Integration of Vision, Force and Tactile Sensing for Grasping

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 report accurate 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. A number of experiments are described that show how these sensors can be combined to increase the capabilities of a robotic hand for grasping and manipulation tasks.

1 Introduction

Grasping arbitrary objects with robotic hands remains a difficult task 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 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.

In this paper, we describe the design and use of a set of additional sensors that augment the capability of a robotic hand. These sensors can be used in a complementary way to integrate vision, force, and tactile information for grasping tasks that require closed-loop, real-time control. Without this sensory feedback, open loop control must be used which requires precise models of the environment to be effective. The experimental hand we are using is the Barrett Hand, which is a three-fingered, four DOF hand. This hand has limited internal sensing capability to which we have added two different sensory systems. The first is a set of real-time vision modules that can be used to track and monitor the hand in real-time as it performs a task, and the second is a set of tactile sensors covering the links of the hand as well as the palm. Our aim is to use this set of internal and external sensors to robustly estimate positions, forces and contacts on the hand. 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

*This work was supported in part by an ONR/DARPA MURI award ONR N00014-95-1-0601, DARPA AASERT awards DAAHO4-93-G-0245 and DAAH04-95-1-0492, and NSF grants CDA-96-25374 and IRI-93-11877.

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 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 monitor a task's progress. 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 quantify contact forces in conjunction with the strain gauge system. An important point is that both of these systems can be retro-fitted with existing, limited-sensing capability hands to improve their capabilities.

The organization of this paper is as follows. In section 2, we describe related research on this problem. Section 3 describes the robotic hand we are using for our research, a kinematic model of the hand, and a force model we have developed for estimating the applied forces from internal strain gauge readings. Section 4 describes the external vision sensing modules we have developed to track the fingers of the hand and compute contact positions. Section 5 describes the tactile sensors we have built and mounted on the robot hand to better estimate contact positions and forces. Section 6 details a series of experiments that utilize these sensors to find contacts and forces on the hand, including a complete task involving integrated grasping and manipulation.

2 Related Research

A number of previous researchers have explored the use of visual feedback and control to assist in the grasping task. Houshangi [9] tracked moving objects for grasping. Hollingshurst and Cipolla [8] 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 [17]. Castano and Hutchinson [4] use visual constraint planes to create compliant surfaces for constrained 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. [13] use perceptual 3D surfaces to represent the workspace of the gripper and object and plan their positioning tasks along these surfaces. Sobh and Bajcsy [15] examined how finite state machines can be used to monitor the grasping process through vision. Smith and Papanikolopolous [14] 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. Two excellent overviews of this field are provided by Dario [5] and Nicholls [12]. Two recent papers that discuss using tactile sensors without vision to estimate forces and contacts for grasping are [10, 11]. Our own related work has explored the capability of vision systems to track and grasp moving objects [1], the use of uncalibrated visual servoing to perform alignment tasks [20], and stereo vision to control an uninstrumented gripper in grasping tasks [21, 19].

3 Barrett Hand: Kinematics and Internal Sensing

The dextrous robot hand used in our research is the Barrett Hand shown in Figure 1. It is an 8-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 8 axes, the hand

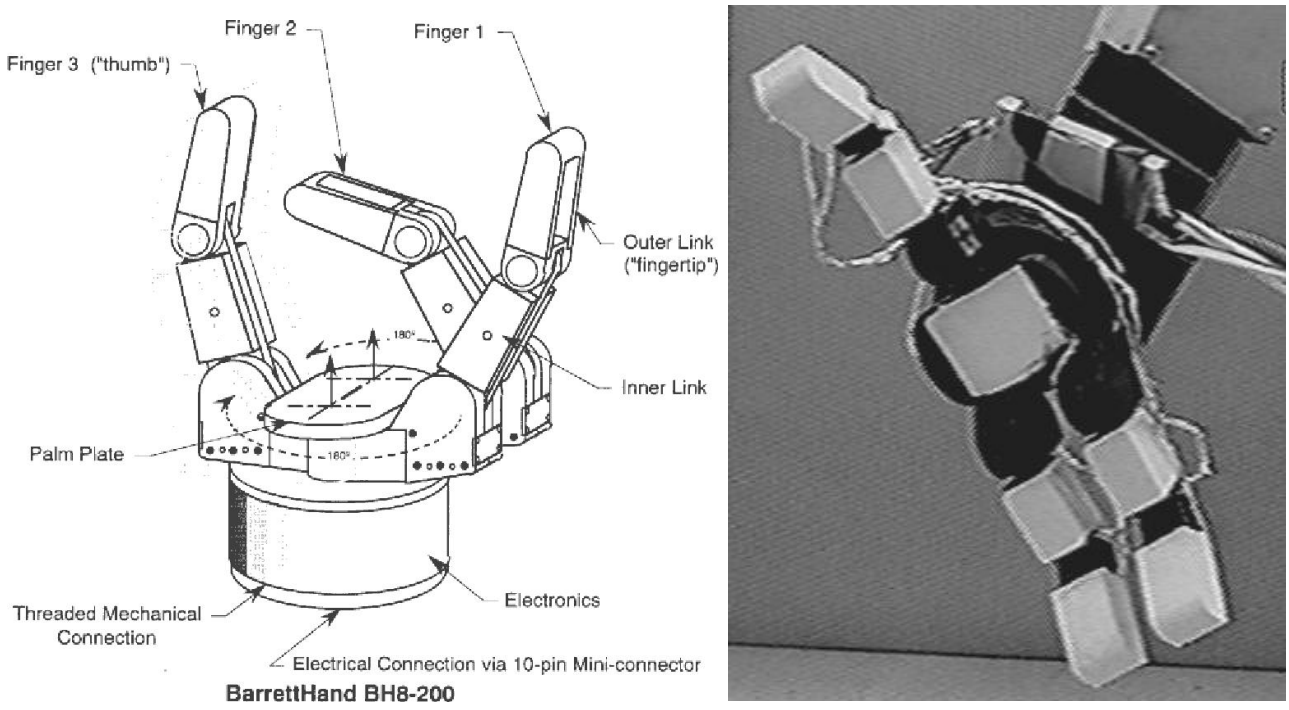

Figure 1: Barret Hand schematic and hand with tactile sensor suite attached. There are sensor pads on the inner and outer links of each finger and a palmar tactile pad

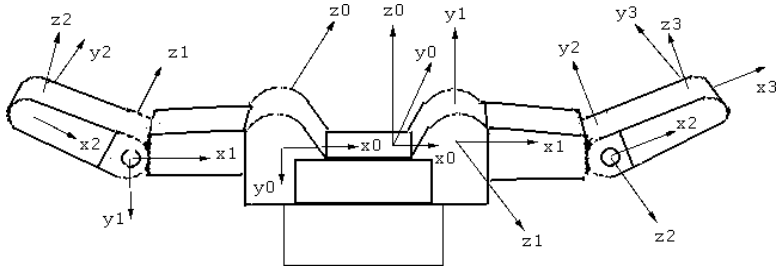

Figure 2: Finger Link Frames

is controlled by 4 motors. Each of the 3 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 2 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. Figure 2 illustrates the location of base, link and tool (finger tip) frames using the conventional Denavit-Hartenberg notation. Finger 2 is not shown for clarity, but will have the same frame orientations as finger 1. Table 1 provides the kinematic link parameters and Figure 3 depicts the link notation and dimensions.

3.1 Internal Force Sensing

A force sensing mechanism consisting of a flexible beam, a free-moving pulley, a pair of cables, and four strain gauges (in a Wheatstone bridge configuration) is completely contained inside each finger (see Figure 4).

Here, the pulley can move vertically with respect to the inner link. Its movement is in

Fingers 1 and 2

Link	θ_k	d_k	a_k	α_k
$k=1$	θ_1	0	l_1	90
$k=2$	θ_2	0	L_1	0
$k=3$	$\theta_3+\psi_{f1}$	0	L_2	-90

Finger 3 =

Link	θ_k	d_k	a_k	α_k
$k=1$	θ_1	0	$-L_1$	0
$k=2$	$\theta_2+\psi_{f3}$	0	$-L_2$	90

Table 1: Kinematic Link Parameters for Barrett Hand

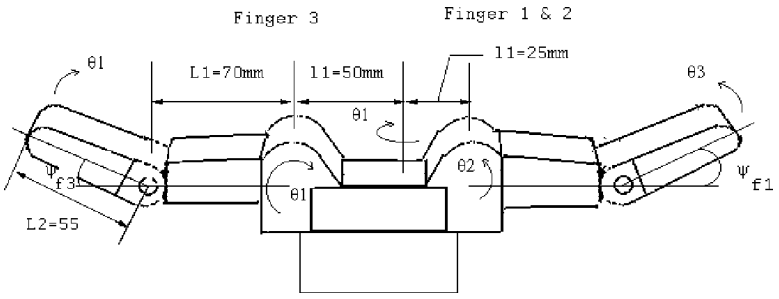

Figure 3: Link Notation and Dimensions

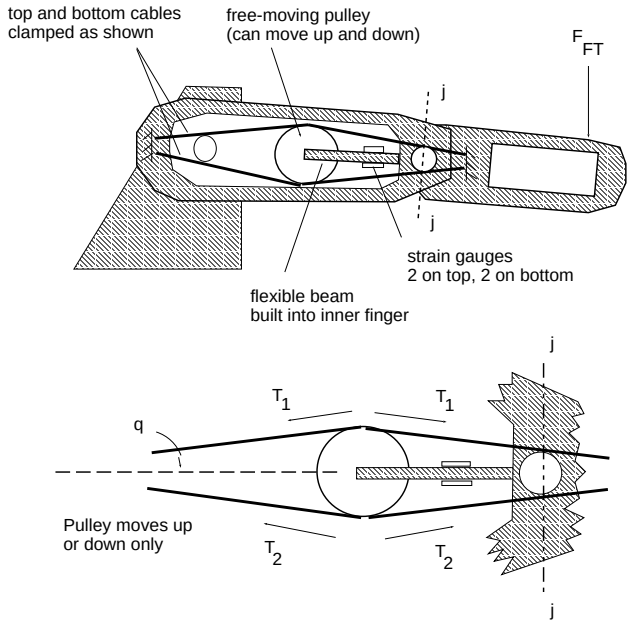

Figure 4: Cutaway diagram (top) shows the location of strain gauges inside the finger's inner link. A blowup figure (bottom) of the flexible beam illustrates that the pulley moves vertically in response a force F_{FT} on the outer link

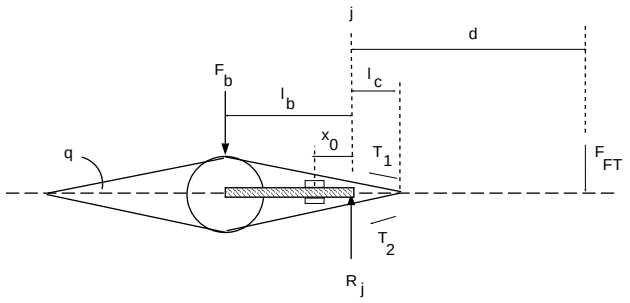

Figure 5: Free body diagram. Forces and moments taken about point j

response to the difference in the tensions of the top and bottom cables due to a force F_{FT} applied to the finger's (finger tip) outer link.

The Barrett Hand was factory installed with strain gauges. The hand's onboard electronics and A/D return an 8-bit number in response to the moments the strain gauges measure on the flexible beam. We would like to have a model for determining F_{FT} given a strain gauge reading s_g and a distance d measured from the joint j along the outer link. We thus performed a calibration in the following manner.

We first collected strain gauge readings by suspending weights ranging from 1.2 lbs (4.72 N) to 4 lbs (18.09 N) at 0.2 lbs intervals along the outer link at 2 mm intervals. Next we performed a least-squares fit of the above data using a model (6) derived from a free body diagram (Figure 5) of the flexible beam in the inner link.

For a cantilever, the strain gauge reading is related to a force, F_b by the following [3].

$$s_g = k_1(l_b - x_0)F_b \tag{1}$$

where x_0 is the distance to the strain gauges, l_b is the distance to where the force F_b is applied. Both distances are measured along the cantilever. k_1 is a constant that depends on the physical properties of the beam, such as its Young's modulus and dimensions.

Assuming point forces, the force F_b is equal to the reaction force R_j. T_1 and T_2 are the tensions along the upper and lower cables respectively. q is constant and is the directional angle these tensions act along. l_c is the length of cable extending past j. Summing forces and moments about the point j, we have

$$F_{FT} - R_j - T_1 \sin q + T_2 \sin q = 0 \tag{2}$$

$$-T_1 l_c \sin q + T_2 l_c \sin q + F_{FT} d = 0 \tag{3}$$

Solving for $(T_1 - T_2)\sin q$ and substituting into (2) yields

$$F_b \equiv R_j = F_{FT}(1 - \frac{d}{l_c}) \tag{4}$$

Using this form of F_b in (1) we have

$$s_g = k_1(l_b - x_0)F_{FT}(1 - \frac{d}{l_c}) \tag{5}$$

l_b, x_0 are constant. In our calibration we performed a least squares fit using the following model

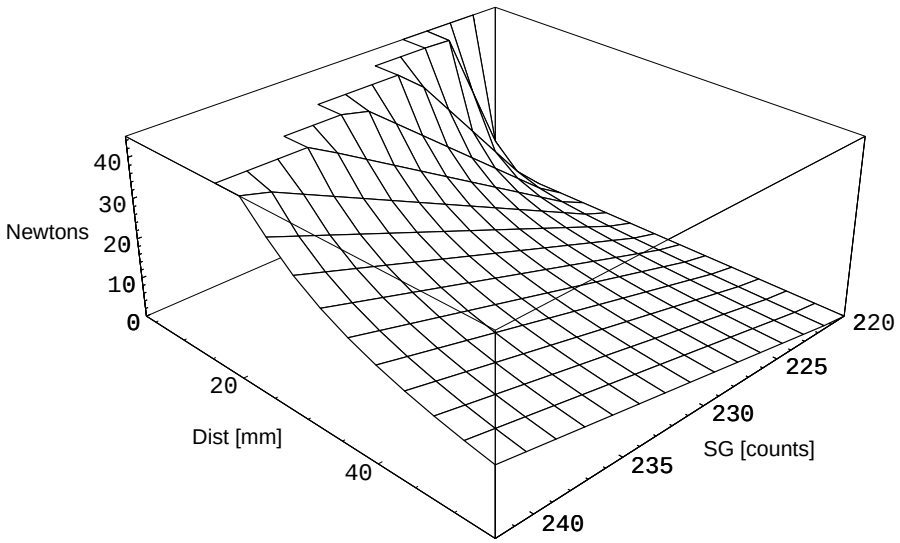

Figure 6: Finger Force as a function of strain gauge readings(s_g) and distance (d)

$$s_g = kF_{FT} + k'F_{FT}d + a \tag{6}$$

to obtain values for the three unknowns k, k' and (offset) a. We thus have a force model with the following form Figure 6 is a plot of the finger force vs. strain gauge and distance function.

$$F_{FT} = \frac{s_g - a}{k + k'd} \tag{7}$$

4 External Vision Sensing

Our goal is to monitor and control the fingers of a robotic hand with vision sensing as it performs grasping and manipulation tasks. Our motivation for this is the general lack of accurate and fast feedback from most robotic hands. Many grippers lack sensing, particularly at the contact points with objects, and rely on open loop control to perform grasping and manipulation tasks. Vision is an inexpensive and effective method to provide the necessary feedback and monitoring for these tasks. Below, we outline some aspects of visual control that are well suited to the grasping problem:

1. Vision can determine grasp points. This is usually a preliminary step before actual grasping takes place, and may not be as time critical as the manipulation task itself.
2. Vision can be very important in dealing with unstructured and/or moving environments, where model-based knowledge may be unavailable or errorful. This is an example of the active vision paradigm.
3. Once a grasp has been achieved, vision can monitor the grasp for stability. By perturbing the fingers, we can measure the amount of displacement and types of displacement in image space of the object. If the object does not move correctly, we can say that the grasp is faulty.
4. Visually monitoring a task will give us the feedback necessary to perform the task or to determine if an error has occurred.

While visual control of grasping can be very helpful, we need to recognize some problems associated with it. The problems listed below need to be adequately addressed in order to successfully control grasping using vision, and are at the crux of why this is a difficult robotics problem.

1. Grasping and manipulation require real-time sensory feedback. Vision systems may not be able to provide the necessary analysis of the image fast enough to meet task and actuator timing requirements.
2. In grasping with a robotic hand, multiple fingers need to be employed. This entails having the vision system follow multiple moving objects in addition to the possible movement of any object to be manipulated.
3. Grasping and manipulation usually require 3-D analysis of relative relationships of fingers and objects. Simple vision systems only provide a 2-D projection of the scene.
4. As fingers close in on an object to be manipulated, visual occlusion of both the object and fingers can easily occur.
5. An important component of most grasping tasks is the knowledge of forces exerted by fingers. Vision systems can not directly compute accurate force measurements.

Visual control of grasping is not a panacea. The integration of vision and local contact/force information is needed for truly precise control of grasping and manipulation. The work described in this paper is aimed at highlighting what vision can provide and how it can be intelligently used in conjunction with other sensors for grasping.

The low-level vision modules we are using are a modification of Hager's X Vision system [7]. X Vision was chosen because it is a software implementation that can be applied to an existing hand system simply by hooking up a camera and frame grabber. It also has the advantage of portability across a number of platforms. Using X Vision, complex trackers can be built up from simple edge and region-based tracking primitives. 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 up to translation, rotation and scaling.

These trackers have several parameters which can be altered by the user. The length of the line tracker, as well as the width of the local window around the line, can be set. In the region based SSD (Sum of Squared Distances) correlation tracker many parameters affect its performance. These include window size and resolution, as well as several flags to enable or disable rotation, scaling, and brightness compensation. By seeding these trackers to follow edges or regions in the image, they will continuously track the features in real time. The trackers are very reliable for reasonable amounts of motion.

In the experiments described in section 6 we have used two different approaches to integrate vision with grasping, each of which relies on the basic tracking primitives described above. In the first approach, vision is used to provide contact point estimation for the fingers. The strain gauges can report the forces acting on the outer links of the fingers, 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. These experiments used a single camera that can image the fingers of the hand and the object to be contacted. We used a scaled orthography camera model which effectively allows us to determine the 3-D position of fingers and contacts from the image plane

directly. This use of vision sensing is not totally passive, as it needs to track the fingers as they move, but vision does not provide explicit feedback to the hand controller

In the last experiment we have implememted an active vision control system that uses real-time visual feedback to control the movement of multiple fingers and grasped objects for a grasping task. We are using an uncalibrated stereo method described by Xie [18] to visually servo the hand within the task space. A calibrated system would find the world space coordinates of two corresponding image points (subject to the quality of the calibration), and command the arm to move to that location. However, in many servoing applications precise world space coordinates are not needed, and an uncalibrated approach proves simpler and more robust. The uncalibrated strategy involves making relative, iterative movements to reduce the error in the image given visual reference features. The method projects the three world axes into image space to determine the mapping between changes in image space features and their world space equivalents.

In essence, this is an online Jacobian-based control. To find the mapping matrix, we move the robot in three non-coplanar directions, and record the image locations of the robot tool at each point. If we have four corresponding image points $\left\{({}^{l}a_i, {}^{r}a_i), \quad i = 1, 2, 3, 4\right\}$, whose coordinates are ${}^{l}a_i = ({}^{l}u_i, {}^{l}v_i)^t$ in the left image and ${}^{r}a_i = ({}^{r}u_i, {}^{r}v_i)^t$ in the right image, Xie derives the following relationship:

$$\Delta I = N \Delta P \tag{8}$$

where

$$N = \begin{bmatrix} {}^{l}u_1 - {}^{l}u_4 & {}^{l}u_2 - {}^{l}u_4 & {}^{l}u_3 - {}^{l}u_4 \\ {}^{l}v_1 - {}^{l}v_4 & {}^{l}v_2 - {}^{l}v_4 & {}^{l}v_3 - {}^{l}v_4 \\ {}^{r}u_1 - {}^{r}u_4 & {}^{r}u_2 - {}^{r}u_4 & {}^{r}u_3 - {}^{r}u_4 \\ {}^{r}v_1 - {}^{r}v_4 & {}^{r}v_2 - {}^{r}v_4 & {}^{r}v_3 - {}^{r}v_4 \end{bmatrix} \tag{9}$$

and ΔI is the image difference between the projections of the robot's current point and target setpoint in the two images, and ΔP is the world space difference between these two points.

Since N is non square (4x3, relating 3 robot coordinate changes to 2 image coordinate changes in each of 2 cameras), we can use a pseudo-inverse technique to invert it and obtain an approximation to the Jacobian:

$$\Delta P = (N^T N)^{-1} N^T \Delta I \tag{10}$$

In fact, this matrix is a constant and can be computed once at setup time. A proportional controller for visual servoing can be designed by multiplying it with a diagonal constant gain matrix. This ensures asymptotic image error convergence. This gain matrix was tuned empirically to give a quick response time without overshoot.

In use, an image error is computed and used to update a series of set points that move the robot to the target feature, at which time the image error is nulled out. Thus, we can effectively perform real-time control using image updates mapped through this matrix to generate control movements of the arm. The mapping matrix is robust to disturbances in the cameras' position and orientation. Small changes in camera position may induce errors in 3-D positioning, but the iterative nature of the control will eventually reduce the positioning error to zero. In the lab, we have moved the cameras laterally up to 12 inches and were still able to achieve convergence.

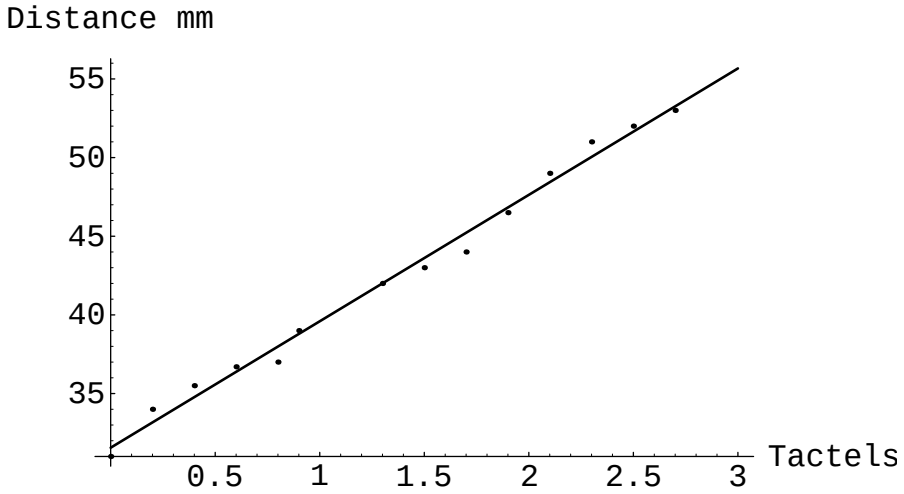

Figure 7: Contact distance vs. tactile localization along a single column of one of the tactile pads. Bold line is actual distance versus tactile reading, and points are sample probes along one of the outer links.

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 [16] that is based upon an earlier design of Fearing's [6]. The sensor is designed to be slipped on to the links of the finger as shown in Figure 1. 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 entire set of sensors can be scanned in approximately 40ms (25Hz). 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. 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 7 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 Experimental Results

A number of experiments have been performed to analyze the capability of a hand system using the three kinds of sensors: internal native joint position and force sensing, external tactile sensors, and external vision sensing. These experiments are aimed at showing how these sensors can be used in a complementary fashion to provide more robust and accurate information than each sensor by itself. We are also interested in how accurately finger contact forces can be

Figure 8: Finger of Barrett Hand applying force to instrumented probe

measured with such a set of sensors.

6.1 Experiment I: Verification of Finger Forces From Vision and Strain Gauge Sensing

The purpose of this experiment was to confirm that 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 model and calibration data. We mounted a spike-like probe on top of an ATI force sensing wrist that provided us with accurate three dimensional force data which we used as ground truth (see Figure 8). 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 reference point marked at 7 mm from the end of the finger (see Figure 9). As the finger closed on top of the spike, the trackers followed it in real-time. A single calibrated camera was used for tracking since the finger motion was in a plane orthogonal to the image plane. The camera, with a 50 mm lens, was placed approximately two meters from the hand, making the horizontal spatial resolution about 0.4 mm per pixel.

Using the state vectors of the trackers (see Figure 9), a point of intersection in image space was computed:

$$y = L1.y - offset, \quad x = \frac{y - C.y}{tan(C.\theta_2)} + C.x \tag{11}$$

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 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. The standard deviation in visually measured position was .4 mm along the finger except at the very tip where it increased to .9 mm. Using the force model in equation 7, we took 10 averaged visual distance measurements and 10 averaged strain gauge readings to predict the force applied to the finger at the point of contact. The results (see Figure 10) showed that as we placed the spike further along the finger our force error dropped until we reached the very end of the

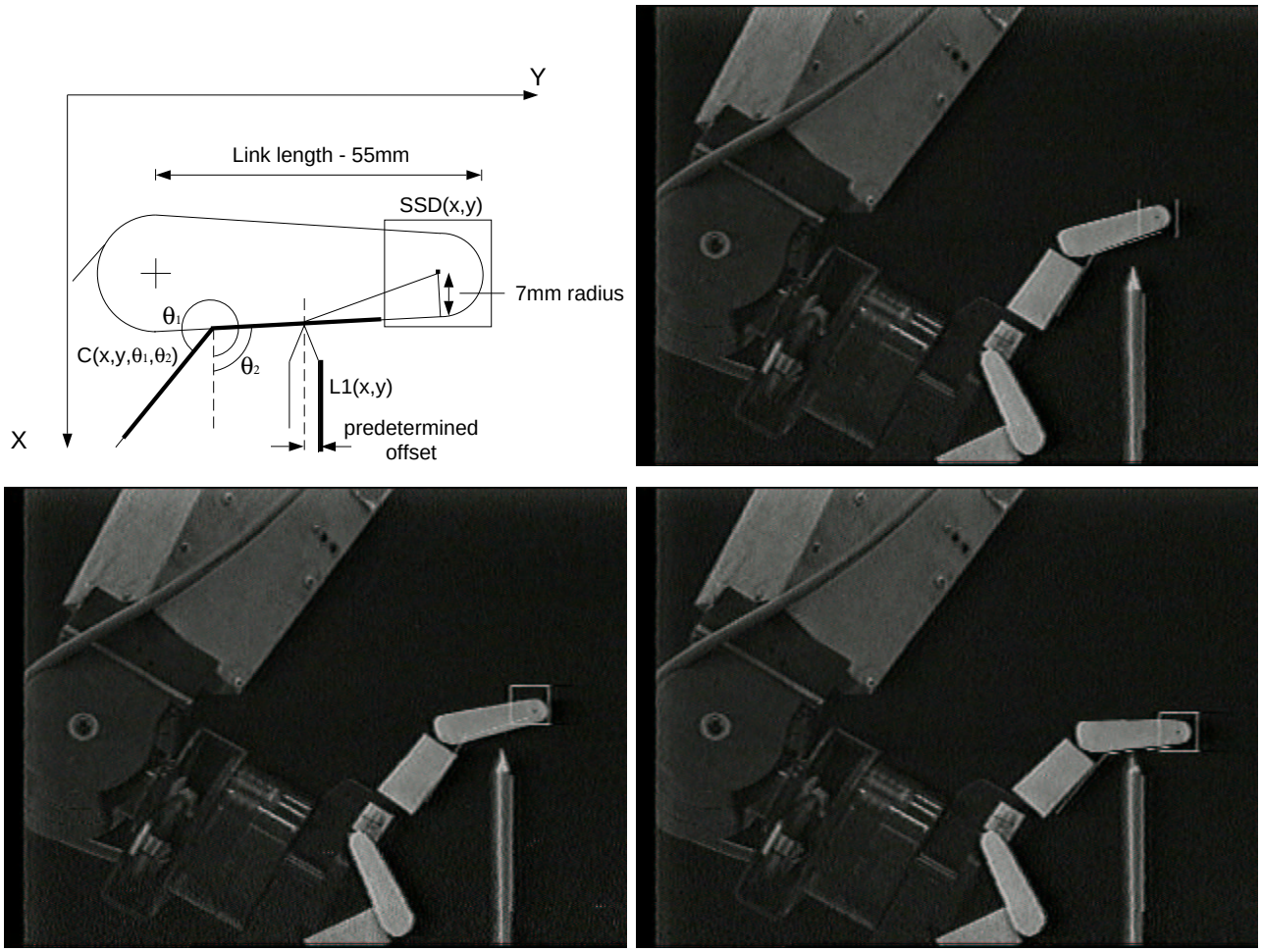

Figure 9: Top Left: Using three feature trackers, the point of contact can be computed. Top Right, Bottom Left and Right: Tracking sequence of Line tracker, Corner tracker, and SSD tracker as the finger is brought into contact with force probe

finger. Generally we found that as more force is applied to the spike, our error percentage dropped significantly. Our force accuracy was somewhat 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. Contacts near the joint between the outer and inner links caused the least deflection and the strain gauge readings were less accurate in this region, with only two bits of information available. By varying the velocity of movement of the finger, we were able to apply a range of forces on the spike, limited by maximum torque available from the hand's motor.

6.2 Experiment II: Finding Contact Points and Estimating Forces After Breakaway

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). Using the kinematic parameters of the hand, this allows us to calculate the endpoint of the inner link. 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 outer link's position in space can be computed. 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 link 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. 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. To find the position in space of the outer link during breakaway, we have used vision sensing to compute the outer link 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 to locate the last link in space.

To test this, we rigidly mounted a small block to the palm of the hand and closed the third finger around 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. A single camera was positioned in the same manner as the previous experiment. However, in this experiment we used two SSD trackers and three corner trackers to track the elements of the scene. 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 Figures 11, 12, and 13). 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.

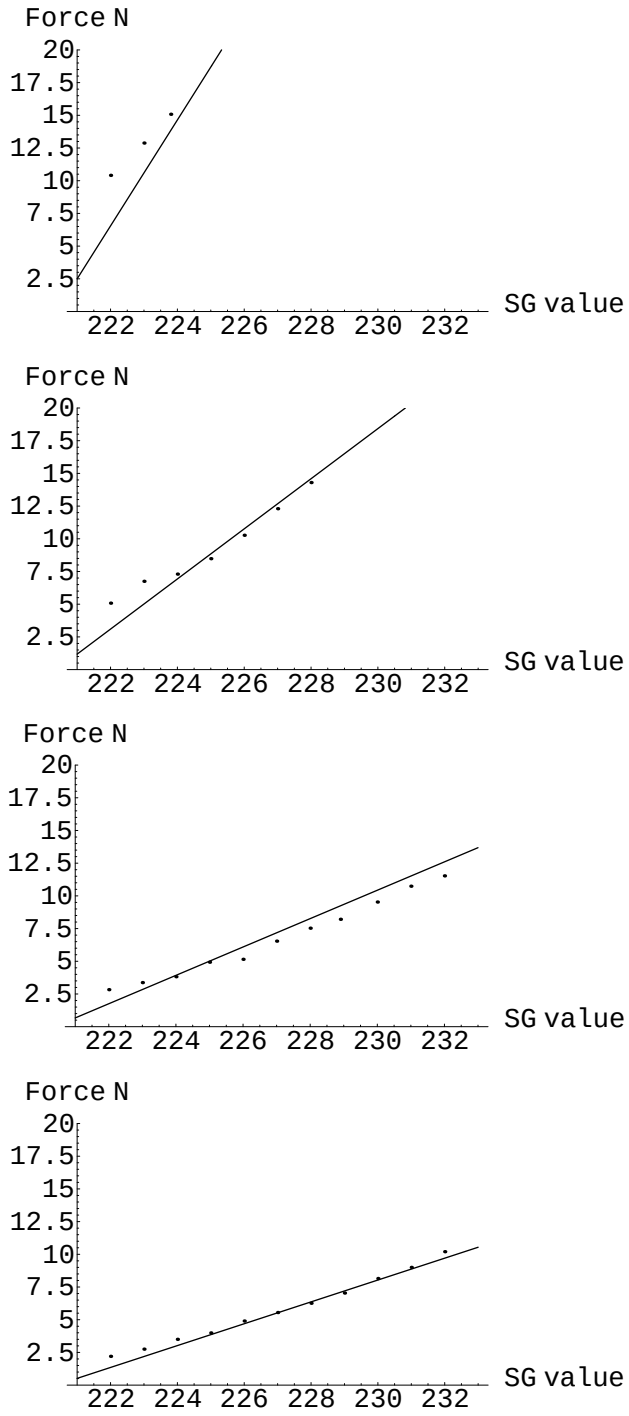

Figure 10: Results of Experiment I: Graphs (top to bottom) show force data when spike is at 13mm from outer joint, 24mm, 40mm, and 51mm. Solid Line represents modeled linear relationship between strain gauge values and forces for a contact at the given distance along the finger. Points are actual force readings from the ATI force sensing wrist for contact at the distance determined by vision. Contacts farther out on the link were modeled more accurately due to the limited range of the finger forces and strain gauge resolution near the finger joint.

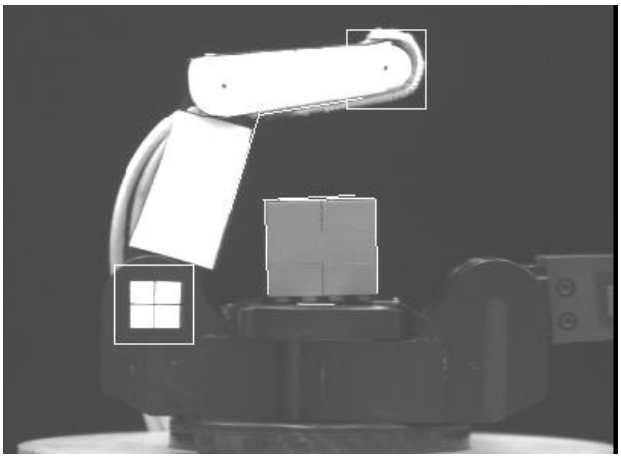

Figure 11: The trackers after initialization.

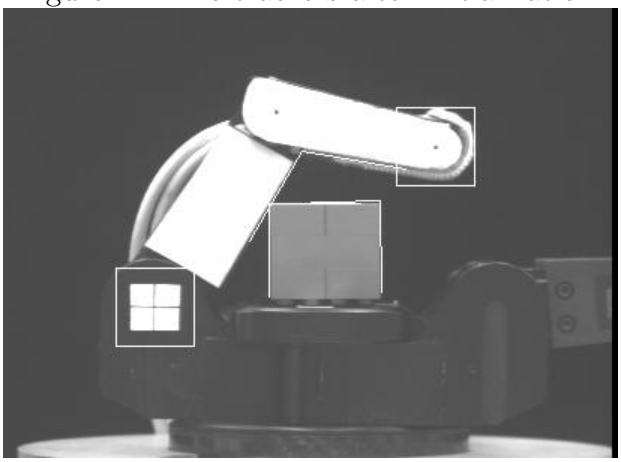

Figure 12: The inner link is stopped, and breakaway occurs.

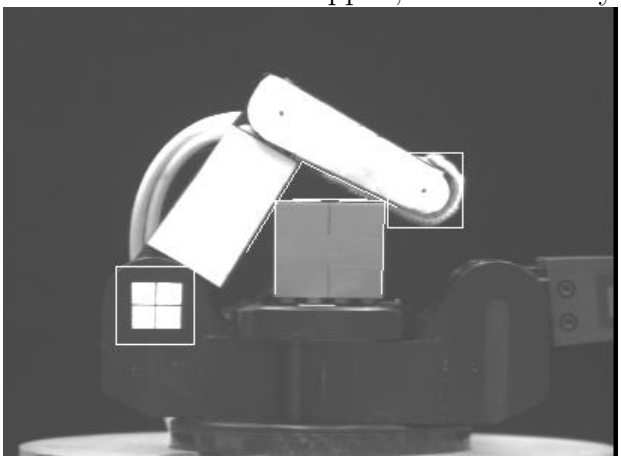

Figure 13: The outer link completes the grasp and the contact points are computed.

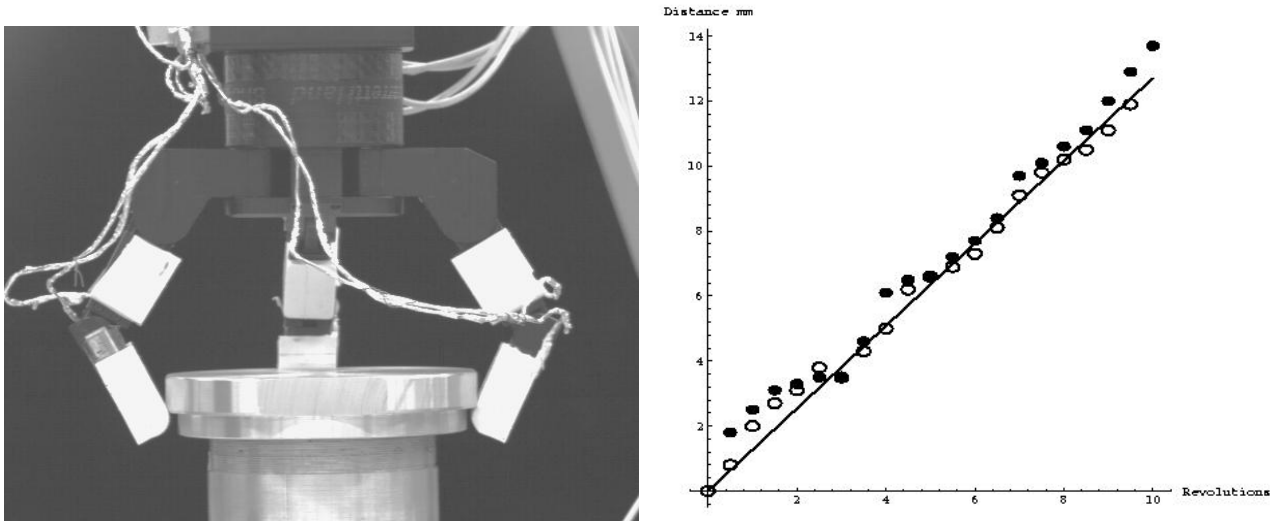

Figure 14: Left: Hand grasping canister top for unscrewing experiment. Right: Vertical displacement of canister top as measured by tactile sensor and kinematic model for finger 2. The bold line is actual screw pitch and the open and closed circles represent the computed start and end vertical displacement for each half revolution.

6.3 Experiment III: Unscrewing a Canister Lid

The purpose of this experiment was to use tacile contacts and our kinematic model to securely grasp the lid of a canister and unscrew it from the base. This requires repeatedly stably 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 14). 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 the total intensity count for each pad. If the total intensity was below the threshold the fingers 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 14 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

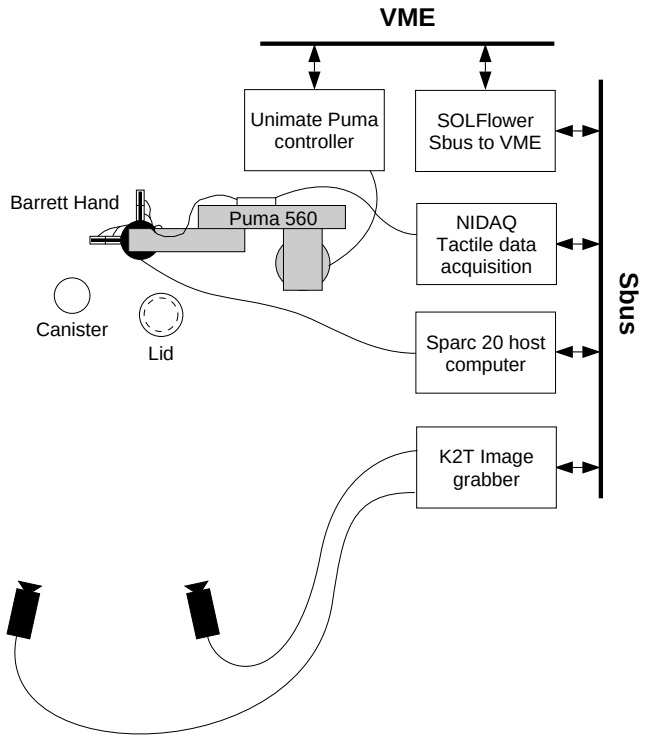

Figure 15: An overhead view of the experimental setup that also shows the components used in our system.

was 11.4 mm.

6.4 Experiment IV: An Integrated Grasping and Manipulation Task

In the previous experiment we verified that the tactile sensors can be used in conjunction with the position encoders and known hand kinematics to make accurate measurements with the fingers. By combining this with an uncalibrated stereo camera setup to control the hand, we can perform the more difficult task of aligning the lid over the canister and screwing it on, where the locations of the objects are variable and the diameters of both objects are unknown at the start. The setup is similar to the last experiment except for the addition of two stationary cameras placed approximately 2.5 meters from the work space and with a baseline of one meter (see figure 15).

We have added circular white fiducial marks to the left and right fingers and to the front surfaces of canister and lid which can be used as visual reference points. At the start of the process, the user initializes centroid trackers on each of these white spots in each image, which the trackers then follow for the remainder of the experiment (see figure 16). After each frame is grabbed, the tracker thresholds a small window of pixels around its center, and computes the new centroid. Since we are tracking eight locations simultaneously (fiducial marks on 2 fingers, 1 mark on the canister, 1 on the lid times 2 cameras) we are using the faster centroid trackers instead of SSD's. After the initialization, the robot enters closed loop control and moves to match the midpoint between the left and right fingers with the mark on the front of the canister using the previously determined uncalibrated positioning matrix. This aligns the fingers with the front plane of the object. Next, the back finger moves until it contacts the back surface of the canister and sufficient pressure is felt on the finger's strain gauge. The computer then reads the tactile sensor on the back finger, and determines the point of contact along the finger.

Using this value and the known kinematics of the hand, the diameter of the canister is calculated (usually within 1mm).

The program then commands the hand to move to a point above the objects, simply to avoid collisions. The location, alignment, and measuring procedure is repeated for the cannister lid. Once this diameter is known, the hand moves back half that distance, and the front fingers close on the lid until it is lifted slightly and the lid's weight is felt on the strain gauges. Then the back finger closes to provide a stable three-fingered grasp. The top is then lifted up, and the robot enters another closed loop to move the top to a position that is 75 pixels up along the projected robot Z axis (determined during the uncalibrated stereo setup). At this point the front faces of the lid and the canister are in the same plane. The hand then moves forward by the difference of the two radii, which is an estimate of where the hand should be positioned to line up the center of the lid with the center of the canister. The robot then moves down slowly until the pressure on the fingers is relieved when the lid comes in contact with the canister (sensed by the strain gauges). A problem that occurs is in aligning the threads of the lid with the canister's threads. We have solved this by slightly biasing the placement of the lid to one side of the canister, and then moving it to the other side with the fingers until resistance is felt on the strain gauges. This occurs when the lid threads engage the canister threads and the fingers register this response. At this point we implement a motion sequence similar to experiment III to screw the lid on the canister. The threads were successfully aligned about 70% of the time. Misalignment generally occurred due to noisy position data from the tactile sensor which resulted in inaccurate diameter measurements. Improving the reliability of the tactile sensors should increase this success rate.

7 Conclusions

In this paper, we have described the design and use of a set of external tactile sensors and external vision modules to extend some of the capabilities of a typical, limited-sensing robotic hand. In the first experiment, vision was used to find the position of contact along a finger in order to estimate applied finger forces. A model of the internal strain gauge force sensing was developed and calibrated, and using this model, we were able to fuse the visual contact position information with the strain gauge values and correctly predict applied finger forces. In the second experiment, vision provided the angular orientation of a finger joint which was used with the hand's kinematic model to estimate fingertip positions and object contacts in space. In addition, tactile sensors were designed, built and used to estimate contact points of the finger on an object. These sensors can report positional contact location and limited forces, and may be used when vision is occluded, as well as confirming any visually determined contact positions. The third experiment showed the use of tactile contact information with forward kinematics to estimate finger positions in space during an unscrewing task. The last experiment used real-time visual feedback to servo the fingers of the hand and monitor task progress for a complete grasping task. This experiment utilized the tactile sensors and internal strain gauge sensing in conjunction with vision to determine finger positions and contacts. Even though we were using an uncalibrated vision system, the combination of vision and touch sensing allowed us to make accurate measurements of the objects to be grasped.

These experiments show in a limited way the utility of providing external sensing to a robotic hand. This sensing may be inexpensive and easy to add to existing hands that lack robust sensing (i.e. the majority of existing hands), and will increase the ability of the hand to operate in a closed loop fashion.

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.

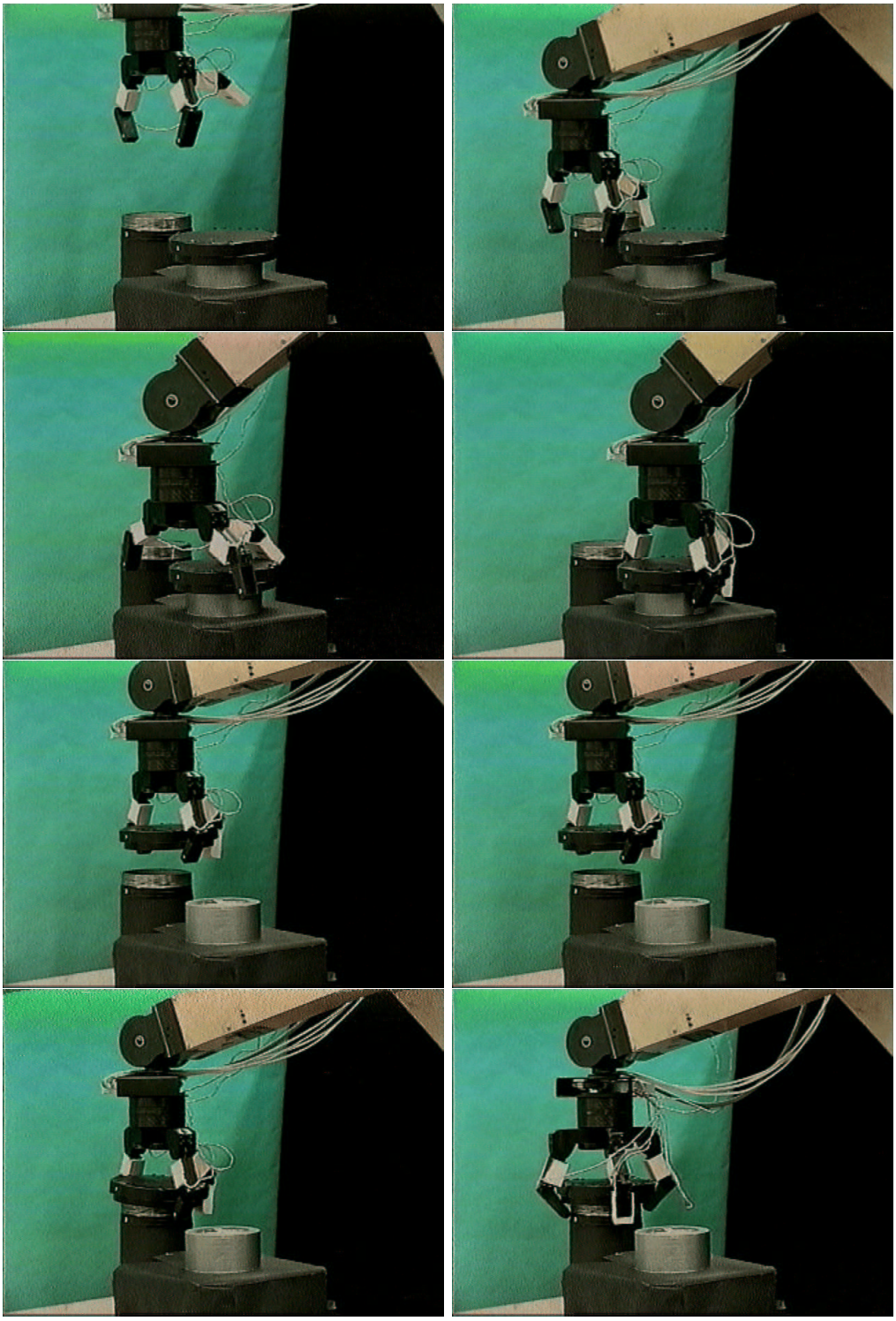

Figure 16: Top to Bottom, Left to Right: A) Hand above canister. B) Fiducial marks on fingers visually servoed to align with canister fiducial mark, with third finger registering tactile contact for diameter measurement. C) Similar procedure for measuring lid diameter. D) Lid is stably grasped with 3 finger grasp. E) Lid fiducial mark is visually servoed to align with canister fiducial mark. F) Centers of lid and canister aligned by moving the lid an amount equal to the half the difference of the measured diameters. G) Lid placed on canister with bias to one side. H) Threads aligned and lid screwed on

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] J. Bentley. *Principles of Measurement Systems*. Longman, New York, 1983.

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

[5] P. Dario. Tactile sensing for robots: present and future. In O. Khatib, J. J. Craig, and T. Lozano-Perez, editors, *The Robotics Review 1*, pages 133–146. M.I.T. Press, 1989.

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

[7] 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.

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

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] M. Xie. Robotic hand-eye coordination: New solutions with uncalibrated stereo cameras. *Machine Vision and Applications*, 10:136–143, 1997.

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

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

[21] 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.