

ロボット-物体インタラクションの予測学習と予測不確実性の最小化による動作生成 Learning of Robot-Object Interaction and Motion Generation by Minimizing Predictive Uncertainty

5116E023-3 山川まどか

指導教員 尾形哲也 教授

YAMAKAWA Madoka

Prof. OGATA Tetsuya

概要：本研究では、動物の行動決定について動作生成モデルを提案し、実験によって振る舞いを観察する。先行研究では行動決定には予測が重要な要因であるということが示されている、本研究では Stochastic-recurrent neural network(S-RNN)を用い、agent の動作と環境変化の予測および予測不確実性を扱う行動生成モデルを提案する。提案モデルでは agent と物体のインタラクションを学習し、この経験に基づいて予測不確実性の最小化により動作を生成する。実験ではロボットがラグビーボールを横から押して転がす動作をタスクとする。この動作と物体挙動を S-RNN を用いて学習し、学習後ネットワークを使って動作生成を行う。この結果、提案モデルはボールの物理的特性、ロボットの身体的特性、システムの経験不足に応じて予測不確実性を評価し動作生成できることがわかった。

Keywords: Motion Generation, Affordance, Free Energy, Stochastic-Recurrent Neural Network

1. 研究背景

動物の行動はどのように決定されるのだろうか？たとえば、我々がコップを掴むとき、コップのどこを持つとか関節の動き方とかを事前に計画して実行しているとは考えにくい。行動は直感的に瞬時に実行されるのである。Fristonによるとagentは自己の内部状態から推測される外部(環境)状態と、センサ情報と外部状態の経験から生成される推測が一致すると期待される行動をとる[1]。ここでは、「agentはセンサ情報から推測される最も予測のしやすい行動をとる」という仮説のもと行動生成モデルを立てその振る舞いを検証する。

2. 自由エネルギー原理(FEP)

FEPではagentは無秩序へ向かう自然的な傾向に抗って存続し続けるために自由エネルギー(FE)を最小化しなければならない。FEは、システムの外部状態の推測と、感覚入力及びそれを引き起こす外部状態についての知識を確率分布で表現したときの距離である。自由エネルギー最小化は(1)システムの内部状態(推測)を変える、(2)行動によって感覚入力を変える、ことで実現される。(2)においてagentは将来FEを下げると期待される行動を選択する。つまり、agentはこれから行う行動によって引き起こされる外部状態の不確実性を仮想している。本研究ではこの仮想的な不確実性の予測が行動を決定しているという知見を取り入れ、予測不確実性の最小化による動作生成をモデル化する。

3. 予測不確実性の最小化による動作生成

外部環境、感覚入力、agentの行動とそれらの予測と予測不確実性を実装するモデルとして Stochastic- Recurrent Neural Network(S-RNN)[2]を用いる。S-RNNは時系列予測学習器で、ある時刻tの入力に対して単位時刻後の予測と予測の分散を出力する。内部に前時刻からの情報を引き継いで文脈を保持する再帰結合構造があるため、時刻t=0の初期状態(IS)はそのデータを特徴付ける値である。

提案手法では、(1)物体に対する動作と物体挙動を学習データとしてS-RNNを最適化、(2)最適化されたS-RNNのISを予測分散が最小となるよう探索する。この手法で探索されたISはシステ

ムの経験から最も予測不確実性が小さいと予測される動作を生成する。

ここで、システムが予測する不確実性には(a)外部環境そのものの持つ不確実性、(b)システムの知識不足による予測困難性、(c)agentのもつ身体的特性による不安定性の3点が考えられる。実験では、動作生成モデルのこれらの不確実性に対する振る舞いを観察する。

4. 動作生成実験

実験タスクとしてテーブルに置かれたラグビーボールを横から押して転がす動作を設定する。ラグビーボールは楕円形であるため横から押した場合、中心部を押すとまっすぐ転がり挙動が安定するが、端を押すと回転を伴って移動し挙動が不安定である。実験では物理シミュレータ gazebo[3]とロボット baxter[4]モデルを用いた。

学習データとして表1に示す3種類のセットを用意した。それぞれ時間長は100step、次元数は腕自由度7とボール位置情報3の計10次元である。ボール位置情報はシミュレータ内のx,y座標と長軸回転角である。図1に示すボールの初期位置は動作を始める前に置かれるボールの位置を表す。ボールの押し方は、ボールの中心と端を長軸方向に±1[cm]のノイズを持って押す場合と、長軸方向に1.5[cm]ずつ断続的に打点をとって押す場合を設定する。

表1: データセットの内容

| | ボール初期位置 | 押し方 |
|-----------|---------|------|
| Dataset A | 0 | 中心と端 |
| Dataset B | 0 | 断続的 |
| Dataset C | 0, 1, 2 | 断続的 |

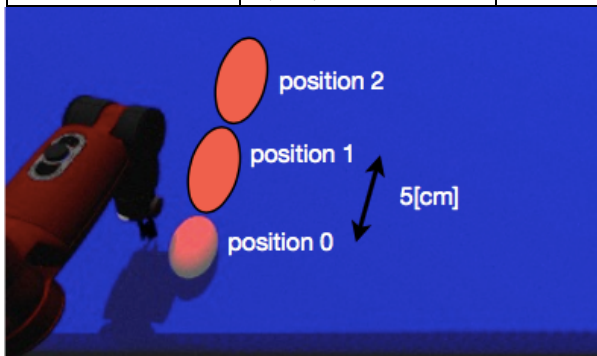


図1: ボール初期位置

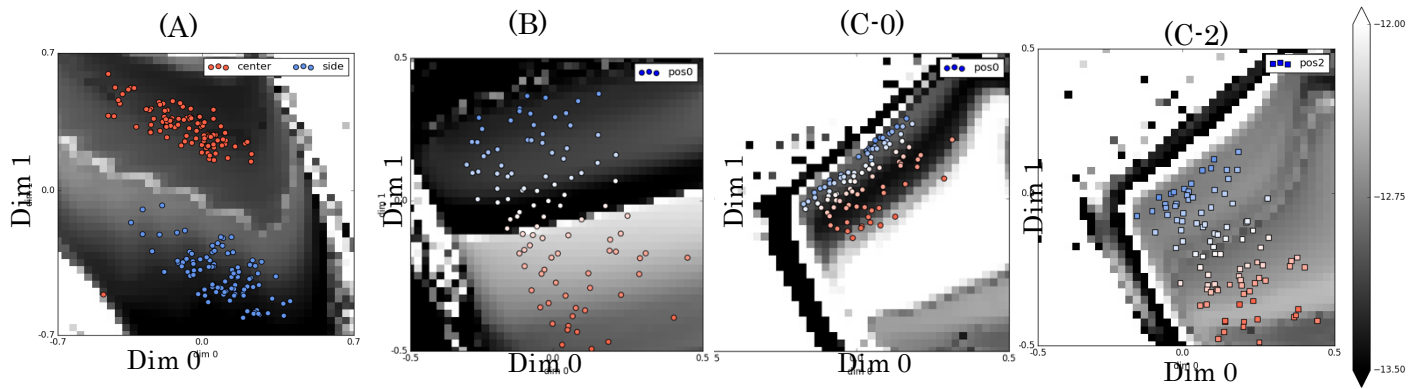


図 2 : IS 空間マップ

5. 実験結果と考察

図 2 に学習で得られた S-RNN の IS と、IS 空間を予測分散で色付けしたものを表す。学習された IS は押し方ごとに色分けされており、(A)では中心と端でそれぞれ赤、青に分かれ、(C)ではロボットの手前側～中心～奥側の順に青～白～赤のグラデーションになっている。空間の色マップは、IS 空間上にグリッドをとり、その格子上の IS を入力したときの S-RNN の予測する分散を色で示したものである。予測分散値が高いほど黄色く低いほど青い。(A)は Dataset A, (B)は Dataset B, (C)は Dataset C をそれぞれ学習した際のマップである。また、(C-0)はボールの初期位置 0, (C-2)は初期位置 2 を初期状態として入力した際のマップである。

ラグビーボールの物理特性をリャブノフ指数の観点から解析すると、中心を押す場合長軸回転角が安定し、端を押す場合 x, y 座標が安定する。これは端を押す場合、ボールはほとんど移動せずその場に止まったまま回転だけするためである。図 2 の (B)ではボールの中心付近を押す IS 付近を底として予測分散値の谷ができており、ラグビーボールの物理特性と一致する。この結果は外部環境そのものに含まれる物理的な予測不確実性を反映している。

(A)のマップでは中心を押す動作の IS 群と端を押す動作の IS 群の間に分散値の山ができていいる。この山はボールの中心と端の間も押すデータが含まれる Dataset B および C を学習した場合には発生しない。Dataset A にはボールの中心または端を押す動作のみでその間を押すデータは含まれていない。そのため経験が不足しており予測自体が困難であるため、予測不確実性が大きくなる。

(C)の 2 つの図を比べると学習された IS のまとめがボールの初期位置が奥になるにつれて広がっていくのがわかる。予測分散値の谷は同じようにボール中心を押す IS 付近を底としてできているが、底の最低値はボール位置が奥になるにつれて高くなっている。これは、ボールの位置が奥になるとそこへ到達する姿勢をとるロボットアームの可操作度が小さくなり、肩関節のぶれが大

きく手先位置に影響することに起因する。目標打点は同じでも実際にボールを押す位置はずれてしまうので、ぶれに応じて分散値は大きく予測されている。

提案手法によって求めた IS によって生成される動作をシミュレータ内で動かして確認すると、どの条件でもボールの中心を押す動作が得られた。

6. まとめ

本研究では動物の直感的な動作の生成モデルを立てその挙動を確認した。このモデルは自身の行動によって引き起こされる環境の変化が最も安定するように動作を選択する。動作生成の手法として、S-RNN による動作と物体挙動の学習と IS の予測分散最小化による探索を提案した。実験ではラグビーボールを押す動作をタスクとし、提案手法によりボールの中心部を押す動作を生成した。この動作はラグビーボールの物理特性上挙動が安定することを示した。また、予測不確実性がどのようにモデルに反映されているか、IS 空間の解析によって示した。本研究では動作の選択肢をあらかじめ設計していたが、今後の研究では探索的に動作を獲得していくモデルを検討したい。

参考文献

- [1] Friston K, Adams RA, Perrinet L, Breakspear M. (2012) Perceptions as hypotheses: saccades as experiments. *Front Psychol.* 3:151.
- [2] S. Murata, J. Namikawa, H. Arie, S. Sugano, and J. Tani, "Learning to Reproduce Fluctuating Time Series by Inferring Their Time-Dependent Stochastic Properties: Application in Robot Learning via Tutoring," *IEEE Transactions on Autonomous Mental Development*, vol.5, issue 4, pp.298-310, 2013.
- [3] N. Koenig and A. Howard. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2149-2154, Sendai, Japan, Sep 2004.
- [4] "Baxter Simulator", 2018/1/22, http://sdk.rethinkrobotics.com/wiki/Baxter_Simulator