

Mutual Localization of UAVs based on Blinking Ultraviolet Markers and 3D Time-Position Hough Transform

Viktor Walter¹, Nicolas Staub^{1,2}, Martin Saska¹ and Antonio Franchi²

Abstract—A novel vision-based approach for indoor/outdoor mutual localization on *Unmanned Aerial Vehicles* (UAVs) with low computational requirements and without external infrastructure is proposed in this paper. The proposed solution exploits the low natural emissions in the near-*Ultra-Violet* (UV) spectrum to avoid major drawbacks of the visible spectrum. Such approach provides much better reliability while being less computationally intensive. Working in near-UV requires active markers, which can be leveraged by enriching the information content through blinking patterns encoded marker-ID. In order to track the markers motion and identify their blinking frequency, we propose an innovative use of three dimensional Hough Transform, applied to stored position-time points. The proposed method was intensively tested onboard multi-UAV systems in real-world scenarios that are very challenging for visible-spectrum methods. The results of our methods in terms of robustness, reliability and precision, as well as the low requirement on the system deployment, predestine this method to be an enabling technology for using swarms of UAVs.

I. INTRODUCTION

The use of swarms of *Unmanned Aerial Vehicles* (UAVs) has extends significantly the capabilities of single UAV, allowing for tasks otherwise impossible for single robot due to payload, actuation or sensory limitations. Typically, small UAVs are used for their cost-effectiveness and commercial availability and they can safely compose compact multi-UAV systems with small relative mutual distances. This raises the importance of mutual relative localization, in order to maintain safety distances, enforce the desired flocking behavior or the decentralized bio-inspired swarm stabilization [1], [2].

A typical challenge of mutual localization for aerial swarms is to present a low-cost infrastructure-independent solution, suitable for both indoor and outdoor settings and reasonable mutual distances. Literature is rich in approaches relaxing these requirements, like indoor work conducted with *Motion Capture systems* (MoCap), e.g. [3], [4], or *Infrared* blinking markers, coupled with an *event-based* ground camera [5] and outdoor setups relying on *Global Navigation Satellite System* (GNSS) [6], [7]. These solutions provide precise mutual localization information (≤ 1 cm, considering *RTK-GNSS*) with the major drawback of requiring pre-installed infrastructures, limiting the usage to known, uncluttered, and easily-accessible environments. Additionally, they are costly and tend to rely on intensive radio-communication

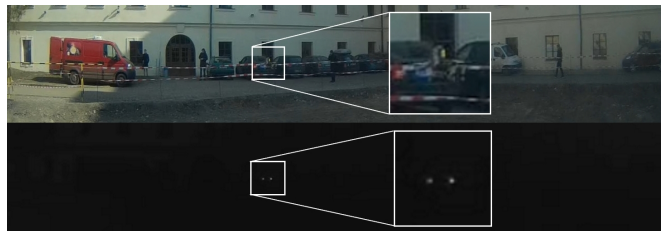


Fig. 1: Far away UAV against urban background in the shade. Barely noticeable in visible spectrum but obvious in UV spectrum.

between the swarm members, which is subject to limited range, interferences and does not scale up for large swarms.

Typical solutions to these issues are visible-spectrum vision-based approaches [8]. In an indoor-only setup, color-based markers can be used, see [9], [10], that are easy to segment under controlled lighting conditions, but not in the extremely unpredictable lighting conditions and the multi-colored outdoor environment. For outdoors, black and white markers are preferred, leading to solutions which combine passive markers and object detection, see [11], [12], used for swarms [2], [13] and heterogeneous groups of robots [14], [15], [16]. The drawbacks of these approaches are the need for large markers, computational complexity and sensitivity to lighting conditions.

Our solution extends our previous research on a novel, vision-based mutual localization in the Ultra-Violet (UV) spectrum [17]. Motivated by the low amount of near-UV radiation in sunlight and most artificial sources, compared to the visible spectrum. The technology uses active UV markers and standard cameras with UV band-pass filters, allowing for fast detection of markers in complex environments.

In order to retrieve the orientation or identity (ID) information, we encoded individual marker ID in blinking patterns. These are retrieved using an unprecedented application of 3D Time-Position Hough Transform. Indeed, the algorithm presented is the first exploiting the Hough Transform for tracking of objects in time. Since, in this case, the precision of the shape fitting is less relevant than the computational speed, this is an exemplary use-case for such algorithm.

The rest of the paper is structured as follows. Sec. II introduces the theoretical background necessary for the proposed algorithm presented in Sec. III. Finally, Sec. IV summarizes the results of our experimental proof-of-concept.

II. THEORETICAL BACKGROUND

A. UV spectrum: properties and motivations

For the sake of brevity, only key properties for our approach are presented, more details are presented in [17].

¹ CTU in Prague, FEE, Department of Cybernetics, Czech Republic {viktor.walter|martin.saska}@fel.cvut.cz

² LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France {nicolas.staub|antonio.franchi}@laas.fr

This research was partially supported by the ANR, Project ANR-17-CE33-0007 MuRoPhen, by CTU grant no. SGS17/187/OHK3/3T/13, by the Grant Agency of the Czech Republic under grant no. 17-16900Y and by project no. DG18P02OVV069 in program NAKI II

The solar spectrum approximates the *black-body radiation model* and has its peak intensity in the visible spectrum [18], while UV radiation is significantly less intense even close to the visible spectrum. This can be leveraged using affordable near-UV LEDs and suitable band-pass filters applied on a monochromatic camera. Tests have shown that the UV radiation refracted through the atmosphere and reflected from matte surfaces can be neglected, see [17]. Artificial radiation sources with strong UV emissions are rare, making this wavelength range very attractive for our application.

The image spots caused by UV LEDs sources are saturated at any time of the day, while the only other bright features are the sun or its specular reflections. Such images can be binarized by *static thresholding*, as opposed to the more computationally intensive *adaptive thresholding* needed for the visible spectrum. This allows for computationally simple detection of markers in the UV-range, robust to outdoor lighting conditions¹.

Following extensive testing in the wavelengths range of commercial UV LEDs. Best results are achieved with near-UV wavelength of 395 nm, compatible with widely available fisheye lenses with a large *Field of View* (FoV).

B. 3D time-position parametrization

With the proposed near-UV vision system, active markers appear in the camera image as white spots against a nearly-black background, see Fig. 1. While easy to locate, they are individually anonymous. In order to enrich their information content we devised blinking patterns encoding ID information (single marker ID or ID common to markers on one side of a UAV). While a non-blinking marker can be tracked among two consecutive camera frames using the nearest distance between two frames, this is not possible for blinking markers which are only visible in their *on-frames*, and periodically disappear in their *off-frames*. The only impact of this addition is a decrease of the *admissible flight dynamics*. On the other hand the introduction of blinking patterns does not only provide additional information, but also increases robustness by allowing 1) to easily filter out the sun and its reflections (non-blinking bright spots) and 2) to detect aliasing or occlusion, if the detected blinking frequency is not among the set of given patterns.

Identification and tracking of these blinking markers calls for a fast algorithm able to accommodate their periodical disappearance, which will use the following time-position parametrization. First, consider that markers observed by a camera are defined by their x - y coordinates in the image plane. These can be stored in an accumulator along time, with t -indexes such that the latest camera image corresponds to $t = 0$. We refer to triplets (x, y, t) as t -points which correspond to observations. Then for a set of markers with limited physical dynamics, their t -points will lie along smooth curves w.r.t. time, intermittent for blinking markers, see Fig. 2-a.

We chose to approximate those curves, around $t = 0$, by lines, see Fig. 2-b. We refer to these lines as t -lines. The time window used for this approximation impacts both the admissible flight dynamics and the range of usable blinking

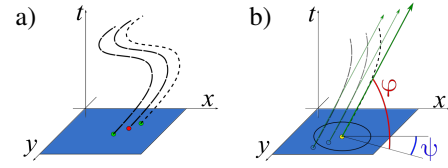


Fig. 2: Basic assumptions for the proposed system. (a) moving points (green) in present camera images (blue plane) follow smooth trajectories w.r.t. time. Due to blinking of the markers these curves are intermittent, making some of the points temporarily invisible (red). (b) considering short enough time-span these curves can be approximated by lines and such lines can be parametrized by their origin-point (yellow) and their pitch φ and yaw ψ .

frequencies. The t -lines can be parametrized by their *origin-point* x - y coordinates and two angles that we call pitch φ and yaw ψ , see Fig. 2-b. The origin-points are the points located in the image plane where the t -lines intersect. In on-frames they coincide with a t -point of $t = 0$, otherwise they are retrieved via Hough Transform, as detailed in Sec. III, and correspond to the theoretical marker position along the curve. The pitch and yaw map to the image speed of the tracked marker, and to the direction of its motion respectively.

Clearly the blinking frequency of a tracked marker can be retrieved via the t -points spacing along a t -line, the construction of which relies on a Hough Transform, see Sec. III. An additional benefit of considering t -lines is that markers fixed to a translating rigid body will have parallel t -lines, thus allowing for association of markers and objects.

C. Hough Transform

Hough Transform is a well-known method used to retrieve regular geometric forms described by a set of parameters. An example application is fitting a line over a set of collinear points. This is done by projecting the points into *image matrix* in the form of curves representing the range of possible lines passing through this point in terms of their parameters. The pixels of the temporary image matrix (Hough space) are incremented along each of these curves. Such projections of a number of collinear points will intersect in the Hough space, creating a local maximum in value representing the most likely parameters of a line common to all the original points. For general 3D line fitting scenarios with reasonable precision the Hough Transform will require a dense discretization of the parameter space consisting of at least four parameters [19]. This means searching for local maxima in a large 4D space which is not feasible for UAV embedded solutions. Instead we use a more purpose-fitted implementation of Hough Transform, relaxing t -lines reconstruction to guarantee a good enough approximation in order to reliably separate adjacent markers. As we aim at a good enough approximation, the discretization steps of the t -line parameters, $\Delta\varphi$ and $\Delta\psi$, can be chosen large enough to enforce robustness against small errors in the origin-point coordinates arising from the camera image pre-processing. Moreover the set of possible t -lines is constrained by the physics of the system, allowing to reduce the size of the Hough space.

¹mrs.felk.cvut.cz/uvddl

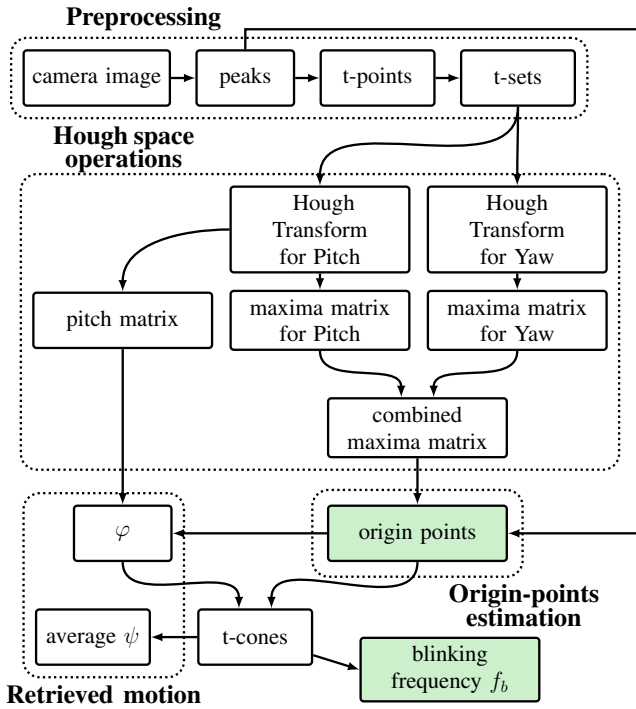


Fig. 3: Flow chart of our tracking algorithm based on Hough Transform. Both origin-point and blinking frequencies are retrieved, by-products are the full t-line parameters.

III. ALGORITHM FOR ORIGIN-POINT POSITION AND BLINKING FREQUENCY TRACKING

This section details the proposed algorithm to retrieve the origin-points and their blinking frequencies from the camera grey-scale image. The overall flow is summarized in Fig. 3.

A. Base algorithm

1) *Image pre-processing*: The pre-processing of the grey-scale image starts by the detection of bright spots, which are used to construct t-points, see Sec. II-B. The t-points are stored in a set, $\mathcal{U} \subset \mathbb{N}^3$. These t-points can be interpreted as points in a bounded 3D space of height F corresponding to the highest t-index, *i.e.* the time window of \mathcal{U} . The set of t-points is updated with each new camera image. We are interested in a way to retrieve for each t-line both the origin-point and the blinking frequency along the t-line.

2) *Hough Space Operations*: A direct approach would consist in applying 4D Hough Transform on \mathcal{U} directly, which proves to be computationally expensive and cumbersome. Therefore, our algorithm is based on two simpler 3D Hough voxel spaces, considering origin-point coordinates combined with pitch and yaw separately. To reduce the search space for the construction of the Hough Transform, the pitch and yaw are discretized such that

$$\psi_j = j \frac{2\pi}{\Delta\psi} \mid j \in \left[\frac{\underline{\psi}}{\Delta\psi}; \frac{\bar{\psi}}{\Delta\psi} \right] \subset \mathbb{N} \quad \text{and} \quad \varphi_i = i \frac{\pi}{2\Delta\varphi} \mid i \in \left[\frac{\underline{\varphi}}{\Delta\varphi}; \frac{\bar{\varphi}}{\Delta\varphi} \right] \subset \mathbb{N},$$

where the range limits and discretization steps are parameters of our algorithm, listed in Tab. I. Our specialized Hough Transform translates t-points into their images in the form of voxelated surfaces in the two aforementioned voxel spaces. If multiple t-points belong to the same marker, their images



Fig. 4: Erroneous peaks (red) in Hough Transform for pitch (left) and yaw (center), are suppressed by element-wise multiplication.

in the Hough spaces will intersect at voxels corresponding to parameters of the t-line on which they lie.

To find the t-line parameters one needs to find local maxima in a 3D voxel space, which is computationally complex. We use the fact that since the curves followed by the t-points are non-retreating w.r.t. time and since the markers are attached around non-transparent objects, it is physically impossible for multiple t-lines to share the same origin-point. With this assumption, we can simplify the search for local maxima into 2D, the Hough spaces are flattened into so-called *maxima matrices*. This operation is done by assigning to each $[x, y]$ element of the maxima matrix the highest value among voxels of the Hough space with the same x, y coordinates. This results in easier detection for origin-coordinates at the expense of an information loss about the associated angle parameters. To keep this information easily accessible, a second matrix, so-called *angle matrix*, is constructed during the flattening process, which stores the angle value (respectively ψ or φ) corresponding of the maxima in the Hough space for each x, y coordinates.

3) *Origin-point retrieving*: From there origin-points can be retrieved as peaks in the maxima matrices. However some aliasing phenomena, as well as ambiguity-based artifacts may be found in the maxima matrix. For pitch, erroneous peaks in between two slow moving markers can appear. For yaw, two different kinds of erroneous peaks can appear; some corresponding to opposite yaw, as well as peaks perpendicular to the connecting lines of neighboring markers due to discretization. To increase robustness against these erroneous peaks, the two maxima matrices are multiplied element-wise, which leads to the suppression of the erroneous peaks as they are not likely to be present in both spaces.

The origin-point coordinates correspond to peaks in the combined maxima matrix, and are retrieved in a two-step method; 1) t-point of index $t=0$ are collected as their coordinate are more reliable than the estimated one, then 2) peaks in the combined maxima matrix are collected. After each located peak, its surroundings are nullified in the combined maxima matrix, allowing for finding further peaks. During the peak search we also consider the number of expected origin-points, L , once L peaks are found the search can stop. The policy to define L can be based on considering: 1) the knowledge of visible origin-points based on the number of UAVs and the average number of visible markers, 2) the maximum number of markers seen simultaneously within the last F frames or 3) any other heuristic.

At this stage of the algorithm, the t-line origin-points are retrieved, *i.e.* the image positions of markers both in on- and off-frames. From the respective angle matrices, it is possible to retrieve the two other t-line parameters estimates.

4) *Blinking frequency retrieval*: As we decided to encode additional information into blinking patterns further processing is needed to retrieve them. A possible way to do so

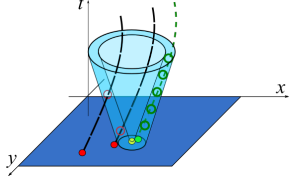


Fig. 5: Cone shell defined by the estimated t-line pitch. The t-lines with distant origin-point (red filled) will intersect the expanded cone shell in few points (hollow red), while t-lines with origin-point (green filled) nearby to the center of the t-cone (yellow) will intersect in most of their points (hollow green). This suppresses their influence on the estimated frequency and yaw.

is to cluster all t-points in the vicinity of the identified t-lines to find their average blinking frequency. To reduce computational requirements induced by exploring two Hough spaces of small granularity, we choose $\Delta\varphi \ll \Delta\psi$ and forgo an angle matrix for ψ . This only gives a reliable estimate of the pitch parameter, while the ψ maxima matrix is still rich enough to reject aliasing. Instead, we now consider t-cones, which are generated by t-lines rotated around their t -axis passing through the origin point, see Fig. 5, and their vicinity to retrieve both blinking frequency and t-line yaw, by averaging the values of the corresponding t-points. The vicinity r_v is a tunable parameter defining the maximum distance from the t-cone where we look for t-points. Note that this method is likely to collect more t-points which are not a part of the desired t-line in the averaging process, see Fig. 5. Nevertheless they are, in practice, outnumbered by the t-points corresponding to the desired t-line. Lastly, the origin-point with blinking frequency under a certain (low) threshold can be disregarded as being the sun or its specular reflections, which are the most significant contaminant of the camera image in the UV spectrum.

This concludes the algorithm as both origin-points and their associated blinking frequency are retrieved, along with the other t-lines parameters.

B. Improvements

In order to increase the robustness to high flight dynamics and computational efficiency, we have designed two refinements of our algorithm.

1) *Weighted Hough space*: First, to increase the admissible dynamics of the tracked markers we propose to introduce weight in the construction of the Hough space. Instead of giving equivalent weight for all t-indices, we introduce the following weighting function,

$$w(t) = \lambda(F - t) + F \in \mathbb{N},$$

where λ is a parameter regulating the weight ratio between the newest t-points, and the oldest ones in \mathcal{U} . In this way the more recent t-points affect the t-lines parameters more, making our algorithm more resilient to abrupt changes of direction, implied by highly dynamic flight. This refinement leads directly to better origin-point estimate.

2) *Pre-computed masks*: The second refinement reduces the computational complexity of the Hough space construction, by applying pre-computed masks to generate it. The constructed masks resemble hollow cones for the pitch

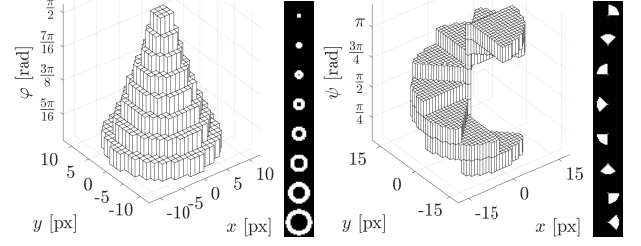


Fig. 6: A mask used in Hough space for pitch, generated for t-points with $t = 10$ (left) and a mask for yaw, generated for t-points with $t = 15$ (right). Side bitmaps show the corresponding mask slices.

Hough space, see Fig. 6-left, and spiral staircases for the yaw Hough space, see Fig. 6-right. These shapes can be explained intuitively. We observed that the possible origin-points of all t-lines passing through a t-point, can be easily expressed w.r.t. the t-line parameters, φ and ψ , for a given t-index t . The potential origin-point for this t-point directly underneath has the parameter $\varphi = \frac{\pi}{2}$ and as the distance of the potential origin-points increases, the corresponding pitch decreases, which leads to a cone shape in the Hough space. Such cones are of the same shape for a given t , while the x and y parameters of the t-point merely shift it to the respective x - y position. Similar reasoning explains the mask shape for the yaw Hough space. In order to prevent discontinuities in the masks, which arise from the angle discretization, we introduce overlap parameters, $\delta\varphi$ and $\delta\psi$.

To construct the Hough space, at each t-point in \mathcal{U} , the introduced masks are used as follow: 1) retrieve the mask associated with the t-index 2) retrieve the x - y coordinates of the t-point and then 3) apply the mask at the coordinates by increasing the corresponding voxel values in the Hough space. This considerably speeds up the construction of the Hough space as instead of calculating all t-lines passing through each t-point we apply a static, pre-computed, mask.

IV. EXPERIMENTAL VALIDATION

Experimental parameters are grouped in Tab. I. The chosen parameters influence the maximal admissible flight dynamics, in our case the maximum linear image speed for a marker is 144 px/s, which translates to a maximum speed perpendicular to the camera axis of respectively 0.6 and 3 m s⁻¹, for marker-camera distance of 1 and 5 m.

The innovative part of the system, the UV active markers and camera, are introduced and discussed in [17]. They can be easily fitted to any UAV platform, which we demonstrate by using a standard hexarotor for outdoor experiments and a standard quadrotor for the indoor ones. Indoor experiment used motion capture system (MoCap) as ground truth.

In the testing phase, our algorithm runs real-time off-board on an Intel NUC 7 (4 cores, 2.6 GHz), a classical embedded computer for UAVs. The prototypical MATLAB implementation loads a single core up to 48.8% on average, validating that our approach can be easily embedded. Indeed there are enough resources left to run typical mission control and planning algorithms. The final C code implementation is expected to run significantly faster. The experiments were recorded on video (mrs.felk.cvut.cz/uvdd2).

Symbol	Meaning	Impact on	Experimental Value [unit]
F	time window size of \mathcal{U}	admissible dynamics	$0.3f$
$\Delta\varphi$	φ discretization step in Hough Transform	robustness to dynamic motion; selectivity; computational complexity	$\pi/64$ rad
$\Delta\psi$	ψ discretization step in Hough Transform	computational load	$\pi/4$ rad
$\delta\varphi$	mask overlap parameter	robustness; overall selectivity	$\pi/68$ rad
φ	upper limit of the pitch value	minimal expected movement	$\pi/2$ rad
$\bar{\varphi}$	lower limit of the pitch value	admissible dynamics; marker selectivity (via $\Delta\varphi$)	$\pi/4$ rad
L	number of expected peaks	number of markers; computational complexity	3 (o) 10 (i)
λ	weighting factor	admissible dynamics	1
r_v	inspected vicinity of t-cones	robustness; marker selectivity	3 px
f	camera frame-rate	maximal blinking frequency \bar{f}_b	72 Hz
t_x	camera exposure rate	selectivity w.r.t. ambient radiation; effective distance range	1000 μ s
$[f_b; \bar{f}_b]$	blinking frequency range	number of ID encodable	3.34 – 40 Hz

TABLE I: Main parameters of the proposed solution, when needed (o) and (i) denotes outdoor and indoor parameters respectively.

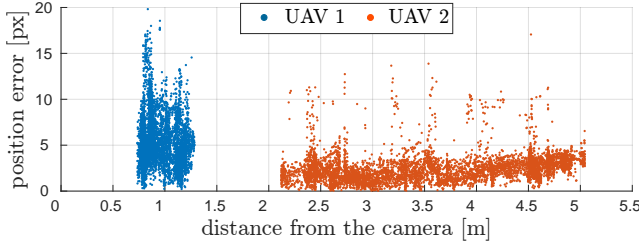


Fig. 7: Indoor precision testing with MoCap as ground truth, the position error in the camera horizontal direction against the distance to the camera is assessed. Marker close have larger (but reasonable) error, due to larger spot in image, making the exact pixel position of the marker ambiguous.

A. Indoor validation against ground truth

To evaluate the accuracy, two quadrotor UAVs are flown in front of a fixed camera, markers are located on the UAV arms, as described in [17]. To identify each UAV, blinking patterns are assigned such that each UAV has two dedicated frequencies, one for front and one for back markers.

In the experiment the *UAV 1* was hovering within 1 m of the camera while performing yaw rotation motion, with blinking frequencies of 6 and 10 Hz. The *UAV 2* was following a zig-zag like trajectory from 5 to 2 m toward the camera, with blinking frequencies of 15 and 30 Hz. The UAVs motion are constrained by the limited size of the flying arena, forbidding tests on distance longer than 5 m.

MoCap information is translated to camera image and MoCap-based image positions are paired with the closest estimated positions of origin points. The position error against the camera-marker distance is used to assess performances, see Fig. 7. Markers close to the camera suffer an error of at most 20 px, while for further away markers the error is mostly below 5 px. This correlates with the size of the bright spots in the image that make t-point detection less precise.

The results were additionally evaluated w.r.t. the blinking frequency of the markers, see Fig. 8. It appears that the blinking frequencies have not detectable influence on the position error, as the trend for each individual frequency follows the aggregated values for their respective UAV.

B. Outdoor validations and characterization

Additional outdoor experiments were conducted to assess our approach performances in operational conditions. Experiments were conducted around noon by clear weather and consisted of flights of two hexarotor UAVs, one equipped

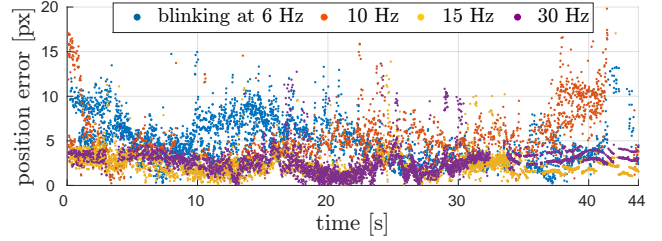


Fig. 8: Impact of blinking frequency on the position error, indoor.

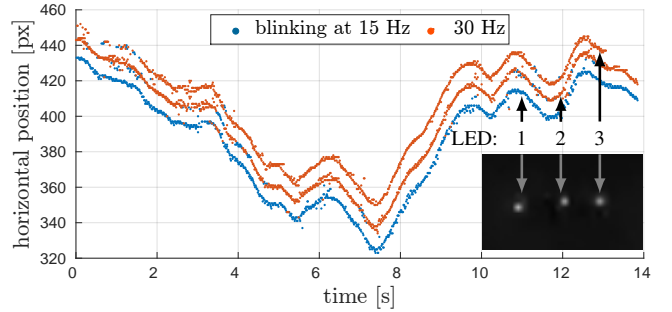


Fig. 9: Example of outdoor tracking for UAV horizontal motion (the most agile). Near the instant $t = 13$ s the UAV was rotated, so that only two markers remained visible.

with a camera and one equipped with markers following [17]. The markers were set to blink with two distinct frequencies, such that two triplets on adjacent hexarotor arms shared frequency. Blinking frequencies were set at 15 and 30 Hz, for back and front respectively. The tracking results presented in Fig. 9 show the good performance of the proposed approach under outdoor light conditions.

In particular, our algorithm was able to keep track of the IDs of the markers encoded in their blinking frequency, while still providing accurate image position estimation even in between the on-frame t-points.

C. Blinking frequency estimation

Both indoor and outdoor data are used to assess the performances of the frequency estimation, as the generated frequency are known, they are compared in Fig. 10. By filtering out the obvious outliers, we compute the average error for all points close to a given frequency. Performances of the blinking frequency estimation, both indoor and outdoor, are good with mean absolute error (MEA) below 3.9%, 2.2%, 3.8% and 3.1% for respectively 6, 10, 15 and 30 Hz blinking frequency. This also demonstrates that the blinking frequency

Method	Resolution	FoV	Range	ID count	Environment	Properties
Color circles [20]	752 × 480	125°	N/A	3+	indoor, well illuminated	large marker size, lighting sensitive
WhyCon [11]	752 × 480	42°	5.5 m	1	indoor/outdoor, illuminated	large marker size (≤ 18 cm)
ALM-DVS [5]	128 × 128	65°	N/A	3+	indoor	requires event-based camera
CNN-YOLO [21]	1280 × 720	132°	15 m	N/A	indoor/outdoor, illuminated	high computational load, marker independent
Proposed approach	752 × 480	180°	15 m	6+	indoor/outdoor	small markers, low computational intensity

TABLE II: Performance comparison with various representative methods.

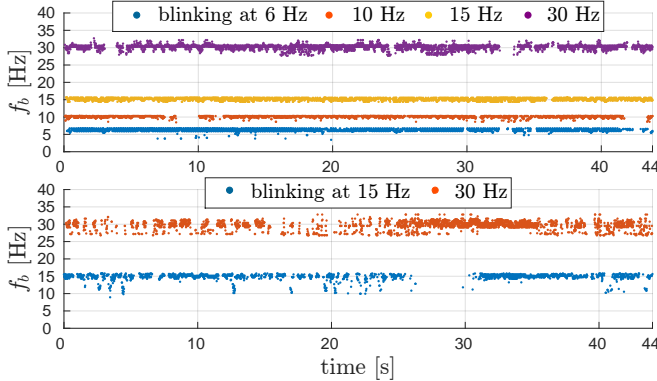


Fig. 10: Estimated Frequencies evolution, in both indoor (top) and outdoor (bottom) experiments. In both cases the frequency estimation performs with a MEA below 3.9%.

estimation performs similarly for all frequencies well inside the admissible blinking frequency range.

D. Comparison with other methods

Our proposed methods is significantly more versatile than state-of-the-art methods, allowing usage both indoor and outdoor without illumination requirements, see Tab.II. Moreover our approach has a low impact in terms of computational power and works with small markers. For each method the precision and range can be tuned by selecting different resolution and FoV, therefore there are no standard metrics to evaluate them. Despite the combination of small resolution and large FoV in our experiments, the performances are comparable or exceeding those of other methods.

V. CONCLUSION

In this paper, we proposed a novel system for outdoor and indoor mutual relative localization using active UV LED markers. It enables significantly better performances in comparison with state-of-the-art methods of UAV mutual localization. Additionally, we have shown how active markers can be leveraged to encode additional information via blinking patterns. Our approach relies on 3D time-position Hough Transform and has been tested in active UAV deployment both indoor and outdoor. Results from outdoor experiments show excellent detection reliability w.r.t. backgrounds such as the sky, trees or even buildings, while still being able to decode the blinking signal. The theoretical predictions, as well as the experimental data presented here, show a lot of promise for deployment in swarm robotics and multi-robot systems in general.

REFERENCES

[1] D. Brandtner and M. Saska, "Coherent swarming of unmanned micro aerial vehicles with minimum computational and communication requirements," in *European Conference on Mobile Robots (ECMR)*, 2017.

[2] M. Saska, "Mav-swarms: Unmanned aerial vehicles stabilized along a given path using onboard relative localization," in *International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2015.

[3] A. Kushleyev, D. Mellinger, C. Powers, and V. Kumar, "Towards a swarm of agile micro quadrotors," *Autonomous Robots*, vol. 35, no. 4, 2013.

[4] P. Stegagno, M. Cagnetti, G. Oriolo, H. H. Blthoff, and A. Franchi, "Ground and aerial mutual localization using anonymous relative-bearing measurements," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1133–1151, Oct 2016.

[5] 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 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[6] G. Vásárhelyi, C. Virágh, G. Somorjai, N. Tarcai, T. Szörényi, T. Nepusz, and T. Vicsek, "Outdoor flocking and formation flight with autonomous aerial robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014. IEEE, 2014.

[7] C. Virágh, G. Vásárhelyi, N. Tarcai, T. Szörényi, G. Somorjai, T. Nepusz, and T. Vicsek, "Flocking algorithm for autonomous flying robots," *Bioinspiration & biomimetics*, vol. 9, no. 2, 2014.

[8] M. Nagy, Z. Akos, D. Biro, and T. Vicsek, "Hierarchical group dynamics in pigeon flocks," *Nature*, vol. 464, no. 7290, p. 890, 2010.

[9] R. Tron, J. Thomas, G. Loianno, J. Polin, V. Kumar, and K. Daniilidis, "Vision-based formation control of aerial vehicles," in *Robotics: Science and Systems*, 2014.

[10] I. Rekleitis, P. Babin, A. DePriest, S. Das, O. Falardeau, O. Dugas, and P. Giguere, "Experiments in quadrotor formation flying using on-board relative localization (technical report)."

[11] J. Faigl, T. Krajník, J. Chudoba, L. Preucil, and M. Saska, "Low-cost embedded system for relative localization in robotic swarms," in *IEEE ICRA*, 2013.

[12] T. Krajník, M. Nitsche, J. Faigl, P. Vaněk, M. Saska, L. Přeučil, T. Duckett, and M. Mejail, "A practical multirobot localization system," *Journal of Intelligent & Robotic Systems*, vol. 76, no. 3-4, pp. 539–562, 2014.

[13] M. Saska, T. Báča, J. Thomas, J. Chudoba, L. Preucil, T. Krajník, J. Faigl, G. Loianno, and V. Kumar, "System for deployment of groups of unmanned micro aerial vehicles in gps-denied environments using onboard visual relative localization," *Autonomous Robots*, vol. 41, no. 4, pp. 919–944, 2017.

[14] M. Saska, V. Vonasek, T. Krajník, and L. Preucil, "Coordination and Navigation of Heterogeneous UAVs-UGVs Teams Localized by a Hawk-Eye Approach," *International Journal of Robotics Research*, vol. 33, no. 10, pp. 1393–1412, 2014.

[15] K. Boudjit and C. Larbes, "Detection and target tracking with a quadrotor using fuzzy logic," in *8th International Conference on Modelling, Identification and Control (ICMIC)*, 2016.

[16] V. Dhiman, J. Ryde, and J. J. Corso, "Mutual localization: Two camera relative 6-dof pose estimation from reciprocal fiducial observation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[17] V. Walter, M. Saska, and A. Franchi, "Fast mutual relative localization of uavs using ultraviolet led markers," in *2018 International Conference of Unmanned Aircraft System (ICUAS)*, 2018, accepted to ICUAS 2018.

[18] M. Iqbal, *An Introduction To Solar Radiation, Chapter 3: The Solar Constant and Its Spectral Distribution*, 1983.

[19] C. Dalitz, T. Schramke, and M. Jeltsch, "Iterative hough transform for line detection in 3d point clouds," *Image Processing On Line*, vol. 7, pp. 184–196, 2017.

[20] R. Tron, J. Thomas, G. Loianno, K. Daniilidis, and V. Kumar, "A distributed optimization framework for localization and formation control: applications to vision-based measurements," *IEEE Control Systems*, vol. 36, no. 4, pp. 22–44, 2016.

[21] M. Vrba, "Relative localization of helicopters from an onboard-camera image using neural networks," Master's thesis, Czech Technical University, Prague, 2018.