Downloaded from orbit.dtu.dk on: Dec 20, 2017

Technical University of Denmark

Pose Estimation of an Uncooperative Spacecraft from Actual Space Imagery

D'Amico, S.; Benn, Mathias; Jørgensen, John Leif

Published in:
Proceedings of 5th International Conference on Spacecraft Formation Flying Missions and Technologies

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
D'Amico, S., Benn, M., & Jørgensen, J. L. (2013). Pose Estimation of an Uncooperative Spacecraft from Actual Space Imagery. In Proceedings of 5th International Conference on Spacecraft Formation Flying Missions and Technologies Munich, Germany.

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

POSE ESTIMATION OF AN UNCOOPERATIVE SPACECRAFT FROM ACTUAL SPACE IMAGERY

S. D'Amico[(1)], M. Benn[(2)], and J.L. Jørgensen[(3)]
[(1)]*German Aerospace Center (DLR), Münchner Str. 20, 82234 Wessling, Germany, simone.damico@dlr.de*
[(2)(3)]*Technical University of Denmark, DTU Space, Elektrovej, Building 327, DK-2800 Lyngby, Denmark, mb@space.dtu.dk and jlj@space.dtu.dk*

Abstract: *This paper addresses the preliminary design of a spaceborne monocular vision-based navigation system for on-orbit-servicing and formation-flying applications. The aim is to estimate the pose of a passive space resident object using its known three-dimensional model and single low-resolution two-dimensional images collected on-board the active spacecraft. In contrast to previous work, no supportive means are available on the target satellite (e.g., light emitting diodes) and no a-priori knowledge of the relative position and attitude is available (i.e., lost-in-space scenario). Three fundamental mechanisms – perceptual organization, true perspective projection, and random sample consensus – are exploited to their full extent to overcome the limitations of monocular passive optical navigation in space. The preliminary design is conducted and validated making use of actual images collected in the frame of the PRISMA mission at about 700 km altitude and 10 m inter-spacecraft separation.*

Keywords: *Pose Estimation, Optical Navigation, Uncooperative Spacecraft, Flight Results, PRISMA.*

1. Introduction

Monocular vision navigation has been identified by various authors as enabling technology for present and future formation-flying and on-orbit-servicing missions (e.g., PRISMA – Sweden [1], DEOS – Germany [2], PROBA-3 – ESA [3]). Indeed simple instruments such as star trackers or infrared cameras can be modified for increased dynamic range to accurately navigate with respect to a target space vehicle at low cost from virtually zero to several tens of kilometers [4].

This paper addresses the implementation of a novel vision system for estimating the pose (relative attitude and position) of a non-cooperative space resident object using single two-dimensional images given its three-dimensional geometric representation (model). Unlike most previous approaches [5-6], the estimation of the six translational and rotational degrees of freedom is accomplished without range measurements and without any a-priori information on the relative motion between the spacecraft. Furthermore the development and verification of the pose estimator are based on actual images collected during the PRISMA mission by two cameras available onboard the Servicer satellite at close separations (5 – 100 m) from the Client.

Although this work is largely inspired by computer vision methods developed in the 80's for terrestrial applications (e.g., Lowe – 1987 [7], Fischler – 1981 [8], DeMenthon – 1995

[9]), a number of improvements have been introduced to cope with the deficiencies encountered when dealing with space navigation. Typically the initialization procedure of the pose estimator is not able to handle lost-in-space configurations without a-priori knowledge. Second space imagery is characterized by low carrier-to-noise ratio and high image contrasts, which cause false or partial detections of physical edges and the extraction of a restricted amount of feature points. Third most artificial satellites have symmetric polyhedral shapes. Although these are easy to model, the perspective equations provide several ambiguous pose solutions in the presence of high symmetry which can not be neglected in the vision system. Furthermore the unknown correspondences between image and model features result in a large search space for ambiguity resolution, and thus in an unacceptable computational load.

A number of requirements for future monocular vision-based systems can be derived based on the aforementioned challenges. Ideally a pose estimator shall rely on a minimum number of image features, minimize the search space for model matching, be robust to ambiguous pose solutions, compensate for image noise and model errors, not rely on the uncooperative satellite dynamics, and offer multi-stage solutions of increasing accuracy from lost-in-space all the way to fine navigation. In order to fulfill these objectives the vision system prototyped in this work makes use of three mechanisms which underlie six sequential functional modules (from image processing to final pose fitting). First a method of perceptual organization is adopted to detect those feature groupings in the image which remain partially invariant over a wide range of viewpoints (i.e., orientation and scale) [10]. These can be matched to corresponding structures in the object model in a probabilistic manner. The probability of accidental instances from randomly distributed segments is evaluated and the size of the matching search space is reduced by combining primitives into larger structures. The software complexity is reduced by the fact that the same perceptual organization functions are used in the spacecraft modeling, image processing, and model matching modules. Furthermore the perceptual groupings of the spacecraft model can be pre-computed once before the mission launch or, if necessary, during the commissioning of the vision-system.

Second the true perspective projection equations are expressed as a function of the six pose state unknowns and the available measurements [11]. The partial derivatives are rigorously computed from the reference state to compute the modeled-image error. Each matched image segment provides three measurements, namely the coordinates of the endpoint and the angle subtended by the corresponding line. This selection of data-types is shown to be beneficial in the presence of unexpected line terminations. The linearization of the image perspective equations is used in two stages of the estimator. During initialization a complete linear system can be solved for the unknown corrections to the reference state through a multi-dimensional Newton-Raphson method [12]. Full convergence with pixel level accuracy is achieved with 3-4 iterations, provided that the initial pose is within approximately 40 deg error. The paper describes an extensive search method which is able to provide such initial pose estimate relying only on the visibility specification attached to the spacecraft model. After the extension of the initial set of matches through projection of the model onto the image, the estimation

problem becomes over-determined. The newly available measurements are processed in a weighted iterative batch least-squares estimator with a-priori information.

Third a random sample consensus method is used in combination with the least squares fit to improve robustness against false model-image matches [13]. In contrast to conventional smoothing techniques, this approach uses as small an initial data set as feasible and enlarges this set with consistent data when possible. The measurement/model compatibility is based on perceptual organization, or equivalently on the probability of non-accidentalness of the considered groups of features. The quality of the estimated state is judged based on the post-fit residuals and the formal output standard deviation of the state parameters.

In addition to the detailed description of the vision-system, the paper focuses on flight results obtained from the processing of actual images taken in orbit. Reference flight dynamics products based on precision carrier-phase differential GPS (relative position) and coarse sun-magnetic attitude estimates (relative attitude) are compared with the pose estimator results [14-15]. Deviations of the order of 10 cm in position and 10 deg in rotation can be inferred at 10 m separation. The paper is concluded by a discussion on the achieved robustness, the potential extensions of the vision-based approach, and the suggested improvements in terms of system architecture and navigation algorithms.

2. Problem Statement

The goal of this research is to estimate the relative position, $\boldsymbol{t}_B$, and orientation, $\boldsymbol{R}_{BC}$, (pose) of a non-cooperative spacecraft (no markers or other specific supportive means) from a single two-dimensional (2D) image given its three-dimensional (3D) geometric representation (model). As depicted in Fig. 1, the relative position vector is expressed in the body or structural frame (subscript B) attached to the Client passive spacecraft, whereas the relative attitude defines the rotation matrix from the Client body frame to the coordinate frame attached to the vision camera head unit (subscript C).

The close-range vision camera is embarked on the Servicer spacecraft and delivers 2D low-resolution (752x580 pixels), gray-scale (8 bit per pixel), monocular (effective focal length, f = $20187 \cdot 10^{-6}$ m) images of the Client. Two examples are shown in Fig. 1, where the image frame is indicated by the subscript D. The algorithm prototyping and testing described in this paper is based on such images collected in October 2011 in the frame of the PRISMA mission. During the considered flight experiments, the Servicer (Mango) approaches the Client (Tango) satellite at about 10 m separation in forced motion control based on relative GPS navigation [1,14-15].

Each image point ρ_D can be expressed as a function of the unknown pose ($\boldsymbol{t}_B$, $\boldsymbol{R}_{BC}$) according to the following 3D-2D true perspective projection equations

$$\boldsymbol{r}_C = (x_C, y_C, z_C)^t = \boldsymbol{R}_{BC}(\boldsymbol{p}_B - \boldsymbol{t}_B)$$
$$\boldsymbol{\rho}_D = (u_D, v_D) = \left(\frac{x_C}{z_C}\frac{f}{du}, \frac{y_C}{z_C}\frac{f}{dv} \right) \tag{1}$$

where $\boldsymbol{p}_B$ represents the corresponding point in body frame (or the matched model point), u_D and v_D are its coordinates in the image frame, du and dv are the different pixel sizes $8.6 \cdot 10^{-6}$ m and $8.3 \cdot 10^{-6}$ m respectively.

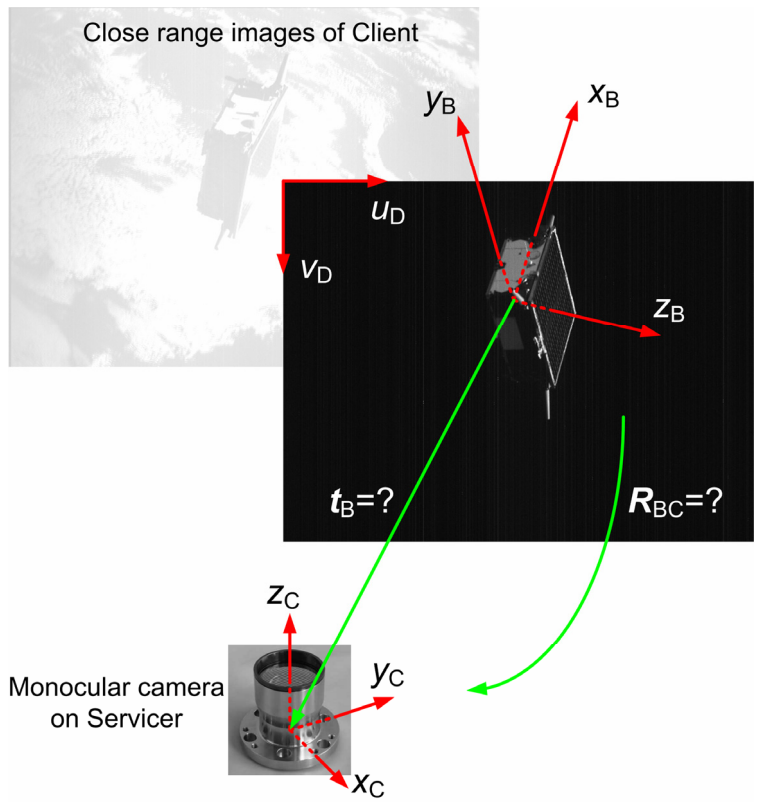

Figure 1. Coordinate frames and unknowns of pose estimation process.

Several considerations can be done on Eq. (1). First of all, the relationship between image points and pose parameters is highly non-linear. Although it is straightforward to describe the projection from a 3D scene to a 2D image, the inverse problem is much more difficult and can have infinite solutions if under-constrained [8]. Second, the correspondence between $\boldsymbol{p}_B$ and ρ_D is unknown. Finally, the image coordinates need to be corrected for non-quadratism and lens distortion before being used in Eq. (1). Despite these obstacles, the pose estimation approach builds on Eq. (1), which constitutes a rigorous observation model and can be easily derived for the computation of partial derivatives w.r.t. the pose state.

Eq. (1) suggests the main ingredients of a pose estimation system. One needs the capability of extracting Client features from the available image (image processing) to obtain measurements, ρ_D. These features have to be matched to correspondent

elements of a Client model (model matching), $\boldsymbol{p}_B$. Finally, the unknown pose translation and rotational parameters need to be estimated based on the available measurements (pose estimation). The following sections are dedicated to each of these fundamental tasks which are highly interconnected.

3. Image and Model Processing

3.1 Perceptual Organization

Most model-based vision systems involve a matching process where a set of k segments extracted from the image need to be matched to a set of j segments of the model. The resulting search tree contains $(j+1)^k$ nodes and requires an efficient strategy to reduce the computational effort. A promising technique is given by the method of perceptual grouping, which is inspired by mechanisms and constraints at the basis of the human visual system [10]. Instead of considering only single image points or segments, these are combined into perceptually relevant structures known to be present in the model, and called perceptual groups. Only those perceptual groups which have the highest probability of not being accidental (i.e., not arising from randomly distributed segments), are considered in the matching process. This leads to a reduction of the search space and to an efficient matching process.

The simplest perceptual groups are adopted in this research, namely segments with endpoints in close proximity, segments which are parallel and collinear segments. These primitive features are then combined into larger structures such as parallelograms and trapezoids in order to reduce the search space even more. The probability, P, of accidental instances from arbitrarily located segments can be evaluated according to the following scale independent expressions [7]

$$\begin{aligned} P_{\mathrm{PR}} &= 2\pi r^2 / l^2 \\ P_{\mathrm{PA}} &= 4\theta s\, l_2 / \left(\pi l_1^2\right) \\ P_{\mathrm{CO}} &= 4\theta s \left(g + l_1\right) / \left(\pi l_1^2\right) \end{aligned} \tag{2}$$

where PR, PA, and CO stands for instances of proximity, parallelism, and collinearity as depicted in Fig. 2 (please refer to figure for symbol explanation). The significance value for the higher-level groups is computed by multiplying the probability of non-accidentalness for each constituent.

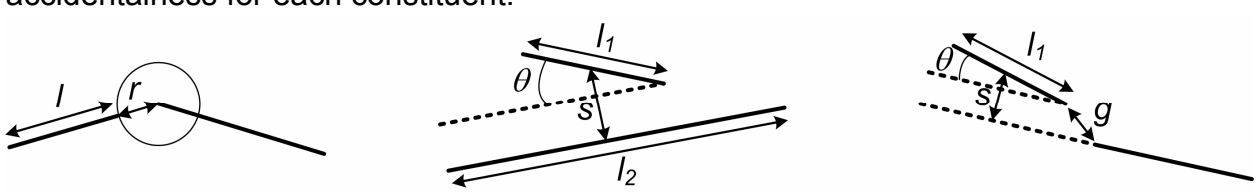

Figure 2 Geometry of proximity (left), parallelism (middle), and collinearity (right) (adapted from [7]).

The matching process considers only those perceptual groups which have the lowest probability of being accidental (i.e., $P << 1$). In addition all perceptual groups are ranked according to their significance to simplify the subsequent matching stage.

3.2 Image Processing

The goal of the image processing stage is to extract line segments which correspond to true edges or at least to edges available in the spacecraft model. Unfortunately the conditions of visual navigation in space make this task difficult and error prone. This is mainly due to the low carrier-to-noise ratio and to the high-contrast illumination which cause the detection of a small number of segments or features. Quite frequently the detected segments identify double edges or partial edges which can induce severe problems in the pose estimation scheme if not accounted for.

Despite these difficulties, here well known off-the-shelf techniques for line detection have been adopted [16]. The approach follows three main steps: i) low-pass filtering, ii) Canny-edge detection [17], and iii) Hough transform [18]. Each of these steps involve a number of process parameters to be set, such as the kernel size for i), the hysteresis and Gaussian smoothing parameters for ii), the tresholding and segment length parameters for iii). In general these functions are difficult to tune for broad applicability, thus the numerical examples shown hereby can be seen as conservative test cases for the pose estimation strategy.

Figure 3 illustrates the aforementioned image processing steps applied to an image of the Client spacecraft taken in orbit. Due to the illumination conditions, the Canny-edge detector can only extract features which basically reside on the two sides of the spacecraft illuminated by the Sun. The subsequent Hough transform can only identify line segments which correspond to multiple and truncated edges (see Fig. 3 – right). Unfortunately a considerable number of false edges have been detected, such as the output segments within the solar panel or on the cylindrical radio-frequency antenna (bottom of image).

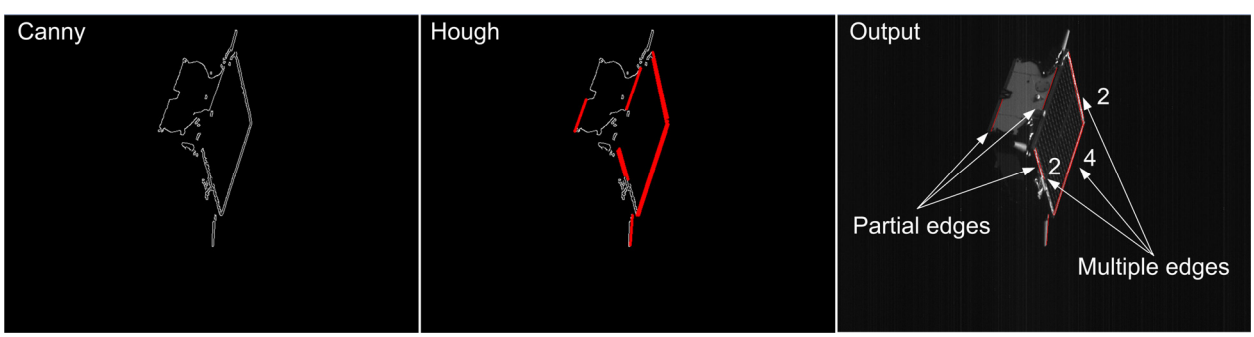

Figure 3 Steps of image processing process: Canny edge detector (left), Hough transform (middle), and output segments superimposed to original image (right).

Considering the design case shown in Fig. 3, the output of the image processing is a list of 10 segments identified by 20 endpoints through their pixel coordinates. Perceptual

groups are computed for approximate instances in 2D (i.e., $P \leq P^{max} = 0.1$). For simplicity and to reduce ambiguities, multiple edges are removed by considering only the most perceptually significant segments (i.e., with the most meaningful connections). This leaves 4 image segments which are combined in 3 parallel and 1 proximity perceptual groups. For completeness, the perceptual groups output of this stage are listed in Table 1 together with their significance scorings. It is noted that a camera calibration model is applied to the coordinates of the extracted segments to correct for distorsion and non-quadratism [19]. This is essential before comparison with the spacecraft model.

3.3 Spacecraft Model

The definition of a proper spacecraft model is a fundamental step of the pose estimation strategy. On one hand, the spacecraft model has to be as minimalist as possible to reduce the system complexity and the search space for matching. On the other hand, the model must be representative to be able to find image matches among the available perceptual groups. A reasonable realism of the spacecraft model is also required to reduce the number of ambiguities associated with the high symmetry of man-made polyhedral shapes.

Figure 4 depicts the Client spacecraft model which has been traded-off in the frame of this research. Basically, it consists of two stacked symmetric and convex polyhedra – one representing the solar panel (56x75x2.5 cm), and one representing the spacecraft body (56x55x30 cm). The centers of the two polyhedra are not aligned, and three additional segments (20.4 cm) represent the radio-frequency antennae. The origin of the body frame is located at the center of the bottom face of the spacecraft body.

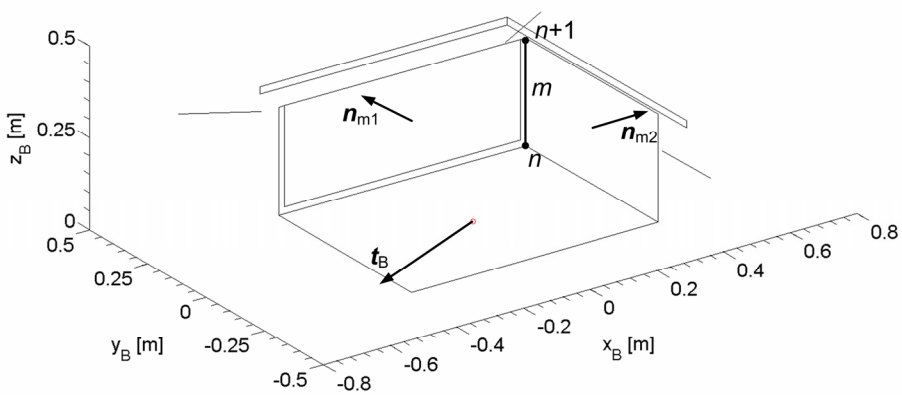

Figure 4 Minimalistic 3D model of Client spacecraft with visibility specification for segment *m* (endpoints *n* and *n*+1) and possible viewpoint t_B.

The spacecraft model must be well organized to boost the efficiency of the search algorithms during matching and pose estimation. To this end, an indexing scheme is introduced which facilitates the finding and manipulation of all segments, including endpoints, orientation, and lengths. In addition a visibility specification is attached to each segment. This is required at various stages of the pose estimation process, including the lost-in-space initialization (Sect. 4) and the extension of the first set of

matches (Sect. 5.2). The visibility is specified by one or two unit vectors normal to the surfaces which contain the considered segment. An example is visualized in Fig. 4, where segment *m* can be viewed from the directions indicated by the unit vectors $\boldsymbol{n}_{m1} = (0,1,0)^t$ and $\boldsymbol{n}_{m2} = (1,0,0)^t$. For simplicity, no partial visibility has been implemented, thus a generic segment *m* is visible under the following algebraic condition

$$t_B \cdot n_{m1} \geq 0 \;\; \| \;\; t_B \cdot n_{m2} \geq 0 \tag{3}$$

As illustrated by Fig.4, the non-rigorous implementation of the visibility constraint makes several segments not visible even if they are partially visible. This is a drawback of the adopted approach in view of the image processing output.

As for the spacecraft model, perceptual groups can be pre-computed once before mission start. In contrast to the image processing output, here exact instances in 3D of the perceptual groups are searched for (i.e., $P = P^{max} = 0$). It is noteworthy that the same perceptual organization functions can be used without changes for image and model processing. Table 1 lists the results of the perceptual organization functions applied to the design case image (Fig. 3), and to the spacecraft model (Fig. 4). The first 4 columns show the organization tree and the size of the available structures, whereas the last column lists the perceptual group rankings or the probability of accidentalness. By definition, this is identically zero for the model groups. The only image segments which are found to be in close proximity are also the most perceptual ones among the available groups (P = 0.01). It is noted that the model groups “Proximity” are less numerous than the groups “Parallel”, “Collinear”, and “III-shaped”, and offer the same amount of measurements (2 endpoints). As a consequence, the usage of these structures can improve the computational efficiency of the search process at no cost. On the other hand, the groups “Trapezoid”, “Parallelogram”, and “Y-shaped” offer 3 to 4 measurements and are of even reduced quantity (13 to 20). These latter’s shall be prioritized in a pose estimation process, but are difficult to extract from space imagery.

Table 1. Perceptual grouping output for design case image and model.

Organization tree			Size	Probability, *P*
Image	Segments	Endpoints	20x2	-
		Length	10x1	
	Groups	Proximity	1x2	0.01
		Parallel	3x2	(0.03,0.03,0.06)
Model	Segments	Endpoints	82x3	
		Length	41x1	-
		Visibility	41x6	
	Groups	Proximity	64x2	
		Parallel	189x2	
		Collinear	235x2	
		Trapezoid	13x4	0.0
		Parallelogram	13x4	
		Y-shaped	20x3	
		III-shaped	626x3	

4 Initial Pose

The estimate of the initial pose is certainly the most challenging task of the pose estimation process. Many authors assume coarse a-priori knowledge of the relative position and orientation to aid the vision navigation system [20], whereas others refer to this task as the weakest and least robust in the estimation chain [11]. Aside the Client spacecraft 3D model, here no assumption on the relative translational and rotational motion is made. The lost-in-space initial pose estimation is solved in two steps and aims at achieving a three-axis attitude error below 30-40° for the subsequent refinement process.

A first guess of the pose follows the approach suggested by [7]. In particular the rotation matrix $\boldsymbol{R}_{BC}$ is computed i) from the most likely viewing direction of the considered perceptual group, and ii) by causing one of the perceptual group model segments to be parallel to the correspondent image segment. The most likely viewing direction is computed from the sum of the unit vectors which specify the visibility of the perceptual group. The translation $\boldsymbol{t}_B$ is obtained i) as a scale factor from the segment pair of the considered perceptual group with minimum ratio model to observed line length, and ii) by aligning the endpoints of model and image segment. Since the correspondence between observed and model segments is unknown, this initial guess has to be computed for all available pairing combinations "model-image segment" of the considered perceptual group. This first step is completed by the computation of the initial pose error as the root-mean-square of the differences between the considered image points and the projected model points.

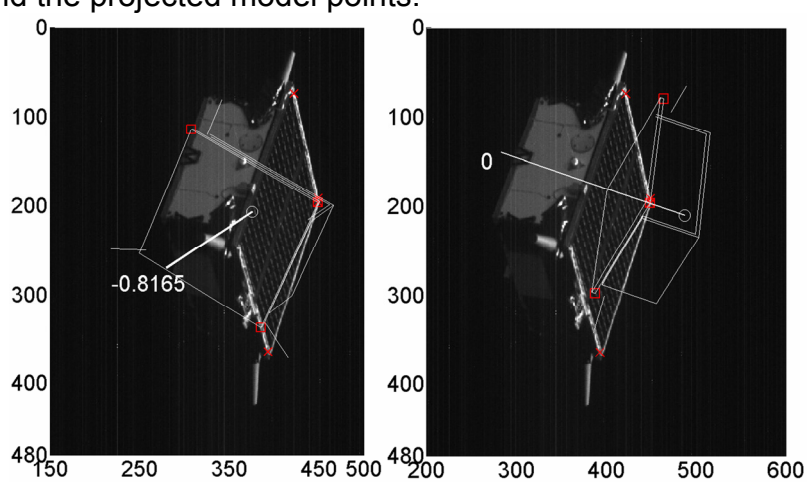

Figure 5 Two out of 32 initial guess pose solutions with total error below 100 pxl (rms). Crosses indicate original measurements. Squares indicate model points.

The application of this procedure to the aforementioned design case provides 32 pose solutions with an error below 100 pixels (rms). All possible pairing combinations provided by the most highly ranked perceptual group (i.e., proximity with P = 0.01) have been considered (i.e., 4x128), cf. Table 1. Fig. 5 illustrates just two of the available pose solutions with fit errors equal to 39.7 (Fig. 5 – right) and 62.3 (Fig. 5 – left) pixels. The vector normal to the projected Client solar panel is also depicted with the indication of its z_C component to improve the readability of the figure. As anticipated this initial

solution is affected by a very large error which prevents any ambiguity resolution or refinement through Eq. (1).

On the contrary this initial guess can be used for a more exhaustive search of the visibility domain. In particular the most likely viewing direction $\boldsymbol{t}_{01}$ is used as a starting point for an efficient grid search whose mechanism is illustrated by Fig. 6. The search is conducted iteratively through cones defined with respect to the viewing direction. These cones are parameterized through their aperture angles, α and β. The chosen search grid and iteration parameters are listed in Table 2. At the first iteration $\alpha = \alpha_1$ is set with respect to $\boldsymbol{t}_{01}$ and the viewing direction $\boldsymbol{t}_{11}$ is searched which minimizes the error E_{11} among all $\beta_k = (\beta_1,\ldots,\beta_p)$. This is repeated using $\boldsymbol{t}_{i1}$ as new viewing direction until the error $E_{i1} \leq E_{i-1,1}$ or until the maximum number of iterations is reached (i.e., $i = n$). The obtained error represents the minimum achievable pose error starting from α_1 as initial search cone. The same procedure can be repeated for different initial cone apertures $\alpha_j = (\alpha_1,\ldots,\alpha_q)$ in order to span the visibility domain as exhaustively as possible. The obtained pose solution is affected by the minimum possible error for all considered α and β.

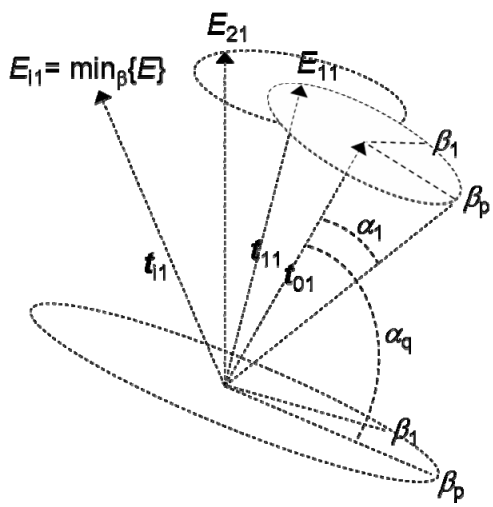

Figure 6 Initial pose search algorithm.

Table 2. Initial pose search grid and iteration parameters

Item	Value
p	36
$\Delta\beta = \beta_k - \beta_{k-1}$	10°
q	5
$\Delta\alpha = \alpha_j - \alpha_{j-1}$	20°
n	4

Representative results of this second step are illustrated in Fig. 7. In particular 4 pose solutions out of the 32 selected at the previous step show an error below 25 pixels (rms). The solutions addressed in Fig. 5 are re-proposed here for better comparison. The projection errors have been reduced to 19.4 (Fig. 7 – left) and 23.5 (Fig. 7 – right) pixels. This has required a rotation as large as α = 100° (i.e., q = 5) in the first case, whereas no rotation was necessary in the second case. Similar results are obtained when processing other images of the Client spacecraft. The resulting initial pose is considered good enough to proceed with a refinement based on Eq. (1). In addition the search space for ambiguity resolution has been reduced to 4 possible pose solutions for the considered perceptual group.

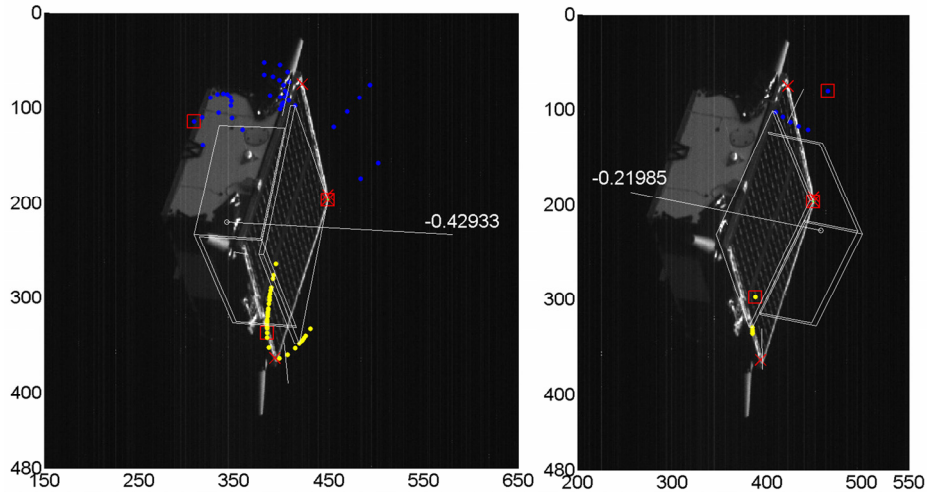

Figure 7 Two out of 4 initial pose solutions with total error below 25 pxl (rms). Crosses indicate original measurements. Squares indicate model points. Blue and yellow dots represent intermediate iterations of search algorithm.

5 Pose Estimation

5.1 Newton-Raphson Refinement

The fit error $\boldsymbol{E}$ between modeled and observed points ρ_{D} can be linearized about the current pose estimate as follows

$$\boldsymbol{E} = \frac{\partial \boldsymbol{\rho}_{\mathrm{D}}}{\partial \boldsymbol{r}_{\mathrm{C}}} \frac{\partial \boldsymbol{r}_{\mathrm{C}}}{\partial \boldsymbol{t}_{\mathrm{B}}} \Delta \boldsymbol{t}_{\mathrm{B}} + \frac{\partial \boldsymbol{\rho}_{\mathrm{D}}}{\partial \boldsymbol{r}_{\mathrm{C}}} \frac{\partial \boldsymbol{r}_{\mathrm{C}}}{\partial \varphi_{\mathrm{BC}}} \Delta \varphi_{\mathrm{BC}} \tag{4}$$

where the partial derivatives with respect to the relative position $\boldsymbol{t}_{\mathrm{B}}$ and to the Euler angles φ_{BC} (rotation matrix $\boldsymbol{R}_{\mathrm{BC}}$) can be simply computed from Eq. (1) as

$$\frac{\partial \boldsymbol{\rho}_{\mathrm{D}}}{\partial \boldsymbol{r}_{\mathrm{C}}} = \begin{pmatrix} f_u/z & 0 & -f_u x/z^2 \\ 0 & f_v/z & -f_v y/z^2 \end{pmatrix}; \quad \frac{\partial \boldsymbol{r}_{\mathrm{C}}}{\partial \boldsymbol{t}_{\mathrm{B}}} = -\boldsymbol{R}_{\mathrm{BC}}; \quad \frac{\partial \boldsymbol{r}_{\mathrm{C}}}{\partial \varphi_{\mathrm{BC}}} = \frac{\partial \boldsymbol{R}_{\mathrm{BC}}}{\partial \varphi_{\mathrm{BC}}}. \tag{5}$$

Each image point provides two error equations [i.e., Eq. (4)] in six unknowns which express small variations $\Delta(\cdot)$ with respect to the reference state. As a consequence a complete linear system of 6 equations and 6 unknowns can be derived from just three point correspondences. This is the amount of observables given by a proximity perceptual group composed of two segments. The unknown corrections to the reference state can be computed through a multi-dimensional Newton-Raphson method [12], provided that the initial pose is affected by an error below about 40°. As shown in Fig. 3, using only point correspondences is weak in the presence of partial edges. In order to cope with uncertain line terminations, the angle θ subtended by the considered segment

can be used as observable as well. Fortunately the partial derivative of θ can be expressed as a linear combination of the partial derivatives of u and v as follows

$$\frac{\partial\theta}{\partial \boldsymbol{r}_{\mathrm{C}}} = \left(\frac{\partial v_2}{\partial \boldsymbol{r}_{\mathrm{C}}} - \frac{\partial v_1}{\partial \boldsymbol{r}_{\mathrm{C}}}\right)\frac{1}{l\cos\theta} = \left(\frac{\partial u_2}{\partial \boldsymbol{r}_{\mathrm{C}}} - \frac{\partial u_1}{\partial \boldsymbol{r}_{\mathrm{C}}}\right)\frac{-1}{l\sin\theta} \tag{6}$$

where 1 and 2 represent the segment start- and end-points. Eq. (6) is derived from the normal form of the equation of a straight line parameterized through its inclination angle θ and its distance from the origin. Independent from the choice of the observables, at each iteration the Newton-Raphson method solves for a vector of six corrections using Eq. (4)

$$\Delta\boldsymbol{x} = \begin{pmatrix}\Delta t_{\mathrm{Bx}} & \Delta t_{\mathrm{By}} & \Delta t_{\mathrm{Bz}} & \Delta\varphi_{\mathrm{BCx}} & \Delta\varphi_{\mathrm{BCy}} & \Delta\varphi_{\mathrm{BCz}}\end{pmatrix}^t \tag{7}$$

and given six computed model-image error components E_{i}. After each iteration, $\Delta\boldsymbol{x}$ shrinks by about one order of magnitude, and no more than four iterations are needed for pixel level accuracy. The powerfulness of this approach is clearly shown by Fig. 8. All four initial pose solutions available from the previous step can be made accurate to below 2.5 pixels (rms). The price to pay is the inversion of a 6x6 Jacobian [composed from Eq. (5)] at each iteration. Here only point correspondences have been used.

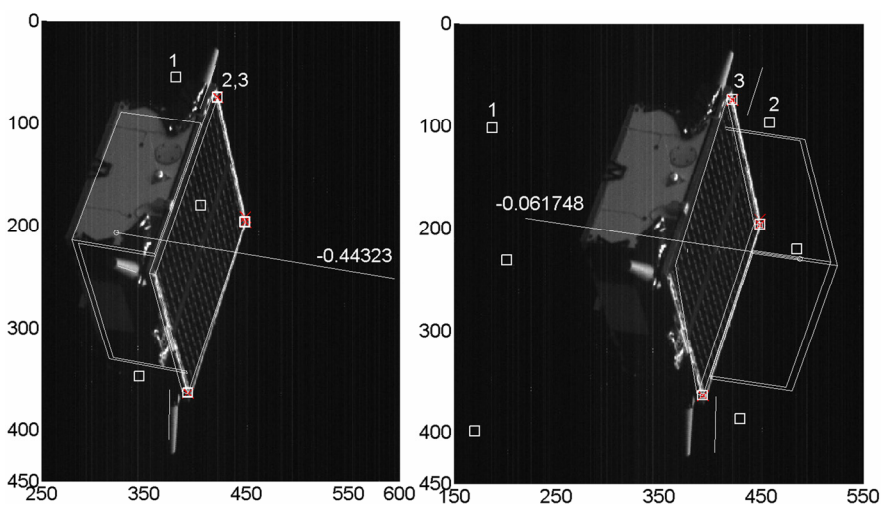

Figure 8 Solutions of Fig. 5 and 7 refined to below 2.5 pxl error (rms). Crosses indicate original measurements. Squares indicate model points upon successive iterations (numbered from 1 to 3).

5.2 Model Matching

Although only based on an initial set of hypothesized matches, the initial pose estimate and its further refinement through the projection equations can be used to predict the locations of other model features in the image and extend the match. The selection of matches in the image which are consistent with predictions from the available pose

follows a probabilistic approach based on the formalism of perceptual organization (see Sect. 3.1). The same perceptual grouping functions first applied to the 3D model (Sect. 3.3) and to the 2D image (Sect. 3.2), can be used to process a new 2D object resulting from the superposition of the detected image segments and the projected 3D model.

Structures equivalent to perceptual groups are formed by comparing the projected segments of the model which are visible with the image segments. Each potential match is assigned a probability of non-accidentalness based on simultaneous instances of proximity and parallelism. In particular, two segments are considered as correct matches if one of the following applies: i) Both endpoints are in proximity (see Fig. 9 - left), ii) One endpoint is in proximity and the segments are parallel with overlapping legs (see Fig. 9 - right). The goodness of match is given by the product of the probabilities of proximity and parallelism as computed from Eq. (2).

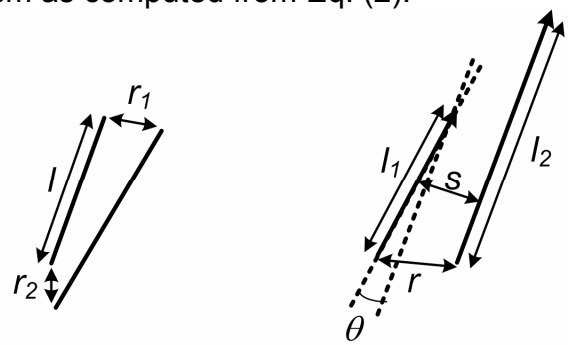

Figure 9 Criteria used for proper model-image matching.

The application of the method described above to the four pose solutions available from the previous step gives the possibility to finally resolve the remaining ambiguity. Only two preliminary solutions give new matches as shown in Fig. 10. The solution with the most highly ranked matches is considered for the final stage (Fig. 10 - left), being the one which best fits the image in a probabilistic sense.

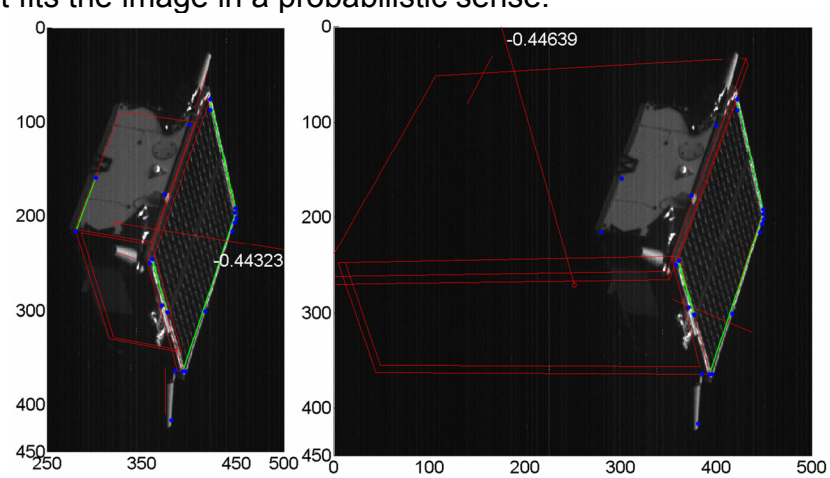

Figure 10 Two solutions which provide new matches (green segments) when projected onto the image plane. Blue dots indicate the original measurements.

It is noted that the alternative solution (Fig. 10 - right) gets high fit scores although obviously wrong to a human eye. This is due to the very limited amount of segments

identified by the image processing. In order to increase the robustness of this ambiguity resolution, an extra image processing stage could be considered which aims at finding specific features of the model in reduced areas of the image according to the available pose solution. In the presence of more image segments, instead of using all gained matches at once in the least-squares fit (described in the next section), one could consider to add these new matches to the best fit solution gradually. By the time a sufficient number of reliable matches have been found, the pose estimation is accurate enough to choose between the few ambiguous alternatives. The set of matches is repeatedly extended in this way until no more can be found. This iterative matching procedure has the appealing property of using the easy cases to provide better pose estimates to disambiguate the more difficult situations.

5.3 Least-Squares Fit

After the matching step, more correspondences between model and image are available than strictly necessary (i.e., $\geq$ 2 segments). The pose estimation problem is now over-determined with more equations (measurements) than unknowns (pose state). This can be used to compensate for measurements and model errors or to estimate extra parameters (e.g., camera parameters). All available measurements are processed in a weighted iterative batch least-squares estimator with a-priori information [21]. For each matched segment, 3 measurements are derived, namely the coordinates, (u,v), of the endpoint in closest proximity and the angle, θ, subtended by the line segment with the u-axis. The sensitivity matrix is computed using the same equations as for the Newton-Raphson method [i.e., Eq. (4-6)]. The a-priori state is inhereted from the initial pose estimate (refined through Newton-Raphson). The a-priori state covariance matrix is diagonal with 1 m and 10° standard deviation for the components of $\boldsymbol{t}_B$ and $\boldsymbol{\varphi}_{BC}$ respectively. A process noise of 10 cm and 1° standard deviation is injected to keep the filter recipient to measurements at each further iteration. The measurement noise is assumed Gaussian and uncorrelated with standard deviations (du,dv) which amount to the different pixel size for (u,v) and du/l for θ. The quality of the estimated state is judged based on the total rms of the post-fit residuals and the formal output standard deviation of the state parameters. After 3-4 iterations no further improvement of the obtained solution is observable.

For the solution under consideration, a total of 4 segments are used providing 12 measurements. The post-fit residuals amount to 0.57 and 1.12 pixels for (u,v), and 0.33° for θ, whereas the formal state standard deviation is (3.9,4.7, 3.3) cm for $\boldsymbol{t}_B$, and (0.40 0.20 0.32)° for $\boldsymbol{\varphi}_{BC}$ after 4 iterations. A closer look at the post-fit residuals shows that the error decreases by 50% in u and θ, whereas it remains constant in v (from first to last iteration). Although the error trend can be largely improved by correcting for the offsets which affect most detected edges (see Fig. 3), it was not possible to remove the bias of 1 pixel in the available v-axis measurements. It is believed that such an error could be absorbed through the identification, matching, and processing of image segments mostly aligned with the u-axis. The availability of more matched segments could greatly improved the standard deviation of the pose error.

In order to get an independent assessment of the pose estimation accuracy, flight dynamics products from the PRISMA mission have been used for comparison, namely the post-facto on-ground precise relative orbit determination based on GPS (accurate to about 2 cm rms) [14], and the on-board coarse Client attitude estimate from sun-sensors and magnetometers (accurate to about 3° rms). As compared with these operational products, the pose estimate is affected by errors of (14.0 -12.7 -6.6) cm and (3.8 -10.6 2.9) ° along and around the camera axis respectively. As expected from the post-fit residuals the largest error components reflect an uncertainty due to rotations around the *u*-axis. This is due to the particular viewing geometry where the Client spacecraft is mostly aligned with the *v*-axis. The obtained results are very promising since the pose estimation error lies within the typical error budget of the applied reference products. It is also noted that the error computation is affected by uncertainties in the location and orientation of the camera. Furthermore other authors have obtained similar errors at about 10% of the inter-spacecraft separation in more optimistic conditions based on pure software simulations [20].

6 Discussion and Conclusions

This paper has described the foundations of a spaceborne model-based pose estimation methodology whose overall layout is depicted in Fig. 11. The detailed description of the applied algorithms has been accompanied by a representative design case to show the functional applicability and the accuracy potential of the developed approach. The architecture of the vision-based navigation system comprises six main functions (cf., Fig. 11), namely image processing (i.e., Canny edge detector and Hough transform, Sect. 3.2), perceptual grouping (applied to model, image and during matching process, Sect. 3.1), initial pose (lost-in-space, Sect. 4), Netwon-Raphson refinement (based on the true perspective projection equations, Sect. 5.1), model matching (based on the projection of the visible model onto the image plane, Sect. 5.2), and the final pose (least-squares fit, Sect. 5.3). The resulting multi-stage estimator is capable of delivering intermediate pose solutions of increasing accuracy (post-fit rms residuals from 10^1 to below 10^0 pixels) and decreasing ambiguity level (possible number of pose solutions from 10^2 to 10^0) throughout the process. Comparisons with independent flight dynamics operational products have shown pose accuracies at the level of 10 cm and 10 °.

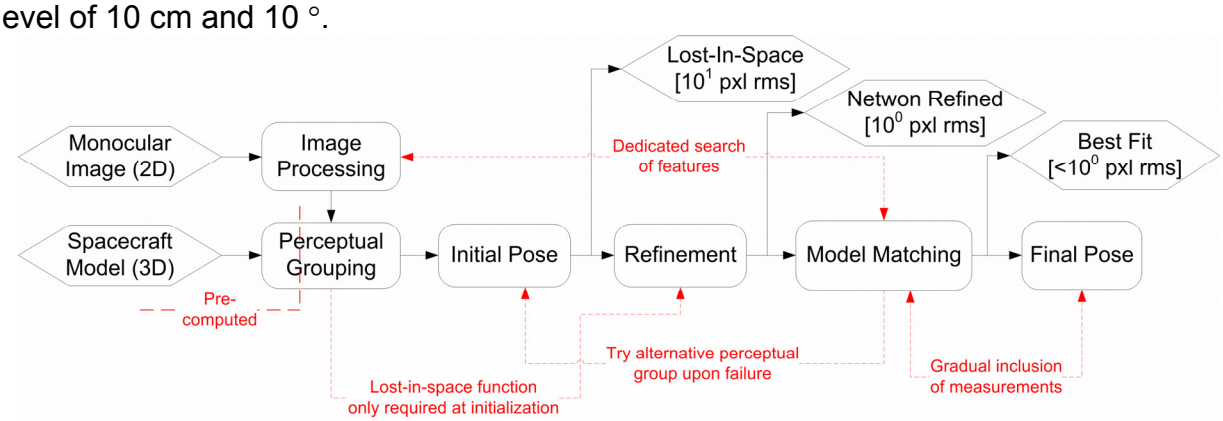

Figure 11. Architecture of pose estimation method, including data flow (arrows), inputs/outputs (hexagonal blocks), and key functions (rounded rectangles).

Despite the promising results, several aspects of the navigation system shall be further analyzed and improved. The major limiting factor for accuracy and reliability has been identified in the image processing module. Aside its overall improvement, efforts shall be put into a dedicated image processing step after the matching process to identify spacecraft features in limited portions of the image (where they are supposed to be). In order to gain new relevant features, the observation model based on line-to-line correspondences (Sect. 5.1) could be extended to match arbitrary curves rather than only straight line segments. In addition the processing and comparison of subsequent images could be exploited. In this case the initial pose estimate does not need to be repeated, provided that the Client spacecraft relative attitude does not change by more than 30-40° between two images. The interplay between model matching and final pose estimation is also considered a major area of research. The gradual inclusion of new measurements and the data editing process need to be studied further. Biased edges have been found to largely affect the performance of the least-squares fit. It has to be understood how such uncertainties in the measurements or model errors could be properly incorporated in the filter scheme.

7. Acknowledgements

This work has been conducted in the frame of the sabbatical program of DLR (Forschungssemester). The authors would like to thank Jan Kleberg (DTU) for the provision of the Tango spacecraft 3D model. Space images used in this work have been collected during a GPS-based flight experiment defined by GMV/ESA and operated by OHB Sweden in the frame of the PRISMA extended mission.

8. References

[1] Bodin, P., Noteborn, R., Larsson, R., Karlsson, T., D'Amico, S., Ardaens, J.-S., Delpech, M., Berges, J.-C., "Prisma Formation Flying Demonstrator: Overview and Conclusions from the Nominal Mission," AAS 12-072; 35th Annual AAS Guidance and Control Conference, 3-8 Feb. 2012, Breckenridge, Colorado, 2012.

[2] Rupp, T., Boge, T., Kiehling, R., Sellmaier, F., "Flight Dynamics Challenges of the German On-Orbit Servicing Mission DEOS," 21st International Symposium on Space Flight Dynamics (ISSFD), 28 Sep. - 2 Oct. 2009, Toulouse, France, 2009.

[3] Landgraf, M., Mestreau-Garreau, A., "Formation Flying and Mission Design for Proba-3", Acta Astronautica, Volume 82, Issue 1, pp 137–145, January, 2013.

[4] Jørgensen, J.L., Denver, T., Jørgensen, P.S., "Using an Autonomous Star Tracker as Formation Flying Sensor", 4th Symposium on Small Satellites Systems and Services, Sept. 20-24, La Rochelle, France, 2004.

[5] Ho, C.J. and McClamroch, N.H., "Autonomous Spacecraft Docking using a Computer Vision System," in Proc. 31st Conf. on Decision and Control, Tucson, Arizona, pages 645–650, December 1992.

[6] Philip, N.K., Kumar, E.K., and Ananthasayanam, M.R., "Sliding Observer for a Robust Relative Position and Attitude Estimation During the Final Phase of an

Autonomous Docking Mission," in Proc. 49th Int'l Astronautical Congress, Melbourne, Australia, October 1998.

[7] Lowe, D.G., "Three Dimensional Object Recognition from Single Two Dimensional Images", Artificial Intelligence, Vol 31, pp 355-395, 1985.

[8] Fischler, M.A., and Bolles, R.C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM, 24(6):381–395, June 1981.

[9] DeMenthon, D.F., Davis, L.S., "Model-Based Object Pose in 25 Lines of Code," International Journal of Computer Vision, Volume 15, Issue 1-2, pp 123-141, June 1995.

[10] Lowe, D.G., "Perceptual Organization and Visual Recognition," MA, USA: Kluwer Academic Publishers, 1985.

[11] McIvor, A.M., "An Analysis of Lowe's Model-Based Vision System," Proc. Fourth Alvey Vision Conference, Univ. of Manchester, pp 73–78, August 1988.

[12] Ostrowski, A.M., "Solutions of Equations and Systems of Equations," 2nd ed. New York:Academic Press, 1966.

[13] Torr, P., and Zisserman, A., "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry," Computer Vision and Image Understanding, 78(1):138–156, 2000.

[14] Ardaens, J.-S., D'Amico, S., Montenbruck, O., "Final Commissioning of the PRISMA GPS Navigation System," 22nd International Symposium on Spaceflight Dynamics (ISSFD), Sao Jose dos Campos, Brazil, 28 Feb. - 4 March 2011.

[15] D'Amico, S., Ardaens, J.-S., Larsson, R., "Spaceborne Autonomous Formation-Flying Experiment on the PRISMA Mission," Journal of Guidance, Control, and Dynamics, 0731-5090, vol.35 no. 3 (834-850), doi: 10.2514/1.55638, May-June 2012.

[16] http://www.opencv.org

[17] Canny, J., "A Computational Approach to Edge Detection", IEEE Trans Pattern Recognition and Machine Intelligence, Vol PAMI-8, No. 6, pp 679-698, November 1986.

[18] Duda, R.O., Hart, P.E., "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Communication of the ACM, 15, 1, January 1972.

[19] Benn, M., Denver, T., "Advanced Stellar Compass - Lens Distortion Correction Method for the µASC," ASC-DTU-MA-3013, Issue 1.0, March 19, 2012.

[20] Cropp, A., "Pose Estimation and Relative Orbit Determination of a Nearby Target Microsatellite using Passive Imagery," PhD Thesis, University of Surrey, UK, 2001.

[21] Montenbruck, O., and Gill, E., "Satellites Orbits – Models, Methods, and Applications," Springer Verlag, Heidelberg, Germany, 2001a.