73rd International Astronautical Congress (IAC), Paris, France, 18-22 September 2022.

IAC–22–C1.3.9

Pose Estimation of a Known Texture-Less Space Target using Convolutional Neural Networks

Arunkumar Rathinam[a]*, Vincent Gaudillière[a], Leo Pauly[a], Djamila Aouada[a]

[a] *Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg, arunkumar.rathinam@uni.lu*
* Corresponding author

Abstract

Orbital debris removal and On-orbit Servicing, Assembly and Manufacturing [OSAM] are the main areas for future robotic space missions. To achieve intelligence and autonomy in these missions and to carry out robot operations, it is essential to have autonomous guidance and navigation, especially vision-based navigation. With recent advances in machine learning, the state-of-the-art Deep Learning [DL] approaches for object detection, and camera pose estimation have advanced to be on par with classical approaches and can be used for target pose estimation during relative navigation scenarios. The state-of-the-art DL-based spacecraft pose estimation approaches are suitable for any known target with significant surface textures. However, it is less applicable in a scenario where the target is a texture-less and symmetric object like rocket nozzles. This paper investigates a novel ellipsoid-based approach combined with convolutional neural networks for texture-less space object pose estimation. Also, this paper presents the dataset for a new texture-less space target, an apogee kick motor, which is used for the study. It includes the synthetic images generated from the simulator developed for rendering synthetic space imagery.
Keywords:spacecraft pose estimation; ellipsoidal modelling; akm dataset; textureless pose estimation

1. Introduction

Over the last two decades, vision-based navigation (VBN) for spacecraft has been pursued to achieve autonomy in mission scenarios where higher time delays restrict the ground segment from having direct control over the spacecraft. With recent technological advancements combined with increased reliability of the software systems, the VBNs are being adapted for a wide range of on-orbit operations to increase autonomy. For example, ISS resupply missions are a well-known example of VBN with a cooperative target, where the target vehicle and the approaching spacecraft (i.e. servicer or chaser) can share information with each other. However, when the target is non-cooperative such as a defunct satellite, considering the mission scenario of On-Orbit Servicing (OOS) and Active Debris Removal (ADR), the navigation scenario poses several challenges to the Visual GNC systems. OOS and ADR are considered key capabilities for spaceflight in this century, and multiple technology demonstration missions are either carried out or planned for the future, including PROBA-3 [1] by ESA and PRISMA [2] by OHB Sweden. The first commercial OOS of a geostationary satellite (IntelSat-901) was launched by Space Logistics using the MEV-1 (Mission Extension Vehicle) satellite platform. MEV-1 docked with the IntelSat satellite, re-positioned it to the designated spot, and continues to provide in-orbit station-keeping services [3].

With more complex missions, the modern Guidance, Navigation, and Control (GNC) solutions need to be more autonomous to allow the robotic systems to perform autonomous relative navigation, rendezvous, and robotic actions like capturing the target or debris [*]. ADR mission may include removing some of the existing launch vehicle parts to clean up the highly sought orbital space or to remove the threat to space assets. With the above mission scenarios, any proximity operation requires on-orbit relative navigation with respect to the target. For spacecraft relative navigation operations, the fundamental requirement is the real-time estimation of the target's position and orientation (*i.e.* pose). This enables the chaser or servicer spacecraft to generate trajectory and perform control updates. In the above scenarios, the target is usually known but non-cooperative. During the approach sequence in every orbital robotic operation, the GNC system needs an estimate of the target's six Degree-of-Freedom (DOF) pose, *i.e.*, the relative position and attitude represent key information for the navigation system. Monocular cameras are widely preferred over other sensors (such as LIDARs and RADARs) for relative navigation scenarios. This is mainly due to their relative simplicity, small size, low weight and power requirements and can be easily adapted to a wide range of spacecraft configurations. The limitations of 2D images from monocular cameras include

[*] https://spaceflight101.com/re-entry-detla-ii-payload-assist-module/

no depth measurements and sensitivity to adverse illumination conditions.

Deep Learning [DL] approaches for spacecraft pose estimation have gained interest in the last few years [4], [5], [6], [7]. The state-of-art approaches are suitable for a known target spacecraft where they have significant textures on the surface. However, it is less suitable when the target is a texture-less and symmetric object like rocket nozzles. For example, ADR missions targeting spent rocket stages or apogee kick motors require an approach that estimates the pose without using keypoints.

In this work, we present a dataset for texture-less space targets, i.e. Apogee Kick Motor (AKM), with relative pose labels. Further, we investigate a novel pose estimation framework for combining position estimation using a convolution neural network and inferring the orientation information from the estimated position information using ellipsoid parameters. The paper is arranged as follows: Section 2 presents background on related spacecraft pose estimation literature works. Section 3 presents the details on the AKM dataset, followed by the description of the Pose Estimation Framework based on an ellipsoidal model in section 4. Section 5 present the details of different experiments and is followed by the analysis of results in terms of orientation accuracies in section 6. Section 7 presents the conclusion and future work.

2. Related Works

Recent advancements in deep learning and the popularity of the ESA's Spacecraft Pose Estimation Challenges [†] enabled new developments with the state-of-the-art performance in the visual pose estimation algorithms. Several research works and their results are published based on this competition, including Spacecraft Pose Network (SPN) [6], Pose Estimation with Deep Landmark Regression [5], Pose Estimation with soft classification [7], segmentation driven approach [8]. Typical deep-learning-based pose estimation for space targets can be classified into two categories. They are **a).** End-to-End or Direct approaches **b).** Hybrid approaches using deep learning in combination with PnP. The direct approach estimates the pose from a single image, which can be done by directly regressing the position and the unit quaternion vectors [9]. However, previous studies indicate that the orientation regression, i.e. a norm-based loss of unit quaternions, doesn't provide good estimates and results in a larger error than the position estimates[7]. This behaviour is attributed to the loss function and its inability to represent the actual angular distance for any orientation representation. A second approach is a keypoint-based approach [5], [10] that relies on estimating interesting target features in the image and is followed using PnP to solve the problem of estimating the target pose. One of the challenges in both approaches is that they heavily rely on the texture of the spacecraft. The algorithms based on the two approaches are mainly validated on the two standard datasets, *SPEED* [11] and *SPEED+*[12].

Pose estimation for textureless space objects by contour points matching was investigated in [13], and the approach utilizes the shape and contour information. Usage of geometric factors was investigated to establish 2D–3D correspondences to calculate confidence probabilities and construct a weight matrix to perform pose estimation. Other works in the field of computer vision including [14], [15], [16] tackles the problem of texture-less pose estimation for objects used in general purpose applications.

3. AKM Dataset

A brief review of the currently available datasets for spacecraft or space objects was summarized in our previous work [17]. One of the limitations of the currently available datasets is that they don't tackle the problem of the textureless object. The target in this dataset addresses two specific characteristics, they are **a).** texture-less **b).** rotational symmetry.

Fig. 1: Sample images of Solid Rocket Motor (Apogee Kick Motor) **Left:** STAR 48 **Right:** STAR 24

The target object in this dataset is the AKM motor, similar to the STAR 24 motor as shown in fig. 1 (Ref: [‡]), a radially symmetric and textureless object. The motor dimensions include a diameter of 24.5 inches (0.62 m) and a length of 42.0 (1.0668 m). The dataset is generated using Blender with lighting conditions similar to the orbital environment. A brief review of different simulators

[†] https://kelvins.esa.int/

[‡] https://www.northropgrumman.com/wp-content/uploads/NG-Propulsion-Products-Catalog.pdf

available for synthetic dataset generation for space objects in orbital lighting conditions is summarised in [18]. The dataset contains 10000 images with an image size of 1024 x 1024 pixels and is generated with a camera model having a field of view of 55 degrees. Sample images from the AKM dataset are shown below in fig. 2. The target positions are obtained from a uniform random distribution between 1m and 10 m, and orientations are obtained from a uniform random distribution. The dataset is available in [19] 10.5281/zenodo.7043325.

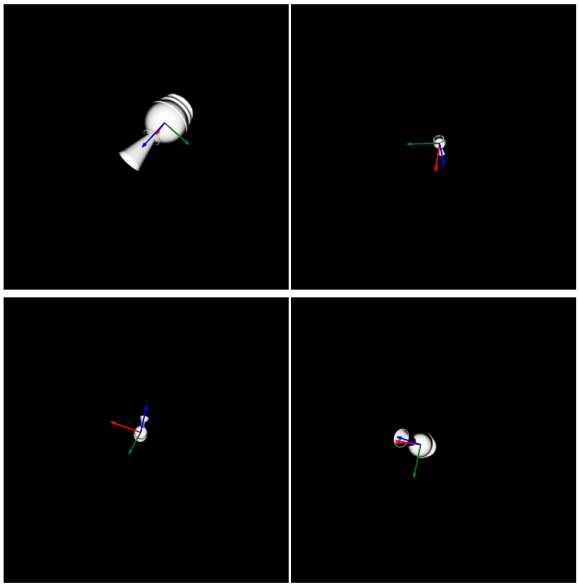

Fig. 2: Sample images from the AKM dataset

4. Pose Estimation Framework

Our model-based pose estimation method leverages an ellipsoidal approximation of the spacecraft. Therefore, it relies on a compact and generic model able to encapsulate the object's position, attitude and coarse dimensions while supporting analytical relations between these quantities. In particular, given the size of the ellipsoid (known model), one can derive the spacecraft attitude from its location up to the model symmetries [20]. Moreover (even if not in the scope of this paper), detecting the ellipsoid projection in the picture (*i.e.* ellipse) does not require any low-level image analysis (*e.g.* keypoint detection), thus making the method suitable at longer ranges and on texture-less targets.

4.1 *Ellipsoidal Modeling*

In Euclidean geometry, and following the notations introduced in [21], an ellipsoid is characterised by its centre position $\mathbf{C} = (C_x, C_y, C_z)^T \in \mathbb{R}^3$ and a 3×3 symmetric positive definite matrix

$$A = R \begin{pmatrix} 1/a^2 & 0 & 0 \\ 0 & 1/b^2 & 0 \\ 0 & 0 & 1/c^2 \end{pmatrix} R^\top$$

encompassing its orientation R and radii a, b, c (all positive), so that its equation is in the form

$$(\mathbf{X} - \mathbf{C})^\top A (\mathbf{X} - \mathbf{C}) = 1, \text{ where } \mathbf{X} \in \mathbb{R}^3.$$

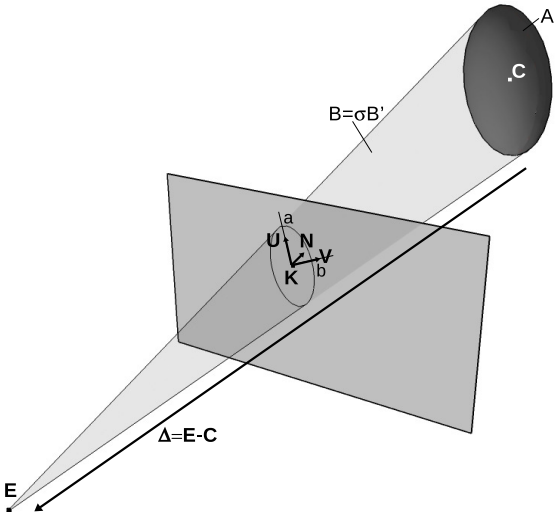

Fig. 3: Illustrating the projection plane, projection centre, ellipsoid and projected ellipse.

Considering a camera with centre $\mathbf{E}$ whose intrinsic parameters are known, and given the ellipse resulting from the ellipsoid projection into the image, one can generate the so-called *backprojection cone* from the lines passing through $\mathbf{E}$ and any point on the ellipse (see Fig. 3). A 3×3 symmetric matrix characterises the shape and orientation of this cone denoted B', so that its equation is

$$(\mathbf{X} - \mathbf{E})^\top B' (\mathbf{X} - \mathbf{E}) = 0.$$

Then, denoting $\mathbf{\Delta}$ the vector connecting the ellipsoid centre $\mathbf{C}$ to the camera centre $\mathbf{E}$, the cone of vertex $\mathbf{E}$ tangent to the ellipsoid, referred to as the *projection cone*, is characterised by the 3×3 symmetric matrix

$$B = A\mathbf{\Delta}\mathbf{\Delta}^\top A + (1 - \mathbf{\Delta}^\top A \mathbf{\Delta}) A,$$

where B plays the same role as B' in the corresponding cone equation [21].

Finally, the projection and back-projection cones being aligned, *i.e.* there exist a scalar σ such that $B = \sigma B'$ [21], the projection equation linking the ellipsoid, ellipse and intrinsic camera parameters is as follows:

$$A\mathbf{\Delta}\mathbf{\Delta}^\top A + \mu A = \sigma B', \tag{1}$$

where $\mu = 1 - \mathbf{\Delta}^\top A \mathbf{\Delta}$.

It has been demonstrated that, given the size of the ellipsoid (*i.e.* given a, b, c), σ, μ and then A can be analytically retrieved from $\mathbf{\Delta}$ and B' [20]. In other words, the orientation of the ellipsoid can be derived from its position, given its size, its projected ellipse and the parameters of the projection (camera intrinsics). This calculation is detailed in Section 4.3.1.

4.2 *Position Regression*

We use a CNN regression model with a ResNet152 [22] backbone to regress the coordinates of the ellipsoid centre (C_x, C_y, C_z) directly from the pictures. The regression model consists of 152 residual layers, followed by a fully connected layer to regress the 3D coordinates. The network is optimised with an SGD solver [23] using the Mean Square Error (MSE) objective loss function. The models are trained for a total of 120 epochs. An initial learning rate of $1e^{-3}$ is used and is reduced by a factor of 0.1 after every 48 (40% of total) epochs.

4.3 *Orientation Derivation*

Given the regressed ellipsoid position, one can analytically compute the corresponding orientation. Such a derivation is presented in Section 4.3.1. Unfortunately, this process is sensitive to noise on the position, even when all the other parameters (i.e. intrinsics, ellipse, ellipsoid size) are known (see Section 6 for more detailed sensitivity analysis). For that, we propose to *learn* the closed-form solution using a MultiLayer Perceptron (MLP) and further propose to use a so-called *Positive Definite* (PD) MLP better to exploit the properties of the A matrix. These strategies are presented in Section 4.3.2.

4.3.1 *Closed-Form Solution*

The scalar parameter σ, involved in Equation (1), is linked to the camera-ellipsoid distance $\|\mathbf{\Delta}\|$ through Equation (2) [20]:

$$\left(\mathrm{tr}(B'^{-1})\right)^2 \sigma = \frac{\det(A)}{\det(B')} \left(\mathrm{tr}(A^{-1}) - \|\mathbf{\Delta}\|^2\right)^2. \quad (2)$$

It is important reminding that, in our problem, A eigenvalues are known, and consequently, the same holds for A^{-1}. Their determinants (det) and traces (tr) are thus also known. Moreover, B' is fully known, since camera intrinsics are known and since detecting the ellipse is not in the scope of this paper (*i.e.* we consider ground-truth projections of the ellipsoidal model as ellipse detection). Therefore, σ can be derived from $\|\mathbf{\Delta}\|$ using Equation (2).

Then, μ can be retrieved from σ using Equation (3) [20]:

$$\mu = -\sqrt{\frac{det(B')}{det(A)}\sigma^3}. \quad (3)$$

Finally, A is computed according to Equation (4) [20]:

$$A = \frac{\sigma}{\mu}\left(B' - \sigma B' \mathbf{\Delta}\mathbf{\Delta}^\top B'\right). \quad (4)$$

The ellipsoid orientation is then obtained by eigenvalue decomposition.

For more information, please refer to [20] (Section 5.2).

4.3.2 *MLP-based Regression*

We created a simple multi-layer perceptron to regress the orientation parameters, in the form of the upper triangular part of A (symmetric), from the ellipse and camera intrinsic parameters, encoded in the upper triangular part of B' (symmetric), and from the regressed position $\mathbf{C}$. There are $6 + 3 = 9$ parameters as input and 6 parameters as output in total.

PD-MLP regression To improve the network's efficiency and to constrain the loss functions based on our prior knowledge about A, we tuned the network layers and defined custom loss functions. In detail, the final layer of the network is adjusted to confirm the Symmetric Positive Definite (SPD) property of the matrix using SPD enforcement layers [§] [24]. The loss function (L_{WS}) was adapted from the Wasserstein loss, with only representations from the covariance term. More precisely, let $\mu_1 = \mathcal{N}(m_1, C_1)$ and $\mu_2 = \mathcal{N}(m_2, C_2)$ be two normal distributions on $\mathbb{R}^n$, with respective expected values m_1, $m_2 \in \mathbb{R}^n$ and symmetric positive definite covariance matrices $C_1, C_2 \in \mathbb{R}^{n \times n}$. Then the 2-Wasserstein divergence between μ_1 and μ_2 is given by

$$W_2(\mu_1, \mu_2)^2 = \|m_1 - m_2\|_2^2 + \mathrm{tr}\left(C_1 + C_2 - 2\left(C_2^{1/2} C_1 C_2^{1/2}\right)^{1/2}\right).$$

Our loss is, therefore

$$L_{WS} = \mathrm{tr}\left(A_{gt} + A_{pred} - 2\left({A_{gt}}^{1/2}\, A_{pred}\, {A_{gt}}^{1/2}\right)^{1/2}\right), \quad (5)$$

where A_{pred} refers to the predicted matrix and A_{gt} to the ground-truth one.

5. Experiments

We have tested our method on two spacecraft pose estimation datasets. The first one is the *AKM* Dataset, introduced in Section 2. The second one is the *SPEED+* Dataset [12], which was introduced for the purpose of an international challenge[¶]. The two sets of experiments are presented in Sections 5.1 and 5.2.

[§] https://github.com/LLNL/spdlayers

[¶] https://kelvins.esa.int/pose-estimation-2021/home/

5.1 *AKM Dataset Experiments*

Since the *AKM* object is not textured, keypoint-based pose estimation methods are likely to fail in most cases. Moreover, its axial symmetry makes its attitude ambiguous, leading direct regression methods such as PoseNet [25] to failure. The results of the AKM dataset experiments are summarised in table 1.

The maximum error is 90° due to the ellipsoidal model symmetries, *i.e.* an axis and a plane orthogonal to that axis.

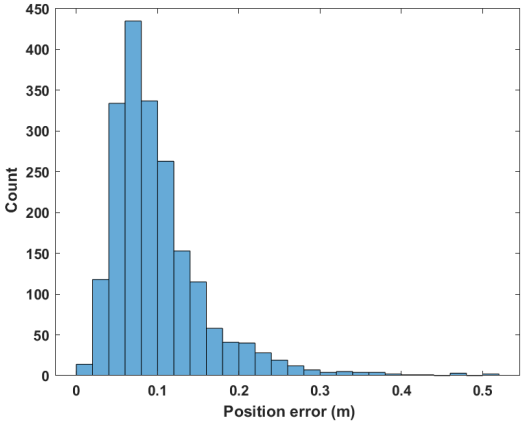

Fig. 4: **AKM** Dataset: **position error** histogram. Positions are obtained by **direct CNN regression**. Mean: **9.97cm**; median: 8.48cm; standard deviation: 5.92cm.

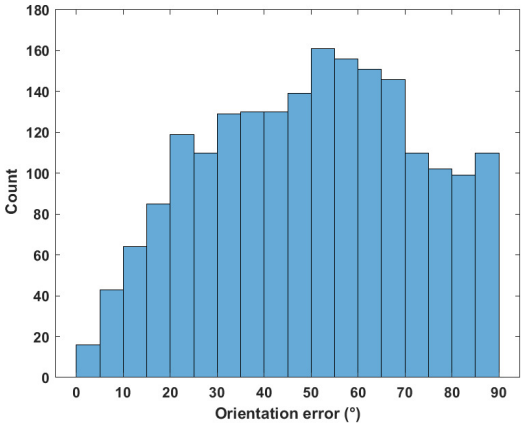

Fig. 5: **AKM** Dataset: **orientation error** histogram. Orientations are analytically derived from **regressed positions** using **ellipsoidal model equations**. Mean: **50.03**°; median: 50.91°; standard deviation: 22.16°.

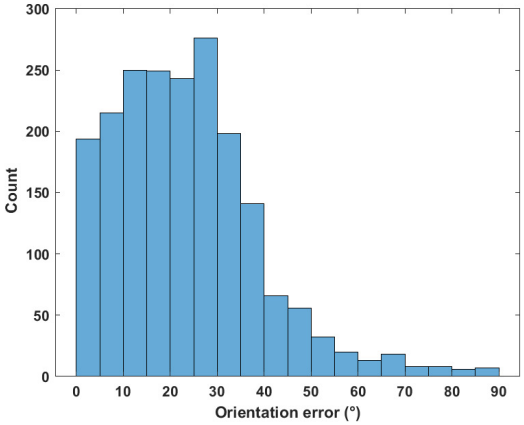

Fig. 6: **AKM** Dataset: **orientation error** histogram. Orientations are predicted from **regressed positions** using a **standard MLP**. Mean: **23.45**°; median: 21.84°; standard deviation: 15.52°.

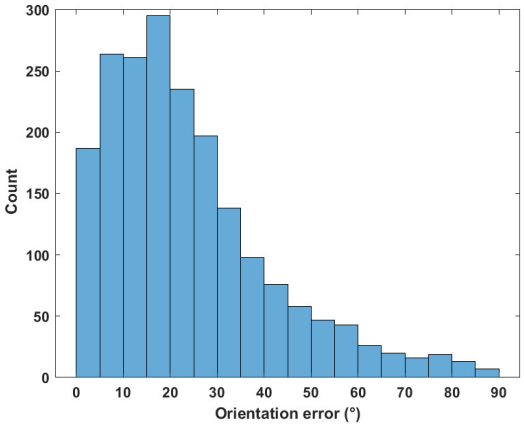

Fig. 7: **AKM** Dataset: **orientation error** histogram. Orientations are predicted from **regressed positions** using a **Positive Definite MLP**. Mean: **23.99**°; median: 19.85°; standard deviation: 17.59°.

5.2 *SPEED+ Dataset Experiments*

The *SPEED+* dataset [12] contains 59960 synthetic images of the TANGO spacecraft** along with their pose labels. The dataset is split into training/test sets following an 80%/20% repartition. To build the validation subset, we apply another 80%/20% split to the training set. As a result, 38374 images were used for training, 9594 for validation and 11992 for testing. The results of the *SPEED+* dataset experiments are summarised in table 2.

**`https://www.eoportal.org/satellite-missions/prisma-prototype`

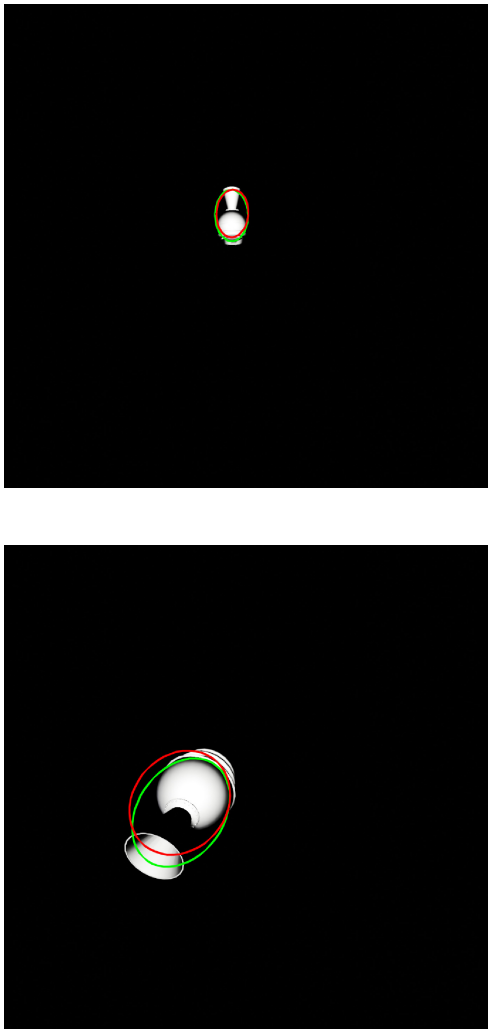

Fig. 8: Reprojections of the ellipsoid into two images from the AKM Dataset, based on the regressed or ground-truth pose. **Red**-predicted ellipse; **Green**-ground-truth ellipse.

Approach	mean	median	std. dev.
ResNet (pos.)	9.97cm	8.48cm	5.92cm
Closed-form (ori.)	50.03°	50.91°	22.16°
Standard MLP (ori.)	23.45°	21.84°	15.52°
PD MLP (ori.)	23.99°	19.85°	17.59°
PoseNet (pos.) [25]	30.66cm	27.06cm	17.36cm
PoseNet (ori.) [25]	105.83°	86.33°	72.84°

Table 1: Summary of position (pos.) and orientation (ori.) errors based on different approaches presented for AKM dataset.

Approach	mean	median	std. dev.
ResNet (pos.)	12.52cm	9.54cm	11.16cm
Closed-form (ori.)	47.86°	46.55°	25.39°
Standard MLP (ori.)	45.63°	42.99°	25.84°
PD MLP (ori.)	49.40°	47.61°	24.43°

Table 2: Summary of position (pos.) and orientation (ori.) errors based on different approaches presented for *SPEED+* dataset.

6. Orientation Accuracy Analysis

In terms of orientation accuracy, the reported results are poor. This section is to analyse the reasons why.

6.1 *Analytical Derivation*

The analytical derivation of the orientation (A) from the position ($\|\boldsymbol{\Delta}\|$), described in Section 4.3, consists in a nutshell in computing the scalar σ (see Equation (2)), then μ (Equation (3)) and finally A (Equation (4)). However, one can observe the norm of the prediction $\|\boldsymbol{\Delta}\|$ is raised to the power of 4 to obtain σ, which is itself raised to the power of 1.5 to compute μ. This drastically heightens the error resulting from the regression of $\boldsymbol{\Delta}$. Moreover, the matrix $\boldsymbol{\Delta}\boldsymbol{\Delta}^\top$ is made of the pairwise products between $\boldsymbol{\Delta}$ elements, here again, multiplying the error. This makes the analytical orientation derivation highly sensitive to noise on the position prediction.

6.2 *MLP-based regression*

To circumvent the difficulties mentioned above, we proposed to learn position-to-orientation mapping. However, the results are also not convincing. We believe the reasons for that lie in the problem structure itself. Indeed, given a single image, there is already an infinite number of ellipsoids that project into the given ellipse [20] (see Figure 9). However, our datasets showcase one solution per image, i.e., the one corresponding to the spacecraft ground-truth pose. Therefore, the network will likely learn a mapping between the ellipse and an average solution to the ellipsoid pose estimation problem (1). One argument to support this explanation is that the performance of the MLP is almost identical whether we input the ground-truth position or the predicted one (see Figures 7 and 10).

7. Conclusion

In this paper, we investigate the use of a generic 3D ellipsoidal model for texture-less spacecraft pose estimation. Despite the interesting mathematical properties of the model, we demonstrate that the set of solutions is large and highly sensitive to noise. Future work will therefore focus on expanding the data to cover the space of solutions better, and on considering a particular case of ellip-

*https://members.loria.fr/VGaudilliere/files/Ellipsoid.mp4

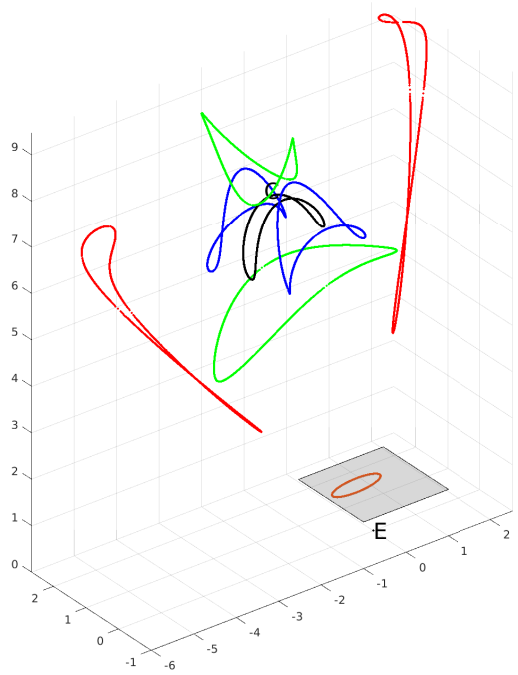

Fig. 9: Loci of the centres (black) and principal axes endpoints (red, green, blue) of the ellipsoids solutions of Equation (1) for a given image. A video is available††.

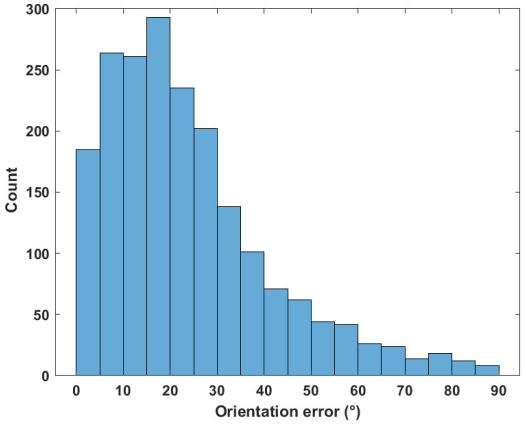

Fig. 10: **AKM** Dataset: **orientation error** histogram. Orientations are predicted from **groundtruth positions** using a **Positive Definite MLP**. Mean: **24.01**°; median: 19.85°; standard deviation: 17.60°. This Figure is to be compared with Fig. 7.

soid having an axis of symmetry (spheroid), since then the pose estimation problem has only two solutions [20].

Acknowledgements

This work was partly funded by the Luxembourg National Research Fund (FNR) under the project reference BRIDGES2020/IS/14755859/MEET-A/Aouada.

References

[1] J. S. Llorente, A. Agenjo, *et al.*, "Proba-3: Precise formation flying demonstration mission," *Acta Astronautica*, vol. 82, no. 1, pp. 38–46, 2013.

[2] P. Bodin, R. Noteborn, *et al.*, "The prisma formation flying demonstrator: Overview and conclusions from the nominal mission," *Advances in the Astronautical Sciences*, vol. 144, pp. 441–460, 2012.

[3] E. IET, *One satellite services another in orbit for the first time*, EN, Apr. 2020.

[4] T. H. Park, S. Sharma, and S. D'Amico, "Towards robust learning-based pose estimation of noncooperative spacecraft," *arXiv preprint arXiv:1909.00392*, 2019.

[5] B. Chen, J. Cao, *et al.*, "Satellite pose estimation with deep landmark regression and nonlinear pose refinement," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 2019, pp. 2816–2824.

[6] S. Sharma and S. D'Amico, "Neural network-based pose estimation for noncooperative spacecraft rendezvous," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4638–4658, 2020.

[7] P. F. Proença and Y. Gao, "Deep learning for spacecraft pose estimation from photorealistic rendering," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 6007–6013.

[8] K. Gerard, "Segmentation-driven satellite pose estimation," Technical Report, Oct. 2019, p. 9.

[9] A. Garcia, M. A. Musallam, *et al.*, "Lspnet: A 2d localization-oriented spacecraft pose estimation neural network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2048–2056.

[10] A. Rathinam and Y. Gao, "On-orbit relative navigation near a known target using monocular vision and convolutional neural networks for pose estimation," in *International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS), Virutal Conference (Pasadena, CA:)*, 2020, pp. 1–6.

[11] M. Kisantal, S. Sharma, *et al.*, "Satellite pose estimation challenge: Dataset, competition design, and results," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 5, pp. 4083–4098, 2020.

[12] T. H. Park, M. Märtens, *et al.*, "Speed+: Next-generation dataset for spacecraft pose estimation across domain gap," in *2022 IEEE Aerospace Conference (AERO)*, 2022, pp. 1–15. DOI: 10.1109/AERO53065.2022.9843439.

[13] X. Zhang, Z. Jiang, *et al.*, "Vision-based pose estimation for textureless space objects by contour points matching," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 5, pp. 2342–2355, 2018. DOI: 10.1109/TAES.2018.2815879.

[14] C. Wu, L. Chen, *et al.*, "Pseudo-siamese graph matching network for textureless objects' 6-d pose estimation," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 2718–2727, 2022. DOI: 10.1109/TIE.2021.3070501.

[15] R. Kaskman, I. Shugurov, *et al.*, "6 dof pose estimation of textureless objects from multiple RGB frames," in *Computer Vision - ECCV 2020 Workshops - Glasgow, UK, August 23-28, 2020, Proceedings, Part II*, A. Bartoli and A. Fusiello, Eds., ser. Lecture Notes in Computer Science, vol. 12536, Springer, 2020, pp. 612–630.

[16] D. Wokes and P. Palmer, "Autonomous pose determination of a passive target through spheroid modelling," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 7493.

[17] M. A. Musallam, A. Rathinam, *et al.*, "Cubesat-cdt: A cross-domain dataset for 6-dof trajectory estimation of a symmetric spacecraft," in *AI4SPACE workshop, European Conference on Computer Vision (ECCV)*, 2022, pp. 1–14.

[18] A. Rathinam, Z. Hao, and Y. Gao, "Autonomous visual navigation for spacecraft on-orbit operations," *Space Robotics and Autonomous Systems: Technologies, Advances and Applications*, vol. 131, p. 125, 2021.

[19] A. Rathinam, V. Gaudilliere, *et al.*, *AKM Dataset: Textureless Space Target Dataset*, Sep. 2022. DOI: 10.5281/zenodo.7043325.

[20] V. Gaudillière, G. Simon, and M.-O. Berger, *Perspective-1-ellipsoid: Formulation, analysis and solutions of the ellipsoid pose estimation problem in euclidean space*, 2022. DOI: 10.48550/ARXIV.2208.12513.

[21] D. Eberly, *Reconstructing an ellipsoid from its perspective projection onto a plane*, https://www.geometrictools.com/, Updated version: March 1, 2008, May 2007.

[22] K. He, X. Zhang, *et al.*, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[23] L. Bottou, "Stochastic gradient learning in neural networks," in *Proceedings of Neuro-Nîmes 91*, Nimes, France: EC2, 1991.

[24] K. Xu, D. Z. Huang, and E. Darve, "Learning constitutive relations using symmetric positive definite neural networks," *Journal of Computational Physics*, vol. 428, p. 110 072, 2021. DOI: https://doi.org/10.1016/j.jcp.2020.110072.

[25] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015.