LONG-RANGE ROVER LOCALIZATION BY MATCHING LIDAR SCANS TO ORBITAL ELEVATION MAPS

by

Patrick J.F. Carle

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Institute for Aerospace Studies
University of Toronto

Abstract

Long-Range Rover Localization by Matching Lidar Scans to Orbital Elevation Maps

Patrick J.F. Carle

Master of Applied Science

Institute for Aerospace Studies

University of Toronto

2009

Current planetary rover localization techniques are lacking in autonomy and accuracy. An autonomous method of globally localizing a rover is proposed by matching features extracted from a 3D orbital elevation map and rover-based 3D lidar scans. Localization can be further improved by including odometry measurements as well as orientation measurements from an inclinometer and sun sensor. The methodology was tested with real data from a Mars-Moon analogue site on Devon Island, Nunavut. By tying 23 real scans together with simulated odometry over a 10km traverse, the algorithm was able to localize with varying degrees of accuracy. Output uncertainties were large due to large input uncertainties, but these could be reduced in future experimentation by minimizing the use of simulated input data. It was concluded that the architecture could be used to accurately and autonomously localize a rover over long-range traverses.

Acknowledgements

I owe greatest thanks to my supervisor Tim Barfoot for his generous guidance and razor-sharp honesty. Thanks to Joseph Bakambu for helpful discussions on feature matching techniques.

I would also like to thank my family for their unwavering support, especially my parents Colleen and François, and my brother Alex. Finally, thanks to my Kingston, UTIAS and other friends for lifting my head out of the books every now and then for the odd adventure.

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Objectives 3

2 Literature Review **4**

2.1 Sensors 4

2.2 Positioning with Satellites 5

2.3 Dead-Reckoning Techniques 6

2.4 Global Map-Matching Techniques 7

2.4.1 Overview 8

2.4.2 Past Work 9

2.5 Literature Review Summary 11

3 Methodology **12**

3.1 Overview 12

3.2 Feature Detection 13

3.3 Feature Matching 18

3.3.1 Overview 18

3.3.2 Hypothesis Search 19

3.3.3 Hypothesis Evaluation 21

3.3.4 Feature Matching Example 27

3.4 Pose Refinement . . . 29
  3.4.1 Multi-frame Odometry-compensated Global Alignment (MOGA) . . . 30
  3.4.2 Local Feature Terms . . . 32
  3.4.3 Global Feature Terms . . . 35
  3.4.4 Odometry Terms . . . 35
  3.4.5 Measured Orientation Terms . . . 38
  3.4.6 Combining Terms and Optimization . . . 39
  3.4.7 RANSAC . . . 41
  3.4.8 MOGA Example . . . 43
3.5 Methodology Summary . . . 43

4 Experimental Setup **45**
4.1 Global and Local Maps . . . 45
4.2 Field Traverses . . . 47
4.3 Ground-Truth . . . 48
4.4 Odometry Measurements . . . 49
4.5 Orientation Measurements . . . 50

5 Results **51**
5.1 Overview . . . 51
5.2 Feature Detection . . . 53
5.3 Feature Matching . . . 54
5.4 Pose Refinement . . . 56
5.5 Single-Frame Localization No Heading . . . 58
  5.5.1 Dataset A, Single No Heading . . . 58
  5.5.2 Dataset B, Single No Heading . . . 62
5.6 Single-Frame Localization With Heading . . . 65
  5.6.1 Dataset A, Single With Heading . . . 65

5.6.2 Dataset B, Single With Heading 69
5.7 Multiple-Frame Localization With Heading 72
5.8 Results Summary 76

6 Conclusion 82

7 Future Work 83

A DEM Properties 90
A.1 DEM resolution 90
A.1.1 Latitude, Longitude 90
A.1.2 Altitude 91
A.2 DEM Positional Accuracy 91
A.2.1 Horizontal 91
A.2.2 Vertical 91

B Search Sphere Shell Thickness Derivation 92

C Axis-Angle Conversion 96

Chapter 1

Introduction

1.1 Motivation

With the signing of the Global Exploration Strategy (CSA et al., 2007), fourteen of the world's leading space agencies made a long-term commitment to the exploration of the solar system. The motivation for this strengthened international cooperation ranged from understanding how to establish a sustainable human presence on the Moon, to inspiring the next generation of scientists and engineers here on Earth. A wealth of newly-gathered scientific data would refine age-old questions on the origins of life and the Earth's history; questions whose answers are scattered, like pieces of a puzzle, in the geologies and atmospheres of other celestial bodies.

Since Yuri Gagarin's historic space voyage, humans have played an important role in the act of exploring space. However, there is an immense cost and risk associated with preserving fragile human life in such extreme environments. For this reason, future space missions will be increasingly supported and led by robots. Indeed, robotic achievements in space have existed for some time, as with the Sputnik satellite of 1957 (Dickson, 2001) or the Lunokhod rovers of the early 1970's (Chaikin, 2004).

A recent milestone in space robotics was the Pathfinder mission of 1997, which delivered the Sojourner rover to Mars (Matijevic, 1998). The success of the planet's first mobile visitor proved that robots could become a feasible and cost-effective alternative to manned planetary exploration. Since then, and with the help of enabling technologies, there has been a boom in space robotics: the Mars Exploration Rovers (Bresina et al., 2005), the Hayabusa asteroid sample-return spacecraft (Yano et al., 2006), the Phoenix Mars lander (Desai et al., 2008), and the planned Mars Science Laboratory (MSL) (Crisp et al., 2008) and ExoMars rover (Vago et al., 2008) are just some examples. While great leaps in rover technology have been made since the Lunokhods first roamed the Moon, these advances pale in comparison to the challenges ahead. One such challenge studied in this thesis deals with the ability for a rover to determine its position and orientation on the surface of a celestial body. This process is known as *localization*.

Present-day rovers are only able to make short-range traverses due to limitations in their mechanical abilities and autonomy. For example, the amount of solar energy available to the Mars Exploration Rovers (MER) restricts the number of operations they can perform on a given day (Leger et al., 2005; Biesiadecki et al., 2005). Their lack of autonomy is also very costly considering the twenty-minute communication delay between Mars and Earth (Norris et al., 2005). As a result, the MERs have journeyed only about 10km in four years (Li, Arvidson, et al., 2007). Even the next generation MSL is planned to drive a mere 6-10km in two years (Palluconi et al., 2003; Volpe, 2003).

It is therefore unsurprising that localization techniques developed to this day are unsuitable for long-range missions on the order of hundreds of kilometres and up. However, as technology improves and demand on a rover's performance increases, rovers will be expected to traverse much greater distances over their lifetimes. For this reason, they will require an accurate and autonomous long-range localization system. The absence of such a system could unknowingly send a rover off its planned path, never to reach its intended destination and unable to accurately label the locations of its scientific finds. This thesis proposes a solution to the problem of long-range localization.

1.2 Objectives

The objectives of this research are:

1. To develop an algorithm to autonomously localize a rover over long ranges with consistent accuracy by matching 3D maps of terrain from ground-based lidar and orbiting satellites.
2. To validate the methodology with a realistic dataset by quantitatively demonstrating the algorithm's ability to consistently localize.

The sections to follow describe and evaluate a novel, long-range localization algorithm developed to satisfy these objectives.

Chapter 2

Literature Review

Past work on rover localization has evolved from manual image comparison to today's more automated image processing techniques. This section gives an overview of the various tools available to solve the problem and how these have been used in past applications.

2.1 Sensors

Two basic sensors found on nearly all mobile robot systems are the wheel odometer and the inclinometer. Wheel odometers measure changes in position and heading based on wheel movement. Inclinometers measure the direction of the gravity vector, which allows for pitch and roll angles to be determined. Sun, Moon or star sensors can measure heading by comparing measurements of a celestial object's actual position in the sky with its expected position based on orbital models (Furgale et al., 2009). These can also be used to determine position to a relatively high degree of uncertainty (Kuroda et al., 2004).

Vision sensors such as stereo cameras and lidar play an equally important role in localization. Stereo cameras produce pairs of images, from which depth can be extracted by detecting and matching features between images. Errors in wheel odometry can then be corrected by tracking these image features over time to estimate changes in pose. Another popular vision sensor is time-of-flight lidar (*Li*ght *D*etection *a*nd *R*anging). This instrument rapidly fires a laser and measures the time for reflections to return, which allows the range of distant objects to be determined very accurately.

Stereo cameras and lidar were evaluated by Se et al. (2004). An advantage of lidar is that it does not require much pre-processing of data, which is not true when extracting depth from a stereo pair of images. Lidar can also work in zero-lighting conditions, operate over a long range of 1-2km, and produce A very accurate, sub-metre representation of the terrain. Its most significant disadvantages are high power draw and a bulky size.

2.2 Positioning with Satellites

The satellite-based Global Positioning System (GPS) has been used to great success on Earth and would be the ideal solution to the localization problem. However, a minimum of four satellites would be required and even this would only provide partial, time-dependent coverage. The cost of constructing this network of satellites around a new planet or moon would at present be unreasonable.

A seemingly more cost-effective approach would be to use ground-based GPS as proposed by LeMaster & Rock (2003). These pseudo-satellites, or pseudolites, could be placed at strategic surface locations and function alone or in cooperation with orbital GPS satellites. Tests showed better than 5cm position error, but were accomplished at ranges of only 30 to 50m. It was claimed that kilometre ranges could be achieved with more powerful pseudolites.

Localization over long distances (i.e. hundreds of kilometres) would require a large array of pseudolites covering the desired exploration path. Furthermore, line of sight would be required between rover and pseudolite. Therefore, the rover would either be restricted to flat land or the pseudolites would need to be deployed with greater difficulty on hills and mountain peaks. This approach could not realistically satisfy long-range localization requirements, and would be better suited for applications in site-planning or rover operation at a permanent base.

Radio positioning was used with the MERs, which were equipped with a radio system to communicate with Earth and orbiting satellites. By comparing a direct-to-Earth signal with a signal relayed through an orbiting satellite (Guinn, 2001), the position of the rovers could be estimated with an accuracy of about 50m (Li et al., 2004). However, to achieve this degree of accuracy, estimates required multiple satellite passes and line of sight with the Earth. Therefore, if this technique were used as the primary localization system, the operating time of the rover would be severely limited.

2.3 Dead-Reckoning Techniques

Dead-reckoning is the determination of pose based on previous pose measurements. The simplest but most error-prone form of rover localization is wheel odometry, a type of dead-reckoning. Wheel odometry is extremely vulnerable to sensor noise and mechanical disturbances such as wheel slippage and azimuthal angle drift. For example, during the MER mission, the Opportunity rover's wheel odometry once underestimated a 19m, three day traverse by over 8% (Li et al., 2005).

Visual motion compensation techniques such as visual odometry (VO) and bundle adjustment (BA) can be used to correct these large errors. VO (Olson et al., 2003) tracks features across pairs of images to measure changes in pose. Features are typically extracted using image processing techniques such as Scale Invariant Feature Transforms (SIFT) (Lowe, 1999; Barfoot, 2005) or Speeded Up Robust Features (SURF) (Bay et al., 2006). Other techniques also exist such as modeling rocks as ellipsoids (Li, Di, & Howard, 2007) and matching them between images. VO is fully automated and operates in real time, although the additional processing required can slow the traverse considerably. BA (Li et al., 2002) is the batch alignment of features between images taken from different perspectives to again determine relative motion. Unlike VO, this method requires a high degree of human intervention but is also more accurate (Li et al., 2006). Because VO and BA require a great deal of processing time, the MERs have reserved these techniques for scenarios with a high risk of wheel slippage such as steep slopes, or when a more accurate rover position was needed in general.

A significant drawback to these dead-reckoning techniques is the unbounded growth of error they produce. Each estimated pose will have some error associated with it. Since each estimate depends on the previous, error may continue to grow without bound. Olson et al. (2003) reduced this error growth to linear rates by including absolute measurements of orientation in the VO algorithm. In field experiments, Konolige et al. (2007) showed a similar VO algorithm to yield less than 0.1% error over a 9km traverse in rough-terrain. Despite such encouraging results, dependence on computationally-taxing visual correction techniques can significantly slow traverses. Furthermore, these dead-reckoning techniques still suffer from unbounded error growth.

2.4 Global Map-Matching Techniques

To determine pose with no dependence on distance traversed, globally consistent imagery, the *global map*, can be matched with local rover imagery, the *local map*. Here, imagery implies either a 2D image or a 3D Digital Elevation Model (DEM). The techniques most relevant to this research generally follow the procedure of feature detection, description and matching. In this process, features are first extracted from the global and local maps. They are then matched to find global-local correspondences. Once correspondences are known, the features can be aligned. In doing so, the pose of the rover with respect to the global frame can be determined.

A number of reviews of matching techniques were consulted to sort through the numerous algorithms developed for these purposes. Mian et al. (2005) and Planitz et al. (2005) evaluated the ability of algorithms to automatically find correspondences between 3D point models. Planitz et al. (2005) also developed a general framework for the common steps involved in these algorithms. Mikolajczyk & Schmid (2005) examined the performance of various feature descriptors in 2D image applications. These reviews served as a top-level reference in determining which feature matching algorithm, or combination of algorithms would be most suitable in overcoming the challenges faced in this thesis. The results are presented in the subsections to follow.

2.4.1 Overview

A feature-based approach is expected to reduce the time needed to match maps when compared to global surface alignment techniques such as Iterative Closest Point (ICP) (Z. Zhang, 1994), which consider the entire surface in the alignment and do not guarantee to converge to a global optimum. A feature-based approach could also be more flexible and allow for the use of higher resolution 2D images from orbit. However, the use of 2D images would present the extra challenge of matching image-based features to features extracted from 3D lidar point clouds.

For this reason, it may be simpler to use 3D matching techniques such as spin-images (Johnson, 1997), harmonic shape images (D. Zhang & Hebert, 1999), point fingerprints (Sun et al., 2003), or local shape descriptors (Taati et al., 2007). These approaches compare feature descriptors between local and global maps to search for matches. For example, spin-image descriptors are 2D histograms representing the distances of nearby points from the tangent plane and normal vector at a chosen interest point. Upon comparison of two features' descriptors, a correspondence is made between the two features if their descriptors are sufficiently similar.

However, this procedure is more difficult to implement when the area around the feature is only partially available as is the case with the rover-based local map, which is scattered with occlusions. Furthermore, these occlusions will always be found near peaks, the most prominent features in the terrain. It therefore becomes difficult to match features that are fully describable from the global map's viewpoint, but only partially from the local map's. An example is shown in Figure 2.1.

A different approach that avoids this problem is to consider features' positions with respect to each other. These groups of features, or *feature constellations*, effectively act as the descriptor. The algorithm used to seek out correspondences in this fashion is the Data-Aligned Rigidity-Constrained Exhaustive Search (DARCES) (Chen et al., 1999). The greatest advan-

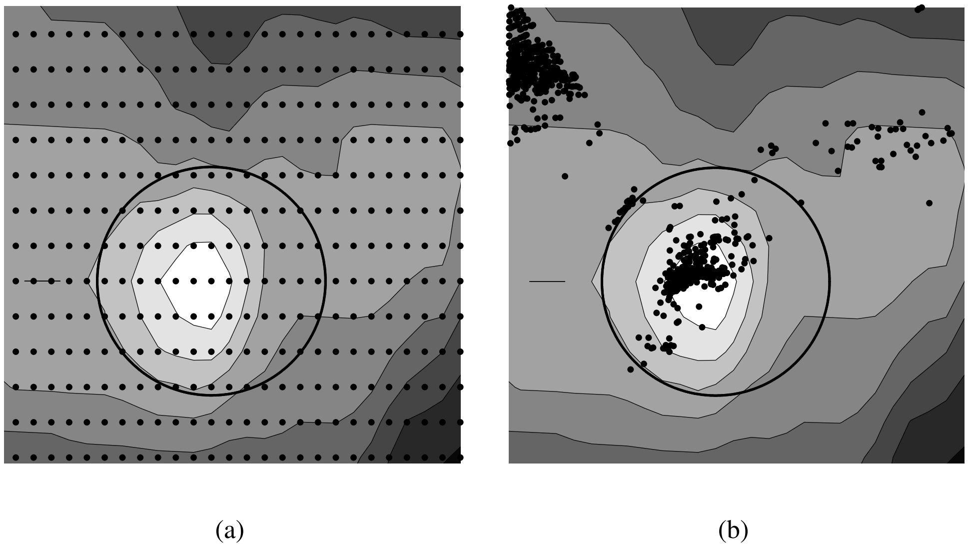

(a) (b)

Figure 2.1: Contour maps of field data showing the difficulty in producing similar descriptors with occluded maps. Elevation increases with lighter shading. Circles indicate the descriptor interest region. (a) Global map with a complete set of evenly spaced data points. (b) Partially occluded, unevenly spaced lidar map. Lidar located off the map in the upper left direction.

tage of DARCES is that it does not require descriptors to be generated, which are not easily obtained from the rover-based, occlusion-ridden local map.

2.4.2 Past Work

Some early Earth-based attempts at 3D feature extraction and matching compared peaks between rover-based laser range data and a high quality global map (Hayashi & Dean, 1989). However, simulations yielded poor results because of problems in accommodating the difference in resolutions between scene and model. Another approach gathered random range measurements and matched them to a DEM, but the method required that the orientation and altitude of the rover be known from onboard equipment (Yacoob & Davis, 1992).

Bakambu et al. (2006) examined the problem of matching two lidar scans. Different feature matchers were evaluated and combined to improve computing efficiency. Interest points on the global and local maps were selected according to curvature change (Bae & Lichti, 2004). Correspondences between images were then found using the spin-image or point fingerprint feature descriptors. Finally, DARCES selected the best correspondences and computed the rigid transformation between the two images. This enhanced algorithm was successfully tested over model resolutions of 0.2, 0.5 and 1.0m. The poorer the model resolution, the more processing time was required. Instances of the computation failing to properly align images were attributed to an insufficient number of interest-points being selected.

Descent imagery provides a series of images at different resolutions during lander descent (Li et al., 2000). The main challenge is to accurately transform a string of descent images to the same orientation. The technique, while achieving a localization accuracy of 1m, is limited to an area within a couple kilometers of the landing zone and requires manual selection of tie points in the descent images.

The problem of matching global aerial maps to local rover maps has been investigated in several papers. A common solution is to compare the skyline from a rover-based image to a topographic reference map (Stein & Medioni, 1995), (Cozman & Krotkov, 1997). However, this generally requires a mountainous horizon and results in poor accuracy on the order of a hundred metres. High resolution imagery from aerobots (Behar et al., 2005) could offer an alternative to coarse orbital imagery, albeit a more problem-prone and expensive one that would clearly not be applicable in environments such as the Moon, where there is no atmosphere.

Vandapel et al. (2006) localized a lidar-equipped rover in heavily vegetated environments with the spin-image technique. To test the method, a helicopter was flown at a height of 400m to create a global map using a lidar with a range resolution of 1cm and positional accuracy of 10 to 30cm. Vandapel outlines the following key points to improve performance in matching. Proper selection of the height and width of the spin-image interest surfaces is crucial. Dead-reckoning should be used as a constraint during localization. Interpolation can be used to accommodate differences in resolution. Interest points should be efficiently selected, for example by filtering out those lying in uninteresting, flat areas.

Attempts to match orbital images to rover-based images are not yet fully autonomous. Li, Di, Hwangbo, et al. (2007) processed stereo imagery from the Mars Reconnaissance Orbiter's HiRISE camera into 3D elevation maps. This global map was then compared to the MER's stereo camera imagery, and was shown to match closely. In future work, they expect to integrate HiRISE imagery directly into the established bundle adjustment process.

Orbital maps vary in quality and coverage. There has been skepticism (Li et al., 2000) in using orbital maps since their resolutions can be poor. However, new high resolution satellite imaging systems entering into operation such as HiRISE are helping to alleviate these concerns (Li, Di, Hwangbo, et al., 2007). Table 2.1 summarizes the most current information for various orbital data sources.

Table 2.1: Summary of coverage and resolutions for various satellite datasets. Sources: For LOLA and LROC see Chin et al. (2007). For MOLA2 see Smith et al. (2001). For HiRISE see Johnston et al. (2005); McEwen et al. (2002). For Devon Island see http://www.geobase.ca. For Typical Earth see Grodecki & Dial (2001).

Target	**Instrument**	**Description**	**Coverage**	**Horiz. Res.**	**Vert. Res.**
Moon	LOLA	Laser altimeter	Total	50-100m	10cm
	LROC	Camera	>10%	50cm	NA
Mars	MOLA2	Laser altimeter	Total	100m	1m
	HiRISE	Camera	2%	<1m	<1m [1]
Typical Earth	IKONOS	Camera	On-demand	1-4m	<10m [1]
Devon Island	NA	NA	Partial	12-23m	1m

2.5 Literature Review Summary

The literature review has shown that autonomous, long-range rover localization of a planetary rover is still an open and relevant problem. This thesis proposes a solution that matches a low resolution, orbital elevation map to high resolution, rover-based lidar maps. The greatest benefit of this approach is the referencing of all local measurements back to a globally consistent orbital map. This will limit the growth of localization error during a rover's traverse.

---

[1] Vertical accuracy of elevation model created from stereo pair of images.

Chapter 3

Methodology

3.1 Overview

The goal is to determine the rover's poses with respect to the global map's frame, $\underrightarrow{\mathcal{F}}_o$. The rover pose at the ℓ^{th} scan site is defined as a transformation, $\mathbf{T}_{\ell o} := \{\mathbf{t}_{\ell o}, \mathbf{S}_{\ell o}\}$, from $\underrightarrow{\mathcal{F}}_o$ to the rover's local frame $\underrightarrow{\mathcal{F}}_\ell$, where $\mathbf{t}_{\ell o}$ and $\mathbf{S}_{\ell o}$ are respectively the translation and rotation from $\underrightarrow{\mathcal{F}}_o$ to $\underrightarrow{\mathcal{F}}_\ell$. An overview of the procedure is presented in Figure 3.1 and summarized as follows:

(a) **Feature Detection**: Features are detected from the global and local maps.

(b) **Feature Matching**: Correspondences are found between global and local features with DARCES (Chen et al., 1999). Another useful output of DARCES is a preliminary estimate of the rover poses. Optimally, local orientation measurements act as filters to increase the chances of finding a good match and improve the efficiency of the DARCES algorithm.

(c) **Pose Refinement**: Initialized with DARCES pose estimates and knowing correspondences between global and local features, the RANSAC-MOGA algorithm refines the rover's poses. Odometry and orientation measurements can be introduced into RANSAC-MOGA to further refine the alignment. Odometry can also allow poses to be estimated in frames for which DARCES has no solution (i.e. when too few feature correspondences are found between global and local maps).

The sections to follow discuss each architecture component in further detail.

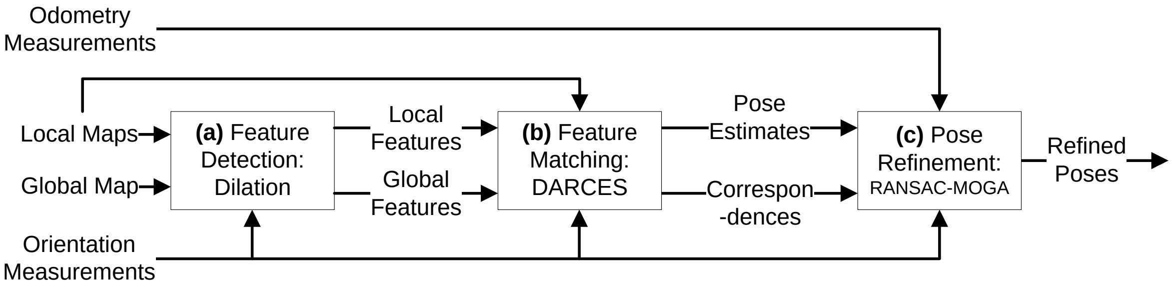


Figure 3.1: Overall architecture.

3.2 Feature Detection

This section discusses feature detection, which is represented as box (a) in the overall architecture diagram of Figure 3.1.

The two fundamental inputs to this step are (i) a low resolution, global DEM taken from an orbiting satellite, and (ii) one or more high resolution, local maps from rover-based lidar scans. The global map is defined as having a resolution of L_{global}. The challenge is to maximize the number of common features that are detectable in *both* maps. A given feature might only be detectable in one map due to the deficiencies in the other map. For example, an occlusion in the local map might hide a feature detected in the global map. Conversely, some small terrain features in the high resolution local map would not be distinguishable in the low resolution global map.

Generally, the most prominent features common to both maps are elevated peaks. These peaks are typically visible in the local map, so long as they are unoccluded and within range. Peaks are detected using a local maxima detector based on morphological dilation[1], a technique

---

[1] Code can be found on Matlab Central repository as 'localMaximum.m' by Yonathan Nativ.

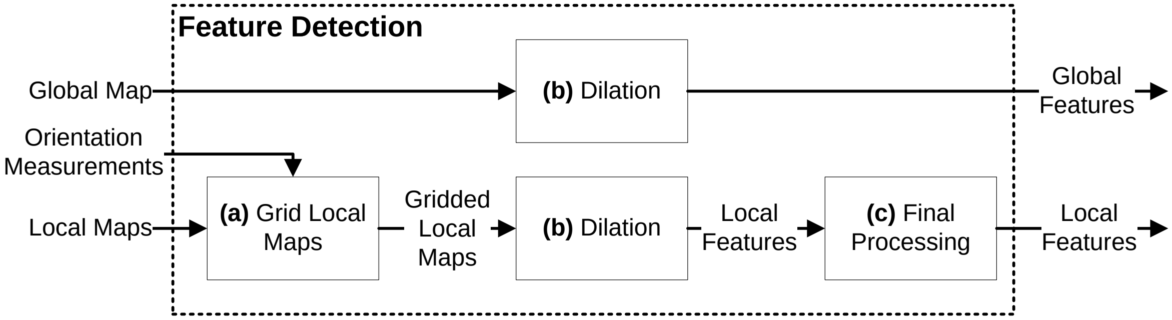


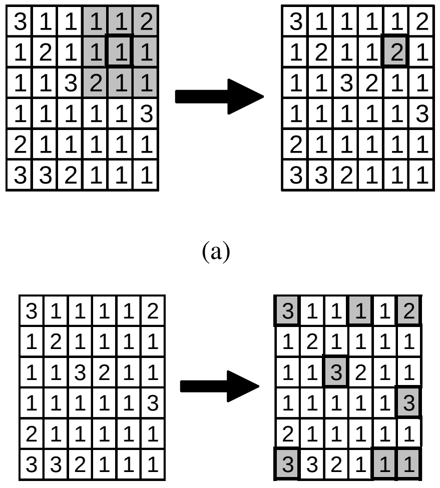
Figure 3.2: Feature detection architecture.

adopted from image processing (van den Boomgaard, R. and van Balen, R., 1992; Haralick & Shapiro, 1992).

Starting from box (a) in Figure 3.2, the lidar scan must be gridded. The local maps are first rotated using pitch and roll measurements from an inclinometer or equivalent device. This ensures the $+z$-direction of the global and local maps coincide. The local map grid is sub-sampled to the global map resolution, L_{global}. Sub-sampling the local map ensures the scale of detected global and local features are the same. Nearest-neighbour interpolation was used since it tended to reduce the number of poor, false peaks detected. Features may now be detected using morphological dilation.

Morphological dilation (see box (b) in Figure 3.2) replaces lower grid values with neighbouring higher grid values. Applied to an elevation map, this will effectively blur out low elevations. The extent of the blurring depends on the size and shape of the window used. Figure 3.3(a) gives an example of a single dilation operation for a square window. The operation replaces the centre point's value with the highest value within the bounds of the window. Once dilation is completed for all points on the grid, the blurred map is compared to the original map. Cells with no change in value are interpreted as local maxima. Figure 3.3(b) gives an example for a small map.

The dilation window is chosen to be a pixelated circle as shown in Figure 3.4. This is done to make the coverage of the window as even as possible in all directions. The radius of this circle limits the size of the detected features, as well as the distance between features. Therefore,



(a)

(b)

Figure 3.3: Morphological dilation used to detect peaks. Numbers represent elevation values. (a) Dilation example at a single cell for a 3x3 square window. (b) Result of local maxima detection using dilation with a 3x3 square window. Cells marked in gray are interpreted as local maxima because their values did not change after dilation.

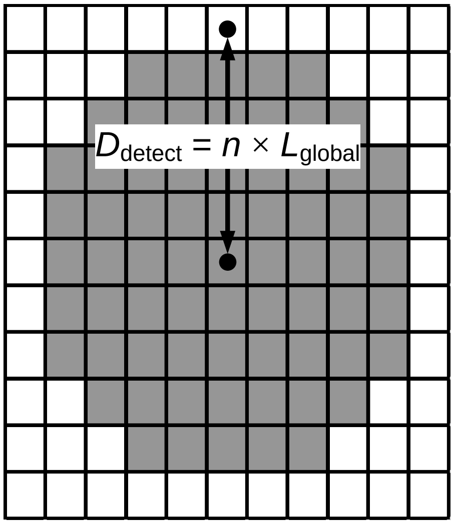


Figure 3.4: Pixelated circle with $n = 5$ cell radius in gray and minimum feature distance, D_{detect}, shown. This is the dilation window used in the feature detection process.

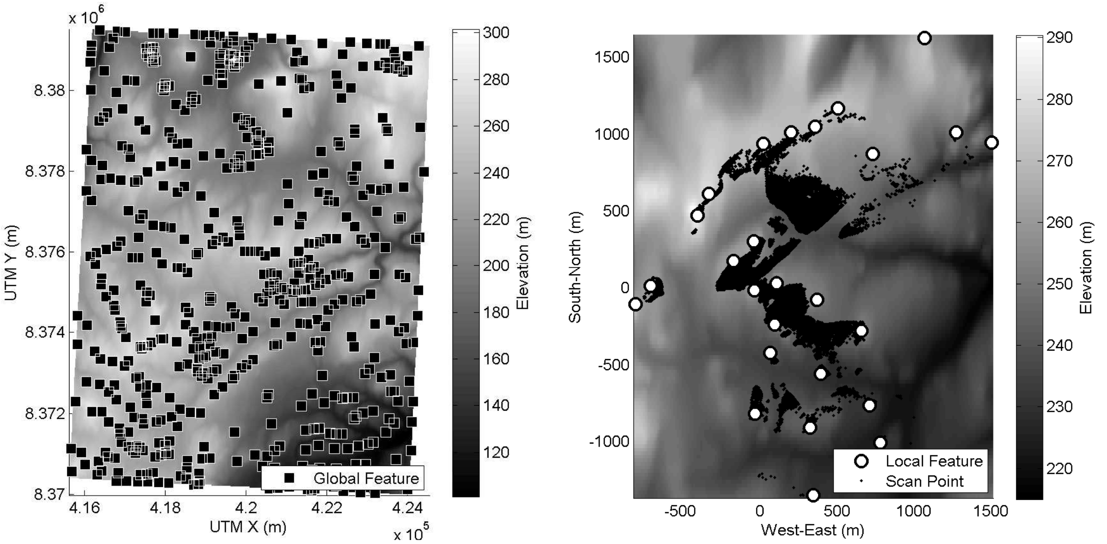
the minimum distance between features, D_{detect}, depends on the global map resolution:

$$D_{\text{detect}} := n \times L_{\text{global}} \tag{3.1}$$

where n is the circle's cell radius. The same window is applied to global and local maps.

Detected local features must now undergo some final processing (see box (c) in Figure 3.2). To ensure each detected local feature corresponds to a measured point in the lidar scan, the nearest measured point is selected as the feature's position. If for a given feature there are no measured points within a distance of the global map resolution, then the point is rejected. A final restriction for both global and local features is that no two features can be within a distance D_{detect} of each other. If two features are within this distance, the least elevated is dropped since the most elevated is more likely to be a peak.

The process of feature detection has been described. Features can now be extracted from the global and local maps. Examples from sample data are presented in Figures 3.5 and 3.6.



(a) Detected global features.

(b) Scan points for A08 and detected local features.

Figure 3.5: Feature detection example. UTM zone is 16X.

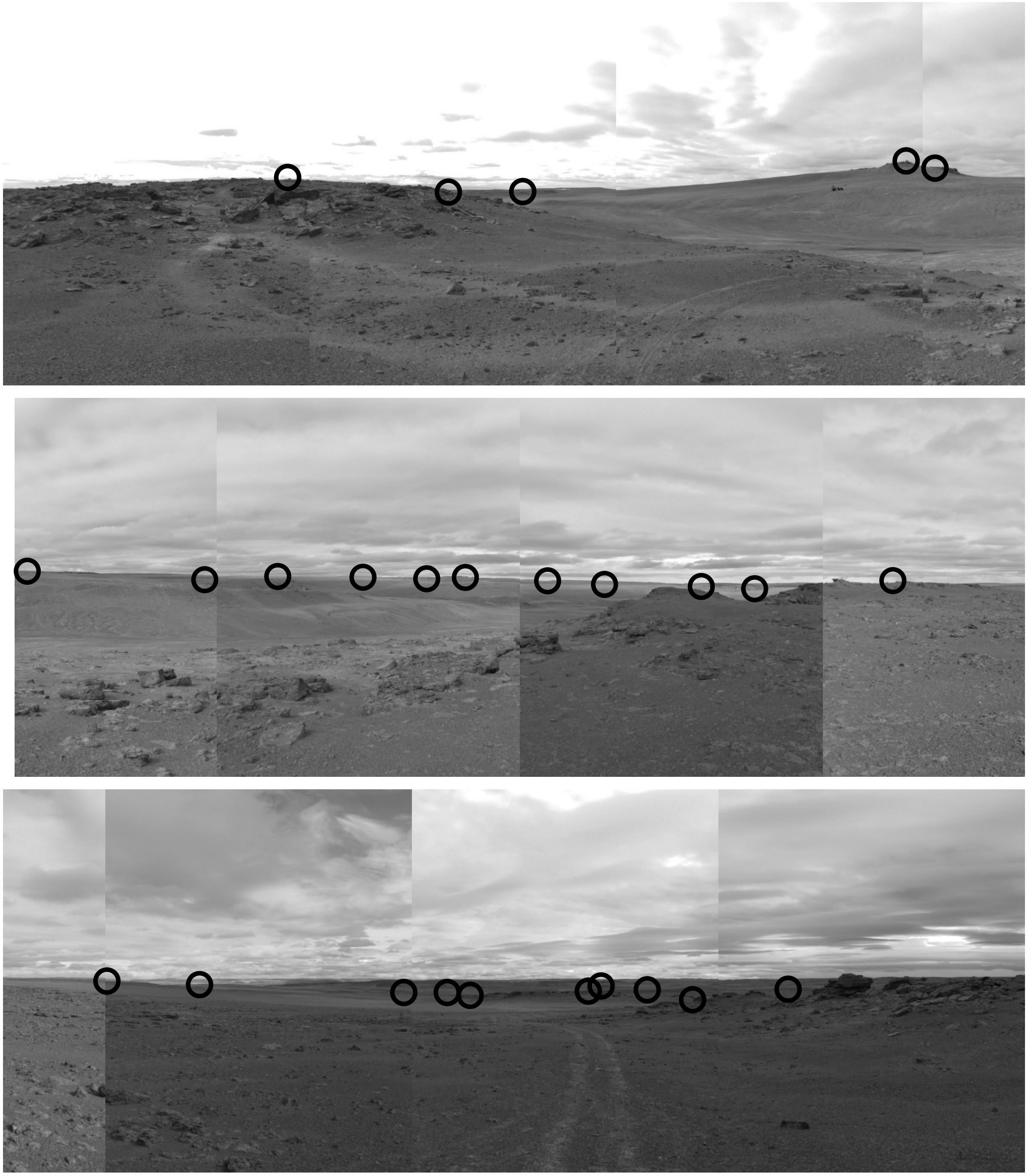


Figure 3.6: Local features overlayed onto panoramic images taken from scan A08. Note the detected hill peaks are not always on the skyline, since a peak can occur anywhere in 3D data.

3.3 Feature Matching

This section discusses feature matching, which is represented as box (b) in the overall architecture diagram of Figure 3.1.

3.3.1 Overview

Global features have been detected from the orbital DEM and local features from rover-based lidar scans, as explained in Section 3.2. The goal is to find a matching global feature for every local feature. These correspondences can then be used to calculate transformation estimates between local and global frames. It should be made clear that this entire block in the algorithm runs on local features for a single lidar scan at a time. If there is more than one scan, the block is looped for each scan. The outputs are then consolidated and passed to the next block of the algorithm.

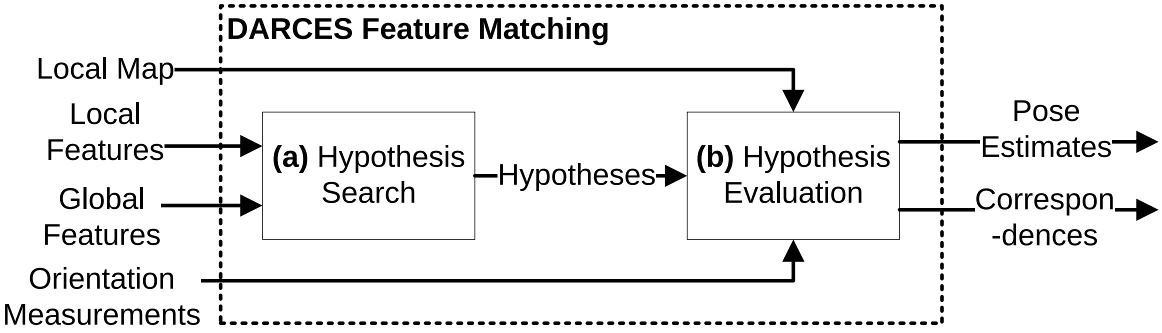


Figure 3.7: Feature matching architecture. DARCES processes scans individually.

The feature-matching methodology is based on the Data-Aligned Rigidity-Constrained Exhaustive Search (DARCES) algorithm (Chen et al., 1999) and is depicted in Figure 3.7. DARCES searches for similarities in features' relative positions. This can be thought of as a search for similar constellations of features between the global and local map. DARCES is chosen over a descriptor-based method because reliable descriptors are difficult to extract given the numerous and large occlusions in the lidar map [2]. Furthermore, occlusions are much more likely to occur in the vicinity of hill-peak features, where one side of the peak is usually hidden from the lidar as previously shown in Figure 2.1.

3.3.2 Hypothesis Search

This section relates to box (a) in Figure 3.7. Given a single scan's local features and all global features as inputs, the first step in DARCES is to generate an initial set of hypotheses. A hypothesis is defined as a group of possible correspondences between three unique local features, called control points, and three unique global features. These three correspondences are needed to produce a 3D transformation estimate between the global and local frames. This transformation is then used to evaluate hypotheses and select the best one.

It will now be explained how to conduct searches for correspondences. Sets of three control points are randomly chosen from a single scan's set of local features. Depending on the number of local features detected in a given frame, a maximum of N_{DARCES} sets of control points are tested because of the possibility that a local feature will have no correspondence. When searching for hypotheses with a single control point set, the three control points are arbitrarily labeled as primary, secondary and auxiliary. The primary control point is initially assumed as having a correspondence with every global feature.

In the noise-free case, for a primary control point correspondence to be valid, the following three constraints must be satisfied:

1. There is a global feature a distance d_{ps} from the global feature assumed to correspond with the primary control point. This would give a potential correspondence with the secondary control point.
2. There is a global feature, different from the last, a distance d_{pa} from the global feature assumed to correspond with the primary control point. This would give a potential cor-

---

[2] In fact, a descriptor-based method was initially attempted but yielded poor results.

respondence with the auxiliary control point.

3. There is a distance d_{sa} between the global features respectively assumed to correspond with the secondary and auxiliary control points.

This amounts to sphere searches as shown in Figure 3.8.

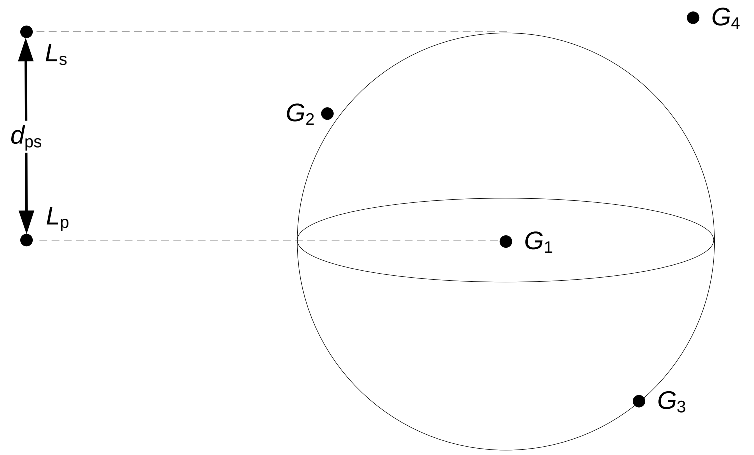


Figure 3.8: DARCES hypothesis search example for a secondary-point correspondence without noise. The primary and secondary control points are respectively L_{p} and L_{s}. The primary-point correspondence under examination is for global feature G_1. The sphere then has a centre that coincides with G_1 and a radius of d_{ps}. Only global feature G_3 qualifies as a secondary-point correspondence since it lies on the surface of the sphere.

In the noise-corrupted case, a search is conducted in the same way, but corresponding points would lie within the shell of a sphere of thickness $2t$, as shown in Figure 3.9. The shell's half-thickness, t, will depend on uncertainties in the positions of global and local features. Its derivation is provided in Appendix B. The result is

$$t \approx 3\sqrt{2}\left(\left(\sigma_{r_{Gxy}}\right)^2 + \left(\sigma_{r_{Lxy}}\right)^2\right)^{1/2} \tag{3.2}$$

where $\sigma_{r_{Gxy}}$ and $\sigma_{r_{Lxy}}$ are respectively one standard deviation of the uncertainties of the x and y position components for global and local features. Note the independence from position measurements. This is due to the simplifying assumption that the distance between two features will be much greater than the features' difference in elevation.

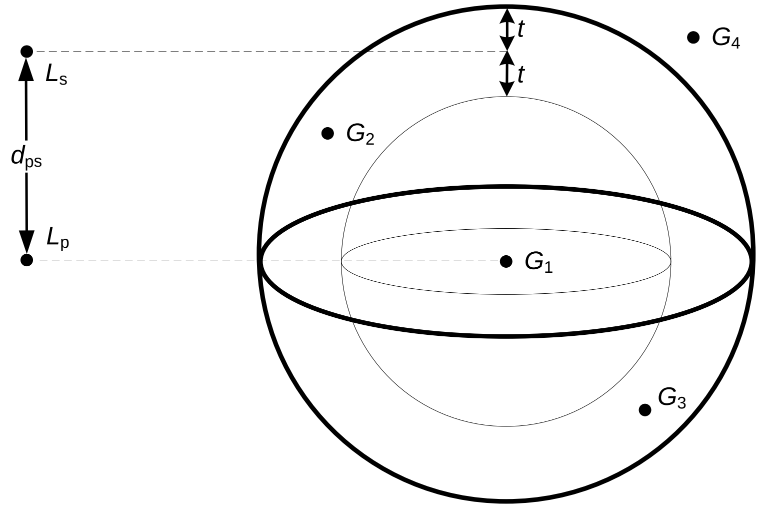


Figure 3.9: DARCES hypothesis search example for a secondary-point correspondence with noise. The primary and secondary control points are respectively L_p and L_s. The primary-point correspondence under examination is for global feature G_1. The sphere then has a centre that coincides with G_1, a radius of d_ps and a shell thickness $2t$. Global features G_2 and G_3 qualify as a secondary-point correspondence since they lie within the search sphere's shell.

Since a hypothesis contains correspondences between three global and three local features, a transformation from the global to the local frame can be calculated. For hypothesis i, the transformation from $\underrightarrow{\mathcal{F}}_o$ to $\underrightarrow{\mathcal{F}}_\ell$, $\mathbf{T}^i_{\ell o} := \left\{\mathbf{t}^i_{\ell o}, \mathbf{S}^i_{\ell o}\right\}$, is obtained using a least-squares point alignment algorithm that minimizes the distance between the three corresponding global and local features (Arun et al., 1987). These hypothesized transformations serve as an initial estimate of the rover pose, and are used to evaluate hypotheses in future stages of the algorithm.

It has been shown how to create an initial set of hypotheses, and how to compute a transformation for each hypothesis. The next step is to evaluate the hypotheses and determine which, if any, are valid.

3.3.3 Hypothesis Evaluation

The input hypotheses are now evaluated (box (b) from Figure 3.7) with the transformation estimate that each hypothesis provides. The best hypothesis can then be used to output the optimal correspondences between global and local features, and an initial estimate of the transformation from global to local frame. Figure 3.10 gives an overview.

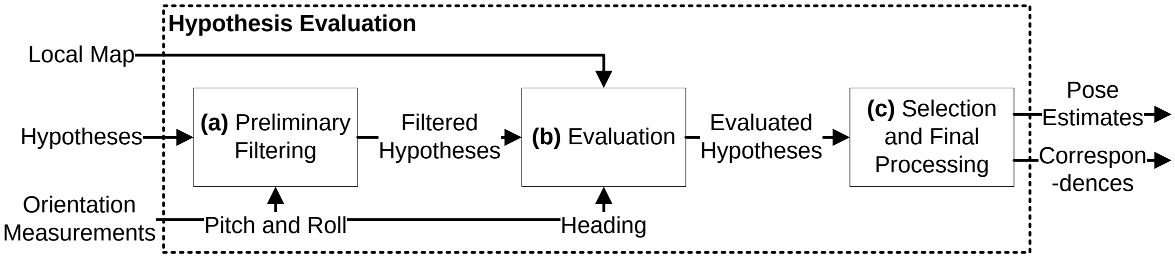


Figure 3.10: Overview of hypothesis evaluation and selection.

Preliminary Filtering

This section relates to box (a) from Figure 3.10. Hypotheses are passed through a preliminary filtering stage to improve the efficiency and robustness of DARCES. A hypotheses is evaluated by comparing its transformation estimate, as described in Section 3.3.2, to a transformation measurement. If an estimate does not agree with the measurement within uncertainties, then the hypothesis is considered invalid and is discarded.

Four filters are used in this step:

1. z-deviation: From the hypothesized x, y translations, t_x and t_y, a DEM z-position measurement is obtained, $\rho_z\left(t_x, t_y\right)$. This measurement is then compared to the hypothesized z-position, t_z. To pass this stage of the filter, the difference between measurement and hypothesis must remain below the threshold E_{zdev}:

$$|\rho_z - t_z| \leq E_{\mathrm{zdev}}. \tag{3.3}$$

2. Map boundary: After transforming all local reference points to the global frame using the hypothesized transformations, a condition is imposed that all reference points must remain within the bounds of the map. Failing to do so, the hypothesis is discarded.

3-4. Roll and pitch: The measured rotation matrix from the global to the local frame, **C**, is compared to the hypothesized rotation matrix from the global to the local frame, **S**, by decomposing both into their $z - y - x$ Euler angles:

$$\mathbf{C} := R_z(\gamma_{\text{meas}})R_y(\beta_{\text{meas}})R_x(\alpha_{\text{meas}}), \quad (3.4)$$

$$\mathbf{S} := R_z(\gamma_{\text{hyp}})R_y(\beta_{\text{hyp}})R_x(\alpha_{\text{hyp}}) \quad (3.5)$$

where α, β, γ are respectively roll, pitch and heading angles. The differences between measured and hypothesized roll and pitch angles are then obtained and verified against their respective thresholds, E_α, E_β. The filter conditions are

$$|\alpha_{\text{meas}} - \alpha_{\text{hyp}}| \leq E_\alpha, \quad (3.6)$$

$$|\beta_{\text{meas}} - \beta_{\text{hyp}}| \leq E_\beta. \quad (3.7)$$

$$(3.8)$$

A heading filter is applied in the “Selection and Final Processing” (box (c) from Figure 3.10) step for reasons discussed later.

Hypotheses have now passed through preliminary filtering.

Evaluation

This section relates to box (b) from Figure 3.10. The remaining filtered hypotheses must now be evaluated using their calculated transformations from the global frame to the local frame as discussed in Section 3.3.3. The local map also plays a key role in this evaluation, but must first be decimated by removing points in highly dense regions of the map. This will increase efficiency and spread scan points more evenly over the map. After decimation, the distance between any two points in the decimated scan should not exceed the global map resolution, L_{global}. These modified local map points are called the *reference points*.

To evaluate the quality of a hypothesis, these reference points are transformed to the global frame using the hypothesis transformation and compared to the global map, as shown in Figure 3.11. Hypotheses are then scored based on the sum of absolute z-errors between the transformed reference points and their interpolated points on the global map. The map boundary constraint, previously discussed in Section 3.3.3, ensures that this metric can be calculated from each hypothesis in a consistent manner.

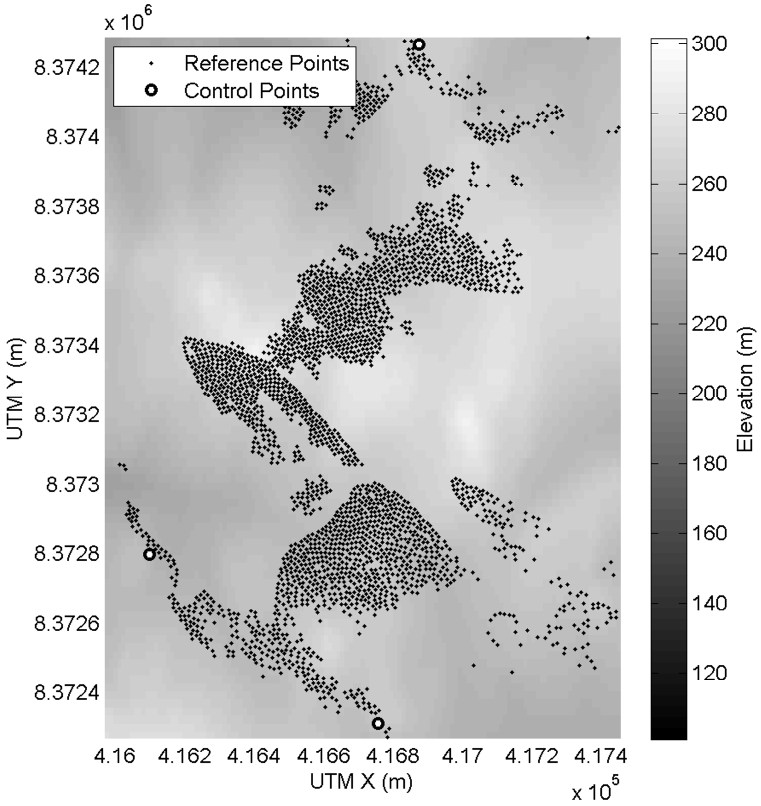


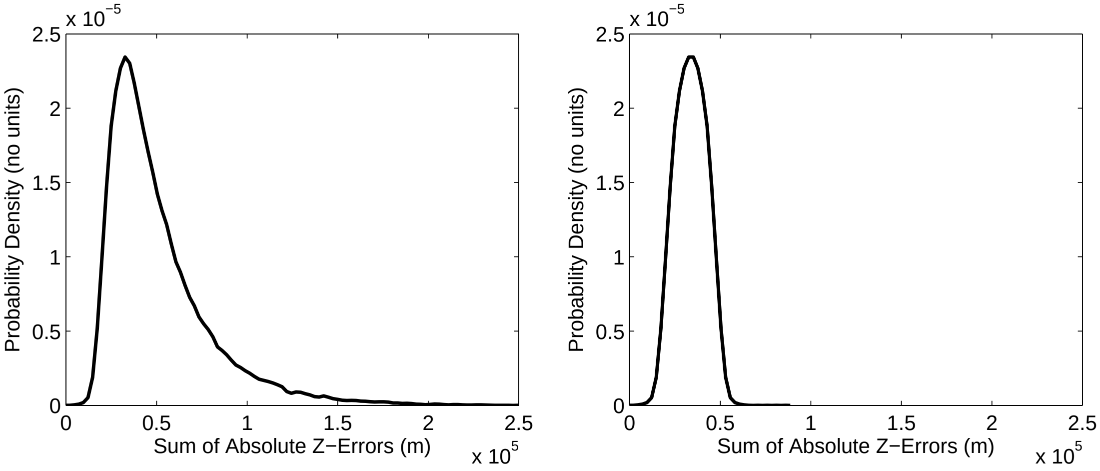
Figure 3.11: Scan A08 reference points transformed to the global frame using a sample hypothesis transformation. UTM zone is 16X.

To ensure a hypothesis is indeed valid, its score must be significantly better than other top-scoring hypotheses. In this circumstance, a good score is a low one. Therefore, hypotheses that are low-score outliers are considered valid. These low-error hypotheses are detected based on a data point's distance from the lower quartile boundary of the set of all hypothesis scores. This is similar to the technique used in spin-image matching (Johnson, 1997). A valid hypothesis is defined as one that scores lower than the threshold E_{valid}:

$$E_{\mathrm{valid}} := \mathrm{Q1}\left(S\right) - \frac{3}{2} \times \mathrm{IQR}\left(S\right) \tag{3.9}$$

where S is the entire filtered dataset, $\mathrm{Q1}$ the lower quartile, and IQR the interquartile range.

This only holds if the data is approximately normally distributed. Since the data tended to be asymmetric, all data scoring greater than the mode score is ignored. This is justified since only low-error scores are of interest. Therefore, only the distribution left of the mode, S_{left}, is used as shown in Figure 3.12. The mode is determined using a kernel density estimator (Botev, 2007) to estimate the probability density of S.



(a) (b)

Figure 3.12: (a) Distribution, S, of scores for scan A08 preliminary-filtered hypotheses. (b) Symmetric distribution, S_{sym} after ignoring hypotheses right of the mode.

The distribution to the left of the mode, S_{left}, is mirrored to the right in order to create a new symmetric distribution, S_{sym}. This has the effect of shifting the lower quartile boundary to the median of S_{left}, $\mathrm{Q1}\left(S_{\text{sym}}\right) = \mathrm{median}\left(S_{\text{left}}\right)$. Consequently, the interquartile range becomes $\mathrm{IQR}\left(S_{\text{sym}}\right) = 2 \times \left[\mathrm{mode}\left(S\right) - \mathrm{median}\left(S_{\text{left}}\right)\right]$. Modifying Equation (3.9), valid hypotheses are then required to score below

$$\begin{aligned} E_{\text{valid}} &= \mathrm{median}\left(S_{\text{left}}\right) - 3 \times \left[\mathrm{mode}\left(S\right) - \mathrm{median}\left(S_{\text{left}}\right)\right] && (3.10) \\ &= 7 \times \mathrm{median}\left(S_{\text{left}}\right) - 3 \times \mathrm{mode}\left(S\right). && (3.11) \end{aligned}$$

It is now possible to determine which of the filtered hypotheses have a sufficiently low error-score to be considered valid.

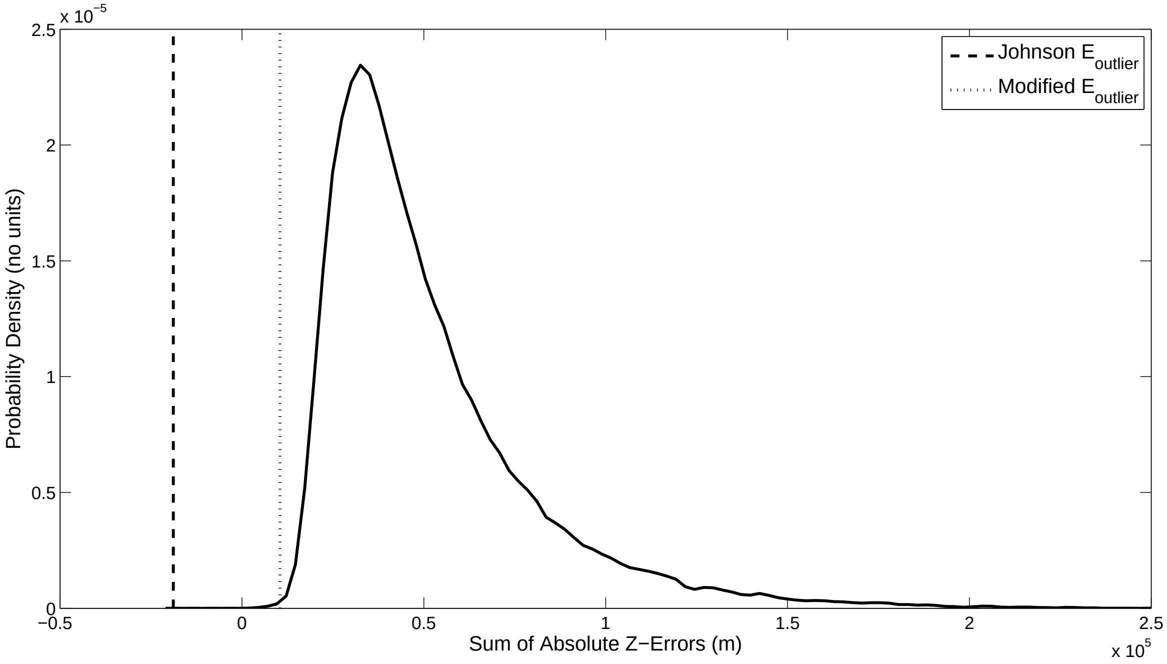


Figure 3.13: Valid hypothesis thresholds for the method by Johnson (1997) (dashed) and the modified method used in this thesis (dotted). These are overlayed on the distribution of scores for scan A08 filtered hypotheses.

Selection and Final Processing

This section relates to box (c) from Figure 3.10. Hypotheses have been filtered and evaluated as valid. Before the best hypothesis is selected, the remaining hypotheses are passed through a heading filter. The measured rotation matrix from the global to the local frame, $\mathbf{C}$, is compared to the hypothesized rotation matrix from the global to the local frame, $\mathbf{S}$, by decomposing both into their Euler angles as in Equations (3.4) and (3.5). The heading error can then be verified against its threshold, E_γ:

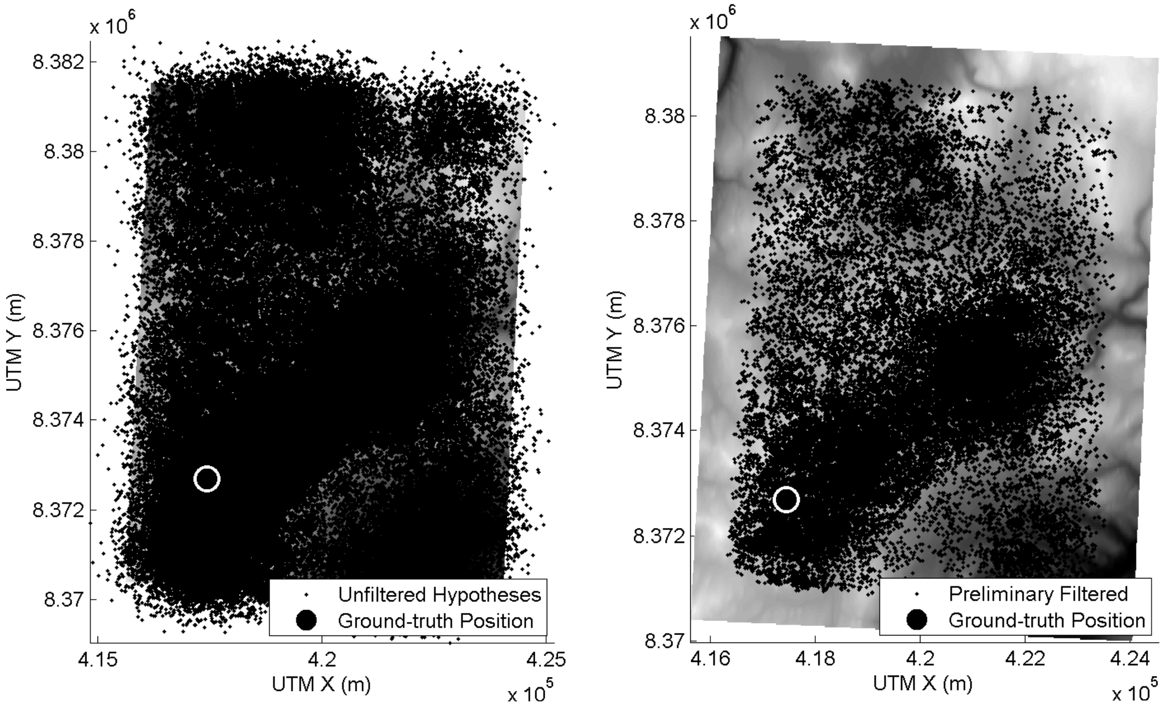
$$|\gamma_{\text{meas}} - \gamma_{\text{hyp}}| \leq E_\gamma. \tag{3.12}$$

A heading filter is used only after the evaluation process because the heading filter tends to eliminate so many hypotheses that the shape of the distribution of fitness scores is destroyed. This distribution is key to the selection of a valid hypothesis, as discussed in Section 3.3.3.

Once this final filtering stage is complete, and if there are any hypotheses remaining, the hypothesis with the lowest error-score is selected as the best. Some final steps are taken to produce the two desired outputs. The initial pose estimate output is the transformation associated with the selected hypothesis, which will be refined in the next section. A full set of correspondences for all local features is obtained by first using this same pose estimate to transform local features into the global frame. The correspondence for each local feature is then to the nearest global feature.

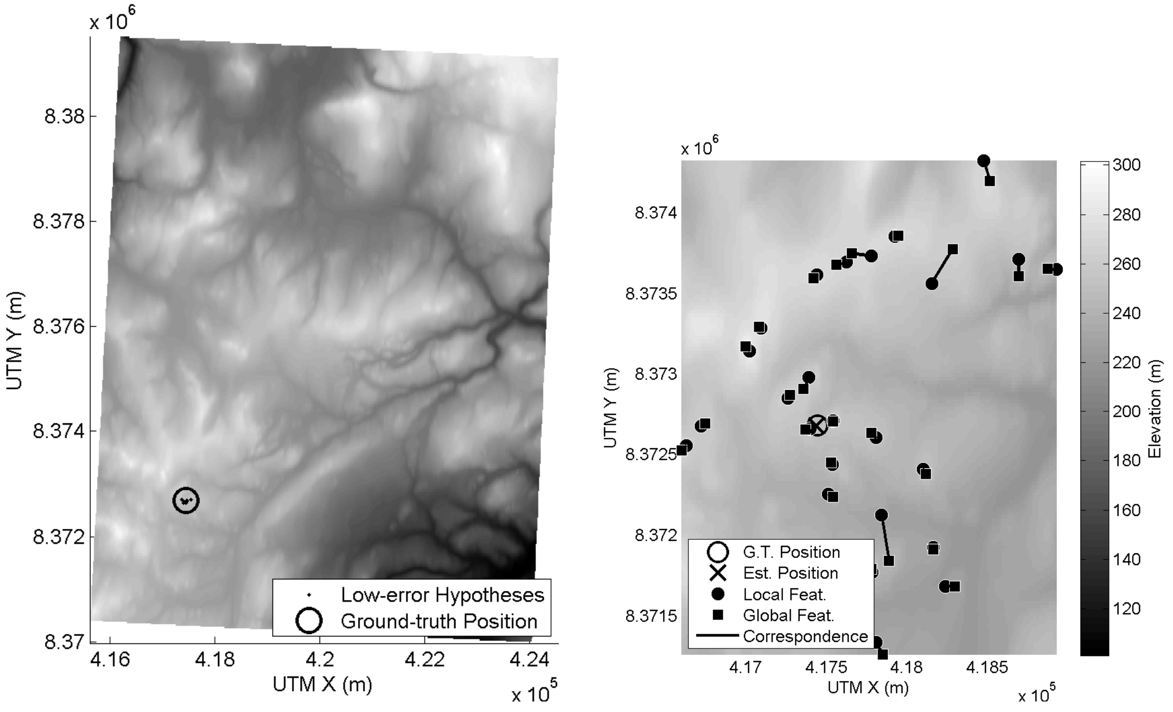
3.3.4 Feature Matching Example

Given as input orientation measurements, all global features, a local map and its detected local features, it has been shown how to output a pose estimate and correspondences between local and global features. Figure 3.14 demonstrates the hypothesized locations of the rover after each stage of the feature matching algorithm, as well as the feature correspondences obtained from the best-scoring hypothesis.



(a) 190,000 unfiltered

(b) 43,000 preliminary filtered



(c) 17 valid

(d) Zoomed-in view of best-scoring hypothesis

Figure 3.14: Decrease in the number of hypotheses at major stages in the algorithm for scan A08. In (d), local features are transformed to the global frame using the transformation from the best-scoring hypothesis. Elevation colourbar applies to all plots. UTM zone is 16X.

3.4 Pose Refinement

This section discusses pose refinement, which is represented as box (c) in the overall architecture diagram of Figure 3.1.

In a long-range traverse, the rover regularly stops to scan the surrounding terrain with its lidar. After some processing, a number of local features are detected from these local scans. Orbital data is processed in a similar manner to produce a set of global features. Correspondences between local and global features and initial estimates of rover pose are then found using DARCES. These correspondences are fed into the pose refinement stage along with odometry and orientation measurements.

The Multi-frame Odometry-compensated Global Alignment (MOGA) algorithm, developed in this thesis, refines poses by simultaneously minimizing the errors between measurements and estimates in each frame. RANSAC is run before the final, full MOGA alignment to discard poor feature correspondences. The sections to follow describe the pose refinement architecture in detail.

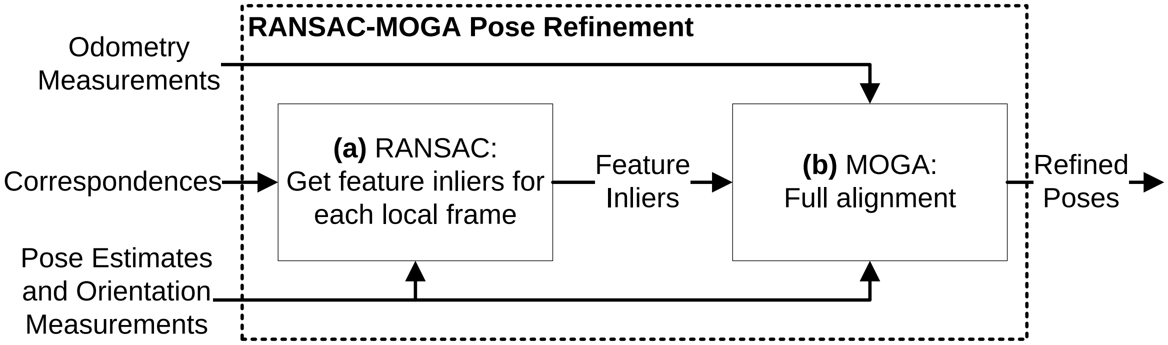


Figure 3.15: Pose refinement architecture.

3.4.1 Multi-frame Odometry-compensated Global Alignment (MOGA)

The main goal is to estimate the transformations from the global frame, $\underrightarrow{\mathcal{F}}_o$, to each of the M local frames, $\underrightarrow{\mathcal{F}}_\ell$. This M-frame alignment problem is depicted in Figure 3.16.

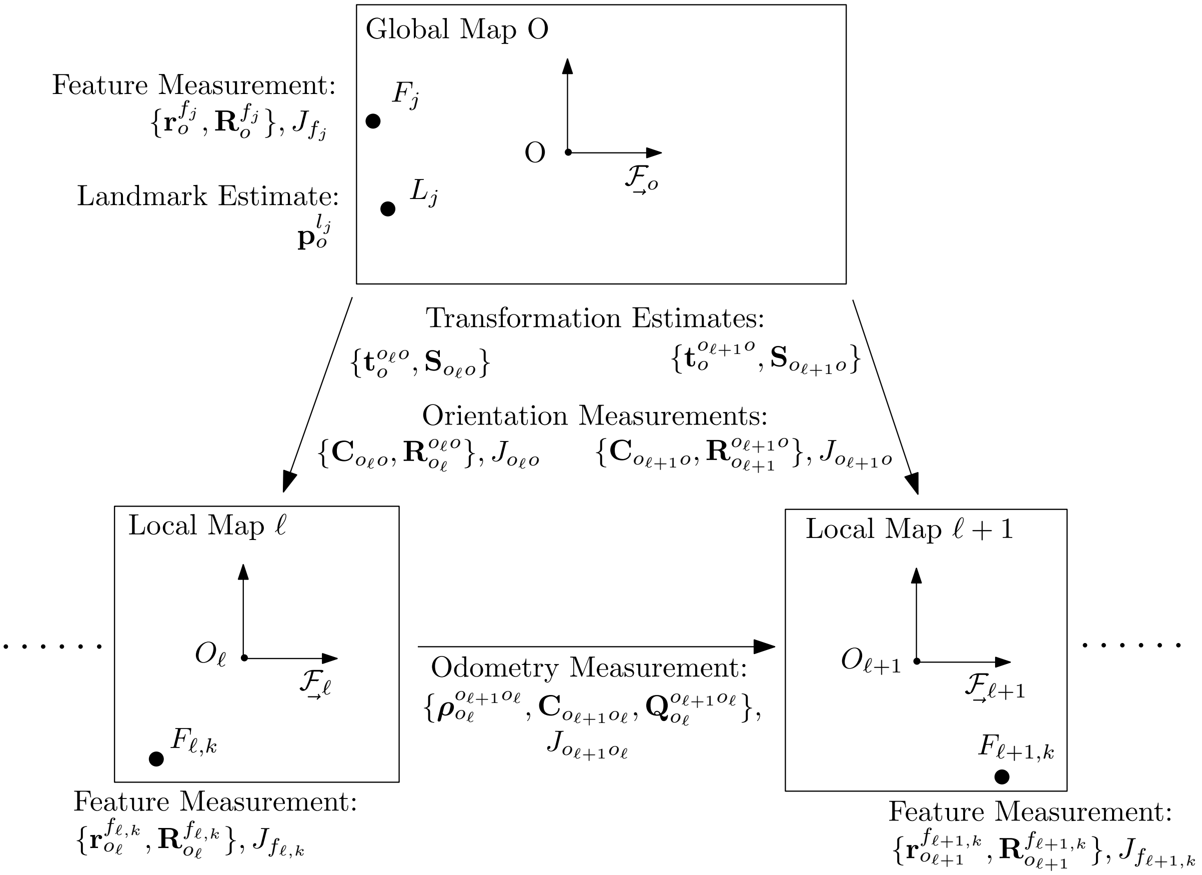


Figure 3.16: The M-frame alignment problem. Inputs: global and local features positions with correspondences, odometry measurements, orientation measurements. Outputs: transformation estimates from global to local frames, feature landmark positions in the global frame.

Input measurements are assumed to be corrupted with white, zero-mean, Gaussian noise. There are four types of measurement:

- In the global frame, $\underrightarrow{\mathcal{F}}_o$, there are N measured feature positions, $\mathbf{r}_o^{f_j}$, with 3×3 covariance matrix, $\mathbf{R}_o^{f_j}$. These are the features detected in the orbital map.
- In each local frame, $\underrightarrow{\mathcal{F}}_\ell$, there are $N_\ell \leq N$ measured feature positions, where $\mathbf{r}_{o_\ell}^{f_{\ell,k}}$ is the position of a feature k in $\underrightarrow{\mathcal{F}}_\ell$. The 3×3 covariance matrix for one of these measurements

is $\mathbf{R}_{o_\ell}^{f_{\ell,k}}$. These are the features detected in the lidar scans. Corresponding global features have been found for each local feature.

- Between local frames there are odometry measurements of rotations, $\mathbf{C}_{o_{\ell+1}o_\ell}$, and translation, $\boldsymbol{\rho}_{o_\ell}^{o_{\ell+1}o_\ell}$, between adjacent local frames $\underrightarrow{\mathcal{F}}_\ell$ and $\underrightarrow{\mathcal{F}}_{\ell+1}$. The combined 6×6 covariance matrix is $\mathbf{Q}_{o_\ell}^{o_{\ell+1}o_\ell}$. These would come from visual odometry, for example.
- For each local frame, $\underrightarrow{\mathcal{F}}_\ell$, there is also an orientation measurement, $\mathbf{C}_{o_\ell o}$, from $\underrightarrow{\mathcal{F}}_o$ to $\underrightarrow{\mathcal{F}}_\ell$ with 3×3 covariance matrix $\mathbf{R}_{o_\ell}^{o_\ell o}$. These would come from a sun sensor and inclinometer, for example.

There are $2M + N$ design parameters that must be estimated. These can be divided into three types:

- Estimated rotations, $\mathbf{S}_{o_\ell o}$, from the global frame $\underrightarrow{\mathcal{F}}_o$ to each of the M local frames $\underrightarrow{\mathcal{F}}_\ell$.
- Estimated translations, $\mathbf{t}_o^{o_\ell o}$, from the global frame $\underrightarrow{\mathcal{F}}_o$ to each of the M local frames $\underrightarrow{\mathcal{F}}_\ell$, and expressed in the $\underrightarrow{\mathcal{F}}_o$.
- Estimated feature positions, $\mathbf{p}_o^{l_j}$, expressed in $\underrightarrow{\mathcal{F}}_o$ for the N unique features input into MOGA. These estimated features, also called feature landmarks, will not only improve the estimate of feature positions, but also the transformation estimates between frames.

There are four types of error between estimated and measured quantities that will be minimized by optimizing the design parameters:

- J_{f_j}, between estimated landmark feature positions and measured global feature positions.
- $J_{f_{\ell,k}}$, between estimated landmark feature positions and measured local feature positions.
- $J_{o_{\ell+1}o_\ell}$, between estimated and odometry-measured transformations from $\underrightarrow{\mathcal{F}}_\ell$ to $\underrightarrow{\mathcal{F}}_{\ell+1}$.
- $J_{o_\ell o}$, between estimated and measured frame rotations from $\underrightarrow{\mathcal{F}}_o$ to $\underrightarrow{\mathcal{F}}_\ell$.

To allow all design parameters to be optimized simultaneously, each of these error terms must be expressed as a function of a common design parameter column, $\mathbf{z}$. The details of this column will be discussed in the next section. The overall objective function, $J(\mathbf{z})$, is then the sum of all these error terms, giving

$$J(\mathbf{z}) := \sum_{j}^{N} J_{f_j}(\mathbf{z}) + \sum_{\ell}^{M}\sum_{k}^{N_\ell} J_{f_{\ell,k}}(\mathbf{z}) + \sum_{\ell}^{M-1} J_{o_{\ell+1}o_\ell}(\mathbf{z}) + \sum_{\ell}^{M} J_{o_\ell o}(\mathbf{z}). \tag{3.13}$$

It is important to note that the local frame sum, $\sum_{\ell}^{M}\sum_{k}^{N_\ell} J_{f_{\ell,k}}(\mathbf{z})$, is over all features visible in a specific frame, ℓ. There may be features in frame $\ell = 1$ that are not visible in frame $\ell = 2$ and vice versa due to topographic occlusions and range limitations.

The optimal design parameter column, $\mathbf{z}^*$, is sought by minimizing the objective function through unconstrained optimization. Therefore,

$$\mathbf{z}^* = \operatorname{argmin}_{\mathbf{z}} J(\mathbf{z}). \tag{3.14}$$

An iterative Gauss-Newton algorithm (Björck, 1996) is used to solve this nonlinear least-squares problem. The sections to follow discuss the details of each error term, how they can be combined, and how to obtain an optimal solution.

3.4.2 Local Feature Terms, $J_{f_{\ell,k}}$

A local feature's error contribution term, $J_{f_{\ell,k}}$, is simply the squared difference between a feature's position measurement in $\underrightarrow{\mathcal{F}}_\ell$, $\mathbf{r}_{o_\ell}^{f_{\ell,k}}$, and its corresponding estimated position in $\underrightarrow{\mathcal{F}}_o$, $\mathbf{p}_o^{l_{\omega(\ell,k)}}$, transformed to $\underrightarrow{\mathcal{F}}_\ell$ with estimated translation and rotation $\mathbf{t}_o^{o_\ell o}$ and $\mathbf{S}_{o_\ell o}$. This produces

$$J_{f_{\ell,k}} := \frac{1}{2}\left(\mathbf{r}_{o_\ell}^{f_{\ell,k}} - \mathbf{S}_{o_\ell o}\left(\mathbf{p}_o^{l_{\omega(\ell,k)}} - \mathbf{t}_o^{o_\ell o}\right)\right)^T \mathbf{R}_{o_\ell}^{f_j\,-1}\left(\mathbf{r}_{o_\ell}^{f_{\ell,k}} - \mathbf{S}_{o_\ell o}\left(\mathbf{p}_o^{l_{\omega(\ell,k)}} - \mathbf{t}_o^{o_\ell o}\right)\right). \tag{3.15}$$

For purposes of iteratively minimizing these error contribution terms with the Gauss-Newton algorithm, a function quadratic in the design parameters is desired. It is therefore necessary to make quadratic approximations to the design parameters. This is done by perturbing each design parameter slightly at each iteration and solving for the perturbations that move the objective function closer to the minimum. This perturbation is straightforward for the landmark position and translation parameters:

$$\mathbf{t}_{o_\ell}^{o_\ell o} := \bar{\mathbf{t}}_{o_\ell}^{o_\ell o} + \delta\mathbf{t}_{o_\ell}^{o_\ell o}, \tag{3.16}$$

$$\mathbf{p}_{o}^{l_{\omega(\ell,k)}} =: \bar{\mathbf{p}}_{o}^{l_{\omega(\ell,k)}} + \delta\mathbf{p}_{o}^{l_{\omega(\ell,k)}} \tag{3.17}$$

where $\bar{\mathbf{t}}_{o_\ell}^{o_\ell o}$ and $\bar{\mathbf{p}}_{o}^{l_{\omega(\ell,k)}}$ are the nominal translation and landmark position estimates, and $\delta\mathbf{t}_{o_\ell}^{o_\ell o}$ and $\delta\mathbf{p}_{o}^{l_{\omega(\ell,k)}}$ are the changes to these estimates. Rotations are slightly more involved because rotation matrices cannot be summed, but must be multiplied in order. Thus, a multiplicative perturbation of the rotation is chosen of the form

$$\mathbf{S}_{o_\ell o} := \delta\mathbf{S}_{o_\ell o}\bar{\mathbf{S}}_{o_\ell o} \tag{3.18}$$

where $\bar{\mathbf{S}}_{o_\ell o}$ is the nominal rotation estimate and $\delta\mathbf{S}_{o_\ell o}$ is a small perturbation rotation. It is also necessary that this small rotation be expressed as a vector so it can be incorporated into the solution column, $\mathbf{z}$. This can be done with the infinitesimal rotation vector, $\delta\boldsymbol{\theta}_{o_\ell o}$:

$$\delta\mathbf{S}_{o_\ell o} \approx \mathbf{1} - \delta\boldsymbol{\theta}_{o_\ell o}^{\times} \tag{3.19}$$

where $\delta\boldsymbol{\theta}_{o_\ell o}$ is a 3×1 vector whose magnitude is the angle of the small rotation, and whose direction is the axis of rotation. A good introduction to this small angle parametrization is given by Hughes (1986). The skew symmetric operator is

$$\mathbf{u}^{\times} := \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}. \tag{3.20}$$

The approximate perturbation for rotations is thus

$$\mathbf{S}_{o_\ell o} \approx \left(\mathbf{1} - \delta\boldsymbol{\theta}_{o_\ell o}^{\times}\right)\bar{\mathbf{S}}_{o_\ell o}. \tag{3.21}$$

The desired rotation vector, $\delta\boldsymbol{\theta}_{o_\ell o}$, can be isolated knowing that the cross product is anticommutative, i.e. $\mathbf{u}^{\times}\mathbf{v} = -\mathbf{v}^{\times}\mathbf{u}$. Substituting Equations (3.16), (3.17) and (3.21) into the objective function of Equation (3.15), using this anticommutative property and expressing the resulting objective function for clarity as $J_{f_{\ell,k}} =: \frac{1}{2}\mathbf{e}^T_{f_{\ell,k}}\mathbf{R}^{f_{\ell,k}{}^{-1}}_{o_\ell}\mathbf{e}_{f_{\ell,k}}$ gives

$$\begin{aligned}\mathbf{e}_{f_{\ell,k}} \;:\approx\; & \mathbf{r}^{f_{\ell,k}}_{o_\ell} - \bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{p}}^{l_{\omega(\ell,k)}}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right) \\ & - \left(\left(\bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{p}}^{l_{\omega(\ell,k)}}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right)\right)^{\times}\delta\boldsymbol{\theta}_{o_\ell o} + \bar{\mathbf{S}}_{o_\ell o}\left(\delta\mathbf{p}^{l_{\omega(\ell,k)}}_{o} - \delta\mathbf{t}^{o_\ell o}_{o}\right)\right). \end{aligned} \tag{3.22}$$

Note that an approximation has been made by dropping terms with higher order perturbations. From the above equation, it is seen that the inter-frame terms contain updates for all involved design parameters. These perturbed quantities make up the $\mathbf{z}$ column, which can now be expressed as

$$\mathbf{z} := \begin{bmatrix} \delta\mathbf{p}^{l}_{o} \\ \delta\mathbf{t}_{o} \\ \delta\boldsymbol{\theta} \end{bmatrix} \tag{3.23}$$

where $\delta\mathbf{p}^{l}_{o}$ represents the column of all N landmark position perturbations in $\underrightarrow{\mathcal{F}}_{o}$, and $\delta\mathbf{t}_{o}$ and $\delta\boldsymbol{\theta}$ are respectively the columns of all M translation and M rotation perturbations from $\underrightarrow{\mathcal{F}}_{o}$ to $\underrightarrow{\mathcal{F}}_{\ell}$. The error contribution term can then be written in the desired form as $\mathbf{e}_{f_{\ell,k}} = \mathbf{b}_{f_{\ell,k}} - \mathbf{A}_{f_{\ell,k}}\mathbf{z}$, where

$$\mathbf{b}_{f_{\ell,k}} := \mathbf{r}^{f_{\ell,k}}_{o_\ell} - \bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{p}}^{l_{\omega(\ell,k)}}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right), \tag{3.24}$$

$$\mathbf{A}_{f_{\ell,k}} := \begin{bmatrix} \mathbf{D}_{f_{\ell,k}} & \mathbf{E}_{f_{\ell,k}} & \mathbf{F}_{f_{\ell,k}} \end{bmatrix}, \tag{3.25}$$

$$\mathbf{D}_{f_{\ell,k}} := \Big[\underbrace{\mathbf{0}\cdots \overbrace{\bar{\mathbf{S}}_{o_\ell o}}^{\omega(\ell,k)^{th}\text{term}} \cdots\mathbf{0}}_{N\ \text{landmarks}}\Big], \tag{3.26}$$

$$\mathbf{E}_{f_{\ell,k}} := \Big[\underbrace{\mathbf{0}\cdots \overbrace{-\bar{\mathbf{S}}_{o_\ell o}}^{\ell^{th}\text{term}} \cdots\mathbf{0}}_{M\ \text{frames}}\Big], \tag{3.27}$$

$$\mathbf{F}_{f_{\ell,k}} := \left[\underbrace{\mathbf{0}\cdots \overbrace{\left(\bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{p}}^{l_{\omega(\ell,k)}}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right)\right)^{\times}}^{\ell^{th}\text{term}} \cdots\mathbf{0}}_{M\ \text{frames}}\right]. \tag{3.28}$$

3.4.3 Global Feature Terms, J_{f_j}

The J_{f_j} error term is derived in the same way, but with the simplification that the landmarks and measured features both belong to the same frame, $\underrightarrow{\mathcal{F}}_o$. This means there is no inter-frame rotation or translation. For a single feature, F_j, the objective function contribution with perturbations included is

$$J_{f_j} := \frac{1}{2}\left(\bar{\mathbf{p}}_o^{l_j} + \delta\mathbf{p}_o^{l_j} - \mathbf{r}_o^{f_j}\right)^T \mathbf{R}_o^{f_j^{-1}} \left(\bar{\mathbf{p}}_o^{l_j} + \delta\mathbf{p}_o^{l_j} - \mathbf{r}_o^{f_j}\right). \tag{3.29}$$

The error contribution in the desired form is $\mathbf{e}_{f_j} = \mathbf{b}_{f_j} - \mathbf{A}_{f_j}\mathbf{z}$, where

$$\mathbf{b}_{f_j} := \bar{\mathbf{p}}_o^{l_j} - \mathbf{r}_o^{f_j}, \tag{3.30}$$

$$\mathbf{A}_{f_j} := \begin{bmatrix} \mathbf{D}_{f_j} & \mathbf{0}_{3\times 3M} & \mathbf{0}_{3\times 3M} \end{bmatrix}, \tag{3.31}$$

$$\mathbf{D}_{f_j} := \big[\underbrace{\mathbf{0}\cdots \overbrace{-\mathbf{1}}^{j^{th}\text{term}} \cdots\mathbf{0}}_{N\text{ landmarks}}\big]. \tag{3.32}$$

3.4.4 Odometry Terms, $J_{o_{\ell+1}o_\ell}$

The objective is to minimize the error, $J_{o_{\ell+1}o_\ell}$, between a measurement and a design parameter estimate of the transformation between two local frames, $\underrightarrow{\mathcal{F}}_\ell$, and $\underrightarrow{\mathcal{F}}_{\ell+1}$. As shown in Figure 3.17, two paths must be compared: one going directly from map ℓ to map $\ell+1$ using an odometry measurement, and one passing through the global map using design parameters.

The error contributions of interest are the translational, $\delta\boldsymbol{\rho}_{o_\ell}^{o_{\ell+1}o_\ell}$, and rotational, $\delta\boldsymbol{\psi}_{o_{\ell+1}o_\ell}$, differences between these two paths. In order to obtain an overall scalar error, $J_{o_{\ell+1}o_\ell}$, the two 3×1 columns are combined using a 6×6 covariance matrix, $\mathbf{Q}_{o_\ell}^{o_{\ell+1}o_\ell}$, as weighting:

$$J_{o_{\ell+1}o_\ell} := \frac{1}{2}\begin{bmatrix} \delta\boldsymbol{\rho}_{o_\ell}^{o_{\ell+1}o_\ell} \\ \delta\boldsymbol{\psi}_{o_{\ell+1}o_\ell} \end{bmatrix}^T \mathbf{Q}_{o_\ell}^{o_{\ell+1}o_\ell^{-1}} \begin{bmatrix} \delta\boldsymbol{\rho}_{o_\ell}^{o_{\ell+1}o_\ell} \\ \delta\boldsymbol{\psi}_{o_{\ell+1}o_\ell} \end{bmatrix}. \tag{3.33}$$

The derivations of the rotation and translation terms follow.

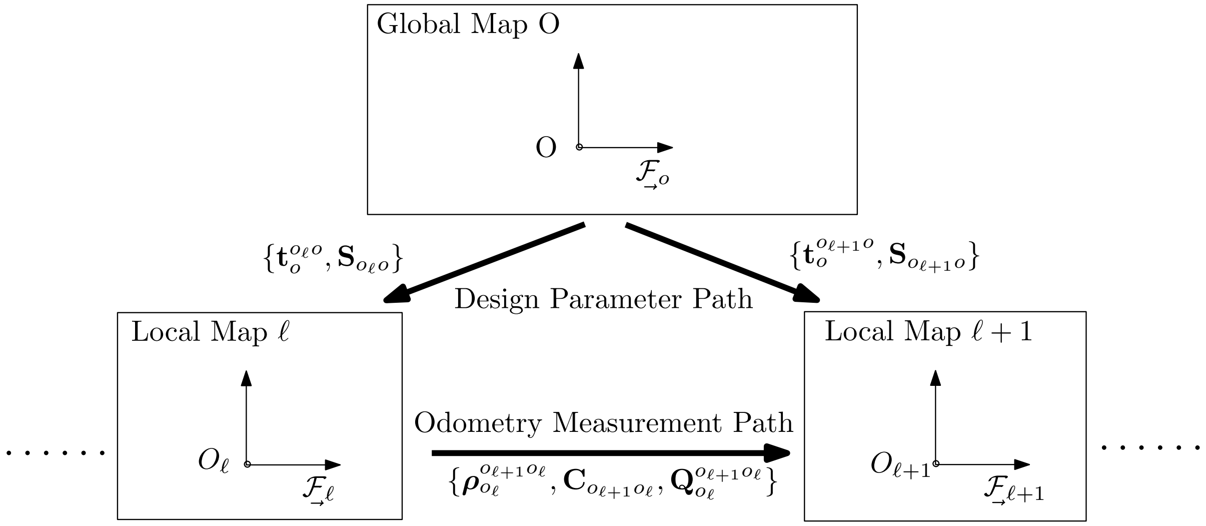


Figure 3.17: Two different transformation paths from map ℓ to $\ell + 1$ in the three frame model.

Rotation

The goal is to have the difference between the measured and estimated rotations be as small as possible. Thus

$$\delta\mathbf{C}_{o_{\ell+1}o_\ell} := \mathbf{C}_{o_{\ell+1}o_\ell}\mathbf{S}^T_{o_{\ell+1}o_\ell} \tag{3.34}$$

should approach identity, where $\mathbf{S}_{o_{\ell+1}o_\ell} := \mathbf{S}_{o_{\ell+1}o}\mathbf{S}^T_{o_\ell o}$ is the path through the orbital map. The rotation $\mathbf{S}_{o_\ell o}$ can be separated into a small rotation, $(\mathbf{1} - \delta\boldsymbol{\theta}^\times_{o_\ell o})$, multiplied by a large one, $\bar{\mathbf{S}}_{o_\ell o}$. Over an iteration, the latter would correspond to the most current estimate for this rotation. The same may be said for $\mathbf{S}_{o_{\ell+1}o}$. Also, $\delta\mathbf{C}_{o_{\ell+1}o_\ell}$ may be assumed to be small, so

$$\delta\mathbf{C}_{o_{\ell+1}o_\ell} \approx \mathbf{1} - \delta\boldsymbol{\psi}^\times_{o_{\ell+1}o_\ell} \tag{3.35}$$

where $\delta\boldsymbol{\psi}_{o_{\ell+1}o_\ell}$ must be isolated. Using these properties,

$$\mathbf{1} - \delta\boldsymbol{\psi}^\times_{o_{\ell+1}o_\ell} \approx \mathbf{C}_{o_{\ell+1}o_\ell}(\mathbf{1} - \delta\boldsymbol{\theta}^\times_{o_\ell o})\bar{\mathbf{S}}_{o_\ell o}\bar{\mathbf{S}}^T_{o_{\ell+1}o}(\mathbf{1} + \delta\boldsymbol{\theta}^\times_{o_{\ell+1}o}). \tag{3.36}$$

Defining $\bar{\mathbf{S}}_{o_{\ell+1}o_\ell} := \bar{\mathbf{S}}_{o_{\ell+1}o}\bar{\mathbf{S}}^T_{o_\ell o}$ and strategically introducing $\mathbf{1} = \bar{\mathbf{S}}^T_{o_{\ell+1}o_\ell}\bar{\mathbf{S}}_{o_{\ell+1}o_\ell}$ gives

$$\mathbf{1} - \delta\boldsymbol{\psi}^\times_{o_{\ell+1}o_\ell} \approx \mathbf{C}_{o_{\ell+1}o_\ell}\bar{\mathbf{S}}^T_{o_{\ell+1}o_\ell}\bar{\mathbf{S}}_{o_{\ell+1}o_\ell}(\mathbf{1} - \delta\boldsymbol{\theta}^\times_{o_\ell o})\bar{\mathbf{S}}^T_{o_{\ell+1}o_\ell}(\mathbf{1} + \delta\boldsymbol{\theta}^\times_{o_{\ell+1}o}). \tag{3.37}$$

It is again assumed that the difference between measured and estimated rotations is small, so

$$\mathbf{1} - \delta\boldsymbol{\phi}^{\times}_{o_{\ell+1}o_\ell} \approx \mathbf{C}_{o_{\ell+1}o_\ell}\bar{\mathbf{S}}^{T}_{o_{\ell+1}o_\ell}. \tag{3.38}$$

Substituting this, expanding, and neglecting all higher order terms in small angles to be zero gives

$$\delta\boldsymbol{\psi}^{\times}_{o_{\ell+1}o_\ell} \approx \delta\boldsymbol{\phi}^{\times}_{o_{\ell+1}o_\ell} + \bar{\mathbf{S}}_{o_{\ell+1}o_\ell}\delta\boldsymbol{\theta}^{\times}_{o_\ell o}\bar{\mathbf{S}}^{T}_{o_{\ell+1}o_\ell} - \delta\boldsymbol{\theta}^{\times}_{o_{\ell+1}o}. \tag{3.39}$$

Using the identity $\mathbf{C}\mathbf{z}^{\times}\mathbf{C}^{T} \equiv (\mathbf{C}\mathbf{z})^{\times}$, for a rotation matrix, $\mathbf{C}$, and knowing that the cross operator is linear,

$$\delta\boldsymbol{\psi}^{\times}_{o_{\ell+1}o_\ell} \approx \left(\delta\boldsymbol{\phi}_{o_{\ell+1}o_\ell} + \bar{\mathbf{S}}_{o_{\ell+1}o_\ell}\delta\boldsymbol{\theta}_{o_\ell o} - \delta\boldsymbol{\theta}_{o_{\ell+1}o}\right)^{\times}. \tag{3.40}$$

Equating the cross-operated terms gives an equation of the desired form:

$$\delta\boldsymbol{\psi}_{o_{\ell+1}o_\ell} \approx \delta\boldsymbol{\phi}_{o_{\ell+1}o_\ell} - \left(\delta\boldsymbol{\theta}_{o_{\ell+1}o} - \bar{\mathbf{S}}_{o_{\ell+1}o_\ell}\delta\boldsymbol{\theta}_{o_\ell o}\right). \tag{3.41}$$

Translation

The difference between measured and estimated translations is expected to be small, so

$$\delta\boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} := \boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} - \mathbf{S}_{o_\ell o}\left(\mathbf{t}^{o_{\ell+1}o}_{o} - \mathbf{t}^{o_\ell o}_{o}\right). \tag{3.42}$$

Perturbing the design parameters gives

$$\delta\boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} = \boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} - (\mathbf{1} - \delta\boldsymbol{\theta}^{\times}_{o_\ell o})\bar{\mathbf{S}}_{o_\ell o}\left(\delta\mathbf{t}^{o_{\ell+1}o}_{o} + \bar{\mathbf{t}}^{o_{\ell+1}o}_{o} - \delta\mathbf{t}^{o_\ell o}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right). \tag{3.43}$$

Expanding the equation above and ignoring higher order terms in small perturbations produces

$$\begin{aligned} \delta\boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} &= \boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} - \bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{t}}^{o_{\ell+1}o}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right) \\ &\quad - \Big(\bar{\mathbf{S}}_{o_\ell o}\left(\delta\mathbf{t}^{o_{\ell+1}o}_{o} - \delta\mathbf{t}^{o_\ell o}_{o}\right) + \left(\bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{t}}^{o_{\ell+1}o}_{o} - \bar{\mathbf{t}}^{o_\ell o}_{o}\right)\right)^{\times}\delta\boldsymbol{\theta}_{o_\ell o}\Big). \end{aligned} \tag{3.44}$$

Overall Odometry Error

Substituting the results from the previous two sections into the objective function of Equation 3.33 and defining $J_{o_{\ell+1}o_\ell} =: \frac{1}{2}\mathbf{e}^T_{o_{\ell+1}o_\ell}\mathbf{Q}^{o_{\ell+1}o_\ell^{-1}}_{o_\ell}\mathbf{e}_{o_{\ell+1}o_\ell}$, the odometry error as a function of the full design parameter column, $\mathbf{z}$, is $\mathbf{e}_{o_{\ell+1}o_\ell} =: \mathbf{b}_{o_{\ell+1}o_\ell} - \mathbf{A}_{o_{\ell+1}o_\ell}\mathbf{z}$, where

$$\mathbf{b}_{o_{\ell+1}o_\ell} := \begin{bmatrix} \boldsymbol{\rho}^{o_{\ell+1}o_\ell}_{o_\ell} - \bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{t}}^{o_{\ell+1}o}_o - \bar{\mathbf{t}}^{o_\ell o}_o\right) \\ \delta\boldsymbol{\phi}_{o_{\ell+1}o_\ell} \end{bmatrix}, \tag{3.45}$$

$$\mathbf{A}_{o_{\ell+1}o_\ell} := \begin{bmatrix} \mathbf{0}_{6\times N} & \mathbf{E}_{o_{\ell+1}o_\ell} & \mathbf{F}_{o_{\ell+1}o_\ell} \end{bmatrix} \tag{3.46}$$

$$\mathbf{E}_{o_{\ell+1}o_\ell} := \begin{bmatrix} \mathbf{0}\cdots\overbrace{-\bar{\mathbf{S}}_{o_\ell o}}^{\ell^{th}\text{term}} \; \overbrace{\bar{\mathbf{S}}_{o_\ell o}}^{\ell+1^{th}\text{term}} \;\cdots\mathbf{0} \\ \mathbf{0}_{3\times M} \end{bmatrix}, \tag{3.47}$$

$$\mathbf{F}_{o_{\ell+1}o_\ell} := \begin{bmatrix} \mathbf{0} \;\cdots & \left(\bar{\mathbf{S}}_{o_\ell o}\left(\bar{\mathbf{t}}^{o_{\ell+1}o}_o - \bar{\mathbf{t}}^{o_\ell o}_o\right)\right)^\times & \mathbf{0} & \cdots\;\mathbf{0} \\ \mathbf{0} \;\cdots & \underbrace{-\bar{\mathbf{S}}_{o_{\ell+1}o}\bar{\mathbf{S}}^T_{o_\ell o}}_{\ell^{th}\text{term}} & \underbrace{\mathbf{1}}_{\ell+1^{th}\text{term}} & \cdots\;\mathbf{0} \end{bmatrix}_{6\times M}. \tag{3.48}$$

The $\delta\boldsymbol{\phi}_{o_{\ell+1}o_\ell}$ parameter can be obtained by converting $\mathbf{C}_{o_{\ell+1}o_\ell}\mathbf{S}_{o_\ell o}\mathbf{S}^T_{o_{\ell+1}o}$ into axis-angle representation, as explained in Appendix C.

3.4.5 Measured Orientation Terms, $J_{o_\ell o}$

Estimates can further be constrained if inclinometer and/or sun sensor orientation measurements are available. These measurements are added to the overall objective function in a similar way to Section 3.4.4. The error contribution is

$$J_{o_\ell o} := \frac{1}{2}\mathbf{e}^T_{o_\ell o}\mathbf{R}^{o_\ell o^{-1}}_{o_\ell}\mathbf{e}_{o_\ell o}. \tag{3.49}$$

Assuming the error between the measured orientation and estimated orientation from $\underrightarrow{\mathcal{F}}_o$ to $\underrightarrow{\mathcal{F}}_\ell$ is small, then

$$\delta\mathbf{C}_{o_\ell o} := \mathbf{C}_{o_\ell o}\mathbf{S}^T_{o_\ell o}, \tag{3.50}$$

$$\mathbf{1} - \delta\boldsymbol{\psi}^\times_{o_\ell o} \approx \mathbf{C}_{o_\ell o}\mathbf{S}^T_{o_\ell o}. \tag{3.51}$$

Approximating the estimated rotation into a small rotation and a nominal rotation gives

$$\mathbf{1} - \delta\boldsymbol{\psi}_{o_\ell o}^{\times} \approx \mathbf{C}_{o_\ell o}\bar{\mathbf{S}}_{o_\ell o}^{T}\left(\mathbf{1} + \delta\boldsymbol{\theta}_{o_\ell o}^{\times}\right). \tag{3.52}$$

The measured orientation multiplied by the transpose of the nominal orientation is also small, therefore

$$\begin{aligned} \mathbf{C}_{o_\ell o}\mathbf{S}_{o_\ell o}^{T} &\approx: \left(\mathbf{1} - \delta\boldsymbol{\phi}_{o_\ell o}^{\times}\right), && (3.53) \\ \mathbf{1} - \delta\boldsymbol{\psi}_{o_\ell o}^{\times} &\approx \left(\mathbf{1} - \delta\boldsymbol{\phi}_{o_\ell o}^{\times}\right)\left(\mathbf{1} + \delta\boldsymbol{\theta}_{o_\ell o}^{\times}\right). && (3.54) \end{aligned}$$

Expanding and dropping higher order terms,

$$\delta\boldsymbol{\psi}_{o_\ell o}^{\times} \approx \delta\boldsymbol{\phi}_{o_\ell o}^{\times} - \delta\boldsymbol{\theta}_{o_\ell o}^{\times}. \tag{3.55}$$

The cross operator is a linear operator, therefore

$$\delta\boldsymbol{\psi}_{o_\ell o} = \delta\boldsymbol{\phi}_{o_\ell o} - \delta\boldsymbol{\theta}_{o_\ell o}. \tag{3.56}$$

The error term can then be written in the desired form as $\mathbf{e}_{o_\ell o} := \mathbf{b}_{o_\ell o} - \mathbf{A}_{o_\ell o}\mathbf{z}$, where

$$\begin{aligned} \mathbf{b}_{o_\ell o} &:= \delta\boldsymbol{\phi}_{o_\ell o}, && (3.57) \\ \mathbf{A}_{o_\ell o} &:= \begin{bmatrix} \mathbf{0}_{3\times N} & \mathbf{0}_{3\times M} & \mathbf{F}_{o_\ell o} \end{bmatrix}, && (3.58) \\ \mathbf{F}_{o_\ell o} &:= \big[\underbrace{\mathbf{0}\cdots \overbrace{\mathbf{1}}^{\ell^{th}\text{term}} \cdots\mathbf{0}}_{M \text{ frames}}\big]. && (3.59) \end{aligned}$$

The $\delta\boldsymbol{\phi}_{o_\ell o}$ parameter can be obtained by converting $\mathbf{C}_{o_\ell o}\bar{\mathbf{S}}_{o_\ell o}^{T}$ into axis-angle form, as explained in Appendix C.

3.4.6 Combining Terms and Optimization

These objective function error terms must be combined in such a way that the overall minimization problem can be easily expressed. The overall objective function becomes

$$\begin{aligned} J = & \sum_{j}^{N}\left(\frac{1}{2}\mathbf{e}_{f_j}^{T}\mathbf{R}_{o}^{f_j\,-1}\mathbf{e}_{f_j}\right) + \sum_{\ell}^{M}\sum_{k}^{N_\ell}\left(\frac{1}{2}\mathbf{e}_{f_{\ell,k}}^{T}\mathbf{R}_{o_\ell}^{f_{\ell,k}\,-1}\mathbf{e}_{f_{\ell,k}}\right) \\ & + \sum_{\ell=1}^{M-1}\left(\frac{1}{2}\mathbf{e}_{o_{\ell+1}o_\ell}^{T}\mathbf{Q}_{o_\ell}^{o_{\ell+1}o_\ell\,-1}\mathbf{e}_{o_{\ell+1}o_\ell}\right) + \sum_{\ell=1}^{M}\left(\frac{1}{2}\mathbf{e}_{o_\ell o}^{T}\mathbf{R}_{o_\ell}^{o_\ell o\,-1}\mathbf{e}_{o_\ell o}\right). \end{aligned} \tag{3.60}$$

It is desired to take a step, $\mathbf{z}^*$, to the minimum of the current local quadratic approximation of J. Thus $\mathbf{z}^*$ is chosen so that $\frac{\partial J^T}{\partial \mathbf{z}} = \mathbf{0}$. Taking the derivative of Equation 3.60 with respect to $\mathbf{z}$, and setting the result equal to zero gives

$$\begin{aligned}
\mathbf{0} \;=\; & \sum_{j}^{N} \left(\mathbf{A}_{f_j}^T {\mathbf{R}_o^{f_j}}^{-1} \mathbf{b}_{f_j} - \mathbf{A}_{f_j}^T {\mathbf{R}_o^{f_j}}^{-1} \mathbf{A}_{f_j} \mathbf{z}^* \right) \\
& + \sum_{\ell}^{M} \sum_{k}^{N_\ell} \left(\mathbf{A}_{f_{\ell,k}}^T {\mathbf{R}_{o_\ell}^{f_{\ell,k}}}^{-1} \mathbf{b}_{f_{\ell,k}} - \mathbf{A}_{f_{\ell,k}}^T {\mathbf{R}_{o_\ell}^{f_{\ell,k}}}^{-1} \mathbf{A}_{f_{\ell,k}} \mathbf{z}^* \right) \\
& + \sum_{\ell}^{M-1} \left(\mathbf{A}_{o_{\ell+1} o_\ell}^T {\mathbf{Q}_{o_\ell}^{o_{\ell+1} o_\ell}}^{-1} \mathbf{b}_{o_{\ell+1} o_\ell} - \mathbf{A}_{o_{\ell+1} o_\ell}^T {\mathbf{Q}_{o_\ell}^{o_{\ell+1} o_\ell}}^{-1} \mathbf{A}_{o_{\ell+1} o_\ell} \mathbf{z}^* \right) \\
& + \sum_{\ell}^{M} \left(\mathbf{A}_{o_\ell o}^T {\mathbf{R}_{o_\ell}^{o_\ell o}}^{-1} \mathbf{b}_{o_\ell o} - \mathbf{A}_{o_\ell o}^T {\mathbf{R}_{o_\ell}^{o_\ell o}}^{-1} \mathbf{A}_{o_\ell o} \mathbf{z}^* \right).
\end{aligned} \tag{3.61}$$

By factoring out $\mathbf{z}^*$ and rearranging, the equation can now be rewritten in the form

$$\mathbf{B}\mathbf{z}^* = \mathbf{y} \tag{3.62}$$

where

$$\begin{aligned}
\mathbf{B} \;:=\; & \sum_{j}^{N} \left(\mathbf{A}_{f_j}^T {\mathbf{R}_o^{f_j}}^{-1} \mathbf{A}_{f_j} \right) + \sum_{\ell}^{M} \sum_{k}^{N_\ell} \left(\mathbf{A}_{f_{\ell,k}}^T {\mathbf{R}_{o_\ell}^{f_{\ell,k}}}^{-1} \mathbf{A}_{f_{\ell,k}} \right) \\
& + \sum_{\ell}^{M-1} \left(\mathbf{A}_{o_{\ell+1} o_\ell}^T {\mathbf{Q}_{o_\ell}^{o_{\ell+1} o_\ell}}^{-1} \mathbf{A}_{o_{\ell+1} o_\ell} \right) + \sum_{\ell}^{M} \left(\mathbf{A}_{o_\ell o}^T {\mathbf{R}_{o_\ell}^{o_\ell o}}^{-1} \mathbf{A}_{o_\ell o} \right),
\end{aligned} \tag{3.63}$$

$$\begin{aligned}
\mathbf{y} \;:=\; & \sum_{j}^{N} \left(\mathbf{A}_{f_j}^T {\mathbf{R}_o^{f_j}}^{-1} \mathbf{b}_{f_j} \right) + \sum_{\ell}^{M} \sum_{k}^{N_\ell} \left(\mathbf{A}_{f_{\ell,k}}^T {\mathbf{R}_{o_\ell}^{f_{\ell,k}}}^{-1} \mathbf{b}_{f_{\ell,k}} \right) \\
& + \sum_{\ell}^{M-1} \left(\mathbf{A}_{o_{\ell+1} o_\ell}^T {\mathbf{Q}_{o_\ell}^{o_{\ell+1} o_\ell}}^{-1} \mathbf{b}_{o_{\ell+1} o_\ell} \right) + \sum_{\ell}^{M} \left(\mathbf{A}_{o_\ell o}^T {\mathbf{R}_{o_\ell}^{o_\ell o}}^{-1} \mathbf{b}_{o_\ell o} \right).
\end{aligned} \tag{3.64}$$

This is just a linear system of equations that has a unique solution iff $\det \mathbf{B} \neq 0$. If all local frames are tied together with odometry, this singularity can be avoided with a minimum of three unique, non-collinear features spread anywhere in the entire chain of frames. Without odometry, at least three unique, non-collinear features would be required for each local frame.

Also of interest is that $\mathbf{B}$ represents the inverse covariance matrix of the estimated parameters. Therefore, the variances of the estimates may be approximated from the diagonal of $\mathbf{B}^{-1}$. This is a conservative approximation since the off-diagonal covariance terms are being ignored.

Using the Gauss-Newton algorithm, the solution obtained represents a step to the minimum of the local quadratic approximation, which over a number of iterations will lead to the local minimum of the objective function.

3.4.7 RANSAC

This section relates to box (a) from Figure 3.15. The RANSAC (*Ran*dom *Sa*mple *C*onsensus) algorithm (Fischler & Bolles, 1981) is used to reject poor feature correspondences before a full alignment with MOGA. A poor correspondence is defined as one where the position error between the global-local feature pair does not obey assumed Gaussian uncertainties. This is expressed mathematically in a future section. RANSAC operates only on individual local frames. The first step in RANSAC is to select N_{RANSAC} sets of three randomly-chosen, unique local features. These are then input into MOGA and evaluated to find the best feature correspondences. The procedure is outlined in Figure 3.18 and detailed in the sections to follow.

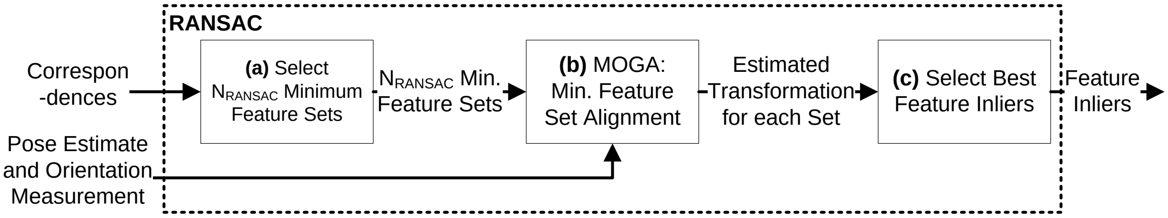


Figure 3.18: RANSAC architecture with odometry measurements for a single local frame.

Selecting Minimum Feature Sets

This section relates to box (a) from Figure 3.18. In three dimensions, at least three unique, non-collinear features are needed to achieve invertibility of $\mathbf{B}$ from Equation (3.63). Choosing three unique features is straightforward, but checking for collinearity is more difficult.

RANSAC repeatedly inputs minimum sets of three features into MOGA. Therefore, it is likely that the three features could become near-collinear at some point in the alignment process. In this manner, it is difficult to choose three features that will never become collinear.

MOGA will thus check that the smallest angle of a group of features never falls below a threshold angle, $E_{\text{collinear}}$. If the angle does fall below this threshold, MOGA fails gracefully and communicates to RANSAC that the features have reached near-collinearity. RANSAC can then proceed to select another set of features.

MOGA for Minimum Feature Sets

This section relates to box (b) from Figure 3.18. The choice of a minimum input of three features allows MOGA to process very quickly, thereby allowing for a thorough search of the best feature correspondences. However, the reason to apply RANSAC on individual local frames is not for efficiency gains.

If odometry were provided, it might be possible for MOGA to tie local frames together and run on all local frames at once. However, this is not implemented since, with only three features, it is expected that potentially large odometry errors would skew the alignment. Therefore, the MOGA block within RANSAC is run on one local frame at a time.

Selecting The Best Feature Inliers

This section relates to box (c) from Figure 3.18. A valid correspondence between a local and global feature is called an inlier. If a given minimum feature set is composed of three inliers, the resulting estimated transformation from MOGA would correctly align other inlier local-global feature pairs. The goal of RANSAC is to find the feature set that produces the most inlier correspondences. These inliers can then be used in a full MOGA run (i.e. box (b) in Figure 3.15). The mathematical definition of an inlier must be developed.

Consider a local feature and its corresponding global feature with respective positions $\mathbf{p}_L := [x_L\ y_L\ z_L]^T$ and $\mathbf{p}_G := [x_G\ y_G\ z_G]^T$. These features are separated by a distance d_{GL}:

$$d_{GL} := \left((x_L - x_G)^2 + (y_L - y_G)^2 + (z_L - z_G)^2\right)^{\frac{1}{2}}. \tag{3.65}$$

In order for this correspondence to be labeled an inlier, d_{GL} must be less than some error threshold E_{inlier}:

$$d_{GL} \leq E_{\text{inlier}} \tag{3.66}$$

where E_{inlier} should relate to the uncertainty in the distance between the two features, $\sigma_{d_{GL}}$:

$$E_{\text{inlier}} := 3\sigma_{d_{GL}}. \tag{3.67}$$

Note that σ is used to represent one standard deviation of uncertainty for the variable in question. With the E_{inlier} threshold set at three standard deviations, 99% of inlier d_{GL} measurements should fall within this threshold and therefore be correctly detected as inliers. It is assumed that other sources of uncertainty are negligible and d_{GL} measurements are normally distributed.

Equation (3.67) must be written in terms of the uncertainty in feature positions. By propagation of uncertainty and assuming zero covariance,

$$\begin{aligned}
\left(\sigma_{d_{GL}}\right)^2 d_{GL}^2 &= \left(\left(\sigma_{x_G}\right)^2 + \left(\sigma_{x_L}\right)^2\right)\left(x_L - x_G\right)^2 \\
&+ \left(\left(\sigma_{y_G}\right)^2 + \left(\sigma_{y_L}\right)^2\right)\left(y_L - y_G\right)^2 \\
&+ \left(\left(\sigma_{z_G}\right)^2 + \left(\sigma_{z_L}\right)^2\right)\left(z_L - z_G\right)^2
\end{aligned} \tag{3.68}$$

where $\sigma_{d_{GL}}$ can be computed since all other variables are known quantities.

3.4.8 MOGA Example

An example is provided in Figure 3.19 to show results after running MOGA, as well as RANSAC's outlier rejection capabilities.

3.5 Methodology Summary

An architecture has been presented to localize a rover over long ranges. The inputs are an orbital map, 3D lidar maps, odometry data, and orientation data. Correspondences between features detected in the orbital and lidar maps, and an initial estimates of the rover pose are determined with DARCES. These quantities, along with odometry and orientation measurements, are then run through RANSAC-MOGA to refine the rover's poses. With a sufficient number of features, this methodology allows for global localization of the rover.

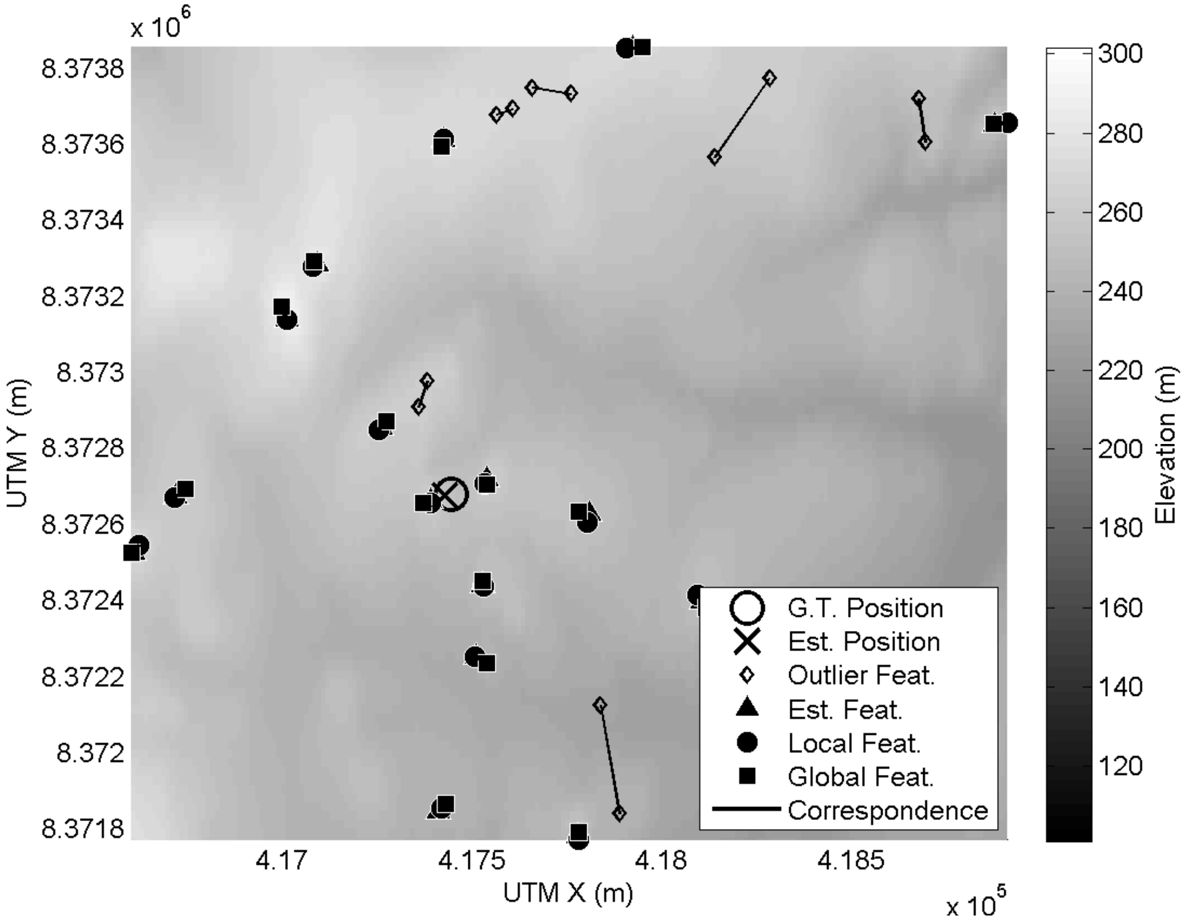


Figure 3.19: Scan A08 features after RANSAC-MOGA. Several outliers were detected and ignored. Estimated feature-landmark positions are also shown. Local features were transformed to the global frame using final MOGA-refined estimates. UTM zone is 16X.

Chapter 4

Experimental Setup

4.1 Global and Local Maps

Field data was collected from Devon Island, Nunavut just North of the Haughton Crater at 75°22'N and 89°41'W. The low-lying terrain and lack of vegetation at this location make it a unique Mars-Moon analogue site. Two critical pieces of information needed for global localization are the local and global maps. The global maps were obtained from GeoBase[1], a repository of DEMs covering all of Canada. The DEM was produced from a stereo image pair collected from orbiting satellite. The segment of the map used for testing was an approximately 10km×10km box as seen in Figure 4.3. Further details on the map resolution and accuracy are presented in Appendix A.

The local maps were constructed from 360° lidar scans with an Optech ILRIS3D-ER lidar mounted on a pan-tilt unit as seen in Figure 4.1(a). The lidar had a maximum range of about 2km in Extended Range mode. Its beam divergence was 0.00974°, which resulted in a range accuracy of 7mm at a range of 100m. The resolution of the scan was set to 0.06° in the horizontal and 0.045° in the vertical. A sample scan is shown in Figure 4.2.

---

[1] Geobase is located online at http://www.geobase.ca

(a) Lidar mounted on pan-tilt unit and tripod. (b) Cart with odometry and heading sensors.

Figure 4.1: Field equipment used to collect Devon Island datasets.

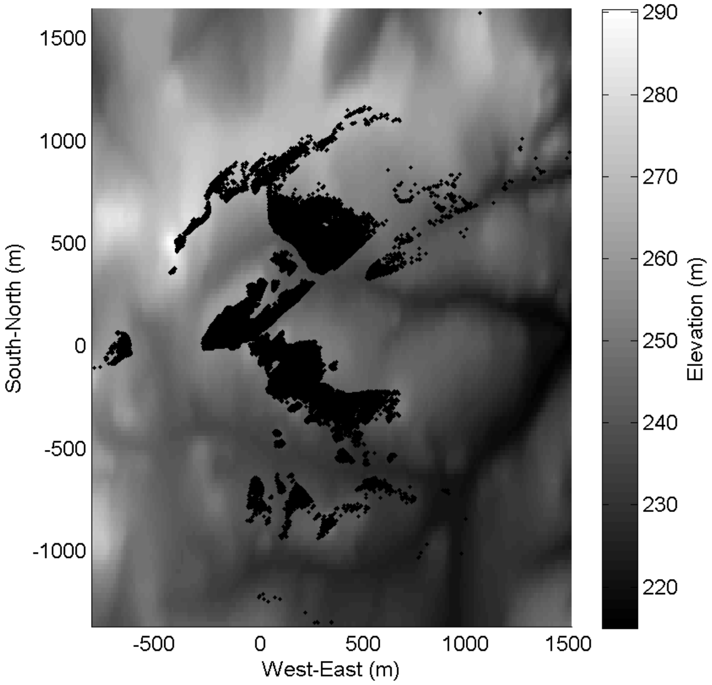


Figure 4.2: Sample scan (A08) overlayed on a section of the orbital DEM. Both maps have been transformed to the same frame of reference using GPS ground-truth. The origins of the lidar scan and of the plot coincide.

4.2 Field Traverses

To simulate a rover's traverse, a cart was outfitted with a suite of rover engineering sensors and collected odometry measurements as it was pushed along a planned path. The cart is shown in Figure 4.1(b). Stops were made along the traverse every 500m to take a $360°$ lidar scan of the terrain.

Two datasets were collected. Dataset A was gathered over a 10km traverse to test the full capabilities of the algorithm. Dataset B was smaller and located about 8km away from Dataset A. It was used to verify the adaptability of the algorithm. The datasets are shown in Figures 4.3 and 4.4.

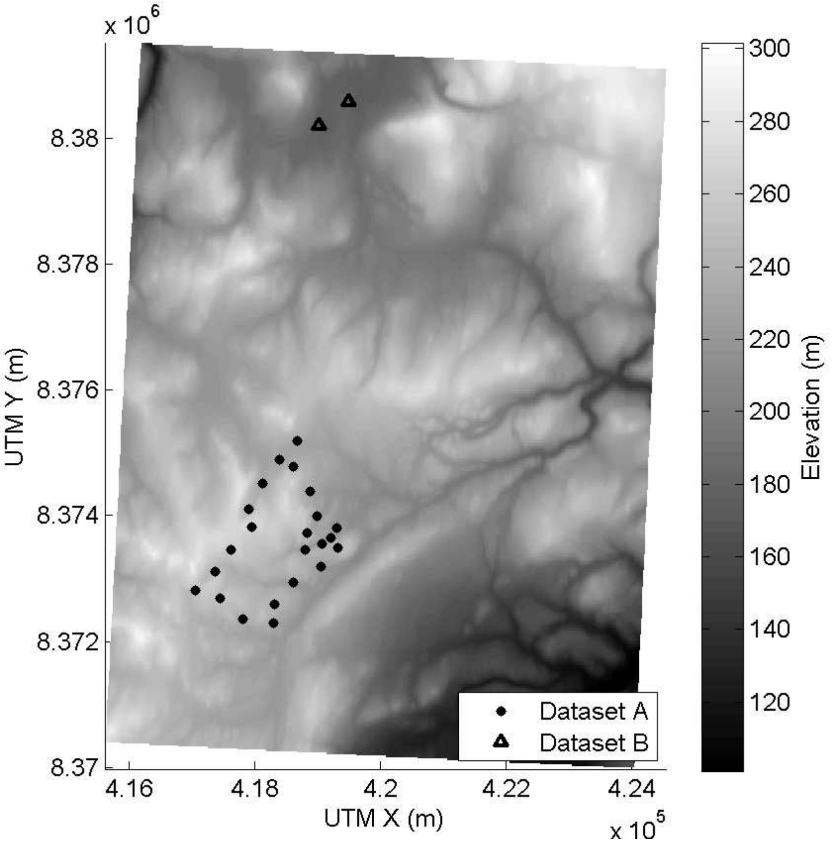


Figure 4.3: Elevation map showing GPS locations of all lidar scans in UTM zone 16X.

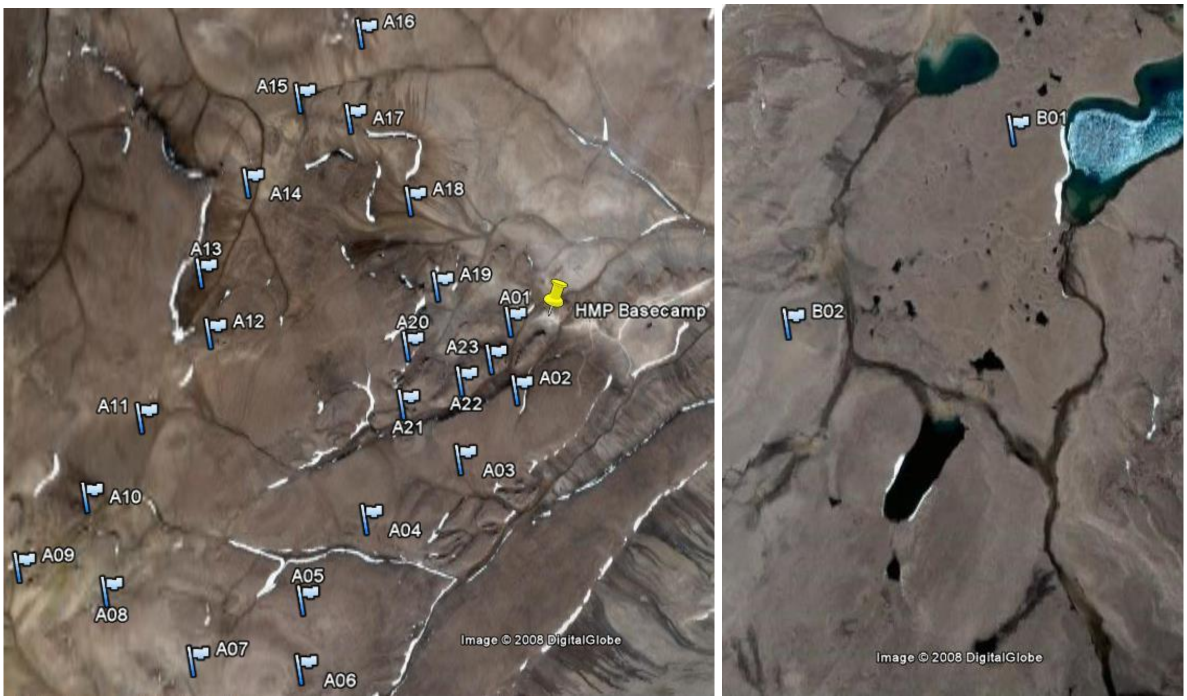


Figure 4.4: Dataset A and B lidar scan positions (GPS) overlayed on orbital images of terrain. Both images are oriented such that the upward direction is approximately North.

4.3 Ground-Truth

Ground-truth x-y position measurements, $\rho_{x,t}$ and $\rho_{y,t}$, were obtained from a consumer-grade handheld GPS. It had positional uncertainty with a standard deviation of $\sigma_{\rho_{x,t}} = \sigma_{\rho_{y,t}} = 5$m. Ground-truth for the z-position was obtained from the global DEM which had an uncertainty with a standard deviation of $\sigma_{\rho_{z,t}} = 12$m. See Appendix A for more detail on DEM accuracies.

Orientation ground-truth for roll and pitch were effectively measured with a two perpendicular bubble levels built into the base of the lidar. Because the lidar was levelled with the bubble levels before each scan, ground-truth roll and pitch were respectively $\alpha_t = \beta_t = 0°$ and had an estimated uncertainty with a standard deviation of $\sigma_{\alpha_t} = \sigma_{\beta_t} = 3°$.

Determining ground-truth for heading, γ_t, was more involved and required knowledge of the position of a second, target reference point. This target setup is shown in Figure 4.5. Before each scan, the target was placed in the lidar's field of view about 100m away such that it did not occlude any important features in the distance. The GPS position of the target was recorded. Since the GPS position of the lidar was also known, it sufficed to determine the position of the target in the lidar's reference frame to calculate the lidar's heading with respect to the global map's reference frame. To obtain this last piece of information, a simple algorithm was created to manually pinpoint the target in the lidar scan. In this manner, the lidar's ground-truth heading could be obtained. This ground-truth heading measurement was estimated to have an uncertainty with a standard deviation of $\sigma_{\gamma_t} = 3°$. For more details on this procedure, refer to Furgale et al. (2009).

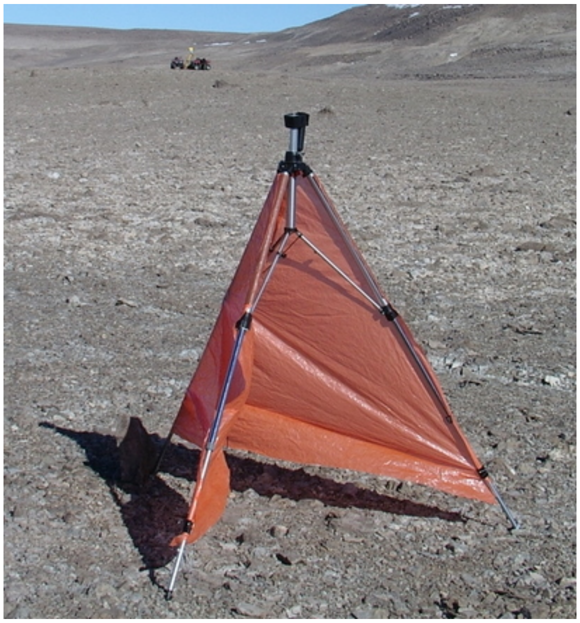


Figure 4.5: Target used to obtain heading ground-truth measurements.

4.4 Odometry Measurements

The cart was equipped with wheel odometry and stereo cameras, which together could produce visually-corrected odometry. These odometry measurements could be used as an additional input to the localization algorithm developed in this thesis in order to refine the alignment between global and local maps. Unfortunately, the processed odometry data was not available in the time-frame of this thesis. However, knowing ground-truth position and orientation at each scan stop, the odometry could be simulated with the assumption that odometry error grew as a function of the distance traversed.

Odometry measurements are simulated using ground-truth transformations between scans and adding error by sampling from a zero-mean normal distribution with covariance matrix, $\mathbf{Q}$. In creating $\mathbf{Q}$, the translational standard deviations in uncertainty are chosen to increase as a proportion, c_ρ, of their respective traversal distance components, ρ_x, ρ_y, ρ_z. Rotational standard deviations in uncertainty grow as a function of the overall radial distance traversed, $\rho := \sqrt{\rho_x^2 + \rho_y^2 + \rho_z^2}$, and at a rate of c_θ radians per metre. The resulting 6×6 covariance matrix is therefore

$$\mathbf{Q} = \operatorname{diag}\left\{(c_\rho \rho_x)^2, (c_\rho \rho_y)^2, (c_\rho \rho_z)^2, (c_\theta \rho)^2, (c_\theta \rho)^2, (c_\theta \rho)^2\right\} \tag{4.1}$$

where the diagonal variables from top left to bottom right are respectively the x, y, z measurement variances of distance traversed and the x, y, z measurement variances of angle displaced.

4.5 Orientation Measurements

Orientation measurements could also be input into the localization algorithm developed in this thesis to improve performance. In a realistic scenario, roll and pitch measurements would be obtained from an inclinometer and heading measurements from a sun sensor or equivalent. However, an inclinometer was not available at the time of experimentation and although the cart was equipped with a sun sensor, the data had not yet been processed for use in this thesis. Therefore, it was necessary to simulate these sensor orientation measurements using ground-truth measurements.

It should be clear that comparing orientation estimates output from the algorithm to ground-truth orientation measurements is merely observing the changes the algorithm made to the ground-truth. However, some useful observations can still be made, such as the effect of these orientation measurements when used as filters in the feature-matching algorithm. This is further discussed in the sections to follow.

Chapter 5

Results

5.1 Overview

Three configurations of the architecture were tested:

1. Single-frame alignment with no heading measurements.
2. Single-frame alignment with heading measurements.
3. Multiple-frame alignment with heading measurements.

The single-frame case only aligned features from a single lidar scan to the global map and assumed no odometry measurements were available. The multiple-frame scenario included simulated odometry measurements to refine the alignment between lidar scans, and to obtain transformation estimates for frames with no DARCES solution. The benefits of inputting heading measurements were also examined.

Results were obtained for Datasets A and B. Dataset A was tested on all three architecture configurations mentioned previously. Dataset B was only run on the single-frame cases, since it was only needed to verify that the architecture worked in different regions. The algorithm was coded in Matlab R2007a and run a system with an Intel Core2 2GHz CPU and 2GB RAM.

Performance of the algorithm is often presented as ρ_x, ρ_y, ρ_z translation and θ_x, θ_y, θ_z axis-angle rotation errors with respect to ground-truth. These axis-angle rotation errors are also respectively referred to as pitch, roll and heading errors.

A summary of the chosen input parameters is presented in Table 5.1. Justification for the selection of these values is explained in the sections to follow.

Table 5.1: Summary of input parameters. All uncertainties correspond to one standard deviation.

Parameter	**Value**	**Description**
L_{global}	12m	Global map resolution.
n	5 cells	Dilation window radius.
D_{detect}	60m	Minimum distance between detected features.
$\sigma_{r_{G_{xy}}}$	14m	Global map x-y uncertainty.
$\sigma_{r_{G_z}}$	12m	Global map z uncertainty.
$\sigma_{r_{L_{xy}}}$	5m	Local map x-y uncertainty.
$\sigma_{r_{L_z}}$	5m	Local map z uncertainty.
t	60m	Half-thickness of hypothesis search sphere shell.
N_{DARCES}	500	Maximum number of control point iterations in DARCES.
E_{zdev}	100m	z-deviation filter threshold.
$E_\alpha, E_\beta, E_\gamma$	9°	Orientation filter thresholds.
c_ρ	5%	Odometry position error growth rate.
c_θ	$5\frac{\circ}{\text{km}}$	Odometry orientation error growth rate.
N_{RANSAC}	500	Maximum number of RANSAC iterations.
R_{orient}	$\left(3^\circ \frac{\pi}{180}\right)^2 \mathbf{1}$	Orientation measurement covariance matrix.
E_{inlier}	45m	Feature inlier threshold for RANSAC.
$E_{\text{collinear}}$	1°	Collinearity threshold.
E_{MOGA}	10^{-16}	MOGA convergence threshold.
η	0.9	Update reduction factor for MOGA line search.

5.2 Feature Detection

The global map covered an area of about $10\text{km}\times 10\text{km}$ on Devon Island, Nunavut. It was located in the UTM 16X zone and bounded by $75.40°$N, $75.50°$N, $90.00°$W and $89.70°$W. Some minor operations were performed on the raw DEM resolution and accuracy to make them compatible with the algorithm (see Appendix A).

The DEM's x and y resolutions were not equal. In order to make use of the extra information provided by the smaller resolution, the global map was interpolated at the smaller resolution to form a uniform grid. Therefore, the global map resolution was $L_{\text{global}} = 12\text{m}$.

The dilation window used to extract features from global and local maps was a pixelated circle with a radius of $n = 5$ global map cells. Therefore, as discussed in Section 3.2, the minimum distance allowed between features was $D_{\text{detect}} = n \times L_{\text{global}} = 60\text{m}$.

With these parameters, 537 global map features were detected in approximately 5 seconds. A typical local frame detection cycle finished in about 25 seconds. Feature detection in local frames was slower mostly due to the final processing steps that improved the chances that a detected feature would be a peak in the local topology.

The distribution of features for each local frame is shown in Figure 5.1. Fewer local features were detected when the lidar was located in a canyon setting such as the first and last few frames of Dataset A. In a canyon, the lidar was surrounded by nearby hills, which occluded long-range terrain features. Even a detected peak may have been a poor representation of the true peak, which was likely to be occluded at close range. A low number of features in a frame tended to make it more difficult to find a DARCES solution.

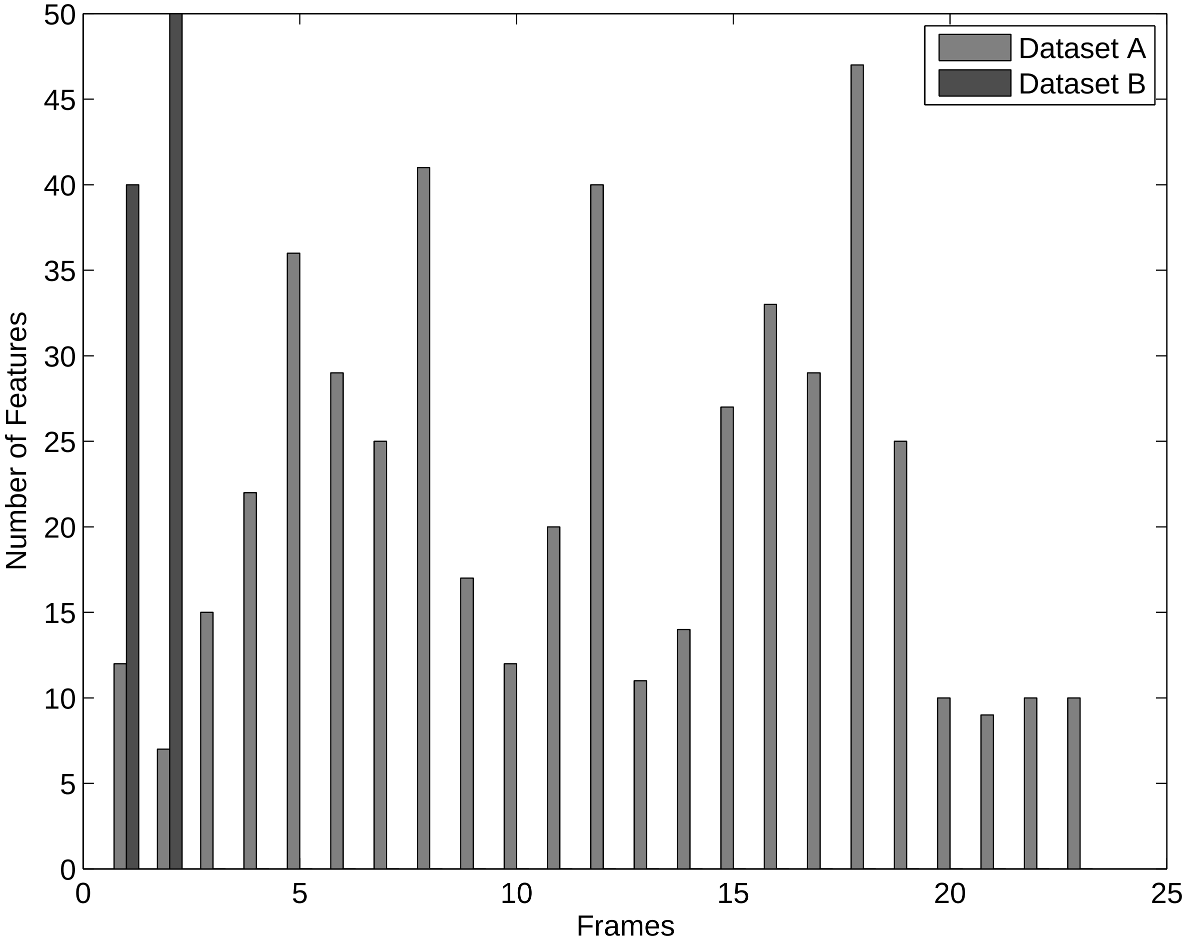


Figure 5.1: Number of features detected in each frame for Dataset A and B.

5.3 Feature Matching

The essential parameter in the search for correspondence hypotheses was the half-thickness of the search sphere's shell, t. As derived in Equation (3.2), the half-thickness depended on the uncertainties in global and local features.

The uncertainties in global feature positions were assumed equal to the uncertainties in the orbital DEM measurements: $\sigma_{r_{G_{xy}}} = 14\text{m}$ and $\sigma_{r_{G_z}} = 12\text{m}$. These DEM uncertainties are derived in Appendix A. To simplify the local feature error model, the local features' position uncertainties were set to $\sigma_{r_{L_{xy}}} = 5\text{m}$ and $\sigma_{r_{L_z}} = 5\text{m}$ based on a conservative estimate of the lidar's accuracy in scanning the position of a topological peak. Therefore, as derived in

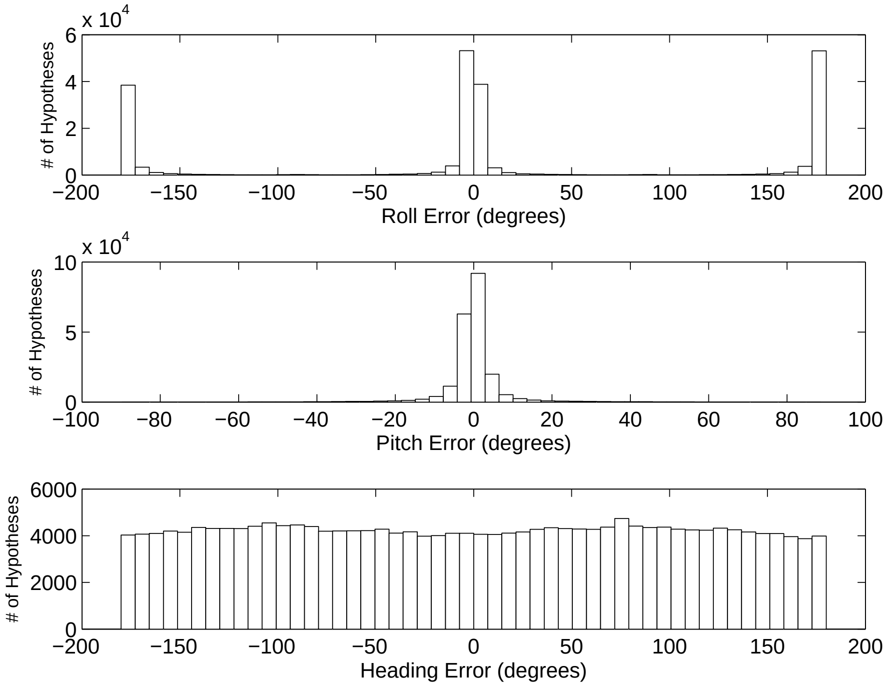


Figure 5.2: Orientation errors for all unfiltered hypotheses from a trial run on scan A08.

Appendix B, a reasonable value for the half thickness of the search sphere's shell is $t \approx 60\text{m}$.

Each DARCES trial was limited to run over a maximum of $N_{\text{DARCES}} = 500$ control points. A number of filters were used to improve computing time and refine the initial set of hypotheses in search of valid hypotheses. The maximum z-deviation threshold was set to $E_{\text{zdev}} = 100\text{m}$. The orientation filters were designed such that errors between measured and hypothesized angles remained three standard deviations below measurement uncertainty. The thresholds were therefore $E_\alpha = 9^\circ$, $E_\beta = 9^\circ$, $E_\gamma = 9^\circ$.

Computing time was approximately one minute per local frame with these settings. Figure 5.2 shows the orientation errors used in filtering the hypotheses. It is interesting to note the distributions of the three orientation errors. Hypotheses tended to favour orientations parallel to the global map x-y plane because of the low-lying nature of the terrain. For example, roll errors tended to be either near-zero (upright) or 180° (upsidedown). Hypotheses producing an upsidedown orientation could therefore easily be filtered. The impact of these filters on the number of hypotheses is significant and can be seen in Figure 5.3.

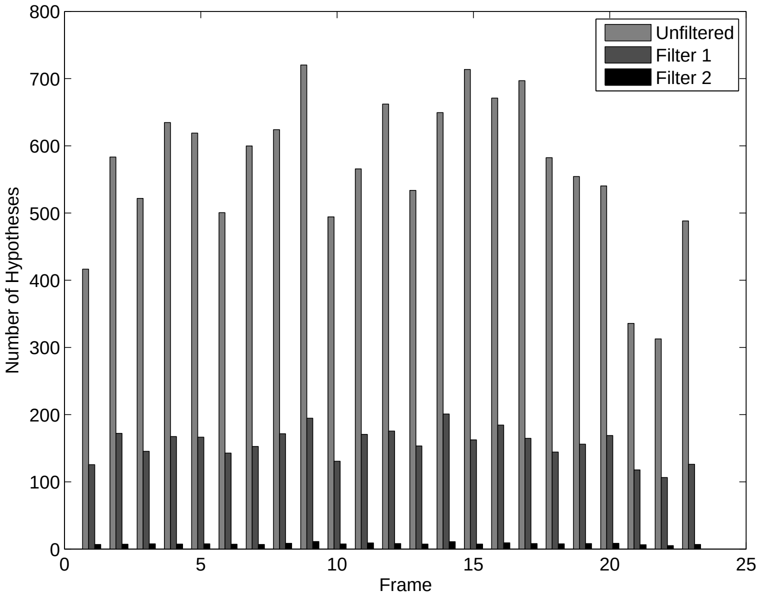


Figure 5.3: Average number of hypotheses over random sample of control points from Dataset A. Shown for three stages of RANSAC. Unfiltered: Hypotheses initially generated. Filter 1: Hypotheses remaining after roll, pitch, z-deviation and map boundary filters. Filter 2: Hypotheses remaining after heading filter.

5.4 Pose Refinement

As explained in Section 4.4, odometry measurements were simulated using ground-truth transformations between scans and adding error by sampling from a zero-mean normal distribution with covariance matrix, $\mathbf{Q}$. Chosen based on typical visual odometry performance, the simulated odometry's translational and rotational error growth rates were respectively $c_\rho = 5\%$ of component-wise distance traversed, and $c_\theta = 5\frac{\circ}{\text{km}}$ of radial distance traversed.

Orientation measurement covariance matrices were estimated from instrumentation uncertainties, $R_{\text{orient}} = \left(3^\circ \frac{\pi}{180}\right)^2 \mathbf{1}$.

RANSAC was run for $N_{\text{RANSAC}} = 500$ iterations. If there was an insufficient number of features to reach N_{RANSAC} iterations, then all possible feature combinations were tested with the realization that the results may not have been as reliable. In the search for inliers, the maximum position error allowed between a pair of corresponding features was E_{inlier}. This was calculated from Equation (3.67) using position measurement values and their uncertainties. A typical value for the inlier threshold was $E_{\text{inlier}} \approx 45\text{m}$.

Global and local features in MOGA were checked for collinearity since this was a particularly likely scenario when fitting the minimum three points in RANSAC. If at any time the smallest angle between the three features was less than $E_{\text{collinear}} = 1^\circ$, the points were considered collinear, forcing RANSAC to attempt a new set of features.

Transformation estimates were obtained from DARCES, which served as initial guesses for MOGA. However, if a particular frame did not have a DARCES solution, its transformation was estimated using odometry to the next-closest DARCES-solved frame along the path.

MOGA was considered to have converged when the difference between the previous and current objective function errors, $\Delta J := J_{i-1} - J_i$, was less than the threshold $E_{\text{MOGA}} = 10^{-16}$. A line search was done on the design parameter updates, $\mathbf{z}$, to ensure the objective function decreased after each iteration. If the initial update produced an increase in the objective, $\Delta J > 0$, then $\mathbf{z}$ was repeatedly reduced by a factor of $\eta = 0.9$ until a decrease, $\Delta J < 0$, was obtained.

With these settings, a full RANSAC-MOGA run on Dataset A took about 5 minutes to complete. A full, 23-frame MOGA run on Dataset A converged in about 10 seconds. Figure 5.4 shows a sample convergence plot for a multi-frame trial over all frames in Dataset A.

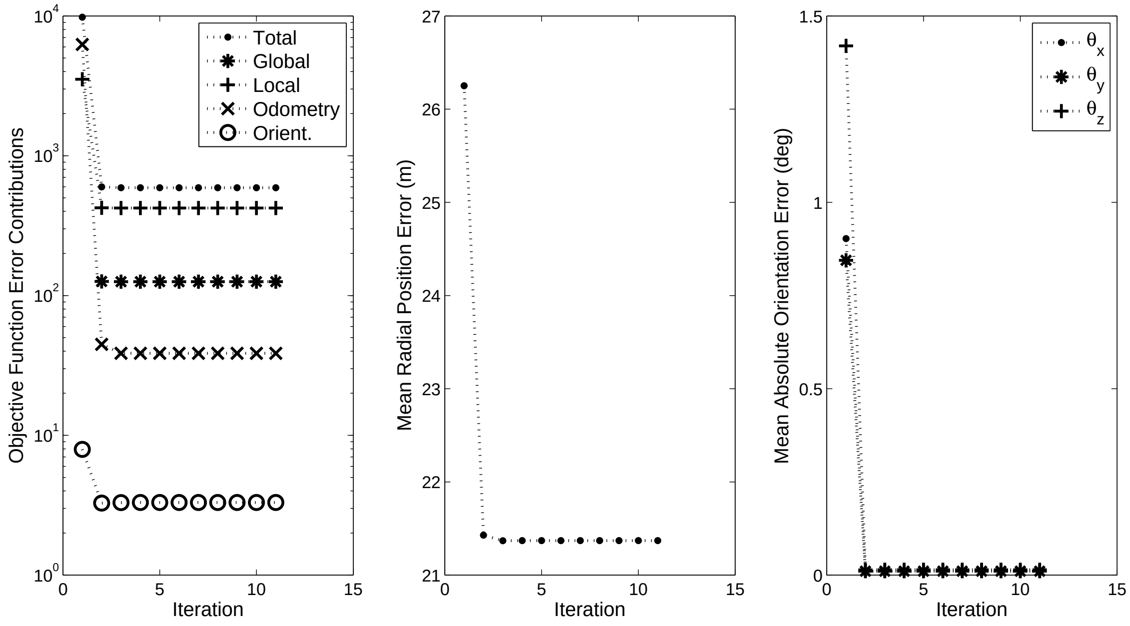


Figure 5.4: Convergence plots for a multi-frame trial run on all frames of Dataset A.

5.5 Single-Frame Localization No Heading

In the single-frame no-heading trials, inclinometer information was available along with local and global maps. Heading data from the sun sensor was not used. Localization was only executed on individual frames. There was no information sharing between frames. This scenario was tested over 30 trials for each of Datasets A and B. In each trial, all inputs were the same except the control point sets in DARCES and the minimum feature sets in RANSAC, which were randomly selected.

5.5.1 Dataset A, Single No Heading

Performance was encouraging, since most trials showed error well below 100m position and 5° heading error as shown in Figure 5.5. The number of trials of frames having one or more estimates falling outside their calculated uncertainties was of slight concern as seen in the error plots of Figures 5.6 and 5.7. Furthermore, calculated uncertainties in the estimates seemed larger than they perhaps should have been. This was possibly due to overestimation of input measurement uncertainties.

These trials clearly suffered from the lack of heading information. In particular, frame A20 demonstrated the risk in localizing over such a large initial search space. The estimated position and heading for this particular frame had an error of about 2km and 180° respectively. A heading filter could have easily caught this poor hypothesis, as will be shown in future sections.

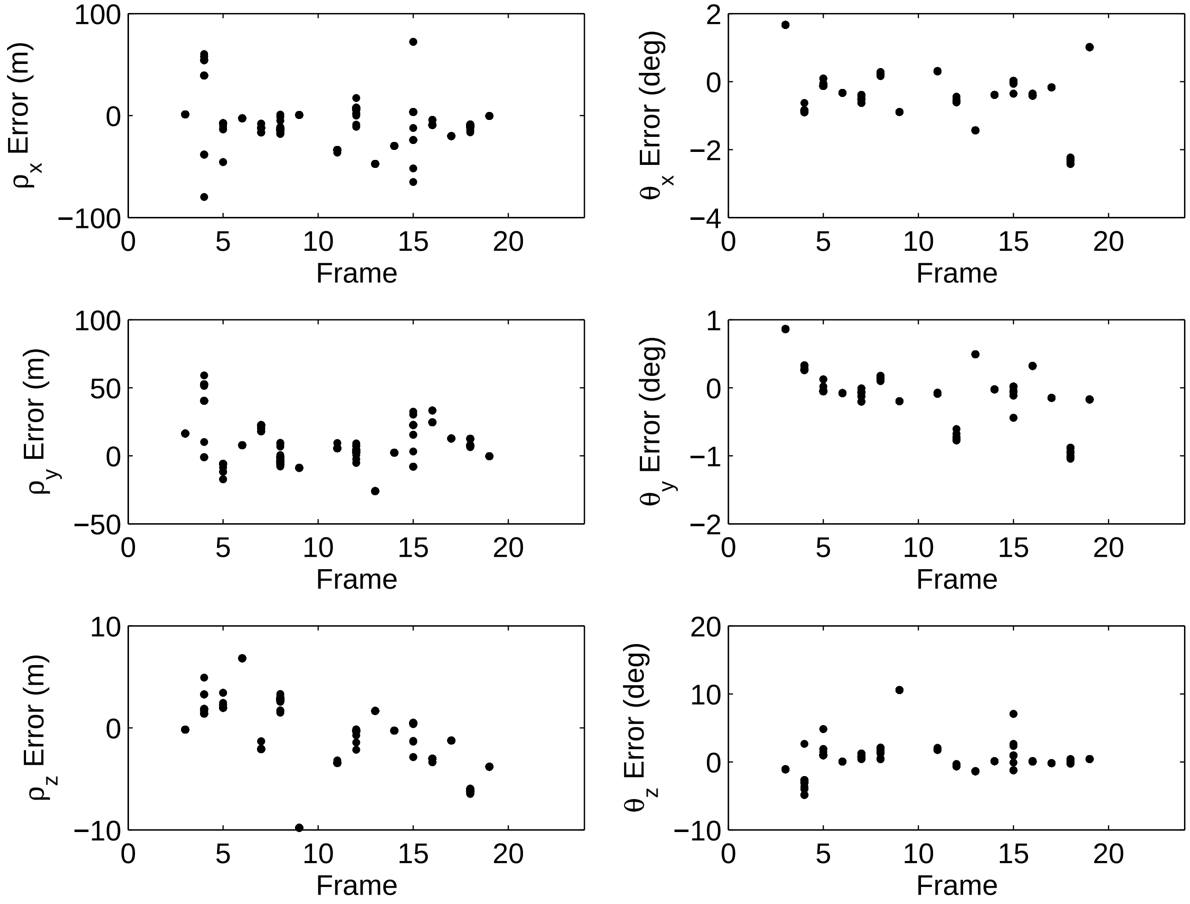


Figure 5.5: Dataset A, no heading input. Translation and rotation errors for all 30 trials. Frame A20 was omitted for clarity since it produced estimates with large errors (e.g. position error on the order of kilometres). The absence of data for a given frame indicates that DARCES could not find a solution for that frame in any of the 30 trials. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

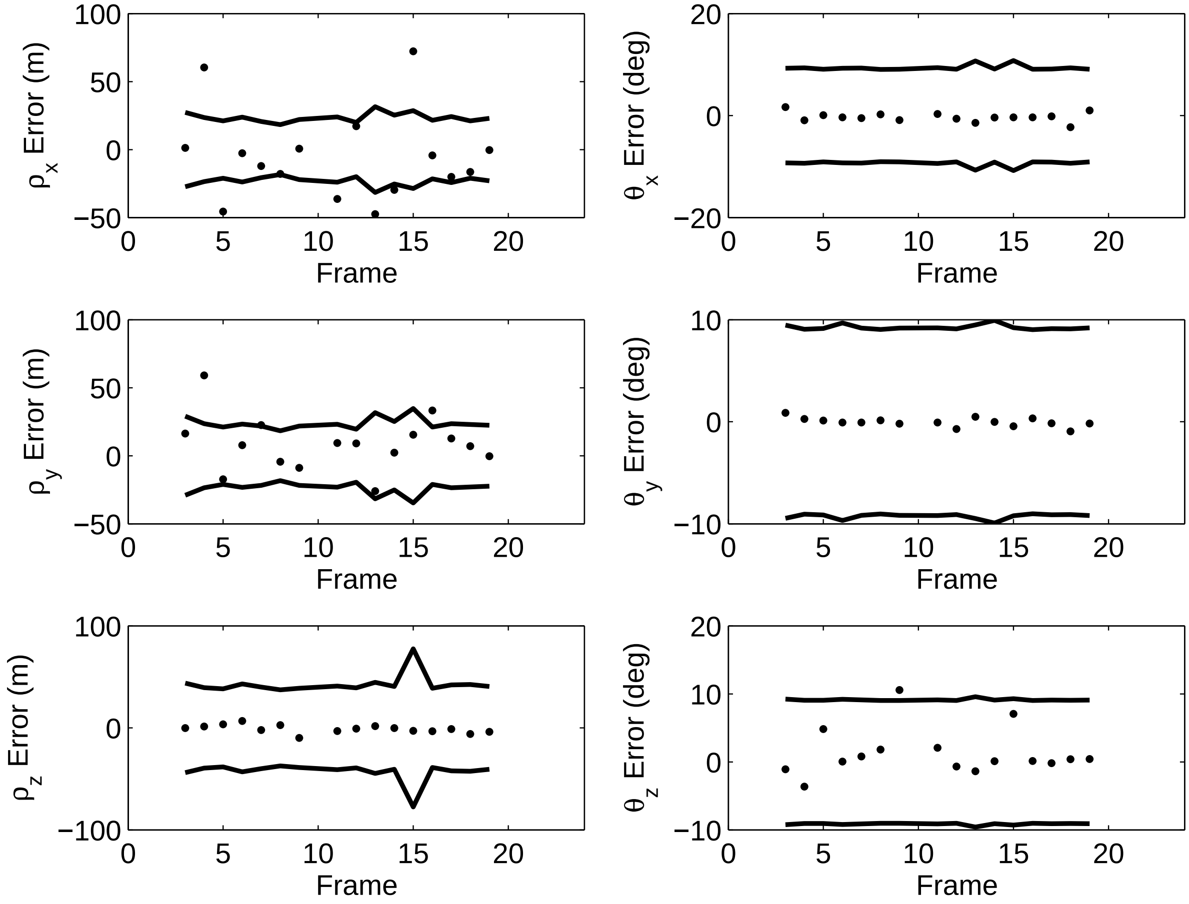


Figure 5.6: Dataset A, no heading input. Errors for the **worst** of all 30 trials (i.e. trials with the worst radial translation error for a given frame). Frame A20 was omitted for clarity since it produced estimates with large errors (e.g. position error on the order of kilometres). The absence of data for a given frame indicates that DARCES could not find a solution for that frame in any of the 30 trials. Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate and ground-truth uncertainty. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

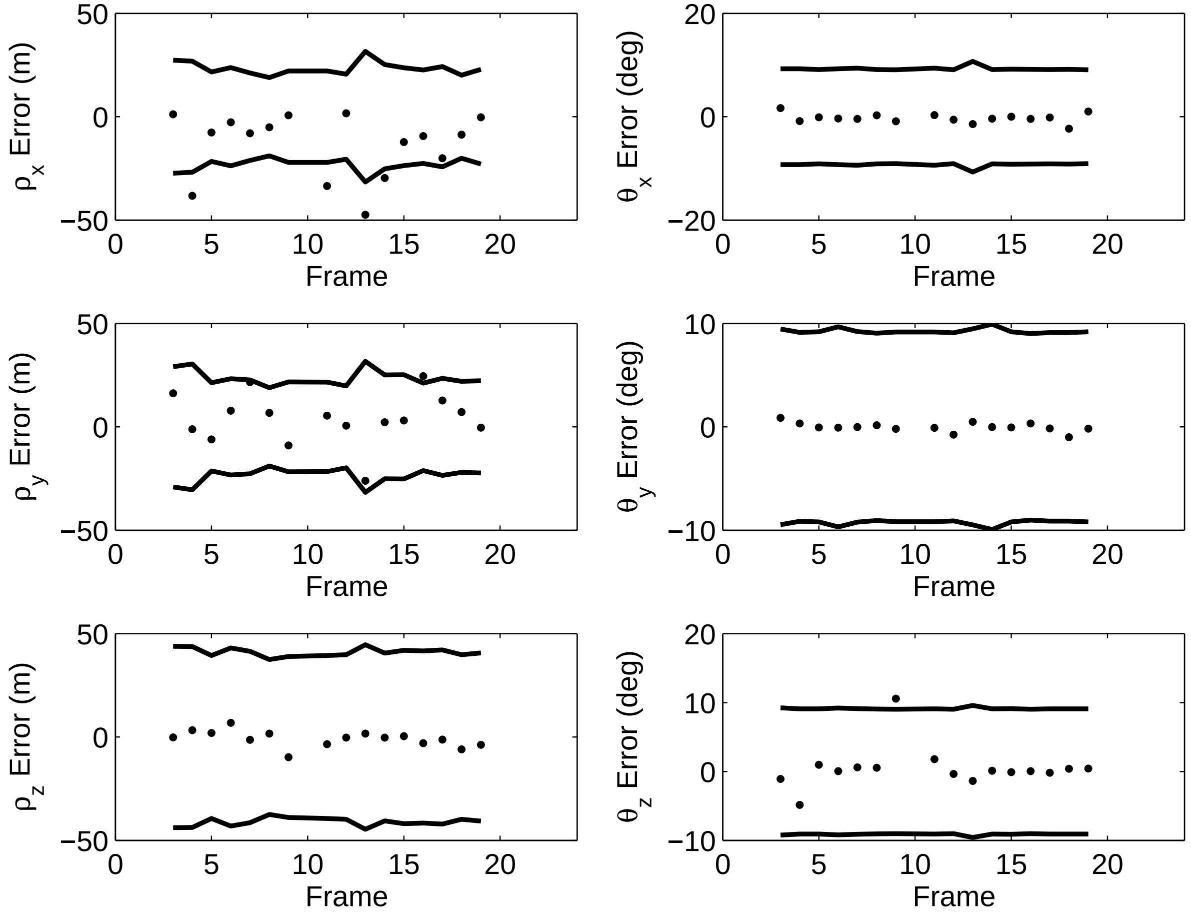


Figure 5.7: Dataset A, no heading input. Errors for the **best** of all 30 trials (i.e. trials with the best radial translation error for a given frame). Frame A20 was omitted for clarity since it produced estimates with large errors (e.g. position error on the order of kilometres). The absence of data for a given frame indicates that DARCES could not find a solution for that frame in any of the 30 trials. Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate and ground-truth uncertainty. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

5.5.2 Dataset B, Single No Heading

Dataset B gave results similar to Dataset A for the single-frame no heading configuration. A solution was obtained for frame B02 only 6 out of 30 times. This was a good reminder of the random nature of DARCES, and was justification for running as many control point sets through DARCES as possible. The otherwise good results presented in Figures 5.8, 5.9 and 5.10 helped verify that the algorithm was robust to different environments, since the exact same input parameters were used for tests on both Datasets A and B, and Dataset B was geographically separated from Dataset A.

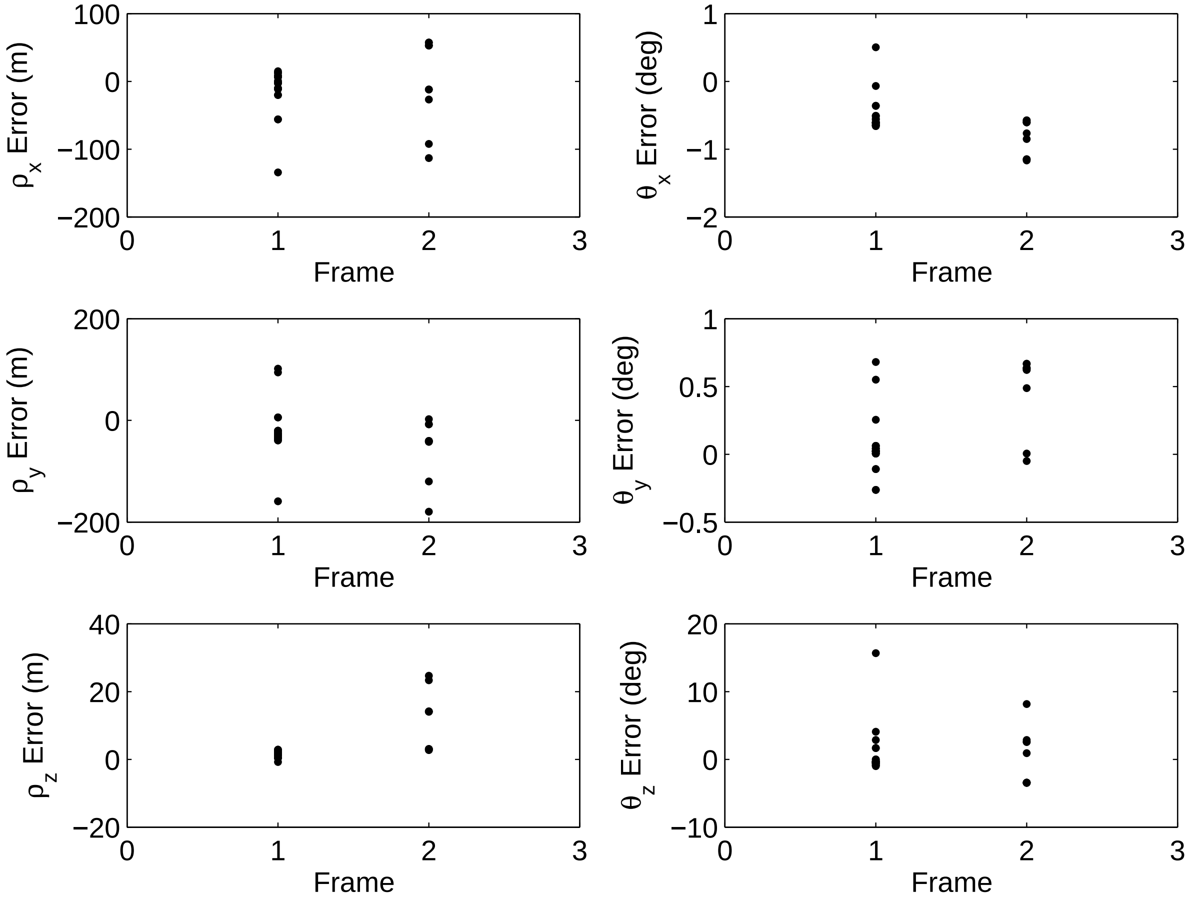


Figure 5.8: Dataset B, no heading input. Translation and rotation errors for all 30 trials. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

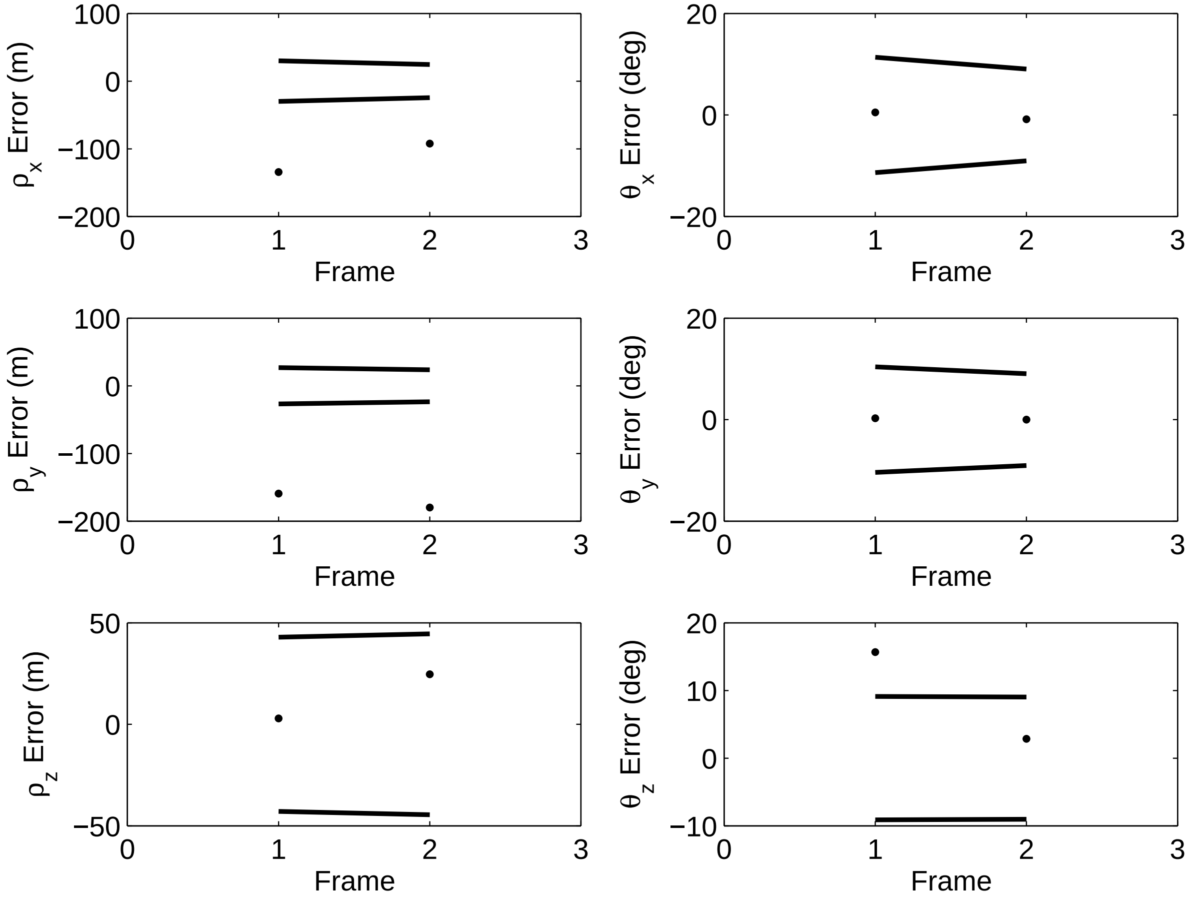


Figure 5.9: Dataset B, no heading input. Errors for the **worst** of all 30 trials (i.e. trials with the worst radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

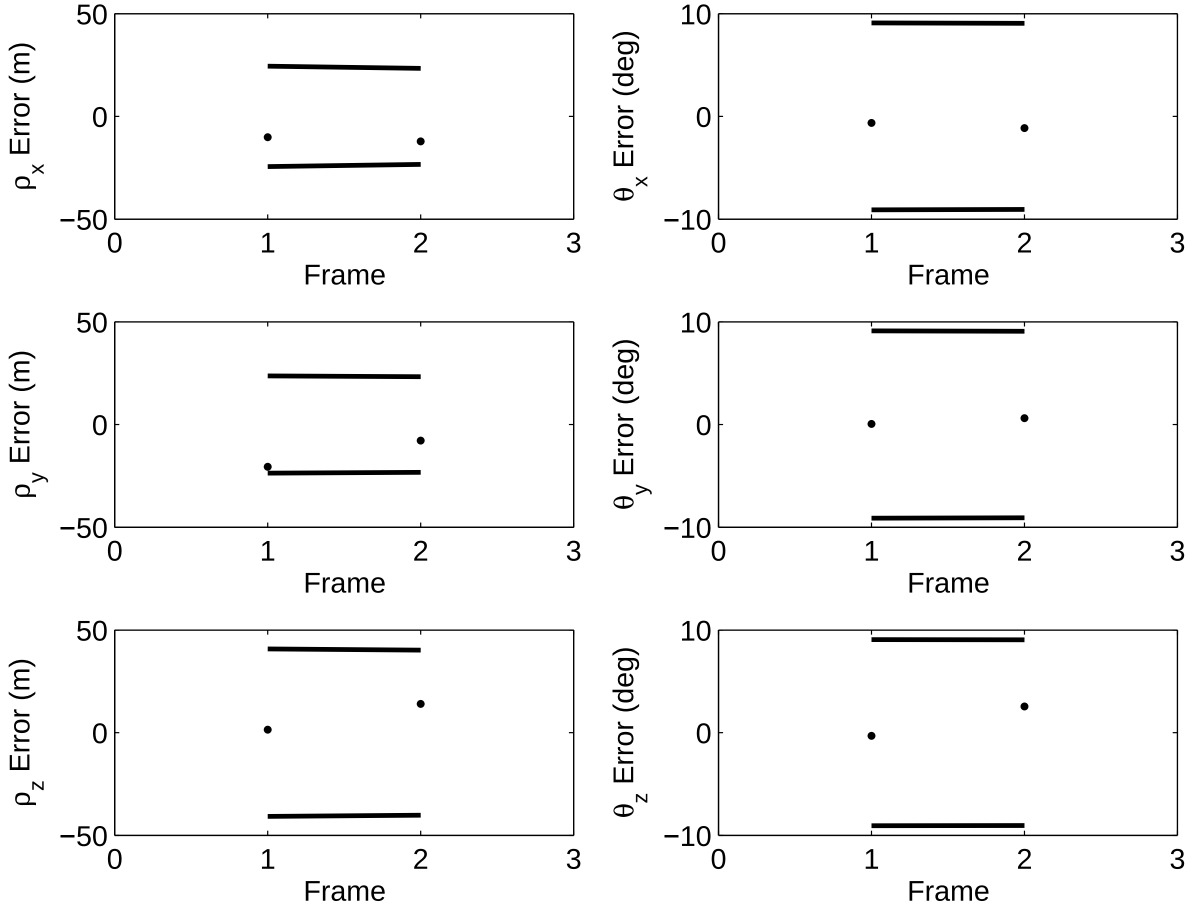


Figure 5.10: Dataset B, no heading input. Errors for the **best** of all 30 trials (i.e. trials with the best radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. Orientation errors, θ_x and θ_y, only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

5.6 Single-Frame Localization With Heading

In this configuration, heading measurements (i.e. simulated sun sensor measurements) were available as an input to the algorithm. Localization was still performed only on individual frames. Results were obtained for 30 trials over each of Datasets A and B. In each trial, all inputs were the same except the control point sets in DARCES and the minimum feature sets in RANSAC, which were randomly selected. The exact same control points were used in these trials as were used in the single-frame no heading case.

5.6.1 Dataset A, Single With Heading

The benefits of the input heading measurements can be seen immediately. As expected, the poor solutions obtained from the previous section in frames A09 and A20 were detected by the heading filter and discarded as shown in Figure 5.11. Accuracies of the estimates were otherwise largely the same as in the single-frame no heading case. The same estimates fell outside uncertainty, as seen in Figures 5.12 and 5.13.

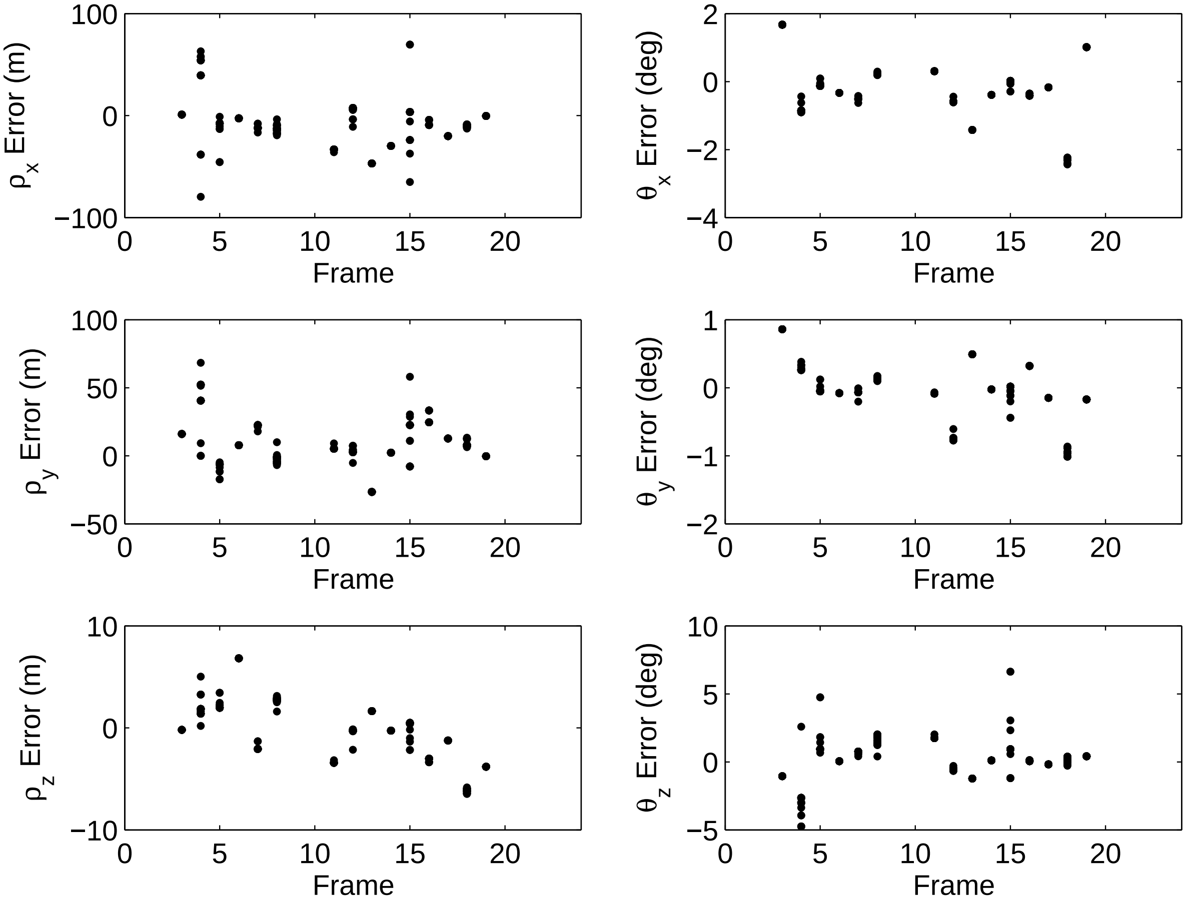


Figure 5.11: Dataset A, with heading input. Translation and rotation errors for all 30 trials. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

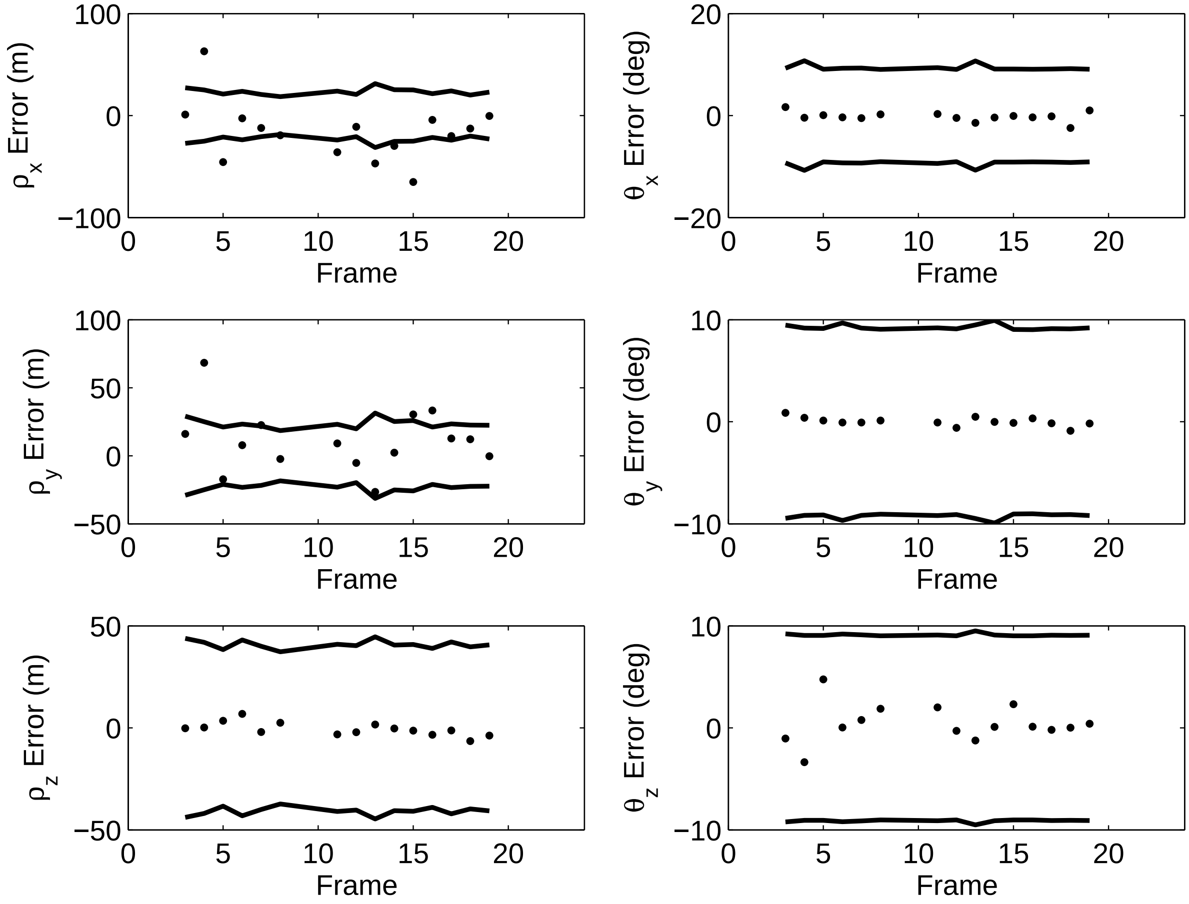


Figure 5.12: Dataset A, with heading input. Errors for the **worst** of all 30 trials (i.e. trials with the worst radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

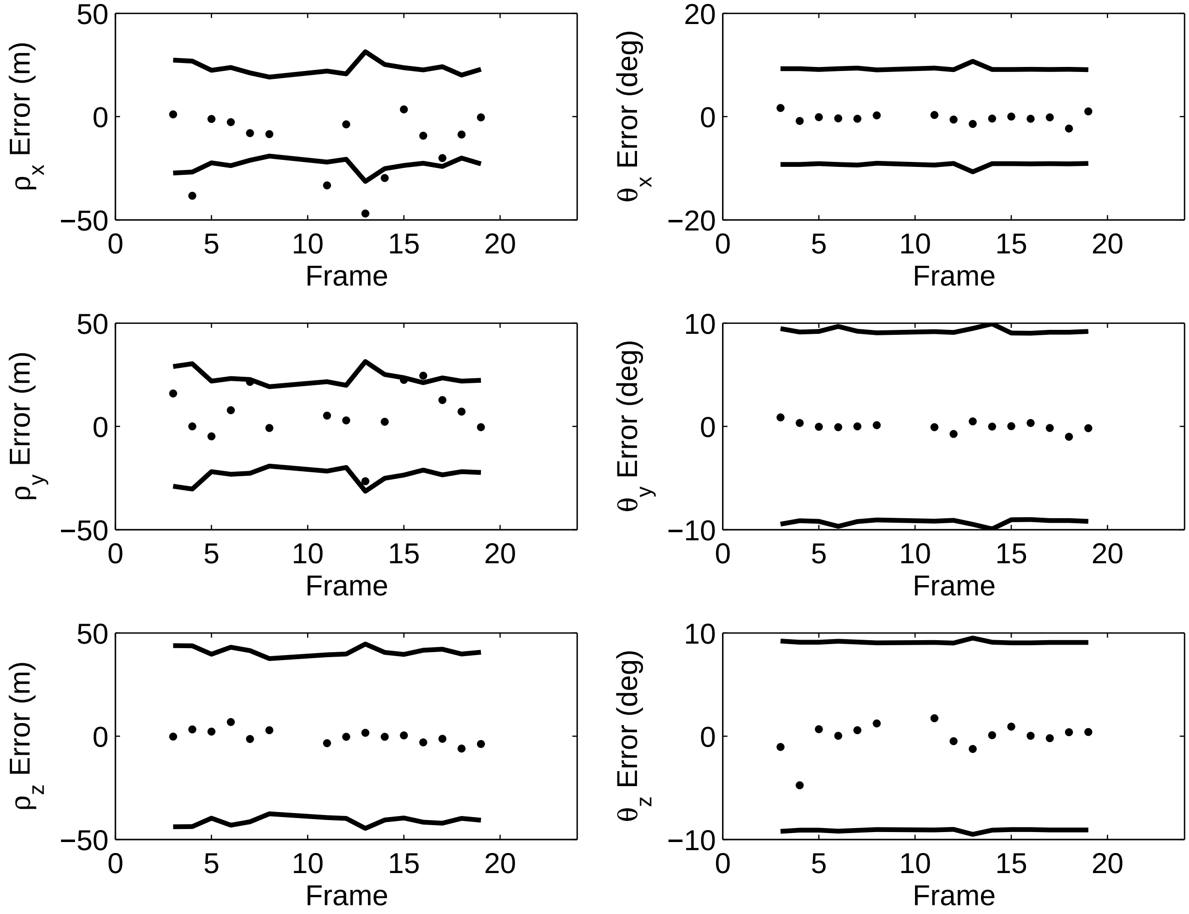


Figure 5.13: Dataset A, with heading input. Errors for the **best** of all 30 trials (i.e. trials with the best radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

5.6.2 Dataset B, Single With Heading

As with Dataset A, the heading information allowed some poor solutions to be detected early and eliminated, as seen in Figure 5.14. The poorest estimates were again observed to fall outside their estimated uncertainties, as shown in Figures 5.15 and 5.16. Otherwise, results were similarly positive. This helped again prove the ability of the algorithm to function in different terrains.

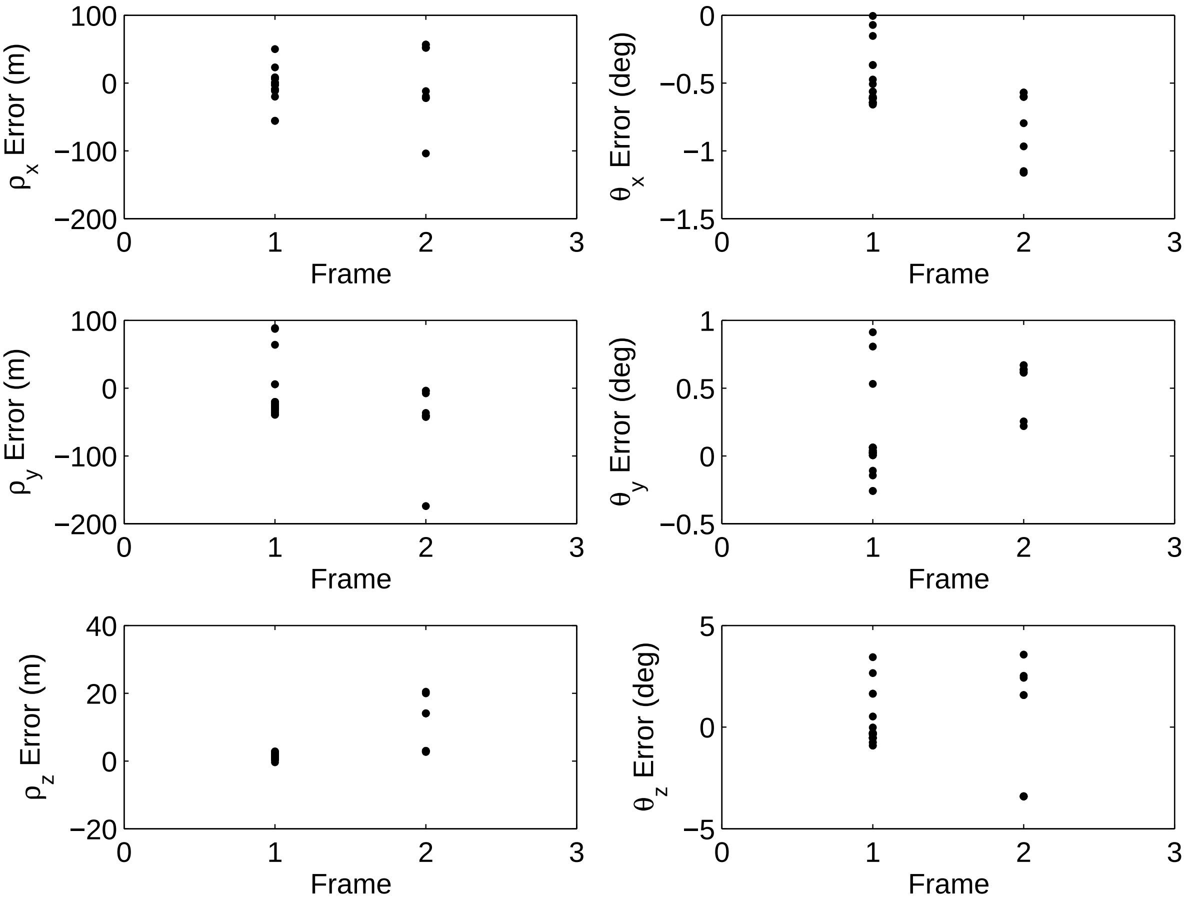


Figure 5.14: Dataset B, with heading input. Translation and rotation errors for all 30 trials. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

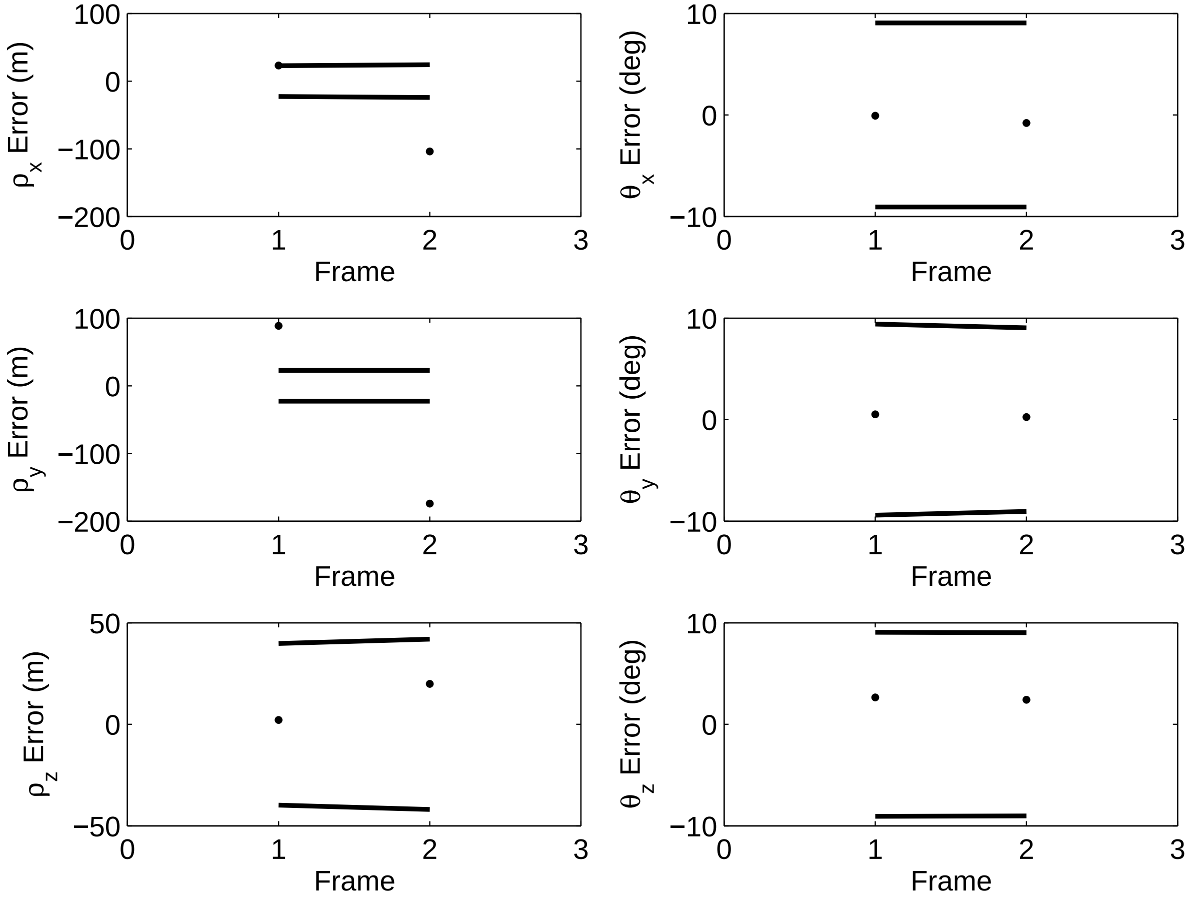


Figure 5.15: Dataset B, with heading input. Errors for the **worst** of all 30 trials (i.e. trials with the worst radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

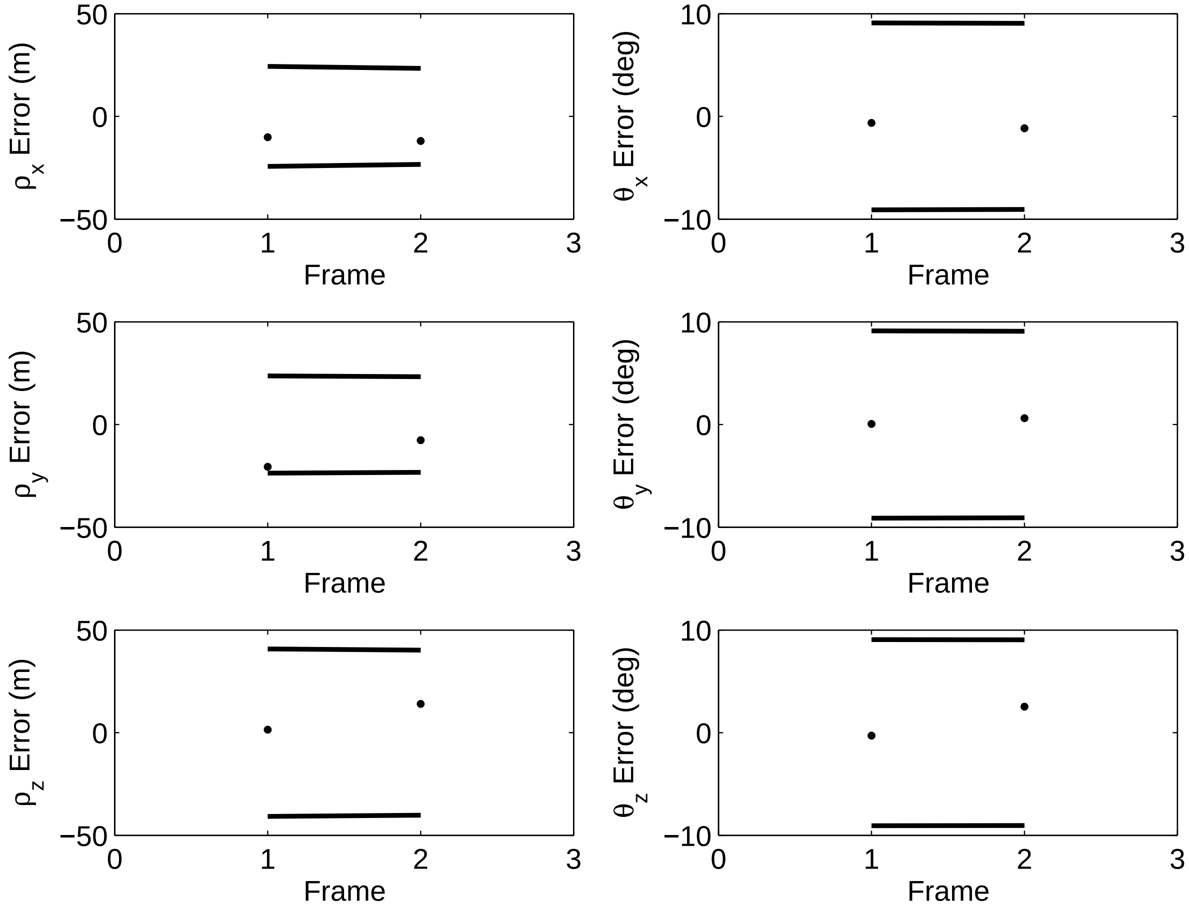


Figure 5.16: Dataset B, with heading input. Errors for the **best** of all 30 trials (i.e. trials with the best radial translation error for a given frame). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

5.7 Multiple-Frame Localization With Heading

In this configuration, odometry and heading measurements were assumed available. Results were obtained for Dataset A over 30 trials in which the simulated odometry was different for each trial. The exact same DARCES correspondences were used as in the single-frame with heading case. A trial finished in about 30 minutes from feature detection to pose refinement.

In this final configuration, estimates for frames with no DARCES solution could now be obtained, as shown in Figure 5.17. Furthermore, more estimates now fell within the calculated uncertainties, as seen in Figure 5.18 and 5.19. However, the ρ_z uncertainties showed growing uncertainties towards the end of the traverse. This was likely a result of the dependence of the final three frames on odometry alone. However, it was unexpected that the growth in uncertainty would propagate to nearby frames such as A19 with DARCES solutions

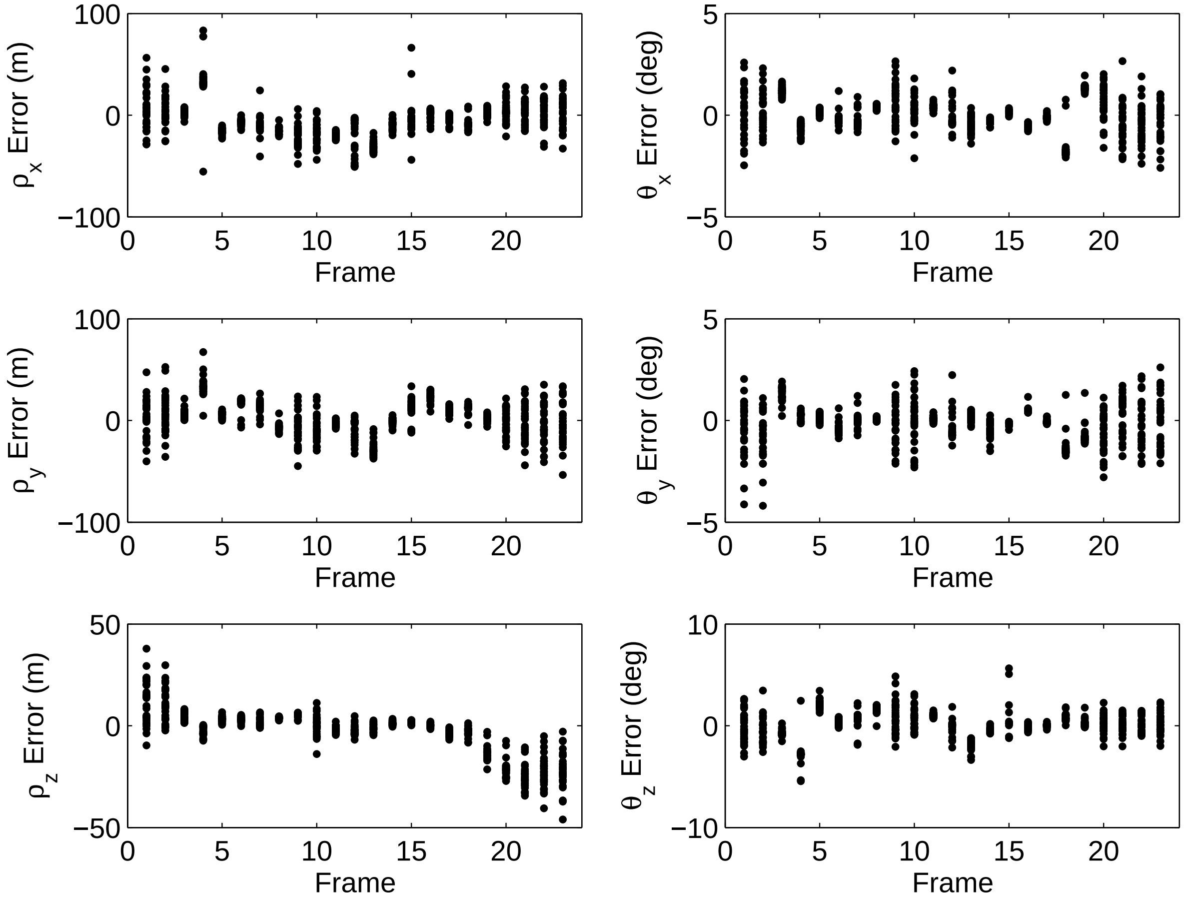


Figure 5.17: Dataset A, multi-frame trials. Translation and rotation errors for all 30 trials. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

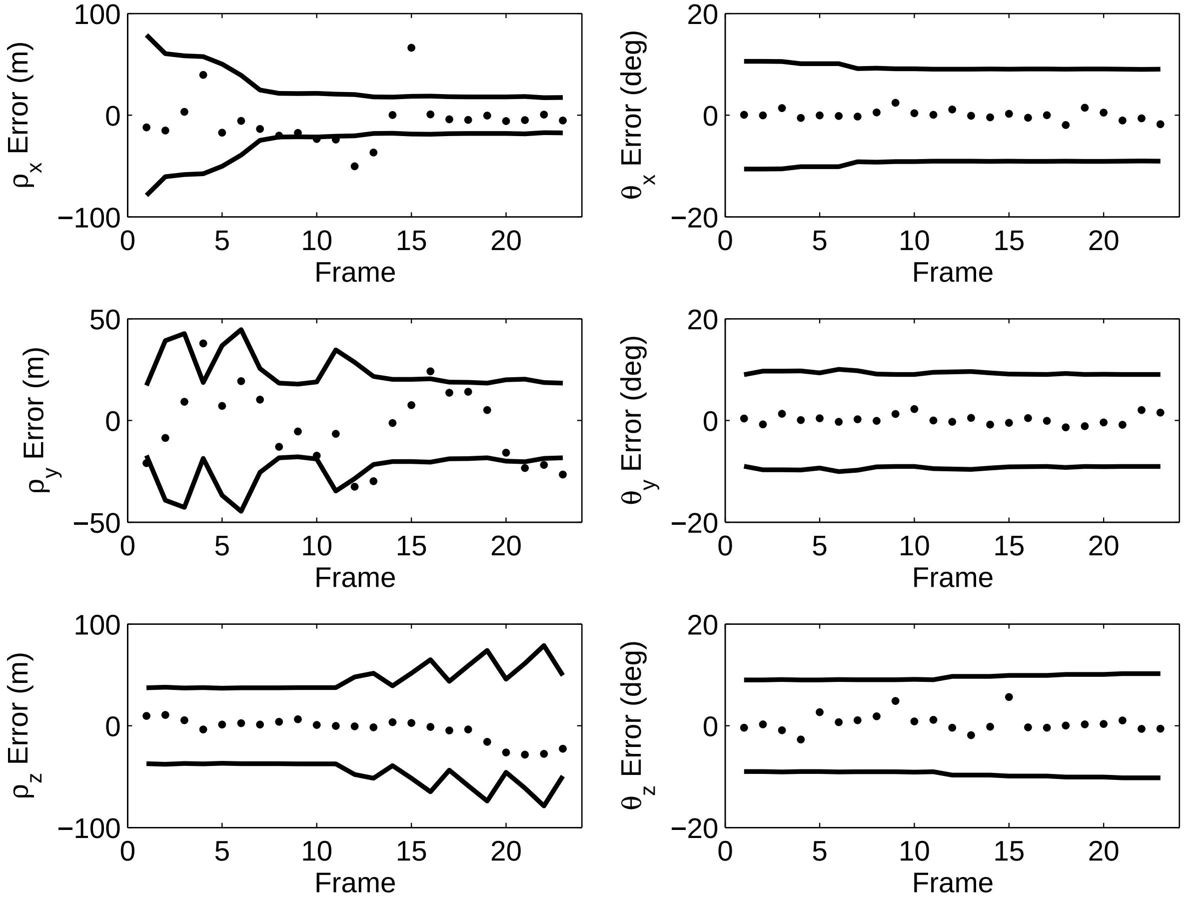


Figure 5.18: Dataset A, multi-frame trials. Errors for the **worst** of all 30 trials (i.e. trial with the worst mean radial translation error). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

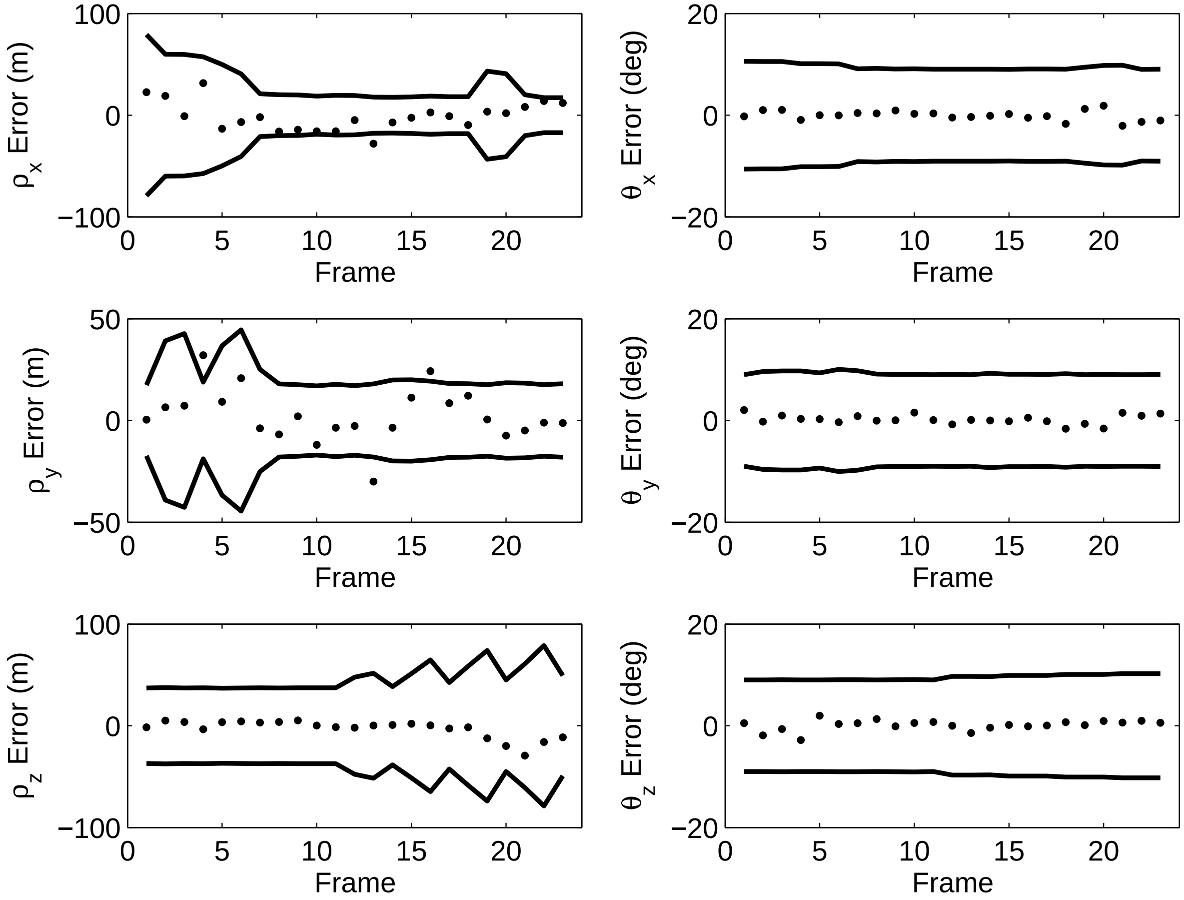


Figure 5.19: Dataset A, multi-frame trials. Errors for the **best** of all 30 trials (i.e. trial with the best mean radial translation error). Uncertainty is shown (solid line) for three standard deviations of combined transformation estimate (i.e. $\mathbf{B}^{-1}$ from Section 3.4.6) and ground-truth uncertainty. All three orientation errors only represent corrections made to their ground-truth measurements, since this same ground-truth was used as an input to the algorithm.

5.8 Results Summary

The results of the three algorithm configurations are compared in this section. Figure 5.20 presents the overall differences in position and heading error, while Figures 5.21 and 5.22 examine this from a frame-by-frame perspective.

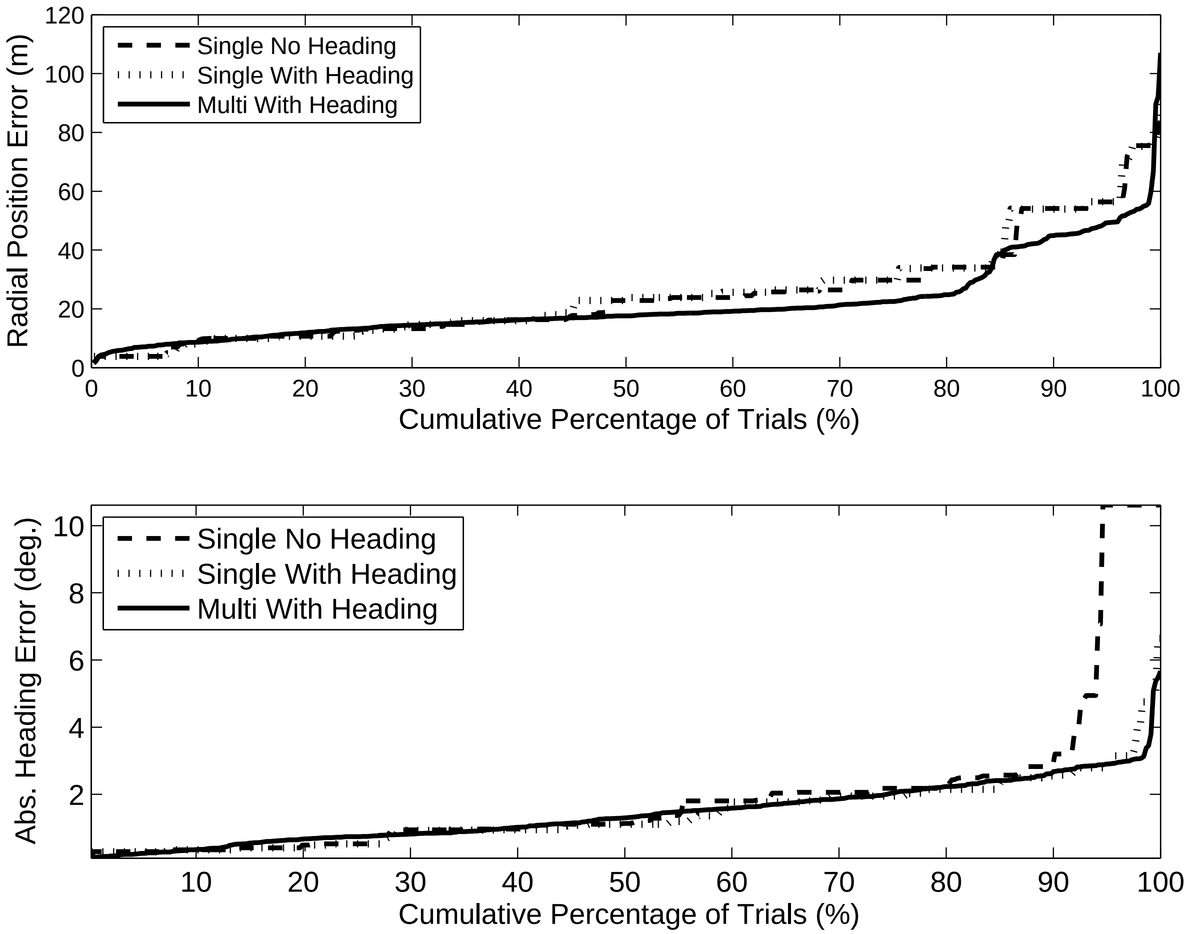


Figure 5.20: Cumulative plots of position and heading error. To properly compare the three scenarios, the multi-frame line was calculated only using frames which had a DARCES solution (i.e. not using frames which depended on pure odometry). Frame A20 for the single-frame no heading case had very large error and was omitted for clarity.

A significant improvement in position error is seen in the multi-frame trials. This is a result of the addition of the odometry measurements, which provide additional information on the desired estimates. Another advantage in the multi-frame case is the use of common features between frames. The same feature may be visible in various frames from different perspectives.

This provides further information for the optimization.

However, there is no guarantee that estimates will always be improved with the incorporation of odometry. For example, position error tended to double in Frame A12 for the multi-frame case. The inclusion of odometry will only provide better estimates *on average*.

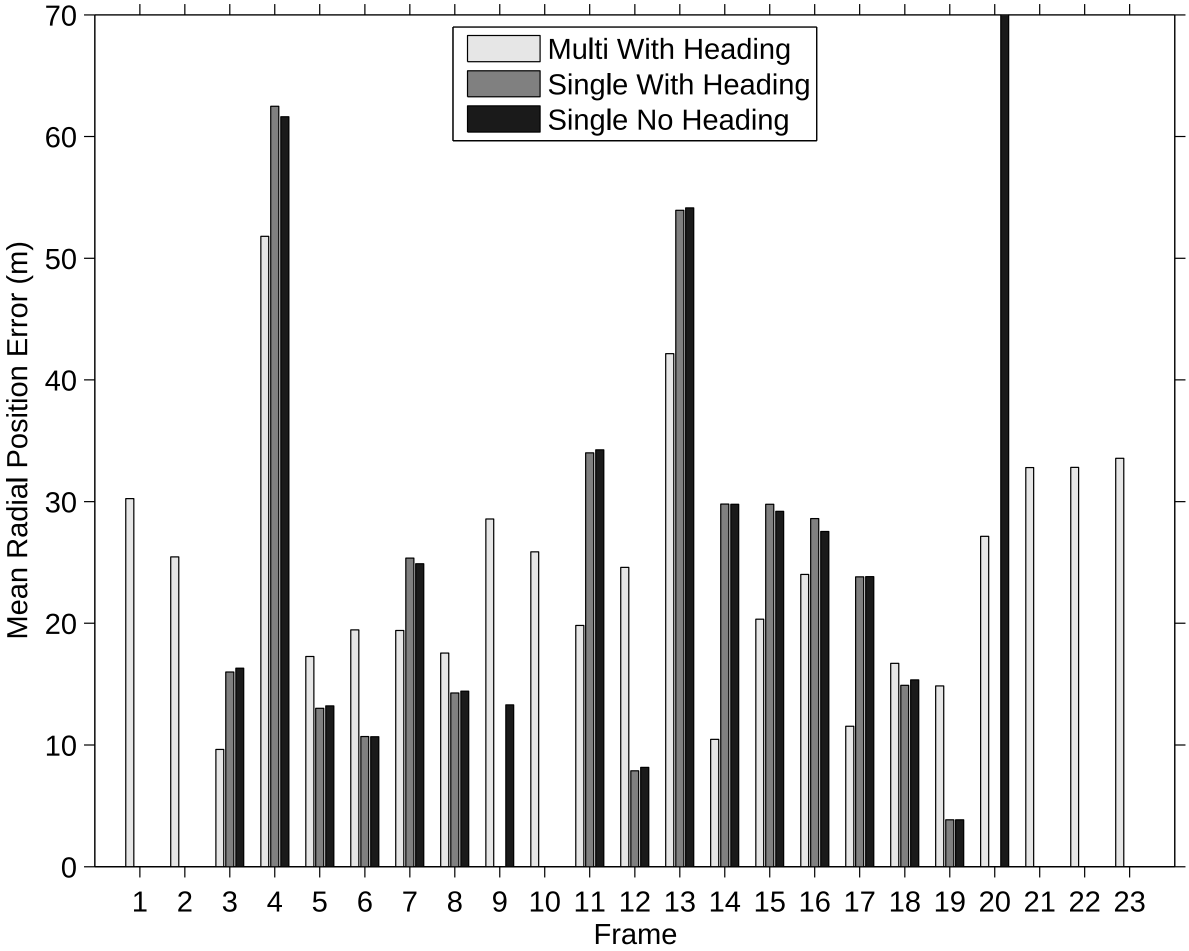


Figure 5.21: The mean radial position errors of translation estimates over all trials. This is shown for each of the three architecture configurations over all Dataset A frames. Note that some frames for the single-frame configurations produced no solution. The bar end for frame A20 for the single-frame no heading configuration extends beyond the plot limits, and is not shown for clarity.

Similarly, when heading data is available, the algorithm becomes much more robust to solutions with poor heading estimates. For example, solutions from A09 and A20 were rejected in the feature matching stage. These would have been discarded in the final DARCES heading filter, since their heading estimates did not fall within uncertainty of heading measurements.

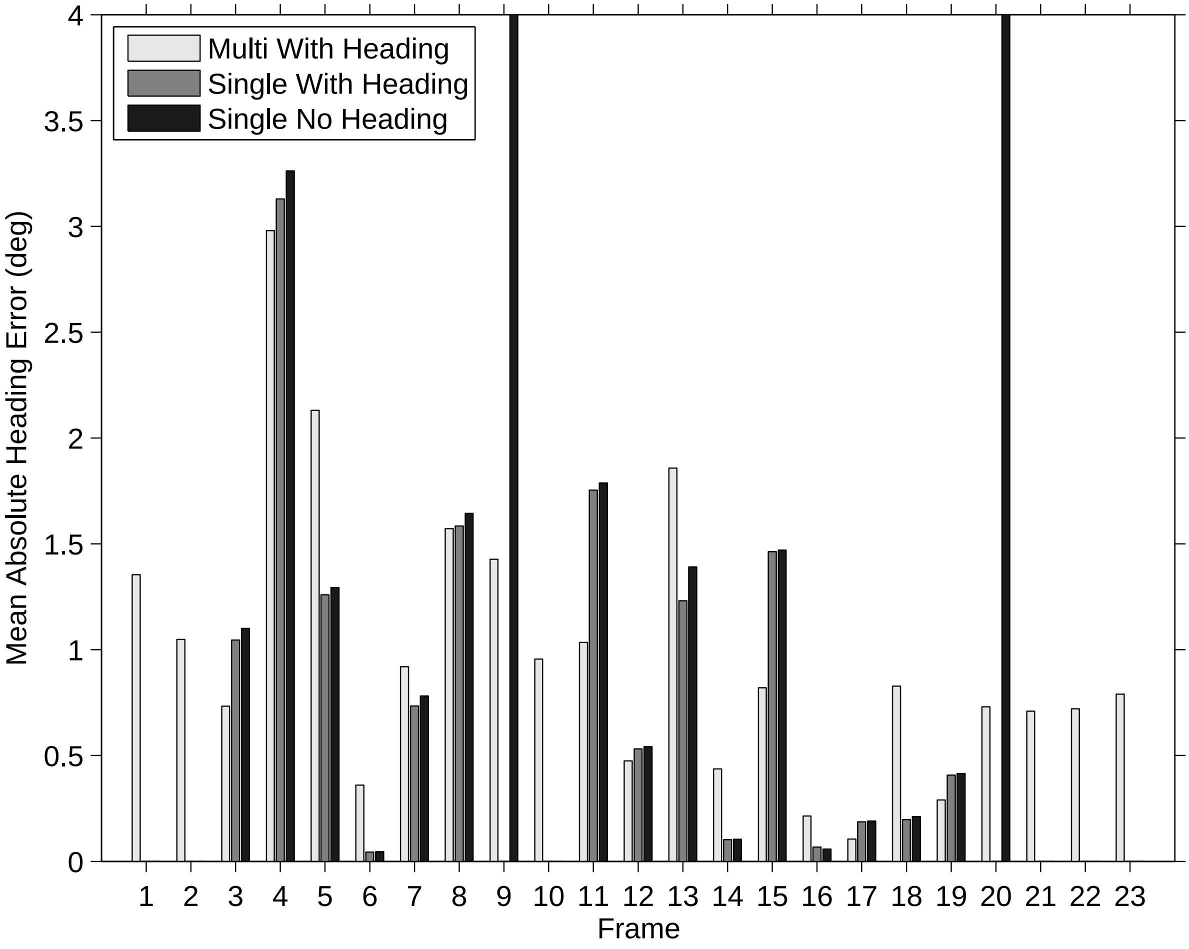


Figure 5.22: The mean absolute heading errors of orientation estimates over all trials. This is shown for each of the three architecture configurations over all Dataset A frames. Note that some frames for the single-frame configurations produced no solution. The bar ends for frames A09 and A20 for the single-frame no heading configuration extend beyond the plot limits, and are not shown for clarity.

Another observation can be drawn from Figure 5.23. There seems to be a weak negative correlation between pose error and the number of feature correspondences input into MOGA. This suggests that the effects of the random error in feature position measurements can be corrected to some degree by having a larger set of well-corresponding features. This may be done by improving the feature detection algorithm and/or using a higher resolution global map.

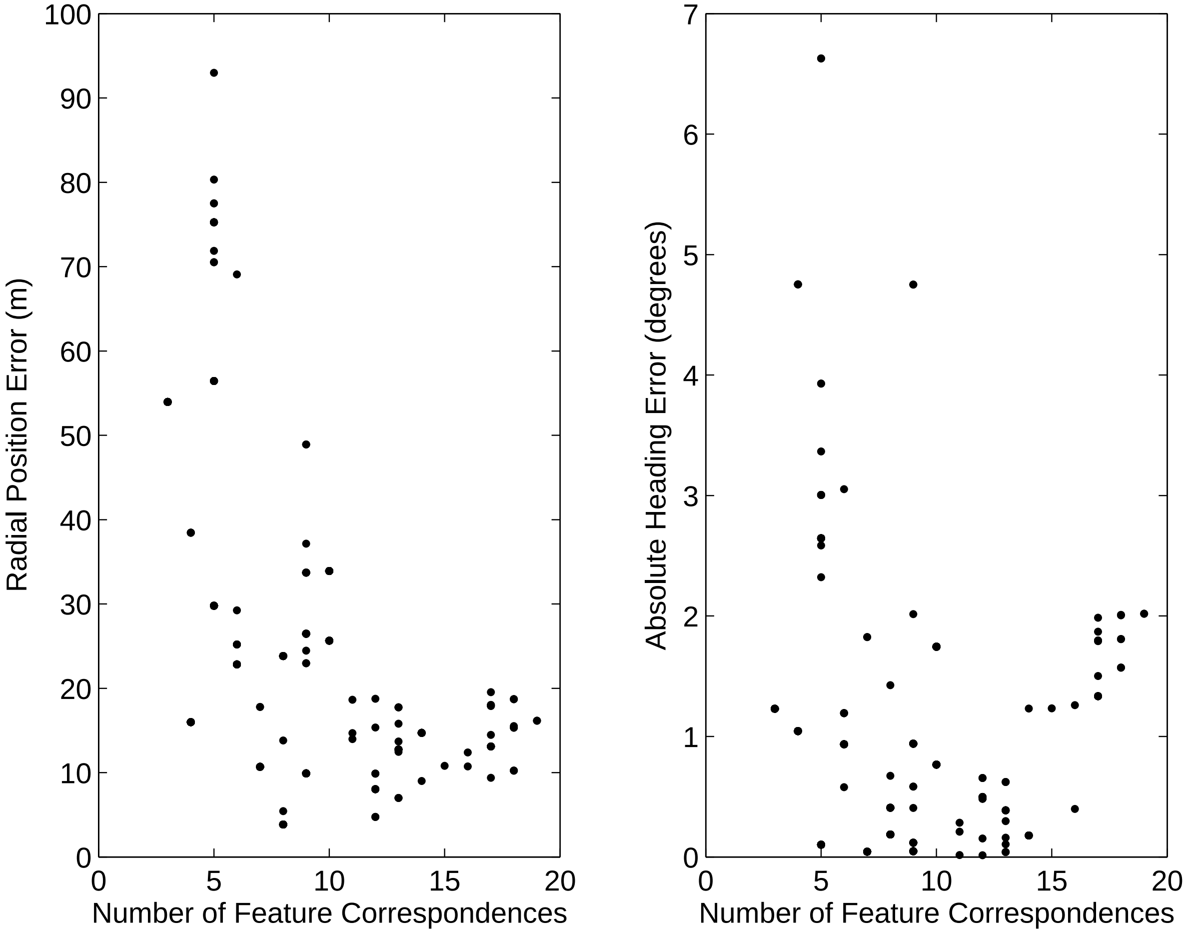


Figure 5.23: Relationship between pose errors and the number of feature correspondences selected by RANSAC. Results are for all dataset A trials of frames in the single-frame with heading configuration.

An important conclusion drawn from the multi-frame configuration was that the algorithm could maintain accurate localization estimates with no dependence on distance traversed. As shown in Figures 5.24 and 5.25, even having the benefit of ground-truth for frame A01, the accumulation of odometry error eventually grew to an unacceptable level. However, the architecture developed in this thesis was clearly able to correct odometry over this 10km traverse without the use of ground-truth to localize in the first frame.

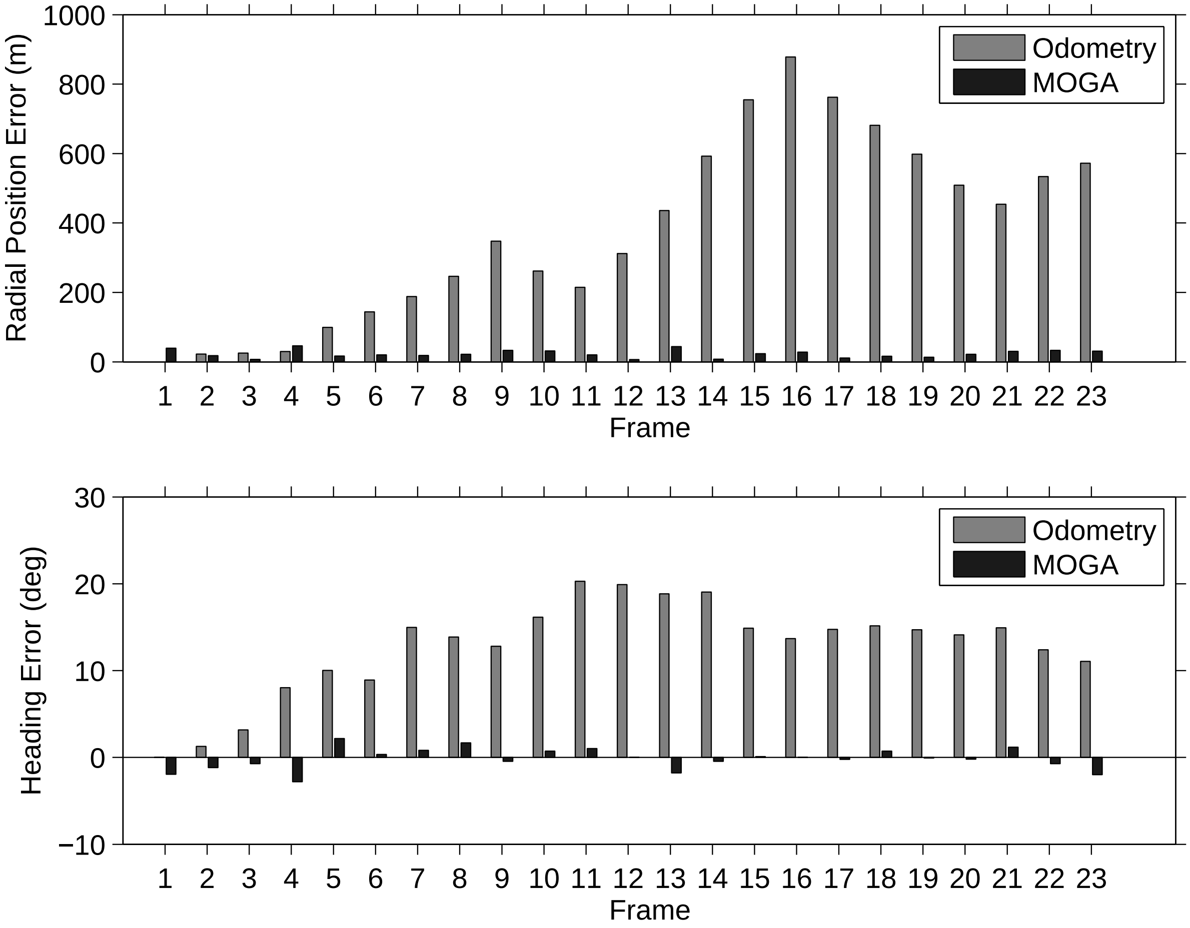


Figure 5.24: Odometry estimates compared to multi-frame MOGA-corrected estimates for a sample trial from Dataset A. Odometry was given ground-truth position and orientation for frame A01.

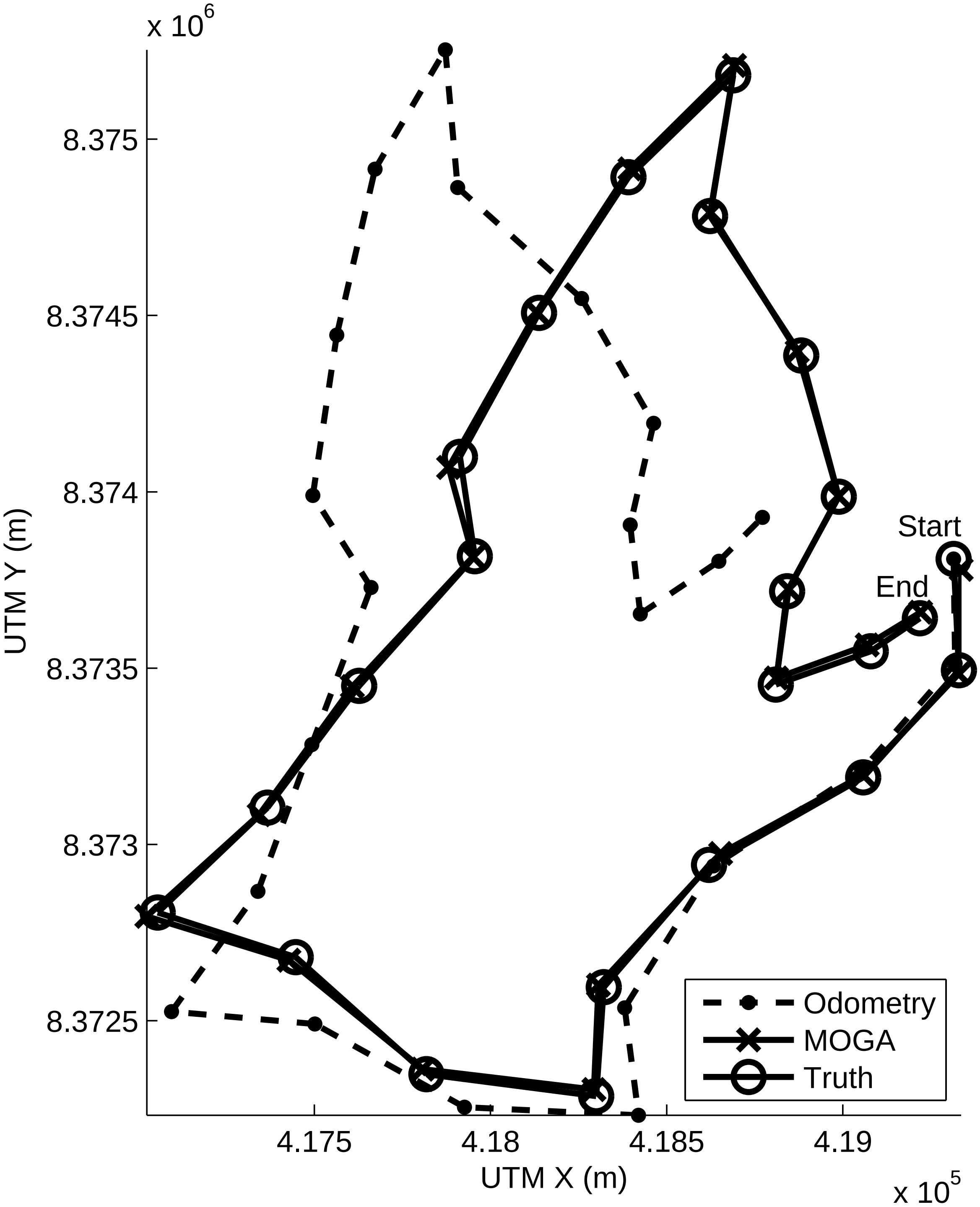


Figure 5.25: Odometry, corrected and ground-truth positions for a sample multi-frame trial run on Dataset A. Odometry was given ground-truth position and orientation for frame A01. Connecting lines do not indicate the intermediate path, but merely serve as a visual aid.

Chapter 6

Conclusion

This research has produced a number of novel contributions. A global localization technique was developed that matches rover-based lidar scans to an orbital elevation map using DARCES feature constellations. MOGA, a multiple-frame, least-squares alignment technique was designed that uses feature position, orientation and odometry measurements to refine pose estimates. The architecture was also validated with a realistic Mars/Moon analogue dataset from Devon Island, Nunavut.

A number of algorithm configurations were tested. The inclusion of heading measurement inputs was shown to greatly reduce the chances of selecting a poor solution. Sun sensors are therefore a valuable localization tool. Odometry information was also found to be very useful. Odometry measurements enabled the algorithm to refine its pose estimates, and to obtain estimates for frames that would otherwise have had no solution.

With the ideas presented in this thesis, a rover would have the ability to traverse over long-ranges with consistent and low localization error.

Chapter 7

Future Work

There are many areas for improvement in future work. The most important item is to more rigorously estimate uncertainties on input parameters and eliminate the use of simulated odometry data. A better model should be developed for lidar uncertainty that takes into account range-dependent errors, such as beam divergence. The estimates on input measurement uncertainties could also be improved by using devices better suited to measuring quantities of interest, such as orientation. In future experimentation, pitch and roll measurements should be acquired with an inclinometer and heading measurements from a sun sensor.

Ground-truth data should also be completely separate from all measurement inputs. In the current implementation, ground-truth with added noise was used to simulate odometry and orientation measurements. As an example of the consequences, the algorithm's output orientation estimates can then only be interpreted as the algorithm's adjustment to ground-truth orientation. A dataset of visual odometry and sun sensor heading measurements was collected in the field, but was not processed in time. Once processed, these will replace all simulated measurements and allow for improved analysis of all output pose estimates.

In terms of improving the architecture, there are many interesting avenues to investigate. A better feature detector could be used to more accurately predict the positions of features, such as a curvature-based detector. The main challenge would be in dealing with the occlusions in the lidar maps. It would also be interesting to quantitatively understand the effect of orbital map resolution on feature detection and localization accuracy.

In the feature matching stage, DARCES could be enhanced by using more than three control points in the initial hypothesis search. As Chen et al. (1999) explains, this would improve efficiency of the initial search which tends to be the bottleneck of DARCES. After many iterations of DARCES there was often several valid hypotheses remaining, where the one with the lowest error-score was chosen as the best hypothesis. However, it was not always the case that the lowest scoring hypothesis corresponded to the best solution. A better approach might be to first find the centroid position of the largest cluster of valid hypotheses using k-means clustering (Bishop, 2006). Assuming the hypotheses in this cluster are distributed evenly around the true position, the best hypothesis is then the one positioned closest to the centroid.

Odometry information could be used by DARCES for further gains in efficiency and accuracy. The translational odometry data could serve as another filter on position between two frames. If carefully implemented, this would greatly reduce the initial hypothesis search space. Odometry would also allow features to be shared between frames. This would be extremely useful for frames with low numbers of features. For example, a frame with only two features would not normally have a DARCES solution. If odometry were available, a third feature could be used from a nearby frame to satisfy the conditions for a solution.

The greatest test of this algorithm would be to integrate it into a fully autonomous rover system. For the rover to successfully localize, it must collect a rich set of global and local features with many valid correspondences. Therefore, the question arises of how to select good scan sites. MOGA benefits greatly from long-range scanning capabilities since position errors of features become less significant when features are well spread out. For this reason, the lidar should not scan in canyons where nearby hills could occlude distant terrain features.

The suggestions made in this section would help to improve the robustness and accuracy of the developed algorithm. With additional work, it is believed this methodology could be used in a practical, long-range rover localization system.

References

Arun, K., Huang, T., & Blostein, S. (1987). Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *9*(5), 698–700.

Bae, K.-H., & Lichti, D. D. (2004, July). Automated registration of unorganised point clouds from terrestrial laser scanners. In *The 20th ISPRS Congress, Istanbul, Turkey.*

Bakambu, J. N., Allard, P., & Dupuis, E. (2006). 3D terrain modeling for rover localization and navigation. In *Proceedings of the The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)* (p. 61). IEEE Computer Society.

Barfoot, T. (2005). Online visual motion estimation using FastSLAM with SIFT features. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 579–585).

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. *European Conference on Computer Vision*, *1*, 404–417.

Behar, A., Matthews, J., Raymond, C., & Means, E. (2005). The JPL PAUSE aerobot. In *The International Conference on System, Man and Cybernetics* (Vol. 4, p. 3939-43).

Biesiadecki, J., Baumgartner, E., Bonitz, et al. (2005). Mars Exploration Rover surface operations: Driving Opportunity at Meridiani Planum. In *IEEE International Conference on Systems, Man and Cybernetics* (Vol. 2).

Bishop, C. (2006). *Pattern recognition and machine learning*. Springer.

Björck, Å. (1996). *Numerical Methods for Least Squares Problems*. Society for Industrial Mathematics.

Botev, Z. I. (2007). A novel nonparametric density estimator.

Bresina, J., Jonsson, A., Morris, P., & Rajan, K. (2005). Activity planning for the Mars Exploration Rovers. In *The International Conference on Automated Planning & Scheduling.*

Chaikin, A. (2004, March). The other Moon landings. *Air & Space Magazine*.

Chen, C.-S., Hung, Y.-P., & Cheng, J.-B. (1999). RANSAC-based DARCES: A new approach to fast automatic registration of partially overlapping range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(11), 1229-1234.

Chin, G., Brylow, S., Foote, M., et al. (2007). Lunar Reconnaissance Orbiter overview: The instrument suite and mission. *Space Science Reviews*, *129*(4), 391–419.

Cozman, F., & Krotkov, E. (1997). Automatic mountain detection and pose estimation for teleoperation of lunar rovers. In *IEEE International Conference on Robotics and Automation* (Vol. 3, p. 2452-7).

Crisp, J., Grotzinger, J., Vasavada, A., & Karcz, J. (2008, April). Mars Science Laboratory: Science overview. *Ground Truth from Mars: Science Payoff from a Sample Return Mission, Albuquerque, New Mexico*.

CSA, ESA, JAXA, NASA, et al. (2007, May). The Global Exploration Strategy: The framework for coordination.

Desai, P., Prince, J., Queen, E., Cruz, J., & Grover, M. (2008). Entry, descent, and landing performance of the Mars Phoenix lander. *AIAA Paper*, *7346*.

Dickson, P. (2001). *Sputnik: The shock of the century*. Berkley Books.

Fischler, M., & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395.

Furgale, P., Barfoot, T., & Enright, J. (2009, March). Sun sensing for planetary rover navigation. In *Proceedings of the IEEE Aerospace Conference. Big Sky, MT.*

Grodecki, J., & Dial, G. (2001). IKONOS geometric accuracy. In *Proceedings of ISPRS Joint Workshop High Resolution Mapping from Space* (pp. 19–21).

Guinn, J. (2001). Mars surface asset positioning using in-situ radio tracking. In *Proceedings of the 11 th Annual AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, CA* (pp. 45–53).

Haralick, R., & Shapiro, L. (1992). *Computer and Robot Vision. Vol. 1.* Addison-Wesley.

Hayashi, A., & Dean, T. (1989). Satellite-map position estimation for the Mars rover. In *Proceedings of the NASA Conference on Space Telerobotics* (Vol. 2, pp. 275–282).

Hughes, P. (1986). *Spacecraft attitude dynamics*. J. Wiley New York.

Johnson, A. (1997). Spin-images: a representation for 3-D surface matching. *Pittsburgh: Carnegie Mellon University*.

Johnston, M., Graf, J., Zurek, R., & Eisen, H. (2005). The Mars Reconnaissance Orbiter mission. In *IEEE Conference on Aerospace* (pp. 1–18).

Konolige, K., Agrawal, M., & Sola, J. (2007). Large scale visual odometry for rough terrain. In *Proc. international symposium on robotics research.*

Kuroda, Y., Kurosawa, T., Tsuchiya, A., & Kubota, T. (2004). Accurate localization in combination with planet observation and dead reckoning for lunar rover. In *IEEE International Conference on Robotics and Automation* (Vol. 2, p. 2092-7).

Leger, P., Trebi-Ollennu, A., Wright, J., et al. (2005). Mars Exploration Rover surface operations: driving Spirit at Gusev Crater. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (Vol. 2).

LeMaster, E. A., & Rock, S. M. (2003, March). A local-area GPS pseudolite-based navigation system for Mars rovers. *Autonomous Robots*, *14*(2-3), 209–224.

Li, R., Archinal, B. A., Arvidson, R. E., et al. (2006, February). Spirit rover localization and topographic mapping at the landing site of Gusev Crater, Mars. *Journal of Geophysical Research*, *111*(E2).

Li, R., Arvidson, R. E., Di, K., et al. (2007, February). Opportunity rover localization and topographic mapping at the landing site of Meridiani Planum, Mars. *Journal of Geophysical Research*, *112*(E2).

Li, R., Di, K., & Howard, A. B. (2007, March). Rock modeling and matching for autonomous long-range Mars rover localization. *Journal of Field Robotics*, *24*(3), 187–203.

Li, R., Di, K., Hwangbo, J., et al. (2007, June). Integration of orbital and ground images for enhanced topographic mapping in Mars landed missions. In *Proceedings of the Annual NASA Science Technology Conference (NTSC), College Park, MD.*

Li, R., Di, K., Matthies, L., Arvidson, R., et al. (2004). Rover Localization and Landing-Site Mapping Technology for the 2003 Mars Exploration Rover Mission. *Photogrammetric Engineering and Remote Sensing*, *70*(1), 77–90.

Li, R., Ma, F., Xu, F., Matthies, L., Olson, C. F., & Xion, Y. (2000). Large scale mars mapping and rover localization using descent and rover imagery. *Proceedings of ISPRS*, *33*.

Li, R., Ma, F., Xu, F., et al. (2002). Localization of Mars rovers using descent and surface-based image data. *Journal of Geophysical Research-Planets*, *10*, 765–760.

Li, R., Squyres, S. W., & Arvidson, R. E. (2005, October). Initial results of rover localization and topographic mapping for the 2003 Mars Exploration Rover mission. *Photogrammetric Engineering & Remote Sensing*, *71*(10), 1129–1142.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *ICCV*, *2*, 1150.

Matijevic, J. (1998). The pathfinder mission to Mars: Autonomous navigation and the Sojourner microrover. *Science*, *280*(5362), 454.

McEwen, A., Delamere, W., Eliason, E., et al. (2002, March). HiRISE - The High Resolution Imaging Science Experiment for Mars Reconnaissance Orbiter. In *Lunar and Planetary Science XXXIII, Houston, TX.*

Mian, A., Bennamoun, M., & Owens, R. (2005). Automatic correspondence for 3D modeling: An extensive review. *International Journal of Shape Modeling*, *11*(2), 253-291.

Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1615–1630.

Norris, J., Powell, M., Vona, M., et al. (2005). Mars Exploration Rover operations with the Science Activity Planner. In *Proceedings of the ieee international conference on robotics and automation* (pp. 4618–4623).

Olson, C. F., Matthies, L. H., Schoppers, M., & Maimone, M. W. (2003, June). Rover navigation using stereo ego-motion. *Robotics and Autonomous Systems*, *43*(4), 215–229.

Palluconi, F. D., Tamppari, L. K., Steltzern, A., & Umland, J. (2003, July). *Mars Science Laboratory (MSL): The US 2009 Mars Rover Mission.* NASA/JPL.

Planitz, B., Maeder, A., & Williams, J. (2005). The correspondence framework for 3D surface matching algorithms. *Computer Vision and Image Understanding*, *97*(3), 347–383.

Se, S., et al. (2004, October). Vision based modeling and localization for planetary exploration rovers. In *55th International Astronautical Congress.* European Space Agency.

Smith, D., Frey, H., Carvin, J., et al. (2001). Mars Orbiter Laser Altimeter - Experiment summary after the first year of global mapping of Mars. *Journal of Geophysical Research*, *106*(E10), 23–689.

Stein, F., & Medioni, G. (1995, December). Map-based localization using the panoramic horizon. *IEEE Transactions on Robotics and Automation*, *11*(6), 892.

Sun, Y., Paik, J., Koschan, A., et al. (2003, August). Point Fingerprint: A new 2-D object representation scheme. *IEEE Transactions on Systems, Man, and Cybernetics*, *33*(4).

Taati, B., Bondy, M., Jasiobedzki, P., & Greenspan, M. (2007). Variable dimensional Local Shape Descriptors for object recognition in range data. *IEEE 11th International Conference on Computer Vision*, 1–8.

Vago, J., Kminek, G., Haldemann, A., et al. (2008, March). Upcoming science activities in support of ESA's ExoMars mission. In *39th Lunar and Planetary Science Conference, League City, Texas.*

van den Boomgaard, R. and van Balen, R. (1992). Methods for fast morphological image transforms using bitmapped binary images. *CVGIP: Graphical Models and Image Processing*, *54*(3), 252–258.

Vandapel, N., Donamukkala, R. R., & Hebert, M. (2006, January). Unmanned ground vehicle navigation using aerial ladar data. *International Journal of Robotics Research*, *25*(1), 31–51.

Volpe, R. (2003). Rover functional autonomy development for the Mars mobile science laboratory. In *Proceedings of the IEEE Aerospace Conference* (Vol. 2).

Yacoob, Y., & Davis, L. (1992). Computational ground and airborne localization over rough terrain. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 781–783).

Yano, H., Kubota, T., Miyamoto, H., et al. (2006). *Touchdown of the Hayabusa spacecraft at the Muses Sea on Itokawa* (Vol. 312) (No. 5778). American Association for the Advancement of Science.

Zhang, D., & Hebert, M. (1999). Harmonic shape images: A 3D freeform surface representation and its applications in surface matching. In *Proceedings of EMMCVPR, Sophia Antipolis, France* (pp. 30–43).

Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, *13*(2), 119–152.

Appendix A

DEM Properties

The global map's DEM resolutions and positional accuracies are given in units of latitude and longitude. These map parameters must to be converted to linear quantities since a UTM grid is used in the algorithm. In the following derivations, the horizontal datum used is the North American Datum 1983 with the ellipsoid of the Geodetic Reference System 1980.

A.1 DEM resolution

A.1.1 Latitude, Longitude

The Earth can be approximated by a sphere of radius equal to the equatorial radius, $R_e = 6,378,137$m. The DEM's resolutions are small, therefore the x resolution is approximated by

$$L_x :\approx R_e \lambda_{\text{lat}} \tag{A.1}$$

where λ_{lat} is latitude resolution in radians. Longitude resolution would change as a function of latitude since the size of the Earth's longitudonal circles change with latitude. Therefore, the y resolution is

$$L_y :\approx R_{\text{e}} \lambda_{\text{lon}} \cos(\Lambda) \tag{A.2}$$

where Λ is latitude. At the DEM's mean latitude of $\Lambda :\approx 75.4°$, latitude and longitude resolutions are respectively, $\lambda_{\text{lat}} := 0.5$ and $\lambda_{\text{lon}} := 1.5$ degree seconds. Knowing a degree second is 1/3600 of a degree, this gives $L_x \approx$ 23m and $L_y \approx$ 12m.

A.1.2 Altitude

The altitude resolution of the DEM is given as a linear quantity and therefore needs no conversion. The z resolution of the global map is $L_z :=$ 1m. The altitude datum is the Mean Sea Level / Canadian Vertical Geodetic Datum of 1928.

A.2 DEM Positional Accuracy

A.2.1 Horizontal

The DEM planimetric accuracy are given in units of Circular Map Accuracy Standard (CMAS):

$$\text{CMAS} := 2.1460\sigma_c \tag{A.3}$$

where σ_c is the standard circular error given by

$$\sigma_c := \frac{1}{\sqrt{2}}\sqrt{\left(\sigma_{r_{G_x}}\right)^2 + \left(\sigma_{r_{G_y}}\right)^2} \tag{A.4}$$

where $\sigma_{r_{G_x}}$ and $\sigma_{r_{G_y}}$ are respectively the standard deviations in x and y global map measurements. These are assumed to be equal. Therefore, given that $\text{CMAS} =$ 30m, the horizontal positional accuracy of the DEM is $\sigma_{r_{G_x}} = \sigma_{r_{G_y}} = 13.979$m.

A.2.2 Vertical

The vertical accuracy is given in units of Linear Map Accuracy Standard (LMAS):

$$\text{LMAS} := 1.6449\sigma_{\text{r}_{\text{G}_\text{z}}} \tag{A.5}$$

Therefore, given that $\text{LMAS} =$ 20m, the vertical positional accuracy of the DEM is $\sigma_{r_{G_z}} =$ 12.159m.

Appendix B

Search Sphere Shell Thickness Derivation

Consider Figure B.1, in which global features G_1 and G_2 are hypothesized to correspond respectively with local feature control points L_1 and L_2. The two global features are separated by a radial distance d_G and the two local features by d_L where

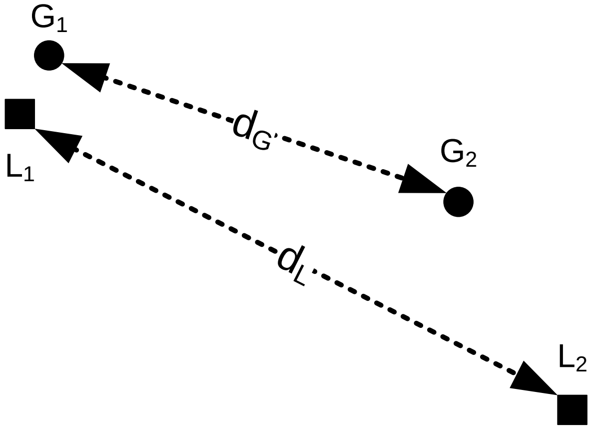


Figure B.1: Comparing distances between global features G_1, G_2 and local features L_1, L_2.

$$d_G^2 := d_{G_x}^2 + d_{G_y}^2 + d_{G_z}^2, \tag{B.1}$$

$$d_L^2 := d_{L_x}^2 + d_{L_y}^2 + d_{L_z}^2 \tag{B.2}$$

and where these distance components are given by the differences in feature positions:

$$\begin{aligned} d_{G_x} &:= \left|r_{G_{1x}} - r_{G_{2x}}\right|, d_{G_y} := \left|r_{G_{1y}} - r_{G_{2y}}\right|, d_{G_z} := \left|r_{G_{1z}} - r_{G_{2z}}\right|, \\ d_{L_x} &:= \left|r_{L_{1x}} - r_{L_{2x}}\right|, d_{L_y} := \left|r_{L_{1y}} - r_{L_{2y}}\right|, d_{L_z} := \left|r_{L_{1z}} - r_{L_{2z}}\right|. \end{aligned} \tag{B.3}$$

The absolute values are necessary to properly define these as distances. If the correspondences were correct and all measurements noiseless, d_G and d_L would be equal. However, noise would produce an error E_{GL}:

$$E_{GL} := d_G - d_L. \tag{B.4}$$

In order for these hypothesized correspondences to pass to the next stage of DARCES, the magnitude of E_{GL} should not exceed the half-thickness of the search sphere's shell:

$$|E_{GL}| \leq t. \tag{B.5}$$

The shell half-thickness, t, is chosen such that

$$t := 3\sigma_{E_{GL}} \tag{B.6}$$

where σ will be used to denote one standard deviation of the uncertainty for the variable in question. With the shell half-thickness set at three standard deviations, 99% of valid correspondence position errors, E_{GL}, should fall within this threshold and therefore be correctly detected as a hypothesis. It is assumed that other sources of uncertainty are negligible and E_{GL} measurements are normally distributed.

Equation (B.6) must now be decomposed into the known global and local feature position uncertainties. In the following derivation, propagation of uncertainties is used with the assumption that zero covariance exists between measurements. The uncertainty in E_{GL} is

$$\left(\sigma_{E_{GL}}\right)^2 = \left(\sigma_{d_G}\right)^2 + \left(\sigma_{d_L}^2\right) \tag{B.7}$$

where σ_{d_L} and σ_{d_G} are given by

$$\begin{aligned} \left(\sigma_{d_G}\right)^2 &= \frac{1}{d_G^2}\left(\left(d_{G_x}\sigma_{d_{G_x}}\right)^2 + \left(d_{G_y}\sigma_{d_{G_y}}\right)^2 + \left(d_{G_z}\sigma_{d_{G_z}}\right)^2\right), \quad \text{(B.8)} \\ \left(\sigma_{d_L}\right)^2 &= \frac{1}{d_L^2}\left(\left(d_{L_x}\sigma_{d_{L_x}}\right)^2 + \left(d_{L_y}\sigma_{d_{L_y}}\right)^2 + \left(d_{L_z}\sigma_{d_{L_z}}\right)^2\right). \quad \text{(B.9)} \end{aligned}$$

The uncertainties of the Cartesian components of distance are related to the uncertainties in position. If the positional uncertainty of every global feature is equal,

$$\begin{aligned}
\sigma_{r_{G_x}} &:= \sigma_{r_{G_{1x}}} = \sigma_{r_{G_{1x}}}, \qquad &(\text{B.10})\\
\sigma_{r_{G_y}} &:= \sigma_{r_{G_{1y}}} = \sigma_{r_{G_{1y}}}, \qquad &(\text{B.11})\\
\sigma_{r_{G_z}} &:= \sigma_{r_{G_{1z}}} = \sigma_{r_{G_{1z}}} \qquad &(\text{B.12})
\end{aligned}$$

and the positional uncertainty of every local feature is equal,

$$\begin{aligned}
\sigma_{r_{L_x}} &:= \sigma_{r_{L_{1x}}} = \sigma_{r_{L_{1x}}}, \qquad &(\text{B.13})\\
\sigma_{r_{L_y}} &:= \sigma_{r_{L_{1y}}} = \sigma_{r_{L_{1y}}}, \qquad &(\text{B.14})\\
\sigma_{r_{L_z}} &:= \sigma_{r_{L_{1z}}} = \sigma_{r_{L_{1z}}} \qquad &(\text{B.15})
\end{aligned}$$

then

$$\begin{aligned}
\left(\sigma_{d_{G_x}}\right)^2 &= 2\left(\sigma_{r_{G_x}}\right)^2, \left(\sigma_{d_{G_y}}\right)^2 = 2\left(\sigma_{r_{G_y}}\right)^2, \left(\sigma_{d_{G_z}}\right)^2 = 2\left(\sigma_{r_{G_z}}\right)^2, \qquad (\text{B.16})\\
\left(\sigma_{d_{L_x}}\right)^2 &= 2\left(\sigma_{r_{L_x}}\right)^2, \left(\sigma_{d_{L_y}}\right)^2 = 2\left(\sigma_{r_{L_y}}\right)^2, \left(\sigma_{d_{L_z}}\right)^2 = 2\left(\sigma_{r_{L_z}}\right)^2.
\end{aligned}$$

These equations can be greatly simplified by assuming that features are well spaced (d_G, d_L are large) and/or the terrain is low-lying (d_{G_z}, d_{L_z} are small):

$$\begin{aligned}
d_{G_z} &\ll d_G, \qquad &(\text{B.17})\\
d_{L_z} &\ll d_L. \qquad &(\text{B.18})
\end{aligned}$$

Therefore,

$$\begin{aligned}
d_G^2 &\approx d_{G_x}^2 + d_{G_y}^2, \qquad &(\text{B.19})\\
d_L^2 &\approx d_{L_x}^2 + d_{L_y}^2 \qquad &(\text{B.20})
\end{aligned}$$

and

$$\begin{aligned}
\left(\sigma_{d_G}\right)^2 &\approx \frac{1}{d_G^2}\left(\left(d_{G_x}\sigma_{d_{G_x}}\right)^2 + \left(d_{G_y}\sigma_{d_{G_y}}\right)^2\right), \qquad &(\text{B.21})\\
\left(\sigma_{d_L}\right)^2 &\approx \frac{1}{d_L^2}\left(\left(d_{L_x}\sigma_{d_{L_x}}\right)^2 + \left(d_{L_y}\sigma_{d_{L_y}}\right)^2\right). \qquad &(\text{B.22})
\end{aligned}$$

If planimetric uncertainties are equal,

$$\begin{aligned} \sigma_{r_{G_{xy}}} &:= \sigma_{r_{G_x}} = \sigma_{r_{G_y}}, && \text{(B.23)} \\ \sigma_{r_{L_{xy}}} &:= \sigma_{r_{L_x}} = \sigma_{r_{L_y}} && \text{(B.24)} \end{aligned}$$

then

$$\begin{aligned} \sigma_{d_G} &= \sqrt{2}\sigma_{r_{G_{xy}}}, && \text{(B.25)} \\ \sigma_{d_L} &= \sqrt{2}\sigma_{r_{L_{xy}}}. && \text{(B.26)} \end{aligned}$$

Substituting back into Equation (B.7):

$$(\sigma_{E_{GL}})^2 \approx 2\left(\sigma_{r_{G_{xy}}}\right)^2 + 2\left(\sigma_{r_{L_{xy}}}\right)^2. \qquad \text{(B.27)}$$

Therefore, the half-thickness of the search sphere's shell is

$$t \approx 3\sqrt{2}\left(\left(\sigma_{r_{G_{xy}}}\right)^2 + \left(\sigma_{r_{L_{xy}}}\right)^2\right)^{1/2}. \qquad \text{(B.28)}$$

Appendix C

Axis-Angle Conversion

Given a rotation matrix, $\mathbf{C}$, it is desired to convert it to an equivalent rotation about an axis $\hat{\mathbf{a}} := [a_1\ a_2\ a_3]^T$ over an angle ω. These are given by

$$\begin{aligned}
\omega &:= \arccos\left(\frac{\mathrm{trace}(\mathbf{C})-1}{2}\right), && \text{(C.1)}\\
a_1 &:= \frac{\mathbf{C}_{23}-\mathbf{C}_{32}}{2\sin(\omega)}, && \text{(C.2)}\\
a_2 &:= \frac{\mathbf{C}_{31}-\mathbf{C}_{13}}{2\sin(\omega)}, && \text{(C.3)}\\
a_3 &:= \frac{\mathbf{C}_{12}-\mathbf{C}_{21}}{2\sin(\omega)}. && \text{(C.4)}
\end{aligned}$$

Therefore, the axis-angle vector representation of $\mathbf{C}$ is $\mathbf{a} := \omega\hat{\mathbf{a}}$. Hughes (1986) discusses this procedure in greater detail.