Article

# Monocular-Based Pose Estimation Based on Fiducial Markers for Space Robotic Capture Operations in GEO 

Roberto Opromolla *(D) Claudio Vela (D) Alessia Nocerino (D) and Carlo Lombardi<br>Department of Industrial Engineering, University of Naples Federico II, 80125 Napoli, Italy<br>* Correspondence: roberto.opromolla@unina.it

Citation: Opromolla, R.; Vela, C.; Nocerino, A.; Lombardi, C. Monocular-Based Pose Estimation Based on Fiducial Markers for Space Robotic Capture Operations in GEO. Remote Sens. 2022, 14, 4483. https:// doi.org/10.3390/rs14184483

Academic Editors: Xiaogong Hu and Jianguo Yan

Received: 4 July 2022
Accepted: 5 September 2022
Published: 8 September 2022
Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:/ / creativecommons.org/licenses/by/ 4.0/).


#### Abstract

This paper tackles the problem of spacecraft relative navigation to support the reach and capture of a passively cooperative space target using a chaser platform equipped with a robotic arm in the frame of future operations such as On Orbit Servicing and Active Debris Removal. Specifically, it presents a pose determination architecture based on monocular cameras to deal with a space target in GEO equipped with retro-reflective and black-and-white fiducial markers. The proposed architecture covers the entire processing pipeline, i.e., starting from markers' detection and identification up to pose estimation by solving the Perspective-n-Points problem with a customized implementation of the Levenberg-Marquardt algorithm. It is designed to obtain relative position and attitude measurements of the target's main body with respect to the chaser, as well as of the robotic arm's end effector with respect to the selected grasping point. The design of the configuration of fiducial markers to be installed on the target's approach face to support the pose determination task is also described. A performance assessment is carried out by means of numerical simulations using the Planet and Asteroid Natural Scene Generation Utility tool to produce realistic synthetic images of the target. The proposed approach robustness is evaluated against variable illumination scenarios and considering different uncertainty levels in the knowledge of initial conditions and camera intrinsic parameters.


Keywords: on orbit servicing; active debris removal; spacecraft relative navigation; monocular pose estimation; Perspective-n-Point Problem

## 1. Introduction

Future space operations such as On Orbit Servicing (OOS) [1,2] and Active Debris Removal (ADR) $[3,4]$ are receiving increasing interest by the aerospace research community for both economical and safety reasons. OOS missions are envisaged to provide services to active spacecraft to extend their operative life (e.g., through refueling and repairing activities), while also potentially improving their performance thanks to technological upgrades [5]. Consequently, OOS represents an opportunity for satellite owners to increase the expected revenue from existing space assets, on the one hand, but also a way to avoid generating new debris that pose a threat to the life of operative spacecraft, on the other hand. Instead, ADR missions are conceived with the idea of addressing the space debris issue by removing the largest, most dangerous inoperative man-made space objects (e.g., dead satellites or rocket bodies) from their orbit [6], in order to help ensure the future sustainability of the space environment while also freeing usable orbital slots.

For both OOS and ADR, an active spacecraft (chaser) with an advanced Guidance, Navigation and Control (GNC) system and a proper docking or berthing mechanism is required to safely approach and capture the target of interest, as well as to control the resulting target-chaser stack after capture. Clearly, the need to achieve a high level of autonomy in these operations is considered a critical requirement to increase repeatability, robustness and reliability for this kind of missions [7]. In this context, the work presented in this paper is part of a research study entitled GNC and Robotic Arm Combined Control (GRACC), which addresses enabling technologies for the capture of a resident space object
using a robotic arm. The study has been conducted by a consortium of three Italian universities (namely the University of Padova, the Polytechnic of Milan, and the University of Naples) under contract with the European Space Agency (ESA) [8]. GRACC research activities have a twofold goal. On the one hand, the development of innovative solutions for (i) relative navigation using Electro-Optical sensors for the final approach (i.e., the final part of the rendezvous when the target-chaser distance is typically below 10 m ); (ii) combined control of the chaser and robotic arm for capture and stabilization. On the other hand, the development of a complete simulation environment, called Functional Engineering Simulator, capable of supporting the design and testing of the mentioned GNC technologies.

With reference to the GRACC project, this paper deals with relative navigation aspects, proposing a pose determination architecture conceived to support OOS of a representative target in Geostationary Earth Orbit (GEO). This scenario is selected since the OOS of large communication satellites is expected to undergo an exponential market growth in the next few years [9,10], also following the recent success of the two Mission Extension Vehicles, which extended the operative life of Intelsat-901 and Intelsat 10-02 in 2020 and 2021, respectively [11]. While in these missions the chaser relied on a heavy and highly power consuming scanning LIDAR as the main relative navigation sensor for operations in proximity [12], this paper focuses on the possibility of using monocular cameras. This solution has significant advantages in terms of reducing the overall system weight, size and cost, but poses complex technical challenges, especially due to the sensitivity of passive sensors facing the harsh illumination conditions that are typical of the space environment [13]. Aiming to address a worst-case OOS mission scenario, the target satellite is assumed to be semi-collaborative, i.e., actively controlled but unable to finely keep an attitude profile to ease approach operations, and passively cooperative, i.e., equipped with artificial markers installed at known locations on the target and designed to be easily recognizable in the data collected by visual systems on board the chaser. In such a scenario, monocular-based pose determination requires three main processing steps: first, the markers' location on the image plane is extracted through ad hoc image processing techniques (detection); then, the extracted markers are matched to their real-world 3D position vectors known in a target-fixed coordinate system (identification); finally, the resulting set of 2D-3D point correspondences is used to compute the relative position and attitude parameters (representing the rigid transformation from a camera-fixed coordinate system to a target-fixed one) by solving the Perspective- $n$-Points ( $\mathrm{P} n \mathrm{P}$ ) problem. While many analytical and numerical solvers have been proposed in the literature to address the PnP problem [14-16] and can be used in this context, the detection and identification steps must be tailored to the specific scenario under study, as they strongly depend on the typology and geometrical configuration of markers $[17,18]$. A detailed overview of monocular-based pose determination approaches for spacecraft close-proximity operations can be found in [19], while a brief survey on fiducial markers used for spacecraft relative navigation is provided below to motivate the selection carried out in this paper.

### 1.1. Fiducial Markers for Visual-Based Relative Navigation of Spacecraft

Fiducial markers can be classified as active or passive depending on whether they require a power source or not. Constantly illuminated and flashing Light Emission Diodes (LEDs), operating in the visible and infrared bands of the electromagnetic spectrum, are typically used as active markers [20,21]. Instead, passive markers can be realized either using objects with high reflectivity thanks to their shape and material (retro-reflective markers) that require proper illumination by an active source on the chaser, or by covering the target surface with black and white paintings or coatings to produce visual features that are characterized by high contrast in the collected images (black-and-white markers). Passive systems are considered in this work since they pose a less significant impact regarding their installation on the target in terms of power and allocation constraints, as opposed to LEDs that require volume-consuming allocation boxes, e.g., for the required power
supply cables and connectors. In addition, passive markers do not suffer failures of the target's power supply system and can thus still be employed even if the entire target has become inoperative.

Concerning retro-reflective markers, Corner Cube Reflectors (CCR) have been frequently used to support relative navigation sensors in technology demonstration missions conducted by space agencies to test autonomous rendezvous and docking capabilities. Relevant examples are given by the Proximity Operation Sensor tested during the Engineering Test Satellites VII mission in 1999 [22], the Advanced Video Guidance Sensor (AVGS) developed for the Orbital Express mission in 2007 [23], and the Videometer, which supports the close-range rendezvous operations of the Autonomous Transfer Vehicle with the International Space Station (ISS) [24]. Following the AVGS legacy, the Smartphone Video Guidance Sensor has been developed for close-range relative navigation with respect to a 3U CubeSat equipped with four CCRs and is currently undergoing experiments on board the International Space Station [25]. Recently, the use of infrared and phosphorescent retro-reflective markers has also been investigated in the frame of the PEMSUN (Passive Emitting Material Supporting Navigation at end-of-life) project [26].

Purely passive, black-and-white markers have been extensively used for the autonomous localization of mobile robotics systems, e.g., for the visual-based landing of unmanned aerial vehicles [27]. One of the first examples of their use in the space domain is given by the Concentric Contrasting Circles (CCC) [28] installed on the ISS as fiducial markers to be detected and tracked by the Canadian Space Vision System [29], and later also tested in a modified configuration during the Synchronize Position Hold Engage Reorient Experimental Satellites (SPHERES) project [30]. Other kinds of black-and-white markers have been designed and tested by means of laboratory experiments on ground. One relevant example is given by the pattern proposed in [18] for the autonomous visualbased pose determination of an eye-in-hand camera installed on a robotic arm, which is based on the combination of circular and linear white features on a black background and demonstrated high accuracy and identification rates at close range (below 2 m ) under variable illumination conditions. Square-shaped black markers on a white background have been proposed in [31]. The four corners of each marker are used to precisely compute its centroid, while identification is carried out by extracting a set of black dots (placed within each marker in a different number). Finally, the possibility of exploiting code-based markers (such as QR-codes) has also been investigated in view of the advantages in the identification process provided by the absence of ambiguity of their inner code [32-34].

Overall, retroreflectors represent a more robust solution with respect to black-andwhite markers, especially concerning the detection process. Indeed, the possibility to perform markers' segmentation by exploiting the principle of subsequent image subtraction (as for the AVGS [23]) makes the detection easier than in the case of black-and-white markers, which instead typically need more complex image processing algorithms (tailored to the specific type and shape of the pattern of markers) to avoid producing too many outliers and, consequently, hindering correct identification [31]. This advantage of retroreflectors clearly comes at the expense of a higher system complexity for the chaser, due to the need of an active illuminator (which is instead not strictly required by purely passive solutions, provided that adequate environmental illumination is available).

### 1.2. Paper Contribution and Organization

To support the final approach and capture of a passively cooperative space target using a chaser equipped with a robotic arm, the relative navigation system must be able to estimate the relative state not only between the two main bodies, but also between the robotic arm's end effector and the selected grasping point on the target. To this aim, this paper proposes an original pose determination architecture relying on two monocular cameras rigidly mounted to the chaser body, e.g., on its approaching face (chaser-fixed camera), and to the robotic arm's end effector (eye-in-hand camera), respectively. In view of the significant criticality of the considered scenario, i.e., the servicing of high-value
space assets such as communication satellites in GEO and based on the considerations on fiducial markers presented in Section 1.1, a set of retroreflectors is selected to support pose determination by the chaser-fixed camera. Instead, to limit the complexity of the robotic arm's visual system, a set of black-and-white markers is placed close to the grasping point on the target to support pose determination by the eye-in-hand camera.

The performance of the proposed pose determination algorithms is thoughtfully assessed within a numerical simulation environment, developed in MATLAB, realistically reproducing a final approach scenario toward a potential OOS target in GEO and the operation of visual sensors on board the chaser. Specifically, the target-chaser relative trajectory and the motion of the robotic arm to reach the selected grasping point are defined using the General Mission Analysis Tool (GMAT) [35] and the MATLAB Robotics System Toolbox [36], respectively. Instead, the target, chaser, and robotic arm modelling, as well as the generation of realistic synthetic images produced by the visual sensors, is entrusted to the Planet and Asteroid Natural scene Generation Utility (PANGU, v5) [37], available under an ESA license.

Overall, the contributions of this work can be summarized through the following points.

1. A pose determination architecture relying on a chaser-fixed monocular camera, an eye-in-hand monocular camera and on artificial markers is proposed to obtain pose information with respect to the main body and grasping point of a passively cooperative space target in GEO during the reach and capture phase.
2. Original approaches are proposed to carry out the detection and identification of retro-reflective and black-and-white markers operating on panchromatic and color images, respectively, which exploit a multi-step outlier rejection strategy. A key innovative point is that these approaches take advantage of an a-priori pose estimate (pose initial guess) to adaptively select the values of the image processing setting parameters (which typically require a complex and not so general tuning procedure to be correctly set).
3. An original implementation of the Levenberg-Marquardt algorithm, which uses a sequence of Euler angles for the relative attitude parametrization and formulates the cost function in normalized image coordinates, is proposed as a numerical least-square solver of the PnP problem.
4. A procedure to select the visual sensors specifications and the geometrical configuration of the corresponding fiducial markers is presented for both the chaser-fixed and eye-in-hand cameras.
5. An analysis of robustness of the proposed pose determination algorithms against variable illumination conditions is carried out, accounting for environmental and artificial light sources. This analysis provides highly valuable hints about the selection of the starting epoch for the final approach to obtain favorable illumination conditions. The presence of shadows cast by both the chaser body and the robotic arm, as well as the occurrence of camera Field of View (FOV) occlusions caused only by the robotic arm, are considered in this analysis.
6. An analysis of the robustness of the proposed pose determination algorithms considering different levels of uncertainty in the knowledge of the camera intrinsic parameters and of the relative state at the start of the reach maneuver is provided.
7. An analysis of the robustness of the proposed pose determination algorithms considering errors in the installation of the artificial markers on the target spacecraft (thus, affecting the correct knowledge of their 3D position vectors in target coordinates) is provided.

The remainder of the paper is organized as follows. Section 2 introduces the adopted mathematical notation and the definition of applicable reference frames, while also giving information on the target's geometry. Section 3 describes in detail the selection of cameras' specifications and fiducial markers' configuration, as well as the pose determination algorithm for both cameras. Section 4 presents the simulation environment and the simulated
scenario. Finally, the achieved results are described in Section 5, while conclusions are drawn in Section 6.

## 2. Preliminaries

### 2.1. Mathematical Notations and Reference Frames Definition

The following mathematical notation is adopted: scalars are indicated with plain italic letters (s); vectors are indicated with italic, underlined letters ( $\underline{v}$ ); matrices are indicated with capital, italic letters with a double underline $(\underset{=}{M})$. Concerning pose parametrization, the following notation is adopted: $\underline{t}^{C} A \rightarrow B$ is the position vector of point $B$ with respect to point $A$ in reference frame $C$ (points $A$ and $B$ can also refer to origins of reference frames, and reference frame indication is omitted if C coincides with A$) ;{\underset{=}{\mathrm{A}}}^{\mathrm{B}}$ is the rotation matrix from reference $A$ to reference $B$.

As for the reference frames employed within the work, two target-fixed coordinate systems are required to define the target/chaser and the grasping-point/end effector poses, respectively: the target geometric fixed frame (TGFF) has its origin at the center of the target launch adaptor ring (LAR), with the $z_{\text {TGFF }}$ axis orthogonal to the plane of the LAR and pointing outwards, and the $x_{\text {TGFF }}$ and $y_{\text {TGFF }}$ axes parallel to this plane; the target attachment point frame (TAPF) has its origin at the selected grasping point on the target with axes parallel to those of TGFF. The grasping point is placed on the LAR (being a common structure for satellites and sufficiently stiff for grasping), at 0.83 m from the TGFF origin along the $y_{T G F F}$ axis.

Sensor-fixed coordinate systems are then defined for the chaser-fixed and eye-in-hand camera. The chaser sensor fixed frame (CSFF) has the origin in the camera optical center, the $z_{\text {CSFF }}$ axis pointing along the sensor's boresight, and the $x_{\text {CSFF }}$ and $y_{\text {CSFF }}$ axes laying in the image plane. The same convention is adopted to define the sensor fixed frame on the robotic arm end effector $\left(\mathrm{CSFF}_{\text {arm }}\right)$. A graphical depiction of these frames in their context of use is provided in Figure 1.


Figure 1. Target and chaser spacecraft as modelled in PANGU, with representation of the reference frames being employed.

### 2.2. Target Modelling

The Maxar's 1300-class GEO satellite is selected as a reference target, being one of the most commonly employed GEO platforms and representative in terms of the mass, size and shape of the whole GEO communication satellites' population [38]. Considering the main available geometrical information on this satellite, it is modelled in PANGU as a parallelepiped featuring extendable solar panels on the faces with $\pm y_{\text {TGFF }}$ as normal directions and two communication antennas mounted on the $\pm x_{\text {TGFF }}$ faces. Details are added to the geometric model to ensure an adequate realism in the target representation,
including the LAR with the apogee motor in its center, a multi-layer insulation (MLI) texture covering all faces of the spacecraft, as well as matte/gloss maps for the definition of the solar panel's reflectivity and appearance, and supports to avoid placing the fiducial markers directly on the MLI blanket. An example of an image obtained by PANGU from this model is provided in Figure 2, also reporting the object's main dimensions.


Figure 2. Target spacecraft model as developed in PANGU, with indication of the main dimensions.

## 3. Pose Determination Architecture

The pose determination approaches employed to process images produced by the chaser-fixed and eye-in-hand cameras are presented in this section. They both operate in tracking mode, meaning that as new measurements are acquired, the pose is updated, taking advantage of at least one previous estimate (pose initial guess). Since the focus of this paper is on the final part of the rendezvous maneuver, such a pose initial guess is also available at the start of the reach and capture phase. Indeed, the target-chaser initial pose can be obtained from the relative state estimate at the end of the previous close-range rendezvous phase, while a first estimate for the pose of the robotic arm's end effector with respect to the grasping point can be obtained by combining the chaser-target pose with the robotic arm's direct kinematics (which allows estimating the pose of the end effector with respect to the base of the robotic arm, knowing the robotic arm's geometry and joints' rotation measurements). So, the pose initialization task is not addressed here. For the sake of completeness, it is worth mentioning that in the absence of a pose initial guess the markers' identification and the pose estimation tasks become partially overlapped and, thus, they are typically addressed simultaneously. One relevant example in the literature is given by the soft-assign technique applied in the SoftPOSIT algorithm [39], while RANSACbased approaches also represent a valid alternative [40]; the knowledge of the geometrical characteristics of the pattern of markers can also be exploited to aid their identification and, consequently, the pose initialization task [18].

Although two separate architectures are required for the two pose determination processes because of the different selection of sensors and fiducial markers illustrated later in this section, the steps of the proposed processing pipeline can be summarized in a single flow diagram, shown in Figure 3, for both cameras. Candidate markers are first found within the imaged scene in a detection step, based on image processing techniques. They are then matched to the real-world fiducials placed on the target, whose 3D position is known in target coordinates. Finally, the found 2D-3D correspondences are exploited within a PnP solver to obtain the pose measurements. The knowledge of a pose initial guess is exploited to adapt the parameters used in the image processing and identification steps to the current distance and relative orientation, together with the known camera intrinsic parameters.


Figure 3. Flow diagram of the general pose determination architecture employed for pose estimation with the chaser-fixed and eye-in-hand cameras.

### 3.1. Pose Estimation with the Chaser-Fixed Camera

The goal is to estimate the relative position vector $\left(\underline{t}_{C S F F} \rightarrow T G F F\right)$ and relative attitude rotation matrix ( $R_{T G F F}{ }^{C S F F}$ ) between TGFF and CSFF. Before entering algorithmic details, information regarding the type of fiducial markers and the optical system is provided, also discussing the strategy to select the markers' pattern and the camera specifications.

Inspired by the AVGS adopted in the OE mission [17,23], the proposed visual system foresees a camera on the chaser body accompanied by two sets of LEDs working at 800 and 850 nm wavelengths, which asynchronously illuminate a set of circular corner-cube retroreflectors on the target. The retroreflectors have a 1.27 cm radius and are assumed to be covered by a band-pass filter capable of blocking and reflecting light at 800 and 850 nm , respectively. Their dimension was chosen considering the typical size of commercial, off-the-shelf retroreflectors [41,42], as well as the heritage from previous missions (e.g., the OE one [43]). While a comprehensive design of the markers is out of the scope of this manuscript, it is worth mentioning that the selection of their size must account for several aspects, including the visual sensor specifications (e.g., camera resolution); the rendezvous operational concept (e.g., operative distance and rendezvous velocity); and the allocation constraints on the target surface. Since all these factors often lead to conflicting design constraints, the value selected here represents a good compromise between markers' detectability, operative range, and volume occupied on the target.

The number of markers and their disposition are selected considering several factors.

- Since the target is semi-collaborative, markers are only required on the target's approach face (i.e., the $+z_{T G F F}$ face).
- The presence of the LAR prevents the markers from being placed inside its circumference.
- The position of the grasping point additionally limits the available area for markers' placement, as it impacts the expected motion of the robotic arm. This means that markers must be placed to reduce the risk of the robotic arm occluding them.
- If a pose initial guess is available, a minimum of three non-collinear markers must be identified to obtain an unambiguous pose estimate [40].
- For a given number of markers, the more they are dispersed in the sensor's FOV, the better the accuracy in the pose estimation process will be. This statement, which can be intuitively understood, is related to the fact that if the markers were imaged very close to each other on the focal plane of the sensor, the coefficient matrix of the system of linear equations which can be built by writing the perspective projection equation for each point would be ill-conditioned. This phenomenon is conceptually similar to the dilution of precision concept, which allows determining the accuracy of GNSS-based positioning starting from the geometry of satellites in view.
- For a given distribution in the FOV, the larger the number of detected markers, the better the achievable pose accuracy level will be [14].
As a result, a pattern of 10 markers was selected, as shown in Figure 4, all placed on dedicated supports on the $-y_{T G F F}$ half of the $x_{T G F F} / y_{T G F F}$ plane to minimize the interferences with the robotic arm approaching the grasping point. Since a pose initial guess is available in the considered scenario, the four markers constituting the group of fiducials to be employed at very short range during the last portion of the final approach (i.e., from \#7
to \#10) could suffice for the purpose. Specifically, the minimum set of three non-collinear markers would be ensured by two in-plane (i.e., \#7, \#8) and one out-of-plane (i.e., \#10) fiducials, while the fourth one (i.e., \#9) is necessary to solve potential pose ambiguities, since it makes the pattern not symmetric with respect to the $y_{T G F F} / z_{T G F F}$ plane. However, a solution based on four markers cannot ensure an adequate pose accuracy during the entire final approach maneuver and, in particular, at the beginning when the camera-target distance is maximum and, consequently, all the markers would be imaged with limited dispersion on the image plane. To address this issue, the pattern must include a larger number of markers. The logic adopted here is to ensure that eight markers can be seen at the minimum camera-target distance considering nominal conditions in terms of chaser attitude pointing. Specifically, six markers (i.e., from \#5 to \#10) are required to ensure an adequate pose accuracy level, while markers \#3 and \#4 provide redundancy and allow further improvement of the achievable accuracy (as the dispersion of markers on the image plane increases). Finally, two additional markers (i.e., \#1 and \#2) are placed on the target to improve the achievable pose accuracy (again by increasing the dispersion of markers on the image plane), especially at the larger distances occurring at the beginning of the final approach. Clearly, the selected pattern represents a redundant solution, which is useful not only to improve the achievable pose accuracy, but also to ensure adequate robustness against the missed detection of markers from the image processing step, the loss of markers from the FOV (which can either be due to the coverage reduction as the camera gets closer to the target, or to significant unexpected deviations from the chaser nominal attitude introducing a misalignment in the camera pointing with respect to the approach face), and the loss of markers due to occlusions (e.g., those produced by the motion of the robotic arm). At the same time, it is worth noticing that the installation of a redundant configuration of markers could have a significant impact on the satellite design and on the associated cost, which needs to be carefully evaluated in the frame of a mission study.


Figure 4. Pattern of markers employed on the target spacecraft for the CSFF-TGFF pose estimation, with indication of the position vectors in the TGFF.

Given the reflective properties of the markers, the chaser-fixed camera must operate in the visible/near-infrared bandwidth. It is placed on the chaser's $+z_{\text {CSFF }}$ face, and its detector and the optics are selected based on coverage and resolution constraints.

First, the camera FOV can be selected based on a coverage constraint. Indeed, it must allow all the markers from \#3 to \#10 at the shortest TGFF/CSFF separation to be seen, i.e., at $r_{t c, \min }=1.8 \mathrm{~m}$. This distance is set considering the typical size and characteristics of robotic arms employed onboard spacecraft, such as those proposed in the studies concerning DEOS [44] and e.Deorbit [45] missions, with high dexterity and robustness to kinematic singularities. This constraint allows the camera to nominally cover eight markers at $r_{t c, \min }$, thus, improving both robustness and accuracy in the last moments of the approach. Since
the distance from markers \#3 and \#4 is $1.6 \mathrm{~m}\left(d_{3,4}\right)$, and given the value of $r_{t c, \text { min }}$, the minimum required FOV can be computed using the pinhole camera model, as in (1). Coherently with their purpose of increasing pose estimation accuracy at higher cameratarget separations and adding further robustness to the process, markers \#1 and \#2 are not directly considered in the definition of the minimum required FOV.

$$
\begin{equation*}
\mathrm{FOV}_{\max }=2 \tan ^{-1}\left(\frac{d_{3,4}}{2 r_{t c, \min }}\right)=47.92^{\circ} \tag{1}
\end{equation*}
$$

Hence, a FOV of $50^{\circ}$ is conservatively selected to ensure an adequate coverage margin. At this point, the detector's selection is carried out by searching for a compromise between angular resolution and computational effort. Indeed, the more pixels that are available (for a given detector's size), the smaller the instantaneous FOV (IFOV) of the camera becomes, resulting in more accurate estimates of the 2D markers' location in the image plane at the cost of an increased computational effort. Hence, a squared four megapixels detector, such as the Teledyne CCD42-40, is selected to ensure an IFOV in the order of the hundredth of degree. This detector is considered in line with cameras used for close-proximity, both in actual operations $[19,20,46]$ and for tests on ground $[47,48]$. Given the number of pixels $\left(n_{d}\right)$ and their physical dimension $\left(d_{p}\right)$ for this detector, the required focal length can be computed again under the pinhole camera model, as in (2).

$$
\begin{equation*}
f=\frac{n_{d} d_{p}}{2 \tan (\mathrm{FOV} / 2)} \tag{2}
\end{equation*}
$$

The camera technical specifications are summarized in Table 1. Considering the markers' size and their minimum separation (i.e., 0.2 m as shown in Figure 4), a single marker would occupy 8.90 pixels and their minimum separation would be 70.09 pixels at the maximum operating distance, i.e., 6.7 m . These numbers ensure that each marker can be correctly detected and distinguished from the others [49]. The maximum distance used for these calculations corresponds to a 10 m separation between the chaser and target's centers of mass (considering the size of the target and assuming a 1.5 m distance between the chaser center of mass and its approaching face where the camera is mounted).

Table 1. Technical specifications of the chaser-fixed camera and its optics, assuming the Teledyne CCD42-40 as detector.

| $n_{p}$ (Pixel) <br> Hor. $\times$ Ver. | $d_{p}(\mu \mathrm{~m})$ <br> Hor. $\times$ Ver. | $f(\mathrm{~cm})$ | FOV $\left.{ }^{\circ}{ }^{\circ}\right)$ <br> Hor. $\times$ Ver. | IFOV ( ${ }^{\circ}$ ) <br> Hor. $\times$ Ver. |
| :---: | :---: | :---: | :---: | :---: |
| $2048 \times 2048$ | $13.5 \times 13.5$ | 2.96 | $50 \times 50$ | $0.0244 \times 0.0244$ |

Before moving to the algorithms' description, it is worth mentioning that, in view of a future hardware implementation of the proposed approach, the design process of the visual sensors to be installed onboard the chaser will have to consider additional aspects about the optical system, e.g., the depth of focus and the f-stop of the lenses, which have not been addressed in this work. This statement is also applicable to the selection of the eye-in-hand camera specifications carried out in Section 3.2.

### 3.1.1. Detection

A block diagram describing the image processing pipeline proposed for markers' detection is shown in Figure 5.


Figure 5. Flow chart for the detection step proposed for the pose estimation algorithm employing images from the chaser-fixed camera.

First, two images are acquired with a very short temporal separation (which has been set to 33 ms considering a camera frame rate of 30 Hz ), illuminating the scene with the 800 nm and 850 nm lights, respectively (Figure 6a). Since the former radiation is absorbed by the markers, they will appear visible only in the second image. Therefore, the pixel-wise difference between the two images (Figure 6b) will ideally contain only the markers. Such image subtraction principle would be unsuited to deal with fast tumbling targets, unless proper synchronization between the target and chaser motion was ensured by the chaser control system. It is instead applicable in this scenario since the target-chaser relative dynamic is slow when maneuvering in close proximity to a semi-collaborative target in GEO (thus, producing only a small pose variation in the time interval between the two image acquisitions), and since the scene is illuminated by laser diodes operating at close wavelengths in which the background has similar reflectivity. Clearly, the slow but nonnegligible motion of the target with respect to the camera, as well as the fact that the target surface materials do not have exactly the same reflectivity at 800 nm and 850 nm , cause the image difference to retain a residual segmentation noise. Consequently, additional processing steps are required to remove this noise and discard potential outliers. First, a global thresholding to binarize the image difference (Figure 6c) based on Otsu's method [50] is employed, with a user-defined threshold $\tau_{b i n}$ that can be set in the interval [0, 1]. Since the background is characterized by much smaller intensities than those returned by the markers, relatively small values of $\tau_{\text {bin }}$ (e.g., from 0.1 to 0.3 ) can be set to remove most of the noise, thus, limiting the risk of discarding suitable candidates too.

The resulting binary image must be further processed to filter out outliers, i.e., highintensity blobs of pixels not corresponding to markers (which, for instance, can be generated by the MLI coating or by the edges of the LAR). A two-step outliers rejection strategy is adopted. First, an area opening operator is applied to the binary mask to discard all the pixel blobs occupying an area smaller than $\tau_{a o}$, adaptively computed as:

$$
\begin{equation*}
\tau_{a o}=\operatorname{nint}\left(\pi r_{p}^{2} s_{a o}\right) \tag{3}
\end{equation*}
$$

where nint is the nearest integer operator, $r_{p}$ is the expected radius (in pixels) of the markers projected on the image, and $s_{a o}$ is a user-defined safety margin. The latter can be set to values smaller than 1 to avoid discarding blobs (potentially corresponding to actual markers) whose area is only slightly smaller than the predicted one (i.e., equal to $\pi r_{p}{ }^{2}$ ). The value of $r_{p}$ can be computed by exploiting the pose initial guess and, consequently, the expected camera-to-target separation $\left(d_{i g}\right)$, as follows:

$$
\begin{equation*}
r_{p}=\left(\frac{r_{r}}{d_{i g}} f\right) \tag{4}
\end{equation*}
$$

where $r_{r}$ is the retroreflector's real-world radius and $f$ is the camera focal length (in pixels). As a second sorting step, all the remaining candidates with a circularity metric smaller than a user-defined threshold $\tau_{\text {circ }}$ are discarded. This metric can be computed as:

$$
\begin{equation*}
C=4 \pi a / p^{2} \tag{5}
\end{equation*}
$$

where $a$ and $p$ are the object's area and perimeter, respectively. Therefore, the equation can only equal 1 for perfect circles. The value of $\tau_{\text {circ }}$ can be set based on the expected direction from which the chaser approaches the target. Therefore, a relatively high value (e.g., 0.5 ) is preferred if an approach is conducted from a direction almost orthogonal to the face where the markers are. The more this direction diverges from the normal to the approach face, the smaller the value this parameter can be given, considering that the circles viewed from askance appear as ellipses (for which $C<1$ ).

After this outliers' rejection process, the coordinates of the centroid of the generic $i$-th blob of pixels $\left(u_{i}, v_{i}\right)$ are computed by weighting each of its $n_{b}$ pixels based on its intensity $I$ on the original difference intensity image as in (6). The computation is applied to each of the $n_{c}$ remaining candidates, and they represent the detected markers on the image plane.

$$
\begin{equation*}
u_{i}=\frac{\sum_{j=1}^{n_{b}} u_{j} I_{j}}{\sum_{j=1}^{n_{b}} I_{j}}, \quad v_{i}=\frac{\sum_{j=1}^{n_{b}} v_{j} I_{j}}{\sum_{j=1}^{n_{b}} I_{j}}, \quad i=1, \ldots, n_{c} \tag{6}
\end{equation*}
$$

When one marker is imaged close to the borders of the camera FOV, only a portion of it may be projected on the image plane. Consequently, although a blob corresponding to the marker is correctly found by this image processing pipeline, its weighted centroid may have a non-negligible offset with respect to the true center of the projected marker. To avoid such a measurement affecting the pose estimation accuracy, the centroids closer than $\tau_{\text {bord }}$ pixels from the image boundaries are discarded. This threshold is adaptively computed as a function of $r_{p}$, defined as in (4), multiplying it by a safety margin $s_{b o r d}$ typically larger than 1:

$$
\begin{equation*}
\tau_{\text {bord }}=\operatorname{nint}\left(r_{p}\right) s_{b o r d} \tag{7}
\end{equation*}
$$

The final output of the detection function is shown in Figure 6d.


Figure 6. Cont.


Figure 6. Image processing pipeline for markers' detection: (a) two subsequent images illuminating the target asynchronously with the two sources at 800 nm and 850 nm are collected; (b) a difference intensity image is obtained by subtracting the two acquired images; (c) Otsu's global thresholding is applied to compute the binary mask of the difference image; (d) the weighted centroids of the candidate markers (in red) are finally detected from the binary mask after a sorting process to discard noise and outliers.

### 3.1.2. Identification

The identification step establishes the 2D-3D correspondences between the candidate markers found within the image and their real-world counterparts, i.e., their position vector in the TGFF.

First, the markers are reprojected on the image plane (direct mapping) based on their known 3D location in target coordinates and the pose initial guess. The reprojected markers are then compared to the detected ones, and matches are found through a Nearest Neighbor $(\mathrm{NN})$ process. To avoid mismatches due to residual false detections, each candidate is associated to the closest reprojected marker only if their pixel distance is smaller than a safety threshold $\tau_{i d}$. This threshold is again computed adaptively based on $d_{i g}$ and considering the minimum real-world distance between two retroreflectors (i.e., 0.2 m ) multiplied by a safety margin $s_{i d}$, which is typically set to values greater than 1 :

$$
\begin{equation*}
\tau_{i d}=\left(\frac{0.2}{d_{i g}} f\right) s_{i d} \tag{8}
\end{equation*}
$$

An example of this matching process is shown in Figure 7.

### 3.1.3. PnP Solver

The 2D-3D correspondences found within the identification step are finally employed for pose determination. A PnP problem defined with all the matched markers is solved by means of the iterative minimization of a cost function through a custom implementation of the Levenberg-Marquardt (LM) least squares method for non-linear parameters' estimation [51]. The LM iterative procedure was preferred with respect to other numerical solvers (e.g., Gauss-Newton as in [14] or Newton-Raphson as in [52]) because of its good compromise between accuracy and rapidity in convergence. The implementation proposed within this work builds upon the work presented by Gavin [53].


Figure 7. Graphical depiction of the matching process through Nearest Neighbor. In red: markers' centroids found through the detection step. In green: markers' centroids reprojected using the pose initial guess. The arrows show the detected markers to which they are matched. Particular of the matching applied to markers \#7 to \#10 (in orange box).

A valid cost function that can be employed with the LM method is the scalar squared reprojection error $\chi^{2}$, which is defined in (9) as the sum of the squares of the distances between each marker found on the image plane and its reprojection.

$$
\begin{equation*}
\chi^{2}(\underline{h})=\sum_{i=1}^{n_{i d}}\left\|\underline{x}_{n}^{i}-\underline{x}_{n, p r e d}^{i}(\underline{h})\right\|^{2}, \quad \underline{h}=\left(\alpha, \beta, \gamma, t_{x,}, t_{y}, t_{z}\right) \tag{9}
\end{equation*}
$$

In (9), $n_{i d}$ is the number of 2D-3D correspondences employed in the computation; $\underline{x}^{i}{ }_{n}=\left(x_{n, i}, y_{n, i}\right)$ and $\underline{x}^{i}{ }_{n, p r e d}=\left(x_{n, \text { pred }, i}, y_{n, \text { pred, }, i}\right)$ are the vectors of the undistorted, normalized image plane coordinates of the $i$-th detected marker and the reprojection of its real-world match, respectively; and vector $\underline{h}$ is the one containing the parameters of the cost function. These correspond to the six elements of the pose vector, namely, the three angles defining the 3-2-1 Euler angles sequence (i.e., $\gamma, \beta$ and $\alpha$, respectively) that is used to parameterize the relative attitude rotation matrix ( $\mathcal{R}_{T G F F}^{C S F F}$ ) and the components $t_{x}, t_{y}$ and $t_{z}$ of the relative position vector ( $\underline{t}_{C S F F \rightarrow T G F F}$ ). The value of $\underline{x}^{i}{ }_{n}$ is constant from one iteration to the other and it is obtained from the coordinates of the markers detected on the observed scene as follows:

$$
\begin{equation*}
x_{n, i}=\frac{u_{i}-c_{u}}{f}, \quad v_{n, i}=\frac{v_{i}-c_{v}}{f} \tag{10}
\end{equation*}
$$

where $c_{u}$ and $c_{v}$ are the image coordinates of the camera principal point. If radial and tangential distortions are considered, Equation (10) provides the distorted normalized image coordinates, which can be transformed into the undistorted ones as explained in [54]. The estimation of $\underline{x}^{i}{ }_{n, \text { pred }}$, instead, exploits the knowledge of the pose parameters $\underline{R}_{T G F F}{ }^{\text {CSFF }}$ and $\underline{t}_{\text {CSFF }} \rightarrow$ TGFF, and is thus updated at each iteration as the parameters are refined. Specifically, first the predicted position vector of the $i$-th marker in CSFF, i.e., $t_{C S F F \rightarrow I}=\left(t_{x, i}, t_{y, i}, t_{z, i}\right)$ is computed starting from the available pose guess and from the knowledge of their position vector in TGFF, i.e., $\underline{t}_{T G F F \rightarrow i}=\left(p_{x, i}, p_{y, i}, p_{z, i}\right)$ :

$$
\begin{equation*}
\underline{t}_{C S F F \rightarrow i}={\underset{=T G F F}{C S F F} t_{T G F F \rightarrow i}+\underline{t}_{C S F F \rightarrow T G F F}}^{{ }_{=}} \tag{11}
\end{equation*}
$$

Then, the components of $\underline{x}_{n, p r e d}^{i}$ can be computed as in (12), where $s_{\bigcirc}$ and $c_{\bigcirc}$ indicate the sine and cosine of " $\bigcirc$ ", respectively.

$$
\begin{align*}
& x_{n, \text { pred,i }}=\frac{t_{x, i}}{t_{z, i}}=\frac{c_{\beta} c_{\alpha} p_{x, i}+c_{\beta} s_{\gamma} p_{y, i}-s_{\beta} p_{z, i}+t_{x}}{\left(c_{\alpha} s_{\beta} c_{\gamma}+s_{\alpha} s_{\gamma}\right) p_{x, i}+\left(c_{\alpha} s_{\beta} s_{\gamma}-s_{\alpha} c_{\gamma}\right) p_{y, i}+c_{\alpha} c_{\beta} p_{z, i}+t_{z}} \\
& y_{n, \text { pred,i }}=\frac{t_{y, i}}{t_{z, i}}=\frac{\left(s_{\alpha} s_{\beta} c_{\gamma}-c_{\alpha} s_{\gamma}\right) p_{x, i}+\left(s_{\alpha} s_{\beta} s_{\gamma}+c_{\alpha} c_{\gamma}\right) p_{y, i}+s_{\alpha} c_{\beta} p_{z, i}+t_{y}}{\left(c_{\alpha} s_{\beta} c_{\gamma}+s_{\alpha} s_{\gamma}\right) p_{x, i}+\left(c_{\alpha} s_{\beta} s_{\gamma}-s_{\alpha} c_{\gamma}\right) p_{y, i}+c_{\alpha} c_{\beta} p_{z, i}+t_{z}} \tag{12}
\end{align*}
$$

The choice of considering the Euler angles as part of the parameters of the cost function instead of a different representation of the relative attitude is related to the update of the parameters' vector at each iteration, which is performed as follows:

$$
\begin{equation*}
\underline{h}^{k}=\underline{h}^{k-1}+\underline{h}_{L M} \tag{13}
\end{equation*}
$$

where ${\underline{h_{L M}}}$ is the vector of the update terms generated at the $k$-th iteration. In fact, if quaternions or rotation matrices were used, the output of (13) would indeed require a normalization step, introducing undesired noise in the process. Nevertheless, it is worth highlighting that, once the method reaches convergence, it is always possible to convert the Euler angles into other attitude parametric representations, considering that, for instance, the use of quaternion is preferable to represent the attitude in the state vector of a Kalman filter for relative state estimation. The update term of (13) is computed as:

$$
\begin{equation*}
\underline{h}_{L M}=\left(\underset{\underline{J^{T}} J}{=}+\lambda \operatorname{diag}\left(\underset{\underline{J^{T}} J}{=}\right)\right)^{-1} \underline{J}^{T}\left(\underline{x}^{i}-\underline{x}_{n, p r e d}^{i}(\underline{h})\right), \quad \underset{=}{J}=\frac{\partial \underline{x}_{n, p r e d}^{i}(\underline{h})}{\partial \underline{h}} \tag{14}
\end{equation*}
$$

where $J$ is the $2 n_{i d}$-by-6 matrix corresponding to the Jacobian of $\underline{x}^{i}{ }_{n, p r e d}$ with respect to the parameter's vector, the diag operation returns a diagonal matrix whose entries are the diagonal terms of the argument, and $\lambda$ is the damping parameter of the update term. The expression in (14) indicates that large values of $\lambda$ lead to a steepest descend update term (and, thus, faster steps towards the minimum of $\chi^{2}$ ); on the other hand, small values allow updating the parameters' vector as in the Gauss-Newton method, leading to more accurate updates. The value of $\lambda$ at the first iteration is user-defined and indicated as $\lambda_{0}$.

The update of $\lambda$ is carried out so that the LM process proceeds with larger steps when it is moving towards the minimum of the cost function (to achieve faster convergence), while reducing the step size either when moving away from the minimum, so that the descent direction can be adjusted to move again towards it, or when moving close to it, so as to converge to more accurate estimates of the parameters. For this reason, once $\underline{h}_{L M}$ is computed, the variation in $\chi^{2}$ from step $k-1$ to step $k$ is evaluated. If $\chi^{2}$ is reduced by a quantity larger than a user-defined threshold $\left(\varepsilon_{0}\right)$, the computed update term is accepted, and thus, $\lambda$ is increased by a factor $\lambda_{U P}$. Conversely, the update term is discarded, and the parameter vector of the previous step is used again, while contextually reducing $\lambda$ by a factor $\lambda_{D N}$ [53]. In this way, a new update term with a smaller step (and closer to the Gauss-Newton update) is computed, starting again from the same point in the parameters' hyperspace. The threshold $\varepsilon_{0}$ is set to a small value (e.g., $10^{-9}$ ) so that the method is sensitive enough to detect even small variations of $\chi^{2}$.

At each iteration, the algorithm evaluates the convergence conditions summarized in Table 2, where $\varepsilon_{1}, \varepsilon_{2}$ and $\varepsilon_{3}$ are user-defined as well, and returns the last computed set of parameters if any of them are satisfied. If none of these situations are verified before the algorithm has performed $n_{m a x, i t}$ iterations, then the last computed parameter vector is assumed as the estimated pose. The value of $n_{\max , i t}$ is set considering a trade-off between accuracy and the rapidity of convergence. A flow chart summarizing the operations of the algorithm is shown in Figure 8.

Table 2. Convergence conditions for the proposed implementation of the LM technique for pose determination.

| Condition | Meaning | Settings |
| :---: | :---: | :---: |
| $\max \left(\left\|\frac{h_{L M}}{\underline{h}}\right\|\right)<\epsilon_{2}$ | The iterative process is stopped <br> if the normalized pose <br> parameters update term <br> becomes negligible. | $\varepsilon_{2}$ is set to $10^{-4}$. Smaller values <br> correspond to variations in the <br> pose estimate to which the <br> algorithm is not sensitive. |
| $\max \left(\left\\|J^{T}\left(\underline{x}_{n}^{i}-\underline{x}_{n, p r e d}^{i}(\underline{h})\right)\right\\|\right)<\epsilon_{1}$ | The iterative process is stopped <br> if the largest component of the <br> gradient in $\underline{h}$ is smaller than <br> threshold $\varepsilon_{1}$. | $\varepsilon_{1}$ and $\varepsilon_{3}$ are set to very small <br> values, i.e., $10^{-10}$ and $10^{-8}$, <br> respectively, to ensure stopping <br> close to a minimum of the cost <br> function in the |
| $\frac{\chi_{k}^{2}}{2 n_{i d}-1}=\chi_{v, k}^{2}<\epsilon_{3}$ | The iterative process is stopped <br> if the cost function goes below <br> a threshold. |  |
| hyper-parameters' space. |  |  |



Figure 8. Flow chart for the proposed implementation of the Levenberg-Marquardt's iterative method for the least squares non-linear estimation of the pose parameters.

Regarding the selection of the damping parameter and its update, $\lambda_{0}$ is set depending on how the LM method should start the iterative process. In this application, small values of the initial parameter (i.e., $10^{-8}$ ) are preferred, because this forces the algorithm to start updating the parameters' vector through the Gauss-Newton rule. Doing so, the parameters will update with smaller steps at the beginning of the process, allowing the correct identification of the direction towards the minimum. Instead, $\lambda_{\mathrm{UP}}$ and $\lambda_{\mathrm{DN}}$ are set to 11 and 9, respectively, as suggested by Gavin [53], to allow the parameters' vector to increase its convergence speed when approaching the minimum, yet promptly decreasing the step size when moving away from it or getting close to convergence.

### 3.2. Pose Estimation with the Eye-in-Hand Camera

The goal is to estimate the pose of the $\mathrm{CSFF}_{\text {arm }}$ within the TAPF. Following the same structure of Section 3.1, a detailed discussion on the selection and placement of the black and white markers on the target spacecraft, as well as of the camera detector and optics' specifications is presented prior to the introduction of the algorithm.

The black and white markers are designed as white dots with a radius of 1 cm placed on a black background granted by the $30-\mathrm{by}-10 \mathrm{~cm}^{2}$ markers' support base on the target's approach face. As for their number and distribution on the target object, it is known that three non-collinear 2D-3D point correspondences can lead to a non-ambiguous pose solution if a pose initial guess is available [43]. Therefore, a set of three, non-coplanar
markers was considered, shown in Figure 9, where the coordinates of their location in the TAPF are reported as well.


Figure 9. Pattern of markers employed on the target spacecraft for the $\mathrm{CSFF}_{\text {arm }}$-TAPF pose estimation, with indication of the position of their centroids in the TAPF.

The same camera specifications selection process as in Section 3.1 is also employed for the eye-in-hand camera. The camera is attached to the robotic arm with an offset with the end effector of 20 cm and 40 cm along the $-y_{\text {CSFFarm }}$ and $+z_{\text {CSFFarm }}$ directions, respectively. The first offset accounts for the vertical separation of the markers with respect to the grasping point, while the along-boresight one is required since, at the end of the reach and capture maneuver, the nominal distance of the end effector from the grasping point (and, consequently, also from the markers on the target surface) approaches zero. At this point, considering that the minimum separation between CSFF $_{\text {arm }}$ and TAPF is equal to the axial offset given to the camera with respect to the end-effector, and that the whole set of markers with their supports must be observed at such distance (i.e., a coverage of 30 cm is required), a minimum FOV of $41.11^{\circ}$ is expected for the camera, ultimately chosen as $45^{\circ}$ to guarantee a margin if the robotic arm is not properly aligned. Furthermore, if the same detector of the chaser-fixed camera is also selected for the eye-in-hand camera, an IFOV of $0.022^{\circ}$ is obtained. This choice is supported by the consideration that, at an assumed maximum operative distance of 1.4 m , each marker has a diameter on the image plane of 37.20 pixels, while their minimum separation is 185.94 pixels, thus, ensuring that they can be correctly detected and distinguished one from another. The technical specifications of the eye-in-hand camera are summarized in Table 3.

Table 3. Technical specifications of the eye-in-hand camera and its optics, assuming CCD42-40 is used as detector.

| $N_{p}$ (Pixel) <br> Hor. $\times$ Ver. | $D_{p}(\mu \mathrm{~m})$ <br> Hor. $\times$ Ver. | $\boldsymbol{F}(\mathrm{cm})$ | FOV $\left({ }^{\circ}\right)$ <br> Hor. $\times$ Ver. | IFOV ( $\left.{ }^{\circ}\right)$ <br> Hor. $\times$ Ver. |
| :---: | :---: | :---: | :---: | :---: |
| $2048 \times 2048$ | $13.5 \times 13.5$ | 3.34 | $45 \times 45$ | $0.0220 \times 0.0220$ |

## Detection

The detection step, which is articulated as in the flow chart of Figure 10, is performed by first identifying a region of interest (RoI) within the original image to which the markers' search is restricted. This RoI is then binarized and the resulting candidate markers are retrieved by applying an outlier rejection scheme. Finally, the centroid for each remaining candidate is computed.


Figure 10. Flow chart for the detection step of the algorithm for pose estimation with the eye-in-hand camera.

The RoI selection allows a reduction of the computational effort and is carried out by exploiting the pose initial guess provided to the algorithm to project the markers' centroids ( $u_{\text {pred, }, I}, v_{\text {pred }, i}$ ) on the image plane. Hence, given the maximum horizontal ( $d_{m a x, u}$ ) and vertical ( $d_{m a x, v}$ ) distances between these predicted markers' positions, the image coordinates of the top-left $\left(u_{l}, v_{t}\right)$ and bottom-right ( $u_{r}$, $v_{b}$ ) corners of the rectangular RoI can be identified as:

$$
\begin{array}{ll}
u_{l}=\min _{i}\left(u_{\text {pred }, i}\right)-s_{x} \max \left(d_{\max , u^{\prime}}, d_{\max , v}\right), & u_{r}=\max _{i}\left(u_{\text {pred }, i}\right)+s_{x} \max \left(d_{\max , u^{\prime}}, d_{\max , v}\right)  \tag{15}\\
v_{t}=\min _{i}\left(v_{\text {pred }, i}\right)-s_{y} \max \left(d_{\max , u^{\prime}} d_{\max , v}\right), & v_{b}=\max _{i}\left(v_{\text {pred }, i}\right)+s_{y} \max \left(d_{\max , u^{\prime}}, d_{\max , v}\right)
\end{array}
$$

where $s_{x}$ and $s_{y}$ are user-defined coefficients which are employed to enlarge the search area depending on the overall dimensions of the pattern on the image plane. As an example, a value of 0.5 for both coefficients allows enlarging the search area, adding half the maximum (horizontal or vertical) distance between the markers symmetrically with respect to the pattern's center. A larger area than the one occupied by the pattern is thus considered, accounting for initialization errors and camera misalignments with respect to the target approach phase due to the motion commanded to the robotic arm.

The RoI cropped from the original image is then converted to grayscale and binarized to distinguish between the elements of the image belonging to the brighter foreground or the darker background. The binarization is performed through an adaptive local thresholding technique, which computes the threshold for each pixel based on the mean of the intensities of neighbors within a window of horizontal and vertical dimensions $w_{u}$ and $w_{v}$, respectively. These are computed as proportional to the expected dimension of the marker's radius on the image plane (in pixel) multiplied by two safety margins $s_{c u}$ and $s_{c v}$, as in (16), where $r_{p}$ is computed considering the radius of the black-and-white (BW) markers in (4) and ceilodd is the operator rounding its argument to the nearest greater odd integer. The safety margins can be set so that a large portion of the marker is within the window when computing the threshold for the edge pixels. Therefore, a value of 2 was considered a plausible choice in this work.

$$
\begin{equation*}
w_{u}=\operatorname{ceilodd}\left(r_{p} s_{c u}\right), w_{v}=\operatorname{ceilodd}\left(r_{p} s_{c v}\right) \tag{16}
\end{equation*}
$$

A user-defined sensitivity coefficient $\tau_{s}$ is also considered in the adaptive thresholding, influencing the threshold to include more pixels within the foreground (for $\tau_{s}>0.5$ ) or the background (for $\tau_{s}<0.5$ ). The adaptive thresholding technique is preferred in this case, as opposed to the global thresholding approach employed for processing the images acquired by the chaser-fixed camera, because the algorithm does not work on an image difference, and thus, the high reflectivity of the MLI coating might produce outliers. On the contrary, the adaptive thresholding technique is known to provide greater robustness to variable illumination conditions [55] and has proven to do the same within images with high contrast.

The obtained binary mask undergoes the same outlier rejection and centroiding processes described in Section 3.1.1, the extracted information related to the markers being very similar in the two cases. Consequently, the identification of the 2D-3D correspondences and their employment in solving the PnP problem can also be carried out in the same way proposed in Sections 3.1.2 and 3.1.3, respectively. Clearly, in this case, the LM-based PnP solver allows the estimation of ${\underset{\sim}{R}}_{\text {TAPF }}{ }^{\text {CSFFarm }}$ and $\underline{t}_{\text {CSFFarm }} \rightarrow$ TAPF .

## 4. Simulation Environment and Scenario Description

The performance of the proposed pose determination architecture is extensively tested within a dedicated simulation environment, summarized in Figure 11 in the form of a block diagram.


Figure 11. Block diagram representation of the simulation environment. The same structure is considered for both the cameras. The trajectory is generated accordingly based on the camera being considered for the simulation.

The target and chaser models are loaded in PANGU, together with the camera intrinsic and extrinsic parameters, as well as the location of the Sun, which is set depending on the epoch at which the approach happens. The true pose information, generated as explained in Section 4.1 for both the chaser-fixed and eye-in-hand camera, is provided in input to PANGU to render the images to be processed by the proposed pose determination algorithms. The processing block also receives in input the target geometric information, the camera intrinsic parameters, and the pose at scenario start.

### 4.1. Reach and Capture Scenario Definition

For the sake of assessing the performance of the proposed pose determination techniques, a reach and capture scenario is simulated with the chaser approaching the target along the radial direction on a free-motion trajectory. Given the orbital parameters of the target at a selected epoch at which the maneuver is assumed to start (see Table 4), a 2-by-1 elliptic relative trajectory which intercepts the target from the required direction is designed as a solution to the non-forced formulation of the Hill-Clohessy-Wiltshire's equations. The semi-major axis is set to 411 m , obtaining a maximum relative velocity of $0.015 \mathrm{~m} / \mathrm{s}$. At this point, the difference between the orbital parameters of the two spacecraft (and, consequently, the chaser orbital parameters at scenario start also reported in Table 4) are obtained by selecting the point along this relative trajectory at which the target-chaser distance is 10 m . Hence, the absolute motion of both the chaser and target is obtained by propagating their orbit using GMAT (including all relevant perturbations).

Table 4. Orbital parameters of the target and chaser spacecraft at the beginning of the simulation of the approach.

| Spacecraft | Semi-Major <br> Axis (km) | Eccentricity | Inclination <br> $\left({ }^{\circ}\right)$ | Right Ascension <br> of the Ascending <br> Node | Argument <br> of Perigee <br> $\left({ }^{\circ}\right)$ | True <br> Anomaly <br> $\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Target | 42,164 | $1.468 \times 10^{-7}$ | 0 | 0 | 0 | 0 |
| Chaser | $42,164.29$ | $7.684 \times 10^{-6}$ | $6.096 \times 10^{-5}$ | 2.565 | 334.093 | 23.34 |

Regarding the rotational motion, the target is assumed to be three-axis stabilized with a pointing accuracy of $2^{\circ}$ in each direction and is constrained to maintain a nadir pointing attitude, in order to preserve its functionality during the servicing operations and ease the reach and capture. On the other hand, the chaser's attitude is constrained to keep the chaser-fixed camera pointed toward the target. The resulting absolute position and attitude of the two spacecraft are combined to obtain the true pose parameters. The resulting approach trajectory is depicted in Figure 12a.


Figure 12. Relative trajectories employed in the simulations: (a) relative trajectory of the chaser toward the target and the corresponding body-fixed reference frames at scenario start; (b) schematic representation showing the end-effector/grasping point pose corresponding to set-points 2 and 3 .

As for the relative trajectory of the end effector with respect to the grasping point, three setpoints are defined for the robotic arm's end effector, corresponding to the position and attitude with respect to the TAPF at the beginning of the robotic arm's extension towards the grasping point, an intermediate condition in which the eye-in-hand camera is constrained to point towards the BW markers and the condition at contact. These latter two setpoints ensure that the markers are within the FOV of the eye-in-hand camera. The arm's inverse kinematics is solved in these setpoints, and a linear interpolation is performed between the obtained solutions to obtain the temporal evolution of the rotations of the robotic arm's joints. Finally, given the robotic arm's geometry, the relative position and attitude of the end-effector within the TAPF can be retrieved. Clearly, only starting from setpoint 2 , that is, at 282.3 s from the approach start, the eye-in-hand camera can guarantee a pose solution: therefore, only that portion of the relative trajectory is considered for testing the algorithm. A schematic representation of set-points 2 and 3 can be appreciated in Figure 12b.

### 4.2. Test Cases Definition

The performance of the two pose determination architectures is assessed by executing four sets of tests.

- Effect of illumination conditions. Different illumination conditions are obtained by changing the starting epoch of the simulation and the target position along its orbit.
- Effect of uncertainty in the knowledge of the pose initial guess at the start of the scenario. An error randomly extracted from a zero-mean Gaussian distribution is added to the true pose.
- Effect of uncertainty in the knowledge of the camera intrinsic parameters. An error randomly extracted from a zero-mean Gaussian distribution is added to the nominal camera focal length and principal point This analysis accounts for residual errors which can still be present even if the calibration parameters are re-computed on orbit after launch.
- Effect of uncertainty in the knowledge of the 3D position vectors of the markers in target coordinates (due to installation errors). An error randomly extracted from a zero-mean Gaussian distribution is added to the true coordinates.

Before entering the details of the four test sets, baseline properties for the simulations are defined. Specifically, the values for the standard deviation of the errors characterizing the knowledge of the pose initial guess and camera intrinsic parameters are reported in Table 5. An uncertainty of one pixel ( $1 \sigma$ ) is considered for the camera focal length $\left(f_{u}\right.$ and $\left.f_{v}\right)$ and for the image coordinates of the principal point. The nominal uncertainty for the knowledge of the pose initial guess is selected by considering the typical performance of an EO-based relative navigation system during close range rendezvous and prior to the final approach's start [56,57].

Table 5. Error level ( $1 \sigma$ ) in the knowledge of pose initial guesses and camera intrinsic parameters considered in the simulations for both the pose determination with chaser-fixed and eye-in-hand cameras.

| Chaser-Fixed Camera |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Pose Initial Guess Uncertainty (1 $\sigma$ ) |  |  |  |  |  | Intrinsic Parameters Uncertainty (1 $\sigma$ ) |  |  |  |
| $T_{x}(\mathrm{~m})$ | $T_{y}(\mathrm{~m})$ | $T_{z}(\mathrm{~m})$ | $\alpha\left({ }^{\circ}\right.$ ) | $\beta\left({ }^{\circ}\right)$ | $\gamma\left({ }^{\circ}\right)$ | $f_{u}(\mathrm{px})$ | $f_{v}(\mathrm{px})$ | $c_{u}(\mathrm{px})$ | $c_{v}(\mathrm{px})$ |
| 0.017 | 0.017 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Eye-in-Hand Camera |  |  |  |  |  |  |  |  |  |
| Pose Initial Guess Uncertainty (1 $\sigma$ ) |  |  |  |  |  | Intrinsic Parameters Uncertainty (1 $\sigma$ ) |  |  |  |
| $T_{x}(\mathrm{~m})$ | $T_{y}(\mathrm{~m})$ | $T_{z}(\mathrm{~m})$ | $\alpha\left({ }^{\circ}\right.$ ) | $\beta\left({ }^{\circ}\right)$ | $\gamma\left({ }^{\circ}\right)$ | $f_{u}(\mathrm{px})$ | $f_{v}(\mathrm{px})$ | $c_{u}(\mathrm{px})$ | $c_{v}(\mathrm{px})$ |
| 0.010 | 0.010 | 0.010 | 0.667 | 0.667 | 0.667 | 1 | 1 | 1 | 1 |

### 4.2.1. Illumination Conditions

The considered test cases are summarized in Tables 6 and 7 for the pose estimation with chaserfixed and eye-in-hand cameras, respectively. In both, the location of the Sun is defined coherently with the target's operative orbit and is indicated through the Sun's azimuth $(A z)$ and elevation ( $E l$ ) in TGFF. These angles are defined coherently with the notation adopted by PANGU, which defines $A z$ as being measured clockwise from the $y_{T G F F}$ direction.

Concerning the tests involving the chaser-fixed camera, the investigation can be restricted to a small set of seven representative cases thanks to the presence of the LEDs used to illuminate the retroreflectors.

On the other hand, the BW markers imaged by the eye-in-hand camera are greatly affected by the natural illumination conditions that can generate shadows which could prevent markers' detection. For this reason, a set of 36 different illumination conditions are considered. The pose estimation accuracy in each case is then compared with an additional test, defined so that the Sun does not strike the approach face directly, but considering an artificial illuminator. This simulation, indicated in Table 7 as "E-SB", allows the effect of the presence of the Sun in the simulation to be assessed, as well as providing a baseline reference for the other tests.

For test cases S 1 to $\mathrm{S7}$ and E-SB, a statistical analysis of the results is carried out over 100 simulations, while for cases E-S1 to E-S36 a single simulation was performed, since those tests mainly aim to verify the approach feasibility under given conditions. In all cases for both the cameras, the uncertainty level in the knowledge of the pose initial guess at scenario start and camera intrinsic parameters is the one indicated in Table 5.

Table 6. Summary of test cases with variable illumination conditions for the pose estimation with the chaser-fixed camera.

| Test Case | $\mathbf{A z}\left({ }^{\circ}\right)$ | E1 ( ${ }^{\circ}$ ) |
| :---: | :---: | :---: |
| S1 | 0 | 90 |
| S2 | 270 | 45 |
| S3 | 270 | 0 |
| S4 | 270 | -60 |
| S5 | 90 | 45 |
| S6 | 90 | 80 |
| S7 | 67 | 75 |

Table 7. Summary of test cases with variable illumination conditions for the pose estimation with the eye-in-hand camera.

| Test Case | Az $\left.\mathbf{(}^{\circ}\right)$ | El $\left.\mathbf{(}^{\circ}\right)$ | Test Case | Az $\left(^{\circ}\right)$ | El $\left(^{\circ}\right)$ | Test Case | Az $\left(^{\circ}\right)$ | El $\left(^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| E-SB | 270 | 0 | E-S13 | 76.50 | -7.01 | E-S25 | 137.25 | -7.33 |
| E-S1 | 16.66 | -7.07 | E-S14 | 77.91 | -22.97 | E-S26 | 137.73 | -22.96 |
| E-S2 | 18.09 | -22.99 | E-S15 | 79.81 | 4.38 | E-S27 | 139.66 | 4.45 |
| E-S3 | 18.31 | 22.38 | E-S16 | 80.14 | 22.15 | E-S28 | 139.98 | 22.18 |
| E-S4 | 19.08 | 4.70 | E-S17 | 81.89 | 14.98 | E-S29 | 141.74 | 15.03 |
| E-S5 | 21.13 | 15.23 | E-S18 | 83.56 | -14.48 | E-S30 | 143.40 | -14.53 |
| E-S6 | 23.72 | -14.42 | E-S19 | 107.33 | -7.36 | E-S31 | 167.18 | -7.30 |
| E-S7 | 46.58 | -7.04 | E-S20 | 107.82 | -22.96 | E-S32 | 168.71 | -23.03 |
| E-S8 | 48.00 | -22.98 | E-S21 | 109.74 | 4.42 | E-S33 | 169.59 | 4.48 |
| E-S9 | 49.01 | 4.74 | E-S22 | 110.06 | 22.16 | E-S34 | 169.90 | 22.19 |
| E-S10 | 50.23 | 22.14 | E-S23 | 111.81 | 15.00 | E-S35 | 171.66 | 15.05 |
| E-S11 | 51.97 | 14.95 | E-S24 | 113.48 | -14.50 | E-S36 | 173.32 | -14.56 |
| E-S12 | 53.64 | -14.45 |  |  |  |  |  |  |

### 4.2.2. Uncertainty in the Knowledge of the Pose Initial Guess at Scenario Start

To prove robustness against a coarser knowledge of the pose initial guess at scenario start, one additional test is carried out for both the cameras by doubling the uncertainty level on the pose parameters reported in Table 5. A total of 100 simulations are conducted for each test case, as in the previous set. A summary of the test cases is provided in Table 8. For both these cases, the uncertainty in the knowledge of the camera intrinsic parameters is the one from Table 5, and the illumination conditions are fixed. Specifically, test case I considers the same illumination condition as in S7, since it features the Sun almost behind the chaser as it approaches the target, which causes significant brightness variation in the camera FOV and, therefore, allows for a more robust testing of the algorithm performance. Test case E-I, instead, features the same illumination condition as in E-SB, in which the Sun is completely behind the target, which allows the influence of shadows on the tests to be eliminated.

Table 8. Summary of test cases with variable accuracy on the initialization of the pose solution for both the pose estimation with the chaser-fixed and the eye-in-hand cameras.

| Pose Estimation with Chaser-Fixed Camera |  | Pose Estimation with Eye-in-Hand Camera |  |
| :---: | :---: | :---: | :---: |
| Test Case | Description of Test Conditions | Test Case | Description of Test Conditions |
| I | Noise on the initial pose guess <br> has twice the standard <br> deviations indicated in Table 5 | E-I | Noise on the initial pose guess <br> has twice the standard <br> deviations indicated in Table 5 |

### 4.2.3. Uncertainty in the Knowledge of the Camera Intrinsic Parameters

To prove robustness against a coarser camera calibration, two test conditions are considered in which the standard deviation of the white Gaussian noise applied to the camera intrinsic parameters increases to two and three pixels. As in the previous set of tests, 100 simulations are performed for each test case. A summary of the different test conditions is provided in Table 9. For all these
cases, the uncertainty in knowledge of the pose initial guess is the one from Table 5, and illumination conditions are fixed as in Section 4.2.2. Therefore, cases C1 and C2 share the same conditions of test S7, while case E-C1 and E-C2 feature those of case E-SB.

Table 9. Summary of test cases with variable accuracy on the knowledge of the camera intrinsic parameters for both the pose estimation with the chaser-fixed and the eye-in-hand cameras.

| Pose Estimation with Chaser-Fixed Camera | Pose Estimation with Eye-in-Hand Camera |  |  |
| :---: | :---: | :---: | :---: |
| Test <br> Case | Description of Test Conditions | Test <br> Case | Description of Test Conditions |
| C1 | Noise on the camera intrinsic <br> parameters has twice the standard <br> deviations indicated in Table 5. | E-C1 | Noise on the camera intrinsic <br> parameters has twice the standard <br> deviations indicated in Table 5. |
| C2 | Noise on the camera intrinsic <br> parameters has thrice the standard <br> deviations indicated in Table 5. | E-C2 | Noise on the camera intrinsic <br> parameters has thrice the standard <br> deviations indicated in Table 5. |

### 4.2.4. Uncertainty in the Positioning of the Markers on the Target

An additional analysis is carried out to demonstrate the robustness of the proposed pose determination architecture in case of incorrect knowledge of the position vectors of the markers in target coordinates, which can be a consequence of installation errors. To this purpose, the two test cases M1 and E-M1 depicted in Table 10 are introduced, respectively, featuring the illumination conditions of cases S7 and E-SB to allow a straightforward comparison of the results. The mounting error is modelled as a Gaussian noise with null mean and a standard deviation of 1 mm for the coordinates of the position vectors of the retroreflectors, and of 0.3 mm for the coordinates of the position vectors of the BW markers. This is considered a conservative choice, as these values are five times and 1.5 times larger than the uncertainty considered in [21].

Table 10. Summary of test cases with errors in the positioning of the markers on the target for both the pose estimation with the chaser-fixed and the eye-in-hand cameras.

| Pose Estimation with Chaser-Fixed Camera | Pose Estimation with Eye-in-Hand Camera |  |  |
| :---: | :---: | :---: | :---: |
| Test Case | Description of Test Conditions | Test Case | Description of Test Conditions |
| M1 | Error in markers' positioning <br> with standard deviation of 1 mm. | E-M1 | Error in markers' positioning with <br> standard deviation of 0.3 mm. |

## 5. Results

A set of metrics is introduced prior to discussing the results. Besides the error in the estimation of each pose parameter (namely, the relative position vector components and the 321 Euler angles defined in Section 3.1.3), e.g., $\Delta \alpha=\alpha_{e s t}-\alpha_{\text {true }}$, an additional metric considered is the error in the position of the detected markers with respect to their reprojection obtained with the true pose, i.e., $\Delta u$ and $\Delta v$. The mean and standard deviations (Std) of these quantities along the duration of the simulation can be computed. When multiple simulations are performed, the mean $(\mu)$ and standard deviation $(\sigma)$ across the different simulations can also be evaluated at each time instant for each metric. Lastly, even more synthetic statistics can be obtained as the temporal mean $\left(\mu_{N}\right)$ and standard deviation $\left(\sigma_{N}\right)$ of the mean error $\mu$ across multiple simulations. These are shown in (17) and (18) for the generic error parameter $h$, where $n_{s}$ is the number of simulations and $n_{t}$ the number of timesteps. Clearly, $\mu_{N}$ and $\sigma_{N}$ provide a valuable understanding of the average behavior of the pose parameters' errors during the simulation rather than a comprehensive description of the pose estimation accuracy, which can instead be achieved by observing the temporal evolution of $\sigma$ over the course of the simulation. The latter quantity is, therefore, reported as well where necessary, specifically, considering the instantaneous interval of three times the standard deviation $(3 \sigma)$ about the instantaneous mean $\mu$ of the error of a given parameter.

$$
\begin{equation*}
\mu_{N}^{h}=\frac{1}{n_{t}} \sum_{t=1}^{n_{t}} \mu^{h}(t), \mu^{h}(t)=\frac{1}{n_{s}} \sum_{k=1}^{n_{s}} h_{k}(t), \quad t=1, \ldots, n_{t} \tag{17}
\end{equation*}
$$

$$
\begin{equation*}
\sigma_{N}^{h}=\sqrt{\frac{1}{n_{s}-1} \sum_{k=1}^{n_{s}}\left(\mu_{N}^{h}-\mu^{h}(t)\right)^{2}} \tag{18}
\end{equation*}
$$

Concerning the execution of the simulations, both the pose determination architectures are assumed to operate at 5 Hz . The user-defined coefficients and safety margins introduced in the discussions of Section 3 are set as reported in Table 11.

Table 11. Summary of the setting parameters employed for the pose estimation algorithms with both the chaser-fixed and eye-in-hand cameras.

| Setting of Parameters-Pose Estimation with Chaser-Fixed Camera |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Detection and Identification |  |  |  | Pose Estimation |  |  |  |
| $\tau_{\text {bin }}$ | 0.2 | $s_{\text {bord }}$ | 1.2 | $\varepsilon_{0}$ | $10^{-9}$ | $n_{\text {max, }{ }_{\text {it }}}$ | 200 |
| $\tau_{\text {circ }}$ | 0.5 | $s_{i d}$ | 2 | $\varepsilon_{1}$ | $10^{-10}$ | $\lambda_{0}$ | $10^{-8}$ |
| $s_{a o}$ | 0.6 |  |  | $\varepsilon_{2}$ | $10^{-4}$ | $\lambda_{\text {UP }}$ | 11 |
|  |  |  |  | $\varepsilon_{3}$ | $10^{-8}$ | $\lambda_{\text {DN }}$ | 9 |
| Setting of Parameters-Pose Estimation with Eye-in-Hand Camera |  |  |  |  |  |  |  |
| Detection and Identification |  |  |  | Pose Estimation |  |  |  |
| $\tau_{\text {circ }}$ | 0.5 | [ $\left.s_{c u}, s_{c y}\right]$ | [2, 2] | $\varepsilon_{0}$ | $10^{-9}$ | $n_{\text {max, }{ }_{\text {it }}}$ | 200 |
| [ ${ }_{\text {c }} \begin{gathered}s_{a o} \\ {\left[s_{x}, s_{y}\right]}\end{gathered}$ | 0.25 | $\tau_{\text {s }}$ | 0.3 | $\varepsilon_{1}$ | $10^{-10}$ | $\lambda_{0}$ | $10^{-8}$ |
|  | [0.5, 0.5] | $s_{i d}$ | 2 | $\varepsilon_{2}$ | $10^{-4}$ | $\lambda_{\text {UP }}$ | 11 |
|  |  |  |  | $\varepsilon_{3}$ | $10^{-8}$ | $\lambda_{\text {DN }}$ | 9 |

### 5.1. Pose Estimation with the Chaser-Fixed Camera

### 5.1.1. Effect of Illumination Conditions

The results of the test cases defined in Section 4.2.1 for the chaser-fixed camera are here presented. The statistics in Table 12 show that, in general, the pose determination architecture can provide sub-mm-level accuracy in the target/chaser relative position, with a standard deviation larger for the along-boresight component than the cross-boresight ones, which is expected considering that monocular cameras do not provide direct target range information. This performance is attainable thanks to the large number of available 2D-3D correspondences, as well as marker \#10 laying out-ofplane with respect to the rest of the pattern. This latter point allows the sensitivity along the boresight to increase, despite the observation geometry. Table 12 also shows that while the cross-boresight ( $\alpha$ and $\beta$ ) attitude parameters are estimated as accurately as hundredths of a degree, the estimation of the along boresight rotation angle $(\gamma)$ has a better accuracy (up to thousandths of a degree). This result can be motivated considering that the rotation along the boresight axis is more observable, i.e., variations of $\gamma$ produce a larger motion of the detected markers on the image plane than variations of $\alpha$ and $\beta$.

Table 12. Mean and standard deviation of the pose estimation errors for test cases S 1 to S 7 .

| Test Case | Statistic | $\Delta t_{x}(\mathbf{c m})$ | $\Delta t_{y}(\mathbf{c m})$ | $\Delta t_{z}(\mathbf{c m})$ | $\Delta \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| S 1 | $\mu_{N}$ | 0.009 | 0.008 | 0.0004 | 0.007 | 0.0002 | 0.0003 |
|  | $\sigma_{N}$ | 0.007 | 0.007 | 0.056 | 0.014 | 0.009 | 0.0005 |
| S 2 | $\mu_{N}$ | 0.009 | 0.008 | -0.00004 | 0.008 | 0.0007 | 0.0004 |
|  | $\sigma_{N}$ | 0.007 | 0.007 | 0.056 | 0.013 | 0.009 | 0.0006 |
| S 3 | $\mu_{N}$ | 0.005 | 0.009 | 0.030 | 0.018 | -0.014 | -0.0003 |
|  | $\sigma_{N}$ | 0.010 | 0.011 | 0.195 | 0.085 | 0.093 | 0.002 |
| S 4 | $\mu_{N}$ | 0.008 | 0.008 | 0.034 | 0.018 | -0.015 | -0.0001 |
|  | $\sigma_{N}$ | 0.010 | 0.011 | 0.193 | 0.084 | 0.092 | 0.002 |
| S 5 | $\mu_{N}$ | 0.009 | 0.007 | 0.001 | 0.007 | -0.002 | -0.0001 |
|  | $\sigma_{N}$ | 0.007 | 0.007 | 0.056 | 0.013 | 0.009 | 0.0005 |

Table 12. Cont

| Test Case | Statistic | $\Delta t_{\boldsymbol{x}}(\mathrm{cm})$ | $\Delta t_{y}(\mathrm{~cm})$ | $\Delta t_{z}(\mathrm{~cm})$ | $\Delta \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\beta}\left(^{\circ}\right)$ | $\Delta \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| S 6 | $\mu_{N}$ | 0.009 | 0.008 | -0.0005 | 0.010 | -0.0003 | 0.0001 |
|  | $\sigma_{N}$ | 0.007 | 0.007 | 0.0006 | 0.014 | 0.009 | 0.0006 |
| S 7 | $\mu_{N}$ | 0.009 | 0.002 | 0.012 | -0.017 | -0.006 | -0.0006 |
|  | $\sigma_{N}$ | 0.014 | 0.008 | 0.105 | 0.071 | 0.048 | 0.004 |

The temporal evolution of both the mean and standard deviation of the errors on the pose parameters for case S1, depicted in Figure 13, clearly shows that the pose accuracy improves as the chaser moves closer to the target. In particular, the standard deviation of the relative position error undergoes a linear reduction for all its components, coherent with the increase in the pixels' spatial resolution. At the same time, sudden variations in the pose error metrics can be observed towards the end of the simulation, which are mostly caused by the appearance/disappearance of markers from the FOV. For instance, considering test case S1, Figure 13 shows an increase in the standard deviation error for $\gamma$ and $\beta$ when markers \#2 and \#4 fall outside of the camera FOV at 267.6 s and 296 s , respectively. This causes a loss of observability, especially for $\beta$ rotations, since both the markers belong to the $y_{T G F F} /+x_{\text {TGFF }}$ half plane. Similarly, both the mean and standard deviation error in the estimates of $\alpha$ and $t_{z}$ increase due to the disappearance of markers \#6 and \#1 at 329 s and 330 s . This can be explained since marker $\# 6$ is one of the farthest from the $x_{T G F F}$ axis, thus, providing a consistent contribution to the sensitivity in estimating $\alpha$; instead, the loss of marker \#1 leaves only a group of closely placed markers in view, thus, affecting the estimation of $t z$. It is interesting to highlight that neither $t_{x}$ nor $t_{y}$ are affected by the markers' losses, as the number of available points is overabundant for their estimation, given that most correspondences lay in a plane orthogonal to the camera.


Figure 13. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test case S1. The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.

It is worth noting that, although the minimum number of detected markers (i.e., 6) results to be smaller than the nominally expected one (i.e., 8) based on the selected coverage constraint
(see Section 3.1), the number and distribution of the markers still ensure that more than enough correspondences are available at every moment of the approach. As a last remark, the statistics on centroiding errors presented in Table 13 over a single simulation for each test case demonstrate that the algorithm detects the markers' centroids with subpixel accuracy, which strongly contributes to the accuracy reached by the estimated poses.

Table 13. Mean and standard deviation of the centroiding error over all markers, computed over a single simulation for each test case, for test cases S 1 to $\mathrm{S7}$.

| Statistic | Test Cases |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | S1 | S2 | S3 | S4 | S5 | S6 |
| $\Delta u$ |  | 0.002 | -0.001 | -0.011 | 0.002 | 0.008 | 0.008 | 0.015 |
| (pixel) |  | 0.076 | 0.075 | 0.094 | 0.093 | 0.075 | 0.074 | 0.159 |
| $\Delta v$ |  | -0.003 | -0.0002 | -0.005 | -0.013 | -0.006 | -0.002 | -0.024 |
| (pixel) |  | 0.078 | 0.077 | 0.091 | 0.091 | 0.077 | 0.077 | 0.127 |

As for the effect of the ambient illumination, Table 12 shows that, overall, the values of $\mu_{N}$ and $\sigma_{N}$ do not vary consistently from one case to another. An exception is given by test cases $\mathrm{S} 3, \mathrm{~S} 4$ and S7, featuring a slightly larger standard deviation for the estimation error on $t_{z}$ and on the attitude angles. In S3 and S4, this is caused by the larger centroiding error at 320 s characterizing marker \#6, whose blob of pixel-as detected by the image processing approach-is not perfectly circular, as shown in Figure 14. Consequently, $\alpha$ and $\beta$ reach a peak error of $0.5^{\circ}$, while $t_{z}$ reaches a maximum error of about 1 cm , while all the other pose parameters are not affected. This issue is caused by the poor illumination of marker $\# 6$, which is projected very close to the borders of the image plane at the end of the reach and capture maneuver. Although the pose determination performance is robust to this phenomenon, a possible solution is to increase the $s_{b o r d}$ safety margin so that the marker is not considered by the LM-based PnP solver.

As for test case S 7 , a larger error on $t_{z}$ is caused by the partial shadowing of some markers during the approach of the chaser towards the target. In fact, the markers illuminated by both the 850 nm length and the sunlight will appear brighter than the ones shadowed by the chaser body, causing the global threshold to shift towards higher values. As a result, the darkest pixels of those markers that are partially shadowed will have a higher chance of being excluded by the thresholding, introducing errors in the centroiding process. However, the peaks in the pose error (Figure 15), which are coherent to those in the centroiding error (Figure 16), are still kept small in entity.

Overall, the results of the tests S1 to S7 demonstrate that, even though sunlight can influence the detection of the markers, the algorithm is robust to a variety of illumination conditions and is able to preserve its performance.


Figure 14. Cont.

(b)

Figure 14. (a) Temporal evolution of the centroiding errors over a single simulation for test case S3; (b) partial loss of marker \#6 at 320 s .


Figure 15. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test case 57 . The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.


Figure 16. Temporal evolution of the centroiding errors over a single simulation for test cases S 7 and I.

### 5.1.2. Effect of Uncertainty in the Knowledge of the Pose Initial Guess at Scenario Start

Figure 17 depicts the temporal evolution of $\mu$ and $\sigma$ for test case I. The plots clearly show that the algorithm is able to converge to a very accurate solution, even though the initial conditions are far away from the ground-truth. This can also be observed by looking at the simulation statistics summarized in Table 14, which are very similar to those reported in Table 12 for S7. It is also worth noting that, under the larger uncertainty on the pose initial guess, not all the markers are immediately identified at the first timestep. Nonetheless, the initially unmatched markers within the camera FOV are quickly recovered as more accurate pose estimates become available as first guesses in the next timesteps.


Figure 17. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test case I. The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.

Table 14. Mean and standard deviation of the pose estimation errors for a higher uncertainty in the knowledge of the initial condition, for test case I.

| Test Case | Statistic | $\Delta t_{x}(\mathbf{c m})$ | $\Delta t_{y}(\mathbf{c m})$ | $\Delta t_{z}(\mathbf{c m})$ | $\Delta \alpha\left({ }^{\circ}\right)$ | $\Delta \beta\left({ }^{\circ}\right)$ | $\Delta \gamma\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| I | $\mu_{N}$ | 0.010 | 0.003 | 0.011 | -0.016 | -0.006 | -0.0007 |
|  | $\sigma_{N}$ | 0.015 | 0.009 | 0.108 | 0.072 | 0.048 | 0.004 |

These results demonstrate the overall algorithm's robustness to uncertainties in the knowledge of the initial pose guess at the start of the simulation up to $2^{\circ}(1 \sigma)$ in the relative attitude angles, as well as up to $5 \mathrm{~cm}(1 \sigma)$ and $10 \mathrm{~cm}(1 \sigma)$ on the cross-boresight and along the boresight position components, respectively.

### 5.1.3. Effect of Uncertainty in the Knowledge of the Camera Intrinsic Parameters

The results of the tests to verify the effect of increasing uncertainties in the camera intrinsic parameters are summarized in Table 15. Although these show no consistent difference among the two cases (besides a slightly larger mean error for the relative position vector components), further insight is given by the temporal evolution of the mean and standard deviation of the errors on the estimated pose parameters, depicted in Figure 18.

Table 15. Mean and standard deviation of the pose estimation errors for different levels of uncertainty in the knowledge of the intrinsic parameters of the camera for test cases C1 and C2.

| Test Case | Statistic | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{x}}(\mathbf{c m})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{y}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{z}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| C 1 | $\mu_{N}$ | 0.018 | 0.010 | 0.003 | -0.017 | -0.005 | -0.0004 |
|  | $\sigma_{N}$ | 0.016 | 0.008 | 0.104 | 0.072 | 0.047 | 0.004 |
| C 2 | $\mu_{N}$ | 0.026 | 0.018 | -0.005 | -0.016 | -0.005 | -0.0003 |
|  | $\sigma_{N}$ | 0.018 | 0.009 | 0.103 | 0.072 | 0.047 | 0.004 |



Figure 18. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for tests C 1 and C 2 . The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.

Apart from the peaks related to the phenomena discussed while presenting case S7, the graphs show how the variability increases from case C 1 to C 2 as a result of the greater uncertainty on the intrinsic parameters, which propagates to the determination of the minimum of the cost function through the computed normalized undistorted coordinates. At the same time, a slight increase in the bias of the position components corresponds to the increasing uncertainty in the knowledge of the image plane's principal point, which also affects the computation of the normalized undistorted coordinates. No further consideration is given on the accuracy in the detection of the markers' centroids, as these are recovered from the image and are, thus, not directly affected by the intrinsic parameters' uncertainties.

Overall, the tests demonstrate the algorithm's capability to deal with uncertainties of up to three pixels $(1 \sigma)$ in the knowledge of camera focal length and principal point in a robust manner, maintaining good pose determination accuracy.

### 5.1.4. Effect of Uncertainty in the Positioning of the Markers on the Target

The results of the simulations evaluating the effect of the uncertainty in the positioning of the markers on the target object are summarized in Table 16, which shows that while most of the statistics appear similar to those of test case $S 7$ presented in Table 12, the standard deviation of the along-boresight position component $t_{z}$ is slightly larger. The trends of $\mu$ and $\sigma$ during the simulations reported in Figure 19 confirm this observation, but additionally show an increase in the instantaneous variability of the error characterizing all pose parameters, except for $t_{x}$ and $t_{y}$. This latter point can be explained since the detection of up to nine coplanar markers ensures higher robustness in the estimation of the cross-boresight position vector components. Instead, the uncertainty in the markers' position makes the fiducials appear closer or farther than they actually are, thus, affecting the estimation of $t_{z}$. A larger increase in the error on $t_{z}$ is also noted due to the reduction in the number of available point correspondences when markers gradually fall out of the camera's FOV. On the other hand, the attitude errors appear to be equally affected, with a $3 \sigma$ about one order of magnitude larger than in S7.


Figure 19. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test M1. The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.

Table 16. Mean and standard deviation of the pose estimation errors in presence of uncertainties in the positioning of the markers on the target for test case M1.

| Test Case | Statistic | $\Delta t_{x}(\mathrm{~cm})$ | $\Delta t_{y}(\mathrm{~cm})$ | $\Delta t_{z}(\mathrm{~cm})$ | $\Delta \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| M 1 | $\mu_{N}$ | -0.008 | 0.003 | 0.017 | -0.0003 | 0.0003 | -0.001 |
|  | $\sigma_{N}$ | 0.015 | 0.009 | 0.129 | 0.087 | 0.046 | 0.005 |

In conclusion, the results show that the presence of a large uncertainty in the positioning of the markers can lead to an increase in the errors up to cm-level for the position components and degree-level for the attitude angles. Nevertheless, the pose estimation process can still provide measurements at mm-level in the cross-boresight position components, as well as at sub-degree level in all attitude angles, confirming the robustness of the proposed method.

### 5.2. Pose Estimation with the Eye-In-Hand Camera

The results of the simulations conducted to evaluate the performance of the algorithm for pose estimation with the eye-in-hand camera are here detailed.

### 5.2.1. Effect of Illumination Conditions

The results of the 100 simulations conducted for test case E-SB, represented in Figure 20, are analyzed first and employed as a baseline. The simulations are conducted in the same fashion as those for the algorithm employing the chaser-fixed camera, assuming the variability on the pose initial guess and camera intrinsic parameters summarized in Table 5. The algorithm demonstrates sub-mm accuracy in the determination of the position components of vector $\underline{t}_{C S F F a r m} \rightarrow T A P F$, as well as errors in the order of the hundredths of degrees in estimating the relative attitude. The graphs also show how the variability in the pose determination accuracy strongly reduces after the first instants of the approach, with the accuracy in the determination of the position components improving as the chaser approaches the target, and the attitude determination accuracy almost immediately reaching small errors. The summary statistics reported in Table 17 show the effect of observing the pattern almost orthogonally. Specifically, $t_{x}, t_{y}$ and $\gamma$ are estimated with higher sensitivity and, thus, are characterized by smaller standard deviations compared to the other parameters. Nonetheless, a smaller mean error affecting the $t_{y}$ component compared to $t_{x}$ can be observed, which is coherent with the adopted pattern of markers being aligned along the $x_{T A P F}$ direction (see Figure 9) and, thus, limiting the sensitivity along this axis. Finally, the temporal evolution of the centroiding error for each marker during a single simulation, reported in Figure 21, shows that the detection and identification steps allow the markers' location in the image plane to be correctly determined with sub-pixel accuracy, with an average detection error over all markers of 0.059 and 0.009 pixels along the horizontal and vertical axes, respectively, and corresponding standard deviations of 0.270 and 0.075 pixels. The graph also shows that the centroiding error on the horizontal coordinate of all the markers ramps up to pixel level during the last moments of the approach. This increase is due to the very close range at which the markers are imaged at the end of the approach. In fact, as the camera moves closer to the set of markers, their dimension on the image plane increases, reaching up to 109.63 pixels in diameter at 0.4 m of separation, and a larger error is expected. Moreover, it should be noted that, at such separation, the spatial resolution of a single pixel corresponds to 0.17 mm , which shows that despite the small separation, accurate information on the markers' centroids is still retrieved. This is confirmed by the errors in Figure 20, which do not show signs of degradation of the pose estimate at the corresponding instants.


Figure 20. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test E-SB. The plots are focused on intervals about the mean values: the arrows indicate the maximum values reached by the pose initial guess at scenario start.


Figure 21. Temporal evolution of the centroiding errors over a single simulation for test case E-SB.

Table 17. Mean and standard deviation of the pose estimation errors for test case E-SB.

| Test Case | Statistic | $\Delta t_{\boldsymbol{x}}(\mathrm{cm})$ | $\Delta t_{y}(\mathrm{~cm})$ | $\Delta t_{z}(\mathrm{~cm})$ | $\Delta \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\Delta \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{E}-\mathrm{SB}$ | $\mu_{N}$ | 0.002 | 0.0003 | 0.031 | -0.004 | -0.007 | -0.0004 |
|  | $\sigma_{N}$ | 0.005 | 0.003 | 0.061 | 0.019 | 0.019 | 0.006 |

The results concerning the remainder of the tests with variable illumination conditions are graphically summarized in Figure 22, showing that the pose estimation with the eye-in-hand camera correctly detects the markers on the scene at every instant of the approach when the Sun illuminates them from above the camera (cases indicated in green). In fact, in those cases, the Sun-chaser-sensor geometry prevents the markers from being occluded by the chaser's shadow. This opposes the other cases in which part of the markers are shadowed from a certain instant during the simulation, not allowing the algorithm to detect them (cases indicated in yellow), or the whole set of markers is obscured from a certain timestep (cases indicated in red). The use of an artificial illuminator would prevent the issue, ensuring the possibility of approaching the target at any epoch. As for the accuracy in pose determination, the simulations show that in all the cases in which the approach was concluded without interruptions in the pose solution, the same level of accuracy as the one observed for case E-SB was maintained. On the other hand, as expected, all the cases in which one marker of the pattern was lost (either partially or completely) present larger errors, which can reach centimeter levels in position (especially along boresight), as well as degree levels in attitude. Moreover, it is worth recalling that a number of correspondences smaller than three leads to an ambiguous formulation of the PnP problem. Nonetheless, these results highlight that the algorithm is still capable of computing a pose solution thanks to the available initial condition allowing disambiguation. However, this is only possible if the evolution of the end-effector's pose parameters does not lead the initial guess too far away from the local minimum of the solution, i.e., the manipulator's dynamics must be smooth enough to allow the pose solution to be tracked by the LM. As a final remark, it is worth mentioning that some of the cases in which markers are lost also feature mismatches caused by the scarce natural illumination: this issue can be removed by tuning the detection and identification parameters.


Figure 22. Qualitative representation of the locations of the Sun in TROF for test cases E-S1 to E-S36 ( $z_{\text {TGFF }}$ exits the figure's plane) and their results. In green: cases in which the camera correctly detects all markers at each timestep. In yellow: cases in which part of the markers are lost from a certain timestep. In red: cases in which all the markers are lost from a certain timestep.

### 5.2.2. Effect of Uncertainty in the Knowledge of the Pose Initial Guess at Scenario Start

This set of tests shows that the algorithm is more sensitive to the uncertainty on the pose initial guess with respect to the one employed with the chaser-fixed camera, as shown by the larger standard deviation errors illustrated in Figure 23. However, the larger variability is also ascribable to the fact that in two simulations out of 100 , the error on the pose initial guess is large enough to cause mismatches at the first timestep, which in turn leads to falling in a local minimum of the cost function (9). The computed pose, therefore, has a larger error on the $t_{x}$ component, which affects both the matching and the computation of the RoI at the next timestep, and causes one of the markers to fall out of the RoI farther in the simulation, as shown in Figure 24.


Figure 23. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test case E-I.


Figure 24. Consequences of the larger uncertainty on initial conditions, for test case E-I (in blue the RoI): (a) marker mismatch at simulation start caused by the larger uncertainties on the pose initial guess; (b) marker cut out of the RoI, at $t=322.10 \mathrm{~s}$.

Nonetheless, it is worth noting that the tests performed in these simulations addressed worstcase conditions, as assuming up to 3 cm and $2^{\circ}$ errors on the pose initial guess is over-conservative for a camera-target separation of 1.5 m . Moreover, the pose estimation by the eye-in-hand camera is expected to work conjointly to the one employing the images of the chaser-fixed camera, exploiting its pose solution to compute its own initial guess. Finally, the mean and standard deviation of the average error on the estimated pose parameters, shown in Table 18, highlight that most of the parameters preserved their accuracy, although an increase in the mean error on $t_{z}$ could be observed,
leading to a bias related to the cases affected by the mismatches issues. Therefore, the discussed results confirm that the algorithm is overall robust to a wide variety of initial conditions.

Table 18. Mean and standard deviation of the pose estimation errors for different levels of uncertainty in the knowledge of the initial condition, for test case E-I.

| Test Case | Statistic | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{x}}(\mathbf{c m})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{y}}(\mathbf{c m})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{z}}(\mathbf{c m})$ | $\boldsymbol{\Delta} \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{E}-\mathrm{I}$ | $\mu_{N}$ | 0.002 | 0.002 | 0.130 | -0.003 | -0.008 | 0.001 |
|  | $\sigma_{N}$ | 0.006 | 0.003 | 0.062 | 0.019 | 0.016 | 0.010 |

### 5.2.3. Effect of Uncertainty in the Knowledge of the Camera Intrinsic Parameters

The tests for the assessment of the effect of increasing uncertainties on the intrinsic parameters are affected by the same phenomena observed in Section 5.1.3. In fact, the temporal evolutions of the mean and standard deviation of the pose parameters' errors, reported in Figure 25, show a similar increase in their variability from case E-C1 to E-C2. Nonetheless, the results show that sub-mm accuracy in estimating the position components is maintained, as well as errors in the order of hundredths of degrees in the determination of the attitude angles, as confirmed by the statistics presented in Table 19. Overall, the algorithm demonstrated robustness to errors up to three pixels $(1 \sigma)$ in the knowledge of camera focal length and principal point.


Figure 25. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for tests E-C1 and E-C2.

Table 19. Mean and standard deviation of the pose estimation errors for different levels of uncertainty in the knowledge of the camera intrinsic parameters, for test cases E-C1 and E-C2.

| Test Case | Statistic | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{x}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{y}}(\mathbf{c m})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{z}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\beta}\left(^{\circ}\right)$ | $\Delta \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{E}-\mathrm{C} 1$ | $\mu_{N}$ | 0.004 | 0.003 | 0.026 | -0.003 | -0.007 | -0.001 |
|  | $\sigma_{N}$ | 0.005 | 0.003 | 0.060 | 0.019 | 0.017 | 0.007 |
| $\mathrm{E}-\mathrm{C} 2$ | $\mu_{N}$ | 0.005 | 0.004 | 0.025 | -0.002 | -0.007 | -0.001 |
|  | $\sigma_{N}$ | 0.005 | 0.004 | 0.060 | 0.019 | 0.017 | 0.007 |

### 5.2.4. Effect of Uncertainty in the Positioning of the Markers on the Target

The results of the simulations for test case E-M1 highlight overall pose estimation errors comparable to those observed in the previous cases. Specifically, Table 20 shows that the standard deviations $\sigma_{N}$ are overall comparable to those observed for test case E-SB in Table 17, although $t_{x}, \alpha$ and $\gamma$ feature average errors that are at least one order of magnitude larger. Nevertheless, these errors remain at sub-mm and sub-degree levels. The comparison between the temporal evolutions of $\mu$ and $\sigma$ for the case under analysis (Figure 26) and for case E-SB (Figure 20) provides a greater understanding of the effect of the uncertainty in the 3 D reference location of the markers on the pose estimation using the eye-in-hand camera. Specifically, the graphs show that the instantaneous standard deviations of the errors on the estimated pose parameters are only slightly increased for the position components (keeping the linear decreasing behavior while the camera approaches the target). Instead, the variability in the error on the attitude angles features an increase of about one order of magnitude compared to case E-SB and a constant trend over the entire simulation period.


Figure 26. Temporal evolution of the mean of the errors on the pose parameters, with representation of the $3 \sigma$ intervals, for test E-M1.

Table 20. Mean and standard deviation of the pose estimation errors in presence of uncertainties in the positioning of the markers on the target, for test case E-M1.

| Test Case | Statistic | $\Delta t_{\boldsymbol{x}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{y}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{t}_{\boldsymbol{z}}(\mathrm{cm})$ | $\boldsymbol{\Delta} \boldsymbol{\alpha}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\beta}\left({ }^{\circ}\right)$ | $\boldsymbol{\Delta} \boldsymbol{\gamma}\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{E}-\mathrm{M} 1$ | $\mu_{N}$ | 0.012 | -0.0003 | 0.026 | -0.027 | -0.005 | -0.03 |
|  | $\sigma_{N}$ | 0.005 | 0.003 | 0.059 | 0.019 | 0.018 | 0.006 |

Overall, these simulations show that the pose estimation process with images from the eye-in-hand camera is robust against uncertainty in the knowledge of the markers position due to installation errors, as it preserves acceptable accuracy up to mm and degree levels on the position and attitude parameters, respectively. The greater effect of such uncertainty on the achieved pose accuracy (compared to the case of the chaser fixed camera) is justified by the fewer number of markers (the pattern of fiducials being not redundant).

## 6. Conclusions

This paper proposed original approaches for image processing and pose determination using monocular cameras to support the servicing of a semi-collaborative space target in GEO equipped with fiducial markers and using a chaser with a robotic arm. The presented architecture relied on two cameras that were rigidly attached to the chaser's main body (chaser-fixed) and to the robotic arm's end-effector (eye-in-hand). It featured dedicated algorithmic solutions to estimate the relative position and attitude of the chaser's body and end-effector with respect to the target's main body and grasping point, respectively. The definition of the specific algorithms to be employed with the two cameras was accompanied by the identification of the most suitable type of markers, as well as the indication of the technical specifications of each camera. These were defined through a specific procedure accounting for coverage and resolution constraints. Both algorithms exploited an original implementation of the Levenberg-Marquardt's non-linear least squares method for the accurate estimation of the pose parameters. A large variety of numerical simulations was conducted for the performance assessment reproducing a reach and capture scenario (with a maximum target/chaser distance of 10 m ) and using the ESA tool PANGU for the realistic generation of synthetic images produced by the visual sensors. The proposed approaches demonstrated the capability of achieving up to sub-mm and hundredths of a degree accuracies in relative position and attitude estimates, respectively. At the same time, their robustness was demonstrated against the coarse initialization of the pose parameters at scenario start, i.e., considering a zero mean Gaussian uncertainty ( $1 \sigma$ ) up to 10 cm for the along-boresight separation and up to $3^{\circ}$ for the relative attitude, as well as uncertainties in the knowledge of the camera intrinsic parameters up to three pixels $(1 \sigma)$ and errors in the positioning of the fiducials on the target spacecraft with a standard deviation of 1 mm and 0.3 mm for the retroflectors and BW markers, respectively. Finally, the comprehensive analysis of the effect of the illumination conditions showed that the solution based on corner-cube retroreflectors adopted for the pose estimation using the chaser-fixed camera is robust to highly variable observation geometries. Moreover, the algorithm employing the eye-in-hand camera was able to successfully estimate the pose without any onboard illumination source, provided that the mission starting epoch was selected to ensure a favorable observation geometry. While all the achieved results constitute a valid proof-ofconcept for the proposed approach, future works will address the coding of the proposed algorithms in C++ and their execution on an embedded processing board, in order to allow software in the loop and hardware in the loop tests, which can be used to evaluate the computational effort and demonstrate real time capabilities. Clearly, the hardware in the loop tests will also require a dedicated laboratory environment, ad hoc calibrated to ensure the availability of an accurate ground truth pose solution (e.g., using motion tracking systems). With foresight regarding the future hardware implementation of the proposed approach, another key point which needs further investigation is related to the feasibility of the installation of a redundant number of fiducials for the operation of the chaser-fixed camera, considering that each additional marker will increase manufacturing, testing and verification constraints, and, consequently, the cost of the client spacecraft. Although the proposed approach would still be applicable if fewer markers were used, this choice would be paid for in terms of a reduced pose estimation accuracy and robustness. In this respect, a careful trade-off analysis must be carried out during mission study to find the best compromise between relative navigation requirements and target-related constraints.


#### Abstract

Author Contributions: Conceptualization, R.O. and C.L.; methodology, R.O., C.L. and C.V.; software, C.L. and C.V.; validation, R.O., A.N. and C.V.; formal analysis, R.O., A.N. and C.V.; investigation, R.O., A.N. and C.V.; resources, R.O.; data curation, C.V.; writing-original draft preparation, R.O. and C.V.; writing-review and editing, R.O., C.V., A.N. and C.L.; supervision, R.O.; project administration, R.O.; funding acquisition, R.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. Acknowledgments: The authors would like to thank Irene Huertas Garcia and Pedro Simplicio of the European Space Agency for the fruitful discussions and interactions during the progress and review meetings of the GRACC project.


Conflicts of Interest: The authors declare no conflict of interest.

## References

1. Stoll, E.; Letschnik, J.; Walter, U.; Artigas, J.; Kremer, P.; Preusche, C.; Hirzinger, G. On-orbit servicing. IEEE Robot. Autom. Mag. 2009, 16, 29-33. [CrossRef]
2. Flores-Abad, A.; Ma, O.; Pham, K.; Ulrich, S. A review of space robotics technologies for on-orbit servicing. Prog. Aerosp. Sci. 2014, 68, 1-26. [CrossRef]
3. Bonnal, C.; Ruault, J.M.; Desjean, M.C. Active debris removal: Recent progress and current trends. Acta Astronaut. 2013, 85, 51-60. [CrossRef]
4. Peters, S.; Pirzkall, C.; Fiedler, H.; Förstner, R. Mission concept and autonomy considerations for active Debris removal. Acta Astronaut. 2016, 129, 410-418. [CrossRef]
5. Li, W.-J.; Cheng, D.-Y.; Liu, X.-G.; Wang, Y.-B.; Shi, W.-H.; Tang, Z.-X.; Gao, F.; Zeng, F.-M.; Chai, H.-Y.; Luo, W.-B.; et al. On-orbit service (OOS) of spacecraft: A review of engineering developments. Prog. Aerosp. Sci. 2019, 108, 32-120. [CrossRef]
6. McKnight, D.; Witner, R.; Letizia, F.; Lemmens, S.; Anselmo, L.; Pardini, C.; Rossi, A.; Kunstadter, C.; Kawamoto, S.; Aslanov, V.; et al. Identifying the 50 statistically-most-concerning derelict objects in LEO. Acta Astronautica 2021, 181, 282-291. [CrossRef]
7. Starek, J.A.; Açıkmeşe, B.; Nesnas, I.A.; Pavone, M. Spacecraft Autonomy Challenges for Next-Generation Space Missions. In Advances in Control System Technology for Aerospace Applications; Lecture Notes in Control and Information Sciences; Feron, E., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; Volume 460. [CrossRef]
8. Pavanello, Z.; Branz, F.; Francesconi, A.; Cenedese, A.; Antonello, R.; Basana, F.; Iob, P.; Vertuani, D.; Massari, M.; Colombo, C.; et al. Combined control and navigation approach to the robotic capture of space vehicles. In Proceedings of the 72nd International Astronautical Congress (IAC 2021), Dubai, United Arab Emirates, 25-29 October 2021.
9. Liu, Y.; Zhao, Y.; Tan, C.; Liu, H.; Liu, Y. Economic value analysis of on-orbit servicing for geosynchronous communication satellites. Acta Astronaut. 2021, 180, 176-188. [CrossRef]
10. Sarton du Jonchay, T.; Chen, H.; Gunasekara, O.; Ho, K. Framework for Modeling and Optimization of On-Orbit Servicing Operations under Demand Uncertainties. J. Spacecr. Rocket. 2021, 58, 1157-1173. [CrossRef]
11. Northrop Grumman and Intelsat Make History with Docking of Second Mission Extension Vehicle to Extend Life of Satellite. Available online: https:/ /news.northropgrumman.com/news/releases/northrop-grumman-and-intelsat-make-history-with-docking-of-second-mission-extension-vehicle-to-extend-life-of-satellite (accessed on 16 February 2022).
12. As Science Fiction Became Reality ... 36.000 km above Earth. Available online: https:/ / www.jena-optronik.de/en/news/reader/ docking-mev-2-is-10-02.html (accessed on 16 February 2022).
13. Clerc, X.; Retat, I. Astrium vision on space debris removal. In Proceedings of the 63rd International Astronautical Congress (IAC 2012), Napoli, Italy, 1-5 October 2012.
14. Lepetit, V.; Moreno-Noguer, F.; Fua, P. EPnP: An accurate O(n) solution to the PnP problem. Int. J. Comput. Vis. 2009, 81, 155-166. [CrossRef]
15. Hesch, J.A.; Roumeliotis, S.I. A direct least-squares (DLS) method for PnP. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6-13 November 2011; pp. 383-390.
16. Li, S.; $\mathrm{Xu}, \mathrm{C} . ; \mathrm{Xie}, \mathrm{M}$. A robust $\mathrm{O}(\mathrm{n})$ solution to the perspective-n-point problem. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 34, 1444-1450. [CrossRef]
17. Howard, R.T.; Heaton, A.F.; Pinson, R.M.; Carrington, C.K. Orbital express advanced video guidance sensor. In Proceedings of the 2008 IEEE Aerospace Conference, Big Sky, MT, USA, 1-8 March 2008.
18. Wen, Z.; Wang, Y.; Luo, J.; Kuijper, A.; Di, N.; Jin, M. Robust, fast and accurate vision-based localization of a cooperative target used for space robotic arm. Acta Astronaut. 2017, 136, 101-114. [CrossRef]
19. Opromolla, R.; Fasano, G.; Rufino, G.; Grassi, M. A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations. Prog. Aerosp. Sci. 2017, 93, 53-72. [CrossRef]
20. Benn, M.; Jørgensen, J.L. Short range pose and position determination of spacecraft using a $\mu$-advanced stellar compass. In Proceedings of the 3rd International Symposium on Formation Flying, Missions and Technologies, Noordwijk, The Netherlands, 23-25 April 2008.
21. Sansone, F.; Branz, F.; Francesconi, A. A relative navigation sensor for CubeSats based on LED fiducial markers. Acta Astronaut. 2018, 146, 206-215. [CrossRef]
22. Mokuno, M.; Kawano, I. In-orbit demonstration of an optical navigation system for autonomous rendezvous docking. J. Spacecr. Rocket. 2011, 48, 1046-1054. [CrossRef]
23. Heaton, A.; Howard, R.; Pinson, R. Orbital express AVGS validation and calibration for automated rendezvous. In Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Honolulu, HI, USA, 18-21 August 2008.
24. Pochard, M. Qualification of videometer an optical rendezvous sensor. In Proceedings of the 54th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law (pp. AP-07), Bremen, Germany, 29 September-3 October 2003.
25. Hariri, N.; Gutierrez, H.; Rakoczy, J.; Howard, R.; Bertaska, I. Performance characterization of the smartphone video guidance sensor as vision-based positioning system. Sensors 2020, 20, 5299. [CrossRef] [PubMed]
26. Pemsun: Passive Emitting Material Supporting Navigation at End-of-Life. Available online: https:/ /blogs.esa.int/cleanspace/20 19/04/17/infrared-and-phosphorescent-markers-to-avoid-future-satellites-to-become-debris/ (accessed on 29 April 2022).
27. Baca, T.; Stepan, P.; Spurny, V.; Hert, D.; Penicka, R.; Saska, M.; Thomas, J.; Loianno, G.; Kumar, V. Autonomous landing on a moving vehicle with an unmanned aerial vehicle. J. Field Robot. 2019, 36, 874-891. [CrossRef]
28. Gatrell, L.B.; Hoff, W.A.; Sklair, C.W. Robust image features: Concentric contrasting circles and their image extraction. In Proceedings of the Cooperative Intelligent Robotics in Space II, Boston, MA, USA, 14-15 November 1991; Volume 1612, pp. 235-244. [CrossRef]
29. Goodwin, D.P.; Hembree, L.E.; Curran, J.P.; Moyer, D.S.; Strachan, R.L.; Mills, I.; Valois, J.S. Orbiter space vision system on space shuttle flight STS-80. In Proceedings of the Visual Information Processing VI, AeroSense '97, Orlando, FL, USA, 21-25 April 1997; Volume 2074. [CrossRef]
30. Tweddle, B.E.; Saenz-Otero, A. Relative computer vision-based navigation for small inspection spacecraft. J. Guid. Control. Dyn. 2015, 38, 969-978. [CrossRef]
31. Pertile, M.; Chiodini, S.; Giubilato, R.; Mazzucato, M.; Valmorbida, A.; Fornaser, A.; Lorenzini, E.C. Metrological Characterization of a Vision-Based System for Relative Pose Measurements with Fiducial Marker Mapping for Spacecrafts. Robotics 2018, 7, 43. [CrossRef]
32. Fiala, M. ARTag Fiducial Marker System Applied to Vision Based Spacecraft Docking. In Proceedings of the IEEE/RSJ IROS 2005 Workshop on Robot Vision for Space Applications, Edmonton, AB, Canada, 2-6 August 2005.
33. Bondy, M.; Krishnasamy, R.; Crymble, D.; Jasiobedzki, P. Space vision marker system (SVMS). In Proceedings of the AIAA SPACE 2007 Conference \& Exposition, Long Beach, CA, USA, 18-20 September 2007.
34. Chan, M.; Bultitude, J.; Faber, D.; Hawes, D.; Fab, O. Productization of CubeSat Rendezvous and Docking Solutions. In Proceedings of the 33rd Annual AIAA/USU Conference on Small Satellites, Logan, UT, USA, 3-8 August 2019.
35. General Mission Analysis Tool. Available online: https:/ / opensource.gsfc.nasa.gov / projects/GMAT / (accessed on 3 May 2022).
36. MATLAB Robotics System Toolbox. Available online: https://it.mathworks.com/products/robotics.html (accessed on 3 May 2022).
37. Planet and Asteroid Natural Scene Generation Utility (PANGU), by STAR-Dundee. Available online: https:/ /www.star-dundee. com/products/pangu-planet-and-asteroid-natural-scene-generation-utility / (accessed on 3 May 2022).
38. 1300-Class Platform I Commercial Spacecraft Platform. Available online: https://www.maxar.com/products/1300-classplatform (accessed on 12 June 2022).
39. David, P.; Dementhon, D.; Duraiswami, R.; Samet, H. SoftPOSIT: Simultaneous pose and correspondence determination. Int. J. Comput. Vis. 2004, 59, 259-284. [CrossRef]
40. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 1981, 24, 381-395. [CrossRef]
41. 4 mm Silver and NIR II Coated, Corner Cube Prims I Edmund Optics. Available online: https:/ /www.edmundoptics.com/p/25 4 mm -silver-and-nir-ii-coated-corner-cube-prism/41585/ (accessed on 30 May 2022).
42. BGR 25.4 Retroreflector. Available online: https://www.newport.com/p/BGR-25.4 (accessed on 30 May 2022).
43. Howard, R.T.; Cole, H.J.; Jackson, J.L.; Kamerman, G.W.; Fronek, D.K. Automatic rendezvous and docking system test and evaluation. In Proceedings of the Laser Radar Technology and Applications II 1997, Orlando, FL, USA, 21-25 April 1997; Volume 3065, pp. 131-139.
44. Rank, P.; Mühlbauer, Q.; Naumann, W.; Landzettel, K. The DEOS Automation and Robotics Payload. In Proceedings of the 11th Symposium on Advanced Space Technologies in Robotics and Automation (ASTRA), Noordwijk, The Netherlands, 1-8 April 2011.
45. Jaekel, S.; Lampariello, R.; Rackl, W.; De Stefano, M.; Oumer, N.; Giordano, A.M.; Porges, O.; Pietras, M.; Brunner, B.; Ratti, J.; et al. Design and Operational Elements of the Robotic Subsystem for the e.deorbit Debris Removal Mission. Front. Robot. AI 2018, 5, 100. [CrossRef] [PubMed]
46. Chabot, T.; Kanani, K.; Pollini, A.; Chaumette, F.; Marchand, E.; Forshaw, J. Vision-based navigation experiment onboard the RemoveDebris mission. In Proceedings of the 10th International ESA Conference on Guidance, Navigation \& Control Systems, Salzburg, Austria, 29 May-2 June 2017; pp. 1-23.
47. Naasz, B.J.; Burns, R.D.; Queen, S.Z.; Van Eepoel, J.; Hannah, J.; Skelton, E. The HST SM4 Relative Navigation Sensor System: Overview and Preliminary Testing Results from the Flight Robotics Lab. J. Astronaut. Sci. 2009, 57, 457-483. [CrossRef]
48. Liu, G.; Xu, C.; Zhu, Y.; Zhao, J. Monocular Vision-Based Pose Determination in Close Proximity for Low Impact Docking. Sensors 2019, 19, 3261. [CrossRef] [PubMed]
49. Benn, M. Vision Based Navigation Sensors for Spacecraft Rendezvous and Docking. Ph.D. Thesis, Danish Technical University, Lyngby, Denmark, 31 October 2010.
50. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man Cybern. 1979, 9, 62-66. [CrossRef]
51. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 1963, 11, 431-441. [CrossRef]
52. D'Amico, S.; Benn, M.; Jørgensen, J.L. Pose estimation of an uncooperative spacecraft from actual space imagery. Int. J. Space Sci. Eng. 2014, 2, 171-189.
53. Gavin, H.P. The Levenberg-Marquardt Algorithm for Nonlinear Least Squares Curve-Fitting Problems; Department of Civil and Environmental Engineering, Duke University: Durham, NC, USA, 2019.
54. Camera Calibration Toolbox for Matlab. Available online: http://robots.stanford.edu/cs223b04/JeanYvesCalib/htmls/ parameters.html (accessed on 17 June 2022).
55. Bradley, D.; Roth, G. Adaptive thresholding using the integral image. J. Graph. Tools 2007, 12, 13-21. [CrossRef]
56. Pinard, D.; Reynaud, S.; Delpy, P.; Strandmoe, S.E. Accurate and autonomous navigation for the ATV. Aerosp. Sci. Technol. 2007, 11, 490-498. [CrossRef]
57. Becker, C.; Howard, R.T.; Rakoczy, J. Smartphone Video guidance sensor for small satellites. In Proceedings of the 27th AIAA/USU Conference on Small Satellites, Logan, UT, USA, 10-15 August 2013.
