物体操作に関する脳の情報処理構造を参考にした運動学習モデル ～身体モデルと目標運動軌道との並行学習～

Brain-inspired model of motor learning mechanism for object manipulation - Co-learning model of body image and desired behavior -

○有江 浩明 (早大) 正 尾形 哲也 (京大) 谷 淳 (理研) 正 菅野 重樹 (早大)

Hiroaki ARIE, Waseda University, arie@sugano.mech.waseda.ac.jp

Tetsuya OGATA, Kyoto University

Jun TANI, RIKEN

Shigeki SUGANO, Waseda University

Abstract— We humans learn behavior skills through repetitions of own behavior experiences. In the brain science perspective, we speculate such skilled behaviors might be acquired in inferior parietal lobe through the repeated sensory-motor experiences and propose a brain model. We utilized two neural networks. One is a Continuous Time Recurrent Neural Network (CTRNN) which learns the trajectory of joint angles to achieve the task. Another is Recurrent Neural Network (RNN) which learns the relation of muscle activity to joint angles. To test the proposed model, we applied it to the task of reaching for an object using real robot with a tendon-based actuation mechanism.

Key Words: Tendon-driven robot arm, Neural Network

1. はじめに

我々人間は試行錯誤を繰り返す中で、自らの経験を通して道具の使用方法や物体操作などの様々な運動スキルを学習してゆく。我々はこの様な発達的な運動スキルの学習に着目し、脳科学の観点から図1に示すモデルを提案している。このモデルは2つの部分によ

Fig.1 brain model

て構成されており、1つはPMv-IPL-M1ネットワークであり、この部分は目的を持った行動の認識・生成を行うミラーシステムである。また、もう1つは評価値を基に学習を行うBG-IPLネットワークである。本研究では後者のBG-IPLネットワークに関して、Actor-Critic法による強化学習でモデル化を行い、実機ロボットを用いて実験を行った。

このモデルでは、ロボットは探索的な行動を行い、そのときに得られた何らかの評価値（報酬）を用いて強化学習を行っており、これによりセンサ・モータの軌跡を繰り返し学習することでInferior Parietal Lobe(IPL)に運動スキーマが獲得されるのではないかと考えている。

従来、Parietal Cortexは多種感覚入力の統合・連想において重要な役割を果たす部位であると考えられていた。しかしながら、近年の神経心理学の研究では、IPLが運動スキル、特に物体操作における運動スキルの表

現に関連しているのではないかと指摘されている。我々は、Ehrssonら¹⁾の実験結果から、この領域が行為に伴うセンサ入力の予測を行っているのではないかとこの仮定をおいている。さらに、IPLはより高次に位置すると考えられるVentral Premotor(PMv)と階層的な相互作用を行うことで、目的に基づいた行為の生成かつ観測した他者の行為の目的の認識も行っているのではないかと考えている²⁾。

2. 提案モデル

図2に本実験で実装するモデルを示す。まず運動指令の流れに関して、ロボットから取得した関節角度情報が上位のCTRNNへと入力され、これに基づきCTRNNは目標関節角速度を下位のRNNへ出力する。これを受け下位のRNNはこの目標角速度を実現するために関節に加えるべきトルクを出力する。但しここで、上位の学習器が強化学習を行っているため、探索のためのノイズをRNNから出力された指令トルクに加えている。

Fig.2 implementation

2.1 BG-IPL Model

本モデルでは、BG-IPLに対応する部分に、TD誤差を用いた強化学習手法の一つであるActor-Critic法³⁾を用いている。このActorとCriticには時系列データを学習することができるCTRNNを用いる。CTRNNはコンテキストニューロンによるフィードバック結合を持っており、この値は学習時に適切な内部状態を表現するように自己組織的に決定される。従って一つのCTRNNにActorとCriticを組み込むこととした。

強化学習は試行錯誤を通して徐々に方策を改善してゆく学習方法であり、一回の試行ごとに追加学習を行う必要がある。そこで、我々の提案するモデルではBGの短期記憶の役割に相当する部分として、複数の試行における運動軌道を記憶しておくデータベースを用いる。CTRNNの学習を行う際に、このデータベースに記憶された複数の試行のデータをあわせて学習することとした。また、各試行の中で得られる報酬の合計値を計算し、過去の試行で得られた最大の報酬値と比較して優れているもののみをデータベースに記憶させている。

2.2 IPL-M1S1 Model

M1S1に相当する部分の学習器には非線形な時系列パターンの学習を行うことができるニューラルネットであるRNNを用いる。RNNの入力には、ロボットアームの状態を表す情報として、現在の関節角度と回転速度を用意し、指令値として目標回転速度を入力する。また、RNNの出力は、目標回転速度を実現するために必要な、ロボットアームの各関節を駆動するトルクとした。RNNの学習には、上位のBG-IPLモデルの探索行動によって得られた運動軌道データを用いる。この軌道を生成する際に目標回転角度は与えられないため、学習に必要な入力情報が不足するが、学習はオフラインで行われるため、学習時には1タイムステップ先の将来の回転速度を目標回転速度の代わりに用いた。RNNの学習アルゴリズムはBackPropagationThroughTime(BPTT)法を用いる。

3. 実機ロボットによる実験

以上のモデルを用いて、本研究では図3に示す実機ロボットで対象物へのリーチング運動の学習を行った。このタスクでは天井に設置したカメラによりロボットの手先位置と対象物（赤いボール）との距離を測定し、その距離が近いほど多くの報酬が得られるように設定した。また、上位・下位のニューラルネットは共にランダムに初期化した状態から学習を行っており、これらのニューラルネットは同時に学習を行っている。具体的には、上位のニューラルネットは強化学習により多くの報酬が得られる運動軌道を学習し、下位のニューラルネットはそのときの運動情報を元に、探索範囲付近でのロボットアームの動かし方を徐々に学習してゆく。このときの実験結果として、強化学習の各試行において得られた報酬の合計値を図4に示す。この例では、学習初期から手とボールの距離が近いため報酬の値が大きくなっているが、その後ボールが大きく動き、下位のニューラルネットが未学習の領域に移動してしまっても徐々に報酬値が増加し、再びボールに触る事が出来ていると確認できる。

Fig.3 robot experiment

Fig.4 reward curve

4. まとめ

本研究では、脳科学の知見に基づき、知覚情報と運動指令に関する高次機能を担う頭頂葉及び大脳基底核と、身体の動力学的特性に関する低次機能を担う第一次運動野及び第一次感覚野を対象に、前者は強化学習を行う人工神経回路網として、後者は自己の身体モデルを学習する人工神経回路網としてモデル化し、それらの並行的な学習を通じて、身体モデルを獲得した上で、目標運動軌道を生成し得るのか検証を行った。実験の結果、学習初期においては自己の身体モデルの学習が未熟なため所望の運動軌道を生成できないのに対し、学習後期においては、身体モデルの習熟により高次の運動指令に従った運動軌道が生成されることが示唆された。

謝辞

本研究の一部は、文部科学省科学研究費補助金特定領域研究「身体・脳・環境の相互作用による適応的運動機能の発現－移動知の構成論的理解－」によって行われた。ここに謝意を表す。

参考文献

- 1) H.Ehrsson, A.Fagergren, R.Johansson, H.Forssberg, "Evidence for the Involvement of the Posterior Parietal Cortex in Coordination of Fingertip Forces for Grasp Stability in Manipulation", Journal of Neurophysiology, 90:2978-2986, (2003).
- 2) L.Fogassi, P.F.Ferrari, B.Gesierich B, S.Rozzi, F.Chersi F. and G.Rizzolatti, "Parietal lobe: from action organization to intention understanding", Science, 308:662-667, 2005.
- 3) Hiroaki Arie, Tetsuya Ogata, Jun Tani, Shigeki Sugano, "Reinforcement learning of a continuous motor sequence with hidden states", Advanced Robotics, vol.21, no.10, pp1215-1229, 2007