Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 11 (2017) 329 – 337

Procedia MANUFACTURING

27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy

Reinforcement Learning for Manipulators Without Direct Obstacle Perception in Physically Constrained Environments

Marie Ossenkopf[a]*, Philipp Ennen[b], Rene Vossen[b] and Sabina Jeschke[b]

[a]*Distributed Systems Group, Universität Kassel, Wilhelmshöher Allee 73, 34121 Kassel, Germany*
[b]*Institute of Information Management in Mechanical Engineering, RWTH Aachen University, Dennewartstr. 25, 52068 Aachen, Germany*

Abstract

Reinforcement Learning algorithms have the downside of potentially dangerous exploration of unknown states, which makes them largely unsuitable for the use on serial manipulators in an industrial setting. In this paper, we make use of a policy search algorithm and provide two extensions that aim to make learning more applicable on robots in industrial environments without the need of complex sensors. They build upon the use of Dynamic Movement Primitives (DMPs) as policy representation. Rather than model explicitly the skills of the robot we describe actions the robot should not try to do. First, we implement potential fields into the DMPs to keep planned movements inside the robot's workspace. Second, we monitor and evaluate the deviation in the DMPs to recognize and learn from collisions. Both extensions are evaluated in a simulation

Peer-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing

Keywords: Dynamic Movement Primitives; Relative Entropy Policy Search; Intelligent Manufacturing; Potential Field; Assembly Robot

* Corresponding author. Tel.: +49 561 804-6280; fax: +49 561 804-6277.
E-mail address: marie.ossenkopf@vs.uni-kassel.de

2351-9789
Peer-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing
doi:10.1016/j.promfg.2017.07.115

1. Introduction

Adapting a robotic assembly system to a new task requires manual setup, which represents a major bottleneck for the automated production of small batch sizes. This paper aims at an alternative future approach to the standard programming procedures like e.g. RobotStudio. This approach reduces manual commissioning time through a self-learning manufacturing system [1,2].

Serial manipulators for assembly or pick and place tasks in industrial environments (henceforth called robots) yield high physical danger for themselves and their environment. The physical danger emerges from potential collisions of the robot with obstacles in the environment or itself. Without knowledge or sensing of the obstacles it is not possible to avoid a collision in the first place. As self-learning robots requires trials for learning a new task, the robot suffers from an increasing learning time if the robot tries movements that cannot lead to suitable solution. For instance, if the robot tries to exceed its mechanical constraints with respect to maximum joint angles, maximum velocity and maximum acceleration. Therefore, learning algorithms for use on industrial serial manipulators need to be adapted to meet these problems.

We present two enhancements for the Relative Entropy Policy Search (REPS) algorithm [3] that enable a robot to: (1) detect collisions during the learning process, (2) react to collisions, (3) learn from collisions and (4) avoid planning movements outside the maximum joint angles. The enhancements utilize Dynamic Movement Primitives (DMPs) as policy representation, which are an established policy representation for serial manipulators [4–6]. DMPs map an acceleration onto a position and a velocity in state space by modeling every dimension as a spring-damper-system. To be independent of the kinematic model, we use the dimensions of the robot's joint angle space instead of world coordinates. In particular, the two enhancements are: (1) We integrate potential fields into the DMPs to lower the possibility of exceeding the maximum joint angles. (2) We monitor the deviation between the planned trajectory and the current position to detect collisions and the exceedance of mechanical constraints. This enables us to interrupt a colliding movement and to use the deviation as an evaluation of the policy.

The new features work on any serial robot with an angular encoder. There is no need of additional sensors, an elaborated vision or modeling of the environment. The approach is independent of the knowledge of the kinematic and dynamic model of the robot, so no exact model has to be determined and the algorithm stays unaffected by model errors. The obstacle avoidance can be learned without knowledge of the obstacles. Hence, these additional properties reduce the requirements imposed on the assembly system and the commissioning time. Though we will test the algorithm on a well known and controlled robot, the approach is explicitly aimed towards reconfigurable robotic systems, whose kinematic model is not fully known at commissioning time and which therefore cannot be programmed in TCP space. We tested the algorithm in a simulation of the ABB IRB120[1] robot. As evaluation task we used a simple reaching task with obstacles. Our results show that the exceedance of maximum joint angles can be significantly reduced, that collisions can be detected instantaneously, and that the algorithm learns to avoid experienced collisions.

1.1. Overview

The remainder of this paper is organized as follows. Section 2 summarizes the related work for making a learning process safe on real robots. Furthermore, Section 3 gives a brief explanation of DMPs and the used algorithm. In Section 4 we describe our algorithm setup. Section 5 and 6 outline the two extensions in detail, which are then justified in the evaluation scenario in Section 7. Finally, the paper is concluded in Section 8.

---

[1] http://new.abb.com/products/robotics/industrial-robots/irb-120