

情報メディア演習B-2 授業レポート

201821636 村松直哉

2018 年 7 月 26 日

1 Related Work

本研究の目的は、汎用的なロボット制御システムを作ることである。

1つのシステムにより、ロボットを制御するためには、「複数のエージェント（ロボット）で動くこと」と「現実空間でロボットを制御できること」が必要である。

シミュレータ環境において、1つのロボットエージェントを制御する研究は多数ある。報酬関数を使わない方法として、[1]がある。事前にメタ学習を行い、エージェントを環境になれさせ、これによりその環境に対して支持されるタスクに対してより早く目標を達成できる。

シミュレータ環境だけでなく、現実空間のロボットを強化学習により制御する手法も多く提案されている [2, 3, 4, 5, 6, 7, 8, 9]。ロボットの種類として、固定された1本のマニピュレータ [3, 6, 8]、2本のアームを使うロボット [5, 4]が利用されることが多い。移動ロボットとしては、自動車を利用したもの [7]や、サーボモータを使った多関節ロボットがある [9]。

1つの学習済みシステムにより、複数のエージェントを制御する問題は、複数タスクを1つのネットワークで行うことと言える。これは、ニューラルネットワークの破滅的忘却のため、学習が困難なことが知られている。この問題を解決したシンプルな方法として、Progressive Neural Network [10]が知られている。新しいタスクを学習をするたびに、ニューラルネットワークを大きくして行き、学習済みタスクのネットワークパラメータを固定した出力を、学習中タスクにも適用することで、多くを学習できる。[10]はAtariのゲームを使って、検証が行われた。

このように、本研究が目標としている複数のロボットに対して、現実世界で制御を行うことを目的とした研究はない。

参考文献

- [1] A. Gupta, B. Eysenbach, C. Finn, and S. Levine, “Unsupervised Meta-Learning for Reinforcement Learning.” [Online]. Available: <https://arxiv.org/pdf/1806.04640.pdf>
- [2] A. A. Rusu, M. Večerík, T. Rothörl, N. Heess, R. Pascanu, and R. Hadsell, “Sim-to-Real Robot Learning from Pixels with Progressive Nets.” [Online]. Available: <https://arxiv.org/pdf/1610.04286.pdf>
- [3] S. James and E. Johns, “3D Simulation for Robot Arm Control with Deep Q-Learning.” [Online]. Available: <https://arxiv.org/pdf/1609.03759.pdf>

- [4] S. Levine, N. Wagener, and P. Abbeel, “Learning Contact-Rich Manipulation Skills with Guided Policy Search.” [Online]. Available: <https://arxiv.org/pdf/1501.05611.pdf>
- [5] S. Levine, C. Finn, T. Darrell, and P. Abbeel, “End-to-End Training of Deep Visuomotor Policies,” *Journal of Machine Learning Research*, vol. 17, pp. 1–40, 2016. [Online]. Available: <https://arxiv.org/pdf/1504.00702.pdf>
- [6] F. Sadeghi, A. Toshev, E. Jang, and S. Levine, “Sim2Real Viewpoint Invariant Visual Servoing by Recurrent Control.” [Online]. Available: http://openaccess.thecvf.com/content/{_}cvpr/{_}2018/papers/Sadeghi{_}Sim2Real{_}Viewpoint{_}Invariant{_}CVPR{_}2018{_}paper.pdf
- [7] A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. B. Amar, and S. Wayve, “Learning to Drive in a Day.” [Online]. Available: <https://arxiv.org/pdf/1807.00412.pdf>
- [8] K. Bousmalis, A. Irpan, P. Wohlhart, Y. Bai, M. Kelcey, M. Kalakrishnan, L. Downs, J. Ibarz, P. Pastor, K. Konolige, S. Levine, and V. Vanhoucke, “Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping.” [Online]. Available: <https://arxiv.org/pdf/1709.07857.pdf>
- [9] S. Ha, J. Kim, and K. Yamane, “Automated Deep Reinforcement Learning Environment for Hardware of a Modular Legged Robot.” [Online]. Available: <https://s3-us-west-1.amazonaws.com/disneyresearch/wp-content/uploads/20180625141830/Automated-Deep-Reinforcement-Learning-Environment-for-Hardware-of-a-Modular-Legged-Robot-Paper.pdf>
- [10] A. A. Rusu, N. C. Rabinowitz, G. Desjardins, H. Soyer, J. Kirkpatrick, K. Kavukcuoglu, R. Pascanu, and R. Hadsell, “Progressive Neural Networks.” [Online]. Available: <https://arxiv.org/pdf/1606.04671.pdf>