Tapping into Touch

Eduardo Torres-Jara* **Lorenzo Natale**** **Paul Fitzpatrick***

*Humanoid Robotics Group, CSAIL
Massachusetts Institute of Technology
32 Vassar St, Cambridge 02139
Massachusetts, USA

**LIRA-Lab, DIST
University of Genova
Viale F. Causa 13
16145 Genova, Italy

Abstract

Humans use a set of *exploratory procedures* to examine object properties through grasping and touch. Our goal is to exploit similar methods with a humanoid robot to enable developmental learning about manipulation. We use a compliant robot hand to find objects without prior knowledge of their presence or location, and then tap those objects with a finger. This behavior lets the robot generate and collect samples of the contact sound produced by impact with that object. We demonstrate the feasibility of recognizing objects by their sound, and relate this to human performance under situations analogous to that of the robot.

1. Introduction

Grasping and touch offer intimate access to objects and their properties. In previous work we have shown how object contact can aid in the development of haptic and visual perception (Natale et al., 2004, Metta and Fitzpatrick, 2003). We now turn our attention to audition: developing perception of contact sounds. Hearing is complementary both to vision and touch during contact events. Unlike vision, hearing doesn't require line of sight – it won't be blocked by the arm, hand, or the object itself. And unlike touch, hearing doesn't require the robot to be the one causing the contact event. We are motivated by an experiment we report in this paper, where human subjects successfully grasped objects while blindfolded, using coarse tactile information and sound.

The extensive use of vision rather than haptic feedback in robotic object exploration may be due to technological limits rather than merit. The robotic hand used in this paper is designed to overcome these limitations. It is equipped with dense touch sensors and series elastic actuators which allow passive compliancy and to measure force at the joints. Force feedback and intrinsic compliance are exploited to successfully control the interaction between robot and environment without relying on visual feedback.

The paper is organized as follows. Section 2. briefly reviews evidence for the importance of augmenting vision with other sensory input for manipulation in human infants and adults, and introduces the notion of *exploratory procedures* in humans and robots. Section 3. describes our robotic platform, designed to enable sensor-rich reaching and grasping (*sensitive manipulation*). Section 4. describes an experiment we carried out with human subjects with their senses interfered to try to simulate our robot. The experiment helps us to understand how humans would solve the kinds of problems with which our robot will be confronted. In section 5., we review our general developmental approach to robot perception, and then apply it to the problem of contact sounds. This motivates us to develop a robot behavior (described in Section 6.) which gives the robot a way to actively probe the sound of objects in its environment in a robust way, by tapping them. Section 7. describes how the experience generated by the robot's behavior is exploited for learning. Section 8. quantifies the accuracy of object recognition enabled by this procedure. Finally, Section 9. discusses the results and places them on a broader perspective.

2. Background

Experimental results suggest that from a very early age, arm movements in infants are influenced by vision. For example, van der Meer and colleagues found that sight of the hand allows infants to maintain the posture of the hand when pulled by an external force (van der Meer et al., 1995). Von Hofsten compared the arm movements of two groups of infants in the presence and absence of an object and found that in the former case arm movements were significantly more frequent. When the infants were fixating the objects the movements were directed closer to it (von Hofsten, 1982). Taken together, these results suggest that in children some sort of eye-hand coordination is already present soon after birth. But on the other hand, continuous visual feedback from the hand is not required for infants to reach for an object (Clifton and D.W. Muir, 1993, Clifton et al., 1994). Indeed it is only at 9 months of age that children seem to be able to exploit visual feedback from the hand during the approach phase (Ashmead et al., 1993). A possible explanation for this could be that in the first months of development the visual system of infants is still rather immature: visual acuity is limited and perception of depth has not developed yet (Bushnell and Boudreau, 1993). Later on during development the role of vision is certainly crucial to control the correct preshape of the hand according to the object's shape and orientation; however, tactile feedback from the contact with an object is an alternative source of information that could initially substitute for the visual feedback.

In adults, several studies have revealed the importance of somatosensory input (force and touch); for example human subjects with anesthetized fingertips have difficulty in handling small objects even with full vision (Johansson, 1991). Humans use a set of strategies collectively called *exploratory procedures* (Lederman and Klatzky, 1987) in their perception of the world around them, such as tracing object outlines with a finger.

This has inspired work on robotics. An analog of human sensitivity to thermal diffusivity was developed by (Campos et al., 1991), allowing a robot to distinguish metal (fast diffusion) from wood (slow diffusion). A robotic apparatus for tapping objects was developed by (Richmond and Pai, 2000) to characterize sounds so as to generate more convincing contact in haptic interfaces. In (Femmam et al., 2001), a special-purpose robot listens to sounds of the surface it "walks" on.

We use a *tapping* exploratory procedure, applied to natural objects by a general purpose, compliant hand (rather than a rigid, special purpose tapping device). Repetitive contact between the fingers and the object (the tapping behavior) allows the robot to collect information about the object itself (the sound produced by the collision of the fingers and the object surface) which is used for object recognition.

3. The robot Obrero

The humanoid robot used in this work, Obrero, consists of a hand, arm and head, shown in Figure 1. Obrero was designed to approach manipulation not as a task mainly guided by a vision system, but as one guided by the feedback from tactile and force sensing – which we call *sensitive manipulation.* We use the robot's limb as a sensing/exploring device as opposed to a pure acting device. This is a convenient approach to operate in unstructured environments, on natural unmodeled objects. Obrero's limb is sensor-rich and safe – it is designed to reduce the risk of damages upon contact with objects.

The arm used in Obrero is a clone of a force-controlled, series-elastic arm developed for the robot Domo (Edsinger-Gonzales and Weber, 2004). The hand consists of three fingers and a palm. Each one of the fingers has two links that can be opened and closed. Two of the fingers can also rotate. Each one of the joints of the hand is controlled using an optimized design for a series elastic actuator (Torres-Jara and Banks, 2004). Series elastic actuators reduce their mechanical impedance and provide force sensing (Williamson, 1995). Summary information about the hand is given in Figure 1.

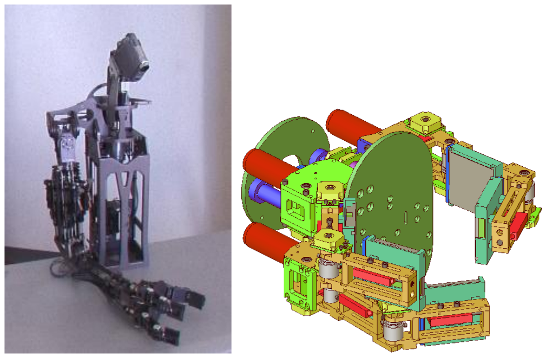

Figure 1: The robot Obrero (left) has a highly sensitive and force controlled hand, a single force controlled arm and a camcorder as a head (used simply as a microphone in this paper). Obrero's hand (right) has three fingers, 8 DOF, 5 motors, 8 force sensors, 10 position sensors and 7 tactile sensors.

4. Simulating our robot with humans

Human haptic perception is impressive, even under serious constraint. In (Lederman and Klatzky, 2004) we can find a review of different experiments done with humans to determine how well they can identify objects using only haptic information. In the experiments mentioned, the individuals wore headphones and a blindfold to make sure that sound and vision did not provide extra information about the objects. Haptic information was also systematically interfered with to explore different aspects of manual exploration. The constraints included: reduced number of end effectors, compliant covering, application of rigid finger splints, rigid finger sheathes, and rigid probes. These constraints reduced either one or many aspects of the cutaneous (spatial, temporal and thermal) and kinesthetic information available to the subjects.

The results showed that by reducing the type of sensing available in the human hand, the subject's recognition performance is reduced. The lowest recognition accuracy for objects was around 40% when the subjects used a probe to explore the object. This recognition task took around 80 seconds. For the researchers who did this work, these numbers may seem low – but for a robotics researcher, they are a cause of envy, and show that human haptic perception is indeed very impressive even under unusually-constrained situations.

To get an "upper bound" of what we could expect from our robot, we evaluated the performance of human subjects when wearing thick gloves that reduced their sensitivity and dexterity to something approaching our robot. We blocked their vision, since we know our robot cannot compete with human visual perception, but let them hear.

We sat 10 subjects in front of a padded desk covered with various objects – a wooden statue, a bottle, a kitchen glove, a plastic box, a paper cup, a desktop phone, a tea bag and a business card. The subjects wore a blindfold and a thick glove which reduced their haptic sensitivity and the number of usable fingers. The glove only allowed them to use their thumb, their index and middle finger. A goal of the experiment was to determine how much and in what way humans can manipulate unknown objects in an unknown environment with capabilities reduced to something approximating our robot (described in Section 3.).

Our subjects were instructed to perform certain tasks starting from a constant initial position, sitting straight with their right arm relaxed and close to their waist. The first task was to find and (if possible) identify objects on a desk. This task was repeated with multiple set of objects. When changing from one set of objects to another, the subjects were moved away and turned around so that their back was facing the desk. The next task extended the challenge further. Along with locating and identifying the objects (an arbitrary name was assigned when an object was not recognized), the subjects were instructed to remember the object's position. Later, they were instructed to move their hand to a named object starting from the initial position.

For the final task, a few objects and a desktop phone were placed on the desk. The hand set and the phone base were disconnected – the phone cord was removed, and the two parts of the phone were placed in separate locations. The subjects initially had no idea a phone was present. They were instructed to find, identify and remember the position of the object on the desk. If they identified the two parts of the phone, they were instructed to grab the hand set and placed in the correct position on the phone base.

Here is a summary of our observations:

- ▷ Exploration strategies vary. Some subjects face their palm in the direction of motion, others towards the desk. The speed at which people swing their arm is generally slow and cautious, with occasional contact with the table.
- ▷ Very light objects were consistently knocked over.
- ▷ Subjects quickly reorient their hand and arm for grasping if either their hand or their wrist makes contact with an object.
- ▷ Subjects exhibited a short-term but powerful memory for object location.
- ▷ Sounds produced by objects and surfaces were used to identify them, compensating partially for the reduction in tactile sensitivity (see Figure 2). This was occasionally misleading: one subject unwittingly dragged a teabag over the desk, and thought from the sound that the surface was covered in paper.

Figure 2: Subjects exploring a desk while blindfolded and wearing a thick glove. Top: light objects were inevitably knocked over, but the sound of their fall alerted the subjects to their presence, location, and (often) identity. Bottom: the sound of object placement was enough to let this subject know where the cup was and suggest a good grasp to use.

Inspired by the last observation, in this paper we focus on exploiting the information carried by sound in combination with tactile and force sensing.

5. Overall developmental approach

We wish to give our robots many ways to learn about objects through action (Fitzpatrick et al., 2003). This contributes to perceptual development, where the robot's experience of the world is filtered by prior experience. This process can be broken down into four steps:

- ▷ Identification of an opportunity to reliably extract some object features
- ▷ Exploitation of that opportunity to extract those features.
- ▷ Use careful generalization to transform the robot's perception of its environment.
- ▷ Transformation of the robot's activity, enabled by its extended perceptual abilities.

In previous work, we have demonstrated this process. In (Arsenio et al., 2003), we showed that poking an object gives us the opportunity to reliably extract visual features of its appearance. By carefully choosing features that generalize, the robot's perception of its environment is transformed, and new activities are enabled (Fitzpatrick, 2003b). Other opportunities we have explored include the use of grasping (Natale et al., 2005) and the integration of multi-modal cues across sound, vision, and proprioception (Fitzpatrick et al., 2005, Arsenio and Fitzpatrick, 2005). Having established this process, we are now seeking to broaden the range of opportunities that can be identified and exploited

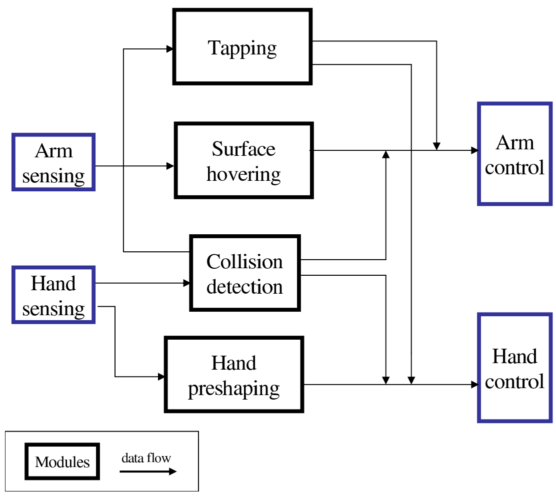

Figure 3: The component elements of the robot's behavior. The modules Arm control, Arm sensing, Hand control and Hand sensing represent the connection with the hardware of the robot.

(steps 1 and 2 above). In the current work, we identify (and in fact create) an opportunity to reliably extract examples of contact sounds involving an object (by tapping that object). We build the appropriate robot behavior and data collection infrastructure to gather those features.

6. The robot's behavior

The behavior of the robot is as follows. It sweeps its hand back and forth over a table, and stops to tap any object (or, indeed, any obstacle) it comes in contact with. This overall behavior is the product of the combined operation of a number of sub-behaviors, shown in Figure 3.

Before we describe how they interact, here is a summary of these component behaviors:

- ▷ *Hand preshaping.* This module places the middle and index fingers together and perpendicular to the palm. The thumb is held up, perpendicular to the other two fingers. For preshaping, the fingers are controlled based on position rather than force.
- ▷ *Collision detection.* This module uses the outputs from the force and tactile sensors in each finger to determine whether a collision has occurred. This is possible because the hand has very low mechanical impedance and consequently the fingers slightly bend upon contact with an object. This bending is detected by the force sensor, often before the force exherted by the finger has greatly affected the object.
- ▷ *Surface hovering.* This behavior hovers the arm and hand over a surface using a predetermined fixed action pattern. The motion can be interrupted at any time.
- ▷ *Tapping.* This behavior moves the fingers back and forward for a given time, in another fixed action pattern.
- ▷ *Arm control.* This module deals directly with the low level motor control of the arm. The arm, for the work described in this paper, uses position control for each of the joints. To produce motion, a smooth trajectory is interpolated between setpoints.
- ▷ *Hand control.* This module provides a connection with the low level controller of the hand. It allows control of parameters such as the gain and the type of controllers, i.e. position and force control.
- ▷ *Arm sensing.* This modules reads the force and position measurements from the low level controller for the arm.
- ▷ *Hand sensing.* This module reads the force, position and tactile measurements from the low level controller for the hand.

The interaction of these parts is as follows. The *hand preshaping* and *surface hovering* modules make the arm and hand sweep over the surface with the middle and index finger extended forward and the thumb up. This is done by sending commands to the *arm control* and *hand control* modules.

When the fingers of the robot come in contact with an object, the *collision detection* module overrides the messages coming from *hand preshaping* and *surface hovering* to the *arm control* and *hand control* modules, commanding the arm to an immediate stop. At the same time the behavior *tapping* sends commands to the *hand control* module to periodically touch the object and to the *arm control* module to keep the arm in position. The tapping lasts a few seconds, after which the *tapping* module relinquishes the control and stop sending commands. At this point the *surface hovering* and *preshaping hand* modules can get their message across to the motor control modules. Consequently, the arm is repositioned and the sweeping behavior reactivated.

These modules run on different machines on the network of computers that control Obrero. The interconnection between modules was done using `YARP` (Fitzpatrick et al., 2004).

During the experiment we recorded vision and sound from the head along with the force feedback from both the arm and hand. The visual feedback was *not used* in the robot's behavior; it was simply recorded to aid analysis and presentation of results. All other sensory information were considered candidates for detecting contact. The force feedback from the hand proved the simplest to work with. Peaks in the hand force feedback were successfully employed to detect the impact of the fingers with the object during both the exploration and tapping behaviors. Force and sound were aligned as shown in Figure 4. Once the duration of a tapping episode was determined, a spectrogram for the sounds during that period was generated as shown in Figure 5. The overall contact sound was represented directly as the relative

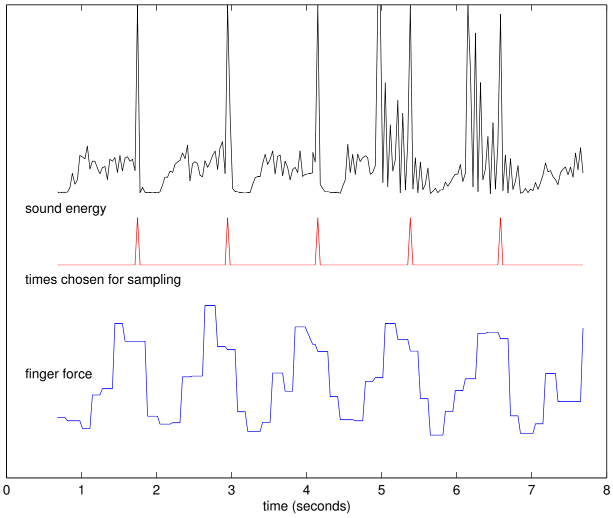

Figure 4: Force readings from the fingers (bottom) reveal when tapping may occur. Swings in the force are compared against sound intensity (top), looking for synchronized sounds. Peaks within one fifth of a period from a force swing are accepted. This process lets the robot filter out environmental sounds that occur when the arm is not moving, and even during tapping. In this example, the first three peaks of sound are clean, but the last two are corrupted by a phone ringing (see Figure 5).

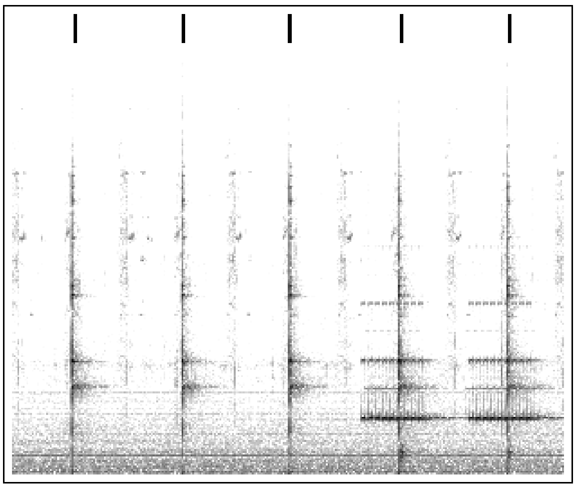

Figure 5: This is the spectrogram of the sounds in Figure 4 (time on the x-axis, increasing frequency on the y-axis, dark color corresponds to activity). The top of the spectrogram is marked to show the five sample times selected automatically. Between these times, there are patches of sound corresponding to the sound of springs in the fingers. The last two samples have the sound of a phone superimposed on them.

distribution of frequencies at three discrete time intervals after each tap, to capture both characteristic resonances, and decay rates. The distributions were pooled across all the taps in a single episode, and averaged. Recognition is performed by transforming these distributions into significance measures (how far frequency levels differ from the mean across all tapping episodes) and then using histogram comparison.

7. Data collection for learning

The robot's behaviors are designed to create opportunities for learning, by finding and tapping objects. The modules that exploit these opportunities for learning are entirely from the modules that control the behavior of the robot. The occurrence of tapping is detected based on sensor data, rather than commanded motion. The only interaction that takes place between these modules is via actions in the world (Brooks, 1990). This improves robustness. We do not have to deal with explicit expectations or their possible failure modes. For example, sometimes the robot fails to hit an object when tapping, so it is good to pay more attention to actual contact rather than commanded motions.

The force measurements from the fingers is summed into a single signal, then classified into "rising", "falling", and "neutral" phases. Classification transitions to "rising" if the signal increases over 10% of the previous range covered by the signal from its highest to lowest point during a rising and falling period. Similarly, the classification transitions to "falling" if the signal falls by over 10% of this range. Since the range is constantly updated, the classification is robust to slow-changing offsets, and the actual gross magnitude of swings. The classifications are scanned for rythmic rising and falling with a period lying between 0.2 and 2 seconds. Then the force signal in these regions is compared with the sound, to find if peaks in the sound line up well (within 20% of a period) of either peaks or troughs in the force signal (the sign depends on the orientation of the fingers during tapping). All going well, a spectrogram of the sound is performed in the appropriate range. Only the spectrogram around the peaks (presumably from tapping) is significant. Three samples are made in quick succession after each peak, to capture not just characteristic resonance but decay properties.

The robot's learning is performed on-line, but not in real-time. Performing data collection and learning in real-time on a robot can lead to research time wasted optimizing code and iterating designs that are otherwise adequate. But simply switching to off-line performance is undesirable, since it offers too many subtle ways for human input to enter the process. Hence we divided the robot's on-line system into two parts, the *real-time* subsystem that controls behavior, and the *near-time* subsystem that continually processes the robot's experience. This follows the design of the robot Cog's object recognition system (Fitzpatrick, 2003a).

Figure 6 shows the time course of an experiment. The key property being illustrated is that the processing of the robot's experience happens at a relatively

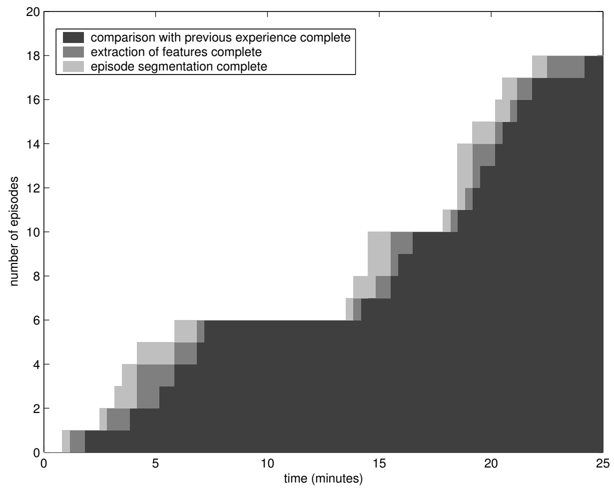

Figure 6: Time course of an experiment, showing the aggregation of experience by the robot. Over this 25 minute interval, 18 tapping episodes are detected. The episodes are first segmented (light color), then analyzed to extract characteristic features of the sound (darker color), then finally compared with previous episodes (darkest color). This process is online but unhurried – each episode can take on the order of minutes to be completely processed. In the meantime, the robot can continue with its normal behavior unimpeded.

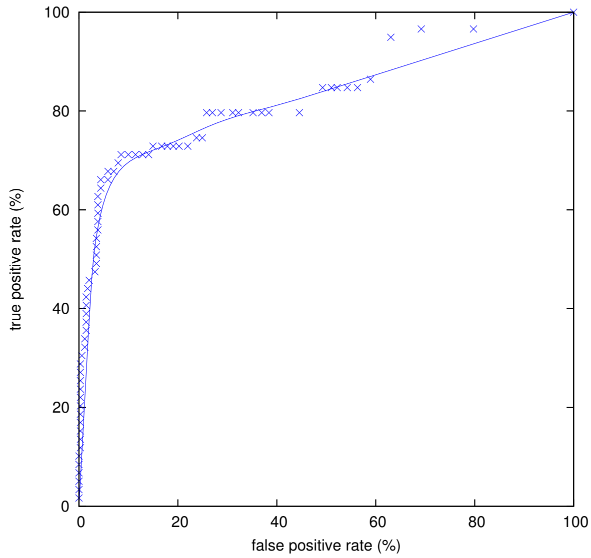

Figure 7: Receiver-operator characteristic curve. Tapping episodes from one day are matched against episodes from a previous day. Matches are ranked, then truncated based on a quality threshold. This plot shows the effect of that threshold on the trade-off between false matches and missed matches.

leisurely pace. This is workable as long as the processing can keep ahead of incoming data. For our robot, a complete rotating log of the robot's sensory input is made that covers about 30 minutes. Technically, this is achieved using a modified version of the open-source tool `dvgrab` for recording from a camcorder, and simple text files for other (much lower bandwidth) proprioceptive and summary data. The logs are maintained on a separate computer from the one controlling the robot's bahavior. These logs are processed using the open-source MATLAB-clone `octave`.

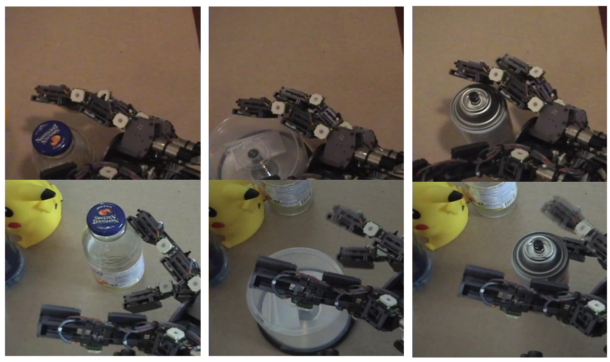

Figure 8: There were four objects in common between the training and test run. Three of them were matched perfectly (using a best-match rather than threshold-based strategy) for every episode: a bottle (left), a CD case (middle), and a spray-can (right). Images on the bottom are from the test run, images on the top are from the best matching episode in the training run. These objects have quite distinctive sounds. A plastic toy (left corner of each lower image) failed to be recognized – it was quiet, and made just a dull thud.

8. Results

We evaluated our work by performing an object recognition experiment. We exposed the robot one evening to a set of seven objects, and then in the morning tested its ability to recognize another set, which had an overlap of four objects with the training set.

Three of these objects were chosen (Figure 8) to represent three different materials, plastic, glass and steel (metal). The idea is that the sound produced by each object depends on its size, shape and the material with which it is made; accordingly we expected the tapping to produce three different distinct sounds. A fourth object (a plastic toy) was relatively silent.

For each run, we placed randomly selected objects on the table in front of the robot, and it was responsible for finding and tapping them. Overall the robot tapped 53 times; of these episodes 39 were successful, meaning that the sound produced by the tapping was significantly loud; in the other 14 cases the tapping did not provoke useful events either because the initial impact caused the object to fall, or the object remained too close to the hand. The high number of successful trials shows that given the mechanical design of the hand, haptic feedback was sufficient to control the interaction between the robot and the environment.

We evaluated the performance of our spectrum comparison method by ranking the strength of matches between episodes on the second day and episodes on the first day. Figure 7 shows what detection accuracy is possible as the acceptable false positive rate is varied. This predicts that we can on average correctly match an episode with 50% of previous episodes involving the same object if we are willing to accept 5% false matches.

9. Conclusions

We have demonstrated a compliant robot hand capable of safely coming into contact with a variety of objects without any prior knowledge of their presence or location – the safety is built into the mechanics and the low level control, rather than into careful trajectory planning and monitoring. We have shown that, once in contact with these objects, the robot can perform a useful exploratory procedure: tapping. The repetitive, redundant, cross-modal nature of tapping gives the robot an opportunity to reliably identify when the sound of contact with the object occurs, and to collect samples of that sound. We demonstrated the utility of this exploratory procedure for a simple object recognition scenario.

This work fits in with a broad theme of learning about objects through action that has motivated the authors' previous work (Fitzpatrick et al., 2003). We wish to build robots whose ability to perceive and act in the world is created through experience, and hence robust to environmental perturbation. The innate abilities we give our robots are not designed to accomplish the specific, practical, useful tasks which we (and our funders) would indeed like to see, since direct implementations of such behaviors are invariably very brittle; instead we concentrate on creating behaviors that give the robot robust opportunities for adapting and learning about its environment. Our gamble is that in the long run, we will be able to build a more stable house by building the ground floor first, rather than starting at the top.

Acknowledgements

This project makes heavy use of F/OSS software – thank you world. This work was partially funded by ABB, by DARPA DABT 63-00-C-10102, and by NTT under the NTT/MIT Collaboration Agreement. Lorenzo Natale was supported by the European Union grants RobotCub (IST-2004-004370) and ADAPT (IST-2001-371173).

References

Arsenio, A. and Fitzpatrick, P. (2005). Exploiting amodal cues for robot perception. *International Journal of Humanoid Robotics*, 2(2):1–19. in press.

Arsenio, A., Fitzpatrick, P., Kemp, C., and Metta, G. (2003). The whole world in your hand: Active and interactive segmentation. In *Proceedings of the Third International Workshop on Epigenetic Robotics*.

Ashmead, D., McCarty, M., Lucas, L., and Belvedere, M. (1993). Visual guidance in infants' reaching toward suddenly displaced targets. *Child Dev.*, 64(4):1111–27.

Brooks, R. A. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6:3–15.

Bushnell, E. and Boudreau, J. (1993). Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development. *Child Dev.*, 64(4):1005–21.

Campos, M., Bajcsy, R., and Kumar, V. (1991). Exploratory procedures for material properties: the temperature perception. In *5th Int. Conf. on Advanced Robotics*, volume 1, pages 205–210.

Clifton, R. and D.W. Muir, D.H. Ashmead, M. C. (1993). Is visually guided reaching in early infancy a myth? *Child Dev.*, 64(4):1099–110.

Clifton, R., Rochat, P., Robin, D., and Berthier, N. E. (1994). Multimodal perception in the control of infant reaching. *J Exp Psychol Hum Perform.*, 20(4):876–86.

Edsinger-Gonzales, A. and Weber, J. (2004). Domo: A force sensing humanoid robot for manipulation research. In *Proc. of the IEEE International Conf. on Humanoid Robotics*, Los Angeles.

Femmam, S., M'Sirdi, N. K., and Ouahabi, A. (2001). Perception and characterization of materials using signal processing techniques. *Instrumentation and Measurement*, 50(5):1203–1211.

Fitzpatrick, P. (2003a). *From First Contact to Close Encounters: A developmentally deep perceptual system for a humanoid robot.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA. Also available as MIT Artificial Intelligence Laboratory technical report AITR-2003-008.

Fitzpatrick, P. (2003b). Perception and perspective in robotics. Proceedings of the 25th annual meeting of the Cognitive Science Society.

Fitzpatrick, P., Arsenio, A., and Torres-Jara, E. R. (2005). Reinforcing robot perception of multimodal events through repetition and redundancy and repetition and redundancy. *Interaction Studies Journal.* accepted for publication.

Fitzpatrick, P., Metta, G., and Natale, L. (2004). YARP: Yet Another Robotic Platform. http://yarp0.sourceforge.net/.

Fitzpatrick, P., Metta, G., Natale, L., Rao, S., and Sandini, G. (2003). Learning about objects through action - initial steps towards artificial cognition. In *Proc. of the IEEE Internat'l Conf. on Robotics and Automation*, Taipei, Taiwan.

Johansson, R. S. (1991). How is grasping modified by somatosensory input? In Humphrey, D. R. and Freund, H. J., (Eds.), *Motor Control: Concepts and Issues*, pages 331–355. John Wiley and Sons Ltd, Chichester.

Lederman, S. J. and Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19(3):342–368.

Lederman, S. J. and Klatzky, R. L. (2004). Haptic identification of common objects: Effects of constraining the manual exploration process. *Perception and Psychophysics*, 66:618–628.

Metta, G. and Fitzpatrick, P. (2003). Early integration of vision and manipulation. *Adaptive Behavior*, 11(2):109–128.

Natale, L., Metta, G., and Sandini, G. (2004). Learning haptic representation of objects. In *International Conference on Intelligent Manipulation and Grasping*, Genoa, Italy.

Natale, L., Orabona, F., Metta, G., and Sandini, G. (2005). Exploring the world through grasping: a developmental approach. In *Proceedings of the 6th CIRA Symposium*, Espoo, Finland.

Richmond, J. L. and Pai, D. K. (2000). Active measurement and modeling of contact sounds. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 2146–2152, San Francisco.

Torres-Jara, E. and Banks, J. (2004). A simple and scalable force actuator. *35th International Symposioum on Robotics.*

van der Meer, A., van der Weel, F., and Weel, D. (1995). The functional significance of arm movements in neonates. *Science*, 267:693–5.

von Hofsten, C. (1982). Eye-hand coordination in the newborn. *Dev. Psychology*, 18(3):450–61.

Williamson, M. (1995). Series elastic actuators. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.