Learning to Detect Doorbell Buttons by Haptic Exploration

Liping Wu

March 10, 2011

1. Introduction

While it is meaningful to learn and get the visual representation of the button so that the model learned can be used to detect the button in visual space, it is also important to learn and get the tactile and proprioceptive representation of the button, so that the robot can find the buttons by tactition and proprioception. Not only the robot can still detect and locate and operate the buttons under the condition where there is no vision available, but also the tactile and proprioceptive representation is more accurate in defining a button, which is invented by human to help human's life.

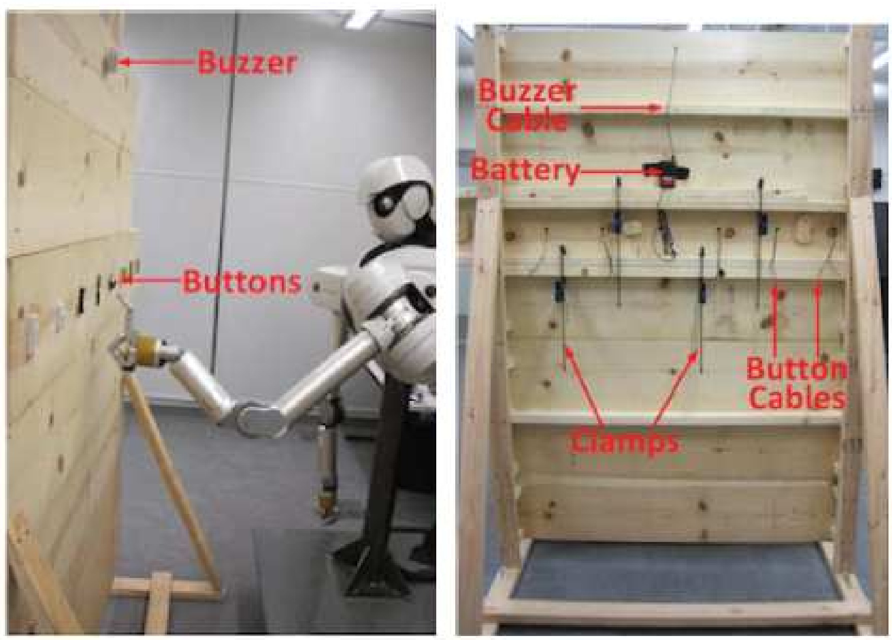
More correctly, this research focuses on the push-button, which has a spring in to return to the un-pushed state. Because of the spring in the push-button, the fingertip can sink in while keep contact with the button's surface and feel the resistance, and also when being released the button's surface will resume automatically. (Sukhoy and Stoytchev 2010) trained a visual model that can detect the button with the button-like texture in vision. However, any object with the button-like texture but without the spring and the tactile and proprioceptive properties above, we still can't say that it is a push-button. Therefore, the representation for a button derived from its tactile property is more accurate and may be able to achieve a more accurate detection result.

From the point of view of haptic exploration, sensing the push-button is also very meaningful. By haptic exploration, humans can learn many characteristics of objects, such as object shape, surface texture, stiffness and temperature. This kind of research is also viewed as the tactile data interpretation, which supports the dexterous manipulation a lot. For example, (Okamura and Cutkosky 1999a) designed a mathematical model based on a differential geometry approach to detect small surface feature of bump.

Therefore, because of the popularity of buttons in the humans' life, interpreting the tactile data when press a push-button and building a mathematical model to represent it is also very meaningful.

2. The Previous Study

This project is based on the previous study by the research group in Developmental Robotics Laboratory at Iowa State University. The robot for the previous study as well as this project proposed is showed in Figure 1. This robot has two Barrett Whole Arm Manipulators (WAMs) with a BH8-Series Barrett Hand as arms. Two Logitech cameras are mounted in its head as his eyes. It also has a microphone mounted on the head and an artificial fingernail attached on the finger 3 on the left arm.



(a) The robot pushing a button (b) Experimental fixture (back)

Figure 1: The robot and fixture for the experiment.

In the previous study, there are mainly two projects. One project (Sukhoy, Sinapov, Wu and Stoytchev 2010) is humanoid robot learning to press doorbell buttons using active exploration and audio feedback. With 5 sampled points in the 7-D joint space, the robot can calculate itself to generate press behaviors of pressing an area on the board. The press behaviors are parameterized by the vector decided by the start position and end position of the behavior in the 7-D joint space. By running the pre-learned classifier on the audio stream, the time when the doorbell is triggered can be detected in real-time. For each behavior, it will be labeled as pressing a button or not pressing a button according to if there is a doorbell detected at the meantime. Finally, k-nearest neighbor algorithm is used to do the learning work and three kinds of active selection strategies—random exploration, uncertainty-driven exploration, and stimulus-driven exploration—are used to speed up the learning.

In another project (Sukhoy and Stoytchev 2010), the humanoid robot learns the visual model of doorbell buttons autonomously. Color tracker is used to track the touch position on the board surface in the image from robot's camera. These touch position is simplified as a pixel in the image and labeled as the functional component or not according to if the associated press behavior is pressing button or not based on the audio feedback. Image is split into 10x10 pixel patches, and each patch is labeled as functional component or not according to the density of functional component touch point falling into the grid. For each patch, the texture, edge and low-frequency color information of itself and neighbors are extracted and logistic regression classifier is used to learn the visual model for detecting patches belonging to the functional component of the doorbell buttons.

3. Related Work

3.1 Button Study

In psychology, (Hauf and Aschersleben 2008) found that a 9-month old infant can anticipate what color of buttons will trigger the light or the ring when he/she presses from experience, and in turn by the anticipation control his/her action to press the working buttons more often. In the experiment, the infants were placed in front of 3 groups of buttons. In the first group, the red button is effective. In the second group, the blue button is effective, and in the third group, none button is effective. The result shows that the infants press red button more often for the first group, blue button more often for the second group and almost the same for the third group.

In robotics, the previous work focuses on the visual feedback more. (Thomaz 2006) used social learning to teach the robot how to turn the button on & off using speech communication. But the robot uses the vision to recognize where the button is and decide if the button is on or off. (Miura, Iwase, and Shirai 2005) made the robot execute an take-an-elevator task based on vision. Vision-based teaching algorithm was used to find the location of the elevator door and elevator button. The origin of the elevator was marked with a red light, and the robot searched the area around the origin to find the image template of the elevator. Similarly, being indicated the rough position of the buttons; the robot finds the position of the button by searching for the area nearby. (Klingbeil, Saxena and Ng 2008) tried a haar classifier using supervised learning algorithm for the robot to detect where the elevator button is.

3. 2 Haptic Exploration

In psychology, haptic exploration is defined as exploratory procedures (EPs) related with the modality of touch. EPs are stereotyped patterns describing the ways of contact and movement between skin and object (Lederman and Klatzky 1987). During exploration, the perceptual system, haptics, incorporates inputs from multiple sensory systems (Loomis and Lederman 1986). Haptics includes a cutaneous system sensing pressure, vibration, temperature, and a kinesthetic system registering position and movement of the muscles and joints. Between EPs and object properties, there are associations describing whether an EP is necessary, optimal, sufficient, or inadequate in exposing a specific property of an object (Klatzky, Lederman, and Matula 1991). By haptic exploration, human can learn these associations, which, in turn, can help the human to choose an optimal EP for obtaining the desired object property. Empirically, the press EP is optimal in obtaining the press feeling of a push-button, which is the reason why a specific press behavior similar to human's is designed in this study.

For the studies of haptic exploration in robotics, most of them focus on detecting the object shape (Caselli et al. 1996, Allen and Roberts 1989, Roberts 1990) and small surface features such as cracks, bumps and ridges (Okamura et al. 2000, Okamura and Cutkosky 2001, Okamura and Cutkosky 1999b). Some papers also designed the models to measure surface toughness, friction, and texture (Okamura et al. 2000, Stansfield 1992, Sukhoy et al. 2009). (Stansfield 1992) used two kinds of pressure EPs to measure the hardness of objects. One is by grasping and squeezing the object, and another one is by probing against the object surface using one finger. In our study, the later kind of EP will be used to detect the push-button, which, to some degree, can be viewed as a soft object that can be probed in.

4. Research Contents

1. The tactile and proprioceptive representation of a push-button here is related with the specific haptic exploration behavior—press, and is how the tactile and proprioceptive feeling is when press a button. The robot will press the doorbell buttons on a board, and a classifier of detecting pressing a push-button using proprioceptive and tactile data will be learned from the experience.

2. Using the classifier got, the robot will go to detect and locate the buttons by pressing the board. The accuracy will be measured to evaluate the classifier learned.

3. A mathematical model to detect the broken doorbell buttons will be built. The broken doorbell buttons are defined as the case that the doorbell is not triggered when they are pressed.

4. Contour maps will be plot as the supplement of the tactile feeling of a doorbell button.

5. Experimental Setup

5.1 Robot and Fixture

See the Fig. 1 for the robot and fixture used in the experiment and previous study for the description.

5.2 Press Behavior

Buttons are associated with the press behaviors, by which they can be manipulated more conveniently. Therefore, we would like to offer the robot the press behavior, and make the press behavior more similar to the humans'. For a human when doing a press, he/she first touches a point on the surface, and presses toward the surface as well as withdraws while keep the touch point unchanged. Fig. 2 is an explanatory drawing for sampling the press behavior in our experiment.

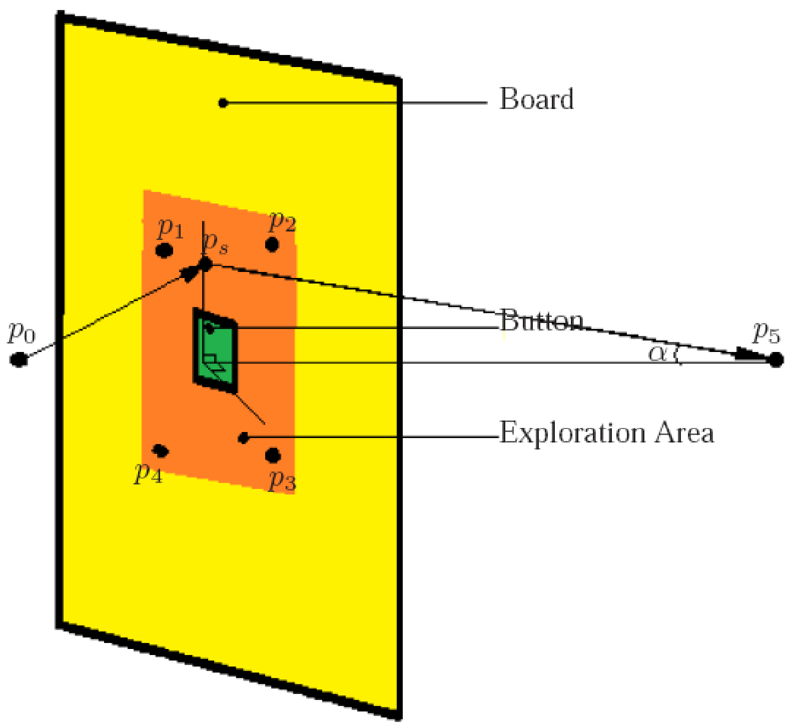


Figure 2: Explanatory drawing for the press behavior.

Keeping the shape of the robot arm like the one in Fig.1 (a), place the fingertip onto six points in the 3-D Cartesian space showed in the Fig. (2) as p_0-p_5. Then, the six associated points (p_0^j-p_5^j) in the 7-D joint space defined by the 7 joint angles of the robot will be used to sample the press behaviors. For each press, the robot arm starts from point p_0^j, and moves to one intermediate point p_m^j sampled from p_1^j-p_4^j, which guarantees that the fingertip will touch a random point on the exploration area of the board. As soon as the fingertip touches the surface, which can be detected by the accelerometer in the artificial fingernail, the robot arm stops and the current position will be taken as the start point (p_s^j, associated to the fingertip position p_s) for the press.

Now, the press begins. Starting from the start point (p_s^j), the robot fingertip presses the surface toward point p_5^j, until any of the joint torque values exceeds the pre-defined the torque limit and the robot arm stops at the end point p_e^j. Because points are sampled in such a way that the intersection angle α between the normal line through p_5 of the board and the line joining p_5 and any touch point is very low, this makes sure that the touch point will keep the same during press because of the friction. Then, the fingertip withdraws back to the start point p_s^j slow enough to keep contact with the surface in the same touch point, finishing the press. Finally, the robot arm withdraws back to p_0^j ready to generate another press on a random point from the exploration surface.

A press consists of the forward press behavior and the subsequent backward behavior. During a press, the contact point keeps the same, so that the displacement of the fingertip position will just be the displacement of the surface. Ideally, the displacement of the fingertip position is zero for the board and the travel distance of the spring for the button during a press. 3-D accelerometer and the fingertip positions will be recorded during each press.

6. Detecting the Press on a Push-button

6. 1. Tactile and Proprioceptive Feeling for a Press on a Push-button

By Wikipedia, a push-button (also spelled pushbutton) or simply button is a simple switch mechanism for controlling some aspect of a machine or a process. Most of the buttons are biased switches. There are two types of biased switch, and they are push-to-make and push-to-break. For a push-to-make button, contact is made when pressed and broken when released. On the contrary, contact is broken when released and made when pressed for a push-to-break button. Most of buttons are push-to-make type, such as computer keyboard and doorbell button, which is the research target of this paper. The function for a push-to-make type button is to make contact by narrowing the distance due to the loading of the external force, and to break contact by broadening the distance due to the unloading of the external force. Consider a press as the combination of a forward press behavior and the subsequent backward behavior. Therefore, the correct tactile and proprioceptive feeling when doing a press on a push-button is, there is considerable displacement change along the force change direction.

We can't say this is a exactly sufficient and necessary, but it can do right decision in most cases, at least in differentiating pressing among a button, a fixed board and a moveable object. More properties of a button, say the visual shape, can be added to boost the result. Because of the spring in the button, we are also interested in finding the associated characteristics, say buffing effect. Also, when the moveable part hits the fixed part of a button, collision will happen to result in the vibration in the SG/tactile values. We are also interested in observing these, but they are minor comparing the property resulting from the function of a button.

6. 2. Methodology

One press consists of one forward press behavior and subsequent backward behavior. Since the forward press behavior is perpendicular to the board and starts from a position contacting the surface, the displacement of the fingertip position from the start position of the press behavior can be used to estimate the travel distance of the button, assuming the board is also a push button with a travel distance of nearly zero. It is the same to the backward behavior.

One way is just considering the travel distance of fingertip during the forward press behavior. In another way , for each behavior, forward press behavior or backward behavior during one press, piecewise aggregate approximation (PAA) will be used to smooth the original accelerometer values and fingertip travel distances as well as feed each kind of values into the same dimension. The second way will be done when time is available.

Data got will be used to learn the model of detecting pressing button. For each press, it will be taken as an independent instance, and labeled by button or not button depending on whether there is doorbell triggered in the mean time. Supervised or unsupervised learning will be used to learn the model of tactile and proprioceptive feeling when pressing a button, depending on whether use the labels or not.

For unsupervised learning, some clustering algorithms like x-means and k-means, will be used to do the learning work. For supervised learning, some classifying algorithms like k-NN and Naive Bayes will be used to do the learning work.

6. 3. Evaluation

Two data sets will be recorded. One set is for training the model, and another set is for evaluation. Two evaluation methods will be used. One method is to do the statistics on the accuracy of classifying each press instance.

Another method is to visualize the detecting result. Track each press point simplified as one pixel by color tracker in visual space, and then label the press points (pixels) as button or non-button by the classifier learned. Based on the labeled pixels, use k-NN algorithm to determine the label of each pixel in the image from the camera of the robot. Finally transfer the image into binary image according to the labels to visualize the detecting result.

7. Detecting the Broken Button

7. 1. Methodology

Use the stimulus-driven strategy to locate the button as well as generate many presses on the button. Record the maximum volume in audio during each press, and use k-means to classify buzzer and non-buzzer press. If the percentage of buzzer presses over the button presses is lower than a learned threshold, then the button is broken, or it is not. The threshold is learned by the experience for getting the classifier and equals to the half of the average percentage of buzzer presses over the button presses for all the buttons.

7. 2. Evaluation

Several broken and non-broken buttons will be offered for the robot to press. The accuracy of detecting will be computed as the evaluation.

8. Experience

Liping Wu is currently working in the Developmental Robotics Laboratory and has been enrolled in the previous study. He has the access to program the robot and is already familiar with programming the robot. Based on the research experience as well as the Machine Learning course and Computational Perception and Artificial Intelligence, he has been equipped with certain knowledge and skills of programming in machine learning algorithms and in C/C++.

9. References

J. Miura, K. Iwase, and Y. Shirai 2005. Interactive teaching of a mobile robot. In IEEE

International Conference on Robotics and Automation, volume 3, page 3378.

P. Hauf and G. Aschersleben 2008. “Action-effect anticipation in infant action control,” Psych. Research, vol. 72, no. 2, pp. 203–210.

Thomaz, A. 2006. *Socially guided machine learning*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Klingbeil, E. and Saxena, A. and Ng, A.Y. 2008. “Learning to Open New doors”, Stanford University.

Allen, P., and Roberts, K. 1989. Haptic object recognition using a multi-fingered dextrous hand. In *IEEE International Conference on Robotics and Automation*, 342–347.

Caselli, S.; Magnanini, C.; Zanichelli, F.; and Caraffi, E. 1996. Efficient exploration and recognition of convex objects based on haptic perception. In *IEEE International Conference on Robotics and Automation*, 3508–3513.

Klatzky, R. L.; Lederman, S. J.; and Matula, D. 1991. Imagined haptic exploration in judgments of object properties. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:314–322.

Lederman, S. J., and Klatzky, R. L. 1987. Hand movements: A window into haptic object recognition. *Cognitive psychology* 19(1):342–368.

Loomis, J., and Lederman, S. 1986. Tactual perception. In Boff, K.; Kaufman, L.; and (Eds.), J. T., eds., *Handbook of human perveption and performance*, 1–41. New Youk: Wiley.

Okamura, A., and Cutkosky, M. 1999a. Haptic exploration of fine surface features. In *IEEE International Conference on Robotics and Automation*, 2930–2936.

Okamura, A., and Cutkosky, M. 1999b. Haptic exploration of fine surface features. In *IEEE International Conference on Robotics and Automation*, volume 4, 2930–2936.

Okamura, A., and Cutkosky, M. 2001. Feature detection for haptic exploration with robotic fingers. *The International Journal of Robotics Research* 20(12):925.

Okamura, A.; Costa, M.; Turner, M.; Richard, C.; and Cutkosky, M. 2000. Haptic surface exploration. *Experimental Robotics VI* 423–432.

Roberts, K. 1990. Robot active touch exploration: Constraints and strategies. In *IEEE International Conference on Robotics and Automation*, 980–985.

Stansfield, S. 1992. Haptic perception with an articulated, sensate robot hand. *Robotica* 10(06):497–508.

Sukhoy, V., Sinapov, J., Wu, L., and Stoytchev, A. 2010a. "Learning to Press Doorbell Buttons," In Proceedings of the 9th IEEE International Conference on Development and Learning (ICDL), Ann Arbor, Michigan, August 18-21, pp. 132-139.

Sukhoy, V., and Stoytchev, A. 2010b. Learning to detect the functional components of doorbell buttons using active exploration and multimodal correlation. In *IEEE International Conference on Humanoid Robots (Humanoids)*, 572–579.

Sukhoy, V.; Sahai, R.; Sinapov, J.; and Stoytchev, A. 2009. Vibrotactile recognition of surface textures by a humanoid robot. In *The 9-th IEEE-RAS International Conference on Humanoid Robots*, 57–60.