Extracting task constraints as a middle layer between low level control and high level planning

Lucia Pais and Aude Billard
Learning Algorithms and Systems Laboratory,
École Polytechnique Fédérale de Lausanne, Switzerland
Email: {lucia.pais, aude.billard}@epfl.ch

***Abstract*—In robot Programming by Demonstration, current approaches for learning motion models rely on encoding the demonstrated trajectories using statistical techniques. In this paper we propose a method for extracting task constraints from demonstrated motions and using them directly as continuously embeddable constraints for controlling the robot. We consider determining the object of interest in each region of the task (frame of reference), and the contribution of the variable of interest, position vs. force on each axis. Furthermore the demonstrated motion can be segmented into meaningful segments based on the change of the task constraints.**

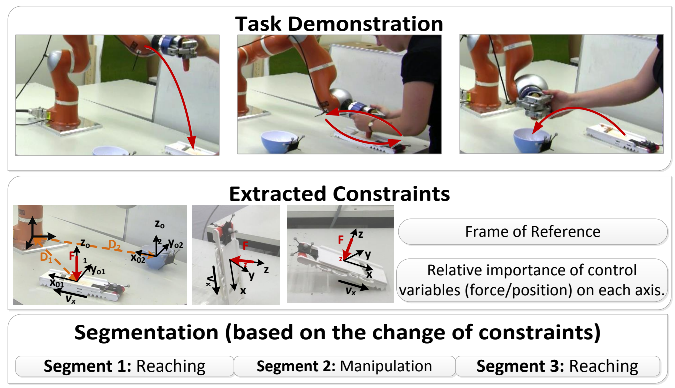

Fig. 1. Encoding a demonstrated task based on extracted constraints and segmenting it into reaching and manipulation parts when constraints change.

I. Introduction

In robotic applications a common way of acquiring skills is Programming by Demonstration (PbD) in which a set of demonstrated trajectories are encoded using statistical techniques. This method has the advantage of allowing the robot to deal with complex motions that are hard to be described analytically, however for achieving a good generalization the task constraints should also be considered.

This paper addresses the problem of constraint-based encoding of high-level tasks, demonstrated to the robot using kinesthetic teaching. These require the completion of a series of actions, such as reaching motions, and manipulation. When analyzing the demonstrated task the difficulty consists firstly in accounting for the large variability that may exist between demonstrations and thus deciding which features of the motion should be reproduced (extracting the task constraints), and secondly expressing these features in relation to the objects involved in the task (extracting the frame of reference (RF)).

Based on the extracted constraints, a model of the motion can be learned in the local frame of reference. The advantage of having a constraint-based encoding of the task ensures flexibility in the task representation, and allows the motion to be reproduced using a single controller and embedding the constraints at run time.

Furthermore we address the problem of deciding whether position control or force control is more suitable for each axis. Usually the control is done in a fixed frame of reference, but determining the right frame of reference for each part of the task can simplify the control, such as performing position and force control on orthogonal axes. In this work we infer the frame of reference from human demonstrations, as well as the axis specific suitable control scheme.

II. Related Work

In our previous work we presented an approach for extracting the frame of reference in a time dependent manner [2], and a more recent approach allows to build a time independent representation of the task using dynamical systems [3]. Here we extend this approach by extracting the importance of variables used for control, such as force and position, and also the frame of reference where they apply.

In our current work we build on these existing approaches in order to extract not only kinematic constraints but also constraints related to forces to be applied on the objects involved in the task. Moreover we consider the problem of segmenting the motion into meaningful sub-parts when the constraints change.

Common approaches for motion segmentation [12] rely on either (1) identifying changes in a variable, like zero-crossings [10]; (2) classification based on existing motion primitives that algorithms use for prior training [7, 11, 6]; or (3) clustering similar motions using unsupervised learning [5]. The downside of most of these approaches is the need of prior knowledge about the task. This may be poor and incomplete according to real-life situations or at times unavailable. Moreover they are sensitive to the encoded variables and raise difficulties when applied to robot control for reproducing the demonstrated task, requiring additional data precessing, or may be specific to a particular context [9].

III. Extracting Task Constraints

While the topics of extracting task constraints and performing segmentation have been addressed previously, our work proposes a one shot algorithm that extracts all the necessary information from human demonstrations of complex tasks. This is designed as a bootstrapping process preceding learning a task model, and it doesn't require any prior information about the type or goal of the motion, see Fig. 1.

The obtained information can be used for both (1) low level control by learning motion models based on the important task variables and (2) for high level planning by determining the right sequence of events in a given situation.

For this we analyze the demonstrated data focusing on the intrinsic variability that exists between demonstrations. We assume that the parts of the task where the demonstrator was coherent (showing consistency between demonstrations) are the features of the motion that should be reproduced [8].

This approach allows us to include at the level of task encoding all the necessary information for successfully completing the task. More specifically we infer from the demonstrations the relative importance of the variables that can directly be used for control, for each axis, such as end effector position and forces or torques to be applied. These variables are expressed relative to the objects involved in the scene. When reproducing the task we use a generic impedance controller in which we modulate the stiffness based on the relative contribution of the control variables determined previously.

All the required information is extracted from the human demonstrations of the task. The advantage of encoding more task information in the low level controllers is that if the objects' position and orientation change while performing the task the robot can adapt on the fly using a reactive behavior, ensuring the successful execution of each of the low level skills, while the planner can combine the demonstrated skills to achieve complex tasks.

Furthermore, in the demonstrated motion a segmentation point can be created segmented when a change occurs in the extracted constraints. This allows learning different motion or manipulation models for each part of the task.

This method of performing segmentation has two main advantages. First using task features ensures that the motion segments are consistent with the human's mental model of the task. Dividing the task into reaching motions and manipulation sub-parts allows the human to easily keep track of the robot's performance while the robot can ask for additional demonstrations of particular segments of the task.

Secondly the individual segments can be used by a high level planner [1] for scheduling more complex tasks. While the problem of using PbD to generate high level planning models has been addressed before [4], the approach presented here allows the planner to reorder the task sequence based on the available objects and actions.

This approach for extracting task features and using them as continuously embeddable constraints in the low level controller, has been validated on a common kitchen task (grating vegetables), as discussed in [8].

IV. Conclusions

In this paper we briefly presented an approach for extracting task constraints from human kinesthetic demonstrations, that allows a unitary representation of the skills to be acquired. The extracted constraints are the frame of reference to be used, and the contribution of position vs. force in each part of the task, based on which a weighting factor can be computed. This can directly be used for reproducing the task using an impedance controller, as described in [8].

The robot can learn how to perform a task without prior knowledge about the type or goal of the task. Also this approach can contribute to abstract action representation by adding proprioceptive information acquired through kinesthetic training. This information is key to successfully executing tasks in which the force to be applied matters, such as grating, cutting, slicing etc. Furthermore the demonstrated set of motions can be segmented when the constraints change, dividing the task into sub-parts that a high level planner can use in executing more complex set of tasks.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no288533 ROBOHOW.COG.

References

[1] M. Beetz, L. Mösenlechner, and M. Tenorth. CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. In *IEEE/RSJ IROS*, 2010.

[2] S. Calinon, F. Guenter, and A. Billard. On learning, representing and generalizing a task in a humanoid robot. *IEEE SMC, Part B*, 37(2): 286–298, 2007.

[3] S. Calinon, Z. Li, T. Alizadeh, N. G. Tsagarakis, and D. G. Caldwell. Statistical dynamical systems for skills acquisition in humanoids. In *Humanoids*, 2012.

[4] R. Jkel, S.W. Rhl, S.R. Schmidt-Rohr, M. Lsch, Z. Xue, and R. Dillmann. Layered programming by demonstration and planning for autonomous robot manipulation. In *Advanced Bimanual Manipulation*, volume 80, pages 1–57. 2012.

[5] D. Kulic, W. Takano, and Y. Nakamura. Combining automated on-line segmentation and incremental clustering for whole body motions. In *ICRA*, 2008.

[6] D. Kulic, C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura. Incremental learning of full body motion primitives and their sequencing through human motion observation. *I. J. Robotic Res.*, 31(3):330–345, 2012.

[7] O. Mangin and P.Y. Oudeyer. Learning to recognize parallel combinations of human motion primitives with linguistic descriptions using non-negative matrix factorization. In *IROS*, pages 3268–3275, 2012.

[8] A. L. Pais, K. Umezawa, Y. Nakamura, and A. Billard. Learning robot skills through motion segmentation and constraints extraction. HRI Workshop on Collaborative Manipulation, 2013.

[9] C. Sminchisescu, A. Kanaujia, and D. Metaxas. Conditional models for contextual human motion recognition. *Comput. Vis. Image Underst.*, 104 (2):210–220, 2006.

[10] W. Takano and Y. Nakamura. Humanoid robot's autonomous acquisition of proto-symbols through motion segmentation. In *Humanoids*, pages 425–431, 2006.

[11] L. Tao, E. Elhamifar, S. Khudanpur, G. Hager, and R. Vidal. Sparse hidden markov models for surgical gesture classification and skill evaluation. In *Information Processing in Computer-Assisted Interventions*, volume 7330, pages 167–177. Springer Berlin / Heidelberg, 2012.

[12] Liang W., Weiming H., and Tieniu T. Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585–601, 2003.