Dynamically Consistent Probabilistic Model for Robot Motion Learning

Diego Pardo
Italian Institute of Technology - IIT
diego.pardo@iit.it

Leonel Rozo, Guillem Alenyà and Carme Torras
Institut de Robòtica i Informàtica Industrial CSIC-UPC
{lrozo,galenya,torras}@iri.upc.edu

***Abstract*—This work presents a probabilistic model for learning robot tasks from human demonstrations using kinesthetic teaching. The difference with respect to previous works is that a complete state of the robot is used to obtain a consistent representation of the dynamics of the task. The learning framework is based on hidden Markov models and Gaussian mixture regression, used for coding and reproducing the skills. Benefits of the proposed approach are shown in the execution of a simple self-crossing trajectory by a 7-DoF manipulator.**

I. INTRODUCTION

Learning by demonstration endows robots with capabilities to acquire skills taught by a human teacher from kinesthetic demonstrations. This work focuses on robot learning of motor skills, i.e., encoding the dynamics of a robotic task. Note that learning approaches to encode robot motions are proposed either from a dynamic systems perspective (e.g., dynamic motion primitives [1], [2]) or by means of probabilistic models (e.g., Gaussian mixture models [3]).

In particular, probabilistic approaches have been extensively used for learning trajectory-following tasks [4] and recently force-based skills [5]. Nevertheless, a gap in the encoding of the tasks is evident in these works, where the dynamics of the robot during the execution of the skills is not modeled. In this context, a learning framework based on Hidden Markov model (**HMM**) and Gaussian mixture regression (**GMR**) is presented in [6] for the encoding and reproduction of the dynamics of a robotic motion. In such work the GMR process performs as a dynamical system. This feature allows to reproduce the task without the explicit dependency of time, which is a key issue for generalization. However, it is worth noting that the GMR process is not a dynamical system by itself, and care need to be taken when selecting the type of variables that are being modeled and the input/output regression used in order to obtain a dynamically consistent behavior. In this abstract we propose a redefinition of the state used for codification in [6]. This new approach allows to model tasks with ambiguities, results are shown on a 7-DoFs manipulator completing a robot trajectory with crossing points.

II. BACKGROUND

The approach in [6] uses the robot operational space positions $\boldsymbol{x} \in \mathbb{R}^3$ and velocities $\dot{\boldsymbol{x}} \in \mathbb{R}^3$, collected during demonstrations, to capture the motion of a given task by means of a joint probability distribution $\mathcal{P}(\boldsymbol{x}, \dot{\boldsymbol{x}})$. Then, a probabilistic model of the dynamics of the task is obtained,

$$\hat{\dot{\boldsymbol{x}}} \sim \mathcal{P}(\dot{\boldsymbol{x}}|\boldsymbol{x}). \tag{1}$$

where a target velocity $\hat{\dot{x}}$ for the system can be derived given the current position of the robot end-effector $\boldsymbol{x}$. Specifically, $\mathcal{P}(\boldsymbol{x}, \dot{\boldsymbol{x}})$ is encoded in an HMM, which defines the task as a sequence of N_s "high-level states" (HMM-states).

When using Eq. (1) as control command, instantaneous changes in velocities are requested irrespective of the current velocity of the system, generating undesired robot behavior and forcing the use of high low-level gains trying to minimize trajectory divergences. This was also addressed in [6] by proposing a complementary GMR, where the roles of the variables (input/output) were swaped,

$$\hat{\boldsymbol{x}} \sim \mathcal{P}(\boldsymbol{x}|\dot{\boldsymbol{x}}), \tag{2}$$

to compute a target position $\hat{x}$ given the current velocity. Eqs. (1) and (2) are then combined in an attractor to generate acceleration commands.

These approaches have poor generalization performance in tasks where the same position is visited along the motion (e.g., motion with crossing points), given that the controller cannot correctly discriminate between states. This problem can be overcome by using high-level information of the task [6], that is encapsulated in the sequence of HMM-states. In contrast to this approach, this work focuses on the adequate modeling of the task dynamics.

III. EXTENDED STATE FOR HMM-MOTION ENCODING

Here we propose to tackle the problem of modeling the dynamics of a robotic motion task using an appropriate representation of the task state. Let us define $\boldsymbol{s}_k$ as the state of the task at instant k. The dynamics of the motion is then given by

$$\boldsymbol{s}_{k+1} = f(\boldsymbol{s}_k), \tag{3}$$

which is a homogeneous system, i.e., no independent input is presented. Given that function f captures the evolution of the task, the output of Eq.(3) can be used as reference for a low-level control system that decides the control action, $\boldsymbol{u}_k$, (usually forces) to be applied to the robot, i.e., $\boldsymbol{u}_k = h(\boldsymbol{s}_k, \boldsymbol{s}_{k+1})$.

Thus, given that a motion is being modeled, the state must include both the position and velocity of the system, namely $\boldsymbol{s}_k = [\boldsymbol{x}_k, \dot{\boldsymbol{x}}_k]^\top$. Moreover, in order to capture its dynamics, the HMM should also encode the evolution of the motion (i.e.,

This research is partially supported by μRALP (FP7-288233) and IntellAct (FP7-269959) projects.

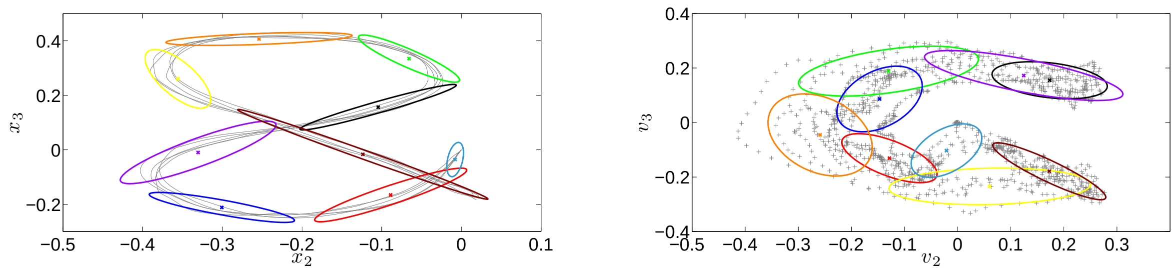


Fig. 1. HMM encoding. *Left*: HMM-States distribution on the position space $\boldsymbol{x}_k$. *Right*: HMM-States distribution on the velocity space $\boldsymbol{v}_k$

the state at $k+1$). The resulting joint probability distribution is given by

$$\mathcal{P}(\dot{\boldsymbol{x}}_k, \boldsymbol{x}_k, \dot{\boldsymbol{x}}_{k+1}, \boldsymbol{x}_{k+1}). \quad (4)$$

Using an extended state and its increment for the generation of the probability density function leads to a multi-output GMR such that

$$\hat{\dot{\boldsymbol{x}}}_{k+1}, \hat{\boldsymbol{x}}_{k+1} \sim \mathcal{P}(\dot{\boldsymbol{x}}_{k+1}, \boldsymbol{x}_{k+1} | \dot{\boldsymbol{x}}_k, \boldsymbol{x}_k). \quad (5)$$

Eq. (5) represents the actual dynamics of the task and the output of this system can be directly used in an attractor similar to the one in [6], but using dynamically consistent values,

$$\hat{\ddot{\boldsymbol{x}}}_{k+1} = \boldsymbol{K}_p(\hat{\boldsymbol{x}}_{k+1} - \boldsymbol{x}_k) + \boldsymbol{K}_v(\hat{\dot{\boldsymbol{x}}}_{k+1} - \dot{\boldsymbol{x}}_k). \quad (6)$$

The attractor imposes the dynamics of the error between the desired behavior, proposed by the GMR, and the actual state of the robot. Thus, gains $\boldsymbol{K}_p$ and $\boldsymbol{K}_v$ can be selected to determine response time and stability for error regulation.

IV. Experiments and Results

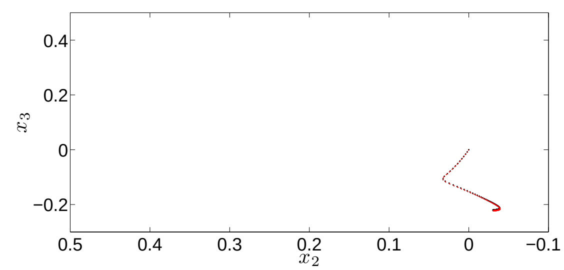
For demonstration, a 2D motion task using the 7-DoFs WAM robot is presented. Fig. 1 shows the motion captured by an HMM from the teacher demonstrations, including positions and velocities. Note that the "ampersand-shape" has two crossing points. This skill is learned using both, the standard and extended frameworks. The corresponding HMM-states are plotted on the input space (see Fig. 1). Note that the projection of the HMM onto the position subspace is similar for both approaches and that the HMM-states are overlapped, nevertheless the extended framework includes the velocity as input, allowing to discriminate among HMM-states of the skill. The performance of both models is presented in Fig. 2. The initial state of the task is provided and the model reproduces the captured dynamics. It can be seen that the extended model completes the skill, adequately resolving the ambiguity.

A video of the experimental results is available at `http://www.iri.upc.edu/groups/perception/RobotMotionLearning`

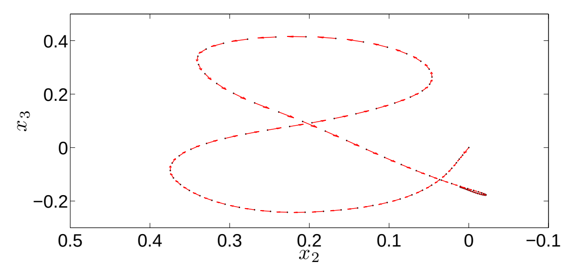
V. Conclusions and Future Work

Compliant motions emerge when a model of the dynamics of the task is used as reference. The target of the motion is the very next state and, in case of human interaction, errors are compensated irrespective of the time evolution. Conversely, when a predefined and time-dependent set of points is provided, the magnitude of the error and the corresponding control actions increase over time.

A straightforward consequence of this adequate modeling of the task dynamics is that skills with ambiguities can be learned.



(a) Standard state representation



(b) Extended state representation

Fig. 2. GMR output corresponding to the "ampersand-shape" motion.*Red* arrows show the direction of the movement.

References

[1] A. Ijspeert, J. Nakanishi, and S. Schaal, "Trajectory formation for imitation with nonlinear dynamical systems," in *IROS*, 2001, pp. 752–757.

[2] D. Pardo, C. Angulo, S. D. Moral, and A. Català, "Emerging motor behaviors: Learning joint coordination in articulated mobile robots," *Neurocomputing*, vol. 72, no. 16–18, pp. 3624–3630, 2009.

[3] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *ICRA*, 2011, pp. 4326–4332.

[4] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitves using the motion refinement tube," *Autonomous Robots*, vol. 31, pp. 115–131, 2011.

[5] L. Rozo, P. Jiménez, and C. Torras, "Robot learning from demonstration of force-based tasks with multiple solution trajectories," in *ICAR*, 2011, pp. 124–129.

[6] S. Calinon, F. D'halluin, D. Caldwell, and A. Billard, "Handling of multiple constraints and motion alternatives in a robot programming by demonstration framework," in *Humanoids*, 2009, pp. 582–588.