

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能システム制御 (杉本 謙二 (教授))		
学籍番号	2011166	提出日	令和 4年 1月 20日
学生氏名	武田 悠佑		
論文題目	最適デモンストレーションを陰的に選択する模倣学習		
要旨			
<p>近年、ロボットの利用が増加しているが、ロボットに要求されるタスクは多様化し、かつ一つ一つのタスクも複雑化が進んでいる。そのため、設計段階で全ての状態と対応する行動を考えることは不可能となってきた。ロボットの利用拡大に合わせて、深層学習などの機械学習技術が発達しており、入力されたデータに対応する適切な解を出力できるモデルの構築が容易になっている。この深層学習技術をロボットに応用した事例が増加しており、特に熟練者の方策を模倣することで事前設計を必要とすることなくタスクを実現できる模倣学習が注目されている。模倣学習の代表的な手法はBehavioral Cloningであり、熟練者が直面したある状態の時に取った行動の組を収集し、模倣者は状態を入力、対応する行動を出力とした回帰問題を解くことで熟練者の方策を獲得できる。教師あり学習であるため、模倣者は不完全な方策で環境と相互作用することなく、熟練者のデータセットから最適な方策を学習可能である。しかしこの手法は、データセットに多種類の軌道が含まれている場合に学習が失敗することが知られている。</p> <p>本研究では、不完全な方策である模倣者が環境との相互作用なしで、多種類の軌道の中から単一の軌道の陰的に選択する模倣学習手法を提案する。具体的には、状態遷移情報から熟練者がおこなった行動を推論するモデルを追加し、軌道を分類する。この分類された軌道から尤もらしいものを陰的に選択するよう、Behavioral Cloningの最小化問題を変更する。提案手法の挙動を定性的に解析し、シミュレーション環境と実機環境で行い、どちらの環境でも最適な一つの軌道を選択して模倣することを確認した。</p>			