

倒立二輪型移動ロボットによる人間の動的動作模倣のための身体部位マッピング

Inverted-Pendulum Mobile Robot Motion Imitating from Human Player Observation and Body Parts Mapping for Imitation

高橋 佐多弥 高橋泰岳 前田陽一郎*¹
Sataya TAKAHASHI Yasutake TAKAHASHI Yoichiro Maeda

中村恭之*²
Takayuki NAKAMURA

*¹福井大学工学部
Faculty of Engineering, University of Fukui

*²和歌山大学システム工学部
Faculty of Systems Engineering, Wakayama University

This paper addresses development of body mapping from a human demonstrator to an inverted-pendulum mobile robot for imitation. An inverted-pendulum mobile robot with torso and body links learns a dynamic kicking motion shown by a human. The robot observes the human demonstration with a camera, extracts the human region in each of images, maps the region to its own two links, estimates the link posture trajectories, and starts kicking motion learning based on the trajectory parameters for imitation. The mapping parameter gives an important role for successive imitation. A reasonable and feasible procedure of learning from observation for an inverted-pendulum robot and development of the body mapping is proposed and investigated in this paper.

1. はじめに

近年、人間共生環境における自律型移動ロボットの研究が盛んに行われている。ロボットが人間と共存していく上で、ロボットが自ら環境に適応していくことが求められるが、我々の生活する環境は常に変動し、それらのデータを事前に設計することは非常に困難である。そこで強化学習に代表される機械学習の手法を用い、試行錯誤を通してロボットに行動学習させる研究[浅田 95]が行われてきた。試行錯誤による学習コストは特に実ロボットの場合大きいため、良い学習初期値を得ることが重要であり、そのため人間や他のロボットが呈示者となり、その動作を模倣する模倣学習が近年盛んに研究[谷口 09][中岡 03]されている。しかし人間の動きを模倣しようとしても、リンク構造や身体ダイナミクスが大幅に違う可能性があるため、観測情報から得られた軌道をそのまま追従し模倣を行うことは困難である。観測データを模倣学習する際の初期値として利用し、試行錯誤を通じた行動学習によって自身に適した動作を獲得し、模倣を実現するのが妥当だと考えられる。

そこで本稿では、低自由度で比較的素早い動作が可能な倒立二輪移動ロボットに、人間の蹴球動作をカメラで観察させ、その観測データを初期値として行動学習を通して観察した蹴球動作を模倣させる。そして観測データを適切にロボットのリンクにマッピングすることでより良い模倣学習を行えることを示し、さらにロボットの模倣学習にとって最適となるように、マッピングパラメータを学習させる手法を提案する。

2. 2リンク倒立二輪型移動ロボットによる模倣学習

模倣学習では、身体構造が異なる呈示者の身体から、模倣するロボットの身体へのマッピングが重要である。人間は身体へのマッピングを適切に行い、リンク構造が異なる動物などが行う動作を模倣することができる。本研究では模倣によって獲得される動作の評価が良くなるように呈示者の身体を自身の身体にマッピングすることによって、この問題を解決する一手法を提案する。図1に模倣学習の流れを示す。

連絡先: 高橋泰岳, 福井大学大学院工学研究科, 910-8507 福井県福井市文京 3-9-1, yasutake@ir.his.u-fukui.ac.jp

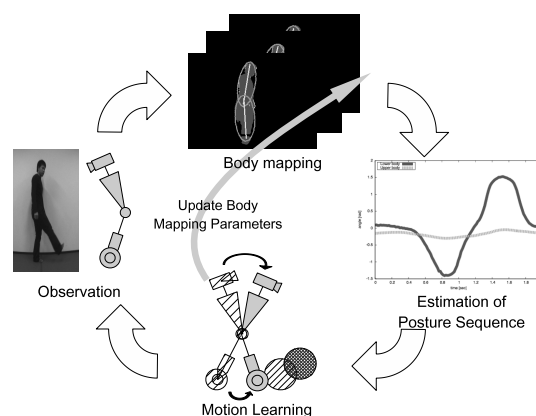


図1: Imitation Learning Flow

2.1 蹴球動作

本研究ではロボットに人間の蹴球動作を模倣学習させる。蹴球動作は2つのモーションで構成されていると仮定し、ロボットは図2のように蹴球動作を行う。蹴球動作を規定するパラメータは、モーションの動作継続時間 t 、上半身の目標姿勢角度 θ_d^u , θ_d^l 、車輪の目標角速度 $\dot{\phi}_d$ 、上半身の姿勢角度誤差、角速度、車輪の角速度、角速度誤差、上半身の積算角度誤差それぞれのゲイン k_1^u , k_1^l , k_2^u , k_2^l , k_3^u , k_4^u , k_{t3}^u の11個で、2つモーションを定義したため総計22個である。

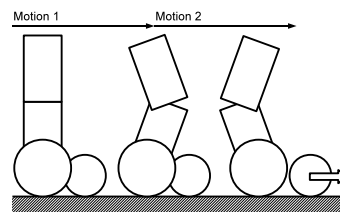


図2: Two Primitive Motions to Kick a Ball

2.2 動作認識

動作認識では、モーションキャプチャや色抽出のためのマーカーを使用せず、背景差分法とフレーム間差分法を用いて取得し

た画像中から呈示者領域を抽出する。抽出した人間の写っている領域を本研究で使用する倒立二輪移動ロボットの各リンクに対応付けさせる。さらにマッピングパラメータを更新することで各リンクの領域が変化し画像モーメントから求まる姿勢角が異なる。ここでマッピングパラメータは人間の写っている領域全体を水平線で上下半身に分けるときの上半身の割合とする。

2.3 行動学習

節 2.1 で定義した蹴球動作を規定するパラメータを方策勾配法を用いて学習する。Kohl and Stone によって定式化されたアルゴリズム [Kohl 04] を利用する。方策勾配法とは、動作パラメータを勾配法により繰り返し修正していくことで局所最適解を求める学習法である。今回使用した評価関数を式 (1) に示す。蹴球動作の評価は、ロボットが初期位置からあまり移動せず、ボールを強く蹴りだせる方が良いとする。また、転倒した場合は模倣に失敗したと判断し、ペナルティを与え評価を低くしている。

$$E = \begin{cases} w_1 v_b - w_2 l_b^2 & \text{転倒しなかった場合} \\ w_1 v_b - w_2 l_b^2 - w_f & \text{転倒した場合} \end{cases} \quad (1)$$

このとき、 v_b はボールの速度、 l_b は積算移動距離、 w_1 、 w_2 は各評価の重みで、 w_f は転倒時のペナルティである。

2.4 マッピングパラメータの学習

前述した動作認識において人間の写っている領域をロボットの上下半身にマップし、得られた各リンクの姿勢角の軌道から蹴球動作模倣のための初期値を得る。マッピングパラメータによって模倣学習の初期値が異なり、方策勾配法で得られる学習結果が異なる。模倣にとってより良い動作を学習するため、マッピングパラメータを学習動作の性能に基づいて勾配法で更新する。

3. 実験

動作認識では実際にカメラで人間の蹴球動作の観察を行った。行動学習ではシミュレータを使用し行動の評価を行った。方策勾配法での勾配方向を推定するための試行回数は 100 回、学習パラメータ更新回数は 5000 回で行動学習を行った。各評価の重みは、 w_1 が 2.0、 w_2 が 1.0、 w_f が 10.0 とした。

3.1 観察による学習初期値の有効性

動作認識で得た初期値の有効性を示すため、人間が手で設定した学習初期値との比較を行った。図 3 に実験結果の平均評価値を示す。この結果より、更新回数が 2000 回付近で両方の学習曲線が緩やかになり、4000 回付近からあまり変化がないので、更新回数を 5000 回と定めたのは妥当だと考えられる。また動作認識で得た学習初期値から行動学習を行った方が、学習後の平均評価値が約 2 倍高く提案する動作認識での模倣が有効であることが示されている。

3.2 マッピングパラメータを変化させたときの学習行動の性能

マッピングパラメータを 0.1 から 0.9 まで変化させたときの学習結果を図 4 に示す。この結果より、マッピングパラメータが 0.6 のとき平均評価値が一番高く、そこを頂上として山なりになっている。これより、現マッピングパラメータ付近で、勾配法を用いて学習結果の評価の高い方へマッピングパラメータの更新を行い、除々に頂上の局所解にたどり着くことにより、ロボットにより適した動作を模倣できることがわかった。

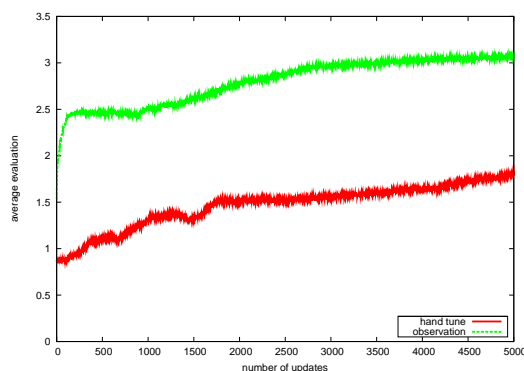


図 3: Motion Evaluation during Learning with Initial Motion Parameters from Observation and Tuned by Hand

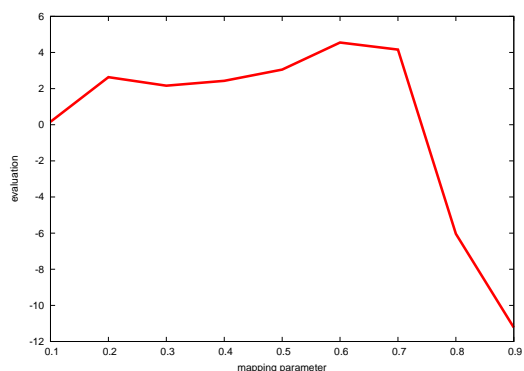


図 4: Acquired Motion Evaluation with Variety of Body Mapping Parameter Values

4. 結言

本稿では、動作呈示者の事前情報をさほど利用せずカメラから得られた画像データより呈示者の身体をロボット自身の身体にマップしリンク構造の異なる呈示者の動作を模倣する手法を提案した。実験よりマッピングパラメータを変化させたときに学習後の行動評価が異なり、勾配法を用いてマッピングパラメータを更新させることで、よりロボットに適した動作を獲得できることがわかった。

参考文献

- [Kohl 04] Kohl, N. and Stone, P.: Policy gradient reinforcement learning for fast quadrupedal locomotion, in *the IEEE International Conference on Robotics and Automation*, pp. 2619–2624 (2004)
- [浅田 95] 浅田 稔, 野田 彰一, 俵積田 健, 細田 耕: 視覚に基づく強化学習によるロボットの行動獲得, *日本ロボット学会誌*, Vol. 13, No. 1, pp. 68–74 (1995)
- [谷口 09] 谷口 忠大, 岩橋 直人: 複数予測モデル遷移の N-gram 統計に基づく非分節運動系列からの模倣学習手法, *日本知能情報フェジィ学会誌*, Vol. 21, No. 6, pp. 1143–1154 (2009)
- [中岡 03] 中岡 慎一郎, 中澤 篤志, 横井 一仁, 池内 克史: シンボリックな動作記述を用いた舞踊動作模倣ロボットの実現, *電子情報通信学会技術研究報告*, Vol. 103, No. 390, pp. 55–60 (2003)