2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Tactile Object Class and Internal State Recognition for Mobile Manipulation

Sachin Chitta[1] Matthew Piccoli[2] Jürgen Sturm[3]

***Abstract*— Tactile information is valuable in determining properties of objects that are inaccessible from visual perception. In this work, we present a tactile perception strategy that allows any mobile robot with tactile sensors in its gripper to measure a set of generic tactile features while grasping an object. We propose a hybrid velocity-force controller, that grasps an object safely and reveals at the same time its deformation properties. As an application, we show that a robot can use these features to distinguish the open/closed and fill state of bottles and cans – purely from tactile sensing – from a small training set. To prove that this is a hard recognition problem, we also conducted a comperative study with 17 human test subjects. We found that the recognition rate of the human subjects were comparable to our robotic gripper.**

I. INTRODUCTION

Humans have a remarkable sense of touch that enables them to explore their environment in the finest detail [8]. Haptic feedback provides subtle cues about the environment that cannot be obtained from any other perceptual sensors. Tactile feedback allows us to localize objects in our hand, determine their rigidity as well as other material properties, and even their identity. Consider, for example, the task of choosing fruit. The response of the fruit quickly lets us figure out whether it is ripe. This is particularly true for fruits like peaches or avocados whose color is often not a good indicator of their ripeness.

Tactile sensors provide robots an additional means of sensing the environment. They range from simple contact sensors that provide binary information about whether the robot is in contact with the environment to more complex arrays of sensors that provide pressure sensing at a resultion comparable to human finger tips. Force-torque sensors mounted on the wrist of a robot are also often used to provide tactile information about the environment, but can only measure forces after an object has been safely grasped.

At the moment, the information provided by contemporary tactile sensors tends to be lower resolution and covers a smaller area compared to visual perception, and therefore requires different algorithms. Furthermore, the sensor must be actively controlled to explore the environment in contrast to visual sensors.

Over the past few years, several promising approaches have been developed on the technological or sensor side.

[1] Sachin Chitta is with Willow Garage Inc., Menlo Park, CA 94025, USA sachinc@willowgarage.com

[2] Matthew Piccoli is with the Grasp Laboratory, University of Pennsylvania, Philadelphia, PA 19104 picolli@me.upenn.edu

[3] Jürgen Sturm is with the Autonomous Intelligent Systems Lab, Computer Science Department, University of Freiburg, Germany sturm@informatik.uni-freiburg.de

Fig. 1. **Left:** A mobile manipulation robot grasping a bottle estimates the internal state of the grasped object from its tactile appearance. **Right:** Comparative study on tactile performance with human test subjects.

Artificial skins that measure orthogonal pressure at spatial and temporal resolutions comparable to human skins are often composed of elastic, conductive or resistive polymers, which change their electrical properties depending on the applied pressure. They can, in principle, be manufactured to cover larger parts of a robot at relatively low cost. Several research groups have reported [7], [9], [10] success in wrapping substantial parts of the surface of their robots using such sensors, for example, to ease human-machine interaction or to improve the robustness of object manipulation tasks. Tactile sensing has most often been used for object recognition or localization.

In this work, we will present a novel approach that uses the tactile information for estimating the state of the object to be manipulated, before the object actually is lifted. We gain this information from the temporal response of the object to an applied force profile during grasping. This information can then be used both for low-level motor control or higher-level motion planning algorithms with the aim to make object manipulation more robust. For example, an appropriate gripping force, manipulation speed, or other constraints can be imposed, in order to neither spill liquid nor crush the bottle.

We apply this approach to the particular case of discriminating among different types of liquid containers, like bottles and cans. We show excellent results in terms of discriminating among the container classes. Furthermore, we show that we can estimate the internal state of the object, indicating whether it is closed or open and empty or full. We compare the results from these experiments to human-study experiments, where human subjects were asked to discriminate among the same set of objects.

A. Related Work

Several studies have shown that humans are very good at modulating the applied grasp force in relation to the expected load force [8]. Even during dynamic motions such as walking or running, humans always apply the minimum force required to hold an object safely. These coordinative constraints simplify the control by reducing several degrees-of-freedom during the manipulation tasks. Tactile perception hereby plays an essential role: In experiments with humans, it was shown that the test subjects exerted much more gripping force than actually was needed when their fingertips are anesthetized, even if visual feedback was available [14].

Different tactile sensor principles for robots have been explored in the past, such as pressure-sensitive conductive polymers [26], piezo-resistive sensors [6], piezo-electric vibration sensors [15], and capacitive sensors which can additionally measure sheer force [3] or temperature [2]. A good survey on the different approaches can be found in [25].

Tactile sensors have been used in the past to explore the 3D shape of objects [1]. Others have used tactile sensors to detect ridges and bumps in the material [16] by sliding the robotic finger over an object to estimate the material from the obtained frequency spectrum [4]. Sensors based on piezo-electric vibration have been used to determine the hardness or softness of probed biomaterials [17]. Force-sensitive fingers have been used to control the robot's position [5], i.e., to continuously keep the finger in physical contact while moving the object. It has also been shown that tactile sensors can be used to estimate the 3D pose of objects with known shapes [18]. Russel *et. al.* used a tactile sensor matrix for object classification [22]. Their approach extracts geometric features like point, line, or area contacts and integrates them over time to classify the objects into generic classes such as boxes, spheres, cylinders, etc. Later, Russel [23] showed that a similar approach can also be used for object classification using an 8-whisker tactile sensor on a robotic gripper. Recent work on object recognition [24] used tactile sensing to gain low-resolution intensity images of an object. These images were then used in a *bag of features* approach from computer vision to recognize objects of different shapes and sizes. The authors also presented an active controller that could determine the next best action to execute to maximize the information gained.

A few prior works [12], [13] exist on estimating the friction coefficients such that slippage and crushing are avoided. Maeno [12] gives a good overview over existing techniques and describes how their system estimates these values from the tangetial forces while pushing a tactile sensor into a surface.

Our approach differs from previous approaches as our aim of estimation is different. To the best of our knowledge, we are the first to estimate the state of a container based on tactile observations. We propose a small set of generic tactile features that can easily be computed for a gripper that is equipped with tactile sensors. Another contribution of this paper is that we provide a human study in which we asked human subjects to perform the same recognition tasks as the robot. The results of this study illustrate the difficulty of the recognition task.

II. Approach

A. Requirements

For our approach, we assume a mobile manipulator with a force-sensitive gripper. At each point in time, we assume that the gripper can measure its position $p(t) \in \mathbb{R}$, velocity $\dot{p}(t) \in \mathbb{R}$ and the force $f(t) \in \mathbb{R}$ sensed by fingertip sensors. In this section, we formulate our approach based on the presence of this set of sensor inputs.

We also assume the existence of a controller that can apply the required force profile and measure the necessary features. An important requirement for this controller is that it should not damage the objects but still should be able to grasp the object firmly. These requirements necessiate the design of a specialized controller, that we describe in Section III.

B. Feature selection

Consider the schematic drawing in Fig. 2 of force-distance profiles that we obtain while grasping objects using our controller described in Section III. From preliminary experiments, we identified two important points in time: the moment the gripper makes first contact with the object t_{first} and the time t_{steady} after which the sensor values have converged (and the grasp is stable). In practice, we require for the first contact detection that both fingers are in contact with the object, i.e., that the force measurement of both fingers is above a certain threshold τ_1. With t_{steady}, we then denote the point in time where the gripper comes to rest, i.e., its velocity drops below a certain threshold τ_2.

$$t_{\text{first}} = \quad \arg\min_{t} |f(t)| < \tau_1 \tag{1}$$

$$t_{\text{steady}} = \quad \arg\min_{t > t_{\text{first}}} |\dot{p}(t)| < \tau_2 \tag{2}$$

At moment t_{first}, we extract the first contact distance $p_{\text{first}} = p(t_{\text{first}})$. Note that this is a measure of the uncompressed size of the object. The second feature is the distance between the two fingertips after the gripper has compressed the object fully. We label this the steady state distance

$$p_{\text{steady}} = p(t_{\text{steady}}). \tag{3}$$

Note that this distance is a function of both the material and geometric properties of the object and of the internal state of the object, i.e. whether the object is open or closed and full or empty.

Another useful feature is the time that it takes between making contact with the object and coming to a rest, denoted by

$$\Delta t = t_{\text{steady}} - t_{\text{first}}. \tag{4}$$

Two other useful features are the average velocity $\Delta p / \Delta t$ of compression and the average rate of change of the fingertip

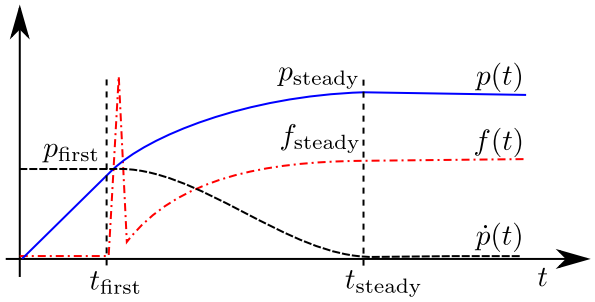

Fig. 2. The set of features chosen for use in the recognition task.

feature	description
p_{first}	the first contact distance
p_{steady}	distance after which grasping is complete
f_{steady}	force sensed after grasping has completed
Δt	duration of the grasping
$\Delta p/\Delta t$	average compression velocity
$\Delta f/\Delta t$	average compression ratio

TABLE I
GENERIC SET OF FEATURES THAT CAN BE USED TO CLASSIFY AN OBJECT BEING GRASPED.

center sensor force $\Delta f/\Delta t$, which can be computed from the features from above as follows:

$$\Delta p/\Delta t = (p_{\text{steady}} - p_{\text{first}})/\Delta t \quad (5)$$
$$\Delta f/\Delta t = (f_{\text{steady}} - f_{\text{first}})/\Delta t. \quad (6)$$

The average velocity $\Delta p/\Delta t$ represents the rate at which the object gets compressed and can differ based on the material properties and the geometry of the object. Equivalently, $\Delta f/\Delta t$ could be thought of as representing an average compression ratio.

We have now defined a set of 6 generic features that can be easily extracted by the robot while grasping an object, see Tab. I. We do not claim that this list is complete, but as we will show in Section III, we were able to recognize the object class and its internal state from these features alone.

C. Training data

We have gathered data for a large number of different objects. For each trial, we obtained measurements for our 6-dimensional feature vector $\mathbf{a} \in \mathbb{R}^6$, i.e.,

$$\mathbf{a} = (p_{\text{first}}, p_{\text{steady}}, f_{\text{steady}}, \Delta t, \Delta p/\Delta t, \Delta f/\Delta t)^T, \quad (7)$$

and a label $c \in C$ containing the object's class and internal state. As a result, we get a training database $\mathcal{D}$ containing a set of attribute-class tuples $(\mathbf{a}, c)$.

D. Decision tree classifier

Subsequently, we have applied a C4.5 decision tree classifier on our training data [27]. We have also tried other supervised classifiers, like support vector machines and neural networks, from which we obtained similar (or slightly worse) results. The reason for this might be that all algorithms are able to extract almost the same amount of data from the training set. The advantage of decision trees over other classifiers is that the learned concepts can intuitively be interpreted.

The C4.5 decision tree classifier [21] is an extension of the ID3 algorithm. In addition to that, C4.5 can deal with continuous attributes.

Decision tree induction is an iterative process: it starts by selecting an attribute that most effectively splits the data according to their data classes. Typically, the information gain (which is the reduction in entropy) is used as a measure for selecting the split. The entropy H of a set $\mathcal{D}$ is defined as

$$H(\mathcal{D}) \quad = -\sum_{c \in C} p(c) \log p(c), \quad (8)$$

where $p(c)$ is the probability of target class c in the training set, i.e.,

$$p(c) = \frac{1}{|\mathcal{D}|} \sum_{(\mathbf{a},c) \in \mathcal{D}} 1. \quad (9)$$

As all our attributes are continuous, a split s is defined by a split value s_{value} for a particular split attribute s_{attr}, i.e., the training set $\mathcal{D}$ is divided into two subsets

$$\mathcal{D}_{\leq} \quad := \{(\mathbf{a}, c) | a_{s_{\text{attr}}} \leq s_{\text{value}}, (\mathbf{a}, c) \in \mathcal{D}\} \quad (10)$$
$$\mathcal{D}_{>} \quad := \{(\mathbf{a}, c) | a_{s_{\text{attr}}} > s_{\text{value}}, (\mathbf{a}, c) \in \mathcal{D}\}. \quad (11)$$

From all possible splits, C4.5 now selects the one with the highest information gain, i.e.,

$$s = \arg\max_{s \in S} IG(\mathcal{D}, s), \quad (12)$$

where the information gain (IG) is defined as the reduction in entropy of the resulting sets compared with the initial set:

$$IG(\mathcal{D}, s) := \quad H(\mathcal{D}) - H(\mathcal{D}|s), \quad (13)$$

where the conditional entropy $H(\mathcal{D}|s)$ is defined as

$$H(\mathcal{D}|s) = \quad H(\mathcal{D}_{\leq})p(\leq) + H(\mathcal{D}_{>})p(>). \quad (14)$$

Each split s corresponds to a node of the decision tree with two children. The same procedure is then repeated for the resulting subsets $\mathcal{D}_{\leq}$ and $\mathcal{D}_{>}$, until the leafs are homogeneous with respect to the target class, i.e., the entropy in the dataset of the leaf is zero.

Another important step after training is pruning, to avoid overfitting to the training data. This is done by replacing a whole subtree by a leaf node if the expected error rate (computed on a test dataset hold out during training) in the subtree is greater than in the single leaf [27].

III. EXPERIMENTS

A. Hardware

The hardware used for the experiments in this paper is part of the PR2 personal robot. The PR2 is a two-armed robot with an omni-directional base. It has an extensive sensor suite useful for mobile manipulation including a tilting laser scanner mounted on the head, two pairs of stereo cameras,

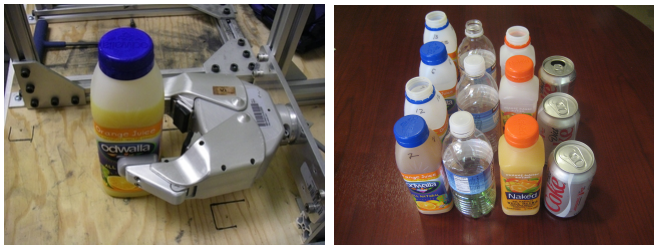

Fig. 3. **Left:** The experimental setup showing the gripper mounted on a stand. **Right:** Some of the bottles and cans used for our experiments. From left to right: Odwalla fruit juice bottle, water bottle, Naked fruit juice bottle, Coke can.

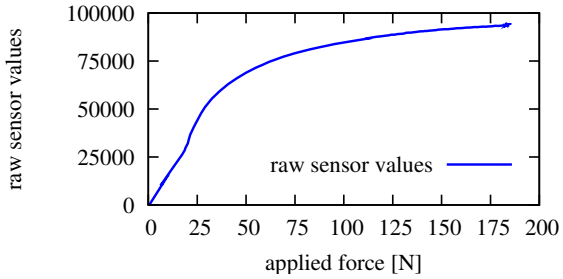

Fig. 4. Calibration data relating raw sensor values to forces calibrated using a load cell.

an additional laser scanner mounted on the base and an IMU mounted inside the body. Encoders on each joint also provide continuous joint angle information. Each arm also has an additional camera mounted on the forearm. The camera is mounted so that an object grasped by the hand is within its field of view.

Figure 3 shows the parallel plate gripper mounted on both arms of the robot. The gripper has a single actuator, a brushless DC motor with a planetary gearbox and an encoder. The transmission function between the rotary displacement of the actuator and the lateral displacement of the two fingers is known. A separate load cell was used to calibrate the gripper. It can apply a maximum force of 200 N but is software limited to 100 N. This was also approximately the amount of force that a human can apply by pinching his/her forefinger and thumb together. Each gripper has a capacitive sensor consisting of 22 individual cells mounted on each fingertip. A 5×3 array is mounted on the parallel gripping surface itself while 2 sensors are mounted on the tip of the fingertip and one on the back. Two sensor arrays with 2 sensors each are mounted on each side of the fingertip.

The capactive pressure sensors respond to the normal forces exerted on them. They were calibrated by using a load cell to measure the net force applied by the gripper and comparing it to the net force sensed by the fingertip sensors. The calibration curve for the two fingertip sensors is shown in Figure 4. Here, the raw sensor values for the fingertip sensors are computed by adding up the sensor values for the individual sensors on the grasping face of the sensor. The calibration data is stored in a lookup table that can be used to convert raw sensor values to the actual force sensed by the fingertip sensors.

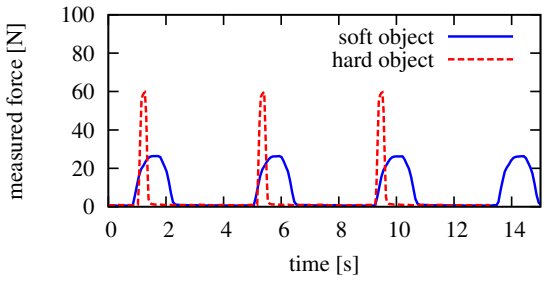

Fig. 5. Measured net fingertip force (N) for grasping a wooden block and a rubber toy when using a pure force controller. The high impact forces can destroy delicate, but rigid objects, like eggs.

B. Controller

We explored different controllers for the gripper to achieve the objective of grasping objects without crushing them. A pure velocity controller $c_{\text{velocity}}(\dot{p}(t), t)$ makes the gripper approach an object slowly, but after it contacts the object, it increases its force output in order to establish a constant velocity $\dot{p}_{\text{target}}$, and thereby crushes the object.

Another option is to use a force controller $c_{\text{force}}(f(t), t)$. Such a controller can hold an object in the hand firmly, by trying to apply a constant force f_{target}. With a constant force controller, the gripper continuously accelerates until contact is achieved. This can lead to high velocities at impact. As an example, see Figure 5, where the the gripper was grasping a very rigid object (here, a wooden block). The significant impact force applied to the object on contact can easily damage rigid, but delicate objects, like eggs [19]. Of course, the applied constant force could be reduced to deal with such cases. In practice, however, if the commanded force is below the force required to overcome static friction, the gripper does not move at all.

Driven by these considerations, we chose to create a compound controller: first, we close the gripper slowly around an object using the velocity controller until it makes contact with the object. Then, we switch seamlessly to the force controller in order to close gently to measure the object's deformability properties. This hybrid controller has two parameters: both the initial velocity $\dot{p}_{\text{target}}$ and its probing force f_{target} have influence on the executed grasp.

$$c_{\text{grasping}}(t) = \begin{cases} c_{\text{velocity}}(t, \dot{p}(t)) & \text{while } f(t) = 0 \\ c_{\text{force}}(t, f(t)) & \text{thereafter} \end{cases} \quad (15)$$

The result of the hybrid velocity-force controller can be seen in Figure 6. Here, a wooden block was grasped by the gripper using the new controller. The peak force acting on the object is significantly lower. Further, this controller was successful in grasping eggs without crushing them. A movie showing the comparison between an open-loop effort controller and the closed-loop grasping controller can be seen at [19].

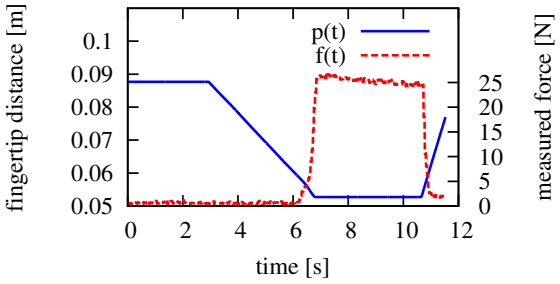

Fig. 6. Fingertip distance and fingertip force vs. time plots showing the reduced impact forces using the hybrid controller. First contact happens at about $t_{\text{first}} = 6$ seconds followed quickly by a steady state at about $t_{\text{steady}} = 6.75$ s where fingertip distance stays constant. The probing effort is set to 23 N. Note that the fingertip force does not spike above the desired probing force on impact.

C. Experimental setup

Using the controller described in Section III, we proceeded to get data on different types of objects. We chose to work with a variety of liquid containers. We chose containers that contained 2 different brands of juices (Odwalla and Naked), water bottles and multiple soda cans (Figure 3(right)). The gripper was mounted on a stand so that it would stay immobile and could grasp each container at a fixed height above a planar table surface, see Figure 3 (left).

The acquisition of training samples started with the gripper fully open. The containers were placed one at a time between the gripper fingertips. The gripper was then closed using the hybrid force velocity controller described earlier. As the gripper closed, it was possible that one of the fingertips came into contact with the object before the other one. This was taken into account in the controller, which always looked for contact on both fingertips before concluding that the object was in the initial state of grasping. Contact was determined by using a threshold value on the net force sensed by each fingertip. On contact, control was switched to a constant force control. Once the gripper came fully to rest, the controller waited for a small interval of time before opening the gripper fully. During each trial, the features described in Sec. II were extracted and written to a file.

D. Objects and Internal States

The container classes present in the training set are the following:

1) Odwalla fruit juice bottles
2) Naked fruit juice bottles
3) Soda cans
4) Water bottles

The internal states that we would like to differentiate are the following:

1) closed and full
2) closed and empty
3) open and full
4) open and empty

TABLE II

CONFUSION MATRIX FOR RECOGNIZING THE CLASS OF THE CONTAINER, WITH $f_{\text{TARGET}} = 20$ N. THE RECOGNITION RATE IS 93.9 %.

a	b	c	d	
58	1	0	1	a = Odwalla fruit juice bottles
8	40	0	0	b = Naked fruit juice bottles
0	0	41	3	c = Softdrink cans
0	0	1	76	d = Water bottles

E. Data acquisition using the robotic gripper

With our robotic gripper, we collected data for each of the internal states for each container class. We carried out a total of 66 trials with 12 Odwalla fruit juice bottles in 4 different internal states, 80 trials with 16 water bottles in 4 different internal states, 42 trials with 12 cans that only had 3 different internal states, and 41 trials with 10 Naked fruit juice bottles. We used different instances of each container class in collecting the data to account for variations within a container class. We also rotated the containers around between taking measurements to take account of variations in surface properties of individual containers with changes in orientation with respect to the gripper fingertips. All this data was collected with the probing force set at 20 N.

We also collected a subsequent dataset just for the Odwalla fruit juice bottles using three different probing forces of 17, 20 and 23 N. This involved conducting 24 trials for each internal state for a total of 96 trials for all the 4 nodes for each probing force.

F. Classification results

The two aims for the classification task were: (a) recognize the different container classes and (b) recognize the internal state within each class that would indicate whether the container is full or empty and open or closed.

To test our classifier we used ten-fold cross validation for each experiment, i.e., first the stratified dataset was divided into 10 parts. Then we learned the classifier on 9 parts, and used it subsequently to classify the test instances from the remaining part. This was repeated each of the ten folds, such that we ended up with target class predictions for all instances in the dataset. Finally, the predictions were compared to the true target class, and the recognition rate was computed as the ratio between correctly and incorrectly instances.

Table II shows the confusion matrix for recognizing the different container classes. Note that the size p_{first} of the different containers is a discriminative attribute for the different containers, yet two sets of containers (Naked and Odwalla fruit juice juice bottles) have similar sizes. They can, however, be distinquished by using f_{steady}, the net force reported by the fingertip sensors at full compression.

As the results show, our approach had a 93.9 % accuracy in recognizing the different sets of liquid containers. The most confusion was in discriminating between the Odwalla fruit juice bottles and Naked fruit juice bottles. Note that this recognition could have been easily performed using image based feature recognition as well. Our approach is not meant

TABLE III

RECOGNITION RATE OF THE INTERNAL STATE, PER CONTAINER CLASS, WITH $f_{\text{TARGET}} = 20\,\text{N}$.

Object Class	Recognition Rate
Odwalla fruit juice bottles	83.3%
Naked fruit juice bottles	58.3%
Softdrink cans	74.4%
Water bottle	32.5%

TABLE IV

RECOGNITION RATE, DEPENDING ON THE PROBING FORCE PARAMETER f_{TARGET}, FOR THE ODWALLA FRUIT JUICE BOTTLES.

f_{target}	Recognition Rate
17 N	69.8%
20 N	83.3%
23 N	94.8%

to compete with such approaches but is meant to complement other approaches and to be used for confirmation of the hypothesis from visual perception.

In the second part of our evaluation, we have evaluated the recognition rate of the internal state of a container, given its class (see Table III). We found, that the recognition rate strongly depends on the particular container. This result is not surprising, as obviously feeling the internal state of a container strongly depends on how well it manifests its internal state to the outside, i.e., in its tactile appearance. Interestingly, we found that the Odwalla bottles were separable the easiest. Their internal state was estimated correctly at 94.8%, compared to 58.3% for Naked bottles, 74.4% for cans and only 32.5% for water bottles. The reason for the low performance on water bottles could be that they are made of very flimsy plastic and tend to deform unpredictably.

We also found, that the recognition rate was a function of the parameters of our hybrid controller. While the influence of the initial grasping velocity $\dot{p}_{\text{target}}$ was negligable, we found that choosing a good probing force f_{target} could improve the recognition substantially, see Table IV. This parameter determines how hard the gripper probes into the object, and should therefore be carefully selected according to the object class. In the case of the Odwalla bottle, we found, for example, the stronger probing force of $\dot{p}_{\text{target}} = 23\text{N}$ to be more informative than weaker ones, yielding a mode recognition rate of 94.8%.

In a combined experiment, where we let the robot estimate both the container class and the object internal state except for water bottles (resulting in 11 possible combinations), we obtained a recognition rate of 63.8%.

The confusion matrix for the specific case of recognizing the internal state of an Odwalla bottle is shown in Table V.

It is interesting to note that the open and full bottle tends to be compressed for the longest time, i.e., Δt is large. The steady state force f_{steady} differentiates between the open and empty bottle and the empty and closed bottles while the steady state distance p_{steady} differentiates the closed and full bottle very easily. However, when we repeated this experiment with bottles that had been subjected to repeated compressions, the recognition rate decreased again to 81 %. This is not surprising considering that the classifier was trained on data from fresh bottles while the testing was now done with bottles that had been subject to repeated stress. A movie showing the system recognizing the internal state of a set of bottles using the learned classifier can be found in [20].

TABLE V

CONFUSION MATRIX OF OUR APPROACH FOR RECOGNIZING THE INTERNAL STATE OF AN ODWALLA FRUIT JUICE BOTTLE FROM THE TACTILE APPEARANCE USING A ROBOTIC GRIPPER ($f_{\text{TARGET}} = 23\,\text{N}$). THE RECOGNITION RATE IS 94.8%.

a	b	c	d	
24	0	0	0	a = full closed
0	20	1	3	b = empty open
0	0	24	0	c = full open
1	2	0	21	d = empty closed

IV. HUMAN STUDY

The experimental results show that the robot could do reasonably well in terms of recognizing both the container class in the first series of experiments and internal state of an object in a second series of experiments.

We designed a human study to compare the performance of the robot to that of humans for the internal state estimation problem. The study was designed to find out if, using only tactile feedback, humans could achieve comparable recognition rates for the task of recognizing the internal state of an object. Figure 1 (right) shows the experimental setup used for the human study. Test subjects were asked to recognize, using only tactile information from squeezing a bottle, the internal state of the bottle. They were provided the opportunity to train beforehand until they were confident about their ability to discriminate between the different internal states of the bottles. Each test subject was then asked to identify the internal state of 16 different bottles sequenced in a random order. The subjects were instructed not to move or pickup the bottles and could not see the bottles while they were grasping them. To simulate the two-fingered grasp used by the gripper, the test subjects were asked to use only their thumb and index finger for the grasping task. Additionally, noise-canceling headphones were used to minimize the sound cues that subjects could pick up. There were a total of 17 test subjects.

Table VI shows the overall confusion matrix for all the trials together. The average recognition rates for all the subjects was 75.2%. The highest recognition rate was for bottles that were full and closed. There was considerable confusion between the empty/closed and full/open bottles. Based on a questionnaire filled out by the subjects at the end of the test, we found that most subjects were using features similar to the ones chosen for the learning approach. The two most cited features by the human subjects were the total compression distance and but also the rate at which the bottle returns to its original shape. The second feature is easier for humans to detect than the robot since the

TABLE VI

OVERALL CONFUSION MATRIX FOR ALL HUMAN SUBJECTS FOR RECOGNIZING INTERNAL STATE OF AN ODWALLA FRUIT JUICE BOTTLE. THE RECOGNITION RATE IS 72.2%.

a	b	c	d	
48	8	5	0	a = empty open
5	41	1	3	b = empty closed
16	11	55	2	c = full open
2	8	7	63	d = full closed

grippers on the robot are not easily back-drivable. The most successful test subjects cited a different set of features to discriminate between the bottles. They used high-frequency feedback from tapping the bottle with their fingers to detect the presence or absence of liquid in the bottle. At present the sensors on the robot are inadequate to gain this kind of information.

V. CONCLUSION

In this paper, we have presented an approach for estimating the internal state of objects being grasped by a mobile manipulator by using tactile sensors. We proposed a set of simple features that can be easily extracted from tactile sensing and proprioception. In experiments carried out on real data, we have shown that both the object class as well as its internal state can be estimated robustly. In a direct comparison experiment, we have shown that the robot's performance is of the same magnitude as human performance. The robot's performance improved with an increase in the magnitude of the probing force used for detection, presumably since this rendered the features used for detection more discriminative.

Despite these encouraging results, there are several directions for future research. Detection would clearly benefit from better sensors, for example with the ability to sense lateral forces with higher temporal resolution. An interesting approach that could be examined in the future includes the use of high-frequency signals and the haptic response from objects to such signals [11]. Temperature sensors in the gripper might also provide additional information to aid the recognition task.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the help of everyone at Willow Garage.

REFERENCES

[1] P. K. Allen. Integrating Vision and Touch for Object Recognition Tasks. *The International Journal of Robotics Research*, 7:15–33, 1988.

[2] F. Castelli. An integrated tactile-thermal robot sensor with capacitive tactile array. *IEEE Transactions on Industry Applications*, 38, 2002.

[3] C. Chuang and R. Chen. 3D capacitive tactile sensor using DRIE micromachining. In C. Cane, J.-C. Chiao, and F. Vidal Verdu, editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5836, pages 719–726, July 2005.

[4] F. de Boissieu, C. Godin, B. Guilhamat, D. David, C. Serviere, and D. Baudois. Tactile texture recognition with a 3-axial force MEMS integrated artificial finger. In *Proc. of Robotics: Science and Systems (RSS)*, Seattle, USA, June 2009.

[5] Z. Doulgeri and S Arimoto. Force position control for a robot finger with a soft tip and kinematic uncertainties. *Robotics and Autonomous Systems*, 55(4):328 – 336, 2007.

[6] Y. Hasegawa, M. Shikida, T. Shimizu, T. Miyaji, H. Sasaki, K. Sato, and K. Itoigawa. Micromachined active tactile sensor for hardness detection. *Sensors and Actuators A: Physical*, 114(2-3):141 – 146, 2004.

[7] T. Hoshi and H. Shinoda. Robot skin based on touch-area-sensitive tactile element. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 3463–3468, 2006.

[8] R. S. Johansson. Sensory and memory information in the control of dexterous manipulation. *Neural Bases of Motor Behaviour*, pages 205–260, 1996.

[9] R. Kageyama, S. Kagami, M. Inaba, and H. Inoue. Development of soft and distributed tactile sensors and the application to a humanoid robot. In *Proc. of the IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, volume 2, pages 981 – 986, 1999.

[10] O. Kerpa, K. Weiß, and H. Wörn. Development of a flexible tactile sensor system for a humanoid robot. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2003.

[11] K. J. Kuchenbecker. Haptography: Capturing the feel of real objects to enable authentic haptic rendering. In *Proc. of the Haptic in Ambient Systems (HAS) Workshop*, 2008.

[12] T. Maeno. Friction estimation by pressing an elastic finger-shaped sensor against a surface. *IEEE Transactions on Robotics and Automation*, 20(2):222–228, 2004.

[13] R. Matuk Herrera. Multilayer perceptrons for bio-inspired friction estimation. In *Proc. of the Int. Conf. on Artificial Intelligence and Soft Computing (ICAISC)*, pages 828–838, 2008.

[14] J. Monzee, Y. Lamarre, and AM. Smith. The effects of digital anesthesia on force control using a precision grip. *Journal of Physiology*, 89(2):672–683, 2003.

[15] K. Motoo, T. Fukuda, F. Arai, and T. Matsuno. Piezoelectric vibration-type tactile sensor with wide measurement range using elasticity and viscosity change. *Journal of the Robotics Society of Japan*, 24(3), 2006.

[16] A. M. Okamura and M. R. Cutkosky. Haptik exploration of fine surface features. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 1999.

[17] S. Omata, Y. Murayama, and C. E. Constantinou. Real time robotic tactile sensor system for the determination of the physical properties of biomaterials. *Sensors and Actuators*, 112(2-3):278 – 285, 2004.

[18] A. Petrovskaya, O. Khatib, S. Thrun, and A. Y. Ng. Bayesian estimation for autonomous object manipulation based on tactile sensors. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 707–714, 2006.

[19] M. Piccoli. Breaking eggs. `http://www.youtube.com/watch?v=fIsMCKOYFEY`, 2009.

[20] M. Piccoli. Recognizing the internal state of a bottle. `http://www.willowgarage.com/blog/2009/09/17/bottle-half-empty-half-full`, 2009.

[21] J. R. Quinlan. Learning with continuous classes. In *5th Australian Joint Conference on Artificial Intelligence*, pages 343–348, Singapore, 1992.

[22] R.A. Russell. Object recognition by a 'smart' tactile sensor. In *Proc. of the Australian Conf. on Robotics and Automation*, Melbourne, Australia, 2000.

[23] R.A. Russell and J.A. Wijaya. Object location and recognition using whisker sensors. In *Proc. of the Australian Conf. on Robotics and Automation*, 2003.

[24] A. Schneider, J. Sturm, C. Stachniss, M. Reisert, H. Burkhardt, and W. Burgard. Object identification with tactile sensors using bag-of-features. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2009.

[25] J. Tegin and J. Wikander. Tactile sensing in intelligent robotic manipulation - a review. *Industrial Robot: An International Journal*, 32, 2005.

[26] K. Weiss and H. Wörn. The Working Principle of Resistive Tactile Sensor Cells. *Proc. of the IEEE Int. Conf. on Mechatronics & Automation*, 2005.

[27] I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2005.