# Visual SLAM for Asteroid Relative Navigation

Mehregan Dor[1] Katherine A. Skinner[2] Travis Driver[1]
Panagiotis Tsiotras[1]
[1]Georgia Institute of Technology [2]University of Michigan
{mehregan.dor,travisdriver,tsiotras}@gatech.edu kskin@umich.edu

## Abstract

*This paper focuses on the application of visual SLAM for the purpose of precise autonomous navigation around an asteroid. We develop a factor graph-based approach allowing for incremental growth and fusion of sensor orientation measurements, Earth-relative inertial position measurements, as well as in-situ monocular camera imagery-based measurements, with an emphasis on the initialization step. Crucially, and in contrast to typical simulated scenarios found in the literature, we validate our approach using real imagery from NASA's DAWN mission to asteroid Vesta, along with navigation comparison data from the NASA NAIF SPICE kernels. Quantitative comparisons show impressive accuracy for a typical target characterization phase segment, both in terms of the estimated trajectory as well as in terms of the tracked estimated landmarks. Based on these results, this paper further supports the viability of autonomous SLAM-based navigation for deep-space asteroid missions.*

## 1. INTRODUCTION

In recent years, with ever improving navigation solutions, space missions have successfully performed daring firsts in near-asteroid navigation. Orbiter NEAR Shoemaker's controlled asteroid touchdown (1996) [31], HAYABUSA's touchdown and successful sample return (2003) [39], DAWN's orbiting of two celestial bodies in a single mission (2007) [19] and the recent OSIRIS-REx's Touch-and-Go (TAG) operation leveraging Natural Feature Tracking (NFT) with high navigation solution accuracy during descent [4], are only few of the most notable feats accomplished thanks to autonomous navigation.

However, the high-risk nature of near asteroid missions, paired with a lack of autonomy in current mission procedures, severely limits the possibilities in mission design [35]. Indeed, ground-segment operators are intimately involved in all in-situ tasks, which ultimately rely on human-in-the-loop verification, as well as ground-based computations for estimation, guidance and control [38]. In addition, long round-trip light times and severely limited bit-rate in communications render ground-in-the-loop processes extremely tedious.

Clearly, precise relative navigation techniques, with increased autonomy, will be a key enabling element of future asteroid orbiter missions [8, 12]. Firstly, good navigation can inform safe and efficient path planning, control execution and maneuvering [5]. In near asteroid missions, achieving fuel-efficiency during non-critical maneuvers and guaranteeing execution of safety-critical maneuvers both require precise knowledge of the whereabouts of the spacecraft with respect to the target. Secondly, precise navigation situates the acquired science data. Indeed, scientists and mission planners design science acquisition phases based on the expected scientific value of instrument data acquired at predetermined times, on specific orbits and with specific spacecraft orientations [24]. Thirdly, precise navigation facilitates the detailed mapping and shape reconstruction of the target asteroid, since good knowledge of the spacecraft relative position and orientation with respect to the target, as well as a good knowledge of the Sun light direction, are crucial in commonly used shape reconstruction solutions [15]. Finally, good estimates of the spacecraft state enable precise characterization of the target asteroid's spin state, mass moment values and gravity model [24].

To contrast with the current human-centric mindset, this paper proposes a viable autonomous navigation approach for near asteroid operations based on asteroid imagery collected by the spacecraft on-board cameras. Specifically, the contributions of this work are the following: (a) we formulate a precise autonomous vision-based navigation scheme based on simultaneous localization and mapping (SLAM) and sensor fusion; and (b) we demonstrate successful SLAM on real imagery obtained in-situ from a previously flown asteroid orbiter mission.

This paper is organized as follows: Section 2 discusses relevant prior work and further delimits the contributions of this paper; Section 3 details the proposed technical ap-

proach; Section 4 presents experiments and results to validate the proposed approach, and, lastly, Section 5 provides a discussion of the results with conclusions and future work.

## 2. RELATED WORK

Filter-based methods [27], such as the Extended Kalman Filter (EKF), have traditionally been applied to perform multi-sensor fusion for precise navigation purposes. Bercovici et al. [3] proposed a Flash-LiDAR-based pose estimation and shape reconstruction approach, by solving a maximum likelihood estimation problem via particle-swarm optimization, followed by a least-squares filter providing measurements for the spacecraft position and orientation in the asteroid frame coordinates. Other recent works in the field have established proof-of-concepts for online implementation of batch optimization and graph-based approaches for precise near-asteroid navigation, like real-time SLAM. Notably, Nakath et al. [26] present an active SLAM framework which also employs Flash-LiDAR as the base measurement of the SLAM formulation, with sensor fusion of data from an inertial measurement unit and star tracker, tested with simulated data. However, the limited range of Flash-LiDAR instruments bounds the spacecraft's orbit to unrealistically small radii, reducing the use scenarios to either navigation near very small asteroids or to the touchdown phase for larger target asteroids. For example, the OSIRIS-REx GoldenEye Flash-LiDAR, which is mentioned by both Nakath and Bercovici, only has a reliable maximum range of 200 m [29]. I contrast, an approach which uses long range optical imagery, like the one we propose in this paper, enables detailed characterization of the asteroid early in the approach phase of the mission, at which point knowledge about the target asteroid is still poor.

Multiple works have applied full visual SLAM solutions for spacecraft relative navigation, notably Tweddle's factor-graph based formulation of stereo SLAM implemented on the SPHERES platforms [37]. However, fewer works have directly applied visual SLAM to the asteroid navigation problem. We denote Cocaud et al. [9], who initially leveraged SURF [2] visual cues and range measurements, and Cocaud et al. who focused on image feature only formulation, and solved using a Rao-Blackwellized particle filter [10]. However, the latter works only tested the algorithm on simulated imagery of asteroid Itokawa. Similarly, Baldini et al. [1] implemented OpenSFM on simulated images of comet 67P. Alternatively, Rathinam et al. [32] adapted Tweddle's original formulation, but failed to tackle the real image feature point detection, feature descriptor computation or matching problems. Takeishi et al. [36] performed a particle filter minimization of the observation error and used both simulated landmarks and SIFT features extracted and tracked across a sequence of real images of a simple asteroid mock-up, with albeit unrealistic motion.

In this paper, we implement a factor graph-based incremental smoothing solution for monocular visual SLAM around an asteroid. Using the GTSAM library [11] and the iSAM2 solver [18], we perform multi-sensor fusion incorporating measurements from a star acquisition device and Earth-relative radiometry to obtain an initial pose prior and then we leverage image-based measurements for subsequent trajectory estimation. In order to capture the real-world effects of the challenging data association problem, the algorithm extracts, describes and matches local image features, thus establishing non-idealized 2D-2D and 2D-3D correspondence mappings. Additionally, and in contrast to current practices that leverage extended and expensive Deep Space Network (DSN) ranging for a navigation solution, we incorporate only sporadic DSN-based measurements for scale ambiguity resolution. Finally, we fully test our algorithm on an image sequence from a real past asteroid probing mission. We compare the estimated solution against archived navigational data from NASA and demonstrate excellent performance. Such a real-word comparison is, to our knowledge, a first among works concerning near-asteroid SLAM algorithms and constitutes a significant contribution of this paper.

## 3. TECHNICAL APPROACH

In this section, we detail the application of a graph-based SLAM smoothing formulation to the problem of asteroid navigation by clearly identifying the connections with typical SLAM methodologies, while underlining the specificities of the asteroid navigation problem.

### 3.1. Definitions

#### 3.1.1 Asteroid Navigation-related Definitions

Let $\mathrm{A}, \mathrm{S}, \mathrm{O} \in \mathbb{E}^3$ and assume that the point A corresponds to the center of mass of the asteroid body, S corresponds to the center of mass of the spacecraft, and O corresponds to an inertial point in space, e.g., the barycenter of the solar system. Initially, we distinguish three frames of interest: the inertial frame $\mathcal{I} \triangleq (\mathrm{O}, \{\bar{\mathbf{n}}_i\}_{i=1}^3)$, the asteroid principal axis frame $\mathcal{A} \triangleq (\mathrm{A}, \{\bar{\mathbf{a}}_i\}_{i=1}^3)$, and the spacecraft body-fixed frame $\mathcal{S} \triangleq (\mathrm{S}, \{\bar{\mathbf{x}}_i\}_{i=1}^3)$. Notice that the definition of the frame $\mathcal{A}$ is based on the a-priori unknown mass moments of the asteroid, i.e., its center of mass and the principal axes of its inertia tensor. It is necessary to initially consider an arbitrarily chosen asteroid-fixed frame $\mathcal{G} \triangleq (\mathrm{G}, \{\bar{\mathbf{g}}_i\}_{i=1}^3)$, and to then estimate the transformation $T_{\mathcal{GA}} \in \mathrm{SE}(3)$ from frame $\mathcal{A}$ to frame $\mathcal{G}$. Note that since frame $\mathcal{G}$ is fixed with respect to asteroid frame $\mathcal{A}$, the transformation $T_{\mathcal{GA}}$ is fixed. An estimate of $T_{\mathcal{GA}}$ is obtained as a result of detailed analysis of the asteroid's shape [13, 15], typically performed at a later stage, and is therefore not a subject of study in this paper. Any reference hereafter to the *relative pose of the*

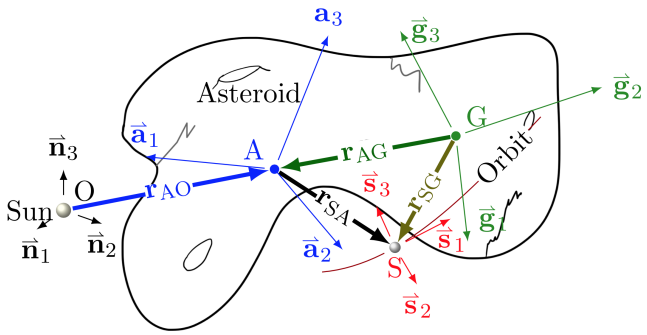


Figure 1: Relative navigation problem frame definitions and vector quantities.

*spacecraft*, for a given spacecraft frame $\mathcal{S} = (\mathrm{S}, \{\vec{\mathbf{s}}_i\}_{i=1}^3)$, designates the transformation

$$T_{\mathcal{GS}} = \begin{bmatrix} R_{\mathcal{GS}} & \mathbf{r}_{\mathrm{SG}}^{\mathcal{G}} \\ 0_{1\times 3} & 1 \end{bmatrix}, \tag{1}$$

which encodes the relative rotation $R_{\mathcal{GS}} = \begin{bmatrix} \vec{\mathbf{s}}_1^{\mathcal{G}} & \vec{\mathbf{s}}_2^{\mathcal{G}} & \vec{\mathbf{s}}_3^{\mathcal{G}} \end{bmatrix}^\top \in \mathrm{SO}(3)$ of the spacecraft with respect to the $\mathcal{G}$ frame and the coordinates $\mathbf{r}_{\mathrm{SG}}^{\mathcal{G}} = \begin{bmatrix} \mathbf{r}_{\mathrm{SG}} \cdot \vec{\mathbf{g}}_1 & \mathbf{r}_{\mathrm{SG}} \cdot \vec{\mathbf{g}}_2 & \mathbf{r}_{\mathrm{SG}} \cdot \vec{\mathbf{g}}_3 \end{bmatrix} \in \mathbb{R}^3$ of the spacecraft position vector relative to the point G as expressed in the $\mathcal{G}$ frame.

Let $t_0$ be an initial time and let $t \geq t_0$. Now let $(t_k)_{k=0}^n \subset [t_0, t]$ be the sequence of sensor acquisition times and let $T_{\mathcal{G}_k\mathcal{S}_k} \triangleq T_{\mathcal{GS}}(t_k)$ describe the pose of the spacecraft as expressed in the $\mathcal{G}$ frame at each time index $0 \leq k \leq n$. Then the sequence $(T_{\mathcal{G}_k\mathcal{S}_k})_{k=0}^n$ describes the discrete trajectory of the relative pose of the spacecraft in $\mathrm{SE}(3)$.

#### 3.1.2 SLAM-related Definitions

We define the camera sensor frame $\mathcal{C} = (\mathrm{S}, \{\vec{\mathbf{c}}_i\}_{i=1}^3)$, with fixed pose $T_{\mathcal{SC}}$ with respect to the spacecraft frame $\mathcal{S}$, and obtain the sequence $(T_{\mathcal{G}_k\mathcal{C}_k})_{k=1}^n$ of all camera poses, also known as *frames*, where $T_{\mathcal{G}_k\mathcal{C}_k} = T_{\mathcal{G}_k\mathcal{S}_k} T_{\mathcal{S}_k\mathcal{C}_k} = T_{\mathcal{G}_k\mathcal{S}_k} T_{\mathcal{SC}}$, $(k = 0, \ldots, n)$.

Denote by $\Psi_k = \{\mathrm{L}_i \in \mathbb{E}^3,\ i = 1, \ldots, m_k\}$ as the set of all *landmarks* accumulated up until time index $k = 0, \ldots, n$, also called the *map* at time index $k$. To each landmark $\mathrm{L} \in \Psi_n$ corresponds a 3D position vector $\mathbf{r}_{\mathrm{LG}} \in \mathbb{R}^3$, whose expression in the $\mathcal{G}$ frame, denoted $\mathbf{r}_{\mathrm{LG}}^{\mathcal{G}}$, is fixed since the asteroid is presumed to be a rigid body. Let $r_k$ be the total number of detected features in the image captured at time $t_k$, $k = 0, \ldots, n$. We collect all detected feature points in the set $\Upsilon_k = \{\mathrm{P}_i \in \mathbb{P}^2,\ i = 1, \ldots, r_k\}$, and to each $\mathrm{P}_i \in \Upsilon_k$ are associated the 2D image coordinates $\mathbf{y}_{ki} \in \mathbb{R}^2$.

### 3.2. Overview of Methodology

Our processing pipeline consists of a front-end system and a back-end system, along with a initialization step and a loop closure detection step.

#### 3.2.1 Initialization

Initialization of the SLAM problem is a two-tier process, in which we generate an initial estimate for the map $\Psi_1$, as well as insert prior factors for at least two camera poses. We leverage sensor information at the approach phase of the asteroid mission to perform the initialization.

Initially, we assume that we have an uncertain measurement of the spacecraft's inertial pose, which, using the exponential map (at the identity) of the SO(3) group of rotations and the *hat* operator (see [7]), can be defined as

$$T_{\mathcal{IS}_0}^{\mathrm{m}} \triangleq T_{\mathcal{IS}_0} \begin{bmatrix} \exp\left([\boldsymbol{\varepsilon}_1]^\wedge\right) & \boldsymbol{\varepsilon}_2 \\ 0_{1\times 3} & 1 \end{bmatrix}, \tag{2}$$

$\begin{bmatrix} \boldsymbol{\varepsilon}_1^\top & \boldsymbol{\varepsilon}_2^\top \end{bmatrix}^\top \sim \mathcal{N}\left(\mathbf{0}_{6\times 1}, \Sigma_{\mathrm{T},0}^{\mathrm{m}}\right)$, with known covariance matrix $\Sigma_{\mathrm{T},0}^{\mathrm{m}}$. Notice that we can write the true spacecraft inertial pose as $T_{\mathcal{IS}_0} = T_{\mathcal{IG}_0} T_{\mathcal{G}_0\mathcal{S}_0}$, and, for simplification, we can choose the arbitrary frame $\mathcal{G}$ such that $T_{\mathcal{IG}_0} = \mathrm{I}_4$, meaning that frames $\mathcal{G}_0$ and $\mathcal{I}$ coincide. It then follows that

$$T_{\mathcal{G}_0\mathcal{S}_0}^{-1} T_{\mathcal{IS}_0}^{\mathrm{m}} = \begin{bmatrix} \exp\left([\boldsymbol{\varepsilon}_1]^\wedge\right) & \boldsymbol{\varepsilon}_2 \\ 0_{1\times 3} & 1 \end{bmatrix}, \tag{3}$$

$\begin{bmatrix} \boldsymbol{\varepsilon}_1^\top & \boldsymbol{\varepsilon}_2^\top \end{bmatrix}^\top \sim \mathcal{N}\left(\mathbf{0}_{6\times 1}, \Sigma_{\mathrm{T},0}^{\mathrm{m}}\right)$. Now, for any $0 \leq k \leq n$, the measurement $T_{\mathcal{IS}_k}^{\mathrm{m}}$ can be obtained by combining an orientation measurement $R_{\mathcal{IS}_k}^{\mathrm{m}} \triangleq R_{\mathcal{IS}_k} \exp\left([\boldsymbol{\nu}]^\wedge\right)$, $\boldsymbol{\nu} \sim \mathcal{N}\left(\mathbf{0}_{3\times 1}, \Sigma_{\mathrm{R}}^{\mathrm{m}}\right)$ and an inertial position measurement $\mathbf{r}_{\mathrm{S}_k\mathrm{O}}^{\mathcal{I},\mathrm{m}} \triangleq \mathbf{r}_{\mathrm{S}_k\mathrm{O}}^{\mathcal{I}} + \boldsymbol{\nu}, \boldsymbol{\nu} \sim \mathcal{N}\left(\mathbf{0}_{3\times 1}, \Sigma_{\mathbf{r}}^{\mathrm{m}}\right)$, to form the pose

$$T_{\mathcal{IS}_0}^{\mathrm{m}} = \begin{bmatrix} R_{\mathcal{IS}_0}^{\mathrm{m}} & \mathbf{r}_{\mathrm{S}_0\mathrm{O}}^{\mathcal{I},\mathrm{m}} \\ 0_{1\times 3} & 1 \end{bmatrix}. \tag{4}$$

Then, for time index $k = 0$, using the logarithm map of the SE(3) group of homogeneous transformations at the identity and the *vee* operator, as detailed in [7], we can write

$$\left[\log\left(T_{\mathcal{G}_0\mathcal{S}_0}^{-1} \begin{bmatrix} R_{\mathcal{IS}_0}^{\mathrm{m}} & \mathbf{r}_{\mathrm{S}_0\mathrm{O}}^{\mathcal{I},\mathrm{m}} \\ 0_{1\times 3} & 1 \end{bmatrix}\right)\right]^\vee = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \sim \left(\mathbf{0}_{6\times 1}, \Sigma_{\mathrm{T},0}^{\mathrm{m}}\right), \tag{5}$$

where, by first-order linear approximation, we have $\Sigma_{\mathrm{T},0}^{\mathrm{m}} = J_{\mathrm{R}} \Sigma_{\mathrm{R}}^{\mathrm{m}} J_{\mathrm{R}}^\top + J_{\mathbf{r}}(R_{\mathcal{IS}_0}^{\mathrm{m}}) \Sigma_{\mathbf{r}}^{\mathrm{m}} J_{\mathbf{r}}^\top (R_{\mathcal{IS}_0}^{\mathrm{m}})$, with Jacobians

$$J_{\mathrm{R}} = \begin{bmatrix} \mathrm{I}_3 \\ 0_{3\times 3} \end{bmatrix}, \quad J_{\mathbf{r}}(R) = \begin{bmatrix} 0_{3\times 3} \\ R^\top \end{bmatrix}, \quad R \in \mathrm{SO}(3).$$

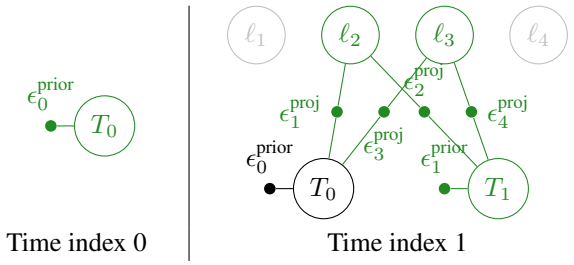


Figure 2: Initialization steps of the SLAM problem.

A prior factor $\epsilon_0^{\text{prior}}(T_{\mathcal{G}_0\mathcal{S}_0})$ is emplaced in the graph. This factor encodes the residual between the pose $T_{\mathcal{G}_0\mathcal{S}_0} \in \text{SE}(3)$ and the measurement $T^{\text{m}}_{\mathcal{I}\mathcal{S}_0}$, with covariance $\Sigma^{\text{m}}_{\text{T},0}$,

In practice, an on-board star acquisition system is leveraged to obtain an orientation measurement $R^{\text{m}}_{\mathcal{I}\mathcal{S}_k}$, which is usually known with very good accuracy and little uncertainty. Position measurements are based on Earth-relative radiometric ranging and bearing measurements, a method of localization widely practiced in deep space mission spacecraft tracking using communication station networks, such as NASA's Deep Space Network (DSN). Ground-based navigation estimates using Earth-relative range and bearing measurements, such as uplink-downlink pulse ranging and delta Differential One-Way Ranging ($\Delta$-DOR), a type of Very Long Baseline Interferometry (VLBI) [22, 23], are very accurate, i.e., in the order of several hundreds of meters for large asteroids and tens of meters for small asteroids [6]. It is important to note that our algorithm requires DSN-type measurements only during the initialization phase, to anchor the initial pose as described above and rectify the scale ambiguity.

Initialization of the map is delayed until time index $k = 1$, at which point at least two images of the target with sufficient parallax are captured. Local image features $\Upsilon_0$ and $\Upsilon_1$ are extracted and undergo data association, with outlier rejection, producing a set of 2D-2D correspondences. A strict outlier rejection criterion is used to obtain a subset of correspondences for landmark triangulation. It is now possible to apply a typical 8-point algorithm [17] using the remaining 2D-2D correspondences to find a guess for pose at time index $k = 1$, followed by a triangulation of landmarks $\Psi_1$ using poses at time indices $k = 0$ and $k = 1$. It is crucial to note, however, that the ambiguity in scale will not be resolved in this fashion. Instead, to triangulate the new landmark positions to scale, we need knowledge about either a pose value and a landmark position, or the true scale values for at least two of the initializing poses. In practice, this can be accomplished by using a combination of pose measurements and rangefinding measurements (like a laser altimeter), for asteroids at close range, or a combination of an initial pose measurement and motion model, with prior knowledge of the asteroid motion. For simplicity, we assume that we have access to a second relative pose measurement $T^{\text{m}}_{\mathcal{G}_1\mathcal{C}_1}$ for this purpose. In addition, a value for the first pose is available from the estimate at time index $k = 0$. With these two pose values, 2D-to-3D triangulation is performed on the set of inlier matched feature points to generate guess values for the estimated landmark positions $\left\{\mathbf{r}^{\mathcal{G}}_{\text{LG}}\right\}_{\text{L}\in\Psi_1}$. At this step, all appropriate factors are inserted based on the 2D-to-3D correspondences, resulting in the factor graphs illustrated in Fig. 2, where the shorthands $\ell_i \triangleq \mathbf{r}^{\mathcal{G}}_{\text{L}_i\text{G}}$, $i \in \Psi_k$ and $T_k \triangleq T_{\mathcal{G}_k\mathcal{S}_k}$, $k = 0, \ldots, n$ are used for brevity.

### 3.2.2 Front-End

The front-end system includes feature detection and matching, and encodes the structure of the SLAM problem in a factor graph using data structures based on the definitions in Section 3.1.2.

We use ORB features [33], as they perform well in practice and they are fast to compute. These are also good placeholders for more robust automatic features to be implemented in the future. For matching, we use brute force nearest neighbor search based on the Hamming distance for binary feature descriptors. We reject 2D-2D matches with a distance ratio greater than 0.85 [21]. We also enforce a fundamental matrix constraint as a geometric check for matched features. We track landmarks by comparing the current frame $n$ to previous frame 2D-to-2D correspondences against past frame 2D-to-3D correspondences. The feature-landmark matched pairs then undergo a reprojection error test, after which the surviving pairs establish new projection factors to be inserted into the graph, relating to the most recent frame $n$. If there are enough tracked features from time index $k = n-1$, then visual tracking is successful and the well-known P$n$P algorithm [20] is used to guess the camera pose $\tilde{T}_{\mathcal{G}_n\mathcal{C}_n}$ value from matched correspondences. Leveraging the latest guess pose $\tilde{T}_{\mathcal{G}_n\mathcal{C}_n}$, a guess value for 3D position for each of the newly detected features is generated using triangulation. Additionally, we delay insertion of new landmarks into the graph based on whether the number of times the associated landmark has been seen is above a predetermined threshold of 3. The front-end system produces incrementally growing factor graphs, as is illustrated in Fig. 3.

### 3.2.3 Back-End

The underlying structure of the navigation problem is thus captured by encoding visual SLAM measurement constraints, star acquisition orientation measurement constraints and Earth-relative inertial position measurement constraints as factors in a single factor graph. The choice of this graph formulation is predicated on the fact that a factor

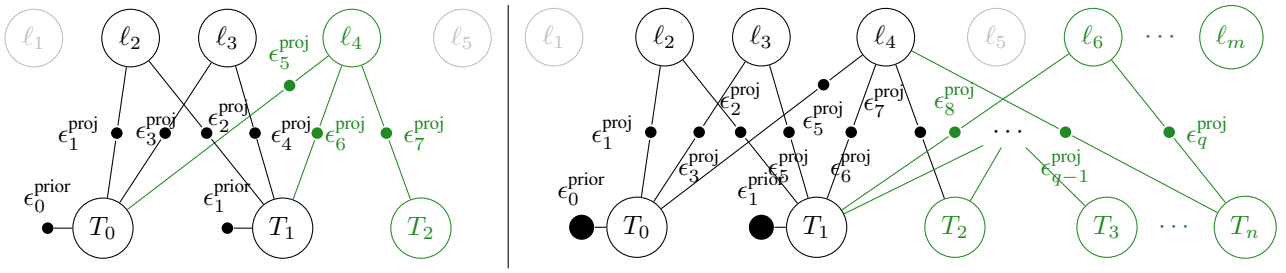


Figure 3: Front-end pipeline factor-graph following the initialization of the SLAM.

graph, as an undirected graph, readily explains the relationships between unknowns, since its incidence matrix directly relates to the $R$ matrix in the $QR$ factorization of the square root information matrix [18]. In addition, the graph lends itself naturally to incremental growth. For every new frame inserted, the graph is incrementally augmented with the new variables and factors by the front-end.

The iSAM2 algorithm integrates these new measurement constraints as new factors and performs inexpensive Givens rotations [18] on the existing $R$ matrix to perform further variable elimination without recalculating the full variable elimination. The back-end system evaluates the losses at each factor, computes the associated Jacobians at the guess values $\left\{\tilde{T}_{\mathcal{G}_k\mathcal{C}_k}\right\}_{k=0}^{n}$ and $\left\{\tilde{\mathbf{r}}^{\mathcal{G}}_{\mathrm{LG}_\mathrm{L}}\right\}_{\mathrm{L}\in\Psi_n}$, and performs the minimization for inference [11], yielding a new estimate solution.

#### 3.2.4 Loop Closure

To perform loop closure, we leverage the bag-of-words representation developed in [14]. We convert each image $i = 1, \ldots, n$ to a bag-of-words vector $\mathbf{v}_i$ and compute the similarity metric

$$s(\mathbf{v}_i, \mathbf{v}_j) \triangleq 1 - \frac{1}{2}\left|\frac{\mathbf{v}_i}{|\mathbf{v}_i|} - \frac{\mathbf{v}_j}{|\mathbf{v}_j|}\right|. \tag{6}$$

We compare all prior images that are at least 10 frames away from the current frame. When the similarity score for two images is greater than a threshold $\eta$, we perform an additional geometric check and then add a factor between the poses corresponding to the detected loop.

## 4. EXPERIMENTAL SETUP & RESULTS

In this section, we discuss the design of the experimental and validation process of the proposed algorithm. Additionally, we provide the results of the estimated trajectory and map of landmarks in tests using real world data, including quantitative results using appropriate metrics.

### 4.1. Experimental Setup

We use real imagery [28] of Asteroid (4) Vesta acquired during the Rotation Characterization 3 (RC3) observation phase of the Dawn mission [34] to validate the algorithm. In the chosen sequence, the $1024 \times 1024$ images were captured while the spacecraft performed one apparent revolution around Vesta in the asteroid body-fixed frame, with a mean orbital radius of 5480 km. The images thus provide a spatial resolution of 0.5 km/pixel of the surface (see Fig. 4 for sample images). This sequence therefore enables possible loop closures to be tested as well. Two tests were conducted for the dataset, each with a different prior uncertainty on poses $T_{\mathcal{G}_0\mathcal{S}_0}$ and $T_{\mathcal{G}_1\mathcal{S}_1}$, i.e., the covariance $\Sigma_{\mathrm{T},0} = J_\mathrm{R}\Sigma^{\mathrm{m}}_R J_\mathrm{R}^\top + J_\mathrm{r}\left(R^{\mathrm{m}}_{\mathcal{IS}_0}\right)\Sigma^{\mathrm{m}}_{\mathbf{r}} J_\mathrm{r}^\top \left(R^{\mathrm{m}}_{\mathcal{IS}_0}\right)$ of the initial pose measurements. The values are chosen as $\Sigma^m_\mathrm{R} = \sigma^2_\mathrm{R}\mathrm{I}_3$ and $\Sigma^m_\mathrm{r} = \sigma^2_\mathrm{r}\mathrm{I}_3$, with $\sigma_\mathrm{R} = 1 \times 10^{-3}\,\mathrm{rad}$ $\sigma_\mathrm{r} = 0.2\,\mathrm{km}$ for the poor prior case, and $\sigma_\mathrm{R} = 1\times10^{-5}\,\mathrm{rad}$ and $\sigma_\mathrm{r} = 0.05\,\mathrm{km}$ for the good prior case. Note that, for simplicity in these trials, the pose discrepancy $T_{\mathcal{AG}}$ is assumed to be known. The intrinsic parameters of the Dawn

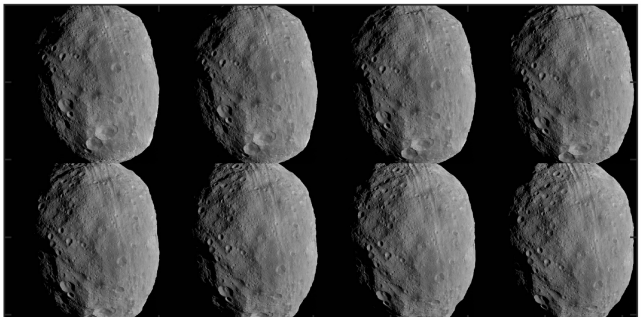

Figure 4: 8 time-consecutive images from the RC3 observation phase of the DAWN mission.

Framing Camera (FC) were taken to be those computed during calibration [34]. In the conducted tests, the algorithm was set to extract 1400 features where 8 scale pyramid levels are explored.

To validate the estimated spacecraft relative trajectory, we use the archival SPICE kernel datasets maintained by NASA's Navigation and Ancillary Information Facility (NAIF). For the missions archived, SPICE kernels can be queried to provide **S**pacecraft ephemeris and asteroid (**P**lanetary) ephemeris as a function of time, as well as **I**nstrument descriptive data, **C**amera orientation matrix data and **E**vents information, such as mission phases.

To validate the reconstructed map, we use an archival shape model of Vesta [30] courtesy of the PDS Small Bodies Node. This shape model was derived using stereo photogrammetry (SPG) from a subset of DAWN mission Framing Camera 2 (FC2) images captured during the High-Altitude Mapping Orbit (HAMO) mission segment. The model comprises approximately 100k vertices and 197k triangular faces and is shown in Fig. 4.

## 4.2. Results

### 4.2.1 Evaluation of Trajectory Estimation

For qualitative evaluation of the relative navigation, we visualize in Fig. 5 the 3D trajectory derived from the relative poses $\{\tilde{T}_{\mathcal{A}_k\mathcal{S}_k}\}_{k=0}^{\mathrm{N}}$ overlayed on the trajectory extracted from the DAWN mission SPICE kernel poses, hereon assumed to be the true poses $\{T_{\mathcal{A}_k\mathcal{S}_k}\}_{k=0}^{\mathrm{N}}$.

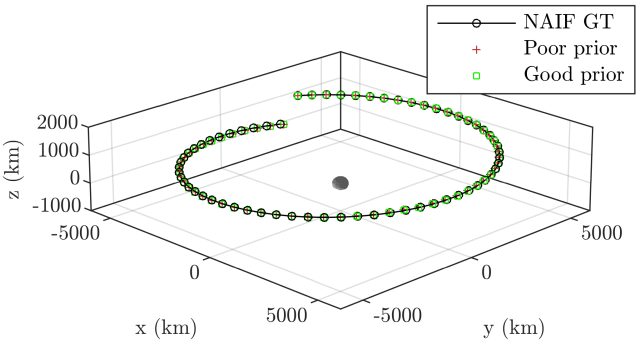


Figure 5: 3D trajectory estimate vs NAIF SPICE kernel ground truth for DAWN RC3 segment.

Although knowledge about the spacecraft's pose in the asteroid body-fixed frame given by $T_{\mathcal{A}_k\mathcal{S}_k}$ is ultimately sought after, we observe that the inverse pose $T_{\mathcal{S}_k\mathcal{A}_k}$ describing the orientation and the position of the asteroid in the rotating spacecraft frame is more revealing when reporting the navigation error, as discussed in Section 5. Decomposing the error along these directions is doubly useful in the DAWN RC3 test case, since the direction of the spacecraft's body-fixed frame unit vector $\vec{\mathbf{s}}_2$ is close to the along-track direction in the relative orbit, while the spacecraft's camera boresight vector $\vec{\mathbf{s}}_3$ points in the radial direction towards the asteroid. Thus, for quantitative evaluation of the estimated trajectory, for each time index $k = 1, \ldots, n$, we compute the error vector

$$\begin{bmatrix}\delta\boldsymbol{\kappa}_k^\top & \delta\mathbf{r}_k^\top\end{bmatrix}^\top \triangleq \left[\log\left(T_{\mathcal{S}_k\mathcal{A}_k}^{-1}\tilde{T}_{\mathcal{S}_k\mathcal{A}_k}\right)\right]^\vee \in \mathrm{se}(3) \qquad (7)$$

between the ground-truth NAIF SPICE pose $T_{\mathcal{S}_k\mathcal{A}_k}$ and the estimated pose $\tilde{T}_{\mathcal{S}_k\mathcal{A}_k}$, where $\log : \mathrm{SE}(3) \to \mathrm{se}(3)$ is the SE(3) logarithm map at the identity, as detailed in [7]. The results of this evaluation are presented in Fig. 6 and discussed in Section 5.

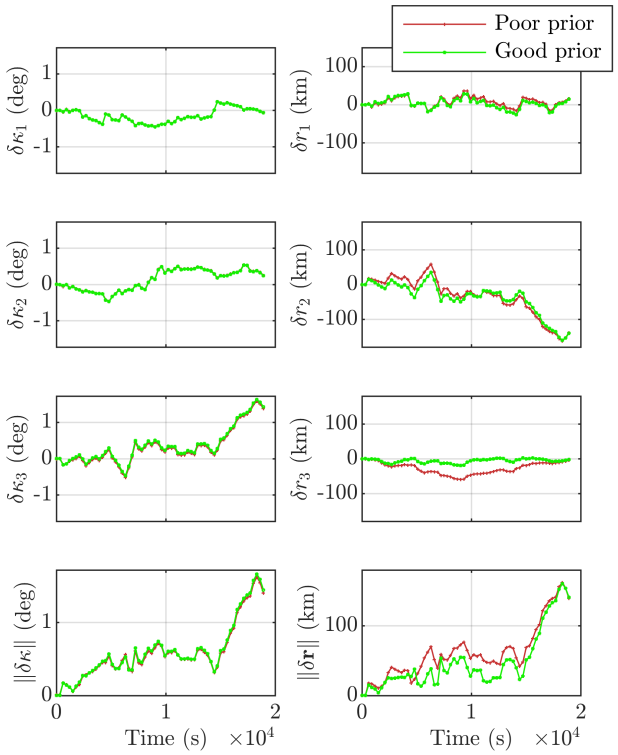


Figure 6: Estimated pose errors with respect to NAIF SPICE kernel ground truth for DAWN RC3 segment.

### 4.2.2 Evaluation of Map Reconstruction

Given the map $\Psi_{\mathrm{N}}$ at the final time $k = \mathrm{N}$ and the set of ground truth 3D shape model vertices $\mathcal{V}$, we evaluate the quality of the estimated landmark by computing the distances $\{d(\mathrm{L}, \mathcal{V})\}_{\mathrm{L}\in\Psi_{\mathrm{N}}}$, where

$$d(\mathrm{L}, \mathcal{V}) \triangleq \min_{\mathrm{V}\in\mathcal{V}} \|\mathbf{r}_{\mathrm{LA}}^{\mathcal{A}} - \mathbf{r}_{\mathrm{VA}}^{\mathcal{A}}\|_2, \qquad (8)$$

which in our case minimizes the 2-norm. The results of this evaluation are presented in Fig. 7, where the estimated landmarks are colored as a function of their distance to the closest point in the ground-truth set of vertices. Note that this distance metric is one-sided. Thus, choosing to instead search over the set $\Psi_{\mathrm{N}}$ would yield different distance values $\{d(\mathrm{V}, \Psi_{\mathrm{N}})\}_{\mathrm{V}\in\mathcal{V}}$. Nevertheless, given the much higher vertex density of the ground truth shape model as compared to the estimated landmarks, we deem the described point-to-set distance $d(\mathrm{L}, \mathcal{V})$, $\mathrm{L} \in \Psi_{\mathrm{N}}$ to be an appropriate measure of the deviation of our solution landmarks from the 3D shape.

### 4.2.3 Comparison to ORB-SLAM2

We further validate our solution by comparing the estimation error performance of our algorithm against that of

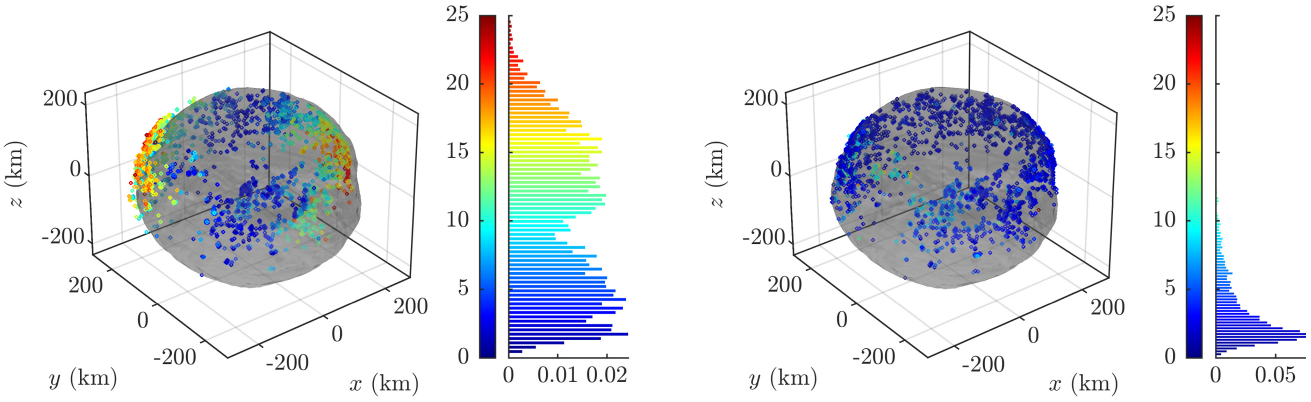


Figure 7: Heat map and distribution of distances of estimated landmarks to true model vertices.
*Poor prior uncertainty (left) and good prior uncertainty (right).*

ORB-SLAM2 [25], which is a modern SLAM algorithm that also exploits ORB features [33] for feature extraction and description and employs g2o [16] for bundle adjustment. ORB-SLAM2 is known to demonstrate robustness to outliers being tracked due to spurious data association by performing a harsh culling of recent map points. A similar approach was employed in our algorithm, where a landmark point only survives in the map if it has been observed in at least 3 different frames.

We compare the performance of both algorithms on the same DAWN RC3 image sequence. As such, the ground truth pose $T_{\mathcal{G}_0\mathcal{S}_0}$ is used to align the first pose in each of the algorithm's resulting trajectories.

It is noteworthy that our algorithm proposes the ability to incorporate a prior on the second initialization pose $T_{\mathcal{G}_1\mathcal{S}_1}$ as well, allowing for a more accurate subsequent map initialization. In comparison, although ORB-SLAM2's g2o engine contains the necessary components for priors to be enforced on specific pose variables in the SLAM problem graph, ORB-SLAM2 does not natively allow for such priors to be included and maintained throughout the optimization. Indeed, ORB-SLAM2 carries out a node culling scheme which removes frames in the graph solely based on whether the said frames fail to satisfy several geometric criteria with respect to their neighboring frames, such as having enough parallax or having a significant change in the detected feature population as compared to the current frame. In the same manner, poses having an associated prior can also be removed from the graph. Thus, in the interest of fairness, no ground truth prior on the second pose of the trajectory is enforced. In terms of the front-end, the same set of parameters related to ORB feature detection and description are used, including the number of features extracted per image, the scale pyramid levels, and thresholds for feature descriptor matching.

In Fig. 8 it is observed that our algorithm outperforms ORB-SLAM2 in the cross-track and radial directions, as well as in terms of attitude error throughout the sequence.

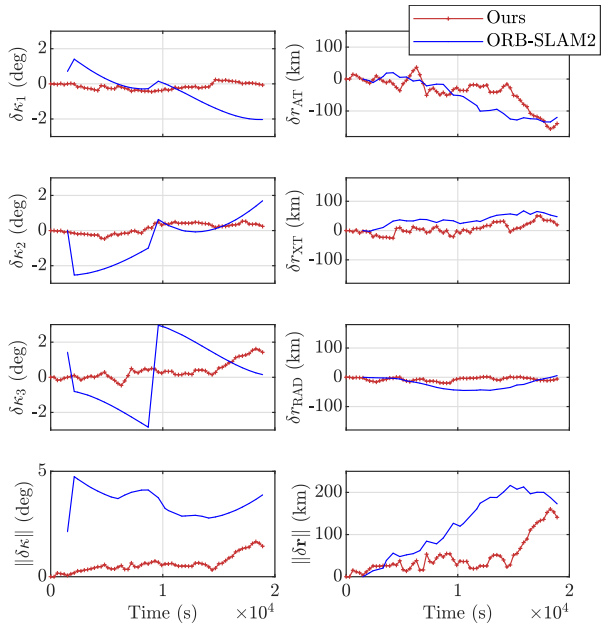


Figure 8: Estimate pose errors with respect to ground truth, comparison against ORB-SLAM2

## 5. DISCUSSION & CONCLUSION

Considering that the orbit radius is roughly 5470 km throughout the sequence, the error in the radial direction ($\delta r_3$), as shown in Fig. 6, yields, for the good prior case, an error of 0.1% (5.8 km) on average, with a worst case of 0.4% (19.2 km), while for the poor prior case, it yields an accuracy of 0.3% (16.3 km) on average and a worst case of 1.1% (59.2 km). Note that the estimated and ground-truth trajectories are close, as illustrated in Fig. 5.

The discrepancy in radial position error is the most significant difference between the poor prior and good prior cases. We observe that the main bulk of the error is, however, in the relative orbit along-track direction, and this er-

ror increases dramatically with the diminishing number of tracked landmarks after $1.4 \times 10^4$ s, as it is evident in Fig. 9.

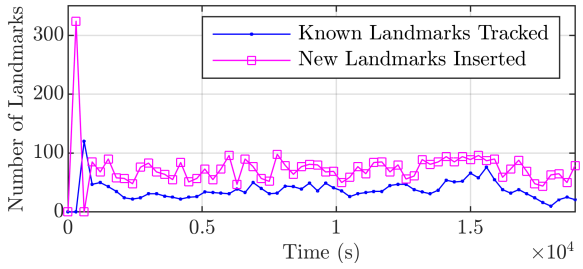


Figure 9: Tally of Known Tracked vs Newly Inserted Landmarks.

Figure 9 also validates the methodology of harsh culling [25] that is commonly used for the less reliable ORB features. Indeed, although there are many landmarks inserted at every time step, the number of tracked landmarks is significantly smaller and does not grow over time in an appreciable manner.

Additionally, the quality of the estimated landmarks, as computed in Section 4.2.2 using a Euclidean distance to true model vertices, is significantly affected by the magnitude of the initial pose uncertainty. Indeed, as can be seen in Fig. 7, in the case with good prior uncertainty, the distribution of the error is tightly concentrated around 3.26 km, which is an error of $\sim 1.2\%$ compared to the target asteroid mean radius of 262 km. In contrast, in the case of poor prior uncertainty, we see a much larger spread of the errors and observe a higher mismatch between estimated landmarks and the target 3D model. This effect is explainable: a low uncertainty on the initialization set of poses makes for a more tightly constrained initialized map. In turn, landmarks in the better constrained initialized map are much less affected by subsequent measurements, which are added incrementally. Indeed, a subsequent change in initialized landmark position would incur a high cost due to the lower imposed uncertainty of the related initial poses.

Finally, it is noteworthy that the apparent revolution trajectory of the DAWN RC3 image sequence lends itself to successful loop-closures, as we observe an increased similarity score $\sim 10^{-2}$ between the last frames of the sequence and the first frame of the sequence. However, we observed little difference in the overall solution once the loop closure constraint was included in the graph at the very last time step, with an ensuing last step of optimization. Further study on this matter is planned in the future using additional real mission images, as well as images produced on an in-lab hardware test-bed.

In summary, in this paper, we demonstrate that a smoothing-based SLAM solution, predicated on a factor graph formulation allowing for fusion of sensor-derived measurements and imagery-derived measurements and incremental growth is well-adapted for autonomous near-asteroid relative navigation problem. We do so by validating the solution with real world data from a previous flown asteroid orbiter mission, producing very accurate results. We establish that satisfactory performance can be achieved in real-world deep-space asteroid orbiter missions, considering only a baseline SLAM architecture. Areas of further work include improved keypoint and landmark management, as well as the integration of a motion model, given the strong priors of the asteroid's rotational motion and the spacecraft's orbital motion.

## 6. ACKNOWLEDGEMENTS

This work was supported by an Early Stage Innovations grant 80NSSC18K0251 from NASA's Space Technology Research Grants Program. The authors would like to thank Kenneth Getzandanner (NASA GSFC) and Frank Dellaert for several fruitful discussions.

## References

[1] F. Baldini, A. Harvard, S. Chung, I. Nesnas, and S. Bhaskaran. Autonomous small body mapping and spacecraft navigation via real-time SPC-SLAM. In *69th IAF International Astronautical Congress (IAC)*, Bremen, Germany, October 1-5 2018. 2

[2] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded up robust features. In *European Conference on Computer Vision*, pages 404–417. Springer, Graz, Austria, May 7-13 2006. 2

[3] B. Bercovici and J. W. McMahon. Robust autonomous small-body shape reconstruction and relative navigation using range images. *Journal of Guidance, Control, and Dynamics*, 42(7):1473–1488, 2019. 2

[4] K. Berry, K. Getzandanner, M. Moreau, P. Antreasian, A. Polit, H. Enos, D. Lauretta, and M. Nolan. Revisiting OSIRIS-REx Touch-And-Go (TAG) performance given the realities of asteroid Bennu. In *43rd AAS Guidance, Navigation and Control Conference*, Breckenridge, CO, USA, January 30 - February 5 2020. Preprint at https://ntrs.nasa.gov/citations/20200000774. 1

[5] S. Bhaskaran and B. Kennedy. Closed loop terminal guidance navigation for a kinetic impactor spacecraft. *Acta Astronautica*, 103:322–332, 2014. 1

[6] S. Bhaskaran, S. Nandi, S. Broschart, M. Wallace, L. Alberto Cangahuala, and C. Olson. Small body landings using autonomous onboard optical navigation. *The Journal of the Astronautical Sciences*, 58(3):409–427, 2011. 4

[7] G. Chirikjian. *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*, volume 2. Springer Science & Business Media, 2011. 3, 6

[8] J. Christian and G. Lightsey. Onboard image-processing algorithm for a spacecraft optical navigation sensor system. *Journal of Spacecraft and Rockets*, 49(2):337–352, 2012. 1

[9] C. Cocaud and T. Kubota. SURF-based SLAM scheme using octree occupancy grid for autonomous landing on asteroids. In *Proceedings of the 10th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS)*, volume 29, Sapporo, Japan, August 29 - September 1 2010. 2

[10] C. Cocaud and T. Kubota. Autonomous navigation near asteroids based on visual SLAM. In *Proceedings of the 23rd International Symposium on Space Flight Dynamics*, Pasadena, CA, USA, October 29 - November 2 2012. 2

[11] F. Dellaert and M. Kaess. Factor graphs for robot perception. *Foundations and Trends in Robotics*, 6(1-2):1–139, 2017. 2, 5

[12] M. Delpech, V. Bissonnette, and L. Rastel. Vision-based navigation for proximity operations around asteroid 99942 Apophis. In *Proceedings of 25th International Symposium on Space Flight Dynamics*, Munich, Germany, October 19-25 2015. 1

[13] T. Driver, M. Dor, K. Skinner, and P. Tsiotras. Space carving in space: A visual-SLAM approach to 3D shape reconstruction of a small celestial body. In *AAS/AIAA Astrodynamics Specialist Conference*, Lake Tahoe, CA, USA, August 9-13 2020. Preprint. 2

[14] D. Galvez-López and J. D. Tardos. Bags of binary words for fast place recognition in image sequences. *IEEE Transactions on Robotics*, 28(5):1188–1197, 2012. 5

[15] R. Gaskell, O. Barnouin-Jha, D. Scheeres, A. Konopliv, T. Mukai, S. Abe, J. Saito, M. Ishiguro, T. Kubota, and T. Hashimoto. Characterizing and navigating small bodies with imaging data. *Meteoritics & Planetary Science*, 43(6):1049–1061, 2008. 1, 2

[16] G. Grisetti, R. Kümmerle, H. Strasdat, and K. Konolige. g2o: A general framework for (hyper) graph optimization. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China*, pages 9–13, 2011. 7

[17] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004. 4

[18] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *The International Journal of Robotics Research*, 31(2):216–235, 2012. 2, 5

[19] A. Konopliv, S. Asmar, R. Park, B. Bills, F. Centinello, A. Chamberlin, A. Ermakov, R. Gaskell, N. Rambaux, and C. Raymond. The Vesta gravity field, spin pole and rotation period, landmark positions, and ephemeris from the DAWN tracking and optical data. *Icarus*, 240:103–117, 2014. 1

[20] V. Lepetit, F. Moreno-Noguer, and P. Fua. Epnp: An accurate O(n) solution to the pnp problem. *International Journal of Computer Vision*, 81(2):155, 2009. 4

[21] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 4

[22] J. Miller. Measurements and calibrations. In *Planetary Spacecraft Navigation*, pages 255–292. Springer International Publishing, Cham, 2019. 4

[23] J. Miller and K. Rourke. The application of differential VLBI to planetary approach orbit determination. *The Deep Space Network Progress Report*, 42:40, 1977. 4

[24] J. K. Miller et al. Determination of shape, gravity and rotation state of asteroid 433 Eros. *Icarus*, 155, issue 1:3–17, 2002. 1

[25] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5):1147–1163, Oct 2015. 7, 8

[26] D. Nakath, J. Clemens, and C. Rachuy. Active asteroid-SLAM. *Journal of Intelligent & Robotic Systems*, pages 1–31, 2019. 2

[27] D. Nakath, J. Clemens, and K. Schill. Multi-sensor fusion and active perception for autonomous deep space navigation. In *21st International Conference on Information Fusion (FUSION)*, pages 2596–2605, Cambridge, UK, July 10-13 2018. 2

[28] A. Nathues, H. Sierks, P. Gutierrez-Marques, S. Schroeder, T. Maue, I. Buettner, M. Richards, U. Chistensen, and U. Keller. DAWN FC2 calibrated Vesta images v1.0. NASA Planetary Data System, 2011. 5

[29] I. Poberezhskiy, A. Johnson, D. Chang, E. Ek, D. Natzic, G. Spiers, S. Penniman, and B. Short. Flash lidar performance testing: configuration and results. In Monte D. Turner and Gary W. Kamerman, editors, *Laser Radar Technology and Applications XVII*, volume 8379, pages 36 – 45. International Society for Optics and Photonics, 2012. 2

[30] F. Preusker, F. Scholten, K. Matz, T. Roatsch, R. Jaumann, C. Raymond, and C. Russell. DAWN FC2 derived vesta DTM SPG v1. 0. NASA Planetary Data System, 2016. 6

[31] L. Prockter, S. Murchie, A. Cheng, S. Krimigis, R. Farquhar, A. Santo, and J. Trombka. The NEAR shoemaker mission to asteroid 433 Eros. *Acta Astronautica*, 51(1):491 – 500, 2002. 1

[32] A. Rathinam and A. Dempster. Monocular vision based simultaneous localization and mapping for close proximity navigation near an asteroid. In *68th IAF International Astronautical Congress*, Adelaide, Australia, September 25-29 2017. 2

[33] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: an efficient alternative to SIFT and SURF. In *IEEE International Conference on Computer Vision*, pages 2564–2571, Barcelona, Spain, November 6-13 2011. 4, 7

[34] C. Russell and C. Raymond. *The DAWN Mission to Minor Planets 4 Vesta and 1 Ceres*. Springer, 2012. 5

[35] J. A. Starek, B. Açıkmeşe, I. A. Nesnas, and M. Pavone. *Spacecraft Autonomy Challenges for Next-Generation Space Missions*, pages 1–48. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. 1

[36] N. Takeishi, T. Yairi, Y. Tsuda, F. Terui, N. Ogawa, and Y. Mimasu. Simultaneous estimation of shape and motion of an asteroid for automatic navigation. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2861–2866, Seattle, WA, USA, May 25-28 2015. 2

[37] B. E. Tweddle. *Computer Vision Based Navigation for Spacecraft Proximity Operations*. PhD thesis, Massachusetts Institute of Technology, 2013. 2

[38] B. Williams. Technical challenges and results for navigation of NEAR Shoemaker. *Johns Hopkins APL Tech. Dig.*, 23, 01 2002. 1

[39] M. Yoshikawa, J. Kawaguchi, A. Fujiwara, and A. Tsuchiyama. HAYABUSA sample return mission. *Asteroids IV*, 1:397–418, 2015. 1