AI-based Sensor Fusion for Spacecraft Relative Position Estimation around Asteroids

Iain Hall*[1], Jinglang Feng[1], Massimilano Vasile[1]

[1]*Aerospace Centre of Excellence, University of Strathclyde, Glasgow, UK*

Asteroid missions depend on autonomous navigation to carry out operations. The estimation of the relative position of the asteroid is a key step but can be challenging in poor illumination conditions. We explore how data fusion of optical and thermal sensor data using machine learning can potentially allow for more robust estimation of position. Source level fusion of visible images and thermal images using Convolutional Neural Networks is developed and tested using synthetic images based on ESA's Hera mission scenario. It is shown that the use of thermal images allows for improved feature extraction. It also demonstrates that the use of source-level sensor fusion achieves better results than just using thermal images. This results in better identification of the asteroid's centroid but has a much smaller effect on range estimation.

1 Introduction

Missions to visit asteroids require the development autonomous navigation for GNC to allow for maximum scientific return. This is due to the long communication times and the uncertainties in the dynamics of the missions limiting ground based operation. Vision based navigation techniques have been the preferred method for previous missions [1–3], allowing for the estimation of state using visible images.

While vision based navigation techniques have been successful for many asteroid missions [1–3] it does still have limitations. Vision based navigation is less effective when in poor illumination such as eclipse, or when there is a large sun phase angle. To improve the robustness of vision-based navigation to poor illumination conditions sensor data fusion can be employed.

Sensor data fusion can allow for a shortcoming of an individual sensor to be addressed by combining it with complimentary data from other sensors[4]. This makes it very applicable for use in vision-based asteroid navigation where the poor illumination conditions will reduce the effectiveness of visible wavelength cameras. To address this, we consider the fusion of visible camera data and thermal infrared (TIR) camera data in this study as they are both sensors onboard the Hera spacecraft[5] of ESA's Hera mission.

Data fusion will be considered for the HERA mission[6], using the Asteroid Framing Camera, laser altimeter (PALT), and the Thermal Infra Red Imager (TIRI). They will be combined in a high-level fusion level, using hand-crafted processing methods to extract features. However, it does not take advantage of any Deep Learning (DL) techniques, which have been used effectively for pose estimation[7], including artificial satellite pose estimation[8, 9].

In this work we will investigate how deep learning and sensor data fusion can be used together for relative navigation around asteroids. Convolutional Neural Networks (CNN) have been applied to sensor data fusion in guidance for Autonomous ground vehicles[10]. Deep Learning methods have also achieved state-of-the-art performance in the related challenge of artificial satellite pose estimation[8]. Sensor fusion is also being considered for the challenge of satellite pose estimation using Infra-red and LiDAR fusion [11].

We will explore how source-level fusion affects the accuracy with which line-of-sight navigation can be achieved. This is done for spacecraft altitudes of 30km to 8.5km around the asteroid Didymos which is the target body of the Hera mission. It covers both the Early characterisation phase and the Detailed characterisation phase of the mission. The lower limit 8.5km is the minimum distance at which line of sight navigation is considered possible in the requirements[12].

2 Method

2.1 CNN Architecture

The CNN architecture used in this study is HRNet [7]. This network consists of multiple sub-networks at different scales allowing the identification of features at different scales, which are recombined several times

*Corresponding author. E-Mail: iain.hall@strath.ac.uk

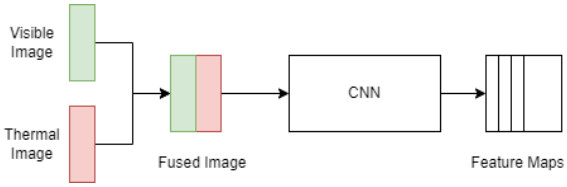

Figure 1: Source level fusion architecture

at multistage fusion layers in the network. Three different HRNet models are trained for 3 different input cases: visible, TIR, and source fusion. The visible and TIR networks both had a single channel input while the source fusion took two channels, one for each data type.

To allow the position and the range to be calculated two outputs were needed: the centroid and the edge of the asteroid. The networks output these as feature maps, which can then be used to find the position and range. The peak pixel in the centroid feature map was taken as the centroid position. In the edge feature map a threshold was applied so only pixels greater than half of the peak pixel (in the edge feature map) were used in the range calculation. The network was trained to minimise the Mean Square Error (MSE) between the label feature maps and the feature maps output by the HRNet models.

The models used transfer learning from the COCO dataset [13] used to originally train HRNet on human pose estimation to speed up the process of training. The models were then trained on a set of 21,000 images for 10 epochs and tested on a separate group of 9,000 test images.

2.2 Fusion Strategies

For source fusion the visible and thermal images are input as a single tensor made from concatenating the two individual image tensors. This approach minimizes computational cost and network complexity for a given network size. This fusion architecture is shown in Figure 1.

2.3 Synthetic Image Generation

Due to the lack of available labelled data from real datasets synthetic data is generated. This is done using Blender [14], an artistic rendering and modelling tool, taking advantage of the physically based Cycles render engine. Blender's open-source nature and python API have led to its use for synthetic training data generation for many other DL-based satellite pose estimation methods[8, 15]. Figure 2 shows the pipeline used for rendering images.

An additional challenge is the need to model the surface temperature of Didymos for producing TIR images. This is done by using MATLAB's partial differential equation library[16] to simulate a transient thermal model of the rotating asteroid. From this the surface temperature can be imported into blender and thermal images rendered. The surface properties and thermal inertia of Didymos used for the thermal simulation were taken from [17].

Figure 3 shows the thermal result when Didymos is at a range of 2AU. The general temperature distribution is similar to [17], with the poles being the coldest but they do not reach as low a temperature. For the purpose of synthetic image generation and testing of the model at this stage it is considered sufficient as it matches the overall temperature behaviour and has a regions where the surface temperature is $< 150^{o}K$, and therefore not visible to the thermal camera. The temperature maps were generated at 3 different distances from the sun, 1AU, 1.5AU, and 2AU.

After the rendering of an initial set of 15,000 image pairs the images undergo affine transforms to produce more variety in orientation and position of the asteroid. This brings the total dataset size up to 30,000 image pairs which was split into a training set and test set. The properties of both cameras are shown in Table 1, with values from [18, 19]. The thermal images are cropped and resized so that they have the same angular pixel resolution ($^{o}/pixel$) as the visible images. The images are generated assuming the visible and thermal images are taken simultaneously

Property	Visible	Thermal	unit
Focal Length, f	106	100	mm
Sensor Width, S	10.2	17.5	mm
Resolution, N_{Pixels}	1020x1020	768x768	pixels
FOV	5.5	10.0	^{o}

Table 1: Camera properties used in image generation.

2.4 Pseudorange Calculation

Calculation of pseudorange was used to estimate the depth of the asteroid. Pseudorange was calculated as defined in [20] based on a pinhole camera model. In Equation (1) we calculate the pseudorange ρ from the focal length, f, the radius of Didymos R, the pixel radius of Didymos in the feature maps N_R, and the size of a pixel in the visible camera, ν. A modification when compared to [20] was to use the average pixel radius from an edge map rather than from a set of keypoint feature maps. The pseudorange and centroid

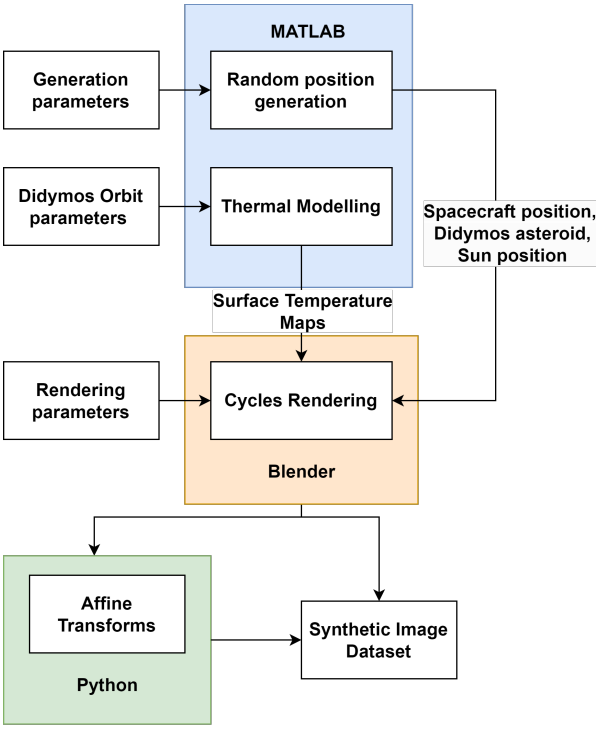

Figure 2: Synthetic Data Generation Pipeline

can then be used to find the position of the asteroid in a camera centred reference frame using trigonometry.

$$\rho = \frac{fR}{N_R \nu} \tag{1}$$

3 Results

For the test set of 9,000 image pairs the overall results of the 3 models are shown in Table 2. The MSE shows how well the different models did at identifying the target features.

Model	Mean MSE	Std. Deviation
Fusion (edge)	0.0973	0.3244
Fusion (centroid)	0.0018	0.0029
Visible (edge)	3.5784	3.7842
Visible (centroid)	0.0112	0.0096
TIR (edge)	0.2009	0.4958
TIR (centroid)	0.0040	0.0037

Table 2: Average MSE and standard deviation in edge and centroid feature maps for the 3 models

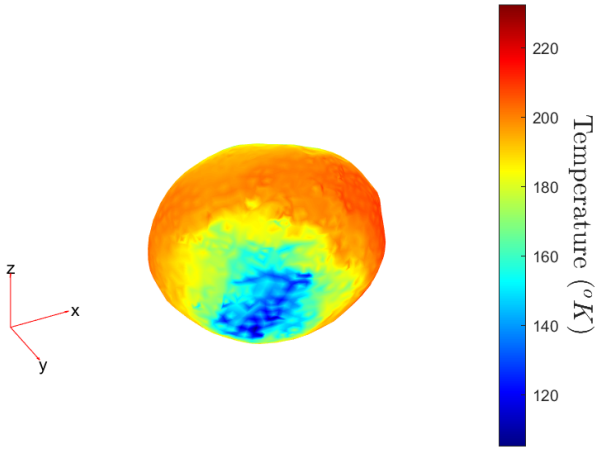

Figure 3: Results of thermal model

3.1 Effect of sun angle

In Figure 4 and Figure 5 it can be seen how the different models perform for different sun angles. Figure 4 shows the variation in centroid prediction accuracy while Figure 5 shows how the accuracy of edge prediction varies. It can be seen from the general trends that the fusion network generally outperforms the visible network and TIR networks for all sun angles. The error of the visible network increases for edge estimation as sun angles increases. We do not see any change with sun angle for centroid estimation error for any of the models.

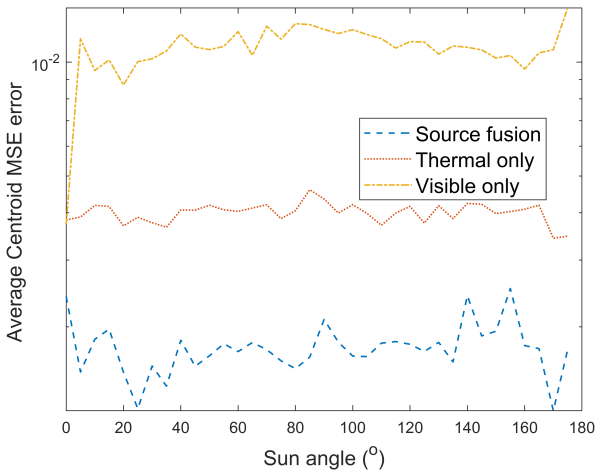

Figure 4: MSE loss for identification of centroid by 3 different networks

This can be seen visually in Figure 8 where the outputs of the different models are shown when tested on an image with a very high sun angle. It can be seen that the visible network struggles to identify the full edge. The Fusion and TIR networks still achieve good results. The visible network does still identify a signif-

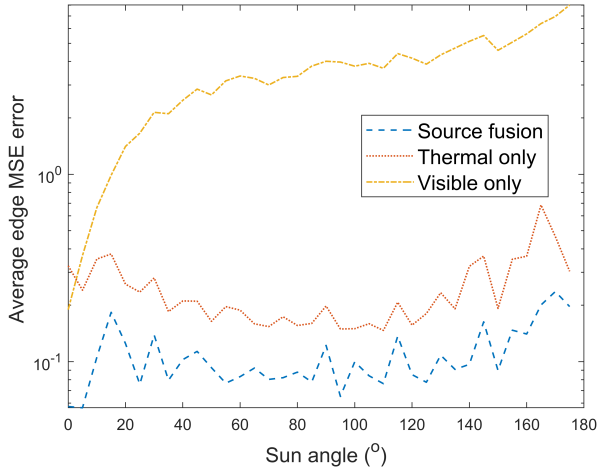

Figure 5: MSE loss for identification of the edge by 3 different networks

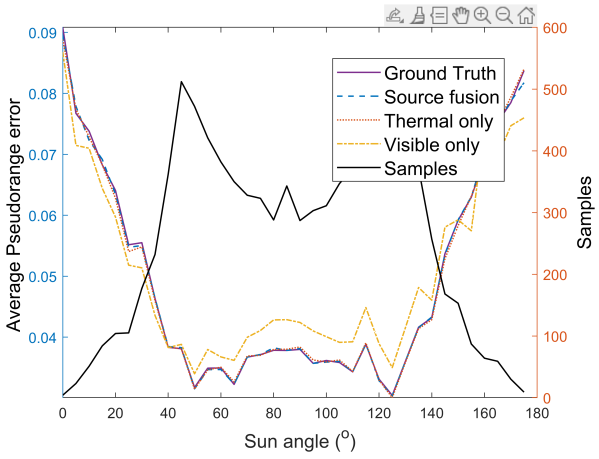

Figure 6: Mean error in Pseudorange estimation at different sun angles

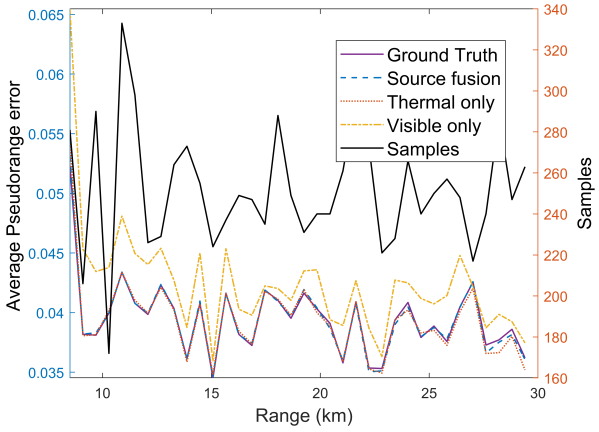

Figure 7: Mean error in Pseudorange estimation at different ranges

icant portion of the edge and identifies the centroid.

3.2 Pseudorange

The pseudorange was calculated for the test images and the results are shown in Table 3, where the pseudorange error presented in the results is normalised to the range the image was generated at, and the centroid pixel error is the distance in pixels between the GT pixel position and pixel position predicted in the feature maps. The ground truth (GT) feature maps were also used to calculate the pseudorange. For the centroid position the Fusion model outperforms the thermal and visible models. For pseudorange there is relatively little difference in the performance of the models with the visible model doing only slightly worse than the fusion and thermal models, which do as well as the GT feature maps. Figure 6 shows how the error in pseudorange varies with sun angle while Figure 7 shows how pseudorange varies with ground truth range. In both the number of samples at a given range or sun angle are shown on the right axis. The pseudorange error shows a slight drop off with range for all 3 models. The models also perform significantly worse at sun angles less than 40^o and greater than 140^o.

Model	Range error	Sd.	Pixel error	Sd.
GT	0.0396	0.0263	0.19	0.22
Fusion	0.0395	0.0264	0.26	0.33
Visible	0.0425	0.0306	0.91	0.63
TIR	0.0395	0.0263	0.32	0.40

Table 3: The Mean error and standard deviation (Sd.) for Pseudorange and centroid pixel error and standard deviation (Sd.) for the 3 models

4 Discussion

From the results shown in Figures 4 and 5 we can see that the use of TIR images in position estimation of an asteroid significantly outperforms centroid and edge estimation than with just using the visible images, especially for higher sun angles. The fusion network also offers better performance than just using just TIR images for finding the edge and centroid.

However, when the edge and centroid are used to estimate the pseudorange the difference in performance is clearly smaller with thermal and fusion models only performing slightly better than visible. When compared to the pseudorange error for the GT feature maps it is clear that the models have reached the minimum error with the current feature map choice. We also see that at sun angles less than 40^o degrees and

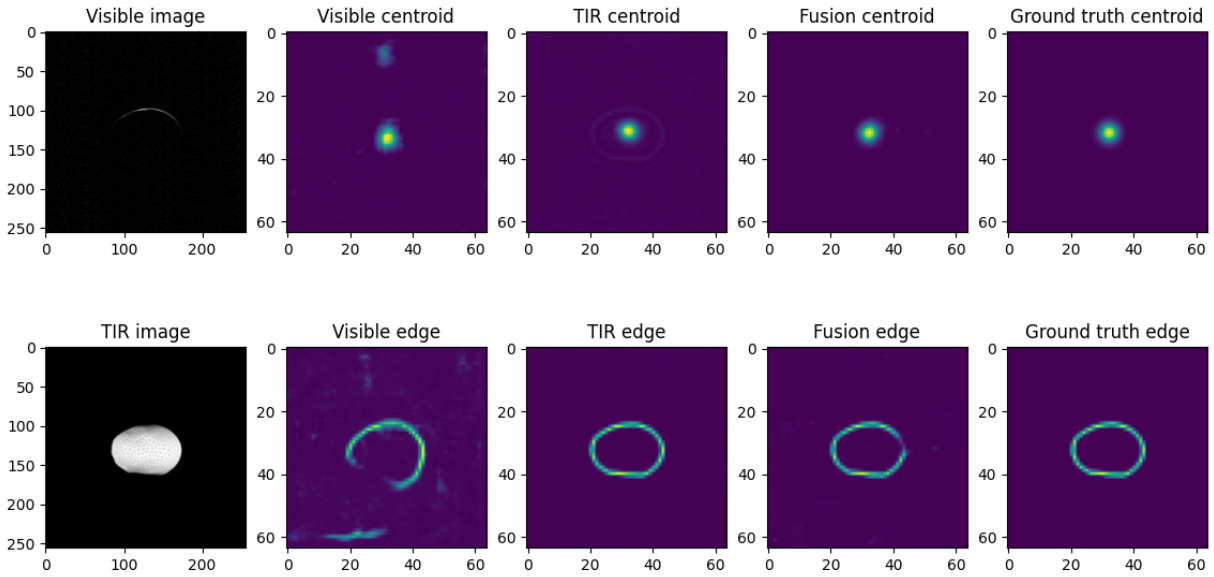

Figure 8: Example results for extreme shadow

greater than 140^o the pseudorange error is higher, and when compared to the number of number of samples we see an inverse relationship, with more samples resulting in better average performance. For pseudorange error against range we see no bias in the sampling and no significant variation in error, reaching a minimum error of approximately 0.04. The choice of a random Cartesian vector for the spacecraft position results in the bias towards angles between 40^o and 140^o, and this disparity in samples is likely what results in the difference in performance.

The inherent error in pseudorange is likely a result of two factors. Firstly is that it relies on the assumption of a spherical asteroid. This will lead to either an overestimation or underestimation of the range depending on the asteroid's orientation. A second limitation arises from the loss in pixel information from the feature maps when compared to the input image. This is because the position of the centroid and edge are quantized in the feature maps, which can lead to a significant error in comparison to the actual position based on how large a physical region is represented by the pixel. At a range of 20km and feature map size of 64x64 the size covered by one pixel is $74m$, resulting in a potential $\pm 34m$ while 1024x1024 it would be $\pm 4.6m$. Further work must be done to assess how feature maps can better capture range.

5 Conclusion

A CNN-based sensor data fusion method has been developed and tested on a range of synthetic images. It has been shown that the use of thermal images has allowed for robustness to lighting conditions when extracting features when compared to visible-only feature extraction. We have also demonstrated that the use of source-level sensor fusion achieves better feature extraction than just using thermal images. This is seen in the extraction of the centroid position but these improvements do not significantly effect the range estimation which is carried out using these features, and exploration of different extracted features is required.

6 Future Work

Further work should develop a data set with a different sampling method to provide more validity to results at high and low sun angles. Other feature map sizes and types of features should be explored which may be able to take advantage of the improved performance in feature extraction achieved by source fusion. Future work should consider employing the centroid and range estimation within a navigation filter to allow for a better assessment of its performance in a real mission scenario.

Validation of the results on real data should be done. This can be done using data from the Hayabusa-2 mission. As it had both visible and thermal images the method can be tested. However as the images are unlabeled the network would need to be tested for position estimation performance using the NASA SPICE kernel data as a ground truth. If this is done further work could consider how the synthetically generated images could be augmented to improve performance on the real image dataset.

References

1. Hashimoto, T. *et al.* Vision-based guidance, navigation, and control of Hayabusa spacecraft - Lessons learned from real operation -. *IFAC Proceedings Volumes* **43,** 259–264. ISSN: 1474-6670. https://www.sciencedirect.com/science/article/pii/S1474667015318504 (2010).
2. ONO, G. *et al.* GNC Design and Evaluation of Hayabusa2 Descent Operations. *TRANSACTIONS OF THE JAPAN SOCIETY FOR AERONAUTICAL AND SPACE SCIENCES, AEROSPACE TECHNOLOGY JAPAN* **19,** 259–265 (Mar. 2021).
3. Lorenz, D. A. *et al. Lessons learned from OSIRIS-REx autonomous navigation using natural feature tracking* in *2017 IEEE Aerospace Conference* (2017), 1–12.
4. Castanedo, F. *A review of data fusion techniques* 2013.
5. *hera instruments* https://www.heramission.space/hera-instruments.
6. Pellacani, A., Graziano, M., Fittock, M., Gil, J. & Carnelli, I. *HERA vision based GNC and autonomy* in *Proceedings of the 8th European Conference for Aeronautics and Space Sciences. Madrid, Spain, 1-4 july 2019* (2019).
7. Sun, K., Xiao, B., Liu, D. & Wang, J. *Deep high-resolution representation learning for human pose estimation* in. **2019-June** (2019), 5686–5696.
8. Park, T. H. *et al.* Satellite Pose Estimation Competition 2021: Results and Analyses. *Acta Astronautica* **204,** 640–665. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576523000048 (2023).
9. Song, J., Rondao, D. & Aouf, N. Deep learning-based spacecraft relative navigation methods: A survey. *Acta Astronautica* **191,** 22–40. ISSN: 00945765 (Feb. 2022).
10. Fayyad, J., Jaradat, M. A., Gruyer, D. & Najjaran, H. Deep Learning Sensor Fusion for Autonomous Vehicle Perception and Localization: A Review. *Sensors* **20.** ISSN: 1424-8220. https://www.mdpi.com/1424-8220/20/15/4220 (2020).
11. Creagh, C. *Pose Estimation of an Unknown Uncooperative Target Using LiDAR & Thermal Infrared Fusion* PhD thesis (Delft University of Technology, 2024).
12. Michael Küppers, J. G. F. *HERA MISSION REQUIREMENTS DOCUMENT* tech. rep. (ESA, Science Operations Department, 2021). https://static1.squarespace.com/static/5fb661af793de226cf27e2a5/t/629f44ddec844b33d2dc5a7e/1654605022569/ESA-HERA-TECSH-RS-022793+Mission+Requirements+v2.10.pdf.
13. Lin, T.-Y. *et al. Microsoft COCO: Common Objects in Context* 2015. arXiv: 1405.0312 [cs.CV].
14. *Blender* https://www.blender.org/.
15. Guthrie, B., Kim, M., Urrutxua, H. & Hare, J. Image-based attitude determination of co-orbiting satellites using deep learning technologies. *Aerospace Science and Technology* **120,** 107232. ISSN: 1270-9638. https://www.sciencedirect.com/science/article/pii/S1270963821007422 (2022).
16. MathsWorks. *MATLAB Partial Differential Equation Toolbox* https://uk.mathworks.com/products/pde.html (2024).
17. Rozitis, B. *et al.* Pre-impact Thermophysical Properties and the Yarkovsky Effect of NASA DART Target (65803) Didymos. **5,** 66 (Mar. 2024).
18. ESA. *Hera Instruments* https://www.heramission.space/hera-instruments (2024).
19. Okada, T. *et al. Thermal Infrared Imager Onboard Hera to Observe S-Type Binary Asteroid Didymos* in *52nd Lunar and Planetary Science Conference* (Mar. 2021), 1449.
20. Kaluthantrige, A., Feng, J. & Gil-Fernández, J. CNN-based Image Processing algorithm for autonomous optical navigation of Hera mission to the binary asteroid Didymos. *Acta Astronautica* **211,** 60–75. ISSN: 0094-5765. https://www.sciencedirect.com/science/article/pii/S0094576523002709 (2023).