Acquisition and Interpretation of 3-D Sensor Data from Touch

PETER K. ALLEN, MEMBER, IEEE, AND PAUL MICHELMAN, STUDENT MEMBER, IEEE

Abstract— Acquisition of 3-D scene information has focused on either passive 2-D imaging methods (stereopsis, structure from motion, etc.) or 3-D range sensing methods (structured lighting, laser scanning, etc.). Little work has been done in using active touch sensing with a multifingered robotic hand to acquire scene descriptions, even though it is a well developed human capability. Touch sensing differs from other more passive sensing modalities such as vision in a number of ways. A multifingered robotic hand with touch sensors can probe, move, and change its environment. This imposes a level of control on the sensing that makes it typically more difficult than traditional passive sensors in which active control is not an issue. Second, touch sensing generates far less data than vision methods; this is especially intriguing in light of psychological evidence that shows that humans can recover shape and a number of other object attributes very reliably using touch alone. Future robotic systems will need to use dextrous robotic hands for tasks such as grasping, manipulation, assembly, inspection, and object recognition. This paper describes our use of touch sensing as part of a larger system we are building for 3-D shape recovery and object recognition using touch and vision methods. It focuses on three exploratory procedures we have built to acquire and interpret sparse 3-D touch data: grasping by containment, planar surface exploration, and surface contour exploration. Experimental results for each of these procedures are presented.

I. INTRODUCTION

ACQUISITION of 3-D scene information has focused on either passive 2-D imaging methods (stereopsis, structure from motion, etc.) or 3-D range sensing methods (structured lighting, laser scanning, etc.). Little work has been done in using active touch sensing with a multifingered robotic hand to acquire scene descriptions, even though it is a well developed human capability [24]. Touch sensing differs from other more passive sensing modalities, such as vision, in a number of ways. A multifingered robotic hand with touch sensors can probe, move, and change its environment. This imposes a level of control on the sensing that makes it typically more difficult than traditional passive sensors in which active control is not an issue. Second, touch sensing generates far less data than vision methods; this is especially intriguing in light of psychological evidence (described below) that shows that humans can recover shape and a number of other object attributes very reliably using touch alone.

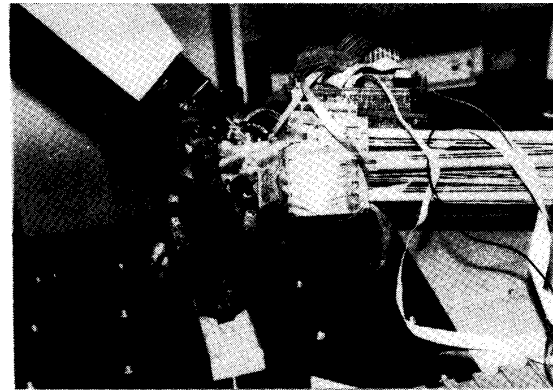Future robotic systems will need to use dextrous robotic

Fig. 1. Utah-MIT hand with tactile sensors mounted.

hands for tasks such as grasping, manipulation, assembly, inspection, and object recognition. This paper describes our use of touch sensing as part of a larger system we are building for 3-D shape recovery and object recognition using touch and vision methods. It focuses on three exploratory procedures we have built to acquire and interpret sparse 3-D touch data. These procedures serve as a front end to an integrated shape recovery and object recognition system that can combine these exploratory procedures into strategies that can derive constraints about an object's most probable shape (described in Roberts [31]).

The outline of this paper is as follows: Section II is an overview of the hardware/sensing environment we have built to perform intelligent hand functions, Section III describes the tactile sensing system we have implemented, Section IV describes three exploratory procedures we have implemented for acquiring and interpreting 3-D touch information, and Section V is a summary that outlines future work to be done with the hand.

II. SYSTEM OVERVIEW

The system we have built consists of a Utah-MIT hand [19] attached to a PUMA 560 manipulator. The hand contains four fingers, and each has four degrees of freedom. It resembles the human hand in size and shape but lacks a number of features that humans find very useful. In particular, it has no palmar degree of freedom (closing of the palm), and the thumb is placed directly opposite the other three fingers with all fingers identical in size (see Fig. 1). The hand has joint position sensors that yield joint angle data and tendon force sensors that measure forces on each of the two tendons (extensor and flexor) that control a joint. The PUMA adds six degrees of
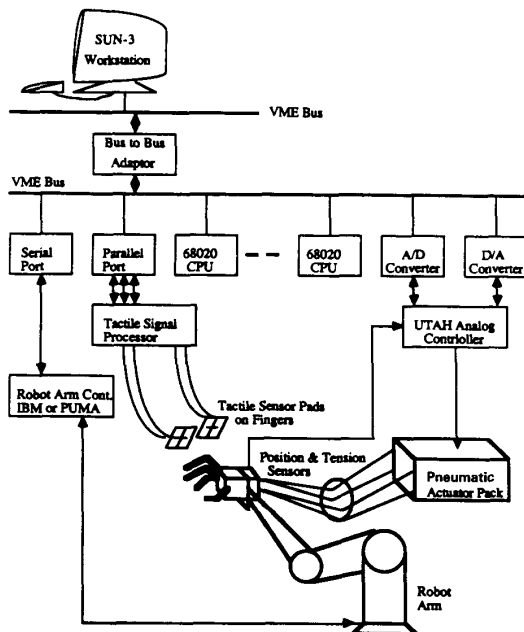
Fig. 2. Hardware overview.

freedom to the system (three translation parameters to move the hand in space and three rotational parameters to orient the hand), yielding a 22 degree-of-freedom system. Clearly, such a system is a nightmare to control at the servo level in real time. Our approach is to use the embedded controllers in each of these systems, controlling and communicating with them through an intelligent, high-level controller that links together the movements of arm, hand, and fingers with the feedback sensing of joint positions, tendon forces, and tactile responses on the fingers.

The hardware structure of the system is shown in Fig. 2. The high-level control resides in a SUN-3 processor. The SUN serves as the central controller and has access to a full UNIX-based system for program development and debugging as well as a set of window-based utilities to allow graphical output and display of the system's various states. The hand is controlled by an analog controller that is commanded through D/A boards from a dedicated 68020 system. The SUN is capable of downloading and executing code on the 68020 and can communicate with it through a shared memory interface [28]. The tactile sensing system is controlled by another dedicated 68020 that monitors the forces on each of the sensor pads. The connection from the SUN to the PUMA is via the VAL-II host control option over a serial interface. We are currently changing the interface to the PUMA to RCCL [14] to make the hand-arm interaction more tightly coupled. The system has been used to perform a number of object manipulation and grasping tasks, including pouring liquids from pitchers and removing lightbulbs from sockets [2].

III. Tactile Sensors

Although the level of sensing provided by the joint position and tendon force sensors on the Utah-MIT hand is better than earlier implemented hands, it still falls far short of the

requirements for a dextrous manipulation system. In particular, what is desired is accurate positional contact information between the hand and a target object and a measure of the forces exerted by the fingers at these contact points. The sensory feedback provided by the hand does not allow for localization of contacts. Hence, a requirement for this system is a robust and accurate tactile sensing capability that utilizes sensors mounted on the links of the fingers. Tactile sensing differs from traditional vision sensing in its active nature. Thus, a robotic system that employs tactile sensors on the fingers of a dextrous hand must deal with three related issues: 1) acquisition and interpretation of tactile sensor data from many sites on multiple fingers, 2) control of the dextrous hand using tactile sensor feedback, and 3) development of sensing strategies using tactile feedback.

To satisfy the first requirement, we have mounted tactile sensors on each of the hand's fingers. The technology being used is a piezoresistive polymeric material manufactured by Interlink, Inc. [33], [36]. The design of the tactile pads we are using sandwiches the polymer between two pliable sheets of Kapton material that contains electrical etching. The application of forces on the pads provides an increased electrical flow channel between the two sheets as the material within is compressed. The piezoresistive polymer is patterned to form rows on one substrate and columns on the other. The rows and columns form a grid in which each intersection acts as a force-sensitive variable resistance whose value decreases approximately exponentially with normal force. The pads consist of 16 rows by 16 columns, providing a sense resolution of 256 points on a 0.5×1.0 in pad.

The 256 sites of each sensor pad are addressed independently by analog circuitry that cancels current flow in all paths of the grid except the one containing the resistive element being measured using a method developed by van Brussel and Belien [37]. A hardware interface board has been developed to perform this operation at high speed. The interface board performs the analog-to-digital conversion task by means of an 8-bit flash A/D converter and allows up to 16 sensor pads to be addressed.

Some of the low-level tactile primitives that have been implemented include the following:

- *Tactile Filters:* A number of useful digital filters have been implemented, including averaging and median filters, which are very useful in processing noisy tactile data [26].
- *Tactile Moments:* A useful technique for quickly getting contact information is central moment analysis [18]. The contact area and centroid of the contact can be determined using moments. The second moments are useful for determining the eccentricity of the contact region and the principal axes of the contact.
- *Edge Detection:* A number of edge detectors have been developed and used for feature extraction from tactile images.
- *Line Detection:* Lines are detected by using the output of the edge detection procedure in a Hough transform [5].

Results with this sensor have been good. The signal is very localized, and by using moment analysis, we have been able to stably determine contact location on the pads.

IV. ACTIVE HAPTIC SENSING FOR OBJECT RECOGNITION TASKS

A focus of our work has been in the use of the hand system described above to recover the shape of objects in a scene. Object recognition has traditionally been associated with vision sensor systems. However, these systems suffer from a number of inherent problems, not the least of which is occlusion. A vision system will be limited to a view that obscures all back-facing areas of the object. In robot manipulation tasks, important areas of the work environment are occluded by the end effector itself. This difficulty is especially acute during the act of acquiring a grasp on an object when the contact areas will be occluded. A number of interesting properties of the human haptic[1] system have been investigated by Lederman and Klatzky and their colleagues [23]–[25]. This work has shown that an important component of the haptic system is its ability to recognize attributes of 3-D objects quickly and accurately. Among these attributes are global shape, hardness, temperature, weight, size, articulation, and function. An outcome of this research is the identification of hand movement strategies that are used by humans in discovering different attributes of 3-D objects. They have labeled these exploratory procedures (EP's) and have reported success rates of 96–99% in identifying different object properties using two-handed, haptic exploration. We have found it natural to extend these human capabilities to our robotic domain.[2] We have implemented three EP's on our robotic hand system, which we describe below.

A. Coarse to Fine Recognition Strategies

In acquiring information about a scene, a hierarchical approach seems intuitive. Information content is often related to scale, and different sensory systems work at different size and detail scales [8], [38]. Our approach is to find gross object shape initially and then use a hypothesis-and-test method to generate more detailed information about an object, as discussed in Allen [1]. This approach is especially relevant with touch sensing, in which there is evidence that the human tactile system serves essentially as a low-pass filter [25]. This

motivates the idea of using an initial global estimate of shape, which can then be further refined by more specific and localized sensing. The problem of generating a good initial hypothesis is central to robust object recognition. If we can generate a good initial shape estimate, we will be much more successful as we try to discover further object structure. The requirements for an initial shape estimator are that it be efficient, stable in the presence of noise and uncertainty, and able to use sparse, partial data. We have implemented such a shape recovery method, which we call grasping by containment. This method was initially discussed in [3], and it is reviewed here since the method serves as a precursor to the other two EP's we have implemented.

B. Exploratory Procedure 1: Grasping by Containment

Grasping by containment is an attempt to understand an object's gross contour and volume by effectively molding the hand to the object. We have chosen to model objects as superquadrics [4], [6], [29] whose surface 3-D vector X is defined below using a latitudinal and longitudinal parameterization expressed in spherical coordinates:

$$X(\eta, \omega) = \begin{bmatrix} a_1 C_\eta^{\epsilon_1} C_\omega^{\epsilon_2} \\ a_2 C_\eta^{\epsilon_1} S_\omega^{\epsilon_2} \\ a_3 S_\eta^{\epsilon_1} \end{bmatrix}$$

$$-\frac{\pi}{2} \le \eta \le \frac{\pi}{2}, \ -\pi \le \omega \le \pi \quad (1)$$

where C_η, S_ω are cosine(η) and sine(ω), ϵ_1, ϵ_2 are the superquadric shape parameters, and a_1, a_2, a_3 are scaling factors along the X, Y, and Z directions.

Superquadrics form a rich set of shape primitives that allows a wide degree of freedom in modeling objects. The parameter space is continuous and allows a smooth change from a cuboid to a sphere to a cylinder with more complex shapes derivable with the addition of bending and tapering parameters. These "lumps of clay" are deformable by the usual linear stretching and scaling operations and can be combined using boolean set operations to create more complex objects.

What makes superquadrics particularly relevant for haptic recognition is the following:

- The models are volumetric in nature, which maps directly into the psychophysical perception processes suggested by grasping by containment.
- The models can be constrained by the volumetric constraint implied by the joint positions on each finger.
- The models can be recovered with sparse amounts of point contact data since only a limited number of parameters need to be recovered. There are five parameters related to shape and six related to position and orientation in space. Global deformations (tapering, bending) add a few more.
- In addition to the use of contact points of fingers on a surface, the surface normals from contacts can be used to describe a dual superquadric, which has the same analytical properties as the model itself [6].

---

[1] An important point to be made in applying hands to robots is that the human perceptual process of interest is *haptic* perception. By this, we mean the interplay of both the cutaneous system (skin, tactile receptors) and the kinaesthetic system (joints, muscle, and bone) of the arm [13].

[2] We must be careful in trying to draw too close a comparison between a human hand and devices such as a Utah-MIT hand. Johansson and Vallbo [20] have reported that there are about 17,000 mechano-receptors in the skin of the human hand; our robotic hand is more limited with 16 joint sensors, 32 tendon force sensors, and four 16 × 16 fingertip tactile sensors. In addition, a human hand has two main differences in structure from our robotic hand. The first is a highly flexible, opposable thumb that is mounted to the side of the other digits. The Utah-MIT hand thumb is identical to the other fingers and is mounted directly opposite the other fingers. The second difference is that a palmar degree of freedom exists in human hands that is missing in the Utah-MIT hand. Humans find this palmar degree of freedom quite useful, especially for encompassing type grasps where the hand is molded to an object and as a grasping mechanism in its own right, which is almost independent of the existence of multijointed fingers.
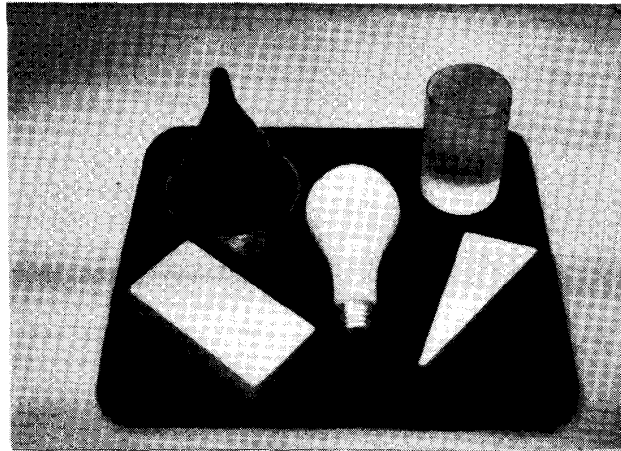
Fig. 3.   Object database.

- The analytic nature of the model created from sparse data allows searching strategies in the model space to proceed in a hypothesize-and-test fashion.

C. Recovery Procedure

For this initial work on recognition, we have used a simplified procedure to gather data points. Our intent is to use the tactile sensors mounted on the finger links to generate contact position data. However, during our initial trials, our tactile sensors were not yet mounted on the hand. Instead, we opted for a method that used the hand's internal joint angle readings and tendon forces to generate Cartesian positions of contact based on fingertip contact. Note that in all the experiments described in this paper, the objects to be explored were rigidly attached to a laboratory table. Exploring objects with active touch sensors can sometimes displace the objects. This is a very difficult problem that is solved at the human level by using two hands, where one hand typically serves as a "fixture." In the absence of a second hand, we have rigidly mounted the objects to allow one-hand exploration.

The PUMA arm moves the hand to a position in which it will close around the object. The fingers are spread wide during approach. The fingers are then closed by position commands until the observed force (which is estimated by the difference between the flexor and extensor tendon tensions) exceeds a given threshold, which indicates that the finger is in contact with the object. The joint angle positions are read, and kinematic models of the hand and the PUMA arm are used to convert them to XYZ positions in world coordinates. The fingers are then opened wide again, and a second containing grasp is executed with the fingers taking different approach paths. The fingers are spread once again, and the PUMA arm moves the hand to the next position.

The sequence of PUMA positions is given in advance. Once the contact points are determined using the forward kinematics of the hand derived from the joint angle sensors, the sparse sets of point data is injected into the recovery algorithm developed by Solina [32]. This algorithm uses a Levenberg–Marquardt nonlinear least squares approximation to fit the superquadric "inside-out function." This is an im-

plicit form of (1), which records if a sample data point lies inside, outside, or on the surface of the superquadric model. By summing the squared distance of each sample data point from the current model, an error of fit measure is generated that is minimized by the algorithm.

Equation (1) is for a canonical superquadric located at the origin. Since our sensor data can exist anywhere in the world coordinate space, the algorithm must recover the six rotation and translation parameters in addition to the five superquadric shape parameters $(a_1, a_2, a_3, \epsilon_1, \epsilon_2)$. In addition, we allow global deformations to include tapering of superquadric forms. The taper is defined to be a linear tapering with two parameters that control the tapering in both the X and Y dimensions. The algorithm must recover a minimum of 11 parameters (13 if the object is tapered).

We tested this procedure against a database of six objects (shown in Fig. 3 plus a smaller cylinder). The database included objects that could be modeled as undeformed superquadrics (block, large cylinder, small cylinder) and deformed (tapered) superquadrics (lightbulb, funnel, triangular wedge). The recovered shapes are shown in Fig. 4 with the sample data points overlaid on them.

The recovered shapes are an accurate representation of the actual shapes, especially considering the sparse nature of the data and the errors in the derived contact points. These errors are a function of the accuracy and calibration of the robotic arm, the hand joint position sensors, and the kinematic model of the hand itself. The data points are overlaid on the recovered shapes to show the closeness of fit and the sparseness of the data. Each object's shape was recovered with extremely sparse amounts of data, which was typically 30–100 points, depending on the object. It is important to note that this is about two orders of magnitude less than typical range data images that try to recover shape with denser data that, unlike touch sensing, is limited to a viewpoint that only exposes half the object's surfaces to the sensor. Not all objects can be modeled adequately as superquadrics; however, we feel that in deriving an initial shape estimate from sparse data, they work quite well. The algorithm typically converges after about 10–15 iterations, and it takes about 10 s on a Sun-4 to

Fig. 5.   Planar surface explorer tactile contacts on two planar surfaces of a rectangular block.
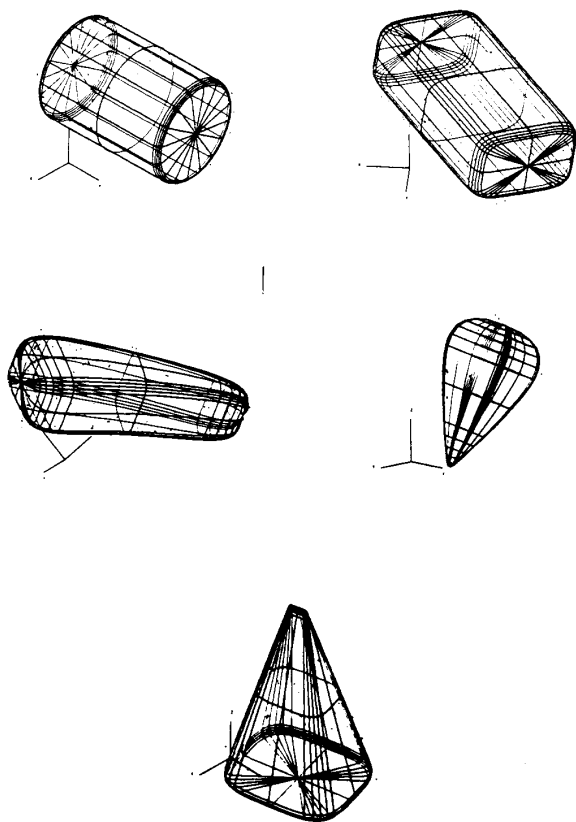
Fig. 4.   Recovered shape of cylinder, block, wedge, lightbulb, and funnel.

recover an undeformed superquadric shape (20 iterations of a 60-contact-point data set). A formal analysis of the convergence properties and accuracy of the recovery method can be found in Solina's thesis [32] and in work by Boult and Gross [9].

D. Exploratory Procedure 2: Planar Surface Explorer

Once a superquadric has been fit to the initial grasp data, we have a strong hypothesis about an object's shape. The shape parameters ϵ_1 and ϵ_2 are of particular importance. The shape of an object can be inferred from these parameters and used to direct further exploration. For example, if the shape parameters appear to be rectangular (ϵ_1, $\epsilon_2 \approx 0.1$), the planar explorer can trace out the plane and perform a least square fit of the trace data to test the surface's planarity. If the shape parameters appear more cylindrical ($\epsilon_1 \approx 1$, $\epsilon_2 \approx 0.1$), the planar faces of the cylinder can be explored with this primitive, and the cylinder's contour can be explored and verified with the contour follower EP (described below). A major benefit of using the superquadric analytic shape description is that it supplies orientation and axis data that are necessary for further active probes of the environment with the hand. Instead of a blind search, we can use the recovered orientation parameters to guide the further exploration of the object. Discovering a planar surface can be a very useful constraint in recognition, particularly if two opposing planar faces are grasped. By discovering multiple planar faces on an object, the recovery methods of Grimson and Lozano-Perez [12] and
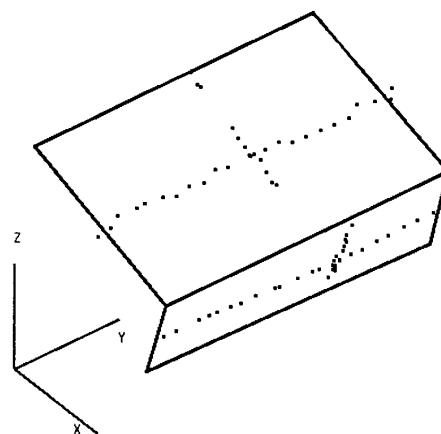
Ellis *et al.* [10] can be invoked, and these have proven to be strong constraints on recognition of an object.

The explorer uses the hand's index finger. Although the index finger is held in an extended position, the PUMA arm is moved until the tactile sensors on the index finger contact a surface (if no contact is detected, the procedure terminates). After the initial contact, the Cartesian position of the contact point is noted. The hand and arm then begin an iterative search for the boundaries of the surface by performing the following sequence: a) lift the finger off the surface until tactile contact is lost; b) move the arm in a direction parallel to the surface; c) if the finger is in contact after the movement, note the new contact location; otherwise, lower the index finger until it makes contact with the surface again; d) repeat steps a)–c) until the finger fails to make contact in step c). In step d), if the finger does not contact the surface, either the finger has moved beyond the edge of the surface, or the surface is too far away from the finger to be detected. To check for the latter case, the arm must be moved toward the surface. After completing the first collection of data points and finding the edge of the surface, the index finger is moved back to the position of initial contact, and a second mapping of the surface is undertaken in a direction 180° opposite. This procedure continues until a second surface edge is detected. The search now continues as before but in a direction perpendicular to the first two traces. This procedure is then able to map out a set of contact points on the surface describing its extent. Each time the fingertip contacts the surface, the Cartesian coordinates of the contact are retained. The acquisition of data points in this method is compatible with the three-point seed method of Henderson and Bhanu for forming planar surfaces from range data [15]. Fig. 5 shows a pattern of traces on two adjacent planar faces of a rectangular block using this EP. Least-square planes were fit to each of the traces, and the computed angle between the recovered planes is 96° (the actual angle is unknown but assumed to be 90°).

E. Exploratory Procedure 3: Surface Contour Following

The third EP we have implemented is surface contour following with a two-fingered grasp. This EP will allow us to

determine an object's contour that has been shown to be a strong shape cue from previous vision research [27], [30], [35]. The contour follower EP in humans is described by Klatzky and Lederman as "a dynamic procedure in which the hand maintains contact with a contour of the object. Typically, the movement is smooth and non-repetitive within a segment of object contour, stops or shifts direction when a contour segment ends, and does not occur on a homogeneous surface [22]." It seems natural that this EP will report information that can be used to recover a shape that can be represented as a class of generalized cylinders. The contours we are able to extract from touch are inherently 3-D. This simplifies recovery of shape since the 2-D image projection used in most contour work entails a loss of information. Since we can recover the 3-D contours, we are able to hypothesize a number of different shapes, including generalized cylinders and solids of revolution, using the 3-D contour alone.

Generalized cylinders have been proposed by many researchers (beginning with Binford [7]) as a shape modeling primitive. Other researchers have expanded on this idea of a swept volume by creating classes of generalized cylinders or cones, depending on the nature of the axis curve, sweeping rule, and cross sectional curve. These primitives have special appeal in the recognition of elongated objects and objects that provide strong visual contours.

Fearing [11] has attempted to recover the shape of a class of these generalized cones, known as right linear straight homogeneous generalized cones (RLSHGC's) using extremely sparse amounts of data. He has characterized the necessary and sufficient conditions for being able to recover the axis and orientation of these cones given limited, multifingered tactile sensor data that includes point contacts, surface normal direction, and surface curvature information.

We have also chosen to use a class of these primitives for shape recovery. The class we are using is surfaces of revolution, which are RLSHGC's with a circular cross-section function (no linear scaling of the cross section is required). These surfaces may be completely described by the rotation of a plane curve about the axis of symmetry. If we take this axis to be the Z axis, the surface will intersect the O–XZ plane in the plane curve.

$$r_c(u) = p(u)\mathbf{i} + z(u)\mathbf{k} \qquad (2)$$

and the surface of revolution has the equation

$$r = p(u)\cos\theta\mathbf{i} + p(u)\sin\theta\mathbf{j} + z(u)\mathbf{k}. \qquad (3)$$

From the definition, we can see that the cross-section curve is circular, but the contour generator curve can be quite arbitrary.

Fearing was able to recover the shape of generalized cylinders from a minimal set of tactile contacts that were static in nature and not dynamic. Our attempt to recover the shape of these objects from touch is based on receiving less-accurate tactile information than the system used by Fearing. In particular, our tactile sensors are not capable of reporting accurate surface normals or surface curvatures. They are, however, capable of producing localized 3-D positional contacts of fin-

gers to the object. Our experience is that information such as surface curvature is very difficult to accurately sense, particularly with sensors that have low dynamic range and are used in an active, exploratory manner. Our approach is to use the contour follower to recover the contour curve described above. If we obtain two such contour curves that are on either side of the object, we can estimate the axis of the surface of revolution and recover the shape. This procedure maps naturally into a two-fingered contour follower EP in which an object's contour on either side is sensed using the thumb and index finger.

The problem of using a tactile device to trace a surface on an object is a complicated one. Previous work by Allen [1], using a one-fingered tactile sensor mounted on a PUMA, traced along a curved surface by calculating a weighted vector of constraint direction that tried to follow the surface curvature while preserving smoothness of the trace and a constraint having to do with creating regions bounded by traces that were equivalent in size. Hor [17] traced contours of planar objects using a planar four-fingered "chopstick" like manipulator. Strain gauge sensors on the fingers of this device would calculate surface normals and move tangentially along a surface, recording the contour. Stansfield [34] experimented with active tactile sensing to determine an object's shape using a single tactile sensor mounted on a robot arm. Following Klatzky and Lederman, Stansfield's system also successfully determined several other object properties that included compliance, elasticity, and texture.

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

The detection of contact and conversion to Cartesian coordinates is a process that requires several steps. The fingers are moved toward the object in a number of discrete intervals. After each movement, two checks are performed. First, did the tactile sensor detect contact, and second, did the finger move the entire distance that was commanded? If the tactile sensor detects contact, the location of the center of the contact region is then found. To find the center of the contact, the first moments of the array are taken. A transformation is then performed from the fingertip coordinate frame to the hand coordinate frame and finally from the hand coordinate frame to world coordinates. The second check is that the finger does not move the entire distance commanded (and there is no tactile contact). This event would signal that something is impeding a finger from moving. In this case, no centroid of the contact
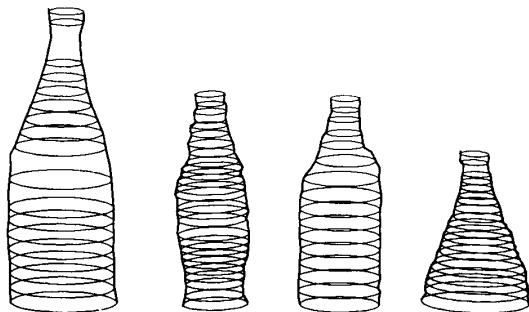
Fig. 6. Recovered solids of revolution from surface contour explorer (left to right; wine bottle, coke bottle, beer bottle, Orangina bottle).

region is found, and the data point is thrown out. Currently, after detecting contact that does not involve the tactile sensor, the exploratory procedure continues looking for valid contact points along the original search axis.

We have performed a series of experiments that try to recover the shape of a number of different solids of revolution including a wine bottle, a beer bottle, a coke bottle, and an Orangina soft drink bottle (a flask-like object). The procedure begins by exploring the object along an exploration axis that is assumed to be perpendicular to the support table (but can be inferred from vision sensing described below). The points generated from these contour traces are then linked into a set of linear contour segments. Circular cross-section curves are then fit perpendicular to the exploration axis, including trace points from each of the contours. The recovered shapes are shown in Fig. 6. The shapes are clearly distinguishable from this sparse data. Additional and important discriminating characteristics are actual 3-D size and volume, and these are calculable from these representations.

V. DISCUSSION

Although the focus of this paper is on the acquisition and interpretation of touch sensor data, our overall approach to the problem of robotic object recognition lies in a multisensor approach; we believe no single sensing modality is currently powerful enough to robustly perceive and recognize its environment. Just as humans exploit a multitude of sensor systems, robotic systems need to use multiple sensors for perception as outlined in Allen [1] and Kak and Chen [21]. A central idea in using multisensor data is that overreliance on one sensor can cause error. It has been empirically observed that trying to extract too much information from a single sensing modality results in a degradation of results; however, using only the most reliable and highest confidence sensor data allows one to proceed along a path that is known to be correct. We call this principle "less is more" in that reduced amounts of reliable data from a single sensor are more useful than large amounts of data that may be spurious. By combining the data that is most reliable from each of a number of sensors, more accurate results may be computed.

An example of this approach is in determining the exploration axis for the contour follower EP. Knowing in which direction to trace the object is important to higher level recovery procedures that need to use this information in the recognition process. Once the hand makes contact with the

object, it explores the contour along a known axis, which we calculate *a priori*. We are currently implementing a vision-based technique to determine this axis. Our method of visual recovery of the exploration axis exploits the recent work of Wolff [39] in stereo line matching. Point-based stereo techniques tend to be unreliable in that multiple correspondences between images can cause mismatches and error. More stable matching can occur using larger primitives such as lines [16]. Even using line-based matching, problems can still occur. Matching the end points of lines can be prone to errors in the output of the line finder, which may break a single line into multiple segments due to differing edge strengths along the line. The problem here is that 3-D depth is being computed, which requires an absolute correspondence of points (whether from point-based or line-based methods).

Our method alleviates this dependence on absolute matching of unstable primitives to generate 3-D depth. All we require of the algorithm is an *orientation* vector in 3-D. We do not need to have its absolute depth, but we do need to generate a match between a family of parallel lines sharing the same orientation. This orientation can then be used by the active hand as the exploration axis. The 3-D depth has already been determined from the contact of the hand with the object. Given this 3-D depth from tactile contact, we can follow the 3-D axis determined by the line-based stereo matcher to continue our exploration.

It is important to note that this method is less sensitive to matching errors and baseline measurement, which is another common cause of stereo error. In addition, it is also less prone to the effects of physical point mismatches as the baseline increases since we are still matching a larger entity, which is the line itself. Intuitively, the method creates a 3-D plane in space from the camera center and any two points on the line. This plane and a similar plane from the other camera are all that are needed to create a 3-D intersection line, which we can use as the exploration axis.

Because there are many lines in a scene, we have to choose a criterion for deciding those lines that constitute the axis of the object. For exploration purposes, we simply want to discover a maximum length line that will serve as an axis. In most cases, this is part of the visual occluding contour of the object, which is exactly the axis we desire for active tactile exploration.

VI. SUMMARY

We have described a set of exploratory procedures using touch sensing that can serve as a front end to a multisensor object recognition system. The EP's can be used in a coarse-to-fine sensing strategy that tries to build shape descriptions at a number of levels. An important feature of this system is the multiple representations used in recovering and reasoning about shape. The first EP, grasping by containment, uses a global volumetric recovery method that is stable and efficient with extremely sparse amounts of data. It can be used as a precursor to more detailed fine shape recovery using either the planar surface explorer or a 3-D surface contour EP that can be used to recover solids of revolution.

In the future, we hope to link all the exploratory procedures

404

IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, VOL. 6, NO. 4, AUGUST 1990

into a fully autonomous system that will be able to use gross object structure as a generator of sensing hypothesis for the finer level EP's. In this way, we hope to be able to recover the shape of more complex objects using tactile and visual processing.

ACKNOWLEDGMENT

Thanks to the referees for improving the paper and K. Roberts for helping acquire the superquadric recovery data and helping to set up and calibrate the hand.

REFERENCES

[1] P. Allen, "Integrating vision and touch for object recognition tasks," *Int. J. Robotics Res.*, vol. 7, no. 6, pp. 15–32, 1988.
[2] P. Allen, P. Michelman, and K. S. Roberts, "An integrated system for dextrous manipulation," in *Proc. IEEE Conf. Robotics Automat.* (Scottsdale, AZ), May 15–19, 1989, pp. 612–617.
[3] P. Allen and K. S. Roberts, "Haptic object recognition using a multi-fingered dextrous hand," in *Proc. IEEE Conf. Robotics Automat.* (Scottsdale, AZ), May 15–19, 1989, pp. 342–347.
[4] R. Bajcsy and F. Solina, "Three dimensional object representation revisited," in *Proc. Int. Conf. Comput. Vision* (London), June 1987.
[5] D. Ballard and C. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
[6] A. Barr, "Superquadrics and angle preserving transformations," *IEEE Comput. Graphics Appl.*, vol. 1, pp. 11–23, 1981.
[7] T. Binford, "Visual perception by computer," in *Proc. IEEE Conf. Syst. Contr.* (Miami), 1971.
[8] R. Bolles and A. Bobick, "Exploiting temporal coherence in scene analysis for automated navigation," in *Proc. IEEE Conf. Robotics Automat.* (Scottsdale, AZ), May 15–19, 1989, pp. 990–996.
[9] T. Boult and A. Gross, "On the recovery of superellipsoids," in *Proc. DARPA Image Understanding Conf.* (Cambridge, MA), April 1988, pp. 1052–1063.
[10] R. Ellis, E. Riseman, and A. R. Hanson, "Tactile recognition by probing: Identifying a polygon on a plane," in *Proc. AAAI-86* (Philadelphia), Aug. 11–15, 1986, pp. 632–637.
[11] R. Fearing, "Tactile sensing for shape interpretation," in *Dextrous Robot Hands* (T. Iberall and S. T. Venkataraman, Eds.). New York: Springer-Verlag, 1989.
[12] W. E. L. Grimson and T. Lozano-Perez, "Model based recognition and localization from sparse three dimensional sensory data," *A.I. Memo 738*, M.I.T. A.I. Laboratory, Cambridge, Aug. 1983.
[13] L. Harmon, "Tactile sensing for robots," in *Recent Advances in Robotics* (G. Beni, Ed.). New York: Wiley, 1985, pp. 389–424.
[14] V. Hayward and R. Paul, "Robot manipulator control under UNIX," in *Proc. 13th ISIR* (Chicago), April 17–21, 1983, pp. 20:32–20:44.
[15] T. C. Henderson and B. Bhanu, "Three point seed method for the extraction of planar faces from range data," in *Proc. IEEE Workshop Ind. Appl. Machine Vision* (Research Triangle Park, NC), May 1982, pp. 181–186.
[16] K. Henriksen, "Line-based stereo matching," MS-CIS-87-52, Grasp Lab 109, Dept. Comput. Inform. Sci., Univ. Pennsylvania, Philadelphia.
[17] M.-K. Hor, "Control and task planning for a four finger dextrous manipulator," Ph.D. thesis, Courant Inst. New York Univ., New York, NY, Oct. 1987.
[18] B. K. P. Horn, *Robot Vision*. Cambridge, MA: M.I.T. Press, 1986.
[19] S. C. Jacobsen, E. K. Iversen, D. F. Knutti, R. T. Johnson, and K. B. Biggers, "Design of the Utah/MIT dextrous hand," in *Proc. IEEE Conf. Robotics Automat.* (San Francisco), Apr. 7–10, 1986, pp. 1520–1532.
[20] R. S. Johansson and A. B. Vallbo, "Tactile sensory coding in the glabrous skin of the human hand," *Trends Neurosci.*, vol. 6, pp. 27–32, 1983.
[21] A. Kak and S. Chen (Eds.), *Proceedings of the AAAI Workshop on Spatial-Reasoning and Multisensor Integration*. Los Altos, CA: Morgan Kauffman, Oct. 1987.
[22] R. Klatzky and S. Lederman, "The intelligent hand," in *The Psychology of Learning and Motivation* (G. Bower, Ed.). New York: Academic, 1987, vol. 21, pp. 121–151.
[23] R. Klatzky, S. Lederman, and V. Metzger, "Identifying objects by touch: An 'expert system,'" *Perception Psychophys.*, vol. 37, pp. 299–302.
[24] S. Lederman and R. Klatzky, "Hand movements: A window into haptic object recognition," *Cognitive Psych.*, vol. 19, pp. 342–368, 1987.
[25] S. J. Lederman and R. A. Browse, "The physiology and psychophysics of touch," in *Sensors and Sensory Systems for Advanced Robots*. New York: Springer-Verlag, NATO-ASI series.
[26] C. Muthukrishnan, D. Smith, D. Myers, J. Rebman, and A. Koivo, "Edge detection in tactile images," in *Proc. IEEE Int. Conf. Robotics Automat.* (Raleigh), 1987, pp. 1500–1505.
[27] V. Nalwa, "Line drawing interpretation: Bilateral symmetry," in *Proc. DARPA Image Understanding Workshop* (Los Angeles, CA), Feb. 1987, pp. 956–967.
[28] S. Narasimhan, D. M. Siegel, and J. M. Hollerbach, "Condor: A revised architecture for controlling the Utah-MIT hand," in *Proc. IEEE Conf. Robotics Automat.* (Philadelphia), April 24–29, 1988, pp. 446–449.
[29] A. P. Pentland, "Recognition by parts," Tech. Rep. 406, SRI International, Menlo Park, CA, Dec. 16, 1986.
[30] J. Ponce, D. Chelberg, and W. Mann, "Invariant properties of the projections of straight homogeneous cylinders," in *Proc. First. Int. Conf. Comput. Vision* (London), 1987.
[31] K. S. Roberts, "Robot active touch exploration: Constraints and strategies," in *Proc. IEEE Int. Conf. Robotics Automat.* (Cincinnati), May 13–18, 1990, pp. 980–985.
[32] F. Solina, "Shape recovery and segmentation with deformable part models," Ph.D. dissertation, Dept. Comput. Sci., Univ. Pennsylvania, Philadelphia, Dec. 1987.
[33] T. Speeter, "Flexible piezo-resistive touch sensing array," in *Proc. SPIE Conf. Optics, Illumination Image Sensing Machine Vision III* (Cambridge), Nov. 1988.
[34] S. Stansfield, "Visually-guided haptic object recognition," Ph.D. dissertation, Dept. Comput. Inform. Sci., Univ. Pennsylvania, Philadelphia, Oct. 1987.
[35] K. Stevens, "Computation of locally parallel structure," *Biological Cybern.*, no. 29, pp. 19–28, 1978.
[36] B. Tise, "A compact high resolution piezo-resistive digital tactile sensor," *IEEE Conf. Robotics Automat.* (Philadelphia), Apr. 24–29, 1988, pp. 760–764.
[37] H. van Brussel and H. Belien, "A high resolution tactile sensor for part recognition," in *Proc. 6th Int. Conf. Robot Vision Sensory Contr.* (Paris), June 1986.
[38] A. Witkin, "Scale-space filtering," in *Proc. IJCAI*, 1983, pp. 1017–1022.
[39] L. B. Wolff, "Measuring the orientation of lines and surfaces using translation invariant stereo," in *Proc. SPIE Conf. Sensor Fusion* (Cambridge, MA), Nov. 1988, vol. 1003.

Peter K. Allen (M'85) received the A.B. degree from Brown University, Providence, RI, in mathematics-economics, the M.S. degree in computer science from the University of Oregon, Eugene, and the Ph.D. degree in computer science from the University of Pennsylvania, Philadelphia.

He is currently Assistant Professor of Computer Science at Columbia University, New York, NY. His current research focuses on using dextrous robotic hands for object recognition and task-level manipulation. Other research interests include developing programming paradigms for sensor-based robotics, 3-D modeling, and real-time motion tracking. In recognition of his work, he has been named a Presidential Young Investigator by the National Science Foundation. At the University of Pennsylvania, he was the recipient of the CBS Foundation Fellowship, the Army Research Office Fellowship, and the Rubinoff Award for innovative uses of computers.

Dr. Allen is a member of ACM, IEEE, and AAAI.

Paul Michelman (S'87) received the B.A. degree in classics from Grinnell College, Grinnell, Iowa, the B.E.E.E. degree from the City College of New York, and the M.S. degree in computer science from Columbia University, New York, NY. He is currently working toward the Ph.D. degree in computer science at Columbia University. His research interests include robotics, tactile sensing and dexterous manipulation.

Mr. Michelman is a member of Tau Beta Pi.