

令和2年2月25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運動学習過程における
シナジー発現プロセスの存在を実証**
- 運動学習メカニズムの解明に貢献 -

【発表のポイント】

- 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いることを初めて定量的に示した。
- 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。
- 人間の運動学習プロセスと関連のある運動シナジーが、深層強化学習でも発現していることから、人間や生物の運動学習メカニズムの理解にもつながる。

【概要】

東北大学大学院工学研究科の林部充宏教授と Jiazheng Chai 大学院生らの研究グループは、深層強化学習^{*1}による運動学習過程において運動シナジー^{*2}の発現プロセスが起きていることを初めて定量的に示しました。深層強化学習アルゴリズムを用いた報酬^{*3}では運動パフォーマンスとエネルギーを考慮しただけで、シナジーについては何も指定していないにもかかわらず、運動習熟レベルが進むほど運動シナジー強度も増大しました。また運動シナジー発現が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ました。運動学習過程で潜在的に起きている現象との関連性が期待されるため、人間や生物の運動学習メカニズムの理解にもつながると期待されます。

本研究成果は、ロボット分野で最もメジャーな国際会議 IEEE ICRA2020 の Journal Option として採択され、科学雑誌「IEEE Robotics and Automation Letters」に2020年1月22日付けで掲載されました。

【詳細な説明】

1. 研究の背景

人間の運動制御問題は工学的には多数の筋骨格筋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は昔からされてきました。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、計算論的に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それが生成されているかは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も生成できますが、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はなっていますが、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うため、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必要があり、これまでは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運動学習を全探索的に行うとき、何が起きているかを調査し潜在的な計算指針がないかどうかを調べました。

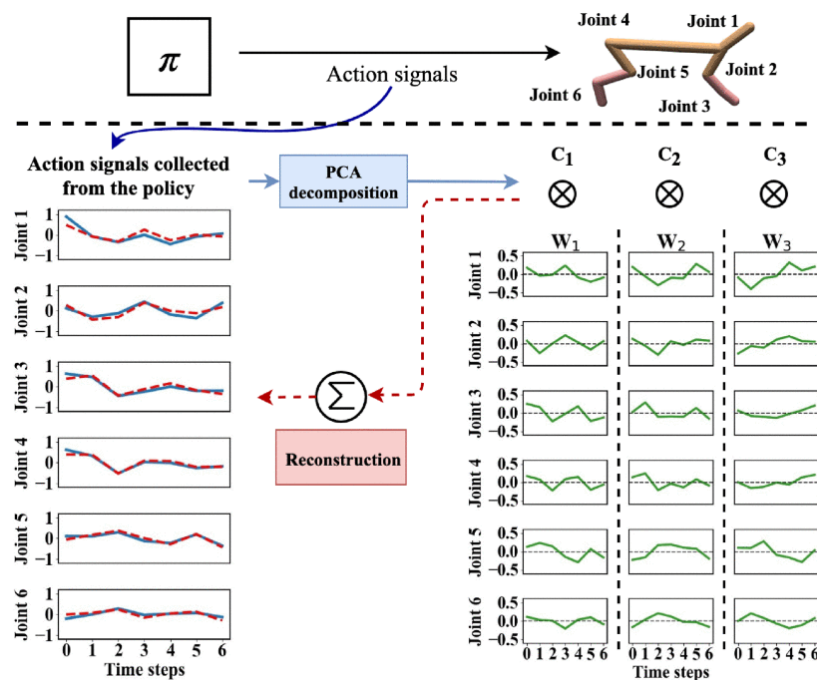


図 1 深層強化学習による歩行タスクを行いながら時空間シナジーを算出する模式図

2. 研究の内容

本研究では図 1 に示すような多関節の歩行エージェントに対し、事前のモデルや環境の情報を全く与えずに、純粋に深層強化学習のみによる繰り返し試行により歩行タスクの学習を行い、関節空間の運動制御信号がどのように変化しているかを調査しました。関節トルク入力空間の時空間パターンを PCA(独立主成分分析)で運動シナジーの発現度合いを各試行ごとに定量化することで、運動習熟度と運動シナジーの発現度合い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変化している運動シナジーの発現度合いを調べると、確かにタスク習熟度が進むにつれて運動シナジーの発現がおきていることがわかりました。また 2 種類の異なる深層学習アルゴリズム SAC(Soft Actor-Critic)、TD3(Twin Delayed Deep Deterministic policy)で調査をすると、より効率的に報酬を高める結果が出た SAC の学習結果の場合の方がより多くの運動シナジーの発現が起きていることがわかりました。2 次元モデル(Half Cheetah)、重い 2 次元モデル(Heavy HC)、3 次元モデル(Full Cheetah)で深層強化学習を行い、それぞれ 300 万時間ステップの運動学習を実施しました。深層学習アルゴリズムはどの関節がどこについているかという予備知識は使っていないが、実際に学習後の運動結果を見ると同一足内の関節間の運動が同位相で駆動されているため、時空間的な低次元化が実際に起きていることが確認されました。また運動シナジーの発現度合いがエネルギーあたりのパフォーマンス(歩行速度)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良く高めるための必要条件として運動シナジーが採用されていることが示唆されました。

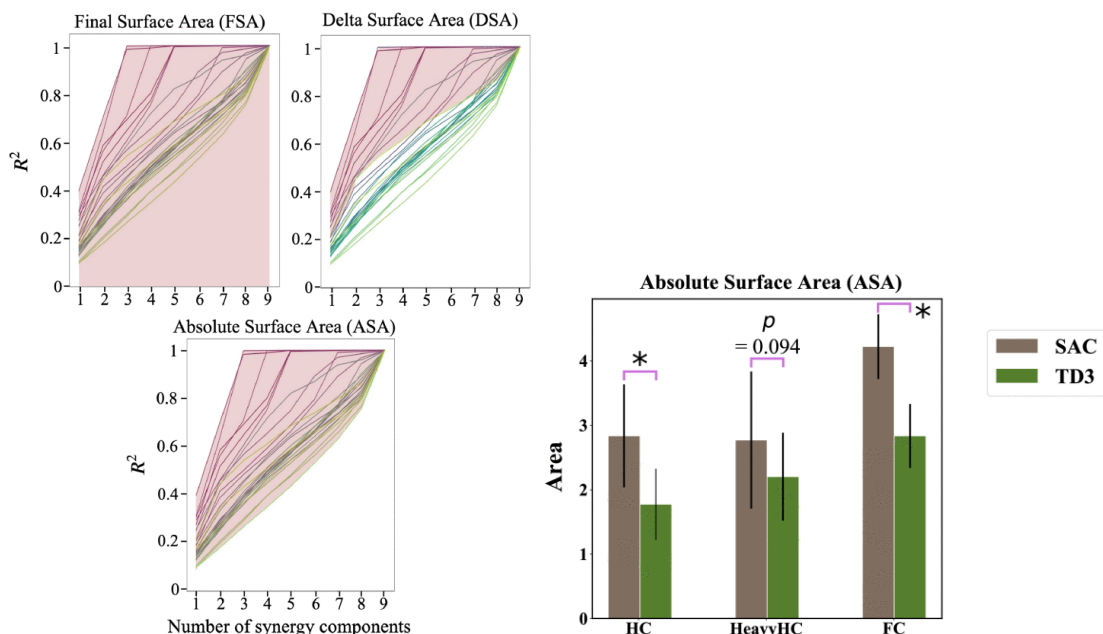


図2(左) 運動シナジーの発現を低次元化度合いで定量化した様子

(右) 運動シナジーの発現度合いを各深層学習アルゴリズムで比較した様子

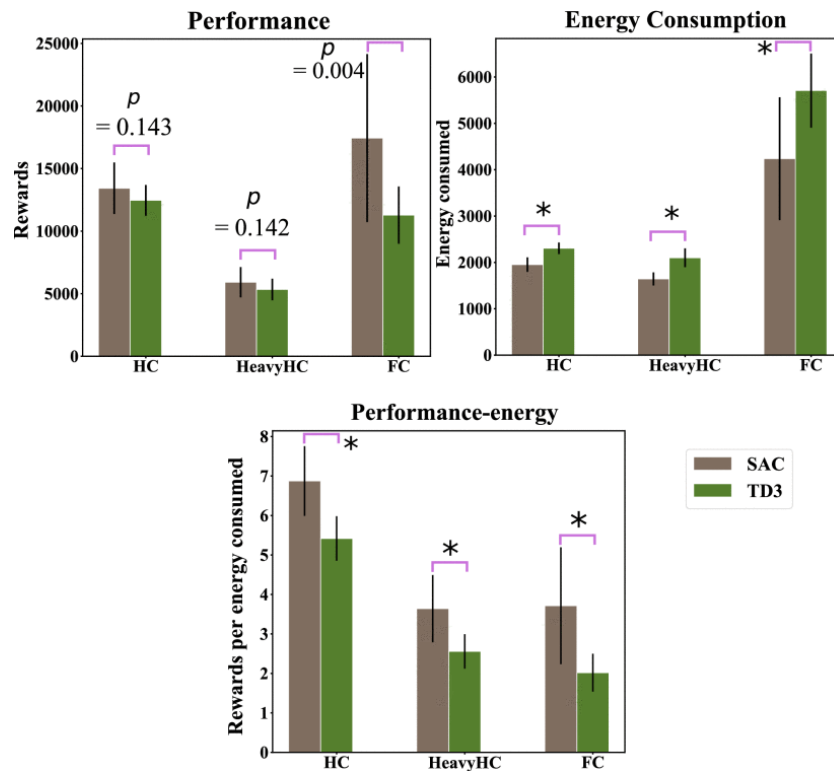


図3 エネルギーあたりのパフォーマンス(歩行速度)を各深層学習アルゴリズムで比較した様子(パフォーマンス単独、エネルギー単独よりもエネルギーあたりのパフォーマンスとよく連動していることがわかった)

3. 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、それがエネルギーあたりのパフォーマンスと高い相関を示したことは、何故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。本論文では歩行運動で検証しましたが、我々の研究室ではリーチング運動でも運動シナジーの発現とエネルギーあたりのパフォーマンスとの相関関係を実証しています。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するが、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げることができるため、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。

【用語解説】

注 1: 深層強化学習

深層学習(ディープラーニング)とは、生物の神経回路を模擬する多層ニューラルネットワークによる機械学習手法で、深層強化学習は一連の行動を通じて報酬が最も多く得られるような方策を学習する最近注目されている AI 計算手法です。

注 2: 運動シナジー

運動が時空間的にある一定の組み合わせで協調的に制御されているという考え方で、人間や生物の運動にはこの協調構造が採用されていると考えられています。

注 3: 報酬

強化学習の枠組みで用いられる計算指針のようなもので、AI が導いた結果の評価として得られるのが報酬です。

【論文情報】

タイトル: Motor Synergy Development in High-performing Deep Reinforcement Learning algorithms

著者: Jiazheng Chai, M. Hayashibe

掲載誌: **IEEE Robotics and Automation Letters**, April 2020, 5(2):1271-1278

URL: <https://ieeexplore.ieee.org/document/8966298>

Journal option として ICRA2020, Paris (May31-June4) でも発表予定

◀ 関連する論文 ▶

タイトル: Synergetic Learning Control Paradigm for Redundant Robot to Enhance Error-Energy Index

著者: M. Hayashibe, S. Shimoda

掲載誌: **IEEE Transactions on Cognitive and Developmental Systems**, 2018, 10(3):573-584

URL: <https://ieeexplore.ieee.org/document/7954741>

【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

◀ 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