

〔研究〕

自律学習機能と事情通口ボットの研究

Research on Autonomous Learning Functions and Jijo-2 Office Robot

松井俊浩

T. MATSUI

In the latter half of the RWC project, autonomous learning function field has been established as one of the R and D program for the Real-World Intelligence Technologies. This report describes the research goals, executive formation and achievements till 1999. Above all, the research at ETL RWI Center to develop the office robot, Jijo-2, which can navigate in office environment without predefined map and interact with people with spoken language, is described.

§ 1 自律学習機能領域の研究の目的と背景

パソコンとインターネットに導かれて情報技術が急速に普及する一方で、情報技術を利用できる者とできない人々の格差、デジタルデバイドが顕在化して来ている。将来誰もが情報機器を縦横に使いこなすことができるようになるために必要な技術として、自律学習機能領域では次の3点が重要と考えている。

1.1 多様な情報の扱い

現在のコンピュータが外界とインタフェースする主な手段は、マウス+キーボードとディスプレイそして機械的な音に限られている。テキストや画像を閲覧するには適しているが、マルチメディア情報を入力、認識することは困難であるし、キーボードに拒否反応を示す人も少なくない。音声や環境音を識別したり、ジェスチャを解釈したり、距離、温度、明るさ、物理的な動きなど実世界に満ちあふれる情報を認識する能力およびこれらを統合的に解釈して総合的な判断を下す機能が必要である。

1.2 学習・適応型の情報機器

現在の情報機器は押しつけ型である。ユーザーがカスタマイズできる部分は限られており、ユーザーが機械操作に習熟することへの依存の度合いが大きく、変化する

環境に適応することができない。たとえば、工場のロボットは、ロボット用に整備された環境での繰り返し作業には適しているが、異なる作業への適応能力が乏しく、したがって困難な作業には熟練者による遠隔作業が必要となる。情報機器が応用のすそ野を広げるためには、機械の側から環境や人に適応するための学習機能の実現が不可欠である。

1.3 身近なコミュニティでの能動的な情報収集機能

携帯電話が普及し、コンピュータを持ち歩けるようにはなったが、届け物をしてくれるわけではない。たいいてい情報は電線を通じて送ることができるが、社会生活は、物流・移動や対面式会話と無縁ではありえない。現在のインターネットは前者をサポートするが、後者には無力である。そのため、地球の裏側の工場で作られる製品の情報はたやすく入手できても、隣のオフィスで誰が何をしているのかといった、身近な情報へのアクセスはあいかわらずの隔靴搔痒の感がある。このような身近な情報を能動的に収集する機能が求められる。

以上の観点から、自律学習領域では、オフィス内を能動的に徘徊し、ソナー、視覚、音声対話、環境音などのさまざまな情報を統合的に利用し、広範なオフィス環境で環境地図や人々の行動を学習することで、人探し、道案内や届け物などの軽作業が達成できるような自律的な

KEY WORDS : オフィスサービスロボット, 移動ロボット, ナビゲーション, 地図学習, 音声対話

移動ロボットシステムの構築を目指す。

§ 2 研究項目

以上に述べた観点から、自律学習機能領域での研究の目標を、「時間的にも変容しうる多様な(オフィスやホーム)環境で、さまざまなセンサ情報に基づいて自律的に移動しながら、環境やそこで働く人々に関する情報を収集・提供するシステム(特に移動ロボットシステム)の研究開発」と定めた。そしてこのようなシステムのイメージに近づくために、次のような重点的研究が必要な要素を決めた。

1. 視覚技術：実環境で場所を同定したり、人を見つけて認識する手法、ナビゲーションのために実時間で動く物体を追跡するための視覚技術。
2. 聴覚技術：実環境でロバストな音声認識を達成するための音源方向の同定、雑音の除去、音場特性の推定、音声以外の環境音を認識することで場所を同定したり音声を抽出する。
3. ナビゲーション技術：視覚、聴覚、ソナー、レーザー距離計などの情報を相補的、総合的に利用することで屋内環境をロバストに航行する手法の開発。
4. オフィス対話技術：人と共存して活動する使いやすいシステムとするために、人と自然な音声会話を交わすことで情報を収集・提供する手法の開発。

5. 環境情報と行動の統合的学習：あらかじめ地図を与えることなく移動ロボットの自律的な探索行動によって上記の各種センサー情報を収集し、それを場所に関連づけることで環境の構造を学習する。その中でタスクを実行するための行動を獲得していく手法の開発。
6. 統合化アーキテクチャ：これらの各機能をモジュールとして構成し、それらの発現パターンを状況と考え、次取るべき行動を学習的に獲得できるような、リアルタイム制御機能と記号的計画機能の両方が並列的に動作するようなマルチエージェントアーキテクチャの開発。

§ 3 研究実施体制

本領域の研究参加機関とその主な役割分担は、以下のとおりである。

電子技術総合研究所RWIセンター

- ・ 視覚による場所、人発見、環境中の雑音低減や音源同定。
- ・ 音声による地図学習、道案内や人捜しの対話。
- ・ 確率ネットワークを用いた知識や環境情報の表現、および確率ネットワークに基づいて行動を生成するアーキテクチャ。
- ・ これらの機能に基づいてオフィスで道案内や人探し

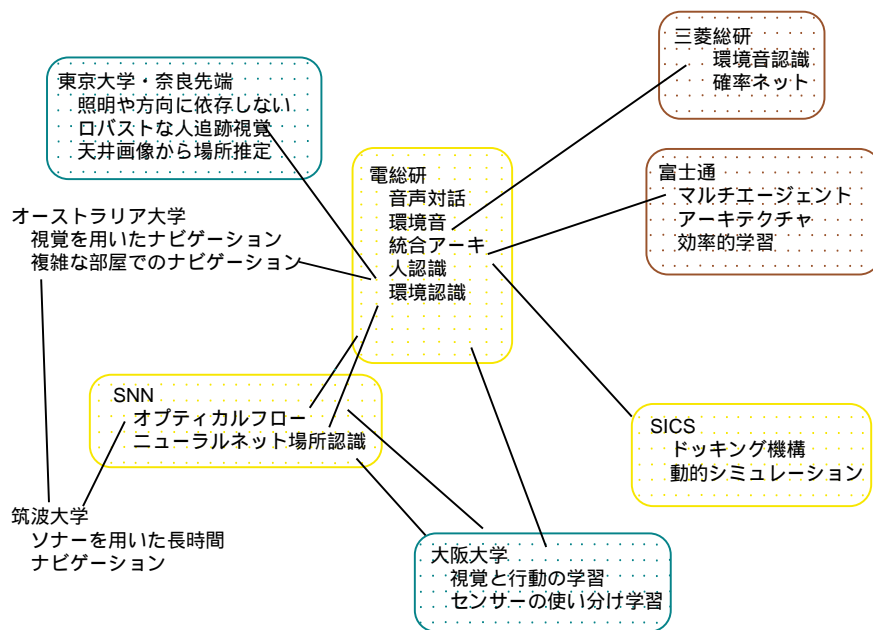


図1 研究実施体制とテーマ分担

を行うロボットを「事情通ロボット」として統合する。
富士通研究所

- ・分散知能アーキテクチャによる移動ロボットの行動および環境の構造化の学習。
- ・状況数に応じて過大になりがちな探索空間の中で効率的な学習を行う手法の研究。

三菱総研(MRI)

- ・環境音の認識、擬音語を用いた環境音に関する対話、その実時間システムの実現。

アムステルダム大学 基礎ニューラルネットワーク研究所(SNN)

- ・ニューラルネットワークを用いて画像情報から場所を同定し、ロバストなナビゲーションを行う手法。

スウェーデン計算機科学研究所(SICS)

- ・移動ロボットが長時間安定して行動するための充電機構やドッキング機構、またそれらの特異な場所へガイドするデバイスの開発。

東京大学工学部 情報システム工学研究室および奈良先端大 情報科学研究科 ロボティクス講座

- ・変化する照明環境下での視覚による安定した人の発見と追跡。
- ・天井画像を用いたナビゲーション。
- ・ビジョンによる床の発見と障害物の同定。

大阪大学大学院工学研究科 創発ロボット工学講座

- ・環境およびその変動に対処する最適なセンサーの利用法、組み合わせを学習する技術の研究。
- ・環境が変わったことを検知し、変化が生じた部分だけを局所的に再学習する方法。

筑波大学 電子情報工学系 知能ロボット研究室

- ・ソナーと画像を用いた長距離のナビゲーション手法。

オーストラリア国立大学

- ・障害物が多い狭隘な環境でのロバストなナビゲーション、人追跡行動。
- ・高速な動作追跡用ビジョンシステム。

システム全体を統合する仕上がりイメージおよびコンセプトは、分散研との議論を統合して電総研RWIセンターがとりまとめた。事情通ロボット(後述)を典型的な目的システムと設定し、そのために必要な要素技術を各分散研と再委託先大学が提供し、それを電総研で検証しながら統合化技術を分散研企業にフィードバックする形態を取る。

1996年に後期計画についての議論を開始し、1997年

に後期基本計画が策定された。1997年と1998年度は、それぞれの参加研究機関で担当の要素技術についての研究開発を行い、1999年度から技術の統合に着手する。統合のプラットフォームとなるのは電総研の事情通ロボットであり、富士通及びオーストラリア国立大学には可能な限りこれと同等なシステムを配備する。これらに対して、ナビゲーション手法、環境音認識、人認識などのモジュールを搭載し、各環境での検証を行う。2000年度はその結果のフィードバックと改善に当て、2001年度はこれらすべてを統合したシステムのデモンストレーションを電総研と富士通が主体となって行う。

§ 4 プロジェクト後期中盤までの成果

1999年度までに、3節であげた研究項目について次のような成果を得た。

1. 視覚技術：距離や回転に依存しにくい顔発見アルゴリズムを開発した(電総研)^{9,10}。画像特徴に基づく場所同定をニューラルネットワークで実現した(SNN)³。トラッキングビジョンにより色情報と動き情報を統合した高速の領域分離手法を開発した(東京大学)⁷。
2. 聴覚技術：多チャンネルマイクロフォンアレイによるビームフォーミングにより音源方向を割り出し、雑音を取り除き、音声認識の初期処理を行うモジュールをDSPを用いて実装した(電総研)⁸。音声以外の環境音のうち単発音を擬音語として表現し認識する方法を提案した(MRI)⁴。
3. ナビゲーション技術：ソナーおよび画像を用いて障害物を回避しながら移動し、その経路を特徴点のシーケンスとして学習する手法を開発した(奈良先端、東大)^{6,18,21}。センサーのエラーを考慮して自己位置の修正を行う手法を開発した(筑波大)。画像から床を抽出し乱雑な室内で安定に移動する手法を開発した(ANU)^{22,23}。オプティカルフローを実時間で求め移動ロボットの進むべき方向をニューラルネットワークで学習する手法を開発した(SNN)³。環境と行動に応じて複数のニューラルネットワークが次第に機能を分化させていくしくみを実現し、環境の分節化とオフィス環境でのナビゲーションのシミュレーションを行った(電総研)¹¹。
4. オフィス対話技術：移動ロボットに音声認識を実

- 装し、場所や移動タスクに関する断片的な日本語発話をもとに会話を行う方法を研究した(電総研)。
5. 環境情報と行動の統合的学習：典型的な環境でのQ学習によって獲得された行動をより一般の環境に適應させる方法を提案した(大阪大学)²⁵⁾。ソナーによる空間の文節と会話によるラベル付けを統合し、確率ネットワークとして環境と行動を記述・獲得する方法を開発した(電総研)^{5,6)}。ロボットが「意識」を持つことで強化学習が加速されることを示した(SICS)⁵⁾。
 6. 統合化アーキテクチャ：移動機能、画像処理、対話、地図獲得などを並列動作するマルチエージェントとして構成し、イベント駆動方式で制御する方法を構築した(電総研)¹⁾。入札機構を備えたマルチエージェントネットワークでアテンションや競合解消を制御する方法を提案した(富士通)²⁾。

§ 5 終盤に向けての展開

終盤に向けて、各機能の高性能化と共に、成果の相互利用を促進し、統合システムの実装を開始する。次のような展開を図る。

- ・ 全方位ビジョンや天井画像から得られる特徴を用いて、また音の情報を付加的に利用することで、より多様な環境で場所を信頼性良く同定できるようにする(奈良先端、ANU、SNN)。
- ・ 動画像から人を発見し、人の動作を認識する(ANU、奈良先端)。
- ・ ソナー、レーザー距離計、視覚など複数のセンサー情報を用いて幾何的地図を獲得し、それに対話から得られる記号的情報を重畳させる地図表現を研究する(電総研)。
- ・ 残響のある実環境で確実な音声認識が行え、文脈情報の処理を加味して、場所やタスクに関するより自然な対話が交わせるシステムを実現する(電総研)。
- ・ 確率的に発生する事象の関係を認識する状況依存エージェントにより行動が制御されるようなマルチエージェントアーキテクチャを研究する(電総研)。

§ 6 電総研RWIセンターでの研究

オフィス内を移動しながら、道案内、人探し、届け物な

どを手伝い、情報を収集することでスケジュール調整をしてくれるような、Jijo-2ロボットシステムの構築を目指している⁴⁾。主な研究のゴールは、さまざまなセンサー情報を統合的に利用した複雑な学習システムを構成する方法の開拓にある。このような方向に向かって、次のような研究成果を上げてきた。

6.1 イベント駆動方式マルチエージェントコントローラ

ロボットは統合システムであり、並列的に実行される多数の機能要素モジュールを結合する必要がある。機能は、障害物の発見や音声の聞き取りなど外界のイベントにリアルタイムで反応する必要があるものと、対話や経路の計画、地図の獲得などリアルタイム性よりも記号的な熟考が必要なものがある。そこで、モジュールを二つの層に分解し、モジュールの独立性を高めることで並列動作とフォルトトレランスを可能にし、通信インターフェースを統一することで段階的な開発を可能にするアーキテクチャを提案した(図2)¹⁾。高レベルのプログラミングには、新情報プロジェクトの初期の成果でもある、ロボット用の機能を多く備えたオブジェクト指向言語 EusLispを用いている。

6.2 距離と回転の変化に強い顔認識

前期では、2次元図形の特徴を学習的に記憶し、特徴の位置に依存しないで図形を高速に発見し識別する局所相関法を開発した。これをオフィスのような環境での顔、あるいは物体認識に適用しようとすると、距離、回転、照明などの条件で初期の性能が得られない。これを改善するために、図3のように画像にLog-Polar変換を施すと、距離の違いはlogarithmicにしか影響せず、回転の影響は位置の変化となって吸収される。この方法により、オフィスにおける顔発見がより頑健になることを検証した^{9,10)}。

6.3 音源方向の同定と環境雑音除去による音声認識性能の改善

音声認識には、電総研が開発したHMMに基づく不特定話者連続音声認識NINJAを用いている。音声認識は一般に、周囲の雑音がない環境でマイクの近くで発声されることを想定しており、オフィス環境で普通にロボットに話しかけた場合認識率が大きく低下しうる。このよ

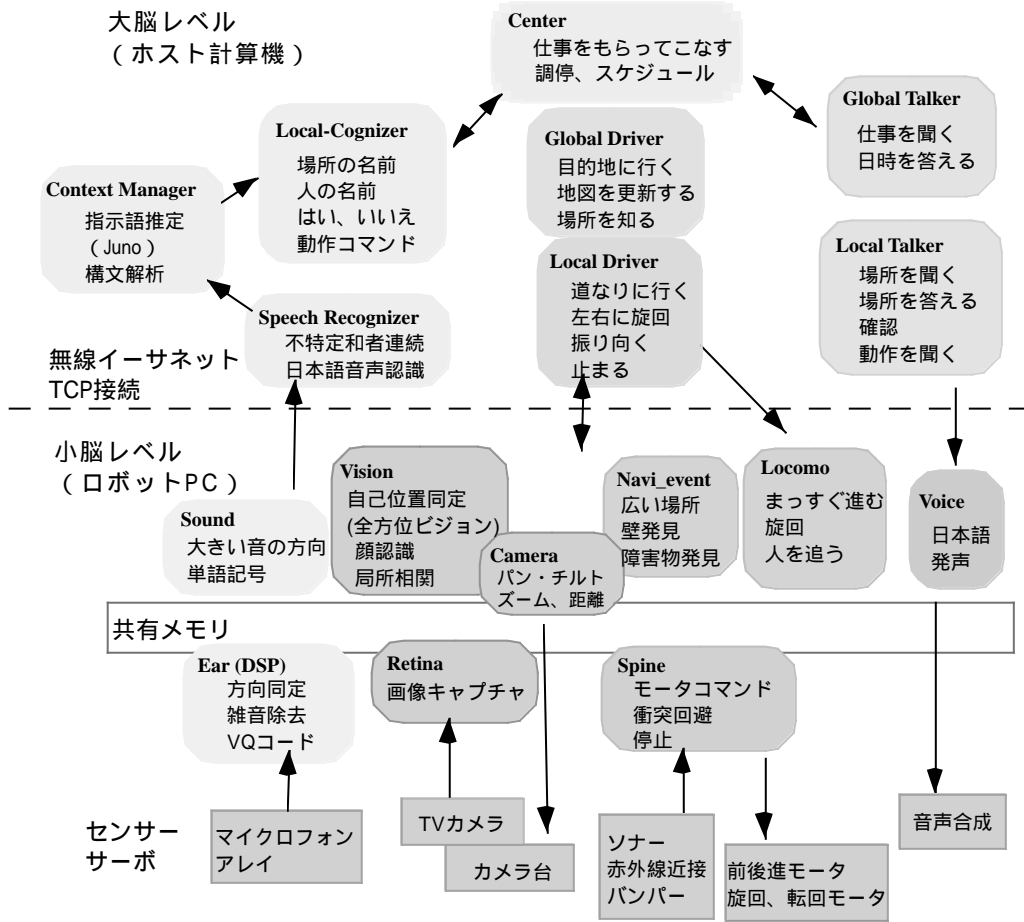


図2 事情通ロボットのマルチエージェントコントローラ

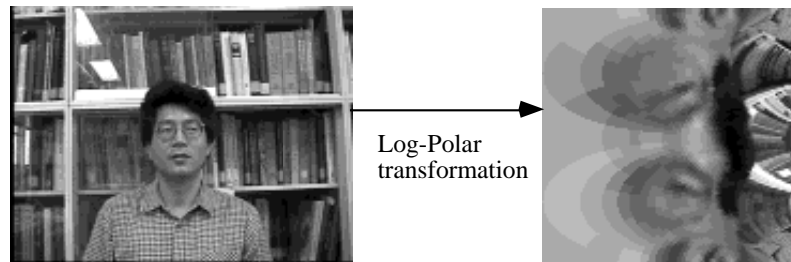


図3 Log-Polar写像による顔画像の変換

うな実世界で生じる問題を解消するため、遅延和方式に基づくビームフォーミングを小型のDSPモジュールによって構成し、高雑音環境下で音声認識率が向上することとDSPによる並列処理によって速度が向上することを確認した^{3,7,8)}。

6.4 確率ネットワークによる地図表現

移動ロボットは、常に自分がどこに置かれているかを

気にしている (localization problem)。ほとんどの移動ロボットは、ソナーと距離・回転計 (odometry) によって自己位置を定めようとするが、図4に示すようにソナーは多くのゴーストを生じ、odometryはわずかの回転誤差で歪曲された地図を形成する。一方、人間は「行き先を指定するのに「道なりに進んで二つ目の角を右に曲がって最初の入り口」というようなトポロジカルな方法を用いる。そこで、地図を場所の名前をノードとし、まっすぐ行

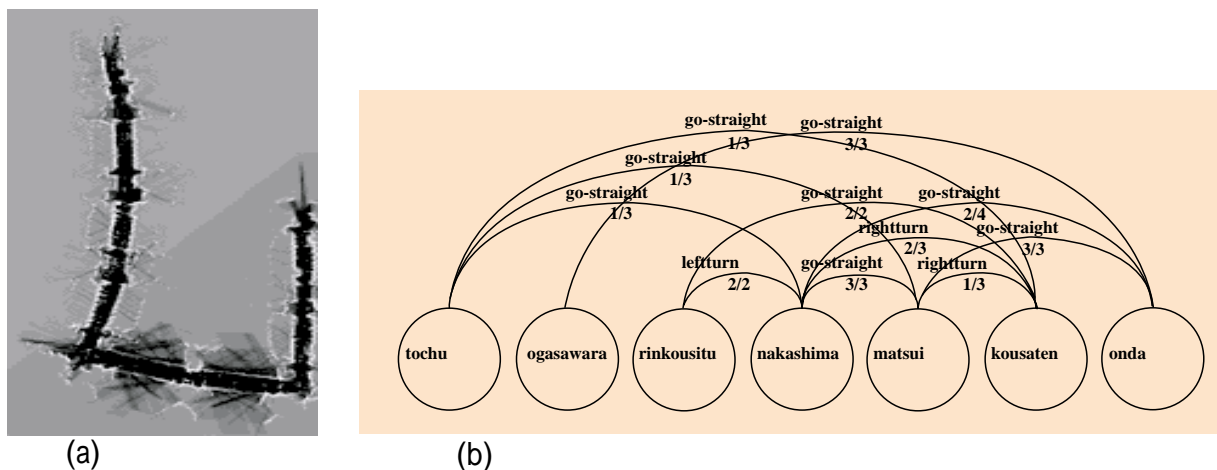


図4 地図の表現

(a)ソナーによって物体の存在する点を白く、存在しない点を黒くプロットした地図。ゴーストが見られる。またオドメトリの誤差によって直線廊下が曲がって認識される。
 (b)場所の名前をノードに、そこに到達する行動をアークにしたトポロジカル地図。

く、左に曲がる、という行動をアークとするネットワークで表現した。角や入り口の発見に失敗する可能性を補うために成功の確率を付加し、オフィス内の廊下を比較的口バストにナビゲートできる手法を開発した^{2,5)}。

これらを用いて、音声会話によってロボットを誘導し、場所を学習させる。実際は、ロボットはソナーによって開けた場所と狭い場所を認識し、前者はオフィスの入り口、後者は廊下の途中だと判断する。ロボットは、未知のオフィスの入り口で停止し、ここはどこかを訪ねるので、インストラクタが名前を発声して教える。何回かの繰り返しの後に場所と移動法が獲得されると、「～さんのところに行って下さい」という行動が実行できるようになる。

§ 7 まとめ

新情報プロジェクト、実世界知能分野、自律学習領域の研究について、実施体制とこれまでの代表的な研究成果について報告した。今後は、理論基盤、インターモーダル学習など、他領域との交流も深め、実世界情報の情報の統合的利用、実環境で使える情報技術の実現を目指して2001年度まで研究を継続する。

謝辞

新情報プロジェクトにおいて事情通ロボットの研究

の機会を与えて下さったことに対し、通商産業省機械情報産業局、電総研RWIセンター大津展之部長、RWCつくば研究所島田潤一所長に厚く感謝いたします。

参考文献

- 1) Matsui T., Asoh H., Hara I. and Otsu N., "An event driven architecture for controlling behaviors of the office conversant mobile robot, Jijo-2", Proc. of IEEE Int. Conf. on Robotics and Automation ICRA'97, Apr. 1997.
- 2) Asoh H., et al., "Socially embedded learning of the office-conversant mobile robot, Jijo-2", Proc. of Int. Joint Conf. on Artificial Intelligenc IJCAI'97, Nagoya, Aug. 1997.
- 3) Matsui T., Asoh H., and Asano F., "Map learning of the office conversant mobile robot, Jijo-2, by dialogue guided navigation", Proc. of IEEE/RSJ Symposium on Field and Service Robots, FSR'97, Canberra, Australia, Dec. 1997.
- 4) 松井俊浩, "おせっかいロボットとも呼ばれる事情通ロボットの計画", bit, Vol.29, No. 12, pp.4-11, 1997.
- 5) H. Asoh, S. Hayamizu, et. al., "Information integration of the office-conversant mobile robot, Jijo-2", Proc. of ICONIP'97, New Zealand, Nov. 1997.
- 6) Motomura Y., Hara I., Asoh H., and Matsui T., "Bayesian network that learns conditional probabilities by neural networks", Proc. of ICONIP'97, New Zealand, Nov. 1997.
- 7) Asano, F. and Hayamizu, S., "Speech ehancement using array

- processing based on the coherent-subspace method", IEICE Trans. on Fundamentals, Vol.E80-A, No.11, 1997.
- 8) 浅野 太, 速水悟, 松井俊浩, "話者方向同定と雑音抑制による音声認識性能の改善", 日本音響学会誌, 第53巻11号, 1997.
- 9) 栗田多喜夫, 堀田一弘, 三島健稔, "Log-Polar画像の高次局所自己相関特徴を用いた大きさに不変な顔画像の認識", 電子情報通信学会論文誌, Vol.J80-D-II, No.8, pp.2209-2217, 1997.
- 10) Kurita, T., et. al., "Scale and rotation invariant recognition method using higher-order local autocorrelation features of log-polar image," poster in the Third ACCV'98, Hong-Kong, Jan. 1998.
- 11) Horikawa K., Asoh H., Tani J., Matsui T. and Kakikura M., "Emergence of Expert Modules for Mobile Robot Navigation from a Mixture of Elman Networks", Proc. of Int. Conf. Neural Information Processing ICONIP-98, Kitakyushu, Sep. 1998.
- 12) 山川宏, 他, "学習のためのマルチエージェントシステムとその入札機構", 第6回マルチエージェントと協調計算ワークショップ (MACC'97), Vol. F7, No. 8, 1997.
- 13) Dev A., Kroese B.J.A., and Groen F.C.A., "Navigation of a mobile robot on the temporal development of the optic flow", Proc. of IROS'97.
- 14) 比屋根一雄, 澤部直太, "単発音のスペクトル構造とその擬音語表現に関する検討", 日本音響学会聴覚研究会, 1998.
- 15) Nilsson, M. and Ojala, J., "Self-awareness in Reinforcement Learning of Snake-like Robot Locomotion", Proc. IASTED 95, pp. 244-247, Mexico, June 1995.
- 16) 松本吉央, 稲葉雅幸, 井上博允, "視野画像列を利用した経路表現に基づくナビゲーション", 日本ロボット学会誌, vol. 15, no. 2, pp.74-80, 1997.
- 17) 稲邑哲也, 稲葉雅幸, 井上博允, "人間のロボット操縦時におけるセンサ時系列からの確率的状況認識", 第15回日本ロボット学会学術講演会, pp. 403-404, 1997.
- 18) Matsumoto Y., Ikeda K, Inaba M. and Inoue H., "Visual Navigation using Omnidirectional View Sequence", Proc. of 1999 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'99), pp.317-322, Kyongju, Korea, Oct. 17-21, 1999.
- 19) 小山剛弘, 松本吉夫, 今井正和, 小笠原司, "射影変換を用いた床面上の障害物検出", 第17回日本ロボット学会学術講演会予稿集, pp.121-122, Sep. 1999.
- 20) Matsumoto Y. and Zelinsky A., "Real-time Face Tracking System for Human-Robot Interaction", Proc. of 1999 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC'99), PP.II-830-II-835, Tokyo, Japan, October 12-15, 1999.
- 21) Sekimoto T., Tsubouchi T. and Yuta S., "A Simple Driving Device for a Vehicle -Implementation and Evaluation", Proc. Int. Conf. on Robotics and Systems (IROS' 97), 1997.
- 22) Ward, K. and Zelinsky, A., "An Exploratory Robot Controller which Adapts to Unknown Environments and Damaged Sensors", Proc. of Int. Conf. on Field and Service Robots (FSR'97), pp. 477-484, Canberra, December 1997.
- 23) Ward, K. and Zelinsky, A., "Learning Mobile Robot Behaviours by Discovering Associations between Input Vectors and Trajectory Velocities", Proc. of Australian Joint Artificial Intelligence Conf. pp. 138-143, Perth, December 1997.
- 24) Heinzmann, J. and Zelinsky A., "Robust Real-Time Face Tracking and Gesture Recognition", Proc. of Int. Joint Conf. on Artificial Intelligence IJCAI'97, pp. 1525-1530, August 1997.
- 25) 浅田稔, 野田彰一, 細田耕, "ロボットの行得獲得のための状態空間の自律的構成", 日本ロボット学会誌 Vol.15, No.6, pp.76-82, 1997.

(2000.5.12 受付)