

物体認識に対する一つのアプローチ

白井良明
(電子技術総合研究所)

1. はじめに

バイオニクス研究室では物体認識の研究を行っており、その全体の紹介は[1]に示されるが、ここでは実際に入力装置を用いて得られる情報から、物体を認識する研究を紹介する。

2. 濃淡画像からの複雑物体の認識[2]

未知のシーンを分析して物体を認識するためには、(1)特徴の抽出と(2)特徴に基づいた認識、が必要である。しかしこの二つの処理を逐次に結合しただけでは高度な認識を行うことは困難である[3]。筆者は多面体認識を、稜の検出と稜の推定の二つのモジュールの協調により、行った[4]。他にも多面体認識に対しては、対象物体のモデルを用いて信頼性のある方法が試みられた[5, 6, 7]。しかし以上の手法を曲面を含んだ物体の認識に適用することは容易ではない。すなわちどのような特徴を抽出したらよいか、また得られた特徴をどのように解釈してシーンに関する推定をしたらよいか、が問題である。

もし色や三次元位置情報が利用できれば、認識を容易に行える場合がある。SRIでは、室内のシーンから机や椅子を見つけようとして研究を行っており[8]。机や椅子の高さは一定の範囲に決まっているので、各点の三次元位置から机の上面の候補をしぼることができ、さらに近傍の点を用いてそれが水平面上にあるかを知ることもできる。したがって限られた候補の中から物体を見つけることになり、簡単な処理となる。

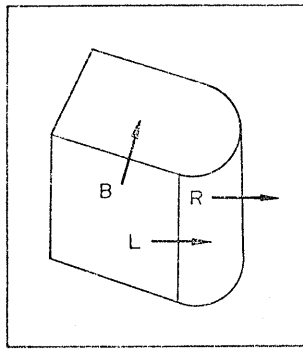
また物体の背景が一様で、物体のりんかくがうまく検出できないことがわかると、それによりりんかくを求めてから、そのりんかくの情報を利用して二つの処理を決定することもできる[4]。この仮定のもとでかなり複雑な機械部品を認識する研究も行われた[9, 10]。

いっぽう机上にある本や電話機などを濃淡情報だけから認識する場合、背景が一様であるという仮定を用いることができない。また一点の明るさだけからはシーンに関する情報を得ることができない。机上の物体の位置も一定ではないため、いわゆるトップダウン的手法を最初から用いることができない。複雑なシーンの特徴抽出の研究は多い[11, 12]が、それを用いてシーンを理解するまでには至っていない。筆者は曲面物体の特徴抽出を提案した[13]が、本研究はその延長である。全体のシステムは(1)特徴線の抽出、(2)特徴線の記述、(3)認識、(4)全体の統括制御、の4つのモジュールから構成される。以下に各部の概要を示す。

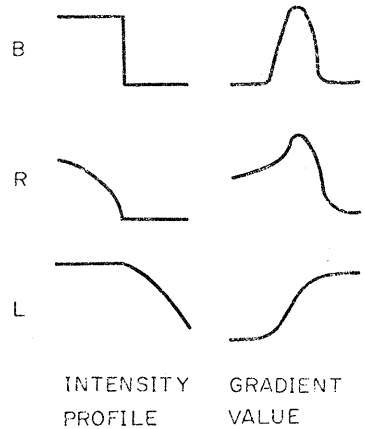
2.1 特徴線の抽出

ここでは曲面物体のりんかくと稜を検出することを目的として特徴線を抽出する。しかし物体表面が一様でなく、影などにより、得られる特徴線は必ずしも目的に完全には合致しない。特徴線を求めるためには、まずそれを構成する特徴点を求めることが必要である。ここでは二種類の特徴点抽出オペレータを用

いる。この場合も明るさや明るさの変化が急激に変わる点を求めるためのオペレータである。もし特徴線の方が未知であれば、二次元オペレータを用い、既知であれば、1次元オペレータを用いて計算の高速化を計る。図1に求めるべき特徴点のタイプ(B, R, L)を示す。オペレータは、この場合も明るさの微分を用いて、この場合、もし特徴線の方が未知であれば、二次元の空間微分による、2微分とその方向 (gradient) を求める。まず微分値が一定以上の点を求め、つぎに特徴線に垂直方向の微分値を調べ、図1のこの場合のタイプであるかを決定する。最初に求める点の微分値が、その特徴線の信頼性レベルを表わすものとして、この場合、



SCENE



INTENSITY PROFILE

GRADIENT VALUE

図1 特徴点の明るさの変化

間微分による、2微分とその方向 (gradient) を求める。まず微分値が一定以上の点を求め、つぎに特徴線に垂直方向の微分値を調べ、図1のこの場合のタイプであるかを決定する。最初に求める点の微分値が、その特徴線の信頼性レベルを表わすものとして、この場合、

図2に特徴線抽出処理の概要を示す。この処理は二つの主要部分で構成される。すなわち (1) 与えられた信頼性レベル以上の特徴線の核を求め、(2) その核を出发点として特徴線に沿ってトラッキングを行う。核は一定範囲内にある同じタイプの特徴点の集合として定義される。

参照マップは核を探し処理を速くするために用いる。原画像を4x4の小領域に分割し、マップの各点は、小領域内の点の微分値の最大値となるようにする。このようにすれば、与えられた信頼性レベル以上の特徴点を求める場合、マップを参照して意味のない領域の探索を避けることができる。しかしマップ作成に時間がかかるため、ここでは簡単な方法で最大値を推定している。

トラッキングは通常の処理と殆んど同じであるが、曲線に沿うために特徴線の方が徐々に変化することを考慮している点と、同じタイプの特徴点だけを追う点が異なる。

特徴線を画面全体から求めるためには、まず参照マップを作り、それから同じめに求める線の信頼性レベルを決定して、図2の処理を行う。終、十分な信頼性レベルを下げ、再び処理を行う。このようなくり返しを行えば、大きな誤りをすることなく線を得ることができ、また途中で処理を中止しても明らかで特徴線が漏れることはない。

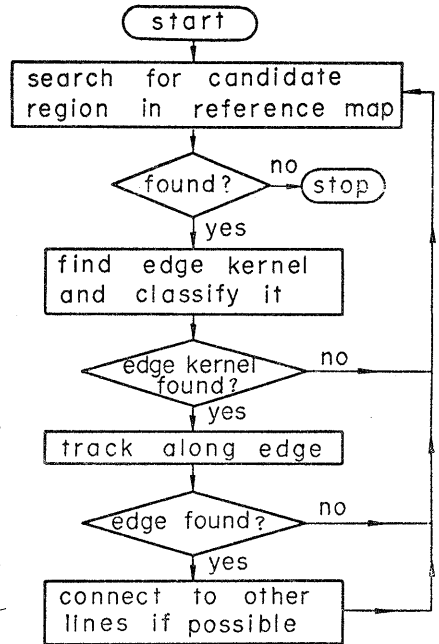
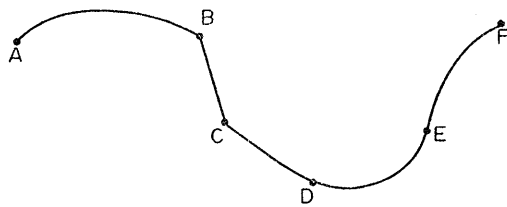


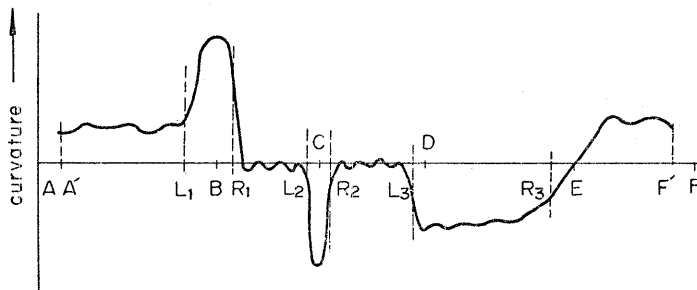
図2. 特徴線抽出の概要

2.2 特徴線の記述

特徴線抽出で得られた各線は、滑らかに連続した特徴点の集合として表わされていゝ。そのまゝでは認識に用ゐることは困難であるので、ここでは特徴線をもい少し意味のある形で記述する。すなわち線を認識に有効ないくつかのパラメータで記述する。そのために曲線を直線と楕円で近似することにした。曲線を直線部分と曲線部分に分割し、曲線部分は楕円で近似できるように分割して方程式に当てはめる。そこで最も重要なことは、適当に曲線を分割することである。分割は線の曲率の変化に基づく。線上の点Pの曲率は、そこ



(a) edge to be segmented



(b) curvature along the edge

図3. 特徴線の分割

から両方向に一定の点だけ離れた点Q, Rを用い、線分QPとQRの正角 θ として定義される(2)。図3に例示するように、(a)の曲線AFの曲率の変化は(b)となる。分割の目的はAFを図のように分け、どの部分も直線か楕円の一部分で近似できるようにすることである。分割の主な手順は、

- (1) 分割点の候補として曲率の大きい部分を探す。
- (2) 曲率の大きい部分が分割点を含むか、あるいは曲線であるかを判定する。
- (3) 分割点を含む部分を除いた区間を直線か曲線かに分類する。
- (4) 曲線部分が一樣な曲率か否かを判定し、必要ならばさらに分割を行う。
- (5) 隣接する曲線部分を結合できないか結合する。

なお各ステップの詳細は(2)に述べられている。

得られた線は Demming の方法(14)で、直線あるいは楕円に当てはめる。

2.3 認識

認識の方式はその目的やシーンに関する拘束条件により異なる。机上シーンの認識で最も困難な問題は、得られた特徴からまず第一に何を推定できるかである。もしある仮定が推定されれば、それを検証することは比較的容易である。たとえば、電話機のダイヤルが推定できれば、果してそれが本当にそうであるのか、あるいは他の丸い物であるかを確かめることはできる。ここでは特徴線を用いて指定された物体を見つける問題を取上げよう。各物体に関して人間がプログラムを書いている。プログラムは、物体で最も特徴的な線を見つけて、それを手がかりにして他の特徴を探して、一定以上の特徴が得られれば物体が求められたという仮定を立てる。仮定の検証は、シーン内にその物体とどの程度類似の物体が存在し得るかに依存する。現在机上の電話機、カップ、電卓スタンド、本を見つけるプログラムがあるが、課題は計算機にプログラムを作らせることである。

2.4 全体の統制制御

このモジュールは、シーンに応じて特徴線抽出の信頼性レベルを設定し、得られた特徴線に応じて別なレベルを設定し直したり、特徴線の記述を行なわせ、また認識結果により再び特徴線抽出に戻ったりすることを制御する簡単なプログラムである。

3. 3次元情報を用いたシーンの記述(15)

曲面を含む物体のシーンの3次元情報を検出し、シーンをわかり易い形で記述する研究である。筆者は当研究室で製作したレンジファインダで距離を入力して多面体を認識した(16)が、ここでは距離情報を直接使って平面上を決定した。Agin は同じような原理で距離を求め、曲面物体を一般円筒で記述した(17)が、はじめに円筒を推定する場合には距離を直接使って、物体のりんかんに基づいた。したがって背景が暗く、物体が細長い部分で構成されている場合以外にはあまり有効な方法ではない。欠長は、はじめから3次元情報を使って平面上を見つけないことにより、20面体のような多面体を認識した(18)。

本研究はこれを発展させ、さらに曲面を含んだシーンも記述できるようにしたものである。平面はその方程式により記述できるが、平面だけから成るシーンを記述することは比較的簡単である。しかし一般に曲面を記述する可能な方法はない。曲面の種類が限られていれば、曲面方程式を用いることが自然であるが一般曲面はこれをいくつかの曲面に分割し、各曲面を平面あるいは一定の曲面で近似することが多い(19)。後者の方法で曲面をいかに正確に近似するかは重点が置かれてくる。シーンを記述し、その記述に基づいて認識を行うためには、多数の曲面のセグメントに関する方程式自体を直接に用いることは困難である。シーンに関する大局的な認識をするには、まずそこにはいかなる面があるかを知る必要がある。すなわち平面や曲面の大きさや位置が最も重要で、さらに詳しく分析する場合には、曲面の形が問題になる。そこで曲面を何らかの方法で記述しなすべからなければならないが、その方法は認識の目的により異なる。

ここではシーンの記述の基礎となる最初の処理である、シーンを平面と滑らかな曲面の領域で記述する方法を述べる。図4にここで行う処理により、2原シーンがどのように変換されていくかを示す。

3.1 面素の作成

図4(1)に示すシーンに対してレンジファインダを用いて距離を求めれば、スリット光が当たり、それがテレビカメラで観測されるような点の3次元位置を(2)のように得ることが出来る。このシーンを分析するのには、3次元位置の得られてくる各点ごとにそれがどのような面に属するかを決定することは可能ではあるがこの決定のためにはその点の近傍にある他の点も考慮しなければならぬ。また扱うべき点の数も多いので能率が悪い。そこで近傍の点を1つにとり、これをユニットとして扱うことにする。ここでは横たて8x8の点をとり、これを面素とよんでいる。全シーンは200x200の3次元の点から構成され、互いに他と横たてた点および点一つ重なるような面素を作れば、全部で40x60の面素が生成される。各面素は平面であるとみなしで最小二乗法でその方程式を求めたおくれ。なお面素は機械的に作られているので、1つの面素は原シーンの1つの面上にあるとは限らず、そのような場合は最小二乗法の分散が大きくなる。分散が大

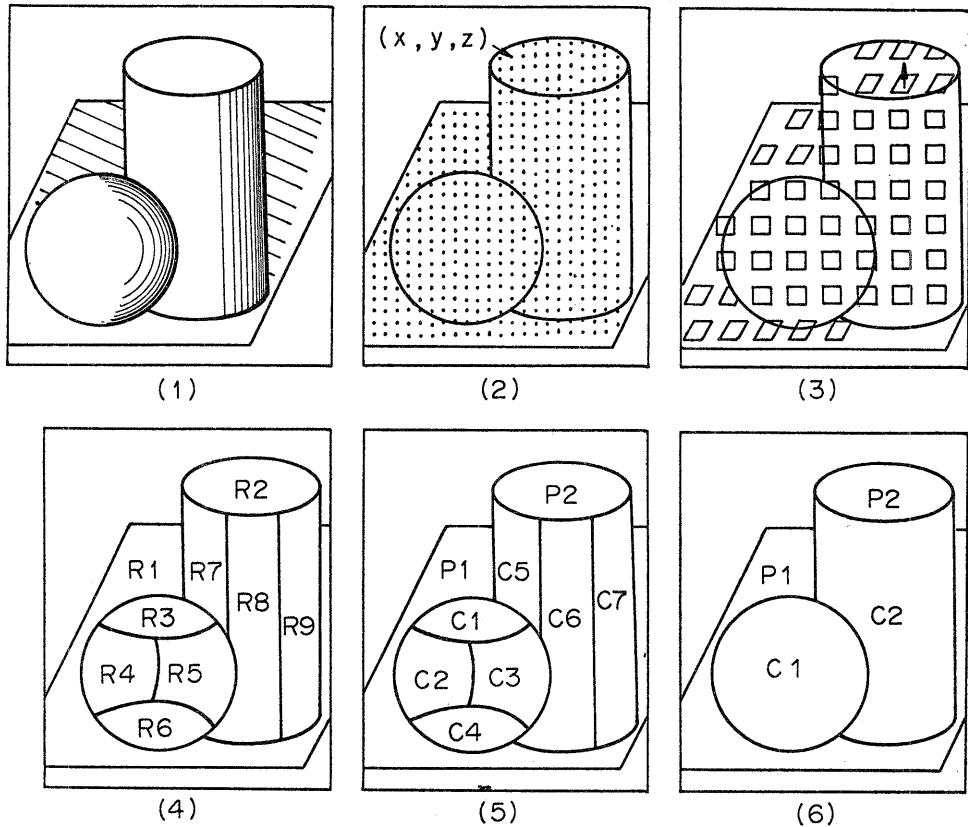


図4. 全体の処理を示す概念図

きい面素は以後の処理の対象から外しておく。図4(3)に面素の概念図を示すが、実際は各面素は互いに重なり合っている。

3.2 基礎領域の作成

ここにはほぼ平面とみなされる領域(基礎領域, Elementary Region)を見つける。平面を求めることは簡単であるので、まず平面部分を確定することは以後の処理の効率を高める。また明らかに不連続な面を分離することも重要である。その方法は、まず平面の一部であると思われる面素(核面素)を探し、つぎにそのまわりに基礎領域を広げていく。このときも面素の傾きと、原点から面素に下した垂線と面素の交点を基準にしている。核面素は、その面素の近傍にある未処理の面素の数とそれらの面素の分散に基づいて選択する。すなわち近傍の面素の数が多く、しかも各分散が小さくなるものが核面素となり、基礎領域となる。つぎにそのまわりの面素の中から、基礎領域の方程式と類似の方程式を持つ面素を結合する。結合後はその領域の方程式を変更しておく。このように基礎領域を拡張できなくなるまでくり返し、基礎領域を決定する。もし核面素の候補が残っているれば他の基礎領域を同様にしして作成する。

このような処理を行えば、平面は一つの基礎領域になり、なめらかな曲面は一つあるいは複数の基礎領域になる(図4(4))。

3.3 基礎領域の分類

ここに得られた基礎領域が、平面の一部であるか曲面の一部であるかを判定する。平面に対応する基礎領域は、それを構成する面素の傾きがほぼ等しい。もし一定以上の大きさの曲面に対応する基礎領域であれば、面素の傾きの分散が大きくなる。また小さな基礎領域であれば、その判定は困難である。基礎領域の大きさに対応する量として、ここでは有効径を用いている。有効径は領域の8方向の最大径を求め、その8個の最小のものとして定めている。図5に平面と曲面の判定基準を示す。有効径の小さいものや、平面と曲面の間の中間的な領域は不明とする。このように基礎領域は、平面、曲面、不明の3種類に分類される。

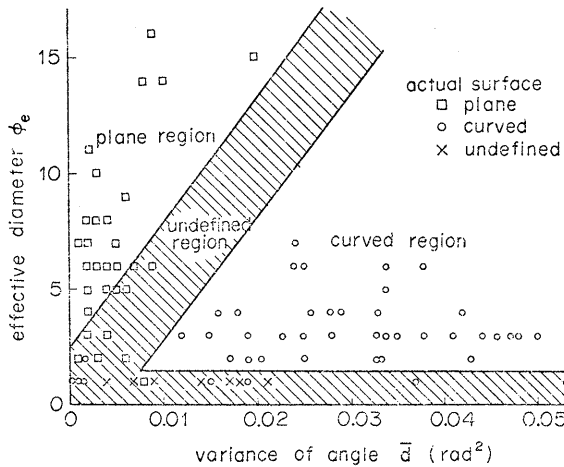


図5. 基礎領域の分類

3.4 連続領域の作成

シーンの記述は、平面と滑らかな曲面の領域であるが、ここでは各領域を連続領域 (Continuous Region) と呼ぶ。前節で分類された平面の基礎領域はその基本平面の連続領域となる。他の基礎領域は曲面の一部となる可能性がある。そこで連続領域作成では、滑らかな曲面を構成する基礎領域を求める。処理手順は、基本曲面の一部とみなされた基礎領域を求め (核領域)、つぎにこれと同じ曲面にある他の基礎領域を求める。

核領域は曲面に分類された基礎領域の中から選ぶ。選択はなるべく大きな領域を優先する。大きさの基準としては前述の有効径および領域を構成する面素の散を用いている。核領域が求められると、そこに隣接する未決定の基礎領域について、これが同じ連続領域となるかを判定する。判定はその基礎領域に隣接する連続領域内の基礎領域と滑らかに連結していかに行う。そのためには二つの基礎領域の連続性が問題である。ここでは各領域の部分領域を平面で近似したときの互いの類似により、連続性を決めよう。

3.5 今後の問題

ここで述べた方法は、球、円筒などのような曲面を含むシーンに対して有効であった。この結果を用いて曲面をある程度分類することもできると考えられる。しかし本章で扱ったような実世界のシーンに対しては問題がある。すなわち同じ分解能で面素を作れば、微細な変化を記述できないことである。またもし入力装置の改善が可能で、分解能を高くできたとしても処理時間が問題となる。すなわちここで述べたような方法は、小さな領域を結合してしばしば大きな領域を作るいわゆる領域法である。領域法はシーン全体を分析するのに適しているが、時間がかかるという欠点がある。そのためには本章で述べたアプローチも必要となる。

参考文献

- 1) 白井良明: 電総研における物体認識の研究, イメージプロセッシング研究会資料2, 情報処理学会(1975)。
- 2) Y. Shirai: Edge Finding, Segmentation of Edges and Recognition of Complex Objects, 4th IJCAI (1975)。
- 3) P. Winston: The MIT robot, Machine Intelligence Vol. 7 (1972)
- 4) Y. Shirai: A Context Sensitive Line Finder for Recognition of Polyhedra, Artificial Intelligence, Vol. 3, No. 2 (1972)。
- 5) G. Falk: Interpretation of Imperfect Line Data as a Three-Dimensional Scene, Artificial Intelligence, Vol. 3, No. 2 (1972)。
- 6) W. Perkins and T. Binford: A Corner Finder for Visual Feedback, Computer Graphics And Image Processing, Vol. 2, No. 3/4 (1973)。
- 7) G. Grape: Model Based (Intermediate Level) Computer Vision, Artificial Intelligence Memo No. 201, Stanford University (1970)。
- 8) T. Garvey and J. Tenenbaum: On the Automatic Generation of Programs for Locating Office Scenes, Proc. 2nd IJCP (1974)。
- 9) S. Tsuji and A. Nakamura: Recognition of an Object in a Stack of Industrial Parts, 4th IJCAI (1975)。
- 10) M. Yachida and S. Tsuji: A Machine Vision for Complex Industrial Parts with Learning Capability, ibid.
- 11) J. Prewitt: Object Enhancement and Extraction, Picture Processing and Psychopictorics, Academic Press, New York (1970)。
- 12) A. Rosenfeld, M. Thurston and Y. Lee: Edge and Curve Detection, IEEE Trans. on Computers, Vol. C-20, No. 5 (1971)。
- 13) Y. Shirai: A Step toward Context Sensitive Recognition of Irregular Objects, Computer Graphics and Image Processing, Vol. 2, No. 3/4 (1973)。
- 14) 木間仁, 春日屋伸高: 次元解析・最小二乗法と実験式, 応用数学講座, 第5巻, コロソ社(昭52)。
- 15) M. Oshima and Y. Shirai: Representation of Curved Objects Using Three-Dimensional Information, 2nd USA-JAPAN Computer Conference (1975)。
- 16) Y. Shirai: Recognition of Polyhedrons with a Range Finder, Pattern Recognition, Vol. 4 (1972)。
- 17) G. Agin: Representation and Description of Curved Objects, Artificial Intelligence Memo No. 137, Stanford University (1972)。
- 18) 久良, 岩本, 白井: 3次元データによる物体認識へのアプローチ, 第4回画像工学コンファレンス論文集(昭58)。
- 19) A. R. Forrest: On Coons and Other Methods for the Representation of Curved Surfaces, Computer Graphics And Image Processing, Vol. 1, No. 4 (1972)。