

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	1911236	提出日	令和 3年 1月 25日
学生氏名	藤石 秀仁		
論文題目	Safe and Efficient Imitation Learning by Utilizing Safety Area Suggested by Expert 熟練者が示唆する安全領域を活用した安全かつ高効率な模倣学習		
要旨			
近年、深層学習などの機械学習技術が発展し、入力に対して適切な出力を行うモデルの構築が可能となり、その技術をロボットに応用している事例が増えている。特に、熟練者の振る舞いをロボットに模倣させることで方策を学習させる模倣学習は環境との作用を必要としない、または比較的少ない環境との作用数で方策を学習することができるため効率的かつ安全な手法として注目されている。模倣学習は一般的に、環境情報である状態とその状態に応じて熟練者が取る行動のペアを取得し、状態を入力として受け取り行動を出力する方策を学習させることで熟練者の模倣を行うが、熟練者の行動情報は実際には筋電位などの内部信号のようにセンサによる計測が必要であることが多く、行動情報を容易に計測できる問題は限られている。そこで従来研究では熟練者の状態遷移時に行われた行動を推定する逆動力学モデルを構築することで熟練者の行動情報がなくとも模倣に成功している。しかしながら、逆動力学モデルを構築するには少なからず環境との相互作用を必要とするため、ロボットや実環境が故障する危険性がある。本研究では、熟練者の行動情報が計測できないという条件下で、ロボットが環境と相互作用する際に起こりうる危険な状態への遷移を防ぐ機構を統合し、安全かつ高効率な探索を可能とする模倣学習の実現を目指す。具体的には、熟練者は危険な状態へ遷移しないことに注目し、ロボットの直面している状態が熟練者が未経験なものかどうかを判定する外れ値検出器を構築する。これにあたって、訓練データが潜在空間における特定の事前分布より生成されたものとみなす変分オートエンコーダを用いて、事前分布に対する尤度に基づき外れ値を判別可能な外れ値検出器を構築した。これにより、安全な領域のみの探索を行うことで安全性を保証し、かつ熟練者の経験した状態空間を重点的に探索することができる。提案手法ではシミュレーション、実機実験において危険な状態への遷移を未然に防ぎ、安全な探索とともに方策を学習できることを確認した。また、熟練者が経験した状態のみに注目して学習を進めることで、従来手法と比較してより少ない環境との作用数で効率的に学習することができ、かつ模倣精度も向上することを確認した。			