



濃淡画像から複雑物体を認識する一手法*

白井良明**

Abstract

This paper describes an approach to the recognition of real-world objects such as books or a telephone on a desk. The system consists of (1) edge finding process which extracts edges of curved objects from light intensity data, (2) segmentation of the edges into straight lines or elliptic curves, (3) recognition of objects by matching lines to the models, and (4) simple supervisor. The module (1), (2) and (3) interact with each other through a simple supervisor so that the system can locate given objects quickly. First, most reliable edges are found and segmented, and then recognition is attempted using segmented lines. If recognition is not successful, less reliable edges are searched for, and recognition is retried. An example of locating several objects on a desk is shown.

1. まえがき

視覚情報を用いて三次元物体を認識するシーンアナリシスの研究では、長い間単純な積木を対象としていたが、最近では複雑なシーンを扱うようになった。このシーンアナリシスは、(1)原データからの特徴抽出、(2)得られた特徴を用いたシーンの解釈、の2つの部分で構成されると考えることができる。しかし信頼性のある認識を行うためには、この2つの部分を単純に直列に組み合わせるだけでは十分ではなく、互いに干渉しながら認識を進める必要があることが指摘されている¹⁾。

多面体認識のために筆者は稜線検出部と稜線予測部の2つを作り、互いに情報を利用して仕事を進めるシステムを作った²⁾。ここでは物体と背景の明るさの差が大きいことを仮定し、物体のりんかくから得られる情報を利用して、ほかにも一定の制限のもとで多面体のモデルを用いて認識を行う試みがあった^{3),4)}。しかし曲面を含む物体の認識にそのまま適用することは困難である。すなわちどのような特徴を抽出すべきか、また特徴をいかに解釈するかが問題である。

いっぽう距離情報を用いて曲面を含むシーンを解釈

する研究も行われている⁵⁾⁻⁷⁾。さらに Garvey 等⁸⁾は、距離と色の情報も得られるものとして、屋内シーンの中から特定の物体を見つける試みを行った。そこでは三次元座標が、机の上面、いすの上面、壁などを決定するのに重要な役割を果たしている。また Yakimovsky 等⁹⁾は色情報を用いて屋外景色を領域法によって解釈した。シーンの領域分けを行う場合に、空、山、道路、自動車などの位置、大きさ、隣接関係などの事前確率を用いたいわゆる意味論を利用したアプローチである。

ところがここで扱う対象はより一般的シーンで、例えば机の上のスタンドの電話機などを含み、しかも濃淡情報だけを用いている。濃淡情報は入力が容易で、データ量も少ないため、未知シーンを最初に分析してある程度の解釈を行うのに適していると思われる。

複雑なシーンの濃淡情報から特徴を抽出する研究も多く行われているが^{10),11)}、得られる特徴を用いて認識を行った例は少ない。もし対象シーンが限られていれば特徴抽出や認識方法も効率の高い方式が考えられる。例えば正面を向いた顔の認識であれば、顔の一部がわかれば、幾何学的制限によって他の部分の範囲を定めることもでき、特徴抽出も予測された部分だけに適した方法を探ることができる¹²⁾。

ここでは扱うシーンは、物体の配置に関する制限が少ないため、モデルを用いてトップダウン的に処理を

* A Method for Recognition of Complex Objects from Light Intensity Data by Yoshiaki SHIRAI (Bionics Section, Information Sciences Division, Electrotechnical Laboratory)

** 電子技術総合研究所パターン情報部バイオニクス研究室

進めることは困難である。したがってはじめでできるだけ信頼性のある特徴を抽出して、それを解釈してつぎの処理に利用するという過程をくり返す必要がある。

曲面を含むシーンの一般的特徴として、筆者は局所の特徴と大局の特徴を提案した¹³⁾。すなわち前者は一点を中心とした一定領域内の明るさのパターンを示し、後者は同じパターンの局所の特徴を連結したものである。もし物体表面の明るさが一様であれば、大局の特徴は物体表面の境界に対応する。本論文はこの研究の延長で、認識システム全体を述べる。システムは、(1)境界線の抽出、(2)境界線の直線と楕円の一部による記述、(3)記述された境界線を用いた認識、(4)全体の統括、の4つの機能を持つ部分で構成される。(1)と(2)の部分は一般性があり、信頼度、最小曲率、分解能などのパラメータを与えることによって望みの機能を発揮させることができる。

本手法の特徴をまとめれば、1)局所オペレータによって面の境界点を求める場合、境界点近傍の濃淡パターンを3種類考慮しているため、平面と曲面の混在するシーンに適している。2)面の境界は同じ種類の特徴点を連結して求めているため、影やノイズの影響を受けにくい。3)境界線抽出を効率よくするため参照マップを用いている。4)境界線の記述は単に境界線を式で表すだけでなく、認識に使い易い方式であり、しかも記述の手続きは試行錯誤を含まず高速処理が可能である。5)認識用モデルは階層的特徴で表わされ照合の手順が与えられるので効率が良い。6)全体の処理は直列的でなく、得られた情報をフィードバックして有効に利用している。

2. 境界線の抽出

ここで境界線とは、曲面を含む物体のりんかくあるいは稜を意味する。境界線はそれを構成する境界点を検出し、境界点を滑らかに連結することによって求められる。

2.1 境界点の検出

境界点として考えなければならない明るさの変化をFig. 1に示す。境界線に直角方向に明るさを調べ、図のB, R, Lの3種類の濃淡波形のいずれかであれば境界点とみなす。あらかじめ境界線の方向(右側が明るいように定める)が予測できる場合には、境界点検出にはつぎのような一次元のオペレータを用いる。いま境界線と直角方向 u に沿って明るさが $I(u)$ で

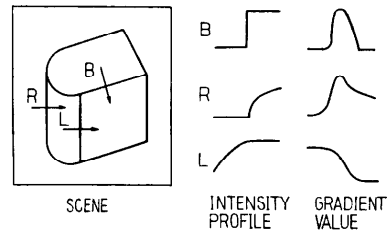


Fig. 1 Intensity profile of edges

表わされるとすれば、明るさの微分は、

$$D(u) = \sum_{i=1}^a I(u+i) - \sum_{i=1}^a I(u-i) \quad (1)$$

ここで a は平均点数、なお平均は境界線の方に行っても行うことがある。点 u_0 が境界点であるためにはつぎの3式を満足しなければならない。

$$D(u_0) \geq D_i \quad (2)$$

$$\exists i (D(u_0+i) < f(D(u_0)) \wedge -w \leq i \leq w) \quad (3)$$

$$\forall i (2D(u_0) > D(u_0-i) + D(u_0+i) \wedge -w \leq i \leq w) \quad (4)$$

ここで D_i は閾値、 w は通常は3以下の自然数のいずれか、 $f(x) = C_0 + C_1 x$ ($C_1 > 0$) である。もし式(3)において正と負の i が存在すればFig. 1のタイプBとみなす。もし $i > 0$ のみであればタイプL、 $i < 0$ のみであればタイプRとなる。平均回数 a とオペレータの領域の大きさを表わす w は、データの分解能によって定める。いずれも小さくすれば分解能は増すがノイズの影響を受け易くなる。また点 u_0 が境界点である信頼度の測度としては $D(u_0)$ を用いる。

もし境界線の方向が未知であれば二次元のオペレータ¹⁴⁾を用いて微分を行って境界線の方向を求める。オペレータの領域は、 3×3 から 5×5 までを使い分ける。境界線の方向が求められれば、境界点の検出は一次元の場合とほぼ同様である。

2.2 境界線の作成

Fig. 2(次頁参照)に境界線作成の簡単な流れを示す。主な処理は、(1)与えられた信頼度以上の境界の核を検出し、(2)その核から出発して境界線を追跡する。(1)ではまず二次元オペレータを用いて境界点を検出する。すなわち信頼度として式(2)における D_i を与え、式(2)~(4)を満足する点を求める。境界点を得られたら、その方向に長方形の小領域(3×5)を設定し、その中にある境界点を調べる。もし同じタイプの境界点で、その方向もほぼ等しいものが一定以上あれば、そこに境界線の核があるとみなす。

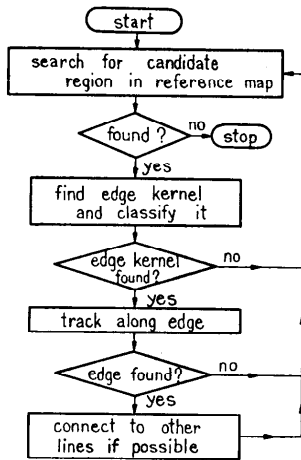


Fig. 2 Flow chart of edges finding

Fig.2 において参照マップは核検出の速度を増すために用いる。マップを作るためには、まず原画像を圧縮した画像 $J(i, j)$ を作り (たとえば縦横 1/2), つぎに (i, j) 近傍の微分値の最大値の目安として

$$D'(i, j) = \max_{-1 \leq \Delta i \leq 1, -1 \leq \Delta j \leq 1} |I(i + \Delta i, j + \Delta j) - I(i, j)|$$

を求める。 $D'(i, j)$ がその対応する原画像の近傍の微分値の最大値を近似するものとみなす。この $D'(i, j)$ を原画像の数点ごとに (4点ごと) 求めておく。核を探す場合には参照マップを見て、 D' が一定以上の点の近傍だけについて境界点の検出を行う。探索が終れば、その対応する参照マップも正確な値に書換えておく、以後の参照のために備えておく。

検出された境界線の核は (2) によって両方向に延長される。追跡の目的は、核と同じタイプの境界点からなる滑らかな線を得ることである。ここでは曲線も追跡する必要があるので、直線の追跡と異なる方式を用いる。追跡には、最後に求められた点を含み、その点と連結する数点の重心に対応する境界線の位置 X_c 、その方向 β_c 、平均的な境界点の微分値 D_c を用いる。以上のパラメータの初期値は、境界線の核の重心、方向および核を構成する境界点の微分値の平均によって与えられる。追跡が進むにしたがってパラメータを更新していく。追跡の主なステップは、

1. X_c と β_c を用いてつぎに求める境界点の位置を予測する。
2. 予測された点の近傍にある境界点を求める。
3. X_c, β_c, D_c を更新する。

以上の手順をくり返し、境界点が検出できなければ

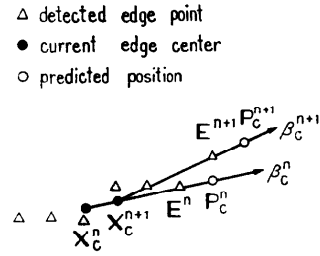


Fig. 3 Tracking parameters

追跡を終了する。

ステップ1で、Fig. 3 のように現在の境界線の位置を X_c^n 、方向を β_c^n 、境界点の予測位置を P_c^n とし、新しく P_c^n の近くの境界点 E^{n+1} が検出されたものとすれば、新しい境界線の位置 X_c^{n+1} は X_c^n から β_c^n 方向に一絵素だけ進んだ点に移り、方向 β_c^{n+1} は X_c^{n+1} と E^{n+1} を結ぶ方向となる。またつぎの境界点の予測位置は、 E^{n+1} から β_c^{n+1} 方向に一絵素進んだ点となる。なお X_c^n と P_c^n までの距離は求めるべき境界線の最大曲率に依存する。ステップ2では前述の一次元オペレータを用いて境界点を検出する。ここでは微分方向 α^n は β_c^n と直角であるとみなし、境界点の予測位置 P_c^n を中心として α^n 方向の近傍の点を調べる。もし望みのタイプの境界点が求まり、その微分値が D_c^n に近ければ、その点を境界線上にあるとみなして追跡を続ける。もしそのような点が検出されなければ以下に示すバックアップを行う。

- a) もしステップ2が P_c^n における最初の境界点検出処理であれば、微分方向を少し変えてステップ2へ行く。
- b) もし直前の境界点が正しく求められていれば、 P_c^n に境界点があるものとみなしてステップ3へ行く。
- c) その他の場合は追跡を終了する。

以上のバックアップは、物体表面の不均一性、ノイズおよび方向 α_c の誤差の影響を避けるためである。

両方向の追跡が終了すれば、得られた点の数によって境界線かノイズかを定める。また両端が他の境界線の近傍にあれば、境界線のタイプ、方向、信頼度に基づいて連結できるかを調べ、もし可能ならば連結する。通常は以上述べた Fig. 2 の手順をくり返して行う。すなわちはじめは信頼度を高く設定して全画面に対して境界線抽出を行い、順次信頼度を下げながら抽出をくり返す。このようにして明らかな境界から順に得ることができ、その結果を用いて何らかの決定を下

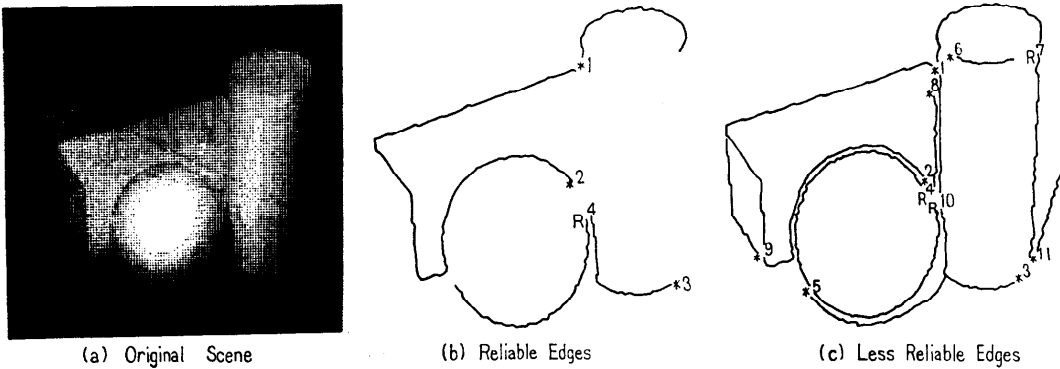


Fig. 4 Example of edges

すこともできる。また他の方法ではノイズによる不要な線が抽出される場合にも、ここではすでに得られた線に交わるためにそれを消去できることが多い。Fig. 4 に境界線抽出の例を示す。

3. 境界線の記述

境界線抽出で得られた各境界線は、滑らかに連結された点の系列である。ここでは境界線を認識するのに適したいくつかのパラメータで記述する。

曲線を記述する研究としては、多角形近似¹⁴⁾、区間に分けて多項式近似を行う方法¹⁵⁾ (スプライン関数近似) などがある。それは誤差の測度を定めておき、最適解を求めようとするものである。しかしこの種の近似は必ずしも認識には適していない。

通常の物体の境界線は、直線と楕円の一部の組み合わせとして記述でき、しかもそれが認識プログラムにおいて直観的に利用できると思われる。1つの境界線を記述するためには、(1)点系列を直線部分と曲線部

分に分割し、(2)各部分を式で近似する。

(1)の目的を例示すれば Fig. 5(a) のように境界線 AF を5つの部分 (AB, BC, ..., EF) に分割することである。分割は境界線の曲率²⁾に基づいて行う (境界線上の点 P における曲率は、P から一定の点 m_a だけ離れた点 Q と R を用いて、直線 QP と PR のなす角度と定義する)。したがって AF の曲率は Fig. 5(b) のように、 m_a だけ内側の区間 A'F' で定義される。分割の主な手順は、

- A. 境界線の節点 (分割の境界) の候補を求める。
- B. 仮に分割された部分を直線、曲線、不明の3種類に分類する。曲線部分はさらに曲線と直線に分割できれば分割する。
- C. 隣り合う部分で連結可能であれば連結する。
- D. 節点の位置を決定する。

まず節点を求めるため、曲率の大きい場所を探す。節点のある場所の候補を Fig. 5(b) に例示する (区間 $[L_i, R_i]$ で表わされる)。この区間は曲率 $\delta(P)$ によってつぎのように定義される。区間 $[L_i, R_i]$ の曲率の絶対値の最大が δ_{max} とすれば、

$$\delta_{max} \geq \delta_i$$

$$|\delta(P)| \begin{cases} \geq h(\delta_{max}) & \text{for } P \in [L_i, R_i] \\ < h(\delta_{max}) & \text{for } P = L_i - 1, P = R_i + 1 \end{cases}$$

ここで δ_i は閾値、 $h(x) (< x)$ は x の線形関数。もしこの区間が短かければ節点区間とし、長ければ曲線部分とし、その両端に節点を挿入する。A によって仮の分割が行われ、各部分は節点間で区切られる (Fig. 5(b) では $[A'L_1], [R_1, L_2], [R_2, L_3], [L_3, R_3], [R_3, F']$)。

B では A で求められた曲線部分を除いた各部分を分類する。まず一定長以下の部分は不明とみなす。残りの分類は、曲率を定義するパラメータ m_a 、曲率の和

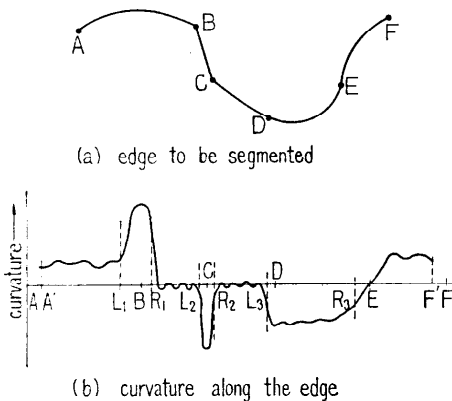


Fig. 5 Initial segmentation of edge

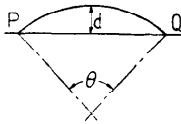


Fig. 6 Arc approximation

S_1 および部分を構成する点の数 N を用いて行う。もし Fig. 6 のように部分 PQ が円弧であると仮定すれば、直線 PQ と円弧との最大距離 d および円弧の中心角 θ は次式で近似される。

$$d = (N/\theta)(1 - \cos(\theta/2))$$

$$\theta = S_1(N-1)/m_d(N-2m_d)$$

分類はこの仮定に基づき、 $d > d_c$ ならば曲線、 $\theta < \theta_c$ ならば直線、 $N < m_d/2$ 、 $d < d_c$ (d_c と d_c は定数で $d_c > d_c$) ならば直線、それ以外は曲線とする。曲線部分はその中に一定長以上の直線部分を含むかを調べて、もしあればさらに分割を行う。

C では隣り合う曲線あるいは不明の部分に関して、その間の節区間の曲率と2つの部分の θ を調べ、もし連結しても楕円の一部とみなされる場合には連結する。以上の手順を適用できなくなるまでくり返し、最後にDで2つの部分の境界を求めて節の位置を決定する。なお連結されずに残った不明の部分は直線とみなす。

境界線記述の後半は、各部分を構成する点 (X_i, Y_i) ($i=1, 2, \dots, n$) を直線あるいは楕円の一部に当てはめる。デミングの方法¹⁶⁾を用いて各点からの距離の2乗和を最小とする式を求める。線の式が m 個のパラメータを持ち、 $f(x, y; a_1, a_2, \dots, a_m) = 0$ で表わされるとすれば、この方法では a_i の近似値から出発し、くり返し法で誤差が一定以下になるまでパラメータを計算する。 a_i の近似値は適当な数の点 (X_i, Y_i) を用いて決定する。曲線部分は楕円の式で当てはめを試みる。パラメータは中心、長軸、短軸の長さ、方向である。もしくり返しが収束しなければ、パラメータの数を減らして再試行する。Fig. 7 に境界線記述の例を示す。

4. 認識とシステム全体の統括

物体認識の方法は対象物やシーンの拘束条件に依存する。ここでは机の上に自由に置かれた日常の物体を認識することを目的としている。この場合にも、シーン全体を分析してどこに何があるかを認識するか、あ

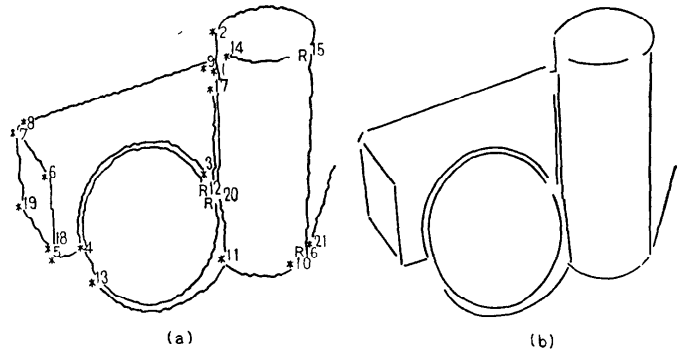


Fig. 7 Segmented edges (left) and fitted line (right)

るいは与えられた物がどこにあるかを知るかによって方式が異なる。前者の場合には、明らかに認識できる部分から順に認識を行い、その結果を利用しながら未知な部分の認識を行う。後者の場合には、目的とする物体の主な特徴を探し、それが真にその物体であるかを検証すればよい。もし主な特徴が見つからなければ、他の明らかな物体を認識して、それを手がかりに再び目的とする物体を見つける。

いずれの場合もシーンを記述する特徴と、物体のモデルの照合が必要である。1つのアプローチとしてモデルをグラフで記述してシーンと照合する方法がある¹⁴⁾。モデルは二次元で、適当な方向から見た物体の線画を直線と楕円の一部で記述し、その相対関係をグラフにする。シーンは一般に複数物体が重なって置かれているため、サブグラフの照合が必要になる¹⁷⁾。この方法で円柱、球、直方体などの簡単な物体の認識は行われているが、物体が複雑になると、モデルに導入する特徴の性質や関係の種類の問題、照合手順の問題などが未解決である。

ここでは手続きによって物体を認識している。すなわちモデルとの照合をプログラムで指定し、能率よく認識を行っている。入力シーンは 256×256 画素 (1画素6ビット) であるため、細い部分を認識するためには十分な分解能ではない。そこで比較的大きな物体の認識の第一段階を示す。

与えられた物体を認識するためには、Table 1 (次頁参照) に例示するような主特徴を探す。もし見つければ主特徴を手がかりとして副特徴を求める。これで十分認識できたと判断できない場合には、さらに他の情報を用いたり、その部分をズームアップして再入力して調べることになる。全シーン解析にも Table 1 のような特徴の階層化は必要である。不十分な分解能の

Table 1 Example of main feature and secondary feature

物体名	主特徴	副特徴
蛍光灯スタンド	長方形状笠の長辺	笠の短辺と笠の支柱
本立	長い垂直線群	長い垂直線群間の短い垂直線
茶わん	左右の垂直なりんかく	左右りんかくに連なる上下面りんかく
電話機	楕円状ダイヤルの外形	ダイヤルを囲むケースのりんかく

シーンから何らかの有用な情報を得て、つぎに行う処理の焦点を定めることは、複雑なシーンの解析に有効である。

Fig. 8 に全システムの構成を示す。与えられた物体を探す場合を例にとってその機能を説明する。まず原画像から、前述の参照マップを作成する。境界線抽出部は、統括部から与えられた信頼度、境界点の上界などに基づいて明らかな順に境界線を求める。つぎに境界線の分割、記述を行い、得られた記述を用いて認識が行われる。もし認識が全部成功しなければ、再び信頼度を少し下げて境界線抽出を続ける。全部認識ができるまでこの過程をくり返す。Fig. 9 に例を示す。処理時間は入力画像に依存するが、5～9分位である。その中で境界線抽出が大部分を占める。したがって

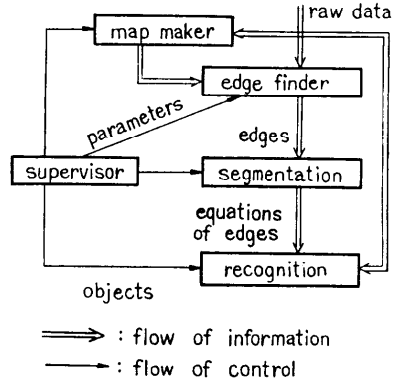
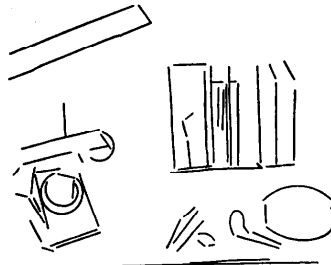


Fig. 8 Total System

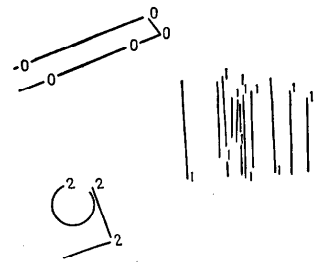
Fig. 8 の方式の利点は、1) 不要な信頼度の低い境界線抽出を避けることができ、2) 途中で認識結果を利用して効率を向上することができる (たとえば既に認識された本立の領域の境界線抽出をする必要はない)。Fig. 10 (次頁参照) にはできるだけ多くの特徴線を求めて、一回で認識した例を示す。同じような結果であるが、不要な線を求めているだけ時間が多くかかっている。



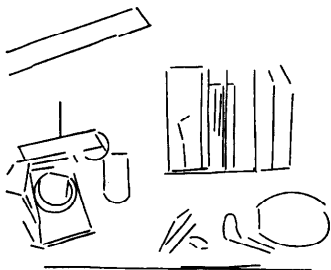
(a) Original scene



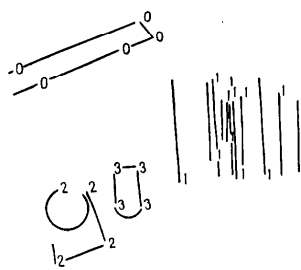
(b) Fitted lines



(c) Lines used for recognition



(d) Fitted lines



(e) Lines used for recognition

(b), (c) : first pass
(d), (e) : second pass

0 : lamp
1 : book stand
2 : telephone
3 : cup

Fig. 9 Example of Locating object

6. むすび

複雑な物体が複数個重なって見える濃淡画像の処理の一手法を示した。曲面を含む物体の面の境界線を抽出し、それを直線と楕円の一部の組み合わせとして記述する。この記述を用いて物体の主特徴を探し、主特徴を手がかりとして副特徴を求める。この方法で机の上に置かれた日常物体の認識を行うことができる。

今後の課題は、ここで行った第一近似としての認識をさらに進めること、および人間が物体ごとに作った認識手続きを効率よく作るための手法の開発である。

参考文献

- 1) P.H. Winston: The MIT Robot, Machine Intelligence, Vol. 7, p. 431 (1972).
- 2) Y. Shirai: A Context Sensitive Line Finder for Recognition of Polyhedra, Artificial Intelligence, Vol. 4, No. 2, pp. 95~119 (1973).
- 3) G. Falk: Interpretation of Imperfect Line Data as a Three-Dimensional Scene, Artificial Intelligence, Vol. 3, No. 2, pp. 101~144 (1972).
- 4) G.R. Grape: Model Based (Intermediate-Level) Computer Vision, Artificial Intelligence Project Memo, No. 201, Stanford University (1973).
- 5) G.J. Agin and T.O. Binford: Computer Description of Curved Objects, Proc. 3rd Int. Joint Conf. Artificial Intelligence, p. 629 (1973).
- 6) M. Oshima and Y. Shirai: Representation of Curved Objects Using Three-Dimensional Information, 2nd USA-JAPAN Computer Conf., p. 108 (1975).
- 7) R.J. Popplestone et al.: Forming Models of Plane-and-Cylinder Faceted Bodies from Light Stripes, Proc. 4th Int. Joint Conf. Artificial Intelligence, p. 664 (1975).
- 8) T.D. Garvey and J.M. Tenenbaum: On the Automatic Generation of Programs for Locating Objects in Office Scenes, Proc. 2nd Int. Joint Conf. Pattern Recognition, p. 162 (1974).

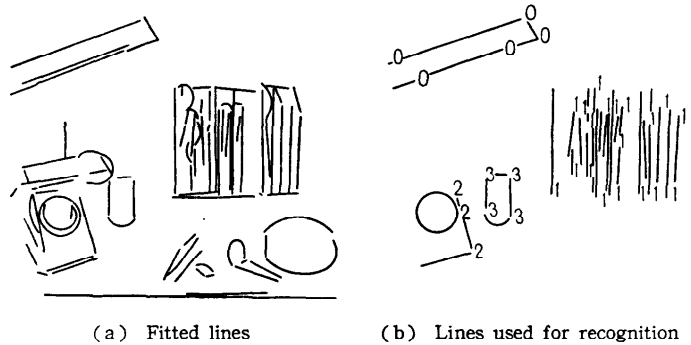


Fig. 10 Example of one pass recognition

- 9) Y. Yakimovsky: On the Recognition of Complex Structures: Computer Software Using Artificial Intelligence Applied to Pattern Recognition, *ibid.* p. 345.
- 10) J.M.S. Prewitt: Object Enhancement and Extraction, Picture Processing and Psychopictorics, p. 75, Academic Press, New York (1970).
- 11) A. Rosenfeld, M. Thurston and Y. Lee: Edge and Curve Detection: Further Experiments, IEEE Trans., Vol. C-21, No. 5, p. 562 (1971).
- 12) 坂井, 長尾, 金出: 計算機による顔写真の解析, 電子通信学会論文誌, Vol. 56, No. 4, p. 226 (1973).
- 13) Y. Shirai: A Step toward Context Sensitive Recognition of Irregular Objects, Computer Graphics and Image Processing, Vol. 2, No. 314, pp. 298~307 (1973).
- 14) T. Pavlidis and S.L. Horowitz: Segmentation of Plane Curves, IEEE Trans., Vol. C-23, No. 8, pp. 860~870 (1974).
- 15) D.D. Braess: Chebyshev Approximation by Spline Functions with Free knots, Numer. Math., Vol. 17, pp. 357~366 (1971).
- 16) 本間, 春日屋: 次元解析・最小2乗法と実験式, 応用数学講座第5巻, p. 255, コロナ社, 東京 (1973).
- 17) 田村浩一郎: ふたつのグラフの共通サブグラフ, 電子技術総合研究所集報, Vol. 37, No. 9 (1973).

(昭和51年2月19日受付)

(昭和51年3月24日再受付)