

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 3 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26730136

研究課題名（和文）多様なセンサを用いた実世界でのロボットの動作学習

研究課題名（英文）Development of motion learning method for a robot with various sensors in a real environment.

研究代表者

中村 泰（Nakamura, Yutaka）

大阪大学・基礎工学研究科・招へい准教授

研究者番号：70403334

交付決定額（研究期間全体）：（直接経費） 2,900,000円

研究成果の概要（和文）：本研究課題では環境についての知識を前提とすることができない実世界で活動するロボットの開発を目的とする。実世界では多様な身体的インタラクションが避けられないため、様々な情報を統合し多様な状況に対応することが必要になる。そこで本研究では多数のセンサモダリティを利用したロボットの運動学習に取り組んだ。初めに制御入力として必要なセンサを自動的に選択する強化学習法、および、多自由度ロボットを制御を目的としたモデルの信頼度を用いたサンプリングベース動作学習法を開発した。そして、実環境で人とインタラクションを行うロボットの開発を行った。

研究成果の概要（英文）：In order for a robot to operate in a real environment, it is inevitable to physically interact with diverse objects including human. Such a robot would face to diverse situation, and it is necessary to integrate various sensory input and take appropriate action based on the observation. For this aim, I developed a reinforcement learning method for a robot where the required sensory input is extracted from multi-modal sensory input and a sampling based roadmap method which considers the confidence of the modelling. I also developed a robot socially interact with the human in a real environment.

研究分野：知能ロボティクス

キーワード：知能ロボット 多様なセンサ 身体的インタラクション

1. 研究開始当初の背景

(1) 実世界で活動するためには多様な状況に直面するだけでなく、人や物体との物理的なインタラクションを行う必要がある。そのため、様々なセンサを用いて状況を観測し、その観測に応じて適切な行動を選択する必要がある。すなわち、その制御則は多様な入力から適切な行動を決めるものとなるため、複雑なものになってしまう。通常、用いられるニューラルネットワークなどの汎用的なパラメトリックモデルにおいても、設計者の知識に基づいて構造を決めるなどの工夫が大きな影響を与えるため、設計者の知見なしに学習を行うことは容易ではない。

(2) 近年、学習データセット自体から直接出力を計算するノンパラメトリックモデルが注目されており、複雑なダイナミクスの推定などの課題で成果を上げている[1]。ノンパラメトリックモデルは、与えられたデータをモデル化するための手法として非常に強力であるが、適切な入出力関係を自律的に学習することが必要な適応的な制御においては、直接応用することができない。

2. 研究の目的

(1) 本研究の目的は、制御対象のモデル化において設計者の知見を用いずに実世界で活動するロボットの複雑な制御則を学習する手法を研究することである。また、このような手法を応用することで、実世界で人間との物理的なインタラクションを行うことができるロボットの開発を目指す。

(2) そこで、近年注目されているベイジアンノンパラメトリックモデルを用いた制御法の開発を行う。ノンパラメトリックモデルは通常、与えられたデータセットを直接利用して入出力関係などのモデル化を行う。本研究課題では、ロボットが経験により得たそれぞれのサンプルに対して有用性を学習したり、サンプル間の関連性を推定するなど、与えられたデータセット自体を操作することで、複雑な環境においても適用可能なロボットの運動学習手法の実現を目指す。

3. 研究の方法

(1) 本研究課題では、ノンパラメトリック法として、ガウス過程回帰を用いる。ガウス過程回帰は、ベイジアンモデルの一種であるため、期待値の推定とその推定に対する信頼度(コンフィデンスインターバル)を計算することができる。また、この信頼度はデータセットに含まれる類似のサンプルが少ないときに低く(コンフィデンスインターバルが大きく)、多いときに高く(小さく)なるという性質を持つ。

(2) そこで、本研究では、サンプルの重要度を変化させることでこの推定値を操作し

たり、サンプル間のつながりを信頼度に基づいて計算するなど、データセットを直接操作することで学習を行うシステムを開発する。

(3) 一方で、通常、ノンパラメトリックモデルはサンプルサイズの増加とともに、計算量が大きくなる。すなわち、実応用を視野にしている場合、現実的な計算時間で計算が可能な手法を用いる必要がある。よって、実用的な高速化手法を用いた制御手法の開発も重要である。

4. 研究成果

(1) 本研究課題では、データセット内のサンプルの重要度や各センサモダリティに対する重要度を方策パラメータとして強化学習を行うことで、重要な情報だけをデータセットから抽出し、適応的な制御を行う手法を開発した。

(2) 具体的には、サンプル自体に重みを与え、得られる報酬の大きさによりサンプルを評価することで、いわば、模倣すべきサンプルを抽出することができる。また、多数のセンサモダリティがある場合には、各サンプルの次元がそのモダリティの数に比例して増えるが、各モダリティに重みを与えることで、どのモダリティを用いてサンプルの類似度を計算するかを学習によって変更することで、適切なセンサモダリティを選択できることを確認した(雑誌論文)。

(3) この成果は、与えられたデータセットのどのような情報を用いて制御を行うかを学習するメカニズムを提案したものとなっている。

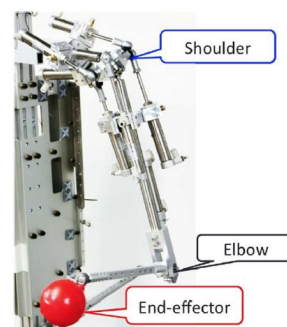


図 1: 多自由度ロボット

(4) さらに、多自由度ロボットを制御するためのサンプリングベース動作生成法の開発に取り組んだ。サンプリングベースの動作生成法は、自動運転技術などで近年盛んに研究されている手法である[2]。通常はモデルが与えられている条件の下で、多数の経路(サンプル)を生成し、最適な動作を探索する手法であるが、図1に示す多自由度のロボットのモデル化は、ロボット単体であっても

容易ではないだけでなく、実環境においては周囲からの影響により動作が変化してしまうため、前もってモデルを与えることは困難である。

(5) さらに、ロボット自身が動作することで収集したデータに基づきモデルを学習する場合においては、十分なサンプルが集まるとは限らないため、そのモデルの信頼性が問題となる。本研究課題では、ガウス過程回帰を用いることでこの問題に取り組んだ。具体的には、ガウス過程回帰モデルでフォワード・インバースダイナミクスのベイズモデルにより推定を行うことで、サンプル間の接続を評価し、信頼度の高い経路を生成する手法を開発した(雑誌論文, , および学会発表)。

(6) 通常はランダムに生成した多数のサンプルにより運動計画を行うが、本手法ではデータセット内のサンプルを繋ぐ経路に対する信頼度に基づく評価を行うこと、可能なハッシュ関数も用いたガウス過程回帰モデルの高速な計算法を用いることで、現実的な時間でロードマップを構成し、実ロボットが制御可能であることを示した。

(7) 多自由度のロボットに対するモデルを名に与えることなく実現しただけではなく、この手法においては信頼度に対する重み付けを変えることでロボットの振る舞いを変えることができることがわかった。図2は制御課題でよく用いられる倒立振り子モデルの制御に応用した場合に生成された経路を示している。縦軸は角速度で横軸は角度を示しており、図中央が倒立状態、左右の端がぶら下がり状態を示している。左図は信頼度にして敏感な基準で動作経路を生成したもので、右図は鈍感にしたものである。図より、信頼度に対して敏感な基準で動作を生成する場合は、再現できる確実性の高い経路を選択するために、比較的長いステップ数で動作する経路が選ばれやすくなった。これは、強化学習における探索搾取問題と同様に、試したことのない経路を試すか、確実に達成できる経路を使うかのバランスが重要になることを示唆している。また、信頼度に鈍感な経路を用いた場合は、計画した経路と実現された経路との間の差が大きくなった。

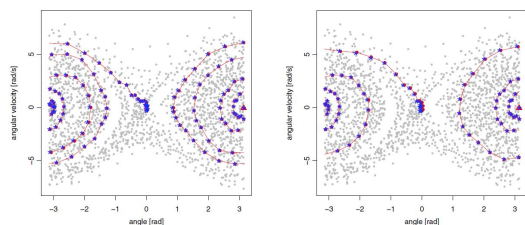


図2: 動作経路

(8) さらに、ロボットと環境中の物体との物理的なインタラクションの推定課題にも取り組んだ。人を含む様々な物体との接触が考えられる実世界での動作生成において、接触によってどのような力が加わるかを推定することは重要になる。我々の提案したハッシュ関数を用いた高速なガウス過程回帰モデルが、サポートベクトル回帰モデルやスパースガウス過程回帰に比べて、精度においても計算速度においても十分な性能を持つことを示した(学会発表)。

(9) 本研究課題を開始後、特に画像処理の分野で深層学習が著しい発展を遂げるとともに、画像を入力とするロボット制御に対する強化学習法が提案されてきた[3]。画像入力の特徴は、センサ情報としては画素値単位であるため、一つのモダリティであっても非常に高い次元を持っていることである。深層学習の強力な特徴抽出メカニズムにより、人手によるチューニングの必要性が少ないことで、様々な問題に応用されてきている。



図3: 実環境での HRI

(10) そこで、本研究課題の発展として、実空間における人とのインタラクション(Human-robot interaction: HRI)の学習課題に取り組んだ。図3に示すように、実環境では多様な状況が発生するため、ロボットの行動ルールを状況ごとに人手で設計する手法を用いることは困難である。本研究では、ロボットが適切に動作することで人との社会的コミュニケーションである握手を行うための方策の学習を行った。ロボットによる14日間の自律的な学習の結果、人間の評価者による行動選択と同様の行動選択則を獲得することができた(学会発表)。

(11) 実環境における HRI の学習においては、ノンパラメトリック法は用いていないが、深層学習の一つのメリットは高性能な特徴抽出技術であると考えられる。パラメトリックモデルを用いる場合、特徴の計算方法が変化すると、それに依存する回帰モデルなどのパラメータもそれに依りて再学習を行う必要性が出てくる。一方で、得られた特徴空間を使いノンパラメトリック法で学習を行う場合、サンプルの重みをそのまま再利用するなど、有効に使える可能性を持つ。すなわち、

これらの知見と本研究課題の成果であるノンパラメトリックモデルを利用した制御法を組み合わせることは今後の課題である。

<引用文献>

[1] NIPS2012 Workshop: Bayesian Nonparametric Models For Reliable Planning And Decision-Making Under Uncertainty. <http://acl.mit.edu/bnpm>

[2] Kuwata, Y., et al. Real-time motion planning with applications to autonomous urban driving. IEEE Transactions on Control Systems Technology, 17(5), 1105-1118. (2009)

[3] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529-533, 2015.

5. 主な発表論文等

〔雑誌論文〕(計 3 件)

Yuya Okadome, Yutaka Nakamura, and Hiroshi Ishiguro, "A confidence-based roadmap using Gaussian process regression," Autonomous Robots, 2016. (DOI: 10.1007/s10514-016-9604-y) <https://link.springer.com/article/10.1007/s10514-016-9604-y>

岡留 有哉, 中村 泰, 石黒 浩, "高速なガウス過程回帰を用いた予測モデルに基づくサンプリングベース動作計画", 電気学会論文誌 C, No. 135, Vol. 5, pp.1-8, 2015.

Jumpei Yamanaka, Yutaka Nakamura, and Hiroshi Ishiguro. A feature selection method for a sample-based stochastic policy. Artificial Life and Robotics, 19(3), pp.251--257. (2014) <https://link.springer.com/article/10.1007/s10015-014-0158-9>

〔学会発表〕(計 3 件)

Ahmed Qureshi, Yutaka Nakamura, Yuichiro Yoshikawa, Hiroshi Ishiguro, "Robot Gains Social Intelligence through Multimodal Deep Reinforcement Learning", 2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids), pp.745--751 (2016)

Yuya Okadome, Yutaka Nakamura, Kenji Urai, Yoshihiro Nakata, and Hiroshi Ishiguro, "HUMA: A Human-like

Musculoskeletal Robot Platform for Physical Interaction Studies", IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS 2015), Seoul, Korea, pp. 676-683, November, 2015.

ガウス過程回帰の信頼度に基づく確率的ロードマップを用いた動作計画. 岡留 有哉, 中村 泰, 石黒 浩. 人工知能学会全国大会 2014 (JSAI2014), 4J1-1, (2014)

〔その他〕
ホームページ等
<http://www.irl.sys.es.osaka-u.ac.jp/home/research/huma>

6. 研究組織

(1) 研究代表者

中村 泰 (NAKAMURA, Yutaka)
大阪大学・基礎工学研究科・招へい准教授
研究者番号: 70403334

(2) 研究協力者

岡留 有哉 (OKADOME, Yuya)