

## 研究速報

## 移動ロボットによるサブゴール間巡回行動の学習

森 実 克<sup>\*1</sup> 山 田 誠 二<sup>\*2</sup> 豊 田 順 一<sup>\*1</sup>

## Robot Learning for Moving between Subgoals

Katsu Morizane<sup>\*1</sup>, Seiji Yamada<sup>\*2</sup> and Jun'ichi Toyoda<sup>\*1</sup>

In this research, our goal is that a mobile robot learns to move between subgoals without human interactions. The robot does not have knowledge on the environment: positions of subgoals and obstacles. Unfortunately, the robot learning is unavoidably influenced by the errors in a real world. We suggest how to learn to move efficiently between subgoals by resetting errors. In our system, we define the distinctive place on which the movement control of a robot changes from sensor-based to coordinates-based. The distinctive place is identified only with local information, and the cumulative errors through the movement between two subgoals are reset on distinctive places. First, the robot moves to search for subgoals. Next, it moves between subgoals repeatedly, and learns to move efficiently. We made some experiments in a real environment. In the experiments, we found out that the robot's movement was gradually improved by learning.

*Key Words:* Mobile Robot, Learning, Navigation, Distinctive Place

## 1. はじめに

移動ロボットのナビゲーションに関する研究では、タスクはロボットが一度目的地に到着した時点で終了するものが多く、経験によるパスの最適化などの改善はない[2]。また、実際にロボットが使用される環境では、障害物やサブゴールの配置状態などの環境要素が動的に変化することも考えられる。さらに、ロボットの移動を手助けする教師の存在をいつも期待できるわけではない。このような背景より、本研究では、サブゴールや障害物の配置が不明で、かつ、変化しうる実環境で、人間とのインタラクションなしに、予め設定されたサブゴール間の定常的で効率的な2次元の移動をロボットに学習させることを目的とする。ここでいうサブゴールとは、環境内で移動ロボットがその上に来た場合のみ、認識・同定できるような場所のことである。また、我々は、上記のようなタスクを、比較的精度の低いロボットで実現することに関心がある。

ロボットの位置決定における累積誤差の解決策として、ロボットの運動を数学的にモデル化する方法やビーコンやランドマーク等を環境に埋め込む方法がよくとられるが、我々の目的は、このようなモデル化や、付加的な誘導システムに頼らずに誤差の影響

を処理することにある。我々のアルゴリズムでは、障害物のあるコーナーをディスタクティブプレース[1]として登録し、累積誤差の影響のリセットおよびパスの学習に利用する。パスの学習は、サブゴール間を巡回しながら累積誤差を徐々に修正するかたちで進められる。

松村らは、プロダクションルールの合成による、障害物回避の際の無駄のない経路学習法を提案している[5]。ただし、この研究では、主眼は障害物回避経路の学習にあり、障害物回避後、目標地点までの移動経路については、誤差について考慮されていないこともあり学習できない。

また、近年、位置評価に関する誤差を扱った、センサベースのパスプランニング・アルゴリズムが数多く開発されてきた。

登尾らは、デッドレコニングエラーを考慮した、デッドロックフリーなパスプランニング・アルゴリズムに関する研究を行った[4]。しかし、そのナビゲーションは、誤差が存在してもゴール近傍のある範囲内には到達できるというもので、ゴールに到達できるように学習を行うことはない。

LOGnet[3]では、誤差の影響への対処として、ナビゲーションの際にパスを外れた場合、ランダムウォークによって既知のLOGに移動し、そこから別のパスを試すという方法をとっており、学習によりLOG間のコマンド系列を改善するようなことはない。これに対し本研究では、失敗したら別のパスを試すのではなく、1つのパスをうまく迎れるように学習する。

鈴木らは、連想データベースを移動ロボットの位置同定や障害物回避に利用している[6]。この手法では、学習過程で人間のオペ

原稿受付 1996年7月19日

<sup>\*1</sup>大阪大学 産業科学研究所

<sup>\*2</sup>東京工業大学 大学院総合理工学研究科 知能システム科学専攻

<sup>\*1</sup>ISIR, Osaka University

<sup>\*2</sup>Department of Computational Intelligence and System Science, Tokyo Institute of Technology

レータの介入が必要となる点で、我々の目指す教師なし学習とは異なる。

## 2. マイクロマウス：実験で使用するロボット

実験に使用するロボットは、マイクロマウスと呼ばれる、8つの赤外線近接センサ、回転角度検出機構および移動距離計を装備した小型の移動ロボットである。ロボットのサイズは、 $10 \times 18 \times 13$  cm である。ロボットは、左右に独立した駆動輪を2つ持つ。前後には、それぞれ逆方向に切れる操舵輪を2つ持ち、これによってロボットは回転することができる。

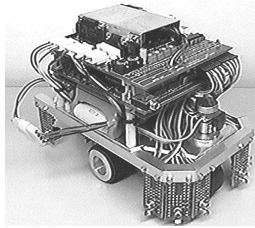


Fig. 1 The robot used for our experiments

## 3. ディスティンクティブプレース

ロボットは、初期移動によるサブゴール探索で得たサブゴールのおよその位置に基づき、サブゴール方向に障害物のセンサ反応がない場合は座標ベースで制御され、反応がある場合は、センサベースの行動制御による壁沿い移動で、これを回避する。

障害物回避中、移動ロボットの行動制御がセンサベースから座標ベースに切り替わるコーナーをディスティンクティブプレース [1](以下 DP) として定義する。移動ロボットは、DP でセンサベースの壁沿い移動による障害物回避を中止し、目的のサブゴール方向への座標ベースの直線移動を再開する。DP は、障害物の凸コーナーで、局所情報を用いて同定される。

本研究では、障害物のコーナーは、障害物方向のセンサエコーの突然の消失によりロボストに検出できると仮定する。この場合、障害物回避中、現在地から見たサブゴールの方向が障害物によって遮られているかどうかは、障害物の凸コーナーに差し掛かった時点でのみ判定すれば充分である。このような判定方法の利点は、壁沿い移動中、常に变化する目的のサブゴール方向を監視する必要がなくなることで、センサの取り付け機構が単純化でき、かつ計算負荷も減る点にある。

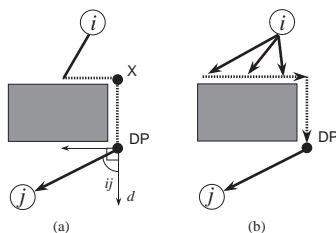


Fig. 2 Distinctive place

本研究では、環境全体に設定されたグローバル座標系と、各サブゴールを原点としたローカル座標系を扱う。グローバル座標系は、ロボットが最初に環境に置かれた際、ロボットの、前方を  $y$  軸の + 方向 (方位  $90^\circ$ ) に、右方向を  $x$  軸の + 方向 (方位  $0^\circ$ ) に、重心を原点にした座標である。ローカル座標系は、サブゴールを原点として方位  $0^\circ$  を  $x$  軸の + 方向に、方位  $90^\circ$  を  $y$  軸の + 方向に設定した座標である。グローバル座標系は、ロボットの現在地及び、初期移動で得たサブゴールの位置を表すためにのみ使用される。

DP の発見と登録について、Fig.2で説明する。Fig.2(a)で、 $i$  と  $j$  はサブゴールを表す。コーナー X では、ロボットから見たサブゴール  $j$  の方向が障害物によって遮られているので、ここから  $j$  方向に直接移動することはできない。つまり、コーナー X は DP ではない。ロボットがあるコーナーに到達し、かつ目的のサブゴール方向が障害物によって遮られていない場合、ロボットはそのコーナーを DP として登録し、付随して以下のデータを記録する。

- (a) スタート地点となるサブゴール (図中  $i$ ) を原点とする DP のローカル座標。
- (b) DP 上での目的のサブゴールの方向と直前に沿ってきた壁の方向のなす角度  $\theta_{ij}$ 。
- (c) サブゴール  $i$  から DP に到達するまでに通過した障害物のコーナーの数  $N_{DP}$ 。

(b) において、壁の方向は、壁沿い移動中のロボットの移動方向を一定間隔で計測し、その平均を計算することで求める。目的のサブゴール  $j$  の方向は、サブゴール  $i$  を原点とするローカル座標系において、ロボットの座標とサブゴール  $j$  の座標から算出する。(c) は、DP の同定のために利用する。

DP に到達したら、現在の座標データと方向データを、前回その DP に到達した際に記録したデータに更新する。これにより、移動中に累積した誤差を、前回到達時の状態に戻すことができる。

また、Fig.2(b) は、 $i$  からの移動の際に、累積誤差の影響で、障害物に行き当たる箇所が前回と異なる場合 ( $i$  からの3つの矢印) でも、そのズレが最初に到達する障害物の1辺の中に収まるものであれば、上記 (c) の情報をもとに同じ DP にロボストに到達できることを示している。

## 4. 学習手続き

学習手続きは、主に3つの段階に分かれる。第1段階では、サブゴールを探索するための初期移動を行う。第2段階では、DP の発見と登録を行う。第3段階では、DP を利用したパスの洗練を行う。以下では、簡単のためサブゴール2個の場合で説明する。

### 4.1 初期移動 (サブゴールの探索)

ロボットには、サブゴールのグローバル座標は与えられていないので、それを獲得するために環境内を探索する必要がある。ただし、ロボットの移動には誤差が含まれるため、ここで得られる位置は不正確である。以下に、サブゴール探索のための初期移動の手続きを示す。

#### Step1: 初期移動

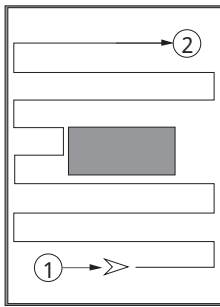


Fig. 3 Initial movement - search for subgoals

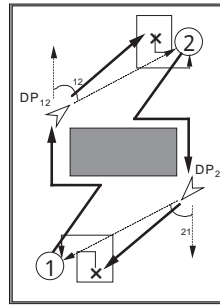


Fig. 4 Discovery and registration of distinctive places

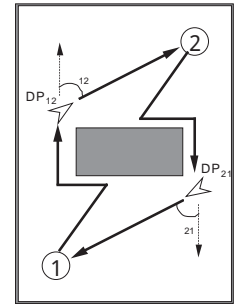


Fig. 5 Path refinement

- (1-1) 2つのサブゴールに到達できたら、次節の Step2 へ。
- (1-2) ロボットは、任意の方向に直進する。
- (1-3) サブゴールに到達したら、そのグローバル座標を記録して (1-1) へ。
- (1-4) ロボットの移動方向が障害物に遮られた場合、壁沿い移動を一定距離行う。回避の際、左右どちら方向に移動するか(ローカルディレクション)は予め決められている。その後、ロボットは壁に対して直角に向きを変え、直進を再開する。最後にローカルディレクションを逆方向に設定し、(1-3) へ。

壁沿い移動を行う距離は予め決められている。実験では、20cm に設定した。Fig.3は、初期移動の様子を示している。なお、この初期移動は完全なものではなく、環境を全探索する保証はない。

4.2 DP の発見と登録

ロボットは、初期移動によって得られた情報から判明した、ロボットから見た目的のサブゴール方向が、障害物で遮られていない場合は、その方向に直進する。遮られている場合は、目的のサブゴール方向が障害物で遮られなくなるまで壁沿い移動によって障害物を回避する。壁沿い移動の際、ローカルディレクションは予め決まっているものとする。Fig.4は、DP の発見および登録作業を示す。詳細な手続きを以下に示す。

Step2 : DP の発見と登録

- (2-1) Step2 で、サブゴール間を 1 往復した場合、次節の Step3 へ。そうでない場合、もう 1 つのサブゴールを目標とする。
- (2-2) 目標のサブゴール方向に直進する。
- (2-3) 障害物に行き当たった場合、障害物を壁沿い移動で回避する。そして、ロボットから見た目的のサブゴール方向が障害物によって遮られないコーナーにさしかかったら、そのコーナーを DP として登録し、(2-2) へ。
- (2-4) 現在位置のローカル座標が目的のサブゴールのローカル座標と一致した場合。
  - (2-4-a) 実際にはそのサブゴール上に到達していない (Fig.4の x 印の場所) とき、近傍探索によってサブゴールの探索を行い、(2-4-b) へ。
  - (2-4-b) 実際に目的のサブゴールに到達したとき、道中で登録された DP における角度情報 (Fig.4の  $\theta$ ) を再計算し更新して、(2-1) へ。

ここで、「近傍探索」とは、距離  $d$  直進、左  $90^\circ$  回転、 $d$  直進、左  $90^\circ$  回転、 $2d$  直進、左  $90^\circ$  回転、 $2d$  直進、左  $90^\circ$  回転、 $3d$  直進、... というように、徐々に探索範囲を広げていく移動を意味する。距離  $d$  は、車体幅と等しくなるように設定した。

4.3 パスの洗練

Step3 でのプロセスは、Step2 とほぼ同じである。Step2 との相違点は、登録された DP の情報をパスの洗練に利用することである (Fig.5)。

Step3 : パスの洗練

- (3-1) 2つのサブゴールへの到達に、近傍探索の必要がなくなった場合、終了。そうでない場合、もう 1 つのサブゴールを目標とする。
- (3-2) 目標のサブゴール方向に直進する。
- (3-3) 障害物に行き当たった場合、障害物を壁沿い移動で回避する。そして、DP として登録されているコーナーにさしかかった場合、ロボットの座標および向きに関するロボットの内部情報を、DP に登録されているものに更新する。さらに、登録されている DP の角度  $\theta$  に基づいて進行方向を変え、(3-2) へ。
- (3-4) 現在位置のローカル座標が目的のサブゴールのローカル座標と一致した場合。
  - (3-4-a) 実際にはそのサブゴール上に到達していないとき、近傍探索によってサブゴールの探索を行い、(3-4-b) へ。
  - (3-4-b) 実際に目的のサブゴールに到達したとき、道中で登録された DP における角度情報 (Fig.5の  $\theta$ ) を再計算し更新して、(3-1) へ。

学習が進むに従って、近傍探索の距離が短くなり、最終的に近傍探索の必要がなくなった時、本手法での最も効率的な移動経路を学習したことになる。

5. 実 験

我々は、サブゴール 2 つ、障害物 4 つの環境 (Fig.9) でいくつかの実験を行った。環境は、 $180 \times 144 \times 13$ cm (縦  $\times$  横  $\times$  高さ) の外壁で囲まれており、障害物として、 $36 \times 18 \times 13$ cm の直方体を 3 つ、 $54 \times 18 \times 13$ cm の直方体を 1 つ配置した。また、サブゴールとして、直径 18cm の円形プレートを使用した。今回の実

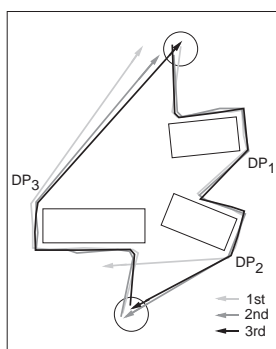


Fig. 6 The trajectories of the robot

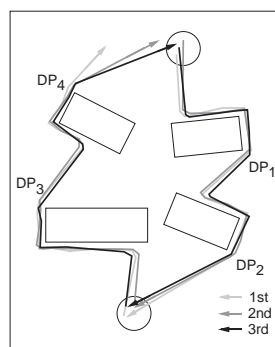


Fig. 7 The adaptation for the appearance of an obstacle

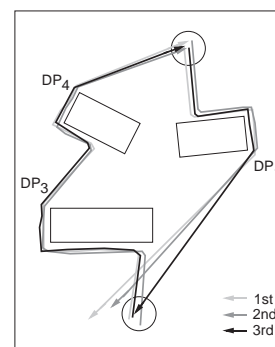


Fig. 8 The adaptation for the disappearance of an obstacle

験では、ロボットは、サブゴールへの到達をオペレーターからの信号によって認識する。なお、障害物の形状は直方体である必要はなく、その他の多角形でもよい。また、数を増やすことも可能である。しかし、初期段階の研究対象としては、比較的単純な環境が適当であると考え、今回は上記のような環境で実験を行った。

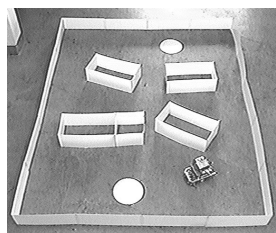


Fig. 9 Experimental environment

Fig. 6は、静的環境におけるロボットの移動軌跡であり、学習によりパスが徐々に短縮されていることがわかる。ただし、図中では近傍探索の軌跡は省略してある。この環境でのサブゴール間の移動に、近傍探索の必要がなくなったところで障害物を1つ追加した (Fig. 7)。障害物が追加されたことにより、図中下側のサブゴールから上側のサブゴールへの移動途中に新たに追加登録された  $DP_4$  の角度が、3周目に近傍探索を必要としなくなる値に収束した。さらに、今度は障害物を1つ除去した (Fig. 8)。これにより、図中上側のサブゴールから下側のサブゴールへの移動途中の2つめの DP (Fig. 7 の  $DP_2$ ) がなくなり、残った DP (Fig. 8 の  $DP_1$ ) の角度は、3周目に近傍探索を必要としなくなる値に収束した。このように、ロボットは、障害物の出現や消失といった環境変化に適応できることがわかる。

## 6. 考 察

初期移動による環境内の探索には、オペレーターによるロボットの内部表現を考慮した環境情報の入力の手間を省くという利点がある。しかし、現在のところ初期移動の手続きはヒューリスティックであるため、サブゴール探索が常に成功する保証はない。

また、本研究では、基本的に同じサブゴール間を移動する際、毎回同じ障害物に到達した場合に学習効果が期待できるが、誤差が大きく、前回移動時に到達したのと異なる障害物に到達した場合、ロボットは登録済みの DP に到達できない。以前に登録したのと異なる DP に到達した場合は、その DP を新しい DP として登録し、その DP での進路変更角度を、以降の学習過程で徐々に修正していく。

現在のところ、誤差のモデル化が困難等の理由により、本アルゴリズムの収束性は証明できていないが、今後、完全性を保証する方向への発展を考えている。

DP をロバストに検出するためには、障害物のコーナーの形状は、センサエコーの突然の消失が起こる程度に鋭角なものに制限される。障害物がそれ以外の形状の場合、DP をロバストに検出するにはより高精度なセンサが必要となる。

## 7. ま と め

本論文では、障害物を含んだ未知の環境内で、移動ロボットがサブゴール間を行き来するための学習手続きを提案し、実機による実験を行った。実験では、ロボットがビーコンのような付加的な誘導システムを用いることなく、サブゴール間の移動を行えるようになった。また、変化する環境にも適応することができた。今後は、初期移動の手続きを改良し、複数の障害物とサブゴールを含むより複雑な環境での実験を行う予定である。

## 参 考 文 献

- [1] B. Kuipers and Y. Byun: "A Qualitative Approach to Robot Exploration and Map-learning," Proc. of Workshop on Spatial Reasoning and Multi-Sensor Fusion, pp. 390-404, 1987.
- [2] J. Latombe: Robot Motion Planning, Kluwer Academic Publishers, 1991.
- [3] P. Malkin and S. Addanki: "LOGnets: A Hybrid Graph Spatial Representation for Robot Navigation," Proc. of the eighth AAAI-90, pp. 1045-1050, 1990.
- [4] H. Noborio et al.: "On-Line Deadlock-Free Path-Planning Algorithms in the Presence of a Dead Reckoning Error," Proc. of the 1995 IEEE Int. Conf. on SMC, pp. 483-488, 1993.
- [5] 松村, 菅沼, 伊藤: "障害物回避における知識の利用", 第6回知識工学シンポジウム, pp. 33-36, 1987.
- [6] 鈴木, 大西, 有本: "連想データベース自己組織化と自律移動ロボットへの応用", 日本ロボット学会誌, vol. 5, no. 6, pp. 3-13, 1987.