Fast and Robust Event-based Optical Communication for Aerial Robots via Active LED Markers

Nafiseh Jabbari Tofighi, Maxime Robic, Jordan Caracotte, Pascal Vasseur, Fabio Morbidi

***Abstract*— Unmanned Aerial Vehicles (UAVs) are popular robotic platforms nowadays and they are gradually entering our everyday lives. For normal operation, a two-way communication link between a UAV and a ground station, has to be established. While radio-frequency communication is a simple and common option, it suffers from major drawbacks, such as interference and multi-path errors. Moreover, it is prone to spoofing/jamming attacks which might have catastrophic consequences in densely-populated or sensitive areas. To overcome these limitations, in this paper, we propose a compact LED-event camera system for fast optical communication via a new *optimized dynamic N-pulse protocol*. A learning-based method guarantees the robust detection and tracking of the active markers and it makes the application of standard 3D pose estimation algorithms possible. The effectiveness of the proposed system is demonstrated via hardware experiments with a DJI Matrice 600 Pro hexarotor and a Prophesee Gen. 4.1 HD camera.**

I. Introduction

Robots are confronted today with increasingly demanding tasks in which communication plays a crucial role, especially in safety-critical applications where decision-making depends on low-latency and reliable data transmission [1]. The field of aerial robotics is particularly sensitive to trustworthy communication [2]: in fact, even minor disruptions (e.g. small time delays) might have a serious impact, possibly resulting in collisions or crashes.

Information exchange can take place via Radio Frequency (RF) communication, visible light communication, and infrared (IR) or wired communication [3]. RF communication is the most prevalent in aerial robotics, but it incurs a number of issues, in practice: interference, multi-path errors (e.g. in urban canyons), congestion, bandwidth saturation, and background noise [4]. Moreover, the power demand could lead to fast battery depletion. To address these limitations, *Optical Camera Communication* (OCC) represents a valid alternative: it is cost-effective and it can be easily integrated into existing robotic platforms. On the downside, most OCC systems rely on the CMOS/CCD technology, which results in high latency and low bit rates. High-speed cameras can partially mitigate these problems, but they are rarely deployed in the field, because of their cost, power consumption, and delicate operation.

To get around the limitations of traditional cameras which buffer absolute intensities and output timed frames, *event-based cameras* represent an attractive option [5]. In fact, these bio-inspired sensors capture per-pixel brightness changes with microsecond resolution. The asynchronous and sparse nature of the events offers distinctive advantages, which are highly desirable in time-constrained communication systems. For example in [6], the authors demonstrated robust communication in noisy and congested environments, highlighting the potential of event cameras for real-time coordination.

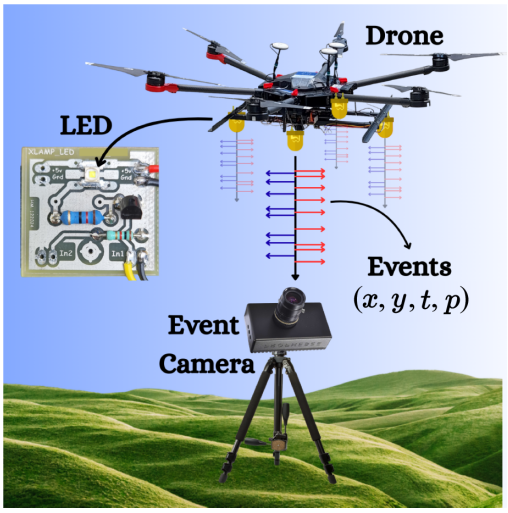


Fig. 1: Overview of the proposed optical communication system for drones, based on an event camera and active markers (LEDs).

Recently, event cameras have been coupled with *optical markers*, which simplified the execution of complex tasks, such as detection, tracking, pose estimation, and communication. These systems achieve high performance since *Event-Based Optical Marker Systems* (EBOMS) [7] only focus on limited regions in space, and they ensure high speed and robustness [8]. Different typologies of markers have been studied in the literature, which can be classified according to their geometry and shape (e.g. ArUco markers, color patterns, bar codes) or to the nature of light produced (visible, IR or ultraviolet radiation). They can also be active or passive, depending on whether they emit or reflect light. Thanks to their compact form factor, markers can be easily mounted on small aerial robots, enabling real-time data transmission [9]. Motivated by these emerging technologies, in this paper, we introduce a communication framework for aerial robots based on EBOMS with active markers. The encoded messages are transmitted from the UAV to the event camera using a lightweight and efficient protocol, providing a robust and real-time communication channel.

The authors are with the MIS laboratory, University of Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens Cedex, France.
Emails: firstname.lastname@u-picardie.fr

This work was supported by the French National Research Agency through the DEVIN project, "*Drones with Omni-Event Vision for Drone Neutralization*" (ANR-23-IAS2-0001).

Foundational work on detection and tracking with EBOMS was done in [10], where the authors proposed two algorithms for real-time 2D visual tracking of blinking LEDs (Light-Emitting Diodes). The first accumulates events within a fixed time window, while the second asynchronously detects event patterns matching LED frequencies, for more accurate and robust tracking. Besides LEDs, geometric markers (e.g. ArUco markers), whose blinking patterns can be easily detected by an event camera in complex scenarios, have also been used in the literature [11], [12], [13]. Although active markers contribute to higher detection accuracy, robust data processing is necessary in non-ideal conditions (i.e. in the presence of vibrations, external perturbations or sharp accelerations). To overcome these limitations, our optical communication system leverages a machine-learning algorithm which makes our marker-detection stage, highly reliable and efficient.

Thus far, only a few papers have exploited EBOMS for communication purposes. Most of the existing approaches build upon the pioneering work of Censi *et al.* [14], which first explored the communication capabilities of EBOMS. A notable contribution in this field is [9], where the authors presented a robust beacon-tracking system that combines density-based clustering, event-driven optical flow, and Kalman filtering to estimate motion and position, followed by reliable message decoding with an integrated error-detection mechanism. Other studies, such as [15] and [16], demonstrated the effectiveness of optical communication systems using LED-based signaling, achieving reliable, high-bandwidth data transmission. In [17], an EBOMS capable of error-free long-range optical communication at 500 bps in bright sunlight, is introduced. It combines compact binary packet encoding, blob detection, high-pass filtering, and threshold-based decoding with high-performance LEDs and optical lenses. Although high accuracy is ensured, using optical lenses with a narrow field of view is problematic with moving platforms, as is typically the case in aerial robotics. A performance degradation over long distances is also observed in [17], without lenses. Moreover, while ASCII decoding can be very effective in stationary conditions, it becomes highly sensitive to noise in the presence of motion. To avoid the drawbacks of ASCII decoding, we designed an optimized dynamic N-pulse protocol. Besides being efficient, fast, and adaptable to a broad range of communication modes, the protocol produces constant pulse patterns that strengthen the detection stage. This has a beneficial effect on the robustness and accuracy of transmission.

EBOMS-based communication systems have received relatively little attention in robotics. The existing implementations have not been specifically designed for aerial robots, where vibrations and variable lighting conditions might dramatically degrade performances. Recently, in [9], the authors considered a simulated scenario where a UAV equipped with an event camera monitors assets on a construction site using blinking beacons for simultaneous communication and tracking. However, no real-world experiments have been conducted to evaluate the robustness of the communication protocol. Moreover, in [9] and related works, a single message or very basic data are transmitted, while real applications often require multiplexing to simultaneously transmit several messages or numerical data, such as position, velocity, or acceleration. Our system has been specifically designed for one-way communication in aerial robotics, and its efficiency has been proved via hardware experiments.

When combined with event cameras, optical markers (and especially the geometric ones) deliver strong performance for *pose estimation* tasks [18], [19], [20], [8]. Once the markers have been detected, a classical optimization problem, such as PnP (Perspective-n-Point), is generally solved [21]. However, the localization of each marker remains an open problem. The state-of-the-art solutions, e.g. ArUco markers or LEDs arranged in specific patterns, often increase the complexity the system and add extra costs. In our case, pose estimation is inherent in the communication system itself. In fact, LED's identification is integrated into the protocol, and it is directly derived from the content of the transmitted message, rather than from external cues, such as distinct blinking frequencies. As a result, marker's ID is conveyed by the communication layer, thus making pose estimation a natural by-product of decoding process, rather than a separate detection step. The LEDs can thus be easily identified and localized in 2D, providing the necessary correspondences for running a 3D pose-estimation algorithm.

In summary, in this work, we present a new real-time optical communication system for UAVs that combines active markers (LEDs) and an event camera (see Fig. 1). The message is encoded by the blinking pattern of the LEDs, observed by the event camera. For LED detection and tracking, we designed a robust and flexible learning-based algorithm. Once the spatial location of the LEDs has been determined, the proposed protocol leverages an optimized dynamic N-pulse strategy, which guarantees high transmission rates and supports a wide range of messages, including numerical data (such as UAV's telemetry).

The main contributions of this paper can be stated as follows:

- A compact event-based optical communication system for UAVs.
- A novel real-time communication protocol that relies on active markers (blinking LEDs) as transmitters and on an event camera as receiver, and a custom learning-based algorithm for reliable LED detection and tracking.
- Evaluation of system's accuracy and robustness via an extensive experimental campaign.

In the remainder of this article, Sect. II presents the proposed system in detail, with a special focus on the event-based LED detection algorithm and communication protocol. The results of real-world experiments are discussed in Sect. III. Finally, Sect. IV concludes the paper with a summary of our contributions and an outline of future research directions.

II. Material and methods

In this section, we describe the main functional blocks of our optical communication system. The exposition begins with a general overview of the system (Sect. II-A), followed by the presentation of our event-based algorithms for LED detection and tracking (Sect. II-B). Finally, Sect. II-C is devoted to our new communication protocol.

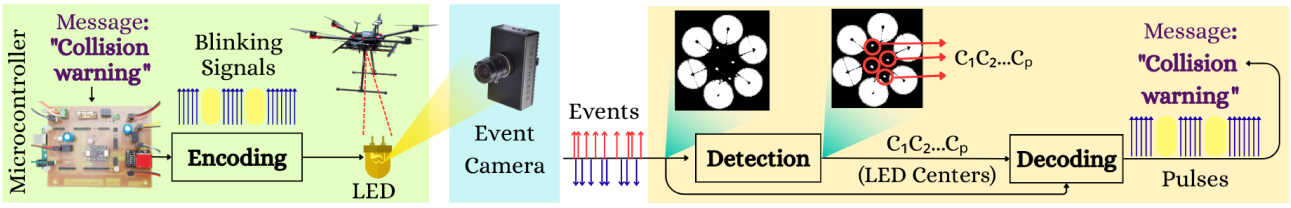


Fig. 2: Functional blocks of our optical communication system.

A. General architecture

Fig. 2 shows an overview of our optical communication system. It includes P LEDs[1] mounted on a UAV (green block in Fig. 2). By switching P transistors, a microcontroller adapts the blinking pattern of the LEDs to the communication protocol. An event camera observes the active markers (Fig. 2, cyan block). The centers of the LEDs are detected and the message, e.g. "*Collision warning*", is finally decoded (Fig. 2, orange block).

B. Event-based LED detection

The raw events generated by the event camera must be processed in order to determine the position of the LEDs (this operation is fundamental for the subsequent message encoding/decoding steps). The corresponding functional block then takes the raw events as input and outputs the coordinates of estimated centers of the LEDs, $C_k(x_k,\, y_k)$, $k \in \{1, \ldots, P\}$. UAV's position is inferred via LEDs tracking-by-detection over successive time steps.

In controlled laboratory conditions, the detection step can be efficiently performed by using a combination of event histograms and clustering techniques (see Sect. III for more details). The process begins by applying a high-pass filter to remove low-frequency events through bias tuning: in this way, we get rid of background noise and random events generated by mechanical vibrations. The events are then randomly downsampled by a factor β to reduce the computational overhead during the construction of the histogram. Downsampling has no impact on the accuracy of decoding step, since downsampled data is only used for the detection of centers $C_k(x_k,\, y_k)$, while raw data is preserved for decoding.

The event histogram is defined as a 2D spatial accumulation map $H(x,\, y)$, where the intensity at pixel $(x,\, y)$ corresponds to the number of events within a temporal window T_w:

$$H(x,\, y) \;=\; \sum_{t_i \,\in\, [t - T_w,\, t]} \delta(x - x_i,\, y - y_i),$$

where $(x_i,\, y_i,\, t_i)$ are the spatial and temporal attributes of the events, $t \geq T_w$ and $\delta(\cdot,\, \cdot)$ is 2D Dirac delta function (the polarity $p_i = \pm 1$ of the events is ignored). From this histogram, regions of high activity are identified by applying an intensity threshold τ. The high-intensity points are then grouped using the DBSCAN algorithm [22], which clusters nearby event peaks, while filtering out noise. The detected positions $C_k(x_k,\, y_k)$ of the LEDs are passed to the decoding stage, for further processing in the communication pipeline.

[1]For 3D pose estimation (see Sect. II-C), $P \geq 3$.

By moving from laboratory to real-life conditions with the LEDs attached to a UAV, we observed a dramatic performance degradation of DBSCAN (due to the motion of the UAV and the high-frequency events generated by the spinning propellers). To address this problem, we designed a *machine-learning algorithm* for the detection of the centers $C_k(x_k,\, y_k)$. To learn the location of the centers from unordered and noisy event data, we adopted a PointNet-style MLP (Multilayer Perceptron) architecture [23]. Each input sample consists of a set of event coordinates extracted within a temporal window, from which a fixed number of points is randomly sampled and normalized to form an unordered point cloud. During inference, the network processes each event independently and outputs the estimated LED centers, which are directly used in the decoding stage. This model is well-adapted to our problem since it treats the input pixel coordinates $(x_i,\, y_i)$ as unordered point clouds without assuming any fixed spatial structure, like conventional CNNs. This helps to preserve the asynchronicity of event data. The network applies a shared MLP across all event points, followed by global pooling to capture spatial patterns. A lightweight regression head predicts 8 normalized values corresponding to the coordinates of LEDs' centers $C_k(x_k,\, y_k)$. To improve robustness, we added a learnable bias term to adjust for systematic shifts. For training, we used a composite loss, combining the Hungarian algorithm for permutation-invariant matching and mean squared error for stability.

C. Communication protocol

While fast and robust optical communication holds great potential for aerial robotics, the subject remains relatively under-explored. In fact, most of the previous works limited to simple fixed-length or static messages [17].

In this paper, we propose a one-way event-based optical communication protocol that not only transmits messages rapidly and efficiently, but also handles a variety of data types, including characters, strings, and numerical values. Note that when dealing with high-frequency, long-distance data transmission, traditional encodings, such as ASCII, might suffer from poor accuracy. In fact, small delays or interruptions might disrupt entire messages. Another drawback of ASCII encoding is its fixed-length representation (e.g. 7 bits/character), which requires messages to be transmitted sequentially (character by character). This results in a low data rate, which is problematic in many applications.

1) ***Dynamic N-pulse****:* In [24], the authors proposed *N-pulse encoding*, a novel protocol for LED-based communication. Instead of using binary or ASCII codes, information is encoded in the number of pulses within a fixed time

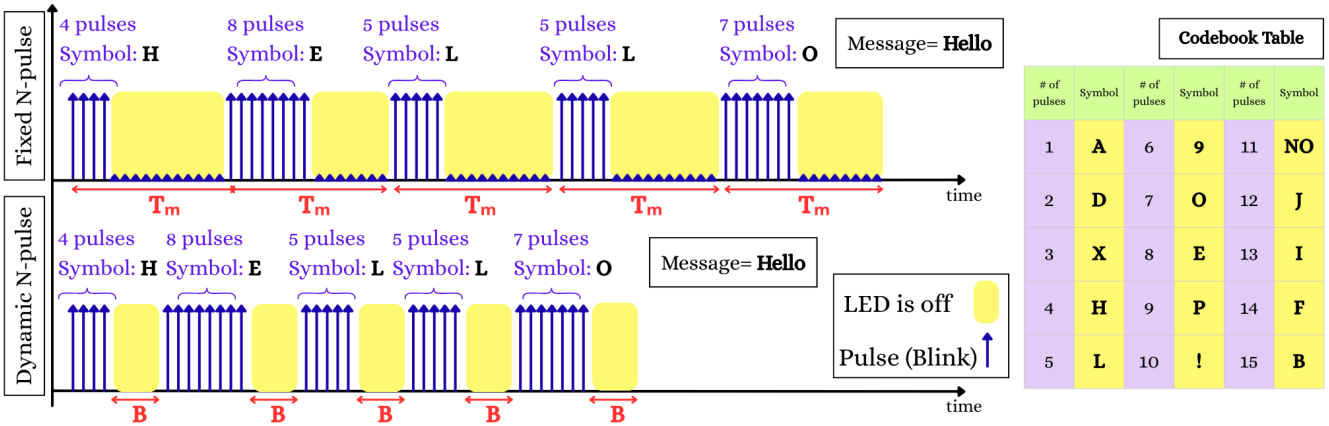


Fig. 3: Comparison between (top left) *fixed N-pulse* transmission, and (bottom left) *dynamic N-pulse* transmission proposed in this paper. The codebook table is shown on the right side.

window T_m: this guarantees low-latency and high-bandwidth data transmission. More precisely, the N-pulse method assigns a unique number of light pulses to each symbol within a fixed time frame (see Fig. 3). To decode the transmitted message, the receiver only needs to count the pulses in that interval. However, the performance of this fixed-frame approach degrades as the number of symbols increases. In all cases, whether a symbol consists of two pulses or twenty, one should wait until the end of the full time frame before transmitting the next message. This has a negative impact on the overall transmission time.

To overcome this limitation, we propose a *dynamic N-pulse* method. Instead of waiting until the fixed time frame has elapsed, the next message is initiated as soon as the previous one has been completed. This approach efficiently uses fixed *OFF periods* B, where the duration of a symbol is defined only by LED's blinking time, proportional to the number of pulses assigned to it. As a result, the time frame does not grow with the number of symbols, and only the fixed OFF period B is maintained between transmissions. Pulse counts are encoded into symbols using a codebook (see the table in Fig. 3). This offers great flexibility: in fact, the codebook can be adapted to the application at hand, and symbols may represent characters, numbers, words, or even complete sentences.

For example, with reference to Fig. 3, consider the transmission of the simple message "*HELLO*". In the fixed N-pulse method, each symbol occupies the same time frame T_m, thereby message's duration scales with the number of symbols, regardless of their pulse count. In contrast, the dynamic N-pulse strategy adjusts the duration of each symbol based on its pulse count, while maintaining a constant OFF period B. This adaptive time mechanism allows the message "*HELLO*" to be transmitted faster, without idle times.

For a comparative analysis of the fixed and dynamic N-pulse methods, let N be the number of distinct symbols, p_m the number of pulses required by message m, and t_p the duration of a single pulse (identical in both strategies). We also assume that each of the N symbols is transmitted once, providing a uniform evaluation scheme for both methods. In the fixed N-pulse method, the duration of a single symbol is constant and independent of the number of pulses assigned to it. The transmission time for one symbol and the total transmission time for all N symbols are $T_{\text{Fixed}} = Nt_p$ and $T_{\text{Total, Fixed}} = NT_{\text{Fixed}}$, respectively. In the fixed N-pulse method, the total transmission time grows quadratically as the number of symbols increases, which is inefficient with large sets of symbols. On the other hand, in the dynamic N-pulse method, the next message is initiated immediately after the pulses of the current symbol have been generated, followed by a fixed OFF period B. In this case, the transmission time for symbol m and the total transmission time for all N symbols are, respectively,

$$T_m = p_m t_p + B, \qquad T_{\text{Total, Dyn.}} = \sum_{m=1}^{N} T_m.$$

Since the expected number of pulses per symbol is $\mathbb{E}[p_m] = \frac{1}{2}(N+1)$, the expected total transmission time simplifies to

$$T_{\text{Total, Dyn.}} = N(\mathbb{E}[p_m]\, t_p + B). \tag{1}$$

In practical scenarios, symbol distribution is typically non-uniform, with certain symbols occurring more frequently than others. In such cases, the dynamic N-pulse method benefits from adapting the transmission length to each symbol, which may result in lower overall transmission times compared to the fixed strategy.

*2) **Optimized Dynamic N-pulse**:* For large values of N, the transmission time can be further reduced by using an *optimized dynamic N-pulse* method. The idea is to assign shorter pulse counts to symbols which are most likely to occur: this ensures that frequently-encountered symbols are transmitted faster (see Fig. 4 for a numerical comparison between the three methods considered in this section). If a symbol belongs to the top-f fraction of most probable symbols, it is transmitted using $p_{\min}$ pulses: otherwise, it is transmitted with $p_{\max}$ pulses. For example, the 0–9 digits frequently appear in the applications (e.g. to represent sensor values), and they are ideal candidates for short pulse assignments. On the whole, compared to the dynamic N-pulse method, the average transmission time decreases, and the efficiency gain

becomes more important when the probability distribution of symbols is highly imbalanced. Formally, the expected number of pulses per symbol is $\mathbb{E}[p_m] = f p_{\min} + (1-f)\, p_{\max}$, where f is the fraction of high-probability symbols, $p_{\min}$ is the pulse count assigned to these symbols, and $p_{\max}$ is the pulse count assigned to the others. The expected total transmission time is again given by equation (1).

Coming back to the previous example, the dynamic N-pulse method requires a total transmission time of $29\, t_p + 5B$ for the "*HELLO*" message. If we now apply the optimized dynamic N-pulse method and assign shorter pulse counts in the codebook to the most frequent symbols (e.g. for the letter "*L*", we pass from 5 pulses to 1 pulse), the total transmission time reduces to $21 t_p + 5B$.

It is worth pointing out here that the problem of finding the *optimal symbol assignment* for a given "alphabet" (i.e. the set of all possible symbols to transmit) is non-trivial, in general. More advanced statistical or corpus-driven encoding schemes could further improve compression efficiency, but this goes beyond the scope of this paper. In fact, our final goal is to devise a simple yet effective encoding scheme for event-based optical communication with UAVs.

3) ***Transmission modes****:* The proposed system supports three transmission modes to adapt to different operational requirements. These modes share the same baseline encoding strategy (optimized dynamic N-pulse), but vary in how messages are distributed across the LEDs: this offers a good balance between accuracy and speed. In *Mode 1*, the P LEDs transmit the same message in parallel (see Fig. 5). This mode is useful when accuracy takes priority over transmission speed (e.g. when critical sensor measurements from the UAV are to be sent to the ground station). In *Mode 2*, each LED transmits a separate message, enabling parallel or multi-channel communication. This guarantees high speed, since multiple data streams are transmitted simultaneously. Finally, *Mode 3* is a fast variant of Mode 2, in which each LED simply transmits its unique identifier using a short message (the identifiers are single symbols and they depend on the number P of LEDs). After the detection and decoding steps, the location of each LED can be efficiently determined: this makes fast 3D pose estimation possible (e.g. by solving the PnP problem with $P \geq 3$, see Fig. 8 in Sect. III-C).

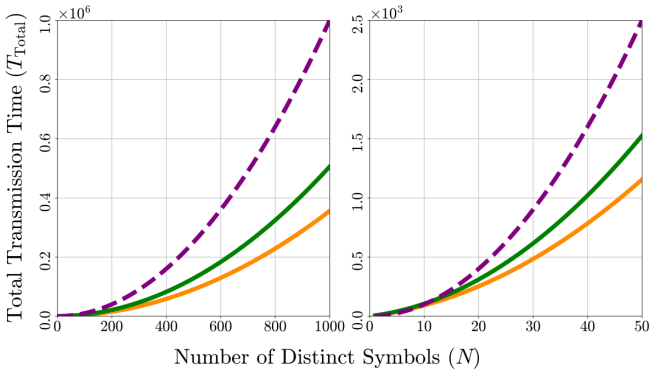


Fig. 4: *Total transmission time (milliseconds) versus number of symbols* N: Fixed N-pulse (dashed purple), dynamic N-pulse (solid green) and optimized dynamic N-pulse (solid orange). (Left) Full range of symbols, $N \in [0,\, 1000]$, and (right) zoom-in for $N \in [0,\, 50]$.

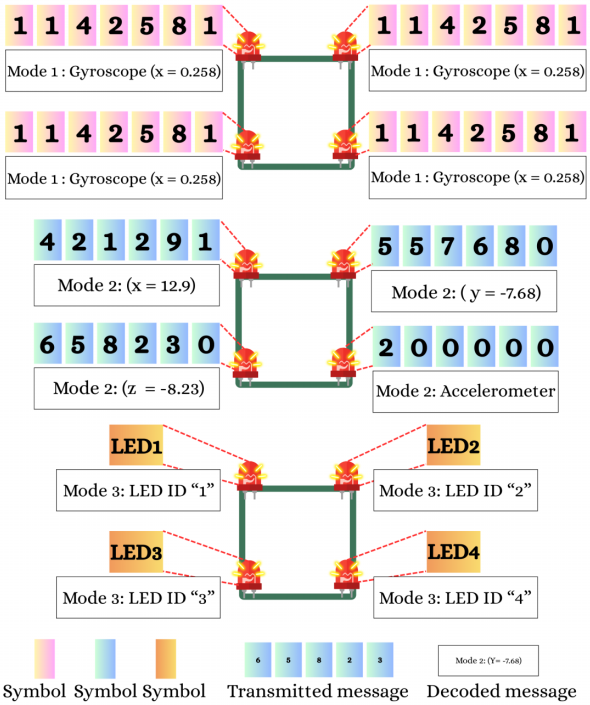


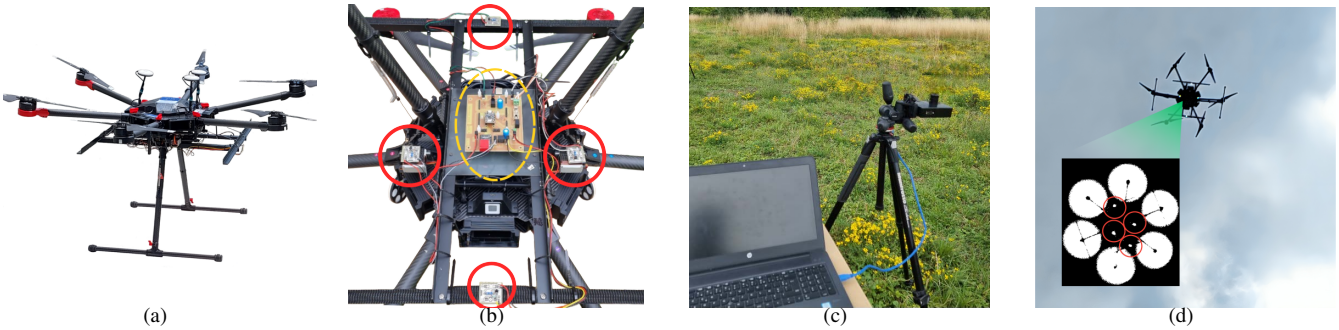
Fig. 5: Graphical illustration of Modes 1, 2 and 3. In this example, $P = 4$.

III. Experimental Validation

A. Hardware setup

To evaluate the performance of the proposed communication system, we conducted a wide range of experiments in both controlled and real-life conditions. In all tests, the number of LEDs was set at $P = 4$, to allow robust 3D pose estimation. In the controlled experiments, the LEDs formed a rectangular (12 cm × 11 cm) pattern on a movable panel detached from the UAV and they were observed by a Prophesee Gen. 4.1 HD camera with a 5 mm AICO lens. In the real-life experiments, the microcontroller and the LEDs that it drives, were mounted on a metallic support underneath the UAV (a DJI Matrice 600 Pro hexarotor, see Fig. 6(a)): the LEDs formed a planar kite pattern (diagonals: 56 cm and 30 cm, see Fig. 6(b)). Special attention was paid to the location of the microcontroller and LEDs (to not overly perturb the center of gravity of the UAV and to minimize the interference of six propellers). The event camera equipped with a 16 mm AVENIR lens was installed on a tripod, pointing upwards towards the UAV (see Fig. 6(c)). In our tests, performed under overcast sky conditions (see Fig. 6(d)), the altitude of the UAV ranged from 15 to 30 m.

To support real-time communication, we chose a custom microcontroller based on the Arduino Mega board. Its primarily role is to switch the 4 high-luminance LEDs on and off (low-luminance LEDs, such as the standard 5 mm, 20 mA types, were tested but they fell short of providing a satisfactory



(a) (b) (c) (d)

Fig. 6: *Experimental setup*: (a) Abeam view of the DJI hexarotor before taking off; (b) Location of the microcontroller (circled yellow) and of the 4 LEDs (circled red), underneath the UAV; (c) The Prophesee event camera connected to a laptop pointing skyward; (d) The hexarotor flying 30 m above the ground, with the 4 LEDs transmitting a message.

performance outdoors). However, for minimum impact on the autonomy of the UAV, a good balance between brightness and energy consumption should be found. The Cree XP-E2 white LEDs are a perfect fit: in fact, they are small, low power and ultra bright (320 lm).

The microcontroller offers extra functionalities as well: the ICM-20948 IMU (Inertial Measurement Unit) provides real-time attitude information and includes a 3-axis gyroscope, a 3-axis accelerometer and a 3-axis compass. These measurements can be transmitted to the ground station with our system, when the hexarotor is in the air.

B. Communication modes

For the three modes, both *offline* and *online* transmission schemes are supported. Offline transmission allows to use predefined, frequently-occurring messages, such as "*Low battery*", "*Collision warning*" or "*Target detected*". In contrast, online transmission dynamically encodes live data (e.g. measurements from on-board sensors), using the optimized dynamic N-pulse method. The encoded signals are optically transmitted by modulating the blinking rate of the LEDs in real time. The codebook used in our experiments consists of 90 symbols, sorted in ascending order of pulse count: 4 LED identifiers, 10 numerical digits, 50 predefined static messages, and 26 characters of English alphabet (see Fig. 3).

The data-encoding structure was conceived to be flexible and application-driven. A 5-digit representation inspired by scientific notation was used for floating-point sensor reading (a sign indicator, a 3-digit mantissa, and a 1-digit exponent). The sign digit jointly encodes the sign of the mantissa and exponent, using four values to represent all possible cases (e.g. the sequence `42581` corresponds to the value 0.258). This encoding scheme for the float numbers is used in Mode 1 and Mode 2. The three operating modes can be distinguished by the total length of the transmitted message. In *Mode 1*, each message consists of 7 symbols: the first digit specifies the sensor type (e.g. a gyroscope or an accelerometer), the second digit indicates which axis (x, y, or z) is considered, and the last five digits represent the encoded sensor value. For instance, `1142581` corresponds to the measurement with respect to the x-axis provided by a gyroscope, whose value is 0.258. In this mode, all four LEDs broadcast the same message in parallel, and redundancy increases reliability. In *Mode 2*, each message consists of 6 symbols: the first digit specifies the axis, and the following five digits encode the value. This allows parallel communication, with three LEDs sending the values corresponding to the x, y, and z axes, and the fourth LED transmitting the sensor type. Compared to Mode 1, the effective data rate in Mode 2, is higher. Finally, in *Mode 3*, each message is a single symbol, corresponding to the identifier of each LED. This mode is apt for rapid LED identification and 3D pose estimation.

Note that the transmission rate of the optimized dynamic N-pulse method depends on the selected mode, number of symbols, pulse duration, and OFF period. In our case, we set the pulse duration at $t_p = 2$ ms and the OFF period at $B = 5\,t_p$. In the fastest mode (Mode 3), each LED sends its identifier using a single symbol: hence, the transmission time is about 12 ms for the four messages and this can be regarded as an upper bound for our system. Modes 1 and 2 lag behind because of the multi-symbol message mechanism, but real-time communication can still be guaranteed. While shorter time pulses have been reported in the literature (e.g. in [21], the authors went down to 3 μs), the experiments have been mostly performed in static scenarios or controlled laboratory conditions. When the active markers are mounted on an aerial robot, the motion and vibrations of the platform make the decoding step harder and ultra-short light pulses further exacerbate this problem. Our choice of $t_p = 2$ ms provides a good trade-off between speed and robustness to motion. Higher data rates could be possibly achieved through massive parallel transmission without detriment to reliability.

C. Discussion of the results

In the controlled conditions, we performed two trials with the communication system off-board the UAV. The first trial was carried out indoors under two lighting conditions: in the presence of artificial illumination and in a dark environment. The goal was to evaluate the effect of neon lighting, which is known to generate multiple artifacts (spurious events caused by the 50 Hz utility frequency). In the second trial, conducted outdoors, we studied the effect of an increasing distance d between the transmitter and the receiver, on the communication error.

A second set of experiments was performed in real-life conditions, with the 4 LEDs mounted underneath the UAV and the event camera serving as a receiver on the ground. In this case, the optical communication system was exposed to the vibrations of the flying robot and to variable light conditions. Despite these difficulties, the detection and decoding steps ran in real time. The reader is referred to the **video accompanying the paper**, for more details on the experimental results.

In the experiments, we tested both the *offline* and *online* transmission modes. In the offline mode, a set of five predefined messages was selected, while in the online mode, the LEDs periodically transmitted sensor data, with the microcontroller encoding the sensor values every 3 s. In each experiment, we repeated the online transmission procedure 100 times for each mode, to ensure statistical significance. For improved reliability, each message was transmitted three times in total (the original message and two redundant copies). The fixed and dynamic N-pulse strategies differ primarily in transmission efficiency, while sharing the same detection and decoding pipeline: hence, their accuracy is expected to be the same (for a comparison of transmission times, see Fig. 4).

For performance evaluation, we considered two metrics. The *Symbol Error Accuracy* (SAR) quantifies the ratio of correctly-received symbols to the total number of transmitted symbols, while the *Message Accuracy Rate* (MAR) measures the ratio of correctly-received messages to the total number of transmitted messages. The slight difference between these two metrics is to be ascribed to the redundancy in the transmission protocol: in fact, even if a symbol is lost, redundant copies of that symbol within the message often ensure that the overall message is still correctly reconstructed.

Our communication system exhibited strong and consistent performance across all experiments (see Table I). In the indoor experiments, 100% accuracy was achieved when the lights were off. Even in the presence of artificial lighting, the system effectively handled the spurious events, maintaining high accuracy (exceeding 96%, on average). In the outdoor controlled experiments, we increased the distance d between the transmitter and the receiver from 2 m to 15 m, in 0.5 m increments: we did not zoom on the panel containing the LEDs, so as not to modify the field of view of the camera.

TABLE I: Summary of the experimental results.

		Specifications	SAR (%) ↑	MAR (%) ↑
Controlled cond.	**Indoor**	Lights on	96.43 ± 2.12	100.00 ± 0.00
		Lights off	100.00 ± 0.00	100.00 ± 0.00
	Outdoor Distance (m)	$d \in [2, 6)$	98.86 ± 1.27	100.00 ± 0.00
		$d \in [6, 9)$	98.07 ± 1.05	100.00 ± 0.00
		$d \in [9, 12)$	95.59 ± 2.64	98.86 ± 0.35
		$d \in [12, 15]$	87.32 ± 2.88	100.00 ± 0.00
Real-life cond.	**Mode 1**	Single message	68.18 ± 3.39	100.00 ± 0.00
	Mode 2	Multi-channel	71.71 ± 2.74	93.48 ± 1.80
	Mode 3	ID-based	84.95 ± 2.53	97.26 ± 0.23
	Mode 2	ASCII code	32.09 ± 5.30	42.73 ± 4.11

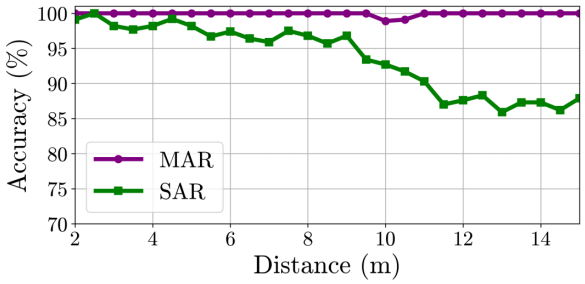


Fig. 7: Effect of distance d (meters) on transmission accuracy: MAR (purple) and SAR (green), in percentage.

The results in Table I and Fig. 7 indicate that the system delivers a quasi-perfect performance up to 10 m, and that the distance d has a negligible effect on transmission accuracy.

The real-life experiments with the LEDs on-board the hexarotor, corroborated our early findings. In fact, despite subtle light variations and spurious events produced by the six propellers, the system successfully operated in the three modes. In Mode 1, we obtained a 100% MAR, since all four LEDs transmitted the same message in parallel (a message is successfully received if it is correctly decoded from at least one LED). In Modes 2 and 3, the accuracy remained high, confirming that our system is able to reliably deliver complex messages. As mentioned in Sect. II, Mode 3 is well suited for 3D pose estimation, since the position of the LEDs is estimated in the detection step and their IDs can be rapidly broadcast in the communication step. We took advantage of the planar arrangement of the four LEDs on the hexarotor, to solve the PnP problem [25] and estimate the relative pose of the UAV with respect to its initial location (see Fig. 8). Since each LED was correctly detected and identified by our system, no rotational ambiguity was experienced.

To further illustrate the benefits of the proposed communication protocol, an ablation study was performed. We compared the optimized dynamic N-pulse method with a classical fixed-length ASCII-based encoding, in Mode 2 (indeed, fixed-length character encoding is relatively common in the existing optical communication systems). ASCII transmission resulted in a drastic drop in both SAR and MAR, confirming its high sensitivity to noise and motion (cf. the last row of Table I). The optimized dynamic N-pulse method emerged as the clear winner, since the constant pulse patterns allow robust LED-based communication, even in the presence of external perturbations.

As a closing remark, it should be noted that while prior works have dealt with EBOMS and OCC, ensuring a fair numerical comparison with the proposed system is not a trivial task. In fact, to the best of our knowledge, no *benchmark* exists today in aerial robotics, to evaluate the performance of competing event-based optical communication frameworks.

IV. Conclusion and Future Work

In this paper, we presented a compact real-time optical communication system for aerial robots, based on an event camera and active markers (LEDs). A novel communication protocol, termed *optimized dynamic N-pulse*, has been

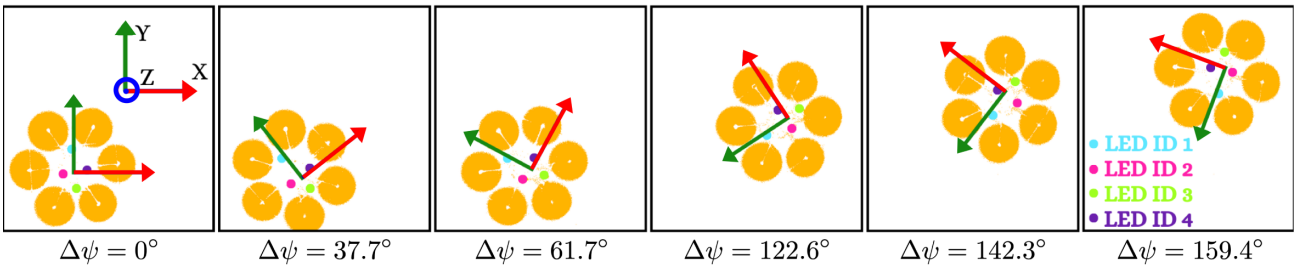


Fig. 8: *Example of 3D pose estimation (Mode 3).* In this sequence (from left to right, top view of the hexarotor), the PnP problem was solved using the four LED markers (colored dots). We obtained a total rotation $\Delta\psi = 159.4°$ about the yaw axis of the UAV, and a constant altitude of 15.95 m. In comparison, the IMU of the aerial robot provided an estimate for the orientation and altitude of 157.1° and 17.59 m, respectively.

designed and successfully tested in real-world conditions using a DJI hexarotor and a Prophesee event camera.

There are several research directions that we would like to further explore in the future. A more sophisticated light emitter, consisting of a matrix of LEDs evenly distributed over the surface of the UAV, would allow the transmission of more complex messages no matter its attitude, but it would make the decoding step more challenging. In this work, we opted for ultra-bright visible-light LEDs, but other wavelengths (e.g. ultraviolet radiation as in [26], [27]) could be considered, thus making the messages invisible to the naked eye. The proposed system has been validated up to 30 m with perfect visibility in daylight conditions: beyond this distance, specialized optics might be required. We also plan to adapt our current prototype to ground-to-multi-UAV and UAV-to-UAV communication. In the long term, our research will lead to stealthy aerial vehicles that, in the absence of RF emissions, are insensitive to jamming attacks.

References

[1] J. Gielis, A. Shankar, and A. Prorok. A Critical Review of Communications in Multi-robot Systems. *Curr. Robot. Rep.*, 3(4):213–225, 2022.

[2] L. Gupta, R. Jain, and G. Vaszkun. Survey of Important Issues in UAV Communication Networks. *IEEE Commun. Surv. Tutor.*, 18(2):1123–1152, 2015.

[3] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra. Visible Light Communication, Networking, and Sensing: A Survey, Potential and Challenges. *IEEE Commun. Surv. Tutor.*, 17(4):2047–2077, 2015.

[4] N. Saeed, S. Guo, K.-H. Park, T. Y. Al-Naffouri, and M.-S. Alouini. Optical camera communications: Survey, use cases, challenges, and future trends. *Phys. Commun.*, 37:100900, 2019.

[5] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza. Event-based vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(1):154–180, 2022.

[6] H. Su, L. Gao, T. Liu, and L. Kneip. Motion-Aware Optical Camera Communication With Event Cameras. *IEEE Rob. Autom. Lett.*, 10(2):1385–1392, 2025.

[7] N. Jabbari Tofighi, M. Robic, F. Morbidi, and P. Vasseur. A Survey on Event-based Optical Marker Systems. *IEEE Rob. Autom. Mag.*, 2026.

[8] G. Ebmer, A. Loch, M. N. Vu, R. Mecca, G. Haessig, C. Hartl-Nesic, M. Vincze, and A. Kugi. Real-Time 6-DoF Pose Estimation by an Event-Based Camera Using Active LED Markers. In *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, pages 8137–8146, 2024.

[9] A. von Arnim, J. Lecomte, N. E. Borras, S. Woźniak, and A. Pantazi. Dynamic event-based optical identification and communication. *Front. Neurorobot.*, 18:1290965, 2024.

[10] G. R. Müller and J. Conradt. A miniature low-power sensor system for real time 2D visual tracking of LED markers. In *Proc. IEEE Int. Conf. Robot. Biomim.*, pages 2429–2434, 2011.

[11] A. Loch, G. Haessig, and M. Vincze. Event-based high-speed low-latency fiducial marker tracking. In *Proc. 15th Int. Conf. Mach. Vis.*, volume 12701, pages 323–330, 2023.

[12] S. Zhang, Y. Huang, X. Pei, H. Lin, W. Zheng, W. Wang, and T. Hou. An Improved LED Aruco-Marker Detection Method for Event Camera. In *Proc. Int. Congr. Commun. Netw. Inf. Syst.*, pages 47–57, 2023.

[13] Y. Huang, T. Hou, S. Zhang, X. Pei, R. Yu, and H. Liu. A Novel ArUco Marker for Event Cameras. In *Proc. Int. Conf. Front. Robot. Softw. Eng.*, pages 271–277, 2023.

[14] A. Censi, J. Strubel, C. Brandli, T. Delbruck, and D. Scaramuzza. Low-latency localization by active LED markers tracking using a dynamic vision sensor. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 891–898, 2013.

[15] J. Perez-Ramirez, R. D. Roberts, A. P. Navik, N. Muralidharan, and H. Moustafa. Optical Wireless Camera Communications using Neuromorphic Vision Sensors. In *Proc. IEEE Int. Conf. Commun. Workshops*, pages 1–6, 2019.

[16] H. Nakagawa, Y. Miyatani, and A. Kanezaki. Linking Vision and Multi-Agent Communication through Visible Light Communication using Event Cameras. In *Proc. 23rd Int. Conf. Auton. Agents Multiagent Syst.*, pages 1436–1444, 2024.

[17] Z. Wang, Y. Ng, J. Henderson, and R. Mahony. Smart Visual Beacons with Asynchronous Optical Communications using Event Cameras. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 3793–3799, 2022.

[18] J. Xu, J. Guo, Y. Wang, Y. Li, Y. Liang, X. Cao, and G. Ye. Method of pose tracking by event camera using LED marker points. In *Proc. 14th IEEE Conf. Ind. Electron. Appl.*, pages 1994–1999, 2019.

[19] G. Chen, W. Chen, Q. Yang, Z. Xu, L. Yang, J. Conradt, and A. Knoll. A Novel Visible Light Positioning System With Event-Based Neuromorphic Vision Sensor. *IEEE Sensors J.*, 20(17):10211–10219, 2020.

[20] M. Salah, M. Chehadah, M. Humais, M. Wahbah, A. Ayyad, R. Azzam, L. Seneviratne, and Y. Zweiri. A Neuromorphic Vision-Based Measurement for Robust Relative Localization in Future Space Exploration Missions. *IEEE Trans. Instrum. Meas.*, 73:5010212, 2024.

[21] L. Bauersfeld and D. Scaramuzza. A Monocular Event-Camera Motion Capture System. *arXiv:2502.12113*, 2025.

[22] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, and X. Xu. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.*, 42(3), 2017. Article n. 19.

[23] C.R. Qi, H. Su, K. Mo, and L.J. Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 652–660, 2017.

[24] J. Aranda, V. Guerra, J. Rabadan, and R. Perez-Jimenez. Enhancing Computational Efficiency in Event-Based Optical Camera Communication Using N-Pulse Modulation. *Electronics*, 13(6):1047, 2024.

[25] T. Ke and S.I. Roumeliotis. An Efficient Algebraic Solution to the Perspective-Three-Point Problem. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 7225–7233, 2017.

[26] V. Walter, N. Staub, A. Franchi, and M. Saska. UVDAR System for Visual Relative Localization With Application to Leader-Follower Formations of Multirotor UAVs. *IEEE Rob. Autom. Lett.*, 4(3):2637–2644, 2019.

[27] T. Lakemann, D. Bonilla Licea, V. Walter, T. Báča, and M. Saska. Towards Agile Multi-Robot Systems in the Real World: Fast Onboard Tracking of Active Blinking Markers for Relative Localization. *Robot. Autonom. Syst.*, 194:105175, 2025.