

解説

エキスパートシステムによる画像解析・理解

松山 隆 司

岡山大学工学部情報工学科 〒700 岡山市津島中 3-1-1

(1990年6月30日受理)

Image Analysis and Understanding by Expert Systems

Takashi MATSUYAMA

Department of Information Technology, Faculty of Engineering,
Okayama University, 3-1-1, Tsushima-Naka, Okayama 700

1. はじめに

エキスパートシステム (以後 ES と略記する) とは、専門家のもつ知識を利用して高度な問題解決を行うソフトウェアシステムで、人工知能研究の実用的有用性を示すものとして 1980 年代に活発にシステム開発がなされた。ES には、現象の分析、診断を目的としたものと、対象の構造設計を目的としたものがあり、前者の例としては血液感染症の診断を目的とした MYCIN があり、後者の例としては顧客の希望に合った計算機の構成を自動的に求める R1 (あるいは XCON) が有名である。

ES (および多くの人工知能システム) では、

知能 = 知識 + 推論機構 (1)

という図式を基本原理としている。すなわち、従来の問題解決システムでは、プログラムという一つの手続きの中に、問題解決に必要な知識とそれを利用した推論・計算が混在しており、知識そのものを独立に取り扱うことができなかった。これに対し、ES では、

IF (起動条件) THEN (動作記述) (2)

といった形式のプロダクション規則の集合として、知識が手続き (推論機構) とは独立なデータとして記述される。これによって知識の追加・修正が容易になる。また、規則の自動生成機能を付加すれば学習も可能となる。さらに、同じ規則集合を用いた場合でも、推論機構を換えることによって、システムの能力や効率を向上させることができる。

(1) の図式に基づいて知能システムを開発する際の最も大きな問題は、知識をどのように表現、記述するかと

いうことである。(2) のような標準的なプロダクション規則を用いた場合でも、起動条件や動作記述を何によって (記述のための基本語彙) どう記述するのかは、対象領域の構造やシステムの目的に応じて考えねばならず、そうした基本語彙の設定や記述形式の設計こそが対象領域に関する知識表現の本質であるといえる。

画像解析、画像理解 (画像に写された対象物を認識し、それらの相互関係の記述を求める研究) の分野においても、(1) の図式に基づいて、システムの解析、認識能力の向上を目指した研究が 1980 年頃から活発に行われるようになってきた。こうした研究分野は知識型ビジョン (knowledge-based vision) あるいはエキスパート・ビジョンと呼ばれている。

本解説では、画像の解析・理解のための知識表現法という観点から、これまでに開発された知識型ビジョンシステムを紹介する。なお、紙面の制約のため紹介が表面的にならざるをえず、また本誌の読者にとってはなじみの薄い記号論理による表現などが現れるが、より深い技術内容に興味がある方は参考文献を参照していただきたい。

2. 画像解析・理解のための知識の分類

画像を解析し、そこに写されている対象を認識、理解するには多種多様な知識が必要となる。それらの知識を大別すると、次のようになる。

(1) 画像世界の知識: 生の画像データから抽出される線や領域、交点や角といった 2 次元の画像特徴のもつ属性 (色、形など) およびそれらの間に成り立

つ幾何学的関係 (平行, 交差, 共直線性など)

(2) 対象世界の知識: 認識対象 (家, 自動車, 道路など) のもつ属性や (3 次元的) 構造および対象物間の空間的, 意味的關係 (たとえば自動車にはセダン, トラック, バスなどの種類があること)

(3) 対象世界と画像世界間の写像關係の知識: 画像を撮影したときのカメラ・パラメータやカメラの位置, 方向, 照明の種類や配置といった光学的知識
 画像解析では, (1) の知識に基づいて, 生の画像データから画像特徴を抽出したり, 単純な画像特徴をグループ化することによって多角形などの複雑な画像特徴を構成する. 対象の認識では, 抽出された画像特徴がどの認識対象に対応するのかを求めればよく, 意味のある対応付けを行うために (3) の知識が利用される.

このほか次節で述べる画像処理エキスパートシステムでは, 第 4 の知識として画像処理技術 (アルゴリズム) に関する知識が利用される. このタイプの知識は, 効果的な画像解析を実現するにはどのような画像処理アルゴリズム (サブルーチンや関数) をいかに組み合わせれば

よいかに関するソフトウェア的なノウハウといえる.

3. 画像処理エキスパートシステム

約 30 年のデジタル画像処理研究の結果, 膨大な種類の画像処理アルゴリズムが開発され, 現在ではその一部がプログラム・ライブラリとしてまとめられ, 誰にでも利用できるようになっている.

しかし, 実際に画像解析を行う場合, どのような画像処理アルゴリズムを選択し, 組み合わせればよいのかといった画像処理技術に関する知識は, 画像処理の専門家のノウハウとして経験的に知られているにすぎない. このため, 画像処理を一つの要素技術として利用しようとする応用分野の技術者は, そうした画像処理技術に関する知識を身につけない限り画像処理プログラム・ライブラリを使いこなせないのが現状である.

こうした問題を解決するために, 画像処理技術に関する知識を計算機上に明示的な形で表現し, それを利用したエキスパートシステムを開発しようとする研究が 1980 年代半ばからわが国で始められた^{1,2)}. 画像処理エ

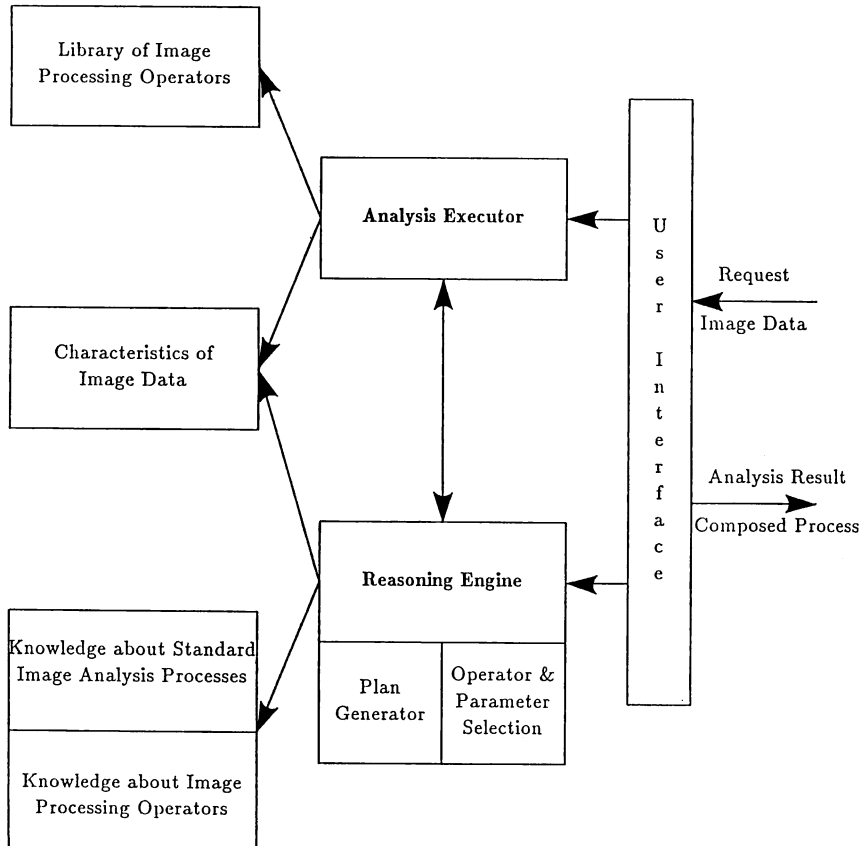


図 1 画像処理エキスパートシステムの構成

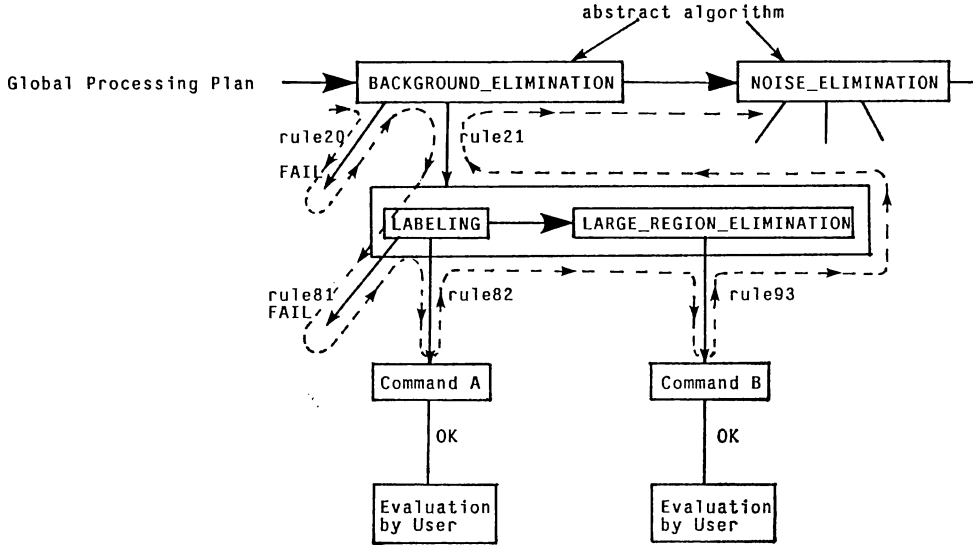


図 2 コンサルテーションのための推論, 探索

キスパートシステムは、いわゆる知識工学の手法によってライブラリ中の個々の画像処理プログラムの仕様や使い方に関する知識を表現し、柔軟で効果的な画像処理機能を実現しようとするものであり、その構成は図1に示すようになっている。

これまでに提案された画像処理エキスパートシステムを目的別に分類すると次の四つのタイプに分けられる²⁾。

3.1 画像処理コンサルテーションシステム

3.1.1 目的

会話型画像処理システムにおけるマン・マシン・インタフェースを柔軟にするために、システムが利用者に画像処理コマンドに関する種々の情報を提案し、画像処理の初心者でも十分にシステムが使いこなせるようにコンサルテーションを行う。

3.1.2 知識

画像を処理する目的（雑音除去や画像特徴の抽出）を画像処理コマンドの系列に展開するためのノウハウおよび、コマンドの適用条件、パラメータ選択法などをプロダクション規則として記述する。

3.1.3 具体例

図2にコンサルテーションのための推論過程を示す³⁾。システムの利用者が処理目的（コンパクトな領域の抽出）を入力すると、まず大まかな処理過程を表す抽象アルゴリズムの系列（図の最上段）が規則によって生成される。次にシステムは系列中の各抽象アルゴリズムを具体化するための規則を再帰的に呼び出し、実行可能

なコマンドの系列を探す。図では、まず BACKGROUND-ELIMINATION という抽象アルゴリズムがルール 21 によって LABELING→LARGE-REGION-ELIMINATION というアルゴリズム系列に展開され、各アルゴリズムはそれぞれルール 82, 93 によって画像処理コマンド A, B として実現できることが推論されている。実行可能なコマンドが見つかったら、順次それを対象画像に適用し、処理結果の良否を利用者に尋ねる。処理結果が満足できる場合は、系列中の次の抽象アルゴリズムの具体化に進むが、そうでない場合は、他の規則を起動してパラメータ値やコマンドを変更し、再び処理を実行する。図ではルール 20, 81 による推論によって生成されたコマンドがうまく動作しなかったことが FAIL と表されている。こうした推論過程はすべてプロダクション規則によってガイドされ、利用者はシステムが示す処理結果の妥当性を判断するだけでよい。

3.2 画像処理プログラムの自動合成

3.2.1 目的

作成したい画像処理プログラムの仕様（機能）を与えると、プログラム・ライブラリ中の基本アルゴリズムを組み合わせることによって、実行可能なプログラムを自動的に合成する。仕様の与え方としては、(a) 画像処理アルゴリズムの名前を表す抽象的なコマンドの系列⁴⁾ (b) 画像から抽出したい画像特徴の属性や構造記述⁵⁾ (c) 画像から抽出したい画像特徴を表す図形⁶⁾、などが考えられている。たとえば、(b) の方式では、

$$d1(\vec{R}, S), d2(\vec{R}, S) = \text{disk}(\vec{R}, S)$$

(d1, d2 はともに円 (disk) 型の構造型変数で, 中心の位置ベクトル \vec{R} と半径 S を属性としてもつという宣言文)

$\text{twodisk}(\vec{R}, \vec{D}, S) \rightarrow [d1(\vec{R}, S), d2(\vec{R} + \vec{D}, S)]$

(twodisk は \vec{R}, \vec{D}, S という属性をもつ構造型変数で, 中心が \vec{D} 離れた二つの円から構成されているという構造型記述)

という仕様を与えると, 中心の位置が \vec{D} 離れた, 半径 S の二つの円から構成される画像特徴を抽出するプログラムが自動的に合成される。

3.2.2 知識

抽象的な仕様を具体的なアルゴリズムの組合せに展開する知識および, ライブラリ中の各プログラムの引数の数やデータ型といったソフトウェア的な制約条件を記述する。

3.2.3 具体例

図3に(c)の方式による推論過程を示す。システムの利用者は処理対象画像(図の右最上段)とそこから抽出したい画像特徴を表すサンプル図形(図の左最下段)を与える。システムは図2と同様, まず大まかな処理過程

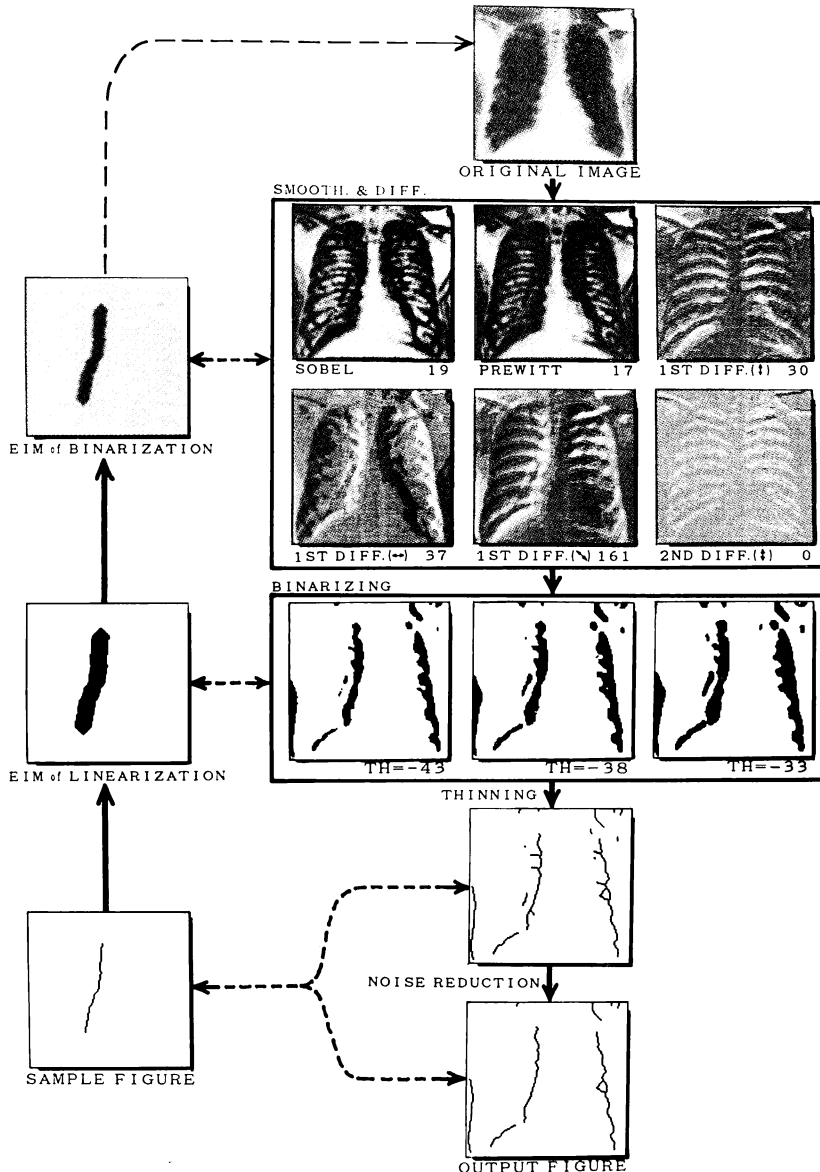


図3 サンプル図形に基づくプログラム合成過程

を表す抽象アルゴリズムの系列を求め、サンプル図形にそのアルゴリズム系列を逆向きに適用することによって、処理の中間結果の予測画像(図の最左列の画像)を生成する。次に処理対象画像に対し、系列中の各抽象アルゴリズムに対応する複数の具体的プログラムを並列的に適用し、それらの処理結果(図の右の一群の画像)を先に求めた中間結果の予測画像と比較する。そして、予測画像と最も一致度の高い結果を出力する具体的プログラムの系列を最終的な合成プログラムとする。

3.3 画像特徴抽出アルゴリズムの設計⁷⁾

3.3.1 目的

画像を対象物を表す意味のある領域に分割するには、領域や境界線の分割、併合を行う必要がある。そうした分割、併合のための知識をプロダクション規則によって表現し、柔軟な制御が可能な画像分割アルゴリズムを実現する。

3.3.2 知識(具体例)

従来プログラム中に埋め込まれていた分割、併合のための知識を

IF REGION-SIZE is SMALL and

REGION-COLOR is SIMILAR,

THEN MERGE two adjacent REGIONS

のような画像解析規則の集合として表し、どの規則を画

像のどの部分に適用すればよいかに関する知識を

IF PREVIOUS-PROCESS was REGION and

PREVIOUS-PROCESS was not ACTIVE,

THEN match LINE-ANALYSIS-PROCESS

のような制御用規則によって記述する。この制御規則は、直前の画像解析(PREVIOUS-PROCESS)が領域指向のものであり、適用可能な解析規則がない(not ACTIVE)ならば、線分指向の解析を起動すればよい(match)という知識を表している。

3.4 画像特徴の自動抽出⁸⁾

3.4.1 目的

画像から抽出したい画像特徴をゴールとして与えることにより、システムが自動的に処理方針を考え、それに従って試行錯誤的に画像を解析し、目的とする画像特徴を抽出する。

3.4.2 知識

濃淡画像から連続した線分を抽出する場合、濃淡画像→(エッジ抽出)→エッジ画像→(しきい値処理)→エッジ点→(連結操作)→線分 といった処理過程が考えられる。そこで、濃淡画像やエッジ点などの画像特徴をノードとし、エッジ抽出やしきい値処理などの処理アルゴリズムをアークと考えると、画像処理に関する知識が図4のようなネットワークとして表現できる。また、処理ア

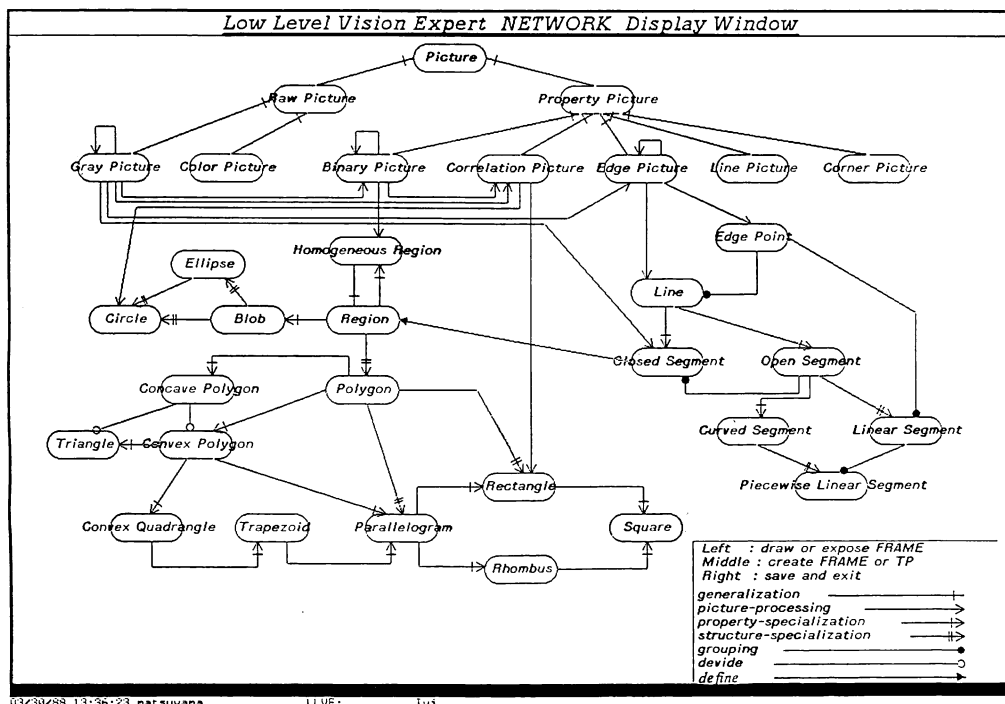


図4 ネットワークによる知識の表現

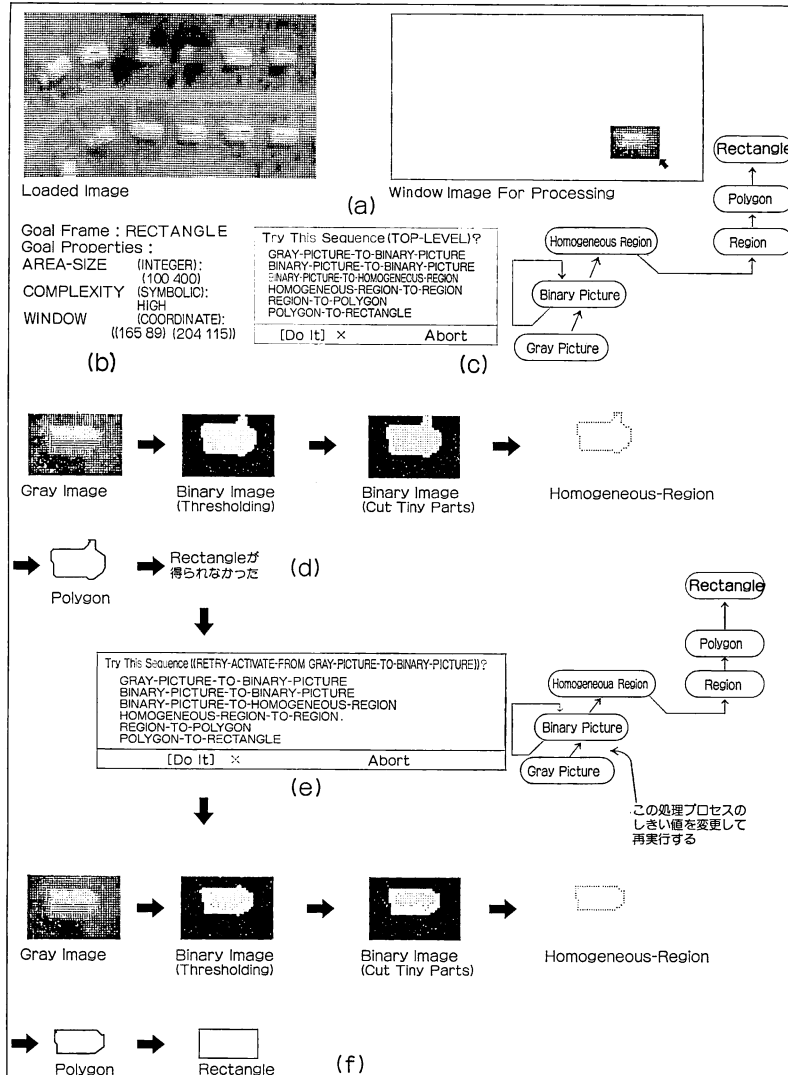


図 5 長方形の自動抽出例

ルゴリズムをいかに組み合わせればよいかといった画像処理技術に関するノウハウはプロダクション規則として別途記述する。

3.4.3 具体例

システムに与えられたゴールは、「図 5 (a) の画像において、(165 89), (204 115) の 2 点で指定される座標軸に平行な矩形領域の中で、面積が 100 から 400 の間の長方形を求めよ」というものである (図 5 (b))。システムはまず図 4 のネットワークを探索して、濃淡画像 (2 値化) → 2 値画像 (連結領域のラベル付け) → 領域 (境界線の折れ線近似) → 多角形 (長方形近似) → 長方形、という処理方針を作成し (図 5 (c))、その方針

に従って実際に (部分) 画像を処理する (図 5 (d))。

しかし、抽出された多角形は長方形とはみなされず、処理が失敗する。そこでシステムは、失敗に対する対策を記述した規則を起動し、その規則に従って別の処理を再実行する (図 5 (e))。今の場合、2 値化のしきい値が変更され、指定された条件を満たす長方形がうまく検出された (図 5 (f))。

4. 画像理解のための知識表現

画像処理エキスパートシステムでは、画像世界の知識がおもな知識源で、対象世界や写像関係に関する知識は利用されていない。画像から対象世界の構造を認識、理

解するには、これらすべての知識が必要となり、そのため知識表現を考案する必要がある。最近では画像の認識、理解のための知識をプロダクション規則のような直観的なものではなく、厳密な数学的、論理的基礎をもった記述法によって表現しようとする研究が試みられている。

4.1 幾何モデルのコンパイル

部品の組立を行うロボットなどの視覚認識機構を考えると、認識対象の幾何学的形状は CAD (computer aided design) による幾何モデルとしてすでに計算機内に蓄えられていると考えてよい。そこで、3次元の幾何モデルと観測した画像中の2次元の画像特徴の対応付けを行えば、対象の種類、位置、姿勢が認識できる。

池内⁹⁾は、認識対象の幾何モデルと画像を撮影したセンサの特性を知識源として入力することにより、その対象を認識するための具体的解析プログラムを自動生成するモデル・コンパイラを開発した。これは、3.2で紹介したシステムを3次元物体認識のために拡張したものと考えられる。

3次元物体の画像中での見え方は、視線方向、照明、センサーの特性によって大きく変化する。そこで、まず物体の幾何モデルを解析することにより、安定な姿勢を計算し、可能な視線方向を限定する(図6(a))。次に、各視線方向から見える物体の面の大きさや形、数によって視線方向の特徴付けを行い、図6(b)のような解析過程をガイドする木構造を作成する。

物体認識過程では、図6(b)の木構造を上からたどり、分岐点に記されている画像特徴の有無や数を調べ、その結果によって指定された枝を選択する。この操作を繰り返すことにより、木構造の最も下に記されている物

体の姿勢(視線方向)が決定できる。実際に開発されたシステムでは、2次元の濃淡画像に加え、3次元の奥行き情報を記録した距離画像が利用されており、正確な視線方向や3次元位置が認識できる。

4.2 論理式による知識表現

工業部品のようにその形状が明確に規定されている対象を認識するには4.1のような方法が有効であるが、地図や航空写真から家や道路を認識するには幾何学的形状のほかに対象間の空間的、意味的關係に関する知識が重要となる。

Reiter ら¹⁰⁾は、線図形として表された地図から、海岸線や道路、川などを認識するシステムにおける知識を第一階述語論理¹¹⁾によって記述することにより、認識、理解過程のもつ論理的意味が明確化できることを示した。

まず図形世界の知識としては、

$$(\forall x) \text{image-object}(x) \equiv \text{chain}(x) \vee \text{region}(x)$$

(すべての画像中の対象は、線分(chain)か領域(region)である)

$$(\forall x) \neg(\text{chain}(x) \wedge \text{region}(x))$$

(線分と領域は異なる)

のような基本図形特徴の定義を表す論理式と、線分(chain)と領域(region)の空間的關係を表す述語

bounds(c, r): 線分cが領域rの境界線である

chi(c, c'): 線分c, c'が交差する

などを用いて認識すべき図形データの構造を記述する。

地図世界の知識は

$$(\forall x) \text{LINEAR-SCENE-OBJECT}(x)$$

$$\equiv \text{ROAD}(x) \vee \text{RIVER}(x) \vee \text{SHORE}(x)$$

$$(\forall x) \text{AREA}(x) \equiv \text{LAND}(x) \vee \text{WATER}(x)$$

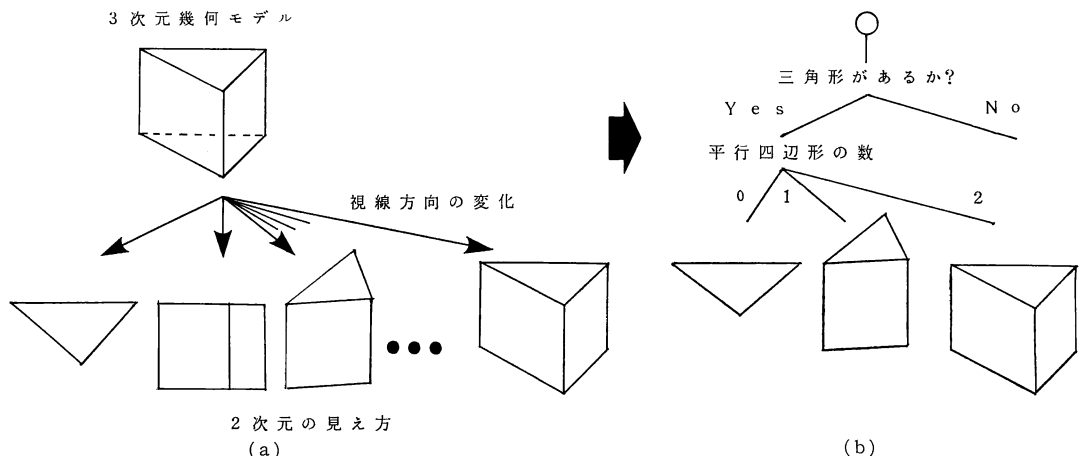


図6 3次元幾何モデルのコンパイル

のような認識対象の定義と、それらの間の空間的関係を表す述語

CROSS (x, y) : 対象 x, y が交差する

LOOP (x) : 対象 x はループ状になっている

を用いて記述される. また, 図形世界と地図世界の間の写像関係は

$(\forall x) \text{image-object}(x) \supset \text{SCENE-OBJECT}(\sigma(x))$

$\wedge \Delta(x, \sigma(x)) \wedge [(\forall y) \Delta(x, y) \supset y = \sigma(x)]$

といった論理式で表現される. ここで, σ は図形世界の要素 x を地図世界の要素に変換する関数, $\Delta(x, y)$ は図形世界の要素 x と地図世界の要素 y の間に写像関係が成り立つことを表す述語である.

以上のような論理式を基に論理的推論を行うことにより, 与えられた線図形の各要素に対して地図世界での対象物としての解釈が求められる. たとえば a を線分とすると (chain (a) が真ならば a が道路を表す ($\Delta(a, \sigma(a)) \wedge \text{Road}(\sigma(a))$ が真である) ことが証明される.

4.3 代数的記述による知識表現

2次元, 3次元の点, 線, 面といった幾何学的対象の位置や形状, それらの間の幾何学的関係, さらに3次元対象世界と2次元画像世界との間の写像関係は, 変数を含む等式によって記述できる. そこで, 画像特徴や認識対象に関する知識や問題の前提条件, さらにそれらをもとに証明したい命題をすべて等式として表現し, 代数的計算によって命題の証明を行うことが考えられる. こうしたアプローチは幾何学的推論¹²⁾(geometric reasoning)と呼ばれ, 1980年代後半から活発に研究が進められている. たとえば,

「線分 AB と AC が互いに直交しているとき, AC と直交する線分 CD は AB と平行であることを証明せよ」

という問題は, 図7のように座標を取ると,

前提条件: $x_1x_2 + y_1y_2 = 0,$

$$x_2(x_3 - x_2) + y_2(y_3 - y_2) = 0$$

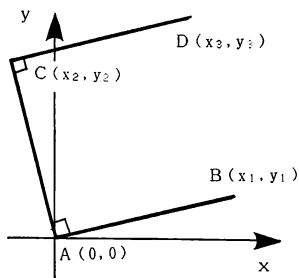


図7 線分の幾何学的関係

証明したい命題: $y_1(x_3 - x_2) - x_1(y_3 - y_2) = 0$

といった等式で表せる.

このような等式集合で記述された問題を自動的に証明するアルゴリズムとしては, 次のものがある (アルゴリズムの詳細については文献 12) 参照).

(1) 多項式の疑似徐算を利用した Wu の方法: 証明したい命題の左辺の多項式を前提条件の左辺の多項式で割って行き, 割り切れれば命題が証明できたことになる.

(2) 等式集合の整合性を調べる Gröbner 基底アルゴリズム: 証明したい命題の否定と前提条件の左辺の多項式からなる多項式集合の Gröbner 基底が1を含んでいれば (矛盾していれば) もとの命題が証明できたとする. たとえば, 上の命題の否定は次のように表せる.

$$(y_1(x_3 - x_2) - x_1(y_3 - y_2))z - 1 = 0$$

(z は新たな変数)

こうした代数的記述を用いると, 論理式ではうまく表現できなかった座標変換に関する知識や, 幾何学的対象の特徴や構造情報が厳密に記述できる. しかし, 実際の画像解析, 理解では, 画像から抽出された画像特徴の位置, 形状は雑音による誤差を含んでおり, 厳密な幾何学的推論は失敗する. 誤差を含む不正確な情報に基づく幾何学的推論をいかに実現するかは今後の課題である.

5. おわりに

本解説では, 画像解析, 理解のための知識表現と推論手法を紹介した. いずれの方法も一長一短があるが, 1で述べたように知識の宣言的記述を目指すことによって初めて, 科学としての画像の認識, 理解の基礎が確立されると考えられ, 今後の研究の進展を期待したい.

文 献

- 1) “画像処理エキスパートシステム特集”, 情報処理学会論文誌, 29, 2 (1988).
- 2) T. Matsuyama: “Expert systems for image processing: Knowledge-based composition of image analysis processes,” Comput. Vision Graphics Image Process., 48 (1989) 22-49.
- 3) 末田直進: “画像処理エキスパートシステム”, 映像情報, 17, 9 (1985) 19-22.
- 4) 坂上勝彦, 田村秀行: “処理モジュールの構造的知識を利用した画像処理プログラム自動合成システム”, 情報処理学会論文誌, 26, 4 (1985) 652-661.
- 5) L. J. de Haas: “Automatic programming of machine vision systems,” 10th International Joint Conf. on Artificial Intelligence (1987) pp. 790-792.
- 6) 長谷川純一, 久保田浩明, 鳥脇純一郎: “サンプル図形提示方式による画像処理エキスパートシステム IMPRESS”, 電子情報通信学会論文誌, J70-D, 11 (1987) 2147-2153.

- 7) A. M. Nadif and M. D. Levine: "Low level image segmentation: An expert system," IEEE Trans. Pattern Anal. Mach. Intell., **PAMI-6**, 5 (1984) 555-577.
- 8) 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, **27**, 2 (1986) 191-204.
- 9) K. Ikeuchi: "Generating an interpretation tree from a CAD model for 3D-object recognition tasks," Int. J. Comput. Vision, **1**, 2 (1987) 145-165.
- 10) R. Reiter and A. K. Mackworth: "A logical framework for depiction and image interpretation," Artif. Intell., **41** (1989) 125-155.
- 11) 野口正一, 滝沢 誠: 知識工学基礎論 (オーム社, 東京, 1986).
- 12) D. Kapur and J. L. Mundy (eds.): *Geometric Reasoning* (MIT Press, Cambridge, 1989).