The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

場所概念に基づく確率推論による音声命令からのパスプランニング

Path Planning by Spatial Concept-Based Probabilistic Inference from Human Speech Instructions

谷口 彰 *1 Akira Taniguchi　萩原 良信 *1 Yoshinobu Hagiwara　谷口 忠大 *1 Tadahiro Taniguchi　稲邑 哲也 *2*3 Tetsunari Inamura

*1立命館大学 Ritsumeikan University　*2国立情報学研究所 National Institute of Informatics　*3総合研究大学院大学 The Graduate University for Advanced Studies

The aim of this study is to enable a mobile robot to perform navigation tasks by probabilistic inference using spatial concepts on a probabilistic model. Specifically, path planning is performed to the target state of a spatial concept estimated through human speech instructions such as “Go to the kitchen”. In the experiment, places instructed by the speech command of the user showed high probability values, and the trajectory toward the target place was correctly estimated.

1. はじめに

人間と共存し様々な生活環境で動作するロボットは，適応的に場所に関する概念や語彙を学習するだけではなく，学習した知識をナビゲーションなどのタスクに活用することが求められる．これまで，我々は教師なし学習によりマルチモーダル情報から自律的・適応的に場所のカテゴリ知識 (場所概念) を獲得する確率的生成モデル SpCoSLAM を提案している [Taniguchi 17, Taniguchi 18]．本研究では，ロボットが自らの経験に基づいて形成された概念を活用する方法について検討する．具体的には,「だいどころに行って」などの人の音声発話命令から，場所概念に基づき推定された目標状態へのパスプランニングを行う．

これまで，我々は Simultaneous Localization And Mapping (SLAM) [Thrun 05] と位置・音声言語・画像のマルチモーダル情報に基づくノンパラメトリックベイズ場所概念獲得モデルを統合した SpCoSLAM を提案している [Taniguchi 17]．また，地図・場所概念・語彙を同時に学習できるオンライン学習アルゴリズムの構築を行ってきた．この手法では，地図を持たない未知環境下からでも場所概念の獲得が可能であり，環境中の未探索の領域に対して追加学習が可能である．また，計算リソースが限られる実ロボットにおいて長期学習を可能とするため，データ数の増加に伴う計算コストの増加，バッチ学習と比較した場合のオンライン学習による推定精度の低下に対処した高精度かつ軽量なオンラインアルゴリズムの開発を行った [Taniguchi 18]． SpCoSLAM では，環境や語彙に関する事前知識なしに地図と場所概念および語彙を獲得することが可能である．しかし，獲得した場所概念の行動決定や意思決定への活用に関する理論的な方法についてはこれまで議論されていなかった．そこで本研究では，SpCoSLAM の確率的生成モデルの上でのロボットの行動決定問題に着目する．

本研究では，場所概念を活用したロボットと人の言語的インタラクションに基づくパスプランニングやナビゲーションを可能とすることを目指す．本稿では，SpCoSLAM の確率的生成モデル上で行動決定に関する確率推論の手法を構築し，実際に獲得された場所概念を用いたナビゲーションタスク実験を行い，提案手法の有効性を検証する．

一般的に，ロボットにおける行動決定には強化学習が用いられることが多い．強化学習では，マルコフ決定過程 (MDP: Markov Decision Process) に基づき試行錯誤的に行動の最適化を行う．一般的な強化学習の枠組みでは，累積報酬の最大化問題を解くが，この枠組みには確率的生成モデルとの理論的な隔たりがあった．しかし近年，強化学習を始めとする行動決定問題を確率的生成モデル上の確率推論として捉える枠組みが注目されている [Toussaint 09, Levine 18]．本研究では，行動決定問題を確率的生成モデル上の確率推論として定式化することによって，教師なし学習における推論と同じ枠組みでパスプランニングやナビゲーションを実現することができると考える．

連絡先: 谷口彰，立命館大学 情報理工学部，a.taniguchi@em.ci.ritsumei.ac.jp

2. 提案手法：SpCoNavi

本章では，SpCoSLAM と同一な生成過程のグラフィカルモデル上での確率推論による，人の音声発話命令からの場所概念に基づくナビゲーション手法 SpCoNavi について説明する．

2.1 概要

本研究の問題設定は，部分観測マルコフ決定過程 (POMDP: Partially Observable Markov Decision Process) とみなせる．確率推論を行うにあたって，音声認識や自己位置推定など様々な不確実性への対処は重要である．SpCoSLAM はベイズフィルタベースの SLAM にマルチモーダルカテゴリゼーションを統合したモデルであるため，本研究の問題設定は多くの不確実性を伴ったより複雑な POMDP であると言える．本研究では，ロボットに確定的な目標地点を明示的に与えるのではなく，ロボットが人の音声命令から自らの内部知識を用いて行動目標を決定することを行う．また，全体が統合されたモデルに基づく推論であるため，確率変数の不確実性を考慮したマルチモーダル情報の相互補完的な推論が可能になる．これにより，モデル上で全体最適なパスプランニングを行うことが可能になると考える．

2.2 定式化と導出

本研究では，環境の地図，場所概念，語彙が既に獲得されているものとする．ある環境上で事前に SpCoSLAM を動作させ，最終時刻での学習結果における各モデルパラメータを固定とする．SpCoSLAM に基づいたナビゲーションのためのパスプランニングを行う際のグラフィカルモデルを Fig. 1 に示す．灰色のノードは観測済み変数および学習済みパラメータを表し，白色のノードは未観測変数を表す．このとき，地図 m,

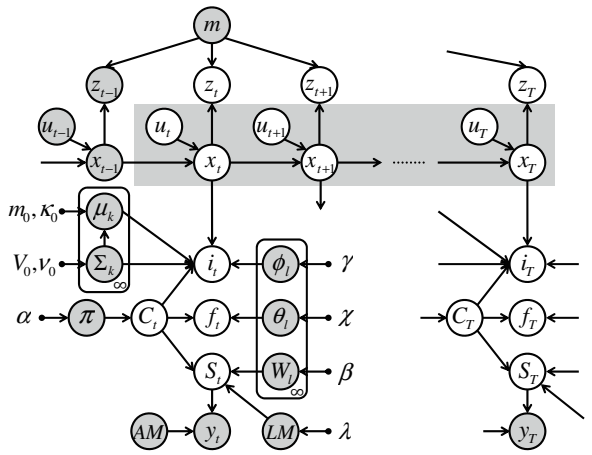

図 1: SpCoNavi のグラフィカルモデル

場所概念を表すモデルパラメータ集合 $\{\mathbf{W}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \theta, \phi, \pi\}$ および，言語モデル LM といった学習済みグローバルパラメータ集合を $\Theta_G = \{m, \mathbf{W}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \theta, \phi, \pi, LM\}$ とする．各変数の詳細については [Taniguchi 17] を参照されたい．灰色の長方形で示した箇所が提案手法で推定対象とする変数の範囲である．下添え字の T はパスプランニングの最終時刻であり，パスプランニング開始時刻を $t = 1$ としたとき，計画区間 (Planning Horizon) を表す．以降では開始時刻 $t = 1$ として数式の表記を行う．

提案手法では，人からの音声発話命令 y_t が与えられたときの動作 u_t と位置 x_t の軌道 (Trajectory) $\tau = \{u_{1:T}, x_{1:T}\}$ を表す確率分布が最大となる動作系列 $u_{1:T}$ (および，地図上のパス $x_{1:T}$) の推定を行う．目的関数となる確率分布が最大値とるときの動作系列を求める式 (1) に示す．

$$u_{1:T} \quad = \quad \underset{u_{1:T}}{\operatorname{argmax}}\, p(\tau \mid y_{1:T}, x_0, \Theta_G) \tag{1}$$

このとき，前時刻 $t-1$ における自己位置 x_{t-1} は与えらえるものと仮定する．実際には，事前に設定した初期位置や Monte-Carlo Localization (MCL) [Dellaert 99] のような自己位置推定により得られる推定値を利用することができる．また，音声発話命令 y_t はゴールシグナルを表す観測変数であるとみなし，現在時刻 t から時刻 T まで同様の観測が得られ続けるという仮定をおく．この仮定は，[Levine 18] における最適か否かを表すバイナリ確率変数 $\mathcal{O}_t$ に対応する．[Levine 18] では，事後分布 $p(\tau \mid o_{1:T})$ における最大事後確率の推論がプランニング問題に対応することが述べられている．ここで，$o_{1:T} = \{\mathcal{O}_t = 1\}_{t=1}^{T}$ である．つまり，式 (1) のように，観測 $y_{1:T}$ を条件とする事後確率分布を最大化することで，指示された場所概念によって示される位置へ移動するための最適な軌道が推定可能となる．

式 (1) を計算するためには，潜在変数 $z_{1:T}$, $i_{1:T}$, $C_{1:T}$, $S_{1:T}$ に関して周辺化する必要がある．このとき，$f_{1:T}$ も未観測な変数であるが，グラフィカルモデルの依存関係より消去されるため，計算上考慮する必要はない．$y_{1:T}$ は実際にはすべて同じ観測情報であるため，音声認識結果 $S_{1:T}$ もすべて同じとみなせる．さらに，音声認識を表す確率分布 $p(S_T \mid y_T, AM, LM)$ の確率値を直接求めることは困難であるため，近似として N-best 音声認識結果 $S_T^N = \{S_T^{(n)}\}_{n=1}^{N}$ を用いる．また，$p(u_t)$ は一様分布を仮定する．加えて，z_t に関する周辺化による足し合わせについては $p(z_t \mid x_t, m)$ の項しか影響を受けないため，$\sum_{z_t} p(z_t \mid x_t, m) = 1$ とできる．

また，地図の形状を考慮し障害物に近づかないような移動を行わせるために，式 (2) のようにマップベースド動作モデル (Map-Based Motion Model) を利用する．これにより，障害物に近づくほど確率値が小さくなり，障害物のあるセルでは確率 0 になるように確率化したコストマップ *1 $p(x_t \mid m)$ を考慮した動作モデルを考えることができる．

$$p(x_t \mid u_t, x_{t-1}, m) \quad \propto \quad p(x_t \mid m) p(x_t \mid u_t, x_{t-1}) \tag{2}$$

以上より，軌道の事後確率分布は式 (3) のようになる．

$$\begin{aligned} & p(\tau \mid y_{1:T}, x_0, \Theta_G) \\ & \approx \prod_{t=1}^{T} p(x_t \mid u_t, x_{t-1}) \prod_{t=1}^{T} \Bigg[p(x_t \mid m) \\ & \quad \sum_{C_t} \Bigg[\mathrm{Mult}(S_T^N \mid W_{C_t}) \mathrm{Mult}(C_t \mid \pi) \\ & \quad \sum_{i_t} \mathcal{N}(x_t \mid \mu_{i_t}, \Sigma_{i_t}) \mathrm{Mult}(i_t \mid \phi_{C_t}) \Bigg] \Bigg] \end{aligned} \tag{3}$$

このとき，状態 x_t は実際には連続値であるが，占有格子地図 m の各セルごとに離散化して表現する．動作 u_t は 1 時刻につき地図上の自己位置から周囲 1 セルを移動するものとする．地図の各セルごとに確率値を計算し，動的計画法に基づく Viterbi アルゴリズムによって軌道の推定を行う．

3. 実験

本実験では，場所概念を獲得したロボットが提案手法によりパスプランニングを行うことが可能であることを示す．

3.1 実験条件

環境内にいるロボットに対し，ユーザは “** に いって” と発話する．** には各場所の名前が入る．実験環境は [Taniguchi 17, Taniguchi 18] で使用したもの *2 と同様である．また，[Taniguchi 18] において学習された Step 50 における最大尤度のパーティクルの地図，場所概念，言語モデルを用いる．音声認識には Julius dictation-kit-v4.4 (DNN-HMM decoding) を使用した．グローバルコストマップは Robot Operating System (ROS) の navigation，costmap_2d パッケージを用いて取得した．計画区間は $T = 100$，N-best 音声認識は $N = 10$ とした．ロボットの初期位置は，地図の中央に設定した．本実験においては，動作モデルは確率モデルではなく，確定的なモデルを仮定する．

3.2 実験結果

提案手法によるパスプランニングの結果の例を示す．図 2，図 3，図 4 に地図および，発話されたときの地図上の各セルにおける出力確率の値をカラーマップで示し，その上に推定された軌道を赤色で示す．図 2 において発話された場所は “きょうゆうせき” (右上の広い部屋)，図 3 において発話された場所は “きゅうけいじょ” (上側右端の狭い部屋と上側左端の部屋)，図 4 において発話された場所は “きょういんけんきゅうしつ” (下側の右の部屋) であった．

*1 Configuration Space (C-Space) の確率的な表現と見なせる．
*2 Robotics Data Set Repository (Radish) [Howard 03] のオープンデータセット

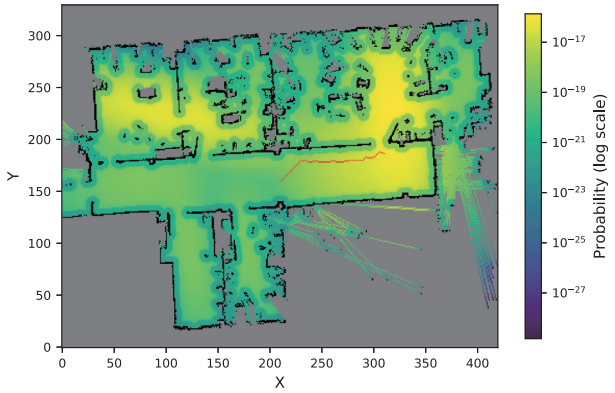

図 2: “きょうゆうせき” における出力確率と推定された軌道

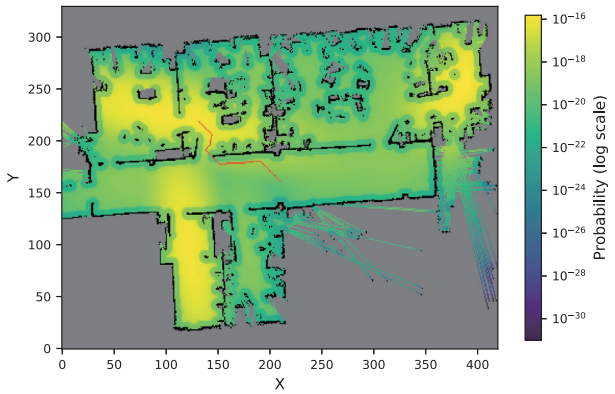

図 3: “きゅうけいじょ” における出力確率と推定された軌道

結果として，ユーザの音声命令によって指示された場所の確率値が高くなっており，指示された場所へ向かう軌道が推定されていることがわかる．地図上の確率値において，対象の場所以外の場所で確率値が高くなっている場合も見られたが，これは獲得された語彙の単語辞書に基づく音声認識結果および，場所概念の学習における誤りが原因であると考えられる．

4. おわりに

本稿では，ロボットが自律的に獲得した場所概念を用いた音声命令からのパスプランニング手法 SpCoNavi を提案した．提案手法は，従来の多くのパスプランニング手法と異なり，ゴールとして地図上の座標点を指定する必要がなく，人との音声によるインタラクションから目的地までのナビゲーションが可能となる．実験では，提案手法が音声命令により指示された場所への軌道を推定できることを示した．しかし，占有格子地図上の移動可能なセルを状態として定義したため，地図の大きさに応じて状態数が多くなり，計算量が増加する問題が見られた．

今後の課題として，A* アルゴリズムなどの従来のパスプランニング手法との定量的な比較評価や，リアルタイムなナビゲーションのための計算量の削減法の検討が挙げられる．

謝辞

本研究の一部は，JST AIP-PRISM JPMJCR18ZC，JST CREST JPMJCR15E3，JSPS 科研費 JP17J07842，JP16H06561，JP16K12497 の助成を受けたものである．

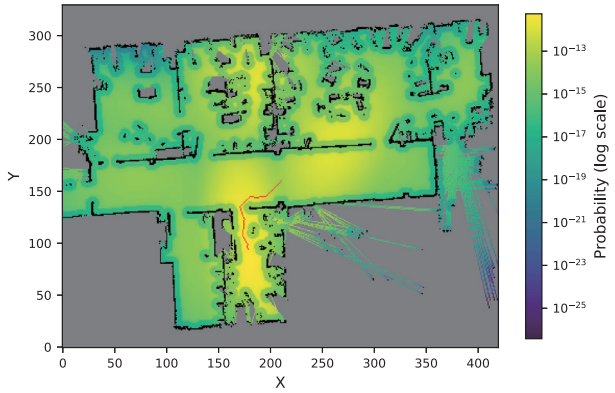

図 4: “きょういんけんきゅうしつ” における出力確率と推定された軌道

参考文献

[Dellaert 99] Dellaert, F., Fox, D., Burgard, W., and Thrun, S.: Monte carlo localization for mobile robots, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Vol. 2, pp. 1322–1328IEEE (1999)

[Howard 03] Howard, A. and Roy, N.: The Robotics Data Set Repository (Radish) (2003)

[Levine 18] Levine, S.: Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review, *arXiv preprint arXiv:1805.00909* (2018)

[Taniguchi 17] Taniguchi, A., Hagiwara, Y., Taniguchi, T., and Inamura, T.: Online Spatial Concept and Lexical Acquisition with Simultaneous Localization and Mapping, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 811–818 (2017)

[Taniguchi 18] Taniguchi, A., Hagiwara, Y., Taniguchi, T., and Inamura, T.: Improved and Scalable Online Learning of Spatial Concepts and Language Models with Mapping, *arXiv preprint arXiv:1803.03481* (2018)

[Thrun 05] Thrun, S., Burgard, W., and Fox, D.: *Probabilistic Robotics*, MIT Press (2005)

[Toussaint 09] Toussaint, M.: Robot trajectory optimization using approximate inference, in *Proceedings of the 26th annual international conference on machine learning (ICML)*, pp. 1049–1056ACM (2009)