Artificial Intelligence for Human-Robot Interaction: Papers from the 2014 AAAI Fall Symposium

Understanding Touch Gestures on a Humanoid Robot

Wallace Lawson
Naval Center for Applied
Research in Artificial Intelligence
Washington, DC

Keith Sullivan[2]
Excelis, Inc.
McLean, VA

J. Gregory Trafton
Naval Center for Applied
Research in Artificial Intelligence
Washington, DC

Abstract

Touch can be a powerful means of communication especially when it is combined with other sensing modalities, such as speech. The challenge on a humanoid robot is to sense touch in a way that can be sensitive to subtle cues, such as the hand used and amount of force applied. We propose a novel combination of sensing modalities to extract touch information. We extract hand information using the Leap Motion active sensor, then determine force information from force sensitive resistors. We combine these sensing modalities at the feature level, then train a support vector machine to recognize specific touch gestures. We demonstrate a high level of accuracy recognizing four different touch gestures from the firefighting domain.

Introduction

Touch gestures are tactile gestures that have a specific interpretation in a given context. They are generally applicable in scenarios where there are well-defined roles. While there has been a lot of previous work on the use of tactile HRI to determine emotional state or to ensure operator safety, there has been little relevant to interpreting touch gestures (Argall and Billard, 2010). While it is difficult to determine the reason for this, two key limitations are readily apparent. First, a significant amount of hardware is required to understand touch in a general way (Billard et al., 2013; Ji et al., 2011). Second, even with the necessary hardware, it can be difficult to determine subtle cues necessary to interpret touch.

We propose a novel solution that is capable of addressing both of these limitations. We combine hand sensing from a Leap Motion active sensor(LeapMotion, 2014) with Force Sensitive Resistors (FSR) placed at key locations to understand how touch was applied.

Our work is similar in intent to (Ji et al., 2011), where the authors present a machine learning approach that is combined with a hardware prototype to distinguish between several similar yet subtly different gestures. Our proposed solution is conceptually much simpler to implement, and takes advantage of the active sensing from the Leap Motion to determine how the hand is touching the robot. Similar work on multi-touch displays (Yuan and Barner, 2006) have shown promising results using hand contours, but they lack the ability to provide any information on the force applied.

While touch gestures may be applied to a wide range of scenarios, we specifically focus on a collaborative firefighting task. In such domains, traditional visual gestures are used very infrequently during firefighting episodes. This is a surprising result; gestures, in general, are extremely prevalent across all forms of human-human communication (Goldin-Meadow, 2007; Alibali et al., 1995; Breckinridge Church and Goldin-Meadow, 1986). However, because of the extremely poor visibility, people cannot assume that teammates can see or interpret their visual gestures. Therefore, they turn to touch as a a more reliable method of communicating in such situations.

The firefighting task is divided into the roles taken by the supervisor and the nozzle operator. For our purposes, we assume that the human takes the role of supervisor and that the robot takes the role of nozzle operator. The touches applied by the human are predictable and have a very narrow and specific interpretation. In some cases, the supervisor will apply touch to let the nozzle operator know that they should continue operating normally (we refer to this as the engage gesture). In other cases, the supervisor will apply touch to instruct the nozzle operator that they should turn left, turn right, exit compartment, etc.

During an initial training period, a human operator interacts with the robot, providing examples of different types of touches that might be applied. We build a feature vector describing this touch, which includes information such as force applied, the number of fingers used, the orientation of the hand, and the spread of the hand. We train a support vector machine to classify different types of touch gestures, as described by these feature vectors.

Touch Sensing Hardware

Force sensing resistors (FSR) are made of a material whose resistance are affected by the amount of force applied. We have an array of FSRs positioned at key points throughout the robot. In total, there are 10 FSRs: 6 are placed along the back, 2 on the side, and 2 on the front. For evaluation, we consider these to be 6 different discrete features: back left, back right, side left, side right, front left and front right.

Table 1: Touch gesture recognition accuracy

Touch Gesture	True Positive Rate	False Positive Rate
Engage	98.9%	0%
Exit	95.5%	0%
Turn Left	95.7%	0.3%
Turn Right	90.9%	1%

Leap Motion is a small (76 mm) sensor that uses structured light to locate hands within it's viewing area. It is designed to operate effectively at close ranges (approximately 1 cm to 1 meter). We place a Leap Motion sensor on the back of the robot, pointing diagonally up and away from the robot, so that the workspace intersects with the back of the robot but also covers a significant portion of the space away from the robot. The purpose of doing this is to ensure that hands can be detected and tracked in a timely manner. This increases the accuracy and stability of the hand during the touch. The Leap Motion developer SDK returns a 6 DoF position of the hand. We determine if the hand is touching the robot by testing for intersection between a point (the hand) and a plane (the robot).

Combining Sensing Modalities

The two modalities are combined at the feature level, producing a total of 16 different features: 6 for force sensing, 5 from the left hand position, 5 from the right hand position. Note that this inherently provides the ability to analyze 2 handed touch gestures vs. 1 handed touch gestures.

Recognizing Touch Gestures

During a training period, the supervisor will provide several examples of each gestural command. A feature vector describing the touch gesture is generated from each of the sensors, at a rate of approximately 30 Hz. In order to recognize the gesture, we train a support vector machine with a RBF kernel.

Experimental Results

Our initial experiments are focused towards learning four different, commonly used firefighting touch gestures. They are “engage”, “turn left”, “turn right”, and “exit”. In practice, these touch gestures will be accompanied by speech commands. For example, a supervisor might instruct the robot to turn to the left, in which case the touch indicates the amount that the robot should turn.

We evaluate during a session, where a supervisor provides several examples of each gesture. We split the session, sequestering a portion for evaluation. We tune the support vector machine using 5-fold cross-validation on the training data.

We report results on an individual frame level, but note that it's possible to increase accuracy by considering results across multiple frames.

Discussion

While we have demonstrated the ability to recognize different classes of touch gestures, it is worth noting the rigid nature of these classes vs. the fluid nature of gesture as a whole. Gestures may have different levels of urgency (e.g., touching with extra force), may be used as a way of attention-getting (e.g., single tap), or to express general spatial information (e.g., pushing). In future work, we propose to further extend the classes of touch gesture to explore different ways in which they are made. For example, can we distinguish between “exit” and “urgent exit”?

The proposed hardware approach has proven to be effective and simple. One limitation is the use of only one Leap Motion sensor per computer system. Future revisions of the Leap SDK are expected to remove this limitation. When such a revision occurs, we intend to place multiple Leap Motion sensors around the humanoid robot in order to sense touch around multiple parts of the robot.

Finally, there is a question of general interpretations of touch gestures. The interpretation of the touch gesture may change depending on context, the way the gesture was made, and other sensory cues (such as speech). Touch must be learned in such a way that can be both sensitive to all of this information, yet flexible enough to permit retraining when it is necessary.

Acknowledgment

This work was supported by the Office of Naval Research (GT). The views and conclusions contained in this paper do not represent the official policies of the U.S. Navy.

References

Alibali, M. W., Bassok, M., Olseth, K. L., Syc, S., and Goldin-Meadow, S. (1995). Gestures reveal mental models of discrete and continuous change. In *Proceedings of the seventeenth annual conference of the Cognitive Science Society*, pages 391–396.

Argall, B. D. and Billard, A. G. (2010). A survey of tactile human-robot interactions. *Robotics and Autonomous Systems*, 58(10):1159 – 1176.

Billard, A., Bonfiglio, A., Cannata, G., Cosseddu, P., Dahl, T., Dautenhahn, K., Mastrogiovanni, F., Metta, G., Natale, L., Robins, B., Seminara, L., and Valle, M. (2013). The roboskin project: Challenges and results. In Padois, V., Bidaud, P., and Khatib, O., editors, *Romansy 19 Robot Design, Dynamics and Control*, volume 544 of *CISM International Centre for Mechanical Sciences*, pages 351–358. Springer Vienna.

Breckinridge Church, R. and Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23(1):43–71.

Goldin-Meadow, S. (2007). Gesture with speech and without it. *Duncan Cassell Levy*, pages 31–50.

Ji, Z., Amirabdollahian, F., Polani, D., and Dautenhahn, K. (2011). Histogram based classification of tactile patterns on periodically distributed skin sensors for a humanoid robot. In *RO-MAN, 2011 IEEE*, pages 433–440.

LeapMotion (2014). http://www.leapmotion.com/.

Yuan, Y. and Barner, K. (2006). An active shape model based tactile hand shape recognition with support vector machines. In *Information Sciences and Systems, 2006 40th Annual Conference on*, pages 1611–1616.