

深層学習を用いたロボット動作生成におけるアクションラベルの活用

○内海 力郎¹ 加瀬 敬唯¹ 尾形 哲也^{1,2}

¹ 早稲田大学基幹理工学研究科表現工学専攻 ² 産業技術総合研究所

1 はじめに

近年、ロボットの動作生成において人間の知識や経験を活用するための様々な手法が研究されている [1]. 人間の知識を用いてロボットが直接動作の生成を行うことができれば、社会におけるロボット導入のハードルが大きく下がることが期待できる. 一方で、ロボットの動作生成に人間が付与 (アノテーション) した知識を活用する研究は、アノテーションの種類からその与え方まで適切な手法を模索している段階にある. そこで本研究ではロボットの行動を人間が主観的にアノテーションしたアクションラベルを深層学習を用いたロボット動作生成に活用するための手法を提案する. 提案手法ではアクションラベルを直接入力として使用するのではなく、ラベル推測モデルを動作生成モデルと共に学習した. 実験の結果、人間の与えたアクションラベルがロボットの動作生成誤差の減少に寄与することを確認した.

2 関連研究

人間の知識をロボットの動作生成に活用する手法は使用モデルによって多様な手法が提案されている. 近年では深層学習において人間の知識を活用する事例が多く、ロボットの行動を人間の行動として説明することで、人間タスク動作から直接ロボットの動作を生成する研究 [2] も行われている. アクションラベルのような簡単なアノテーションの活用では、強化学習の学習過程で人間による評価ラベルを活用する手法 [3] が知られている. しかし、これらの研究は学習の補助としてラベルを活用しており、そのほかのセンサ情報と同様に入力として扱った事例は確認できなかった. これはセンサ情報が連続的かつ確実なものであるのに対し、ラベル情報が非連続的かつ不確実であることが一つの理由として考えられる. 本研究では深層学習を用いたロボット動作生成手法の一つである模倣学習においてアクションラベルを活用する手法を提案する. 提案手法ではラベルをモデル内で連続的なベクトル表現に変換し、そのほかのセンサ情報と並列的に扱っている.

3 提案手法

本研究で使用する動作生成システムの概要図を図 1 に示す. 動作生成には時系列情報を扱うのに適した Long Short

Term Memory(LSTM) モデルを使用する. LSTM は内部情報と時刻 t における視覚運動情報からロボットの時刻 ($t+1$) の姿勢を予測し、ロボットは予測された姿勢に移動することを繰り返すことによって一連の動作を実現する. ロボットの視覚画像情報は事前に学習された Convolutional Auto Encoder(CAE) を用いて次元圧縮することで学習効率を向上させている.

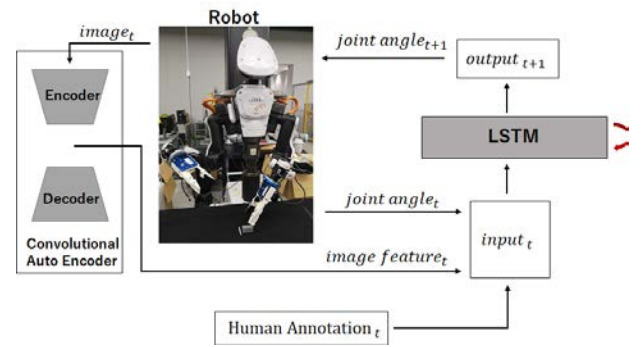


図 1 動作生成システム

比較するモデルを図 2 に示す. ラベルの有効性の確認のため、ラベルを入力として使用しない LSTM モデル, 正解ラベルを直接入力として使用する Teacher Label LSTM(TL-LSTM) モデル, 下位モデルで連続的なベクトル表現としてラベルを変換して入力する Cross Entropy Label LSTM(CEL-LSTM) モデルの 3 種類を比較する. 3 つのモデルは事前に学習データの予測誤差を用いて学習される. TL-LSTM モデルでは人間が与えたアクションラベルを one-hot-vector に変換し、視覚運動情報と同様に入力情報として扱う. CEL-LSTM モデルではラベルを直接入力とするのではなく、下位モデルで推測したラベル情報を入力として使用する. 下位モデルはラベルを連続的に扱うため、動作生成と同様に LSTM モデルを使用し、学習時には教師ラベル T と推測ラベル P を用いて式 3 を損失関数として学習を行う.

$$L_{P,T} = -T_i \log P_i - (1 - T_i) \log (1 - P_i) \\ \hat{i} = \arg \max_i P_i \quad (3.1)$$

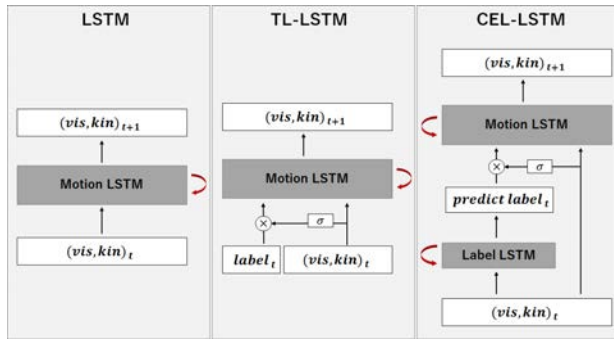


図2 比較モデル

4 実験と結果

4.1 タスクデザイン

長さの異なる複数の立方体を異なる初期位置からゴール位置に移動させるタスクデータ (図3) を用いて各モデルの予測誤差の比較を行う。立方体の長さは0.5[cm] 刻みで4.5～8.5[cm] の9種類あり、このうち5種類をトレーニングオブジェクト、4種類をテストオブジェクトとして用いる。物体ごとに初期位置を0.5[cm] 刻みで13か所から移動させ、計116シーケンスのデータが撮影された。

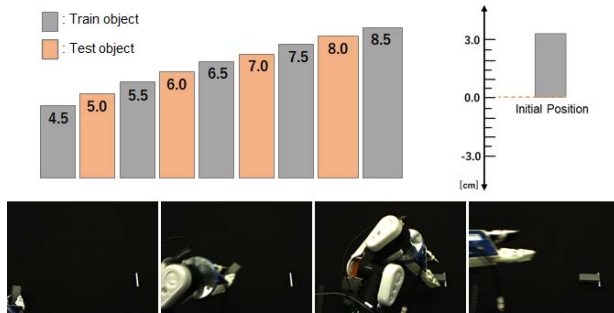


図3 使用データセット

ロボットには産業用双腕型ロボットであるNextage, エンドエフェクタには開閉のみ可能な1次元グリッパーであるEZGripperを使用した。姿勢情報は左右の手先位置, 姿勢, グリッパーの14次元からなり, 頭部カメラから撮影された128×128のRGB画像と共に10Hzで記録されている。また, 全てのデータにはフレームごとに表1のいずれかのアクションラベルが与えられている。

Action number	Description
1	待機状態
2	物体位置への移動
3	手を下げる
4	物体を持ち上げる
5	目的位置へ移動
6	物体を置く
7	初期位置へ戻る

表1 与えられたアクションラベル

4.2 実験結果

3つのモデルの動作生成誤差の比較を図4に示す。平均値で見ただけの場合, アクションラベルを直接入力として扱ったTL-LSTMモデルはLSTMモデルと大きな差は無いものの, 下位モデルで共同学習されたラベル表現を入力として用いるCEL-LSTMモデルは他の2つのモデルに比べて動作生成誤差が小さくなることが確認された。このときCEL-LSTMモデルのラベル推測出力は人間の与えたラベルと境界の異なるベクトル表現として獲得されていた。このことから人間の与えたラベルを都合の良い多次元情報として変換することが動作生成上有利に働くことが推測できる。

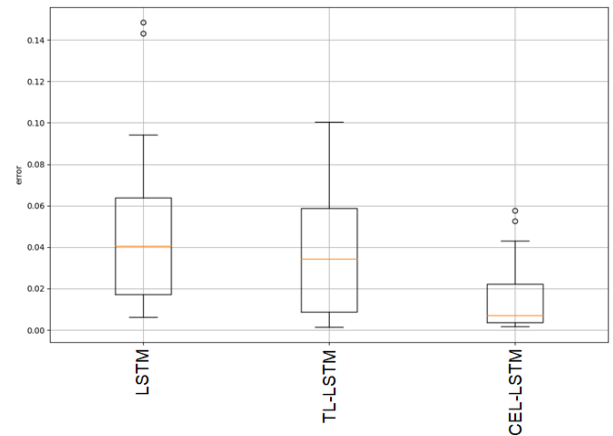


図4 モデルごとの動作生成誤差の比較

5 まとめ

本研究では模倣学習を用いたロボット動作生成において人間の与えたアクションラベルを活用する方法を提案した。実験の結果, アクションラベルは直接入力として用いた場合には動作生成誤差に大きな変化がなかった。一方で, 提案手法であるCEL-LSTMモデルではラベルを連続的なベクトル表現として共に学習することで動作生成誤差を減少させることができた。今後は実際にロボットを動かした場合の物体位置精度を用いて比較実験を行う予定である。

参考文献

- [1] Celeste Mason, Konrad Gadzicki, Moritz Meier, Florian Ahrens, Thorsten Kluss, Jaime Maldonado, Felix Putze, Thorsten Fehr, Christoph Zetzsche, Manfred Herrmann, Kerstin Schill, and Tanja Schultz. From human to robot everyday activity. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 8997–9004, 2020.
- [2] Tianhe Yu, Pieter Abbeel, Sergey Levine, and Chelsea Finn. One-shot composition of vision-based skills from demonstration. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2643–2650, 2019.
- [3] Anis Najjar and Mohamed Chetouani. Reinforcement learning with human advice: A survey. *Frontiers in Robotics and AI*, Vol. 8, p. 74, 2021.