(Preprint) AAS 24-392

MULTIVIEW OPTICAL NAVIGATION FOR SPACE MANIPULATOR SYSTEMS DURING IN-ORBIT SERVICING MISSIONS

Niccolò Faraco[*], Michele Maestrini[†], Mauro Massari[‡], and Pierluigi Di Lizia[§]

In-orbit servicing missions will play a substantial role in the aerospace industry in the upcoming decades. These missions could potentially revive numerous satellites through refueling or minor repairs, thereby reactivating valuable assets in space. Among the various techniques envisioned as possible solutions to the problem, using Space Manipulator Systems (SMS), i.e. servicing spacecrafts mounting manipulators on-board, is particularly interesting, as it offers great flexibility. Designing this kind of missions, however, is often challenging. As the system moves in close proximity of the target spacecraft, correctly estimating its position and attitude is paramount in order to avoid undesirable collisions. The task is particularly arduous when the target is uncooperative or only passively cooperative and, therefore, relative navigation has to be performed relying on vision systems. On this matter, studies have also shown that relying on the direct kinematics of the manipulator to reconstruct the pose of the end-effector as a function of the spacecraft base pose and joint variables of the arm lacks robustness and does not meet the error thresholds required to enable this kind of technology. This problem can be mitigated using the direct measurements of a sensor mounted in the vicinity of the end-effector of the manipulator.

The use of LIDAR sensors and stereo cameras has extensively being investigated, in the last few years, as an option to tackle the problem. While such kinds of sensors necessitate higher power and mass budgets, opting for monocular cameras can mitigate this issue, albeit at the cost of reduced accuracy. This paper proposes a novel approach in which an UnScented QUaternion Estimator (USQUE) is used to concurrently perform non-linear filtering on the attitude and position of both the base of the chasing spacecraft and the end-effector of the manipulator within a fully relative framework. This is achieved exclusively through information derived from two cameras mounted, respectively, on the spacecraft base and the manipulator end-effector, effectively creating a variable-baseline stereo camera system. Also, the dynamical model of the system embedded in the filter directly includes states representing the position and attitude of the end-effector, effectively tackling the shortcomings of the robustness of their kinematic reconstruction. The results demonstrate the efficacy of the method, showcasing high accuracy in the reconstruction of the position and orientation of both the reference frames of interest.

As a secondary outcome for this study, a high-fidelity model of the SMS is generated using Modelica and employed to obtain the simulated reference trajectory against which the filter is tested. The model incorporates a fully relative representation of gravitational acceleration in the Local-Vertical Local-Horizontal (LVLH) frame centered on the target. This approach circumvents potential numerical challenges and preserves the possibility of incorporating gravitational perturbations into the modeling, a consideration often overlooked in the related literature.

[*]Ph.D. student, DAER department, Politecnico di Milano, Via La Masa 34 Milano.
[†]Assistant Professor, DAER department, Politecnico di Milano, Via La Masa 34 Milano.
[‡]Associate Professor, DAER department, Politecnico di Milano, Via La Masa 34 Milano.
[§]Associate Professor, DAER department, Politecnico di Milano, Via La Masa 34 Milano.

INTRODUCTION

Designing new space missions continues to face significant challenges. One of the most pressing issues today is the overcrowding of heavily used orbit regimes due to the proliferation of man-made Resident Space Objects (RSOs). This increases the risk to both new and existing satellites and complicates access to space for manned missions. Although various techniques, such as the use of graveyard orbits and controlled atmospheric reentry, have been developed and recommended to mitigate this problem,[1,2] there is a crucial need to develop In-Orbit-Servicing (IOS) missions and Active Debris Removal (ADR) services for the refueling and repairing of otherwise functioning assets and the de-orbiting of irremediably compromised ones.
Central to these missions are robotic manipulator systems mounted on the servicing spacecraft, often called chaser, which offer flexibility and precision for complex tasks in the harsh environment of space. The development of these systems, usually referred to as Space Manipulator Systems (SMSs), has benefited from significant advancements in robotics, control systems, and space technology. Initial implementations, such as the Shuttle Remote Manipulator System (SRMS) or Canadarm, showcased the feasibility of using robotic arms for satellite deployment and retrieval missions.[3] These systems paved the way for more advanced manipulators like the European Robotic Arm (ERA) and the Canadian Space Agency's Dextre, both crucial to maintaining the International Space Station (ISS). These robotic systems have been extensively documented, emphasizing their contributions to mission success through enhanced dexterity and operational reach in microgravity environments.[4–6]

Integrating autonomous systems in IOS missions, however, is crucial due to communication delays and finite visibility time windows associated with spacecraft control from ground. Autonomous navigation and control systems can enhance mission efficiency and reduce human error risks, as already demonstrated by test flights.[7–9] Enabling spacecraft to make real-time decisions and adjustments ensures more reliable and accurate execution of complex tasks in space's unpredictable environment,[10] especially when moving in close proximity to another object, when unwanted collisions may cause catastrophic consequences.
A critical challenge in IOS missions is, in fact, the navigation and control of the robotic manipulator during the final approach phase to the target object. Accurate positioning and orientation are essential for successful capture and manipulation. Traditional navigation techniques rely on sensor data, including LIDAR, radar, and stereo-vision systems, to achieve precise relative positioning. Despite advancements in sensor technology and image processing, navigation systems during the final approach phase have several shortcomings.

LIDAR and radar systems provide high accuracy and reliability but come with significant power and computational demands.[11,12] Additionally, these systems can be bulky, increasing the overall weight of the spacecraft, a crucial consideration for space missions. Also, they are susceptible to interference and may not perform optimally in cluttered environments or with complex target geometries.[13] Stereo vision systems offer a more compact, lower-power solution but require complex image processing algorithms to extract depth information, which can be computationally intensive and prone to errors under varying lighting conditions and in the presence of space debris.[14–17] They also require precise calibration to maintain accuracy, a rather challenging task in the dynamic environment of space and IOS missions specifically.[18–20]
Furthermore, as a general remark, reliance on a single sensor modality can be limiting. Single-sensor systems lack redundancy and robustness, making them vulnerable to failures or malfunctions, particularly critical in space missions where repair and replacement are infeasible. Moreover, even

when fully operative, this kind of configuration is not free from limitations. Studies have shown that the positioning error thresholds required to enable IOS missions can not be met when using the vision-based system to estimate the position of the base of the chaser with respect to the target and relying solely on the manipulator's direct kinematics to reconstruct the pose of the end-effector (EE).[21]
Innovative approaches combining multiple sensors and sensing modalities are needed to enhance navigation system reliability and accuracy during the final approach phase.

Monocular cameras present a promising alternative to traditional sensors due to their lightweight, low power consumption, and high-resolution imaging capabilities. Various navigation techniques utilizing monocular cameras have been proposed and implemented with varying degrees of success. Feature-based methods detect and track distinctive image features, such as corners or edges, to estimate relative motion between the camera and the target. The Scale-Invariant Feature Transform (SIFT) algorithm is widely used due to its robustness to changes in scale and orientation, though it can be computationally expensive. Faster alternatives like ORB (Oriented FAST and Rotated BRIEF) offer better performance but may be less accurate under certain conditions.[22] Optical Flow techniques estimate object motion in the camera's field of view by analyzing brightness pattern changes between consecutive frames. The Lucas-Kanade method is popular in real-time applications,[23] but optical flow methods can be sensitive to noise and struggle in environments with low texture or repetitive patterns.[24] Structure from Motion (SfM) techniques reconstruct the 3D structure of a scene from a series of 2D images taken from different viewpoints. These methods can provide accurate 3D models but are typically computationally intensive and unsuitable for real-time applications.[25] Pose Estimation methods determine the camera's position and orientation relative to the target object. One approach involves using known geometric models of the target and matching these to observed image features. Techniques like Perspective-n-Point (PnP) solve for the camera pose given a set of 3D points and their 2D projections.[26,27] These methods can be highly accurate but depend on a reliable 3D model of the target.

Monocular camera-based techniques, as mentioned, offer several advantages but also have inherent limitations. First and foremost, monocular cameras do not provide direct depth information, complicating accurate distance estimation to the target object. This often necessitates integrating additional sensors or sophisticated algorithms to infer depth, increasing complexity and computational demands.[28] Furthermore, monocular cameras can perform poorly under varying lighting conditions, such as changes in illumination or shadows,[29] as well as in cluttered environments or when occlusion occurs (e.g. when the manipulator enters the field of view of the sensor). Both conditions can lead to sub-optimal feature detection and tracking and, consequently, to navigation errors; this, in turn, calls for the necessity to devise complex guidance algorithms capable of avoiding such unfavorable situations.[15] Finally, many monocular camera-based techniques require intensive image processing and computational resources, which can be a limiting factor for real-time applications in space missions.[30,31] Ensuring these algorithms run efficiently on the limited hardware available on spacecraft is a critical challenge.[32]

Other shortcomings of the investigation carried out so far in the literature concerning the IOS missions and SMS systems in particular involve the modeling of the system. Specifically, the environmental disturbances are often underrepresented in the literature[13,19] and the system-level motion is often decoupled from the manipulator reconfiguration.[33] The latter in particular is a limiting assumption, especially for missions requiring the chaser to move around a bit or randomly tumbling target. In such cases, in fact, the forced motion of the chaser around the target would be required,

causing the information coming from the sensors to evolve faster than it is usually assumed, possibly exacerbating the limitation of the discussed vision systems and navigation techniques.

To address these challenges, a dual monocular camera system, mounted on the chasing spacecraft's base and the manipulator's end-effector, is proposed in this work. This configuration offers several advantages over traditional single-sensor systems. Monocular cameras are lightweight, low-power, and provide high-resolution images, making them suitable for space applications. Utilizing such two monocular cameras, effectively constituting a variable-baseline variable-orientation stereo system, enables a more robust and flexible navigation solution. While the base-mounted camera provides a wide field of view, allowing the system to maintain situational awareness and track the target object during the approach phase, the end-effector-mounted camera may offer a close-up view, facilitating precise manipulation once near the target. This dual-camera setup allows for continuous monitoring and control, reducing occlusion likelihood and enhancing the system's ability to adapt to dynamic environmental changes.[34] Furthermore, monocular cameras enable advanced image processing techniques, such as the aforementioned feature tracking, optical flow, and machine learning-based algorithms, to enhance navigation accuracy.[35] Leveraging the complementary strengths of the two cameras achieves higher redundancy and robustness, addressing existing navigation system shortcomings.

As a first step towards a more widespread investigation of this idea and the aforementioned techniques and in order to provide some baseline validation of the proposed approach, this work focuses specifically on designing a navigation filter that estimates the pose of the two camera frames in the scene directly from the x and y pixel positions of the target spacecraft features in the image planes, which are assumed to be extracted using relevant image processing techniques, outside the scope of this work. Unlike the common practice in the literature of filtering the pose extracted using Simultaneous Localization And Mapping (SLAM) algorithms, this approach, combined with an internal description of the system dynamics that inherently accounts for uncertainties in the measurements from the manipulator joint encoders, enables a more robust reconstruction of the end-effector's position. This is crucial for the mission's efficiency compared to the typically used end-effector pose reconstruction from the spacecraft base position and manipulator kinematics, as it relies on direct measurements for the estimation of the end-effector pose.

The remainder of this paper is organized as follows. The next section will introduce the different ways employed in this work for the modeling of the space manipulator system and the simulation architecture used for the validation of the approach. The adopted filtering technique will then be briefly discussed. The following section will then analyze the results in more detail, including the Monte-Carlo simulation carried out. Possible future developments will then be discussed in the conclusive section.

SPACE MANIPULATOR SYSTEM MODELING

In order to verify the performance of the filter, a simulation is orchestrated as depicted in Figure 1. The desired trajectories of the spacecraft base and end-effector are arbitrarily chosen as to represent the motion of the chasing spacecraft around the target. A simple open-loop guidance and control algorithm based on the flying-base space manipulator inverse dynamics is then used in order to infer the control actions to be exerted on the base of the spacecraft and on the joints of the manipulator for them to follow the prescribed trajectories. These controls are then fed as inputs to a Functional Mockup Unit (FMU) describing a high-fidelity model of the multibody system representing the Space Manipulator System (SMS). The latter provides, as an output, the evolution in

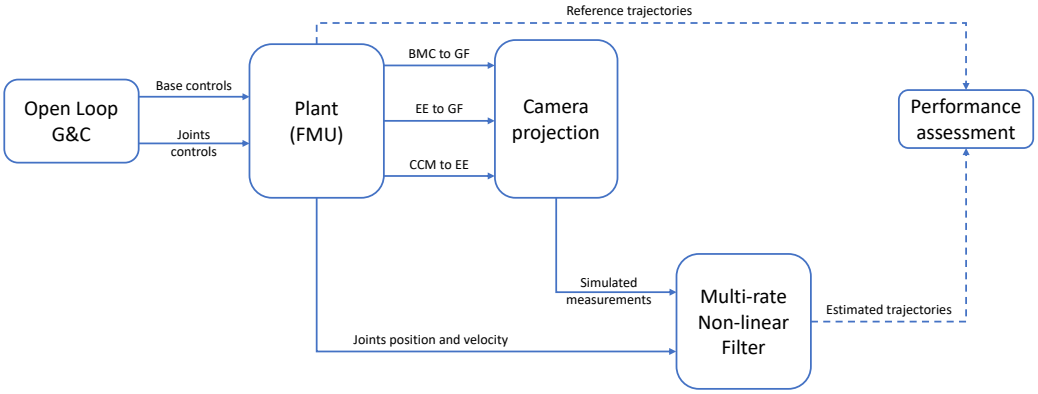


Figure 1. Overall simulation and testing architecture.

time of the pose of the body-mounted camera and end-effector one with respect to the body frame of the target spacecraft. On the one hand these constitute the true trajectories related to the provided control inputs, to be used as the reference truth for the evaluation of the performance of the filter. On the other hand, they are used, together with a wire-frame model of the target (which provides information on the recognizable features of the spacecraft), to project the position of each feature on the two image planes for the two cameras. This information, after being added with random noise, constitutes the measurements the sequential filter relies on for the correction phase. The filter also relies on the measurements of the manipulator joints' positions and velocities, which are assumed to be exact, as the precision in the measurements coming from the electric motors' encoders is usually orders of magnitudes higher than that of typical image feature extraction techniques such as corners and edges detection, blob detection or texture analysis.

As shown, a relatively simple setup like this still relies on two different representations of the dynamical system at hand, with varying degrees of accuracy. This is necessary, as we want a reference representation of the system which as faithful as possible to the real system, while we need a simpler dynamical description within the filter, as to avoid high computational costs and undesirable filter bias. The two models are described hereafter.

High fidelity model

As already mentioned, the high fidelity model is constituted by an FMU which is created using Modelica* (and the Dymola Modelica tool†). Modelica is an open-source object-oriented modeling and simulation language particularly suited for the modeling of physical systems as it uses an acausal programming paradigm. This means that the code does not follow a strict flow of inputs and outputs, like in other tools that serve a similar purpose, like for example Simulink, but rather allows the modeling engineer to write modules that describe the physical behavior of the system in a more natural way, focusing on the high-level mathematical description of the systems. The equal sign is not an assignement operator like in imperative programming languages, but it instead represents the equivalence operator, just like one is used to do when using pen and paper. At compile time, all of the equations constituting the overall model are then rearranged and inverted, if needed, in order to reduce the Differential Algebraic Equations (DAE) system to the minimum possible order.

*`https://modelica.org/` [last accessed on July, 17th, 2024]
†`https://www.3ds.com/products/catia/dymola` [last accessed on July, 17th, 2024]

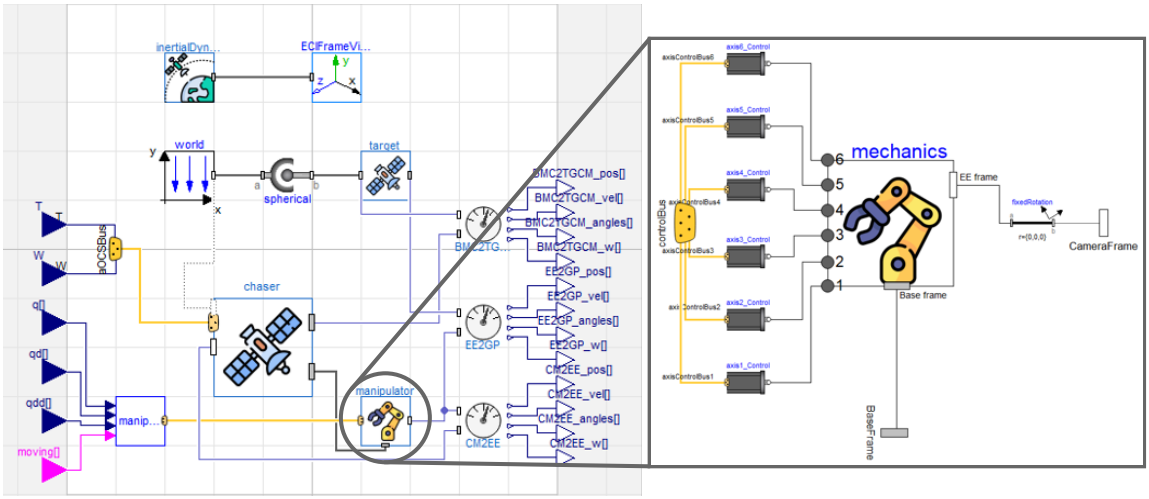
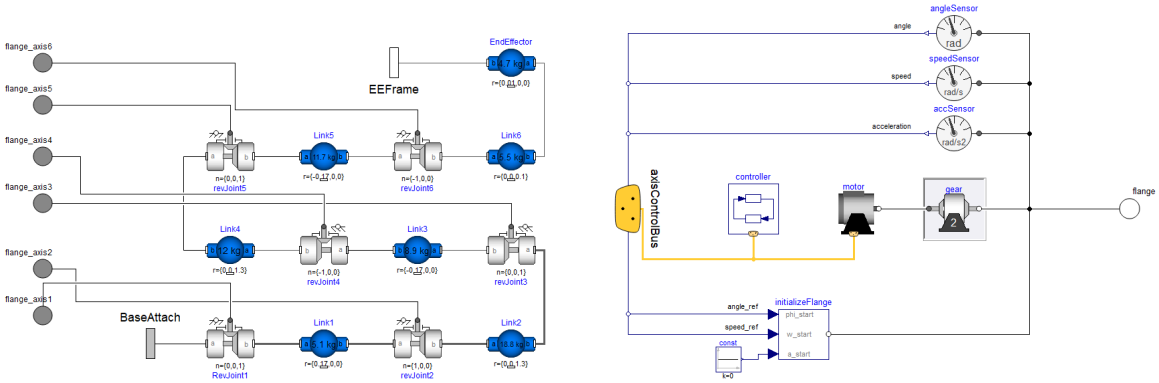


Figure 2. Full Modelica model of the chasing spacecraft system and dynamical environment (left) and detail of the manipulator model (right).

Modelica also supports the FMI standard, allowing the execution of FMU or the creation of new ones that can be executed from within other tools supporting FMI (here Simulink is used to execute the FMU).

The way Modelica is designed enables the user to create comparatively more complex models with a lower effort with respect to what would be required in other tools, partially due to the different programming paradigm, partially because the standard library is already pretty extensive and often renders writing code by hand unnecessary. Figure 2 shows, on the left, the full model for the simulation: it accounts for the dynamics of the target and chasing spacecrafts under gravitation and provides the dynamical response of the SMS depending on the provided control inputs. Thanks to the Modelica library, the manipulator shown in Figure 2 is modeled together with the controlling assembly of the joints; each of them, not only includes the control logic and the mechanical response of the gear but also the measuring sensor model and the actual circuitry of the electric motor too, as shown in figure Figure 3b.

The manipulator mechanical assembly shown in Figure 3a is composed of a series of joints and bodies representing the links. The Body objects from the standard Modelica library are built so that they experience the acceleration due to the gravity function defined in the ”world” model shown in Figure 2. However, this only allows the definition of acceleration models as a function of the position of the object expressed in a geocentric inertial frame. To address this limitation, a custom Body model, inheriting from the standard library one, was developed and used throughout the system, whose acceleration is derived from the equations specified in the ”inertialDynamics” block seen in the top left corner of Figure 2. This block implements the relative acceleration model by Franzini and Innocenti.[36] Such an approach offers greater flexibility for adding perturbation effects and ensures that the simulation is fully relative, thereby avoiding numerical issues that can arise due to the difference in magnitude between the absolute positions of the target and chaser and their relative position.



(a) Detail of the mechanical structure of the manipulator.

(b) Detail of the model of the motors controlling the manipulator joints.

Figure 3. Manipulator assembly modeling in Dymola.

Filter low fidelity model

As mentioned, a simpler model of the SMS has to be included in the filtering algorithm. Using, for example, the Lagrangian approach, the dynamics of an SMS can be written as:[37,38]

$$\begin{bmatrix} M_0 & M_{0m} \\ M_{m0} & M_m \end{bmatrix} \begin{bmatrix} \ddot{\theta}_0 \\ \ddot{\theta}_m \end{bmatrix} + \begin{bmatrix} C_b & C_{0m} \\ C_{m0} & C_m \end{bmatrix} \begin{bmatrix} \dot{\theta}_0 \\ \dot{\theta}_m \end{bmatrix} + \begin{bmatrix} N_0 \\ N_m \end{bmatrix} = \begin{bmatrix} \tau_0 \\ \tau_m \end{bmatrix} \tag{1}$$

often shortened as:

$$M\ddot{\theta} + C\dot{\theta} + N = \tau \tag{2}$$

for convenience. Here:

- θ is the vector of generalized coordinates $\{\theta_0; \theta_m\}$, associated to the base and manipulator respectively
- M is the Generalized Inertia Matrix (GIM)
- C is the Convective Inertia Matrix (CIM) or Coriolis Matrix
- N accounts for the gravitational effects (and other potentials)

Other terms can be added to include the effects of external disturbances, friction, etc.

Differently from typical on-ground robotic applications, an SMS base is not constrained in any way and this causes the off-diagonal terms of the GIM and the CIM to be non-zero, as a consequence of the dynamics coupling effects between the base and manipulator motion. When actuating the manipulator disturbing forces are transmitted to the base (if not actively controlled) due to the angular momentum conservation, altering its motion.
Equation (1), while offering interesting insights on the dynamical response of the system, is not well suited for simulation purposes, as its complexity is of the order $\mathcal{O}(n^3)$, where n is the number of links constituting the manipulator. Resorting to a Newton approach for the representation of the manipulator, fortunately, recursive techniques of order $\mathcal{O}(n)$ can be devised, such as the ones based on Screw Theory and the Articulated Body Model (ABM).[37]

In this work, in particular, the SPART software has been used for the modeling of the SMS in the dynamics of the filter.[39] In this way, the forward dynamics of the system can be easily retrieved. In particular, the accelerations of the generalized coordinates related to the base and end-effector, $\ddot{\theta}_0$ and $\ddot{\theta}_m$, can be expressed as a function of the control actions applied, τ_0 and τ_m, the forces (wrenches) acting on the system, W_{F_0} and W_{F_m}, and the current configuration and velocities, θ_0, θ_m, $\dot{\theta}_0$ and $\dot{\theta}_m$.
The state-space representation of the system can therefore be formulated as follows:

$$\begin{cases} \dot{q}_0 = \frac{1}{2}[w_0; 0] \otimes q_0 \\ \dot{r}_0 = v_0 \\ [\dot{w}_0, \dot{v}_0] = \ddot{\theta}_0 \\ \dot{q}_{EE} = \frac{1}{2}[w_{EE}; 0] \otimes q_{EE} \\ \dot{r}_{EE} = v_{EE} \\ [\dot{w}_{EE}, \dot{v}_{EE}] = J_{0_{EE}}\ddot{\theta}_0 + \dot{J}_{0_{EE}}\dot{\theta}_0 + J_{m_{EE}}\ddot{\theta}_m + \dot{J}_{m_{EE}}\dot{\theta}_m \end{cases} \tag{3}$$

where q_0 and q_{EE} are the quaternions representing the orientation of the base of the chasing spacecraft and the end-effector of the manipulator, respectively. Similarly, r_0, r_{EE}, v_0 and v_{EE} represent the position and velocity of the base and end-effector. $\theta_0 = \{q_0; r_0\}$ collects both the rotational and translational states of the base of the manipulator.

Differently from what has already largely been done in the literature, the state definition is here based on the end-effector quantities rather than on the joints ones. This causes the flexibility of the joints, and therefore the uncertainty on their displacement, to be implicitly accounted for in the filter and hence increase its robustness and improve its performance.

FILTERING TECHNIQUE

Many different non-linear filtering techniques and algorithms have been devised, throughout the decades, and applied to the pose estimation problem. Among them, the family of the Kalman filters is largely used for estimation purposes and the Extended Kalman Filter (EKF), in particular, is very common when dealing with non-linear systems. However, it suffers a significant degradation in performance when dealing with highly non-linear systems as it linearizes the dynamics in the neighborhood of the current system state. A good idea, in such cases, is resorting to the Unscented Transform during the propagation step, because it allows to retain the non-linearity and capture the statistics of a transformed distribution with higher accuracy, while not increasing the computational cost thanks to a careful selection of the samples (sigma points) to be propagated. This is the approach used by the Unscented Kalman Filter (UKF).[40]

For the problem at hand, however, this may still not be enough, as we are dealing with quaternions to represent the attitude of the base and end-effector. Although quaternions are often treated as 4D Euclidean vectors, they are indeed elements of the $SU(2)$ Special Unitary group of order 2 (or, equivalently of the $Q(8)$ Quaternion group of 8 elements) which is not closed with respect to the Euclidean addition operation, such as the one that has to be performed in the correction step of the UKF. In other words, summing two quaternions may not result in a quaternion as the constraint on unitary norm might be violated. This causes problems both in the recombination of the sigma

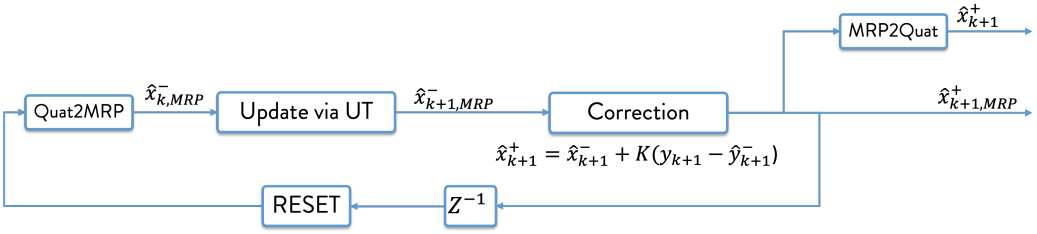


Figure 4. Schematic representation of an UnScented QUaternion Estimator. Here, x represents the states, y the measurements, k the time instant, and K the Kalman gain. The minus and plus signs refer to quantities before and after the correction step, respectively.

points (which can somewhat be easily solved by introducing a normalization step), but also in the computation of the correction factor. This problem is common to all the filters based on Kalman theory, which share the same correction equation, and historically led to the development of various alternatives.

Among these, there's, for example, the Multiplicative EKF (MEKF), which retains however the shortcomings of the traditional EKF, and the UnScented QUaternion Estimator (USQUE), which could basically be interpreted as a multiplicative UKF.[41] By resorting to a multiplicative definition of the error difference between two quaternions, it allows for a correct representation of the orientation errors. In practice, the internal state of the filter contains the error of the orientation estimate, rather than the orientation itself, and it temporarily moves to a representation of such errors via Generalized Rodriguez Parameters (GRPs),[42] which allows for a safer reconstruction of the state after the UT and the automatic compliance with the unitary norm constraint of the quaternions. A simple schematic representation of the filter architecture is shown in Figure 4.

Measurement function

As already briefly discussed in the introduction, another peculiar aspect of the approach proposed in this paper concerns the measurement function used in the filter. The typical approach, largely used in the literature, to the navigation problem using optical instruments, consists of three steps:[30,43,44]

1. First, the images provided by the sensors are fed to a feature extraction algorithm that identifies in them the markers or generic features of interest, possibly matching them to a prescribed ID.

2. Then, PnP algorithms and alike are employed to infer the pose of the camera in the scene based on the distribution of the features in the image, either relying on some known model of the target or on one reconstructed online using SLAM techniques, as in Reference 15.

3. Finally, the poses thus obtained are used as-is, possibly fused with the information coming from other sensors, to correct the state of the filter coming from the update step.

The work here presented uses a different approach: it avoids the use of the PnP algorithms by directly embedding the raw measurements from the feature recognition step inside the filter. Specifically, the measurement function of the filter includes a simple model of the monocular camera, through which, based on the predicted state of the filter, the expected positions (i.e. x and y coordinates) of the target features in the two cameras are retrieved. The correction step then compares these to the actual measurements coming from the sensors to correct the state.

The camera model, based on the simple pinhole projection, includes visibility checks to only account, at each time step, for the features that are illuminated and not occluded. Nonetheless, the dimensionality of the problem tends to grow quite rapidly as 4 measurements are associated with each visible feature (the x and y coordinates in each of the two images). The technique may then need further refinement to reduce the computational workload and improve reproducibility in practical applications. It may be beneficial to select a fixed-dimension subset of visible features at each time step, rather than always using all of the available ones. However, determining the criteria for feature selection and considering factors such as aberration and coma's impact on filtering performance remains, at the moment, an open research question.

RESULTS

Methodology

The target model is assumed to be a simple rectangular prism whose corners are identified by fiducial markers. Visibility checks are introduced during the projection in the image plane, based on occlusion and illumination. The lighting conditions (i.e. the relative orientation, in time, of the lighting direction with respect to the camera-target line of sight) are chosen so that they are representative of what could reasonably be expected in an operational scenario, as to avoid particularly favorable situations and the consequent introduction of an unwanted bias in the results. As a consequence, the number of visible markers, whose detail is reported in Figure 5, varies in time and is particularly unfavorable in the beginning stages of the pose reconstruction, when the initialization error of the filter has the most influence.

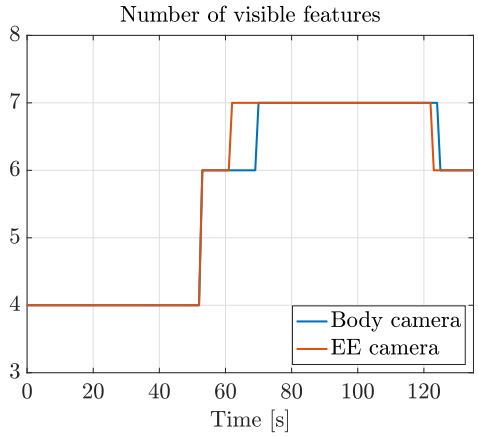


Figure 5. Number of visible features from the two cameras as a function of time.

The simulated trajectory of the chaser around the target is chosen arbitrarily: the base is assumed to be moving with non-uniform velocity in a straight line at distances from the target that range from 12 to 20 meters. Some keypoint poses are defined for the end-effector camera with respect to the CoM of the spacecraft base throughout the simulation and interpolated with a minimum-jerk trajectory to have the evolution in time of the pose with respect to the target. Without any loss of generality for the study at hand, given the approach just described for the definition of the two cameras' trajectories, the target is assumed to be axis stabilized and therefore spinning at a constant rate (in the range of 0.5 to 1 deg/s) around the Z-axis of the LVLH frame (which is not coincident with any of the principal axis of the target spacecraft).

The measurements are assumed to be only available at intervals larger than one second, which is a reasonable assumption when considering the necessary time for the processing of the images and extraction of the features positions.

Results

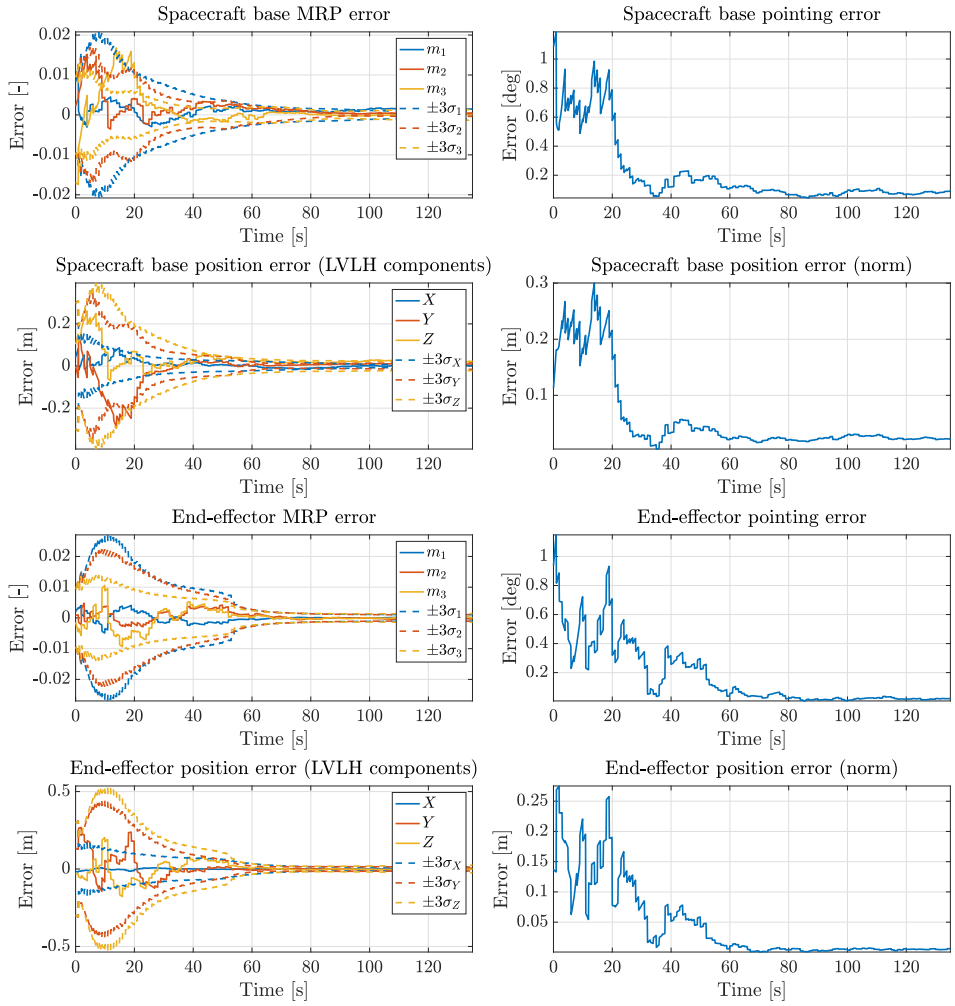


Figure 6. Error components (left column) and norm (right column) for the spacecraft and end-effector frames attitude and position. Measurements are assumed to be available once per second and with a 1-pixel standard deviation from reality.

The results for the simulation, shown in Figure 6, demonstrate the efficacy of the method, showcasing high accuracy in the reconstruction of the position and orientation of both the reference frames of interest. As can be seen in the picture, the steady-state error on the norm of the position vectors of the two frames of interest with respect to the LVLH one is in the order of millimeters and below the 3-σ threshold derived from the state covariance matrix. The pointing error is also particularly low, indicating an efficient reconstruction of the relative orientation of the base and

end-effector frames with respect to the target, despite the large initialization errors. This highlights how the approach effectively mitigates the lack of a direct range measurement and how the two-camera system, coupled with a suited navigation algorithm, effectively constitutes a variable-baseline variable-orientation stereo camera.

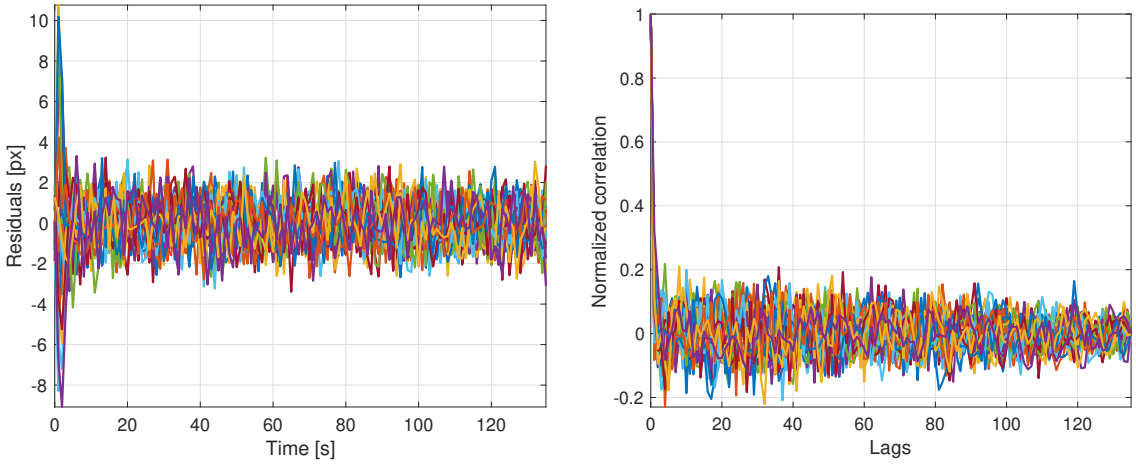


Figure 7. Innovation (left) and normalized cross-correlation of the innovation (right). Each line refers to one of the measurement components.

Tests have also been conducted to testify the consistency of the filter: Figure 7 shows how the magnitude of the innovation of the filter throughout the time-span of the simulation is randomly distributed around zero and much smaller in magnitude with respect to the typical dimension of the measurements (the x-y position in pixel of the detected features in the acquired images, assumed to be 1920-by-1200 pixels large); the cross-correlation of the innovation is also zero-mean and small in magnitude, as expected.

Results also show how a relatively low number of detected features is enough to keep the filter from diverging, strengthening the idea that the introduction of a feature screening/selection step, as described in the previous section, could prove beneficial in operative situations without undermining the robustness of the filter.

A comparative test has also been carried out with respect to the use of a simple UKF. The improvement in the reconstruction of the rotational states is only marginal. This is probably to be identified as the effect of the measurements choice. Using the image coordinates of the features in the pictures as measurements for the filter, causes the filter to always rely on a relevant number of features, effectively rendering the choice of the actual filtering scheme less relevant.

Monte-Carlo simulation

To further validate the robustness of the approach, a Monte-Carlo simulation has been run, testing the performance of the filter with varying degrees of initialization errors. In particular, the initial state of the filter is randomly chosen, at each run, from a normal distribution centered around the actual state of the system and variances linked to each of the elements of the state. Such variances have been chosen based on the typical performances of sensors and algorithms used for the guidance, navigation, and control of spacecrafts during mid-range approaches to targets. The values are:

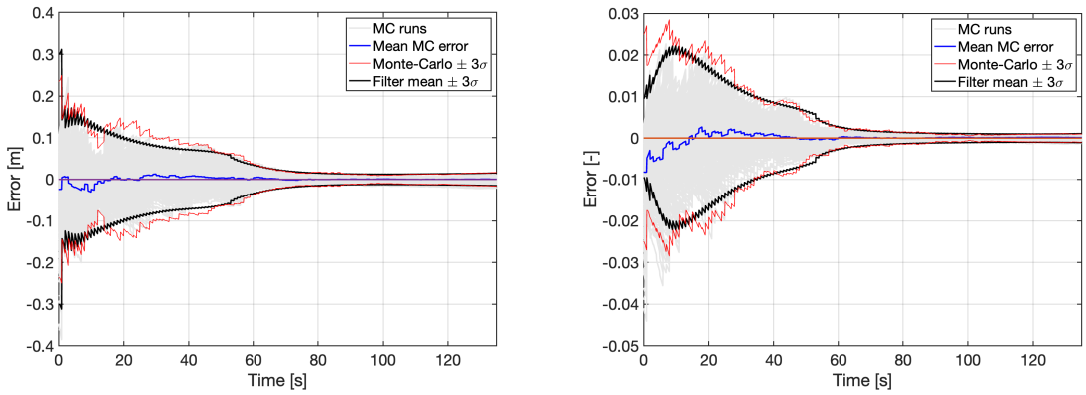


Figure 8. Results of the Monte-Carlo simulation for the end-effector position (left) and orientation (right). For readability reasons, only one representative component is shown for each quantity.

- 1 cm for the positional states;
- 1 mm/s for the linear velocities;
- 10^{-5} for the components of the GRP representing the rotational state of the two cameras, corresponding to a maximum initial pointing error of around 2 degrees;
- 10^{-5} rad/s^2 for the angular velocity of the two camera frames.

The results for a 900-samples simulation are shown in Figure 8. Here the red line represents the 3σ threshold associated with the error distribution across the different test runs at each time step. As it can be seen, it closely resembles the one coming from the covariance history output from the filter (black line, computed as the mean of the filter covariance across the samples at each time step). Moreover, it can be noticed how the error rarely exceeds the 3σ threshold, no matter how large the initialization error is. This proves the feasibility, robustness, and effectiveness of the proposed approach: the filter does, in fact, converge in each simulation despite the assumed adverse conditions, and the error is reduced by an order of magnitude at the end of the simulation.

CONCLUSIONS

The presented paper demonstrated that the integration of a dual monocular camera system for navigation during the final approach phase in IOS missions could represent a significant advancement over traditional methods. By combining the strengths of base-mounted and end-effector-mounted cameras, this approach offers enhanced reliability, accuracy, and flexibility. The error in the reconstruction of the pose of the base and end-effector has, in fact, been shown to be in the order of millimeters and tenths of a degree when the chaser distance from the target is still in the 10-20 meters range.

The present work relied on simulated data to retrieve the measurements to be fed to the navigation algorithm. Future work will include a more refined simulation, in which the measurements are extracted from images using feature extraction algorithms. Given the results shown in this paper, the feature extraction and matching one is, in fact, foreseen as the most critical step in the process and needs further investigation. The focus will then shift to the development and testing of advanced

image processing algorithms to fully exploit the potential of this system, paving the way for more effective and reliable in-orbit servicing missions.

As it has been demonstrated that a low number of detected features is enough for the algorithm to provide good results, future works will also investigate criteria to be used to select which measurements to discard in the filter, to keep the dimensionality of the problem within reasonable limits while ensuring a marginal loss of accuracy, granting computational loads compatible with the deployment on flight hardware.

REFERENCES

[1] K. Wormnes, R. Le Letty, L. Summerer, R. Schonenborg, O. Dubois-Matra, E. Luraschi, A. Cropp, H. Krag, and J. Delaval, "ESA technologies for space debris remediation," *6th European Conference on Space Debris*, Vol. 1, ESA Communications ESTEC Noordwijk, The Netherlands, 2013, pp. 1–8.

[2] H. Klinkrad, P. Beltrami, S. Hauptmann, C. Martin, H. Sdunnus, H. Stokes, R. Walker, and J. Wilkinson, "The ESA Space Debris Mitigation Handbook 2002," *Advances in Space Research*, Vol. 34, No. 5, 2004, pp. 1251 – 1259.

[3] K. Yoshida, B. Wilcox, G. Hirzinger, and R. Lampariello, *Space Robotics*, pp. 1423–1462. Springer International Publishing, 2016, 10.1007/978-3-319-32552-1_55.

[4] K. Yoshida, "Engineering test satellite VII flight experiments for space robot dynamics and control: theories on laboratory test beds ten years ago, now in orbit," *The International Journal of Robotics Research*, Vol. 22, No. 5, 2003, pp. 321–335.

[5] G. Hirzinger, B. Brunner, J. Dietrich, and J. Heindl, "ROTEX-the first remotely controlled robot in space," *Proceedings of the 1994 IEEE international conference on robotics and automation*, IEEE, 1994, pp. 2604–2611.

[6] A. Flores-Abad, O. Ma, K. Pham, and S. Ulrich, "A review of space robotics technologies for on-orbit servicing," *Progress in aerospace sciences*, Vol. 68, 2014, pp. 1–26.

[7] S. Chien, J. Doubleday, D. R. Thompson, K. L. Wagstaff, J. Bellardo, C. Francis, E. Baumgarten, A. Williams, E. Yee, E. Stanton, *et al.*, "Onboard autonomy on the intelligent payload experiment cubesat mission," *Journal of Aerospace Information Systems*, Vol. 14, No. 6, 2017, pp. 307–315.

[8] M. Maestrini, P. Di Lizia, and F. Topputo, "Analytical Impulsive-to-Continuous Thrust Conversion in Linearized Relative Dynamics," *Journal of Guidance, Control, and Dynamics*, Vol. 44, No. 4, 2021, pp. 862–871, 10.2514/1.G005520.

[9] C. Zhao, M. Maestrini, and P. Di Lizia, "Low-Thrust Optimal Control of Spacecraft Hovering for Proximity Operations," *Journal of Guidance, Control, and Dynamics*, Vol. 47, No. 7, 2024, pp. 1457–1469, 10.2514/1.G008063.

[10] J. A. Starek, B. Açıkmeşe, I. A. Nesnas, and M. Pavone, "Spacecraft autonomy challenges for next-generation space missions," *Advances in Control System Technology for Aerospace Applications*, pp. 1–48, Springer, 2015.

[11] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on robotics*, Vol. 32, No. 6, 2016, pp. 1309–1332.

[12] F. Aghili and C.-Y. Su, "Robust Relative Navigation by Integration of ICP and Adaptive Kalman Filter Using Laser Scanner and IMU," *IEEE/ASME Transactions on Mechatronics*, Vol. 21, No. 4, 2016, pp. 2015–2026, 10.1109/TMECH.2016.2547905.

[13] E. Papadopoulos, F. Aghili, O. Ma, and R. Lampariello, "Robotic Manipulation and Capture in Space: A Survey," *Frontiers in Robotics and AI*, Vol. 8, 2021, 10.3389/frobt.2021.686723.

[14] A. Petit, E. Marchand, and K. Kanani, "Vision-based space autonomous rendezvous: A case study," *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2011, pp. 619–624.

[15] M. Maestrini, M. A. De Luca, and P. Di Lizia, "Relative navigation strategy about unknown and uncooperative targets," *Journal of Guidance, Control, and Dynamics*, Vol. 46, No. 9, 2023, pp. 1708–1725.

[16] W. Xu, B. Liang, C. Li, and Y. Xu, "Autonomous rendezvous and robotic capturing of non-cooperative target in space," *Robotica*, Vol. 28, No. 5, 2010, pp. 705–718.

[17] M. Maestrini and P. Di Lizia, "COMBINA: Relative Navigation for Unknown Uncooperative Resident Space Object," *AIAA Scitech Forum, 3-7 January 2022, San Diego, CA*, AIAA, 2022, pp. 1–19. Paper No. AIAA 2022-2384, 10.2514/6.2022-2384.

[18] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations," *Progress in Aerospace Sciences*, Vol. 93, 2017, pp. 53–72, https://doi.org/10.1016/j.paerosci.2017.07.001.

[19] B. M. Moghaddam and R. Chhabra, "On the guidance, navigation and control of in-orbit space robotic missions: A survey and prospective vision," *Acta Astronautica*, Vol. 184, 2021, pp. 70–100.

[20] B. Tweddle, A. Saenz-Otero, and D. Miller, "Design and development of a visual navigation testbed for spacecraft proximity operations," *AIAA SPACE 2009 Conference & Exposition*, 2009, p. 6547.

[21] Z. Pavanello, F. Branz, A. Francesconi, A. Cenedese, R. Antonello, F. Basana, P. Iob, D. Vertuani, M. Massari, C. Colombo, M. Lovera, D. Invernizzi, P. Ghignoni, L. Ticozzi, G. Borelli, R. Opromolla, M. Grassi, G. Fasano, A. Nocerino, and P. Simplicio, "Combined control and navigation approach to the robotic capture of space vehicles," *72nd International Astronautical Congress (IAC 2021)*, 10 2021.

[22] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *2011 International conference on computer vision*, Ieee, 2011, pp. 2564–2571.

[23] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *IJCAI'81: 7th international joint conference on Artificial intelligence*, Vol. 2, 1981, pp. 674–679.

[24] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, 1994.

[25] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.

[26] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EP n P: An accurate O (n) solution to the P n P problem," *International journal of computer vision*, Vol. 81, 2009, pp. 155–166.

[27] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," *2011 International Conference on Computer Vision*, IEEE, 2011, pp. 383–390.

[28] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *IEEE robotics & automation magazine*, Vol. 18, No. 4, 2011, pp. 80–92.

[29] X. Mei, X. Sun, W. Dong, H. Wang, and X. Zhang, "Segment-tree based cost aggregation for stereo matching," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 313–320.

[30] D. Kaidanovic, M. Piazza, M. Maestrini, and P. Di Lizia, "Deep learning-based relative navigation about uncooperative space objects," *Proceedings of the 73rd International Astronautical Congress, 18-22 September, 2022, Paris, France.*, Vol. 2022-September, IAF, 2022, pp. 1–11. Paper No. IAC-22.C1.IPB.38.x70501.

[31] B. Razgus, M. Maestrini, and P. D. Lizia, "Initial Attitude Acquisition for Uncooperative Resident Space Objects Using Principal Lines," *AIAA SCITECH Forum, 8-12 January 2024, Orlando, Florida*, AIAA, 2024, pp. 1–8. Paper No. AIAA 2024-0201, 10.2514/6.2024-0201.

[32] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 2938–2946.

[33] J. Virgili-Llop, C. Zagaris, R. Zappulla, A. Bradstreet, and M. Romano, "A convex-programming-based guidance algorithm to capture a tumbling object on orbit using a spacecraft equipped with a robotic manipulator," *The International journal of robotics research*, Vol. 38, No. 1, 2019, pp. 40–72.

[34] B. Sumetheeprasit, R. Rosales Martinez, H. Paul, R. Ladig, and K. Shimonomura, "Variable Baseline and Flexible Configuration Stereo Vision Using Two Aerial Robots," *Sensors*, Vol. 23, No. 3, 2023, 10.3390/s23031134.

[35] N. Faraco, M. Maestrini, and P. Di Lizia, "Instance Segmentation for Feature Recognition on Noncooperative Resident Space Objects," *Journal of Spacecraft and Rockets*, Vol. 59, No. 6, 2022, pp. 2160–2174, 10.2514/1.A35260.

[36] G. Franzini and M. Innocenti, "Relative Motion Dynamics with Arbitrary Perturbations in the Local-Vertical Local-Horizon Reference Frame," *The Journal of the Astronautical Sciences*, Vol. 67, No. 1, 2020, pp. 98–112, 10.1007/s40295-019-00185-0.

[37] K. M. Lynch and F. C. Park, *Modern Robotics: Mechanics, Planning, and Control*. Cambridge, UK: Cambridge University Press, 2017.

[38] R. Featherstone, *Rigid body dynamics algorithms*. Springer, 2014.

[39] J. Virgili-Llop *et al.*, "SPART: an open-source modeling and control toolkit for mobile-base robotic multibody systems with kinematic tree topologies," `https://github.com/NPS-SRL/SPART`.

[40] E. Wan and R. Van Der Merwe, "The Unscented Kalman Filter for Nonlinear Estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, Lake Louise, Alta., Canada, IEEE, 2000, pp. 153–158, 10.1109/ASSPCC.2000.882463.

[41] J. L. Crassidis and F. L. Markley, "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, July 2003, pp. 536–542, 10.2514/2.5102.

[42] H. Schaub, J. L. Junkins, *et al.*, "Stereographic orientation parameters for attitude dynamics: A generalization of the Rodrigues parameters," *Journal of the Astronautical Sciences*, Vol. 44, No. 1, 1996, pp. 1–19.

[43] M. Piazza, M. Maestrini, and P. Di Lizia, "Monocular Relative Pose Estimation Pipeline for Uncooperative Resident Space Objects," *Journal of Aerospace Information Systems*, Vol. 19, No. 9, 2022, pp. 613–632, 10.2514/1.I011064.

[44] T. H. Park and S. D'Amico, "Adaptive Neural-Network-Based Unscented Kalman Filter for Robust Pose Tracking of Noncooperative Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 46, No. 9, 2023, pp. 1671–1688, 10.2514/1.G007387.