Skills Learning in Robots by Interaction with Users and Environment

Sylvain Calinon[1,2]

[1]Idiap Research Institute, Centre du Parc, Rue Marconi 19, PO Box 592, CH-1920 Martigny, Switzerland.
[2]Department of Advanced Robotics, Istituto Italiano di Tecnologia (IIT), Via Morego 30, 16164 Genova, Italy.
(E-mail: sylvain.calinon@idiap.ch)

1. Introduction

The fast technological evolution and dissemination of multimodal sensors and compliant actuators bring a new human-centric perspective to robotics. The variety of human-robot interactions that stem from these new capabilities unveil compelling challenges for machine learning.

An attractive approach to the problem of transferring skills to robots is to take inspiration from the way humans learn by imitation, adaptation and self-refinement. Such learning strategies require various types of interaction with the end-users and with the robot's environment. The overall skill acquisition process can hardly be segmented or sequenced in a specific way in advance. This indicates the importance of finding a representation of skills that can be shared by different learning strategies and that can accommodate multimodal continuous data streams for both analysis and synthesis purposes.

I will show that a variety of approaches such as dynamic movement primitives, optimal control and stochastic optimization can contribute to the skill transfer problem, and that they can be combined together by ensuring that a shared statistical treatment of these processes is provided.

The aim is to provide robots with a representation of rich motor skills able to handle recognition, prediction, synthesis and refinement in a continuous and synergistic way. It also requires to be robust to various sources of perturbation, persistently arising from the environment, from the user, and from the robot. One important challenge in this direction is to devise an encoding scheme that is able to generalize tasks to new situations, that can potentially act in multiple coordinate systems, and that can exploit the modern compliant control capabilities of robots to generate natural, efficient and safe movements for the surrounding users.

I will present an approach exploiting the variability of multiple demonstrations and the co-variability of sensorimotor signals to extract the important characteristics of a task/skill. This information is used within an optimal control strategy to provide the robot with a minimal intervention controller regulating the stiffness and damping characteristics of the robot's actions according to the estimated precision and coordination requirements. For example, the robot is controlled stiffly only when it requires to be accurate, and otherwise remains compliant (for safety and energy efficiency).

Examples of applications with a compliant humanoid, a continuum flexible surgical robot and a set of gravity-compensated manipulators will be showcased.

2. Proposed approach

Tasks in robotics are traditionally decomposed into motion phases (in which robots freely move in the air), and contact phases (in which robots apply desired force patterns to objects/environment with hard contacts). Different controllers are often used to treat one or the other case. This decomposition is however incompatible with applications requiring contacts with soft objects, tools or surfaces. It could restrain the development of versatile systems that could adapt to a wider range of applications, within environments that would be better characterized by local levels of pressures rather than binary contact states (surgical robots moving in-between organs, robots for ocean seabed/space exploration, humanoids walking on carpets and sitting on sofas, etc.).

From this observation, the proposed approach aims at representing movement/force profiles as the behavior/path of a virtual spring-damper system pulling the robot, where a Gaussian mixture model (GMM) is used to encode the attractor manifold, with the variability being used to infer the impedance parameters of the spring-damper system (with the noticeable difference that stiffness and damping parameters are represented as full matrices in order to react to perturbations in a coordinated manner).

We introduced in [1] a probabilistic formulation of dynamic movement primitives (DMPs) [2] encoding the joint evolution of the input (decay term) and the output (forcing terms) within a multivariate GMM, where Gaussian mixture regression (GMR) is used to retrieve at each iteration the forcing terms corresponding to the current input (that can be either time-dependent or time-invariant). In this approach, the movement is not considered as a set of univariate outputs as in standard DMPs. Instead, it also models local correlations among and in-between inputs-outputs (e.g., to discover and re-use local sensorimotor patterns or synergies).

We showed in [3] that the above model can be extended to a *task-parameterized model*, in which several frames of reference interact together to describe the path of virtual springs in several coordinate systems, where the variations from these different perspectives are used to determine how strong (or correlated) these springs should be. In the resulting model, the predicted task variations and couplings are exploited to regulate the impedance of the virtual spring-damper systems acting in several
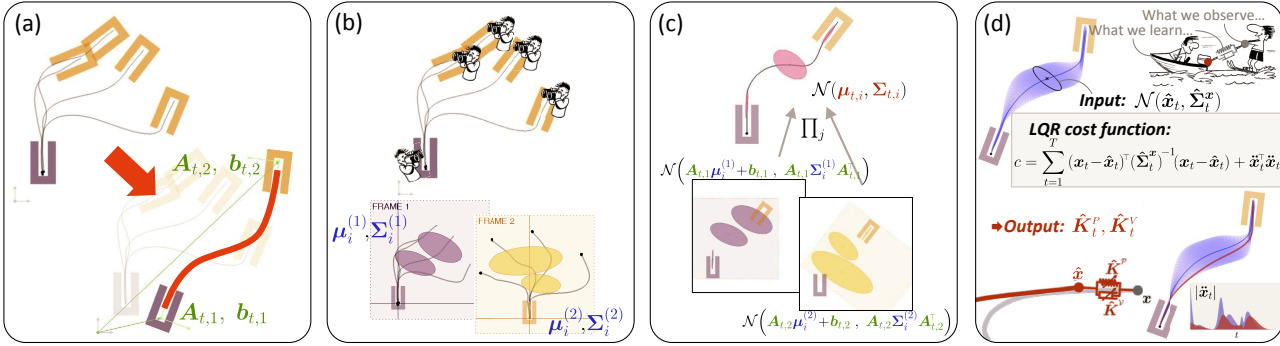
Fig. 1 Illustration of the overall approach combining statistical mixture models, dynamical systems and optimal control to learn and reproduce movement skills. The complete procedure consists of a *demonstration* phase, a *learning* phase, an *online planning* phase and an *impedance regulation* phase.

frames of reference. The approach shares links with optimal feedback control strategies in which deviations from an average trajectory are corrected only when they interfere with task performance, also known as *minimal intervention principle* [4]. Similarly to the solution proposed by Medina *et al.* in the context of risk-sensitive control for haptic assistance [5], the predicted variability can be exploited to form a linear quadratic regulator (LQR) in task space or in joint space, providing a formal way of adapting the impedance parameters, see [3] for details.

Fig. 1 illustrates the overall approach. In the demonstration phase in Fig. 1 - *(a)*, a set of movements is recorded as position and orientation of the robot end-effector. The position and orientation of a set of candidate frames (related to objects in the robot workspace) is also collected. The bottom part illustrates the generalization challenge. Fig. 1 - *(b)* shows the movements observed from the different frames (recording of the same movement from multiple viewpoints). Input and output variables are concatenated for each frame, forming a third order tensor dataset. In this example, time is used as input variable, but a decay term, the robot state or other external object position variables can similarly be employed, see [1] for details. In this learning phase, a task-parameterized mixture model is fit to the tensor dataset by following an *expectation-maximization* (EM) procedure. This training set can then be discarded.

Fig. 1 - *(c)* shows the reproduction phase, for a situation involving new position and orientation of objects, the learned model is first used to estimate a temporary Gaussian mixture model (GMM), that is automatically updated if the position/orientation of objects changes. Depending on the application, this temporary GMM either needs to be updated at each time step (e.g., adapting movements to moving targets), or for each new reproduction attempt (planning approach). In Fig. 1 - *(d)*, *Gaussian mixture regression* (GMR) is then used to retrieve information about the reference to track in a statistical manner, represented as the expected path of a virtual spring-damper system. This strategy of encoding the attractor manifold can be used in free space to encode a movement, but it can also be used to encode reactive behaviors and skills

requiring a desired force profile to be applied, as well as other skills in-between (e.g., soft contacts). Finally, the predicted variability information is used by a linear quadratic regulator to form a minimal intervention controller, see [3] for details.

3. Conclusion

This abstract presented an approach capable of adapting the centers and covariance matrices of a GMM to external task parameters represented as frames of reference. This *task-parameterized model* can be applied to different learning contexts, such as learning from demonstration or direct policy search, with the goal of compactly encoding and generalizing a task to new situations with both interpolation and extrapolation capability. The model can then be extended to a virtual spring-damper system with variable impedance gains. For new position and orientation of the frames, the system generates a probabilistic flow tube predicting the path of the virtual spring and its variations. This covariance information is exploited within an optimal control strategy to locally reduce the control commands according to the varying precision and coordination required throughout the task. The general goal of the approach is to devise learning and control strategies that allow end-users and robots to safely collaborate and interact in a shared workspace.

References

[1] S. Calinon, Z. Li, T. Alizadeh, N. G. Tsagarakis, and D. G. Caldwell, "Statistical dynamical systems for skills acquisition in humanoids," in *Proc. IEEE Intl Conf. on Humanoid Robots (Humanoids)*, Osaka, Japan, 2012, pp. 323–329.

[2] A. Ijspeert, J. Nakanishi, P. Pastor, H. Hoffmann, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2013.

[3] S. Calinon, D. Bruno, and D. G. Caldwell, "A task-parameterized probabilistic model with minimal intervention control," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Hong Kong, China, May-June 2014, pp. 3339–3344.

[4] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, vol. 5, pp. 1226–1235, 2002.

[5] J. R. Medina, D. Lee, and S. Hirche, "Risk-sensitive optimal feedback control for haptic assistance," in *IEEE Intl Conf. on Robotics and Automation (ICRA)*, May 2012, pp. 1025–1031.