

相互感覚地図の学習を通じたロボットによる身体像の獲得

大阪大学・工・知能機能創成工学専攻 吉川 雄一郎, 細田 耕, 浅田 稔

1 はじめに

知的人工物創成のための機能デバイス・システムインテグレーションプロジェクトでは、最終的な統合の象徴の一つとして、ヒューマノイドロボット開発があげられている。本研究では、そのヒューマノイドの感覚・知覚・認知実装技術をデバイスが出来上がってから行うのでは遅いので、本年度から並行して実施すべき項目と考え、ロボットの身体像獲得研究を行う。

2 本研究の目的と背景

従来のロボティクスにおいて、ロボットの身体表現は設計者がカーテシアン座標系で定義することが多く、またそのため、ロボットのセンサ・モータ系とカーテシアン座標系との関係を設計者が校正する必要があった。従って、ロボットの身体あるいは環境に変化が起こったとき、設計者の手によらないで、身体表現を適応させることは難しい問題であった。この問題に対して、行動を通して得られるセンサ情報から身体表現を学習する能力をロボットに与えるというアプローチが有効であると考えられる。

ロボットにとっては、身体表現の獲得は難しい問題であるが、生物は誰に教えられるでもなく、いとも簡単に身体表現を獲得しているように見える。神経心理学や発達心理学などの分野において生物の身体表現は、身体像あるいは身体図式などと呼ばれ、認知発達過程解明の重要な要素としてその構造あるいは獲得機構の研究が注目されている [1, 2, 3, 4] が、非侵襲計測の限界のため、有力なモデルは見つかっていないのが現状である。Asada et al. [5] が指摘するように、ヒトの認知発達仮定を再現するロボットの実現と検証を繰り返すことは、知能ロボットの設計論に対する示唆とヒト知能の構成論的モデルを得るための有効なアプローチであると考えられる。すなわち、これらの研究分野からの示唆は、身体像を自分自身で獲得するロボットの実現の助けになることが期待され、また同時に、身体像を獲得するロボットの実現を通じて、ロボットの設計論の導出と生物の身体像獲得のメカニズムの構成論的モデルの両方に貢献することが期待される。

身体像獲得するロボットの実現を考える際、設計者によるセンサ・モータ系の解釈なしに、ロボットが感覚受容野に自分の身体を発見することは、最も基本的な問題の一つである。Asada et al. [6] はロボットの運動系への運動指令と相関のある変化を生じる画像領域は、ロボットの身体であるか、あるいは静止環境であるとみなすことができることを指摘したが、ロボットの身体と静止環境をどのように区別するかという問題は扱っていなかつ

た。Fitzpatrick and Metta [7] も同様の考え方で、外見についての先見的知識なしに、視覚においてロボット自身のアームの位置同定を実現したが、身体と環境の区別のために、カメラ運動に関する DOFs についての先見的知識をロボットに与えておく必要があった。

そこで本研究では、ロボットのセンサが物理的な身体に埋め込まれていることに起因する多様な感覚様式間の不変性によって身体が定義することができることを提案する。運動は感覚との相関をとるのに用いられるのではなく、不変性を知覚するための経験を導く役割を果たす。ロボットの身体がいくつかの様式の感覚受容野に捉えられているとき、ロボットの身体構造は短期間で変化しないとすると、それらのある種の関係は環境によらず不変である。一方、ロボットの身体を捉えていない場合、すなわち環境を捉えている場合、当然のことながら、感覚様式間の関係は環境に依存する。従って、ロボットは多様な感覚様式間の関係が不変であるか否かによって、感覚受容野が身体を捉えているか否かを判断する、すなわち、身体を発見することができる。

この考え方に基づき、Yoshikawa et al. [8] は、自分を触る経験 (ダブルタッチ) を通じて、視覚、触覚、固有感覚の相互感覚地図の学習によって身体表面の触覚センサの受容野と視覚パターンを対応づけることで、身体表面についての身体図式を獲得する手法を提案している。しかし、視覚パターンの分節化は設計者が行っており、完全にセンサ情報のみから身体図式の構成は実現されていない。そこで本研究では、はじめの課題として、視覚における注意対象が身体であるか否かを識別する問題を扱う。不変性を表現する構造として視覚と固有感覚の相互感覚地図を導入し、学習によって注視領域が身体であるか非身体であるかの識別を実現する手法を提案する。そしてこの相互感覚地図を上半身人型ロボットに適用し、手法の有効性を示す。

3 相互感覚地図の学習

ステレオカメラを持つ m 自由度のロボットを考える。左カメラの中心領域を注視領域とし、注視領域の視差を d 、ロボットの姿勢を $\theta \in \mathcal{R}^m$ とする。ロボットがある姿勢 θ をとったとき、もし注視領域に身体がある場合には、異なる環境においても同じ視差 d となる (図. 1 参照)。一方、注視しているのが身体でない場合には、異なる環境では異なる d となる。この性質を利用すると、ある θ について時間経過に対して常に一定の d をとる部分が身体、そうでない部分が非身体であると言える。

そこで、時間経過に対する不変性を発見するための構造として、姿勢を表現するノード群と視差を表現するノード群の全結合ネットワーク (相互感覚地図) を導入

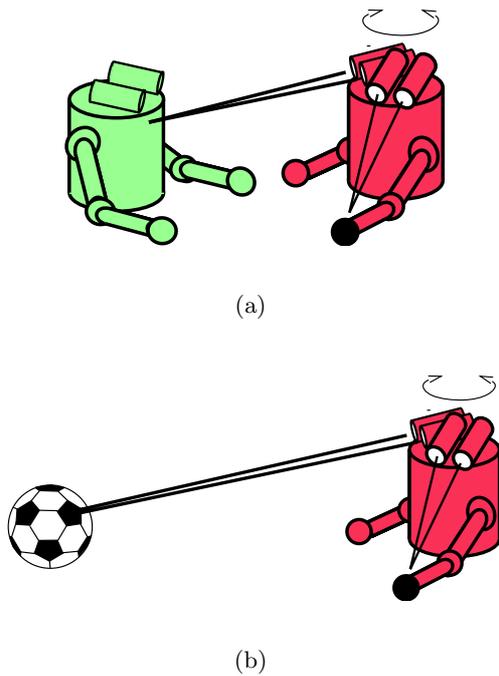


図 . 1
環境変化 ((a) および (b)) に対する視差と姿勢の間の不変性 .

する (図 . 2) . 1 から M_θ 番目のノードは姿勢を知覚するセンサに, また $M_\theta + 1$ から $M_\theta + M_d$ 番目のノードは視差を知覚するセンサに結び付けられており, i 番目のノードは, それぞれ対応するセンサ値の次元の代表ベクトル x_i を持っている . i 番目のノードは対応するセンサからセンサ値 x を受け取り, 代表ベクトルとの距離に応じて活性度 a_i を

$$a_i = \exp(-\|x - x_i\|/\sigma^2), \quad (1)$$

のように出力する . ここで, $\|x - x_i\|$ はベクトル x と x_i の距離を表わし, σ は定数である . ノード i から j へのシナプス強度 w_{ij} は

$$\tau \dot{w}_{ij} = -w_{ij} + ca_i a_j \quad (2)$$

のように更新する . ここで τ は学習の時定数, c は学習率である . 式 (2) の更新則により, シナプス強度は

$$w_{ij} = cE\{a_i a_j\} \quad (3)$$

に収束する [9] . ここで $E\{x\}$ は x の平均値を表わす . 実際には, 式 (2) を離散化した式,

$$w_{ij}(t+1) = w_{ij}(t) + \frac{1}{\tau}(a_i(t)a_j(t) - cw_{ij}(t)), \quad (4)$$

によってシナプス強度の更新を行なう . ここで, t は時間ステップを表わす . 従って, ロボットを動的な環境下で移動させながら様々な姿勢をとらせて, 相互感覚地図を学習させると, 同じ姿勢 θ のときに必ず見える身体についての視差との結合が強められる . すなわち, ある姿勢のノードが視差のノードと結合しているか否かで, 注視領域が身体か否かを知ることができる .

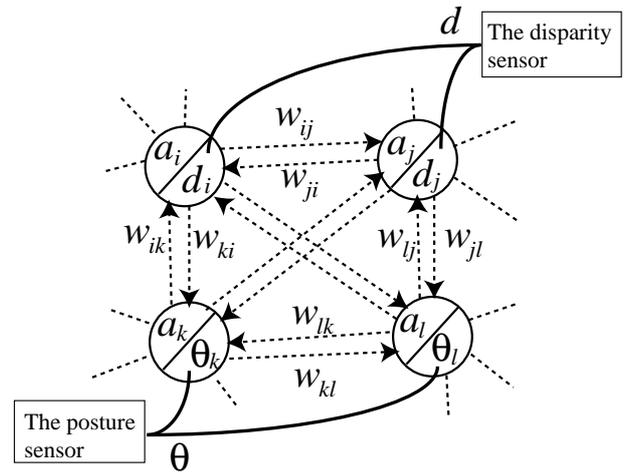


図 . 2 相互感覚地図の構造

4 実験

上半身人型ロボット (図 . 3) に相互感覚地図を実装し, ランダムな注視と移動経験を通じて, 多様な感覚様式間の不変性の学習を行なった .

4.1 実験設定

ロボットは2つのカメラ (SONY, CCB-EX37), パンチルト軸に回転可能なステレオカメラヘッド (三菱エンジニアリング, CT-301), 4 自由度アーム, 移動台車 (Nomadic, Nomad150) からなり, PC (Pentium II 400MHz) によって制御される . 図 . 4 はロボットの視野の一例である . 2つのカメラが同じ物体を捉えたフレームにおいて, 注視点の視差が計算され, ロボットのポテンシオメータの値と共に, 相互感覚地図に入力される (図 . 5) .

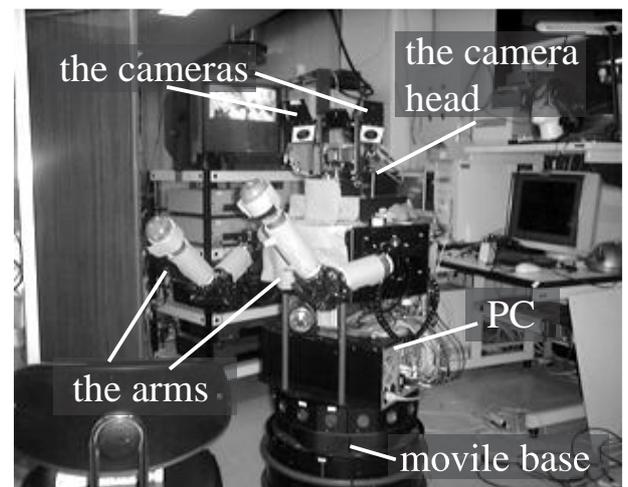


図 . 3 上半身人間型ロボットの概観 .



図 . 4 ロボットの視野の一例.

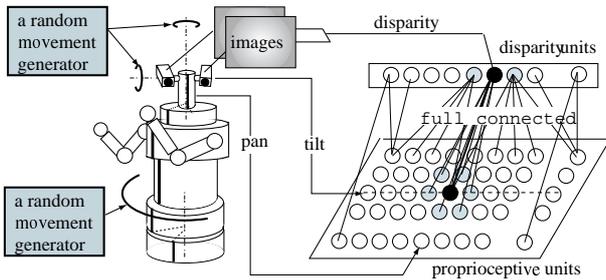


図 . 5 相互感覚地図.

4.2 学習のプロセスと結果

移動台車とカメラヘッドにランダムな入力を与え、約 6 分間学習を行なわせた後に獲得された身体図式を、簡単のためアームを適当な姿勢で固定したときの一断面を示す (Fig.6) . これは、学習後のカメラヘッドのノードを一つ選んで発火させたとき、結合を通して最も活性化した視差ユニットの代表ベクトルをカメラヘッドの関節角の関数として示したものである . カメラヘッドの関節角はパンとチルトの可動域 ($\text{pan} = -45 \sim 45$ [deg], $\text{tilt} = 10 \sim 70$ [deg]) を 20×15 個に、視差 d の定義域 ($-128 \sim 128$) は 15 個に量子化されている .

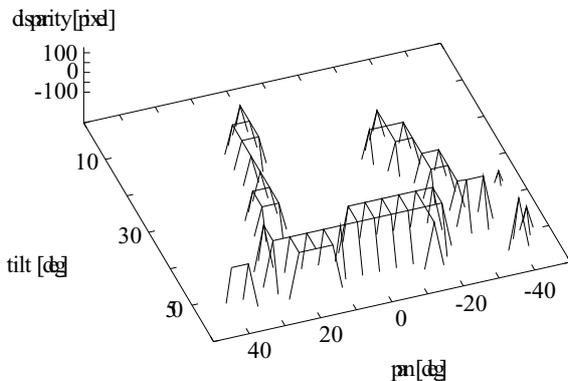


図 . 6 獲得された相互感覚地図の活性度.

図 . 6 とロボットの視野 (Fig. 4) を見比べると、腕と胸に対応する部分に視差ユニットとの結合が残っていることが読み取れる . 従って、獲得された相互感覚地図はロボットの身体を表わしていると考えられる .

5 議論と結論

本研究では、多様な感覚様式間の不変性が身体を表現するという考えに基づき、不変性を学習可能な構造として相互感覚地図を導入した . これを上半身人型ロボットに適用し、先見的な知識なしに、注視領域が身体であるか非身体であるかを識別可能な表現を、相互感覚地図に獲得することができた . 視差の変わりに、カメラ画像や触覚センサなどの他の様式を用いても同様の結果が得られると考えられる .

2 章で述べたように、本研究は生物の身体像獲得のメカニズムの構成論的モデリングも目的のひとつにしている . 現段階では十分説得力のあるモデルであると言うことは難しいが、常に観測されるものの表現が身体像を構成する一つの要素であると推測している . 神経科学の研究 [2] において、道具使用を学習したサルの脳に、道具が手の延長として一体化した像を表現していると考えられるパイモダルニューロンが発見され、身体図式との関わりが指摘されている . 道具使用時には、道具は常に手の先に観測されるものであるから、我々の推測はこの発見と矛盾しないと考えられる . しかし、このパイモダルニューロンの受容野が、サルの道具を使うという意図に依存して変化することから、タスクに応じて相互感覚地図を修正可能であるような構造に変更することを検討する必要がある .

また、提案した相互感覚地図に獲得された表現は注視領域が身体であるか否かを識別することにしか用いることができない . 獲得した表現をどのように様々なタスクに適用するのか、は今後の課題である . さらに、発見された身体を分節化して身体部位を同定することも今後の課題としてあげられる . これらの問題に対しては、触覚を含んだ多様式感覚、タスク達成および評価のシステムを統合することが必要であると考えられる .

参考文献

- [1] V. S. Ramachandran and S. Blakeslee. *Phantoms in the Brain: Probing the Mysteries of the Human mind*. William Mollow, 1998.
- [2] A. Iriki, M. Tanaka, and Y. Iwamura. Coding of modified body schema during tool use by macaque postcentral neurons. *Neuroreport*, Vol. 7, pp. 2325–2330, 1996.
- [3] A. Iriki, M. Tanaka, S. Obayashi, and Y. Iwamura. Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research*, Vol. 40, pp. 163–173, 2001.
- [4] M. S. A. Graziano, D. F. Cooke, and C. S. R. Taylor. Coding the location of the arm by sight. *Science*, Vol. 290, No. 5498, pp. 1782–1786, 2000.
- [5] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive developmental robotics

as a new paradigm for the design of humanoid robots. *Robotics and Autonomous System*, Vol. 37, pp. 185–193, 2001.

- [6] M. Asada, E. Uchibe, and K. Hosoda. Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development. *Artificial Intelligence*, Vol. 110, pp. 275–292, 1999.
- [7] P. M. Fitzpatrick and G. Metta. Toward manipulation-driven vision. In *Proc. of the 2002 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pp. 43–48, 2002.
- [8] Y. Yoshikawa, H. Kawanish, M. Asada, and K. Hosoda. Body scheme acquisition by cross map learning among tactile, image, and proprioceptive spaces. In *Proc. of the 2nd Intl. Workshop on Epigenetic Robotics*, pp. 181–184, 2002.
- [9] 甘利俊一. 神経回路網の数理 -脳の情報処理様式-, 神経細胞の学習理論. 産業図書, 1978.