



令和3年6月7日

報道機関 各位

東北大学大学院工学研究科

### 深層強化学習による自然なリーチング運動パターン生成 - 人間の計測データを用いず7自由度アームの運動シナジーを発現 -

#### 【発表のポイント】

- 人間の計測データを使わずに深層強化学習により、自然なリーチング運動パターンを生成できることを示した。
- 運動シナジーの発現度合いが運動学習の習熟度と連動しており、特にエネルギーあたりの運動パフォーマンスと高い相関があることを示した。
- フィードバック制御器と共に用いることで深層学習の速度を3割程度速くできること、またエネルギー効率性とシナジー度合いもさらに高まることを定量的に示した。
- フィードバック制御依存からフィードフォワード制御にシフトする運動学習プロセスを深層強化学習のフレームワークで再現した。

#### 【概要】

これまでの深層強化学習による高自由度関節モデルを用いた運動生成では、タスクは実行できても多くは不自然な運動結果となるため、人間の計測データを用いた模倣学習によって自然な運動パターンを生成するアプローチがとられてきました。

東北大学大学院工学研究科の林部充宏教授と Han Jihui 大学院生(研究当時)らの研究グループは、人間の計測データを使わずに深層強化学習によって自然なリーチング運動パターンを生成する手法を提案しました。人間の計測データを一切用いずに、運動習熟レベルが進むほど運動シナジー強度が増大していくプロセスを定量的に再現することに成功しました。真の意味で未知の物理的環境下での運動学習の方法としての解決策やシナジー生成メカニズムを明らかにすることは容易ではなく、どのような計算指針でシナジーが生成されるのかを扱うものがこれまでほとんどありませんでした。本研究では深層強化学習において環境適応性を確保しつつ運動シナジーが発現するプロセスを再現できるかどうかを検証しました。

本研究成果は、Journal 誌「IEEE TRANSACTIONS ON MEDICAL ROBOTICS AND BIONIC」に2021年5月20日付けで掲載されました。

関連動画: <https://www.youtube.com/watch?v=juU5RNWZQ9k>

**【問い合わせ先】**

**< 研究に関して >**

東北大学大学院工学研究科 ロボティクス専攻

教授 林部 充宏

電話 022-795- 6970

E-mail: [mitsuhiro.hayashibe.e6@tohoku.ac.jp](mailto:mitsuhiro.hayashibe.e6@tohoku.ac.jp)

**< 報道に関して >**

東北大学工学研究科情報広報室 担当 沼澤 みどり

TEL: 022-795-5898

E-mail: [eng-pr@grp.tohoku.ac.jp](mailto:eng-pr@grp.tohoku.ac.jp)

## 【詳細な説明】

### ＜研究の背景＞

人間の運動制御問題は多数の筋骨格を制御してはじめて成立することから、多数の関節の冗長性問題と多数の筋肉の冗長性問題を解決する必要があります。しかしながら、実際には我々人間は容易に問題を解決していることから、どのようなメカニズムでこの多自由度空間問題を解決しているのかという議論が昔からなされてきました。ベルンシュタイン問題と言われ、Nicholai A. Bernstein の階層的運動制御の考え方に基づいて運動シナジーの存在が示唆されました。その後の研究で運動シナジーが人間や生物の運動制御で用いられていることが確認されましたが、中枢神経がどのような法則に基づいて、どのようなメカニズムで生成されているかの計算論的数理モデル構築には至っていないのが現状です。

これまでの計算論的神経科学では、何らかの評価関数を最小にする(最適化する)ような計算方法が提案されています。実際に最適化計算を行うと人間らしい動きを生成できますが、それは人間がある指針を最適にするように動いていることの証明になってはいるものの、その数学的最適化計算には環境と身体の数学的モデルが事前に必要となってしまいます。真の意味で未知の物理的環境下での運動学習の方法としての解決策やシナジー生成メカニズムを明らかにすることは容易ではなく、どのような計算指針でシナジーが生成されるのかについて扱うものがこれまではほとんどありませんでした。本研究では深層強化学習において環境適応性を確保しつつ運動シナジーが発現するプロセスを再現できるかどうかを検証しました。

本研究は科学研究費補助金（新学術領域）超適応プロジェクト(B05-01)20H05458 の支援を受けて行ったものです。

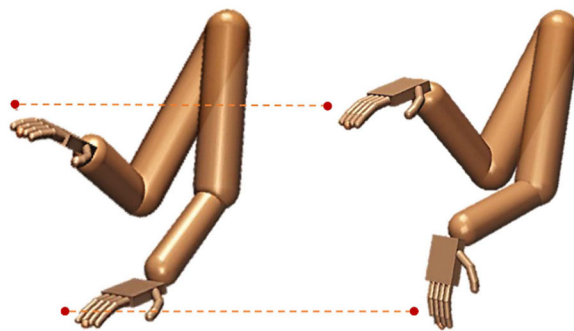


図 1 深層強化学習により環境の違いに適応している様子(左側は軽い手、右側は重い手の条件で発現されたリーチング運動の様子)

### ＜研究の内容＞

本研究では高自由度多関節アームを使用し、事前のモデルや環境の情報を全く与えずに、純粹に深層強化学習のみによる繰り返し試行によりリーチング運動の学習を行い、関節空間の運動制御がどのように変化しているかを調査しました。また、

関節運動の時空間パターンに PCA (独立主成分分析) を用いることで、運動シナジの発現度合いを試行ごとに定量化し、運動習熟度との連動性を調べました。図 2 のように学習が進むにつれて、運動シナジの発現により関節の連動性が高まること、またそれがフィードフォワードのパターンを形成していることを確認しました。PD フィードバック制御と深層強化学習(DRL: Deep Reinforcement Learning)を同時に使用する PDRL フレームワークを新たに提案し、フィードバック制御からフィードフォワード制御に遷移していく運動学習の様子を再現しました。さらに、運動シナジの発現度合いがエネルギーあたりのパフォーマンスと高い相関関係にあることがわかりました。すなわち、エネルギーあたりのパフォーマンスを効率よく高めるための必要条件として運動シナジが採用されていることが示唆されました。

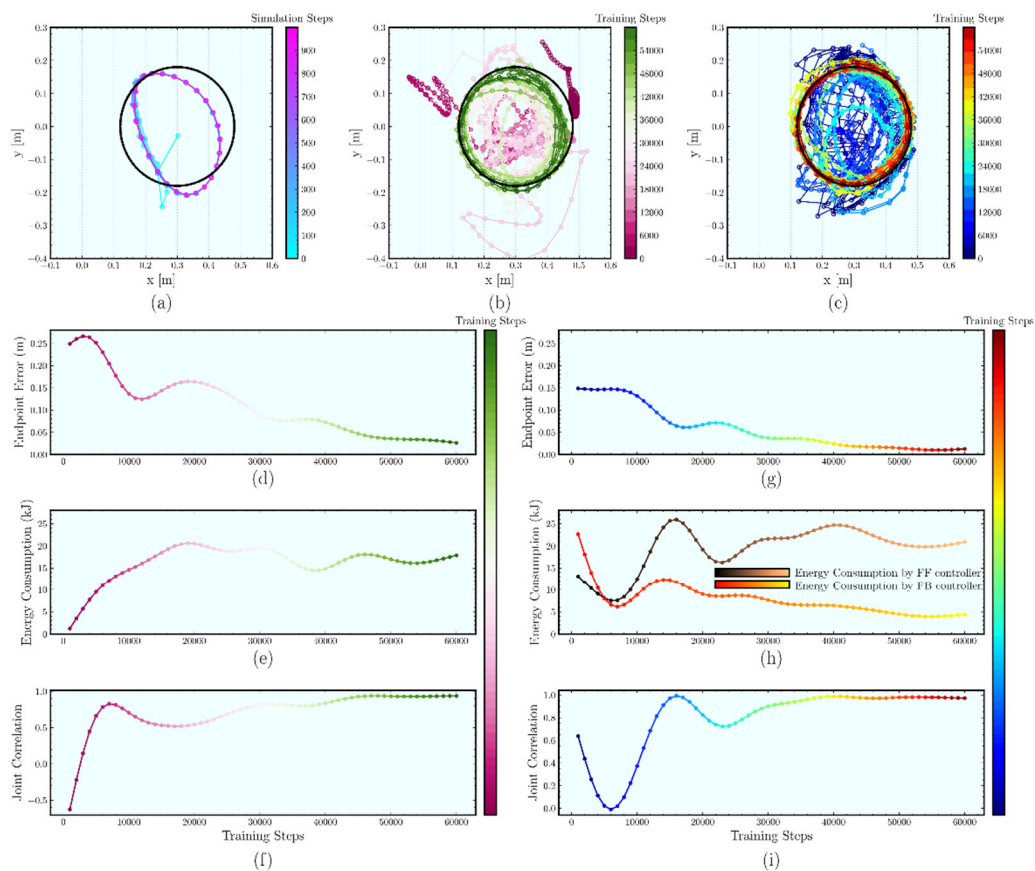


図2 7 自由度アームの終点遷移(a)は PD フィードバック制御、(b)は DRL、(c)はフィードバック制御と併用した PDRL 制御の場合。終点誤差の推移(d)DRLと(g)PDRL。エネルギー消費量の推移(e)DRL, (h)PDRL。エネルギー消費がフィードバックからフィードフォワードに遷移している。関節連動性の変化 (f) DRL と (i) PDRL。関節連動性がフィードフォワードのエネルギー消費量に対応していることが確認できる。

#### <研究の意義・今後の展望>

深層強化学習による運動学習タスクにおいて運動シナジーの発現プロセスが起きており、それがエネルギー当たりのパフォーマンスと高い相関を示したことは、なぜ人間や生物が運動シナジーを活用しているのかという問いの答えにつながるため、科学的な意義が高いと考えられます。2020年1月に発表した論文[1]では、歩行運動についても同様の知見を得ています。また工学的な応用としては現在の深層学習は膨大な計算コストを要しますが、効率的な運動学習における潜在的な方策として運動シナジーを用いることができれば大幅な計算の効率化につながるため、本論文は新しい深層強化運動学習アルゴリズムに向けて示唆に富む情報となることが期待されます。

- [1] Jiazheng Chai, M. Hayashibe, “Motor Synergy Development in High-performing Deep Reinforcement Learning algorithms”, **IEEE Robotics and Automation Letters**, April 2020, 5(2):1271-1278  
<https://ieeexplore.ieee.org/document/8966298>

#### 【論文情報】

タイトル: Synergy Emergence in Deep Reinforcement Learning for Full-Dimensional Arm Manipulation

著者: Jihui Han, Jiazheng Chai, M. Hayashibe

掲載誌: **IEEE Transactions on Medical Robotics and Bionics**, May 2021, 3(2): 498-509

DOI: 10.1109/TMRB.2021.3056924

URL: <https://ieeexplore.ieee.org/document/9345796>