# HIPS-Assisted Target Handoff between Camera Pairs

Kevin Nickels

*Abstract*— The requirements for planetary manipulation systems are becoming more stringent with time. Missions are moving beyond simple driving and instrument placement to include sample handling, target selection from meters away, and precision re-sampling of a location.

In one of these tasks, remote target selection, the target is selected in a high-viewpoint camera pair. The rover begins tracking the target in these cameras as the rover approaches the target. When the rover gets close enough that the target is in view of the low-viewpoint cameras, *target handoff* occurs. Target handoff is the process of identifying the same target in each set of cameras so that the low-viewpoint cameras can commence tracking. Tracking is then done in the lower cameras until the rover is within optimal range for manipulation operations. Target handoff also has applications in areas such as surveillance, security, and multi-viewpoint tracking.

Planetary manipulation systems typically employ two sets of calibrated stereo cameras to achieve precision on the order of one centimeter or ten pixels. As imprecision in handoff translates to location error of the final sample, this level of precision is not sufficient for upcoming missions. By utilizing the method of Hybrid Image-Plane/Stereo (HIPS), initially developed for visually guided manipulation, handoffs have been achieved with errors of about one and a half pixels.

**Keywords:** Stereo Vision, Triangulation, Target Tracking, Target Handoff

## I. INTRODUCTION

Planetary manipulation systems are becoming more agile and capable. From the fixed instruments of Pathfinder in 1990 [9] to the four instruments and five degree of freedom Instrument Deployment Device (IDD) of the Mars Exploration Rovers (MER) in 2003 [4], scientists and engineers are demanding ever increasing levels of scientific productivity and precision from these systems.

One of the most exciting lessons learned, from an engineering systems design standpoint, as a result of the MER project is the scientific value of increasing the level of autonomy of the rover [6].

### A. Driving and Target Approach

As scientists have become familiar with operating a "robotic field geologist" on Mars, the science flow involving manipulation has evolved into several distinct types of command sequences.

In a *drive sol*, the rover "moves tens of meters in the direction of some selected target" [19]. An *approach sol* attempts to move the rover close enough to a target to place it within the imaging volume of the Hazard Avoidance Cameras (hazcams). In order to use the IDD for any in situ observations, two sets of front hazcam image pairs are acquired at the end of any approach drive. The penultimate pair is taken at most 85cm from the final rover position, and the ultimate pair is taken at the final rover position [21]. The penultimate pair is used to verify free space for IDD operations, and the ultimate pair is used for final target selection. A *touch and goal* sol comprises an in-situ analysis (microscopic imagery, Mössbauer spectroscopy, or Alpha Particle X-ray Spectroscopy) and a drive or approach to the next target [21].

Therefore as many as three sols are needed to analyze a novel target: a drive sol to enter a new area and identify the specific target, an approach sol to position the specific target within the IDD work volume and acquire the penultimate and ultimate hazcam image pairs, and finally a touch and go sol in which the specific target is analyzed with an IDD-mounted science instrument and the rover is driven to the area of the next target. Certainly, in some cases fewer sols are needed, as operations can be combined and targets of opportunity can be exploited instead of predetermining a specific target [20].

The degree of autonomy in IDD-related approach and analysis operations has been increasing along with the degree of autonomy of other rover operations. In particular, Visual Odometry, a process that uses image measurements to correct the reported ego-motion of the rover, is utilized to ensure precise target approaches [6]. Several conditional drives of decreasing duration are

K. Nickels is with the Department of Engineering Science, Trinity University, San Antonio, TX 78212, USA knickels@trinity.edu
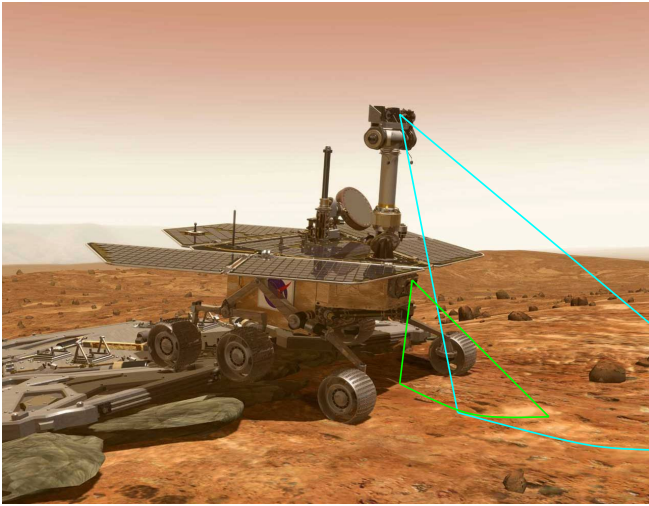
Fig. 1. An illustration of the fields of view of the navcams (upper) and hazcams (lower) on the Mars Exploration Rovers. Background image courtesy NASA/JPL-Caltech.

sequenced at the end of the approach sol, ensuring a final rover position at a range of 1.0 - 1.1 m, the optimal range for IDD operations. This has made approaches to targets much more robust than previously.

### B. Single Cycle Instrument Placement

Due to the time involved in sequencing IDD activities and the potential increased science return, single-sol instrument placement has been studied over the past few years [17], [15], [1], [2], [10], [13]. In the most recent of these, Madison identifies navcam to hazcam target handoff as an area needing more refinement [13].

The mid-2006 software updates to the MER vehicles include a step in this direction: the ability to perform autonomous in situ instrument placement following a successful drive, a process referred to as "go and touch" [6]. However, this activity only offers the opportunity to analyze targets of opportunity in a general region, not the analysis of a specific operator-indicated target.

In the framework considered in this work, a distant target is designated in one set of cameras, typically the Navigation Cameras, or navcams (see Figure 1.) This target is tracked in the navcams as the rover approaches the target. As soon as the target enters the workspace that is shared by the navcams and the hazcams, tracking is transferred from the navcams to the hazcams, and the target is subsequently tracked in the hazcam image pair. This process, known as *target handoff* [2], is the subject of this paper. When the target has been tracked to a range of 1.0 - 1.1 m, the IDD can be used to analyze the target with one of the three available instruments.

### C. Target Handoff

One traditional method for target handoff has been the use of highly-calibrated systems [13]. In this method, both sets of cameras are calibrated, and the target is tracked either in the image-planes or in stereo in the first set of cameras. When handoff occurs, the most up-to-date location of the target is triangulated into 3D by the first set of camera models, and projected back to 2D by the second set of camera models to act as a seed for the tracking to continue. If the viewpoints of the two camera sets do not differ significantly, correlation techniques can be used to refine the target in the second set of images. Other methods for target handoff utilize 3D structure of the environment (e.g. ground plane constraint) to mitigate the need for explicit calibration [7], [12].

In the context of planetary robotics, target handoff between the high-resolution cameras (pancams) and the medium-resolution cameras (navcams), both typically located on a rover's mast, has been reported [11], [13]. However, Madison notes that several methods for handing targets off from the navcam to the hazcam "produced average handoff error of 1 pixel when successful, but only worked in about half of the tests." In addition, "two other methods reliably produced about 10 pixels of error, with none reliably reaching the goal of 1cm (about 4 pixels) of handoff accuracy" [13].

### D. Hybrid Image Plane/Stereo

The Hybrid Image Plane/Stereo, or HIPS, technique, was proposed by Baumgartner et el. [5] and refined by Robinson et el. [18]. HIPS generates camera models based on visual sensing of the manipulator's end-effector, then uses these models to precisely position the manipulator at a target location specified in the image-plane of each camera of the stereo pair. It has been used to increase the precision of instrument placement activities on several different platforms [18], [16].

In the normal formulation, $n$ manipulator-generated targets are used to refine all parameters of a camera model, typically the CAHVOR or CAHVORE models described in [8]. In this case, the equation

$$J(\mathbf{CAHVORE}) =$$
$$\sum_{i=1}^{n} W_i \left[ \left\{ u^i - f_x(\mathbf{P}(\mathbf{\Theta}^i), \mathbf{CAHVORE}) \right\}^2 \right.$$
$$\left. + \left\{ v^i - f_y(\mathbf{P}(\mathbf{\Theta}^i), \mathbf{CAHVORE}) \right\}^2 \right] \quad (1)$$

where the point $\mathbf{P}(\Theta)$ is the position of of a manipulator-mounted fiducial when the manipulator is at joint angles $\Theta$ or, more simply, the forward kinematics of the manipulator. Point $\mathbf{P}(\Theta)$ projects to the image plane location $(u^i, v^i)$. The 3D to 2D mapping function $f()$ is determined by the camera model. Finally, the weighting factor $W_i$ is typically set to unity, but could be used to bias the model fit, for example, towards the expected work volume of a particular instrument.

In some cases, it is helpful to update only the extrinsic parameters of the cameras. Robinson et el. [18] describe the use of this mode in the health analysis of the MER IDD subsystem. The extrinsic-only update is utilized in this work.

## II. RESULTS

In this experiment, several targets are selected in navcam images. The HIPS technique is utilized to modify the nominal camera models, and the combination of triangulation utilizing navcam camera models and projection utilizing hazcam models is compared against manual selection of target points.

### A. Experimental Setup

A planetary manipulator mock-up at the Jet Propulsion Laboratory, shown in Figure 2, was utilized for this work. This five degree of freedom arm is one of the manipulators in JPLs Modular Robotic Test-bed, where robotic arms of various kinematic configurations can be quickly and easily constructed from AMTEC PowerCubes and steel piping. This system has the approximate kinematic structure and camera locations of the Mars Exploration Rover (MER) vehicle [4],[14].

### B. Camera Calibration

Both the navcams and the hazcams were calibrated utilizing unsurveyed methods [18], [8] to create CAHVORE camera models. The resultant camera-centered coordinate system was then translated and rotated according to (crude) manual measurements. This step certainly introduces error into the calibrations, and amplifies the errors introduced by the "standard" kinematic handoff procedure. However, the relatively poor initial camera models do not impact the final accuracy of the HIPS-assisted handoff — similar results to the HIPS results presented below have been achieved with much more well-calibrated systems (e.g. The Field Integrated Design and Operations, or FIDO [22]).
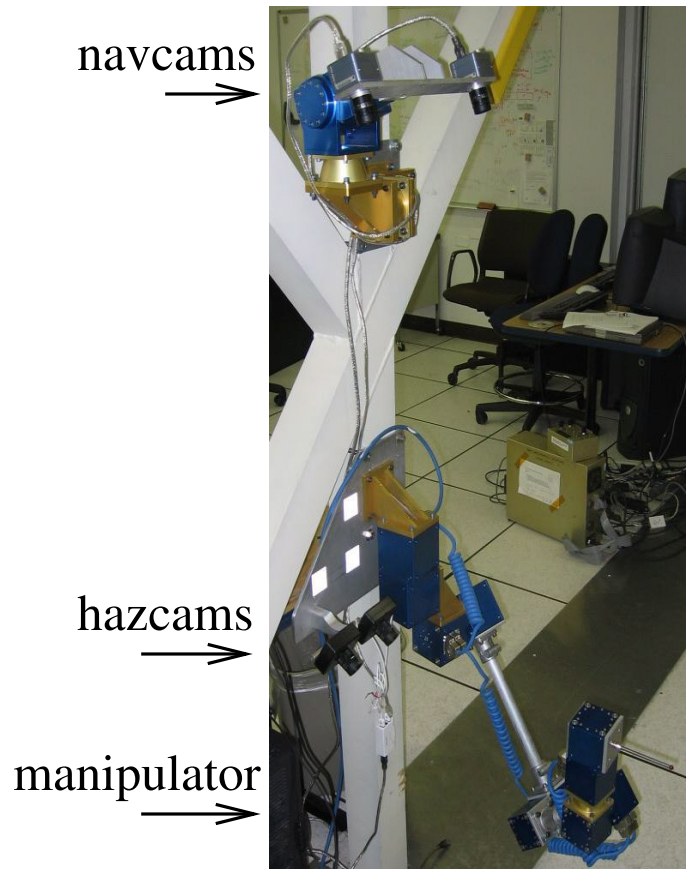


Fig. 2. The Five Degree of Freedom Modular manipulator

### C. HIPS Training

Before the target handoff procedure, a series of 27 images of the manipulator were taken. The manipulator is imaged from both the navcams and the hazcams, and HIPS is utilized to adjust the extrinsic parameters of the nominal camera models to achieve an optimal least-squares fit according to Equation 1. This process, described in detail in [18], is referred to as the preplan trajectory. An example set of images can be seen in Figure 3. The correlation-based fiducial detector used was developed to locate the Mössbauer Contact Plate on MER [3], but any fiducial and its appropriate detector could be used.

As Robinson et el. point out [18], the resulting camera models are *optimized for the workspace region from which the preplan targets are drawn*. In particular, our handoff targets are all outside this region of the workspace. However, as only the extrinsic parameters of the camera models are modified, it is probable that this local optimality will generalize better than models generated from an optimization of all camera parameters.
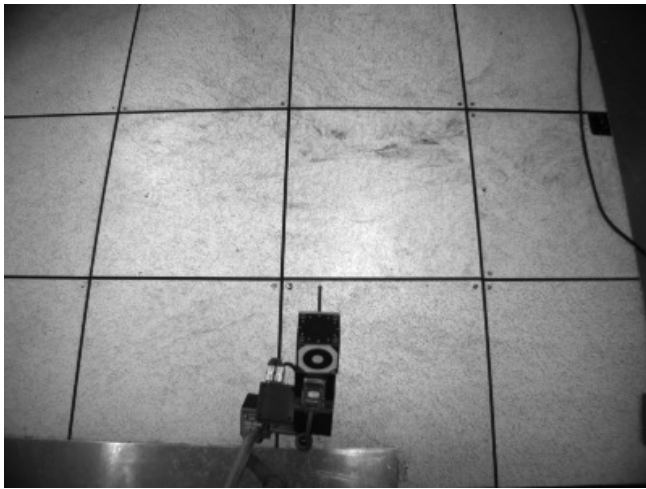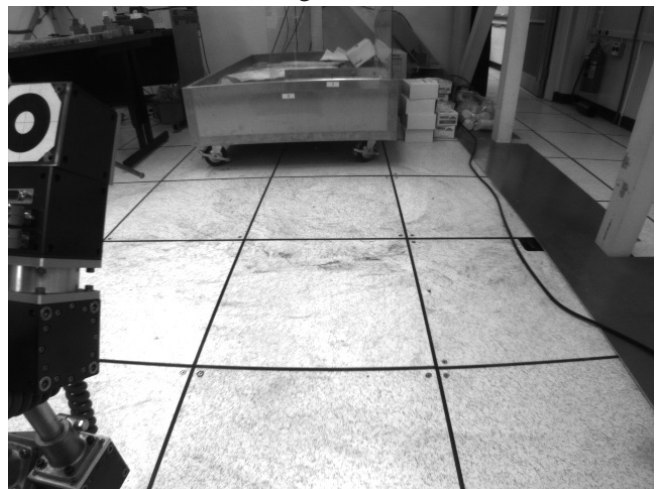
(a) Left navcam
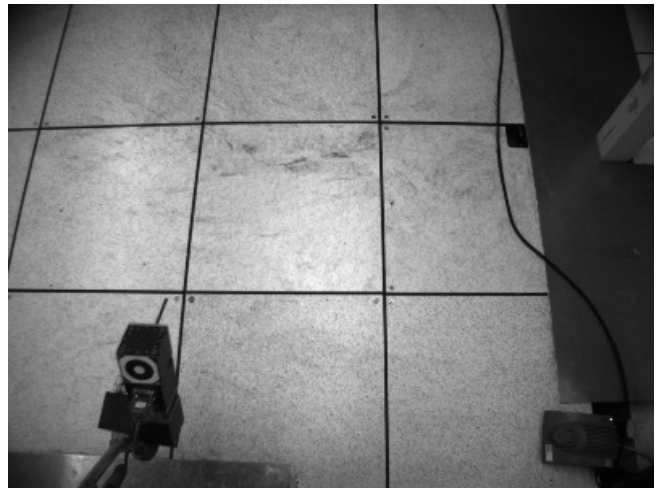
(b) Right navcam
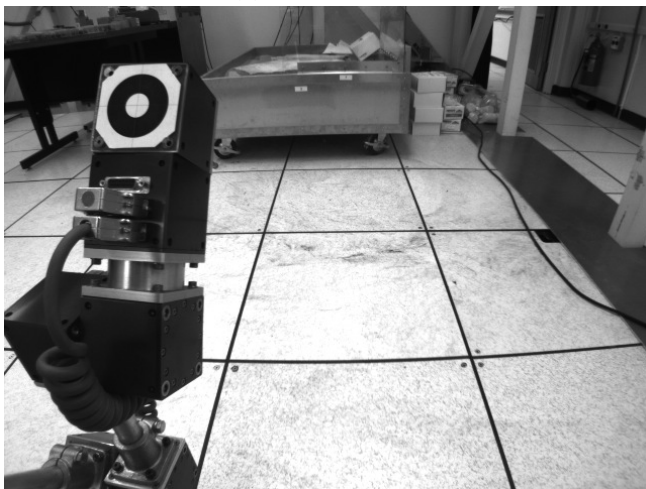
(c) Left hazcam

(d) Right hazcam

Fig. 3.   An Example Training Pose

## D. Handoff Experiment

With the manipulator moved out of the center of the field of view, the five targets shown in Figure 4 were manually selected in the navcams. The selected locations in the left image are accurate to the nearest pixel. The selected locations in the right image are adjusted to maximize the cross-correlation, to sub-pixel accuracy, within a window, to reduce error induced by the manual selection.

Each 2D location was triangulated to a 3D location via either the nominal navcam camera models or HIPS navcam models described in Section II-C. Each of these two 3D locations were then projected to 2D image-plane locations via either the nominal hazcam camera models or HIPS hazcam models, resulting in four possible automatically selected target locations in the hazcams.

The selected target points are then manually located in the hazcam images. Again, the selected locations in the left image are accurate to the nearest pixel, while the selected locations in the right image are adjusted to sub-pixel accuracy. This step introduces some error into the analysis, as it is unlikely that the projection of a pixel center in the left navcam will project to a pixel center in the left hazcam.

The manually selected points are then compared against the automatically selected points. The results are summarized in Table I and Figure 5.

## III. CONCLUSION & FUTURE WORK

In this experiment, the HIPS technique, originally developed to improve the positioning of robotic manipulators, has been used to improve the precision of target handoff between two pairs of stereo cameras. While previous methods have generated errors of up
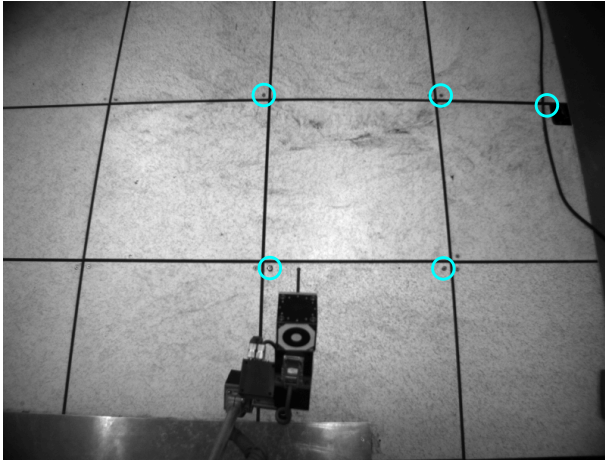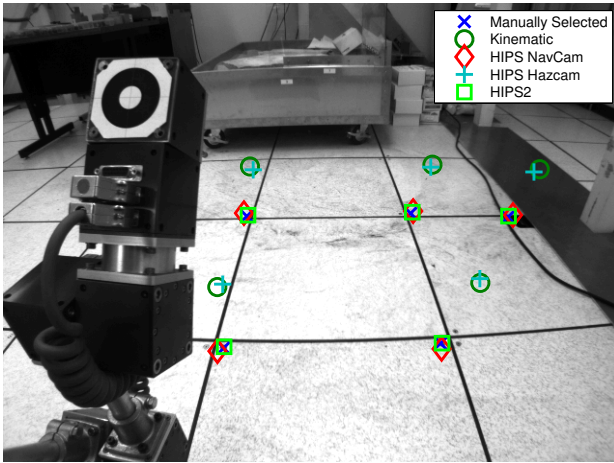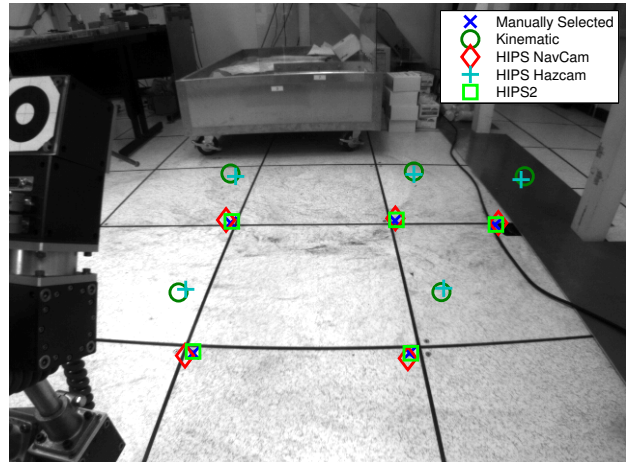
(a) Left navcam
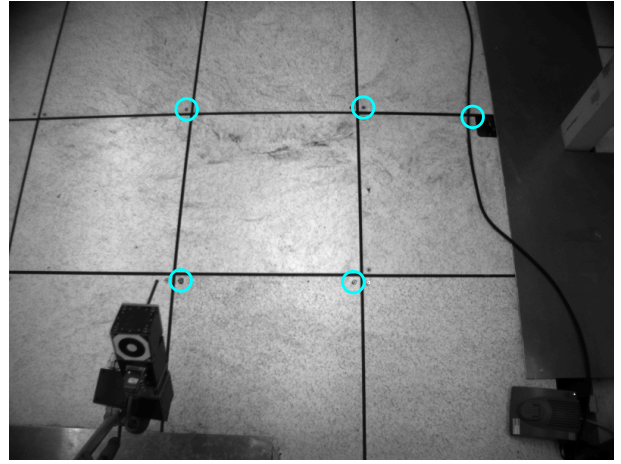
(b) Right navcam

Fig. 4.    Targets, selected in navcams



(a) Left hazcam

(b) Right hazcam

Fig. 5.    Handoff Results

TABLE I

HANDOFF RESULTS. THE MEAN AND STANDARD DEVIATION OF
THE REPROJECTION ERROR ARE GIVEN IN PIXELS, FOR EACH
CAMERA.

| navcam | hazcam | $\mu_L$ | $\mu_R$ | $\sigma_L$ | $\sigma_R$ |
|---|---|---|---|---|---|
| nominal | nominal | 61.62 | 60.68 | 9.00 | 8.05 |
| nominal | HIPS | 60.46 | 59.93 | 12.39 | 11.53 |
| HIPS | nominal | 5.38 | 5.34 | 1.41 | 2.41 |
| HIPS | HIPS | 1.65 | 1.53 | 0.33 | 0.42 |

to four pixels, this method re-projects navcam-selected targets into the hazcams with an average error of about one and one-half pixels. The tools utilized in this study only have selection accuracy of one pixel (in the left image).

This is an extremely small data set, done in a lab environment. Extensive experimentation is needed to confirm these preliminary results. In future work, for example, the algorithm should be implemented and tested on a rover platform, and statistics generated over many more targets. In addition, the experiment should be repeated on natural terrain.

However, this technique for automatically correcting for errors in the extrinsic parameters of pairs of cameras is computationally efficient, straightforward to implement, and shows great promise. It shows particular promise in situations where an extremely rigid and well-calibrated set of cameras is not available, and dynamic in-situ calibration is required.

## ACKNOWLEDGMENT

## REFERENCES

[1] P. Backes, A. Diaz-Calderon, M. Robinson, M. Bajracharya, and D. Helmick. Automated rover positioning and instrument placement. In *Proc. IEEE Aerosp. Conf.*, March 2005.

[2] M. Bajracharya, A. Diaz-Calderon, M. Robinson, and M. Powell. Target tracking, approach, and camera handoff for automated instrument placement. In *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, March 2005.

[3] Max Bajracharya, Matt DiCicco, Paul Backes, and Kevin Nickels. Visual end-effector position error compensation for planetary robotics. *J. Field Robot.*, 2007. , to appear.

[4] E. T. Baumgartner, R. G. Bonitz, J. P. Melko, L. R. Shiraishi, P. C. Leger, and A. Trebi-Ollennu. Mobile manipulation for the Mars Exploration Rover. *IEEE Robot. Automat. Mag.*, 13(2):27–36, June 2006.

[5] E. T. Baumgartner, P. C. Leger, P. S. Schenker, and T. L. Huntsberger. Sensor-fused navigation and manipulation from a planetary rover. In *Proc. SPIE Sym. on Sensor Fusion and Decentralized Cntrl. in Rob. Sys.*, Boston, November 1998.

[6] J. Biesiadecki, C. Leger, and M. Maimone. Tradeoffs between directed and autonomous driving on the mars exploration rovers. *Intl. J. Robot. Res.*, 26(1):91–104, January 2007.

[7] A. R. Dick and M. J. Brooks. Issues in automated visual surveillance. In *Proc. VIIth Digital Image Computing: Techniques and Applications*, pages 195–203, December 2003.

[8] D. B Gennery. Generalized camera calibration including fisheye lenses. *Intl. J. of Computer Vision*, 68(3):239–266, July 2006.

[9] M. P. Golombek et al. Overview of hte Mars pathfinder mission: Launch through landing, surface operations, data sets, and science results. *J. Geophysical Research*, 104:8523–8553, 1999.

[10] T. Huntsberger, Y. Cheng, A. Stroupe, and H. Aghazarian. Closed loop control for autonomous approach and placement of science instruments by planetary rovers. In *Proc. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2005.

[11] W. S. Kim, A. Ansar, and R. D. Steel. Rover mast calibration, exact camera pointing, and camera handoff for visual target tracking. In *Intl. Conf. Adv. Robot.*, Seattle WA, July 2005.

[12] L. Lee, R. Romano, and G. Stein. Monitoring activities from multiple video streams: Establishing a common coordinate frame. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8), August 2000.

[13] R. Madison. Improved target handoff for single cycle instrument placement. In *Proc. IEEE Aerosp. Conf.*, March 2006.

[14] J. N. Maki et al. Mars Exploration Rover engineering cameras. *Journal of Geophysical Research*, 108(E12):8071, November 2003.

[15] I. Nesnas, M. Bajracharya, R. Madison, E. Bandari, C. Kunz, M. Deans, and M. Bualat. Visual target tracking for rover-based planetary exploration. In *Proc. IEEE Aerosp. Conf.*, March 2004.

[16] K. Nickels et al. Vision-guided self-alignment and manipulation in a walking robot. In *Proc. 2006 IEEE International Conference on System of Systems Engineering*, Los Angeles, CA, USA, April 2006.

[17] L. Pedersen, R. Sargent, M. Bualat, M. Deans, C. Kunz, S. Lee, and A. Wright. Single cycle instrument deployment. In *Proc. International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS)*, 2003.

[18] M. Robinson, E. Baumgartner, K. Nickels, and T. Litwin. Hybrid Image Plane/Stereo (HIPS) Manipulation for robotic space applications. *Autonomous Robots*, 2007. , to appear.

[19] S. Squyres et al. Athena Mars rover science investigation. *J. Geophysical Research*, 108(E12):8062, November 2003.

[20] S. Squyres et al. The spirit rover's Athena science investigation at Gusev crater, mars. *Science*, 305(5685), August 2004.

[21] Ashitey Trebi-Ollennu, Eric T. Baumgartner, P. Chris Leger, and Robert G. Bonitz. Robotic arm in-situ operations for the mars exploration rovers surface mission. In *Proc. IEEE Int. Conf. Syst. Man Cybern.*, volume 2, pages 1799–1806, October 2005. also as mer-arm-tr.

[22] E. Tunstel et al. FIDO rover field trials as rehearsal for the NASA 2003 Mars Exploration Rovers Mission. In *Proc. of 9th Intl. Symp. on Robotics & Applications, 5th World Automation Congress*, pages 320–327, Orlando, FL, June 2002.