

令和 2 年 6 月 30 日現在

機関番号：32503

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00313

研究課題名(和文) 実世界を行動するエージェントの空間認識能力と行動決定則の同時学習

研究課題名(英文) Simultaneous learning of spatial knowledge and decision making for real world agents

研究代表者

上田 隆一 (Ueda, Ryuichi)

千葉工業大学・先進工学部・准教授

研究者番号：20376502

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では移動ロボットに教えた動作を再現させる課題について取り組んだ。この課題では、移動ロボットは自身の移動やセンシングの雑音、他者からの妨害に適切に対応する必要がある。この問題に対し、ロボットが記憶をたどって動作を再現するparticle filter on episode (PFoE)というアルゴリズムを発表した。また、記憶をフェーズ分けし、より雑音や妨害に強くするアルゴリズムや、当初実験で用いた測距センサの他、レーザスキャナや加速度センサなどを試し、得られた効果を公表した。

研究成果の学術的意義や社会的意義

本研究で開発したPFoEは、ロボットが教えられた動きを再現するための、単純かつそれまでなかったアルゴリズムである。これが簡単なタスクながらも実世界の移動ロボットで機能することを確認したことで、「教えられたものを再現する」ということの原理のひとつを示したことになる。また、人のために仕事をするというレベルまでPFoEでできることを増やすためには、今後さらに研究を重ねる必要があるが、玩具程度であれば現状のもので実用可能である。

研究成果の概要(英文)：We have tackled the motion teaching problem for mobile robots.

In this problem, the robot must replay a sequence of motion that is taught against noises and interference on motion and sensing. We have proposed the particle filter on episode (PFoE) method for this problem. This method makes a mobile robot replay motion sequence based on episodes memorized in the robot.

After publishing it, we have then tried algorithms that divide a sequence of memory for giving meaning to each part. We have also examined PFoE with a laser scanner and an acceleration sensor.

研究分野：移動ロボット

キーワード：PFoE 教示 teach-and-replay 移動ロボット 確率ロボティクス

1. 研究開始当初の背景

2014年のノーベル生理学・医学賞は、齧歯類の脳の海馬周辺から「場所細胞」等、特定の機能を持つ細胞が発見されたことが受賞理由となった。場所細胞は、齧歯類（以後はラットと表記）が慣れ親しんだ環境にいる場合、活動パターンと環境中での場所が一一対一対応するという特徴があり、ラットが脳内に「環境の地図」を持つという証拠とされている[1]。この知見はすでにロボットに応用されており、例えば Milford らは場所細胞の機能を真似て環境の地図を生成する「RatSLAM」という手法を提案している[2]。

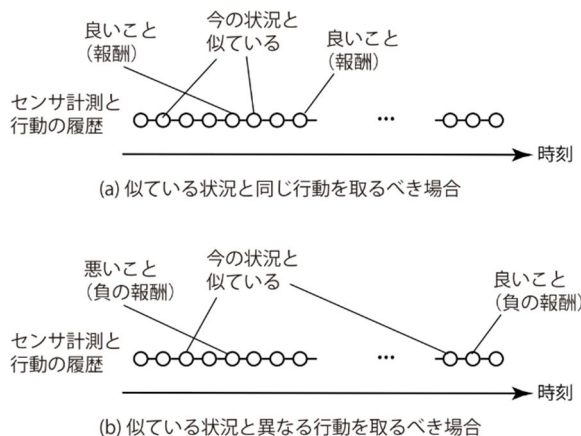


図 1: エピソード上のパーティクルフィルタによる行動決定の原理

申請者らも同様に海馬の機能に注目し、海馬とロボットのナビゲーション手法の関連性を調査している。この研究での議論の中で、地図を用いてラットがナビゲーションするのは、経験したことを無限に記憶できないからなのではないかという仮説が出た。また、ラットが少し前の記憶を思い出す時、ラットが過去通った場所の順番に対応して場所細胞が活動するという現象（記憶の回顧）がある。これらの知見をヒントに、地図にこだわらず、経験のシーケンス（エピソード）を無限に蓄積してそれを思い出す計算をパーティクルフィルタと呼ばれる手法で実装し、実機実験したところ、学習アルゴリズムとして働くことを発見した[3]。このアルゴリズムは「エピソード上のパーティクルフィルタ（Particle Filter on Episode, PFoE）」と呼ばれる。

図 1 に PFoE の考え方を示す。(a), (b)に見られる丸を連結したものはロボットの記憶（センサ値やとった行動、得た報酬）を示す。PFoE では、パーティクルフィルタを用いて、今の状況が記憶のどの部分と似ているかを探し、(a)のようにその先にあった報酬がよければその時の行動をとり、(b)のように報酬が悪ければ別の行動を行っていく。学習が進むと記憶上に良い報酬を得た経験が蓄積され、行動選択が適切になって行く。図 2 は実機に PFoE を適用している様子であり、このタスクではロボットは「左右交互にゴールが置かれる」というルールを 20 回程度の試行で学習した。

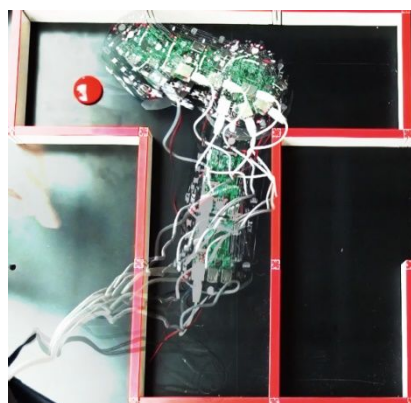


図 2: 学習後のロボットの動作

(丸がゴール地点で交互に置かれ、ロボットは交互にゴール地点に行く)

一方で PFoE は簡単な仕事しかロボットに行わせることができない。ここで言う簡単か、難し

いかの基準は、ロボットが自分の行動を決める数である。現状の PFoE ではこの決定の数が高々 1, 2 回であり、動作のようにわずかの時間に無数の出力の決定を行う用途にはこのまま利用することが難しい。

そこでやはり地図に相当するものが必要となる。ただし、海馬の話に戻ると、動物の場合はロボットで行われるような「地図を作って（空間に対する知識をロボットに与えて）、その後に行動を決める（学習や制御則を実装する）」という方式よりも、空間に対する知識の構築と行動の決定が同時並行的に進むのではないかと考えられる。また、それがロボットで実現すると、ロボットの行動生成の際の人間の介在が少なくなり、実装のコストが下がる。

2. 研究の目的

そこで申請時、(1)、(2)の二つの点を明らかにすることを目的とした。また、研究期間中、他に(3)、(4)のように新たな目的を設定した。

(1) 記憶のクラスタリング

移動ロボットが何か（例えば「ある場所に行く」等の仕事）をして移動すると、記憶のシーケンスができるが、このシーケンスをグループ分けすることで、地図状のデータができることを示す。このクラスタリングは「タスクに基づいて行われる」ことが特徴で、環境をセンサで計測した時に似たような値が得られた状況をただ単にクラスタリングするよりも、そのタスクの文脈やその環境の持つ特徴に即した地図ができることを示す。

(2) 地図を用いた PFoE

PFoE に、上記の地図生成のためのクラスタリングアルゴリズムを付加することで、移動ロボットが各種のタスクを学習できるようにする。タスクについては、環境での適切な移動という基本的なナビゲーションタスク、ナビゲーションのゴール地点を交互に別の場所に置く（同じ場所が餌の位置によって別の意味を持つ）タスク、何か環境中の目印からゴール地点の手がかりを得て、そこに移動するタスク等を考える。

(3) 実用性の向上や多種多様なセンサの利用

PFoE を、学習よりもはやく行動を獲得できる教示 (teach-and-replay) で用いる方法を確立する。

(4) 多種多様なセンサの利用

図 2 の実験では、赤外線測距センサの値をエピソードに記録して PFoE に用いたが、PFoE は近年用いられるレーザスキャナ（いわゆる LiDAR）や加速度センサ、ジャイロセンサなどとも組み合わせることができると考えられる。これらのセンサを PFoE に利用して評価する。

3. 研究の方法

(1) teach-and-replay のための PFoE

まず、2(3)の目的のために、PFoE を強化学習だけでなく、teach-and-replay タスクに適用することにした。teach-and-replay とは、人がコントローラを通じて教えた動きをロボットが再現することである。マニピュレータと比較すると、実験で使う移動ロボットの場合は動きに再現性が乏しく、センサフィードバックが必要な難問となる。この問題に対し[4]に記述されているアルゴリズムを考案し、実験を行った。なお、このアルゴリズム自体に関しては研究期間前に考案しており、実験により実世界で機能することを確認していた。研究期間中は、より詳細に実験を行った。

(2) 記憶のクラスタリング

記憶のクラスタリングについて、機械学習で用いられているベイズ統計学に基づく手法をいくつか試した。このクラスタリングは、ロボットが teach-and-replay 中に、突然タスクで通る別の場所に移動されたときに（移動ロボットの用語で「誘拐」されたとき）、それを知覚してその場所からタスクを再現するためのアルゴリズム（リセット）を利用して評価した。

(3) 地図生成のためのクラスタリング

2(2)について、teach-and-replay に SLAM (ロボットの位置推定と地図生成を同時に行う方法) を用いるソフトウェアを構築した。

(4) 様々なセンサの利用

2(4)について、密な2次元のレーザスキャナ (以後 LiDAR と呼ぶ) を用いた PFoE, 加速度センサを利用した PFoE の2つのアルゴリズムを構築し、実験を行った。

4. 研究成果

次のような成果が得られた。[4]-[8], [10], [11], [13]-[16]が本研究の成果である。

(1) teach-and-replay のための PFoE

[8]の実験で、ロボットが動きで数を数えるタスク、壁沿い走行をするタスク、人のあとを追いかけるタスクなどを PFoE で実現できることを確認した。これにより、PFoE はロボットに teach-and-replay を行わせるための全く新しいアルゴリズムであるが、これが実世界で機能することを示せた。詳細は[8]に記載してある。

(2) 記憶のクラスタリング

まず、[5,6]では GMM (Gaussian Mixture Model, 混合ガウス分布モデル) でセンサデータをクラスタリングすることで、ロボットが teach-and-replay タスクの中で見せるサブタスク (壁に沿って走る、角を曲がる、引き返す、など) の分類を行った。また、[7]では時系列情報を分類できる HDP-HSMM (Hierarchical Dirichlet Process Hidden Semi-Markov Model) を使用した。

GMM, HDP-HSMM とともに、ロボットがタスクの一連の動作をサブタスクに分類することに有効であることが示せた。これにより、ロボットがタスク中に妨害を受けた際に、どのサブタスクのどこから動作を再開させるかを適切に選べるようになった。また、後に試した HDP-HSMM のほうが、GMM よりも事前に設定しなければならないパラメータの数を少なくすることができ、より汎用的に利用できるということが分かった[7]。

(3) 地図を用いた PFoE

SLAM (simultaneous localization and mapping) アルゴリズムである Cartographer[9]を用い、次のようなソフトウェアを構築した。いまだ文章化していないが、GitHub においてソフトウェア[11], YouTube において実験動画[10]を公開した。

構築したソフトウェアは次のような teach-and-play 機能を実現した。

- ・操作者がロボットを走らせると、Cartographer が地図とロボットの経路を推定
- ・一通り地図ができたあと、操作者がロボットのボタン操作で教示の終了を知らせると、ロボットは作られた地図で自己位置推定しながら、記録された経路を順に移動

[10]にあるようにこのソフトウェアは、一つの環境での実験ながら実世界において機能することが確認された。このように地図を用いる移動ロボットの teach-and-replay については[12]にあるように、地図と整合性のある経路を記録することが重要であり、難しい。例えば軌跡が地図中の壁にめり込む等、teach-and-replay が不可能になる出力が得られる場合があるが、Cartographer を用いると、そのような問題は起こりにくいことが確認できた。

一方、自己位置推定の不確かさを考慮しながら無駄のない動きで軌跡を追従することまでは teach-and-replay の枠組みでは実現できず、新たな研究課題を確認した。また、単に機能を実装しただけで、動物の脳との関連性については深く考察ができなかった。研究期間中は、地図を用いないで teach-and-replay を実現するアルゴリズムのほうの開発が進んだが、この成果と地図との併用が今後の課題として残った。

(4) 様々なセンサによる PFoE

LiDAR による PFoE

LiDAR からは1回の観測で数十から数百の値が得られる。これらの値は、どの方向にレーザを反射するもの (主に壁) がどの距離で存在するかを示す。このようなセンサは、[4]で用いたよ

うな、1 個の値しか返さない測距センサを数個並べただけのセンサとの比較で、「密 (dense) なセンサ」と呼ばれる。

まず、センサの返す多くの値の中から、4 つの値 (4 方向の壁の距離) を選んで PFoE が機能することを確認した[13]。これより多くのセンサ値を使おうとすると計算量が増えてしまうため、センサ値から壁の形状とロボットの壁からの姿勢を求め、これらのデータをセンサ値に反映させることを試みた。実験結果は[14]にまとめられている。ただ、[14]の結果は[13]の方法に比べて良好ではなく、密なセンサを PFoE で用いる難しさが確認されることとなった。また、実装した壁の形状と姿勢の比較は、少しのロボットの向きの違いが過大に評価されるものであったので、この点を改良することが課題として残った。

加速度センサによる PFoE

次に、加速度センサを用いた PFoE を試した[15,16]。実験によって、測距センサなしで教示した動きを再現できることを確認した。ただし、これは教示で与えた制御指令をそのまま再生するだけでも同じ効果が得られるため、有効性を示すには、測距センサや LiDAR と組み合わせ、より複雑なタスクに適用できることを示さなければならない。

参考文献

- [1] G. Buzsáki and E. I. Moser, “Memory, navigation and theta rhythm in the hippocampal-entorhinal system,” *Nature Neuroscience*, vol. 16, no. 2, pp. 130-138, 2013.
- [2] Milford, M. J. and Wyeth, G. F.: “Mapping a Suburb With a Single Camera Using a Biologically Inspired SLAM System,” *IEEE Trans. on Robotics and Automation*, vol. 24, no. 5, pp. 1038-1053, 2008.
- [3] R. Ueda, K. Mizuta, H. Yamakawa and H. Okada: “Particle Filter on Episode for Learning Decision Making Rule,” *Proc. of Intelligent Autonomous Systems 14 (IAS-14)*, Shanghai, 2016.
- [4] Ryuichi Ueda, Masahiro Kato, Atsushi Saito, Ryo Okazaki: Teach-and-Replay of Mobile Robot with Particle Filter on Episode, *Proc. of IEEE International Conference on Robotics and Automation*, pp.3475-3481, Brisbane, Australia, 2018.
- [5] 齊藤篤志, 上田隆一: “Particle Filter on Episodeにおける記憶のクラスタリング”, 日本機械学会ロボティクス・メカトロニクス講演会2018講演論文集, pp.1A1-C13, 北九州, 2018.
- [6] 齊藤篤志, 上田隆一: Particle Filter on Episodeにおける尤度関数の自動決定, 第24回ロボティクスシンポジウム予稿集, pp.213-218, 富山, 2019.
- [7] 齊藤篤志, 上田隆一: Particle Filter on EpisodeにおけるHSMM-HDPを用いたリセット法, 日本機械学会ロボティクス・メカトロニクス講演会2020講演論文集, 2A1-L09, 2020.
- [8] Ryuichi Ueda, Masahiro Kato, Atsushi Saito: Particle Filter on Episode, arXiv:1904.08761, 2019.
- [9] W. Hess, D. Kohler, H. Rapp, and D. Andor, Real-Time Loop Closure in 2D LIDAR SLAM, in *Robotics and Automation (ICRA)*, 2016 IEEE International Conference on. IEEE, 2016. pp. 1271-1278.
- [10] Ryuichi Ueda: “teach-and-replay with Cartographer,” <https://www.youtube.com/watch?v=fk8Y7kWhSQ>, YouTube, 2019. (last visit: 2020-06-19)
- [11] Ryuichi Ueda: “raspimouse_map_based_teach_and_replay,” https://github.com/ryuichiueda/raspimouse_map_based_teach_and_replay, GitHub, 2019. (last visit: 2020-06-19)
- [12] 原 祥亮, 坪内 孝司, 油田 信一: “教示再生による自律走行のための ROS stacks を活用した地図と経路の獲得”, ロボティクスメカトロニクス講演会2012講演論文集, 2A1-L09, 2012.
- [13] Ryuichi Ueda: “PFoE with URG,” <https://www.youtube.com/watch?v=89IsMelhzug>, YouTube, 2017. (last visit: 2020-06-24)
- [14] 芦埜遥平: 密なセンサを用いたParticle Filter on Episode, 2019年度千葉工業大学先進工学部未来ロボティクス学科卒業論文, 2020.
- [15] 佐藤大亮, 齊藤篤志, 上田隆一: 加速度情報を用いたEpisode on Filter Particle, 日本機械学会ロボティクス・メカトロニクス講演会2019講演論文集, 1A1-L06, 広島, 2019.
- [16] 佐藤大亮: 教示のためのParticle Filter on Episodeにおける自律ロボットの動作再現性の向上, 2019年度千葉工業大学大学院未来ロボティクス専攻修士論文, 2020.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計5件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 齊藤 篤志, 上田 隆一
2. 発表標題 Particle Filter on Episodeにおける尤度関数の自動決
3. 学会等名 第24回ロボティクスシンポジア
4. 発表年 2019年

1. 発表者名 Ryuichi Ueda, Masahiro Kato, Atsushi Saito, Ryo Okazaki
2. 発表標題 Teach-and-Replay of Mobile Robot with Particle Filter on Episode
3. 学会等名 IEEE International Conference on Robotics and Automation (国際学会)
4. 発表年 2018年

1. 発表者名 齊藤篤志, 上田隆一
2. 発表標題 Particle Filter on Episodeにおける記憶のクラスタリング
3. 学会等名 日本機械学会ロボティクス・メカトロニクス講演会
4. 発表年 2018年

1. 発表者名 加藤 正紘, 上田 隆一, 岡崎 亮
2. 発表標題 教示のためのエピソード上のパーティクルフィルタの応用 -人間の操縦で得られた移動ロボットの経路選択の再現
3. 学会等名 第35回日本ロボット学会学術講演会
4. 発表年 2017年

1. 発表者名 上田 隆一
2. 発表標題 教示のためのエピソード上のパーティクルフィルタ
3. 学会等名 第35回日本ロボット学会学術講演会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>Particle Filter on Episode https://arxiv.org/abs/1904.08761</p> <p>実世界を行動するエージェントの空間認識能力と行動決定則の同時学習 https://lab.ueda.tech/?page_id=3301</p> <p>raspimouse_gamepad_teach_and_replay GitHub https://github.com/ryuichiueda/raspimouse_gamepad_teach_and_replay</p> <p>芦埜遥平: 密なセンサを用いたParticle Filter on Episode, 2019年度千葉工業大学先進工学部未来ロボティクス学科卒業論文, 2020.</p> <p>佐藤大亮: 教示のためのParticle Filter on Episodeにおける自律ロボットの動作再現性の向上, 2019年度千葉工業大学大学院未来ロボティクス専攻修士論文, 2020.</p> <p>齊藤篤志: Particle Filter on Episodeにおける尤度関数の自動決定, 2018年度千葉工業大学先進工学部未来ロボティクス学科卒業論文, 2019.</p>

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考