COGNITION AND BRAIN THEORY, 1984, 7(2), 199-214

Active Touch and Robot Perception

Kenneth Y. Goldberg
Ruzena Bajcsy
Department of CIS, University of Pennsylvania

ABSTRACT

Psychologists distinguish between *active* and *passive* touch. The latter arises when objects are brought into contact with a passive tactile surface, such as the palm of the hand. Active touch describes a dynamic exploration of objects involving receptors located in both the skin (cutaneous) and the joints (kinesthetic). Research in the area of robotic tactile perception has focussed on passive touch, developing cutaneous grids with increasingly improved resolution. A robot developed at the University of Pennsylvania, however, suggests that the most efficient way to achieve tactile recognition is to process kinesthetic information gained from active exploration. The results may be of interest to researchers in both psychology and robotics.

INTRODUCTION

One of the primary tools which humans use for object recognition is the sense of touch. Although investigations have been conducted by psychologists since William James (James, 1890), a comprehensive theory of touch has yet to be formulated. Working at the University of Pennsylvania, we have produced a robot which models tactile form perception and may provide insight into this process.

The psychological and physiological literature acknowledges that human tactile perception is achieved through the interaction of two sensory networks (Gordon, 1978). The cutaneous network includes sensors located in the skin which respond to contact and texture. The kinesthetic network includes sensors located in the joints and muscles which monitor limb position. These two networks are often combined under the term *haptic perception*, although tactile perception is commonly used as a synonym (Loomis & Lederman, 1983).

Requests for reprints should be sent to Professor Kenneth Y. Goldberg, Computer Science Department, Carnegie-Mellon University, Pittsburgh, PA 15213.

The role played by each network has been a subject of debate since J. J. Gibson divided touch into two categories: active (touching) and passive (being touched) (Gibson, 1962). During active touch, both networks interact as the hand moves over an object, responses from the skin guiding motions of the arm and fingers (i.e., an afferent-efferent feedback loop). In passive touch, the hand is held motionless while an object is pressed into it. Gibson found that object recognition was drastically impaired under the latter conditions (49% correct vs. 98% for active touch), suggesting that cutaneous information alone is not rich enough to support object recognition.

Other studies (Lawrence, 1978; Derevensky, 1979; Heller, 1980) support the conclusion that active touch is superior to passive touch for object recognition. In one experiment, the cutaneous network was removed by anesthetizing the epidermis. Subjects were still capable of successfully recognizing objects using only kinesthetic receptors (Katz, 1925).

While the last result lends support to the argument that the kinesthetic network plays a dominant role in object recognition, this conclusion has been attacked on the grounds that these experiments did not effectively isolate the networks, that different regions of skin were tested in Gibson's experiment and that subsurface cutaneous receptors were still active in Katz'. Still needed, then, is a model for the interaction between the cutaneous and kinesthetic networks.

Interest in this subject has recently come from a different quarter. In the field of robotics, where object manipulation is of primary importance, the tactile sense is being reexamined. Robotics researchers have primarily focussed on the cutaneous sense, developing grids of sensors with increasingly improved resolution (Sato, 1977; Briot, 1979; Purbrick, 1981). Techniques borrowed from robot vision have traditionally been applied to processing such grids, although when compared to automated vision sensors, current tactile sensors are crude and inaccurate. Cutaneous systems have been successful with 2-dimensional objects such as washers and cotter pins (Hillis, 1982)[1], but they have not produced satisfactory results with objects in the so-called "kitchen" domain, which range in size from a thimble to a breadbox (Goodenow, Abramowitz, & Paul, 1983).

Such multi-faceted objects must be explored from several perspectives. It is possible to press objects against a cutaneous grid, generating grey-scale "images" analogous to pixel arrays (sometimes called *texel* arrays). In order to effectively discriminate between objects, the texel array must be wider than the object under study; either the grid must be large (Briot, 1979) or the object must be small (Hillis, 1982). Alternatively, a small grid can be stepped

[1]In his paper, Hillis uses the term "active touch" to denote a top-down, or hypothesis-driven approach to recognition in contrast to a bottom-up, or data-driven strategy. The approach presented in this paper is data-driven but active in that it makes use of kinesthetic inputs.

across an object and the series of images pieced together to build a large image. In all of these methods, as in vision, the task of reconstructing the 3D source from planar images is highly dependent on the spatial resolution of the grid.

Here Gibson's distinction between active and passive touch becomes relevant. Processing a set of cutaneous images is analogous to passive touch. Object characteristics must be determined solely from impressions upon the palm or cutaneous grid. Like Gibson's subjects, robotics researchers have found this approach unsatisfactory.

Borrowing from psychology, then, it may be desirable to incorporate kinesthetic data into the recognition process. For example, if a robot hand were known to be grasping an object, joint angles of the fingers could be processed to determine points of contact in three dimensional space. This approach has been implemented (Kinoshita, 1975; Okada, 1977; Briot, 1978; Ozaki, 1982), but most pattern recognition techniques have proven unreliable for discriminating between objects. Recently, however, a model-based method has been developed which prunes an interpretation tree based on geometric constraints in order to discriminate between known objects (Grimson & Lozano-Perez, 1984). Taking three points of contact as input, this method converges quickly for objects with few faces, demonstrating that such objects can be differentiated on the basis of extremely sparse spatial data.

In order to capture the essence of *active* touch, however, the cutaneous and kinesthetic networks must interact in a feedback loop. That is, cutaneous data should guide the collection of kinesthetic data. A search of the literature shows that this approach was proposed as early as 1974, but was in one case not implemented (Ivancevic, 1974) and in another only partially realized due to sensor limitations (Kinoshita, 1975). Below, we describe a mobile cutaneous sensor which can dynamically explore objects under computer control. Three-dimensional positional coordinates are taken from the sensor's trajectory. We have found that this kinesthetic information is sufficient for recognizing a large class of objects.

The Goal

The goal is to build a machine capable of recognizing objects in the real world. In order to avoid epistemological controversy, *recognition* will be narrowly defined as the act of obtaining, processing, and matching sensory information with a database in order to identify current input with a previously-examined object. Although work is currently underway to incorporate vision into our system (Allen, 1984), the present discussion is concerned only with the sense of touch.

We have restricted our domain to include only rigid, static objects ranging in volume from 10 to 1000 cubic centimeters. Linear, curved, convex, and concave surfaces are allowed with the restriction that all sides be vertical.

The system uses a single tactile probe, the "finger," which can be moved by stepper motors within an XYZ coordinate system. Finger location is taken relative to an origin, (0,0,0), which is located in the upper corner of the frame (see Figure 1).

IMPLEMENTATION

Hierarchical Organization

We have divided the task into two levels. On the lower level, interaction between the tactile device and the mobile arm is used to dynamically circumscribe the object. This process—"tracking"—yields a list of contact points which derive from the arm's trajectory. On the higher level, the list of contact points can then be processed to compute physical parameters such as volume and centroid which lead to identification.

Breaking the task into two levels was originally done for pragmatic reasons—to avoid overloading our computer with real-time data acquisition and motor control. Delegating such functions to specialized components is a fundamental technique for improving system efficiency, and is often found in biological organisms.

System Hardware

The hardware in our system can be loosely partitioned under four regions of activity: Sensing, moving, controlling, and identifying. The first is accomplished by means of a cylindrical "finger" capable of detecting tactual contact almost anywhere on its surface (donated by G. Giralt, LAAS, France—see detail, Figure 1). The finger has a grid of 133 electrodes covered by a variable-resistance rubber "glove." As pressure is applied, a proportional voltage drop occurs at the local electrode. This drop is measured by an external device which samples all 133 electrodes and converts the results into 8-bit digital values. The sensor is susceptible to drift and hysteresis in the low 4 bits (Bajcsy, 1982). To reduce error, the data is thresholded at 17 (decimal) determine if the finger is either "touching" or "not touching." The finger is thus treated as a binary sensor.

The finger is moved within an XYZ frame by three stepper motors: one for each axis. The range of movement in each direction is about 30 centimeters, divided into 1200 steps. Arm position can thus be resolved to within .25 mm. The frame is also equipped with six end-of-travel switches which are used to zero the position registers and to avoid overstepping. Some error is introduced by motor slippage, especially when the sensor is pressed against an ob-

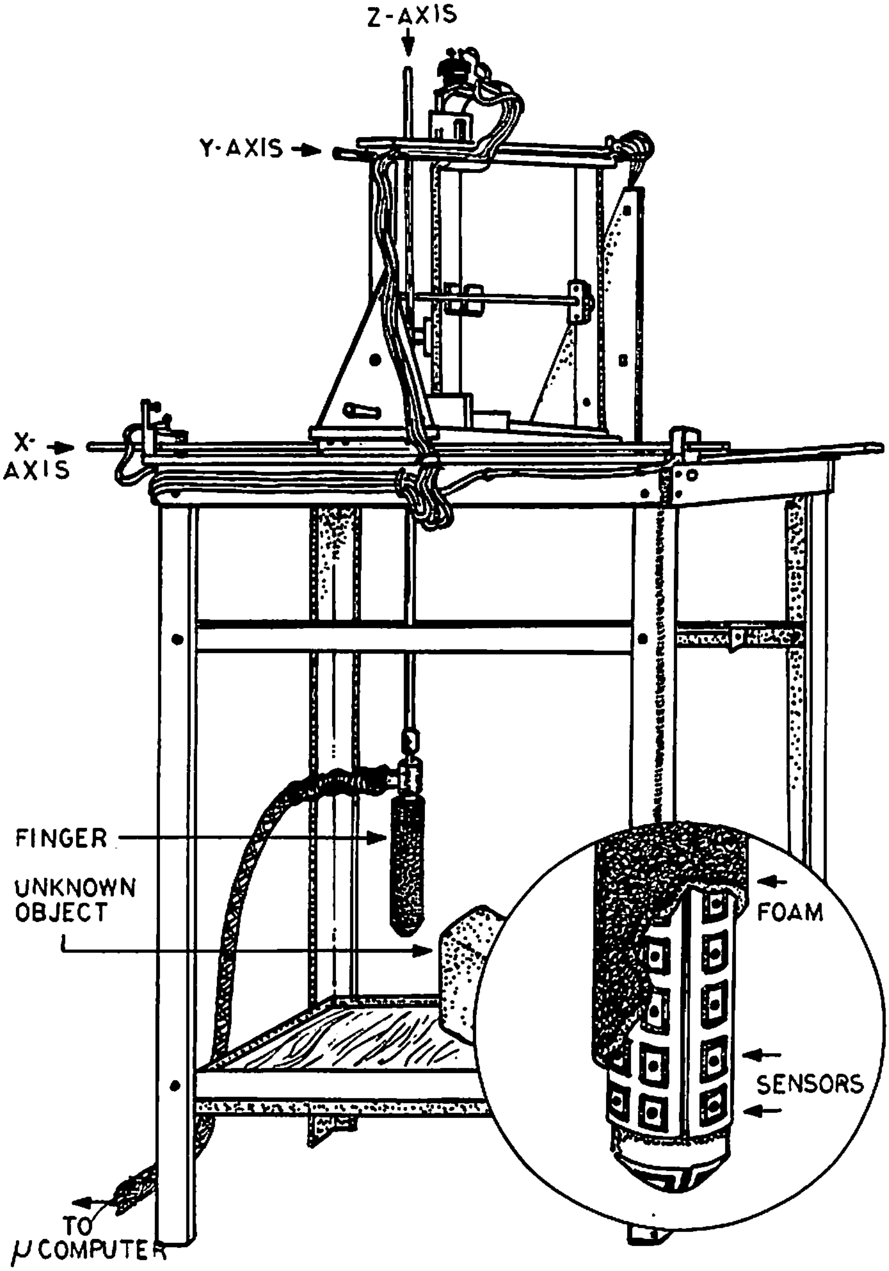

Z-AXIS
Y-AXIS
X-AXIS
FINGER
UNKNOWN OBJECT
FOAM
SENSORS
TO μ COMPUTER

Figure 1

ject. These errors are cumulative and appear as "fuzziness" on the output described below.

Two dedicated Z80 microprocessors allow for independent control of tactile and mobile subsystems. The tactile processor continually monitors the sensors, filling a table with values and comparing with the threshold to detect contact. The motor processor computes direction vectors and keeps track of finger position. Interprocessor communication is achieved via interrupts: the tactile processor can inform the motor processor if contact is made. The processors are in turn connected by a serial interface to a PDP 11/60 minicomputer.

The 11/60 directs simple commands to the processors and receives data. It is here that high-level processing, recognition, and generation of visual models takes place.

Low-Level Software

Software for the system exists on two levels. At the low level, assembly language routines reside in the microprocessors for fast execution of commands such as:

100X	- relative move: move 100 steps in X direction.
@500,500,500	- absolute move: move to location (500,500,500).
=	- position: send current position to 11/60.
S	- snapshot: send current sensor values to 11/60.
G900,900,900	- guarded move: move toward (900,900,900) but stop if contact is made enroute.

Most interesting here is the guarded move. This command takes advantage of interprocessor communication to have the tactile processor stop the motors if contact is detected. This concise directive is used throughout the tracking routine and has eliminated the need for clumsy trial-and-error search commands.

The two routines which follow run on our minicomputer. The first of these is a search routine which begins at (0,0,0) and sweeps out a prescribed path until an object is encountered. This process assumes that gravity will hold objects down; hence it sweeps only the bottom of the XYZ domain.

The second routine directs the system in tracking the object. Whenever tactile contact is registered, arm position is recorded in memory as a function of three coordinates (XYZ). The intuitive algorithm uses guarded moves to move the finger around an object in a clockwise direction. While tracking, the system can handle two contingencies: losing contact or becoming blocked (see Figure 2). When the finger arrives within 5 mm of its initial point of contact, it backs off, approximates the object's center, and approaches it from

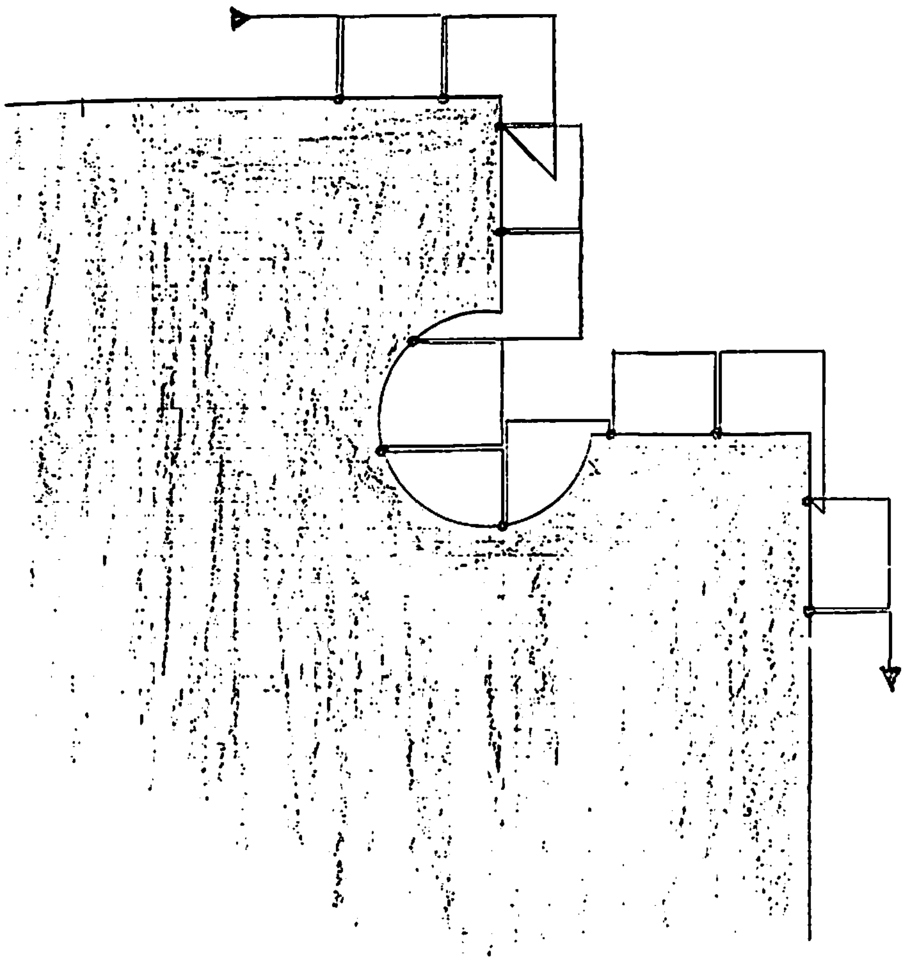
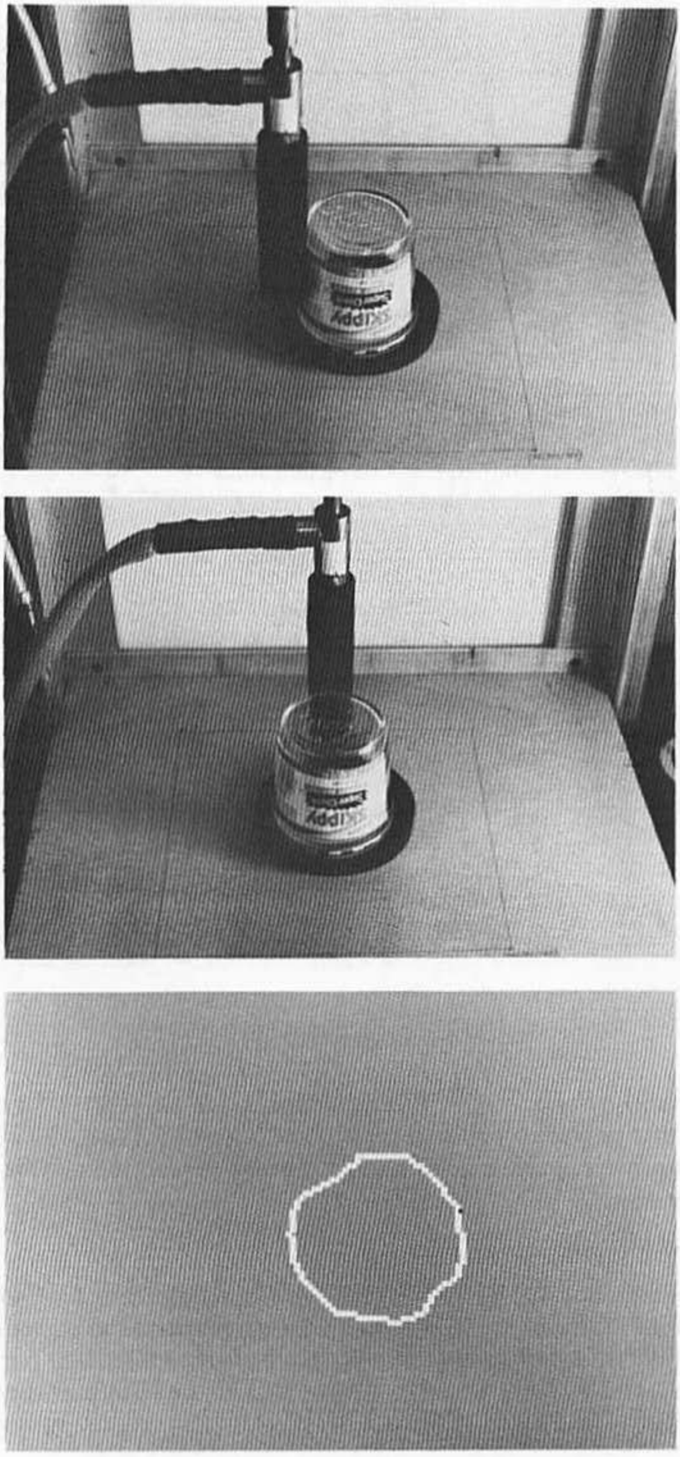


Figure 2

the top to establish object height. This algorithm, originally written for convex cuboids, has proved surprisingly general and has worked well with cylindrical and concave objects, suggesting that the task of tracking may be domain independent.

High-Level Software

High level recognition is then carried out by processing the list of points which represent locations in three-dimensional space where the object was encountered. First, a routine connects vectors between subsequent points. Since the arm travels about 10 mm between contacts, the resulting "slice" is somewhat jagged (see Photos 1). Next, a trapezoidal approximation is used



Photos 1

to integrate the area enclosed by these points and to produce parameters such as centroid, volume, and the moment of inertia.

Parameters such as these are then compared against a file of known objects. If a "match" is found, the system identifies the object by name and in some cases can generate a shaded model of the object for display (see Photos 2). If no match is found, the system requests the user to input a name which is then stored for subsequent identification. In this way, the system "learns" about its environment.

RESULTS

Within the narrow definition of recognition, our tactile system, dubbed "Orion" after the mythological blind hunter, has been able to consistently recognize objects in its environment. We have used common household items such as jars, coffee cups, and cardboard boxes as well as wooden shapes created specifically to test difficult surfaces (see Photos 3, 4, 5).

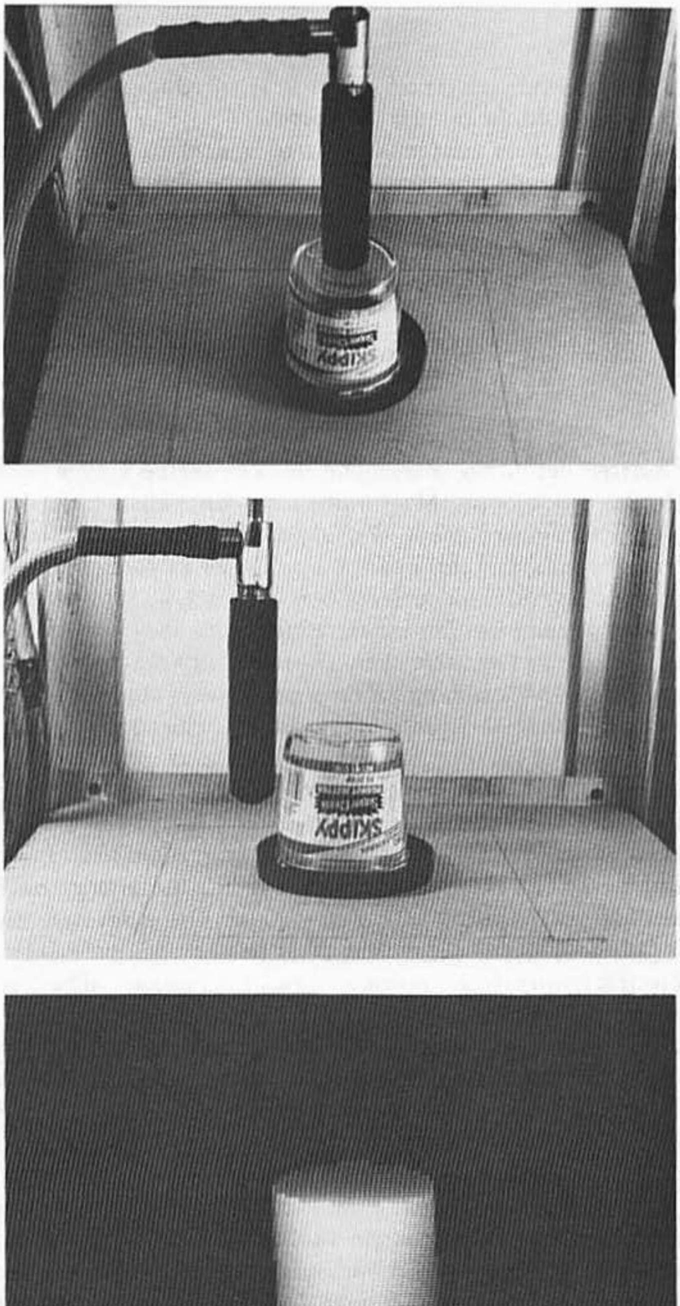
Once the object is placed within the cartesian frame, the operator invokes Orion, whereupon the system takes over. The probe moves across the object, investigating corners and searching for recessed surfaces until the object has been circumscribed. Secondary routines then process kinesthetic data to produce visual displays and perform database comparisons.

The kinesthetic output produced by Orion has produced volume estimations which vary by less than 6% (see Table 1). Errors can be attributed to slippage and variation in sensor sensitivity over time. It has been suggested that a calibration trial before exploration would reduce the latter effect. Note that the true volume of the peanut butter jar is 200 cc. Orion does not attempt to correct for probe diameter (3cm) and other fixed system parameters because the objective is only to gain repeatable estimates which can be matched with previously-stored inputs.

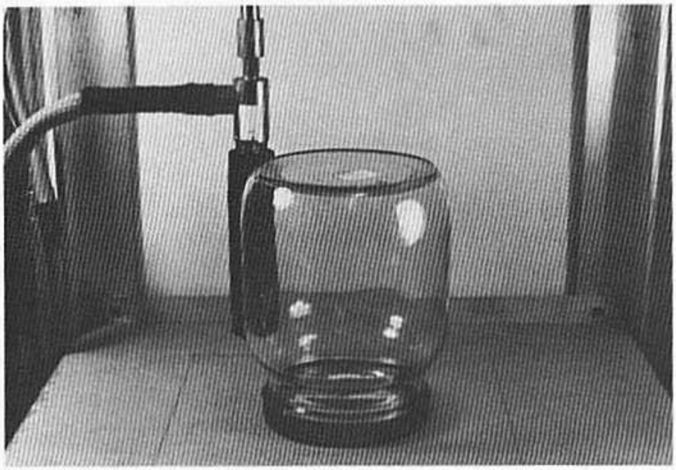
Orion can discriminate between a dozen or so objects currently represented in its database. When input cannot be matched to within 5% tolerance of existing models, Orion queries the operator for a label and appends the new object (label and parameters) to its database.

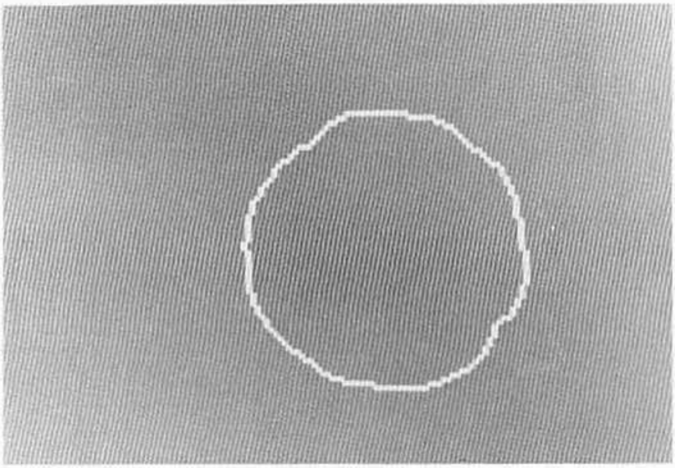
The sensor has provided several problems: (1) a "bad" sensor, which continually registers contact and must be masked out with software, (2) blind spots, which occur at the 45 degree angles of the finger requiring that objects have vertical sides to avoid contact with this part of the finger, and (3) insensitivity to light contact, which requires that the object be held stationary with fasteners to avoid it being pushed by the manipulator.

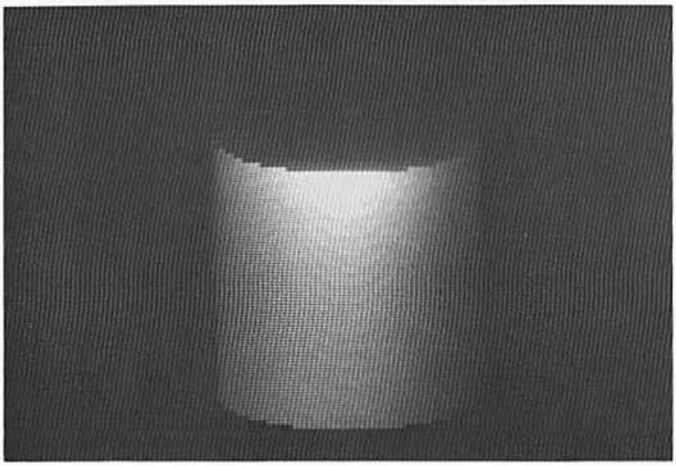
Our cartesian positioning device is limited to three degrees of freedom. This makes it impossible, for instance, to explore recessed regions which are blocked by overhanging surfaces. Similarly, blind spots on the sensor can



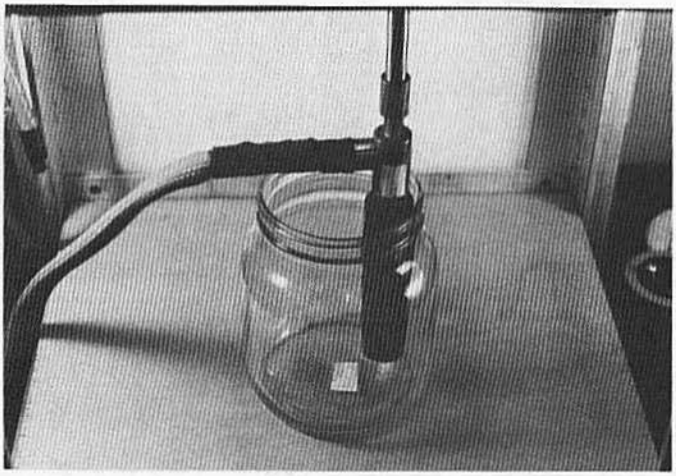
Photos 2

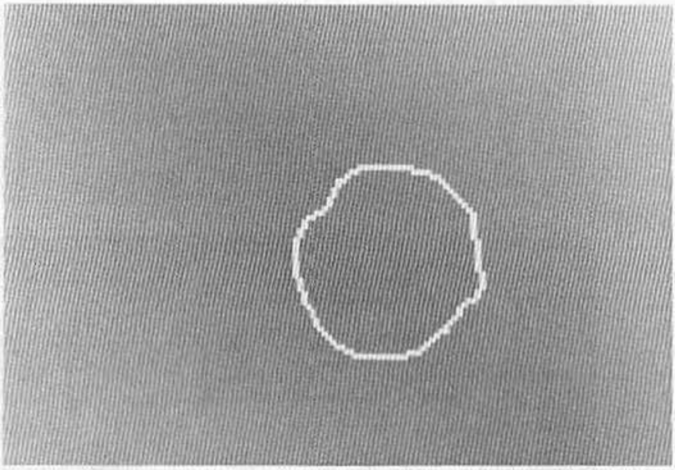




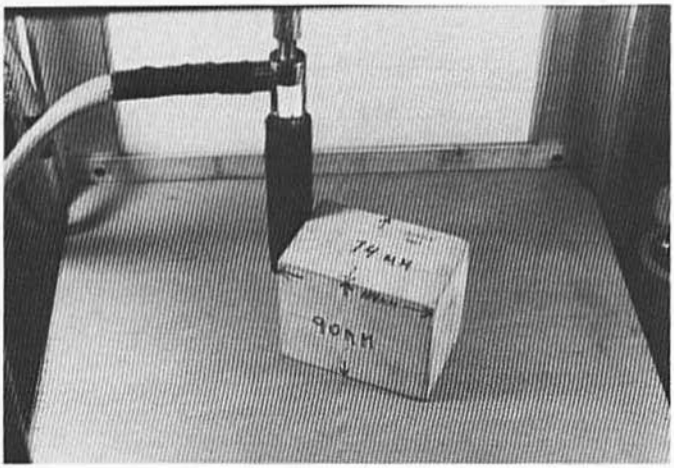


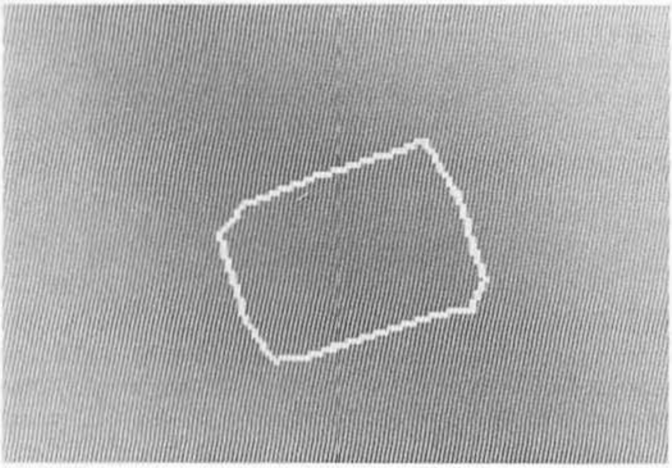
Photos 3





Photos 4





Photos 5

TABLE 1
Ten Trials on Peanut Butter Jar

1:	570.5 cc	
2:	588.0	
3:	604.9	
4:	586.6	
5:	589.3	
6:	574.2	
7:	548.2	
8:	556.0	
9:	584.2	
10:	580.1	
$\overline{X}$	=	578.2cc
σ	=	15.8
max err (trial 7)	=	5.2%

only be avoided by putting restrictions on the objects being sensed. The finger is now being interfaced to a PUMA robot arm with seven degrees of freedom where it will be integrated with a vision system (Allen, 1984).

In addition, the three-fingered Pennsylvania Articulated Mechanical Hand (PAMH) is currently being fitted with tactile sensors for research into how grasping attempts may lead to identification.

CONCLUSIONS

Psychologists have found that people can recognize an object better if they move their hands over it rather than have it pressed into their palms from several orientations. This suggests that another mechanism, the kinesthetic sense, might work in conjunction with the skin to achieve object recognition. The terms active and passive were applied to touch in the early sixties to convey this cooperative relation between the cutaneous and kinesthetic senses.

Roboticists, meanwhile, have struggled to produce machines which can mimic human touch. These efforts have been directed primarily toward the cutaneous sense; the kinesthetic sense has never been fully integrated into a system designed to imitate active human touch. The Orion system uses a tactile sensor in a feedback loop to physically circumscribe an object. Object parameters are then reconstructed from the motion of the sensor. We believe that this represents a crude model of active touch.

While Orion's cutaneous sensor conveys very little information (one bit), the system is nonetheless able to consistently distinguish between objects in its environment. From a psychological perspective, Orion's design suggest that human beings may prefer active touch over passive for the same reason that Orion monitors its positioning device: the information provided by kinesthetic receptors is more reliable than that provided by the skin.

Joint coordinates on a robot arm can be accurately monitored by counting discrete increments which are inherently digital and do not require analogue conversion. Further, the signal-to-noise (SNR) of kinesthetic information is higher than that of cutaneous, since the dynamic range of limb positions greatly exceeds that of the tactile values. While the psychophysical characteristics of human skin are outside the scope of this paper, spatial sensitivity is low, being on the order of 1 mm. The cutaneous layer is also susceptible to variations in insensitivity, with scars and bandages being analogous to the blind spots on our tactile sensor. Perhaps the human mechanism has evolved in a manner that puts emphasis, pragmatically, on the best data.

From the standpoint of robotics, the success of the Orion system, albeit limited, suggests that an effective tactile machine, which can pick up and identify objects automatically, must utilize both cutaneous and kinesthetic sensors to accomplish this task.

ACKNOWLEDGMENTS

The authors are indebted to David Brown, Jeff Wolfield, and Dan Peters for their significant contribution to the development of the tactile system, and to Peter Allen for suggesting the guarded move. (Figure 1: Avrum Goldberg).

This work was in part supported by: NSF grant MCS-82-07294, Airforce grant AFOSR 82-NM-299, NIH grant 5-RO1-HL-29985-02, NIH grant NS-10939-11 as part of Cerebro Vascular Research Center, DEC corp, IBM corp. and LORD corp.

REFERENCES

Allen, P. Object recognition for vision and touch. University of Pennsylvania PhD. Proposal, 1984.

Bajcsy, R. What can we learn from one-finger experiments? University of Pennsylvania CIS Department Working Paper, 1982.

Briot, M., M. Renaud, & Z. Stojilkovic. An approach to spatial pattern recognition of solid objects. *IEEE Transactions on Systems Man, and Cybernetics*, vol. SMC-8, #9, September, 1978.

Briot, M. The utilization of an "artificial skin" sensor for the identification of solid objects. 9th International Symposium for Industrial Robotics, 1979.

Derevensky, J. L. Relative contributions of active and passive touch to a child's knowledge of physical objects. *Perceptual and Motor Skills, 48*(3), June (1979) 1331–1346.

Gibson, J. J. Observations on active touch. *Psychology Review*, 1962, *69*, 477–490.

Giralt, Prof. G. LAAS, Toulouse Laboratory, France. Developed the finger in 1981.

Goodenow, J., J. Abramowitz, & B. Paul. Pennsylvania Articulated Mechanical Hand. ASME International Conference on Robotics 1983.

Gordon, G. (Ed.) *Active touch.* Pergamon Press, 1978.

Grimson, W. E., & T. Lozano-Perez. Model-based recognition and localization from tactile data. *Proceedings of the IEEE International Conference on Robotics*, Atlanta, GA, March 13–16, 1984.

Heller, M. A. Reproduction of Tactually Perceived Forms. *Perceptual and Motor Skills, 50*(3), June (1980) 943–946.

Hillis, W. D. A high-resolution imaging touch sensor. *International Journal of Robotics Research I.*, 2. Summer 1982.

Ivancevic, N. Stereometric Pattern Recognition by Artificial Touch. *Pattern Recognition, 6* (1974), 77–83.

James, W. *The principles of psychology.* Henry Holt and Co. 1890. (see especially Chapter XX).

Kandel, E. R., & J. H. Schwartz. *Principles of neural science.* Elsevier, 1981.

Katz, D. Der Aufbau der Tastwelt. Zeitschrift fur Psychology. Erganzungband 11 (1925). (see Krueger for a synopsis).

Kinoshita, G., S. Aida, & M. Mori. A pattern classification by dynamic tactile sense information processing. *Pattern Recognition, 7*, 1975, p. 243.

Krueger, L. E. David Katz's der aufbau der tastwelt (The world of touch): A Synopsis. *Perception and Psychophysics, 7;* 6, (1970).

Lawrence, D. M., N. J. Cobb, & J. I. Beard. Influence of active encoding on tactile recognition memory for common objects. *Perceptual and Motor Skills, 47*, Oct (1978) 596–598.

Loomis, J. M., & S. J. Lederman. Tactual pattern perception. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of human perception and performance.* New York, Wiley (in press).

Okada, T., & S. Tsuchiya. Object recognition from grasping. *Pattern Recognition* 9, (1977) 111–119.

Ozaki, H., et al. Pattern recognition of a grasped object by unit-vector distribution. *IEEE Transactions on Systems, Man, and Cybernetics. 12*, 3, May-June (1982) 315.

Purbrick, J. A. A force transducer employing conductive silicon rubber. 1st Robot and Vision Conference, Stratford-On-Avon, England, April, 1981.

Raibert, M. H., & J. E. Tanner. Design and Implementation of a VLSI Tactile Sensing Computer. *International Journal of Robotics Research* I, 3, 1982.

Sato, N., W. B. Heginbotham, & A. Pugh. A method for three dimensional part identification by tactile transducer. 7th International Symposium for Industrial Robotics, 1977.

Weinstein, S. Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D. R. Kenshalo (Ed.), The skin senses. Charles C. Thomas, Springfield, Ill. 1968.