ABSTRACT
This paper proposes several Computer Vision methods for Vision-Based Navigation with the estimation of the relative pose (position and orientation) of a service spacecraft (chaser) with respect to a target object at different ranges, for use in Rendezvous manoeuvres, On-Orbit Service (OOS) operations and Debris Removal, in this case, with non-cooperative targets. At far distances (around 30 km) a visual navigation can be performed using a star catalogue when separating the stars and camera hotspots from the initial candidates for the target spacecraft in the identified points of interest within the camera image.

For manoeuvres at close distances (as for instance in the case of initial de-spinning of non-collaborative targets and debris removal) and in Rendezvous and Docking (RvD) operations, results are provided using two camera set-ups: Monocular (with the known ambiguity in depth/scale solved when fitting to the available 3D model description of the target object) and binocular (for better mapping of the identified features with the model). In some cases complementary active illumination of the target (pulsed, structured, etc) has been taken into account in order to improve the accuracy of the estimation of the relative attitude and position. These methods are accommodated to exploit the multiple view geometry in a sequence of images resulting from the change in spatial configuration of the two space objects following their relative movement. Sensitivity analysis is also done in order to highlight the limitations in the accuracy on the dynamics parameters that can be extracted from the cameras(s) following specific trajectories.

Simulations are performed with 3D models and mock-ups of actual satellites which allow to evaluate the achievable performance in realistic scenarios. These methods can be applied in other areas of Vision-Based Navigation when orbiting space objects like asteroids or comets, where certain features of the target could be tracked at the same time, providing an initial rough model of the shape of the object which could be gradually refined.

1. INTRODUCTION
Vision Based Navigation has been put in practice already for actual missions that involve Rendezvous and Docking (RvD) operations between two Spacecrafts in orbit, and while the situations arising from manoeuvres with cooperative targets have been tested in Space, limited experience is available in the case of non-cooperative targets, for instance for debris removal. This triggers the need for dedicated activities, first by simulations and by representative conditions in a laboratory (see for instance [1], [2] and [3]) and then in In-Orbit-Demonstrators, to develop techniques, algorithms, SW and HW, that could fit in the more complex situations within those conditions. Extensive literature exists (see for instance [7]) in the case of combination of Visual sensors (cameras, in the visible spectrum and in the Infrared (IR)) and LIDARs, and this can be used as a first step by means of a Technology demonstrator and for calibration/evaluation purposes in orbit, for the new techniques and algorithms, assessing the performance of these devices alone, when comparing with the other sensors in some specific missions.

Special attention has to be devoted to the set-up of the Laboratory on-ground (including the simulation environment) for having testing configurations that would be meaningful of those ones in orbit, because for instance the illumination conditions (changing over time) are driven, in the case of Low-Earth orbit by the Sun (when not in eclipse), the Earth albedo, reflections of other parts of the S/C (from those light sources and from other potential active illumination coming from the chaser S/C).

The results described in this paper are simple and effective methods that could be combined with other well established more sophisticated techniques from the Computer Vision area. They exploit the Multiple View Geometry (see for instance [13]) resulting from the consecutive positions and orientations of a camera mounted on a moving chaser S/C. The binocular condition for two images arises as a special case for fixed conditions between the cameras (rotation and translation are known) and the remaining problem would be the feature matching (point correspondence) between the two images.

2. COMPUTATION OF ROTATION AND TRANSLATION FROM THE FUNDAMENTAL MATRIX
As an example of this type of computations, a couple of
images are taken of a Satellite mock-up (extracted from a sequence of them) in a laboratory set-up and are used to compute the Fundamental Matrix (for details on the theoretical background for this see for instance [13]) and from there the Rotation and Translation matrices between the camera position and pointing when taking those pictures, simulating the camera installed in the chaser looking at the target:

![Figure 1 Two images taken from a Satellite mockup (test set-up) at different positions of the camera](image1)

![Figure 2 Composite overlaid Image (Red – Left image; Cyan – Right image)](image2)

![Figure 3 SURF features](image3)

![Figure 4 Matched points in the two images](image4)

![Figure 5 Inlier points (RANSAC)](image5)

![Figure 6 Inlier points in the rectified stereo images](image6)

Extending the above process, a sequence of images of the relative movement of the camera with respect to the camera can be processed in a similar manner, repeating the computations for pairs of images.

3. COMPUTATION OF RELATIVE ROTATION AND TRANSLATION BY USING PnP (OR SIMPLIFIED P3P) COMPLEMENTED WITH FEATURE TRACKING

The Perspective-n-Point (PnP) problem, also known as pose estimation consists on the determination of the position and orientation of the camera, whose intrinsic parameters are considered known, with respect to an object in a 3D scene for which $n$ correspondence points (between marks in the image plane and points in the 3D world related by a camera projection) are given. This is summarized in [10] as follows:

“Given the relative spatial location of $n$ control points, and given the angle to every pair of control points from an additional point called the Center of Perspective (CP), find the lengths of the line segments joining CP to each of the control points.”

In practical terms a Reference (“Model”) image is taken and then points in that picture are linked to points in a 3D model and this image is taken to initialize the calculations for the matching correspondences (see for
instance [6]).

Being the PnP one of the fundamental problems of photogrammetry, a lot of effort has been devoted to optimizing algorithms that provide solutions to this problem, and for instance the methods described in [4] give results for which the computational time depends linearly with the number of points (thus O(n)). For further details on the PnP see other sources mentioned in the References (section 10).

For simplicity and for its intuitiveness, we will consider in this paper the minimal problem P3P where only 3 point correspondences (in fact 4 would be better in general to solve ambiguities in some cases) are considered with a calibrated camera (an algorithm for solving this P3P problem is described for instance in [11]).

In the example presented in Figure 7 it is possible to compute the position and orientation of the camera by marking manually 3 initial points in the image (green triangle) for which the sides of the triangle are known (in the S/C mock-up in this case, but also in the actual satellite). With the solution of the P3P problem the pose of the camera, and thus the Coordinate System associated to the Chaser (“observing”) Spacecraft can be expressed in the Reference Frame defined by that triangle (attached to the target S/C). In all this calculations it is assumed that the camera is calibrated (focal length known).

When a sequence of images is given, resulting from the relative movement of the chaser around the target (as commented in section 2) once the solution has been initiated and a 3D model is linked to the first image it is possible to associate a number of tracking points to the corresponding points in the 3D model in such a way that always some of them will be visible. This represents an iterative process for which features in the images are better fitted into the 3D model.

4. ILLUMINATION OF THE TARGET: STRUCTURED LIGHT

Using similar techniques as the ones described in section 3 for the P3P problem, it is possible to compute the position and orientation of both the camera and a projector of structured light with respect to the calibration pattern and thus the relative pose of the projector with respect to the camera. A minimum number of 3 points are retrieved as x-y coordinates in the image (see Figure 8). As commented before for the PnP or the general camera calibration methods, more points in the overdetermined problem could provide more accuracy in the calibration.

This is a direct calibration method (introducing potential effects on the pointing of the projected grid, etc) as compared with the indirect method of calibration by measuring directly the actual positions of both the projector and the camera in the chaser S/C (both calibration methods can be complementary).

Figure 8 Calibration pattern for camera and projected structured light

Figure 9 Image of the S/C mock-up with Calibrated Grid of structured light and projected sub-grid onto the S/C model

By frame differencing (in practice rapid flashing with and without illumination with a pulsed light; in a static configuration in controlled conditions in the laboratory, just by occluding the projector), and using the Green channel of the image (being in this case the colour of the structured light) to filter the rest of the colour content, the illuminated and non-illuminated image can
be considered at the same relative position in 3D and we obtain the image with the differences in Figure 10 (the actual projected grid is presented in blue colour).

It should be noted that other parts of the S/C, like the dish antennas at the top are also illuminated by unwanted reflections, but as long as the relative orientation with respect to the S/C (represented here as a 3D model) is roughly known it is possible to isolate the projected subgrid as a Region of Interest (RoI) and assess the perspective deformation (also in terms of size) of the corresponding quadrilaterals to refine the estimation of the relative position (including range) and orientation of the camera (in the chaser S/C) to the target the S/C.

It is possible to project consecutively different subgrids (or squares within the grid) in a known sequence in order to improve the process, refining the estimation by comparing the subgrid onto the S/C with the estimated (modelled) projected one, taking advantage of the fact that the relative configuration of the projector and the camera is known very accurately.

The advantage of projecting specific subgrid of the pattern or individual quadrilaterals is that it allows for identifying and isolating them within the pattern, while the general grid, as a combination of all of them, covers a wider projection covered angle.

Well-known techniques for 3D scanning for the generation of depth maps work when the relative geometry between the object and the camera (with the attached laser plane/line) is changing (in this case because of the relative movement of both S/Cs) in a known way. Thus these computations are complementary with those ones associated to the relative attitude and position of both Spacecrafts.

5. MATCHING TO 3D COMPUTER MODELS

As an additional example of the techniques described in section 2, Figure 12, Figure 13 and Figure 14 present the different stages in the calculation of the rotation angle in a 360 degree pitch manoeuvre (retrieved from [21]). Figure 15 and Figure 16 depict the results at a later moment in the manoeuvre, which is used in Figure 17 to manually associate a number of points to a computer graphic 3D model of the Space Shuttle in the way that it was described in section 3 for the Reference ("Model") image that could link the images from the video sequence with the 3D model from that moment onwards. This 3D model is particularly useful for improving the results in the estimation of the position and orientation of the Target S/C for relative movements that have poor observability (not well conditioned in the direct Computer Vision problem). For relating the two camera models (the one that has taken the sequence of images and the one that rendered the 3D model) it is necessary to know (or to estimate) the calibration parameters (focal length, etc) of both.
In these calculations there is a trade-off in the time step between consecutive images in the sequence, which should be apart enough for representing a significant movement that can be estimated (measurable, thus distinguishable from the image noise) and frames that should be close enough that a significant number of features are present in both images so as to have a robust tracking of features (continuous during a reasonable time). These conditions depend on the relative dynamics of both S/Cs and the characteristic times (translation and rotation) of their movements, which in any case have to be taken into account for pointing accurately the camera in the chaser spacecraft.
6. TOOLS FROM THE COMPUTER VISION FIELD

In the last years the Computer Vision field has experienced a vast and general progress, where major breakthroughs alongside with optimized algorithms, improvements in SW and miniaturization and affordability of the HW allows for general utilization of these technologies as common-place in areas that some time ago were considered impractical.

In this sense it is worth mentioning how for instance commercial applications can be used to generate on-the-fly 3D models by taking a number of pictures at different positions around an object (see for instance the example in Figure 19) and the techniques used in SLAM (Simultaneous Localisation and Mapping), as for instance presented in [28], [29] and [30] provide an insight on the possibilities of these technologies.

Additional applications are related to the use of Plenoptic cameras for the computation of Depth from De-focusing but they seem to be only a plausible option in close proximity when the relative range between the two S/Cs is comparable with the characteristic lengths to be measured in the target S/C.

![3D model reconstruction from sequence of pictures and computation of camera position using 123DCatch™](image)

Figure 19

7. POSE ESTIMATION ERRORS IN THE P3P PROBLEM: SENSITIVITY ANALYSIS

For simplicity and for the geometrical usefulness, an error analysis of the P3P problem is provided in this section (analysis of the impact on the solution of the more general PnP problem for the errors associated to the features location can be found in [5]).

Assuming that the lengths of the sides of the reference triangle ABC in the 3D model are perfectly known, errors in practice on the solution of the P3P problem are the result from inaccuracies in the measurement of the projected points in the image, which in turn is translated into slightly wrong angles $\alpha$, $\beta$, and $\gamma$ from the point D (see Figure 20) at the camera center (assuming a pinhole camera model) to the Reference points A, B and C.

![P3P geometry for errors in the angles $\alpha$ and $\beta$](image)

Figure 20

Variations in the position of point D can be related to changes (errors) in the set of angles $(\alpha, \beta, \gamma)$ or ranges $(r_A, r_B, r_C)$ as described by the formulas in [Eq.1], [Eq.2] and [Eq.3]:

$$\Delta D = (\Delta r_A, \Delta r_B, \Delta r_C)$$

[Eq.1]

$$J_{\Delta (\alpha, \beta, \gamma)} = J^{-1}_{(\alpha, \beta, \gamma)}$$

[Eq.2]

$$J_{\Delta (r_A, r_B, r_C)} = J^{-1}_{(r_A, r_B, r_C)}$$

[Eq.3]

allows to compute, given some small variations in the angles $(\delta \alpha, \delta \beta, \delta \gamma)$, the infinitesimal variations in the ranging distances $(\delta r_A, \delta r_B, \delta r_C)$ and from there the variations in the position and orientation of the camera, located at vertex D in Figure 20. The orientation of the camera with respect to the triangle ABC can be computed from there due to the fact that in a calibrated camera with a pinhole camera model, the projected points Ap, Bp and Cp lie in the lines AD, BD and CD, respectively, at distances from D that are known from the position of those pixels in the image.

In the example shown there is an amplification factor of the errors in the markers for the estimation of the
position of the camera. As depicted in the configuration of Figure 21 the errors are usually roughly parallel to the triangle ABC, which correspond in fact to an error distribution approximately in the directions defined by the vectors $\frac{\partial \mathbf{D}}{\partial r_A}$, $\frac{\partial \mathbf{D}}{\partial r_B}$ and $\frac{\partial \mathbf{D}}{\partial r_C}$, which are perpendicular at D to the planes DBC, DAC and DAB, respectively.

![Figure 21 Typical distribution of the errors in the estimation of the position of the camera in the P3P geometry](image1)

![Figure 22 Error distribution in the estimation of the Pointing of the camera in the P3P geometry (computed for the point A in the Reference triangle)](image2)

8. ACKNOWLEDGEMENTS

The author would like to thank Luis Felipe Peñín and Aleke Nolte for several fruitful discussions on related topics.

9. DISCLAIMER

The work presented in this paper is related with the technical research of the author as a Ph.D. candidate in Computer Vision, it is not necessarily related to any ESA project and might not reflect the current official view of ESA on these matters.

10. REFERENCES


[19] Thomas Werner, Solution of P3P problem, Center for Machine Perception, Czech Technical University, Prague


