S4-54
44776

OBJECT RECOGNITION AND POSE ESTIMATION OF PLANAR OBJECTS FROM RANGE DATA

Michael Magee
Computer Science Department
University of Wyoming
Laramie, Wyoming 82071-3682

Chiun Hong Chien
Mark L. Littlefield
Lockheed Engineering and Sciences Company
Mail Code C-19
2400 NASA Road 1
Houston, Texas 77058

Thomas W. Pendleton
NASA Johnson Space Center
Engineering Directorate
Automation and Robotics Division
Robotics Intelligence Section
Mail Code ER-4
Houston, Texas 77058

Abstract

The Extravehicular Activity Helper/Retriever (EVAHR) is a robotic device currently under development at the NASA Johnson Space Center that is designed to fetch objects or to assist in retrieving an astronaut who may have become inadvertently de-tethered. The EVAHR will be required to exhibit a high degree of intelligent autonomous operation and will base much of its reasoning upon information obtained from one or more three-dimensional sensors that it will carry and control. At the highest level of visual cognition and reasoning, the EVAHR will be required to detect objects, recognize them, and estimate their spatial orientation and location. The recognition phase and estimation of spatial pose will depend on the ability of the vision system to reliably extract geometric features of the objects such as whether the surface topologies observed are planar or curved and the spatial relationships between the component surfaces. In order to achieve these tasks, three-dimensional sensing of the operational environment and objects in the environment will therefore be essential.

One of the sensors being considered to provide image data for object recognition and pose estimation is a phase-shift laser scanner. The characteristics of the data provided by this scanner have been studied and algorithms have been developed for segmenting range images into planar surfaces, extracting basic features such as surface area, and recognizing the object based on the characteristics of extracted features. Also, an approach has been developed for estimating the spatial orientation and location of the recognized object based on orientations of extracted planes and their intersection points. This paper presents some of the algorithms that have been developed for the purpose of recognizing and estimating the pose of objects as viewed by the laser scanner, and characterizes the desirability and utility of these algorithms within the context of the scanner itself, considering data quality and noise.

1. Introduction

There has been considerable recent research devoted to the development of intelligent free-flying robots that can assist in space operations.[1,2] One such robotic device, the Extra Vehicular Activity Helper/Retriever (EVAHR), is intended to operate in relatively close proximity to a human operator, assisting with tasks such as fetching a tool, retrieving objects that may have drifted away from the primary work area, or even retrieving an astronaut who may have inadvertently become de-tethered. Early results from tests using a Manned Maneuvering Unit (MMU) propelled EVAHR on a Precision Air Bearing Floor (PABF) to simulate the frictionless environment of space demonstrated that it was possible to retrieve both large and small objects using computer vision to sense the operational environment and to employ a speech recognition system for understanding human voice commands to direct the robot's actions.[3,4] Studies are currently underway to assess the operational characteristics of the sensors and robot control mechanisms in microgravity with experiments on NASA's KC-135 aircraft.

The ability of the EVAHR to sense its operational environment is central to its functionality as an autonomous or semi-autonomous device since it must be able to recognize objects, track them, estimate their spatial poses, and estimate their motion parameters over time.[5,6,7] Because of the heterogeneous nature of these tasks, it is ultimately likely that several sensors with complementary capabilities will be employed to achieve different goals depending upon the current state of the world (the world model), the task to be achieved, and the characteristics of the sensors themselves.[8] For example, images from a color camera are useful for identifying objects based on their visible spectral characteristics but it difficult to estimate pose from two-dimensional images. Conversely, a laser scanner can provide three-dimensional coordinates for points on a scanned object, but no color information is available. The remainder of this paper focuses on processing actual image data from a laser scanner, and documents a method for segmenting objects into their primary planar regions, recognizing them, and estimating their spatial poses.

2. Laser Scanner Characteristics

The sensor employed for the studies whose descriptions follow is a laser range scanner that measures distances based on the phase shift of a modulated signal carried on an infrared laser beam. The range values returned by the scanner are represented by 12 bit integers that span a single ambiguity interval of approximately 15.2 meters. This means that a difference of one range unit (out of 4096) represents a distance change of about 4 mm. The scanner is able to produce a dense range image by employing a rotating mirror whose rotation axis can be tilted. The scanner simultaneously provides two separate range and reflectance (intensity) images that are fully registered.

The quality of the range data provided by the scanner is affected by several factors which generally relate to the composition of the surface material, its reflectivity characteristics, its geometry, and the orientation of surface normals relative to the scanner itself. The most influential among these factors is the reflectivity of the surface material. For extreme cases in which a scanned region is composed of a highly specularly reflective material, reliable range estimates are not expected since the laser beam will be reflected away from the sensor.

For less extreme cases involving diffuse reflective surfaces, however, the quality of the data is highly dependent on the albedo of the surface. These dependencies can best be illustrated by examining the quality of the range images acquired by scanning black and white planar surfaces (sheets of paper) that were oriented perpendicular to the optical axis of the scanner. As a measure of data stability, the local standard deviation (sigma) for range values was computed within a row. This local standard deviation was based on the center range value and the nearest 8 neighbors within the row. It was observed that the local sigma varied by as much as 3 range units. For such cases, in excess of 99% of the range samples could be expected to fall within 3 sigma (± 9 range units) of the mean value. For the test case under discussion, this translates into a local variation of approximately ±33 mm over a distance of 8 mm. For the black surface, the quality of the data was significantly worse. Local standard deviations as high as 9 range values were observed meaning that a 3 sigma test would include range values as far as ± 100 mm over this limited region of a scan line. The local standard deviations for reflectances varied up to 30 units for the white surface and up to 8 units for the black surface.

The implications of these observed local variations are very important when designing algorithms that attempt to segment the image into component regions such as planes and curved surfaces. For example, the magnitude of the local variations in range values makes it extremely difficult to segment planar surfaces based on a local geometric constraint such as surface normal consistency. Furthermore, even on white objects, it is difficult to recognize the curvature of objects smaller than 100 mm since the magnitude of local range variation is large relative to surface size. If the data is smoothed by a classical filtering mechanism, finer details that are necessary to recognize an object and/or estimate its pose may be lost. Hence, algorithms that depend on local geometry are less likely to succeed than those that take a more global approach to object analysis. The results of both local and global algorithms that were developed are presented in the next section.

3. Finding Planes, Recognizing Objects and Estimating their Spatial Poses

The local instability of range values observed for the laser scanner makes scene segmentation using locally computed surface normals exceptionally difficult unless the range values are smoothed using a reasonably large filter. Applying such a filter, of course, results in a loss of scene detail but does make it possible to find planes that are large relative to the size of the filter.

An approach that was found to be both more computationally efficient and robust was to grow surfaces based on local range and reflectance difference constraints. It was determined that after applying a 7X7 mean filter, planes that were not highly oblique to the sensor axis could be successfully grown by adding to regions neighboring image elements whose smoothed reflectance and range values did not differ by more than 40 and 1.5, respectively. This provided the basis by which planar regions could be segmented and the segmented planes used for object recognition and pose estimation. Figure 1 shows one object, a simulated Orbital Replacement Unit (ORU), to which the plane segmentation algorithm was applied. This ORU consists of a rectangular solid to which an H-shaped handle is attached by an intermediate short cylindrical section. When viewed by the laser scanner and rendered as a solid model, the ORU appears as in Figure 2. With respect to the observed noise characteristics that had to be dealt with algorithmically, Figures 3 and 4 are more revealing, however.

Figure 3 shows a wireframe rendering of the scanned ORU with the bright line profile across the main body and H-shaped handle being isolated in Figure 4. It should be noted that the raw data across the major left surface should be linear, but is extremely "busy". It is this effect that makes the segmentation of planes using surface normals difficult since inconsistent directions based on local patches are computed unless large smoothing filters are applied. On the other hand, using a region growing approach based on propagating the local constraints of reflectance and range similarities, it is possible to successfully segment the scene into its planar regions as shown in Figure 5. It should be noted, however, that these successfully segmented planar regions are somewhat deceptive since when viewed from the perspective of the laser scanner the true variation of the original data is not evident. Figure 6 shows the same segmentated image data but from a different viewpoint. It should be noted that there are several areas of high variation. In particular, the H-shaped handle has range values that vary by as much as the width of the handle's vertical substructures. Hence, the level of noise is relatively large compared to the feature itself.

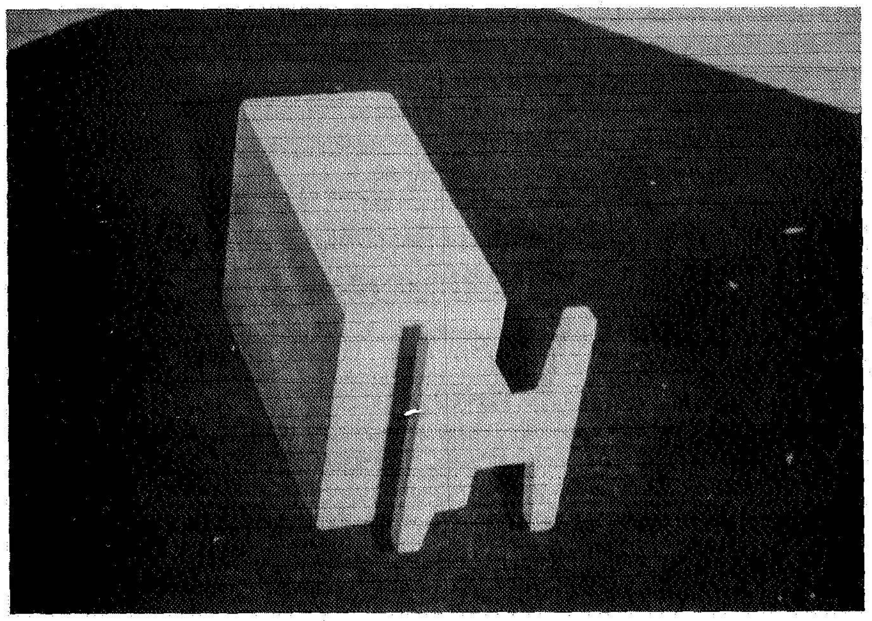

Figure 1: simulated orbital replacement unit (ORU)

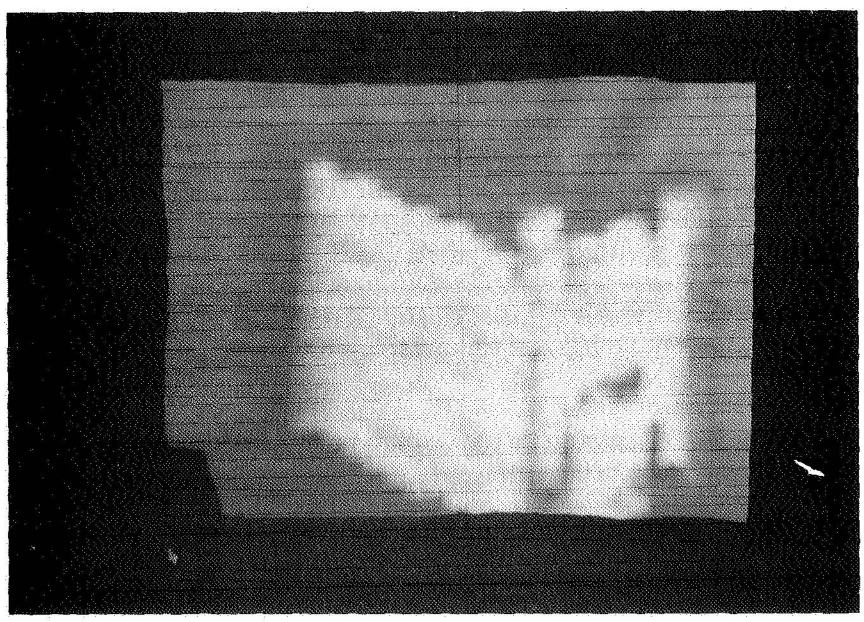

Figure 2: ORU as a shaded model graphically reconstructed from range data

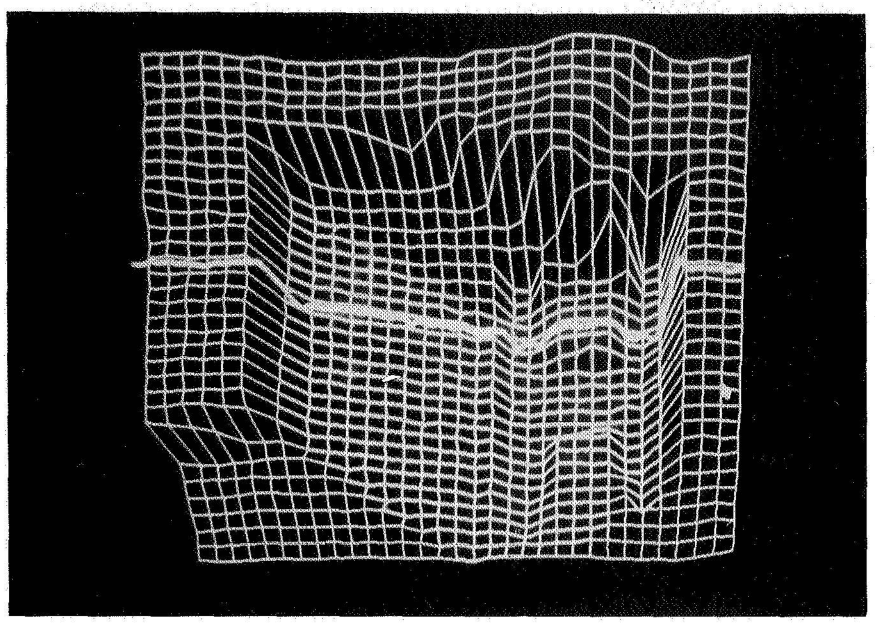

Figure 3: ORU as a wireframe model graphically reconstructed from range data

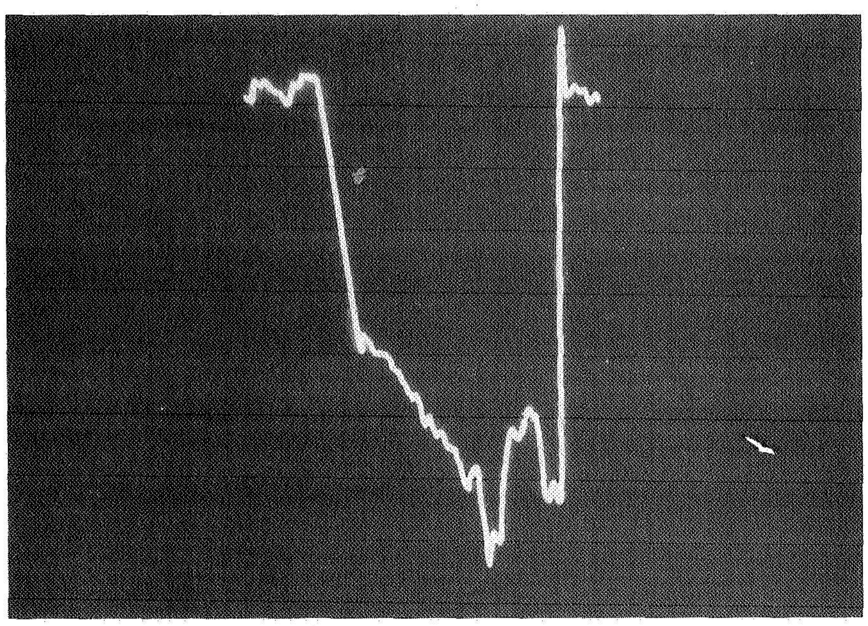

Figure 4: a single line (profile) of laser range data across the ORU

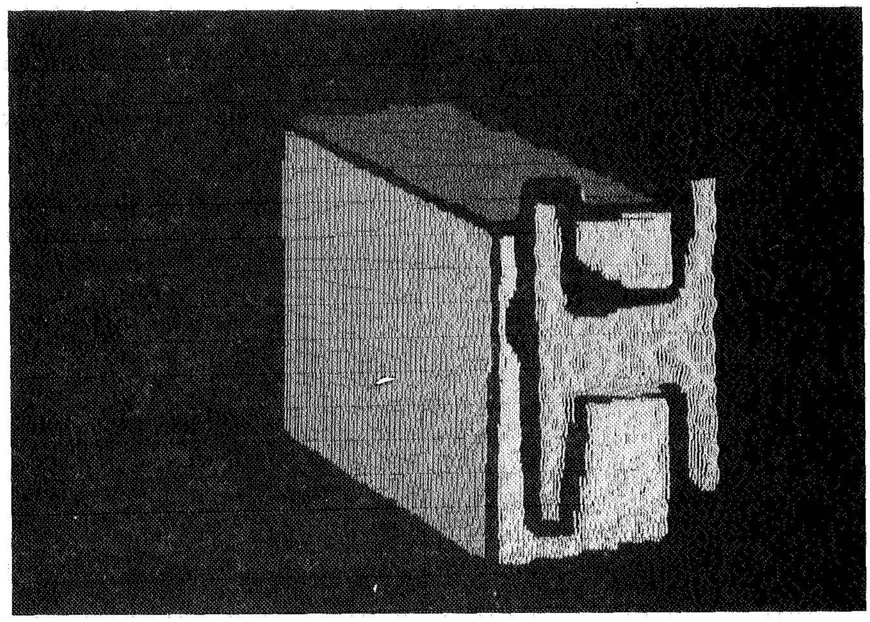

Figure 5: laser range data segmented into planar regions

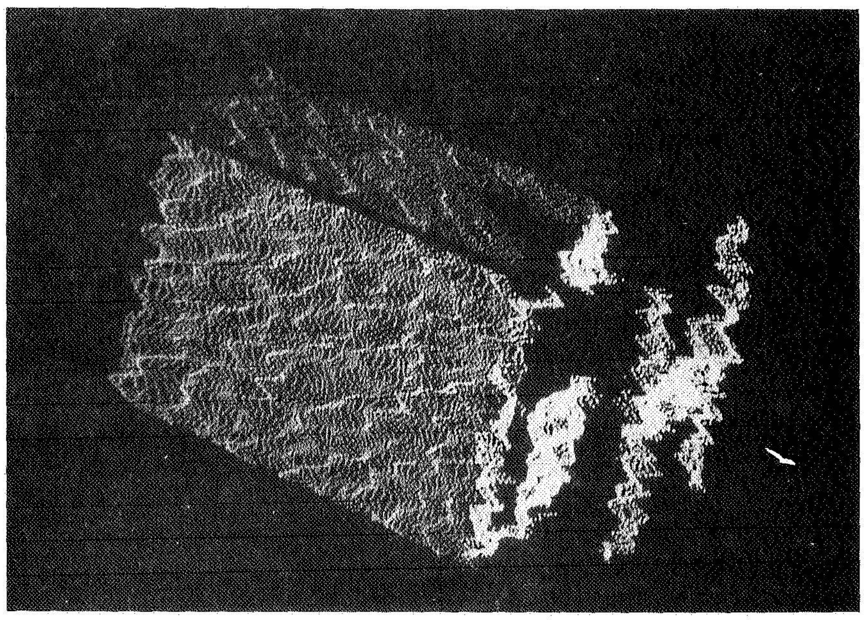

Figure 6: variations in laser range data from planar regions

The method by which the primary planar features of the ORU were recognized was based on the areas of the observed planes. Since the original sensor data as shown in Figure 6 has extreme variations in the form of hills and valleys, however, incorrect areas for the planar features would be computed unless the data were forced to conform to the best plane equation that fits all of the range points belonging to a segmented feature. This was achieved by computing the plane equation using a least squares fit of all the points in each segmented planar feature and backprojecting each point in the segmented planar feature onto the computed plane. Figure 7 shows the points in the adjusted three-dimensional range image that results when this process is applied to the data in Figure 6.

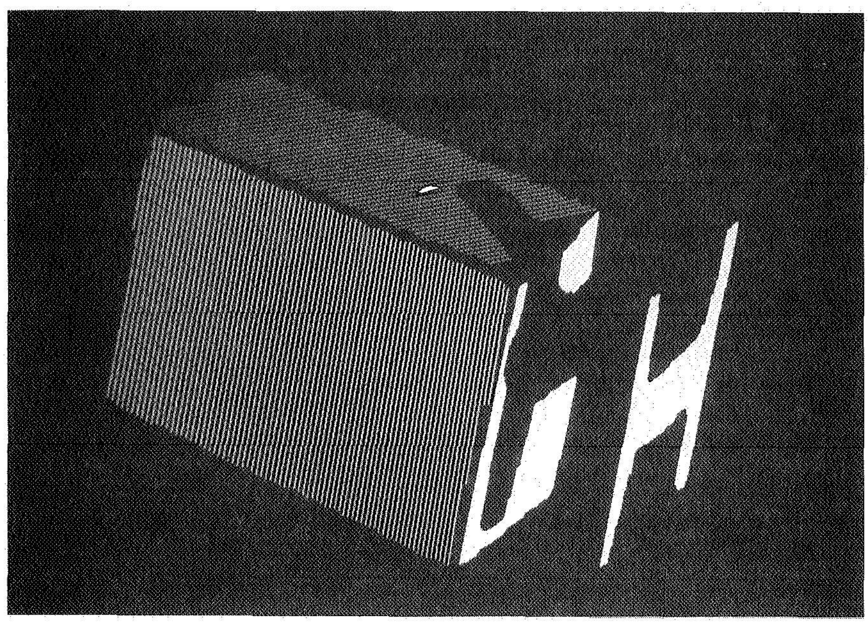

Figure 7: laser range data after conformal mapping to extracted planes

After the planar conformal mapping of the original data has been achieved, the area of each feature is computed and compared against the areas of planar features in the model base, and correspondences between observed and model features are established. Since, this feature matching method is based on computed surface areas, it is necessarily sensitive to occlusion. However, once surfaces have been grown, it is possible to compute other features that would be useful for recognition such as the vertices and line segments that result from the intersections of planes. Four or more features are sufficient to provide the basis for feature matching and pose estimation.

For the current study, pose is estimated by orienting the model such that three of its surface normals match the orientations of the corresponding planes in the observed data and such that the intersection point of these three planes is translated to be consistent with the analogous observed intersection point. The wireframe overlay in Figure 8 demonstrates that the proper spatial pose for the ORU model is computed such that its features correspond to those in the original range image data.

4. Conclusions

A method has been presented for segmenting planar regions from laser range and reflectance data which is useful for recognizing objects and estimating their spatial poses. The method, which is based on local constraint propagation, permits successful planar segmentation even in the presence of significant noise, but postprocessing of the three-dimensional data in the segmented regions is required to accurately characterize and use the planar regions.

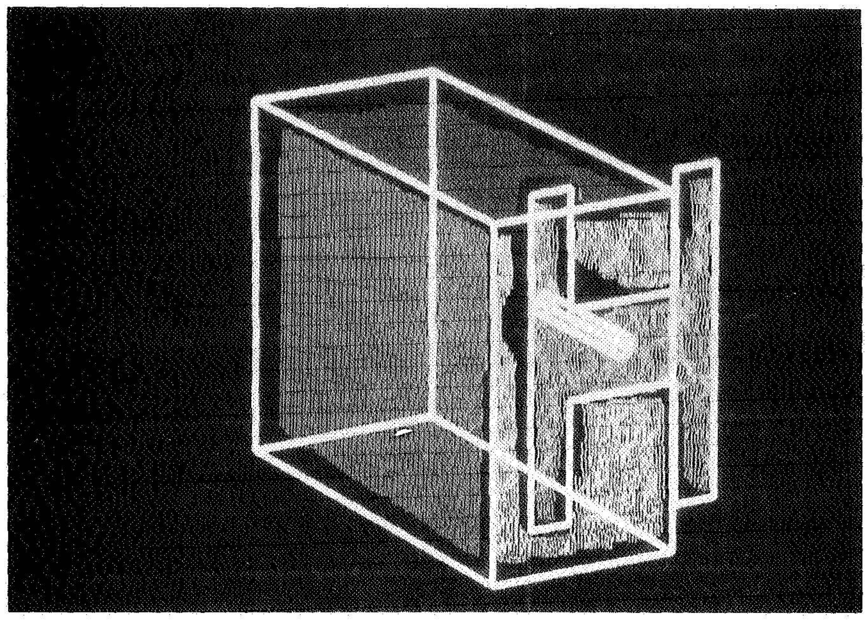

Figure 8: overlay showing correctly estimated pose for ORU model

5. Acknowledgements

The support of the NASA Johnson Space Center Summer Faculty Fellowship Program and the University of Wyoming Planetary and Space Science Center is gratefully acknowledged.

6. References

1. G.J. Reuter, C.W. Hess, D.E. Rhoades, L.W. McFadin, K.J. Healey, and J.D. Erickson, "An Intelligent Free-Flying Robot", *SPIE Symposium on Advances in Intelligent Robotic Systems, Space Station Automation IV*, SPIE 1006-03, Cambridge, Massachusetts, Nov. 6-11, 1988.
2. J.D. Erickson, C.R. Price, D. Cooke, "Future Needs for Space Robots for SEI", *Proceedings of the SPIE Symposium on Cooperative Intelligent Robotics Systems in Space II*, Boston, Massachusetts, November 10-14, 1991.
3. J. Erickson, R. Goode, K. Grimm, C. Hess, R. Norsworthy, G. Anderson, L. Merkel, and D. Phinney, "Technology Test Results from an Intelligent Free-Flying Robot for Crew and Equipment Retrieval in Space", *SPIE Symposium on Cooperative Intelligent Robotic Systems in Space II*, Boston, MA, November 10-14, 1991.
4. K.A. Grimm and J.D. Erickson, "An Experiment in Vision Based Autonomous Grasping within a Reduced Gravity Environment", *Spie Cooperative Intelligent Robotics in Space III*, Cambridge, Massachussetts, November, 1992.
5. C.H. Chien, "Multi-view Based Pose Estimation from Range Images", *Spie Cooperative Intelligent Robotics in Space III*, Cambridge, Massachussetts, November, 1992.
6. M.L. Littlefield, "Adaptive Tracking of Objects for a Mobile Robot Using Range Images", *Spie Cooperative Intelligent Robotics in Space III*, Cambridge, Massachussetts, November, 1992.
7. L. Hewgill, "Motion Estimation of a Freely Rotating Body in Earth Orbit", *Spie Cooperative Intelligent Robotics in Space III*, Cambridge, Massachussetts, November, 1992.
8. M. Magee, C.H. Chien, and T.W. Pendleton, "A Vision System Planner for the Extravehicular Activity Retriever", *Proceedings of the International Conference on Intelligent Autonomous Systems*, Pittsburgh, Pennsylvania, February, 1993, pp. 696-705.