

# ロボットの行動獲得のための視覚

浅田 稔

大阪大学工学部電子制御機械工学科

電子メール: asada@robotics.ccm.eng.osaka-u.ac.jp

## 1 はじめに

生物は、動的に変化する環境に適応しながら、種々のタスクを上手に遂行しているように見える。このような能力を人工のエージェント、すなわちロボットに付与することは、ロボティクス及び人工知能研究の究極の目標と考えられる。環境に適応するためには、感覚器と動作部を介して環境と相互作用し、内部構造を組織化する能力、すなわち学習能力をもつことが必要不可欠と考えられる。

生理心理学の分野では、2匹の小猫を使った以下のような実験が行われた[1](図1参照)。一方の猫Aは、自分の身体的運動を伴って視覚情報を獲得する。他方の猫Bは、自分の身体的運動を伴わずに、Aの運動のもとに視覚情報を得る。Aでは、視覚誘導型の行動が獲得され正確な知覚が構成されたのに対し、Bでは、失敗した。但し、小猫Bも実験後、能動的運動を可能にすることによって、通常の知覚が構成される。これらの実験は、能動的な身体運動と知覚が密に結合することにより、種々の行動が獲得されることを示唆する。但し、環境との相互作用で全てが決定されるわけではなく、事前に内部構造をもち、相互作用によって環境に適応する過程が必要であると解釈できる。生物学の分野[2]でも、蜂などに見られるように、生物にとって運動は知覚構成に本質的に係わっているとの見方が強い。

図 1: 2匹の小猫による視覚誘導型行動獲得の実験

それでは、人工エージェントとその環境との係わりの研究について触れて見よう。人工知能研究の分野では、最

近、環境とエージェントの相互作用を計算論的アプローチによって定式化しようとする動きがある[3]。とくに、機械学習の分野では、環境との相互作用に基づいた強化学習(後述)が、ロボット学習として持て囃され、精力的に研究されている[4]。但し、シンボル化された環境やエージェントの定式化、学習の高速化、収束証明などの理論的考察が多く、実ロボットを用いた研究例は少ない。それらは動的に変化しない単純な環境を対象にしたものがほとんどで、パンパやソナーなどの局所的なセンサーが多く、そのため反射的な行動に限定されている。[5, 6]。

環境との相互作用を実現する感覚器として「視覚」は、非常に大きな役割を担うと考えられる。観測者中心の相対的ではあるが、大局的な(遠方の)情報収集が可能である。従来、ロボットの視覚の役割は、ロボットの作業空間での正確な3次元計測、すなわち、位置、稜線の方向、面の傾きの推定と考えられてきた。機械(ロボット)の視覚の実現を目標とするコンピュータビジョンの分野でも、2次元画像からの3次元情報(構造、運動)再構成問題が、現在も精力的に研究されてきている。但し、コンピュータビジョンでの成果が直接的にロボットの視覚として役立っているとのコンセンサスはあまりない。これは、3次元再構成問題の困難さ、表現の妥当性、実時間性などの問題からと考えられるが、一方、ロボティクスの側からも、正確な3次元情報の獲得のみが視覚に期待すべき役割であるかどうかは、再考しなければならない。近年、コンピュータビジョンの新しいパラダイムとして、「能動視覚」、「定性視覚」、「目的性視覚」が着目され、ロボットのタスク遂行プロセスにおける視覚の役割との関連が明確になりはじめた[7]。更に「目的性視覚」は「視覚」を孤立したものととらえず、環境と相互作用する複雑なシステムの一部であると考えられ始めている[8]。但し、残念ながら、多くの研究者が、まだ3次元再構成問題の単純化の手段としてとらえており、視覚情報とモーターコマンドの係わりを研究している例は少ない[9]。

本稿では、環境との相互作用をとりながら、自分の行動を獲得していくプロセスに視覚情報が大きな役割を果たしていることを、移動ロボットや視覚サーボの研究例を通して示す。いずれも、環境やロボットの構造に対す

る事前知識を前提とせず、視覚情報のみを通して、タスクを遂行する枠組を提案している。そしてタスク遂行過程において、必要なパラメータ推定を実施しており、従来必要とされてきた、カメラキャリブレーション、3次元情報の再構成を一切必要としないという特徴がある。

## 2 視覚に基づく強化学習による行動獲得

最近、反射的かつ適応的な行動を獲得できるロボットの学習法として、強化学習が注目を浴びている [10]。この学習法の最大の特徴は、環境やロボット自身に関する先験的知識をほとんど必要としないところにある。強化学習の基本的な枠組みでは、ロボットと環境はそれぞれ、離散化された時間系列過程で同期した有限状態オートマトンとしてモデル化される。ロボットは、現在の環境の状態を感知し、一つの行動を実行する。状態と行動によって、環境は新しい状態に遷移し、それに応じて報酬をロボットに渡す。これらの相互作用を通して、ロボットは与えられたタスクを遂行する目的行動を学習する。(図2参照)。

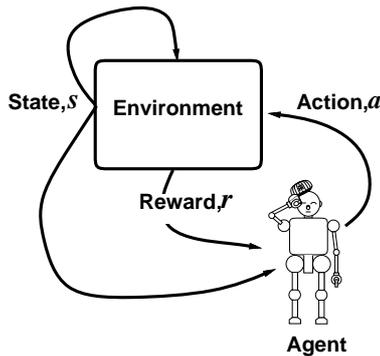


図 2: 強化学習の基本的枠組

従来、コンピュータシミュレーションよるものがほとんどで、実ロボット、特に視覚を用いた強化学習による行動獲得研究は、視覚情報処理のコストの高さから敬遠されてきた。ここでは、視覚に基づいて環境の変動を感知し、行動を決定する例として、サッカーロボットによるシューティング行動獲得の例 [11] を示す (図3参照)。

この研究の最大の特徴は、世界に関する知識、例えば、ボールやゴールの3次元位置や大きさ、ロボットの運動学・動力学的パラメータなどの知識を一切必要とせずに、ボールをゴールにシュートする行動を獲得することである。ロボットが利用できる情報は、TVカメラから得られるボールやゴールの映像情報だけであり、それらはロボットが選択した行動により変化する。ロボットの行動は前後進や回転であるが、それらの物理的意味もロボットは知る必要がない。

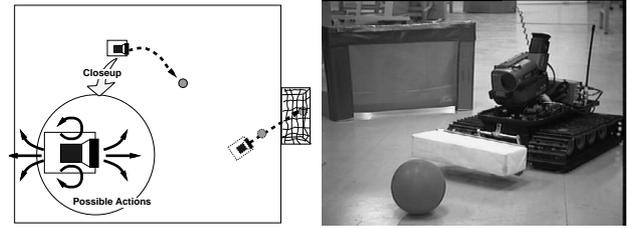


図 3: タスクとロボット

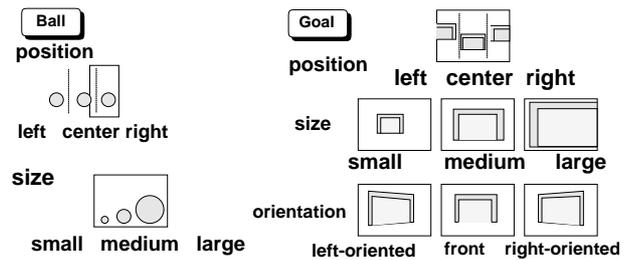


図 4: ボールとゴールからなる状態空間

視覚に基づく強化学習では、視覚情報を基本として状態空間を構成する (このタスクではボールやゴールの位置、大きさ、向きなどの組合せ: 図reffig:statsp参照) ので、画像上で識別可能な空間と、ロボットの3次元空間とは必ずしも一致しない。TV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、ロボットの行動は元の3次元世界で、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に対応するとは限らず、「状態と行動のずれ」が生じる。ここでは、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、学習を効率的に進めるために、やさしいタスクから始めて徐々に困難なタスクを遂行する枠組も提案されている (詳細は文献 [12] 参照)。行動の例を図5に示す。

## 3 オプティカルフローを用いた行動獲得

オプティカルフローの抽出に関しては、コンピュータビジョンの分野では、これまで多大な論文が発表されてきているが、画像全体に渡って正確にフローを抽出することは、非常に困難で、また多大な計算時間を要する [13]。そのため、ロボットへの応用は、問題視されてきたが、局所相関によるリアルタイムトラッキングビジョン [14] が開発されたことにより、時間をかけて正確に求める代わりに、多少ノイズがあっても、ビデオレート (33[ms]) で連続出力できることが、ロボットの常動性を確保でき、タスク遂行に有効である事が示されて来た。ここでは、こ

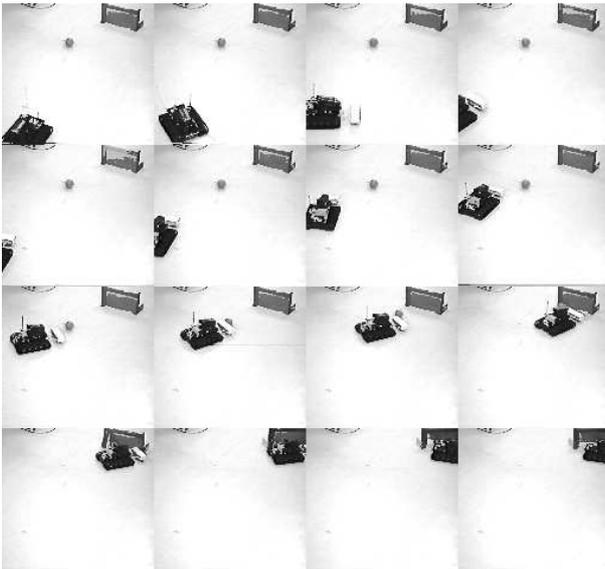


図 5: シューティング行動の例

れにより算出したオプティカルフローを用いて、移動ロボットの行動学習問題を扱った例を紹介する [15] .

エージェントと環境との相互作用を表す表現手法として、「運動スケッチ」が提案されている (図 6 参照) . 運動スケッチは、床面追跡器、目標追跡器、障害物追跡器からなる視覚行動と、左右のモーターに送るコマンドの系列のモーター行動とからなり、これらの密な結合により、視覚誘導型の行動 (目標追跡や障害物回避及びそれらの統合) が獲得される .

まず、運動コマンドと視覚情報の密な結合を獲得するために、運動コマンドを左右モーターに送る単純なボタンスイッチとし、床面のオプティカルフローとこれらのスイッチの間関係 (sensorimotor apparatus) を得る . 画像全体にフロー抽出用ウィンドウを設定し、得られるフローパターン (図 7(a) 参照) とそのとき左右のモーターに送出したモーターコマンドの組合わせを集めて、相関関係を求める (フローパターンの主成分解析) ことにより、ロボットの行動を表現する . ロボットは基本的な運動成分として回転と直進を持つ . 解析結果から得られる基本的な運動は図 7(b) に示すように、回転と直進のフローパターンに対応している . そして、全ての運動コマンドが、これら二つの主成分の線形和で表現できる . 注意しなければならない事は、ロボット自身は、3次元運動の物理的意味は、理解しておらず、自分の運動とそれによる視覚情報の変化の対応を獲得したに過ぎない .

次に、強化学習を用いて、目標追跡行動、障害物回避行動、及びそれらの統合行動を獲得できる . 図 8 に植木鉢を目標として進むが、障害物 (人間の脚) を発見して後退し、障害物が見えなくなったら、もう一度目標に向かっ

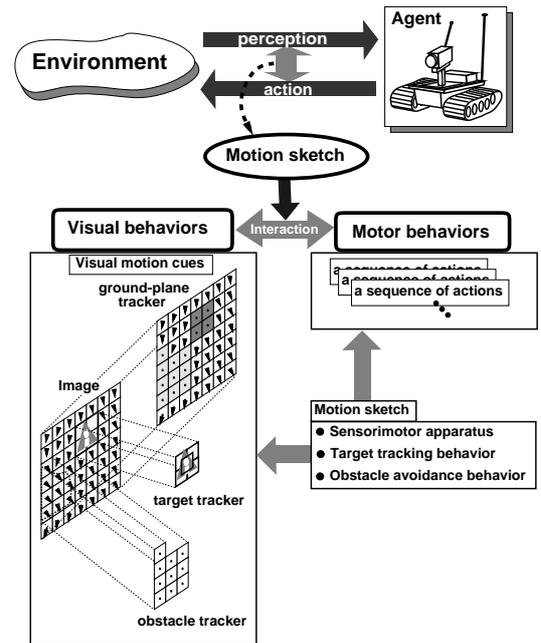
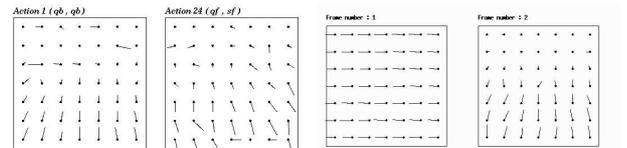


図 6: 運動スケッチ



(a) フローの例

(b) 2種類の主成分フロー

図 7: フローパターンからの主成分運動の獲得

ている様子を示す . 上段がロボットの視野を示し、下段が上から眺めた様子を示している .

#### 4 追跡視を利用した視覚サーボ

近年、視覚情報を利用したロボットアームの制御法として「視覚サーボ」に関する研究が盛んに行われている [16] . これらの研究では、視覚目標からのずれによるフィードバックによるサーボ系を構成しており、反射的な行動制御を主体としている . また、ほとんどのシステムが、ロボットアーム系やカメラ系の構造パラメータなどを既知としている . これに対し、細田ら [17, 18] は、以下の特徴をもつ汎用ビジュアルサーボ系を提案している .

- カメラパラメータやロボットアームの構造パラメータ等を一切必要とせず (よって、カメラの台数や、リンク数に依存しない) , 画像上の特徴量と関節速度の



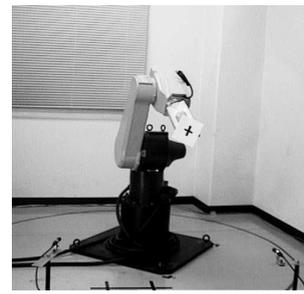
図 8: 追跡視により獲得された行動の統合

関係(ヤコビアン)をオンラインで推定することにより、サーボ系を構成する。

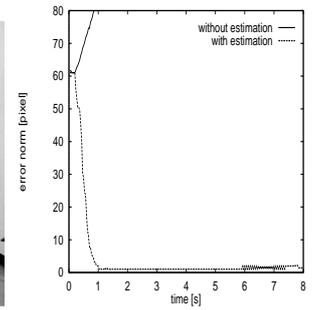
- フィードバックのみでなく、フィードフォワードを考慮する事により、(1)系の安定化をはかる、(2)反射的行動だけでなく、連続的な経路を追従させるような合目的行動の制御も可能になる。

前者の意味は、ロボット自身のもつセンサー(この場合、視覚情報と関節速度情報)のみから、視覚目標に追従するための行動を獲得するために必要な情報を推定しながら制御を行う点で、先に示した二つの例のコンセプトと合致している。後者では、3次元空間で実現可能な軌道が画像情報として与えられたときに、オンライン推定で得られたヤコビアンを用いて予測制御することが可能であり、フィードバック項と合わせてシステムのパフォーマンスの向上に寄与している。連続軌道が与えられる事の意味は、単なるPTP制御による反射的な制御(よってフィードバック主体)ではなく、事前に目標となる一種の行動系列が獲得されているという意味で、合目的な行動(purposive behavior)の制御が可能である事を主張している。

実験結果を図9に示す。(a)ではカメラが環境に固定され、ロボットアームの先端に設定された視覚目標を追跡している。(b)にステップ応答の場合の誤差の変化を示す。縦軸が誤差[画素]、横軸が時間[s]である。ヤコビアンのオンライン推定を行わない場合(without estimation)、特異姿勢に達し、目標に到達できないが、本手法(with estimation)により到達している様子が分かる。



(a) カメラ固定のセットアップ



(b) 結果

図 9: 汎用ビジュアルサーボ

## 参考文献

- [1] R. Held and A. Hein. "Movement-produced stimulation in the development of visually guided behaviors". *Journal of Comparative and Physiological Psychology*, Vol. 56:5, pp. 872-876, 1963.
- [2] G. A. Horridge. "The evolution of visual processing and the construction of seeing systems". In *Proc. of Royal Soc. London B230*, pp. 279-292, 1987.
- [3] Philip E. Agre. "Computational research on interaction and agency". *Artificial Intelligence*, Vol. 72, pp. 1-52, 1995.
- [4] R. S. Sutton. "Special issue on reinforcement learning". In R. S. Sutton(Guest), editor, *Machine Learning*, Vol. 8, pp. -. Kluwer Academic Publishers, 1992.
- [5] P. Maes and R. A. Brooks. "Learning to coordinate behaviors". In *Proc. of AAAI-90*, pp. 796-802, 1990.
- [6] J. H. Connel and S. Mahadevan. "Rapid task learning for real robot". In J. H. Connel and S. Mahadevan, editors, *Robot Learning*, chapter 5. Kluwer Academic Publishers, 1993.
- [7] Y. Aloimonos. "Introduction: Active Vision Revisited". In Y. Aloimonos, editor, *Active Perception*, chapter 0. Lawrence Erlbaum Associates, Publishers, 1993.
- [8] Y. Aloimonos. "Reply: What I have learned". *CVGIP: Image Understanding*, Vol. 60:1, pp. 74-85, 1994.
- [9] G. Sandini. "Vision during action". In Y. Aloimonos, editor, *Active Perception*, chapter 4. Lawrence Erlbaum Associates, Publishers, 1993.
- [10] J. H. Connel and S. Mahadevan, editors. *Robot Learning*. Kluwer Academic Publishers, 1993.
- [11] 浅田, 野田, 依積田, 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*, Vol. 13:1, pp. 68-74, 1995.
- [12] M. Asada, S. Noda, S. Tawaratsumida, and K. Hosoda. Vision-Based Reinforcement Learning for Purposive Behavior Acquisition. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 146-153, 1995.
- [13] M. Tarr and M. Black. "Dialogue: A computational and evolutionary perspective on the role of representation in vision". *CVGIP: Image Understanding*, Vol. 60:1, pp. 65-73, 1994.
- [14] 福葉雅幸. "局所相関を用いたトラッキングビジョン". *日本ロボット学会誌*, Vol. 13:3, pp. ??-??, 1995.
- [15] T. Nakamura and M. Asada. Motion Sketch: Acquisition of Visual Motion Guided Behaviors. In *Proc. of IJCAI-95*, pp. 126-132, 1995.
- [16] 橋本浩一. "視覚フィードバック制御一静から動へ". *システム制御情報学会誌 システム/制御/情報*, Vol. 38:12, pp. 659-665, 1994.
- [17] 細田, 浅田. "構造やパラメータに関する知識を用いないビジュアルサーボ系の構成". *第4回ロボットシンポジウム予稿集*, pp. 37-42, 1994.
- [18] K. Hosoda and M. Asada. Versatile Visual Servoing without Knowledge of True Jacobian. In *Proc. of IEEE/RSJ/GI International Conference on Intelligent Robots and Systems 1994 (IROS '94)*, pp. 186-193, 1994.