

建設現場における資材搬送ロボットのための単眼カメラ画像（視覚情報）を用いた構造理解とトポロジカルナビゲーション

Structure Understanding and Topological Navigation Using Monocular Camera Images (Visual Information) for Material Transport Robots in Construction Sites

システム工学群

知能ロボティクス研究室 片山 大輝

1 緒言

現在,日本では少子高齢化に伴い,深刻な働き手不足に陥っている.その中でも建設業は総務省の「労働力調査」⁽¹⁾によると現場で働く技能職の就業者数が減少傾向にある.そこでロボットによる作業の代替が必要であると考え.その中でも単純作業かつ重労働である資材の運搬作業はニーズが多く研究も進められている.

先行研究⁽²⁾ではARマーカーを用いた方向誘導による資材搬送ロボットのナビゲーション方法が提案されており,建設現場で有用性の実証を行っている.しかし,ARマーカーを設置する方法では現場が大きくなるに伴い,設置するARマーカーの数が増え,資材搬送ルート変更のたびに回収・再設置の手間がかかる.更に照明による影響や物に隠れて見えなくなるなど視認性の問題がある.本報告では,ARマーカーの代わりに資材搬送ルートの動画を撮影する.そこから通路構造を判断しナビゲーションする方法を提案する.

2 トポロジカルナビゲーション

本研究ではトポロジカルマップを用いてナビゲーションをおこなった.トポロジカルマップ⁽³⁾とは,環境中の特徴的で見分けが付きやすい地点とその繋がりをグラフ状に表現したものである.本研究では特徴的な地点(ノード)を通路構造とし,そのノード同士の繋がりをエッジとした.図1はT字路をトポロジカルマップにしたものである.番号のついた円は各ノードであり,それぞれ通路構造を有している.ノード同士を結ぶ線がエッジである.

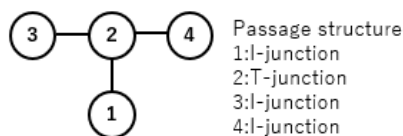


Fig. 1 Topological maps and passage structures

ノード1からノード4に移動したい時はノードのI字路を直進して,ノード2のT字路を右に曲がることでノード4にたどり着く.このように通路構造を表すノードとノード上での移動動作があればナビゲーションが可能である.

3 ナビゲーション情報の生成

3.1 ナビゲーション情報を生成する流れ

以下のような流れでナビゲーション情報を生成した.

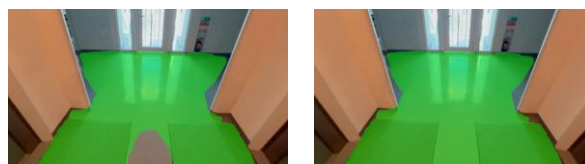
1. セマンティックセグメンテーションにより床マスクを得る.
2. 床マスクのノイズ低減及び形状単純化処理
3. スケルトン化により床マスクの中心線を抽出
4. 分岐点と分岐方向の抽出
5. オプティカルフローによる回転の検出

6. 通路構造を判別

7. 通路構造とそれに対応する移動動作の判別

3.2 セマンティックセグメンテーション

床領域を得るためにセマンティックセグメンテーションを用いた.本研究では,NVIDIAが公開する事前学習済みモデルsegFormer-B0⁽⁴⁾(ADE20K⁽⁵⁾で学習済み)を用いた.図2はセマンティックセグメンテーションにより得た床マスクを元画像に重ねたものがある.緑色の部分が床マスクである.左がモデルから得た床マスクであり,右側がその床マスクにノイズ低減や形状単純化処理を行ったものである.

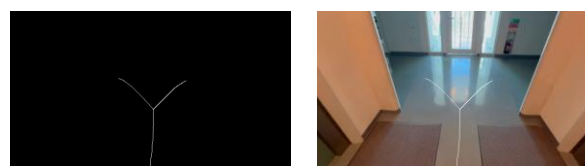


Before processing After processing

Fig. 2 Original image with floor mask

3.3 スケルトン化

図3はスケルトン化により得た床マスクの中心を表すスケルトンである.白い線がスケルトンである.左がスケルトンのみ,右が元画像にスケルトンを重ねたものである.



Skeleton Original image with skeleton overlaid

Fig. 3 Skeleton and original image

スケルトン化は以下の式より,画素ごとに床以外の画素までの最短距離を求める.その距離値局所的に最大値を取る画素を取り出すことでスケルトンを抽出できる.

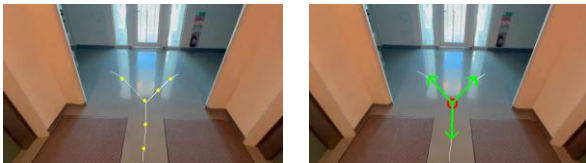
$$f(i,j) = \begin{cases} 0 & (\text{床以外}) \\ \infty & (\text{床}) \end{cases} \quad (1)$$

$$g(i,j) = \min_{1 \leq k \leq N} ((i-k)^2 + f(k,j)) \quad (2)$$

$$d^2(i,j) = \min_{1 \leq l \leq N} ((j-l)^2 + g(i,l)) \quad (3)$$

3.4 分岐点と分岐方向の抽出

スケルトンを基に分岐点と分岐方向を抽出した。スケルトンの画素をクラスタリングにより減らし、その後画素から最も近い画素を繋いでいく。繋いでいる画素が最も多いものを分岐点とした。その分岐点からドーナツ状の円と円の内側にある画素へのベクトルを調べ、そこから分岐の向きを抽出した。図4の左側がスケルトンの画素をクラスタリングにより減らしたものである。図4の右側が分岐点の位置と分岐の方向を示したものである。白い線がスケルトン、黄色い点がクラスタリングにより減った画素の位置を表す。赤い円は分岐点、緑の矢印が分岐の方向である。これにより、写真の通路構造はT字路だと分かる。



After clustering Branching points and branching directions

Fig. 4 Extraction of branching points and branching directions

3.5 通路構造とそれに対応する移動動作の判別

1. 連続する同じ通路をまとめる。(回転イベント前後は別)
2. 分岐5以上または連続数が小さい通路構造は省く。
3. 短い範囲で繰り返す通路構造パターンは片方の通路構造に合わせる。(分岐が多い方)
4. 回転開始の直前にある分岐路に回転イベントを割り当てる。ない場合はI字路がL字路だったとする。
5. 回転が終わるまでは通路形状がブレやすいから削る。
6. 回転イベントがないL字路をI字路にする。
7. 割り振られた回転イベントを基に移動動作を判別する。

- 例) 回転イベントありの十字路
→十字路を右/左に曲がる。
回転イベントなしの十字路
→十字路を直進

3.6 ナビゲーション情報の生成実験

ナビゲーション情報の生成実験を行った。図5は撮影したルート上の地図である。青線が動画を撮影したルートである。赤点から撮影を開始して黄色い点で撮影を終了した。動画はiPhone13 超広角レンズ (×0.5) の720p, 30fpsで撮影した。

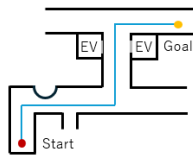


Fig. 5 Shooting route

1326フレームの動画を取得した。ここから6フレームごとに解析して、ナビゲーション情報を生成した。表1は実験結果をまとめたものである。通路構造とその通路での移動動作を記している。移動動作の“r”, “l”, “st”はそれぞれ「右に曲がる」, 「左に曲がる」, 「直進」を表している。

結果より、通路構造とそこでの移動動作を抽出したので、ナビゲーション情報を生成したと言える。

Table 1 Navigation information

Step	1	2	3	4	5	6
Passage structure	L	I	T	I	T	I
Action	r	st	l	st	r	st

r:Turn right, l:Turn left, st:Go straight

4 ナビゲーション実験

4.1 実験方法

ナビゲーション情報とカメラのリアルタイム映像を基に指定のルートを移動する実験を行った。カメラで映像を取り込み、ナビゲーション情報の生成と同様に通路構造を判断した。次にその通路構造での移動動作を指示した。移動の様子を6秒撮影して、指示を出し、移動の様子を6秒撮影、このサイクルを繰り返し目的地に向う実験を行った。指示される移動動作は1. 前進, 2 右旋回, 3 左旋回, 4. 一時停止の4つである。目的地に着くとナビゲーションは終了する。

Table 2 Navigation experiment success rate

Achievement rate	6/7(85.7%)
------------------	------------

4.2 ナビゲーション実験の実験結果と考察

表2はナビゲーション実験の結果である。実験は7回行い、そのうちの6回は目的地にたどり着いた。

失敗したときはT字路の分岐点より前に旋回の指示を出してしまった。この原因は「カメラを体より前に構えたことで実際の位置より前にいると認識したこと」、「トポロジカル表現では分かれ道の始まりも終わりも同じと扱われること」の2点だと考えられる。

全体を通して通路構造の判別の正解率は87%であった。

5 結言

本報告ではルート動画を基に通路構造を判別してナビゲーションする方法を提案し、検証した。ARマーカーを用いた方法と比べ、少ない手間と準備物でナビゲーションに成功した。今後は建設現場でロボットをナビゲーションすること、道を間違えた時の修正方法、動画撮影者によらないナビゲーションを作ることが課題となる。

文献

- (1) 総務省統計局 “労働力調査” 2002-2024
<https://www.stat.go.jp/data/roudou/index.html>
閲覧日:(2026/02/05)
- (2) 楊光, 王碩玉, 岡村甫, 猪野真吾, 奥畑一男, 溝渕宣誠, “資材自動搬送ロボットの自立化制御”, 日本ロボット学会誌, Vol. 39, No8, (2021), pp755-758
- (3) 田崎勇一, “トポロジカルな地図表現を用いた移動ロボットのナビゲーション”, 日本ロボット学会誌, Vol33, N010(2015), pp773-778
- (4) Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P., “SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers”, Advances in Neural Information Processing Systems, Vol. 34 (2021), pp. 12077–12090.
- (5) Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., and Torralba, A., “Semantic Understanding of Scenes through the ADE20K Dataset”, International Journal of Computer Vision, Vol. 127, No. 3 (2019), pp. 302–321.