

感覚統合と能動的認識に関する工学的研究

阪 口 豊

①

感覚統合と能動的認識に関する工学的研究

平成6年1月

阪口 豊

## 目次

<b>1 序論</b>	<b>1</b>
1.1 本研究の目的	1
1.2 感覚統合と能動的認識の思想	2
1.3 本研究の方法論	3
1.4 工学的技術との関わり — センサフュージョン	5
1.5 本論文の構成	7
<b>2 感覚統合と能動的認識のモデル</b>	<b>10</b>
2.1 感覚統合と能動的認識	10
2.1.1 感覚統合	10
2.1.2 能動的認識	13
2.1.3 知識に基づく内部像の形成	14
2.1.4 内部像に基づく制御	16
2.2 感覚統合と能動的認識の数理モデル	17
2.3 ボトムアップな感覚器選択メカニズム	20
2.4 内部像の予測を伴う能動的認識	22
2.4.1 状態変化モデルが単一である場合	23
2.4.2 状態変化モデルが複数ある場合	24
<b>3 能動的認識アルゴリズムに基づく触覚認識システム</b>	<b>27</b>
3.1 触覚の生理学	27
3.1.1 触覚受容器の構造と機能	27
3.1.2 温覚受容器の構造と機能	29
3.1.3 触覚, 温覚受容器の特徴	31
3.2 従来の触覚センサ	32
3.3 触覚認識システムの構成	33
3.3.1 システムの構成	33
3.3.2 計測と特徴量抽出の過程	37
3.3.3 材質の識別実験	38

3.3.4	紙質、布質の識別実験 . . . . .	42
3.3.5	考察と評価 . . . . .	43
3.4	触知覚感の定量的表現 . . . . .	44
3.4.1	感覚量の客観化 . . . . .	44
3.4.2	特徴量の組み合わせによる触知覚感の定量的表現 . . . . .	44
3.4.3	触知覚感の類似性 . . . . .	48
3.4.4	風合い評価との対照実験 . . . . .	50
3.5	第3章のまとめ . . . . .	50
4	局所的情報を統合する図形認識システム . . . . .	52
4.1	基本的な考え方 . . . . .	52
4.2	システムの構成と観測位置決定アルゴリズム . . . . .	53
4.2.1	システムの構成 . . . . .	53
4.2.2	特徴素子間の相互結合 . . . . .	54
4.2.3	観測位置の決定方法 . . . . .	54
4.2.4	トップダウン刺激 . . . . .	56
4.3	数値実験 . . . . .	57
4.3.1	実験の設定 . . . . .	57
4.3.2	結果 . . . . .	60
4.4	ボトムアップな観測位置の決定法 . . . . .	62
4.5	表現層の階層化 . . . . .	64
4.6	視覚探索モデルへの発展 . . . . .	65
5	内部像の信頼度を考慮した運動計画 . . . . .	70
5.1	基本的な考え方 . . . . .	70
5.2	なぞり運動に関する心理物理学的知見 . . . . .	72
5.3	運動計画過程のモデル化 . . . . .	74
5.4	確率モデルを用いたアルゴリズムの定式化 . . . . .	76
5.4.1	運動計画のアルゴリズム . . . . .	76
5.4.2	内部像の表現形式 . . . . .	76
5.4.3	内部像の学習 . . . . .	77
5.4.4	信頼度の制御 . . . . .	78
5.4.5	図形形状の内部像を含めた運動計画のアルゴリズム . . . . .	79
5.5	数値実験 . . . . .	81
5.5.1	実験の設定 . . . . .	81
5.5.2	実験結果 . . . . .	81
5.6	考察 . . . . .	85

5.6.1	長期的な学習と短期的な適応の組合せ	85
5.6.2	基底関数の学習	86
5.6.3	ステイフネス調節との関係	87
5.6.4	Fittsの法則との関係	88
<b>6</b>	<b>考察</b>	<b>91</b>
6.1	聴覚知覚過程における能動的認識	91
6.1.1	音源分離とカクテルパーティ効果	91
6.1.2	聴覚知覚におけるストリーム形成	92
6.1.3	能動的認識アルゴリズムに基づく音源分離システムの構想	95
6.2	知識と内部表現の獲得	97
<b>7</b>	<b>結論</b>	<b>100</b>
	謝辞	102
	参考文献	103

## 第 1 章

### 序論

#### 1.1 本研究の目的

本研究は、感覚統合と能動的認識という視点から人間の感覚情報処理について考察し、そのメカニズムを工学的に明らかにすることを目的とするものである。

人間の脳は、多数の神経細胞による並列情報処理、豊かな学習能力、高度な思考能力などの特徴を有し、生理学者、心理学者のみならず、工学者の関心を引きつけてきた。実際、多くの工学者が、脳のもつ様々な知能のメカニズムを明らかにし、またそれと同様な機能を有するシステムを構成すべく研究を行ってきた。人間のシンボル操作や推論の過程を論理演算によって実現した人工知能や、脳の神経回路網に類似した構造のネットワークを用いて脳機能を説明しようとする神経回路モデルは、そのような工学的研究の代表例である。また、これらに加えて、超並列処理や非線形力学、自己組織化なども脳における情報処理の基本原則と考えられており、それぞれの切り口から脳のメカニズムを解明しようとする研究が盛んに行なわれている。本論文の主題である感覚統合と能動的認識も、そのような脳の情報処理原理の一つである。

感覚統合 (sensory integration) とは、文字どおり、複数の感覚情報を統合することによって、外界を理解することである。すなわち、脳は、視覚、聴覚、触覚など各感覚系で捉えた対象に関する情報を統合することによって、対象に対する「像」を内部に形成し、対象を理解する。例えば、ワインを鑑賞するときには、グラスに注いだときの色や透明度などを視覚によって、香りを嗅覚によって、そして口に含んだときの味を味覚によってそれぞれ捉え、それらの情報をあわせることによってそのワインの全体像を作り上げる。また、対象の形状や材質を認識する際には、その対象をさまざまな角度から眺めて立体形状を視覚的に知り、対象を触ったときの感触から材質や形状を触覚的に知り、また、対象をはじいたときの音の感じから聴覚的に材質を知るといったように、さまざまな観測行為によって得られた情報を組み合わせて対象を理解する。このように、さまざまな感覚情報を組み合わせて対象を総合的に理解する機能は、人間が日常的活動を営む上で重要な役割を果たしている。

一方、能動的認識 (attentional perception<sup>1</sup>) とは、対象に関する情報を収集する際に、感覚受容器が捉えた情報を受動的に受け取るだけでなく、自分にとって必要な情報を能動的に獲得することによって対象を認識することである。例えば、対象をいろいろな方向から眺めてその形を知ろうとするとき、人間は、無為に頭を動かすわけではなく、適切な位置に視点を定めて効率的にその立体形状を理解する。また、文章を読む際には、視野全体の情報を一様に処理するのではなく、視野の一部に（つまり今読んでいる文字に）注意を集中し、注意を走査することによって文字を逐次的に処理していく。

卓近な例を取れば、「網膜には像として映っていても自分にとって関心のあることしか見えていない」「盲人では眼のつけどころが違う」といったことは、人間が受け取った情報をすべて平等に処理しているのではなく、自分にとって有用な情報に重きを置いて処理を行っていることを示唆している。このように、人間は、自分の意図や目的に応じて、必要な情報だけを抽出することによって対象を理解している。以下の章で詳しく論じるが、能動的認識は、感覚統合と一体となって人間の感覚情報処理過程を支えている。

このように、感覚統合と能動的認識は脳の情報処理を特徴づける重要な要素であり、また、特定の感覚や状況のみではなく、あらゆる感覚情報処理過程において共通して機能する基本的メカニズムである。にもかかわらず、これらはいずれも従来の知能機械研究において正面から捉えられなかった。本研究は、これら二つの原理を情報処理的な視点から整理して、そのメカニズムを工学的アルゴリズムとして記述するとともに、構成したアルゴリズムを具体的な感覚情報処理過程に適用してその有効性を実験的に検証するものである。

## 1.2 感覚統合と能動的認識の思想

人間が外界を理解する上で、種々の感覚情報を能動的に統合することが重要であることは、古くから哲学や心理学の分野で指摘されてきた。

人間の感覚を視覚、聴覚、味覚、嗅覚、触覚の五感に分けたのは Aristotelés であると言われている [78]。しかし、彼は、感覚がそれら五つに分解されることだけでなく、それらの間に共通して働く機能やそれらを統合する機能があることを同時に指摘し、それを共通感覚と呼んだ。そして、運動や静止、大きさ、数といった複数の感覚に共通する概念は共通感覚のもとに成立し、これが欠けると世界を積極的に構成できなくなると述べている。このように、感覚を統合する機能が、単に受け取った感覚情報を組み合わせるためだけでなく、人間が外界を理解した外界に働きかける上で不可欠な機能であることは、すでにギリシャ時代から指摘されていた。

異種感覚間の関係や相互作用については、その後さまざまな議論がなされてきた。例えば、18世紀の視覚学者 Barkeley は、「触覚によって立方体と球を区別できることを学んだ生来の視覚障害者が閉眼したとき、視覚によってそれらを判別できるか」という「Molyneuxの疑問」を題材として視覚と触覚のどちらが優位であるかについて論じ、視覚によって得られる空間知覚

<sup>1</sup>この用語は筆者の造語である。能動的認識は active perception の邦訳であるが、active perception の元来の意味が筆者の用いる意味と若干異なるので、混同を避けるためこの用語をあてた。

は、筋肉や関節を含めた身体感覚としての触覚が視覚に教えたものだと述べている [122]。

20 世紀に入ると、知覚における能動性の意味は現象論の枠組みの中で論じられるようになる。Husserl は、すべての心的活動は必ず何か対象に向かっている（志向性という）と指摘し、意識は志向を次々と向けかえることによって対象の全体像を構成すると主張している [116]。彼の議論の根底には、物体そのもの（対象）とそれが知覚されるもの（表象）とは別のものであり、表象が構成されていく場こそが意識であるという考え方が流れている。また、Merleau-Ponty は、「知覚の現象学」[67]において、対象を知覚する過程においては、自分の身体そのものがそれに影響を及ぼすことを指摘し、身体を抜きにして対象と意識との関係を論じることはできないと述べている。

心理学の分野においても、著名な研究者はすべてといってよいほど知覚における能動性について言及している。発達心理学者 Piaget は、感覚の発達は運動によってもたらされるものであり、動作能力の発達にしたがって感覚機構は作られていくと述べている [27]。つまり、運動能力が発達し、見たり聞いたりしているものを実際に手にとっていろいろと操作できるようになると、物体の見え方が向きによって変わることやいろいろな音がすることを眼や耳で体験することができ、その結果、感覚相互の関係が獲得できると主張している。ただし、彼が指摘したのは、知覚メカニズム獲得過程における能動性の重要性であり、認識過程における能動性ではない。

Neisser [80] は、知覚過程が「感覚受容→モダルの同定→予測→感覚受容」の繰り返しによって成立していると指摘している。知覚循環論と呼ばれるこの考え方は、知覚の過程が、刺激の受容を出発点としてそれを処理し運動が発現するという古典的な「感覚反応図式」によって説明されるものではなく、各時点での心の状態に応じて感覚受容の内容が変化する能動的な過程であることを表している。Neisser はまた、種々の感覚系からの情報が最終的に一つの表象としてまとまることによって感覚統合の過程は完結すると述べている。

また、Gibson [24] は、アフォーダンス (affordance) という言葉を用いて、認識対象はそれに対する行為を内包していると主張している。これは、例えば、「そこに座るもの」が椅子であると考えれば、道ばたの丸太に腰をかけようと考えた時点で、その丸太は「椅子」として捉えられる、という考え方であり、対象の知覚がそれをどのように用いるかということの影響を受けることを指摘しているものである。これは、「道具としての意味」という視点から対象の認識について論じたものであるといえる。

このほか現在でも、知覚における能動性の果たす役割について心理学的研究が精力的に行なわれている [88, 103]。

### 1.3 本研究の方法論

前節で述べたように、認識過程における感覚統合と能動性の重要性は思想の分野において強く指摘されてきたにもかかわらず、そのメカニズムに関しては長い間暗箱のまま放置されてきた。これらのメカニズムを明らかにするためには、その情報処理過程を、思想としてではなく客観的な理論やアルゴリズムとして記述することが不可欠である。そして、そのための方法論として、

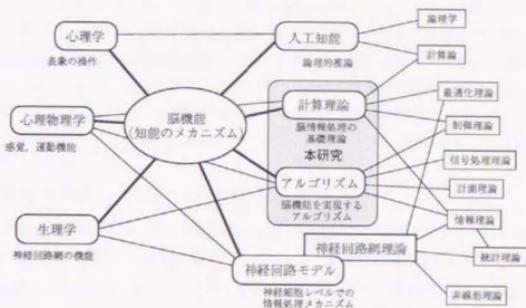


図 1.1 脳機能解明へのアプローチと本研究の方法論

工学的（あるいは構成的）手法は有益な手段を提供すると考えられる。本研究では、人間と同じような振舞いをするシステムを構成することを通じて、そのメカニズムやアルゴリズムを情報処理レベルで明らかにするという方法論を用いる。以下では、本研究でのアプローチについて脳研究における位置づけも含めて述べる（図 1.1）。

冒頭でも述べたように、脳の情報処理メカニズムを取り扱う工学的手法の代表的な例として、人工知能と神経回路モデルが挙げられる。

人工知能は、論理学やアルゴリズム理論を背景とした技術であり、その能力や性質は理論的に明らかにされている [10, 105]。しかし、人工知能はあくまで人間の論理的推論過程を模倣した手法であり、感覚統合や能動的認識といった感覚情報処理過程を記述するための表現手段として適切であるとはいえない。実際、人工知能のアルゴリズムを用いて知能ロボット制御などが行なわれているが、その振舞いからは知能的というよりむしろ機械的な印象を受けることが多い。

一方、神経回路モデルは、脳がそれ自体神経回路網で構成されていることを根拠に、脳のモデルとして広く用いられてきた [8, 79, 85]。確かに、神経回路モデルは、その非線形入出力特性ゆえに情報表現能力が豊かであり、種々の情報変換や連想記憶、動的パターン生成などさまざまな興味深い現象を実現する潜在的能力を有している。しかし、その非線形性のために性質が複雑であり、神経回路モデル自身の性質を明らかにしようとする神経回路網理論が、力学理論や統計理論をもとにして現在構築されている段階にある。実際、特定の脳機能に対する神経回路モデルを構成しようとするとき、そのような現象が起こるような特別な構造のネットワークを用意することになるため、モデルの構造に恣意性が生じやすい。そのため、複雑な機能を実現しようとするとき、現実の脳の構造から乖離してモデルが一人歩きしてしまう危険をはらんでいる。このような理由から、神経回路網はその潜在能力にもかかわらず、現時点では、高度な知能を表現する土台として利用することは難しいと思われる。

以上の理由から、本研究では、人工知能や神経回路モデルといった人間の知能を扱うための表現土台から離れて、人間を一つの「機械」として捉えた上で、それと同じような振舞いをする機械を、情報理論、制御理論、信号処理論といった工学的な理論に立脚して設計し、製作するという方法を用いる。これは、生物やさまざまな機械に共通する情報処理原理を明らかにしようとするサイバネティクス (cybernetics) [134] の考え方に基づくものといえる。あるいは、Marr [66] が用いた計算理論的アプローチを、よりアルゴリズムに近いレベルで行なうものだともいえるであろう。

このような方法論には、次のような特徴がある。

1. 工学理論の表現法を用いて対象を記述することによって、暗箱を覗かず思想的、抽象的なモデルを越えて、実際に動作する機械として脳機能を実現することができる。
2. 表現形態に特に制限を設けないので、場合に応じて適切な表現形式を利用できる。
3. モデルを構成する際に、これまで蓄積されてきた工学理論の結果を利用することができる。これによって、見通しよく議論を進めることができるほか、モデルの性能を客観的に評価することができる。

このようにして構成した機械は明らかに人間とは異なる方法で計算を行なっているが、構成した機械の振舞いが人間の振舞いとよく似ていれば、情報処理アルゴリズムのレベルで脳機能を模擬しているとみなすことができる。神経回路レベルでのモデル化は、今後、生理学的知見が蓄積されるに従って徐々に進めていけばよい。

感覚統合と能動的認識は感覚情報処理における共通原理であるから、その本質は抽象的なアルゴリズムとして表現することができると考えられる。本論文では、まず、このような一般的なアルゴリズムを数理モデル上に実現し、それによって、これらのメカニズムの本質を抽出することから始める。続いて、そのような考察を踏まえた上で、具体的な感覚知覚過程を根拠するシステムの構成を通じて個々の具体的なアルゴリズムを実現し、その性能を実験的に評価することにする。具体的な題材については、1.5節で述べる。

#### 1.4 工学的技術との関わり — センサフュージョン

さて、感覚統合のメカニズムを工学的に応用しようという研究は、センサ統合 (multi-sensor integration) [64] やセンサフュージョン (sensor fusion) [136] という名の下ですでに始められている。例えば、自律移動車や自律ロボットにおいて、テレビカメラ、超音波センサ、レンジファインダといった複数のセンサ情報をいかにして統合するか [44, 86]、レーダーシステムにおいて、多数のレーダー基地からの情報をいかにして統合するか [119]、対象の三次元形状を知るためにはどのような位置から対象を観測すればよいか [101] といったことについては、すでに現実の問題として研究が行なわれている。また、種々の場所に配置されたレーダーからの情報を統合する問題については、分散検定などの統計的問題として理論的研究が進められている [16, 30, 87, 119, 133]。

センサフュージョンによってもたらされる工学的な利点としては、次のようなものが挙げられる。

**情報の補完** 例えば、受容域の異なるセンサをいくつか用意することによって、全体として広い受容域を実現するといったように、種々のセンサからの情報を相補的に組み合わせることで、検出できる信号の範囲を広げることができる。異なる波長の光を受容する光感知器を置いて、光のスペクトル分布を得ることもその一例である。

**異種情報の統合** 異なる種類の信号や異なる観測方法によって得た信号を統合することによって、新たな形態の情報を得ることができる。例えば、複数のカメラ画像を組み合わせる対象の3次元形状を構成したり、圧力センサと流速センサからの情報を組み合わせる流量を求めたりすることがこれに相当する。

**信頼性の向上、故障の検知** 同一の情報を担う複数の信号を観測することで、測定値の信頼性を向上させることができる。また、それらの情報を互いに監視することによってセンサの故障を検知することができる。

**効率的な計測** まず、レンジファインダで対象の位置を知った後、カメラをそこに向けて対象の形状を知るといったように、一方のセンサで獲得した情報をもとに他のセンサの動作を規定することによって、効率的な情報獲得が実現できる。

複数の情報を組み合わせることによって、単一のセンサでは得られなかった機能がもたらされるのはある意味で当然のことであり、このような計測手法はそれら一つ一つを取り上げれば、これまでも用いられてきたに違いないであろう。しかしその一方で、このような計測手法は、問題ごとにそれぞれ *ad hoc* な形でしか利用されてこなかったともいえる。工学におけるセンサフュージョン研究の意味は、このようなさまざまな技術を統一的な視点から見直すことによって、その設計過程や手順を体系的に定式化することにある [36,37]。

そのためにはまず、従来用いられてきた種々の方法やアルゴリズムを整理し、その中から目的を実現する上で適切なものを選択する手法を生み出さなければならない。すなわち、受け取ったセンサ情報を単純に処理するという「受動的な情報処理」ではなく、システムの目的に応じてどのようなセンサをどのように使用してどのように処理すべきかという意志決定の問題を含んだ「能動的な情報処理」を実現することが求められる。このように、センシングの目的にまで踏み込んでセンシング全体を統一的に考える枠組みは「意図的センシング (intentional sensing)」と呼ばれる [136]。

このように、センサフュージョン研究の要点は、解くべき問題とそれを解くためのアルゴリズムとの関係を客観的に定式化することにあると考えられている。しかし、このことに加えて、次の二つの意味で、この研究は今後さらに重要な意味をもつと考えられる。

一つ目は、システムのもつ計算資源と計測アルゴリズムの関係である。

システムにおいてセンサの数を増やせば、その分だけ対象に関する情報が増加するのであるから、システムの性能が向上することは当然のことである。しかし、センサの数が増えるに従って処理すべき情報は増え、それに伴う計算量も増加する。しかも、それらの情報を組み合わせる処

理することを考えれば、それに要する計算量は組合せ論的に増加する。このような状況に対処する方法として二つの方向が考えられる。

一つは、増加する計算量をこなすだけの計算資源を投入することである。これは、すべての情報を処理するという意味で完全な方法であるが、それを実現するためには大容量の計算資源を用意しなければならない。もう一つの方法は、すべてのセンサ情報を処理するのではなく、その中から必要な情報だけを取り出して処理することによって、適度の計算資源のもとで適度な性能を実現する方法である。これは、以上で繰り返し述べてきた能動的認識の考え方にほかならない。この方法では、情報処理戦略を定めるための計算資源が必要になるが、システム全体としての計算資源は少なく抑えることができる。ただし、計算の一部を省略しているのであるから、すべての場合に万全の性能を示すものではない。

このように、センサフュージョン・システムにおいては、システムの扱う情報量の増加に伴う計算量の増加が本質的な問題の一つとして現われ、これを解決する上で、能動的認識の考え方は有効な鍵になると思われる。

二つ目は、並列処理と直接処理の統合である。

いま、視覚情報処理を例にとって考えてみる。画像処理プロセスにおいては、多くの場合、エッジ抽出、細線化などの局所的な前処理を行なった後、それらの結果を組み合わせることによって処理を進めていく。これらの前処理の多くは、画面の各部分ごとに独立に実行できるので、並列演算による処理が有効であると考えられる。しかし、これら局所的な情報を統合するメカニズムをすべて並列に用意するためには、Minsky [70] が指摘したように膨大な数の計算要素が必要になる。これに対し、人間は画像に含まれる特徴を抽出する単純な処理は低次の感覚野において並列に行ない、高度な認識処理は注意を逐次的に振り向けることで実現している。このことは、脳が、直接処理のメカニズムを用いることによって並列処理機構が巨大化するのを抑えていると解釈することができる。そして、この並列処理と直列処理をつなぐ部分が、まさに感覚統合と能動的認識の舞台である。

このように、センサフュージョンの研究は、並列情報処理と直列情報処理の橋渡しについて論じる上でも種々の題材を提供する分野であると考えられる。そして、人間の感覚統合、能動的認識のアルゴリズムを明確に記述することは、これらの工学的な問題を解決する上で有益な示唆を与えるものと考えられる。

## 1.5 本論文の構成

先に述べたように、本論文では、いくつかの感覚情報処理過程を題材として取り上げ、ケーススタディを重ねる中で感覚統合と能動的認識について具体的に検討する。題材として取り上げるのは、触知覚認識、図形認識、運動計画の三つである。

人間の触知覚は、皮膚の機械受容器だけではなく、筋紡錘、温覚受容器などさまざまな感覚受容器からの情報を統合することによって得られる感覚であり、また、押す、こする、つまむといった種々の触運動を通じて得られることから [22, 42, 63]、感覚統合と能動的認識について議論

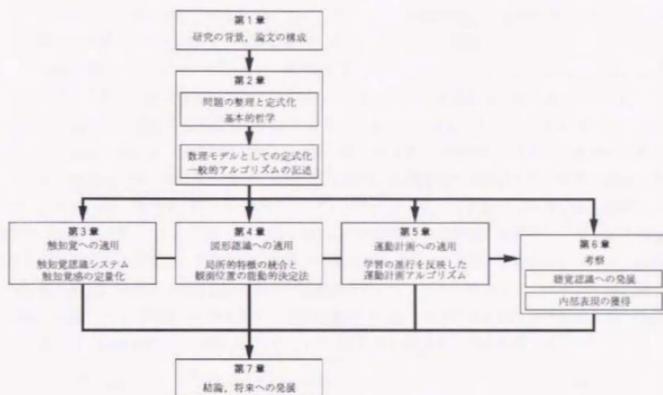


図 1.2 本論文の構成

するための題材として適切である。本論文では、種々のセンサ情報を統合して対象の材質を能動的に識別するシステムを実際に構成することによって、能動的認識の性能について実験的に検討する。また、製作したシステムによって得られた触知覚感の表現が人間の触知覚感とよく対応していることを示すとともに、触知覚感の定量的表現についても考察する。

人間の視覚認識過程においては、網膜に投影された画像を並列的に解析するだけでなく、視野の各部分を走査することによって逐次的に解析を行なっていることが、心理物理実験によって明らかにされている [34]。また、顔の写真や絵などを見るとき、人間が注視点を次々と移動させていくことは、よく知られた事実である [137]。これらの事実は、人間が局所的な情報を逐次的に統合して対象を理解していることを強く示唆しているが、このとき、脳の中では、観測した局所的情報をもとに対象の全体像が構成されていると思われる。以上の考えに基づき、二つ目の題材として、対象に関する知識を利用して、観測した局所的な特徴をもとに図形の全体像を形成する問題を取り上げる。この問題を解決する中で、図形の像が安定的にかつ速やかに形成されるような、観測位置の能動的決定法を明らかにする。

人間がある運動をしようとする際、脳は、運動器官に関する内部モデルを利用して、種々の運動指令に対して身体がどのように動くかを予測した上で、適切な運動指令を選択していると考えられる。このことは、運動計画が内部モデルに依存して定められることを表しており、「内部モデルに基づいて判断を下す」という意味で、運動計画の過程が能動的認識と共通する側面を有していることを示している。このような考えから、三つ目の題材として、内部モデルに基づく運動計画を取り上げる。そして、内部モデルがどれだけ確実に運動器官の特性を表現しているかを表

す「内部像の信頼度」という概念を設けることによって、目標の運動を適切に実現しつつ、運動に習熟するに従ってそのパフォーマンスを向上させるシステムを構成する。

本論文の構成を図 1.2 に示す。以下、第 2 章の前半では、感覚統合と能動的認識の問題を整理した上で、それらが感覚情報処理において果たす役割について考察する。第 2 章の後半では、この考察に基づいてそれらの過程を一般的な数理モデルとして定式化し、そのアルゴリズムを記述する。本章において、以降の章で論じる問題を取り扱うための基本的な考え方を系統的に述べる。第 3 章以下では、それぞれ触覚認識、図形認識、運動計画を題材として具体的な議論を行なう。まず、第 3 章では、第 2 章で構成したアルゴリズムを用いて対象の材質や質感を識別する触覚認識システムについて述べる。第 4 章では、外界のもつ情報構造を利用して局所的な特徴から図形の全体像を構成し認識するシステムを構成する。第 5 章では、運動計画における内部像の役割を論じ、内部像の信頼度を考慮した運動計画アルゴリズムを構成する。第 6 章では、聴覚認識の問題、および認識の土台となる内部表現獲得の問題についてそれぞれ考察を加える。最後に、第 7 章では本論文の結論を述べるとともに、今後の研究課題について論じる。

## 第 2 章

### 感覚統合と能動的認識のモデル

本章では、感覚統合と能動的認識の過程を独自の視点から定式化した上で、その数理モデルを構成する。

まず 2.1 節において、これらのメカニズムが感覚情報処理において果たす役割について考えるとともに、問題を整理して本論文で取扱う問題を明確にする。また、この節において「内部像に基づく対象の理解」という本論文における基本的考え方を述べる。続いて 2.2 節において、感覚統合を Bayes 推定として、能動的認識を逐次実験計画としてそれぞれ定式化し、そのアルゴリズムを具体的に記述する。さらに、この基本的アルゴリズムの拡張として、2.3 節ではボトムアップな感覚器選択方法について、2.4 節では時間的に変化する対象に対する能動的認識アルゴリズムについて述べる。

#### 2.1 感覚統合と能動的認識

##### 2.1.1 感覚統合

人間は、感覚器を通じて外界や対象の情報を獲得し、自分の頭の中にその内部イメージでも呼ぶべきものを構成する。本論文では、この内部イメージ（あるいは表象）に対応するものを「内部像」と呼び、人間の知覚過程を「対象の内部像を構成する」過程と定める。知覚過程をこのように定めると、感覚統合は「複数の感覚情報を組み合わせることによって、対象の内部像を構成する」過程であると定義することができる。

しかし、人間の知覚過程においては、種々の情報がさまざまな形式でさまざまな段階を経て統合されており、それら全体をひとまとめにしたままでは明確な議論ができない。以下ではまず、感覚統合の過程をいくつかの様式に分類し、本論文で扱う問題がその中のどこに位置するかを明確にする。

まず、感覚統合の過程は、大きく次の二つの様式に分けることができる（図 2.1）。

一つは、複数の感覚器によって並行して捉えた感覚情報を統合するというものである。この様式はさらに、空間的に分布した同種の感覚受容器からの信号を統合する形式と、異種の感覚受容器からの信号を統合する形式に分けられる。前者の例として、網膜に分布した多数の視覚細胞で

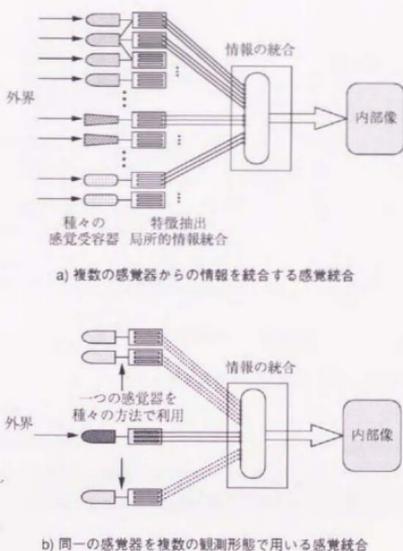


図 2.1 感覚統合の二つの様式

光の強度分布を並列的に捉えて画像情報を得ることや、皮膚内に分布した機械受容器の情報を統合して皮膚表面への刺激の分布を知ること、また後者の例として、聴覚情報と視覚情報を統合して対象の位置を知ることが挙げられる。

もう一つは、複数の観測行為によって逐次的に得られた情報を脳の中で組み合わせる形式である<sup>1</sup>。例えば、視覚によって自分の周囲の状況を理解する際、眼や頭を動かして視野を変化させることや、触覚によって対象を知る際、指を使って対象を押ししたりこすったりすることがその良い例である。

これら二つの様式は、それぞれ次のような意味をもっていると考えられる。前者は、多数の受容器を空間的に並べることによって分布情報を獲得する、あるいは、画像情報、音響情報、機械的情報などそれぞれの情報形態に合った受容器を用意し、それらの情報を統合することによって対象を多種多様な面から捉えるという形で機能する。一方、後者は、多数の受容器を並べることができない、あるいは専用の感覚器を設けることができないといった物理的、生物的制約の下

<sup>1</sup>一般には、感覚統合という用語は前者のみを指すことが多いようである。しかし、後者の様式も感覚統合の枠組みの中で統一に取り扱うとわかりやすいので、本論文ではこの用語を両者を含めた広い意味で用いる。

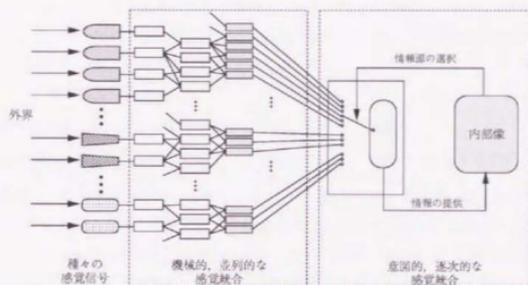


図 2.2 感覚統合の二つの段階

で、感覚器を動かすことによって一つの感覚器を種々の形態で利用し、情報を収集する機能であると考えられる。

このように、感覚統合の意味は、

1. 空間的に分布した情報を同時並行的に収集すること。
2. 種々の物理、化学現象を検知する多種類の感覚器からの情報を組み合わせて、対象に関する多様な情報を収集すること。
3. 同一の感覚器を多様な形態で利用することにより、生物的、物理的な制約の下で種々の情報を獲得すること。

の三点に集約することができる。

さらに、感覚統合の過程は次の二つの段階に分けることができる (図 2.2)。

具体的な例として、視覚情報処理過程を取り上げて考える。網膜で捉えられた映像情報は、外側膝状体を経て視覚一次野に送られ、そこでさまざまな特徴抽出がなされる。ここで行なわれる特徴抽出は、複数の網膜神経節細胞からの情報を統合することによってなされるので、一種の感覚統合過程である。このような低次の情報処理過程は、人間の意識に上ることなく、多数の神経細胞によって並列的、独立的に実行される。また、上丘や大脳 MST 野などで見出される poly-sensory な細胞 (ある特定の領域に与えられた視覚、聴覚、触覚のいずれの刺激にも反応する細胞) も、このようなボトムアップな統合メカニズムによってその活動が定められているといえる。

一方、人間は、眼球を動かしたり、同一視野内で注意を動かしたりすることによって、いろいろな位置にある情報を逐次的に組み合わせて対象を理解する。この過程もまた、複数の情報を統合するという意味で感覚統合過程である。しかし、この過程は前者と異なり、どこに眼球を向けるか、どの部分に注意を向けるかが意識的に定められる<sup>1)</sup>。したがって、後者の過程は、逐次

<sup>1)</sup>ただし、注意を向ける過程が常にトップダウンな過程であるとは限らない。

的、意図的である点で前者と大きく性質が異なる。なお、後者の過程は、すでに並列に捉えている情報の中からその一部だけを取り出す形式（注意を動かす場合）と、同時に捉えることのできない情報を逐次的に観測する形式（眼球を動かす場合）の二つに分類できる。

このように、感覚統合の過程は、

1. 複数の感覚器で捉えた情報を並列的、受動的に統合する過程
2. 注意の移動や観測を逐次的に行なって、情報を直列的、能動的に統合する過程

の二つの段階からなると考えられる（あるいは、前者だけで統合が完了する場合も考えられる）。Neisser [8] が唱えたように、知覚過程を、注意を向けなくとも処理が行なわれる「前注意過程 (preattentive process)」と注意を集中することによって処理を進める「注意集中過程 (focal attention process)」とに分類すれば、前者は前注意過程における情報統合、後者は注意集中過程における情報統合に相当するといえる。

本論文で中心的に論じるのは、後者の段階である。前者に関しては、6.2節において内部表現の獲得という視点から考察する。なお、修士論文 [89] で論じた統合的空間地図形成のモデルは前者を扱ったものである。

### 2.1.2 能動的認識

前節で述べたように、感覚統合は、人間が対象を理解する際に、多種多様な情報を収集するメカニズムである。しかし、実際に感覚統合を行なう際には次のような問題が生じる。

まず、同時に多数の感覚器から情報を受け取ったとしても、それらをすべて処理するためには膨大な計算資源、すなわち神経細胞が必要であり、脳の限られた計算資源でそれらをすべて同時に処理することは実質的に不可能であることである [70]。実際、「眼には入っているのに気がつかない」といった現象がおこるのは、人間が自分の取り込んだ情報をすべて処理していないことの傍証である。また、自動車の運転者の有効視野（視野のうち実際に認知に関与する部分）の大きさが道路の状況に応じて変化する。すなわち、交差点付近など情景が複雑なときには有効視野が狭まり、高速道路など情景が単純なときには広がるという事実は、視覚の情報処理（あるいは通信）容量が一定のレベルで制限されていることを示唆している [72]。

また、一つの感覚器をさまざまな形で利用する場合においても、実際にどのような観測行為をとるかという選択肢は無数に存在する。例えば、いろいろな視点から眺めることによって、対象の立体形状を知ろうとする際、視点すなわち観測位置をどこに定めるかの解は無数にある。しかし、これらの観測位置をすべて試して対象を眺めることは、非効率であるばかりでなく実現不能である。現実には、人間は、適切な視点をいくつか選択し、それによって得た情報をもとに対象の立体像を把握するのである。

このような適切な観測行為の選択という機能は、観測の効率を向上させるという意味だけでなく、目的的な情報処理を行なうためにも不可欠な機能である。なぜならば、仮に感覚器が捉えた情報がすべて中枢に伝えられ意識に上ったとすれば、おそらく脳は、自分の目的と全く関係のない情報が満たされてしまい、意味のある情報処理が実現できないことになるからである。

脳からみれば、各受容器や各観測行為は、すべて対象に関する情報を提供する「情報源」であると考えられる。脳は、対象の内部像を構成するにあたり、無数の情報源の中から適切なものを選択し、そこから得られた情報を統合していく。感覚統合を進めるには、内部像の構成がスムーズに進行するような情報源を選択することが不可欠であり、そのためのメカニズムが能動的認識であると考えられる。すなわち、感覚統合は、能動的認識のメカニズムを伴ってはじめてその優れた機能を発揮できるのである。

さて、どのような感受器あるいは観測行為を選択するかは、人間がそのときに何を知らうとしているかに依存している。例えば、対象の重さを知ろうとすれば、対象をつかんで持ち上げるといふ行為を選択し、また、対象が腐敗しているかどうかを知ろうとすれば、対象の匂いを嗅ぐという行為を選択する。さらに、選択される行為は、人間がその時点で対象についてどれだけの情報をもっているかにも依存する。対象の位置を知る際に、あらかじめ知りたい位置がわかっている場合は、その位置の周辺を集中的に探索するであろうし、全くわかっていない場合は、自分の周囲を一様に探索することになる。

このように、能動的認識は、「認識の目的」と「その時点での内部像」に依存して決まる過程であるといえる。視点を変えて見れば、能動的認識とは、自分にとって不要な情報処理を積極的に「さぼる」ことによって、有限な情報処理機構を目的の動作だけに集中するメカニズムであるともいえるだろう。

「認識の目的」とは、何に関する内部像を構成するかを定めるものであり、その意味で「内部像の場」を定めるものであるといえる。例えば、対象の材質認識が目的であれば、自分の知っている材質の集合が内部像の場になる。人間が日常的に活動する際には、内部像の場は時々刻々変化していくと考えられるが、本論文では内部像の場はあらかじめ設定されていると考え、認識の目的がいかにして与えられるかという問題には踏み込まない。

### 2.1.3 知識に基づく内部像の形成

上で述べてきたように、感覚統合や能動的認識の過程について議論する際には、「内部像を介して対象を理解する」という考え方が本質的である。すなわち、以上の議論は、「人間が対象を認識するときには、対象に関する完全な情報をもとに理解するのではなく、感覚情報をもとに脳内に構成した内部像を拠り所に対象を理解する」という考え方に基いている（あるいは、形成された内部像そのものが対象を理解した結果であると考えられることもできる）。このような考え方の下で、能動的認識は、真の対象に忠実な内部像をなるべく効率的に構成するためのアルゴリズムであると捉えられる。

ところで、内部像は対象の表象であるから、元来、対象と一対一対応の関係にあると考えべきものであろう。しかし、内部像はあくまで脳の中に一時的に保持されている「像」にすぎないから、常に対象と対応関係にあるわけではなく、対象から乖離することがありうる。このように内部像と対象が別個のものであることを認めることによって、実際の観測によらずに脳の中だけで内部像を形成することを考えることができる。これは、過去の経験から得られた対象に関する

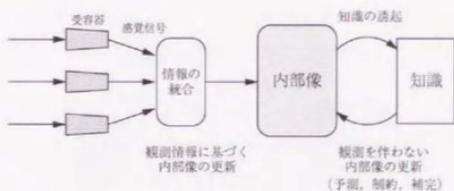


図 2.3 知識による内部像の形成

知識を用いて、観測を伴わずに内部像を形成するに相当する(図2.3)。このようないわば「知識による内部像の形成」は、限られた感覚情報を補完して対象を理解する上で重要である。すなわち、対象に関する知識を活用することによって、無駄な観測行為を省いたり状況を速やかに判断したりすることができるのである。また、逆に考えれば、知識だけからでは導くことのできない情報を得ることに観測行為を集中しているということもできる。

さらに、内部像を介した対象の理解という考え方は、いわゆるモデルベースト(model based)な認識を自然に包含している。

認識の過程は、一般に次のような形で定式化されることが多い。すなわち、対象の状態や性質が何らかの物理現象として現れ、それが感覚器によって捉えられる。脳は、獲得した感覚情報をもとにして対象の状態を求めることによって対象を認識する。対象の状態から感覚信号が定まる過程が因果律に従った「順方向」の過程であるのに対し、認識の過程は、因果関係を遡って対象の性質を求めなければならないことから、「逆問題を解く」過程として捉えられる(図2.4)。このとき一般に、感覚器の捉える情報と対象の状態との間には一対多の関係があるため、感覚情報だけでは一意に対象の状態を定めることができない。このようないわゆる解不定問題において解を一つに定めるためには、解に関する制約を与えること、すなわち、対象のもつ性質を利用して解の空間を狭めたり、文脈や知識に基づいて解を絞り込んだりすることが必要である。このように、認識の解を得るために外から与える枠組みがモデルであり、与えられたモデルに基づいて認識結果を得るのがモデルベーストな認識である。

知識による内部像の形成は、解に関する制約を付加するメカニズムの一つであることから、これをモデルベーストな認識過程として捉えることもできよう。しかし、これを内部像を「形成する」過程として捉えることによって、さらに次のような機能を考えることができる。

知識に基づいて内部像を構成する際には、構成した内部像が真の対象からかけ離れたものになっていないかどうかを検証する必要がある。内部像が与えられると、内部像と感覚信号との関係から、感覚器で実際に観測した際のような信号が観測されるかを内部像に基づく「順方向」の計算によって予測することができる(図2.4)。そして、このようにして予測した信号が実際に観測した信号と一致しているかどうかを調べることによって、内部像の「検定」を行なうことができる。すなわち、予測と実測との間に大きな差異が生じた場合には、内部像が対象と異

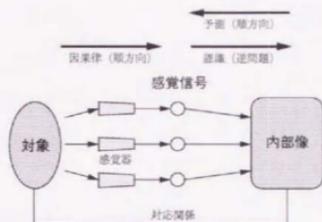


図 2.4 認識における順方向と逆方向の計算

なっていることを意味するので、それをもとにして積極的に内部像を修正することができるのである。このように、知識に基づく推定と観測による検定を繰り返すことによって、内部像を確かなものにしていくことができる。このように予測と観測が相互作用しながら内部像を形成するメカニズムについては、第4章の図形認識システムにおいて具体的に扱う。

#### 2.1.4 内部像に基づく制御

さて、内部像は認識過程のみならず、運動発現過程においても重要な役割を演じる。

人間がある運動をしようとするとき、脳はこれから出力する運動指令に対して自分の身体がどのように動くかを予測した上で、目標とする行動を間違いなく実現するような運動指令を選択している。そして、この予測を行なう上で、脳は自分の内部に構成した運動器官に関する内部像を利用している。

しかし、筋骨格系の力学的特性が代謝や疲労によって刻々と変化していることを考えると、事前に自分の運動器官の特性を完全に知った上で運動結果を予測しているとは考えにくい。むしろ、脳は、運動器官に関する内部像が一定の誤差を含んでいることを認めた上で、その誤差を考慮しながら運動の結果を予測していると考えられる。

第5章では、内部像がどれだけ対象を確実に表現しているかを表す「内部像の信頼度」という概念を与え、これを運動計画に反映させることによって、自分の運動器官に関する内部像が確実でない場合でも目標行動をほぼ間違いなく実現し、内部像が確実になるにつれて運動のパフォーマンスを向上させるアルゴリズムを構成する。

以上2.1節では、本論文における基本的な考え方を述べてきた。次節では、以上の議論の基づいて、感覚統合と能動的認識のメカニズムを最も一般的な形で定式化する。また、以降の章では、本節で述べた考え方を具体的な題材に対して適用し、その中で個々のアルゴリズムを具体化する。

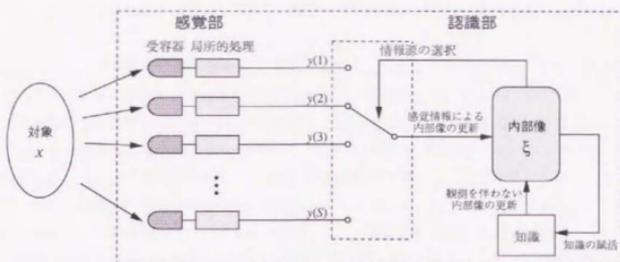


図 2.5 感覚統合と能動的認識のモデル

## 2.2 感覚統合と能動的認識の数理モデル

本節では、感覚統合と能動的認識の過程を数理モデルとして定式化し、それらが Bayes 推定と逐次実験計画としてそれぞれ表されることを述べる [94, 95].

前節で示した感覚統合の図式を抽象的に描くと、図 2.1 のようになる。図に示したように、このモデルは、外界から情報を受け取り局所的な情報処理を加える感覚部と、情報を逐次的に統合して内部像を構成する認識部に分けられる。この図において、感覚部は、感覚受容器および観測行為がそれぞれもたらす情報源の集合として表され、一方、認識部は、感覚部が提供する情報源（あるいは通信路）を選択し、得られた情報を用いて内部像を更新する推定器として表される。

以下ではまず、このモデルの下で、種々の情報源からの情報を統合するアルゴリズムについて考える。

いま、内部状態空間（内部像の場）として対象の状態集合に対応する集合  $\Xi$  を考え、 $\Xi$  の各要素  $\xi$  に対して時刻  $t$  における確率  $\pi_t(\xi)$  を割り当てる。この確率は、 $t$  回観測時点で対象がどの状態にあるらしかを表す量である。したがって、このモデルにおいて、内部像は状態空間における確率分布として実現されることになる。また、情報源  $s (= 1, 2, \dots, S)$  の集合  $\mathcal{S}$  を考え、情報源  $s \in \mathcal{S}$  が出力する値（すなわち観測値） $y(s)$  の集合を  $\mathcal{Y}_s$  と表す。また、対象が状態  $\xi$  にあるとき、情報源  $s$  が信号  $y(s)$  を出力する条件付き確率を  $p_s(y(s)|\xi)$  と表す。条件付き確率の値は、経験によってすでに得られているものと仮定する。また、情報源の間には交互作用がないものとする。

さて、時刻  $t$  に情報源  $s$  を選択した結果、信号  $y_t(s)$  が得られたとする。このとき、観測後の内部状態空間の確率分布  $\pi_t(\xi)$  は、Bayes の定理によって、

$$\pi_t(\xi; y_t(s)) = \frac{p_s(y_t(s)|\xi)}{\sum_{\eta \in \Xi} p_s(y_t(s)|\eta)\pi_{t-1}(\eta)} \pi_{t-1}(\xi) \quad (2.1)$$

として与えられる。すなわち、観測を行なうごとに内部状態空間の確率分布を更新することに

よって、種々の情報源からの情報を逐次的に統合して、対象に関する内部像を組み上げていくことができる。

このように、感覚統合の過程は Bayes 推定として定式化することができる。次に、この定式化の下で、いかなる情報源を選択すれば内部像の構成が速やかに進行するかについて考える。

この問題について考えるためには、内部像構成の進み具合を表す評価基準が必要である。ここでは、その基準として内部状態空間のエントロピーを採用する [113]。エントロピーは内部状態空間のあいまいさを表す量であり、内部像の確かさを表す量として適切な量である。ただし、エントロピーの値は内部像が対象の真の状態と一致していなくとも小さくなり得るので、この基準は内部像のいわば「確信度」を表すものであって、「正確さ」を表すものではない。

内部像の構成を速やかに進める一つの方法として、状態更新後のエントロピーの期待値が最も小さくなるような情報源を選択する方法が考えられる。いま、仮に時刻  $t$  に情報源  $s$  を選択した結果、信号  $y_t(s)$  が得られたとすると、状態更新後のエントロピー  $H_t(y_t(s))$  は、式 (2.1) で求められる  $\pi_t(\xi; y_t(s))$  を用いて、

$$H_t(y_t(s)) = - \sum_{\xi \in \Xi} \pi_t(\xi; y_t(s)) \log \pi_t(\xi; y_t(s)) \quad (2.2)$$

と表される。しかし、 $y_t(s)$  の値は、実際に観測を行なってみない限り知ることができない。そこで、ここでは、状態更新後のエントロピーの期待値  $H_t(s)$  を用いて、それを平均評価することを考える。

$$\begin{aligned} H_t(s) &= \sum_{y_t(s) \in \mathcal{Y}_s} P(y_t(s)) H_t(y_t(s)) \\ &= \sum_{y_t(s) \in \mathcal{Y}_s} \sum_{\xi \in \Xi} p_s(y_t(s)|\xi) \pi_{t-1}(\xi) H_t(y_t(s)) \end{aligned} \quad (2.3)$$

内部像の形成を速やかに進めるためには、この値が最も小さくなるような情報源を選択することが、少なくとも平均評価では最適であると考えられる。あるいは、状態更新によるエントロピーの減少量の期待値  $I_t(s)$  を基準にとり、これが最大値をとる情報源を選択しても同じことである。

$$\begin{aligned} I_t(s) &= H_{t-1} - H_t(s) \\ &= \sum_{\xi \in \Xi} \pi_{t-1}(\xi) \sum_{y_t(s) \in \mathcal{Y}_s} p_s(y_t(s)|\xi) \log \frac{p_s(y_t(s)|\xi)}{\sum_{\eta \in \Xi} p_s(y_t(s)|\eta) \pi_{t-1}(\eta)} \end{aligned} \quad (2.4)$$

この量は、情報理論において相互情報量と呼ばれるものである [7, 12]。

以上の議論では、内部状態空間を離散的な集合としたが、これを連続的な集合としても同様の議論ができる。特に、内部状態分布および条件付き確率分布が正規分布で表されるときには、更新に伴う処理が平均と共分散行列の漸化式という形で表され、また、エントロピーや相互情報量の値も簡単に計算することができる。

いま、時刻  $t$  における内部状態を  $M$  次元ベクトル  $\xi_t$  で表し、その平均を  $\mu_t$ 、共分散行列を  $D_t$  とする。また、情報源  $s$  から得られる観測値  $y_t(s)$  が、

$$y_t(s) = c_s \cdot \xi_t + v_t(s) \quad (2.5)$$

と表されるものとする。ただし、 $v_t(s)$  は平均 0、分散  $\sigma_s^2$  のガウス性雑音とする。また、異なる情報源の間で  $v_t(s)$  は無相関であるとする。

このとき、時刻  $t$  で情報源  $s$  を選択し、その結果信号  $y_t(s)$  を受け取ったとすると、内部状態分布は、

$$D_t(s) = \left[ D_{t-1}^{-1} + \frac{1}{\sigma_s^2} c_s c_s^T \right]^{-1} \quad (2.6)$$

$$\mu_t(s) = D_t(s) \left[ D_{t-1}^{-1} \mu_{t-1} + \frac{y_t(s)}{\sigma_s^2} c_s \right] \quad (2.7)$$

として更新される。また、 $M$  次元正規分布のエントロピーが、

$$H_t = \frac{M}{2} (1 + \log 2\pi) + \frac{1}{2} \log |D_t| \quad (2.8)$$

と与えられることから、時刻  $t$  で情報源  $s$  を選択したときのエントロピーの期待値は、

$$\begin{aligned} H_t(s) &= \frac{M}{2} (1 + \log 2\pi) + \frac{1}{2} \int \log |D_t(s)| P(y_t(s)) dy_t(s) \\ &= \frac{M}{2} (1 + \log 2\pi) + \frac{1}{2} \log |D_t(s)| \end{aligned} \quad (2.9)$$

として与えられる。この式の第 1 項は定数であり、また、対数関数が単調増加関数であることから、状態更新後のエントロピーの期待値を最小化する情報源を選択するためには、状態更新後の共分散行列の行列式  $|D_t(s)|$  を最小化する情報源  $s$  を選択すればよいことになる。

一方、相互情報量は、

$$\begin{aligned} I_t(s) &= H_{t-1} - H_t(s) \\ &= -\frac{1}{2} \log |D_t(s)| + \frac{1}{2} \log |D_{t-1}| \end{aligned} \quad (2.10)$$

と計算されるので、相互情報量最大化として定式化した場合も、上と同様にして状態更新後の共分散行列の行列式  $|D_t(s)|$  を最小化する情報源を選択すればよいことが導かれる<sup>1</sup>。なお、 $\frac{1}{\sigma_s^2}$  が十分小さいとき、相互情報量は、

$$I_t(s) \simeq \frac{1}{2\sigma_s^2} c_s^T D_{t-1} c_s \quad (2.11)$$

という近似式で与えられる [113]。この式を用いることにより、式 (2.6) を実行するのに比べて少ない手間で評価基準を得ることができる。

<sup>1</sup>相互情報量を最大化する逐次実験計画は、非 Bayes 統計学における  $D$ -最適な実験計画に一致することが知られている。

このように、能動的認識の過程は「相互情報量を評価基準として用いた逐次実験計画」として定式化することができる。以上で述べたアルゴリズムでは、単一の観測によって得られる平均的情報量を最大化する場合のアルゴリズムであるが、場合によっては、何回か連続して観測した時点で得られる情報を最大化したいことも考えられる。この場合は、上述の計算を再帰的に行なって「先を読む」ことによって、その時点での内部状態空間のエントロピーが最小化されるような観測手順を選択するという戦略も考えられる[91]。ただし、このような方法は、先を読む段階数が増えるに従って、探索する場合の数が急激に増加し、計算量が爆発することに注意しなければならない。観測行為選択にかかるコストが観測にかかるコストよりも大きくなってしまえば、意味がなくなってしまうからである。

また、相互情報量の代わりに対称化ダイバージェンス (J-divergence)

$$g(\xi_1, \xi_2; s) = \sum_{y \in \mathcal{Y}_s} (p_s(y(s)|\xi_1) - p_s(y(s)|\xi_2)) \log \frac{p_s(y(s)|\xi_1)}{p_s(y(s)|\xi_2)} \quad (2.12)$$

や、Hellinger 距離

$$h(\xi_1, \xi_2; s) = \sum_{y \in \mathcal{Y}_s} \sqrt{p_s(y(s)|\xi_1) p_s(y(s)|\xi_2)} \quad (2.13)$$

などの確率分布間距離を基準量として用いることもできる[118]。対称化ダイバージェンスを用いた場合、情報源の選択基準は、上式で定めた  $g(\xi_1, \xi_2; s)$  を用いて、

$$J_I(s) = \sum_{\xi_1, \xi_2 \in \Xi} \pi_I(\xi_1) \pi_I(\xi_2) g(\xi_1, \xi_2; s) \quad (2.14)$$

で与えられる。この式において、 $g(\xi_1, \xi_2; s)$  の値は確率分布  $\pi_I(\xi)$  によらずに定まるので、この値を表として保持しておくことにより、評価基準を求める際の計算量を減らすことができる。

### 2.3 ボトムアップな感覚器選択メカニズム

前節で述べたアルゴリズムは、内部像のあいまいさの減少量が最も大きいと期待される情報源を選択するという戦略を実現したものであった。しかし、このような方法では、選択すべき情報源の数が増加するに従って、それらすべてを探索することが事実上不可能になる。このような場合、選択すべき情報源の数を減少させる一つの方法として、互いに相関の強い情報源をあらかじめグループ分けし、その中から代表的な情報源を選ぶという方法が考えられる。この方法は、認識メカニズムを構成する段階で情報源をグループ分けするという点で静的な方法であるといえる。一方、これに対する動的なメカニズムとして、認識部がトップダウンに情報源を選択するだけでなく、情報源の方からボトムアップに自分を選択するように「立候補」するようなメカニズムを考えることができる。本節では、網膜に映った画像情報の一部が取り出されて処理される場合のように、感覚情報が同時に取り込まれた後その一部が選択される状況を考え、その中で後者のメカニズムについて考える。

さて、情報源の方から立候補するためには、各情報源が独立して計算できる評価基準を用いる必要がある。ここでは、各情報源が内部像を構成する上で有効な情報を供給するように、以下に挙げるような意味で有用な情報源を選択することを考える。

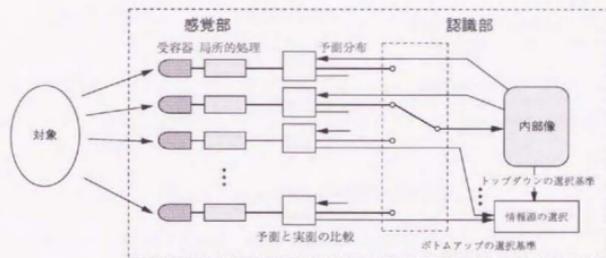


図 2.6 ボトムアップな情報源選択を考えるためのモデル

1. 非日常的信号 普段あまり経験することのない信号は、特別な意味をもっている可能性が高い。
2. 予測との差異 内部像をもとに予測された信号と実測した信号との差が大きな場合は、予測が誤っていること、あるいは、内部像が真の対象と遊離していることを意味しているので、そのような情報は内部像を訂正する上で重要である。
3. 予測の一様性 内部像に基づく予測だけではその信号値が絞れないような情報源は、内部像を形成する上で新たな情報を提供する可能性が大きい。ただし、その情報源のもつ信号が内部像と無関係である場合には、これは成り立たない。
4. 空間的非一様性 信号が空間的に分布しているような場合、その中の一部だけが他の部分と比べて大きく異なるときは、何か特別なものがそこにあることを意味している。例えば、一様なテキストチャの中に異質なものが一つだけあるとき、それが何であるかを知ることは重要である。

非日常的な信号とは普段予測されない信号であると解釈できるので、最初の二つは同じ枠組みで説明することができる。四つ目の条件は、空間全体にわたって一定の規則性や構造が見いだされたとき、その構造から逸脱している部分を選択的に観測するという点で、ゲシュタルト(Gestalt)の考え方と関係するものである。ここでは四つ目の条件については考えず、残りの三つの条件について議論する。

以上の問題について考えるために、図 2.5 のモデルに若干変更を加えて、図 2.6 のモデルを設ける。このモデルには、内部像から各情報源に対してどのような信号が来るかを予測する信号と、予測した信号と観測した信号を比較する比較器が新たに加えられている。内部像から各情報源への予測は、情報源ごとに独立して行なうことができるので、各情報源に付随した比較器は並列的に動作すると考えてよい。

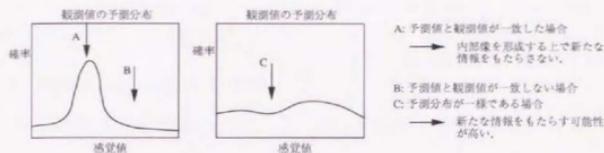


図 2.7 ボトムアップな情報源選択基準

信号の値を予測する分布の計算は、条件付き確率分布を用いて次のように行なう。

$$P_i(\hat{y}(s)) = \sum_{\xi \in \Xi} p_{s_i}(\hat{y}(s)|\xi) \pi_i(\xi) \quad (2.15)$$

ただし、 $\hat{y}(s)$  は情報源  $s$  から得られる信号の予測値を表す。このようにして得られた分布のエントロピーが大きな情報源  $s$  は、上記の三番目の条件に適合する (図 2.7C)。

次に、比較器では、このようにして得られた予測分布と各感覚器が捉えた実測値とを比較する。このとき、観測値に対する予測確率の値が十分に大きければ、観測値が予測値と一致していたことを意味するので、その観測値は新たな情報を担っていないことになる。一方、観測値に対する確率が小さな感覚器については、その予測が間違っていることを意味するので、内部像を更新する上で有用な情報を担っているものと考えられる。これは、上記の二番目の条件に適合する (図 2.7B)。

以上の考察から、次の二つの評価基準を考えることができる。

$$K(s) = - \sum_{\hat{y}(s) \in \mathcal{Y}_s} P(\hat{y}(s)) \log P(\hat{y}(s)) \quad (2.16)$$

$$L(s) = P(\hat{y}(s) = y(s)) \quad (2.17)$$

ただし、 $\hat{y}(s)$  は予測値、 $y(s)$  は実測値である。これらの評価基準が大きな値をとる情報源が立候補することによって、認識部は選択すべき情報源の選択肢を絞り込むことができる。

## 2.4 内部像の予測を伴う能動的認識

前節までは、静止している対象の認識という時間的な変化を伴わない問題について議論を行ってきた。しかし、対象が運動している場合や制御を伴う問題の場合には、時間的に変化する対象の状態を予測しながら対象を理解する必要がある。本節では、カルマンフィルタ (Kalman filter) の議論を参考にして、そのような場合について議論する。なお、この過程は、対象の時間的な変化に関する知識を用いて内部像を形成する過程として捉えることができる。

カルマンフィルタは、対象の状態や雑音が正規分布に従う条件の下で用いられる一種の最適予測フィルタである [11, 48]。カルマンフィルタの状態推定アルゴリズムは、時間の進行に伴う対象の状態変化を予測する状態予測過程と、観測情報を用いて内部像を更新する観測更新過程に分

けられるが、いずれの過程も Bayes 推定の考え方に基づいて構成されているので、前節までの議論から容易に移行することができる。

#### 2.4.1 状態変化モデルが単一である場合

いま、時間的に変化する対象  $x_t$  があり、これがある規則  $f(x)$  に従って変化しているとする。

$$x_{t+1} = f(x_t) \quad (2.18)$$

システムの内部では、これを

$$x_{t+1} = Ax_t + Bu_t \quad (2.19)$$

という線形モデルを用いて表現する。ここで、 $u_t$  は平均 0、共分散行列  $Q$  のガウス性雑音である。以下、 $x_t$  に対する内部像を  $\xi_t$  をもって表す。一方、センサー  $s$  による観測値  $y_t(s)$  は、

$$y_t(s) = c_s \cdot x_t + v_t(s) \quad (2.20)$$

で与えられるとする。ただし、 $v_t(s)$  は平均 0、分散  $\sigma_s^2$  のガウス性雑音で、 $u_t$  と独立とする。 $u_t$  と  $v_t(s)$  は、それぞれ予測と観測に伴う誤差であると考えられることができる。なお、2.2 節のモデルは、式 (2.19) において  $A = I, B = 0$  とした場合に相当する。

ここで、 $x_t$  の振舞いのある決められた精度でなるべく少ない観測回数で知りたいという状況を考える。前節までの議論に従い、本節でも、システムは内部像  $\xi_t$  をもって対象の状態  $x_t$  の振舞いを把握するとし、また、内部像が対象をどれだけ確実に表しているかを内部像のあいまいさをもって評価するものとする。

このとき、内部像のあいまいさが定められた精度に比べて十分小さいときには、システムはわざわざ対象を観測しなくとも、自分の内部で予測を繰り返すことによって、対象の状態をある程度の誤差の範囲で把握し続けることができる。しかし、一般に状態変化モデルは対象の性質を完全に記述しているわけではないので、予測を繰り返すに従って、内部像のあいまいさは次第に増加し、やがて定められた許容限度を超える。

内部像のあいまいさが許容限度を超えたときには、その時点であいまいさが最も小さくなるような情報源を選択することによって、観測の効率を最適化することができる。このことを実現するためには、2.2 節の議論と同様に、観測後の内部状態空間のエントロピーが最小になることが期待される情報源を選択すればよい。ここでは、そのための基準として内部像と情報源との相互情報量を取り、これが最大になる情報源を選択する。相互情報量基準は、式 (2.10)、あるいは式 (2.11) によって求められる。

以上の議論をまとめると、次のようなアルゴリズムが得られる。

1. まず、以下の式に従って、内部状態分布を更新する。

$$\mu_t^- = A\mu_{t-1}^+ \quad (2.21)$$

$$D_t^- = AD_{t-1}^+A^T + BQB^T \quad (2.22)$$

この式において、添字の記号  $-$ ,  $+$  はそれぞれその時刻での観測前、観測後を表す。

2. 更新した時点で、内部状態空間のエントロピー（あるいは共分散行列の行列式）が定められた値を超えた場合は、式(2.10)または式(2.11)に従って最適な情報源を選択し、対象を観測する。エントロピーが定められた範囲に収まっている場合は、内部像は十分確かであるとして観測を行わない。

3. 情報源  $s$  から得た信号  $y_t(s)$  を用いて内部状態を更新する。

$$\mu_t^+ = \mu_t^- + [y_t(s) - c_s \cdot \mu_t^-] k_t \quad (2.23)$$

$$D_t^+ = D_t^- - k_t c_s^T D_t^- \quad (2.24)$$

$$k_t = [c_s^T D_t^- c_s + \sigma_s^2]^{-1} c_s \quad (2.25)$$

ここで、 $k_t$  はカルマンゲイン (Kalman gain) と呼ばれる中間変数である。時刻  $t$  で観測を行わないときは、観測前の分布をもって観測後の分布とする。すなわち、この場合は、

$$\mu_t^+ = \mu_t^- \quad (2.26)$$

$$D_t^+ = D_t^- \quad (2.27)$$

となる。

4. 1. へもどる。

以上が、対象の状態が時間的に変化する場合の能動的認識のアルゴリズムである。

一般のカルマンフィルタでは、サンプリング周期一定という条件の下で状態予測誤差を最小化する形で定式化されている。これに対し、本節のアルゴリズムでは、対象の状態をある決められた精度で推定できればよいという状況の下で、必要な観測精度に応じて観測頻度を能動的に設定する形で定式化したことになる。

#### 2.4.2 状態変化モデルが複数ある場合

前節では、対象の状態変化を示すモデルが一つあり、そのモデルに基づいて対象の状態を予測するという問題について議論した。次に、状態変化モデルが複数ある場合に、そのモデルの選択も含めた状態予測メカニズムについて検討する。

いま、状態変化モデルの集合を  $\{f_{\alpha}\}$  とし、対象の状態はその内のいずれかに従って変化しているものとする。また、各状態変化モデルは次のような線形モデルであるとする。

$$\xi_{t+1}^{\alpha} = A_{\alpha} \xi_t^{\alpha} + B_{\alpha} u_t \quad (2.28)$$

ここで、 $\xi_t^{\alpha}$  は状態変化モデル  $f_{\alpha}$  に関する内部状態、 $u_t$  は平均 0、共分散行列  $Q$  のガウス性雑音である。また、各モデルにおいて観測値は、

$$y_t^{\alpha}(s) = c_s^{\alpha} \cdot \xi_t^{\alpha} + v_t(s) \quad (2.29)$$

によって与えられるとする。ただし、 $v_t(s)$  は平均 0、分散  $\sigma_s^2$  のガウス性雑音で、 $u_t$  と独立である。また、 $\xi_t^{\alpha}$  の平均と共分散行列を、それぞれ  $\mu_t^{\alpha}$ 、 $D_t^{\alpha}$  と表す。

次に、時刻  $t$  において対象の状態変化がモデル  $f^\alpha$  に従っている確からしさを  $\pi_t(\alpha)$  と表す。このとき、システム全体としての観測値  $y_t$  に対する予測値  $\eta_t$  は、各モデルに関する出力を混合することによって得られる。

$$\eta_t = \sum_{\alpha} \pi_t(\alpha) c_{\alpha}^{\alpha} \cdot \mu_t^{\alpha} \quad (2.30)$$

以上の準備の下で、状態の予測と観測による更新のアルゴリズムを記述する。

まず、状態予測の操作は各状態変化モデルに対して独立に行ない、それを各時刻で組み合わせることによって新しい内部像を得る。

$$\mu_t^{\alpha+} = A_{\alpha} \mu_{t-1}^{\alpha+} \quad (2.31)$$

$$D_t^{\alpha+} = A_{\alpha} D_{t-1}^{\alpha+} A_{\alpha}^T + B_{\alpha} Q B_{\alpha}^T \quad (2.32)$$

一方、観測更新の操作は、各状態変化モデルに対する内部像の更新とモデル選択に関する分布  $\pi_t(\alpha)$  の更新の二つの部分から構成される。まず、各モデルに対する内部像の更新は、各モデルに対しそれぞれ次式に従って行なう。

$$\mu_t^{\alpha+} = \mu_t^{\alpha-} + [y_t(s) - c_{\alpha}^{\alpha} \cdot \mu_t^{\alpha-}] k_t^{\alpha} \quad (2.33)$$

$$D_t^{\alpha+} = D_t^{\alpha-} - k_t^{\alpha} c_{\alpha}^{\alpha T} D_t^{\alpha-} \quad (2.34)$$

$$k_t^{\alpha} = [c_{\alpha}^{\alpha T} D_t^{\alpha-} c_{\alpha}^{\alpha} + \sigma_s^2]^{-1} D_t^{\alpha-} c_{\alpha}^{\alpha} \quad (2.35)$$

ここで、 $y_t(s)$  は観測値である。一方、 $\pi_t(\alpha)$  の更新は、

$$\pi_t(\alpha) = \frac{P(y_t(s)|\alpha)}{\sum_{\beta} P(y_t(s)|\beta) \pi_{t-1}(\beta)} \pi_{t-1}(\alpha) \quad (2.36)$$

として行なう。なお、 $P(y_t(s)|\alpha)$  は次式で与えられる。

$$\begin{aligned} P(y_t(s)|\alpha) &= \int P(y_t|\xi_t^{\alpha}) \pi(\xi_t^{\alpha}) d\xi_t^{\alpha} \\ &= [2\pi \{ \sigma_s^2 + (d_t^{\alpha})^2 \}]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{e_t^{\alpha}}{\sigma_s^2 + (d_t^{\alpha})^2} \right\} \end{aligned} \quad (2.37)$$

ただし、

$$e_t^{\alpha}(s) = y_t(s) - c_{\alpha}^{\alpha} \cdot \mu_t^{\alpha} \quad (2.38)$$

$$d_t^{\alpha}(s) = c_{\alpha}^{\alpha T} D_t^{\alpha-} c_{\alpha}^{\alpha} \quad (2.39)$$

である。時刻  $t$  で観測を行わないときは、 $\pi_t(\alpha)$  の値は更新せずに古い値を継承する。

さて、時刻  $t$  で観測を行なうかどうか、また、観測を行なう場合にどの情報源を選択するかは、それぞれ内部状態空間のエントロピーと相互情報量を評価基準として決定する。状態変化モデルが複数ある場合の内部状態空間のエントロピーは、

$$H_t = \sum_{\alpha} \pi_t(\alpha) \left\{ - \sum_{\xi \in \Xi} \pi_t(\xi|\alpha) \log \pi_t(\xi|\alpha) \right\} - \sum_{\alpha} \pi_t(\alpha) \log \pi_t(\alpha) \quad (2.40)$$

すなわち、各モデルのもつエントロピーの荷重和とモデル集合のエントロピーの和として与えられるので、時刻  $t$  において内部状態分布を予測したとき、この値が定められた閾値以下であれば、時刻  $t$  では観測を行わないことにする。逆に、エントロピーが閾値を超えたときは、次式で近似的に与えられる相互情報量基準が最も大きな情報源を選択して観測を行なう。

$$I_t(s) \simeq \frac{1}{2\sigma_s^2} \left[ \sum_{\alpha} \pi_t(\alpha) \left\{ c_s^{\alpha} \cdot \mu_s^{\alpha} - \sum_{\beta} \pi_t(\beta) c_s^{\beta} \cdot \mu_s^{\beta} \right\} + \sum_{\alpha} \pi_t(\alpha) d_t^{\alpha}(s) \right] \quad (2.41)$$

このように、相互情報量基準は、モデル選択に関する部分（第1項）と各モデルに対する内部像に関する部分（第2項）の和として与えられる。この式は、いずれか一つのモデルについて  $\pi_t(\alpha) = 1$  となり、他のモデル  $\beta (\neq \alpha)$  に対して  $\pi_t(\beta) = 0$  となれば、式(2.11)と一致する。したがって、状態変化モデルの判別が進行し最終的にモデルが一つに絞られた時点で、このシステムの振舞いは状態変化モデルが単一である場合に帰着される。

なお、以上のアルゴリズムを運動する物体の位置をカメラで追跡する問題に適用したところ、予想通りの性能が得られることが数値実験によって示されている [92]。

対象の動きが的確に予測できるようになると対象に注意を払わなくなることは、日常生活でよく経験することである。これは、人間が同時に複数の目的をもって行動するためでもあり、また、人間の「好奇心」が移り変わるためでもある。いずれにせよ、人間は、一つの問題に脳のすべての資源を投入するのではなく、感覚情報の取込みやそれに伴う情報処理を省略して、限られた情報処理資源をさまざまな情報処理に振り分けていることに間違いない。複数の事柄を同時に気にかけてながら仕事をしているとき、一方のことが適度に予測できれば他方の処理に注意を集中できるといったこともまた、同じメカニズムの上に実現されていると思われる。本節で述べたアルゴリズムは、人間のこのような振舞いを説明する原理モデルであるといえるだろう。

## 第 3 章

### 能動的認識アルゴリズムに基づく触知覚認識システム

本章では、前章で構成したアルゴリズムを触知覚認識システムに適用し、その性質について実験的に検討する。

対象の質感や手ざわりを知ろうとするとき、人間は、皮膚の機械受容器だけでなく、温覚受容器や関節内の固有受容器など種々の受容器からの情報を統合することによって総合的な判断を下す。また、人間は、押す、こするといったさまざまな触運動を行なって、識別に有用な情報を収集する [22, 42, 63]。このような触覚に関わる総合的な感覚を、皮膚刺激によって引き起こされる狭義の触覚 (tactile sense) と区別して、触知覚 (haptics) と呼ぶ [23]。この定義からわかるように、触知覚はまさに感覚統合、能動的認識の場であるといえる。

筆者は、このような触知覚過程の特徴に着目し、前章で述べたアルゴリズムに基づいて、種々の触運動を選択的に行ないながら複数のセンサからの情報を統合して対象の材質や手ざわり感を識別するシステムを構成した。

以下、3.1節で触覚や温覚に関する生理学的知見について、3.2節で従来の触覚センサについてそれぞれ簡単にまとめたのち、3.3節で製作したシステムの構成や動作について述べる。また、3.4節では、このシステムを用いて触知覚感の定量化を試みた結果について述べる。

#### 3.1 触覚の生理学

本節では、文献 [41, 104] に従って触覚および温覚に関する生理学的知見についてまとめる。

##### 3.1.1 触覚受容器の構造と機能

皮膚は、外界に近い方より表皮 (epidermis)、真皮 (corium)、皮下組織 (subcutaneous tissue) の3層からなり、表皮はさらに3ないし5層から構成されている (図 3.1)。

受容器の主なものとして、表皮の最深部にあるメルケル細胞 (Merkel cell)、真皮の最外層にあるマイスナー小体 (Meissner corpuscle)、深層にあるパチニ小体 (Pacini corpuscle)、ルフィニ終末 (Ruffini ending) がある。このほか、有毛部皮膚では、毛包に神経終末がからみついた毛包受容器 (hair follicle receptor)、メルケル細胞が集合してできたピンカス (Pinkus) の毛盤な

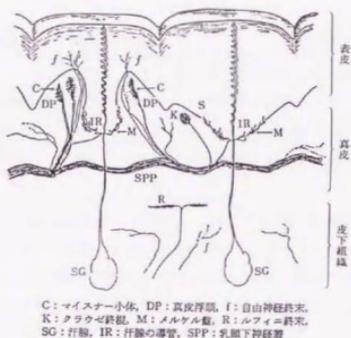


図 3.1 皮膚の機械受容器 (文献 [41] より引用)

どがある。また、腱や皮下組織にはパチニ小体によく似たゴルジ・マツツォニ小体 (Golgi-Mazzoni corpuscle) がある。これらの受容器は、神経繊維の終端が特殊化したものであり、それ自身が外界からの機械的刺激を神経膜の興奮に変えるエネルギー変換器となっている。

ヒトの掌には約 17,000 個の感覚受容器があるが、矩形波状に持続する皮膚の変形刺激に対する応答の順応からそれらを二つに区別することができる。一つは、刺激のある間応答がほぼ持続するもので、遅順応型 (SA 型: Slowly Adapting) と呼ばれる。もう一つは、刺激の始めと終わりに応答するもので、速順応型 (RA 型: Rapidly Adapting) と呼ばれる。数の割合は、前者が 44%、後者が 56% である。

これらは、さらに受容野の大きさにより I 型、II 型に分けられる。I 型の受容野はその大きさが小さく境界も鮮明であるのに対し、II 型の受容野は広く境界が不鮮明である。上述の受容器と対応させると、RAI タイプはマイスナー小体、SAI タイプはメルケル細胞、RAII タイプはパチニ小体、SAII タイプはルフィニ終末であることがわかっている。

RAI, SAI タイプは、直径 2~4mm のほぼ円形の受容野をもち、その中に感度が最大になる点が複数個ある。また、受容野の境界は鮮明で、受容野をはずれると急激に感度が落ちる。これらの受容器は、皮膚の部位によってその分布密度が大きく異なり、特に指先での分布密度が高い。例えば、RAI タイプの密度は、掌で 25 個/cm<sup>2</sup> であるのに対し指先で 140 個/cm<sup>2</sup>、また、SAI タイプは掌で 10 個/cm<sup>2</sup> であるのに対し指先では 70 個/cm<sup>2</sup> である。また、これらの受容器は、一般に鋭いエッジに対して良く反応する。

一方、RAII, SAII タイプは、受容野のほぼ中心に感度最大点が一点だけあり、周辺にむかって感度がゆるやかに減少している。受容野が広く境界が不鮮明であるのは、これらの受容器が皮下深くにあるためである。また、分布密度は指先、掌によらず一定であり、全体として密度は

表 3.1 触覚受容器の特性

タイプ	受容器	位置	受容野	応答特性
RAI	マイスナー小体	浅部	小さい	周波数が40Hz付近の振動によく反応する。
SAI	メルケル細胞	浅部	小さい	一定圧刺激に反応する。
RAII	パチニ小体	深部	大きい	100-300Hz付近の振動によく反応する。
SAII	ルフィニ終末	深部	大きい	皮膚の引っ張りに反応する。

低い。RAIIタイプは皮膚の振動によく反応する。特に100~300Hzの振動に対する感受性が高く、この範囲では1 $\mu$ m程度の皮膚の変形で十分興奮する。なお、100Hzより低い周波数の振動に対してはRAIタイプの方がよく反応し、特に40Hzくらいの振動に対し最も良く反応する。SAIIタイプは、振動よりも皮膚の持続的変形、つまり引っ張りなどに良く反応する。この受容器は、対象を握る時の制御などに利用されていると考えられている。

これらの機械受容器の特性を表3.1にまとめた。

### 3.1.2 温度受容器の構造と機能

温度感覚は、温度刺激に対する温度受容器の活動によって引き起こされる。温度受容器は、次のような性質をもっていることが知られている。

1. 皮膚温度が一定であれば、活動の大きさは一定である（静的反応）。
2. 皮膚温度が変化すると、細胞の興奮程度は上昇あるいは下降する（動的反応）。
3. 熱に関係しない刺激に対しては無反応である。
4. 皮膚への温度刺激に対して、感覚の閾値と受容器の閾値はほぼ同じである。
5. 受容野の大きさは1mm<sup>2</sup>以下であり、一つの神経線維は1個あるいは2、3個の温点、あるいは冷点につながっているにすぎない。
6. 求心性線維の伝達速度は高々20m/sにすぎず、遅いものでは0.4m/sのものもある。

温覚と冷覚に対する受容器は別々に存在し、それぞれ温線維と冷線維と呼ばれる。その実体は完全には明らかになっていないが、おそらく自由神経終末であると考えられている。冷線維は表皮のすぐ下に、温線維は真皮の中層や上層に数多く見られる。また、冷線維はミエリン鞘をもつが、温線維はもたない。

一定温度の刺激に対して、温線維、冷線維はいずれも最も反応しやすい温度を中心にしてベル型の反応特性を示す。中心温度は、冷線維については約30°C（17°Cから36°Cの範囲）、温線維については約43°C（41°Cから47°Cの範囲）である。31°Cから36°Cの範囲ではいずれの受容器も神経パルスを出す。この範囲では人間は温かいとも冷たいとも感じない。

皮膚の温度が36°C以上になると、人間は常に暖かいと感じる。これは、43°Cまでの範囲では、温線維が温度の上昇とともに受容器の活動が活発になるためである。しかし、31°C以下の

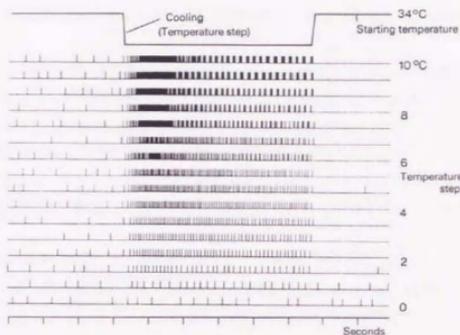


図 3.2 冷線維の過渡的な応答 (文献 [104] より引用)

冷温に対してはそれほど単純ではなく、 $25^{\circ}\text{C}$  と  $33^{\circ}\text{C}$  の間では冷線維の活動がほとんど変化しないにもかかわらず、低い温度に対しては冷たく感じ、高い温度では何も感じない。このような感覚が生じるのは、中枢神経系が何らかの補償をしているためであると考えられている。なお、受容器レベルでの適応は数秒以内に終了するのに対し、主観的な温度の適応は何分にもわたることが知られており、温度感覚に中枢神経系が関与していることはほぼ間違いない。実際、求心性神経の活動変化によって引き起こされた中枢神経系の活動は非常にゆっくりと減衰していく。

皮膚温度の変化に対する冷線維の反応を図 3.2 に示す。この信号は、サルの正中神経から記録したものである。この図からわかるように、皮膚温度を急激に冷やすと、その後数秒間にわたって過渡的な応答を示し、やがて安定状態に到達する。過渡応答の大きさは、温度変化の大きさによって定まる。次に、皮膚を元の温度にもどすと、しばらくの間興奮が停止したのち、もとの状態に戻る。温線維の反応はこれと鏡像関係にある。

温度に関する特別な現象として、次のようなものがある。一つは、温度刺激を取り除いても温度感が引き続いて感じられる現象 (after-sensation) である。これは、例えば、冷たい金属を皮膚にしばらくあてておくと、それを離して皮膚の温度がもとどおりになっても、依然冷たい感覚が残るというものである。この現象は受容器レベルで確認されている。このほか、非常に高い温度の刺激を与えると、かえって冷たく感じられるという現象も見出されている。これは、冷線維が  $45^{\circ}\text{C}$  以上の刺激に対して活動するためである。なお、 $45^{\circ}\text{C}$  以上の温度に対しては「熱い」感覚が生じるが、これは、温度感覚というよりはむしろ痛覚と呼ぶべきものである。

## 3.1.3 触覚、温覚受容器の特徴

以上で述べた生理学的知見をもとに、触覚覚感認識システムを構成する上での要点をまとめておく。

触覚受容器については、4種類の機械受容器は大きく分けて次の二つの役割を担っていると考えられる。まず、SAIとRAI、RAIIの3種類は、対象に接触したときの皮膚の局所的変形や、対象表面をこすったときの皮膚の振動を捉える役割を果たしている。一方、SAIIは、皮膚の局所的な振動ではなく、皮膚全体がどちらかの方向に引っ張られたときに生じる歪みを捉えている。言い換えれば、これらの受容器は、対象と接触した部分の「マイクロな変形や振動」と、対象表面と接触面全体との摩擦による「マクロな変形」の二つの情報を獲得していることを示している。

次に、触覚受容器の密度について考えてみると、比較的高密度なRAIタイプの受容器でさえ、その分布密度は最も高い指先で百数十個/cm<sup>2</sup>程度である。このことは、触覚受容器が1mm<sup>2</sup>あたり高々1個程度しかないことを示しており、触覚受容器の密度がそれほど高くないことを示している。にもかかわらず、人間は指先で対象のかなり微妙な差異を知ることができる。このことは、このような微妙な識別が、指を単に対象に押しつけただけでは不可能であり、指を対象とこすり合わせることによって初めて可能になることを考えれば納得できる。すなわち、触覚覚感には、皮膚表面に並んだ機械受容器からの出力を面状に分布した情報として捉えるのではなく、指を動かす際に得られる種々の振動、変形情報を合成することによって得られると考えられる。さらに、人間が微妙な触覚覚感の違いを判断する際には、対象にかける力を変えたり、こする速度を変えたりするなどして、それぞれの条件のもとでの感じ方を比較するという戦略をとっている。このことから、触覚システムを実現するには、

1. 高分解能な面状センサを用いるよりも、低密度であっても数種の異なったセンサを用意すること
2. さまざまな形態、すなわち接触力やすべり速度を種々に設定して、センサを対象にこすり合わせること

が重要であるといえる。

温覚情報については、対象に触れた瞬間の皮膚温度の変化が最も大きな情報を提供すると考えられる。前節で述べたように、温覚受容器は、皮膚の急激な温度変化に対して過渡的な応答を示し、反応の大きさは温度変化の大きさに応じて定まる。皮膚表面の温度変化が対象の熱容量や熱伝導率の影響を受けることを考えれば、センサが対象に接触した瞬間の、センサ表面の温度変化の大きさや変化速度は、材質を識別する上で有用な情報をもたらすと考えられる。

ただし、対象の温度は常に一定であるとは限らず、また温度変化の大きさは対象の体積や表面形状にも依存するので、温覚情報は対象の材質を判断する上で絶対的な情報にはならない。したがって、温度変化の情報は、最終的な判断を下すためではなく、対象に触れた瞬間に対象の候補を絞り込むという補助的な役割を果たしていると考えられる。

本節では述べなかったが、触覚や温覚のほかに、対象を押ししたときの対象の変形の具合や反

力の感覚なども、対象を識別する上で有用な情報を提供する。したがって、触知覚感を得るためには、対象に圧力を印加するとともに対象の変形の大きさを計測するメカニズムも必要であると考えられる。

### 3.2 従来の触覚センサ

本節では、これまでに種々の開発されたセンサについて手短かにまとめ、その能力と限界について述べる。

1. **スイッチ** スイッチは最も単純な接触センサであり、工業的に最もよく使われている触覚センサである。レバーなどの部品と組み合わせることにより、さまざまな形で利用できる。
2. **シリコンウェハ上の圧力センサ** シリコンウェハ上にダイヤモンドを形成し、その圧力変化を観測することによって接触を知るものである [58, 110]。シリコンの微細加工技術を用いることによって、同一ウェハ上に周辺回路を集積したり、ダイヤモンドを多数並べて圧力分布を得たりすることができる。
3. **感圧導電ゴム** 感圧導電ゴムとは、シリコンゴム中に炭素粒子や金属粒子を分散させて、圧力をかけると抵抗値が変化するようにした材料であり、電卓のスイッチとして広く用いられている。特に、炭素粒子を埋め込んだものは、圧力に応じて抵抗値が連続的に変化するため、圧力の大きさを連続的に取り出すことができる。初期に提案された面状触覚センサの多くは、これを絶縁体と交互にストライプ状に並べて実現されている [26, 29]。また、シリコンゴム面上への荷重の重心位置を少ない配線で計測するための信号処理手法が提案されている [35, 38]。
4. **圧電フィルム** ポリフッ化ビニリデン (PVDF) など圧電効果を生じるフィルムを用いることによって、圧力を電圧信号として取り出すものである [31]。また、これをシリコンゴム弾性体に埋め込むことによって、シリコンゴム内部の局所的な応力を計測するものもある [74, 102]。
5. **光反射型** スポンジ、ゴムなどをガラス板の上におき、下側から光を当ててガラス面の反射率の変化を測定することで、対象との接触状態を知るものである [75]。
6. **分布投影型** 圧力分布を透明膜中の光の減衰率に変換し、無圧時からの光量変化を計測することによって、接触部分の圧力分布を知るものである [52]。
7. **ゴム弾性体** シリコンゴム中に微小な空洞を設け、その空洞の変形を音響信号として取り出すことにより、シリコンゴムの歪を計測するものである。シリコンゴムの弾性体としての性質を利用して対象の運動方向や対象表面の空間周波数を測定することができる [106-109]。
8. **圧電振動子** 圧電振動子に固定した検出素子の出力を振動子に帰還させて発振させ、振動子が対象に接触したときの発振周波数の変化を知ることで、対象物の硬さを測定しよう

とするものである [83]。硬さと周波数変化の大きさが比例関係にあることが実験的に確かめられており、人間の手と同様に対象の硬さを知覚することができる。

このほかにも種々の触覚センサが開発されているが [56]、上の説明からわかるように、ほとんどの触覚センサは、対象との接点の分布や接触圧力の分布を知覚することを目的に作られている。このような形態のセンサが多く見られるのは、触覚感覚を「皮膚表面への力の分布」として捉えているためである。このことは、皮膚表面の刺激分布を「画像」として捉えようとする試みが数多くなされていることから裏付けられる [39]。

ここに挙げたセンサは、いずれも対象との接触状態を知るデバイスとしてそれぞれの特徴を有しており、ロボットの手先において対象の把持状態を知る際などに、その有効性を発揮すると思われる。しかし、その計測原理はあくまで「触覚」の領域にとどまったものであり、「触知覚」の領域にまで踏み込んだものではないといわざるを得ない。次節では、前節で行なった考察をもとに、人間の触知覚過程を忠実に模擬したシステムを構成する。

### 3.3 触知覚認識システムの構成

#### 3.3.1 システムの構成

3.1.3 節の検討より、人間の触知覚過程を模擬するためには、次のような条件を満たしている必要がある。

1. 皮膚の単純な接触や圧迫を測定するセンサを有する。
2. 皮膚表面の振動を測定するセンサを有する。
3. 皮膚表面のゆがみや旋断応力の大きさを捉えるセンサを有する。
4. 対象表面をセンサ先端でこする機構をもつ。
5. 対象をさまざまな大きさの力で押し、そのときの変形の大きさを捉える機構をもつ。
6. 対象に接触したときの皮膚表面の温度変化を捉えるセンサを有する。

筆者らは、以上の仕様を満たすような触知覚認識システムを構成した [61,62,90,95,111,112]。図 3.3、図 3.4 にそれぞれ試作したシステムの構成と外観を示す。

本システムは、ステージ部とセンサヘッド部の二つから構成される (図 3.5)。ステージ部は対象を載せて上下に移動する台であり、対象はネジによってステージに固定される。ステージ部の中央には、ロードセル (荷重センサ) が取り付けられており、これによって対象表面にかかる力の大きさを監視する。測定の際には、ステージに対象を載せたのち、ロードセルで捉えた力の大きさが定められた値になるように台を引き上げる。

センサヘッド部は、左右に移動して対象表面をこすりながら、対象表面の状態を計測する部分である。センサヘッド部はさらに、左右に移動する移動台 (マウント) とセンサを搭載したヘッドからなる (図 3.6、3.7)。

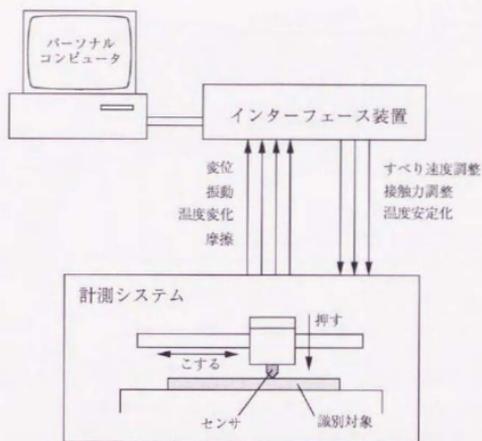


図 3.3 試作したシステムの構成

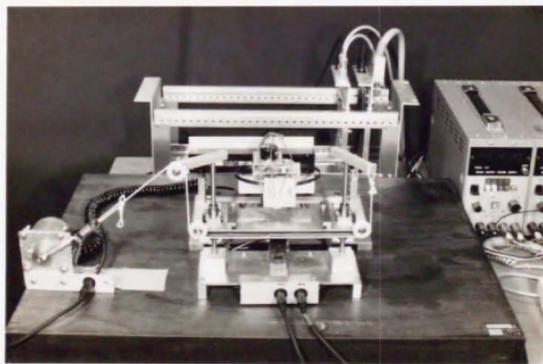


図 3.4 試作したシステムの外観

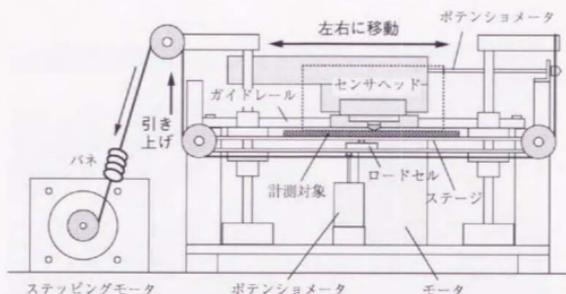


図 3.5 試作したシステムの構造

振動センサは、小型マイクロフォンの先端部をシリコンゴムで覆ったものであり、対象表面を擦る際のシリコンゴムの振動をマイクロフォンで捉える。また、シリコンゴムの中にはサーミスタが埋め込んであり、これによって対象と接触した瞬間のシリコンゴムの温度変化を捉える。

振動センサを固定しているアルミ製の筐体には小型セラミックヒータが取り付けられており、その発熱量を調節することによって、シリコンゴムの温度がほぼ一定に保たれるようにしている。アルミ製筐体を介して加熱するのは、熱容量を大きくすることによって、接触した際にシリコンゴム先端部以外の温度変化が小さくなるようにするためである。これは、人間の体温に対応するものである。

センサヘッド部は、移動台に対して平行リンクで結合されており、移動台を中心に回転するように作られている。また、センサヘッド部は板バネによって移動台に結合されており、常にその位置が中心位置に復元するよう力が働くようになっている。このような機構により、センサ先端が対象表面に接触してこすられるときに、センサ先端と対象表面との間のマクロな摩擦力の大きさをヘッドの回転角として捉えることができる。すなわち、このセンサは摩擦による皮膚の変形を捉える SAIH ユニットに対応する。

センサヘッド部を左右に移動させる部分には、回転モータの軸にゴム車輪を巻き、これを介して移動台を運動させる。移動台が運動する際の振動を抑えるため、移動台はベアリングを用いた直動台に固定してある。モータはパルス幅変調方式によって駆動し、パルス幅を変化させることによって、運動速度を変化させる。

以上のメカニズムにより、上述の条件をすべて満たすシステムを構成することができた。次節では、本システムを用いた具体的な測定方法について述べる。

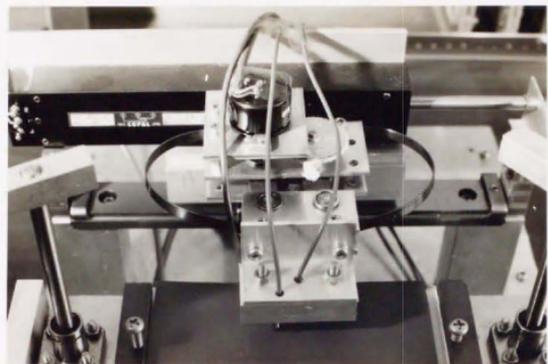


図 3.6 センサヘッドの外観

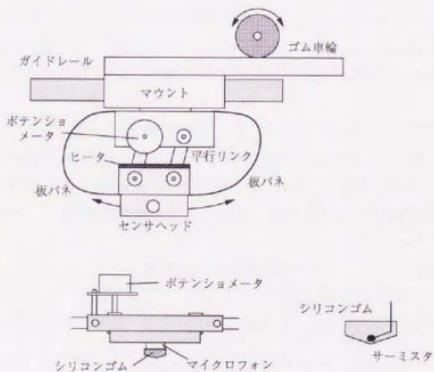


図 3.7 センサヘッドの構造

## 3.3.2 計測と特徴量抽出の過程

実験にあたっては、表 3.2 に示したなぞり速度 3 種類、接触力 3 種類を設定して計 9 回の測定を行なってデータを採取した後、計算機で特徴量を計算した。観測方法の選択は、計算機上でシミュレートする形で行なった。

測定の具体的な手順は以下の通りである。

まず、測定対象をステージの上に固定した後、規定の接触力に達するまでステージを引き上げる。この間に、センサの先端部が対象に接触するが、接触した瞬間はマイクロフォンの信号として取り出すことができる。そして、接触した瞬間から規定の接触力に達するまでにセンサがステージが移動した量から、規定の力をかけたときの対象（およびシリコンゴム）の変形量を知ることができる。また、接触してからの温度センサ出力の変化速度を測定することによって、熱流量に対応する値を知ることができる。

続いて、センサヘッドを左右に動かす。このとき、シリコンゴムが対象表面とこすれあう際の振動をマイクロフォンによって観測するとともに、センサ先端と対象表面のマクロな摩擦の大きさをヘッド回転角として捉える。

マイクロフォンで捉えた信号は、サンプリング周期 500Hz で計算機に取り込み、窓時間幅を 128ms として Hanning フィルタをかけた後、窓時間間隔 64ms で 64 点 FFT によって周波数解析を行なった。周波数解析の結果を低周波数成分 (0-86Hz)、中周波数成分 (94-164Hz)、高周波数成分 (172-250Hz) の 3 つの成分に分解し、このうち中周波数成分、高周波数成分のパワーを計算した。低周波数成分は、材質間の変化が少なかったため、認識の処理には用いなかった。

測定した信号をもとに計算した種々の特徴量の中から、実験で取り扱った 46 種類の材質について比較的大きな差が現われるものを 16 種類抽出した。抽出した特徴量とその特徴量を得るための測定条件を示す (表 3.3)。表において、測定条件欄に示した記号 (S, M, F および 1, 2, 3) は表 3.2 に示した条件を表している。実験では、これらの特徴量を情報源として選択する形で識別動作を進めた。

識別実験を行なう前に、あらかじめ各材質に対して 10 回ずつ測定し、そのデータを用いて各特徴量の平均値と分散を推定した。また、観測値は正規分布に従うものとして、相互情報量と確率分布間距離 (対称化ダイバージェンス) を計算した。そして、2.2 節で述べたアルゴリズムを用いて特徴量を選択し、識別動作を進めた。以下では、選択基準としては相互情報量を用い、観

表 3.2 測定の条件

なぞり速度		接触力	
S	33mm/s	1	0.23N (23g 重)
M	43mm/s	2	0.32N (33g 重)
F	60mm/s	3	0.46N (46g 重)
(誤差)	1.5mm/s	(誤差)	0.04N (4g 重)

表 3.3 識別に用いた特徴量

特徴番号	特徴量の内容	測定条件	
		なぞり速度	接触力
1	マイクログフォン出力の高周波成分	S	2
2	マイクログフォン出力の高周波成分	S	3
3	マイクログフォン出力の中周波成分	S	1
4	マイクログフォン出力の中周波成分	S	2
5	マイクログフォン出力の中周波成分	M	2
6	マイクログフォン出力の中周波成分	F	3
7	回転角の振動数	F	1
8	回転角の振動数	F	2
9	回転角の振幅	S	2
10	回転角の振幅	M	1
11	回転角の振幅	M	3
12	回転角の振幅	F	3
13	温度変化速度	-	1
14	温度変化速度	-	2
15	温度変化速度	-	3
16	対象の変形量	-	2

測後の内部状態空間のエントロピーが0.1以下になった時点で最大確率を示す内部状態をもって識別結果としたときの実験結果を示す。

### 3.3.3 材質の識別実験

識別すべき材料として、表 3.4 に示した 20 種類の材質を設定した。実験では、それぞれは 10cm 四方の試料を用意して上記の方法で計測を行なった。なお、皮革や布地については、試料をプラスチック板に貼りつけて計測した。

図 3.8 に認識過程の例を示す。これは、試験試料として AL と WD1 を提示したときの結果である。図からわかるように、初期状態では特徴量 1 (高周波成分) が最も情報をもたらすものとして選択された。AL を対象とした場合、この特徴量を選択して内部状態分布を更新すると、分布は図に示したように変化した。続いて特徴量 15 (温度変化速度) を選択して分布を更新したところ、エントロピーは十分に減少し、識別結果として AL が得られた。対象としてが WD1 を与えたときも、同様にして 3 回の観測を行なった結果、正しい識別結果が得られた。

図 3.9 は、対象の集合を AL, AC, SF, SP, RB1, CR, WD1, LT1 の 8 種類に限定したときの結果である。この場合は、先の場合と異なり、初期状態では特徴量 15 (温度変化速度) が選択さ

表 3.4 実験に用いた試料

試料番号	略称	材質	備考(板厚)
1	AL	アルミニウム	アルミ板(3mm)
2	TL	陶器	壁用タイル(5mm)
3	CR	コルク	コルク板(5mm)
4	PL	塩化ビニル樹脂	プレート(1.2mm)
5	AC	アクリル樹脂	アクリル板(3mm)
6	SF	発泡スチロール	発泡スチロール板(3mm)
7	SP	スポンジ	グレースポンジ(5mm)
8	WD1	木1	さくら材(5mm)
9	WD2	木2	けやき材(5mm)
10	RB1	ゴム1	ブレンゴムのシート(1mm)
11	RB2	ゴム2	NBRゴムのシート(3mm)
12	PP1	紙1	上質紙
13	PP2	紙2	コート紙
14	LT1	なめし皮	牛皮革
15	LT2	クロム皮	
16	LT3	スエード皮1	
17	LT4	スエード皮2	
18	CL1	布地	デニム
19	CL2	布地	
20	CL3	布地	カーペット

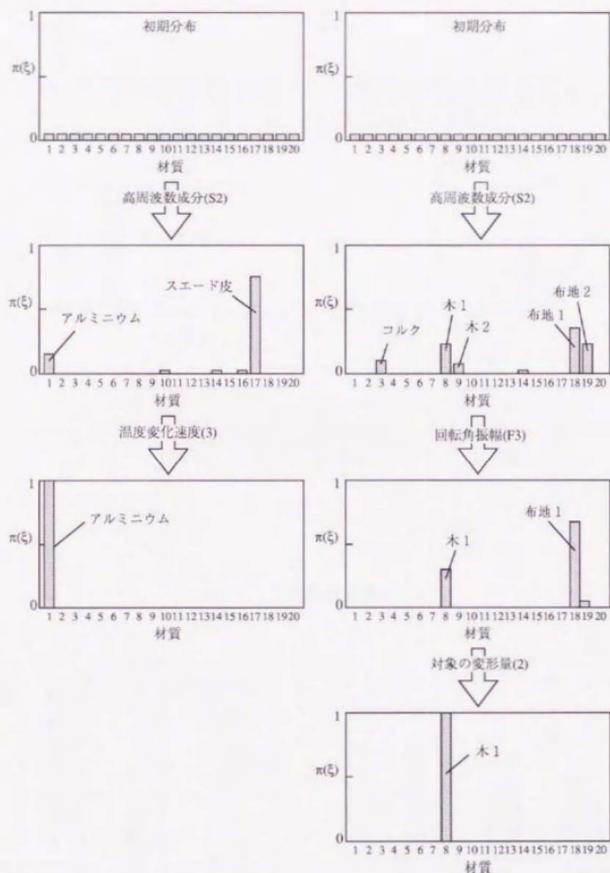


図 3.8 特徴量選択の過程 1

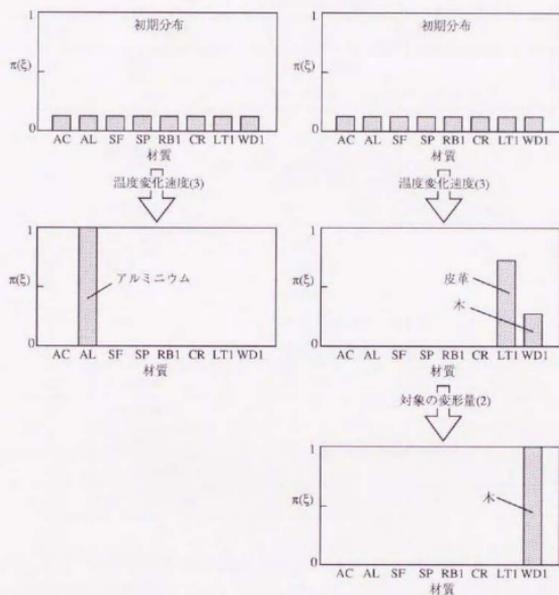


図 3.9 特徴量選択の過程 2

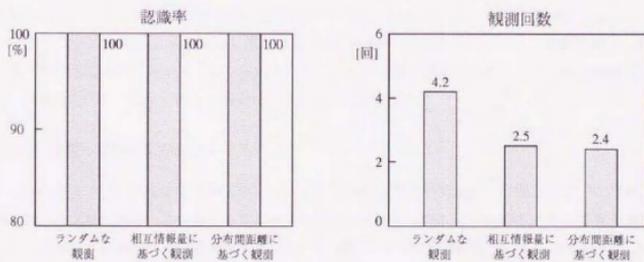


図 3.10 認識率と観測回数

表 3.5 実験に用いた紙質

試料番号	材質	備考
1	上質紙	コピー用紙
2	コート紙	パンフレットに用いられるもの
3	画用紙	
4	新聞紙	
5	ケント紙	製図に用いられるもの
6	剥離紙	紙シールの合に用いられる剥離紙
7	ティッシュペーパー	JKワイパー
8	封筒 1	事務封筒
9	封筒 2	事務封筒 (厚手)
10	封筒 3	コートされているもの
11	レザック紙	表紙などに用いられるデザイン紙
12	タント紙	◇
13	ベルネ紙	◇
14	ランヤ紙	◇
15	カイゼル紙	◇
16	マーメイド紙	◇

れた。このように、識別すべき対象の母集団が異なるとそれに応じて選択される特徴量が変化する点に、能動的認識の特徴が現れている。また、図 3.8 と比較すると、識別問題の内容が異なると、同じ材質を対象としても観測回数が増えることがわかる。

図 3.10 に、センサをランダムに選択した場合、センサを相互情報量基準に基づいて選択した場合、および確率分布間距離基準に基づいて選択した場合について、認識率と識別までに要した特徴観測回数を示す。この問題では、特徴量の中に識別に十分なだけの情報が含まれていたため、すべてのアルゴリズムに対して識別は完全に (正答率 100%) 行なわれた。観測回数については、能動的認識のアルゴリズムを用いることによって無駄な観測が減り、ランダムに観測を行なう場合に比べて平均して約 1.7 回少なくなることがわかった。

### 3.3.4 紙質、布質の識別実験

前節の実験では、温度感覚や弾力感なども含めて、対象間で特徴量が十分異なっていたため、認識動作は比較的スムーズに進行した。そこで次に、純粋に手ざわりの違いしかない材料を対象として選んで、前節の同様の実験を行なった。具体的には、対象の材質として、表 3.5 に示した 16 種の紙質、および 16 種の布質を用意して実験を行なった。なお、これらの試料の中には互いに微妙な差しかないものも含まれており、数人の被験者に対して実験した限りにおいて、人間が

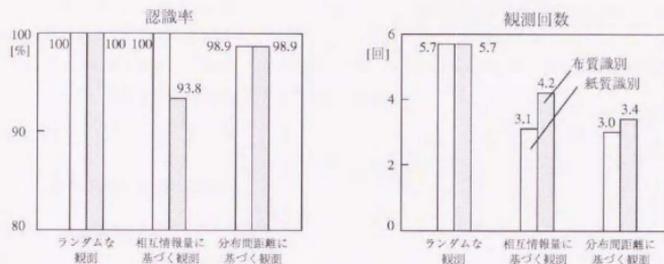


図 3.11 認識率と観測回数

容易に識別できる（少し触っただけですぐにどれであるかがわかる）ようなものではないことを述べておく。

図 3.11 に識別率と観測回数を示す。このように、上記の実験と比べて観測回数は増加するものの、識別率はほぼ 100% に保たれていることがわかる。また、相互情報量基準と確率分布間距離基準とを比べると、後者の方が優れた性能を示すことがわかった。

### 3.3.5 考察と評価

以上の実験結果から、種々のセンサ情報を統合することにより、普通の人間が識別できる程度のさまざまな材質を識別できることが示された。また、識別過程において統合する特徴量を能動的に選択することによって、識別過程が効率的になることが実験的に示された。

ここで考察すべき問題として、「対象のカテゴリ分類」が挙げられる。人間が対象を識別（あるいは分類）する場合、分類の数が増えたときには、対象をその細かな分類に直接分けることはなく、まず大まかな分類をしたのちに細かな分類をするというように、段階を踏んで認識動作を進めることが多い。本システムも、このような人間の戦略と同様の識別戦略をとるように拡張することができる。

そのためには、識別対象をいくつかのカテゴリに分類し、生成されたカテゴリの集合を対象の集合と考えた上で上述のアルゴリズムを適用する。そして何回かの観測を行なってカテゴリの数が絞られた時点で、そのカテゴリに属する対象の集合をあらためて対象の集合と考え、再び同じアルゴリズムを適用する。同じことを繰り返すことによって、本アルゴリズムを階層的なカテゴリ構造をもった識別問題に対して適用することができる。なお、対象をカテゴリに分類するには、天下りのようにカテゴリを指定する方法や、条件付き確率  $P(\xi|y(s))$  の距離に基づいて各対象をグループ分けする方法が考える。

このような段階的な識別を行なうことによって、

1. 内部状態空間の要素数を減らすことができるので、特徴量選択基準に要する計算量<sup>1</sup>を大幅に減らすことができる。
2. カテゴリ間の識別とカテゴリ内での識別とでは、識別に有効な特徴量が異なる可能性が高いので、能動的認識の利点を生かすことができる。

といった利点が生じると思われる。

### 3.4 触知覚感の定量的表現

#### 3.4.1 感覚量の客観化

さて、色、味、匂といった人間の感覚に関わる性質を調べる場合、専門家による官能検査が行なわれることが多い。しかし、いかなる専門家が検査を行なったとしても、このような方法は、検査者個人の性質や心理状態に左右される点で理想的な方法とはいえない。したがって、感覚量を人間の感覚を経ることなく客観的に計測し表現することは、品質の規格化などの点で有用性が大きい。

このような感覚量の定量化の優れた例として色度図 (chromaticity diagram) が挙げられる。すなわち、色度図上で位置を定めることによって対象の色を一意に定めることができる。近年、化学センサを用いて味覚や嗅覚の定量的表現を構成しようとする研究が行なわれており、興味深い結果が出されている [82, 120, 121]。

触覚については、物理的センシングだけで解決されるということから、すでに十分センサの開発が進んでいるかのように考えられている。しかし、3.2節でも述べたように、これまで作られてきたセンサは、接触や表面圧力分布、すべり速度といった物理現象もしくは物理量の計測に主眼がおかれており、質感や手ざわりといった感覚量の計測を意識して作られたものではない。人間の感覚を意識した研究としては、かつて、服飾生地の客観的な評価をめざした研究 [50] が行なわれたが、これも人間の触知覚過程を模擬したものではなかった。

以下では、前節で述べた計測システムを応用することによって、触知覚感を客観的に表現する実験を行なった結果について述べる [111, 112]。

#### 3.4.2 特徴量の組み合わせによる触知覚感の定量的表現

3.3節の識別実験で用いた特徴量の中から選択される頻度の高い8個の特徴を選んだ(表3.6)。これらをもとに主成分分析を行なった結果を図3.12に示す。この図は、主成分分析を行なった後、その第1軸と第2軸を取り出して図示したものである。この図は8次元空間に並んだ点を2次元平面に投影して表示したものである。これだけを見て有用な情報が得られるわけではないが、このように2次元平面に投影したのもでも各材質がある程度分離されていることから、選択した8つの特徴量がこれらの材質を区別するのに十分な情報を担っていることが期待される。

<sup>1</sup>相互情報量を求めるのに要する計算量は要素数の3乗に比例する。

表 3.6 定量的表現に用いた特徴量

特徴量	特徴量の内容	測定条件	特徴量の担う情報
a	マイクロフォン出力の高周波成分	S 2	対象表面の「さらさら感」
b	マイクロフォン出力の中周波成分	S 2	対象表面の凹凸感
c	マイクロフォン出力の中周波成分	S 1	対象表面の凹凸感
d	マイクロフォン出力の中周波成分	M 2	対象表面の凹凸感
e	回転角の振動数	F 2	指を動かしたときの皮膚の自励振動
f	回転角の振幅	M 3	指を動かしたときの皮膚変形
g	回転角の振幅	F 3	指を動かしたときの皮膚変形
h	温度変化速度	- 3	対象に触れたときの温度変化



図 3.12 主成分分析の結果

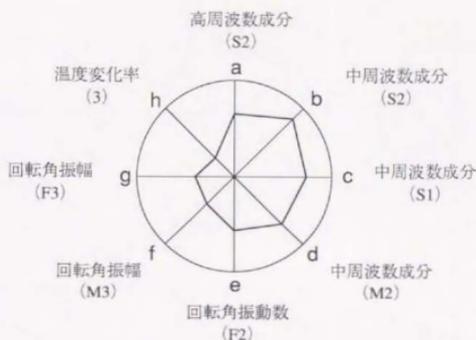


図 3.13 レーダーチャートによる触知覚感の表現

主成分分析によって得られた表現では、各軸にさまざまなバタメータが混じって表されているため、その意味を理解することが難しい。ここでは、各軸に表された情報が人間の観測する信号となるべく対応するようにするため、以上8種の特徴量を直接軸に取ってレーダーチャートとして表す方法を用いることにする。図3.13は、以下で用いるチャートの見方を説明したものである。各軸において、円の半径の midpoint が平均値を表し、円周上および円の中心は、標準偏差を $\sigma$ としたとき $2\sigma$ の値に対応している。

図3.14に、種々の材質をチャートにして表したものを示す。このチャートは人間の触知覚感をよく表現している。例えば、スポンジや発泡スチロールのように触ったときに暖かい感じがするものは、すべてh軸（温度変化速度）上で小さな値を示しており、逆にアルミニウムのように冷たい感じがするものは、h軸上で大きな値を示している。また、スポンジやスエード皮のように、こすったときに指の皮膚が引っ張られる感じのするものは、f, g軸（回転角振幅）の値が大きくなる。発泡スチロールやカーペット地のように、こすったときにカサカサしたり、シャリシャリしたりするものは、a軸（高周波成分）の値が大きくなる。ざらざらした感じのするものは、一般にb, c, d軸（中周波成分）上で大きな値を示す。

このように、レーダーチャートは人間の感じる感覚をほぼ忠実に表していることがわかる。このことは、逆に、製作したシステムが人間が感じているものと同じ情報を捉えていることを意味しており、このシステムが人間の触知覚メカニズムをよく模擬していることを示している<sup>1</sup>。

<sup>1</sup>このシステムで捉えられない特徴量として、対象の「湿り感」が挙げられる。人間は指先の汗を介して対象の湿り具合を感じているようである

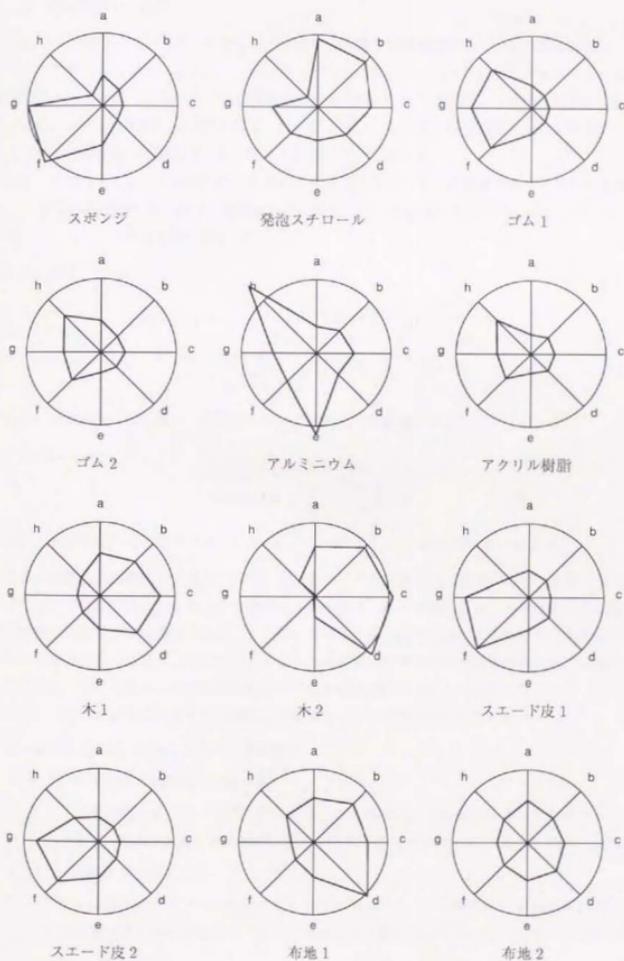


図 3.14 種々の材質のレーダーチャートによる表示

## 3.4.3 触知覚感の類似性

前節では、構成したレーダーチャートの各軸が、人間の触知覚感を構成する特徴と対応していることについて検討した。次に、二つの対象を特徴量の組合せによって表現したとき、それらの表現が似ているか似ていないかが、人間がその対象を触ったときに似ていると感じるか感じないかに対応しているかを調べる実験を行なった。以下では、8次元の特徴空間において計算した距離と人間の主観に基づく類似度とを比較した結果について述べる。

まず、システムにおける距離の計算法について説明する。いま、試験試料の*i*番目の特徴量を  $f_i^*$ 、*k*番目の標準試料の*i*番目の特徴量の平均値を  $f_i^{(k)}$ 、共分散行列を  $D^{(k)}$  と表したとき、以下のようにして二つの距離を計算する。

a)  $\text{dist}_a(k)$ 

$$\begin{aligned} \text{dist}_a(k) &= \sqrt{(f^* - f^{(k)})^T [D^{(k)}]^{-1} (f^* - f^{(k)})} \\ &= \sqrt{\sum_{i,j} (f_i^* - f_i^{(k)}) [D^{(k)}]_{ij}^{-1} (f_j^* - f_j^{(k)})} \end{aligned} \quad (3.1)$$

これは、標準試料に関する共分散行列を用いて正規化した距離である。

b)  $\text{dist}_b(k)$ 

$$\text{dist}_b(k) = \sqrt{\sum_i \frac{(f_i^* - f_i^{(k)})^2}{D_{ii}^{(k)}}} \quad (3.2)$$

これは、各特徴量の間に相関がないとき ( $D_{ij} = 0$  for  $i \neq j$ ) の  $\text{dist}_a(k)$  と一致する。

ここで定義した距離はいずれも対称でなく、また三角不等式を満たさないで、厳密な意味での「距離」ではない。これらは、その定義からわかるように、未知の試料から得られた特徴量が標準試料の特徴量の平均値からどれくらい離れているかを、標準試料ごとの特徴量分布の広がり考えて測った量である。したがって、ここでの距離は、未知試料の各標準試料に対する類似度を与えるもので、任意の二試料間の類似度を与えるものではないことを述べておく。

一方、人間の触知覚感に関する心理物理実験は次のような手続きで行なった。

1. 被験者に目隠しを施し、椅子に座らせる。
2. 机の上に標準試料と試験試料を並べる。
3. まず、標準試料をすべて一度ずつ触らせる。これは標準試料の範囲をあらかじめ知らせるためである。なお、試料に触る際は、左右いずれの手のいずれの指を使ってもよいとした。
4. 次に、試験試料を触らせる。
5. 以降は、試験試料とすべての標準試料を自由に触らせる。この過程で、試験試料に最もよく似た標準試料の一つを選ばせる。なお、判断に迷う場合には、二つ以上の標準試料を回答してもよいとした。

表 3.7 人間の主観に基づく類似度と特徴空間における距離の関係

試験試料		被験者による回答			距離の小さな標準試料						
紙	1	A, H, I	A, H	A	a	A (8.2)	F (12.6)	E (13.2)			
					b	A (4.4)	D (5.2)	H, J (6.7)			
	2	B, A	B, A	B	a	J (16.0)	A (19.4)	E (23.0)			
					b	B (8.2)	A (8.5)	J (12.2)			
	3	H, A, F	H, F	H, A, F	a	F (8.9)	E (12.8)	G (13.1)			
					b	F (5.5)	G (7.8)	D (10.3)			
	4	H, F	H, F, A	H, A	a	A (16.6)	D (20.2)	J (23.5)			
					b	H (5.7)	D (7.4)	J (7.8)			
布	1	B, C	B, C	B, H	a	C (50.5)	A (59.3)	F (128.0)			
					b	C (36.2)	A (49.1)	B (61.3)			
	2	B, A	B, A	B, A	a	A (7.7)	C (14.5)	J (15.3)			
					b	A (4.9)	B (6.2)	C (8.7)			
	3	E, D, H	F, D	I, H, E	a	E (8.2)	D (12.8)	J (15.6)			
					b	E (4.5)	F (7.9)	H (8.7)			
	4	I, D	D, I, J	C, D, I	a	J (5.5)	F (9.3)	D (11.0)			
					b	I (3.5)	J (4.2)	D (6.7)			

なお、実験を繰り返すと、指先の感覚が次第に麻痺してしまい正確な判別ができなくなるので、感覚の麻痺を感じた場合には適宜休息をいれるようにした。

具体的には、紙質と布質を試料として用いて実験を行なった。実験の結果を表 3.7 に示す。この表は、4 種類の未知の試験試料 (1, 2, 3, 4) を被験者に提示し、それが 10 種類の標準試料 (A, B, C, D, E, F, G, H, I, J) のどれと最も近いかを答えさせた結果と、上の二つの方法で計算した距離がそれぞれ小さな標準試料を表したものである。表中で、かっこ内の数字は計算した距離の大きさを表す。

この表から、二つの距離計算法を比較すると、 $dist_b$  による距離に基づく類似性の方が人間の感覚の類似性によく一致していることがわかる。このことは、人間は複数の観測行為によって得た情報を組み合わせて類似性を判断しているのではなく、各観測行為で得られた情報をそれぞれ個々の情報として取り扱っていることを示唆している。

実際、筆者自身が被験者となったときの感覚では、類似性を判断する際には、何か一つの手ざわり特徴に神経を集中して、その特徴に関して対象を比較しようとするのがわかった。また、別の被験者の一人は、実験中に「これは、こちらよりさらさらしている点で似ているが、毛が立っている感じがする点で違う」としており、このことは、人間が手ざわりを判断する上で、いくつかの特徴を軸として設定し、その軸上での距離をもって類似性としていることを示してい

る。人間が類似性を判別する際にどの軸に重点をおいているかについては未検討であるが、距離を計算する際の各軸の重み付けに被験者の感じ方を反映させることによって、より人間の感覚を忠実に反映した類似度の計算法を構成することができるものと思われる。

#### 3.4.4 風合い評価との対照実験

最後に、川端 [50] の製作した服飾生地風の標準試料を、本システムを用いて計測した結果を示す。川端は、生地の風合いを「こし」「ふくらみ」「はり」「しゃり」「ぬめり」の5つの指標を用いて表現するとともに、その測定方法を示している。ここでは、標準試料の中から、表面手ざわり感として捉えることのできる「ふくらみ」(6種)、「しゃり」(4種)、「ぬめり」(6種)のサンプル(紳士服用生地)計16種類を測定し、川端による指標と上の表現との関係調べた。

図3.15に、それぞれの風合いに対するレーダーチャートを示す。この図より、「しゃり」に関しては、明らかに振動成分の大きさとして表れており、指標が大きくなるにつれて、チャート上の数値も大きくなるのがはっきりとわかる。「ふくらみ」や「ぬめり」については単純な傾向は現われていないが、いずれも振動よりは熱移動や摩擦に依存した指標であるものと判断される。なお、これらのチャートにおいて、指標に対して一様な傾向が現れていないのは、各標準試料について注目している指標のほかの指標の値が一定でないためであると考えられる。また、服飾生地の風合い表現という制限した枠組みの中で考えれば、改めて風合いの表現に適切な特徴物を取り出すことができるものと思われる。

### 3.5 第3章のまとめ

本章では、前章で提案したアルゴリズムを触知覚認識過程に適用し、種々の感覚情報を能動的に統合して対象の材質を識別するシステムを構成した。本システムを通じて、能動的認識アルゴリズムの効用について実験的に検証することができた。また、構成したシステムが微妙な触知覚感の違いを識別できることが、実験の結果示された。このような性能は、触知覚における感覚統合と能動的認識に着目し、その過程を忠実に模擬したシステムを構成することによって、はじめて可能になったものと考えられる。

さらに、同システムを用いて触知覚感を定量的に表現する試みを行なった。実験の結果、製作した触知覚感システムは、人間の触知覚感をかなり忠実に実現していることが示された。このことは、人間の触知覚感がこのような多種類の受容器からの情報を統合することによって実現されていること、また、種々の接触力、すべり速度で対象をこすることによって対象の情報を得ていることを実験的に裏付けており、これにより、人間の触知覚メカニズムが本システムと同様のメカニズムによって実現されていることが示された。

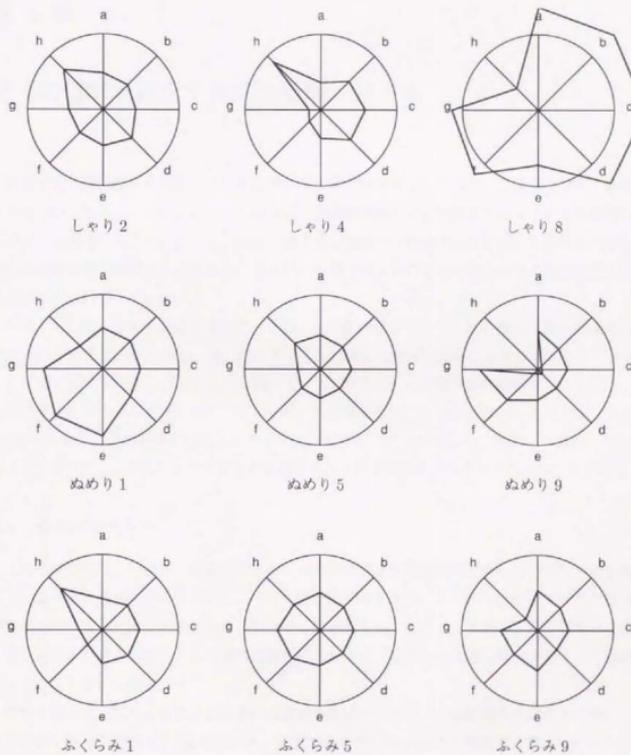


図 3.15 レーダーチャートによる「風合い」の表示

## 第 4 章

### 局所的情報を統合する図形認識システム

本章では、観測位置を能動的に決めながら局所的情報を逐次的に統合して図形を認識するシステムについて論じる。このシステムの要点は、物理的制約という外界が内包している情報構造を知識として利用することによって、図形の部分的な情報だけを観測しながら、システム内部に図形全体の内部像を構成する点にある。さらに、本章では、このような設定の下で観測位置を適切に定める方法について検討する。

以下、4.1節で基本的な考え方を述べた後、4.2節においてシステムの構造と観測位置の決定法について述べる。その中で、以上のメカニズムが競合と協調の原理に基づくネットワークモデルとして実現できることが示される。続いて、4.3節において数値実験の結果について述べ、上記の観測アルゴリズムを用いることによって、内部像が安定して構成されることを示す。また、4.4節ではもう一つの観測位置決定法とその性質について、4.5節ではシステムの改良について述べる。最後に、4.6節において本章の議論をもとに視覚探索メカニズムのモデル化を試みる。

#### 4.1 基本的な考え方

人間の視覚系は、一般に、空間的に広がった画像情報を多数の神経細胞を用いて並列的に処理していると考えられている。しかし、2.1.2節で述べたように、目の前に広がるすべての情報を同時に処理することは不可能であり、またその必要もない。そこで、人間は、自分にとって必要な部分にのみ注意を集中し、その他の情報を捨てた上で、収集した情報を処理するという戦略をとっていると考えられる。

情報を逐次的に取り込みながら対象の内部像を構成する際に、対象に関する知識を利用することが有効であることは2.1.3節で述べた。図形形状に関する内部像を構成する際に利用できる知識の一つとして、外界のもつ物理的な制約が挙げられる。画像情報はデータ量そのものは大きいものの、連続性をはじめとする種々の構造を内包しているため、全体としてデータの冗長性が高い。実際、雑音で乱された画像を復元することができるのは、画像データが冗長であるためにほかならない。ここでは、このように対象が本来有している情報構造を利用することによって、部分的な情報から内部像を組み立てるメカニズムについて考える。

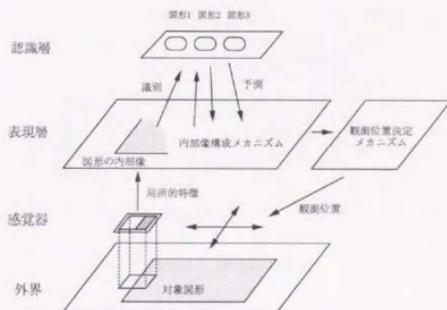


図 4.1 システムの構成

議論を具体化するため、以下では、隣接する局所の特徴（縁、角など）同士のつながりの関係を利用することによって、いくつかの局所の特徴から図形全体の内部像を形成し、さらに、形成した内部像をもとに図形を識別するシステムを考える。次に、このような設定の下で、内部像を効率的に構成するにはどこを観測すればよいかについて考え、観測位置を定めるための評価基準を設ける。

人間が図形の特定の部分に眼を向けるメカニズムには、対象を理解する上で重要な位置を観測するという合目的なものだけでなく、図形そのものに「目を引く」部分があってそこに注意が引きつけられるというものもある。すなわち、画像の中に含まれる局所的な特徴の性質に基いて観測位置が決定される仕組みも存在すると考えられる。そこで、上のようなトップダウンの観測位置決定法に加えて、ボトムアップの観測位置決定法についてもあわせて議論する。

## 4.2 システムの構成と観測位置決定アルゴリズム

### 4.2.1 システムの構成

構成したシステム [93, 96, 98] は、感覚器、表現層、認識層の三つの部分からなる (図 4.1)。

感覚器は、観測位置が与えられると、その位置へ移動し図形の局所の特徴を捉える。そして、観測した特徴とその観測位置を表現層に伝える。したがって、システムは図形全体を一度に見渡すことができない。観測位置の決定法については後述する。

表現層は、図形の内部像を構成、保持する場であり、本システムの中心をなす部分である。表現層は、特徴表現素子の集合体 (以下「特徴カラム」と呼ぶ) が格子状に並んだものである。そして、特徴表現素子の活動パターンをもって図形の内部像を表現する (図 4.2)。

各カラムは外界の特定の位置と一対一対応しており、感覚器から観測した特徴とその観測位置が伝えられると、対応する特徴カラムに属する対応する特徴表現素子に一定の大きさの入力が与

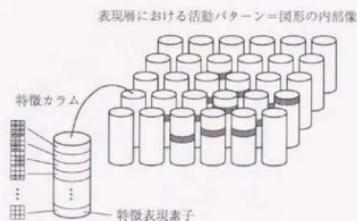


図 4.2 表現層の構造

えられる。また、隣接するカラムに属する特徴素子間には、次節で述べるように相互結合が設けられ、この結合を通じて各素子の活動が周囲の特徴カラムへ伝播するようになっている。また、後述するように、トップダウンの刺激が加えられる場合には、識別層からも入力を受ける。

識別層は、表現層の出力を受けて図形の識別結果を出力する。識別層は各図形に対応する識別素子から構成される。各識別素子はそれぞれ表現層の出力の荷重和を計算し、それが閾値を超えたときに1を出力する。なお、識別素子間には相互作用は設けない。

#### 4.2.2 特徴素子間の相互結合

画像のもつ物理的な連続性を利用して、観測した局所の特徴を基に内部像を構成するため、隣接する特徴カラムに属する特徴ユニットの間に相互結合を設ける。

いま、図 4.3 a) の左側のカラムに示した特徴を観測したとする。このとき、この特徴に隣接する位置で観測されるはずの特徴は、画像の連続性からある程度予測することができる。このように、形状情報が有する物理的な制約を利用することによって、観測した局所の特徴をもとにその周囲についても特徴の分布を定めていくことができる (図 4.3 b))。

このような性質を実現するため、表現層内部に、互いに隣合って存在することの多い特徴を表現する素子間には興奮性の結合を、逆に、互いに隣合って存在することの少ない特徴を表現する素子間には抑制性の結合を設定する。具体的には、隣接する特徴素子間の活動の相関を計算し、その大きさに応じて結合の符号や強度を定めることになる。なお、同一カラム内部の特徴素子間には相互結合を設けない。

#### 4.2.3 観測位置の決定方法

センサーが図形のどの部分を観測するかは、内部像を構成する効率に大きな影響を与える。そこで、能動的認識の考え方を適用して内部像を効率的に進めるための観測方法を考える。

2.1.2 節では、システムの目標を「内部像のあいまいさをなるべく早く減少させる」ことと設定したが、本節では、内部像のあいまいさに代わるものとして「特徴カラムのあいまいさ」を考

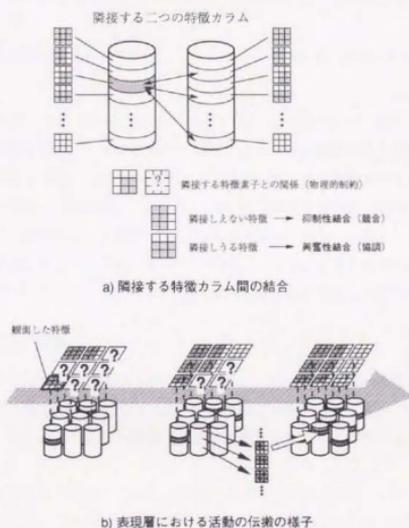


図 4.3 特徴の並びにおける物理的制約と活動の伝播

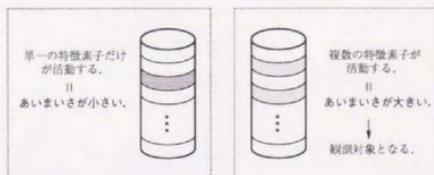


図 4.4 特徴カラムのあいまいさ

える。これは、

1. 特徴カラムに属するどの特徴素子も発火していないとき（その位置の特徴が何であるかわからないとき）
2. 複数の特徴素子が大きな出力を出しているとき（その位置の特徴の候補が複数存在するとき）

に大きな値をとるような量である（図4.4）。そして、このようにして定めたあいまいさを表現層全体に対して総和をとったものなるべく小さくなるように観測位置を定めることを考える。

このような条件の下では、あいまいさがすでに小さくなっている特徴カラムに対応する位置を観測しても新たに得られる情報は少ないので、内部像の構成を効率的に進めるには、あいまいさが大きな特徴カラムに対応する位置を観測するのが効果的である。したがって、ここでは「特徴カラムのあいまいさ」を観測位置決定基準として定め、この評価量が大きな値をとるカラムに対応する位置を観測することにする。基準の具体的な定義については4.3節で述べる。

#### 4.2.4 トップダウン刺激

表現層に組み込まれている内部像構成メカニズムは、特徴間の局所的な接続関係だけを基にしたものである。構造的にいわゆる「極小解」に陥りやすい性質をもっている。このような問題を超えるには、外部から、より大域的な内部像の構造を定めるような制約を与えることが有効である。

大域的な制約を与える一つの方法として、識別層から表現層に対してトップダウンに刺激を与える方法が考えられる。これは、識別素子の活動度がある程度大きくなった時点で、その識別素子から表現層に向けて、その間の荷重に比例した刺激を加えるというものである。これは、「対象はその図形である」という仮説の下で、各位置に出現しやすい特徴素子に対して興奮性の刺激を、逆に出現しにくい特徴素子に抑制性の刺激を加えることに相当する。すなわち、図形が何であるかについてある程度確信がもてた時点で、その図形に関する知識をもとに大域的な予測を行って、内部像の構成を促進することになる。

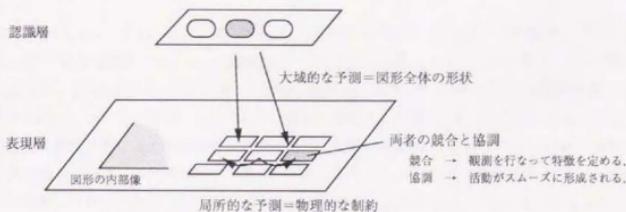


図 4.5 大域的な予測と局所的な予測の相互作用

トップダウンの予測機構を組み込むことによって、次のような現象が起こることが期待される(図4.5)。表現層においては、観測した特徴をもとに表現層内部の結合を通じて局所的な予測が繰り返されているので、この局所的予測とトップダウン予測がうまく整合するときには、内部像の構成は速やかに進行する。一方、これら二つの予測が相矛盾する場合には、矛盾が生じたカラムにおいて、上で定めた「あいまいさ」が増加することになる。このときに、あいまいさが大きくなったカラムに対応する位置を観測することによって、この矛盾を積極的に解消することができる。

この過程は、局所的な予測と大域的な予測という二つの内部操作が互いに矛盾する場合に、その矛盾を解消するために観測を行なうという意味で、まさに合目的であるといえる。すなわち、予測だけでは解決できない部分を優先的に観測することによって、効率的な観測行為を実現することができる。この仕組みは、内部像を形成するメカニズムが多数存在しても全く同じように機能するので、普遍的な観測行為選択アルゴリズムとして用いることができる。

### 4.3 数値実験

#### 4.3.1 実験の設定

構成したシステムの振舞いを調べるため、数値実験を行なった。本節ではまず、実験での設定について述べる。

まず、図形は  $xy$ -座標系で  $(0,1) \times (0,1)$  の領域に提示され、感覚器は 0.05 間隔の  $3 \times 3$  の合計 9 点を観測するものとした。表現層においては、観測したパターンを 46 種類の  $3 \times 3$  の特徴パターンと比較して、最も近い特徴に対応する特徴素子に刺激が加えられるとした。

実験に用いた図形は、三角形、四角形、円、十字、鍵型の 5 種類で、上記の領域をほぼ覆うような大きさの図形をさまざまな位置に提示した(図4.6)。

また、表現層のカラム数は  $21 \times 21 = 441$  とし、各カラムは 0.05 間隔の格子点に対応するものとした。隣接するカラムに属する特徴素子間の結合強度は、次式に従って定めた。

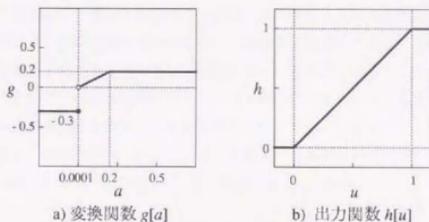
$$w(i_x, i_y, f; j_x, j_y, f') = g \left[ \frac{\text{Freq}(\sigma(i_x, i_y) = f \wedge \sigma(j_x, j_y) = f')}{\text{Freq}(\sigma(i_x, i_y) = f)} \right] \quad (4.1)$$

ここで、 $w(i_x, i_y, f; j_x, j_y, f')$  は、位置  $(j_x, j_y)$  にある特徴カラムに属する特徴素子  $f'$  から位置  $(i_x, i_y)$  にある特徴カラムに属する特徴素子  $f$  への結合係数、 $\sigma(i_x, i_y)$  は位置  $(i_x, i_y)$  で観測された特徴、 $\text{Freq}(A)$  は  $A$  が真であった回数、関数  $g[a]$  は図 4.7 a) で与えられる変換関数である。この関数は、ほとんど隣接したことのない特徴に対応する特徴素子間では強力な抑制をかけ、その他の特徴素子間ではその隣接する確率に応じた強さで促進するという形をしている。隣接カラム間の相互作用は、4 近傍に対してのみ作用するものとした。また、荷重の値は  $(i_x, i_y)$  の絶対的な位置には依存せず、 $(i_x, i_y)$  と  $(j_x, j_y)$  の相対的な位置関係によって定まるものとした。

各特徴素子は、次式に従って自身の内部状態  $u(i_x, i_y, f; t)$  を更新する。



図 4.6 識別実験に用いた図形

図 4.7 相互結合の変換関数  $g[a]$  と出力関数  $h[u]$ 

$$\begin{aligned}
 u(i_x, i_y, f; t) = & (1 - \alpha) u(i_x, i_y, f; t - 1) \\
 & + r(i_x, i_y, f; t) \\
 & + \sum_{(j_x, j_y) \in U(i_x, i_y)} w(i_x, i_y, f; j_x, j_y, f') h[u(i_x, i_y, f; t - 1)] \quad (4.2)
 \end{aligned}$$

ここで、 $r(i_x, i_y, f; t)$  はセンサーから与えられる刺激の大きさで、特徴の観測位置に対応するカラムに属する特徴素子については、

$$r(i_x, i_y, f; t) = \begin{cases} 1 & \text{時刻 } t \text{ で特徴 } f \text{ が観測されたとき} \\ -1 & \text{時刻 } t \text{ で特徴 } f \text{ 以外の特徴が観測されたとき} \end{cases} \quad (4.3)$$

その他の素子については 0 と定めた。また、 $\alpha (= 0.1)$  は内部状態の減衰定数、 $U(i_x, i_y)$  は相互作用を及ぼす範囲、 $h[u]$  は各特徴素子の出力関数で図 4.7 b) で表されるものを用いた。また、各素子の出力には、 $[0, 0.001]$  の一様乱数がノイズとして加わるものとした。

観測位置を選択するための基準として、各カラムに含まれる特徴素子の出力の和が 1 になるように正規化したのち、確率空間のエントロピーに相当するものを計算したのを用いた。すなわち、各カラムに対して、

$$p(i_x, i_y, f; t) = \frac{h[u(i_x, i_y, f; t)]}{\sum_{f'=1}^F h[u(i_x, i_y, f'; t)]} \quad (4.4)$$

$$I(i_x, i_y; t) = \sum_{f=1}^F -p(i_x, i_y, f; t) \log p(i_x, i_y, f; t) \quad (4.5)$$

を計算してこれをカラムの評価量とする。そして、この評価量が最大値をとったカラムに対応する位置を観測位置として選択する。

なお、上記の基準量に準ずるより簡便な基準として、カラム内の素子出力の和の符号を反転したものの

$$J(i_x, i_y; t) = - \sum_{f=1}^F h[u(i_x, i_y, f; t)] \quad (4.6)$$

も用いた。この評価量は、いずれの特徴素子も観測していないカラム（未観察の部分）や周囲のカラムとの競合によって活動が抑制されたカラム（競合が生じている部分）において大きな値をとるので、上の評価量  $J$  に代わるものとして機能するものと予想される。

上述の内部像の更新は、一回の観測に対して二回の割合で行なうものとした。また、内部像の構成は、各カラム内の最大出力がすべて 0.6 を越えた時点で完了するものとした。ただし、実験では計算量の都合上、観測回数が 150 回に到達した時点で強制的に計算を打ち切った。

時刻  $t$  における識別素子  $k$  の出力  $y(k; t)$  は、次式によって定めた。

$$v(k; t) = \sum_{i_x, i_y, f} s(k; i_x, i_y, f) h[u(i_x, i_y, f; t)] - s_0(k) \quad (4.7)$$

$$y(k; t) = \phi[v(k; t)] \quad (4.8)$$

ここで、 $v(k; t)$  は識別素子の内部状態、 $s(k; i_x, i_y, f)$  は位置  $(i_x, i_y)$  にある特徴素子  $f$  からの結合係数、 $s_0(k)$  は抑制性入力からの結合荷重、 $\phi[v]$  は階段関数

$$\phi[v] = \begin{cases} 0 & v \leq 0 \\ 1 & v > 0 \end{cases} \quad (4.9)$$

である。また、結合係数  $s(k; i_x, i_y, f)$ 、 $s_0(k)$  の学習は、教師信号  $z(k)$  を用いて

$$\Delta s(k; i_x, i_y, f) = (z(k) - y(k; t_0)) h[u(i_x, i_y, f; t_0)] \quad (4.10)$$

$$\Delta s_0(k) = -(z(k) - y(k; t_0)) \quad (4.11)$$

として行なった。なお、 $t_0$  は内部像の構成が完了した時刻を表す。

トップダウン刺激を行なう際は、次式に従って各特徴素子の状態を更新した。

$$\begin{aligned} u(i_x, i_y, f; t) &= (1 - \alpha) u(i_x, i_y, f; t-1) \\ &+ r(i_x, i_y, f; t) \\ &+ \sum_{(j_x, j_y) \in U(i_x, i_y)} w(i_x, i_y, f; j_x, j_y, f') g[u(i_x, i_y, f; t-1)] \\ &+ \beta \sum_{k=1}^K \psi[s(k; i_x, i_y, f)] y(k; t-1) \end{aligned} \quad (4.12)$$

ここで、 $\beta (= 0.01)$  はトップダウンの刺激の大きさを定める定数であり、 $\psi[v]$  は、

$$\psi[v] = \begin{cases} 0 & v \leq 0 \\ v & v > 0 \end{cases} \quad (4.13)$$

で定められる関数である<sup>1</sup>。また、トップダウン刺激は、識別素子の内部状態  $v(k; t)$  が閾値 80 を越えたときに加えるものとした。なお、複数の識別素子の内部状態が閾値を越えたときは、それぞれの識別素子から刺激が与えられるものとした。

#### 4.3.2 結果

まず、図 4.8 に四角形を提示した場合の内部像構成過程を一例を示す。これらの図は、各カラムにおける出力最大の素子に対応する特徴を表したものである。ただし、出力の最大値が 0.6 以下であるカラムは表示していない。図 a) は観測位置をランダムに定めた場合、図 b), c) はそれぞれ基準  $I, J$  に基づいて定めた場合の結果を表している。図中で左下に  $\circ$  を付けた位置は、システムが実際に観測した位置を表している。これらの図より、観測位置評価基準に基づいて観測位置を定めることによって、内部像の構成がスムーズに進行することがわかる。

図 4.9 に図形の識別率と内部像が構成されるまでの観測回数をまとめた。この結果は、5 種類の図形をランダムに合計 300 個提示して学習を行なった後、200 個の図形を 3 種類の観測方法で観測したときの結果を示したものである。観測回数のグラフ中の数字は、150 回観測しても内部像が収束しなかった図形数を示しており、棒グラフはそれらを除いた図形に対する平均観測回数を表している<sup>2</sup>。これらの結果からわかるように、観測位置をランダムに定めた場合は内部像が安定して構成されないのに対し、観測位置を能動的に定めた場合は内部像が速やかに構成され、また図形を正しく識別することができる。また、観測位置選択基準として  $I, J$  のいずれを用いても、性能はほとんど変化しないことがわかった。

また、図 4.9 からわかるように、トップダウン刺激を与えると、内部像が収束する図形数が増加するとともに、内部像構成までに必要な観測回数が減少する。このことは、トップダウンの刺激を加えることによって内部像の形成がより安定的に進むことを示している。図 4.10 に、トップダウンに刺激を与えたときの内部像構成過程の一例を示す。

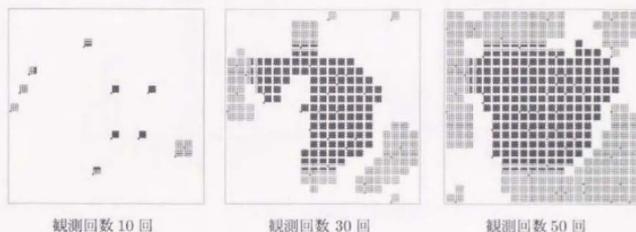
以上の結果より、

1. 観測位置を能動的に定めることによって、内部像の構成が促進され、図形の識別率が向上すること。
2. 認識層から表現層に対してトップダウンに内部像を予測するメカニズムを組み込むことによって、内部像の構成がさらに促進されること。

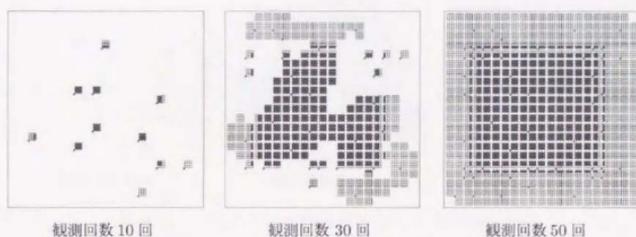
が示された。

<sup>1</sup>これは、表現層へ加えられるトップダウン刺激が興奮性刺激に限られることを意味する。このような設定を行なったのは、抑制性のトップダウン刺激によって表現層内部の活動伝播が妨げられるのを避けるためである。

<sup>2</sup>内部像が収束しにくいのは、図形が三角形と円の場合、すなわち図形に斜めのエッジが存在する場合である。これは、局所的特徴として  $3 \times 3$  パターンを用いたための限界であると考えられる。



a) ランダムに観測位置を定めた場合



b) 選択基準 I を用いて観測位置を定めた場合



c) 選択基準 J を用いて観測位置を定めた場合

図 4.8 内部像の構成過程 (四角形を提示したとき)

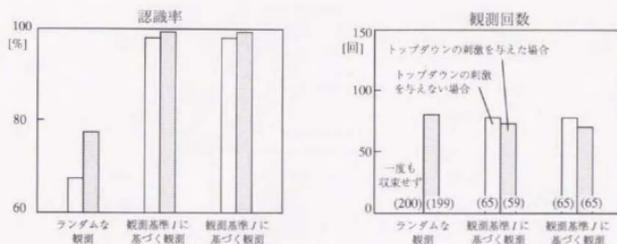


図 4.9 認識率と観測回数

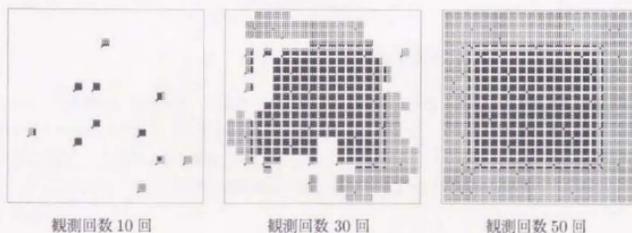


図 4.10 トップダウン刺激を加えたときの内部像構成過程

#### 4.4 ボトムアップな観測位置の決定法

人間は、眼球運動を伴わなくとも、同一の視野の中で注意を移動させることが知られており、mental scanning などと呼ばれている [34]。この場合、網膜が受け取った画像の特定の領域の情報だけを取り出して、情報を処理しているものと考えられる。

このような注意移動の過程を取り扱うために、前システムの設定を変更して、視野全体にわたる画像情報がいったんバッファに取り込まれ、特徴解析がなされた後、その中から選択された特徴が逐次表現層に送られるという構造を考える (図 4.11)。

このような設定の下で、前節で述べた「特徴カラムのあいまいさ」に基づく観測位置決定基準に加えて、画像の特徴分布を観測位置決定に反映させることについて考察する。

画像情報を基に観測位置を定めるためには、画像に含まれる局所的な特徴に対して何らかの評価基準を与える必要がある。画像中の特徴に基づいて視点を決定するメカニズムについては、道と飯島 [140, 141] が図形認識の理論という視点から議論しているが、ここでは、内部像を速やかに形成するという視点から統一的に考察する。

内部像構成を促進するためには、それを観測することによってその周囲の特徴分布がほぼ決定

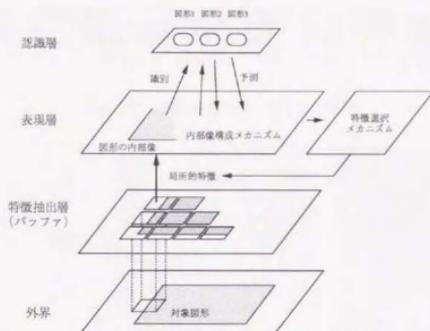


図 4.11 特徴抽出層を介した観測位置の選択

されるような特徴、すなわち、隣接する特徴が特定のものに限定されるような特徴を選択的に観測することが肝要である。このような特徴が確定することによって、その部分を核として内部像の構成は速やかに進行すると思われる。

いま、この条件を満たす特徴を選択するため、特徴量  $f$  に対する評価基準  $L(f)$  を次のようにして定める。

$$L(i_x, i_y, f) = \min_{(j_x, j_y) \in U(i_x, i_y)} \max_{f'} w(i_x, i_y, f; j_x, j_y, f') \quad (4.14)$$

この値は、四つの近傍それぞれに関する隣接する特徴素子への結合加重の最大値の中で最も小さいものの数値である。一つの隣接方向に対する結合加重の和は1であるから、この値が1に近いほど、周囲の特徴分布を決定する力が大きいことになる。図 4.12 に  $L(f)$  の値が大きな特徴を10個選んで示した。この図からわかるように、この基準を満たす特徴は、画像の角を検出する特徴であることがわかる。

以上で定めた観測基準を用いた場合のシステムの振舞いについて調べるため、数値実験を行った。

実験では、ボトムアップの評価基準  $L$  ともとの評価基準  $I, J$  を次のように組み合わせ、全体的な評価基準  $I', J'$  を設定した。

$$I'(i_x, i_y) = I(i_x, i_y) + \gamma L(i_x, i_y, f) \quad (4.15)$$

$$J'(i_x, i_y) = J(i_x, i_y) + \gamma L(i_x, i_y, f) \quad (4.16)$$

このように評価基準を定めることにより、システムは、特徴量評価  $L(i_x, i_y, f)$  が大きな位置の中で、特徴カラムのあいまいさが残っている部分を選択的に観測することになる。なお、実験では係数  $\gamma$  は0.01とした。以下に、この基準に基づいて観測位置を定めたときのシステムの振舞いを示す。

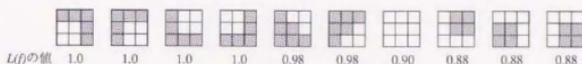
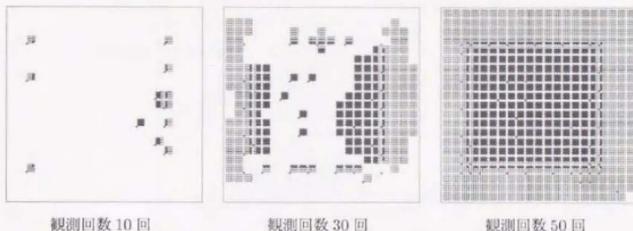
図 4.12 評価基準  $L(f)$  の大きな特徴

図 4.13 ボトムアップの評価基準を反映させて観測位置を定めたときの内部像構成過程

図 4.13 に、四角形を提示したときの内部像形成過程を、図 4.14 に、内部像形成までに必要な観測回数と図形の認識率を示す。これらの結果からわかるように、ボトムアップの観測位置決定基準を加えることによって観測効率が向上することがわかる。

#### 4.5 表現層の階層化

上記システムの性質をさらに高める方法として、内部像表現機構の階層化が考えられる。

上述のシステムにおいては、内部像形成メカニズムの本質だけを抽出して議論するため、図形の表現層を単一層で実現したが、これを多層化することによってさらに効率的に内部像を構成することができる。

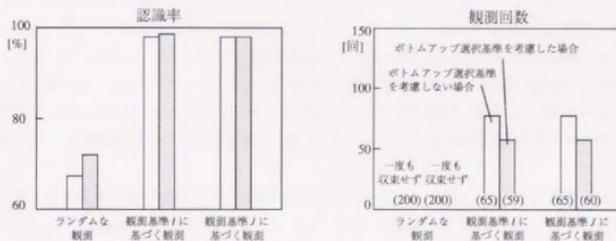


図 4.14 ボトムアップの評価基準を反映させたときの観測回数と認識率

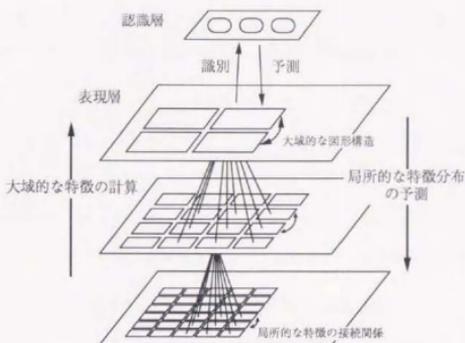


図 4.15 表現層の階層化

多層構造の表現層の概念図を図 4.15 に示す。上位の表現層は、下位の表現層の一定の範囲から信号を受けて、より大域的な図形の構造を表現する。上位の表現層へ行くほど大域的な図形構造に基づく相互作用が働くため、図形全体の構造が速やかに形成されることになる。そして、そのような大域的な表現層の活動を基に下位の層の活動を定めることによって、下位層における像の構成が促進される。

このように、上述のシステムに階層的構造を取り込むことによって、内部像の構成が速やかにかつ安定して進行するようになることが予想される。なお、このような階層的システムは、福島[21]による選択的注意のモデルと類似した振舞いを示すものと思われる。

#### 4.6 視覚探索モデルへの発展

近年、視覚探索 (visual search) と呼ばれる研究が、視覚心理学の領域で注目を集めている [59, 138, 139]。これは、複数の刺激からなる視覚パターンを与えたときにその中の特定の刺激が浮かび上がって見える現象に着目して、視覚情報処理のメカニズムを解明しようとする研究である。特定の刺激が浮かび上がる際には注意がその特徴に引きつけられると考えられるので、視覚探索の現象は、注意移動のメカニズムという点で本章で論じた主題と密接に関係しているといえる。本節では、この研究の中から本章に関係する知見を拾いだし、本論文の視点からそれらの解釈を試みる。

視覚探索の研究において扱われる代表的な現象に、ポップアウト (popout) と探索非対称性 (search asymmetry) の二つがある。ポップアウトとは、多数の妨害刺激の中から特定の目標刺激が浮かび上がって見える現象である。例えば、図 4.16 a) において、円に線分がつけ加えられたものは、円だけのパターンの中から浮かび上がって見えるが、この現象をポップアウトとい

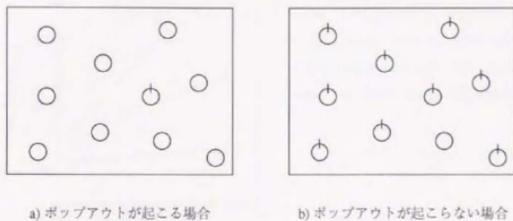


図 4.16 視覚探索に関わる心理物理現象

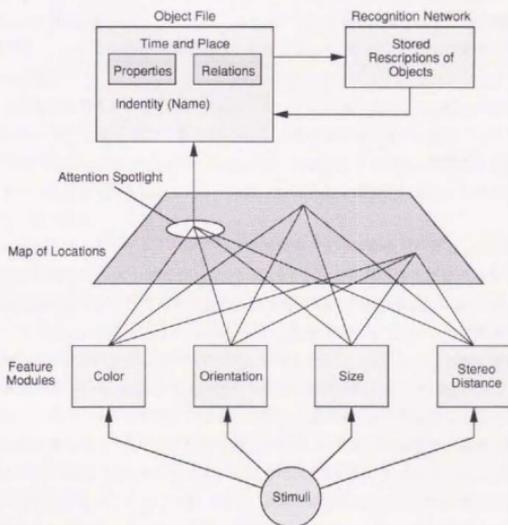


図 4.17 Treisman による特徴統合理論 (文献 [124] より引用、改変)

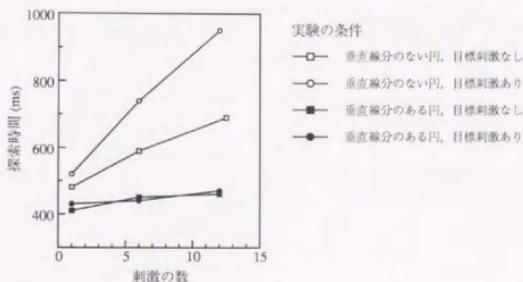


図 4.18 特徴探索と結合探索に対する反応時間 (文献 [51] より引用, 改変)

う。一方, b) において一つしかないのは円だけからなるパターンであるが, a) の場合と異なつて, このパターンは浮かび上がりにくい。このように, 妨害刺激と目標刺激との間に対称性が成立しないことを探索非対称性という。また, ポップアウトの手がかりとなるような特徴は初期特徴と呼ばれており, いかなる特徴が初期特徴となりうるかということについて数多くの実験が行なわれている [59]。

このような現象を説明する心理学的モデルとして, Treisman による特徴統合理論 (feature integration theory) がよく知られている [123-131]。このモデルは, 知覚過程を「前注意過程」と「注意集中過程」の二つに分けて捉える Neisser [81] の考え方に基づき, 特徴解析を並列的に行なう特徴マップ生成過程と, そのようにして生成された特徴を逐次的に統合する特徴統合過程から構成されている (図 4.17)。

Treisman によれば, ポップアウトするのは目標刺激が単一の特徴だけによって妨害刺激から区別できるときに限られ, ポップアウトしないのはそれらが複数の特徴を組み合わせることによってはじめて区別されるときである。前者の探索過程を特徴探索 (feature search), 後者を結合探索 (conjunctive search) という。Treisman は, 特徴探索の場合は, 一つの特徴地図の出力だけから目標刺激を取り出すことができるため, 注意を介在させることなく目標刺激を検出できるが, 結合探索の場合には, 注意を順次走査して複数の特徴地図に属する特徴を統合するため, 探索に時間がかかるという解釈を与えている。実際, 特徴探索と結合探索のそれぞれの場合について, 目標刺激が見つかるまでの反応時間を測定すると, 特徴探索の場合は, 妨害刺激の数が増加しても探索時間がほぼ一定であるのに対し, 結合探索の場合は, 妨害刺激の数に比例して探索時間が増加することがわかっており (図 4.18), 後者の過程において, 人間が個々の刺激を逐次的に検査していることを示している。

特徴統合理論のほかにも Wolfe, Cave による活性化マップ・モデル [15, 135] などが知られているが, ここでは, これらのモデルとは独立に, 前節までの議論をもとに視覚探索過程の原理モデルを構成する。

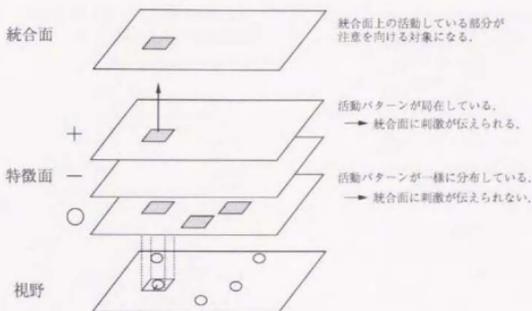


図 4.19 視覚探索メカニズムの原理モデル

Treisman は、注意の走査が必要になるのは結合探索の場合だけであり、特徴探索の場合は不要であると述べているが、特徴探索の場合であっても、目標刺激が見つかるのと同時に注意はその刺激に向くものと考えられる<sup>1</sup>。ここでは、ポップアウトという現象そのものが何であるかということには踏み込まず、「ポップアウトする場合には、目標刺激に対して直接注意が向けられ、そうでないときには、何らかの探索過程を経て目標刺激に注意が向けられる」とみなした上で、そのような注意誘引のメカニズムについて考察する。

図 4.19 は、このメカニズムに対するモデルを示したものである。このモデルは、種々の局所的特徴を表現する特徴面とこれらの情報を統合する統合面から構成される。

各特徴面の活動は画像情報からボトムアップな計算だけで得られるとする。このとき、各特徴面には対応する特徴の視野全体にわたる分布が現れるので、対応する特徴が視野全体に対して一様に分布していれば、特徴面の活動パターンは面全体に広がり、逆に、対応する特徴がただ一つしかなければ、活動パターンは孤立局在興奮となる。ここで、各特徴面に対して活動パターンの分散を考えると、前者の場合は分散が大きくなり、後者の場合は小さくなる。なお、特徴面上に興奮が全くない場合は、分散は無限に大きいとみなす。

統合面は、特徴面の出力に各特徴面の分散に反比例した荷重をかけた荷重和を受け取るとする。そして、統合面上で大きな出力を出す位置が注意を引きつけると考える。このようなメカニズムを考えると、視野全体にわたって分布している特徴は、特徴面の興奮パターンの分散が大きいために観測位置決定にほとんど関与せず、逆に、視野の中で一カ所だけに現れている特徴は、観測位置決定に大きな影響を与える。

このようにして統合面にまとめられた活動パターンが、ただ一つの局在興奮になっているときは、注意の向かう位置はその局在興奮の位置によって一意に定められる。一方、統合面の活動パ

<sup>1</sup>このことについては現在議論が分かれている。

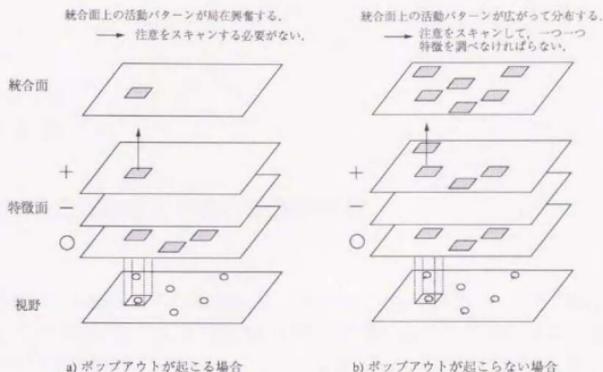


図 4.20 視覚探索モデルの振舞い

ターゲが複数の局在興奮からなっているときには、それら一つ一つを順次調べることによって目標の刺激を探索することになる。

このモデルに基づいて、ポップアウトと探索非対称性が生じる原因について説明を試みると次のようになる(図 4.20)。

ポップアウトが生じる条件の下では、統合面には局在興奮がただ一つだけ現われるので、注意は必然的に目標刺激のある場所に向けられ、その結果ポップアウトが起こることになる。一方、目標刺激と妨害刺激を逆にした場合には、目標刺激を表す特徴面も、妨害刺激を表す特徴面も活動パターンの分散が広がるため、統合面には局在興奮が多数現われることになる。その結果、刺激一つ一つに注意を向けなければならないことになり、反応時間が長くなることになる。このように考えると、探索非対称性の原因は、ある特徴が「ある」ことを表す特徴面はあるが「ない」ことを表す特徴面がないこと、すなわち、脳の情報表現様式に起因していると考えられる。

以上、視覚探索過程の原理モデルについて述べてきた。近年、視覚探索については多数の実験結果が報告されており、ここで述べたような単純なモデルでそのすべてを説明することはできない。しかし、このような心理現象は、注意が特定の特徴に引きつけられるという点で、2.3節で述べたボトムアップの観測方法選択の実例として興味深い。また、上で述べたように、探索非対称性は脳における情報表現メカニズムを明らかにする上で大きな示唆を与えてくれる。このような意味で、視覚探索は、注意や情報表現の問題を議論する上で有効な題材であり、今後、この現象に関する知見を詳細に検討することによって、これらの問題を解決する上でさらに大きな示唆が得られるものと期待される。

## 第5章

### 内部像の信頼度を考慮した運動計画

本章では、内部像に基づく運動計画について議論する。議論の題材として「なぞり運動」を取り上げ、この運動課題を実行する際の人間の振舞いを模倣するシステムを構成することを通じてアルゴリズムを具体化する。

まず5.1節で、本章で取り扱う問題を明らかにし、それを解決するための基本的概念である「内部像の信頼度」について述べる。次に、5.2節でなぞり運動に関する心理物理実験について紹介した後、5.3節においてその運動過程を定式化し、5.4節で具体的なアルゴリズムを記述する。続いて、5.5節で構成したアルゴリズムに関する数値実験の結果を示す。5.6節では、構成したアルゴリズムの発展や関連する問題について考察する。

#### 5.1 基本的な考え方

人間の運動の中で、自己の意志で自発的に起こす運動は、外部刺激に対して反射的に起こす運動と区別して、随意運動 (voluntary movement) と呼ばれる。そして、随意運動過程において身体がどのような軌道を描くか、またそのような軌道がどのようなメカニズムによって生成されるのかが、多くの研究者によって調べられてきた。

例えば、十分に習熟した運動を行なう場合、人間の腕が一定の軌道を描くことが知られているが、このときに腕が描く軌道がどのような規則に従っているかを明らかにしようとする心理物理学的な研究が行なわれている [2, 20, 132]。また、得られた実験データに基づいて、腕の加速度の総和を最小とする「躍度最小モデル」 [20]、関節まわりのトルク変化の総和を最小にする「トルク変化最小モデル」 [132] などの仮説が提案されており、それぞれ人間の運動軌道をよく模倣することが知られている。また、このような軌道生成のための神経回路網モデルも提案されている [55, 65]。

一方、人間が運動に習熟する過程についても、いくつかのモデルが提案されており、視覚座標における目標軌道から必要な運動指令を計算する逆モデルの学習 [6, 60, 69]、運動指令から軌道を定める順モデルと軌道から運動指令を定める逆モデルを同時に獲得する順逆モデルの学習 [43]、また、フィードバック系の出力を教師信号に用いるフィードバック誤差学習 [25, 54] など

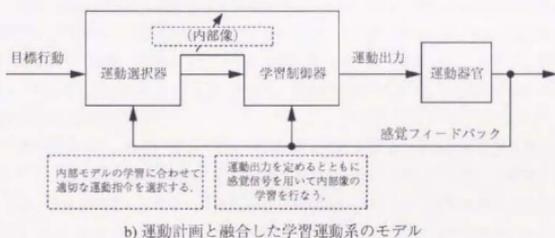
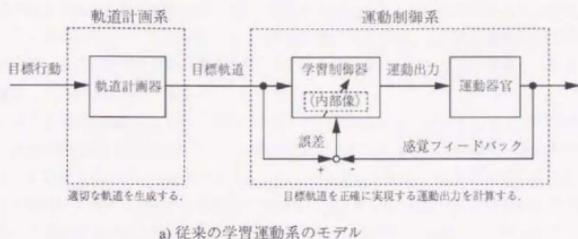


図 5.1 学習運動系のモデル

がよく知られている。これらはいずれも、外部から与えられた目標軌道を忠実に実現するように学習を行なうものである。

このように、これまでの随意運動のモデル化においては「軌道生成過程」と「学習制御過程」が分離されており、前者によって適切な軌道が生成されたのち、後者によってそれを実現するための運動指令が計算されるという図式が用いられてきた(図 5.1a)。そして、学習過程については、軌道生成過程が生成した理想的な軌道を目標に繰り返し学習するという設定で論じられることが多かった。

これらの研究は、随意運動の発現過程においてこれら二つの問題が重要であることを明確にした点、また、それぞれの問題に対する解答を明らかにした点で有意義な研究である。しかし、人間が慣れない運動を行なう場合を考えると、これら二つの過程は完全には分離されていないと考えるのが自然である[53, 100]。すなわち、不慣れな運動を行なうとき、人間は、頭の中で理想的な軌道を計画したのち、それをそのまま実現しようとするようなことはしない。仮にそのようなことをしたとすると、学習が終了するまでのあいだ、人間の身体は目標とする軌道から大きくはずれて動くことになる(実際、多くの学習運動系のモデルは、学習過程で目標軌道から大きくはずれた軌道を描く)。

人間が随意運動を行なうときには、ふつう何らかの目的をもっているはずであり、それを実現するために身体を動かす。したがって、このときの第一の目標は、この目的を確実に達成することであって、身体を効率よく動かすことではない。効率的な運動は、運動に習熟するに従って次第に獲得していけばよいのである。

人間は、不慣れな運動を行なうときは、身体をゆっくり動かすことによって（速度は犠牲にしても）大きな誤りを犯さないようにし、学習や適応が進むにしたがってすばやい運動を行なっているように思われる。上で述べたように、ここで鍵となるのは、1) 目標行動を誤りなく実現することが第一であり、最初から理想的な運動をしようとする点、2) 学習や適応の進み具合に応じて運動指令が変化する点である。つまり、人間の運動計画は適応や学習の進行を反映していると考えられる（図 5.1b）。

さて、学習の進行に応じて適切な運動指令を構成するためには、脳は運動器官に関する内部像（内部モデル）の学習がどの程度進んでいるかを評価する必要がある。しかし、内部像がどの程度正確であるかは、実際に運動を行なって、予測した軌道が実際の軌道からどれだけずれているかを測るまで知ることができない。したがって、運動を行なう前には、内部像の確かさを「頭の中で」評価しなければならない。

この問題を解決するために、本章では、内部像がどの程度信頼できるかを自分自身の内部で評価したもとして「内部像の信頼度」という概念を与える [99, 100]。そして、内部像の信頼度が低いときは、大きな誤差を生まないように小さな運動しか行わず、学習が進んで内部像の信頼度が向上すれば、予測を生かして効率的な運動を行なうという形で、運動計画を定めることとする。

以下では、このように内部像とその信頼度を組にして用いることによって、目標運動をはば間違いなく実現しつつ、学習の進行とともに運動の性能を向上させるアルゴリズムを構成する。具体的な題材として、自分の示指で図形の縁をなぞる「なぞり運動」を取り上げる。なぞり運動においては運動軌道が図形の縁として与えられるので、軌道生成の細かい議論を避けて運動計画過程に議論を集中することができる。

## 5.2 なぞり運動に関する心理物理学的知見

アルゴリズムを具体化する前に、赤松ら [3-5] によるなぞり運動に関する心理物理実験の結果についてまとめておく。これらの実験は、視覚と体性感覚の関係が変形された状態で「なぞり運動」を行なったとき、視覚・体性感覚・運動間の変換がどのように適応していくかを調べるために行なわれた実験である。

実験では、厚さ 4mm、辺の長さが 30mm ~ 150mm、長短比 3/10 ~ 10/3 の長方形の亚克力板を用意し、被験者に利き手の示指で図形をなぞるように指示する。実験は暗室中で行なわれるため、被験者は自分の指や図形を直接見ることができない。その代わりに、被験者は指先に装着した赤外 LED の光を PSD カメラで捉え CRT に表示したもの (Sell Spot system) を見ながら図形をなぞる。このとき、図形の輪郭も CRT 上に表示される（図 5.2）。

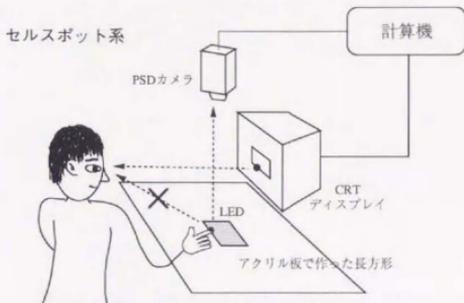


図 5.2 実験の状況

このように CRT を介して指の位置を提示することによって、実験者は、被験者の視覚・体性感覚間関係を任意に変形することができる。実験では、カメラで捉えた情報を縦あるいは横方向に伸縮したものを画面に表示する（それに合わせて図形の輪郭も伸縮させて表示する）。したがって、画面を見ながら図形の縁を間違いなくなぞるためには、被験者は未知の視覚・体性感覚・運動間関係に適応する必要がある。

赤松らは、この実験装置を用いて、画面に指の位置が表示されないとき（視覚フィードバックが欠けたとき）、図形を CRT にだけ表示し触覚的に提示しないとき（触覚フィードバックが欠けたとき）、指を他人に動かしてもらったとき（自分の運動系が使えないとき）など種々の条件の下で、適応の速さや適応によるなぞり運動の変化について調べている。ここでは、指の位置が表示される場合（視覚フィードバックが働くとき）について取扱う。

図 5.3 になぞり運動中の典型的な指の動きを示す。この図からわかるように、なぞり運動中の指の動きは、一度に大きく移動する「跳躍」の部分と、頂点を探しながら細かい加減速が繰り返される「探索」の部分に分けることができる。

さて、なぞり運動を繰り返すうちに、適応が進行し被験者の運動はスムーズになっていく。このときの指の動きの変化を模式的に示したものが図 5.4 である。図に示したように、適応が進むにつれて、跳躍部分での指の速度が増すとともに、跳躍部分の行程が長くなる。以上の結果から、赤松らは「運動の前半はフィードフォワード制御、後半はフィードバック制御が行なわれており、適応が進むにつれてフィードバック制御に切り替わるタイミングが遅くなる」と結論づけている。また、眼球運動を計測すると、適応が進むにつれて辺の終端付近を注視する回数が減少することも報告されている [3]。

次節では、このような人間の振舞いを模擬するシステムを構成する。

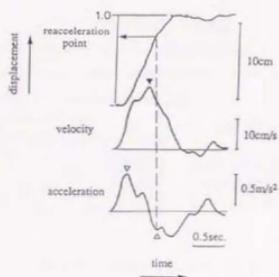
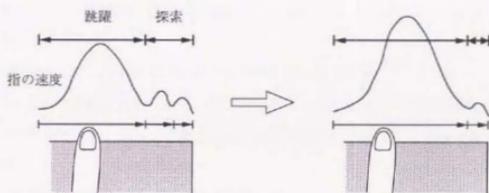


図 5.3 なぞり運動中の指の動き (文献 [3] より引用)



運動に慣れるにつれて

- ・跳躍部分の距離が長くなる。
- ・跳躍部分の速度が大きくなる。

図 5.4 適応による指の動きの変化

### 5.3 運動計画過程のモデル化

図形をなぞる際には、指を図形の線に引っかけてよいために、指の運動方向は物理的に拘束される。そこで、ここでは指の運動方向については考慮せずに、指の運動距離だけを取り扱うことによって問題を単純化する。また、運動の加速減速プロフィールも取り扱わず、各運動指令に対し結果として指が動く距離（以下跳躍距離と呼ぶ）だけを考えることにする。これによって、細かい軌道生成の過程に触れることなく、運動計画の問題に議論を集中することができる。

以上の条件の下で、なぞり運動を誤りなくかつなるべく早く行なうためには、運動の結果指が頂点を通り過ぎず、かつ頂点に近い位置まで動くような運動指令を出力すればよいことになる。そのために必要なことは、頂点までの距離を正確に把握すること、運動指令の結果を的確に予測することである。逆に、頂点がどこにあるのかわからない（視覚情報が遮断されている）場合や、頂点に到達するために必要な運動指令が計算できない（運動の予測が不確実である）場合で

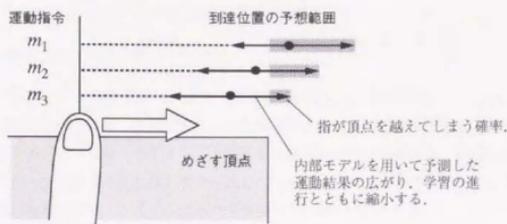


図 5.5 運動計画のアルゴリズム

は、運動をすぐに修正できる程度のゆっくりとした速度で頂点を探索しながら、指を動かさなければならぬ（「てさぐり」の状態！）。

以上の考察から、なぞり運動を効率よく実行するためのアルゴリズムは、次のように書き下すことができる（図 5.5）。

1. 視覚情報をもとに、現在の指の位置と次の頂点の位置を知る。
2. 内部像を用いて運動結果を予測し、指が頂点を行き過ぎないように運動指令を計算する。
3. 的確な予測ができないときは、誤りを生じない程度のゆっくりとした速度で指を動かす指令を出す。
4. 計算した運動指令に基づき、実際に指を動かす。
5. 頂点に達したことがわかれば、次の頂点を新たな目標とする。

このような運動戦略の下で、視覚・運動変換の適応が進む状況を考えて、次のようになる。なぞり運動を始めた時点では、視覚・運動変換の信頼度が低いので、縁から指が離れないようにするには、必然的に指の運動速度を小さくせざるを得ない。さらに、自分の指がどこにあるか、またどこへ動いたかを知るために視覚フィードバックの手助けが不可欠であり、そのために視覚情報を参照する頻度が高くなる。しかし、運動を繰り返すうちに、運動指令と結果として生じた指の運動との関係がわかり、これをもとに内部像の学習を行なうことができる。学習が十分に進行すると、内部像の信頼度が向上し、運動指令によって生じる跳躍距離を正確に予測できるようになる。そのため、次の頂点まで指を動かすのに適切な運動指令を生成できるようになり、結果としてすばやい運動が可能になる。また、それに伴って視覚フィードバックの頻度も減少する。

以上のアルゴリズムを用いることによって、未知の状況においては誤りを犯さない程度の小さな運動を行ない、学習が進むにつれて大きな運動をすることができる。次節では、このアルゴリズムを確率モデルを用いて具体的に記述する。

## 5.4 確率モデルを用いたアルゴリズムの定式化

## 5.4.1 運動計画のアルゴリズム

まず、内部像として順システムモデル（各運動指令に対して跳躍距離を与えるモデル）を考え、これを確率モデルによって記述する。具体的な記述法については後述する。指や頂点の位置情報は視覚のみによって得られるとし、それをもとに運動指令を計算する。運動の結果、指と頂点との距離が一定値以下であれば、頂点に達したことがわかるものとする。また、取り得る運動指令の種類は有限個とし、システムは其中から一つを選んで出力する。

以上の条件の下で、次のようにアルゴリズムを定める。

1. 視覚情報によって現在の指の位置と頂点の位置を知り、指を動かすべき距離を計算する。
2. 自分がとり得るすべての運動指令に対して、それぞれ指の跳躍距離の期待値、および頂点を行き過ぎてしまう確率を内部像を用いて計算する。
3. 頂点を越えてしまう確率が閾値  $\theta$  以下である運動指令のうち、指の到達点の期待値が頂点に最も近いものを選択する。
4. 3. の条件をみたす運動指令がない場合は、最小の運動を生成する運動指令を選択する。
5. 運動指令を出力し、実際に指を動かす。
6. 出力した運動指令と運動前後の指の位置から、内部像の学習を行なう。
7. 頂点に達すれば、次の頂点を目標とする。

## 5.4.2 内部像の表現形式

条件で定めたように、内部像は、指が位置  $\mathbf{x} = (x, y)$  にある状況で運動指令  $m$  を出したときの指の跳躍距離  $d(\mathbf{x}, m)$  という形で与える。  $d$  は確率変数であるので、その確率分布  $p(d|\mathbf{x}, m)$  が内部像の実体となる。具体的な表現方式として、ここでは次の二つを考える。

## a. 離散表現

指の位置  $\mathbf{x}$  を量子化して離散値  $\bar{\mathbf{x}}$  にし、その各点と運動指令  $m$  の組合せに対して、それぞれ一つずつ条件付き分布  $p(d|\bar{\mathbf{x}}, m)$  を割り当てる。分布の形としては、平均  $\mu_{\bar{\mathbf{x}}, m}$ 、分散  $\sigma_{\bar{\mathbf{x}}, m}^2$  の正規分布を考える。

## b. 関数表現

指の移動量  $d$  を、  $J$  個のパラメータ  $\xi_j$  ( $j = 1, 2, \dots, J$ ) によって定まる関数として表す。また、パラメータベクトル  $\xi (= (\xi_1, \xi_2, \dots, \xi_J)^T)$  は、平均  $\mu$ 、共分散行列  $D$  の  $J$  次元正規分布に従うとする。

$$d = f(\mathbf{x}, m; \xi) \quad (5.1)$$

取扱いを簡単にするため、ここではパラメータに関する線形モデルを考える。

$$d = \sum_{j=1}^J \xi_j f_j(\mathbf{x}, m) = \xi \cdot \mathbf{f}(\mathbf{x}, m) \quad (5.2)$$

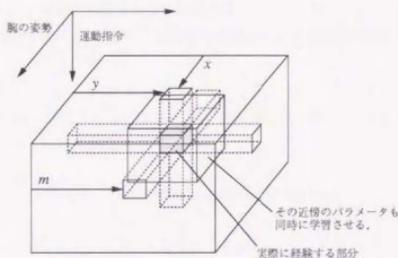


図 5.6 隣接する分布の学習

## 5.4.3 内部像の学習

内部像の学習は Bayes 法に従って行なう。すなわち、学習前、学習後のパラメータ  $\alpha$  の分布をそれぞれ  $p^-(\alpha)$ ,  $p^+(\alpha)$  としたとき、位置  $x$  において運動指令  $m$  を出したとき跳躍距離が  $d$  であった場合には、Bayes の定理に従って、

$$p^+(\alpha) = \frac{P(\alpha|d)}{\int P(\alpha|d)p^-(\alpha)d\alpha} p^-(\alpha) \quad (5.3)$$

として確率分布を更新する。確率分布が正規分布である場合には、更新の操作が分散と平均に關する単純な漸化式の形で表されるので、二つの表現法に対する具体的な学習法は、それぞれ次のようにして与えられる。

## a. 離散表現の場合

離散化した指の位置  $\tilde{x}$  と運動指令  $m$  の組合せに対する分布を次のようにして更新する。

$$\sigma_{\tilde{x},m}^2(t+1) = \left[ \frac{1}{\sigma_{\tilde{x},m}^2(t)} + \frac{1}{\sigma_d^2} \right]^{-1} \quad (5.4)$$

$$\mu_{\tilde{x},m}(t+1) = \sigma_{\tilde{x},m}^2(t+1) \left[ \frac{\mu_{\tilde{x},m}(t)}{\sigma_{\tilde{x},m}^2(t)} + \frac{d(t)}{\sigma_d^2} \right] \quad (5.5)$$

ただし、 $d(t)$  は指の跳躍距離の観測値であり、 $\sigma_d$  は  $d(t)$  がもつ誤差の標準偏差である。

しかし、この学習法では、それまでに経験した部分に対してしか学習が進まないで、運動量の大きな部分に対しては学習が進行せず、いつまでたっても小さな運動を繰り返すことになる。そこで、 $x$ ,  $m$  と  $d$  との関係が連続的であると仮定して、トポグラフィ形成のモデル [57] と同様に、隣合う離散値  $\tilde{x}'$ ,  $m'$  についても次のようにして分布を更新する (図 5.6)。

$$\sigma_{\tilde{x}',m'}^2(t+1) = \left[ \frac{1}{\sigma_{\tilde{x}',m'}^2(t)} + c \frac{1}{\sigma_d^2} \right]^{-1} \quad (5.6)$$

$$\mu_{\tilde{x}',m'}(t+1) = \sigma_{\tilde{x}',m'}^2(t+1) \left[ \frac{\mu_{\tilde{x}',m'}(t)}{\sigma_{\tilde{x}',m'}^2(t)} + c \frac{d(t)}{\sigma_d^2} \right] \quad (5.7)$$

ここで、 $c$  ( $0 < c < 1$ ) は学習の強さを表す定数である。

このような技巧を用いることにより、内部像の学習は小さな運動に対応する部分から徐々に進行する。

#### b. 関数表現の場合

関数表現の場合には、パラメータベクトル  $\xi$  の平均  $\mu$  と共分散行列  $D$  を次式に従って更新する。

$$D(t+1) = \left[ D^{-1}(t) + \frac{1}{\sigma_d^2} \mathbf{f}(\mathbf{x}, m) \mathbf{f}^T(\mathbf{x}, m) \right]^{-1} \quad (5.8)$$

$$\mu(t+1) = D(t+1) \left[ D^{-1}(t) \mu(t) + \frac{d(t)}{\sigma_d^2} \mathbf{f}(\mathbf{x}, m) \right] \quad (5.9)$$

関数表現の場合には、関数の大域性から自然に汎化が行なわれるため、離散表現で用いた技巧は不要である。

#### 5.4.4 信頼度の制御

さて、Bayes 推定によってパラメータの学習を繰り返すと、確率分布の分散は一律に減少していくが、これは、確率モデルが対象の性質を完全に記述しているという仮定の下で推定を行なうためである。しかし、現実には、離散表現、関数表現のいずれの場合も、用意した内部表現の土台の上に対象を完全に表現できるとは期待できない。このように、表現ベースの能力が不足しているときに、内部像の学習が正しく進行しなくても分散が小さくなってしまふことを避けるため、分散が必要以上に小さくならない（内部像の信頼度を適切なレベルに保つ）ように、学習アルゴリズムを次のように変形する（図 5.7）。

1. 運動指令を出力して、跳躍距離を観測する。
2. 1. と並行して、内部像を用いて跳躍距離の期待値とその標準偏差を計算する。
3. 跳躍距離の予測値と実測値との差が、標準偏差（あるいはその定数倍）の範囲に収まっているかどうかを調べる。
4. 予測した範囲内に収まっている場合には、前述のアルゴリズムに従って平均と分散を更新する。そうでない場合は、分散の値を予測値と実測値の差の大きさに応じて増加させる。

具体的には以下の手続きを踏む。まず、跳躍距離の標準偏差は、離散表現の場合は、

$$\Delta \hat{d}(t) = \sigma_{\mathbf{z}, m}(t) \quad (5.10)$$

関数表現の場合は、

$$\Delta \hat{d}(t) = \sqrt{\mathbf{f}^T(\mathbf{x}, m) D(t) \mathbf{f}(\mathbf{x}, m)} \quad (5.11)$$

として与えられる。そして、実測値  $d(t)$  と予測値  $\hat{d}(t)$  との差、すなわち予測誤差を

$$\Delta d(t) = |d(t) - \hat{d}(t)| \quad (5.12)$$

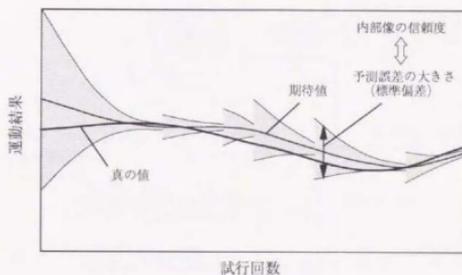


図 5.7 信頼度の制御

と表したとき、 $\Delta d(t)$  と  $\Delta \hat{d}(t)$  の比

$$r = \frac{\Delta d(t)}{\Delta \hat{d}(t)} \quad (5.13)$$

が、閾値  $r_0$  より小さいときには学習を継続し、そうでないときには分散を増加させる。分散を増加させる場合は、離散表現のときは単純に

$$\sigma_{x,m}(t+1) = kr\sigma_{x,m}(t) \quad (5.14)$$

として新たな分散を定める。ただし、 $k$  は増加率を定める定数である。一方、関数表現のときは、共分散行列の対角成分に誤差の増加分を負わせる。

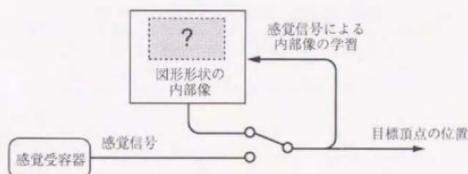
$$D_{jj}(t+1) = kr^2 D_{jj}(t) \quad (j = 1, \dots, J) \quad (5.15)$$

このような拡張を加えることによって、内部像の表現能力を補うだけでなく、内部像を運動系の特性の時間変化に追従させることも可能になる。

#### 5.4.5 図形形状の内部像を含めた運動計画のアルゴリズム

さて、実際になぞり運動を行なう場合には、運動を繰り返すにつれて、運動器官に対する内部像だけでなく、図形の形状に対する内部像の学習も進行する。そして、形状の内部像が形成されると、視覚情報がなくてもそれを利用して運動指令が生成できるようになる。人間が図形をなぞるとき、視覚情報が遮断されていてもパフォーマンスが向上するのは、運動を繰り返す過程で体性感覚情報をもとに図形形状の内部像が形成されるためである。すなわち、図形形状の内部像のあいまいさが小さくなると、感覚フィードバックに頼らずに運動指令を出力できるために、すばやい運動が可能になるのである。

本節では、図形の内部像を頂点位置の期待値と分散という形で表現することによって、上述のシステムに図形の内部像を組み込むことを考える。いま、扱った図形は長方形に限定し、4つの頂



図形状の内画像の信頼度が向上すると、感覚信号がなくても頂点の位置を知ることができる。

図 5.8 情報源としての内部像とその選択

点の座標を  $(x_i, y_i)$  ( $i = 1, \dots, 4$ )、その内部像の期待値、標準偏差をそれぞれ  $(\mu_{x_i}(t), \mu_{y_i}(t))$ 、 $(\sigma_{x_i}(t), \sigma_{y_i}(t))$  と表す。

形状の内画像の更新は、運動器官の内画像の更新と同様に行なう。

$$\sigma_{x_i}^2(t+1) = \left[ \frac{1}{\sigma_{x_i}^2(t)} + \frac{1}{\sigma_{x_i}^2} \right]^{-1} \quad (5.16)$$

$$\mu_{x_i}(t+1) = \sigma_{x_i}^2(t+1) \left[ \frac{\mu_{x_i}(t)}{\sigma_{x_i}^2(t)} + \frac{\tilde{x}_i(t)}{\sigma_{x_i}^2} \right] \quad (5.17)$$

ただし、 $\tilde{x}_i(t)$  は頂点の位置  $x_i$  の観測値を表す。 $y_i$  についても全く同様である<sup>1</sup>。

正規分布に従う二つの変数の和もまた正規分布に従うことから、運動計画のアルゴリズムは、上述のアルゴリズムにおける内部像の分散として、運動器官の内画像と形状の内画像の分散の和を用いることによって同様にして構成できる。

頂点の情報を感覚信号から得る（実際に観測行為を行なう）か、内部像から得る（観測を行わずに頭の中だけで判断する）かの選択は、図形に対する内部像のあいまいさに応じて定められよう。すなわち、内部像が提供する情報量が十分であれば（内部像のあいまいさが十分小さければ）観測を行なう必要はなく、逆に、不十分であれば、感覚信号から情報を得、それと同時に内部像の学習を行なえばよい。したがって、最初のうちは、視覚情報を用いて内部像の学習を行なうが、そのあいまいさが十分減少した時点で、視覚フィードバックを取り去ることが可能になる。

なお、図形の頂点の位置を知るために内部像を選択するか、現実の感覚信号を選択するかは、情報源の選択の問題、すなわち能動的認識の問題として捉えることができる。つまり、内部像のあいまいさが減少し、制御を行なう上で十分な情報を提供できるようになった時点で、内部像が選択されるようになると解釈すれば、以上の過程は能動的認識の枠組みの中で理解することができる（図 5.8）。

<sup>1</sup>  $x_i$  と  $y_i$  を組にして考えるときは、2次元の正規分布を考えればよい。

表 5.1 基底として用意した関数の種類

$f_1: 1$	$f_2: m$	$f_3: x$	$f_4: y$
$f_5: mx$	$f_6: my$	$f_7: xy$	$f_8: m^2$

## 5.5 数値実験

前節で構成したアルゴリズムの振舞いを調べるため、数値実験を行なった。

### 5.5.1 実験の設定

実験において、図形は  $xy$ -平面上で  $(0,1) \times (0,1)$  の領域に提示されるとし、4 辺が  $xy$ -軸に平行な長方形とした。取り得る運動指令  $m$  の種類は 0.05 から 1.0 まで 0.05 間隔の 20 種類とした。頂点を越える確率の許容値  $\theta$  は 0.05 を用いた。観測誤差の標準偏差  $\sigma_d$  は 0.01 で一定とし、また、指先中心から頂点までの距離が 0.05 以下になったとき、頂点に達したことがわかるものとした。

離散表現の場合は、上記の領域を  $20 \times 20 = 400$  個の小正方形領域に分割し、それぞれに対して一つの確率分布を割り当てた。また、隣接する分布に関する学習定数  $c$  として 0.015 を用いた。これにより、隣接する分布に関しては、等価的に観測誤差約 0.08 で学習が行われることになる。

分散を増加させる閾値は  $r_0 = 1$  とし、また、分散の増加倍率は  $k = 1$  とした。関数表現の場合は、基底となる関数群  $\{f_i\}$  として表 5.1 に示した 8 個の関数を用意した。

以下、指の位置  $(x, y)$ 、運動指令  $m$  と跳躍距離  $d$  の関係が、

$$d = 0.5(m - 0.05)(1 - 0.3x + 0.3y) \quad (5.18)$$

である場合の結果を紹介する<sup>1</sup>。

### 5.5.2 実験結果

図 5.9、図 5.10 にそれぞれ、内部モデルとして離散表現、関数表現を用いたとき、1 周目、2 周目、3 周目、10 周目に指が停留した位置を示す。いずれの場合も、指が頂点を行き過ぎてしまうことはなく、未学習時でも指が図形の頂点を行き過ぎないという目標を確実に実現していることがわかる。実際、10 個の図形に対してそれぞれ 10 周ずつなぞるという実験を行なったところ、指が頂点を行き過ぎたのは離散表現の場合に 6 回、関数表現の場合は 1 回に過ぎなかった。また、学習が進むにつれて 1 回の跳躍距離が大きくなるため、1 周に要する時間が少なくなることもわかる。

1 回の跳躍距離が大きくなるということは、その分だけ頂点を探る道程が短くなることを意味しており、この点で、システムの振舞いは赤松らによる心理物理実験の結果とよく適合してい

<sup>1</sup>基底関数や運動・跳躍距離間関係として人間の腕の特性を反映した形を設定しても同様の議論ができる。

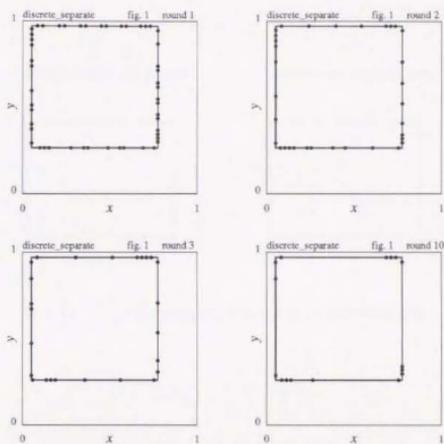


図 5.9 離散表現を用いたときの指の停留位置

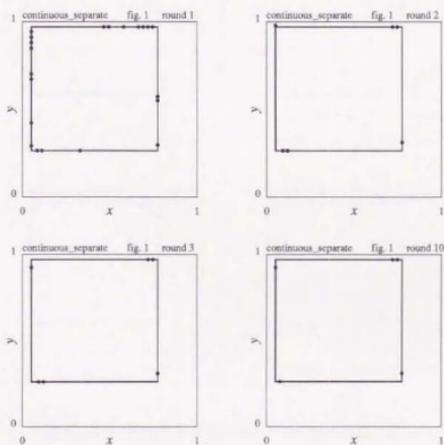


図 5.10 関数表現を用いたときの指の停留位置

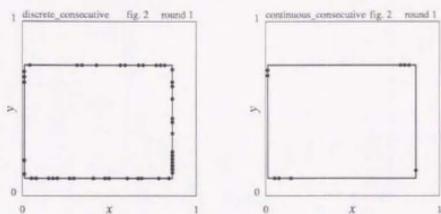


図 5.11 二つ目の図形をなぞらせたときの指の停留位置

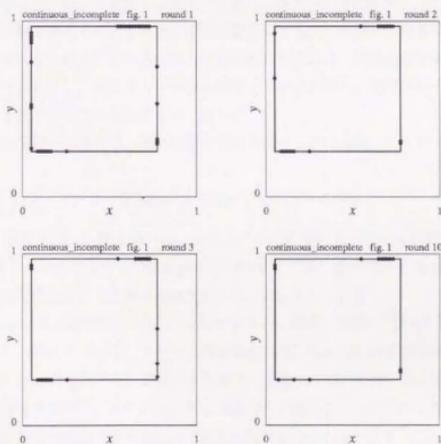


図 5.12 不完全な内部表現の下での指の停留位置

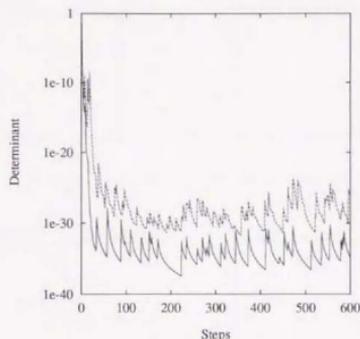


図 5.13 行列式の変化

る。また、内部像が確実なものになるにつれて、1周に要する時間が少なくなるという事実は、運動タスクに慣れるにつれて動きがすばやくなるという我々の日常的経験と一致する。

離散表現法と関数表現法との大きな違いは汎化能力にある。この違いが顕著に現われるのは、一つの図形をなぞらせたのち、別の図形をなぞらせた場合である。関数表現を用いた場合は、二つ目の図形をなぞるとき、1周目から停留点の数が少ないのに対し、離散表現を用いた場合は最初の図形のときと比べてほとんど差がない(図 5.11)。

次に、関数表現を用いた場合に、基底関数が完全でないときの振舞いについて紹介する。跳躍距離が、

$$d = 0.5(5m^2 - 0.02)(1 - 0.3x + 0.3y) \quad (5.19)$$

で与えられるとき、1周目、2周目、3周目、10周目の指の停留点を示したものが図 5.12 である。図 5.10 と比較するとわかるように、基底が不十分な場合でも、指が頂点を行き過ぎることはないが、十分な学習を繰り返しても指の停留点の数は一定以上減少しない。

また、図 5.13 は、真の跳躍距離が式 (5.18) で表される場合(実線)、式 (5.19) で表される場合(破線)の二つの場合について、それぞれ共分散行列の行列式  $|D|$  の変化の様子を示したものである。この図からわかるように、基底が不完全である場合には、完全である場合に比べて分散が高いレベルで留まっており、その結果大きな運動指令が出力されないものと考えられる。

最後に、図形形状の内部像もあわせてもつ場合の結果について紹介する。

ここでは、視覚情報が完全に遮断されたとして、体性感覚情報だけから頂点の位置がわかるものとする。実験条件は上と同じように設定するが、位置観測の誤差は  $\sigma_d = 0.05$  とする。

図 5.14 に、1周目、2周目、3周目、10周目に指が停留した位置を示す。この結果は、内部表現形式として関数表現を用いたときの結果である。1周目は、頂点の位置が全くわからない状態で

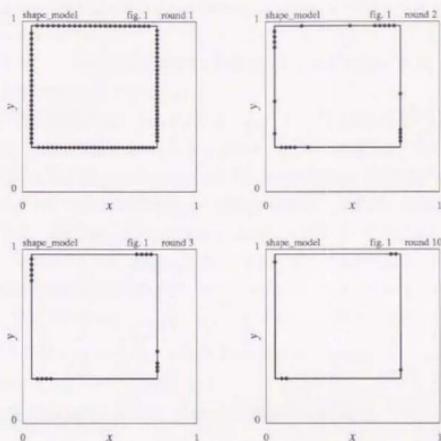


図 5.14 図形の内部像を考慮した場合の指の停留位置

あるので、指は非常にゆっくりした速度でしか動かせない。しかし、周回を繰り返していくうちに内部像の信頼度が向上し、それによって跳躍距離が長くなっていく。

## 5.6 考察

### 5.6.1 長期的な学習と短期的な適応の組合せ

人間の視触覚変換はきわめて柔軟性に富んでおり、視触覚間の関係に変形を加えても短い時間でそれに適応する一方、しばしば自分の手を自分の眼でみないとその変換が容易に失われてしまうことが知られている [28]。

しかし、視触運動間の関係がすぐに失われてしまうといっても、その関係が完全になくなってしまうことはない。これらの事実を考慮すると、視覚・運動変換は、精度が粗く適応も遅い成分と、精度が高く適応も速い成分の二つの要素からなっており、前者が大まかな傾向を与え、後者が細かな対応関係を実現していると考えられる。また、後者は減衰が速く、学習が中断すると短い時間でその情報を失ってしまうものと思われる。

このように視覚・運動変換が柔軟に作られているのは、筋内の代謝や疲労などによって、制御対象の特性が時々刻々変化していくためではないかと考えられる。特性が変化するのであれば、ある時点で精確な関係を学習したとしても、それを使う段ではずれが生じてしまうため、その学習は無駄になってしまう。もう一つの理由として、高精度の変換を長期間記憶しておくためには

それだけ多くの記憶資源を必要とすることが挙げられる。脳はこの変換にそれだけ多くの記憶資源を投入しないのかもしれない。

上記のシステムでは、このような長期成分と短期成分を分けずに議論したが、このことを考慮した形に拡張するには次のようにすればよい。

まず、パラメータの分散が時間とともに大きくなるように、すなわち忘却するように学習則を変形する。ただし、分散が無限に大きくなると情報が完全に失われてしまうので、指数関数的にある限界値に近づいていく形にするのが適当である。この限界値が、長期成分のもつ精度に相当する。一方、パラメータの期待値を長期成分と短期成分に分け、長期成分を短期成分の移動平均によって修正しつつ、過渡的に変化した短期成分は減衰させていく。これによって、大まかな関係は長期成分として保持する一方、過渡的な適応は短い時間で忘却するようにできる。

この過程を具体的に数式を用いて表せば、次のようになる。いま、パラメータの短期成分は、長期成分の期待値、分散をそれぞれ  $(\mu_S, \sigma_S^2)$ 、 $(\mu_L, \sigma_L^2)$  と表し、これらを合成したパラメータは、期待値  $\mu_S + \mu_L$ 、分散  $\left(\frac{1}{\sigma_S^2} + \frac{1}{\sigma_L^2}\right)^{-1}$  の正規分布に従うものとする。このとき、短期成分、長期成分をそれぞれ次のようにして更新する。

a) 短期成分の更新 短期成分の更新は、前節と全く同様に行なう。

$$\sigma_S^2(t+1) = \left[ \frac{1}{\sigma_S^2(t)} + \frac{1}{\sigma_d^2} \right]^{-1} \quad (5.20)$$

$$\mu_S(t+1) = \sigma_S^2(t+1) \left[ \frac{\mu_S(t)}{\sigma_S^2(t)} + \frac{d(t)}{\sigma_d^2} \right] \quad (5.21)$$

一方、短期成分の忘却は次式のようにして行なう。

$$\sigma_S^2(t+1) = \sigma_S^2(t) + \sigma_f^2 \quad (5.22)$$

$$\mu_S(t+1) = (1 - \epsilon_S)\mu_S(t) \quad (5.23)$$

ここで、 $\sigma_f$ 、 $\epsilon_S$  は忘却の大きさを表すパラメータである。これらの式は、短期成分の期待値が時間とともに減衰し、また分散が増加することを示している。また、短期成分の分散は時間とともに無限大に発散するが、合成パラメータの分散を上のように定めたので、パラメータ全体としての分散は漸的に  $\sigma_L^2$  に近づく。

b) 長期成分の更新 長期成分の期待値は、短期成分の変動にゆっくりと追従するように微量だけ変化させる。また、長期成分の分散は別の条件によって定まるとし、このような動的な適応によっては変化しないと考える。

$$\sigma_L^2(t+1) = \sigma_L^2(t) = \text{const.} \quad (5.24)$$

$$\mu_L(t+1) = \mu_L(t) + \epsilon_L \mu_S(t) \quad (5.25)$$

### 5.6.2 基底関数の学習

一般に、学習の問題は内部表現様式と切り離して論じることができないが、このシステムにおいても適応の進行速度や最終的に達成される誤差の大きさは、どのような関数を基底関数をとる

かに依存する。本章の議論の要点は、基底関数が不完全であっても、運動計画がそれをカバーすることで目標行動を間違いなく実現できることにあるが、これとは別に、基底関数を充実させることによって本質的にパフォーマンスを向上させることも可能である。実際の人間もまた、練習を繰り返す過程で有用な基底関数を獲得し、運動能力を向上させているものと考えられる。

これまで提案されてきた種々の運動学習モデルは、まさにこの基底関数獲得過程を捉えたものであるといえる。したがって、従来の学習運動系のモデルを基底関数の学習モデルとして取り込むことによって、本章のシステムは総合的な学習運動系のモデルとして発展させることができる。

この問題については、6.2節において内部表現の獲得という視点からあらためて考察する。

### 5.6.3 スティフネス調節との関係

人間は、運動に不慣れなあいだは、拮抗筋を使って身体を硬くする（スティフネスを大きくする）一方、その運動に慣れてくると、余計な力を使わずにスムーズに運動できるようになる[40]。

身体を硬くすることは弾性定数を大きくすることに相当するので、それによって身体の動きを抑えるような作用が働き、運動時の誤差を小さく保つことができるようになる。このことから、光山ら[71]は、不慣れな運動をする際に身体を硬くするのは、軌道学習時の運動軌道を安定化するためであるという解釈を与えている。本節では、この現象についてあらためて考察を加える。

以上で述べてきたように、人間は、不慣れな運動を行なう際、「小さな運動指令を出力し、その結果が目標とする軌道からはずれていないかを視覚フィードバックによって評価する」ということを頻繁に繰り返しながら運動を遂行する。このとき、拮抗筋を働かせて身体を硬くすれば、その分だけ運動指令に対する身体の感度が低くなるので、運動指令の変動に対して生じる身体の動きを小さく抑えることができる。これにより、人間は自分の身体が不意に動いてしまうのを抑えているといえる。

一方、運動を繰り返す過程で、自分の描く軌道が目標軌道からはずれていないことが確実に従って、視覚フィードバック、すなわち自分の描く軌道を監視する必要性は低くなり、その結果一度に大きな運動ができるようになる。この段階では、身体を柔らかくし小さな運動出力で効率よく動くようにする方が、エネルギー効率の点から見て望ましい。そこで、自分の運動が確実に従って、拮抗筋の働きを弱くし身体の感度を高くしていると解釈できる。

このように考えると、人間は、自分の身体が目標軌道からはずれていないかということに関する「自信」すなわち「信頼度」に応じて、自分の身体の感度を変化させていると考えられる。言い換えれば、人間は、身体の運動インピーダンスを能動的に定めることによって、目標の運動を間違いなく行ないながら、学習の進行に従って効率的な運動を実現しているといえる。

このように、拮抗筋を用いた腕のインピーダンス制御も、「許された範囲で目標軌道を守るような運動指令を出力する」という考え方の下で統一的に理解することができる。具体的なアルゴリズムについては未検討であるが、運動指令生成のメカニズムと注意のメカニズムをあわせて検討することによって、上述のアルゴリズムを拡張する形で実現できると思われる。

## 5.6.4 Fittsの法則との関係

人間がリーチング運動を行なう際にかかる時間に関する知見として、Fittsらによる心理物理実験[18, 19]がよく知られている。彼らは、移動距離を  $A$ 、目標の大きさを  $W_S$ 、運動に必要な時間を  $t$  とすると、

$$I_p = -\frac{1}{t} \log \frac{W_S}{2A} \quad (5.26)$$

がほぼ一定になることを示し、 $I_d \equiv -\log \frac{W_S}{2A}$  を運動の難しさを表す指数と定義している。以下、この法則について本章の考え方に基づいて考察を加える。

まず、「移動距離一定のとき目標の大きさが小さいと運動にかかる時間がかかる」ことについて考える。上の実験を本章の視点から捉えると、「目標の大きさ」は「運動に際して許容される誤差」に相当するので、目標が小さいことは許容誤差が小さいことを意味する。したがって、目標が小さいときには、目標からはずれる可能性が高い大きな運動よりも、誤りの修正ができる程度の小さな運動を選択することになるため、結果として時間がかかると解釈できる。逆に、目標の大きさが十分大きな時は、許容される誤差が大きいため、運動結果の精度を気にすることなく大きな運動を起こすことができ、その結果運動時間が短くなるものと考えられる。

また、Fittsの法則によれば、「移動距離と目標の大きさが比例するとき、運動にかかる時間が等しくなる」が、このことを本章のアルゴリズムと対比させると、「移動距離の大きさに応じて内部像のあいまいさが大きくなる」という結論が導かれる。ただし、この問題については、その原因を単純に内部像のあいまいさに帰着することはできない。

この問題に関係する一つの因子として、目標までの距離の評価誤差が挙げられる。本章で述べたアルゴリズムでは、距離の評価誤差として目標位置の観測誤差  $\sigma_d$  のみを考えたが、目標までの距離が大きくなると、目標位置を把握するのに眼球運動や首の運動を伴うようになるので、距離を評価する際の誤差が大きくなる可能性が高い。距離評価のあいまいさが大きくなれば、それに伴って誤りを犯す確率が大きくなるので、等価的に許容誤差を小さくする効果が生じる。逆に、目標の大きさが距離の評価誤差を相殺する程度に大きくなる状況では、運動距離の大きさに依存せずに運動指令を生成することができ、そのため運動に要する時間もあまり変化しないと考えられる。

さて、Fittsの法則を説明するモデルとして、従来 deterministic iterative correction model と呼ばれるものを用いられることが多かった[17]。このモデルは、リーチング運動は視覚フィードバックに基づいて定められる離散的な部分運動の列から構成されるというものである。

このモデルではまず、各部分運動は常に一定の時間  $t$  をかけて実行され、部分運動によって移動する距離は目標の中心までの距離の一定割合  $p_d$  であるという仮定をおく。したがって、目標の中心までの距離が  $A$ 、目標の大きさが  $W_S$  である場合、最初の部分運動の距離は  $p_d A$ 、二つ目は  $p_d A(1 - p_d)$ 、三つ目は  $p_d A(1 - p_d)^2$  というようにして与えられる(図 5.15 a)。このようなモデルを用いることによって、Fittsの法則をはじめとして、いくつかの現象を説明できることが知られている。

しかし、このモデルの振舞いは、以下の点で人間の行動と一致しないことが指摘されてきた。

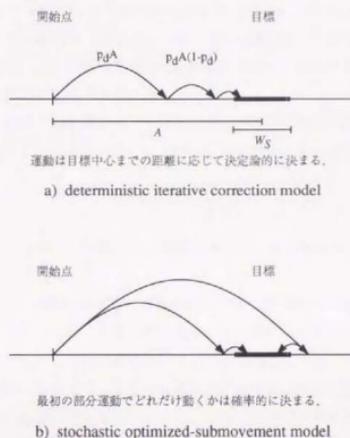


図 5.15 Fitts の法則を説明する二つのモデル

1. このモデルでは部分運動に要する時間を常に一定としているが、人間の場合は、最初の部分運動に要する時間が、目標までの距離や目標の大きさに依存して変わること。
2. このモデルでは、各部分運動によって指が跳躍する距離が比例的に減少するとしているが、人間の運動ではそれが成り立たないこと。
3. このモデルによれば、最初の部分運動だけで目標に達するか否かは決定論的に定まるのに対し、人間の運動ではその可否が確率的にばらつくこと。
4. 人間の運動の中には、部分運動に分解できないような運動があること。
5. このモデルでは身体が目標を行き過ぎることはありえないが、人間の行動ではしばしばそのようなことが見られること。
6. このモデルでは、視覚フィードバックのない状況での運動を全く説明できないが、人間の場合は、視覚フィードバックがなくとも運動にかかる時間が  $D/W$  の対数に比例すること。

これらの問題を解決するため、Meyerら [68] は、stochastic optimized-submovement model という考え方を提案している。これは、運動による移動距離が雑音の影響によってばらついていくという仮定の下で、運動にかかる時間の期待値を最小にするような運動指令を選択するというものである。

このモデルではまず、運動制御系に含まれる雑音のために、運動結果にばらつきが生じるという仮定から出発する。また、リーチング運動は、目標までの距離や目標の大きさにかかわらず、

高々二つの部分運動から構成されているとする(図 5.15 b)。すなわち、最初の部分運動で目標に達した場合はそれで運動が終了するとし、そうでない場合は、二つ目の部分運動で目標に達するものとする。さらに、各部分運動での移動距離は正規分布に従ってばらつき、その標準偏差は平均速度に比例して決まるという仮定をおく。

以上の条件の下で、最初の部分運動と二つ目の部分運動の所用時間の和の期待値が最小になるように、各部分運動の所用時間を定める。以上の設定の下で求められる運動全体の所用時間は、

$$T = k_1 + k_2 \sqrt{\frac{A}{W_S}} \quad (5.27)$$

という形で表されるが、この式から得られる運動時間は、一定の領域で Fitts の法則とはほぼ一致する。

Meyer らの考え方は、運動計画において確率的に最適な運動指令を選択するという点で本章で述べた考え方に近いといえるが、確率モデルを用いる対象が全く異なっている。すなわち、本章では、運動器官の特性に対する内部像のあいまいさを表現するために確率モデルを用いたのに対し、彼らは、同じ運動をしようとしても神経系内部の雑音のために運動結果がばらついてしまうことを表現するために確率モデルを用いている。なお、運動器官の特性自体が確率的であるとした場合でも、特性のばらつきを内部像のあいまいさに吸収することができるので、本章で述べた運動計画アルゴリズムは同様にして機能する。

運動計画は、目標位置、身体の姿勢、運動器官の特性などさまざまな因子が総合されて行なわれるものと考えられるが、いずれにせよ、人間は、それらについて完全に知りえない状況で、足りない情報を頭の中で補いながら判断を下し、運動指令を生成していると考えて間違いない。この点で、「内部像に基づく対象の理解」という考え方は、運動計画や学習制御のメカニズムについて議論する上でも有効であるといえるだろう。

## 第6章

### 考察

本章では、以上の章で扱うことのできなかった問題のうち、聴覚認識と内部表現の自己組織化の二つを取り上げ、それぞれについて考察を加える。

#### 6.1 聴覚知覚過程における能動的認識

##### 6.1.1 音源分離とカクテルパーティ効果

本節では、能動的認識のアルゴリズムを聴覚知覚過程に適用することについて考察する。

聴覚過程を考える上で新たに生じる問題は、時系列信号の取扱いである。時間の変化を伴う能動的認識のアルゴリズムについては2.4節でも取り扱ったが、そこでの議論は対象の状態変化が鏡形モデルで記述できる場合に限られていた。これに対し、聴覚認識を実現するためには、対象の状態がシンボルで表される場合も含めてより一般的な時系列を取扱う方法を明らかにしなければならない。

議論を行なうための題材として、ここでは「音源分離」の問題を取り上げる。音源分離とは、同時に鳴っている複数の音を分離して知覚することである。音源分離の日常的な例として「カクテルパーティ効果」が挙げられる。これは、パーティ会場のようにいろいろな人の声が交錯する中で、自分が注意を向けた特定の人の声だけが浮かび上がって聞こえてくる現象である。これは、自分にとって関心のある情報だけを選択的に処理するという意味で、能動的認識について論じる上で格好の題材であるといえる。同様のことは、オーケストラの中から特定の楽器の旋律が聞こえてくるという場面でも経験される。そこで、以下では、同時に鳴っている複数の楽器の音を分離する問題について議論する。なお、最近、音響情報をもとに環境を認識するメカニズムを明らかにしようとする研究が、“auditory scene analysis”の名の下で、心理学的、および工学的視点から行なわれている[13,14]。

音源分離の問題には、時系列認識の問題のほかにもう一つ新しい問題が含まれている。すなわち、これまでの議論では、認識すべき対象がただ一つ与えられている状況で、対象を効率よく識別するための感覚情報を選択する方法について議論してきたのに対し、音源分離の問題では、対象からの信号が雑音の中に埋もれていたり複数の対象が同時に提示されたりする状況で、特定の

対象の状態を推定する方法を考えなければならない。具体的には、捉えた情報の中のどの部分が同一の対象から生成されたのかを定める、群化あるいはセグメンテーションのアルゴリズムを構成する必要がある。そして、群化のメカニズムと能動的認識のメカニズムが相互作用しながら働く仕組みを考えることによって、特定の楽器の音に注意を集中するという人間の高度な感覚機能をモデル化することができると考えられる。

さて、楽器音の分離については、自動採譜システムという純粋に工学的な視点から研究が行なわれているほか[32, 33, 49, 115]、「視聴覚融合」という考え方に基づいた方法が提案されている[114]。

これは、聴覚信号と相関のある視覚信号(例えば、口の動き、楽器の動き、楽器を弾く指など)を手がかり信号(cue signal)として、聴覚信号の中でその手がかりと相関のある成分だけを通すような適応フィルタを構成することによって、目標とする音響信号だけを取り出そうとするものである。実際、人間は、口の動きを見ている方が、見ていないときに比べて相手の話す言葉をよく理解できることが知られており、その点で、この研究は人間の感覚統合の一面を捉えたものであるといえる。なお、原理からわかるように、手がかりとして用いる信号は、目標とする聴覚信号と相関のある情報であれば視覚情報でなくてもよい。

この方法の特色は、相関処理と適応フィルタリングという信号処理レベルで閉じていることである。このため、1) 入出力とも音響信号である、2) 音源の種類を選ばない、という特徴を有しており、「S/N比を向上させる前処理」として優れた方法であるといえる。ただし、この方法は、音源分離の過程を純粋に信号処理過程として実現したものであり、認識の過程にまで踏み込んだものではない。

この方法とは対照的に、音源の性質を積極的に利用する、すなわち、各音源から出される音の性質の違いをきっかけにして音源を分離する方式も考えられる。この方式の特徴としては、1) 出力されるのは音響信号そのものではなく認識結果である、2) 手がかり信号を必要としない、3) 音源の性質をあらかじめ知っておく必要がある<sup>1</sup>といったことが挙げられる。

このような考え方に基づいて実際に構成されたシステムとして、植野[45]によるものがある。このシステムは、楽器音を周波数分析した後、倍音構造や立ち上がり、立ち下りの同期性などに基づいて各周波数成分間の距離を計算し、その距離に基づいて周波数成分をグループ分けすることによって異なる楽器音を分離するもので、電子楽器のよる二重奏の音を楽譜化できることが実験的に示されている。このシステムは基本的にボトムアップだけの処理に基づくものであるが、後に、トップダウンの処理を加えることによってシステムの拡張がはかられている[46, 77]。

### 6.1.2 聴覚知覚におけるストリーム形成

聴覚知覚過程において注目すべき特徴は時間方向のつながりであり、それが最もよく現われている現象が「ストリーム形成」と呼ばれるものである。

<sup>1</sup>オンラインで音源の性質を同定することも可能である。

図 6.1 ストリーム形成の例 (文献 [47] より引用)

「ストリーム」とは、音響事象に対応する知覚的なまとまりのことである。これは、物理的なつながりではなく、知覚上のつながりを表す概念である。ストリームが形成される例を図 6.1 に示す。左側の図は、同じ音の系列でもその速度に応じて聞こえ方が異なることを示しており、音情報はその順番だけが意味をもつ単純な時系列として捉えられているわけではないことを示している。一方、右側の図は、左右の耳に交互に音を入れ替えて聞かせても (図 b)、音は連続的に変化しているように聞こえること (図 c) を示している。このような現象は、耳に加えられた信号が物理的に不連続であっても、知覚上は連続して聞こえることを意味している。

ストリーム形成の重要性を表す例として、「ストリーム形成によるマスキング解除」が挙げられる。マスキングとは、強い音と弱い音を同時に鳴らすと弱い音の知覚が妨げられる現象である。この現象には、1) 妨害音の帯域幅が狭いときは、信号音と妨害音の周波数差が小さいときにマスキングの効果が大きい、2) 中心周波数が一致している場合は、妨害音の帯域幅が大きい方がマスキングの効果が大きい、あまり (100~200Hz 以上) 大きくしても効果はない、といった性質があるため、従来比較的狭い範囲の信号音と妨害音との強度比によって説明できるとされてきた。しかし、最近これに加えて、「雑音の振幅が異なる帯域を通じてゆっくりと変動するとき、帯域が広がるにつれてマスキングの効果が減少する」という現象 (共変調マスキング解除) が見いだされている。これは、雑音が一つのストリームとして知覚されることによって、「雑音でなくなる」ために生じる現象であると考えられている。

このように、人間の聴覚には、ストリームを形成しそれを単位として音響現象を理解することによって、他の原因で生じる音響信号による干渉を排除する機能が備わっている。このことは、聴覚認識モデルを構成する上でストリーム形成メカニズムを組み込むことが本質的であることを示唆している。

また、人間は、会話をしている最中に雑音によってその音がマスクされても、相手の音声がか断されたような印象を受けることなく、その内容を正しく聞き取ることができる。音声の場合に

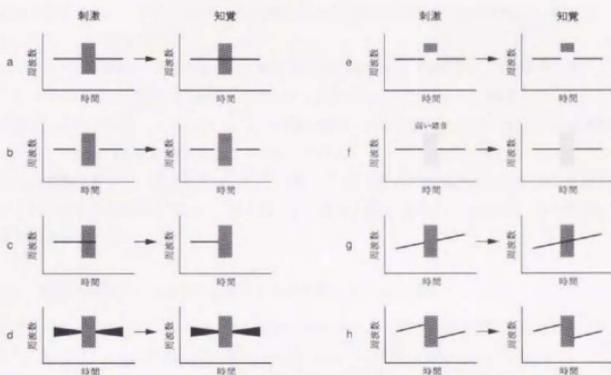


図 6.2 音の知覚の連続性 (文献 [47] より引用, 改変)

は文脈など非音響的な情報を用いて情報の補完を行なっている可能性が高いが、文脈をもたないより単純な音響現象についても聴覚的な補完現象が起こることが報告されている。

いま、一定周波数の純音に対して、広い周波数帯域をもつ雑音がかぶさっている状況を考える。このとき、

1. 雑音の長さが 300ms 程度あっても、純音が連続して聞こえることがある。
2. 信号音と雑音の間に隙間があると、連続して聞こえない。
3. 信号音が雑音と同時に鳴りやんだ場合には、信号音は雑音の始まるの時点で鳴りやんだように聞こえる。
4. 雑音が信号音をマスクする程度の帯域や強度をもっていなければ、純音は連続して聞こえない。
5. 妨害音の前後で信号音の周波数が変わっていてもよいが、このとき、信号音の周波数変化が連続的である場合の方が補完されやすい

といった現象が起こることが報告されている (図 6.2)。さらに興味深い現象として、人間の会話を録音したものを 50ms おきに無音部分をいれて再生すると聞き取りにくくなるのに対し、無音部分の代わりにホワイトノイズをいれて再生すると無音の場合ほど聞き取りにくい、という現象が見いだされている。

このようなことから、信号音が補完されるためには、

1. 信号音が妨害音の前後で一連の音響事象であること、
2. それが中断していることを示す積極的な証拠がないこと

が重要であることがわかる。特に、2.の条件は、視覚探索における探索非対照性と類似した性質としてたいへん興味深い。

以上で述べた現象は、人間が耳で捉えた物理的な信号をそのまま解釈しているのではなく、前後のつながりなどの情報構造を利用して頭の中に形成した内部像を介して解釈していることを強く示唆するものである。このことは、「内部像に基づく対象の理解」という考え方が聴覚知覚過程をモデル化する上でも有効であることを示している。

以上の考察に基づき、現在筆者は、同時に鳴っている複数の楽器音の中から特定の楽器音に注目し、その音程を識別するアルゴリズムについて検討を進めている。次節では、その構想について簡単に紹介する。

### 6.1.3 能動的認識アルゴリズムに基づく音源分離システムの構想

以下ではまず、楽器が一つの場合に、楽器の種類とその音程を識別する能動的認識システムを構成し、続いて、それを複数楽器の場合に拡張するという段階を踏むことにする。

まず、楽器音を分離するための手がかりとして、以下の情報が有効であると考えられる。

**音源の位置** 両耳時間差や両耳強度差から音源の位置に関する情報が得られる。

**音源の性質** 音の倍音構造、その時間的変化が楽器ごとに決まっているので、その構造を用いる。

**周波数成分の同期性** 同じ楽器から出た周波数成分はその振幅が同期するので、それらをグループ化することで周波数成分を分離できる。

**音程の連続性** 同一の楽器からは同じ音程の音がしばらく連続して出力されること、音程はしばらく連続的に変化することなどから、時間方向に音程変化を予測することで音程の候補を絞り込むことができる。

ここでは、楽器の位置は一定とし、楽器の性質を表す情報として音の倍音構造やその時間的変化だけが利用できるものとする。そこで、本システムにおいては、まず音を周波数分解した後、各周波数成分の強度やその時間的変化を基本的特徴量として計算する。これは、視覚系における初期特徴の計算に相当する。

このようにして得られた特徴情報を統合することによって、楽器の種類を識別することが可能になる。しかし、音響情報においては、特徴量が時間的に一定でなく、楽器ごとに一定の規則に従って時間的に変化すると考えられる。そのため、各楽器に対してそれぞれ特徴の時間的変化を記述するモデルを用意する必要がある。具体的には、楽器音の内部表現として隠れマルコフモデルを用いることを考えている。

以上の設定において、観測した音響信号を最もよく説明する楽器モデルを判別する際に、能動的認識のアルゴリズムを適用することについて考える。すなわち、楽器音を識別するために有効な情報をもたらす特徴量を逐次的に取り込むことによって、隠れマルコフモデルの選択を行なう。この段階において第2章で述べたアルゴリズムが適用できると思われるが、そのためには、あらかじめ交互作用のない特徴量を抽出しておく必要がある。

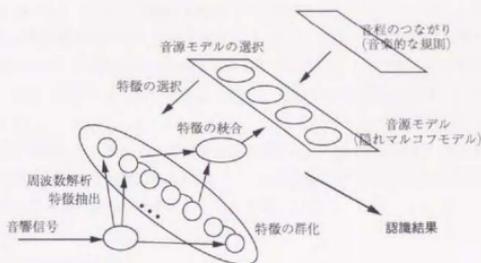


図 6.3 能動的聴覚認識システムの概念図

さて、楽器音を認識する際には、音が立ち上がる瞬間に多くの情報が含まれていることが実験的に指摘されているが [1]、このことは、楽器音を識別する上で、各周波数成分の立ち上がりを優先的に観測するようなメカニズムを組み込むことが有用であることを意味している。したがって、何らかのボトムアップな観測選択基準を定め、立ち上がり部分の特徴を優先的に選択するメカニズムについても考える必要があろう。

次に、以上のシステムを用いて複数音源からの音を取り扱う場合について考える。

この段階で必要なものとして、先に述べた群化のメカニズムが挙げられる。これは、振幅や立ち上がりの同期性など、音の各成分が有する共通の構造に基づいてそれらをグループ化するメカニズムであり、ストリーム形成を実現する上で鍵となるものである。なお、人間の聴覚過程において、雑音という全く未知の音響情報に対してもストリーム形成が起こることを考慮すると、群化のアルゴリズムは、時間的相関など単純かつ普遍的な原理に基づいて実現すべきであろう。

さて、複数の楽器が同時に鳴っている場合には、おそらく各楽器が生成する信号成分に対してそれぞれ群化のメカニズムが作用することになる。しかし、この過程で複数の群化メカニズムが相互作用すると、異なる楽器の成分が一まとめにされたり単一の楽器成分が複数の群に分けられたりするなどして、正常な群化が行なわれないことが予想される。このような群化の衝突は、目標音源を特徴づける特徴量の計算を妨げるので、その結果、楽器の認識はうまく行なわれないことになる。

この段階で、特定の楽器に関する情報処理を選択的に進行させる仕組みとして機能するのが、「注意」のメカニズムであると考えられる。その形態として、例えば、特定の楽器に関する群化を優先して進行させるメカニズムなどが考えられるが、その具体的な実現法については未検討である。

そして、注意を向ける際に重要な意味をもつのが、音響情報のもつ時間方向の連続性である。例えば、システムが捉えた複数の音程のうち、どれが着目している楽器から出たものであるかが判断できない場合には、その楽器がその前に出していた音程からのつながりで最も自然だと思

われる音程を選択するといった仕組みを検討する必要がある。もちろん、このようにして求めた音程が実際の音程と異なっている可能性はあるが、図 6.1 に示したように、人間の聴覚系が音程が連続的である音と同じ音源から出た音とみなしやすい性質をもっていることを考えると、むしろ、そのような誤りを犯すようなシステムの方が、人間の聴覚知覚過程のモデルとして有効であると考えられる。

以上の方針で構成したシステムの概念図を図 6.3 に示す。いずれにせよ、システムの具体的な現法についてはまだ構想の段階にあるので、ここでの議論は以上で止めることにする。詳細については、実験結果がまとまった段階であらためて報告したい。

## 6.2 知識と内部表現の獲得

本論文では、対象に関する知識あるいは内部表現を獲得する過程については触れずに、対象に関する知識がすでに得られているという仮定の下で、動的な感覚統合過程について論じてきた。しかし、認識の過程がそれまでに得られた知識から大きな影響を受けることは間違いないことである。本節では、認識過程の土台ともいえる知識や内部表現の獲得過程について考察する。

まず、知識の獲得過程には、次の二つの段階があると考えられる。

一つは、確率分布のパラメータの学習である。本論文で用いたアルゴリズムにおいては、各対象に対して信号や特徴量がどのような値をとるのかを表す条件付き確率分布をあらかじめ与えた上で、内部状態空間の更新や情報源選択基準の計算を行なった。これらのアルゴリズムを用いるためには、その条件付き分布を事前に獲得しておく必要がある。

確率分布の学習を行なう方法は、経験の積み重ねによって分布のパラメータを同定することによって尽きる。対象に関する知識がない場合は、いろいろな観測方法を試行錯誤的に用いることによって、分布のパラメータを定めていくことになる。学習の開始時点では、試行錯誤が重なるために認識のパフォーマンスは低くならざるを得ないが、経験を繰り返して分布のパラメータが定まるにつれて、次第に適切な観測行為だけが選択されるようになる。

ただし、確率分布の学習過程において、そのときまでに得られた知識だけをもとに能動的認識アルゴリズムを適用し最適な感覚器だけを利用していったのでは、そのほかの感覚器に関する知識が得られず、知識に偏りが生じることになる。したがって、学習の進行如何にかかわらず、さまざまな観測方法で対象を観測することが学習を平均的に進行させる上で重要である。この点において、「遊び」が経験に幅を与えるといったことが成立するといえるだろう。

二つ目は「内部表現」の生成である。ここでいう内部表現とは、対象を記述するための特徴量や対象と特徴量の関係を記述する確率モデルのことであり、このような表現のベースを経験を通じて獲得することが、知識獲得のもう一つの段階である。第 2 章で述べた前注意過程での感覚統合、すなわち選択されるべき特徴としていかにして生成するかという問題や、第 5 章で触れた基底関数生成の問題は、この段階に対応するものである。

表現のベースとなる特徴量を形成するための基本的方針として、ボトムアップ、トップダウンの二つの考え方が挙げられる。ボトムアップの考え方は、信号源に含まれている情報をなるべく

忠実に表現した情報表現機構を形成するという考え方であり、一方、トップダウンの考え方は、識別などのタスクを実現する上で必要な特徴を生成するという考え方である。

これらの考え方には、それぞれ長所と短所がある。前者は、信号に含まれる性質をもとにして下から組み上げていく形で表現を発展させることができるので、表現機構を確実に拡張していくことができるが、その反面、上位の特徴は下位の特徴の組合せの数だけ存在し得るため、層の数が増えるに従って特徴量の種類が爆発的に増加し、有用な特徴量を効率的に生成できないという問題がある。したがって、複雑な特徴量がこのようなメカニズムだけで生成されることは期待できない。一方、後者は、与えられたタスクの実現に必要な特徴だけを生成するので、うまく機能すれば効率的な表現機構を実現することができる。しかし、その実現過程は、結果から原因を探る逆問題を解くことに相当するので、その関係が複雑な場合には、そのような特徴量を確実に生成できる保証はない。

神経回路モデルにおける学習を例にとれば、前者はHebb学習[8]を用いて多層回路網を組み上げていく場合、後者は多層回路網において逆伝搬学習[85]を行なう場合に相当する。Hebb学習だけを用いては有用な特徴量がなかなか形成されないこと、逆伝搬学習では実質的に複雑な問題を学習できないことが、それぞれの限界をよく示しているといえる。

単純なアイデアとしては、これらの長所を組み合わせ、**「ボトムアップに生成しトップダウンに削除する」**方法が考えられる。これは、ボトムアップな考え方に基いて特徴表現機構を下側から積み上げていながら、その過程で不要な表現機構をトップダウンの評価に従って削っていくという考え方である。このような方法を用いることによって、システムは表現機構が爆発的に拡大するのを抑えつつ、複雑な情報表現を段階的に構築していくことができる。

内部表現獲得について議論する際に忘れてはならないもう一つの視点は、学習内容と内部表現の整合性である。

人間が物事を身につけるときには、一般に、学習を続けてもしばらくの間表面的には能力が向上せず、ある時期を境に階段状に向上するという現象が見られる。また、学習が効率的に進むのは学習の内容と学習者の能力がうまく整合しているときであり、学習内容に応じて適切な予備知識を身につけていないと、学習を効果的に行なうことができない。逆に、いったん基本的な知識や概念を身につけてしまえば、それを土台にしてさらに高度な概念を獲得することができる。同様のことは、運動能力の獲得過程についても成立する。

このような事実を踏まえると、内部表現獲得のメカニズムとしては、段階的な構造を順次構築しながら、単純なものから複雑なものを次第に獲得するようなものを考えることが妥当である[97]。このメカニズムの下では、新しい概念が身につくまでに必要な時間は、問題を表現するのに適切な内部表現を探るための時間であると考えられることができる。このとき、その時点で人間がすでに有している表現能力に対して与えた問題が難しすぎると、有用な内部表現を見つけないことができず、逆に、すでに自分のもっている概念を組み合わせることによって表現することのできる程度の問題であれば、それを単純な探索で見つけ出すことができる。このように考えれば、新たに獲得する表現とすでに獲得している表現との整合性が重要であることが説明できる。

適切な内部表現を選択する研究としては、統計学の分野において「モデル選択」という名の下で研究が行われており、モデルの良さを表すものさしとしてAIC[117]やMDL[84]などの情報量基準が知られている。また、神経回路モデルにおいても、このような視点からネットワークの良さを判断する研究が行われている[76]。内部表現の獲得について具体的に議論するためには、このような理論的研究を参考にすることも重要であると思われる。

最後に、学習と表現の線形性の関係について一つの考えを述べて、考察を終わりにする。

本論文では、すでに形成された表現のベースの上で認識や運動制御の問題を扱ったが、議論を進める上で問題の定式化を線形の範囲で行なった。これは、問題を単純化するという意味だけでなく、「非線形の部分を担うのは表現のベースであり、その上にある認識や適応の処理は線形の範囲にとどめる」という考え方に基づいている。線形性が成立する範囲においては逆問題を一意に解くことができるので、適切な表現の土台さえできあがっていれば、その上への適応の過程はスムーズに進行する。一方、表現の土台の部分は、一度獲得すれば、その上への適応機構が環境に適応しても半永久的に表現のベースとなり続ける。また、表現ベースが不完全な場合には、適応機構がそれを「あいまいさ」として吸収する一方、必要に応じて表現ベースを発達させることによって、対象をより正確に表現することができる。

このように、非線形変換を実現する表現ベースの部分と、線形処理に基づく認識や適応のメカニズムを分けることによって、それぞれの長所を生かした認識系や制御系を構成することができると思われる。このような考え方の有用性を示すためにも、今後、内部表現の獲得の問題についてより具体的に検討したいと考えている。

## 第7章

### 結論

本論文では、人間の感覚情報処理過程において本質的な役割を担っている感覚統合と能動的認識のメカニズムを情報処理的な視点から定式化し、その原理アルゴリズムを Bayes 推定、および情報量基準に基づく逐次実験計画として構成した。また、具体的な題材として、触知覚認識、図形認識、運動計画の三つの問題を取り上げ、それぞれについて実験的な検討を行なった。

本論文ではまず、これまで種々の過程が混同されて議論されていた感覚統合の過程を独自の視点から整理し、その上で本研究で取り扱う問題を明らかにした。続いて、能動的認識の過程をモデル化するために「内部像に基づく対象の理解」という基本的考え方を提案し、この考え方が人間の知覚過程を説明する上で有用であることを示した。さらに、この考え方に基づいて、感覚統合と能動的認識のアルゴリズムを一般的に定式化した後、上に挙げた三つの題材について具体的なシステムを構成した。

これらの問題について検討を進める上で、人間と同一ような振舞いを示す機械を工学的な理論に基づいて設計するという方法を用いた。これにより、感覚統合と能動的認識のメカニズムを、思想や哲学としてではなく工学的アルゴリズムとして具体的に書き下すことができた。また、三つの問題それぞれについて次のような成果が得られた。

まず、触知覚過程については、人間の触知覚系を忠実に模倣したシステムを実際に製作し、対象の材質や表面状態を識別するシステムを実現した。本システムにおいて能動的認識のアルゴリズムを用いた結果、状況に応じて適切な触運動を選択することによって、微妙な手ざわりの違いを識別できることが実験的に示された。また、システム内部の表現に基づく類似度が、人間の感覚の類似度とよく一致することが示された。このように、構成したシステムが人間とほぼ同じ挙動を示したことは、人間の触知覚メカニズムおよび触知覚情報の表現様式がシステムのもつそれらとよく一致していることを示唆しており、本システムが触知覚メカニズムのモデルとして優れたものであることを表している。

次に、図形認識過程については、人間が対象を眺めるときに、注視点や注意を向ける先を次々と移動させることに着目して、局所的な情報を逐次的に収集して図形の全体像を構成するシステムを構成した。このシステムにおいては、対象の内部像を構成する際に、感覚情報だけでなく対象に関する知識を利用することについて検討した。その結果、知識だけでは予測できない部分を

選択的に観測するという形で、能動的認識のアルゴリズムを導くことができた。構成したシステムの構造は、視覚系の構造と直接対応はつくものではないが、図形の線や角を視点が移動するという人間の視覚系の性質を、単に画像の微分に帰着させるのではなく、対象の内部像を構成するために有用な特徴を選択的に獲得するための機能として意味づけすることができた。

また、運動計画過程については、従来別個の問題として扱われてきた運動計画と学習制御を一つにまとめて考察することによって、未学習の状態でも目標の運動をほぼ間違いなく実行し、学習の進行とともにより効率的な運動を行なうシステムを構成した。運動制御の問題は、一見すると認識の問題と全く異なる問題であるように思えるが、「内部像に基づいて対象を捉える」という点で共通する構造を有している。このシステムの要点は、内部像に基づく運動予測の誤差を自分の内部で評価し、それに基づいて運動を計画することにあり、これは、内部像に応じて観測行為を定める能動的認識アルゴリズムと共通する考え方から生まれたものである。

ところで、本論文で述べた種々のアルゴリズムやシステムに共通する考え方は、卓近な言葉を使えば「思いこみ」という言葉で説明することができる。

能動的認識とは、「対象はこうであるに違いない」ということを示す内部像に基づいて、その仮定の下で適切な観測行為を選択するメカニズムである。このような思いこみがあるために、人間は、迷うことなく適切な観測行為を選択することができる。知識による内部像の構成は、実際に観測を行わずに知識だけを頼りに内部像を組み立てるという点で、まさに思いこみのものである。また、運動計画においては、実際に運動してみなければわからない運動結果とその誤差を頭の中で思いこむことによって、誤りが大きくならない指令を選択している。

ただ、このような思いこみのために、かえって間違いを犯してしまうことも確かである。しかし、人間は、間違いを犯しても、別の情報をもとにそれを修正する機能を備えている。そして、この間違いを修正する機能は、多種多様な情報を統合する感覚統合メカニズムによって実現されている。このように考えると、感覚統合と能動的認識は、人間が日常的な活動を行なう上で不可欠なメカニズムであることがあらためて感じられる。

さて、本研究の今後の発展として次の三つの課題が挙げられる。

第一に、第6章で述べたように、複数の物体が同時に提示されたときに、それらを分離して理解する仕組みを明らかにすることが挙げられる。このメカニズムは、音源分離に限らず画像認識においても機能する基本的なメカニズムであると考えられるが、その中で「注意」は本質的な役割を果たしていると思われる。今後、群化と注意の関係について考察を進め、そのアルゴリズムを明確に記述することが第一の課題である。

第二の方向として、本論文の議論を生理学的な知見と対照させることによって、感覚統合と能動的認識のメカニズムを神経回路レベルで明らかにすることである。本研究では、これらのメカニズムについてアルゴリズムレベルで議論を進めてきたが、脳機能の解明を進めるためには、最終的に神経回路レベルでそれを明らかにしなければならない。そのためには、生理学的知見や心理学的知見を取り込んで、生体のモデルとして妥当な形にしていく必要がある。

近年神経生理学の分野では、感覚刺激に対する単純な反応だけでなく、脳の高次機能に関わる実験が行なわれるようになっており、これまで心理学の世界でのみ論じられてきた「注意」につ

いても、実験結果が多数報告されるようになってきた。特に、注意の向きに応じて受容野が変化する細胞が視覚野に存在するという報告 [73] は、これまで実体を伴っていなかった「注意」を神経回路レベルで捉えた点で特記に値する。今後、このような実験事実を解釈しつつ、実験結果を予測するような精密なモデルを構成することが第二の課題である。

最後に、第6章で議論した内部表現の獲得を今後取り組むべき大きな課題として挙げておく。感覚統合と能動的認識のメカニズムそのものは、遺伝的に組み込まれた情報処理原理であるといえるが、それらは、外界の構造を頭の中に適切な形で表現することによってはじめて有効に機能する。したがって、内部表現獲得の問題は、認識の問題を考える上で避けて通ることのできない問題である。

人間は、そのような内部表現を生まれたときから長い時間をかけて構築していく。そして、新たな内部表現を獲得することに、その情報表現能力、情報処理能力を段階的に拡大させていく。このような内部表現獲得過程を解明するためには、脳の自己組織化、すなわち神経回路網の成長メカニズムを探ることと並行して、階層的表現機構の能力やその構築アルゴリズムについて情報処理的な視点から明らかにしていかなければならない。今後、「発達」の問題とも関連づけて取り組んでいきたい課題である。

いずれにせよ、脳は複雑かつ巨大なシステムであり、今後短い時間でそのすべてが解明されることはないであろう。しかし、この巨大な知能機械の設計図を完成させられないまでも、それを少しずつ書き進めていく作業を続けていきたいと考える。

## 謝辞

本論文は、筆者が東京大学工学部計数工学科在籍中に行なった研究の一部をまとめたものである。本研究を進めるにあたり、多くの方々から有形無形の援助を頂戴した。

まず、中野馨助教授には公私にわたり大変お世話になった。ここにそれらを一つ一つ記すことはできないが、特に、研究の枠組みの作り方、題材の選び方など研究者として仕事を進めていく上での考え方について、数知れないご教示をいただいた。鈴木良次教授には、第8講座の論議を通じてさまざまな指摘やご意見をいただき、また有益な資料をいただいた。甘利俊一教授には、神経回路網理論や逐次実験計画法に関してご教示いただいた。また、氏の学習理論や情報幾何学の考え方からは間接的に種々の示唆を受けた。吉澤修治教授には、旧南雲研究室の頃よりいろいろとお世話になった。特に、力学論や非線形理論、確率論について論議の場などを通じて教えていただいた。石川正俊助教授には、日常的に種々の議論の相手をしていただいたほか、センサフュージョンの研究会を通じてたくさんのご意見をいただいた。これらの研究会での議論は、感覚統合と能動的認識の研究を進める上で大きな刺激となった。また、以上の先生方には、論文審査の労をお取りいただいた。ここに感謝の意を表したい。

第8講座研究室の助手、大学院生、学部学生の皆さんからは、論議やその他の場を通じてさまざまなご意見やご協力をいただいた。博士課程在学中の下平英寿氏には、情報理論の立場からさまざまなご批判をいただいた。脳機能のモデル化、知能情報処理という境界領域の議論を行なう上で氏の合理的な考え方には大きな影響を受けた。同じく博士課程在学中の池田思朗氏には、日常的な議論を通じていろいろとご示唆をいただいた。また、松下技研(株)の金道敏樹氏には、物理学の視点から刺激的なご意見をいただいた。このほか、計数工学科事務室の皆さんをはじめとして、ここでは紹介することのできない多くの方々から援助をいただいた。合わせて謝意を表したい。

第3章で述べた触覚認識システムの実現においては、卒業研究を通じて黒川朋也氏、住友謙一氏のご協力をいただいた。この場を借りて御礼申し上げる次第である。

## 参考文献

- [1] 安部素嗣, 阪口豊, 中野馨: “楽器音の学習認識システム”, 第30回計測自動制御学会学術講演会予稿集, 605-606, 1991.
- [2] Abend, W., Bizzi, E. and Morasso, P.: “Human arm trajectory formation”, *Brain*, **105**, 331-348, 1982.
- [3] Akamatsu, M.: “The influence of combined visual and tactile information on finger and eye movements during shape tracing”, *ERGONOMICS*, **35**, 5/6, 647-660, 1992.
- [4] 赤松幹之: “形状のなぞりにおける視覚と触覚と運動との統合”, 電子情報通信学会春季大会予稿集, 6-291-292, 1992.
- [5] 赤松幹之, 石川正俊: “形状知覚における視-触覚の感覚統合過程の解析”, *バイオメカニズム*, **10**, 23-32, 1990.
- [6] Albus, J. S.: “A new approach to manipulator control: The cerebellar model articulation controller (CMAC)”, *Transactions of the ASME Journal of Dynamic Systems, Measurement, and Control*, **97**, 271-277, 1975.
- [7] 甘利俊一: 情報理論 II, 共立出版, 1968.
- [8] 甘利俊一: 神経回路網の数理, 産業図書, 1980.
- [9] 安西祐一郎: 知識と表象, 産業図書, 1986.
- [10] 安西祐一郎: 認識と学習, 岩波書店, 1989.
- [11] 有本卓: カルマン・フィルター, 産業図書, 1977.
- [12] 有本卓: 確率・情報・エントロピー, 森北出版, 1980.
- [13] Bregman, A.: *Auditory scene analysis — the perceptual organization of sound*, MIT Press, 1990.
- [14] Brown, G. J.: *Computational auditory scene analysis: A representational approach*, Ph. D. thesis, University of Sheffield, 1992.
- [15] Cave, K. R. and Wolfe, J. M.: “Modeling the role of parallel processing in visual search”, *Cognitive Psychology*, **22**, 225-271, 1990.

- [16] Chair, Z. and Varshney, P. K. : "Optimal data fusion in multiple sensor detection systems", *IEEE Transactions on Aerospace and Electronic Systems*, **AES-22**, 98-101, 1986.
- [17] Crossman, E. and Goodeve, P. : "Feedback control of hand-movement and fits' law", *Quarterly Journal of Experimental Psychology*, **35A**, 3, 251-278, 1983, Original work was presented at the meeting of the Experimental Psychology Society in 1963.
- [18] Fitts, P. M. : "The information capacity of the human motor system in controlling the amplitude of movement", *Journal of Experimental Psychology*, **47**, 6, 381-391, 1954.
- [19] Fitts, P. M. and Peterson, J. R. : "Information capacity of discrete motor responses", *Journal of Experimental Psychology*, **67**, 2, 103-112, 1964.
- [20] Flash, T. and Hogan, N. : "The coordination of arm movements: An experimentally confirmed mathematical model", *Journal of Neuroscience*, **5**, 1688-1703, 1985.
- [21] Fukushima, K. : "Neural network model of selective attention in visual pattern recognition", *Applied Optics*, **26**, 4985-4992, 1987.
- [22] Gibson, J. J. : "Observation on active touch", *Psychological Review*, **69**, 477-490, 1962.
- [23] Gibson, J. J. : *The senses considered as perceptual systems*, Houghton Mifflin, 1966.
- [24] Gibson, J. J. (古崎敬, 古崎愛子, 辻敬一郎, 村瀬曼共訳) : 生態学的視覚論 — ヒトの知覚世界を探る, サイエンス社, 1985.
- [25] 五味裕章, 川人光男 : "フィードバック誤差学習による閉ループシステム学習制御", システム制御情報学会論文誌, **4**, 37-47, 1991.
- [26] Harmon, L. D. : "Automated tactile sensing", *International Journal of Robotic Research*, **1**, 2, 3-32, 1982.
- [27] 波多野完治 (編) : ビアジェの発達心理学, 国土社, 1965.
- [28] 樋口正浩, 旭敏之, 笠井健 : "手の視覚情報が手動作の位置制御に及ぼす影響", 第1回生体・生理工学シンポジウム論文集, 65-68, 1986.
- [29] Hillis, W. D. : "A high-resolution image touch sensor", *International Journal of Robotic Research*, **1**, 2, 33-44, 1982.
- [30] Hoballah, I. Y. and Varshney, P. K. : "Distributed Bayesian signal detection", *IEEE Transactions on Information Theory*, **IT-35**, 995-1000, 1989.
- [31] 伊福部達 : "材質感の認識センサ", センサ技術, **4**, 1, 31-33, 1984.
- [32] 井口征士 : "音楽情報の処理 — 電算機を用いた自動探譜", 計測と制御, **19**, 3, 314-319, 1980.

- [33] 井口征士：“探舘と音楽情報処理”，コンピュータと音楽，共立出版，1987。
- [34] Inui, T., Kawato, M. and Suzuki, R.: “The mechanism of mental scanning in foveal vision”, *Biological Cybernetics*, **30**, 147-155, 1978.
- [35] 石川正俊：“マトリクス状センサからの出力分布の中心の位置と総和の検出方法”，計測自動制御学会，**19**, 5, 381-386, 1983.
- [36] 石川正俊：“センサフュージョンシステム — 感覚情報の統合アルゴリズム —”，日本ロボット学会誌，**6**, 3, 251-255, 1988.
- [37] 石川正俊：“センサフュージョンの課題”，日本ロボット学会誌，**8**, 6, 735-742, 1990.
- [38] 石川正俊，下条誠：“感圧導電性ゴムを用いた2次元分布荷重の中心の位置の検出方法”，計測自動制御学会論文集，**18**, 7, 730-735, 1982.
- [39] 石川正俊，下条誠：“ビデオ信号出力を持つ圧力分布センサと触覚パターン処理”，計測自動制御学会論文集，**24**, 7, 662-669, 1988.
- [40] 伊藤宏司，伊藤正美：生体とロボットにおける運動制御，計測自動制御学会，1991.
- [41] 岩村吉晃：“触覚受容器の構造と特性”，日本ロボット学会誌，**2**, 5, 54-60, 1984.
- [42] 岩村吉晃：“つまむ — 指の運動と触覚”，*Japanese Journal of Sports Sciences*, **6**, 10, 612-616, 1987.
- [43] Jordan, M. I.: “Motor learning and the degrees of freedom problem”, In Jeannerod, M., editor, *Attention and Performance VIII*, 796-836, Lawrence Erlbaum Associates, 1990.
- [44] 金出武雄：“米国における自律移動ロボットの研究の動向”，日本ロボット学会誌，**5**, 5, 375-383, 1987.
- [45] 柏野邦夫：音源分離に関する研究，東京大学大学院工学系研究科修士論文，1992.
- [46] 柏野邦夫，田中英彦：“モノラル楽器音の音源分離のための知覚の手がかりの検討と処理モデルの実装”，日本音響学会聴覚研究会資料，H-93-84, 1993.
- [47] 柏野邦夫：“音の流れを聞きとる — 聴覚による環境の解析”，科学，**62**, 374-379, 1992.
- [48] 片山徹：応用カルマンフィルタ，朝倉書店，1983.
- [49] 片寄晴弘：“知的探舘システム”，人工知能学会誌，**5**, 1, 59-66, 1990.
- [50] 川端季雄：風合い評価の標準化と解析，日本繊維機械学会，1980.
- [51] 川津茂生：“探索非対称性”，数理科学，**349**, 71-75, 1992.
- [52] 川嶋稔夫，青木由直：“圧力分布投影型触覚センサによる局所特徴の認識”，電子情報通信学会論文誌，**J72-D-II**, 4, 526-534, 1989.
- [53] Kawato, M.: “Optimization and learning in neural networks for formation and control of coordinated movement”, In Meyer, D. and Kornblum, S., editors, *Attention and Performance XIV*, 821-849, MIT Press, 1993.

- [54] Kawato, M., Furukawa, F. and Suzuki, R.: "A hierarchical neural-network model for control and learning of voluntary movement", *Biological Cybernetics*, **57**, 169-185, 1987.
- [55] Kawato, M., Maeda, Y., Uno, Y. and Suzuki, R.: "Trajectory formation of arm movement by cascade neural network model based on minimum torque-change criterion", *Biological Cybernetics*, **62**, 275-288, 1990.
- [56] 木下源一郎: "触覚センサの開発現状", 日本ロボット学会誌, **2**, 5, 46-53, 1984.
- [57] Kohonen, T.: "Self-organized formation of topographically correct feature maps", *Biological Cybernetics*, **43**, 56-69, 1982.
- [58] Kowalski, C.: "Silicon sensors for tactile arrays and distributed touch sensing", SME Technical Papers, MS84-1040, 1984.
- [59] 熊田孝恒: "ポップアウト", 数理科学, 345, 78-82, 1992.
- [60] Kuperstein, M.: "Neural model of adaptive hand-eye coordination for single postures", *Science*, **239**, 1308-1311, 1988.
- [61] 黒川朋也: 能動的触覚センシングの研究, 東京大学工学部卒業論文, 1992.
- [62] 黒川朋也, 阪口豊, 中野馨: "意図的触覚センシングシステム", 第31回計測自動制御学会学術講演会予稿集, 671-672, 1992.
- [63] Lederman, S. J. and Klatzky, R.: "Hand movements: A window into haptic object recognition", *Cognitive Psychology*, **19**, 342-368, 1987.
- [64] Luo, R. C. and Kay, M. G.: "Multisensor integration and fusion in intelligent systems", *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-19**, 901-931, 1990.
- [65] 前田芳晴, 川人光男, 宇野洋二, 鈴木良次: "カスケード構造の神経回路モデルによるヒト上肢多関節運動軌道の生成", 電子情報通信学会技術報告, NC88-169, 1989.
- [66] Marr, D. (乾敏郎, 安藤広志共訳): ビジョン - 視覚の計算理論と脳内表現 -, 産業図書, 1987.
- [67] Merleau-Ponty, M. (竹内芳郎, 小木貞孝共訳): 知覚の現象学, みすず書房, 1974.
- [68] Meyer, D., Abrams, R., Kornblum, S., Wright, C. and Smith, J.: "Optimality in human motor performance: Ideal control of rapid aimed movements", *Psychological Review*, **95**, 3, 340-370, 1988.
- [69] Miller, W. T.: "Sensor-based control of robotic manipulators using a general learning algorithm", *IEEE Journal of Robotics and Automation*, **RA-3**, 157-165, 1987.
- [70] Minsky, M. and Papert, S. (中野馨, 阪口豊共訳): パーセプトロン, パーソナルメディアア, 1993.

- [71] 光山訓, 西井淳, 宇野洋二, 鈴木良次: “筋スティフネス調節機構を持った上腕の運動制御モデル”, 電子情報通信学会技術報告, NC92-122, 1993.
- [72] 三浦利章: “外界情報の獲得・処理様式”, 数理科学, 354, 53-58, 1992.
- [73] Moran, J. and Desimone, R.: “Selective attention gates visual processing in the extrastriate cortex”, *Science*, **229**, 782-784, 1985.
- [74] 森本直樹, 篠田裕之, 安藤繁: “局所応力テンソル検出に基づく触知覚アルゴリズム—基礎的解析および接触状態分類への適用—”, 第33回計測自動制御学会学術講演会予稿集, 1313-1316, 1993.
- [75] Mott, D. H., Lee, M. H. and Nicholls, H. R.: “An experimental very high resolution tactile sensor array”, In *Proceedings of 4th International Conference on Robot Vision and Sensory Control*, 241-250, 1984.
- [76] 村田昇: 学習の統計的漸近理論, 東京大学大学院工学系研究科博士論文, 1992.
- [77] 中森一博, 柏野邦夫, 田中英彦: “音楽音響信号を対象とする音源分離システム—音モデルに基づくアプローチ”, 情報処理学会研究報告, SIGMUS 1-1, 1993.
- [78] 中村雄次郎: 共通感覚 (中村雄次郎著作集第V巻), 岩波書店, 1993.
- [79] 中野馨, 阪口豊, 森田昌彦, 新妻素直, 脇本良則, 田中宏一良, 藤田啓介, 衛作人: ニューロコンピュータの基礎, コロナ社, 1990.
- [80] Neisser, U. (古崎敬, 村瀬夏共訳): 認知の構図, サイエンス社, 1978.
- [81] Neisser, U. (大羽藁訳): 認知心理学, 誠信書房, 1981.
- [82] 岡畑恵雄: “嗅覚—水晶発振子をデバイスとする匂いセンサ”, 数理科学, 356, 38-42, 1993.
- [83] 尾股定夫, 尾崎伸自: “硬さ知覚用触覚センサとその応用”, 第7回日本ロボット学会学術講演会予稿集, 573-574, 1989.
- [84] Rissanen, J.: “Universal coding, information prediction, and estimation”, *IEEE Transactions on Information Theory*, **IT-30**, 4, 629-636, 1984.
- [85] Rumelhart, D. E., McClelland, J. L. and PDP Group: *Parallel distributed processing*, MIT Press, 1986.
- [86] Ruoff, C., et al.: “Designing a mars surface rover”, *Aerospace American*, **23**, 11, 54-56, 1985.
- [87] Sadjadi, F. A.: “Hypotheses testing in a distributed environment”, *IEEE Transactions on Aerospace and Electronic Systems*, **AES-22**, 134-137, 1986.
- [88] 佐伯祥, 佐々木正人 (編): アクティブ・マインド—人間は動きの中で考える, 東京大学出版会, 1990.

- [89] 阪口豊：神経回路網における情報変換機構の自己形成，東京大学大学院工学系研究科修士論文，1988。
- [90] 阪口豊：“触知覚における感覚統合と能動的認識”，電子情報通信学会誌，76，11，1222-1227，1993。
- [91] Sakaguchi, Y.：“Haptic sensing system with active perception”，*Advanced Robotics*, in printing.
- [92] 阪口豊：“動きの予測を伴う能動的認識のアルゴリズム”，日本ロボット学会誌，in printing.
- [93] 阪口豊，中野馨：“物理的な制約にもとづく空間の内部表現”，第5回生体・生理工学シンポジウム論文集，157-160，1990。
- [94] 阪口豊，中野馨：“能動的認識の数理モデル”，第6回生体・生理工学シンポジウム論文集，373-376，1991。
- [95] Sakaguchi, Y. and Nakano, K.：“Active perception with intentional observation”，In *Proceedings of ISMCR-92*, 241-248，1992。
- [96] Sakaguchi, Y. and Nakano, K.：“Figure recognition with intentional observation”，In *Proceedings of IJCNN-92-Baltimore*, III, 792-797，1992。
- [97] 阪口豊，中野馨：“ボトムアップ学習による階層ネットワークの自己組織化”，第31回計測自動制御学会学術講演会予稿集，745-746，1992。
- [98] 阪口豊，中野馨：“外界からの制約にもとづく図形の内部表現と認識”，計測自動制御学会論文集，28，3，392-400，1992。
- [99] Sakaguchi, Y. and Nakano, K.：“Motor planning according to reliability of internal model”，In *Proceedings of IJCNN-93-Nagoya*, II, 1321-1324，1993。
- [100] 阪口豊，中野馨：“内部モデルの信頼度を考慮した学習運動系のモデル”，電子情報通信学会技術報告，NC92-121，1993。
- [101] 坂根茂之：“ロボット用能動的視覚センシングシステムの設定計画に関する研究”，電子技術総合研究所研究報告，903，1989。
- [102] 先田力哉，篠田裕之，安藤繁：“3次元構造を有する触覚センサ(II)”，第34回自動制御連合講演会予稿集，529-530，1991。
- [103] 佐々木正人：からだ—認識の原点，認知科学選書，15，東京大学出版会，1987。
- [104] Schmidt, R. F.：“Somatovisceral sensibility”，In Schmidt, R. F., editor, *Fundamentals of sensory physiology*, 81-125，Springer Verlag，1981。
- [105] 志村正道：機械知能論，昭見堂，1983。
- [106] Shinoda, H. and Ando, S.：“A tactile sensing algorithm based on elastic transfer function of surface deformation”，In *Proceedings of IEEE ICASSP-92*, 3, 589-592，1992。

- [107] Shinoda, H., Uehara, M. and Ando, S.: "A tactile sensor using three-dimensional structure", In *Proceedings of 1993 IEEE International Conference of Robotics and Automation*, 435-441, 1993.
- [108] Shinoda, H., Uehara, M., Okada, M. and Ando, S.: "A fingertip tactile sensor with three dimensional structure", In *Proceedings of Transducers '93*, 614-617, 1993.
- [109] 篠田裕之, 安藤繁: "3次元構造をもつ触覚センサ—変形パターンへのヒルベルト変換の知覚とその利用—", 電気学会センサ技術研究会資料, 69-78, 1991.
- [110] Sugiyama, S., Takigawa, M. and Igarashi, I.: "Integrated piezoresistive pressure sensor with both voltage and frequency output", *Sensors and Actuators*, **4**, 1, 113-120, 1983.
- [111] 住友謙一, 坂口豊, 中野繁: "触知覚感の定量化に関する研究", 第32回計測自動制御学会学術講演会予稿集, 33-34, 1993.
- [112] 住友謙一: 能動的触覚センシングシステムの研究, 東京大学工学部卒業論文, 1993.
- [113] 鈴木教夫: 応答曲線推定のための逐次実験計画法の研究, 東京大学大学院工学系研究科修士論文, 1983.
- [114] 高橋弘太: 音響センサの知能化と視聴覚融合, 東京大学大学院工学系研究科博士論文, 1993.
- [115] 高澤薫光: "計算機による探譜", 情報処理, **29**, 6, 593-598, 1988.
- [116] 竹田青嗣: "現象学から実存主義へ", 現代思想入門, JICC 出版局, 1984.
- [117] 竹内啓: "AIC 基準による統計的モデルの選択をめぐって", 計測と制御, **22**, 5, 445-453, 1983.
- [118] 滝山竜三: "距離の概念の導入によるパタン認識系の評価法", 九大工学集報, **37**, 318-324, 1965.
- [119] Tenney, R. A. and Sandell, Jr., N. R.: "Detection with distributed sensors", *IEEE Transactions on Aerospace and Electronic Systems*, **AES-17**, 4, 501-510, 1981.
- [120] 都甲潔: "味覚—人工脂質膜", 数理科学, 356, 25-30, 1993.
- [121] 都甲潔, 山藤馨: 美味しさを測る—舌を超えた驚異の味センサ, 講談社, 1991.
- [122] 鳥居修晃: 視覚の心理学, サイエンス社, 1982.
- [123] Treisman, A.: "Preattentive processing in vision", *Computer Vision, Graphics, and Image Processing*, **31**, 156-177, 1985.
- [124] Treisman, A.: "Features and objects in visual processing", *Scientific American*, **254**, 114-125, 1986.
- [125] Treisman, A.: "Spreading suppression of feature integration? A reply to Duncan and Humphreys (1992)", *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 2, 589-593, 1992.

- [126] Treisman, A. M. : "Search similarity, and integration of features between and within dimensions", *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 3, 652-676, 1991.
- [127] Treisman, A. M. and Gormican, S. : "Feature analysis in early vision: Evidence from search asymmetries", *Psychological Review*, **95**, 1, 15-48, 1988.
- [128] Treisman, A. and Gelade, A. : "A feature integration theory of attention", *Cognitive Psychology*, **12**, 97-136, 1980.
- [129] Treisman, A. and Sato, S. : "Conjunction search revisited", *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 3, 459-478, 1990.
- [130] Treisman, A. and Schmidt, H. : "Illusory conjunctions in the perception of objects", *Cognitive Psychology*, 107-141, 1982.
- [131] Treisman, A. and Souther, J. : "Search asymmetry: A diagnostic for preattentive processing of separable features", *Journal of Experimental Psychology: General*, **114**, 285-310, 1985.
- [132] Uno, Y., Kawato, M. and Suzuki, R. : "Formation and control of optimal trajectory in human multijoint arm movement —minimum torque-change model—", *Biological Cybernetics*, **61**, 89-101, 1989.
- [133] Viswanathan, R., Thomopoulos, S. C. A. and Tumhuri, R. : "Optimal serial distributed decision fusion", *IEEE Transactions on Aerospace and Electronic Systems*, **AES-24**, 366-376, 1988.
- [134] Wiener, N. (池原止戈夫, 彌永昌吉, 室賀三郎, 戸田巖共訳) : サイバネティックス (第2版), 岩波書店, 1962.
- [135] Wolfe, J. M., Cave, K. R. and Franzel, S. L. : "Guided search: An alternative to the feature integration model for visual search", *Journal of Experimental Psychology*, **15**, 3, 419-433, 1989.
- [136] 山崎弘郎, 石川正俊 (編) : センサフュージョン, コロナ社, 1992.
- [137] Yarbus, A. : *Eye movements and vision*, Plenum Press, 1967.
- [138] 横澤一彦 : "一目でわかること", *科学*, **62**, 6, 356-362, 1992.
- [139] 横澤一彦 : "視覚探索とは", *数理科学*, 344, 66-69, 1992.
- [140] 趙南元, 飯島泰蔵 : "安定視点木法による画像の特徴抽出理論", *電子通信学会論文誌*, **J68-D**, 5, 1125-1132, 1985.
- [141] 趙南元, 飯島泰蔵 : "図形観測における視点, 視野の決定法に関する理論", *電子通信学会論文誌*, **J68-D**, 4, 508-514, 1985.

