

総合論文

群ロボット環境における局所的通信に基づく衝突回避

新井 義和^{*1} 藤井 輝夫^{*2} 浅間 一^{*3}
 鈴木 昭二^{*4} 嘉悦 早人^{*5} 遠藤 勲^{*5}

Collision Avoidance in Multi-Robot Environment based on Local Communication

Yoshikazu Arai^{*1}, Teruo Fujii^{*2}, Hajime Asama^{*3},
 Sho'ji Suzuki^{*4}, Hayato Kaetsu^{*5} and Isao Endo^{*5}

In this paper, we propose a new collision avoidance method among multiple autonomous mobile robots. Generally, it is difficult to apply conventional collision avoidance methods to robots in multi-robot environment because those methods are only applicable to a single robot environment. For this problem, we propose a new method using the LOCISS (LOCally Communicable Infrared Sensory System) by which a robot is able to communicate with other robots locally to exchange information necessary for their mutual collision avoidance. To realize the collision avoidance based on these information, we introduce a learning method by which a robot is able to acquire the behavior adaptively and autonomously. A learning curriculum is divided into multiple layers to reduce a number of situation for the learning. Finally, we implemented this system to execute acquired behaviors, so that robots can avoid other robots and obstacles at the same time. It is confirmed by conducting experiments that this method is effective in the multi-robot environment.

Key Words: Collision Avoidance, Local Communication, Reinforcement Learning, Multi-Layered Learning, Infrared Sensor, Multiple Robots

1. はじめに

ロボットへの作業要求の多様化に伴い、単純な機能を備えた複数のロボットの協調に基づく群ロボットシステムの研究が盛んに行われている[1]~[3]。このようなシステムでは、多くのロボットが環境内に存在し、それらが互いに移動の妨げとなることから、衝突回避は群ロボットシステムを実現する上で最も重要な基本動作である。

一般に、衝突回避を行うためには、ロボットは周囲に存在する物体を認識し、その物体の種類(移動物体か否かなど)に合わせて回避行動を決定し、これを実行する必要がある。物体を

検出あるいは認識するための手法としては、これまで大きく分類してセンシングに基づく手法と通信を用いる手法の二つのアプローチが試みられてきた。

センシングに基づく手法においては、超音波センサなどのアクティブセンサもしくは視覚センサなどのパッシブセンサが用いられてきたが、それぞれに利点欠点が考えられる。アクティブセンサはほぼ実時間で物体との距離情報を取得することが可能であり、衝突回避のための情報獲得手法としては非常に有効である[4]~[6]。ところが、これらのセンサを複数のロボットが同時に用いる場合、複数のセンサが発する信号同士が干渉し、認識を正確に行うことができない。また、超音波センサなどで測定した距離情報のみを用いて物体との衝突を回避する場合、ロボットなどの移動している物体の行動を予測することは困難であり、必ずしも柔軟に対応できるとはいえない。パッシブセンサのうち最も広く用いられている視覚センサを用いる手法[7][8]では、複数の移動ロボットを認識するための画像処理に大きな計算コストを要するため、小型ロボットに搭載しうるハードウェアを想定すると、実時間で情報を得ることは難しい。

ロボット間の通信を用いることによりお互いの位置情報や移動情報を交換し、それらの情報に基づいて回避手法を交渉する手法[9]~[11]では、大局的な通信を用いるため、群ロボット環境においてはロボットの台数の増加に伴い、それらの間の通信

原稿受付 1999年2月8日

^{*1}岩手県立大学ソフトウェア情報学部

^{*2}東京大学生産技術研究所

^{*3}理化学研究所工学基礎研究部

^{*4}公立ほこだて未来大学システム情報学部

^{*5}理化学研究所化学システム研究室

^{*1}Faculty of Software and Information Science, Iwate Prefectural University

^{*2}Institute of Industrial Science, University of Tokyo

^{*3}Instrumentation Project Promotion Division, The Institute of Physical and Chemical Research (RIKEN)

^{*4}Dept. of Media Architecture, Future University-Hakodate

^{*5}Biochemical Systems Laboratory, The Institute of Physical and Chemical Research (RIKEN)

量が増大し、通信の停滞なしに実時間で周囲の情報を獲得するのに十分な通信容量を確保することが困難となる。

以上の議論をふまえて、筆者らは、群ロボット環境における情報獲得手法として赤外線を用いた局所的通信を提案し、開発を進めてきた [12]。赤外線は、発光強度を調節することにより到達範囲を一定の距離に限定し、容易に通信の局所性を実現できる。それによって、同一環境内に複数のロボットが存在する場合でも、特定のロボット同士が他のロボット間の通信の影響を受けることなく情報をやりとりすることが可能となる。また、その通信によって得られる情報を用いれば、周囲の状態、すなわちロボットおよび障害物の存在ならびにロボットの行動を認識することが容易に行えるとともに、それらを衝突回避に利用することができるようになる。

センサなどから得た情報に基づいて認識した周囲の状態に対して適切な行動をロボット自身が選択するためには学習手法の導入が不可欠である。特に、強化学習 [13] を用いた場合、ロボットは自分自身で行動の評価を行うことができるため、外部から教師信号を与えることが困難であるような複雑な環境に対応することができる [14] [15]。しかし、これらの研究は、学習の際に入力する情報を制限することにより、比較的低い次元の状態空間を想定しており、本研究で扱うような高い次元の状態空間をもつ問題に適用することは困難である。

本論文では、まず局所的通信の概念に基づいて開発した赤外線通信センサシステム LOCISS (LOCally Communicable Infrared Sensory System) について述べ、ロボットの周囲の状態認識の原理とともにそのハードウェア構成を示す。また、LOCISS を用いて認識された周囲の状態に対し、あらかじめ定義したルールに基づいて回避行動を行う実験を通して LOCISS の有効性を確認する。さらに、実際の群ロボットシステムで想定されるような多数台のロボット間での相互の衝突回避という複雑な環境に対応するために、強化学習を用いて実行すべき回避行動を適応的に獲得する方法を提案する。さらに、この手法を実際のロボットに搭載するために、階層化学習の概念を提案する。学習を階層化することにより、学習の際に入力する情報を制限することなく状態空間を削減することができ、周囲のロボットの行動などの衝突回避を行うために非常に有効な数多くの情報を学習に反映することが可能となる。最後にこの手法を実ロボットに搭載し、実際に多数のロボットを動作させることにより、提案した一連の手法の有効性を示す。

2. 赤外線通信センサシステム LOCISS の開発

本章では、赤外線を用いた局所的通信システムとして筆者らが開発した LOCISS について、その基本概念、ハードウェア構成および基本性能について述べる。赤外線は周囲の物体に対して透過性がなく、ロボットが接触する恐れのある見通し範囲内にもみ到達可能であるという点で、群ロボット環境における衝突回避を目的とした局所的通信に適した媒体である。

2.1 LOCISS による状態認識の原理

LOCISS による局所的通信に基づく状態認識の概念を Fig. 1 に示す。おのおののロボットの ID および移動方向・速度などの情報を局所的通信によってロボット間で交換すれば、通信の局

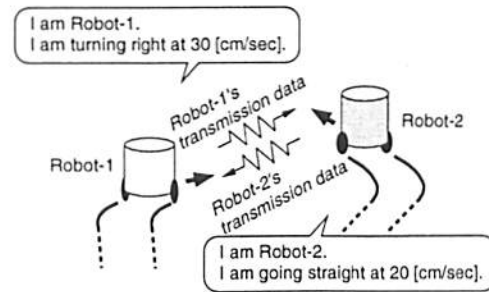


Fig. 1 Recognition of surrounding situation

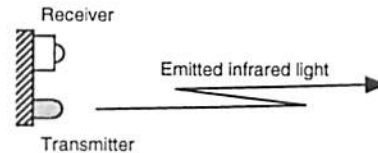


Fig. 2 Infrared communication unit

所性により付近に他のロボットあるいは障害物が存在することならびにロボットに対してはその行動を容易に確認できる。すなわち、ロボット 1 の通信可能範囲内にロボット 2 が接近すると、ロボット 1 はロボット 2 の信号を受信する。このとき、ロボット 1 は受信した ID 情報が自分自身の ID と異なることから接近した物体が自分以外のロボットであることを認識するとともに、それがロボット 2 であることを識別する。また、ロボットが移動して障害物に接近した場合、ロボットは障害物に反射した自分自身の ID を受信することから、障害物であると識別できる。検出した物体がロボットであった場合、同時に移動方向および速度を受信することにより相手のロボットの行動を知ることができる。

2.2 センサシステムの構成

赤外線による通信は太陽光や照明などの外光が雑音となる。これに対処するため、伝送媒体として赤外線を 38 [kHz] の周波数で変調する。送信部には赤外発光ダイオード (SHARP GL537)、受信部にはリモコン用受光素子 (SHARP IS1U60) を用いた。受光素子には上記周波数のバンドパスフィルタが内蔵されており、受信した信号を容易に復調することができる。送受信のコントローラとしては、Z80 CPU ボード (秋月電子 AKI-80 12M) を用いた。赤外線の到達距離は、筆者らのロボット [18] の移動速度 0.3 [m/s] およびサイズ $W \times D = 0.43 [m] \times 0.43 [m]$ を考慮し 1.2 [m] とした。なお、この距離は送信回路中の抵抗の値により調整できる。

上記の発光および受光素子を用いて Fig. 2 に示すように送受信ユニットを構成する。このとき、それぞれの素子の特性からユニット全体としての指向性は約 45 [deg] である。本システムでは 360 [deg] のすべての方向に対して局所的通信を実現するために、八つのユニットを Fig. 3 に示すように放射状かつ等間隔に配置することとした。したがって、ロボットは周囲の 8 方向に対して前述の情報を送受信することができる。

八つの各ユニットには図中に示すようにロボットの移動方向を基準としたセンサ番号を反時計方向に割り当てる。この番号は、ロボットが移動方向を変更するたびにそれに応じて更新さ

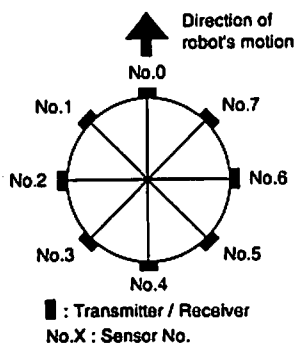


Fig. 3 Configuration of sensor units and sensor number

れる。このセンサ番号をセンサコードとして伝送することにより、ロボットは信号を受信した自分自身の“センサ番号”および相手のロボットが送信した“センサコード”の組み合わせを用いてお互いの移動方向の相対的関係を認識できる。すなわち、これは相対的移動方向の情報をセンサ番号によって伝えることに対応している。移動速度は0.3, 0.2, 0.1, 0.0 [m/s]の4段階に変化できることとし、それぞれの速度を3, 2, 1, 0とコード化して2ビットで表現する。ロボットIDは、ここでは最大8台とし、3ビットで表現する。センサコード8方向分の3ビットを含め、伝送データを合計8ビットすなわち1バイトのデータ長とする。現状ではハードウェアの制約から伝送データの最大長は1バイトであるが、これを改良することにより移動速度およびロボットIDを必要に応じてより自由にコード化することが可能である。

2.3 送信データの符号化および干渉検出

送信データは Fig. 4 に示すようにパルス符号化する。単位パルス幅を a とすると、 $6a$ 幅のヘッダおよびそれに続くデータを一つのフレームとして送受信する。図のように各ビットは P_a : 同期パルス, P_b : データパルスおよび P_c : 干渉検出パルスの三つのパルスによって表し、これらを連続的に並べてデータとする。これにより、データには最大でも $2a$ 幅の ON 状態しか現れないため、 $5a$ 幅の ON 状態を観測することによりヘッダが検出でき、フレーム同期をとることができる。干渉パルスは基本的には常に OFF 状態であるが、複数のロボットからの信号が同時に受信される場合などはヘッダ部分やデータ部分とオーバーラップすることにより、ON 状態となる可能性がある。したがって、干渉パルスの状態を監視することにより複数の入力の干渉の検出が可能である。

前述の Z80CPU ボードを用いる場合、その性能から単位パルス幅 $a = 630 [\mu s]$ となり、例えばヘッダ付き1バイトデータの通信が $18 [ms]$ で行うことができる。また、パルスの立ち上がりの同期がとれないため、受信時の ON/OFF 状態の判別は $a/2 = 315 [\mu s]$ の周期で行う。以上のコーディングにより、次の4通りの通信状態の検出が行える。

- (1) No signal : まったく信号を受信していない状態
- (2) Correct frame : フレームを正しく受信している状態
- (3) Frame error : 信号を受信しているがデータが復号不能な状態
- (4) Interference : 干渉が発生している状態

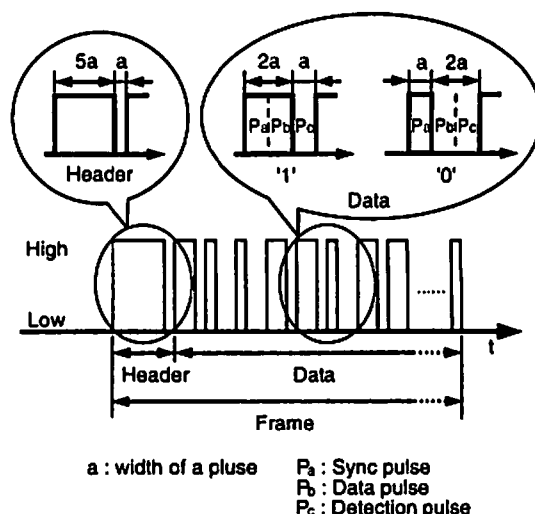


Fig. 4 Pulse coding of transmission data

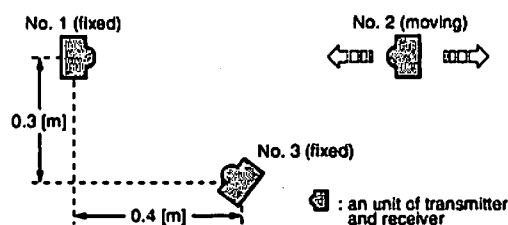


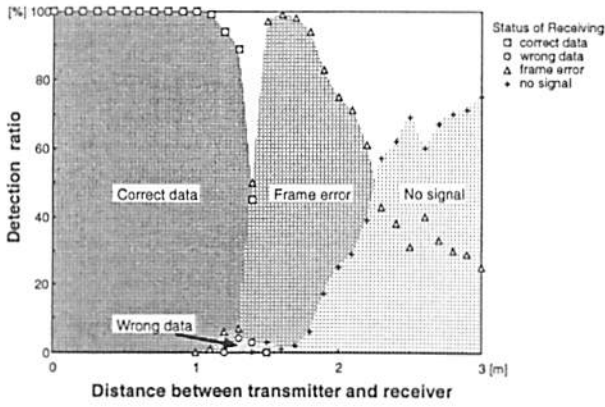
Fig. 5 Experimental environment of performance evaluation

2.4 基本性能の評価

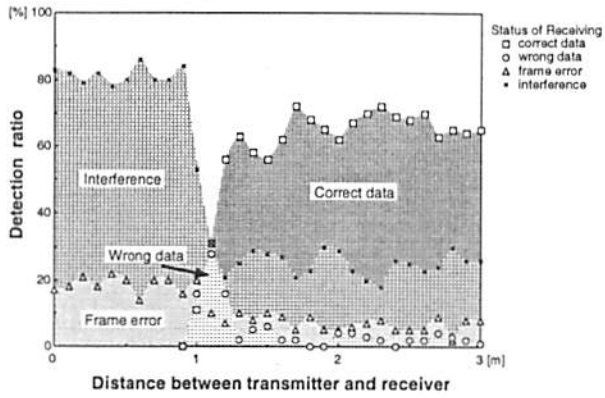
開発した LOCISS の基本性能を評価するために、1対1の通信状態の検出および干渉の検出の二つの実験を行った。Fig. 5 に実験環境を示す。

1対1の通信状態の検出では、図中の No.1 および No.2 の送受信ユニットを向かい合わせて配置し、No.1 を受信用、No.2 を送信用とした。このとき、No.1 を固定し、No.2 を $0.1 [m]$ ずつ動かすことによりユニット間の距離を変化させ、それぞれの距離ごとに受信側のコントローラ上で $0.5 [s]$ 間隔で 100 回ずつ受信データの読み取りおよび通信状態の判別を行った。実験結果を Fig. 6 (a) に示す。ただし、ここでは前述の4通りの通信状態の中の (2) Correct frame を、送信データと同じデータを受信できた場合 (Correct data) とそうでない場合 (Wrong data) に分けている。横軸がユニット間の距離、縦軸が各状態の割合を示している。この結果から、送信側のユニットとの距離に応じて判別される通信状態の変わり目では、その割合が急激に変化しており、正しい受信ができていない範囲が設計上の赤外線の到達距離とほぼ一致していることが分かる。

干渉の検出では、通信状態の判別の実験環境に No.3 の送信用ユニットを加える。No.1 および No.3 のユニットは固定され、それらの間では常に通信が成立している。このとき、No.2 の送信用ユニットを $0.1 [m]$ ずつ動かし、No.1 の受信用ユニットにより受信データの観測を行った。観測の条件は通信状態の検出実験と同様である。実験結果を Fig. 6 (b) に示す。No.1-No.2 間の距離が $0 \sim 1.0 [m]$ の間では実際には常に干渉が発生してお



(a) Detection of reception status by one to one



(b) Detection of interference

Fig. 6 Result of data exchange experiments

り、そのうち約 80%が干渉 (Interference) として正しく検出されている。残りの約 20%はフレームエラーとして検出された。1.0[m] 以上離れると、干渉の影響が約 20%ほど残ってはいるものの、No.3 が送信したデータが約 60%の割合で正しく受信できていることが確認できる。

本システムでは前述のように八つの送受信ユニットをロボットの外周上に配置していることから、ロボットのほぼ全方位に対して上記の結果のような認識が可能である。また、障害物の検出においては赤外線の利用しているため信号が減衰し、ロボットの検出に比べて検出可能範囲が約半分となる [12]。

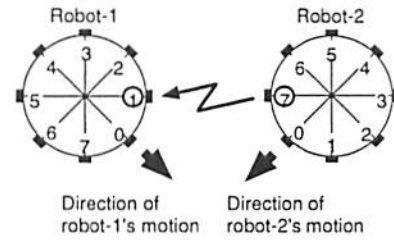
以上から、赤外線通信センサシステム LOCISS の基本性能として、通信の局所性ならびに干渉の検出を含む各通信状態の判別が実現可能であることが確かめられた。

3. 相手の行動を考慮した衝突回避

前章で述べた LOCISS を用いることにより、周囲のロボットと障害物を見分け、さらにロボットの行動を考慮した衝突回避を実現できる。ここではまず、干渉が発生する恐れがなく、かつロボット間の相互関係を容易に把握することができる 2 者間の衝突回避について、あらかじめ衝突回避のルールを定義することにより回避行動を決定する問題を考える。実験には、ルールの定義を簡略化するために、取りうる回避行動の選択肢が比

		Sensor No. : S_n							
		0	1	2	3	4	5	6	7
Sensor Code : S_c	0	T(L)	I	I	I	I	I	I	I
	1	S	T(R)	I	I	I	I	I	I
	2	S	I	I	I	I	I	I	S
	3	F	I	I	I	I	I	I	F
	4	F	I	I	I	I	I	I	I
	5	F	F	I	I	I	I	I	I
	6	S	S	I	I	I	I	I	I
	7	S	S	I	I	I	I	I	T(L)

(a) Rule matrix



(b) Example of detection

Fig. 7 Collision avoidance rule for robot

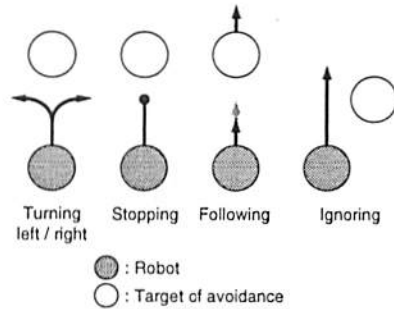


Fig. 8 Definition of avoiding behaviors

較的限られているステアリング型の移動ロボットを用いる。

3.1 衝突回避ルールの定義

ここでは、回避対象がロボットであるのか障害物であるのかによって選択すべき回避行動が異なるため、ルールは対ロボットおよび対障害物の 2 種類に分割する。

Fig. 7 (a) 対ロボットのルールマトリクスを示す。列は注目しているロボットの八つのセンサユニットのうち、信号を受信したセンサユニットのセンサ番号 S_n を、行は受信した相手のロボットが送信したセンサコード S_c を示す。ロボットは、これらの二つの情報に基づいて、相手のロボットとの位置関係およびその行動を認識し、マトリクスの要素 (S_c, S_n) に示す行動により相手のロボットを回避する。ここでは、ロボットは目的地に向かって通常時は最大速度で移動することを前提とし、ロボットがとるべき行動として“右への軌道変更 (T (R))”、“左への軌道変更 (T (L))”、“追従 (F)” および “停止 (S)” に何も回避行動をとらない “無視 (I)” を加えた 5 種類を用意する (Fig. 8)。ここで “追従 (F)” とは、2 台のロボットが同じ方向に移動していて、後ろのロボットの方が最大速度が速い場合

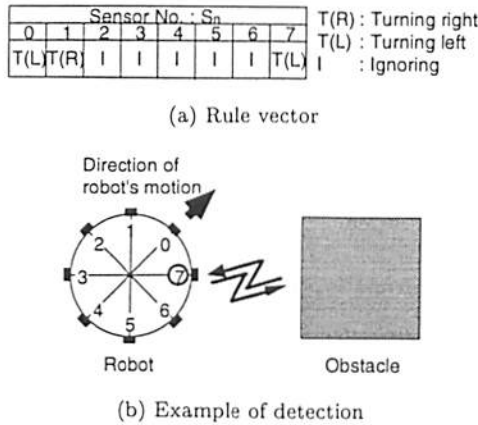


Fig. 9 Collision avoidance rule for obstacle

に、後ろのロボットが減速することを意味する。

例えば、Fig. 7 (b) に示すような状況においてロボット 1 とロボット 2 が接近した場合、ロボット 1 はセンサ番号 1 のセンサユニットでロボット 2 のセンサコード 7 を受信する。したがって、マトリクスの $(S_e, S_n) = (7, 1)$ の要素を参照し、回避行動として“停止 (S)”を選択する。同時に、ロボット 2 は $(S_e, S_n) = (1, 7)$ の要素である“無視 (I)”を選択する。

Fig. 9 (a) に対障害物のルールベクトルを示す。列方向は 8 組のセンサのうち、障害物に反射した自分自身の信号を受信したセンサユニットのセンサ番号 S_n を示す。ロボットは、この情報に基づいて障害物との位置関係を認識し、ベクトルの要素 (S_n) に示す行動により障害物を回避する。例えば、Fig. 9 (b) に示すような状況で、ロボットが障害物に接近した場合、ロボットはセンサ番号 7 のセンサユニットから自分自身が発した信号を受信する。したがって、検出した物体が障害物であることを認識すると同時に、ベクトルの $(S_n) = (7)$ の要素を参照し、回避行動として“左への軌道変更 (T (L))”を実行して障害物との衝突を回避する。

3.2 衝突回避実験

LOCISS による状態認識およびその認識に基づく衝突回避の有効性を示すために実機に対して前述のルールを適用し、ロボットおよび障害物に対する衝突回避実験を行った。それぞれの実験の結果を Fig. 10, 11 に示す。なお、図中では 5 [sec] 間隔で記録したロボットの位置を、軌跡としてプロットした。

ロボットに対する衝突回避実験では、2 台のロボットを直角の向きに配置し、ゴールへ向かう途中で衝突する状況を想定した。2 台のロボットはお互いに接近すると同時にそれぞれを認識し、ルールマトリクスに基づいてロボット 1 は“無視 (I)”を、ロボット 2 は“停止 (S)”を選択した (Fig. 10 (a)). また、ロボット 1 がロボット 2 の前方を通り過ぎて衝突の危険がなくなると、ロボット 1 はそのまま“無視 (I)”を選択するとともにロボット 2 もまた“無視 (I)”を選択し、ゴールに向かって移動を再開した (Fig. 10 (b)). これは全体から見るとロボット 2 がロボット 1 に道を譲り、ロボット 1 が先に通過したこと以外ならぬ、それぞれのロボットが相手の動きを認識したからこそ実現できる衝突回避である。

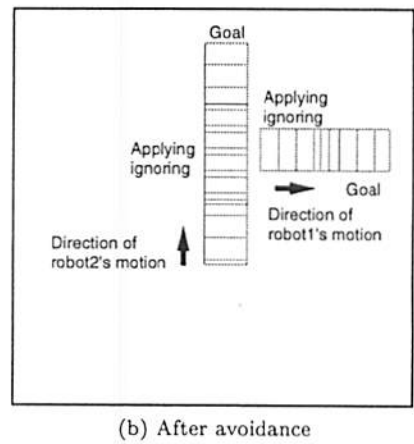
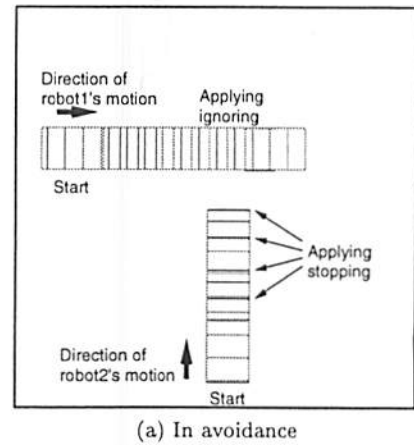


Fig. 10 Result of rule based collision avoidance for robot

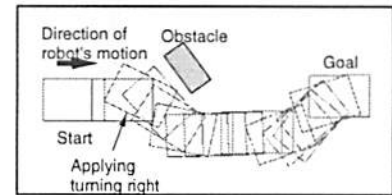


Fig. 11 Result of rule based collision avoidance for obstacle

障害物に対する衝突回避実験では、ロボットのスタートとゴールの途中で障害物を配置した。ロボットは、障害物に接近すると同時にそれを検出し、“右への軌道変更 (T (R))”を選択した。また、衝突の危険がなくなると“無視 (I)”を選択してゴールへ向かって移動を再開したことが確認できる (Fig. 11)。

以上により、LOCISS を用いて周囲のロボットや障害物を検出し、ルールに基づいた 2 者間の衝突回避が実現できることが確認された。これは、周囲の状況に対して回避行動を選択するためのルールを得ることさえできれば、ロボットは複数のロボットが存在するような複雑な環境においても柔軟に衝突回避を行うことができることを意味する。ここでは、ルールの定義を単純に行うためにステアリング型の移動ロボットを用いたが、後述する群ロボット環境においては複数の物体を同時に回避するために高い移動能力が必要であることから、全方向移動ロボッ

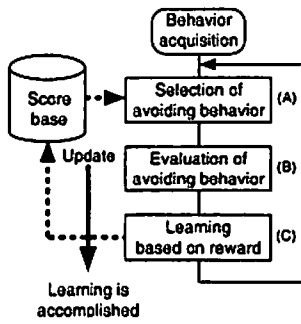


Fig. 12 Procedure of adaptive behavior acquisition

トを用いることとする。上記の LOCISS の性能はロボットの移動能力には依存しないので、全方向移動ロボットを用いた実験においても、同様の性能を得られることが期待できる。

4. 適応的な回避行動獲得

複数の物体が存在する環境において前述のような周囲の状態の認識を行った場合、認識される複雑かつ膨大な数の状態に対して適応的な回避行動を特定するのは困難であり、人間が衝突回避のルールを記述することは事実上不可能である。そこで、このような問題に対して、学習手法をロボットに搭載することにより、様々な状態に対する経験に基づいてロボット自身が自律的に回避行動を習得することを考える。

学習手法を実際のロボットに搭載するには、ロボットは取得した情報に基づいて、行動を選択・評価し、学習を行うという一連の作業を移動と同時に進めなければならない。そのためには、外部からの教師信号が不要であり、なおかつ学習に必要とされる計算量が少ないことが望ましい。ここでは、これらの条件を満たす学習手法として強化学習を導入し、LOCISS によって認識した周囲の状態に対する学習に基づく適応的な回避行動の獲得手法を提案する。また、シミュレーション環境において学習の収束性を示すことにより、提案手法の有効性を示す。

4.1 強化学習に基づく回避行動獲得

Fig. 12 に強化学習に基づく回避行動獲得の手順を示す。ここではロボットが、後述のスコアベースに基づいて回避行動を選択し、その行動を自分自身で評価するとともに、そこから得られた報酬に基づいてスコアベースを更新することにより学習を行う。以下にそれぞれの手順について詳しく述べる。

4.1.1 学習のための状態認識

ロボットの周囲の状態を認識し、それに対して適切な回避行動を学習するために入力される情報を Fig. 13 に示す。同図はロボット 1 がロボット 2 を検出した状況を示している。ロボット 1 は、ロボット ID (Robot ID) によって検出した物体がロボット 2 であることを識別し、ロボット 2 の移動方向 (Moving Direction) および速度 (Moving Speed) によりその行動を認識する。ただし、物体が障害物である場合はこれらの情報は用いず、その存在のみを認識する。また、これらの認識に加えて、物体を検出した方向 (Direction of Detected Object) および自分自身の移動速度 (Moving Speed) により、自分自身の物体に対する相対的な行動を認識する。ゴールの相対的な方向 (Direction

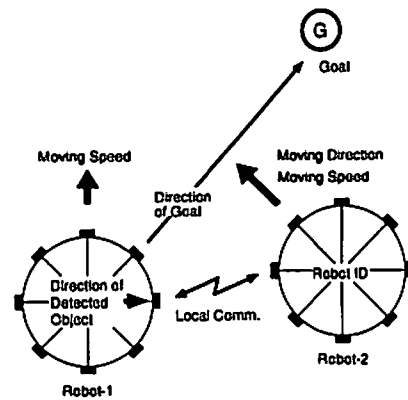


Fig. 13 Parameters for recognition

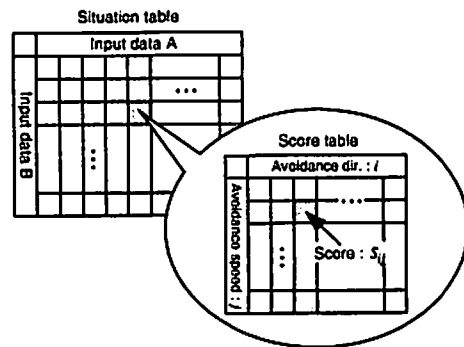


Fig. 14 Concept of score and score base

of Goal) はロボットが本来目指すべき方向を認識するために用いられる。

4.1.2 スコアの概念

学習を行うために、回避行動の選択確率を決定するための指標としてスコアを導入する。スコアの概念を Fig. 14 に示す。状態表は前述の入力情報の組み合わせによって特定の状態を規定する。図では状態表が二次元のマトリクスとして表現されているが、実際には図中の入力 A, B の部分に具体的に移動方向 (Moving Direction) や速度 (Moving Speed) が割り当てられ、入力情報の数と同じだけの次元を持つ。状態表の各要素には各状態に対するスコア表が割り当てられる。スコア表では、ロボットの周囲の特定の状態に対して、ロボットが選択しうる回避行動を回避方向および回避速度の 8 × 4 通りの組み合わせによって表現することとし、回避方向および回避速度のすべての組み合わせに対して一つずつスコアを割り当てる。これらの膨大な数のスコアの集まりをスコアベースとする。

4.1.3 回避行動の選択 (Fig. 12 (A))

ある状態において、回避方向 i 、回避速度 j の組み合わせで表現される回避行動に対するスコアを S_{ij} 、その行動の選択確率を r_{ij} とする。このとき、 r_{ij} は、その状態におけるすべての回避行動のスコアに対して S_{ij} を正規化することにより次のように算出することとする。

$$r_{ij} = \frac{S_{ij}}{\sum_m \sum_n S_{mn}} \quad (1)$$

ロボットは、この選択確率 r_{ij} に基づいて回避行動を選択する。

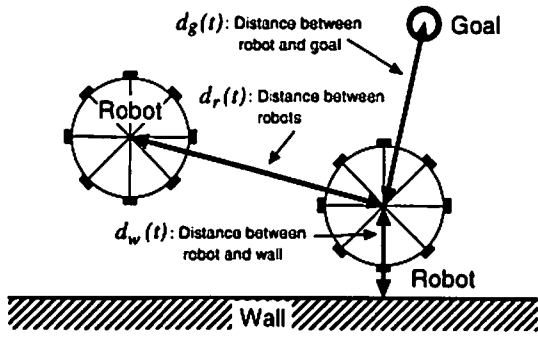


Fig. 15 Evaluation of avoidance behavior

すなわち、大きなスコアを持つ回避行動ほど、選択される確率が高くなる。

4.1.4 回避行動の評価 (Fig. 12 (B))

外部からの教師信号を必要とせずにロボットが適応的な回避行動を獲得するためには選択した行動を自分自身で評価し、その評価を報酬として学習に反映させることが必要である。ここでは、Fig. 15 に示すように、時間 t におけるロボットから障害物までの距離 $d_w(t)$ 、ロボットまでの距離 $d_r(t)$ およびゴールまでの距離 $d_g(t)$ の3種類の距離を用いて次のように評価関数 $E(t)$ を定義する。

$$E(t) = \alpha \Delta d_w(t) + \beta \Delta d_r(t) - \gamma \Delta d_g(t) \quad (2)$$

$$\Delta d_w(t) = d_w(t) - d_w(t - \Delta t)$$

$$\Delta d_r(t) = d_r(t) - d_r(t - \Delta t)$$

$$\Delta d_g(t) = d_g(t) - d_g(t - \Delta t)$$

ただし、 $\Delta d_w(t)$ 、 $\Delta d_r(t)$ 、 $\Delta d_g(t)$ は前述の三つの距離の時間的変化量である。また、 α 、 β および γ はそれぞれの距離に対する重み係数である。実際には、注目しているロボットの周囲には複数の物体が存在するので、障害物までの距離およびロボットまでの距離はそれらの総和で表す。したがって、評価関数 $E(t)$ は、基本的にロボットが周囲の障害物およびロボットからより遠ざかり、なおかつゴールにより近づいた回避行動に対してより大きな値をとる。また、上記のおおの距離の評価の際の影響度は重み係数によって調節することができる。すなわち、 α 、 β を大きくすればそれぞれ周囲の障害物あるいはロボットを優先的に回避する行動が学習され、 γ を大きくすれば物体を回避することよりもゴールに向かうことに優先度をおいた行動が学習される。

4.1.5 報酬に基づく学習 (Fig. 12 (C))

前項の評価関数 $E(t)$ を強化学習の報酬としてそのまま用いることにより、以下の式に従ってスコアの更新を行う。

$$S_{new} = S_{old} + E(t) \quad (3)$$

ここで、 S_{old} は更新前のスコア、 S_{new} は更新後のスコアである。これは $E(t)$ が正であれば正の強化を、負であれば負の強化を意味する。また、回避行動を実行することによりロボットが周囲の物体と衝突した場合は、その回避行動は実行すべきではないという判断から、無条件にスコアを0として再びその行

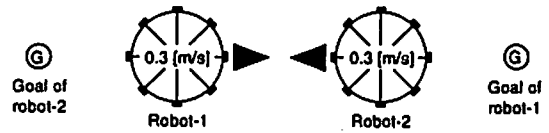


Fig. 16 Environment of simulation experiment

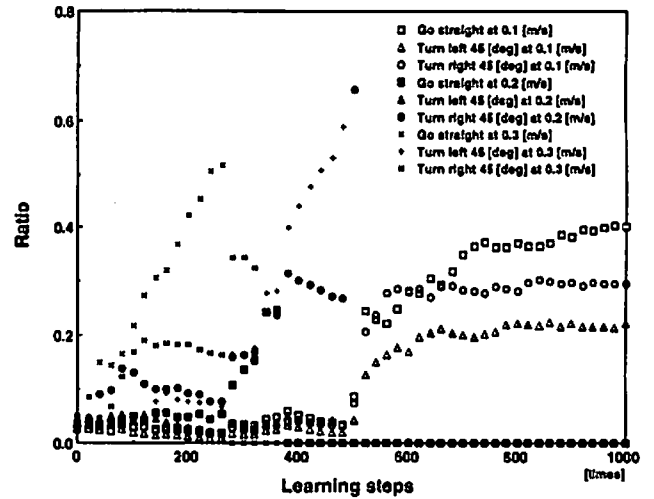


Fig. 17 Transition of selecting probability for avoidance behavior based on learning

動を選択することを禁止する。

以上の手順を繰り返すことにより、認識されたロボットの周囲の状態にふさわしい回避行動のスコアが大きくなる。

4.2 学習の収束性

本手法による学習の収束性をシミュレーション実験によって確認した。実験における前提条件は以下の通りである。

- ロボットはデッドレコニングにより自己位置を認識する
- ロボットはゴールの位置を与えられる
- ロボットは自分自身の姿勢を一定に保ちながら移動する
- 回避速度は0.3, 0.2, 0.1, 0.0 [m/s] の4段階とする
- 回避方向をセンサユニットが搭載されている8方向とする

ただし、デッドレコニングは、実際のロボットでは、車輪に付属したエンコーダの出力に基づいて行われるが、シミュレーション環境では、算術演算によって得られるロボットの位置を示す座標を用いて実現される。

初期状態においてロボットがより早い回避速度を選択でき、回避方向としては8方向を同じ確率で選択できるように、各回避行動のスコアの初期値は、回避速度が早いほど大きく、各回避方向に対しては一定とした。ここでは、評価を算出する際の係数を $\alpha = 50, \beta = 50, \gamma = 100$ とした。

実験の対象とした環境を Fig. 16 に示す。2台のロボットが向かい合った状態で存在し、それぞれのゴールは相手のロボットの向こう側に設定されているとき、2台のロボットが最大速度の0.3 [m/s] でお互いに正面から接近する。ロボットは、この状態を認識するたびにそれに対するスコア表を参照し、一連の手順によりスコアを更新する。この状態における回避行動の選択確率の変遷を Fig. 17 に示す。ロボットがスタート地点からゴールに到達する間に行われる一連の学習を1回と数えるとき、

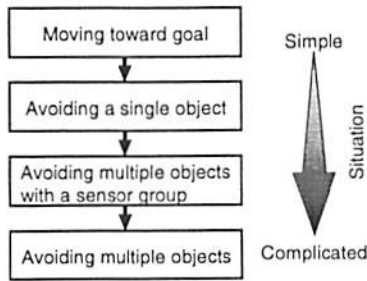


Fig. 18 Concept of multi-layered learning

学習回数約 500 回までは、いくつかの回避行動の選択確率が 0.5 よりも大きな値になったが、それらの行動によってロボットが衝突を起こしたためスコアは 0 となった。結果として、約 800 回の学習によって、回避行動の選択確率が一定の分布へと収束したことが確認できる。すなわち、0.1 [m/s] で直進 (Go straight at 0.1 [m/s]) が 0.4, 0.1 [m/s] で右 45 [deg] 方向 (Turn right 45 [deg] at 0.1 [m/s]) が 0.3, 0.1 [m/s] で左 45 [deg] 方向 (Turn left 45 [deg] at 0.1 [m/s]) が 0.2 の確率を得ている。これらは、この局面においてはまず速度を下げたから右または左に回避する行動が有効であることを示している。

以上の結果から、提案手法による学習が収束し、妥当な回避行動が得られることを確認した。

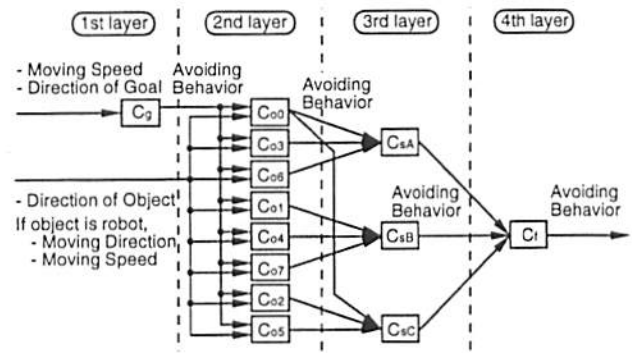
5. 実ロボットへの実装

これまで行われてきた強化学習を用いた実ロボットの衝突回避の研究 [16] [17] は、回避対象物が静止しているなどの前提から、学習の際に入力すべき情報の数は比較的少なくできた。ところが、前述のような数多くの情報によって規定されるロボットの周囲の状態に対して前述のような学習アルゴリズムを適用する場合、それらによって形成される状態空間は非常に高い次元を持つこととなり、それらのすべてを管理するためには膨大なメモリサイズおよび計算コストを要する。一般に実際のロボットに搭載できるメモリのサイズおよび CPU の処理能力は有限であり、それらにより上記のような条件を満たすことは困難である。したがって、実際のロボットに学習による回避行動の獲得手法を搭載するためには、状態空間、すなわち状態数の削減が必要不可欠となる。ここでは、これらの問題に対して階層化学習の概念およびその具体的方法を提案する。

5.1 階層化学習

Fig. 18 に階層化学習の概念を示す。階層化の概念とは、ロボットに最初から難しい課題を与えて学習させるのではなく、簡単な課題から始めて徐々に課題の複雑さを増していく。すなわち、課題および学習の過程を階層的に構造化しようとするものである。これにより、学習もスムーズに進行するとともに、各階層ごとに扱うべき入力次元を小さく抑えることができる。すなわち、図中では、まず最も単純な“ゴールへの移動”という課題から始まって、“一つの物体の回避”、“センサグループごとの複数物体の回避”および“複数物体の回避”の 4 段階の階層化を示している。

Fig. 19 により具体的に示した実際の階層化学習の流れを示



C_g : Controller for movement toward goal
 C_{oi} : Controller for avoidance of single object ($i = 0 - 7$)
 C_{sj} : Controller for avoidance of multiple objects with sensor group ($j = A, B, C$)
 C_i : Controller for avoidance of multiple objects

Fig. 19 Structure of multi-layered learning controller

す。ここでは、学習を前述の四つの階層に分割することとし、各階層における行動の選択機構をコントローラとして表現している。本論文では、学習とは周囲の状態に応じた適切な回避行動を選択するコントローラを構築することに対応する。各コントローラ間の矢印は情報の流れを示しており、入力された複数の情報に基づいて学習が行われ、その結果得られた回避行動が出力される。以下に各階層について詳細を述べる。

5.1.1 第 1 層：ゴールへの移動

第 1 層の学習の目的はゴールへ向かうための行動を獲得することである。Fig. 20 (a) に第 1 層における学習を行う環境を示す。周囲に移動の障害となる物体が一つも存在しない環境において、ロボットは何の干渉も受けずにゴールに向かって移動する。コントローラ C_g の入力はゴールの相対的な方向および自分自身の移動速度である。出力はゴールへ向かうために選択された行動である。

5.1.2 第 2 層：一つの物体の回避

第 2 層の学習の目的はロボットの周囲に存在する一つの物体を回避するための行動を獲得することである。ここでは、八つのセンサユニットに対して八つのコントローラ $C_{o0} \sim C_{o7}$ が存在し、実際には各コントローラはロボット ID による物体の識別に基づいて Fig. 21 に示すように三つの動作を行う。このとき、物体を検出したセンサユニットのセンサ番号を i とすると、コントローラ C_{oi} において学習が行われる。具体的には、検出された物体が障害物であればコントローラ C_{oi} に第 1 層において学習したゴールに向かうための行動が入力される。Fig. 20 (b) に第 2 層において障害物の回避を対象とした学習を行う環境を示す。ロボットの経路上には必ず障害物が存在し、ロボットはその障害物を迂回するとともに、ゴールに向かって移動する。また、物体がロボットならば、コントローラ C_{oi} にその行動に加えて、検出されたロボットの移動方向および速度が入力される。Fig. 20 (c) に第 2 層においてロボットの回避を対象とした学習を行う環境を示す。2 台のロボットがそれぞれの経路の途中でお互いに接近し、進路を譲り合うことにより、ゴールに向かって移動する。障害物あるいはロボットが検出されたセンサユニットに対するコントローラからはそれらを回避する行動が出力さ

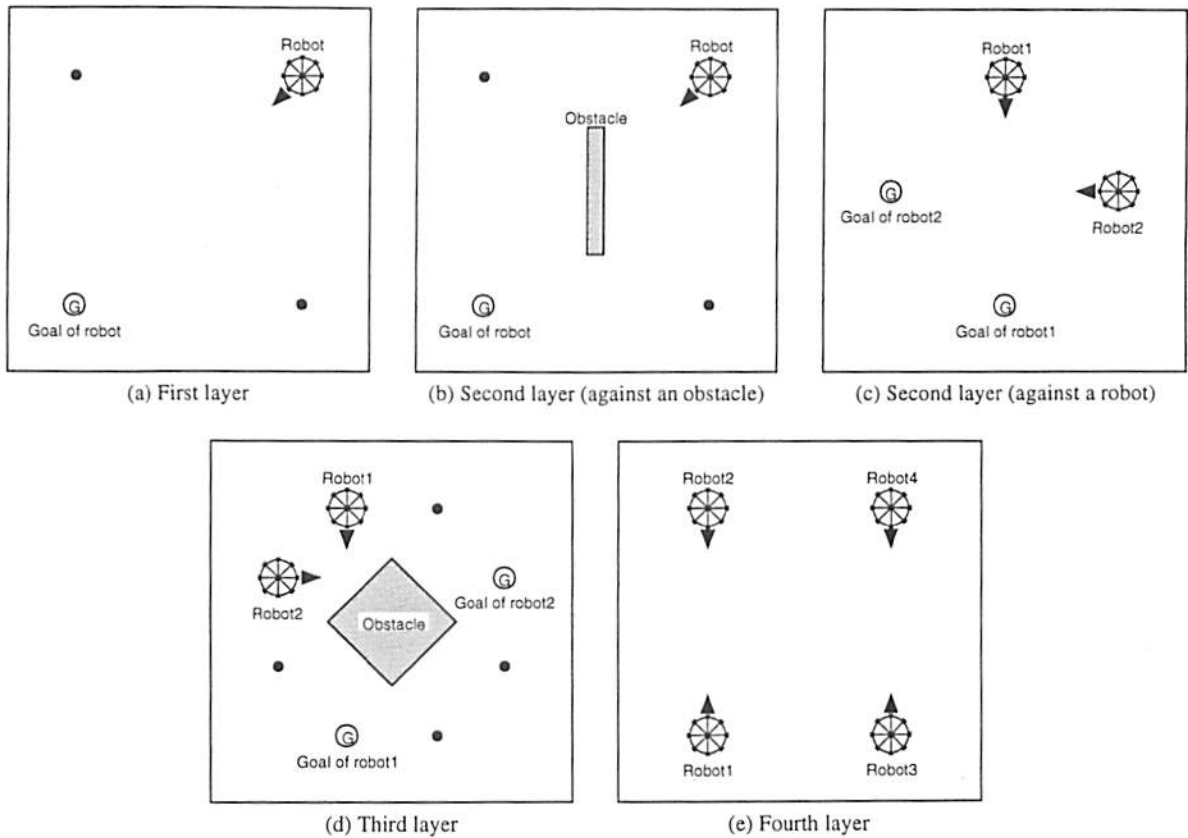


Fig. 20 Environment of learning for each layer

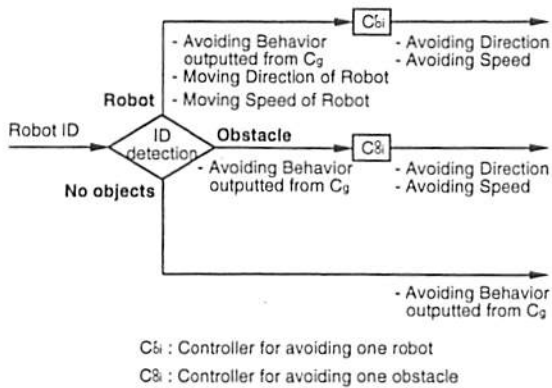


Fig. 21 Structure of learning controller for single object avoidance C_{O_i}

れる。一方、物体を検出なかったセンサユニットに対するコントローラにおいては、特に回避すべき物体がないことから第1層からのゴールへ向かうための行動がそのまま出力される。

5.1.3 第3層：センサグループごとの複数物体回避

複数の物体を回避するための行動を学習する際に、一度に多数の物体を回避対象とすると急に問題が複雑になり、学習が終了するまでに多くの時間を要する。第3層の学習では、回避対象となる物体の数を制限するために8組のセンサユニットを三つのグループに分割し、そのグループごとに認識された複数の物体を回避する行動を獲得することを目的とする。ここでは、

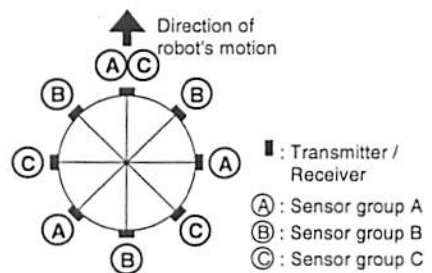


Fig. 22 Configuration of sensor group

Fig. 22 に示すようにセンサユニットをロボットの移動方向を基準にして A, B, C の三つに分割し、そのそれぞれをセンサグループと呼ぶ。ロボットの移動方向にある物体は特に衝突を起こしやすいと思われるので、移動方向のセンサユニットをセンサグループ A および C に割り当てる。これにより、移動方向のセンサによって得られる情報は学習の結果へ強く反映される。したがって、回避の対象となる物体の数は最大で三つである。この層には三つのセンサグループに対して三つのコントローラ C_{sA} , C_{sB} , C_{sC} が存在する。各コントローラの入力各センサグループに属する三つのセンサユニットによって検出された物体それぞれに対して第2層のコントローラからの出力となった行動である。Fig. 20 (d) に第3層における学習を行う環境を示す。中央には障害物が存在し、その周辺において2台のロボットが接近する。それぞれのロボットは、ゴールに到達するため

に、他のロボットおよび障害物を同時に回避しなければならない。出力はセンサグループごとに検出された複数の物体を回避するための行動である。

5.1.4 第4層：総合的な複数物体の回避

第4層は最終層であり、すべての入力情報を考慮した総合的な衝突回避を目的とする。コントローラの入力第3層の各センサグループに対するコントローラからの出力となった回避行動である。Fig. 20(e)に第4層における学習を行う環境を示す。4台のロボットが2組みに分かれてそれぞれ向かい合っており、それぞれのゴールは向かい合った相手のロボットの位置に設定されている。すなわち、4台のロボットは中央付近ですれ違うこととなり、ゴールに到達するためには複数のロボットを同時に回避する必要がある。以上により、本来の目的であるすべてのセンサユニットによって検出された物体のすべてを回避しながらゴールへ向かうための回避行動を学習する。

一連の学習では、前述の Fig. 20 の各図に示すロボットとゴールの位置関係に対する衝突回避に特化せずに、より汎用的なコントローラを構築するべきである。したがって、第4層を除く各階層では、Fig. 20 に示す位置を含む図中の黒点で示した点上にロボットのスタートおよびゴールをランダムに設定することとした。

5.2 状態数の改善

階層化学習における各階層の状態の数を Table 1 に示す。第1層では、そこで認識される状態の数はゴールの方向およびロボット自身の速度によって表現されるので、この層における状態数は 3.2×10^1 である。第2層では、検出された物体がロボットであるときにその移動方向および速度によって表現される32通り、そして物体が障害物であるときの1通りの計33通りの状態が LOCISS により認識される。さらに、第1層から出力された回避行動を考慮し、各センサユニットにおける状態数は約 1.1×10^3 である。LOCISS は八つのセンサユニットを持つので、最終的に状態数は8倍の約 8.4×10^3 となる。第3層では、第2層のうち各センサグループに属する三つのコントローラから出力された回避行動に基づいて各センサグループにおける状態数は 32^3 となる。したがって、この層の全体の状態数はその3倍の約 9.8×10^4 である。第4層では、第3層の三つのコントローラから出力される回避行動に基づいて状態数は約 3.3×10^4 となる。システム全体に対する状態数はこれらの各層の数の総和として表され、約 1.4×10^5 である。一方、階層化をせずに学習を行う際の状態数は、ゴールの方向、ロボット自身の移動速度、各センサから得られる情報（検出されたロボットの行動、検出された障害物の存在および物体が検出されない状態）の8チャンネル分のすべてを積算して

$$8 \times 4 \times (32 + 1 + 1)^8 = 5.7 \times 10^{13} \quad (4)$$

のように表され、これと比較すると階層化手法における状態数が激減していることが分かる。各状態には回避方向および速度の組み合わせで構成される32通りの回避行動が存在し、各層に必要なスコアの数は状態数の32倍となるため、スコアの総数は約 4.5×10^6 である。このとき、スコアを1バイトで表すと、システム全体に要するメモリのサイズは約4.5Mバイトとなる。これはロボット上に搭載することが十分に可能な容量であると思われる。

6. 衝突回避実験

群ロボット環境における一連の提案手法の有効性を示すために、実際のロボットを用いた衝突回避実験を行う。適応的に回避行動を獲得するために前述のように学習を行う場合、実際のロボットを用いることは時間的および物理的に現実的ではない。ここでは、実機による実験を行う前に、シミュレーション環境においてスコアベースを構築し、それを実ロボットに移すことにより実機による衝突回避を実現する。

6.1 スコアベースの構築

シミュレーション環境における衝突回避実験の条件を次のように設定した。前述のコントローラによる行動の選択の後、その行動の実行および一連の学習を経て、次の行動を選択するに至るまでの時間を動作周期とする。このとき、ロボットの動作特性に対して動作周期が短すぎる場合、選択された回避行動が実現する前に次の行動が選択されてしまうことから、結果的にどの行動も実行されず学習がうまく進行しない。また、周期が長すぎる場合、一回の行動時間が長くなりすぎてリアルタイムに衝突回避を実現することができない。そこで、動作周期は実験で用いる全方向移動ロボットの動作特性を考慮して200[msec]とした。このとき、動作周期と比較して、LOCISSによるロボット間の一度の通信に要する時間は非常に小さいことから、一度の動作周期において複数の通信を行うこととする。したがって、後述の実際のロボットを用いた実験においても、通信の信頼性を向上させることができ、シミュレーション環境で想定している理想の通信状態に近づけることができる。また、LOCISSによって送信するセンサコードは、動作周期ごとにロボットの移動方向に基づいて変化するため、同一の動作周期内であれば、複数の通信を通して同じコードを示すこととなり、安定して伝送することができる。一方、各階層における学習の回数は、学習の収束性を考慮して1,000回とした。以上の条件で各階層の学習計算を行ったところ、ロボットは与えられた状態に対して柔軟に衝突を回避することが可能となった。そのときの各階層における衝突回避の例を Fig. 23 に示す。

Fig. 23(a)は第1層におけるゴールへ向かう行動の学習の結果を示している。この環境では、ロボットの移動経路上には障害物は存在せず、ロボットはゴールに向かってほぼ一直線に移動することが確認できる。

Fig. 23(b)は第2層における一つの障害物に対する衝突回避の学習の結果を示している。ロボットは、回避の対象となる障害物を検出するまでは、ゴールに向かって移動する。LOCISS

Table 1 The number of situations for each layer

	The number of situations	
1st layer	8×4	$= 3.2 \times 10^1$
2nd layer	$(32 + 1) \times 32 \times 8$	$= 8.4 \times 10^3$
3rd layer	$32 \times 32 \times 32 \times 3$	$= 9.8 \times 10^4$
4th layer	$32 \times 32 \times 32$	$= 3.3 \times 10^4$
Total		1.4×10^5

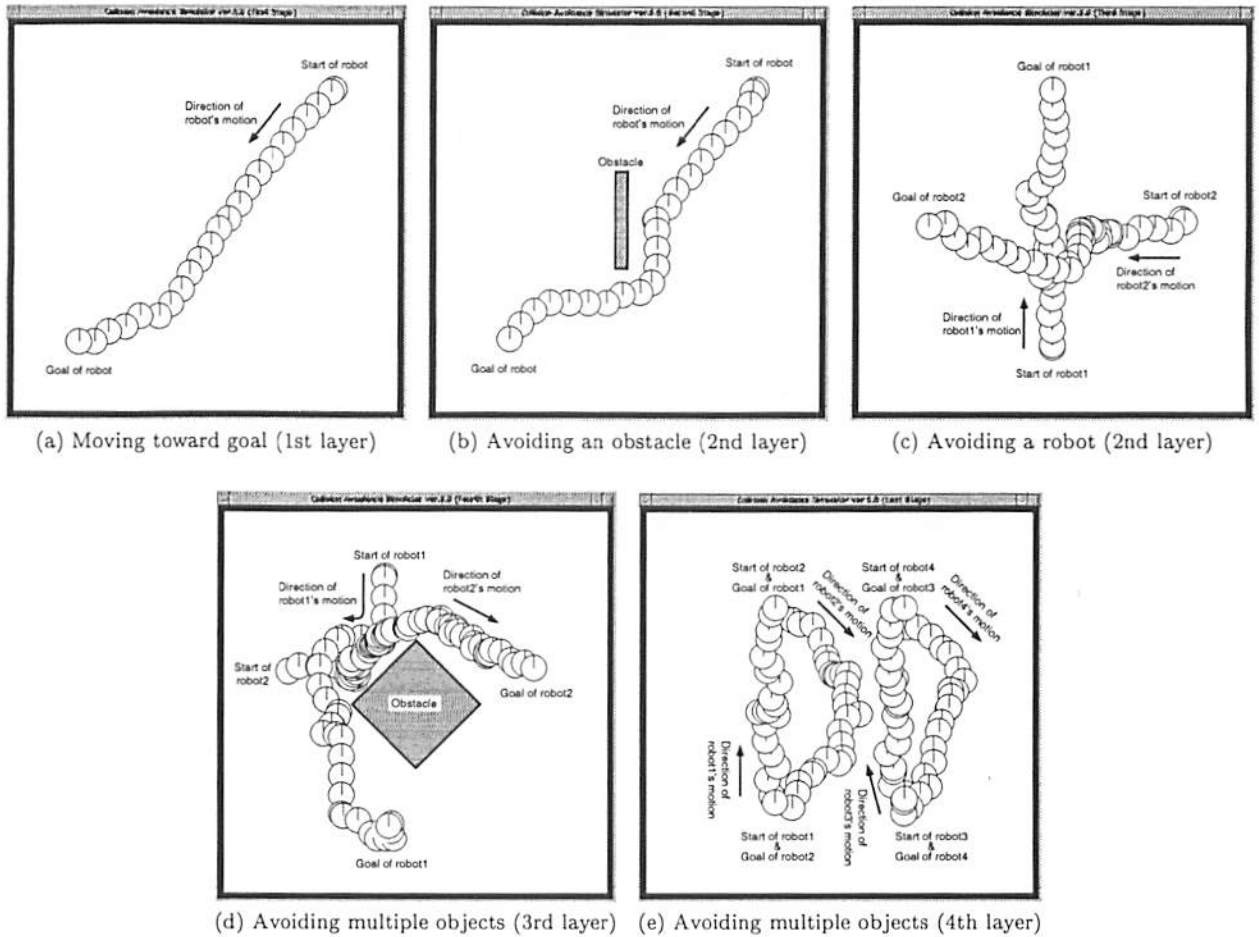


Fig. 23 Results of collision avoidances in simulation environment

によって障害物を検出すると、ロボットは移動経路を左に変更することにより衝突を回避し、また衝突の恐れがなくなると、再び直線的な経路をとりゴールに向かって移動する。以上から、第2層における学習は第1層の学習の結果を踏まえて行われていることが分かる。

Fig. 23 (c) は第2層における1台のロボットに対する衝突回避の学習の結果を示している。一つの障害物に対するときと同様に、2台のロボットは環境の中央付近で遭遇するまではそれぞれのゴールに向かって移動する。接近して互いを自分以外のロボットであると識別すると、お互いに相手の行動を認識し、それらを考慮した衝突回避を実現する。すなわち、ロボット1はロボット2が停止して待つことを認識し、ゴールに向かって直線的に移動している。このときロボット2は停止し、ロボット1が通り過ぎると、ロボット2はロボット1の後ろに回り込むように回避し、完全に衝突の恐れがなくなると直線的にゴールに向かって移動している。

Fig. 23 (d) は第3層における複数の物体に対する衝突回避の学習の結果を示している。この環境では、2台のロボットと障害物が一つ存在する。2台のロボットが障害物の近辺で遭遇したとき、ロボット1はロボット2が停止している間に障害物およびロボット2を回避し、ロボット2はロボット1が通り過ぎ

た後に障害物を回避しながら移動を開始したことが確認できる。その後、2台のロボットは障害物を回避し、それぞれゴールに向かって移動する。

Fig. 23 (e) は第4層における複数の物体に対する衝突回避の学習の結果を示しており、得られるすべての情報を反映した4台のロボット間の衝突回避が実現されていることが確認できる。ここでは、ロボット1と2およびロボット3と4をそれぞれ向き合うように設置し、それぞれのロボットのゴールを向かい合ったロボットのスタート位置に設定した。それぞれがゴールに向かって移動し、お互いを検出したとき三つの衝突回避が同時に行われている。これは群ロボット環境における衝突回避の典型的な例であり、複数物体が存在する複雑な状態において、提案した手法が有効であることを示している。この層における衝突回避については、後述の実機による衝突回避実験において同様の環境を用いるので、そこで詳細に述べる。

以上のことから、単純な環境から複雑な環境へと階層的に学習を行うことが可能であり、最終的に実機に搭載するべきスコアベースを構築することができた。

6.2 全方向移動ロボット

Fig. 24 に本実験で用いる自立型全方向移動ロボットを示す。このロボットは前後方向、左右方向および回転方向の三つの自

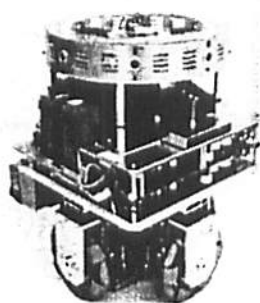


Fig. 24 Omni-directional mobile robot and LOCISS

由度をそれぞれ別のアクチュエータで独立に駆動することが可能であり、それらの動きを組み合わせることにより自由な方向に移動できる [18]。各アクチュエータにはエンコーダが内蔵されており、その回転数からデッドレコニングを行うことができる。大局的通信デバイスとして無線 LAN を、局所的通信デバイスとして LOCISS を搭載している。

6.3 実ロボットによる衝突回避

衝突回避実験を行うために、シミュレーション環境において構築したスコアベースをすべての実ロボットに搭載する。これは、すべてのロボットの特定の状態に対する回避行動の選択確率は均一なものとなるが、実際の選択は確率的に行われるため、まったく同一の動作を意味するものではない。

Fig. 25 に実験環境を示す。ロボット 1 から 4 までの 4 台のロボットが存在する。ロボット 1 と 2 およびロボット 3 と 4 をそれぞれ 4.0 [m] 離れて向かい合わせ、ロボット 1 と 3 およびロボット 2 と 4 の間隔を 2.0 [m] とした。各ロボットの三角形の頂点はそのロボットの正面を示す。ロボット 3 および 4 の側面方向 1.5 [m] の距離には障害物として壁を設置した。また、それぞれのロボットのゴールを前方 4.0 [m] の地点、すなわち向かい合った相手のロボットのスタート地点とした。

以上のような環境において、4 台のロボットが衝突回避をした際の各ロボットの軌跡を Fig. 26 に示す。この図は、1 [sec] ごとに各ロボットの位置を計測し、プロットすることにより作成した。4 台のロボットは、本来それぞれのゴールに向かって直進するが、他のロボットあるいは障害物を検出することにより、左側に軌道を変更することにより衝突を回避していることが確認できる。図中の A~H は Fig. 27 に示すロボット間の通信記録と実際の回避行動の対応をつけるために付加した記号である。Fig. 27 は、時間の経過に対してロボット間で行われた通信を矢印で表している。すなわち、各ロボットの軸線（二重線）に達する矢印は、どちらかの隣接するロボットから信号を受信したことを示している。ロボット 4 の記録に現れる、軸線上の塗りつぶし部分は自分自身の信号の受信を示し、障害物が検出されたことを意味する。図中の A~H は軌跡上の記号に対応し、その通信が行われたときにロボットが Fig. 26 の軌跡上のどこに存在したかを示す。

Fig. 28 に、複数のロボット間の衝突回避の例として Fig. 26 におけるロボット 2 の A~D 部分の拡大図を示す。A 地点では、

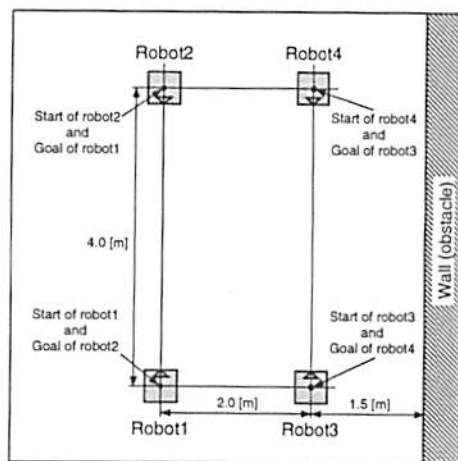


Fig. 25 Environment of collision avoidance experiment

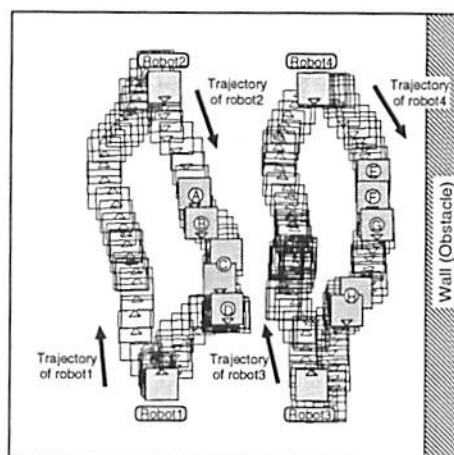


Fig. 26 Result of collision avoidance experiment

ロボット 2 は Fig. 28 の矢印で示す方向からロボット 1 の信号を受信し、他のロボットからの信号が存在しないので (Fig. 27 参照)、ロボット 1 を回避するために左前方に移動した。B 地点では、ロボット 1 および 3 からの信号を受信し、これらを回避するための行動として、双方のほぼ中央を通過する軌道へ向かって移動する行動を選択したことが確認できる。C 地点では、ロボット 1 はすでに後方に遠ざかっており、ロボット 2 はロボット 3 からの信号だけを進行方向の左側から受信した。ここでは、ロボット 2 および 3 とともに並行して移動し、お互いの衝突を回避した。D 地点では、ロボット 3 も後方に通り抜けており、ロボット 2 は周囲に物体が存在しないため、ゴールの方向に向けて移動した。

Fig. 29 に複数の物体に対する衝突回避の例として Fig. 26 におけるロボット 4 の E~H 部分の拡大図を示す。Fig. 27 によれば、E、F および G 地点ともに、ロボット 4 は自分自身が発した信号を受信することにより壁を検出した。このとき、F 地点に見られるように、壁の反対側から他のロボットを検出した場合においても障害物に接近しすぎることなく、同時にロボットとの衝突を回避する行動を選択したことが確認できる。H 地点では、ロボット 3 からの信号を受信したが、衝突を起こす恐

以上のように、本手法を用いることにより、複数のロボットに
 対する衝突回避および複数の物体に対する衝突回避とともに実
 現できた。本学習手法では、LOCISSによる通信に干渉が発生
 する場合は想定していない。干渉が発生する状況として、ある
 ロボットに対してほぼ同じ方向、かつそのロボットのLOCISS
 による検出可能範囲内に複数のロボットが存在する場合などが
 考えられる。しかし、本実験で対象としているような、ロボット
 の密度が疎な環境においては、本衝突回避手法を適用すること
 により、そのような状況は起こらないと考えられる。また、ロ
 ボットの密度が密な環境においては、干渉の発生も考慮すべき
 であるが、LOCISSによって検出した干渉状態をロボットの周

Fig. 29 Collision avoidance for multiple objects: Robot4

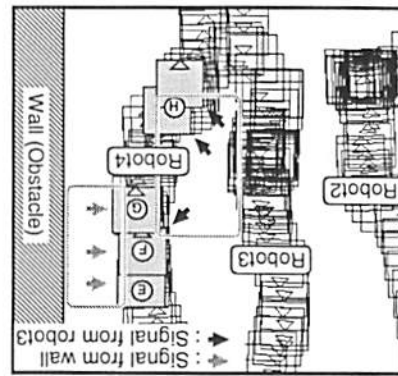


Fig. 28 Collision avoidance for robots: Robot2

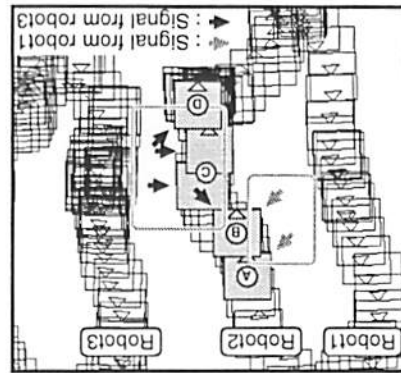
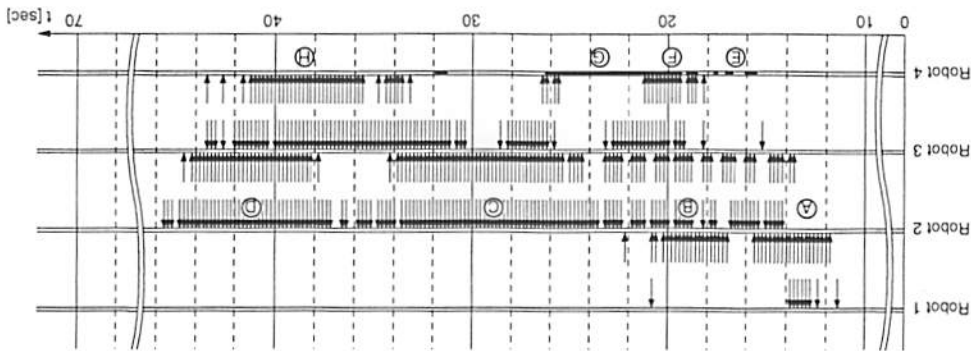


Fig. 27 Log of communication between robots



群ロボット環境における局所的通信に基づく衝突回避

本論文では、群ロボット環境における衝突回避を実現するた
 めに、以下の一連の検討を行った。局所的通信に関しては、赤
 外線の特長を利用したロボットの周囲の状態を認識する手法を
 提案すると同時に、群ロボット環境において問題となる赤外線
 の干渉を検出する機能を含んだ伝送情報の符号化手法に基づい
 てLOCISSを開発した。また、LOCISSを用いることにより、
 干渉の検出ならびに通信の局所性が実現可能であることを確認
 するとともに、LOCISSによって認識される周囲の状態に応じ
 てルールに基づいた衝突回避を実現した。回避行動の適応的な
 獲得手法の開発では、強化学習を導入することにより、複数の
 ロボットが存在する複雑な状態に対して衝突を回避する行動を
 ロボット自身が学習する手法を提案した。また、シミュレーショ
 ン環境において実験を行うことにより、提案手法による学習の
 収束性を確認した。学習手法の実ロボットへの実装では、学習
 における状態の数の削減を目的として、学習の階層化手法を提
 案した。これにより、システム全体の学習における状態の数は
 10^5 オーターまで激減し、実際のロボットに搭載することが可
 能となった。最終的に、これらの一連の手法を組み合わせ、実
 機を用いて群ロボット環境における衝突回避を実現した。

提案した手法は、ロボットの周囲の局所的な情報だけに基
 づいているが、オペレータによる経路計画など、大局的な手法と
 の組み合わせによる柔軟なシステムの構築が今後の課題である。
 謝辞 赤外線通信センサシステムLOCISSの開発に当たり
 て尽力頂いた科学技術振興事業団川入学習動態脳プロジェクト
 の琴坂信哉氏に感謝する。

参考文献

[1] H. Asama, T. Fukuda, T. Arai, I. Endo, Eds.: Distributed Au-
 tonomous Robotic Systems, Springer Tokyo, 1994.
 [2] H. Asama, T. Fukuda, T. Arai, I. Endo, Eds.: Distributed Au-
 tonomous Robotic Systems 2, Springer Tokyo, 1996.
 [3] T. Lueh, R. Dillmann, P. Dario, H. Wern, Eds.: Distributed
 Autonomous Robotic Systems 3, Springer Berlin, 1998.
 [4] 久保田孝, 橋本孝紀, 原島文雄: "ローカルフロンティアによる移動ロ
 ボットの経路計画", 日本ロボット学会誌, vol.7, no.4, pp.267-274,
 1989.
 [5] J. Borenstein and Y. Koren: "Real-Time Obstacle Avoidance

- for Fast Mobile Robots." The IEEE Trans. on Systems, Man and Cybernetics, vol.19, no.5, pp.1179-1187, 1989.
- [6] 石川繁樹, 浅香俊一: "移動障害物を含むような動的変化を伴う走行環境における自律移動ロボットの走行誘導方式", 日本ロボット学会誌, vol.11, no.6, pp.856-867, 1993.
- [7] 八木康史, 川戸慎二郎, 辻 二郎: "全方向視覚センサ COPIS を用いた移動物体との衝突回避", 電子情報通信学会誌, vol.J74-D-II, no.7, pp.908-917, 1991.
- [8] 新井民夫, 木村久幸, 前田健太郎, 太田 順, 梅田和昇: "移動ロボット相互の位置・姿勢実時間計測システムの開発", 日本ロボット学会誌, vol.12, no.3, pp.472-478, 1994.
- [9] 尾崎功一, 浅間 一, 石田慶樹, 松元明弘, 遠藤 勲: "自律分散型ロボットシステムのための無線通信システムの開発", 精密工学会誌, vol.61, no.8, pp.1137-1140, 1995.
- [10] 尾崎功一, 浅間 一, 石田慶樹, 松元明弘, 遠藤 勲: "通信を用いた複数移動ロボットの相互衝突回避", 日本ロボット学会誌, vol.14, no.7, pp.961-967, 1996.
- [11] S. Yuta and S. Premvuti: "Coordination Autonomous and Centralized Decision Making to Achieve Cooperative Behaviors between Multiple Mobile Robots," The Proc. of International Conference on Intelligent Robots and Systems, pp.1566-1574, 1992.
- [12] 鈴木昭二, 新井義和, 琴坂信哉, 浅間 一, 嘉悦早人, 遠藤 勲: "マルチ移動ロボット環境における衝突回避のための局所的な通信を利用したセンサシステムの開発", 日本機械学会論文集 (C 編), vol.62, no.602, pp.3752-3758, 1996.
- [13] 畷見道夫: "実例に基づく強化学習", 人工知能学会誌, vol.7, no.4, pp.697-707, 1992.
- [14] 浅田 稔, 野田彰一, 依積田健, 細田 耕: "視覚に基づく強化学習によるロボットの行動獲得", 日本ロボット学会誌, vol.13, no.1, pp.68-74, 1995.
- [15] 山口智浩, 三浦正宏, 谷内田正彦: "適応型模倣による複数個体の強化学習", 人工知能学会誌, vol.12, no.2, pp.323-331, 1995.
- [16] 黒山和宏, 山田和明, 中村陽一郎, 大谷和博, 上田完次: "実例に基づく強化学習の手法 (第 2 報, 小型移動ロボットによる行動学習実験)", 日本機械学会ロボティクス・メカトロニクス講演会 '97 講演論文集, pp.1159-1160, 1997.
- [17] 港 隆史, 浅田 稔: "環境の変動に適應する移動ロボットの行動獲得", 第 16 回日本ロボット学会学術講演会予稿集, pp.399-340, 1995.
- [18] 浅間 一, 佐藤雅俊, 嘉悦早人, 尾崎功一, 松元明弘, 遠藤 勲: "3 自由度独立駆動型全方向移動ロボットの開発", 日本ロボット学会誌, vol.14, no.2, pp.249-254, 1996.
- [19] 新井義和, 鈴木昭二, 琴坂信哉, 浅間 一, 嘉悦早人, 遠藤 勲: "赤外光局所的通信センサシステムを用いた複数自律移動ロボットの衝突回避", 第 13 回日本ロボット学会学術講演会予稿集, pp.847-848, 1995.
- [20] 新井義和, 片岡靖志, 藤井輝夫, 浅間 一, 嘉悦早人, 松元明弘, 遠藤 勲: "強化学習に基づく複数自律移動ロボットの衝突回避行動生成", 日本機械学会ロボティクス・メカトロニクス講演会 '97 講演論文集, pp.413-414, 1997.
- [21] 新井義和, 藤井輝夫, 浅間 一, 嘉悦早人, 遠藤 勲: "階層的強化学習に基づく複数移動ロボット間の衝突回避", 日本機械学会ロボティクス・メカトロニクス講演会 '98 講演論文集, 2A12-1, 1998.



新井義和 (Yoshikazu Arai)

1970年8月3日生。1998年埼玉大学大学院理工学研究科博士後期課程修了。1996-1998年理化学研究所ジュニア・リサーチ・アソシエイト。1998年岩手県立大学ソフトウェア情報学部助手となり現在に至る。博士(工学)。自律分散型ロボットシステムの研究に従事。1998年4th Int. Symp. on Distributed Autonomous Robotic Systems Best Paper Award 受賞。IEEE, 日本機械学会などの会員。(日本ロボット学会正会員)



浅間 一 (Hajime Asama)

1959年1月18日生。1984年東京大学大学院工学系研究科修士課程修了。1986年理化学研究所化学工学研究室研究員補, 生化学システム研究室研究員, 先任研究員を経て, 現在研究基盤技術部副主任研究員。1998年埼玉大学地域共同研究センター客員教授。自律分散型ロボットシステム, 複数移動ロボットの協調, 群知能, ロボットによる保全技術, バイオプロセスの知能化技術の開発に従事。1995年日本機械学会ロボメック賞, 1996年日本機械学会ロボメカ講演会 Best Poster 受賞。RoboCup に, UTTORI United Team の一員として参加。RoboCup-98 Japan Open 人工知能学会賞受賞。工学博士(東京大学)。Distributed Autonomous Robotic Systems (Springer-Verlag, Tokyo) の第1巻および第2巻の編者。IEEE, ニューヨーク科学アカデミー, 日本機械学会, 精密工学会などの会員。(日本ロボット学会正会員)



嘉悦早人 (Hayato Kaetsu)

1948年10月30日生。1971年東京理科大学中退。1971年理化学研究所同位元素研究室勤務。1981年同化学工学研究室へ移籍。同位体分離の研究。自律分散型ロボットシステムの研究に従事。現在, 先任技師として理化学研究所生化学システム研究室に在籍。日本原子力学会, 日本機械学会, 精密工学会の会員。



藤井輝夫 (Teruo Fujii)

1964年4月5日生。1988年東京大学工学部船舶工学科卒業。1991-1993年日本学術振興会特別研究員。1993年東京大学大学院工学系研究科船舶海洋工学専攻博士課程修了。博士(工学)。同年東京大学生産技術研究所客員助教授。理化学研究所勤務を経て, 1999年より東京大学生産技術研究所海中工学研究センター助教授。IEEE, ACS, 電気学会などの会員。(日本ロボット学会正会員)



鈴木昭二 (Sho'ji Suzuki)

1964年4月10日生。1993年筑波大学大学院博士課程工学研究科修了。博士(工学)。日本学術振興会特別研究員, LAAS, 理化学研究所を経て, 1995年大阪大学助手となる。2000年より公立はこだて未来大学助教授となり現在に至る。移動ロボットの研究に従事。日本機械学会, IEEE 会員。(日本ロボット学会正会員)



遠藤 勲 (Isao Endo)

1940年8月14日生。1970年東京大学大学院工学研究科博士課程修了。1970年理化学研究所化学工学研究室勤務, 生化学システム研究室に改名。現在同主任研究員。1989年より埼玉大学理工学研究科客員教授兼任。バイオプロセスの異常診断, バイオセパレーションエンジニアリング, 新しい固体培養プロセスシステムの開発, 光感応性酵素, 群知能ロボットの研究に従事。1978年化学工学協会論文賞受賞。1993年市村産業賞功績賞受賞。1994年ヘルシンキ工科大学名誉工学博士。1999年科学工学会賞受賞。2000年科学技術庁長官賞研究功績賞受賞。ニューヨーク科学アカデミー, アメリカ化学工学会, アメリカ化学会, 化学工学会などの会員。(日本ロボット学会正会員)