DOI: 10.13009/EUCASS2019-961

Multispectral Sensing for Relative Navigation

A. Fiengo[1], J. I. Bravo[1], P. Rosa[2], N. Guercio[3], J. Christy[3], V. Dubanchet[3]
[1]DEIMOS Space S.L.U., Tres Cantos – Madrid, Spain
[2]DEIMOS Engenharia, Lisbon, Portugal
[3]Thales Alenia Space, France

Abstract
A relative navigation system based on multispectral sensing device, for non-cooperative rendezvous between a chaser and a target, as well as for the descent and landing on an asteroid, is presented. Firstly, the advantages of using a multispectral camera capable of detecting different spectral signatures are discussed. Subsequently, the developed pose estimation algorithm is presented, giving details on the strategies for fusing information from the different bands and presenting our codebook-based attitude retrieval mechanism. Finally, the adopted navigation filter, based on Extended Kalman Filtering (EKF), is described. Simulations are conducted to demonstrate the advantages and limitations of the proposed solution.

1. Introduction

Autonomous relative navigation between a chaser and a target is clearly becoming a mandatory function in many of the future space missions, and the associated challenges have been already the subject of many studies in the past. A key element of a successful navigation is obviously the navigation sensor: the optimal selection of the sensor is hindered by the great variety of and uncertainties associated to the targets, as well as the large variability of distances from these targets to the camera, therefore typically leading to locally optimal solutions, tailored for the scenarios at hand. Many developments have been performed and are still under way for navigation sensors working in the visible wavelengths, mainly because this uses a very mature technology, it is relatively low cost, and the database of targets in these wavelengths is quite complete. For example, for one of the next ESA interplanetary mission, JUICE, the navigation camera uses the well-known star tracker technology. Others technologies are under study, and for example TAS-F is leading a contract with ESA for a Time-Of-Flight camera for typically RdV type of missions, but still in visible wavelengths.

Multispectral and Hyperspectral imaging has been used in space but at instrument level, and this always has resulted in quite complex instruments, because of the number of detectors needed, the required accurate temperature control, etc. It is clear, however, that using a larger part of the spectrum is a way to bring much more inputs to the navigation function and therefore make the navigation solution more robust and efficient. The challenge is thus to keep the resulting sensor at a reasonable level of complexity and cost. Relevant progress has been achieved in the last decade, as multispectral technologies are now widely used within different realms, such as drones for fields and forests observation. The progress made in this area has allowed a miniaturization of the sensing devices and it is therefore expected that it will also impact on space applications.

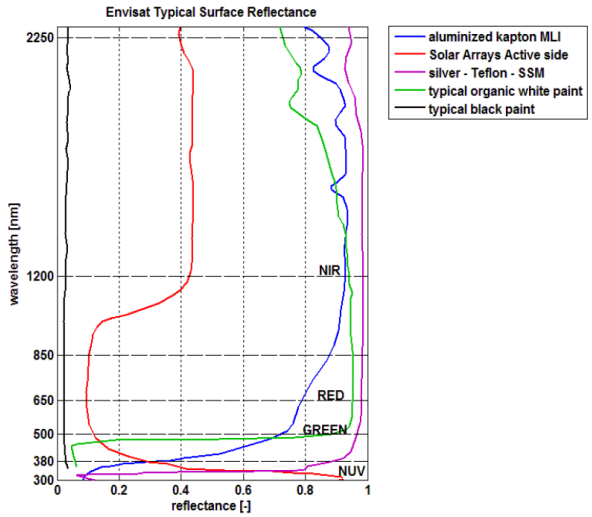
This paper provides an overview of the methodology adopted within the ESA-funded Multispectral Sensing for Relative Navigation (MSRNAV) project to design and assess such navigation systems, in the case where the main sensor providing relative measurements between the chaser and a known target is a multispectral camera. Within the scope of the project, two operational scenarios have been identified: the first scenario is a non-cooperative rendezvous with a dead satellite (Envisat) in which the chaser performs a rendezvous with the target object, evaluates its attitude dynamics and CoM position and, if required, performs a forced translation in order to reduce the relative motion between chaser and target to levels adequate to initiate the capture; the second scenario comprises the descent and landing on one of the components of the binary asteroid Didymos with the objective of estimating the asteroid's relative angular velocity, assuming that the spacecraft is static and, hence, all the motion is due to the rotation motion of asteroid. The main focus of this paper is the design and validation of the image processing and navigation filter developed in the scope of this

DOI: 10.13009/EUCASS2019-961

study; due to space limitations, focus is given to the autonomous rendezvous scenario, while only brief results are shown for the asteroid scenario. MSRNAV weights, in particular, the benefits of a multispectral sensing device with regards to additional complexity and costs, and when compared to the usual visual-spectrum sensing.

2. Multispectral Sensing for Relative Navigation

Relative navigation between a chaser and a target is a key S/C capability for future Space missions, due to the relevance of close-proximity operations within the realm of Active Debris Removal (ADR) or Space Servicing Vehicles (SSVs), to cite a few examples. The target for the scenario is a large, dead satellite that is tumbling; for the purpose of the study, ESA's civilian Earth observation satellite Envisat has been selected. A radiometric analysis was carried out in order to identify the spectral bands of interest for the study. The spectral bands identified are shown in the following Table. The values have been chosen to observe a maximum of spectral features while minimizing the number of spectral bands, to avoid introducing complexity to the instrument. The proposed bands can be seen as maximum bandwidths, and they could be reduced for algorithm and instrument specific constraints.



ID	Colour	Wavelength (nm)
1	NUV	300-380
2	GREEN	500-650
3	VIS	650-850
4	NIR	1200-2250
ID	**Band**	**Temperature (K)**
5	TIR	200-400

Possible spectral bands of interest

The following section provides an overview of the methodology adopted within MSRNAV to design and assess such navigation systems, in the case the main sensor providing relative measurements between the chaser and the target is a multispectral camera with the proposed spectral bands. The device to be designed shall be generic enough for the two scenarios addressed within the project (although only one of the scenarios is considered in this paper), in order to be regarded as a potential GNC sensor.

2.1 High-level architecture

The high-level simulation architecture adopted within MSRNAV is illustrated in Figure 1. The simulator encompasses all the chaser and target dynamics, guidance algorithms, and sensor information generation. The latter data is used by the Image Processing (IP) algorithms and the navigation filter (where applicable) to generate relative state estimates, which are subsequently used to assess the performance of the overall estimation approach.

DOI: 10.13009/EUCASS2019-961

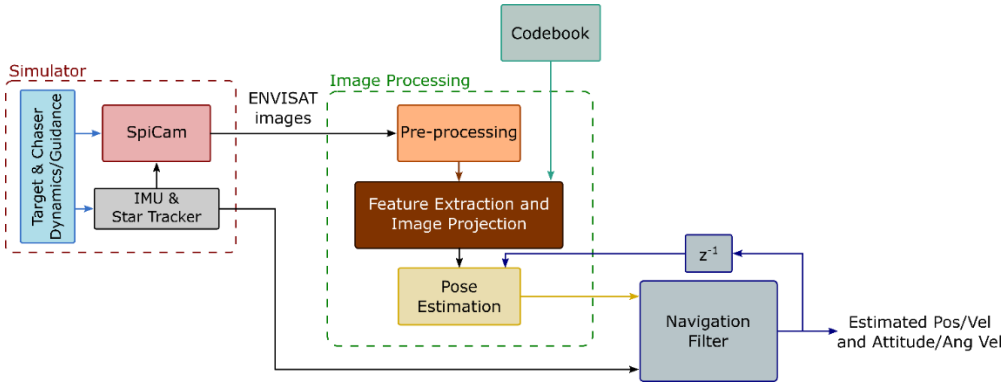


Figure 1: High-level simulation architecture

2.2 Navigation Solution

The architecture of the navigation filter, which is based on Extended Kalman Filtering (EKF), was developed according to four main design concerns:

- **Accurate modelling** of the state dynamics – the accurate modelling of the chaser's relative motion and attitude kinematics is critical for state propagation. Accurate state prediction is necessary to keep the KF linearization point valid, and to increase the system's robustness to errors in aiding sensor readings. State prediction should also provide reliable estimates of the state at the whenever the G&C commands are applied.
- **Integration of sensor information** – the strategies for integrating the measurements from the available sensors (IMU, star tracker, and camera) must be carefully considered. The nonlinear model of the sensor reading must be correctly described as a function of the estimated state, and relevant sensor non-idealities, such as noise and bias, should be compensated for in the navigation system.
- **Numerical stability and computational optimization** – besides the analytic performance of the Kalman filter, the implementation of the navigation filter should account for computational issues. Namely, numerical methods should be adopted guaranteeing that the filter variables remain numerically-well defined, while the most appropriate state update techniques should be analysed.
- **Algorithmic updates motivated by sensor limitations** – the navigation algorithm should be adapted to account for the image processing delay and of the GNC algorithm itself. The GNC algorithm should be designed such that G&C commands are applied at the proper time instant, irrespectively of the GNC processing delay. Also, a filtering strategy to integrate delayed measurement integration must be implemented, due to the delay between image capture and image processing

2.3 Image Processing Solution

The IP solution is designed to provide the attitude and position (pose) of a non-cooperative target with a monocular multispectral camera. The image processing pose estimation algorithm has been designed as a retrieval problem: the input image is compared against a codebook with image representations (aggregate of visual features) associated to poses; the pose is retrieved from the entry of the codebook with the minimum distance match to the aggregate of visual features extracted from the input image. A data dimensionality reduction technique (2D-PCA with L1 norm and sparse regularization) is applied in this case to reduce the size of the descriptors, so the matching process against the codebook can be performed with a reasonable amount of computational time. To exploit the information provided by the different bands of the multispectral camera, a codebook is built for each band, and the fusion of measurements is based on majority voting (i.e. the mode of the retrieved poses)

DOI: 10.13009/EUCASS2019-961

3. Simulation Architecture Overview

Based on the design considerations presented in Section 2, the navigation architecture depicted in Figure 2 : Scenario description was adopted.

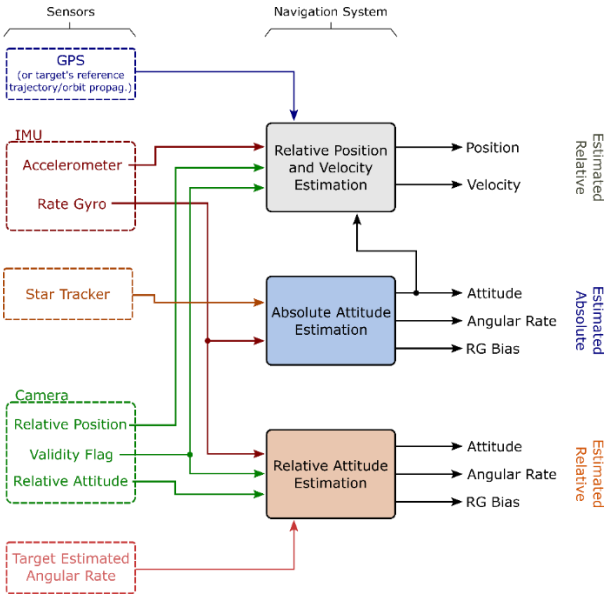


Figure 2: Overview of the navigation architecture

The main rationale for the selection of this loosely coupled architecture is the modularity of the approach. For the integration of the camera measurements, a tightly coupled design approach would depend on the IP itself. For example, if a feature detection and tracking were adopted, the EKF could directly take into account the position of the feature in the camera frame. However, this would mean that a change in the IP would lead to a change in the navigation filter. Moreover, for the selected IP approach within the ADR scenario, the natural output is the pose, thus requiring no transformations.

3.1 Relative Position and Velocity Estimation

This filter estimates the relative position and velocity of the chaser in Local-Vertical/Local-Horizontal (LVLH) coordinates (P_{LVLH} and v_{LVLH} respectively). This estimation is performed by using a relative sensor providing relative pose measurements, an inertial measurement unit, and an estimate of the absolute position (for orbital elements determination).
The position and velocity are propagated using the so-called Yamanaka-Ankersen state transition matrix proposed in [5], augmented with a forced term that accounts for the thrust commanded by the control system. The bias of the inertial sensor is modeled as a slowly time-varying process, i.e. a random walk signal driven by white noise with very small covariance.
For the measurements integration, the accelerometer model considered in the filter is given by:

$$a_r = {}^{RSF}a + b_a + n_a$$

where n_a is the zero-mean measurement noise. Finally, the camera measurement is introduced in the filter by transforming the measured relative position from the camera reference frame to the LVLH reference frame.

3.2 Absolute Attitude Estimation

The absolute attitude (and angular rate) estimation is performed, within the scope of MSRNAV, by exploiting the star tracker measurements, as well as the angular rate readings provided by the rate gyros, within the IMU. An EKF was adopted, with the dynamics used for the propagation steps being:

- Rate gyros measurements, used for the propagation of the estimated state, by the EKF
- Star tracker measurements, accounted for by the EKF during the update cycle

DOI: 10.13009/EUCASS2019-961

3.3 Relative Attitude Estimation

In terms of relative attitude estimation, an approach similar to the one described in the previous subsection (i.e. an EKF-based solution) was adopted. The key difference, however, is that the sensor providing attitude measurements is a camera instead of the star tracker. The IMU is used for the same reasons as before.

4. Image Processing Architecture

4.1 Background

This section is devoted to the introduction of the concepts and theoretical background specific to the image processing algorithm, based on the approach from [6]. Firstly, the Histogram of Oriented Gradients (HOG) image descriptor is introduced,, followed by a dimensionality reduction technique based on two-dimensional Principal Component Analysis, the subsection ends with the overall IP algorithm definition.

4.1.1 Histogram of Oriented Gradients

Histogram of Oriented Gradients is an image descriptor proposed by Dalal and Triggs in [7]. The descriptor encodes the object appearance by describing the orientation of its edges (gradients) along the image region.
To achieve so, first the image is divided into small spatial regions known as cells. In each cell, the orientations of the edges are computed and used to build a one-dimensional histogram that encodes the cell orientations. Then, to achieve better invariance to illumination changes, the histogram cells are normalized by the energy of the neighbourhood (block). The set of normalized blocks conforms the HOG descriptor.

4.1.2 Sparse 2DPCA-L1

Principal Component Analysis [8] is a statistical procedure used to reduce the data dimensionality while preserving the variance. The process is designed to extract the directions (or vectors) of the most variability of the data, with the constraint that each direction, known as principal component, must represent the highest variance being orthogonal to the preceding. The data dimensionality is reduced by projecting the data onto the subspace spanned by the principal components.
Two dimensional PCA (2DPCA) [9] is an extension of PCA introduced by Yang et al. in the scope of face reconstruction and recognition. The idea behind this method is the same as standard PCA: find the optimal projection that maximizes the dispersion of the data, but in 2DPCA the data is arranged in matrices instead of in column vectors (e.g. images).
The relation between PCA and 2DPCA is that 2DPCA is essentially PCA performed on the rows of the data matrices, understanding each row as a data unit. In other words, 2DPCA of a data matrix can be seen as PCA on the set of rows of the matrix, where the covariance matrix is evaluated using the rows of all the centered training samples [10], [11]. The introduction of L1-Norm-based 2DPCA [12] aimed to obtain basis vectors that maximize the data variance including the L1 norm to provide robustness against outliers. The square nature of L2 norm implicit in standard PCA entails that large differences between the data and its projection are more penalized compared to the L1 norm. Hence, the L1 norm increases the method's robustness to outliers.
The sparsity (measured by the ratio of input variables equal to zero) is included in [13] by means of elastic-net regularization [14]. Compared to standard PCA methods, where each principal component is a linear combination of all the input data features [14], the use of sparse methods encourages the reduction of the number of features in the generation of the principal components, selecting a few salient features among all the input data, and allowing a more efficient encoding of the information.
The description of the process used to obtain the set of k principal components that maximize the variance for 2DPCA with sparse regularization is out of the scope of this paper – the interested reader is referred to [13].

4.1.3 Attitude Sampling

The representativeness of a dataset increases with the number of rendered views, but also with the diversity of these views. For a given number of rendered attitudes N, the dataset will be more representative if the attitudes are not clustered but evenly distributed.

DOI: 10.13009/EUCASS2019-961

The uniform sampling of attitudes, expressed in Euler angles or quaternions, does not lead to a uniform sampling of the possible views of the 3D object. Differently, uniform sampling generates distributions heavily biased towards regions of the sphere according to the set of rotation axes [15], [16]. However, if the set of Euler angles or quaternions representing the set of possible attitudes is sampled with a random uniform distribution, the resulting subset is uniform.
An algorithm for generating uniformly distributed random quaternions is presented in [15]. Assuming rand() is a function that generates pseudo random numbers from a uniform distribution between [0,1]. The set of quaternions $Q = (w, x, y, z) = (\cos\left(\frac{\theta}{2}\right), vx \sin\left(\frac{\theta}{2}\right), vy \sin\left(\frac{\theta}{2}\right), vz \cos\left(\frac{\theta}{2}\right))$ described by a rotation angle θ and a three-dimensional vector (vx,vy,vz), can be uniformly sampled with the following procedure:

1. Define a number of random orientations N.
2. For each orientation
   a. Get a sample from the uniform random distribution $rand()$
   b. Define $\sigma_1 = \sqrt{1 - s}$ and $\sigma_2 = \sqrt{s}$
   c. Randomly generate two rotations by sampling the random distribution again:

$$\theta_1 = 2\pi \cdot rand()$$
$$\theta_2 = 2\pi \cdot rand()$$

   d. Compute the quaternion values as

$$w = \cos(\theta_2) \cdot \sigma_2$$
$$x = \sin(\theta_1) \cdot \sigma_1$$
$$y = \cos(\theta_1) \cdot \sigma_1$$
$$z = \sin(\theta_2) \cdot \sigma_2$$

   e. Set $Q_i = (w, x, y, z)$

4.2 Architecture

Figure 1: IP architecture shows a block diagram of the proposed IP algorithm, divided into two parts: codebook training, and classification. In the training stage, the rendered images are processed to extract a set of features describing the associated attitude and information used to retrieve the position. To obtain a computationally efficient version of the features describing the attitude, these are reduced in dimensionality by applying the sparse 2DPCA-L1 algorithm.

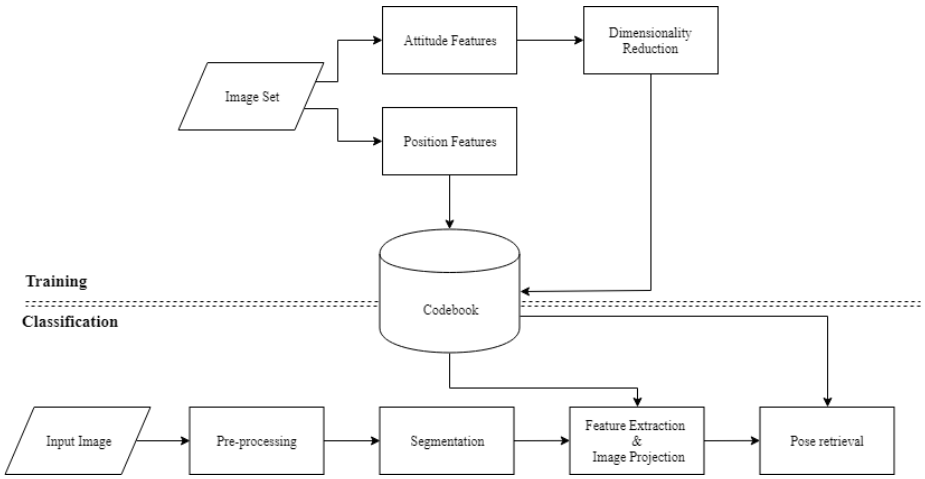


Figure 1: IP architecture

DOI: 10.13009/EUCASS2019-961

The classification part follows a similar procedure: each input image is processed to generate an image containing only the target spacecraft. Then, features are extracted from the image, which are projected onto the principal components obtained in the training stage, to retrieve the attitude by finding the nearest-neighbour in the codebook.

4.2.1 Training

This stage is designed to populate the codebook with image representations associated to attitudes and the information used to retrieve the position.

The dataset with N images corresponding to N sampled attitudes is processed to obtain descriptors related to the attitude (Attitude Features) and position (Position Features). The attitude descriptors are obtained by describing the appearance of the target spacecraft in a given attitude. To do so, the concatenation of the pixel intensity values of the grey-scale image, and Histogram of Oriented Gradients is used. The relative position of the spacecraft, including the distance, is referred to a fixed-point relative to the spacecraft body. This fixed point is extracted from the ground-truth generated in the training set, and it is assumed to be the pivoting point of rotation of the target spacecraft in the training images, whose position remains unchanged for any attitude. The fixed point is expressed relative to the origin of the minimum bounding-box enclosing the object and stored into the codebook for each relative attitude.

The position information is directly stored in the codebook, but the attitude descriptors present a large dimensionality, which makes unfeasible their direct storage on-board for both memory and computational reasons. For this reason, after the extraction of the whole set of descriptors, their principal components are computed using the 2DPCA-L1 algorithm described in [13]. The final set of descriptors is built by projecting the data over the first k principal components spanned by the basis vectors representing the principal components. Then, the projected descriptors associated to each pose are stored to populate the codebook along with the position information. In addition, the projection matrix is also stored, as it is necessary to project the new input descriptors from the classification stage.

The holistic nature of the descriptors allows to uniquely describe each possible attitude. This is advantageous with respect to widely used local methods that model each view as a set of descriptors that can be present in different attitudes. The rotation invariance removes the ambiguity caused by two similar rotated views of the target spacecraft. In addition, the use of HOG descriptors and the application of sparse 2DPCA-L1 over the pixel intensity values provide robustness against illumination conditions not present in the training stage

4.2.2 Classification

In this stage, input images from the on-board camera are processed to retrieve the best match for the associated attitude.

The input images to the system cannot be assumed to present ideal conditions in terms of noise or illumination. The introduction of a **pre-processing** stage is mandatory to improve the image in terms of signal-to-noise ratio and illumination:

- **Image denoising** is performed to filter the noise from the image signal while keeping the image details unchanged. The technique used for image denoising is the LLSURE operator [17]. This is an edge-preserving technique that assumes the output pixel is a linear transformation of the input pixel's neighbourhood. The linear transformation is driven by the SURE estimator, a mean-squared error unbiased estimator. It has O(N) complexity and it adapts automatically to the noise variance in the image, working under Poisson and Gaussian distributed noise.
- **Image Equalization** is performed to even the illumination conditions on the image. The image equalization technique used is CLAHE: a contrast limited adaptive histogram equalization [18]. The contrast limit avoids over-contrasted images that would lead into a binary output. Adaptive means the algorithm is designed to consider the histogram over small patches. In this way, the contrast is optimized on a local basis, leading to even illumination conditions in the output.

DOI: 10.13009/EUCASS2019-961

After the image pre-processing, the pixels representing the target spacecraft are extracted by means of image segmentation**. Image segmentation** is the process of subdividing the image into a set of regions that share similar characteristics. The goal is to divide the image into two regions: foreground and background. Foreground is understood as the region containing the object of interest, i.e. target spacecraft, whereas background comprises any other object present in the image. This step, previous to pose estimation, provides knowledge about the location of the object and avoids a (computationally expensive) full search of the object over the image.

Large variations in illumination conditions pose a challenge for image segmentation algorithms due to the casting of shadows that may occlude different parts of the target spacecraft, and thus leading into an incorrect segmentation. Multi-spectral imagery captures the radiance reflected by the object according to an external illumination source, whereas thermal imagery captures radiance emitted by the object, providing similar image conditions under different illumination scenarios including the absence of light in the scene. To minimize the effects of illumination conditions not considered in the development stage and possible segmentation errors that may arise, the thermal image has driven the development of the image segmentation algorithm as it presents similar characteristics over all the scenarios considered, providing robustness against variations in the source of light. As such, the foreground-background segmentation is implemented by using the information provided by the thermal band. The nature of the images allows a clear separation of the foreground in terms of pixel intensity levels under the illumination conditions present in the test scenarios. The segmentation is local-threshold-based in combination with morphological operators and a further optional refinement step with level-set minimization [19].

Once the target spacecraft is segmented, the **pose is retrieved** by extracting the features described in Section 4.1.The extracted features are projected onto the principal components matrix stored in the codebook and used to retrieve the attitude of the spacecraft by finding the nearest-neighbour in the codebook under the L1 norm. This distance metric proved to be the most discriminative under preliminary tests. Once the attitude is retrieved, the position, including the distance to the object is computed using the position information associated to the pose, and information retrieved from the image.

5. Active Debris Removal scenario results

In order to assess the validity of the proposed methods and the advantages of using a multispectral camera, a simulation scenario has been identified that envisages the “Rendezvous Phase” of an Active Debris Removal (ADR) mission: the chaser performs a rendezvous with the target object, evaluates its attitude dynamics and CoM position and, if required, performs a forced translation in order to reduce the relative motion between chaser and target to levels adequate to initiate the capture. A first hold point is assumed at a distance from the target of 100 m, and another one at 50 m.
The target is a large, dead satellite, which is tumbling. The tumbling target is assumed to have the following orbital parameters:

- Semi-major axis: 7145 km
- Type of orbit: polar, Sun-synchronous.

The target object is assumed uncooperative (no inter-satellite or target-ground communications, no hardware actively helping the rendezvous and capture shall be assumed to be available on the target) and non-passivated.
The chaser is assumed to be 3-axis controlled and able to perform relative navigation w.r.t. the target object during the full target orbit, anytime of the year. This relative navigation shall be fully autonomous without any ground intervention.
In order to verify the performance of the Navigation and Image processing algorithms, they have been integrated into a simulation framework that, according to the scenario description, is defined as follow:

- The relative motion trajectory is an ellipse centered in the target with semi-axis of 100 m and 50 m.
- At the beginning of the maneuver the chaser is 100 m behind the target, then a ΔV in the radial direction, toward the Earth, is applied.
- The camera line of sight direction with respect to the inertial frame changes with the mean anomaly of the spacecraft when the ΔV is applied. Then, it remains nearly constant during the scenario as shown Figure 2.

DOI: 10.13009/EUCASS2019-961

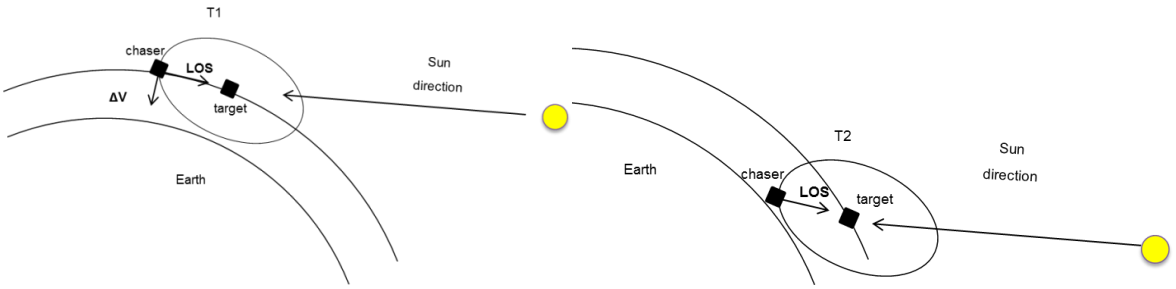


Figure 2 : Scenario description

The simulation starts from an arbitrary point in the orbit, changing the initial mean anomaly, maintaining the same mean anomaly of application of the ΔV.
The performance of the Navigation and IP have been tested in two configurations:

- Lost in space configuration: no previous knowledge on the target attitude or position; this test aims assess the influence on the performance of the IP and navigation filter of parameters such as the number of bands available, the sun phase angle and the presence of Earth in the background.
- Feedback configuration: the initial pose of the target is assumed to be known with a certain confidence level; this test aims assess the capacity of the IP and navigation filter to estimate the pose of the target during the whole trajectory.

5.1 Lost in space configuration

The first set of experiments aims to characterize the behavior of the IP algorithm in nominal conditions, i.e. without the navigation filter and without any prior knowledge about the pose of the target spacecraft.
The performance of the IP algorithm, in terms of mean attitude and position error, is measured under different combinations of bands, Sun-phase angles, distance to the target, and presence of the Earth in the background. To do so, 100 images of the target in random attitudes are generated for each combination of:

- Sun phase angles, ranging from 0 to 160 in steps of 10 degrees
- Distances to the target: 50 m
- Presence/absence of the Earth in the background

And these tested over four different configurations of the algorithm in terms of band usage:

- B1: All bands
- B2: VIS-TIR (2 bands)
- B3: NUV-VIS_TIR (5 bands)
- B4: NUV-VIS (4 bands)

The statistics obtained, in particular in terms of standard deviation, are used to feed the EKF with a surrogate of the confidence level of each IP output. To provide a more intuitive graphic representation, the average attitude and position error as a function of the Sun phase angle are represented in Figure 5-3 to Figure 5-8 for different distances to the target and Earth in the background conditions. The average position error is shown separately for the x, y, and distance components, as the latter is usually one degree of magnitude larger.

No Earth in the background

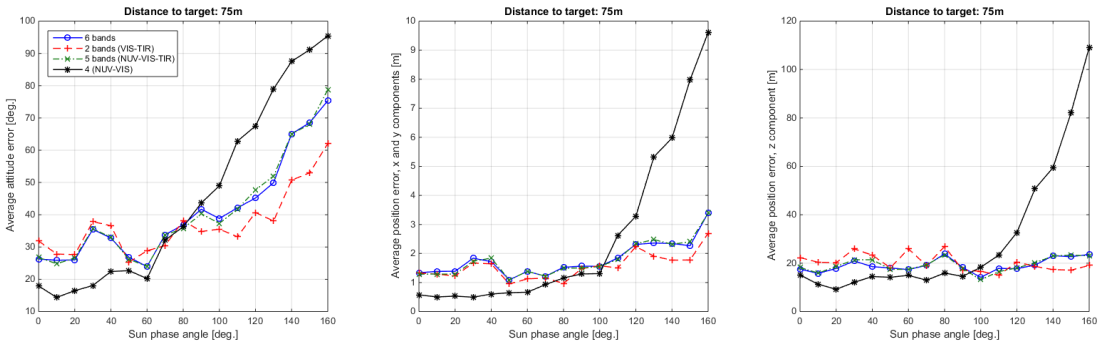


Figure 3 : Average attitude and position error as a function of the Sun phase angle. Distance to target 50m. No presence of Earth in the background

DOI: 10.13009/EUCASS2019-961

Earth in the background

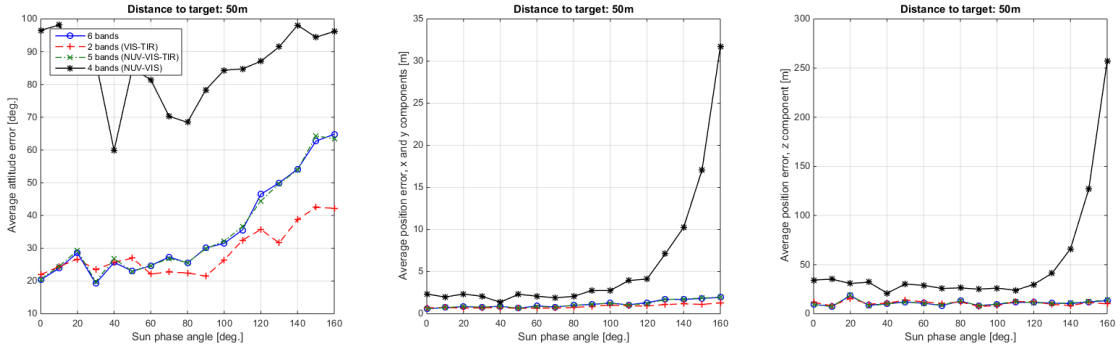


Figure 4: Average attitude and position error as a function of the Sun phase angle. Distance to target 50m. With presence of Earth in the background

A first conclusion on the results is the **importance of the thermal band for image segmentation** when the Earth is in the background (Figure 3), and for large Sun phase angles when the Earth is not in the background (Figure 3), where the multispectral-based segmentation errors caused by the strong shadows lead to large estimation errors.
In terms of robustness to illumination, the **multispectral bands are more affected by illumination changes**, namely larger Sun phase angles conditions. This is reflected in the results, where an increase of the Sun phase angles leads to bigger errors in configurations with a large number of multispectral bands, causing a loss of influence of the thermal infrared. It can be noted that, when the Sun phase angles are larger than about 100 to 120 degrees, the results worsen for the multispectral bands, being improved when the thermal band has more influence, providing more stable results in more difficult illumination conditions and in the presence of the Earth in the background.
However, in favorable illumination conditions (small Sun phase angles), **the configurations with a larger number of multispectral bands provide better results** in both terms of angle and position errors. It is of particular interest the case in which the influence of the thermal band is completely removed when the Earth is not in the background (Figure 3 NUV-VIS configuration), in which the lowest errors are achieved.

5.2 Feedback configuration

In this configuration, a feedback loop is considered between the navigation filter and the IP block. Two cases of study are presented in the following subsection:

- All six spectral bands, no Earth in the background
- All six spectral bands, presence of Earth in the background.

In all the simulation tests, an initial position estimation error of 10% is considered, in addition to an initial error angle of about 10 deg. This helps demonstrate the ability of the filter to recover from a wrong initial estimate, while also providing more meaningful results in the case no position and attitude measurements are used (i.e. propagation-only cases), as otherwise they could artificially lead to very small estimation errors.

No Earth in the background

For this case of study, particularly challenging illumination conditions have been considered, as depicted in Figure 5, with the target sub-exposed which leads to a lack of information in some critical areas of the spacecraft for the multispectral images. Figure 6 shows the estimated and measured relative position, relative velocity and estimated and measured relative attitude with the associated errors.

DOI: 10.13009/EUCASS2019-961

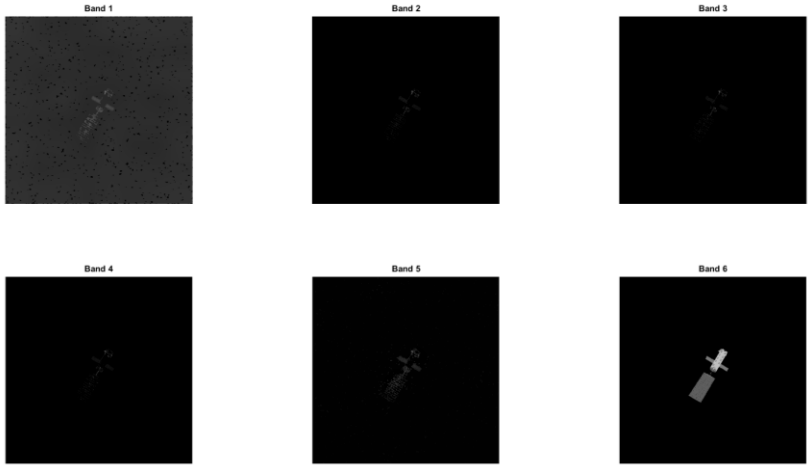


Figure 5: IP input for case 1 after denoising and equalization

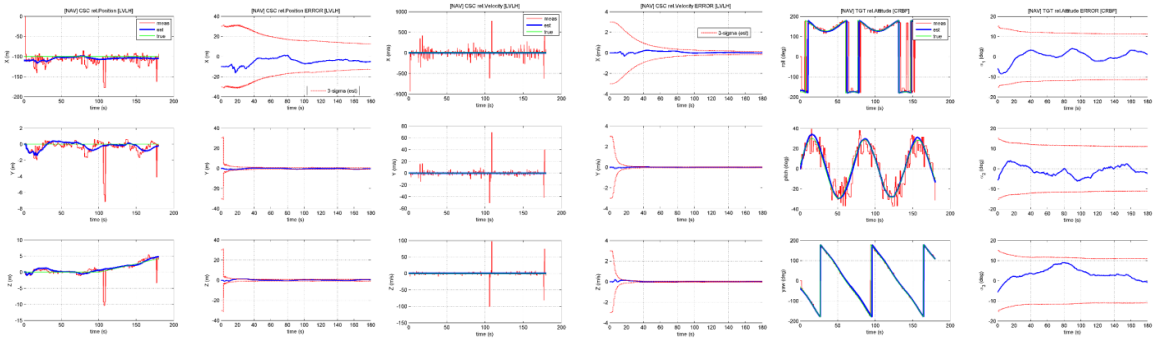


Figure 6: Estimated and measured relative position and associated error (left), estimated and measured relative velocity and associated error (centre), estimated and measured relative attitude and associated error (right)

Earth in the background

The features chosen for the IP are a concatenation of pixel intensity values and HOG: when the fine details of the image cannot be relied on, only a part of the descriptor loses relevance, while the part based on HOG carries enough information provided by the gradient of the contour. This can be observed in this case where, even if the light source is behind the target, its contour can be retrieved leading to results that are qualitatively similar to the ones in the previous case, as seen in Figure 8, showing the estimated and measured relative position, relative velocity and estimated and measured relative attitude with the associated errors. This proves the capabilities of the algorithm to work in conditions where features based on fine detail of the image cannot be relied on.

DOI: 10.13009/EUCASS2019-961

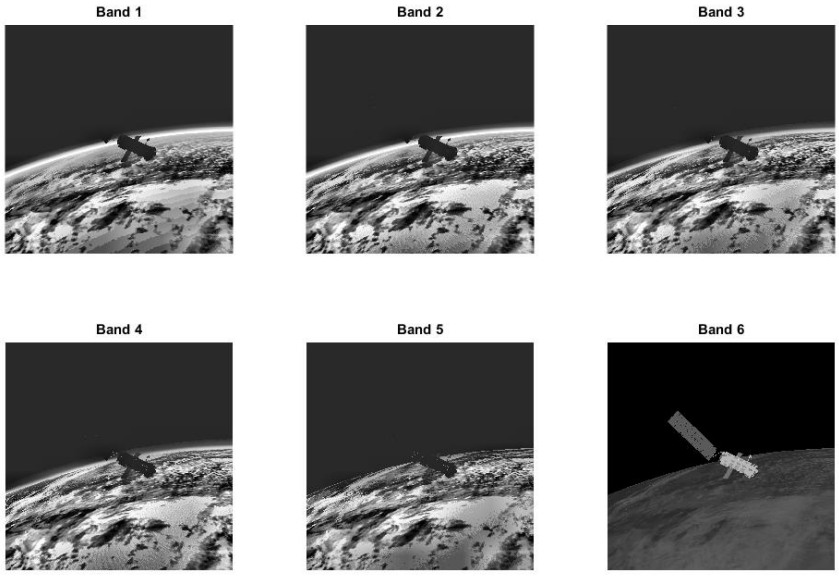


Figure 7: IP input for case 1 after denoising and equalization

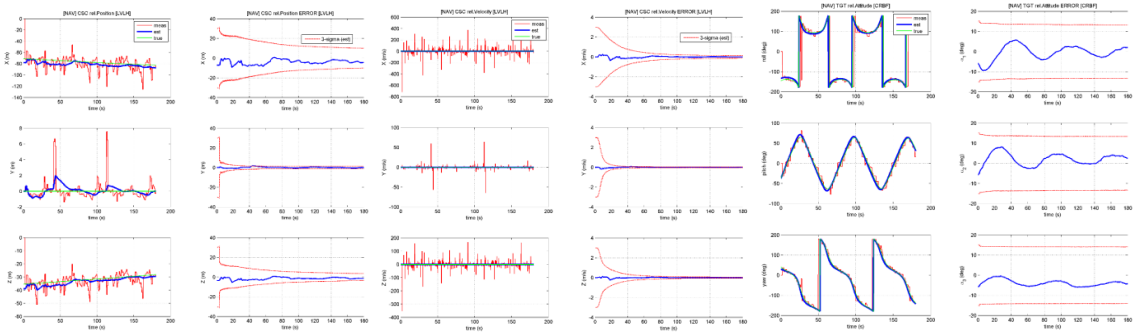


Figure 8: Estimated and measured relative position and associated error (left), estimated and measured relative velocity and associated error (centre), estimated and measured relative attitude and associated error (right)

5.3 Discussion

The IP algorithm chosen for the Rendez-vous scenario presented in the previous sections have been evaluated in different conditions and with multiple combinations of spectral bands. The algorithm does not require depth-information coming from a set of stereo-cameras, LIDAR, or time-of-flight sensors. Differently, it is designed to work on multi-spectral mono-cameras by exploiting the capabilities of each band with a high degree of modularity in terms of band usage. The proposed method has the capability of working without any prior knowledge of the spacecraft's pose (thus being able to operate in the lost-in-space condition), and converging to a reasonable good estimate of the attitude in a short period of time, or use prior knowledge to improve the estimation and reduce the computational load.

The thermal-infrared band, potentially with lower resolution and less capability of acquiring fine details, can be mainly used for segmentation. Then, the different spectral bands with larger resolution and capability of acquiring fine details may be used for the extraction of features and retrieval of the pose. Therefore, the first key advantage of a multi-spectral device is the availability of thermal-infrared measurements, which have been demonstrated to play a major role in proper image segmentation.

Another relevant feature of the IP algorithm implemented is that it does not require an on-line evaluation of a target spacecraft 3D model, and the posterior resolution of complex PnP problems, as the matching scheme is based on the L1 distance. This allows an efficient and parallel implementation of the matching process, with a straightforward way of introducing previous knowledge of the target attitude.

DOI: 10.13009/EUCASS2019-961

The features chosen for the pose estimation algorithm are designed to provide high robustness against illumination conditions and to directly capture the information provided by the different spectral bands. On the one hand, the multispectral information is introduced by using the principal components of each image band as a feature. On the other, HOG descriptors provide robustness to different illumination conditions as they rely on information provided by the gradient. However, the direct usage of the feature combination is unfeasible due to two main different reasons:

- The direct usage of pixel intensity values is not robust to changes in illumination and it is too dependent on the fine detail of the image.
- The large size of the images combined with the HOG descriptor would result in large image matching times for each of the different bands.

The introduction of a principal component analysis technique for reducing the size of the features allows an efficient manipulation of the dataset. The features are reduced in dimensionality by projecting them onto the principal components, retaining the meaningful information at a fraction of the original dimensionality. The two-dimensional nature of the principal component algorithm chosen (Sparse 2DPCA-L1) permits to set a larger number of images in the training stage, whereas sparsity encourages that only meaningful features are kept.

The fusion of the features is done at the decision level (high-level fusion), as the creation of image composites (low-level fusion) for training the algorithm presents a main inconvenience: the modularity in terms of usages of bands is lost. It is not possible to add or remove layers without explicitly training a codebook with that given configuration, which would lead to a very large number of codebooks to account for all the possibilities, giving rise to large memory requirements.

Differently, training one codebook per image band permits to set any desired configuration with a small number of trained codebooks. In conditions where the illumination is favorable, the usage of multispectral bands can improve the detection thanks to the capability of acquiring more details. In other scenarios, where shadowing can affect the capability of the multispectral images, the detection can be thermal-based.

The fusion scheme is based on majority voting between codebooks corresponding to each active band. If a given attitude has majority across codebooks, it is set as the final attitude. In the case no majority exists, the retrieved attitude with a smaller feature distance to the respective codebook is returned.

Hence, the second clear advantage of a multi-spectral device is the availability of several bands which, when combined with high-level fusion approaches like majority voting, can improve the overall robustness of the estimation.

6. Asteroid scenario results

The algorithm developed for this scenario estimates the target's relative angular velocity by measuring the visual displacement of visual features (feature tracking) extracted from two consecutive input images corresponding to two consecutive time instants. The visual features are extracted in each spectral band and the displacement information (in terms of pixels) is processed by an error model that provides the rotation error of the algorithm against a ground truth. Three cases have been considered representing different illumination conditions; for each case, different band combinations have been tested. Table 1 shows the asteroid’s angular rate determination error for all the cases studied and all the band combinations, presenting the total error and the isolated IP error contribution (in brackets, evaluated assuming the asteroid shape knowledge error and the rate knowledge error equal to zero), as well as the average number of tracked features per image, presented in Figure 9 for case 1. The use of TIR offers the best performance for 3 out of the 4 cases and the asteroid rate determination error for this band alone is the most stable one, confirming the robustness of TIR band to the illumination conditions.

It is also remarked that, for cases 1 and 2, which present the best illumination conditions, the errors are similar for all the combination of spectral bands, while for case 3 the combination of different spectral bands has a strong impact on the overall performance of the algorithm, resulting in a clear reduction of the error. In fact, VIS-NIR-TIR is the only combination of bands that ensures an error roughly below 25% in all of the scenarios considered, thus fully illustrating the clear advantages in terms of robustness and performance of the multi-spectral device.

DOI: 10.13009/EUCASS2019-961

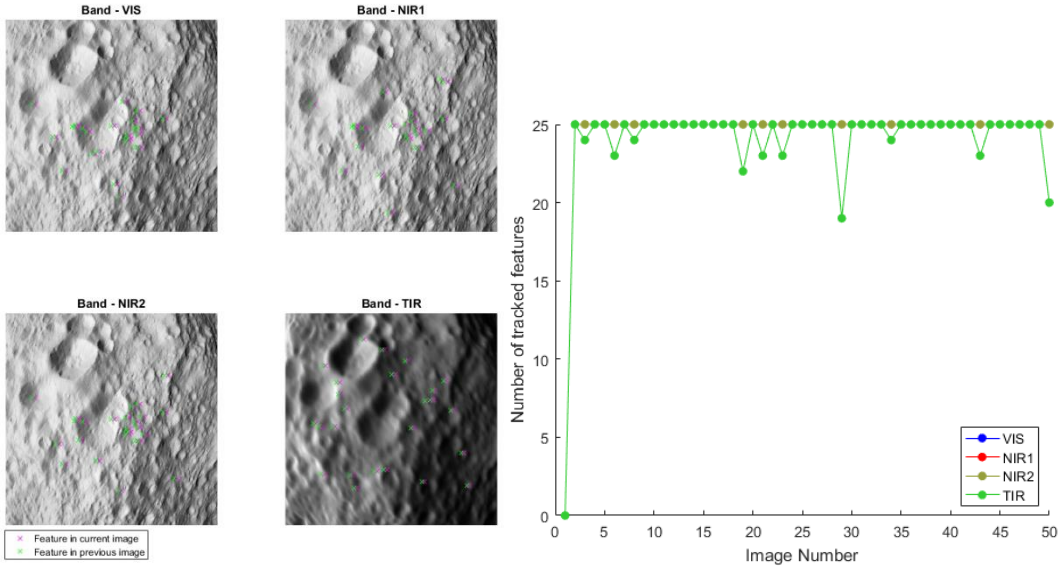


Figure 9: Tracked features for each band, case

Table 1: Asteroid Rate Determination Error for all the cases and all the band combinations considered. Error is a percentage of the actual rotation rate of the asteroid; in brackets, the error due to the IP

Case ID	Band Combination	Asteroid Rate Determination Error (3σ) (%, around inertial Z axis)	Av. number of tracked features per image
1	VIS	26.18 (15.15)	25
	TIR	**21.79 (10.99)**	24.5
	VIS-NIR	24.88 (14.72)	75
	VIS-NIR-TIR	22.97 (13.01)	99.5
2	VIS	25.98 (14.96)	25
	TIR	**23.29 (12.38)**	23.6
	VIS-NIR	25.37 (15.19)	75
	VIS-NIR-TIR	23.71 (13.72)	98.6
3	VIS	116.25 (89.79)	3.7
	TIR	51.34 (38.23)	24.7
	VIS-NIR	51.26 (37.37)	11.9
	VIS-NIR-	**22.06 (11.39)**	29.0

7. Conclusions

A novel, mono-camera approach for non-cooperative spacecraft pose estimation has been designed, implemented and successfully tested under several challenging scenarios with unfavourable illumination conditions and, in most of the cases, with the presence of the Earth in the background. The results of the test campaigns lead to the conclusion that the use of a multispectral camera for autonomous navigation presents several advantages over a standard optical camera, while certain limitations have also been identified. For both scenarios, the use of the thermal infrared band allows the IP to have robustness to illumination conditions. In the first scenario, the presence of the Earth in the background represents a challenge for the monocular visual-based segmentation that can be directly overcome with the use of a thermal infrared sensor; in the second scenario, the number of features that can be tracked is sensitive to the illumination conditions: if the scene is overexposed or covered by shadows, the number of visual-features is drastically reduced. The usage of the TIR band allows tracking features on those extreme cases. In addition, it helps the algorithm achieve more consistent results across different illumination conditions. However, when the thermal layer is not available, more complex techniques are required in order to obtain a reliable segmentation. A possible approach is the use of Deep Learning techniques: since the target is well-known and a large number of images can be synthetically generated, a Convolutional Neural Network may provide improved segmentation performance with a relatively low computational cost. Moreover, such a network can, in principle, be trained and validated on-ground and tested in-flight.

DOI: 10.13009/EUCASS2019-961

This study also provided a trade-off between the several possible IP solutions for multi-spectral devices. IP algorithms that are sensitive to small intensity variations are, by construction, less robust to changes in the illumination conditions. Hence, an approach based on descriptors robust to illumination conditions was adopted and shown to generate the expected results under a wide range of configurations. Moreover, the IP algorithms were designed in such a way so as to directly exploit the information present in the different multispectral bands, and to allow a modular usage in terms of band configuration without a significant increase in terms of computational complexity. Finally, the algorithms were selected taking into account also the associated computational limitations, as approaches that use, for instance, the 3D model of the target, typically require the resolution of complex PnP problems, thus not being regarded as baseline for this study.

References

[1] Degrez, G., P. Barbante, M. de la Llave, T. Magin, and O. Chazot. 2001. Determination of the catalytic properties of TPS materials in the VKI ICP facilities. In: *3rd ECCOMAS Computational Fluid Dynamics Conference*. 162–167.

[2] Magin, T., and G. Degrez. 2004. Transport algorithms for partially ionized and unmagnetized plasmas. *J. Comput. Phys.* 198:424–449.

[3] AGARD. 1998. A selection of test cases for the validation of large eddy simulations of turbulent flows. Agard Advisory Report 345. North Atlantic Treaty Organization.

[4] Rini, P. 2006. Analysis of differential diffusion phenomena in high enthalpy flows, with application to thermal protection material testing in ICP facilities. PhD Thesis. Université Libre de Bruxelles, Faculté des Sciences Appliquées.

[5] Yamanaka and Ankersen. New State Transition Matrix for Relative Motion on an Arbitrary Elliptical Orbit. In JCGD v.25, n° 1, Jan-Feb 2002

[6] Bravo, J. I, Spacecraft Pose Estimation for Non-Cooperative Rendezvous Based on Feature Codebook, Universidad Autónoma de Madrid, 2017.

[7] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886-893. IEEE, 2005.

[8] Lindsay I Smith et al. A tutorial on principal components analysis. Cornell University, USA, 51(52):65, 2002.

[9] Jian Yang, David Zhang, Alejandro Frangi, and Jing-yu Yang. Two-dimensional pca: a new approach to appearance based face representation and recognition. IEEE transactions on pattern analysis and machine intelligence, 26(1):131-137, 2004.

[10] Hui Kong, Xuchun Li, Lei Wang, Earn Khwang Teoh, Jian-Gang Wang, and Ronda Venkateswarlu. Generalized 2d principal component analysis. In Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on, volume 1, pages 108-113. IEEE, 2005.

[11] Parinya Sanguansat. Two-dimensional principal component analysis and its extensions. In Principal Component Analysis. InTech, 2012.

[12] Xuelong Li, Yanwei Pang, and Yuan Yuan. L1-norm-based 2dpca. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 40(4):1170-1175, 2010.

[13] Haixian Wang and Jing Wang. 2dpca with l1-norm for simultaneously robust and sparse modelling. Neural Networks, 46:190-198, 2013.

[14] Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse principal component analysis. Journal of computational and graphical statistics, 15(2):265-286, 2006.

[15] James J Kuffner. Effective sampling and distance metrics for 3d rigid body path planning. In Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on, volume 4, pages 3993-3998. IEEE, 2004.

[16] Xavier Perez-Sala, Laura Igual, Sergio Escalera, and Cecilio Angulo. Uniform sampling of rotations for discrete and continuous learning of 2d shape models. In Robotic Vision: Technologies for Machine Learning and Vision Applications, pages 23-42. IGI Global, 2013.

[17] Qiu, T., Wang, A., Yu, N., & Song, A LLSURE: local linear SURE-based edge-preserving image filtering. IEEE Transactions on Image Processing, 22(1), 80-90. (2013).

[18] S. M. Pizer, E. P. Amburn, J. D. Austin, et al.: Adaptive Histogram Equalization and Its Variations. Computer Vision, Graphics, and Image Processing 39 (1987) 355-368.

[19] Chan, T. F., & Vese, L. A. (2001). Active contours without edges. IEEE Transactions on image processing, 10(2), 266-277.