Mars Microrover Navigation:
Performance Evaluation and Enhancement[*]

Larry Matthies, Erann Gat, Reid Harrison,
Brian Wilcox, Richard Volpe, and Todd Litwin
Jet Propulsion Laboratory - California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109

Abstract

In 1996, NASA will launch the Mars Pathfinder spacecraft, which will carry an 11 kg rover to explore the immediate vicinity of the lander. To assess the capabilities of the rover, as well as to set priorities for future rover research, it is essential to evaluate the performance of its autonomous navigation system as a function of terrain characteristics. Unfortunately, very little of this kind of evaluation has been done, for either planetary rovers or terrestrial applications. To fill this gap, we have constructed a new microrover testbed consisting of the Rocky 3.2 vehicle and an indoor test arena with overhead cameras for automatic, real-time tracking of the true rover position and heading. We create Mars analog terrains in this arena by randomly distributing rocks according to an exponential model of Mars rock size frequency created from Viking lander imagery. To date, we have recorded detailed logs from over 85 navigation trials in this testbed. In this paper, we outline current plans for Mars exploration over the next decade, summarize the design of the lander and rover for the 1996 Pathfinder mission, and introduce a decomposition of rover navigation into four major functions: goal designation, rover localization, hazard detection, and path selection. We then describe the Pathfinder approach to each function, present results to date of evaluating the performance of each function, and outline our approach to enhancing performance for future missions. The results show key limitations in the quality of rover localization, the speed of hazard detection, and the ability of behavior control algorithms for path selection to negotiate the rock frequencies likely to be encountered on Mars. We believe that the facilities, methodologies, and to some extent the specific performance results presented here will provide valuable examples for efforts to evaluate robotic vehicle performance in other applications.

Submitted to the Autonomous Robots Journal.

Contents

List of Figures

List of Tables

1   Introduction

In planetary exploration, cost constraints are shrinking the size of spacecraft and reducing the scope of missions that can be undertaken. For Mars rovers, the result has been a shift from the large rovers and long missions envisioned in the 1980's to the small rovers and short missions planned now. For example, the Mars Rover Sample Return (MRSR) mission scenario envisioned a 1000 kilogram (kg) rover that would navigate autonomously for 100's of kilometers (km) in a mission lasting much of a year [1]; in contrast, the Mars Pathfinder mission, scheduled to launch in 1996, will carry an 11 kg rover that will move at most a few 100 meters (m) in a mission lasting 1 to 4 weeks. Future Mars missions must broaden the range of exploration within even more stringent mass constraints; hence, rover technology development must put more capability into lighter, smaller packages.

For this development to proceed effectively, we must have a quantitative understanding of both the requirements of future missions and the limitations of current rovers. Unfortunately, quantitative information about the navigation performance of current rovers is severely lacking. This is true not just for Mars rovers, but for robotic vehicles in general. Robotic vehicles, large and small, have been built, tested, and used for a number of applications [2, 3, 4]. However, most published performance data is anecdotal, in that it reports speeds attained, distances covered, or missions performed over a few test courses; it does not quantify reliability as a function of sensor suite or terrain characteristics. Although performance evaluation for obstacle detection sensors is advancing [5, 6, 7], there has been little experimental evaluation of terrain traversability; in particular, we have almost no knowledge of actual failure rates. This stems from both a lack of models to characterize performance and, even more so, a lack of the large number of experimental trials needed to measure failure rates and validate performance models. Such information is necessary to establish the reliability of given rover designs, as well as to trade-off design alternatives when size constraints strongly limit the sensing and computing capabilities that can be placed onboard.

To redress this deficiency, we are developing methodologies and experimental facilities to enable systematic evaluation of rover navigation performance. An important tool in this work is a new microrover test arena that uses overhead cameras to provide automatic, real-time tracking of the true rover position and heading. Mars analog terrains are created in this arena by distributing rocks according to parameterized models of rock distributions at the Viking lander sites. We are using this facility to quantify rover navigation performance as a function of rock frequency and have recorded detailed logs of over 85 navigation trials to date. This work has taken important steps forward in defining performance evaluation methodologies for rover navigation and in measuring the performance of current rover prototypes. This is a key step in validating rover navigation performance for the 1996 Pathfinder mission and in guiding future rover development.

This paper is organized as follows. Section 2 gives an overview NASA's Mars Exploration Program, including a description of the lander and rover being built for the Mars Pathfinder mission and a discussion of how rover requirements are likely to evolve for future missions. Section 3 describes the facilities built for rover performance evaluation and introduces a decomposition of the rover navigation task into the functions of goal designation, rover localization, hazard detection, and path selection. For each of these functions in turn, sections 4, 5, 6, and 7 describe the approach being taken for that function by the Pathfinder mission, our approach to evaluating performance of that function, and the results of evaluation to date. Section 8 summarizes the principal results and our plans for future work.

2   Overview of Mars Exploration Program

Current plans call for two to three launches every launch window (roughly every 25 months), beginning in November, 1996. The missions for 1996 are the Mars Global Surveyor orbiter and the Mars Pathfinder lander. Follow-on missions are anticipated in 1998, 2001, and 2003. In 1998, the spacecraft will be a second orbiter and a lander; subsequent missions are still being defined. The missions in 1998 and beyond are likely to include international collaboration; whether or not rovers are part of these missions depends on scientific, technical, and international programmatic trade-offs that have yet to be made.

2.1   Mars Pathfinder Mission

Mars Pathfinder is a new spacecraft designed to be affordable enough to enable the frequent launches described above. It will perform a "hard" landing, using a heat shield, parachute, and air bags to decelerate and to cushion the impact of landing, instead of the usual "soft" landing performed with thrusters. The lander will carry a small rover, known as the Microrover Flight Experiment (MFEX), a stereo camera pair, and other scientific instruments. The total launch mass of the spacecraft, including the rover, is 712 kg. The primary goals of the Pathfinder mission are to develop this spacecraft and to demonstrate that it can serve as the basis for subsequent Mars missions.

The chosen landing site, known as Ares Tiu, is at the confluence of the Ares and Tiu valleys, which are believed to be catastrophic flood channels. This site was chosen because it is expected to contain rocks from several geologic units, which would have been swept along the flood channels and deposited in this area. Remote sensing data indicate that this site has a rock abundance comparable to that observed at the Viking Lander 2 site. This information is being used to define terrain models for rover navigation testing, as discussed in section 7. Pathfinder is expected to land in July, 1997. The rover primary mission lasts one week, with a goal of four weeks; the lander primary mission is one month, with a goal of one year.

2.1.1   Key Lander Characteristics

The lander is a tetrahedron approximately 1.5 m wide and 1.0 m high; its mass is 264 kg, including a 23 kg payload of the rover, stereo cameras, and other science instruments (figure 1). Three sides of the tetrahedron are "petals" that open with actuators strong enough to push the lander upright, regardless of which side is originally down. The lander computer is the RAD 6000, a radiation-hardened version of the IBM RS 6000, with 128 MB of mass memory on a VME backplane, running the VxWorks operating system. Imaging uses a stereo camera pair with a pan/tilt mount on a mast 1.5 m above the ground. A filter wheel provides 12 spectral channels in the range from 450 to 1050 nm. In addition to scientific observations, this imaging system will be used for designating rover goal locations and for updating the rover position; similar imaging systems are likely to be part of future landers. The cameras have a stereo baseline of 15 cm, CCD imagers with 256×256 pixels, and an angular resolution of 0.001 radian per pixel. The maximum data communication rate from the lander to Earth is 1260 bits/second with a latency of 11 minutes; data transmission will take place for roughly 2 to 3 hours per day.
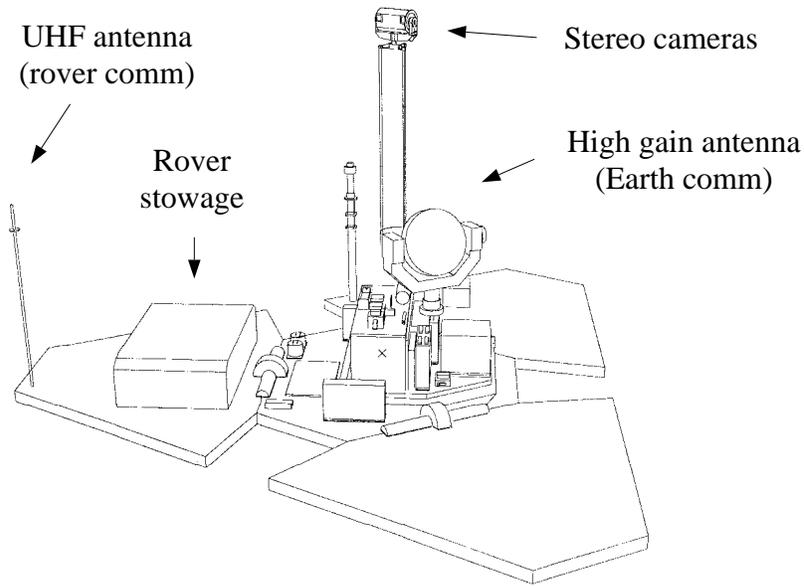
UHF antenna
(rover comm)

Stereo cameras

Rover
stowage

High gain antenna
(Earth comm)

Figure 1: Pathfinder lander in deployed configuration

2.1.2 Key Rover Characteristics

The rover is 65 cm long, 48 cm wide, and 30 cm high; its mass is 11.5 kg, including 1.3 kg of science instruments (figure 2). The chassis is a passive rocker bogey mechanism designed to allow the rover to climb vertical steps of 1.5 wheel diameters; the wheels in this case are 13 cm in diameter. There are six drive motors and four steering motors. The rover computer is an Intel 8085 with 0.5 MB of mass memory; its software is a custom executive with a single thread of control. Obstacle detection is done with a light-stripe triangulation system, angle encoders on the bogeys, and other sensors. Science instruments include an alpha/proton/x-ray spectrometer (APXS), which is on a hinge mechanism at the rear of the rover, and a dust adhesion sensor. The rover communicates with the lander over a UHF radio with a clear-field radial range of 500 m and a bandwidth of 2000 bits/second. Primary power for the rover comes from solar cells, which have a peak output of 16 W at noon; non-rechargeable batteries provide power through the night. Power consumption by subsystem is illustrated in table 1. The power budget is sufficiently tight that only one major function (eg. driving, steering, transmitting, and hazard detection) is performed at a time.

   The rover mission will be controlled by human operators on Earth, who will use stereo image pairs from the lander to designate waypoints leading to places where experiments are to be performed. Between waypoints, the rover will use onboard hazard detection sensors and behavior control algorithms to avoid obstacles (sections 6 and 7). The rover will operate for approximately four hours per day, centered around noon, due to solar power and temperature issues. Also for power and temperature reasons, the driving speed is limited to 0.67 cm/sec; hence, the maximum traverse distance per day would be about 95 m if the rover drove continuously. Since the rover will stop frequently to do hazard detection, which itself is a slow process, the daily traverse will be limited to under 40 m. At the end of each day, the lander will transmit a stereo image pair of the rover to Earth; human operators will use these images to update the position and orientation of the rover prior to the next day's operation. The nominal rover mission includes taking APXS

3
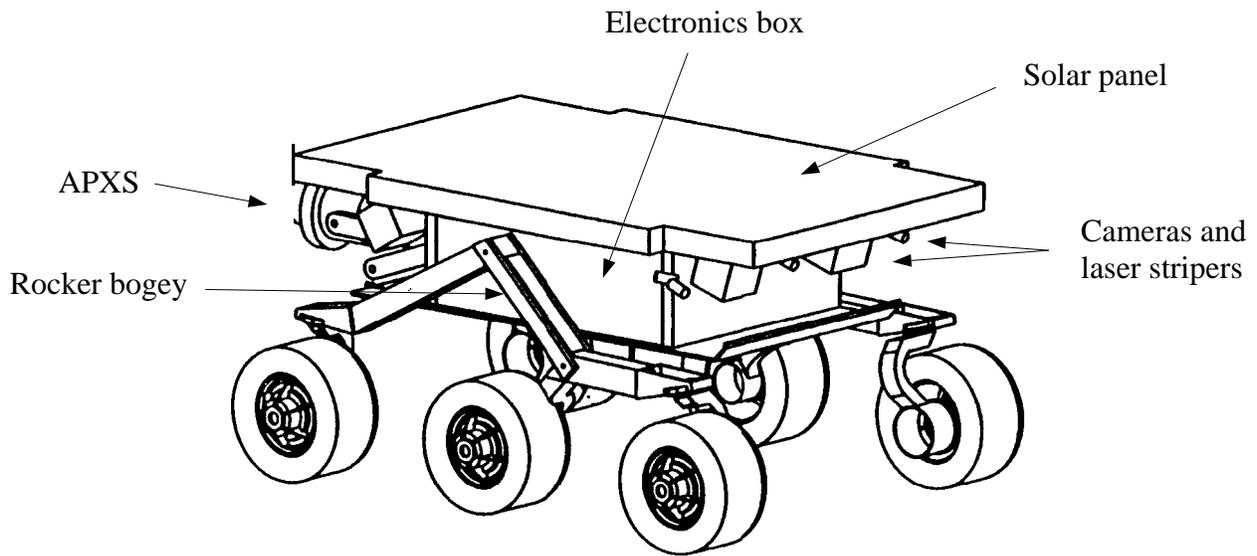
Figure 2: Pathfinder rover

Subsystem	Power usage (W)
CPU and I/O	3.77
CCD's and clocks	0.75
Gyro	0.80
Accelerometers	0.82
Light stripes (each)	1.0
Modem (transmitting)	2.50
Driving, -40C, extreme terrain	8.28
Steering, -40C, extreme terrain	6.70

Table 1: Best estimate power usages for major subsystems of MFEX rover, as of May, 1994.

Sol	Traverse distance (m)	Activity
1	3	Deploy from lander
2	3	Approach a rock
4	10	Image the lander
8	40	Test radio every 10 m
18-21	10 to 30 m/day	Image the landing path

Table 2: Pathfinder rover nominal operational scenario for selected days. A "sol" is a Martian day, which is 24 hours, 37 minutes. The first seven sols constitute the primary mission.

spectra from rocks, making dust adhesion measurements, conducting soil mechanics experiments by spinning its wheels, testing the performance of the radio as a function of the distance from the lander, and imaging its own tracks, the lander, and the impact trail made by the lander and airbags (table 2).

2.2 Objectives of Potential Follow-on Missions

Current plans for U.S. missions in 1998 include a second Mars Global Survey orbiter and a lander. The payload for the lander has not been decided. For cost reasons, plans for the lander call for a smaller launch vehicle than for Pathfinder, which would force the lander itself to be as much as 50% lighter than Pathfinder. International missions planned for 1998 are a Japanese orbiter and a Russian lander with a rover and a balloon. Plans beyond 1998 are less well defined, but include the possibility of multiple landers per launch vehicle; this would require each lander to be lighter still. Exploration objectives for these missions include landing at higher latitudes than Pathfinder, exploring the polar ice caps, and carrying a broader suite of instruments for sample acquisition, sample analysis, and seismic and meteorological observation.

In this general picture, rovers compete for payload allocation with other instruments and instrument deployment methods; hence, rover development must recognize the need to offer higher pay-off than alternative payloads. This includes designing rovers to provide adequate mass and power budgets for rover-mounted instruments, equipping rovers with power and thermal control systems adapted to the temperatures and solar illumination conditions of higher latitudes, and making rovers light enough for the ever smaller landers. In the area of rover navigation, it also includes increasing the speed and navigation range of rovers, to enable a greater number of science operations per day and exploration over much greater distances from the lander. To effectively address these requirements, we first need a firm assessment of the limitations of current rovers. The balance of this paper presents our first step in this assessment.

3 Navigation Functions, Evaluation Methodology, and Facilities

We decompose the task of rover navigation into the following four functions:

- Goal designation. (Where does the rover need to end up?)

- Rover localization. (Where is the rover now?)

- Hazard detection. (Where are the obstacles?)

- Path selection. (Which clear path should the rover take to the goal?)

Inevitably, each of these functions is performed with some level of uncertainty. Therefore, to understand overall rover navigation performance, we must understand the uncertainties associated with each function. This is reflected in our evaluation methodology, which assesses the performance of each function separately, then tests the overall navigation system in Mars analog terrain.

Our testbed vehicle is Rocky 3.2, a rebuilt version of Rocky 3 [8] that has essentially the same chassis design, size, computer, and sensor suite as are planned for MFEX. Principal differences between Rocky 3.2 and MFEX that are pertinent to navigation performance are that Rocky 3.2

Figure 3: The indoor test arena with Rocky 3.2 vehicle. Monitors in the corner display images from the overhead tracking cameras. The mural is an image of Mars from Viking Lander 1.

weighs more (about 20 kg), uses rubber instead of metal tires, and lacks some of the contact and motor current sensors planned for MFEX. Testbed facilities also include a mock-up lander of approximately the same size as the Pathfinder lander, with color stereo cameras[1] on a pan/tilt mount and the same stereo baseline and angular resolution as the Pathfinder camera system.

Indoor testing is conducted in a 4×12 m arena filled with landscaping sand mixed with crushed brick to provide a reddish hue (figure 3). Rocks of various sizes and types are placed in the arena to provide obstacles. To create Mars analog terrain, these rocks are selected according to size distributions that have been estimated for rocks in various Martian terrain types (see section 7). During navigation trials, the true rover position and heading can be tracked with a set of four cameras mounted in the ceiling above the arena. The tracking system locates a black and white target placed on top of the rover and runs at 15 Hz on a Motorola 68040 CPU. Since tracking is done monocularly, the rover 3-D position is determined from the measured image position using prior knowledge of the heights of the rover and the cameras. Systematic tracking errors (ie. due to camera calibration) are less than 2 cm in position and 0.4 degrees in heading; random tracking errors (ie. repeatability) are approximately 0.3 mm RMS error in rover position and 0.2 degrees RMS error in rover heading. For outdoor testing, the rover, lander, and a Sun workstation are hauled to the test site in a small trailer equipped with a gasoline-powered generator and an air conditioner; at the test site, the trailer serves as a portable operator control station and data-logging facility. Outdoor tests have been conducted in the Arroyo Seco, next to JPL.

These facilities are being used to evaluate performance for each of the four navigation functions listed above. In the following sections, for each function we outline the approach being taken for the Pathfinder rover, the evaluation approach and results to date, and the approaches being taken to enhance performance.

---

[1] Sony XC-999 single-chip color cameras.

4   Goal Designation

The Pathfinder approach to goal designation is for mission operators on Earth to specify 3-D waypoints using stereo imagery from the lander. Waypoints will be specified using a stereographic display to view the scene, a spaceball to input 3-D coordinates, and subpixel cross-correlation to refine the operator's estimate of the distance to the waypoint. The waypoint coordinates are defined relative to the lander coordinate frame and uplinked to the rover. The rover also models its own position and plans its steering maneuvers in the lander coordinate frame.

Waypoint coordinates are subject to systematic errors caused by miscalibration of the cameras, mast, and pan/tilt axes and to random errors caused by noisy estimation of the stereo disparity of the waypoint. Experimental evaluation of the likely size of these errors has not yet been undertaken; however, reasonable bounds can be established from prior experience with similar camera systems. For convenience, we express waypoint uncertainty in terms of the crossrange and downrange errors in the estimated coordinates, as a function of the true range to the waypoint. For random errors, we assume conservatively that waypoints are designated in the image with a standard deviation of one pixel and that the stereo disparity of waypoints is estimated with a standard deviation of 0.1 pixel [7]. As noted in section 2.1.1, the angular resolution of the lander cameras is 0.001 radian/pixel. Therefore, these assumptions about image plane designation error imply approximate crossrange and downrange uncertainties of [7]:

$$\sigma_c \approx 0.001\ R \tag{1}$$
$$\sigma_d \approx 0.0001\ R^2/B\ , \tag{2}$$

respectively, where R is the true range and B is the stereo baseline (15 cm). Systematic errors, in the form of translational and rotational errors in the location of the camera in the lander coordinate frame, will compound the designation uncertainty. Of these, rotational errors are likely to be most significant. To a first approximation, rotational errors introduce only crossrange error in the waypoint coordinates, which take the same form as equation (1):

$$\Delta_c = \Delta_\theta\ R \tag{3}$$

If the orientation of the cameras is known only to within one degree ($\Delta_\theta = 0.017$ radian), then the systematic error in the camera orientation greatly dominates the crossrange uncertainty due to the angular resolution of the camera. Figure 4 shows the random and systematic errors predicted by (1) through (3).

The meaning of these results depends on the type of goal the rover is trying to reach, the accuracy of rover localization, and the ability of the rover to overcome localization errors by recognizing the goal as it draws near. For the goal of placing a spectrometer against a rock, the size, spacing, and individual recognizability of rocks become relevant concerns. These issues are examined in subsequent sections. For now, we simply note that figure 4 predicts that designation error will be less than half the width of the vehicle (48 cm) within the range of the rover primary mission, or roughly 10 m from the lander (see table 2); hence, designation performance is likely to be adequate within that range. For distances beyond 20 m, the combined designation error is likely to exceed the dimensions of the vehicle, making it difficult for the rover to find goal rocks at such ranges based on 3-D coordinates alone. We discuss potential solutions to this problem at the end of the following section, after examining rover localization performance.
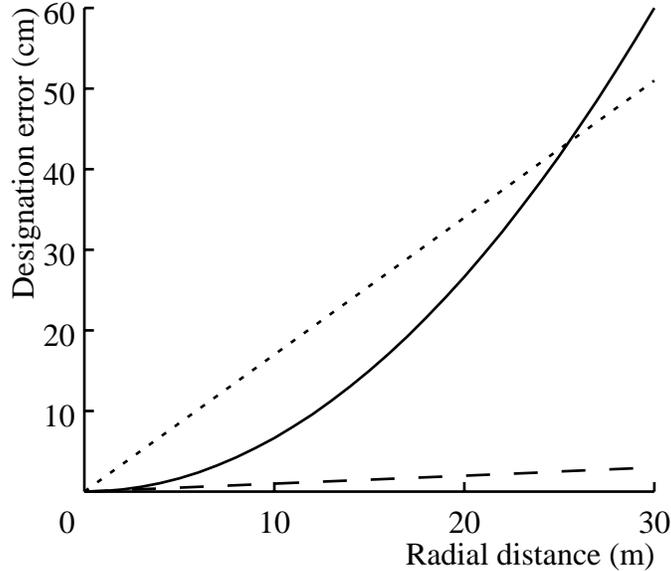
Figure 4: Designation error vs. distance. Solid: standard deviation of downrange random error; dashed: standard deviation of crossrange random error; dotted: crossrange systematic error assuming one degree miscalibration of pan angle.

5 Rover Localization

1

The Pathfinder approach to rover localization combines deadreckoning with a daily location update mediated by mission operators on Earth. Deadreckoning is performed with wheel encoders and a solid state turn rate sensor[2]. Since these sensors are subject to drift, the rover location is updated once daily by imaging the rover from the lander, sending the images to Earth, localizing the rover by interactively registering a graphical rover model to the imagery, and uplinking the observed rover position and heading prior to the next day's operations.

Deadreckoning performance is being evaluated by measuring and extrapolating the performance of Rocky 3.2, which uses the same deadreckoning sensors as MFEX. Since deadreckoning performance is a function terrain characteristics, we have established a base case by measuring deadreckoning error with the rover traversing level sand with no rocks or other obstacles. Error in heading (θ) and total distance traveled (r) was estimated as the difference between the deadreckoned estimates and those measured by the overhead tracking system. Twenty runs of approximately 7 meters each were observed, consisting of ten runs each direction up and down the course. Assuming that the dominant noise sources for r and θ are random wheel slip and random walk in the turn rate sensor, respectively, and that these sources manifest themselves as additive, independent errors over time, we expect the error variances for r and θ to grow linearly with the true distance traveled. Figure 5 shows the error variances measured from the twenty runs. Heading error shows the expected linear trend, with a best fit slope of 0.70 deg^2/m. Since error in the turn rate sensor is a function of time, this result is a function of the vehicle velocity. For Rocky 3.2 the velocity is 15 cm/sec, whereas for MFEX it is 0.67 cm/sec; therefore, directly scaling the heading error

---

[2]The QRS-11 from Systron Donner Co.
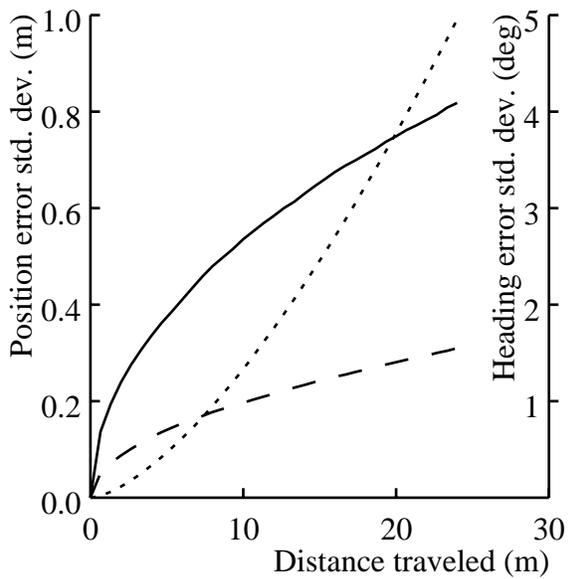
8
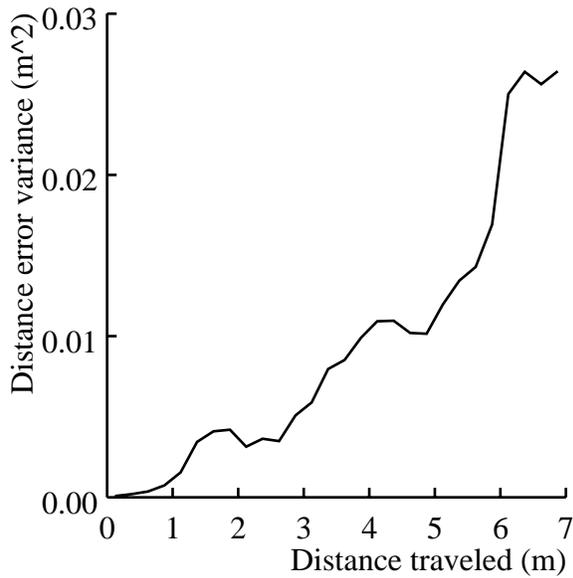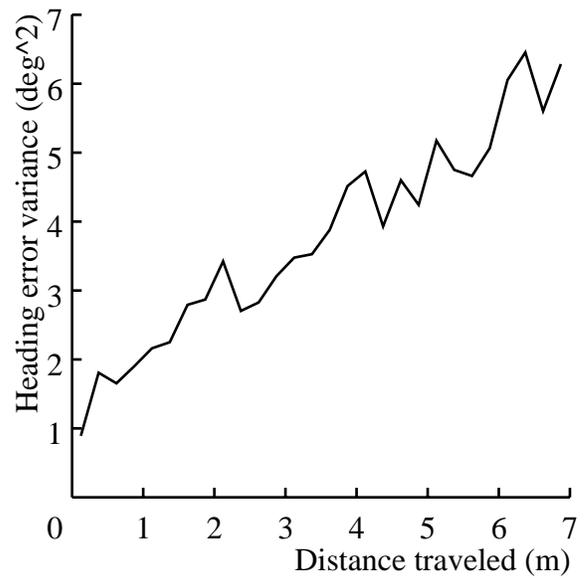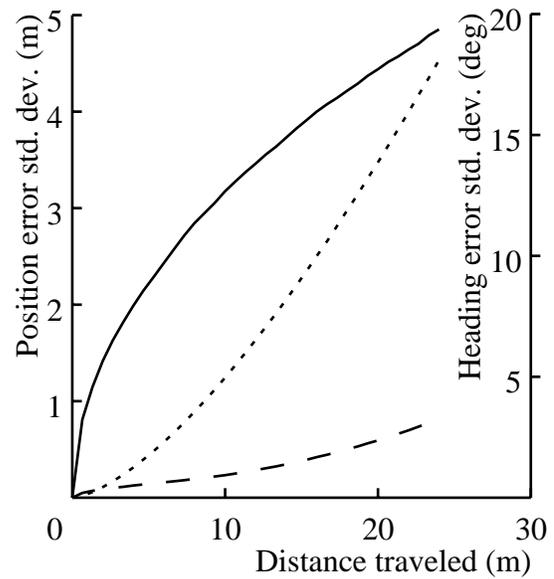
Figure 5: Position and heading error vs. distance traveled, as measured with Rocky 3.2 on sand

Figure 6: Extrapolations of position and heading error vs. distance traveled: (a) using variances for Rock 3.2 on sand, (b) using heading variance extrapolated to MFEX velocity.

to the MFEX velocity produces a variance of $0.70 * 15/0.67 \approx 16$ deg^2/m. It remains to be seen whether this rate will be observed in practice[3]. The variance of the distance error (figure 5a) appears to grow faster than linearly. The reason for this has not yet been determined; for the time being, we continue to assume a linear trend and use the best fit slope of 0.0039 m^2/m obtained from this data. We assume that factors contributing to this error, such as wheel slip, are not a function of the vehicle velocity.[4]

To predict deadreckoning error over longer distances, we used the measured results to define noise terms in the following rover state equations:

$$
\begin{aligned}
x_{i+1} &= x_i + \Delta r_i \cos \theta_i \\
y_{i+1} &= y_i + \Delta r_i \sin \theta_i \\
\theta_{i+1} &= \theta_i + \Delta \theta_i
\end{aligned}
$$

That is, we assume that at each time step i the rover moves straight in the current direction for distance Δr_i, then turns in place through angle $\Delta \theta_i$. Monte Carlo simulations were run with $\Delta r_i = 6.7$ cm and $\Delta \theta_i = 0$ deg; at each time step, Δr_i and $\Delta \theta_i$ were corrupted by Gaussian white noise with variances obtained by scaling the empirical results above to 6.7 cm per step. Figure 6a shows the standard deviation of x, y, and θ as a function of the mean distance traveled, using the error variances measured for Rocky 3.2. Error in the lateral direction (y) grows much more quickly than error in the forward direction (x), because over moderate distances heading error affects lateral position to a much greater degree than forward position. Figure 6b shows the results using the heading error variance extrapolated to the MFEX velocity.

The significance of these results is what they suggest about the ability of the rover to reach its intended goal. Consider first the results for Rocky 3.2 shown in figure 6a. At 10 m, the standard deviation of lateral error is 27 cm, or about half the width of the vehicle. The standard deviation of forward error is 20 cm. From figure 4, designation errors at this distance are likely to be less than 17 cm. Hence, for a rock of any reasonable size within a 10 m traverse, Rocky 3.2 should be able to reliably stop near enough the rock to find it with contact sensors or the light stripe sensor. This was confirmed by success in a dozen outdoor trials conducted in September, 1994, in which goal rocks were typically about 5 m from the lander. After 20 m of total traverse, figure 6a predicts the standard deviations to be 76 cm for lateral position and 29 cm for forward position; designation errors are likely to be in the 20 to 30 cm range. This is enough error that the light stripe sensor and/or local search are likely to be necessary to reliably find the goal. When we consider the results obtained with heading error scaled to the MFEX velocity (figure 6b), at 10 m the standard deviations are 125 cm for lateral error and 24 cm for forward error. This suggests that reaching a goal 10 m from the lander will require external updates of rover position and heading to cancel deadreckoning error. At 5 m, the corresponding values are 45 cm and 15 cm. At this distance, it is likely that the rover will be able to deadreckon close enough to goal rocks to find them with the hazard detection sensor.

The results so far are based on deadreckoning in roughly straight lines on level sand with no rocks; hence, the experiments need to be extended to cases where the rover is turning and/or driving

---

[3]The relative importance and the scaling behavior of various factors contributing to heading error are unknown at present (eg. vehicle vibration effects on the turn rate sensor).

[4]Rocky 3.2 deadreckoning uses a scale factor that was calibrated to account for wheel slip on level sand. For these runs, the minimum, mode, and maximum errors in distance traveled were -4%, 0.2%, and 10%, which indicate good calibration.

over rocks. For example, figure 7 shows a trace of vehicle position for a run involving three rocks. Figure 8 shows heading as a function of distance traveled for this run, with arrows marking points at which the rover thought it saw obstacles. The first arrow corresponds to a false alarm, the second to detection of rock A, and the third to detection of rock C. Rock B was not detected on this run; however, the rover scraped its hubs considerably against rock B and scraped slightly against rock C. The deviation at 2.6 meters into the run from near 0 degrees to near -5 degrees of heading error occurred while the rover turned in place to avoid rock A. The drift from -5 to $+10$ degrees heading error occurred while the rover executed an arcing turn to the right and scraped against rock B. Heading error stayed around $+10$ degrees while the rover drove in a straight line from rock B to rock C. At rock C, the rover turned in place about 20 degrees, then executed an arcing turn around the rock. Near the end of this turn, heading error jumped abruptly to $+15$ degrees; this corresponded to scraping rock C. Over the 8.1 meter run, the final errors were 1.0 meter in position and 15.3 degrees in heading.

The results of this run imply (1) that the rover systematically underestimates its rate of turning and (2) that scraping against rocks introduces large heading errors. The underlying causes of these problems have not yet been explored, but it is likely that the former is a calibration error and the latter is the result of vibrations transmitted to the turn rate sensor. These results also underscore the importance of further evaluation of deadreckoning performance on rocky ground. We return to this issue in section 7.

Collectively, all of the deadreckoning results demonstrate the need for periodic updating of the rover location from other measurments. As noted earlier, the Pathfinder approach to this is to update localization once a day by interactively matching rover models to imagery sent to Earth. With deadreckoning errors like those observed above, this will limit sorties to specific locations to under 10 m/day. To simplify mission operations and to enable longer traverses, for future missions it is highly desirable to update the rover location automatically, using sensors onboard the rover, the lander, or even an orbiter. Approaches being pursued include:

- integrating a sun sensor on the rover, based on one developed in [9], to provide absolute heading measurements;[5]

- automatically locating the rover in lander imagery, with algorithms running on the lander computer, to provide position and heading measurements when the rover is in view from the lander;

- developing a radio beacon system, with components on both the rover and the lander, to provide lander-relative bearing and range measurements when the rover is beyond visual range from the lander.

Visual localization methods are being developed at JPL and are reported on below. The sun sensor and radio beacon are being developed under contract to JPL by the Martin Marietta Astronautics Group; results of this work will be reported in the future.

As an initial approach to visually locating the microrover in the vicinity of the lander, we placed a cylindrical target on the rover that has different colors in each of its four quadrants [11]. By judicious selection of the colors, segmentation of imagery from the lander can locate the centroid of the cylinder, as well as the visible color quadrants. Using a pair of stereo images, the distance

---

[5]The magnetic field strength of Mars is $\leq 10^{-4}$ of Earth's [10], making compasses impractical.
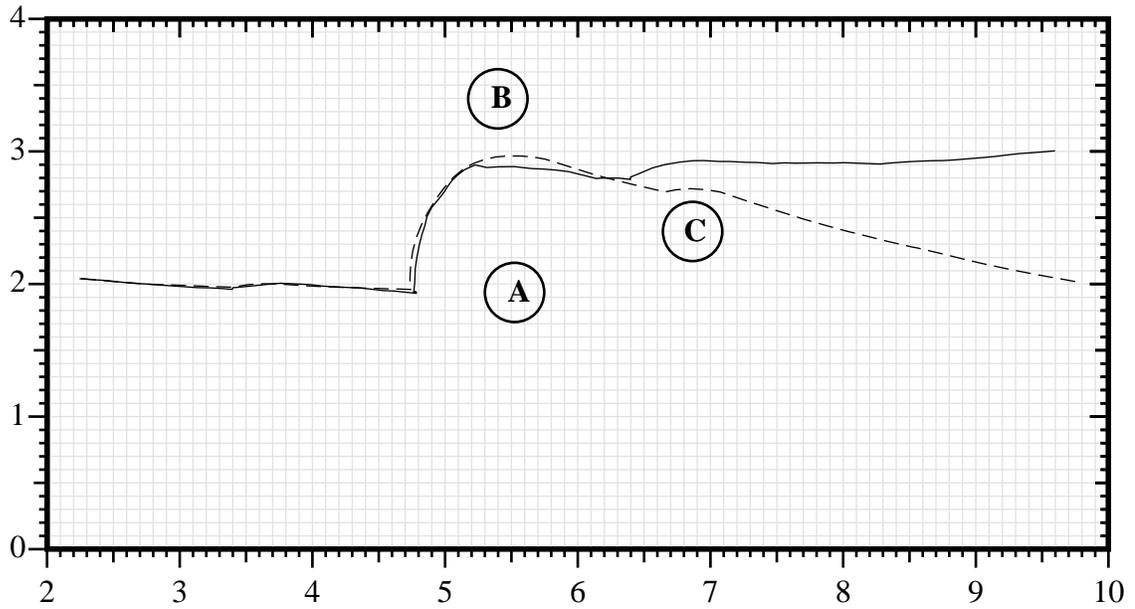
Figure 7: Planview plot of vehicle trajectory for a run with three rocks. Axis units are meters. Solid line: true rover trajectory; dashed line: deadreckoned trajectory. The rover scraped its hubs against rocks B and C.
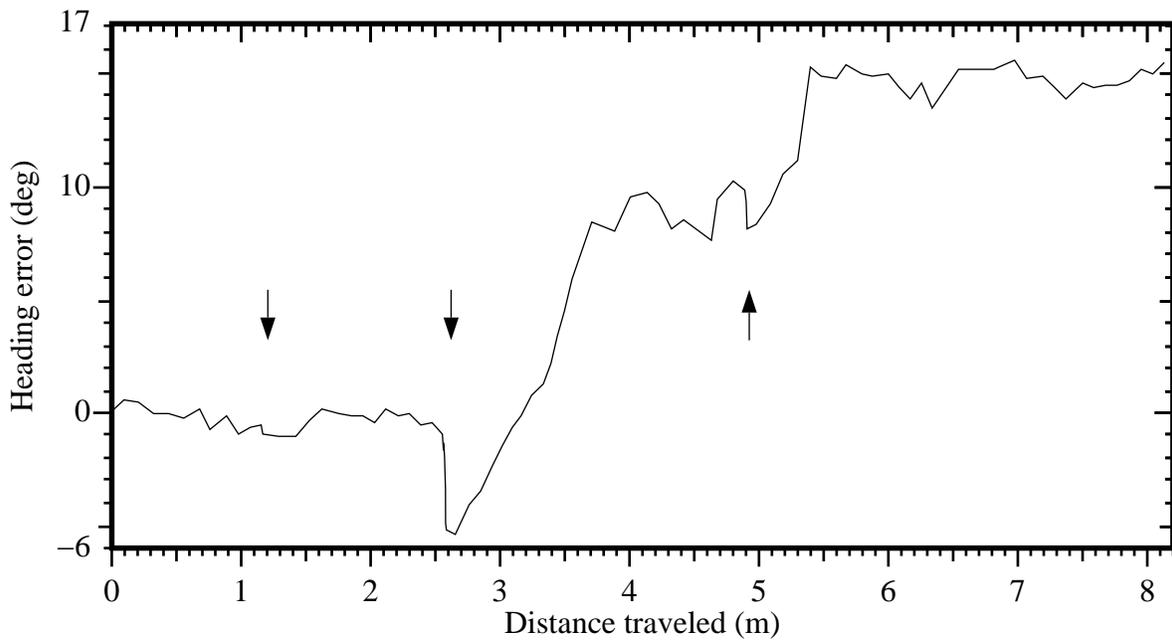

Figure 8: Corresponding plot of heading vs. distance traveled. Arrows show when the rover thought it saw an obstacle. The first arrow was a false alarm, the second was rock A, the third was rock C; rock B was not seen.

to the rover is determined by stereo triangulation on the centroid of the cylinder. The angle of the cylinder, hence the heading of the rover, is determined by analyzing which color quadrants are visible and to what degree. Mathematical details of this approach can be found in Appendix A. In the balance of this section, we discuss experimental results on the accuracy of this localization scheme and relate these results to Pathfinder and future rover missions.

To test this approach, a cylinder 19.2 cm in diameter by 10 cm high was constructed out of fluorescent colored cardboard, with the four quadrants colored blue, pink, orange, and pink again. The colors were originally selected so that adjacent colors have similar spectral response in one channel of the color cameras, helping prevent boundary pixels from being excluded in the segmentation of the cylinder from the full image. To evaluate the accuracy of the approach, the cylinder was placed with its center along the line midway between a pair of Sony XC999 cameras, set up for stereo viewing with a baseline of 12 cm. Calibration procedures indicate that the field of view of the cameras is $29.33°$. Images of the cylinder were collected at $20°$ increments and four different distances. The zero position of the cylinder was aligned with the center of rotation of the cameras by bore-sighting, and the cylinder was rotated with a machinists' turntable. This procedure was repeated at each of the chosen distances of 0.914, 1.829, 4.572, and 10.058 meters (3, 6, 15, and 33 feet, respectively) from the center of rotation of the stereo camera fixture. The front of the cameras was offset from this origin by .076 m (3 inches). Ten measurements were taken at each position/orientation setting.

Figures 9 and 10 show the mean measured angular error and mean measured relative distance error as a function of the actual angle. The means of the angular errors are typically within $\pm 5°$, especially when the cylinder is close. The means of the relative distance measurement errors are within 5% of correct and are are usually short, possibly indicating a small calibration bias. Figures 11 and 12 show the standard deviations of the measured angular positions and normalized distances. The graphed results can be compared against the theoretical errors for single-pixel changes in the cylinder images (see Appendix A), which are shown the following table.

d	$\Delta\theta_{min}$	$\frac{\Delta d}{d}\vert_{stereo}$
0.838 m	0.50°	0.0071
1.753 m	1.06°	0.015
4.496 m	2.70°	0.038
9.982 m	6.09°	0.085

It is apparent that for both orientation and distance, the standard deviations indicate that the segmented images usually deviate by less than one or two pixels. Therefore, once calibrated for bias, this method will provide results that are quite precise and are at the limit of measurement of the given cameras.

Of notable exception to the generally good quality of measurements is the angular error and standard deviation in the $-180 \leq \theta \leq -90$ range for the 10 m data. This is attributed to problems in segmenting the pink and orange quadrants at this distance. Subsequent experiments at the 15 meter range further confirmed this conclusion by providing reasonable results except in this angular domain, where no consistent measurements were possible.

Two improvements to the implementation of this technique would be a selection of different colors and use of a more robust segmentation algorithm. One possibility for new colors would be a simple replacement of the orange panel with a green one. A more strongly contrasting color appears to be desirable since the contrast between the pink and blue enabled more accurate cylinder
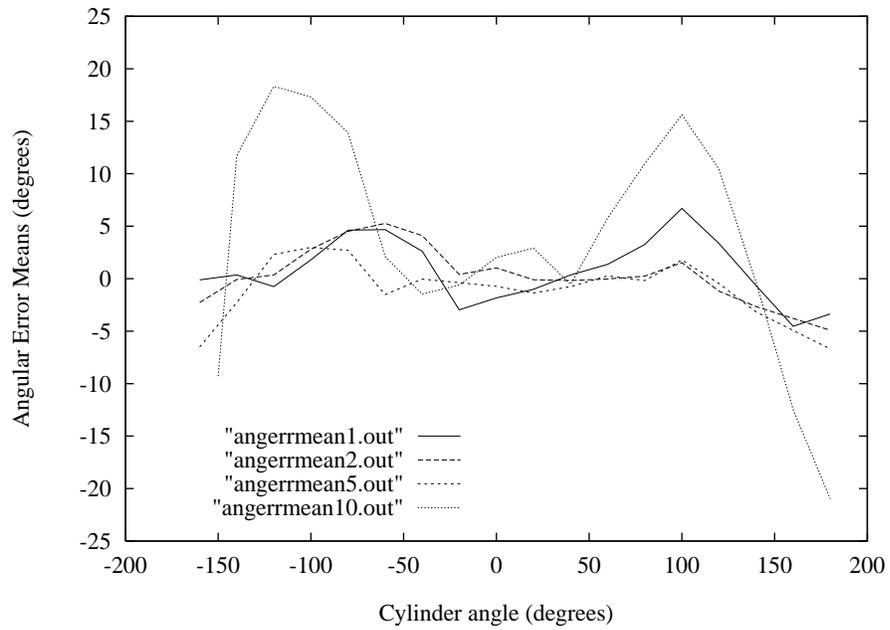
Figure 9: Rover localization with a colored cylinder: means of the angular error at the four tested distances.

Figure 10: Rover localization with a colored cylinder: means of the relative error at for stereo distance determination at the four tested distances.

14

Figure 11: Rover localization with a colored cylinder: standard deviation of the angular error at the four tested distances. The one data point off the graph is at (-140,16).
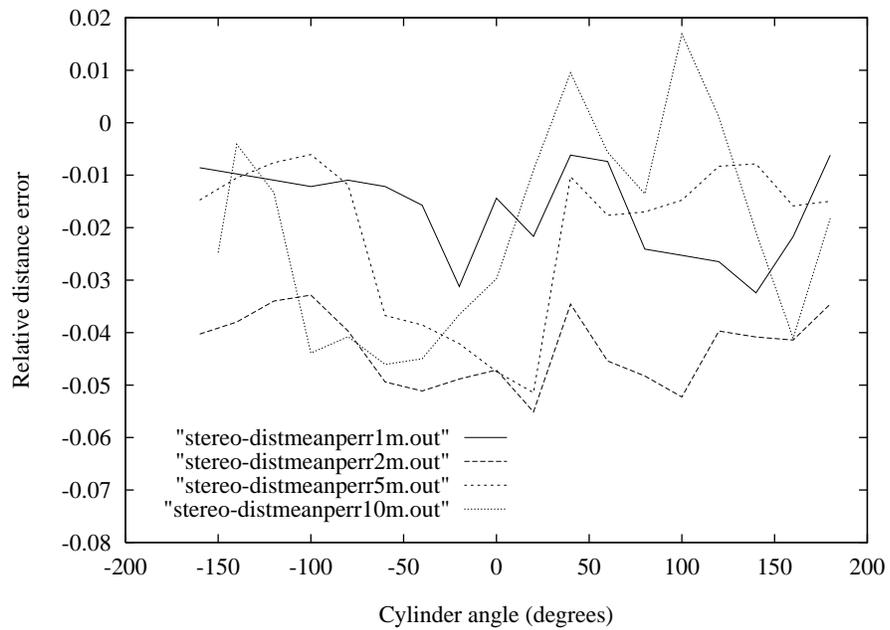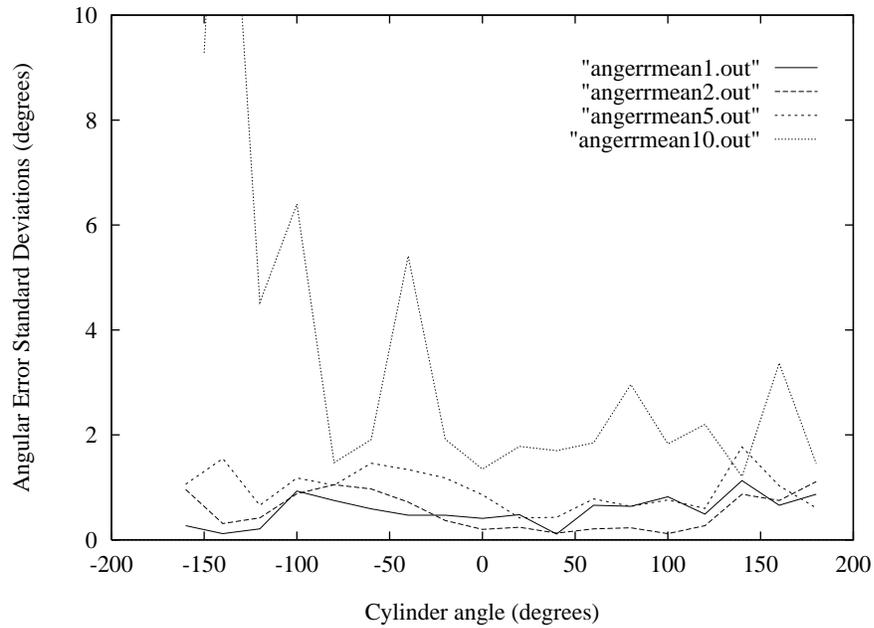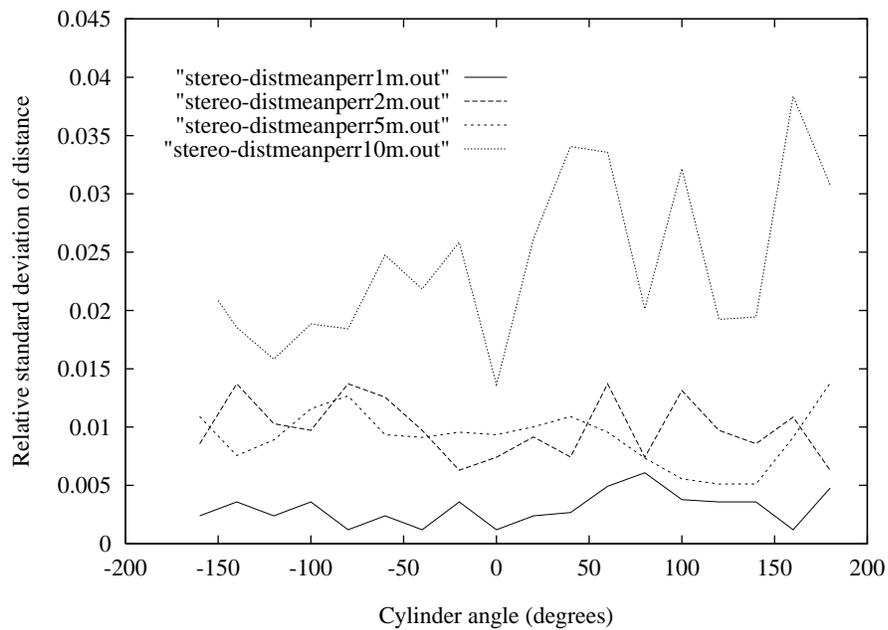

Figure 12: Rover localization with a colored cylinder: relative standard deviations for stereo distance determination at the four tested distances.

orientation determination. This can be seen in the 10 meter angular error data — the smallest errors were in the range of $-90°$ to $+90°$, when the blue quadrant plays a significant role. (Partial data taken at 15 meters confirmed this trend.)

The second improvement, more robust segmentation, would increase the accuracy of both the angle and distance measurements. First, by refining the segmentation of the individual colored regions on the cylinder, the angle measurements would improve. Second, by fine tuning the the accuracy of the cylinder centroid measurement, stereo disparity calculations and, therefore, distance determination would improve slightly. Also, episodes of erroneous segmentation of regions of the image that are not part of the cylinder could be reduced by an improved algorithm.

Finally, it important to point out that these results are pertinent to Pathfinder MFEX. Although human orientation determination of the rover will be used, the visual clues will be limited by similar image resolution problems. Therefore, the results shown here suggest that stereo pair viewing by humans for rover localization will not provide results better than $\pm 5°$ in heading and $\pm 5\%$ of distance. Future work on visual localization will tailor the approach described here to use targets that are compatible with the size and form constraints of flight rovers.

6 Hazard Detection

The Pathfinder approach to hazard detection is to detect rocks, pits, and excessive slopes with a light stripe ranging sensor, rocker bogey angle encoders, inclinometers, motor current sensors, and contact sensors on the edges of the solar panel and the leading edge of the undercarriage. The light stripe sensor has been implemented previously on Rocky 3.2, so hardware and software exist for this sensor. Most of the other sensors are not implemented on Rocky 3.2; therefore, with MFEX still under construction, software to use the other sensors is pending completion of MFEX prototypes. Since the Pathfinder philosophy is to count on lookahead sensors (ie. the light striper) as the primary means for hazard detection, we concentrate here on the design and performance of the light stripe sensor.

The light striper performs active triangulation using two cameras and five laser diode-based light stripe emitters (table 3). The rover's 8085 CPU performs all computing and control functions for the sensor, including clocking and readout of the CCD, exposure control for the image, and all image processing for hazard detection. The system detects images of four points along each stripe to produce an array of 20 measurements of terrain elevation (figure 13). On level ground, the farthest point on the central stripe projects about 30 cm in front of of the rover; the farthest points on the outermost stripes project about 13 cm beyond each side of the rover. A hazard is declared if any of the following criteria are met:

- either of the nearest two points for any laser are not detected in the images;

- the elevation difference between any two 8-connected neighbors in the 4×5 measurement array exceeds a threshold $t_{\Delta h}$;

- the difference between the highest and lowest elevation of the whole array exceeds another threshold.

Currently, Rocky 3.2 scans for hazards once every wheel radius of forward motion (6.5 cm). As discussed below, the rover has to stop to do each hazard scan.

16

Imagers	Kodak KAI-0370NI CCD's
Resolution	768 H $\times$ 484 V pixels
Active area	8.9 H $\times$ 6.6 V mm
Camera focal length	4 mm
Field of view	2.2 H $\times$ 1.6 V radians
Angular resolution	2.9 H $\times$ 3.4 V mr/pixel
Laser diode input power	0.66 W each
Wavelength	850 nm
Beam spread	0.003 H $\times$ 1.0 V radians
Camera separation	12 cm
Camera-to-emitter baseline	5 to 10 cm

Table 3: Basic specifications of the MFEX light stripe sensor. Rocky 3.2 uses a 4.8 mm lens, giving it a slightly narrower field of view.
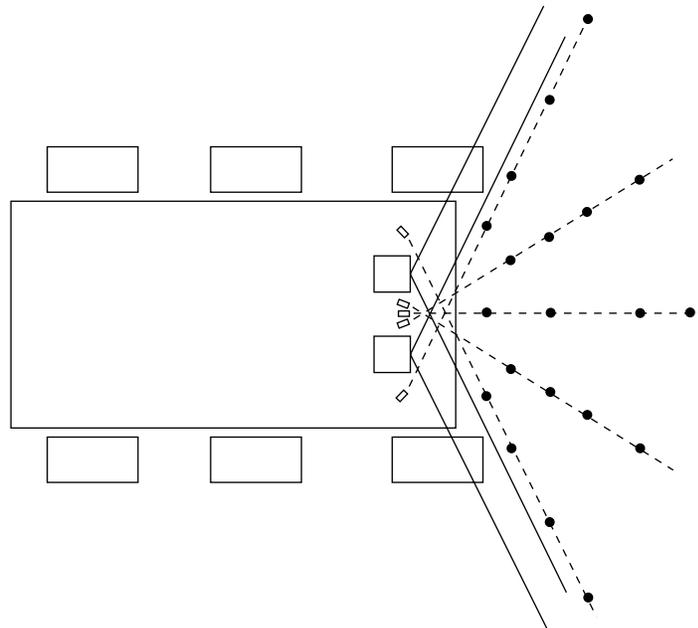

Figure 13: Geometry of the MFEX light stripe-based lookahead sensor, approximately to scale, showing camera fields of view (solid lines), light stripes (dotted lines), and where the array of 4×5 measured elevation points land on flat ground (filled circles).

Two groups of performance issues are important for hazard sensors: (1) the quality of hazard detection and (2) the mass, power, volume, and runtime requirements of the sensor. Since an outdoor version of the light stripe sensor has only recently become fully operational, quantitative evaluation of its detection performance has just begun; therefore, we will discuss this issue based on qualitative experience to date. Three factors affecting detection performance are:

- reliability of detecting the stripe itself in the images, in the presence of full solar background illumination and within acceptable power levels for the laser diodes;

- precision of the elevation estimates, given that the stripes are detected correctly in the images;

- potential for aliasing of obstacles, by failing to see small obstacles in the spaces between the measured elevation points.

From table 3 and equation (2), range resolution is approximately 2 mm at 30 cm, given that the stripes are detected to pixel resolution in the imagery; therefore, precision of the elevation measurements should not be a limiting factor. Potential for aliasing of obstacles is a moderate concern for the expected terrain profiles (we will return to this issue later). For outdoor operation, the limiting factor on detection performance has been reliability of stripe detection in the images. Several stages of filtering are used to enhance this reliability, including:

- narrow optical bandpass filtering around the wavelength of the lasers, to reduce the sunlight content in the images;

- 5×5 averaging of the acquired images to reduce noise;

- differencing of two consecutive images, one with the lasers on and one with the lasers off, to subtract out remaining image power due to sunlight.

Stripes are detected by thresholding the difference images. Two lasers are on for each image, to balance runtime against power consumption and the difficulty of associating detected stripes with the correct laser. Approximately one of every hundred points is detected in the wrong place in the image, which gives a false alarm rate of one in every five hazard scans. This is unacceptably high, so to reduce the false positive rate the system requires that detected hazards be seen consistently in two consecutive scans before they are believed. This is implemented by requiring all of the elevation measurements in two consecutive scans to agree to within a threshold, which makes the false alarm probability vanishingly small. The rate of false negatives (missed detections) is estimated to be about one in every thousand hazards. Since hazards are expected to occur on the order of once per meter (section 7), this rate is acceptable for Pathfinder, but would need to be improved for exploration over multiple kilometers.

Returning to the issue of obstacle aliasing, the relatively sparse pattern of elevation measurements can miss the top of a rock. Experience indicates that on average rocks must be about 1.5 times higher than $t_{\Delta h}$ to be seen as hazards. On Rocky 3.2, this threshold is currently set at 5 cm to provide a comfortable clearance margin for the undercarriage, which is 13 cm off the ground; therefore, rocks typically seen as hazards are 7 to 8 cm or more high. In some quarters this threshold is considered unduly conservative, given that the step-climbing capability exceeds 13 cm. This issue is subject to ongoing debate and will have to be resolved experimentally.

Turning to the resources required by the light stripe sensor, the mass and volume of the sensor components are relatively insignificant. The MFEX power budget for hazard detection mode is 7.33 W, which consists of 2.01 W for two lasers at a time (including voltage conversion losses), 3.77 W for the CPU and I/O functions, 0.75 W for the CCD's, and 0.8 W to run the attitude sensors. Thus, the lasers are a relatively significant part of the power budget. The time required to perform hazard detection is 20 seconds per scan. Since MFEX will drive about 6.5 cm between scans at a speed of 0.67 cm/sec, the net traversal rate will be around 0.22 cm/sec; in a word, painstaking. Still, this should be enough to reach rocks a few meters from the lander within an hour.

As with all hazard detection sensors, the goals for improving the performance of this system are to increase its speed and detection reliability while reducing its power consumption, in order to enable safe exploration of much more territory per day. Approaches to doing so include:

- Reducing the power consumption by changing the beam fanout from about 300 scanlines to 3 scanlines. The saving of a factor 100 in power requirement per emitter could be used to increase the detection reliability, to increase the maximum detection range, or to increase the number of emitters turned on at once, which could increase the speed of the system.

- Replacing the CCD imagers with CMOS imagers [12]. This offers the potential of a factor of 10 power reduction for imaging alone, plus the potential for much faster processing by incorporating image processing electronics onto the imager chip itself.

- Switching from active triangulation to passive triangulation with stereo vision. In previous work [13, 7], we have developed and demonstrated reliable stereo vision-based hazard detection systems for automobile-sized vehicles. In one instantiation, these algorithms acquire images at 512×480 resolution, reduce the resolution to 64×60 by image pyramid transformation, and produce roughly 45×45 range pixels by stereo matching at the 64×60 level of the image pyramid. This version has recently been tuned and benchmarked on a Sparc 10 to run in 100 ms for the image pyramid and 3 ms/scanline for the rest of the processing. Our benchmarks also suggest that a 68040 CPU could do all of the processing in one second. Morever, if the image pyramid portion were eliminated, by using low resolution imagers or special purpose convolution hardware, even the 8085 could do the remaining computations at rates very competitive with the current light striper. Stereo has the potential of providing higher spatial resolution and being much more flexible than the light striper. Therefore, it appears that stereo vision is now both practical and attractive for microrovers.

7   Path Selection

The Pathfinder approach to path selection is a behavior control algorithm [14, 8] that uses very simple steering logic based on the instantaneous state of the hazard detection sensors; that is, it uses no internal map or memory of previously encountered hazards. Basically, the logic is as follows:

IF there is no hazard,
    move forward and turn toward the goal,
ELSE IF there is a hazard on the left,
    turn in place to the right until no hazard is detected;

ELSE IF there is a hazard on the right,
     turn in place to the left until no hazard is detected.

Currently, hazards in the center are resolved by turning right if that is clear, otherwise turning left. In "move forward" mode, a small amount of additional logic chooses a turning radius based on the bearing to the goal. The simplicity of the algorithm makes it practical to implement on the 8085 flight computer. Experience to date shows that the algorithm is quite effective as long as obstacle frequencies are not too high; as discussed below, ongoing performance evaluation is quantifying what is meant by "too high".

To evaluate the performance of a path selection algorithm, we must define the terrain type(s) on which the rover will be tested and the metrics against which performance will be measured. We would like to use terrain types that closely match what is known about Martian terrain. Knowledge of Mars terrain on the scale of a rover is available from images and other measurements made by the two Viking landers. Lower resolution knowledge, on scales from a few meters to many kilometers per pixel, is available from Viking orbiter imagery and Earth-based radar observations [10]. The lander imagery has been used to derive a model (known as Moore's model) of the rock size frequency distribution for the Viking Lander 2 (VL-2) site; together with rock abundances estimated from orbiter thermal imagery, this model can be used to predict rock size distributions for other areas of the planet [15, 16]. Moore's model gives the cumulative number N of rocks/m^2 with diameters of D and larger as

$$N = kD^{-2.66} \tag{4}$$

where $k = 0.013$ for the VL-2 site. This model predicts that 19% of the surface around the lander is covered by rocks 0.1 m in diameter and larger, which is in agreement with rock abundance estimates made from global thermal inertia data. These rock abundance estimates give the modal value of surface rock cover over the whole planet as 6%. From this value, one can obtain $k = 0.00415$ for the modal rock abundance; in what follows, we refer to this case as "nominal" terrain. Thermal inertia and other data indicate that the Pathfinder Ares Tiu landing site has a rock abundance approaching 20%, which suggests a rock size frequency distribution comparable to VL-2[6]. Figure 14 shows an image from the VL-2 site for reference.

One approach to creating Mars analog terrains for rover testing is to generate randomly placed rock fields with sizes randomly chosen from Moore's model. A simpler approach is to define obstacles as all rocks with diameters greater than or equal to some fixed value D_0, to ignore all rocks with $D < D_0$, to use Moore's model to predict the number of obstacles/m^2, and to generate randomly placed obstacles with that frequency. We have used both approaches in experiments conducted to date.

One class of performance metric is to estimate the probability that the rover will reach its goal, for an ensemble of test runs over a given terrain type, and to catalog the nature of the failures that occur when the rover does not reach the goal. A second class of metric is to estimate the distribution of some parameter of the runs as a function of terrain type, such as the total distance traveled to reach goals a given radial distance from the lander. We have conducted tests with each metric; the first has been used for terrains generated in terms of obstacles/m^2, the second for terrains with rocks distributed according to Moore's model with k defined for nominal terrain. We plan to

---

[6]Since different geologic processes may have been in effect at the two sites, this interpretation may be wrong. Nevertheless, it provides a reasonable starting point for performance evaluation. Equation (4) overestimates the frequency of rocks with diameter less than 14 cm; see [15, 16] for details.

Figure 14: View of Mars looking northeast from Viking Lander 2. The rock in the lower right corner is about 0.25 m wide and 2.5 m away; the largest rock in the center of the image is about 1.0 m wide and 6.5 m away. The horizon is about 3 km away; its slope is due to the 8-degree tilt of the lander.

extend the latter set of tests to terrains with k ranging up to the value for the VL-2 site. Since these tests are very laborious to conduct, we are also developing a computer simulation of rover runs and attempting to validate the simulation by comparing its predictions to the results of real runs. Ultimately, we also intend to evaluate path selection performance as function of the onboard sensor suite, for example by varying the set of hazard detection sensors, to make trade-offs of capability versus mass and power budgets.

For terrains defined in terms of obstacles/m^2 (obs/m^2), one set of twelve runs was performed with 5 obstacles in a 17.6 m^2 area and two sets of six runs each were performed with 10 and 15 obstacles in the same area (0.28, 0.57, and 0.85 obs/m^2, respectively). All obstacles were rocks of between 25 and 35 cm diameter, which was large enough to be detected unambiguously by the light stripe sensor. In each case, the rover was given a goal 7.6 m away. A run was declared successfully completed when the rover's deadreckoned position estimate implied that it had reached the goal. At 0.28 and 0.57 obs/m^2, all runs were successful. At 0.85 obs/m^2, four runs were successful, one run failed when the rover got jammed between two rocks and depleted its batteries, and one run was terminated when it became clear that the rover couldn't find a way to the goal.

To interpret these results in terms of obstacle densities likely to be encountered on Mars, we need to estimate what rock diameter D_0 corresponds to an untraversable obstacle. Taking the ground clearance of the MFEX undercarriage ($\approx$ 15 cm) as the minimum obstacle height, and
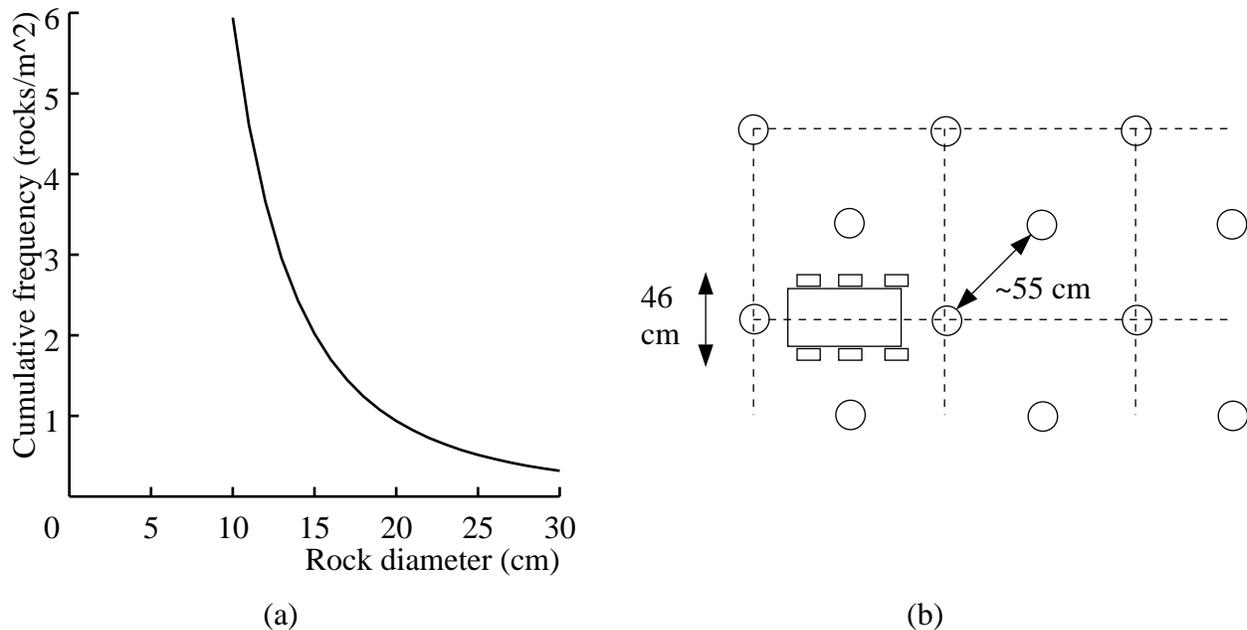
21

Figure 15: (a) Moore's model for VL-2 site. (b) An illustration of two 15 cm obstacles per square meter.

using the heuristic that on average rocks are 1/3 buried, we obtain $D_0 \approx 23$ cm. Moore's model for the VL-2 site then implies an obstacle density of 0.65 obs/m^2. These results suggest that the rover may be able to navigate fairly well through VL-2 class terrain. However, the model is quite sensitive to the value of D, and $D_0 = 23$ cm may be optimistic; for example, for $D_0 = 15$ cm the model predicts 2.0 obs/m^2 (figure 15a), which would be quite difficult to traverse (figure 15b). Since there is controversy over the size of rock that can be climbed safely, this underscores the importance of experimentally evaluating this aspect of performance.

Although the above results provide insight into rover performance, there are limitations to characterizing terrain in terms of obstacles/m^2. In particular, this effectively models navigation as taking place in a 2-D world and largely ignores the 3-D geometry of the terrain. To address this limitation, we are conducting a larger set of trials with rock frequencies generated from Moore's model. To date, we have completed a set of 40 trials for "nominal" terrain ($k = 0.00415$). For these trials, we have tabulated whether or not the run was successful, the reason for failure, if any, and the total distance traveled. A complication with these runs, as well as those described above, is that the 4 m width of the test arena was too narrow to allow much maneuvering room for the rover. In order to simulate a wider area, a "virtual" test arena procedure was devised in which the layout of the rock field was made symmetric along lines 0.9 m in from the edge of the arena. Whenever the rover's position reached this line, it was turned in place to the mirror-image heading back into the arena; likewise, its internal state estimate was reflected about this line. This procedure effectively simulated a rock layout that was random across the central 2.2 m of the course width, and mirror-image periodic beyond that.

Figure 16 shows a schematic map of the rock layout and the true and deadreckoned rover positions for one of the runs. The rover avoided large rocks several times during the run and drove over several small rocks. The run was declared successful, but owing to deadreckoning error the true final position of the rover was 1.4 m from the goal. Out of the 40 runs, only one was unsuccessful; in this case, the run was terminated because the rover could not find a path through

22

	Distance error (%)	Heading error (deg)	Position error (m)
Min	-3.0	0.2	0.22
Mode	4.8	8.0	0.69
Max	9.0	37.2	2.45

Table 4: Deadreckoning error statistics for 40 runs in Mars nominal terrain. The mode of the true distance traveled was 8.39 m. Two runs showed a negative distance error (true distance $>$ deadreckoned distance); this happens because deadreckoned distance estimates include a scale factor used to calibrate for wheel slip on level, rock-free sand.

the rock field to the goal. Higher rock frequencies will be tested in future work, including the VL-2 distribution being used as a model for the Pathfinder landing site.

Figure 17 shows the cumulative distribution function of the total distance covered in the 40 runs. That is, each diamond plots the distance covered for one run, and the vertical axis shows the probability that a run covered less than or equal to the associated distance. The shape of the curve is generally consistent with an exponential distribution for distance traveled; this in turn is consistent with the random rock layout being a Poisson process. The dotted curve shows the distribution of distances produced by a computer simulation that is under development. The shape of the simulation curve is in good agreement with the experimental data; the offset between the two curves occurs because the simulator does not yet include deadreckoning error. For the real runs, the shortest distance traveled was 7.0 m, whereas the true shortest distance from start to goal was 7.6 m; the difference reflects the impact of wheel slip on deadreckoning error. Table 4 summarizes deadreckoning error statistics for these runs. These errors are consistent with the errors seen in section 5 for a run involving rocks and turns, but significantly larger than seen for straight runs on level sand.

Approaches to improve the path selection algorithm will include using maps generated from lander imagery, rover sensors, or both. Their value will be evaluated by measuring their impact on the variables used above, including failure rates and mean distance traveled over large numbers of trials.

8   Summary and Conclusions

Autonomous navigation systems for robotic ground vehicles have not been subjected to systematic, quantitative performance evaluation as a function of terrain difficulty. Such work is sorely needed, for both planetary exploration and for terrestrial applications. We have begun to fill this gap by developing facilities and methodologies for evaluating the performance of microrovers for Mars exploration.

In this paper, we outlined current plans for Mars exploration, summarized the design of the lander and rover for the 1996 Mars Pathfinder mission, and described our rover evaluation testbed. Nominal mission scenarios for the Pathfinder rover ("MFEX") call for traversing between 3 and 40 m per day at distances out to 40 m or more from the lander. The evaluation testbed includes the Rocky 3.2 vehicle, which in most respects is functionally equivalent to MFEX, and a 4×12 m indoor test arena with overhead cameras for tracking the true position and heading of the rover. Mars analog terrains are created by randomly distributing rocks according to the exponential model
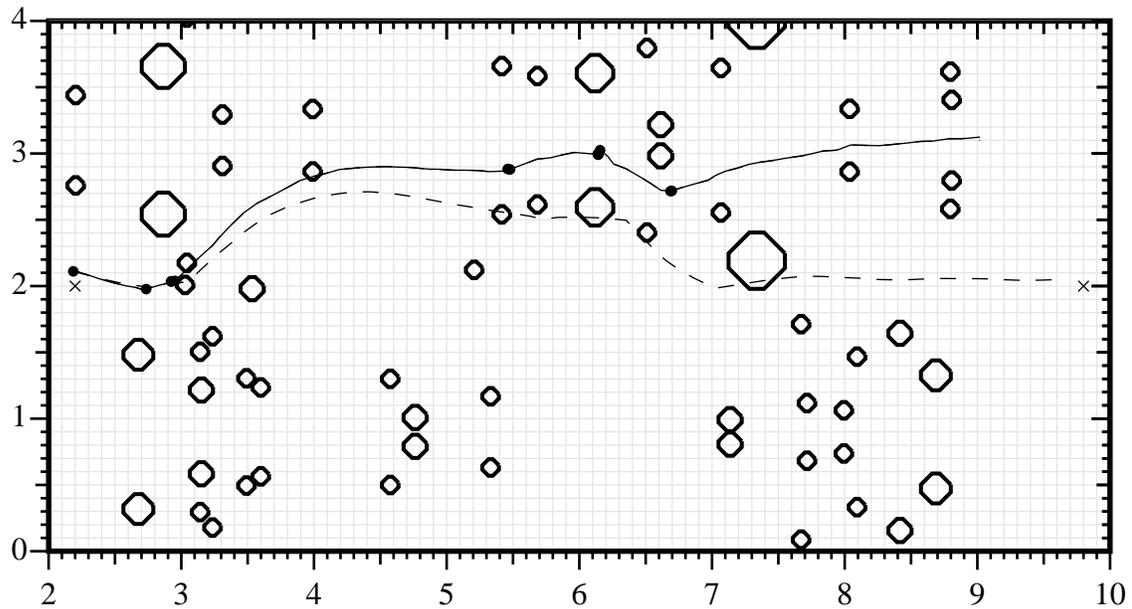
Figure 16: Schematic map of one run in nominal terrain. Octagons show positions and sizes of rocks placed in the test arena. The solid line traces the rover's actual position; the dotted line traces its deadreckoned position. Filled circles on the traces mark rover positions at which obstacles were detected. The starting point was the $\times$ on the left, the goal point was the $\times$ on the right.
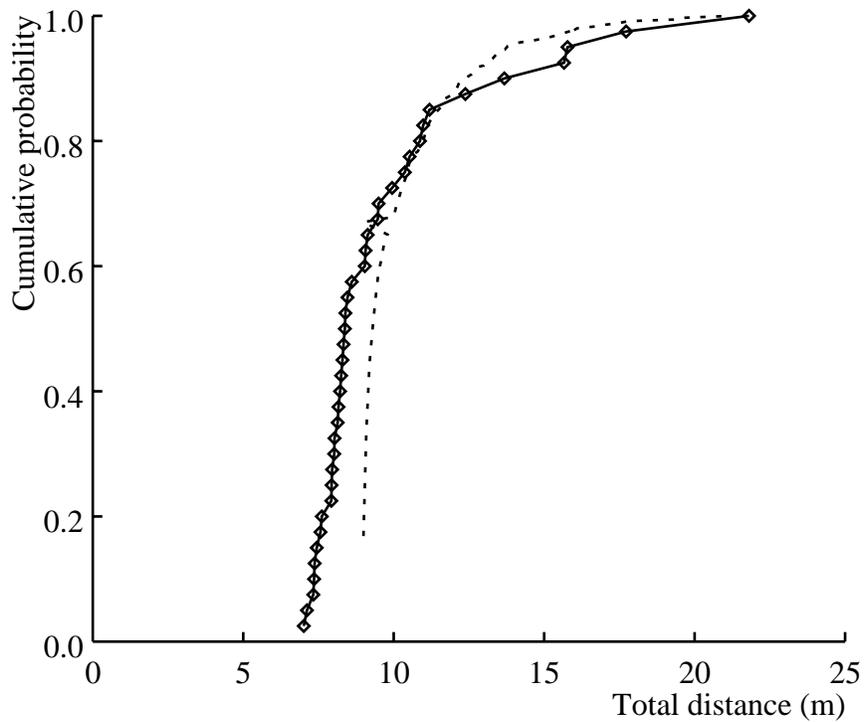

Figure 17: Cumulative distribution function for total distance traveled in nominal terrain. Solid curve: 40 real runs; dotted curve: computer simulation of 300 trials.

24

of rock size frequency derived for the Viking Lander 2 site by Moore [15, 16]. Initial testing has been done with a "nominal" distribution that covers 6% of the terrain with rocks, which corresponds to the modal rock abundance estimated from global thermal inertia data.

We decomposed the navigation task into four functions: goal designation, rover localization, hazard detection, and path selection. For each function, we outlined the approach being taken for the Pathfinder mission, the results of performance evaluation to date, and the technologies we are developing to improve performance for future missions.

Goals for the rover will be defined as 3-D waypoints designated with lander stereo imagery. Simple sensitivity analyses show that designation errors reach one vehicle length (60 cm) at about 20 m from the lander and grow quadratically with distance beyond that. Since the rover uses onboard sensors to recognize prominent rocks from about a vehicle length away, this approach to goal designation should be adequate within 20 m; new methods will be necessary at greater distances.

Rover localization will be done by deadreckoning with wheel encoders and a solid state turn rate sensor. Rover position and heading will be updated once a day by interactively locating the rover in lander imagery that is sent to Earth. To date, deadreckoning performance has been evaluated with Rocky 3.2 driving at 15 cm/sec. On level, rock-free sand, the variance of heading error grows linearly at 0.70 deg^2/m. Since heading error is a function time, naively scaling this rate to the MFEX velocity of 0.67 cm/sec predicts a heading variance of 16 deg^2/sec. This should be viewed as an upper bound, since the relative importance of various contributing noise terms has not been determined. On top of this, Rocky 3.2 underestimates turn angles by about 6% and incurs large heading errors (≈ 5 deg) when it scrapes against rocks. Likely causes are a calibration error and vibrations perturbing the turn rate sensor, respectively. The variance of error in deadreckoned distance traveled is approximately 0.0039 m^2/m. In 40 trials on the nominal rock distribution, with a true, straight line distance from start to goal of 7.6 m, typical wheel slip estimates were 5% worse than on rock-free terrain. Typical errors in position and heading at the end of a trial were 0.69 m and 8.0 deg, respectively; worst case errors were 2.45 m and 37.2 deg. This performance is somewhat worse than suggested by the results on rock-free terrain; in general, we expect that performance will degrade with increasing rock frequency. Overall, the deadreckoning results suggest that localization updates will be needed at least every 10 m if the rover is to successfully reach specific rocks.

Hazard detection will be done primarily with a light stripe sensor that measures an array of 4×5 elevation points spread over 30 cm in front of the rover and 13 cm to each side of the rover. On Rocky 3.2, this system is currently tuned to detect rocks ≈ 7 cm high. The algorithms are designed to make the probability of a false alarm vanishingly small; the probability of missed detection is estimated to be about 1 in every 1000 hazards. Speed is a key limitation of the sensor, which takes 20 sec per hazard scan. This stems from a combination of factors, including the slowness of the 8085 processor, the wide fan-out of the light stripes, and the constrained power budget on the rover. At this rate, the current plan of doing hazard scans once per wheel radius of forward motion leads to a maximum traverse of 30 m in the four hours available per day.

The current approach to path selection is a very simple behavior control algorithm that chooses steering directions based on the instantaneous readings of hazard sensors. Maps are not employed. In terrains characterized by the number of obstacles/m^2, Rocky 3.2 succeeded in reaching goals 7.6 m away in 18 out of 18 trials for obstacle frequencies up to 0.57 per square meter; at 0.85 obstacles/m^2, only 4 of 6 trials were successful. Obstacle frequencies for the Pathfinder landing

site are hard to predict, but plausible estimates range from 0.65 to 2.0 obstacles/m^2. In terrains characterized by the nominal 6% rock abundance, 39 out of 40 trials were successful. The rock abundance predicted for the Pathfinder landing site is close to 20%; it remains to be seen how well the rover will perform in such terrain.

In summary, the most significant limitations revealed to date are poor heading estimation and slowness of the hazard detection system. In addition, there is uncertainty about how well behavior control algorithms for path selection will work at the rock frequencies likely to be encountered on Mars. Objectives for future missions include alleviating these bottlenecks and enabling exploration over much greater distances with lighter, lower-power vehicles. To achieve these objectives, we are:

- building a new testbed vehicle with a lighter, lower-power mobility subsystem;

- developing a sun sensor with a field of view of greater than 150 degrees to provide an absolute heading reference;

- developing RF navigation aids for position estimation when the rover is beyond visual range from the lander;

- developing a faster, lower power, higher resolution hazard detection sensor based on CMOS cameras and stereo vision;

- evaluating both behavior-based and map-based path selection algorithms at higher rock frequencies;

- incorporating a 68040-based processor on the new testbed vehicle to enable far more rapid development and testing.

In related work, we are also developing dexterous mechanisms for sample acquisition.

In conclusion, this work represents one of the few efforts to systematically measure the performance of autonomous navigation systems as a function of the difficulty of the terrain. Such work is essential for identifying the limitations of current rovers and focusing future work on issues that matter. The new testbed facilities have been invaluable in this process, in particular the automatic rover tracking system and the parameterized family of Mars analog terrains. We believe that the facilities, methodologies, and to some extent the specific performance results presented here will provide valuable examples for efforts to evaluate robotic vehicle performance in other applications.

Acknowledgements

References

[1] B. H. Wilcox, D. B. Gennery, and A. Mishkin. Mars rover local navigation and hazard avoidance. In *Proc. SPIE Conf. 1007, Mobile Robots III*, pages 72–76. SPIE, November 1988.

[2] B. Wilcox et al. Robotic vehicles for planetary exploration. In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pages 175–180, May 1992.

[3] E. G. Mettala. The OSD Tactical Unmanned Ground Vehicle Program. In *Proc. DARPA Image Understanding Workshop*, pages 159–172. Morgan Kaufmanm Publishers, January 1992.

[4] D. Wettergreen, C. E. Thorpe, and W. Whittaker. Exploring Mount Erebus by walking robot. In *Proc. 3rd Int'l Conf. on Intelligent Autonomous Systems (IAS-3)*, pages 72–81. IOS Press, February 1993.

[5] M. F. Reiley, D. C. Carmer, and W. F. Pont. 3-D laser radar simulation for autonomous spacecraft landing. In *Proc. SPIE Int'l Symposium on High Power Lasers*. SPIE, January 1991.

[6] M. Hebert and E. Krotkov. 3D measurements from imaging laser radars: how good are they? *Image and Vision Computing*, 10(3):170–178, April 1992.

[7] L. H. Matthies and P. Grandjean. Stochastic performance modeling and evaluation of obstacle detectability with imaging range sensors. *IEEE Transactions on Robotics and Automation*, 10(6):783–791, December 1994.

[8] E. Gat, R. Desai, R. Ivlev, J. Loch, and D. P. Miller. Behavior control for robotic exploration of planetary surfaces. *IEEE Journal of Robotics and Automation*, 10(4):490–503, August 1994.

[9] W. N. Kaliardos. Sensors for autonomous navigation and hazard avoidance on a planetary micro-rover. Master's thesis, Massachusetts Institute of Technology, May 1993. CSDL-T-1186.

[10] H. H. Kieffer (ed.). *Mars*. University of Arizona Press, 1992.

[11] R. Volpe, T. Litwin, and L. Matthies. Mobile robot localization by remote viewing of a colored cylinder. In *Proc. 1995 Int'l Conf. on Intelligent Robots and Systems (IROS'95)*, page (to appear), August 1995.

[12] E. R. Fossum. Active pixel sensors: are CCD's dinosaurs? In *Proceedings of SPIE Conf. 1900: Charge-coupled devices and solid-state optical sensors III*, pages 1–13, 1993.

[13] L. H. Matthies. Stereo vision for planetary rovers: stochastic modeling to near real-time implementation. *International Journal of Computer Vision*, 8(1):71–91, July 1992.

[14] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation*, RA-2(1), March 1986.

[15] H. J. Moore and B. M. Jakosky. Viking landing sites, remote-sensing observations, and physical properties of Martian surface materials. *Icarus*, 81:164–184, 1989.

[16] P. R. Christensen and H. J. Moore. The Martian surface layer. In H. H. Kieffer, editor, *Mars*, chapter 21, pages 686–729. University of Arizona Press, 1992.

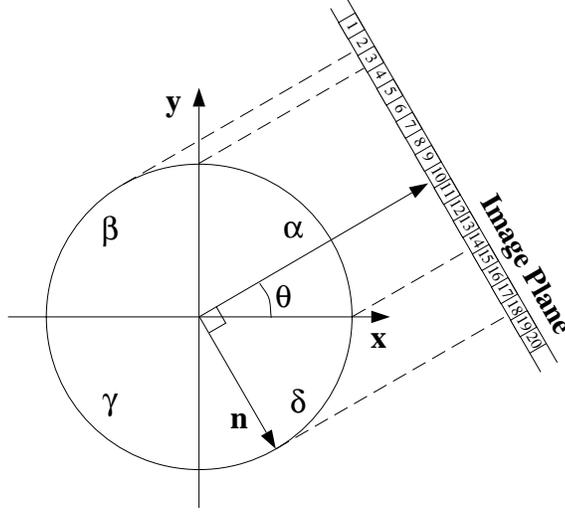[17] D. H. Ballard and C. M. Brown. *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, 1982.

Figure 18: Geometry of the viewing of the colored cylinder. See text for details.

A   Rover Localization with a Colored Cylinder

This appendix gives the detailed derivation of rover position and orientation estimates from images of a colored cylinder mounted on the rover, as well an analysis of the sensitivity of the estimates to image segmentation errors and a description of the image segmentation algorithm that has been used to date.

A.1   Position and Orientation Determination

Figure 18 shows a top view of a short cylinder of radius $\rho = |\mathbf{n}|$. The external wall of each of the four quadrants is painted in a different color $\kappa \in \{\alpha, \beta, \gamma, \delta\}$. When viewed from a distance, each color will be orthographically projected onto the image plane in quantity $k \in \{A, B, C, D\}$. The image width of the cylinder is m pixels, and the total image width from the camera is M.

A.1.1   Orientation

Initially, it is assumed that the viewing direction is in the first quadrant at an angle θ, and the vector $\mathbf{n}$ is normal to this direction as shown in the figure. Therefore, the amount of each color seen is:

$$
\begin{align}
A &= (\mathbf{n} \cdot \hat{\mathbf{x}}) - (\mathbf{n} \cdot \hat{\mathbf{y}}) \tag{5} \\
B &= \rho + (\mathbf{n} \cdot \hat{\mathbf{y}}) \tag{6} \\
C &= 0 \tag{7} \\
D &= \rho - (\mathbf{n} \cdot \hat{\mathbf{x}}) \tag{8}
\end{align}
$$

where the circumflex accent indicates unit vectors. The quantity C is zero since it can not be seen. To remove the dependence on the cylinder radius, it is useful to take the ratio of the quantities: A/B and A/D. Recognizing that

$$
\begin{align}
\hat{\mathbf{n}} \cdot \hat{\mathbf{x}} &= \sin\theta \tag{9} \\
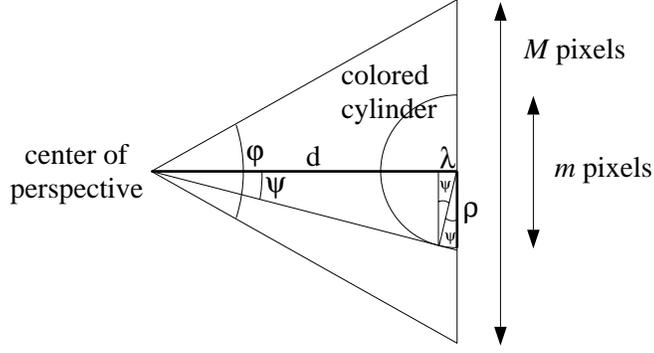\hat{\mathbf{n}} \cdot \hat{\mathbf{y}} &= -\cos\theta \tag{10}
\end{align}
$$

28

Figure 19: Geometry for determining the distance to the cylinder. See the text for details.

the ratios may be reformulated as a pair of equations:

$$
\begin{bmatrix} B & A+B \\ A+D & D \end{bmatrix} \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} A \\ A \end{bmatrix}
\tag{11}
$$

Solving this set of equations yields:

$$
\sin\theta = \frac{A+B-D}{A+B+D}
\tag{12}
$$

$$
\cos\theta = \frac{A-B+D}{A+B+D}
\tag{13}
$$

For viewing directions in other quadrants, there will be different colors involve, and an angular offset (θ_0 equal to a multiple of $\frac{\pi}{2}$) from the geometry described above. A simple test to determine the quadrant of interest is to measure which color is not present. A zero value for this color quantity specifies the row in the following table:

A	B	C	D	θ_0
k_1	k_2	0	k_3	0
k_3	k_1	k_2	0	$\frac{\pi}{2}$
0	k_3	k_1	k_2	π
k_2	0	k_3	k_1	$\frac{3\pi}{2}$

Using the proper row for the zero color, the angular offset is determined and the measured color quantities are mapped to intermediate variables, k_i, for use in the following orientation solution:

$$
\theta = \theta_0 + \text{atan2}\left(\frac{k_1+k_2-k_3}{k_1+k_2+k_3}, \frac{k_1-k_2+k_3}{k_1+k_2+k_3}\right)
\tag{14}
$$

where $\text{atan2}(x,y)$ is the unambiguous form of $\tan^{-1}(x/y)$.

A.1.2 Position by Size

Figure 19 shows the geometry of viewing at a distance d, with a camera field of view φ. Again assuming

29

an orthographic projection of the scene on to the image plane, the image will consume m/M of the width. Therefore,

$$d = \frac{M\rho}{m \tan \frac{\phi}{2}} \tag{15}$$

When the camera is close to the cylinder, the perspective can introduce a foreshortening error. Figure 19 shows that $sin\psi = \frac{\rho}{d} = \frac{\lambda}{\rho}$. Therefore, the correction quantity, λ, must be added to the previous expression for the cylinder distance:

$$d \approx \frac{M\rho}{m \tan \frac{\phi}{2}} + \frac{\rho^2}{d} \tag{16}$$

In this approximation, Equation (15) can be used to iteratively provide a solution, or Equation (16) can be solved as a quadratic equation.

For example, the 17 pixel image from the previous section corresponds to a distance of 11.24 m, assuming a 0.1 m radius cylinder with a nominal 512 x 480 image and a 30° field of view.

A.1.3   Position by Stereo Image Triangulation

The geometry for stereo triangulation is very similar to that described in the previous section. Instead of relying on the edges of the imaged cylinder to form a triangle with the camera location, two images are used to triangulate on the centroid of the cylinder. In this case, the distance is dependent on the *disparity* of the centroid in the images, measured in pixels μ:

$$d = \frac{Mb}{2\mu \tan \frac{\phi}{2}} \tag{17}$$

where b is the baseline distance between the stereo cameras. Since the cylinder is radially symmetric, no view angle effects are encountered to effect the centroid location.

A.2   Pixel Resolution Effects

A.2.1   Orientation Errors

The precision of this orientation measuring scheme is dependent on the spatial resolution of the pixels in the image of the cylinder. There are two extremes to the resolution of orientation determination by this scheme, depending on the amount of rotation needed to move the color boundaries from one pixel to the next.

The *best case* of orientation resolution occurs when one of the color boundaries is in the center of the image. In this orientation, a minimal amount of rotation is needed, since the motion of the color boundary is approximately parallel to the image plane. In the general case of m total pixels, the angular resolution is:

$$\Delta\theta_{min} = \frac{max \ image \ plane \ motion}{radius \ of \ cylinder} = \left(\frac{2\rho}{m}\right)\left(\frac{1}{\rho}\right) = \frac{2}{m} \tag{18}$$

The *worst case* of orientation resolution occurs when the color boundary motion is most out of the image plane. Given the four quadrant design of the cylinder, this occurs for $\theta \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$. Therefore,

$$\Delta\theta_{max} = \frac{min \ image \ plane \ motion}{radius \ of \ cylinder} = \left(\frac{2\rho}{m \cos \frac{\pi}{4}}\right)\left(\frac{1}{\rho}\right) = \frac{2\sqrt{2}}{m} \tag{19}$$

For the example of $m = 17$, the resultant resolution bounds are: $\Delta\theta_{min} = 6.7°$ and $\Delta\theta_{max} = 9.5°$.

It is useful to eliminate m from the equations above, by employing Equation (16):

$$\Delta\theta_{min} \;\;=\;\; \frac{2}{M}\left(\frac{d}{\rho}-\frac{\rho}{d}\right)\tan\frac{\phi}{2} \tag{20}$$

$$\Delta\theta_{max} \;\;=\;\; \sqrt{2}\Delta\theta_{min} \tag{21}$$

A.2.2 Relative Position Error by Size

For distance measurements, *relative* error is used to normalize the results and allow more direct comparison of them. The relative error in the distance calculation of Equation (15) caused by a one pixel change in the size is:

$$\left.\frac{\Delta d}{d}\right|_{size} = -\frac{M\rho}{dm^2\tan\frac{\phi}{2}} = -\frac{d\tan\frac{\phi}{2}}{M\rho} \tag{22}$$

For the example parameters used above, a one pixel change causes an error is 0.66 m, or about 5.8% of the total distance.

A.2.3 Relative Position Error by Stereo

Similarly, using Equation (17) the relative error for distance by stereo caused by one pixel change in the centroid is:

$$\left.\frac{\Delta d}{d}\right|_{stereo} = -\frac{Mb}{2d\mu^2\tan\frac{\phi}{2}} = -\frac{2d\tan\frac{\phi}{2}}{Mb}. \tag{23}$$

If the stereo camera baseline is equal to the diameter of the cylinder, the resolution for our example configuration will be same as for estimating position by size: 5.8%.

A.3 Segmentation Algorithm

The segmentation of the image to locate the quadrants of the cylinder is performed in four steps: (1) the image pixels are classified by color, (2) nearly adjacent regions are connected, (3) the cylinder is located, and (4) the visible colored quadrants of the cylinder are located.

Three color classes are defined, one for each of the three distinct colors of the cylinder's quadrants: blue, magenta, and red. Determining to which class each pixel belongs is done by taking the ratio of the primary colors of the image (for the red and magenta classes: green/red and blue/red; for the blue class: red/blue and green/blue). For each ratio an experimentally determined range of values is used to determine class membership. Intensity ratios are used because of their relative insensitivity to ambient lighting.

As mentioned previously the choice of colors is made to minimize rejection of pixels on the transition between the colors. Even so, many pixels on such boundaries are rejected as belonging to neither color class. As will become clear in the next step, it is important that the three visible color quadrants of the cylinder need to touch each other in the image. It is therefore necessary to filter the image in order to join almost touching groups of pixels of different colors. This is done by finding any pixel separated on its right by only background (rejected) pixels from a pixel of a different color. The gap considered is from one to four background pixels. The gap is filled in symmetrically by growing the colors on the left and the right of the gap toward each other.

Before searching for the cylinder in the image, the area of interest in the image is restricted so that none of the sky is visible when operating outdoors. Otherwise, the sky is often confused with one of the quadrant colors. Region growing is then performed on the selected part of the image using local blob coloring to identify the four-connected regions of the image which contained any mixture of the cylinder's colors [17].

31

Without any analysis as to shape, the largest such region is taken as corresponding to the cylinder. This is not unreasonable since the colors were quite artificial in appearance, and the background is typically a natural outdoors scene. Also, in a later error-checking stage, the bounding boxes of the cylinder in the left and right stereo images are compared to make sure that they are similar to each other.

The bounding box of the cylinder is then analyzed to find the individual quadrants. The same region-growing technique used above is again applied here several times more. But this time, instead of growing regions of any mixture of the colors, only regions of single colors are considered. The largest such region is labeled the "middle" region. Looking to the left of the middle region, the largest region of a different color is labeled the "left" region. Similarly the largest region to the right of the middle is labeled the "right" region. Consistency checking is performed to make sure that the regions found make physical sense. Note that it is possible for only one or two regions to be visible.

The final result of this process is the center of mass of the entire cylinder, as well as the bounding boxes and pixel counts for each of the quadrants.