

(続紙 1)

京都大学	博士（情報学）	氏名	山森 聡
論文題目	Variational Inference-based Skill Learning for Robot Control （ロボット制御のための変分推論に基づくスキル学習）		
（論文内容の要旨）			
多自由度ロボットで人間のような運動動作を再現することは、ロボット制御分野の主要な目標である。本学位論文では、多様な状況に適応しつつ、ロボットが効率的に運動スキルを学習できる新たな枠組みを提案している。近年の計算機工学の進展もあり、多自由度人型ロボットの制御器設計において深層学習が実用的かつ有望なアプローチとなりつつある。しかし、深層学習には膨大な数のパラメータと大規模な学習データ、そして複雑な学習プロセスが必要である。さらに、訓練環境と実際の状況との不一致などによって、学習済みの制御方策を実際に適用する際に困難が生じる。新しい状況に対して方策を再学習するには膨大なデータが必要だと考えられている一方で、人間はわずかな試行回数で実世界に適応できる。人間のような適応性を実現するには、ロボットには単なる制御パラメータの調整を超えて動作を修正させるための特別な表現が必要となる。 上記の背景・動機の下、本論文では、1章において、ロボットにおける適応能力を有する運動スキル獲得の背景と課題について論じ、その解法の基礎となる変分推論の問題設定を示した。また、2章では、本論文全体で用いるKL最小化問題に基づく変分推論の枠組みを導入した。さらに、3章では、リズムミクな人間の運動スキルが有する力学的特性を抽出し、人間からロボットへの運動スキル移転を可能にする位相振幅縮約表現を活用した模倣学習法を提案した。提案法は、位相振幅縮約に基づくリミットサイクルの位相と振幅表現を、変分推論によるエンコーダ・デコーダ学習によって獲得する。さらに獲得した位相振幅縮約表現を利用した多自由度ロボットの軌道追従制御も示した。具体的には、6軸腕ロボットによる人からロボットへの動作転移の検証によって提案法の有効性を検証した。つづいて、4章では、少数の試行回数による新たなタスク適応を促進するためのマルチタスク学習による運動スキル生成の手法を提案した。ロボットが未知の状況に対する適応性を獲得するためには、方策を状況に即して変化させる潜在的な表現が必要である。この章では、変分推論による表現学習とマルチタスク強化学習を統合した枠組みにより、このような潜在表現を獲得し、さらにベイズ最適化による未知タスクへの適応手法について提案した。加えて、5章では、マルチエージェント環境を通じた経験から多様で柔軟なスキル表現を獲得する手法を提案した。複数のエージェントによる経験を、マルチタスク強化学習に分解して獲得することによって、様々な状況に対処できるスキル表現を獲得できる。また、同章では獲得したスキル表現を競争的マルチエージェントタスクに適応する手法についても提案し、提案法により獲得された運動スキルの柔軟性について評価した。最後に6章では、本論文のまとめ、運動スキル獲得によるロボット制御の今後の展望について述べている。本論文を通じて、変分推論をKL最小化問題と			

して定式化することにより、模倣学習・強化学習の双方を統合した一貫性のある運動スキル表現の獲得とそのスキル表現を活用した学習型制御を実現した。

(論文審査の結果の要旨)

本論文では、変分推論の枠組みを応用した動作スキルの表現学習と方策学習の手法を以下のように提案した。1) 位相振幅縮約による模倣学習：人間の運動スキルの力学的性質に注目した上で、ロボットによる人のスキルの模倣学習の手法を提案、2) マルチタスク学習によるスキル生成：マルチタスク学習に基づくことで、少量の追加試行で新しいタスクへ適合可能な運動スキル生成手法を提案、3) マルチエージェント学習によるスキル生成：マルチエージェントによる協調・競合学習に基づくことで、多様なスキルの生成を可能とする手法を提案。表現学習と模倣学習および強化学習（方策学習）を変分推論問題として定式化することで、一貫性のある実装を可能にし、多自由度ロボットの実システム、あるいは、シミュレーションに基づく評価を通して、各手法が有効に機能することを示した。

1. 位相振幅縮約による模倣学習

変分推論に基づき、位相振幅縮約による人間の潜在力学モデルの同定とそのモデルを利用した相互フィードバックの方法を提案した。人間が生成した動作をロボットが追従するという動作プリミティブを用いた模倣学習では、リミットサイクル上の定常的な軌道の再現に限定され、過渡的な動作軌道の模倣が容易でないという問題点があった。提案法における潜在力学系モデルでは位相と振幅を表す潜在空間を有し、それぞれ定常的な成分と過渡的な成分に分解して表現するため、過渡的な動作スキルの軌道について良い模倣学習が可能となる。ロボットマニピュレータを用いた評価では、標準的に用いられる動作プリミティブ法と比較して、過渡的な動作軌道も含め人間の運動スキルへの追従性能が改善されることが示された。

2. マルチタスク学習による運動スキル生成

変分推論に基づくマルチタスク強化学習の手法を提案し、それに基づくことで新しい運動スキルへの汎化を可能とする方策の構成に成功した。人間は既知の運動スキルの汎化により新しい運動スキルの獲得が可能であるが、そのためには、多彩なタスクに対応した運動スキル生成の方策のセットの学習と、新しいタスクに対応する方策を学習済の方策セットから迅速に構成する機能が必要になる。提案法によれば、変分推論に基づき、タスクの違いを表現する潜在空間とそれを用いた方策のセットを同時に学習する。また、新しいタスクに対応する際には、ベイズ最適化による潜在変数の同定により、学習済の方策セットを汎化させることができる。一脚ロボットモデルを用いた評価では、獲得された潜在空間を利用したスキル適応により、方策セットの学習時とは異なる、新しい運動スキルを遂行可能な方策が構成可能であり、さらに、関連手法であるドメイン適応法よりも高効率であることが示された。

3. マルチエージェント環境を利用した人間のような多様なスキル生成

2で開発したマルチタスク強化学習手法をマルチエージェント環境に導入し、エージェント間の協調あるいは競合により多様な経験を生成、潜在空間で統合することでスキル学習に用いるとした学習方式を提案した。まず、複数の協力タスクを通じた経験から潜在的スキル表現を学習し、その潜在空間を用いて新しい運動スキルに対しても汎化を可能とする。この枠組みにより複数のエージェントはスキルの共有による不確定性の低減を図ると同時に、獲得した複数の動作を素早く生成することができた。2体の一脚ロボットモデルによる協力動作継続タスクを用いた評価によって、提案法はシングルエージェント学習を単純に組み合わせる場合に比べて協力動作生成の性能を改善することが示された。また、同様の手法を競合タスク環境に

適用することで、新しい運動スキルの創発が可能となることを示唆した。

以上、本論文では、変分推論の枠組みを応用した動作スキルの表現学習と方策学習の手法をおおきく三つの課題を通じて提案した。第一に、人間の運動スキルの力学的性質に注目した人からロボットへの模倣学習手法を提案、第二に、マルチタスク学習に基づくことで、少量の追加試行で新しいタスクへ適合可能な運動スキル生成手法を提案、第三に、マルチエージェントによる協調・競合学習に基づくことで多様なスキルの生成を可能とする手法を提案した。表現学習と模倣学習および強化学習（方策学習）を変分推論問題として定式化することで、一貫性のある実装を可能にした。本論文での成果は、ロボットが柔軟で適応的な運動スキルを効率よく獲得し、複雑な環境やタスクに対応する能力を向上させるためのアルゴリズムの創出につながるものであり、機械学習、および知能ロボティクス分野での幅広い貢献が期待できる。

よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、令和7年2月19日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。なお、本論文は、京都大学学位規程第14条第2項に該当するものと判断し、公表に際しては、（令和8年3月24日までの間）当該論文の全文に代えてその内容を要約したものとすることを認める。

要旨公開可能日：令和7年6月24日以降