

# 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\*

山田 誠二†

豊田 順一\*

\*大阪大学 産業科学研究所

†東京工業大学大学院 総合理工学研究所

## Robot Learning for Moving between Subgoals

○Katsu MORIZANE\*

Seiji YAMADA†

Jun'ichi TOYODA\*

\*Osaka University

†Tokyo Institute of Technology

### 1 はじめに

移動ロボットのモーションプランニングに関する最近の研究では1), 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, 経験による改善はない. 実際の応用の観点では, このようなタスクは不十分である. ロボットには定常的な環境内の移動が要求されることもありうるので, 探検を通してその移動をより効率的なものにできることが望ましい. さらに実環境では, 学習の手助けをしてくれる教師はいない. 以上のようなことから, 我々は, 人間とのインタラクションなしにロボットにサブゴール間の効率的な移動を学習させることを目的とする.

実環境では, センサーデータは誤差を含み, さらに, 例えば障害物やサブゴールの位置などの環境の要素は変化しうる. ロボットには, このような誤差を扱い, かつ, 環境の変化に適応することが求められる. 我々は, 効率的に移動するための学習は, 累積誤差の影響をある時点でリセットすることによってなされるものと考え. 本研究は障害物の存在する道の実環境内で行動する移動ロボットがサブゴール間を行き来するようになるための学習手続きに関心がある. サブゴールは, 環境内で移動ロボットがその上に来た場合のみ, 認識し同定できるような場所のことである. しかし, サブゴールの位置はあらかじめ知ることはできず, また, 変化しうる.

実験で使用する実機は, マイクロマウス (Fig.2) と呼ばれる小型の移動ロボットである. マイクロマウスは, 8つの赤外線近接センサー, 回転角度検出機構, および距離計を装備している. ロボットの現在位置の評価はデッドレコニングによって行われる.

近年, ロボティクス分野で, 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.

LOGnet 3) では, DPは全方向のソナーパターンによって表されている. システムは, 回転を伴うパターンマッチングによってDPを同定する. 実環境での実験が行われているが, パターンマッチングのコストは高い. その上, LOGnetではデッドレコニングによる誤差を扱っておらず, また, 学習により行動を改良することもしない.

登尾は, 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5). これらのアルゴリズムは, 任

意の形状の障害物を扱うことができる. しかし, デッドレコニングエラーの存在を仮定しているにも関わらず, スペシャルポイントへの再到達がわかるとしている. 一般に, ロボットの位置をデッドレコニングで評価するシステムの場合, 誤差修正の手掛かりとなるような付加的な仕組みなしに, ある地点に再び戻ったことを知ることはできない.

実世界では, 移動ロボットに関していくつかの評価誤差の影響がある. このうちのいくつか, 例えば, 距離や方向に関する評価誤差が, デッドレコニングによる位置評価に大きく影響する. それゆえ, ロボットの位置評価には, ビーコン4)等の追加手段を必要とする. 我々の目的は, このような付加的な誘導システムを必要としない学習手続きの開発である.

学習手続きの大筋は以下の通りである. まず, ロボットは初期の移動ルールに従って, 環境内を移動し, サブゴールを探索する. 次に, 初期移動で発見したサブゴール間の移動を試みる. しかし, 初期移動によって得られたサブゴールの座標は, デッドレコニングエラーを含むため, ロボットは目的のサブゴールに正しく到達することができない. 我々の学習手続きでは, この累積誤差の影響をディスティンクティブプレース2)でリセットする.

§2ではディスティンクティブプレースについて, §3では実験で使用した移動ロボットについて説明する. また, 学習手続きに関しては§4で説明する. §5では実験の設定と結果に関して述べる. §6では我々の学習手続きの制限に関して述べる. 最後に, まとめと今後の課題を§7で述べる.

### 2 ディスティンクティブプレース

移動ロボットの行動制御には, 位置ベースとセンサベースの2つがある.

位置ベースの行動制御では, ロボットの移動方向は, 現在の位置とゴールの位置によって決められる. ゴール方向に障害物のセンサ反応がなければ, ロボットの行動は位置ベースで制御される.

センサベースの行動制御では, ロボットの移動方向はセンサデータによって障害物に接触しないように決められる. ゴール方向に障害物のセンサ反応があり, かつ障害物との距離がある値を下回ったら, ロボットはセンサベースで制御される. センサベースの行動制御は, 障害物を壁沿い移動6)で回避する際に利用される. 一般に, センサベースの行動制御はロバーストで

あることが知られている。我々の使用するロボットは、状況に応じてこれら2つの行動制御を使い分ける。

本研究では、移動ロボットの行動制御がセンサベースから位置ベースに切り替わる地点をディスティンクティブレース2)(以下 DP)とする。つまり、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直進を再開する。DP は、通常、障害物の凸コーナーで、局所的な情報のみで同定可能である。障害物のコーナーは障害物方向のセンサエコーが突然なくなることでロボバストに検出できる。Fig.1では、A と B はサブゴール

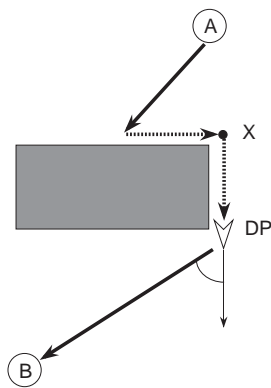


Fig.1 DP

ル、を表す。コーナー X では、サブゴール B の方向が障害物によって遮られているので、ここから B 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、ロボットはそこを DP として登録し、付随して以下のデータを記録する。

- (a) DP の座標.
- (b) *dw-angle*  $\theta$  (Fig.1) : 目的のサブゴールの方向といままで沿ってきた壁の方向のなす角度.
- (c) DP に到達するまでに通過した障害物のコーナーの数.

(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、その平均をとることで求める。(a)と(c)は、後で DP を同定するために利用する。DP に到達したら、累積誤差を含む座標データと方向データを、前回その場所に到達した際のデータにリセットする。

### 3 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 (Fig.2) と呼ばれる。マイクロマウスは、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。ロボットの向きは、初期

状態からの回転運動毎の角度変化を積算することで求める。シミュレーションとは異なり、実際に回転した角度とセンサを通して得た回転角度は等しくならないので、回転運動を繰り返すうちに内部情報としてもつ向きには累積誤差が生じる。ロボットの位置の評価は、デッドレコニングによって行われる。移動距離の評価自体は比較的正確で、ロボットの位置の評価は、向きの評価誤差に大きく依存する。

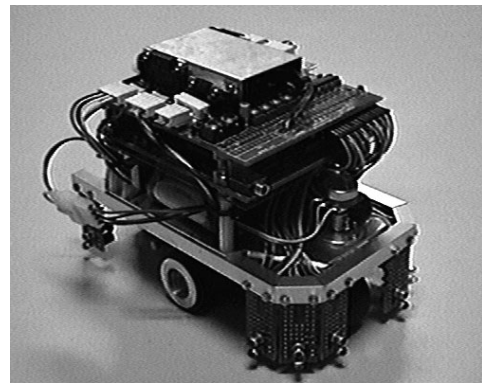


Fig.2 実験に使用する移動ロボット

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。

### 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つの場合を説明する。

#### 4.1 初期移動 - サブゴールの探索

初め、ロボットにはサブゴールの位置に関する情報は与えられていない。そのため、ロボットは、サブゴールの位置情報を獲得するために環境内を探索する必要がある。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情報は不正確なものである。

#### Step1 : 初期移動

- (1-1) 2つのサブゴールに到達できたら、次節の Step2 へ進む。
- (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、その座標を記録して (1-1) へ進む。
- (1-4) ロボットの移動方向が障害物に遮られた場合、壁沿い移動による障害物回避を一定距離行う。回避の際、左右どちら方向に回避するか(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

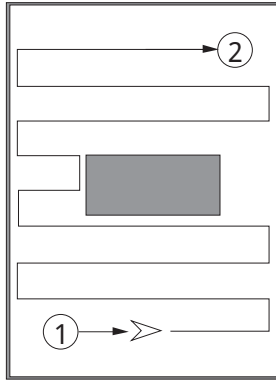


Fig.3 初期移動 - サブゴールの探索

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)へ進む．

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．

#### 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目的のサブゴール方向が，障害物で塞がれていない場合は，その方向に直進する．サブゴール方向が障害物によって遮られている場合は，再び目的のサブゴール方向が開け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，左右どちらの向きに周回を行うか(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．Fig.4では，ローカルディレクションは左である．詳細なプロセスを以下に示す．

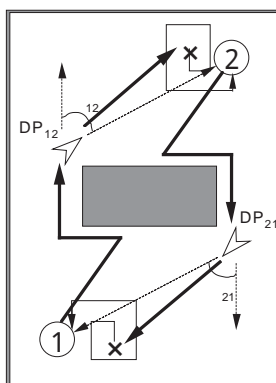


Fig.4 DPの発見と登録

#### Step2 : DPの発見と登録

- (2-1) 本プロセス開始時点にロボットが存在したサブゴール(2)上に，もう一方のサブゴール(1)に到達した後再び到達した場合，次節のStep3へ進む．

- (2-2) 目的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，以下の条件の一つが満たされるまで障害物を壁沿い移動で回避する．
  - (2-3-a) 目的のサブゴール上の到達した場合，(2-5)へ進む．
  - (2-3-b) 目的のサブゴール方向が開けるコーナーにさしかかった場合，そのコーナーをDPとして登録し，(2-2)へ進む．
- (2-4) 現在位置の座標が目的のサブゴールの座標と一致し，かつ，実際にはそのサブゴール上に到達していない(Fig.4において×印で示される場所に存在する)場合，渦巻状の近傍探索によってサブゴールの探索を行い，(2-5)へ進む．
- (2-5) 目的のサブゴールに到達した場合，道中で登録されたDPにおける角度情報(Fig.4の $\theta$ )を再計算し更新する．その後，もう一方のサブゴール方向に直進を開始し，(2-1)へ進む．

#### 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と，サブゴールへの到達に近傍探索の必要がなくなるまでプロセスが継続することである(Fig.5)．

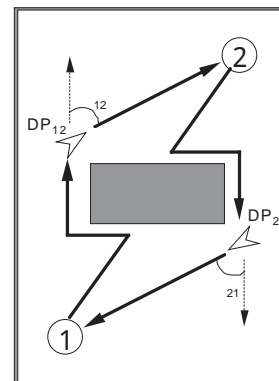


Fig.5 パスの洗練

#### Step3 : パスの洗練

- (3-1) 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．
- (3-2) 目的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，以下の条件の一つが満たされるまで障害物を壁沿い移動で回避する．
  - (3-3-a) 目的のサブゴール上の到達した場合，(3-5)へ進む．

- (3-3-b) DP として登録されているコーナーにさしかかった場合、座標および向きに関するロボットの内部情報を、登録されているものにリセットする。さらに、登録されている角度情報 (dw-angle) に基づいて進行方向を変え、(3-2) へ進む。
- (3-4) 現在位置の座標が目的のサブゴールの座標と一致し、かつ、実際にはそのサブゴール上に到達していない場合、渦巻状の近傍探索によってサブゴールの探索を行い、(3-5) へ進む。
- (3-5) 目的のサブゴールに到達した場合、道中で登録された DP における角度情報 (Fig.5の $\theta$ ) を再計算し更新する。その後、もう一方のサブゴール方向に直進を開始し、(3-1) へ進む。

## 5 実験

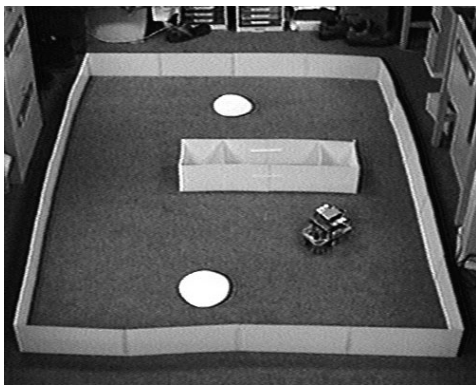


Fig.6 Experimental environment

我々は、サブゴール2つ、障害物1つの環境でいくつかの実験を行った (Fig.6)。環境は、180×144×13cmの外壁で囲まれており、障害物は、72×18×13cmの直方体である。サブゴールは、直径18cmの円形のプレートを使用した。今回の実験では、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。

### 5.1 静的環境

Fig.7は、環境とロボットの移動軌跡の俯瞰図である。見易くするために、初期移動と近傍探索の軌跡は省略した。初期移動を終了した時点では、ロボットは図中上側のサブゴール状に存在する。ロボットは、既に登録済みの DP への到達をロボパストに行い、dw-angle を徐々に修正し、サブゴール間の移動を近傍探索なしに行うことに成功した。

### 5.2 動的環境

次に、環境変化への適応性を調べるために、静的環境での軌跡が収束した後に、サブゴールと障害物の位置をずらして実験を続けた (Fig.8, 9)。環境変化によって、前の実験の際に登録した DP の情報は、信頼できないものになる。それゆえ、1周目、ロボットは目的のサブゴールに近傍探索なしでは到達できなかった。2周目は、1周目で更新された DP の情報を利用する

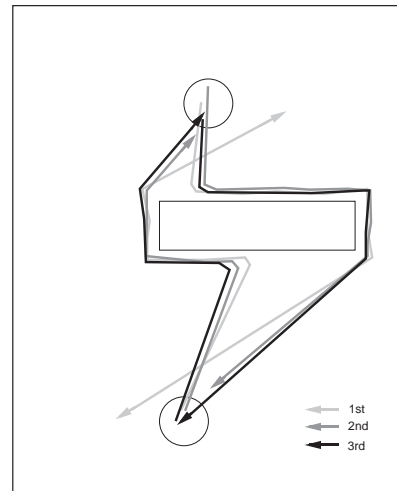


Fig.7 ロボットの軌跡

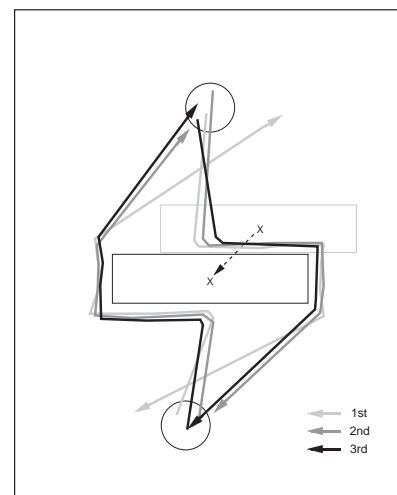


Fig.8 障害物の位置変化への適応

ことにより、近傍探索に入る直前のロボットの最終的な位置は、1周目よりも目的のサブゴールに近づいた。ロボットは、3周目で近傍探索なしに目的のサブゴールに到達することに成功した。このように、ロボットは、環境変化に適応し、近傍探索なしにサブゴール間の移動を行えるようになった。

さらに、障害物を取り除いた環境で10周分実験を継続した (Fig.10)。図中の数字は、各周回で近傍探索に入る直前のロボットの位置 (近傍探索を必要としなかった周回については、最終的なロボットの位置) を示している。障害物のない環境では、DPでの累積誤差のリセットが行えず、軌跡は収束しなかった。

また、障害物の出現や消失といった環境変化への適応性を調べるため、複数の障害物が存在する別の環境で実験を行った。

Fig.11は、障害物の出現や消失がおこる前の環境におけるロボットの移動軌跡を示している。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

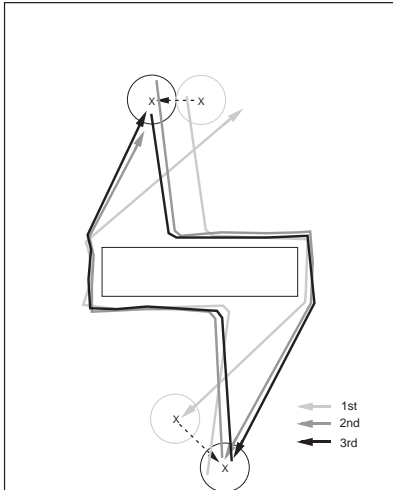


Fig.9 サブゴールの位置変化への適応

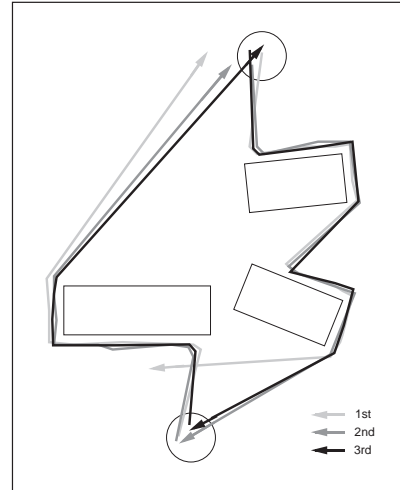


Fig.11 複数の障害物のある環境での挙動

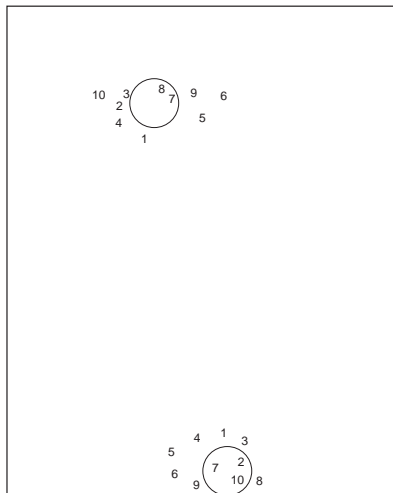


Fig.10 障害物のない環境での挙動

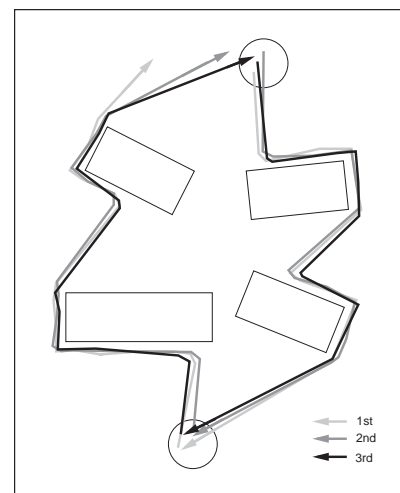


Fig.12 障害物の出現への適応

たところで障害物を1つ追加した (Fig.12)．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 での dw-angle 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.13)．障害物が除去されたことにより，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つの DP の内1つがなくなり，残った DP での dw-angle が，3 周目に近傍探索を必要としなくなる値に収束した．ロボットは，障害物の出現や消失といった環境変化への適応に成功した．

## 6 考察

初期移動の手続きには問題点がある．これは，初期移動の手続きがあくまでヒューリスティックなアプローチであるため，サブゴール探索が常に成功する保証のない点である．また，現時点では，初期移動で獲得されるのはサブゴールの位置に関するデータのみで，障害物に関するデータは獲得していない．それにも関わ

らず環境内を実際に探索してデータを獲得する必要性は，ロボットの内部表現を考慮した環境情報のオペレーターによる入力手間を省くことにある．

本研究では，基本的に同じサブゴール間を移動する場合，毎回同じ障害物に到達することを仮定しているが，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．この場合，ロボットは新しい DP を登録し，学習によって再び dw-angle を収束させる必要がある．同様に，次の周回でこの新しく登録された DP に到達できるという保証もない．最悪の場合，いつまでもたってもパスが収束しないことも考えられる．このような弱点を補うために，我々はサブゴールや DP 以外の環境データを利用する方法を考案中である．具体的には，壁沿い移動中に獲得した障害物の部分的なデータを反映したポテンシャルフィールドを形成し，現段階では固定されている障害物回避方向 (ローカルディレクション) を，目的のサブゴールまでのパスがより

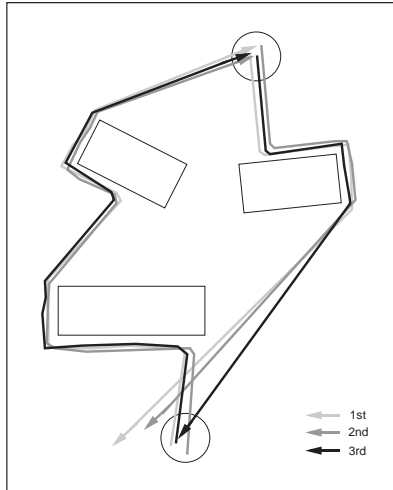


Fig.13 障害物の消失への適応

短くなる方向に決定するために利用するというものである。

DPは障害物のコーナーに現れる。障害物のコーナーは、障害物方向のセンサエコーが突然なくなることで検出されるため、DPをロバストに検出するためには、障害物の形状が多角形に制限される。

## 7 まとめ

本論文では、障害物を含んだ未知の動的環境内で、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た。我々は、位置や向きの評価に誤差のある実環境で、いくつかの実験を行った。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が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。

将来、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) Latombe, J.C., Robot Motion Planning, *Kluwer Academic Publishers*, 1991
- 2) Kuipers, B.J., Byun, Y.T., A Qualitative Approach to Robot Exploration and Map-learning, *In Proc. of Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp.390-404, 1987
- 3) Malkin, P.K., Addanki, S., LOGnets: A Hybrid Graph Spatial Representation for Robot Navigation, *In Proc. of the eighth AAAI-90*, pp.1045-1050, 1990
- 4) Skewis, T. and Lumelsky, V., Experiments with a Mobile Robot Operating in a Cluttered Unknown Environment, *In Proc. of the 1992 IEEE Int. Conf. on Robotics and Automation*, pp.1482-1487, May 1992

- 5) Noborio, H., Yoshioka, T., and Hamaguchi, T., On-Line Deadlock-Free Path-Planning Algorithms in the Presence of a Dead Reckoning Error, *In Proc. of the 1995 IEEE Int. Conf. on System, Man, and Cybernetics*, pp.483-488, 1993
- 6) Ando, Y. and Yuta, S., Following a Wall by an Autonomous Mobile Robot with a Sonar-Ring, *In Proc. of the 1995 IEEE Int. Conf. on Robotics and Automation*, pp.2599-2606, 1995