Sensory Fusion for Planetary Surface Robotic Navigation, Rendezvous, and Manipulation Operations

Terry Huntsberger, Yang Cheng, Eric T. Baumgartner, Matthew Robinson, Paul S. Schenker
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109
(Terry.Huntsberger, Yang.Cheng, Eric.T.Baumgartner, Matthew.Robinson, Paul.S.Schenker)@jpl.nasa.gov

Abstract

The upcoming missions to Mars planned by NASA/JPL in 2009 and 2013 are more ambitious than any previously flown. Included in these missions are extended multi-kilometer traverses, and autonomous rover rendezvous with science targets and man-made structures such as landers. This paper reports some of the ongoing work at JPL in the areas of autonomous sensory fusion of both raw and derived inputs for better localization during long traverses, precision rendezvous operations with both labeled (man-made) and unlabeled (natural) targets, and precision manipulation of targets. We also present the results of some experimental studies done in laboratory and external environments.

1. Introduction

The sensor suites on the JPL technology prototype rovers such as FIDO (Field Integrated Design and Operations) and SRR (Sample Return Rover) shown in Figure 1 span a wide range of spectral bands, stereo baselines, and fields of view. Traditionally, these sensor suites have been used in isolation for such tasks as planetary surface traversal. For example, although distant obstacle information is known from the narrow FOV navigation camera (NavCam) suite on SRR or FIDO, it is not explicitly used at this time for augmentation of the wide FOV hazard camera (HazCam) information for localization. Such augmentation coupled with internal state sensors such as gyros/accelerometers and sun-sensors can greatly improve the localization of the rovers and enable precision autonomous operations. Due to the long round-trip time delay for communication with Mars during a mission (up to 45 minutes), teleoperation is not a control option in such cases.

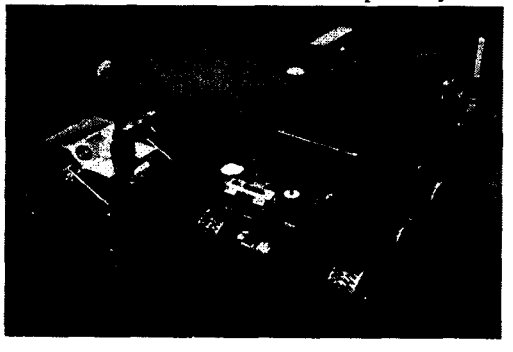

Figure 1. JPL technology prototype rovers: left -Sample Return Rover (SRR); right – Field Integrated Design and Operations (FIDO).

During the past six years, members of the Planetary Robotics Laboratory (http://prl.jpl.nasa.gov) at JPL in Pasadena, CA have been developing and field testing a suite of algorithms that run onboard rovers to address performance issues with respect to rover localization and manipulation capabilities. In this paper, we report on a technique for improving rover localization that fuses the results from a map registration technique, that yields a measure of the change in rover's pose (position and orientation) from one stereo frame to another, with rover wheel odometry, IMU data, and sun-sensor measurments via an extended Kalman filter framework.

Precision navigation from relatively great distances (~125-200 meters) to rendezvous points such as a lander for operations during the upcoming 2013 Mars Sample Return (MSR) mission, as well as from shorter distances to science targets prior to instrument placement on the upcoming 2009 Mars Science Laboratory (MSL) mission, is key to a successful completion of the science objectives. The main differences between the lander and the science target rendezvous operations are the types of targets (man-made vs. natural), and the relative localization of the targets (possible unknown vs. selected from imagery). We have developed a multi-feature fusion algorithm suitable for long-range rendezvous operations that integrates the outputs of horizontal line and wavelet-based visual area-of-interest operators [9] that fuses horizontal, vertical, and diagonal texture characteristics at multiple scales for detection of a man-made target such as a lander and rendezvous with the same from distances greater than 125 meters. This technique is coupled with 3D visual terminal guidance algorithms that extract and utilize cooperative features of the lander to accurately, iteratively estimate structure range-and-pose estimates, and then steer the rover to the rendezvous position. For the shorter range science target rendezvous operations, we have developed a technique that fuses the relative rover pose returned from a multi-target tracking algorithm with the current onboard camera model output to update the location of the

science target relative to an approaching rover in natural terrain.

After navigating to the designated science targets, the primary science objective of the rovers is the robust and reliable acquisition of data using instruments mounted on the end of an instrument arm. The baseline approach for this manipulation task is the localization of the science target with respect to the manipulator arm via the hazard-avoidance stereo cameras located on the rovers. Once the science target has been identified in the image plane of one of the stereo cameras, the 3D range to the target is determined via stereo correlation and triangulation. From this 3D range information, the joint rotations that take the manipulator to the appropriate location in 3D space are determined using the arm's inverse kinematics.

The difficulty associated with this approach is that sources of error tend to accumulate in the stereo system and the manipulator kinematics that ultimately reduce the ability to accurately place the manipulator's end-effector at the desired location. Such error sources include stereo calibration and stereo ranging errors that can yield a range uncertainty of 1.5 cm at 0.75 meters from the stereo pair used on the rover. Errors are also associated with accurate knowledge of the transformation between the stereo reference frame to the base reference frame of the manipulator as well as kinematic errors such as link length uncertainties and joint position uncertainties. Finally, degradations and/or changes in the system configuration due to environmental factors such as launch and landing vibrations and thermal expansion will also effect the ability to accurately position a rover-mounted manipulator at a target of interest. We have developed a precision manipulation technique called HIPS (Hybrid Image Plane Stereo) that fuses the visual feedback from the rover stereo camera pair with manipulator arm joint angles for the robust and reliable positioning of the rover manipulator's end-effector.

Section 2 describes the improved rover localization technique for autonomous long-range navigation. Section 3 describes the suite of algorithms for long-range rendezvous operations. Section 4 presents the multi-step algorithm for precision rendezvous with natural targets. Section 5 describes the HIPS algorithm for precision manipulation. This is followed by an experimental studies section, and finally a summary section including current directions.

2. Fusion of Range Map and Internal State

This section describes the development of a state estimation approach for surface rovers that derives the pose transformation between two range maps of the terrain, and then fuses this information with wheel odometry, inertial measurement sensors, and a sun sensor to generate accurate estimates of the rover's position and attitude. A linear Kalman filter is used to estimate the rate sensor bias terms associated with the inertial measurement sensors and then these estimated rate sensor bias terms are used to compute the attitude of the rover during a traverse. The estimated attitude terms are then combined with the wheel odometry to determine the rover's position and attitude through an extended Kalman filter approach. Finally, the absolute heading of the vehicle is determined via a sun sensor which is then utilized to initialize the rover's heading prior to the next planning cycle for the rover's operations. A more detailed discussion of this technique can be found in papers by Hoffman, et al. [13] and Baumgartner, et al. [3, 4, 5].

2.1. Range Map Registration & Motion Estimation

Significant errors may occur in rover localization when executing obstacle avoidance maneuvers during a long traverse. In such cases, 3D registration of successive range maps can be used to estimate rover motion. Previous work in this area includes maximum likelihood estimation of motion coupled with a Kalman filter update of rover position [19, 20]. A recent method using probabilistic map matching based on a maximum likelihood map similarity estimator was developed by Olson [22, and references therein] for rover localization. His studies, which only included translations in the environment, indicated that sub-pixel accuracy within a discretized pose space was achievable. The only concern raised was that of scaling for pose determination from a totally unknown state, which usually wouldn't be an issue due to other sensors onboard the rover.

We use a range map registration technique that blends Zhang's Iterative Point Matching (IPM) [42] and the surface interpolation of Besl and McKay's Iterative Closest Point (ICP) [6] algorithms to compute a 3D rotation and translation which aligns one set of data points (S_1) with another (S_2). Initially, the data points in S_1 are transformed according to estimated rover motion taken from the commanded movement. Next, the algorithm constructs a k-dimensional (or k-D) tree for the points in S_2, then alternates between two phases: closest-point computation, and estimation of the transformation between corresponding points. The k-D tree enables efficient determination of the closest point in the data set [25].

In the first phase, the algorithm uses the k-D tree to compute the closest point p_2 in the fixed point set S_2 to each point p_1 in the data set being transformed S_1. Triangular faces are then created, with each face containing p_2 and two neighbors in the data set. If the distance between p_1 and the closest point on the interpolated triangles is less than a threshold, then p_1 and

the closest point are added as a pair to a list of point correspondences. We use ICP's closest-point-on-surface approach, rather than using p_2 as the corresponding point for p_1.

The second phase is performed after the correspondences have been computed. The dual number quaternion method [38] computes a rotation matrix and translation vector which minimize the sum of the squared distances between corresponding points. Our previous studies using this technique have shown that the rover's change in heading and lateral position can be accurately estimated, while the estimates of the rover's change in forward position are not as accurate [13, 15].

2.2. Kalman Filter

The rover localization approach is based on the work presented in Baumgartner and Skaar [2] and Yoder et al. [41] that describes a time-independent extended Kalman filter formulation for the determination of the rover's state estimates. Many researchers have also studied the problem of integrating inertial measurement units on mobile robots for improved rover state estimation. These researchers include Barshan and Durrant-Whyte [1], Vaganay, et al. [37], Goel, et al. [11], and Chung, et al. [8]. The IMU modeling approach presented in this paper closely matches the techniques described by Goel, et al. [11] and Chung, et al. [8].

For the IMU sensor, the attitude of the vehicle is described by the roll, pitch, and the yaw (heading) angles where the yaw angle (ϕ) is defined as shown in Figure 2, and the roll and pitch angles are specified relative to the local gravity vector about the rover's local X and Y coordinate frame axis, respectively. The roll and pitch angles of the rover are determined by the 3-axis accelerometer and the heading of the vehicle during motion is determined using the yaw gyroscope rate sensor. When the rover is stationary, the IMU rate sensor bias terms are estimated via a linear Kalman filter. When the rover begins its motion, the rate sensor bias estimator is deactivated and the most recent rate sensor bias term is then used to integrate the yaw rate sensor measurements in real-time to determine the heading of the vehicle throughout the rover's traverse.

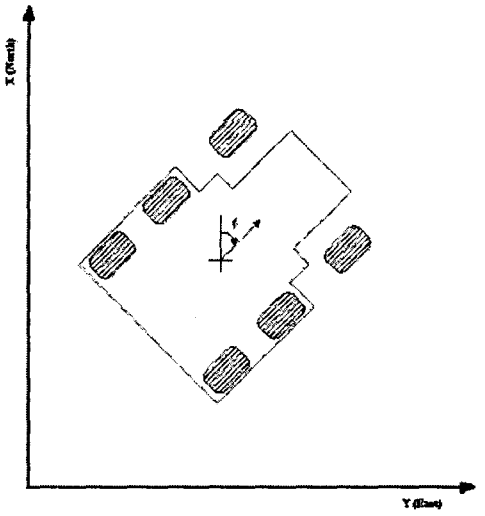

Figure 2. Local rover coordinate system with X aligned with true North, and Y aligned with true East, and yaw angle ϕ defined as the angle between the lateral axis of the rover and true North.

3. Multi-feature Fusion

The nature of features that can be extracted from visual imagery for goal/target-based navigation has a strong dependence on the finite resolution of the sensors on a mobile platform. In addition, the types of information required will vary depending on the relative distance of the platform to the target. At longer distances only heading information will be required, but as the rover approaches closer to the target, rough relative pose and orientation and finally precision alignment information is required. We have developed a suite of algorithms that derive distance and heading information from visual images acquired during a 3 phase approach to a man-made target, in this case a lander (shown in Figure 3).

For the long distance approach phase, fine details on the target are difficult to distinguish and target signatures oftentimes are masked by the surrounding terrain features. In the majority of cases for planetary rover operations, a distinctive texture measure of the target (lander) can be defined for discrimination. We have previously defined a wavelet-based texture signature operator that has been used successfully for automatic target recognition [9], generation of semi-sparse range maps based on texture analysis [13], and closed loop science instrument deployment [16]. A horizontal line detection algorithm that localizes possible lander candidates based on detection of the lander deck is used to constrain the texture signature extraction to a local window of the wavelet coefficient space for each of the lander candidates. The multi-feature fusion algorithm eliminates the false positives in this process arising from features in the surrounding terrain.

The wavelet-based texture signature is derived from the horizontal, diagonal, and vertical coefficients that are generated from the wavelet transform of an image. An example for an image of a lander taken during an approach is shown in Figure 4, where the transform has been taken to three levels. The texture signature uses the strength of the wavelet coefficients in local windows that are sampled at each of the levels, with a normalization

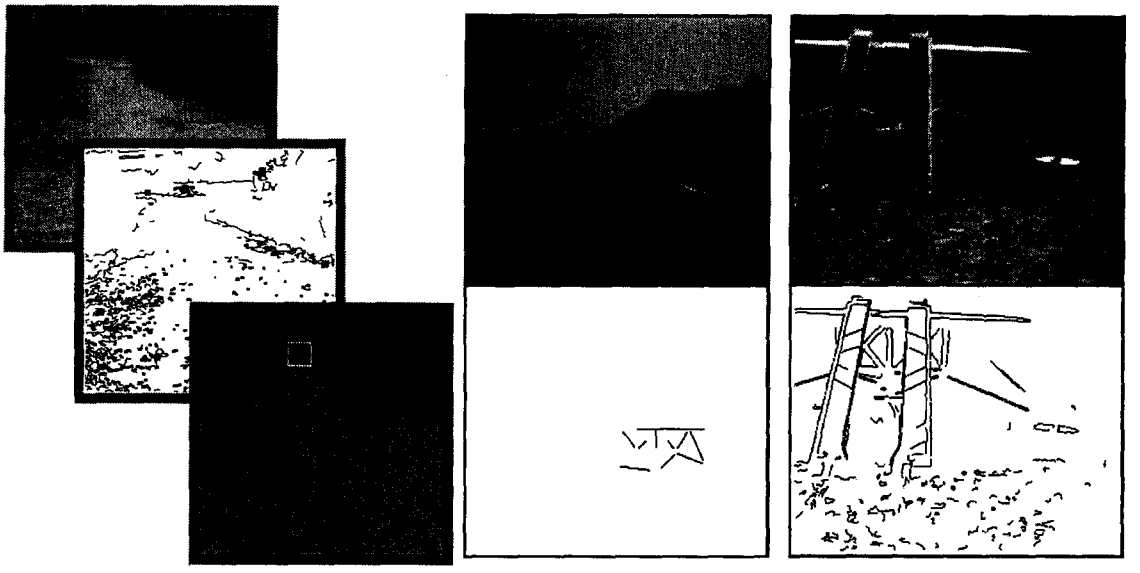

Figure 3. The phase approach to lander from 125+ meters away. From left to right: image montage showing from top to bottom - long-range approach with lander candidates from wavelet texture signature, horizontal line detection, and fused result (shown in box); mid-range montage with original image on top and extracted lander deck and truss structure features on bottom; close-range montage with original image on top and extracted ramp and lander features.

based on local window size. The expression for the signature is:

$$TS(a,b) = \frac{1}{L}\sum_{i=0}^{L} \frac{\sum_{j\in n}\sum_{k\in m} |W_i(a+j,b+k)|}{\text{window_size}}$$

where L is the number of levels in the wavelet coefficient

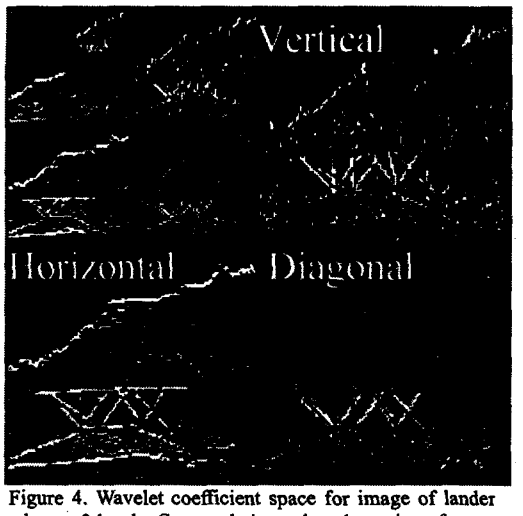

Figure 4. Wavelet coefficient space for image of lander taken to 3 levels. Gray scale is used to show sign of terms with light gray being positive, white being negative, and black being zero.

space, n and m are the sampling window's height and width respectively, *window_size* is the product of the two, and $W_i(a,b)$ is the wavelet coefficient centered at (a,b) for each of the horizontal, diagonal, and vertical wavelet channels. This measure fuses the contributions from the multiple scales through the normalized summation over L, can be used to generate a single signature by averaging over the three channels, or can be used to generate a vector signature that contains both the channel and scale resolution.

During the mid-range phase of the rendezvous approach, parallel line features are extracted from the lander truss structure. In order to correctly distinguish the truss structure, the lander deck is detected first, which appears as a pair of nearly horizontal parallel lines. Any parallel lines below the deck are considered part of the truss structure. The engineering model of the lander is known and constrains the expected geometry of the truss structure. The average distance between the parallel pairs is used to estimate the distance from the target, and the center of mass of these parallel lines is used to compute the heading direction.

The close-range ramp alignment algorithm includes three major steps: feature extraction, feature match, and pose estimation. Features in this case include six stripes that were deliberately arranged on the ramps leading up to the lander deck so that any two stripes have a unique combination spatially and topologically. This unique configuration greatly reduced uncertainty and computational complexity. A Canny edge detection

algorithm is applied, followed by straight-line segment extraction. In order to find the stripes, we first look for the ramps, which are defined by a set of long straight and nearly parallel lines.

With a single detected stripe in the image, a linear affine transformation is constructed based on the four corners. If this match is correct, the transformation is used to find other matches. Because there are relatively few stripes (six on the ramps) an exhaustive search is used to pick up the best match. Once the matches are found, the pose and orientation are estimated by fusing the estimates from the outside corners of all detected stripes. A minimum of four stripes is used for safe navigation.

4. Multi-view, Multi-feature Fusion

The previous section described algorithms that exploited the man-made aspects of something like a lander using a single view fusion of information from derived features. The lack of fiducial marks on natural targets obviates the need for autonomous generation of interest points in the images, coupled with tracking of these points between views for the recovery of 3-D information during the traverse to the science targets. Such techniques have been used previously for updating rover localization information in the absence of a specific goal during traverses [23, 24] and for the approach to and manipulation of science targets [16, 18]. Both of these previous efforts used short range sensor information derived from the HazCams on the rovers and thus are limited in use to a range of less than 4-5 meters. We have developed a multi-step algorithm based on images from the mast mounted NavCams (range of 20-50 meters) that the scientists use for initial target selection, an example of which is shown in Figure 5.

Figure 5. Sample NavCam panorama used for selection of a science target for instrument placement operations. Labeled target is about 5.9 meters away.

The navigation and tracking portion of the precision rendezvous algorithm consists of an iterative application of multi-feature extraction [10, 33] followed by stereo range determination of the features [21] that have been matched using a fusion of maximum likelihood cross correlation [23, 24] and homographic transformation [12] estimates of position before and after each leg of the traverse. Since obstacle avoidance is active during the approach, it is not guaranteed that the rover will be facing the target after each leg of the traverse. A 13 step algorithm for this process has been developed and tested in the field [16]. We previously reported on a continuous approach sequence with target tracking but without obstacle avoidance running on two rovers transporting an extended container [31, 32].

The camera models for the first and second image pairs are transferred to the global coordinate frame according to the mast joint angles, rover attitude and position prior to the application of the Förstner interest operator [10] in a pre-defined region about the center of the initial left stereo image taken before the rover moved. The initial position of the science target in the global coordinate frame was given to the system by the ground operator (derived from stereo NavCam images). Since both sets of stereo pairs are in the same global coordinate frame, each point returned by the interest operator can be used to project a strip of the first image along its epipolar line to the second image based on the first and second camera models and range. A strip is projected rather than a single line in order to mitigate small uncertainties in the newly derived camera models. The rough location of the point in this strip in the second image is determined using maximum likelihood cross correlation [23, 24], followed by the derivation of homographic transform parameters [12] between the two images along this strip to determine the precise location of the feature point in the second image. The range of the feature point is then recomputed at the finest level in the stereo pyramid [21]. All of the features ranges are fused in a least square error sense with outlier rejection to determine the rover pose relative to the target, and the relative position of the target is corrected giving the new drive target and mast pointing parameters for the next leg of the traverse.

5. Hybrid Image Plane Stereo (HIPS)

We have developed a vision-based manipulation technique known as Hybrid Image-Plane/Stereo (HIPS) manipulation that addresses the issues raised in the introductory section with respect to pointing precision. This technique is an extension of the work by Skaar, et al. [7, 34] which has been come to be known as Camera Space Manipulation (CSM). In the CSM approach, widely separated cameras are utilized to determine the direct relationship between the joint rotations of a manipulator and the image-plane (or camera-space)

appearance of a "cue" located on the manipulator's end-effector. The inverse problem is then solved to determine the joint rotations that drive the end-effector to a particular image-plane target location in the participating cameras without regard to any physical reference frame. This method achieves excellent terminal pointing precision (sub-millimeter) when the participating cameras are spaced nearly perpendicular to each other. Unfortunately, widely-spaced cameras on a rover platform is difficult to achieve due to the finite size of the rover and due to the use of existing rover cameras which are configured as stereo pairs for rover navigation needs (i.e. the addition of manipulator-specific cameras is unattractive due to mass, volume, power, and complexity issues).

The HIPS manipulation technique makes use of the basic idea underlying the CSM method - the generation of camera models using visual sensing of the manipulator's end-effector and the use of these camera models to drive the end-effector to a target location without regard to any "real" physical reference frame. In HIPS, the calibration of each of the stereo cameras is conducted using a visual "cue" located on the manipulator's end-effector. This is implemented by positioning the end-effector throughout a set of joint rotations that span the workspace of the manipulator as well as the image plane of each of the stereo cameras. The 3D location of the end-effector at each joint rotation set is computed using the nominal arm kinematics. Once this data set has been collected, the parameters in the stereo camera models are determined. It is important to note that the camera model developed may or may not have any relationship to the camera model generated using an accurate calibration fixture. However, the camera models generated do accurately reflect the relationship between the position of the end-effector (known via the manipulator's forward kinematics) and image-plane appearance of the end-effector in each of the stereo cameras.

Once these manipulator-generated camera models have been determined as described above, a target in the image plane is selected and these camera models are utilized to solve for the 3D location of this target via stereo correlation and triangulation. Again, the actual 3D location of this target may be quite different from the 3D location determined by the manipulator-generated camera models. The important point is that the 3D location determined by the manipulator-generated camera models is accurate with respect to the manipulator which is the means by which the end-effector positioning will be realized. Thus, with the 3D location of the target known, the nominal inverse kinematics are used to solve for the joint rotations of the manipulator that places the end-effector at the desired target location.

This approach addresses the systematic errors that are present in the baseline technique (separate camera calibration and manipulator kinematics), however, stochastic errors that occur due to finite image-plane cue detection precision, inaccurate knowledge of joint angles, etc. are not necessarily accounted for using the HIPS manipulation technique. Therefore, the manipulator-generated camera models are updated throughout the trajectory that takes the end-effector to the target by identifying the image-plane location of the visual cue on the end-effector and by using this additional information to re-compute the parameters in the camera model. In this way, the manipulator-generated camera models are refined such that they are quite accurate in the vicinity of the terminal target location.

6. Experimental Studies

This section reports on a number of experiments that have been run to test the algorithms described in the previous sections. These experiments have been done using the SRR and FIDO technology platforms, shown in Figure 1, in both laboratory and field environments.

SRR is a four-wheel drive, independently four-wheel steered rover platform with a split differential and independently-controlled shoulder (or rocker) joints. FIDO is a six-wheel drive, independently six-wheel steered rover platform with a rocker-bogie suspension system. FIDO has a 4-degrees-of-freedom (DOF) mast that extends to 1.94 m when deployed. The mast-head houses a stereo PanCam (panorama camera), a stereo NavCam (navigation camera), and an Infrared Point Spectrometer (IPS). The NavCam is a low spatial-resolution (640X480), monochrome, wide field-of-view (45°) stereo imaging (25 cm baseline) system used for traverse planning. In addition to the mast, FIDO has a 4-DOF Instrument Arm with a color microscopic imager mounted on the end-effector. SRR has a 3-DOF manipulator arm that is used for manipulation and doubles as a sensing mast using the same NavCam stereo setup as FIDO. FIDO has two other sets of wide field-of-view (112°), narrow baseline (15 cm) stereo cameras, the Front HazCam and the Rear HazCam. The HazCams are used to provide range data for autonomous hazard avoidance algorithm for obstacle detection during rover traverses. The front HazCam is also utilized to choose science targets for in-situ instruments mounted on the Instrument Arm. SRR has a single front HazCam system with the same parameters as FIDO. Both rovers carry a commercial Crossbow three axis accelerometer and gyroscope IMU package The computing platform on

both rovers is a PC/104 266Mhz, Pentium-class CPU with a VxWorks 5.4 real-time operating system.

The first series of experiments was done to test the rover localization algorithms discussed in Section 2. To test the IMU rate sensor bias estimation technique, data was collected from the FIDO's IMU for approximately 10 minutes while the rover remained stationary. Nine separate data sets were collected on different days at different times so that the temperature of the IMU varied for each of the data sets. The bias estimator operates on the IMU data at 200 Hz (the same rate as the IMU sampling). Results from two of the nine data sets are shown in Figure 6. In these two data sets, the filter converges rapidly within 0.5 seconds from the start of the run. Also due to the collection of the data sets at

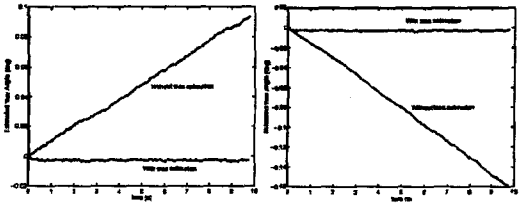

Figure 6. Estimated yaw angle results from two data sets out of nine acquired during different times of day for a stationary FIDO rover. The yaw angle is plotted vs. time in secs over a 10 minute time frame for each data set. The horizontal line in each plot is the Kalman corrected yaw angle and the sloped line is the uncorrected yaw angle.

different days and times, the rate sensor bias terms converge to different values (approximately 0.01 degrees per second for the first case and -0.02 degrees per second in the second case). Despite the different values for the estimated rate sensor bias terms, the estimate of the yaw angle remained quite stable when compared to the rate sensor integration without bias estimation.

The second study tested the fusion of the motion information derived from matching of range maps between successive views as described in Section 2.1 and the wheel odometry and IMU readings using an extended Kalman filter described in Section 2.2. This study was done indoors and thus the sun sensor was not used as an input. A birds-eye-view of the experimental runs is shown in Figure 7. The results for these runs indicated the mean localization error was decreased by a factor of 4 (mean error of 1.38 m for dead reckoning vs. 0.4 m for Kalman filtered) over a series of 12 meter runs. Further details can be found in Hoffman, et al. [13] and in Baumgartner, et al. [4, 5] for the FIDO rover in field trial situations.

The next series of experimental studies field tested the long range rendezvous sequence shown in Figure 3 in the Arroyo Seco at JPL. A representative image from the

Figure 7. Aerial view of experimental laboratory runs for comparison of dead-reckoned and Kalman-filtered approaches to a designated goal.

sequence taken at 25 meters from the lander is shown in Figure 8, where the texture signatures in a localized window about the lander for the first three levels of wavelet decomposition are illustrated. The average error in placement for the rover in relation to the ramps over 15 runs was 1.8 cm in the lateral and 2.4 cm in the longitudinal directions, with a average heading error of 2.7°. Representative images of the rover taken at the long, mid and close range offsets to the lander are shown in Figure 9.

The next series of experimental studies tested the multi-view, multi-feature precision rendezvous with natural targets using FIDO in the Arroyo Seco at JPL. A series of 11 runs was done with a target remotely selected with NavCam images under the WITS (Web Interface for TeleScience) interface that is used for training of the 2003 MER (Mars Exploration Rover) science team. Some representative NavCam images taken mid-way through the autonomous traverse with feature points shown as white crosses are shown in Figure 10. Over the 11 runs, there was an average instrument arm placement error of 7.5cm (1.3%) and RMS of 2.5cm with an average autonomous approach of 5.9 meters. The spread of errors on the science target is shown in Figure 11, where crosses mark the end-effector placement position, and a 1 cm radius circle gives the desired range of accuracy around the nominal target point.

Our final experimental study tested the HIPS algorithm for precision manipulation. The experimental testbed in the Planetary Robotics Lab at JPL is shown on the left in Figure 12, where the target cues are the

black/white circular decals on a mockup of a taskboard with a circular grip point. The end-effector cue is the same type of target. On the right side of Figure 12, a graph displays the experimental results. The preplanned cue predictions are shown throughout the workspace as light colored circles, and the preplanned cue samples are shown as dark colored circles. These are used to calibrate the system as described in Section 5. A sample trajectory is shown as a connected path with the predictions and samples shown as small and larger circles respectively. There was an average error of 1.5mm in lateral offset and 1mm in distance for 15 tests.

7. Summary and Conclusions

We have described the development and testing of a number of sensory fusion algorithms for planetary surface robotic operations that would be common to upcoming NASA/JPL missions in 2009 and 2013. These include an extended Kalman filter fusion of range map, wheel encoder, IMU, and sun sensor information for better rover localization during long traverses, fusion of multiple view and multiple feature information for precision rendezvous operations, and fusion of visible and manipulator joint angle information for precision manipulation. Experimental results using the FIDO and SRR technology prototype rovers at JPL were presented for flight-like operational scenarios. The localization results indicated an error of ~3% of the distance traveled for position, and an error ~2-3° in heading [4, 5]. The heading error is greatly dependent on the availability of the sun sensor and time of day [35]. For the long range rendezvous with the lander the average error in placement for the rover in relation to the ramps over 15 runs was 1.8 cm in the lateral and 2.4 cm in the longitudinal directions, with a average heading error of 2.7°. For the short range science target rendezvous, there was an average instrument arm placement error of 7.5cm (1.3%) and RMS of 2.5cm over 11 runs with an average autonomous approach of 5.9 meters. Finally in our HIPS manipulation experiments, there was an average error of 1.5mm in lateral offset and 1mm in distance for 15 tests.

Our current research directions include the integration of the short range science target rendezvous and HIPS algorithms into a system that will reduce the placement error (desired accuracy as specified by the Science Team for the 2009 Mars Science Laboratory mission is 1 cm). This approach combines the accuracy of precision placement found in HIPS with the accuracy of precision rendezvous found in the multi-view, multi-feature navigation algorithm.

Acknowledgment

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The work involves important contributions from many colleagues at both JPL and collaborating institutions. We gratefully acknowledge these interactions and note many of the specific developments in references that follow.

References

[1] B. Barshan and H. F. Durrant-Whyte, "Inertial navigation sensors for mobile robots," *IEEE Transactions on Robotics and Automation*, Vol. 11, pp. 328-342, 1995.

[2] E. T. Baumgartner and S. B. Skarr, "An autonomous vision-based mobile robot," IEEE Transactions on Automatic Control, Vol. 39, 3, 1994.

[3] E. T. Baumgartner, P. S. Schenker, C. Leger, and T. L. Huntsberger, ``Sensor-fused navigation and manipulation from a planetary rover,'' in *Proc. Symposium on Sensor Fusion and Decentralized Control in Robotic Systems*, SPIE Vol. 3523, pp. 58-66, 1998.

[4] E.T. Baumgartner, H. Aghazarian, A. Trebi-Ollennu, T.L. Huntsberger, and M.S. Garrett, "State estimation and vehicle localization for the FIDO rover," in *Proc. Symposium on Sensor Fusion and Decentralized Control in Robotic Systems III*, SPIE Vol. 4196, pp 329-336, 2000.

[5] E. T. Baumgartner, H. Aghazarian, and A. Trebi-Ollennu, "Rover localization results for the FIDO rover," in *Proc. Conf. Sensor Fusion and Decentralized Control in Autonomous Robotic Systems IV*, SPIE Vol. 4571, 2001.

[6] P. Besl and N. McKay, "A method for registration of 3-D shapes," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 6, 2, pp. 116-130, 1999.

[7] W. Z. Chen, U. Korde, and S. B. Skaar, "Position-Control Experiments Using Vision," International Journal of Robotics Research, Vol. 13, 3, pp. 199-208, 1994.

[8] H. Chung, L. Ojeda, and J. Borenstin, "Accurate mobile robot dead-reckoning with a precision-calibrated fiberoptic gyroscope," *IEEE Transactions on Robotics and Automation*," Vol. 17(1), pp. 80-84, 2001.

[9] F. Espinal, T.L. Huntsberger, B. Jawerth, and T. Kubota, "Wavelet-based fractal signature analysis for automatic target recognition," *Optical Engineering, Special Section on Advances in Pattern Recognition*, Vol. 37, No. 1, pp. 166-174, 1998.

[10] Förtsner and Pertl, Photogrammetric Standard

Methods and Digital Image Matching, 1986.

[11] P. Goel, S. I. Roumeliotis, and G. S. Sukhatme, "Robust localization using relative and absolute position estimates," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1134-1140, 1999.

[12] R. Hartley & A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, UK, 2000.

[13] B. D. Hoffman, E. T. Baumgartner, T. Huntsberger, and P. S. Schenker, "Improved rover state estimation in challenging terrain," *Autonomous Robots*, Vol. 6, No. 2, pp. 113-130, 1999.

[14] X. Hu and N. Ahuja, "Matching point features with ordered geometric, rigidity, and disparity constraints," IEEE Trans. PAMI, Vol. 16, 10, pp. 1041-1048, 1994.

[15] T. L. Huntsberger, E. T. Baumgartner, H. Aghazarian, Y. Cheng, P. S. Schenker, P. C. Leger, K. D. Iagnemma, and S. Dubowsky, "Sensor fused autonomous guidance of a mobile robot and applications to Mars sample return operations," in *Proc. Symposium on Sensor Fusion and Decentralized Control in Robotic Systems II*, SPIE Vol. 3839, Boston, MA, Sept. 1999, pp. 2-8.

[16] T. Huntsberger, H. Aghazarian, Y. Cheng, E. T. Baumgartner, E. Tunstel, C. Leger, A. Trebi-Ollennu and P. Schenker, "Rover Autonomy for Long Range Navigation and Science Data Acquisition on Planetary Surfaces," *Proc. IEEE International Conf. on Robotics and Automation (ICRA2002)*, pp. 3161-3168, 2002.

[17] C. B. Madsen, "Viewpoint variation in the noise sensitivity of pose estimation," Proc. IEEE CVPR, pp. 41-46, 1996.

[18] M. Maimone, I. Nesnas, H. Das, "Autonomous Rock Tracking and Acquisition from a Mars Rover," Proc. 5th International Symposium on Artificial Intelligence, Robotics, and Automation in Space (iSAIRAS'99), 1999.

[19] L. Matthies, Dynamic Stereo Vision, Ph.D. dissertation, Carnegie Mellon University, 1987.

[20] L. Matthies & T. Kanade, "Kalman filter-based algorithms for estimating depth from image sequences," *International J. Computer Vision*, Vol. 3, pp. 209-236, 1989.

[21] L. Matthies, "Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation," *Int. J. Computer Vision*, Vol 8, No. 1, pp. 71-91, 1992.

[22] C.F. Olson, "Probabilistic Self-Localization for Mobile Robots," *IEEE Trans. on Robotics and Automation*, Vol. 16, pp. 55-66, 2000.

[23] C. F. Olson, L. H. Matthies, M. Schoppers, and M. W. Maimone, "Robust stereo ego-motion for long distance navigation," in *Proc. IEEE CVPR*, pp 453-458, 2000.

[24] C.F. Olson, L. H. Matthies, M. Schoppers, M.W. Maimone, "Stereo ego-motion improvements for robust rover navigation," *Proc. IEEE International Conf. on Robotics and Automation (ICRA2001)*, pp. 1099-1104, 2001.

[25] F. Preparata and M. Shamos, Computational Geometry, An Introduction, Springer, New York, 1986.

[26] S. Roumeliotis and G. Bekey, "An Extended Kalman Filter for frequent local and infrequent global sensor data fusion," Proc. SPIE, Vol. 3209, 1997.

[27] P. S. Schenker, L. F. Sword, A. J. Ganino, D. B. Bickler, G. S. Hickey, D. K. Brown, E. T. Baumgartner, L. H. Matthies, B. H. Wilcox, T. Balch, H. Aghazarian, M. S. Garrett, "Lightweight rovers for Mars science exploration and sample return," *Proc. Symposium on Intelligent Robots and Computer Vision XVI: Algorithms, Techniques, Active Vision, and Materials Handling*, SPIE Vol. 3208, pp. 24-36, 1997.

[28] P. S. Schenker, E. T. Baumgartner, S. Lee, H. Aghazarian, M. S. Garrett, R. A. Lindemann, D. K. Brown, Y. Bar-Cohen, S. S. Lih, B. Joffe, S. S. Kim, B. H. Hoffman, and T. Huntsberger, ``Dexterous robotic sampling for Mars *in-situ* science," *Proc. Symposium on Intelligent Robots and Computer Vision XVI: Algorithms, Techniques, Active Vision, and Materials Handling*, SPIE Vol. 3208, pp. 170-185, 1997.

[29] P. S. Schenker, E. T. Baumgartner, R. A. Lindemann, H. Aghazarian, A. J. Ganino, G. S. Hickey, D. Q. Zhu, L. H. Matthies, and B. D. Hoffman, T. L. Huntsberger, "New planetary rovers for long-range Mars science and sample return," *Proc. Symposium on Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision*, SPIE Vol. 3522, pp. 2-15, 1998.

[30] P. S. Schenker, E. T. Baumgartner, L. I. Dorsky, P. G. Backes, H. Aghazarian, J. S. Norris, T. L. Huntsberger, Y. Cheng, A. Trebi-Ollennu, M. S. Garrett, B. A. Kennedy and A. J. Ganino, R. E. Arvidson, S. W. Squyres, "FIDO: a Field Integrated Design & Operations rover for Mars surface exploration," in *Proc. 6th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS-'01)*, 2001.

[31] P.S. Schenker, T.L. Huntsberger, P. Pirjanian, A.

Trebi-Ollennu, H. Das, S. Joshi, H. Aghazarian, A.J. Ganino, B.A. Kennedy, and M.S. Garrett, "Robot work crews for planetary outposts: Close cooperation and coordination of multiple robots," *Proc. Symposium on Sensor Fusion and Decentralized Control in Robotic Systems III,* SPIE Vol. 4196, pp. 210-220, 2000.

[32] P. S. Schenker, T. L. Huntsberger, P. Pirjanian, E. T. Baumgartner, and E. Tunstel, "Planetary Rover Developments Supporting Mars Exploration, Sample Return and Future Human-Robotic Colonization," *Autonomous Robots*, Vol. 14, pp. 103-126, 2003.

[33] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE CVPR*, 1994.

[34] S. B. Skaar, W. H. Brockman, and W. S. Jang, "Three-Dimensional Camera Space Manipulation," The International Journal of Robotics Research, Vol. 9, 4, pp. 22-39, 1990.

[35] C. Tomasi and T. Kanade, "Detection and tracking of point features," Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[36] Trebi-Ollennu, T. Huntsberger, Y. Cheng, E. T. Baumgartner, B. Kennedy and P.S. Schenker, "Design and Analysis of a Sun Sensor for Planetary Rover Absolute Heading Detection," *IEEE Trans. Robotics and Automation*, Vol. 17, No. 6, pp. 939-947, 2001.

[37] J. Vaganay, M. J. Aldon, and A. Fourinier, "Mobile robot attitude estimation by fusion of inertial data," *Proceedings of the International Conference on Robotics and Automation*, pp. 277-282, 1993.

[38] R. Volpe, "Mars rover navigation results using sun sensor heading determination," in *Proc. IEEE/RSJ International Conference on Intelligent Robot and Systems*, vol. 1, pp. 460-467, 1999.

[39] M. W. Walker, L. Shao, and R. A. Volz", "Estimating 3-D location parameters using dual number quaternions," CVGIP: Image Understanding, Vol. 54, 3, pp. 358-367, 1991.

[40] Y. Wu, S. S. Iyengar, R. Jain, and S. Bose, "A new generalized computational framework for finding object orientation using perspective trihedral angle constraint," IEEE Trans. PAMI, Vol. 16,10, pp. 961-975, 1994.

[41] J.-D. Yoder, E. T. Baumgartner, and S. B. Skaar, "Initial results in the development of a guidance system for a power wheelchair," *IEEE Transactions on Rehabilitation Engineering*, Vol. 4, (3), pp. 143-151, 1996.

[42] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," International Journal of Computer Vision, Vol. 13, 2, pp. 119-152, 1994.

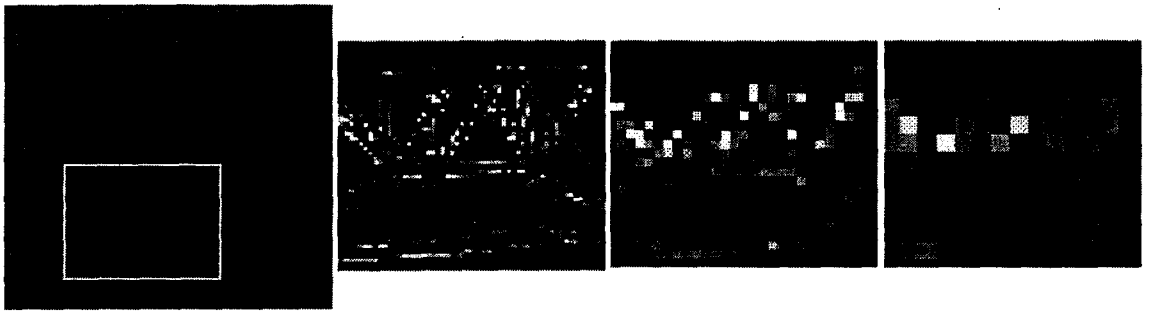

Figure 8. Example of texture signature values for local window in image of lander taken with mast mounted NavCams from 25 meters. From left to right: Original image with lander area boxed, Level 0 texture signature values, Level 1 texture signature values, Level 2 texture signature values. Gray scale is used to give relative values of texture signature, which is the fused value of the horizontal, diagonal, and vertical values for each pixel in the window. The target detail decreases with increasing levels of wavelet decomposition (see Figure 3). The total fused value for the texture signature window (see Equation 1) at Level 0 is 0.82, Level 1 is 0.86, Level 2 is 0.94 as compared to the background range of 0.0 to 0.45 in the same sized window.

Figure 9. Representative images of FIDO rover during long range lander rendezvous sequence in the Arroyo Seco at JPL. From left to right: Range of about 125 meters; range of about 20 meters; rover aligned with ramps at range of 50 cm.

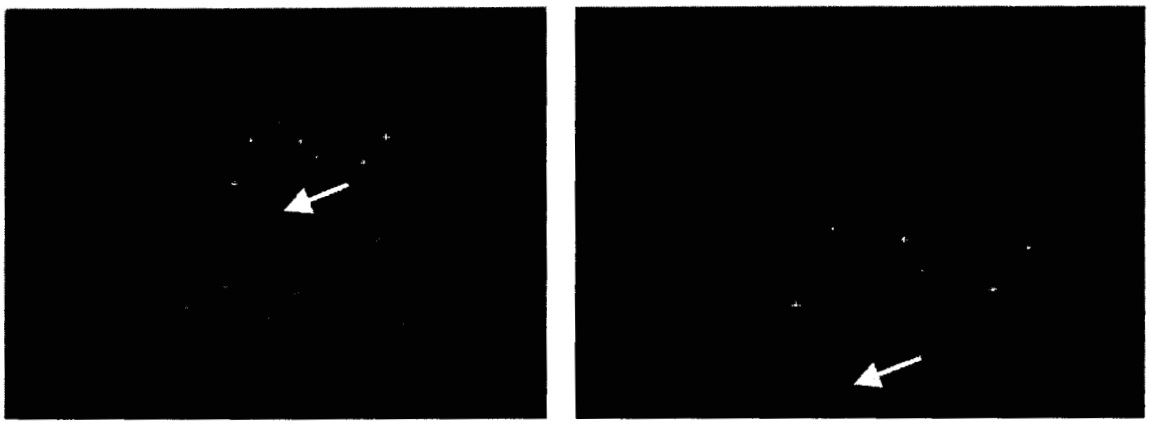

Figure 10. NavCam view of scene with feature points marked with white crosses and designated target shown with arrow. Left image) 3.83 meters from target; right image) 2.66 meters from target

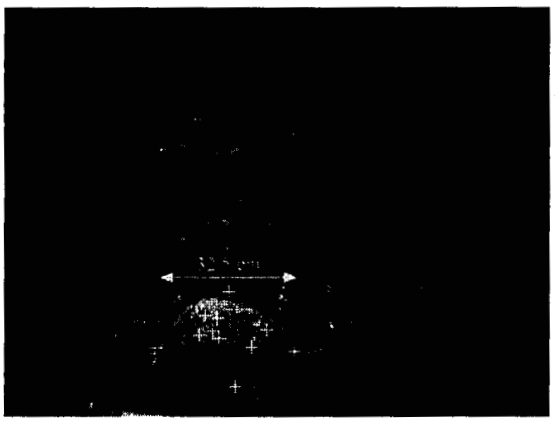

Figure 11. Instrument arm placement results for 11 trials in the Arroyo Seco at JPL. HazCam view of the target taken from 85 cm away. Crosses are arm end-effector positions from all trials, and circle is nominal 1 cm radius area around designated target.

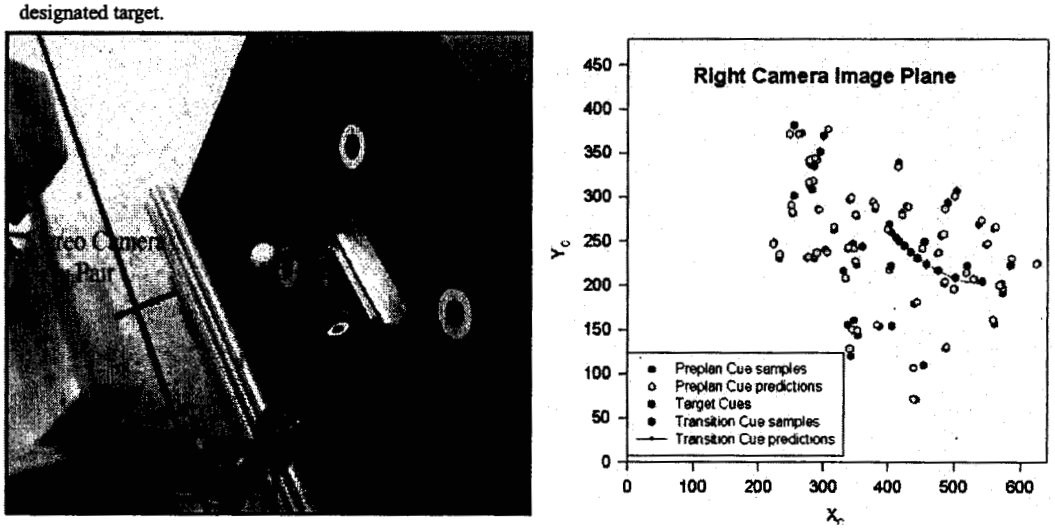

Figure 12. HIPS test setup and results. Left: Task Board with target and end-effector cues; Right: Experimental results plotted in the X/Y image plane.