ArTuga: A novel multimodal fiducial marker for aerial robotics

Rafael Marques Claro *, Diogo Brandão Silva, Andry Maykol Pinto

DEEC FEUP - Department of Electrical and Computer Engineering, Faculty of Engineering, University of Porto, FEUP Campus, Rua Dr. Roberto Frias, 4200 - 465, Porto, Portugal
INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, FEUP Campus, Rua Dr. Roberto Frias, 4200 - 465, Porto, Portugal

ABSTRACT

For Vertical Take-Off and Landing Unmanned Aerial Vehicles (VTOL UAVs) to operate autonomously and effectively, it is mandatory to endow them with precise landing abilities. The UAV has to be able to detect the landing target and to perform the landing maneuver without compromising its own safety and the integrity of its surroundings. However, current UAVs do not present the required robustness and reliability for precise landing in highly demanding scenarios, particularly due to their inadequacy to perform accordingly under challenging lighting and weather conditions, including in day and night operations.

This work proposes a multimodal fiducial marker, named ArTuga (Augmented Reality Tag for Unmanned vision-Guided Aircraft), capable of being detected by an heterogeneous perception system for accurate and precise landing in challenging environments and daylight conditions. This research combines photometric and radiometric information by proposing a real-time multimodal fusion technique that ensures a robust and reliable detection of the landing target in severe environments.

Experimental results using a real multicopter UAV show that the system was able to detect the proposed marker in adverse conditions (such as at different heights, with intense sunlight and in dark environments). The obtained average accuracy for position estimation at 1 m height was of 0.0060 m with a standard deviation of 0.0003 m. Precise landing tests obtained an average deviation of 0.027 m from the proposed marker, with a standard deviation of 0.026 m. These results demonstrate the relevance of the proposed system for the precise landing in adverse conditions, such as in day and night operations with harsh weather conditions.

1. Introduction

Multicopter Unmanned Aerial Vehicles (UAVs) are autonomous aerial robots that have been used in several applications due to the ability to maintain a very stable flight and to approach not easily accessible areas, which results in the collection of high quality data. The main application areas of UAVs include several industries, such as agriculture [1,2], surveillance (including fire prevention) [3,4], search and rescue operations [5], and also the cinematographic and military industries [6]. Furthermore, several inspection tasks are being carried out by UAVs [7,8], representing a reduction of the costs and of the total time consumed [6,9]. The usual cycle of a UAV mission includes its takeoff from a base station, the execution of its task and the returning to the launching area. An important feature that allows this autonomous operation is the capability of the UAV to land with precision on its target, which requires centimetric accuracy. Fig. 1 depicts a UAV data acquisition cycle that follows the aforementioned description.

A crucial task for enabling a VTOL (Vertical Take-Off and Landing) UAV to successfully land on the desired landing area is its ability to autonomously detect and recognize the landing target in real-time. If the UAV misses the landing area it can damage its aerial equipment, as well as compromise the execution of the following missions. The landing accuracy of 1–3 m of traditional GPS-based methods does not satisfy the high precision needed for several UAV applications [11] that require a centimeter level accuracy, due to the restricted size of landing areas, such as for package delivery or boat landing. In addition, several operations are conducted in GPS-denied areas. Nevertheless, the performance of traditional visual approaches do not feature the required robustness and precision to operate in challenging weather and lighting conditions, including in day and night operations. A perception system that collects heterogeneous data (with both 2D and 3D information) improves the navigation skills of the vehicle by creating a more complete and reliable map of the surrounding scene [12–15], enhancing its understanding in real time [16–19].

* Corresponding author.
E-mail addresses: rafael.marques@inesctec.pt (R.M. Claro), diogo.a.silva@inesctec.pt (D.B. Silva), amgp@fe.up.pt (A.M. Pinto).
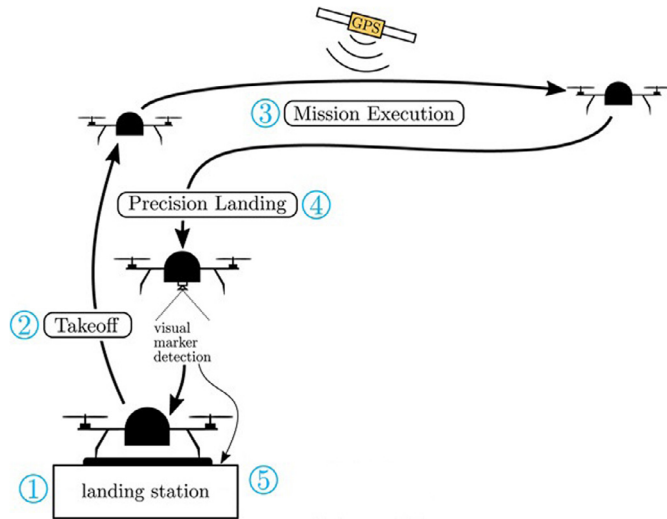
Fig. 1. Autonomous UAV data acquisition cycle [10]: (1) vehicle on the landing station to start the operation; (2) takeoff; (3) mission execution; (4) landing station detection and precision landing; (5) landed.

Therefore, this work proposes a multimodal-based perception system for landing VTOL UAVs with high accuracy and precision, capable of operating under severe environment conditions. As a result, the contributions of this work are the following:

1. A multimodal fiducial marker, named ArTuga,[1] whose pose can be extracted in real-time by analyzing thermal, visual and point cloud data. This active marker enables relative positioning in challenging environments;
2. A heterogeneous perception system, named TriOPS, that collects both photometric and radiometric data, by resorting to multiple cameras and range sensors. This perception system improves the situational awareness of the vehicle, increasing its navigation and precise landing abilities;
3. A procedure to intrinsically and extrinsically calibrate multimodal perception systems, formed by visual, thermal and range sensors;
4. Extensive experiments for the error characterization in a controlled scene, for the robustness and redundancy analysis, and performing of multiple precise landing tests in a real scenario using a real VTOL UAV. Results demonstrate the complementarity and accuracy of the developed system.

The article is organized as follows: Section 2 presents a brief review of artificial markers for landing area detection in aerial robotics. Afterwards, Section 3 presents the developed multimodal system for precise landing. Experimental evaluation is demonstrated in Section 4, which includes experiments conducted in different testing conditions. Experimental considerations show that the multimodal fiducial marker and the heterogeneous perception system perform satisfactorily better and, therefore, the combination of ArTuga and TriOPS proposed in this research represents a robust, redundant and reliable landing target detection and precise landing solutions. Finally, Section 5 clarifies the most important conclusions of this research.

2. Related work

Several robotic applications resort to artificial markers to label positions of interest and to help the robot retrieve its relative

position. The most common types of artificial markers used include visual markers, infrared (IR) beacons, thermal markers and retroreflective markers.

To achieve precise landing of a VTOL UAV, it is necessary to accurately retrieve the relative pose between the vehicle and the detected landing area. Independently of the type of markers and of the sensors used, the relative pose estimation requires the awareness of the pose of the markers in a fixed coordinate frame. Usually, the knowledge of the real size of the marker is also necessary for the scale factor. With this information, combined with the correct association between the detected marker and its real pose, it is possible to estimate the relative distance from the UAV to the landing target. When a calibrated camera sensor is used, the task of calculating its position relative to a marker (3D reference points) and its projection toward the image plane (2D projection points) is known as the Perspective-n-Point (PnP) problem [20–22]. There are different solutions to solve this problem and whose performance depends on the type of markers used [23]. Each of these methods allows to estimate the transformation from the camera to the marker frames. The pose of the UAV is then calculated by knowing the transformation from the camera to the UAV frames, and from the marker to the landing target frames. Unlike cameras, using range sensors, such as 3D LiDAR, allows direct perception of the depth of the surrounding environment. This characteristic enables to easily calculate the relative pose between the UAV and a marker, which can be detected by analyzing the LiDAR's point cloud.

Many UAV systems resort to visual information to identify visual markers for a multiplicity of tasks, including for landing areas detection. Several studies have been performed in the area of computer vision for UAVs [24], which are able to increase the landing accuracy. The most common approaches use monochromatic markers placed on the landing target. Each marker presents a layout that encodes a unique identification by a binary code. They are generally based on regular geometric shapes (such as squares), making any of them distinguishable from any other marker. In robotics, some of the most popular fiducial markers are the ARTag [25], AprilTag [26,27] and ArUco [28] (depicted in Fig. 2). The quadrangular shape of these markers allows the extraction of the camera pose from the four corners, as long as the camera is properly calibrated [28]. In summary, these markers are detected by firstly extracting the edges of an image collected by a visual camera, followed by filtering the contours that make up a polygon with four vertices, and finally by extracting its respective binary code [28]. Circular shaped markers, such as the CCTag [29] and the STag [30], are also used in several robotics applications. Other examples of visual and monochromatic markers that combine both corner and circular shapes identification, used in different UAV applications for landing areas detection, are the RAPiD marker [31], the marker used in the MBZIRC challenge [32–35], and the marker presented by Nguyen et al. [36,37]. Although there are optimization techniques to improve the detectability of these markers [38], a major drawback is that they are dependent on lighting and environment conditions, only performing efficiently indoors or in daylight with favorable weather [39].

Another type of marker used for the detection of landing areas consists in the utilization of Light-Emitting Diodes (LEDs). Therefore, the UAV must be equipped with a sensor that is able to capture the light emitted by LEDs. The most common approaches propose the use of a beacon with IR LEDs and of an IR camera for the landing area detection [21,40,41]. The major drawback of these IR marker approaches is that they are only suitable for indoor operations, given its sensibility to the sunlight [24]. Another disadvantage is related with the short range of the beams [21,24] and with the limitation of the operation range that the angle of radiation implies [21,40], given that the tilting of the camera produces high localization errors [41].

(a) ARTag [25].   (b) AprilTag [26, 27].   (c) ArUco [28].
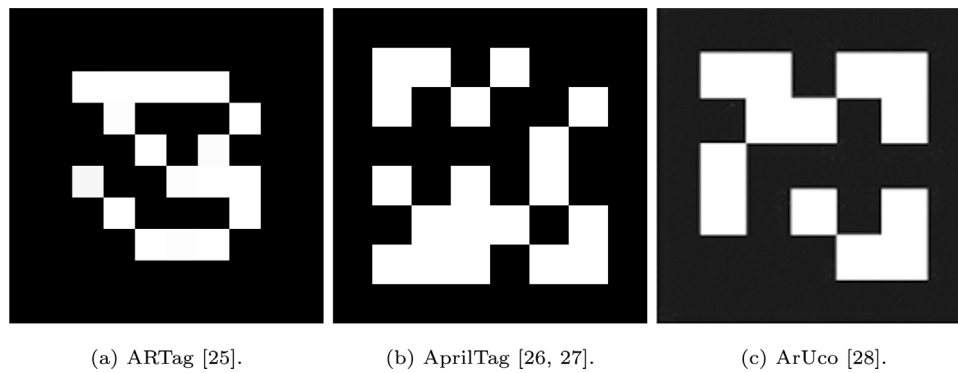
Fig. 2. Most used fiducial markers: ARTag, AprilTag and ArUco.

One possibility to overcome these limitations consists in the utilization of markers that present a considerable temperature difference compared with its background, allowing its recognition with a thermal camera sensor [42–45]. While an image collected by a visual camera degrades significantly in the presence of obstructions, such as fog and smoke [44], the use of a thermal camera to detect a heated or cooled marker is a solution suitable for all weather conditions [43]. Some examples of proposed solutions for thermal markers consist of electrically heated T-shaped markers [42,43], and of dark rubber markers that are heated by the sun radiation, contrasting with a reflective aluminum sheet [45]. Another approach consists of a marker made of low thermal conductivity material, cut to shape a traditional ArUco marker, that contrasts with the cooled or heated surroundings [44]. This method has the advantage of enabling the application of common detection algorithms used with visual cameras and fiducial markers, with the significant improvement of representing a more robust solution in low visibility conditions.

The previously mentioned techniques lean on camera sensors, meaning that they are not able to sense depth directly [39]. The use of 3D range sensors, such as a LiDAR (Light Detection And Ranging), can overcome this limitation. These 3D sensors are used in robotics to improve the quality of distinct tasks, such as object detection [46], scene reconstruction and navigation [47]. This sensor collects a point cloud of the environment and has a performance that is resilient against adverse weather and lighting conditions, representing an accurate and occlusion resilient solution [39]. A proposed approach uses a reflective marker whose layout is composed by a beacon with retro-reflective and matte areas [39]. This design allows to detect a reflective pattern with a 3D LiDAR, being easy to isolate it by applying a reflective intensity filter. Although the utilization of 3D LiDAR sensors for precise landing and landing area detection for UAV applications is practically non-existent, it represents a high potential and accurate solution [39].

State-of-the-art approaches do not present the required robustness for highly demanding autonomous UAV operations, in particular for the precise and accurate landing maneuver. By using a multimodal marker where each of its components complement each other, it is possible to increase its reliability and accuracy. The different sensors used to collect data from this marker, such as cameras and LiDARs, require proper calibration. Calibration boards used for visual camera-LiDAR extrinsic calibration vary from using a planar generic calibration pattern, such as a regular chessboard [48] or a charuco [49,50], to other custom made solutions [51]. Although these methods are quite precise, they expect a set of prerequisites to obtain a reliable calibration [52], such as of the calibration board's background and of the segmentation approach. Another solution consists in using a planar fiducial marker for the visual camera, complemented by

a retroreflective target for the LiDAR sensor. This solution makes it easier to extract the relevant points for the extrinsic calibration process of the LiDAR, by applying an intensity filter [52]. Nevertheless, none of the solutions previously presented also address the extrinsic calibration with thermal cameras, when combined with LiDARs and visual cameras.

3. Multimodal perception system for precise landing of UAVs
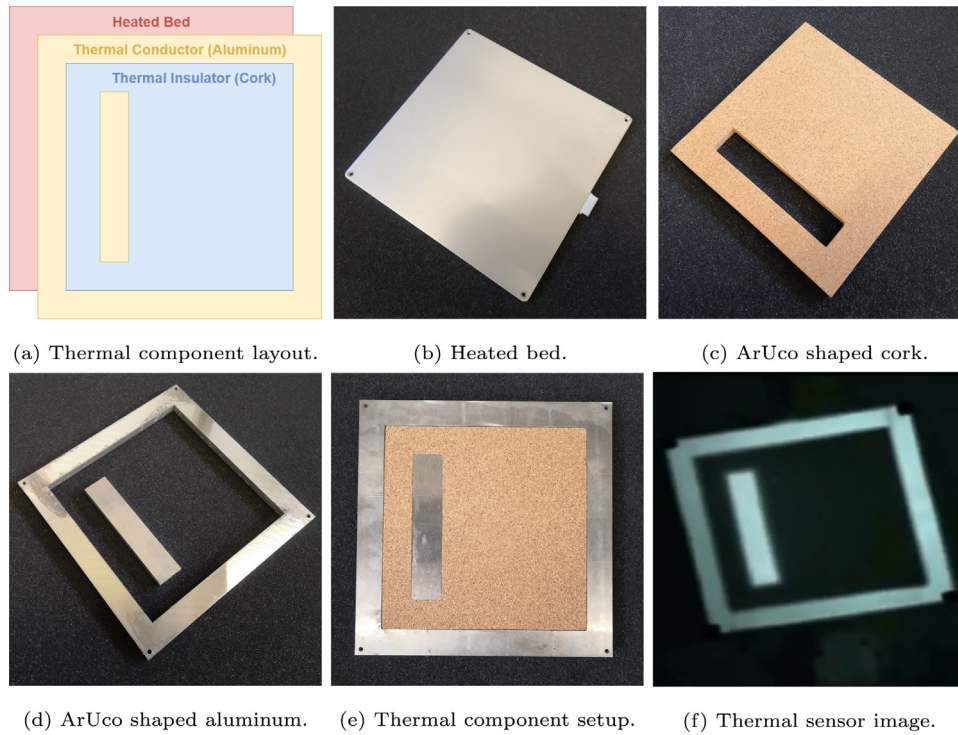
For autonomous UAV operations, the system has to operate in real-time and has to be robust and resilient against adverse lighting and weather conditions. Conventional solutions consist of unimodal methods with very limited applicability to daily operations in day and night activities. Thus, UAVs are being used in a limited range of scenarios under favorable weather and daytime conditions. A multimodal solution that resorts to photometric and radiometric data is able to provide a reliable detection of the landing target under severe environmental and lighting conditions, such as rain, fog and night operations.
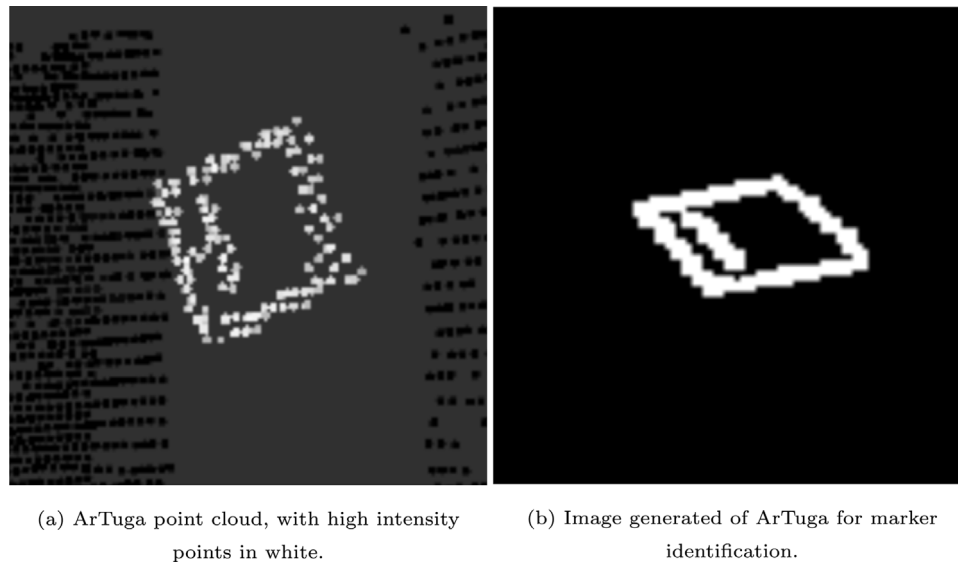
3.1. Multimodal fiducial marker

This research proposes ArTuga (Augmented Reality Tag for Unmanned vision-Guided Aircraft), which is an active fiducial marker that enables multimodal detections using heterogeneous sensors. This multimodal marker extends the concept of ArUco marker [28] by combining the advantages of photometric and radiometric detections conducted by visual, thermographic and LiDAR devices. Thus, the proposed marker aims to be detected and identified by multiple sensors whose performance is usually complementary under distinct weather conditions. The identification of ArTuga is secured by a distinguishable thermal, reflectivity-based and visual binary encoding, meaning that a single code is achieved by combining different material properties:

- low and high temperatures;
- low and high reflectivity indexes;
- light and dark colors (traditionally, black and white).

For the coding of the thermal component of this active marker, different temperatures have to be produced in its surface. To achieve this, Artuga resorts to a heated source and to two components with distinct thermal conductivity coefficients. For convenience, this paper assumes an electronically heated bed with controllable temperature up to 100 °C, that can be powered up by battery, solar panels, or electric plugs. The material selected for the good thermal conductor was aluminum (thermal conductivity of 239 W/(m·K)), while cork was selected as the good thermal insulator material (thermal conductivity of 0.045 W/(m·K)). Both materials are in contact with the heat source. However, the aluminum will create hot areas (close to 100° C) contrasting with

(a) Thermal component layout.

(b) Heated bed.

(c) ArUco shaped cork.

(d) ArUco shaped aluminum.

(e) Thermal component setup.

(f) Thermal sensor image.

Fig. 3. Thermal component of ArTuga fiducial marker, for a thermal sensor detection.

(a) ArTuga point cloud, with high intensity points in white.

(b) Image generated of ArTuga for marker identification.

Fig. 4. Retroreflective component of ArTuga marker, for 3D LiDAR detection.

the regions made of cork (cold areas, close to the ambient temperature). The hot areas represent the white bits of the regular ArUco, while the cold areas represent the black bits. Thereby, it is possible to detect and localize this marker as a thermal ArUco. Fig. 3 depicts all the elements that constitute the thermal component of ArTuga, as well as its layout and an example of an image collected by a thermal sensor.

The reflectivity-based detection of ArTuga is achieved by LiDAR sensors and by selecting two materials with distinct reflectivity indexes. In this way, the signals emitted and received by the LiDAR will be of different intensities, which increases the contrast level for signals that correspond to the white and black bits of the regular ArUco. The material selected for the high reflectivity index surface (white bits) was white retroreflective

tape, that was attached over the aluminum, while the cork of the thermal component was used for the matte surface (black bits). The detection and identification of ArTuga in this data modality can be performed in real-time by isolating the marker's points in the LiDAR's point cloud by applying an intensity high pass filter. To recognize the identification (ID) of the marker, a range image can be created from the filtered point cloud. Fig. 4 depicts the point cloud obtained by the LiDAR and the correspondent image generated.

The pose extraction is obtained by calculating the centroid of the maximum and minimum points of the filtered point cloud, enabling to estimate the center of the marker.

At last, by painting the cork in black, it is possible to recreate the regular binary pattern of ArUco marker which can be detected
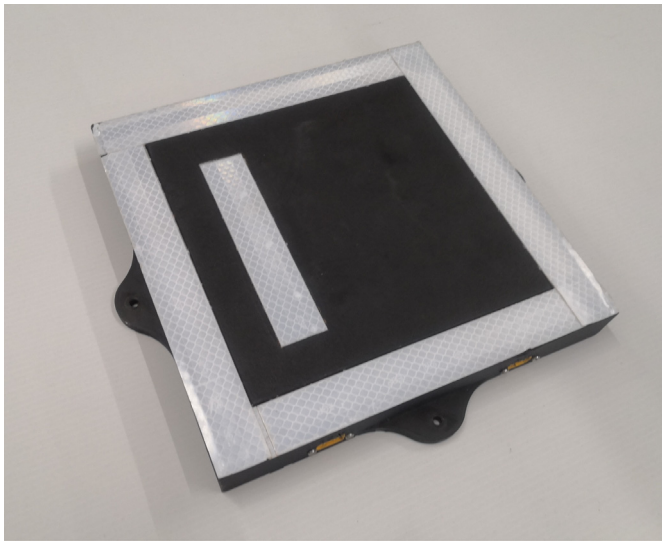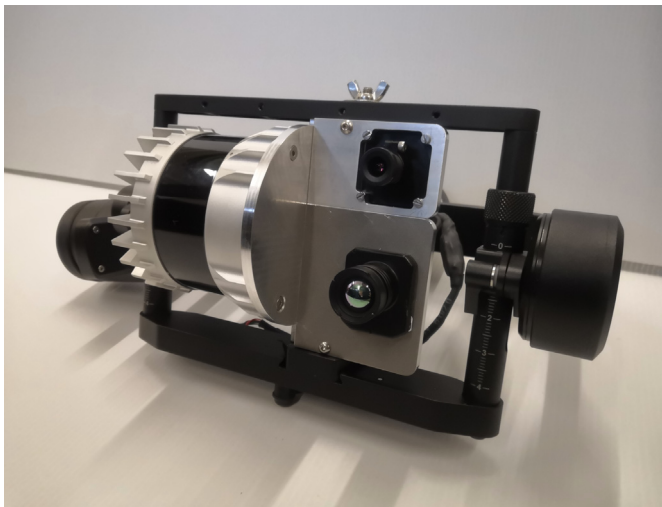
Fig. 5. ArTuga final design.

Fig. 6. Extrinsic configuration of sensors for TriOPS.

and radiometric data, such as visual, thermal and point cloud information, and is designed for the severe offshore environment. Fig. 6 depicts the developed perception system, mounted on a gimbal. The visual camera is a well-known and widely used perception sensor that collects images in the visible light spectrum, while the thermal camera gathers both thermal and radiometric information of the scene, representing a more robust sensor that does not depend so much on the lighting conditions. On the other hand, the 3D LiDAR directly acquires range data from the surrounding environment using laser beams, that are represented in point clouds. While the camera sensors collect denser data for close range procedures, the 3D LiDAR has a considerably higher field of view and range, making it suitable for long range operations, as it increases the awareness of the perception system for detecting the landing area. TriOPS not only allows to collect multimodal and complementary information about the landing target, but also plays an important and significant role in the navigation of the UAV, increasing its situational awareness of the scenario of operation, contributing to a safer operation of the UAV.

3.2.1. System calibration

The calibration of TriOPS can be performed in a two-step methodology. The first step consists in the intrinsic calibration of the camera sensors (both visual and thermal), that relates the pixel coordinates of the image with the 3D space perceived by the camera. This procedure enables to obtain the parameters of focal length, principal point and the distortion coefficients of the sensor. The second step comprises the extrinsic calibration of all sensors to obtain the transformations between each sensor frame.

For the intrinsic calibration of the visual and thermal cameras, a multimodal calibration board is proposed. For a pattern to be recognized by a thermal sensor, it has to include two components at a considerable temperature difference. To produce the calibration board, the same method explained in Section 3.1 was followed. The calibration board has an electric heat bed as the heat source, the aluminum as the good thermal conductor and the cork as the good thermal insulator. By painting the cork in white and the aluminum in black, the thermal pattern is also detected by the visual camera. Given that the heat diffusion does not allow a precise detection of the corners in a traditional calibration chessboard, an asymmetric circleboard calibration layout was selected. The circular patterns minimize the influence of the image sharpness in the calibration process, in particular in the thermal camera. Fig. 7 displays the proposed calibration board, including an image collected with the thermal camera. Note that in Fig. 7 the hot areas are represented with the black color, while the cold areas are represented with the white color. The procedure used for the intrinsic calibration of the cameras was based on the method proposed by Gramkow [53], and on the implementation presented by Wang et al. [54].

With the intrinsic calibration of the camera sensors completed, it is possible to proceed to the extrinsic calibration of TriOPS. This method used the ArTuga itself. The relative pose estimation between TriOPS and ArTuga is calculated with the ArUco algorithm [28] in both thermal and visual cameras (by using the PnP algorithm [22]), while with the LiDAR sensor the 3D positions of the collected point cloud are directly analyzed. By measuring several real positions of ArTuga, and by estimating its position relative to each of TriOPS' sensors, it is possible to perform an extrinsic calibration to obtain the transformations (translations and rotations) between each sensor frame.

by visual cameras, fulfilling the visual component of ArTuga. Both visual and thermal components of ArTuga are detected using the ArUco library [28], while the relative pose is retrieved recurring to the PnP algorithm best suited for pose estimation given four coplanar object points (the four ArUco marker's corners) [22]. The final design of the marker is depicted in Fig. 5. The developed multimodal fiducial marker allows the detection and identification of a single marker with three heterogeneous sensors, by analyzing both photometric and radiometric data. This characteristic increases the robustness and reliability of the marker, and also the complementarity of the whole system. This marker represents a compact and easy to use solution for multiple robotic applications.

3.2. Heterogeneous perception system

A multimodal perception system formed by heterogeneous sensory devices such as LiDAR, visual and thermographic cameras is proposed in this research to allow the UAV to successfully land in adverse environments. This multimodal system, named TriOPS (TRIple Offshore Perception System), obtains both photometric

(a) Asymmetric circleboard.
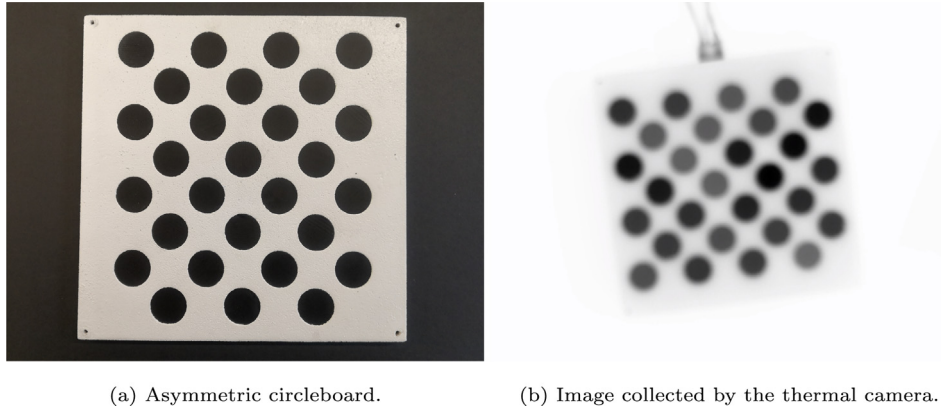
(b) Image collected by the thermal camera.

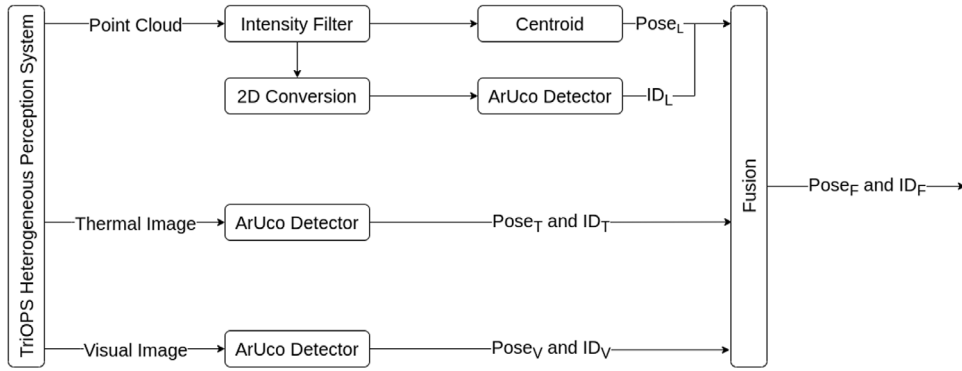**Fig. 7.** Calibration asymmetric circleboard for visual and thermal cameras.

Fig. 8. Scheme of position estimation and identification of ArTuga, for each of TriOPS' data types.

3.2.2. Multimodal data fusion

A late fusion approach is proposed in this manuscript to combine the three pose estimations of ArTuga, to output a single and redundant localization of the detected marker. The selected method consists in applying a weighted average that ensures a short processing time and increases the computational efficiency for an embedded system, guaranteeing a real time detection and relative pose estimation. Fig. 8 depicts the scheme for position estimation and identification of ArTuga, for each of TriOPS' data types.

Eqs. (1) and (2) demonstrate how the position estimation of the system is obtained:

$$X_{F_j} = \frac{\sum_{\forall i} \lambda_i w_{i,j} X_i}{\sum_{\forall i} \lambda_i w_{i,j}}, \quad i = \{V, T, L\}, \quad j = \{x, y, z\}, \tag{1}$$

$$\begin{cases} \lambda_i \in \{0, 1\}, \\ w_{i,j} \in [0, 1], \\ \sum_{\forall i} w_{i,j} = 1 \; \forall j. \end{cases} \tag{2}$$

where $X_F = (X_{F_x}, X_{F_y}, X_{F_z}) = (x_F, y_F, z_F)$ is the position estimation of the system, w_i and $X_i = (x_i, y_i, z_i)$ are the weights and the estimation of the relative position of ArTuga given by each TriOPS' sensor: visual camera (V), thermal camera (T) and 3D LiDAR (L), respectively. λ_i is a Boolean variable that takes value 1 when its correspondent sensor detects the marker, and takes value 0 otherwise. The weight of each sensor in the fusion method is calculated according to Eq. (3), where σ and $\bar{\epsilon}$ represent the standard deviation and the mean error, respectively:

$$w_{i,j} = \frac{1}{|\bar{\epsilon}_{i,j} + \sigma_{i,j}|}, \quad i = \{V, T, L\}, \quad j = \{x, y, z\}. \tag{3}$$

These values can be empirically obtained in a controlled environment (as discussed in Section 4.1.1), where the real position of

ArTuga is measured and compared with the pose estimation given by each TriOPS' sensor.

4. Results

The analysis on the accuracy and robustness of the developed system, which includes both TriOPS' perception system and ArTuga fiducial marker, comprises several experiments:

1. Characterization of the system's error in a controlled environment;
2. Several precise landing tests in a real scenario with a UAV;
3. Detection tests in distinct troublesome scenes (such as with different altitude and lighting conditions).

The experiments have been conducted with TriOPS perception system, with a set of sensors with the following specifications:

- **Visual Camera** - *The Imaging Source DFM 37UX273-ML* - Frame Rate: 5 Hz, Resolution: 1440 × 1080 pixels,1u Field of View: 45° horizontal;
- **Thermal Camera** - *FLIR Boson 640 Radiometry* - Frame Rate: 5 Hz, Resolution: 640 × 512 pixels, Field of View: 50° horizontal, Temperature Measurement Accuracy: ±5 ° C;
- **3D LiDAR** - *Ouster OS1-64* - Frame Rate: 10 Hz, Resolution: 64 × 512 channels, Range: 120 m, Accuracy: ±0,05 m, Field of View: 360° horizontal and 45° vertical.

The precision landing tests featured INESC TEC's Robotic Aerial Vehicle for any-Environment Navigation (RAVEN). This VTOL UAV is based on a hexacopter frame (with a wingspan of 1.6 m and a maximum payload of 10 Kg) with a PixHawk controller running the ArduPilot software. Its navigation system is composed of a GNSS/RTK sensor, that has a positioning accuracy of 2.5 m and
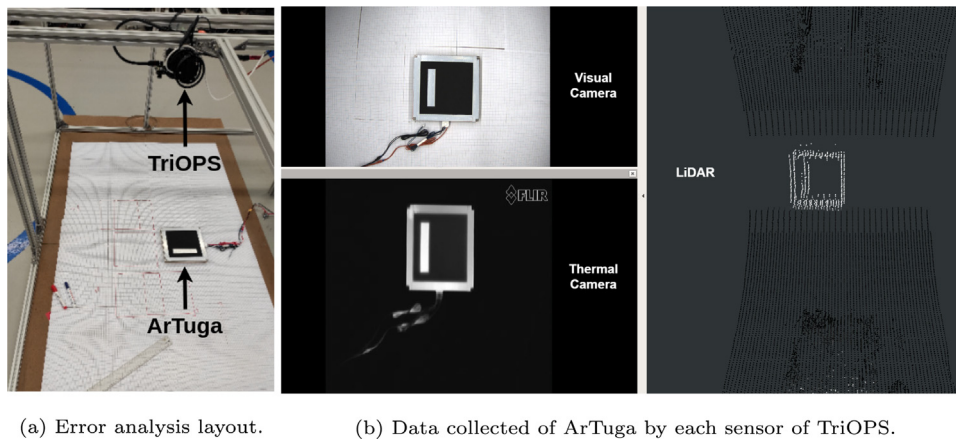
(a) Error analysis layout.

(b) Data collected of ArTuga by each sensor of TriOPS.

Fig. 9. Error analysis of ArTuga and TriOPS, at 1 m height.

of 0.01 m when paired with a base station. It has an onboard processing unit that corresponds to an Aetina AN110-XNX with a NVIDIA Jetson Xavier NX Module. RAVEN is supplied by four 6 S 30,000 mAh semi-solid Li-Ion batteries, which enables a total flight time of at least 50 min (depending on the payload). TriOPS perception system was mounted on a Gremsy T3 V3 3-axis gimbal, which allows it to obtain more stable data and to increase the degrees of freedom of the sensors. This setup enables to detect ArTuga that is placed on the landing target.

The results obtained with ArTuga and TriOPS where compared with the state-of-the-art approach that resorts to a single visual camera and to a binary planar marker, such as an ArUco [28]. The marker pose estimation is then retrieved by applying the PNP algorithm [22].

4.1. Error characterization

The characterization of the error on the relative localization of ArTuga by TriOPS is analyzed in a solid $2 \times 2 \times 1$ m cage, which allows to study the stationary precision and accuracy of the perception system. TriOPS was attached to the top of the structure, facing downward, while ArTuga was placed in its bottom and over a grid. This layout is depicted in Fig. 9(a), while Fig. 9(b) displays the data of ArTuga collected with TriOPS. The grid allows to measure the real position of ArTuga, relative to TriOPS, and to compare the obtained value with the position estimation given by each sensor. More than 400 samples were collected to analyze the error of the system, with ArTuga in diverse positions inside the cage. The positions were selected to analyze the error in close range, given that the requirements on the precision of the system increases with the proximity to the landing target.

Table 1 presents the error characterization for relative pose estimation of the marker for each sensor, exposing the average of the maximum and mean error, as well as of the standard deviation, for all the samples collected. Given that the landing target plane corresponds to the xOy, the error in both x and y axis were analyzed by the Euclidean distance from the real pose to the estimated one. The other variable in analysis is the relative altitude, that corresponds to the z axis absolute error to the landing target plane. The obtained results show that the camera sensors, at 1 m height, are considerably more precise than the 3D LiDAR in the position estimation in the xOy plane. This can be explained by the fact that the LiDAR method estimates the center of the marker based on a centroid calculation. In addition, the lower resolution of the LiDAR also justifies its lower precision, when compared with camera sensors with higher resolution. At close range, the camera resolution and data distribution are

denser that in a LiDAR sensor, which can explain its increased precision on the estimation. The main advantage of using a LiDAR corresponds to its larger field of view and range, increasing the awareness of the landing target at long range. The errors obtained in the z axis show that the 3D LiDAR has a higher accuracy than the cameras, being capable of compensating the error in the vertical measurements that are obtained with the cameras. This result is explained by the fact that the cameras have a limited ability to perceive depth, as they are 2D sensors. The standard deviation values demonstrate that the position estimations are stable, which is important for the navigation system to maintain a steady behavior.

4.1.1. Data fusion

As presented in Eq. (3), the weights used in the fusion method of the weighted average are calculated recurring to the mean error and to the standard deviation values. The weights of each sensor were calculated and are exhibited in the footnote of Table 1. The weights are distinguished between the xOy plane and the z axis, as the precision of each sensor varies with the estimation variable. Note that the presented weights are optimized for close range (at 1 m height) detections and position estimations, due to the more demanding precision requirements in the final stage of the precise landing. Using the calculated weights for the data fusion method, it is possible to evaluate this approach. Table 1 presents the error characterization, including for the data fusion method, exposing the maximum and mean error, as well as the standard deviation. Fig. 10 depicts the charts of the Euclidean error in the xOy plane and the absolute error in the z axis, for the relative pose estimation of ArTuga.

Although the mean and maximum errors in the xOy plane of the fusion method are higher than the results obtained with only the visual camera, the inclusion of the thermal camera and LiDAR increases robustness, range and field of view to the whole perception system. On the other hand, the standard deviation in the xOy plane and the error in the z axis present a considerable improvement when compared to the position estimation obtained by each sensor only by itself. Note that the error obtained with the fusion method in the z axis is lower than the error obtained with each sensor, due to the fact that TriOPS' sensors produce estimations with opposite signals, of which each error compensates the other.

The precision of the fusion method of 0.006 m, in the xOy plane (at 1 m height), consists in a considerable improvement in the landing target detection and pose estimation, when compared with traditional methods that only rely on GNSS sensors. The results obtained with the standard visual approach, that relies on

Table 1

Error characterization for the relative pose estimation of ArTuga (for more than 400 samples), for each of TriOPS sensors and for the fusion method[a].

Sensors	Maximum error		Mean error		Standard deviation	
	xOy (m)	z (m)	xOy (m)	z (m)	xOy (m)	z (m)
Visual Camera	0.0057	0.0699	0.0056	0.0660	0.0004	0.0013
Thermal Camera	0.0108	0.0648	0.0099	0.0605	0.0007	0.0017
3D LiDAR	0.0763	0.0291	0.0694	0.0247	0.0047	0.0017
Fusion	0.0064	0.0156	0.0060	0.0128	0.0003	0.0011

[a] $w_{V,x} = w_{V,y} = 0.6053$, $w_{T,x} = w_{T,y} = 0.3596$, $w_{L,x} = w_{L,y} = 0.0351$;
$w_{V,z} = 0.2518$, $w_{T,z} = 0.1702$, $w_{L,z} = 0.5780$.
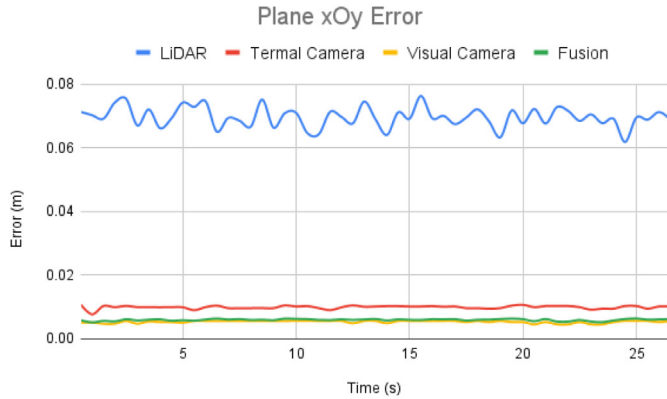
(a) Euclidean error in the xOy plane.
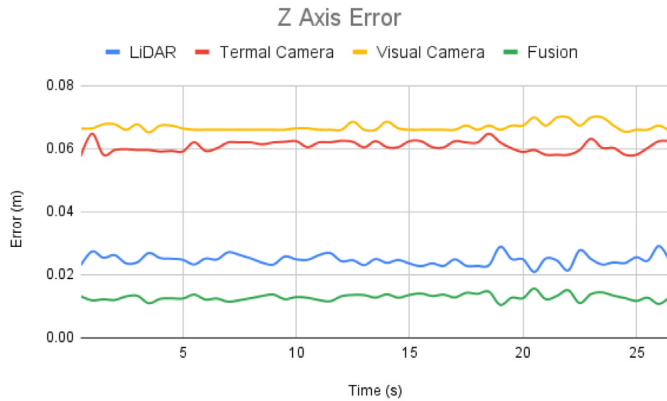
(b) Absolute error in the z axis.

Fig. 10. Euclidean error in the xOy plane and absolute error in the z axis of the relative pose estimation of ArTuga, for each of TriOPS' sensors and for the fusion method.

a single camera and a binary marker, are presented in the first row of Table 1. Despite the fact that the error on the xOy plane is similar to the one obtained with the fusion method, the error on the z axis is considerably higher (80.6% greater when the mean error is analyzed). These results validate the improvement that is obtained in the precise landing target detection when ArTuga and the TriOPS system are used, when compared with state-of-the-art approaches. Also, the inclusion of the thermal camera and the 3D LiDAR guarantees redundancy and that this system works properly under severe weather and lighting conditions.

4.2. Precise landing tests

The accuracy of the landing tests was analyzed after performing 12 successful landing maneuvers with TriOPS mounted on

Table 2

Euclidean deviation from the center of RAVEN's frame to ArTuga, for 12 precise landing samples.

Trial	Euclidean error (m)	Trial	Euclidean error (m)
1	0.004	7	0.008
2	0.018	8	0.027
3	0.017	9	0.019
4	0.029	10	0.059
5	0.002	11	0.048
6	0.003	12	0.085

Maximum error (m)	Mean error (m)	Standard deviation (m)
0.085	0.027	0.026

RAVEN, in the *Faculty of Sports of University of Porto* (FADEUP) with clear sky and light wind (wind speed of 2.5 m/s ±0.5). Once again, ArTuga was placed over a precision grid that allows to measure the deviation between the spot where the UAV landed and the marker. Fig. 11 depicts the precise landing test scene with RAVEN, TriOPS, ArTuga and the precision grid.[2] After each landing, the distance from the center of RAVEN's frame to ArTuga was measured. The obtained results are exhibited in Table 2, which presents the Euclidean error of each trial, as well as the maximum and mean deviation from the marker, along with the standard deviation, for all the collected samples.

Even with the wind conditions during the precise landing tests that forced a higher deviation from the marker, the obtained results are significantly better than the accuracy obtained with a traditional GNSS-based method. Given that the landing experiments were conducted under favorable lighting conditions, it is expected that the performance of the standard approach of a single visual camera and a binary marker would be identical. However, it would not perform appropriately under adverse conditions. This field experiment depicts the potentiality of heterogeneous perception systems and multimodal fiducial markers for the landing target detection and precise landing tasks. The complementary of TriOPS and ArTuga increases robustness and redundancy to this problematic maneuver, when compared with other standard and limited systems.

### 4.3. Redundancy and robustness analysis

The landing target detection by a UAV may be troublesome in several complex scenarios of operation. Variables such as the altitude of the UAV, the lighting condition and marker occlusions, caused by the environment conditions, degrade the detection rate of standard and limited systems. For the redundancy and robustness analysis of this system, several tests were conducted. To assess the impact of the altitude in ArTuga detection, TriOPS was mounted on RAVEN, which successively landed and took off

---

[2] A video of a precise landing test is available at https://youtu.be/ETL2BCwhnhM (3rd person view), and at https://youtu.be/7_ugB5YbNBQ is available a video of the data collected by TriOPS during another landing.

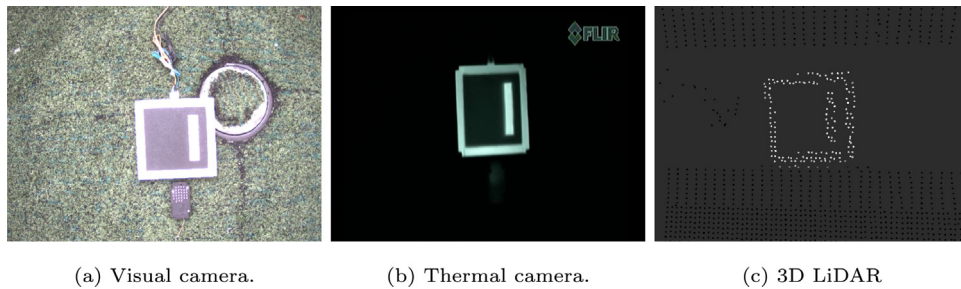**Fig. 11.** TriOPS perception system mounted on RAVEN UAV, above ArTuga and the precision grid.

(a) Visual camera.　　　　(b) Thermal camera.　　　　(c) 3D LiDAR
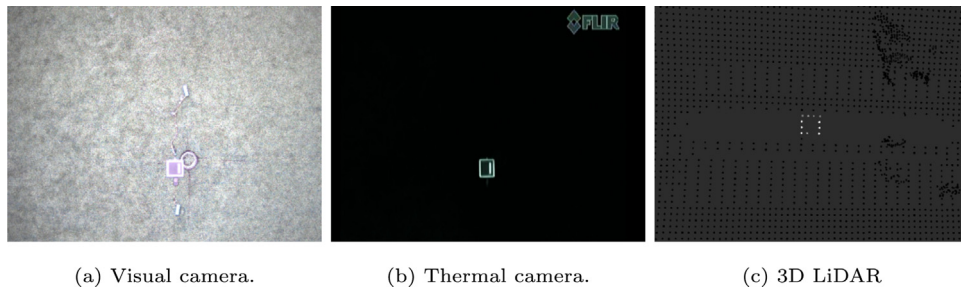
Fig. 12. Data collected of ArTuga by each TriOPS' sensor at an altitude of 5.5 m. Position returned by the fusion method: $(-0.6; -0.1; 5.5)$ m.

(a) Visual camera.　　　　(b) Thermal camera.　　　　(c) 3D LiDAR

Fig. 13. Data collected of ArTuga by each TriOPS' sensor at an altitude of 2.5 m. Position returned by the fusion method: $(0.1; -0.1; 2.5)$ m.

above the marker, in a day with favorable weather and lighting conditions. To evaluate the influence of the lighting conditions, two distinct scenarios were used. To test the behavior of the system in a dark scene, both TriOPS and ArTuga were placed in a closed locker without any light source. For the opposite scenario, several landing maneuvers were performed with RAVEN in FADEUP in a day with high intensity of sunlight.

### 4.3.1. Assessing the impact of altitude

Figs. 12 and 13 depict the data collected of ArTuga from each TriOPS' sensor in two distinct situations. In Fig. 12 the UAV was at an altitude of 5.5 m, while in Fig. 13 the UAV was at an altitude of 2.5 m.

It is noticeable that at the higher altitude the marker can be detected and identified by both cameras. Although the LiDAR detects the marker, it does not have sufficient information to obtain its ID. In this situation, the three sensors will be able to detect and localize the marker, while only the cameras can retrieve the landing target identification. At the lower altitude, every sensor is capable of detecting and identifying the marker, as the LiDAR already has the required points density. Fig. 14 depicts the detection rate of TriOPS as a function of the altitude, with a 22 cm length ArTuga, for more than 800 samples.

This graphic shows that the detection rate of TriOPS is over 99.5% until 8 m of altitude from the landing target, dropping to over 94% from 8 m to 10 m of altitude. The detection rate of TriOPS is always superior to the detection rate of each sensor by
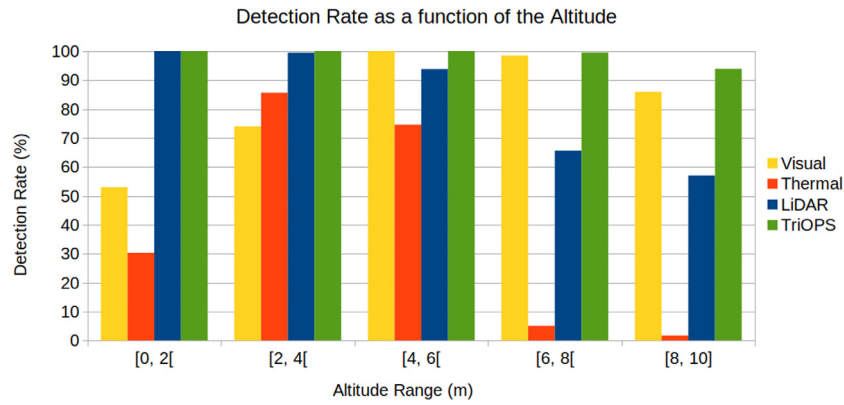
Fig. 14. Detection rate of TriOPS as a function of the altitude, with a 22 cm length ArTuga, for more than 800 samples.

(a) Visual camera.   (b) Thermal camera.   (c) 3D LiDAR
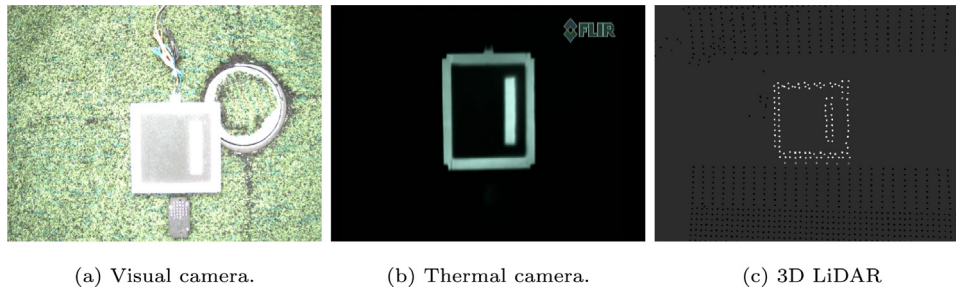
Fig. 15. Data collected in an intense sun light reflection environment, at 1.5 m height. Position returned by the fusion method: $(-0.2;\ 0.0;\ 1.5)$ m.

(a) Visual camera.   (b) Thermal camera.   (c) 3D LiDAR
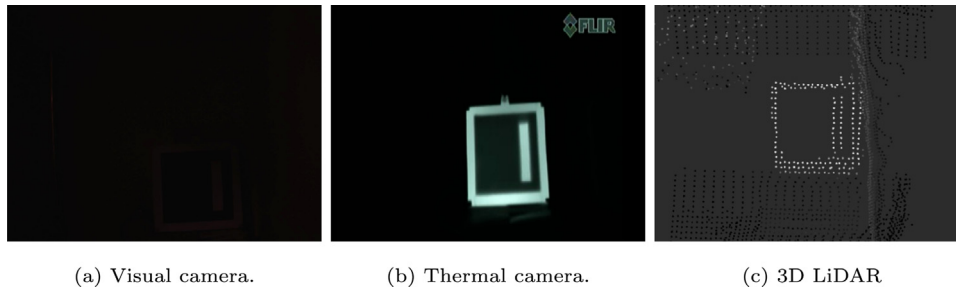
Fig. 16. Data collected in a dark environment, at 1 m of distance. Position returned by the fusion method: $(-0.2;\ -0.1;\ 1.0)$ m.

itself. Considering only the visual camera, the total detection rate is significantly lower, particularly at close range. These results prove the complementarity of TriOPS perception system. At lower altitudes, the camera sensors show a lower detection rate due to the fact that the marker's dimensions exceed the cameras' Field of View (FOV). At higher altitudes, the visual camera show better results due to its greater resolution, when compared with the thermal camera and with the 3D LiDAR. The heat diffusion also degrades the sharpness of the thermal image, decreasing the detection rate of the thermal camera for altitudes higher than 6 m. Note that during this experiment no false positives were detected.

4.3.2. Evaluating the influence of the lighting conditions

Figs. 15 and 16 depict the data collected of ArTuga from each TriOPS' sensor in distinct lighting conditions. Fig. 15 was collected in an intense sun light reflection scene, while Fig. 16 was obtained in a closed locker without any lighting source. In Fig. 15 it is noticeable the effect of the sun light reflection in the visual camera, which does not allow the correct detection of the marker. Furthermore, Fig. 16 also depicts an example where the visual camera was unable to detect the marker, due to the absence of

illumination in the scene. These two situations demonstrate the limitations of a visual camera sensor to detect the marker with non-favorable lighting conditions. In this case, the state-of-the-art approach that only recurs to a visual camera and a binary marker would not be able to detect the marker, due to its dependency on the external light conditions. On the other hand, both the thermal camera and the LiDAR are able to successfully detect and identify the marker, being robust against the illumination of the scene. These results demonstrate the redundancy and complementarity of TriOPS system, which is able to recognize ArTuga under extreme lighting conditions.

5. Conclusion

This work presents a novel multimodal fiducial marker, named ArTuga, that can be detected through photometric and radiometric data. This marker improves significantly the robustness and redundancy in the landing target detection and precise landing tasks of a VTOL UAV. A heterogeneous perception system, named TriOPS, was also developed to match ArTuga capacities. The complementarity between these two systems showed its ability to successfully operate in several distinct and challenging

conditions, such as at different heights, with different lighting and also in harsh environment conditions, where standard visual approaches are not capable of performing accordingly. Real experiments achieved a precision of 0.006 m for relative position estimation, with TriOPS and ArTuga distanced 1 m from each other. At last, field experiments of landing tests with a real UAV validated the developed system, which performed several precise and successful landings with an average precision of 0.027 m.

As future work, it is intended to improve the position estimation method used with LiDAR point clouds, as well as upgrade the data fusion method to also consider UAV altitude.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

References

[1] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, I. Moscholios, A compilation of UAV applications for precision agriculture, Comput. Netw. 172 (2020) 107148, http://dx.doi.org/10.1016/j.comnet.2020.107148.

[2] C. Franco, G. Buttazzo, Coverage path planning for UAVs photogrammetry with energy and resolution constraints, J. Intell. Robot. Syst. 83 (2016) 445–462, http://dx.doi.org/10.1007/s10846-016-0348-x.

[3] B. Trenčanová, V. Proença, A. Bernardino, Development of semantic maps of vegetation cover from UAV images to support planning and management in fine-grained fire-prone landscapes, Remote Sens. 14 (2022) 1262, http://dx.doi.org/10.3390/rs14051262.

[4] T. Giitsidis, E. Karakasis, A. Gasteratos, G.C. Sirakoulis, Human and fire detection from high altitude UAV images, in: 2015 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, 2015, pp. 309–315, http://dx.doi.org/10.1109/PDP.2015.118.

[5] E. Lygouras, A. Gasteratos, K. Tarchanidis, A. Mitropoulos, ROLFER: A fully autonomous aerial rescue support system, Microprocess. Microsyst. 61 (2018) 32–42, http://dx.doi.org/10.1016/j.micpro.2018.05.014.

[6] J.P. Aquilina, R.N. Farrugia, T. Sant, On the energy requirements of UAVs used for blade inspection in offshore wind farms, in: 2019 Offshore Energy and Storage Summit, OSES, 2019, pp. 1–7, http://dx.doi.org/10.1109/OSES.2019.8867145.

[7] D.F. Campos, M. Pereira, A. Matos, A.M. Pinto, DIIUS - Distributed perception for inspection of aquatic structures, in: OCEANS 2021: San Diego – Porto, 2021, pp. 1–5, http://dx.doi.org/10.23919/OCEANS44145.2021.9705939.

[8] A.M. Pinto, J.V.A. Marques, D.F. Campos, N. Abreu, A. Matos, M. Jussi, R. Berglund, J. Halme, P. Tikka, J. Formiga, C. Verrecchia, S. Langiano, C. Santos, N. Sá, J.-J. Stoker, F. Calderoni, S. Govindaraj, A. But, L. Gale, D. Ribas, N. Hurtós, E. Vidal, P. Ridao, P. Chieslak, N. Palomeras, S. Barberis, L. Aceto, ATLANTIS - The atlantic testing platform for maritime robotics, in: OCEANS 2021: San Diego – Porto, 2021, pp. 1–5, http://dx.doi.org/10.23919/OCEANS44145.2021.9706059.

[9] P. Cao, Y. Liu, C. Yang, S. Xie, K. Xie, MEC-driven UAV-enabled routine inspection scheme in wind farm under wind influence, IEEE Access 7 (2019) 179252–179265, http://dx.doi.org/10.1109/ACCESS.2019.2958680.

[10] D. Malyuta, C. Brommer, D. Hentzen, T. Stastny, R. Siegwart, R. Brockers, Long-duration fully autonomous operation of rotorcraft unmanned aerial systems for remote-sensing data acquisition, J. Field Robotics 37 (1) (2020) 137–157, http://dx.doi.org/10.1002/rob.21898.

[11] J. Wubben, F. Fabra, C.T. Calafate, T. Krzeszowski, J.M. Marquez-Barja, J.-C. Cano, P. Manzoni, Accurate landing of unmanned aerial vehicles using ground pattern recognition, Electronics 8 (12) (2019) http://dx.doi.org/10.3390/electronics8121532.

[12] D.F. Campos, A. Matos, A.M. Pinto, Multi-domain mapping for offshore asset inspection using an autonomous surface vehicle, in: 2020 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC, 2020, pp. 221–226, http://dx.doi.org/10.1109/ICARSC49921.2020.9096097.

[13] D. Campos, A. Matos, A. Pinto, Multi-domain inspection of offshore wind farms using an autonomous surface vehicle, SN Appl. Sci. 3 (2021) http://dx.doi.org/10.1007/s42452-021-04451-5.

[14] P.N. Leite, R.J. Silva, D.F. Campos, A.M. Pinto, Dense disparity maps from RGB and sparse depth information using deep regression models, in: A. Campilho, F. Karray, Z. Wang (Eds.), Image Analysis and Recognition, Springer International Publishing, Cham, 2020, pp. 379–392, http://dx.doi.org/10.1007/978-3-030-50347-5_33.

[15] A. Pinto, A. Matos, Maresye: A hybrid imaging system for underwater robotic applications, Inf. Fusion 55 (2019) http://dx.doi.org/10.1016/j.inffus.2019.07.014.

[16] P.N. Leite, A.M. Pinto, Exploiting motion perception in depth estimation through a lightweight convolutional neural network, IEEE Access 9 (2021) 76056–76068, http://dx.doi.org/10.1109/ACCESS.2021.3082697.

[17] A.M. Pinto, P.G. Costa, M.V. Correia, A.C. Matos, A.P. Moreira, Visual motion perception for mobile robots through dense optical flow fields, Robot. Auton. Syst. 87 (2017) 1–14, http://dx.doi.org/10.1016/j.robot.2016.08.014.

[18] A.M. Pinto, M.V. Correia, A. Paulo Moreira, P.G. Costa, Unsupervised flow-based motion analysis for an autonomous moving system, Image Vis. Comput. 32 (6) (2014) 391–404, http://dx.doi.org/10.1016/j.imavis.2014.04.003.

[19] A. Maykol Pinto, P.G. Costa, M.V. Correia, A. Paulo Moreira, Enhancing dynamic videos for surveillance and robotic applications: The robust bilateral and temporal filter, Signal Process., Image Commun. 29 (1) (2014) 80–95, http://dx.doi.org/10.1016/j.image.2013.11.003.

[20] X. Liu, S. Zhang, J. Tian, L. Liu, An onboard vision-based system for autonomous landing of a low-cost quadrotor on a novel landing pad, Sensors 19 (21) (2019) http://dx.doi.org/10.3390/s19214703.

[21] K. Wenzel, A. Masselli, A. Zell, Automatic take off, tracking and landing of a miniature UAV on a moving carrier vehicle, J. Intell. Robot. Syst. 61 (2011) 221–238, http://dx.doi.org/10.1007/s10846-010-9473-0.

[22] T. Collins, A. Bartoli, Infinitesimal plane-based pose estimation, Int. J. Comput. Vis. 109 (2014) http://dx.doi.org/10.1007/s11263-014-0725-5.

[23] O. Araar, N. Aouf, I. Vitanov, Vision based autonomous landing of multirotor UAV on moving platform, J. Intell. Robot. Syst. 85 (2017) http://dx.doi.org/10.1007/s10846-016-0399-z.

[24] A. Al-Kaff, D. Martín, F. García, A. de la Escalera, J. María Armingol, Survey of computer vision algorithms and applications for unmanned aerial vehicles, Expert Syst. Appl. 92 (2018) 447–463, http://dx.doi.org/10.1016/j.eswa.2017.09.033.

[25] M. Fiala, ARTag, a fiducial marker system using digital techniques, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, CVPR'05, 2005, pp. 590–596, http://dx.doi.org/10.1109/CVPR.2005.74.

[26] E. Olson, AprilTag: A robust and flexible visual fiducial system, in: 2011 IEEE International Conference on Robotics and Automation, 2011, pp. 3400–3407, http://dx.doi.org/10.1109/ICRA.2011.5979561.

[27] J. Wang, E. Olson, AprilTag 2: Efficient and robust fiducial detection, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2016, pp. 4193–4198, http://dx.doi.org/10.1109/IROS.2016.7759617.

[28] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, M. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, Pattern Recognit. 47 (6) (2014) 2280–2292, http://dx.doi.org/10.1016/j.patcog.2014.01.005.

[29] L. Calvet, P. Gurdjos, C. Griwodz, S. Gasparini, Detection and accurate localization of circular fiducials under highly challenging conditions, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 562–570, http://dx.doi.org/10.1109/CVPR.2016.67.

[30] B. Benligiray, C. Topal, C. Akinlar, STag: A stable fiducial marker system, Image Vis. Comput. 89 (2019) 158–169, http://dx.doi.org/10.1016/j.imavis.2019.06.007.

[31] T.S. Richardson, C.G. Jones, A. Likhoded, E. Sparks, A. Jordan, I. Cowling, S. Willcox, Automated vision-based recovery of a rotary wing unmanned aerial vehicle onto a moving platform, J. Field Robotics 30 (5) (2013) 667–684, http://dx.doi.org/10.1002/rob.21467.

[32] T. Baca, P. Stepan, V. Spurny, D. Hert, R. Penicka, M. Saska, J. Thomas, G. Loianno, V. Kumar, Autonomous landing on a moving vehicle with an unmanned aerial vehicle, J. Field Robotics 36 (5) (2019) 874–891, http://dx.doi.org/10.1002/rob.21858.

[33] M. Beul, M. Nieuwenhuisen, J. Quenzel, R.A. Rosu, J. Horn, D. Pavlichenko, S. Houben, S. Behnke, Team NimbRo at MBZIRC 2017: Fast landing on a moving target and treasure hunting with a team of micro aerial vehicles, J. Field Robotics 36 (1) (2019) 204–229, http://dx.doi.org/10.1002/rob.21817.

[34] R. Bähnemann, M. Pantic, M. Popović, D. Schindler, M. Tranzatto, M. Kamel, M. Grimm, J. Widauer, R. Siegwart, J. Nieto, The ETH-MAV team in the MBZ international robotics challenge, J. Field Robotics 36 (1) (2019) 78–103, http://dx.doi.org/10.1002/rob.21824.

[35] Z. Li, C. Meng, F. Zhou, X. Ding, X. Wang, H. Zhang, P. Guo, X. Meng, Fast vision-based autonomous detection of moving cooperative target for unmanned aerial vehicle landing, J. Field Robotics 36 (1) (2019) 34–48, http://dx.doi.org/10.1002/rob.21815.

[36] P.H. Nguyen, M. Arsalan, J.H. Koo, R.A. Naqvi, N.Q. Truong, K.R. Park, LightDenseYOLO: A fast and accurate marker tracker for autonomous UAV landing by visible light camera sensor on drone, Sensors 18 (6) (2018) http://dx.doi.org/10.3390/s18061703.

[37] P.H. Nguyen, K.W. Kim, Y.W. Lee, K.R. Park, Remote marker-based tracking for UAV landing using visible-light camera sensor, Sensors 17 (9) (2017) http://dx.doi.org/10.3390/s17091987.

[38] D. Khan, S. Ullah, D.-M. Yan, I. Rabbi, P. Richard, T. Hoang, M. Billinghurst, X. Zhang, Robust tracking through the design of high quality fiducial markers: An optimization tool for ARToolKit, IEEE Access 6 (2018) 22421–22433, http://dx.doi.org/10.1109/ACCESS.2018.2801028.

[39] S. Davis, K.G. Ricks, R.A. Taylor, Reflective fiducials for localization with 3D light detection and ranging scanners, IEEE Access 7 (2019) 45291–45300, http://dx.doi.org/10.1109/ACCESS.2019.2909467.

[40] K. Wenzel, P. Rosset, A. Zell, Low-cost visual tracking of a landing place and hovering flight control with a microcontroller, J. Intell. Robot. Syst. 57 (2010) 297–311, http://dx.doi.org/10.1007/s10846-009-9355-5.

[41] I. Kalinov, E. Safronov, R. Agishev, M. Kurenkov, D. Tsetserukou, High-precision UAV localization system for landing on a mobile collaborative robot based on an IR marker pattern recognition, in: 2019 IEEE 89th Vehicular Technology Conference, VTC2019-Spring, 2019, pp. 1–6, http://dx.doi.org/10.1109/VTCSpring.2019.8746668.

[42] G. Xu, X. Qi, Q. Zeng, Y. Tian, R. Guo, B. Wang, Use of land's cooperative object to estimate UAV's pose for autonomous landing, Chin. J. Aeronaut. 26 (6) (2013) 1498–1505, http://dx.doi.org/10.1016/j.cja.2013.07.049.

[43] G. Xu, Y. Zhang, S. Ji, Y. Cheng, Y. Tian, Research on computer vision-based for UAV autonomous landing on a ship, Pattern Recognit. Lett. 30 (6) (2009) 600–605, http://dx.doi.org/10.1016/j.patrec.2008.12.011.

[44] S. Khattak, C. Papachristos, K. Alexis, Marker based thermal-inertial localization for aerial robots in obscurant filled environments, in: 2018 International Symposium on Visual Computing, 2018, pp. 565–575, http://dx.doi.org/10.1007/978-3-030-03801-4_49.

[45] K. Ribeiro-Gomes, D. Hernández-López, J.F. Ortega, R. Ballesteros, T. Poblete, M.A. Moreno, Uncooled thermal camera calibration and optimization of the photogrammetry process for UAV applications in agriculture, Sensors 17 (10) (2017) http://dx.doi.org/10.3390/s17102173.

[46] M.I. Pereira, R.M. Claro, P.N. Leite, A.M. Pinto, Advancing autonomous surface vehicles: A 3D perception system for the recognition and assessment of docking-based structures, IEEE Access 9 (2021) 53030–53045, http://dx.doi.org/10.1109/ACCESS.2021.3070694.

[47] A.M. Pinto, P. Costa, A.P. Moreira, L.F. Rocha, G. Veiga, E. Moreira, Evaluation of depth sensors for robotic applications, in: 2015 IEEE International Conference on Autonomous Robot Systems and Competitions, 2015, pp. 139–143, http://dx.doi.org/10.1109/ICARSC.2015.24.

[48] L. Zhou, Z. Li, M. Kaess, Automatic extrinsic calibration of a camera and a 3D LiDAR using line and plane correspondences, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2018, pp. 5562–5569, http://dx.doi.org/10.1109/IROS.2018.8593660.

[49] A.S. Pinto de Aguiar, M.A. Riem de Oliveira, E.F. Pedrosa, F.B. Neves dos Santos, A camera to LiDAR calibration approach through the optimization of atomic transformations, Expert Syst. Appl. 176 (2021) 114894, http://dx.doi.org/10.1016/j.eswa.2021.114894.

[50] M. Oliveira, E. Pedrosa, A.P. de Aguiar, D.F.P.D. Rato, F.N. dos Santos, P. Dias, V. Santos, ATOM: A general calibration framework for multi-modal, multi-sensor systems, Expert Syst. Appl. 207 (2022) 118000, http://dx.doi.org/10.1016/j.eswa.2022.118000.

[51] M. Faraz, D. Kottas, S. Roumeliotis, 3D LIDAR–camera intrinsic and extrinsic calibration: Identifiability and analytical least-squares-based initialization, Int. J. Robot. Res. 31 (2012) 452–467, http://dx.doi.org/10.1177/0278364911435689.

[52] L. Grammatikopoulos, A. Papanagnou, A. Venianakis, I. Kalisperakis, C. Stentoumis, An effective camera-to-lidar spatiotemporal calibration based on a simple calibration target, Sensors 22 (15) (2022) http://dx.doi.org/10.3390/s22155576.

[53] C. Gramkow, 2D and 3D Object Measurement for Control and Quality Assurance in the Industry (Ph.D. thesis), Technical University of Denmark, 1999.

[54] Y.M. Wang, Y. Li, J.B. Zheng, A camera calibration technique based on opencv, in: The 3rd International Conference on Information Sciences and Interaction Sciences, 2010, pp. 403–406, http://dx.doi.org/10.1109/ICICIS.2010.5534797.

Rafael M. Claro received the M.Sc. degree in electrical and computer engineering from the Faculty of Engineering, University of Porto, Portugal, in 2020. He is currently pursuing the Ph.D. degree in electrical and computer engineering, in Faculty of Engineering of University of Porto. He is a Researcher with the Centre for Robotics and Autonomous Systems, INESC TEC. Currently, his research interests include robotics, aerial vehicles, distributed perception and perception techniques, mainly focused on maritime scenarios.

Diogo B. Silva received the M.Sc. degree in electrical and computer engineering from the Faculty of Engineering, University of Porto, Portugal, in 2022. He was a Researcher with the Centre for Robotics and Autonomous Systems, INESC TEC. Currently, he is working in the area of autonomous driving.

Andry M. Pinto received the Ph.D. degree in electrical and computer engineering from the Faculty of Engineering, University of Porto, Portugal, in 2014. He is currently an Assistant Professor with the Faculty of Engineering, University of Porto, and a Senior Researcher with the Centre for Robotics and Autonomous Systems, INESC TEC. He is the Principal Investigator of national and international research and development projects related to robotic-based operation and maintenance (O&M) activities for offshore infrastructures. His main research interests include multi-domain perception, underwater imaging, artificial intelligence, and mobile robotics.