

# 入力画像に基づく転移強化学習のモデル選択方式

磯田卓万<sup>†</sup> 野口博史<sup>†</sup> 新井誠亮<sup>†</sup>

NTT ネットワークサービスシステム研究所<sup>†</sup>

## 1. はじめに

Internet of Things (IoT) が急速の拡大しており [1], 将来は大量のセンサで環境のリアルタイム情報を取得し, デバイスを制御するサービスが将来増加すると我々は想定している [2]. センサの取得情報を解析し, 実環境の状況全てでの自律的なデバイス制御が求められる. 制御結果をフィードバックして制御を修正するシステムを構築するために強化学習が研究されている [3]. 本研究では普及するセンサの中で最も多くの割合を占めるカメラの映像を入力情報にデバイスの制御信号を出力するシステムを対象とする. 一例として, 複数のカメラが撮影する領域における掃除ロボット制御システムが考えられる.

## 2. 関連研究

強化学習は多様な環境に応じて制御を調整できる反面, 時間を要する. 学習時間を短縮することでサービス提供までの待ち時間短縮, 計算リソースの効率化による省エネルギーが期待できる.

学習時間を短縮するために, 近い特徴を持つ環境に対し学習済の強化学習モデル転用する転移学習の研究が進んでいる [4]. 学習モデルを別の環境に適用する転移学習において, 不適な学習モデルを用いた場合, 学習時間は短縮しないため, 適切な学習モデルを選択することが課題となる.

## 3. 提案手法

実環境を表す特徴量は膨大であることと, 不適な学習モデルを用いた場合は学習時間の短縮にならないことから, 転用するモデルを選択する際に有効な指標を明らかにする必要があると我々は考えている.

本研究では, カメラの映像情報を用いる強化学習モデルに置いて有効と考えられる指標を用いて, 適切な学習済モデル選択を判断する手法を3つ提案する.

- 画像分類による転移学習モデル選択
- 画像に占める制御対象の大きさによる転移学習モデル選択
- A. および B. の複合

提案手法 A は, CNN [5] を利用した画像分類器の画像類似度スコアを基に最もスコアが高いカメラの学習モデルを転用する手法である. 画像類似度を用いることで, 制御対象だけでなく, 周辺環境の特性を考慮した転移を期待できる. また, 類似度の閾値を用意することで転移学習の要否の判断も実施可能である.

提案手法 B は, カメラ画像における制御対象のデバイスを囲うバウンディングボックスの大きさを基に最もバウンディングボックスの面積が大きいカメラの学習モデルを転用する手法である. 高精度な制御対象の入力画像によって学習効率が向上すると考えている. また, バウンディングボックスの面積に閾値を設けることで, 転移学習の要否の判断も実施可能である.

## 4. 評価実験

3D シミュレーション環境で評価実験を実施した. 移動ロボット1台, ゴールエリア1箇所, カメラ7台を設置する (図1). 移動ロボットは回転, 直進移動をする. 1台のカメラ映像の入力に

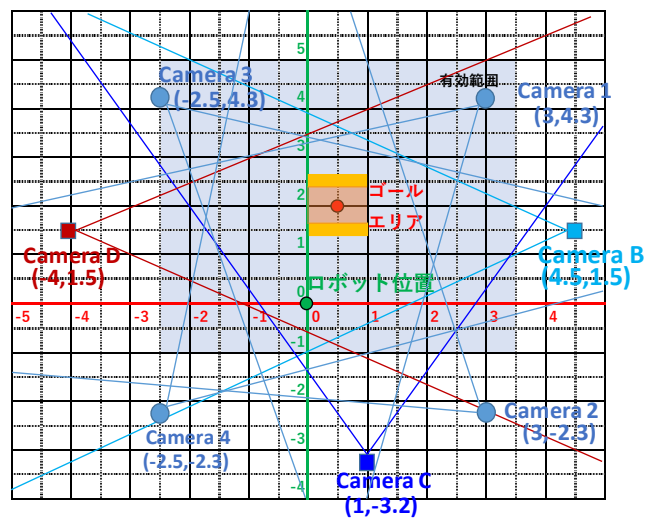


図1. カメラ配置図

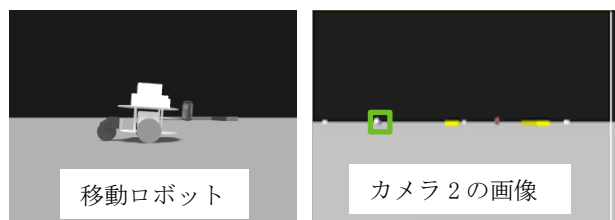


図2. 入力画像例

Model Selection Method for Transfer Reinforcement Learning based on Input Images

Takuma Isoda<sup>†</sup> Hirofumi Noguchi<sup>†</sup> Seisuke Arai<sup>†</sup>  
NTT Network Service Systems Laboratories.

対し、移動ロボットの制御信号を出力する強化学習モデルでゴールエリア到達を目指すシナリオで実験した。入力情報は解像度  $320 \times 240$  (QVGA)、グレースケールの画像とした (図 2)。出力情報は移動ロボットの制御値で、回転速度:  $-0.5 \sim 0.5$  の  $0.1$  刻み、移動速度:  $0.1 \sim 0.5$  の  $0.1$  刻みとした。ロボットは、回転速度の値に従い、回転した後直進移動を行う。それぞれの制御時間は  $5$  秒とした。

強化学習では、1回の試行 (エピソード) に対し  $60$  秒間の時間制限を設ける。指定エリア外・カメラ撮影範囲外への移動を  $-1$  の報酬とし、ゴールエリアへの到達時は残時間を正の報酬とした。2000 エピソードで学習を実施し、得られた報酬総和を出力する。

転移学習を行うために  $7$  台のうち、 $3$  台 (カメラ A, B, C) で学習済モデルを用意する (図 1)。残りの  $4$  台のカメラを用いて、新規学習及び転移学習を実施する。 $4$  台のカメラ (カメラ 1, 2, 3, 4) による新規学習及び  $3$  台の学習済モデルの転移学習の合計  $16$  パターンの実験を実施した。

提案手法 A の画像分類には VGG-16 [5] を利用した。図 1 に示した格子点の全てでロボットを  $90^\circ$  ずつ回転させた  $4$  方向の画像を収集し、学習: 検証 =  $8:2$  の割合で画像を分割し、ファインチューニングで  $1$  時間程度学習した画像分類器を用いた。各カメラの撮影する移動ロボットの初期位置の画像  $1$  枚使って画像類似度スコアを算出した。提案手法 B の物体検出では手動でバウンディングボックスを設置した。 $7$  台のカメラにおいて、移動ロボットの初期位置の画像に対してバウンディングボックスを描画した。実システムにおいては YOLO v3 [6] で数時間程度学習したモデルを用いて自動生成が可能である。

## 5. 結果と考察

提案手法 A においては、カメラ B, C, D 及び新規の平均と比較して、報酬が収束するまでのエピソード数が  $11\%$  減少した。提案手法 B では収束するまでのエピソード数が  $6\%$  減少した。また、提案手法 C では  $13\%$  の改善が見られた。

画像分類での類似度スコアが高いカメラ、および制御対象を捉える画像の分解能が高いカメラの学習モデルほど、転移学習による学習効率が良く確認できた。

デバイス制御を対象とする強化学習では学習時間が数時間以上要するため、 $10\%$  程度の短縮でも、サービス提供までの時間短縮や省エネルギーに効果があると考えている。

提案手法 B は今回の実験条件では入力画像が低

解像度 (QVGA) であることが影響し、ロボットの向き等の情報が判断できなかった可能性があるため、入力画像の解像度を上げた実験を検討している。学習効率の改善が期待できる反面、画像分類モデルの学習時間の延伸を懸念している。

実用する際の提案手法 A, B の特性を検討する。提案手法 A は転移学習の時間短縮の効果は高い反面、カメラ数が多いほど画像分類モデルの学習時間が長くなる、かつ、カメラが追加されるたびに画像分類モデルの再学習が必要となる。提案手法 B は提案手法 A と比較して、転移学習の効率化の効果は小さいが、制御対象の学習を一度実施すれば良いため、カメラ数には依存せず、推論を実行できる。提案手法 B はカメラが不特定多数存在する場合の一次フィルタリングとして有効である。以上から、C の手法を B  $\rightarrow$  A の順の適用が実用時には有効だと考えられる。

## 6. おわりに

画像の類似度、および制御対象のバウンディングボックスの面積による転用モデルの選択により、学習効率が良くなることが判明した。また、両者を複合することで更に結果の向上が見られた。今後は転用モデルの選択だけでなく、転移の要否判断を含めた実験を実施し、提案手法の有効性確認を進めていく。

## 参考文献

- [1] HIS Market, The Internet of Things, 2017.
- [2] M. Kataoka, N. Hosikawa, H. Noguchi, T. Demizu, and Y. Yamato, "Tacit computing and its application for open IoT era," CCNC, 2018.
- [3] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," Nature, 2015.
- [4] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," ICML, 2017.
- [5] K. Simonyan, and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," ICLR, 2015.
- [6] J. Redmon, and A. Farhadi, "YOLOv3: An Incremental Improvement," <https://pjreddie.com/media/files/papers/YOLOv3.pdf>
- [7] Gazebo, <http://gazebosim.org/>