

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

近年、深層ニューラルネットワーク (DNN) を使用した機械学習手法の発展を背景とし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いる。ロボット学習において未来の観測、特に自らの行動に対するフィードバックを適切に予測するよう学習することは、環境の遷移モデル (内部モデル) を獲得すること、そしてそれを用いたプランニングや行動結果の予測、安全性の評価などに欠かせない。時間変化する環境の DNN を用いたモデルとして深層状態空間モデル (Deep State Space Model) があり、これは昨今深層強化学習の分野で取り入れられているが、深層状態空間モデルを用いて映像予測を学習することで内部モデルが獲得できるとされている [1]。しかし現在一般的に使われて深層状態空間モデルとその学習方法は各時刻の状態が低次元のベクトルで表現できるという仮定を暗に置いており、高次元な表現を仮定すると学習が進まず、複雑な環境の予測問題にスケールさせる方法は自明でない。

本研究ではまずこのような問題があることを実験的に指摘した上で、状態の表現に階層性を考えることで深層状態空間モデルにより高次元な状態表現を仮定できるように拡張し、低次元な表現を仮定した場合よりも良い表現が獲得できることを示す。評価実験では様々な物体が置かれた机上でのマニピュレーションを題材にしたデータセットを用いてロボットの行動で条件付けた映像予測を行う。単純な深層状態空間モデルをベースラインに設定し、提案手法の定性的・定量的な有効性を示す。

# 目次

概要	i
第 1 章 序論	1
1.1 本研究の背景 . . . . .	1
1.1.1 ロボット学習 . . . . .	1
1.1.2 状態表現学習 . . . . .	1
1.1.3 映像予測 . . . . .	1
1.2 本論文の構成 . . . . .	1
第 2 章 提案手法	2
2.1 問題設定 . . . . .	2
2.2 確率モデル・最適化 . . . . .	3
2.3 モデルアーキテクチャ . . . . .	4
2.3.1 推論ネットワーク . . . . .	4
2.3.2 生成ネットワーク . . . . .	4
2.4 類似手法との差分の整理 . . . . .	5
第 3 章 まとめ	7
謝辞	8
参考文献	9

# 図目次

2.1	提案手法のグラフィカルモデル . . . . .	3
2.2	推論ネットワークで用いる CNN([3] より引用) . . . . .	5

# 表目次

2.1	提案手法と類似手法の比較 . . . . .	6
-----	------------------------	---

# 第 1 章

## 序論

### 1.1 本研究の背景

#### 1.1.1 ロボット学習

多様な環境でさまざまなタスクを実行することのできる，汎用的なロボット（generalist robots）（Escudero Rodrigo et al., 2015; Finn, Yu, et al., 2017）の開発は，ロボット工学の最重要課題の一つである．

#### 1.1.2 状態表現学習

DNN は特徴抽出が得意．表現学習と呼ばれる．時間変化する環境に拡張した表現学習のことで、各時刻の生の観測データやその系列からこの状態表現（特徴）を見つけることが目的となる。

#### 1.1.3 映像予測

上記のような強化学習や実機タスクへの応用から（行動条件付き）映像予測だけを切り取った研究も多い．

提案手法の有効性について定性的・定量的な評価を行う．

最後に実験結果を踏まえて，今後の課題と社会応用について述べる．

### 1.2 本論文の構成

本論文の構成は以下のとおりである．

## 第 2 章

# 提案手法

本章では，生成クエリネットワークのメタ学習としての解釈に関する前章での議論を踏まえて，改善手法の提案を行う．前章で述べた通り，既存手法である生成クエリネットワークはシーン固有の知識をもつ変数  $\mathbf{r}_i$  をコンテキスト  $D_i$  から決定論的に推論しているため，コンテキストから推論可能なシーン表現の不確実性を考慮できていない．また，個別のデータに固有の潜在変数  $\mathbf{z}_i^k$  と，その事後分布  $p(\mathbf{z}_i^q | \mathbf{x}_i^q, \mathbf{v}_i^q, \mathbf{r}_i; \theta)$  を近似する  $q(\mathbf{z}_i^q | \mathbf{x}_i^q, \mathbf{v}_i^q, \mathbf{r}_i; \phi)$  の存在により，最適化すべきパラメータが増大してしまっているという問題点がある．そこで，提案手法では，生成クエリネットワークにおけるシーン表現  $\mathbf{r}_i$  と潜在変数  $\mathbf{z}_i^k$  を統合して，1つの確率的な潜在変数  $\mathbf{z}_i$  とすることで，モデルの冗長性を排除し，シーン表現の確率的な推論を可能にすることを目指す．

はじめに，提案手法で扱う問題設定について整理し，その後提案手法の具体的な説明として，確率モデル・最適化とモデルアーキテクチャについて述べていく．

### 2.1 問題設定

提案手法では生成クエリネットワークと同様の問題設定を対象とする．モデルは多数あるシーンの中からランダムにサンプリングされたあるシーン  $i$  について，いくつかの視点座標と観測画像のペア群  $D_i = \{(\mathbf{v}_i^k, \mathbf{x}_i^k)\}_{k=1}^M$  をコンテキストとして受け取り，さらに別の視点座標  $\mathbf{v}_i^q$  をクエリとして入力される．そしてモデルはクエリに対応する観測画像  $\mathbf{x}_i^q$  を予測して出力する．今回はモデルの評価指標として，3つのコンテキストが与えられた状況におけるクエリ画像  $\mathbf{x}_i^q$  の予測精度を用いる．これは，コンテキストが3つ与えられた状況においては，シーン内の物体の性質（色・形状など）はほぼ一意に定まるため，モデルが正しく学習されていればクエリ画像の完全な予測が可能であると考えられるからである．

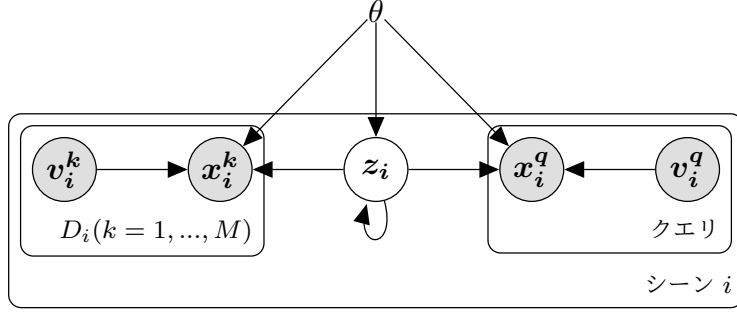


Fig. 2.1 提案手法のグラフィカルモデル

## 2.2 確率モデル・最適化

ここからは、提案手法で用いるモデルの詳細について述べる．まずデータの生成過程を Fig. 2.1 のようなグラフィカルモデルで表現される確率モデルで定義する．既存手法の生成クエリネットワークでは、観測不可能な潜在変数として  $\mathbf{r}_i$  と  $\mathbf{z}_i^k$  の 2 つを定義していたのに対し、提案手法ではシーン固有の知識をもつ 1 つの確率的な潜在変数  $\mathbf{z}_i$  のみを用いて、観測データの生成過程を表現している．ただし、グラフィカルモデル上のノード  $\theta$  はニューラルネットワークで定義されるモデルのパラメータを表しており、変数ではない点に注意されたい．このように生成過程をモデル化すると、提案手法で最大化する目的関数は以下のような変分下限となる．

$$\log p(\mathbf{x}_i^q | \mathbf{v}_i^q, D_i; \theta) = \log \int p(\mathbf{x}_i^q | \mathbf{v}_i^q, \mathbf{z}_i; \theta) p(\mathbf{z}_i | D_i; \theta) d\mathbf{z}_i \quad (2.1)$$

$$\simeq \log \int p(\mathbf{x}_i^q | \mathbf{v}_i^q, \mathbf{z}_i; \theta) q(\mathbf{z}_i | D_i; \phi) d\mathbf{z}_i \quad (2.2)$$

$$\geq \mathbb{E}_{\mathbf{z}_i \sim q(\mathbf{z}_i | D_i; \phi)} [\log p(\mathbf{x}_i^q | \mathbf{v}_i^q, \mathbf{z}_i; \theta)] \quad (2.3)$$

$$q(\mathbf{z} | D; \phi) = \prod_{l=1}^L q_l(\mathbf{z}_l | \mathbf{z}_{<l}, D; \phi) \quad (2.4)$$

それぞれの分布  $q_l(\mathbf{z}_l | \mathbf{z}_{<l}, D; \phi)$  は、正規分布に従うと仮定し、その母数 (平均, 標準偏差) を出力するニューラルネットワークで定義する．

$p(\mathbf{x}^q | \mathbf{v}^q, D; \theta)$  は推論分布からサンプリングされた潜在変数  $\mathbf{z} = \{\mathbf{z}_l\}_{l=1}^L$  とクエリ  $\mathbf{v}^q$  を入力として、クエリ画像  $\mathbf{x}^q$  を生成する分布であり、これを生成分布と呼ぶ．ここでは、生成分



布は標準偏差 1 の正規分布  $\mathcal{N}(\mu(\mathbf{v}^q, \mathbf{z}), \mathbf{I})$  に従うと仮定し、その尤度を計算することで、式 2.3 の目的関数をモンテカルロ近似する。生成分布もその平均  $\mu(\mathbf{v}^q, \mathbf{z})$  を出力するニューラルネットワークで定義する。

これが提案手法の確率モデルと最適化アルゴリズムの概要である。

## 2.3 モデルアーキテクチャ

本節では前項で定義した確率モデルを表現するニューラルネットワークの具体的なアーキテクチャについて述べていく。提案手法のモデルアーキテクチャは大きく 2 つに分けて考えることができる。1 つはコンテキスト  $D_i = \{(\mathbf{v}_i^k, \mathbf{x}_i^k)\}_{k=1}^M$  を入力として潜在変数  $\mathbf{z}_i$  の分布の母数 (平均, 標準偏差) を出力する推論分布  $q(\mathbf{z}_i | D_i; \phi)$  のネットワーク (推論ネットワーク) であり, もう 1 つが推論分布からサンプリングされた潜在変数  $\mathbf{z}_i$  とクエリ  $\mathbf{v}_i^q$  を入力として, クエリに対応する観測画像の予測  $\hat{\mathbf{x}}_i^q$  を出力する生成分布  $p(\mathbf{x}_i^q | \mathbf{v}_i^q, \mathbf{z}_i; \theta)$  のネットワーク (生成ネットワーク) である。ここから, それぞれのネットワークのアーキテクチャについて説明する。

### 2.3.1 推論ネットワーク

推論ネットワークには, 既存手法で表現ネットワークとして用いられているアーキテクチャをベースに, 確率的な潜在変数  $\mathbf{z}_i$  を出力するための変更を加えたものを使用する。まず, コンテキスト  $D_i = \{(\mathbf{v}_i^k, \mathbf{x}_i^k)\}_{k=1}^M$  を視点と観測のペアごとに分け, それぞれのペア  $(\mathbf{v}_i^k, \mathbf{x}_i^k)$  について, Fig. 2.2 のような 8 つの畳み込みニューラルネットワーク (CNN) を使って圧縮した特徴表現  $\mathbf{r}_i^k$  を作る。ただし, Fig. 2.2 中では, 矢印が畳み込み層と活性化関数 ReLU のブロックを表現しており,  $k, s$  はそれぞれ畳み込み層のカーネルサイズ, ストライド, 「+」はスキップコネクション [2] を意味する。そして,  $M$  個の特徴表現  $\{\mathbf{r}_i^k\}_{k=1}^M$  の要素ごとの和をとったものを  $\mathbf{r}_i$  とする。なお, ここからは簡単のため, シーンの表記  $i$  を省略する。

このようにして, 潜在変数  $\mathbf{z}$  のサンプリングを推論ネットワークによって行う。

### 2.3.2 生成ネットワーク

生成ネットワークは, 推論ネットワークを用いて得られた潜在変数  $\mathbf{z} = \{\mathbf{z}_l\}_{l=1}^L$  とクエリである  $\mathbf{v}^q$  からクエリに対応する観測  $\mathbf{x}^q$  を予測する。ここでも DRAW と同様に畳み込み LSTM と 1 層の畳み込みニューラルネットワークを用いて  $L$  回に分けて画像の生成を行う。

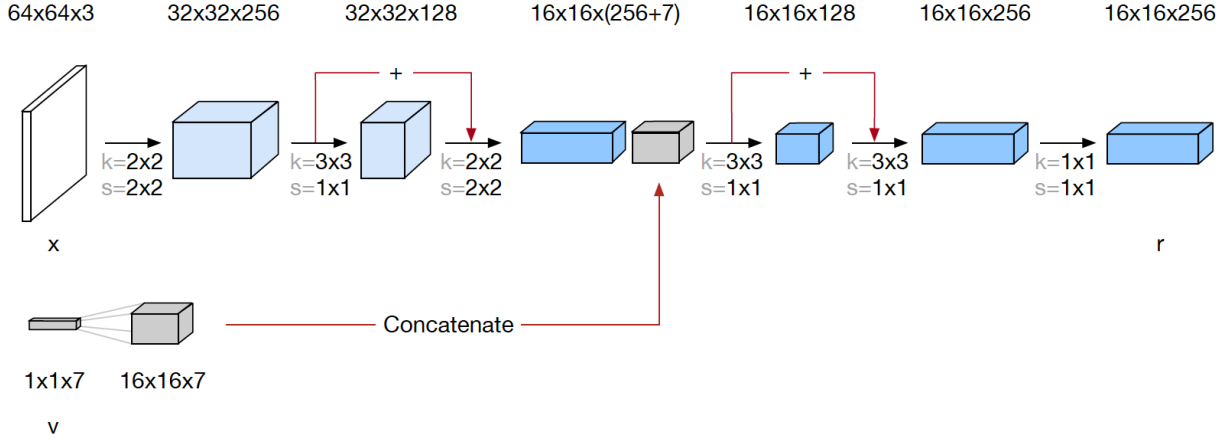


Fig. 2.2 推論ネットワークで用いる CNN([3] より引用)

具体的には、 $v^g$  と  $z_l$  を畳み込み LSTM の入力とし、その出力となる隠れ変数  $h^g$  をさらに 1 層の畳み込みニューラルネットワークへの入力として、その出力を変数  $u_l$  とする．これを  $L$  回繰り返す、最後に  $\{u_l\}_{l=1}^L$  の要素ごとの和を入力とする畳み込みニューラルネットワークの出力を観測画像の予測  $\hat{x}^g$  とする．

このようにして、生成ネットワークは画像の生成を行う．

## 2.4 類似手法との差分の整理

本節では、提案手法と類似手法の差分について、特に確率モデルの観点から整理する．提案手法と同様に Fig. 2.1 のような確率モデルを仮定し、それぞれのデータセットに固有の変数をコンテキストから amortized inference により推論するニューラルネットワークを定義するメタ学習の手法として、Conditional Neural Processes(以下, CNPs)[4] と Neural Processes(以下, NPs)[5] がある．

CNPs は、最適化に用いる目的関数は提案手法と同様であるが、生成クエリネットワークの場合と同様に、データセットに固有の変数  $r_i$  を決定論的に推論しており、不確実性を考慮していない点が提案手法と大きく異なる．

ここまでの内容を整理すると、提案手法と生成クエリネットワーク、CNPs, NPs の関係は Table 2.1 のようになる．

Table 2.1 提案手法と類似手法の比較

手法	データセット固有の変数への推論	データ固有の変数	潜在変数の近似事後分布
生成クエリネットワーク	決定論的	あり	あり
CNP <sub>s</sub>	決定論的	なし	なし
NP <sub>s</sub>	確率的	なし	あり
提案手法	確率的	なし	なし

## 第 3 章

# まとめ

本論文では、はじめに について取り上げ、その問題点として である点を挙げた。そして、その問題点の原因として、 を指摘し、 を行なった。

そこで、本論文では、これらの問題点を解消する改善手法の提案を行なった。提案手法では、 することで、 し、 を可能にした。

実験では、BAIR Push Dataset を用いてベースライン手法を比較し、定性的・定量的な評価を行なった。提案手法は、 においてベースライン手法である を上回る結果となり、提案手法が が実験的にも示された。

最後に考察として、今後の課題や社会応用について議論し、本研究が深層生成モデルを用いて観測情報から環境をモデル化する世界モデルの研究の 1 つとして位置付けられることを述べ、今後の研究の方向性についての展望を整理した。

本研究は、既存の工学的な技術の代替としてのみならず、より汎用的な人工知能技術の達成のために重要な技術の 1 つであると考えており、今後も実環境への提案手法の適用や、時系列遷移を考慮したモデルの考案、強化学習技術との融合などを見据えて、人工知能分野の研究の発展に貢献していきたいと考えている。

# 謝辞

本論文を作成するにあたり，多くの方々にご協力をいただきました．

最後に，配属からの一年間，あらゆる面でサポートをしてくださった松尾研究室の皆様に御礼申しあげて，謝辞とさせていただきます．

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 30 年 2 月 谷口尚平

## 参考文献

- [1] Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Jimenez Rezende, and Daan Wierstra. Draw: A recurrent neural network for image generation. *arXiv preprint arXiv:1502.04623*, 2015.
- [2] SM Ali Eslami, Danilo Jimenez Rezende, Frederic Besse, Fabio Viola, Ari S Morcos, Marta Garnelo, Avraham Ruderman, Andrei A Rusu, Ivo Danihelka, Karol Gregor, et al. Neural scene representation and rendering. *Science*, Vol. 360, No. 6394, pp. 1204–1210, 2018.
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [4] Marta Garnelo, Dan Rosenbaum, Chris J Maddison, Tiago Ramalho, David Saxton, Murray Shanahan, Yee Whye Teh, Danilo J Rezende, and SM Eslami. Conditional neural processes. *arXiv preprint arXiv:1807.01613*, 2018.
- [5] Marta Garnelo, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J Rezende, SM Eslami, and Yee Whye Teh. Neural processes. *arXiv preprint arXiv:1807.01622*, 2018.