High-Fidelity Planetary Route Determination Using Computationally Efficient Monocular Fisheye Odometry and Sun Compass

Eugene Fang

CMU-RI-TR-16-14

Submitted in partial fulfillment of the requirements
for the degree of Masters of Science in Robotics

May, 2016

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Masters Committee:
William “Red” Whittaker
David Wettergreen
Christopher Cunningham

Abstract

Accurate route determination is essential for robotic exploration on other planetary bodies. Today's rovers are limited in their speed and autonomy due to the computational expense of current route determination methodologies. Small, high-cadence, minimalist rovers with greater autonomy are poised to break new ground by expanding exploration capabilities. The ability to quickly and efficiently estimate a rover's route becomes even more crucial as the size, mass, computation, and power budgets continue to shrink.

This thesis develops a computationally efficient method for planetary route determination using a downward-facing monocular fisheye camera for visual odometry and a sun compass for a measure of absolute bearing. The method is tested in a series of lunar analog field experiments and shows the viability of such an approach. This research lays the groundwork that will enable safer, faster, and smarter navigation for the planetary rovers of tomorrow.

Table of Contents

1. Introduction and Motivation

Today's planetary robotic exploration is carried out by large, lumbering rovers. Due to the expense of such rovers, the resulting missions are risk averse. Areas that offer the greatest scientific payout such as pits, lava tubes, and icy craters are avoided because they are too dangerous; getting stuck in the rugged terrain characteristic of these areas means the end of a mission. Even less rugged areas of the moon are currently avoided due to communication shadows that prevent direct communication to Earth. Small, high-cadence, minimalist rovers with more autonomy are poised to break new ground by expanding space exploration capabilities. Whether by decreasing overall mission costs, enabling symbiotic exploration among multiple rovers, or being able to operate autonomously for extended periods of time, these minimalist rovers will allow for more rapid exploration of areas previously avoided.

As with existing rovers, one of the primary challenges of minimalist planetary rover exploration is route determination. Determining a rover's route is crucial for path planning, operational awareness, and quantifying scientific discoveries (e.g. to determine the relative locations and distances among interesting findings). This becomes increasingly important as rovers gain more autonomy, venturing into areas without communication with Earth where a human operator would otherwise be able to assist if the rover ever became lost. Many sensors and methods that are used in terrestrial systems are too large, too heavy, or too power hungry for space. Existing methods of localization and route determination of planetary rovers are expensive, both in computational time and power requirements. Current optical methods used in planetary rovers that utilize stereo imaging are slow due to limited computational power and are limited in their speed and performance even on today's large, expensive rovers. As a result, they are not always performed and are at most used as an aid to teleoperation [1] – not for long-distance fully autonomous missions. Not only does the heavy computation required by existing visual odometry make implementation on minimalist rovers impractical, the array of stereo imaging sensors on a complex pan-tilt camera head is undesirable for such rovers with strict mass and power budgets. The ability to quickly and efficiently estimate a rover's route becomes even more crucial as the size, mass, computation, and power budgets continue to shrink. A method for inexpensive route determination will enable safer, faster, and smarter navigation of these minimalist rovers.

This research presents a novel approach for computationally efficient visual odometry with an unactuated downward looking monocular fisheye camera which may be feasible for minimalist rovers from both a computation and electromechanical configuration standpoint. Furthermore, this work shows how combining this potentially imperfect visual odometry with a measure of absolute bearing from a sun compass and pose graph optimization results in high fidelity planetary route determination.

2. Priors

Visual odometry, or the estimation of a camera's egomotion, is a widely studied problem in robotics. Methods range from simple incremental estimation using optical flow to full 3D triangulation with 6-DOF motion estimation approaches and utilize both monocular and stereo camera configurations.

Perhaps the most famous instances of stereo visual odometry utilized on planetary rovers come from the Mars Exploration Rovers and Mars Science Laboratory (Figure 1). For the Mars Exploration Rovers, terrain features are tracked between two pairs of stereo images in both pixel and world coordinates to estimate 6-DOF rover motion using a maximum likelihood estimator [1]. 2D features with large triangulation errors are thrown out, 3D features that don't undergo similar rigid motion between two frames are thrown out, and additional outliers are removed during motion estimation using RANSAC. The result is motion estimation error as low as ~2.5% and 1.5% in terrestrial tests over 24 meters and 29 meters respectively. However, visual odometry is not always performed due to computation time and power constraints since it can take around 3 minutes per frame on the Mars Exploration Rovers [1] [2], though this was reduced to ~40 seconds for the Mars Science Laboratory. As a result, most of the time only wheel odometry and inertial measurement are used, errors or 10% are common, and this error increases drastically when the wheels slip [2] [3]. To further improve accuracy, ground-based methods of orbital localization and bundle adjustment can be performed manually on Earth well after the fact. However, the actual use of this is not well documented, nor is this practical for high-cadence autonomous rovers that cannot communicate with Earth for extended periods of time.

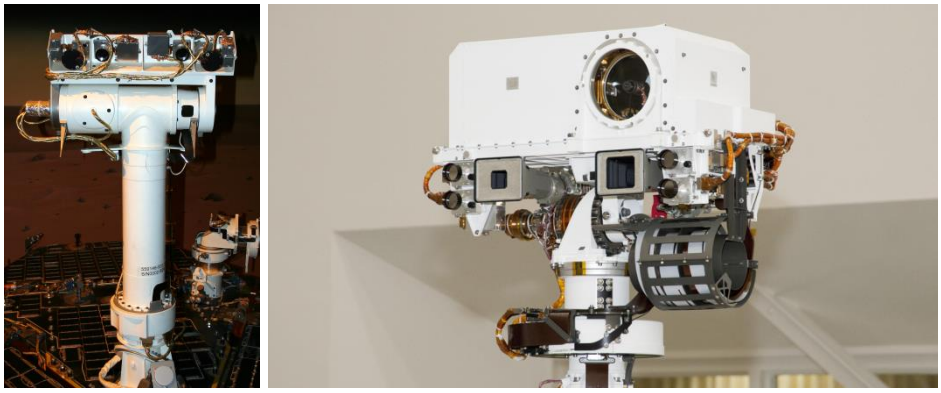

Figure 1: The pan-tilt stereo camera heads from the Mars Exploration Rovers (Spirit & Opportunity) (left) and the Mars Science Laboratory (Curiosity) (right)

The greater computational cost of stereo visual odometry, 3D triangulation, and 6-DOF motion estimation over a simpler monocular optical flow approach for 3-DOF motion estimation is known and explored. Methods that rely on the eight-point algorithm [4], which efficiently estimates the relative 6-DOF camera pose between two images given eight or more image point correspondences, fail in largely planar scenes – typical of those experienced by a planetary rover. The more computationally intensive five-point algorithm

(Figure 2) does not suffer from this same issue, a more efficient solution is presented by Nistér [5], and is shown to work with a monocular approach [6]. However it is shown that the result of the five-point algorithm still degrades in planar scenes [7]. More importantly, neither the eight-point nor five-point algorithms take advantage of the known ground plane to improve estimation accuracy. These approaches estimate unconstrained motion in six degrees of freedom, which is not necessary in many cases for ground vehicles.

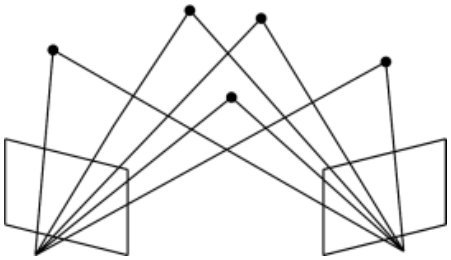

Figure 2: A sketch of the problem of estimating the relative pose of two cameras from five image point correspondences

Taking advantage of constrained motion in three degrees of freedom on the ground plane decreases computation time and improves estimation accuracy when compared to unconstrained six degree of freedom motion [8] [9]. Furthermore, the assumption that the camera is a fixed height above the ground plane allows for the scale of motion to be recovered, something that is atypical for general monocular visual odometry algorithms. These ideas are incorporated heavily in this research.

The use of wide field-of-view monocular cameras for odometry and its benefits have been explored in many contexts. In particular, it has the benefit over traditional lensing in that it is not affected by ambiguous motion [10]. One of the earliest instances of comes from Vassallo et al. [11], which uses a catadioptric vision system on an indoor mobile robot. Their approach projects optical flow vectors from images onto a unit sphere and calculate ego-motion (up to a scale factor) directly from the motion field on the sphere. However, it requires knowing the camera motion in advance to choose between two motion models and has greatly increased error if the camera undergoes both translation and rotation.

One instance where this type of camera was taken outdoors into the planetary setting was by Corke, who has experimentally seen benefits in trading camera image resolution for an increased field of view and presents two methods for visual odometry on a planetary rover [12]. The first is an optical flow method that takes advantage of a known ground plane to map the expected motion of a point on the ground back into the image plane, and then finds the motion of the camera through nonlinear optimization over a function of three motion parameters and five intrinsic camera parameters. However, this optimization step can be expensive and reliably achieving a solution is not guaranteed even when imposing constraints. The second method presented is a structure from motion method, but it had inconclusive results. Scaramuzza explores visual odometry of urban ground vehicles [13] and optimizes it to take advantage of the Ackermann steering constraint of such vehicles [14], but does not explore barren planetary environments. However, it utilizes appearance-based visual compassing for rotation estimation, which is the inspiration behind the weighting regions presented by this research.

One approach for efficient optical flow based odometry from a downward-facing camera is explored by Seegmiller [15]. The research specifically addresses the issue of computational cost by avoiding stereo triangulation and focusing on monocular optical flow, with additional focus on high vehicle speed and robustness to shadows. In this method, pyramidal Lucas-Kanade is used to track features that can have large displacements. These features are transformed into the robot frame under the assumption that the ground is flat so all features lie in the same plane, where 3-DOF velocity is then extracted after outlier rejection using RANSAC.

More recently, direct SLAM methods – those that do not rely on extracting feature keypoints – have gained traction [16] and have been extended to monocular omnidirectional cameras [17]. These direct approaches operate on the entire image to incrementally compute dense or semi-dense depth maps instead of sparse features, but require more computational resources as a result. Again, these approaches estimate unconstrained motion in six degrees of freedom and does not take advantage of a known ground plane. As a result, this research does not explore these methods too deeply, as optical flow feature based methods seem more promising.

The merits of a sun compass as a measure of absolute bearing for localization of planetary rovers is well understood and is shown to improve route determination even with imperfect odometry. This research incorporates Boirum's work in sensor design [18], but does not require the same high rate (~1 Hz) measurement.

3. Technical Approach

The merit of this method of route determination will be evaluated by comparing the determined routes with and without fisheye odometry, a sun compass, and pose graph optimization with a baseline route using only wheel and inertial odometry. A two dimensional environment is assumed, where the rover's route is represented as a series of states (x, y, θ) over time. The dataset used to evaluate this method was generated on an existing skid-steer lunar rover prototype in a 500m traverse over lunar-like terrain.

3.1. Wheel and Inertial Odometry

The baseline for comparing route determination uses a simple two dimensional odometry model using wheel turns. Figure 3 shows how a four wheel skid-steer rover is modeled as a two wheel differential drive rover.

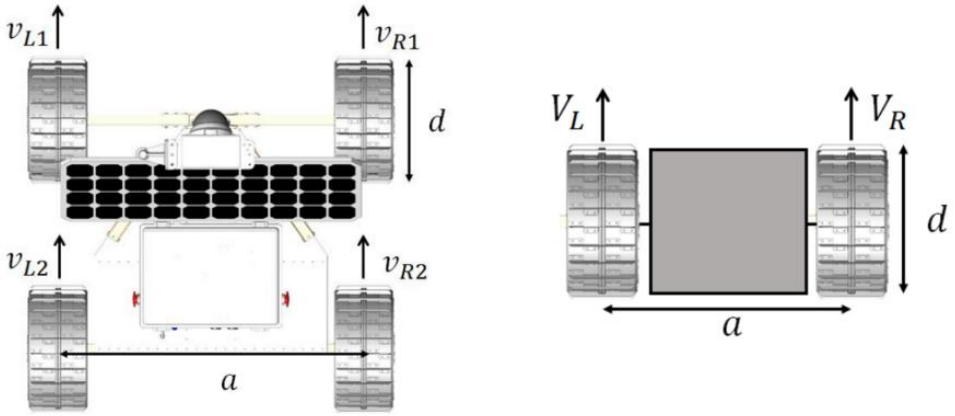

Figure 3: The four wheel skid-steer rover as viewed from above (left) is modeled as a two wheel differential drive rover for the purposes of odometry (right).

In the simplified model, V_L and V_R are the average velocities (in encoder ticks per second) of the left two and right two wheels respectively, a is the rover's track width, and d is the diameter of the rover's wheels. Furthermore, let n denote the number of encoder ticks per revolution of each wheel.

From this model, the change in left and right rover wheel ticks can be converted to left and right distances traveled as follows:

$$\Delta s_L = \Delta ticks_L \frac{\pi d}{n}$$

$$\Delta s_R = \Delta ticks_R \frac{\pi d}{n}$$

The change in overall rover distance traveled and heading is modeled as follows:

$$\Delta s = \frac{1}{2}(\Delta s_L + \Delta s_R)$$

$$\Delta\theta = \frac{1}{a}(-\Delta s_L + \Delta s_R)$$

This gives the rover's two dimensional position (x, y) and heading (θ) at time t:

$$\theta_t = \theta_{t-1} + \Delta\theta$$

$$x_t = x_{t-1} + \Delta s\ cos(\theta_t)$$

$$y_t = y_{t-1} + \Delta s\ sin(\theta_t)$$

Inertial yaw measurements are incorporated simply by ignoring the wheel odometry yaw and using the value from inertial measurement. While the more mathematically accurate approach is to utilize a Kalman filter to combine both measurements, it was experimentally determined that the estimated yaw from wheel odometry is extremely unreliable due to wheel slip. The resulting covariance corresponding to wheel odometry yaw is so large that its contribution is negligible, thus making it more computationally efficient to forego the filtering altogether.

3.2. Monocular Downward-Facing Fisheye Visual Odometry

While this work focuses on using a downward-facing fisheye camera, any camera that can be modeled as a spherical camera [19] and provides a large (~180°) field of view can be used. The spherical camera model maps a point X in the world viewed at pixel (u, v) to a ray that starts at the camera frame's origin and points to that point as shown in Figure 4. These rays are normalized on the unit sphere.

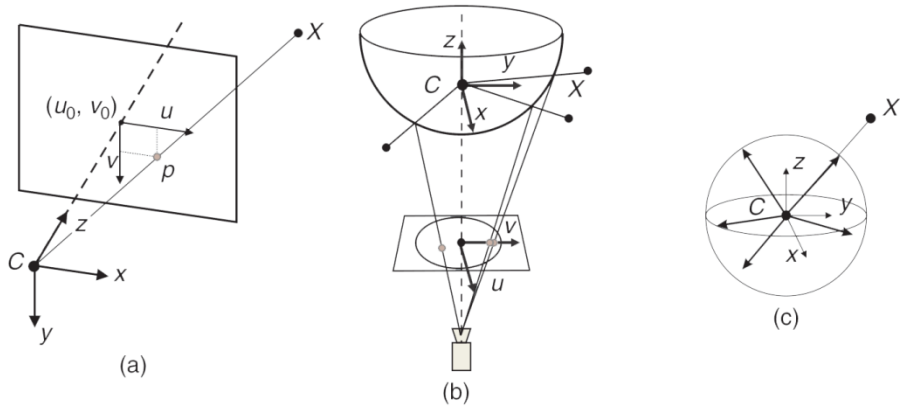

Figure 4: (a) Perspective projection, (b) catadioptric projection, (c) spherical model for both the perspective and catadioptric cameras. From [19]**.**

Cameras that meet these criteria include fisheye cameras and catadioptric cameras (camera with a hyperbolic mirror) (Figure 5). Simplified models of these cameras are shown in Figure 6.

Figure 5: Left: the fisheye lens and camera used in this research, Right: an example catadioptric lens and camera

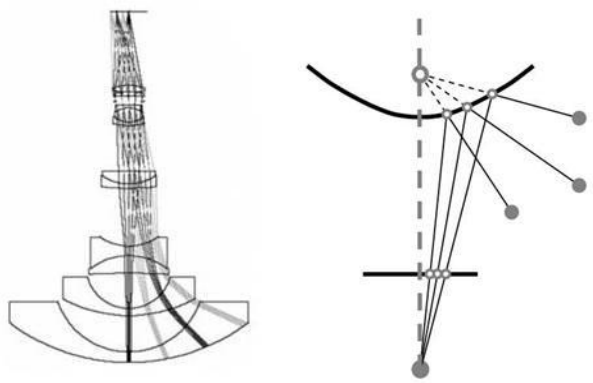

Figure 6: Left: downward-facing fisheye lens model, Right: catadioptric lens model

Fisheye cameras are more common than catadioptric cameras, but have implementation challenges when used in a downward-facing configuration. Because the lens is below the camera, the mechanical structure supporting the camera as well as wires from the camera are in its field of view. The structure must be designed to be minimally obstructing while being robust enough to survive planetary travel. In this work, a minimally obstructing camera support structure was designed while maintaining structural integrity by using two orthogonal plates positioned in planes that intersect the camera center. The view obstruction of this structure is compared to that of a naïve structure in Figure 7.

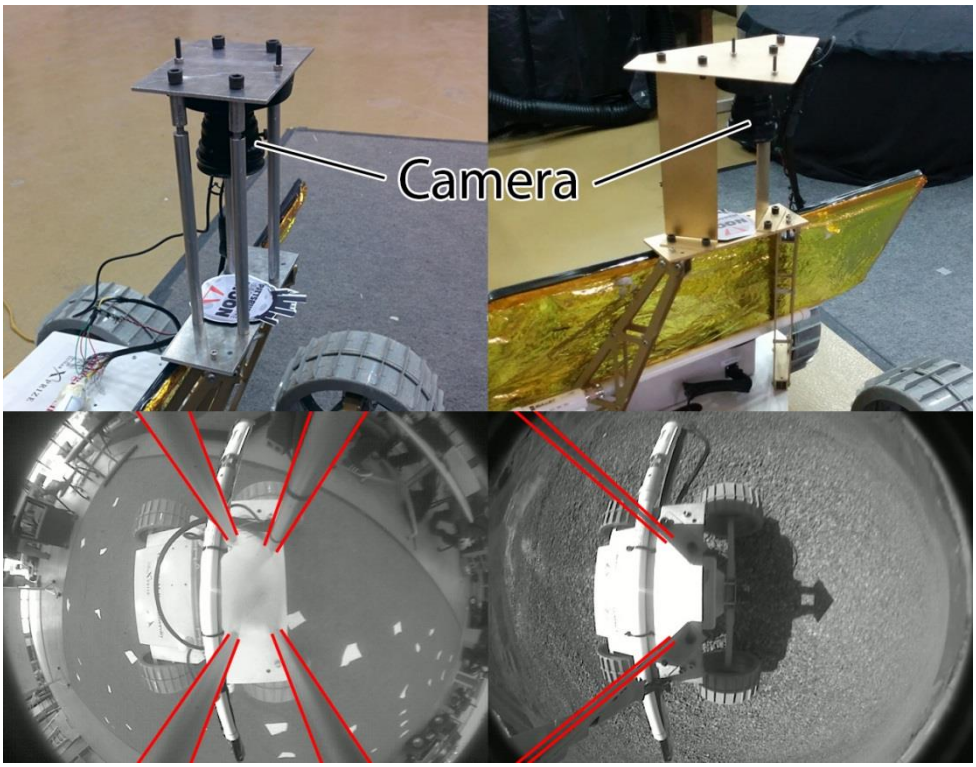

Figure 7: Naïve structure (top left) obstructs much of the camera image (bottom left). Orthogonal plate structure (top right) minimally obstructs the camera image (bottom right). The structure is outlined in red for clarity.

Catadioptric cameras offer the advantage over downward-facing fisheye cameras in that the camera is mounted below the lensing, lowering the center of gravity (which is an important point of consideration especially for small rovers where a camera system is a larger mass fraction of the rover) and eliminating wires or mechanical structure from obstructing the field of view. However, a catadioptric camera was not used in this work due to the difficulty in obtaining such lensing and because the visual odometry algorithm presented is largely indifferent either camera configuration.

3.2.1. Downward-Facing Wide Field of View Advantages

Downward-facing wide field of view cameras have several benefits over traditional outward-looking narrow field of view cameras in terms of estimating motion. It produces translation and rotation operators that are more robust due to geometric advantages. Because translational motion of the rover is generally perpendicular to the optical axis of a downward-facing camera, motion is more apparent and therefore estimated motion is more accurate. This geometry is illustrated in Figure 8.

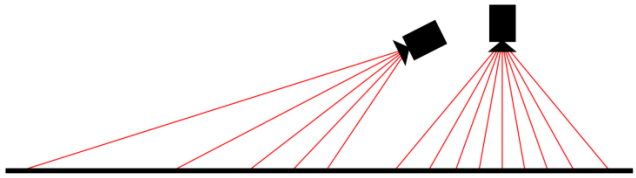

Figure 8: Red lines represent light rays emanating from the camera focal point, intersecting the image plane at constant pixel spacing ∂u, and projected onto the ground plane. The distance between ground plane projections is ∂x. $\frac{\partial u}{\partial x}$ gets larger as the rays approach the vertical axis, meaning the motion is more apparent.

Additionally, visual odometry often requires sufficient translation in order to accurately estimate motion; it is generally difficult to differentiate translation from rotation when the axis of rotation is perpendicular to the camera's optical axis. Because the optical axis of a downward-facing camera is mostly parallel to the rotational motion of the rover, it is not susceptible to these issues of degenerate motion.

Finally, this camera configuration allows the tracking of features and landmarks over long periods of time. This improves the effectiveness of bundle adjustment and allows for using the camera as a "visual compass" [20]. However, neither of these is incorporated in this work.

3.2.2. Optical Flow Method for Visual Odometry

Development on an optical flow method for visual odometry was chosen over a full structure from motion approach due to the generally less expensive computation required. At its core, this novel method directly operates on raw unrectified fisheye images and relies on feature matching and 2D rigid alignment to estimate incremental motion. Scale can be recovered if the height of the camera is known. Different weighting regions for translation and rotation are then used to improve accuracy.

3.2.2.1. Feature Matching and Rigid Alignment

Estimating incremental motion – the change in rover pose $(\Delta x, \Delta y, \Delta\theta)$ – between two images begins with finding corresponding features between the images. These feature correspondences are found in the raw unrectified fisheye images and both SURF and SIFT feature descriptors seem to give similar performance results for this algorithm. Figure 9 shows the matched SURF features between two sequential fisheye images.

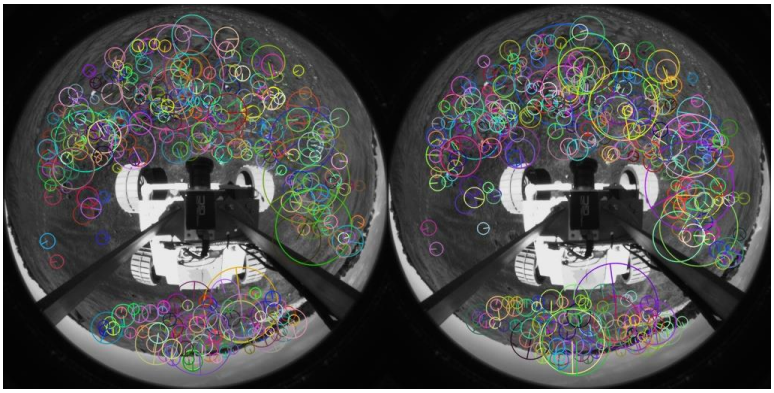

Figure 9: Feature matches between two sequential images

Once corresponding features are matched, they can be projected to onto the ground plane using the known spherical camera model and the height of the camera off the ground (Figure 10).

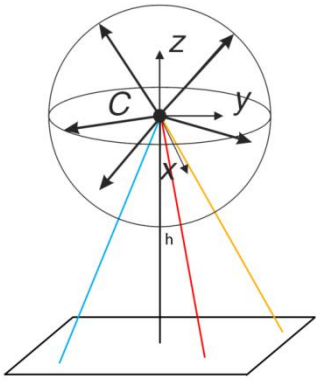

Figure 10: Image features projected onto the ground plane knowing the spherical camera model and camera height

First, following the spherical camera model, a pixel (u, v) is projected to point (x, y, z) on the unit sphere centered at the camera's focal point. In this particular implementation, the camera model and toolbox described by Scaramuzza [21] is used for calibrating the camera. This model first takes into account lens misalignments by using an affine transform with parameters (c, d, e) and the camera center $(u_c,\ v_c)$. Given a pixel at (u, v):

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} u_c \\ v_c \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix}^{-1} \left(\begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} u_c \\ v_c \end{bmatrix} \right)$$

The distance of $(x',\ y')$ from the image center is then mapped to the appropriate z' using a 4th order polynomial fit for the radial distortion of the camera with coefficients $(ss_0, ss_1, ss_2, ss_3, ss_4)$:

$$r = \sqrt{x'^2 + y'^2}$$

$$z' = ss_0 + ss_1 r + ss_2 r^2 + ss_3 r^3 + ss_4 r^4$$

Finally,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = norm \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

The intersection with the ground plane is computed by finding the intersection point between the line that passes through $(0, 0, 0)$ and (x, y, z) with plane $z = -h$, where h is the height of the camera above the ground.

Next, find the optimal rotation and translation to align the corresponding ground plane points (Figure 11).

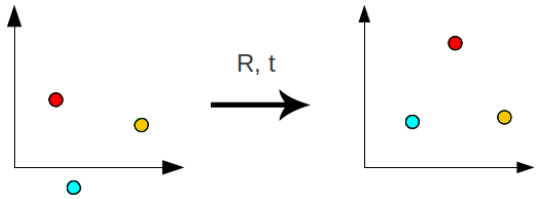

Figure 11: Given two sets of features projected onto the ground plane, find the optimal rotation and translation to align these corresponding points

To do this, a common algorithm for aligning point clouds using SVD is used. This algorithm has three main steps: 1) Find the centroids of both sets, 2) Move the centroid of each set to the origin to compute the optimal rotation, and 3) Compute the optimal translation. The full algorithm is outlined below:

$$centroid_A = \frac{1}{N} \sum_{i=1}^{N} P_A^i$$

$$centroid_B = \frac{1}{N} \sum_{i=1}^{N} P_B^i$$

$$H = \sum_{i=1}^{N} \left(P_A^i - centroid_A\right)\left(P_B^i - centroid_B\right)^T$$

$$[U, S, V] = SVD(H)$$

$$R = VU^T$$

$$t = -R * centroid_A + centroid_B$$

RANSAC is then used for outlier rejection. Figure 12 shows the motion of features between two frames with inliers in green and outliers in black as well as those features projected onto the ground plane. However, RANSAC for outlier rejection is not sufficient if a large portion of the image is the rover's self-shadow. If this poses to be a problem, a dynamic mask for feature selection by avoiding shadow edges can be used [15]. However, this was found to not be necessary in this research where the rover's shadow occupies a small fraction of the scene, but may be necessary if the number of ground features is low where shadow features make up the largest consistent set.

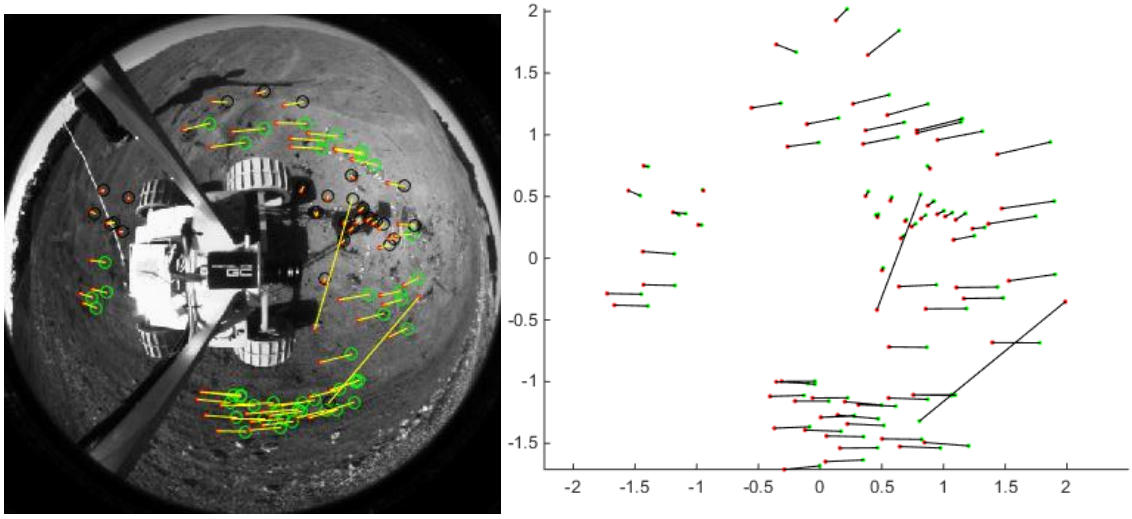

Figure 12: Left: Motion of features between two frames, with inliers in green and outliers in black. Notice the features detected in the shadow of the rover mast that are rejected. Right: Features projected onto the ground plane.

One key thing to notice is that instead of performing alignment or optimization in pixel coordinates which most visual odometry algorithms do, this approach performs alignment in world coordinates. This is beneficial from a computational efficiency standpoint since the alignment optimization is performed in a space with 3 degrees of freedom (x, y, θ) without needing to include camera intrinsics as part of the optimization. However, the side effect is that even small errors in feature detection near the periphery of the image (those that get projected far away, near the horizon) have a large effect on the estimated translation. In order to remedy this, different regions of the image are considered when estimating translation vs. rotation.

3.2.2.2. Weighting by Translation and Rotation Regions

Due to the geometry of the lensing and positioning of the camera, different areas of the image are more greatly influenced by translation and rotation of the rover. Figure 13 shows how a pure forward translation has insignificant optical flow when compared to a pure rotation in the blue regions, making it ideal for extracting rotation. The red region is less affected by near-horizon inaccuracies and makes it more ideal for extracting translation. Note that the blue region does not extend around the entire image – it is only at the front and back of the rover. This is to reduce the effect of parallax due to forward and backward rover motion [20].

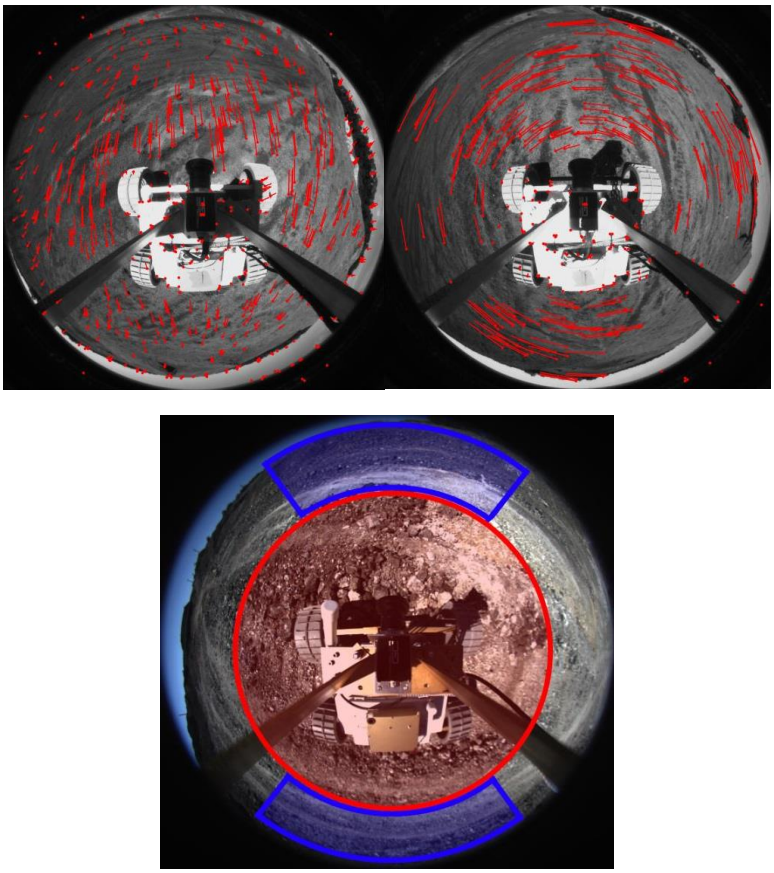

Figure 13: Top left: optical flow from a pure translation. Top right: optical flow from a pure rotation. Bottom: Approximate illustration of the different weighting regions for translation and rotation.

To make use of these different regions, the 2D rigid alignment is performed twice, once using only features within the red region and once using only features from the blue regions. The translation motion estimate comes only from features within the red region, while the rotation estimate is a simple weighted combination of the rotation from each rigid alignment, with the estimate using features from the blue region being weighted twice as heavily.

The choosing of these regions and their weighting is not extensively explored by this research. Only its effect is compared in Figure 14. Experiments have shown that for appearance-based visual compassing, a blue region that is ~60° wide strikes a good balance between

reducing parallax error and keeping as much of the image as possible [22]. It is reasonable to assume that this feature based rigid alignment within the blue region is geometrically similar to the appearance-based visual compassing, so a 60° region was used in this research as a baseline. However, the effect of different geometries, sizes, and perhaps non-discrete weighting regions is an area of great research interest and needs to be further investigated.

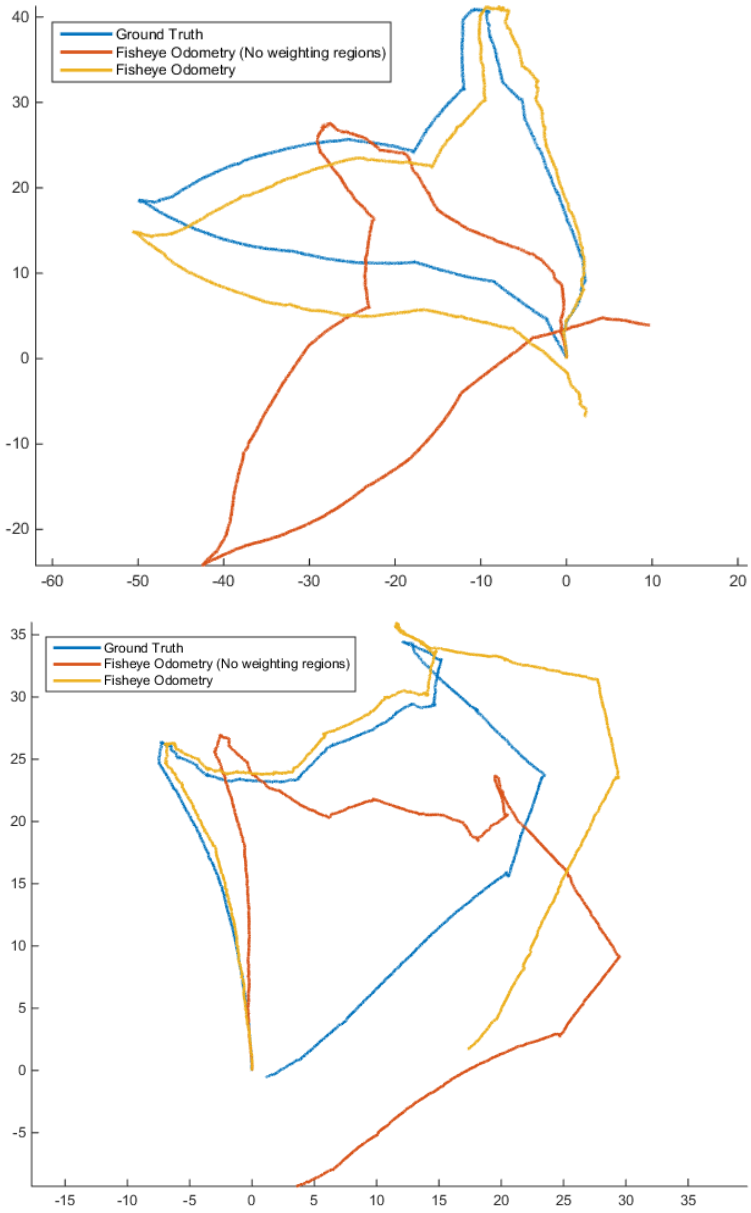

Figure 14: A comparison of the effect of weighting regions for two estimated routes. Notice that with weighting regions (yellow), the motion estimation is more rotationally robust than without (red) due to the reduced parallax effect.

3.3. Sun Compass for Absolute Bearing

Obtaining a measure of absolute bearing by looking at the sun is largely based on the methodology developed by Boirum [18]. A differentiating factor is that Boirum relies on continuously tracking the sun at a high rate (~1 Hz), while this research can utilize any sparse measure of absolute bearing. The sun compass functional block diagram from Boirum's work is outlined in Figure 15.

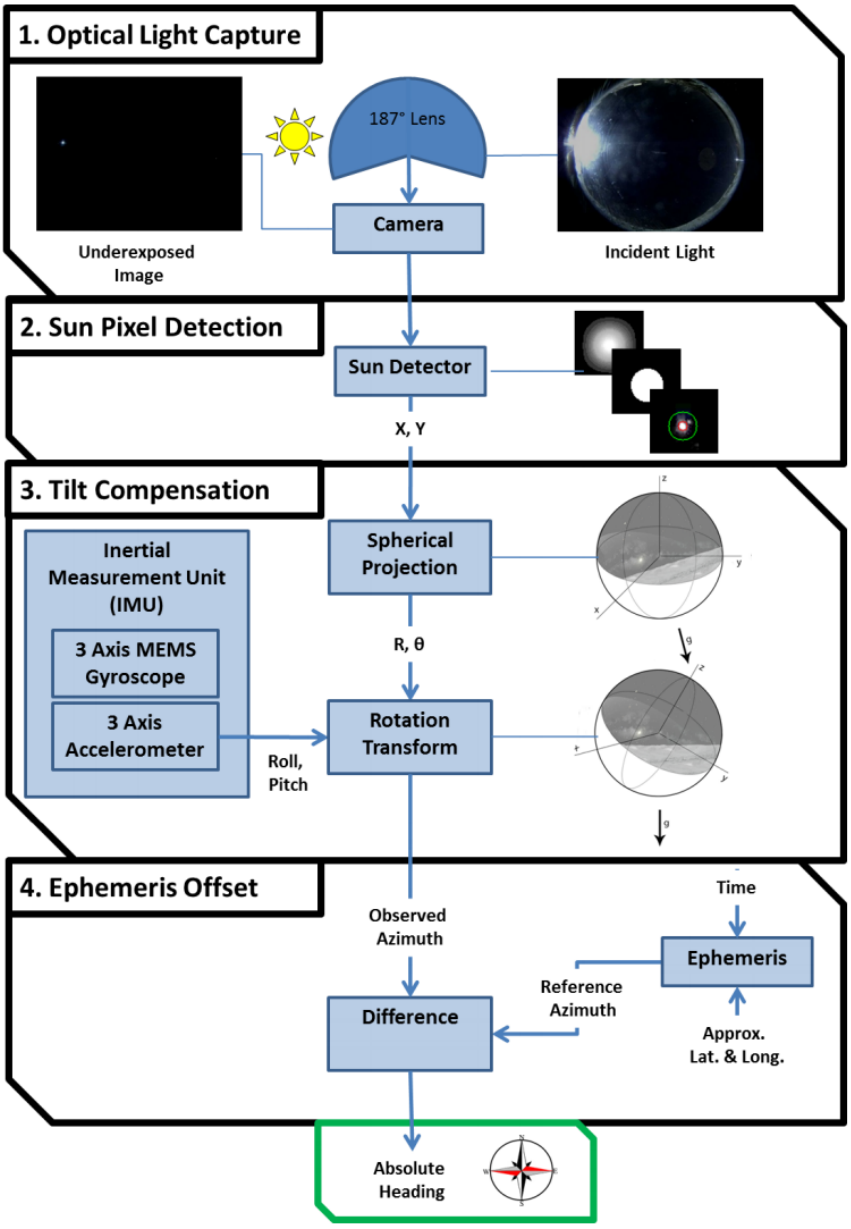

Figure 15: Sun Compass Functional Block Diagram. From [18]**.**

3.4. Pose Graph Optimization

In order to determine a rover's route, not just its current position, simple filtering to combine various measurements is not adequate. Filtering finds the most probable current state given all evidence up to the present. In contrast, the rover's determined route is the most probable sequence of states given all evidence up to the present. Pose graph optimization is used to combine various measurements and find a globally consistent solution. Figure 16 shows how a graph is constructed using measurements from visual odometry, sun bearing, and loop closure. Relative odometry measurements are added as edges between vertices, while absolute measurements of bearing and position are added as edges between a vertices and the root node.

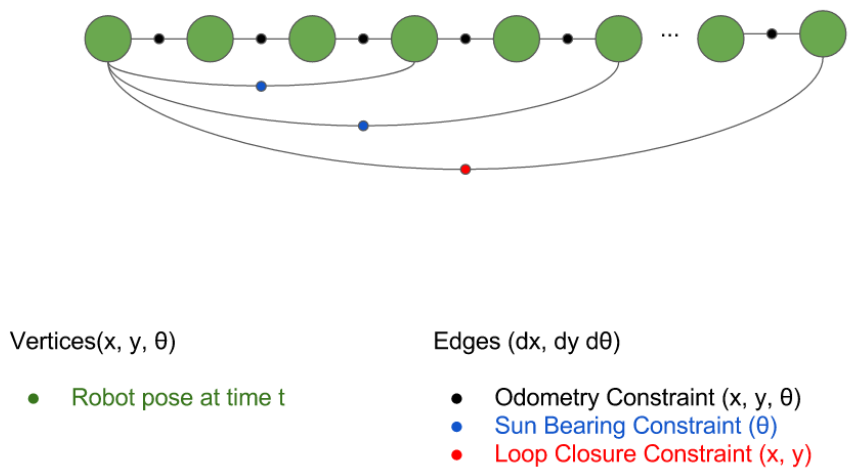

Figure 16: The various vertices and edges in the pose graph

There are many incremental and global methods for pose graph optimization. In this work, the open source g2o graph optimization framework [23] is used, which is able to optimize a graph from a 500m traverse containing ~3,000 vertices and ~6,000 edges in under 200ms on a desktop class processor. Whether or not this computation is feasible on a minimalist rover has not been investigated. Regardless, a graph of this size can be represented and transferred using only ~70KB with basic compression, making on-Earth computation possible, albeit with some latency. Figure 17 shows an example of such a graph.

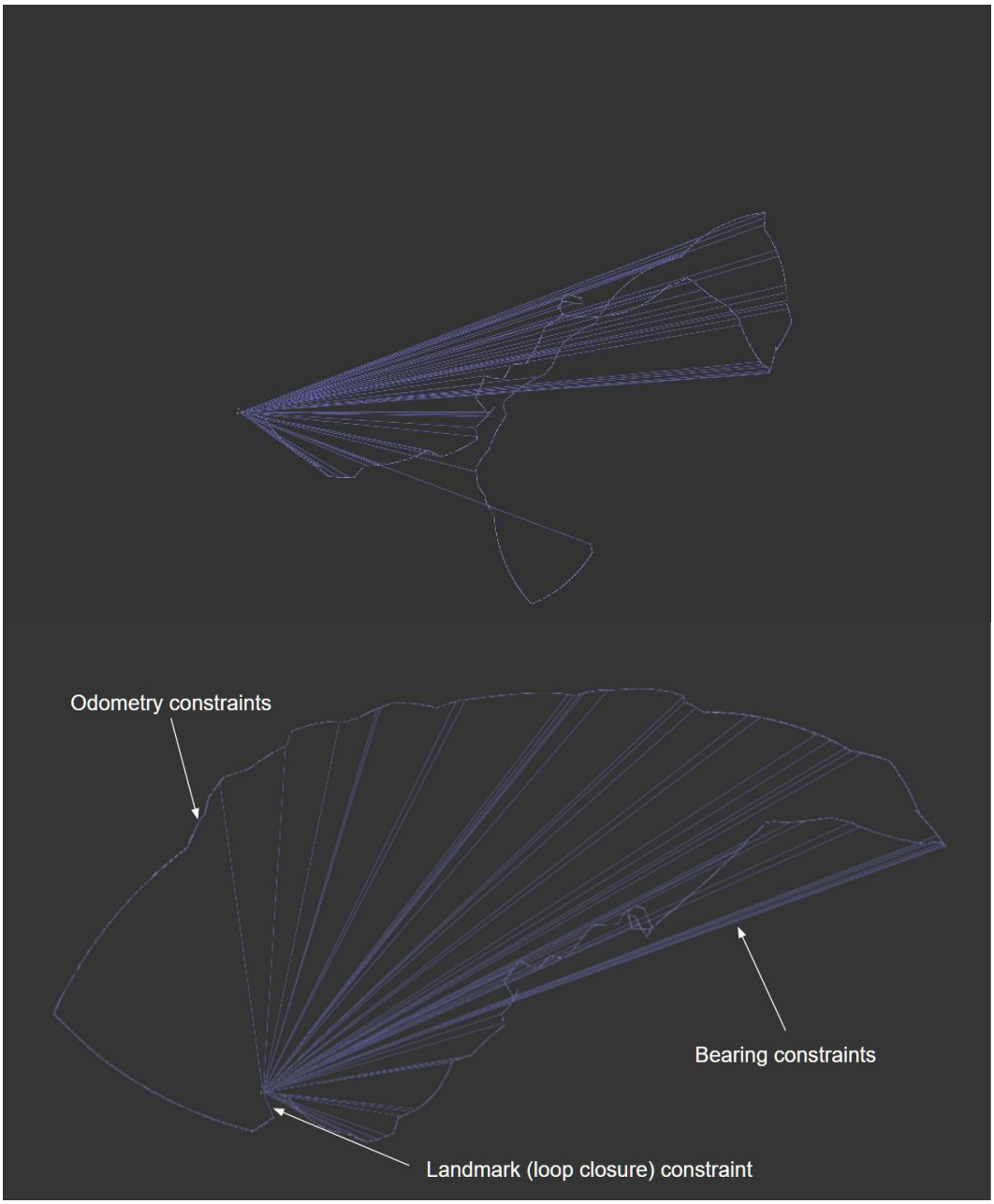

Figure 17: Initial pose graph of a 500m traverse using wheel odometry as the initialization (above). The same pose graph a few iterations into optimization (below)

4. Lunar Analog Field Experiments

The performance of fisheye odometry alone as well as combined fisheye odometry, sun compass, and pose graph optimization for route determination were evaluated in two teleoperated field tests carried out in April and July of 2015 totaling 781 meters of traverse. Both tests were carried out on "Andy," a four-wheel differential drive prototype lunar rover. Ground truth position was tracked with millimeter precision using a robotic total survey station and a corner cube mirror prism attached to the top of the rover's mast. The dataset from the April experiment (Figure 18) does not include inertial or sun compass data. Thus it is only used to evaluate the performance of fisheye odometry. Figure 19 shows Andy in its configuration during the July field test.

Figure 18: The version of Andy used for the April field experiment

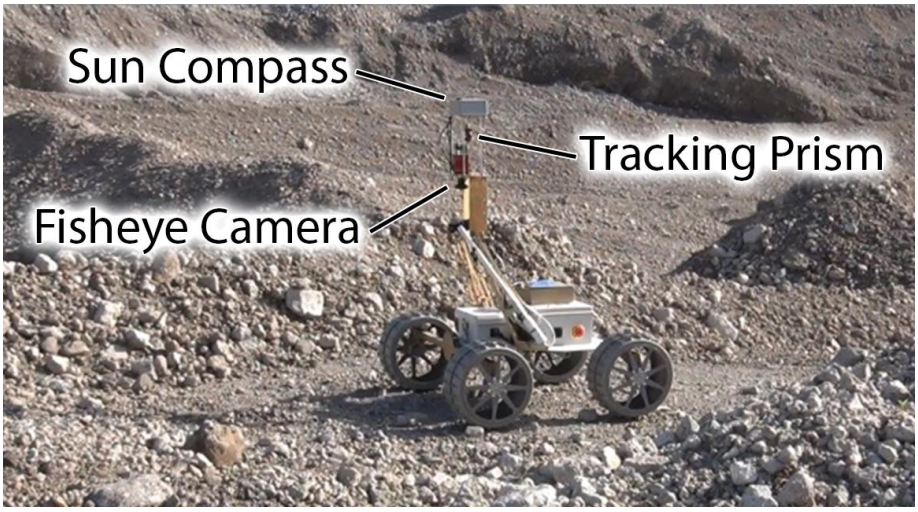

Figure 19: Andy with a sun compass used in the July field experiment

5. Results

Field experiments using fisheye odometry, sun compass, and pose graph optimization demonstrates route determination accuracy and simplicity not achieved by previous research. While fisheye odometry alone is sufficient in generally planar environments, the incorporation of absolute bearing and loop closure is necessary in environments where the planar assumption does not hold as well.

5.1. Wheel Odometry vs. Fisheye Odometry Only

Even in generally planar environments, such as shown in Figure 20, wheel odometry alone is a poor method for route determination.

Figure 20: Open, planar environment (April 2015)

Figure 21 shows how single instances of incremental errors due to wheel slip compound and results in path reconstruction divergence. Fisheye odometry is also an incremental measure, so it also suffers from the same problem of divergence. However, this divergence is more gradual as there are fewer singular erroneous measurements.

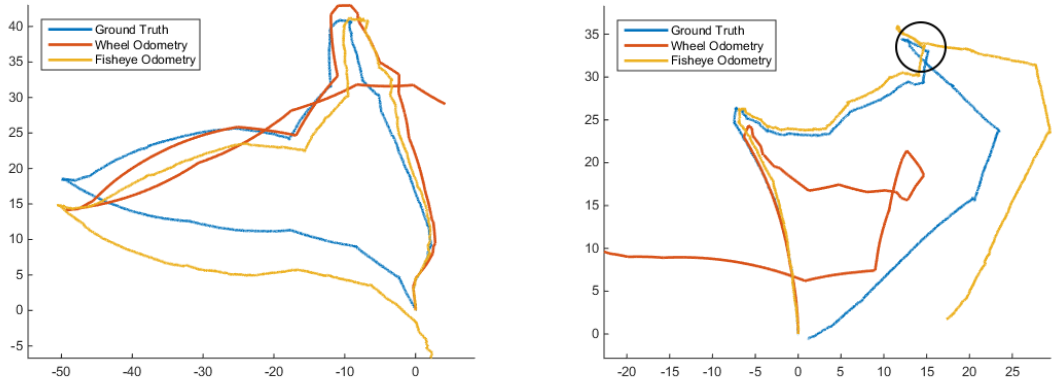

Figure 21: Comparison of wheel odometry, fisheye odometry, and ground truth in a planar environment. The circled area in the right image indicates where the rover climbs partially up a hill. Units are meters.

One large divergence in the fisheye odometry route in Figure 21 is circled. This occurs when the rover climbs partially up a hill, invalidating the ground plane assumption. This issue is even more pronounced in a particularly hilly environment as shown in Figure 22.

Figure 22: A hilly environment that invalidates the ground plane assumption (July 2015)

In this environment, fisheye odometry still outperforms wheel odometry alone, but the resulting path is is far from the ground truth (Figure 23).

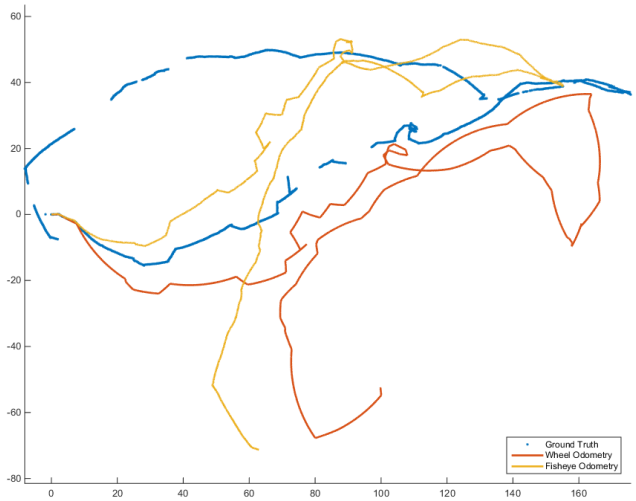

Figure 23: Comparison of wheel odometry, fisheye odometry, and ground truth in a hilly environment. Units are meters.

Notice that most of the errors are due to an incorrect heading, not distance. Incorporating absolute bearing fixes this problem.

5.2. Incorporating Absolute Bearing

Incorporating absolute bearing measurement at a rate of ~1 measurement per minute greatly improves route determination accuracy. Figure 24 shows how absolute bearing ensures all sections of the route are parallel to their respective sections in the ground truth. The maximum and final errors are 3-4% of the total distance, which is on par with the 2-3% errors seen by the Mars Exploration Rovers in terrestrial tests but with a more lightweight algorithm.

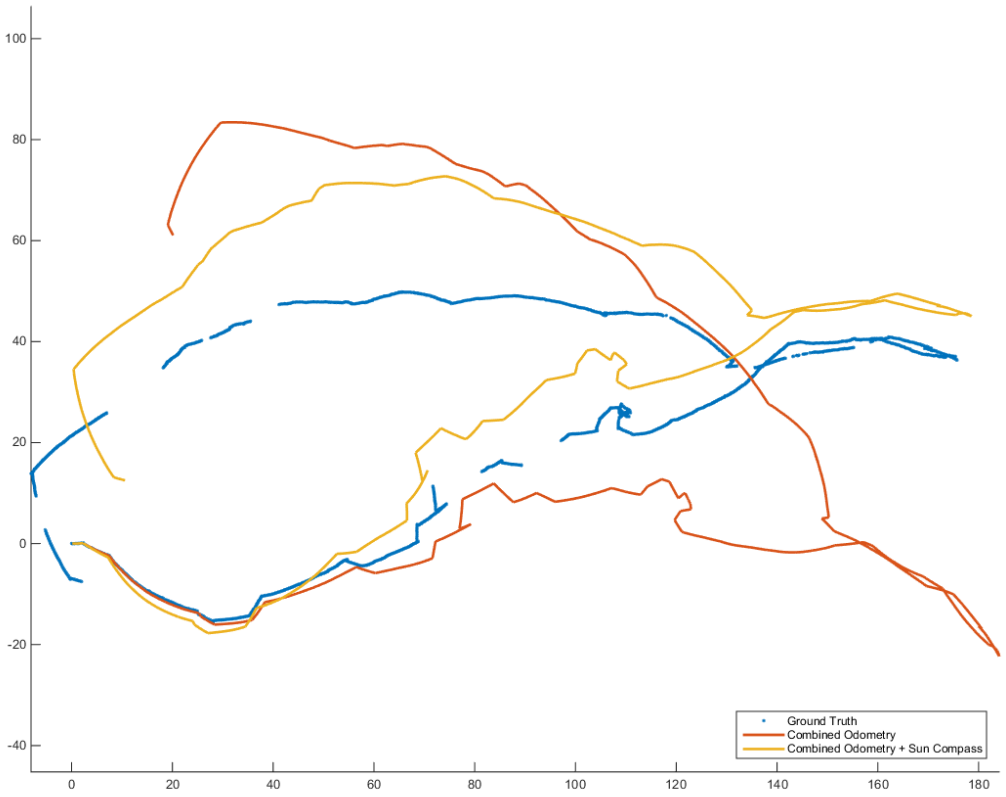

Figure 24: A measure of absolute bearing on estimated route drastically improves route determination, even when combined with imperfect combined odometry. Units are meters.

5.3. Including Single Point Loop Closure

A single point of loop closure was introduced simply by a human operator recognizing that the rover had returned to its starting point – the lander. Figure 25 shows the effect of loop closure, which reduces the maximum error over a 500 meter treacherous traverse to just 8 meters, or 1.6% of the total distance, which is below the 2-3% errors of the Mars Exploration Rovers.

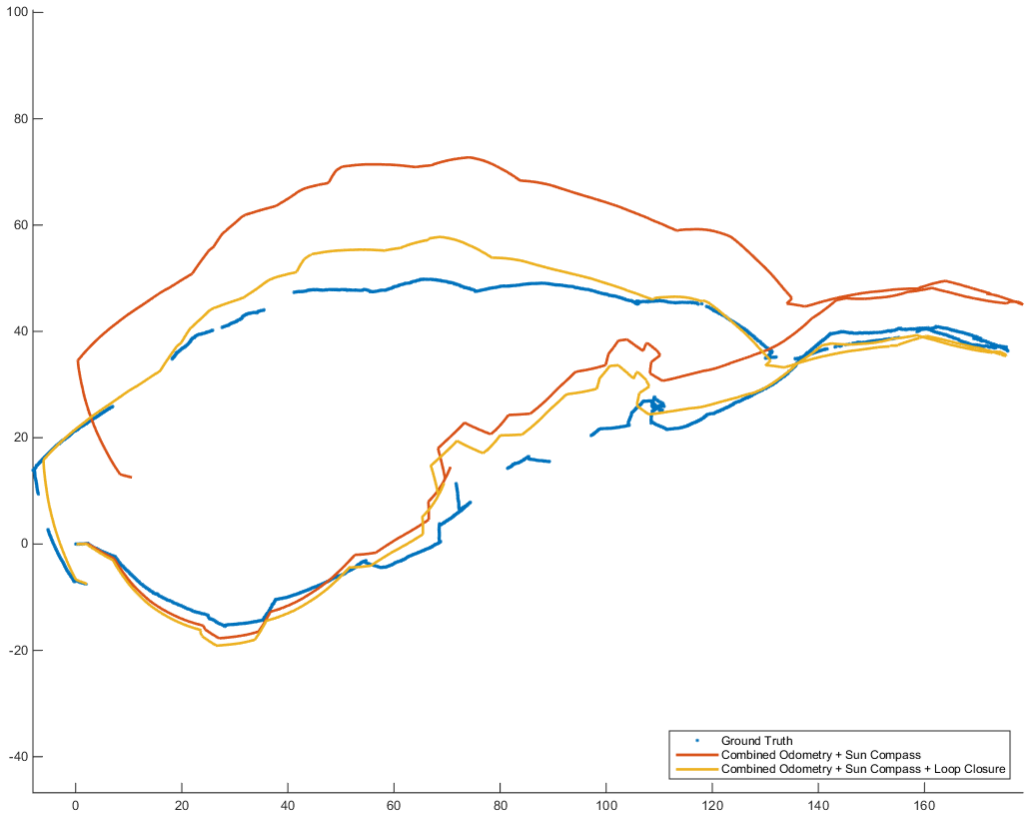

Figure 25: A single point of loop closure results in an 8m maximum error over a 500m traverse = 1.6% error

5.4. Comparison with Priors

Similar work to this research is that of Corke et al. [12]. Their approach defines a function that estimates the displacement of a pixel given the pixel's coordinate, rover's motion, and camera intrinsic parameters, defines a cost function to be the median of the error norm between estimated and observed pixel displacement, and then optimizes over Δx, Δy, $\Delta \theta$, and 5 camera intrinsics. Thus, this optimization takes place in a space of 3 + 5 = 8, which is both time consuming and does not guarantee the optimal or correct solution even when imposing constraints on the optimization problem. In contrast, the 2D point alignment using SVD is a well-conditioned problem and always finds the correct optimal solution.

This research yields visual odometry results and errors similar to that of Seegmiller [15], but does not attempt to tackle the problems of high speed driving and robustness to self-shadowing. Incorporation of the solutions to those problems in this research is straightforward and may improve results. In both cases, the largest source of odometry error in this method comes from breaking the planar terrain assumption, where this research corrects for using absolute bearing.

Furthermore, assuming feature extraction time is constant across feature-based methods, those that utilize the five-point algorithm or its variants also perform SVD, but SVD is performed on a much smaller matrix in the algorithm presented in this research. While this research does not quantify the algorithmic efficiency of the approach presented, it is not unreasonable to believe that this is one of the most lightweight visual odometry methods for in-plane motion estimation.

6. Conclusions and Future Work

This research presents a novel approach for visual odometry that is shown in several field experiments to perform very well in generally planar environments and adequately in those that do not satisfy the planar assumption. This approach is able to exploit the geometric advantages of a downward-facing monocular fisheye camera over pan-tilt stereo for odometry and results in a simple and efficient algorithm not yet used on planetary rovers. This is in stark contrast to the visual odometry systems of existing rovers which use complex pan-tilt stereo camera heads and algorithms that cannot be used all the time due to limited computation and power budgets. The exact accuracy and runtime of this approach as well as its impact on route determination error need to be more fully characterized and quantified in future research. The effect of different feature descriptors and weighting regions should also be further explored.

In addition, a low-rate measure of absolute bearing using a sun compass was demonstrated to significantly improve route determination fidelity. A method of combining these various measurements to estimate route using pose graph optimization was shown to produce accurate results and be practical to offload computation if necessary due to the low data transfer required. Still, whether or not pose graph optimization is feasible onboard the rover should be investigated.

The results achieved with the methodology for utilizing fisheye odometry and absolute bearing from the sun for planetary route determination opens new possibilities for high-cadence,

risk tolerant, minimalist rovers. Under certain conditions, route reconstruction accuracy is on par with or better than that of the Mars Exploration Rovers and Mars Science Laboratory (not including ground-based methods) without the need for stereo visual odometry algorithms or the hardware, power, and computation to support it. Furthermore, the possibilities of utilizing a single set of optics for both odometry and sun sensing for destinations with grazing sunlight – such as the lunar poles – is worth investigating for further simplification aligned with minimalist rover ideals. These research developments lay the groundwork that will enable safer, faster, and smarter navigation for the planetary rovers of tomorrow.

7. Acknowledgements

I would like to thank my advisor Red Whittaker for his support and mentorship, in research and in life. My committee members David Wettergreen and Chris Cunningham have been invaluable with their advice and insight on my research. Thanks to Curt Boirum, Dan Arnett, Chuck Whittaker, and other members of the Planetary Robotics Lab for spending long hours in the field doing anything needed to help get the job done. Finally, thanks to my family for giving me opportunities throughout life and for all the ways they support me.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1252522.

7. Bibliography

[1] M. Maimone, Y. Cheng and L. and Matthies, "Two Years of Visual Odometry on the Mars Exploration Rovers," *Journal of Field Robotics,* 2007.

[2] Y. Chang, M. Maimone and L. Matthies, "Visual Odometry on the Mars Exploration Rovers," *Systems, Man and Cybernetics, 2005 IEEE International Conference on.,* 2005.

[3] R. Deen, "Mars Science Laboratory Software Interface Specification Release 1.2," *Pasadena, CA: Jet Propulsion Laboratory,* 2016.

[4] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature Vol. 293,* 1981.

[5] D. Nistér, "An efficient Solution to the Five-Point Relative Pose Problem," *Pattern Analysis and Machine Intelligence, IEEE Transactions on.,* 2004.

[6] D. Nistér, O. Naroditsky and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics,* 2006.

[7] S. Segvic, G. Schweighofer and A. Pinz, "Performance evaluation of the five-point relative pose with emphasis on planar scenes," 2007.

[8] B. Liang and N. Pears, "Visual Navigation Using Planar Homographies," *Robotics and Automation, 2002. ICRA '02. IEEE International Conference on.,* 2002.

[9] H. Wang, "Visual Odometry Based on Locally Planar Ground Assumption," *Information Acquisition, 2005 IEEE International Conference on. IEEE,* 2005.

[10] P. Baker, "New eyes for robotics," *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on.,* 2003.

[11] R. Vassallo, J. Santos-Victor and H. Schneebeli, "A General Approach for Egomotion Estimation with Omnidirectional Images," *Omnidirectional Vision, 2002. Proceedings. Third workshop on. IEEE,* 2002.

[12] P. Corke, D. Strelow and S. Singh, "Omnidirectional Visual Odometry for a Planetary Rover," *Intelligent Robots and Systems, 2004. (IROS 2004), Proceedings, IEEE/RSJ International Conference on.,* 2004.

[13] D. Scaramuzza and R. S, "Appearance-Guided Monocular Omnidirectional Visual Odometry for Outdoor Ground Vehicles," *Robotics, IEEE Transactions on.,* 2008.

[14] D. Scaramuzza, F. Fraundorfer and R. Siegwart, "Real-Time Monocular Visual Odometry for On-Road Vehicles with 1-Point RANSAC," *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on.,* 2009.

[15] N. Seegmiller and D. Wettergreen, "Optical Flow Odometry with Robustness to Self-shadowing," *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on.,* 2011.

[16] J. Engel, T. Schöps and D. Cremers, "Semi-Dense Visual Odometry for a Monocular Camera," *Computer Vision, 2013. (ICCV 2013), IEEE International Conference on.,* 2013.

[17] D. Caruso, J. Engel and D. Cremers, "Large-Scale Direct SLAM for Omnidirectional Cameras," *Intelligent Robots and Systems, 2015. (IROS 2015), IEEE/RSJ International Conference on.,* 2015.

[18] C. Boirum, "Improving Localization of Planetary Rovers with Absolute Bearing by Continuously Tracking the Sun," *tech. report CMU-RI-TR-16-01, Robotics Institute, Carnegie Mellon University,* 2016.

[19] D. Scaramuzza and F. Fraundorfer, "Visual Odometry: Part I – The First 30 Years and Fundamentals," *IEEE Robotics and Automation Magazine,* 2011.

[20] F. Labrosse, "The visual compass: Performance and limitations of an appearance-based method," *Journal of Field Robotics,* 2006.

[21] D. Scaramuzza, A. Martinelli and R. Siegwart, "A Toolbox for Easily Calibrating Omnidirectional Cameras," *Intelligent Robots and Systems, 2006. (IROS 2006), IEEE/RSJ International Conference on.,* 2006.

[22] F. Labrosse, "Appearance-based heading estimation: the visual compass," *Technical Report UWADCS-06-048, Computer Science Department, University of Wales, Aberystwyth, UK,* 2006.

[23] G. Grisetti, "g2o: A general framework for graph optimization," *IEEE International Conference on Robotics and Automation,* 2011.