estimating at 100 Hz.

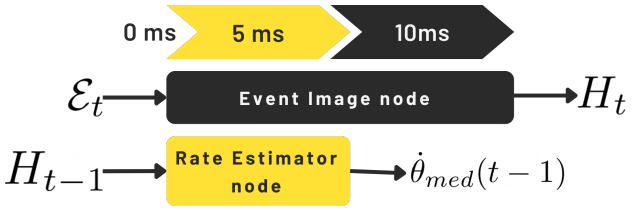


Fig. 7. Flow of data from events to angular rate estimate for each computation cycle

In this experiment, the camera was oriented to face the IKEA storabo rug. This rug was chosen as play rugs provide several none repeating features in a small area useful for camera ego-motion estimation. Similar rugs are also featured in prior event camera literature [2]. The rotation rate was set at 2160 °/s (6 Hz), just above the gyro saturation limit of 2000 °/s. Fig. 8 shows the rotation rate estimated by the proposed algorithm closely tracking the US Digital E6 with an average absolute percentage error of 2.50 %.

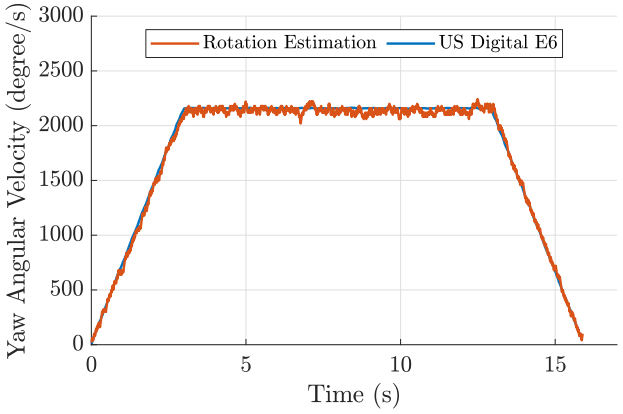


Fig. 8. Estimation of angular velocity without a priori for each 10ms event packet on the rotation rig

The event rate for the entire experiment is plotted in Fig. 9 with an average event rate of 13×10^6 event/s. As the aggregation function is the rate-limiting factor of the overall algorithm, the histogram of the time between successive event images is plotted in Fig. 10. The graph shows that

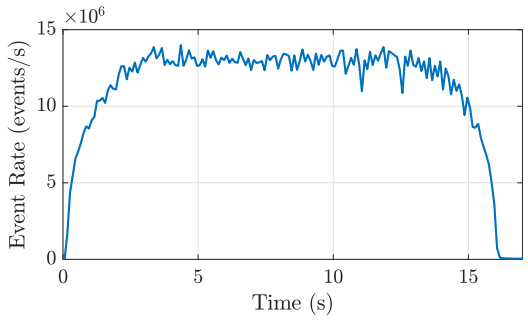


Fig. 9. Event rate of DVXplorer Mini rotating at 2160 °/s

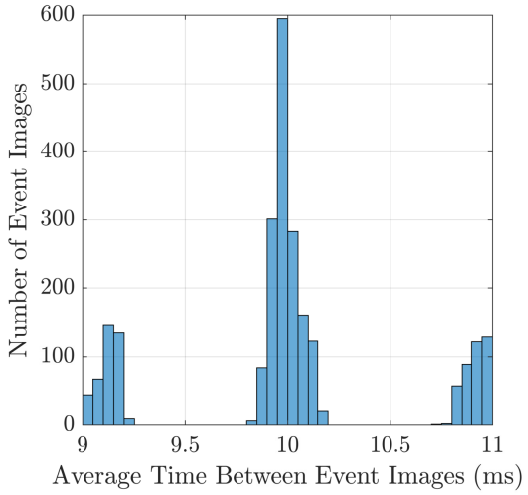


Fig. 10. Histogram of the time between succcessive event images generated by the event image node

the first node can satisfy the computation deadline of 10 ms most of the time, and the whole algorithm performs within expectations.

B. Validation in dynamic flight

The second experiment validates the performance of the algorithm on a dynamic platform. As the rig constrains the rotation to one axis, this test exposes the algorithm to 6-DOF motion. The rotorcraft used in this experiment is a tri-wing freerotor (see Fig. 11). The DVXplorer Mini and Jetson Xavier NX are mounted to its underbelly with the DVXplorer Mini oriented downwards. The UAV is controlled remotely via a Crazyradio PA and Crazyflie Bolt. A system of Optitrack Prime 41 cameras provides the ground truth of the flight.

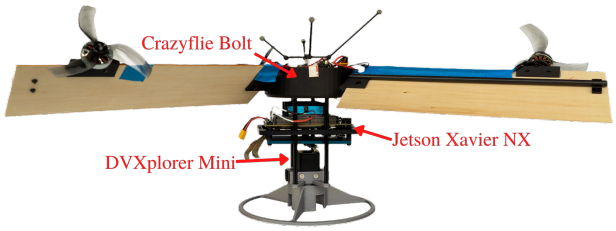


Fig. 11. Side view of the flying experimental setup

The in-flight estimation of angular velocity and yaw rate from the IMU and Optitrack are plotted in Fig. 12. As shown, the algorithm continued to estimate the yaw rate of the freerotor above the gyro saturation of $2000\,^\circ/\mathrm{s}$ with an average absolute percentage error of $4.68\,\%$. The necessity of alternative angular rate estimators becomes evident by plotting the relative distance of the motion capture markers to the ground. Fig. 13 shows the gyro saturating before the freerotor reaches its required takeoff rotation velocity.

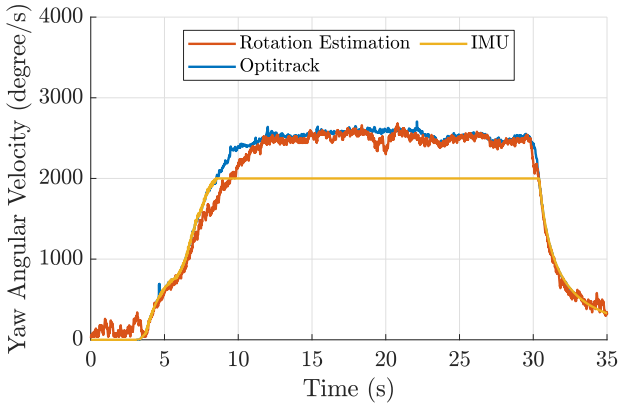


Fig. 12. In-flight estimation of angular velocity without a priori for each 10ms event packet

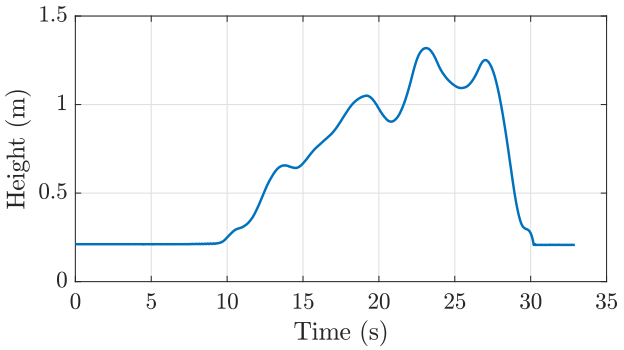


Fig. 13. Relative distance of Optitrack markers on the freerotor to ground

Similarly, the event rate and histogram of time between successive event images are plotted in Fig. 14 and Fig. 16 respectively. Fig. 14 shows a large dip in events from approximately 15×10^6 event/s to 7×10^6 event/s around the 15 s to 20 s mark due to the freerotor temporarily flying away from the rug.

It can be seen in Fig. 3, the freerotor flying at the edge of the rug around 16 s mark, and Fig. 15 shows the freerotor flying outside the rug. As there are very little texture outside the rug, the event rate plummets to between 20×10^3 event/s to 40×10^3 event/s. Despite the low texture, the algorithm continues to estimate well, as the median filter is able to reject any failed estimate. Lastly, Fig. 16 confirms that the algorithm can meet the computation deadline of 10 ms, matching the DVXplorer Mini frequency of approximately 100 Hz.

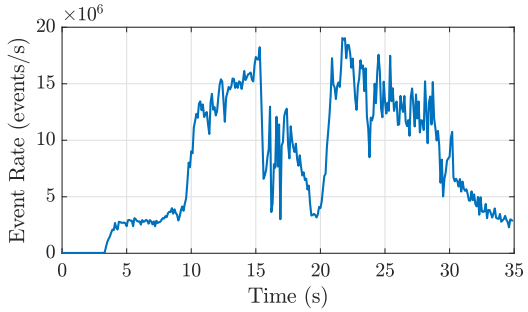
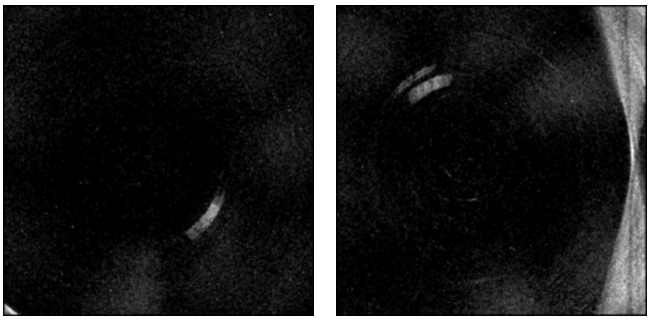


Fig. 14. Event rate of DVXplorer Mini during experimental flight



(a) Frame 1863 (b) Frame 1928

Fig. 15. Two event image frames extracted from 18 s and 19 s mark

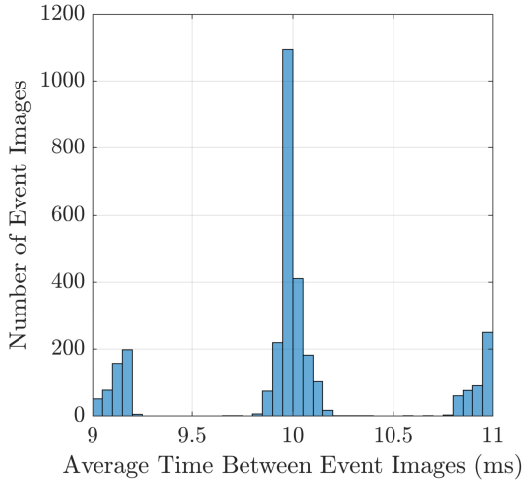


Fig. 16. Histogram of the time between succcessive event images generated by the event image node during the flight experiment

The algorithm's overall performance also confirms the estimator invariance to translation and scale change during the flight. The use of the frequency domain also demonstrates an alternative approach to state estimation that can be robust to significant changes in the number of events compared to feature-based approaches in the spatial domain.

C. Demonstration of upper limit

Dynamic angular rates in freerotors have been demonstrated up to $4160\,^\circ/\mathrm{s}$ [15]. This high rotational rate is largely due to the yaw angular rates necessary to achieve sustained lift. Consequently, various rate estimators were developed to

operate at the high angular rates needed by the freerotor [16], [17], [18]. However none of these methods can be leveraged or adapted for use on an event camera. A demonstration of performance at twice the gyro saturation limit qualifies the operating region to match that of prior literature. For operational safety, the rotation rig is physically limited to $3960\,^\circ/\mathrm{s}$ (11 Hz) to minimize vibrations due to rotational imbalances in the camera-computer setup.

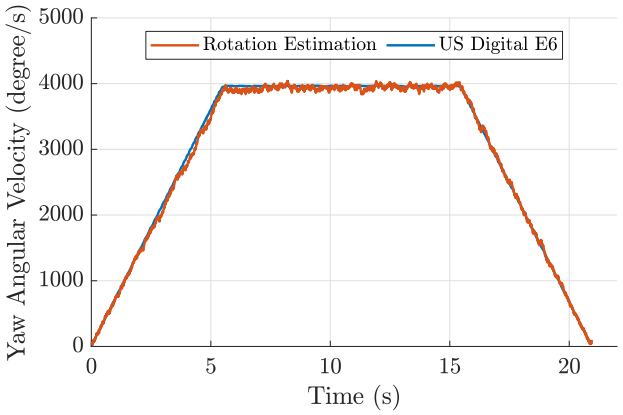


Fig. 17. Estimation of angular velocity without a priori for each 10ms event packet on the rotation rig

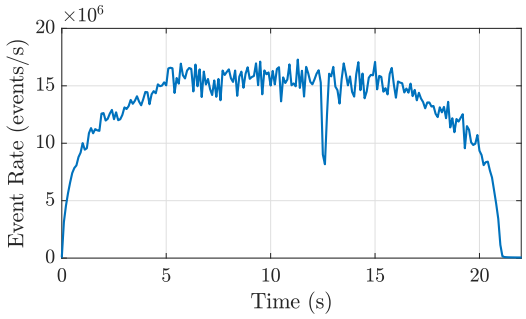


Fig. 18. Event rate of DVXplorer Mini rotating at $3960\,^\circ/\mathrm{s}$

Fig. 17 shows the close tracking performance of the proposed algorithm to the US Digital E6 with an average absolute percentage error of 2.33 %. This performance is expected as the theoretical sensing upper limit of the proposed algorithm is $18\,000\,^\circ/\mathrm{s}$. Specifically, the rotation estimator is limited to 180° per estimation cycle as only the upper half quadrants (0 to π) of the Fourier magnitude spectra are used in the log-polar transform. The corresponding event rate of the experiment is plotted in Fig. 18.

IV. Conclusion

This paper proposes a novel angular rate estimator for event cameras. It was shown to be capable of estimating in-plane rotations directly on non-motion compensated events through an event image. The algorithm is unaffected by motion blur in event images that typically cause feature-based methods to fail at high rotational speeds. The fast estimation speed and parallel computing of the two nodes enable the algorithm to perform in real-time using the Jetson Xavier NX, showing promising use as a motion primer in event frameworks dependent on motion-compensated events.

References

[1] M. Ng, Z. M. Er, G. S. Soh, and S. Foong, "Aggregation functions for simultaneous attitude and image estimation with event cameras at high angular rates," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4384–4391, 2022.

[2] 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.

[3] E. Mueggler, B. Huber, and D. Scaramuzza, "Event-based, 6-dof pose tracking for high-speed maneuvers," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 2761–2768.

[4] G. Gallego, H. Rebecq, and D. Scaramuzza, "A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3867–3876.

[5] M. Gehrig, S. B. Shrestha, D. Mouritzen, and D. Scaramuzza, "Event-based angular velocity regression with spiking networks," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 4195–4202.

[6] M. Ng, E. Tang, G. S. Soh, and S. Foong, "Shift: Selective heading image for translation an onboard monocular optical flow estimator for fast constantly rotating uavs," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 8525–8531.

[7] S. Sun, G. Cioffi, C. De Visser, and D. Scaramuzza, "Autonomous quadrotor flight despite rotor failure with onboard vision sensors: Frames vs. events," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 580–587, 2021.

[8] S. Raman and U. B. Desai, "2-d object recognition using fourier mellin transform and a mlp network," in *Proceedings of ICNN'95-International Conference on Neural Networks*, vol. 4. IEEE, 1995, pp. 2154–2156.

[9] B. S. Reddy and B. N. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE transactions on image processing*, vol. 5, no. 8, pp. 1266–1271, 1996.

[10] G. Wolberg and S. Zokai, "Robust image registration using log-polar transform," in *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, vol. 1. IEEE, 2000, pp. 493–496.

[11] S. Nagashima, K. Ito, T. Aoki, H. Ishii, and K. Kobayashi, "A high-accuracy rotation estimation algorithm based on 1d phase-only correlation," in *International Conference Image Analysis and Recognition*. Springer, 2007, pp. 210–221.

[12] T. Bylander and B. Rosen, "A perceptron-like online algorithm for tracking the median," in *Proceedings of International Conference on Neural Networks (ICNN'97)*, vol. 4. IEEE, 1997, pp. 2219–2224.

[13] E. Mueggler, H. Rebecq, G. Gallego, T. Delbruck, and D. Scaramuzza, "The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and slam," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 142–149, 2017.

[14] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A 240× 180 130 db 3 μs latency global shutter spatiotemporal vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, 2014.

[15] E. R. Ulrich, D. J. Pines, and J. S. Humbert, "From falling to flying: the path to powered flight of a robotic samara nano air vehicle," *Bioinspiration & biomimetics*, vol. 5, no. 4, p. 045009, 2010.

[16] C. H. Tan, D. S. bin Shaiful, E. Tang, G. S. Soh, and S. Foong, "High angular rates estimation using numerical phase-locked loop method," in *2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2020, pp. 1516–1521.

[17] F. Fries, S. K. H. Win, E. Tang, J. E. Low, L. S. T. Win, P. V. y Alvarado, and S. Foong, "Design and implementation of a compact rotational speed and air flow sensor for unmanned aerial vehicles," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10 298–10 307, 2019.

[18] T. S. Lembono, J. E. Low, L. S. T. Win, S. Foong, and U.-X. Tan, "Orientation filter and angular rates estimation in monocopter using accelerometers and magnetometer with the extended kalman filter," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 4537–4543.