Model-free Contact Localization for Manipulated Objects using Biomimetic Tactile Sensors

Artem Molchanov, Oliver Kroemer, Zhe Su and Gaurav S. Sukhatme

Abstract— **Manipulation tasks often require robots to make contact between a grasped tool and another object in the robot's environment. The ability to detect and estimate the positions and directions of these contact points is crucial for monitoring the progress of the task, and detecting failures. In this paper, we present a data-driven approach for detecting and localizing contacts between a grasped object and the environment using tactile sensing. We explore framing the contact localization as both a regression and a classification problem and train neural networks accordingly to estimate the contact parameters. We also compare the neural networks with Gaussian process regression and support vector machine classification with spatio-temporal hierarchical matching pursuit feature learning. We evaluate the presented approach using hundreds of contact events on eighteen objects with different shapes, sizes and material properties. The experiments show that the neural network approach can learn to localize contact events for individual objects with a mean absolute error of less than 2.5 cm for the positions and less than 10° for the directions.**

I. Introduction

Many manipulations in unstructured environments require a robot to use a grasped object, i.e., a tool, to interact with other objects. Often, a specific part of the tool, such as the tool's tip, needs to make contact with another object to perform the manipulation. For example, a hammer should make contact with a nail on the flat surface of its head when performing a hammering task. By sensing if and where the tool has made contact, the robot can verify that it is performing the skill correctly and otherwise adapt the skill accordingly.

The task of localizing contacts has usually been approached using either wrist-mounted force-torque sensors or joint torque sensors. However, approaches based on these sensors face challenges such as bias drift and, in some cases, they require an accurate model of the robot and the grasped object. Tactile sensors provide another sensor modality that can be used to estimate contact parameters. Recent developments in the tactile sensor technologies have provided robots with human-like tactile signals [1]. The large amount of data provided by these sensors could potentially result in significantly more robust manipulation skills for robots. However, in order to fullfil this potential, the robots will also need suitable estimation methods to process the tactile data.

In this work, we explore using machine learning methods to estimate contact parameters between grasped objects and the environment based on data from biomimetic tactile sensors (BioTacs) [1]. In particular, we investigate estimating the positions and (force) directions of contacts using neural network (NN) classification and regression, Gaussian process regression (GPR), and support vector machine (SVM) classification with features learned using spatio-temporal hierarchical matching pursuit (ST-HMP) [2]. We evaluate the methods using data collected from 18 objects with different shapes, sizes and materials. In our work, we rely on a few assumptions. First, we restrict our contacts to the transient type, i.e., short taps or bumps. We also restrict our investigation to a single contact point and perform object dependent learning.

The key contributions of this work are: a) a model-free approach for estimating contact positions and directions between the environment and a grasped object based on tactile signals, b) an accurate labelled dataset[1] for evaluating and benchmarking contact localization methods, and c) an evaluation of the presented approaches using real robot experiments.

II. Point-of-Contact Estimation

Our contact learning pipeline consists of two main parts: contact detection and contact localization.

A. Contact Detection and Feature Extraction

We detect contact events by applying threshold on the high-pass filtered pressure vibration (PAC) signals extracted from the BioTac sensor. In order to remove jitter, closely located events are reduced into a single event using DB-SCAN [3] algorithm. The resulting clusters in time dimension define the beginning and the end of every contact event.

To form a feature vector for localizing the contact point, we extract the values of the tactile signals (electrodes (Electr), pressure vibration (PAC), DC pressure (PDC) after a contact event is detected. We use a window of $t = 25$ consecutive time steps at a sampling rate of 100Hz from the beginning of the event. In order to reduce the influence of gravity on the baseline values of tactile readings, we subtract the average signal values of the three time steps immediately before the contact event.

B. Regression

We parametrize the location of the contact point using its Cartesian position (*x, y, z*) relative to the robot's palm. The direction of the contact is parametrized using *yaw* and *pitch*.

Artem Molchanov, Oliver Kroemer and Gaurav S. Sukhatme are with the Department of Computer Science; Zhe Su is with the Department of Biomedical Engineering, University of Southern California, Los Angeles. {molchano, okroemer, zhesu, gaurav}@usc.edu.

[1]The dataset is available at `http://bicl.robotics.usc.edu`

Fig. 1: The 18 objects and an example of a grasp used for data collection.

We compare two different machine learning techniques for learning the contact estimation function: NN and GPR.

The NN architecture consists of two fully connected hidden layers with 900 neurons each. The training of the NN is performed in a supervised manner using stochastic gradient descent with the Euclidean quadratic loss function.

We compare NN regression with GPR, a state-of-the-art non-parametric Bayesian supervised learning approach with automatic relevance determination (ARD).

Due to the high computational cost of GPs, we compute the average signal values during contact event window and use them as features.

C. Classification

In the classification approach, we represent the point-of-contact in the form of a distribution over the discretized contact pose parameters.

For NN we use the same architecture for the estimator, however for classification the NN output is converted to the distribution over labels using softmax. Furthermore, we use RMSprop adaptive learning rate with cross-entropy classification loss for training. We train separate classifiers for each contact pose parameter (*x,y,z,yaw,pitch*) to reduce number of classes for learning.

We also apply linear SVM in combination with ST-HMP [2] to perform the classification.

III. Evaluation

A. Data Collection Setup

The goal of this experiment is to evaluate the accuracy of the contact localization using the proposed methods. The experiments were performed using a three-fingered Barrett hand. Each finger tip is equipped with a BioTac tactile sensor.

Eighteen objects with a variety of sizes and materials were chosen for this experiment (see Fig. 1). During the data collection, the robot held one of the objects (see Fig. 1 for an example grasp) while a person tapped the object with a Vicon-tracked plastic rod.

The BioTac sensor readings were then extracted using the contact detection method described in Section II-A. Using this approach, we collected ≈ 15100 samples for all 18 objects.

B. Results and Discussion

Fig. 2 shows the mean absolute error (MAE) calculated from errors of all 18 objects. We evaluated NN regression using electrode features (*NN:Electr*) and using the full feature set (*NN:Electr+PAC+PDC*). We also evaluated GPR

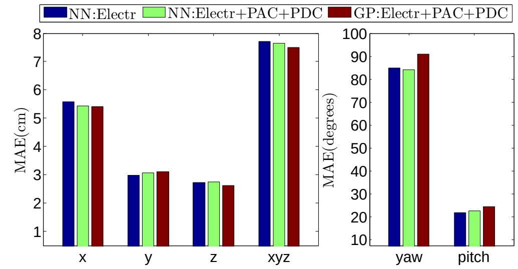

Fig. 2: Results of regression using different sensor modalities for NN and GP regressors.

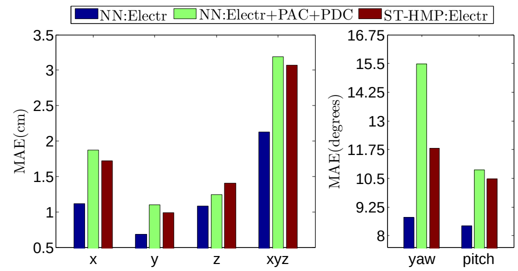

Fig. 3: Results of classification using different sets of sensor modalities for NN and SVM classifiers.

using the full feature set (*GP:Electr+PAC+PDC*). The three regression approaches resulted in considerable errors that, in some cases, exceed 50% of the object's size. Such significant errors are probably caused by ambiguities in the mapping between features and the estimated contact parameters, which can not be represented properly by the regression. These results motivated us to approach our problem from the point of classification.

Similar to the regression approach, we also evaluated NN classifiers using only the electrode features and using the full set of features. For this experiment, we pick a 1cm/5° grid with 25 time steps as a baseline parameter set. Fig. 3 presents the MAE across the test sample sets of all 18 objects using our classification approach described in Section II-C. We also combine predictions of individual dimensions in Cartesian coordinates for every sample in order to calculate the average Euclidean norm of the error vector for all location predictions (*xyz* on the figure). The results indicate that electrodes are the most relevant features for contact localization. Incorporation of PAC and PDC injects additional noise and leads to overfitting. Fig. 3 also shows that NN outperforms ST-HMP approach for both Cartesian and angular coordinates (for more details about the work please see [4]).

References

[1] N. Wettels, V. J. Santos, R. S. Johansson, and G. E. Loeb, "Biomimetic tactile sensor array," *Advanced Robotics*, vol. 22, no. 8, pp. 829–849, 2008.

[2] M. Madry, L. Bo, D. Kragic, and D. Fox, "ST-HMP: Unsupervised Spatio-Temporal Feature Learning for Tactile Data," in *IEEE International Conference on Robotics and Automation (ICRA)*, may 2014.

[3] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Kdd*, vol. 96, no. 34, 1996, pp. 226–231.

[4] A. Molchanov, O. Kroemer, Z. Su, and G. S. Sukhatme, "Contact localization on grasped objects using tactile sensing," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2016.