

機械部品の認識

谷内田正彦 辻 三郎
(大阪大学 基礎工学部)

1. まえがき

積み木などの簡単な物体を認識する機械の目は種々研究されており、すでにいくつもの単純な作業に実験的に応用されている^{(1)~(3)}。しかし、このような機械の目を組立てるとの検査などの実際の産業用の目的に利用するには、複雑な実物体を認識できる視覚システムが必要となる。

実際の産業用部品の認識を扱った例として、Olsztyń 等⁽⁴⁾の研究がある。しかし、このシステムは自動車の車輪をハブに挿入するという特定の用途に開発されたもので、多くの種類の産業用部品を認識できるような一般性はない。

各種の複雑な機械部品を認識できる一般性を機械の目に与えるには、次のような問題点を解決する必要がある。

- (1) 機械部品はその構造が複雑であり、しかも油、ゴミ、ハイライト(光源の反射)などの雑音が多い。したがって、このような複雑な入力画像から本質的な特徴を抽出できる有効な認識方式が必要となる。
- (2) 認識すべき対象の種類が多く、例えば一つの製品を組立てるのに必要な部品の種類は30種以上にのぼることもあり、これらの部品をすべて識別できなければならない。しかも、機械はそれらを分類するだけでなく、その位置、方向、姿勢と組立てに必要なネジ穴の位置などを計測する必要がある。
- (3) 機械の目に実用性を持たせるには、ある特定の製品だけでなく、各種の製品を扱える汎用性が必要である。したがって、新しい対象に対しても、容易にその構造を学習し適応できる柔軟性が必要である。

ここでは、以上のような点を中心に、汎用性を持った機械の目のシステムについて述べる。しかし、積み重ねた部品を認識するのは、物体の分離(segmentation)などの困難な問題がある。物体を上方から見ると、しかもそれらが互いに重なり合っているという条件をつける。この条件は物体を上方から見た場合、重ねて見えることが少なく、またたとえ重ねていてもマニピュレータによりそれらを分離できるので、実用上はあまり厳しい条件ではない。

2. 認識システムの概要

従来、パターン認識には前処理、特徴抽出、認識というBottom-upな処理方法が多く用いられて来た。例えば、多面体について言えば、まず入力画像を線図形に変換し、その頂点の形などの特徴から認識が行われる。しかし、この方法は簡単な対象に対しては有効であるが、機械部品のような複雑な対象に対しては、認識に必要な特徴を抽出することが困難であり適用できない。一方、顔写真⁽⁵⁾、X線写真^{(6),(7)}などの複雑な画像を分析する有効な方法としてTop-downな処理方法が用いられている。例えば、人間の顔写真の場合、その各部分構造(目、鼻、口など)の位置関係はほぼ一定である。したがって、まず簡単に見出される特徴を求め、次にそれに基づいて顔の別の部分の位置を予測して求めることができる。

このように概略の位置とか性質の解っている特徴を抽出することは、河らの知識なしに特徴を抽出する場合と比較して、信頼は良く特徴を抽出することが出来る。この方法は対象が顔写真というように限定されている場合に、その各部分構造を固定するのに有効である。しかし、機械部品のように対象の種類が多い場合に、ある入力物体が“何である”かを認識するには適用できない。すなわち、入力物体を認識するにはその特徴が必要であるが、一方これらの特徴を信頼性良く抽出するにはその物体が何であるかが解っている必要があるという矛盾が生じる。

人間がこのような複雑な物体を認識する場合を考えると、Bottom-upな方法とTop-downな方法を組合せたような処理方法をとっていると思われる。すなわち、まず入力物体から認識の手掛りとなるような主要な特徴を見つけ、それに基づいていくつかの候補を想定する。次に、これらの候補についての知識を用いて、入力物体を識別するにはどのような特徴が必要であるかを予想しつ特徴を抽出し、候補の数を減らして行く。

ここでは上に述べたような処理方式を用いている。図1にそのシステム構成を示すが、図中の太線は学習に於ける情報の流れを示す。モデルはシステムがあらかじめ知っている各機械部品についての知識であり、その構造などが書かれている。シーン解析プログラム(Scene Analyzer)は、入力物体に関して現在得られている特徴とモデルとを比較して、次にどのような特徴を抽出するのが最も有効であるかを解析するプログラムである。また、スペシャリストは解析プログラムから要求された特徴を入力画像から抽出するための特徴抽出ルーチンの集まりである。

次に、認識に於ける処理の流れを述べる。まず入力画像から、認識への糸口となるような主要な特徴として、物体の大きさ、外形などの特徴が抽出される。次に、これらの特徴と各モデルに記載されている特徴とを比較し、類似した特徴を持つモデルを候補として残す。機械部品は類似した外形を持つものが比較的多いので、このような一次処理で4~10個のモデルが候補として残る。解析プログラムはこれらの候補モデルの構造を調べて、次にどのような特徴を抽出するものが認識に最も有効であるかを予測し、特徴抽出ルーチンの名前とその予測位置をスペシャリストに送る。適当なスペシャリストにより抽出された特徴は解析プログラムにより評価され、再び次に抽出すべき特徴が決定される。このプロセスは候補の数が一つになるまで、すなわち入力物体が認識されるまで続けられる。

認識が終了と、次に同定

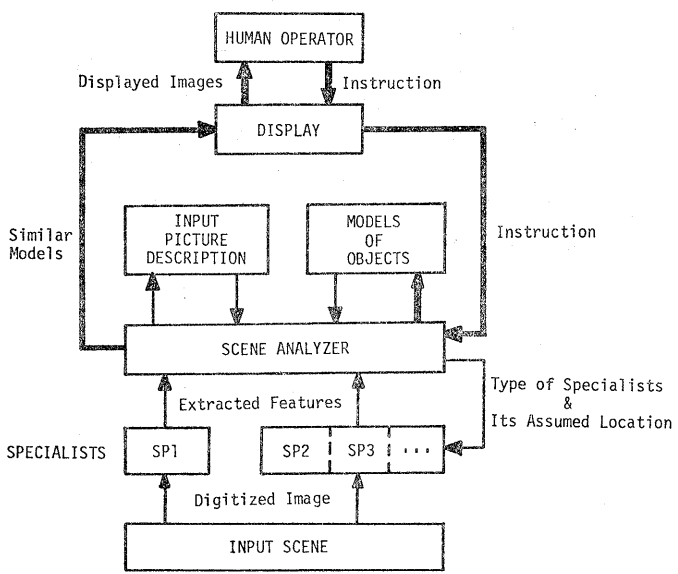


図1. システム構成

の段階に移る。これは認識には必要ではないが、機械部品を組立てるのに必要なネジ穴などを固定して、その位置を求めるプロセスである。すでに入力物体は識別されているので、そのモデルの記述に従って Top-down 的に処理が行われる。

機械の目に実用性を持たせるには、特定の対象を処理できるだけでなく、多種多様な対象にも簡単に適用できる汎用性が必要である。本システムでは、問題解析プログラム、スペシャリストなどの処理プログラムは対象に依存しない構成をとっているので、対象が変わってもプログラムは変更する必要はなく、そのモデルをデータとして与えるだけで良い。しかし、それでも対象物体の数が多くなると、モデルを逐一作成するのは容易ではない。そこで、本システムでは機械に実例物体を示すと、そのモデルが自動的に作成される構成をとっている。図2に新しい物体を学習する時の処理の流れを示す。

実例物体が提示されると、すでに述べた認識システムにより、すでに学習している知識(モデル)を用いて、新しい物体の認識を試みる。ここで、新しい物体を識別できれば、次にネジ穴などの組立てに必要な特徴の位置を教示する。しかし、すでに学習している知識だけでは、新しい物体を識別できない場合、それらを区別するのに重要な特徴を人間が教示する。すなわち、識別できない類似物体がディスプレイ上に表示されるので、それらを識別するのに重要な特徴をカーソルを使って教示する。教示された特徴は認識システムにより抽出され、再び新物体の認識が試みられる。このループは新物体が完全に識別されるまで続けられる。識別が終了と、組立てに必要なネジ穴などの位置をディスプレイを通して教示する。この過程で抽出された特徴は新しい物体のモデルとしてモデルのリストに追加される。

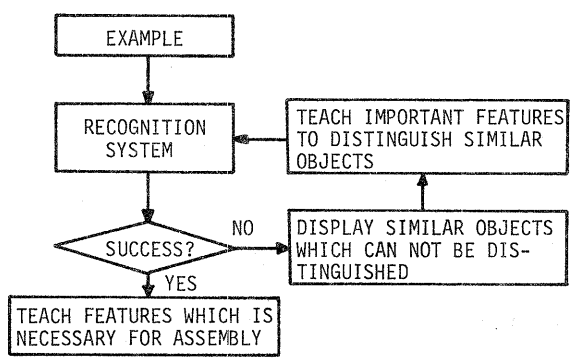


図2. 学習に於ける処理の流れ

3. 認識システム

3.1 前処理、一次処理

図3に本システムで対象としている機械部品の例を示す。同図は工業用エンジンを構成している部品の例である。入力画像はTVカメラにより上方から入力され、AD変換器により各点もビットにデジタル化される。サンパル方式として、全画面を低分解能でサンパルするモードと、画面の4分の1の部分を高分解能でサンパルするモードの2つが用意されており、両モードとも128×128点のデジタル画像がハッ

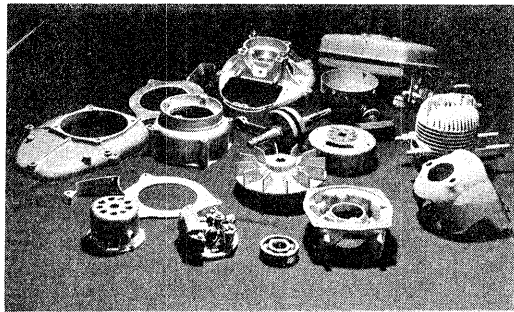


図3. 機械部品の例

ファ・メモリに記憶される。

前処理では、まず各物体の概略の位置を知るため、データを低分解能で入力し、それらの輪郭を求める。輪郭は、全体の画面の明るさのヒストグラムからしきい値を決め、そのしきい値以上の明るさを持つ点を追跡することにより求めることができる。

一次処理では、前処理で検出された各物体のデータを高分解能で入力し、その輪郭をより正確に求め直す。なお一次処理では輪郭を追跡する時、前処理のように一定のしきい値を使わずに、背景と物体の明るさの変動に適合して、しきい値を動的に決定している。一次処理への入力画像例と検出された輪郭を図4に示す。

入力物体の輪郭が抽出されると、次にその大きさ、丸さ、形などの性質が計算される。輪郭の形はその重心を原点とした極座標で表される。^(9,10) すなわち、図5に示すように、角度ごとの $r(\theta)$ の値に対して、重心から輪郭までの距離 $r(\theta)$ が計算される。こうして入力物体の大きさ、丸さ、形などの性質が計算されると、それらはモデルと比較される。まず計算の簡単な大きさ、丸さが比較され、類似した値を持つモデルに対して、 $r(\theta)$ グラフが比較される。

入力物体とモデルの $r(\theta)$ グラフをそれぞれ $r_I(\theta)$ 、 $r_M(\theta)$ とすると、これらは $r_M(\theta)$ を左にシフトすることにより比較される。ここで、 $r_M(\theta)$ のそれぞれのシフト角 θ_s について、2つのグラフの距離 $D(\theta_s)$ を

$$D(\theta_s) = \sum_{\theta=0}^{120} |r_I(3n) - r_M(3n + \theta_s)|$$

とすると、 $D(\theta_s)$ の最小値が一定のしきい値より小さい時、2つのグラフは類似しているとして定義する。こうして入力物体の輪郭の諸性質と類似した値を持つモデルが候補として選択される。

図4の入力物体に対する候補を図6に示す。図から解るように、輪郭の情報では識別できない物体が候補として残る。なお、この場合は4個の候補が残っているが、70個程度の候補が残るような物々もある。入力物体と図6(b)の物体との $D(\theta_s)$ グラフを図7に示す。このように外形が類似していると $D(\theta_s)$ グラフは深い谷を持つ。この時のシフト角をマッチ角 θ_M と呼ぶと、 θ_M

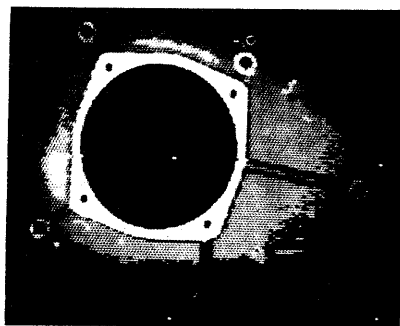


図4. 一次処理における入力画像と輪郭を抽出した結果

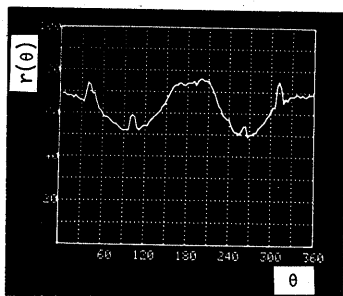


図5. $r(\theta)$ グラフ

図6. 図4の入カ物体に対する候補

- (a) 部品番号20の底面
- (b) 部品番号20の上面
- (c) 部品番号21の底面
- (d) 部品番号21の上面

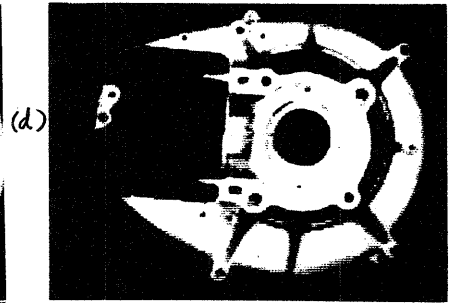
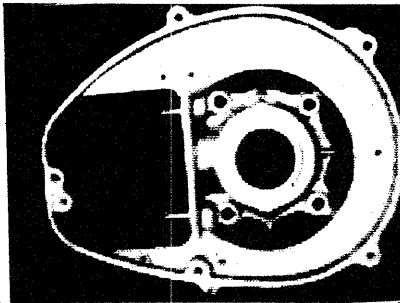
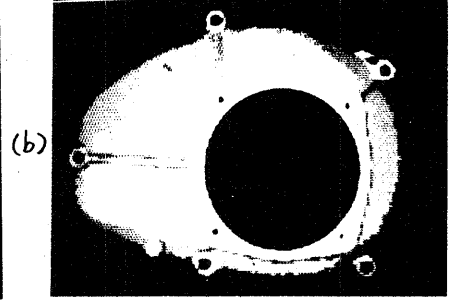
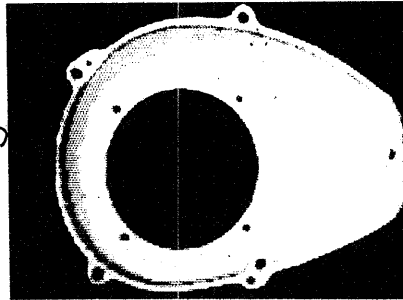


図7. 入カ物体と図6(b)の物体とのD(0s)グラフ

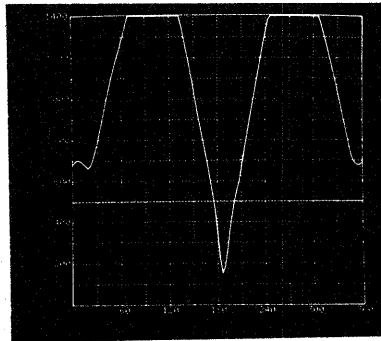
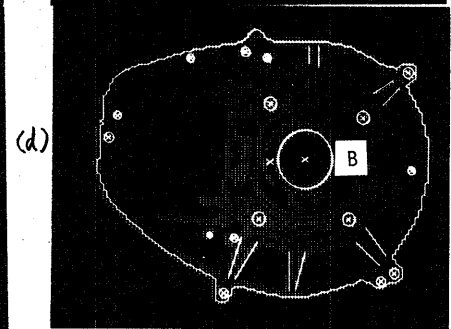
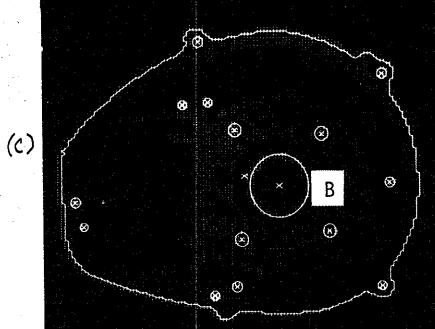
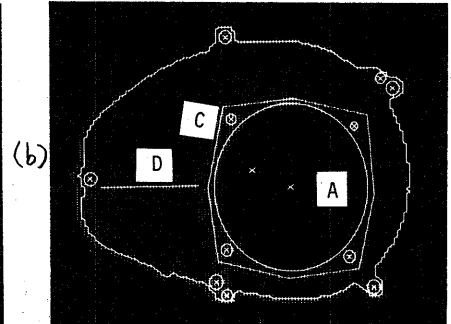
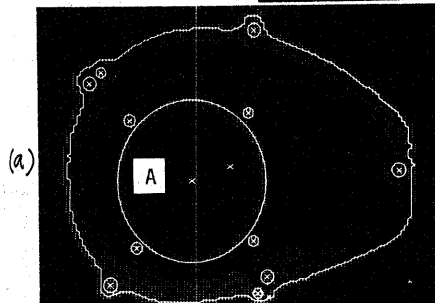


図8. 図6の物体に対するモデル一解り易いように一定の微分値以上の画像と共に示してある。

- (a) モデル40
- (b) モデル41
- (c) モデル42
- (d) モデル43



はモデルに対する入力物体の相対的回転角を示している。

3.2 モデルの構造

モデルは本システムにおいて中心的役割を果たすので、まずモデルの構造を明らかにしておく。各物体にはいくつかの面があり、それぞれ面のパターンは異なっている。各面に対してモデルが記述されている。ここでモデルの集合を M とすると、 M は各物体の各面に対するモデル M_1, M_2, \dots のリストで記述されている。また各モデル M_i は次のようなリストで構成されている。

$$M_i = (\text{PARTNO} \quad \text{FACE} \quad F_1 \quad F_2 \quad \dots)$$

PARTNO は部品番号を示し、FACE はどの面であることを示す。また F_1, F_2, \dots はその物体を記述する特徴である。なお、能率良く認識を行うには、各物体を識別するのに必要十分な特徴のみを、より重要なものから詳細なものへと順番にモデルに記述しておくのが得策である。したがって F_1, F_2, \dots はこのような基準に沿ってモデルに記述されている。

各特徴 F_i には $F_i = (\text{FTYPE} \quad \text{PLIST} \quad \text{ALIST} \quad \text{RSCORE})$ などの記述が付している。ここでは特徴として輪郭、穴、特徴線、模様などの数種類の特徴を使い、FTYPE は特徴の種類を示す。PLIST は各特徴の位置を示すパラメータで、図9に示すようにその代表的な点が重心を原点とした極座標で記述されている。ALIST は特徴の性質を示すリストで、各特徴の種類に対するALISTは表1に示されている。RSCORE はその特徴を抽出する時の信頼性を表わし、どの程度信頼性良く抽出できるかによって1~5の値が与えられている。

図6の物体に対するモデルを図8に示す。なお、図中の小円は認識には必要でないが、組立てに必要なネジ穴を示している。また表2に図8(b)のモデルの各特徴に対する実際の値を示す。

OUTLINE	AREA THINNESS $r(\theta)$ GRAPH
HOLE	AREA THINNESS RADIUS
LINE	AVERAGE GRADIENT VALUE
TEXTURE	DIRECTION

表1. ALIST

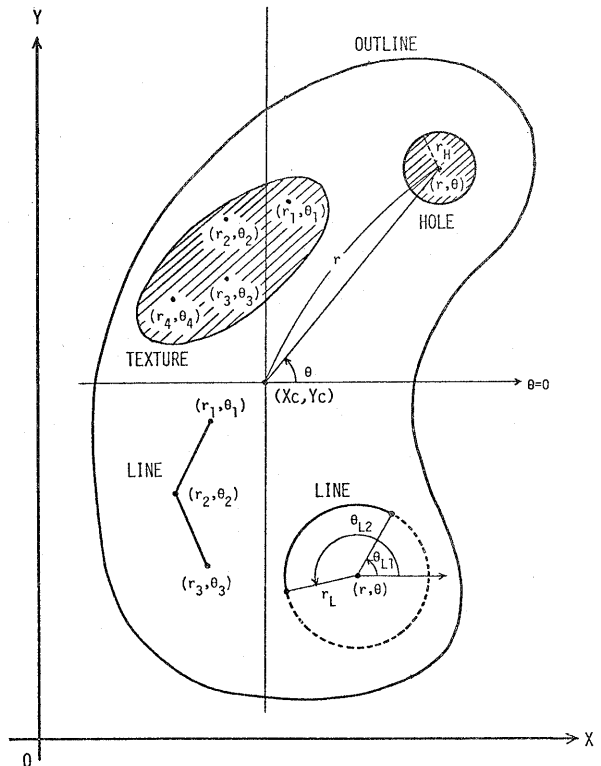


図9. PLIST

3.3 認識と同定

図10に認識における処理の流れを示す。まず一次処理で得られた輪郭情報から候補モデルが決定されるが、その手順を次に示す。

(1) Difference Analyzerは各モデルに対して、入力物像に対する類似度を示すスコア表を持っており、それらには最初一定の値が与えられている。

(2) 新しく特徴が抽出されると、それはモデルと比較され、類似した特徴を持たないモデルに対して、その特徴に付かないしている RSCORE の値がそのモデルのスコア表から差し引かれる。

(3) 次にスコア表を調べ、ある一定の値以上のスコアを持つモデルを候補モデルとする。

こうして候補モデルが決定されると、それらのモデルを調べ、次にどのような特徴を抽出するのが入力物像を認識するのに最も有効であるかが特徴提案プログラムにより提案される。この決定は次のような優先順位で行われる。

- (1) 候補モデルの内、最高のスコアを持つモデル
- (2) 上のモデルの内、より多くの共通した特徴を持つもの
- (3) より高い信頼性を持つもの (RSCORE から解る)
- (4) より処理時間の早い特徴 (穴、特徴線、模様 の順)

こうして提案された特徴は適当なスペシャリストにより抽出される。ここで、使用するスペシャリストの種類は FTYPE により、またその予想位置は PLIST とマッチ角 θ_M により与えられる。抽出された特徴の位置と性質はモデルに記述されている PLIST と ALIST の値と比較され、その結果により次の3通りの処理が行われる。

(a) もし、それらが類似していたら、観測された位置パラメータに基づいてマッチ角を補正して特徴提案プログラムに行く。特徴提案プログラムはそれらのモデルのさらに下の特徴を調べて、より細かい特徴を提案する。

(b) そうでない時は、その特徴の RSCORE の値をスコア表から引く。もしスコアがなお、大きい

MODEL 41 PART NO. 22 TOP FACE

COMPONENT	FTYPE	SCORE	PLIST	
1	OUTLINE	5	-	SIZE (882) THINNESS (7) r(0) GRAPH
2	HOLE	5	(18,338)	SIZE (259) THINNESS (9) RADIUS (34)
3	LINE	3	(29,116), (34,58) (55,23), (55,349) ⋮	AVERAGE GRADIENT (48)
4	LINE	3	(67,185), (25,194)	(15)
5	SHOLE		(56,102)	SIZE (22) THINNESS (9) RADIUS (3)
			⋮	

表2. モデル41

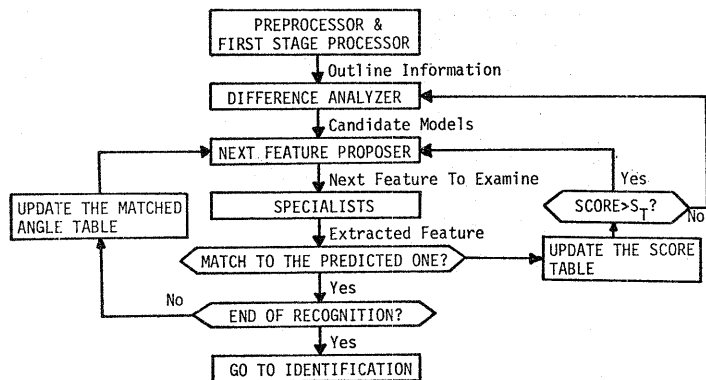


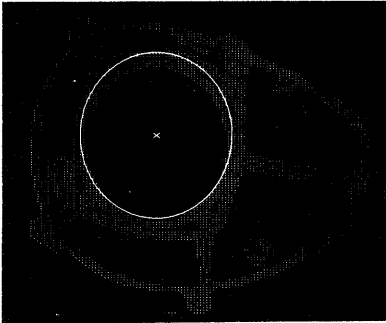
図10. 認識における処理の流れ

値より大きい値を持っている場合は特徴提案プログラムに行き、そうでない時は次の(c)へ移る。

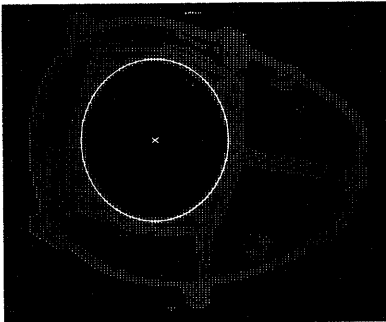
(c) 現状までに観測された情報をもとに、候補モデルを選択し直し、他の特徴を提案する。

このループはあるモデルが終端ノードに到達するまで、すなわち入力物体が認識が終るまで続けられる。こうして入力物体が認識されると、次にネジ穴の位置の同定に移る。この時、すでに物体は識別されているので、そのモデルの記述に

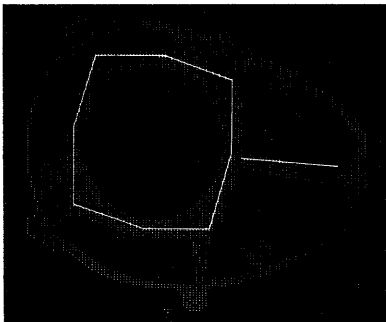
(a) 特徴提案プログラムにより提案された特徴とその予測位置



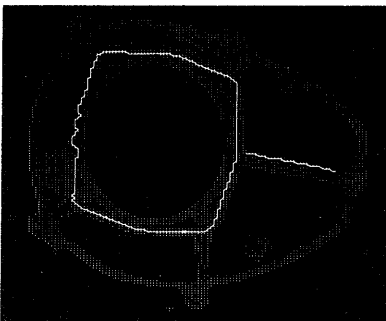
(b) HOLE detectorによる抽出結果



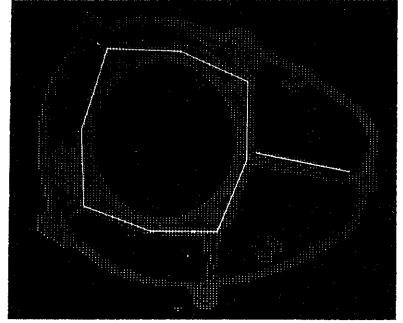
(c) 特徴提案プログラムにより提案された特徴とその予測位置



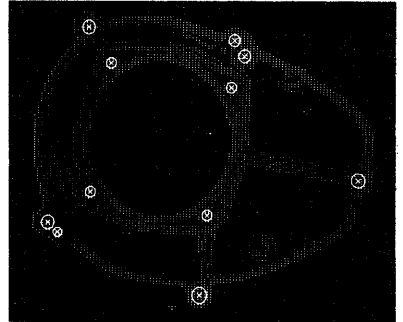
(d) LINE finderによる抽出結果



(e) 線分で近似した結果



(f) SMALL HOLE detectorによるネジ穴の抽出結果



(g) 最終結果

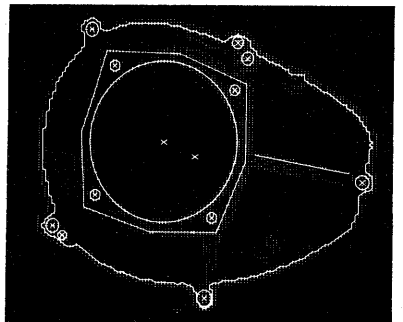


図 11 図 4 の入力物体に対する認識過程の例

したがって、ネジ穴を検出すれば良い。

次に、図4の入カ物体に対する実験結果を図11に示す。一次処理が終了段階での候補モデルは図8に示したが、次に抽出すべき特徴として、A、Bの穴が特徴提案プログラムにより与えられる。しかし、この2つは全く同じ優先順位であり、どちらを選択するかは決定できない。このような場合、ランダムに選択されるが、たとえBの穴が選択されても、このような穴は入カ物体に存在しないのといわれればともAの穴が提案されるようになる。穴の場合、その重心が予想位置として与えられ、その点を出発点として、一次処理と同様な方法でその輪郭が検出される。次に、輪郭から表1のような諸性質が計算され、モデルと比較される。モデル41から予想される穴を図(a)に、また抽出された穴を図(b)に示す。これらの値は類似しているので、マッチ角を補正した後、モデル40と41の特徴がさらに調べられる。図8から解るように、これら2つを区別する特徴として、C、Dのような特徴線が特徴提案プログラムにより提案される。図(c)にその予想位置を示すが、マッチ角が補正されているので、実際の位置と比較的に良く合っている。

特徴線の抽出には、まず予想位置を囲む帯状の探索領域を設定し、動的計画法⁽¹⁾を用いてその中での最適の曲線を決定している。次にそれらを線分で近似し、その時の端点を位置パラメータとし、その平均値を特徴値として抽出する。図(d)は検出された特徴点を示し、(e)はそれらを線分で近似した結果を示す。モデル40はこのような特徴線を持つといっているので、結局モデル41として認識される。すなわち部品番号20の上面と認識され、同時にその位置、方向も計測されている。入カ物体が認識されると、そのモデルの記述に従って、ネジ穴が検出される。図(f)はこうして抽出されたネジ穴を示し、(g)は抽出された特徴をすべて表示したものである。

4. あとがき

各種の複雑な機械部品を認識する機械の目について述べた。ここでは、比較的簡単な部品を例としてあげたが、かなり複雑な機械部品も認識することが出来る。認識時間は対象により変わるので一概には言えないが、小型計算機PDP-8/Eで実行した場合約30秒であった。処理時間は、高速パターン処理装置⁽²⁾などのハードウェアでスペシャリストを実行することにより、かなり改善されると思われる。

ここで述べたシステムの最大の制約は物体を上方から見て、しかもそれらが重なり合っていないという点である。この制限はすでに述べたように実際のシステムではあまり問題とはならないが、積み重なった部品を任意の方向から見て認識するのは、人工知能的な観点からも興味ある問題であり、今後に残された課題である。

なお本研究は文部省 科学研究費の補助を受けた。

[参考文献]

- (1) M. Ejiri et al: "An intelligent robot with cognition and decision-making ability", Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI), 2, p. 350 (1971)

- (2) P. H. Winston: "The MIT Robot", Machine Intelligence, 7, Edinburgh Univ. Press, p. 431 (1972).
- (3) A. P. Ambler et al: "A versatile computer-controlled assembly system", Proc. IJCAI, 3, p. 298, (1973).
- (4) J. T. Olsztyjn et al: "An application of computer vision to a simulated assembly task", Proc. Int. Conf. on Pattern Recognition, 2, p 505 (1974)
- (5) M. D. Kelly: "Visual identification of people by computer", Ph.D Thesis, Stanford Univ. (1970)
- (6) C. A. Harlow et al.; "The analysis of radiographic images", IEEE Trans. on Computers, C-22, 7, p. 678 (1973)
- (7) D. H. Ballard et al; "Hierarchic recognition of tumors in chest radiographs", Proc. Int. Conf. on Pattern Recognition, 2, p 258 (1974)
- (8) C. K. Chow et al; "Boundary detection of radiographic images by a threshold method", S. Watanabe (Ed.), Frontiers of Pattern Recognition, p. 61, Academic Press (1972)
- (9) A. Klinger et al; "Computer Analysis of chromosome patterns; Feature encoding for flexible decision making", IEEE Trans. on Computers, C-20, 9 (1971)
- (10) 依田, 江尻; "マッチング手法による物像認識の一手法", 昭47年信学会全国大会 (1972)
- (11) U. Montanari: "On the optimal detection of curves in noisy pictures", ACM, 14, 5, (1971)