ISSUES IN BUILDING INTELLIGENT GRASPING SYSTEMS

Peter K. Allen

Department of Computer Science
Columbia University
New York, NY 10027

ABSTRACT

Intelligent robotic systems will need to have the capability to grasp, recognize, inspect and manipulate objects. Our experience in building an intelligent system with these capabilities has forced us to confront a number of important issues that we have found to be important in building an intelligent grasping system. The issues are 1) the need for multiple sensing capability in hands, 2) synchronization of actuation devices, and 3) coordinated contact movement on a surface. We demonstrate how these issues are overcome with an example of haptic sensing with our Utah/MIT hand system.

1. INTRODUCTION

The focus of our work is in creating a multi-fingered robotic hand system that can be used for manipulation, grasping, inspection and object recognition. This is a formidable task that involves many open research issues in robotics. In particular, such a system requires attention to the following areas:

- Research in mechanical design of complex linkages with many degrees-of-freedom.
- Research in actuation devices that provide large forces with low mass and size.
- Research in sensors that provide accurate high bandwidth force and position information.
- Research in developing tactile capabilities for hands.
- Research in real-time control of multiple degree-of-freedom devices.
- Research in programming environments that allow a robotic hand to be used as an integral part of a robotic system
- Research in task-level specification of grasping and manipulation
- Research in active sensing strategies with an arm mounted hand.

Despite this seemingly formidable research agenda, great strides have been made in developing multi-fingered robotic hands such as the Salisbury hand [15] the Utah-MIT hand [13], and the Belgrade hand [23]. These hands provide varying levels of low-level control and sensing to allow using them at the task level. Our approach is to look at robotic grasping and manipulation from a task level perspective, moving the focus away from low level control issues toward the higher levels of control. This paper's focus is on a number of issues that we have found to be quite important in building an intelligent grasping system. Sections 2-4 outline these issues, section 5 describes our experimental system, and section 6 gives an example in haptic shape recovery that discusses our particular solution to these problems.

2. ISSUE 1: THE NEED FOR MULTIPLE SENSING CAPABILITY IN HANDS

In order to build hands for dexterous manipulation, progress must be made with respect to the hand's capability to perceive object features, such as weight, texture, hardness, and shape. If we were to model the haptic systems of robotic hands on the human system, we would see that the system would need to have the following classes of sensors: (a) tactile sensors that form a skin-like surface, (b) joint angle sensors to judge the relative positions of the different links (bones), and (c) muscle or tendon force sensors. Sensing joint angles and tendon tensions is possible with current technology, and several dexterous hands have incorporated these sensors successfully [13, 15]. The design and use of tactile sensing has proved more difficult, however. The ideal tactile sensing system for a haptic system -- one that mimics the human system -- would be capable of detecting position, orientation, velocity, forces and temperatures of contacts [6, 7, 12]. Many types of tactile sensor have been designed in recent years, but none has all of the properties required for a rich haptic system (see Nicholls and Lee [17] for an excellent survey of this area). Most of the sensors are capable of yielding only information about the locality and magnitude of contacts normal to point of contact, even though their sensitivity to forces and spatial resolution are good [9, 20, 22]. Howe and Cutkosky [11] have developed a sensor that can detect acceleration -- which is important if a hand is to sense slip and texture. But no single sensor or sensor system can extract all of the features that a human haptic system can. Furthermore, one of the most important features required of a tactile sensing system is that it provide complete coverage of the area being sensed and that it conform to the hand's shape. Bicchi, Salisbury, and Dario [3] have experimented with an idea they call "intrinsic tactile sensing," in which force sensors located in the finger joints can be used to calculate the direction and magnitude of an applied force. Likewise, thermally sensitive skins are only beginning to be developed [19]. Most tactile sensors are in the development stage and have not yet been tested in a working haptic system.

A haptic system requires a convergence of many sensory inputs to form inferences about object properties. When a hand makes contact with an object, there are many cues that an intelligent planner can use and must be looking for. For example, suppose that a single finger is moving toward a hard surface, but the height of the surface is unknown. When the finger touches the surface, several changes in the system state occur. If the finger has tactile sensors at the point of contact, these sensors will provide the exact location of the contact region. Even without tactile contact, it is possible to realize that the finger is touching the surface by means of other sensors. By interrogating the joint angle sensors on the finger, the controller detects that further movement commands produce no effect. Therefore, an obstacle must be hindering motion and contact can be hypothesized. Similarly, if the finger's motion is controlled by flexor and extensor tendons, and the forces on these tendons are monitored, a sudden change in the values of these tendon forces is a certain indication of contact. Another technique would be to use the intrinsic tactile sensors mentioned above. The human haptic system uses all of these cues, as well as expectations about the environment, to formulate hypotheses about object features, and it will be important for dexterous robot hands to fuse the available sensory inputs. In addition, the fusion of many types of sensory input makes the sensing more robust and reliable.

Related to this multiple sensing capability is the need for a number of different coordinate frames, each related to the particular sensing modality. In a pick-and-place type task, we need to reason about the Cartesian position of the fingers, the localization of tactile contact on the fingertip array for stable grasping in the center of the array, and finally, the differential in tendon space tensions to properly monitor the forces transmitted to the joints on the finger. Each of these frames of reference is needed and used to perform the overall task of grasping the object with the fingers.

3. ISSUE 2: SYNCHRONIZATION OF ACTUATION

Most complex robotic systems are really amalgams of separate processing devices, connected by a number of different methods: The simplest and most reliable way to connect these systems (in our case this is the robotic arm system, the Utah-MIT hand system, and the tactile sensor system) is via a bus level connection. This allows maximum bandwidth via shared memory, allowing the systems to be thought of as fully interconnected. However, many applications cannot support this due to hardware problems (incompatible bus protocols), distance limitations on system components, or desire to use existing, incompatible systems. Non-bus level integration of these sensors necessitates synchronization of the systems. However, it is usually the case that each subsystem can operate semi-autonomously, with low bandwidth communication for synchronization. As an example, our robotic arm is controlled via a serial interface line. In order to coordinate the movement of our hand and arm together, we need to build delays into our system equal to the communication delay. As we approach human speed of movement for grasping tasks, this becomes more difficult to accurately predict. A better solution (which we are currently implementing) is to use a high speed network interface that can support hand/arm synchronization at approximately 50 HZ. This rate appears to be adequate for the grasping task we are interested in, is easily sustainable over the network, and matches well with our vision systems processing speed (60 HZ frame rate processor) to allow vision to be tightly interconnected to the hand/arm system. By keeping each system logically separate, different algorithms can be tested easily, and integration can be done at ever increasing levels of coupling and communication bandwidth.

4. ISSUE 3: COORDINATED CONTACT MOVEMENT ON A SURFACE

The problem of moving a hand along an arbitrary surface is a difficult one. It necessitates a hybrid force/position control scheme [18]. The forces can be measured by the tendons on the hand, but they will not localize contacts, which may be critical to the particular task at hand (no pun intended!). Tactile feedback may be used in such a hybrid scheme to better localize the contact forces and positions. In addition to the sensor feedback, the movement needs a high level strategy or direction of movement. This vector then can be communicated to the actuation system to effect the movement (postionally) while maintaining the force/contact constraints imposed by surface following.

Our previous work [1] using a one-fingered tactile sensor mounted on a PUMA traced along a curved surface by calculating a weighted vector of constraint directions that tried to follow the surface curvature while preserving smoothness of the trace and a constraint having to do with creating regions bounded by traces that were equivalent in size. Hor [10] traced contours of planar objects using a planar four-fingered "chopstick" like manipulator. Strain gauge sensors on the fingers of this device would calculate surface normals and move tangentially along a surface, recording the contour. Stansfield [21] used a planar LORD tactile sensor mounted on a PUMA to trace edges and other features on objects, again using local geometry to effect movement vectors.

Our particular interest in this is motivated by recovering surface shape from touch, particularly from curved surface objects. In the example in section 6 we will describe our method of following a curved surface with our multi-fingered Utah/MIT dextrous hand.

5. SYSTEM OVERVIEW

The system we have built consists of a Utah-MIT hand attached to a PUMA 560 manipulator, and is described in detail in [2]. The hand contains four fingers, each with four degrees of freedom. It resembles the human hand in size and shape, but lacks a number of features that humans find very useful. In particular, it has no palmar degree of freedom (closing of the palm) and the thumb is placed directly opposite the other three fingers, with all fingers identical in size (see Figure 1). The hand has joint position sensors that yield joint angle data and tendon force sensors that measure forces on each of the two tendons (extensor and flexor) that control a joint. The PUMA adds 6 degrees of freedom to the system (3 translation parameters to move the hand in space and 3 rotational parameters to orient the hand), yielding a 22 degree of freedom system. Clearly, such a system is a nightmare to control at the servo-level in real-time. Our approach is to use the embedded controllers in each of these systems, controlling and communicating with them through an intelligent, high-level controller that links together the movements of arm, hand and fingers with the feedback sensing of joint positions, tendon forces, and tactile responses on the fingers.

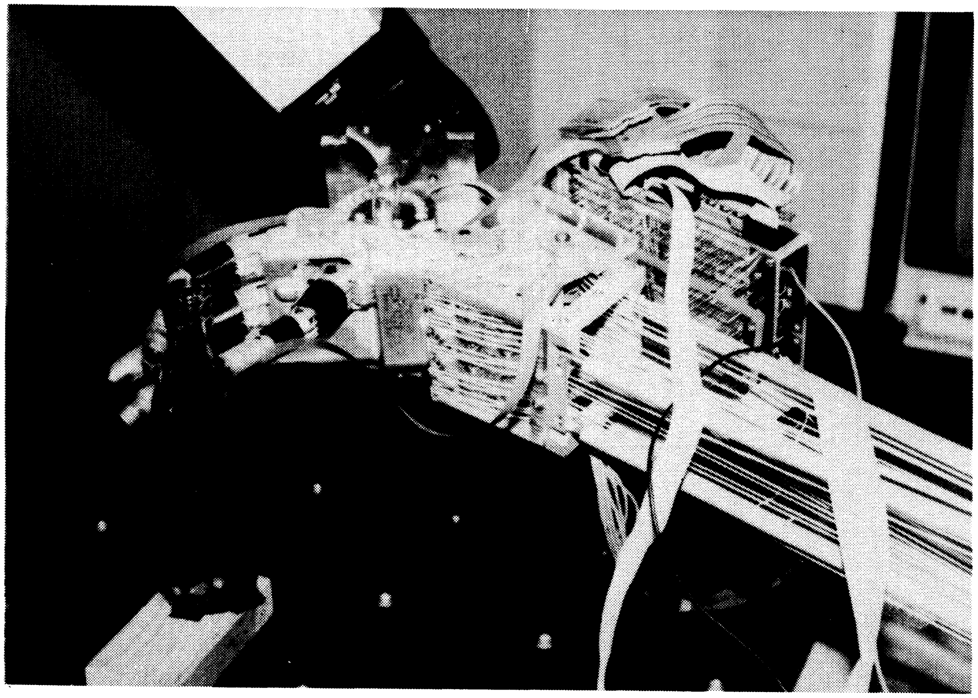

Figure 1: Hand/Arm system.

The hardware structure of the system is shown in Figure 2. The high-level control resides in a SUN-3 processor. The SUN serves as the central controller, and has access to a full UNIX based system for program development and debugging as well as a set of window-based utilities to allow graphical output and display of the system's various states. The hand is controlled by an analog controller that is commanded through D/A boards from a dedicated 68020 system. The SUN is capable of downloading and executing code on the 68020 and can communicate with it through a shared memory interface [16]. The tactile sensing system is controlled by another dedicated 68020 that monitors the forces on each of the sensor pads. The connection from the SUN to the PUMA is via the VAL-II host control option over a serial interface; however we have just changed this interface to allow real-time control of the arm via RCCL [8].

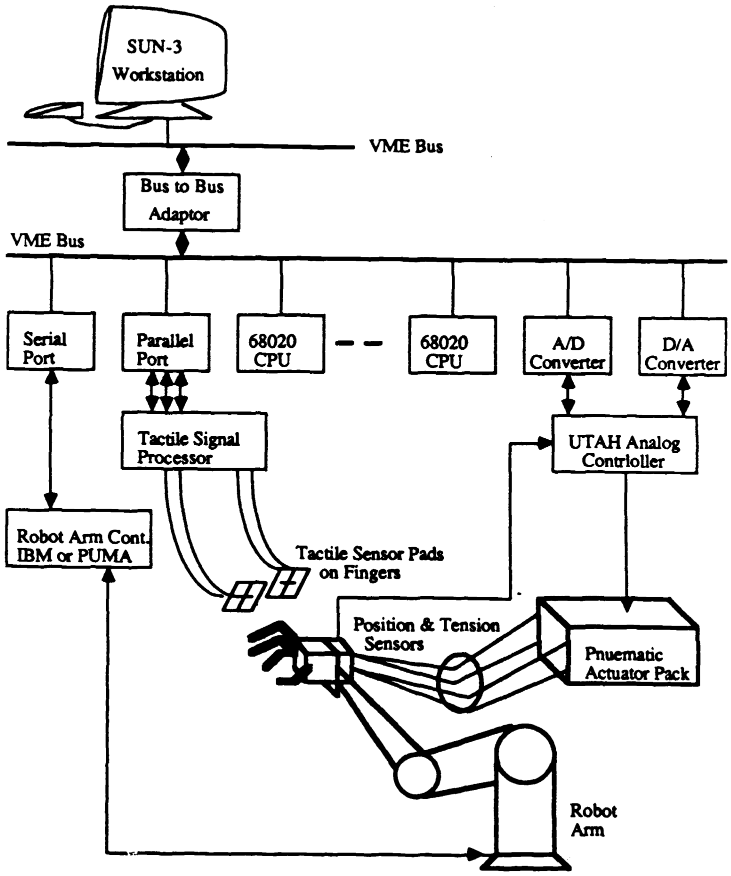

Figure 2: Hardware Overview.

6. EXAMPLE: CONTOUR FOLLOWING WITH DEXTROUS HAND

Our interest in human haptic capabilities has led us to implement a contour follower Exploratory Procedure (EP). The Contour Follower EP in humans is described by Klatzky and Lederman as " a dynamic procedure in which the hand maintains contact with a contour of the object. Typically, the movement is smooth and non-repetitive within a segment of object contour, stops or shifts direction when a contour segment ends, and does not occur on a homogeneous surface [14]." It seems natural that this EP will report information that can be used to recover a shape that can be represented as a class of generalized cylinders. Generalized cylinders have been proposed by many researchers beginning with Binford [4] as a shape modeling primitive. Other researchers have expanded on this idea of a swept volume by creating classes of generalized cylinders or cones, depending upon the nature of the axis curve, sweeping rule and cross sectional curve. These primitives have special appeal in the recognition of elongated objects and objects that provide strong visual contours.

Fearing [5] has attempted to recover the shape of a class of these generalized cones (RLSHGC - Right Linear Straight Homogeneous Generalized Cones) using extremely sparse amounts of data. He has characterized the necessary and sufficient conditions for being able to recover the axis and orientation of these cones, given limited, multi-fingered tactile sensor data that includes point contacts, surface normal direction, and surface curvature information.

We have also chosen to use a class of these primitives for shape recovery. The class we are using is surfaces of revolution, which are RSHGC with a circular cross section function (no linear scaling of the cross section is required). These surfaces may be completely described by the rotation of a plane curve about the axis of symmetry. Our approach is to use the Contour Follower to recover the contour curve described above. If we obtain two such contour curves that are on either side of the object, we can estimate the axis of the surface of revolution, and recover the shape. This procedure maps naturally into a two-fingered Contour Follower EP, in which an object's contour on either side are sensed using the thumb and index finger.

The problem of using an active tactile device to trace a surface on an object is a complicated one. Our previous work [1] using a one-fingered tactile sensor mounted on a PUMA traced along a curved surface by calculating a weighted vector of constraint directions that tried to follow the surface curvature while preserving smoothness of the trace and a constraint having to do with creating regions bounded by traces that were equivalent in size. Hor [10] traced contours of planar objects using a planar four-fingered "chopstick" like manipulator. Strain gauge sensors on the fingers of this device would calculate surface normals and move tangentially along a surface, recording the contour. Stansfield [21] used a planar LORD tactile sensor mounted on a PUMA to trace edges and other features on objects.

Our Contour Follower EP is now described. First, the PUMA is moved to a location near one end of the explored object, and the thumb and index finger are opened enough to allow them to encompass the object without making contact with it. Then the thumb is slowly moved toward the object until the sensors detect contact between the thumb and the object. Next, the index finger follows the same movement. After detecting contact, the positions of the two contact locations are noted, and the fingers are backed off the object so that they are no longer in contact. The arm and hand are moved a small amount along the axis of the explored object, and the process is repeated. This exploratory procedure ends when one of the fingers moves toward the object and fails to make contact. (The location of the object and its axis are not currently determined autonomously, but with human aid.)

The detection of contact and conversion to Cartesian coordinates is a process that requires several steps. The fingers are moved toward the object in a number of discrete intervals. After each movement, two checks are performed. First, did the tactile sensor detect contact? And second, did the finger move the entire distance that was commanded? If the tactile sensor detects contact, then the location of the center of the contact region is found. To find the center of the contact, the first moments of the tactile array are taken. Then a transformation is performed from the fingertip

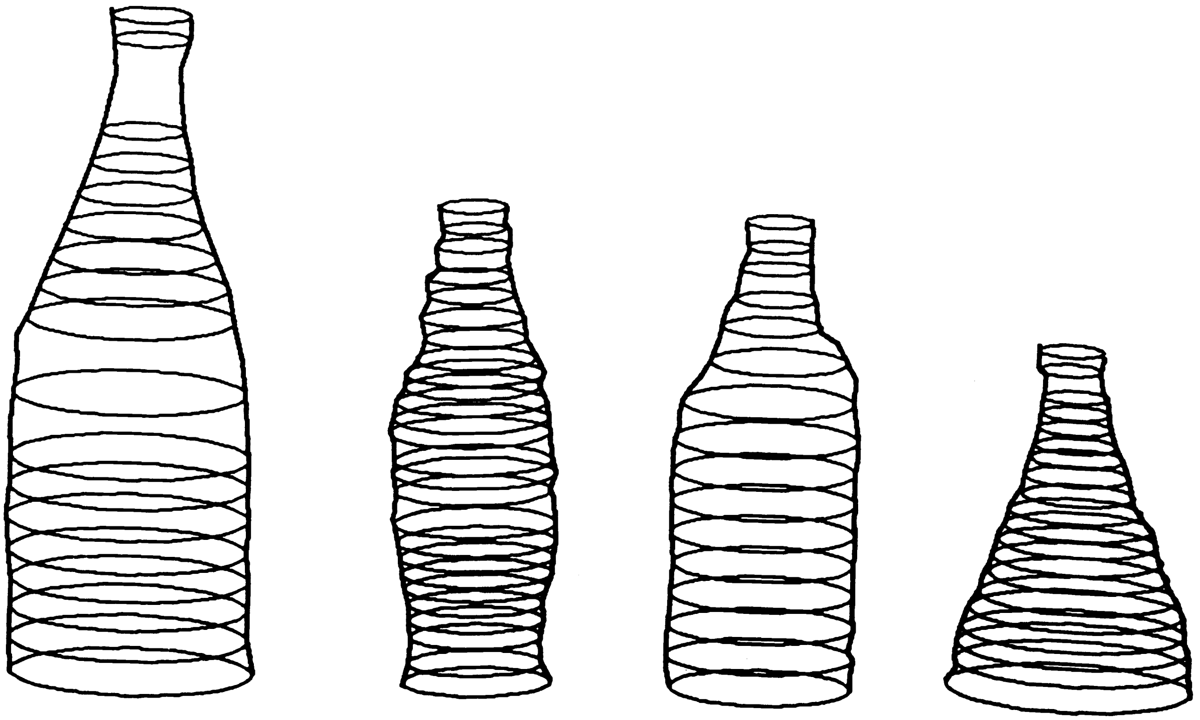

Figure 3: Linked contour points and recovered surfaces of revolution from Contour Explorer EP (left to right wine bottle, coke bottle, beer bottle, Orangina bottle).

coordinate frame to the hand coordinate frame, and finally, from the hand coordinate frame to world coordinates. The second check is to see if the finger has moved the entire distance commanded (and there is no tactile contact). This event would signal that something is impeding a finger from moving. In this case, no centroid of the contact region is found and the data point is thrown out. After detecting contact that does not involve the tactile sensor, the exploratory procedure continues looking for valid contact points along the original search axis.

We have performed a series of experiments that try to recover the shape of a number of different surfaces of revolution including a wine bottle, a beer bottle, a coke bottle and an Orangina soft drink bottle (a flask like object). The procedure begins with exploring the object along an exploration axis that is assumed to be perpendicular to the support table (but can be inferred from vision sensing described below). The points generated from these contour traces are then linked into a set of linear contour segments. Circular cross section curves are then fit perpendicular to the exploration axis that include trace points from each of the contours. The recovered shapes are shown in Figure 3. The shapes are clearly distinguishable from this sparse data. An additional and important discriminating characteristic is actual 3-D size and volume which are calculable from these representations.

In this example, we can see the importance of the three issues discussed in this paper. First, we need to have an integrated sensing capability on the hand in order to ascertain contacts from both contact on the tactile sensor itself, and contacts not on the sensor but in which the fingers are prevented from moving further. Secondly, we need to synchronize the actuation of the fingers making contact with the arm which moves the hand to a new sensing position. These motions can be overlapped using our high level control scheme describe in [2]. Finally, the coordinated movement along the surface of the contour is needed to accurately recover the shape of the surface or revolution.

ACKNOWLEDGEMENTS

This work was supported in part by DARPA contract N00039-84-C-0165, NSF grants DMC-86-05065, DCI-86-08845, CCR-86-12709, IRI-86-57151, North American Philips Laboratories, Siemens Corporation and Rockwell Inc. Special thanks to Paul Michelman and Ken Roberts without whom there would be no experimental system to explore these issues.

References

1. Allen, Peter, "Integrating vision and touch for object recognition tasks," *International Journal of Robotics Research*, vol. 7, no. 6, pp. 15-32, 1988.
2. Allen, Peter, Paul Michelman, and Kenneth S. Roberts, "An integrated system for dextrous manipulation," *IEEE Conference on Robotics and Automation*, pp. 612-617, Scottsdale, AZ, May 15-19, 1989.
3. Bicchi, A., J. K. Salisbury, and P. Dario, "Augmentation of grasp robustness using intrinsic tactile sensing," *IEEE Conference on Robotics and Automation*, pp. 302-307, Scottsdale, AZ, May 15-19, 1989.
4. Binford, T., "Visual perception by computer," *Proc. IEEE Conference on Systems and Control*, Miami, 1971.
5. Fearing, Ronald, "Tactile sensing for dextrous manipulation," *Proc. IEEE Workshop on Dextrous Robotic Hands*, Philadelphia, April 1988.
6. Grupen, Roderic and Thomas Henderson, "A survey of dextrous manipulation," Technical Report, Department of Computer Science, University of Utah, 1987.
7. Harmon, Leon, "Automated tactile sensing," *Int. Journal of Robotics Research*, vol. 1, no. 2, pp. 3-32, Summer 1982.
8. Hayward, Vincent and Richard Paul, "Robot manipulator control under UNIX," *Proc. of the 13th ISIR*, pp. 20:32-20:44, Chicago, April 17-21, 1983.
9. Hillis, W. D., "A high resolution imaging touch sensor," *Int. Journal of Robotics Research*, vol. 1, no. 2, pp. 33-44, Summer 1982.
10. Hor, Maw-Kae, "Control and task planning for a four finger dextrous manipulator," Ph.D Thesis, Courant Institute, New York University, October 1987.

11. Howe, Robert, "Mark Cutkosky," *IEEE Conference on Robotics and Automation*, pp. 145-150, Scottsdale, AZ, May 15-19, 1989.
12. Iggo, A., "Cutaneous sensory mechanisms," in *The senses*, ed. J. D. Mollon, Cambridge University Press, 1982.
13. Jacobsen, S. C., E. K. Iversen, D. F. Knutti, R. T. Johnson, and K. B. Biggers, "Design of the Utah/MIT dextrous hand," *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 1520-1532, San Francisco, April 7-10, 1986.
14. Klatzky, Roberta and Susan Lederman, "The intelligent hand," in *The psychology of learning and motivation*, ed. G. Bower, vol. 21, Academic Press.
15. Mason, Matthew T. and J. Kenneth Salisbury, *Robot hands and the mechanics of manipulation*, M. I. T. Press, Cambridge, 1985.
16. Narasimhan, Sundar, David M. Siegel, and John M. Hollerbach, "Condor: A revised architecture for controlling the Utah-MIT hand," *IEEE Conference on Robotics and Automation*, pp. 446-449, Philadelphia, April 24-29, 1988.
17. Nicholls, H. and M. Lee, "A survey of robot tactile sensing technology," *International Journal of Robotics Research*, vol. 8, no. 3, pp. 3-30, June 1989.
18. Raibert, Marc and John Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 102, June 1981.
19. Shrader, E., R. Kornbluh, and G. B. Andeen, "A thermal slip detector," Technical Report, Engineering Sciences Laboratory, SRI International, June 1986.
20. Speeter, Thomas, "Flexible piezo-resistive touch sensing array," *SPIE Conference on Optics, Illumination and Image Sensing for Machine Vision III*, Cambridge, November 1988.
21. Stansfield, Sharon, "Visually-guided haptic object recognition," Ph.D. dissertation, Department of Computer and Information Science, University of Pennsylvania, October 1987.
22. Tise, B., "A compact high resolution piezo-resistive digital tactile sensor," *IEEE Conference on Robotics and Automation*, pp. 760-764, Philadelphia, April 24-29, 1988.
23. Tomovic, Rajko, George Bekey, and Walter Karplus, "A strategy for grasp synthesis with multi-fingered robot hands," *Proc. IEEE International Conference on Robotics And Automation*, pp. 83-89, Raleigh, N. C., March 31-April 3 1987.