This paper has been accepted for publication at the
IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Abu Dhabi, 2020. ©IEEE 2020

Towards UAS Surveillance using Event Cameras

J.R. Martínez-de Dios, A. Gómez Eguíluz, J.P. Rodríguez-Gómez, R. Tapia and A. Ollero

***Abstract*— Aerial robot perception for surveillance and search and rescue in unstructured and complex environments poses challenging problems in which traditional sensors are severely constrained. This paper analyzes the use of event cameras onboard aerial robots for surveillance applications. Event cameras have high temporal resolution and dynamic range, which make them very robust against motion blur and lighting conditions. The paper analyzes the pros and cons of event cameras and presents an event-based processing scheme for target detection and tracking. The scheme is experimentally validated in challenging environments and different lighting conditions.**

***Index Terms*— event camera, surveillance, aerial robots.**

Fig. 1: Experimentation of event-based target surveillance with aerial robots.

I. Introduction

Aerial robot perception in unstructured and complex environments typical of surveillance and search and rescue applications pose very challenging problems. Lighting conditions, vibrations and motion blur or presence of specular reflections due to glass or metallic structures often constrain the feasibility and performance of traditional aerial robot sensors such as cameras or LIDARS. One of the most widely used approach in these cases is to exploit the perception synergies that can be established between different sensors.

In the last years event cameras have received high interest in the robotics and perception communities. Event cameras have high temporal resolution and dynamic range. They are insensitive to motion blur, lighting conditions and have low weight and low power consumption, which make them adequate for small aerial robot perception. Many successful event-based techniques have been proposed in the last years evidencing their capabilities, see e.g. [1].

This paper proposes the use of event cameras for surveillance and search and rescue using aerial robots. It first analyses the advantages and constraints of event cameras and other traditional aerial robot sensors for surveillance. Then, it briefly presents an event-based processing scheme for detecting and tracking moving targets in unstructured environments, for instance to prevent proximity to a dangerous area or in intruder surveillance. Finally, the presented processing scheme is validated experimentally, see Fig. 3.

The rest of the paper is structured as follows. The main related work is briefly presented in Section II. Section III summarizes the pros and cons of the main aerial robot sensors for surveillance applications. Section IV describes an event-processing scheme for aerial robot-based surveillance. The validation experiments are presented in Section V. The conclusions are summarized in Section VI.

II. Related Work

Recent advances in pedestrian detection, face recognition, motion segmentation, target tracking and robot fleet coordination have resulted in the increasing use of robots for surveillance applications. For instance, some works have developed specific techniques to be used for video surveillance such as crowd density estimation and people counting [2], scene background removal [3], or detection different anomalous situations such as fighting, vandalism or road accidents. The capabilities of aerial robots for detecting intrusions have been also explored from the point of view of wireless sensor networks [4] in border surveillance applications. Other works explored the coordination of multiple UAS in order to cover an area in an efficient manner [5].

Detection and tracking of moving objects is a core functionality of autonomous intrusion monitoring in surveillance tasks. A moving object detection method for surveillance tasks using an UAV was proposed in [6]. The authors performed feature extraction and matching from images to further remove the background using Otsu's thresholding and, thus, segment the moving objects in the scene. However, the limitations of traditional cameras at adapting to different lightning conditions restricts the application of existing methods to indoor scenarios, and their motion blur sensitivity hampers their effectiveness at capturing fast motions.

The research interest on event cameras in the robotics and computer vision communities has significantly grown in recent times [1] due to their low latency, high dynamic range, and robustness to motion blur. Some event-based methods

This work was supported by the European Research Council as part of GRIFFIN ERC Advanced Grant 2017, Action 788247 and ARM-EXTEND (DPI2017-8979-R) project funded by the Spanish National R&D Plan.

The authors are with the GRVC Robotics laboratory, University of Seville, Seville 41092, Spain email: {jdedios, ageguiluz, jrodriguezg, raultapia, aollero}@us.es

have been developed for human detection. One of the main approaches for event-based vision rely on building the called *event images*; frames created by accumulating events either by time or a fixed number of samples. The work in [7] used image frames and event images to detect pedestrians fusing the output of two YOLO object detectors from a fixed-position DAVIS camera. Two methods for face detection using event cameras were presented in [8] and compared the performance of using traditional and event cameras at detecting human faces, showing the viability of using only events for face detection. The work in [9] segments the scene by clustering motion-compensated *event images* into objects that fit similar motion models. Further, the work in [10] also relied on *event images* to segment independent moving objects in indoor scenes using a deep learning approach.

Although the above methods provide event-based solutions for surveillance related problems none of them evaluated their performance onboard real robots. The work in [11] used a combination of *event images* and grayscale images (i.e. conventional camera) to train a Convolutional Neural Network that guides a predator non-holonomic Unmanned Ground Vehicle to catch a prey robot. *Event images* of fixed number of events were used in [12] to detect and track a ball using a Hough transform approach and optical flow information in order to guide the gaze of an iCub robot. An extension of their method was presented in [13] to enhance the robustness of the target tracking using a particle filter.

The use of event cameras on-board UAS has recently gained the attention of the aerial robotics community. The work in [14] proposed a method to detect moving objects from *event images* by compensating the global motion of a Micro Aerial Vehicle (MAV). Recently, another motion segmentation method was used in [15] to allow a multirotor UAV avoiding collisions when objects are thrown at it. Their proposed method uses a stereo event camera set-up to segment the moving object by grouping the events in *event images* of 10 ms.

Above methods use event frames and not fully exploit the sequential and asynchronous nature of event cameras. In fact, motion blur cancellation mechanisms for *event images* have been used in some works such as [16]. Contrarily, asynchronous *event-by-event* processing do not suffer from these limitations. Although a number of works have explored event-by-event processing for localization [17], feature detection and tracking [18], and clustering [19], their application on real robots navigating in realistic, complex and unstructured scenarios is still an under-researched area. In particular, asynchronous *event-by-event* processing has not been fully exploited in aerial robotics such as, for instance, surveillance tasks. As computational requirements of processing asynchronous data are high, powerful onboard computers are required, which often cannot be equipped due to the weight limitations of UAVs, i.e. payload constraints.

In our baseline work [20], a first approach to *event-by-event* intrusion monitoring onboard a UAV was presented. The performance and computational cost of the method were evaluated, concluding the full event stream in not always needed to achieve high accuracy. In this work, we propose a surveillance system for detecting intrusions using event cameras onboard aerial robots. Our approach uses *event-by-event* processing reducing the computational cost through event stream packaging control. Additionally, dynamic adaptation to the scene conditions is achieved by using only a percentage of the incoming events when it is required.

III. Event Cameras for Aerial Robot Perception

A wide variety of sensors for aerial robots have been researched [21]. Traditional frame-based cameras have been very commonly used both as navigation sensor and also as application-related sensor. Frame-based cameras are small, lightweight and inexpensive. They provide texture and color and other information of the object visual appearance. Besides, stereo vision systems and RGB-D cameras can retrieve depth information at short distance (mainly indoors in the case of RGB-D cameras). However, cameras are sensitive to illumination conditions and suffer from motion blur, which can severely constrain their use in many cases.

Infrared cameras are frequently used in surveillance and search and rescue applications [22]. They capture the infrared energy in the scene to create frame images. Some cameras provide as output images with only qualitative information of the level of infrared energy. Others integrate software tools that can estimate the temperature of objects and provide images with quantitative temperature measurements, which after camera radiometric calibration can be very accurate (lower that 0.1 Celsius degrees [23]). There are infrared cameras covering the different main spectral windows: near-infrared (0.75–1.4 μm), short-wave-infrared (1.4–3 μm), mid-infrared (3–8 μm), and long-wave-infrared (8–15 μm), each of them with different perception properties. In the last years significant reductions in infrared camera size, weight and cost have motivated their popularization in an increasing variety of civilian applications. Infrared cameras can be very interesting in cases where the targets have different temperature from the background. However, infrared camera detectors usually have low sensitivity, which require using large exposure times, frequently causing motion blur when mounted onboard aerial robots.

LIDAR sensors are also very widely used mainly as navigation sensor [24] [25]. They use the reflections of pulsed laser light on object surfaces to provide distance measurements. 2D/3D LIDAR sensors rotate the pulsed laser light to obtain point clouds that represent the environment geometry. LIDARS are immune to lighting conditions but are very much affected by vibrations and often have moderate or high weights, although the recent solid-state LIDARS are significantly lighter than traditional electro-mechanical LIDAR. However, their data acquisition rate can be insufficient for aerial robots performing high-speed maneuvers.

Range sensors provide distance measurements between the robot and a set of radio beacons deployed in the environment. They have been used for aerial robot navigation, see e.g. [26], and for application-related perception applications, such as object search [27]. Among others, it is possible to estimate

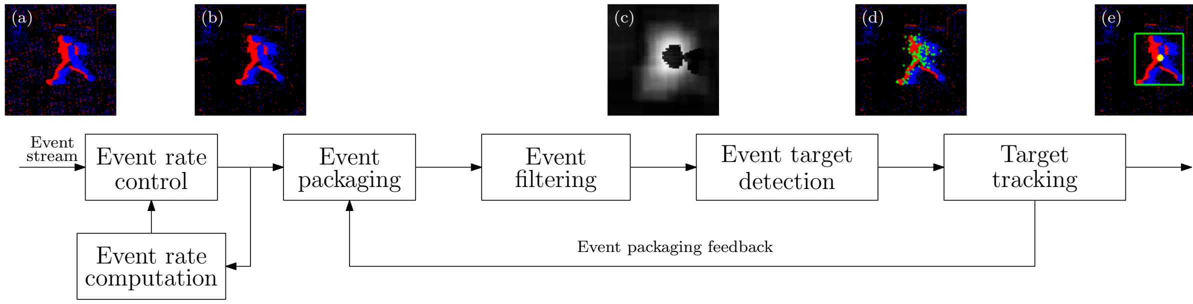

Fig. 2: Event-based target detection and tracking scheme: a) event stream representation, b) event stream representation after random discard procedure, c) APM obtained for event filtering, d) corner tracking, e) event clustering and, f) target tracking.

distance using the Radio Signal Stregth Indicator (RSSI) or Time-of-Flight (ToF) measurements obtained from received radio signals using technologies such as WiFi, Bluetooth or Zigbee. Also, Ultra-WideBand (UWB) provides long measurement range and good accuracy even in indoors [27]. Range measurements are very efficiently computed, naturally prevent the data association problem but, metallic structures significantly can reduce their accuracy, hence they are usually combined with outlier rejection methods. They provide 3D estimations (not orientation) and require having radio beacons in the environment, which is not the case in many cases.

Event cameras are neuromorphic sensors that intend to mimic biological retinas. They capture visual information in the form of events, which represent changes of illumination in the scene. Events are triggered asynchronously with high temporal resolution (in the order of μs). Besides, event cameras have high dynamic range, which provides high insensitivity against illumination conditions, being able to perceive scenes both in total darkness and with high illumination. Event generation require movement. A static event camera in an static scenario would not produce any events, except for those originated by noise, but this is not often an issue when mounted on aerial robots due to its natural vibrations and movement. The low latency, low power consumption and high dynamic range of event cameras motivated the use of these sensors for robotics. Currently, event cameras are being combined with IMU to improve performance in visual odometry, 3D reconstruction, and SLAM, among others [1]. Event-based cameras have good properties for aerial robot perception in highly unstructured scenarios. They can cope with high-speed maneuvers and vibrations due to their very low latency and the absence of motion blur [28]. Besides, their high dynamic range ($>$ 100 dB) provide robustness against lighting conditions changes. However, event-based processing involves a change of paradigm w.r.t. traditional image processing that requires the development of new vision processing algorithms.

Table I summarizes the main features of a typical 3D LIDAR (Velodyne HDL-64E), frame-based camera (DAVIS-346 AVS), infrared camera (FLIR Vue Pro R), UWB (Decawave DWM1000) and event camera (DAVIS-346 DVS) for their perception capabilities in aerial robots.

TABLE I: Summary of sensor details.

Sensor	Dynamic range	Max. rate /bandwidth	Power consumption	Weight
Velodyne VLP-16 Lite	-	5-20Hz	900 mA/9V	$\sim$ 590 g
DAVIS-346 (AVS)	56.7 dB	40 Hz	200 mA/5 V	170 g
FLIR Vue Pro R	-	8.3 Hz	2.1 W V	113 g
Decawave DW1000	-	40 Hz	160 mA/3.6 V	2 g
DAVIS-346 (DVS)	120 dB	Async.	160 mA/5 V	170 g

IV. Event-based target detection and tracking

This section describes an event-processing scheme for aerial robot-based surveillance. The scenarios are assumed static except for the targets or objects of interest. We assume that targets in motion originate nearby corners in the event stream. Hence, targets create groups of events close to corners with globally consistent motion in the scenario.

The developed scheme consists of six main processing modules: *Event rate computation*, *Event rate control*, *Event packaging*, *Event filtering*, *Target detection*, and *Target tracking*. They only use as input the event stream received asynchronously from the event camera. Figure 2 shows results from the event-based processing modules: a) event stream, b) event stream result of controlling the event rate c) APM obtained for event interest filtering, d) corner tracking, e) event clustering and, d) intruder tracking.

A. Event rate control and packaging

Event cameras provide asynchronous measurements in the form of events, which are often fed to the event-based vision algorithms in packages in order to avoid system overflow when the event rate is too high, i.e. too many events are captured in a short period of time. However, this is inefficient when the algorithm is capable of processing the full event stream and reduces the method responsiveness.

Our surveillance system include two mechanisms to allow efficient real-time computation without the risk of overflowing. First, the *Event packaging* mechanism uses a close-loop approach to adapt the size of the event packages by considering the time devoted by the *Event filtering*, *Event target detection* and *Target tracking* modules to process the last packages. Thus, the time encapsulated in each package is synchronized (made equal) to the time required by the algorithm to process the last event packages. Second, *Event rate control* mechanism randomly discards some events when the hardware limitations preclude the algorithm to process all of them. The work in [20] showed that the event stream can be reduced by randomly discarding up to 20% of the events without significantly affecting the performance. The developed modules dynamically adapts the percentage of events that are randomly discarded so that the risk of the algorithm overflowing is reduced.

B. Event filtering

In static event cameras all triggered events correspond to objects of interest. On board aerial robots, triggered events can be caused either by moving objects or by static background with robot motion. The objective of event filtering is to differentiate between both cases.

The adopted approach uses spatio-temporal information to differentiate between both types of events assuming that the regions that correspond to moving objects cause significantly more events than static objects. It uses the *Attention Priority Map* (APM) to capture the regions that trigger more events within a time frame. The events generated by moving objects are assigned with higher value –higher attention priority– in the APM than those created by static objects in the scene.

The APM is defined by matrix $\mathbf{\Omega}$, which size is similar to that of the event camera, to represent the number of events triggered within a region during a dynamic time frame. $\mathbf{\Omega}$ is updated with each event by increasing the values in $\mathbf{\Omega}$ in a window centered at coordinate $\mathbf{x} = (u, v)$ of the event. Similarly, old events are removed from $\mathbf{\Omega}$. An event at coordinate $\mathbf{x} = (u, v)$ is considered of interest (corresponding to a moving object) if $\mathbf{\Omega}_{\boldsymbol{u},\boldsymbol{v}}$ is greater than a threshold $\omega \in [0, 1]$. Otherwise, the event is considered caused by static background.

C. Event-based target detection

This module detects regions with nearby corners and events with high priority. The first module performs corner detection. We adopted *eFast [18] due to its trade-off between accuracy, performance and computational efficiency. The corners detected are separated by polarity and tracked to remove inconsistent and noisy corners. All tracked features are stored in the list $\hat{\mathbf{F}} = [\hat{\mathbf{f}}_1, \dots \hat{\mathbf{f}}_n]$, where $\hat{\mathbf{f}}_i$ is the i-th feature tracked and is defined by a timestamp t of the event that generated the feature and its coordinates (u, v).

The candidate evaluation consists of analyzing the occurrence of previous tracks $\hat{\mathbf{f}}_i$ in the neighborhood of the candidate $\mathbf{f}$. Previously tracked features within the neighbourhood area of a candidate feature are added to the list of feature matches $\mathbf{M}$. If the candidate results in $\mathbf{M} \neq \emptyset$, the oldest track is updated by the candidate $\mathbf{f}$. If $\mathbf{M} = \emptyset$, the candidate is considered a new track and $\mathbf{f}$ is added to $\hat{\mathbf{F}}$. Thus, only features with continuous occurrence are tracked.

Finally, clusters with both events and corners should be created. Events are generally triggered at the contour of the objects. The method finds clusters analyzing the distance of each new event to a random sample of the cluster events. It evaluates the spatio-temporal distance of each new event to the existing clusters. If the event is close to only one cluster, it is assigned to it and the cluster is updated. If the event is close to more than one cluster, the clusters are merged and the resulting cluster is updated. If the event is not close to any existing cluster, a new cluster is created. If a cluster contains less than a certain number of corner tracks, it is filtered out.

D. Target tracking

Intruders create groups of events close to corners with consistent motion. Finally, the obtained clusters represented by their centroid are tracked, see Fig. 2-d) where the intruder is marked with a green square. If a cluster is not updated in consistently, it is filtered out.

V. Experiments

Two sets of experiments were performed. The first one validates the event cameras motion blur and lighting robust performance. The second analyzes the event-based processing presented in Section IV.

A. Motion blur and lighting conditions robustness

We first analyzed the sensitivity of event cameras to motion blur. We installed a DAVIS 346 sensor [29] on a flapping-wing aerial robot (ornithopter) prototype. The objective is to analyze the event camera performance while the ornithopter flapped its wings at different rates. The greater the flapping frequency, the higher the motion amplitude over the fuselage and the event camera. The DAVIS 346 includes a monocular camera, an Inertial Measurement Unit (IMU) and an event camera. It was mounted to provide frontal views and connected through USB to a computer that recorded all the measurements provided: event stream, monocular camera, and IMU. The sensor pointed towards a fixed target such that a white screen with the GRIFFIN logo as the only object on the scene, see Fig. 3. The experiments consisted of flapping at three different frequencies (i.e. 1.6, 2.1 and 2.5 Hz.). Flapping frequency impacted on both cameras. Table II summarizes the results obtained. The number of blurred images highly increased as the flapping frequency increases. The number of events generated highly increased with the flapping frequency, which enables capturing more perceptual information at higher flapping rates.

Regarding robustness to lighting conditions, Fig. 4 shows the overlapping of the frame-based image and the events obtained in two UAS-based moving target detection and tracking experiments performed in the day (left) and in the night in total darkness (right). The event camera detected the moving target in both lighting conditions.

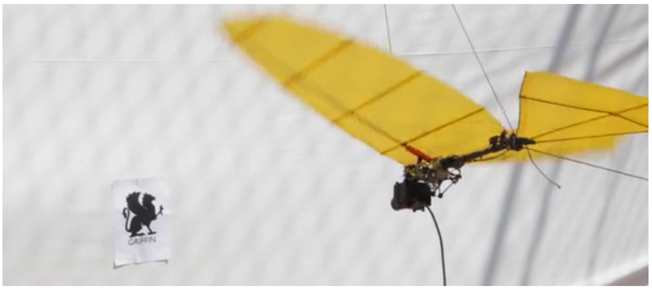

Fig. 3: Set-up of the motion-blur experiment with a DAVIS 346 sensor onboard a flapping-wing ornithopter.

TABLE II: Effect of flapping at different frequencies.

Flapping frequency	**Percentage of blurred images**	**Number of events**
1.6 Hz	15.9 %	994981
2.1 Hz	26.5 %	1832151
2.5 Hz	40.2 %	2716686

B. Moving target detection and tracking

The target detection and tracking scheme proposed in Section IV was validated in sets of experiments in challenging scenarios. The robot used was a DJI Flamewheel F550 Drone shown in Fig. 5 equipped with a DAVIS 346 and an INTEL NUC for online computation and logging. The event-based scheme was implemented in ROS Kinetic.

Different experiments were performed in which a person moved in the environment. The objective is to detect and track the person, for instance to prevent proximity to a dangerous area in case of accident or simply in a intruder detection and tracking problem. Figure 6 shows some results in experiments performed during the day. The target is tracked with a green window. The processing was performed using only the event stream. The frame-based images from the DAVIS APS, which are overlaid with the events, are shown only result visualization. To evaluate the scheme detection success rate and noise rejection, we computed:

$$Accuracy = \frac{TP + FP}{T + N}, \quad (1)$$

$$Precision = \frac{TP}{TP + FP}, \quad (2)$$

$$Recall = \frac{TP}{TP + FN}. \quad (3)$$

where T is $TP + TN$, N is $TN + FN$ and TP, TN, FP, and FN stand for the true positive, true negative, false positive and false negative rates, respectively .

The average results obtained in day experiments were *Accuracy*=0.96, *Precision*=0.98 and *Recall*=0.95, see Table III. Similar sets of experiments were performed with different lighting conditions: a) dark lighting in night experiments and d) dynamic lighting in experiments performed in the night with strong temporal lighting changes caused by moving lights in the scene. Figure 7 shows the results obtained in different sets of experiments. The obtained metrics are shown in Table III. The method performed as expected in all conditions regardless of the lighting conditions. The best performance was obtained in day experiments. In night experiments, the low illumination caused low signal-to-noise ratio slightly reducing performance. Dynamic lighting originate noisy events all around the scene that hampered performance. Despite these effects, the presented results show the robustness of the event-based scheme to challenging lighting conditions and validate its use for surveillance.

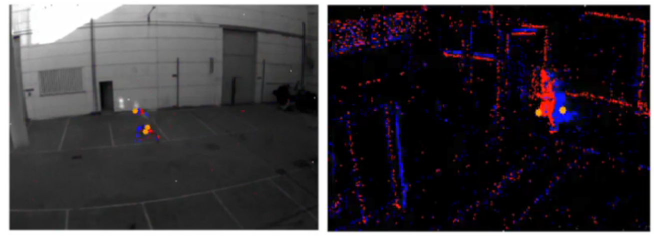

Fig. 4: Frame-based image and event stream obtained in surveillance experiments performed in the day (left) and in the night in total darkness (right).

TABLE III: Performance results.

Experiment Set	**Precision**	**Recall**	**Accuracy**
Day	0.98	0.95	0.96
Night	0.95	0.94	0.95
Dynamic lighting	0.95	0.87	0.90

VI. Conclusions

This paper analyses the use of event cameras for surveillance in challenging outdoor environments. Many of the most widely-employed sensors have constraints due to lighting changes and motion blur. This paper first analyzes their pros and cons and compare them to those of event cameras. Also, an asynchronous event-based processing scheme composed by six main modules –event rate computation, event rate control, event packaging, event filtering, event-based target detection and target tracking– is presented. The processing scheme has been implemented and experimented, validating its robustness against motion blur and lighting conditions.

Fig. 5: Robot used in the shown experiments: a DJI Flamewheel F550 platform with a DAVIS 346 and an INTEL NUC.

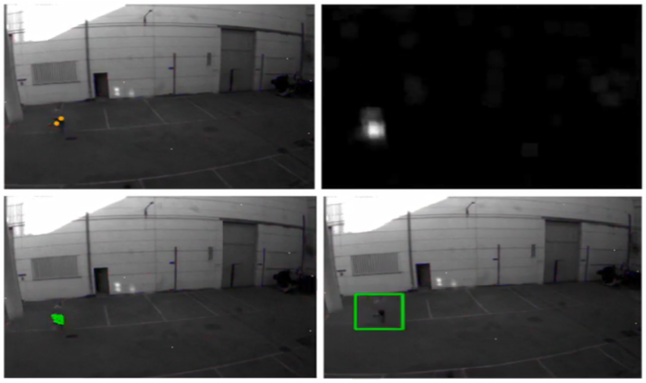

Fig. 6: Results of day experiments: top) events over frame-based images (left) and resulting APM (right); bottom) result of target detection (left) and tracking (right).

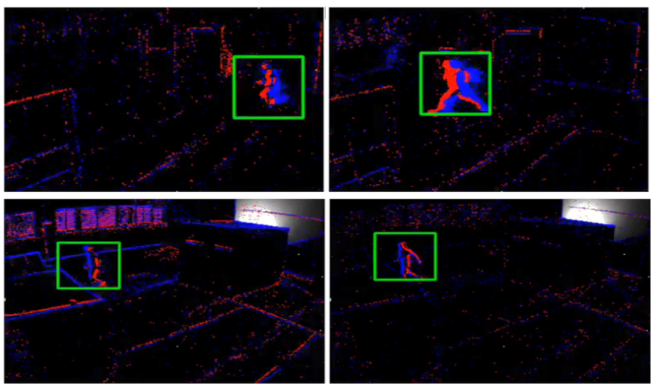

Fig. 7: Detection and tracking results in dark (top) and dynamic (bottom) lighting conditions experiments.

Although event cameras provide very interesting properties, the event stream is not as informative as the measurements provided by other sensors such as cameras or LIDARS. A clear trend is the combined processing (fusion) of the event stream with other measurements such as images and IMU data. The integration of the presented scheme together with frame-based images provided by the DAVIS sensor is object of current research.

References

[1] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, *et al.*, "Event-based vision: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[2] S. A. M. Saleh, S. A. Suandi, and H. Ibrahim, "Recent survey on crowd density estimation and counting for visual surveillance," *Engineering Applications of Artificial Intelligence*, vol. 41, pp. 103–114, 2015.

[3] T. Bouwmans and E. H. Zahzah, "Robust pca via principal component pursuit: A review for a comparative evaluation in video surveillance," *Computer Vision and Image Understanding*, vol. 122, pp. 22–34, 2014.

[4] S. Berrahal, J.-H. Kim, S. Rekhis, N. Boudriga, D. Wilkins, and J. Acevedo, "Border surveillance monitoring using quadcopter uav-aided wireless sensor networks," *Journal of Communications Software and Systems*, vol. 12, no. 1, pp. 67–82, 2016.

[5] N. Nigam, "The multiple unmanned air vehicle persistent surveillance problem: A review," *Machines*, vol. 2, no. 1, pp. 13–72, 2014.

[6] S. A. Cheraghi and U. U. Sheikh, "Moving object detection using image registration for a moving camera platform," in *2012 IEEE Intl. Conf. on Control System, Computing and Engineering*, 2012, pp. 355–359.

[7] Z. Jiang, X. Pengfei, K. Huang, W. Stechele, G. Chen, Z. Bing, and A. Knoll, "Mixed frame-/event-driven fast pedestrian detection," in *IEEE ICRA*, 2019.

[8] S. Barua, Y. Miyatani, and A. Veeraraghavan, "Direct face detection and video reconstruction from event cameras," in *IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 2016, pp. 1–9.

[9] T. Stoffregen and L. Kleeman, "Simultaneous optical flow and segmentation (sofas) using dynamic vision sensor," *arXiv preprint arXiv:1805.12326*, 2018.

[10] A. Mitrokhin, C. Ye, C. Fermuller, Y. Aloimonos, and T. Delbruck, "Ev-imo: Motion segmentation dataset and learning pipeline for event cameras," *arXiv preprint arXiv:1903.07520*, 2019.

[11] D. P. Moeys, F. Corradi, E. Kerr, P. Vance, G. Das, D. Neil, D. Kerr, and T. Delbrück, "Steering a predator robot using a mixed frame/event-driven convolutional neural network," in *Intl. Conf. on Event-based Control, Communication and Signal Processing*, 2016, pp. 1–8.

[12] A. Glover and C. Bartolozzi, "Event-driven ball detection and gaze fixation in clutter," in *IEEE/RSJ IROS*. IEEE, 2016, pp. 2203–2208.

[13] ——, "Robust visual tracking with a freely-moving event camera," in *IEEE/RSJ IROS*. IEEE, 2017, pp. 3769–3776.

[14] A. Mitrokhin, C. Fermüller, C. Parameshwara, and Y. Aloimonos, "Event-based moving object detection and tracking," in *IEEE/RSJ IROS*, 2018, pp. 1–9.

[15] D. Falanga, K. Kleber, and D. Scaramuzza, "Dynamic obstacle avoidance for quadrotors with event cameras," *Science Robotics*, vol. 5, no. 40, 2020.

[16] N. J. Sanket, C. M. Parameshwara, C. D. Singh, A. V. Kuruttukulam, C. Fermuller, D. Scaramuzza, and Y. Aloimonos, "Evdodge: Embodied ai for high-speed dodging on a quadrotor using event cameras," 2019.

[17] A. Z. Zhu, N. Atanasov, and K. Daniilidis, "Event-based visual inertial odometry," in *2017 IEEE CVPR*. IEEE, 2017, pp. 5816–5824.

[18] I. Alzugaray and M. Chli, "Asynchronous corner detection and tracking for event cameras in real time," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3177–3184, 2018.

[19] F. Barranco, C. Fermuller, and E. Ros, "Real-time clustering and multi-target tracking using event-based sensors," in *IEEE/RSJ IROS*, 2018, pp. 5764–5769.

[20] J. P. Rodríguez-Gómez, A. Gómez Eguíluz, J. R. Martinez De-Dios, and A. Ollero, "Asynchronous event-based clustering and tracking for intrusion monitoring in uas," in *IEEE ICRA*, 2020.

[21] A. Ollero and L. Merino, "Control and perception techniques for aerial robotics," *Annual reviews in Control*, vol. 28, no. 2, pp. 167–178, 2004.

[22] M. Vollmer and K. Möllmann, *Infrared Thermal Imaging: Fundamentals, Research and Applications*. Wiley Online Library, 2017.

[23] R. Hudson, *Infrared System Engineering*, ser. Wiley series in pure and applied optics. Wiley-Interscience, 1969.

[24] J. L. Paneque, J. Martínez-de Dios, and A. Ollero, "Multi-sensor 6-DoF localization for aerial robots in complex GNSS-denied environments," in *IEEE/RSJ IROS*, 2019.

[25] B. Schwarz, "Lidar: Mapping the world in 3d," *Nature Photonics*, vol. 4, no. 7, p. 429, 2010.

[26] A. Torres-González, J. R. Martinez-de Dios, and A. Ollero, "Range-only slam for robot-sensor network cooperation," *Autonomous Robots*, vol. 42, 07 2017.

[27] J. R. Martinez-de Dios et al., "GRVC-CATEC: aerial robot co-worker in plant servicing (ARCOW)," in *Bringing Innovative Robotic Technologies from Research Labs to Industrial End-users*, ser. Springer Tracts in Advanced Robotics, F. Caccavale, C. Ott, B. Winkler, and Z. Taylor, Eds. Springer, 2020, vol. 136, pp. 211–242.

[28] A. Gómez Eguíluz, J. P. Rodríguez-Gómez, J. R. Martinez De-Dios, and A. Ollero, "Asynchronous event-based line tracking for time-to-contact manuevers in uas," in *IEEE/RSJ IROS*, 2020.

[29] 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 J. of Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, 2014.