

# 電統研における物体認識の研究

白井良明  
(電子技術総合研究所)

## 1. はじめに

当研究室は、生物の情報処理を研究するバイオニクスグループと、計算機による2次元のシーンを処理する物体認識グループで構成されるが、ここでは後者を紹介する。物体認識の研究を開始した1968年には、知能ロボットの眼を再現することを目指してゐた。1970年には簡単な多面体と曲面体を認識し、2次元ピクセルで物体をハンドリングする試みが行われた。その後はとくにロボットと関係付けることなく、視覚情報を用いた物体認識を研究してゐる。対象物体に対する拘束条件もなすべく緩やかなものとし、物体認識共通の問題を解決することに重点を置いてゐる。なお当研究室は現在通産省の大形プロジェクト「パターン情報処理システム」に参加してゐる。したが、電統研内の音声認識、図形認識などのグループと共に、共通の計算機システムを使用することに加ふる。以下に研究室のシステム構成と現在行つてゐる研究の概要を示す。

## 2. システム構成

現在使用の実験システムの構成を図1に示す。

### 2.1 計算機システム

HITAC-8250は、共通の計算機TOSBAC-5600と結合されており、パターン情報処理研究用の複合大形計算機システム(EPICS)の1つのノードシステムである。このEPICSは中央の大形計算機に、周辺の6個の小形集中形計算機を4ヤネルレベルで星形に結合した集中制御方式のインハウスコンピュータネットワークである。TOSBAC-5600とHITAC-8250間では、(1)双方向のリモートバックサービス、(2)プログラム間通信、および(3)ファイルデータの直接転送の各機能をサポートするソフトウェアが完備してゐる。

TOSBAC-5600は256K wordsの主メモリを持ち、バックジョブサービスと、タイムシェアリングサービスとを同時に行なう。バックジョブは最大70K~80K wordsのジョブが、タイムシェアリングサービスでは1ジョブ最大64K wordsまでのジョブが実行可能であるが、しばしば両方のユーザ間で、メモリとCPU時間の争奪合戦が起り、割当スケジュールの変更が行なわれる。端末はテレタイプ、文字ディスプレイおよびグラフィック端末が、現在47端末設置されており、ユーザは端末から、エディタによるファイルの編集、プログラムの作成および改善作業、プログラムの起動/停止、実行中のプログラムの制御/状態検知ができる。このシステムを使用した実感として、強力なエディタの機能と、大容量ファイルシステム(約120M words)が、作業能率に大きく貢献してゐる。特にHITAC-8250とリモートバックサービスがあることで、HITAC-8250のプログラムも、TOSBAC-5600の端末から作成することが出来る。

HITAC-8250は、物体認識研究グループが優先的に使用してゐる集中形計算機(256K bytes)で、オペレーティングシステムは、マルチジョブをサポートしてゐる。マルチジョブの内1つは、常時TOSBAC-5600との通信をこなしてゐる。

ユーザは残りの2つのクラスを使用できる。この2つのクラスには優先順位がつけられており、長時間の大ジョブはバックグラウンドジョブで走り、ダイバックのためジョブをフォアグラウンドで優先して走らせるといったルールと大まかに取決めで運用している。

### 2.2 実験装置接続用ケーブル

HITAC-8350に種々の実験装置を接続するためのケーブルで、計算機側のインターフェースは1つのセレクトチャンネルへ接続されたおり、実験装置側のインターフェースは、当研究室の標準仕様になっている。各個々の装置が接続できる。この仕様によれば、入出力は8ビット並列で、入出力以外にセンス信号が8ビットあり、転送速度は最高250 K bytes/secで実験装置側の転送速度に同期がとれ、60msで接続ケーブルを延長しても安定に動作する。このケーブルにより、ミニコンピュータとのDMAなども行なっている。

### 2.3 データ入力装置

物体認識の研究では、テレビカメラから入力したデータをを用いている。テレビカメラの映像信号をサンプリングしてA/D変換して計算機へデータを送り込むのが図1のケーブル接続セッションダイヤグラムである。距離情報はレンジファインクを用い

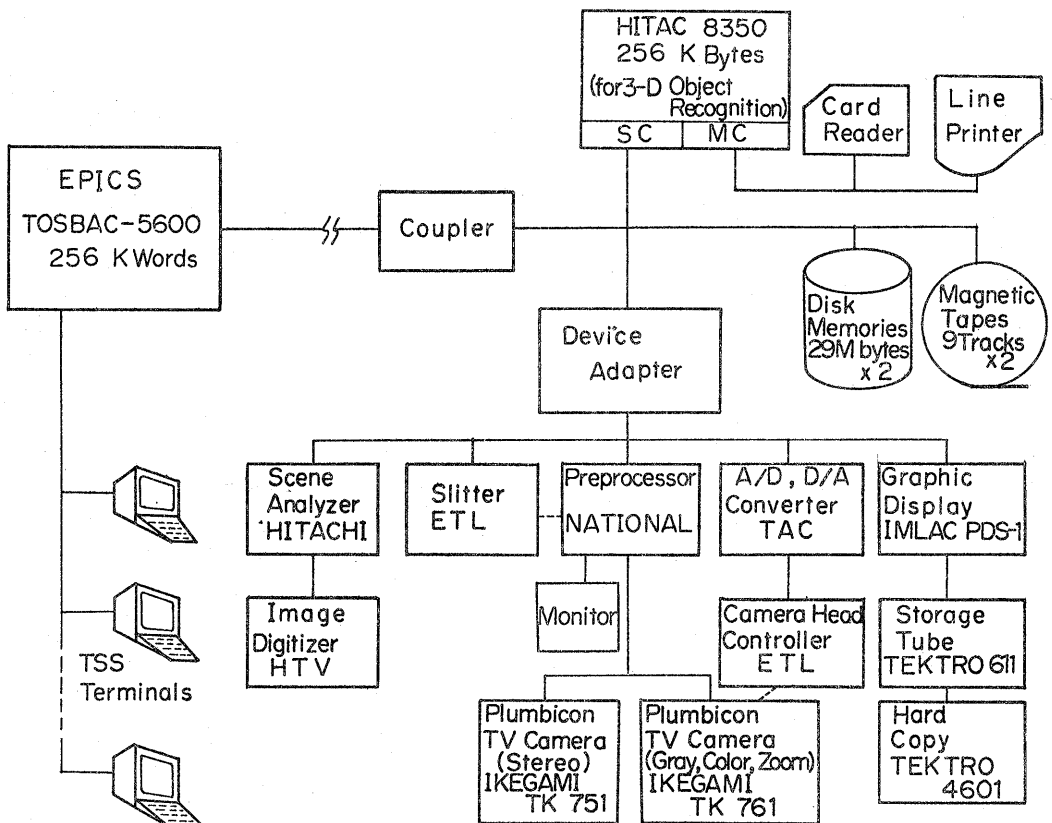


図1 実験システムの構成

2. スリットから計算機へ採り込まれる。

(1) TVカメラ

ガンビコン撮像管は、阻止型ターゲットであるために、暗電流が低く、 $\gamma$ -特性が1に近いので、カラーゲータを得るのに適していることが採用された。

走査方式は、走査線数256本をインターレースで行なうことである。

1インチのガンビコンを使用しているが、同じ径のビジコンと比べると、分解能が劣っている。更に高分解能で、1画面を1000×1000点以上でサンプル可能な入力装置を必要としているが、市販の撮像管あるいは固体素子などが存在しないのが実情である。

カメラの分解能を補うためには、対象物体を拡大して撮影する必要から、レンズ系にはズームレンズを用い、カメラ全体をパン・ティルト台に載せている。このレンズ系と、パン・ティルト台は計算機から制御することができるようになっている。

色情報を入力するために、赤青緑の3色のラッテンフィルタを使用している。ラッテンフィルタは時間とともに特性が変化しやすいため、取扱いに注意が必要である。特に青フィルタの退色が著しい。

距離情報はレンジファインダによる方式(後述)と、ステレオ方式による2得られられる。ステレオ方式では、右側のカメラを用いると両カメラの位置差(レジストレーション)が難しい。そこで、ここには単眼鏡に左右の画像を切換えて、中央に置かれた1台のカメラから画像を入力している。

イメージダイセクタ撮像管は任意走査が可能であり、ゲータサンプルのための周辺装置が単純なためという利点があるため試用している。現在使用中のカメラは、分解能が低いのが最大の欠点である。

(2) コリコセッサヒューマンアナライザ

コリコセッサは、ラスタ走査カメラ用で、1画面を256×256画素について明暗を6ビット(64レベル)に量子化する高速A/D変換器と、その256×256点から64×64点をサンプルして記憶できるICメモリから成っている。サンプル点の間隔の違いによる種類のサンプルモードがある。ICメモリはHITAC-8050から直接読み書きができるので、サンプルされたデータを計算機へ読み込ませることができる。

このICメモリの内容は、再びD/A変換して、モニターに濃淡図形として表示されている。従って、計算機からICメモリへデータを書き込むことにより、ビデオディスプレイとして使うことができるようになっている。

ヒューマンアナライザは、中に組み込まれたミニコンピュータで任意の走査モードで画像データを入力し、さらに処理することが出来る。

(3) スリット

レンジファインダとスリットで1つの距離情報入力システムを構成している。レンジファインダは、線状のスリット光を或る方向へ出し、それが対象物体へ投影されたものをステレオカメラで撮って、その物体までの距離を測定するシステムで、三角測量の原理を応用したものである。

スリットは画像の中から、スリット光が当たっている位置をオンラインで検出する。その検出は、相対変動度60%を1000 TVlineで実現できない入力装置であるが、市販の撮像管(1.5インチのビジコン)では、60%を得るには高さ400 TVlineを必要とし、1000 TVlineは数%以下に下がる。特殊な撮像管としては、NHK技研が開発したリターニームサコンがあり、これは1000 TVlineで80%以上の変動度特性を有するものがある。

また、の装置である。処理は全てデジタルで行ない、各走査線について、パルス状の波形とした信号の中で最大の明さのものの位置を検出して、画面に割振られた座標値を出力する。一般には、これがスリット光の位置に対応していきと云える。

## 2.4 テータ出力装置

物体認識の研究に扱うテータの性質上、表示するテータ量も多量である場合が多いので、そのような時には、蓄積型ディスプレイが適している。一方、イテミツクに変化する図形の表示や、インタラクティブな処理には、リフレッシュ型のグラフィックディスプレイが良い。この二つの異なる要求を満すために、別々の装置を導入したのでは、費用の点だけではない、プログラムの互換性から得策ではない。そのために我々は図1に示すように、IMLAC社のPDS-1というグラフィックディスプレイを選んだ。この機種には、蓄積型ディスプレイのインターフェースをオプションとして付加することが出来る。従って、リフレッシュ型のグラフィックディスプレイの特性を生かしながら、必要な図形を次々と蓄積型へ書き加えていくことにより、多量のデータの表示が容易で、さらにハードコピーユニットを付けて比較的密着なコピーも出来る融通性の富んだシステムとして、活用している。

## 2.5 画像処理用ソフトウェアパッケージ

研究の能率向上に大きく貢献している道具の一つに、ストアードデータシステムがある。これは、テレビ画像のデータの幾種類かをディスクファイルへストアして置き、通常のプログラムのデバッグの時、このデータを使えるようにしたソフトウェアシステムである。テレビカメラがポリプロセッサを介して、直接計算機へ繋がっている中で、一つ一つデータを採り込めようとするのではあるが、実際は、対象物体を設定し、適当な照明を行ない、装置の電源を入れたらと、かなりの手続きを必要とする。さらにプログラムのデバッグの場合、結果を再現する必要がある場合が往々にしてあるが、データが微細に変り、2しきるとそれが出来なくなるとも有り得る。このような場合、ディスクに予めストアしてあるストアードデータが高速に取り出せば、効率が向上するであろうことは想像に難くない。

このストアードデータにアクセスするには、関数型オペレータをCALLすればよいので、ユーザはまたかまた次元のレイに画像データが入っていると思え、プログラムの書き方はよい。このソフトウェアシステムでは、ディスクへのアクセスの際、データのバッファリングを行ない、スピードの向上を図っている。

現在20個の画像データが常時使用可能になっている。

距離情報のデータに関しても同様の工夫がなされている。

## 3. 明さデータを伴った認識

色や距離の情報を用いずに、明さだけに基いて複雑な物体を認識する研究がある。認識の方法は対象物体に依存するが、その方法の複雑さは必ずしも対象物体の見かけの複雑さに比例するものではない。一つの物体が他と区別できれば、その物体を分析して詳しい情報を得ることは比較的容易である。最も困難な点では、個々の物体を簡単に分離できない場合である。これは、いくつかの物体が重なって、

2 見えるシーンを扱ひ、手元の水に物体をレーンの中から見つける問題を想定してゐる。図2に全体の処理のフロー図を示す。以下に各フローの概要を示す。

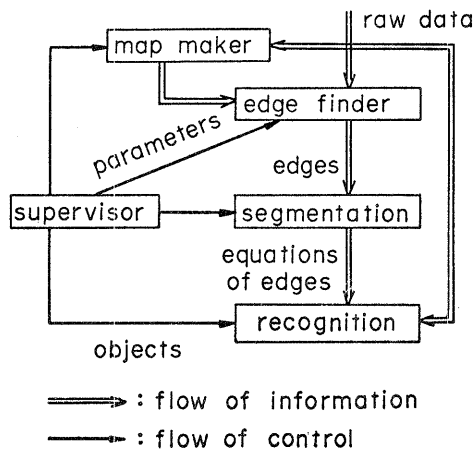


図2 明きデータを用いた認識

### 2.1 稜の抽出

ここでは多面体や曲面体の稜を求め、手元の水に明きデータの急変する点、あるいは左右の明きデータの変化が異なる点を探し、同じような点を滑らかに追跡して稜と思われし曲線を求める。手元の水に明きデータは、明きデータの物体の稜に付くのではなく、稜と同じような明きデータの変化をする曲線を念頭にとする。

### 2.2 マップ作成

稜の抽出に付いて候補点を探する場合、シーン全体をすべて走査することは避けるため、レーンを圧縮した後、その微分を行い、その結果をマップとして蓄えておく。一定の明きデータの差を閾値とする稜を求める場合には、マップを参照しながら候補点のありとない領域だけを探索。

### 2.3 セグメンテーション

抽出された稜は、その曲線を構成する点の集合で表わされたい。ここには各稜を、直線あるいは楕円の一部分で近似する。そのためには曲線を適当な点で分割する必要が有る。その曲線の曲率を用いて、曲線の分割(セグメンテーション)は、各セグメントの直線と楕円への分類を行なひ、さらに Deming の方法でセグメントの式を求める。また各セグメントは、その明きデータを示す向きが手元の水に示される。

### 2.4 認識

対象レーンと水の間に限り、その上に本、電気スタンド、電話機などが、普通の状態で置かれてゐるものと認識を行う。入力データの分解能の制限により、詳しく分析することはできない。そこで顕著な特徴を見つけ、それと類似して他の特徴を探索方法を用いてゐる。たとえば電話機を見つけるときには、手元の水にイヤルを求め、その代わりにあるレーンの境界を探索。場合により、手元水

マップをとして詳しく分析する必要もあるが、コンピュータの入力は試みている。

#### 4.5 スーパーダイザ

上述の処理をレーンレシカルに行うだけでは効率が悪い。適当な閾値を指定して様子を求め、もし十分な様子が得られれば、セグメンテーションと認識を行う。もし与えられた物体が認識できなければ、閾値を下げて再び様子を求め直す。このようにして、時間のかかる様子の抽出をなるべく少なくして全体の効率を上げる。

#### 4.6 距離データを用いた認識

認識すべきシーンの距離を検出し、距離情報から認識に役に立つ情報を抽出しようとする研究がある。距離データはシーンの幾何学的特徴を抽出するのに有効である。距離の検出には各種の方法があり、本研究でも両眼立体視の研究を行って、この場合は前述のレンジファインダを用いている。シーンの分析には又通りの方法が考えられ、一つは一樣な平面や曲面を抽出する領域法で、他は距離の不連続点あるいは傾きの異なる面の境界点を抽出する方法である。前者のアプローチは図5のようにして行っている。図において対象シーン(1)にスリット光を当てて距離データを求めた(2)を得る。この次元の点をいくつか集めて小さな領域(面素とよぶ)を作る。各面素は機械的に生成されるため、必ずしも一樣な面上に

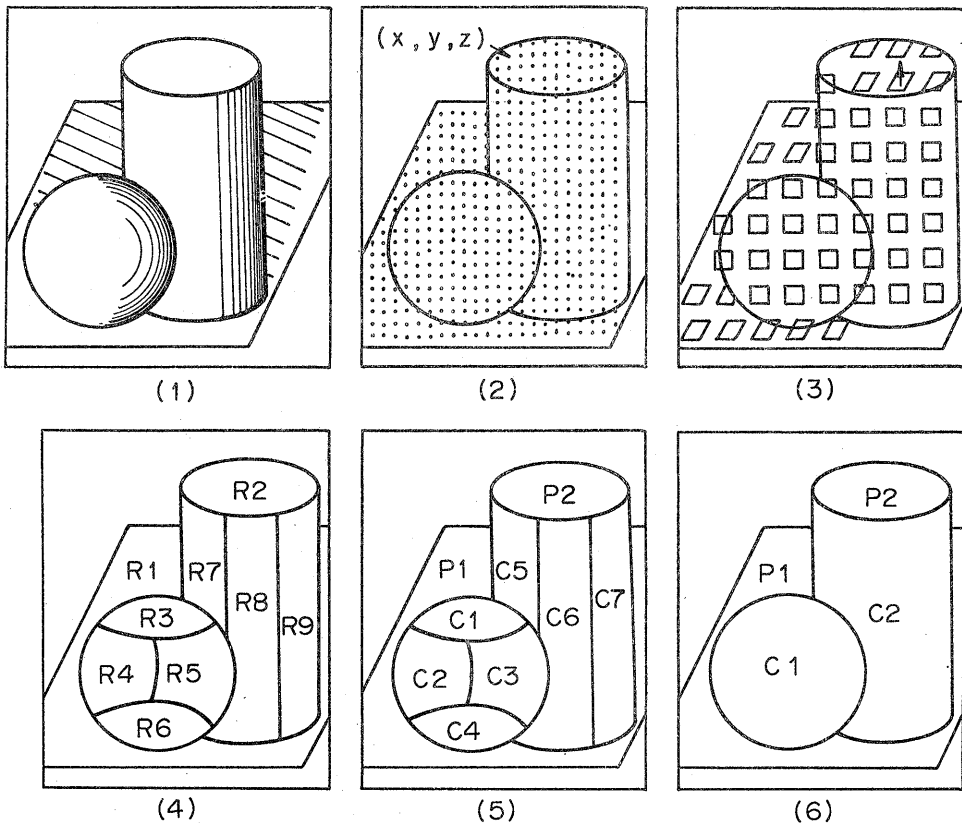


図5 距離データを用いた物体の記述

ではなく、面と面の境界に生成されることもある。面素を構成する点から面素の平面方程式を最小二乗法により求め、(3)を得る。つぎに隣接する面素の方程式がよく似ているものを結合して平面に近い領域を作る。この場合に最小二乗法の分散が大きい面素は異なる面の境界となるため、結合の対象とはしない。このようにして(4)を得る。各領域を構成する面素の方程式の係数と領域の大きさに基づいて、この領域を平面、曲面、不定の3種類に分類し、(5)を得る。つぎに曲面領域の近くにある他の曲面領域および不定領域に関して、それと滑らかに接続していかどうかを今エックル、もしそうであれば領域どうしを結合する。このようにして(6)を得るが、これは平面と滑らかに接続する曲面により、シーンを記述したことになる。物体認識にはこのように記述を用いることが有効であろうと思われる。

い、ほう列なアプロアチとして、一様な面上にばい点に着目する方法を研究している。まずシーンの中から次元位置の不連続点を探し、そこから不連続点を追跡して一連の不連続点を求める。この処理をくり返してシーンの不連続点をすべて求める。つぎに異なる面が交わり、2つの場所を探す。このためには一定領域の左右の面の傾きの差を計算してオペレータを適用する。このオペレータは $90^\circ$ 間隔で各方向に作用させて特徴点を探す。一点が見つければ追跡により、同じような特徴点を得ることができ、本方法は物体の稜を検出できるだけでなく、その稜が凸か凹か物体の境界であるかが同時にわかる利点がある。したがって、明瞭なエッジから抽出した線画より情報量が多く、線画の解釈も容易である。現在特徴点を求める場合に、それまでに求められた情報を利用して効率を上げることが考えられている。また曲線を含む線画もこの方法によれば求めることができるのではないかと思われている。

### 5. グラフを用いたパターンマッチング

入力データから何らかの特徴が得られたものとし、この特徴を用いて認識を行うための研究がある。特徴として明示データから抽出した稜線、距離データから抽出した領域などが考えられる。特徴の持つ情報は、特徴を構成する要素と各要素間の関係を表わすことができる。これをグラフで記述し、各要素を点に、要素間の関係を枝に対応させる。たとえばセグメントされた稜線を点とし、稜線の方程式と長さをもとの点の特性とする。また稜線の接続、平行などの関係も枝とする。このようにして分析すべきシーン下、レベルのつけられたグラフで表わされる。認識すべき物体も同様にグラフで表わされる。すなわち物体が1つだけシーンにあるものとし、理想的に稜線が抽出された場合のグラフを作ることもできる。このようにして認識すべき物体のモデルを作ったとき、未知のシーンの中からモデルと同じ物体を探し出す。シーンは多数の物体を含み、しかも物体が重なると見えのため、モデルと全く同じグラフを求めるだけでは認識が難しい。与えられたモデルと同じ物体を見つけるためには、モデルのグラフとシーンのグラフの部分マッチングを行う必要がある。すなわちモデルのグラフ  $M_G$  のグラフを  $M_G$ 、シーンのグラフ  $S_G$  のグラフを  $S_G$  としたとき、 $M_G$  と  $S_G$  が同形となることである。このようにしてグラフは一般には多数存在するが、ここには点の数が最大となるようなグラフ ( $M$  と  $S$  の最大共通サブグラフ) を求めたい。2つのグラフの最大共通サブグラフを求める手数は、2つのグラフの同形を調べる手数よりも複雑である。このアルゴリズムは「ウィラムの予想」を用い、能

率の下のものとして、<sup>10)</sup>

本手法を適用する場合、入力となる特徴の品質が認識の性能を左右する。たと  
えば、本来一つの要素となるべき線が、二つ以上に分離されたらば、その部  
分の単純なマッチングは成功しない。また交わらざるべき線が分離されたらば、表  
示は全く異なり、たまたまのちがいで済む。そこでマッチングを用いたパターンマッ  
チングを行う前に、入力を整形しなくてはならない。その処理は一般的には決定さ  
れない。たまたま人間の直観に合わせ、線を延長したり、切断したり、位置を  
修正してやる。

## 6. 物体の記述

物体認識には、未知のシーンを分析し、あらかじめ蓄えられている物体  
のモデルと照合する過程がある。この節では物体の最も重要な特徴を手かりとし  
て認識を進めたが、そこではモデルはプログラムの手録として蓄えられている。  
この節の前節ではグラフによる記述を用いて照合を行なった。物体が複雑なれば  
モデルを作るのも手数がかかる。人間がすべてを行うことは必ずしも得策で  
はない。グラフによる物体の記述では、モデルを作るために物体をいろいろな方  
向から見せ、シーンプラットフォームに、自動的に生成することを望みたい。こ  
こではある程度複雑な物体の記述に関する研究を紹介する。この研究が始ま  
った間もないうちに、具体的な結果は得られた。

第一のアプローチは二次元のモデル作成である。物体は二つの簡単な要素  
で表わされるものとする。すなわち物体を、多角柱、円柱、円錐、円板とを適当  
に組合わせ、組立てることにより、記述する。各要素は機械部品のように機械加  
工により、得られるものとする。したがって、物体をいくつかの方向から見れば、  
その形を推定することはできる。モデルを作るべき物体をいくつかの方向から見  
てその線画を作り、簡単な要素を見つけ、各要素間の幾何学的関係を決定してモ  
デルを作る。このモデル作成は必ずしも全自動である必要はなく、人間と会話的  
に行うことも考えられる。未知のシーンを分析してモデル作成と同じように簡単  
な要素の結合を記述し、モデルの記述とマッチングを行う、認識しようとするの  
が本研究の目的である。単純にシーンを分析して自動的に合致する記述を得るの  
は困難であるため、そこには何らかのトップダウン的手法が必要となる。すなわ  
ちモデルにより、シーンプラットフォームの制御されることとなる。

他のアプローチは完全に三次元の記述を行うのとはなく、自動的に認識に有  
用な記述を作り出すこととするのである。物体の像をいくつか見ると、そこから重  
要な特徴を抽出する。特徴として線画の単純な線（直線、円弧など）を用いる方  
法はすでに研究されたことが、二つ以上の関係を考慮し、自動的に物体固有の特  
徴を抽出する試みはあまり知られていない。以上の研究を入力として理想的な線  
画を用いるにはある程度実現可能であるが、入力も実際に得られる視覚情報であ  
れば、特徴抽出も完全に自動化することはできないため、柔軟性に富んだ方法を  
見つけなければならない。



### 参 考 文 献

- (1) 電子技術総合研究所彙報, ETL ロボット MK I 特集, 才35巻, 才3号 (1971)。
- (2) 川谷: パターン情報処理研究システム EPICS, 情報処理, 才16巻, 才2号, pp.145-158 (1975)。
- (3) 大島, 高野: 3次元物体認識システムのための特殊ハードウェア, 電統研彙報, 才37巻, 才5号, pp.483-501 (1973)。
- (4) 越川, 山上: 小型マイクロコンピュータ計算機用ソフトウエア, 電統研彙報, 才36巻, 才11号, pp.777-802 (1972)。
- (5) 白井, 藤原, 鎌山, 久良: 物体認識レポートプログラムの, 電統研彙報, 才37巻, 才8号, pp.785-805 (1973)。
- (6) Y. Shirai: Edge Finding, Segmentation of Edges and Recognition of Complex Objects, 4th IJCAI (1975)。
- (7) Y. Shirai: A step toward Context Sensitive Recognition of Irregular Objects, Computer Graphics and Image Processing, Vol. 2, pp.298-307 (1973)。
- (8) 安江, 白井: 物体認識のための両眼立体視, 電統研彙報, 才37巻, 才12号 (1974)。
- (9) M. Oshima and Y. Shirai: Representation of Curved Objects Using Three-Dimensional Information, USA-JAPAN 2nd Computer Conference (1975)。
- (10) 田村浩一郎: ふたつのグラフの共通サブグラフ, 電統研彙報, 才37巻, 才9号 (1973)。
- (11) 田村浩一郎: 線画の構造化, S50信学全大, 1173。