Onboard Detection and Correction of Orbital Registration Errors Using Rover Imagery

G. Foil, C. Cunningham, D. S. Wettergreen, W. L. Whittaker

Robotics Institute, Carnegie Mellon University, USA
e-mail: gfoil@cs.cmu.edu, ccunningham@cmu.edu, dsw@ri.cmu.edu, red@cmu.edu

Abstract

Orbital imagery is a key component in planning and executing rover missions. It can be used to predict hazardous terrain, target science objectives, or plan traversal paths, yet the utility of these products is crucially tied to their registration accuracy relative to a rover’s location. Misregistrations of orbital data can lead to problems ranging from unsafe navigation to incorrect or misleading science findings. As rovers become more autonomous it is important that techniques are developed to perform onboard reasoning about terrain information in real-time.

This paper outlines a multiresolution approach for detecting and correcting for orbital misregistration errors. Using a path generated from a series of navigation camera images, a mapping is made to high-resolution orbital imagery. This high-resolution imagery is then linked to lower-resolution multispectral imagery, and a direct transformation from the rover’s ground location to high- and low-resolution imagery is recovered.

1 Introduction

As rover autonomy increases, the amount of terrain covered in a mission as well as the data products needed to support a mission will increase. It is inevitable that some of these data products will have inaccuracies, misregistrations, or will simply be at a lower resolution than the rover’s sensors. It is imperative, therefore, that methods are developed that detect errors and misregistrations and account for them without hindering or stopping rover operations.

Errors in registration can have detrimental effects on a mission outcome. Inaccuracies in local or global localization can lead to suboptimal or even dangerous planning. In other situations, such as the scenario described by Wettergreen et al. [1], in which the rover queries orbital information for adaptive science purposes, incorrect registration can lead to inaccurate, biased, or faulty science findings.

In the scenario addressed here, we are concerned with correctly locating a rover’s position within orbital data, but not necessarily correctly locating the rover in world coordinates. In a perfect system with no misregistration and accurate localization, locating a rover in orbital data is a simple coordinate look-up. In a misregistered or inaccurate system this becomes more difficult, as a direct look-up provides inaccurate and erroneous results.

We describe an approach that uses a multi-resolution approach to localizing a rover relative to orbital data. It begins by combining a series of rover navigation camera (navcam) measurements as the rover moves, collecting a path of dense, sub-centimeter resolution color data projected onto the $x - y$ plane. It then uses this data to calculate the transform between rover-based data and high-resolution orbital data such as GeoEye images. Then, in order to correctly register to lower-resolution information, such as ASTER images, a transformation is found between the high-resolution orbital imagery and the low-resolution imagery, allowing the rover to self-localize within both orbital images.

This method is also applicable to both the Moon and Mars where orbiting satellites contain similar instruments to GeoEye and ASTER. For Mars, the HiRISE camera provides information at a similar resolution to GeoEye. Various spectrometers, including CRISM and THEMIS, return lower resolution imagery of Mars in longer-wavelength spectral bands. Similarly, for the Moon, LROC produces high-resolution visual imagery, and DIVINER captures images in nine different visible and infrared spectral channels.

We evaluate on a number of traverses taken from a rover in the Atacama Desert in Chile. We show that the position of the rover can be recovered relative to an orbital frame even when the resolutions at each scale differ by large amounts or the initial position of the rover is not known precisely.

The paper is organized as follows. Section 2 describes background and related work in rover-based registration and localization. Section 3 describes the details of our approach. Sections 4 and 5 discuss our experimental setup and our final results.

2 Background and Related Work

Rover localization relative to overhead imagery has been a topic of interest for many years. The Sojourner Mars rover localized using wheel odometry and a gyroscope, but accumulating errors limited the distance the
rover could travel between communications with Earth. Once each sol, operators would give an absolute update to position by analyzing rover stereo imagery to correct for accumulated errors [2]. To reduce the amount of human intervention needed, Olson and Matthies proposed an approach that used imagery from Sojurner to improve localization by comparing local stereo point clouds to previously generated point clouds to determine relative position [3].

The Mars Exploration Rovers improved localization by combining several different methods. Initially, the landers were localized using a combination of two-way Doppler radiopositioning and triangulation using landmarks visible in both orbital and ground based images. Then once the rovers left the landers, an inertial measurement unit (IMU) and wheel odometry were used to provide initial estimates of traverses. These were improved upon using visual odometry to correct for slip that occurred due to loose terrain especially in and around craters. Finally, bundle-adjustment was used to build an image network containing all panoramas and traverse images leading to highly accurate localization of rover positions [4]. The bundle adjustment step was later improved using an automatic cross-site tie point selection method that extracted, modeled, and matched rocks across images [5].

There have been many approaches to automatic registration of orbital data products, yet most of them focus on alignment of orbital and aerial products of similar scales [6, 7]. A handful of approaches focus on rover localization within orbital data. Hwangbo finds networks or constellations of rocks that are visible in Mars Exploration Rover (MER) navcam images, then searches within HiRISE imagery for similar networks of visible rocks [8]. Similarly, Di et al. extract rocks and rock image features from MER navcam images, then search for similar image features within HiRISE images [9]. However, both of these rely on the fact that rocks or other visual features are visible in both navcam and orbital images, an assumption which is sometimes valid in orbital imagery of as high resolution as HiRISE, yet falls apart with most Earth-based orbital imagery.

Sheshadri et al. project rover LIDAR and camera panoramas into the frame of high-resolution orbital imagery. Precise global rover localization is then determined using normalized cross-correlation. This was only demonstrated using simulated lunar imagery matched to LRO data [10]. Li et al. use a similar method for determining the position of an astronaut with respect to orbital imagery. A local digital terrain model is taken using stereo cameras mounted on a suit. Then this model is compared to a digital terrain model determined from orbital imagery to determine an absolute initial positions. Relative position updates are used to track astronauts after this initial calibration [11].

Carle and Barfoot developed MOGA, a similar method to globally localize a robot by matching features detected from orbital elevation models and rover LIDAR scans [12]. Unfortunately, this requires long range LIDAR for 3D mapping, which is not always feasible. In similar work, Cozman, et al. developed the VIPER algorithm, which provided a global estimate of position by matching horizon. The horizon is extracted from a 360° visual panorama and compared to the estimated horizon from digital elevation models of the surrounding terrain. Skylines were chosen because they are the most prominent features in a panorama and the lunar environment is generally devoid of texture. In practice, this algorithm had localization errors on the order of hundreds of meters but performed very quickly [13]. Barfoot et al. tested both VIPER and MOGA on Devon Island in the Canadian High Arctic getting average errors of 36 and 30 meters for VIPER and MOGA, respectively [14].

Unlike many of these approaches, our proposed method works completely in image space. Instead of using 3D features of the terrain, such as rocks or LIDAR scans, our approach keys off changes in ground cover to more naturally match to an orbital image. We furthermore allow for the registration of data sources that span resolutions orders of magnitude different from each other, all by using terrain context observed along an extended traverse.

3 Approach

Directly matching from rover imagery to lower-resolution orbital imagery is difficult, at best. Navcam images often contain high-frequency information not present in orbital imagery, such as individual rocks or small-scale terrain features like tire tracks. Furthermore, even ignoring the wavelength differences that may be present between navcam and orbital images, rover-based imagery suffers from self-shadowing, lens glare, and other lighting artifacts depending on the time of day.

Because of these factors we instead opt to use an intermediate registration step, first registering to high-resolution orbital imagery much more similar to the resolution and wavelengths observed by the rover. This high-resolution imagery is then registered to a low-resolution, multispectral image, and the transformation from the rover’s ground position to the correct location in the multispectral image can be recovered.

3.1 Navcam Path Generation

A single image captured by the rover has insufficient information to reliably match to an orbital image. Instead, we propose creating a navcam path, or a series of navcam images integrated together and placed into a global refer-
ence frame using onboard rover odometry and an initial location estimate.

To create the path, the rover first projects navcam observations into a colorized point cloud in a local coordinate frame using stereopsis. As the rover moves, odometry estimates from onboard sensors are used to track the relative motion between frames, so successive navcam images can be integrated together. Because each image contains information at centimeter or sub-centimeter resolution, the path is voxellized and stored as a lower resolution point cloud as the rover stops observing that location. Given an initial rover location and orientation, this path can be projected into a global coordinate frame and treated as an overhead image of the traverse.

Using a path such as this both reduces noise and allows for greater correlation with terrain statistics in the orbital image. As most terrain is observed over multiple frames, irregularities such as glare or rover shadowing are averaged across multiple images and have less of an impact on the resulting path.

Note that we make two assumptions here. The first is that the odometry error in the rover is small relative to the length of the path. This means that our approach assumes a series of navcam images can be integrated strictly using odometry estimates. In reality, this is not always true, and a more robust solution might use Iterative Closest Point (ICP) [15] or a SLAM-based graph relaxation technique to reduce odometry error between a sequence of navcam observations. Odometry errors in the path can also be minimized by limiting the path length itself, which is the approach we have taken in this work.

The second assumption we make is that the initial location and orientation of the rover is known fairly accurately. This is a somewhat ambiguous statement, and we explore the robustness of our system to initial location accuracy in Section 4.

### 3.2 Registration to High Resolution Imagery

Once a navcam path has been generated and projected into global coordinates, the visual information accumulated across the path can be used to align and refine the location of the path in the frame of a high-resolution orbital image. Given a starting position and orientation, we would ideally like to find the transformation $M$ that minimizes the objective function:

$$
M^* = \arg\min_M \sum_i (Mp_i - q_i)^2
$$

where $p_i$ and $q_i$ are corresponding points in the path and orbital images, respectively, and $M$ is a rigid transformation composed of a yaw rotation, $R_{yaw}$, and translation, $T = [t_x, t_y]^T$:

$$
M = \begin{bmatrix}
R_{yaw} & T \\
0 & 1
\end{bmatrix}
$$

There are a number of approaches to solve for this transformation, and most work either in image space or in feature space. In image space common statistics between the two images are used to align them, either with an exhaustive search over some set of possible transformations and rotations, or with a transformation of the space, such as log-polar coordinates, that allows for matches to be made more easily. In feature space unique features of each image are calculated and features similar to both images are used to calculate the above transformation.

In this scenario the size and resolution of the navcam path makes feature detection implausible. The width of the path is approximately three to five meters, so in an overhead one meter per pixel image the number of features that can be extracted from the path is very limited.
We focus on statistics in image space. Our solution uses the weighted sum of two approaches: a normalized cross-correlation score computed over a set of transformations close to the given initial location and an ICP-based approach in which the path and the orbital image are treated as point clouds to be aligned. We have found that individually each method struggles with overcoming local optima, yet both highly score the true transformation. By normalizing and combining both scores the overall solution is more robust and less prone to finding sub-optimal transformations.

In more detail, our ICP-based approach models the orbital data and path as 3D point clouds in which their $x$ and $y$ coordinates are their global locations and their $z$ coordinate is the greyscale color intensity at that point. Solving ICP on this effectively finds the transformation of the navcam path that minimizes the difference in color. Our solution computes the error in ICP alignment, $S_{ICP}$, after convergence for each of a given number of candidate transformations $M$ about the initial location estimate:

$$S_{ICP} = \sum_i (M p_i - q_i)^2$$

We then compute the normalized cross-correlation score between the navcam path under a candidate transformation and an overlapping region of the orbital image. Formally, for a navcam path $f$ at a given location $(u,v)$ in the orbital image $g$, the cross correlation score $S_{Corr}$ is computed as

$$S_{Corr} = \frac{\sum_{x,y} (f(x,y) - \bar{f})(g(u + x, v + y) - \bar{g})}{\sqrt{\sum_{x,y} (f(x,y) - \bar{f})^2 \sum_{x,y} (g(u + x, v + y) - \bar{g})^2}}$$

where $\bar{f}$ and $\bar{g}$ are the means of the path and orbital regions, respectively.

For each candidate initialization location we then find the weighted sum between the ICP error $S_I$ (normalized over all evaluated transitions and inverted so larger scores have lower error) and the normalized cross-correlation score $S_C$ (also normalized over all evaluated transitions):

$$S = \alpha S_{ICP} + \beta S_{Corr}$$

The best candidate location is thus the transformation that maximizes this score, effectively having the least color alignment error and the highest cross-correlation. We have found values of 0.5 and 0.8 for $\alpha$ and $\beta$, respectively, provide good results. An example of the scoring process is shown on the following page in Figure 3.

### 3.3 Registration to Low Resolution Imagery

In some applications this initial mapping is all that is needed. However, in others the resolution of orbital data products can differ enough from the navcam image path that a direct mapping is unlikely to be found. In this case our method uses a series of orbital data products of decreasing resolution to perform the mapping. We first align the navcam image path to a high resolution data product, such as 1 meter per pixel GeoEye imagery. Once we are localized within the higher resolution imagery we down sample the high-resolution orbital data and find the transformation that maps the higher resolution data to the lower resolution data.

This mapping is calculated by finding a homogeneous transformation between the two image frames by matching image keypoints. First, both images are converted to greyscale and their intensities are normalized. This helps improve the success of keypoint matching by compensating for the fact that the image channels in each orbital data product do not necessarily correspond to the same wavelength and are thus not necessarily comparable. Next, SURF keypoints [16] are extracted from both images, although other feature descriptors (e.g. SIFT [17]) could prove equally successful. Keypoints are then matched between the images based on the similarities in their feature descriptors. The best matches are used to find a homography between the two images. Finally, the homography is used to transform the low-resolution orbital data to match the perspective of the high-resolution orbital data. Example keypoint matches are shown in Figure 2.

### 4 Experiments

We test our approach on data collected during multiple rover traverses in the Atacama Desert in Chile during June, 2013. The rover used for the experiments, Zoë, is equipped with stereo navigation cameras, onboard odometry, and a GPS sensor for ground truth information. In tests Zoë’s odometry system has shown 2-4% translational odometry error and approximately 5% error in yaw over an extended traverse. Images are captured at approximately one frame per second, and the rover has a maximum speed of one meter per second.
4.1 High Resolution Registration Tests

Our experiments focus on characterizing the robustness of the proposed solution to errors in the rover's initial pose estimate. We begin by generating a navcam path over our traverse and voxelize it to lower the resolution to approximately 0.25 meters per pixel. We further apply a blurring to the color values of the path by averaging over the set of points within a radius of 0.5 meters of a candidate voxel. This has the effect of removing some of the high frequency information, which improved matches between ground and orbital imagery.

Satellite imagery used are one meter per pixel GeoEye images. The images are preprocessed to increase the resolution from 1.0 to 0.25 meters per pixel. Color values in this preprocessing are interpolated as an average of nearby points weighted by their distance to the pixel.

When creating the \((x, y, \text{color})\) point clouds of both the navcam path and orbital imagery the color information is scaled down to be more in line with the scale of the terrain. In tests we have found that an unscaled color or color not scaled enough tends to dominate the ICP alignment, leading to poor results with very little ICP adjustment. Our best performance has been with a scaling factor 0.01 - 0.03. For example, this essentially equates a gradient from pure white to pure black (0 to 255 in pixel values) to a 2.5 to 7.5 meter tall feature.

Once point clouds are created, a dense, uniform sampling of transformations is evaluated, mimicking an initial pose error of a corresponding amount. For each of these transformations ICP is run and the resulting pose adjustment is applied to the navcam path. The normalized cross-correlation score of this path with the corresponding locations in the orbital data is then obtained, and a final score for each transformation is calculated.

In order to test the robustness of our approach to inaccuracies in the estimation of the initial location, we simulate a number of initialization values for both position and orientation. In each test we evaluate initialization points successively farther from the true location of the scan, then track how often our approach recovers the correct location. For each evaluation point we consider it correctly localized if it’s estimated localization adjustment places it within 2 meters of the correct location and within 15 degrees of the correct orientation.

Test data is taken from three sections of a traverse in the Atacama Desert. Traverse 1 is approximately 150 meters, covering an approach up to and around a small hill, and is shown in Figure 1. Traverse 2 is a challenging 50 meter straight-line traverse in which the rover transitions from lighter-colored clays to darker desert pavement. Traverse 3 is a 100 meter loop in which the rover moves from light-colored clays to darker desert pavement, then back to lighter clays. All images were collected under favorable lighting conditions, although there are periods of the traverse with discolorations due to high glare.

4.2 Low Resolution Registration Tests

The focus of our experiments in matching high- and low-resolution orbital images is in characterizing our approach’s robustness to wavelength differences. We test by registering one meter per pixel GeoEye images to 15 meter per pixel ASTER image. The GeoEye images used in testing contain three bands in the 450-510, 510-580, and 655-690 nanometer ranges. For comparison, the ASTER images used in testing contain nine bands in the ranges of 520-600, 630-690, 760-860, 1600-1700, 2145-2185, 2185-2225, 2235-2285, 2295-2365, and 2360-2430.
nanometers. Even in the simplest case, in which we compare all three bands of the GeoEye image to the first three bands of the ASTER image, there is not complete overlap across all bands.

To test the effect the degree of wavelength overlap has on keypoint matching and image transformation, an experiment was run to match ASTER orbital imagery to high-resolution GeoEye imagery. In the ASTER orbital image, all possible permutations of the 9 wavelengths were assigned to red, green, and blue channels to create 504 different images. Several points were selected from the ASTER images for testing the resulting transformations. Their ground-truth correspondences in the GeoEye image were determined in order to evaluate the error of the resulting transformation. Each of the 504 ASTER image permutations were matched to the GeoEye image in both normalized and unnormalized forms in order to evaluate the effect of normalization.

5 Results

We analyze two main aspects of our approach: the robustness of the approach to inaccuracies in initial location when matching to a high resolution image, and the robustness of our approach to differences in wavelengths when matching orbital imagery.

5.1 Registration to High Resolution Imagery

We find that the approach handles recovery from position and orientation initialization inaccuracies very well, shown in Figure 5. For two of the three test traverses, Traverses 1 and 3, it is able to localize correctly when within a distance of 7 meters from the true location, and gradually falls off as the rover initializes farther from the true position.

For the third traverse, Traverse 2, the path is a straight line, as shown in Figure 4. It suffers from the fact that the approach matches well to a large amount of terrain in the local area. In this traverse the straight-line transition from light clays to darker desert pavement is not unique to the true initial location, so true location cannot be easily recovered.

Overall we find that, as one would expect, longer traverses with more scene context result in better localization. As shown in Figure 3, the combined scores for Traverse 1 result in a peak centered nearly exactly at the true location of the rover. For the other traverses tested, the resulting score function was not as distinct, and in the case of Traverse 3 there were a large number of areas with high scores, leading to poor performance as the initialization became less accurate. While there is some computation cost incurred when working with larger navcam paths, they appear to offer better performance when compared to shorter, less informative paths.
5.2 Registration to Low Resolution Imagery

As expected, the normalized ASTER images were matched more consistently than the unnormalized images. Out of the 504 permutations of wavelengths, only 21 found a sufficient number of keypoint matches to compute a homography when unnormalized. In contrast, 236 of the 504 permutations were able to find a transformation when normalized. The average error of a point in the ASTER image when compared to the ground truth was 80 pixels. However, the median was only 8 pixels, which shows that a majority of the transformations were very accurate, with a few outliers with very poor results. These results confirmed that the wavelengths assigned to image channels have a strong influence on the quality of the keypoint matches between images and the resulting image transformation.

The number of matching keypoints is a strong function of the wavelengths that are used in the normalized ASTER images. Figure 6 shows the percentage of matches when different wavelengths were used in the ASTER image. There was a much higher chance of finding a match when one of the two visual wavelengths of the near infrared wavelength was used. The probability of finding a match increases as two or all three of those wavelengths are used. It was much harder to find correspondences with the longer six wavelengths. In fact, only 1% of the permutations matched that did not involve any of the visual or near infrared wavelengths.

6 Conclusion and Future Work

We have demonstrated a technique for detecting and accounting for misregistration errors in orbital imagery. By combining a series of rover-based measurements into a single data product, a more robust match can be made to high-resolution imagery. This high-resolution imagery can then be used to find an appropriate translation into lower-resolution imagery, allowing for accurate retrieval of data from orbital data products, even when misregistration is present.

While we have framed the problem as the calculation of a rigid transformation between data products, in reality that is not always true. Due to constantly changing errors in rover odometry, the true rover path as viewed from overhead will differ from the navcam paths we have discussed here. For future work we would like to explore this more general case, potentially treating the path as a graph of segments of smaller paths.

We would also like to investigate the inclusion of additional scoring functions in our proposed approach for mapping navcam paths to high-resolution orbital imagery. Specifically, we look for alignment metrics that can be calculated on the images themselves and used to characterize the fitness of a candidate transformation. An example would be the Chamfer Matching score, in which edge contours are extracted from an image’s color space and the total distance between edges in a path and orbital image is calculated. While we have tested this approach as a registration technique and it performed poorly on its own, it likely could augment our proposed scoring function and improve performance in some ambiguous regions.

7 Acknowledgments

This work was supported by NASA Office of the Chief Technologists Space Technology Research Fellowships, as well as ASTEP grant NNX11AJ87G.
References


