**Abstract zl语法修正版0.16**

Nowadays robotic manipulators are widely utilized for milling tasks due to their flexibility as well as relatively low cost. For a robot manipulator, the capability of tracking reference trajectory precisely is one of the most important performance factors, especially for a robot that is applied for industrial use such as welding and laser cutting. However, it is a challenge to guarantee high absolute accuracy in the milling task because of the low stiffness of the robot arm.

The conventional model-based method for controlling the robot manipulator contains a feedforward and feedback controller to compensate for the residual error. This method, however, has a drawback in obtaining a good model of the robot changing over time. The identification of the system model could be very time-consuming. Therefore, this class of controller may not be the best option for such a condition.

Given this situation, a compensation method using Reinforcement Learning is proposed in the thesis to solve the aforementioned problem from the point of view of a model-free, learning controller which improves the tracking performance safely in an online fashion. We focused on implementing the Reinforcement Learning (RL) method realized by the Actor-Critic (AC) algorithm to compensate for the trajectory deviation of the robot arm under external milling force in the simulation environment. Improvement to the control performance is done by feeding this additive RL-based compensation signal given by a learned policy function to the robot’s nominal PID controller.,

This method is implemented on UR10 in the simulation environment, which is a real 6 degrees of freedom (DoF) 3D robotic. Three different reference tracking tasks are designed and executed to evaluate the method. A comparative study is conducted between the pure PID controller and the proposed RL-based controller which contains the PID controller with the RL compensator.

The experimental result shows that the RL-based controllers improve the performance of the nominal control significantly. It indicates that the RL-based compensator is able to meet the requirement of improving the milling accuracy of the robot manipulator.

Keywords: reinforcement learning (RL), actor-critic, reference tracking, robot manipulator.

**cpt1. Introduction**

Nowadays robotic manipulators are widely utilized for milling tasks to replace the Computerized Numerical Control (CNC) machine due to their flexibility as well as relatively low cost. Statistic by the International Federation of Robotics (IFR) indicates that the global sales of industrial manipulators increase steadily [1]. The survey points out that more and more robot manipulators have come into industrial application in the markets. The advantage of robot manipulators over other methods lies in their flexibility, easy-to-adapt feature, and higher number of degrees of freedom (DoF) which enable the execution of difficult configuration in 3-dimension (3D) space.

As one of the most important performances of the robot manipulator, reference or trajectory tracking is widely investigated. The requirement of the task is: Given the desired path, the robot should track the reference path as closely as possible to yield a minimum deviation. The capability to perform this precise tracking is especially crucial for robots that are to be applied in manufacturing industries such as electronics (e.g. parts assembly) and automotive (e.g. spot welding, laser cutting) since a minuscule error could lead to a defect in the product. This thesis is interested in the application of robot manipulators for milling. Nevertheless, it is a challenge to guarantee high absolute accuracy in milling tasks because of the low stiffness of the robot arm. Therefore, it is very essential to investigate the method of controlling the robot manipulator with high precision.

The conventional method for controlling the robot manipulator contains a model-based feedforward controller combining with a feedback controller to compensate for the residual error. This method, however, has a drawback in obtaining a good model of the robot. Furthermore, a robot’s physical properties are changed over time. Such a problem will require the robot to re-perform a system identification which is often unacceptable due to the time it takes and the possibly dangerous movement the robot must execute. Naturally, a solution to this problem is to introduce a controller that is capable of adjusting itself over time. By doing so, the controller will have an extra degree of freedom to compensate for the unknown properties and thus improve the tracking quality. The controllers with self-adjusting characteristics are called learning or adaptive controller. Therefore, a compensation method using Reinforcement Learning is proposed in the thesis to solve the aforementioned problem from the point of view of a model-free, learning controller which improves the tracking performance safely in an online fashion.

This paper proceeds as follows. Section 2 briefly makes an overview of the existing control method used in the field of robot manipulator’s trajectory tracking. We also discuss the advantages and disadvantages of these methods. Then some RL algorithms that can be used are shown in this part. In section 3, the method of implementing the Reinforcement Learning (RL) method realized by the Actor-Critic (AC) algorithm to compensate for the trajectory deviation of the robot arm under external milling force is presented. As a comparison study, pure PID control is also conducted in this part. In Section 4, the results for the previously mentioned 2 methods are presented and analyzed. In Section 5, the meaning and reasons behind the results are discussed. Section 6 draws the conclusion and analyzes the limitations of this study. An outlook on future research activities is given at the end of this research.

**1) Motivation**：

* why do we need to compensate for the deviation of trajectory?

-> Milling force causes deviation from reference trajectory…

* why we use AI to control.

->generalized controller!

（放在背景里 pid是什么 写稍微详细写，对比下优点不足缺陷什么的）

**2) Currently commonly used methods**：

* traditional PID controller.
* other widely-used learning controllers (but not RI-based)，e.g. iterative learning control (ILC) and repetitive control (RC)

**3) Gap**：

The limitations of current methods：e.g.,

* PID is not generalized for another robot arm or workpiece with more complex geometry.
* ILC or RC works badly with any generic nonlinear system.

**4) Introduction in sentences & Goal:**

the compensation method based on reinforcement learning and our goal.

**cpt2. State of art**

1. **Literature review:**

* about AI-Compensation methods based on reinforcement learning

-> list their advantages and limitations

1. **introduce the principle** of Reinforcement learning.
2. **introduce some algorithms** usually used in Reinforcement learning.

**cpt3. Method**

1. **About simulation of milling force**

* How do I simulate the milling force?
* What model do I use?
* In which way do I simplify the real milling condition?

1. **About simulation of the UR10 robot arm**

* How to model UR10 robot arm and milling force on Gazebo?

Loading force as a function of time onto the tool

1. **How to control the trajectory?**

-> The framework/flow chart of control system （core）

1. **Which AI method do I use?**

-> neural network control / fuzzy control / genetic algorithm…

1. **Introduce the specific algorithm I use.**

**cpt4. Results**

1. **Evaluation of the controlling performance**

* ~~target order of magnitude of the error: 0.5mm~~

the real order of magnitude of the error:

* Some graphics to show the target and real trajectory

做出了一定的效果

写一份报告最终证明不是一个很好的方法，训练难度大，不好确定，也不是非常实时的方法，online需要很大的资源，训练的时间过长。不利于成为最终选定的好方法的东西

贴几副图出来

没有任何外界补偿，有pid补偿下，偏移的

加强化学习，偏移有什么作用

1. **~~Comparison with PID and other methods~~**

之前想的是纯pid 和 纯 强化学习

或者和 完全没有控制器下做比较

多做几个案例，正弦力，白噪声 力，轨迹 平滑自然，轨迹2d 矩形（角发生振动）

**cpt5. Conclusion & further research**

1. Is the performance of the RI-based control method better?
2. Further study: from simulation to realization

**References**

**Appendix**

Pid可以先不加

Hiwi

266-274行