

## 画像学習システム MIRACLE-IV における 機能的特徴と視覚的特徴の対応付け†

松原 仁<sup>††</sup> 坂上 勝彦<sup>†††</sup>  
山本 和彦<sup>†††</sup> 山岸 健太郎<sup>††††</sup>

異なった形状を持つ同一物体のシルエット画像群から、その物体の内部構造ならびにその物体に適切な画像処理戦略とを学習する、画像学習システム MIRACLE-IV を作成した。MIRACLE-IV はモデル獲得部と処理戦略部の2つのサブシステムから成り立つ。物体の明示的なモデルや画像処理戦略はあらかじめ固定されたものではなく、学習の過程を通じてそれぞれに生成される。MIRACLE-IV は物体がヒンジ・スライド・ソリッドからなるという仮定に基づき、シルエット画像の形状の変化から物体の内部構造を推定しモデルを構築する。さらに、視覚的特徴が似ていれば機能的特徴も似ているであろう、という類似性に関する思い込みに基づいて機能的特徴と視覚的特徴の間の対応付けを試みる。その一連の画像群における視覚的特徴を抽出するための画像処理の手順は、前もって与えられるのではなく、試行錯誤を通じて経験的に学習される。本論文では特に MIRACLE-IV が類似性に基づき機能的特徴と視覚的特徴の間の対応付けを学習する過程を中心に述べる。

### 1. はじめに

現在の画像認識システムの多くは、あらかじめ決められた範ちゅうの画像の入力に対し、ある一定の手順で前処理を施して、その処理結果と固定されたモデル群との間で定型的な照合を行い、その分類の結果を最終的な出力とする、という枠組みになっている。人工知能の一環として人間の高度な視覚機能を機械化しようとする立場から見ると、この現状は決して満足すべきものではない。人間の視覚機能のような柔軟性に欠ける。その原因の一端はこれまでのシステムに学習機能がないためであると考えられる。

これからの新しい画像認識システムの研究は、システム自体が概念やモデルなどの各種の知識を経験によって獲得し、獲得した知識から目的や状況に応じて必要な情報を取り出して柔軟に認識を行うという枠組みが目標の1つになる。我々は目標とするこのような枠組みをとりあえずインテリジェント・ビジョンと呼ぶことにした。

我々は、異なった形状を持つ同一物体のシルエット

画像群から、その物体の内部構造および適切な画像処理戦略を学習する、画像学習システム MIRACLE-IV (Multiple Image Recognition system Aiming Concept LEarning—Intelligent Vision) の設計・試作を行った。このシステムは上記のインテリジェント・ビジョンに向けてのプロトタイプ・システムを目指したものである。

MIRACLE-IV はモデル獲得部と処理戦略部の2つのサブシステムから成り立つ。物体の明示的なモデルや画像処理戦略は前もって与えられず、学習の過程を通じてそれぞれに生成される。ここではまず MIRACLE-IV の概要について述べ、次にモデル獲得部と処理戦略部との対話を通じて機能的特徴と視覚的特徴の間の対応付けが学習される様子を例を挙げて説明する。

パターン認識と記号処理とを有機的に統合しようという試みは人工知能における重要な研究テーマであり<sup>1)</sup>、本研究もその中のひとつに位置付けられる。画像を対象とした概念学習をここでは画像学習と呼ぶ。従来からも画像学習を題材としたさまざまな研究がなされてきた<sup>2)~6)</sup>が、その多くは最初の段階でパターン情報を高次の記号に変換し、以後はその記号だけを処理の対象としている。それに対し、MIRACLE-IV では高次の記号処理の段階でも必要に応じて原画像のパターン情報が参照される。すなわち、パターン情報と高次の記号との間で情報が相互参照されていることが MIRACLE-IV の大きな特徴である。

† Correspondence between Functional Features and Visual Features on Visual Learning System MIRACLE-IV by HITOSHI MATSUBARA (Machine Inference Section, Machine Understanding Division, Electrotechnical Laboratory), KATSUHIKO SAKAUE, KAZUHIKO YAMAMOTO (Image Understanding Section, Machine Understanding Division, Electrotechnical Laboratory) and KENTARO YAMAGISHI (Electronics & Control Systems Research & Development Laboratory, Nippon Steel Corp.).

†† 電子技術総合研究所知能情報部推論研究室

††† 電子技術総合研究所知能情報部画像研究室

†††† 新日本製鐵(株)電子・制御研究開発センター

## 2. MIRACLE-IV の概要

MIRACLE-IV は、入力として提示される同一物体の異なる形状の一連のシルエット画像群から、対象物体の内部構造（モデル）と画像処理の戦略（ストラテジ）とを自動的に学習するシステムである。本論文で例とするシルエット画像群を図1に示す。シルエット画像の形状の変化から物体の内部構造を推定し、その推定に基づいてモデルを構築する。さらに、内部構造としての機能的特徴と視覚的特徴（「見え方」に相当する）との間の対応付けを試みる（図2参照）。これは、視覚的特徴が似ていればその機能的特徴も似ているであろう、という一種の類推<sup>7)</sup>（ここでは思い込みと呼ぶ）をしていることになる。思い込みによって内部構造の推定が効果的に行われる。その一連の画像群における視覚的特徴を抽出するための画像処理の手順は、前もって与えられるのではなく、試行錯誤を通じ

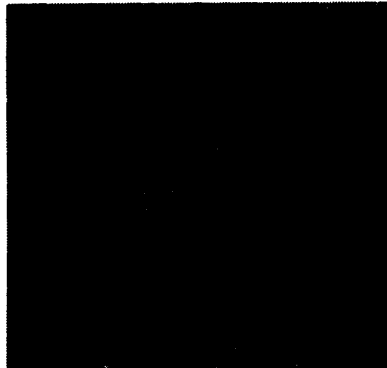


図1 例とするシルエット画像群  
Fig. 1 A series of silhouette images.

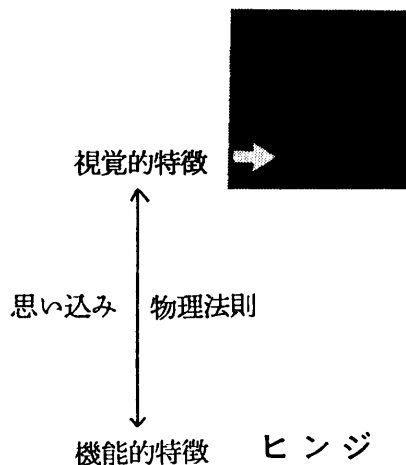


図2 機能的特徴と視覚的特徴の対応付け  
Fig. 2 Correspondence between a functional feature and a visual feature.

て経験的に得られる。

コンピュータによる学習を考える際には、何を既知としてそこから新たに何を学習させるかということが研究の鍵となる。何かを新しく学習するためには、あらかじめそれについて多くのことを知っている必要がある。形状が任意に変化するような物体の内部構造をコンピュータに学習させるのは極めて難しいので、変化の自由度を何らかの形で制限しておかなくてはならない。本研究では複数の可動接続部を持つ構造物<sup>8)</sup>を処理対象として取り上げ、その形状バリエーションをシステムに提示する。具体的には、対象とする物体はコンパス、はさみ、定規などの身近な器具である。MIRACLE-IV は、物体の形状を変化させる要因はヒンジ（支点）とスライドという2つの内部構造だけであることをあらかじめ知っているものとする。すなわち、対象物体は基本的にソリッド（剛体）から成り立ち、ソリッド間はヒンジもしくはスライドで結ばれている、ということを知っている。これを HSS 仮定（Hinge-Slide-Solid Assumption）と呼ぶ。本論文で機能的特徴と呼んでいるのは具体的にはヒンジでありスライドである。また、MIRACLE-IV は基本的な画像処理モジュール（ほぼ画像処理ソフトウェア・パッケージ SPIDER<sup>9)</sup>に含まれているものである）の存在とそれらの間の接続に関する物理的拘束を知っているものとする。

MIRACLE-IV はこれらの知識を用いて、入力の一連のシルエット画像群から、対象物体がいくつのヒンジ・スライド・ソリッドから成っているか、それらがどのような関係にあるかを学習し、同時に対象物体の視覚的特徴を処理するためのモジュール列の適切な並びを学習するシステムである。その結果として、機能的特徴と視覚的特徴の間の対応付けを学習する。

## 3. モデル獲得部と処理戦略部

MIRACLE-IV は2つのサブシステムから構成されている。モデル獲得部と処理戦略部である。前者は物体の内部構造を推定し機能的特徴を抽出する。後者は機能的特徴に相当する実際の画像データ上の視覚的特徴を抽出する。2つのサブシステムは互いにモデルの専門家と画像処理の専門家として働き、両者の間では既知とする知識が大きく異なっている。

モデル獲得部は、HSS 仮定ならびに対象物体の形状の変化の自由度が2次元に限定されていることを知っている。そこから対象物体がいくつのヒンジ・スラ

イド・ソリッドから成り立ち、それらがどのような関係にあるのかを学習する。ただし実際の画像データの情報を直接参照することはできない。また画像処理モジュールのことについては何も知らない。

処理戦略部は実際の画像データの情報を管理し、個々の画像処理モジュールの間の接続に関する物理的拘束の知識を持っている。試行錯誤の繰り返しによって偶然に得られる成功例を発見することでシステム内に画像処理戦略に関する知識を蓄積する。ただし HSS 仮定や 2 次元性仮定などについては何も知らない。

後で例を挙げて説明するように実際の学習は両者の対話を通じてなされる。このように 2 つの独立したサブシステムから成る構成にしておくことによって、処理戦略部で蓄積される画像処理戦略は特定の物体の構造とは独立なものになり、MIRACLE-IV の一般性が保証されるわけである。処理戦略部は実画像を取り扱い、モデル獲得部は高次の記号（線素集合）を取り扱い、両者は必要に応じて適宜情報を交換する。最初の段階でパターンを記号に変換して以後はその記号の上で学習をするというシステムと比較すれば、MIRACLE-IV は処理戦略部とモデル獲得部で相互に情報を参照するので、実画像をデータとしていることに本質的な意味のある概念学習システムと言える。

モデル獲得部の実際は対象物体を何にするかに強く依存する。ここではコンパス、はさみ、定規などの身近な器具を対象物体に限定したシステムを作成した。これらの実用的な器具は一般にその機能に無関係の飾りなどがついていないので、機能的特徴と視覚的特徴の対応が付きやすい。機能に無関係の飾りがついてる物体では、飾りという視覚的特徴に対応する機能的特徴は存在しないので、MIRACLE-IV の枠組みは成り立たない。

#### 4. モデル獲得部の機能とその実現

たとえば対象物体として、あるコンパス（円を描く道具）を考えることにしよう。コンパスという物体は人間が手で操作することによって形状がさまざまに変化する。その中のいくつかのシルエット画像をある適切な順番で MIRACLE-IV に示す。提示の順番は重要な意味を持つので後で議論する。提示の際にそれらの画像が同一物体であることは教えるが、コンパスというものがどのような物体であるかは一切教えないものとする。このような状況設定のもとで、モデル獲得部は処理戦略部との対話を通じて、そのコンパスに

はヒンジがいくつ存在してそれらがどのような関係にあるのか、という内部構造のモデルを獲得する。

モデル獲得部は以下の仮定に基づいてモデルを構成する。

- 1) 対象物体は少数のソリッドから構成されている。
- 2) ヒンジの方がスライドよりも出現頻度が高い。
- 3) 形状の変化は 2 次元平面上である。

1) と 2) の仮定はシステムの作成に先立って我々が身近な器具を大量に集めて分析した結果得られたものである。1) は構成ソリッドの数が多くなると使いにくくなるという理由によるものと思われる。2) はヒンジの方がスライドよりも作りやすいあるいは使いやすいという理由によると想像される。3) は問題を簡単にするために我々が与えた作爲的な仮定である。

提示する順番は学習に都合のよいものでなければならぬ。すなわち、1 つ前に見せた形状と少しだけ異なる形状を順々に提示する。具体的には対象物体のヒンジもしくはスライドの部分を 1 つ前から 1 箇所だけ変化させたものを見せる。ヒンジの場合は関連するソリッド間の角度を変え、スライドの場合は関連するソリッドをずらして長さを変える。一度に複数箇所を変化させるような順番で提示すると学習はうまくいかない。第 6 章で学習がうまくいく提示順序とうまくいかない提示順序の例を示す。

モデル獲得部は原則として処理戦略部から線素の集合を受け取ってそれを入力とする。ある画像（線素集合）のある部分が他の画像のどの部分に対応しているのかを決定するのは、モデル獲得部の重要な役割である。特にこの場合は比較すべき 2 つの画像の形状が一致していないので、柔軟な対応付けを行う必要がある。

MIRACLE-IV では柔軟な対応付けを行うために弛緩整合法 (relaxation method)<sup>10)</sup> を用いている。弛緩整合法では、線素間の拘束条件を伝播させることによって、長さや角度が類似した線素同士の対応を取ることができる。我々が以前に弛緩整合法によって手書き漢字の認識を行った研究<sup>11)</sup>においては、対象とした画像について、方向はほぼ一定であること、形状が極端に変化しないこと、という 2 つの条件が成立していた。しかし MIRACLE-IV ではこれらの条件は成り立たない。物体は形状が変わるので方向の一定性という概念がない。そこで、絶対的な座標値・角度ではなく相対的な座標値・角度を用いる、1 本の線素に高々

2本の線素を対応付けるのではなく $m$ 本の線素に $n$ 本の線素を対応付ける( $m$ と $n$ は任意の自然数), という拡張を行った。

さらに, 拡張した弛緩整合法の上に, HSS 仮定を埋め込んだ推論システムを作成した。この推論システムは, 弛緩整合法のパラメータをいろいろと変化させて収束の状況を観察することによって, 線素間の対応付けと並行してヒンジ・スライドの存在を推定する。弛緩整合法には大きく分けて角度と距離という2種類の要素のパラメータが存在する。基本的には, 角度差によってヒンジの存在が, 距離差によってスライドの存在が, それぞれ推定されることになる。対応付けすべき2つの線素集合が与えられると, モデル獲得部はまず標準的なパラメータ値の弛緩整合法による対応付けを試みる。一般には形状の変化により対応がうまく取れないので収束しない。推論システムはその結果に応じて弛緩整合法の角度や長さのパラメータ値を変更する。パラメータ値は要素の相対的な重要性を表しているため, たとえば角度のパラメータ値を小さくすると, 弛緩整合法は角度の変化に鈍感になり, 相対的に長さの類似性を重視した対応付けをすることになる。推論システムは試行錯誤によって角度と長さのどちらが変化したかを推定する。それがヒンジ・スライドの存在の推定につながる。その推定の過程で, 対象物体に対する上記の3つの仮定を利用している。

モデル獲得部の線素集合の対応付けの能力は理論的には不明である。しかし高々1箇所のヒンジ・スライドしか動かしていないという条件のもとでは, 経験的にはほぼ正しく機能している。一度に複数箇所のヒンジ・スライドを動かした場合や線素集合の線素数が多くなった場合などは一般に正しく機能しない。パラメータ値を変更しても収束しなかったり, 間違った対応付けをしたりする。

モデルはヒンジ・スライド・ソリッドの集合として表現される。ソリッドは線素集合の部分集合として, ヒンジとスライドは複数のソリッド間の関係として表現される。これらの表現の例は第6章に現れる。また各画像におけるヒンジ・スライド・ソリッドの位置は具体的な座標値によって参照可能である。

## 5. 処理戦略部の機能とその実現

モデル獲得部によって得られる物体の機能的特徴を, 実際の画像データに対する画像処理によって視覚的な特徴(すなわち見え方)として検出するのが処理

戦略部の役割である。

ゴールを目指して自動的に試行錯誤を行うシステムとしては LLVE<sup>12)</sup>, IMPRESS<sup>13)</sup> などが既に開発されており, 前もってシステムに与えられた高度な知識によって試行錯誤過程が制御されている。これに対し, 我々の MIRACLE-IV の処理戦略部では, 偶発的な成功例を蓄積することによって, 画像処理戦略を計算機内に形成する機構を実現した。

処理戦略部は, 接続して画像処理アルゴリズムを構成することのできる画像処理モジュール(たとえば2値化モジュール, 輪郭抽出モジュール)群<sup>14)</sup>を持っている。それらの接続に関する物理的拘束, すなわち各モジュールの入出力のデータ型を処理戦略部は知っている。たとえば, ヒンジという内部構造の検出を考えてみる。コンパスなどを見てわかるように, ヒンジは円の形状を持つことが多い。しかし初期の段階においては, ヒンジの持つ視覚的特徴について処理戦略部は何も知らない。この場合にはモデル獲得部の主導によって処理が進められる。すなわち提示されたシルエット画像群よりモデル獲得部はヒンジの存在とそのおおよその位置とを推定し, 処理戦略部に対し視覚的特徴によるヒンジ検出を要求する。処理戦略部は, 与えられたヒンジの予測点を検出できるまで画像処理の試行錯誤を繰り返す<sup>15)</sup>。具体的には, 濃淡画像を入力データとし, 点というデータ型を出力する画像処理モジュール系列を乱数に基づいて生成する。この系列は実行されて, 処理結果と予測点との誤差が評価される。誤差が大きければ新しい系列が再度生成される。誤差が許容範囲内であれば, その系列に対し乱数による変更(モジュールの削除, 挿入, パラメータ変更, 等)を加え, より良い結果が得られる処理系列を探索する。もし検出に成功すれば, ヒンジという概念に対する画像処理手順が処理戦略部内に形成されたことになる。

機能的特徴と視覚的特徴との対応は対象物によらない一般的な概念である。いろいろな器具の中に円の形をしたヒンジが存在する。処理戦略部が十分に獲得されれば, 未知の物体に対しても処理戦略部はモデル獲得部の助けなしに一目でその構造を推論することができるようになるであろう。

## 6. 学習過程の例

ここではある対象物体(いわゆる「コンパス」)の異なる形状のシルエット画像を順番に1枚ずつ計4枚

MIRACLE-IV に提示したケースを例にとって、機能的特徴と視覚的特徴の対応付けの学習過程を説明する。まず学習がうまくいった例を詳しく述べる。MIRACLE-IV に入力したシルエット画像群(原画像はまず処理戦略部に渡され、そこで処理された線素集合としてモデル獲得部に渡される)は既に図1に示した。

(1) 処理戦略部から渡された1枚目の線素集合(図3(1)参照)は情報をそのまま保持する。まだこれだけでは対象物体について何もわからないので、特に処理は行わない。モデル獲得部は、ヒンジのところ

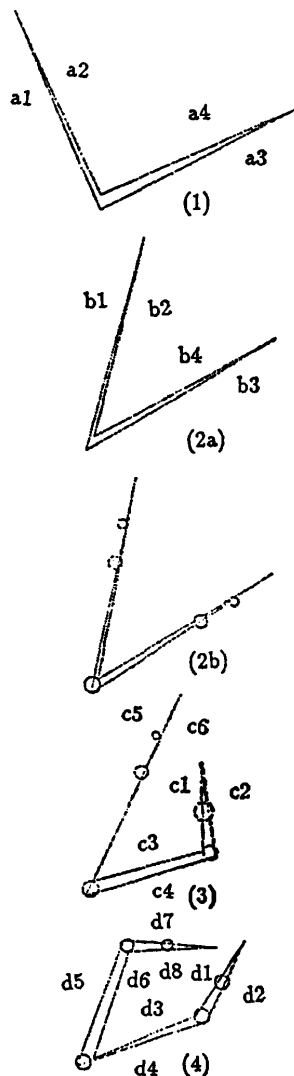


図3 学習過程の例

Fig. 3 An example of learning flow.

で曲がっているところがヒンジであるという知識は持っていない。第4章で述べたように、ある場所の回りの角度が変わればその場所にヒンジが存在すると判断する。ヒンジは曲がるという知識はモデル獲得部において手続き的に表現されている。その表現については(2)の段階で構築されるモデルを例にしてそこで説明する。この段階では「コンパス」なるものは全体として1つのソリッドで成り立つ、というモデルを構築する。

$$\text{SOLID-A1}=(a1 \ a2 \ a3 \ a4)$$

(2) 2枚目の線素集合が与えられる(図3(2-a)参照)と、1枚目との間で対応付けを行うことになる。その結果、曲がっている場所で角度が変わっているのでその場所にヒンジが1つ存在すると判断し、まずその判断に基づいた内部構造を形成する。コンパスは2つのソリッドが1つのヒンジで結ばれているものである、というモデルである。

次にモデル獲得部は、処理戦略部に対して「ヒンジ」なる概念の存在とその大体の位置を教示し、それと形状(視覚的特徴)が似た箇所がほかにもあるかどうかを尋ねる。この際に処理戦略部はヒンジの機能的な意味(「曲がる」ということ)については何も知らないという事実注意到意されたい。処理戦略部はそのヒンジの位置を視覚的特徴として出力するような手順が見つかるまで画像処理手順の生成を繰り返す。この場合には円形状を検出する画像処理手順によってヒンジの視覚的特徴を検出できることを処理戦略部が最終的に発見する。処理戦略部からの回答を図3(2-b)に示す。円で囲んだ箇所は処理戦略部がその形状が教えられたヒンジに似ていると判断した部分である。

モデル獲得部としては、それらの箇所が曲がっているような画像を実際に見たわけではないので、ヒンジであるという確信を抱くまでには至らない。しかし、視覚的特徴が似ていればその機能的特徴も似ているであろうという思い込みによって、これらの4箇所もヒンジである可能性が非常に高いと判断し、有力なヒンジ候補として記憶する。この際に今度はモデル獲得部はヒンジの視覚的な意味(この例では「円」ということ)については何も知らないという事実注意到意されたい。

$$\text{SOLID-B1}=(b1 \ b2)$$

$$\text{SOLID-B2}=(b3 \ b4)$$

$$\text{HINGE-B1}=(\text{SOLID-B1}, \text{SOLID-B2})$$

$$[b1 \leftrightarrow a1, \ b2 \leftrightarrow a2, \ b3 \leftrightarrow a3, \ b4 \leftrightarrow a4]$$

b1 と b3 の接続点と b2 と b4 の接続点の midpoint がヒンジ HINGE-B1 の存在位置である。ヒンジは曲がるという知識は、この位置を中心として SOLID-B1 と SOLID-B2 の間の角度が変化するという形で表現されている。

(3) 3枚目の線素集合 (図3(3)参照) と、2枚目までを見た時点で得られている構造記述との間で対応付けを行う。ヒンジの候補となっている箇所についてはあらかじめ曲がる可能性 (そこで角度が変化することが可能) を前提とした対応付けをすることができるので、形状が変化しているにもかかわらず線素間の対応付けは比較的スムーズになされる。ヒンジ候補の情報が効果的な弛緩整合法を可能にしているのである。結果として、新たに1つのヒンジの存在が確認されたので、それにしたがって構造記述も更新される。コンパスは3つのソリッドが2つのヒンジで結ばれているものである、というモデルである。

SOLID-C1=(c1 c2)  
 SOLID-C2=(c3 c4)  
 SOLID-C3=(c5 c6)  
 HINGE-C1=(SOLID-C1 SOLID-C2)  
 HINGE-C2=(SOLID-C2 SOLID-C3)  
 [(c1, c3)↔b4, (c2, c4)↔b3, c5↔b1, c6↔b2,  
 SOLID-C3↔SOLID-B1, HINGE  
 -C2↔HINGE-B1]

(4) 4枚目の線素集合 (図3(4)参照) に対して (3)と同様の処理が行われる。また新たに1つのヒンジの存在が確認されたので、さらに構造記述が更新される。コンパスは4つのソリッドが3つのヒンジで結ばれているものである、というモデルである。

SOLID-D1=(d1 d2)  
 SOLID-D2=(d3 d4)  
 SOLID-D3=(d5 d6)  
 SOLID-D4=(d7 d8)  
 HINGE-D1=(SOLID-D1 SOLID-D2)  
 HINGE-D2=(SOLID-D2 SOLID-D3)  
 HINGE-D3=(SOLID-D3 SOLID-D4)  
 [d1↔c1, d2↔c2, d3↔c3, d4↔c4, (d5, d7)↔c5,  
 (d6, d8)↔c6, SOLID-C1↔SOLID-D1,  
 SOLID-C2↔SOLID-D2, HINGE-D1↔  
 HINGE-C1, HINGE-D2↔HINGE-C2]

もしもこれらの4枚の画像だけで教示を終了することになると、(4)で得られた構造記述が対象物体「コンパス」なるもののモデルである。言い換えれば、

MIRACLE-IV (のモデル獲得部) は「コンパス」の概念を学習したことになる。なお最後まで未確認のヒンジ候補はあくまで候補としてモデルの中に残っている。実際にはヒンジでない箇所をヒンジと思い込んでしまったのである。一方の処理戦略部は「円」という視覚的特徴を抽出するための知識を学習したことになる。この知識は意味のある画像処理系列としてシステム内に蓄積される。

図4は他の物体に対する処理戦略部の回答である。コンパスの例によって処理戦略部は円弧状の部分を検出するための画像処理系列を獲得している。この場合は、視覚的特徴から直接ヒンジ候補の位置が推定できる。

上でも強調したように、モデル獲得部は具体的な視覚的特徴の意味は知らないし、処理戦略部は具体的な機能的特徴の意味は知らない。この例では、モデル獲得部は「自分の知っているヒンジという機能的特徴は、このコンパスとかいう物体においては円とかいう知らない視覚的特徴と対応しているらしい」と思っているに過ぎない。一方の処理戦略部は「自分の知っている円という視覚的特徴は、このコンパスという物体ではヒンジとかいう知らない機能的特徴と対応しているらしい」と思っているに過ぎない。この状態をシステムの外から眺めると、MIRACLE-IV はコンパスにおけるヒンジと円の間結び付きを学習したということになるのである。これが機能的特徴と視覚的特徴の間の対応付けの学習例である。

ここまで説明したのは学習がうまくいくように提示

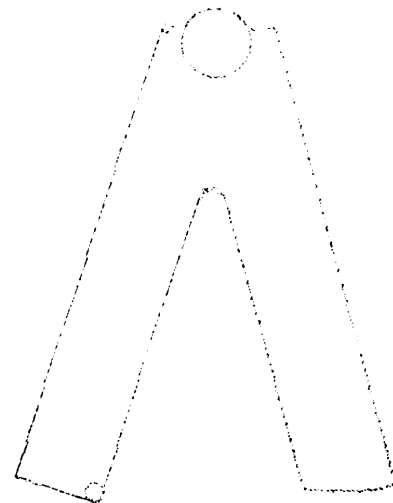


図4 定規への適用  
 Fig. 4 Application to a folding ruler.

の順番を考慮した例である。図3の例で(4), (3), (2), (1)と逆順に提示しても同様にうまくいくことは実験で確認したが、たとえば(1), (4), (2), (3)という順番で提示すると学習はうまくいかなかった。(1)と(4)とでは3箇所ヒンジの角度が一度に変化しているために、モデル獲得部における線素集合間の対応付けが正しくできないのである。(1)と(4)を見せた段階では整合的なモデルは構築できない。次に(2)を見せても1つ前の(4)と2箇所ヒンジの角度が異なるので対応付けが正しくできない。最後に(3)を見せると1つ前の(2)との間でようやく対応付けが成功し、ヒンジがひとつ存在することが判明する。その後たとえば再び(4)を見せると新たにヒンジの存在が判明することになる。このように MIRACLE-IV では提示の順番が大きな意味を持っている。

コンパスの例はヒンジを対象とした学習であったが、図5の一連のシルエット画像群を MIRACLE-IV に提示するというケースを例にとり、スライドを対象とした学習について考察することしよう。この対象物体は「比例コンパス」である。なおこの比例コンパスなるものにはヒンジも含まれているが、ここでの考察ではスライドのみに学習対象を限定する。また、話をわかりやすくするために単純化して説明する。

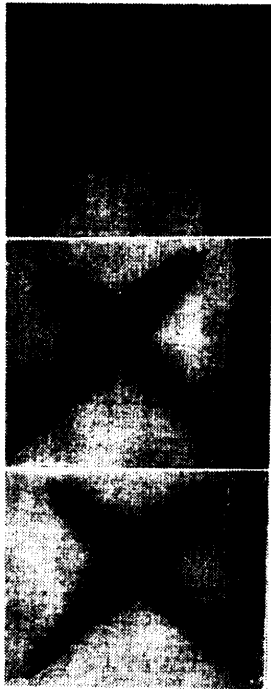


図5 比例コンパスのシルエット画像群  
Fig. 5 Silhouette images of proportional compasses.

(1) 処理戦略部から渡された1枚目の線素集合(図6(1)参照)は情報をそのまま保持する。コンパスの1枚目の段階と同様に、比例コンパスも1つのソリッドから成るものである、というモデルを構築する。

(2) 2枚目の線素集合が与えられる(図6(2)参照)と、1枚目との間で対応付けが行われる。その結果、スライドが1つ存在すると判断し、まずその判断に基づいた内部構造を形成する。比例コンパスは2つのソリッドが1つのスライドで結ばれているものである、というモデルである。この段階でモデル獲得部が学習したのは一方のソリッドがもう一方のソリッドに対してスライドするという知識のみであることに注意されたい(相互にスライドするという知識はまだ得られていない)。

次にモデル獲得部は、処理戦略部に対して「スライド」なる概念の存在とその大体の位置を教示し、それと形状が似た箇所がほかにあるかを尋ねる。処理戦略部はたとえば(スライドした)ソリッドの中の細長い穴の存在がスライドの視覚的特徴になっていることを発見する。さらに処理戦略部は、もう一方のソリッド

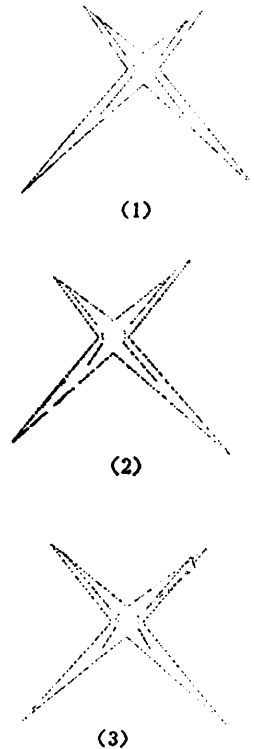


図6 比例コンパスの線素集合表現  
Fig. 6 Line segments of proportional compasses.

の中にも同じように細長い穴が存在することを見つけて、その旨をモデル獲得部に回答する。

モデル獲得部は、前述したような思い込みによって、もう一方のソリッドもスライドする可能性が非常に高いと判断し、有力なスライド候補として記憶する。

(3) 3枚目の線素集合(図6(3)参照)と、2枚目までを見た時点での構造記述との間で対応付けを行う。ここでもスライド候補の存在が効果的な弛緩整合法を可能にする。結果として、もう一方のソリッドもスライドすることが確認されたので、それにしたがって構造記述も更新される。比例コンパスは2つのソリッドが1つのスライドで結ばれている(2つのソリッドは相互にスライドする)ものである、というモデルである。

この学習によって、モデル獲得部は比例コンパスのモデルを獲得し、処理戦略部は「細長い穴」という視覚的特徴を抽出する戦略の知識を学習したことになる。MIRACLE-IVは全体として、比例コンパスにおけるスライドという機能的特徴と細長い穴という視覚的特徴との間の結び付きを学習したことになる。

### 7. 概念学習システムとしての MIRACLE-IV

コンピュータによる学習はさまざまな観点から分類することができる。以下にMIRACLE-IVの概念学習システムとしての特徴を述べる。

(1) 具体的な事例から一般的な規則を導き出す帰納的学習である。

ひとつひとつの画像(形状バリエーション)が具体的な事例に相当する。ヒンジ・スライド・ソリッドの組み合わせによる構造記述が一般的な規則に相当する。したがって、複数の具体的な事例からそれらに共通する単一の概念を一般的な規則という形で導き出すという典型的な帰納的学習である。

(2) 正の事例のみを対象とした学習である。

概念学習では、学習すべき概念の例を正の事例と呼び、反例を負の事例と呼ぶ。同一物体の形状バリエーションのシルエット画像群を提示するのであるから、正の事例のみを対象としていることになる。特にモデル獲得部における画像間の対応付けの過程では、正の事例のみであるということをも前提とした処理を行っている。また既に述べたように学習に都合のよい順番で実例が与えられること<sup>16)</sup>を仮定している。すなわち、

1つ前の実例と大幅に異なるものではなく、少しだけ異なるものが順々に与えられる。

(3) 類推による学習に基づいている。

Winstonの定義によれば、類推とは「類似性は因果関係を保持する。すなわち、似た状況は似た結果を生じやすい。」という原則に基づき、既知の状況から未知の状況へ関係を写像することである。MIRACLE-IVは複数の例を用いた類推による学習をしている。MIRACLE-IVにおいては視覚的特徴が「状況」に相当し、機能的特徴がその因果的な「結果」に相当する、と仮定している。対象物体とした身近な器具では形状が用途に直接的に対応していることが多いので、この仮定は妥当であると考えている。機能に関係ない飾りなどが付属している物体ではこの仮定は成立しない。

(4) 画像・記号間で情報が相互参照されている学習である。

ひとたび画像を記号に変換したらその先はすべてその記号の上で学習を進めるという枠組みでは、画像というパターンをデータとしていることがその学習システムにとって本質的な意味を持たないことになる<sup>1)</sup>。MIRACLE-IVではモデル獲得部と処理戦略部の間で相互に情報が参照されている。MIRACLE-IVにとっては画像というパターンをデータとしていることが本質的である。

### 8. システム構成

図7にMIRACLE-IVのシステム構成を示す。記号処理の部分はすべてSymbolics 3640というリスパマシン上でなされている(プログラムはUranus<sup>17)</sup>とLispで記述されている)。信号レベルの画像処理は画像処理専用プロセッサIP8500を伴ったμVAX-II上で実行する。2つの計算機はEthernet(TCP/IP)でつながっている。画像処理の命令はSymbolicsからμVAX-IIに送られる。画像処理の結果(たとえば成功・失敗の情報、線素データなど)は逆方向に送られ

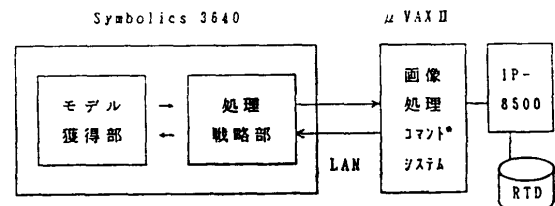


図7 システム構成図  
Fig. 7 System configuration.



る。試行錯誤の過程で生じる膨大な中間結果を蓄える画像データベース装置として RTD (リアルタイム磁気ディスク) を利用する。

## 9. おわりに

MIRACLE-IV における機能的特徴と視覚的特徴の間の対応付けの学習について例を挙げて説明した。形状が変化する物体の内部構造を推定するというのは新しい試みと思われる。また画像を対象としていることが学習において本質的な意味を持っていることも大きな特徴である。モデル獲得部では画像モデル、処理戦略部では画像処理戦略をそれぞれ学習している<sup>15), 18), 19)</sup>。

MIRACLE-IV の「思い込み」が誤りである可能性もありうる。図 8 がその例である。最後の画像を見るまでは 2 つのソリッドから成るハサミであると思ひ込む。人間が同様の状況で同じように誤った「思い込み」をするのであれば、それは MIRACLE-IV が人間の学習のモデルとして正当であることを意味しているものと考えられる。学習という過程は探索の問題に密接に関わっている。物理的・数学的に厳密に考えれば無限もしくは非常に多くの可能性の枝を探索しなくては

ならないはずなのに、人間はほとんどの場合そのような探索空間の爆発には煩わされていないように見える。そこではここで述べたような「思い込み」による枝刈りが大きな役割を果たしていると思われる。

MIRACLE-IV では機能的特徴と視覚的特徴との間の対応付けを行っているが、ここで言っている機能的特徴とはヒンジ・スライドという形状を変化させる部分のことである。本論文ではコンパスを例に挙げたが、コンパスの本来の機能である「円を描く」ことは MIRACLE-IV における学習では取り扱っていない。このような器具としての機能の特徴は必ずしも視覚的特徴と対応しているとは限らないので取り扱いが困難であることが想像される (たとえば Winston らの研究<sup>2)</sup> では自然言語によって対応を明示的に教示することでその困難さを回避している) が、MIRACLE-IV の次のステップとして興味深い研究テーマである。

**謝辞** 研究の機会を与えていただいた電総研の中島隆之知能情報部長、日頃ご討論いただく画像研究室の皆様へ感謝します。また Symbolics Lisp Machine についていろいろと教示いただいた、言語システム研究室の戸村哲氏ならびに認知科学研究室の錦見美貴子氏に感謝します。最後に適切なコメントをいただいた査読者の方に感謝します。

## 参 考 文 献

- 1) 松原 仁, 坂上勝彦, 横矢直和, 山本和彦: 概念学習を題材とした画像理解と記号処理の統合の試み, 人工知能学会誌, Vol. 3, No. 5, pp. 572-580 (1988).
- 2) Winston P.H., Binford, T.O., Katz, B. and Lowry, M.: Learning Physical Descriptions from Functional Definitions, Examples, and Precedents, AAAI-83, pp. 433-489 (1983).
- 3) Connell, J. H. and Brady, M.: Generating and Generalizing Models of Visual Objects, *Artif. Intell.*, Vol. 31, No. 2, pp. 159-183 (1987).
- 4) 伊藤幸宏, 高木 朗, 六沢一昭, 清水正朗, 北岡和憲, 小原啓義: 2次元図形世界における日本語文と視覚情報の付き合い処理, 信学論(D), Vol. J69-D, No. 7, pp. 1083-1094 (1986).
- 5) 安部憲広, 西田好広, 辻 三郎: 類似例を用いたマンマシ系による教示, 情報処理学会論文誌, Vol. 25, No. 5, pp. 755-763 (1984).
- 6) 神田享子, 小川 均, 北橋忠宏: 画像理解のための機能に基づく物体モデルの記述, 電子情報通信学会人工知能研究会, AI 88-29 (1988).
- 7) 半田剣一, 松原 仁, 石崎 俊: 学習におけるアナロジー, 人工知能学会誌, Vol. 2, No. 1, pp. 44-51 (1987).

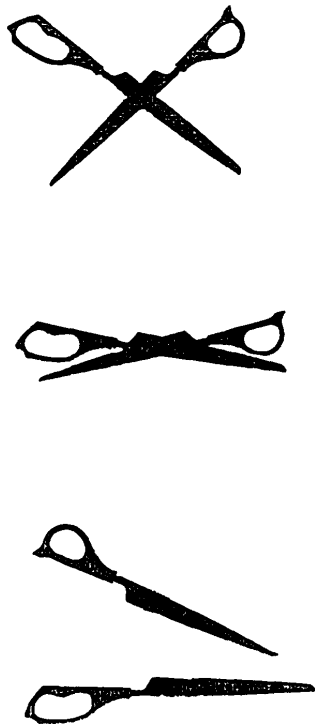


図 8 ハサミのシルエット画像群  
Fig. 8 Silhouette images of scissors.

- 8) 富田文明: 3次元物体の境界線に基づく記述の生成と認識, 信学論 (D), Vol. J71-D, No. 2, pp. 343-349 (1988).
- 9) 田村秀行, 坂根茂幸, 富田文明, 横矢直和, 金子正秀, 坂上勝彦: ポータブル画像処理ソフトウェア・パッケージ SPIDER の開発, 情報処理学会論文誌, Vol. 23, No. 3, pp. 321-328 (1982).
- 10) Rosenfeld, A., Hummel, R. A. and Zucker, S. W.: Scene Labeling by Relaxation Operations, *IEEE Trans. SMC*, Vol. SMC-6, No. 6, pp. 420-433 (1976).
- 11) 山本和彦: 弛緩整合法による手書き教育漢字認識, 信学論 (D), Vol. J65-D, No. 9, pp. 1167-1174 (1982).
- 12) 松山隆司, 尾崎正治: LLVE: トップダウン・セグメンテーションのための画像処理エキスパートシステム, 情報処理学会論文誌, Vol. 27, No. 2, pp. 191-204 (1986).
- 13) 長谷川純一, 久保田浩明, 高須晶英, 鳥脇純一郎: 画像処理エキスパートシステム IMPRESS における画像処理手順の集約化機能について, 情報処理学会論文誌, Vol. 29, No. 2, pp. 126-133 (1988).
- 14) 山岸健太郎, 坂上勝彦, 山本和彦, 松原 仁: 輪郭図形からの注目候補点の検出方法, 信学論 (D-II), Vol. J72-D-II, No. 5, pp. 750-759 (1989).
- 15) 坂上勝彦, 山本和彦, 松原 仁, 山岸健太郎: 認識機構学習システム MIRACLE-IV における画像処理戦略の学習, 電子情報通信学会パターン認識・理解研究会, PRU 88-20 (1988).
- 16) Winston, P. H.: Learning Structural Description from Examples, in Winston, P. H. (ed.), *The Psychology of Computer Vision*, McGraw-Hill (1975).
- 17) Nakashima, H.: Uranus Reference Manual, 電総研彙報, Vol. 50, No. 8, pp. 829-910 (1986).
- 18) 松原 仁, 山本和彦, 坂上勝彦, 山岸健太郎: 認識機構学習システム MIRACLE-IV における内部構造モデルの獲得, 電子情報通信学会パターン認識・理解研究会, PRU 88-19 (1988).
- 19) Yamamoto, K., Sakaue, K., Matsubara, H. and Yamagishi, K.: MIRACLE-IV: Multiple Image Recognition system Aiming Concept LEarning - Intelligent Vision, *Proc. of 9th ICPR*, pp. 818-821 (1988).

(平成元年7月31日受付)  
(平成2年6月4日採録)

#### 松原 仁 (正会員)



1959年東京生. 1981年東京大学理学部情報科学科卒業. 1986年同大学院工学系研究科情報工学専門博士課程修了. 工学博士. 1986年電子技術総合研究所入所. 現在知能情報部推論研究室ならびに協調アーキテクチャ計画室に所属. 人工知能, 特に学習と知識表現に興味を持つ. AIUEO, 日本ロボット学会, 日本認知科学会, 人工知能学会各会員.

#### 坂上 勝彦 (正会員)



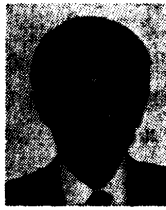
昭和29年生. 昭和51年東京大学工学部電子工学科卒業. 昭和56年同大学院博士課程修了. 同年電子技術総合研究所入所, 現在に至る. パターン認識, 画像処理などの研究に従事. 工学博士. 昭和61年本学会論文賞受賞. IEEE, 電子情報通信学会, 電気学会各会員.

#### 山本 和彦 (正会員)



昭和44年東京電機大学工学部電気通信卒業. 昭和46年同大学院修士課程修了. 同年電子技術総合研究所入所, 手書き文字認識の研究に従事. 昭和54~55年米国メリーランド大学のコンピュータビジョン研究所で並列分散形処理による画像理解の研究. 現在電子技術総合研究所知能情報部画像研究室室長. 人工知能, パターン認識, 特に図形の柔軟な対応付けの研究に従事. 工学博士. 電子情報通信学会, 電気学会, AVIRG 各会員.

#### 山岸健太郎 (正会員)



昭和34年生. 昭和56年東京大学工学部計数工学科卒業. 昭和58年同大学院修士課程修了. 同年新日本製鐵(株)入社. 東京大学工学部助手を経て, 平成元年より住友電工(株)情報電子研究所勤務. 電子情報通信学会, IEEE 各会員.