POINT CLOUD VISUAL SLAM FOR AUTONOMOUS NAVIGATION AND MAPPING AROUND SMALL CELESTIAL BODIES

Jacopo Villa[1*]; Jay W. McMahon[1]; Issa A. Nesnas[2]; [1]Colorado Center for Astrodynamics Research, University of Colorado Boulder, 3775 Discovery Dr, Boulder, CO 80303, [2]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109. [*]Jacopo.Villa@colorado.edu

Abstract. *In this paper, we present a landmark-free technique to simultaneously localize a spacecraft around a small celestial body and map the target shape, solely using optical data. By triangulating surface features observed through a small stereo baseline, a surface point cloud is estimated. Multiple point-cloud observations can be registered with each other and used to both build a reference shape model and localize the camera with respect to such a reference map. This method does not rely on identifying and matching surface landmarks through multiple images, and hence it is more robust to lighting conditions and observing geometries.*

Introduction. Navigating in the vicinity of small celestial bodies, such as asteroids and comets, is a crucial yet demanding task required for flybys, close approach, proximity operations, and landing. Optical data from the onboard cameras are commonly used in this context as they provide accurate surface-relative observations. The latter are usually angle measurements of surface reference points, also known as landmarks, relative to the observing camera. The identification, selection, and tracking of known landmarks over multiple images is a standard and effective approach which has been used for decades to estimate the spacecraft motion as well as the target shape model, notably, using Stereophotoclinometry (SPC).[1] However, the surface appearance from the camera perspective is subject to rapid and continuous evolution due to changes in surface lighting and observing geometry, caused by the spacecraft-surface-Sun relative motion. As such, identifying the same landmark locations over multiple images is known as a challenging task. Typically, this is performed by expert ground operators by making use of prior knowledge such as predicted images and the spacecraft orbit, and by assessing the quality of the downloaded images. Hence, current landmark-based navigation largely relies on human support and supervision.

The OSIRIS-REx mission pioneered a semi-autonomous landmark-matching technique for visual navigation during the Touch-and-Go maneuver, called Natural Feature Tracking.[2] However, this method relies on a very high-resolution shape model of the landing site and its surroundings, as well as the selection of a high-quality set of landmarks, both obtained through months of ground-based operations. In previous work, Simultaneous Localization and Mapping (SLAM), and in particular its feature-based[1] implementations, repeatedly emerged as a successful framework for surface-relative navigation at small bodies.[3,4] However, the proposed approaches were only tested on short time frames (e.g., a single small-body rotation), where changes in lighting and observing geometries are small. For example, image sets from the approach phase were used, where changes in spacecraft latitude and Sun phase angle are small. While such feature-based approaches may successfully perform loop closure and localization in the above scenarios, previous work did not tackle the more general problem involving surface observations from arbitrary viewpoints and lighting conditions, which is commonly encountered during proximity operations (e.g., for surface flybys, landing, and most non-hovering orbits). Additionally, it has been shown that matching performance using state-of-the-art visual features (such as SURF and ORB, used in several asteroid-SLAM studies) is poor for severe changes of the surface appearance.[5] Given the challenges and limitations associated to using traditional feature matching for autonomous localization in this environment, we propose a landmark-free SLAM method based on vision-based point-cloud registration. Point clouds are very common datasets in robotics, and LIDAR-based point clouds have been shown to enable robust and autonomous navigation and mapping at small bodies in the presence of uncertainty in the environment.[6] The objective of this work is to develop similar capabilities for vision-based point clouds.

Method.

Map Building. A set of images, to be processed in pairs of consecutive frames, is given. The stereo baseline between images in each pair should be small enough so that accurate feature tracking can be performed. In the first image of each pair, Shi-Tomasi[7] features are detected and then tracked to the second image in the pair, using the Kanade-Lucas-Tomasi (KLT) algorithm.[8,9] The idea is that Shi-Tomasi features are usually very dense in textured images such as small-body surfaces, and hence the corresponding tracks can be used to estimate a dense point cloud. The relative camera pose between the two views is estimated and the 3D point cloud is obtained by triangulating the tracked features. Features with large reprojection errors are rejected. Each image pair provides an estimated point cloud, which is registered to the previous one using the Iterative Closest Point (ICP) algorithm.[10] Each view is added to the pose graph and loop closure is currently performed by assuming knowledge of the small-body rotation rate. After loop closure, pose-graph optimization is executed, and the point clouds for each image pair are aligned, downsampled, and denoised.

[1]In this context, a *feature* represents the optical data corresponding to a surface landmark.

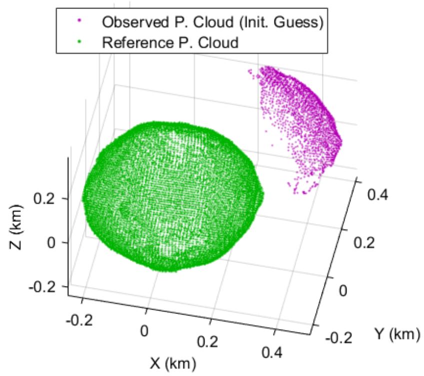

Figure 1. ***Reference point cloud estimated during the map-building step compared to a single point cloud observed from a camera view. A large error between the point-cloud positions is assumed, i.e., a bad initial guess in camera position is used for point-cloud registration.***

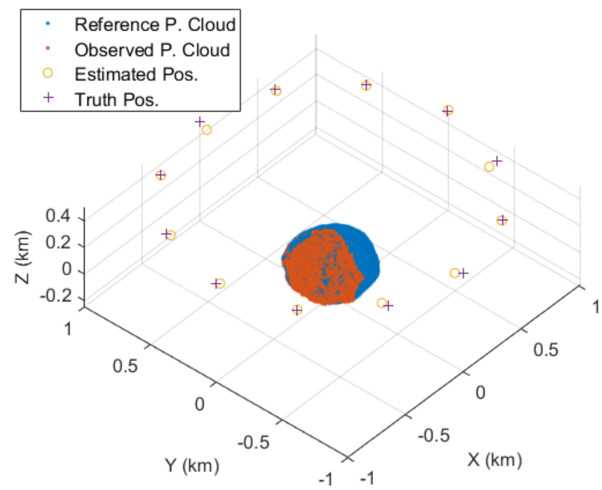

Figure 2. ***Camera position estimates compared to the ground truth, for the localization step. The observed point cloud from a single view is overlaid with the reference point cloud.***

Localization. Once the reference point cloud is constructed, the spacecraft is localized by estimating a point cloud in view from any pair of images, as previously described. Then, ICP is performed to compute the rigid transformation between the observed and the reference point cloud, which in turn is used to localize the observing camera with respect to the map. In this process, the spacecraft attitude is assumed to be known, i.e., ICP only estimates the translation between the two point clouds and not on the rotation, which is given. Current results suggest that similar performance is obtained if this assumption is relaxed.

Results. We tested the point-cloud SLAM algorithm using a simulated scenario around the asteroid Bennu. Two image sets with very different appearance are generated: a lower-latitude, lower Sun phase hovering trajectory for the mapping step (Figure 4) and a higher latitude, higher Sun phase hover for the localization step (Figure 3). A longitudinal displacement of 1° is used between consecutive frames. A dense set of tracked features is obtained, as shown in Figure 5. The obtained reference and observed point clouds are compared in Figure 1. For the localization step, a bad initial guess in the camera position is assumed, to assess the robustness of the technique. Finally, localization performance is shown in Figure 4: despite the difference in appearance between the mapping and localization image sets, the camera position is accurately estimated. Notice that only the higher-latitude portion of the point cloud is visible in the localization image set, suggesting that this technique is suitable for scenarios where only small portions of the map are visible.

Acknowledgements. A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. ©2022 California Institute of Technology. Government sponsorship acknowledged.

References.

[1] R. Gaskell, O. Barnouin-Jha, D. J. Scheeres, A. Konopliv, T. Mukai, S. Abe, J. Saito, M. Ishiguro, T. Kubota, T. Hashimoto, *et al.*, "Characterizing and navigating small bodies with imaging data," *Meteoritics & Planetary Science*, vol. 43, no. 6, pp. 1049–1061, 2008.

[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 *2017 IEEE Aerospace Conference*, pp. 1–12, IEEE, 2017.

[3] M. Dor, K. A. Skinner, T. Driver, and P. Tsiotras, "Visual slam for asteroid relative navigation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2066–2075, 2021.

[4] D. Nakath, J. Clemens, and C. Rachuy, "Active asteroid-slam," *Journal of Intelligent & Robotic Systems*, vol. 99, no. 2, pp. 303–333, 2020.

[5] B. J. Morrell, J. Villa, S. Bandyopadhyay, D. Lubey, B. Hockman, S. Bhaskaran, D. Bayard, and I. Nesnas, "Autonomous feature tracking for autonomous approach to a small body," in *ASCEND 2020*, p. 4151, 2020.

[6] B. Bercovici, *Mapping and Navigation of Small Bodies in the Presence of Uncertainty.* PhD thesis, University of Colorado at Boulder, 2019.

[7] J. Shi *et al.*, "Good features to track," in *1994 Proceedings of IEEE conference on computer vision and pattern recognition*, pp. 593–600, IEEE, 1994.

[8] B. D. Lucas, T. Kanade, *et al.*, "An iterative image registration technique with an application to stereo vision," Vancouver, 1981.

[9] C. Tomasi and T. Kanade, "Detection and tracking of point," *Int J Comput Vis*, vol. 9, pp. 137–154, 1991.

[10] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611, pp. 586–606, Spie, 1992.

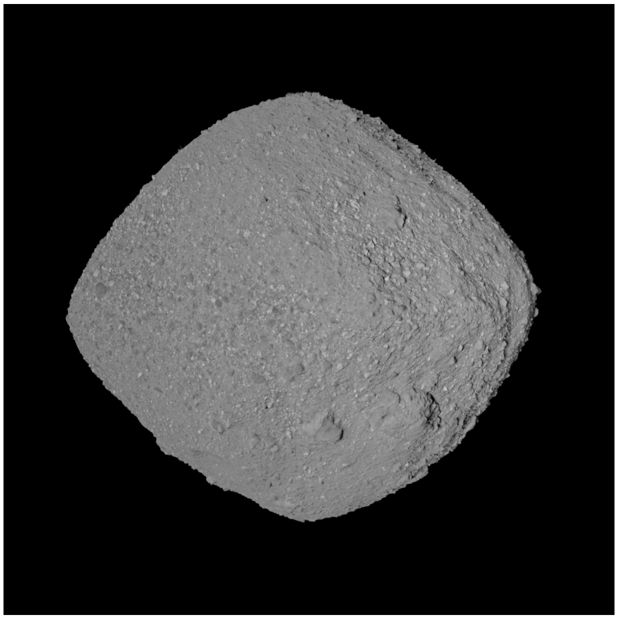

Figure 3. Sample simulated image of the asteroid Bennu used for the map-building step. The observing camera radius from the asteroid center is 1 km, its latitude is 0° *and the Sun phase angle is* 20°*. The camera field-of-view is* 40°*.*

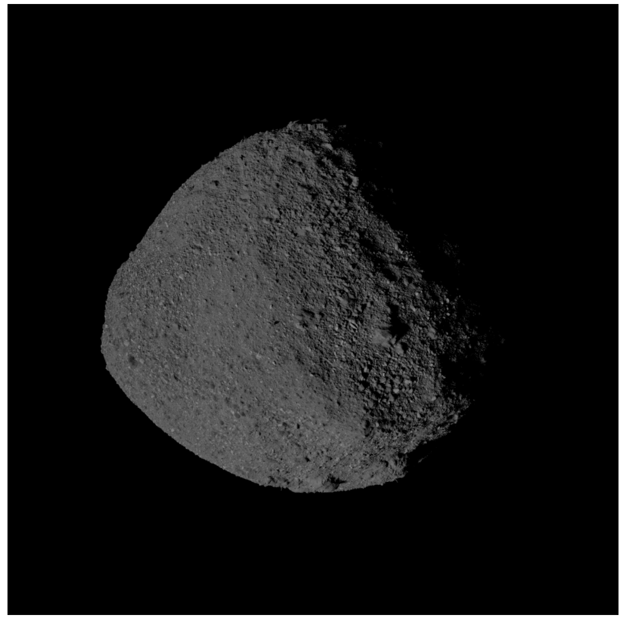

Figure 4. Sample simulated image of the asteroid Bennu used for the localization step. The observing camera radius from the asteroid center is about 1.1 km, its latitude is about 27° *and the Sun phase angle is* 40°*. The camera field-of-view is* 40°*.*

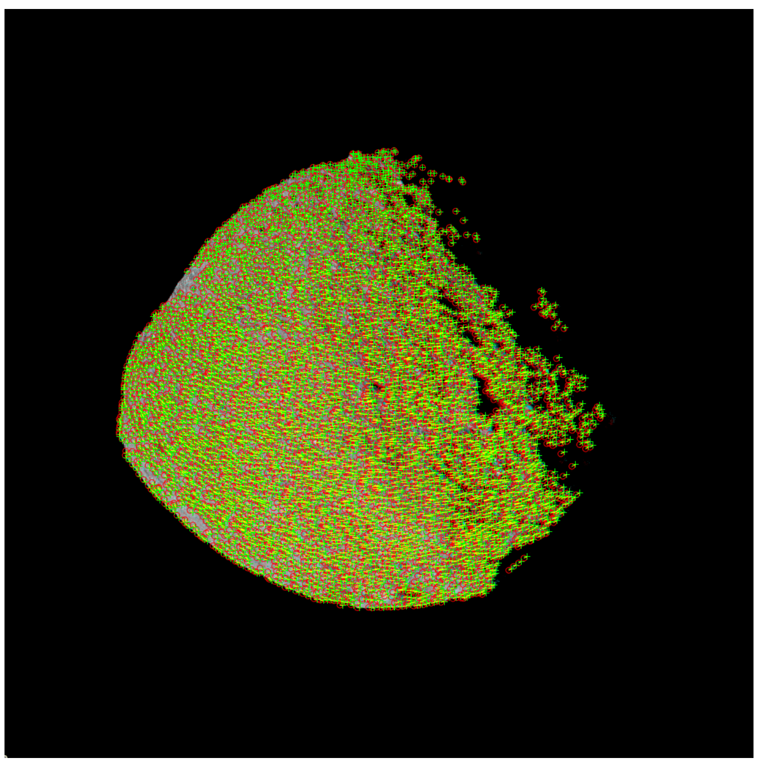

Figure 5. Feature tracks (yellow lines) obtained using the KLT algorithm to track features from the first frame (red circles) to the second frame (green crosses), using an image pair from the localization dataset. Such feature tracks are used to estimate the surface dense point cloud. This figure highlights the high density of Shi-Tomasi features for this type of small-body surfaces.