Using Tactile and Visual Sensing with a Robotic Hand

Peter K. Allen Andrew T. Miller Paul Y. Oh Brian S. Leibowitz
Department of Computer Science, Columbia University, New York, NY 10027 *

Abstract

Most robotic hands are either sensorless or lack the ability to accurately and robustly report position and force information relating to contact. This paper describes a robotic hand system that uses a limited set of native joint position and force sensing along with custom designed tactile sensors and real-time vision modules to accurately compute finger contacts and applied forces for grasping tasks. Three experiments are described: integration of real-time visual trackers in conjunction with internal strain gauge sensing to correctly localize and compute finger forces, determination of contact points on the inner and outer links of a finger through tactile sensing and visual sensing, and determination of vertical displacement by tactile sensing for a grasping task.

1 Introduction

Robotic dextrous manipulation remains a difficult problem with many open research problems. While there have been a number of detailed analyses of the kinematic and dynamic constraints necessary to effect stable grasps, most require a high level of sensory input and feedback from the grasping device (i.e. robotic hand) to perform dextrous manipulation. The sensory information required typically includes contact point estimation, surface normal and curvature measures, and knowledge of both applied and induced forces on the fingers of the hand. While great strides have been made in robotic hand design and a number of working dextrous robotic hands built, the reality is that the sensory information required for dextrous manipulation lags the mechanical capability of the hands. Accurate and high bandwidth force and position information for a multiple finger hand is still difficult to acquire robustly.

This paper describes the design and use of a set of additional sensors that augment the capability of a robotic hand. The hand we are using is the Barrett Hand, which is a three-fingered, four DOF hand with limited sensing capability. We have added two different sensory systems to the hand, one internal and the other external. The first is a set of tactile sensors covering the links of the hand as well as the palmar surface, and the second is a set of real-time vision modules that can be used to track and monitor the hand as it performs a task. Our aim is to use this set of internal and external sensors to accurately and robustly estimate forces and contacts on the hand. The tactile sensor system can be used to localize contacts on the surfaces of the hand, as well as determine contact forces. The Barrett hand has a limited amount of internal strain gauge force sensing capability built into it, and the tactile system can be used to accurately quantify contact forces in conjunction with the strain gauge system. Vision can be an effective sensing modality for grasping tasks due to its speed, low cost, and flexibility. It can serve as an external sensor that can provide control information for devices that lack internal sensing or that would require extensive modification and re-engineering to provide contact and force sensing. Using a vision system, a simple uninstrumented gripper/hand can become a precision device capable of position and possibly even force control. Additionally, when vision is coupled with any existing internal hand sensing, it can provide a rich set of complementary information to confirm and quantify internal sensory data, as well as monitoring a task's progress.

2 Related Research

A number of previous researchers have explored the use of visual feedback and control to assist in the grasping task. Houshangi [7] tracked moving objects for grasping. Hollingshurst and Cipolla [6] have developed a system for positioning a gripper above an object in the environment using an affine stereo transform to estimate the object's position. Taylor et al. have used 3-D vision to guide the grasping task [15]. Castano and Hutchinson [3] use visual constraint planes to create compliant surfaces for constraint robot movement in the real world. Bendiksen and Hager [2] have used vision in conjunction with gripper forces to achieve stable grasps. Sharma et al. [11] use perceptual 3D surfaces to represent the workspace of the gripper and object and they plan their positioning tasks along these surfaces. Sobh and Bajcsy [13] examined how finite state machines can be used to monitor the grasping process through vision. Smith and Papanikolopolous [12] have recently extended their visual servoing and control algorithms to create a hand-eye tracker capable of grasping static and moving objects. There have been many previous efforts to include a robust set of tactile sensors on a robotic hand. An excellent overview of this field is provided by Nicholls [10]. Two recent papers that discuss using tactile sensors without vision to estimate forces and contacts are [8, 9]. Our own work has explored the capability of vision systems to track and grasp moving objects [1] and use uncalibrated visual servoing to perform alignment tasks [17]. This work motivated us to use stereo vision to control an uninstrumented gripper in simple grasping tasks [18, 16]. The robotic hand used in that research had no internal sensing, and thus vision system could not report forces being applied to the grasped objects.

*This work was supported in part by DARPA contract DACA-76-92-C-0007, NSF grants CDA-90-24735 and IRI-93-11877, and an ONR MURI Grant

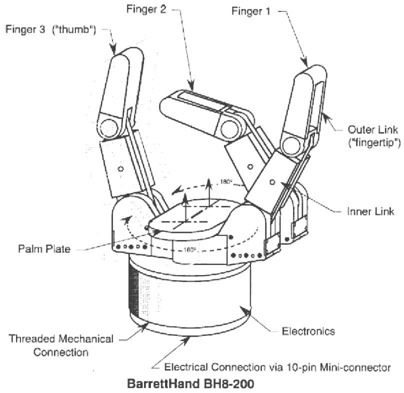

Figure 1: The Barrett Hand

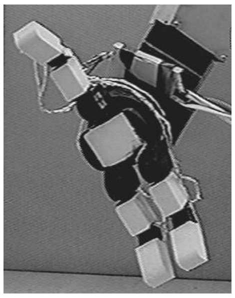

Figure 2: Hand with tactile sensor suite attached. There are sensor pads on the inner and outer links of each finger and a palmar tactile pad

3 The Barrett Hand

The dextrous robot hand used for our investigation is the Barrett Hand depicted in Figures 1 and 2. It is an eight-axis, three-fingered mechanical hand with each finger having two joints. One finger is stationary and the other two can spread synchronously up to 180 degrees about the palm (finger 3 is stationary and fingers 1 and 2 rotate about the palm). Although there are eight axes, the hand is controlled by four motors. Each of the three fingers has one actuated "inner" link, and a coupled "outer" link that moves at a fixed rate with the inner link. A novel clutch mechanism allows the outer link to continue to move if the inner link's motion is obstructed (referred to as *breakaway*). An additional motor controls the synchronous spread of the two fingers about the palm. Various grasp classifications capable with the hand include but are not limited to: power, hook, capture, cylinder-tip, spherical, and cylinder grasps.

3.1 Native Joint Position and Force Sensing

Each of the four motors in the hand is equipped with an optical position encoder that can supply angular data for the finger joint where it meets the palm (referred to as the inner joint). A kinematic model of the hand has been developed and calibrated so that joint level information can be mapped to finger position in space via forward kinematics. There is no encoder at the joint between the inner and outer links (referred to as the outer joint). In normal operation, the outer joint of each finger is driven at a 4:3 ratio with respect to the inner joint, and using the kinematic equations and this angle ratio, the finger's position in space can be computed. Often, however, due to backlash, this information is not exact. Moreover, in a breakaway situation, the angle of the outer joint cannot be derived from the optical encoder in the inner joint since the clutch has disengaged the links and the 4:3 ratio no longer is valid. This is exactly one of the situations where additional sensors are needed to compute contacts and forces.

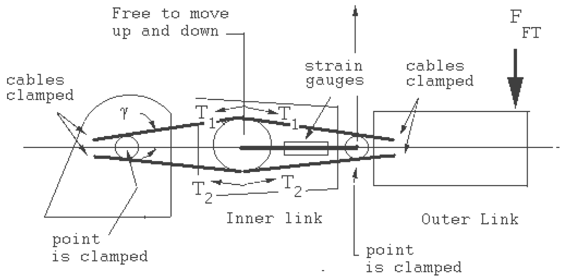

Figure 3: Location of Strain Gauges and Tensions

Each finger is equipped with a strain gauge. The gauges measure the axial strain in response to loads applied to the outer link due to coupled cable tensions (see Figure 3). The beam to which the strain gauges are attached to can be modeled as a cantilever beam under static equilibrium as follows.

$$-T_1 \sin\gamma - T_1 \sin\gamma + T_2 \sin\gamma + T_2 \sin\gamma - F_{FT} = 0 \quad (1)$$

$$F_{FT} = 2(T_2 - T_1)\sin\gamma = 2\Delta T \sin\gamma \quad (2)$$

where F_{FT} is the force along the finger's outer link, T_1 and T_2 are the tensions along the upper and lower cables respectively, and γ is the directional angle these tensions act along. Thus, given a strain gauge reading ΔT, F_{FT} can be calculated.

The point forces act in the x-y plane. The General Beam Method Deflection at any point along the finger's outer link is determined with the General Beam Method. Beam curvature is related to a bending moment M_b as

$$\frac{d^2y}{dx^2} = \frac{M_b}{EI_{zz}} \quad (3)$$

E is Young's modulus and I_{zz} is the moment of inertia about the z-axis (pointing out). Vertical deflection is given by

$$y(x) = \frac{F_{FT}\,x^3}{6EI_{zz}} - \frac{F_{FT}\,Lx^2}{2EI_{zz}} \quad (4)$$

where L is the length of the beam. Maximum deflection occurs at $x = L$, that is, at the finger tip:

$$y(L) = -\frac{F_{FT}\,L^3}{3EI_{zz}} \quad (5)$$

3.2 Predicting Forces from Strain Gauges

It is desired to solve for the applied forces on the finger, F_{FT} as introduced in (4) without explicitly measuring the vertical finger displacement $y(x)$. This can be

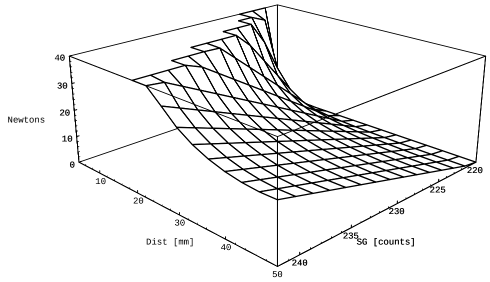

Figure 4: Finger Force as a function of strain gauge readings(s_g) and distance (x)

done by assuming that the strain gauge output s_g and $y(x)$ are linear such that

$$s_g = k_1 F_{FT} x^3 - k_2 F_{FT} x^2 + a + \varepsilon \quad (6)$$

where a is an offset and error ε. We can use a least-squares solution to find the parameters above and relate s_g to x. In this calibration process, fixed weights ranging from 1.2 lbs (4.72 N) to 4 lbs (18.09 N), at 0.2 lbs intervals, were suspended along the finger at 2 mm intervals. Since the strain gauges as illustrated in Figure 3 are configured in a conventional Wheatstone bridge, it is reasonable to assume linearity between s_g and $y(x)$. Figure 4 is a plot of the finger force vs. strain gauge and distance function. By kinematically deriving the predicted beam behavior, and empirically determining the parameter estimates, a predictive force model can be solved for without explicitly measuring the vertical displacement as:

$$F_{FT} = \frac{s_g - a}{k_1 x^3 - k_2 x^2} \quad (7)$$

Furthermore the moment of inertia I_{zz} for the finger does not need to be explicitly calculated. This would pose to be challenging given both the non-uniform cross sectional area and non-homogeneous composition of the finger. Experimental data in Section 6 demonstrates the validity of the above predictive force model.

4 External Vision Sensing

The strain gauge output can report the forces acting on the outer link, but cannot localize them. The strain gauges only provide us with torque readings about the outer joints - it is necessary to find the point of contact along a finger to determine the force normal to that finger. Vision sensing can be used to provide this contact point estimation, and thereby calculate the actual finger tip forces in conjunction with the strain gauge readings and the kinematic model developed in the previous section. The effective requirements of any vision sensing system to perform this integration task include real-time position estimation of multiple contacts as the fingers are moving, and estimation of the object position.

We are using a scaled orthography single camera model which effectively allows us to determine the 3-D position of fingers and contacts from the image plane directly. The methods described here can be easily extended to a stereo perspective camera model that can provide full 3-D depth recovery as well [16]. The vision system we are using is a modification of Hager's X Vision system [5]. Each tracker has a state vector consisting of position and orientation information which is updated after each iteration of the tracking loop. Once a line or region tracker is initialized on an edge or window within the image it will track the feature at frame rates. In our experiments we used three different trackers: a corner tracker that is composed of two line trackers that reflect the intersections of the lines along the inner and outer links of a finger, another line tracker to track a spike-like force probe that contacts the finger, and an SSD tracker for the endpoint of the finger. Using the scaled orthographic camera model, we were able to track a finger and determine a point of contact to within 1 mm of the actual contact point.

5 Tactile Sensor Suite

While vision can track fingers and help to determine contact points, it is important to note that occlusion can prevent the vision sensor from reporting this information. This can be overcome by the use of a finger mounted tactile sensor that can estimate contact point localization. We have designed a set of tactile sensors that can be mounted on the fingers of the hand, covering the active surfaces of the fingers and the palm. The pads use a capacitive tactile sensor designed by Howe [14] that is based upon an earlier design of Fearing's [4]. The sensor is designed to be slipped on to the links of the finger as shown in Figure 2. The electronics package is mounted on the robot wrist with wiring to each pad on the fingers and palm. The tactile sensor geometry on each finger link is a 4x8 grid with each capacitive cell approximately 3 mm by 3 mm and 1 mm spacing between tactile elements (tactels), and the sensor can bend to the curve of the fingertip. This provides 64 sensors per finger, plus an 8x8 grid on the palm for a total of 256 sensors on the hand. The sensor is covered with a compliant elastomer that allows force distributions to be spread over the sensing surface. While the intensity values need to be carefully calibrated to provide accurate force information, it is possible to use their relative responses to compute the weighted center of contact of the applied force using moments. We have recently constructed a pressure vessel that can be used to supply a constant force over the entire surface of the sensor. To calibrate the sensor for position contact, we probed one column of the sensor which was aligned along the length of the finger at different distances along its length and it reported the contact center with a precision of about 1.2 mm. Figure 5 shows the predicted relationship between tactel location and distance along the finger and sample probes along one of the sensors attached to the outer link.

6 Experiment I

The purpose of this experiment was to see if vision sensing could track finger and object movements, and then localize contact along the finger to estimate actual applied finger forces using the strain gauge calibration data. We mounted a spike-like force probe on top of an

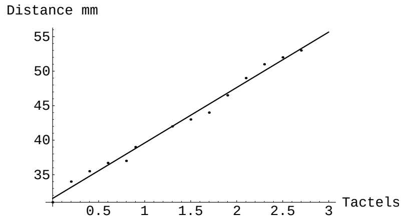

Figure 5: Contact distance vs. tactile localization along a single column of the tactile sensor. Bold line is actual distance versus tactile reading, and points are sample probes along one of the outer links.

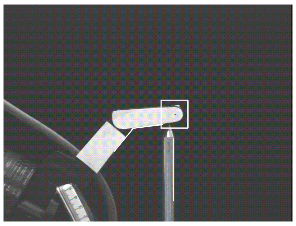

Figure 6: Visual Sensing: Line tracker, corner tracker, and SSD tracker applied to finger

ATI force sensing wrist that provided us with accurate three dimensional force data which we used as ground truth. The hand was mounted on a PUMA-560 robot and positioned in the vicinity of the spike. One finger of the hand was positioned above the spike as the trackers were initialized. To find the point where the spike contacts the finger we used one line tracker initialized on the right side of the spike, one corner tracker placed along the inside edges of the the two links of the finger, and an SSD tracker initialized so that it is centered on a point marked at 7 mm from the end of the finger which is used as a reference (see Figure 6). As the finger closed on top of the spike, the trackers followed it.

Using the state vectors of the trackers (see Figure 7), a point of intersection in image space was computed:

$$y = L1.y - offset, \qquad x = \frac{y - C.y}{tan(C.\theta_2)} + C.x \qquad (8)$$

where $L1.y$ is y-coordinate of the line tracker following the spike, $(C.x, C.y)$ is the position of corner tracker, $C.\theta_2$ is the orientation of the upper line in the corner tracker, and $offset$ is the distance in pixels from the center of the spike to the edge. The point of finger contact

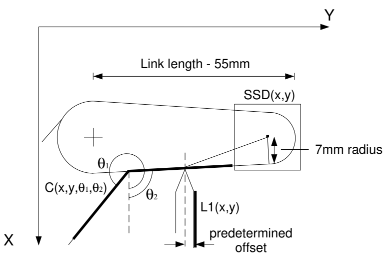

Figure 7: Using three feature trackers, the point of contact can be computed.

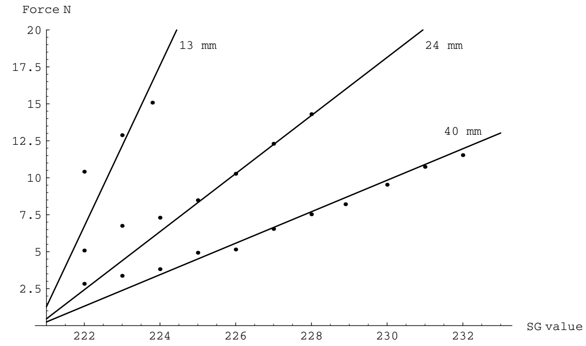

Figure 8: Results of Experiment I: Graph shows force data when spike is at 13mm, 24mm, and 40mm from outer joint. Solid Line represents modeled linear relationship between strain gauge values and forces for a contact at the given distance along the finger. Points are actual force readings from the ATI force sensing wrist for contact at the distance determined by vision.

and the point tracked by the SSD were then transformed to world coordinates using the calibration matrix.

This method resulted in fast, reliable measurements within 1mm of the actual point of contact. Near the end of the finger where the distance between the contact point and the tracked endpoint reference was small, our accuracy dropped to about 1.5mm from 1mm. Using 10 averaged distance measurements and 10 averaged strain gauge readings taken from the hand we were able to predict the force applied to the finger at the point of contact. The standard deviation in position was .4 mm along the finger except at the very tip where it increased to .9 mm. Our accuracy was also limited by the low resolution of the strain gauges in this force range. The maximum range was near four bits at the end of the finger where the maximum deflection of the beam occurs. Our results (see Figure 8) showed that as we placed the spike further along the finger our error dropped until we reached the very end of the finger. Generally we found that as more force is applied to the spike, our error percentage dropped significantly.

7 Experiment II

One of the Barrett Hand's unique features is its ability to complete a grasp of an object after the inner link's motion has been obstructed. During this breakaway situation, the clutch mechanism in each finger allows the outer links to continue closing after its inner link has been stopped. This feature is especially useful for grasping large objects or irregularly shaped objects where the inner links may be blocked early, and the outer links finish the grasp. Vision sensing can provide the joint angle by calculating the difference in orientation angle for the two line trackers along the inner and outer links of the finger. Knowing this joint angle and the link geometry allows us to use forward kinematics and locate the last link in space. During breakaway we cannot localize the contact along the outer link without utlizing additional sensing.

In this experiment we rigidly mounted a small block to the palm of the hand and closed the third finger around

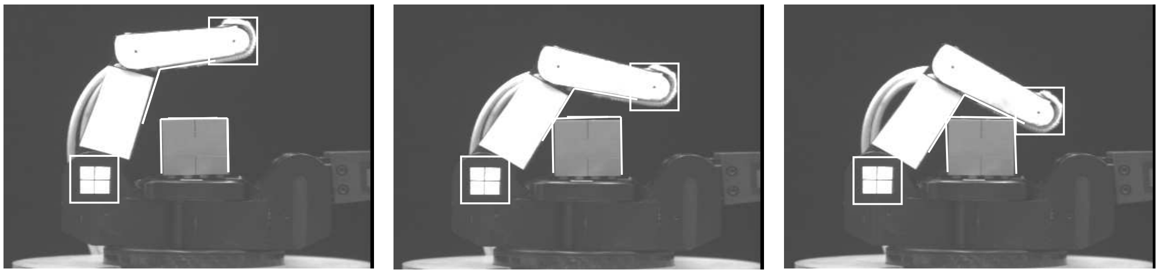

Figure 9: Left to right: The trackers after initialization. The inner link is stopped, and breakaway occurs. The outer link completes the grasp and the contact points are computed.

it. By securing the block from sliding, we were able to better ensure the obstruction of the first link, and cause the grasp to result in two contacts by the inner and outer links on the block which we attempted to locate. The motion trackers in this experiment included two SSD trackers and three corner trackers. One SSD tracker was initialized on the end of the finger and a second one was centered on a cross hair marking the inner joint's axis. Two corner trackers were placed on the block, one on each side, and the third corner tracker was placed along the inside edges of the finger as before. After the trackers had been initialized, the finger was closed around the block (see Figure 9). Using this method, we were able to track the angle of the joint between inner and outer links at all times. We found that the vision system reported contact points that were within 2 mm of the actual contact points. We encountered some difficulty in keeping the first SSD tracker centered on the mark near the end of the finger. We also used the tactile sensor on the outer link of finger 3 to provide us with additional contact information. In this experiment, the actual contact distance along the outer link was determined to be 38 mm as read by a ruler. The vision system reported 40.0mm as the distance and the tactile sensor reported it as 36.6mm. The actual contact distance along the inner link was determined to be 49 mm as read by a ruler and the vision system reported 51.3 mm as the distance.

8 Experiment III

In this experiment we attempt to securely grasp the lid of a canister and unscrew it from the base using the tactile and kinematic information. This requires repeatedly grasping the lid, rotating it 180 degrees, releasing the lid, and rotating back. Each time the lid is grasped, the tactile sensors on the fingertips report the points of contact along the length of the finger, as well as an indication of the force applied. For this experiment it is sufficient to verify that the normal force of each finger tip is over an experimentally derived threshold.

The canister was rigidly mounted to the table, while the Barrett hand was mounted on a PUMA arm that suspended it directly over the lid. Fingers 1 and 2 were rotated so that they directly opposed each other and were each 90 degrees from Finger 3 (see Figure 10). At the start, the lid was screwed down 10 revolutions, and fingers 1 and 2 closed on the bottom edge. This formed a line contact on the tactile sensors which reported the weighted centroids of the contacts, as well as

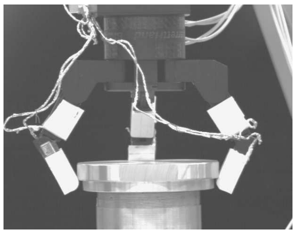

Figure 10: Hand grasping canister top for unscrewing experiment.

the total intensity count for each pad. If the total intensity was below the threshold the finger were commanded to close more tightly until the threshold was passed. At that point the values of the joint encoders were recorded and the vertical distance of the contact points from the palm could be computed using the hand kinematics. The PUMA then rotated the hand 180 degrees counterclockwise, and the height of the lid was recomputed using the same procedure. The fingers then released the lid and rotated back clockwise 180 degrees, to perform the procedure 19 more times until the lid was removed. The third finger did not actually apply force to the lid as it was unopposed on the other side, but served to steady the lid when it was time to remove it. The vertical displacement of the lid was recorded at the start and end of each half revolution, and was plotted against the true pitch of the threads (20 threads per inch or 1.27 mm/thread). Figure 11 shows the height of finger 2's contact as measured by the tactile sensor and the hand kinematics. The open circles are the beginning contact position and the closed circles are the ending contact after a half revolution. The figure shows 10 full revolutions of the canister lid which was when the lid began to come off. After 9 revolutions of the hand, the tactile sensors reported a vertical displacement of 11.1 mm while the actual displacement was 11.4 mm.

9 Conclusions and Future Work

Humans appear to have very easily integrated position, force and vision sensing during grasping. In this paper, we have described the design and use of a set of internal tactile sensors and external vision modules to extend some of these capabilities to a robotic hand system. The tactile sensors can report positional contact location also, and may be used when vision is occluded,

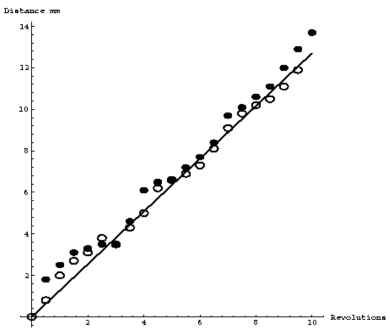

Figure 11: Vertical displacement of canister top as measured by tactile sensor and kinematic model for finger 2. The bold line is actual screw pitch and the open and closed circles represent the computed start and end vertical displacement for each half revolution.

as well as confirming any visually determined contact positions. The vision modules can be used in a control loop to 1) determine if contact has occurred and 2) estimate the position of contact along a finger. Vision has the advantage of being fast and simple to add to an existing hand system that may be lacking in sensory capability. We have also calibrated the internal force sensors of the hand to fuse the visual contact position information with the modeled strain gauge values and correctly predict applied finger forces. This will allow a grasping task to 1) visually determine the grasping points of contact and 2) modify the control of the finger to apply pre-determined forces on an object.

We are currently extending the system described here in a number of ways. First, we are extending the vision system to a 3-D binocular stereo setup. By using pairs of feature trackers, one in each image, we will be able to compute depth information in the scene. Also, by using more than one camera for external visual sensing, we can reduce the occlusion problems during grasping. Second, we are applying these methods to a set of grasping tasks that include handling deformable and fragile objects, where finger forces need to be pre-determined and monitored. We still need to characterize the finger forces from the tactile responses.

Acknowledgment: We would like to thank Rob Howe, Jae Son, Bill Peine, and Parris Wellman from the Harvard Robotics Lab for their help in building the tactile sensor system.

References

[1] P. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Trans. on Robotics and Automation*, 9(2):152–165, 1993.

[2] A. Bendiksen and G. Hager. A vision-based grasping system for unfamiliar planar objects. In *IEEE International Conference of Robotics and Automation*, pages 2844–2849, May 1994.

[3] A. Castano and S. Hutchinson. Visual compliance: Task-directed visual servo control. *IEEE Trans. on Robotics and Automation*, 10(3):334–342, June 1994.

[4] R. S. Fearing. Tactile sensing for shape interpretation. In S. T. Venkataraman and T. Iberall, editors, *Dextrous Robot Hands*. Springer-Verlag, 1989.

[5] G. D. Hager and K. Toyama. X vision: A portable substrate for real-time vision applications. Technical report, Department of Computer Science, Yale University, 1995.

[6] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. Technical Report CUED/F-INFENG/TR126, Department of Engineering, University of Cambridge, 1993.

[7] N. Houshangi. Control of a robot manipulator to grasp a moving target using vision. In *IEEE International Conference on Robotics and Automation*, pages 604–609, Cincinnati, May 13-18 1990.

[8] D. Johnston, P. Zhang, J. Hollerbach, and S. Jacobsen. A full tactile sensing suite for dextrous robot hands and use in contact force control. In *Proc. of the 1996 IEEE International Conference on Robotics and Automation*, pages 3222–3227, 1996.

[9] H. Maekawa, K. Tanie, and K. Komoriya. Dynamic grasping force control using tactile feedback for grasp of multifingered hand. In *Proc. of the 1996 IEEE International Conference on Robotics and Automation*, pages 2462–2469, 1996.

[10] H. R. Nicholls. *Advanced Tactile Sensing for Robotics*. World Scientific Press, 1992.

[11] R. Sharma, J. Herve, and P. Cucka. Analysis of dynamic hand positioning tasks using visual feedback. Technical Report CAR-TR-574, Center for Auto. Res., University of Maryland, 1991.

[12] C. Smith and N. Papanikolopoulos. Vision-guided robotic grasping: Issues and experiments. In *Proc. of the 1996 IEEE International Conference on Robotics and Automation*, pages 3203–3208, 1996.

[13] T. M. Sobh and R. Bajcsy. Autonomous observation under uncertainty. In *Proc. of IEEE International Conference on Robotics and Automation*, pages 1792–1798, May 1992.

[14] J. Son and R. Howe. Tactile sensing and stiffness control with multifingered hands. In *Proc. of the 1996 IEEE International Conference on Robotics and Automation*, pages 3228–3233, 1996.

[15] M. Taylor, A. Blake, and A. Cox. Visually guided grasping in 3d. In *IEEE International Conference of Robotics and Automation*, pages 761–766, May 1994.

[16] B. Yoshimi. *Visual Control of Robotics Tasks*. PhD thesis, Dept.of Computer Science, Columbia University, 1995.

[17] B. Yoshimi and P. Allen. Alignment using an uncalibrated camera system. *IEEE Trans. on Robotics and Automation*, 11(4):516–521, August 1995.

[18] B. H. Yoshimi and P. Allen. Visual control of grasping and manipulation tasks. In *MFI '94: 1994 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 575–582, 1994.