

論文内容の要旨

博士論文題目

Gaussian Process Policy Search with Latent Variables in Uncertain Environments

(不確実な環境下における潜在変数を持つガウス過程方策探索)

氏名 佐々木 光

(論文内容の要旨)

ロボットの制御方策を自律的な試行錯誤から学習する方法として、方策探索型の強化学習が提案されている。特に、ガウス過程回帰を方策モデルとして用いたガウス過程方策探索は、高次元で冗長なセンサー値を入力としたロボットの制御方策が学習可能である。一方で、ガウス過程方策探索の実世界タスクへの応用は実世界タスクの環境が持つさまざまな不確実性によって、依然として困難である。不確実性の高い環境下において性能の高い方策を獲得するためには、方策が状態から行動への複雑な関数関係を捉える必要がある。しかし、シンプルな方策モデルを仮定してガウス過程方策探索は複雑な関数を学習できない。このような方策の学習に対する問題を克服するために、教師あり/教師なし学習における潜在変数モデルに着目する。潜在変数モデルは、観測可能な変数と観測できない潜在変数を関連付ける確率モデルで、潜在変数を適切に設計することで、アルゴリズムのシンプルさを維持しながら複雑なデータを捉えるモデルを学習することができる。この博士論文は、不確実な実世界環境で引き起こされる方策学習の困難さに対して、ガウス過程方策探索に潜在変数を取り入れるアプローチを探索する。不確実性をもつ実世界タスクへの方策探索を応用することを目的とし、潜在変数の推論と方策の学習を同時に行うアルゴリズムを導出する。特に、実世界タスクが持つ、1) 曖昧に設計された報酬関数によって生じる複数の最適行動と 2) 環境の情報が十分に得られない観測の2つの複雑さ、3) タスク性能の評価に対する予測できない影響に焦点を当てる。1) と 2) に対して、潜在変数をガウス過程方策モデルに導入することにより、実世界タスクが持つ複雑さに対する新たなガウス過程方策モデルを設計し、変分ベイズ学習に基づいて潜在変数付き方策の更新則を導出する。提案された方策探索手法の性能をシミュレーションとロボットマニピュレータを使用したタスクによって

氏名	佐々木 光
----	-------

(論文審査結果の要旨)

本論文では、ガウス過程回帰を方策モデルとして用いる強化学習手法であるガウス過程方策探索に関するアルゴリズム開発と、実世界ロボットタスクへの応用に関する研究を扱っている。ガウス過程方策探索は、そのノンパラメトリックな性質により、高次元で冗長なセンサデータを入力とする方策の学習に適している。一方で、実世界タスクへの応用については、環境が持つ様々な不確実性の影響によりあまり進んでいない。不確実性の高い環境下において性能の高い方策を獲得するためには、状態から行動への複雑な関数関係を捉える能力が方策モデルに要求される。しかし、ガウス過程方策探索では状態から行動への関数としてガウス分布を仮定しており、その要求に応えることができない。

本論文では、そのような制限を緩和するため潜在変数モデルに着目した。潜在変数モデルは、観測可能な変数と観測できない潜在変数を関連付ける確率モデルで、潜在変数を適切に設計することで、アルゴリズムのシンプルさを維持しながら複雑なデータを捉える能力を実現できる。よって、ガウス過程方策探索において適切な潜在変数を方策モデルに導入し、実世界タスクの不確実性の問題に対処するアプローチを探求している。具体的には、実世界タスクが持つ 1) 曖昧に設計された報酬関数によって生じる最適行動の多峰性、2) 環境の情報が十分に得られない限定的な観測、3) タスク評価に対する予測における不確実性、の3つの課題に焦点を当てた。1) と 2) については、潜在変数をガウス過程方策モデルに導入することで、実世界タスクが持つ複雑さに対処可能なガウス過程方策モデルを設計し、変分ベイズ学習に基づいて潜在変数付き方策の更新則を導出した。提案手法の性能をシミュレーションとロボットマニピュレータを使用したタスクによって検証した。3) については、潜在変数を導入したベイズ最適化に基づく方策学習フレームワークを提案し、ゴミ焼却施設のゴミクレーンタスクへの有効性を検証した。

これらの結果から、本論文は潜在変数を導入した柔軟で応用範囲の広いガウス過程方策探索アルゴリズムの提案、およびゴミクレーンのような実世界ロボットタスクへの稀少な適用事例として、新規性および有効性の観点から一定の学術的意義があるものと評価できる。よって、本論文は博士(工学)の学位論文として価値あるものと認める。