

ヒューマノイドロボットによるキネティックチェーン投球運動学習

西部 慎一† 石田 稔† 加藤 昇平† 伊藤 英則†

†名古屋工業大学 大学院工学研究科 情報工学専攻

1 はじめに

様々な球技において、投球は重要な運動である。人間の投球運動は、主に腰、肩、肘、手首各関節を連動させる複雑な運動によって実現する。近年、ヒューマノイドロボットの運動制御に関する研究においても注目されている運動の一つである。ロボットにおいても効率良くボールにエネルギーを伝えるには運動連鎖（キネティックチェーン）が必要である [1]。先行研究として、倒立二輪型ロボットの投球動作獲得 [2] がある。この研究は身体ダイナミクスを巧みに利用した投球運動を実現しているものの、2 関節 2 自由度のロボットを用いており、複雑なキネティックチェーン投球運動を獲得しているとはいえない。そこで本稿では、シミュレータ上の多自由度ヒューマノイドロボットに、強化学習を用いて投球運動を獲得させた。

2 ヒューマノイドロボット

図 1 にロボットの概観およびリンクモデルを示す。同図において、ロボットは全身に 20 自由度を有しており、各関節は粘弾性特性を有する。また、ロボットの身長および体重はそれぞれ 1.50[m], 56.0[kg], ボールは 0.4[kg] とした。ロボットの右腕部先端にグリッパを有し、ボールを把持する。本稿では、投球時の上半身の挙動に着目するため、腰の 1 自由度、右肩の 2 自由度、右肘の 1 自由度、グリッパの把持、非把持の計 5 自由度を学習対象とした。左腕の関節にはトルクの印加せず、下半身の関節および首関節は固定した。

3 学習アルゴリズム

本研究ではロボットの投球運動学習に、強化学習の一種であるオフライン Q 学習を採用する。ここでは、行動価値関数 (Q 値) 推定のために関数近似器として RBF ネットワーク (Radical Basis Function Network) を用いる。Q 値は以下の式で表現される。

$$Q(s_t, a_t) = \sum_i \omega_{a_t} \phi_i(s_t) \quad (1)$$

$$\phi_i(s_t) = \exp\left(-\frac{d_i(s_t)}{2\sigma_i^2}\right) \quad (2)$$

$$d_i(s_t) = \|s_t - c_i\| \quad (3)$$

ここで、 a は行動、 s は観測された n 次元状態ベクトル、 c_i , σ_i , ω_i は n 次元ガウス関数の中心座標、分散、結合重み、 t は時刻、 i ガウス関数識別子である。 ϕ_i は状態 s_t における基底関数 i の出力を返し、 d_i は状態 s_t と基底関数の中心との n 次元ユークリッド距離を表す関数である。

*Acquisition of kinetic chain throwing motion for humanoid robots based on reinforcement learning, Shinichi NISHIBE†, Minoru ISHIDA†, Shohei KATO† and Hidenori ITOH†
†Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
{nishibe, ishida, shohey, itoh}@juno.ics.nitech.ac.jp

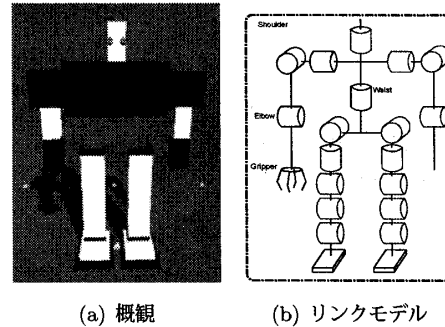


図 1: ヒューマノイドロボット

表 1: 関節印加トルク

部位	腰	肩 (pitch)	肩 (roll)	肘
印加トルク [Nm]	35.0	20.0	17.0	7.0

3.1 Q 関数更新手続き

状態空間のガウス関数の結合重み ω の更新手続きを以下に示す。

$$\omega_t \leftarrow \omega_t + \frac{\alpha \delta_t}{d_i(s_t)} \quad (4)$$

$$\delta_t = r + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \quad (5)$$

ここで、 δ は TD 誤差、 α は学習率、 r は報酬、 γ は割引率を表す。結合重み ω の更新幅は観測状態 s_t と各基底関数の中心の距離 d に反比例する。

3.2 状態空間

ロボットの状態空間を表現する n 次元状態ベクトル s の要素として以下の 7 次元空間を選択した。

$$s = (\theta_w, \theta_{sp}, \theta_{sr}, \theta_e, v_x, v_y, v_z)^T \quad (6)$$

ここで、 θ_w , θ_{sp} , θ_{sr} , θ_e はそれぞれ腰、肩 (pitch), 肩 (roll), 肘の関節角度を表す。 $(v_x, v_y, v_z)^T$ はボールの速度ベクトルを表す。ボールへのエネルギー伝導性を評価するために、速度ベクトルを状態ベクトルとして採用した。

3.3 行動選択

行動選択アルゴリズムとして Softmax 法を採用した。ロボットの各関節は 1 ステップ毎にトルクの印加または非印加のいずれかの行動を選択する。グリッパは開放の行動を持つ。したがって、制御関節数を m とすると、1 ステップ毎に計 $2^m + 1$ 通りの行動が存在する。印加するトルクは関節毎に一定とする。

4 投球運動生成実験

キネティックチェーンを考慮した投球運動学習の効果を確認するために、動力学シミュレータ ODE (Open

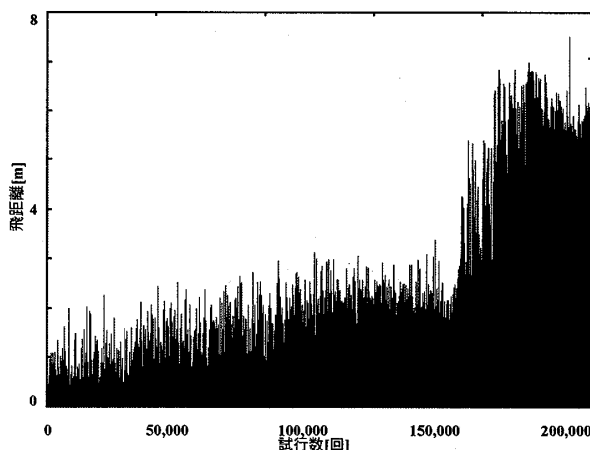


図 2: 飛距離の遷移

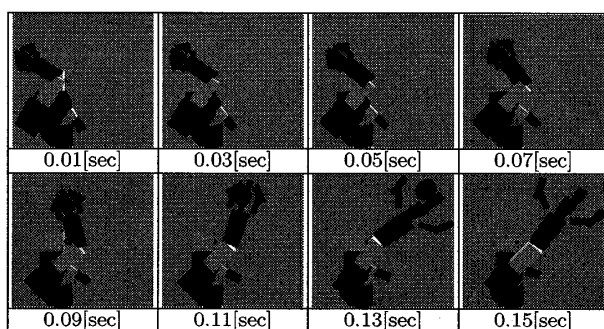


図 3: 投球モーション

Dynamics Engine) を用いて実験を行った。報酬 r はボールの飛距離、1 ステップは 0.01[sec] とし、200,000 試行の投球学習を行った。人間の各関節トルク値を参考に、ロボットの各関節毎の印加トルク値を表 1 のように設定する。初期姿勢は、オーバースローのテイクバック姿勢とした。ロボットが転倒した場合はその試行を失敗とした。

4.1 実験結果

図 2 にボールの飛距離の推移を示す。同図において、学習の経過とともに飛距離が増加しているのが確認できる。学習により得られた投球運動を用いて投げたボールの最高飛距離は 7.44[m] であった。図 3 に同飛距離を記録した投球モーションを示す。また、図 4 に図 3 の投球運動における印加トルクのタイミングチャートを示す。実線はトルクの印加、点線はトルクの非印加を表す。同図より、投球運動開始直後から中盤までは肩 (pitch) に連続してトルクを印加し、中盤からリリース時までには肘に連続してトルクを印加していることがわかる。これはエネルギーをボールに効率よく伝えるために、トルク印加関節を切り替えるキネティックチェーンを学習により獲得していることを示している。また、腰関節に関してはトルクの印加を断続的に繰り返していることがわかる。これは連続でトルクを印加した場合、過剰な遠心力によりバランスを崩し飛距離が伸びない、あるいは、ロボットが転倒することを学習した結果と推察される。

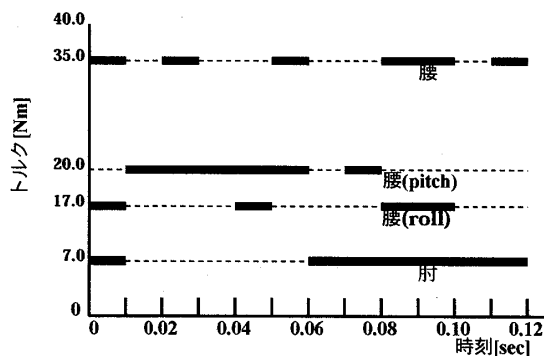


図 4: タイミングチャート

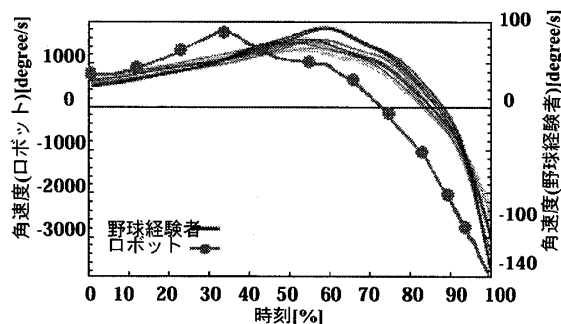


図 5: 肩内外旋角速度

4.2 人間の投球との比較

人間の投球において肩内外旋動作がボール速度生成、および、肘伸展屈曲相互作用トルクに重要な役割を果たしていることがわかっている [3]。図 5 に野球経験者とロボットの投球時の肩内外旋角速度の軌跡を示す。同図より、ロボットの軌跡が野球経験者の軌跡と類似していることが確認できる。これは肘伸展に有効な肩内外旋角速度を学習できていることを示している。

5 おわりに

ヒューマノイドロボットに腰、肩、肘関節を順次運動させるキネティックチェーンを考慮したオーバースロー投球運動を強化学習によって獲得させた。今後の課題として、下半身も含めた運動学習、印加トルクの連続化、および、実機への適用が挙げられる。

参考文献

- [1] 妹尾 拓, 並木 明夫, 石川 正俊: 高速スローイング動作におけるエネルギー伝播の解析. 第 7 回システムインテグレーション部門講演会, 2006.
- [2] 西川 剛広, 高橋 泰岳, 中村 恭之, 浅田 稔, 石黒 浩: 倒立二輪型移動ロボットの全身ダイナミクスを生かした投球動作の学習. ロボティクス・メカトロニクス講演会 2009 講演論文集, 2A2-D18, 2009.
- [3] 高木 博康, 丸山 剛生, 内藤 耕三: 球速の異なる野球の投球動作における肘伸展運動と他関節運動の相互作用関係. 第 20 回日本バイオメカニクス学会大会論文集, p. 106, 2008.