

# 頂点辞書を利用した距離画像解析†

杉原厚吉††

3次元シーンに関する一般的知識を頂点辞書の形にまとめ、これを利用しながらシーンまでの距離のデータを効率よく解析する方法を提案する。ここで扱うシーンは、平面または2次元曲面で囲まれた物体がいくつか積重なってできたものである。また、距離データは、スリット光をこのシーンに投影し、その像を別の方向からテレビカメラで捕え、それに三角測量を適用することによって得られる。本解析法では、まずはっきりわかる稜線と頂点を抽出し、各時点で見つかっている頂点を辞書と比較して、どの方向にどのような物理的性質の稜線がなければならないか（あるいはありそうか）を予測し、その予測に従ってデータ解析を進めていく。また、辞書には、2個以上の物体が接触してできる頂点に関する知識も含まれており、それを利用してシーンを単純物体に分割する。そして、最終的には、シーン中の各物体を、頂点・稜線・面の関係として記述する。この結果は、物体の同定や、物をつかむための機械の手の動きの計算等に利用できる。

頂点の形に関する知識は、今までは完全な線画を解釈するという目的に使われていた。しかし、雑音を含む実際のデータの処理で最もむずかしいのは、完全な線画を作ることである。本論文は、距離情報の特質を生かして、今までとは異なる目的（すなわち欠けた線の予測）に頂点知識を利用し、実際のシーンから得られる生データを解析し完全な線画を作る方法を開発した、という点に最大の意義がある。

## 1. まえがき

3次元物体をテレビカメラでとらえ、得られる画像から物体の形状や姿勢を自動解析することは、工業用ロボットに高度の視覚処理機能を付加するための重要な課題である。その際、単に外界の明るさを濃淡画像として得るよりは、三角測量等の方法で物体までの距離を測定し距離画像を得た方が、形状の識別には都合が良い。そのための距離測定装置とそれで得られた距離情報の解析手法とが、今までにいくつか研究されてきた<sup>1)~3)</sup>。その結果、単純な処理で解析できることはほぼ解明され、処理効率を高めるために対象物体に関する知識をいかに記述し解析に利用するかが、現在の大きな課題の一つとなっている。

一方、3次元物体を描いた線画に現われる頂点の形は、物体の3次元形状に関して多くの情報を持っていることが知られており、線画の意味解釈等に利用されてきた<sup>4)</sup>。しかし、そこでは濃淡画像からあらかじめ完全な線画が得られたものと仮定されており、実際の画像から完全な線画を得ることは困難なので、理論的にはおもしろいものであっても実用性はなかった。

本稿では、3次元物体に関する一般知識を頂点辞書の形にまとめ、この辞書を利用して効率良く距離画像から線画を抽出する方法を提案する。これは、距離画像

解析の観点からは、対象世界に関する一般的で高度な知識を利用して処理の効率と信頼性を高めたという意義をもち、一方、線画の頂点形状から3次元物体に関する情報を抽出するという観点からは、今まで線画の意味解釈のために利用されていた性質を、雑音を含む実際のデータから線画を検出するために利用する方法を開発したという意義をもつ。

まず、距離画像から稜線を抽出する局所演算を構成する(2章)、次に頂点辞書を作成し(3,4章)、それを利用した画像解析法を提案する(5,6章)。最後に実験結果を報告する(7,8章)。

## 2. 距離画像とその線画化

三角測量の原理に基づいた距離測定の一方法を図1(a)に示す。光源 $O$ から対象物体に向けて細い光線を発し、照らされた物体表面の点 $P$ をカメラ $O'$ で観測して、カメラ撮像面 $I$ 上での $P$ の像の横座標 $d(P)$ を得る。光源と対象との間に平面 $S$ を固定すると、光線の方向は光線と $S$ の交点 $p(i, j)$ で決まる。従って、 $d(P)$ を光線方向の関数とみなして $d(i, j)$ と書くことができ、光線方向を変えて得られる観測値集合 $\{d(i, j); i=1, \dots, m, j=1, \dots, n\}$ を距離画像と呼ぶ。実用上は $O'$ としてテレビカメラを使用し、水平走査線を直線 $OO'$ と平行に設置すると、 $d(i, j)$ は水平走査の開始からテレビ映像信号が最大値をとるまでの時間として測定でき、また光線は垂直なスリット光でおきかえられて同時に多数の点を観測できる<sup>5)</sup>。

† Range Data Analysis Guided by a Junction Dictionary by KOKICHI SUGIHARA (Information Sciences Division, Electro-technical 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

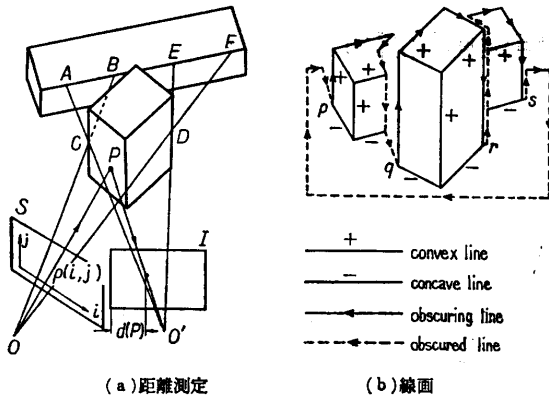


図 1 距離測定と線画

Fig. 1 Range observation and a range picture.

$d(P)$  が測定できるのは、物体表面上の点  $P$  が  $O$  と  $O'$  の両方から見えるときだけである。従って、たとえば図 1 (a) の区間  $AB$  は  $O$  からは見えるが  $O'$  からは見えないため、距離画像の対応する領域は値が定まらない。そこで、距離画像から線画を作るとき、図 1 (b) のように線を 4 種類に分類すると便利である。すなわち、凸稜線 (実線に + 印; 両側の面が凸に交わる)、凹稜線 (実線に - 印; 両側の面が凹に交わる)、輪郭線 (実線に 矢印; 矢の右側が物体で左側は背景)、影線 (点線に 矢印; 矢の右側は物体だが左側は手前にある他の物体の影に隠れて不明) である。また、画面の端も影線で表わす。

距離画像から稜線を検出するためには、次のような 2 階差分が利用できる:

$$r = [d(i-l, j) + d(i+l, j) - 2d(i, j)] / l$$

$r$  の値は、3 点  $p(i, j)$ ,  $p(i-l, j)$ ,  $p(i+l, j)$  に対応する物体表面上の点  $P_0, P_1, P_2$  が同一平面にあるときは零、 $P_0$  を通って  $P_1P_2$  にほぼ直角な凸稜があるときは負、凹稜があるときは正となる。また、 $r$  の絶対値が大きいほど確かに稜線であるという信頼性が高い。従って、奥行きに跳躍のない領域で上記の局所演算あるいは差分の方向を変えたものをほどこせば、稜線の種類と方向を同時に検出できる。そして、稜線の一部がみつければ、その稜線の方向に沿って両側へ線を追跡し、稜線全体を検出することができる。

### 3. 可能頂点

シーンに関して次のような仮定をおく: (1) 対象物体は多面体でその各頂点ではちょうど 3 個の面が交わっている (この仮定はのちに曲面物体に拡張する);

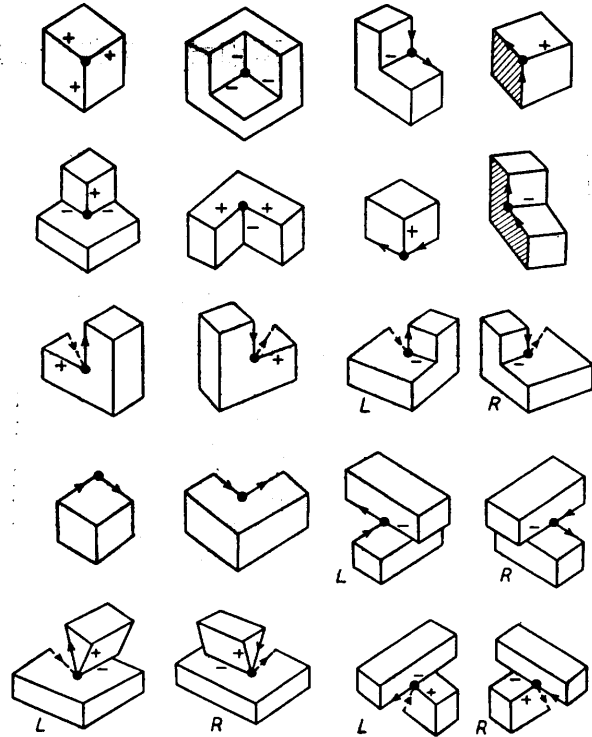


図 2 可能頂点の一覧表

Fig. 2 Permissible junction types.

(2) シーン中の多面体は互いに一般の位置にある。すなわち 2 物体が接するとき稜線や頂点同志が一致しない。(3) 眼は一般の位置にある、すなわち眼の位置を少し動かしたとき線画が急激に変わることはない。以上の仮定のもとで線画に現われ得る頂点 (これを可能頂点と呼ぶ) の種類を数え上げると、図 2 のような結果になる<sup>6)</sup>。ただしここでは、線画平面上での平行移動と回転、および隣り合う線のなす角が  $\pi$  より大きいか小さいかを不変に保つ変形 (これは眼の位置を少し変えたときの線画の変化に対応する) で互いに移り得る頂点は同じものとみなし、また、自明の方法で単純な複数個の頂点に分割できるもの (6 章の分割規則 1 を参照) も省略した。さらに、影線上の頂点は、影の原因となる輪郭線がその頂点に接続しているもの (たとえば図 1 (b) の  $p, q, r, s$ ) のみを数え上げた。図中、斜線をほどこした面は光源からは見えるがカメラからは見えない面を表わす。また、 $L$  または  $R$  で示した頂点は、2 物体が接触したところで生ずる可能性のある頂点で、シーン中の物体を一つ一つ切り離すための重要な手掛かりとなる。

ここで注目すべきは、可能頂点の種類が非常に少ないという点である。実際、図 2 に現われる頂点の線の

符号を取り替えてあらゆる組合せを機械的に作ると、その種類は 2000 を越す。そのうち（影線上の頂点を除けば）わずか 20 個のみが線画の中に許される頂点であり、従って図 2 は対象物体の世界に関する一般知識を効率良く凝縮した表現法だということがわかる。

#### 4. 頂点辞書

可能頂点および可能頂点から線を一部除いてできる頂点の集合を  $N$  とする。頂点辞書とは、 $N$  に属す各頂点についてその属性および他の頂点との関係を記述したもので、次の 5 つ組で定義される： $(N, f, B, g, h)$ 。ここに、 $f$  は  $N$  から  $\{0, 1\}$  への関数で、頂点  $x (\in N)$  が可能頂点のとき  $f(x)=1$ 、それ以外のとき  $f(x)=0$  である。 $B$  は  $N \times N$  の部分集合で、頂点  $x, y (\in N)$  に対して  $(x, y) \in B$  となるのは、頂点  $x$  に線を 1 本加えて頂点  $y$  を作る事ができるときかつそのときに限る。 $g$  と  $h$  は  $B$  を定義域とする関数で、 $g(x, y)$  は頂点  $x$  から頂点  $y$  を作る時追加すべき線の種類を表わし、 $h(x, y)$  は追加すべき線の許される方向範囲（反時計回りの角度区間で）表わす。辞書中の  $N$  と  $B$  に着目すると、節点集合を  $N$  とし枝集合を  $B$  とする有向グラフ  $(N, B)$  を作る事ができる。

頂点辞書の一部を図 3 に示す。図中、影の原因となる輪郭線と結果となる影線の間は、弧でつないで表現した。ここに描かれた 7 個の頂点のうち、長方形で囲んだ 3 個は不可能頂点、円で囲んだ 4 個は可能頂点である。可能点のそばには、それを実現する線画の例を示した。2 個の頂点を結ぶ矢印は  $B$  の要素を表わしている。たとえば、頂点  $t$  を構成する影線と輪郭線の（ある基準方向から反時計回りに測った）方向角度を  $\theta_1, \theta_2$  とすると、図中の斜線部で示した角度区間  $[\theta_2 - \pi, \theta_1]$  に凸稜線を加えると頂点  $u$  が得られる。従って  $(t, u) \in B$ ,

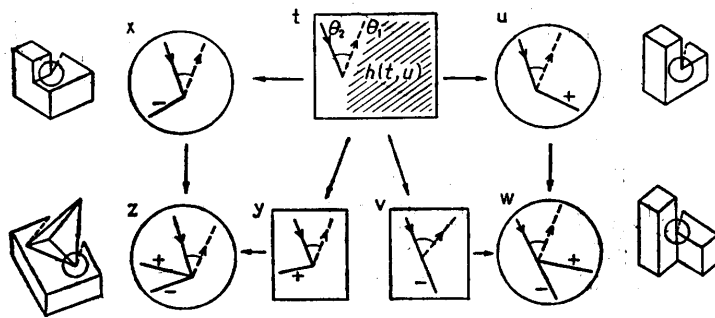


図 3 頂点辞書の一部  
Fig. 3 Part of the junction dictionary.

$$g(t, u) = \text{凸稜線},$$

$$h(t, u) = [\theta_2 - \pi, \theta_1]$$

である。頂点  $u$  にさらに凹稜線を加えて頂点  $w$  を作ることができ、 $(t, u, w)$  という頂点変形の経路ができる。このような経路が、まだみつかっていない稜線の種類と方向を示唆し、距離画像解析のための案内役となる。さらにこの示唆は強さを持つ。たとえば、解析の途中で頂点  $u$  がみつかったとすると、これはすでに可能頂点だから  $w$  に変更されるかも知れないし変更されないかも知れない。従って  $(u, w)$  は弱い予測である。これに反して、頂点  $v$  がみつかった場合には、 $v$  は不可能頂点だから必ず  $w$  へ変更されることがわかり、 $(v, w)$  は稜線の強い予測である。

#### 5. 頂点改善手順

距離画像からある程度の稜線が抽出されたとき、そこに現われる任意の頂点  $x$  に着目し、 $x$  のまわりにまだみつかっていない稜線を探し頂点  $x$  を改善する方法について考える。

頂点辞書から作られる有向グラフ  $(N, B)$  において、 $x$  から出ている枝の集合を  $G(x)$  とする： $G(x) = \{(x, y) | (x, y) \in B\}$ 。前章でみたように、 $G(x)$  に属す各枝  $(x, y)$  は  $x$  からさらにのびている稜線の可能性を示唆し、その稜線の種類と存在範囲はそれぞれ  $g(x, y)$ ,  $h(x, y)$  である、 $h(x, y)$  は一般には扇形領域である。そこで、距離画像の中の  $h(x, y)$  で指定される領域に、2章で定義した稜線検出演算をほどこし、演算値の絶対値の  $h(x, y)$  における最大値を  $(x, y)$  で予測される稜線候補の信頼度とみなし  $r(x, y)$  と記す。そして、 $r(x) = \max \{r(x, y) | (x, y) \in G(x)\}$  を計算する。 $r(x)$  が十分大きければ新しい稜線がみつかったとみなせるが、あまり大きくないときは判断の確実性を増すためにさらに次のような検査を行う。グラフ

$(N, B)$  における有向経路  $p = (x, z_1, z_2, \dots, z_k, y)$  ( $0 \leq k, k=0$  ときは  $p = (x, y)$  とみなす) が、 $f(z_i) = 0$  ( $1 \leq i \leq k$ ),  $f(y) = 1$  を満たすとき、極小可能経路と呼ぶ。経路  $p$  の信頼度を  $r'(p) = \min \{r(x, z_1), r(z_1, z_2), \dots, r(z_k, y)\}$  で定義する。 $x$  を始点とする極小可能経路の集合を  $H(x)$  とするとき、 $s(x) = \max \{r'(p) | p \in H(x)\}$  と定義すると、 $s(x)$  は頂点  $x$  に稜線 (1 本とは限らない) を付け加えて他の可能頂点へ移る可能性の強さを表わすとみなすことがで

きる。

以上を考慮して、頂点  $x_0$  が与えられたとき次のような頂点改善手順を考えることができる。ここに現われる  $T_1, T_2, T_3$  は稜線を採用するか否かを判定するための域値で  $T_1 > T_2 > T_3$  である (特に指定のないときはすぐ次のステップに移るものとする)。

- 1)  $x = x_0$  と  $x$  を定める。
- 2)  $r(x)$  を計算する。  $r(x) \geq T_1$  なら 4) へ行く。
- 3)  $s(x)$  を計算する。  $f(x) = 1, s(x) \geq T_2$  なら 4) へ行く。  $f(x) = 0, s(x) \geq T_3$  なら 4) へ行く。他の場合は 6) へ行く。
- 4)  $r(x)$  の値を実現する稜線候補 (ステップ 2 から来た場合) または  $s(x)$  の値を実現する稜線候補集合 (ステップ 3 から来た場合) を新しい稜線として採用し、頂点  $x$  に接続する。稜線の追加によって頂点の種類は  $x$  から  $y$  に変わったものとする。
- 5)  $x = y$  と  $x$  を変更し、2) へ行く。
- 6) 処理を終わる。

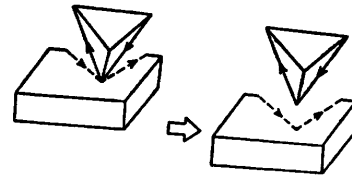
実際のシステムでは、線面の頂点を 1 個 1 個取り出して上記の処理を行う。そこでは、みつかった線の追跡、その線上に新しくみつかった頂点の登録、稜の終端の処理等が加わるが、それらはのちに例で見ることにする。

### 6. 物体分割規則

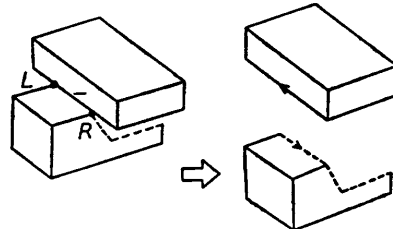
シーンを単純な物体に分割することは、シーン中の物体の同定等に役立つ。しかし、分割は一意に解の決まる問題ではなく、シーンを記述するための便法に過ぎない。ここでは、頂点の形を手掛かりとした次の 2 つの分割規則を考える。

**【分割規則 1】** 頂点に入る矢線 (輪郭線または影線) と頂点から出る矢線の対が 2 組あったら、入る矢線とその右隣りの出る矢線を対にして分割する (図 4 (a))。

**【分割規則 2】** 図 2 で定義された  $R$  に属す頂点と  $L$  に属す頂点がある、または  $R$  に属す頂点と影線上の頂点がある、または影線上の頂点と  $L$  に属す頂点がある、凹稜線の列で結ばれているとき、その凹稜線列に沿って物体を分割し、凹稜線列を、 $R$  から  $L$  に向かって右側の物体では輪郭線でおき



(a) 分割規則 1



(b) 分割規則 2

図 4 物体分割規則

Fig. 4 Body partitioning rules.

かえ、左側の物体では影線でおきかえる (図 4 (b))。

### 7. 距離画像解析

前章までに見てきた稜線検出演算・頂点辞書・物体分割規則を組み合わせた距離画像解析システムの振舞いを処理例で示す。

2 個の角柱の相貫物体が直方体の上に載っているシーンの距離画像を図 5 (a) に示す。ここでは、垂直なスリット光をシーンに当て、テレビカメラでとらえたその像を重ね合わせて表示してある。

システムは最初に、物体の輪郭線と影線を抽出し (距離画像では輪郭線と影線とを容易に区別できる<sup>6)</sup>)、それらの線上の頂点をみつける (b) (みつかった線は

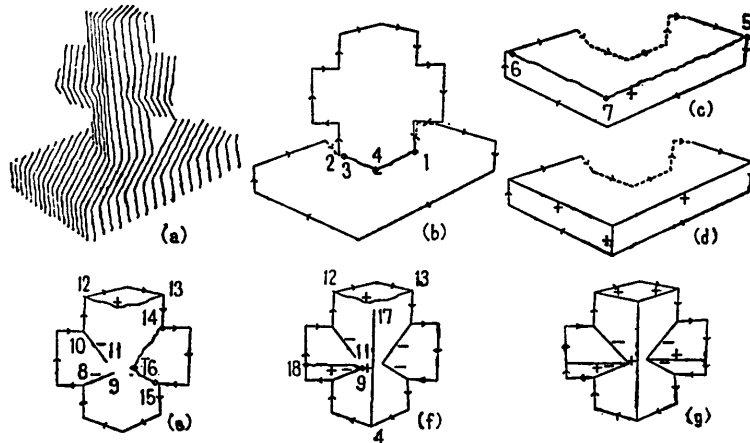


図 5 多面体シーン 1 の解析

Fig. 5 Analysis of polyhedral scene 1.

初めは点列の形をしているが、直線の一部でおきかえられるものは、どんどんおきかえていく。

次に、現在得られている頂点を1個1個辞書と比べ、頂点改善手順に従って稜線を見つけていく。まず、頂点1をとりあげ(これは図3の頂点 $t$ と同種)、辞書からの示唆に従ってそのまわりに線を探すと凹稜線が見つかる(図3の稜線予測 $(t, x)$ )。これを追跡し、稜線1~3(頂点 $a, b$ を結ぶ稜線を $a \sim b$ と記す)を抽出する。この線の終端3と途中にみつかった頂点4とを、新しい頂点として登録する。頂点1のまわりには他の稜線が見つからないので次の頂点に移る。頂点2を調べたとき、この頂点に接続可能な凹稜線3~4がすでにみつかったことがわかるので、頂点3を頂点2に一致させる。ここで頂点1, 2とそれを結ぶ凹稜線列1~4, 4~2とに分割規則2が適用でき、シーンは2個の物体に分けられ、それぞれについて処理が進む。一方の物体(c)では、頂点5を調べたとき凸稜線5~6とその上の頂点7が見つかる。同様の処理を繰り返して、直方体の記述が構成される(d)。もう一方の物体も、途中結果(e), (f)を経由して最終結果(g)に達する。図中の番号は、稜線や頂点のみつかった順を表わしている。

一般のシーンでは、解析の出発点である輪郭線に連結していない稜線もあるため、上記処理のあとシステムは画像全体を粗く調べ、新しい頂点や稜線が見つかることを新たな出発点として同様の処理を繰り返す。

図6は別の処理例で、3個の直方体からなるシーンの距離画像と解析結果である。

上記両例で示すように、解析の結果、シーンは単純な物体に分割され、それぞれが正しく記述される。このような記述が得られれば、シーン中の物体を同定したり、位置・姿勢を知ったりすることは容易である。

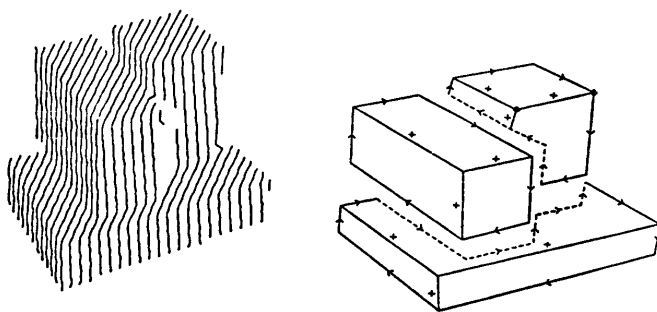
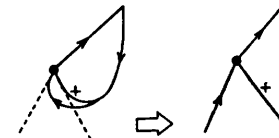
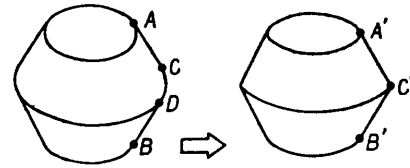


図6 多面体シーン2の解析  
Fig. 6 Analysis of polyhedral scene 2.



(a)接線による置替え



(b)擬似的な頂点

図7 曲面物体のための便法

Fig. 7 Conventions for curved objects.

## 8. 曲面物体への拡張

シーンが曲面物体を含むと、その線画には曲線が現われ、可能頂点の種類も急激に増加する。しかし、図7(a)のように、曲線を頂点における接線方向の半直線でおきかえ、また図7(b)のように、接線の一致する2曲線の方向を少し変更し擬似的な頂点におきかえると、可能頂点の数はわずかに増加するだけで済む<sup>6)</sup>。限られた分解能の画像から抽出された線の一部を局所的に見たとき、直線と曲線を区別したり接線方向を正確に知ることはむずかしいため、図7のおきかえは実用的立場からも望ましいものである。

3章で述べたシーンに関する3つの仮定のうち第1のものを、「(1') 物体表面は平面および2次曲面から構成され、各頂点ではちょうど3個の面が交わる」という仮定でおきかえ、仮定(1'), (2), (3)のもとで頂点辞書を作り、それを利用した解析例を図8に示す。円柱・球・立方体・直方体が正しく分離されていることがわかる。

## 9. むすび

頂点辞書を利用した距離画像解析法を提案し、その有用性を実験例で示した。本解析法の特徴は次のとおりである：

- (1) 辞書によって次に探すべき稜線が予測できるため、稜線抽出の効率と信頼性が高い。
- (2) 辞書には物体に関するかなり一般的な知識が集約されるため、シーン中の物体をあらかじめいくつかの形(直方体、三角柱等)に限る必要はない。
- (3) シーンを単純物体に分割できる。

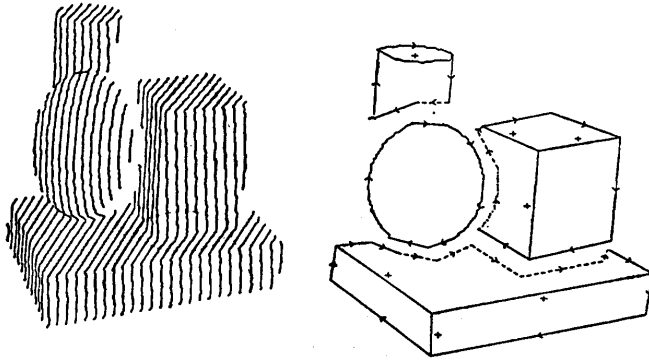


図 8 曲面物体シーンの解析

Fig. 8 Analysis of plane and curved objects.

- (4) 解析の結果得られる線画は、単なる線の集まりでなく、線の種類および頂点の種類が正しく記述されたものであるため、物体の同定等にただちに役立つ。

本解析法はかなり一般的な物体を扱えるというものの、面は平面と2次曲面に限る等の制約がある。これらの制約をはずし、このままの形で任意の物体を扱おうとすると、辞書が膨大となり、その結果、稜線の探索効率が低下して本解析法の良さが生かされなくなるであろう。しかし、実際の場面ではシーンの性質がある程度わかっているのが普通だから、(1)各頂点の出現頻度を調べて、探す稜線に優先順位を設けるか、あるいは、(2)シーンの性質が変わるごとに辞書を入れ替えられるよう、辞書の自動作成システムを開発する方向で解決できるであろう。

謝辞 有益な御批判と御助言をくださった電子技術

総合研究所バイオニクス研究室長杉江昇博士、および田村浩一郎博士・白井良明博士をはじめとする同研究室の皆様に感謝申し上げます。

### 参 考 文 献

- 1) Y. Shirai and M. Suwa: Recognition of Polyhedrons with a Range Finder, Second International Joint Conference on Artificial Intelligence, pp. 80-87, London (1971).
- 2) G. J. Agin: Representation and Description of Curved Objects, Stanford Artificial Intelligence Project Memo, AIM-173 (STAN-CS-72-305), p. 134 (1972).
- 3) M. Oshima and Y. Shirai: Representation of Curved Objects Using Three-Dimensional Information, Second USA-Japan Computer Conference Proceedings, pp. 108-112 (1975).
- 4) D. Waltz: Understanding Line Drawings of Scenes with Shadows, The Psychology of Computer Vision, P.H. Winston ed., pp. 19-91, McGraw-Hill, New York (1975).
- 5) 大島正毅, 高野善雄: 3次元物体認識システムのための特殊ハードウェア, 電子技術総合研究所集報, Vol. 37, No. 5, pp. 493-501 (1973).
- 6) K. Sugihara: Dictionary-Guided Scene Analysis Based on Depth Information, Report on Pattern Information Processing System, No. 13, pp. 48-122, Electrotechnical Laboratory, Tokyo (1977).

(昭和53年2月6日受付)

(昭和53年5月23日採録)