

確率的自己位置推定法における機械学習の併用

○赤井 直紀 (名古屋大学) 平山 高嗣 (名古屋大学) 村瀬 洋 (名古屋大学)

1. はじめに

確率的推論の様なモデルベースの自己位置推定法がこれまでに多数提案され、様々な成功を収めてきた。しかし、移動ロボット実用化などという観点から見ると、これらの手法には限界がある様に思える。一方で機械学習、特に畳み込みニューラルネットワーク (CNN) などの深層学習の発展は著しく、モデルベース法を凌ぐ性能が多数報告されている。しかし、学習器の内部状態に対する説明性の不足や、未学習データに対する出力の不安定性などが指摘されており、機械学習の安定性が懸念されている。本稿では、自己位置推定を題材として、モデルベース法と機械学習を併用して性能向上を計ろうとする、著者らの近年の研究 [1] を報告する。

2. 関連研究

深層学習の発展により、様々なモデル化困難な入出力の関係が、データ駆動的に End-to-End (E2E) で学習・モデル化できる様になってきた。自己位置推定の分野では、E2E 学習による自己位置推定である PoseNet が 2015 年に Kendall らによって提案された [2]。さらに Kendall らは、E2E 学習による自己位置推定結果の不確かさの推定方法も提案した [3]。これらを境に CNN の自己位置推定への適用例は増え続け、深層 Autoencoder の様な次元圧縮を活用する例 [4]、自己位置推定モデルをネットワークにより代替する例 [5, 6] などが提案されている。さらに、点群に対しても深層学習の適用が可能となり [7]、点群位置合わせに CNN が活用される例も提案された [8]。興味深い例として、E2E 学習で車両制御のための行動を学習し、この行動の予測値を尤度として用いる自己位置推定法も提案されている [9]。

また、自己位置推定の失敗を機械学習により検知する方法も提案されている [10]。自己位置推定の失敗は一般にモデル化困難であるため、機械学習の適用は極めて効果的であるといえる。著者らも、自己位置推定の失敗を CNN を用いて検知する方法を提案した [1]。また提案法は、CNN による予測が不確かさを含むことを前提とし、予測結果を可観測変数として利用する。そして、自己位置とその位置推定結果の信頼度推定を行う問題として定式化している。これにより、モデルベース法と CNN の協調的な働きが誘発できる。

3. モデルベース法と機械学習の併用

3.1 信頼度付き自己位置推定

上述の通り提案法は、現時刻 t の自己位置 \mathbf{x}_t と信頼度 r_t を同時推定する問題として定式化を行っている。なお信頼度とは、自己位置推定が成功している確率を表した値であり、 $r_t = p(s_t = \text{success})$ と記述される。ここで s は自己位置推定状態であり、success と failure の離散的な 2 状態を取る。

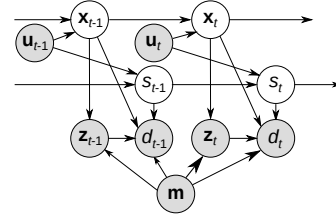


図 1 信頼度付き自己位置推定のグラフィカルモデル [1].

図 1 には、提案法で利用されるグラフィカルモデルを示す。白色のノードが隠れ変数、灰色のノードが可観測変数である。可観測変数としては、制御入力 $\mathbf{u}_{1:t}$ 、センサ観測 $\mathbf{z}_{1:t}$ 、地図 \mathbf{m} 、そして自己位置正誤判断 $d_{1:t}$ を有する。正誤判断は CNN により予測される 0 から 1 の連続値であり、1 に近い程位置推定に成功している可能性が高いことを意味する確率である。本モデルに対する事後分布は、以下の様に分解できる。

$$\begin{aligned} p(\mathbf{x}_t, s_t | \mathbf{u}_{1:t}, \mathbf{z}_{1:t}, \mathbf{m}, d_{1:t}) \\ = p(\mathbf{x}_t | \mathbf{u}_{1:t}, \mathbf{z}_{1:t}, \mathbf{m}, d_{1:t}) p(s_t | \mathbf{x}_t, \mathbf{u}_{1:t}, \mathbf{z}_{1:t}, \mathbf{m}, d_{1:t}) \end{aligned}$$

提案法ではこの分布を Rao-Blackwellized particle filter (RBPF) を用いて推定する。具体的には、自己位置に対する分布を PF で求め、自己位置推定状態に対する分布を解析的に求める。ここで、自己位置に対する分布を $\text{Prob}(t) \stackrel{\text{def}}{=} p(\mathbf{x}_t | \mathbf{u}_{1:t}, \mathbf{z}_{1:t}, \mathbf{m}, d_{1:t})$ と定義すると、この分布は以下の様に展開できる。

$$\begin{aligned} \text{Prob}(t) = \eta p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{m}) \sum_{s_t} p(d_t | \mathbf{x}_t, s_t, \mathbf{z}_t, \mathbf{m}) p(s_t) \\ \times \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) \text{Prob}(t-1) d\mathbf{x}_{t-1} \end{aligned}$$

ここで η は正規化係数、 $p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{m})$ は観測モデル、 $p(d_t | \mathbf{x}_t, s_t, \mathbf{z}_t, \mathbf{m})$ は正誤判断モデル、 $p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t)$ は動作モデルである。

通常 PF で自己位置推定を行う場合には、観測モデルのみを用いて尤度計算を行う。一方で提案法では、観測モデルと正誤判断モデルを用いて尤度計算を行う。ここで判断モデルは、例えば信頼度の高い (低い) パーティクルに対して、CNN が自己位置推定結果が誤っている (正しい) と判断した場合に、尤度を減少させる働きを持つ。そのため、CNN の出力が含む不確かさの影響を低減させる効果を持つ。また、センサ観測と地図が誤対応して観測モデルの尤度が高くなった場合にも、CNN が自己位置推定に誤っていると判断すれば、正しく信頼度が低い状態にあると認識できる。

3.2 学習のタスクから見た併用の利点

画像上の物体を CNN により検知・識別するためには、大量のデータが必要となる。文献 [2] においても、

E2E で自己位置推定を行うためには、転移学習の様な効率化を行う必要があると述べられている。また E2E 自己位置推定は、画像と自己位置の関係を学習するため、当然ながら学習した環境でしか適用できない。

一方で提案法は、自己位置推定の正誤の学習を行う。これは、センサ観測と地図の間に誤対応があるかどうかを学習することと等しい。すなわち、通常の物体検知・認識と比較して、学習に対するタスクは圧倒的に簡略化されるため、学習データが少なくても問題にならない。さらに、センサ観測と地図の間の誤対応は、環境の違いによって大きく変化しないため、未学習環境でも機能する。提案法は、従来のモデルベース法で行われていたことすべてを機械学習で置換しようとしないうえ、機械学習が行うタスクを簡略化させることができ、かつ学習の効率を向上させることができる。

4. 検証結果

故意に自己位置推定に失敗するパラメータを用いて行った検証の結果を図2に示す。上図は、真値(赤)と推定値(青)の軌跡であり、中図は真値と推定値間の位置誤差(赤)と角度誤差(青)である。下図は、提案法により推定された信頼度(赤)、RBPFにおける最尤パーティクルに対するCNNの正誤判断の値(緑)、および比較法として利用された拡張モンテカルロ位置推定(AMCL) [11]において位置推定失敗の判断として用いられる値(青)である。AMCL内で利用される値は、ゼロを越えると位置推定に失敗していると判断される。

本図からわかる様に、提案法は推定誤差の増大に伴い、滑らかに、かつ適切に信頼度が増減していることが確認できる。一方でCNNによる正誤判断の値は、大きく変動していることが確認できる。このことから、提案法がCNNによる予測の不確かさの影響を低減させていることが確認できる。またAMCLと比較して、推定された信頼度がより正確に自己位置推定の正誤を説明していることが確認できる。

本実験を行うにあたり、実験環境とは異なる環境で、位置推定成功・失敗データをそれぞれ約5000個作成した。またネットワークは転移学習などを行わずに、スクラッチの状態から学習を行った。結果として、CNNが約90%の精度で正誤の判断を行えることを確認した。この結果からも、学習のタスクが簡略化され、少ないデータで高精度な認識が行えることが確認できる。

5. おわりに

本稿では、著者らが提案した、従来の確率的推論を行うモデルベース法において機械学習を併用した枠組みである、信頼度付き自己位置推定法について述べた[1]。今後は、モデルベース法とE2E学習を融合する様な枠組みの構築を目指す。

参考文献

[1] N. Akai et al. Simultaneous pose and reliability estimation using convolutional neural network and Rao-Blackwellized particle filter. *Advanced Robotics*, 32(17):930–944, 2018.
 [2] A. Kendall et al. PoseNet: A convolutional network for real-time 6-DOF camera relocalization. In *Proc. of the IEEE ICCV*, pages 2938–2946, 2015.

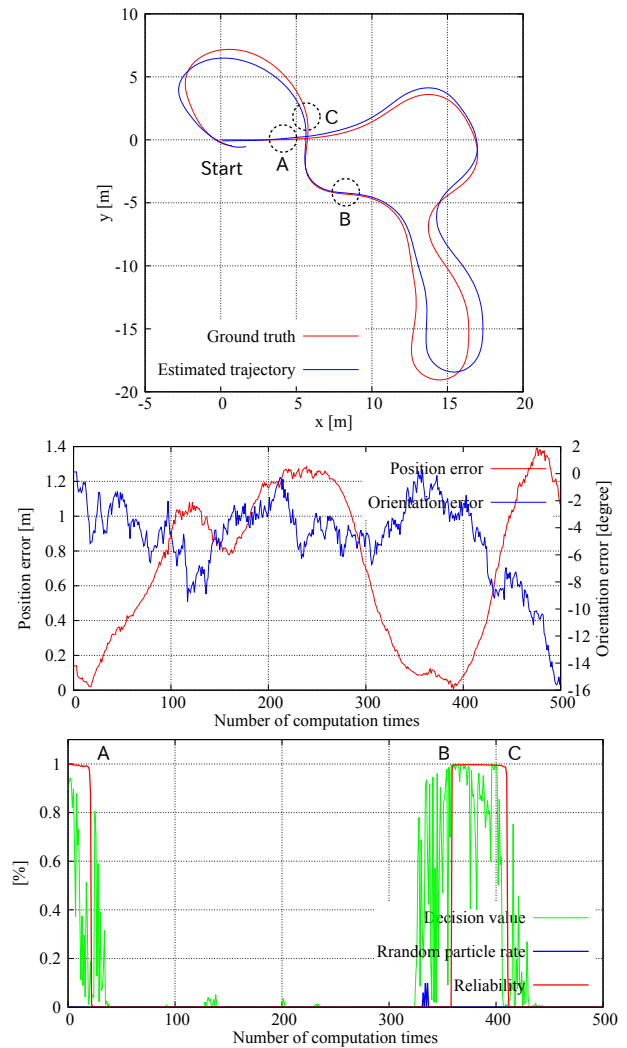


図2 実験結果 [1].

[3] A. Kendall et al. Modelling uncertainty in deep learning for camera relocalization. In *Proc. of the IEEE ICRA*, pages 4762–4769, 2016.
 [4] G. Elbaz et al. 3D point cloud registration for localization using a deep neural network auto-encoder. In *Proc. of the IEEE CVPR*, pages 2472–2481, 2017.
 [5] P. Karkus et al. Particle filter networks: End-to-end probabilistic localization from visual observations. *arXiv:1805.08975*, 2018.
 [6] W. Lu et al. L³-Net: Towards learning based LiDAR localization for autonomous driving. In *Proc. of the IEEE CVPR*, pages 6389–6398, 2019.
 [7] C. R. Qi et al. PointNet: Deep learning on point sets for 3D classification and segmentation. In *Proc. of the IEEE CVPR*, pages 652–660, 2017.
 [8] W. Yuan et al. Iterative transformer network for 3D point cloud. *arXiv:1811.11209*, 2018.
 [9] A. Amini et al. Variational end-to-end navigation and localization. In *Proc. of the IEEE ICRA*, 2019.
 [10] H. Almqvist et al. Learning to detect misaligned point clouds. *J. of Field Robotics*, 35:662–677, 2018.
 [11] J. Gutmann et al. An experimental comparison of localization methods continued. In *Proc. of the IEEE/RSJ IROS*, pages 454–459, 2002.