Submitted to IROS 96

# Preliminary Results on Grasping with Vision and Touch

**Jae S. Son and Robert D. Howe**

Harvard University, Division of Applied Sciences
Pierce Hall, Cambridge, Massachusetts, 02138 USA
Email: jae@hrl.harvard.edu

**Jonathan Wang and Gregory D. Hager**

Yale University, Dept. of Computer Science
P.O. Box 208285, New Haven, Conneticut, 06520 USA
Email: hager@cs.yale.edu

## Abstract

*This paper presents initial results in integrating touch with vision for delicate manipulation tasks. A generalizable framework of behavioral primitives for tactile and visual feedback control is proposed. Since vision provides position and shape information at a distance, while tactile provides small-scale geometric and force information, we focus on the complimentary roles of vision and touch. We demonstrate that visual feedback can perform the rough positioning needed for tactile sensor feedback, and that grasp force and object orientation angles can be sensed and controlled with tactile sensing. A force sensor based approach provided a comparison measure, and we observed that the use of tactile sensing results in a much more gentle grasp.*

## 1. Introduction

Human performance of delicate manipulation tasks is governed by a task-dependent combination of visual, kinesthetic and tactile cues. Delicate manipulation in particular, demands that the forces transmitted to the object during grasping be minimized which suggests tactile or force sensing should play a primary role in controlling motion. However, in an unstructured environment, vision is also an essential means of determining an initial location to place the hand, as well as a source of positional feedback once the object is grasped.

We are studying the interaction of vision and touch in the control of manipulation in unstructured environments. In many ways, the integration of tactile and visual sensing in this context is reminiscent of prior work in hybrid control (Raibert and Craig 1981) or LMT-style plan execution (Lozano-Pérez et al. 1984). In hybrid control, position control and force control are combined to exert forces in some directions of motion, while motions are executed in an orthogonal set of directions.

Likewise, vision and touch play largely complimentary roles: vision provides position and shape information at a distance, while touch provides small-scale geometric and force information. Vision is thus particularly effective for free-space motions by quickly locating and recognizing the object, it enables a fast and direct approach before a manipulator contacts an object. Once the fingers make contact with the object, however, touch sensing becomes preeminent, as it avoids the obstruction and resolution limits of vision. Tactile sensing reveals the exact locations of the contacts between the object and the fingers, as well as the object's local surface shape, which permits optimization of the grasp. In addition, touch sensing can be used to minimize contact forces, thus preventing unwanted disturbance of the object before the grasp is secured.

Most previous work combining vision and touch has concerned perceptual issues. One focus has been the use of touch sensing to supplement visual information for object recognition or exploration (Allen 1987, Stansfield 1988, Boshra and Zhang 1995). Other work has developed representations for the integration of visual and tactile spatial information (Rafla and Merat 1990). Rucci and Bajcsy (1995) have examined coordination of active visual and tactile perception, and Ono et al. (1995), have begun to apply these sensing modalities to manipulation by combining vision with a simple form of tactile sensing for handling flexible materials.

In this paper, we discuss the integration of vision and touch at the feedback level for delicate manipulation tasks. We begin by outlining the conceptual framework for combining visual and tactile feedback in control of manipulation, and describe the visual and tactile primitives used in this paper. The subsequent section describes the experimental hardware and task configuration. Experimental results are then presented, and performance is evaluated in terms of net forces applied to the object. We conclude with a discussion of the results, our plan for future experiments, and system integration issues.

## 2. Combining Vision and Touch

Effective combination of vision and touch requires consideration of many factors, especially the capabilities and limitations of each sensing modality and the task context. As noted above, the unique strength of vision is its ability to sense motion over a large space. The fundamental measures are distances and directions, and resolution is limited to a fraction of the field of view. Touch, particularly using tactile arrays, senses local geometry and pressure, providing both length and force measurements (Son et al. 1996). Sensing takes place at contact locations that are hidden from view by the hand, and in some situations may provide better spatial resolution than vision.

These considerations suggest that a variety of combined sensing modalities will prove useful. In some situations both visual and tactile information provide spatial information, which can be "fused" to enhance resolution. One example is finding finger tip contact locations for grasping a complex object: vision can provide global shape information at limited resolution, while tactile sensing can measure local curvature and surface normals. Finger placement can be optimized by using both measurements to ensure full kinematic restraint while minimizing the chance of slips at each finger tip.

In other situations, the sensed information will be complementary, and the robot controller may be configured to use the best sensing modality. The experiments described below involve grasping a target object at an arbitrary location in the workspace with a two-fingered gripper using minimal forces. At the beginning of the task, vision provides the only source of information, and the robot must be visually servoed. When the fingers of the gripper make contact with the object, contact location and force information are provided only by touch, and the controller must use tactile feedback to minimize forces and servo the contact location.

In this paper, we focus on the latter situation, with complimentary visual and tactile information. Our approach is based on sets of visual and tactile "primitives," which are basic sensing and control modes appropriate to various segments of a task. These primitives are then combined, either serially or in parallel, for task execution. In the following sections we describe these vision and touch primitives, and our framework for their integration.

## 2.1 Vision Primitives

Hager (1995a), proposes a conceptual framework for static hand-eye coordination tasks. The central notion in this framework is that of a task-space kinematic constraint. Any robot task is specified in terms of the kinematic constraints that should be achieved between the robot and a target object. For example, if the goal is to place a screwdriver onto a screw, the desired kinematic constraints are the alignment of the axis of the screwdriver with the axis of the screw, and the tip of the screwdriver touching the head of the screw. Notice that these constraints only determine 5 degrees of freedom of motion – they do not specify the orientation of the screwdriver about its axis.

The central idea in all of our visual feedback primitives is to translate task-space kinematic constraints into sensor-level constraints. Given sensor-level constraints, an image-based error function is defined such that the error function is zero if, and only if, the task-space constraints are satisfied. Standard linearization techniques and PID control methods are then used to drive the manipulator into the desired configuration. Details on this process can be found in (Hager 1995b). Throughout this paper, we treat the robot as a velocity-controlled device which acts as a pure integrator. This approach neglects many potentially important dynamic effects, but for low-velocity tasks it has been experimentally validated as a reasonable model of the system (Hager et al. 1995). This permits us to treat the control problem as one of generating velocity commands from sensor data.

Here, we review the essential ideas behind the theory using simple point-to-point motion. Let and be points in space attached to the robot manipulator and a target object, respectively, expressed in robot base coordinates. These points are imaged into a camera using the usual perspective map which we denote $\pi : \Re^3 \rightarrow \Re^2$. Suppose $r$ and $t$ are the projections of these points in the camera image, and the robot translates with some velocity $V$. Then the projection of $R$ will move with some velocity $\upsilon$. The relationship between these velocities is given by the so-called *image Jacobian* $J = \partial\pi / \partial R$. Suppose now that we have two cameras observing $R$ and $T$. It is easy to show that $R$ is coincident with $T$ if and only if their projections are coincident in both cameras. Let $r_i$, $t_i$, and $J_i$ denote the projections of $R$ and $T$ and the image Jacobian for cameras $i = 1, 2$. Define

$$e = \begin{bmatrix} r_1 - t_1 \\ r_2 - t_2 \end{bmatrix}, \text{ and } J = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix}. \qquad (1)$$

Let $J^+$ denote the generalized inverse of a matrix. It follows that applying feedback

$$u = -J^+ e \qquad (2)$$

to the robot will stabilize it so that $R \rightarrow T$, as $time \rightarrow \infty$ (Hager et al. 1995). In general, a manipulator is controlled by a velocity screw $s$ which controls six degrees of freedom of motion. Thus, visual kinematic constraints up to and including six degrees of freedom can be specified and carried out by this type of feedback arrangement. Note that the Jacobian will generally be $m \times 6$, where $m$ is the degree of the visual constraint.

## 2.2 Tactile Primitives

Tactile sensing can serve two functions in manipulation task execution. The first is to provide feedback for continuous control of the manipulator, as does vision in the primitives described above. The second is to detect events, the discrete transitions between mechanical states of the hand-object system. Common events include making or breaking contact between the fingers and the object, and the onset of the fingers sliding against the object's surface. Tactile sensing is uniquely suited to this type of event detection, and a number of workers have used tactile sensing to increase the robustness of task programming and execution (e.g., Cutkosky and Hyde 1993, Brock 1993). In the following, we describe the simple control and event detection primitives used in the grasping task, which are representative of the range of tactile-based behaviors that can be constructed.

A vast number of tactile sensing devices have been presented in the literature, and many could perform the functions described below. We have elected to use tactile array sensors, which consist of grids of pressure sensors located in the grasping surface of the fingers. The pressure distribution signal from these sensors directly provides contact location and contact force information. With more elaborate signal processing, many other parameters may be extracted, including object curvature (Fearing 1990b). Intrinsic contact sensing, which uses multi-axis force and torque sensors, could provide similar information about contact location and forces. A detailed comparison of the performance of both sensing schemes can be found in (Son et al. 1996).

### 2.2.1 Make Contact

Sensing first contact between the hand and an object is essential to minimizing forces on the object. The wrist force sensor's sensitivity for detecting small forces is limited due to dynamic range issues, and it is affected by gear noise due to the gripper mass in front of the sensor. The array sensor response is slow because of the multiplexing required, but it can be used to determine many parameters including contact. A more suitable sensor for detecting contact, such as the stress rate sensor (Son et al. 1994), could provide more sensitive and timely response.

This primitive begins with the hand approximately centered on the object by visual control, and the fingers are commanded to close. When the sum of the readings from all elements of an array exceeds a threshold, the gripper is commanded to stop. Although it is possible for contact to occur at both fingers simultaneously, we assume that contact occurs at one finger first.

### 2.2.2 Reorient Gripper

This primitive senses the object's precise orientation w.r.t. the gripper and reorients the hand to permit the desired grasp. Determining object pose has been an important area of research for tactile sensing. Fearing (1990a) has analyzed this problem in the general 3–D case using contact ellipsoid models, and Chen et al. (1995) have determined orientations using moment area methods. Here we focus on the rotational axis parallel to the cylindrical sensor axis, where the contact centroid is directly proportional to the object orientation angle.

The primitive described here is specialized to the one degree-of-freedom two-fingered cylindrical gripper used in the experiments, although it illustrates the issues in the general case. The primitive starts with one finger in contact with the object. From a cylindrical sensor contacting a parallel flat sided object, we can determine the orientation of the object (Son and Howe 1996). In the experimental configuration, the object orientation angle relative to the finger is simply $\phi = \alpha\rho$, where $\alpha$ is the contact pressure centroid location on the sensor surface and $\rho$ is the sensor radius. In this task, the goal is to achieve contacts at symmetric locations on opposite sides of the object, which will minimize the net applied torque and maximize grasp stability. This pertains when $\phi = 0$.

Once the object's orientation has been determined, there are many ways to reorient the hand to ensure that the fingers reach the desired final contact locations. With a multijointed, multifingered hand, it may be possible to move only the finger joints to achieve the goal configuration. Since our gripper has one degree-of-freedom, the joints of the robot arm must also be used. As a start, we use a regrasping approach and rotate the gripper through an angle opposite the measured object orientation angle. Because the manipulator rotates about a tool coordinate frame located at the midpoint between the gripper finger tips, the sensor temporarily breaks contact with the object.

Another approach for compensating object orientation angle involves directly servoing the contact force, while at the same time, servoing the contact centroid to the center of the sensor which corresponds to . This method would maintain contact with the object, but introduces sliding between the object and the sensor since the rotation is about the tool frame. The advantage of this scheme, which uses a feedback loop to maintain contact, is that if the object moves during the grasping process, the controller can compensate for these changes. The disadvantage of this process is that sliding can introduce disturbances, and sliding repeatability will depend on friction, which can vary significantly.

Alternatively, to avoid sliding and breaking contact during reorientation, the robot can be commanded to roll the finger tip against the object, by rotating about the instantaneous contact centroid location. A PID controller can set the magnitude of the velocity, which is directed to bring $\phi \rightarrow 0$. This requires a coordinate transformation to transform the pure rotation from the contact centroid frame to tool frame coordinates. The transformation can be calculated from the contact location on the sensor and gripper opening position. This transformation and $\phi$ must be updated from tactile data as the contact location rolls over the finger tip.

### 2.2.3 Make Two Finger Contact

After the correct orientation is obtained, the gripper must close to bring both fingers into contact. This again requires coordinated motion of the arm and gripper to avoid applying large forces at the finger in contact. Tactile array information can be used in the feedback loop to minimize the disturbances during the grasping process. The pressure sensors' readings are summed to find, $F_i^{actual}$, the total contact force on the finger tip $i$, and these are used in a simple proportional gain control law:

$$v_i = K_i(F_i^{desired} - F_i^{actual}) \tag{3}$$

Transforming these velocities into tool frame and gripper closing velocities,

$$v_{tool} = (v_1 + v_2)/2, \text{ and } v_{gripper} = (v_1 - v_2)/2. \quad (4)$$

This tactile sensing primitive could be used directly to acquire objects if compensating for object orientation is not important.

### 2.2.4 Increase grasp force

Once stable two-finger contact has been made, we can use the tactile force information again to control the grasping force. Here, instead of commanding velocities at each finger to generate a desired grasp force, we use one gripper velocity. One problem with using the "make two finger contact" primitive here is that any sensor gain differences will generate tool frame velocities which will cause the gripper to drift. Therefore, the controller sends gripper velocity commands based on the average grip force measured.

$$v_{gripper} = K(F^{desired} - F^{average}) \quad (5)$$

## 2.3 Integration for task control

Although tactile sensing is force-based, it is often used to attain or improve the kinematic relationship between the robot manipulator and a target. Hence, tactile feedback can also be conceptually modeled as a means to achieve a predetermined set of kinematic constraints. The immediate issue that arises is how vision and tactile feedback should be combined. Here, we propose and discuss two simple and natural combinations which are *serial combination* and *prioritized parallel combination.*

Serial combination can be described as "vision then touch." This is a useful mode of interaction when vision can be used to roughly place the robot end-effector near the desired goal configuration, and then tactile sensing is used to perform a final fine positioning. In order for this operation to be guaranteed to perform correctly, two conditions must hold: (1) there is a set of initial robot configurations such that tactile sensing can be guaranteed to achieve the desired constraint from any $q \in I$, and (2) the visual constraints are such that it can be guaranteed that the system will move to into a configuration $q_v \in I$. These concepts closely parallel the basic ideas underlying the LMT-style planning framework (Lozano-Pérez et al. 1984). This is not surprising, as we are using visual and tactile constraints in a manner reminiscent of the contact constraints in LMT-style planning. This is also closely related to the "funneling" idea proposed by Burridge et al. (1995) for composition of dynamic behaviors. The advantage of serial composition is that the two modalities operating independently makes their combination extremely simple. The disadvantage is that it is often the case that tactile sensing is not rich enough to guarantee that all task constraints can be met. Hence, it is quite possible that feedback from tactile sensing may undo constraints that have been achieved by visual sensing.

Prioritized parallel composition provides for more direct combination of tactile and visual feedback by superimposing feedback from one modality onto degrees of freedom which the other modality does not constrain. For example, suppose that the manipulator is performing point-to-point motion as describe above. Then the image Jacobian will be $3\times6$. Suppose that $s_t$ is a motion commanded by tactile sensing. If the goal is to maintain the visual constraint, but superimpose tactile-based motion, then a feedback of the following form can be used:

$$s = -J^{\tau} e + (I - J^{\tau} J) s_t \quad (6)$$

where as before $J$ is the image Jacobian for the visual constraint, and $e$ is the image error for that constraint. The term on the right projects the tactile motion into the kernel of the image Jacobian. Physically, this causes the robot to move under tactile control only in directions which are uncontrolled by visual sensing. Note that to the extent which tactile sensing can be written as a feedback system of this form, it is also possible to give tactile priority over visual feedback.

Note that a simple case of this feedback law is *orthogonal combination* in which tactile and visual feedback control independent degrees of freedom. For example, the vision system positions the gripper using a point-to-point primitive, and the tactile sensing system is used to orient the gripper.

# 3. Experimental Methods

## 3.1 Hardware

The entire system architecture required the use of four separate processing platforms: dedicated motor controllers, host PC for the Robot, another PC for the tactile signal processing, and a workstation for the vision signal processing. At the lowest level, each joint of the Zebra-Zero manipulator is controlled by a dedicated motor control IC. These controllers can be configured for velocity or position control which leaves the host computer PC to focus on coordination issues such as

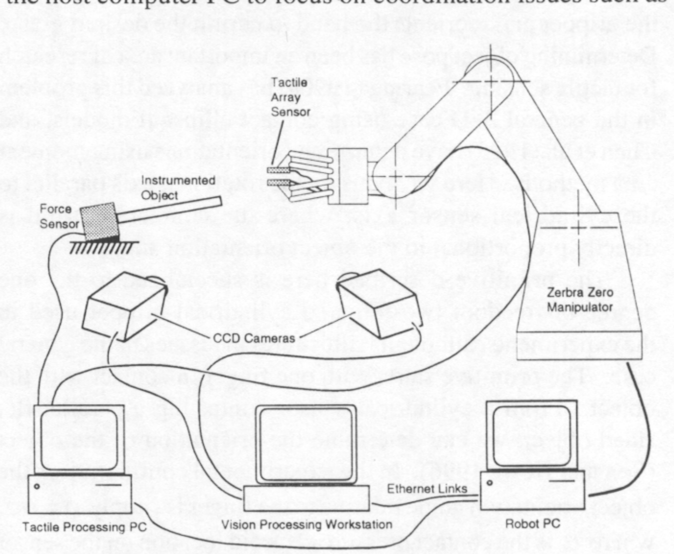


Figure 1. Schematic of the experimental setup.

calculating the kinematics, Jacobian, coordinate transformations, etc. The signal processing computers send tool frame velocities to the robot host PC via ethernet connections. Figure 1 shows a schematic of the experimental setup.

The advantage of this type of distributed system is that it is possible to segment the problems into manageable pieces such as vision, tactile, kinematics, and motor control. This allows independent process development and simplifies debugging the entire system. In addition, computational power is increased through the use of multiple processors. The disadvantage is that communication between the subsystems is limited. In our implementation, we are only able to send desired velocity commands to the host PC.

The vision system consists of two Sony XC-77 CCD cameras with 12.5 mm lenses connected to a Sun 10/42 workstation via two Imaging Technologies FG101 framegrabbers. Visual tracking for both cameras and visual feedback calculations are performed on the Sun host using the XVision tracking system (Hager 1995b) and the Servomatic hand-eye coordination system (Hager 1995a).

In the experiments reported here, initial positioning of the gripper is performed using the point-to-point motion primitive reported above. The tracking system follows the motion of the two fingers of the gripper using a simple pattern matching method, which tracks by locally optimizing the sum of the squared differences between an area of the live image and a stored reference image (Hager et al. 1995). A setpoint midway between the fingers is computed in each camera, and then the distance from this "virtual" feature to the object endpoint (which is also tracked) is used to implement point-to-point positioning.

The tactile sensors used in the experiment are capacitive tactile array sensors. Each sensor containing 64 elements in an 8 x 8 matrix equally spaced at 2 mm. The entire sensor is encapsulated in silicone rubber to protect the sensor and increase friction and compliance. The gain of each element was calibrated by applying uniform pressure to the sensor surface. Details of the sensor design can be found in (Son et al. 1995). The sensors are mounted on cylindrical gripping surfaces which permits better contact localization. Figure 2 shows a close-up of the tactile array sensors mounted on the gripper.

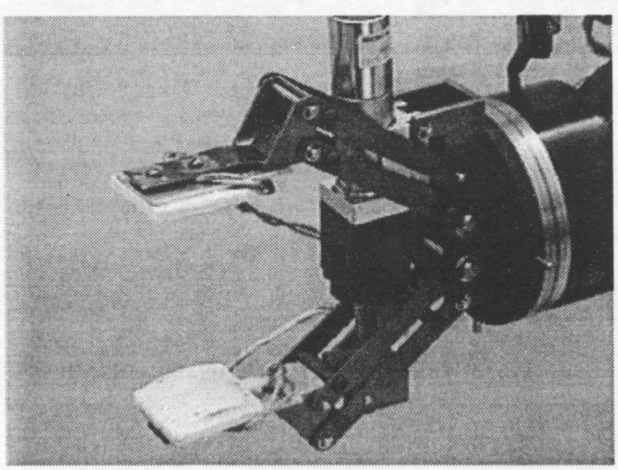

Figure 2. Tactile array sensors mounted on the gripper.

The target object in these experiments was a 10 mm diameter aluminum tube. This tube was mounted on a force-torque sensor that measured two force components and one torque component in the plane of the gripper and object. This sensor was mounted on an aluminum frame which was securely clamped in place.

## 3.2 Experimental Procedure

The experimental task required the robot to gently grasp the object, which was fixed within the visual field of view and the manipulator workspace. To simulate an unstructured environment, the controller had no a priori information about object position or orientation. Task execution followed the order of primitives described in Section 2: once the vision system moved the gripper into position, tactile feedback minimized disturbances while grasping the object. Although numerous methods for combining vision with tactile sensing could be used, our initial efforts were focused on integrating the entire system to complete the task successfully.

We elected to use minimization of forces and torques applied to the object to evaluate performance of the system. Alternatively, displacement of the object could have been measured if the object had been compliantly mounted. Minimization of disturbance forces, however, is a realistic constraint in handling delicate objects. It also emphasizes the differences between vision and touch, since vision cannot provide force information to the controller. Tactile sensing is used to detect contact, compensate for the difference between the orientation of the object and the gripper, make contact on both fingers, and finally increase the grasp force to the desired value. Not compensating for the orientation angle would impose a torque on the object through the difference in the contact locations where the grasp force was applied.

As a comparison benchmark, we also executed the task using an alternative technique to minimize disturbance forces. One of the built-in force control modes of the Zebra-Zero Robot is "sfloat" or "servo float," which controls the robot joints to minimize the forces and torques measured at the wrist force sensor. In this mode, during the grasping operation the robot attempts to compensate for the orientation angle difference and the translational errors of the object through sensed forces at the wrist. This wrist-based compliant motion scheme provides a simple and immediate means of verifying the benefits of tactile sensing.

# 4. Experimental Results

As noted above, our preliminary experiments have focused on integrating the tactile and visual system, and developing a set of tactile feedback strategies. To this end, we have chosen situations where vision and tactile sensing perform independently. Thus, in our experiments, we use visual feedback to perform rough positioning of the hand, and then switch to tactile sensing to perform hand reorientation and repositioning during closing. Our goals are to demonstrate that vision is sufficiently accurate to provide an initial set of conditions in which tactile feedback operates correctly, and to determine the effectiveness of direct tactile feedback in minimizing forces.

The top three images in figure 3 show how the vision system is able to move the gripper allowing it to grasp the end of a rod. Feedback computed from the stereo camera images provides sufficient information for a complete three-axis positioning relative to the tip of the rod. The object orientation angle could in principle also be controlled by visual feedback, but we avoided doing so to study the capabilities of the tactile system. Once the visual feedback to the robot endpoint dropped below 0.5 mm/sec in all axes, we switched to tactile feedback and invoked the strategies described above.

The bottom three images in figure 3 show how the tactile array sensor is able to detect contact, and compensate for object orientation. Figure 4 shows measurements from the array sensor during each of these phases. Without tactile sensing or finger tip force sensing, the grasp force on the object is difficult to obtain. Some crude measure of the grasp force can be estimated though the motor current, but errors associated with this approach are large due to friction. Figure 5 shows the total grasp force computed by adding all the array elements. By using this information, the grasp force can be controlled more precisely.

Finally, we look at the forces imparted to the object with tactile sensing and compare it to servo float mode. Figure 6 shows that with tactile sensing, the contact forces and torques are reduced considerably despite the fact that the sensitivity of the tactile array sensors used was low due to the extra long cables required for the experiment (the sensor gain is approximately a function of sensor capacitance divided by the line capacitance as explained in (Fearing 1990a)). The final orientation angle reached using tactile feedback was also observed to be much closer to the desired value than that achieved by "servo float" mode. The torque generated by the small differences in the contact locations of the gripper was not sufficient to overcome the friction and completely reorient the gripper.

## 5. Conclusions and Discussion

We have proposed a generalizable framework for touch and vision in terms of behavioral primitives which can be combined in a variety of ways. In our first experiments combining tactile with vision, we have used a sequential combination of the two. These experiments demonstrate two main points: (1) visual feedback is sufficiently precise to perform the rough positioning needed to invoke tactile sensor feedback, and (2) tactile feedback can compensate for uncertainties in object orientation angles and control grasp force. We observed that the use of tactile sensing results in a

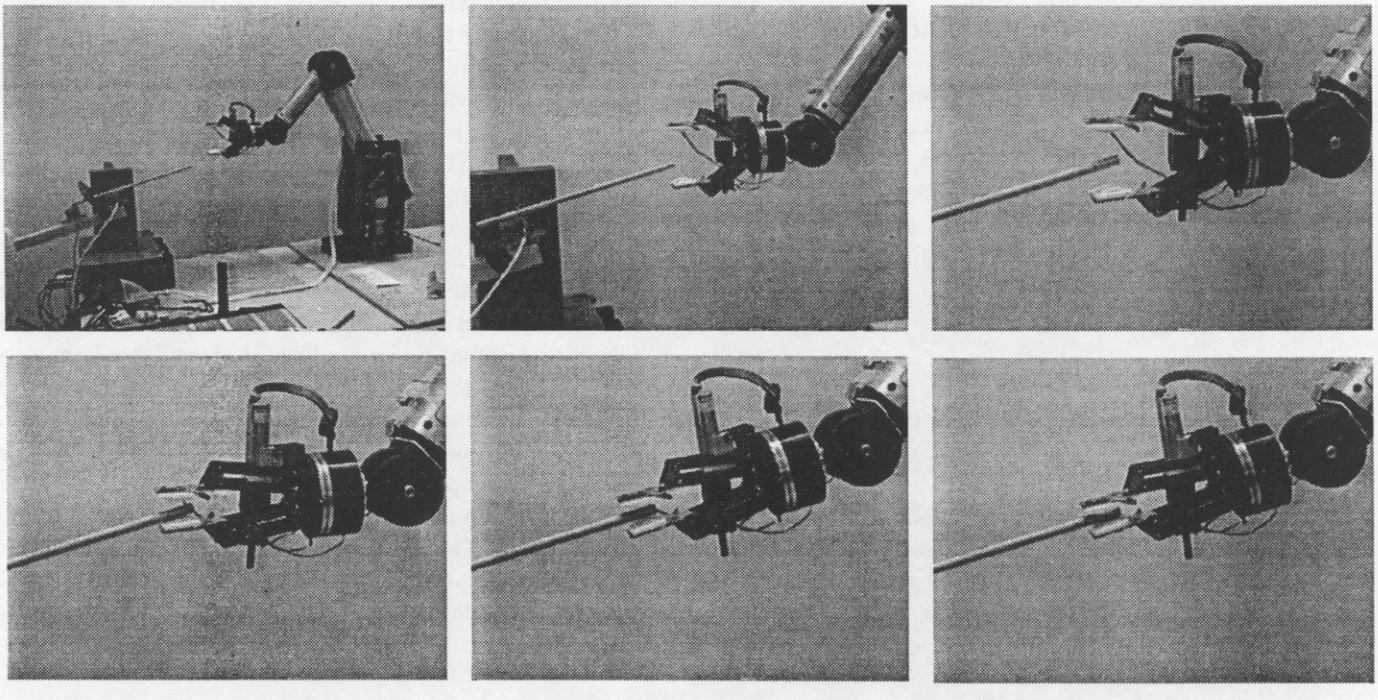

Figure 3. In the top three photos, the vision system brings the gripper to the object using a point-to-point primitive. Later, the tactile sensors are used to grasp the object by detecting contact, compensating for orientation angle, and controlling the grasp force.

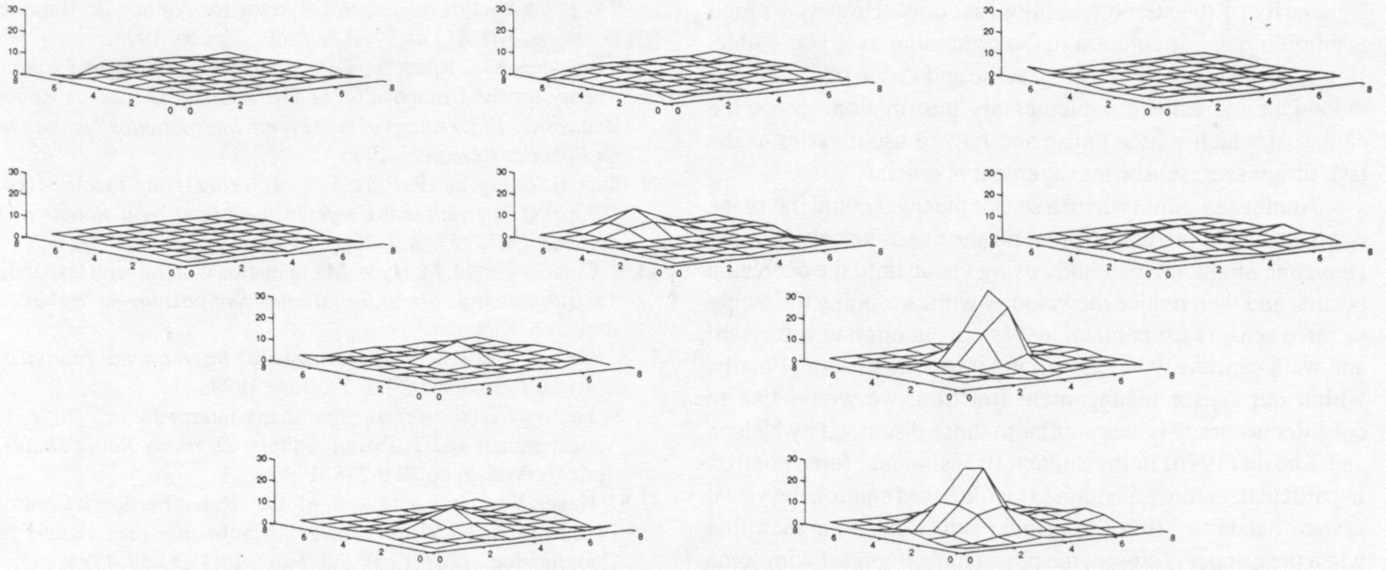

Figure 4. Tactile array sensor response during the grasping process. The response is flat until contact occurs at the edge of the bottom sensor. As the manipulator compensates for this orientation and starts closing the gripper, only the bottom sensor is in contact. Contact forces are minimized until both sensors make contact, where upon the grasp force is increased to the desired value. The vertical axis units are in psi.

much more gentle grasp in comparison to a wrist force sensor based approach.

The tactile sensing scheme also exhibited much larger time delays than the force sensor. This is due to several factors, most notably the time required to scan and process the array information, the time required to compute the velocity commands, and the time consumed in transmitting the commands to the host PC. Despite these limiting factors, all of which worked unfavorably for the tactile array sensor, the array sensor is able to improve the performance of the grasping process.

Finally, it is interesting to note that one of the concerns before performing the experiment was manipulator controllability. In the force control literature, it is well known that transition to contact and maintaining contact both require fast, high bandwidth feedback. For tactile sensing, it was unclear how much positional accuracy was available, and how much accuracy was needed. Backlash and manipulator stiffness were also thought to be issues. A very stiff manipulator could generate high forces, which could damage

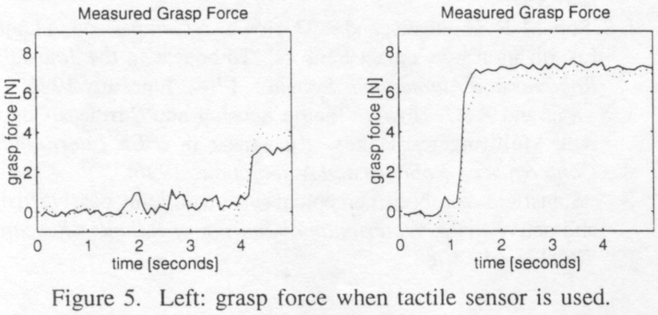


Figure 5. Left: grasp force when tactile sensor is used. Right: grasp force without tactile sensing.

the sensor. However, none of these concerns proved to be problems during the experimental execution. The tactile sensor combined with the compliance of the manipulator and the ability to actively control grasp forces produced a reliable and robust feedback mechanism. In this sense, the tactile based procedures proved to be quite complementary to the visual feedback strategies which are also robust to these phenomena.

# 6. Future Work

Our next goal is to demonstrate a closer coupling between vision and touch, and to develop an intelligent *sensor manager* which controls the character and interaction of sensor feedback as a task develops. In particular, we plan to perform experiments where it would not be possible for either vision or touch to achieve the task alone. One example would be grasping the same object, where the unknown orientation is in the plane defined by the camera optical axes. In this configuration, visual control fails because this case is a

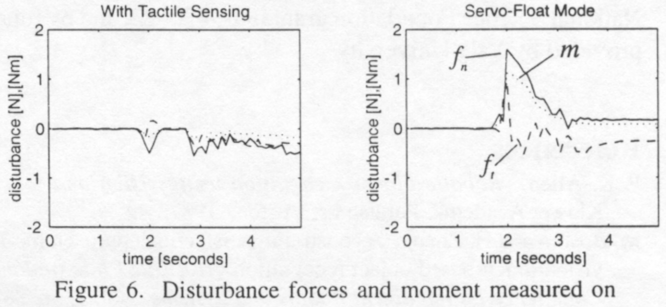


Figure 6. Disturbance forces and moment measured on the instrumented object. Left: tactile array sensor is used to minimize disturbances. Right: wrist force sensor is used to minimize disturbances.

singularity of the stereo mapping function. However, rough positioning and orientation in the remaining axes is possible. Hence, this is a situation where tactile and vision provide some redundant and some complementary information. Since the choice of which sensor to use and how to use it varies as the task progresses, sensor management is crucial.

Another example where a sensor manager could intervene is when visual obstruction prevents the use of visual servoing. Here, one might move quickly using vision until the occlusion occurs, and then reduce the velocity while scanning the tactile or force sensors for contact, just as we put our hands forward and walk cautiously when walking into a dark room. Finally, within our sensor management function, we would like to consider accuracy issues similar to those discussed by Nelson and Khosla (1996) in the context of vision and force control. In particular, before transition to tactile based control, the visual system needs to bring the gripper into a position such that when the gripper is closed, the object makes contact with some active portion of the array sensor. This establishes the visual servoing accuracy requirements. One interesting aspect to explore in this vein is to move to an active camera so that visual servoing accuracy can be improved at certain points in the task execution.

There are several architectural issues that we are also planning to resolve in future experiments. In order to reduce tactile feedback delays, we are developing a micro-controller interface so that array sensor values can be quickly read from memory at a convenient time for the interface PC. Also, our current network architecture only provides for one-way information flow from the various sensor processors to the robot host computer. We believe that a more reasonable architecture is to use one management computer interface that concentrates feedback from all three subsystems.

## Acknowledgments

J. Son and R. Howe were supported by the Office of Naval Research under ONR Grant No. N00014-92-J-1814. Support for J. Son was also provided by a Doctoral Fellowship from GM Hughes Electronics Company.

J. Wang and G. Hager were supported by ARPA grant N00014-93-1-1235, Army DURIP grant DAAH04-95-1-0058, National Science Foundation grant IRI-9420982, and by funds provided by Yale University.

## References

P. K. Allen, *Robotic object recognition using vision and touch.* Kluwer Academic Publishers, Boston. 1987.

M. Boshra and H. Zhang, A constraint-satisfaction approach for 3D vision/touch-based object recognition. *IEEE/RSJ International Conference on Intelligent Robots and Systems.* vol. 3, pp. 368-73. 1995.

D. Brock. A Task-level Control System for Automatic Robotic Grasping. *MIT AI Lab Number 1469.* January 1993.

R. R. Burridge, A. A. Rizzi, and D. E. Koditschek. Toward A Systems Theory for the Composition of Dynamically Dexterous Robot Behaviors. *Proceedings of the Seventh International Symposium on Robotics Research.* 1995.

N. Chen, H. Zhang, and R. Rink. Edge Tracking Using Tactile Servo. *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS '95.* vol. 2, pp. 84-89, 1995.

M. R. Cutkosky and J. M. Hyde. Manipulation Control with Dynamic Tactile Sensing. *6th International Symposium on Robotics Research.* October 1993.

R. S. Fearing. Tactile sensing mechanisms. *International Journal of Robotics Research*, 9(3):3–23, June 1990a.

R. S. Fearing. Tactile sensing for shape interpretation. In S. T. Venkataraman and T. Iberall, editors, *Dextrous Robot Hands.* Springer-Verlag, pp. 209-238, 1990b.

G. D. Hager, W. Chang, and A. S. Morse. Robot Feedback Control Based on Stereo Vision: Towards Calibration-Free Hand-Eye Coordination. *IEEE Cont. Sys. Mag.* 15(1):30-39, 1995.

G. D. Hager. A Modular System for Robust Hand-Eye Coordination Using Feedback from Stereo Vision. To appear in the *IEEE Transactions on Robotics and Automation,* submitted Jan. 1995a.

G. D. Hager. The "XVision" System: A General Purpose Substrate for Real-Time Vision-Based Robotics. *Proceedings of the Workshop on Vision for Robotics.* pp. 56-63, 1995b.

T. Lozano-Pérez, M. T. Mason, and R. H. Taylor. Automatic Synthesis of Fine-Motion Strategies for Robots. *International Journal of Robotics Research.* 3(1):3-24, 1984.

B. J. Nelson and P. K. Khosla. Force and Vision Resolvability for Assimilating Disparate Sensory Feedback, To Appear in *IEEE Trans. Robot. Automation.* Sept. 1996.

E. Ono, N. Kita, and S. Sakane. Strategy for unfolding a fabric piece by cooperative sensing of touch and vision. *IEEE/RSJ International Conference on Intelligent Robots and Systems.* vol. 3, pp. 441-5. 1995.

N. I. Rafla and F. L. Merat. Vision-taction integration for surface representation. *IEEE International Conference on Systems Engineering.* pp. 519-22. 1990.

M. Raibert and J. Craig. Hybrid position/force control of manipulators. *ASME J. of Dyn. Sys., Measurement, and Control.* 103(2):126-133. 1981.

M. Rucci and R. Bajcsy. Learning visuo-tactile coordination in robotic systems. *IEEE International Conference on Robotics and Automation.* vol. 3, pp. 2678-83. 1995.

J. S. Son, E. A. Monteverde, and R. Howe. A Tactile Sensor for Localizing Transient Events in Manipulation, *IEEE International Conference on Robotics and Automation.* vol. 1, pp. 471-476. May 1994.

J. S. Son, M. R. Cutkosky, and R. D. Howe. A Comparison of Object Localization by Contact Sensors. To appear in the *Journal of Robotics and Autonomous Systems.* 17(4), June/July 1996.

J. S. Son and R. D. Howe. Tactile Sensing and Stiffness Control with Multifingered Hands. To appear in *IEEE International Conference on Robotics and Automation.* 1996.

S. A. Stansfield, A robotic perceptual system utilizing passive vision and active touch. *International Journal of Robotics Research.* 7(6):138-61. 1988.