〈一般研究課題〉　人工知能とともに探求するチーターの
運動知能
助 成 研 究 者　名古屋工業大学　　上村　知也

人工知能とともに探求するチーターの運動知能

上村　知也
(名古屋工業大学)

Investigating locomotion intelligence of cheetah with artificial intelligence

Tomoya Kamimura
(Nagoya Institute of Technology)

Abstract：

Animals use dynamic interaction between body, cranial nerves and environment to produce sophisticated locomotion, such as the fast and adaptive running of cheetahs. In this study, we constructed a control schematic for a robot with an animal-like hierarchical structure using deep reinforcement learning and neural oscillators and verified its effectiveness. As a result, the obtained controller showed robustness to disturbances while utilizing the system's inherent limit cycles. The results of this research play a significant role not only in robotics, but also in understanding the intelligence that actuates animal bodies.

1. はじめに

動物は全身の多くの自由度を制御し，巧みな運動を実現している．その運動は身体・脳神経・環境の力学的相互作用からなる自然な運動を活用していることが特徴であり，一般的なロボット制御で用いられる逆モデルを明示的に構築しない．逆モデルがないため，脳神経系は時間をかけて繰り返し計算を行う必要がなく，極めて高速に運動したり，状況に応じて適応的な行動を取ったりすることができる．本研究では，動物の中で最も高速に走行するチーターの適応的な走行運動に着目する．

動物の神経系は階層的な構造を持ち，上位中枢となる脳は抽象的な司令を脊髄に送り，下位に存在する脊髄や神経系は具体的な司令を担い全身の筋肉を駆動する役割を持つ(図1A)．これまで多

くの研究において，筋肉を制御するための神経系と，身体・環境の力学的相互作用について研究が行われてきた．例えば運動を司る神経振動子(Central Pattern Generator, CPG)をモデル化した位相振動子によって，4足動物の多様な歩容が創発されることや，外乱に対する適応的な振る舞いが得られることなどが示されてきた [1]．しかしこれらの研究の多くは，上位中枢による司令を持たず，下位の神経による制御が大きな役割を果たしている．上位中枢を導入することによって，より優れた適応性や多様な動作の生成が可能になると予想される．

近年は機械学習，とくにロボットに経験則から制御則を獲得させる深層強化学習が大きく着目されている．これらは多くの場合，多次元の入力変数からアクチュエータの指令値を導くEnd-to-endな学習を行うが，状態-行動対の総数が大きくなりすぎて学習に時間がかかりすぎるという問題がある．この問題に対して様々な解決策が提案されているが，本研究では，動物のように動物の階層的な制御構造を導入することで，制御のための次元を削減し，効果的な学習を行うことを考えた．具体的には，強化学習エージェントを上位中枢として，下位にあたるCPGのパラメータを上位からの司令によって制御する(図1B)．さらに，モータ制御を下位の神経振動子に任せることによって，上位中枢にあたる学習器の計算負担が低減することも期待される．

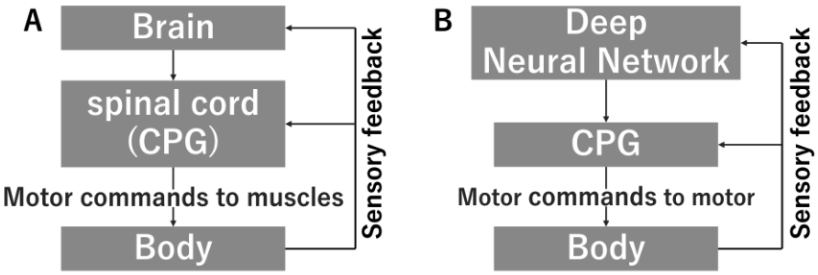

図1　身体を駆動するための階層的な制御構造．

2. 手法

2.1 モデル

本研究では，チーターの高速走行をシンプルな力学モデルと深層強化学習によって実現する．まず，チーターの走行における矢状面上の運動に着目するため，図2に示すシンプルな前後2足ロボットを動力学シミュレータMuJoCo内に構築した．ロボットは体幹部が矢状面に拘束されていて，2次元平面内を移動する．ロボットの脚は2リンクからなり，それぞれの関節は位置制御・トルク制御が可能なサーボモータが取り付けられている．モデルの質量・慣性モーメント・各部の長さは実際のチーターの測定データから決定した．

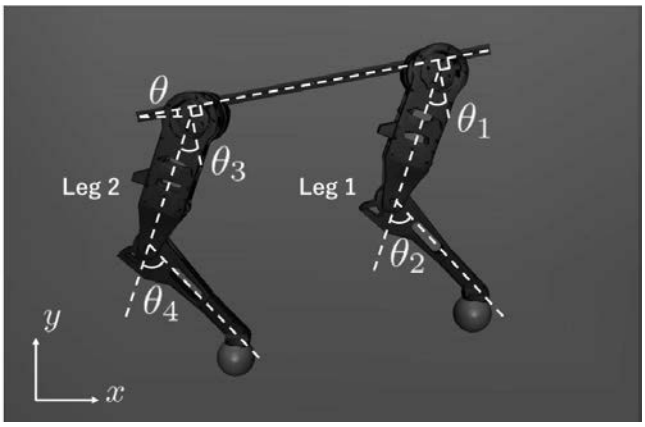

図2　動力学シミュレータ内に構築したロボットモデル．

脚に取り付けたモータは，ダイナミックな走行を効率よく行うため，近位側と遠位側で異なる制御を行った．近位側(肩関節・股関節)のモータは以下で定義する位相振動子の位相に応じて位置制御を行う．

$$\dot{\phi}_1 = \omega + \varepsilon \sin(\phi_2 - \phi_1 - \pi) - \phi_1^{\mathrm{td}} \delta(t - t_1^{\mathrm{td}}) \tag{1}$$

$$\dot{\phi}_2 = \omega + \varepsilon \sin(\phi_1 - \phi_2 + \pi) - \phi_2^{\mathrm{td}} \delta(t - t_2^{\mathrm{td}})$$

ここで，ϕ_1とϕ_2はそれぞれ前後脚の位相であり，ωは標準位相角速度，$\delta(\cdot)$はディラックのデルタ関数である．右辺第２項は前後脚を逆相に保つ効果を持つ．右辺第３項は位相リセットと呼ばれる効果を表していて，脚が接地した瞬間に位相をゼロに戻す．上記で定義される位相振動子に対して，デューティ比βを用いて，接地相($0 \leq \phi_i < 2\beta\pi$)と遊脚相($2\beta\pi \leq \phi_i < 2\pi$)を定義する．近位関節は，接地相ではあらかじめ定義された離地角Γ_i^{lo}，遊脚相では接地角Γ_i^{td}を目標角として位置制御される．上記の制御則をまとめて本研究で用いるCPGとする．

遠位側のモータは，動物の脚が示す弾性的な振る舞いを再現させるために，位相振動子の位相に応じて，前後脚ともに常に目標角度γ_i^{td}，γ_i^{lo}を目標角として柔らかいPD制御を行う．

2.2 深層強化学習

本研究では，CPGを調節する上位中枢として，深層強化学習エージェントを用いる．具体的には，エージェントは現在のロボットの姿勢や関節角度を観測し，それに基づいてCPGのパラメータω，β，Γ_i^{td}，Γ_i^{lo}，γ_i^{td}，γ_i^{lo}を徐々に調整する．

あらかじめ発見しておいた，安定して走行するCPGパラメータを初期値として，そこから外乱環境下での学習を行う．具体的には，定常走行から外れた高さでロボットを宙吊りにして，それを初期値としてシミュレーションを開始し，安定した走行(リミットサイクル)を実現させるように学習させる．報酬rは以下のように定義した．

$$r = w r_{\mathrm{v}} + r_{\mathrm{h}} + r_{\mathrm{c}} + r_{\mathrm{p}} \tag{2}$$

ここで，r_v，r_h，r_c，r_pはそれぞれ，目標速度に対する報酬，転倒時のペナルティ，関節が接地したときのペナルティ，体幹ピッチ角が大きくなりすぎないようにするペナルティを表す．wは重みである．強化学習には，世界モデルベースの深層強化学習アルゴリズムであるDreamerV2 [2]を用いた．

3. 結果

複数の目標速度に対して学習を行った結果，それぞれ結果が収束した．得られた運動の様子を図3に示す．学習前(CPGのみ)と学習後(CPGと学習器の組み合わせ)において，いずれも同様の運動に収束していることから，学習によって大きく運動が変化せず，元になったリミットサイクルを活用している．

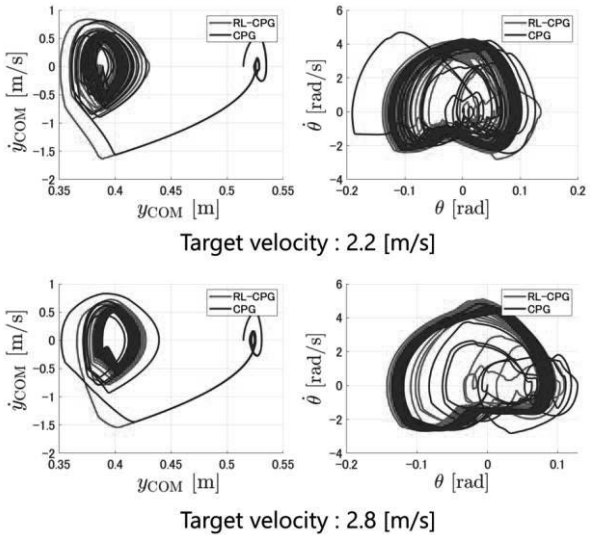

図3　学習前(CPG)と学習後(RL-CPG)における運動の様子．

得られた運動に対して，外乱に対する応答性を比較するため，ロボットの質量中心にランダムな外力を加え続ける実験を行った．その結果を図4に示す．前進速度$\dot{x}$がゼロになっているとき，ロボットは転倒している．独立した5回の実験を行った結果，CPGのみを用いる学習前は4から5回転倒しているのに対して，学習後は一度も転倒しなかった．学習後は学習前と比較して，上位中枢である学習器が適応的にCPGのパラメータを変化させるため，高いロバスト性を示すと言える．

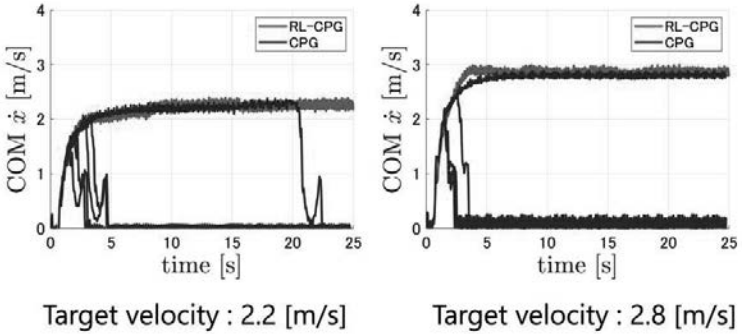

図4　学習前(CPG)と学習後(RL-CPG)のそれぞれに外乱を加え続けたときの応答．

4. 考察

本研究では，ロボットの制御則を構築するにあたって，CPGと深層強化学習を用いて動物の持つ階層的な制御構造を再現した．CPGのみによる制御とCPGと深層強化学習を組み合わせた制御の結果を比較すると，両者はほぼ同じ定常状態に収束していた(図3)ことから，上位中枢は制御によって新しい運動を生成するわけではなく，システムがもつリミットサイクルを活用し，CPGを

補助することでロバスト性を向上させる役割を果たしていると考えられる．この，CPG が運動を生み出し，上位中枢がそれを補助するという制御構造は，動物が持つ神経系の制御構造と類似している．複雑な運動のすべてを脳で制御するのは多大な計算量を要するため，非常に効率が悪い．これに対して，CPGなどの低次の神経系が運動制御の主要な部分を担い，高次の神経系がそれを補助することで，計算コストを抑えながらも高い適応性能を実現できたと考えられる．

また，CPGと深層強化学習を組み合わせることで，外力が加えられた環境でも走行を維持できた．ここから，深層強化学習を用いて上位神経系のモデルを獲得することで，CPG のみによる制御よりもロバスト性の高い制御則を獲得できたと言える．訓練時に初期値からリミットサイクルに収束する際，この差異を外乱とみなして調整する方法を学習器が見出したことで，転倒を防ぐための適切な制御が獲得されたと考えられる

CPGと強化学習エージェントを組み合わせることによる効果は，計算資源の節約とロバスト性以外にも存在すると期待される．本研究ではロボットの自由度を制限していたほか，あらかじめリミットサイクルを発見しておくなど，学習による運動探索の自由度を大きく制限していたため，得られたメリットが限定的であったと考えられる．今後の研究においては，より一般化したモデルや学習対象に提案手法を適応していきたい．

参考文献

[1] S. Aoi, P. Manoonpong, Y. Ambe, F. Matsuno and F. Wörgötter, ”Adaptive control strategies for interlimb coordination in legged robots: A review,” Frontiers in Neurorobotics, vol. 11, no. 39, p. 39, 2017.

[2] D. Hafner, T. Lillicrap, M. Norouzi and J. Ba, ”Mastering Atari with Discrete World Models,” in ICLR 2021 - 9th International Conference on Learning Representations, 2020.