

01-R 014

新情報処理技術に関する総合的調査研究

認識・理解WG報告書

平成 2 年 3 月

JIPDEC

財団法人 日本情報処理開発協会



この報告書は、日本自転車振興会から競輪収益の一部である
機械工業振興資金の補助を受けて平成元年度に実施した「新情
報処理技術に関する総合的調査研究」の成果をとりまとめたも
のであります。





序

過去30余年にわたる情報処理技術の進歩は著しく、その成果を支えに社会の情報化も急速に進展し、コンピュータは今やわれわれの社会活動にとって不可欠のものとなった。しかしながら、その利用範囲の拡大や多様化にともない、情報処理機能の一層の高度化が求められ、特に最近では、固定化した逐次処理型の情報処理機能から、人間にとって親和性に富む、より柔軟性の高い知的な情報処理機能への変革が、強く期待されるようになった。

このため、従来のコンピュータが不得手とする、人間の脳が行っているような、高度かつ多彩な情報処理の実現に向けて、必要とされる機能、計算原理および実装技術、さらには、革新的な情報処理技術と社会の係わり等を含めた総合的な調査研究を、平成元年度から開始した。

本年度は、通商産業省における新情報処理技術調査研究委員会の下部機構として、当協会に基礎技術、新機能、社会応用の3分科会および超並列・超分散処理、学習、光技術・新デバイス、3次元情報処理、認識・理解、自律・協調の6ワーキンググループを設置して調査研究を実施するとともに、大学・研究所への研究委託、海外調査等により、21世紀を目指す革新的な情報処理技術がいかにあるべきかを、ニーズ、シーズの両面から追求し、新しいナショナルプロジェクトとしての可能性を検討した。

最後に、本調査研究に多大なご協力を頂いた各分科会およびワーキンググループの委員各位に厚く御礼申し上げる次第である。

平成2年3月

「認識・理解WG」委員名簿

(順不同)

	氏 名	役 職
主査	杉江 昇	名古屋大学工学部電気工学科教授
幹事	山本 和彦	工業技術院電子技術総合研究所 知能情報部画像研究室長
幹事	河原 英紀	日本電信電話(株) NTT基礎研究所 情報科学研究部主幹研究員
委員	小林 啓二	三菱電機(株) 情報電子研究所知識処理開 発部知的インタフェース 研究所グループマネージャー
	斎藤 秀昭	玉川大学工学部 情報通信工学科教授
	下沢 楯夫	北海道大学応用電気研究所 感覚情報工学部門教授
	白井 克彦	早稲田大学理工学部 電気工学科教授
	杉原 厚吉	東京大学工学部計数工学科助教授
	杉本 軍司	(株)豊田中央研究所研究7部 機械認識研究室長
	田中 厚夫	シャープ(株) 情報システム 研究所第一開発部主任研究員
	中野 康明	信州大学工学部情報工学科教授
	松島 整	(株)日立製作所 中央研究所 第6部主任研究員

目 次

1. はじめに	1
1.1 調査の背景	1
1.2 調査研究項目	2
2. 調査検討経過	3
3. 検討分野の現状と概念軸・パラダイム軸の検討	5
3.1 主要概念の抽出	5
3.1.1 パラダイム軸	5
3.1.2 記述レベル軸	5
3.2 全体イメージ	6
3.2.1 認識・理解の問題の性質	6
3.2.2 認識・理解における新機能	7
4. 分野の現状と展望	8
4.1 生理・心理	8
4.1.1 時空間処理総論	8
4.1.2 感覚統合	9
4.1.3 空間認知の神経メカニズム	13
4.2 感覚・知覚（視覚、聴覚、その他）	28
4.2.1 最適化問題の流れ（正則化、弛緩法）	28
4.2.2 NP問題と能動的解釈	30
4.2.3 能動性について	31
4.2.4 新しい音声情報処理	35
4.3 感性・文化・応用	37
4.3.1 感性情報処理	37
4.3.2 センサ融合	38
4.3.3 ヒューマンインタフェース	41
4.3.4 コンピュータビジョン	45
4.3.5 ロボットビジョン（移動ロボット）	47
4.3.6 OA（マルチメディア、ハイパーメディア）	52
5. 新情報として取り上げるべき研究課題の提案	56
5.1 プロジェクトイメージ	56
5.2 研究課題の提案（状況理解における新機能）	58
5.2.1 状況パターンの能動的総合理解	58
5.2.2 不良設定問題・組合せ論的問題の新解法	58
5.2.3 情報表現の階層化・モジュール化と異種モダリティの統合	59
5.2.4 動機付け機構のモデル化と認識・行動のモデル化	61
5.2.5 認識理解機構の自己組織的学習	62
5.2.6 感性情報処理	64
6. 研究成果の展開	66
6.1 利用者の状況を把握して応答を変える学習型の情報システム	66
6.2 高度な映像情報の編集・加工システム	67
6.3 教育・娯楽機器	69
6.4 人工的に様々な体験を現実感を伴って提示しセンシングすることのできる 超環境実現装置	70
6.5 障害者・高齢者等の低下した機能の補綴システム	71
6.6 心やさしきコンパニオン	72

参考資料

関連するキーワード表

1. は じ め に

1.1 調査の背景

我々が日常の活動の中で行っている知的な活動には、大別すると、

- (1) 言語で記述できる処理形式と、
- (2) 言語化される以前の状態でやっている処理形式

とがある。これらは、それぞれ知性の論理的側面と直観的側面を代表しているといえよう。

(1)のタイプの処理形式の具体例としては、数値計算や論理的推論等がある。このタイプの処理は、数理的、論理的な扱いに馴染み易く、技術的にも実現が比較的容易であったことから、ディジタル計算機が出現してからの40年程の短い期間に飛躍的な発展が達成された。

(2)のタイプの処理の例としては、画像、音声、自然言語の認識・状況理解等、現実の人間の知的処理の大部分を占める活動がある。人間の知性の論理的側面も、生物進化の長い歴史を通じて洗練されてきた直観的な処理の集積に支えられて初めて出現し得たものと考えられる。知性の直観的な側面は、機能的には柔軟さと重層性、機構的には超多重並列処理で特徴づけられる。これらは、人間にとっては非常に容易であるにも関わらず、いざ機械的に実現しようとするとは定式化が難しく、簡単な機能であっても膨大な計算量を必要とするなどの問題があり、進歩は遅々としていた。

しかし、厳密な定式化と逐次的論理的な処理に代表される情報処理技術の偏った発展は、現実的な問題への対応、人間とのインタフェースの機会の増大に伴い、その硬直的な側面が持つ様々な問題を顕在化させている。また、高度化する21世紀の社会を支えるために、今後ますます多数の知能を有する機械が我々の生活している環境の中に入ってくることが予想される。このような機械を我々の生活している環境の中に取り入れて我々と共存させていくためには、従来の情報処理技術に欠けていた柔軟かで重層的な処理機能を実現することが必須である。

最近のデバイス技術の進歩は、従来では考えられなかった程の多くの素子を集積したシステムの実現を可能とし、計算機などの情報処理能力の飛躍的向上も、従来では困難であった複雑性、規模の問題の検討を可能にしている。また、研究手法の高度化、理論的解析

の進歩に支えられて、生物の有する機能やそれを実現する機構の解明が大きく進展している。このような状況を考慮すると、認識・理解の諸機能に代表される柔らかで重層的な情報処理機能の本格的な検討を推進する条件は整ってきたと言えよう。

1. 2 調査研究項目

このような状況を背景として、生体の有するパターン認識、状況理解の仕組みの把握とそれを工学的に実現する方法についての調査検討を行った。具体的には、認識や理解の様々な機能とそれを実現する構造に関する生理・心理学的知見の調査を行うとともに、時間的空間的に広がった外界の状況を限られた手がかりから迅速に把握し適切な応答を形成する新しい機能を実現する方法について、工学的な観点からの検討を進めた。

2 . 調 査 検 討 経 過

調査検討に際しては、生体をとるべく状況の認識・理解を通じた適応行動に本質的な概念の抽出およびそれらに必要な新機能の抽出のために、調査およびヒアリングを平行して行った。また、これらの結果に基づいた集中討論により、新情報処理機能として取り上げべき研究課題とプロジェクトについてのイメージの練り上げを行った。

以下、検討の経過と、主要な検討結果について説明する。

第1回 平成元年9月21日

- ・WGの狙いについてのイメージ合わせ。自己紹介。
- ・WGでの検討方針と今後の進め方についての意見交換。

記事：

認識・理解に関するキーワード、ヒアリングを希望するトピック、WGの進め方に対する提案のアンケート調査を次回のWGまでに実施。

第2回 平成元年10月9日

- ・ヒアリング 下沢委員 「昆虫のセンシングと行動」
- ・キーワードの整理
- ・ヒアリング項目の整理
- ・WGの進め方についての意見交換。

記事：

キーワードを2つの観点から整理し、表形式にまとめた。

合宿による集中的な議論の日程を決定。

第3回 平成元年10月30日

- ・ヒアリング 松島委員 「人工的ニューラルネットのエンジニアリング応用研究の事例
・・・日立における研究」
- ・認識・理解のキーワード表、検討方法についての意見交換。

記事：

パラダイム軸、記述レベル軸の組合せで表された各項目について各委員で分担して、

合宿における話題提供、プロジェクトの提案を行う。

第4～6回 平成元年11月18日～19日

- ・ヒアリング 斎藤委員 「パターンの運動情報の脳内表現」
- ・ヒアリング 白井委員 「新音声情報処理」
- ・プロポーザルの検討

テーマ：能動的理解、不良設定

テーマ：記憶・学習、記号処理とパターン処理の融合

テーマ：感覚統合、モジュール化・階層化

- ・全体の取りまとめについて。

第7回 平成元年12月18日

- ・新機能分科会第1回中間報告について他
- ・認識・理解WG検討結果のまとめ方、スケジュールについて

第8回 平成2年1月29日

- ・ヒアリング 理化学研究所 藤田氏 「フクロウにおける聴覚による空間認識」
- ・報告書の作成について

第9回 平成2年2月26日

- ・報告書（案）の検討

3. 検討分野の現状と概念軸・ パラダイム軸の検討

3.1 主要概念の抽出

学習・認識に関連するキーワードの調査結果を巻末の付録に示す。これらのキーワードの中に隠されている本質的な構造を見つけ出すため、幾つかの概念に基づく軸を設定し、整理を試みた。

一つの軸は、問題をとらえる枠組を表すものであり、「パラダイム軸」と呼ぶこととする。もう一つの軸は、問題をどのような情報媒体、レベルでとらえるかを表すものであり、「記述レベル軸」と呼ぶこととする。

これらの軸はキーワードに含まれる概念のすべてを網羅している訳ではない。例えば、キーワードに対応する機構の構造的なレベル（「システム軸」とでも呼ばれる。）という観点は欠落している。しかし、本WGでの検討は、新情報処理機能の抽出を主な狙いとしているため、今回の整理では、上で説明した2つの軸のみを用いることとする。

3.1.1 パラダイム軸

問題を捉える枠組としては、当初6つの観点が抽出されたが、報告書のまとめでは、以下の3つに大分類する。

- (1) 能動的理解、不良設定問題
- (2) 記憶・学習、記号処理とパターン処理の融合
- (3) 感覚統合、モジュール化

3.1.2 記述レベル軸

問題を情報媒体、検討のレベルの観点から整理し、以下の3項目に大分類する。

- (1) 生理・心理
- (2) 感覚・知覚（視覚、聴覚、その他）
- (3) 感性・文化・応用

3. 2 全体イメージ

3. 2. 1 認識・理解の問題の性質

認識・理解の諸機能の大部分は、言語化される以前の無意識的なレベルで行われている。これらの機能は、人間以外の生物にも存在しており、人間の知的処理も、この無意識的なレベルに支えられたものである。広い意味で認識・理解の処理をとらえれば、外界あるいは内部状態に関して得られる様々なデータを解釈して、適切な応答を生み出すために利用できるような形に変換することであるといえよう。このような処理は、次のような特徴を持っている。

- (1) 問題の全体を明確に定義することが難しい (ill-defined problem)。
- (2) 問題を分解して部分問題を定式化できたとしても、得られるデータだけでは情報が不足し、解が一意に決まらなかったり、また、データが誤りを含んでいるために、データと全く矛盾しない解が存在しなかったりする。
- (3) 解釈の可能性をすべて調べて厳密解を求めようとすると計算量の爆発を招く。
- (4) 解の精度のみならず解が求められるまでの時間が短いことが重要であり、両者のトレードオフを調整できることが必要である。

従って、問題の全体に対する解を逐次的な手続きや規則に基づいた論理的な処理により求めることは難しく、従来の情報処理には不向きである。さらに、逐次的なアルゴリズム化や論理的な規則の構成が困難であるという問題に加えて、情報そのものが不完全であるという問題を抱えている。すなわち、

- (5) 対象を完全に観測することは不可能であり、また何を対象とするかを完全に記述すること (対象のモデル化) 自体も一般には不可能である。
- (6) 得られた情報からの特徴抽出も不完全である。(例えば物理的条件だけでは、対象を分節化できない。)

といった問題がある。このため、従来は、ad hoc なプログラムを書くということになり、対象あるいは実現する機能を極端に制限せざるを得なかった。

このような現実の認識・理解の問題に対しては、「そもそも完全な情報が得られることはなく、問題の解空間を限定する“部分的な制約=不完全なデータと知識と仮定”しか得られない」という認識に立つ必要がある。これは、生物における認識・理解の諸機能の構成の基本デザインでもある。

3.2.2 認識・理解における新機能

このように、対象に対する知識・情報が不完全な場合の認識・理解の問題の定式化については、方法論が確立されておらず、現在の認識・理解システムは「注意深く限定された理想的な環境の下」でしか実用にはなっていない。そこで、今後の認識・理解研究においては、不完全な情報の表現方法（柔らかな情報表現）を確立するとともに、認識・理解の問題をとらえる新しい視点を確立することが重要となる。

生物はこれらの問題に対応するために、様々な戦略を用いている。生物の戦略を検討することにより、この視点がどのようなものであるかについてのヒントが得られる。詳しくは次章で紹介されるが、以下のような特徴が認められる。

- (1) 能動性、志向性、
- (2) 適切な内部表現、
- (3) 異種感覚の統合、
- (4) 学習、記憶

まず、感覚細胞レベルから認知過程のレベルに至るまでの多くのレベルにおいて、意味のある情報を獲得するための能動的機能が存在している。これらは、低いレベルでは、ポジティブフィードバックによる信号選択機構であったり、高いレベルでは、知識に基づく次入力の予測機構であったりする。また、適切な内部表現を獲得するため、特殊な感覚器官を発達させているものがある。これらの様々な感覚は統合されて、対象の捕食や、対象からの逃亡の反応を引き起こす。さらに、これらのいわば「組み込まれた偏見」と現実とのずれを補うものとして学習・記憶の機能が存在している。このように生物の戦略は、重層的であり、高度に階層的並列的である。

生物に見られる複雑に絡み合った重層的な諸機能すべてを統一的に説明できる原理を見いだすことは困難であるが、そのうちの多くは、「制約充足問題」として整理することができる。すなわち、情報の不完全性を許容する形で問題を厳密に定義し、制約充足問題を解くことによって（準）最適解を求めるというアプローチである。このアプローチは、多数の要素的な知的機能の相互作用を通じて新しい情報処理を実現しようとする「群情報処理」の本質に呼応している。

4. 分野の現状と展望

4.1 生理・心理

4.1.1 時空間処理総論

生体が時々刻々変化する外界で生きていくために不可欠な処理、それが時空間処理である。

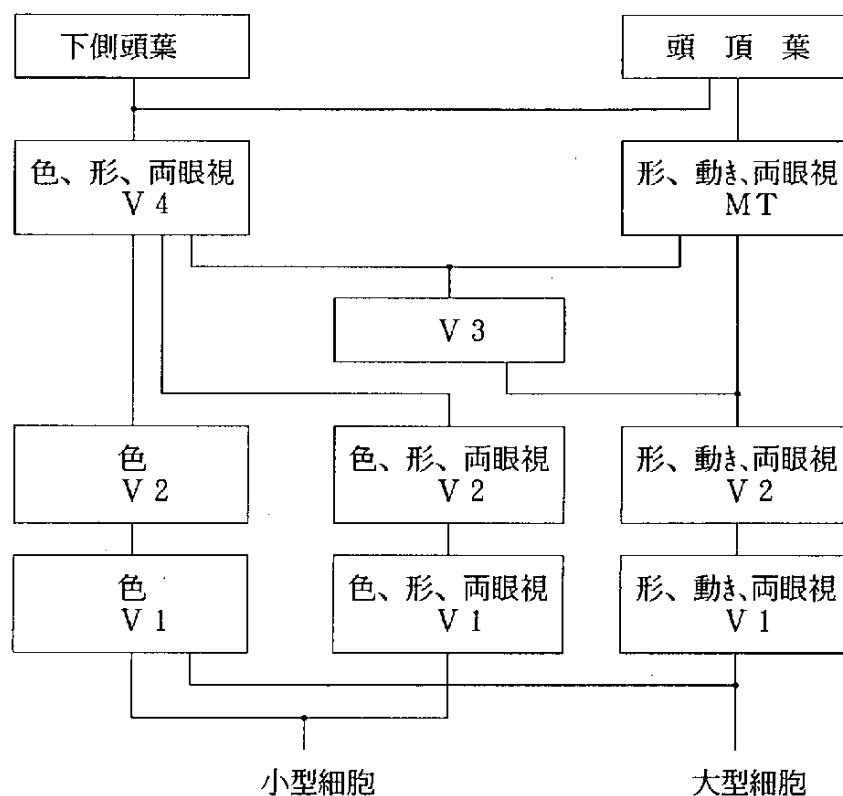
視覚情報処理の場合、2次元の網膜像から外界に存在する物体の3次元形状、3次元運動、表面反射率を、照明条件や視線方向の変化を克服して、安定に知覚する仕組みが存在する。これは恒常性という言葉でくることができる。鳥、猫などのように、同一個体でありながら、さまざまな3次元形状をとる物体が、常に安定して、鳥や猫として知覚される仕組みも恒常性の延長線上にあると言える。

これらの仕組みについては、心理学レベルでは、両眼立体視、運動立体視、陰影からの立体視、色の恒常性などの名前でかなり研究が行われている。

神経生理レベルでは、ネコやサルなどの高等動物を対象にして、両眼立体視、運動立体視、色の恒常性に関連したニューロンが発見されているが、機構全体を説明できるところまではいっていない。興味深いことに、大脳皮質には図表4.1-1に示すように視覚に関連したモジュールが十数個あり、各モジュールは両眼視、運動視、色覚などに専門化していることが分かってきた。これらモジュールは、並列的かつ階層的に相互接続されている。これは各機能モジュールの結合により、効率的かつ高信頼性での処理を実現しているものではないかと思われる。

以上の諸機能のモデルは、並列処理ネットワークのレベル、計算理論のレベルで、両眼立体視や運動立体視を中心に1975年頃から幾つか提案されている。これらの機能は生得的に存在するわけではなく、生後次第に発達するものであることが知られている。このような自己組織化ニューラルネットワークモデルが、両眼立体視において提案されている。

聴覚系は、普通、時間情報処理系として音声認識についての研究が中心課題となっているが、最近フクロウやコウモリの空間定位機構の神経生理学的研究により興味深い成果が得られている。これと心理学的知見とを組合せて複数存在する音源の方位を数度以内の誤差で決定する方式が開発された。方位が決定できると、複数音源から同一期間中に到来する混合音を、各音源に分離して出力できることも実証されている。



図表4.1-1 視覚系の構成 (Van Essen 1983)

さらに、視覚と聴覚、視覚と触覚の情報を統合しているニューロンも報告されていて、これは感覚統合に関与している知見と思われる。

4.1.2 感覚統合

我々は、自分の棲んでいる物理的世界を正しく認識していると思っている。しかし必ずしもそうではないということは、エッシャーの騙し絵をはじめとする様々な錯視図形があることでも分かる。物理学者Ernst Machは1865年、ヒトの視覚ではコントラストの境界でその差が強調されて認識されること（マッハバンド）を見いだして、我々の認識が物理的世界と必ずしも一致しないことを警告し実証主義を唱える基となっている[1]。カプトガニの視神経でもマッハバンドが起り、その機構は隣合う神経細胞への側方抑制（Lateral inhibition）であることをHartlineが見つけたのは80年後の1949年であった[2]。それ以来、ヒトの認識も神経回路の情報処理として扱われるようになった。現在のニューラルネット、ニューロコンピューティングの研究分野はHartlineの側方抑制模型のさらに40年後の姿である。

新しい情報技術とは我々の意識がより広い世界に触れることを可能にするものであろう。我々を取り巻く世界の一部として、将来どのような情報メディアが望ましいのかを考えるに当たって、我々以外の生物が外界をどのように見ているか、外界にどの様に働きかけているか、その生理学的機構と生物学的背景を述べ、生物の採った戦略と物理的世界のとらえ方の多様性を考察する[3]。

ゴキブリやコオロギは空気の急な動きで逃避行動を起こす。彼らは体の後端の尾葉に気流の速度と加速度を検出する感覚毛を持っている。ゴキブリは50%位の成功率でヒキガエルの舌から逃れる。尾葉に糊を塗ると、成功率は8%に落ちる。カエルはゴキブリを見つけると頭を標的に向け、3~4 cmのところまで顔を突き出してから舌を打ち出す。ゴキブリはカエルの頭が押しのけた空気の動きを感じて逃げる。気流に流速だけがあって、加速度が低いのは、遠くで大きなものが動いていることに相当する。このとき、ゴキブリは動きを止めて背景にとけ込み、カエルに発見される確率を下げる。流速だけでなく、加速度も大きい気流はすぐ近くに動くものがあることを意味する。この時には、とにかく逃げる。

蛾の幼虫には、スズメバチが近づくと、体を固くして食草から落ちるものがある。胸部にある細長い感覚毛を除くと、この行動は起こらない。コオロギ、バッタなどを狩って針で麻痺させ、卵を産みつけるアナバチがいる。コオロギも黙って狩られるわけではなく、ハチの羽音に対して前中胸の4肢で体を支え、浮かせた後肢で後ろから襲いかかるハチを蹴り返す。この防御姿勢はコオロギの尾葉を取り除くと起こらない。ハチのように波長よりはるかに小さな振動板は $1/6$ 波長以内の空気をゆり動かすだけで、外部へは圧力波を放射しない。感覚毛で流れを測っていれば、小さな音源の接近のみを検出できる。我々の耳のような受圧膜では流れをさえぎることによる動圧と遠くからくる圧力波を区別できない。トンボの幼生（ヤゴ）は生き餌だけを捕らえて食べる。ヤゴは夜行性で、泥質の水底に棲む。視覚は役に立たない。ヤゴは小動物が作る水流に反応して捕食行動を起こす。低レイノルズ数の世界に棲む昆虫と、より大きなレイノルズ数の世界に棲む我々とは意味のある情報も、その情報を担う物理量も違う。

アリやミツバチは餌場や巣の方向を定位するのに太陽コンパスを使う。この時、太陽が直接見える必要はなく、青空の一部が見えるだけで良い。これらの動物に偏光を与えると方向が狂う。青空から来る光は太陽光がレイリー散乱されたもので紫外部に富み、太陽、天空の一点（散乱体）、観測者がなす面に垂直な電場振動面を持つ偏光を多く含む。つまり、天空上で太陽を中心とした同心円に沿った偏光が多い。従って天空の任意の2ヵ所で

最大偏光面を測れば、太陽はそれぞれに垂直な2つの大円の交点にあるはずである。しかし、アリやハチはこの問題を球面幾何学で解いてはいない。彼らの複眼の背縁部の視細胞は紫外に感度が良く体の左右方向の偏光の度合を検出できる。ハチの眼の背方に左右方向の偏光フィルタを付け、巣板の上で仲間に餌場の方向を伝える8の字ダンスをしているところをTVカメラで写す。テレビモニタ上のハチの向きのコサイン成分に従って明るさを変える光（非偏光）を与える。ハチは見事に騙されて、人工光が一番明るくなる時の「自分の向き」を太陽の方向とみなして振舞う。ハチは青空から来る光の偏光方向を測っているのではなく、生れつきの左右方向の偏光メガネをかけて、体を回して天頂部を走査し、空が一番明るく見える方向を太陽方向とみなすのである。ハチが眼の左右対称性を使って、天頂部の偏光の対称軸を見つけていると言っても良い。

昆虫には複眼の他に正面、左、右の3つの単眼がある。単眼のレンズはひどく弱くて、その下に並ぶ視細胞には広い範囲のひどくぼけた像が写る。単眼の視細胞は紫外部に感度が良いが、少数の神経に収束しており脳へ送られる信号は空間情報も偏光情報も持ち得ない。昆虫は機械的な平衡器を持たないにもかかわらず、水平飛行できる。昆虫を背中で軽く支えて前から風を当て、肢の接触を断つと、飛び始める。まわりの上半分が明るく下半分が暗ければ、飛び方は正常である。左右どちらかの単眼に影を作ると、影の側を上げる。正中の単眼に影をつくると、頭部を上げる。自然界では紫外光は空からしか来ない。夜空も地表に比べれば十分明るい紫外光源である。前と左右の地平線を含む大きな空間を見て、紫外光量が減った側を持ち上げるように飛行運動系に負帰還をかければ水平を保って飛べる。単眼は飛行の際の地平線検出器で、弱い光学系と強い神経収束は、光量の空間積分を意味している。

神経系は感覚入力を行動出力に変換しさえすれば良い。コオモリが超音波パルスを発射してそのエコーを分析するとき、外界の特徴をいちいち脳の上に並べ上げる必要はない。エコー入力がある分布をした活動を脳表面に作り出し、その分布模様に従って飛行運動系が動いて標的を捕捉すればそれで良い。捕捉すべき標的の時に、単一の“エサ細胞”が発火する必然性は無い。脳では情報を三次元の構造空間へ展開、投射を繰り返して、その中の相関（秩序）を浮き彫りにしている。作り上げられ易い秩序は動物毎に偏っている。その動物にとって生存上意味のある外界の像と我々ヒトのそれとが一致すれば“オバアサン細胞”にみえ、ヒトのそれとはかけ離れているときは、単に脳表面に分布した神経活動に見える。

生物の進化は盲目的である。遺伝子の複合体（生物の種）はその組合せを変えながら、コピー数の多い状態へと局所勾配に従ってすべり落ち、小さなくぼ地の一つ（局所的安定解）にはまり込む。何かの力が加わって別のくぼ地に向かうにしても、小さな変更の積み重ね（変異と淘汰）しか方法はない[4]。ある形質の前途有望さ、つまり遠い子孫が受け取るであろう利益に感激して自身がコストを先払いする祖先種はいないから、基本設計にまで遡るような大きなモデルチェンジは起こらない[5]。将来出会うであろう課題のためにより一般的な問題解決能力を身につけたり、めったに起こらないことに対処する神経系が進化することは無い。自然が淘汰を通して情報処理装置である神経系を作り上げるとき、それぞれの動物に科された生物学的タスクを、その動物が棲んでいる条件の中（ニッチェ）で正解と同じ結果を生みさえすればどんなトリックで解いても良い。トリック解は、しかし、その設計条件の外へ引き出されれば、馬鹿げた結果になる。夜間飛行をしている単眼水平器が強い点光源の近くを通りかかれば「飛んで火に入る夏の虫」になる。

進化には方向性がなく、神経系も一般的方法で生物学的課題を解いてはいないが、生物の機能が無原則に出てきたと考えるのは正しくない。それぞれの動物は、淘汰つまり有限の資源を巡るゲームを通して、多次元の遺伝的空間のある局所安定点に落ち着いている。つまりそのニッチェでの必要条件を満足している。空間認識に光を利用できない条件下で、コオモリはエコーロケーションを発明したが、同じ方式は少なくとも二度、イルカとウミツバメで独立に成立している。われわれも、周りの空間の認識に多少なりとも音の反響に頼っている。一方で、アマゾンやアフリカの弱電気魚は泥水中での空間認識タスクを全く別の方式、電流場の生成とその変形の検出で達成している[6]。神経系は、自然が淘汰を通して作り上げた情報処理装置であるが、適応の過程で蓄えられた生き方に関する遺伝情報を、感覚入力に従って読み出す装置でもある。トリック解は、遺伝情報に基づいた、感覚情報の能動的処理と見なせる。能動的処理は勝手な解釈を生み出すが、動物もヒトも勝手に解釈しても間違いの無い所（ニッチェ）にのみ棲む。勝手な解釈は、一般性を求めてすべての場合を調べ尽くそうとするときの、計算量や計算資源の爆発的増加を回避する最善の戦略であろう。

一般にヒトの脳は計り知れない能力を持っていると考えられている。しかし、自然から要求されもしない機能が進化するはずはない。かつて我々の祖先は、この脳の機能が無ければ生き延びれない厳しい淘汰（おそらく食料や繁殖を巡っての駆引き、競争、騙し合い）を経験しているはずである[7]。現在は、我々の脳を進化させたこの淘汰圧が緩んでおり、

我々はこの構造を他の目的に転用できる幸運を手に入れている。この物理的世界の解釈は一通りである必要はない。将来の情報メディアは、そこに棲み得る動物の世界(Umwelt)を、ページをめくるように実感させてくれる夢違観音であって欲しい。

【参考文献】

- [1] Mach Ernst: Über die Wirkung der räumlichen Verteilung des Lichtreizes auf die Netzhaut, I. Akademie der Wissenschaften in Wien Sitzungsberichte, math.-nat Cl., 1865, 52, Abt. 2, 303-322 (1865).
- [2] Hartline H. K.: Inhibition of activity of visual receptor by illuminating nearby retinal areas in the Limulus eye. Federation Proceedings, 8, 69 (1949).
- [3] 下澤 楯夫: 日本ロボット学会誌, 6, 240-244, (1988).
- [4] R. ドーキンス: 延長された表現型 (日高、遠藤、遠藤訳) 紀伊國屋書店, (1987).
- [5] S. J. ゲールド: 個体発生と系統発生 (仁木、渡辺訳)、工作社, (1987).
- [6] 川崎 雅司: 電気魚のニューロエソロジー、一混信回避行動の神経機構一、科学, 59, 437-445, (1989).
- [7] 澤口 俊之: 知性の脳構造と進化—精神の生物学序説一、海鳴社, (1989).

4.1.3 空間認知の神経メカニズム

自分の意志で移動するという能力を身につけた動物が、安全な移動、目的に沿う移動を行うために必然的に備えなければならなかった能力、すなわち外空間の構造を認識・理解する能力は、近い将来開発が期待される移動ロボットにも不可欠である。外界認識のための入力情報として、動物が一般的に用いているのは視覚情報と聴覚情報である。ここでは視覚および聴覚による外空間認知の神経情報処理について概観してみよう。

(1) 視覚による空間認知の神経メカニズム

外空間の構造を認識するための視覚的手がかりには、

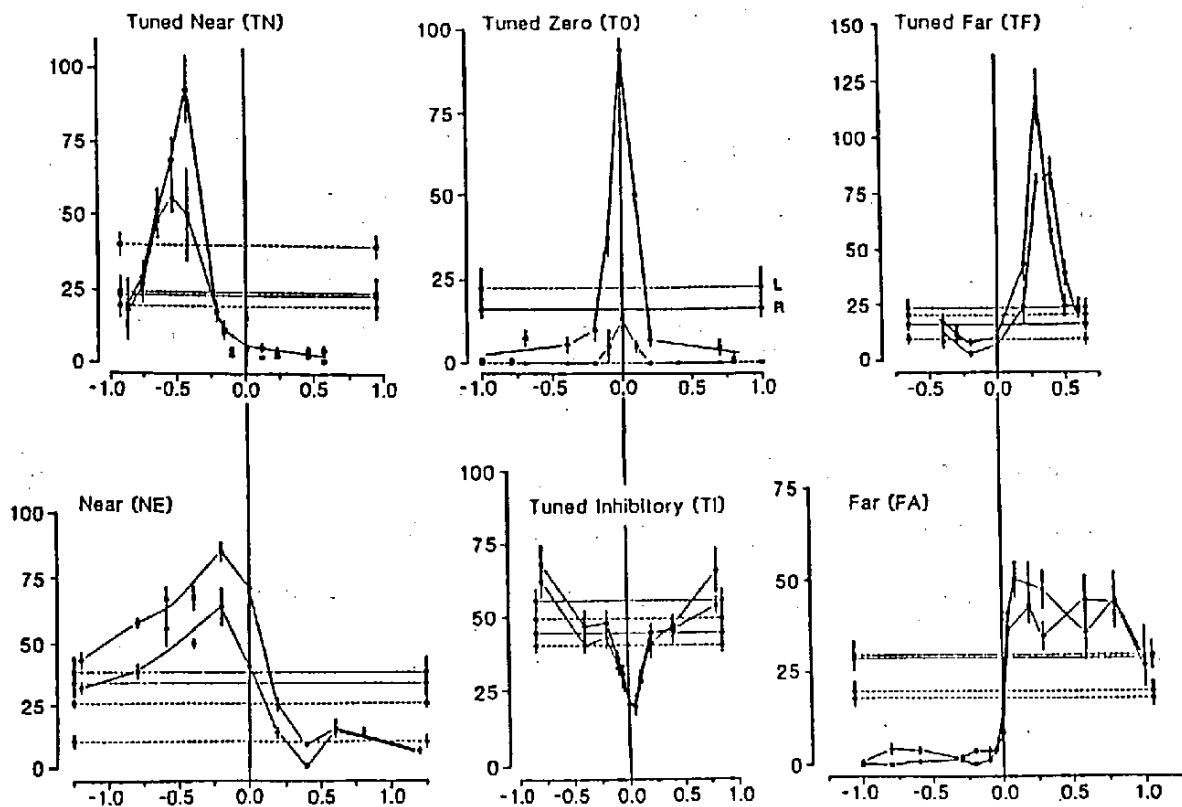
- a) 静的手がかり (両眼視差、陰影、パースペクティブなど) と
- b) 動的手がかり (運動視差、空間速度勾配、視流線など)

がある。

(1.1) 静的手がかりを用いた空間情報の分析と脳内表示

(1.1.1) 両眼視差

パターンの左右眼視差の値に選択的に反応する神経細胞（視差選択性細胞）は、両眼情報が初めて一つの細胞に集束する大脳皮質第一次視覚野（V1野）にみつかると、V2野、V3野と階層的に上位（図表4.1-4）に行くにつれその割合が増えてくる[1]。どの領野にも次の6種類の視差選択性細胞がある。すなわち、視差がゼロ（注視面）付近のある値に急峻な選択性をもって興奮または抑制反応を示す細胞群（tuned excitatory, TE細胞；tuned inhibitory, TI細胞）、注視面より近く又は遠くの視差に急峻な選択性をもって興奮反応を示す細胞群（tuned near, TN細胞；tuned far, TF細胞）、注視面より近くの視差には興奮反応、遠くの視差には抑制反応をする細胞群およびその逆の反応選択性を示す細胞群（near, NE細胞；far, FA細胞）の6種類である（図表4.1.2）。



図表4.1-2 V1野～V3野にみられる視差選択性細胞の反応特性[1]

いろいろな視差平面（横実には注視点より遠方を正にとって視度で表す）で左右に動く、幅 $0.1 \sim 0.3^\circ$ 、長さ $1 \sim 3^\circ$ のスリットに対する反応（縦軸にインパルス/秒で示す）。図中の粗い点線は左眼単独刺激の反応を、細かい点線は右眼単独刺激の反応を表す。両眼視差刺激に対する反応は、それぞれスリットの左方向および右方向の運動に対して得られた2年のチューニングカーブで示されている。

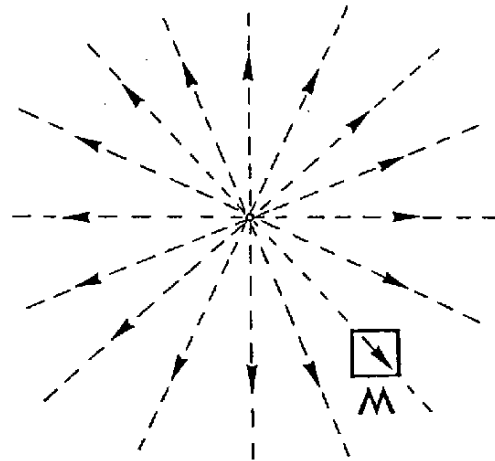
これらの細胞により奥行きを脳がどのようにコードしているかについては、「脳には最適視差が規則的に変わる細胞配列があり、奥行きはどの細胞がもっとも強く反応したかで脳に表示される」という‘おばあさん細胞仮説’的な考え方、「注視面付近の精密な立体構造はTE細胞とTI細胞の活動で、視野の奥行き構造のような粗い立体構造は、NE細胞とFA細胞の活動のバランスで表示される」という‘細胞集団仮説’的な考え方の2つがある。

(1.1.2) 陰影とパースペクティブ

両眼視差は視野の局所ごとに計測できる局所情報であるが、陰影やパースペクティブは、ある程度広い視野領域の情報を統合して初めて検出できる情報である。従って、それらをコードする細胞はそれに見合う広さの受容野をもつ必要がある。神経生理学的研究では、そのような情報を担っている細胞の存在は明らかにされていない。しかし LehkyとSejnowski[2]は、誤差逆伝搬学習を取り入れた3層神経回路モデルに陰影による物体表面の凹凸の判断を学習させたところ、モデルの中間層の素子はV1野の単純型細胞に、終層の素子は複雑型細胞に反応特性が似てきたことを報告している。

(1.2) 動的手がかりを用いた空間情報の分析と表示

自ら動くことができるという特典をもつ動物は、動くことによって外界に積極的に働きかけ、その3次元構造を知るのに有用な動きの情報を作り出すことができる。その一方で、自ら動くという特典はパターンの動きを複雑にする。すなわち網膜上のパターンの動きは、物体の動きと観察者である自分の動きの組み合わせで生じている。人間を代表とする視覚性の動物は、そのような複合された動きの観測データに基づき、物体や外空間の3次元構造、物体の空間内での運動状態、外空間に対する自分の運動状態などを認識理解することができる。このことは、脳に少なくとも「身体運動に影響されずに物体の動きを計算するサブシステム、すなわち視野の局所とその背景の相対運動を計算するサブシステム」および「物体の運動に邪魔されずに身体運動による広視野の動きを計算するサブシステム」が備わっていることを示す。どちらのサブシステムも、“小窓問題”（図表4.1-3）および動きの原因に関する“不良設定問題”をうまく処理しなければならない。小窓問題については、局所運動の分析→統合による大局的運動の識別という階層処理で脳はそれを解決しているのである。

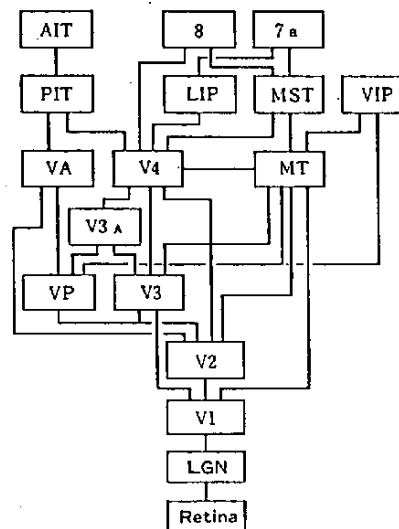


図表4.1-3 小窓問題

小さな窓から覗き見た視野のごく一部の運動には、大局的運動の違いを知る手がかりが含まれていない。この図は、視野の拡大によるパターンの流れ（視流線）を小窓（W）から覗いた場合を示しているが、W内の動きは視野全体が斜め左上45°方向に動いた場合と区別できない。このような局所運動情報からどのようにして大局的運動の様相を正しく決定するかという問題を小窓問題と呼ぶ。

(1.2.1) MT野における局所的運動の分析[3]

マカクザルのMT野は、網膜に始まる視覚分析路の中で図表4.1-4 に示すような階層的位置にあり、おもな信号入力をV1野から直接受けている[4]。



図表4.1-4 マカクザルの網膜から中枢に至る視覚経路の階層構造

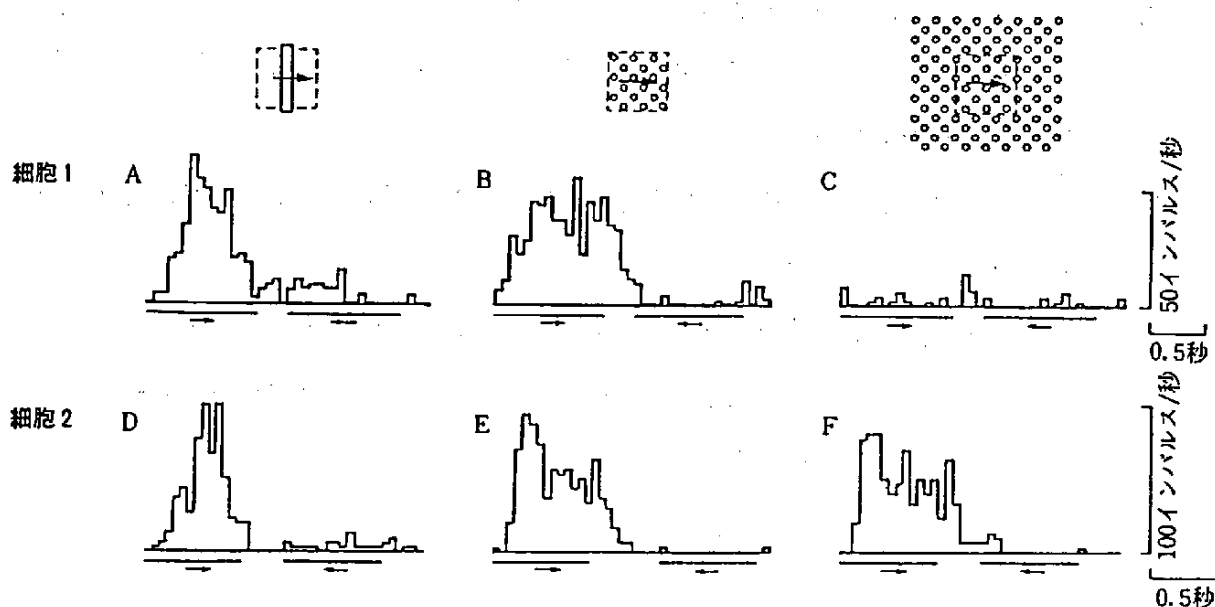
視覚中枢は多くの独立した小領域から成る。それぞれの小領域には名前が付けられており、各小領域間の階層的結合関係が明らかにされている。V1野からMT野、MST野を経て7a野に至る経路ではパターンの運動情報がV2野、V4のを経てIT野に至る経路では色や形の情報が分析・結合されている。
Retina: 網膜。 LGN: 外側膝状体。
(Van Essen[4] の図を改変)

MT野は次の理由で局所運動の分析・表示システムであると認定される。

- ① MT野細胞の受容野は小さく（ $1 \sim 10^\circ$ のオーダ）、左右MT野内に対側視野の視野地図がある。
- ② MT野の細胞のはぼ90%がパターンの直線運動に方向選択的に反応する。
- ③ MT野の皮質表面に沿って最適運動方向が規則的に変わる細胞配列がある。
- ④ MT野の細胞の反応は動きのスปีドにも選択性を示す（最適スปีドは $2 \sim 256^\circ$ /秒のレンジにある）。

MT野はパターンの動きの原因（物体運動と自己運動）分析の開始点となっている。

というのは、MT野の方向選択性細胞が次の2種類に等分されるからである[5]。一方の M^- 型細胞では、受容野（興奮性受容野）内に与えた最適運動刺激に対する反応が、その周囲に方向もスปีドも同じ運動刺激を与えると周囲刺激がない場合の半分以下に抑えられる。もう一方の M^+ 型細胞では周囲の動きによる反応の抑制が50%未満である（図表4.1-5）。



図表4.1-5 MT野の細胞の2つの型[5]

細胞1、細胞2のどちらも受容野の中を動くパターンに対してはその形に無関係に反応するが（AとB、DとEを比較）、細胞1ではパターンが受容野の周囲にまで拡がると反応が消失するのに対し、細胞2ではそのようなことがない（CとF）。細胞1は M^- 型細胞の典型、細胞2は M^+ 型細胞の典型である。

縦軸は細胞が1秒当り発射するインパルス数、横軸は時間を表す。反応ヒストグラムの上の挿絵の四角の点線の囲みは受容野を、刺激パターンの中の矢印は細胞にとっての最適運動方向に示す。ヒストグラムの下に添えた直線はパターンが運動している期間を、矢印は刺激パターンの運動方向を示す。

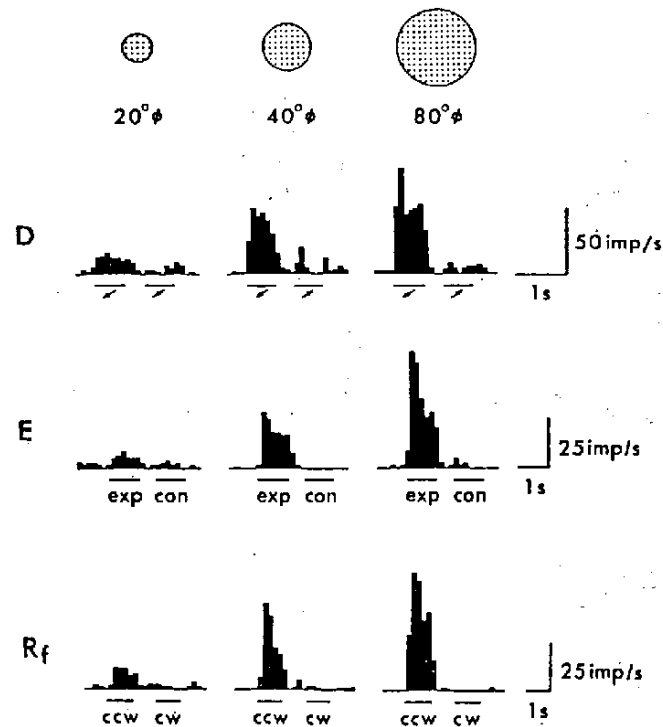
M⁻ 型細胞は受容野内のパターンの動きが周囲の動きとは異なるときに初めて大きな出力を出す。つまりこのような細胞は、運動の不連続の程度とそれが生じている視野上の位置を計算し、その結果を地図表示していることになる。これは、物体の存在、その3次元空間内での位置や運動状態を認識するための情報となる。M⁻ 型細胞の約半数では、受容野中心部と周辺部それぞれの単独刺激に対して反応する運動方向が逆転する。このような細胞は真の相対運動をコードしているといえる。

M⁺ 型細胞は、刺激の面積を広げても反応が抑制されないので広視野運動の知覚へ関与するが、局所運動と視野全体の一様な運動とを区別する能力を持たない。それを区別して分析する細胞は、MT野から入力を受ける MST野に発見された。

(1.2.2) MST野における広視野運動の分析[3]

MST 野背側部は、次の理由により自己運動の様相を認識・理解するための広視野の動きを分析・表示するサブシステムであると認定される[6]。

- ① 細胞の受容野が大きい（平均直径40°）。
- ② 運動方向選択性細胞が全細胞の約90%を占める。
- ③ 自己運動に伴って生じる広視野テクスチャパターンの4種類の基本的運動に選択的に反応する細胞が揃っている。
 - i) 等距離平面上を直線運動したときに方向選択的に反応する細胞（directionally selective cell, D細胞）、
 - ii) 奥行き方向に直線運動したとき、すなわち接近によるパターンの拡大（Expansion）、又は離反によるパターンの縮小（Contraction）に選択的に反応する細胞（E細胞、C細胞）
 - iii) 等距離平面上での回転（Rotation in frontoparallel plane、時計方向又は反時計方向）に選択的に反応する細胞（Rf 細胞）
 - iv) 奥行き方向の回転（Rotation in depth）に選択的に反応する細胞（Rd 細胞、希）で、D細胞>E細胞>C細胞>Rf 細胞の順に多い。
- ④ D、E、C、Rf 細胞のほとんどはパターンの局所的動きには感度が低い（図表4.1-6）[7]。
- ⑤ 反応の大きさがテクスチャの形によらない（図表4.1-7）[7]。

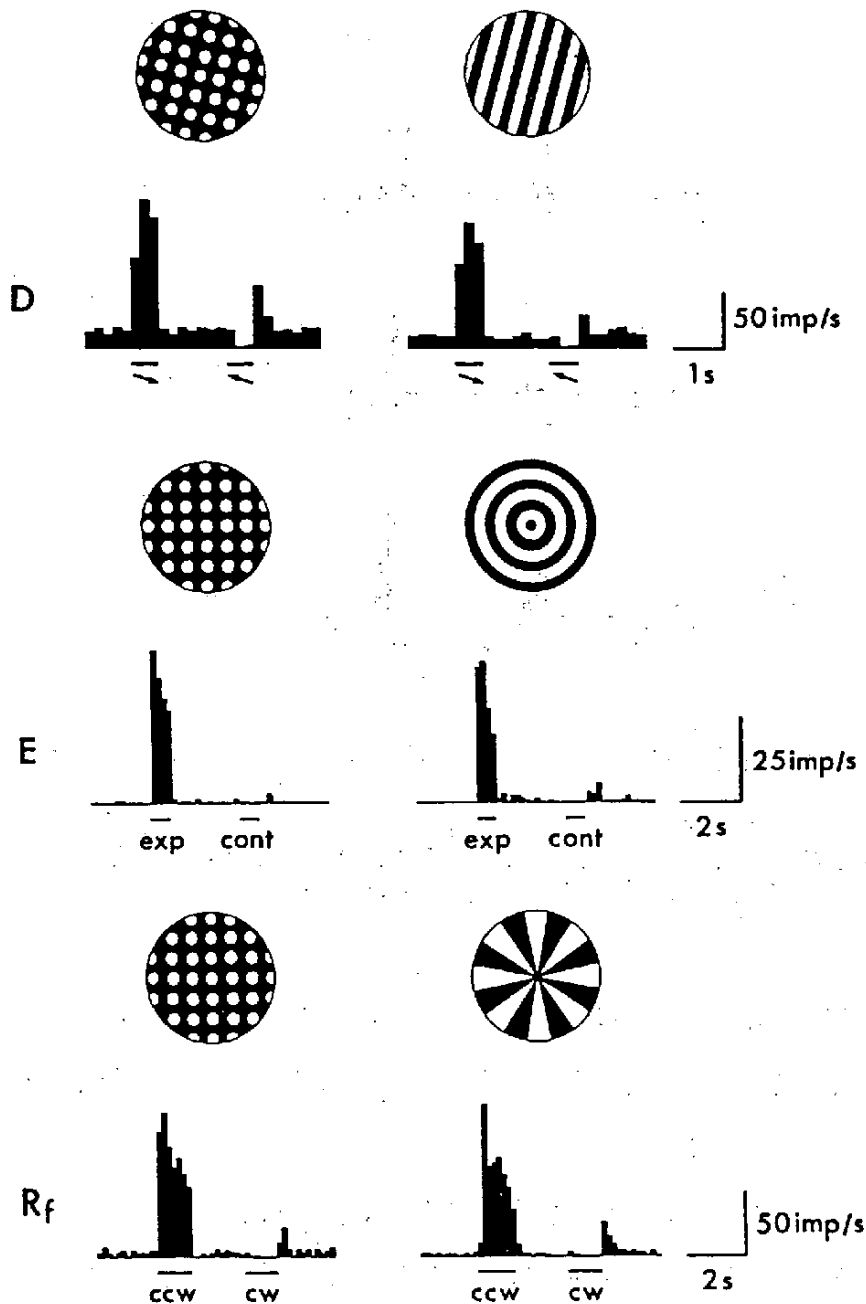


図表4.1-6 MST野背側部のD、E、R_f細胞の刺激面積と反応の大きさの関係[7]

〔上の挿絵に示す大きさの窓の中に運動するドットパターンを与えた。
exp: 拡大、con: 縮小、ccw: 反時計回回転、cw: 時計回回転。
単位: 縦軸、インパルス/秒; 横軸、秒。〕

眼前に展開されるパターンは時と場所により千差万別である。反応の大きさがテクスチャの形によらないことは、自己の動きによる広視野パターンの動きを分析するのに経済的であるといえる。

図4.1-8 はD、E、R_f細胞の反応選択性を決める神経回路モデルである。MT野のM⁺型細胞から受け取る局所運動方向の情報を空間的に統合することにより方向選択性が決まると考えたものである。局所運動方向の情報を入力する素子としてM⁺型細胞を選んだ理由は、M⁻型細胞は広視野のコヒーレントな動きには反応しないからである。MST野背側部の細胞が広視野の動きにしか反応しない理由としては、閾処理型の非線形があると考えればよい。同時に興奮するMT野細胞の数は、物体が局所的に動いた場合に比べ広視野パターンが動いた場合の方が圧倒的に多い。すなわち、物体の局所運動によってもたらされる興奮はこの領野の細胞を活動させるには至らないと考えるのである。

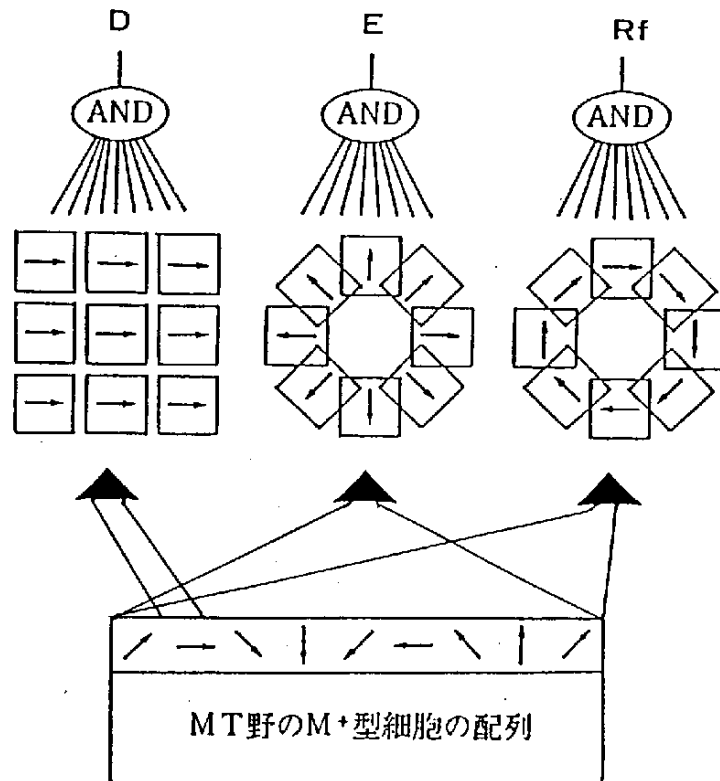


図表4.1-7 MST野背側部のD、E、R_f細胞の反応の強さはテク
スチャ要素の形に無関係である[7]

挿絵に示すように、切れ切れのテクスチャとしてはどの運動に対してもドットパターンを用い、長い線で構成されるテクスチャとしては、直線運動、拡大縮小、回転それぞれに対して運動方向が線分の長さ方向と直角をなすようなパターンを選んだ。

exp: 拡大、con: 縮小、ccw: 反時計回転、cw: 時計回転。

単位: 縦軸、インパルス/秒; 横軸、秒。



図表4.1-8 MST野の背側部のD、F、R_f細胞の反応選択性を作る神経回路モデル

これらの細胞の反応選択性は、いずれもMT野の細胞が検出した局所運動方向の情報の統合により説明される。矢印の入った囲みはMT野の細胞の受容野の相対的位置と検出する局所的運動方向。背側部の細胞はM⁺型MT野細胞の運動方向と視野の位置の細胞配列から中段に示す受容野特性をもつ細胞を選んで入力を受け取ると考えられる。

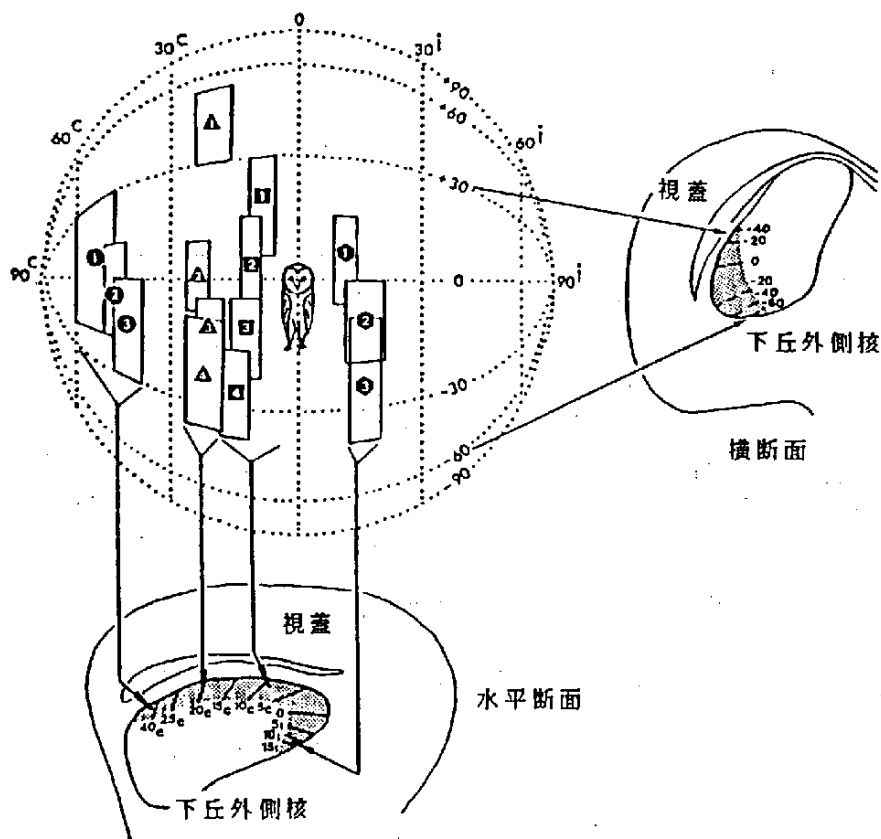
MT野のM⁻型細胞、MST野のD、E、C、R_f細胞は、局所運動を空間的に統合することにより“小窓問題”を解決し、運動の不連続性（M⁻型細胞）や統一性（D、E、C、R_f細胞）をコードしていることを示した。しかし動きの原因に関して解が一意に決定されたわけではない。例えば、広視野が特定の視流線を伴う統一的な動きをしたとき、それを自分が動いたための視野の動きと判定するのは、「観察者である自分が動かなければ広視野が統一的動きをすることはほとんどありえない」という思い込み的仮定を採用しているからである。隣の電車が動き始めたただけなのに自分の乗っている電車が発車したと錯覚するのは、このために生まれる誤った解である。しかしこの誤りを冒すこと自体、私達の外空間認識がいかに強く視覚に依存しているかの証であり、またMST野背側部の細胞の表示する情報が不完全ではあっても使いものになっている理由でもある。

(2) 聴覚情報を用いた空間情報のコーディングと脳内表示

(2.1) 左右耳時間差情報と音圧差情報を用いた音源方向表示

音源の方向を知る主な手がかりには、左右の耳に入る音の時間差の情報と強度差の情報の2つがある。音源の方向知覚を高度に発達させている動物はフクロウである。夜行性で、暗闇の中で動きまわる小動物を捕える面フクロウは、獲物のいる方向を音だけで正確に把握していることが行動実験で示されている。

小西らの一連の研究（文献[8]、[9] 参照）によると、面フクロウの聴覚神経系には左右耳に入る音の時間差と音圧差を別々に処理する独立の経路があり（図表4.1-10）、2つの情報は最終的に統合されて中脳の下丘に2次元の聴空間の地図が表示されることが明らかになっている（図表4.1-9）。驚くべきことに、面フクロウの耳の位置は左右で段違いになっており、顔面にある飾り羽根でできた扇状の音の反射板の向きも左右非対称になっている。これらは、特別な進化がもたらした音圧差を精度良く捕えるための構造として注目に値する。

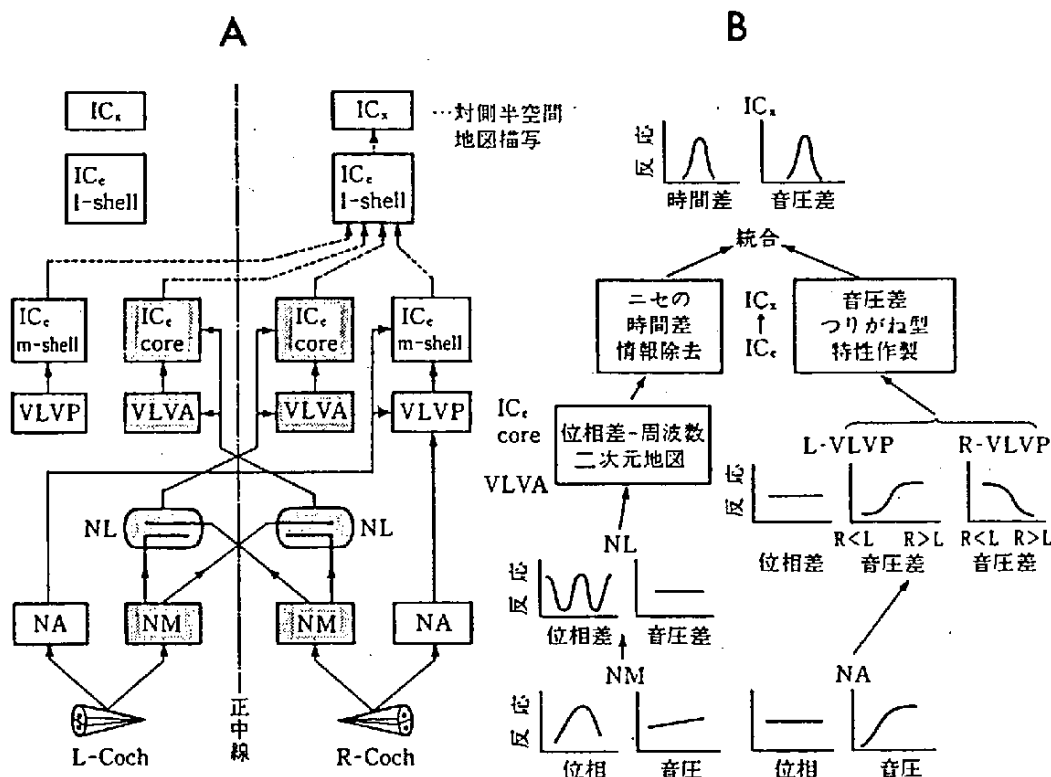


図表4.1-9 面フクロウの下丘細胞による音源方向の地図表示

下丘外側核に電極を刺入して記録される細胞の性質を調べると、個々の細胞は聴空間の特定の範囲の方向に音源があるときにのみ反応することがわかる。その範囲を受容野と呼ぶと、下丘外側核内には受容野の位置が規則的に変わるように細胞が配列されている。

聴空間内にちりばめられた四角で囲まれた領域は、矢印の部位の電極刺入で記録される細胞の受容野を示している。下丘水平断面に記入されている数字は聴空間の水平方向の角度（cは反応側空間、iは同側空間）を、横断面に記入されている数字は聴空間の上下方向の角度を示す。（小西[8]による）

図表4.1-10は、小西らが明らかにした聴空間地図作成の神経システムを斎藤が簡単なブロックダイアグラムとして描いたものである[9]。まず蝸牛神経核の2つの核のうち大細胞層（NM）では音刺激の位相が、角状核（NA）では音圧が左右耳それぞれ別々に分析される。左右耳からの位相情報は次段の層状核（NL、上オリーブ核に相当）に送られる。図表4.1-11Aに示すように、左右の大細胞核からの神経線維は互いに層状核の別々の方向から入り、核の短軸に沿って反対側まではほぼ直線状に伸びている。この線維はパルス信号の遅延線の役目を果たす。層状核の細胞は両側から入ってくる神経線維から入力を受け、左右からの信号がほぼ同時に到着して入力信号の時間加算が生じたときに反

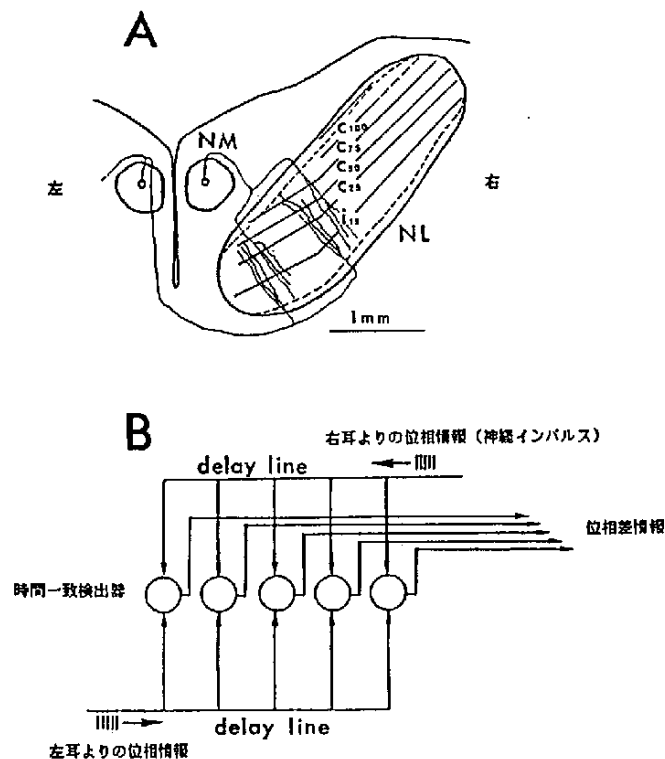


図表4.1-10 面フクロウの聴覚神経系における音源方向分析の主要経路の統計図[9]

- A : 音源の左右方向を表示するための両耳時間差情報と音源の上下方向を表示するための両耳音圧差情報は、それぞれ別々の経路を経て分析され下丘の外側核 (ICx) に至るまでに統合されていることを示す。影を施した神経核群では時間情報が、影を施していない神経核群では音圧情報が処理されてゆく。点線は推定される結線。
L-Coch、R-Coch : 左蝸牛、右蝸牛、NM : 蝸牛神経核・大細胞核、NA : 蝸牛神経核・角状核、NL : 層状核 (上オリブ核に相当)、VLVA : 外側毛帯核・前核、ICc core : 下丘中心核・芯核、VLVP : 外側毛帯核・後核、ICc m-shell : 下丘中心核・内側蓋核、ICc l-shell : 下丘中心核・外側蓋核、ICx : 下丘外側核。
- B : 時間差情報、音圧差情報の処理経路の各神経各における単一神経細胞の反応特性の模式図。
反応特性の変容を辿ると、情報が音源定位に向けて洗練されていく過程がわかる。蝸牛神経核の2つの核 (NMとNA) で音波の位相と音圧が独立に検出される。左右の蝸牛神経核で検出された位相と音圧の情報はそれぞれ比較され、最後に統合されて特定の左右時間差と特定の左右音圧差の組合せに反応する細胞が下丘外側核 (ICx) にできあがる。ICx には時間差と音圧差の組合せで表現された聴空間の地図を示す細胞配列がある (図表4.1-9 参照)。

応する仕組みになっている。左右の大細胞核細胞から同時に信号を受け取るために必要とされる左右耳刺激の時間差は、核内での短軸に沿う細胞の位置により規則的に変わる。つまり、層状核内には左右耳時間差が規則的に変わる細胞配列があることになる。図表4.1-11Bは Jeffress[10] によって提案された左右時間差を計算表示するモデルである

が、面フクロウの層状核ではまさにこれと同じ仕組みで左右耳刺激の時間差が計算・表示されていることになる。層状核の中で観測される最大遅延は $180\mu\text{sec}$ で、これはちょうど音源が真横にあるときの左右耳時間差に相当する。



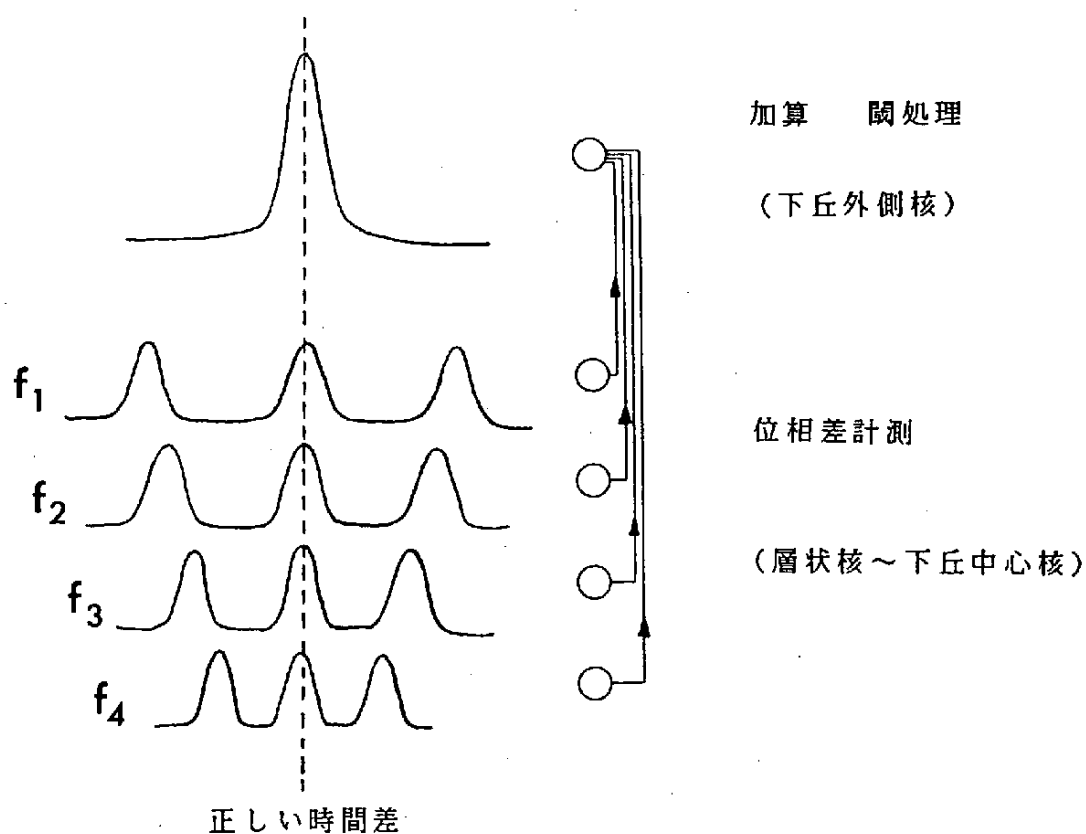
図表4.1-11 面フクロウの層状核における音信号の左右耳位相差計測表示システム[8]、[9]

- A : 層状核 (NL) 内には、反応する音信号の左右耳位相差が規則的に変わる細胞配列がある。C100、75、50、25は対側の耳に到達した音波の位相が同側の耳に到達した音波の位相よりそれぞれ 100、75、50、25 μsec 進んでいるときに反応する細胞が並んでいる線。i 15は同側の方が15 μsec 進んでいるときに反応する細胞が並んでいる線。線に沿う方向には反応する音の周波数が変わる。互いに向き合う方向から層状核に入る左右の蝸牛神経核・大細胞核 (NM) の細胞の神経線維が層状核内で遅延線の役目を果たしている。
- B : 左右耳時間差検出回路のモデル[10]

層状核の長軸は周波数軸になっている。つまりいろいろな周波数の信号に対して時間差が計測され、それらが層状核の直交する軸に沿って2次元状の地図として表示されているのである。この情報は下丘の中心核のコアと呼ばれる部位に送られ、時間差と周波数に対する選択性がさらに尖鋭化され直交する軸に沿って整然と地図表示されている。

さて、個々の層状核細胞によって計測された時間差信号がそのまま音源の左右方向を知るための時間差を表すわけではない。なぜなら、音刺激に周期的成分が強く含まれて

いる場合には、層状核に送られてくる大細胞核からの位相情報にも周期的成分が現れ、層状核の細胞が検出しているのは、実は“にせ”の時間差信号をも含んだ位相差情報になっているからである。周波数と位相差の2次元マップは、この“にせ”の時間差信号を消去するのに必要なのである。というのは、“にせ”の時間差は周期波の周波数によって異なるが、真の時間差は周波数に依存しないので、いろいろな周波数で計測した位相差情報の中から共通なものを選び出すことにより真の時間差を求めることができるからである（図表4.1-12）。共通信号の選び出しは、抑制をまじえた非線形処理でなされているらしい。そしてそのような処理過程は下丘の外側核にあると考えられている。



図表4.1-12 位相差信号に含まれる時間差多義性の除去法[8]

いろいろな周波数の音で計測して得られる左右耳位相差信号を加算すると、周波数によらずに一定の時間的位置にでる信号が最大のピークとなる。この時間が真の左右耳時間差である。したがって、加算→閾処理という非線形処理により周期的位相差信号より真の時間差信号がとりだせる。横軸は左右耳時間差、縦軸は細胞の反応強度。

左右耳に到達した音の強さの関係は別のルートで同様に何段かの処理を経て音圧差情報として下丘の外側核に送り込まれ、そこで時間差情報と統合されて音源の方位を表す情報となる（図表4.1-10）。

【参考文献】

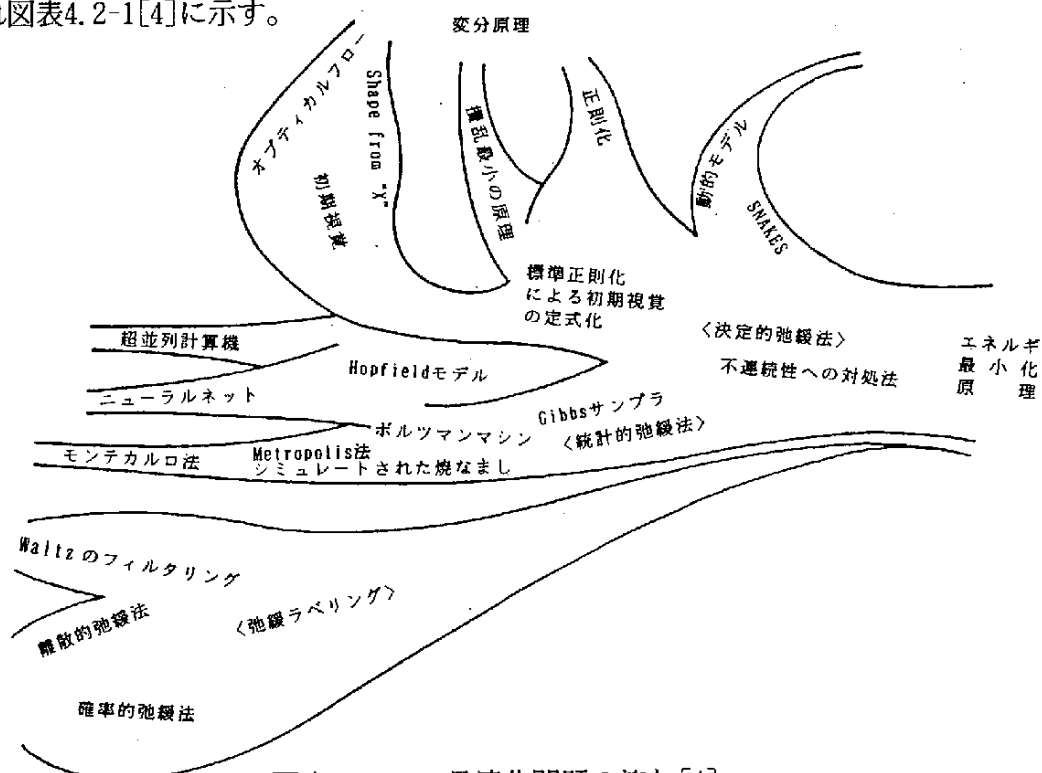
- [1] Poggio, G. F., Gonzalez, F. and Krause, F. : Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *J. Neurosci.*, 8, 12, 4531-4550. (1988).
- [2] Lehky, S. R. and Sejnowski, T. J. : Network model of shape-from-shading: Neural function arises from both receptive and projective fields. *Nature* 333, 452-454, (1988).
- [3] 斎藤 秀昭 : 高次視覚中枢におけるパターンの運動情報のコーディング、遺伝, 1989, 11, 24-33 (1989)
- [4] Van Essen, D. C. : Functional organization of primate visual cortex. in *Cerebral cortex*, 3, ed. A. Peters and E. G. Jones, pp.259-329, Plenum Press, New York & London. (1985)
- [5] Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y. and Iwai, E. : Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurosci.*, 6, 1, 134-144. (1986)
- [6] Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. and Iwai, E. : Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci.*, 6, 1, 145-157. (1986)
- [7] Tanaka, K. and Saito, H. : Analysis of motion of the visual field by Direction Expansion/Contraction, and Rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.*, 62, 3, 626-641. (1989)
- [8] 小西 正一 : フクロウの音源定位の脳機構、科学、60, 1, 18-28. (1990)
- [9] 斎藤 秀昭 : 脳における視覚・聴覚の情報処理、
松本 元編 「バイオエレクトロニクス」、pp.103-145. 丸善 (1989)
- [10] Jeffress, L. A. : A place theory of sound localization. *J. Comp. Physiol. Psychol.*, 41, 35-39. (1948)

4.2 感覚・知覚（視覚、聴覚、その他）

4.2.1 最適化問題の流れ（正則化、弛緩法）

認識・理解の多くの問題は、従来の論理的な積み上げだけでは解けない。なぜなら認識・理解、総合的判断、予測、学習といった処理は、問題が明確に定義し難く、「部分的な制約＝知識しか与えられない」ため、従来の情報処理形式には本来不向きである。従来の様に、情報が完全に与えられるとの前提で、与えられたデータを受動的に解釈するだけでは不十分であり、言語化される以前の状態ではあるが積極的に解釈しようとする無意識的な処理形式が要求される。このような問題を解くために、例えば初期視覚を最適化問題として定式化し、弛緩法や、正則化などの変分法による解法が提案されている。特にPoggioらが初期視覚を正則化理論で説明し[1]、統一的な枠組みを提唱してからこの流れは大きくなっている。

またニューラルネットの分野では、ボルツマンマシン[2] と呼ばれる統計力学へのアナログを基本原理としたネットワークが提案され、最適化マシンとしてある種のエネルギー最小化原理に基づく問題解法に有効であることが示されている。このアプローチはGeman & Geman[3]によって拡張され、統計的弛緩法として画像処理に適用され、その後の研究に大きな影響を与えている。こうした一連の弛緩法や正則化からエネルギー最小化原理への研究の流れ図表4.2-1[4]に示す。



図表4.2.1 最適化問題の流れ[4]

(正則化理論)

全体イメージの所でも述べたが聴覚や視覚の観測系を通して得られた情報から元の対照の物理的性質を推定する逆問題は、一般的に不良設定問題であることが多い。例えば、雑音除去、両眼ステレオ、距離データからの曲面の再構成、明るさの変化から傾きを求める問題などでは、データ不足のため、1) 一般的には、データと全く矛盾しない解釈は存在しなかったり、2) 与えられたデータだけでは解釈が一意には決まらなかったりする。このため画像理解の問題は不良設定問題となる。このため対象に関する何らかの仮定や拘束条件を用いる事によって解を特定する必要がある。正則化理論はこれらの問題を統一的に取り扱おうとする動きである。

正則化とは、求めたい物理量 x から線形操作 A を受けてデータ

$$y = Ax$$

が得られているとして、データ y から元の x を推定する問題の解法として、

$$E = \|Ax - y\| + \lambda \|Px\|$$

を最小にする x を求める最小化問題に持ち込むことである。

ここで $\|Ax - y\|$ はデータと解との差を表すペナルティ関数、 $\|Px\|$ は安定化汎関数と呼ぶ現実世界の一般的な拘束条件（例えば滑らかさの拘束）を表す。

すなわち、ある適切な拘束条件を用いて、解の存在空間を制限し、問題を解く手法である。

【参考文献】

- [1] Poggio, T., Torre, V. and Koch, C. : Computational Vision and Regularization Theory, Nature, Vol. 317, No. 6035, pp. 314-319 (1985)
- [2] Farlman, S.E., Hinton, G.E. and Sejnowski, T.J. : Massively Parallel Architectures for AI : NETL, Thistle, and Boltzman Machines, Proc. AAAI-83, pp. 109-113 (1983)
- [3] Geman, S. and Geman, D. : Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, IEEE Trans., Vol. PAMI-6, No. 6, pp. 721-741 (1985)
- [4] 坂上他 : 弛緩法と正則化、情報処理学会誌、Vol. 30, No. 9 (1989)

4.2.2 NP問題と能動的解釈

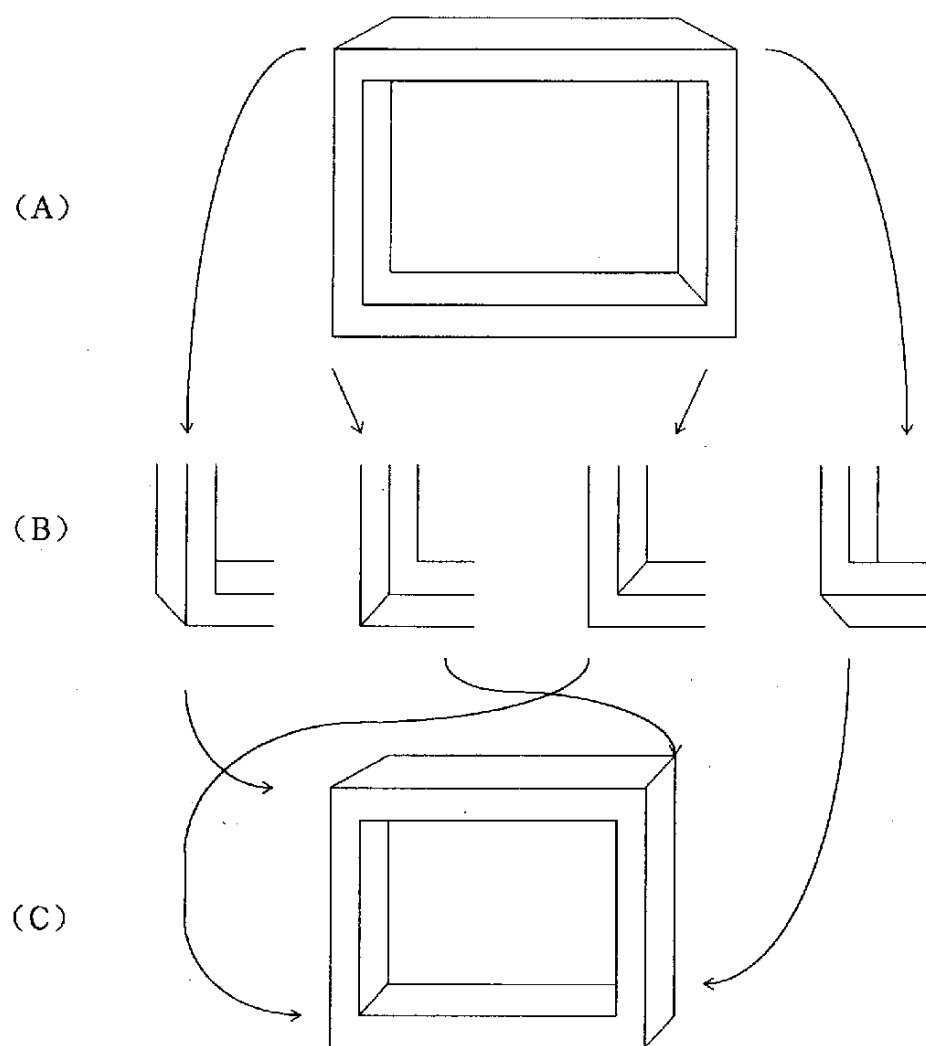
計測量の理論によれば、必要に応じて幾らでも多くのプロセッサが供給される並列計算機（そのような“夢”の計算機は現実には存在しない）が使えるというのなら話は別であるが、現実の計算機（逐次型計算機や有限個のプロセッサしか持たない並列計算機）ではどれほど工夫を施しても効率良く解けそうにない問題が多数存在することが知られている。NP完全、NP困難などと呼ばれる問題群がその例である。

一方、人間が身の周りの環境を認識・理解するために行っている情報処理の能力の中には、上の意味の難問をいともたやすく解いているかのように見える局面がある。線画からそこに描かれている立体の構造を抽出する線画解釈能力などが、その典型例である。しかし、このような人間の能力を難問を効率良く解く方法が存在することの証拠であるとみなして、ニューロコンピュータなどに過大の期待を寄せることは危険である。なぜなら、実際に人間が解いているのは、大きな問題群の中のほんの一部であり、それらは計算機によっても効率良く解けるものである場合が多いからである。

線画解釈問題を人間がいかにして解いているかは、錯視現象から推察できる。図表4.2-2 Aは、視覚心理学の分野で変則図形あるいは不可能物体と呼ばれている線画の例である。人間は、このような線画を見ると、大域的には何か変だと感じながらも、部分部分では立体構造を知覚してしまう。大域的に何か変だと感じるのは、意識にのぼる後期の情報処理であるのに対して、各部分に対して立体感を持つのは無意識に近いレベルで行われるより早期の情報処理であると思われる。このことから、人間は、線画を見たとき、図表4.2-2 Bに示すように部分的につじつまの合う立体的解釈を見つけ、基本的にはただそれらをつなぎ合わせているだけであると考えられる。実際、図表4.2-2Cに示すような正しく立体を表している線画に対しては、このような処理だけで正しい解釈が得られることが多いから、効率良く解けているように見えるのであろう。すなわち、情報全体を総合的に解釈するという困難な作業は避け、部分的につじつまが合う解釈が見つかったら、それに違いないと積極的に思い込む能動的な情報処理が行われていると考えられる。

このような解釈の能動性は、従来から、トップダウン、モデル駆動型などと呼ばれてきた情報処理の方式と同様のものである。うまくいく場面では効率の良い情報処理ができるが、いったん失敗すると回復が難しい。失敗した場合の危険を考慮すると、能動的解釈に処理をまかせることはできない。しかし、もう一方では、計算量の理論から、失敗することのない情報処理を実現しようとするすると極端に効率が落ちることが示唆されている。従っ

て、人間と同様の認識・理解能力を持った情報処理機構を作るためには完璧な解法を追求するという方向以外に、錯視のような“失敗”をある程度は許容するという前提で解法を考えることも大切であろう。



図表4.2-2 線画解釈

4.2.3 能動性について

感覚・知覚の様々なレベルにおいて、能動性は普遍的に認められる現象である。感覚・知覚の性質を調べて見ると外界の混沌とした不完全な情報の中から、自分にとって意味のある構造・情報を何とかして取りだそうとする強い傾向が認められる。

聴覚を例に取り、能動性がどのように現れているかを説明する。

音声信号は、1次元の信号であり、両耳からの入力があることを考慮しても、3次元の

外界を表現するには情報は全く不足しており、このような信号から、複数の音源の種類、位置、自分の置かれている空間の広さを初めとして、話し言葉の伝える意味内容、話し手の感情にいたる多様な情報が取り出される。この不良設定問題を解くために、様々なレベルにおいて信号の物理的構造、対象の性質等に関する制約条件が利用され、能動的な解釈・変換が行われている。

音声信号から意味をくみ取る努力は、聴覚末梢系において既に開始されている。耳から入った音声信号は、内耳の蝸牛で神経のインパルスに変換される。蝸牛の中では、基底膜と呼ばれる振動膜が、入力された音に応じて振動し、その動きに伴うリンパ液の流れが基底膜上のコルチ器の有毛細胞の毛を変形させることで細胞の電位の変化を生み出し、接続している神経細胞にインパルスを発生させる。内耳における基底膜は、分散性を持った伝送線路として働き、一方から入った音波は、基底膜上を伝播しながら指数関数的に伝播速度を低下させ、周波数に応じて決まる特定の位置で最大の振幅を示した後は急速に減衰して消滅する。従って、基底膜上には音声信号の周波数成分がほぼ周波数の対数に比例する位置に展開される。この展開は、外界の現象の時間的スケール（周波数）が物理的な寸法に対応することを利用している。また、小さな音でも聴き取ることのできる能力は、生物の生存にとって大きな価値がある。外から入ってくる音が小さな場合には、そのような音を聴き逃さないように、外有毛細胞に起因する機械的ポジティブフィードバックの量が増加して、感度が上昇する。この外有毛細胞による感度の上昇は、エネルギーの比では、10,000倍以上に達し、場合によっては、機械的な発振現象を生ずこともある。

心理学的なレベルにおいても、同様に能動的な機能が働いている。

正弦波を断続的に雑音で置き換えた音声を聴取すると、雑音に置き換えられてしまって存在しないはずの正弦波があたかも存在するかのように、連続した音として聴こえる現象が知られている。周波数が連続的に変化する正弦波を同様に断続的に雑音で置き換えた場合にも、欠落した部分が存在しているように聴こえる。

また、ピアノ曲で良く認められる現象であるが、1秒間に4～5回以上の速さで継続的に離れた音程の音符が提示される場合には、ある時点では1つの音しか出ていなくとも、複数の旋律に分離してそれらが同時に演奏されているように聴こえる。

これらは、時間的空間的に物理的性質の近い信号は同一の実体に起因するものであるという先入観（知識）が心理的機能として組み込まれており、信号の解釈を色づけしていることを意味している。

言語情報の処理でも、意識化される以前のレベルにおいて、同様な現象が認められる。

音声中の一部分の音韻を雑音によって置き換えても元の音韻が聴こえるという音韻修復の現象は、知覚レベルの類似の現象のような信号の物理的類似性からは説明出来ず、単語辞書や統語構造の制約に基づく予測・補完機能の存在を示唆する。

単語の中の特定の音韻を検出するまでの反応時間を調べた実験の結果、単語の後半部分では、意味のある単語を用いた場合の方が人工的に作成した無意味な単語を用いた場合よりも反応時間が短くなることが見い出されている。この現象は、音声を聞き取る場合に、単語の知識に基づく自動的な予測機構が働いていることを示唆する。

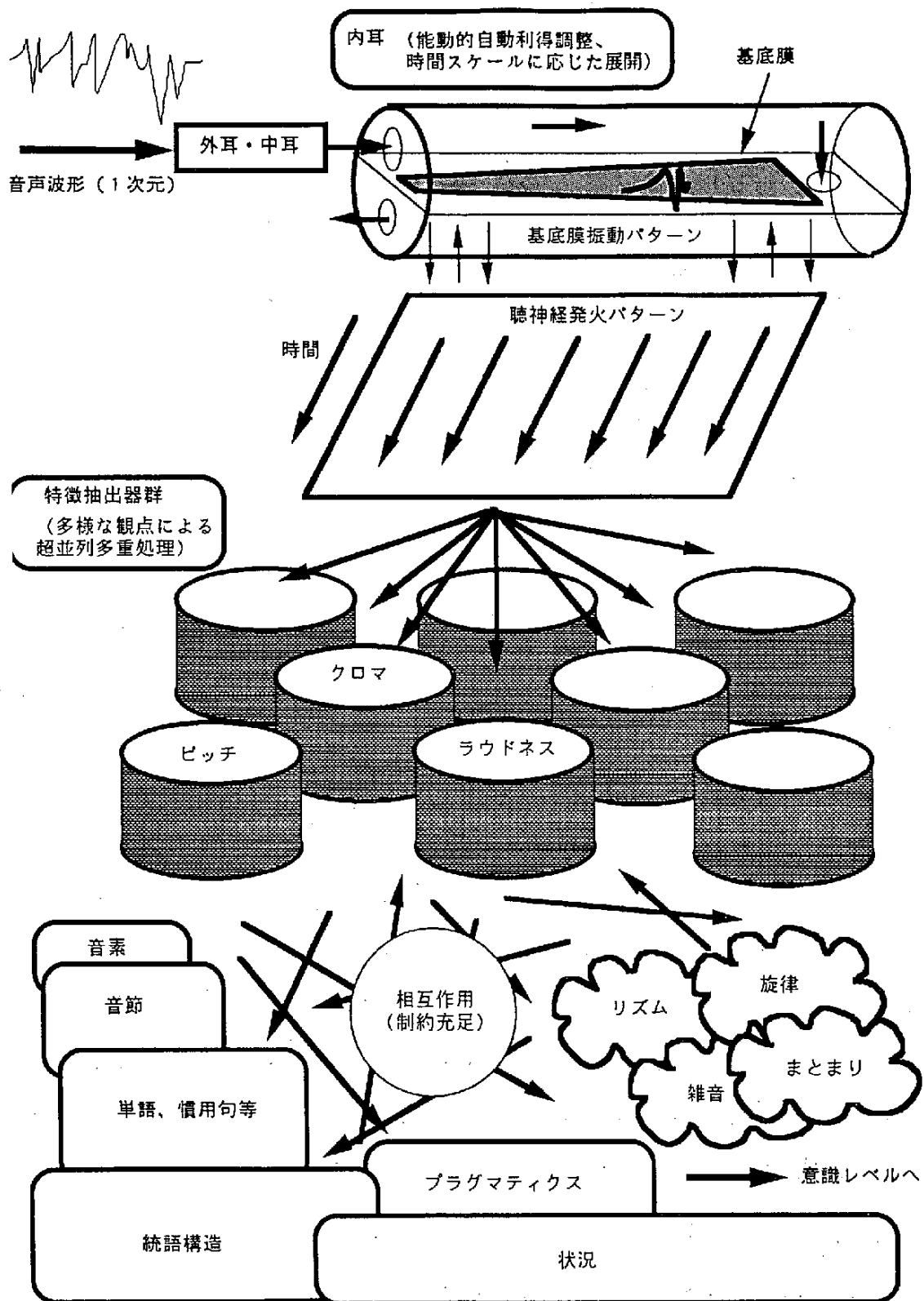
同様な実験で、検出すべき音韻を音節とした場合と音素（音節の構成要素）とした場合を比較した結果は、全体（音節）を認識する時間の方が部分（音素）を認識する時間よりも短いことを示している。これは、言語音声の認識過程が単純な階層構造モデルでは表せないことを意味しており、多様な要素的処理機能の相互作用の結果として認識・理解機能が実現されるという「新情報処理」モデルがこのレベルの処理の特徴を良くとらえていることを示唆している。

最終的には、音声からの情報の抽出には意識のレベルや他の感覚からの入力も動員される。これらの意識下の自律的処理による情報の抽出が破綻した場合には、自動的に意識レベルに対する注意が喚起され、聞き返しや、前後の文脈からの推論、相手の顔を見る（唇の動きを見ることにより言葉の聴こえ方が変化する現象が知られている）といった反応が起動される。

このように、聴覚という一つのモダリティにおける処理においても目的とする情報の抽出のためには、あらゆるレベルの制約条件が利用され、場合によっては、他のモダリティや意識レベルの処理までも動員され、能動的・統合的に情報の獲得が行われていることが分かる。一方、これらに対応する音声の機械処理においては、各レベルにおける知識の利用は進められているものの、基本的には論理的情報処理原理に立脚しており、階層設定の粒度、能動性・統合性は遥かに低いレベルにあるのが現状である。

【参考文献】

- [1] 勝木他：新版聴覚と音声、電子情報通信学会誌、（1982）
- [2] 難波編：聴覚ハンドブック、ナカニシヤ出版、（1984）
- [3] 入来、外山編：生理学1、文光堂、（1986）
- [4] Warren: Auditory Perception, Pergamon, (1982)



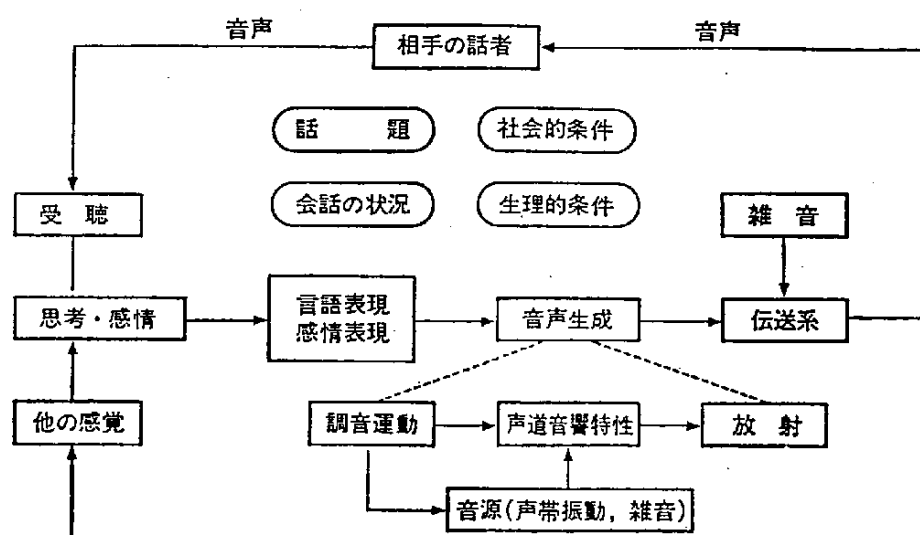
図表4.2-3 聴覚における音声処理過程のモデル

4.2.4 新しい音声情報処理

音声言語および文字言語に対する情報処理技術は、今後、情報処理機械が一層人間の生活全般に関係してくる中で極めて重要性の高い技術課題となっている。従来、音声認識は簡単なパターンマッチングによる単語認識しか実用にされておらず、コマンドの入力というイメージしかなかった。しかし、いうまでもなく言語は人間の通信と思考の根本的道具であるから、音声による対話の実現しなくては、音声情報処理を実現する意義は少ない。

これまでの音声認識、理解システムの研究も図表4.2-4のような音声対話の全プロセスを考慮しようと試みてきたのであるが、ある程度の限定を置いたとしても、人間との自由な音声会話を可能とする技術はまだ存在しない。

新音声情報処理で開発されるべき課題をあげてみる。



図表4.2-4 音声通信過程のモデル

(1) 聴覚に見られる高精度で柔軟な音響処理

音の時間・周波数パターンに対して、多様な時間スケール、周波数スケールで極めて広いダイナミックレンジをもった処理が行われている。音声情報の入口として、話者適応と雑音除去機能が重要であるが、この高度化のためには相当部分で聴覚機構と同等な音響処理の実現が必要となる。従来の前処理では、スペクトルの時間変化パターンを求めて一様に処理するのみであるが、音韻性の情報抽出には、全体のパターンの中で埋もれてしまいがちな小さな短時間のイベントが問題となる一方で、話者適応のためにはかなり長時間の音声からその話者性を抽出することが不可欠である。このような多様な処

理は、これまで主に演算量の制約から未開発に置かれている分野である。

(2) 音声対話の情報処理

まず、音声言語の言語情報処理が問題となる。文字言語に対する自然言語処理も多くの問題を残しているが、機械翻訳などを目的としてかなり研究が進められてきた。これに対して、音声言語では構文規則が強い制約条件を与えるとは限らず、意味レベルあるいは対話進行上の法則を合わせてとらえることが重要である。また、音声会話は本質的に相互に必要な十分な、矛盾のない情報交換をしながら進行するものではない性質を持っている。従って、常に不良設定問題を解かねばならない課題を負っている。もち論、音響処理の不完全さや雑音による影響も不良設定問題を生ずる大きな原因となっているが、本来、不十分で不正確な情報の中から、現在の状況に合わせた最大限の理解と対話行動が要求される。この意味では、音素、単語、文という階層構造を基本にとらえて音響処理と言語処理を分けて考えてきた従来の音声認識の枠組では無理があることになる。すなわち、このような枠組では、それぞれの知識源が実用に耐えるだけの大きさを持ったときには、当然、処理量が莫大になり實際上動作しなくなる。従って、音声対話においては、人間の思考レベルと同レベルの情報単位の処理を中心に置く、トップダウン処理が適することになる。しかし、この場合に、単純なトップダウンとボトムアップの融合的手法では問題は解決できず、確実性の高いキー情報から加速度的に推論する加速推論機構が必要である。

加速推論機構とは、制約充足型問題を解く具体的手段であって、大規模な不良設定問題において状況に応じて可能性の高いあるいはある種の評価関数の評価の高い解を、その解の根拠を与えるキー情報を探り出すことによって発見的に見出す機構である。このような情報処理は、単純な要素の並列動作だけでは無理があり、それぞれの部分で実行される処理内容に応じた部分構造とそれらを接続する結合構造の両方が問題である。そして、それぞれの構造には、様々な学習、適応機能が付与されている必要がある。

さらに、この問題のもう一つの特徴は、1 : 1 または 1 : n の複数の人間あるいは機械が関与する点で、実用上は比較的簡単な場合でも、人間行動の心理的、社会的側面も重要となる。

(3) 一般的知識の表現と利用

音声言語の情報処理の背景には可能な限り大きな一般的知識が存在し、それが音声理解上有効に利用されるべきことは今更言うまでもない。しかし、莫大な一般知識を一樣

な形で蓄積しておき検索する従来の方法では、実際の規模の問題については計算不可能になることは明らかである。状況に依存した、能動的な知識の利用と組立てが行われる機構として、加速推論機構は、この部分でも重要である。

(4) 知識の獲得、学習、データベース

現在の音声認識を初めとするパターン認識技術の中では、統計的手法は中心的役割を果たしている。しかし、構文規則のように、必ずしも統計的に学習されていないものもある。現状では、大きい実際のデータベースがあれば、そこから統計的に得られる知識ベースに基づいて統計的決定を行うことで性能の向上が得られている。上述した加速推論機構の中には、当然、知識の獲得のためのメカニズムが含まれていなくてはならない。しかも、それは単純な確率の学習では不十分で、小数データあるいは組合せデータに基づく効果的学習方法が必要である。

学習機構の重要性は、人間も会話できるようになるまでに豊富な経験を積む必要があることを考えれば当然なことである。ここでは、適切な学習機構と十分なデータベースが問題ということになるが、最後に重要なことは、徹底した学習を積んで完成された機械を作れば問題解決ということではないことである。望まれる機械は多様であって、様々な性格を持ちしかも発展変化することが本質的に要求されている。

(5) 音声中の個性、感情の理解と表現

従来、音声理解の研究では音声中の言語表現の要素が中心に扱われてきたが、音声通信の役割として、感情表現や個人性も大きな部分を占めている。感情表現には、文の形式、韻律だけでなく声色や身振りなどが重要な要素で、それらの総合的な処理が不可欠となる。

4.3 感性・文化・応用

4.3.1 感性情報処理

機器の発展過程は、まず必要な機能の実現の段階から、使いやすさ、見てくれのよさ、個別ユーザへの適合性へと進むのが普通である。

情報機器も今の機械中心の研究開発の段階から、上述の意味での良質な機器開発の段階へと移行しつつある。感性情報処理はこのような良質の機器開発にかかわる新しい技術の開発を目ざすものであり、今後の情報処理機器のコストの大部分は、付加価値としての、

感性情報処理のために費やされるものと予想されている。

感性情報処理とは、ユーザの気持ちを察知するための新しいパターン情報処理技術と、ユーザに楽しく使ってもらうためのパターン情報生成とに大別できる。その具体例を以下にあげる。

視覚に関しては、ユーザの表情からその喜怒哀楽を認識したり、コンピュータの向う側にいる相手の表情を生成したりする技術についての研究が始まっている。またコンピュータグラフィックスの分野で、美しい形や色を持った製品をデザインする研究も行われている。これに関連して、達筆になるための書道教育システムの提案もなされている。

音声に関しては、ユーザの声の表情の認識や声の美しさの判別と、心のやすらぐ声や、励ましの声、ソプラノやテナーの声の生成、作曲などが、この分野のテーマとして考えられよう。このうち、声を通して、怒っているとか笑っているとかを認識する予備的な研究はなされているようである。

自然言語に関しては、文を解析して係り受けなどから分かりやすい構造であるか否かを判定する研究が行われている。さらに進んで、一連の文からなる文章を解析して、筆者の意図を理解する試みもなされている。この逆に分かりやすい文を組合せて、明確に意図を伝えるための文章の構成法、さらに高度なものとして、文学作品の分析や創作などをおこなうこともできるであろう。

ユーザの個性に合わせた機器の開発というのも、重要な課題である。ワープロで同音異義語の優先順位の決定法は、個性化のごく初歩的な例である。

4.3.2 センサ融合

センサ融合とは複数の異種センサからの情報を融合して必要な情報を信頼性高く抽出する技術を言っており、特に最近、重要視されるようになってきている。例えば、触覚センサとイメージセンサを備えたマニピュレータ、各種のイメージセンサを備えた移動ロボットなどで用いられており、自律型のロボットの実現に対しては重要な要素となる[1]。人間の場合、五感（視覚、聴覚、触覚、味覚、嗅覚）など利用できるあらゆる感覚機能を活用して環境の把握に努めたり、様々な情報の抽出・伝達に利用している。センサ融合はこのような人間の持つ感覚統合機能を工学的に実現しようとするものであるが、まだ研究が開始されたばかりといえる状況にあり、その技術の確立が望まれる。

センサ融合の利点としては、それぞれのセンサからの不確実な情報を集めて総合的に信

頼度を上げたり、相補的な情報を組合せて、今まで得られなかった特徴が抽出できる[2]などが挙げられる。認識・理解そのものが不完全な特徴に基づいて、多くの弱い制約条件下での解を求めるプロセスであり、探索空間の爆発が生じるような問題になっている。人間は、予測し仮説を立てる作業を無意識的に行っていて、これにより強力な知覚・認識機能を持たせている[5]。センサ融合においても、このような能動的な観点から、世界モデルに基づいて適当なセンサ群を選択する制御機構をもったもの[3]、センサ融合を論理センサのモジュールとして扱うものなど、一般的な枠組みの提案もある[4]。

しかし、具体的な問題に対しては、いろんな意味でかなり制限された状況の中でその成果を発揮できている。各種のセンサモジュール間での相互作用を通してどのような反応が得られるかは、今後の大きな課題であり期待されるところでもある。特に、センサ融合では、システムの動きを含めて、あらかじめ、あらゆる環境変化を完全に想定できなかったり、そのモデルのインプリメントに膨大な時間や作業が必要となることが考えられる。従って、より一般化された柔らかい体系[6]や自律学習型の融通の効くシステムの実現が一層待望される。

センサ融合により、例えば、直感的な部分や第6感的な判断力、識別能力、エキスパート的な判断機能などを実現できる可能性がある。人間の場合、

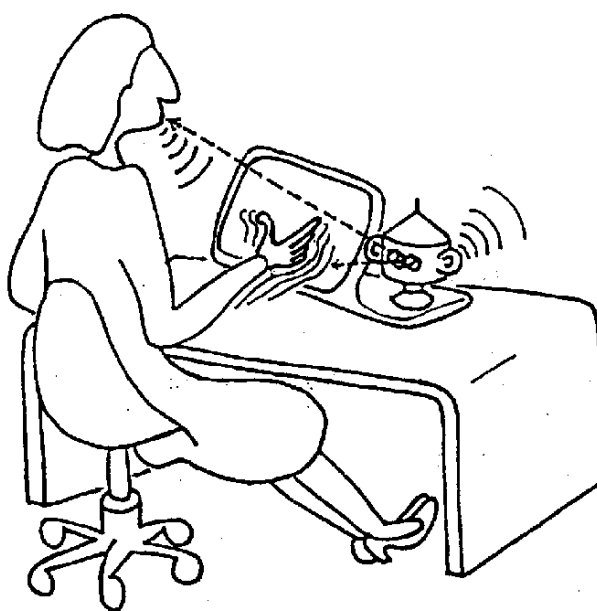
- ・あまり意識しないで、ノブを回してドアを開けたり、鍵をかけたり、
- ・触覚や視覚をフルに使ってバランスをとったり、
- ・耳、目、手や指の感覚を使って、楽器を演奏したり、
- ・色艶、触った感覚、臭い、味などから、腐ったものかどうかを判断したり、
- ・声、顔形、外見、筆跡、指紋、匂いなどから、だれなのかを見分けたり、
- ・野球で、打った音と、打者のスイングなどから飛ぶ方向・距離を推測したり、
- ・顔の表情、声の調子、仕草などから感情を推し量ったり、
- ・匂い、音楽、周りの景色などから、昔のことを思い出したり、
- ・雰囲気という言葉で代表されるような、微妙な状況を感じたり、

などといった、極めて高級なことをやってのけている。このようなことの一部、または、部分的な機能でもシステムにインプリメントできたら、どれくらいシステムの融通性、守備範囲が広がることであろうか。

システムが利用者の表情、仕草、身振り、手振り、声の調子などから総合的に解釈して、利用者の感情的な側面や意図を推し量れるのであれば、利用者の意図を別途パラレルにシ

システムに知らせるために利用できる。ボディアクションの解釈はシステムにリアルタイムな反応のチャンネルを広げる。自動ディクテーションシステムや国際会議などの議事記録システムや自動通訳システム、通訳電話など、言語音声を扱うものを考えた場合、伝えたい情報は、言語情報だけではなく、感情を含めたニュアンスも伝えたいし、また、言葉では表現しにくい感情や大きさや程度などの情報も同時に伝えているからである。教育システムでは、対話のチャンネルが豊富になり、例えば、子供達に対して彼らの反応に応じて、感情を含めて表現・応答ができ、適切な指導を行うのに有用である。ゲームでは、音声とアクションを取り入れた臨場感のあるスケールの大きいゲームが考えられる。

人間の感性とは、多くの知覚チャンネルをうまく使って、より多くの信頼性の高い情報を融通のある形で収集・分析できる感受性・能力を指していると考えられる訳であり、システムに感性という多様で柔軟な統合能力を与えるためには、この機能のインプリメントは必須であると言える。



図表4.3-1 センサ融合

【参考文献】

- [1] 石川正俊：“センサーフュージョンシステムー感覚情報統合のメカニズムー”，
日本ロボット学会誌，Vol.6, No.3, pp.251-255 (Jun.1988)。
- [2] 浅田稔：“センサ統合と環境モデルの構築”，信学技報，PRU89-101, pp.1-8,
(1990)

- [3] R. C. Luo and M. K. Kay: "Multisensor Integration and Fusion in Intelligent Systems", IEEE Trans. Syst. Man Cybern., Vol.SMC-19, No.5, pp.901-931、(1989)
- [4] T. Henderson and E. Shilcrat: "Logical Sensor Systems", J. Robot. Syst., Vol.1, No.2, pp.169-193, (1984)
- [5] 「新形態情報処理技術の動向調査報告書ーインテリジェントビジョンの動向ー」昭和62年度調査資料（日本電子工業振興協会）.
- [6] 大津展之: 認識と理解のための柔らかな論理, 電子情報通信学会誌, Vol.71, No.11, pp.1231-1240 (Nov.1988) .

4.3.3 ヒューマンインタフェース

(1) 概 要

ヒューマンインタフェースの目的は、ユーザは機器を解釈者として情報を送り、機器はユーザを解釈者として情報を送り返すことにより、ユーザの生産性と満足を最大にするよう人間の知的活動や問題解決を支援することである。

80年代以降パソコン、ワークステーションなど情報機器が急速に増加するとともにその機能が高度化されてきた。オフィスに限らず一般家庭や公共の場所においても広く使用され始め、今後熟練者のほかコンピュータに不慣れなユーザの使用が急増することが予想される。ユーザにとっては、誰もが容易に使用できる『柔らかいヒューマンインタフェース』が求められる。またシーズ側から見ても、情報機器のソフトウェアリソースに占めるヒューマンインタフェース部が増大し、その役割内容も大きく変わるであろう。

現状の情報機器のヒューマンインタフェースは

- a) 専門家の効率化を目指してマルチウィンドやハイパーテキスト,
- b) 誰もが容易に使用できるものを目指してアイコン、メタファ、オブジェクト指向の2つの方法が融合されたものが多くなっている。これらは視覚的であること、現実世界のモノと対応が付く『Look & Feel』のインタフェースが特長となっている。これらを支える基本概念には『直接操作性』[1]がある。例えばユーザが文書を作成する場合、コンピュータを使うのはプログラムを作成するためではなく、自分の意図する仕事を実行するためである。この時、ペンや消しゴムのメタファを表示し、文房具の感覚で作業を進める。すなわち現実世界の道具と同様な操作ができ、使いやすくなる。

今後は人間の思考に沿った自然なヒューマンインタフェースが求められている。すなわち人間の持つ心の原理や特性に沿った人間中心のヒューマンインタフェースにより生産性の効率化や使いやすさが求められる。さらには『使用感』『美しさ』など感覚、感性レベルまで含めた使いやすさや『楽しい』など感情レベルのヒューマンインタフェースが期待される。

なお視覚的なインタフェースに限らず、音声、音響、言語、画像メディアなど、人間の現実世界の情報メディアと同様、情報メディアの融合利用が自然なヒューマンインタフェースとして欠かせないであろう。

(2) 技術の現状と展望

ヒューマンインタフェースに関して以下の技術について考える。

- a) 人間とコンピュータの問題解決の場としてのヒューマンインタフェース
 - b) 認知的側面からのヒューマンインタフェース
 - c) 直接操作性（ダイレクトマニピュレーション）とメタファ
 - d) 自然なマルチメディアの理解、表現
- a) 問題解決者（ユーザ）とコンピュータとの処理の流れは図表4.3-2 に示す①～⑤の5つの階層に整理される[2]。本来の問題解決は対象に対する知覚、運動だけでなく、ユーザの思考が重要となる。④、⑤は知覚運動操作の支援、③は最小限の作業支援、①、②は現在研究の中心的なテーマになりつつあるユーザモデル、システムモデル（後述）、あるいは適応型のインタフェースである。解決すべき課題すなわち注目点が遷移し、これに従って①～⑤の階層を遷移する。コンピュータは課題の把握と対話の教示を実行しながら適応しようとする。しかしユーザのレベル、タスクの種類など動的に変化するので問題を難しくしている。今後は状況に適応するヒューマンインタフェースが必須であろう。

図表4.3-2 インタフェースの階層構造

	計算機システム	←情報の流れ→	ユーザ
①	知的なシステム	←曖昧なアイディア← →助言・補佐→	問題生成・合成
②	認知的情報処理	←問題領域← →利用可能ツール→	問題の規定
③	定型的ソフトウェア	←データ・指図← →回 答→	問題解放・求解
④	データ処理	←コマンド群← →メッセージ→	思考的操作
⑤	デバイス・装置	←打鍵操作など← →状態表示→	機械的操作

- b) ユーザはコンピュータシステム概念に対してメンタルモデルを持つ。またデザイナーはコンピュータシステム概念に対してデザインモデルを持つ。ユーザのメンタルモデルは物理的システムやマニュアルが表すシステムイメージとのインタラクションを通じて形成される[3]。デザイナーの設計はユーザに分かりやすいシステムイメージを形成することが重要である。しかし今後、1つのシステムを複数のユーザが使用した場合や、システムの巨大化にともない複数のデザイナーが設計した場合など、メンタルモデルとデザインモデルが大きく異なる問題が発生するであろう。この問題を解決する方法としてユーザ適応のヒューマンインタフェースが必須であろう。
- c) 従来のコマンド中心のインタフェースからオブジェクト指向のヒューマンインタフェースに変わりつつある。画面に意味のある操作のアイコンを表示し、現実世界のモノに対応つける。『はさみ』は切るもの、『紙』は文書を示し、『はさみ』を『紙』に重ねれば、『はさみで文書を切る』即ち『文書を切り抜いて編集する操作』を意味する。しかし抽象対象や複雑な操作を的確に表すことが難しくなること、状況、時代、個人によってアイコンの持つ意味が変化することなどの課題がある。最近デスクトップメタファを更に拡大し、作業環境メタファをオフィスにまで広げた興味深い検討が行われている[4]。作業環境メタファは広くは状況の制約と見なすことができる。

d) 現実世界で用いているメディアを使うことにより、最も自然なヒューマンインタフェースを実現することができる。まず音声メディアは日常会話では最も自然である。しかし認識の難しさに加え、表現の曖昧さ、文法の不整合、省略文が多く、文脈依存からの理解が必須となっている。言語メディアはコマンドを学習する必要もなく即座に使用できる。しかし音声メディアの理解と同様、表現の曖昧さ、話題の焦点抽出などの課題がある。画像メディアについては、人間は視覚の認識には非常にすぐれていると言われている。音声、言語のようにその体系を学ぶ必要もなく簡単に意味することが理解できる。今後は各メディアの利用技術の性能向上を行うことにより、その特長を生かした自然なコミュニケーションが期待される。

(3) 応 用

状況理解や制約などにより、コミュニケーションにおける曖昧性、多義性の解消、省略部分の補間、部分的誤りの修正などが期待できる。

ユーザとコンピュータとのコミュニケーションの一例を以下に紹介する。

.....

ユーザ : 『例のものの頼むよ』

コンピュータ : 『例のもののと言いますと、昨日行われた会議の議事録ですね』

ユーザ : 『そうだ』

コンピュータ : 『かしこまりました』

昨日の議事録が、プリントアウトされてくる。

.....

と言うような曖昧性を含んだ対話で自然なコミュニケーションが行えるであろう。

【参考文献】

- [1] Shneiderman, B.: Direct manipulation: A step beyond programming languages, IEEE Computer, Vol. 16, NO. 8, pp. 57-69 (1983)
- [2] Thesen (1986) : 人工知能ハンドブック (人工知能学会編) : pp. 751 (1990)
- [3] Norman, D. A. and Draper, S. W. : USER CENTERED SYSTEM DESIGN, Lawrence Erlbaum Associates, pp. 45-48 (1986), (邦訳, インターフェースの科学, pp. 28-29 (1987))
- [4] 市川、平川 : HI-VISUAL'89、第3回公開シンポジウム知的ヒューマンインタフェース、PP. 27-38 (1990)

4.3.4 コンピュータビジョン

コンピュータビジョン研究の中で生産工程の自動化、図形のベクトル化等の応用は実用例が多い。これらは一定の照明条件対象の限定がやりやすいものが多く、そこでの画像処理も雑音除去・特徴抽出も比較的実現しやすく、かなり人間に近い、あるいは人間より勝る性能が達成されている。しかし、一般には計算機は依然として画像情報から環境を理解できる段階ではない。コンピュータビジョンへの期待はますます増大しており、技術の進展が望まれている。

認識・理解の面から見たコンピュータビジョンの研究は、基本的には環境理解技術[1]である。環境は幾つかの対象の集合体からなっており、この技術は情景や対象（候補）の空間的配置などから対象物のモデルを想定し、これを用いて対象物を（能動的に）理解するものである。そのための開発技術は知識モデルの設定と画像解釈アルゴリズムとなろう。これに対して、認識対象の記述・モデル化、環境表現技術、情報入力技術、画像認識技術、認識のための推論・学習技術、並列分散アーキテクチャ技術などが課題として上げられている。これらの課題の多くは、パターン情報を扱うことに起因するもので、いずれも3.2項で触れている各概念と深いかわりがある。特に群情報処理が大きなインパクトをもたらすと思われる事項について以下に個別に説明する。

(1) 認識モデル

単純な画像認識では、センサから入力された信号は、まず前処理・特徴抽出を施される。次にそれらの特徴群を使って個々の対象物が認識され、最終的に環境が理解されていく。通常この流れをボトムアップ処理と呼んでいる。しかし環境は、前述のように対象物の複合から成り立っている。そのため、このような一連のボトムアップ処理のみで理解できることはほとんど無い。候補となる対象の存在をあらかじめ、または処理の過程で想定し、その想定の良い否をトップダウンで検証する処理と、ボトムアップ処理とを混ぜて使用するのが普通である。このトップダウンとボトムアップ処理をいかに融合させるかの方法に関する研究開発が今後ますます必要である。パターン情報が本来持っている多義性などのために、各種検証はあいまいさや、不確かさ、不完全性が付きまとう。このため、本融合方式は柔らかな論理に基づく必要がある。

また、この際に使用されるいろいろな知識に関して、どのような知識を与えれば問題が解けるのか、また、個々の解決すべき問題と画像が与えられた場合、持っている知識群のどれを具体的に選択すべきかが問題となる。まずこれを実現する手法を明らかにし

ていくことが重要であろう。

(2) 前処理・特徴抽出等の単体処理機能

上記モデルの探索と共に、一連の処理の流れの各階層での、新機能の実現や既存機能の向上も欠かせない。まづ前処理についてであるが、この中には2値化などの閾値処理、画質の改善、幾何学補正や濃度補正、強調・雑音除去・平滑化などが挙げられる。これらはいずれも実用段階にある。しかし実際には、それぞれ複数個存在する手法のうち、どれを選択するか、あるいは閾値などのパラメータをどうするかは、まだ担当者に任されており、専門的な経験を要求される場合が多い。これを自動化する研究も進んでいるが、今後学習によって自動的にこれらの機能選択やパラメータ設定が行えるシステムが必要となろう。

次に前処理された情報は、その後に特徴抽出・解析が行われる。この中には、マッチング、セグメンテーション、幾何学的特徴抽出、テクスチャ抽出などが含まれる。これらの機能は生体では簡単に実現されているものであるが、その機能を参考にし、より精度の高い、かつ高速（実時間処理）なパターン特徴抽出能力が求められている。このためには並列処理が必須となるとともに、さらに一般に機能の高い機能ほど、具体的なパラメータ指定が難しくなるので、サンプル例示によって自動的に機能が構成されるような自己組織能力が要請される。さらに特徴抽出に続く段階の処理として、抽出された対象画像領域の輪郭、濃淡変化、テクスチャ変化などから、対象の3次元的形状や、動きを判定する機能も求められる。

これらの機能の実現、高度化のほかに、各種機能を手軽に利用できるためのモジュール化や、情報表現の統一などが重要課題である。さらには機能間の自由な組合せを可能とする（オブジェクト指向などの）プログラムパラダイムの導入が必要となる。

このように、機能の高速化のための並列処理、複数の機能を同時並列的に実行するための分散処理機構、学習によるパラメータ等の自動修得によるユーザに使いやすい機能など新情報処理機構への期待は大きい。

(3) 感覚統合

人は五感を通じて総合的に外界を把握している。計算機も、複数のセンサの情報を複合・融合し、必要情報をより信頼度高く取得する機能が開発されれば、外界の把握能力が飛躍的に向上するものと期待できる。具体的に、生体では、動き、形状、テクスチャ、色、空間配置などの感覚機能の統合が並列的に実行されていることが分かってきており、

これらの多次元情報処理機能の工学的実現が期待される。技術的には視覚・聴覚・触覚など各種のセンシング技術のモジュール化と融合化の研究が始められており、今後、さらに異種情報間の照合、2つの耳、両眼のように複数の聴覚や視覚センサから同時に入ってくる情報を同時並列的に処理できる機能の進展が望まれる。また、この感覚統合によって、知覚のアクティブな活用も可能となろう。

(4) モデルベース認識

人は多くの経験を通じて、脳の中に多くのモデルを築いていると考えられる。環境を理解するとき、それらが有機的につながり合わされて使用されるものと考えられる。しかしこれについては、実際に脳の中でどのような記述の方法があるのか、どれくらいの量のモデルが必要か、どのように更新されていくのかなどがまだ分かっていない。これら生体機能の解明は別途期待するとして、モデルの記憶性、モデルの想起・連想法、モデルとのマッチング等に新しい情報処理機構の活用が望まれる。例えば、相互結合型ニューラルネットがモデルとの対応関係を高速に解いた研究例も現れてきている。例えば3次元世界の把握、隠れた点の対応関係などの探求が高速化され、処理時間の上で事実上不可能だった動画像理解も徐々に可能となるであろう。

以上述べた要素技術の実現によって、今までは単一の情報としてしか利用できなかったパターン情報が、計算機で内容が理解されるようになり、コンピュータビジョンのより広範な利用が図られることが期待される。

4.3.5 ロボットビジョン（移動ロボット）

(1) 視覚の担う機能

極限作業ロボットが原子炉の中で保守作業をする。無人工場の中をロボットが走りまわり製造作業を行う。農作業ロボットが農場を動きまわる。自動操縦車が長距離長時間走行で活躍する。これらのロボットが与えられた指示に従って移動するためには、ロボットは、視覚を使って周囲環境をセンシングし、周りの状況がどうなっているかを知らねばならない。すなわち、移動ロボットは、センシングに基づいて

- ・自分がどこにいるか、またどちらを向いているかを知る（自己定位）
- ・接近すべき目標がどこにあるか知る（目標物認識、定位）
- ・進路上に障害物（穴、立体物、接近物等）はないかを知る（安全確認）

を実現する機能を備えていなければならない。現在、用途に応じて各種のロボットビジ

ジョンの研究開発が行われている。しかし、これらは使用環境への制約が大きい上、用途向けに特化したいわば単能ビジョンの段階にあり、広範に使用できる視覚の実現に向けて飛躍的な技術レベルの向上が望まれる。ここでは、ビジョンについての現状と将来展望を主な環境を屋内、屋外に分けて述べる。

(2) 自己定位

環境に既存の、特定のものを手がかりとして、自分の現在位置を割り出すことである。ロボットが内蔵する、地図に記載されている手がかりの位置と、視覚で見た手がかりの位置とをもとに算出する。以下の研究例がある。

① 屋 内

- ・画像処理のエッジ抽出手法を利用して、既定通路を他から区分している白線を検出したり[1]、通路である廊下の床と壁の境界線を明暗差から検出したり[2]することにより、通路に対する自分の位置・方向を知るもの
- ・廊下内の移動に伴うオプティカルフローを利用して、位置と動きを知るもの[3]
- ・廊下の壁面に存在する、色で区別される各種形状の面を手がかりとして、色画像処理により抽出した上記領域の形状・位置を利用して、廊下の中での定位を行うもの[4]
- ・単純な形状をした、いわば積木の世界のような、実験環境（窓、ドアなどのある部屋、階段、廊下）と、これを線画で記述した環境モデルを用意し、画像処理により得た線画と、環境モデルとのマッチングにより定位を行うもの[5]

② 屋 外

- ・画像処理のエッジ抽出手法を利用して、道路の境界線やレーンマークを検出し、道路に対する定位を行うもの[6]
- ・走行すべき道路を、画像の色やテクスチャの差異に基づいて、道路外領域から分離し、得られた領域の形や位置などから道路に対する定位を行うもの[7]
- ・ステレオ画像処理により路側のガードレールを検出し、これに対する定位を行うもの[8]

図表4.3-3 は、これらのロボットの視覚を、何を手がかりにどう定位しているか、という観点から整理したものである。これらの研究例に共通するのは、

- ・廊下や道路などの環境ごとに、定位のために使えるような手がかりを、まず、決め
- ・その検出に有効であるような視覚（画像）処理アルゴリズム、ハードシステムを

開発する
という方法である。

図表4.3-3 自己定位用視覚の現状

手がかり	手がかりの抽出法	位置の求め方	適用環境	課題
構造物の輪郭線	エッジ抽出 →線画	線画のモデル マッチング	画像処理に よる整備は室内	一般環境へ の適用
構造物輪郭線 の特徴点の動き	エッジ抽出 →特徴点	オプティカルフロー（特 徴点の動き）	廊下	
通路、道路の 路側線	エッジ抽出	エッジ線の位置 傾きから算出	廊下、道路	各種環境下 での、手が かりの確実 な抽出
路面、壁面の 色やきめ	色、テクスチャによる 領域分割	特定領域の形 位置から算出		
ガードレール	ステレオ立体視線	ガードレールの位置 傾きから算出	道路	

この方法で開発を進めてゆくと、特定の環境に特化した、その環境では強力な視覚が
でき上がる。しかし、今後、移動ロボットの用途を拡大していくためには、広範な環境
を対象としなければならない。特定環境では有効であった手がかりがいつも使えるとは限
らない。今後望まれるのは、複数の手がかりを利用し、広範な環境で使用できる視覚を
開発することであろう。そのためには、(a)個々の手がかりを確実に抽出する、(b)複数の
手がかりを融合して使う、ことが重要課題となる。

(3) 目標物定位

ロボットの周りに存在するものの中から、特定のもの（目標物）を発見し、その位置
を割り出すことである。目標物のモデルや位置が、ロボットが内蔵する地図に記載され
ている場合と、記載されていない場合とがある。記載されている場合は、視覚から得た
情報と地図の情報との照合により、目標物の位置を算出する。記載されていない場合に
は、視野の中から目標物を探し出し、位置を算出する。以下の研究例がある。

① 屋 内

- ・原子炉内のバルブのコックの位置を、ロボット内蔵の線画モデルと、画像処理によ
り得た線画とのマッチングにより推定するもの[1]

② 屋 外

- ・果物収穫作業では、目標物である果物の位置はあらかじめ分かっている。視野の中から色などを手がかりに果物をさがし、カメラを移動させた時の2枚の画像から、立体視法によりその位置を推定するもの[9]
- ・撮像した目標物の輪郭に複数のウィンドウを固定することにより、自分も目標物もともに動いている場合でも、目標物の位置を算出するもの[6]

移動ロボットは、単に移動するだけでなく、目標物を対象に作業を行う。今後、移動ロボットの用途を拡大していくためには、多種類の目標物を対象にする必要がある。位置があらかじめ分かっている（地図に記載されている）かどうかによって目標物を分類すると、以下のようになる。

- (a)位置は固定しており地図に記載されている …………… 機械装置、ドア、コック
- (b)存在領域が地図に記載されている …………… 机の上の本、収穫する果物
- (c)動きまわる目標のため地図に記載なし …………… 追跡すべき前方の自動車

(a)、(b)の目標の定位には、画像理解や図表4.3-2の最初の視覚などと共通の、モデルマッチング手法が有効であろう。しかし、ロボットが移動すると目標物の見え方は変化する。このため、静止画像を対象にした従来のモデルマッチングではなく、いわば動的なモデルマッチング法が今後重要になる。 (c)の目標の定位は、移動ロボット固有の問題であり、動的なモデルマッチングがキーになる。動きの予測に基づく、トップダウン的な見え方の予測、物体認識が必須である。

(4) 安全確認

既定の進路（廊下、道路）上に存在する障害物の発見や進路上へ侵入が予想される物の発見、また、ごちゃごちゃと物が置いてある場所や部屋での安全な通路の発見である。以下の様な研究例がある。

① 屋 内

- ・明るく照明された廊下で、前方廊下画像を明暗に基づいて領域分割し、得られた結果から、障害物の検出とそこまでの測距を行うもの[12]
- ・ステレオ立体視法により前方通路上の障害物の検出と、そこまでの測距を行うもの[1]
- ・ソナーと光センサの出力を融合することで、部屋の中の立体物やドアの位置を正確に求めるもの[13]

② 屋 外

- ・前方道路画像を、色や明暗などに基づいて分割して得た領域から、停止車両の存在とそこまでの距離を推定するもの[10]
- ・前方道路画像からエッジ抽出することにより、直方体の静止障害物の輪郭線、即ち障害物、を発見し、そこまでの距離を推定するもの[6]
- ・前方道路上に存在する静止立体物を、ステレオ立体視法により検出し、そこまでの距離を測るもの[8]、[11]
- ・前方道路上の立体物の検出と、そこまでの測距にレーザレンジファインダを用いるもの[7]

移動ロボットの実用のためには、安全確認技術の完成は必須である。基本は、進路上の障害物あるいは進路上への進入が予測されるものの確実な検出と、その位置の推定である。現在は、進路上の静止障害物を、特定の単一手がかりに基づいて検出している。しかし、多様な環境の中で確実に障害物を検出するためには、①単一手がかりを確実に抽出した上で、②複数の手がかりを融合して使うことが重要になる。移動障害物については、動きの予測に基づく、トップダウン的な検出が重要になる。

(5) 期待する技術

移動ロボットを広範な分野で使用するためには、多様な環境の中で有効な視覚が必須である。その実現には、複数の手がかりの融合使用（感覚融合）と、予測に基づいたトップダウン的な認識・位置推定能力（能動的理解）の付加が重要になってくる。これと共に、地図ないし、環境モデルは、3次元空間や移動物体などが、容易に表現できるものにならなければならない。さらに、「黒い影が動いた。ハッとして止まってよく見たら犬だった。」という体験から推測されるように、環境モデルは、パターンで表現されパターンのまま判断・行動に至る情報の流れと、パターンから記号に変換され認識・理解に至る情報の流れの両方に深くかかわっている。パターンと記号の統合が重要になる。

【参考文献】

- [1] 穂積、他：“視覚情報を用いた移動経路生成および移動実験”、第7回日本ロボット学会学術講演会、3102、(1989)
- [2] 恩田、他：“移動ロボットのための廊下の3次元環境認識”、信学論(D)、J70-D、12、PP. 2510～2516、(1987)

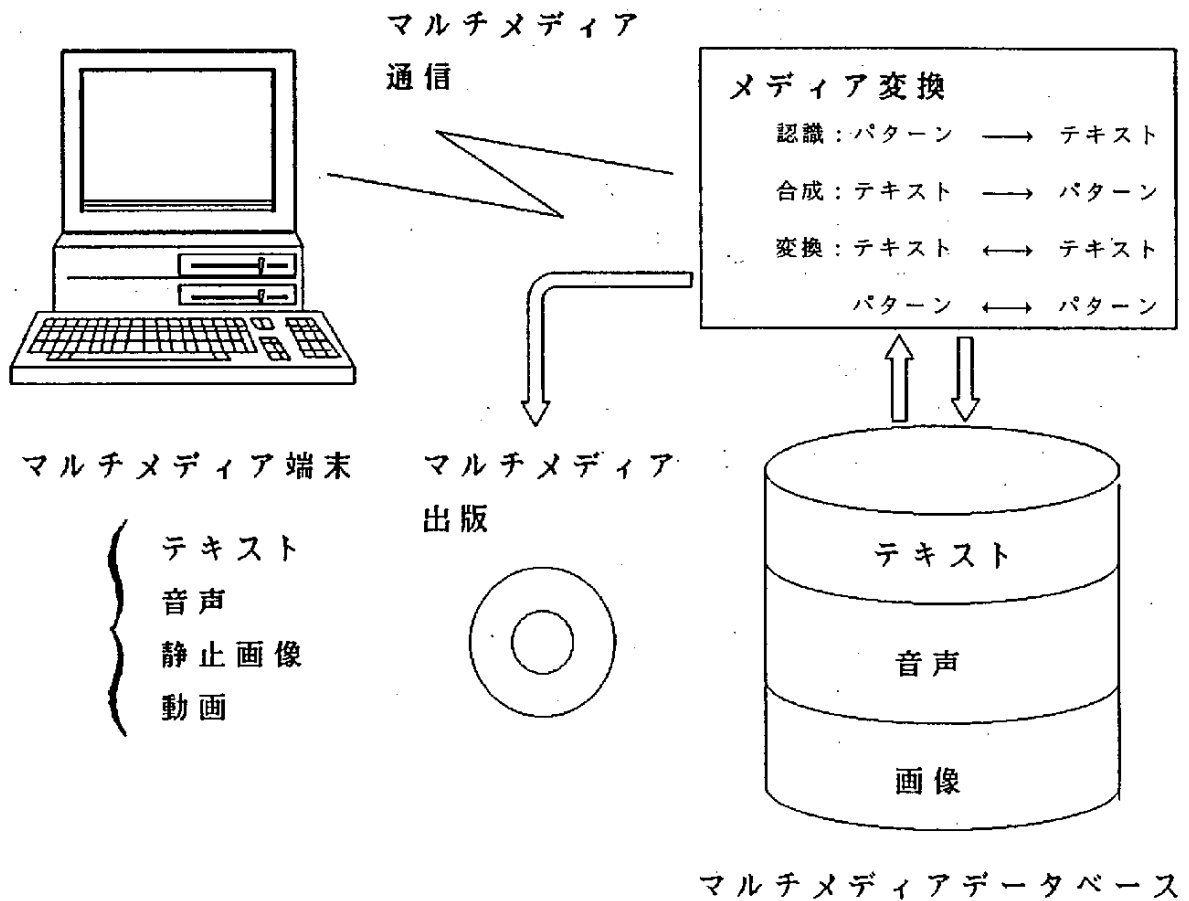
- [3] 八木、他：“動画像処理を用いた移動ロボットの環境認識”、信学論(D)、J69-D, 6, PP. 967~974, (1986)
- [4] 坪内孝司：“屋内環境を移動するロボットの視覚システムに関する研究”、筑波大学大学院博士課程工学研究科博士論文, (1987, 12)
- [5] 亀島、他：“再帰型画像処理機構を用いた移動ロボットの視覚誘導”、日本ロボット学会誌、vol. 5 No.5, PP. 343-350, (1987)
- [6] E. D. Dickmanns: “4D-dynamic scene analysis with integral spacio-temporal models”, Proceedings 4-th ISRR, Santa Cruz, Aug. 9-14, (1987)
- [7] Charles Thorpe, et. al: “Vision and Navigation for the Carnegie-Mellon Navlab”, IEEE, Trans. on Pattern Analysis and Machine Intell. vol. 10, No. 3, PP. 362-373, (1988)
- [8] 谷田部照男：“知能自動車”、自動化技術、第13巻第10号, PP. 22-28
- [9] 川村、並川：“農業部門におけるロボット”、日本ロボット学会誌、vol. 4 No.2, PP. 126-130, (1986)
- [10] 森、他：“構内道路を走行する自律移動ロボット”、日本ロボット学会誌、vol. 5, No.5, PP. 361-374, (1987)
- [11] 越正毅監修：“車が変わる交通が変わる”、PP. 203-205, 日刊工業新聞社
- [12] 小林、他：“自律走行実験車の開発”、第7回日本ロボット学会学術講演会、3101, (1989)
- [13] A. M. Flynn : “Combining Sonar and Infrared Sensors for Mobile Robot Navigation”、The International Journal of Robotics Research, vol. 7, No. 6 PP. 5 ~14, (1988)

4.3.6 OA (マルチメディア、ハイパーメディア)

事務処理でのコンピュータ利用が勘定系だけでなく、文書系にも発展しているため、数値・文字データ、文章、文書画像、グラフ、動画、音声といった、多種多様の情報媒体(メディア)を統一的に扱うマルチメディア情報処理が要求されている。

マルチメディア情報処理を実現するための技術的背景は、光ディスクなどの安価で大容量の記憶媒体の出現、光ファイバによる通信路容量の増大、ワークステーションの能力向上などがあげられる。特に、CD-ROMのマルチメディア出版における意義は大きい。

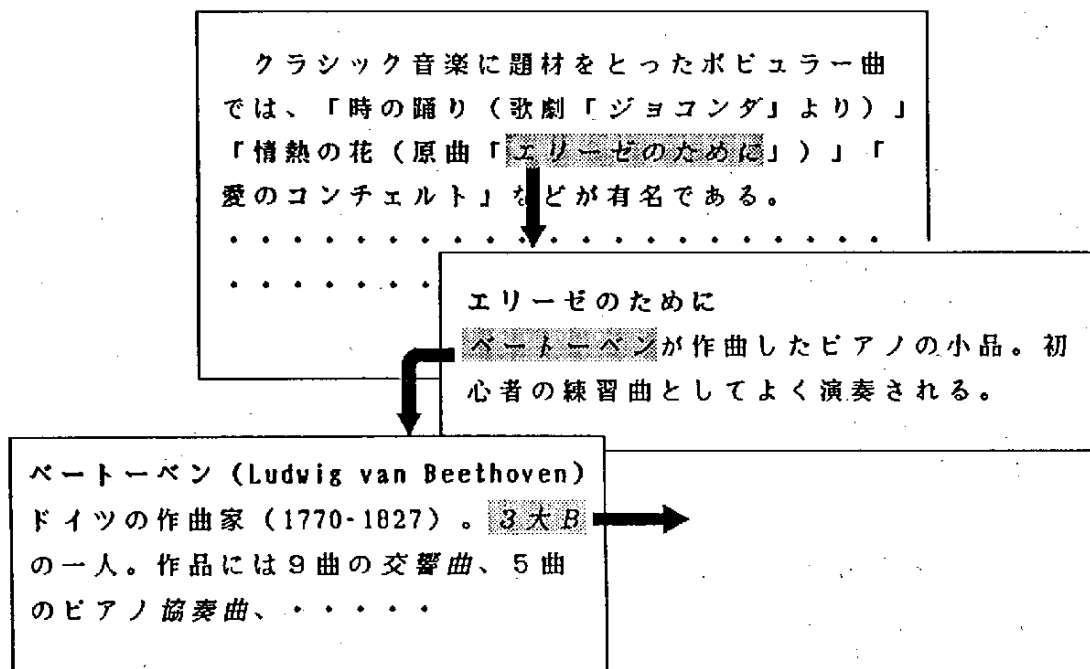
マルチメディア情報処理応用システムは、データベース、出版、通信に大きく分類することができる。これらの関係は図表4.3-4に示す。



図表4.3-4 アマチメディア情報処理

ここでは、特に出版に重点をおいて述べる。マルチメディア出版は、電子百科事典、受験参考書、ゲームなど、多方面で今後の発展が期待される。特にCD-ROMの大容量を生かして、音声・画像・テキストの有機的結合ができる利点である。

マルチメディア情報を人間に提示するとき、ハイパーテキスト技術が不可欠となろう。従来のテキストファイルは線形なファイルであった。一方、ハイパーテキストでは、テキストを幾つかの部分（ノードと呼ぶ）に分け、各ノードの間にリンクを張る。このリンクを通してノードの間の往来を可能にしている。図表4.3-5はハイパーテキスト化文書を読んでいる例で、詳しく知りたい事項に出会えば、その単語をクリックすると説明文が別のウィンドウに現れる。これは、紙の文書では実現困難だったことで、革命的な進歩といえる。



図表4.3-5 ハイパーテキストの例

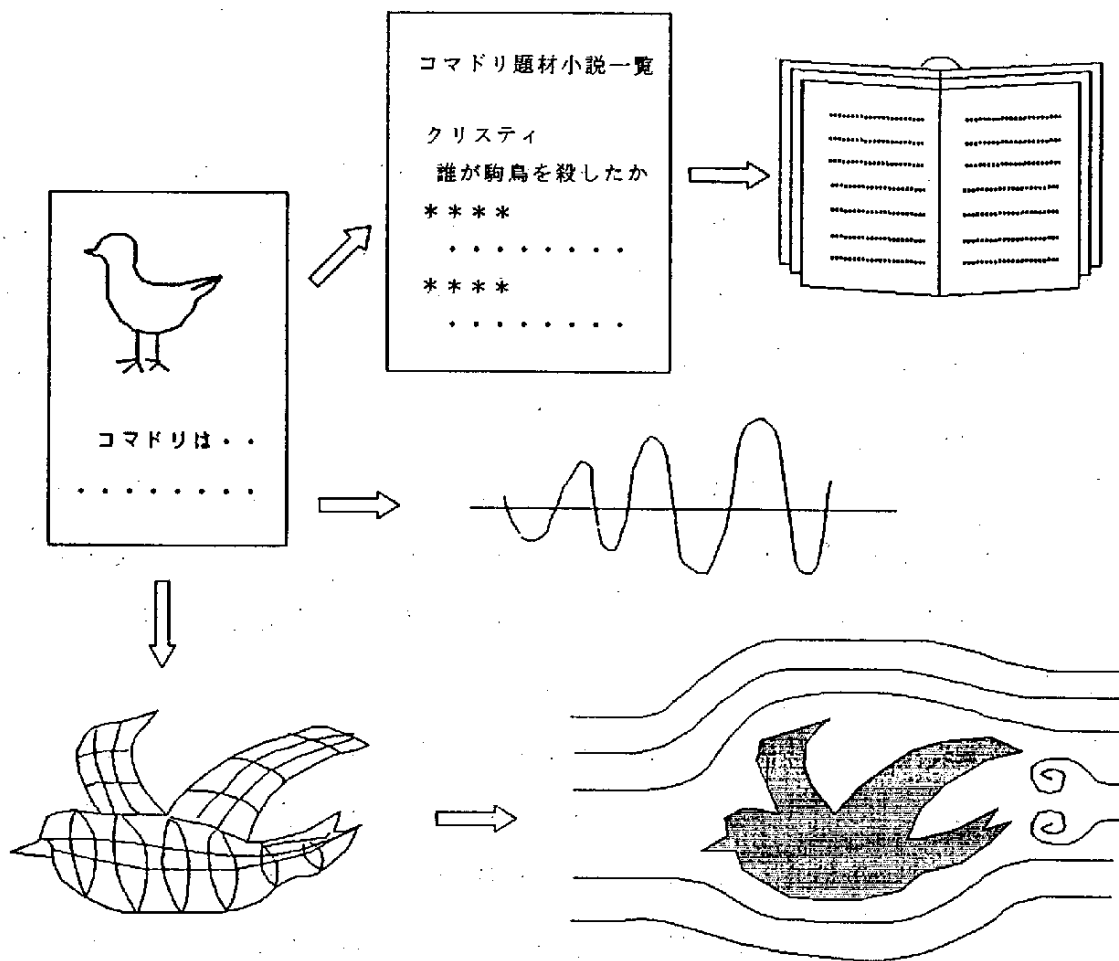
〔各ウィンドウが、ハイパーテキストのノードである。斜体の単語から、別ノードへのリンクが張られている。網かけの単語は、クリックされたことを示す。〕

ハイパーテキストの考え方を異種メディアに拡張したのが、ハイパーメディアである。その一つの例示として、電子百科事典の将来像を描いたのが図表4.3-6である。

ハイパーメディアないしハイパーテキストを実現する上で問題になるのが、そのソースをどうして作るかである。過去の情報は、紙の文書の形で保存されていることが多い。この画像として表現されている文書をハイパーテキスト化しようとする、文書画像をテキスト、図表、写真の各メディアに分離し、その中を意味的に分類して、リンクを張る必要がある。この作業はこれまで人手により行われてきたが、ハイパーテキスト普及に伴って自動化が迫られており、現在研究が進められている文書理解技術の実用化が要請される。

ハイパーメディアでは、これに加えて音声や動画といった異種メディアを統合する必要があるため、パターン認識・理解技術の一層の高度化が必要である。

高度なパターン理解のためには、対象パターンの中に内在する各種の制約を利用することの重要性が指摘されているが、従来の逐次処理手法では実現困難なことも認識されつつある。そのため、従来のパターン情報処理を越える新しい枠組みが要求されている。



図表4.3-6 ハイパーメディアの利用例

鳥の説明 → 文学作品一覧 → テキスト
 → 鳴き声の波形
 → ワイヤフレーム → 流体力学

【参考文献】

- [1] マルチメディアデータベース特集、情報処理、Vol.28、No.6 (1987.6)
- [2] 坂井利之：マルチメディア学内情報システム、第2回「大学と科学」公開シンポジウム (1987.6)
- [3] マルチメディア時代のユーザー・インタフェース、日経コンピュータ別冊、日経BP社 (1989.7)

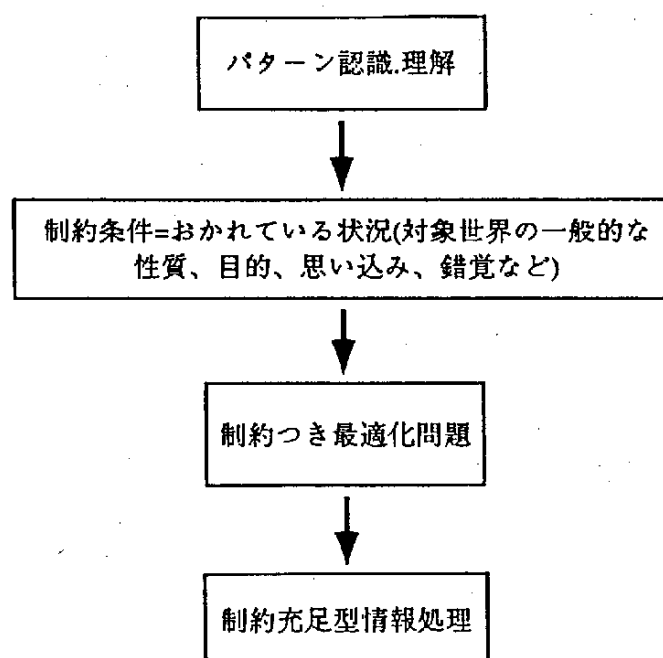
5. 新情報として取り上げるべき 研究課題の提案

5.1 プロジェクトイメージ

パターンの認識・理解においては、得られたデータを解釈することとなるが、現実の状況の下では、得られたデータだけでは解釈が定まらない。このような状況では、あらゆる可能性を想定してそれらの各々に対する評価を論理的に組合せることにより判断を下すという従来の方法の延長上の方策は、組合せ論的爆発により処理時間についての要請を満たすことは出来ない。

そのため、「新情報処理」プロジェクトの一環としての認識・理解のプロジェクトにおいては、「認識・理解機能をモジュール化し、各モジュールでの問題を明確に定義し、それに対する解法を考える。」というアプローチを採り、制約充足問題として扱うことにより統合を図る。

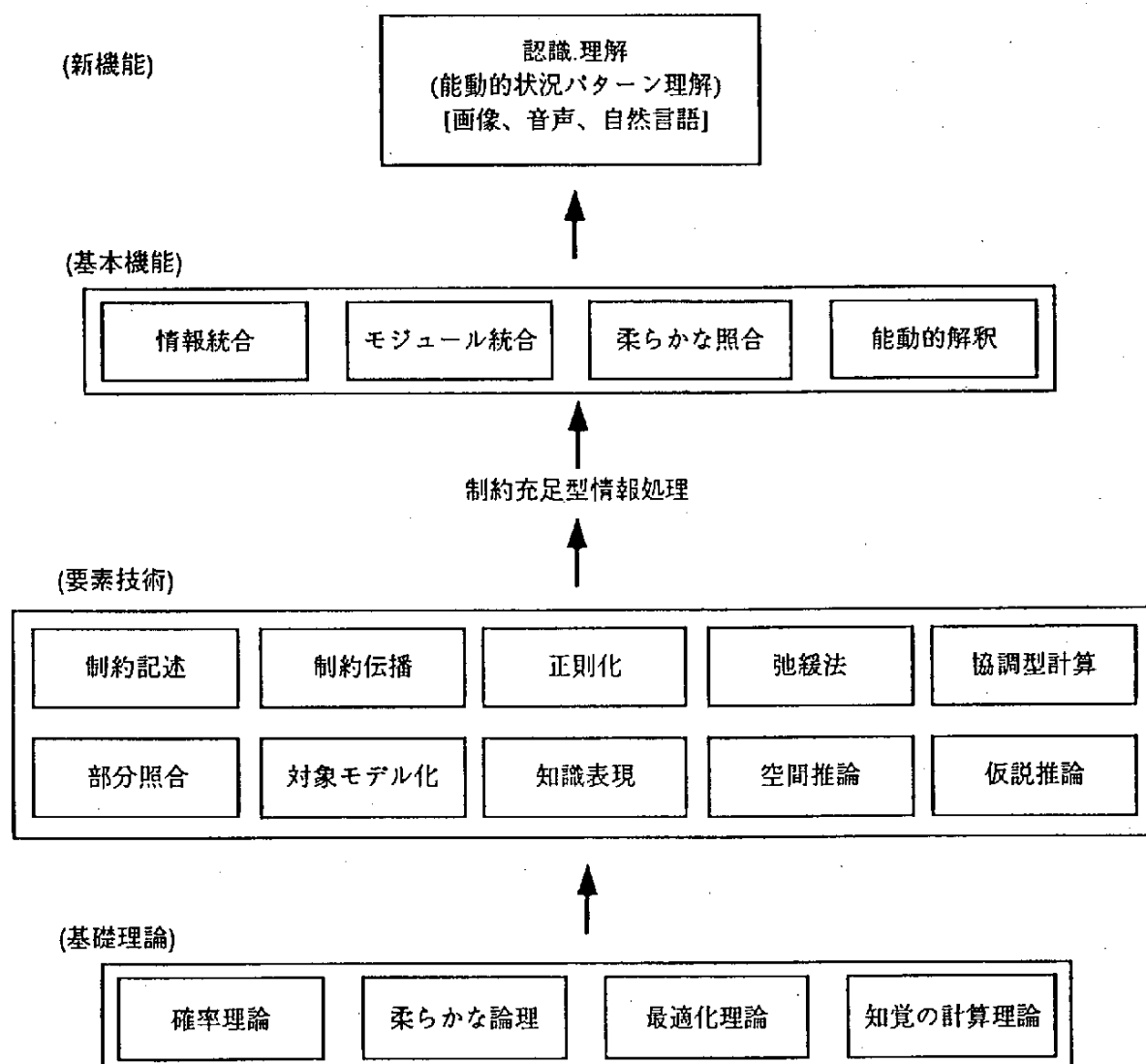
このアプローチでは、モジュール化された認識・理解の問題を、おかれている状況（対象世界の一般的な性質、目的、思い込み、錯覚、文脈等）を制約条件とした制約充足問題としてとらえ、情報の不完全性を許容する形で厳密に定義し、外界の状況を積極的に予測し、入力だけでは得られない情報をも駆使する、能動的な側面に焦点を当てた状況パターン理解の研究を行う（図表5.1-1）。



図表5.1-1 パターン認識・理解の図式

このようにして定式化された問題は、広い意味での制約つき最適化問題であり、多数の処理ユニットの相互作用に基づく制約充足型情報処理が基本となる。

制約充足型情報処理による認識・理解におけるメディア（画像、音声等）に共通した研究開発課題を図表5.1-2 に示す。



図表5.1-2 制約充足型情報処理による認識・理解のための技術課題

制約充足型情報処理は、超並列化の可能性を持ち、漸近的な近似処理の観点からNP困難な問題にも対処し得る可能性があるという点で、既存技術とは異なっている。

5.2 研究課題の提案（状況理解における新機能）

5.2.1 状況パターンの能動的総合理解

我々を取り巻く周囲の状況は多数の要素を含み時々刻々変化しており、それらすべてを把握することは出来ない。また、外界の構造に対するモデルも、不完全な形でしか構築することは出来ない。新しい世代の知的システムでは、「周囲の状況について完全に把握することはできず、外界の構造についてのモデルも不完全である。」ことを前提として、限られた時間の中で何等かの方法で状況の重要なポイントを総合的に把握して適切な行動をとる機能を実現することが要求されている。

この要求条件を満たすためには、以下の技術的課題の研究が必要である。

- (1) 各感覚モダリティの様々な抽象化のレベル（処理レベル）における物理的拘束条件・文脈依存性等を利用した、不完全なパターンの修復機能の実現。
- (2) 個別には不完全で断片的な複数の感覚モダリティからの情報に基づき、相互作用を通じて欠落した情報を補完する機能の実現。
- (3) 状況理解を、外界との相互作用による情報獲得・操作活動としてとえられることにより、不足情報の獲得行動や不適当な状況の改変行動を起動させる機能の実現。
- (4) これらの修復された情報に基づいて、自己の機能の遂行・保全、許容反応時間等の観点から総合的な評価を下す機能の実現。

これらの機能の実現により可能になる行動の具体例は、以下のようなものである。

「突然大きな音がすると、あらかじめ予定されていなくとも、その機械は音の種類を大まかに判断して作業を一時中断し、音の到来方向に視線（注意）を向けて、必要な場合には回避行動や救援活動を開始する。」

5.2.2 不良設定問題・組合せ論的問題の新解法

人間が身の周りの環境を認識・理解するために解いている問題の中には、観測データに誤差が入っているため厳密には解が全く存在しなかったり、あるいは観測データが不十分のため解が多数存在するというような性質をもった不良設定問題、すべてを調べ尽くそうとすると組合せ論的爆発を生じてしまうような問題などの難問が多数存在する。これらの問題を人間と同様に効率良く解くことができる情報処理機構の枠組みを確立することは、重要な研究課題である。ただし、これらの問題の多くに対しては、厳密な解を効率良く求

める解法を追求することがほぼ絶望的であることが、計算量の理論によって明らかにされてきている。従って、何らかの意味で、“近似解で我慢する”あるいは“少々の失敗の危険性は恐れない”という、新しい観点に立った解法の追求を行わざるを得ないであろう。

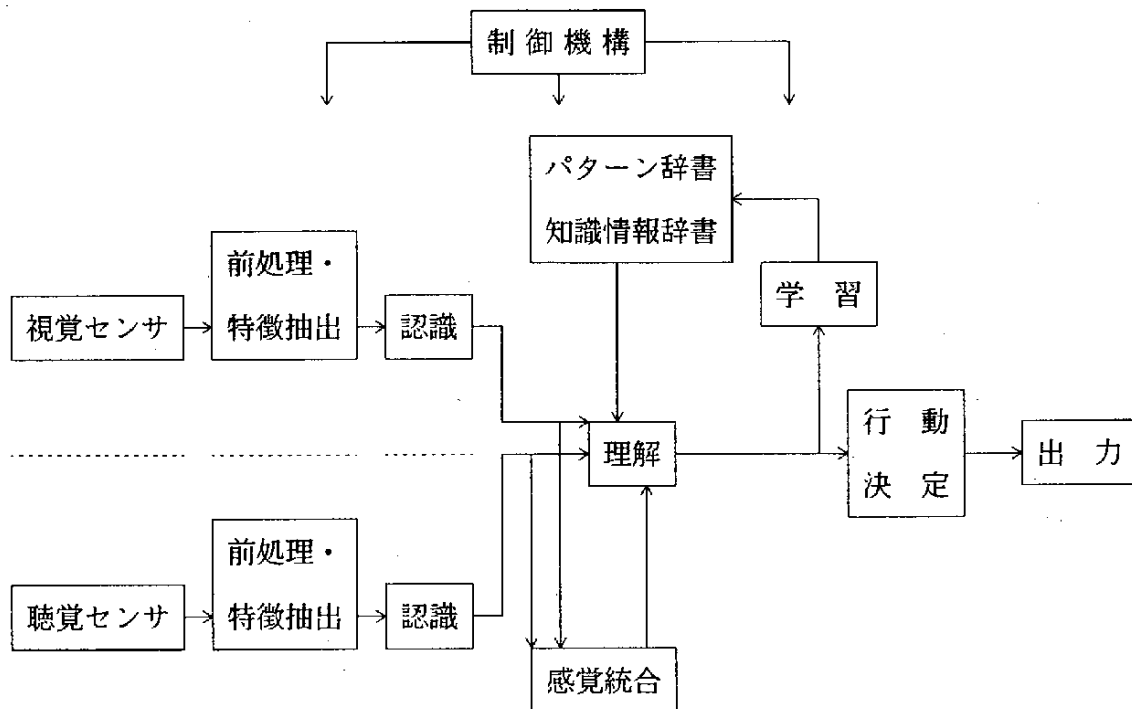
不良設定問題に対しては、従来から最小二重法などのオーソドックスな手段があり、最近では、正則化という名称のもとでより広い範囲の問題を扱う枠組みとして定着しつつある。組合せ論的爆発をもたらす問題に対しては、単純に近似解で我慢しようとする発見的方法、精度保証付きの近似解を得る方法、どの時点で処理を止めても処理に費やした時間に見合った解が得られる分枝限定法、などが考えられている。しかしながら、これらは未だ、人間の認識・理解の能力に達しているとは言いがたい。

より性能の優れた解法を求めるためには、問題自身を見直してみる必要がある。なぜなら、人間の情報処理の各側面を数理的な問題として設定すると、一般に人間が解いていない問題を含めた形で問題群がとらえられてしまうことが多いからである。従って、そのように定式化した問題群の中で人間が実際に解いているものとそれ以外のものとの境界をできるだけはっきりさせ、前者が確実に解けるなら後者に対しては間違った解を作ってしまう危険性があってもかまわないという観点に立った解法の研究も可能であり必要でもあろう。このような研究の方向転換によって、初めて、計算量の理論に反することなく効率の良い解法を追求する道が開けるように思われる。

5.2.3 情報表現の階層化・モジュール化と異種モダリティの結合

次世代の認識・理解システムを実現するための、モジュール階層構造化を検討する。システム全体として柔軟な群情報処理に基づく認識・理解機能を実現するための、モジュール構成法、各モジュールの主要機能、モジュール間情報形態、モジュール制御方法等を研究する。

具体的に想定するシステム構成例を図表5.2-1に示す。これは、人のパターン情報処理の方法と、それを基にした行動を応用工学的な観点から模式化してみたものである。



図表5.2-1 新情報処理の流れ

これらの各モジュール群に関して、主に以下の内容を検討する。

① パターン群情報処理モジュールの構成技術

学習付きパターン認識モジュール、自己組織型パターン識別モジュール、制約不充足問題解決モジュール、パターン連想等の新機能を持つ要素モジュールの実現法を検討する。

② モデル記述・検索モジュール技術

対象モデルの記述・記憶・検索方法の検討とその実現方法を検討する。

③ モジュール高速処理技術

学習、自己組織化、問題解決、などを高速に実行するための並列、分散処理アーキテクチャを検討する。

④ モジュール制御技術

各モジュール間の情報伝達、分散処理技術、信頼性技術など全体として統一を持ったシステムとする手法を検討する。

⑤ 複数モダリティ実現技術

基本モジュールを複数のモダリティに使用するための汎化技術を開発するとともに、これを用いて視覚、聴覚、臭覚等の異なった感覚モダリティを統合したり、形、色、

奥行き、動きなどの知覚機能を統合するための基本技術を開発する。

5.2.4 動機付け機構のモデル化と認識・行動のモデル化

現在の技術を使って、周囲の状況を認識し、適切に行動するシステムを実現するには、生起する状況を予測し、これに対して適切なプログラムを作り、インプリメントすることになる。しかし、あらかじめあらゆる状況を把握するのは不可能であり、この方法では、システムのデッドロックは避けられない。解決策は、状況に応じた適切な行動を発現する機構を組み込むことである。例えば、単に、周囲の状況のみに依存するのではなく、システムの行動目的や、行動結果に関する予測評価や、過去の経験などを動機付けとして、デッドロックなしに妥当な行動を発現する方法が考えられよう。

人やサルにはこの機構が備わっている。前頭連合野に、適切な行動を企画する機能がある、と言われている。この脳部位は、感覚情報や動機付けに関する情報を統合し、それに基づいて判断や意志決定を行ったり、随意運動の統御を行ったりする“最初中枢”とされている[1]。

この前頭葉の機能のうち、注意のメカニズムが、神経回路網により模擬・研究されている[2]。“注意”は、認識や行動に密接に関係する意識的ないし無意識的な精神活動であり、各種の動機付けが関与している。この事例では、注意する事象ごとに、相互には抑制統合された出力細胞を用意し、その興奮により、どの事象に注意しているかを表している。出力細胞の興奮状態には、多数の介在細胞を含む巧妙な神経結合を介して、感覚入力、覚醒入力、動機入力に影響している。この回路を用いて、パラメータの数値を変更するのみで、①新規に生起した事象に次々注意を移す特性と、②一旦注意を向けたら、新規の事象の生起には頓着せず、同一事象に注意を向け続ける特性、とが実現されている。この研究は、神経回路網による前頭葉の一部の機能の模擬可能性を示している。この考え方を発展させることにより、前述の機構のモデルが得られると期待できる。

このように、①②の注意特性を積極的に用いて、適切にコントロールできれば、システムがどういう場面で、何に関心を持ち、何に重きを置いて作業を行うかという行動のモデルを導入できるという期待がもてる。これによって、システムに知識・情報不足からくるデッドロックを防ぐ上での高次の制約条件が導入でき、ある種のシステムの個性をインプリメントできることにもなる。すなわち、システムの行動パターンへの傾向を与えるメン

タルモデルを提供できることにもなる。これにより、利用者側で、ある程度システムの行動が予測できるようになり、利用者に対して一種の安心感を与える要素にもなる。また、システムに利用者と同様の動機付け要因を一部に備えさせれば、利用者の要求をくみ取って作業するシステムに変身させられるという期待も持てる。従って、ここでは、動機付けを支配する要因[3]を与えるメカニズムの研究開発、これらを統合して動機付けとして機能させるための研究開発を課題として挙げる。

【参考文献】

- [1] 渡辺正孝：“前頭連合野と知的機能”、脳科学の新しい展開、pp.106-117、岩波書店、(1986)
- [2] Daniel S. Levine and Paul S. Prueitt：“Modeling Some Effects of Frontal Lobe Damage-Novelty and Preservation”, Neural Networks, Vol.2, pp.103-116, (1989)
- [3] D. A. Norman：“認知科学のための12の主題”, 認知科学の展望(D. A. Norman編), pp295-342, 産業図書, (1984)

5.2.5 認識理解機構の自己組織的学習

パターン認識・理解研究の歴史は長いが、現状では音声・文字・画像など対象ごとに別々の手法を用い、統制された環境のもとで成功を納めているに過ぎない。その理由は、現在の技術では個々のパターンに応じて人間が場合分けを行い、個別処理を行っているが、統制なしでは、多種多様のパターンの変形すべてに対応することが困難だからである。また、問題の規模が増大すると、組合せの数が指数関数的に増えるため、逐次形の処理では実現困難な場合も多い。

これからの社会で認識・理解の重要性は益々増大し、対象の拡大や規模の増大が厳しく迫られることになろう。従って、人間が場合分けを行って個別にプログラムを書くのではなく、認識理解機構を自己組織的に構成する技術の研究が極めて重要となる。

この研究は、次の各段階を追って進むと考えられる。

第1段階：既存の認識理解機構を利用し、その中のパラメータを学習によって定める。

第2段階：原始的・要素的な認識理解系を用意し、その組合せや内部パラメータが最適になるように学習し、自己組織的に認識理解機構を構成する。

第3段階：数種類の原始的・要素的な認識理解系から出発し、それらを多数重複させたり、要素機能を改良して、高度な認識理解機構が構成されるように自己組織化する。

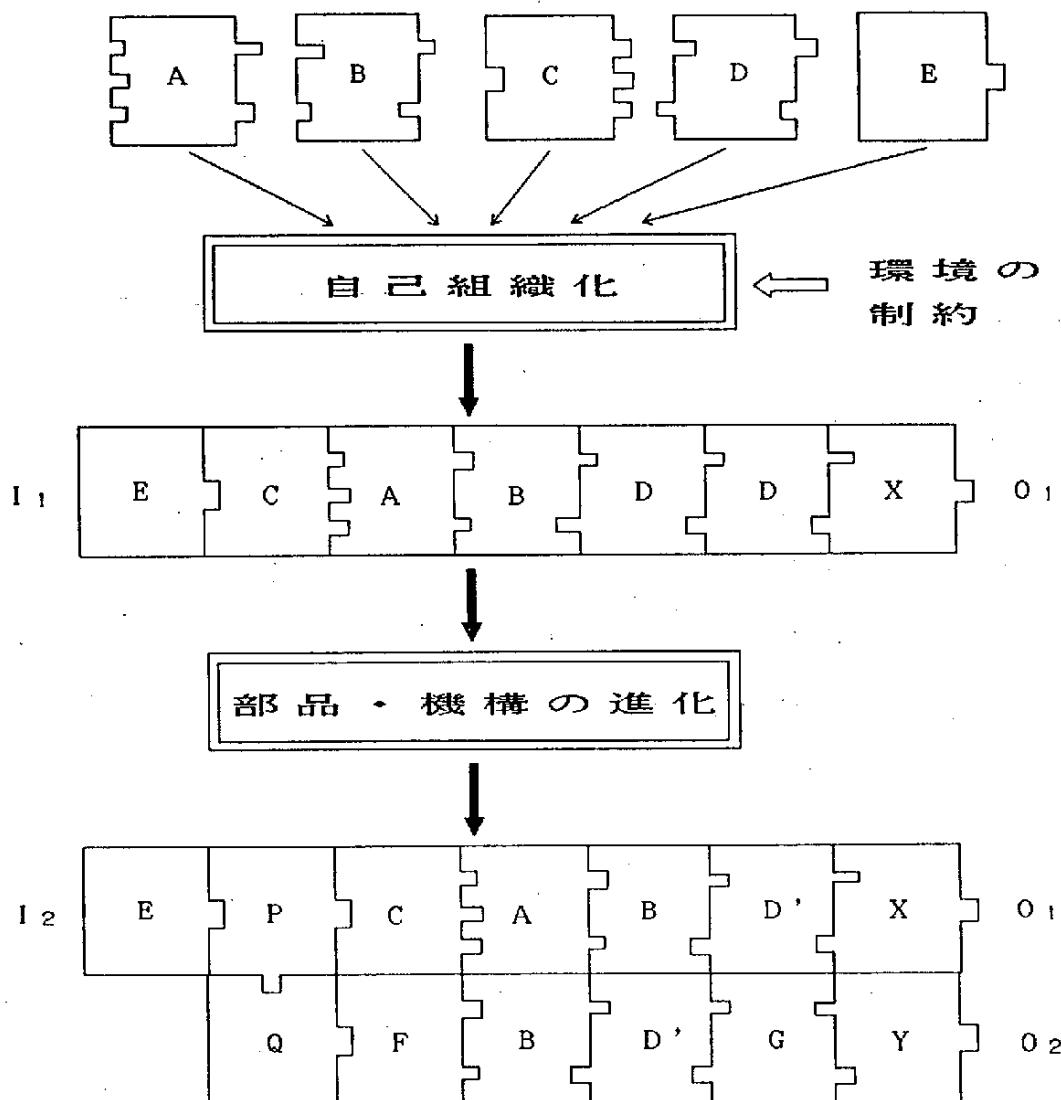
いずれの段階でも、認識理解系と環境との相互作用によって、機能・性能を進化させて行くものであり、生物の進化になぞらえることができる。

現在は、第1段階で各種の試みがなされている状況にある。しかし、既存の認識理解機構を利用するのでは飛躍的な発展は望めず、認識理解機構自身の学習的形成が必要となる。

図表5.2-2 は、第3段階で実現が望まれる自己組織化過程を模式的に描いたものである。最初に用意する機能要素は、種類も少なく機能も低い。これらの入出力インタフェースは数種類に規格化され、組合される相手が限定されている。認識・理解したい環境は、この初期機構に対する淘汰圧力として働き、初期機構は、環境の変化に対してうまく適応するように、これらの機能要素の組合せを自己組織的に最適化して行く。自己組織化が進むにつれて、機能要素の部分集合を重複させたり、機能やインタフェースを確率的に変化させたりして、より洗練された高度な認識理解機構に進化し、より厳しい環境に耐えられるようになる。

21世紀を目前にして、上に述べた認識・理解機構の自己組織化の研究を誘導・加速すべき時期に来ていると思われる。

認識・理解機能部品



図表5.2-2 機能部品の結合による認識・理解機構の自己形成

5.2.6 感性情報処理

(1) 知にはたらけば角がたつ

これまでの情報処理技術は、様々な知的機能の実現を目指して発展してきた。

しかし、人間と知的情報処理機器との円滑なコミュニケーションのためには、知的機能だけでは不十分で、ユーザの気分や能力などに応じた応対が必要となる。そこで、以下のような、新しく重要な研究課題を提案する。

(2) ユーザの気分を読み取る

① 画像処理による表情理解

画像処理により、ユーザの喜怒哀楽を判定する。これは、時間的な変化も重要な手

がかりとなるため、時間的空間的な文脈を考慮した処理が必要となる。

② 音声処理による気分理解

音声を聴いて、われわれは話し手の気分をある程度知ることができる。

これは音の強弱、高低などが重要な手がかりとなることが予想される。

③ 自然言語処理によるユーザの意図理解

ユーザの発話を分析することにより、ユーザの知識の程度、問題解決能力を知るだけでなく、陽には表れてないユーザの意図を見てとることを可能とする。また、一連の文章から、ユーザの性格、例えば積極—消極、淡白—粘液質などを推定する手法を開発する。

④ 総合的ユーザ理解システム

①～③の研究成果を統合することにより、全体としての処理の信頼性と効率を向上させる研究を行う。

(3) パターンの美しさの個性的理解

絵や書のような視覚パターンの美しさを人間のように判定する処理法を研究する。また音楽のような聴覚パターンの快さ、文章の美しさなどについてもその判定法を開発する。

美の判定は、多くの人に共通な部分と、個人に強く依存した部分とからなるものであるから、個性のよってきたところも明らかにする。

(4) ユーザに合わせた出力法

これまで述べたことはシステムへの入力に関するものである。システムから人への出力として以下の課題が考えられる。

ユーザの気分に合わせて、快活に、あるいは遠慮がちに音声を出力したり、文や文章の表現にも気くばりをはたらかせたりするための研究が一つのテーマである。

つぎに、ユーザの個性に合わせた、美しいグラフィックスの生成も重要な課題である。

6. 研究成果の展開

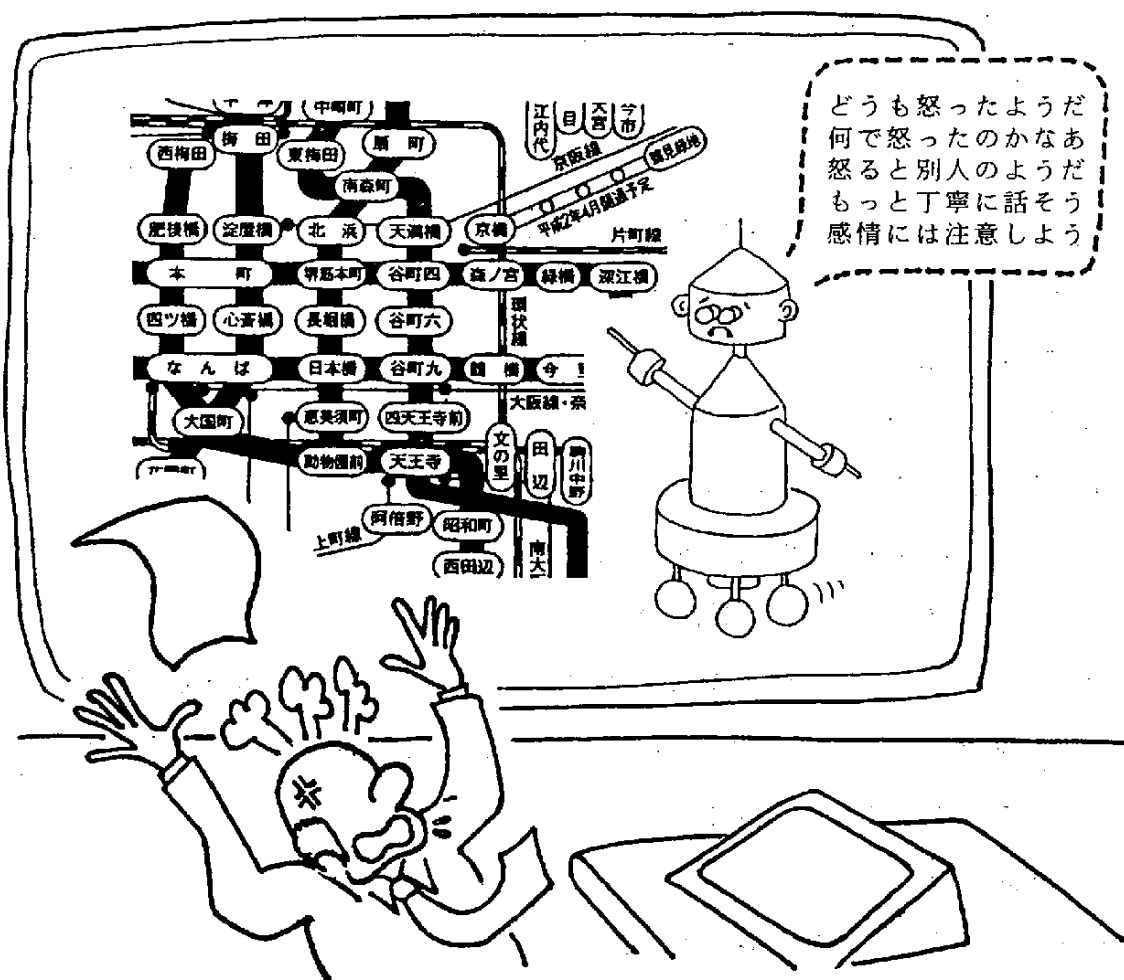
6.1 利用者の状況を把握して応答を変える学習型の情報システム

従来、コンピュータやワープロなどの情報機器には、利用者側から、「使いこなす」という感じで、慣れ親しんでいくものであった。今後、システムが高度になり、普及するにつれて、利用者も、その要求も多種多様になってくると考えられ、システム側の適応力や表現力も豊富にならなければならない。このため、マルチメディアの各種のチャンネルを利用した双方向のコミュニケーションが重要になるとともに、利用者自身の個性や状況を多面的に把握でき、その変化にも自律的にうまく学習・適応していくシステムが望まれる。ここでは、音声による対話をベースとしたシステムについて考えてみることにする。

- ・声から利用者を聞き分け（話者同定）、その人の、年令、性別、社会的立場などの利用者の背景、既に学習した個性（利用者モデル）に応じて音声応答者モデル（年令、性別、敬語、話し方など）を変えることが出来る。また、音声のみならず、図やアニメも使って、臨場感豊かに表現できる。マルチメディアをフルに活用して、高齢者や障害のある人にも使いやすい入出力環境を整備する。
- ・うなずいたり、手で示す方向や「待った」の合図、「困った」という表情といった音声以外の表現をも認識し、即座に反応できるリアルタイム性の優れたものにする。システムが急いで結論を出さなければならないと判断したら、内容の吟味をほどほどにしても、結果を出せる。声の調子などから感情的側面を推し量りつつ、場合によっては、元気づけたり、感情に訴えたり、説得したり、なだめたりもできる。
- ・状況に応じて、例えば、音声応答を止めたり、表現・話し方を変えたりできる。利用者が動き回っても、その声の位置を追跡して、求める音声やイメージだけをピックアップし、周囲の騒音や利用者の位置に応じて音声の大きさを換えたりできる。音声で説明できないと判断したら、図をディスプレイするなど機転を利かす。
- ・利用者の話し癖、言い回し方などのほか、会話・表現パターンを学習し、これを高次の制約条件として利用して予測・適応していける。また、以前からのシステムの使われ方の履歴から、どう対処すべきか、学習していけるような、融通のきくものにする。
- ・システム自身は、いろんなタイプの人間・環境のモデルを持っているが、システム・利用者互いの知識不足や情報不足の状況、利用者自身の意図もはっきりしない状況なども

生じることがあり、場合によっては、システム側が指導権をもって質問したり、行動することもできる。システムは、利用者との関係を想定しながら、利用者のメンタルモデルを学習し、利用者にとって最適な状態を推定して、それに適合していくようなメカニズムを持っている。

これらにより、やさしく、物分かりのよい、親しみやすいシステムに仕上げ、この結果、システムと人間との豊かな共働作業の環境を提供できると考える。



図表6.1-1 学習型情報システム

6.2 高度な映像情報の編集・加工システム

通信手段の発達、データベースの整備等で、21世紀には全世界のデータがいつでも、どこでも利用可能になると期待される。しかし、利用者の立場からは、その豊富な情報をど

のように利用するか、また利用できるかが課題となってくる。本システムは、そのような情報を人に見やすく整理する技術を開発することによって、個々の利用者にとっては一見無秩序に入ってくる映像情報を、利用しやすい形に変換し、提供することを目指すものである。

データベースの提供や放送などのメディアは不特定多数の人に向けた情報を提供する。これに対して各個人、又は企業は、その中から真にためになる情報を選択して聞こうとする。情報の中には、それによって具体的な対応が求められるために聞き漏らしてはならないもの、知識として知っておけばよいもの、あるいは聞かなくてもよいものといろいろ存在する。この選択の基準は個人によって変わり、かつ、時と共に変化していくものである。そのため、情報の中身を理解すると共に、個人が何を欲しているかを学習で知る必要がある。さらに、そのようなシステムはマンマシンによって対話的に答えを出してくれると考えられるので、使用者のあいまいな問いかけに対しても柔軟な発想、時には連想機能を用いて答えるような柔軟な情報処理が求められる。これらをまとめると、主な研究開発項目は以下のとおりである。

(1) 動画像理解技術

放送メディア等の自動整理、知識情報とメディア情報の自動対応を行うために、

- ・映像情報を理解し、シーンを自動分離、整理し、
- ・各シーンと音声情報との関連を抽出し、
- ・テキスト情報等、他の関連情報との関連付け

を行う。このために、人の知覚、特に画像理解、自然言語理解、等の機能を融合した認識・理解アルゴリズムを開発する。例としては、多段多層ニューロネット型認識機構、モデルベースド認識機構が挙げられる。

(2) 学習・知識獲得技術

データベース検索の質を高めるために、

- ・ユーザの検索の過程を学習し
- ・興味事項を推定し、
- ・それに合うように情報の加工、インデキシング

を行う。そのためのユーザに関する知識を獲得する技術を開発する。

(3) 加工・編集技術

上記開発技術を基に、検索した情報を基に、新しい情報に編集加工するためのシステ

ム編成技術を開発する。

6.3 教育・娯楽機器

マルチモーダルな状況パターン（音声、文字、イメージ、概念、シーン、など）の能動的理解や感覚知覚系の統合処理により、人間中心のコミュニケーションが実現される。ここでは従来の人間－コンピュータ間のコミュニケーションの枠を越え、

- ・『美しさ』『楽しさ』など感性、感情の理解、表現の実現。
- ・日常世界をリアルにシミュレートする仮想環境や、非日常的世界を創作・シミュレートする仮想環境の実現。

などの『豊かな表現』が期待できる。特にモノの生産をとまなわない教育・娯楽機器においては、この技術が大きな原動力となる。

ここで巨大な状況パターン（音声、文字、イメージ、概念、シーン、など）データベースを背景とした状況パターンを知覚するセンサとリアルな表示装置を備えたシミュレーション環境を想定し、その応用システムを考える。

(1) 個人学習を豊かにするサポートシステム

地理の時間の風景。データベースからニューヨーク街角の風景やソ連の田園風景などが現われ、学習者の視線に沿って場面がどんどん展開して行く。また時間を戻したり、高度1,000mからの視線にしたり、サイズを変えたりすることなど実際に経験できないもの（マジック）までも経験できる。

(2) 協調作業のサポートシステム

授業や会議の風景。操作者は家庭やオフィスなどそれぞれ別の場所においてスクリーンを観測している。教室の授業や討議を行いながら実施する会議と同様に、スクリーンには教室や会議場が表示され、かつ参加者の動作などがその中に表示される[1]。

臨場感を持って『直接参加者』あるいは『スクリーンの主人公』として参加できる。

(3) トレーニングのサポートシステム

スポーツのトレーニング風景。身体に取り付けた器具や感覚センサ系と人間の動作が連動されている。実際に走ってみる。スピードや視線に沿って道路や周りの風景など現実風景が展開していき、自然の中を走っている気持ちになる。ドライブやスキーのトレーニングも同様に行える。また森や田園をゆっくり散歩することもでき、爽快感をも味わうことができる。

【参考文献】

- [1] 岸野：知的通信と臨場感通信、第3回公開シンポジウム知的ヒューマンインタフェース、PP. 55-64(1990)

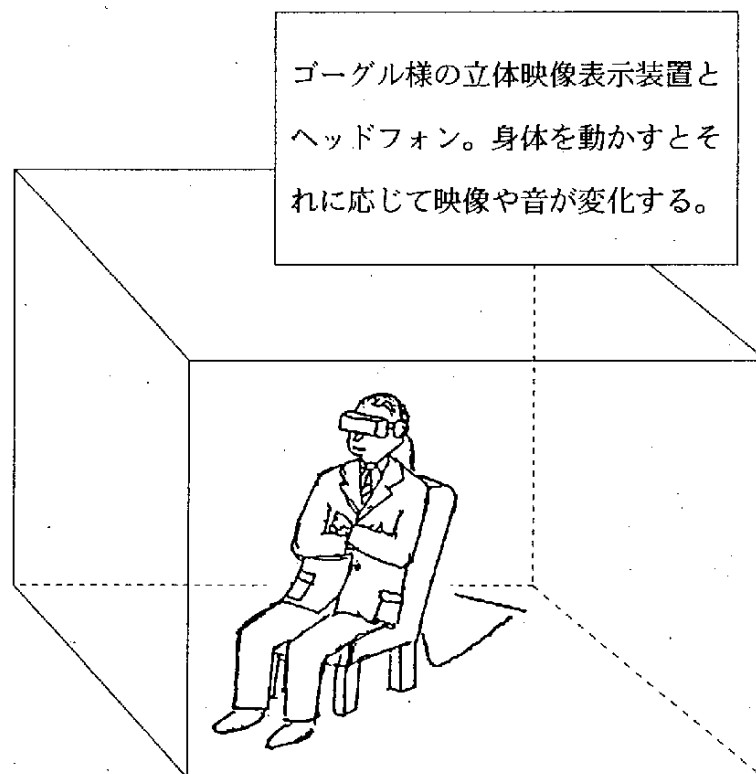
6.4 人工的に様々な体験を現実感を伴って提示し センシングすることのできる超環境実現装置

複雑な事象に関する、多面的なデータベースのアクセス結果や、多面的なシミュレーション予測結果等を、総合的に現実感を伴って、分かりやすく人に見せるために、その人を包む空間の中に、人が見るべき情報を擬似環境として作りだし、体験的に理解させるための情報提示装置。図表6.4-1 にイメージを示す。

例えば、未知の他国の状況を、この装置が提示する情報によって体験的に理解する場合を考えてみる。旅行者として、車で街の中を走っていると仮定する。図表6.4-1 に示す部屋の中に車のシートを置き、それに座した人は、データベースをアクセスすれば、その後はあたかもそこにいるかのような景色が見え、感じがする。これを楽しむこともできるし、他の国に対する理解を深めることもできる。

また、例えば、酸性雨が環境に与える影響を理解するため、他国のデータベースにアクセスし、あたかもその国の森の中にいるような状況や、水中にいるような状況を、この装置で再現すれば、森林の様子、あるいは水中生態系の変化の様子を、体系的に理解できる。

データベースにアクセスする代わりに、シミュレーション結果をこの装置を用いて提示すれば、予測された未知状態を現実感高く理解できる。例えば、現在、いろいろ懸念されているCO₂による地球温暖化の問題も、シミュレーション結果を本装置を用いて提示すれば、無味乾燥な数値の代わりに、視聴覚ないし肌で感じる情報であるため、誰でも容易に理解できる。



気温、湿度、風、香りなどを制御できる部屋

図表6.4-1 超環境実験装置イメージ

6.5 障害者・高齢者等の低下した機能の補綴システム

経済的に高度な発展をとげつつあるわが国において、これまで生産指向の技術が主流をなしてきた。しかしながら、今後の技術は、世界規模での人類の福祉を目指した人間指向のものを取り上げていくことがわが国に課せられた責務であろう。

このような人間指向の技術の適用の場として、障害者や高齢者用の支援システムが考えられる。支援すべき機能としては、感覚、運動、高次思考の諸機能があげられる。

では個々の応用システムを列挙してみよう。

(1) 視覚障害者用補綴システム

文字だけでなく、図面をTVカメラを通して読み取り、図面の空間的構成の概要を触覚ディスプレイにより、図面がどのような構成要素からなるかなどの記号的情報を音声により、それぞれ障害者に呈示する。このシステムをさらに高度にして、放送局のアナウンサーによるスポーツ、演劇などの実況中継のかわりをするシステムを開発することが考えられる。この場合、非常に強力な動画像理解の開発が前提となる。

(2) 聴覚障害者用補綴システム

相手の発話内容を、音声認識、自然言語理解の処理手法により抽出し、障害者に伝えるシステムをまず開発する。さらに進んだシステムでは、相手の気分、意図などの情報も抽出、呈示できるようにする。また、音源の方向も検出し、呈示する。

(3) 高次中枢障害者用補綴システム

このようなシステムを開発することは難しい。しかし、次のようなシステムでもある程度意味があると思われる。例えば、軽度の老人性痴呆患者と、同じような会話を、あくことなく繰り返し、お相手をしてくれるシステムである。このようなシステムの“はしり”はELIZAである。

(4) 発声、発話障害者用補綴システム

喉頭がんなどのため、発声器官を失った患者の発声・会話などの意図を、口の動きなどから読み取り、合成音声で代用するシステムを開発する。“意図”が分かれば、システムが、有能な秘書のように、“気を利かせて”代行してくれるシステムの用途も大きいと思う。

6. 6 心やさしきコンパニオン

一人暮らしの老人や、病人の相手をしてくれるシステムのことである。

このようなシステムは、感性情報処理技術の重要な応用分野である。

老人や、病人の気持ちを汲みとって、それに合わせて、励ましたり、なぐさめたり、笑わせたりする機能を持たせる。

また、顔には現れていない意図を、表情や、発話内容から理解し、気を利かせたお世話をすることも重要である。

時には、一緒になって、ゲームをしたり、読書をしてあげ、その内容について話し合うといったこともできれば申し分ない。

『老人とコンパニオン』

老 「おーい。ちょっと来てくれないか。」

コ 「はい。为什么呢か。」

(老人の声を認識し、また声のした方向から老人のいる所を判断して老人の所へ来る。)

老 「腹が減ったな。」

コ 「じゃ、軽い食事を用意しましょうか。」

老 「ウン」

(老人の意図を理解し、好みの料理を用意する。)

.....

コ 「今日はいいいお天気ですね。上野あたりは花見で大変な人出でしょうね。」

老 「そうだな。それでは神代植物園にでも行ってみるか。」

コ 「結構ですね。タクシーを呼びましょう。」

(老人が退屈そうなのを見てとり、気晴らしに外出を勧める。)

.....

コ 「見事な花ですね。」

老 「まったくだ。丁度見頃だな。」

コ 「深大寺の方へ回って見るのもいいですね。」

老 「そうしよう。」

(花の美しさを判断する。隣接した深大寺が老人の好みであることを知っている。)

.....

7. む す び

認識・理解という概念を、技術の歴史的発展と生理学・心理学の新しい知見を踏まえて見直すことにより、新情報処理体系の鍵となる機能の抽出を図った。この検討の過程は、これまでどちらかといえば意識的、明示的な側面が強調されてきた知的機能が、実は多数の巧妙な無意識的、自律的下部システムの相互作用に支えられているものであることを明らかにしてきた。このような意味で、認識・理解の諸機能は、「新情報処理」が探求すべき方向の指針を与えるものと考えることが出来る。

生物に見られるような重層性を特徴とする認識・理解の諸機能のすべてを統一的に特徴付ける原理を見いだすことは困難であるが、制約充足処理の概念によりその中の多くの機能を理解でき、実現できる可能性が示された。いわば制約という共通語を介して多様な視点を代表する要素的な知的素子間の調停を行うことで目的とする情報処理機能の実現を図るわけである。これらを実際に具体化していくことは、今後の課題である。報告書では、具体的な研究項目の提案を行った。

このような検討を通じて実現される「新情報処理技術」は、これまで、逐次的、論理的な側面が偏って発展してきたコンピュータ関連技術に新しい展開をもたらし、感性の領域にまで踏み込むための手がかりを与えることが期待される。報告書では、これらから生みだされる様々な応用についての提案を行った。

以上、認識・理解の観点から次世代の情報処理に要請される新機能の検討結果について報告した。ここで見いだされた方向、様々な提案は、21世紀を真に人間的で豊かなものとするための鍵を与えるものと期待したい。

参考資料：関連するキーワード表

〈キーワードからの概念の抽出〉

1. 経 過

委員へのアンケートに基づいて、キーワードを集めた。それらのキーワードの意味するものが近いものをまとめてグループを作成し、そのグループの集まりをうまく記述できるように概念軸を抽出した。こうして最初に抽出された概念を、次に示す。また、付録の最後に、集められたキーワードを概念毎にまとめて示す。

2. 抽出された概念と具体的内容

- 検討の枠組
 - 感覚統合
 - 機能分散処理
 - 記号処理とパターン処理の融合（照合）
 - 不完全情報問題の解法
 - 能動的理解（active interpretation）
 - 自己学習
- 検討のレベル
 - 生理・心理
 - 視覚
 - 聴覚
 - 感性・文化・応用

〈概念分類項目と具体的なキーワードの例〉

	感覚統合	機能分散	記号処理とパ ターンの融合	不完全情報	能動的理解	自己学習
生理・心理	A 1	B 1	C 1	D 1	E 1	F 1
視覚	A 2	B 2	C 2	D 2	E 2	F 2
聴覚	A 3	B 3	C 3	D 3	E 3	F 3
感性・文化・応用	A 4	B 4	C 4	D 4	E 4	F 4

- 1 A 感覚統合
 - 2.1.1 A1 感覚統合：生理・心理
キーワード Fusion（感覚融合）／共感覚／異種情報間の照合／連合野／情報の生物学的意味
 - 2.1.2 A2 感覚統合：視覚
キーワード Fusion（感覚融合）／動き、形状、テクスチャ、色、空間配置の統合／3次元情報伝送・表示／立体視／パターン情報の入力／異種情報間の照合／ホログラフィ
 - 2.1.3 A3 感覚統合：聴覚
キーワード 異種情報間の照合／音源方向、ピッチ、音色、韻質、個人性の統合
 - 2.1.4 A4 感覚統合：感性・文化・応用
キーワード モーションによるインタフェース（高齢化、福祉）
- 2 B 機能分散処理
 - 2.2.1 B1 モジュール化・階層化（並列分散）：生理・心理
 - 2.2.1.1 B1 情報の表現方法
キーワード 認識細胞／分散表現／空間の表現／射影 情報の脳内表現地図／自然な範疇分け
 - 2.2.1.2 B1 機構と機能の関連
キーワード 皮膚内多重並列投射（認識機構）／階層的認識システム／並列処理モデル／ホロニック／アナログ機能モジュール／汎化 柔らかな認識／連想／音源方向、ピッチ、音、韻質、個人性の表現／動き、形状、テクスチャ、色、空間配置の表現／昆虫などにおける出力系（行動？）／脳磁界計測／脳波解析
 - 2.2.2 B2 モジュール化・階層化（並列分散）：視覚
 - 2.2.2.1 B2 画像理解アーキテクチャ
キーワード オブジェクト指向／スケジューリング・制御／モデルベース／階層的認識システム／知識とのリンク／処理の分配方式
 - 2.2.2.2 B2 構造化 認識・理解の計算理論
キーワード 正規化・正則化／object consistency／color consistency／動き検出／モーション抽出／モーション理解／異種情報間の照合／空間の表現／時空間認識
 - 2.2.2.3 B2 視覚システム
キーワード マンマシンコミュニケーション／画像通信／資源探索／製造検査工程への応用／3次元素子／形状意味論／認識のソフトウェア／汎化 柔らかな認識／文脈（シーン）の理解
 - 2.2.3 B3 モジュール化・階層化（並列分散）：聴覚
 - 2.2.3.1 B3 構造化 認識・理解の計算理論
キーワード スケジューリング・制御／モデルベース／識別単位／認識のソフトウェア
 - 2.2.3.2 B3 階層的認識システム
キーワード 音声合成・認識／言語の理解／常識など／知識とのリンク／処理の分配方式
 - 2.2.3.3 B3 聴覚システム
キーワード マンマシンコミュニケーション／雑音除去／汎化 柔らかな認識／両耳効果
 - 2.2.4 B4 モジュール化・階層化（並列分散）：感性・文化・応用
 - 2.2.4.1 B4 感性処理
キーワード 感性情報処理 グラフィックス／感性情報処理 音楽／感性情報処理 書字システム／感情／自然な範疇分け／楽しみの追求／マンマシンコミュニケーション／モーション理解／美的なものの追求／表情の理解
 - 2.2.4.2 B4 文字・音声・図形・画像等、既存分野への応用
キーワード 気象予測／障害者の社会復帰／職人芸の修得／画像診断／通訳／防災 異常検知／理解の入力（確認）手段
- 3 C 記号処理とパターン処理の融合（照合）
 - 2.3.1 C1 記号処理とパターン処理の融合（照合）：生理・心理
キーワード PDP／Society of Mind／コネクショニズム／分散表現／認識機能モジュール／認識細胞
 - 2.3.2 C2 記号処理とパターン処理の融合（照合）：視覚
キーワード PDP／Society of Mind／コネクショニズム／アイコロジック／画像検

- 索／識別関数／手書き文字認識／特徴抽出／メディア変換
- 2.3.3 C3 記号処理とパターン処理の融合（照合）：聴覚
キーワード PDP / Society of Mind / コネクショニズム / 音声認識・理解 / 音声の個人性処理 / 音声入出力（認識・合成） / 特徴抽出 / 識別関数 / 言語解析 / 通訳・翻訳電話 / メディア変換 / 抑揚 / 方言・なまり
- 2.3.4 C4 記号処理とパターン処理の融合（照合）：感性・文化・応用
キーワード 感性情報処理 美声・悪性 / 感性情報処理 表情認識 / 感性情報処理 名文・悪文
- 4 D 不完全情報問題の解法
- 2.4.1 D1 不良設定問題解法：生理・心理
キーワード 運動軌道の生成 / 汎化 柔らかな認識
- 2.4.2 D2 不良設定問題解法：視覚
キーワード 3次元形状の復元 / 拘束伝播と弛緩法 / 正則化 不完全な情報の再生 / 「高速性」計算の複雑さへの対応 / スループットの高い処理方式 / 状況理解 / 汎化 柔らかな認識 / 文脈（シーン）の理解
- 2.4.3 D3 不良設定問題解法：聴覚
キーワード 正則化 不完全な情報の再生 / 「高速性」計算の複雑さへの対応 / スループットの高い処理方式 / 音声からの調音器官形状の推定 / 音声知覚の運動説 / 状況理解 / 対話処理 / 汎化 柔らかな認識 / 話者の背景の利用
- 2.4.4 D4 不良設定問題解法：感性・文化・応用
キーワード デザイン
- 5 E 能動的理解（active interpretation）
- 2.5.1 E1 能動的理解（active interpretation）：生理・心理
キーワード 知覚循環 相互作用としての認識 / 注意、選択的注意 情報の取捨選択 / 耳音響放射
- 2.5.2 E2 能動的理解（active interpretation）：視覚
キーワード 知覚循環 相互作用としての認識 / Active Interpretation / 能動的理解 / 概念駆動型処理（目的意識を持った処理） / 注意、選択的注意 情報の取捨選択 / 注意（の制御） / 注視の機構 / 注視 認識対象の検出 / インテリジェントビジョン / 状況の限定 / 知的ロボット / 宇宙作業ロボット
- 2.5.3 E3 能動的理解（active interpretation）：聴覚
キーワード 知覚循環 相互作用としての認識 / Active Interpretation / 注意、選択的注意 情報の取捨選択 / 概念駆動型処理 / 言語情報の利用による予測 / 混合音声 / カクテルパーティ効果 / 状況の限定 / 対話処理 / 不特定話者対応 / 話者の背景の利用
- 2.5.4 E4 能動的理解（active interpretation）：感性・文化・応用
キーワード 車のエレクトロニクス化
- 6 F 自己学習
- 2.6.1 F1 学習（発達の）：生理・心理
キーワード 可塑性 / 自己組織化 / 連想記憶 / 意味付け / 志向性 不良設定問題の解法 / 帰納的推論のできる学習方式 / 論理、類推
- 2.6.2 F2 学習（発達の）：視覚
キーワード 帰納的推論のできる学習方式 / 志向性 不良設定問題の解法 / 知識獲得 / 論理、類推 / 進化
- 2.6.3 F3 学習（発達の）：聴覚
キーワード 志向性 不良設定問題の解法 / 帰納的推論のできる学習方式 / 言語の獲得 / 言語の学習 / 論理、類推
- 2.6.4 F4 学習（発達の）：感性・文化・応用
キーワード 感覚代行、リハビリテーション / 情報機器の個人対応化 / 帰納的推論のできる学習方式 / 多様な利用を可能とする、アルリズム自動生成とカスタマイズ技術 / 論理、類推

—— 禁無断転載 ——

平成 2 年 3 月 発行

発行所 財団法人 日本情報処理開発協会

東京都港区芝公園 3 丁目 5 番 8 号

機械振興会館内

TEL (432) 5405 (代表)

印刷所 三協印刷株式会社

東京都渋谷区渋谷 3 丁目 11 番 11 号

TEL (407) 7316

01-R014

