An Optical Spatial Localization System for Tracking Unmanned Aerial Vehicles Using a Single Dynamic Vision Sensor

Hunter Stuckey, Amer Al-Radaideh, Leonardo Escamilla, Liang Sun, Luis Garcia Carrillo, and Wei Tang

***Abstract*—This paper reports a novel optical localization method, including both the hardware design and algorithm design, to track mobile Unmanned Aerial Vehicles (UAVs). The method relies on a circle-shaped blinking LED marker installed on the UAV and uses a single Dynamic Vision Sensing (DVS) camera to sense the temporal difference of the video streams. A temporal-filtering algorithm processes the video stream and detects the target marker by filtering out the background image. The triangulation-based spatial localization algorithm captures the trace of the target with the help of the prior knowledge of the physical size of the marker. The proposed system was evaluated in flight tests and compared with ground truth data provided by a motion capture system. The proposed system provides a simple and accurate localization solution for UAV tracking with a low computing overhead.**

***Index Terms*—Aerial Systems: Perception and Autonomy, Localization, Visual Tracking, Event-based Camera.**

I. Introduction

Spatial localization is critical in applications involving Unmanned Aerial Vehicles (UAVs). As UAVs can reach heights and places where humans could not normally go without the same safety concerns or cost [1], UAV technologies have been advancing at a rapid pace with many important applications, such as aerial photography, building safety inspection, precision agriculture, storm tracking, and weather surveillance [2]–[4]. In particular, UAVs have been playing important roles in the COVID-19 pandemic due to the safety and medical concerns of quarantine [5]. In such applications, identifying the spatial location of UAVs is one of the inevitable tasks for precise track and control.

Spatial localization is defined as the capability to obtain the location information of a UAV relative to a reference point. Such location information could be calculated either using its onboard sensors to estimate its own location (e.g., cameras, RF transceivers, and GPS) or through off-board sensing technologies installed in surrounding infrastructure, i.e., a motion capture system. These methods are also referred to as self-localization and cooperative localization [6]. UAV localization typically requires the location accuracy of the UAV in the level of centimeters and the distance between the UAV and the reference point in the level of tens of meters [7]. Novel localization methods are anticipated to be able to replace or compensate for the shortcomings of the conventional UAV localization methods that use satellite or Wi-Fi signals, especially in application scenarios where those signals are unavailable.

Spatial localization methods can be categorized according to signal sources and types [8]. Based on the signal source, there are active and passive methods. In an active method, the signal source is generated on a UAV and tracked by some sensing technologies attached to a base station on the ground [6]. For example, the UAV can take images or videos of the environment to calculate its own location, or it can transmit the data to the base station to let it calculate the location of the UAV. In a passive method, none of the localization tasks are performed by the UAV while the base station is in charge of tracking the location of the UAV using various types of signals, such as radio/RADAR signals [9], acoustic/sonar signals [10], [11], optical/LiDAR signals [12], and a taut tether [13], [14]. These methods have different advantages and limitations in terms of their performance in UAV spatial localization.

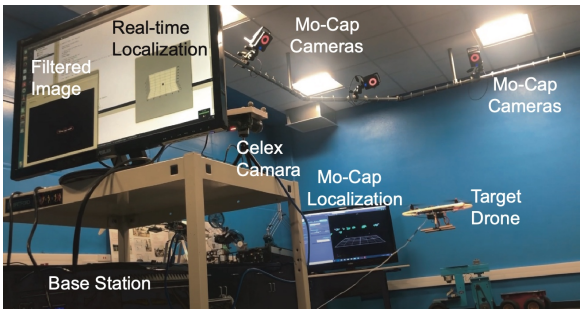

Fig. 1: Experimental environment for real-time 3D localization testing, where a flying quadcopter was being tracked by both a motion-capture system and a single dynamic vision sensor running the proposed temporal 3D localization algorithm.

The main design specifications in UAV-localization methods include accuracy, detection range, computing overhead, power consumption, reliability, latency for real-time applications, the dependency of landmarks, and the dependency of GPS or Wi-Fi signals. For instance, the main consideration in UAV delivery applications is accuracy, while in aerial photography the interest may shift to computing overhead, which is directly related to battery life. The design specification determines the system implementation methods. For example, if the data-collection devices and algorithms for localization are implemented on a UAV, the communication power can be saved but the sensing and data processing power increases. On the other hand, the implementation of a localization system on a base station allows for more computational overhead and alleviates extra weight for onboard sensing and processing, but limits the range of tracking.

The primary design challenges of UAV localization methods are limited power consumption and resolution. For instance, the resolution of regular GPS- and Wi-Fi-based localization methods is at the level of meters, which would not be accurate enough for many applications. Moreover, there are scenarios where GPS or Wi-Fi signals become unavailable. Image-based localization methods face challenges from limited power in either communication or processing. These methods usually have high computing overhead, which can limit the system battery lifetime. For instance, in the case that the computing is performed on a base station, transmitting raw videos from the UAV to the base station would be expensive in communication power and limits the detection range. An image-processing algorithm implemented on the UAV usually consumes considerable power and shortens the flight time [15]. Some image-based methods rely on specific ground landmarks, which limit the detection range and could not be applied in new areas.

To address the aforementioned challenges, this paper presents a novel single-camera-based optical spatial localization system that consists of a circle-shaped blinking LED ring as the marker on a UAV, an event-based dynamic vision sensing (DVS) camera for image sensing including video capturing and pre-processing, a band-pass optical filter for object detection, and a 3D localization algorithm running on a base station. The circle-shaped LED blinking at a predetermined frequency is mounted to the perimeter of the UAV as the marker. The image stream of blinking events sensed from the DVS camera is loaded into a buffer in the base station processor. Then a band-pass filter identifies the pixels with the target blinking frequency in the image streams so that the circle-shaped marker from the image can be detected and filtered out from the background. The 3D location is finally calculated using the known actual physical diameter of the circle-shaped marker and the Field of View (FOV) of the image sensor. The proposed localization system has been validated in real-time flight tests using a motion capture system, as shown in Fig. 1.

The main contribution is the development of an accurate real-time 3D UAV localization system without environmental knowledge except for the physical parameters of the LED marker. The system does not spend onboard communication or computing power of the UAV. All of the sensing and processing tasks are performed at the base station. The event-based DVS camera provides a unique approach to identify the blinking LED marker from the background of the image stream so that the UAV can be detected to achieve high-speed localization. This method is based on an assumption that the background does not contain objects that have the same blinking pattern as the LED marker. This method has a higher fidelity compared to other methods that use color or spectrum of the marker to identify the target, which is more susceptible to background noise. The detected blinking LED marker facilitates the tracking of the position of the UAV relative to the camera, which is described by the relative distance and the azimuth and elevation angles (i.e., the coordinates of the UAV in the spherical frame centered at the camera). It can also help to estimate the attitude (roll and pitch) angles of the UAV. The proposed localization system does not demand extra power from the UAV except for lighting the blinking LED marker. Thanks to the DVS output features, the localization algorithm running at the base station also has a low computing overhead compared to existing image-based localization methods. This makes it suitable for implementation on a mobile station in the future.

This paper is organized as follows. Section II introduces a summary of related work. In Section III, the hardware system is presented including the circle-shaped blinking LED marker, the event-based DVS camera, and the base station multi-thread operation. Section IV describes the signal-processing approach and the algorithms applied for object detection and spatial localization. Section V presents the experimental setup, results, and performance analysis. Section VI discusses the advantages, limitations, and future scope of the proposed work. Finally, Section VII concludes the paper.

II. Related Work

The proposed system is related to optical UAV localization methods, high-speed event-based dynamic vision sensors, and single-camera image processing algorithms. The adoption of UAVs in the workforce has greatly increased the number of localization solutions being created. The primary considerations of choosing localization methods depend on the specific application regarding the localization accuracy, detection range, and computing overhead. Several UAV localization methods have recently been proposed based on tracking the optical markers on a UAV using cameras as the reference point with image sensing and processing algorithms. These methods apply the prior knowledge of the markers, such as shape, color, and size, to obtain the information of a UAV's attitude and its location relative to the camera. Using a constellation of markers, pose-estimation can be achieved with high accuracy. For example, Walter *et al.* [16] implemented a leader-follower system using ultraviolet (UV) markers with a hexagonal orientation for recognition and tracking of the leader UAV by the follower UAV. Teixeira *et al.* [17] implemented a UAV localization system with Infrared (IR) LED markers that were oriented in a pentagonal fashion around the UAV. The small size of LED markers makes them easily attachable to UAVs, with markers emitting in certain sections of the UV/IR spectrum being usable for outdoor testing. However, the marker constellation increases the computing overhead of image processing. It is also challenging to correctly identify the markers if the background environment also contains the same UV/IR light source.

In our prior work [18] [19], we proposed alternative solutions of applying circle-shaped markers so that the major axis of the marker's ellipse shape in the image can always be available in the image for measuring the UAV distance while avoiding the problem of identifying the constellation. We also proposed and validated localization methods using blinking LED markers with temporal filters in the image processing algorithm to remove the background noise pixels. However,

due to the lack of high-speed DVS cameras, these methods were not implemented on a UAV for flight testing.

High-speed event-based image sensing is crucial to perform real-time UAV localization and tracking. In event-based sensing, the pixel intensity change in an image is directly linked to calculating the object location [20]. Using temporal-difference images to detect moving objects while filtering out the background can be used as an effective means of tracking [21]. Tracking objects by their moving edges using a fast event-camera with compensation for motion blur has been implemented in [22], where an accumulation of events and a motion-compensation model are used to greatly improve performance. In this paper, we applied these methods for enabling UAV tracking using an event-based dynamic vision sensor. The event-based image sensor provides high-speed event sensing, which is critical to accurately identify and track the location and attitude of a UAV.

Single-camera or monocular-vision-based localization systems [23]–[26] have become popular in recently reported literature compared to multi-camera systems [27]. For instance, [28] presents an example of building a 3D model of the environment and then using convolutional neural networks to accurately map and navigate a UAV in an arena. The authors in [29] proposed a method using image masks of an object with the monocular vision algorithm running on graphic processing units (GPU) in conjunction with a high frame-rate color camera for pose-estimation of robots. Similarly, [30] applies monocular vision along with Radio Frequency Identification (RFID) and antennas to obtain the 3D position from a 2D image using triangulated methods. These methods achieve reliable localization using a single camera combined with other types of sensors or prior-knowledge of the target or the environment. The common challenges in the above-mentioned monocular vision methods come from the computing overhead of image processing algorithms for real-time tracking, power cost for communication and processing, and the environmental noise that affects the accuracy of localizing the color-based markers.

Considering the advantages and shortcomings of the above-mentioned methods and systems, in this work we designed a DVS-based single-camera system with blinking markers and temporal filter algorithms to achieve UAV localization with low computing overhead. The DVS directly computes the temporal difference image in the hardware to save post-processing power and time. The following sections describe the implementation, verification methods, and results.

III. Hardware Design and Implementation

The system contains the blinking marker, the DVS camera, and a base station. In this system, UAV localization is achieved based on a minimal modification of a UAV by adding a circle-shaped blinking LED as the marker that is tracked by a high-speed event-based DVS camera. The base station runs an image processing algorithm to remove environment background noise, detect the marker, and calculate the relative location of the UAV in real-time. The primary goal of designing the hardware system is to meet the target latency, detection range, and localization accuracy. Since the UAV can move fast, the timing resolution is set at 0.1 seconds, with a target range of 30 meters and a resolution in tens of centimeters.

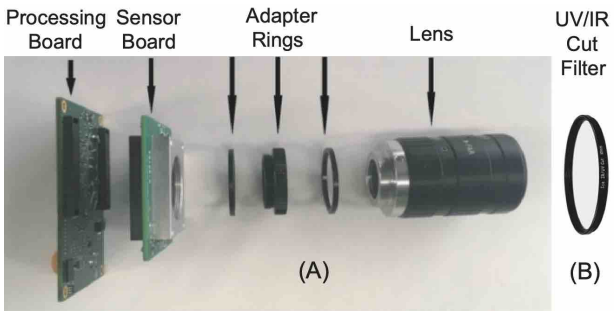

Fig. 2: (A) Picture of the Celex4 event camera, retrieved from the Celepixel SDK manual [31], (B) Picture of 37 millimeter UV/IR cut filter placed in front of the event-camera.

A. Event Camera

The event-based camera applied in the system is the Celex4 dynamic vision sensor [31] from Hillhouse Technology Ltd as shown in Fig. 2 (A). Unlike a conventional rolling-shutter-based image sensor, the DVS camera is an address-even-representation (AER) image sensor that generates a stream of pixel events. A pixel event is defined as the temporal difference of the same pixel in the image stream between two consecutive frames. An event is generated by subtracting the current pixel value from the previous pixel value. The returned value reflects either a zero if the difference is below a predefined threshold, or a one, otherwise. The event image shows if a large enough change has taken place in a corresponding pixel. This is performed using internal pixel-level comparison circuits between frames, which also compare the result with the threshold. Therefore, the sensor can achieve an event rate of up to 500 frames/second. This feature saves the post-processing computing overhead for detection and localization. Since we applied a white LED marker ring, an optical UV/IR cut filter is placed in front of the camera lens to eliminate the infrared energy that may create noise, as shown in Fig. 2 (B). Based on pixel events, temporal filters and optical localization algorithms are implemented on the base station. The Celex4 camera is connected to the base station through a Spartan-6 FPGA board of XEM6301 provided by Opal Kelly Inc. and a Universal Serial Bus (USB) 3.0 cable. In our experimental setup, the Celex camera is set in the event mode with the event frame time of the camera set as 4 ms.

B. Circle-shaped Blinking LED Marker

The optical marker on the UAV composes a blinking LED ring with a diameter of 2 feet, as illustrated in Fig. 3. The LED ring is mounted on the outside circle of the UAV. The optical marker is expected to fulfill two requirements. First, it should be easily distinguished from the background image

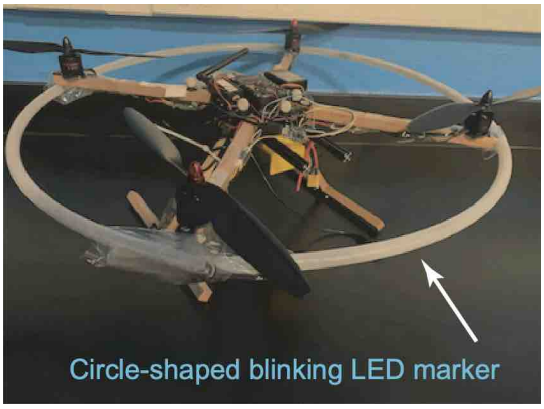

Fig. 3: Picture of UAV with LED marker ring.

for object identification with minimum interference from the noise source. In our design, this is achieved by the predefined blinking pattern of the LED. Second, the marker should help the localization algorithm to obtain the 3D coordinates of the UAV using the known physical dimension of the marker. In this design, the system uses the known diameter of the LED ring and the FOV of the DVS camera. The LED ring is powered and controlled by an Arduino Nano Board that can provide a current of 100 mA. The Arduino Nano board is programmed so that the LED is blinking at a fixed frequency of 40 Hz. It is difficult for human eyes to sense the blinking so the ring looks like a solid white light. The LED ring is powered by a portable USB battery charger.

The main advantages of using a circle-shaped blinking LED marker include that (1) the event-based sensor can easily identify the ring by removing the background from the image by applying a counter-based temporal band-pass filter, and (2) the diameter of the ring can be easily measured from the image in terms of the number of pixels regardless of the attitude of the UAV. The system uses the diameter of the ring in the image to calculate the distance of the UAV. Compared with other shapes of markers, such as a hexagon [16] or a pentagon [17] constellation, the center and diameter of the ring are easier to be obtained from the image.

C. Base Station

The base station collects the image data from the DVS camera, runs the algorithm for object identification and 3D localization, and plots the real-time trace of the UAV in a graphical user interface (GUI). The base station is a desktop running Ubuntu 16 LTS. The base station system application including the GUI is written in C++. Two threads were programmed to implement the algorithms. One thread is dedicated to perform a temporal band-pass filter for the blinking LED marker identification and to display the image showing the LED ring without a background. The other thread runs the post-processing and optical-localization algorithm, as well as plots the 3D trace of the UAV on the GUI showing the 3D coordinates, as illustrated in Fig. 1. The measured base station processing latency, including both identification and localization, is under 100 ms so that the timing resolution of the system can achieve 10 measurements/second in real-time.

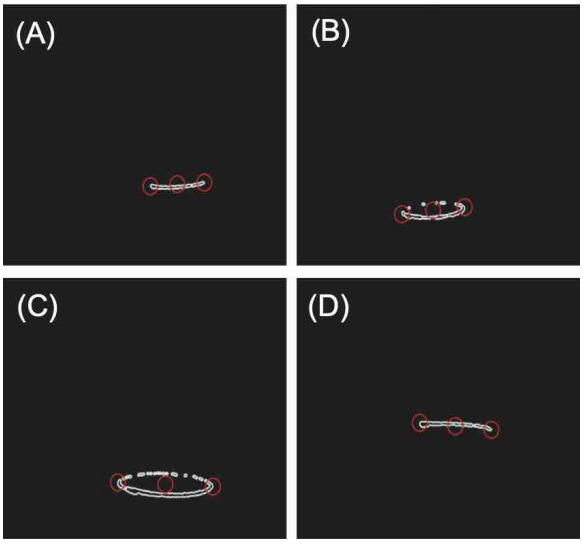

Fig. 4: Binary images obtained from the temporal band-pass filter with the marked left, right, and center points of the target. (A) The UAV is about level with the event camera. (B,C) The UAV is below the event camera and the ring of LEDs is visible. (D) The UAV is at a slightly higher altitude above the camera.

IV. Algorithm Design and Implementation

The objective of the signal-processing algorithms is to track the markers attached to the UAV. In this section, a detailed explanation of a temporal band-pass-filter-based image-processing algorithm and a triangulation-based optical spatial localization algorithm is presented. The temporal band-pass filter, elaborated in the first subsection, detects the object marker from the image and removes the background and noise from the image. The output of the temporal band-pass filter is forwarded to the spatial localization algorithm, which is explained in the second subsection. The spatial localization algorithm calculates the 3D coordinates of the UAV relative to the camera. The two algorithms are both implemented on the base station with separated threads and work in conjunction with each other to create an accurate localization method.

A. Temporal Band-Pass Filter

The temporal band-pass filter takes the output stream of events from the event camera and processes the stream into an image output that only contains the target markers without the image background. The temporal band-pass filter uses the prior knowledge of the blinking rate of markers to identify the UAV. The output of the temporal band-pass filter is a binary image where the outline of the markers (white color) presents as the foreground and the rest of the image (black color) represents the environment background that has been removed, as shown in Fig. 4.

The operation of the temporal band-pass filter begins from the event camera input. The system requires the camera to

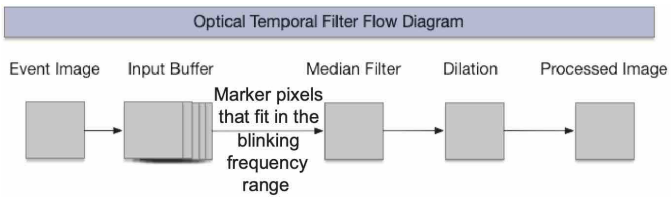

Fig. 5: Flow diagram of event-image processing.

be fast enough to take images with a much higher frequency than the blinking rate of the target. This is because the sampling rate should be at least two times higher than the blinking rate to guarantee that each on and off phase of the blinking pixel is captured. Then the events are generated by subtracting consecutive frames. Therefore, the sampling rate of the camera should be at least four times higher than the blinking rate of the marker. Finally, in order to track the movement of the target, the sampling rate should be even higher. In our system, the camera sampling rate is 250 frames/second for the 40 Hz blinking rate of the marker. Therefore, the algorithm should be able to handle a large amount of data of event images from the high-speed event camera. The incoming event images during a certain time period are stored in a fixed-length First-In-First-Out (FIFO) buffer as an event sampling window. The window size is the number of frames of the event images stored in the buffer. All the event images in the buffer are added together pixel by pixel to generate a combined image to calculate the event frequency of each pixel. The event frequency is compared to an upper threshold and a lower threshold values. Only pixels with an event frequency between the two thresholds are selected. By doing so only the pixels blinking at the desired frequency are identified and the background image is removed. The event sampling window pushes out the old frames when the buffer is full.

This temporal band-pass filter adopts a sliding-window operation that uses a window shifting through time and thresholding the number of events for each pixel in the current window. The decision of whether a pixel in the image belongs to the marker or the background image depends on the frequency thresholds. The event sampling window determines the number of times that a pixel has changed during the given window of time. For example, if the sampling rate is N frames/second and the window size is W frames, a pixel in the combined image has n events, we can estimate that the event rate is $(N \cdot n)/W$ per second and the blinking rate of the pixel is $(N \cdot n)/(2W)$ Hz. Since the camera sampling and the marker blinking are not synchronized, a fixed event sampling window is not able to precisely determine the blinking rate of a pixel. Therefore, two frequency thresholds should be applied instead of using only one frequency to decide if the pixel is blinking at the desired frequency. The temporal band-pass filter eliminates the pixels with an event rate outside the window between the two thresholds. Only the pixels with event rates between the upper and lower threshold are passed and identified as pixels of the markers. A median filter and a dilation filter are also applied to remove pixel noise. Fig. 5 illustrates a flow diagram of the processing steps of the temporal band-pass filter.

The design parameters of the temporal band-pass filter include the size of the event window buffer and the threshold values, which are tuned for the blinking rate of the target markers and the latency of the temporal band-pass filter. For example, a larger size of the event window buffer implies a higher resolution of detecting event rate but a higher latency to process the image, which may cause a problem to track a fast-moving object. On the other side, if the selected event window size is too small, the lower timing resolution may cause unwanted pixels in the image and create a problem for the optical localization algorithm. Another consideration is the blinking rate of the target. A higher blinking rate creates more events and helps track the movement of the target while it may require a higher sampling frequency of the camera and more computing power. In our implementation, the blinking rate of the LED marker is set at 40 Hz. The output of the temporal band-pass filter is forwarded to the spatial localization algorithm through a data buffer to avoid data access conflict. A multi-threading process allows for higher usage of the computing resource for both the temporal band-pass filter and the spatial localization algorithm to run concurrently.

B. Spatial Localization

After the event stream is processed by the temporal band-pass filter, the 3D location of the UAV is calculated relative to the location of the image sensor. Since the circle-shaped marker on the UAV is captured as an ellipse in the filtered image, the diameter of the circle can be measured as the major-axis of the ellipse in the image, regardless of the UAV's position. The circle-shaped marker allows for the UAV's position to be estimated from various angles/altitudes with high accuracy and low computing overhead.

The location of the UAV is calculated from the known parameters of the system and the measured data during tracking, as illustrated in the geometry representation of the spatial localization system shown in Fig. 6. The x-axis and y-axis are the horizontal and vertical directions while the z-axis is normal to the x-y plane. The known parameters include the diameter, D, of the LED ring marker, the horizontal field of view (HFOV), Φ, and the vertical field of view (VFOV), Θ, of the camera. We assume that the resolution of the camera is L pixels (horizontal) by H pixels (vertical). The measured data are the most left (X_L, Y_L) and most right point (X_R, Y_R) of the ellipse in the image, with respect to the center of the image. The physical location of the UAV is then obtained. The physical location of the target refers to the center of the LED ring marker relative to the camera. The physical location data include the radial distance, ρ, the azimuth angle, ϕ, and the elevation angle, θ.

The geometry localization method has been proposed and implemented in [18]. The center of the UAV in the image, (X_O, Y_O), can be calculated as

$$X_O = (X_L + X_R)/2 \tag{1}$$

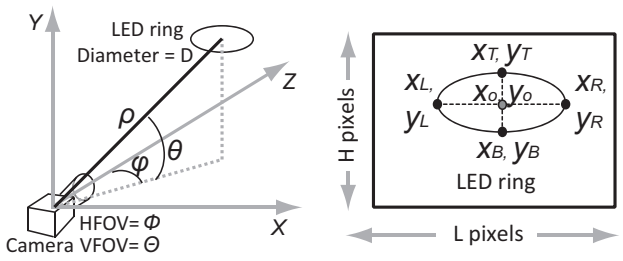

Fig. 6: Geometry for localization algorithm (left), Measurement setup of the target ring (right), retrieved from [18].

$$Y_O = (Y_L + Y_R)/2 \tag{2}$$

Then the tangent value of azimuth angle ϕ and elevation angle θ are obtained as

$$\tan\phi = 2 \cdot X_O \cdot \frac{\tan(\Phi/2)}{L} \tag{3}$$

$$\tan\theta = 2 \cdot X_O \cdot \frac{\tan(\Theta/2)}{H} \tag{4}$$

Finally, the radial distance ρ, which is the distance between the centroid of the target ring and the lens of the camera is obtained as

$$\rho^2 = \left(\frac{D}{2} \cdot \frac{X_O}{|X_O - X_R|}\right)^2 \cdot \left(1 + \frac{1}{\tan^2\phi}\right) \tag{5}$$

In the case that the centroid of the target is located at the middle of the image, X_O is zero, ρ is calculated as

$$\rho = \frac{D}{2 \cdot |X_R|} \cdot \frac{L}{2 \cdot \tan(\frac{\Phi}{2})} \tag{6}$$

Note that $\tan\phi$ in (5) can be obtained from (3). Since Φ, Θ, L, and H are constant parameters, the above equations can be calculated using multiplication without calculation of trigonometric functions, which saves the computing overhead. In the system implementation, the output of the spatial localization algorithm is sent to a buffer that is accessed by a mutual exclusion object (MUTEX) to prevent the conflict of data being accessed at the same time, similar to the data being sent to the spatial localization algorithm by the temporal band-pass filter.

The spatial localization implementation includes a post-processing filtering mechanism, in which the previous location of the target is used to deal with flickering noise that could disrupt the location tracking of the target. For example, if the pixel passing through the band-pass filter is too far away from the previous location of the target, which is determined using a preset threshold, the system would determine that pixel as a noise pixel, and ignore it as a part of the target. The system effectively sets the pixel value to zero in the new image.

This localization algorithm is designed based on the following assumptions. First, the target circle-shaped marker is assumed to be fully contained within the FOV of the camera, as shown in Fig. 6. Second, since the image sensor has a finite resolution, it is assumed that UAV is identifiable in the images captured by the camera. In other words, the target should not be too far away from the camera, otherwise if the number of the pixels of the ellipse's major-axis in the image is small, the accuracy of localization decreases. Increasing the detection range requires a higher resolution of the camera but may increase the system processing time. In summary, the system design trade-offs are among the resolution of the camera, the detection range, and the relative error, and the system processing time.

V. Experimental Setup and Results

The system performance is evaluated by comparing the tracking result from the proposed single DVS system and the benchmark result from a VICON Mo-Cap system that has 10 infrared cameras tracking reflective markers in a 6 m $\times$ 8 m flight-testing arena, as shown in Fig. 1. An optical band-pass UV/IR cut filter was placed on the front of the Celex event-camera to block the strong infrared energy from the VICON system. During the experiments, the IR reflective markers were attached to both the UAV and the Celex Camera. The Mo-Cap system then tracks the reflective markers to calculate the 3D position of the UAV relative to the Celex Camera. The experimental environment is presented in Fig. 1.

The relative error of the proposed localization method is calculated by subtracting the localization results of the single-camera system from the value of the benchmark Mo-Cap system, then divided by the benchmark value. The two relative error metrics calculated are the error over time and error over distance. The relative error for each axis and over radial distance are also evaluated. Three experimental evaluations were carried out to assess the performance of the optical localization with temporal band-pass filtering. The Celex event-camera was mounted on a tripod above the ground, so any Y-axis measurements showing a negative value means that the target is located below the camera. The average relative error for each axis was below 6% for each trial, with the highest error being on the X-axis, which is the horizontal direction that the target has the most movement.

In Fig. 7, the proposed single-camera temporal filtering localization 3D traces are plotted in correspondence with the Mo-Cap traces. The 3D traces from the proposed method align with the benchmark traces retrieved from the Mo-Cap system. The proposed system can capture the UAV trajectory and accurately identify the 3D location in real-time (100 ms latency) relative to the camera's position. Fig. 8 presents the errors over time for the 3 trials. The error appears to be relatively larger when the UAV was making a sharp turn or at a higher velocity. This measurement depends on both the X-axis and Y-axis edge locations of the markers. The scatter plots in Fig. 9 show the relative error over distance for each of the three trials. The distance between the UAV and the camera (ρ) in three trials ranged between 2 and 5.5 m, which is limited by the space size that the Mo-Cap system covers. The maximum range of the localization algorithm is over 30 m, which has been tested in [18], [19].

VI. Discussion

The experimental results show that the proposed localization algorithm and system can be applied in UAV localiza-

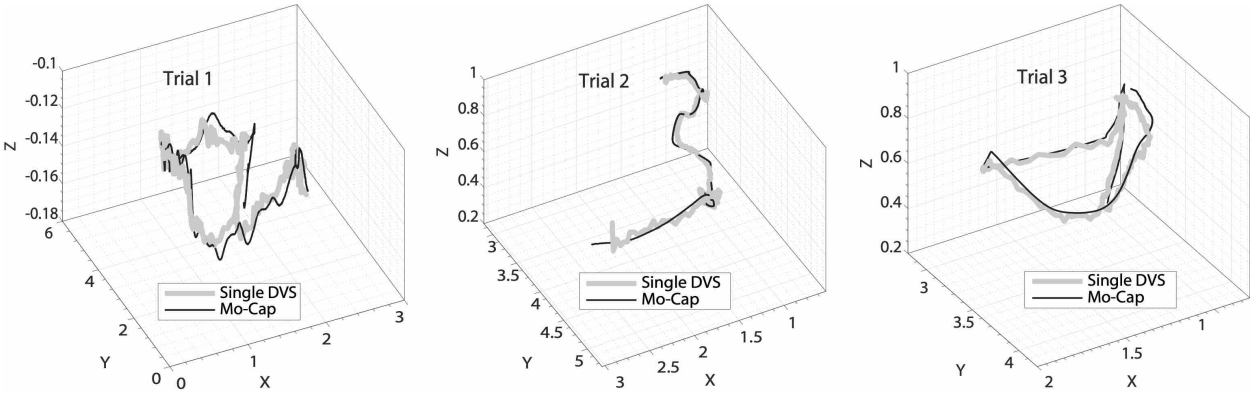

Fig. 7: 3D traces of three trials tracked by the single DVS system and the benchmark Mo-Cap system. The units of measurements are in meters.

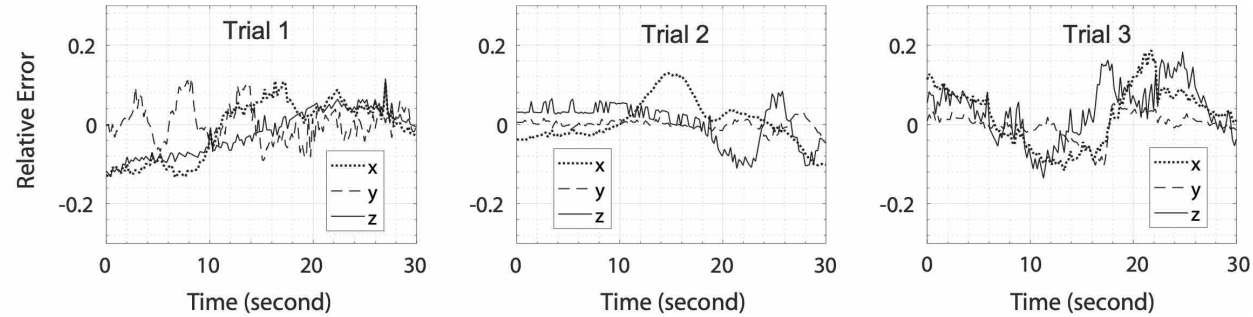

Fig. 8: The relative error over time with respect to the X,Y, and Z axes for each of the three trials.

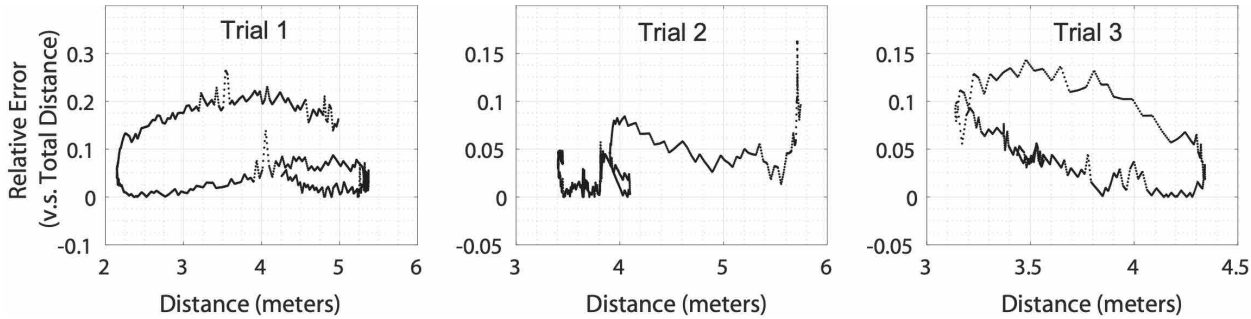

Fig. 9: The relative error over distance for each of the three trials.

tion. The main advantage of the proposed method is that it only uses a single camera with low computing overhead for localization, thanks to the circle-shaped blinking marker and the temporal image filter. Compared to other similar works [16], [17], our proposed method avoids potential noise interference from other infrared sources in the background image. This is because the proposed system applies a temporal image filter to match the blinking rate of the LED marker. Since the temporal filter is targeting at a specific blinking frequency, in a natural environment, it would be a much lower probability to have a noise source with the same blinking rate of the LED marker, compared to the chance of having a noise source with a similar spectrum of the color-based marker such as UV/IR/RGB-based markers. Another advantage of the system is that the algorithm of localizing the circle-shaped marker demands lower computing overhead than a marker constellation. A comparison of the proposed method with other recently published methods is summarized in Table I. Compared to other similar technologies, the proposed blinking marker with the DVS camera achieves a higher detection range with lower relative errors. It also brings the advantage of low computing overhead for image processing thanks to the DVS built-in pixel-level comparison. The primary limitations of the proposed system include the finite resolution of the image sensor and the system processing time. The resolution of the image sensor determines the spatial localization error and the detection range. The system processing time limits the maximum trackable speed of the UAV, which could be improved with event-based image processing algorithms.

TABLE I: Comparisons of monocular-based localization methods tracking a single target.

Method	This Work	[17]	[29]	[30]
Application	UAV	UAV	Robot	Robot
Range	30m	15m	1m	2.3 m
Sensor Type	Dynamic Vision Sensor	2xIMU + PF-MPE Camera	High-speed RGB-D Cameras	Camera+ Antenna Array
Number of Sensors	1	3	2	2
Marker	Blinking	RGB	RGB	RGB
Processing Algorithm	Temporal Band-pass + Optical Localization	Particle Filter with Pose Estimation PFMPE	Pixel Wise Posterior 3D (PWP3D)	Coarse-to-Fine Tag Position
Processing Time/Rate	100mS (10Hz)	20 Hz	400 Hz	0.629 Second
Range	30m	15m	N/A	10m
Average Error	2-6 cm (Range 6 m)	2-15 cm	5.45 mm	8 cm

VII. Conclusion and Future Scope

We presented a novel method for UAV localization using a single dynamic vision camera with a temporal band-pass filter and a circle-shaped blinking LED marker. Compared to existing optical localization methods, the proposed method achieves a higher accuracy with a larger detection range. The proposed method has the advantage of only using one image sensor with a low computing overhead, which has the potential in low-power moving target tracking applications. In future work, we plan to perform outdoor flight testings under different lighting conditions, study the motion effects on event noise regarding the temporal band-pass filter using the optical flow method, develop Kalman filters, increase the processing speed, and test more frequency bands of the LED markers.

References

[1] L. Apvrille, Y. Roudier, and T. J. Tanzi, "Autonomous drones for disasters management: Safety and security verifications," in 2015 1st URSI Atlantic Radio Science Conference (URSI AT-RASC), 2015, pp. 1–2.

[2] M. Alwateer, S. W. Loke, and W. Rahayu, "Drone services: An investigation via prototyping and simulation," in 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), 2018, pp. 367–370.

[3] D. Joshi, "Drone technology uses and applications for commercial, industrial and military drones in 2020 and the future," Business Insider, Dec 2019. [Online]. Available: https://www.businessinsider.com/drone-technology-uses-applications

[4] S. Banker, "Is the future of drones now?" Forbes, Jun 2020. [Online]. Available: https://www.forbes.com/sites/stevebanker/2020/06/11/is-the-future-of-drones-now/#acf0eb032840

[5] L. Schroth, "The drone market size 2020-2025: 5 key takeaways," Droneii, Jun 2020. [Online]. Available: https://droneii.com/the-drone-market-size-2020-2025-5-key-takeaways#:~:text=1.,away,is42.8billionUSD.&text=Fromgenerating22.5billionUSD,almostdoublethatin2025.

[6] N. Iliev and I. Paprotny, "Review and comparison of spatial localization methods for low-power wireless sensor networks," Sensors Journal, IEEE, vol. 15, pp. 5971–5987, 10 2015.

[7] J. Skoda and R. Barták, "Camera-based localization and stabilization of a flying drone," 2015.

[8] C. Shinde, R. Lima, and K. Das, "Multi-view geometry and deep learning based drone detection and localization," in 2019 Fifth Indian Control Conference (ICC), 2019, pp. 289–294.

[9] T. Martelli, F. Murgia, F. Colone, C. Bongioanni, and P. Lombardo, "Detection and 3d localization of ultralight aircrafts and drones with a wifi-based passive radar," in International Conference on Radar Systems (Radar 2017), 2017, pp. 1–6.

[10] F. Guarato, V. Laudan, and J. F. C. Windmill, "Ultrasonic sonar system for target localization with one emitter and four receivers: Ultrasonic 3d localization," in 2017 IEEE SENSORS, 2017, pp. 1–3.

[11] J. Kim, C. Park, J. Ahn, Y. Ko, J. Park, and J. C. Gallagher, "Real-time uav sound detection and analysis system," in 2017 IEEE Sensors Applications Symposium (SAS), 2017, pp. 1–5.

[12] J. Han, D. Kim, M. Lee, and M. Sunwoo, "Enhanced road boundary and obstacle detection using a downward-looking lidar sensor," IEEE Transactions on Vehicular Technology, vol. 61, no. 3, pp. 971–985, 2012.

[13] A. Al-Radaideh and L. Sun, "Self-localization of a tethered quadcopter using inertial sensors in a GPS-denied environment," in IEEE International Conference on Unmanned Aircraft Systems, 2017, pp. 271–277.

[14] ——, "Observability analysis and Bayesian filtering for self-localization of a tethered multicopter in GPS-denied environments," in IEEE International Conference on Unmanned Aircraft Systems, 2019, pp. 1041–1047.

[15] Chungki Woo, Sanha Kang, Hojin Ko, Hochan Song, and Jae Ook Kwon, "Auto charging platform and algorithms for long-distance flight of drones," in 2017 IEEE International Conference on Consumer Electronics (ICCE), 2017, pp. 186–187.

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

[17] L. Teixeira, F. Maffra, M. Moos, and M. Chli, "Vi-rpe: Visual-inertial relative pose estimation for aerial vehicles," IEEE Robotics and Automation Letters, vol. 3, no. 4, pp. 2770–2777, 2018.

[18] I. White, E. Curry, D. K. Borah, S. J. Stochaj, and W. Tang, "An optical spatial localization algorithm using single temporal difference image sensor," IEEE Sensors Letters, vol. 3, no. 3, pp. 1–4, March 2019.

[19] I. White, D. K. Borah, and W. Tang, "Robust optical spatial localization using a single image sensor," IEEE Sensors Letters, vol. 3, no. 6, pp. 1–4, June 2019.

[20] J. Huang, S. Wang, M. Guo, and S. Chen, "Event-guided structured output tracking of fast-moving objects using a celex sensor," IEEE Transactions on Circuits and Systems for Video Technology, vol. PP, pp. 1–1, 05 2018.

[21] Zheng Yi and Fan Liangzhong, "Moving object detection based on running average background and temporal difference," pp. 270–272, Nov 2010.

[22] J. Xu, M. Jiang, L. Yu, W. Yang, and W. Wang, "Robust motion compensation for event cameras with smooth constraint," IEEE Transactions on Computational Imaging, vol. 6, pp. 604–614, 2020.

[23] T. Brox, B. Rosenhahn, J. Gall, and D. Cremers, "Combined region and motion-based 3d tracking of rigid and articulated objects," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 3, pp. 402–415, 2010.

[24] V. Lepetit and P. Fua, Monocular Model-Based 3D Tracking of Rigid Objects: A Survey, 2005.

[25] S. Srigrarom, S. M. Lee, M. Lee, F. Shaohui, and P. Ratsamee, "An integrated vision-based detection-tracking-estimation system for dynamic localization of small aerial vehicles," in 2020 5th International Conference on Control and Robotics Engineering (ICCRE), 2020, pp. 152–158.

[26] R. Mur-Artal and J. D. Tards, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255–1262, 2017.

[27] Y. Liu, X. Yu, S. Chen, and W. Tang, "Object localization and size measurement using networked address event representation imagers," IEEE Sensors Journal, vol. 16, no. 9, pp. 2894–2895, May 2016.

[28] O. Moolan-Feroze, K. Karachalios, D. N. Nikolaidis, and A. Calway, "Improving drone localisation around wind turbines using monocular model-based tracking," in 2019 International Conference on Robotics and Automation (ICRA), 2019, pp. 7713–7719.

[29] Y. Liu, P. Sun, and A. Namiki, "Target tracking of moving and rotating object by high-speed monocular active vision," IEEE Sensors Journal, vol. 20, no. 12, pp. 6727–6744, 2020.

[30] Z. Wang, M. Xu, N. Ye, F. Xiao, W. n. Ruchuan, and H. Huang, "Computer vision-assisted 3d object localization via cots rfid devices and a monocular camera," IEEE Transactions on Mobile Computing, pp. 1–1, 2019.

[31] Celepixel User Manual. [Online]. Available: https://github.com/CelePixel/CeleX4-OpalKelly