IAC–19–6.6.7

RemoveDebris Vision-Based Navigation preliminary results

Eric Marchand
Univ Rennes, Inria, CNRS, IRISA, Rennes, France

François Chaumette
Inria, Univ Rennes, CNRS, IRISA, Rennes, France

Thomas Chabot
Airbus Defense and Space, France

Keyvan Kanani
Airbus Defense and Space, France

Alexandre Polinni
CSEM, Switzerland

The EC FP7 RemoveDebris is a mission whose goal is to demonstrate Active Debris Removal (ADR) key technologies. Among these technologies a vision-based navigation (VBN) experiment was carried out in October 2018. VBN is considered as a tool to observe and quantify the relative dynamics between an uncooperative debris and the platform preparing for its capture and retrieval. To test VBN, the RemoveDebris spacecraft released a CubeSat. The VBN experiment goal was to localize the CubeSat. The experiment validated the vision-based navigation algorithm (along with VBN hardware) validated through ground-based processing of actual images acquired in flight.
keywords: VBN, RemoveDebris, Visual tracking, Detection, Localization,

1. INTRODUCTION

Since the beginning of the space era in 1957, the number of missions has incredibly increased. Unfortunately, it has also progressively generated a huge amount of debris that are still orbiting Earth. Today, space agencies feel more and more concerned about Earth orbital environment cleaning since they realize those debris represent a big threat due to the high risk of collision that could, among safety issues, damage operational satellites [3]. To investigate solutions to this issue, various projects aim to actively remove debris objects by catching them. Among these projects, is the RemoveDebris project co-funded by the European Union and members of European space industry. In a first stage, the RemoveDebris mission aimed at catching two miniaturized satellites (CubeSats DSat-1 and DSat-2 produced by the Surrey Space Centre) previously ejected from a microsatellite called RemoveSat, itself released from the International Space Station (ISS). The CubeSats were used as targets instead of actual debris. This debris removal demonstration included net capture, harpoon capture and vision-based navigation using standard camera and LiDAR [9, 1].

In this paper we describe the results of the Vision-based navigation experiments. Along with the development of VBN Harware [5, 9], it consists in tracking and localizing the DSat-2 using two model-based tracking algorithm provided by Inria and Airbus, and using the Lidar provided by CSEM.

In the literature, several works propose approaches based on active sensors, such as LiDAR, for space rendezvous and docking missions [10]. Those approaches have proved to be efficient in situations where lighting conditions can be tricky due to directional sunlight resulting in high specularities and shadows. So far, LiDAR were known to be generally expensive and rather heavy sensors. Lidar has a limited range and accuracy, and it cannot take the advantage of using potential texture information available on the target surface in order to improve the navigation process. Based on the knowledge of the 3D model of a target and from the images acquired by the camera, estimating the pose (i.e., the 3D position and orientation) between the camera and the target has always been an ongoing issue in computer vision and robotics [14]. For instance, regarding space applications, [15, 17] use model-based tracking approaches for space rendezvous with space target or debris. More generally, common approaches address this problem by using either texture [2], edge features [8, 7, 17, 15] or color and

intensity features [16, 20, 18].

In this paper, we focus on the vision-based navigation part of the RemoveDEBRIS mission using a single standard B&W or RGB camera on one hand, and a LiDAR on the other hand.

The paper is organized as follow: in a first stage, the model-based tracking methods developed by Inria and Airbus are recalled. Then, we shortly give an overview of the Lidar experiment. Finally, results are presented, first on synthetic image sequences to assess the proposed methods, then on the actual data acquired during the VBN experiment achieved in October 2018. A comparison of these results is then proposed.

2. 3D MODEL-BASED TRACKING

Within the Removedebris projet, two localization algorithms have been proposed by Inria and by Airbus respectively. Both are model-based methods, as described in the next subsections and have been tested on various space robotics scenarii. They consist in determining the pose that minimizes the error between observations in the image and the forward-projection of the model of the target in this image (see Fig. 1 for an overview of the localization process). The applications are miscellaneous, from debris removal to in-orbit servicing or air-to-air refueling (see [18, 11]). Airbus and Inria?s approaches are very similar. The main differences lay in their implementation. Indeed, Airbus version of the tracking is compatible with space missions (HW limitations, trajectories, background,...). Furthermore, Airbus method also features successive detections of the target.

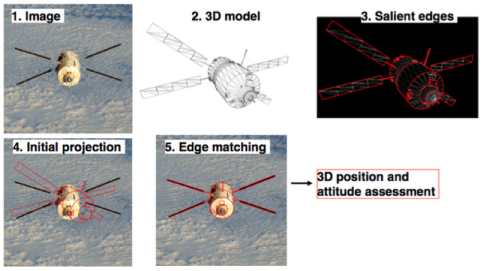


Fig. 1: Overview of the localization process

2.1 *Model-based tracking: a local non-linear optimization problem*

The goal of model-based tracking is to determine an estimate $\widehat{\mathbf{r}}$ of the camera pose $\mathbf{r}, \mathbf{r} \in SE(3)$, by minimizing, with respect to $\mathbf{r}$, the forward projection error $\Delta(\mathbf{r})$:

$$\widehat{\mathbf{r}} = arg\min_{\mathbf{r}} \Delta(\mathbf{r}) \quad [1]$$

with

$$\Delta(\mathbf{r}) = \sum_i \rho(e_i(\mathbf{r})) \quad [2]$$

$\Delta(\mathbf{r})$ accounts for errors $e_i(\mathbf{r})$ between a set of visual features extracted from the image and the forward projection of their 3D homologues in the image plane according to the pose. ρ is a robust estimator [13], which reduces the sensitivity to outliers. This is a non-linear minimization problem with respect to the pose parameters $\mathbf{r}$, and we follow a Gauss-Newton minimization framework to handle it, by iteratively updating the pose $\mathbf{r}$.

At each iteration k, a displacement is performed in the parameter space which is $SE(3)$, the displacement $\delta\mathbf{r}$ is computed through:

$$\delta\mathbf{r} = -\lambda(\mathbf{DJ})^{+}\mathbf{De} \quad [3]$$

$$\text{with } \mathbf{J} = \frac{\partial \mathbf{e}(\mathbf{r})}{\partial \mathbf{r}} \quad [4]$$

and where $(\mathbf{DJ})^{+}$ is the pseudo inverse of $\mathbf{DJ}$, $\mathbf{J}$ being the Jacobian matrix of the error vector $\mathbf{e}(\mathbf{r})$ with respect to the pose. λ is a proportional gain and $\mathbf{D}$ is a weighting matrix associated to the Tukey robust estimator. The pose $\mathbf{r}$, represented by its homogeneous matrix $\mathbf{M}_{k+1}$, can be updated as follows, using the exponential map [12]:

$$\mathbf{M}_{k+1} = exp([\delta\mathbf{r}])\ \mathbf{M}_k. \quad [5]$$

where $exp([\delta\mathbf{r}])$ is the exponential map enabling to express the rigid motion $\delta\mathbf{M}$ generated by $\delta\mathbf{r}$.

Cost function. For a better accuracy and robustness, we have chosen to integrate, within this framework, geometrical edge-based features with geometrical features: keypoints extracted from the image and tracked through the KLT algorithm [21]. The objective is to fuse in the criterion Δ to be optimized a geometrical information provided by distances between edges with distances between keypoints. The goal is to benefit from the complementarity of these features and to overcome the limitations of classical single cue approaches.

By combining these features, Δ can be rewritten as:

$$\Delta = w^g\Delta^g + w^p\Delta^p \quad [6]$$

Δ^g refers to the geometrical edge-based error function and Δ^p their corresponds to the keypoint feature. w^g and w^p are the respective weighting parameters.

2.2 *Edge-based visual features*

From the knowledge on the 3D model of the target, a potential option is to use its silhouette to define features to rely on. Thus and as defined in (5), the first features we consider for this method are edge-based features. As in

[3, 4], the edges that are considered correspond to the projection of the CAD model of the target, using the current camera pose $\mathbf{r}$. More precisely, the edges corresponding to the visible polygons of the model are considered in the minimization process. That is why this approach requires a simplified version of the 3D model that only contains polygons with visible edges (in term of gradient in the image).

The edge-based function Δ^g is computed in a similar way to [18]. From the model edge points we perform a 1D search (see Fig. 8) along the normal of the underlying edge of each $\mathbf{x}_i(\mathbf{r}_k)$. A common approach is to choose on the scan line the pixel with the maximum gradient as the matching edge point $\mathbf{x}'_i$ in the new image.

Once correspondences between the set of control points $\{\mathbf{x}_i\}_{i=1}^{N_g}$ and the set of image edge points $\{\mathbf{x}'_i\}_{i=1}^{N_g}$ are established, our approach considers the distance between the projected 3D line $l_i(\mathbf{r})$ underlying the projected control point $\mathbf{x}_i(\mathbf{r})$ (projected from the 3D point $\mathbf{X}_i$) and the selected matching point $\mathbf{x}'_i$ in the image. The error function $\Delta^g(\mathbf{r})$ to be minimized with respect to the pose $\mathbf{r}$ can be written as:

$$\Delta^g(\mathbf{r}) = \frac{1}{N_g} \sum_i \rho^g(e_i^g(\mathbf{r})) \qquad [7]$$

where

$$e_i^g(\mathbf{r}) = d_\perp(l_i(\mathbf{r}), \mathbf{x}'_i))$$

and $d_\perp(l_i(\mathbf{r}), \mathbf{x}'_i))$ is the distance between the point $\mathbf{x}'_i$ and the corresponding line $l_i(\mathbf{r})$. ρ^g is a Tukey robust estimator.

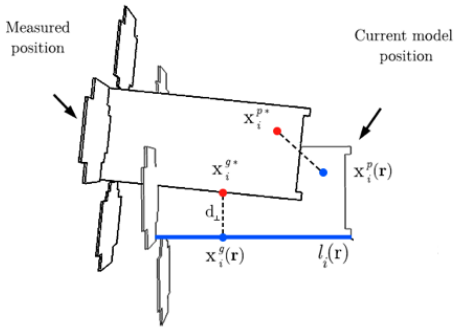


Fig. 2: Model-based tracking principle considering the case of edge and keypoint features

2.3 *Geometrical keypoint based features*

Another class of visual features which can be used are keypoints tracked across the image sequence. As previously suggested by [4] or by [22, 19] within their hybrid approaches, the idea is to design a texture-based objective function Δ^p accounting for geometrical distances between keypoints extracted and tracked over successive images. But in contrast to [19], which process 2D-2D point correspondences to estimate the 2D transformation from $\mathbf{I}_k$ to $\mathbf{I}_{k+1}$ of planar local regions underlying the points, we use 2D-3D correspondences to directly minimize Δ^p w.r.t. the pose $\mathbf{r}$.

More specifically, let us denote $\{\mathbf{x}_i\}_{i=1}^{N_p}$ a set of keypoints detected in frame $\mathbf{I}_k$. Assuming the pose $\mathbf{r}_k$ has been properly estimated, we can restrict these points to be lying on the projected 3D model with respect to $\mathbf{r}_k$. Since we rely on a complete 3D model, the depth of the points in the scene can be accurately retrieved, and using $\mathbf{r}_k$, we can back-project these points on the 3D model, giving a set $\{\mathbf{X}_i\}_{i=1}^{N_p}$ of 3D points of the 3D model.

We employ the Harris corners detector inside the silhouette of the projected model in the image to extract $\{\mathbf{x}_i\}_{i=1}^{N_p}$ in $\mathbf{I}_k$. Then, the KLT tracking algorithm, for which Harris corners are optimal [21], enables to track this set of points in frame $\mathbf{I}_{k+1}$, resulting in a corresponding set $\{\mathbf{x}'_i\}_{i=1}^{N_p}$ (see Fig. 3, right image, red and blue (tracked) dots).

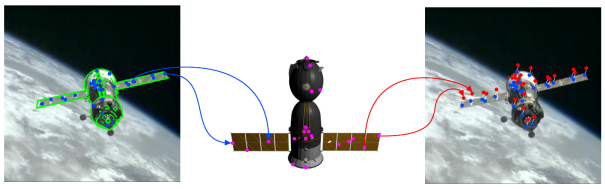


Fig. 3: Determination of correspondences between tracked keypoints and their 3D homologues. On the left are represented Harris points detected on a frame (blue dots). These keypoints are back-projected on the 3D model (pink dots) using the pose computed for the frame. The keypoints are then tracked on a new frame (right image, blue dots) for determining the new pose.

From the correspondences between $\{\mathbf{X}_i\}_{i=1}^{N_p}$ and $\{\mathbf{x}'_i\}_{i=1}^{N_p}$, Δ^p can be computed as follows:

$$\Delta^p(\mathbf{r}) = \frac{1}{N_p} \sum_i^{N_p} \rho^p(e_i^p) \qquad [8]$$

with $e_i^p = (\mathbf{x}_i(\mathbf{r}) - \mathbf{x}'_i)$ and $\mathbf{x}_i(\mathbf{r}) = pr(\mathbf{X}_i, \mathbf{r}))$. ρ^p is the Tukey robust estimator associated to these errors.

2.4 *Target detection*

Target pose estimation is divided into two phases: detection of the target and tracking. Detection enables to find the target in the image, without any initialization (or

a very coarse one) and to assess the target pose. It is generally used to initialize the tracking, which is faster and more accurate. Tracking, on its side, assesses the pose of the target from one image to the other. It generally requires a small motion of the target between two consecutive images.

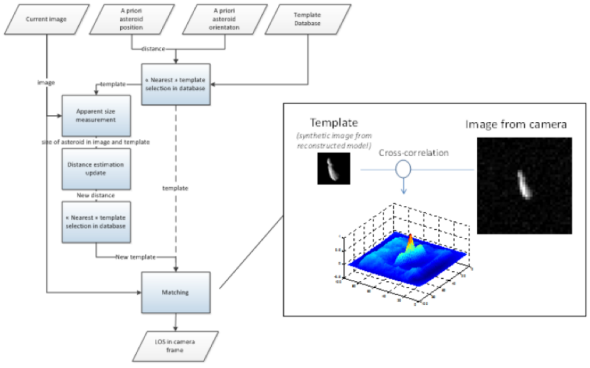


Fig. 4: Target detection with template matching approach

Detection principle is different. It consists in building a database of elements describing the target viewed with various positions and attitudes, and to find the element closest to the current acquired image using correlation criterion.

The main difficulty of detection is to find the descriptor which limits the size and the number of elements in the database, while covering all the possible expected poses of the target. For RemoveDebris mission, we have chosen to use a template matching technique previously used for asteroid pose estimation (see [6]). This solution enables to estimate the pose of the target for a various set of attitudes, knowing the sun phase angle, and with a removed background. It has been applied to RemoveDebris RGB camera images, which could not be successfully processed by tracking given their low acquisition rate (0.1Hz).

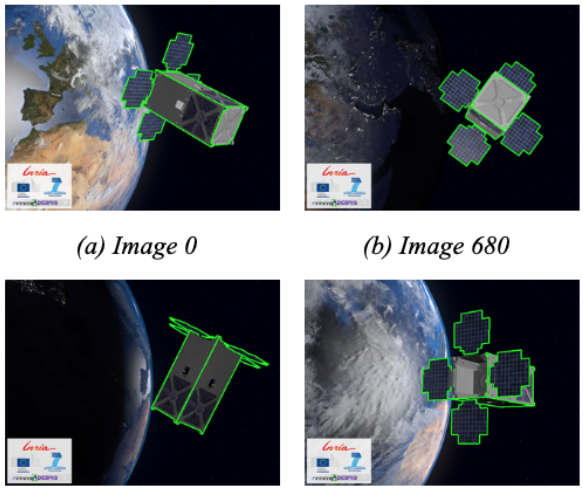
3. REMOVEDEBRIS LIDAR

LiDAR is considered as an enabling technology for missions with rendezvous phases. The LiDAR architecture used for RemoveDebris is based on the flash principle where the object or the scene of interest is illuminated at once without any form of scanning. There is a long term strategy to develop this type of LiDAR for space applications to facilitate not only rendezvous but also rover navigation and precise landing.

While today the keyword LiDAR is known by many people due to its use in sensor solutions for level 3, 4 and 5 autonomous vehicles, it was not like that at the time of the project elaboration. A limited number of commercial critical components (e.g. VCSEL laser diodes array and time-of-flight detector) existed and no qualified ones were available. Unconventionally, the hardware TRL was at 2 at the start of the project (paper concept) and at the end it achieved TRL8, 4 years later with the in-orbit demonstration. This fast development pace was only possible by using a "cubesat" philosophy consisting in using commercial off-the-shelf components and assessing them for low earth orbit operation. The overall development budget is estimated to have been divided by 10-20 and the cost of one unit is around 20-40% of the cost of LiDAR developed according to the traditional stepwise space industry approach.

4. VBN EXPERIMENT USING SIMULATED DATA

To assess the VBN experiment and prepare the actual in-orbit mission we first considered synthetic images. In the sequence depicted on Fig. 5, the considered target, the foreseen RemoveDebris 2U Cubesat, is on a geosynchronous orbit. It has an angular speed that creates rotational motions in the camera frame. However, the distance to the camera remains the same over the sequence. We have also considered different lighting conditions to simulate day and night (where the background also changes) in order to prove the robustness of our methods whatever the situation (see Fig. 5). The full video result is available at: https://youtu.be/Gr78pmRPKOA.



(a) Image 0 *(b) Image 680*

(c) Image 2145 *(d) Image 2602*

Fig. 5: Vision-based navigation of a CubeSat DS-2 for the first sequence of a geosynchronous orbit, using the CPU-based approach.

The proposed method has been able to perfectly track the target. The estimated trajectory is accurate and smooth

(see Fig. 7). Furthermore, according to the collected ground truth data, the vision-based navigation provides precise pose estimation with an RMSE that is less than 1cm on the translation and about 2deg on the rotation. Finally, on Fig. 6, we can see the norm of the translational and rotational errors to the ground truth, which also shows where the algorithms encounter difficulties. Around frame 2200 for example, the precision of the estimated pose decreases a bit. The reason is that the panels of the target are orthogonally aligned with the focal axis of the camera (see Fig. 5.c), which results in increasing the incertitude on the target orientation.

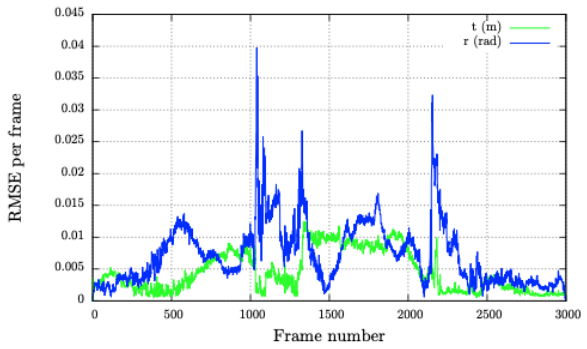


Fig. 6: Error to ground truth.

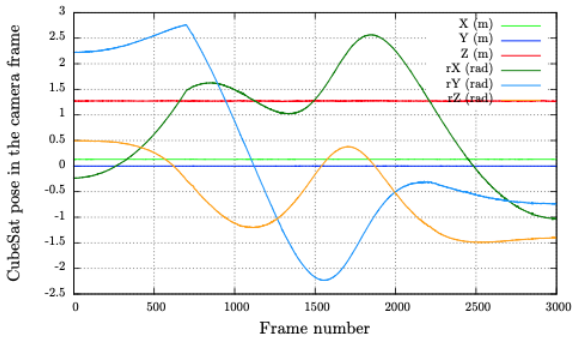


Fig. 7: 3D poses estimated for the first sequence of a geosynchronous orbit (translation in meters and rotation in radians).

5. VBN EXPERIMENT FROM ON-ORBIT DATA

The vision-based navigation (VBN) experiment was carried out in October 2018. The VBN hardware and software is considered as a tool to observe and quantify relative dynamics between an uncooperative debris and the platform preparing for its retrieval. To test the VBN, the RemoveDebris spacecraft released a CubeSat. The experiment goal was to localize a 2U CubeSat using optical images acquired by cameras and a LiDAR on-board the spacecraft.

5.1 *RemoveDebris VBN Sensors*

The VBN is made of 2 vision-based sensors: a color camera and a flash-imaging LIDAR (see Fig. 8). It weights 1.5 kg and is 10x10x15 centimeters large. The device includes a laser source that lights up the object in the field-of-view (FoV). The FoVs are respectively 8^ox6^o for the LiDAR and 21^ox15^o for the color camera. 3D (160x120 pixels), color (2048x1536 pixels) and B&W (1280x1024 pixels, 12mm lens with 12mm square pixels) images were acquired at respectively 0.1, 0.08, and 1Hz, for the B&W camera with successively 3 different aperture times, meaning that successive images at same aperture were acquired at 0.3Hz. The VBN captures images independently after a start trigger from the satellite computer according to pre-programmed sequences defining the respective series of sampling times and integration times for the LiDAR and the camera. Both 3D and color images are stored in non-volatile flash memory. There are downloaded by the satellite computer whenever possible and forwarded down to Earth.

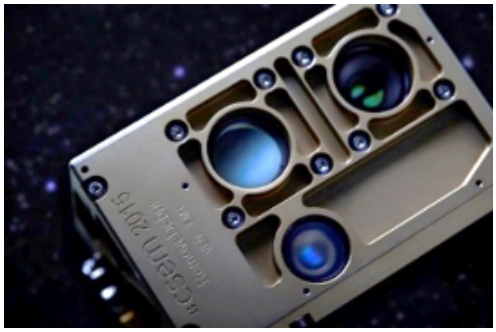


Fig. 8: Vision-based sensor (VBN). Top right: color camera optics. Bottom center: laser illumination head and above LiDAR receiver optics.

The VBN generates 3 images: standard RAW or JPEG color images with the camera, and gray-level and depth map images with the LiDAR. The images on Fig. 9 were captured during the on-orbit calibration phase above Australia.

5.2 *In-orbit operations overview*

The first fact to be highlighted is that DSAT2 remained in the cameras field of view over the entire sequence, meaning that the attitude guidance profile was correctly generated and followed by RemoveSat, and that disper-

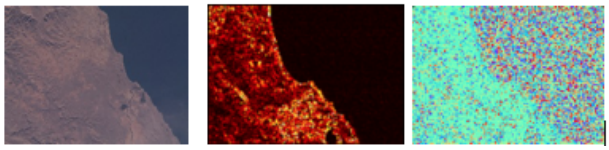


Fig. 9: from left to right: color camera JPEG image, gray-level and depth map/time-of-flight LiDAR images.

sions at ejection remained at an acceptable level. This is illustrated on Fig. 10.

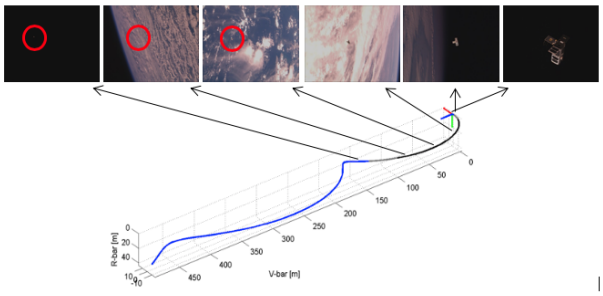


Fig. 10: DSAT2 within VBN camera field of view over the reference trajectory

The planned acquisition timeline was perfectly followed and images were acquired as planned. A late decision to increase SV camera acquisition frequency allowed the use of 3 different integration times, maximising the likelihood of correct exposure all along the demonstration.

Different VBN algorithms provide 3-DoF (position only) and 6-DoF (position and attitude) relative navigation depending on the rendezvous phase:

- At far and medium range, only 3-DoF (position only) relative navigation is needed and performed
- At short range, 6-DoF (position and attitude) relative navigation is required for successful capture. In this paper we only focused on these approaches.

Examples of image processing algorithms and corresponding transitions are illustrated on Fig. 11.

5.3 *Tracking results*

Both Airbus and Inria tracking solutions rely on a 3D model of the DSAT2 target (see Fig. 12).

5.3.1 *Results on the actual RemoveDebris data using Inria algorithm*

The experiment validated the VBN algorithm (along with the hardware) through ground-based processing of

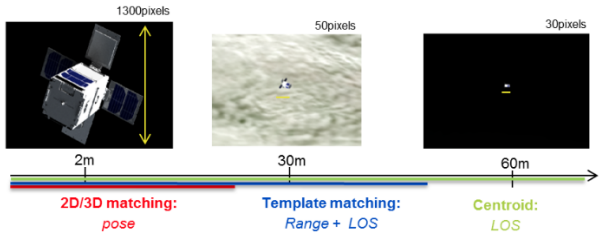


Fig. 11: Different image processing algorithms are used depending on relative distance and navigation needs

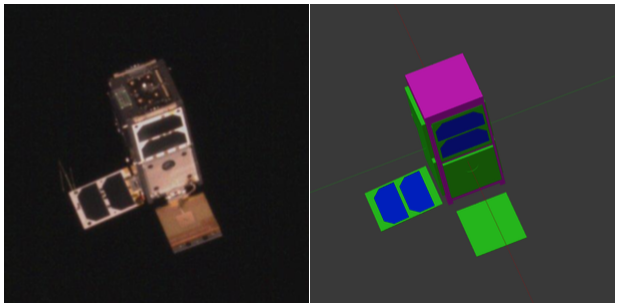


Fig. 12: DSAT2 in orbit image and 3D model

the actual images acquired in flight by the supervision camera. Images (1280x1024 pixels, 12mm lens with 12mm square pixels) were acquired at 3 Hz with successively 3 different aperture times, meaning that successive images at same aperture were acquired at 1 Hz. In particular, the CubeSat has been successively tracked and localized (see Fig. 13 and Fig. 14) on 210 successive images (that is, 210 s or 3.5 minutes). Note that the time processing of each frame is less than 100 ms. After the DSAT2 target is to small in the image to be localized (position and orientation) by the model-based tracking approach.

5.3.2 *Results on the actual RemoveDebris data using Airbus algorithm*

The Airbus in-house tracking solution has been tested on the VBN camera images. This solution had to be improved in order to deal with the low acquisition frequency of the VBN camera (0.1 Hz), showing good results as long as the background is dark as illustrated on Fig. 15. The green contours illustrate the 3D model of the target which has been matched by the algorithm to the actual DSAT2 image.

5.3.3 *Results of the localization using the Lidar*

After several trade-offs mostly between satellite safety and experiments operation conditions requirements, in-orbit experiments time windows were chosen to favor a strong illumination of the target by the Sun. These opera-

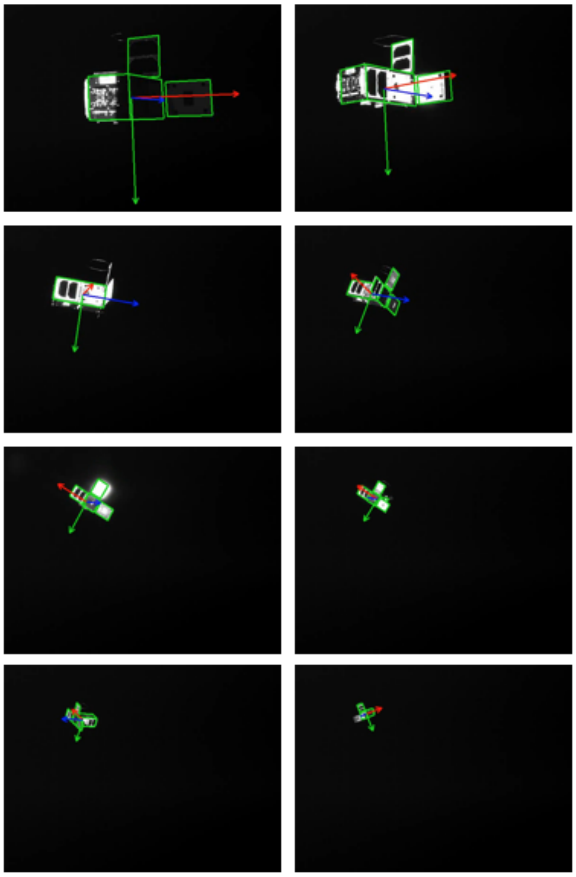


Fig. 13: Vision-nased navigation of the Cubesat for the actual Remodebris sequence. The Cubesat has been successively tracked and localized on 210 successive images (that is during 201s, ie, 3.5 minutes)

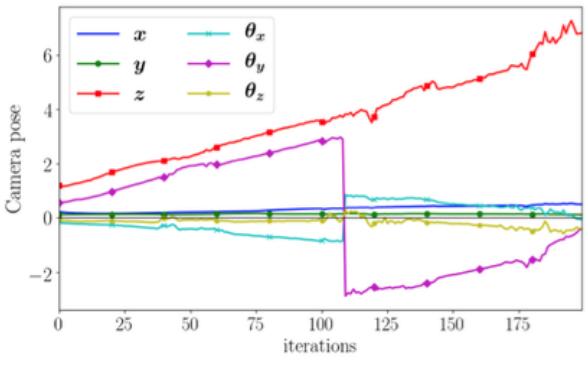


Fig. 14: Successive 3D pose estimated for the actual Remodebris sequence (translation in meters and rotation in radians). One iteration correspond to one second.

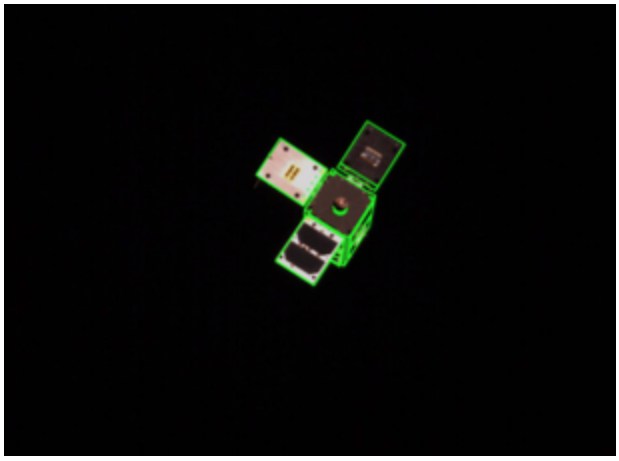


Fig. 15: Example of tracking performed during the VBN demonstration using the Airbus tracking approach.

tion conditions are the worst case for a sensor with active illumination capability. Fig. 16 illustrate the most significative added values of LiDAR measurements in the RemoveDebris context. The sensor settings defining the image dynamic range (aperture, integration time) were defined on Earth, several days before each experiment. The first illustration are images taken respectively by the color camera and the LiDAR a few tenth of seconds before the cubesat capture by the NET.

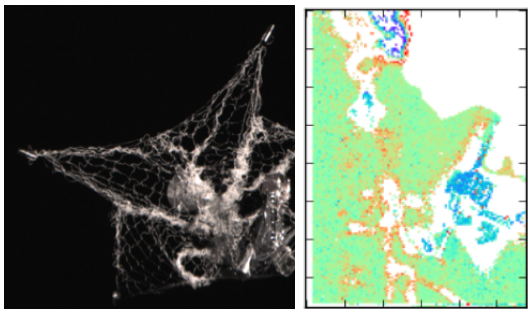


Fig. 16: Left, cropped color camera image, right, time-of-flight LiDAR image

According to the camera picture it is difficult to say whether the cubesat (on the left picture's right hand side) has been or will be catched by the NET. Looking at the LiDAR image it is instantaneously clear that it is just before the capture as the NET (the green structure in the right picture) is on average at 12 m while the cubesat (the blue structure on the right hand side of the right picture with 2 arms of the sail deployed) is at 16 m. At this distance, there is not a complete overlap of the respective FoV of the camera and the LiDAR, this is why the images appear to have been taken for two different positions. Fig. 17 ex-

emplify the effects of the strong illumination by the Sun of the cubesat with deployed panels for the VBN experiment.

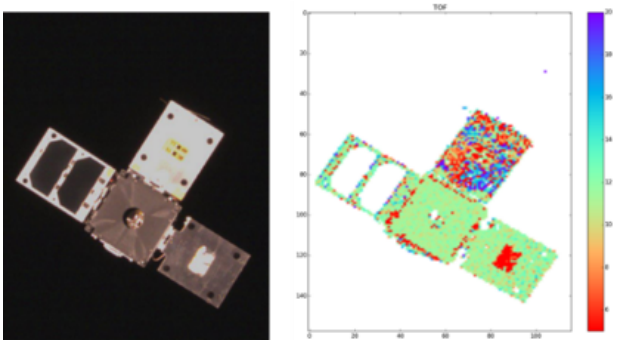


Fig. 17: Left, cropped color camera image, right, time-of-flight LiDAR image, captured respectively 272 s and 268 s after ejection.

Particularly bright areas can be seen in the color camera image. The left and top panels, and center area of bottom right panel are particularly bright. The effect of the strong sunlight is visible in the LiDAR image where the same areas appear either with only red or mixed colors spots. They show the detector saturation and noisy measurements. The panels hit with less sunlight (center and bottom right ones) appear with mostly green spots in the LiDAR image indicating a cubesat that is at 12 m. The un-saturated areas provides measurements allowing to get correct TOF and distances. From these measurements several secondary figures can be calculated such as: 3D target image, relative attitude, velocity and rotation of the target, etc.

In relation with long term flash imaging LiDAR development strategy, the RemoveDebris experience is unvaluable as it shows what it takes in a real situation with strong Sun and Earth background perturbations to get proper visual measurements that can be considered with confidence to control the approach of space debris by waste collection crafts or more generally for in space rendezvous between two objects. The RemoveDebris VBN captured color and LiDAR images over more than 100 m of a 3U cubesat size target (10x10x30 cm) within the expected distance accuracy of a few tens of cm. The strong light background either due directly to the Sun and the sunlight reflected by the Earth is the most challenging factor. It can be addressed through several ways including faster image capture rate, shorter integration time, smaller bandwidth optical filter, matched filtering, etc.

5.4 *Comparison of the different methods*

The results from the image processing algorithms based on SV0 images using the Inria approach (section 5.3.1) and VBN camera images using the Airbus approach (section 5.3.2) together with LiDAR data), GPS measurements and the expected nominal DSAT2 trajectory are shown hereafter on Fig. 18.

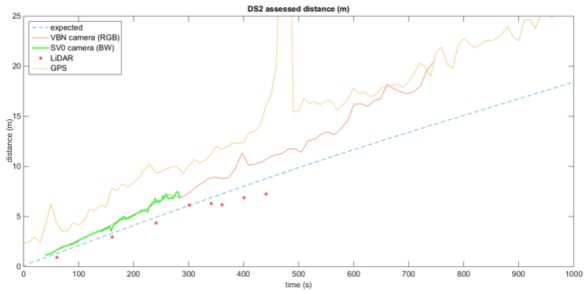


Fig. 18: Tracking results synthesis: DSAT2 assessed distance (expected distance in blue).

As can be seen SV0 and VBN camera measurements are consistent (2.3 cm/s ejection) using both Inria and Airbus approaches. It is also consistet with the expected trajectory. Both approaches allow the estimation of the position and orientation for almost 300s. Thanks to its tracking by detection ability, Airbus allows a longer estimation of the target position. Although not reported in this paper, it has to be noted that GPS measurements seem not reliable for short range (<10m) and that LiDAR measurements (provided by CSEM) are also close to expected trajectory (2 cm/s ejection).

6. Conclusion

Main objectives were mostly met with the VBN demonstration, therefore considered a success by Airbus / CSEM / Inria. Indeed, a very rich and useful set of images and LiDAR data was gathered, allowing preliminary characterization of algorithms performance and identification of possible improvements. Future activities should focus on:

- Analysis of the remaining data, in particular, medium/long range data
- Improvement of ground truth data, including investigation of sensors cross-correlation
- Work on algorithms current limitations and improvements e.g. medium range behaviour, influence of background, navigation aiding for improved tracking, robustness to 3D model errors, LiDAR/camera fusion.

For a future ADR demonstration, next steps would include real-time, onboard VBN algorithms and eventually closed-loop GNC.

7. ACKNOWLEDGEMENT

This research is supported by the European Commission FP7-SPACE-2013-1 (project 607099) "RemoveDebris - A Low Cost Active Debris Removal Demonstration Mission", a consortium partnership project consisting of: Surrey Space Centre (University of Surrey), SSTL, Airbus GmbH, Airbus SAS, Airbus Ltd, Ariane Group, Innovative Solutions in Space (ISIS), CSEM, Inria and Stellenbosch University.

REFERENCES

[1] G. Aglietti, B. Taylor, S. Fellowes, T. Salmon, I. Retat, A. Hall, T. Chabot, A. Pisseloup, C. Cox, A. Zarkesh, A. Mafficini, N. Vinkoff, K. Bashford, C. Bernal, F. Chaumette, A. Pollini, and W. Steyn. The active space debris removal mission RemoveDebris. Part 2: in orbit operations. *Acta Astronautica*, 2019.

[2] G. Bleser, Y. Pastarmov, and D. Stricker. Real-time 3d camera tracking for industrial augmented reality applications. *Journal of WSCG*, pages 47–54, 2005.

[3] C. Bonnal, J.M. Ruault, and M.C. Desjean. Active debris removal: Recent progress and current trends. *Acta Astronautica*, 85:51–60, 2013.

[4] T. Brox, B. Rosenhahn, D. Cremers, and H.-P. Seidel. High accuracy optical flow serves 3-D pose tracking: exploiting contour and flow based constraints. In A. Leonardis, H. Bischof, and A. Pinz, editors, *European Conf. on Computer Vision, ECCV'06*, volume 3952 of *LNCS*, pages 98–111, Graz, Austria, May 2006. Springer.

[5] T. Chabot, K. Kanani, A. Pollini, F. Chaumette, E. Marchand, and J. Forshaw. Vision-based navigation experiment onboard the removedebris mission. In *Int. ESA Conf. on Guidance, Navigation & Control Systems, GNC 2017*, Salzburg, Austria, May 2017.

[6] M. Chapuy, N. Despré, P. Hyounet, F. Capolupo, R. Brochard, and C. Robin. Neoshield-2: design and end-to-end validation of an autonomous closed-loop gnc system for asteroid kinetic impactor missions. In *10th International ESA Conference on Guidance, Navigation & Control Systems*, Salzburg, Austria, May 2017.

[7] A.I. Comport, E. Marchand, M. Pressigout, and F. Chaumette. Real-time markerless tracking for augmented reality: the virtual visual servoing framework. *IEEE Trans. on Visualization and Computer Graphics*, 12(4):615–628, July 2006.

[8] T. Drummond and R. Cipolla. Real-time visual tracking of complex structures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(7):932–946, July 2002.

[9] J. Forshaw, G. Aglietti, S. Fellowes, T. Salmon, I. Retat, A. Hall, T. Chabot, A. Pisseloup, D. Tye, C. Bernal, F. Chaumette, A. Pollini, and W. Steyn. The active space debris removal mission RemoveDebris. Part 1: from concept to launch. *Acta Astronautica*, 2019.

[10] R. Howard, A. Heaton, R. Pinson, and C. Carrington. Orbital express advanced video guidance sensor. In *IEEE Aerospace Conference*, pages 1–10, 2008.

[11] K. Kanani, C. Robin, A. Masson, R. Brochard, P. Duteis, and R. Delage. Dual technologies to mature vision based navigation: the example of space rendezvous and air to air refueling. In *10th International ESA Conference on Guidance, Navigation & Control Systems*, Salzburg, Austria, May 2017.

[12] Y. Ma, S. Soatto, J. Košecká, and S. Sastry. *An invitation to 3-D vision*. Springer, 2004.

[13] E. Malis and E. Marchand. Experiments with robust estimation techniques in real-time robot vision. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'06*, pages 223–228, Beijing, China, October 2006.

[14] E. Marchand, H. Uchiyama, and F. Spindler. Pose estimation for augmented reality: a hands-on survey. *IEEE Trans. on Visualization and Computer Graphics*, 22(12):2633–2651, December 2016.

[15] N. Oumer, G. Panin, Q. Mülbauer, and A. Tseneklidou. Vision-based localization for on-orbit servicing of a partially cooperative satellite. *Acta Astronautica*, 117:19–37, 2015.

[16] G. Panin, Al. Ladikos, and A. Knoll. An efficient and robust real-time contour tracking system. In *IEEE Int. Conf. on Computer Vision Systems, ICVS'06.*, 2006.

[17] A. Petit, K. Kanani, and E. Marchand. Vision-based detection and tracking for space navigation

in a rendezvous context. In *Int. Symp. on Artificial Intelligence, Robotics and Automation in Space, i-SAIRAS 2012*, Torino, Italia, September 2012.

[18] A. Petit, E. Marchand, and K. Kanani. A robust model-based tracker for space applications: combining edge and color information. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'2013*, pages 3719–3724, Tokyo, Japan, November 2013.

[19] M. Pressigout, E. Marchand, and E. Mémin. Hybrid tracking approach using optical flow and pose estimation. In *IEEE Int. Conf. on Image Processing, ICIP'08*, pages 2720–2723, San Diego, California, October 2008.

[20] V. Prisacariu and I. Reid. Pwp3d: Real-time segmentation and tracking of 3d objects. In *British Machine Vision Conference*, September 2009.

[21] J. Shi and C. Tomasi. Good features to track. In *IEEE Int. Conf. on Computer Vision and Pattern Recognition, CVPR'94*, pages 593–600, Seattle, Washington, June 1994.

[22] L. Vacchetti, V. Lepetit, and P. Fua. Combining edge and texture information for real-time accurate 3d camera tracking. In *ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'2004*, volume 2, pages 48–57, Arlington, Va, November 2004.