Online Behavior Adaptation utilizing a Data Driven Control Policy

Erik Berger[1], David Müller[1], David Vogt[1], Bernhard Jung[1], Heni Ben Amor[2]

***Abstract*— Autonomous robot behaviors need to be adjusted for perturbations caused by changing environmental conditions. For measuring such perturbations, robots often require special-purpose sensors, e.g. force-torque sensors, that are however often heavy, expensive and prone to noise. We propose a machine learning approach for measuring external perturbations of robot behavior that uses a wide variety of commonly available, low-cost sensors only. During a training phase, behavior-specific statistical models of sensor measurements, so-called perturbation filters, are constructed using Principal Component Analysis, Transfer Entropy and Dynamic Mode Decomposition. During behavior execution, perturbation filters compare measured and predicted sensor values for estimating the amount of external perturbations. Such perturbation filters can therefore be regarded as virtual sensors that produce continuous estimates of environmental changes which are relevant for the execution of the behavior and can be used to establish a control policy to adjust the robot's actions.**

I. Introduction

Autonomous robots require accurate sensing capabilities in order to act in an intelligent and meaningful way within their environment. In particular such robots require sensors for measuring external perturbations, e.g., physical contact with human or environmental changes. Recorded measurements can be used by a robot to ensure safety during behavior execution and to react to physical perturbations. To this end, it is important that both the magnitude of an external perturbation and its relevance to the current behavior is reliably detected. Various researchers have proposed the so-called *soft robotics* paradigm: compliant robots that "can cooperate in a safe manner with humans" [1]. An important robot control method for realizing such a compliance is impedance control [2]. Impedance control can be used to allow for touch based interaction and human guidance. To this end, impedance controllers require accurate sensing capabilities, in the form of force-torque sensors. The main disadvantages of such sensors are that precise ones are typically heavy and expensive while cheap ones suffer from significant noise. However, there are numerous affordable low-cost sensors available which, while not directly measuring perturbation forces, can be used to generate estimates of external perturbations.

In this paper, we present an approach for perturbation estimation which is based on a combination of low-cost sensors and machine learning techniques. During a training phase, we extract a compact representation, called a *perturbation filter*, which specifies the evolution of sensor readings during an undisturbed execution of a motor skill. The extraction is guided by information-theoretic measures such as Transfer Entropy, that determine the relevance of a specific sensor w.r.t. the executed robot behavior. In contrast to our previous work [3] [4], we will not use any higher level stability parameters, such as the center-of-mass, center-of-pressure, or zero-moment-point for learning. Instead, we will learn the perturbation filter from low-level sensor data, solely. As a result, no knowledge about the robot kinematics or dynamics is required.

After a perturbation filter has been learned, it is used to generate a continuous estimate of the external perturbation magnitude. A huge advantage of our approach is that the perturbation value is measured in the parameter space of the executed behavior and therefore can be used to facilitate a control policy for the robot's behavior. The presented perturbation filter can be regarded as a virtual force sensor on the one hand and a data-driven control policy for online reaction on the other hand.

Fig. 1: A NAO robot estimates the influence of external perturbations applied by a human interaction partner to its current behavior execution.

II. Approach

The objective of the presented method is to learn a behavior specific control policy by inferring estimates of

[1]Institute of Computer Science, Technical University Bergakademie Freiberg, Bernhard-von-Cotta-Str. 2, 09599 Freiberg, Germany

[2]Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, 801 Atlantic Drive, Atlanta, GA 30332-0280, USA

external perturbations from low-cost sensor readings such as joint angles, current or an inertial measurement unit (IMU). First, we record training data for a behavior with different parameter configurations, e.g., varying step lengths during walking. In this data acquisition phase, no external perturbations are applied. Thereafter, the training data is used to create a *Feature Space data model* during feature extraction. Linear combinations of different sensors are weighted by their relevance to the observed parameter and projected into the low-dimensional Feature Space. In the following, the configuration parameter will be referred to as the *target vector*. The relevance of a specific sensor to the target vector is extracted using Transfer Entropy [5] (TE). In this context, TE is used as a measure of predictability and information flow between the target vector and the conduct of sensors. Sensors that have a high TE w.r.t. the robot's behavior are deemed more influential and relevant.

During behavior execution, an external perturbation is detected by comparing the recorded training data to the current sensor data within the low-dimensional Feature Space. Dynamic Time Warping (DTW) [6] is used as a distance function in order to include the temporal pattern for the comparison. The estimation of a *perturbation value* is performed by comparing the current sensor readings to the sensor readings acquired during training. The perturbation value is then inferred from the difference between the currently configured behavior parameter, e.g., the currently employed step length, and the estimated behavior parameter which produced similar sensor readings during training. A detailed step by step description of the approach can be found in a previous publication [7]. However, in this paper we utilize the resulting perturbation value to control the robots behavior.

In the following section, we will depict our control policy in more detail.

III. Control Policy

Since we extracted only features which are relevant to the corresponding behavior, we constrain the perturbation estimation to changes in the environment which are directly influencing the execution of the behavior.

Accordingly, we can generate an estimate for the amount of possible perturbations by calculating the difference between the configured behavior parameter used to control the robot and the estimated behavior parameter identified by the learned model. This difference is being used to generate a control policy to respond to external changes which influence the execution of the behavior in real time. For instance, our approach can be used in scenarios where a robot has to detect and react to external perturbations in order to fulfill a specified task as shown in a video which can be found here: `http://youtu.be/wHZYx6Dzswk`.

Furthermore, certain groups of sensors are suitable to qualify certain perturbations, allowing additional conclusions about the characteristics of a possible perturbation. For instance, lifting up a robot while executing a walking gait does not directly disturb the joints of the robot but is recognized by the group of pressure sensors attached to the robots feet. In consequence, the combination of sensor groups further increases the variety of recognizable perturbations.

IV. Conclusion

In this paper, we implied an approach for estimating external perturbations. Instead of using expensive sensors, e.g., force-torque sensors, we leverage the available information from different low-cost sensors. We introduced a machine learning approach that can learn software based behavior-specific perturbation filters. Transfer Entropy, an information-theoretic measure, is used to guide the feature extraction process. Given a set of low-level sensor data, our approach allows for the automatic identification and scaling of relevant sensor values by calculating the influence of sensors on future robot states. In turn, these filters can be used to generate continuous estimates for perturbations which are influencing the execution of the behavior. We have shown that the computed perturbation value can be used to control a robots walking behavior in a cooperative human-robot interaction task.

However, other schemes of responses for autonomous robots are feasible. For example, competitive soccer robots have to react to their opponents behavior differently than industrial robots to perturbations in assembling tasks. For that, we are currently investigating mathematical models that allow varying control policies. Furthermore, in our system we assume a Gaussian distribution over the noise. In the future, we want to investigate the use of more complex models in order to cope with highly nonlinear noise in every time step.

References

[1] A. Albu-Schäffer, O. Eiberger, M. Fuchs, M. Grebenstein, S. Haddadin, C. Ott, A. Stemmer, T. Wimböck, S. Wolf, C. Borst, and G. Hirzinger, "Anthropomorphic soft robotics from torque control to variable intrinsic compliance," in *Robotics Research*, ser. Springer Tracts in Advanced Robotics, C. Pradalier, R. Siegwart, and G. Hirzinger, Eds. Springer Berlin Heidelberg, 2011, vol. 70, pp. 185–207.

[2] S. Haddadin, *Towards Safe Robots - Approaching Asimov's 1st Law*. Springer, 2014.

[3] E. Berger, D. Vogt, N. Haji-Ghassemi, B. Jung, and H. Ben Amor, "Inferring guidance information in cooperative human-robot tasks," in *Humanoids'13*, 2013.

[4] E. Berger, M. Sastuba, D. Vogt, B. Jung, and H. Ben Amor, "Estimation of perturbations in robotic behavior using dynamic mode decomposition," *Advanced Robotics*, 2014.

[5] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, 2000.

[6] H. Sakoe, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, pp. 43–49, 1978.

[7] E. Berger, D. Müller, D. Vogt, B. Jung, and H. Ben Amor, "Transfer entropy for feature extraction in physical human-robot interaction: Detecting perturbations from low-cost sensors," *Humanoids'14*, 2014 (to appear).