Vision-Based End-Effector Position Error Compensation

Max Bajracharya, Paul Backes, Matthew DiCicco
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109, USA
818-393-6613
maxb@robotics.jpl.nasa.gov

Abstract— This paper describes a computationally efficient algorithm that provides the ability to accurately place an arm end-effector on a target designated in an image using low speed feedback from a fixed stereo camera. The algorithm does not require high fidelity calibration of either the arm or stereo camera, and is robust to visual occlusion of the end-effector. The algorithm works by maintaining an error vector between the location of a fiducial on the arm's end-effector as predicted by a kinematic model of the arm and detected and triangulated by a stereo camera pair. It then uses this error vector to compensate for errors in the kinematic model and servo to the target designated in the stereo camera pair.

TABLE OF CONTENTS

1. INTRODUCTION

The ability to autonomously place an arm end-effector on a target selected from a camera image is becoming a common requirement in many space applications, including all foreseeable future Mars landed missions (2007 Phoenix Scout Mission, 2009 Mars Sample Laboratory, and Mars Sample Return), lunar landed mission concepts (construction, etc.), and telerobotic space applications (Robonaut, robotic inspection, etc.). However, past extraterrestrial methods have generally required extremely high fidelity calibration of both the arm and stereo pair and have been directly sensitive to errors in calibration or required computationally intensive methods. Effective industrial solutions have been developed but tend to use extremely fast image feedback and control, which is generally not available in flight systems.

In response to these needs and the given constraints, a computationally efficient but effective algorithm has been developed and implemented. This algorithm locates and triangulates the position of a single fiducial on the end-effector and servos the end-effector to a position designated in the stereo image pair. In doing so, the algorithm maintains an error correction vector between the fiducial's position as predicted by the forward kinematics of the arm and as found by the stereo camera pair. If the fiducial is occluded or imaging is no longer available, the previous error vector can be used locally to reduce positioning error in the final placement. The algorithm requires that the stereo cameras be calibrated well enough to triangulate the position of a point found in the two cameras (although weakly calibrated cameras would be sufficient since metric reconstruction is not necessary, convergence will be significantly faster with more information), and that the arm be calibrated well enough to do roughly accurate small cartesian motions. The transformation between the stereo camera frame and the arm base frame must be approximately known in order to facilitate predicting the fiducial position (see sensitivity study; section?).

2. PRIOR AND RELATED WORK

The rovers used in the 2003 Mars Exploration Rover (MER) mission routinely deploy their Instrument Deployment Device (IDD) and place one of four end-effectors on a target designated by scientists in a front mounted body stereo pair ("hazcam"). However, the placement is generally done using only the kinematic model of the arm and can achieve an accuracy of approximately 1cm [3].

The state of the art for vision-based instrument placement on a planetary surface was the recent use of the Hybrid Image-Plane/Stereo (HIPS) algorithm [4] [10] for the MER rovers. HIPS is based on the Camera Space Manipulation (CSM) [11] concept, but modified to work for very wide field of view cameras with a target close to the camera. However, because HIPS updates its camera model on every iteration (a nonlinear least-squares adjustment), the algorithm is quite computationally expensive. But the benefit of this approach is that it does not require well calibrated cameras and can handle occlusion of the end-effector fiducial.

Other space related work includes Robonaut, a robotic as-

0-7803-7231-X/01/$10.00/
IEEEAC paper # XXX

tronaut's assistant being developed at Johnson Space Center, for which a hand-eye calibration method has been proposed for more accurate grasping [7]. This technique uses an end-effector mounted spherical target that is detected in the robot's head mounted stereo cameras to calibrate its 7DOF arm's kinematic model.

In addition, many other standard visual servoing techniques also exist [5] and have been applied to industrial robots. These include both monocular and stereo vision systems, with either static or end-effector mounted cameras. However, many of these approaches assume a fast image acquisition rate to close the loop around the end-effector. (expand this)

This new algorithm is an attempt to combine standard visual servoing techniques while taking advantage of a reasonably well calibrated stereo camera pair and arm to reduce computation time and image acquisition rate.

3. Algorithm Description

The input to our algorithm is a target selected in the stereo camera pair that provides the position of the target in the stereo camera frame. The arm is then placed in a configuration that allows the fiducial to be seen by the camera; the fiducial is detected, and its position triangulated in the stereo camera frame. The difference between the arm's position as determined by forward kinematics (in the arm frame) and by stereo (in the camera frame) can then be used to reposition the end-effector at the desired location in the camera frame. Because this error is a function of the arm configuration and location in the image, it is a local error, and is repeated as the arm approaches its final position. If the arm and cameras were calibrated perfectly, the compensation error being calculated would be zero and kinematic placement would be sufficient. A block diagram of the algorithm is shown in Figure 1.

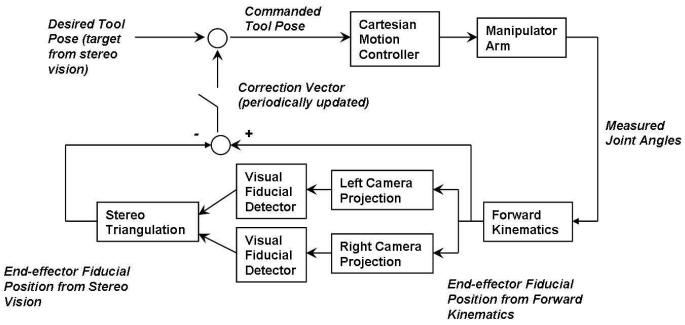

Figure 1. Control system block diagram

Target Selection

The end-effector target is selected in one of the images of the stereo pair and standard subpixel correlation based stereo is used to determine the 3D location of the target. For our camera configuration, there is a range error associated with the selection due to correlation error (generally around 0.3 pixels) leading to less than 1cm of range error. This error is compensated for by using a final approach along the normal of the target until a contact sensor is triggered. Lateral error due to selection is better than 1mm for our camera configuration.

Fiducial Detection

In order to servo the end-effector to the target, the algorithm must be able to localize the end-effector in the stereo camera pair. This is done by detecting the position of a fiducial on the arm's end-effector (Figure 2). Several ways of detecting the end-effector fiducial have been tested. These include the Harris corner detector [12] and a simple correlation based method [6], as well as a global but pyramidal image search and a local search using the arm's forward kinematics and camera projection to predict the location of the fiducial in the image. While the Harris corner detector is fast and invariant to affine transformations, it is prone to false positives, particularly with direct sunlight on the specular metal end-effector. If the arm model is known accurately, then the corner detector in a limited area is effective, but degradation of the arm model causes this approach to fail. The correlation approach (searching for the maximum correlation score of a fiducial template across the image) is more robust for a known fiducial configuration, but is more computationally expensive and not affine invariant, breaking down when the template does not match the projected appearance of the fiducial (for example, when the rotation of the fiducial is not known). A circular fiducial has also been considered, which might allow more accurate subpixel localization due to its larger area, but would require foreshortening compensation and would be harder to detect.

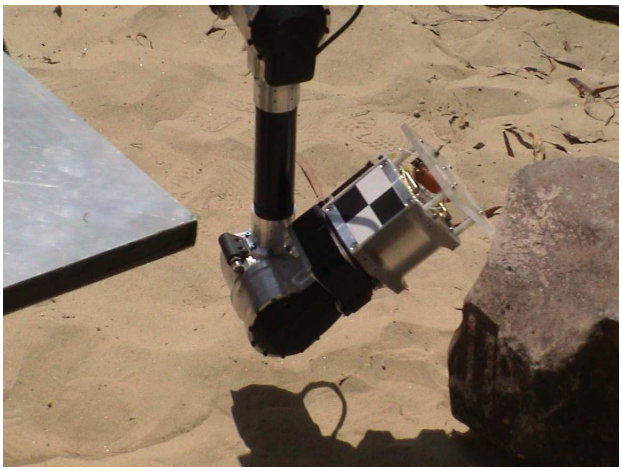

Figure 2. Rocky 8 arm end-effector fiducial

Stereo Triangulation

With the fiducial located in each of the stereo images, the position of the fiducial in 3D is triangulated by finding the point of closest intersection between the two rays projected from the camera. The pixel at which the fiducial was detected corresponds to a ray $r_{\{left,right\}}$ passing through the focal

point of the camera lens. For a baseline b between the two cameras, the distance $d_{\{left,right\}}$ of the fiducial along each ray can be calculated with:

$$[d_{left}\ d_{right}] \begin{bmatrix} -r_{left} \\ r_{right} \end{bmatrix} + b = 0 \quad (1)$$

The camera models are the JPL standard CAHVOR model [13] and account for the CCD and lens parameters, including distortion. The camera are calibrated offline using an unsurveyed calibration technique [14]. The calibration provides the relative position of the two cameras but their location in the vehicle frame are based only on the vehicle's CAD model. As a result, this is the largest source of error in the triangulation of the fiducial.

Motion Correction

The difference between the fiducial's position based on the arm's forward kinematics and the triangulation from the stereo cameras provides an error vector used to compensate for inaccuracies in the arm and camera models during end-effector positioning. This error is a function of the fiducial location in the images and the joint angles of the arm, which makes it valid only in a local region. As a result, as the end-effector is moved closer to the target location, the error vector is recomputed. If the fiducial is ever occluded, the previous error vector can be used and should be valid for small motions.

In general, placing the end-effector relatively close to the target using purely kinematics would be feasible, and consequently only a single iteration of applying the error vector would be required. However, because the initial arm model was not accurate, the algorithm was used in a closed loop control system that allows the arm to start at any location and updates the correction vector as it approaches the target.

Instrument Contact

In order to enforce an approach normal to the target, the end-effector is servoed to a position that is offset along the target normal and then driven along the normal until contact. The target normal is the normal to the best plane fit of a stereo data point cloud surrounding the target. Because the fiducial only provides position information, end-effector orientation is based solely on the arm's kinematics. Once positioned along the target normal, the end-effector is driven along that normal using the last error correction vector until a contact switch triggers contact with the target.

4. TESTBED AND SYSTEM INTEGRATION

This algorithm was implemented and tested on two 5DOF arms. One, labeled "Bluestreak" (Figure 3), is wall-mounted and built with commercial joints and links. The other, mounted on the Rocky 8 research rover (Figure 4), is entirely custom built at JPL. Both have the same DH table and camera configuration, but because Bluestreak is wall-mounted its motion does not affect the camera position (unlike the rover, where arm motion can affect the rover pose due to its suspension system). It also has the advantage of being completely built out of commercial components and is thus low cost and easily duplicated. While the joints are also more accurate, they draw significantly and prohibitively more power than the rover arm joints.

The camera configuration has two cameras with an 8cm baseline pointing 45 degrees down. They have 2.8mm lenses and a 640x480 CCD with 4.65um pixel size. This configuration corresponds to 1 pixel being approximately ??mm in lateral error and a 0.3 pixel error in stereo disparity matching error being ??mm in range error. However, because the algorithm uses a target designated in the same stereo pair that it uses to compute the error vector, it should be able to achieve even better placement performance.

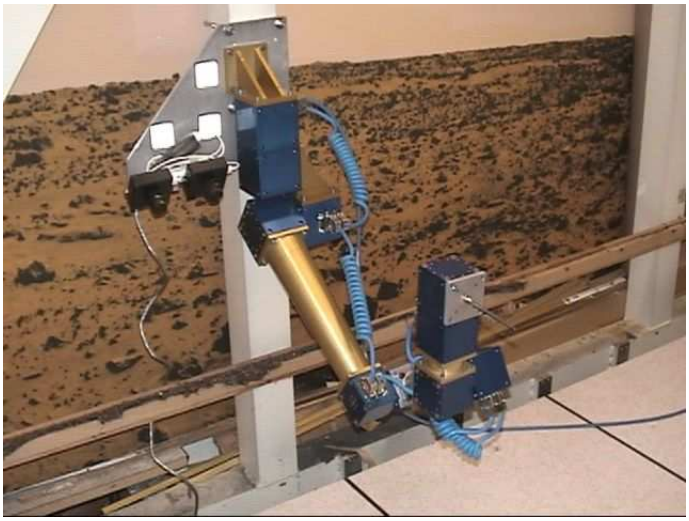

Figure 3. The Bluestreak manipulator

Figure 4. Rocky 8 research rover

The algorithm has been integrated into a complete end-to-end system that demonstrates "single cycle instrument placement" for a planetary rover ([1], [2], [9], [8]). In this scenario, the rover takes an image of a scene and a scientist selects a target on which the arm's end-effector instrument will be placed. The rover then autonomously approaches the target, places itself so that the target is in the arm workspace, deploys the arm and places the instrument on the target.

5. Testing and Results

Initial analysis of the EPEC algorithm was performed via simulation. This simulation consisted of two robots, one with no errors in kinematic and camera parameters and another robot with errors. The analysis scheme is outlined in figure 5 and consists of two stages. In stage one, a target pose is chosen and inverse kinematics is performed with the perfect robot model to create a set of desired joint angles. To simulate a real robot moving to these desired angles, forward kinematics calculations are performed on this joint set with a robot model that contains errors in the DH parameters. The resultant pose is then adjusted by an erroneous camera transform to return a perceived arm location. The difference between the perceived arm location and the initial target pose is calculated and stored as the EPEC error correction vector.

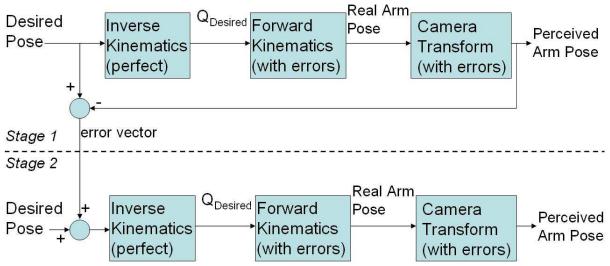

Figure 5. block diagram of epec analysis

This process is intended to simulate what would occur in a real robot when servoing to a position designated in camera space. If the DH parameters used in the forward kinematics matched perfectly with those used in the inverse kinematics, the real location of the robot would match the desired. If the camera transform was perfect, the perceived location would match perfectly with the real location. Introducing errors into the parameters of the forward kinematics and the camera transform

In stage 2, the EPEC error correction vector is used to modify the desired pose and the same process of simulating the robot motion is employed to produce a new perceived final location. The desired pose used in stage 2 does not need to be the same one used in stage 1 to calculate the error vector. In fact, a major goal of this analysis is to determine the effects of using a correction vector calculated at one position to correct errors at an entirely different position in the workspace.

All tests were performed with a 5 DOF arm of similar kinematics to the two experimental systems mentioned above. The 5 degrees of freedom in the arm correspond to 20 DH parameters (4 per link) and the camera transform corresponds to an additional 6 parameters (3 translation and 3 rotation). Not considered in this analysis are errors produced by the camera models used to determine 3D location from the stereo images and errors introduced by poor fiducial detection algorithms. These two error sources will be incorporated into later versions of this analysis.

EPEC Effectiveness

To evaluate the EPEC algorithm's ability to correct kinematic errors, it was first tested at the target. In these tests, the arm was commanded to a desired pose, the error vector was calculated at that location, and then the arm was commanded again at the same spot with the correction vector added. To observe the overall effect of kinematic errors, this process was performed at numerous targets spanning the workspace. At each target, the outlined test was performed 100 times with random sets of parameter variations. Table 1 outlines the results of these tests. The end-effector error was averaged over all 100 trials at each of the 196 targets spread throughout the workspace. These tests were performed three times which varying levels of random error distribution on the parameters. Not surprisingly, when the error vector is calculated at the target and the errors are purely kinematic, the placement error was reduced significantly.

Error (std. dev)	Average Error before EPEC	Average Error after EPEC
0.5deg, 1.0mm	12.2mm	0.3mm
1.0deg, 1.0mm	23.0mm	0.9mm
1.0deg, 2.5mm	25.7mm	1.2mm

Table 1. Summary of simulation results

Locality of Correction Vector

Of great interest in the development of this algorithm is the question of locality. It may not always be possible to observe an end-effector fiducial at the final target location. If error vectors are valid within a wide range of the workspace, an image with a clear fiducial from somewhere in the general neighborhood of the final target could be used to improve final positioning accuracy.

To test this error vector locality, the test outlined above was performed again with one major difference: one initial target location is used to calculate the correction vector and then this correction vector is applied to numerous other targets spread throughout the workspace. The final positioning error can then be observed in relation to the distance from the location where the error vector was calculated.

This test was performed three times with three different initial locations. The first location was chosen in the center of the

arm workspace, the second was chosen in the center of the stereo camera pair field of view, and the final location was chosen at the extreme edge of the workspace.

Again all tests were performed with 100 different parameter sets, each with a normal distribution of error on the 26 parameters. The standard deviation of error on the length and camera translation parameters was 2.5mm and the standard deviation on all angular parameters was 1 degree.

As expected, the targets close to the location where the error vector was calculated show error values in line with the results gained in the first analysis. As the target was placed further away, the final positioning error grew linearly, as shown in figure 6. In addition to being linear, this locality shows a shallow slope, meaning that an error vector can be calculated far from final target location. In order to achieve .5mm positioning accuracy, the error vector can be calculated as far as 16cm away from the target.

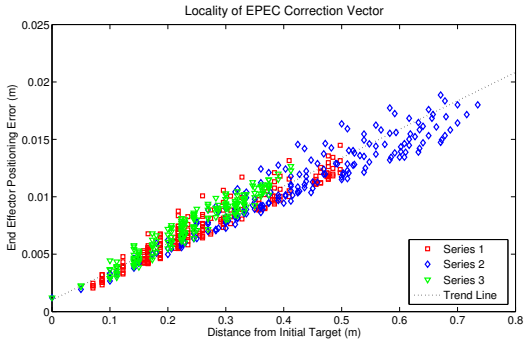

Figure 6. EPEC correction vector applied to different parts of the workspace

Experimental Results

In order to assess the impact of factors not included in this analysis, it was desirable to perform gather experimental data as well to support the above claims. Using the Bluestreak manipulator system described earlier (Figure 3), an experiment was performed similar to the locality analysis described above. The arm was moved into a neutral starting location in front of the stereo cameras. Images were taken and the 3-D position of a fiducial on the end effector was calculated. The error vector was calculated by comparing the calculated 3-D position with the 3-D position calculated through forward kinematics on the robots current joint angles.

By placing a target fiducial inside of the camera space, 3-D target positions were designated and numerous locations throughout the workspace. The arm was commanded to these locations first without using the EPEC correction and then with. Locations of the target and the arm with and without the EPEC correction were recorded with a ground truth laser ranging tool. Figure 7 shows the positioning errors plotted versus distance from the location of the initial error vector calculation.

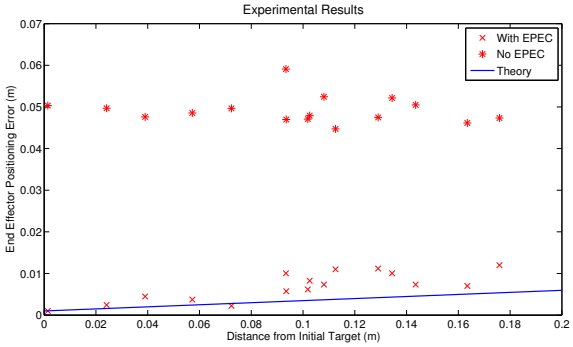

Figure 7. Experiment vs. Theory

As expected, the positioning error without the EPEC correction was significant and was essentially constant for all points tested. All trials where the EPEC correction was used showed a significant reduction in the final positioning error, even with the correction vector taken from as far as 20cm away. Again a linear trend is observable in the correlation between positioning error and distance from the initial error vector calculation. The slope of the experimental data is slightly steeper than the predicted theory from the analysis outlined above. The predicted slope is plotted along with the experimental data for comparison. This is not a surprising result as the camera modeling errors (particularly lens distortion) and imperfect fiducial detection algorithms are likely to introduce additional error.

6. Future Work

Further analysis of this algorithm is currently underway. This expanded analysis will include investigating the sensitivity of the algorithm to changes in the arm and camera models and accuracy of the fiducial detector. This same setup will also be used to compare this algorithm to existing methods for vision guided manipulation. Of particular interest is classification of the computational efficiency of this algorithm.

In addition to algorithm analysis, various ways of improving the current algorithm are being considered. These range from implementation details, like improving the fiducial detection's subpixel accuracy, to algorithmic changes, like modifying the fiducial marker to provide both position and orientation. Because the location of the tip of the end-effector is calculated using the end-effector orientation based solely on the arm's forward kinematics, errors in the arm model could significantly effect tip placement accuracy.

7. Conclusions

A new, computationally efficient algorithm for placing an arm end-effector on a target designated from a stereo camera pair has been developed, implemented, and tested experimentally. Because of its simplicity, the algorithm is extremely easy to

understand, implement, test, and analyze (both analytically and experimentally). As it appears to converge well given a reasonably well calibrated camera pair and arm, it has the potential to be a very robust and effective algorithm, being an ideal candidate for a flight system.

8. ACKNOWLEDGEMENTS

The research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract from the National Aeronautics and Space Administration (NASA), with funding from the Mars Technology Program, NASA Science Mission Directorate.

REFERENCES

[1] P. Backes, A. Diaz-Calderon, M. Robinson, M. Bajracharya, and D. Helmick. Automated Rover Positioning and Instrument Placement. IEEE Aerospace Conference Proceedings, March 2005.

[2] M. Bajracharya, A. Diaz-Calderon, M. Robinson, and M. Powell. Target Tracking, Approach, and Camera Handoff for Automated Instrument Placement. IEEE Aerospace Conference Proceedings, March 2005.

[3] E. T. Baumgartner, et al. The Mars Exploration Rover Instrument Positioning System. IEEE Aerospace Conference Proceedings, March 2005.

[4] E. T. Baumgartner and P. S. Schenker. Autonomous Image-Plane Robot Control for Martian Lander Operations. Proceedings of the IEEE International Conference on Robotics and Automation, April 1996.

[5] S. Hutchinson, G. Hager, and P. Corke. A tutorial on visual servo control. IEEE Trans. Robotics and Automation, vol. 12, no. 5, pp 651-670, Oct. 1996.

[6] J. P. Lewis, "Fast normalized cross-correlation," Vision Interface, 1995.

[7] K. Nickels. Hand-Eye Calibration for Robonaut. NASA Summer Faculty Fellowship Program 2003 Final Report, Johnson Space Center.

[8] L. Pedersen, R. Sargent, M. Bualat, M. Deans, C. Kunz, S. Lee, A. Wright. Single Cycle Instrument Deployment. International Symposium on Artificial INtelligence, RObotics and Automation in Space (i-SAIRAS), 2003.

[9] T. Huntsberger, Y. Cheng, A. Stroupe, H. Aghazarian. Closed Loop Control for Autonomous Approach and Placement of Science Instruments by Planetary Rovers. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2005.

[10] M. Seeling, J.D. Yoder, E. T. Baumgartner and S. B. Skaar. High-Precision Visual Control of Mobile Manipulators. IEEE Transactions on Robotics and Automation, Vol 18, No. 6, December 2002.

[11] S. B. Skaar, W. H. Brockman, and R. Hanson. Camera space manipulation. Int. Journal of Robotics Research, vol. 6, no. 4, pp. 20-32, Winter 1987.

[12] J. Shi, C. Tomasi, "Good Features to Track", IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), June 1994.

[13] D. Gennery. Least-Squares Camera Calibration Including Lens Distortion and Automatic Editing of Calibration Points. Calibration and Orientation of Cameras in Computer Vision, A. Grun and T. Huang, editors, Springer-Verlag, ISBN 3-540-65283-3, Chapter 5, pp. 123-136, 2001.

[14] A. Ansar. Survey Free Camera Calibration. JPL Technical Report NTR# 42336, 2005.

Max Bajracharya *is a member of the Mobility and Manipulation Group at the Jet Propulsion Laboratory, Pasadena, CA. His research includes vision-based tracking, manipulation, navigation, and pose estimation for Mars rovers. He is currently the task lead for several tasks focusing on autonomous target approach and instrument placement. Max received his Bachelors and Masters degrees in computer science and electrical engineering from MIT in 2001.*

Paul Backes *is a technical group leader in the Autonomy and Control section at Jet Propulsion Laboratory, Pasadena, CA, where he has been since 1987. His current activities focus on planetary manipulation, sample acquisition, and sample handling. He received the BSME degree from U.C. Berkeley in 1982, MSME in 1984 and Ph.D. in 1987 in Mechanical Engineering from Purdue University. Dr. Backes received the 1993 NASA Exceptional Engineering Achievement Medal for his contributions to space telerobotics, 1993 Space Station Award of Merit, Best Paper Award at the 1994 World Automation Congress, 1995 JPL Technology and Applications Program Exceptional Service Award, 1998 JPL Award for Excellence, 1998 NASA Software of the Year Award Sole Runner-up, and 2004 NASA Software of the Year Award. He has served as an Associate Editor of the IEEE Robotics and Automation Society Magazine.*

Matthew DiCicco *is an associate member of technical staff at the NASA Jet Propulsion Laboratory. He received a masters degree from MIT where he studied high power robotic manipulators for the Navy. At JPL he has continued his work on robotic manipulators and is currently performing research tasks for the Mars Science Lab rover mission as a member of the Mobility and Manipulation Group.*