

博士論文

視覚を有するエージェントのための
触覚に基づく内部表現獲得手法に関する研究

寺田 和憲

2001年 3月 23日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(工学) 授与の要件として提出した博士論文である。

論文番号： NAIST-IS-DT9761012

提出者： 寺田 和憲

審査委員： 小笠原 司 教授
横矢 直和 教授
木戸出 正継 教授
武田 英明 助教授
今井 正和 助教授

提出日： 2001年 3月 23日

視覚を有するエージェントのための 触覚に基づく内部表現獲得手法に関する研究*

寺田 和憲

内容梗概

本研究の目的は、環境の変化に柔軟に対応する知的エージェントの構築である。エージェントが環境を行動するためには、入力から出力へのマッピングを記述した内部表現を持つ必要があるが、この内部表現を設計者の知識に基づいて記述すると、想定された以外の状況には適応できないという問題がある。この問題を解決するために、エージェントは入力に対して自身の経験に用いて意味づけを行う必要がある。本論文では、経験に基づいた内部表現獲得の問題を課題として、入力システムとして視覚を有するエージェントが、身体性に基づいた汎用的な内部表現を獲得するための手法について考える。そのような内部表現獲得のために、環境との物理的相互作用を直接知覚可能な唯一のセンサである接触センサに注目し、接触センサに基づく内部表現獲得手法に関して、以下の三つの研究を行った。

まず、エージェントの内部表現を記述するための基本となる状態空間を自動的に分割する手法について研究した。この研究では、自己組織化マップアルゴリズムに基づき、視覚入力パターンをその類似性に基づいてオフライン競合学習を行い、状態空間を自動的に分割した。この手法により、状態空間を構成する入力次元を削減でき、より少ない状態数で行動学習できるようになった。

しかしながら、この研究では状態空間は入力ベクトルの類似性のみに基づいて構築されたため、より複雑なタスクを行うためにはタスクに応じて状態空間を構築する必要がある。そこで、二番目の研究として行動経験に基づいて状態空間を獲得する方法について研究を行い、接触信号に基づいてタスク遂行のために必要

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 博士論文, NAIST-IS-DT9761012, 2001年3月23日.

な視覚的特徴軸の抽出と特徴軸を分割する手法を提案した．実世界における行動では環境との接触が重要な意味を持つので，エージェントのタスクは接触信号によって意味付けられると考えられる．本手法では，この考えに従って，行為の終端で知覚される接触信号を行為系列に対して割り引いて与え，その割り引き信号を統計的に解析することで，タスク遂行のために必要な視覚的特徴軸の抽出と特徴軸の分割を行った．実験の結果あらかじめ準備した複数の特徴軸の中から目標到達行動のために必要な特徴軸が選択された．また，細かく分割されていた状態を統合し，状態数を削減することができた．得られた状態空間を用いて強化学習を行ったところ，エージェントは対象に到達する行動を達成することができた．

三番目の研究として効用関数を用いた対象認識と行動生成の研究を行い，エージェントの身体性を考慮した対象の内部表現の獲得手法を提案した．二番目の研究において，タスク遂行という条件に基づいて状態空間の生成を行えることを示したが、この状態空間はエージェントのもつ物理的特徴・能力を暗に組み込んだものである．そこで、エージェントのもつ物理的特徴・能力を明示的にとりこむために、視覚入力に対して接触信号に基づいて効用関数を生成することで，エージェントの身体性を考慮した対象の内部表現を獲得する方法を提案した．この方法では，視覚入力中に見える対象は効用関数によって表現される．効用関数は身体表面と対象表面の相対関係をその間で予測される行動によって表現するものである．実験の結果，エージェントは視覚入力から効用関数を生成し，仮想的ななぞり行動によって6種類の対象の識別を行うことができた．また，通り抜けられるすきまと通り抜けられないすきまを判別し適切な行動を行うことができた．

これらの研究を通じて，内部表現構築のための手がかりとして接触センサを用いると，従来設計者が埋め込んでいた知識を可能な限り排除した自律的なエージェントが構築可能であることがわかった．

キーワード

自律エージェント，内部表現，視覚，触覚，身体性，対象認識，効用関数

Development of Methods for Acquiring Internal Representation for Vision based Agents by Tactile Information*

Kazunori Terada

Abstract

The goal of this research is to develop autonomous mobile agents, which can tolerate change of environment. Agents that act according to the surrounded environment should have some internal representations of the environment. In traditional robotics research, such representations are usually given a priori, e.g., generation of three-dimensional model by interpretation of images. But it is not realization of autonomous agents because it means that their cognitive abilities are carefully designed by human designers according to the environment and their physical capabilities. In order to realize autonomous agents, they should build their internal model of the environment by themselves not by human designers. There exist two problems to acquire internal representation, especially mapping between visual input and action; 1) how to construct internal state space automatically and 2) how to consider the embodiment of agent. In this paper, in order to solve these problems, we show three types of methods of acquiring internal representation based on tactile sensors, which can only perceive physical interaction between agent's body and object directly.

Firstly, we show a method to construct a state space with self-organization map. The state space is generated by classifying input patterns according to their

*Doctor's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DT9761012, March 23, 2001.

similarity that is performed as off-line competitive learning with self-organization map. An experimental result shows that the agent can learn appropriate behavior with the generated state space of which number is reduced in comparison with the number of the original state space.

In the second study, the state space is generated by experiences in achieving the tasks. In the first study, the state space is generated only by similarity of input pattern. But it is not sufficient for agents to perform more complicated tasks, i.e., the state space should be generated according to the tasks. Since tasks in this study is to reach some objects, achievement of tasks is perceived as the tactile stimulus received at the end of action series. We give a reward for an action in the action series by discounting the end of the action series. By analyzing these tactile stimulus statistically, we could discover visual feature and could discriminate state spaces. An experimental result shows that one visual feature is selected from some prepared features, and the size of state space is reduced by merge process. The agent can learn behavior of reaching with this state space.

In the third study, we represent internal representation depending on embodiment as a utility function based on tactile sensor. Utility function represents objects in visual input, and the physical relation between agent's body and object by estimating the behavior between them. By integrating these utility functions according to tasks, the agent can discriminate various objects and determine whether a gap between two objects is to pass or not.

Through these studies, we can conclude that using tactile sensors as cue of construction of internal state space, we can realize complete autonomous mobile agents with less a priori knowledge on their tasks, bodies and environment.

Keywords:

autonomous agent, internal representation, vision, tactile sensor, embodiment, object recognition, utility function

目次

1	はじめに	3
2	内部表現 -視覚から行動へのマッピング-	9
2.1.	知的エージェント	10
2.2.	マッピングの表現方法	11
2.2.1	記号に基づくシステム	11
2.2.2	サブサンクションアーキテクチャ	13
2.3.	効用関数に基づくアプローチ	15
2.3.1	効用関数	15
2.4.	相互作用と身体性	17
2.4.1	アフォーダンス	17
2.5.	身体性	20
2.6.	視覚と身体性の関係	21
2.6.1	モリヌークス問題	21
2.6.2	開眼手術者に対する認知心理実験	22
2.6.3	Held の動物実験	23
2.7.	対象記述のための内部表現	24
2.8.	本研究におけるアプローチ	25
3	ロボット	27
3.1.	タイプ I のロボット	27
3.2.	タイプ II のロボット	29
3.2.1	システムの概要	31
3.2.2	駆動システム	31

3.2.3	視覚センシングシステム	32
3.2.4	接触センサシステム	33
3.2.5	モーターコントロールシステム	34
3.3.	タイプ III のロボット	34
4	自己組織化マップを用いた状態空間の構築	37
4.1.	視覚から行動へのマッピングの学習	37
4.2.	Visual State Action Map	39
4.2.1	自己組織化マップを用いた画像のカテゴリー化	41
4.2.2	入力画像の前処理	42
4.2.3	状態遷移確率モデルの獲得	43
4.2.4	学習	44
4.3.	実機を使用した実験	44
4.3.1	VSAM の構築	45
4.4.	まとめと考察	48
5	接触信号に基づく視覚的特徴の選択と状態分割	51
5.1.	視覚と触覚の関係	51
5.2.	状態空間の分割問題	54
5.3.	提案手法	55
5.3.1	特徴軸の選択	55
5.3.2	軸の分割	56
5.4.	実験結果	57
5.4.1	特徴軸の選択	58
5.4.2	特徴軸の分割	60
5.4.3	行動の学習	60
5.5.	まとめ	60
6	接触に基づく効用関数を用いた行動生成と対象表現	65
6.1.	視覚入力と効用関数	66
6.2.	提案手法の概要	66
6.2.1	遷移モデルを用いた効用値の更新	68

6.2.2	報酬の与え方	69
6.2.3	視覚入力の扱い方	70
6.3.	効用関数の生成方法	70
6.3.1	前提とするエージェントと環境	71
6.3.2	効用関数の計算	71
6.3.3	効用関数の重ね合わせ	74
6.4.	身体性を考慮した行動生成	75
6.4.1	身体性を表現するための効用関数	76
6.4.2	タスク遂行のための状態判断	76
6.4.3	接触状態への到達	77
6.4.4	行動生成	77
6.5.	障害物回避	77
6.5.1	状態遷移確率モデルの獲得	78
6.5.2	グローバルビジョンによる通り抜けタスク	78
6.5.3	ローカルビジョンによる障害物回避	82
6.6.	身体性の変化と行動出力に関する実験	84
6.7.	実機による実験	88
6.8.	対象の認識と判別	91
6.8.1	仮想的なぞり行動の生成	92
6.8.2	チェイン符号化	93
6.8.3	大きさ, 形状特徴の抽出	94
6.8.4	対象認識と判別の実験	95
6.8.5	形状特徴ベクトルの抽出	95
6.8.6	対象の判別	96
6.9.	まとめ	97
7	おわりに	101
7.1.	本研究のまとめ	101
7.2.	それぞれの手法に対する考察	104
7.2.1	画像処理における設計者の知識	104
7.2.2	内部表現の汎用性	105

目次

7.2.3 接触信号の役割	106
7.3. 結論	106
7.4. 今後の展望	107
謝辞	109
参考文献	111
研究業績	117

目 次

2.1	Probabilistic environment	15
2.2	An architecture of acquiring a mapping between vision and action based tactile information	25
3.1	Type-I robot	28
3.2	System architecture of type-I robot	28
3.3	Type-II robot	30
3.4	Software architecture	31
3.5	Driving system	32
3.6	Tactile sensors by utilizing a key board	34
3.7	Motor drive board	35
3.8	Type-III robot	36
4.1	Architecture	40
4.2	Two line figures that are orthogonal	42
4.3	Transition probability model	43
4.4	Configuration of system	44
4.5	Generated map	45
4.6	Learned Q value: When agent is going along the wall, Q value of the action toward wall is low and Q value of the action leaving from the wall is high.	47
4.7	The action avoiding the wall. 1,4,6: The agent turns the corner.; 2,3,5: The agent goes along the wall.	47

5.1	Cognitive Robot Architecture. A circle means a information and a rectangle means the processing of the information.	52
5.2	Simulator field and Captured image	58
5.3	Robot's action and distribution of tactile sensors	59
5.4	discounted tactile stimulus class histogram for feature 3	61
5.5	Behavior of reaching ball	61
6.1	Overview of our method	67
6.2	Assumed environment and agent	71
6.3	Input image	72
6.4	Expected touch area, receptor area, and reaching action path . . .	72
6.5	An example of input image	74
6.6	An example of utility function	75
6.7	Camera model used in our experiment	78
6.8	Input images used in our experiment	79
6.9	Utility function for reaching a goal	79
6.10	Utility function of narrow case	80
6.11	Utility function of wide case	80
6.12	A generated path of narrow case	81
6.13	A generated path of wide case	81
6.14	Input image of local vision	83
6.15	Utility function of local vision	83
6.16	A generated path of avoidance	84
6.17	Agents used in this experiment	85
6.18	Generated paths of agent A	85
6.19	Generated paths of agent B	86
6.20	Generated paths of agent C	86
6.21	A generated path in complex environment	87
6.22	The robot and environment used in real world experiment	88
6.23	The experiment environment	89
6.24	The input image and processed image	89

6.25 Comparison of generated paths of the real robot and the simulated robot without obstacle	90
6.26 Comparison of generated paths of the real robot and the simulated robot with obstacle	90
6.27 Trace action and chain coding	93
6.28 Characterizing of a size and shape of an object	94
6.29 Shape characterizing value of f3-f4	96

表目次

4.1	Example of VSAM	46
5.1	Relations among object, color and physical property	57
5.2	The correlation coefficient rate	59
5.3	The variance of \mathcal{F}	60
5.4	ChiMerge discretizations for feature 3	62
6.1	Chain codes corresponding to actions	93
6.2	Objects used for learning	95
6.3	Average of shape characterizing value	95
6.4	Object recognition rate	96
7.1	Image processing methods applied each study	104

第 1 章

はじめに

本研究で対象とするのは、環境で自律的に振舞う知的エージェントである。エージェントはセンサによって環境を知覚し、環境に対してはアクチュエータを介して動作を出力する。エージェントがタスクを遂行するためには各場面において入力状態を判断し適切な出力を行わなければならない。そのために、エージェントは入力から出力の関係を記述した表現 (マッピングと呼ぶ) を持っている必要がある。旧来の人工知能ではマッピング記述方法として、設計者の知識を記号によって表現し、記号操作によって入出力関係を定義するというように、ほとんど設計者の知識に基づいて記述されていた。しかし、このような方法では、設計者の想定した以外の状況に適応出来ないという問題がある。

このような問題に対して Brooks[8] が環境との相互作用を重視するシステムを提案した。しかし、環境との相互作用を重視するという着眼点は評価されるものの、このシステムでは、旧来の知識の記述方法をモジュールごとに独立に書き換えただけのものであり、依然として想定外の状況に適応できなかった。そこで、次に環境との相互作用から内部表現を学習するという強化学習によってマッピングを獲得する手法が注目され、多くの研究が行われるようになった。このような人工知能の歴史は、多様な環境に適応するためにできるだけ自律性の高いシステムを作ろうという試みによって積み重ねられてきたと言える。

エージェントの自律性は設計者による知識の組み込み度合に関係している。エージェントは、エージェント自身の経験によって得られた知識と、それが動く環境に合わせて設計者によって組み込まれた知識の両方に基づいて行動するが、エージェントが自律的であるためには、設計者によって組み込まれる知識は少ない方がよい。またエージェントの自律性は、獲得された内部表現の汎用性・再利用性

にも関係する．エージェントが自律的であるためには，特定のタスクに特化した内部表現を持つのではなく，あらゆる異なった状況にも対応可能な内部表現を持つべきである．本論文では，自律エージェント，特に視覚を有する自律エージェントにおいて，設計者の埋め込み知識をいかに少なくして汎用的なマッピングを獲得するかに焦点を当てて研究を行う．そのために，視覚的特徴として何を入力として用い，どんな方法でマッピングを獲得するかが問題となる．

エージェントが視覚を入力として用いると，環境の様々な情報を採り入れることが出来るため，有効であるが，一方でその膨大な情報量をいかに取り扱うかが問題となる．多くの場合，こういった情報を入力として用いるかは設計者によって決定される．設計者の埋め込み知識を少なくすることを目標としている強化学習においても，膨大な情報を含む視覚入力を取り扱うためには，様々な工夫が必要であり，多くの手法が提案されている．しかし，提案されている多くの方法では，一般的に強化学習を用いるエージェントが遭遇する状態空間の構築時の問題点と，視覚を用いることにおける固有の問題点を同じ問題として考えているために，エージェントが視覚情報をいかに用いるかという問題を本質的に解決するものではない．そこに足りない議論は，視覚という固有のセンサモダリティがエージェントに対して本質的にどのような情報を提供し得るかということである．そこで視覚のあり方について原点にもどって考え直す必要がある．

それは，15世紀の哲学者 Berkeley によって提案された仮説を受け入れるところから始まる．その仮説とは，視覚情報には本質的には環境内の物理的情報が全く含まれないというものであり，この極端な仮説は，視覚を用いて環境の全てを理解していると考えられるわれわれ人間にとって俄に理解しがたいものである．この仮説は，モリヌークス問題として知られる，アイルランドの哲学者モリヌークスの提案した次のような問題に対する解答として提案されたものである．その問題とは「ある生来の盲人が，今は成人して，同じ金属からできたほとんど同じ大きさの立方体と球とを触覚によって識別するように教えられ，その立方体と球とがテーブルの上におかれ，その盲人が見えるようになった，と想像してほしい．疑問：彼は両者に触れないうちに，視覚によってどちらが球でどちらが立方体であるかを識別し言うことができるであろうか」というものである．この問題に対して，モリヌークス自身とバークレーは両方とも「見えるようになった」盲人は視

覚によって二つの対象を見分けることができないと結論づけた．モリヌークスが問題提起した時代には実際にこの問題の真偽をはっきりさせることはできなかったが，彼らの結論は後に開眼手術を受けた先天盲人に対する心理実験によって，正しいことが確かめられている．その結果として言えることは，次の事柄である．視覚情報を用いてわれわれが様々なものを理解出来るのは，身体を用いて実際に環境を行動した結果と視覚入力中での見え方の関係を経験的に理解しているからである．すなわち，視覚情報に本質的に環境内の物理的情報が含まれるのではなく，そこに存在するのは，身体を用いた行動経験へのラベルとしての意味だけである．これは，視覚入力から行動出力へのマッピング獲得の際に，身体と物理的環境との直接的相互作用が重要な役割を果たすということである．

そこで，本研究では自律エージェント構築のために，物理的相互作用を直接知覚可能な唯一のセンサである接触センサに焦点を当て，接触センサに基づいて，視覚入力から行動出力へのマッピングを獲得することを旨とする．本論文では，マッピング獲得に関する3つの研究を示し，その中で，どのような視覚的特徴を用いれば設計者の埋め込み知識を少なくし汎用的な表現を獲得できるかについて議論する．マッピングの表現形式としては，効用関数を用いるものとする．

まず視覚入力として，自己組織化マップによって状態数を圧縮し，状態空間を構成した．この方法では，教師無し分類学習を用いているので，少ない埋め込み知識によって入力とする視覚的特徴を決定することができる．しかし，入力ベクトルの類似性に基づいてクラス分類しているため，タスク遂行のために注目しなければならぬ微小な特徴量の変化も，一様に無視してしまうという問題がある．この問題を解決するために次に2番目の研究として，接触信号に基づいてタスク遂行に必要な視覚的特徴の選択と特徴軸の分割を行った．この手法は，Berkeleyの主張する視覚形成における触運動の重要性を考慮するものである．しかし，この手法では，あらかじめ用意したいいくつかの視覚的特徴を接触信号に基づいて選択するという方法をとるため，用意する特徴軸の選定に設計者の知識を要する．また，複雑な環境では，用意する特徴軸が爆発的に増加することが予想される．ここでの問題点は，考え得る視覚的特徴を多く用意してその中からタスクに必要なものを選択するという，予備的処理が重要な役割を果たすということである．そこで，タスク遂行のために本質的に必要な視覚的特徴について考察を行った．

その結果, Berkeley の仮説にもあるように, 視覚の役割は接触を予測することであり, したがって, 実世界で行動するエージェントにとって接触状態に関係する対象の物理的境界が本質的に必要とされる視覚的特徴であるという結論を得た. そこで, 視覚入力から抽出する特徴として対象の物理的境界を用いて, その特徴に対して接触信号をもとに行動を意味づけすることを考えた. この方法によって, 様々な対象に関して任意の行動を生成可能な汎用的なマッピングを獲得することができた.

これらの研究によって得られる結論は, 視覚から行動へのマッピングを獲得する際に接触信号を用いると, 接触信号を介したフィードバックループを形成することができ, 設計者の知識をより少なくできるということである. また, 視覚の本質は接触を予測することであり, タスク遂行のためには対象の境界に相当する視覚的特徴が重要な役割を果たすということである.

本論文では, まず 2章で, 人工知能の歴史の中で入力から出力へのマッピングがどのように記述されてきたかを振り返り近年注目されている効用関数がその表現方法として有効であることを説明する. また, 本研究で一貫して主張する接触信号の重要性を, 認知心理学実験を例に挙げて説明する.

3章では, 本研究で仮定した 3 種類のロボットについて設計方針と仕様を説明する.

4章では, 最初の研究として, エージェントの内部表現を記述するための基本となる状態空間を自動的に分割する手法について説明する. ここでは状態数の削減のために, 入力画像ベクトルの類似性にもとづいてオフライン競合学習を行う自己組織化マップアルゴリズムを用いた. この手法により, 少ない埋め込み知識によって状態数を削減することができ, 構成された状態空間を用いてロボットは障害物回避の行動を獲得できた. しかし, この手法では, 入力ベクトルをユークリッド距離を用いて比較するため距離に応じて状態が分割され, 細かい状態の変化は一様に無視される. しかし, 特定のタスク遂行のためは, エージェントは特定の視覚的特徴の細かい変化を見なければならない.

そこで, 5章では, タスク遂行のため必要な視覚的特徴の選択と選択された特徴軸の分割の問題を取り扱った. ここでは, 軸の選択と分割において埋め込み知識を用いないために, 環境との物理的相互作用を直接知覚可能な接触センサに注

目し、接触信号に基づいた軸の選択と分割の手法を提案した。これは実世界におけるタスクを次のように捉えることで導かれる発想である。すなわち、実世界における行動はほとんどの場合接触を伴うので、エージェントのタスクは接触信号によって意味付けられると考えられる。本手法では、この考えに従って、行為の終端で知覚される接触信号を行為系列に対して割り引いて与え、その割り引き信号を統計的に解析することで、タスク遂行のために必要な視覚的特徴軸の抽出と特徴軸の分割を行った。実験の結果あらかじめ準備した複数の特徴軸の中から目標到達行動のために必要な特徴軸が選択された。また、細かく分割されていた状態を統合し、状態数を削減することができた。得られた状態空間を用いて強化学習を行ったところ、エージェントは対象に到達する行動を達成することができた。

エージェントが環境との相互作用によってマッピングを獲得するときに、身体性の問題は避けて通れない問題となる。二番目の研究において、タスク遂行という条件に基づいて状態空間の生成を行えることを示したが、この状態空間はエージェントのもつ物理的特徴や運動能力能力を暗に組み込んだものである。そこで、6章では、マッピング獲得時にエージェントの物理的特徴や運動能力、すなわち身体性を明示的に取り込む方法について議論を行行い、効用関数を用いた対象認識と行動生成の研究を行行い、エージェントの身体性を考慮した対象の内部表現の獲得手法を提案した。ここでは、対象認識に関して、次の仮説を用いた。すなわち、従来対象の内部表現と行為の内部表現は別のものであるとして考えられてきたが、対象認識の問題はその対象についての行動生成の問題である。この仮説に基づき、エージェントは視覚入力から見えている対象について可能な全ての行動を予測し、それらの行動系列の特徴によって対象を判別する。行動系列を予測するために、エージェントは視覚入力を微小領域に分割し、それぞれの微小領域中に観測される対象の境界について効用値の計算を行う。対象が効用関数によって記述されていると、エージェントは対象に関するあらゆる行動を生成することが可能になる。また、効用関数は身体表面と対象表面の相対関係をその間で予測される行動によって表現するものである。実験の結果、エージェントは視覚入力から効用関数を生成し、仮想的ななぞり行動によって6種類の対象の識別を行うことができた。また、通り抜けられるすきまと通り抜けられないすきまを判別し適切な行動を行うことができた。

最後に7章では本研究のまとめを行い、それぞれの提案手法が設計者の埋め込みの少ない汎用的な表現になっているかについて考察を行い、本論文を結論づける。また、本研究が今後とるべき方向性について展望を述べる。

第2章

内部表現 - 視覚から行動へのマッピング

エージェントはセンサによって環境を知覚し，環境に対してはアクチュエータを介して動作を出力する．知的なエージェントは，タスクを滞りなく遂行するために，直面する各場面において適切な判断を下す必要がある．エージェントが遭遇する各場面を定義するのは，それまでの知覚の履歴，すなわち知覚列である．エージェントが各場面において適切に振舞うためには，知覚列に応じて動作を対応づける表，すなわち知覚列から行動へのマッピング持っていればよい．原理的には，考え得る全ての知覚列に対する行動へのマッピングを持っていれば，エージェントは合理的に振舞うことが出来るが，現実的にはそのような表を作ることは不可能である．そのため，このマッピングはなんらかの方法で記述される必要がある．本章では，このマッピングを記述するための手法について，まず，知的エージェントの定義を行い，自律エージェントにおけるマッピングの表現手法についてこれまでの人工知能の分野で用いられてきた手法をいくつか紹介する．次に，環境との相互作用からマッピングを獲得する手法として強化学習の説明を行う．また，エージェントと環境との相互作用を説明する上で重要な概念であるアフォーダンスと身体性について述べ，視覚と身体性の関係において認知心理実験を引用して，視覚形成に身体運動が欠かせないことを説明する．最後に，これらの事項に基づいて本論文で提案する三つの研究の方向性を述べる．

2.1. 知的エージェント

知的なエージェントは環境で自律的に振舞う。エージェントの自律性は設計者による知識の組み込み度合に関係している。エージェントは、それが動く環境に合わせて設計者によって組み込まれた知識と、エージェント自身の経験によって得られた知識の両方に基づいて行動するが、エージェントが自律的であるためには、設計者によって組み込まれる知識は少ない方がよい。組み込み知識のレベルはアーキテクチャのレベル、アルゴリズムのレベル、プログラミングのレベルなど様々なレベルで存在するが、エージェントが人間によって作られるシステムである以上、設計者の知識を全く組み込むことなく動くエージェントは存在しない。すなわち、何も設計されなければエージェントは一步も動くことができず、エージェントの知識となる経験の獲得もできない。しかし一方で、設計者の知識に依存しすぎるエージェントも、組み込まれた行動規則以外の状況に直面すると一步も動くことはできない。そこで、ある程度の自由度を持って自律的に行動できる妥当なレベルで知識を埋め込む必要がある。

ここで、エージェント設計における設計者の組み込み知識としてどのようなものがあるか考える。エージェントに組み込む知識とは、エージェントの設計において前提とするエージェントの能力のことであり、その能力に応じて可能なタスクのレベルが異なる。例えば、オフィスにおける人間の活動に関する知識を持っていることを前提としたエージェントとして、事情通口ロボットがある [46]。このロボットは、オフィスに訪ねて来た人を目的の人の所に案内することができる。そのために、ロボットは普段からオフィスを巡回し、誰がどこに居るかという情報を蓄積する。このような高度なタスクを行うためには、様々な組み込み知識が必要となる。例えば、音声認識では音声データからテキストデータへの変換を行うが、どの音がどのテキストに対応しているかをあらかじめ知っている必要がある。また、目的地に到達するために、地図データと人物や交差点、戸口などのランドマークとの対応を知っている必要がある。事情通口ロボットでは、このような知識を学習によって獲得しているが、アーキテクチャそのものは設計者によって埋め込まれたものである。このように、抽象度の高いタスクすなわち人間の日常生活に近いタスクを行うためには、設計者が様々なレベルの知識を組み込む必要がある。

また、自律性の高いエージェントは環境への適応性も高い。自律性の高いエージェント、すなわち組み込み知識の少ないエージェントはエージェント自身が経験によって獲得した知識に基づいて行動する。すなわち、エージェントの知識はエージェントと環境の相互作用の経験に基づいているため、環境の構造を反映したものとなっており、その結果エージェントは環境に適応しているといえる。

また、知的エージェントを考える上でエージェントの持つマッピングの汎用性・再利用性の議論をする必要がある。マッピングが設計者によって記述されたものであっても経験によって獲得されたものであっても、その内部表現が特定の環境やタスクに対してしか適用できないと、異なった環境やタスクに対して新たに内部表現を獲得する必要がある。取得のためのコストが必要となる。したがって、内部表現は様々な環境やタスクに対して汎用的に利用でき、再利用可能であることが望ましい。

2.2. マッピングの表現方法

本節では、人工知能の歴史の中でマッピングの表現方法がどのように取り扱われてきたかについて説明する。

2.2.1 記号に基づくシステム

人工知能の歴史の中で、記号に基づくシステムがマッピングを表現する方法として長く使われてきた。このようなシステムの例としては、マッピングを反射的な条件-動作 (*if-then* ルール、プロダクションルールと呼ばれる) の形で記述したルールを入力状態に応じて探索するシステムや、ゴールを達成するために、記述された行為と状態を用いてプランニングを行うシステムなどがある。

スタンフォード大学で作られた STRIPS[12] は建物内を移動するロボットの行動計画を作成するプログラムである。STRIPS では、状態と動作に関する知識が、オペレータによって記述され、オペレータによって記述されたりリストの追加と削除を行うことで、プランニングを行う。

記号接地問題

このような記号に基づくシステムの限界を指摘し批判する問題として Harnad は記号接地問題 (Symbol Grounding Problem) を提起した [16]。シンボルグラウンディング問題は、システム内で記述される記号によるモデルをシステムの外の現実世界に関係させる必要があるということを指摘している。記号システムにおいては、記号間の関係をどのように記述するかについて、また記号をどのように処理するかを問題としているため、記号が実世界とどのように関係するかについて議論されることはない。システム内の記号が実世界と関係がなくて良い場合は、システムの記号処理結果を人間が解釈する場合である。WWW の検索システムは、インターネット上にある膨大なテキストに記述されている内容を現実世界の事象に関連させて理解しているわけではないが、ユーザの要求に対して適切な検索結果を出力することができる。出力結果は、人間に解釈されることによって実世界に接地するのである。しかしながら、ひとたび出力結果を解釈する人間が取り除かれると出力結果は全く意味を持たないものとなる。もし記号がシステムの中にあれば、それらの意味はシステム自身が実世界とインタラクションを行なって得た経験に基づいていなければならない。記号が他の記号を参照しているようなモデルは、システムと実世界との関連がないので、実世界に接地しているとは言えない。記号はデザイナーやユーザのための意味を持つだけで、システム自身が持つべき意味を持たないのである。

フレーム問題

また、記号に基づくシステムを批判する問題としてフレーム問題が知られる、フレーム問題は、どの条件を考慮に入れ、どの条件を考慮に入れないか、という問題である。この問題は McCarthy と Hayes によって最初に指摘された [32]。フレーム問題の主要となるポイントは、モデルがどのように変化するかということに関係している [20]。連続的に変化する環境のモデルが与えられたとすると、モデルが実世界と同調を保つのは困難である。いま、モデルが論理的命題のセットで書かれているとすると、どの命題もあらゆる時点において変化することができる。しかしながら物理世界は本質的に物理法則に拘束されている。オブジェクトが消失したり、理由なしに飛びはじめたりしない。しかし、氷は温度が上昇すれ

ば融けて消失する。問題は、起こり得る可能性が潜在的に非常に多く存在することである。Denett[10] の例で考えてみる。

ロボット R1 は、部屋の中にバッテリーと爆弾があって、爆弾が爆発する前にバッテリーを部屋の外に出さなければならないと教えられている。バッテリーと爆弾は同じ台車の上に乗っている。R1 は台車を部屋の外に出せばバッテリーも部屋の外に出せることを知っている。R1 は実際にそうして台車を部屋の外に出したが、爆弾は爆発した。R1 は台車を部屋の外に出すことによって、爆弾も部屋の外に出ることが分からなかったのである。設計者は、行動に関する含意だけでなく、プランを実行する上で派生的に現れる含意も考慮に入れる必要があることに気づき、次のロボット R1D1 を設計した。R1D1 は、自分の起こす行動によって他に与えられる影響を推論することができる。ロボット R1D1 は台車を外に出すことによって発生する含意を推論しはじめた。しかし、台車を外に出すことによって部屋の壁の色が変わらないことを推論したところで、爆弾は爆発してしまった。問題点はもう明らかである。ロボットはどの含意が関係していて、どれが関係していないかを教えられる必要がある。そこでロボット R2D1 が設計された。しかし、R2D1 は、部屋の中に爆弾を放置したまま止まったままだった。そこでデザイナーが「何かしろ」と R2D1 に言うと、R2D1 は次のように答えた。「私は今、何千もの関係ない含意を無視することで忙しい。今も、関係のない含意を見つけたところです。これを無視する方のリストに入れなければ...」爆弾は爆発してしまった。

2.2.2 サブサンクションアーキテクチャ

伝統的な記号に基づくシステムが抱える記号接地問題とフレーム問題は、システム内のモデルが人間によって記号を用いて記述されることによって生じる。これらの問題を解決するために Brooks はサブサンクションアーキテクチャを提案した [7][8][18]。サブサンクションアーキテクチャでは、旧来の AI で用いられて来たような、知覚 - モデル - プランニング - 行動出力というモデルを中心としたシステムではなく、モデル - プランニングの部分をも、振舞の抽象度にしたがって分割し、階層化されたモジュールに置き換えた。それぞれのモジュールは独立に動き、階層の上位のモジュールは下位のモジュールに対して抑制することができ

る．例えば，最下層から順に障害物回避，さまよう，探索などのモジュールが存在し，探索モジュールはさまようモジュールに対して，コントロールできる．このシステムは次のような特徴を持つ

1. 知覚と行動が直接結び付いている．このシステムでは、各モジュールがセンサや他のモジュールからの入力に対して反射的に決められた行動を出力するので、全体ではセンサ入力から直接行動を出力しているように見える．
2. 抽象化された内部表現を持たない．システムを構成してるモジュールが，センサ入力に対する行動を決定するので、従来のシステムのように記号によって環境を記述する必要がない．
3. 環境との相互作用によって挙動が決まる．センサ入力に対して直接行動が決定され，行動を実行した結果新たなセンサ入力を得る．このサイクルの繰り返すことにより，行動は環境を変化させ，変化した環境は行動をエージェントにフィードバックする．

サブサンクションアーキテクチャは、抽象化された内部表現を持たず、環境との相互作用によって行動を決定することで、前述した2つの問題を回避した．このシステムは、行動によって変化した環境からフィードバックを受けながら行動を行なうので、行動は環境に影響され発現的である．環境に対してタスクを細かく分割し入出力関係をモジュール化したことによって、複雑な環境に対してロバストに動くことができる．しかし、各モジュールが独立にゴールをもって機能しているため、全体のシステムとしての明示的なゴールを扱うことができない．また、入出力関係が固定的行動要素から構成されるため、既知の行動要素に状況が分解できないと適応性や有意味性が崩れてしまう．

このアーキテクチャでは、モデルは明示的な記号によって表現されていないが、設計者がモジュールという形に分解してモデル化を行っているため、各モジュールの構成の段階で設計者の知識は埋め込まれている．

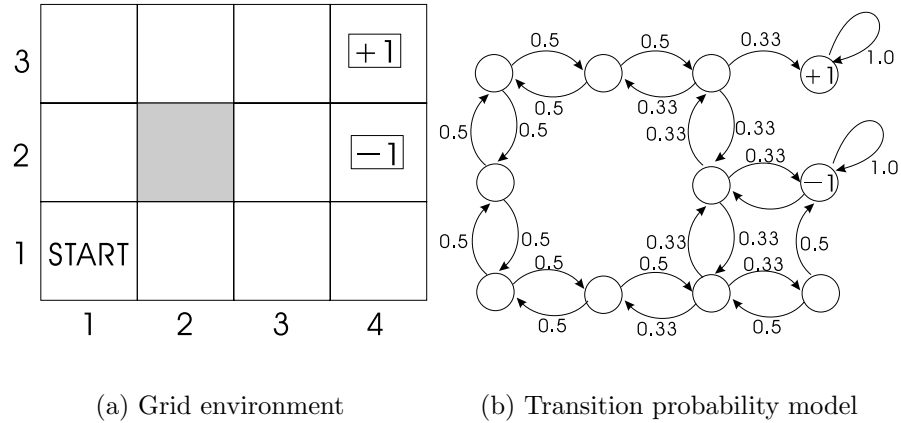


図 2.1 Probabilistic environment

2.3. 効用関数に基づくアプローチ

先に述べたように、エージェントの持つマッピング獲得の際に設計者の手による記述は少ないほど良い。Brooks は、マッピングを記号として記述する伝統的な手法を批判し、表象を持たないシステムを目指した。しかし、結果として出来上がったものは、環境を直接モデル化したものではないものの、機能レベルに分割された振舞いとして環境をモデル化したものであり、設計者の組み込み知識を必要とするものであった。組み込み知識を少なくするためには、エージェントの経験を有効に利用する必要がある。より環境との相互作用を重視し、経験によってマッピングを獲得するための方法として強化学習がある [21]。強化学習では、目的とされるタスクは報酬としてエージェントに与えられる。エージェントはその報酬を手がかりに、環境との相互作用の中から適切な行動を学習する。強化学習で学習するのは、環境の構造そのものではなく、タスクを遂行するための最適政策である。政策は、入力に対してどのようなコマンドを出力するかを規定するためのルールである。エージェントは試行錯誤的に環境との相互作用の中から最適政策を学習するので、環境に対して適切な行動を学習することが出来る。

2.3.1 効用関数

環境との相互作用を重視するエージェントの研究において、学習アルゴリズムとして強化学習がしばしば用いられる。強化学習とは、不確実な環境において意志決定を行うための効用関数を探索的に求める手法である。ここでは、探索過程を含まずに効用関数を求める動的計画法について説明する。

決定的な環境における意志決定は、初期状態と動作の効果がわかっている場合、探索アルゴリズムによって解くことができる。非決定的な環境においては、個々の動作の結果は確率的になる。例えば、図 2.1 のような非決定的環境においては、(1, 1) から上に行くという動作は確率 0.8 でエージェントを (1, 2) に運ぶが、確率 0.2 で (2, 1) に移動させる。与えられた動作によって発生する状態間の可能な遷移に付随した確率の集合を状態遷移モデル (transition probability model) と呼ぶ。エージェントがセンサを通じて環境の状態を知覚できるなら、環境はエージェントにとってアクセス可能であると言う。アクセス可能で不確実な環境において、状態遷移モデルが既知のときに最適方策を計算する問題はマルコフ意思決定問題 (Markov Decision Problem: *MDP*) と呼ばれる。マルコフ決定問題では、マルコフ性が成立する環境における意思決定問題を扱う。ある状態における状態遷移確率が、それまでの履歴に依存せず現在の状態にのみ依存する場合、マルコフ性が成立していると言う。

値反復法

マルコフ決定問題において、最適政策を計算するアルゴリズムについて述べる。確率的環境においては、エージェントは動的計画法を用いて効用値の計算を行い、最大期待効用原理に基づいて最適な行動をとることができる。状態から動作へのマッピングは政策 (policy) と呼ばれる。最大期待効用原理とは次のような原理である。マルコフ決定過程において効用関数 U が与えられると、最適な動作はその結果の状態の期待効用が最大となるものとして決められ、状態 s におけるエージェントの政策 f は次のようになる。

$$f(i) = \arg \max_a \sum_s M_{ss'}^a U(s') \quad (2.1)$$

ここで、 $\arg \max_a f(a)$ は $f(a)$ の値が最大となるような a の値を返す。記号 $M_{ss'}$ は動作 a が状態 s においてなされたときに状態 s' に到達する確率を表す。ある状

状態の効用 $U(s)$ は、全ての状態遷移の期待効用の和と報酬によって表される。

$$U(s) = R(s) + \max_a \sum_{s'} M_{ss'}^a U(s') \quad (2.2)$$

$R(s)$ は状態 s において与えられる報酬である。単純な動的計画法のアプローチでは、行為系列の終端における効用値を用いてそのひとつ前の全ての状態の効用値の計算を行うことから始める。そして、次にそのまた前の可能な全ての状態を計算する。しかし、多くの場合環境履歴はループになっている場合が多く、終端状態がないため、上に述べた方法を適用することはできない。そこで、反復手続きを用いて値の更新を行う。各状態における効用値は、隣接する状態で式 (6.1) を何度も適用することにより更新される。これは次の式で表される。

$$U_{t+1}(s) \leftarrow R(s) + \max_a \sum_{s'} M_{ss'}^a U_t(s') \quad (2.3)$$

ここで、 $U_t(s)$ は t 回反復後の状態 s の効用値を表す。 $t \rightarrow \infty$ とすると、効用値は与えられた条件下で安定した値に収束することが知られる。このアルゴリズムは値反復法 (Value Iteration) として知られる。状態の効用関数が与えられると、式 (6.5) を用いて容易に政策を決定することができる。

2.4. 相互作用と身体性

これまでに、エージェントと環境との相互作用を重視した内部表現獲得手法について説明してきたが、環境との相互作用の重要性は認知科学の分野においても指摘されている。本節では、環境との相互作用を説明する上で重要な概念であるアフォーダンスについて説明し、身体性の問題についても言及する。

2.4.1 アフォーダンス

強化学習などの、今日の相互作用を重視する研究に影響を与える概念としてアフォーダンスがある。アフォーダンス理論は、アメリカの知覚心理学者 Gibson によって、1960年代に完成された [13]。アフォード (afford) は「～ができる、～を与える」などの意味を持つ動詞であるが、英語にはアフォーダンス (affordance)

という名詞はなく、Gibson の造語である。アフォーダンスとは簡単に言うと、知覚者の受け取る情報を「環境が提供 (アフォード) している」というように、環境の方に情報発信者としての能動性をもたせようというものである。また、知覚者の存在にかかわらず、環境には潜在的に情報が含まれており、知覚者は単にその情報を受動的に受け取っているだけなのである。また、Gibson はアフォーダンスとは、環境が動物に提供する「価値」のことであると述べている。アフォーダンスは環境のあらゆる場面で存在する。例えば、Gibson の著書に書かれた例を引用すると、もしも陸地の表面がほぼ水平 (傾斜しておらず) で、平坦 (凹凸がなく) で、十分な広がり (動物の大きさに対して) を持っていて、その材質が堅い (動物の体重に比して) ならば、その表面は支えることをアフォードするのである。また、椅子は人間に対して座ることをアフォードし、崖は落ちることをアフォードする。椅子のような人間が創り出した道具の場合、アフォーダンスは機能と置き換えて考えることができる。

先の陸地の表面の例では、水平、平坦、広がり、堅さという属性をアフォーダンスを表現するための尺度として用いたが、これらは、面の物理的性質を指示しているものではない。これらの属性のいずれも、支えられる動物との関係を測るための尺度として用いられている。すなわち、アフォーダンスは環境の物理的な性質を表すのではなく、動物にとっての環境の性質を相対的に表現するものなのである。椅子は大人の人間に対しては、座るということをアフォードしているが大人と身長異なる小さい子供に対しては座ることをアフォードしていない。身体機能の大きく異なる動物例えば、象に対しても座ることをアフォードしない。このように、アフォーダンスは環境と主体との相対的關係に強く依存して存在するものなのである。

アフォーダンスに基づく工学的研究

ここで、アフォーダンス理論に基づく研究を紹介する。塩瀬ら [44] は移動ロボットの L 字路や十字路を通過する際の、注視点の選択を GA に基づいて行った。彼らは、環境が注視点形成のために情報をアフォードしていると述べている。しかし、環境のどのような要素が注視点の選択に寄与したかについての考察は行っていない。McDorman は [28] アフォーダンスを知覚することはセンサ - モータの

不変項を認識することと同じであり、エージェント自身が環境に対して何が出来るかを予想することが、アフォーダンスの知覚であると考えた。実際には、あらかじめ獲得された行動に対する画像入力モデルを用いてメンタルにパスを生成した。田川ら [54] はエージェントの頭脳を有限状態機械 (FMS) で記述し、環境内で適切に行動を行うための FMS を GA によって獲得した。エージェントのタスクは、異なった構造を持つ 2 つの部屋で障害物を避けながら、スタートからゴールまでたどりつくことである。エージェントが、適切な行動を学習すると、環境の構造の違いは、エージェント内部での表現 (FMS) の違いとなってあらわれる。田川らは、エージェントが通過することができた領域を正のアフォーダンス、衝突した場所を負のアフォーダンスと呼び、適切な FMS を獲得すること、すなわち環境へ適応することがアフォーダンスの発見であると主張している。エージェントが通過できるかどうかは、環境の構造、エージェントの身体の形状、大きさ、運動能力によって制約を受けるが、この文献では、これらの関係に対する詳細な考察を行っていない。羽倉ら [22][43][15] は、不変項をエージェントにとって有用な環境中の機能的属性であるとし、エージェント自らが学習する手法を提案した。不変項はエージェント内部における環境のモデル *Inner Perceptual Model* をエージェントが更新することによって獲得される。しかしながら、*Inner Perceptual Model* は、環境を反映した単なる内部モデルであり、しかも、行動にしたがってその表現が変化するため、不変項を獲得しているとは言い難い。塩瀬、田川の研究におけるアフォーダンスの解釈は、環境からなんらかの情報がアフォードされた結果として、行動が生成出来るという「結果に」注目したものである。また、羽倉の研究では、環境の構造を反映した内部表現の生成に環境からアフォードされた情報が用いられているという解釈である。これらの研究に共通することは、アフォードされた結果として何かが形成されるということであり、環境とエージェントの間のアフォーダンスの受渡しの方法やその情報などについての定義は行われていない。すなわち、環境における「どんな情報」を「どのようにして」エージェントが知覚するとアフォーダンスが知覚されたことになるのかが不明である。

もともとのアフォーダンスの定義に翻って考えてみると、アフォーダンスは環境内の物とそれを知覚する主体の間に発生するものであり、そこでは、対象の幾何的構造や物理的特性、主体の身体の大きさや形、運動能力が制約となっている。

すなわち，この制約こそアフォーダンスを定義するものであると言える．そこで次の節では，主体における身体の大きさや形，運動能力が認知に対してどのような影響を与えるかについて議論する．

2.5. 身体性

近年人工知能や認知科学の分野においてエージェントが身体をもつことの重要性が指摘されている [30] [35][49]．國吉ら [56][6] は，知能は質の異なる多様な相互作用系から構成されると考え，相互作用郡の中から有意義な秩序を抽出・生成する過程でその相互作用の制約となる身体性が重要な役割を果たすと主張し，人間型能動両眼視覚システム ESCHeR を用いて，首と目の冗長系において身体を通じて環境と相互作用を行うことによって協調構造が発生することを示した．佐々木はこのような相互作用からの秩序の抽出を説明するためにアフォーダンス理論を引用し，その抽出過程は行為主体が環境のアフォーダンスを探索する過程であると述べている [45]．國吉による身体性の定義は，相互作用において制約となるものという比較的大きな枠組を示すものであるが，本研究では身体性をより具体的に「エージェント固有の身体の広がりと運動能力そして，それらを知覚するセンサ」と定義する．

身体性が主体の認知活動に与える影響は大きい，例えば「椅子」は人間にとって座るための道具であるが，人間と全く異なる身体性を持つ蟻にとってはそれは巨大な構造物であり，もはや道具としての椅子の機能は存在しない．これは，物理空間において同一の対象であってもエージェントの身体性が異なれば，全く異なった意味を持つものとして認識されるということである．

エージェントが身体をもつということは，実世界においてある領域を占めることであり，そのために実世界との相互作用を生じる．アフォーダンスの項でも述べたように，身体はエージェントと環境の間に起こる相互作用における制約となり，特に身体の形による制約は大きい．例えば移動ロボットが回転した場合，身体の各部の描く軌跡はその大きさや形状によって異なる．ロボットは身体と環境の相互作用を通じて次の2つの事柄を知ることができる．一つは身体が環境内に占めている体積である．もう一つはロボットの行動の結果何が起こるかについて

の予測である．接触センサはそれらの事柄を知るための唯一のセンサであるので，本研究では自律エージェントが接触センサを持つことの重要性を強調する．

エージェントは身体性に基づいて環境を知覚するが，これは言い換えるとアフォーダンスを知覚していると言える．ここで，重要なのはアフォーダンスが身体性に依存して変化するということである．環境を知覚するためにエージェントは，センサを備える必要があるが，それ以前に身体を持つ必要がある．エージェントが広がりのある身体を持つと，その身体の広がり具合によって環境は異なって知覚される．したがって，アフォーダンスの知覚においても身体性は非常に重要な役割を果たす．身体性を用いて表現された対象は，その表現自体が物理空間における対象そのものを表現しているので，不変なものでありエージェント自ら構築した環境のモデルであると言える．

2.6. 視覚と身体性の関係

視覚は環境に関する多くの情報を含んでいるためにエージェントが環境知覚のために視覚センサを用いることの利点は多い．実世界を行動するエージェントは視覚を通じてどんな情報を得ているのだろう．我々は普通，視覚入力から環境内の対象の様々な情報を得て理解していると考えられる．しかし，この一見疑う余地のない見解に対して真向から反対する仮説がある．すなわち，我々は本質的に視覚を用いて「何も理解できない」というものである．

2.6.1 モリヌークス問題

Molyneux, W.(1656-1698) はアイルランドの哲学者であり，1693 年に Locke. J.(1632-1704) に宛てた書簡の中で次のような問題を提起したことで知られる．その問題は，Locke の「人間知性論」の中で紹介されており，Berkley, G.(1685) も「視覚新論」の 132 節でそれをそのまま引用している．その問題は次のようなものである．「ある生来の盲人が，今は成人して，同じ金属からできたほとんど同じ大きさの立方体と球とを触覚によって識別するように教えられ，その立方体と球とがテーブルの上におかれ，その盲人が見えるようになった，と想像してほしい．疑問：彼は両者に触れないうちに，視覚によってどちらが球でどちらが立方

体であるかを識別し言うことができるであろうか」この問に対し、鋭敏で賢明な提題者 Molyneux は次のように答える。「否．なぜなら、その人は球や立方体がどのように彼の触覚に作用するかについての経験はすでに得ているが、かくかくに触覚に作用するものが視覚にもかくかくに作用しなければならないという経験にはいまだ到達していないからである．あるいはまた、立方体の突起した角は、かれの手を不均一に圧迫するが、それは彼の眼にとっても、立方体においてそうであるように見えるであろう、という経験にいまだ到達していないからである。」(Berkeley 視覚新論より)

Molyneux 問題に対する Berkeley の見解は以下のようなものである．距離、大きさ、位置、運動などの空間に関する緒観念は、本来触覚によってのみ直接知覚され得るものであり、それらが視覚によって感知されるのは、経験によって、視覚的観念と触覚的観念との間に結合が生じているからだということである．つまり、われわれはしばしば、視覚によって直接対象の距離や大きさの観念をあたえられているように思い込む、しかしそれは、それらの触覚によって知覚される観念を、特定の視覚による観念と結びつけられたものとして、長期間経験してきたからにすぎないのである．すなわち、視覚の諸観念と触覚の諸観念の間にはそもそも共通する観念が存在しないということである．また、3次元空間の知覚は、光学的な計算によって成立するのではなく運動感覚を含む触覚によって訓練されることによって成立すると述べている．これは、Marr の視覚認識における計算論的アプローチ [29] を否定するものである．

2.6.2 開眼手術者に対する認知心理実験

Molyneux の問題提起時には、この問題の是非を実際に確かめることはできなかったが、その後の医学の進歩によって白内障の患者に対する水晶体摘出手術や、角膜疾患患者に対する角膜移植などが可能となり、この問題に対する解答を実証する機会が生まれた．鳥居らは、開眼手術を受けた患者に対する認知心理実験を通して次のように述べている [53]．視・運動系以外の感覚系を基礎として日常の認知活動を行ってきた先天盲の場合、「ものの形」やその他の属性の弁別・識別ならびに「もの」そのものの識別には、依然として触・運動系が中心的役割を果たし、そこにはまだ錬成を経ていない視・運動系の活動を新しく行動体制の中に組

み込むことは、想像以上に困難な課題であるように思われる。ある被験者は厚紙を切り取って作られた3種類の図形を識別することを求められたとき、眼で見ようとせずすぐに指で図形のふちに触れてその形の名前を答えようとする。そして、このような対象識別における触運動はその後の鍛錬によって徐々に消えて行くと述べている。開眼手術を受けた患者は、自分が実際に触れて知覚していた対象のイメージと見え方との相違すなわち身体的な表象と視覚の不一致に戸惑う。例えばある被験者は、3次元の対象の弁別実験において「立方体は手で触った感じでは裏の方も底の方も上もみんな同じ形なのに、“見る”ときには見る形によって形が違って見えるのだと驚いた」と述べ、また日常生活において「同じ幅の道でも遠くに行くほど細く見えていく、そういうことが私にとっては本当に驚きだった」と述べている。また、紙上に描画された対象に関して次のように述べている。「見る学習をする以前には立体感、特に紙の上に描かれたものに対する立体感がまるっきりなかった。どうして一枚の紙の上に建物とか山とかそういう高さのあるものが表現できるのかが大きな疑問であった」これは、Berkeley が主張するように通常視覚に属するとされる、距離、大きさ、空間内の位置などの知覚が、そもそも視覚とは全く関係ない他の感覚系に依存して成立していることを示すものである。

2.6.3 Held の動物実験

われわれの見えるの世界が運動系に基づいていることを示す例をもう一つ挙げる。Held & Hein[17] は、生後間もない子ネコに初めて視覚経験を与える際、対のうち一方は自発的な移動ができ(能動ネコ)、他方は自発的に移動ができない(受動ネコ)状況においた。その方法は、円形の小部屋に2匹のネコを入れ、能動ネコの首輪に取り付けた装置でその移動方向と距離を検出し、それを受動ネコが乗せられたゴンドラに伝えるというものであった。これにより、移動とその視覚的フィードバックはほぼ等しいが、その移動が自発的なものか受動的なものかという点だけが異なる対を作ったのである。このような視覚 - 運動経験を生後10週間まで暗室で育てた小ネコに1日3時間、計30時間与えた結果、10対のうち能動ネコでは正常な視覚行動を見せはじめたが、受動ネコではそれが形成されなかった。例えば、実験者がネコの体を持ってテーブルの上に下ろすと、能動ネコは足を伸

ばしてテーブルに触れようとするが，受動ネコは足を伸ばさなかった．また，透明な床の下が深く落ち込んで見える視覚的断崖に置かれると，能動ネコは回避行動をとったが受動ネコにはそのような行動は見られなかった．この動物実験は，正常な視覚行動が生じるためには，自発的な運動による視覚 - 運動経験が必要であることを示唆している．

2.7. 対象記述のための内部表現

本節の前半では，行動を行うためのマッピングの獲得方法，すなわち行為を記述するための内部表現の獲得方法について議論してきた．ここで，「実世界にあるもの」を記述するための内部表現について考える．人工知能の世界では，対象の内部表現と行為の内部表現が全く別であるという見解に基づいて研究が行われてきた．STRIPS においては，物を押してある場所から別の場所に移動させるという行為は $Push(box, x, y)$ のように記述され，行為は物や場所を引数とする関数として取り扱われている．このように行為と操作対象を全く別のものと仮定したことによって，その後のコンピュータビジョンやロボットビジョンの研究では，画像中に写っている対象が何であるかを言い当てることが目標とされ，Marr の視覚の計算理論 [29] に代表されるように，2 次元の画像データから 3 次元の幾何学的構造を定量的に再構成することが主な研究課題となってきた．

このような，対象認識と行為を分離して取り扱うという考えに対して，Aloimonos ら [2, 36, 11] は合目的ビジョン (Purposive Vision, Active Vision) を提案し，その中で「視覚は独立したシステムとして存在するのではなく，環境に働きかけるシステム全体の一部として機能する」と述べ，認識と行為が不可分であることを主張した．これは，認識は行動のためにあり，行動の結果認識も変化するというように，視覚の存在意義が環境との相互作用のループの中に存在するというものである．より極端な言い方をすれば，対象認識の結果として，エージェント内部に中間的な対象の内部表現が存在せず，相互作用のループそのものの中に暗に対象認識が含まれるということである．さらに言うと，対象の内部表現と行為の内部表現が全くの別ものであるという旧式の AI 考え方とは正反対に，対象の内部表現と行為の内部表現が全く同一であるということである．本研究では，

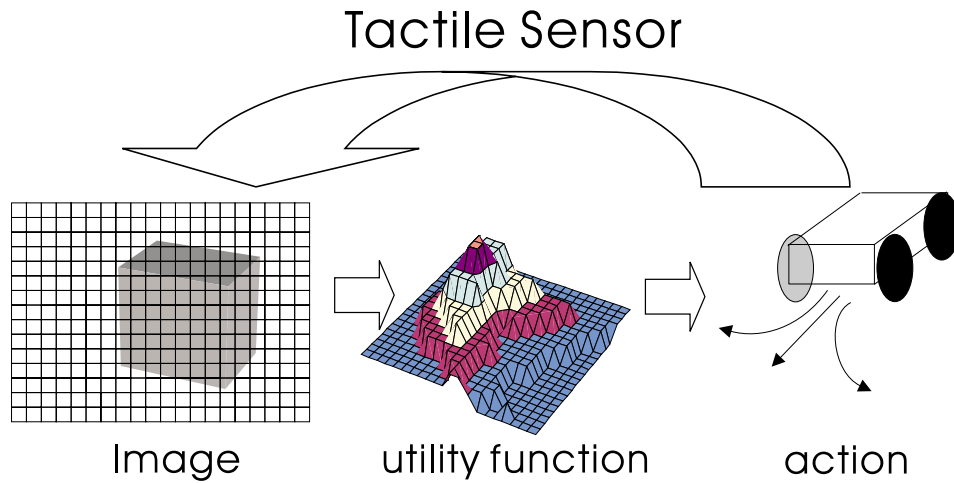


図 2.2 An architecture of acquiring a mapping between vision and action based tactile information

この立場を強く意識し，対象認識の問題をその対象についての行動生成の問題として考える．

2.8. 本研究におけるアプローチ

これまでに説明したことに基づいて本研究の方向性について述べる．本研究では環境内で自律的に振舞う知的エージェントを構築することを目指す．知的なエージェントはセンサからの入力を適切に判断し状況に応じた行動出力を行う必要がある．そのためにエージェントは入出力関係を記述したマッピングを内部表現として持つ必要がある．内部表現記述のためには，設計者の埋め込み知識は出来るだけ少ないほうがよく，エージェントは自らが環境と相互作用して得た経験を基に内部表現をエージェント自身によって獲得するべきである．また，内部表現の獲得時のコストを考えると，一旦獲得された内部表現を再利用出来ることが好ましい．視覚入力には環境における多くの情報が含まれているのでエージェントが視覚を用いる利点は多い．しかし，一方でその膨大は情報をどのように取り扱うかが問題となる．そこで，本論文では，どのように視覚を取り扱えば，設計者による知識の埋め込みを少なくして汎用的な表現を獲得できるかということ

問題として3つの研究を行う。すなわち、視覚入力から行動出力への内部表現獲得に関して次の事項を目標とする。

1. 設計者の埋め込み知識を少なくする
2. 汎用的な表現を獲得する

内部表現の記述方法として効用関数が有効であることが知られているため、本研究においても内部表現として効用関数を用いる。内部表現獲得における前提として、「視覚には本来物理世界に関する情報は本質的に含まれず、視覚は触覚を基準とする身体運動によって形成される」という、Berkeleyの仮説を採用する。そこで、内部表現獲得時に接触信号を有効に活用してマッピングを獲得することを目指す(図2.2参照)。

まず、4章では、視覚入力から状態空間を構成する際に埋め込み知識を少なくするために、自己組織化マップを用いる。この方法では、入力ベクトルの類似性に基づいてクラス分類しているため、タスク遂行のために注目しなければならない微小な特徴量の変化を見逃してしまうという欠点がある。そこで次に、5章では接触信号に基づいてタスク遂行に必要な視覚的特徴の選択と選択された特徴軸の分割を行う。この手法では、実世界におけるタスクを全て接触信号に関連するものであるという仮定のもとに、視覚的特徴に接触信号を関連づけるものである。あらかじめ用意した視覚的特徴は対象の領域情報などであったが、これらは直接的に接触情報に関連するものではない。そこで、6章では、本質的に接触信号に関連する視覚的特徴が対象の境界であることに着目し、微小領域内に検出される対象の境界に接触信号を関連づけることを行った。このようにして獲得されたマッピングは本章で議論した身体性の問題を考慮するものである。

第3章

ロボット

研究を行うために3種類のロボットを使用した。ひとつは Applied AI System の市販ロボット，後の二つは筆者によって製作されたものである。以下では，それぞれタイプ I，タイプ II，タイプ III と呼ぶ。本章ではそれぞれのロボットについて仕様を説明する。また製作したものに関しては設計指針，実現方法について述べる。タイプ II タイプ III はロボットのサッカー競技 (RoboCup) のレギュレーションに準ずるものである。三種類のロボットの共通点はいずれも二輪駆動型の移動ロボットであり，視覚センサとしてカメラと触覚センサを備える点である。一方異なる点は，接触センサの分布である。タイプ I では接触センサはロボット前部の一部に装備しているだけであるが，タイプ II ではロボットの前部に分布させて配置され，タイプ III ではロボットの全周に配置される。また，タイプ II，タイプ III はタイプ I と異なりリモートブレインシステムではなく全ての制御をロボットの上に搭載したノート PC で行っている。

3.1. タイプ I のロボット

タイプ I のロボットとして IS Robotics 社の 2 輪駆動型ロボット R-3 Robot を用いた。本実験で用いられたロボットを図 3.1(a) に示す。センサシステムとしては，視覚センサとして CCD カメラを搭載し，また接触センサとしてグリッパの先端に赤外線センサを装備している。ロボットは Powered Wheel System によって移動する。片側のモータに送るコマンドはそれぞれ $-1, 0, +1$ であり，それら

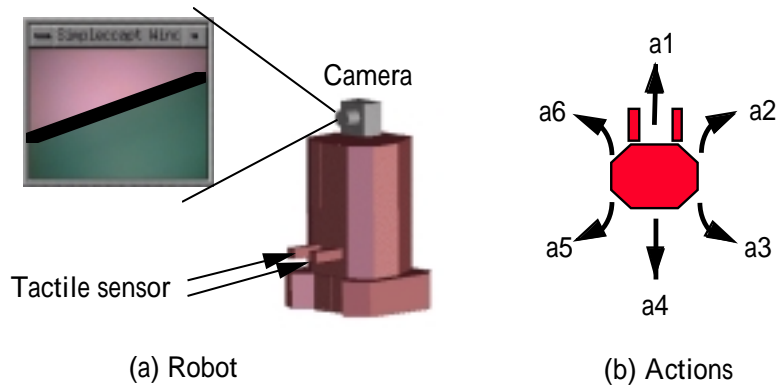


図 3.1 Type-I robot

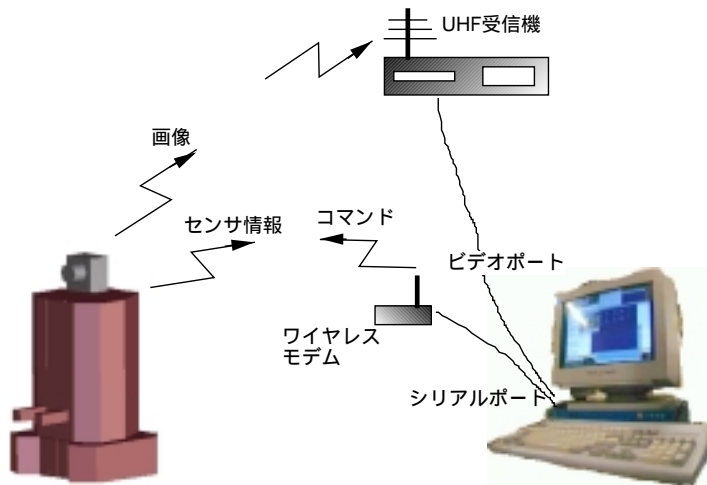


図 3.2 System architecture of type-I robot

の組合せによって次の6つを行動要素とする。

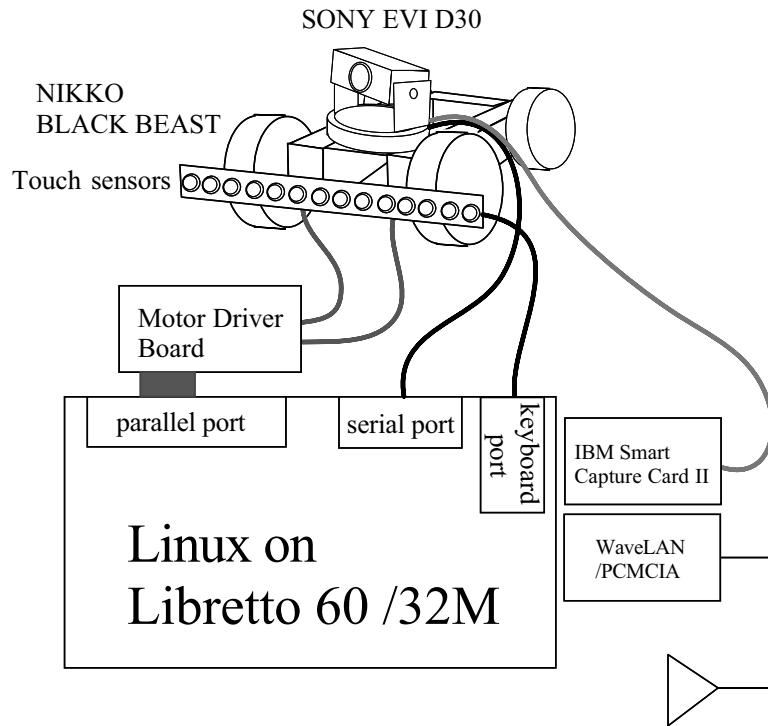
$$a_1(1, 1), a_2(1, 0), a_3(-1, 0) \\ a_4(-1, -1), a_5(0, -1), a_6(0, 1)$$

従ってロボットは6方向に移動可能である(図3.1(b)参照)。CCDカメラの出力する画像は 160×120 ピクセルである。

また、ロボットのコントロールシステムとしてはリモートブレイン方式を採用した。システム構成を図3.2に示す。ロボットの制御は外部のSilicon GraphicsのIndyワークステーションからワイヤレスモデムを介して行われる。ロボットは近接センサとして赤外線センサを高さ5センチのバンパーの周囲に12個備えている。また、ロボットにはCCDカメラが備えられており、キャプチャーされた画像はUHFトランスミッターを用いてワークステーションに入力される。

3.2. タイプ II のロボット

タイプIでは、視覚と触覚の関連付を行うために十分な機構ではないので、タイプIIの製作を行った。タイプIIの目的は、与えられた環境の中でタスクを遂行するために環境を認識するために必要十分なセンサ、アクチュエータ、コントローラを組織化することである。なお、タイプII、タイプIIIはロボットのサッカー競技(RoboCup)に参加することを前提として作成されたものであり、RoboCupのレギュレーションに準ずる。以下にRoboCupについて簡単に説明する。RoboCupは人工知能とロボティクスの研究の両方における問題を現実的なアプリケーションを通じて解決しようという試みである[25, 3, 24]。RoboCupは機械学習やエージェントアーキテクチャなどのさまざまな理論を評価するのに適した環境を与えるものである。RoboCupのフィールドとゲームはロボットが適合しsurviveするために適している。我々は、ロボットがどのようにあるべきか、また環境のもとでどのように振舞うかについて考えなければならない。すなわち、環境と知覚、とアクチュエータの機能である。RoboCupの環境は実際のものとは限定的で簡単であるという意味で大きくことなる。しかし、サッカー競技では敵の動きが動的かつ予測困難なので、動的かつ非決定的であるという実際の環境を反映してい



(a) Hardware configuration of Type-II robot



(b) Appearance of type-II robot

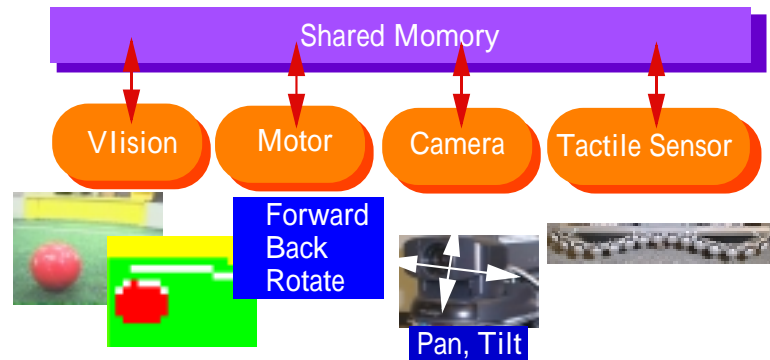


図 3.4 Software architecture

る．さらに，環境が限定的かつ簡単であるという事とは対比的に，センサやアクチュエータは比較的不正確で限定されていると考える．

3.2.1 システムの概要

ロボットのハードウェアはノート型 PC(東芝リブレット)，視覚システム，接触センサシステム，台車の 4 つのコンポーネントから形成される．それぞれのコンポーネントの間を関係を図 3.3 に示す．各ハードウェアはそれぞれモータコントローラ，カメラコントローラ，接触センサコントローラ，視覚モジュールの 4 つのソフトウェアモジュールによって制御される．ノート型 PC に搭載されるオペレーティングシステムは Linux [1] である．

これらのモジュールは共有メモリ [38] を介して通信を行い非同期にコントロールされる．図 3.4 はソフトウェア構成である．それぞれのソフトウェアモジュールは共有メモリを非同期に読み書きすることによってモジュール間の通信を行う．

3.2.2 駆動システム

台車として玩具の 4 輪駆動型のラジオコントロールカー，NIKKOH の BLACK BEAST を用いた．図 3.5 参照．駆動方法は，2 つの独立したモータが左右でそれぞれ前後の 2 輪を駆動する PWS (Power Wheeled Steering) 方式である．この方

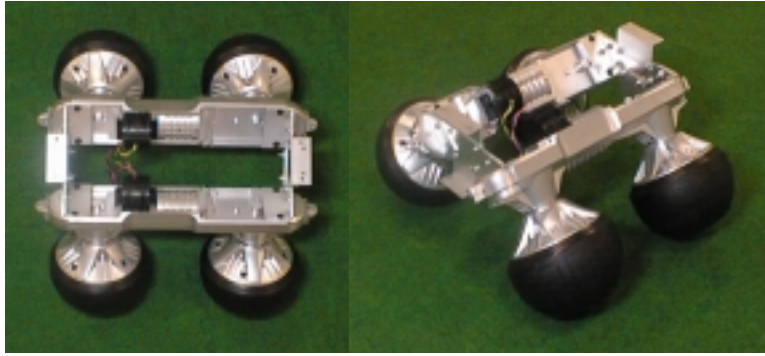


図 3.5 Driving system

式のため、タイヤの滑べりを伴ってその場で回転することができる。市販状態ではギア比が高く全てのシステムを搭載して移動するのに十分なトルクが得られないため遊星ギアを用いて必要なトルクを発生するようにした。その結果ペイロードは 3 kg となった。

3.2.3 視覚センシングシステム

視覚システムは CCD カメラ (SONY EVI D30)、PCMCIA 規格のビデオキャプチャーカード (IBM Smart Capture Card II)、画像処理を行うソフトウェアによって構成される。CCD カメラによって撮影された画像はカメラ内で NTSC 信号に変換され、キャプチャーカードによってデジタル信号に変換される。ソフトウェア部分では、画像処理モジュールによっていくつかの色領域に分割される。カラーフォーマットとして、YUV を用いた。

EVI-D30 は、PC から RS232C ポートを経由してパン・チルトの制御を行うことができる。パンの角度は -100 から $+100$ 、チルトの角度は -25 から $+25$ である。RoboCup フィールドでは、種類、機能に応じて環境内の対象が色分けされている。例えば、地面は緑、2つのゴールはそれぞれ青、黄色、ボールは赤などである。

各ピクセルの色の値を5つのカテゴリーに分類するために超楕円体を用いた判別方法を用いた。あらかじめ各色について楕円モデルを準備しておき、各ピクセルの値が楕円内に入っているかどうかを判別し、入っていればその色のタグを割

り当てる．楕円モデルは，重心とマハラノビス距離によって定義される．各色の楕円モデルは，単色にのみによって構成される理想的なサンプルデータから計算される．マハラノビス距離の2乗 D^2 は次の式によって計算される．

$$D^2(x, y, z) = (x - \bar{x}, y - \bar{y}, z - \bar{z}) \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}^{-1} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \\ z - \bar{z} \end{bmatrix} \quad (3.1)$$

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} = \begin{bmatrix} \frac{w_{11}}{N-1} & \frac{w_{12}}{N-1} & \frac{w_{13}}{N-1} \\ \frac{w_{21}}{N-1} & \frac{w_{22}}{N-1} & \frac{w_{23}}{N-1} \\ \frac{w_{31}}{N-1} & \frac{w_{32}}{N-1} & \frac{w_{33}}{N-1} \end{bmatrix} \quad (3.2)$$

$$\begin{bmatrix} \frac{w_{11}}{N-1} & \frac{w_{12}}{N-1} & \frac{w_{13}}{N-1} \\ \frac{w_{21}}{N-1} & \frac{w_{22}}{N-1} & \frac{w_{23}}{N-1} \\ \frac{w_{31}}{N-1} & \frac{w_{32}}{N-1} & \frac{w_{33}}{N-1} \end{bmatrix} = \begin{bmatrix} \Sigma x^2 - \frac{(\Sigma x)^2}{N} & \Sigma xy - \frac{\Sigma x \Sigma y}{N} & \Sigma xz - \frac{\Sigma x \Sigma z}{N} \\ \Sigma xy - \frac{\Sigma x \Sigma y}{N} & \Sigma y^2 - \frac{(\Sigma y)^2}{N} & \Sigma yz - \frac{\Sigma y \Sigma z}{N} \\ \Sigma xz - \frac{\Sigma x \Sigma z}{N} & \Sigma yz - \frac{\Sigma y \Sigma z}{N} & \Sigma z^2 - \frac{(\Sigma z)^2}{N} \end{bmatrix} \quad (3.3)$$

ここで $\bar{x}, \bar{y}, \bar{z}$ は各色のサンプルデータの平均値である．また N はサンプルデータ数である．

3.2.4 接触センサシステム

接触センサはボールや壁などの対象との接触状態を検出するのに用いられる．サッカーロボットは頻繁に他のロボットや壁，ボールと接触するので接触センサが果たす役割は大きい．さらに，接触センサは視覚の制限を補うことができる．ロボットに搭載されたカメラの視界は限定されているためロボット同士の接触，ロボットとボール，壁との接触が視界の外で起こるとロボットは画像を用いて接触状態を検出することはできない．この問題は接触が予想される場所に接触センサを配置することによって解決される．

接触センサはPCのキーボードを利用して製作した(3.6参照)．接触センサは帯状に配置する．キーボードを分解し基板だけ取り出し，配置された接触センサの出力を基板に入力する．キーボードはON/OFFのスイッチがマトリクス状に配置されたのである．日本語キーボードの場合は109個のスイッチによって構成さ

れる．キーボードはキーに応じて ASCII コードを出力するので，コードを読み取ることでどの位置のセンサに接触したかを判別することができる．

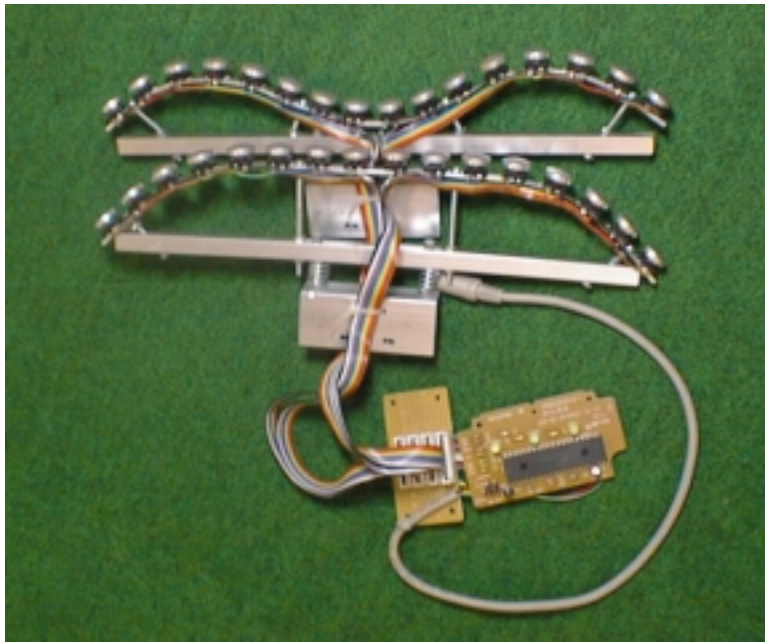


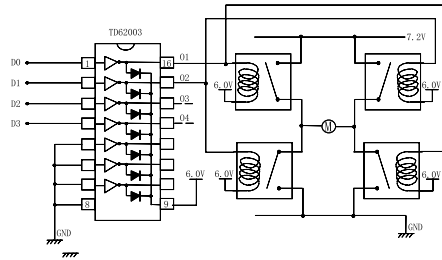
図 3.6 Tactile sensors by utilizing a key board

3.2.5 モーターコントロールシステム

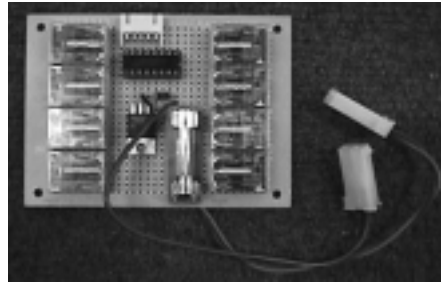
モーターコントロールシステムは PC とモータの間のインタフェースボードであり，2つの DC モータを制御する (図 3.7 参照)．このボードは PC のパラレルポートに接続され 2 5 6 段階のモータコマンドに応じて PWM 制御により 2 5 6 段階の回転数を出力できる．

3.3. タイプ III のロボット

タイプ II のロボットでは，接触センサが前後にしか配置されていないため側面が環境に接触してもそれを知覚することができず，スタックに陥ってしまうとい



(a) caption



(b) caption

図 3.7 Motor drive board

う問題があった，そこで，全周囲に接触センサを設置したタイプ III のロボットを製作した．形状がほぼ円形で接触センサを全周囲に配置したこと以外はハードウェア，ソフトウェア構成はタイプ II と同じである．タイプ III のロボットを図 3.8に示す．

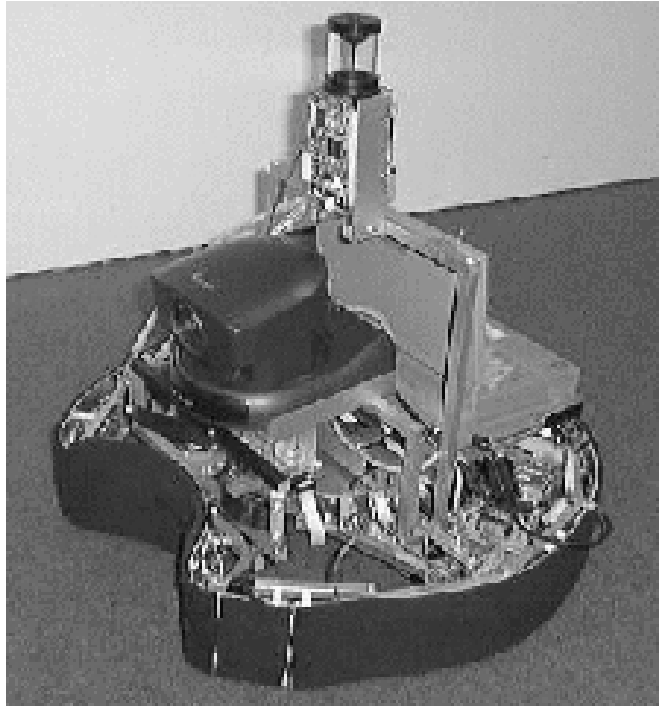


図 3.8 Type-III robot

第4章

自己組織化マップを用いた状態空間の構築

本章では、視覚入力から行動出力のマッピングを記述するための内部表現として *Visual State Action Map* (以下 *VSAM*) を提案する。またその中で視覚を用いて行動を学習するエージェントが回避できない問題である状態数削減の問題を議論する。提案した手法は実機を用いた実験によってその有効性を確かめられた。

本章の構成は次のとおりである。まず、画像から行動へのマッピングを学習するエージェントについて概説し、その中で画像の状態数削減が問題となることを示す。次に、本手法で提案する *VSAM* の説明を行い、その中で、*SOM* を用いた状態数削減と強化学習を用いたマッピングの獲得について説明する。次に実機を用いた実験の説明を行い、最後に本章のまとめと提案手法の考察を行う。

4.1. 視覚から行動へのマッピングの学習

視覚情報は環境に関する多くの情報を含んでいるため、エージェントが効果的に視覚を用いると複雑な行動が可能になるが、その半面多すぎる情報を使いこなせないとエージェントは行動できない。したがって、エージェントが視覚を有効に用いて行動するためには、膨大な情報を含む視覚入力をいかに取り扱うかが問題になる。例えば、カメラからの画像サイズが 10×10 で2値の場合でさえ、エージェントが考慮すべき状態数は 2^{100} になる。そこで、設計者は様々な先見的な知識を用いて入力画像の状態数を減らそうと試みる。例えばそれは、2値化や

エッジ検出，領域抽出などの画像処理技術であるかもしれないし，主成分分析などの統計的な解析手法かもしれない．しかし，それらの手法は状況に応じて設計者が決定するアドホックなものである．例えばエッジ検出や領域抽出では勾配を計算するためのオペレータがいくつも提案されており，その選択は設計者の主観に委ねられる．最も簡単な画像処理である2値化の処理でさえ，閾値の決定は設計者によって決定される．例えば， 10×10 の2値画像でさえ，可能な状態数は 2^{100} となる．

視覚に基づく強化学習においても，膨大な情報を含む入力をいかに取り扱うかが問題になる．視覚に基づく強化学習では視覚情報を基準として状態空間を構成するので，画像上で識別可能な空間と，ロボットの三次元空間とは必ずしも一致しない．テレビカメラの近傍では分解能が高く，逆に遠方では低い．これに対して，ロボットの行動はもとの三次元空間でほぼ同じ量の運動として実現されるので，ロボットによる行動と状態遷移が1:1に対応するとはかぎらず「状態と行動のずれ」生じる．浅田ら [50] は，ロボットの実際の行動を行動要素として定義し，状態変化を伴うまでの同一行動要素の集合を行動として再定義することにより，このずれ問題に対処した．また，視覚を入力として用いた強化学習では，学習可能な状態空間構築のために何らかの画像処理を行って状態数の削減を行うのが普通である．中村らは視覚から行動へのマッピングを表現するために運動スケッチを提案した [33]．運動スケッチでは，画像データから抽出される視覚的特徴としてオプティカルフローを用い，そのデータに対して主成分解析を行うことで，あらかじめタスク遂行のために必要なパラメータ推定を行い，状態空間として用いて学習を行った．単眼カメラを有する実ロボットを用いた実験によって，運動スケッチを獲得したエージェントは，障害物回避や目標追従などの行動を達成できることが示されている．浅田らは視覚を持つサッカーロボットのシュート行動の学習のために，Q学習を用いた [50]．この中で浅田らは，あらかじめ画像を行って抽出したボールやゴールの領域の大きさや位置を状態空間として用いた．Ishiguroらは全方位視覚を有する移動ロボットによるナビゲーションタスクに強化学習を適用した [19]．Ishiguroらは状態空間として，主成分分析によって抽出された画像特徴を用いた．MaCallumは自動車の自動走行を行うための強化学習アルゴリズム *U-Tree* を提案した [31]．この中では遠くの赤いものなど，画像処理によって

抽出された比較的抽象度の高い特徴が状態空間として用いられた。

これまでに紹介した手法はいずれも、学習可能な状態空間構築のために何らかの画像処理手法を用いている。しかし、自律エージェント構築の立場からは、状態数削減のため用いる設計者の知識はできるだけ少ない方が良く、できるだけ画像処理を行わないことが望まれる。柴田らはこのような立場から Direct Vision based Reinforcement Learning を提案し、視覚入力を直接ニューラルネットワークに入力し状態空間を自動的に生成する方法を提案した [37]。我々も、直接の画像をシステムへの入力として用いることを目指す。

これまでに述べた問題点を整理すると以下の2点である。

1. 強化学習を行うために適切な状態数までどうやって減らすか。強化学習を行うための計算時間は状態数の指数オーダーで増加するため、学習を効率良く行うためには状態数は少ないほど良い。
2. 自律エージェント構築の立場から、状態数削減のために用いる設計者の知識はできるだけ少ないほうがよい。

本章では、これらの問題を考慮に入れて、視覚入力から行動出力へのマッピングの獲得のために、*Visual State-Action Map (VSAM)* を提案する。VSAM では、自己組織化マップ (Self-Organizing Map)[26] によって類似した入力画像を一つのクラスに分類することによって入力次元を減らし、生成された状態空間上で強化学習を行うことで、視覚と行動の関係を獲得するものである。

4.2. Visual State Action Map

VSAM では、視覚入力から状態数を削減して状態空間を構築する際の画像処理において、設計者の埋め込み知識出来るだけ少なくするために自己組織化マップを用いる。自己組織化マップでは、入力ベクトルの類似性に基づいて教師無し学習を行うので、設計者の知識に頼ることなく状態空間を構築することができる。自己組織化マップを使う利点は、類似した入力を同じ状態として考えることが出来るため、画像内での微小な変化を無視することができることである。また、

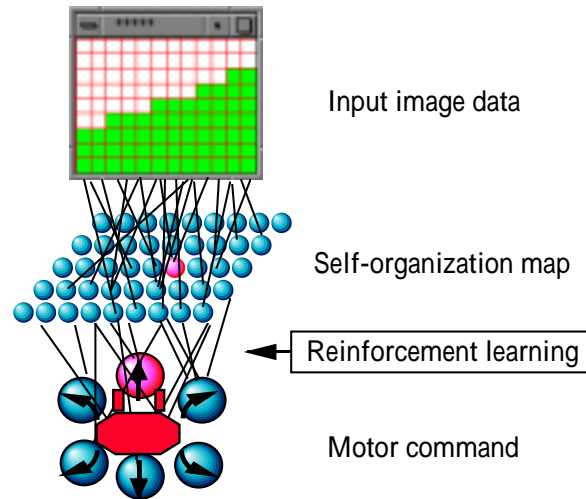


図 4.1 Architecture

VSAM では、自己組織化マップによって構築されたノードを状態空間として用いて強化学習を行うことによってそれぞれの状態についての適切な行動を学習する。

VSAM は、3層のニューラルネットワークによって説明することができる(図4.1参照)。入力層は離散化された画像であり、2次元ベクトルである。入力ベクトルは中間層である自己組織化マップ上の2次元のノードの配列にマッピングされる。自己組織化マップは状態空間として用いられ、出力層の各ユニットに関連づけられる。出力層の各ユニットはモータコマンドを表している。中間層と出力層の結合重みベクトルは、強化学習によって決定され、その結合の強さによってモータコマンドが選択される。

VSAM を構築するための過程は次の三つのステップからなる。

1. 環境をランダムに行動してあらかじめ収集された画像データに対して自己組織化マップを用いて画像のカテゴリ化を行う。自己組織化マップのノード群によって状態空間が構築される。
2. 環境をランダムに行動し、強化学習を行うために遷移確率の推定と報酬値の記録を行う。
3. 構築された状態空間と遷移確率、報酬を用いてオフラインで強化学習を行

い，適切な行動の学習を行う．

以下にそれぞれのステップについて詳しく述べる．

4.2.1 自己組織化マップを用いた画像のカテゴリー化

T. コホネンが開発した自己組織化マップ (Self Organization Map : SOM) は、ニューラルネットワークアルゴリズムであり、高次元のデータを2次元のマップ上に投影できる解析法である．マップは2次元でなくても良いが，以下の説明では2次元を前提とする．自己組織化マップは、SOMの特徴としてあげられる2次元のマップ上に視覚化することを利用し、様々な多変量のデータ解析に用いられ、その有効性も高まってきている．本研究では、ピクセルの微小な変化を無視し、類似した状態を一つの状態として取り扱うために、この手法を採用する．また、可能な画像を生成して状態空間を構築すると、実際には存在しない状態が作られ状態数が増えてしまうため、実際に環境内を行動して得られるデータから状態空間を構築する．

SOM は入力データ空間 R^n から任意の次元のノードへのマッピングとして定義される．それぞれのノードは参照ベクトル $m_i = [\mu_{i1}, \dots, \mu_{in}]^T \in R^n$ に関連づけられている．入力ベクトル $X = [x_1, \dots, x_n]^T \in R^n$ は、可変のスカラー重み μ_{ij} を介して並列に全てのノードと結合しており、全ての m_i と比較される．ユークリッド距離 $\|X - m_i\|$ を最小にするようなノードが勝者ノードとして定義され添字 c によって表される．

$$\|X - m_c\| = \min_i \{\|X - m_i\|\} \quad (4.1)$$

ノード i の出力 y_i は

$$y_i = \begin{cases} 1 & i = c \\ 0 & \text{otherwise} \end{cases}$$

である．SOMの学習は、勝者ノード m_i とその近傍のノードの重みを更新することによって行われる．勝者ノードの近傍のノードに対する重みの更新は局所的な平滑効果と大局的な順序づけの効果を生じる．

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{ci}(t)[\mathbf{X}(t) - \mathbf{m}_i(t)] \quad (4.2)$$

ここで $t = 0, 1, 2, \dots$ は整数すなわち離散時間座標である．関数 $h_{ci}(t)$ は近傍関数と呼ばれ，格子点上に定義される平滑カーネルとして振舞う．SOM の学習が収束するためには， $t \rightarrow \infty$ のとき $h_{ci}(t) \rightarrow 0$ であることが必要である．普通 $h_{ci}(t)$ は次のように定義される．

$$h_{ci}(t) = h(\|r_c - r_i\|, t) = \alpha(t) \cdot \frac{(N_c + 1 - \|r_c - r_i\|)}{N_c} \quad (4.3)$$

ここで $r_c \in \mathbf{R}^2$ と $r_i \in \mathbf{R}^2$ はそれぞれの配列の中でノード c と i の位置ベクトルである． $\alpha(t)$ は学習率であり， $(0 < \alpha(t) < 1)$ である． N_c は c の近傍の配列である．通常 $\alpha(t)$ と N_c の半径は学習が進むにつれて単調減少する．

$$\alpha(t) = \alpha_0 \left(1 - \frac{t}{T}\right) \quad (4.4)$$

$$N_c = (d_0 - 1) \left(1 - \frac{t}{T}\right) + 1 \quad (4.5)$$

ここで α_0 は定数， d_0 は N_c の半径の初期値， T は全ての学習時間である．

4.2.2 入力画像の前処理

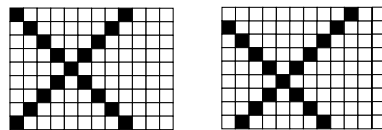


図 4.2 Two line figures that are orthogonal

SOM のアルゴリズムでは，入力パターンの重みベクトルをユークリッド距離を用いて比較する．このため，パターンに対する基本的な操作，例えばパターンの並進，回転，拡大などにおいて微小な変化に対してさえ耐え得ることができない．図 4.2 のような入力パターンを考える．2 つの 2 値パターンは我々にとって非常に似ているように見えるが直交している．これは SOM において全く異なる

パターンであると識別されるということである．この問題を解決するために前処理として最も簡単なぼかしを用いる．ぼかしは以下の式にしたがって行われる．

$$C(d) = \begin{cases} 1 & d = 0 \\ 0.66 & d = 1 \\ 0.33 & d = 2 \\ 0 & d > 2 \end{cases}$$

ここで d はセル (i, j) からの距離である．すなわちパターンのある点 (i, j) の値が 1 のとき，近傍の点の値をその距離にしたがって減衰させた値とする．

4.2.3 状態遷移確率モデルの獲得

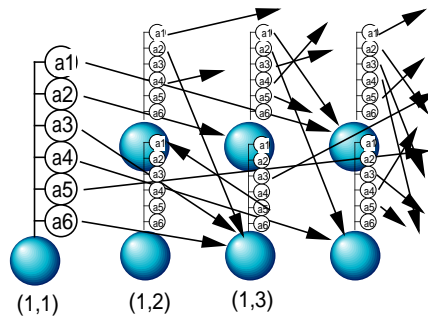


図 4.3 Transition probability model

VSAM のそれぞれのノードは 6 つのサブノードを持つ．サブノードは強化学習を行うために状態遷移確率モデルと報酬を保持している (図 4.3 参照)．次に自己組織化マップの、ユニット間状態遷移確率と報酬の獲得を次のように行なった．以下は VSAM における状態遷移確率モデルと報酬獲得の手順である．

1. $a \leftarrow$ ランダムに行動を選択する
2. $U_t \leftarrow$ 状態が変化し勝者ユニットが決定する
3. $r \leftarrow$ 報酬を与えられたら報酬を記録する
4. U_{t-1} に a, U_t, r を記憶させる

4.2.4 学習

学習方法は Q 学習を用いた . Q 学習では、状態 s_t で行動 a_t をとり、それ以降最適な行動をとったときの評価の見積もりを Q 値とし、再帰的に

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_{b \in A} Q(s_{t+1}, b) - Q(s_t, a_t)) \quad (4.6)$$

と定義される . r_t は状態 s_t で行動 a_t をとったときの報酬である . また α は学習率であり、 $0 < \alpha \leq 1$ なる定数である . γ は割引率である .

学習は以下のステップで行なう .

1. $s \leftarrow$ 現在の状態
2. $a \leftarrow$ 行動をランダムに選択
3. $s' \leftarrow$ 次状態、 $r \leftarrow$ 状態 s' における報酬
4. $Q(s, a)$ を更新

4.3. 実機を使用した実験

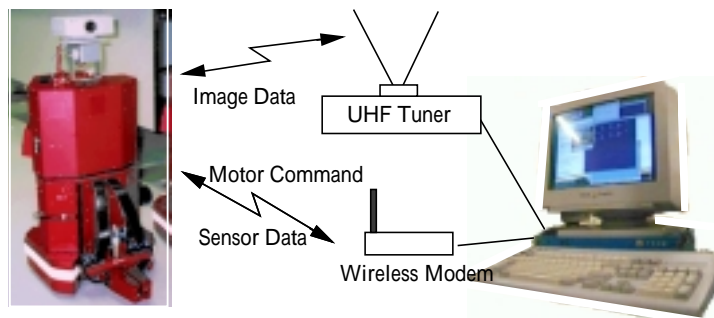


図 4.4 Configuration of system

提案した手法の有効性を示すために VSAM を実際の移動ロボットに適用した実験を行った . 実験にはタイプ I のロボットを用いた (図 4.4 参照) . 実験環境と

して、高さ 40cm の壁に囲まれた $90\text{cm} \times 180\text{cm}$ の矩形領域の人工的な環境を使用した。また、壁と床の境界は視覚認識可能なように 5cm の黒いラインでマーキングした。

4.3.1 VSAM の構築

自己組織化マップの生成

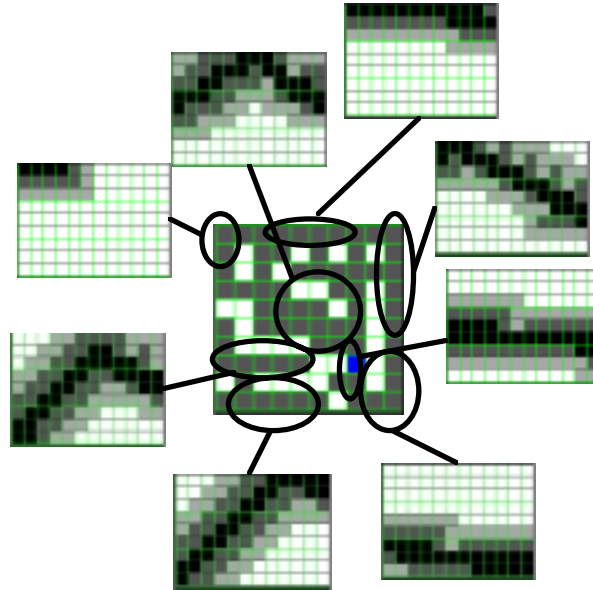


図 4.5 Generated map

まず、環境内を動き回ってサンプルデータの獲得を行った。設定された環境中全域に渡って、壁に衝突しないようにロボットをランダムに環境内を行動させた。その結果 201 枚の画像を収集することができた。これらの画像に対して、先に述べた自己組織化マップアルゴリズムを用いてクラス分類を行った。自己組織化マップのノードとして 10×10 の 2 次元配列のノードを用意した。また、学習回数は 200 ステップとした。自己組織化マップの入力として用いた画像は 160×120 ピクセルを次のルールにしたがって圧縮し 16×12 ピクセルにした 2 値画像である。もしオリジナル画像中の 10×10 ピクセルの領域内の 20% が黒、すなわち床と壁の境界が検出されたらそのピクセルの値は 1、それ以外は 0 とした。

図 4.5は学習された自己組織化マップである．入力ベクトルがマッピングされたユニット数は 70 であった．図 4.5中の円は人間の目で見ると類似したパターンの集合を示している．この図から，自己組織化マップは類似した画像を近傍のノードに配置していることがわかる．隣接するノード間のユークリッド距離の平均は 1.4 である．

遷移確率の獲得

表 4.1 Example of VSAM

	a_1	a_2	a_3	a_4	a_5	a_6	reward
1,9	3,3	1,7	0,4	3,3		0,6	0
2,9	2,2	0,6	6,9	3,3	2,5	5,5	0
3,9	3,6	0,5	6,8	3,4	3,5	5,6	0
4,9		0,5	7,9	2,5		7,6	0
5,9	3,9	8,9	8,7	2,6	5,6	7,5	0
6,9				2,9	3,9	,6,8	-1

次に，先と同様にランダムな行動を行うことによって各ノード間の状態遷移確率を求めた．報酬は、ロボット前面にあるグリッパの先端にある赤外線センサによって壁が検知されたら -1 を与え、それ以外は 0 を与えるものとする．600 試行行なった結果、選択可能な状態と行動 420 組のうち、317 組の関係が獲得された．600 回のランダムな行動の結果 317 個の遷移を獲得することができた．ノード数が 70 で選択可能な行動が 6 なので、可能な遷移は 420 個ある．したがって全ての遷移のうちで、約 80% の遷移が獲得されたことになる．表 4.1 は獲得された状態と行動の例である．例えば、(1, 9) において行動 a_4 をとると状態は (3, 3) に遷移する．エージェントが壁にぶつかり手の先に付いている接触センサが反応すると報酬が得られる．報酬の値は -1 である．表中の空欄は遷移が観測されなかったことを示す．

行動学習

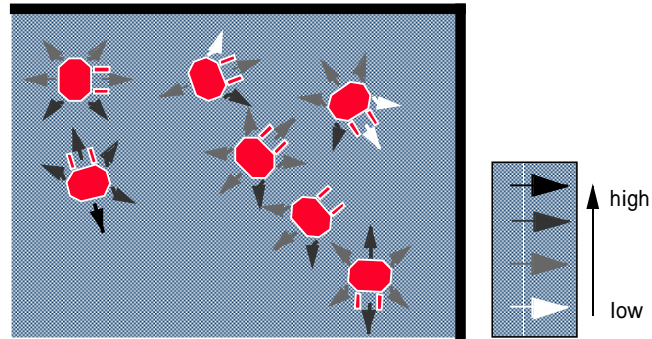


図 4.6 Learned Q value: When agent is going along the wall, Q value of the action toward wall is low and Q value of the action leaving from the wall is high.

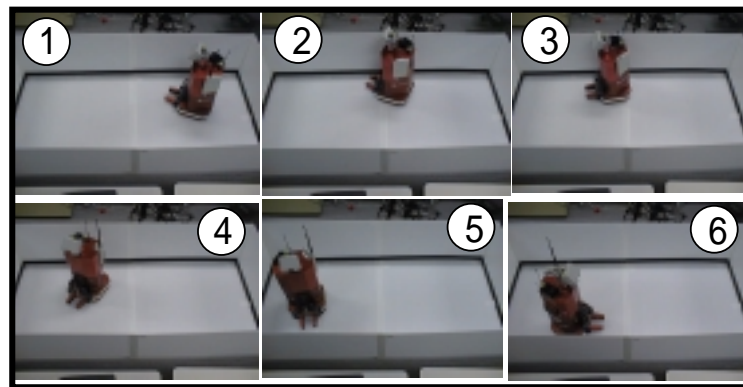


図 4.7 The action avoiding the wall. 1,4,6: The agent turns the corner.; 2,3,5: The agent goes along the wall.

次に、獲得された状態間の遷移確率と報酬を用いてオフラインでQ学習を行った。獲得された状態空間を用いてVSAMを構築するために図4.6は学習が十分収束した後のQ値の例を示している。図中の濃淡のつけられた矢印はQ値の大きさを示している。すなわち色が濃いほど高いQ値を表し、色が薄いほど低いQ値をあらわしている。図から、壁に向かう行動ほどQ値が低いことがわかる。したがって、高いQ値を選択するような政策のもとではエージェントは壁を避ける行動を行う。

図4.7に獲得されたVSAMを用いてロボットに実際に行動させた様子を示す。図よりエージェントは壁に衝突することなくコーナーで方向を変え、壁に沿って走行していることがわかる。

4.4. まとめと考察

本章では、視覚入力から行動出力へのマッピングを表現するために、自己組織化マップと強化学習を用いてマッピングの学習を行う *Visual State Action Map: VSAM* を提案した。その中で、特に入力として視覚を用いる場合の問題点として、1) 状態数削減の問題と、2) 状態空間構築における設計者の介在の問題、を挙げ、それを解決する方法として、自己組織化マップを用いて解決することを試みた。提案した手法を実機に適用して実験を行ったところ、削減された少ない状態数で壁を回避する行動が学習された。

実験では、入力状態は35%に圧縮されたが、200状態から70状態というようにサンプル入力データの数が少ないため、35%という数字が妥当であるかどうかは分からない。しかし、一つ言えることは、実験において環境のあらゆる場面において、エージェントは壁を回避して行動を行うことができたので、実験で用いた限定された環境においては、35%の圧縮で十分であったということである。しかし、実験で用いた環境は直線で構成される簡単な環境であるため、より複雑な環境において圧縮率は低くなると予想される。提案手法をより複雑な環境に適用するためには、VSAM上に構築された状態を適当な視点からさらに圧縮(抽象化)する必要がある。さらに状態数を削減する方法としては次のような方法が考えられる。提案手法では、自己組織化マップアルゴリズムによって入力画像をマップの

ノード上にマッピングしただけであったが、類似したノードは近傍に配置されるという自己組織化マップの性質を利用して、何らかの指標に基づいて近傍のノードをまとめより状態数を削減できることが予想される。

本手法では状態遷移は決定的としたが、より複雑な環境においては、遷移を非決定的なものとして扱う必要がある。しかしながら、自己組織化マップの効果によって、ある一定範囲にある、知覚入力と同じノードのマッピングされ、ある程度の誤差を含んだ遷移を決定的なものとして扱うことができると考えられる。

VSAM を動的環境に適用することを考える。動的環境では静的環境に増して状態空間は膨大になる。なぜなら、VSAM のノードは画像全体を入力としているからである。この問題を回避するためには、環境内の対象が移動すると画像は異なったものとして解釈され新たなノードを作る必要がある。したがってエージェントは対象を画像パターンの中から切り出して独立したものであると考える必要がある。

自己組織化マップを用いて状態数を削減する方法として、我々と同様のアプローチが提案されている。Króse [27] らはセンサを用いてナビゲーションを行うために自己組織化マップを用いて環境の大域的な表現を形成する方法を提案した。彼らは、センサとして 16 個のレンジセンサを用いて実験を行ったが、シミュレーションのみの実験にとどまっている。

最後に自己組織化マップを用いることで発生する問題点について述べる。自己組織化マップでは画像全体を一つの入力ベクトルとして取り扱うため、環境に部分的に変化が生じた場合には、実際にはそれほど状況が変化していないにもかかわらず、異なった状況と認識されてしまうことである。状態空間の軸を明示的に決定するために、エージェントは行動に伴って現れる入力画像の変化に注目する必要がある。すなわち、画像の特徴もしくは構築された状態空間におけるある状態の重要度は行動を通じて重みづけをする必要がある。

第5章

接触信号に基づく視覚的特徴の選択と状態分割

前章で提案した手法では、画像全体を一つの入力ベクトルとして取り扱い、入力ベクトルの類似性に基づいてクラス分類しているため、タスク遂行のために注目しなければならない微小な特徴量の変化を見逃してしまうという問題がある。また、環境に部分的に変化が生じた場合には、実際にはそれほど状況が変化していないにもかかわらず、異なった状況と認識されてしまうという問題点がある。この問題に対するために、エージェントはタスク遂行のために適切な視覚的特徴を選択し、関係ないものは無視する必要がある。そこで本章では、タスク遂行のために適切な視覚的特徴の選択を問題として、特徴選択のために接触信号を用いた手法の提案を行う。

また、本手法ではマッピング表現方法として効用関数を用い、獲得のために強化学習を用いることを前提とするため、視覚的特徴の選択問題は、強化学習における状態空間分割の問題として位置付けることができる。強化学習を行うためには状態空間が適切に分割されている必要があるため、選択された特徴軸の分割問題も取り扱う。この分割問題に関しても接触信号を用いた手法を提案する。

5.1. 視覚と触覚の関係

視覚情報には色、色領域、エッジなど様々な特徴情報が含まれているが、それらの中には与えられたタスク遂行のために関係ないものが多く含まれている。エー

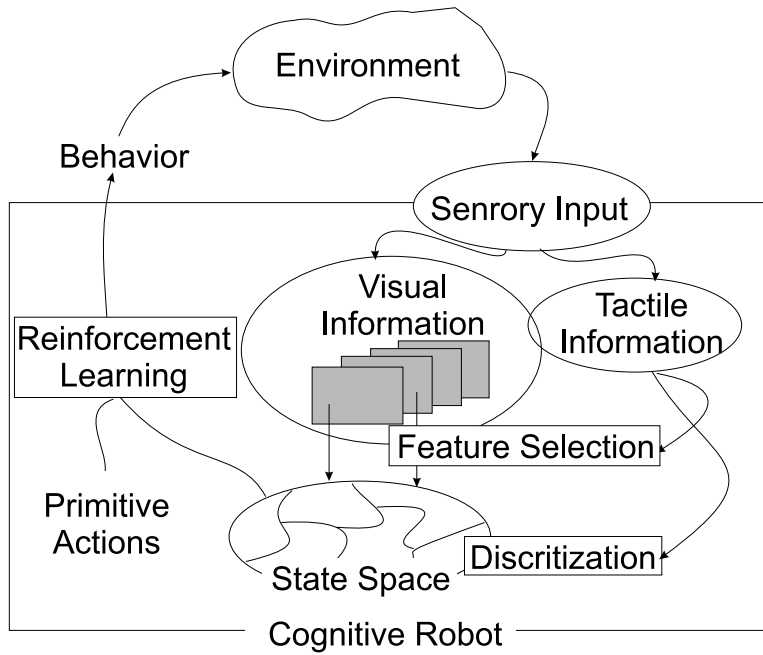


図 5.1 Cognitive Robot Architecture. A circle means a information and a rectangle means the processing of the information.

エージェントがこのような不必要な情報を除外せずに全ての情報を考慮に入れて行動することは計算コストの面で好ましくない。そこで、エージェントはタスク遂行のために必要十分な情報だけに注目する必要がある。本章では、このタスク遂行のために必要な特徴軸を選択するという問題に対して、接触信号を用いる方法を提案する。

2章では、人間の視覚を用いた認知機構において接触が重要であることを述べた。バークレーの主張に基づくと、視覚形成に関して次のことが言える。

1. 視覚情報には本質的に環境内の物理的な情報が含まれない
2. 見えの世界を作り出すためには対象に実際に接触するという経験が必要

彼の理論の中で、対象認識において対象の物理的属性を見分けることが重要であり、接触センサが対象との接触を知覚可能な唯一のセンサであることが強調されている。さらに、彼は視覚センサから入力された画像に対して物理的意味を与えることができる唯一のセンサであると述べている。従って、接触センサのみが視覚情報を特徴づけられるということである。視覚による対象認識において重要なことは物理的存在である対象と視覚情報を関連させることである。言い替えると対象が視覚入力中でどのように表現されているかを知ることである。また、エージェントが実世界を行動するときには、必ず接触が伴う。そこで本章では、エージェントのタスクを接触を終端とする行動として考え、行為の終端において知覚される接触情報をもとにタスク遂行のために必要な視覚的特徴を選択する方法を提案する。また、本手法ではマッピング表現方法として効用関数を用い、獲得のために強化学習を用いることを前提とするため、視覚的特徴の選択問題は、強化学習における状態空間分割の問題として位置付けることができる。強化学習を行うためには状態空間が適切に分割されている必要がある。そこで、選択された特徴軸の分割に対しても接触信号を用いた手法を提案する。

提案手法を図 5.1 のような認知エージェントのアーキテクチャによって説明する。このアーキテクチャでは接触センサが特徴選択と選択された特徴軸の分節のために重要な役割を果たす。ここで視覚的特徴軸とは、画像から抽出された領域情報などの特徴量をパラメータ化して表現するものである。エージェントは行動系列の終端から接触情報を受け取る。行動の終端で与えられる接触センサは対象

までの距離を表現するために、割り引かれて行動系列の各要素に割り当てられる。行動している間、画像系列もまた観測され、それぞれの特徴の変化が観測される。割り引かれた接触刺激はどの視覚的特徴が行動系列に関係しているかを判断するのに用いられる。エージェントが対象に向かって移動したときにある視覚的特徴量が変化したとするとその特徴が対象へ向かう行動を表現するのに適した特徴軸であると言える。さらには、対象を表現するのに適した特徴軸であると言える。また行動系列に対して与えられる割り引かれた接触信号を用いて特徴軸の分割が行われる。こうして対象についての行動を表現可能な状態空間が構築される。構築された状態空間を用いて強化学習が行われる。このアーキテクチャにおいて行われる処理について以下の節で詳しく述べる。

接触信号によってタスクを評価するということは、自律エージェント構築の立場における、マッピング獲得時に設計者の知識を減らすという趣旨に合致する。すなわち、完全に自律的なエージェントの立場からは、タスク達成に関する評価もエージェント自身によって行われるべきである。

5.2. 状態空間の分割問題

タスク遂行のために必要な特徴軸の選択と分割の問題は、しばしば強化学習における状態空間構築における問題として取り扱われる [31][9]。強化学習では、学習時にエージェントは何度も試行を重ねる必要があるため、状態数が増加すると、学習時間が指数関数的に増加し、膨大な学習時間を要する。しかし、一方で、状態の離散化を粗くし、状態数を少なくすると、異なった状態を同じ状態と知覚してしまう「知覚的見せかけ問題」が生じる。したがって、状態空間を適切に分割する必要がある。この適切な状態分割問題を解決する方法として、報酬を手がかりに状態の分割を行う方法が提案されている [19][51][47][9][34]。強化学習を用いてタスクを学習する際に、エージェントは報酬を与えられ、この報酬を用いて入出力関係を効用関数という形で学習する。報酬は、通常タスクの達成度に応じて与えられ、達成度にしたがって高い報酬が与えられる。強化学習の枠組では、エージェントは最大期待効用原理に従って期待効用が最大となるような行動を選択する。このため、そのような選択をするエージェントにとっては、次の状態に同じ

効用値をとるものがあれば、それらを同じ状態と考えるのが自然である。報酬を基準として状態空間の分割を行う方法は、この考えにもとづく。石黒ら [19] は主成分分析を用いて視覚的特徴を発見し遅れ報酬の分布に基づいて逐次的に状態空間の分割を行った。本手法で提案する割り引かれた接触信号も、石黒らの遅れ報酬と同じような働きをする。

特徴軸の選択と状態分割は同じ問題として取り扱われ、明示的に取り扱われない場合が多い。すなわち、あらかじめ準備した状態空間をある尺度にしたがって分割した際に、分割されなかった軸は結果的に、タスク遂行において不必要なものとなるからである。しかし、早い段階で特定の特徴軸のみを選択し、必要な特徴軸に対してのみ分割を行うことは計算コストの削減につながる。そこで本研究では、まず特徴軸の選択をし、その後に軸の分割を行うという方法を採用する。

5.3. 提案手法

本節では、接触信号に関連した視覚的特徴を画像の中から見付ける方法について述べる。また、発見された特徴軸を分割する方法についても述べる。

5.3.1 特徴軸の選択

我々が選択すべき視覚的特徴軸は、対象との接触状態を予測させるだけでなく、接触信号を受け取るまでの距離を表現できる必要がある。従って、対象そのものと対象までの距離の両方を表現できる視覚的特徴軸を発見することが重要な課題となる。重要なことは、対象の物理的属性に関する視覚的特徴の学習と距離の学習を同時に行なうということである。対象とそれを表す視覚的特徴軸との関係はエージェントが実際に対象に接触したときに獲得される。対象と身体との接触状態を見ずに接触信号が得られた瞬間の画像から対象を表現するための視覚的特徴を発見することは困難である。さらに、エージェントは視覚入力空間において身体がどのように見えるかを知っている必要がある。

特徴軸を発見するための基本的な着眼点は、距離の変化がある特定の特徴量の変化によって表されていることである。従って我々は、特徴軸の発見のために特徴量と行動系列の間の相関を利用する。 $\mathcal{F} = \{f_1, f_2, \dots, f_n\}$ を特徴軸の集

合とする．エージェントは，ある時刻 t において行動を選択できる．行動系列 $\{a_t, \dots, a_1, a_{0(end)}\}$ の最後において，強化学習の報酬のようにエージェントは接触刺激 s_0 を受け取る．そして，接触刺激は割り引かれて各タイムステップへ分配される． $\{s_t, \dots, s_1, s_0\}$ ここで， $s_t = 0.9^t$ である．今，ある特徴軸 f_m に注目し，各時刻において観測された特徴の値を $\{f_{(m,t)}, \dots, f_{(m,1)}, f_{(m,0)}\}$ とすると，割り引き接触信号と各時刻における特徴量との相関係数 r_m は次のように計算される．

$$r_m = \frac{\sum_{j=0}^t (f_{(m,t)} - \bar{f})(s_t - \bar{s})}{\sqrt{\sum_{j=0}^t (f_{(m,t)} - \bar{f})^2 \sum_{j=0}^t (s_t - \bar{s})^2}} \quad (5.1)$$

相関係数は全ての特徴軸に対して計算され，相関係数 r の値が $r_{threshold}$ を越えるものを対象を表現するための特徴軸として選択することにする．

5.3.2 軸の分割

前節では，特徴軸の選択方法に関して説明した．本節では，選択された特徴軸を適切に分割する方法について説明する．分割するために我々は，Kerber によって提案された ChiMerge アルゴリズム [23] を用いる．ChiMerge では， χ^2 分布を用いて離散的な属性の値の計算を行う．ChiMerge は初期状態とボトムアップな統合の過程からなる．統合過程では，終了条件が満たされるまで間隔は連続的に統合される．ChiMerge では，まず，適当な間隔で分割されたサンプルデータを属性の値にしたがってソートし，初期状態を作る．そして，それぞれのサンプルデータは各クラスに分類される．間隔の統合は次の 2 つのステップからなる．(1) 隣接する間隔間で χ^2 の値を計算する．(2) χ^2 の値が最も低い隣接する間隔は統合する．望ましい有意水準を選択することによって $\chi_{threshold}^2$ が決定される． χ^2 の値を得るためには自由度もまた決定する必要がある．自由度はクラス数より 1 少ない数である． χ^2 の値を計算する式は次のとおりである．

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \quad (5.2)$$

ここで

$m = 2$ (比較されるべき間隔の数)

$k =$ クラス数

object	color	physical property
ground	green	horizontal plane
wall	white	heavy, vertical plane
ball	red	light, globe

表 5.1 Relations among object, color and physical property

A_{ij} = number of examples in i 番目の間隔, j 番目のクラスにおけるデータ数

$R_i = i$ 番目の間隔におけるサンプルデータ数 . = $\sum_{j=1}^k A_{ij}$

$C_j = j$ 番目のクラスにおけるサンプルデータ数 . = $\sum_{i=1}^m A_{ij}$

$N =$ 全てのサンプルデータ数 . = $\sum_{j=1}^k C_j$

$E_{ij} = A_{ij} = \frac{R_i \times C_j}{N}$ の期待周波数

である .

以上をまとめると次のような手順となる

1. 環境内をランダムに行動して接触状態を終端とする行動系列を獲得する
2. 軸の各状態に対して接触信号を分配する
3. 視覚特徴と行動系列との相関を計算する
4. 相関係数の値の高いものを注目する特徴として採用する
5. ChiMerge によって軸の分割を行う

5.4. 実験結果

この手法の有効性を示すためにコンピュータシミュレーションによって実験を行った . エージェントのタスクは , 赤く塗分けられたボールに到達することである . エージェントは我々によって開発されたタイプ II のロボット [52] , 環境は RoboCup のフィールドを想定した . RoboCup のフィールドの特徴は , 対象の色がその物理的特徴によって塗分けられていることである . 表 5.1 は対象名と色 , 物理的な意味の関係をあらわしたものである .

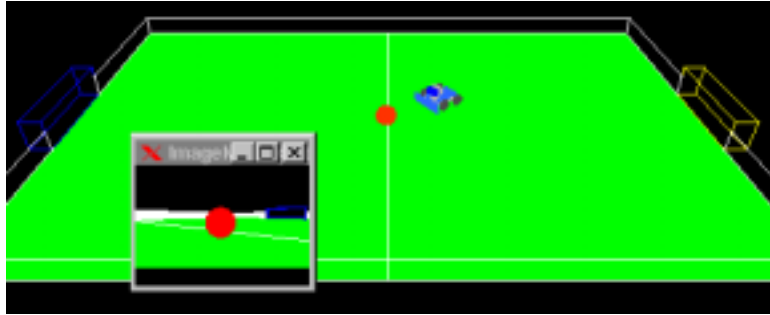


図 5.2 Simulator field and Captured image

図 5.2 はフィールド内のエージェントが見ている風景である．エージェントは a_1 から a_5 までの 5 つの行動を出力することができる．すなわち $a_1(0, 1)$, $a_2(0.5, 1)$, $a_3(1, 1)$, $a_4(1, 0.5)$, $a_5(1, 0)$ である．(図 5.3 参照).

エージェントにおける接触センサの分布と配置は，タスク遂行中においてエージェントが身体と物理環境との接触を完全に知覚するために重要である．従って，物理環境との接触が予想される全ての身体上の点に接触センサを配置する必要がある，接触センサの分布と配置に関する要件は次の 3 点である．(1) エージェントは 2 次元平面内のみを移動する．(2) フィールドの壁は垂直な平面である．(3) ボールの直径は h である．(1) (2) から接触センサが水平に分布する必要があることを意味する．(3) からは接触センサを配置する高さが h である必要があることがわかる．従って我々は，16 個のセンサをエージェントの前面に配置した．また，エージェントの能力として色領域の抽出が可能，かつその領域の重心座標と面積が計算可能であることを前提とした．

5.4.1 特徴軸の選択

我々は，あらかじめ画像処理を行い 9 個の特徴軸を用意した．それらは赤，緑，白の各色について領域の面積と重心を計算したものである．我々は， \mathcal{F} を次のように定義した．すなわち $f_1 = (\text{red } x)$, $f_2 = (\text{red } y)$, $f_3 = (\text{red } a)$, $f_4 = (\text{green } x)$, $f_5 = (\text{green } y)$, $f_6 = (\text{green } a)$, $f_7 = (\text{white } x)$, $f_8 = (\text{white } y)$, $f_9 = (\text{white } a)$. ここで， x, y は色領域の重心の座標， a は面積をあらわす．

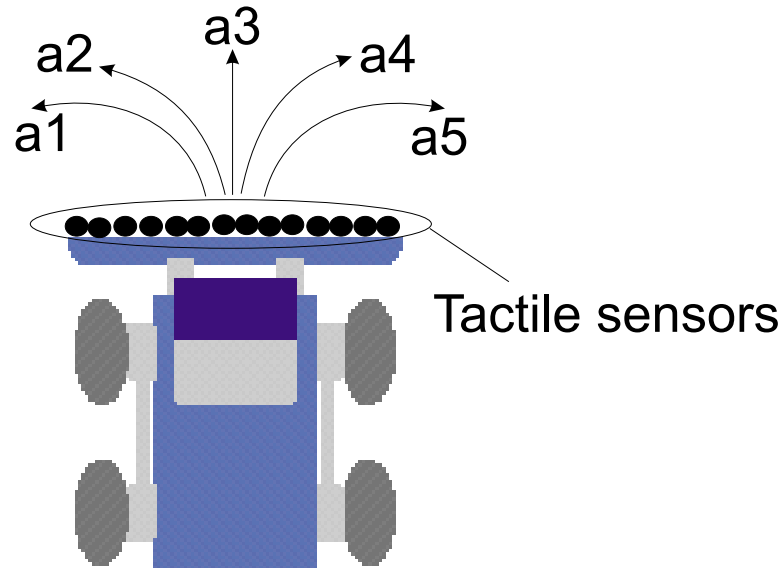


図 5.3 Robot's action and distribution of tactile sensors

f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
0.03	0.62	0.60	-0.12	-0.48	-0.39	-0.14	0.30	-0.47

表 5.2 The correlation coefficient rate

58,223 回のランダムな行動の結果 459 の接触を終端とする行動系列を獲得することができた。それぞれの行動系列に対して割り引かれた接触刺激 $\{s_t, \dots, s_1, s_0\}$ の計算を行った。ここで $s_t = 0.9^t$, また s_t は 10 のクラスに分割した; class1=[0 → 1), class2=[1 → 2) ... class10=[9 → 10]. 表 5.3 はそれぞれの \mathcal{F} の分散を示している。

表 5.2 は割り引き接触刺激と特徴量の間の特関係数を表している。

ボールに到達するタスクを遂行するのに必要な特徴軸として、特関係数の値が高い f_3 が選ばれた。 f_2 の特関係数も同様に高い値を示しているが分散が 2.9(表 5.3 参照) しかない。これは、 f_2 が行動系列を表現するのに適切でないということ

f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
271.8	2.9	25600.1	49.2	3.3	2498	8.0	0.3	14004.8

表 5.3 The variance of \mathcal{F}

を意味するので，ここでは， f_3 のみを必要な特徴軸として選択した．

5.4.2 特徴軸の分割

図 5.4 は特徴軸 3 の各間隔におけるサンプルデータのヒストグラムである．

図 5.4 の特徴軸の分割は ChiMerge において有意水準が 0.95 ($\chi^2 = 16.92$) の場合の最終結果を示している．図 5.4 中の矢印はそれぞれの間隔の境界を示している．

5.4.3 行動の学習

次に選択された特徴軸が適切であるかどうかを示すために，ボールに到達する行動を Q 学習を用いて獲得する実験を行った．

図 5.5 はボールの到達する行動系列である．図より，エージェントがボールに左から近付き，最後にボールに接触していることがわかる．

5.5. まとめ

本章では，我々は，エージェント自身が自ら内部表現を組織化するという認知ロボットにおける認知アーキテクチャを示し，その中で視覚的特徴軸の選択と軸の分割を接触信号を用いて行う方法を提案し，実験によってその有効性を確かめた．具体的には，行為の終端において知覚される接触情報を行為系列に対して割り引いて与え，この割り引き接触信号を用いて各特徴軸が行為達成に果たす寄与率を計算し特徴軸の選択を行った．また，選択された軸の分割には同様に，割り引き接触信号に対して 分布を計算し，それにもとづいて軸の分割を行った．実験の結果あらかじめ準備した 10 種類の特徴軸の中から到達行動のために必要な特徴軸ひとつが選択され，またもともと 80 状態あった状態が 14 状態にまとめら

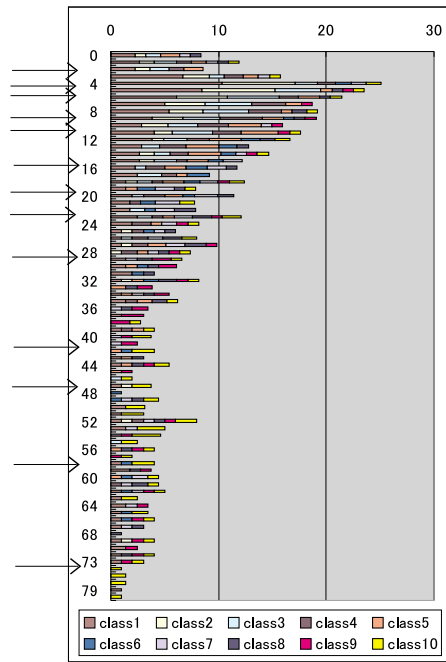


図 5.4 discounted tactile stimulus class histogram for feature 3

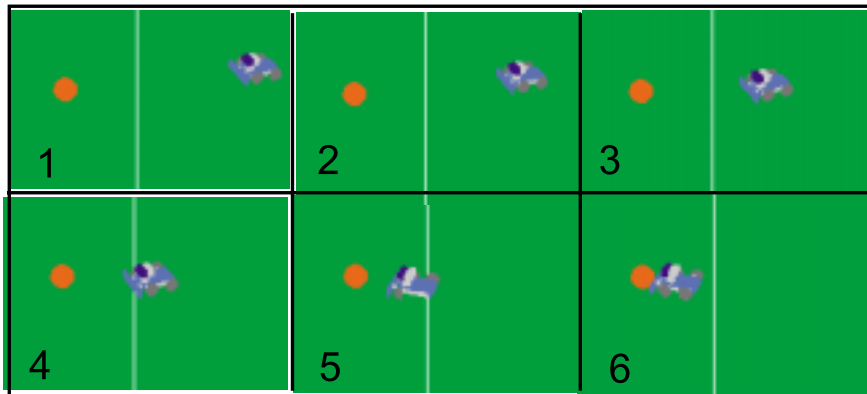


図 5.5 Behavior of reaching ball

Int	frequency										chi^2
0	40	13	11	4	10	0	3	2	0	2	73.1
3	281	81	12	8	2	2	4	0	0	3	37.4
5	147	89	25	1	2	1	0	1	1	1	65.7
6	170	72	106	32	8	1	0	3	1	2	42.9
9	53	25	36	32	16	1	2	1	2	3	45.6
11	97	13	42	40	77	20	7	4	6	4	47.9
16	37	5	14	9	18	29	14	5	2	5	29.1
20	14	3	7	12	10	7	24	13	2	4	20.0
23	27	5	6	14	13	4	13	25	13	12	19.9
29	22	4	2	8	12	7	8	23	39	17	27.9
42	4	0	2	0	6	2	0	2	4	14	19.5
47	13	2	1	3	2	6	7	4	12	59	23.9
59	21	2	1	1	7	5	8	11	14	23	22.6
74	1	0	0	1	1	0	0	0	0	12	

表 5.4 ChiMerge discretizations for feature 3

れた．この軸によって構築される状態空間を用いて強化学習を行うことによって対象の物理的な属性に対して適切な行動が獲得され，エージェントは視覚を用いて行動することができた．

実験結果から次のことが言える．エージェントは対象の視覚的特徴である色と物理的属性の関係を自律的に獲得することができた．また，エージェントは行動と接触を通じた経験によってどの視覚的特徴に注目すべきかを学習することができた．またこの結果は，エージェントは視覚の特徴に関してあらかじめ情報を与えないでも行動を獲得することができることを示すものである．すなわち提案したアーキテクチャによって，環境に関する情報はほとんど事前に定義することなしに接触情報のみにもとづいて内部表現を構築することができたと言える．その結果，エージェントは環境の変化に対して柔軟に適応することができる．

エージェントが実空間における対象を認識するためにはエージェントは様々な視覚的特徴を知る必要がある．視覚的特徴は対象の物理的特徴に非常に強く関連しているため接触センサと視覚入力の間関係を獲得することは重要である．本手法においてはエージェントの持つ接触センサの数は限られていたが，より多くの種類の対象を識別するためにはエージェントは多くの接触センサを用いて対象に接する必要がある．

第6章

接触に基づく効用関数を用いた行動生成と対象表現

先に説明した2つの研究では、タスク遂行のための視覚入力からの状態空間の構築に焦点を当てていたため、対象の記述そのものに関しては言及していなかった。そこで、対象をより一般的に表現するための内部表現を構築することを目標として、触覚に基づく効用関数を用いた対象認識と行動生成の研究を行った。内部表現は環境との相互作用によって構築され、特に行為の終端などを表す接触信号がマッピング獲得に重要な役割を果たす。そこで、エージェントの身体性と身体と環境との物理的相互作用を知覚する接触センサに焦点を当て視覚入力に対して接触信号に基づいて効用関数を生成することで対象の表現を行った。

バークレーの視覚新論 [5] によると視覚による対象（環境）認識とは、視覚から接触行動を予測することである。ここでは、対象は身体との間に生じる行動によって相対的に記述され、視覚は相対的關係を予測する役割を果たす。この關係を一般化して記述するものが効用関数である。画像から効用関数を生成するための基本的なアイデアは画像を微小領域に分割し、それぞれの領域について独立に効用値を計算することである。独立に計算された効用値を足しあわせることで、対象全体に関する効用関数を生成することができる。これによって、画像中の任意の対象について効用関数を生成することができる。また、対象が効用関数で記述されているとエージェントは（最大期待効用原理を用いて）対象についてのあらゆる行動を生成することができる。実験の結果、エージェントは視覚入力から効用関数を生成し、仮想的ななぞり行動によって6種類の対象の識別を行うことができた。また、効用関数を用いて、通り抜けられるすきまと通り抜けられないす

きまを識別し適切な行動を行うことができた。

本章の構成は以下のとおりである。2節では対象表現を行うための効用関数を視覚入力から生成するための方法の概略について述べる。3節では効用関数生成のための具体的な方法について述べる。4節では行動生成時に身体性を考慮することと効用関数の関係について述べる。5節では提案手法に基づいた行動生成の実験について述べる。6節では提案手法を応用して身体性に基づく対象の認識と判別について述べ、最後に本研究のまとめを行う。

6.1. 視覚入力と効用関数

本章におけるエージェントの視覚の役割は、接触を予測すること、すなわち現在の視覚入力から適切な接触状態へ到達経路を発見することである。もしエージェントが視覚入力から効用関数を生成することができれば、期待効用を最大にする行動を選択するという、最大期待効用の原理に基づいてエージェントは接触へ到達するための最適な行動を生成することができる。そこで、視覚入力から効用関数を生成することで、対象への触れ方を予測することを考える。視覚入力に対して効用関数を生成する方法として、強化学習が良く知られている [50][19][37]。強化学習を用いた手法では、設計者はエージェントの身体性や環境の構造を考慮する必要がなく、ゴール状態を与えるだけでエージェントは自ら環境を探索し、タスクを遂行するために最適な効用関数を生成する。従来の手法は特定のタスク遂行のために効用関数を生成するためのものであるが、本研究における効用関数の役割は、接触を予測することであり、目的が異なるのでこれらの手法を用いると様々な問題が生じる。以下にそれらの問題と我々の目的に従った効用関数の生成方法について議論を行う。

6.2. 提案手法の概要

具体的な議論に入る前にまず、提案手法の概要を示す。図 6.1は提案手法の概要を図式化したものである。本手法の目的は視覚入力から環境やタスクに依存した効用関数を生成することである。効用関数の生成時に考慮すべき点は以下の

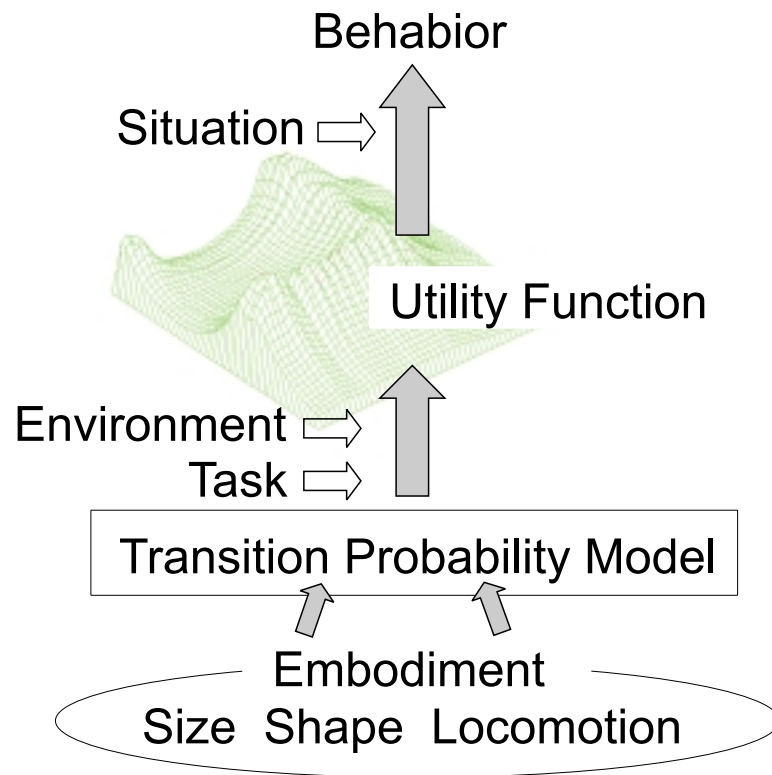


図 6.1 Overview of our method

3点である。

1. 状態遷移確率モデル (以下では単に遷移モデルと呼ぶ) 持つ。すなわち自身の行動の結果が視覚入力でどのような変化となって現れるかを事前に知っているということである。この遷移モデルは身体性に依存して変化する。保持した遷移モデルを用いて、環境とタスクの変化に対応して効用関数の生成を行う。
2. ゴール状態を接触状態として与える。タスクを身体性と関連させて記述するために、実世界における全てのタスクを接触を基準として考える。
3. 視覚入力を微小領域の集合として取り扱う。視覚から抽象化した情報を取り出して入力とするとタスクや環境に依存した恣意的なものとなるので、視覚入力に対しては簡単な処理のみを行う。

以下にそれぞれについて詳述する。

6.2.1 遷移モデルを用いた効用値の更新

効用関数の獲得時に最初に考慮すべき点は遷移モデルを用いるかどうかである。遷移モデルを用いない方法としては、TD法 [39] や Q 学習 [41] が知られる。遷移モデルを用いる方法としては、予め獲得された遷移モデルを用いる Value Iteration [4] や Policy Iteration, 環境を探索しながら遷移モデルと効用値を同時に更新する Dyna [40] が知られる。強化学習の多くは、未知環境において効用関数を求めることを問題としているため、環境を探索しながら効用値を更新すれば良いので、遷移モデル自体を獲得する必要はない。ここで、環境が未知であるということについて考える。強化学習を用いた研究の多くでは環境は未知であると仮定しているが、生物において環境が全くの未知であるのは生後のある一定の期間だけである。物に触れるとか障害物を避けるなどの接触を基準とする基本的なタスクは一旦獲得されると、どんな環境においても再利用することができる。より抽象的なタスクはこの基本的なタスクをもとにして遂行されると考えることができる。基本的なタスクを遂行するための効用関数の獲得時において、環境が変化する度にその都度環境を探索しては非常に効率が悪い。そこで遷移モデル

を保持しておき、環境が変化した際には遷移モデルを用いて環境やタスクに応じて Value Iteration を用いて効用関数を生成することを考える。

一般に遷移モデルは、行動出力に対する知覚の変化をあらわす。モデルはカメラが環境に固定されている (以下グローバルビジョンと呼ぶ) か、エージェントの身体に固定されている (以下ローカルビジョンと呼ぶ) によって、遷移モデルの意味が異なる。すなわちグローバルビジョンでは、環境が静的な場合、遷移モデルがあわらすのは自己身体の見え方の変化である。一方、ローカルビジョンではエージェントが行動を行っても自己身体の見え方は変化しない。変化するのは環境の見え方なので、遷移モデルは環境の変化をあらわす。しかしどちらの場合も、接触を予測する効用関数を生成出来ると言う点で差はない。

後述するように、視覚入力中でモデルは微小領域内における、身体や環境の見え方の変化としてあらわされる。Value Iteration を用いて効用関数を生成するためには、全ての微小領域についてモデルを持つ必要がある。

6.2.2 報酬の与え方

次に考慮すべき点は、実世界におけるあらゆるタスクを接触を基準としてとらえることである。すなわち、タスクを「接触状態への到達」と「接触状態の回避」の組合せとして考える。強化学習では、普通タスクは報酬として与えられるが、タスクに合わせて設計者が適当に設定する必要がある。

実世界におけるエージェントのタスクはほとんどの場合環境との直接的な接触を伴う。例えば、ロボットサッカーにおけるシュート行動では適切な姿勢からボールにアプローチして、接触を行なう必要がある。また、ロボットアームにおける対象の把持動作においては、ハンドを適切な姿勢から対象に近づけ、適切な形で接触させる必要がある。したがって、実世界におけるタスクは身体のある部分を対象のある部分に接触させることであると言える。また、多くの場合ゴール状態に到達するまでに接触を回避することが要求される。そこで我々は、接触状態をゴール状態と考え、接触が予測される状態に報酬を与えることで効用関数を生成する。エージェントが各状態におけるモデルを持っていると、画像入力の任意の形状に対して、対象の境界部分に報酬を置くことによって Value Iteration を用いて効用関数を計算できる。

6.2.3 視覚入力への扱い方

三番目に考慮すべき点は視覚入力に対して出来るだけ情報抽出を行わないと言うことである。そのために我々は、画像入力を微小な領域の集合と考える。効用関数は、見え方に対する効用値を出力するが、この関係はもともと未知であるのでエージェントは学習によって獲得する必要がある。このときに見え方の変化を適切に表現するような状態空間を規定しておかなければ、タスクを遂行するために適切な効用関数を獲得することはできない。視覚入力から状態空間を作るときに、あらかじめ特定のタスクに依存して状態空間を用意しておく方法がある。浅田ら [50] はサッカーロボットのシュート行動の学習において、視覚入力から抽出されたボールとゴールの大きさを適当に離散化した状態空間を用いた。しかし、設計者が予め設定した状態空間がタスク遂行のために適しているかどうかの保証はない。そこで、エージェントが自ら状態空間を規定する方法が提案されている。Ishiguro ら [19] はタスクとして与えられる報酬の分布に従って状態空間の分節を行った。この手法では、タスク遂行のために必要な状態空間をエージェント自らが構築するものであるが、直接の視覚入力を扱うのではなく、主成分分析を行うなど統計的手法を用いてあらかじめ入力の次元を落している。このような手法は、膨大な情報を含む視覚入力からタスク遂行のために必要十分な情報を抽出するという注視点の抽出問題と複合されている。柴田ら [37] は Direct Vision based Reinforcement Learning を提案し、視覚入力を直接ニューラルネットワークに入力し状態空間を自動的に生成する方法を提案した。特定の目的の下で画像処理を行って対象のパラメータなどを抽出し状態空間を構築する方法では、タスクや環境に依存した状態空間になってしまうため、我々は画像入力を、なるべくタスクに依存しないように直接取り扱う方法を選択する。

6.3. 効用関数の生成方法

以下の節では、前節で述べた方針に基づいた具体的な方法について説明を行なう。

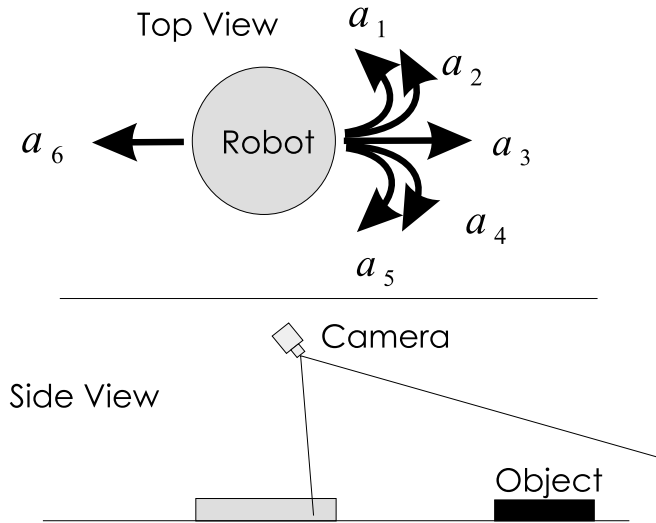


図 6.2 Assumed environment and agent

6.3.1 前提とするエージェントと環境

以下の説明では、視覚入力を単純化するために、エージェントとエージェントの存在する環境は2次元平面であるとし、エージェントは2次元平面内を移動するものとする(図 6.2参照)。エージェントには、環境との接触を知覚するための接触センサが備えられている。また、身体と環境が接触する状態を観察するために、身体と対象の両方が視野に入る位置にカメラを備えている。カメラは、グローバルビジョンでもローカルビジョンでもどちらでもよいが、以下の説明では、グローバルビジョンを仮定する。エージェントの行動は、 a_1 から a_6 の6個の行動とし、行動の単位は観測する状態が変化するまでとする。

6.3.2 効用関数の計算

図 6.3のような入力画像が得られたとする。入力画像に関しては次の仮定を用いる。

1. 画像は縦横に適当に離散化され微小領域に分割されている。
2. エージェントは微小領域内に対象もしくは身体の境界が存在しているかど

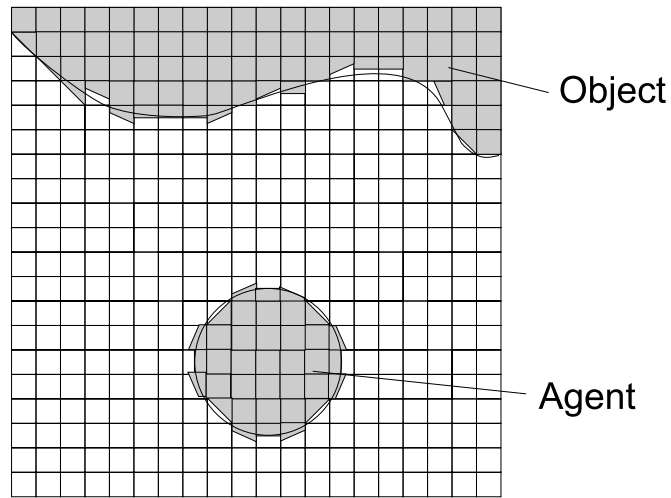


図 6.3 Input image

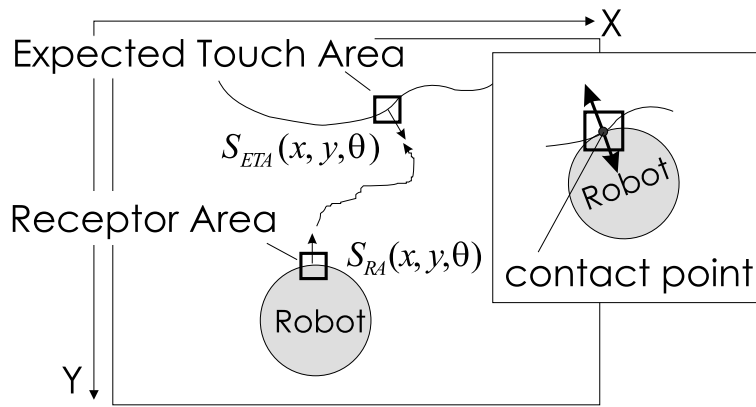


図 6.4 Expected touch area, receptor area, and reaching action path

うかの判断が可能である．微小領域内に存在する境界は角度方向に適当に離散化される．

3. エージェントは微小領域内の対象が自分の身体かどうかの判断が可能である．

ここで，対象の境界が含まれる微小領域を Expected Touch Area(以下 ETA) と呼ぶことにする．また，身体の境界が含まれる微小領域を Receptor Area(以下 RA) と呼ぶ．ETA と RA の状態は，画像内の位置 x, y と微小領域内の境界の法線ベクトルの傾き θ を用いて $s(x, y, \theta)$ として表現する．

現在の視覚入力に対して生成される効用関数は，RA が ETA に接触することを予測するための状態評価を行うので，ETA に報酬を置くことによって計算される．各状態における効用値は動的計画法の基礎として知られる次の式を解くことで求められる．

$$U(s) = R(s) + \max_a \sum_{s'} M_{ss'}^a U(s') \quad (6.1)$$

ここで，記号 $M_{ss'}^a$ は動作 a が状態 s においてなされたときに状態 s' に到達する確率を表す．グローバルビジョンでは，行動 a を出力すると RA が変化するので，モデル M は RA に対して求める．モデル $M_{ss'}^a$ は，同じ状態から何度も同じ行動を行なうことによって計算される．すなわち

$$M_{ss'}^a = \frac{n_{s'}}{n_s^a} \quad (6.2)$$

によって計算される．ここで n_s^a は，状態 s において動作 a を行なった回数， $n_{s'}$ はその結果 s' に到達した回数である． $R(s)$ は状態 s において与えられる報酬である．本手法では RA が ETA に接触する状態に対して $R(s)$ を与える．そのときの接触条件は，

$$(x, y)_{ETA} = (x, y)_{RA} \wedge \theta_{ETA} = -\theta_{RA} \quad (6.3)$$

である．本手法では，各状態における効用値を計算するために隣接する状態で式 (6.1) を何度も適用する Value Iteration を用いる．Value Iteration では各段階で隣接する前の状態の効用値をもとにして各状態の効用値の更新を行う．これは次の式で表される．

$$U_{t+1}(s) \leftarrow R(s) + \max_a \sum_{s'} M_{ss'}^a U_t(s') \quad (6.4)$$

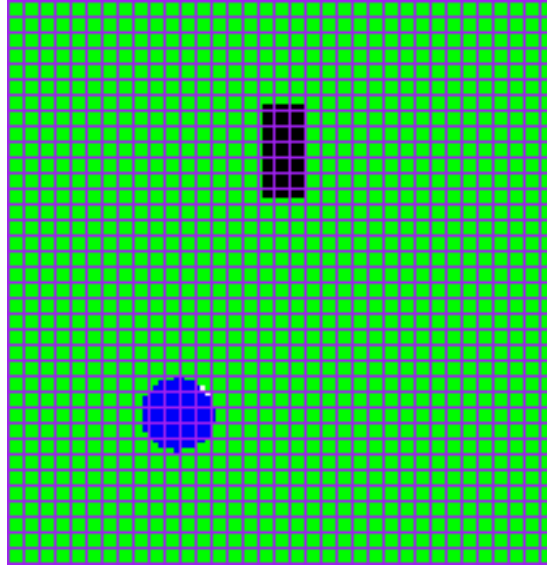


図 6.5 An example of input image

ここで、 $U_t(s)$ は t 回反復後の状態 s の効用値を表す。 $t \rightarrow \infty$ とすると、効用値は与えられた条件下で安定した値に収束することが知られる。

マルコフ決定過程において効用関数 U が与えられると、最適な動作はその結果の状態の期待効用が最大となるものとして決められ、状態 i におけるエージェントの政策 f は次のようになる。

$$f(i) = \arg \max_a \sum_j M_{ij}^a U(j) \quad (6.5)$$

ここで、 $\arg \max_a f(a)$ は $f(a)$ の値が最大となるような a の値を返す。エージェントはこの政策を用いることによって、実際に動くことなく仮想的に最適な接触行動を生成することができる。

6.3.3 効用関数の重ね合わせ

先に述べた方法で求められるのは、一つの ETA に関する効用関数である。ETA は通常画像から複数検出されるので、複数の ETA に関する効用関数を求める必要がある。そのために効用関数の重ね合わせを考える。画像から n 個の ETA が

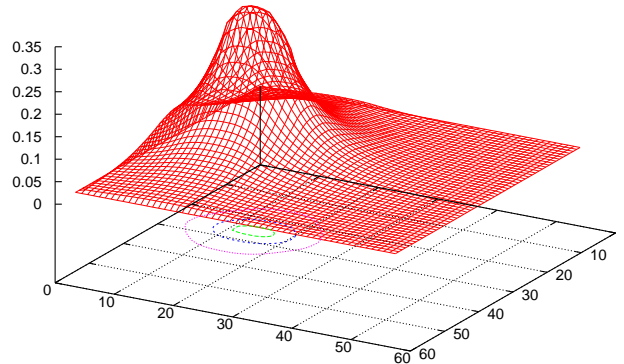


図 6.6 An example of utility function

観測されたとき，各状態の効用値はそれぞれの ETA における効用値の中の最大値をとるように決める．これは次式で表される．

$$U(s) = \max U_i(s); U_i(s) \in U_1, \dots, U_n \quad (6.6)$$

このように，それぞれの ETA における効用関数を重ね合わせることで任意の形状の対象についての効用関数を生成することができる．

図 6.5, 図 6.6は，環境に矩形の対象が置かれた場合の入力画像と，それに対する効用関数の例である．この図では，角度方向の状態に関して平均をとっており， $x - y$ 方向のみに関する表示となっている．以下の図においても同様の表現方法を採用する．

6.4. 身体性を考慮した行動生成

前節では，身体上のある 1 点のみについて任意の形状の対象に依存した効用関数を生成する方法について説明した．本節では，タスク遂行のために，エージェントの身体の大きさと形状を考慮に入れて状態判断を行なうことを考える．

6.4.1 身体性を表現するための効用関数

エージェントは固有の身体の大きさと形状を持つので、タスク遂行時には常に身体表面と環境の対象の表面の関係を考慮しながら行動する必要がある。具体的には以下の2点である。

1. 身体の各部において状態遷移が異なる。身体が広がりを持っていると、同一の行動を行なった際に身体の各部は異なった状態遷移をする。したがって、身体の各部で異なったモデルを保持する必要がある。
2. 身体の各部において効用が異なる。本研究における効用関数の役割は、接触状態を基準として、状態の価値を見積もることである。すなわち、効用関数は接触への近さを出力する。身体が広がりを持っていると、身体の各部で、接触状態への近さが異なり、したがって状態価値が異なる。

上記を考慮して、エージェントの状態を次のように定める。入力画像から、エージェントの身体上の点、すなわち RA が n 個観測されたとする。各 RA_i において観測される状態 $s_i(x, y, \theta)$ はモデル M_i に従って遷移する。画像から対象が発見されると、前節で述べた方法により、各 RA_i に対して効用関数 U_i を生成する。この効用関数によって計算される状態価値を v_i とすると、エージェントの状態は効用関数の集合により

$$V = v_1, v_2, \dots, v_n \quad (6.7)$$

のように定義される。

6.4.2 タスク遂行のための状態判断

先に述べたように我々は実世界におけるエージェントのタスクを接触を基準として考える。具体的には、実世界におけるタスクを「接触の回避」と「接触状態への到達」として考え、先に説明した効用関数を用いてこれらを実現することを考える。

6.4.3 接触状態への到達

接触状態へ到達するという事は、身体のある点を環境内の対象のある点に接触させることである。到達行動は、画像から検出されたある ETA から、身体上の接触させたい RA におけるモデルを用いて計算された効用関数を用いて式 (6.5) の政策にしたがうことによって生成できる。このときの政策を f_{goal} とする。

6.4.4 行動生成

エージェントの身体が環境に接触しないための条件は、身体上の各 RA_i における効用関数 U_i によって計算される全ての状態の値 v_i が閾値を越えない、すなわち、

$$v_i < v_{threshold}; i = 1 \cdots n \quad (6.8)$$

となることである。

目的を持たない接触の回避は存在しないので、ここでは目的遂行のもとでの接触回避について考える。目的遂行のための政策が f_{goal} であるとき、接触しないで目的を達成するためには、政策決定時に選択された行動をとったときの各 RA における期待値が、式 (6.8) を満たす必要がある。

6.5. 障害物回避

上述した提案手法の有効性を検証するために、2つの課題を設定して、コンピュータシミュレーションによって実験を行った。まず一つ目は行動生成で、エージェントのタスクは障害物を回避して目的地に到達することである。二つ目は対象の認識と判別で、身体を用いた対象認識を行うためにこれまでに述べた手法を応用したものである。

本節では、効用関数を用いて身体性を考慮した行動生成について述べる。実験には、図 6.7 のようにグローバルビジョンとローカルビジョンの2つのカメラモデルを用いた。図 6.7(a) では、カメラは地面に対して垂直を環境に固定されている。図 6.7(b) では、カメラは斜めにエージェントに固定されている。実験に用いたエージェントの身体の形状は円形とし、直径は3とする。画像のサイズは

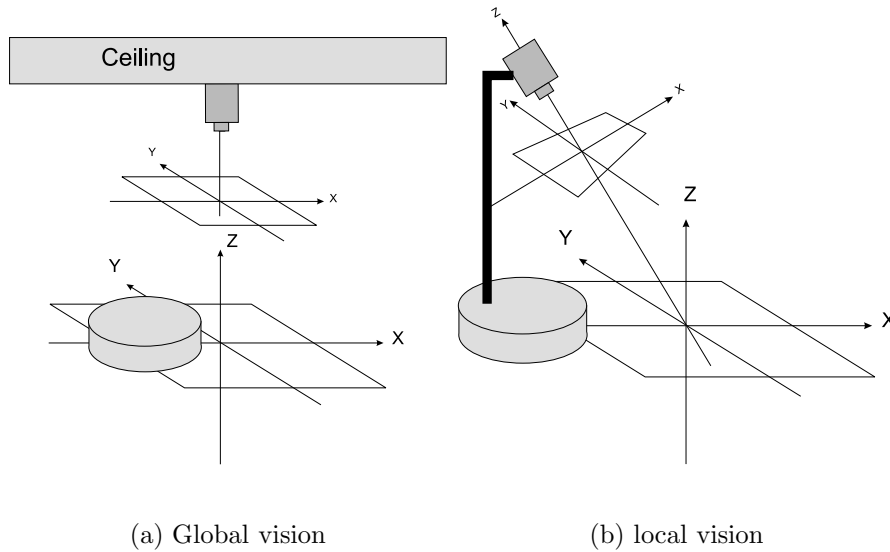


図 6.7 Camera model used in our experiment

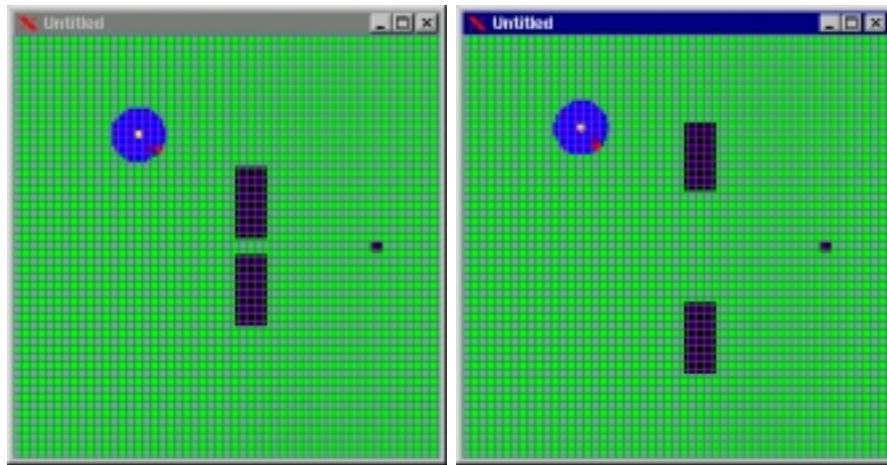
150 × 150 であり，3 ピクセルずつ 50 × 50 の微小領域に離散化されている．微小領域内で観測される境界の傾きは，−157 ~ 180 度の範囲で 16 段階に離散化されている．したがって観測される状態の全ての数は，50 × 50 × 16，すなわち 40,000 状態である．

6.5.1 状態遷移確率モデルの獲得

任意の対象についての効用関数を生成するための準備としてモデルの獲得を行った．モデルは観測される 40,000 全ての状態について，各動作 $a_1 \cdots a_6$ を行ない， n_s^a が 100 を超えるまで行動を行なうことによって獲得した．

6.5.2 グローバルビジョンによる通り抜けタスク

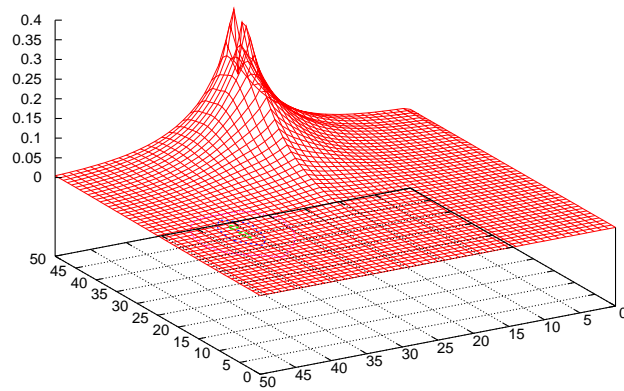
次に，身体性を考慮した行動生成の例として通り抜けタスクを行った．目的地に到達するタスクにおいて，その最短到達経路上に大きさ 3×6.4 の矩形の対象を 2 つ並べて配置する．対象間の間隔は，狭い場合には 2 に，広い場合は 10 とした．エージェントの直径が 3 なので，広い場合すき間を通り抜けることができ，



(a) Narrow case

(b) Wide case

☒ 6.8 Input images used in our experiment



☒ 6.9 Utility function for reaching a goal

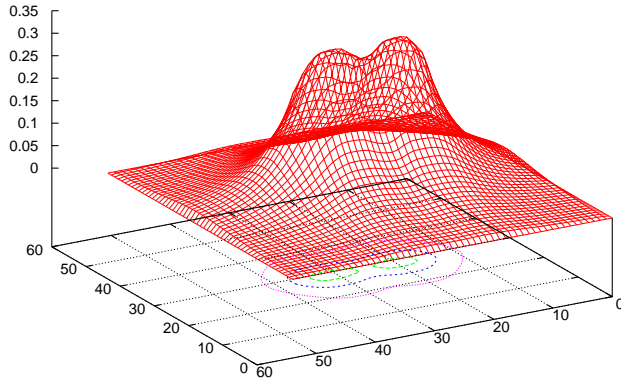


図 6.10 Utility function of narrow case

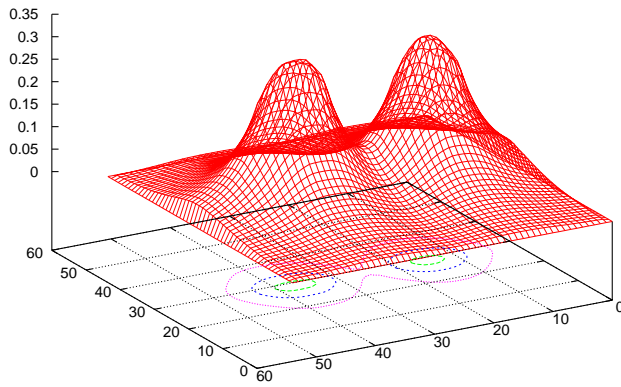


図 6.11 Utility function of wide case

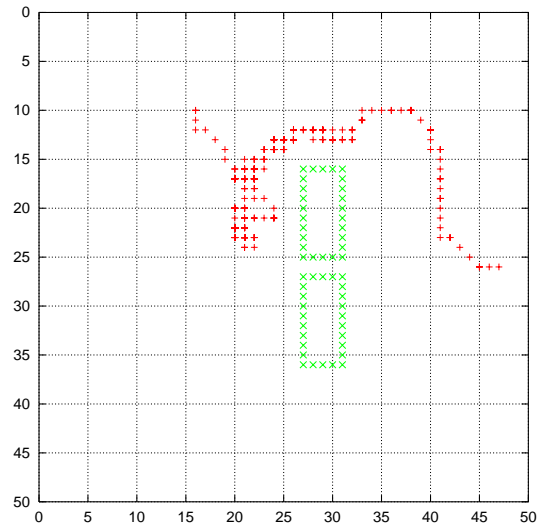


図 6.12 A generated path of narrow case

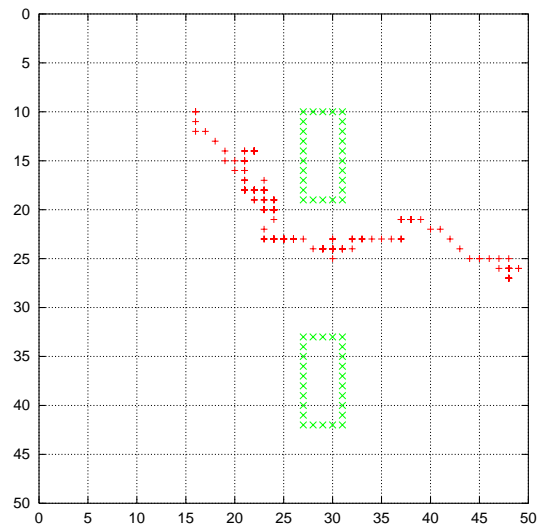


図 6.13 A generated path of wide case

狭い場合には通り抜けられずに回避する必要がある．ふたつの場合のそれぞれの入力画像を図 6.8(a)(b) に示す．両図中左上にある円がエージェントを表し，スタート地点もこの位置とする．入力画像から，それぞれの対象に関して 18 個，計 36 個の ETA が検出された．また RA については 20 個検出された．ゴール状態は，図 6.8 中の右中央の小さな対象に，身体正面の RA が左から接することとして与える．ゴール状態における RA の画像上での状態は，(47, 26, 0) である．

ゴールに到達するための効用関数は図 6.9 のようになる．障害物回避のための効用関数を，狭い場合について図 6.10，広い場合について図 6.11 に示す．それぞれの効用関数は 36 個の ETA について重ね合わせることによって生成された．また，この効用関数は正面の RA において計算されたものである．各 RA について接触状態を評価する必要があるので，同様の効用関数をそれぞれの RA に対して生成した．

ゴールに到達するための効用関数と，対象から生成された効用関数を用いて生成された行動の経路を，狭い場合について図 6.12，広い場合について図 6.13 に示す．図から，狭い場合には二つの対象を回避しゴールに到達する経路が生成されたことが分かる．また，広い場合には二つの対象の間を通り抜けてゴールに到達する経路が生成されたことがわかる．

6.5.3 ローカルビジョンによる障害物回避

次に，提案手法がグローバルビジョンだけでなくローカルビジョンにも適用可能であることを示すために，ローカルビジョンを用いて障害物を回避しながらゴールに到達する実験を行った．ローカルビジョンでは，グローバルビジョンとは異なり，行動出力に対して状態が変化するのは ETA である．したがって，モデルは ETA のにおける状態遷移確率をあらわす．図 6.14 に入力画像の例を示す．中央下の円がエージェントの身体であり，左上の長方形が障害物である．また中央上の点がゴールである．カメラは図 6.7(b) のように真下ではなく斜め前方下を向いているため，行動出力に対する画像の変化は遠く (画像中では上部) ほど小さい．本来ならこの点を考慮して一様な状態の分割でなく，状態の変化を適切に表現出来るような分割をするべきであるが，ここでは簡単化のため一様な分割とした．検出された RA は計 42 個であり，それらに報酬を与えて計算した効用関数を

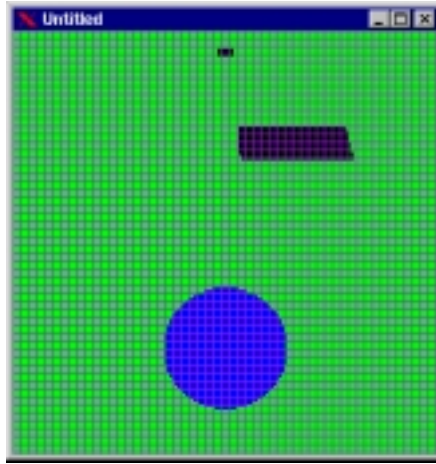


図 6.14 Input image of local vision

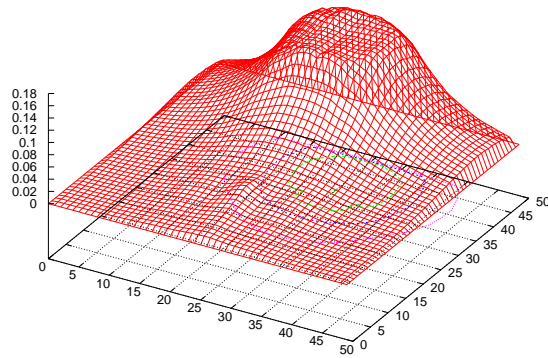


図 6.15 Utility function of local vision

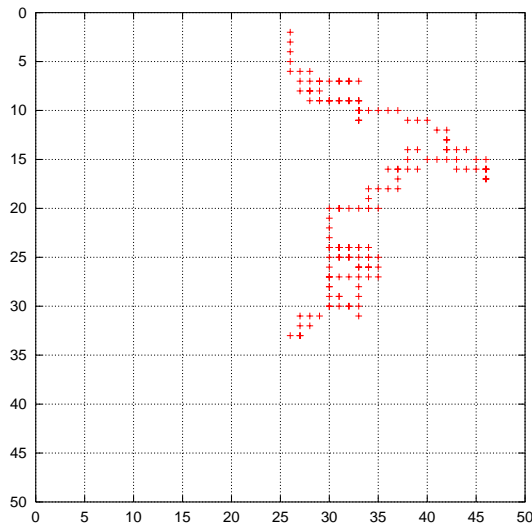


図 6.16 A generated path of avoidance

図 6.15に示す．タスクは，障害物を回避しながら身体の正面の点 $(26, 33, -90)$ をゴール地点 $(26, 2, -90)$ 移動させることとした．図 6.16に生成された行動の軌跡を示す．軌跡はゴール地点の画像中での座標の系列を示している．図より，エージェントが障害物を左に回避してゴールに到達していることがわかる．

6.6. 身体性の変化と行動出力に関する実験

次に、同一のタスクを行う際に、エージェントの身体性が異なった場合の行動出力の違いを調べる実験を行った．実験はグローバルビジョンを用いて行った．図 6.17(a),(b),(c) に本実験で用いたエージェントの形状とモータの配置を示す．エージェント A は円形をしている．エージェント B,C は，ともに形状大きさは同じであり，モータの配置のみ異なる．また $p00$ を正面とする．agent B は身体の前方にモータを配置し，agent C は身体の後方にモータを配置した．タスクはともに，身体上の点 $p00$ を画像中で $(33, 8, -90)$ に移動させることであり，その途中で矩形の障害物を置く．図 6.19は agent B の場合に生成された経路である．エージェントは図中の左上から右下に向かって移動する．図 6.19(a) は障害物を

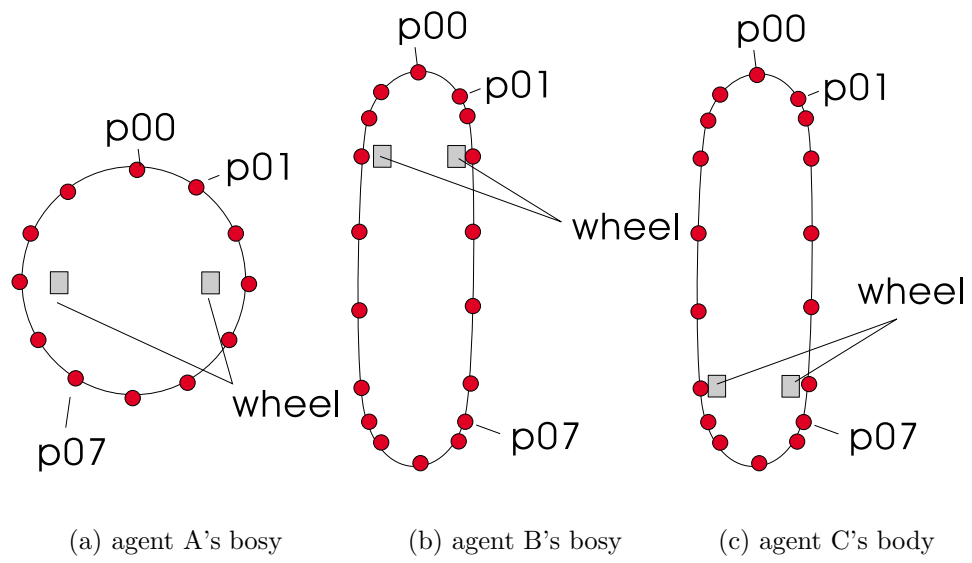
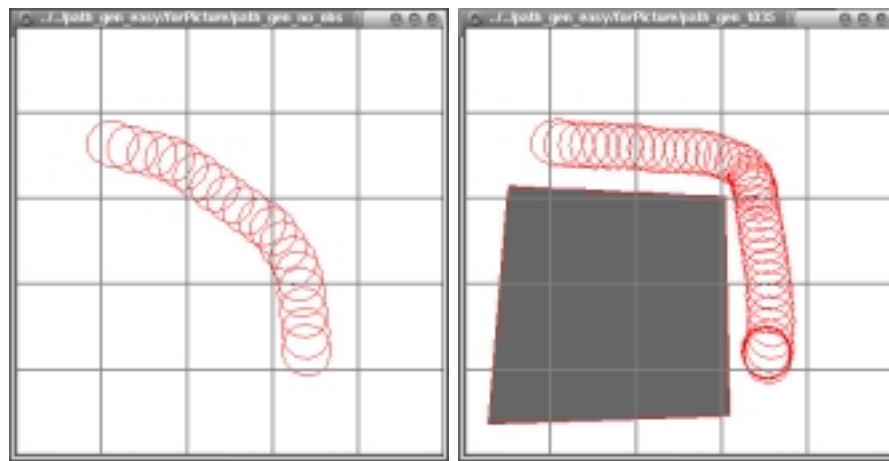


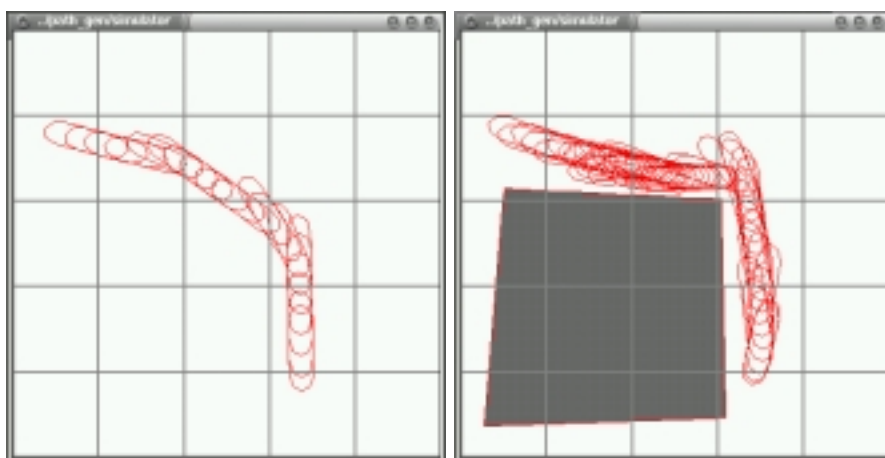
図 6.17 Agents used in this experiment



(a) Without obstacle

(b) With obstacle

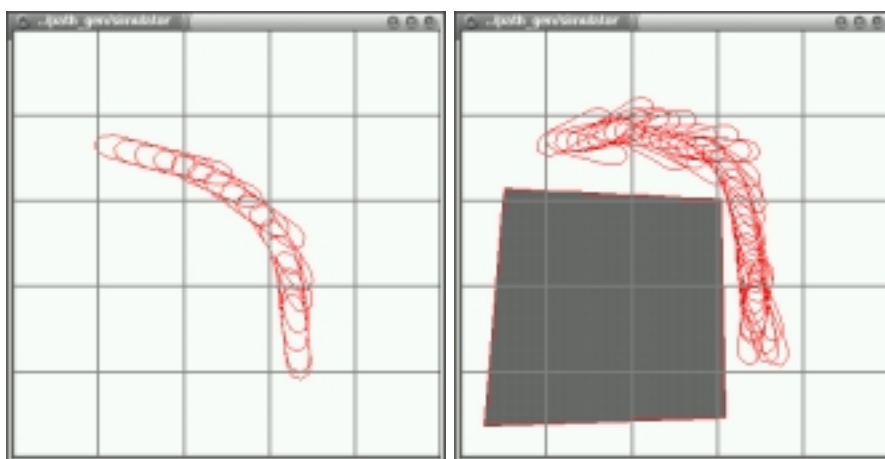
図 6.18 Generated paths of agent A



(a) Without obstacle

(b) With obstacle

図 6.19 Generated paths of agent B



(a) Without obstacle

(b) With obstacle

図 6.20 Generated paths of agent C

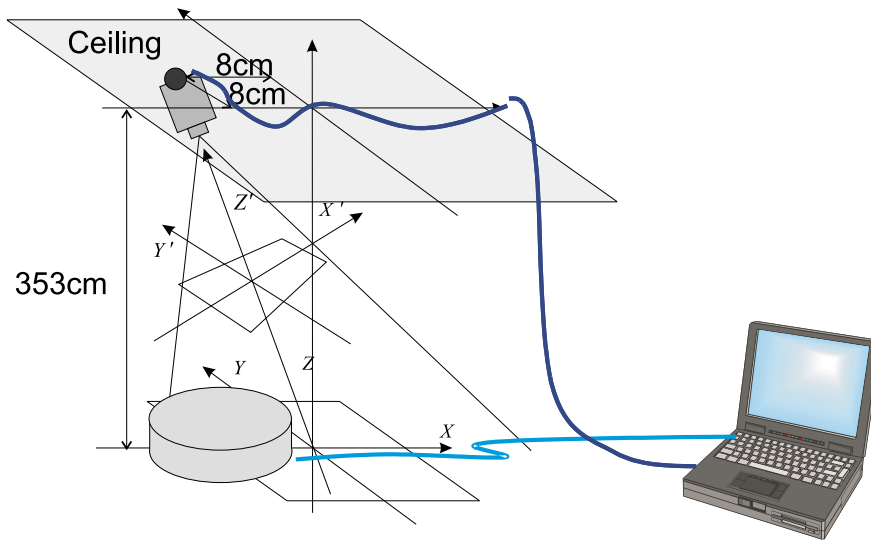


図 6.22 The robot and environment used in real world experiment

置かない場合に生成された経路である．図 6.19(b) の左下にある矩形領域は障害物である．形状が歪んでいるのは，カメラを地面に対して傾けて設置したためである．図 6.19(b) から，エージェントは障害物に沿って移動し，障害物の角の部分ではほぼ直角に曲がっていることがわかる．

図 6.20は agent C の場合に生成された経路である．図 6.20(b) から，エージェントは角を直角に曲がらずにゆるやかに曲がっていることが分かる．これは，モータを身体の後方に配置しているため，障害物の左側面に沿って進入するためには，大きくまわらなければならないからである．

図 6.21に agent B を用いて複雑な環境において行動生成を行った実験結果を示す．このように，複雑な環境においても効用関数を生成し，適切な行動が可能である．

これらの実験によって，モータの配置という身体性が異なった場合に，それぞれに適した行動が生成されることがわかった．

6.7. 実機による実験

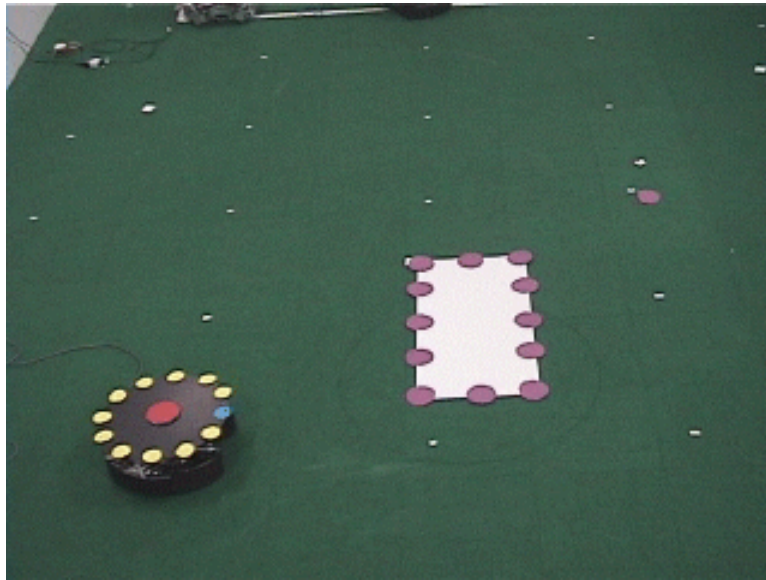
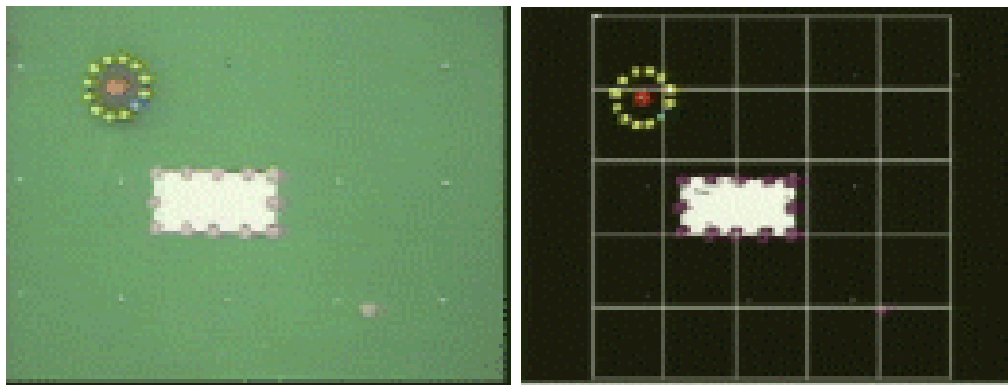


図 6.23 The experiment environment



(a) The raw image

(b) The processed image

図 6.24 The input image and processed image

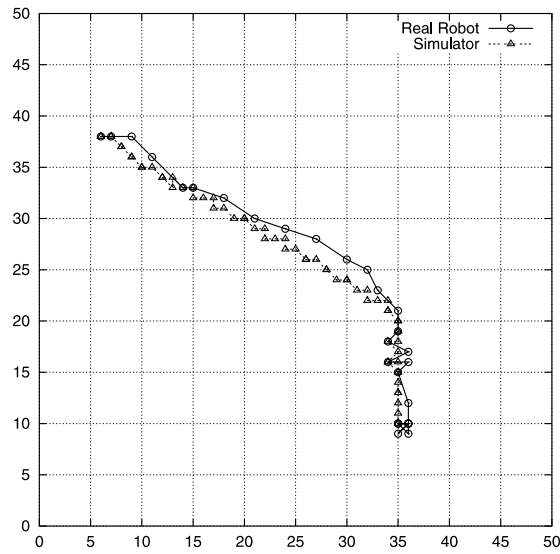


図 6.25 Comparison of generated paths of the real robot and the simulated robot without obstacle

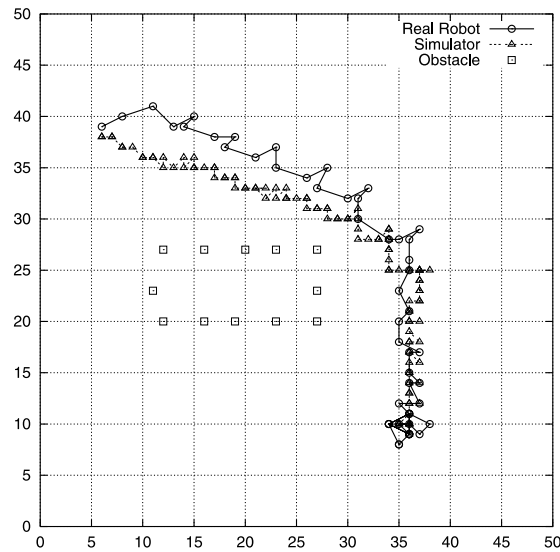


図 6.26 Comparison of generated paths of the real robot and the simulated robot with obstacle

次に、提案手法の実環境に対する適用可能性を調べるために、実機を用いた簡単な実験を行った。使用したロボットはタイプ III のもので、カメラは環境に取りつけた、すなわちグローバルビジョンシステムを用いて実験を行った (図 6.22 参照)。図 6.23 は実験環境の全体図である。障害物には画像認識しやすいように境界に相当する部分にマーカーを取り付けた。図 6.24(a) は環境に取り付けられたカメラからの入力画像であり、図 6.24(b) は、その入力画像を処理したものである。タスクは、身体上の正面の点を画像上の目標点 $(35, 10, -90)$ に到達させることとし、途中に障害物を置く場合と置かない場合を設定した。実験は、実環境と同様なシミュレーション環境を構築し、遷移確率の計算、効用関数の計算はシミュレータ上で行い、その結果を用いて実機を動かす方法をとった。行動生成の際には、入力画像から効用関数を用いて現在の効用値を計算し、政策決定を行った。

実験の結果生成された軌跡を図 6.25(障害物なし) と図 6.26(障害物あり) に示す。それぞれの図中において実線で描かれたものは実際のロボットのとった軌跡、破線で描かれたものはシミュレータ上で計算された軌跡を示している。障害物がない場合は概ね同じ軌跡をたどっているが、障害物を置いた場合には経路の前半では実ロボットとシミュレーション結果がずれている。これは、実ロボットとシミュレーションでエージェントの行動単位が異なるため、最初のステップでの回転行動により、実ロボットの方が大きく回転してしまったためである。しかし、この最初のずれはゴールに近づくにつれてすくなくなり最終的には同一地点に到達することができている。これは、効用関数という汎用的な表現を持ち、状況に合わせて政策決定を行っているからである。

6.8. 対象の認識と判別

本節では、生成された効用関数を用いて対象への行動生成を行い、出力する行動によって対象認識と判別を行なう方法を説明する。抽象度の低いタスクにおいては、対象の形状を意識して行動する必要はない。しかし、人間の行なうような抽象レベルの高いタスクにおいては、対象そのものの形状を視覚によって判別する必要がある。そこで、画像認識の問題を我々の手法に従って説明する。

画像からの対象認識において、従来の 3 次元再構成を行なう方法では、画像入

力を3次元モデルという視覚に対して不変なモデルに変換することによって見え方の違いを吸収する。我々は、3次元再構成を行わないという立場から、視覚入力を直接身体性を用いた表現に関連づけることによって、視覚入力に対して不変な認識を行なう。視覚入力に変化しても、恒常性を保つのは身体を基準とした表現である。そこで、対象の見え方から触れ方を予想し、触れ方によって対象を記述し、判別することを考える。具体的には以下の手順で行なう。

1. 視覚入力から対象の形状に応じた効用関数を生成する。
2. 効用関数を用いて仮想的になぞり行動を生成する。
3. なぞり行動をチェーン符号化し、形状特徴の抽出を行なう。
4. 形状特徴空間において対象の判別を行なう。

(1) に関しては前節で述べたので(2)以降について以下に説明する。なお、(3),(4)については、画像からの形状認識の手法として従来用いられてきたものを行動空間における対象の記述のために適用したものである。画像中で対象のチェーン符号化を行うと、見る位置や角度によって異なったものとして認識される。しかし、見え方が異なっても触れ方は同じであるので、行動空間で対象を記述すると見え方に関わらず同一の対象として認識される。

6.8.1 仮想的なぞり行動の生成

なぞり行動は部分的に見ると、図6.27左のように隣接するETAを始点と終点とする行動である。今、 ETA_k, ETA_{k+1} において観測された状態が s_k, s_{k+1} であるとき、RAを s_k から s_{k+1} に移動させる行動は、 s_{k+1} において報酬を与えることによって計算された効用関数を用い、式(6.5)の政策に従うことによって仮想的に生成することができる。このときに、次節で述べる、なぞり行動の符号化において、始点を終点で身体上の別の点が接すると、符号化の結果が異なるために身体上の同一の点で、 ETA_k, ETA_{k+1} に接触する必要がある。本研究においては、エージェントの運動能力を制限しているため、RAを隣のETAに移動させるための行動が複雑であるが、全方向に移動できるエージェントであれば、移

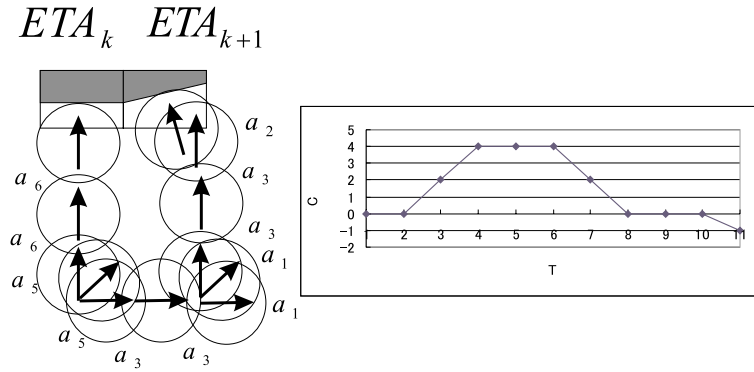


図 6.27 Trace action and chain coding

表 6.1 Chain codes corresponding to actions

行動	a_1	a_2	a_3	a_4	a_5	a_6
符号	-2	-1	0	1	2	0

動経路はより単純なものとなる．その場合でも同様に，モデルと報酬から効用関数を計算し，式 (6.5) の政策に従って経路を求める必要がある．

6.8.2 チェイン符号化

画像中の対象の形や大きさを特徴づけるための手法として，対象の境界を追跡して，方向により符号化を行なうチェイン符号化が知られる [42]．この方法では，画像上での対象の境界の変化の仕方に注目するため，実空間における対象の境界の変化の仕方を表すものではない．そこで，画像上での境界の変化の仕方を，実空間における変化の仕方，すなわちなぞり行動に変換することを考える．この隣接する ETA の間の実空間における相対角度を，行動によって符号化することを考える．

各行動単位に対してに，Table 6.1のように符号の値を決めておく．隣接する二つの ETA_s, ETA_g が与えられたときに，その間を結ぶ行動系列 a_1, a_2, \dots, a_u のような行動系列が存在したとする．そのときのコード系列は $c = \{c_1, c_2, \dots, c_u\}$

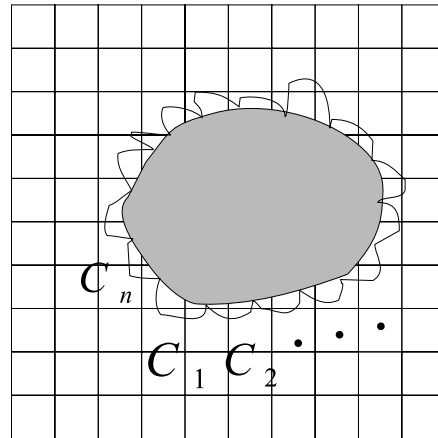


図 6.28 Characterizing of a size and shape of an object

となり，その積算は，

$$C = \sum_{i=1}^u c_i \quad (6.9)$$

となる．積算 C は，隣接する ETA 間における実空間での相対角度をあらわしている．

6.8.3 大きさ，形状特徴の抽出

次に，先に求めた C を用いて対象の形状と大きさを特徴付けることを考える．図 6.28 のような画像入力があったとき，それぞれの隣接する ETA から C の系列 $C = \{C_1, C_2, \dots, C_n\}$ が計算される．これは，対象の形状と大きさによって異なる．例えば，四角形のように直線と直角で構成される図形の場合は， $C = \{0, 0, 4, 0, 0, 4, 0, 0, 4, 0, 0, 4\}$ のようになる．また円のような場合は $C = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$ のように C の値が一定となり，対象の形状に応じた表現が可能である．大きさに関する特徴は，符号列の長さに相当する．また形状に関する特徴は，また， C を

$$C = \sum X_k e^{-i \frac{k\pi n}{|C|}} \quad (6.10)$$

のように，フーリエ展開したときの， X_k の係数部分によって表すことができる．

表 6.2 Objects used for learning

対象名	形状	大きさ
obj1	円	半径 2.1
obj2	円	半径 2.8
obj3	円	半径 3.0
obj4	矩形	3 × 5
obj5	矩形	1.5 × 4
obj6	矩形	3 × 3

表 6.3 Average of shape characterizing value

name	f_1	f_2	f_3	f_4	f_5
obj1	16	-1.2	-1.2	0.9	0.7
obj2	22	-0.4	-0.4	-0.9	0.1
obj3	24	-0.6	0.2	0.3	0.1
obj4	24	-0.8	1.9	0	-1.4
obj5	16	-1.0	2.9	0	-2.3
obj6	16	-1.0	-1.1	-0.2	-1.4

6.8.4 対象認識と判別の実験

次に、画像入力から仮想的になぞり行動を生成し、行動によって対象を認識、判別する実験を行なった。この実験におけるエージェントの目標は、環境内に配置された対象を、その位置・姿勢に関係なく同一のものとして認識することである。この実験に用いた対象を Table 6.2に示す。表中の大きさはシミュレーション環境中での大きさであり、画像上での大きさとは異なる。

6.8.5 形状特徴ベクトルの抽出

それぞれの対象についてなぞり行動を生成し、なぞり行動の系列を符号化し、形状特徴ベクトルを計算した。形状特徴ベクトルの f_1 は、符号の系列の長さである。また、符号の系列をフーリエ展開したときの、 X_1 の実数部を f_2 、虚数部を f_3 、同様に、 X_2 の実数部を f_4 、虚数部を f_5 とする。

エージェントを異なった4箇所に配置し、それぞれの位置から形状特徴ベクト

表 6.4 Object recognition rate

	obj1	obj2	obj3	obj4	obj5	obj6	average
rate	0.75	0.91	0.75	0.91	0.83	1.0	0.85

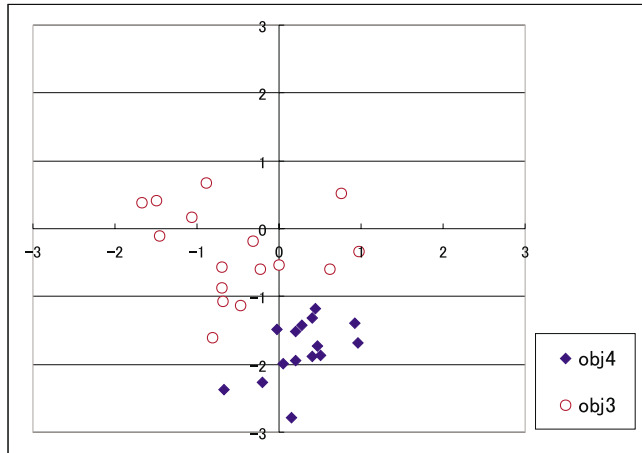


図 6.29 Shape characterizing value of f3-f4

ルの計算を行なった．各位置からはそれぞれ，20 回形状特徴ベクトルの計算を行なった．そのときのそれぞれの対象についての形状特徴ベクトルの値の平均を Table 6.3 に示す．

6.8.6 対象の判別

先に求めた，それぞれの対象についての形状特徴ベクトルを用いて対象の判別を行なった．対象の判別には，それぞれの対象について 4 つの教師事例を抽出し NN 識別法 (nearest neighbor classification) を用いた．obj1 から obj6 をそれぞれ 1 2 の異なった場所に配置し，認識の試験を行なったところ認識率は Table 6.4 の通りであった．認識率を低下させる原因としては，遷移確率が分散しているために最適経路の計算に失敗する場合があります，チェーン符号の積算値を正しく求められず，その結果として形状特徴ベクトルの値がばらついてしまうことが考えら

れる。

図 6.29に obj2 と obj4 の f_3, f_4 について，教師事例と試験に用いた事例の特徴値をグラフで示す．このグラフから，形状特徴ベクトルを用いた表現によって，同一の対象では分散が少なく，異なった対象間には分離可能であることがわかる．これは，対象を異なった位置・姿勢から観測しても同一の対象として認識可能であることを示している．

6.9. まとめ

本章では，事前知識なしにエージェントの主観に基づいて世界を表現するために，「視覚による対象認識」は「接触行動を予測すること」という仮説をもとに，効用関数によって対象表現を行うことを試みた．従来のコンピュータビジョンにおいては視覚入力から3次元再構成が可能であることが前提となっていたが，本研究では，Berkeley の主張や認知心理学実験の結果を踏まえ，視覚そのものに3次元再構成可能な幾何学的情報は含まれていないという仮定を採用した．このような仮定のもとでは，視覚入力はエージェントに対して意味を持つ情報を提供しないので，エージェントは，視覚入力に対して自己身体を用いて自ら意味づけを行う必要がある．そこで本研究では，環境を記述するために身体の属性（身体性）として身体の大きさ，形状，運動能力を挙げ，これらを用いて視覚入力に対する意味づけを行うことを試みた．提案手法において，エージェントは視覚入力が見られると環境に応じた効用関数を生成する．この効用関数はあらかじめ獲得された状態遷移モデルを用いて生成されるが，状態遷移モデルは，身体性を反映したものになっているので，効用関数によって表現された環境は，エージェントの主観によるものだと言える．

提案手法の有効性を確認するために，いくつかの実験をコンピュータシミュレーションと実機によって行った．2つの物体の間の通り抜けのタスクでは，エージェントは自己身体と比較して広いすき間はすりぬけ，狭いすき間は回避する行動を生成することができた．異なる身体を持つエージェントに対する行動生成の実験では，身体形状に応じた適切な行動を行うことができた．この2つの実験から，エージェントは自己身体に関する情報と環境に対する情報を明示的に与えられな

くても、自身の経験に基づいて環境を記述し、適切な行動生成が出来たと言える。対象認識では、異なった大きさ形の対象を行動空間において分類することによって、適切に判別することができた。これらの実験を通して、エージェントは、一旦状態遷移確率モデルを獲得するとそれを用いて任意の環境に対して効用関数を生成し行動生成できることが確認できた。このことから提案手法における環境のモデル化方法は汎用的なものであると言える。また、カメラを環境に対して傾けて設置しても適切な行動を獲得できたことから、提案手法を用いると視覚系に依存しない環境のモデル化を行うことができると言える。これはすなわち、視覚入力から幾何学的情報を取り出さなくても自己身体を基準として用いると環境を認識可能であることを示しており、これまでのコンピュータビジョンで用いられて来た3次元再構成を用いた手法とは異なる視覚認識手法の可能性を示唆するものである。

提案手法では、障害物回避のために、身体表面と対象の表面の関係を全て計算しているが、実際には多くの部分は関係ないので、効用関数の中でタスクに関連したものだけに注目する必要がある。本研究では、環境に2次元を仮定したが、提案手法は同様に3次元に拡張可能である。効用関数を生成するためには、モデルとして行動出力に対する見え方の変化を保持すればよいので、ステレオ視覚を用いて、奥行き方向の変化を検出し、モデルとして用いればよい。また、本論文における実験ではエージェントの身体と対象の形状を簡単なものとしたが、視覚入力中でのエージェントの身体と対象を微小領域の総和として取り扱い効用関数生成しているので、提案手法は複雑な形状についても適用可能である。

実世界において行動するエージェントにとって必要な情報は、3次元の幾何学的な情報ではなく対象について何が出来るかということである。このエージェントにおける可能な行動の知覚は、ギブソンの言葉を借りるならアフォーダンスの知覚である [13]。例えば、椅子は人間に対して座るというアフォーダンスを持つが、アリにとっては巨大な構造物として、またゾウにとっては単なる障害物として認識され、座るというアフォーダンスはない。すなわち、椅子の持つ「座れる」という機能をエージェントが実現するためには、座れるための身体の形状と大きさ、そして座るという行為を実現するための運動能力が必要となる。本手法で示した対象の認識はこの意味でまさにアフォーダンスの知覚である。言葉をかえれ

ば本手法はこのようなアフォーダンスの知覚を工学的手法によって実現可能であることを示しているといえる。

第7章

おわりに

本章では、これまでに行ってきた研究のまとめを行い、それぞれの提案手法が設計者の埋め込みの少ない汎用的な表現になっているかについて考察を行ってから、本論文を結論づける。また、本研究が今後とるべき方向性について展望を述べる。

7.1. 本研究のまとめ

本論文では、環境で合理的に振舞う自律エージェントを設計することを目標として、視覚を有するエージェントの視覚入力から行動出力へのマッピング獲得時に設計者の知識をいかに少なくし、自動的に獲得するかということに関して三つの研究を行ってきた。

まず2章では、マッピング記述のために旧来の人工知能がとってきた方法について概説した。旧式の人工知能では、マッピング記述方法として、人間の知識を記号によって表現し、記号操作によって入出力関係を定義するというように、ほとんど人間の知識を埋め込んで記述されていた。しかし、このような方法では、人間の想定した以外の状況に適応出来きないという問題がある。

このような問題に対して Brooks が環境との相互作用を重視するシステムを提案した。しかし、環境との相互作用を重視するという着眼点は良かったものの、このシステムでは、旧来の知識の記述方法をモジュールごとに独立に書き換えただけのものであり、依然として想定外の状況に適応できなかった。そこで、次に環境との相互作用から内部表現を学習するという強化学習の手法がクローズアップされ、多くの研究が行われるようになった。本研究でも、環境との相互作用を

重視し、強化学習の手法を用いてマッピングを表現する方法を採用し、そこで生じる問題について議論することにした。問題点は、内部表現、特に視覚入力から行動出力へのマッピングの自動的獲得においては、1) どのようにして状態空間を自動的に構築するか、2) エージェントの身体性に基づいた内部表現をどのように獲得するかということである。

強化学習は、報酬を与えるだけでエージェントが探索によって最適な行動を学習するので、かなり少ない知識の埋め込みで、マッピング獲得が可能であるが、視覚を入力として用いると、視覚に含まれる膨大な情報をいかに取り扱うかという問題に遭遇する。強化学習では、学習時間が状態数の指数関数的に増加するので、学習が可能な状態数に状態空間を削減することは重要な問題である。これが1番目の問題である。この問題に対して、4章では、状態数削減のために設計者の知識を出来るだけ少なくするために、自己組織化マップを用いた状態空間の構築方法を提案した。ここでは状態数の削減のために、入力画像ベクトルの類似性にもとづいてオフライン競合学習を行う自己組織化マップアルゴリズムを用いた。この手法により、設計者の埋め込み知識なしで状態数を削減することができ、構成された状態空間を用いてロボットは障害物回避の行動を獲得できた。しかし、この手法では、入力ベクトルをユークリッド距離を用いて比較するため距離に応じて状態が分割され、細かい状態の変化は一様に無視される。しかし、特定のタスク遂行のためは、エージェントは特定の視覚的特徴の特徴の細かい変化を見なければならない場合がある。

そこで、5章では、タスク遂行のために視覚的特徴の選択の問題を取り扱った。ここでは、同時に、選択された特徴軸を適切に分割して状態数削減も行ったのだが、この特徴選択と状態分割のために、埋め込み知識を用いないために、環境との物理的相互作用を直接知覚可能な接触センサに注目した。認知心理学実験によって、視覚が行動に密接に関連していることが指摘され、特に開眼手術を受けた患者に対する認知心理学実験では、視覚が形成されるためには、実際に対象に触れるという行動が必要不可欠であることが報告されている。これは、コンピュータビジョンの分野で Aloimonos[2] らによって提案された合目的視覚 (Purposive Vision) の概念に一致する。合目的視覚での主張点は「視覚は独立したシステムとして存在するのではなく、環境に働きかけるシステム全体の一部として機能す

る」である。そこで、我々は、接触情報に基づいてタスク遂行のために必要な視覚的特徴軸の抽出と特徴軸を分割する手法を提案した。実世界における行動はほとんどの場合接触を伴うので、エージェントのタスクは接触信号によって意味付けられると考えられる。本手法では、この考えに従って、行為の終端で知覚される接触信号を行為系列に対して割り引いて与え、その割り引き信号を統計的に解析することで、タスク遂行のために必要な視覚的特徴軸の抽出と特徴軸の分割を行った。実験の結果あらかじめ準備した複数の特徴軸の中から目標到達行動のために必要な特徴軸が選択された。また、細かく分割されていた状態を統合し、状態数を削減することができた。得られた状態空間を用いて強化学習を行ったところ、エージェントは対象に到達する行動を達成することができた。

エージェントが環境との相互作用によってマッピングを獲得するときに、身体性の問題は避けて通れない問題となる。二番目の研究において、タスク遂行という条件に基づいて状態空間の生成を行えることを示したが、この状態空間はエージェントのもつ物理的特徴や運動能力を暗に組み込んだものである。そこで、6章では、マッピング獲得時にエージェントの物理的特徴や運動能力、すなわち身体性を明示的に取り込む方法について議論を行い、効用関数を用いた対象認識と行動生成の研究を行い、エージェントの身体性を考慮した対象の内部表現の獲得手法を提案した。ここでは、対象認識に関して、次の仮説を用いた。すなわち、従来対象の内部表現と行為の内部表現は別のものであるとして考えられてきたが、対象認識の問題はその対象についての行動生成の問題である。この仮説に基づき、エージェントは視覚入力から見えている対象について可能な全ての行動を予測し、それらの行動系列の特徴によって対象を判別する。行動系列を予測するために、エージェントは視覚入力を微小領域に分割し、それぞれの微小領域中に観測される対象の境界について効用値の計算を行う。対象が効用関数によって記述されていると、エージェントは対象に関するあらゆる行動を生成することが可能になる。また、効用関数は身体表面と対象表面の相対関係をその間で予測される行動によって表現するものである。実験の結果、エージェントは視覚入力から効用関数を生成し、仮想的ななぞり行動によって6種類の対象の識別を行うことができた。また、通り抜けられるすきまと通り抜けられないすきまを判別し適切な行動を行うことができた。

1 番目の手法	二値化
2 番目の手法	領域抽出
3 番目の手法	微小領域における境界抽出

表 7.1 Image processing methods applied each study

7.2. それぞれの手法に対する考察

以下では、画像処理方法を比較することで設計者の埋め込み知識がどの程度用いられたかについての考察を行う。また、獲得された内部表現の汎用性と、それぞれの手法において接触センサの果たした役割についての考察を行う。

7.2.1 画像処理における設計者の知識

本研究の目的は、マッピング獲得においていかに設計者の知識を用いることなく自動的に獲得するかであった。その中で、本研究では特に、視覚入力の取り扱い方に焦点を当てて議論してきた。この目標のために、どのような手法を適用しどの程度達成されたかについて3つの研究を比較してみる。3つの研究ではいずれも、視覚入力に対して効用関数を生成し、効用関数をマッピングの表現方法として用いた。表 7.1はそれぞれの手法において入力画像に対して適用した画像処理を比較したものである。これらの手法に対して、多様な環境への適用性の観点から考察を行う。

1 番目の手法では、入力画像に対して二値化を行い自己組織化マップへの入力とした。実験では、二値化を行う際に床と壁の境界を示すマーカの線を適切に識別できるように二値化するときの閾値を決定した。二値化の手法では、状況に応じて閾値が変動することが予想されるので、多様な環境への適用性は低い。ただ、この手法での着眼点は物理的な境界を表す特徴を見付け入力として用いることなので、別の方法を用いても境界に相当する特徴をうまく抽出することができれば、適用範囲は広がる。

2 番目の手法では、視覚入力中に見える実世界の物が判別できると仮定し、領域抽出を行い、その領域の画像内での重心の座標と面積を計算した。領域抽出で

は、画像内で連続とされる領域を抽出する際に恣意的要素を要求されるため多様な環境に適用することは難しい。実験においては、環境中の対象がその物理的特性にしたがって色分けされていたため、領域抽出も容易であったが、現実的な環境では一つの対象について多くの領域が抽出されるなど、タスク遂行のための特徴としては冗長な抽出が行われることが予想される。

3番目の手法では、画像を微小領域に分割し、微小領域において対象もしくは身体の境界が検出できることを仮定し、境界の傾きを計算した。これは1番目の手法と同様に、物理的境界を表す特徴を見つけ出すことを目標としたものであるが、ピクセルレベルで画像を取り扱うという意味で汎用性は高い手法であると言える。

以上をまとめると、視覚を有するエージェントにおいて最低限要求される視覚情報処理方法は、物理的な境界に相当する特徴を発見することであると言える。物理的境界は、実世界で行動するエージェントにおいて重要な意味を持つのでこれは妥当な前提であると考えることができる。

7.2.2 内部表現の汎用性

マッピングの獲得時に設計者の知識を少なくすることができても、内部表現が汎用的なものでなければ、エージェントは環境が変わるたびに最初から学習しなおす必要がある。そこで、3つの手法において一度獲得された内部表現がどの程度汎用的(再利用可能)であるかについて考察を行う。

1番目の手法では、自己組織化マップを用いて環境のスナップショットそのままひとつの入力ベクトルとして用いた。自己組織化アルゴリズムによって微小な変化を無視することができ、状態数を削減することができるが多様な環境では、状態数が環境の多様性にしたがって増加する。したがって、一度ある環境で獲得された内部表現は他の環境へそのまま適用することが難しく、汎用性は低い。

2番目の手法では、あらかじめ画像処理を行い領域情報を抽出して準備したいくつかの特徴軸からタスク遂行のために必要な軸の選択と分割を行った。もともと、特定の環境において特定のタスク遂行のために特徴を選択しているため、一度獲得されたマッピングを他のタスクに適用することは難しく、汎用性は低い。

3番目の手法では、微小領域内の対象の傾きに対して効用関数を計算し、対象

全体に関する効用関数は、それらを足し合わせることによって生成される。したがって、多様な対象に対しては効用関数を様々に組み合わせることで対応でき、汎用性は高い。

1番目の手法では、画像を巨視的視点から眺め、1枚の画像全体を入力としたことが汎用性を下げる理由であった。2番目の手法では、画像特徴の意味づけのために接触情報を利用したが、接触情報を直接接触に関連しない画像特徴に関連づけてしまったために接触情報と画像特徴の関連づけがタスクに依存したものとなった。3番目の手法では、接触情報を直接接触に関連する画像特徴に関係づけることを行い、また画像を微視的な視点から眺め、対象を微小領域の集合としてとらえることで汎用性を高めた。

7.2.3 接触信号の役割

本論文では一貫して接触センサの重要性を指摘し、接触信号に基づいて内部表現を獲得することを行ってきた。ここでは、それぞれの手法において接触信号果たした役割について考察を行う。

1番目の手法では接触センサの役割はそれほど重要ではなく、障害物回避というタスクの学習のために強化学習における報酬として用いらただけである。2番目の手法では、接触信号はタスク遂行に必要な視覚的特徴の選択と分割のために用いられ、視覚認識に果たす接触信号の重要性をより意識したものであった。3番目の手法では、環境における行動が全て接触を基準とするものという立場から、逆に接触信号に関連する視覚的特徴を選び、それに対して接触信号を関連づけるようにした。これによって、接触信号と視覚の密な関係を構築することができた。

7.3. 結論

これらの研究によって得られる結論は、視覚から行動へのマッピングを獲得する際に接触信号を用いると、接触信号を介したフィードバックループを形成することができ、設計者の知識をより少なくできるということである。また、視覚の本質は接触を予測することであり、タスク遂行のためには対象の境界に相当する

視覚的特徴が重要な役割を果たすということである。

7.4. 今後の展望

近年の脳科学の研究から、脳内に対象物とそれに対する行動の両方に反応する Bimodal Neuron の存在が報告されている [14]。これは、視覚における対象認識はそれに対する行動そのものを予測することであるという 6章で用いられた仮定を神経生理学の立場から説明するものである。脳科学の分野では、人間の認知行動に対する脳内での活動を MRI などを用いた観測を通して体系化しようとしている。そのような実験結果にたいして計算論的神経科学の分野では数学的に説明づける試みを行っている [48]。本研究で取り扱ってきた自律エージェントにおいても、そのような視点が必要であり、その中で、身体の形態や接触を明示的に説明するような理論を構築することが重要である。

本研究では、マッピングそのものの獲得を問題としたで、マッピングから生成される抽象表現に関しては考慮してこなかった。しかし、人間は認知的な負担を減らすために、多くの事象を抽象化した表現として持っている。機械学習における抽象化の研究が行われており [47][55]、これら研究では、従来の強化学習の枠組のもとで、ある程度柔軟な表現を獲得することが示されているが、抽象化のプロセスに特化した研究であるため、自律エージェントが自律的に内部表現を獲得するための統一的な枠組を提案するものではないが、今後は、このような抽象化を行う手法を用いて、より一般的なマッピングを獲得することを目指したい。

上に述べたことを実現することによって、全く人間の介在しない真の自律エージェントが構築されることが期待される。

謝 辞

主指導教官である小笠原司教授には、博士過程の3年からロボティクス講座に受け入れ、御指導頂きまして深く感謝致します。

今井正和助教授には、研究会などで適切な御助言を頂きました。ここに深く感謝致します。

中村恭之助手には、研究だけでなくロボカップの競技を通じても様々な御指導を頂きました。ここに深く感謝致します。

松本吉央助手には、普段からディスカッションして頂きました。ここに深く感謝致します。

国立情報学研究所の武田英明助教授(前本学知能情報処理学講座助教授)には、修士過程から6年にも渡って未熟な筆者を懇切な御指導と御助言によって、ここまで導いて頂きました。心より感謝致します。

知能情報処理学講座の上野敦志助手には、研究会等で厳しくも適切な御助言を頂きました。ここに深く感謝致します。

東京大学工学部の西田豊明教授(前本学知能情報処理学講座教授)には、博士過程に進学するきっかけを与えて頂き、また東京大学に移られてからも、何かと面倒を見て頂きました。ここに深く感謝致します。

知能情報処理学講座の木戸出正継先生には、個別ミーティングを設けて頂いて厳しくも適切な御助言を頂きました。ここに深く感謝致します。

ソフトウェア基礎講座の横矢直和教授には、お忙しいところ博士論文の審査委員になって頂きまして、適切な御助言を頂きました。ここに深く感謝致します。

知能情報処理学講座の谷村優香里さんには、研究活動の様々な面を支えて頂きました。ここに深く感謝致します。

参考文献

- [1] Linux. <http://www.linux.org>.
- [2] Y. Aloimonos. What I have learned. In *CVGIP: Image Understanding*, Vol. 60, pp. 74–85, 1994.
- [3] M. Asada, Y. Kuniyoshi, A. Drogoul, H. Asama, M. Mataric, D. Duhaut, P. Stone, and H. Kitano. “The RoboCup Physical Agent Challenge:Phase I(Draft)”. In *Proc. of The First International Workshop on RoboCup*, pp. 51–56, 1997.
- [4] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.
- [5] G. Berkeley. *Essay Towards a New Theory of Vision*. Everyman’s Library, No. 483. Dent., 1709.
- [6] L. Berthouze and Y. Kuniyoshi. Emergence and categorization of coordinated visual behavior through embodied interaction. *Machine Learning*, 31:187–200, 1998.
- [7] R. A. Brooks. “A robust layered control system for a mobile robot”. *IEEE J. Robotics and Automation*, RA-2:14–23, 1986.
- [8] R. A. Brooks. “Intelligence without representation”. *Artificail Intelligence*, 47:139–160, 1991.
- [9] D. Chapman and L. P. Kaelbling. Input generalization in delayed reinforcement learning: An algorithm and performance comparisons. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI91)*, pp. 726–731, 1991.

- [10] D. Denette. *Cognitive wheels: The frame problem of AI*. Cambridge University Press, 1987.
- [11] S. Edelman. “Reply: Representatin without reconstruction”. *CVGIP: Image Understanding*, 60:1:92–94, 1994.
- [12] R. E. Fikes and N. J. Nilsson. Strips: A new approach to theorem proving in problem solving. *Artificial Intelligence*, 2:189–208, 1971.
- [13] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin Company, 1979.
- [14] M. Graziano, G. Yap, and C. Gross. Coding of visual space by premotor neurons. *Science*, 266(11):1054–1057, 1994.
- [15] J. Hakura, H. Yokoi, and Y. Kakazu. Affordance in autonomous robot:grounding mechanism of sensory inputs. In H. Asama, T. Fukuda, T. Arai, and I. Endo eds., *Distributed Autonomous Robotic Systems 2*, pp. 156–167. Springer-Verlag, 1996.
- [16] S. Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.
- [17] R. Held and A. Hein. Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative & Physiological Psychology*, 56:872–876, 56.
- [18] I. Horswill. *Specialization of perceptual processes*”. PhD thesis, Masachusetts Institute of Technology, Cambridge, May 1993.
- [19] H. Ishiguro, R. Sato, and T. Ishida. Robot oriented state space construction. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS96)*, pp. 1496–1501, 1996.
- [20] L. Janlert. *Modeling change - the frame problem*. Z.W. Pylyshyn, 1987.
- [21] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A aurvey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

-
- [22] Y. Kakazu and J. Hakura. Exploration of perceptual mechanism for adaptive agents. In *Proc. of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1699–1706, 1996.
- [23] R. Kerver. Chimerge: Discretization of numeric attributes. In *AAAI-92, Proceedings Ninth National Conference on Artificial Intelligence*, pp. 123–128, 1992.
- [24] H. Kitano, M. Tambe, P. Stone, and et.al. The robocup synthetic agent challenge 97. In *Proc. of The First International Workshop on RoboCup*, pp. 45–50, 1997.
- [25] H. Kitano, M. Tambe, P. Stone, and et.al. “The RoboCup Synthetic Agent Challenge 97”. In *Proc. of The First International Workshop on RoboCup*, pp. 45–50, 1997.
- [26] T. Kohonen. *Self-Organizing Maps*. Springer, 1997.
- [27] B. Krose and M. Eecen. A self-organizing representation of sensor space for mobile robot navigation. In *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems*, pp. 9–14, 1994.
- [28] K. F. MacDorman and Y. Miyazaki. Robots that recognize affordances: A predictive approach. In *16th Annual Conference of the Robotics Society of Japan (vol. 2)*, pp. 901–902, 1998.
- [29] D. Marr. *Vision*. W. H. Freeman, San Francisco, 1982.
- [30] M. J. Mataric. Studying the role of embodiment in cognition. *Cybernetics and Systems, special issue on Epistemological Aspects of Embodied AI*, 28, 1997.
- [31] A. McCallum. Learning to use selective attention and short-term memory in sequential tasks. In *Proc. of The Fourth International Conference on Simulation of Adaptive Behavior (SAB96)*, 1996.
- [32] J. McCarthy and P. Hays. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*, Vol. 4, pp.

- 463–502. B. Meltzer and D. Michie(eds.), 1969.
- [33] T. Nakamura and M. Asada. Motion sketch: Aquisition of visual motion guided behaviors. In *Proc. of Int. Joint Conference on Aritificial Intelligence*, pp. 126–132, 1995.
- [34] A. N.J. Dubrawski and P. Reingnier. Learning to categorize perceptual space of a mobile robot using fuzzy-art neural network. In *Proc. of IEEE/RSJ/GI International Conference on Intelligent Robots and Systems 1994 (IROS '94)*, pp. 1272–1277, 1994.
- [35] E. Perm. The implications of embodiment for cognitive theories. In *Cybernetics and Systems '98, Proceedings of the 14th European Meeting on Cybernetics and Systems Research*, pp. 687–692, 1998.
- [36] G. Sandini and E. Grosso. “Reply: Why purposive vision”. *CVGIP: Image Understanding*, 60:1:109–112, 1994.
- [37] K. Shibata and Y. Okabe. Reinforcement learning when the visual signals are directly given as inputs. In *Pros. of ICNN '97*, pp. 1716–1720, 1997.
- [38] W. R. Stevens. *UNIX NETWORK PROGRAMMING*. Prentice Hall, Inc., 1990.
- [39] R. S. Sutton. Learning to predict by the method of temporal differences. *Machine Learning*, 3(1):9–44, 1988.
- [40] R. S. Sutton. Planning by incremental dynamic programming. In *8th International Workshop on Machine Learning*, pp. 353–357, 1991.
- [41] C. Watkins and P. Dayan. Technical note: Q-learning. *Machin Learning*, 8:39–46, 1992.
- [42] P. Zingaretti, M. Gasparroni, and L. Vecchi. Fast chain coding of region boundaries. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 20(4):407–415, 1998.
- [43] 嘉数侑昇. アフォーダンス理論に基づく自律エージェントの環境認識機構. 第5回マルチエージェントと協調計算ワークショップ, 1995.

- [44] 塩瀬隆之, 榎木哲夫, 山下剛史, 片井修. 進化型計算によるアフォーダンス概念の獲得と非言語コミュニケーションに関する考察. 計測自動制御学会第 23 回知能システムシンポジウム講演論文集, pp. 53–58, 1996.
- [45] 佐々木正人. オーガニゼーション・オブザード. 日本ロボット学会誌, 14(4):496–500, 1996.
- [46] 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*, 29(12):4–11, 1997.
- [47] 上野敦志, 中須賀真一, 堀. 自律エージェントのための状況認識と行動規則の同時学習. 人工知能学会誌, 15(2):297–308, 2000.
- [48] 川人光男. 脳の計算理論. 産業図書, 1996.
- [49] 浅田稔, 石黒浩, 國吉康夫. 認知ロボティクスの目指すもの. ロボット学会誌, 17:2–6, 1999.
- [50] 浅田稔, 野田彰一, 俵積田, 細田. 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, 13(1):68–74, 1995.
- [51] 浅田稔, 野田彰一, 細田耕. ロボットの行動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, 15:886–892, 1997.
- [52] 中村恭之, 寺田和憲, 柴田篤志, 岩瀬昌治, 森本純, 足立秀和, 武田英明. ロボットサッカー研究のための安価な視覚移動ロボットの開発. 第 3 回 JSME ロボメカ・シンポジウム講演論文集, pp. 31–36, 1998.
- [53] 鳥居修晃, 望月登志子. 視知覚の形成 1 -開眼手術後の定位と弁別-. 培風堂, 1992.
- [54] 田川聖治, 川口俊介, 井上克己, 羽根田博正. 遺伝的アルゴリズムとアフォーダンスを用いた知能ロボットの創発. 日本ロボット学会誌, 17(7):1023–1030, 1999.
- [55] 矢入健久, 中須賀真一, 堀浩一. 異種冗長なセンサ情報に基づく自律的状态抽象化法. 人工知能学会誌, 14(4):667–678, 1999.
- [56] 國吉康夫, ベルトゥーズリュク. 身体性に基づく相互作用の創発に向けて. 日本ロボット学会誌, 17(1):29–33, 1999.

研究業績

論文誌

1. Kazunori Terada, Takayuki Nakamura, Hideaki Takeda, and Toyooki Nishida. A cognitive robot architecture based on tactile and visual information. *Advanced Robotics*, Vol. 13, No. 8, pp. 767–777, 2000.
2. Kazunori Terada, Takayuki Nakamura, Hideaki Takeda, and Tsukasa Ogasawara. Embodiment based object recognition for vision-based mobile agents. *Journal of Robotics and Mechatronics* (accepted).
3. 寺田和憲, 中村恭之, 武田英明, 小笠原司. 身体性に基づく外界の認識-接触に基づく効用関数による環境の認識と行動生成-. 日本ロボット学会誌 (投稿中).

国際会議

4. Kazunori Terada, Takayuki Nakamura, Hideaki Takeda, and Tsukasa Ogasawara. Towards cognitive agents: Embodiment based object recognition for vision-based mobile agents. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 2000 (IROS 2000)*, pp. 2067–2072, 2000.
5. Kazunori Terada, Hideaki Takeda, and Tyoaki Nishida. An acquisition of the relation between vision and action using self-organizing map and reinforcement learning. In *Second International Conference on Knowledge-based Intelligent electronic systems*, pp. 429–434, 1998.
6. T. Nakamura, K. Terada, A. Shibata, J. Morimoto, H. Adachi, and H. Take-

- da. The robocup-naist: A cheap multisensor-based mobile robot with on-line visual learning capability. In *Proceedings of the second RoboCup Workshop.*, pp. 291–303, 1998.
7. Takayuki Nakamura, Kazunori Terada, Atsushi Shibata, Jun Morimoto, Hidekazu Adachi, and Hideaki Takeda. “Development of a Cheap On-board Vision Mobile Robot for Robotic Soccer Research”. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1998 (IROS '98)*, pp. 431–436, 1998.

研究会・シンポジウム

8. 寺田和憲, 中村恭之, 武田英明, 小笠原司. 対象認識のための接触に基づく効用関数の生成. 人工知能基礎論研究会 (第 42 回) 予稿集, pp. 45–50, 2000.
9. 寺田和憲, 中村恭之, 武田英明, 小笠原司. 身体性に基づく環境モデリングと行動生成. 第 5 回ロボティクスシンポジウム予稿集, pp. 92–97, 2000.

全国大会

10. 寺田和憲, 中村恭之, 武田英明, 小笠原司. 身体性に基づく対象認識. 人工知能学会全国大会 (第 14 回) 論文集, pp. 508–511, 2000.
11. 寺田和憲, 上野敦志, 武田英明, 西田豊明. 触覚と視覚の統合に基づく認知ロボットのアーキテクチャー. 第 17 回日本ロボット学会学術講演会予稿集, pp. 997–998, 1999.
12. 寺田和憲, 武田英明, 西田豊明. 触覚と視覚の統合に基づく移動ロボットの環境への適応. 人工知能学会全国大会 (第 12 回) 論文集, pp. 220–223, 1998.
13. 寺田和憲, 武田英明, 西田豊明. 自己組織化マップと強化学習を用いた視覚情報と行動の関係の獲得. 第 15 回日本ロボット学会学術講演会予稿集, pp. 177–178, 1997.
14. 中村恭之, 寺田和憲, 柴田篤志, 岩瀬昌治, 森本純, 足立秀和, 武田英明. ロボットサッカー研究のための安価な視覚移動ロボットの開発. 第 3 回 JSME

ロボメカ・シンポジア講演論文集, pp. 31-36, 1998.