The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

# Hopping Odometry: Motion Estimation with Selective Vision

Edmond Wai Yan So, Tetsuo Yoshimitsu, and Takashi Kubota

*Abstract*— **We present a two-step iterative algorithm to estimate the trajectory of a hopping rover. In the first step, a monocular scheme of visual odometry is adapted to estimate an initial portion the hopping trajectory. From this, the parameters for the ballistic motion are recovered, and the trajectory is extrapolated to predict the positions of the rover for the remainder of the hop. In the second step, we devise a scheme called "*selective vision*", combining the ideas of active vision and guided search. An envelope lying between the start and end of a hop is defined, within which features most likely to be re-observed across a hop are detected and matched. Performing pose estimation on the these matched features allow the relative motion between a camera frame within the visual odometry step and a camera frame within the extrapolated trajectory to be estimated. The newly determined camera frame in the extrapolated trajectory can then be used to refine the parameters of the ballistic motion, and the trajectory can be re-extrapolated to predict future positions of the hopping rover. Following this scheme, it is possible to estimate the trajectory of a hopping rover undergoing continuous rotational motion with only one set of cameras without continuous tracking of terrain features.**

## I. INTRODUCTION

In the microgravity environment on the surface of an asteroid, the effectiveness of a wheeled rover can be drastically reduced due to the loss of traction. An attractive alternative for locomotion is hopping, because only a small force is needed to generate vertical motion, and a hopping rover can navigate over obstacles with little path-planning. Numerous hopping rovers with different designs for the actuating mechanism have been proposed [1]–[4]. Most significantly, the MINERVA rover [5] was a part of the Hayabusa mission to the asteroid 25143 Itokawa, although its deployment to the surface was unsuccessful.

When exploring on the remote surface of an asteroid, the operations and control of a hopping rover can be greatly improved by autonomous navigation, for which accurate localization is a prerequisite. Absolute localization refers to the positioning of a mobile robot with respect to an asteroid-fixed coordinate frame. Higo [6] has proposed the use of radio ranging with an orbiter for this purpose. Relative localization refers to the positioning of a mobile robot with respect to some local coordinate frame, such as a nearby landmark or the beginning of a long traverse. Accurate relative localization enables a mobile robot to navigate to a certain target with high precision, allowing measurements or sampling to be made of a particular feature. On wheeled robots, relative localization is usually accomplished with wheel odometry; this is obviously not applicable to a hopping rover. On aerial vehicles, relative localization is typically accomplished with inertial sensors such as accelerometers and gyroscopes. Fiorini [7] has applied this technique to perform relative localization on hopping rovers. However, as they have noted, this is subject to accumulation error over time. In this paper, we consider the use of visual odometry.

Visual odometry refers to the use of images to estimate the motion of a mobile robot. This has been successfully demonstrated by the Mars Exploration Rover [8], and an updated version is being developed for the Mars Science Laboratory [9]. A real-time implementation has also been demonstrated for terrestrial robots [10].

In this paper, we adapt the technique of visual odometry to a hopping rover to estimate its initial motion. We then propose a method called "*selective vision*", in which the features most likely to be re-observed across a hop are used for feature matching and pose estimation. In Section II, we describe the kinematics of hopping motion. In Section III, we demonstrate the adaptation of visual odometry to a hopping rover. In Section IV, we explain our concept of selective vision, where the hopping kinematics described in Section III is used to guide our visual observations. We give conclusions in Section V.

## II. HOPPING KINEMATICS

The motion of a hopping rover can be roughly divided into three different phases, as shown in Fig. 1: initial hop, secondary bounces, and settling motion.

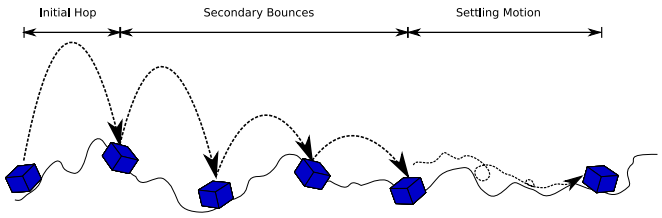



Fig. 1. Phases of Hopping Motion

At the beginning of the initial hop, the takeoff velocity is largely controlled by the hopping mechanism, although gravity and friction will limit the achievable hopping angle [11]. At the end of the initial hop, the rover rebounds from the surface, beginning another shorter hop. The magnitude and direction of the rebounding velocity is determined by the coefficients of restitution and friction, and the surface

E. So is with the School of Physical Sciences, Department of Space and Astronautical Science, The Graduate University for Advanced Studies, 3-1-1 Yoshinodai, Sagamihara-shi, 229-8510, Japan edmond@nnl.isas.jaxa.jp

T. Yoshimitsu and T. Kubota are with the Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, 3-1-1 Yoshinodai, Sagamihara-shi, 229-8510, Japan kikko@nnl.isas.jaxa.jp, kubota@isas.jaxa.jp

geometry. The rebounding will continue at the end of every hop until the vertical velocity of the rover has been dampened out. After these bounces have subsided, residual horizontal and angular momentum causes sliding and rolling, resulting in tumbling motion. These kinetic energies are eventually dissipated through surface friction, at which time the rover will come to a complete stop.

### A. Coordinate Frames

Across an undulating terrain, we define a ground plane that approximates the average height of the terrain. The local vertical is defined as the vector perpendicular to this ground plane. We define a world coordinate frame $\mathcal{F}_W = [\mathcal{X}_W, \mathcal{Y}_W, \mathcal{Z}_W]^T$ with the $\mathcal{X}_W - \mathcal{Y}_W$ plane aligned with the ground plane and the $\mathcal{Z}_W$-axis aligned with the local vertical (Fig. 2).

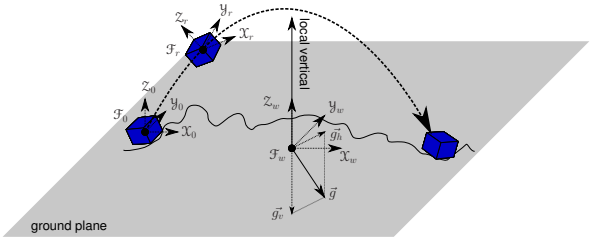



Fig. 2. Ground Plane, Local Vertical, and Gravity Vector

We define a rover coordinate frame $\mathcal{F}_r = [\mathcal{X}_r, \mathcal{Y}_r, \mathcal{Z}_r]^T$ fixed to the rover. The initial coordinate frame $\mathcal{F}_0 = [\mathcal{X}_0, \mathcal{Y}_0, \mathcal{Z}_0]^T$ is defined to coincide with the rover coordinate frame $\mathcal{F}_r$ at the beginning of a hop. The $\mathcal{X}_r$-axis is aligned with the forward direction of the rover, and the $\mathcal{Z}_r$-axis is aligned with the top of the rover, so that $\mathcal{Z}_0$ is aligned with the surface normal $\hat{n}$ of the takeoff surface, and the $\vec{v_0}$ lies in the $\mathcal{X}_0$-$\mathcal{Z}_0$ plane (Fig. 3).

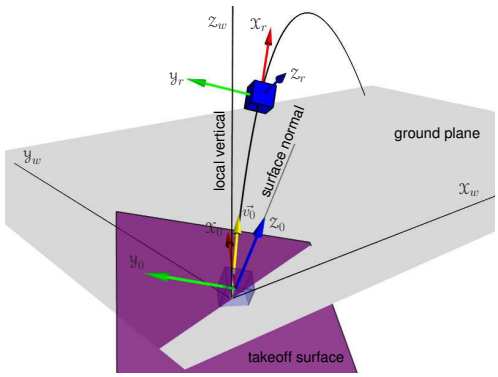



Fig. 3. World, Initial, and Rover Coordinate Frames

### B. Translational Motion

On an asteroid, due to irregular shape and uneven mass distribution, the gravity vector will not be aligned with the local vertical in general. Thus, with respect to the world coordinate frame $\mathcal{F}_W$, the gravity vector $\vec{g} = [g_x, g_y, g_z]^T$ will have both a vertical component $\vec{g}_v = g_z\hat{k}$ and a non-zero horizontal component $\vec{g}_h = g_x\hat{i} + g_y\hat{j}$.

If we assume the gravity field to be constant across a single hop, the trajectory of a hopping rover can be described by the equations of ballistic motion. With initial position $\vec{x}(0) = [x_0, y_0, z_0]^T$ and initial velocity $\vec{v}_0 = [v_{0x}, v_{0y}, v_{0z}]^T$, the trajectory of the hopping rover $\vec{x}(t)$ is:

$$\begin{aligned} \vec{x}(t) &= \vec{x}_0 + \vec{v}_0 t + \frac{1}{2}\vec{g}t^2 \\ \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} &= \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} v_{0x} \\ v_{0y} \\ v_{0z} \end{bmatrix} t + \frac{1}{2}\begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} t^2 \end{aligned} \tag{1}$$

Thus, all three components of the position vector $\vec{x}(t)$ are affected by gravitational acceleration.

The trajectory $\vec{x}(t)$ is a linear combination of the initial velocity $\vec{v_0}$ and the gravity vector $\vec{g}$. Thus, it lies in a plane defined by these two vectors (Fig. 4).

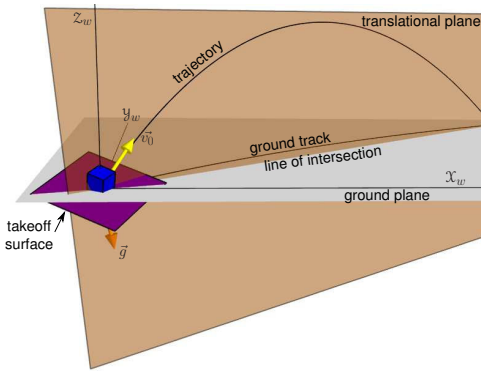



Fig. 4. Plane of Translational Motion

If the height of the ground is the same at the point of takeoff and at the point of landing, the traversal time $T$ of the initial hop is determined by the vertical components of the initial velocity and gravitational acceleration:

$$T = -2\frac{v_{0z}}{g_z} \tag{2}$$

The maximum height $z_{max}$ attained is:

$$z_{max} = -\frac{{v_{0z}}^2}{2g_z} \tag{3}$$

The total distance traversed $D$ along the ground plane is:

$$\begin{aligned} D &= |\vec{x}_{landing} - \vec{x}_0| \\ &= \sqrt{(x(T) - x_0)^2 + (y(T) - y_0)^2} \\ &= \sqrt{\left(v_{0x}T + \frac{1}{2}g_xT^2\right)^2 + \left(v_{0y}T + \frac{1}{2}g_yT^2\right)^2} \end{aligned} \tag{4}$$

### C. Rotational Motion

A hopping rover typically attains its initial velocity $\vec{v_0}$ through surface reaction forces. This consists of a surface normal force $\vec{f_n}$ that gives the rover its vertical motion, and friction force $\vec{f_{fr}}$ that gives the rover its horizontal motion

(Fig. 5). In particular, the friction force $\vec{f}_{fr}$ will act along the surface, and will be offset from the rover's center of mass. This will result in rotational motion.

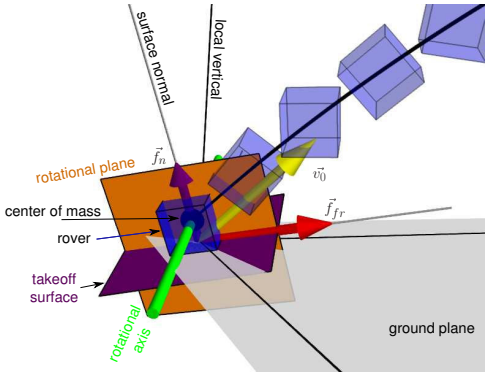



Fig. 5. Rotational Motion of Hopping Rover

Without attitude control, a hopping rover will undergo continuous rotational motion with constant angular velocity throughout a hop. Given the initial orientation (expressed as a quaternion) $q_0 = [q_{0w}, q_{0x}, q_{0y}, q_{0z}]^T$ with respect to the world coordinate frame $\mathcal{F}_W$, and the initial angular velocity $\vec{\omega} = [\omega_x, \omega_y, \omega_z]^T$, the orientation $q(t)$ of a hopping rover at time $t$ during the initial hop is:

$$q(t) = q_0 \otimes q(\vec{\omega}t)$$

$$\begin{bmatrix} q_w(t) \\ q_x(t) \\ q_y(t) \\ q_z(t) \end{bmatrix} = \begin{bmatrix} q_{0w} \\ q_{0x} \\ q_{0y} \\ q_{0z} \end{bmatrix} \otimes \begin{bmatrix} cos\frac{|\vec{\omega}|t}{2} \\ \frac{\omega_x}{|\vec{\omega}|} sin\frac{|\vec{\omega}|t}{2} \\ \frac{\omega_y}{|\vec{\omega}|} sin\frac{|\vec{\omega}|t}{2} \\ \frac{\omega_z}{|\vec{\omega}|} sin\frac{|\vec{\omega}|t}{2} \end{bmatrix} \quad (5)$$

Ideally, across the initial hop, the axis of rotation $\hat{\omega}$ will be perpendicular to the plane defined by the initial velocity $\vec{v}_0$ and the surface normal $\hat{n}$ of the takeoff surface. On uneven terrain, the takeoff surface will generally not be parallel to the ground plane, so the surface normal $\hat{n}$ will not be aligned with the local vertical $\hat{k}$ . Also, the surface normal $\hat{n}$ will generally not be aligned with the gravity vector $\vec{g}$. Thus, in general, the plane of translation motion, which is defined by the initial velocity $\vec{v}_o$ and the gravity vector $\vec{g}$, will not coincide with the plane of rotational motion, which is defined by the initial velocity $\vec{v}_0$ and the surface normal $\hat{n}$ of the takeoff surface (Fig. 6).

## III. HOPPING ODOMETRY

In this section, we describe the adaptation of conventional visual odometry to a hopping rover. The motion estimation is based solely upon a perspective projection camera model, independent of the hopping kinematics described in the previous section.

### A. Stereo Depth Resolution

We choose to use a monocular approach for visual odometry, because for a hopping rover, a limited stereo baseline and large vertical displacement results in very poor stereo depth

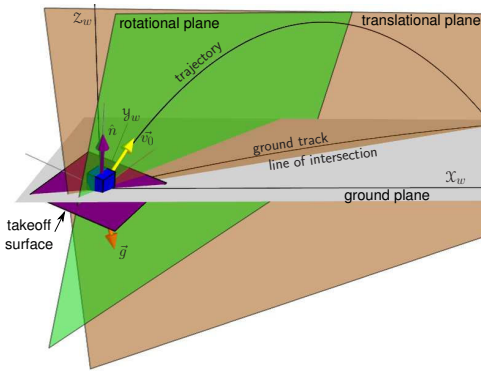



Fig. 6. Planes of Translation and Rotational Motion

resolution. The depth resolution of a pair of stereo cameras (Fig. 7) can be approximated by the following equation [12]:

$$res_d \approx \frac{s_{px} {Z_{scene}}^2}{fb} \quad (6)$$

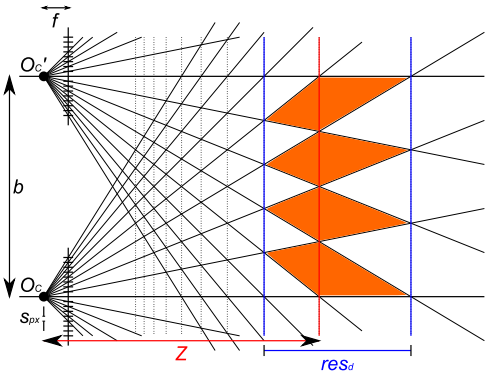



Fig. 7. Stereo Depth Resolution

Thus, for a pair of stereo cameras with focal length $f$ and pixel size $s_{px}$, the depth resolution $res_d$ is inversely proportional to the baseline $b$, and quadratically proportional to the scene distance $Z_{scene}$.

On a hopping rover, the baseline $b$ available for stereo vision is typically very limited. For example, the MINERVA hopping rover was a hexagonal prism with a diameter of 12 cm and a height of 10 cm, and the stereo cameras had a baseline of $b = 3\ cm$ [13].

On the other hand, the scene distance $Z_{scene}$ varies over a large range across a hopping trajectory. On a planetary surface, the observed scene is the terrain on the ground plane; thus, the scene distance is effectively the distance of the rover from the ground plane, $Z_{scene} = z(t)$. From Eq. 3, the maximum height attained during a hop is a function of the initial vertical velocity $v_{0z}$ and the vertical component of the gravitational acceleration $g_z$. In turn, the initial velocity $\vec{v_0}$ of a hopping rover on an asteroid surface is primarily limited by the escape velocity $v_e$. Using the dynamic parameters on the asteroid 25143 Itokawa [14], $|\vec{v_0}| = v_e \approx 0.1\ m/s$, $g_z = 1 \times 10^{-4}\ m/s^2$, for a hopping elevation angle $\theta_e = 45°$

so that $v_{0z} = 0.0707\ m/s$, the maximum height is $z_{max} = 25\ m$.

Using the parameters of the stereo cameras on the MINERVA rover [13], $s_{px} = 7.2\ \mu m$, $f = 2.8\ mm$, and $b = 3\ cm$, its depth resolution $res_d$ as a function of scene distance $Z_{scene}$ is illustrated by Fig. 8. When resting on the surface of an asteroid, with the cameras at a height of $h = 5\ cm$, and a vertical field-of-view of $34.3°$, the nearest observable terrain will be at a distance of $\approx 17\ cm$, giving a stereo depth resolution of $res_d = 2.5\ mm$. As the rover hops away from the ground, the stereo depth resolution deteriorates quadratically. At the maximum height of the hop, $z_{max} = 25\ m$, and $res_d = 53\ m$.

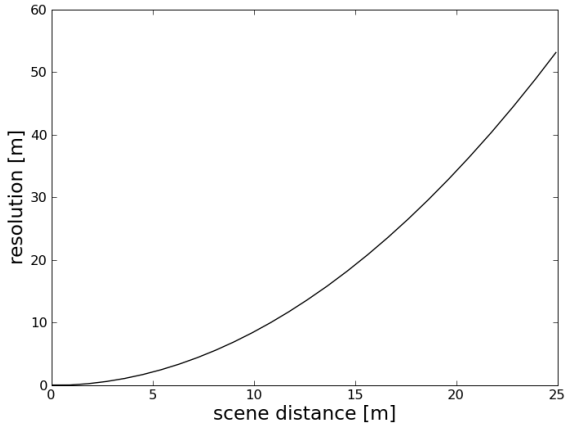



Fig. 8. Stereo Depth Resolution on MINERVA

Thus, our solution is to use small-baseline stereo vision at the beginning of a hop to triangulate features and establish the absolute scale of the scene, and use monocular vision throughout a hop to estimate the motion.

### B. Monocular Visual Odometry

We follow the monocular scheme for visual odometry as described by Nister [10]. The algorithm is outlined in Fig. 9 and illustrated in Fig. 10.

We first initialize the scene scale by using a pair of small-baseline stereo cameras to triangulate feature points in the nearby terrain. The Harris corner detector [15] is used to select features in the two images, they are matched by normalized correlation, and their 3D coordinates are estimated using linear triangulation [16]. This gives a set of features $\boldsymbol{x_0}$ with their corresponding structure $\boldsymbol{X_0}$ with respect to the initial coordinate frame $\mathcal{F}_0$.

A set of new features $\boldsymbol{x_1}$ distinct from $\boldsymbol{x_0}$ is then selected. Both set of features, $\boldsymbol{x_0}$ whose structure is known and $\boldsymbol{x_1}$ whose structure is unknown, are simultaneously tracked using the Lucas-Kanade feature tracker [17] to the next camera frame $\mathcal{F}_1$. Pose estimation [18] is then performed with the tracked features $\boldsymbol{x'_0}$ with known structure $\boldsymbol{X_0}$ within a RANSAC framework [19] to robustly recover the relative motion $\{\vec{t_0^1}, R_0^1\}$ between camera frames $\mathcal{F}_0$ and $\mathcal{F}_1$. This in turns allows the features $\boldsymbol{x_1}$ and $\boldsymbol{x'_1}$ to be triangulated,

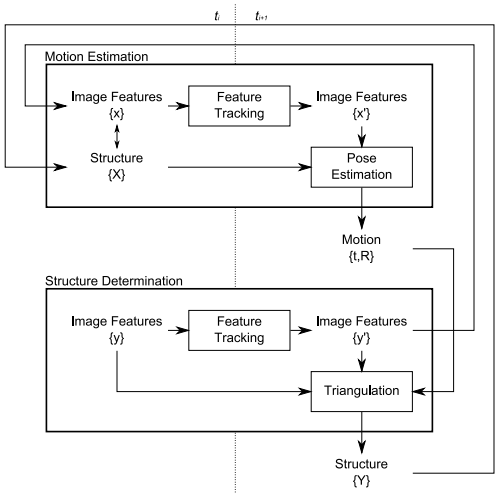



Fig. 9. Monocular Visual Odometry Algorithm

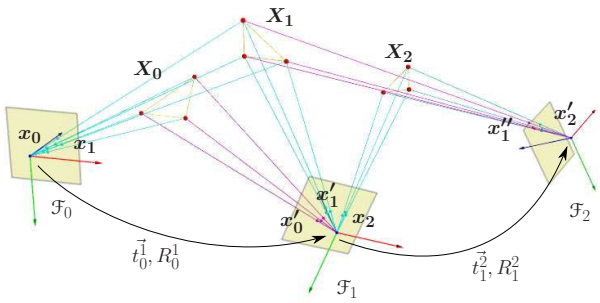



Fig. 10. Monocular Visual Odometry

giving its structure $\boldsymbol{X_1}$ with respect to camera frame $\mathcal{F}_1$. The algorithm is then repeated, with a new set of features $\boldsymbol{x_2}$ being selected, tracked along with the set of features $\boldsymbol{x'_1}$ with known structure $\boldsymbol{X_1}$ to the next camera frame $\mathcal{F}_2$, and pose estimation and triangulation are then performed.

## IV. SELECTIVE VISION

Using the monocular visual odometry described in Section III, each pose estimation produces a motion estimate $\{\vec{t_{k-1}^k}, R_{k-1}^k\}$ giving the relative motion from the previous camera frame $\mathcal{F}_{k-1}$ to the current camera frame $\mathcal{F}_k$. These motion estimates can be combined recursively to give the position $\vec{x}(t_k)$ of the hopping rover relative to the initial coordinate frame $\mathcal{F}_0$:

$$\begin{aligned} \vec{x}(t_k) &= R_0^{k-1} T_{k-1}^k + \vec{x}(t_{k-1}) \\ R_0^k &= R_0^{k-1} R_{k-1}^k \end{aligned} \tag{7}$$

However, due to the continuous rotational motion of a hopping rover, the terrain will eventually leave the field-of-view of a camera. As a result, multiple cameras pointing in different directions will be needed to provide continuous tracking of the terrain in order for the hopping trajectory to be fully reconstructed.

Alternatively, with a single camera, we propose to reconstruct the entire hopping trajectory through an iterative process (Fig. 11). First, monocular visual odometry is used to estimate the hopping motion until the terrain moves out of view of the camera. The parameters for a ballistic trajectory, namely, the initial velocity $\vec{v_0}$ and the gravitational acceleration $\vec{g}$, can then be recovered. This ballistic trajectory is then extrapolated, so that the position of the rover for the remainder of the hop can be predicted. This prediction is used to guide a selection process, where if the rover is predicted to be oriented towards previously observed features, the image taken at that position is selected for feature matching and pose estimation. This generates a new position estimate in the hopping trajectory, so the parameters of the ballistic motion can be refined, and further images are selected for feature matching and pose estimation based on the re-extrapolated trajectory.

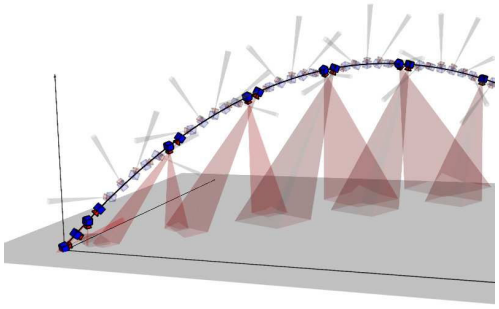


Fig. 11. Selective Vision

We estimate the two parameters of the ballistic trajectory, the initial velocity $\vec{v_0}$ and the gravitational acceleration $\vec{g}$, using only the position estimates $\vec{x}(t_1), \ldots, \vec{x}(t_n)$. Eq. 1 (with $\vec{x}_0 = \vec{0}$ because the position estimates $\vec{x}(t_k)$ are relative to the initial coordinate frame $\mathcal{F}_0$) can be rewritten as:

$$\begin{bmatrix}\vec{x}(t_k)^T\end{bmatrix} = \begin{bmatrix}t_k & \frac{1}{2}t_k^2\end{bmatrix}\begin{bmatrix}\vec{v_0}^T \\ \vec{g}^T\end{bmatrix} \tag{8}$$

A least-squares solution for the ballistic parameters can then be obtained:

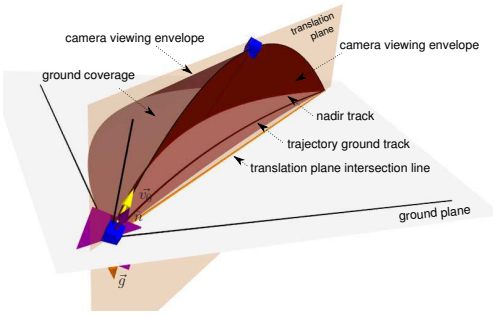
$$\begin{bmatrix}\vec{v_0}^T \\ \vec{g}^T\end{bmatrix} = (A^TA)^{-1}A^T\begin{bmatrix}\vec{x}(t_1)^T \\ \vdots \\ \vec{x}(t_n)^T\end{bmatrix}, A = \begin{bmatrix}t_1 & \frac{1}{2}t_1^2 \\ \vdots & \vdots \\ t_n & \frac{1}{2}t_n^2\end{bmatrix} \tag{9}$$

This batch least-squares estimation can be reformulated as a recursive filter [20] to reduce computation during the iterative process. The constant angular velocity vector $\vec{\omega}$ can be estimated as the average of all the relative rotations between successive frames, $R_0^1, \ldots, R_{n-1}^n$ [21]. The hopping trajectory is then extrapolated according to Eq. 1 and Eq. 5.

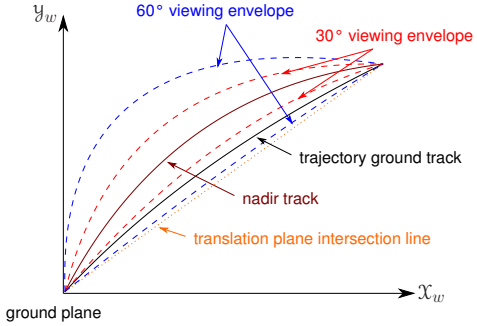
Next, at each time $t_k$ in the extrapolated trajectory, we determine if previously observed features can be re-observed by re-projecting them onto the image plane at $t_k$. However, instead of attempting to perform this re-projection test for every image, we limit it to portions of trajectory where the rover is oriented towards an envelope lying between the start and end of a hop, where features that can be re-observed across a hop are most likely to exist.

This is akin to combining the ideas behind the active vision approach [22], [23], where a camera is actively controlled to fixate on certain features, and the guided search approach, where some knowledge of the motion model is used to predict where features will be reprojected to improve feature matching. However, for a hopping rover, because the viewing angle is assumed to be not controllable, we act as a passive observer, and instead rely on the known motion model of a ballistic trajectory to predict when desirable observations can be made.

The envelope that is used to limit the guided search (Fig. 12) is the projection onto the ground plane of the viewing envelope of a camera pointing in the nadir direction, which is defined to be opposite to the surface normal $\hat{n}$ of the takeoff surface. This viewing envelope represents the ground coverage, or the observed portion of the ground plane, for a camera that is constantly oriented towards the nadir direction throughout a hop. The nadir track, or the projection along the nadir direction of the hopping trajectory onto the ground plane, dissects the viewing envelope. Because the takeoff surface is generally not parallel to the ground plane, the nadir track will be distinct from the ground track of the hopping trajectory.




(a)




(b)

Fig. 12. Viewing Envelope of a Nadir-Pointing Camera

To determine this envelope, the orientation of the nadir direction relative to ground plane is required. The nadir direction is simply $-\hat{n} = -\hat{k}_0$, the negative $\mathcal{Z}_0$-axis of the initial coordinate frame $\mathcal{F}_0$, with respect to which the

position estimates $\vec{x}(t_1), \ldots, \vec{x}(t_n)$ are defined. However, the orientation of the ground plane is generally unknown, because the ground plane and the takeoff surface are in general not parallel, and the gravity vector $\vec{g}$ is in general not aligned with the local vertical. However, because the ground plane was defined to be the average height of the terrain, after two or more hops, we can the fit a plane to all the landing positions to estimate this plane.

During the initial visual odometry step, whenever the hopping rover is oriented towards the envelope described above, features are detected and saved for matching against later images. For this guided searching, instead of reusing the Harris corners used for feature tracking in visual odometry, we make use of SURF features [24], which are more costly to compute, but produces unique signatures for more distinct features. This enables us to match features more reliably between two points far apart in a hopping trajectory, where the difference in viewpoint can produce significant perspective distortion. To establish feature correspondence, we compare the Euclidean distance between the 64-dimensional feature descriptor associated with each SURF feature. A ratio test between the two closest neighbors is used to determine matches [25].

Once features matches are established between a camera frame within the initial visual odometry step and a new camera frame in the extrapolated trajectory, together with the 3D coordinates of the features established in the camera frame within the visual odometry step, pose estimation can be performed, and the relative motion between the two camera frames can be estimated.

## V. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a two-step iterative algorithm to estimate the trajectory of a hopping rover. A monocular scheme of visual odometry is used to estimate an initial portion the hopping trajectory, from which the parameters for the ballistic motion are recovered, and the trajectory is extrapolated to predict the positions of the rover for the remainder of the hop. Then, a guided search of features lying in an envelope between the start and end of a hop is used to estimate the position of the rover in the extrapolated trajectory. The newly determined camera frame in the extrapolated trajectory can then be used to refine the parameters of the ballistic motion, and the trajectory can be re-extrapolated to predict future positions of the hopping rover.

Following this scheme, it is possible to estimate the trajectory of a hopping rover undergoing continuous rotational motion with only one set of cameras without continuous tracking of terrain features.

The constituent parts of this algorithm have been tested individually, and implementation of the complete system is in progress. It is expected to be validated using simulation and experimental results.

## REFERENCES


[1] K. Yoshida, "The Jumping Tortoise: A robot design for locomotion on micro-gravity surface," in *5th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, vol. 440, 1999, p. 699.

[2] Y. Nakamura, S. Shimoda, and S. Shoji, "Mobility of a microgravity rover using internal electro-magnetic levitation," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, 2000, pp. 1639–1645.

[3] S. Shimoda, T. Kubota, and I. Nakatani, "New mobility system based on elastic energy under microgravity," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, 2002, pp. 2296–2301.

[4] J. Burdick and P. Fiorini, "Minimalist jumping robots for celestial exploration," *The International Journal of Robotics Research*, vol. 22, no. 7-8, pp. 653–674, 2003.

[5] T. Yoshimitsu, "Autonomous navigation and observation on asteroid surface by hopping rover MINERVA," in *6th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, 2001.

[6] S. Higo, T. Yoshimitsu, and I. Nakatani, "Localization on small body surface by radio ranging," *Advances in the Astronautical Sciences*, vol. 124, pp. 727–738, 2006.

[7] P. Fiorini, C. Cosma, and M. Confente, "Localization and sensing for hopping robots," *Autonomous Robots*, vol. 18, no. 2, pp. 185–200, 2005.

[8] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the Mars Exploration Rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.

[9] A. Johnson, S. Goldberg, Y. Cheng, and L. Matthies, "Robust and efficient stereo feature tracking for visual odometry," in *Proceedings of IEEE International Conference on Robotics and Automation*, 2008, pp. 39–46.

[10] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics*, vol. 23, no. 1, pp. 3–20, 2006.

[11] S. Shimoda, A. Wingert, K. Takahashi, T. Kubota, and I. Nakatani, "Hopping direction controllability for small body exploration robot," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, 2004, pp. 2987–2992.

[12] B. Cyganek and J. P. Siebert, *An Introduction to 3D Computer Vision Techniques and Algorithms*. Wiley, 2009.

[13] T. Yoshimitsu, T. Kubota, I. Nakatani, T. Adachi, and H. Saito, "Micro-hopping robot for asteroid exploration," *Acta Astronautica*, vol. 52, no. 2-6, pp. 441–446, 2003.

[14] D. Scheeres, S. Broschart, S. Ostro, and L. Benner, "The dynamical environment about asteroid 25143 Itokawa, target of the Hayabusa mission," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2004.

[15] C. Harris and M. Stephens, "A combined corner and edge detector," *Alvey Vision Conference*, vol. 15, p. 50, 1988.

[16] R. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, pp. 146–157, 1997.

[17] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *International Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.

[18] B. M. Haralick, C. Lee, K. Ottenberg, and M. Nlle, "Review and analysis of solutions of the three point perspective pose estimation problem," *International Journal of Computer Vision*, vol. 13, no. 3, pp. 331–356, 1994.

[19] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[20] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, 1st ed. Wiley-Interscience, 2001.

[21] C. Gramkow, "On averaging rotations," *International Journal of Computer Vision*, vol. 42, no. 1, pp. 7–16, 2001.

[22] A. Blake and A. Yuille, *Active Vision*. The MIT Press, 1992.

[23] A. J. Davison and D. W. Murray, "Mobile robot localisation using active vision," in *Computer Vision ECCV98*, 1998, p. 809.

[24] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[25] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.