

DEEP FEATURE DETECTION AND DESCRIPTION FOR SMALL BODY RELATIVE NAVIGATION

Travis Driver^{1*} and Panagiotis Tsiotras¹; ¹Georgia Institute of Technology, Atlanta, GA, *travisdriver@gatech.edu

Abstract. *Missions to small bodies rely heavily on visual feature tracking for characterization and relative navigation of the target body. While deep learning has led to great advancements in computer vision, training and validating data-driven models for space applications is challenging due to the limited availability of relevant annotated data. Therefore, this paper introduces a large-scale dataset comprised of over 100,000 annotated, real images of 15 different small bodies captured during past and ongoing missions. First, we perform an exhaustive evaluation of both handcrafted and data-driven methods for feature detection and description on small body imagery. Finally, we train a deep feature detection and description network using our data and demonstrate increased performance.*

Introduction. There has been an increasing interest in missions to small bodies due to their great scientific value. In addition to planetary protection and resource utilization, small bodies (e.g., asteroids, comets) are believed to be remnants from the solar system’s formation, and studying their composition could provide insight into the solar system’s evolution and the origins of organic life on Earth.¹ Due to their irregular shape, variable surface properties, and long round-trip communication time, accurate *autonomous* navigation is critical for safe and precise maneuvering in the vicinity of a small body.

The current state-of-the-practice in small body relative navigation relies on a manual approach where human operators extract salient surface features from images acquired during an extended characterization phase to estimate a collection of digital terrain maps (DTMs), local topography and albedo maps. These topographic features are then uplinked to the spacecraft and correlated with onboard images to produce a navigation solution.² This lack of autonomy in current mission procedures limits mission capabilities and increases operational costs.

While automated feature tracking methods have been investigated for autonomous small body relative navigation,³ researchers have had trouble capitalizing on recent advances in deep learning due to the unavailability of relevant, annotated data.^{4,5} To the best of our knowledge, there exists no large-scale, annotated dataset comprised entirely of *real* small body images. Indeed, previous work has either relied entirely on simulated data⁶ or uncomprehensive (i.e., <150 images or restricted to a single body) sets of real imagery.^{4,7} Moreover, operation in space presents unique environmental (e.g., dynamic hard lighting) and operational (e.g., significant scale and perspective changes) challenges that are likely not adequately captured in datasets based on terrestrial imagery.

Therefore, this paper presents a large-scale dataset comprised of over 110,000 densely annotated, real images

of 16 different small bodies from legacy and ongoing deep space missions to facilitate the study of deep learning for autonomous navigation in the vicinity of a small body. The contributions of this paper are as follows: (i) we develop a novel benchmarking suite and evaluate both handcrafted and data-driven feature detection and description methods on real small body imagery; (ii) we perform end-to-end training of a state-of-the-art, deep feature detection and description network on small body imagery and demonstrate increased performance on our benchmarks.

Data Generation. We leverage publicly available images and ancillary data (i.e., camera pose, camera calibration, shape models) from both legacy and active small body missions provided through NASA’s Planetary Data System (PDS).⁸ Shape models (i.e., watertight, triangular surface meshes) are developed as part of the relative navigation pipeline of small body missions, typically constructed using stereophotoclinometry.⁹ We leverage high-resolution shape models to generate dense depth maps for each image, which are computed by backward-projecting rays at each pixel in the image and recording the intersection depth between the ray and shape model. We also estimate a mask of the occluded regions of the imaged surface for training (see Figure 1).

Evaluation. Performance is evaluated on a per image pair basis using the metrics $Precision = \#Inlier\ Matches / \#Putative\ Matches$, $Recall = \#Inlier\ Matches / \#Ground\ Truth\ Matches$, and $Accuracy = \#Correct\ Matches \& \ Nonmatches / \#Features$. Each method is limited to extract 5,000 features, and putative matches are computed using mutual nearest neighbors (MNN). Matches are verified by projecting each keypoint in the first image into the second using the ground truth pose, calibration, and depth map, and matches are verified by checking that the projected coordinates are within 5 pixels of its match. Ground truth matches are registered if there exists a keypoint within 5 pixels of the projected image coordinate.

We benchmark both *handcrafted* (i.e., ORB¹⁰ and SIFT¹¹) and *data-driven* (i.e., SuperPoint¹² and ASLFeat¹³) features. These results are summarized in Table 1. SIFT demonstrates competitive performance, but suffers when applied to datasets with harsher illumination (i.e., Rosetta, OSIRIS-REx). SuperPoint achieves high recall but low precision and generally underperforms with respect to all other methods except ORB. ASLFeat consistently ranks among the top performing methods with respect to all datasets. Therefore, we selected ASLFeat for end-to-end training.

Learning Features from Small Body Imagery. We train ASLFeat using a procedure similar to the original implementation,¹³ where we use an approxi-

Table 1. Feature performance with respect to precision (P), recall (R), and accuracy (A).

Dataset	# Images	Feature	# Matches	P	R	A
Cassini @ Epimetheus (Saturn XI)	133	ORB	890	14.2	9.2	5.9
		SIFT	193	28.7	26.1	16.4
		SuperPoint	<u>405</u>	11.8	<u>21.6</u>	4.9
		ASLFeat	348	<u>24.3</u>	20.2	<u>7.9</u>
Cassini @ Mimas (Saturn I)	307	ORB	1115	10.8	6.3	5.2
		SIFT	<u>505</u>	<u>21.6</u>	<u>16.9</u>	<u>11.5</u>
		SuperPoint	178	14.5	17.4	8.1
		ASLFeat	487	28.0	16.6	13.3
Dawn @ 1 Ceres	4182	ORB	1648	23.7	23.9	9.4
		SIFT	1912	<u>43.8</u>	57.0	<u>20.7</u>
		SuperPoint	513	42.6	71.5	20.3
		ASLFeat	<u>1865</u>	57.1	<u>64.7</u>	24.2
Dawn @ 4 Vesta	1922	ORB	<u>1262</u>	21.4	19.0	8.3
		SIFT	1097	<u>41.1</u>	46.7	<u>18.5</u>
		SuperPoint	381	39.7	55.7	18.3
		ASLFeat	1636	57.4	<u>53.7</u>	23.5
Hayabusa @ 25143 Itokawa	603	ORB	660	2.2	1.1	0.8
		SIFT	207	4.1	3.0	2.0
		SuperPoint	75	<u>6.3</u>	10.3	<u>3.6</u>
		ASLFeat	<u>272</u>	9.9	<u>6.3</u>	4.0
OSIRIS-REx @ 101955 Benu	2613	ORB	1631	7.2	4.8	3.5
		SIFT	<u>1491</u>	12.9	11.1	6.9
		SuperPoint	635	<u>15.4</u>	<u>18.1</u>	<u>9.3</u>
		ASLFeat	1365	22.5	18.9	12.1
Rosetta @ 67P	1465	ORB	1132	1.8	0.8	0.6
		SIFT	<u>724</u>	5.3	3.6	2.2
		SuperPoint	218	<u>6.3</u>	8.1	<u>2.8</u>
		ASLFeat	480	13.0	<u>7.3</u>	3.7
Rosetta @ 21 Lutetia	40	ORB	<u>467</u>	12.1	6.8	4.6
		SIFT	261	23.7	25.6	10.8
		SuperPoint	382	<u>26.7</u>	29.1	<u>13.9</u>
		ASLFeat	971	42.9	<u>26.2</u>	17.2

mate 90/10 train/test split. We withhold data corresponding to 4 different small bodies with variable surface characteristics from training, i.e., Epimetheus, Mimas, 25143 Itokawa, and 21 Lutetia. This will test the network’s ability to reliably compute features upon arrival at a previously unexplored body. The network is also tested on held-out images of small bodies it observed during training to emulate a scenario in which images obtained during earlier stages of a mission could be used to train the network for feature extraction in later phases. The ASLFeat model trained on small body images, i.e., ASLFeat-CVGBESM, is compared against the pretrained model. These results are shown in Table 2. ASLFeat-CVGBESMU consistently outperforms the pretrained model with respect to our metrics, with the exception of Rosetta @ Lutetia. Importantly, our model achieves higher matching precision and recall on the novel testing instances.

Acknowledgements. This work was supported by a NASA Space Technology Graduate Research Opportunity. The authors would like to thank Kenneth Getzandanner (NASA GSFC) for several helpful discussions.

References.

- [1] M. A. Barucci, E. Dotto, and A. C. Lvasseur-Regourd, “Space missions to small bodies: asteroids and cometary nuclei,” *Astronomy and Astrophysics Review*, vol. 19, no. 48, pp. 1–29, 2011.

Table 2. AstroVision-trained model compared to pretrained.

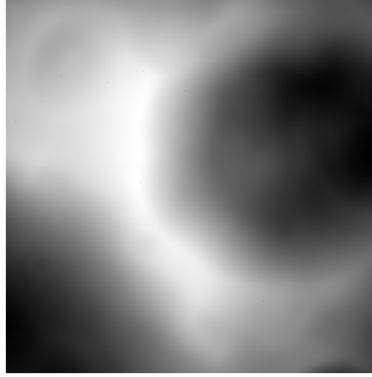
Dataset	Feature	# Matches	P	R	A	AUC		
						@5°	@10°	@20°
Cassini @ Epimetheus (Saturn XI) [†]	ASLFeat	386	27.4	29.0	74.7	2.7	8.2	13.7
	ASLFeat-CVGBESMU	396	28.9	27.5	74.1	2.7	8.6	14.0
Cassini @ Mimas (Saturn I) [†]	ASLFeat	372	21.8	15.7	65.3	0.2	0.2	0.3
	ASLFeat-CVGBESMU	328	23.6	14.9	67.1	0.0	0.1	0.2
Dawn @ 1 Ceres	ASLFeat	1535	48.4	67.8	80.2	12.9	27.1	42.4
	ASLFeat-CVGBESMU	1514	52.8	71.5	82.1	15.9	31.6	46.9
Dawn @ 4 Vesta	ASLFeat	1524	59.0	66.1	84.3	17.5	31.9	46.0
	ASLFeat-CVGBESMU	1412	70.3	69.7	87.4	17.5	33.0	48.7
Hayabusa @ 25143 Itokawa [†]	ASLFeat	338	13.5	11.3	47.5	2.2	4.2	7.6
	ASLFeat-CVGBESMU	363	15.2	11.0	53.7	2.9	5.0	8.8
OSIRIS-REx @ 101955 Benu	ASLFeat	1378	33.1	30.9	68.7	8.0	14.4	20.9
	ASLFeat-CVGBESMU	858	34.2	28.4	79.5	6.7	12.6	19.3
Rosetta @ 67P	ASLFeat	1147	25.0	24.0	62.8	3.4	6.4	10.6
	ASLFeat-CVGBESMU	837	30.4	23.9	69.8	4.2	7.9	13.4
Rosetta @ 21 Lutetia [†]	ASLFeat	970	42.9	35.0	71.9	6.0	12.1	23.8
	ASLFeat-CVGBESMU	778	41.3	31.1	76.3	8.4	13.2	22.3

[†] No images of this body were included in the training set

- [2] D. A. Lorenz, R. Olds, A. May, C. Mario, M. E. Perry, E. E. Palmer, and M. Daly, “Lessons learned from OSIRIS-REx autonomous navigation using natural feature tracking,” in *IEEE Aerospace Conf.*, (Big Sky, MT, USA), pp. 1–12, 2017.
- [3] B. J. Morrell, J. Villa, and A. Havard, “Automatic feature tracking on small bodies for autonomous approach,” in *ASCEND*, pp. 1–15, 2020.
- [4] T. Fuchs, D. R. Thompson, B. D. Bue, J. Castillo-Rogez, S. A. Chien, D. Gharibian, and K. L. Wagstaff, “Enhanced flyby science with onboard computer vision: Tracking and surface feature detection at small bodies,” *Earth and Space Science*, vol. 2, no. 10, pp. 417–434, 2015.
- [5] J. Song, D. Rondao, and N. Aouf, “Deep learning-based spacecraft relative navigation methods: A survey,” *Acta Astronautica*, vol. 191, pp. 22–40, 2022.
- [6] M. Pugliatti, M. Maestrini, P. Di Lizia, and F. Topputo, “On-board small-body semantic segmentation based on morphological features with U-Net,” in *AAS/AIAA Space Flight Mechanics Meeting*, pp. 1–20, 2021.
- [7] H. Lee, H.-L. Choi, D. Jung, and S. Choi, “Deep neural network-based landmark selection method for optical navigation on lunar highlands,” *IEEE Access*, vol. 8, pp. 99010–99023, 2020.
- [8] “NASA PDS.” <https://pds.nasa.gov/>.
- [9] R. W. Gaskell, O. S. Barnouin-Jha, D. J. Scheeres, A. S. Konopliv, T. Mukai, S. Abe, J. Saito, M. Ishiguro, T. Kubota, T. Hashimoto, J. Kawaguchi, M. Yoshikawa, K. Shirakawa, T. Kominato, N. Hirata, and H. Demura, “Characterizing and navigating small bodies with imaging data,” *Meteoritics & Planetary Science*, vol. 43, no. 6, pp. 1049–1061, 2008.
- [10] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” in *ICCV*, pp. 2564–2571, 2011.
- [11] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *IJCV*, vol. 60, pp. 91–110, Nov. 2004.
- [12] “SuperPoint: Self-supervised interest point detection and description,” in *CVPR Workshops*, 2018.
- [13] Z. Luo, L. Zhou, X. Bai, H. Chen, J. Zhang, Y. Yao, S. Li, T. Fang, and L. Quan, “ASLFeat: Learning local features of accurate shape and localization,” in *CVPR*, pp. 6589–6598, 2020.



(a) Image



(b) Depth map



(c) Mask

Figure 1. Example data products.

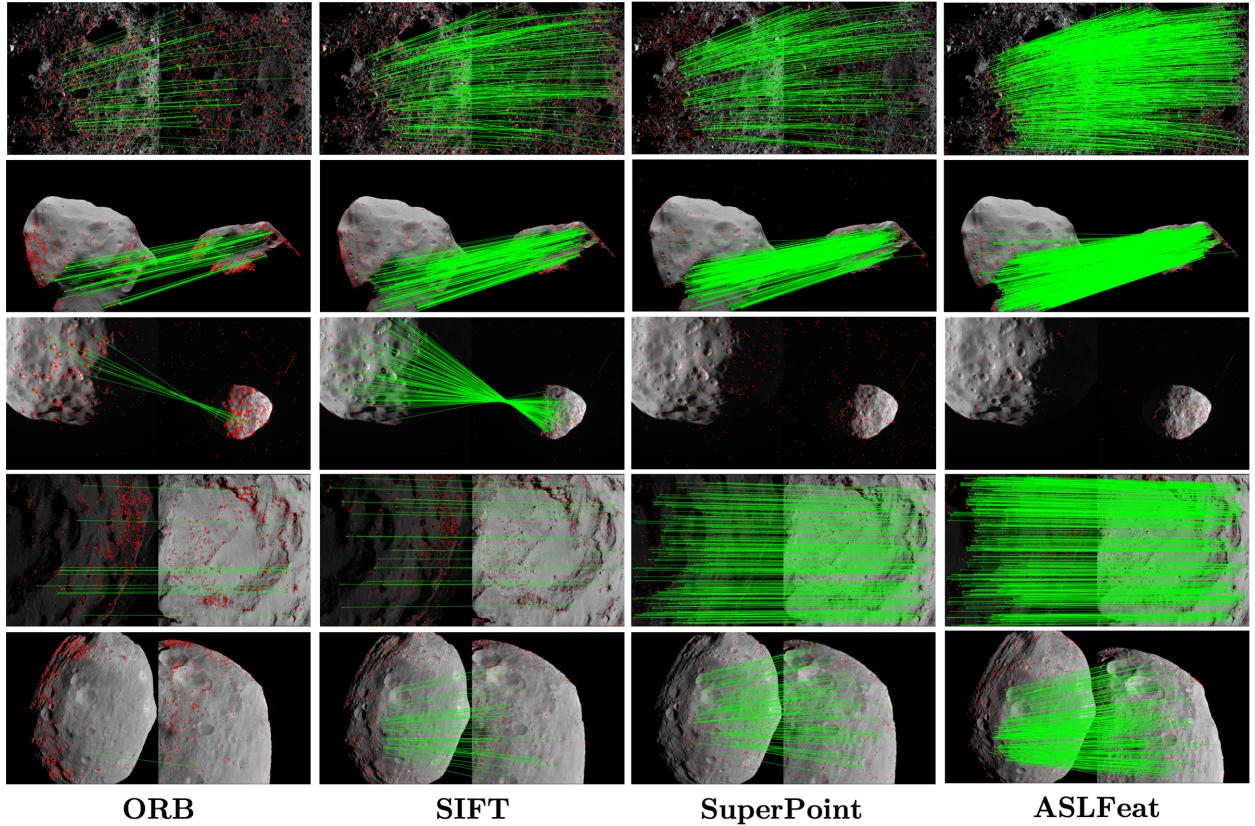


Figure 2. Qualitative comparison of matching performance on OSIRIS-REx @ 101955 Bennu (1st row), Rosetta @ 21 Lutetia (2nd row), Cassini @ Epimetheus (3rd row), Rosetta @ 67P (4th row), and Dawn @ 4 Vesta (5th row). Correct matches are drawn in green and the keypoints of incorrect matches are drawn in red.