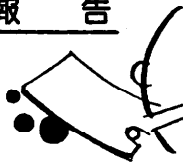


報 告



パネル討論会

コンピュータビジョンと視覚情報処理

昭和 58 年度後期第 27 回全国大会† 報告

パネリスト

鳥脇純一郎¹⁾, 杉原 厚吉²⁾, 辻 三郎³⁾
 福島 邦彦⁴⁾, 外山 敬介⁵⁾
 司会 杉江 昇⁶⁾

コンピュータ・ビジョン (CV) の研究が始められてから、20 年以上の歳月が経過し、多くのすぐれた成果が生みだされた。他方、人間やサルなどの高等動物における視覚情報処理 (ヒューマン・ビジョン, HV) の仕組みについての研究も、心理学や神経生理学、バイオサイバネティクスなどの分野で着実に進展している。

本パネル討論会では 2 つの分野での研究の発展の経緯をふり取り、問題提起と討論を行い、今後の研究への指針をえることを目的とした。

まず、司会者から序論として、2 つの分野での研究の流れと、両者のかかわりについて述べられた。ついで CV の分野から、鳥脇純一郎氏は 3D 画像の特徴抽出の手法を中心に論じ、生体の視覚情報処理との比較の資料を提供した。杉原厚吉氏は、CV における、数学的原理に基づいた処理方式を取りあげ、コンピュータショナルなレベルでの HV 研究との深いかかわりを指摘した。辻三郎氏は、CV における知識の有用性を指摘し、画像解析モジュールと推論・知識モジュールとの分離の重要性を論じた。福島邦彦氏は、視覚情報処理の解明には、バイオサイバネティクス研究が不可欠であり、またその成果は工学的目的にたいして大きな可能性を秘めていると論じた。最後に外山敬介氏は、神経科学の立場から、高次視覚中枢や連合野で発見された高度な視覚情報処理について紹介し、処理の階層性につき CV のそれとの対応を指摘した。

回顧と展望

杉江 昇

ここで取りあげる 2 つの分野は、これまであるときは密接に、あるときは比較的疎遠にかかわりあいながら、それぞれに発展をとげてきた。現在は、後に杉原氏が論じるように、両者の関係が、表面的ではなく深層でこれまでになく強くなっている段階にあるものと思われる。とはいえ、その関係は、2 つの分野の比較的限られた部分においてであり、今後両者の関係がどのように発展するのかは、予測しがたい。

2 つの分野の発展の経過とかかわり合いについて、3 つのレベルにつき整理してみよう。3 つのレベルとは、局所の特徴抽出、特殊な知識に依存しない大域的特徴抽出、知識関連プロセスである。このような処理の階層性は、最近の知見にもとづく高等動物の脳での機能分散的、階層的な視覚情報処理のアーキテクチャとも対応するものである。

(1) 局所の特徴抽出

60 年代前半においては、HV の分野では、大脳視覚領において、明暗のエッジ抽出細胞などが発見され、CV の分野で始まったエッジ抽出オペレータの研究との関連が意識されていた。その後しばらくは HV と CV の研究はたがいにほぼ独立に推移した。すなわち HV では色のエッジ抽出、両眼視差抽出、動きの抽出、奥行運動抽出、空間周波数チャネルなどについて知見が積み重ねられた。また CV では、細線化、明暗・色の一様な領域の抽出法などが開発されるとともに、HV には対応するもののないような処理法も考案された。すなわち、スポット光やストライプ光照射による能動的距離情報抽出法や、CT 画像などに関連した 3D 画像データからの特徴抽出法などである。70 年

† 日時 昭和 58 年 10 月 18 日 (火)、12:30~14:45

場所 名古屋大学豊田講堂

1) 名古屋大学、2) 名古屋大学、3) 大阪大学、4) NHK、

5) 京都府立医科大学、6) 名古屋大学

代後半からは HV と CV とのかかわりは、ふたたび密接になる。網膜でのエッジ抽出機構を F^2G オペレータにより定式化したことが、CV の側で大きな関心をもって迎えられた。HV の側の研究にも大きな刺激をあたえている。

(2) 大域的特徴抽出

HV と CV とのかかわりは、最初は比較的疎遠であったが、最近になって、両者の関係はきわめて密なものになった。すなわち、HV の分野では、両眼立体視、運動立体視、テクスチャ知覚、色や大きさの恒常性などについて、多くの心理学的知見が蓄えられてきたが、局所的特徴抽出のレベルでの神経生理学的知見の飛躍的増大を背景として、70年代後半からその協調と競合をとまなう処理機構について、コンピュータショナルなレベルや、神経回路網のレベルでのバイオサイバネティクス研究が急速に進んでいる。他方 CV においても、2次元画像から3次元情報を復元する各種理論（線画解釈、勾配空間、フォトメトリック・ステレオ、緩和法）や、テクスチャ領域の抽出などが、大きな関心事となっており、本質的には HV との差がないと見せる場面もすくなくない。

(3) 知識関連プロセス

知識の獲得、表現、利用に関するプロセスについては、60年代前半に、HV の分野で提案された学習機械パーセプトロンなどが CV の側での大きな関心と呼んだ。しかし、パーセプトロンの能力には一定の限界があることが指摘されるにおよび、CV の側での関心は冷却し、今にいたっている。しかし、HV の側では、パーセプトロン以降、自己組織化神経回路網について、神経生理学的にも、バイオサイバネティクスのにも多くの貴重な研究が積み重ねられ、福島氏のネオングニトロンや中野氏のアソシアトロンのようなすぐれた成果がえられている。CV の側では、実例からの対象物の概念形成、知識表現のためのフレーム理論、ブラックボード方式の知識利用などの研究がある。今後、高度な認知のレベルで、HV と CV とがどのように影響しあっていくかが、見ものである。

コンピュータ・ビジョンにおける 画像特徴の抽出

鳥脇純一郎

1. ま え が き

生体の視覚情報処理の本質的機能は、3次元空間情報の取得にあり、人工知能における画像理解ではごく

少数の2次元画像からの3次元空間の状態の理解を主目的とする。一方、電算機による画像処理においても、計算機断層像や顕微鏡像、などの3次元画像データの普及、増加に伴い、3次元画像処理への要請が急速に高まってきている。しかし、これら三者の性格やそこで用いる手法は、互いにかなり異なるものである。

本講演では、3次元デジタル画像(3D画像)の電算機による認識の手法—特に濃淡画像のセグメンテーションの段階での手法を、ごく簡単に紹介する。

2. 3D画像処理の特徴

3D画像処理の対象となる画像は次の形をもつ¹⁾。

$$F = \{f_{ijk}\}, f_{ijk} = \text{点}(i, j, k) \text{の濃度値}$$

人間は、これを $F_k = \{f_{ij,k}\}, k=1, 2, \dots, K$ なる K 枚の2次元(2D)画像の組として観測し、頭の中で3次元空間の情報を組立てる。一方、電算機では F を3次元データとして直接に処理して、3次元的情報を得る。しかし、処理システムや手法の性能評価を行う人間は上記の2D画像を経由したやり方を取るため、それをやり易くする研究開発用ツールとしての3D画像の表示法(グラフィクス)も重要な問題となる¹⁾。

3. セグメンテーション

3D画像のセグメンテーション(あるいは対象物検出)の考え方は、基本的には2D画像の場合と同様であり、輪郭検出、領域形成、テンプレートマッチング、等がある。ここでは、輪郭検出のみを扱う。その基本方針も2D画像の場合と同様である。すなわち、**前処理(平滑化等)—差分型フィルタ—しきい値処理—連結成分処理—薄面化、距離変換とスケルトン抽出、または境界面追跡**(あるいは、この部分列)である。以下、この各段の処理のあらましを紹介する。

4. 3D画像の輪郭検出

(1) 差分型(エッジ検出)フィルタ: 2次元の各種のフィルタが容易に3次元に拡張される²⁾。ただし、方向特性のとり方や重み係数の具体的な値については、2次元の場合よりも自由度がかなり大きい。また、複雑な重みをもつものは、処理時間の増加が著しい。

(2) しきい値処理と連結成分抽出: 上記フィルタ出力を適当なしきい値で2値化し、連結成分を抽出する。両処理とも2次元の場合の手順が直ちに拡張できる¹⁾。

(3) 薄面化: 厚みをもつ3D連結成分の中心面(線)を表わす曲面(曲線)を抽出する¹⁾。2次元の細線化に対応するが、3D画像の幾何学的性質が著しく

複雑になるため、2次元の場合との差違が最も大きくなる部分である。基本的な考え方は次のようになる。すなわち、2値3D画像中の値1の画素の消去可能条件(図形トポロジ保存条件)をテストして、消去可能な画素を逐次的*に0に変えていく。この消去作業を連結成分の上下、左右、前後から順に一層ずつ行い、連結成分が3次元単体を含まなくなったら(消去可能画素がなくなったら)停止する¹⁾。

現在2種類のアルゴリズムがあるが、その性質には不明の点が多く、また、処理時間も極めて長い¹⁾。

(4) 距離変換とスケルトン抽出: 定義と並列型アルゴリズムは、2次元の場合の直接的拡張でよい¹⁾。逐次型手順もほぼ2次元の場合と同様であるが、その処理速度は複雑な特性を示す。これらの前提となる距離関数については3次元独自の考察が必要であるが、基礎的性質はほぼ説明されている^{1), 5)}。

(5) 境界(輪郭)面追跡: 連結成分の表面にある画素を一定順序で辿って、その位置をリストに出力する。2D画像の境界追跡(border following)に対応する処理であるが、追跡対象が閉曲線から閉曲面に変わるため、2次元の場合とはまったく異なるアルゴリズムになる。基本的手順はほぼ明らかにされている^{3), 7)}。

5. 処理時間の実験例

対象画像 F の大きさ(画素数)を $M \times N \times L$ とする。これは、大きさ $M \times N$ の2D画像 L 枚分に等しいデータ量を持つから、 F の処理時間が上記2D画像1枚の処理時間の L 倍にまで増加するのは当然であり、それ以上の増加分があれば、それが3D画像固有の複雑さに基づく増加分とみなされる。そこで、画像 F に3Dアルゴリズムを適用したときの処理時間を t_{3D} 、この F を大きさ $M \times N$ の2D画像 L 枚とみて、その各々に、対応する機能の2次元処理アルゴリズムを適用したときの処理時間の合計を t_{2D} として、 $r = t_{3D}/t_{2D}$ を3D処理アルゴリズムの処理時間の増加因子とよぶ。現在利用可能な3Dアルゴリズムに対する増加因子 r の実測例を示すと、次のようになり、薄面化が際立って複雑な処理であることがわかる⁴⁾。

差分フィルタ: $r = 2.1 \sim 3.8$,

薄面化: $r = 6.9 \sim 20.2$, 距離変換: $r = 1.8 \sim 2.9$

輪郭追跡: $r = 1.1 \sim 1.7$, 図形収縮: $r = 1.9 \sim 2.9$

6. むすび

セグメンテーションの後には形状特徴抽出, パター

ン分類, 記述の導出, 等の処理がくるが、そこでは各問題領域に固有の知識が用いられることが多い。知識の利用については、本パネル討論の別の講演にゆずる⁶⁾。

3D画像処理は、特にデータ量の大きさと対象の構造の複雑さゆえに、電算機のみでなく人間にとっても2D画像処理よりもはるかに難しい。それだけに電算機特有の処理への期待も大きい。しかし、現在は最小限の処理のツールが整えられたにすぎず、アルゴリズムの性質の解明は今後の研究に待たねばならない。

参考文献

- 1) 鳥脇純一郎: 高次元デジタル画像の構造解析に関する基礎的研究, 昭57年文部省科研費研究成果報告書(課題番号56460105)(Mar. 1983).
- 2) 鈴木, 鳥脇: 3次元デジタル画像上のエッジ検出オペレータの導出とその能力の評価, 情報処理学会論文誌, Vol. 25, No. 2, pp. 243-250 (Mar. 1984).
- 3) 松本, 横井, 鳥脇, 福村: 3次元デジタル画像の境界追跡, 信学技報PRL研, PRL 83-10 (May. 1983).
- 4) Toriwaki, J. and Yokoi, S.: Algorithm for Skeletonizing Three-dimensional Digital Binary Pictures, Proc. the SPIE Symp. No. 435, pp. 2-9 (Aug. 1983).
- 5) Yamashita, M. and Honda, N.: Distance Functions Defined by Variable Neighborhood Sequences, Tech. Rep., Toyohashi University of Technology, Toyohashi (Sep. 1983).
- 6) 辻 三郎: コンピュータビジョンと知識, 情報処理 Vol. 25, No. 9 p. 967 (1984).
- 7) Srihari, S.N.: Representation of Three-dimensional Digital Images, Computing Surveys, Vol. 13, No. 4, pp. 399-424 (Dec. 1981).

ヒューリスティックスと数理との分離

杉原 厚吉

コンピュータ・ビジョンとは、身の回りの世界を見て認識するという人間の視覚機能を機械に代行させようとする学問である。人間の視覚機能は、人が人生経験を通して蓄積してきた“知識”を総動員した複雑な処理過程によって実現されているように見える。そのため、これを機械で代行させようとするヒューリスティックな処理にたよらざるを得ないという印象が、特に研究の初期には強かった。しかし、この一見複雑な過程の中にも、比較的単純な数学的原理のみに基づいた機構で代行できる部分が数多く存在することが次

* アルゴリズムの逐次型, 並列型は2次元の場合と同様にして定義される。

第にわかってきた。このような数理的裏付けを伴った処理単位を一つ一つ分離し、基本モジュールとして蓄積するという努力が、コンピュータ・ビジョンの主要な研究方向の一つである。

一方、生体の視覚情報処理の研究において、computational なレベルで機能を考えるという立場が、近年盛んになってきた。これは、網膜に投影された像の中に原理的にどのような情報が含まれているのかを数学的にまずはっきりさせようとする立場である。これは、最終的には、神経回路モデルの模索や生理学的研究のための指針を与えようとするものであるが、このcomputational な立場のみを切り離してみると、生体とは独立に工学独自の計算機構を構成しようとするコンピュータ・ビジョンの立場とほとんど差がないことがわかる。今や、数理的裏付けのある視覚モジュールを蓄積しようとする努力は、コンピュータ・ビジョンの研究者と、computational な立場での生体視覚情報処理の研究者との共同作業となっている。

基本モジュールの蓄積は、画像中の2次元の特徴から立体形状を決定するという処理段階で、特に盛んに行われている。これは、この段階での機能の入出力関係が、数学的に明確な形で定義しやすいからであると思われる。これらのモジュールは、構造的情報を抽出するものと計量的情報を抽出するものの二つに大別される。

構造的情報を抽出するモジュールの代表例は、線画からそこに描かれている多面体の構造（面・稜線・頂点がどのように接続しているかという位相的な性質）を抽出する計算機構である。約15年前に、GuzmanがMITで、線画の線で区切られた領域を同一物体ごとにまとめるプログラムを作ったが、これは数理的な処理からはほど遠いものであった。このプログラムは根拠の薄弱な多数の規則のよせ集めであったばかりでなく、設定した問題そのものが数学的に明確でなかった（2個の物体が積み重なった状態と、もともとこの形でくっついている1個の物体とは、線画からは区別できないから、プログラムがどのような判定をしたとき正しいとみなすべきかはあいまいであった）。その結果、どのような条件下でうまく行くのかという適用範囲が不明確なため、安心して使えるものではなかった。

これに対して、Huffman等は、頂点のみえ方の一覧表を使って、線画中の線の物理的屬性（凹、凸など）を決定する方法を提案した。これは、線画が正しく多面体を表すための一つの必要条件を明らかにしそ

れを利用したものであるため、この方法で処理が行き詰まる線画に対しては、誤った線画であると“自信”をもって判定できるようになった。しかし、この方法でうまく処理できても、結果が正しいという保証はなかった。その意味で、これは半歩だけ数理的処理に近づいたものだけということができる。

その後、線画が正しく多面体を表すための必要十分条件が線形代数を使って定式化され、その結果、線画が正しいか否かの判定作業は、線形計画問題の実行可能解の存在を調べる作業に帰着できることがわかった。また、この方法はこのままでは数学的に厳密過ぎて誤差を含むデータに対しては使えなかったが、この過剰な厳密さを克服する方法も与えられ、その意味で実用的にも解決されている。

構造的情報の抽出では、このほかに輪郭線や表面模様から曲面の凹凸を決定するものなどがある。

一方、計量的情報を抽出するモジュールは、2次元画像に現れる各種の手がかりから立体の計量的情報（たとえば面がどちらの方向に何度傾いているかという情報）を抽出しようとするもので、多数のモジュールが蓄積されている。

既知の照明環境で光学的性質のわかっている物体の明るさを観測すれば、面上の各点の傾きに関する自由度を1だけ減らすことができる。そこで、照明状態を変えて撮った複数の画像を同時に利用したり、明るさの値と物体の大域的情報（表面がなめらかな曲面から成り立っているという情報など）を組み合わせたりすることによって、立体形状を抽出することができる（濃淡からの形状決定）。また、面が既知の小図形でおおわれている場合には、小図形のみかけのひずみから面の傾きを復元できる（表面模様のひずみからの形状決定）。さらに、表面がランダムな模様でおおわれていても、その統計的性質から面の向きを決定することもできる。たとえば、表面が多くひっかききずでおおわれ、そのきずの向きの分布があらかじめわかれば、みかけのきずの向きの分布から立体を復元することができる（テクスチャの統計的性質からの形状決定）。

そのほかにも、両眼立体視、焦点合わせによる距離計測、遠近画法の効果からの形状決定、鏡面反射効果を利用した形状決定、動く物体の複数画像からの動きと形状の決定、動く物体の速度場（オプティカル・フロー）からの動きと形状の決定、などの数理的手法がある。

このような基本視覚モジュールが数多く蓄積されれば、実際に機械で視覚システムを構成しようとする際に、メニューとして役立つであろう。実際の場面では、画像の分解能や雑音、処理対象の複雑さ、処理の目的などに応じて、多くのヒューリスティックな処理も必要であろうが、数理的裏付けのある視覚モジュールをその中に少しでも多く組み込むことができれば、それだけシステム全体の信頼性も高くなるであろう。

また、基本モジュールの蓄積によって、2次元画像の中には立体形状に関して非常に多くの手がかりが含まれていることがわかってきた。このことは、見方を変えれば、立体の絵を不注意に描くと現実には存在しない絵を描いてしまう危険性が大きいことを意味している。このような観点からの成果の再整理は、絵画の創作、写真の撮影、コンピュータ・グラフィックス・ユーティリティの使用などに際して、有益な基本技術を提供してくれるであろう。

コンピュータビジョンと知識

辻 三郎

対象世界を撮した画像をコンピュータが解析して、その構造・状態を記述する世界モデルを作成するのがコンピュータビジョンのひとつの目標である。最近の傾向としては、画像理解という言葉で代表されるように、知識を利用して画像解析、モデルの作成を行う人工知能的な色彩が濃くなった。

(1) 知識の種類

われわれの住んでいる三次元世界では、カメラの角度、光源の方向、物体の面の向きや反射率などからカメラの写した像は決定される。このような物体の見え方の法則を覚えておくと、コンピュータはこれを用いてシーンの特徴を調べることができる。

コンピュータビジョンで用いる知識は、この他多種多様なものがある。まず物に対する知識がある。例えば、椅子には大きさ、形など種々のものがあるが、人間はそれらに共通した椅子としてのイメージを持っている。これに対応する共通モデルの知識がある。航空写真を調べるには地図があると便利である。地図とか構造図のようなデータベース、さらにどのようにデータベースを用いるかは重要で有用な知識である。われわれが判定できない医用写真も、経験の豊富な専門医は病状を推定できる。これは、専門分野の知識で、最近特にその重要性が指摘されている。

(2) 知識の構造

従来の画像処理でも、これらの知識の一部は利用されていた。しかし、その多くは解析プログラムの各所に埋込まれており、その対象以外に応用することは難しい。また画像解析の中間段階で推論をし、その結果に基づいて処理の制御をするといった柔軟な処理の例は極めて少ない。

これらの欠点を補うには、画像解析モジュールと推論・知識の部分を分離させることが有効である。利用分野別の知識を独立した構造にし、知識獲得を重視する知識工学的アプローチが次々と試みられている。

(3) 世界モデルの表現

世界モデルの記述を得るというコンピュータビジョンの目標に対して、現在までの研究はまだ不十分である。

個々の物体の3-Dモデルは、現在までは階層構造を用いた幾何モデルが一番よく利用されている。また最近では、距離画像を利用したモデルの作成が盛んになった。このような幾何モデルを基にして、各種の推論をするコンピュータビジョンは興味深いテーマであるが、そのモデルが柔軟性を欠くことは否めない。樹木や岩石など複雑な形の物体が多数存在する自然環境をどのように表現するかについては、ほとんど手がつけられていない状態である。コンピュータグラフィックスで用いられているフラクタル理論を、コンピュータビジョンに応用しようという試みはこの意味では興味深い。

コンピュータ・ビジョンと バイオサイバネティクス

福島 邦彦

生物の脳は、コンピュータを始めとする現在の情報処理装置にはない高度の情報処理能力をもっている。このような情報処理能力がいったいどのようなメカニズムに帰因しているかを知ることは、生理学や心理学者だけでなく、工学者にとっても大きな関心事である。生物の脳の中で行われている情報処理のメカニズムを知ることができれば、それは直ちに、新しい情報処理装置の開発につながるからである。

とは言っても、脳の中には膨大な数の神経細胞があり、これらが複雑につながり合って一つの大きな神経回路を構成している。脳の中の神経細胞の数は人間の場合には百億個以上にも達すると言われている。したがって微小電極を用いて単一神経細胞の反応を分析している現在の神経生理学的手法だけでは、断片的な知

識は数多く得られても、脳の神経回路の全容を知ることとはほとんど不可能に近い。

そこで、数学的あるいは工学的立場からの理論的研究、すなわちバイオサイバネティクス的研究の重要性が目されるようになってきた。従来の生理学や心理学では主として分析的手法を用いるのに対し、バイオサイバネティクスでは合成的手法によって脳を解明しようとするのである。すなわち、調べようとする機能に関して脳と同じ反応を示す神経回路モデルを構成することから始める。モデルの構成に際しては、生理学や心理学の実験によって解明されている事実はできるだけ忠実に取り入れるが、まだ解明されていない部分に関しては大胆な仮説を導入する。このようにして構成したモデルの性質を計算機シミュレーションや数学的解析によって調べる。もしこのモデルが脳と違う反応を示した場合には、採用した仮説に誤りがあったと考えて、仮説を修正していく。このような手順を何回も繰り返すことによって脳における情報処理の基本原則を探求していく。一旦このようなモデルが出来上がれば、脳における情報処理にとって本質的な役割を果たしている要因が何であるかを容易に見極めることができ、脳研究の発展に役立つのみならず、それはそのまま、神経系の長所を取り入れた情報処理装置の設計原理の開発にもつながる。

生物の優れた性質を新しい機械の設計原理として工学的に利用しようとする研究としては、1960年にSteeleが提唱したバイオニクス(生体工学)を連想される方も少なくないと思う。しかし初期のバイオニクスと現在のバイオサイバネティクスとは、そのアプローチがだいぶ変わってきている。バイオニクスの初期には、生理学や心理学の文献を工学者にわかる用語に翻訳しさえすれば新しい機械の設計図が得られると安易に考えた工学者が少なくなかったように思う。このような安易な考えでバイオニクス研究に入ってきた工学者の多くは、生理学や心理学においても、工学者の欲するような形での脳の解明はそれほど進んでいないことがわかってくるにつれて、バイオニクス研究から去っていった。これに対して現在のバイオサイバネティクス研究者は、脳のメカニズムの解明を生理学や心理学者だけにまかせておくのではなく、合成的手法を武器に、自らも積極的にその解明にたずさわろうとしているのである。

このようなバイオサイバネティクス研究の結果、脳の種々のメカニズムに関する神経回路モデルが提唱さ

れた^{1)~4)}。そのうちのいくつかのものを以下に列挙してみよう。網膜や眼球に関しては、順応特性を考慮した視細胞のモデル、側抑制機構に基づいたコントラスト検出のモデル、網膜の中心部と周辺部とで解像度の異なる不均一網膜のモデル、眼球運動や注視点移動のモデルなどがある。一方大脳の視覚野に関しては種々の特徴抽出機構のモデルが作られている。例えば、入力パターンに含まれる直線成分の抽出、曲率の抽出、直線の端点・交点・折点などの抽出、あるいは運動情報の抽出などの能力をもったモデルや、両眼視による奥行き情報の抽出、両眼像の対応関係を求める機構のモデルなどもある。また、成長にともなって神経回路が次第に形成されていく機構、損傷を受けた神経回路が自己修復していく機構、外界の状況に適応して自らの構造を少しずつ変更していく自己組織化の機構、記憶や学習の機構、条件反射の機構などに関するモデルも少なくない。またこのような機能をいくつか組み合わせ、パターン認識機構のモデルや連想記憶のモデルを作る試みも広くなされている。

このようなモデル構成を進めていっても、すべての面で脳と完全に同じモデルを完成するまでには、この先長い年月を要するであろう。しかしこのことは、決して、脳と完全に同じモデルが出来上がるまでは工学的に役立つ成果が得られないことを意味しているのではない。それどころか、最終目標に到達する過程で次々と生まれてくる個々のモデルは、ある特定の機能に関しては脳と同じ反応特性を示すものであるから、そのモデルの考え方は、新しい情報処理装置の設計原理としてその都度工学的に利用していけるのである。

例えば、パターン認識機構の神経回路モデルとして筆者らが提唱したネオコグニトロン^{5),6)}も、バイオサイバネティクス研究の成果の一つである。このアルゴリズムを用いれば、従来よりも高度のパターン認識がより単純な操作で実現できる。さらにネオコグニトロンは、学習能力をもっているため、従来の機械よりも、もっと人間的な融通性を示す。

バイオサイバネティクス研究は、工学的目的のためには一見遠回りのように感じられるかもしれないが、通常の研究手段では思いつかないような高度の情報処理方式を生み出す可能性を秘めた有望な研究分野である。

参考文献

- 1) 福島邦彦: 視覚の生理とバイオニクス, 電子通

信学会編 (1976).

- 2) 福島邦彦: 神経回路と自己組織化, 情報科学講座, 共立出版 (1979).
- 3) 福島邦彦, 大串健吾: 視覚系のモデルによる研究, 計測と制御, Vol. 18, No. 1, pp. 22-29 (1979).
- 4) 福島邦彦: 学習のモデル, 医用電子と生体工学 (日本 ME 学会誌), Vol. 19, No. 5, pp. 319-330 (1981).
- 5) Fukushima, K. and Miyake, S.: Neocognitron: A New Algorithm for Pattern Recognition Tolerant of Deformations and Shifts in Position, Pattern Recognition, Vol. 15, No. 6, pp. 455-469 (1982).
- 6) 福島邦彦, 三宅 誠, 伊藤崇之: ネオコグニトロンの原理を用いた数字パターン認識, 電子通信学会論文誌 D, J66-D [2], pp. 206-213 (1983).

高次視覚中枢の情報処理

外山 敬介

目で捉えられた視覚情報は視神経、中継核(外側膝状体)を経由して、第一次大脳視覚中枢に送られる。第一次の視覚中枢では図形の位置、形、運動方向などの幾何学的特性や色(光の波長)などの物理学的特性が分析されることが知られている。

第一次の視覚中枢で処理された情報は高次の視覚中枢で処理され、更に連合野で統合されて視覚認知に至ると考えられている。最近、高次の視覚中枢や連合野の情報処理の機構が進展し、第一次の視覚中枢には見られない高度の情報処理が行われ、第一次視覚中枢、高次視覚中枢、連合野の間に階層的な機能分化があることが明らかにされている。ここでは、高次の視覚中枢の働きについての最近の知見を紹介し、computerによる視覚情報処理との対応について議論することにした。

高次の視覚中枢について先ず最初に述べておきたいことは、機能分化があることである。つまり高次の視覚中枢はいくつかの領野(5~6カ所)に区分され、それぞれの領野は別個に第一次視覚中枢から視覚情報を受け、独自の視覚情報の分析を行っている。高次視覚中枢の中でもサルのV4野とネコClare-Bishop野の研究が特に進展している。前者は色の認知、後者は三次元の運動の認知に関係するが、両者とも第一次視覚中枢から送られてくる情報の中から、ある種のパラメータの変化に対して不変で、より本質的な視覚情報を取り出す働きがあることが知られている。

V4野の細胞は受容野(神経細胞が反応する神野の

部分)に示された視覚刺激の色に選択的に反応する。例えば“赤”細胞は白色光の照明の下で赤い色紙を受容野に示すと強く反応する。次に受容野に白い紙を置き、赤い光で照らすと、細胞は全く反応しない。受容野で受け取られた光は物理学的にはまったく同一であるのに、細胞の反応は大きく異なっている。われわれにも前者の条件下では紙は赤色に、後者の条件では白色に見える。つまりわれわれは紙の色を反射されてくる光の波長によって判断するのでなく、いろんな波長の光に対する紙の反射率によって判断するのである。このような機構により、照明の光のスペクトルにまどわされることなく、視覚対象の色を判断することができる。この働きは“色の恒常性”と呼ばれる。雨の日にカラー写真を撮ると思ったよりもはるかに色が悪くがっかりすることが多いが、これも“色の恒常性”により、雨天の日でも晴天と同じように風物がわれわれには色鮮やかに見えるためである。

V4野の細胞の反応もこれと同じ“色の恒常性”を備えている。このことからわれわれの“色の恒常性”はV4野の働きから生まれてくるのではないかと考えられる。このような働きはどのようにして生まれるのであろうか。V4野の細胞の受容野は実は中心部と周辺部の二重の構造を持ち、中心部で視覚対象から反射される光を、周辺部では背景から反射される光を捉える。赤色細胞は中心部で捉えられた光のスペクトル分布が周辺部の光のスペクトル分布よりも長波長例に大きくずれている時に反応する。つまり“赤”細胞は視覚対象と背景の光のスペクトル分布を比較することによって、“赤い色”に対する反応の恒常性を獲得するのである。V4野には“赤”細胞の他にも、青、緑、紫などいろいろな色に対して恒常性を備えた細胞がある。

ネコのClare-Bishop野には三次元の運動の認知に関与していると考えられる細胞がある。三次元の運動の認知には二つの手がかりが利用されている。その一つは左右の目で捉えられる運動のずれ(運動視差)である。例えば視覚対象が近づいて来る時には網膜の映像は互いに発散する方向(右眼は右、左眼は左)に移動する。逆に視覚対象が遠ざかれば、映像は互いに収斂する方向(左眼は左、右眼は右)に移動する。等距離で左右に移動した時には左右の眼の映像の動きは全く一致し、このような運動視差は生じない。

もう一つの手がかりは、映像の大きさの変化である。視覚対象が近づけば、距離に逆比例して網膜の映

像は大きくなる。しかしながら距離を考慮して大きさを判断する大きさの恒常性により、視覚対象の大きさは一定であるように見える。

筆者らはこの二つの視覚的手がかりを再現する三次図形刺激装置を用いて、Clare-Bishop 野を調べ、二つの手がかりのいずれか一つあるいは二つの組合せに反応する細胞を見出した。

例えば視覚対象の接近を認知すると思われる細胞としては、接近の運動視差に反応する細胞がある。この細胞は左右それぞれの目からの視覚入力に対して、左方向と右方向の運動方向の検出器とその出力の論理積をとる“AND”ゲートを備えているので、左右の目に発散方向（接近の）運動視差が与えられた時に反応する。また大きさの増大に反応する細胞、接近の運動視と大きさの増大の組合せに反応する細胞もある。

これらの細胞は大きさの変化の検出器やその出力と運動視差の出力の論理積をとる“AND”ゲートを備えていると考えられる。

ここで特に指摘しておきたいことは、Clare-Bishop 領の細胞が、左右の目からの視覚情報を組合せる（論理積をとる）ことにより、運動視差という三次元運動の認知のための重要な手がかりをとり出していることである。運動視差を更に大きさの増大と組合せる細胞があり、ここで三次元の運動の認知が完成する。また運動視差と大きさの増大を“NAND”ゲートで組合せることにより、前後方向の運動に伴わない純粋の大きさの増大に反応する細胞もある。この細胞は真の大きさ変化を見かけ上の大きさの変化から識別して反応するので大きさの恒常性に関係する細胞と考えられる。

この意味で Clare-Bishop 領の細胞もまた、左右の目からの情報を組合せることによって、視覚対象の運動や大きさの変化についての普遍的な情報をとり出していると言える。つまり V4 野や Clare-Bishop 領は異種の視覚情報を組合せることにより、普遍性を持つ情報を取り出すという点で第一次視覚領より高度な視覚情報処理を営むと考えられる。

各パネリストから説明のあった computer vision の手法にもやはり同様の情報処理の階層性があるように思われる。まず極めて local で同種の情報の処理を行い、次いで情報処理の範囲を地理学的にも、また種類という点でも逐次拡大して、見かけ上の情報からより普遍性を持つ情報をとり出すことが、工学的な視覚情報処理にも不可欠であるように思われる。

あとがき

CV と HV とのあいだの相互の関心が、これまでになく高まっている時期に、このようなパネル討論会を行うことができ幸いであった。本企画を立案された、名大福村教授に謝意を表したい。パネリストはいずれも、それぞれの分野で、第一線の研究者であるだけでなく、CV と HV の両方に深い関心と知識を持っている方たちで、有益なご意見をうかがうことができた。

なお、限られた時間ではあったが、総合討論において、記号的処理とアナログ的処理、並列処理、能動的処理、フィードバック処理などについて、パネリスト間、パネリストとフロアとの間で討論が行われたことを付記する。