Dynamic Visual Simultaneous Localization and Mapping for Asteroid Exploration

*Naoya Takeishi[1] and Takehisa Yairi[1]

[1]*The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, E-mail: {takeishi, yairi}@ailab.t.u-tokyo.ac.jp*

ABSTRACT

In asteroid explorations, it is indispensable to estimate the shape of the target asteroid, which can be solved in a manner similar to one of simultaneous localization and mapping (SLAM). This work proposes a SLAM framework dedicated to the asteroid exploration, which considers both rigid-body dynamics of the asteroid and motion of the spacecraft, estimating asteroid's shape, centroid, rotational axis, angular velocity and phase, as well as spacecraft's attitude and position relative to the asteroid. Experimental results with artificially generated data show that the proposed method can accurately estimate these quantities using images of a monocular camera and measurements of attitude and inertia sensors.

1 INTRODUCTION

In asteroid explorations, it is indispensable to estimate the shape of the target asteroid for the spacecraft to land on its surface, and for us to understand its scientific state. This estimation, termed *global mapping*, is conducted above the asteroid after rendezvous, and needs information on spacecraft's attitude and position as well as visual cues of the asteroid.

Since spacecraft's position relative to the asteroid cannot be measured directly, we have to estimate it together with the shape of the asteroid. Hence the problem we are going to solve is a kind of simultaneous localization and mapping (SLAM), where geometry of both environment and an observer is to be reconstructed from observation.

If optical images of the asteroid are available, as is often the case, just reconstructing its shape and camera's position is not so challenging, because of advancement on SLAM solutions and development of related software and hardware in this decade. For example, Mori and Hirata [1] reported successful results of applying an open-source library[1] to images of an asteroid for the shape estimation. However, several issues specific to asteroid exploration cannot be considered well by the current general solution and tools on SLAM.

[1] Mori and Hirata [1] used a library for the structure from motion, which is a problem almost equivalent to SLAM.

There are three major points at SLAM in asteroid exploration. First, the relative position between the asteroid and spacecraft is insufficient because we would like to know asteroid's motion (rotation) and spacecraft's motion separately. Second, information we can utilize is not only images; measurements of attitude sensors and inertial sensors should be incorporated into the solution for better estimation. Third, a photometric appearance of asteroid's surface changes as the sunlight shifts, and thus we must be careful about landmark tracking.

To tackle the problems stated above, this work proposes a monocular SLAM framework dedicated to the spacecraft above the rotating asteroid. It considers both rigid-body dynamics of the asteroid and motion of the spacecraft, estimating asteroid's shape, centroid, rotational axis, angular velocity and phase, as well as spacecraft's pose and position relative to the asteroid. The proposed framework mainly utilizes images of the asteroid taken by a monocular camera of the spacecraft, because stereo cameras will not work due to a long distance between the asteroid and spacecraft. Moreover, measurements of an attitude sensor and inertial sensors can be taken into account simultaneously.

Though we mainly focus on offline estimation, where the estimation is done after we obtain a batch of measurements, the proposed method can be run also in an online (real-time) manner, that is, the estimation can be updated given a new set of measurements incrementally. Note, however, that we think primarily of processing on the ground with enough resources, and improving computational speed and efficiency for on-orbit processing remains as an open problem for our method.

We present details of the proposed method in Section 2 and its experimental results in Section 3. Related studies important to understand well the proposed framework are introduced in Section 4, and this paper is ended with conclusion in Section 5.

2 PROPOSED METHOD

2.1 Notation and Definitions

In the following, a subscript i represents timestamps and $1 \leq i \leq N$, and a subscript k is a number of landmarks and $1 \leq k \leq L$. We represent the total number of landmarks by L, but use another notation

$\{L(i)\}$ to denote the numbers (IDs) of landmarks that can be reconstructed within the first i images, that is, the landmarks that have appeared more than once by the i-th timestamp. Also in this manner we utilize notation $\{L^{\text{new}}(i)\}$ to denote a set of the numbers of landmarks that were observed twice at first at the i-th timestamp. A subscript j $(1 \leq j \leq M_i)$ denotes the number of feature points observed in the i-th image. We use notation $M = \max_i M_i$ in the following.

Coordinate Frames Five types of coordinate frames are considered. A frame termed $\boldsymbol{B}$ is the *asteroid frame*, which is defined with principal axes and the center of asteroid's inertia. $\boldsymbol{G}$ denotes the *landmark frame*, where true and estimated location (in 3-D) of landmarks on asteroid's surface lies. Their location in 2-D images are treated in the *image frame* $\boldsymbol{D}$. The *camera frame* $\boldsymbol{C}$ is defined with the focal plane and the optical axis of the camera equipped to the spacecraft. Moreover, the inertial frame of reference is denoted by $\boldsymbol{S}$.

Unknown Quantities We would like to estimate: asteroid's pose[2] $\mathbf{x}_i$ and its temporal derivative $\dot{\mathbf{x}}_i$, ratio of inertia moments of the asteroid $\mathbf{k}$, the transformation from the landmark frame to the asteroid frame $\mathrm{T}_{\boldsymbol{G}\to\boldsymbol{B}}$, landmark's position $\mathbf{l}_k$, and spacecraft's pose $\mathbf{z}_i$. The pose $\mathbf{x}_i$ consists of the transition part $\mathbf{t}_i$ and the rotation part $\mathbf{r}_i$, and their time derivatives are denoted by $\dot{\mathbf{t}}_i$ and $\dot{\mathbf{r}}_i$.

Observed Quantities We are given: measurements of an attitude sensor $\mathbf{s}_i$, measurements of an IMU $\mathbf{u}_i$, and images of the asteroid I_i. For simplicity, we assume that $\mathbf{s}_i$ and $\mathbf{u}_i$ are already calibrated and transformed into the camera frame $\boldsymbol{C}$, and $\mathbf{u}_i$ denotes odometry of the spacecraft between the $(i-1)$-th and the i-th timestamps, while this assumption can be easily eliminated. Also we utilize feature points at $\mathbf{y}_{i,j}$ in the images and their descriptions $\mathbf{d}_{i,j}$ that are extracted using feature detectors and descriptors.

2.2 Models

The following is the models in which relation between the unknown and observed quantities are specified.

Attitude Observation Since we assume that the attitude measurements $\mathbf{s}_i$ are already calibrated, a very simple model suffices:

$$\boldsymbol{s}_i = R(\boldsymbol{z}_i) \cdot exp(\boldsymbol{e}_{sensor}),$$

where $\mathrm{R}(\cdot)$ denotes the rotation part of a pose, $\mathbf{e}_{\text{sensor}} \in \mathbb{R}^3$ is Gaussian noise, and $\exp(\cdot)$ is the exponential map from $\mathbb{R}^3$ into $\mathrm{SO}(3)$.

Inertia Observation Like the attitude observation, measurements of the IMU $\mathbf{u}_i$ are described as follows.

$$\mathbf{u}_i = \mathbf{z}_i \cdot (\mathbf{z}_{i-1})^{-1} \cdot \exp(\mathbf{e}_{\text{inertia}}),$$

where $\mathbf{e}_{\text{inertia}} \in \mathbb{R}^6$ is also Gaussian noise. Note that the exponential map here is applied only to a part of the vector $\mathbf{e}_{\text{inertia}}$ that corresponds to rotation, and is an identity mapping on the rest of the vector (transition part).

Landmark Observation The main source of information is landmarks configured and tracked on the surface of the asteroid, from the observed images $\mathrm{I}_{1:N}$. We adopt a simple pinhole camera model:

$$\mathbf{y}_{i,j} = \left(\mathrm{K} \circ \mathbf{z}_i^{-1} \circ \mathbf{x}_i \circ \mathrm{T}_{\boldsymbol{G}\to\boldsymbol{B}}\right)\mathbf{l}_{c_{i,j}} + \mathbf{e}_{\text{landmark}},$$

where K is the internal matrix of the camera and $c_{i,j} \in \{1, \ldots, L\}$ denotes data association, that is, which landmark is refereed by the j-th feature point in the i-th image. The noise $\mathbf{e}_{\text{camera}} \in \mathbb{R}^2$ is assumed to follow a zero-mean Gaussian.

Asteroid Dynamics While dynamics of the spacecraft are described quite simply by the attitude and inertia observations, we have to elaborate on the model of asteroid's dynamics. We follow the model of a rotating object by Tweddle et al. [2][3]:

$$\begin{bmatrix}\mathbf{x}_{i+1}\\ \dot{\mathbf{x}}_{i+1}\end{bmatrix} = f(\mathbf{x}_i, \dot{\mathbf{x}}_i) + \exp(\mathbf{e}_{\text{asteroid}}),$$

where $f(\mathbf{x}_i, \dot{\mathbf{x}}_i)$ is the prediction of $\mathbf{x}_{i+1}$ and $\dot{\mathbf{x}}_{i+1}$ with numerical integration, and $\mathbf{e}_{\text{asteroid}} \in \mathbb{R}^{12}$ is a noise vector, whose four subvectors in $\mathbb{R}^3$ correspond to rotation, transition and its derivatives respectively, and $\mathbf{e}_{\text{asteroid}}$ follows a zero-mean Gaussian with covariance Λ_i. The prediction $f(\mathbf{x}_i, \dot{\mathbf{x}}_i)$ is obtained as follows.

$$\log\left(f(\mathbf{x}_i, \dot{\mathbf{x}}_i)\right) = \log\left(\begin{bmatrix}\mathbf{x}_i\\ \dot{\mathbf{x}}_i\end{bmatrix}\right) + \int \frac{\mathrm{d}}{\mathrm{d}t} \log\left(\begin{bmatrix}\mathbf{x}_i\\ \dot{\mathbf{x}}_i\end{bmatrix}\right),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\left(\begin{bmatrix}\mathbf{x}_i\\ \dot{\mathbf{x}}_i\end{bmatrix}\right) = \begin{bmatrix} \dot{\mathbf{t}}_i \\ \frac{1}{2}(\gamma_i \dot{\boldsymbol{\alpha}}_i + \dot{\boldsymbol{\alpha}}_i \times \boldsymbol{\alpha}_i - \eta_i \boldsymbol{\alpha}_i) \\ \frac{1}{m}\mathrm{W}_1 \\ \mathrm{J}^{-1}(-\dot{\boldsymbol{\alpha}}_i \times \mathrm{J}\dot{\boldsymbol{\alpha}}_i + \mathrm{W}_2) \end{bmatrix},$$

where W_1 and W_2 denote random disturbance force and torque, and γ_i and η_i are calculated as:

$$\theta_i = |\boldsymbol{\alpha}_i|,$$

[2] In the following parts of this paper, we use the term *pose* to denote both translation and rotation in each reference frame, whereas in some literatures the former is referred to as position and the latter as pose or attitude.

[3] Tweddle et al. [2, 3] adopted modified Rodrigues parameters to represent small rotation, while we used the logarithm map and representation in the tangent space, just for theoretical simplicity.

$$\gamma_i = \begin{cases} \theta_i \cot\left(\frac{\theta_i}{2}\right) & (\theta_i \gg 0) \\ \frac{12 - \theta_i^2}{6} & (\theta_i \to 0) \end{cases},$$

$$\eta_i = \begin{cases} \frac{\dot{\boldsymbol{\alpha}}_i \cdot \boldsymbol{\alpha}_i}{\theta_i}\left(\cot\left(\frac{\theta_i}{2}\right) - \frac{2}{\theta_i}\right) & (\theta_i \gg 0) \\ \frac{60 + \theta_i^2}{360}\dot{\boldsymbol{\alpha}}_i \cdot \boldsymbol{\alpha}_i & (\theta_i \to 0) \end{cases}.$$

We used a new notation $\boldsymbol{\alpha} = \log(\mathbf{r})$, and $\log(\cdot)$ is the logarithm map of rotation from $\mathrm{SO}(3)$ into $\mathbb{R}^3$.

We adopted the numerical integration with adaptive step-size RK4 described in the literature [2], where $\dot{\boldsymbol{\alpha}}_i$ is assumed to be constant within each time step to ensure that above equation is linearized and temporally discretized. This assumption is to estimate the covariance Λ_i numerically, utilizing the fact that it follows Lyapunov equation. For the detailed description, see [2] or [3].

Priors For stable computation, prior distributions are set. In this paper every prior is modeled with zero-mean Gaussian, and corresponding quantities are initialized with zero. For the first asteroid's pose $\mathbf{x}_1$, we should set small variances, say 10^{-6}, to make the first pose to be the origin of the asteroid frame, and variances of $\dot{\mathbf{x}}_1$ to almost zero for transition and π or 2π for rotation. Variances of a prior on $\mathrm{T}_{\boldsymbol{G}\to\boldsymbol{B}}$ are should be large enough.

2.3 Main Algorithm

In Algorithm 1, we show the main algorithm to infer the unknown quantities given the observed ones, under the model presented in the previous section. Its overview is as follows: First, correspondences between the observed feature points and landmarks are computed in Line 1, as correctly as possible. Then some quantities are initialized with zero in Line 2, as stated in the previous section. Afterward, iterative updates, which consist of two stages, start in Line 3. At the first stage in Lines 3–5, a new frame is added at each iteration and the solution so far is updated, using only the initial data association of Line 1. At the second stage in Lines 6–8, we add new data association if possible, utilizing both the descriptors and the poses of the asteroid and the spacecraft, with the whole solution updated.

Initial Data Association The data association of Line 1 starts with feature-point matching with the descriptors $\mathbf{d}_{1:N,1:M}$, as in the literature [4]. We matched a feature point y_{1,j_1} in the image I_1 and another feature point y_{2,j_2} in I_2 if and only if the following condition holds:

$$1.5 \times \mathrm{dsim}(\mathbf{d}_{1,j_1}, \mathbf{d}_{2,j_2}) \leq \max_{j' \in \{1,\ldots,M_2\}/j_2} \mathrm{dsim}(\mathbf{d}_{1,j_1}, \mathbf{d}_{2,j'}),$$

Algorithm 1 Asteroid Visual SLAM.

```
AsteroidSLAM(s_{1:N}, u_{2:N}, y_{1:N,1:M}, d_{1:N,1:M})
1 c_{1:N,1:M} ←InitAssoc(y_{1:N,1:M}, d_{1:N,1:M})
2 x_1, ẋ_1, z_1, T_{G→B}, k
    ←InitFirstFrame(s_{1:2}, u_2, y_{1:2,1:M}, c_{1:2,1:M})
3 for i = 2, ..., N   // First Stage
4   x_i, ẋ_i, z_i, l_{L^new(i)}
      ←InitNewFrame
        (s_{i−1:i}, u_i, y_{i−1:i,1:M}, c_{i−1:i,1:M},
         x_{i−1}, ẋ_{i−1}, z_{i−1}, l_{L(i−1)}, T_{G→B}, k)
5   x_{1:i}, ẋ_{1:i}, z_{1:i}, l_{L(i)}, T_{G→B}, k
      ←SolutionUpdate
        (x_{1:i}, ẋ_{1:i}, z_{1:i}, l_{L(i)}, T_{G→B}, k)
6 for i = 2, ..., N   // Second Stage
7   c_{i,1:M_i}
      ←AddAssoc(y_{i,1:M_i}, d_{i,1:M_i}, x_i, z_i, T_{G→B})
8   x_{1:N}, ẋ_{1:N}, z_{1:N}, l_{1:L}, T_{G→B}, k
      ←SolutionUpdate
        (x_{1:N}, ẋ_{1:N}, z_{1:N}, l_{1:L}, T_{G→B}, k)
9 return x_{1:N}, ẋ_{1:N}, z_{1:N}, T_{G→B}, k, l_{1:L}
```

where $\mathrm{dsim}(\cdot,\cdot)$ denotes dissimilarity between two descriptors, and we use Euclidean distances for the dissimilarity. This condition means that feature points y_{1,j_1} and y_{2,j_2} are matched when the dissimilarity between y_{1,j_1} and y_{2,j_2} is smaller enough than one between y_{1,j_1} and feature points in the image I_2 other than y_{2,j_2}.

After computing the matching for possible pairs of feature points, its errors are eliminated using RANSAC [5] with estimation of a homography. Finally we assigned a landmark number to each connected component of the graph build with the feature-points matching.

First Stage Update At each iteration in the first stage (Lines 3–4), given observations at a new timestamp, unknown quantities with regard to them are initialized (Line 4), and all the variables initialized so far are updated (Line 5).

The initialization of the new frame is done either with the five point algorithm [6] or just integrating the previous $\dot{\mathbf{x}}$ numerically. We adopted the former for $i = 2$ and the latter for $i > 2$, because the latter made the succeeding updates more stable. Also we found that initializing the new frame as a P n P problem led the solution to be unstable because of inaccurate estimation of landmarks' geometry. The landmark position $\mathbf{l}$ can be initialized with a simple triangulation.

The solution update is done within a framework of the incremental smoothing and mapping (iSAM) [7], where an incremental optimization on the unknown quantities is carried efficiently by factorization of a

matrix that is maintained sparse. In other words, within the iSAM framework, the unknown quantities are represented with a factor graph as shown In Figure 1, and the corresponding joint distribution is linearized and maximized as least squares. We utilized an iSAM implementation in the library GTSAM[4] [8] with slight modification for asteroid's pose, the ratio of inertia moments, and the numerical integration.

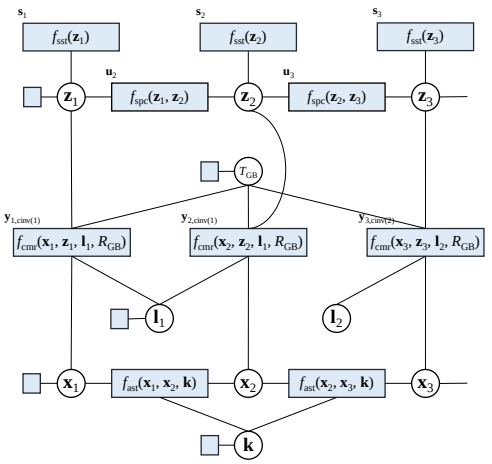


Figure 1 A factor graph representing the joint distribution. The circles correspond to unknown quantities, and the squares correspond to relations between the connected quantities, which are specified by the models presented in Section 2.2.

Second Stage Update After going through all the timestamps once in the first stage, they are revisited in another loop in Lines 6–8. Purposes of this second loop are to salvage feature points that were not labeled as any landmarks in the first data association (Line 1), and to eliminate feature points that were erroneously assigned to the landmarks in Line 1. This procedure will improve density the reconstructed shape of the asteroid and accuracy of the estimation.

The way to add data association (Line 7) is similar to that in the first data association, but the range of candidate feature points is geometrically limited now. We search for a feature point y_{2,j_2} in image I_2 that is matching with a feature point y_{1,j_1} in I_1, only around the point where y_{1,j_1} is anticipated to be in I_2. The position of the feature points can be anticipated since poses of the asteroid and spacecraft are already obtained in the first stage of estimation. Incorporating geometric information in data association will be useful to make the descriptors more distinctive, because the main reason of failure of the feature-points matching only with the descriptors is distraction by totally unrelated feature points in distant regions.

[4] https://collab.cc.gatech.edu/borg/gtsam/

2.4 Tips on Implementation

Here we list empirical tips to implement the algorithm stated above.

- Storing large rotation: Since the exponential map $\exp(\boldsymbol{\alpha})$ has singularities on spheres $|\boldsymbol{\alpha}| = 2n\pi \ (n = 1,2,3,\dots)$, the rotation part of $\mathbf{x}$ should be stored within two parts, that is, a reference part and a small rotation part. During the solution updates, the former is kept constant and the latter is adjusted; sometimes they are merged into the reference part and the small part is reset to be zero.
- Unstable triangulation: The optimization often gets unstable as a large value of depth is obtained at the triangulation because of insufficient amount of observation. We found that simply forcing the average of the depth zero helps to make the optimization more stable.
- Favorable methods for initialization: As is often the case with monocular visual SLAM, the best methods to initialize new frames is not obvious and depends on the nature of problems, so a flexible mechanism for operators to empirically choose the initialization method is essential.

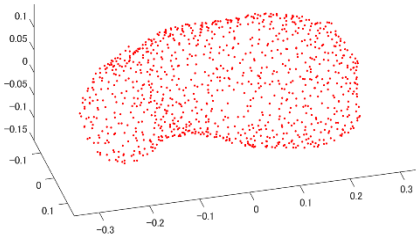
3 EXPERIMENT

This sections presents results of an experiment we conducted with artificially-generated data. The experiment is preliminary, but enough to confirm availability of our framework for the SLAM above an asteroid.

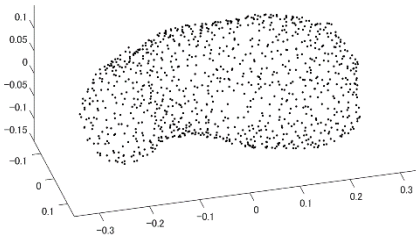
3.1 Setup

We used a shape model [9] of the asteroid 25143 Itokawa as the ground truth, with random subsampling to 1,000 vertices as shown in Figure 2 (a). Though we used a model of Itokawa, physical conditions except size (the longest axis ~ 535 [m]) are configured slightly differently from its actual situation for experimental clarity.

To generate observations of the landmarks, the shape of the asteroid was rotated around an axis by 0.06 [rad] per image, and its vertices were observed by a camera that was randomly moving following Gaussians with $\mathrm{std} = 10$ [m] for translation and $\mathrm{std} = 0.001$ [rad] for rotation. The observations were disturbed by a Gaussian noise with $\mathrm{mean} = 0$ and $\mathrm{std} = 0.1$ [px]. Also, measurements of the attitude and inertia sensors were generated with zero-mean Gaussian noises.



(a)



(b)

Figure 2 (a) True and (b) reconstructed shapes of the asteroid.

3.2 Results

We run only the first stage of Algorithm 1 with the data generated as stated in the previous section, since these parts will dictate validity of the solution.

The reconstruction of asteroid's shape is shown in Figure 2 (b), which we can hardly distinguish from the ground truth. The RMS error after configuration by the ICP algorithm was 0.036 [m].

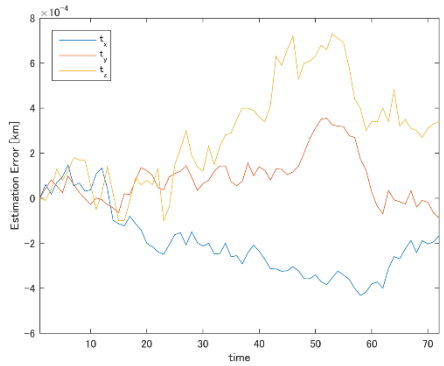
In Figure 3, estimation errors of spacecraft's pose are plotted. We can see that in Figure 3 (a) the maximum error is about 0.8 [m], and the errors of both translation and rotation do not diverge.

3.3 Discussion

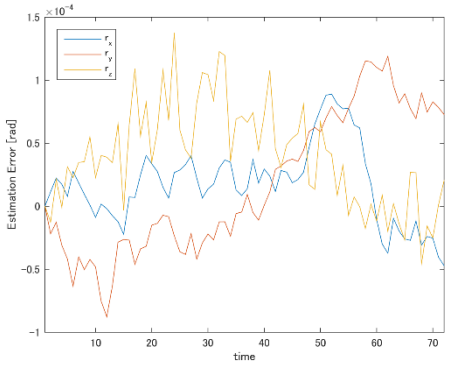
Since the experiment conducted in this paper is preliminary and not enough to claim that our method is the best, we are successful at accurately estimate the quantities needed in asteroid exploration, such as asteroid's geometry and spacecraft's pose, using a SLAM framework based on visual observations and attitude and inertia sensors.

We must proceed experiments with attention to the following points: We should follow more severe physical conditions such as larger noises, nutation, and absence of attitude/inertia sensors. Also, tests on optical images are essential, while we resorted to direct generation of landmarks [5] . Moreover, evaluating how the second stage improves estimation is necessary of course.

[5] We obtained successful results with optical images using image descriptors, on a different algorithm [10].



(a)



(b)

Figure 3 Estimation errors of spacecraft's pose: respectively for (a) translation and (b) rotation.

4 RELATED WORK

In previous missions of exploration of small celestial bodies, optical images taken nearby a target were utilized. However, they depended much on human operators in landmark annotation on images and in information fusion on different methods and sensors. On the other hand, mainly in areas of robotics and computer vision, solutions to SLAM and structure from motion (SFM) have been more and more sophisticated in this decade, which are fully automatic procedures. Below is briefly noted related work on these two distinct topics.

Previous Explorations In the mission of MUSES-C (Hayabusa) of JAXA, the shape of the asteroid Itokawa for spacecraft navigation was build using the limb profile of the asteroid and the stereo-photogrammetry (ordinal stereo method) [11, 12], and the navigation was conducted with manual tracking of landmarks [13]. Another model was created using a photometric stereo technique [9].

In the mission of Rosetta of ESA, reconstruction of the shape of the asteroid Lutetia [14] and of the comet Churyumov-Gerasimenko [15] was conducted with the stereo-photogrammetry method.

Also, the photometric stereo technique as in [9] was applied to the images of Churyumov-Gerasimenko [16].

SLAM and SFM Though there are a large number of SLAM/SFM solution in general, there are not so many ones that explicitly focus on a rotating target and estimation of its parameters: Lichter and Dubowsky [17], Aghili [18] and Tweddle et al. [2] for example. Note that these previous studies assume range sensors such as laser cameras and stereo cameras, which are not necessarily applicable to the asteroid exploration.

There have been attempts to apply the SLAM/SFM to the asteroid exploration. For example, Cocaud and Kubota tested a SLAM solution based on the Rao-Blackwellized particle filter [19], and Mori and Hirata utilized an open-source software for bundle adjustment SFM and the patch-based stereo [1]. Ideas of these studies look similar to ours, but they do not explicitly model and estimate parameters of asteroid's rotation, hence needs additional analyses.

5 CONCLUSION

In this paper, a framework of the simultaneous localization and mapping (SLAM) dedicated to asteroid explorations is proposed. This framework is to estimate asteroid's shape, centroid, rotational axis, angular velocity and phase, as well as spacecraft's pose and position relative to the asteroid. The estimation is done utilizing observations by optical camera, attitude sensors and inertia sensors equipped to the spacecraft.

We showed preliminary results of the proposed method, and confirmed its validity, while we need much additional experiment as stated in Section 3.3. For further experiments, we are planning to create new data that follow actual physics using a high-speed camera with rotating object, because it is available in an ordinal environment on the ground, and the proposed framework is easily modified to consider the gravity.

One of points to extend the proposed framework is a more sophisticated technique of initialization, which is now based on empirical adjustment. A rough estimation using non-iterative methods such as matrix factorization might be a solution. More challenging but interesting solution would be to incorporate ground-based observation of the target celestial body.

The most important point to be improved is computational efficiency. Since our framework can be run in an online (real-time) manner, making the computation much efficient will open the way to fully on-board shape estimation and optical navigation.

Acknowledgement

We greatly appreciate members of Hayabusa-2 Global Mapping Team at Japan Aerospace Exploration Agency (Yuichi Tsuda, Fuyuto Terui, Naoko Ogawa and Yuya Mimasu) who gave us much insightful discussion.

This work was supported by JSPS KAKENHI Grant Number 15J09172.

References

[1] Mori Y, Hirata N and Hayabusa-2 shape reconstruction study team (2014) Asteroid Shape Reconstruction by Open-source Structure from Motion Tools. In the 45th Lunar and Planetary Science Conference.

[2] Tweddle BE et al. (2015) Factor Graph Modeling of Rigid-body Dynamics for Localization, Mapping, and Parameter Estimation of a Spinning Object in Space. *Journal of Field Robotics* 32(6):p897–p933.

[3] Tweddle BE (2013) Computer Vision-based Localization and Mapping of an Unknown, Uncooperative and Spinning Target for Spacecraft. Ph.D. dissertation, The Massachusetts Institute of Technology.

[4] Lowe DG (2004) Distinctive Image Features from Scale-invariant Keypoints. *International Journal of Computer Vision* 60(2):p91–p110.

[5] Fischler MA and Bolles RC (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communication of the ACM* 24(6):p381–p395.

[6] Nistér D (2004) An Efficient Solution to the Five-point Relative Pose Problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(6):p756–p770.

[7] Kaess M, Ranganathan A and Dellaert F (2008) iSAM: Incremental Smoothing and Mapping. *IEEE Transactions on Robotics* 24(6):p1365–p1378.

[8] Dellaert F (2012) Factor Graphs and GTSAM: A Hands-on Introduction. Technical Report, GT-RIM-CP&R-2012-002.

[9] Gaskell R et al. (2006) Landmark Navigation Studies and Target Characterization in the Hayabusa Encounter with Itokawa. In AIAA/AAS Astrodynamics Specialist Conference and Exhibit, AIAA 2006-6660.

[10] Takeishi N et al. (2015) Simultaneous Estimation of Shape and Motion of an Asteroid for Automatic Navigation. In *Proceedings of 2015 IEEE International Conference on Robotics and Automation* p2861–p2866.

[11] Maruya M, Ohyama H and Uo M (2006) Navigation Shape and Surface Topography Model of Itokawa. In AIAA/AAS Astrodynamics Specialist Conference and Exhibit, AIAA 2006-6659.

[12] Demura H et al. (2006) Pole and Global Shape of 25143 Itokawa. *Science* 312(5778):p1347–p1349.

[13] Shirakawa K et al. (2006) Accurate Landmark Tracking for Navigating for Navigating Hayabusa prior to Final Descent. In *Advances in the Astronautical Sciences* 124:p1817–p1826.

[14] Preusker F et al. (2012) The Northern Hemisphere of Asteroid (21) Lutetia—topography and Orthoimages from Rosetta OSIRIS NAC Image Data, *Planetary and Space Science* 66(1):p54–p63.

[15] Preusker F et al. (2015) Shape model, reference system definition, and cartographic mapping standards for comet 67P/Churyumov-Gerasimenko—Stereo-photogrammetric analysis of Rosetta/OSIRIS image data, *Astronomy & Astrophysics* 583.

[16] Jorda L et al. (2014) The Shape of Comet 67P/Churyumov-Gerasimenko from Rosetta/OSIRIS Images. In AGU Fall Meeting 2014.

[17] Lichter MD and Dubowsky S (2004) State, Shape and Parameter Estimation of Space Objects from Range Images. In *Proceedings of 2004 IEEE International Conference on Robotics and Automation* p2974–p2979.

[18] Aghili F (2012) A Prediction and Motion-planning Scheme for Visually Guided Robotic Capturing of Free-floating Tumbling Objects with Uncertain Dynamics. *IEEE Transactions on Robotics* 28(3):p634–p649.