SELECTING KEY FEATURES OF 3D OBJECT MODEL FOR RELATIVE POSE ESTIMATION

K. Klionovska[(1)] **and M. Burri**[(1)]

[(1)] *German Aerospace Center (DLR)/German Space Operations Center (GSOC), 82234, Weßling, Germany, ksenia.klionovska@dlr.de, matthias.burri@dlr.de*

ABSTRACT

This paper addresses the discussion of the complexity and level of detail of the 3D model of a target satellite that is required in order to get a stable and accurate relative pose estimation with monocular cameras during an On-Orbit Servicing (OOS) mission scenario. We assume to know the target 3D mesh with thousands of vertices before the mission takes place. Nevertheless, for on-board pose estimation we have to use compact 3D models of the targets with only some key points. In this paper we compare the pose estimation results using three models for tests. The first model has manually extracted key features, where the second one includes only features extracted with Harris3D technique. The third model includes a part of Harris3D key features and some manually selected points. The offline pose estimation tests were using the data from European Proximity Operations Simulator (EPOS) facility.

1 INTRODUCTION

Visual navigation is part of complex procedures during rendezvous and proximity operations in On-Orbit Servicing (OOS) missions. In particular, the estimation of the space target's pose (position and orientation) in close range presents a challenge. Different factors influence the accuracy of a visual pose estimation. One of them is the uncooperativeness of the space object, it does not have a communication link or any navigation aid reflectors. Moreover, if this space debris has been staying for a long time in space, we have to consider an aging process and maybe also a loss of some object's parts.

Usually monocular cameras and LIDARs are used for pose estimation in real missions. Researchers all over the world develop and improve image/point cloud processing techniques for pose estimation [1][2][3][4]. There are two main groups: Deep Learning (DL) and traditional image processing (TIP). The trendy first group DL has gained a big popularity in the last years and has shown its pros and cons in different applications. O'Mahony et al. [5] give an imposing comparison of DL versus TIP techniques. DL techniques are trained and not programmed, therefore require less fine tuning. Neural networks even discover the common underlying pattern of the input dataset and automatically provide the set of features with respect to the specific object. In case of TIP, the feature extraction process and parameters tuning are a long trial for a development team. On the other side, if training data of DL is poor or differ from the real (testing) images, the trained model could not perform well and it is almost impossible to tweak its parameters. DL requires high computing power, which is not available on the on-board computers of the current generation of satellites. For example, in the experiments that have been flown to demonstrate the autonomous visual rendezvous, a rough target identification process

using a histogram analysis with FPGA was performed [6]. Recently, Ekblad et al. [7] have already experimented to run DL techniques for future missions with space-grade Xilinx Versal FPGA. They proposed some ideas for performance improvements. First one is an acceleration of post processing using parallelization techniques. Second one is a pipelining, when the on-board's DPU and CPU work both on something in parallel.
In current research, we rely on TIP techniques and develop software for visual navigation with monocular cameras. In our previous work [8], we have introduced a robust feature extraction technique for relative pose estimation. Since we want our method be applicable for different approaches (such as fly-around or straight-line approach) with tumbling targets, we have done many closed-loop experiments with EPOS. Through some experiments, it has been found, that not only image processing affects the final result of the relative pose. The configuration of the key features of the known 3D model of the target has also its impact. In this paper we are going to present the experimental results of the offline tests of pose estimation technique with three setups of target key features. The images and ground truth were collected at the EPOS laboratory.

2 MATERIALS AND METHODS

2.1 Key Points Extraction

The process of key points extraction from 3D models is a very challenging task. There is no strict definition how the interesting key features should be defined. In most cases it is dependent on the features, which are going to be detected on the space object. Very common features are corners or lines of the spacecraft, which can be detected from the visual data for further matching. Considering these parts, we have to build a 3D key features of the target accordingly. Moreover, the following issues listed below make the task even more difficult.

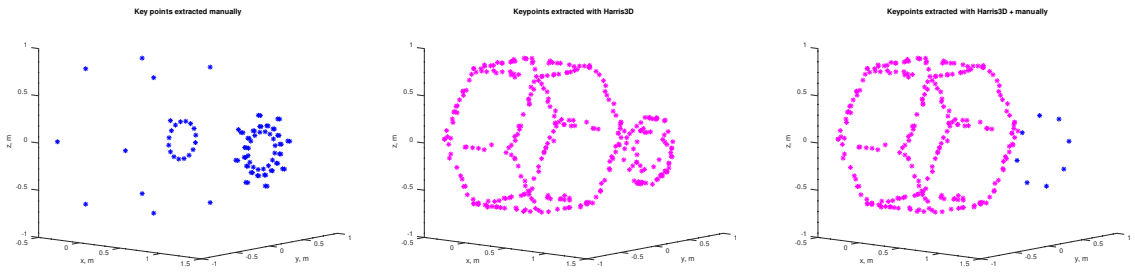
- The space debris might have a tumbling motion in space. The completeness of the key features model must be sufficient for the developed pose estimation technique to track it from different approach angles, while a subset of the features could be hidden behind the target.
- The up-to-date target model may be only available during the mission, e.g. reconstructed after an inspection phase. When we are still on the ground, we can only have an approximate 3D model of the target. Then the question arises – how do we select the model key points on the ground before the OOS mission takes place, if these points could be no more existent on the real space object? Ideally, re-selection of the feature points based on the updated model would happen in an automatic way.
- The image processing task is a challenge in space, because of the limited computing power of an On-Board Computer (OBC). Ideally the feature points extraction from a 2D image with its following matching should be executed with a few times per second.

2.2 Simplified Models

Let us present a 3D model we are working currently with in Figure 1. The mockup geometry is similar to that one, which was planned to use in DEOS mission [9]. The corresponding model is in STL format and has more than 42k vertices and 78k faces. The shape of the space target mockup is complex. It consists of the hexagon part, a nose cone and a disc with an octagon frame. The surface properties of this mockup are representative for the real satellite. The MLI material is used on the frontal hexagon part, whereas six side panels are solar panels.

Figure 1: Both images: initial 3D model in .stl format is presented from different views.

Having a look at the constrains described above in section 2.1, we are not able to take the current huge 3D mesh for further processing. Therefore, we have to choose only some main key points from this 3D mesh.
Three simplified 3D models of key points are presented in Figure 2 and used to track the target object. The pose estimation results for the tracking are presented in section 3.



(a) (b) (c)

Figure 2: Three types of key points: (2a) manually extracted; (2b) with Harris3D algorithm; (2c) Harris3D + manually

.

The model with manually extracted key points in Figure 2a considers the corners of the hexagon and octagon, as well as the cone part. In Figure 2b the key points have been extracted using HarrisKeypoint3D [10] function from PCL library [11]. This function needs two input parameters besides the original model - radius for normal estimation and threshold for filtering out weak corners. The key features from the Figure 2b were extracted with radius = 35.0 mm and threshold = 0.000001. The third model of key points in Figure 2c contains the whole hexagon part extracted with HarrisKeypoint3D and a minimized set of manually selected points of the octagon. The idea to create the third model appeared after processing and analyzing the pose estimation errors with the first and second models. As it was already mentioned in introduction, the final decision for extraction of key features in traditional image processing techniques is up to the developer.

2.3 Pose estimation

In this work we consider the pose estimation only with a monocular camera. The process to estimate the position and orientation of the space object via gray scale images consists of several steps and already was presented in [8]. The image is processed with Hough Line Transform (HLT) [12] and Shi-Tomasi (ST) [13] feature detector algorithms. The HLT transform delivers the straight lines, where only the end points of these lines are taken for the following matching. The end points of HLT and corners from ST detector organize a big set of 2D image points. The next step is feature matching. In the feature matching step, the simplified 3D model from Section 2.2 is projected onto image plane and is aimed to find the nearest image point. Thereafter the EPNP solver [14] calculates the relative position and orientation of the space object.

3 RESULTS

The Hardware-in-the-Loop (HiL) European Proximity Operations Simulator (EPOS) simulator located at German Aerospace Center [15] is used for simulation of different approach scenarios. We present here experiments of the straight line approach and a fly-around. The sketch of the experiment is presented in Figure 3. Let us define Test I as a straight-line approach rendezvous from 10 m to 4 m, with an approach velocity of 2 cm/sec; Test II is a fly-around with an azimuth α changing from $86°$ to $50°$ and back to $77°$. The target mockup rotates with 1 deg/sec along it's principal axis (see Figure 3) during all simulations.

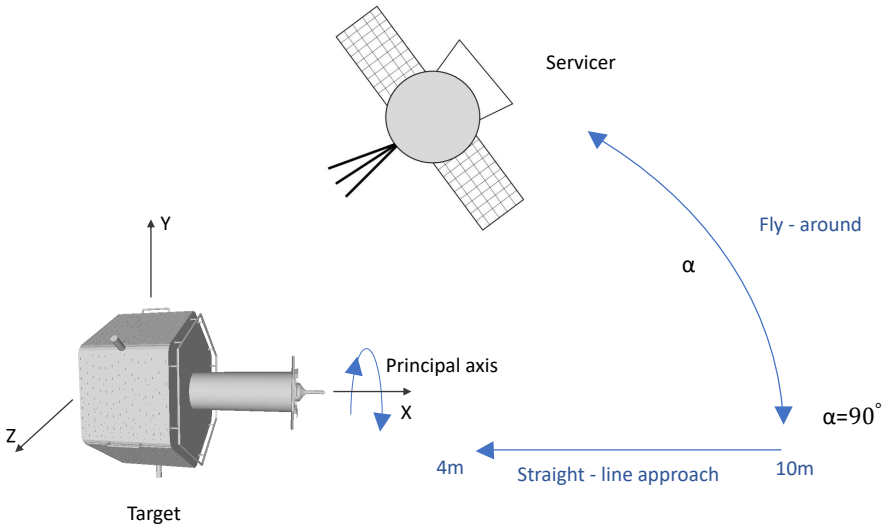


Figure 3: Sketch of the experiment.

The image data was collected from wide field-of-view (WFOV) camera Prosilica GT205. The characteristics of this visual sensors are presented in Table 1.

Table 1: Cameras characteristics

Name	**WFOV**
Model	Prosilica GT2050 (02-2626A)
Pixel size [μm]	5.5 x 5.5
Sensor field [pixel]	2048 x 2048
Focal length [mm]	12.5
FOV [°]	47 x 47

Test I: straight-line approach

In Figure 4 the position (first three rows) and angular (last row) errors for three different configurations of key features during straight-line approach are presented. The dark red (blue) line in the plot for the position (angular) errors is a moving average. It is plotted over an error band limited by the moving minimum and moving maximum, which are colored in black. The data were filtered with a 20-points filter window. In all cases, the position error slightly drops with the decreasing distance between chaser and target. In the test with *Manual* and *Harris3D+Manual* extracted points, the jumps of position errors occur and are around 0.4 m. Nevertheless, it has not affected the stability of the tracker. Angular errors are calculated as an angular difference between two quaternions - the ground truth orientation and estimated one. During the whole straight-line approach, the average angular errors are not higher than 3 deg. There is a jump of 7 deg with *Harris3D+Manual* extracted points at 9.5 m relative distance, which caused an error of X position error around 0.4 m.

Test II: fly-around

The fly-around case is more sensitive to errors than the straight-line approach. During the fly-around the relative distance is kept constant to 5.9 m. It can be noticed (see Figure 5) during the test with *Manual* and *Harris3D* models, that the peaks in angular errors lead to the position errors. For example, the angular error with a jump up to 14 deg with *Harris3D* model provoked errors with a magnitude of 0.4 m in distance component along X axis, 0.1 m along Y axis and also 0.05 m along Z axis. From the plots with *Harris3D* model, it is evidently shown, that the fly-around motion before this case was stable with an average 0.1 m errors along X axis. The motion with the *Manual* model has higher magnitudes of maximum errors up to 0.2 m. (see e.g. by azimuth $83°$ and $57°$), even when the angular error is not higher than 3 deg. When the angular error jumped up to 9 deg, increased the position errors respectively: X error is 0.58 m, Y error is 0.2 m and Z absolute error is 0.15 m. The fly-around with *Harris3D+Manual* model provided a better and stable results. The angular error of 8 deg led to 0.2 m the absolute position error X, 0.05 m Y error and 0.03 m Z error. The cross correlation between the maximal observed angular error and the position error is minimal with *Harris3D+Manual* model compared to the other models.

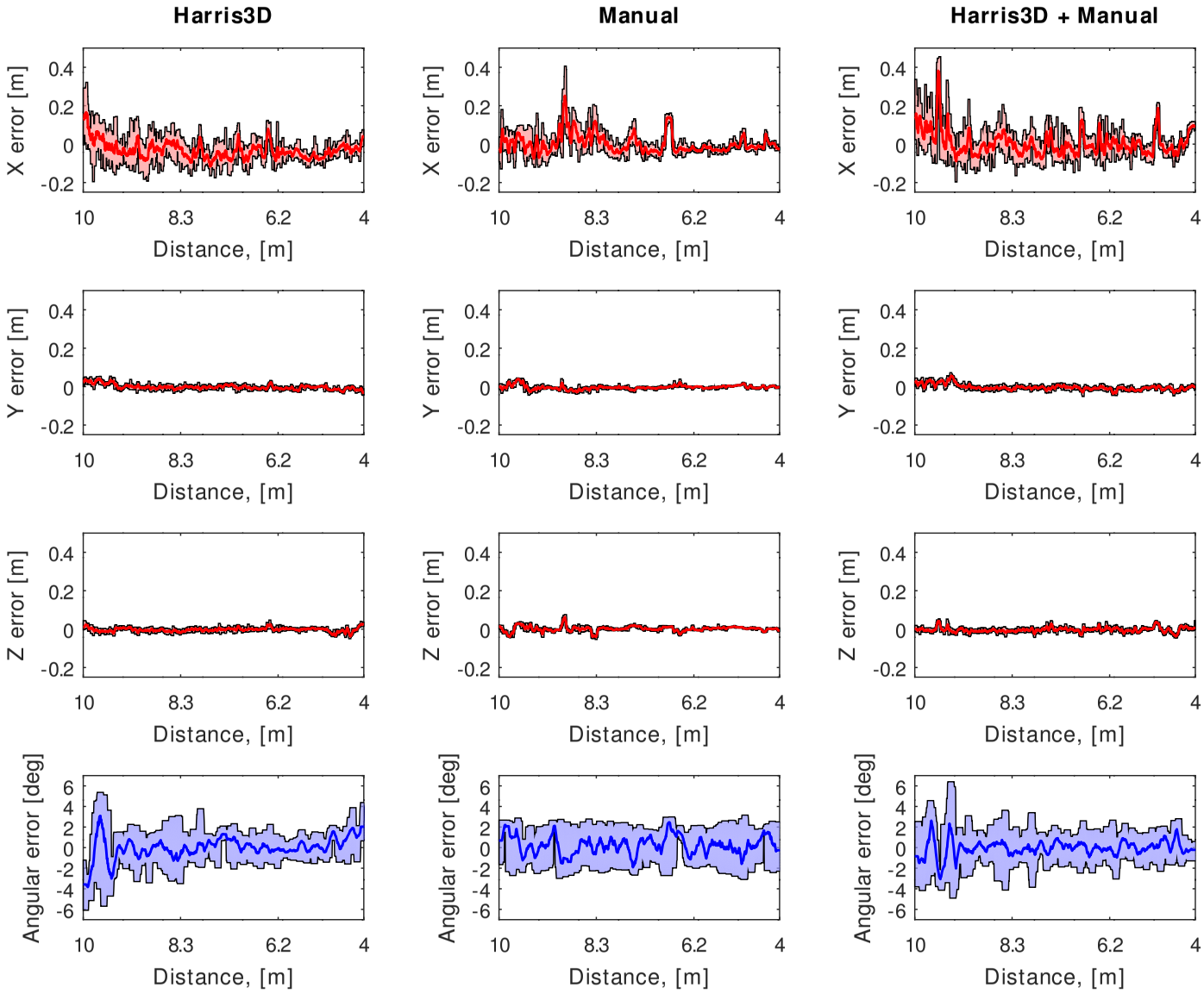


Figure 4: Position errors for Test I.

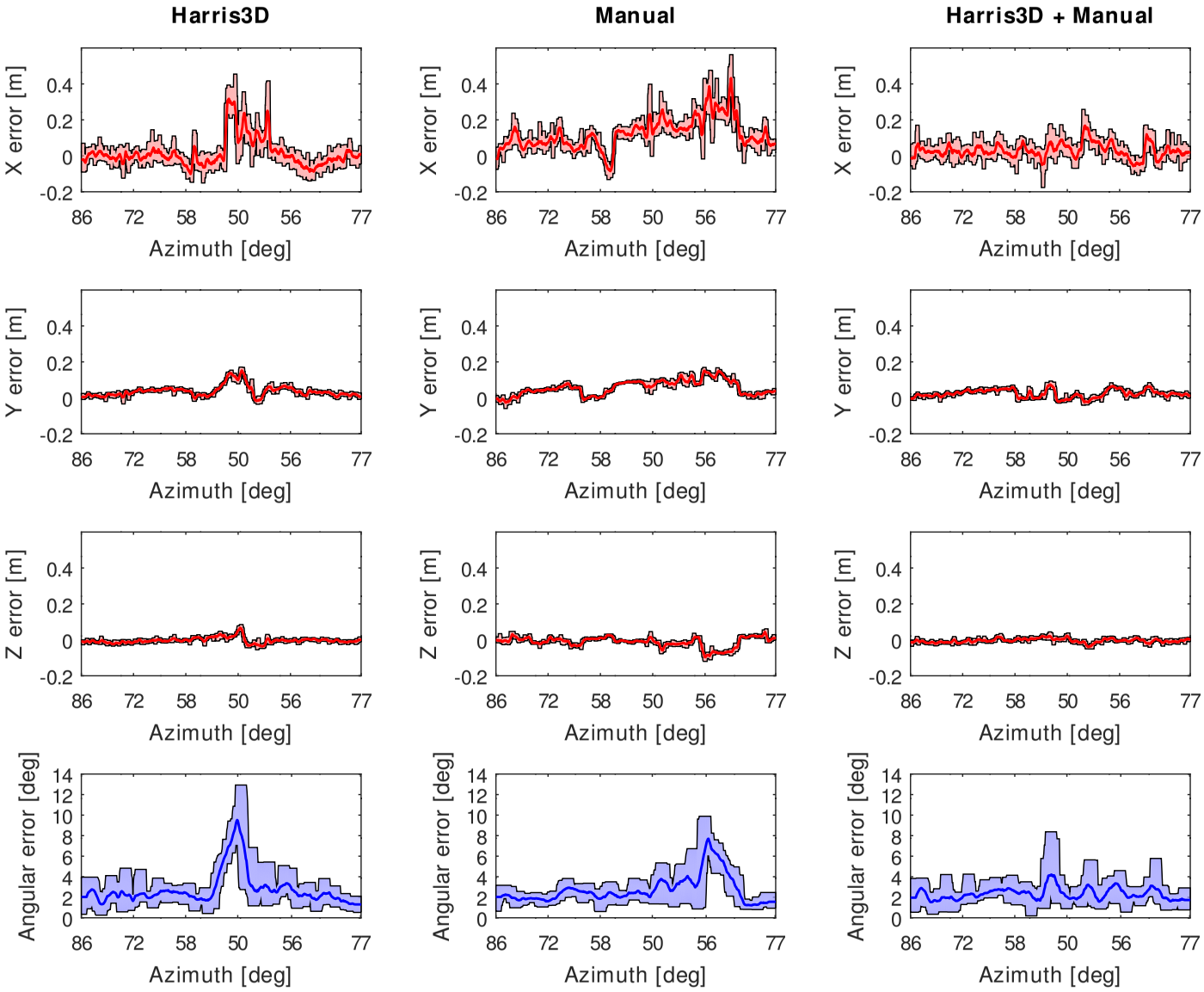


Figure 5: Position errors for Test II.

In Table 2 we provide statistical analysis of estimated distances with respect to the recommended accuracy according to [16], which is 1% of range. For the Test I, the maximum and minimum percentages of acceptable distances were identified with *Manual* model and *Harris3D* models, 91.1 % and 77.3% respectively. Nevertheless, for Test II, we observe completely different results. The maximum percentage of safe measurements belongs to *Harris3D+Manual* model and is 81.7%. The *Manual* key features model gained 17.2%. It is a minimum percentage of measurements, which are inside of the tolerance range. The total statistic for Test I and Test II shows that the more accurate estimation of relative distance occurred with *Harris3D+Manual* key features model.

Table 2: Percentage of measurements within a safe region (1% of range)

Model Name	**Test I, [%]**	**Test II, [%]**	**Total, [%]**
Harris3D	77.3	76.7	77.1
Manual	91.1	17.2	64.7
Harris3D + Manual	85.6	81.7	84.2

4 CONCLUSIONS

In this paper we focused on the problem, how complex should be a 3D model of key features for pose estimation. We arranged two experiments (straight-line approach and fly-around), where we tested three configurations of model points. The first model is *Harris3D* used the automatic Harris3D key extraction technique; the second one is *Manual* with hand extracted key points; and *Harris3D+Manual* model contains both (automatically and manually key points). The most important conclusion from the results: not in all simulation cases the same 3D key features model gives good results. However, the tracking with *Harris3D+Manual* model showed the best performance for a fly-around and straight-line approach scenario used in this paper.
Since we want our technique for pose estimation be applicable and stable for different approach scenarios, there is a big need for automated testing. We are going to work on it in the near future. It will help us to test all three key features models with other image data sets for a more accurate assessment.

REFERENCES

[1] K. M. Kajak, C. Maddock, H. Frei, and K. Schwenk, "Domain randomisation and cnn-based keypoint-regressing pose initialisation for relative navigation with uncooperative finite-symmetric spacecraft targets using monocular camera images," *Advances in Space Research*, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0273117723001436

[2] A. Lotti, D. Modenini, P. Tortora, M. Saponara, and M. A. Perino, "Deep learning for real-time satellite pose estimation on tensor processing units," *Journal of Spacecraft and Rockets*, vol. 0, no. 0, pp. 1–5, 0. [Online]. Available: https://doi.org/10.2514/1.A35496

[3] L. Renaut, H. Frei, and A. Nüchter, "Smoothed normal distribution transform for efficient point cloud registration during space rendezvous," in *Proceedings of the 18th International Joint Con-*

ference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP, INSTICC. SciTePress, 2023, pp. 919–930.

[4] G. Napolano, C. Vela, A. Nocerino, R. Opromolla, and M. Grassi, "A multi-sensor optical relative navigation system for small satellite servicing," *Acta Astronautica*, 2023.

[5] N. O. Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. A. Velasco-Hernández, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Computer Vision Conference*, 2019.

[6] S. Kimura, Y. Horikawa, and Y. Katayama, "Quick report on on-board demonstration experiment for autonomous-visual-guidance camera system for space debris removal," *Transactions of the Japan Society for Aeronautical and Space Sciences, Aerospace Technology Japan*, vol. 16, pp. 561–565, 09 2018.

[7] A. Ekblad, T. Mahendrakar, R. T. White, M. Wilde, I. Silver, and B. Wheeler, "Resource-constrained fpga design for satellite component feature extraction," *arXiv preprint arXiv:2301.09055*, 2023.

[8] K. Klionovska, M. Burri, and H. Frei, "Robust feature extraction pose estimation during fly-around and straight-line approach in close range." in *16th Symposium on Advanced Space Technologies in Robotics and Automation (ASTRA 2022)*, 2022. [Online]. Available: https://elib.dlr.de/187326/

[9] T. Wolf, D. Reintsema, B. Sommer, P. Rank, and J. Sommer, "Mission deos proofing the capabilities of german's space robotic technologies," in *International Symposium on Artificial Intelligence, Robotics and Automation in Space–i-SAIRAS*, 2012.

[10] I. Sipiran and B. Bustos, "Harris 3d: A robust extension of the harris operator for interest point detection on 3d meshes," *The Visual Computer*, vol. 27, pp. 963–976, 11 2011.

[11] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China: IEEE, May 9-13 2011.

[12] P. V. Hough, "Method and means for recognizing complex patterns," 12 1962. [Online]. Available: https://www.osti.gov/biblio/4746348

[13] J. Shi and Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[14] V. Lepetit, F. Moreno-Noguer, and P. Fua, "Epnp: An accurate o(n) solution to the pnp problem," *International Journal of Computer Vision*, vol. 81, no. 2, p. 155, Jul. 2008.

[15] F. Rems, H. Frei, E.-A. Risse, and M. Burri, "10-year anniversary of the european proximity operations simulator 2.0 - looking back at test campaigns, rendezvous research and facility improvements," *Aerospace*, vol. 8, no. 9, p. 235, August 2021. [Online]. Available: https://elib.dlr.de/143624/

[16] W. Fehse, *Automated rendezvous and docking of spacecraft*, ser. Cambridge Aerospace Series, M. J. Rycroft and W. Shyy, Eds. Cambridge University Press, 2003.