

[研究助成 (A)]

ヒトとロボットの安全な共生のための
人工筋肉駆動ロボットの強化学習による制御
Control of artificial muscle-driven robots by reinforcement learning
aimed to for safe coexistence of humans and robots

2191004



研究代表者

中央大学
理工学部 精密機械工学科

助教 奥井 学

[研究の目的]

本研究は人工筋肉駆動ロボットのための強化学習を用いた制御システムの開発を目標とする。既存のロボットがヒトと同じ環境で動作する場合、柵によって動作範囲を制限したり、動作速度を制限するなどして安全を確保している。これはロボットの活動領域を抑制してしまう。この根本的な要因は、モータ駆動が生物の筋肉駆動と比較し、重く堅い動きとなり衝突時の危険が大きいからである。そこで軽量で柔軟な人工筋肉の開発が盛んであるが、応答の遅さや非線形性といった特性が制御において課題であり、実用に至っていない。この制御性の悪さが解決されれば、人工筋肉駆動ロボットの実用化が近づくと考えられる。

そこで、ヒトが特定の動作に対して反復練習により成功時の感覚を体に覚えこませていることに着目した。この習熟過程で獲得された動作は、生体筋の特性変化や環境変化に対してロバスタであるため、人工筋肉の制御には強化学習が適していると考えた。これまでにも、ソフトロボットや人工筋肉に対して実機を用いた強化学習に関する研究が行われているが、ほとんどが経路計画やフィードバックパラメータを学習

するものであり、人工筋肉駆動ロボットの可変剛性特性を陽に扱っていない。可変剛性特性を活用すれば、ヒトのように運動時に関節剛性を変化させ必要に応じて脱力状態などを活用するなど、ハードウェア性能を最大限活用するような動的な運動生成が期待できる。

そこで本研究では打楽器演奏を対象とし、人工筋肉駆動ロボットのための強化学習を用いた制御システム開発を目標とする(図1)。打楽器演奏を対象とした理由は、動作の中に脱力状態を含む剛性変化が存在するためである。



図1 人工筋肉駆動ロボットシステムの概要

[研究の内容, 成果]

打楽器演奏動作の解析

実ロボットを用いた強化学習では、ロボットのハードウェア性能が所望の動作を再現できる性能を有する必要がある。そこで、ロボット設計に向けて、人間の打楽器演奏動作の解析

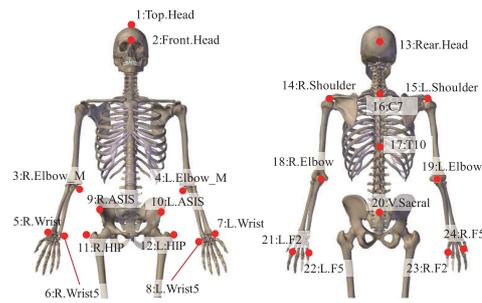


図2 マーカーセット

を行った。解析では、光学式モーションキャプチャシステム (MAC3D) で動作を取得したのち、筋骨格解析ソフトウェア nMotion Musculous (ver. 1.58, ナックイメージテクノロジー株式会社) を用いて関節角度やトルクを算出する。計測では、被験者は高さ 520 mm の椅子に座り、打面高さ 680 mm, 椅子との中心間距離 630 mm に置かれたスネアドラム (YAMAHA, AMS1460RAU) を演奏した。図2に示すように、被験者の各体節の骨格特徴点上にマーカーを 23 点貼り、8 台の計測用カメラを用いて 100 Hz で演奏動作を計測した。計測結果から筋骨格モデルを用いて、肘関節の屈曲、内旋、内転および手首関節の屈曲、撓屈の角度を算出する。また、手首の屈曲についてはトルク算出も行う。なお、トルク算出時にスティックの重さは考慮しない。計測動作は 60 bpm, 音量 *mf* での四分音符とした。被験者はメトロノームに合わせ、測定動作を 2 小節、すなわち 8 打演奏する。被験者は演奏歴 10 年以上の打楽器奏者である。

測定結果を図3に示す。図は8打のうち4打目から6打目を含みよう、測定データのうちの3秒間を示している。肘関節について屈曲、内旋、内転方向を正方向とした。手首関節は屈曲、撓屈方向を正とした。また、実験時に撮影した動画からスティックが打面に当たる瞬間を確認し、グラフ中に赤いラインで示した。

測定データより、演奏では1秒を1周期として、待機状態から振り上げ (バックスイング) 動作を経て、振り下げ動作によって演奏するこ

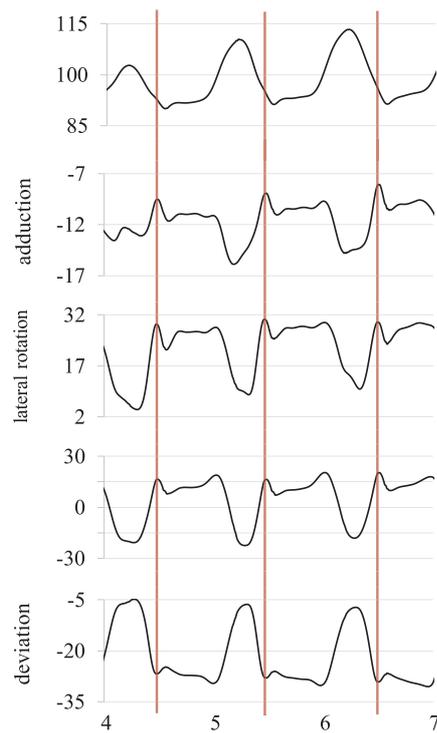


図3 測定角度. 縦軸は角度 [deg] を、横軸は時間 [s] を表す。スティック先端が打面に当たる瞬間を赤線で示した

とがわかる。加えて、降り下げ動作の最後には打面からの反発による関節角度の変動が観察できる。また、振り上げ振り下げ動作については肘関節が手首よりわずかに早く動作を開始していることが観察できる。

各データの振幅に着目すると、ほとんどの関節の関節角度の振幅が 20-30 deg 程度であるのに対して、手首屈曲は 45 deg 程度と大きい。これは、手首屈曲が最も演奏に寄与していることを示唆する。またこの結果は、演奏者へのヒヤリング結果とも一致する。

図4に手首の屈曲トルクを示す。トルクは常に負の値であり、常に重力に対して手首を支える方向に働いている。待機状態では自重を支え、振り上げ動作では、負の方向にトルクを増大し

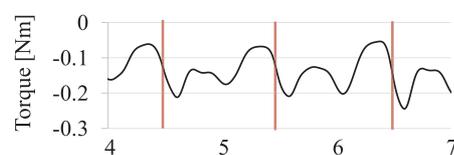


図4 手首の屈曲方向トルク

手首を振り上げ、振り下げにおいては力を抜くことで、自重を活用して手首を動かしている。また、打面をたたいた直後は、振り下げた手首を待機状態とするため動作を止める方向にトルクを発生している。

これより、本実験条件での打楽器演奏においては、打面を叩くために手首を積極的に屈曲方向に動かしていないことがわかる。すなわち、振り上げによって蓄積した位置エネルギーを脱力によって速度エネルギーに変換し、演奏していると推定できる。

ドラムロボットの設計・製作

演奏動作解析結果より、手首屈曲の演奏時の角度変化幅が最も大きく動作に最も寄与していたため、打楽器演奏ロボットは手首のみアクティブな平面動作ロボットとした。その構成を図5に示す。前腕部分に人工筋肉を拮抗配置し、手首関節において可変剛性を実現している。またスティックを握る手の部分はパッシブな関節となっており、手首関節の動きによりスティックが上下に動く。本構造によって、上の人工筋肉の収縮により振り上げ、弛緩により振り下げることによってドラムを叩く動作ができる。

手部は長さ 120 mm、重さは 390 g である。ドラムスティックの持ち手となる部分にパッシブな関節を配置した。前腕部分に McKibben 型人工筋肉 2 本を拮抗配置させ、収縮力がワイヤ（ダイニーマ）により、手首関節を模しているプーリに伝達されてトルクが発生する。

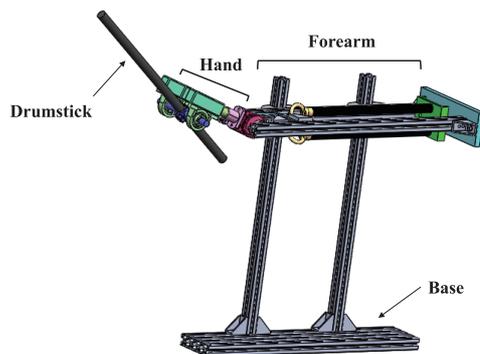


図5 開発したドラムロボット

McKibben 型人工筋肉は Festo 社の Fluidic Muscle DMSP (10) の内径 10 mm、長さ 180 mm の人工筋肉を使用した。プーリ径は 16 mm であり、手首関節部分の可動域は 100 deg、全可動域で 1 Nm のトルクを発生できる。そのため、動作解析の結果より打楽器演奏動作を生成できる。

実ロボットのための強化学習方法

ヒトの動作獲得では反復練習により運動スキルが向上することは経験的に良く知られている。このとき、筋単体は独立に制御されておらず、符号化された動作パターンの組み合わせによって運動していると言われている。そこでロボットの強化学習においても、ハードウェアである筋骨格構造を活用して符号化を行うことで、学習スピードの向上や身体構造を生かした動作獲得ができると考えた。

提案する学習方法の概要を図6に示す。提案手法はモンテカルロ法を参考に開発した。モンテカルロ法はエピソードごとの報酬に基づいて方策を更新する手法である。学習方法における行動は脱力、保持、振り下げ、振り上げの4通り、状態は行動開始からステップ時間ごとの経過時間である。 Δt を学習におけるステップ時間とし、0.1 s とした。 T_i は episode ごとの行動開始からの経過ステップ時間である。また、目標とする発音タイミングを T_d とする。方策である Q table の値に基づき、各ステップ時間での行動を決定する。行動の選択では ϵ -greedy 法を用いた。

Q-table は episode が終了する毎に一度、以下の式に基づき更新される。

$$Q(T_i, a) = (1 - \alpha) Q(T_i, a) + \alpha [r(T_i) \gamma^{(T_r - T_i)}]$$

ただし

$$r(T_i) = r_{\text{success}}(T_i = T_d)$$

$$r(T_i) = r_{\text{failure}}(T_i \neq T_d)$$

Q-table

		Action a			
		Keep	Swing up	Swing down	Relax
State	T_0				
	T_1				
	T_2				
}					
State	T_{r-2}				
	T_{r-1}				
	T_r				

Repeat every time step ΔT

1. Select action a according to Q table using ϵ -greed
2. Time step is updated $T_i \leftarrow T_i + \Delta T$

Until Reward $\neq 0$

Update Q-table according to the equation below

$$Q(T_i, a) = (1 - \alpha) Q(T_i, a) + \alpha [r(T_i) + \gamma^{(T_r - T_i)}]$$

図6 学習方法の概要

である。ここで α は学習率, r は報酬, γ は割引率を表す。また, T_i は報酬を得たステップ時間である。

強化学習実験

実ロボットを用いた学習実験により, 身体をうまく使った動作を獲得できるかどうか検証する。図7に実験環境を示す。制御システムとして dSpace 社製 MicroLabBox を用いる。人工筋肉に印加する圧力は MicroLabBox に接続した電空比例弁で制御する。コンプレッサからの空圧の変化を空気圧センサで検出する。手首部分とドラムスティックの回転角度はエンコーダで計測する。ドラムの打面に取り付けた加速度ピックアップ (NP-3418; 小野計測, 神奈川県, 日本) で音量と発音タイミングを計測する。振動量を振動加速度に比例した電気信号として取り出す際にアンプ (SR-2210; 小野計測, 神奈

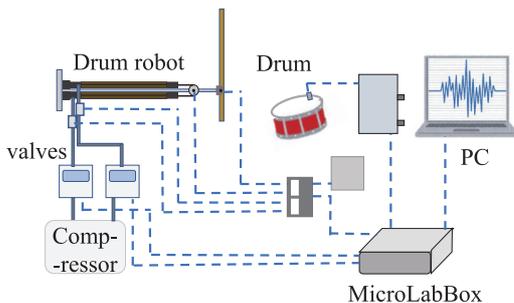


図7 実験環境

表1 行動と人工筋肉へ印加圧力

	Extensor	Flexor
Keep	0.35	0
Swing up	0.5	0
Swing down	0	0.5
Relax	0	0

川県, 日本) を用いる。制御, 計測周期はともに 100 Hz とした。ロボットの関節角度の定義は, Fig と同様の設定である。手首関節の可動域は ± 40 deg となるようにワイヤを設定した。スティック部については $-5 \sim 90$ deg となるように, バンプラバーを取り付け可動域を制限した。

状態行動空間は図6と同様である。4つの行動と人工筋肉への印加圧力の関係は表1に示す通りである。エピソードは, 2秒間の Keep 状態ののちに開始する。打面を叩くか, エピソード開始から 1.2 s 経過で試行を終了する。エピソード終了時に報酬を受け取り, Q table を更新する。加速度ピックアップの出力の絶対値が 1 を超えた際に打面を叩いたと判定する。報酬は, エピソード開始から 3 ± 0.05 s で叩いた際に, 加速度ピックアップで計測した加速度の立ち上がりを報酬とした。3 \pm 0.05 s 以外のタイミングで叩くか, 叩かずに 3.5 秒経過した際の報酬は -1 とした。 $\epsilon = 0.01$, $\alpha = 0.1$, $\gamma = 0.9$ とした。

実験の結果, 約 1000 エピソードでの学習の収束を確認した。収束時の関節角度および圧力, 加速度ピックアップで計測した値を図8に示す。角度のグラフより, 振り上げ, 降り下げから大きな音量を適切なタイミングで出すことに成功した。学習後のデータから, 振り上げから降り下げをへて, 骨格構造を活用した運動生成を確認できる。しかし, 振り上げ動作が2回確認でき, 人間の動作とは異なる結果となった。これは, 人間の場合は消費エネルギーを少なくするために運動を改善するが, 今回の報酬設定の場合にはただ大きな音を出せばよい条件のためだと考える。なお, 打面を叩く直前での急峻な角度の方向変化は, 関節可動域の限界に達してス

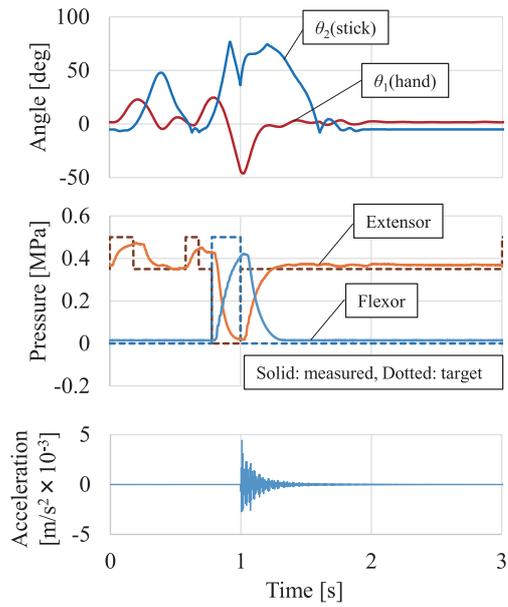


図8 獲得した動作

ティックがラバーにぶつかり、跳ね返っているためである。

[今後の研究の方向, 課題]

実ロボットを用いた実験により, 提案手法する学習方法によってハードウェア特性を活用した動的な運動生成が可能なことを示した。今後は学習方法の理論的な考察および他手法との比較を行う。

[成果の発表, 論文等]

- [1] 奥井 学, 西濱里英, 鈴木隆二, 中村茉莉, 木村成吾, 中村太郎, “空気圧人工筋肉を用いた打楽器演奏ロボットの開発に向けた打楽器演奏動作の解析”, 第38回日本ロボット学会学術講演会, 1H3-5 (2020.10)
- [2] 奥井 学, 中村茉莉, 木村成吾, 中村太郎, “実機を用いた強化学習のための空気圧人工筋肉駆動打楽器演奏ロボットの開発”, ロボティクス・メカトロニクス講演会 2021, E07-2P2, (2021.6)