Progress in Aerospace Sciences 110 (2019) 100548

Contents lists available at ScienceDirect

Progress in Aerospace Sciences

journal homepage: www.elsevier.com/locate/paerosci

Review of the robustness and applicability of monocular pose estimation systems for relative navigation with an uncooperative spacecraft

Lorenzo Pasqualetto Cassinis*, Robert Fonod, Eberhard Gill

Delft University of Technology, Kluyverweg 1, 2629 HS, Delft, the Netherlands

ARTICLE INFO

Keywords:
Relative pose estimation
Active debris removal
In-orbit servicing
Monocular cameras
Image processing
Visual-based navigation filters

ABSTRACT

The relative pose estimation of an inactive target by an active servicer spacecraft is a critical task in the design of current and planned space missions, due to its relevance for close-proximity operations, i.e. the rendezvous with a space debris and/or in-orbit servicing. Pose estimation systems based solely on a monocular camera are recently becoming an attractive alternative to systems based on active sensors or stereo cameras, due to their reduced mass, power consumption and system complexity. In this framework, a review of the robustness and applicability of monocular systems for the pose estimation of an uncooperative spacecraft is provided. Special focus is put on the advantages of multispectral monocular systems as well as on the improved robustness of novel image processing schemes and pose estimation solvers. The limitations and drawbacks of the validation of current pose estimation schemes with synthetic images are further discussed, together with the critical trade-offs for the selection of visual-based navigation filters. The state-of-the-art techniques are analyzed in order to provide an insight into the limitations involved under adverse illumination and orbit scenarios, high image contrast, background noise, and low signal-to-noise ratio, which characterize actual space imagery, and which could jeopardize the image processing algorithms and affect the pose estimation accuracy as well as the navigation filter's robustness. Specifically, a comparative assessment of current solutions is given at different levels of the pose estimation process, in order to bring a novel and broad perspective as compared to previous works.

1. Introduction

In the past years, advancements in the field of Distributed Space Systems (DSS) have been made to cope with the increasing demand for robust and reliable engineering solutions in challenging scenarios for Guidance, Navigation, and Control (GNC), such as Formation Flying (FF) missions, In-Orbit Servicing (IOS), and Active Debris Removal (ADR).

Previous research in the context of FF has led to robust and reliable real-time estimation of the position and velocity of a target object with respect to the main spacecraft. Navigation architectures which combine absolute and relative measurements have been designed and implemented in past and current missions that rely either on Radio Frequency (RF), Global Positioning System (GPS) sensors or on cameras. As an example, the PRISMA mission provided the first in-orbit demonstration of non-GPS RF-based metrology instruments for relative navigation [1], and recent improvements have been made to use a Visual-Based System (VBS) as the main navigation system in more recent missions [2]. Moreover, additional effort has been made in the recent years on IOS and assembly and Debris Removal [3,4]. For these close-proximity scenarios, the relative position and orientation, herewith referred to as pose, of a target spacecraft with respect to a servicer spacecraft, represent a key information for the navigation system. A proper characterization of the target spacecraft is essential to determine its status and to plan the final strategy of the approaching orbit during autonomous close-proximity operations. Notably, the pose estimation problem is in this case complicated by the fact that the target satellite is, especially in the context of ADR, uncooperative, namely retained as non functional and/or not able to aid the relative navigation. In particular, the additional flexibility required to deal with a non-functional and/or freely tumbling target has an impact on the navigation system. Compared to FF missions or more commonly to cooperative close-proximity missions, vision-based sensors should be preferred over RF sensors when the satellite is uncooperative. Additionally, the navigation system cannot rely on known visual markers installed on the target spacecraft, and requires advanced Image Processing (IP) and pose estimation algorithms in order to cope with the lack of knowledge of the initial relative position and attitude. Moreover, if the target is tumbling at a relatively high rate, additional challenges arise in the tracking of the relative pose due to the fast relative dynamics.

* Corresponding author.
E-mail addresses: L.PasqualettoCassinis@tudelft.nl (L. Pasqualetto Cassinis), R.Fonod@tudelft.nl (R. Fonod), E.K.A.Gill@tudelft.nl (E. Gill).

https://doi.org/10.1016/j.paerosci.2019.05.008
Received 18 February 2019; Received in revised form 1 May 2019; Accepted 25 May 2019
Available online 14 June 2019
0376-0421/

List of abbreviations

ADR	Active Debris Removal
BRIEF	Binary Robust Independent Elementary Features
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
DA	Differential Algebra
DQ-MEKF	Dual Quaternion Modified Extended Kalman Filter
DSS	Distributed Space Systems
EDL	Edge Drawing Lines
EO	Electro-Optical
ESA	European Space Agency
EKF	Extended Kalman Filter
FF	Formation Flying
FREAK	Fast Retina Keypoint
GEO	Geostationary Earth Orbit
GFTT	Good Feature to Track
GNC	Guidance, Navigation and Control
GNFIR	Goddard Natural Feature Image Recognition
GPS	Global Positioning System
HCD	Harris Corner Detection
HEO	High Elliptical Orbit
HIL	Hardware-In-the-Loop
HST	Hubble Space Telescope
HT	Hough Transform
ICP	Iterative Closest Point
IMU	Inertial Measurement Unit
IOS	In-Orbit Servicing
IoU	Intersection-Over-Union
IP	Image Processing
IRLS	Iteratively Re-Weighted Least Squares
KF	Kalman Filter
LEO	Low Earth Orbit
LIDAR	LIght Detection And Ranging
MEKF	Multiplicative Extended Kalman Filter
MRP	Modified Rodrigues Parameters
MSRN	Multi-Spectral Sensing for Relative Navigation
MWIR	Mid-Wave Infra-Red
LPF	Low Pass Filter
LSD	Line Segment Detector
LWIR	Long-Wave Infra-Red
NIR	Near Infra-Red
NRM	Newton Raphson Method
RF	Radio Frequency
PCA	Principal Component Analysis
PC-P	Phase Congruency Point
PnP	Perspective-n-Point
RANSAC	RANdom SAmple Consensus
ROE	Relative Orbital Elements
ROI	Region Of Interest
RCM	Roberts Cross Method
RPN	Region Proposal Network
SIL	Sotware-In-the-Loop
SIFT	Scale-Invariant Transform
ST	Shi-Tomasi
SURF	Speeded Up Robust Features
SNR	Signal-To-Noise Ratio
TIR	Thermal Infra-Red
TOF	Time-Of-Flight
UKF	Unscented Kalman Filter
VBS	Visual-based System
VNIR	Visual-Near Infra-Red
WGE	Weak Gradient Elimination

From a high-level perspective, visual-based sensors can be divided into active and passive devices, depending on whether they require power to function, i.e. LIght Detection And Ranging (LIDAR) sensors and Time-Of-Flight (TOF) cameras, or if they passively acquire radiation, i.e. monocular and stereo cameras. Spacecraft relative navigation usually exploit Electro-Optical (EO) sensors such as stereo cameras [5,6] and/or a LIDAR sensor [7] in combination with one or more monocular cameras, in order to overcome the partial observability that results from the lack of range information in these latter [8]. However, systems based solely on monocular cameras are currently being investigated given the fact that monocular navigation ensures rapid pose determination under low power and mass requirements [9], which is an asset given the constraints in the processing power available for in-flight pose estimation, while on the other hand, stereo cameras and LIDAR sensors are less flexible and less convenient in terms of operational range, mass, power consumption and processing power [10]. The range unobservability problem of monocular cameras can indeed be tackled if a wireframe 3D model of the target is included in the pose estimation, by matching it with features extracted from the 2D monocular image and solving for the full relative pose, or alternatively if an offline database of images of the target is available together with their associated pose label. However, given the low Signal-To-Noise Ratio (SNR) and the high contrast which characterize space images, a significant effort is still required to comply with most of the demanding requirements for a robust and accurate monocular-based navigation system.

In the presented framework, the aim of this paper is to provide a detailed overview of the robustness and applicability of state-of-the-art monocular-based pose estimation systems for the relative navigation with an uncooperative target. Recent surveys on the topic focused on a comparative assessment of the pose estimation solvers [11] or provided a broader review on cooperative as well as uncooperative targets by including monocular-as well as stereo- and LIDAR-based systems [10]. Furthermore, only monocular cameras operating in the visible spectrum where reviewed, and recent estimation methods based on deep learning techniques were not included. The novelty of this work stands in extending the previous surveys in mainly three directions. Firstly, focus is put on the applicability and robustness of multispectral monocular cameras. Secondly, both IP systems and pose estimation algorithms are analyzed with particular emphasis on the relative range they were tested on, the robustness with respect to the image background, and on the synthetic and real images database adopted for their validation. Furthermore, novel pose estimation schemes are reviewed which are based on Convolutional Neural Networks (CNN). Finally, a review is presented for the navigation filters currently adopted. A distinction is made between known targets, for which mass and inertia properties as well as a 3D model of the target are known and available, and partially known targets, for which the uncertainty is constrained to the target center of mass and moment of inertia, while a 3D model of the target is available. Notably, this distinction impacts on the internal dynamics of the navigation filter rather than on the image processing and pose estimation prior to the filter. The reader is referred to Opromolla et al. [10] for an overview of the pose estimation of uncooperative unknown targets, for which neither the target mass and inertia properties nor a 3D model of the target are available prior to the on-line estimation.

The paper is organized as follows. Section 2 presents a review of the robustness and applicability of monocular cameras operating in the visible (VIS), Near Infrared (NIR) and Mid/Long Wave Infrared (MWIR/LWIR), the latter type of cameras being also referred to as Thermal Infrared (TIR) cameras. Section 3 contains a detailed review of IP algorithms as well as pose estimation algorithms which have been developed for uncooperative targets. Section 4 provides a review of

Table 1
Advantages and disadvantages of TIR/NIR/VIS cameras for space applications, based on the reviewed papers. Here, the characteristics of VIS cameras are referred to as 'Nominal' for clarity of the comparison.

	Saturation due to the Sun	Robustness w.r.t. Eclipse	Robustness w.r.t. Earth in background	Image quality	Robustness w.r.t thermal dynamics
VIS	Nominal	Nominal	Nominal	Nominal	Nominal
TIR	Superior	Superior	Superior	Inferior	Inferior
NIR	Nominal	Superior	Nominal	Nominal	Inferior

visual-based navigation systems with focus on the navigation filters currently adopted. Finally, Section 5 lists the main conclusions and recommendations.

2. Review of monocular EO sensors

One of the first applications of VIS cameras for the pose estimation of an uncooperative target is represented by the Relative Navigation Sensor which flew as part of the Hubble Space Telescope (HST) Servicing Mission 4 (SM4). The camera suite consisted of three monocular cameras operating at long (28 m - 260 m), medium (6 m - 40 m) and short (2 m - 5.5 m) range [12] to aid the estimation of pose of the target telescope, assumed to be unknown. Subsequently, inspired by the promising applications of existing visual-based systems for present and future FF missions and in-orbit servicing missions, many authors continued with the investigation of the feasibility of VIS cameras for the pose estimation of uncooperative spacecraft. Du et al. [13] proposed a scheme which combines a singular VIS camera, in the closing (15 m - 300 m) and mid-range (5 m - 15 m) phases, with two collaborative monocular VIS cameras in the final approach phase (1 m - 5 m), in order to increase the camera FoV and aid the feature extraction within the IP system. The cameras were used to estimate the pose of large non-cooperative satellites in Geostationary Earth Orbit (GEO). Liu and Hu [14] evaluated the performance of a pose estimation method for cylinder-shaped spacecraft which makes use of single images from a monocular VIS camera, whereas D'Amico et al. [15], Sharma and D'Amico [16] and Sharma et al. [9,17] used images collected by the monocular VIS camera onboard the PRISMA mission to investigate the robustness of several pose estimation schemes with respect to image noise, illumination conditions and Earth in the background geometries. Furthermore, Schnitzer et al. [18] included two monocular VIS cameras in the sensors suite adopted in their on-ground testing of image-based non-cooperative rendezvous navigation, and Pesce et al. [19] adopted a single passive monocular camera to reconstruct the pose of an uncooperative, known target. Despite the differences in the experimental setup, as well as in the pose estimation schemes, a common feature that was found for VIS cameras, even for cooperative pose estimation, is their strong dependency on the Solar or Earth illumination, which becomes more severe when the target does not have any fiducial marker.

On the other hand, TIR cameras are infrared cameras sensitive to the mid- and far-infrared spectral ranges (3 μm - 14 μm). Due to size, complexity, and power consumption of cryogenically-cooled infrared sensors, the current state-of-the-art on TIR cameras for spacecraft relative navigation relies on uncooled microbolometers operating in the range 8 μm - 14 μm, as they can provide sufficient sensitivity at low cost [20]. This type of sensor was flight-tested as part of the LIRIS demonstrator during the ATV5 Mission [21] as well as part of the Raven ISS Hosted Payload [22], and it has been used in Ref. [23] as well as in Ref. [24] and in Ref. [18] to assess the robustness of a TIR-based navigation system for ADR and to validate a pose estimation method based on feature extraction, respectively. Also, Shi et al. [25–27] used synthetic and real TIR camera images to validate a model-based and an appearance-based pose estimation methods, respectively. Notably, the TIR camera in Ref. [22] was fused with a visual camera and a flash LIDAR in order to improve the overall sensors performance.

When compared to VIS cameras, TIR cameras do not depend on external light sources but rather on the emitted thermal radiation of the target spacecraft, thus avoiding any saturation due to Sun presence in the camera FoV or Earth in the background. This makes the sensor more robust against the different illumination conditions, typical of an ADR scenario [28]. On the other hand, their image resolution is usually much lower than VIS camera. As reported in Ref. [23], the amount of blur in the image can significantly affect the performance of feature detection algorithms within the IP system. Also, the results of the tests with real TIR camera images in Ref. [18], in which a scaled model of the Envisat was heated through resistors mounted on the rear of the plates and a Halogen lamp was used for the illumination, demonstrated that real TIR images clearly differ from synthetic images. More in particular, Barrel distorsion was found to be more severe than the one modelled in the synthetic dataset, and the edges of the spacecraft silhouette were found more faded in the real images compared to the synthetic ones. Furthermore, the different thermal dynamics encountered during an ADR mission due to varying temperature profile of the target over one orbit, as well as the different thermal surface coatings of the target, introduce some challenges in the imaging. As an example, the performance of the method proposed in Ref. [25] cannot be evaluated due to the too optimistic assumptions of the thermal environment of the target. Furthermore, as stated in Ref. [27], the resolution of TIR images sensibly affects the accuracy of the pose determination in the training phase of a non-model based method.

Finally, NIR cameras are cameras which operate in the spectral range from 780 to 2500 nm. As such, current CMOS/CCD technologies can be adopted to sense the incoming NIR radiation, and a superior image quality compared to TIR microbolometers can be achieved. To the best of the authors' knowledge, the only pose estimation scheme so far tested with NIR images is based on a model-based IP in which the camera suite combines VIS/NIR/TIR images to increase the robustness of the pose estimation.[1] This work was part of a Technology Research Programme (TRP) study, sponsored by the European Space Agency (ESA) and called Multi-spectral Sensing for Relative Navigation (MSRN), which focused on the design of a multispectral camera that can be used for navigation purposes in a wide variety of scenarios. This activity focused on increasing the accuracy and robustness of normal multispectral cameras by combining a Visual-Near Infra-Red (VNIR) spectral channel to a TIR spectral channel [29]. In this way, the benefits of each single camera type, listed in Table 1, can be combined to return a superior performance of the camera suite. Fig. 1 illustrates the different coupling schemes proposed. Data fusion both at image and image processing levels was investigated in order to comply with the requirements of a robust and computationally fast IP prior to the navigation filter.

The current state of the art on monocular cameras is further reviewed by focusing on the applicability of the proposed camera suites for the desired operational range, considering the requirement to have a robust pose estimation of an uncooperative target from several hundreds of meters down to docking, which characterises most of the close-proximity rendezvous missions. Table 2 lists some relevant characteristics of the camera suites and reports the tested range of the pose estimation simulations. Naasz et al. [12] and Cavrois et al. [21] tested

[1] https://www.esa.int/Our_Activities/Space_Engineering_Technology/Shaping_the_Future/Multispectral_Sensing_for_Relative_Navigation.

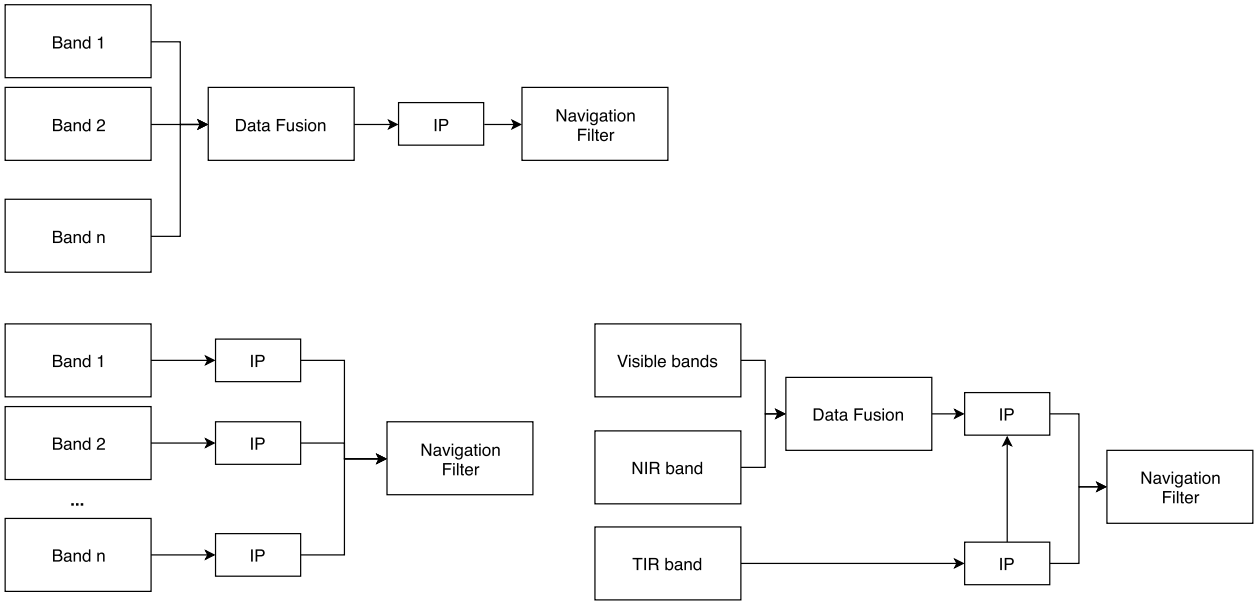

Fig. 1. Illustration of the cameras coupling schemes investigated during ESA's MSRN programme. The selected third scheme combines the advantages of relying on data fusion prior to the IP (Scheme 1) with the benefits of having separate channels, which improves the system robustness in case of failure in either the VNIR or in the TIR band (Scheme 2).

Table 2
Characteristics of the camera suites adopted in different pose estimation schemes and their tested range.

Ref.	Camera Suite	Tested range	FoV [deg]
[12]	3 monocular VIS cameras	150 m - 1 m	11/23/23
[13]	monocular + collaborative VIS cameras	300 m - 1 m	55
[14]	Monocular VIS camera	40 m - 5 m	–
[9,15–17]	Monocular VIS camera	13 m - 8 m	22.3–16.8
[21]	3 Monocular VIS/TIR cameras	70 km - 8 km 3.5 km - docking	60 × 45
[25–27]	Monocular TIR camera	~5 m	40
–	Monocular VNIR/TIR camera[1]	far range - 7 m	40 × 40 VNIR 40 × 30 TIR
[23]	Monocular TIR camera	–	30
[18]	2 Monocular VIS/TIR cameras	100 m - docking	–
[24]	Monocular TIR camera	70 m - 21 m	–
[19]	Monocular VIS camera	< 30 m	–

monocular cameras down to 0.5 m from the target and down to actual docking, respectively. However, the challenges of feature extraction within the IP at close range were not investigated. As an example, with a FoV of around 23° and a distance from the target of around 0.5 m, the IP would need to extract features from a portion of the spacecraft as small as a 0.2 m-by-0.2 m rectangle, which can be challenging if the satellite is relatively large. On the other hand, the claim in Ref. [13] that collaborative cameras are strictly required for the close approach phase relates to the fact that their selected IP scheme is based on the extraction of large rectangular features of large communication GEO satellites. Other authors investigated several different pose estimation schemes which rely on more flexible feature extractions. However, their pose estimation systems were not tested for relative ranges below 5 m. It can be concluded that some effort is still required to assess whether a single monocular camera can be used for close-proximity pose estimation of an uncooperative target or if collaborative cameras are needed. As a general remark, it should in principle be possible to rely on a single monocular camera when the target is fully in the camera FoV, and switch to the feature tracking of the desired docking port for closer ranges, as performed in Ref. [18]. Furthermore, several orbit scenarios should be recreated in future tests in order to investigate the robustness and applicability of each type of monocular camera as well as a combined VNIR/TIR camera suite for multispectral imaging. The scheme in Fig. 1, as well as the one proposed in Ref. [22] provided that no LIDAR systems are considered, shall be investigated. Finally, the infrared characteristics of the target spacecraft should be fully understood in order to maximize the performance of the NIR/TIR cameras. Although Yilmaz et al. [30] proposed an infrared signature estimation method capable of characterizing the dynamical thermal behaviour of space debris, some effort is still required to assess its validity and to confirm whether an exact infrared appearance model of the target is needed for a robust relative navigation solution which relies on IR images.

3. Monocular pose estimation

Monocular pose estimation consists in estimating the relative pose of a target spacecraft with respect to the servicer spacecraft by only using 2D images, either taken by a monocular camera or fused from more monocular cameras (Fig. 1), as measurements. In other words, monocular pose estimation is associated to the computation of *pseudomeasurements* of the relative pose from the input image, prior to the navigation filter. From a high level perspective, the architecture of the pose estimation process usually involves an acquisition step, or initialization, in which there is no a-priori information on the target pose, and a tracking step, in which knowledge from the previous estimates is used when new images of the target are acquired. In both cases, estimation methods can be divided into model-based and non-model based. Model-based pose estimation makes use of a simplified wireframe 3D model of the target and it is described in detail in Section 3.1. On the other hand, non-model based methods estimate the spacecraft pose

without using an existing 3D model of the target. In this review, appearance-based and feature-based methods are considered. In appearance-based methods, the pose estimation is performed by comparing the 2D image with a pre-stored database of images and by minimizing the matching error between the in-flight image and each of the images in the database. As such, no feature extraction is required and thus no IP system is needed. Appearance-based methods are reviewed in Section 3.2.

In addition to the above-mentioned methods, CNNs are recently becoming a promising solution for the pose initialization of a target spacecraft. In a CNN-based method, the monocular image is fed into a pre-trained neural network, which solves a regression and/or a classification problem to return the predicted pose. Depending on the selected architecture adopted to solve for the relative pose, these methods can either rely on a wireframe 3D model of the target spacecraft or solely on the 2D images used in the training, and hence they can either be referred to as non-model based or model-based. Fig. 2 illustrates a high level representation of the monocular pose estimation methods reviewed in this paper. Feature-based methods are included beside the other pose estimation methods to underline that the features extracted by the IP algorithms could also represent input measurements for the navigation filter.

3.1. Model-based pose estimation

Model-based monocular pose estimation methods receive as input a 2D image and match it with an existing wireframe 3D model of the target spacecraft to estimate the pose of such target with respect to the servicer camera by extracting some features from the 2D image (IP system, described in Section 3.1.1) and by matching these features to the corresponding elements of the 3D model. Then, the relative pose is obtained by solving the Perspective-n-Points (PnP) Problem described in Section 3.1.2. Interested readers are referred to Ref. [10] for a more detailed overview on *template matching* as an alternative to solving the PnP problem.

3.1.1. IP algorithms

The IP system is a fundamental step for feature-based pose estimation, and several methods exist in literature to extract and detect target features from a monocular 2D image, based on the specific application. From a high-level perspective, the target features can be divided into keypoints (or interest points), corners, edges and depth maps. Table 3 provides a list of the IP schemes reviewed in this Section.

Naasz et al. [12] accomodated two different IP within their Relative Navigation Sensor (RNS) system: a Sobel edge-enhancing image filter to process a 10-bit camera image and perform the edge extraction, also adopted in Ref. [22], and a digital correlation image processing technique which computed the position of certain features of the target spacecraft. These two methods were used separately by different pose estimation systems which were tested during the HST-SM4. Several realistic lighting conditions were recreated to validate the robustness of the IP algorithms with respect to illumination. Du et al. [13] included a median filter before the other steps of the IP to cope with image noise and smooth the data. The Canny edge detection algorithm was selected to detect edges in the image, and a subsequent Hough transform (HT) [31] was used to extract the detected lines. Several tests were conducted to assess the robustness of the IP with respect to image noise at different variance levels. However, a limitation of their method was that it focused on the extraction of rectangular structures on a large target spacecraft. Liu and Hu [14] presented a robust method based on ellipses extraction for cylinder-shaped spacecraft, but its application is not feasible for the pose estimation of a spacecraft of generic shape.

D'Amico et al. [15] used the same feature detection and extraction methods in Ref. [13] in combination with a Low-Pass Filter (LPF). Its method was tested with the PRISMA image dataset and proved to be flexible with respect to the spacecraft shape, but it lacked of robustness to illumination and background conditions. Furthermore, it did not prove to be robust with respect to the spacecraft symmetry. Shi et al. [25] selected the Roberts Cross Method (RCM) in combination with the Harris Corner Detection (HCD) method to improve the computational time of the IP. However, the limitations of the RCM in producing less edges than the Canny's were not assessed. Shi et al. [26] implemented a Contrast Limited Adaptive Histogram Equalization (CLAHE) to clean and restore blurred TIR images. A Scale Invariant Feature Transform (SIFT) [32], in combination with the Binary Robust Independent Elementary Features (BRIEF) method [33], was used to extract the target interest points from the denoised image. The RANdom SAmple Consensus (RANSAC) [34] algorithm was further included in the IP scheme in order to quickly extract image features and descriptors by using some internally pre-stored test image features for feature matching.

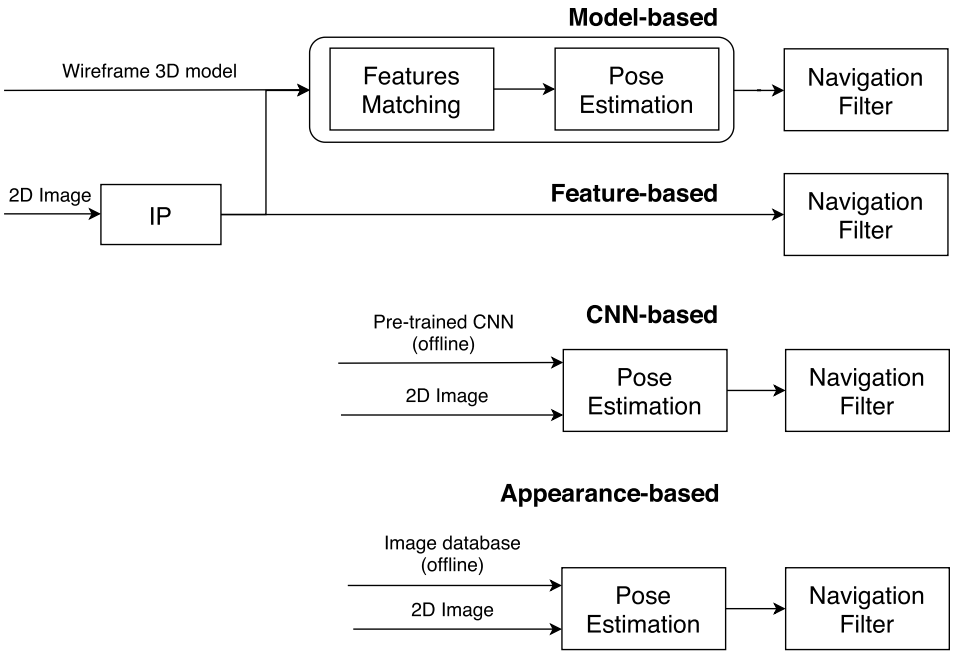

Fig. 2. High level architecture of monocular pose estimation methods reviewed in this paper.

Table 3
Characteristics of state-of-the art IP algorithms. Here, NA refers to the fact that no robustness test could be found in the reference. Notice that no information on the robustness is reported for TIR-based systems, given the negligible Earth's emittance in the TIR band.

Ref.	IP	Tested Range	Robust w.r.t. Earth in the background	Offline Database required
[12]	Digital corr./Sobel	150 m - 1 m	NA	No
[13]	Canny + HT	300 m - 1 m	NA	No
[14]	Ellipses extraction	40 m - 5 m	NA	No
[15]	LPF + Canny + HT	13 m- 8 m	NA	No
[25]	RCM + HCD	~5 m	–	No
[22]	Sobel	NA	NA	No
[26]	CLAHE + SIFT/BRIEF + RANSAC	–	–	Yes
[24]	Canny	100 m - 21 m	–	Yes
[37]	FREAK + EDL	NA	NA	Yes
[9]	WGE + S/HT	13 m - 8 m	Robust when Earth's horizon is not in the FoV	No
[19]	GFTT	< 30 m	NA	No
[38]	Prewitt + gradient filter ST + HT + LSD	45 m - 5 m	Robustness proven (including Earth's horizon in FoV)	No

Yilmaz et al. [35] performed an evaluation of the invariance of edge and corner detectors applied to TIR images. The Good Feature to Track (GFTT), Speeded Up Robust Features (SURF) and Phase Congruency Point (PC–P) edge algorithms, as well as edge detectors such as the Sobel, were traded-off based on their robustness under different thermal conditions representative of the dynamic space thermal environment. Their results showed that thermal variations can cause significant variation in the thermal signatures, and thus challenge the robustness of pose estimation methods based on feature extraction. Rondao et al. [36] also investigated the performance of several keypoint detectors applied to VIS/TIR synthetic images. In his work, the combination of the Fast-Hessian feature detector with the Binary Robust Invariant Scalable Keypoints (BRISK) descriptor proved to have comparable performance in both spectra, resulting in a promising option when reduced memory usage represent a key requirement.

Gansmann et al. [24] adopted the Canny algorithm to extract edges from TIR images and from a 2D rendered representation of the target, obtained by projecting a 3D model. The variation in brightness and the variation in depth were used to extract the edges from the TIR images and from the render, respectively. Furthermore, Rondao and Aouf [37] adopted a Fast Retina Keypoint (FREAK) descriptor in combination with the Edge Drawing Lines (EDL) detector to extract keypoints, corners, and edges to find the correspondence between features. In their method, a depth mapping was further performed which aided the features extraction. The limitation of these two latter methods is that they require an offline database for image matching. More recently, Sharma et al. [9] proposed a novel technique to eliminate the background of images, called Weak Gradient Elimination (WGE). After using a Gauss filter to blur the original image and aid the feature extraction, the image gradient intensities were computed, and the WGE was used to threshold the weak gradient intensities corresponding to the Earth in the background. In the next step, the Sobel algorithm and the Hough Transform (S/HT) were used to extract and detect features. Notably, the WGE technique can also be used to identify a rectangular region of interest (ROI) in the image which can allow an automated selection of the hyperparameters required by the HT. In this way, the hyperparameters are automatically scaled based on the varying distance from the target. By creating two parallel processing flows, the method proved to be able to extract main body features as well as particular structures such as antennas, and thus to solve the symmetry ambiguity which characterized other IP schemes. Furthermore, the implementation of the WGE method returned a much higher robustness with respect to Earth in the background compared to the other methods. However, scenarios in which the Earth horizon is present in the background represented a challenge for the IP due to an improper ROI detection.

Alternatively, Capuano et al. [38] introduced a new IP scheme in which three different parallel processing streams, which use the Shi-Tommasi (ST) corners detector, the HT, and the Line Segment Detector (LSD), are exploited in order to filter three sets of points and improve the robustness of the feature detection. This was performed in order to overcome the different drawbacks of each single method. Feature fusion was then used to synthesise the detected points into polylines which resemble parts of the spacecraft body. By including a background removal step similar to the WGE in Ref. [9], which makes use of a Prewitt operator in combination with a gradient filter, the authors could also demonstrate the robustness of their IP with respect to the Earth in the background. Furthermore, the scenarios with the Earth horizon were tackled by tuning the threshold of gradient filter to a more selective value. The last three feature extraction schemes [9,37,38], which combine several keypoints, edges and corners detectors, are depicted in Fig. 3.

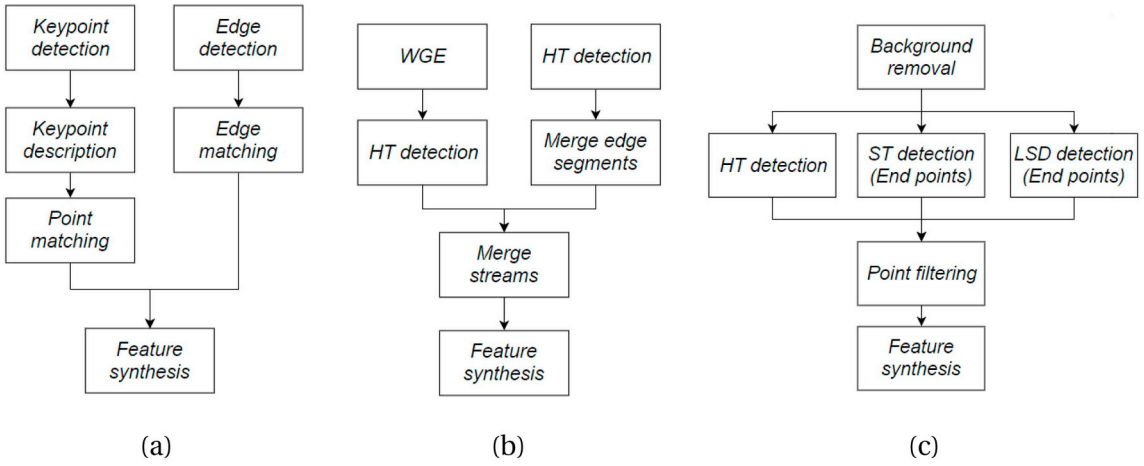

Fig. 3. Examples of feature synthesis schemes. (a) [37], (b) [9], (c) [38].

Finally, Pasqualetto et al. [39] investigated the potentials of using a hourglass neural network [40] to extract the corners of a target spacecraft prior to the pose estimation. In this method, the output of neural network is a set of so called *heatmaps* around the features used in the offline training offline. The coordinates of each heatmap's peak intensity characterize the predicted feature location, with the intensity indicating the confidence of locating the corresponding keypoint at this position. Despite a lack of actual space imagery to test the network performance, the proposed method proved to be capable of detecting features which are either not visible due to adverse illumination or occulted by other parts of the target, when trained and tested with synthetic images. Due to these characteristics, the proposed method could emerge as a promising alternative to state-of-the-art IP algorithms. However, the robustness of the features extraction with respect to the Earth in the background was not fully proven, and the impact of an inaccurate detection on the pose estimation accuracy was not assessed.

As a general remark, IP algorithms based on keypoint features detectors present some advantages compared to algorithms based on edge and corner detectors, given their invariance to perspective, scale and illumination changes [32,41]. However, they could still be sensitive to extreme illumination scenarios. Moreover, their robustness with respect to outliers, which would be present when the Earth is in the image background, has not been fully proved yet in the framework of relative pose estimation in space. On the other hand, the recent advancements in the IP algorithms based on corners/edges detection showed an improvement in the robustness of such methods with respect to the Earth in the background [9]. Furthermore, edges and corners detectors are retained to be more robust than features detectors in case of partial occlusion of the target, especially during tracking [42]. Future works should focus on the assessment of the robustness of keypoint features detectors to outliers in space imagery, as well as in combining such IP methods with edges/corners detectors in order to benefit from the advantages in both algorithms, similarly to what has been proposed in Ref. [37].

3.1.2. Pose estimation methods

The features detected by the IP algorithms described in Section 3.1.1 can be directly used as measurements in a navigation filter to solve for the pose of the target spacecraft. This is usually performed when the extracted features are represented by points. However, pseudomeasurements of the relative pose are usually computed from the extracted features and a wireframe 3D model of the target by solving a pose initialization problem. Referring to Fig. 4, the pose initialization problem consists in determining the position of the target's center of mass $\mathbf{t}^C$ and its orientation with respect to the camera frame C, represented by the rotation matrix $\mathbf{R}_B^C$. The 3D/2D true perspective equations,

$$\mathbf{r}^C = \mathbf{R}_B^C \mathbf{q}^B + \mathbf{t}^C, \tag{1}$$

$$\mathbf{p} = (u_i, v_i) = \left(\frac{x^C}{z^C} f_x + C_x, \frac{y^C}{z^C} f_y + C_y \right), \tag{2}$$

relate the unknown pose with the corresponding point $\mathbf{p}$ in the image plane. Here, $\mathbf{q}^B$ is a point in the 3D model, expressed in the body-frame coordinate system B, whereas f_x and f_y denote the focal lengths of the camera and (C_x, C_y) are the principal points of the image. Since solving the P*n*P problem requires an image processing suite that extracts target features from a given image, Eqns. (1) and (2) do not have to be solved for non-model based estimators such as CNN-based or appearance-based.

Several methods exist in the literature to solve for the initial pose of an uncooperative target. Based on two different surveys by Opromolla et al. [10] and Sharma and D'Amico [11], the most commonly used solvers can be identified as the PosIt [43] and Coplanar PosIt [44], the SoftPOSIT [45], the EP*n*P [46] and the Newton Raphson Method (NRM). In Ref. [37], the EP*n*P solver was used to initialize the relative pose, which was further refined by means of an M-Estimator minimization to increase the robustness with respect to erroneous correspondences between features. In their method, the Rodrigues parameters were used to represent the relative attitude in order to handle a 6×1 pose vector. In a recent effort, Sharma et al. [9] further proved that the EP*n*P method has the highest success rate and offers a superior performance in terms of both pose accuracy and runtime when compared with other state-of-the-art PnP solvers. In their estimation scheme, the NRM was also used after the EP*n*P to refine the final pose estimation. The idea behind such P*n*P solver switch is that, since EPnP has the lowest runtime, it can be used when large number of correspondence hypotheses need to be validated within the first iterations. Once the search space for correct feature correspondence has been reduced, NRM can be used due to its better accuracy in the presence of outliers and noise [11]. Furthermore, Pesce et al. [19] proposed a novel pose estimation scheme in which the RANSAC algorithm is used in combination with the Principal Component Analysis (PCA) to generate subsets of image-model correspondences, so called *consensus sets*. For this purpose, the features extracted with the GFTT algorithm were compared with an off-line feature point classification of a simplified 3D model. Once the correspondences are set, the EP*n*P is used to solve for the pose initialization. The SoftPosIt algorithm was further included to solve for the pose tracking. Due to the capability to detect particular spacecraft components, their estimation scheme proved to be robust with respect to spacecraft symmetry.

Aside from the listed solvers adopted to solve the pose initialization problem, other authors [12,22] implemented the technique proposed in Ref. [47] and the ULTOR engine [48] in their Goddard Natural Feature Image Recognition (GNFIR) and ULTOR algorithms, respectively, for the pose tracking. As opposed to PnP solvers, this technique makes use of the Lie group SO(3) to find and measure the distance between a rendered model of the target and the matching nearby edges in the image. In their works, the GNFIR algorithm was adopted to perform edge tracking once the pose initialization is acquired, whereas ULTOR could be used for both pose initialization and tracking. Additionally, Gansmann et al. [24] assumed the initialization to be known and implemented a tracking method based on [47] which uses an Iteratively

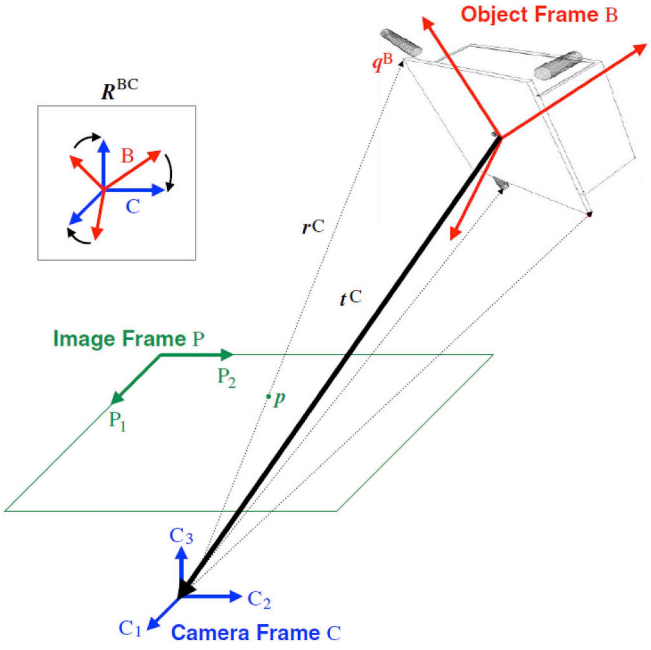

Fig. 4. Schematic representation of the pose estimation problem using a monocular image [9].

Re-Weighted Least Squares (IRLS) to get an a-posteriori pose via the interframe motion. Their algorithm minimized the squared residuals of model template edges, extracted from a 3D rendering of the target, to image query edges, extracted from each TIR image. Their tracking algorithm was tested for the distance of 100 m until 21 m and proved to return centimetric and sub-degree accuracy for the relative pose. However, convergence to local minima associated to a wrong pose represented an issue with the algorithm. A proposed solution to this problem was to perform a re-initialization of the pose estimation with an acquisition algorithm, as a sudden jump in the estimated pose would be easily detected due to the smoothness of the relative motion.

The comparative assessment of the different P*n*P solvers in Ref. [11] is reported in Table 4. Table 5 lists some characteristics of the different pose estimation solvers in relation to the IP methods described in Section 3.1.1. From the comparison, it can be concluded that the pose estimation scheme proposed in Ref. [9] is a good candidate for the pose initialization, given the robustness of its IP system and the fact that it has been tested for several illumination conditions as well as with the Earth in the background. The proposed system is in fact robust to the background of the images due to the WGE, it requires no a-priori knowledge of the target spacecraft's pose, and it is computationally efficient. In particular, this architecture shows improvements with respect to previous IP and pose estimation techniques [10,11,15]. Fig. 5 illustrates the main steps of the pose determination subsystem. However, some remarks shall be made about the images used for the validation of the pose estimation schemes. As reported in Table 4, most of the pose estimation schemes were tested with synthetic images in which the different reflectivities of spacecraft materials were not included. As such, the robustness of the algorithms with respect to realistic illumination conditions could not be assessed. Also, the limited amount of realistic space images available in Refs. [9,15,24] could not represent all the challenging orbital scenarios for which a specific camera-target-Sun-Earth geometry would affect the pose estimation accuracy.

Following the recommendations in Ref. [18], real VIS/TIR/NIR images should be accounted for early in an activity to avoid validating navigation algorithms with synthetic images which considerably differ from the ones taken in orbit. In the future, image acquisition tests should be conducted on ground with real cameras and S/C mock-ups, in order to solve both the low representativeness of synthetic images and the limited amount of actual space imagery. Furthermore, since the generation of representative TIR images in a laboratory environment requires the spacecraft model to have thermal signatures which are usually difficult to reproduce, an additional effort will be required in order to account for thermal effects as well as to hide the image background. It is worth mentioning that, due to the fast variation in the space thermal environment, a model-based method could be unfeasible when using TIR images. As anticipated in Refs. [23,35], the different thermal inertia of spacecraft materials could result in a mismatch between the off-line TIR model and the time-varying extracted features and could thus lead to inaccurate relative pose estimates. An idea could be to adopt a model-based pose estimation which uses images from a VIS camera in combination with a non-model based method which uses images from a TIR camera. In this way, the limited observability which results from the TIR-based estimation could be solved, and both the robustness and the accuracy of the pose estimation improved.

3.2. Appearance-based pose estimation

Compared to feature-based methods, in which the IP is used to extract features such as corners and edges, only the spacecraft *appearance* is used in appearance-based methods. Depending on whether a 3D model of the target spacecraft is used or not, appearance-based methods can be classified as model-based and non-model based, respectively. Opromolla et al. [49] proposed a model-based pose framework for spacecraft pose estimation. However, the framework was designed to process 3D point clouds and thus its application was constrained to LIDARs or stereovision systems. To the best of the author's knowledge, the only appearance-based method for spacecraft pose estimation based on a monocular camera was proposed by Shi et al. [27], and it is based on PCA.

The pose matching algorithm is separated into an off-line training portion and a testing portion that computes the pose of the spacecraft in-flight. The PCA algorithm matches the object from the camera image (test image) to a stored matrix of images that has been transformed to its eigenspaces during the training phase. The advantage of PCA stands in the fact that the dimension of the training dataset can be drastically reduced by considering only the principal eigenvectors of the training data matrix. However, the test image needs to be compared to each image of the training dataset at each pose solution, which still requires a considerable computational effort if the number of training frames is large. In Ref. [27], the validation of the algorithm was performed with $M = 12.660$ frames as a result of a trade-off between the computational time and the estimation accuracy. The resulting mean search time was found to be approximately 62.8 ms, which is relatively low for uncooperative pose estimation. However, the PCA algorithm performance was proved to degrade with the image noise, which is unwanted due to the noisyness of actual space imagery. Furthermore, one of the assumptions for the PCA is that the object must be completely visible, which might not be the case if part of the spacecraft falls outside the camera FoV. Finally, as the validation was not performed with the Earth in the background, it is unclear whether the pose estimation is robust against other objects present in the camera image, as one of the main requirements of PCA is that each image shall contain a single, non-occulted object.

3.3. CNN-based pose estimation

From a high-level perspective, CNNs are neural networks built from multiple dual-layers of convolutional masks which were inspired by the human visual cortex. Given their capability of classifying images, their implementation in monocular pose estimation has become attractive in recent years [50]. A pose estimation architecture based on CNNs does not distinguish between an IP subsystem and a pose estimation subsystem, but rather between an off-line training phase and an in-flight test phase. The advantage of CNNs over feature-based algorithms is an increase in the robustness for adverse illumination condition, as well as a reduction in the computational complexity. However, compared to terrestrial applications, space imagery are characterized by high contrast, low signal-to-noise-ratio and low sensor resolution. As such, their accuracy is expected to be lower. Usually, due to the lack of a large synthetic dataset of space images, which is usually required to fully train a CNN, a network which has been pretrained on a dataset of terrestrial images is used, and transfer learning is applied to train only a limited number of layers of the convolutional network.

A CNN architecture for pose estimation for uncooperative spacecraft has been proposed in Ref. [17]. Synthetic datasets of up to 125.000 space images were created, for which a 3D texture model of the target spacecraft was required. The architecture of the *AlexNet* network [51] was then adopted as the baseline architecture, and a classification problem was solved to return the relative pose of the target spacecraft associated to each image. Transfer learning was used to train the last

Table 4
Comparative assessment results from simulations as a qualitative decision matrix in Ref. [11]. Here, PosIt + refers to a solver that can switch between Coplanar PosIt and PosIt.

Solver	Number of Features	Noise	Outliers	Distance to Camera
PosIt	Nominal	Superior	Inferior	Nominal
EP*n*P	Superior	Par	Inferior	Inferior
PosIt +	Nominal	Superior	Inferior	Nominal
NRM	Superior	Superior	Nominal	Nominal

Table 5
Characteristics of state-of-the art model-based pose estimation schemes. Here, NA refers to the fact that no robustness tests could be found in the reference.

Ref.	IP	Pose Initialization/Tracking	Tested Range	Robust w.r.t. symmetry	Validation Database
[12]	Digital corr./Sobel	ULTOR/GNFIR	150 m - 1 m	NA	Flight spare cameras/Lab pictures
[13]	Canny + HT	Analytical	300 m - 1 m	NA	Synthetic images Realistic camera model No materials' reflectivity
[14]	Ellipses extraction	NRM	40 m - 5 m	Yes	Synthetic images Ideal camera model No materials' reflectivity
[15]	LPF + Canny + HT	Perceptual Groups + NRM	13 m - 8 m	No	Actual space imagery (PRISMA)
[25]	RCM + HCD	SoftPosIt	~5 m	NA	Synthetic images Camera model not given No materials' reflectivity
[22]	Sobel	GNFIR	NA	–	–
[26]	CLAHE + SIFT/BRIEF + RANSAC	EPnP/SoftPosit	–	NA	Synthetic and lab TIR images Camera model not given No materials' reflectivity
[24]	Canny	IRLS	100 m - 21 m	NA	Actual space imagery (ISS)
[37]	FREAK + EDL	EPnP/RANSAC + M-estimator	NA	Yes	Synthetic images Camera model not given Materials' reflectivity included
[9]	WGE + S/HT	EPnP + NRM	13 m - 8 m	Yes	Actual space imagery (PRISMA)
[19]	GFTT	RANSAC + PCA + EPnP/SoftPosIt	< 30 m	Yes	–

fully-connected layers using a subset of up to 75.000 images from the synthetic datasets (Fig. 6), while the first layers were trained with the ImageNet dataset. This was performed by means of transfer learning on the last three fully-connected layers. Shi et al. [52] used two state-of-the-art CNNs, namely *Inception-ResNet-V2* [53] and *ResNet-101* [54], in combination with an object detection engine [55] to improve their reliability. Synthetic images generated in the 3DS-Max software were used in combination with real images to train and test the two networks, specifically 400 and 100 images, of which 8% were real images, were used for training and testing the networks, respectively. Transfer learning was also performed to adapt the pre-trained networks to the pose classification of a target spacecraft.

In a recent effort, Sharma and D'Amico [56] proposed a novel network based on five convolutional layers and three separate branches (Fig. 7). In the first branch, the Region Proposal Network (RPN) proposed in Ref. [55] detects a 2D bounding box around the target spacecraft. In the other two branches, three fully-connected layers are used to solve a classification and a regression problem, respectively, and to output the relative attitude of the target spacecraft. Then, the bounding box information is used together with the attitude information to solve for the relative position by minimizing the distance between the corners of the bounding box and the extremal points of a wireframe 3D model of the target. The training was performed with 12.000 synthetic images of the TANGO spacecraft, whereas two test sets were created with 3.000 synthetic images and 300 actual camera images, respectively. Furthermore, half of the synthetic images included the Earth in the background.

The CNN-based algorithm in Ref. [17] has been extensively tested against the number of synthetic images used in the training, different levels of image noise and the amount of displacement of the target from the center of the image plane, which has not been tested in the validation of other pose estimation algorithms. However, several improvements are proposed in the paper. First of all, the CNN should be trained with actual space imagery. This can be clearly seen in Table 6, in which the pose errors considerably increase when the network is tested with real images. Also, larger datasets shall be considered for a comprehensive comparative assessment of the CNN architecture with the conventional pose determination architectures. Furthermore, assumptions on the illumination environment, target texture and reflectance properties shall be investigated to increase the robustness of the pose estimation, and different CNNs, such as the GoogLeNet, the ResNets and the DenseNet, shall be traded-off with respect to computational time and accuracy in the pose estimation, following the promising results reported in Ref. [52] for the *Inception-ResNet-V2* and *ResNet-101*. The scheme proposed in Ref. [56] proved to return better pose estimates than the AlexNet scheme while at the same decreasing the size of the training set, as well as a comparable accuracy in the 2D bounding box detection compared to the architecture in Ref. [56]. Furthermore, it proved to be robust with respect to the Earth in the background. However, its performance was found to drop-off at relatively close distances for which the target is not fully in the camera FoV as well as during poor illumination conditions close to eclipse, due to inaccurate box detections. Notice also that, since the training in Refs. [17,56] has been performed with relative distances from 3 up to 50 m as labels, the estimation system for close-proximity operations down to docking could not be validated.

Despite the relatively coarse accuracies in the pose estimation, especially in the relative attitude, neural networks could still improve the pose initialization. As mentioned in Ref. [17], a feature-based algorithm with a CNN-based pose estimation, which provides a coarse initial guess, could increase the robustness of the pose initialization with respect to scenarios in which the IP fails in extracting the target features from the image background.

Finally, none of the previous CNN-based pose estimation methods were tested in a navigation filter, and some effort is still required in the modeling of the measurement noise when neural networks are adopted

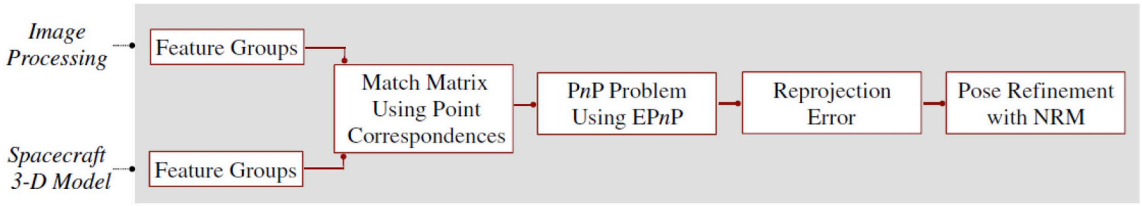

Fig. 5. Novel pose determination subsystem proposed in Ref. [9].

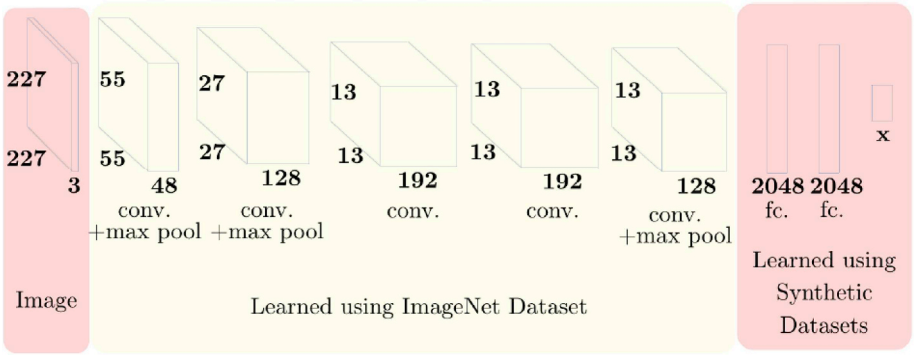

Fig. 6. Illustration of the AlexNet architecture adopted in Ref. [17].

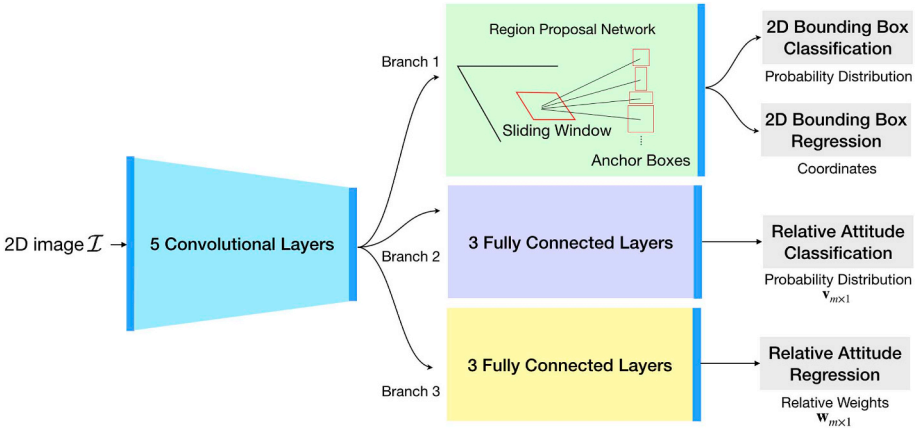

Fig. 7. Illustration of the CNN architecture adopted in Ref. [56].

prior to the filter estimation. It is also important to notice that, if the target shape during operation considerably differs from the one assumed during the training phase, the reliability of CNNs might be affected. Future works shall assess the impact of such uncertainty in the target shape on the pose estimation accuracy, as well as investigate the benefits of CNN-based schemes over feature-based schemes.

4. Visual-based navigation filters

The relative pose estimation schemes described in Section 3 provide an initial estimate of the relative position and attitude of a target spacecraft with respect to the servicer spacecraft for lost-in-space scenarios, in which no a-priori information of the relative state is available. This is referred to as the *pose initialization subsystem*. Once the initial guess on the relative state is computed from the estimation scheme, *pose tracking* can be performed by collecting a new camera image and using the previous state as the new initial state for a subsequent pose initialization problem. However, the pose initialization routines are not well suited to produce pose estimates at high frequencies, especially due to the computationally expensive IP in combination with the PnP solvers. Therefore, a relative navigation filter shall be used in combination with the camera measurements and the pose estimation suite in

Table 6
Comparison of CNN architectures for relative pose estimation. Here, the mean position and attitude errors, E_T and E_R, are reported together with the Intersection-Over-Union (IoU) metric, which measures the accuracy of the 2D bounding box detection.

Ref.	Architecture	Training/Test Set Images	E_T [m]	E_R [deg]	IoU
[17]	AlexNet	75.000/50.000 synthetic	0.12	11.94	–
	(3.000 pose labels)	75.000/25 real	1.12	30.75	
[52]	101-layer ResNet Inception ResNet V2 (with RPN)	400/100	–	–	0.88 0.88
[56]	Convolutional layers +	12000/3000 synthetic	[0.055, 0.046, 0.78]	8.4	0.8582
	RPN + Fully-connected layers	12000/300 real	[0.036, 0.015, 0.189]	18.19	0.8596

Table 7
Comparison of navigation filters for relative pose estimation, together with the adopted performance validation method. Here, NS refers to papers in which the adopted filters were not specified.

Ref.	Translational filter	Rotational filter	Performance Validation Method
[12]	Linear KF	MEKF	Ground-based test on HST mockup
[16]	MEKF	MEKF	Numerical simulations
[67]	Linear KF	Linear KF	HIL in closed GNC loop
[22]	MEKF/Schmidt KF	MEKF/Schmidt KF	Numerical simulations
[69]	D-Q MEKF	D-Q MEKF	Ground-based experimental test
[70]	NS	NS	SIL/HIL in closed GNC loop
[71]	DA filters	DA filters	Numerical simulations
[72]	–	Minimum Energy Filter Attitude Observer 2nd Order Minimum Energy Filter MEKF	Numerical simulations
[19]	H_∞ filter	2nd Order Minimum Energy Filter	Numerical simulations

order to return relative state solutions at high frequency [16]. Furthermore, the internal dynamics of the filter improve the accuracy of the predicted relative state from measurements and allow a more robust pose tracking. From a high level perspective, two different relative navigation architectures are usually exploited in the framework of the relative pose estimation of an uncooperative target. A *tightly-coupled* architecture, where the extracted features are directly processed by the navigation filter without exploiting any model-based method, and a *loosely-coupled* architecture, in which the relative pose is already determined prior to the navigation filter, i.e. by adopting a model-based method. When dealing with uncooperative tumbling targets, a loosely-coupled approach is usually preferred since the fast relative dynamics could jeopardize the robustness of features tracking, provided that a simplified geometrical model of the target is available. On the other hand a tightly-coupled approach is the best option when dealing with unknown targets, since it does not rely on any a-priory knowledge of the target geometrical model.

In the framework of spacecraft relative motion, several representations of a linearized relative state exist based on the intersatellite range, orbital eccentricity and perturbation forces involved. Linearized models are required when the filter internal dynamics needs to be linearized, as it is the case for linear Kalman FIlter (KF) and Extended Kalman Filter (EKF). Ref. [57] provides a detailed overview on closed-form dynamics model suited for onboard relative navigation. Notice that, for ADR and On-orbit servicing, the target orbit can usually be assumed to be circular, thus simplifying the computational burden that results from not neglecting the orbital eccentricity of satellite orbits. Generally, a distinction is made between models which make use of a Cartesian representation of the relative state (position and velocity) and models which consider a set of the Relative Orbital Elements (ROE). Notably, perturbation models can be easily accommodated in the filter dynamics in the latter case [58–60]. Clearly, a linearized model is not required if nonlinear filters are adopted. On the other hand, in the context of spacecraft relative attitude, several linear and nonlinear models exist based on either Euler angles, quaternions and Modified Rodrigues Parameters (MRP) [61–63].

Navigation systems for close-proximity operations have been extensively validated in the context of RF and monocular vision navigation for FF and on-orbit servicing, when the target is cooperative [1,61,64–66]. However, there is still a lack of a comprehensive validation of navigation systems for the pose estimation of an uncooperative target. As an example, the EKF and the Unscented Kalman Filter (UKF) presented in Refs. [61,66], respectively, rely on the availability of gyro measurements from each spacecraft, which is usually not the case for uncooperative spacecraft in ADR scenarios. When the uncooperative target is known, it is assumed that a simplified geometrical model of the target is available and representative of the target state in orbit. As such, when a model-based pose estimation method is adopted prior to the navigation filter, the 3D model of the target can be assumed to be reliable, and the navigation system can estimate the relative pose based on the pseudomeasurements derived from the extracted features of the target without including uncertainty in the geometrical model. However, if the shape of the target has changed due to orbit degradation and/or due to unforeseen events, the assumptions on its state made in the simplified geometrical model might differ from its real conditions in orbit. Furthermore, the target's mass and moment of inertia, together with other relevant parameters, might differ from the assumed values. As such, the navigation filter might have to estimate additional parameters aside from the relative pose.

4.1. Design and validation of monocular navigation systems: known targets

When dealing with uncooperative known targets, the state vector to be estimated in the navigation filter consists in the relative position, velocity, attitude and angular velocity between the chaser and the target. Additionally, if the relative dynamics between the servicer and the target spacecraft, modelled in the relative navigation system, account for perturbation models which might be inaccurate, key perturbation parameters should be included given the uncertainty of the dynamics models. As already mentioned, loosely-coupled navigation architectures are usually preferred when the target is known.

Table 7 lists the state-of-the-art for the navigation filters adopted in the framework of pose estimation of uncooperative known targets. Naasz et al. [12] implemented a Multiplicative Extended Kalman Filter (MEKF) [63] for attitude estimation and a linear KF for translation to estimate the pose of the HST, assumed to be uncooperative. Furthermore, Sharma and D'Amico [16] proposed a reduced-dynamics pose estimation in which a MEKF is formulated, validated and stress-tested with the PRISMA dataset. The measurement model was computed from pseudomeasurements, derived from the line segments detected from the image by the IP, by expressing each line segment as a function of the ROE and of the relative attitude quaternion. However, in both implementations the filter dynamics were highly simplified and no perturbation models were included. Moreover, the initial conditions for the relative state in Ref. [16] were assumed from the separate results of the pose initialization subsystem, without modeling the interface between the initial pose estimation and the filter itself, and no SIL/HIL tests were conducted. Gasbarri et al. [67] performed a Hardware-In-the-Loop (HIL) experiment in a closed GNC loop using the camera as a standalone sensor. However, no perturbation models were included in the filter dynamics and only a simplified linear KF was implemented. Galante et al. [22] proposed the fusion of several measurements from different types of monocular sensors and a LIDAR in a MEKF. Their navigation filter was designed assuming that no information about the servicer absolute position and velocity is available. As such, they neglected orbital dynamics in the filter propagation step, and considered a Schmidt KF [68] to counteract the limited system observability, which results from the lack of sufficient richness in the relative motion dynamics. Furthermore, the filter state was augmented with sensor biases to account for the different optical spectra of the pose measurement sensors. Filipe et al. [69] validated experimentally a Dual Quaternion MEKF (DQ-MEKF) [63] suitable for uncooperative satellite proximity operation scenarios, in which the pose measurements are rearranged in a dual quaternion form and fed into the navigation filter. Their filter proved to be fast enough for operational use and insensitive to singularity problems, due to its error formulation. However, only limited scenarios were simulated in the tests. Colmenarejo et al. [70] performed a comprehensive ground testing to investigate system, as well as

subsystems, level considerations related to several ADR scenarios. A complete GNC model designed in a FES was Software-In-the-Loop (SIL)/HIL-tested, thus accounting for the interfaces between the navigation filter, the IP and the initial pose estimator. Results validated several aspects of the filter robustness, such as information about the illumination quality and sensitivity to blackouts. However, several challenges behind fusing different absolute and relative sensors in the navigation filter were not solved, and the robustness of the navigation filter was not fully investigated. Furthermore, the testing did not account for recent IP methods, and the robustness of the filter with respect to a tumbling scenario was not assessed. Cavenago et al. [71] proposed two innovative nonlinear filters based on Differential Algebra (DA) to limit the computational time while preserving the filter performance. Their design included relative rotational dynamics which account for the apparent torques, the servicer-inertial torques and the target inertia matrix, thus improving other models which assumed simplified, unperturbed relative rotational motion. However, only a simplified software was used for the validation of the navigation system. In a recent effort, Pesce et al. [72] decoupled the translational and rotational motion, and compared nonlinear filtering techniques to a MEKF for the relative attitude estimation of an uncooperative target. Nonlinear filtering algorithms such as the Minimum Energy Filter, the Attitude Observer [73,74], and the 2nd Order Minimum Energy Filter [75] were adapted for the specific application. Compared to the analysis conducted in Ref. [71], the filters performance was assessed by considering limited knowledge on the target inertia matrix by neglecting the relative dynamics in their formulation. Their results showed that, despite a quicker convergence in transient, the MEKF has a lower performance at steady-state when compared to the nonlinear filters. Furthermore, the second-order minimum energy filter without dynamics was proposed as the best option in scenarios where neither the angular velocity nor the inertia matrix of the target are fully known. Furthermore, Pesce et al. [19] proposed a novel navigation system in which a H_∞ Filter [76] was selected for the translational motion estimation and the 2nd Order Minimum Energy Filter for the rotation motion estimation, respectively. The translational filter implemented the Yamanaka-Ankersen [77] formulation of satellite relative motion, and it was chosen based on the claim that assumptions of KF are usually not satisfied when dealing with optical systems, and on the fact that the absolute position of the servicer, together with the illumination conditions, can strongly affect the process and measurement noise if a KF is selected. Their design returned a navigation system for which filter robustness is preferred rather than filter optimality. On the other hand, the selected rotation filter was characterized by a null derivative of the angular acceleration, in order to avoid the dependence of the filter accuracy on the knowledge of the inertia matrix of the target spacecraft. Despite the worse performance compared to filters that include the relative dynamics, and thus the inertia matrix of the target, the proposed formulation could be extended for the pose estimation of partially known targets. Results obtained by considering Low Earth Orbit (LEO), Highly Elliptical Orbit (HEO) and GEO scenarios showed a steady state relative position and attitude Root-Mean-Square Error (RMSE) lower than 3 cm (except for HEO) and 1°, respectively. Notice also that no perturbation models were included in both filters.

An important aspect of the relative navigation filter reviewed so far relates to whether the absolute state of the servicer spacecraft is required to estimate the relative state between the servicer and the target. Except for the design in Ref. [22], the reviewed filter designs assumed that the absolute state of the servicer is known, which implies that absolute sensors such as GPS and/or Inertia Measurement Units (IMU) shall be included in the absolute filter. However, GPS can increase complexity to the system and it is not being considered in some of the current designs for close-proximity rendezvous missions. On the other hand, the limited accuracy of the absolute position and velocity information from an IMU onboard the servicer would probably result in a decreased accuracy in the estimated relative state, when compared to the estimation accuracy results obtained by assuming no noise in the absolute position and velocity. It can be stated that the interface between the relative and absolute navigation filters onboard the servicer spacecraft still presents open issues. Future research should investigate more filter designs which do not rely on the servicer absolute position (and velocity) by solving the challenges of a simplified orbital dynamics model. At the same time, the impact of the measurement noise of the servicer position on the relative navigation filter should be assessed for those designs which include the servicer absolute position (and velocity).

4.2. Design and validation of monocular navigation systems: partially known targets

During close-range rendezvous, the relative attitude dynamics is strongly dependent on the target's moment of inertia, which might be partially unknown for inactive satellites. At the same time, the knowledge of the location of the center of mass is critical for a safe approach to the target. As such, it is important to include the estimation of these parameters in the navigation filter, in order to improve the knowledge of the target state as well as of its orbit relative to the servicer.

The position and velocity of the center of mass can be estimated by solving a least squares problem in which the position and velocity of the geometrical center, or of a feature point, on the target body are measured by a monocular camera [78,79]. Alternatively, Al-Isawi and Sasiadek [80] calculated the location of the center of mass using kinematic equations and an Iterative Closest Point (ICP) algorithm, and Meng et al. [81] implemented an EKF and additionally estimated the target body mass by applying an impulse to the target.

Several approaches exist in literature to estimate the target moment of inertia with stereo cameras or other active sensors such as LIDARs. The interested readers are referred to the survey in Ref. [10] for a comprehensive overview. However, there are more restrictions on system observability when monocular cameras are adopted. Sheinfeld and Rock [79] presented a framework for rigid body inertia estimation for torque-free and non torque-free motion applicable to monocular vision. Following these findings, Benninghoff and Boge [78] and Qiu et al. [82] proposed two methods based on kinematic equations and the conservation of angular momentum, in combination with a constrained least squares method, to ensure positive diagonal values of the inertia matrix. Additionally, Hou et al. [83] proposed a dual vector quaternions-based EKF and a dual vector quaternions-based adaptive fading factors EKF to estimate the ratios of the inertia parameters of a free-floating tumbling space target. In all these methods, only normalized moments of inertia were estimated, since no external torques were applied on the target spacecraft. Setterfield et al. [84] proposed a method to additionally estimate the three principal axes together with the inertia ratios through the analysis of the target object's polhode in an arbitrary target-fixed geometric frame. Felicetti et al. [85] analyzed the estimation of the full inertia matrix by exerting a control torque on the object and by adopting an EKF. However, their method is applicable only to estimate the moment inertia of the multibody system once the chasing and the grasping phases have occurred. Xu and Wang [86] investigated the possibility to estimate the target inertia by using the information of the mass and velocity of a bullet shot to the target to change its angular momentum. Recently, Meng et al. [81] proposed a different method based on the application of a number of impulses to the target in order to observe the resulting motion changes and solve for all the inertia parameters. An EKF was used to estimate the normalized inertia matrix together with the target mass, and a least squares method was added to estimate the full set of inertial parameters.

5. Conclusions and recommendations

This paper presented a detailed review of the robustness and applicability of state-of-the-art monocular pose estimation systems for the

relative navigation with an uncooperative spacecraft. The research is motivated by the applicability of relative pose estimation in future space missions, i.e. ADR and IOS, which involve close-proximity operations of a servicer spacecraft around a target. Monocular systems were reviewed due to the strict power, mass, and operational range requirements driving the current design of these missions, which are usually killer requirements for active, as well as stereo, systems.

First, a review of monocular EO systems is given in which VIS, TIR and NIR camera suites are traded-off against image quality and robustness with respect to the space environment. Due to the limited robustness of VIS/NIR cameras against harsh illumination conditions and the presence of the Sun or the Earth in the background, and the limited image quality which characterises TIR cameras, multispectral systems are identified as a promising solution capable of increasing the overall system robustness, while at the same time preserving system accuracy. Furthermore, the applicability of the state-of-the-art camera suites to operational ranges from several hundreds of meters down to docking is analyzed in order to assess whether only a single monocular camera could be used during close-proximity operations. In principle, collaborative cameras could be avoided by switching to feature tracking as soon as the target is not fully in the camera FoV.

Monocular pose estimation is analyzed by firstly focusing on the IP algorithms adopted prior to the actual estimation. Three main feature synthesis schemes are identified which are able to combine the advantages of several feature detectors into a more robust system. Furthermore, it is foreseen that the combination of keypoint detectors with edge and corner detectors will represent an additional step forward in the design of robust and reliable IP systems, provided that keypoint detectors are validated against scenarios in which the Earth is in the image background. Besides, it is expected that feature detection methods based on neural networks will improve the system robustness against adverse illumination conditions as well as partial occultation of the target.

The different techniques adopted for the pose initialization and tracking are then reviewed. A comparative assessment of several PnP solver is presented, from which it is concluded that a combination of different solvers should, in principle, improve the pose estimation accuracy. Furthermore, the challenges involved in VIS-based and TIR-based estimation systems are listed in terms of the image database adopted for the validation, the robustness against image background and spacecraft symmetry, and the associated IP system adopted. The comparison suggests to investigate a pose estimation system in which model-based and non-model based methods are used to estimate the pose from VIS and TIR images, respectively. This follows from the challenges in relying on an off-line TIR model of the target spacecraft, due to fast variations in the space thermal environment.

A review of recent pose estimation systems based on CNNs is provided in order to investigate the level of accuracy that could be achieved by exploiting them during pose initialization. Three novel methods are reviewed which adopt transfer learning of pre-trained networks to solve for the relative pose. In particular, the comparative assessment showed that a relatively small training database could be used without affecting the network performance, provided that suitable network layers are selected. Additionally, it is suggested that the coarse accuracy, which characterises the networks reviewed in this paper, could be compensated by including a PnP solver which uses the CNN solution as initial guess for the relative pose. Still, the drop in performance when the target spacecraft is not fully in the camera FoV, together with the amount of realistic images to use during training and/or testing, represent unanswered questions which will require further analyses.

Finally, visual-based navigation filters are reviewed by assessing their applicability to scenarios in which the target spacecraft is fully or partially known. The comparison between different filters shows that filter selection for the pose estimation of an uncooperative target is, from a high-level perspective, driven by a trade-off between filter robustness and filter optimality. In particular, when the target is partially known, the dependence of the filter on the target inertia matrix could be tackled by simplifying the filter internal dynamics, or by estimating the target mass and inertia in-flight. Furthermore, the qualitative comparison suggests that the impact of the absolute filter's solution on the relative pose estimation should be accounted for in the design of the navigation filter.

Acknowledgements

This study has been funded and supported by the European Space Agency and Airbus Defence and Space under Network Partnering Initiative (NPI) program with grant number NPI 577 - 2017.

References

[1] S. D'Amico, et al., Prisma, Distributed Space Missions for Earth System Monitoring, 2013, pp. 599–637, , https://doi.org/10.1007/978-1-4614-4541-8.
[2] L. Tarabini-Castellani, J. Salvador Llorente, J. Fernandez Ibarz, M. Ruiz, Proba 3 mission, Int. J. Space Eng. 1 (No.4) (2013) 349–366.
[3] A. Tatsch, N. Fitz-Coy, S. Gladun, On-orbit Servicing: a brief survey, Proceedings of the 2006 Performance Metrics for Intelligent Systems Workshop, 2006, pp. 21–23.
[4] M. Wieser, H. Richard, G. Hausmann, J.-C. Meyer, S. Jaekel, M. Lavagna, R. Biesbroek, e. deorbit mission, OHB debris removal concepts, ASTRA 2015-13th Symposium on Advanced Space Technologies in Robotics and Automation, Noordwijk, The Netherlands, 2015.
[5] J. Davis, H. Pernicka, Proximity operations about and identification of non-cooperative resident space objects using stereo imaging, Acta Astronaut. 155 (2019) 418–425.
[6] V. Pesce, M. Lavagna, R. Bevilacqua, Stereovision-based pose and inertia estimation of unknown and uncooperative space objects, Adv. Space Res. 59 (2017) 236–251.
[7] R. Opromolla, G. Fasano, G. Rufino, M. Grassi, Uncooperative pose estimation with a lidar-based system, Acta Astronaut. 110 (2015) 287–297.
[8] S. Segal, P. Gurfil, K. Shahid, In-orbit tracking of resident space objects: a comparison of monocular and stereoscopic vision, IEEE Trans. Aerosp. Electron. Syst. 50 (No.1) (2014) 676–688.
[9] S. Sharma, J. Ventura, S. D'Amico, Robust model-based monocular pose initialization for noncooperative spacecraft rendezvous, J. Spacecr. Rocket. 55 (No.6) (2018) 1–16.
[10] R. Opromolla, G. Fasano, G. Rufino, M. Grassi, A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations, Prog. Aero. Sci. 93 (2017) 53–72.
[11] S. Sharma, S. D'Amico, Comparative assessment of techniques for initial pose estimation using monocular vision, Acta Astronaut. 123 (2015) 435–445.
[12] B. Naasz, R. Burns, S. Queen, J. Van Eepoe, J. Hannah, E. Skelton, The HST SM4 relative navigation sensor system: overview and preliminary testing results from the flight robotics lab, J. Astronaut. Sci. 57 (Nos. 1 & 2) (2009) 457–483.
[13] X. Du, B. Liang, W. Xu, Y. Qiu, Pose measurement of large non-cooperative satellite based on collaborative cameras, Acta Astronaut. 68 (Nos.11 & 12) (2011) 2047–2065.
[14] C. Liu, W. Hu, Relative pose estimation for cylinder-shaped spacecrafts using single image, IEEE Trans. Aerosp. Electron. Syst. 50 (No. 4) (2014) 3036–3056.
[15] S. D'Amico, M. Benn, J. Jorgensen, Pose estimation of an uncooperative spacecraft from actual space imagery, Int. J. Space Eng. 2 (No.2) (2014) 171–189.
[16] S. Sharma, S. D'Amico, Reduced-dynamics pose estimation for non-cooperative spacecraft rendezvous using monocular vision, Adv. Astronaut. Sci. Guid. Navig. control 159 (2017).
[17] S. Sharma, C. Beierle, S. D'Amico, Pose estimation for non-cooperative spacecraft rendezvous using convolutional neural networks, IEEE Aerospace Conference, Big Sky, MT, USA, 2018, https://doi.org/10.1109/AERO.2018.8396425.
[18] F. Schnitzer, A. Sonnenburg, K. Janschek, M. Sanchez Gestido, Lessons-learned from on-ground testing of image-based non-cooperative rendezvous navigation with visible-spectrum and thermal infrared cameras, 10th International ESA Conference on Guidance, Navigation, and Control Systems, Salzburg, Austria, 2017.
[19] V. Pesce, R. Opromolla, S. Sarno, M. Lavagna, M. Grassi, Autonomous relative navigation around uncooperative spacecraft based on a single camera, Aero. Sci. Technol. 84 (2019) 1070–1080.
[20] L. Kozlowsky, W. Kosonocky, 3 ed., Handbook of Optics vol. 2, McGraw-Hill, 1995.
[21] B. Cavrois, A. Vergnol, A. Donnard, P. Casiez, U. Southivong, O. Mongrard, F. Ankersen, C. Pezant, P. Breteker, F. Kolb, M. Windmüller, LIRIS demonstrator on ATV5: a step beyond for european non cooperative navigation system, AIAA Guidance, Navigation and Control Conference, 2015.
[22] J. Galante, J. Van Eepoel, C. D' Souza, B. Patrick, Fast kalman filtering for relative spacecraft position and attitude estimation for the raven ISS hosted Payload, 39th AAS Guidance and Control Conference, Breckenridge, CO, USA, 2016.
[23] O. Yilmaz, N. Aouf, L. Majewski, M. Sanchez Gestido, G. Ortega, Using infrared based relative navigation for active debris removal, 10th International ESA Conference on Guidance, Navigation, and Control Systems, Salzburg, Austria, 2017.
[24] M. Gansmann, O. Mongrard, F. Ankersen, 3D Model-Based Relative Pose Estimation for Rendezvous and Docking Using Edge Features, Salzburg, Austria, 2017.
[25] J. Shi, S. Ulrich, S. Ruel, M. Anctil, Uncooperative spacecraft pose estimation using

an infrared camera during proximity operations, AIAA SPACE 2015 Conference and Exposition, Pasadena, CA, USA, 2015, https://doi.org/10.2514/6.2015-4429.
[26] J. Shi, S. Ulrich, S. Ruel, Spacecraft pose estimation using a monocular camera, 67th International Astronautical Congress, Guadalajara, Mexico, 2016.
[27] J. Shi, S. Ulrich, S. Ruel, Spacecraft pose estimation using principal component analysis and a monocular camera, AIAA Guidance, Navigation, and Control Conference, Grapevine, TX, USA, 2017, https://doi.org/10.2514/6.2017-1034.
[28] J. Deloo, E. Mooij, Active debris removal : aspects of trajectories, communication and illumination during final approach, Acta Astronaut. 117 (2015) 277–295.
[29] CleanSpace, e.deorbit Implementation Plan, Technical Report, European Space Agency, 2015.
[30] O. Yilmaz, N. Aouf, E. Checa, L. Majewski, M. Sanchez Gestido, Thermal analysis of space debris for infrared based active debris removal, Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, SAGE Publications, 2017, pp. 1–13, , https://doi.org/10.1177/0954410017740917.
[31] R. Duda, P. Hart, Use of the hough transformation to detect lines and curves in pictures, Commun. ACM 15 (No. 1) (1972) 11–15.
[32] D. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (No. 2) (2004) 91–110.
[33] M. Calonder, V. Lepetit, C. Strecha, P. Fua, BRIEF: binary robust independent elementary features, European Conference on Computer Vision, 2010, pp. 778–792, , https://doi.org/10.1007/978-3-642-15561-1_56.
[34] M.A. Fischer, R. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (No. 6) (1981) 381–395.
[35] O. Yilmaz, N. Aouf, L. Majewski, M. Sanchez Gestido, Evaluation of Feature Detectors for Infrared Imaging in View of Active Debris Removal, (2017) Darmstadt, Germany.
[36] D. Rondao, N. Aouf, O. Dubois-Matra, Multispectral image processing for navigation using low performance computing, 69th International Astronautical Congress, 2018 Bremen, Germany.
[37] D. Rondao, N. Aouf, Multi-view monocular pose estimation for spacecraft relative navigation, 2018 AIAA Guidance, Navigation, and Control Conference, Kissimmee, FL, USA, 2018, https://doi.org/10.2514/6.2018-2100.
[38] V. Capuano, S. Alimo, A. Ho, S.-J. Chung, Robust features extraction for on-board monocular-based spacecraft pose acquisition, AIAA Scitech 2019 Forum, San Diego, CA, USA, 2019, https://doi.org/10.2514/6.2019-2005.
[39] L. Pasqualetto, R. Fonod, E. Gill, I. Ahrns, J. Gil Fernandez, Comparative assessment of image processing algorithms for the pose estimation of an uncooperative spacecraft, International Workshop on Satellite Constellations & Formation Flying, Glasgow, UK, 2019.
[40] G. Pavlakos, X. Zhou, A. Chan, K. Derpanis, K. Daniilidis, 6-DoF object pose from semantic keypoints, IEEE International Conference on Robotics and Automation, ICRA, 2017.
[41] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (SURF), Comput. Vis. Image Understand. 110 (No. 3) (2008) 346–359.
[42] V. Lepetit, P. Fua, Monocular model-based 3D tracking of rigid objects: a survey, Found. Trends® Comput. Graph. Vis. 1 (No.1) (2005) 1–89.
[43] D. Dementhon, L. Davis, Model-based object pose in 25 lines of code, Int. J. Comput. Vis. 15 (No.1–2) (1995) 123–141.
[44] D. Oberkampf, D. Dementhon, L. Davis, Iterative pose estimation using coplanar feature points, Comput. Vis. Image Understand. 63 (No.3) (1996) 495–511.
[45] P. David, D. DeMenthon, R. Duraiswami, H. Samet, SoftPOSIT: simultaneous pose and correspondence determination, Int. J. Comput. Vis. 59 (No.3) (2004) 259–284.
[46] V. Lepetit, F. Moreno-Noguer, P. Fua, EPnP: an accurate O(n) solution to the PnP problem, Int. J. Comput. Vis. 81 (2009) 155–166.
[47] T. Drummond, R. Cipolla, Real-time visual tracking of complex structures, IEEE Trans. Pattern Anal. Mach. Intell. 24 (No.7) (2002) 932–946.
[48] S. Hannah, ULTOR Passive Pose and Position Engine for Spacecraft Relative Navigation, (2008), https://doi.org/10.1117/12.777193.
[49] R. Opromolla, G. Fasano, G. Rufino, M. Grassi, Pose estimation for spacecraft relative navigation using model-based algorithms, IEEE Trans. Aerosp. Electron. Syst. 53 (2017) 431–447.
[50] J. Ball, D. Anderson, C. Chan, Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community, J. Appl. Remote Sens. 11 (No.4) (2017) 568–584.
[51] A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks, 26th Annual Conference on Neural Information Processing Systems, vol. 1, 2012, pp. 1097–1105 Lake Tahoe, NV, USA.
[52] J. Shi, S. Ulrich, S. Ruel, CubeSat simulation and detection using monocular camera images and convolutional neural networks, 2018 AIAA GUidance, Navigation, and Control Conference, Kissimmee, FL, USA, 2018, https://doi.org/10.2514/6.2018-1604.
[53] C. Szegedy, S. Ioffe, V. Vanhoucke, Inception-v4, inception-ResNet and the impact of residual connections on learning, AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 2017.
[54] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[55] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (No.6) (2017) 1137–1149.
[56] S. Sharma, S. D'Amico, Pose estimation for non-cooperative spacecraft rendezvous using neural networks, 29th AAS/AIAA Space Flight Mechanics Meeting, Ka'anapali, HI, USA, 2019.
[57] J. Sullivan, S. Grimberg, S. D'Amico, Comprehensive survey and assessment of spacecraft relative motion dynamics models, J. Guid. Control Dyn. 40 (No.8) (2017) 1837–1859.
[58] J. Hamel, J. de Lafontaine, Linearized dynamics of Formation Flying spacecraft on a J2-perturbed elliptical orbit, J. Guid. Control Dyn. 30 (No.6) (2007) 1649–1658.
[59] T. Guffanti, S. D'Amico, M. Lavagna, Long term analytical propagation of satellite relative motion in perturbed orbits, Adv. Astronaut. Sci. Spaceflight. Mech. 160 (2017) 355.
[60] A. Koenig, T. Guffanti, S. D'Amico, New state transition matrices for spacecraft relative motion in perturbed orbits, J. Guid. Control Dyn. 40 (No.7) (2017) 1749–1768.
[61] S. Kim, J. Crassidis, Y. Cheng, A. Fosbury, Kalman filtering for relative spacecraft attitude and position estimation, J. Guid. Control Dyn. 30 (No. 1) (2007) 133–143.
[62] E. Lefferts, F. Markley, M. Shuster, Kalman filtering for spacecraft attitude estimation, J. Guid. Control Dyn. 5 (No. 5) (1982) 417–429.
[63] F. Markley, Attitude error representations for kalman filtering, J. Guid. Control Dyn. 26 (No. 2) (2003) 311–317.
[64] J. Branco, V. Barrena, D. Escorial Olmos, L. Tarabini-Castellani, A. Cropp, The Formation Flying navigation system for proba 3, Annu. Rev. Earth Planet Sci. 24 (2015) 37–47.
[65] G. Allende-Alba, S. D'Amico, O. Montenbruck, Radio Frequency sensor fusion for relative navigation of formation flying satellites, Int. J. Space Eng. 3 (No.2) (2009) 129–147.
[66] L. Zhang, T. Li, H. Yang, S. Zhang, H. Cai, S. Qian, Unscented kalman filtering for relative spacecraft attitude and position estimation, J. Navig. 68 (No. 3) (2015) 528–548.
[67] P. Gasbarri, M. Sabatini, G. Palmerini, Ground tests for vision based determination and control of formation flying spacecraft trajectories, Acta Astronaut. 102 (2014) 378–391.
[68] S. Schmidt, Applications of state space methods to navigation problems, Adv. Control Syst. 3 (1966) 293–340.
[69] N. Filipe, M. Kontitsis, P. Tsiotras, Extended kalman filter for spacecraft pose estimation using dual quaternions, 2015 American Control Conference, Chicago, IL, USA, 2015, pp. 3187–3192, , https://doi.org/10.1109/ACC.2015.7171823.
[70] P. Colmenarejo, M. Graziano, G. Novelli, D. Mora, P. Serra, A. Tomassini, K. Seweryn, G. Prisco, J. Gil, Fernandez, On ground validation of debris removal technologies, In Press, Acta Astronaut. 158 (2018) 206–219. Available online January 2018.
[71] F. Cavenago, M. Massari, S. Servadio, A. Wittig, DA-based nonlinear filters for spacecraft relative state estimation, 2018 Space Flight Mechanics Meeting, 2018.
[72] V. Pesce, M. Haydar, M. Lavagna, M. Lovera, Comparison of filtering techniques for relative attitude estimation of uncooperative space objects, Aero. Sci. Technol. 84 (2019) 318–328.
[73] R. Mortensen, Maximum-likelihood recursive nonlinear filtering, J. Optim. Theory Appl. 2 (No.6) (1968) 386–394.
[74] M. Zamani, J. Trumpf, R. Mahony, Minimum-energy filtering for attitude estimation, IEEE Trans. Autom. Control 58 (No.11) (2013) 2917–2921.
[75] M. Zamani, J. Trumpf, R. Mahony, On the Distance to Optimality of the Geometric Approximate Minimum-Energy Attitude Filter, Portland, OR, USA, (2014), pp. 4943–4948, https://doi.org/10.1109/ACC.2014.6858915.
[76] D. Simon, State Estimation: Kalman, H_∞ and Nonlinear Approaches, John Wiley Sons, 2006.
[77] K. Yamanaka, F. Ankersen, New state transition matrix for relative motion on an arbitrary elliptical orbit, J. Guid. Control Dyn. 25 (No.1) (2002) 60–66.
[78] H. Benninghoff, T. Boge, Rendezvous Involving a Non-cooperative, Tumbling Target - Estimation of Moments of Inertia and Center of Mass of an Unknown Target vol. 25, (2015) Munich, Germany.
[79] D. Sheinfeld, S. Rock, Rigid body inertia estimation with applications to the capture of a tumbling satellite, Adv. Astronaut. Sci. 134 (2009) 343–356.
[80] M. Al-Isawi, J. Sasiadek, Guidance and control of a robot capturing an uncooperative space target, J. Intell. Robot. Syst.& (2018) 1–9.
[81] Q. Meng, J. Liang, O. Ma, Estimate of all the inertial parameters of a free-floating object in orbit, 2018 AIAA Guidance, Navigation and Control Conference, Kissimmee, FL, USA, 2018, https://doi.org/10.2514/6.2018-1606.
[82] S. Qiu, Y. Guo, J. Xing, G. Ma, Inertia parameter and attitude estimation of space noncooperative tumbling target based on a two-step method, IECON 2017-43rd Annual Conference of the IEEE Industrial Electronics Society, Beijing, China, 2017, https://doi.org/10.1109/IECON.2017.8217143.
[83] X. Hou, C. Ma, Z. Wang, J. Yuan, Adaptive pose and inertial parameters estimation of free-floating tumbling space objects using dual vector quaternions, Adv. Mech. Eng. 9 (No. 10) (2017) 1–17.
[84] T. Setterfield, D. Miller, A. Saenz Otero, E. Frazzoli, J. Leonard, Inertial properties estimation of a passive on-orbit object using polhode analysis, J. Guid. Control Dyn. 41 (No. 10) (2018) 2214–2231.
[85] L. Felicetti, M. Sabatini, A. Pisculli, P. Gasbarri, G. Palmerini, Adaptive thrust vector control during on-orbit servicing, AIAA SPACE 2014 Conference and Exposition, 2014.
[86] B. Xu, S. Wang, Vision-based moment of inertia estimation of non-cooperative space object, 10th International Symposium on Computational Intelligence and Design, Hangzhou, China, 2017, https://doi.org/10.1109/ISCID.2017.68.