

屋内動的環境における準動的物体の考慮による逐次 SLAM の改善

Improvement of Sequential SLAM by Considering Semi-dynamic Objects in Indoor Dynamic Environments

○学 押久保 雄平 (中央大学) 秋庭 啓吾 (中央大学)
Sarthak Pathak (中央大学) 正 池 勇勳 (JAIST)
正 梅田 和昇 (中央大学)

Yuhei OSHIKUBO, Chuo University, oshikubo@sensor.mech.chuo-u.ac.jp

Keigo AKIBA, Chuo University, akiba@sensor.mech.chuo-u.ac.jp

Sarthak PATHAK, Chuo University, pathak@mech.chuo-u.ac.jp

Yonghoon JI, JAIST, ji-y@jaist.ac.jp

Kazunori UMEDA, Chuo University, umeda@mech.chuo-u.ac.jp

In this paper, we propose sequential SLAM considering semi-dynamic objects that is robust in dynamic indoor environments. Generally, accuracy of SLAM decreases in dynamic environments and environments with few geometric features. Therefore, this method provides semantic information in addition to geometric information via object detection, and extracts point clouds that are effective for localization by considering the attributes and relationships of the objects. Moreover, this method can extract more features even in environments where there are few static objects and be clearly identify how to use each of object point cloud for SLAM by adopting the concept of semi-dynamic objects. This is expected to improve the accuracy of SLAM.

Key Words: SLAM, Object Detection, Scan Matching, Mobile Robot, Dynamic Environments

1 序論

近年、労働人口の減少に伴い、人からロボットへの代替が進んでいる。空港や工場などの限定された空間では既に自律移動ロボットが数多く導入されており、搬送、警備、清掃などの様々な用途がある。これらの環境でロボットが自律移動するには、事前に作成された高精度な地図を利用することが多い。しかし、工事現場等の頻繁に環境が変化する場所ではその都度地図を作り直す必要があると予想される。

地図構築を行うために Simultaneous Localization and Mapping (SLAM) がしばしば用いられるが、動的環境において、フレーム間での点群の誤対応が多くなることで精度が低下したり、構築した地図に不要な点群が残ってしまったりすることがある。また、形状特徴量に乏しい環境では退化 [1] によって位置合わせがうまくいかず、自己位置推定が失敗することがある。退化とは例えば形状が 2 本の直線の場合、直線に平行な位置合わせは一意に確定できないことである。文献 [2] では、動的物体・静的物体を事前に定義し、深層学習による検出によって動的点群を削除、静的物体を目印として動的環境に対応できるシステムを構築した。しかし、定義した静的物体に PC や椅子などの動く可能性があるものが含まれており、静的である保証が得られないものもあった。動く可能性があるものを全て除去し静的なものだけを活用する方法もあると考えられるが、位置が時系列で完全に変化しない静的な物体は一般に少ないと考えられる。

そこで本研究では、動く可能性があるものでも静的なときは利用するため、準動的物体 [3] という概念を導入し、屋内の動的環境においてロバストかつ高精度な SLAM を提案する。また、準動的物体の定義については、2.2 節で詳しく述べる。この手法により頻繁に環境が変化する場所でも手軽に地図を作成することができ、作成した環境地図は自律移動ロボットの自己位置推定のための事前情報としての活用が期待できる。

2 提案手法

2.1 概要

本手法全体の大まかな流れを図 1 に示す。最初に、RGB-D カメラによって複数フレームで画像と 3 次元点群を取得する。また、異なるフレーム間で得られたホイールオドメトリも取得し、これを位置合わせの計算の初期位置に用いる。次に、画像による物体検出を活用して取得した点群から物体の点群をそれぞれ抽出

する。この物体検出方法については、2.3 節で説明する。その後、動的点群の除去処理を行い位置合わせに効果的な静的点群のみを残す。この方法については、2.5 節で説明する。ここで、静的点群は 3 次元座標情報だけでなく、何の物体であるかの情報と色情報を持っている。この静的点群を用いて ICP (Iterative Closest Point) [4] による位置合わせを行い、回転行列 R 、並進ベクトル T を算出する。最後に R 、 T を用いて自己位置推定、地図構築を行う。

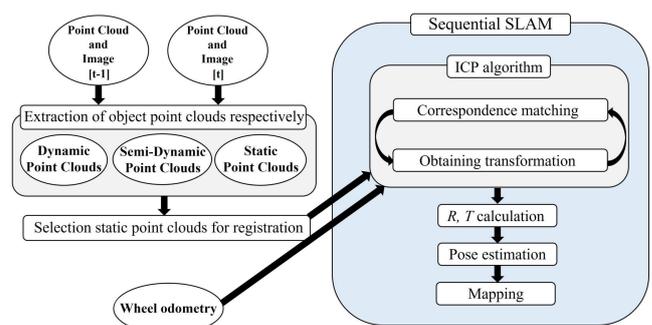


Fig.1 Flow of the proposed method

2.2 定義

前提として、扱う物体を物体の移動する頻度に基づいて、静的物体、準動的物体、動的物体に分ける [3]。人のように頻繁に動くものを動的物体、椅子のように時々動くものを準動的物体、壁などの位置が基本的に不変なものを静的物体と定義する。本研究における実験では人を動的物体、段ボールを準動的物体、ドアを静的物体としている。屋内環境を想定していることから風などの影響は無いとして、準動的物体は動的物体のみによって動くことと仮定する。実際に工事現場で、段ボールや台車が勝手に動くことはなく、動くときは人やロボットなどの動的物体が作用しているといえる。

2.3 物体検出

本手法では、RGB-D カメラで取得した RGB 画像に対して Alexy らが開発した YOLOv4[5] を用いて、物体検出を行った。本研究で検出したい 3 つの物体はあらかじめ学習されていなかったため、新しく学習を行った。筆者が撮影した 175 枚の画像に対して左右反転、ごましおノイズ、明るさの変化などを加えて拡張した 4200 枚の画像を用いて人、段ボール、ドアを学習させた。出力結果の例が図 2 であり、人、段ボール、ドアをそれぞれ検出できていることがわかる。YOLOv4 の推論結果から得られた情報を 2.4 節の物体点群の抽出に用いる。

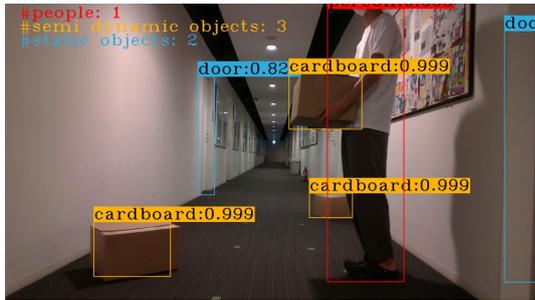


Fig.2 Detection result by YOLO in customized model

2.4 各物体点群の抽出

距離画像は 3 次元点群に変換できる。よって、2.3 節で求めた RGB 画像と距離画像の対応を取ることで、画像中における物体領域（背景を含む矩形領域）に相当する点群を抽出することが可能となる。しかし、そのままでは点群に後ろの背景も含まれてしまうのでダウンサンプリング後にクラスタリングを行う。RGB-D カメラで得られる点群は密であり、そのまま扱うとクラスタリングや位置合わせの計算時間が長くなってしまいうため、計算量削減のためダウンサンプリングを行う。本手法では点群の密度が均等になるように、ボクセルグリッドダウンサンプリングを用いた。また、クラスタリング手法は点間のユークリッド距離を用いるユークリッドクラスタリングを採用し、これを行うことで、ある程度、壁の点群を除外することができた。

画像中において物体が検出されなかった領域に対しても点群に変換をしておき、ダウンサンプリングを行ったあと、メモリに保存しておく。この点群は位置合わせには用いないが、地図構築時に用いる点群として残している。

2.5 準動的物体を考慮した位置合わせに用いる点群の選択

2.2 節でも述べたように、屋内において準動的物体が動いているかは動的物体が起因していると考えられる。よって、準動的物体の近くに動的物体がある場合、準動的物体は動いている可能性がある。本手法では、準動的物体と動的物体の距離に着目することで、準動的物体が各フレームで静止しているかの判断を行う。

3 次元空間を上から見た図 3 で、具体的に説明していく。点がセンサーで得られた点群の各点を表しており、青点がドア、オレンジ点が段ボール、赤点が人に該当する。最初に、物体点群の 3 次元座標の平均をとることでそれぞれ重心を求める。その後、動的物体点群の重心から準動的物体の重心までの距離を求める。動的物体点群の重心と準動的物体点群の重心のユークリッド距離があるしきい値 λ 以上であれば、準動的物体点群を静的点群として位置合わせの計算に用いる。反対にしきい値以下であった場合は、準動的物体点群を動的点群とみなして除去し、位置合わせにも地図構築にも用いない。しきい値 λ の値は経験的に 1.0 m としている。図 3(b) から準動的物体を、重心間距離が λ より小さい d_1 のときは動的、 λ より大きい d_2 のときは静的と判断していることが分かる。ここで、静的と判断された点群は位置合わせ、すなわち自己位置推定に活用する。各フレームで同じ処理を順に行い、隣接するフレームで得られた静的点群を入力とした ICP による自己位置推定、その自己位置を元にした地図構築を繰り返すことで逐次 SLAM ができる。また、ICP を行う前の初期位置

合わせには隣接するフレーム間で得られたホイールオドメトリを使用する。

実データでの、位置合わせに用いる点群の選択前、後をそれぞれ図 4(a) と (b) に示す。図 2 の情報と距離画像を元に、2.4 節の処理をしたものが、図 4(a) である。また、本節の処理によって、図 4(a) で存在していた人の点群と人が持っている段ボールの点群が、処理後には図 4(b) に示すように、取り除かれていることが分かる。これにより、動的点群による SLAM への悪影響を取り除くことができる。

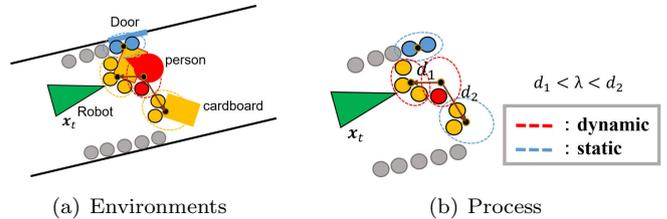


Fig.3 Explanation of selection of static point clouds for registration

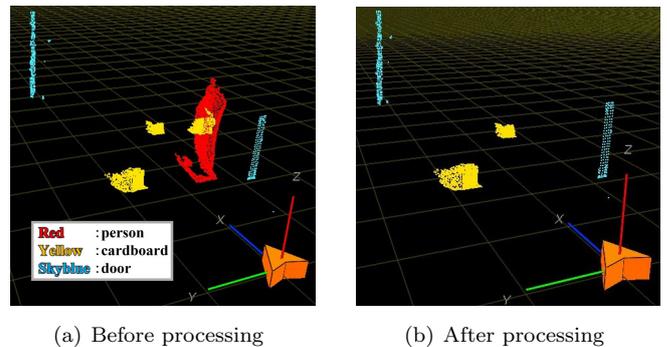


Fig.4 Selection of static point clouds for registration

3 実験

3.1 実験条件

本実験は、動的で退化が起きやすいような場所を想定し、図 5(a) に示す直線的な廊下において人に段ボールを運ばせるシナリオで行った。事前に 2 つの準動的物体と定義した段ボールが設置されており、人は途中で段ボールを床に置く。図 5(c) の番号は各フレームでのそれぞれの物体の位置を表している。ロボットを 1.0 m ずつ、事前に床に貼った目印を元に 5 回直進させ、計 6 回計測を行った。ICP を用いた自己位置推定による修正量が極端に大きい場合は点群の位置合わせが失敗と判断し、ホイールオドメトリのみを位置合わせに用いた。

評価指標はロボットの自己位置推定誤差と地図構築結果とした。自己位置推定誤差は真の位置とのずれをユークリッド距離で、定量的に評価し、人除去処理が行われていない場合、人除去処理が行われている場合 [6]、提案手法により動的点群の除去及び位置合わせに効果的な静的点群の抽出が行われている場合で比較を行った。いずれも位置合わせには Point Cloud Library (PCL) で実装されている ICP を利用した。地図構築結果は、不必要な点群が含まれているかどうかを定性的に評価し、人除去処理が行われていない場合、提案手法を比較した。提案手法と人除去 SLAM [6] の地図構築結果の違いは動的・準動的物体をどのように定義していくかによっても変わるため、ここでは省略する。また、動的と判断された準動的物体に関しては位置合わせ、地図構築どちらでも除去するという点で問題ないが、静的と判断された準動的物体を地図構築に用いるかどうかに関して、次のように

考えられる。準動的物体は、地図構築時には静的で任意の位置に存在したとしても、時系列の変化と共に場所が移動する可能性があるため、準動的物体が存在する地図でロボットの自己位置推定を行ってしまうと準動的物体がSLAM時と異なる位置に移動していた場合、自己位置推定精度が低下してしまう恐れがある。以上の理由から、本手法で準動的物体であるものは、SLAM時の静的、動的に関わらず地図構築の点群には用いなかった。

本研究では、移動ロボットとして Adept MobileRobots 製の Pioneer 3-AT を使い、図 5(b) に示すように RGB-D カメラである Intel 製 RealSense LiDAR Camera L515 を固定した。制御はこの移動ロボットにノート PC を載せ、筆者が行った。

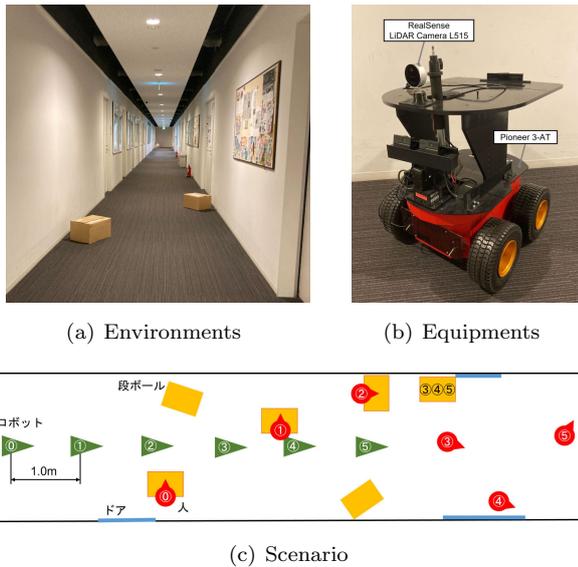


Fig.5 Experimental conditions

3.2 実験結果

まず、ロボットの自己位置推定誤差の結果を図 6 と表 1 に示す。この結果から、本実験のシナリオにおいて、提案手法は比較手法に対して自己位置推定の精度が高いといえる。これは、提案手法が他手法に比べてより多くの静的点群を自己位置推定に活用できたためと考えられる。しかし、どの手法においても 5 フレーム目で、ICP による点群の位置合わせが失敗している。提案手法以外の 2 つは、累積誤差が大きくなり、ローカルミニマムに陥り、正しい位置合わせができなくなったと考えられる。提案手法に関しては、5 フレーム目でセンサの視野角のため段ボールが観測できず、位置合わせに用いることができる物体の数が減ってしまったことが原因と考えられる。

次に、構築地図結果を比較する。人除去処理をしていない場合と提案手法における地図構築結果を図 7 に示す。図 7(a) は構築した点群に人と段ボールが含まれており、図 7(b) から人と段ボールの点群はほぼ除去できていることがわかる。このことから、提案手法は動的環境でもロバストに地図構築できることが分かった。

Table 1 Robot Pose Error for each method

Method	0	1	2	3	4	5
Only ICP	0.000	0.173	0.271	0.455	0.533	0.440
Akiba's[6]	0.000	0.177	0.280	0.189	0.276	0.181
Ours	0.000	0.103	0.143	0.106	0.056	0.077

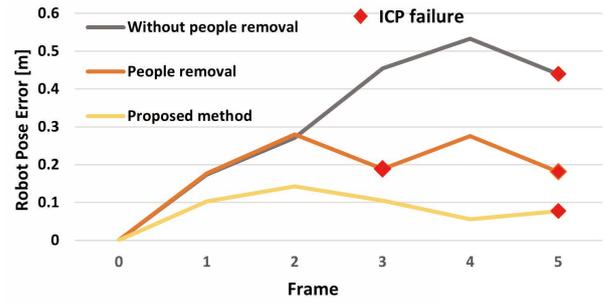


Fig.6 Robot pose error

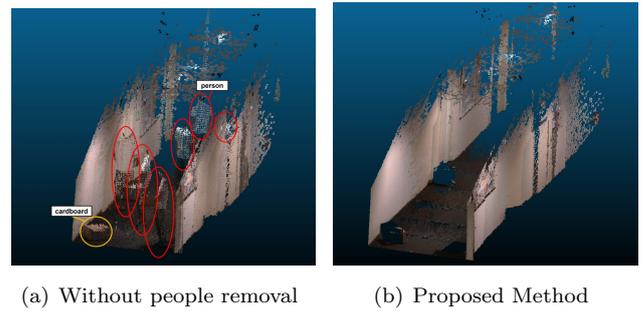


Fig.7 Maps by each method

4 結論

本研究では、屋内の動的環境において、YOLO を用いて検出した物体の属性と 3 次元空間における物体間の距離の関係から位置合わせに効果的な静的点群を抽出する ICP-SLAM を提案し、SLAM に用いることの有用性を示した。

一方、逐次的な処理のため長距離使用では誤差の累積、物体のオクルージョンによる点群地図の欠損などが本手法の問題として挙げられる。以上より、今後の展望として 2 つ存在する。1 つめは、スキャンマッチングを工夫したり、準動的物体の扱いを考慮した上で逐次的処理に加えてループクローリング等の全体を包括的に最適化したりすることである。2 つめは、欠損を補完するような処理を加えるかオクルージョンに対応した経路生成を行っていくことを検討する。また、様々なシナリオでの実験を重ね、本手法の有効性を更に検証していきたい。

参考文献

- [1] 友納 正裕, "SLAM 入門", 株式会社オーム社, p.20, 2018.
- [2] Fanguwei Zhong, Sheng Wang, Ziqi Zhang, China Chen, Yizhou Wang, "Detect-SLAM: Making Object Detection and SLAM Mutually Beneficial", *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.
- [3] Hongjun Zhou, Shigeyuki Sakane, "Localizing objects during robot SLAM in semi-dynamic environments", *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2008.
- [4] Paul J. Besl, Neil D. McKay, "A Method for registration of 3-D shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp.239-256, 1992.
- [5] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection", arXiv:2004.1093, 2020.
- [6] 秋庭 啓吾, 鈴木 龍紀, 池 勇動, Sarthak Pathak, 梅田 和昇, "YOLO を用いた人除去処理により動的環境に適用した屋内 SLAM," 第 40 回人間情報学会オーラルセッション講演集, pp.5-6, 2021.