

A2-2 倒立二輪型移動ロボットによる模倣学習のための身体部位マッピング

福井大学 工学部 知能システム工学科 進化ロボット研究室

木村 竜也 (指導教員 : 高橋泰岳、前田 陽一郎)

1. 緒言

人間との共生環境下で活動する自律移動ロボットを実現する上で、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、後で必要とされる全ての動作を事前に設計することは困難である。そこで強化学習を始めとする機械学習の手法を用い、試行錯誤の中でロボットが与えられたタスクに対する行動学習を行う研究 [1] が行われている。しかし、状態変数の数が増えるほど探索空間が増大し、特に実ロボットでの利用は難しくなる。学習時間を抑えるため良い学習初期値を得ることが重要であり、人間や他のロボットが提示者になり模倣学習させることが盛んに研究されている [2]。

ヒューマノイドロボットによる動作模倣が研究されているが、現在広く利用されているヒューマノイドロボットは多くのサーボモータを用いるためロボットの製造コストが高く、自由度が多いため制御をより複雑にしている。そこで本研究では低自由度で比較的素早い動作が可能な倒立二輪型移動ロボットを用いる。

倒立二輪型移動ロボットは動的安定性を実現した上で、地面への投影面積を小さく、人間と同等程度の大きさ、速さで動作できよりダイナミックな行動が可能である。人間とリンク構造や身体ダイナミクスが異なる倒立二輪型移動ロボットが模倣学習を行うには、人間の動きの観測データからロボットの各リンクに適切にマッピングし、マッピング後の各リンク姿勢の目標軌道を初期値として行動学習を行い、ロボット自身に適した動作を獲得する必要がある。そこで高橋ら [3] は簡便に模倣学習を行うために、カメラ動画像から背景差分、時間差分を用い領域抽出、各リンクのマッピング及び姿勢角の検出を行い提示者に負担をかけずに動作を比較的簡便でかつロバストな手法で動作認識をしている。認識した動作を適切にロボット自身のリンクにマッピングすることで目標姿勢角を推定し、行動学習の初期値として利用する手法を提案している。しかしこの手法では提示者の動作を一点方向から観察することを期待している。

本研究では観察する場所と方向の条件を緩和するために、マーカやセンサ類を装着せずに距離と画像を同時に取得できる Microsoft 社の XBOX360 Kinect センサを用い、提示者の関節座標を取得し、各関節リンクから目標姿勢角を推定する手法を提案する。各関節リンクの姿勢角が似ているものはクラスタリングを行い、倒立二輪型移動ロボットの模倣に適したマッピングを行う。ロボットのリンク数に合わせた目標姿勢角の初期値を取得し、行動学習後に評価する。

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提示者の身体から、模

倣するロボットの身体へのマッピングが重要である。提案する模倣学習の流れを図 1 に示す。まずロボットが提示者の動作を観察し 11 個の姿勢角の軌道を抽出する。姿勢角の軌道が似ているリンクはクラスタリングを行い、ロボット自身の各リンクに適したマッピングを行う。マッピングされたリンクから学習初期値の抽出し、学習を行いロボットに適した行動を取得する。

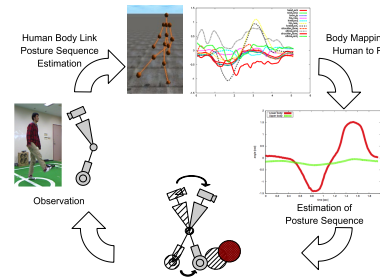


図 1: 模倣学習の流れ

2.1 動作認識と身体部位マッピング

Kinect センサを用い、人間の蹴球動作を認識する。人間の関節 16 点の座標を取得し、11 個のリンクの姿勢角を求める (図 2)。11 個のリンクの姿勢角の軌道を図 3 に示す。似ている姿勢角の軌道をクラスタリング



図 2: Kinect センサで取得する関節点とリンク

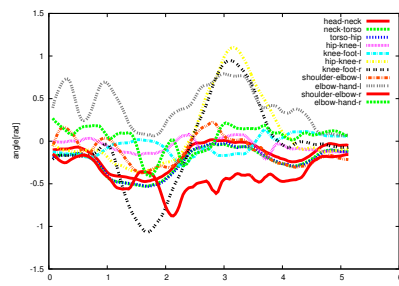


図 3: 11 個のリンクの姿勢角軌道

するため 2 つのリンク間の軌道の距離を求める。2 つのリンク間の軌道の距離を式 (1) で定義する。2 つのリンクが似ていれば、距離 d_{ij} は小さくなる。求めた値から閾値 T_e を決め、クラスタリングした例を図 4 に示す。クラスタリングしたリンクを倒立二輪型移動ロボット

を上半身，下半身，その他に分類する．

$$d_{ij} = \frac{1}{T} \sum_{t=1}^T \sqrt{(\theta_i(t) - \theta_j(t))^2} \quad (1)$$

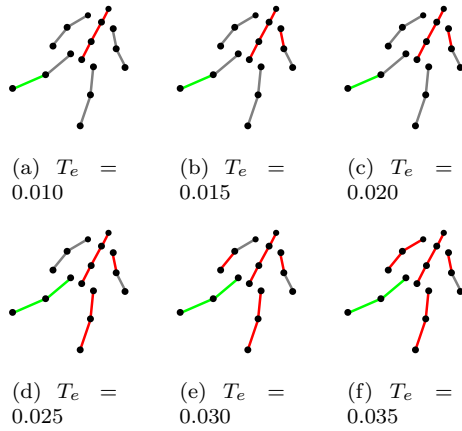


図 4: クラスタリング例：赤が上半身，緑が下半身，灰色がその他

2.2 蹴球動作の設計

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5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 ，上下半身の目標姿勢角度 θ_d^u, θ_d^l ，車輪の目標角速度 $\dot{\varphi}_d$ ，上下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 $k_1^u, k_1^l, k_2^u, k_2^l, k_3^w, k_4^w, k_{t3}^u$ 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．

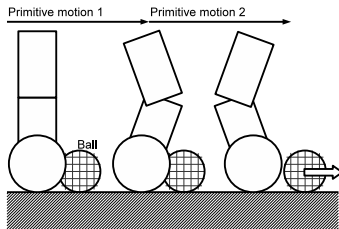


図 5: 倒立二輪型移動ロボットによる蹴球動作

2.3 行動学習

節2.2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[4]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(2)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．

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

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．

3. 実験

動作認識では実際に Kinect センサで人間の蹴球動作の観察を行い，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．模倣の観点から下半身をボールがぶつかる右脚を含むリンクとし，上半身は頭を含むリンクとした．実験結果を図6に示す．複数のリンクをクラスタリングされている $T_e = 0.025, 0.030, 0.035$ の時は学習初期段階は良いが学習後は評価が上がらない．クラスタリング数が少ない $T_e = 0.015, 0.020$ の時は学習初期段階は評価が低い，学習終了後の評価は高くなった． $T_e = 0.010$ は学習途中の評価の上がり下がり激しく不安定でだが高い評価で収束した．

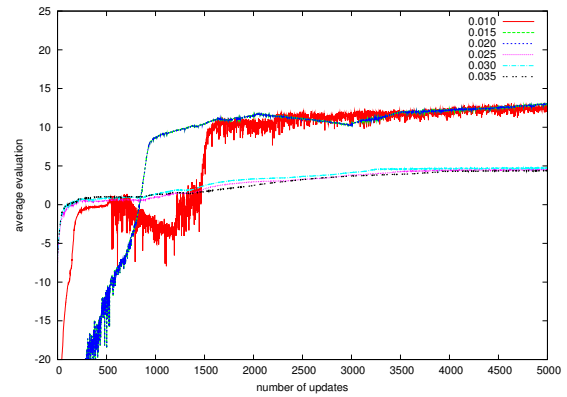


図 6: 実験結果

4. 結言

本研究では倒立二輪型移動ロボットによる模倣学習のための身体部位マッピングの1手法を提案した．実験により人間のリンクを大雑把に分類し，模倣学習した時は，学習の初期段階では評価が良いが，最終評価は悪くなる．逆に細かく分類・マッピングした場合は，学習初期の評価は低く，安定しないが最終評価はよくなることがわかった．今後自動的に上下半身のマッピングを行えるようにしたい．

参考文献

- [1] 浅田 稔: “ロボットの行動学習・発達・進化,” 共立出版, 2002
- [2] 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,” 日本知能情報ファジィ学会誌, pp.1143-1154 (2009)
- [3] 高橋 佐多弥, 高橋 泰岳, 前田 陽一郎, 中村 恭之: “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マッピング,” JSAI2011 講演番号 3B1-OS22b-8 オーガナイズドセッション投稿, (2011)
- [4] Nate Kohl, Peter Stone: “Policy gradient reinforcement learning for fast quadrupedal locomotion,” *Proceedings of the IEEE International Conference on Robotics and Automation 2004*, pp.2619-2624 (2004)