

Hybrid Image-Plane/Stereo (HIPS) for Orientation Control of Manipulators

Kevin Nickels

Abstract—Planetary exploration systems, operating under severe environmental and operating conditions, have thus far successfully employed carefully calibrated stereo cameras and manipulators to achieve desired precision in instrument placement activities. However, the environmental and functional restrictions imposed by the remote operation of semi-autonomous robots in a harsh planetary atmosphere for long periods of time limit the ultimate operational accuracy of this approach.

This paper builds on an algorithm, referred to as Hybrid Image-Plane/Stereo (HIPS), developed to optimize the positioning accuracy of a manipulator. The HIPS method generates camera models through direct visual sensing of the manipulator end-effector. It estimates and subsequently uses these models to position the manipulator at a target location specified in the image-planes of a stereo camera pair using stereo correlation and triangulation.

While positioning control of manipulation systems is important, orientation control of these systems is also crucial. Many planetary manipulation tasks being considered for the Mars Science Laboratory rover, due to launch in 2009, and subsequent missions, will require precise orientation control of manipulators. This paper studies the effect of using HIPS models to control the position and orientation of manipulator end-effector.

As seen in previous position control experiments, the static version of the HIPS technique reduces position error by almost an order of magnitude. This paper additionally shows that orientation error is reduced by almost a factor of two.

I. INTRODUCTION

The use of vision to aid in manipulation has a long and rich history. Some methods, such as *visual servoing*[1], include visual information directly in the control loop, enabling the system to become robust to inaccuracies in modeling the cameras, the manipulator, or the environment. These adaptive methods have some degree of risk associated with them, however. This risk includes, but is not limited to, fiducial detection failure.

K. Nickels is with the Department of Engineering Science, Trinity University, San Antonio, TX 78212, USA knickels@trinity.edu

Extremely risk-averse applications, such as planetary robotics, have typically utilized careful calibration of the cameras and manipulator to achieve the desired precision. Some *in-situ* image-based corrections have been studied [2], [3], but the analyses to date have only addressed positioning accuracy.

Many planetary manipulation tasks now being developed or contemplated require the control of the *orientation* of the end-effector, as well as the position. For example, core sampling requires the re-insertion of the coring bit into a previously drilled hole[4].

This paper briefly reviews one approach to position control, then studies its effect on the orientation of the end-effector.

A. Hybrid Image Plane/Stereo (HIPS)

The Hybrid Image Plane/Stereo, or HIPS, manipulation technique, was proposed by Baumgartner et al. [5] and refined by Robinson et al. [2]. 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 positioning precision of instrument placement activities on several different platforms [2], [6].

In the initial formulation, referred to as *static HIPS*, the manipulator is moved to n positions, and n image pairs are obtained and analyzed. The 3D position of the manipulator as reported by the manipulator itself is compared against the 2D image of the manipulator in each of the images. This information is used to refine all parameters of a camera model, typically the CAHVOR or CAHVORE models described in [7]. In this case, the equation

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

is minimized with respect to **C,A,H,V,O,R**, and **E**. The point $\mathbf{P}(\Theta)$ is the position of a manipulator-mounted fiducial when the manipulator is at joint angles Θ or, more simply, the forward kinematics of the manipulator. Point $\mathbf{P}(\Theta)$ projects to the image plane location (u^i, v^i) . Thus the image-processing to find the manipulator results in the location (u^i, v^i) . The 3D to 2D mapping function $f()$ is determined by the camera model - thus the predicted location of the manipulator would be (f_x, f_y) . 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 [5]. This process is, at its heart, a nonlinear minimization. While the version used in this work has been optimized for CAHVORE models, and has a long history of use for these models, a more general non-linear minimization could be used on different camera models. Since the pre-existing camera models are used to initialize this fit, this is not a particularly computationally intensive process.

An on-line version of HIPS has also been formulated that acquires and analyzes additional manipulator images as the manipulator approaches the target. For these additional images, the same procedure is utilized, except that the weighting factor W_i is increased on latter images, to bias the models to fit these data better. In this way, the HIPS camera models are dynamically modified to optimize positioning accuracy in the region surrounding the target [2]. This *on-line* version of HIPS is not addressed in this paper.

To use the static HIPS camera models to control the manipulator to a novel target, the target is first designated by an image-plane location in the right and left images and the HIPS camera models are used to triangulate a 3D location. Then the manipulator is then commanded to this 3D location.

II. RESULTS

In this experiment, a manipulator with three visual fiducials on the end-effector was moved in front of a stereo pair of cameras. Images of the fiducials are analyzed to modify the nominal camera models according to the static HIPS technique. The ability of these new camera models to predict the observed orientation of the manipulator in previously unseen positions and orientations is then examined.

A. Experimental Setup

A planetary manipulator mock-up at the Jet Propulsion Laboratory (JPL), shown in Figure 1, was utilized

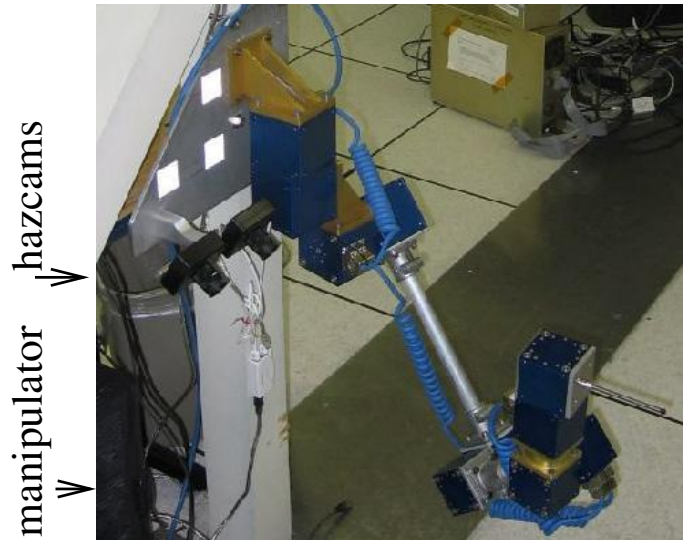


Fig. 1. The Five Degree of Freedom Modular manipulator. The Hazard Avoidance Cameras (hazcams) are used in this work.

for this work. This five degree of freedom arm is one of the manipulators in JPL's 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 [8],[9].

B. Camera Calibration & Data Gathering

The cameras used in this experiment were initially calibrated utilizing JPL-standard techniques [2], [7] to create CAHVORE camera models.

In preparation for this experiment, the manipulator was moved to 69 distinct workspace locations. In each location, the manipulator was moved to nine different orientations in space (azimuth angle $\alpha \in [-0.25, 0, 0.25]$, elevation angle $\epsilon \in [0, 0.25, 0.50]$). Azimuth/elevation angles are equivalent to roll/pitch angles, but are more frequently used at JPL. 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. The resultant data were manually inspected for false alarms and modeling errors, resulting in 456 usable pairs of images.

This set of images was randomly assigned into a training set of 213 images and a testing set of 243 images. The training images were only used for the HIPS training described in Section II-C, and the testing images used for the subsequent experiments.

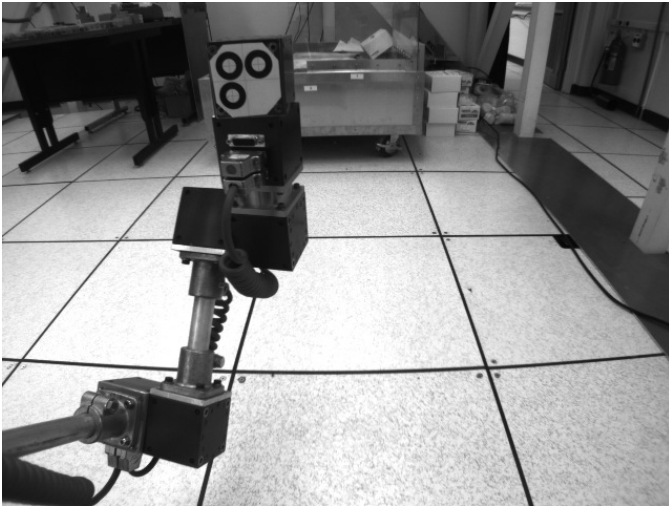


Fig. 2. Left Hazcam Image

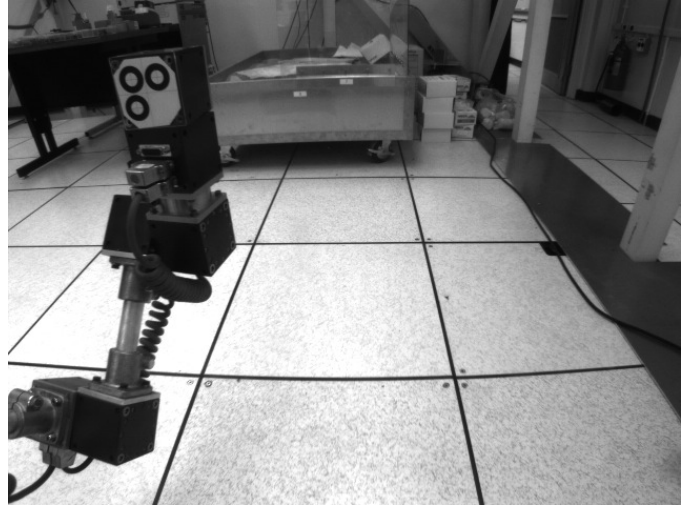


Fig. 3. Right Hazcam Image

Three fiducials are arranged in a triad on the end-effector, to make a local coordinate system that provides both location and orientation information about the end-effector. For details on triad processing, see [10] or [2].

C. HIPS Training

The location of the end-effector as reported by the manipulator control software and as reported by triangulating the detected locations using the nominal camera models is utilized in the HIPS technique to derive new camera models. These models are optimized to achieve an optimal least-squares fit according to Equation 1. This process, described in detail in [2], is referred to as the preplan trajectory. An example set of images can be seen in Figures 2 - 5. Only the location of the upper left fiducial is used for HIPS training.

As Robinson et al. point out [2], the resulting camera models are optimized to reduce the workspace positioning error of the manipulator in the workspace region from which the preplan targets are drawn.

D. Position Control

In order to verify the utility of the HIPS generated camera models, the manipulator is commanded to a variety of novel positions and orientations. The original and HIPS camera models are then utilized to compute the achieved position and orientation of the end-effector. The error metric used below is the difference between the manipulator-reported position and orientation and those computed by the respective camera models, based on visual sensing.

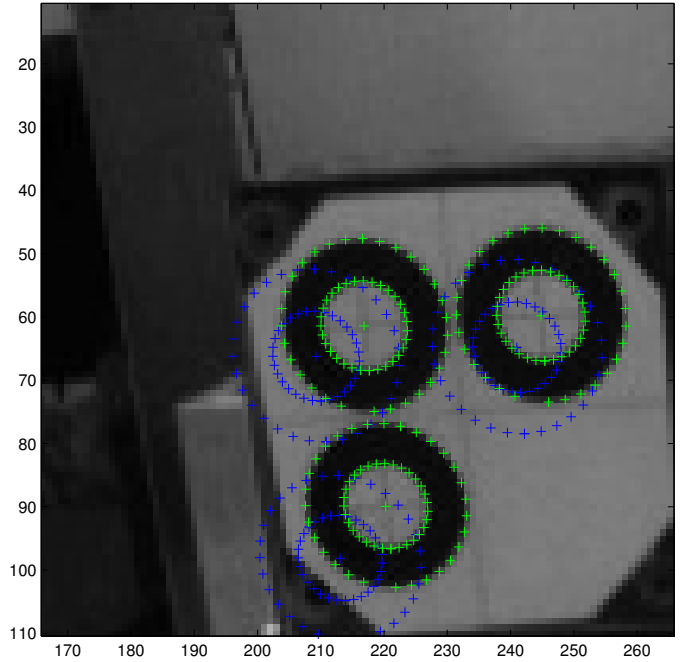


Fig. 4. Left Fiducial Triad. The blue ellipses (or dark, if this paper is printed in greyscale) indicate the predicted location of the fiducials based on manipulator-reported end-effector location and nominal camera models. The green (or light) ellipses indicate the detected locations of the fiducials. Only the position of the upper left fiducial is used in training.

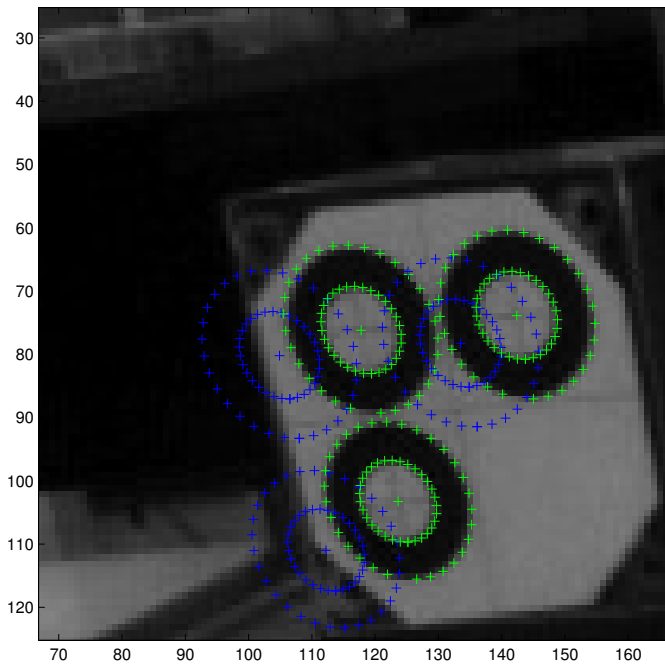


Fig. 5. Right Fiducial Triad

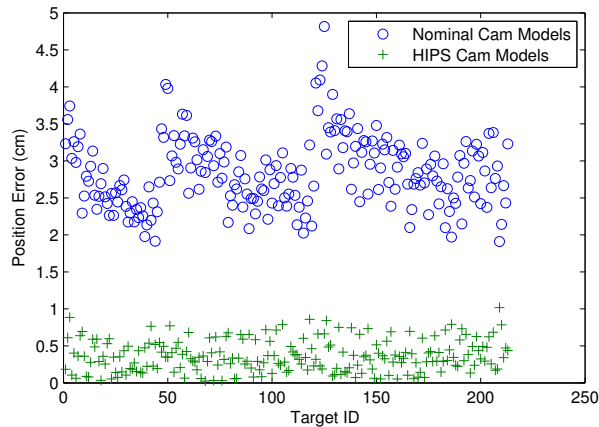


Fig. 6. Workspace Position Error: Difference between manipulator-reported position and as computed with nominal or HIPS camera models

A summary of position error is given in Figure 6. As reported in other work, use of the static HIPS technique reduces workspace positioning error in novel, but similar, locations to the training set by almost an order of magnitude: in this case from 2.84 cm with the nominal camera models to 0.36 cm with the HIPS models. Standard deviation is reduced from 0.48 cm to 0.21 cm.

E. Orientation Control

Many planetary manipulation tasks require control, or at least monitoring, of the orientation as well as the

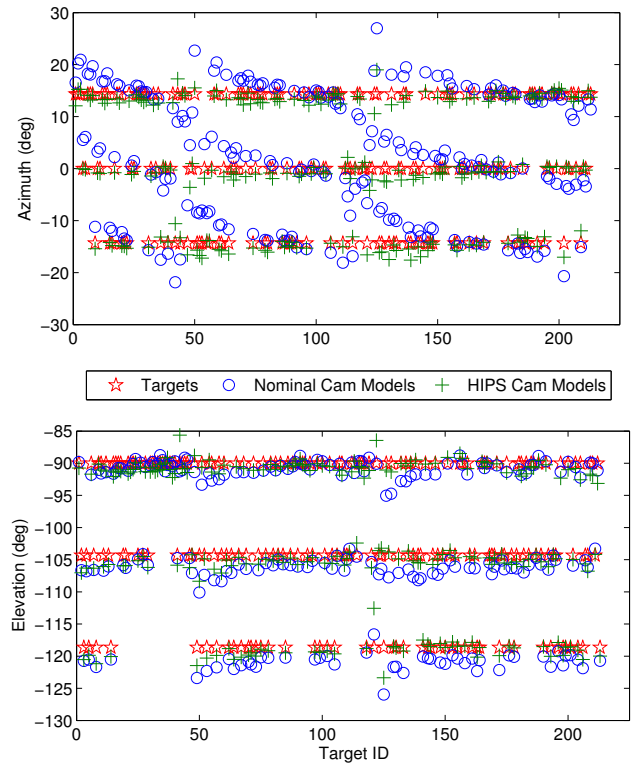


Fig. 7. For each target, the commanded Azimuth and Elevation Angles are compared with the observed angles computed from nominal or HIPS camera models

position of a robotic end-effector. Thus, the impact on orientation control of utilizing vision-guided manipulation algorithms such as HIPS is important.

The azimuth and elevation angles for each point described above, as commanded and as observed, are illustrated in Figure 7 (vs Point ID) and Figure 8. For the testing set, the difference between manipulator and visually reported orientations is reduced from a mean of 2.61 degrees in azimuth and 1.61 degrees in elevation with the nominal camera models to a mean of 1.07 degrees in elevation and 1.03 degrees in elevation with the static HIPS models.

Finally, the solid angle between the commanded and visually computed surface normals is analyzed and shown in Figure 9. The angle between the commanded and visually computed surface normals was reduced from a mean of 3.04 degrees with the nominal camera models to a mean of 1.65 degrees with the static HIPS models.

A summary of each of these error metrics is given in Table I.

III. DISCUSSION

It is clear from this and previous work that the use of camera models derived from manipulator-generated positioning information and visual sensing can improve the position control of a manipulator, at least as measured by visual inspection from the cameras [3], [11], [12], [2], [13]. Increased orientation control has been observed qualitatively on several tasks, such as [10]. To date, however, no study has quantitatively analyzed the impact of these models on the orientation control of the end-effector.

Unfortunately, both the position control and this orientation control experiment suffer from a lack of available absolute position measurement apparatus with sufficient accuracy and precision to observe improvements on the order of millimeters. Thus, strictly speaking these studies only definitively show that, for a novel target specified in the image-plane, the models can precisely predict the manipulator command that will position the end-effector at the desired image-plane locations. Similarly, these data show that a given configuration of three sets of image-plane locations, corresponding to a desired orientation of the end-effector, can be achieved. The quantitative mapping from this precision to an absolute precision in the workspace is not contained in this work.

However, data do show that the relative improvements seen in position control also appear in relative measurements of orientation control. Since (in this work, at least) measurement of orientation is derived from measurements of position, this is expected. However, it has not been shown before.

This work also only utilizes the *static* formulation of the HIPS manipulation technique. There is an *on-line* formulation that improves positioning precision significantly over the static HIPS by gathering more manipulator/visual samples as the manipulator approaches the desired location.

IV. CONCLUSION & FUTURE WORK

In this experiment, the HIPS technique, originally developed to improve the positioning of robotic manipulators, has been analyzed to quantify its impact on the observed orientation of the end-effector.

This is a single data set, done in a lab environment. Extensive experimentation is needed to confirm these preliminary results, and to test the generalizability of these results to using the on-line version of HIPS for orientation control of a rover-mounted manipulator. A practical implementation of this technique would also

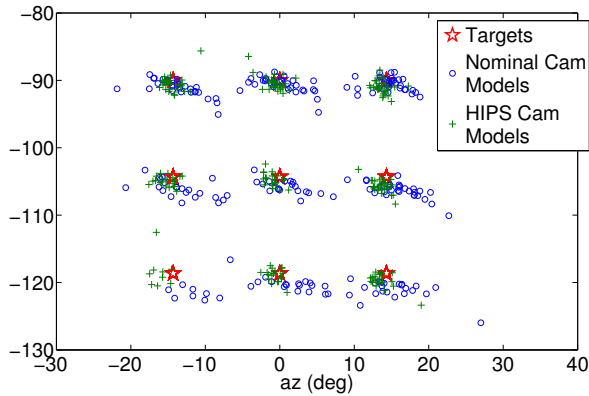


Fig. 8. Azimuth vs Elevation Angles: commanded and computed

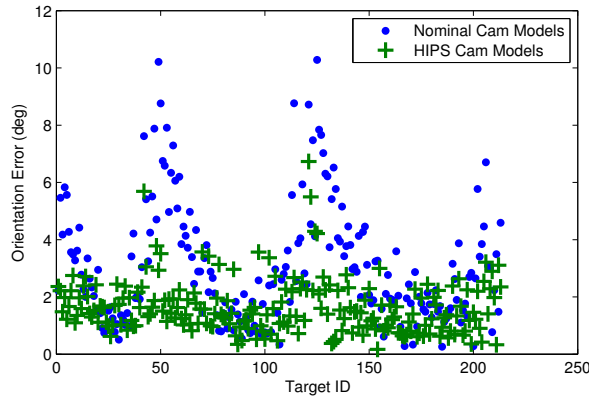


Fig. 9. Predictability of Manipulator Surface Normals. The solid angle between commanded and visually computed surface normals with nominal or HIPS camera models

TABLE I
SUMMARY OF RESIDUALS FOR TEST DATA

Measure	Nominal Camera Models	HIPS Camera Models
Mean Position Error (cm)	2.84	0.36
StdDev Position Error (cm)	0.48	0.21
Mean Azimuth Error (deg)	2.61	1.07
StdDev Azimuth Error (deg)	2.14	0.83
Mean Elevation Error (deg)	1.61	1.03
StdDev Elevation Error (deg)	1.16	0.87
Mean Orientation Error (deg)	3.04	1.65
StdDev Orientation Error (deg)	2.05	0.95

utilize natural features of the end-effector to compute orientation instead of the artificial fiducials used here.

However, this technique for automatically correcting for imprecision in the manipulator/camera system by modifying camera models is effective, 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.

The level of precision in position and orientation control of planetary manipulators in harsh environments needs to increase in order to meet current and future missions. Visually guided manipulation techniques such as HIPS may have an important place in helping to meet those needs.

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. These interactions are gratefully acknowledged, particularly many helpful discussions with Eric Baumgartner, Matthew Robinson, Terrance Huntsberger, and Max Bajracharya. Many of the specific developments are noted in references that follow.

This work was carried out while Dr. Nickels was on an academic leave from Trinity University. Trinity's contribution to his stay at the Jet Propulsion Laboratory during the 2005-2006 academic year is gratefully acknowledged.

REFERENCES

- [1] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.
- [2] M. Robinson, E. Baumgartner, K. Nickels, and T. Litwin, "Hybrid Image Plane/Stereo (HIPS) Manipulation for robotic space applications," *Autonomous Robots*, vol. 23, no. 2, pp. 83–96, Aug. 2007.
- [3] M. Bajracharya, M. DiCicco, P. Backes, and K. Nickels, "Visual end-effector position error compensation for planetary robotics," *J. Field Robot.*, vol. 24, no. 5, pp. 399–420, May 2007.
- [4] C. Weisbin, G. Rodriguez, P. Schenker, H. Das, S. Hayati, E. Baumgartner, M. Maimone, I. Nesnas, and R. Volpe, "Autonomous rover technology for mars sample return," in *Proc. Int. Symp. Artificial Intelligence, Robotics, and Automation in Space*, 1999, pp. 1–10.
- [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, Nov. 1998.
- [6] 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, Apr. 2006.
- [7] D. B. Gennery, "Generalized camera calibration including fish-eye lenses," *Intl. J. of Computer Vision*, vol. 68, no. 3, pp. 239–266, July 2006.
- [8] 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.*, vol. 13, no. 2, pp. 27–36, June 2006.
- [9] J. N. Maki *et al.*, "Mars Exploration Rover engineering cameras," *Journal of Geophysical Research*, vol. 108, no. E12, p. 8071, Nov. 2003.
- [10] 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, Apr. 2006.
- [11] M. Bajracharya, M. DiCicco, and P. Backes, "Vision-based end-effector position error compensation," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, Mar. 2006.
- [12] M. DiCicco, M. Bajracharya, K. Nickels, and P. Backes, "End-effector pose error compensation – error analysis," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, Mar. 2007.
- [13] M. Seelinger, J. D. Yoder, E. Baumgartner, and S. Skaar, "High-precision visual control of mobile manipulators," *IEEE Trans. Robot. Automat.*, vol. 18, no. 6, pp. 957–964, 2002.