Robot Compliant Behavior through Reinforcement Learning

Jie Liu, Matteo Saveriano and Dongheui Lee

In tasks which involve contact or require a specific response to physical perturbations, position control alone is insufficient to achieve the goals of the task. In these cases, the robot should have the ability to adapt the impedance for different task requirements in order to achieve better versatility and robustness. Impedance control is one of the most adopted methods controlling the interaction between a manipulator and the environment. For redundant manipulators, different impedance behaviors can be realized for the end-effector and the null-space of the main task [1].

In this work, we proposed a method to learn variable stiffness gains by experience. We build a reinforcement learning system to learn the new trajectory and corresponding null-space stiffness in robot body based on some tasks. In contrast to [3], we use the PoWER algorithm [2] in this system. Furthermore we use this system to learn the variable stiffness in null-space impedance control so that the Cartesian space execution will not be affected.

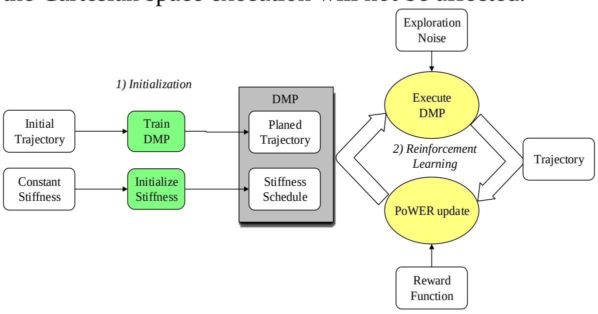

Figure 1: Overview of reinforcement learning system. After initialization, the planed trajectory and gain schedules of a Dynamic Movement Primitive (DMP) are optimized with respect to a reward function with reinforcement learning.

Null-space impedance control

For redundant manipulators it is possible to have some joint-space impedance and task-space impedance simultaneously by projecting the joint space stiffness in the null-space of the Jacobian of the main task.

$$\ddot{q}_c = J^{\dagger}(\ddot{x}_c - \dot{J}\dot{q}) + N(\ddot{q}_d + M^{-1}{}_d(B_d(\dot{q}_d - \dot{q}) + K_d(q_d - q) - \tau_{ext})) \quad (1)$$

where $J^{\dagger}$ is the pseudo inverse of the Jacobian, $\ddot{x}_c$ is the task-space commanded acceleration, N is a $(n \times n)$ matrix which projects joint velocity into the null-space of Jacobian, q are the joint variables, M_d, B_d and K_d are the desired inertia, damping and stiffness matrices respectively.

*This work has been partially supported by the European Community within the FP7 ICT-287513 SAPHARI project.
Authors are with Technische Universität München, Munich, Germany.
{jie.liu, matteo.saveriano, dhlee}@tum.de

Maintaining the Integrity of the Specifications

To learn the null-space stiffness, we built a reinforcement learning system. The system is shown in Figure 1. The initial trajectory is obtained by training Dynamic Movement Primitive (DMP). Initial stiffness is set constant. Then we use PoWER algorithm to optimize the planed trajectory and stiffness according to the reward function. The modified trajectory will be updated through

$$\tau\ddot{y} = \alpha_z(\beta_z(g - y) - z) + b_{tra}^T(\theta + \mathrm{E}) \quad (2)$$

where y is the modified trajectory on robot body, g is a known goal state, α_z and β_z are constants, τ is a temporal scaling factor. θ is the policy parameter, which is updated by PoWER and E is the exploration noise. b_{tra} is nonlinear function composed of Gaussian basis functions of the trajectory. The stiffness update rule is

$$\dot{K}_p = \alpha_K(b_{sti}^T(\theta + \mathrm{E}) - K_P) \quad (3)$$

where K_P is the leaned stiffness, θ and E share the same definition with (2). α_K is a constant. b_{sti} is the vector of basis functions of stiffness.

We performed an experiment on a KUKA LWR4+. A spherical obstacle is located in the environment. We set the reward function so that it leads to good tracking far from the obstacle and to avoid obstacle and to decrease stiffness near the obstacle. Figure 2 shows the task-space error with constant stiffness and learned variable stiffness. One can see that with the learned impedance the robot achieved better task-space execution.

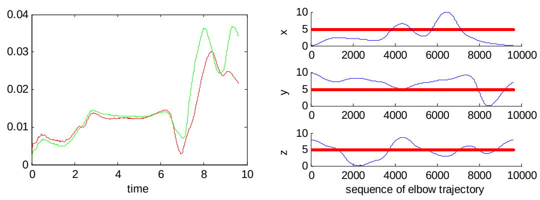

Figure 3: (a) Norm error in task-space. Green line is the task-space error with constant stiffness, red line is task-space error with learned variable stiffness. (b) Constant null-space stiffness (red line) and learned variable null-space stiffness (blue line).

References

[1] H. Sadeghian, L. Villani, M. Keshmiri, and B. Siciliano, Task-space control of robot manipulators with null-space compliance, *in IEEE Transactions on robots*, Vol.30, NO.2, April 2014.

[2] Kober, J., Peters, J. Learning Motor Primitives for Robotics. In *IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, Japan, 12–17 May 2009; pp. 2112–2118.

[3] J. Buchli, F. Stulp, E. Theodorou and S. Schaal. Learning variable impedance control, in *The International Journal of Robotics Research* vol. 30, no. 5, pp. 556-573, 2011.