网络训练问题描述

问题：控制器设计及仿真验证

从强化学习的角度看，控制器即智能体，我们的目的即训练智能体。在已经完成环境建模的基础上，设计神经网络控制器。该控制器的输入为4个加速度传感器信号，输出为1-4号电机的输入信号。0号电机不接入控制器。控制器的作用是根据0-3号传感器的信号计算1-4号电机的输入信号，以此抑制0号电机引起的系统振动，即目标是使得0-3号传感器信号尽可能接近于0。



我们提供了pyTorch 1.10.0环境下编写的DDPG强化学习程序代码。（请将DDPG\_VC.py\motor0on.xlsx \Wts0126.xlsx放在一个目录下）。

1. 环境部分已经写好。除Reward函数以外，不要修改。
2. 智能体部分采用了DDPG算法，可以该AC网络结构，但是state\_dim（状态空间维度）, action\_dim（动作空间维度）先不要修改。
3. 优先修改超参数、AC网络结构。

可能需要修改的程序行：

1. 超参数

parser = argparse.ArgumentParser()

parser.add\_argument('--mode', default='train', type=str) # mode = 'train' or 'test'

parser.add\_argument('--load\_nn', default=False, type=bool) #是否导入已有网络

parser.add\_argument('--tau', default=0.002, type=float) # target smoothing coefficient

parser.add\_argument('--lr\_A', default=1e-4, type=float) #A网络学习率

parser.add\_argument('--lr\_C', default=1e-3, type=float) #C网络学习率

parser.add\_argument('--gamma', default=1, type=int) # discounted factor

parser.add\_argument('--capacity', default=20000, type=int) # replay buffer size

parser.add\_argument('--batch\_size', default=64, type=int) # mini batch size

parser.add\_argument('--episode\_length', default=20000, type=int) # 回合长度

parser.add\_argument('--save\_interval', default=1, type=int) # 相隔n回合存储一次网络参数

parser.add\_argument('--max\_episode', default=201, type=int) # 回合数

parser.add\_argument('--update\_iteration', default=300, type=int) # 每回合更新网络参数的次数

parser.add\_argument('--exploration\_noise', default=0.1, type=float) #探索噪声初值

parser.add\_argument('--state\_scale\_factor', default=1.0, type=float) #状态放大系数（放大之后输入网络）

parser.add\_argument('--action\_scale\_factor', default=2.0, type=float) #动作放大系数（网络输出后放大）

args = parser.parse\_args()

（2）网络结构

class Actor(nn.Module):

class Critic(nn.Module):

（3）奖励函数

self.reward = -(self.y1)\*\*2 -(self.y2)\*\*2 -(self.y3)\*\*2 -(self.y4)\*\*2

（4）探索与利用

args.exploration\_noise \*= 0.99 #每回合噪声衰减

如图1，每回合结束可以得到一张图片。第1排是0-3号传感器该回合的时间序列，红色是没有智能体时的信号，蓝色是加入智能体后的信号。如果智能体训练效果好，蓝色信号会比红色信号小很多。第2排是第1排的FFT变换。第3排是该回合智能体的4个输出信号。第4排是各回合的回报。

现阶段存在的问题是，网络容易陷入局部最优解，训练结果不佳。希望通过修改超参数和网络结构达到较好的训练结果。

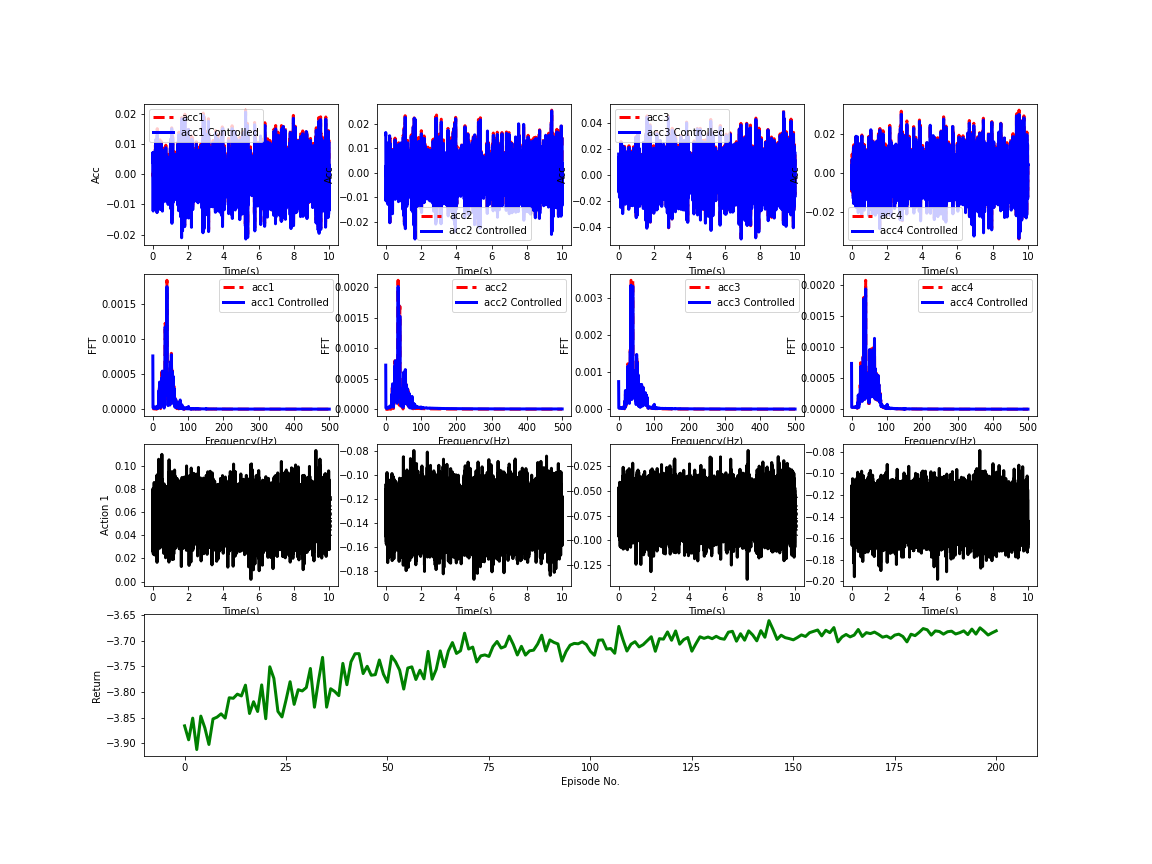


图1

用图2来说明较好的训练结果的含义。下图2是针对单根线谱的智能体训练结果，蓝色信号明显比红色信号小。如果图1可以达到类似图2效果，则说明智能体有一定的效果。

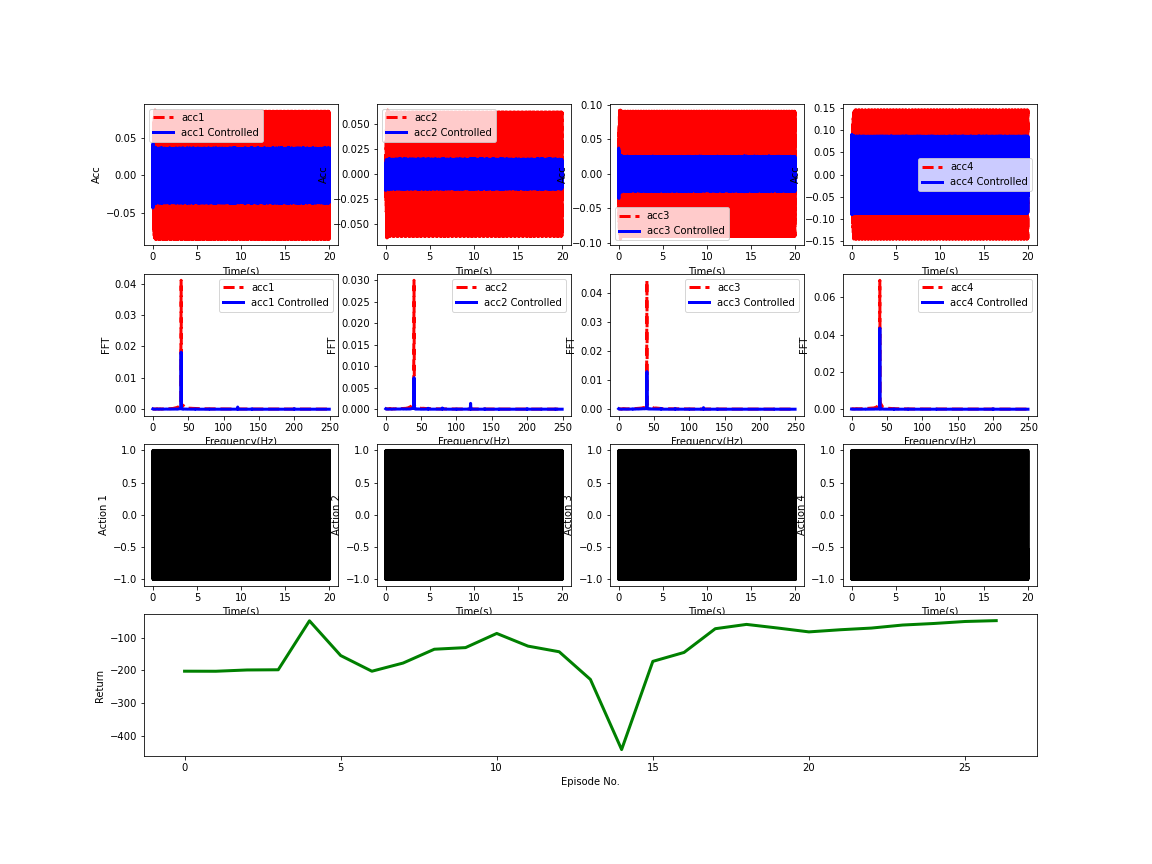


图2