

平成 31 年度

卒業論文

複雑環境下でのロボット学習に向けた
深層状態空間モデルを用いた映像予測

平成 32 年 2 月

指導教員 松尾豊教授

東京大学工学部システム創成学科

知能社会システムコース

03-180961 近藤生也

概要

目次

概要	i
第 1 章 序論	1
1.1 本研究の背景	1
1.1.1 行動条件付き映像予測	1
1.1.2 既存の行動条件付き映像予測手法の問題点	2
1.2 本研究の目的	3
参考文献	4

図目次

表目次

第 1 章

序論

1.1 本研究の背景

1.1.1 行動条件付き映像予測

多様な環境で様々なタスクが遂行可能な汎用的なロボット（generalist robots）の開発はロボット工学の最重要課題の一つである。ロボットハードウェアの低価格化，汎用的なロボットソフトウェアの普及に加え，近年の急速な深層学習技術の発展を受けてロボットの制御方策を自ら学習させるロボット学習の研究が進んでおり，ロボットで遂行可能なタスクは着実に増えてきている。

ロボット学習において，将来予測，特に映像予測を明示的に学習することは，① ロボット自身が映像予測を用いた方策をたてることが可能になる，② 映像予測結果を人が評価することでロボットの行動を予め評価できる，という大きく二つの点から重要であると言える。① の映像予測を用いた方策の例として，Hafner ら [1] は，強化学習の問題設定において明示的に学習した映像予測モデルを用いることで簡単なアルゴリズムで効率的なプランニングが可能であることを示した。② の映像予測を人が評価する例として Ebert らによる研究 [2] では学習した映像予測モデルを用いて，ロボットの操作によって予想される物体の移動の軌跡を確率分布として出力することができ，これを用いて人がロボットの行動の正しさを予め判断することができる。

このようにロボット学習における映像予測は重要であるが，映像予測だけを切り取って研究されることも多い。映像予測の中でも，ロボットの行動の結果として観測される映像を予測する問題設定を行動条件付き映像予測と呼び，様々な研究がなされてきている。近年高精度な行動条件付き映像予測手法がいくつか提案されており，ロボット学習研究で扱うタスクの高度化

を背景にしてこれらの映像予測手法をより複雑な問題設定に対して適用していきたいと考えられているが、いくつかの研究で既存の行動条件付き映像予測は上手く機能しない可能性があることがわかってきている。ただしここでいう「より複雑な問題設定」とは具体的には環境中に複数の操作対象の物体が隣接し合っていてある場合、操作対象が布などの非剛体物である場合などを想定している。次に既存の行動条件付き映像予測手法の問題点について示す。

1.1.2 既存の行動条件付き映像予測手法の問題点

行動条件付き映像予測手法は大きく再帰型ニューラルネットワーク (RNN) ベースの手法と深層状態空間モデル (DSSM) ベースの手法に分けられる。Hafner ら [1] の研究など深層強化学習の問題で映像予測を明示的に行う場合は後者の DSSM ベースの手法が多く採用されるが、映像予測の問題では前者の RNN ベースの手法が多く使われている [3][4]。

RNN ベースの手法は予測した 1 ステップ先の画像を入力にして更に 1 ステップ先の画像を出力するというような、自らの出力を逐次入力する構造を持つ。RNN ベースの手法は DSSM と比較して高精度な映像を生成に長けている反面、近年提案されている RNN ベースの手法には以下のような問題点がある。

- 「確率的な遷移を考慮できない」、は確率的な遷移 + 自己回帰な論文もあるので入れませんでした
- 誤差が蓄積しやすい
- 文脈を必要とする

一点目について、RNN ベースのモデルを用いると常に直前のフレームを参照して次のフレームを予測するため短い期間の予測であれば精度は高くなるが、予測誤差が蓄積していくために長期の予測には向かないことが示されている [1]。

二点目について、RNN はモデルの内部状態を十分に更新した後でないと適切に予測が行えず、予測を始める前に文脈としてそれより前の数フレームを与える必要があり、この文脈として与えられるフレーム数が少ないと予測が悪化することが知られている [4]。ロボット実機への応用を考えた場合、現在の状態から未来を予測する際に文脈を得るために先に数ステップ行動することは、予測してから行動するという目的意識に反しており実用的できでない。このため RNN ベースの手法をロボット実機に応用する際には、文脈として現在の観測という 1 フレームのみ与れば十分機能するように改善する必要がある。このように、RNN ベースの手法は制約がありそもそも実ロボットへの応用に向いていない可能性がある。

一方, DSSM は各時刻の状態をベクトル (状態ベクトル) で表現し, 毎時刻ロボットの行動によって状態ベクトルが遷移し, その時刻に観測される画像は状態ベクトルからの写像であると考えて遷移モデルと写像のモデルを学習する. DSSM は強化学習の分野で長期の予測にも用いられているなど安定した未来の予測に長けているが RNN と比較して画像の生成時に直前の画像を用いないことから高精度な生成は難しく, また映像生成自体を目的にして DSSM を用いた研究は現状少なくなっている.

1.2 本研究の目的

これらの研究背景を踏まえ, 行動条件付き映像予測の問題をより複雑な問題設定にスケールさせることを目指し, 本研究では特に実ロボットへの応用を重要視して DSSM ベースの行動条件付き映像予測に取り組む. まず DSSM で高精度な生成が難しいことを確認しその理由を簡単に考察する. その上で特に複雑な問題設定に広く取り入れることが可能な帰納バイアスを提案しモデルに組み込むことで DSSM を拡張, DSSM を使った際にも高精度な生成を可能にすることを目指す.

東京大学工学工学部システム創成学科

知能社会システムコース

松尾研究室学部 4 年

平成 31 年 2 月 近藤生也

参考文献

- [1] Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In *International Conference on Machine Learning*, pp. 2555–2565, 2019.
- [2] Frederik Ebert, Chelsea Finn, Sudeep Dasari, Annie Xie, Alex Lee, and Sergey Levine. Visual foresight: Model-based deep reinforcement learning for vision-based robotic control, 2018.
- [3] Emily Denton and Rob Fergus. Stochastic video generation with a learned prior, 2018.
- [4] Ruben Villegas, Arkanath Pathak, Harini Kannan, Dumitru Erhan, Quoc V. Le, and Honglak Lee. High fidelity video prediction with large stochastic recurrent neural networks, 2019.