

階層的な形状プリミティブの並列的発見

小泉 正彦*

*松下電器産業(株)東京研究所

富田 文明**

**電子技術総合研究所

包括的な物体の概念を表現する方法として、物体を物体に共通的なパターン(プリミティブ)によって構成し、認識過程では、オクルージョンや照明の加減によって物体の一部が見えなかったり、物体の表面に各種の模様、影、ノイズがある場合でも、画像から各プリミティブを並列的に発見する方法を提案する。そして、この方法を、個別的な物体の幾何モデルと画像データを照合する場合にも、認識の手がかりとして、画像データから上記のプリミティブを最初に発見することによって、それに関連するモデルだけを効率よく呼び出す方法として考える。また、ステレオ法における対応問題への適用についても述べる。

Parallel Detection of Hierarchical Shape Primitives

Masahiko Koizumi*

Fumiaki Tomita**

* Tokyo Lab., MATSUSHITA Electric Ind. Co., Ltd.
3-10-1 Higashimita, Tama-Ku, Kawasaki 214, Japan

** Electrotechnical Laboratory
1-1-4 Umezono, Tsukuba 305, Japan

This paper presents a method to represent objects by generic patterns (primitives) of objects as comprehensive concepts of objects, and to detect them parallelly in images in which some portions of objects are not observable because of occlusions and lighting conditions and objects may be covered by textures, shadows, or noises. This method can be used even for matching individual geometric models of objects with image data by first detecting primitives in image data and efficiently invoking only the models related to them. This method is also applied to the correspondence problem of stereo vision.

1. まえがき

これまで、「物体認識」を物体の幾何モデルと画像データとの間の対応点を求め、その座標変換を求める問題として考えることが多かった^{(1)~(5)}。この場合、座標変換を決定するためには、少なくとも3個の対応点を求める必要がある。ところが、一般的な環境、すなわち、未知の物体を含む複数の物体が存在し、オクルージョンや照明の加減によって物体の一部が見えなかったりする状況においては、たとえ認識しようとする物体が単純なものであったとしても、正しい対応点を求めることは容易ではない。さらに、実際には多くの物体のモデルと多くの姿勢があり、その中から適切なモデルと姿勢を選択しなければならない^{(6)~(8)}。いま、モデルの数をM、モデルiの頂点の数を m_i 、データの頂点の数をnとすると、調べるべき対応点の組み合わせは、単純に計算すると

$$\sum_i^M m_i C_3 \cdot n C_3$$

となる。したがって、現実問題として、できるだけこの探索空間を減らすことが必要である。しかし、従来の多くの方法がそうであるように、対応の単位として頂点のような局所的で数の多いものを用い、総当たり探索を基本とするモデル駆動のトップダウンによる逐次的で定量的な照合を行なうと、処理の時間的爆発が生じる。この問題を解決する方法は、対応の単位を大きくして数を減らし、データ駆動のボトムアップによる並列のおよび定性的な照合を行なうことである。

さらに、「物体認識」の広い意味からは、個々の物体に対して正確な幾何モデルを用意することはむしろ特殊である。例えば、道路シーンにおいて、道路上の自動車を認識するためにすべての車種の幾何モデルまで用意する必要はなく、せいぜい「自動車」であることが認識できればよいのである。そこで、本論文では、包括的な物体の概念（例えば、上記の「自動車」）を表現する方法の一つとして、物体を物体に共通的な形状パターン（以下、プリミティブと呼ぶ）によって構成し、それらを画像から並列的に発見する方法を提案する。そして、この方法を、個別的な物体の幾何モデルと画像データを照合する場合にも、認識の手がかりとして、画像データから上記のようなプリミティブを最初に発見し、それに関連するモデルだけを効率よく呼び出す方法として考える。また、ステレオ法における対応問題への適用についても

述べる。

2. プリミティブの階層

物体を構成するプリミティブには、図1のように、構造体—素立体—面—辺のような階層がある。例えば、カメラを考えてみよう。カメラ（構造体）の外形は「直方体」（素立体）と「円柱」（素立体）からなる。直方体は3個の「平行四辺形」（面）から、円柱は「楕円」（面）と「円柱側面」（面）からなる。平行四辺形は「平行線」（辺）と「凸角」（辺）からなり、楕円は「楕円弧」（辺）からなる。これらはすべてプリミティブである。個々の物体のすべての要素をプリミティブとして登録する必要はなく、むしろ、特殊な要素は排除されるべきである。そして、プリミティブは、原則的に、絶対的な寸法を考慮しないで定性的に表現される。したがって、下位のプリミティブは上位のプリミティブによって共有される。

認識過程では、下位のプリミティブほど発見するのが容易であり、またそのように構成されなければならない。下位のプリミティブが発見されると、それらを要素とする上位のプリミティブの発見ルーチンが発火する。

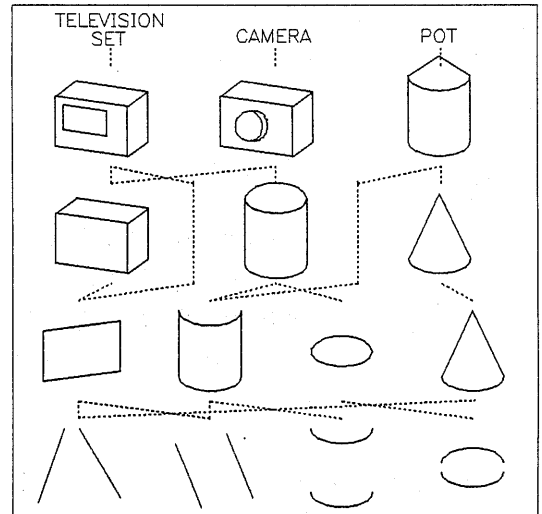


図1 物体のプリミティブの階層構造

3. プリミティブの発見

そこで、画像からどのようなプリミティブをどのように発見するかが最も重要な課題となる。本論文では、画像データとして、画像の境界表現⁽⁹⁾

を用いる。そして、最下位のプリミティブ（原子または一次プリミティブ）として、一様な領域の境界線を構成し、領域を右に見るように方向付けられた線分または楕円弧を用いる。したがって、楕円弧には、領域の凸部分を表す楕円弧と凹部分を表す楕円弧が存在する。例えば、図5(a)は、ジュースの缶とポットの画像で、(b)はそれから得られた一次プリミティブである。

上位のプリミティブは、これらの一次プリミティブを組み合わせることで構成されるが、利用する一次プリミティブの個数にしたがって、二次プリミティブ、三次プリミティブ、…と分類することができる。二次プリミティブは、平行線、凸角、同一楕円上の楕円弧対などのように、一次プリミティブの対として発見される。三次以上の高次プリミティブとしては様々なものが考えられるが、ここでは、平行四辺形、楕円、円柱側面などの面型プリミティブと、直方体、円柱などの素立体型プリミティブ、そして、テレビ、ポットなどの構造体型プリミティブを発見する方法について述べる。

要求される仕様は、オクルージョンや照明の加減によって物体の一部が見えなかったり、物体の表面に各種の模様、影、ノイズがある場合でも、それぞれのプリミティブを安定して発見できることである。

3.1 二次プリミティブ

任意の一对の一次プリミティブが、指定された条件（関係）を一定誤差以下で満足する場合に、それぞれのリストに登録される。

3.1.1 PLリスト

平行かつ逆方向の線分対のリストである。

3.1.2 CLリスト

同一直線上の同方向の線分対のリストである。

3.1.3 CVリスト

凸角を形成する、またはその可能性のある（したがって、必ずしも連結している必要はない）線分対のリストである。いま、図2のように、2つの線分を表すベクトルを a と c 、 a の終点から c の始点へ向かうベクトルを b 、 c の終点から a の始点へ向かうベクトルを d とすると、次の条件が満足されるとき、 (a, c) は CV リストに登録される。

$$\begin{aligned} a \times b &< 0 \\ b \times c &< 0 \\ c \times d &< 0 \\ d \times a &< 0 \end{aligned}$$

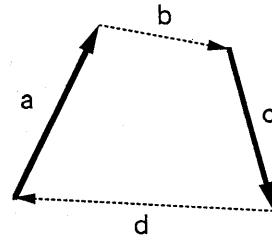


図2 凸角を形成する可能性のある線分対

3.1.4 EAリスト

同一楕円上の同方向の楕円弧対のリストである。ここで注意すべきは、本来同じ楕円上の楕円弧であっても、画像上でそれぞれが分割されて楕円近似されると、デジタル化の影響もあって、同じ方程式で表されるとは限らず、むしろまったく異なる楕円で近似されるものと考えた方がよい。したがって、各楕円弧対に対して楕円当てはめを新たに試みなければならない。

3.2 面型プリミティブ

面型プリミティブは閉ループ型プリミティブである。したがって、構成要素は異なるが同一とみなせるプリミティブの重複を検証することができる。また、全長が定まるので、発見されたプリミティブの構成要素の長さの合計の全長に対する比を、そのプリミティブの確からしさの指標として用いることができる。

3.2.1 平行四辺形

平行四辺形は、PLリストおよびCVリストを用いて発見することができる。

- 1) いま、`parallel(x, y)` という手続きは、線分対 (x, y) が PL リストに入っているとき真の値を返すものとする。また、`convex(x, y)` という手続きは、線分対 (x, y) が CV リストに入っているとき真の値を返すものとする。4本の線分の組 (a, b, c, d) は、図3のように、次の値を真とすると、平行四辺形の候補となる。

```
parallel(a, c)
parallel(b, d)
convex(a, b)
convex(b, c)
convex(c, d)
convex(d, a)
```

- 2) 上記の手続きは並列的に行われるため、図3のように、平行四辺形の1つの辺が2つ以上の線分に分かれている場合、同じ平行四辺形

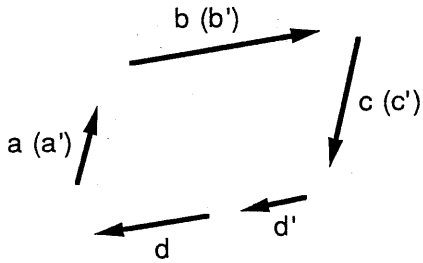


図3 平行四辺形を構成する線分の組

が異なった線分の組として発見されることになる。したがって、平行四辺形候補が作成された後、すべての2つの平行四辺形候補、例えば(a, b, c, d)と(a', b', c', d')について、それらが同じ平行四辺形を表しているかどうかを検証しなければならない。もし、次の値を真とするとき、2つの平行四辺形候補は統合される。

- equal (a, a') \vee colinear (a, a')
- equal (b, b') \vee colinear (b, b')
- equal (c, c') \vee colinear (c, c')
- equal (d, d') \vee colinear (d, d')

ここで、equal (x, y) という手続きは、xとyが同じものであるとき真の値を返すものとする。また、colinear (x, y) という手続きは、線分対(x, y)がCLリストに入っているとき真の値を返すものとする。

- 3) 最後に、各平行四辺形候補に対して平行四辺形としての確からしさが検証される。いま、平行四辺形候補がN本の線分で構成されており、各線分の長さを l_i ($i=1, \dots, N$)、平行四辺形の周長をLとすると、確からしさc.f.は

$$c.f. = \sum_i^N l_i / L$$

によって定義される。そして、平行四辺形候補の中で、このc.f.の値が閾値を越えるものを平行四辺形リストに登録する。

3.2.2 楕円

楕円は、EAリストまたは楕円弧リストを用いて発見することができる。

- 1) いま、EAリストの2つの楕円弧対が、(a, b)と(b, c)のように楕円弧bを共有するとき、(a, b, c)の組み合わせに対して楕円を当てはめ、その誤差が小さい場合には合併する。以下同様にして、他の楕円弧対を併合する。そして、他の組み合わせに含まれない最大の組み合わせをすべて作成し、それらを楕円候補

とする。

- 2) また、EAリストに含まれない楕円弧と、EAリストの中で他の楕円弧対と合併することのできない楕円弧対も、そのまま楕円候補とする。

- 3) 楕円候補の中で(図5(c))、確からしさの高いものを楕円リストに登録する(図5(d))。

3.2.3 円柱側面

円柱側面は、PLリストと楕円弧リストを用いて発見することができる。いま、PLリストの各線分対(a, b)に対して、次の条件を満たす楕円弧対(c, d)を組み合わせて円柱側面の候補を求める(図5(e))。

- ・楕円弧cとdは、線分aとbの直線に挟まれた領域に存在する。
- ・楕円弧cとdのうち一方は凸、他方は凹である。

しかし、上記の条件だけでは、図4(a)のように、正しい楕円弧が選ばれる場合もあるし、同図(b)のように、偽の楕円弧も選ばれる場合もある。また、すでに述べたように、楕円弧の方程式は非常に不安定であるので、楕円弧cとdが円柱側面の辺として妥当かどうかを調べる次の手続きが必要となる。

- 1) 線分aとbの直線の中心線を求める(同図(c), (d))。
- 2) 楕円弧cとdに対して、中心線を軸とする線対称な楕円弧c'とd'を求める(同図(e), (f))。
- 3) 楕円弧対(c, c')と(d, d')のそれぞれに対して楕円を当てはめる。当てはめの誤差が大きい場合は、円柱側面候補とはなり得ない。したがって、同図(b)の組み合わせは除去される。
- 4) そして、楕円当てはめの誤差が小さく、楕円の長軸が側面の幅に対して十分に長く、長軸の直線がaまたはbと交差しない場合に、(a, b, c, d)を円柱側面候補とする(同図(g))。
- 5) 円柱側面候補の中で、同じ円柱側面を表す円柱側面候補を統合した後、確からしさの高いものを円柱側面リストに登録する(図5(f))。

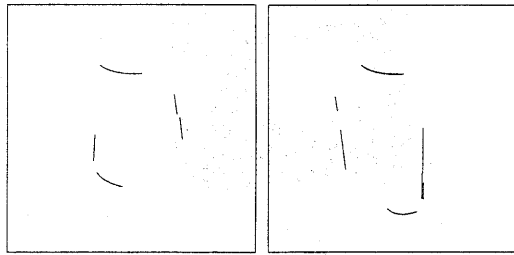
3.2.3 円錐側面

円錐側面は、CVリストと楕円弧リストを用いて、円柱側面の場合と同様にして、図5(g)の円錐側面候補から同図(h)の組み合わせが発見される。

3.3 立体型プリミティブ

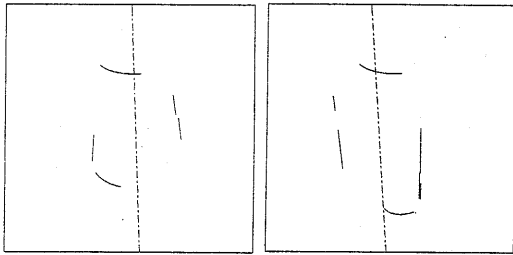
3.3.1 直方体

3個以上の平行四辺形が発見された場合に、その中のどれか3個の平行四辺形の組が直方体を構成するかを調べる。いま、3個の平行四辺形をA、



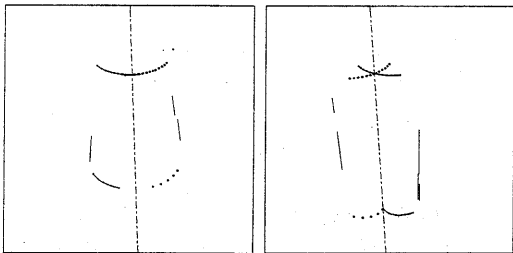
(a)

(b)



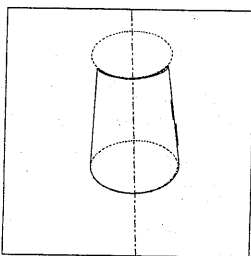
(c)

(d)



(e)

(f)



(g)

図4 円柱側面候補の検証

B、Cとすると、次の値を真とするとき、その3個の平行四辺形は直方体を構成する。

adjacent (A, B)

adjacent (B, C)

adjacent (C, A)

ここで、adjacent (X, Y) は、Xを構成する辺 x_i ($i = 1, \dots, 4$) とYを構成する辺 y_j ($j = 1, \dots, 4$) の中に、equal ($x_i, -y_j$) を真とするものが存在する場合に真となる。

例えば、図6 (a)はテレビの画像で、(b)はそれから得られた一次プリミティブ (線分と楕円弧)、そして、(c)はその中で発見された直方体である。

3.3.2 円柱

円柱は、楕円リストと円柱側面リストを用いて発見することができる。いま、楕円Pと円柱側面Qが次の値を真とするとき、それらは円柱を構成する (図5 (i))。

adjacent (P, Q)

3.4 構造体型プリミティブ

3.4.1 ボット

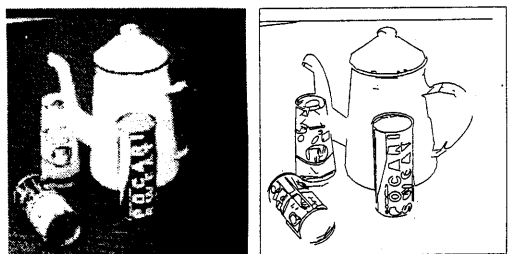
図5 (j) は、同図 (b) から発見されたすべての高次プリミティブを表すが、ボットは、楕円弧を介して隣接する円柱 (側面) と円錐 (側面) の組み合わせとして発見される。

3.4.2 テレビ

テレビの発見のためには、直方体のプリミティブ以外に、図6 (d) のように、「入籠の平行四辺形」というプリミティブを発見する。そして、テレビは、同図 (e) のように、直方体と入籠の平行四辺形の組み合わせによって発見される。

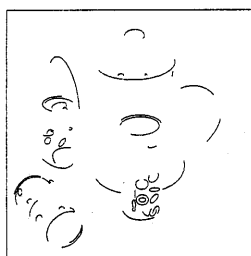
4. ステレオへの応用

二台のカメラを用いるステレオは、三角測量の原理に基づいた最も簡便な距離測定法である。しかし、ステレオには、一方の画像の点が他方の画像のどの点に対応するのかを決定する対応問題がある。対応の単位として、従来のエッジ点から、最近では、直線や曲線のセグメントを用いるようになってきている。これは、対応の単位が大きいほど単位の個数が減少するとともに、対応の曖昧さも減少し、対応の探索空間が減少する利点があるからである。本章では、対応の単位をさらに拡大し、画像から発見的に得られるプリミティブの構造体—素立体—一面—辺の階層構造を用いる方法について述べる。

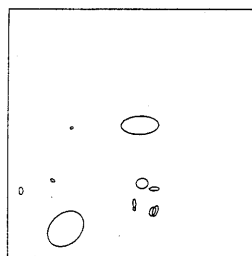


(a)

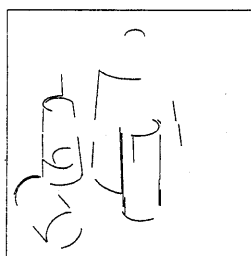
(b)



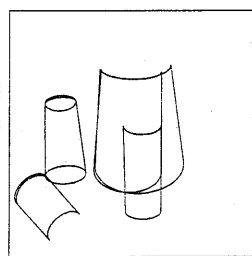
(c)



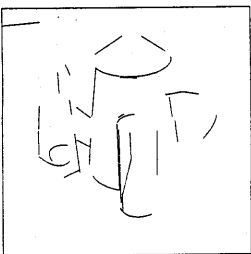
(d)



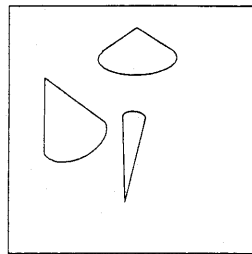
(e)



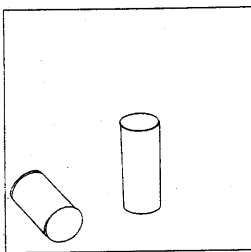
(f)



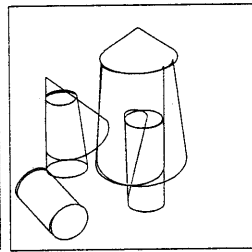
(g)



(h)

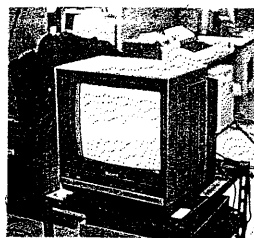


(i)

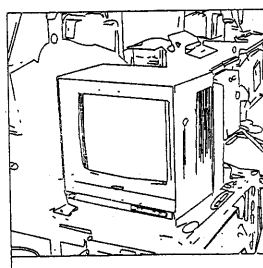


(j)

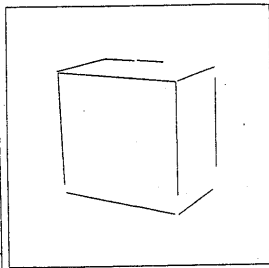
図5 ジュースの缶とポットのあるシーン



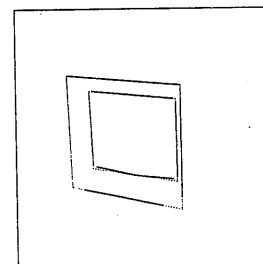
(a)



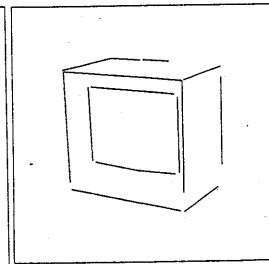
(b)



(c)



(d)



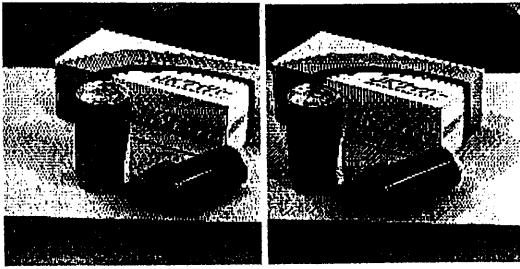
(e)

図6 テレビのシーン

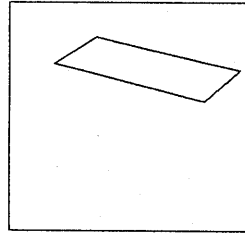
4.1 階層的対応法

左右のステレオカメラから得られる画像は、線分と楕円弧によって表現される。そして、左右の画像からそれぞれ独立に上位のプリミティブが発見される。発見したプリミティブは、前述のように、階層的に表現されている。そして、左右の画像で発見されたプリミティブの対応探索は、上位のものから下位に向かって行なわれ、最終的には辺（線分または楕円弧）を対応の最小単位とする。上位のプリミティブほど拘束条件が多く、個数も少ないので対応が容易になる。また、上位の安定した対応結果が下位に継承されるので、下位のプリミティブの対応探索を省略することができる。プリミティブの対応に用いる拘束条件は、

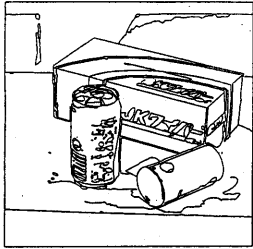
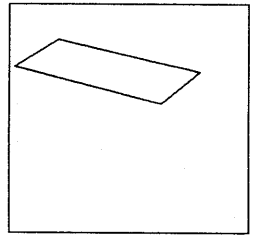
- (1) プリミティブのレベル、
- (2) プリミティブの種類、
- (3) エピポーラ拘束



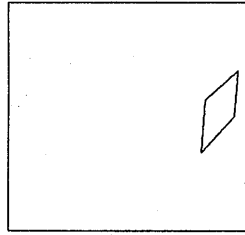
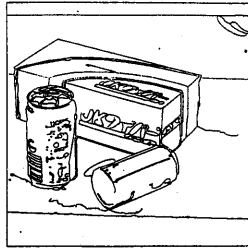
(a)



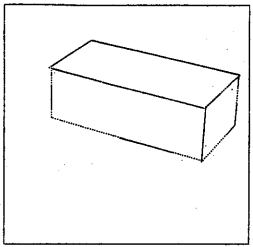
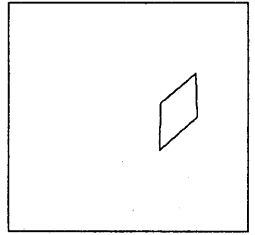
(f)



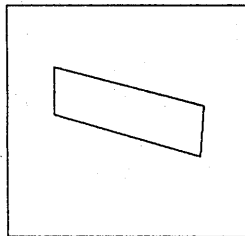
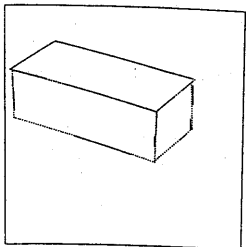
(b)



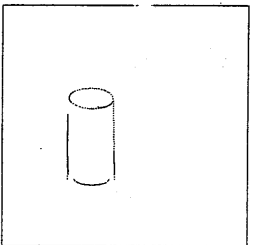
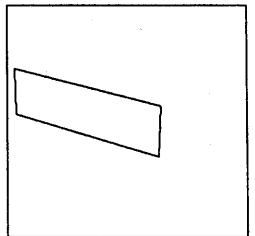
(g)



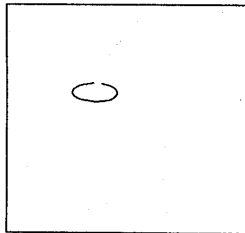
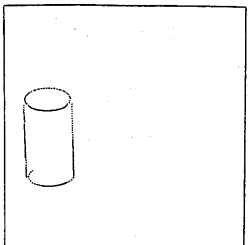
(c)



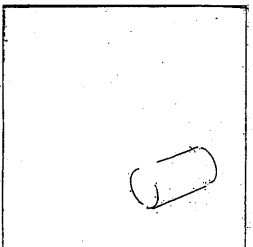
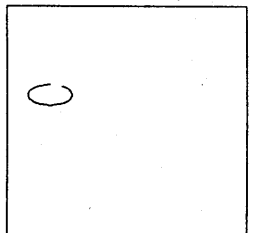
(h)



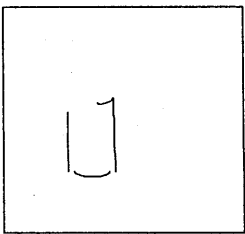
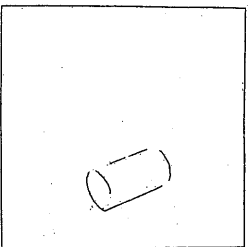
(d)



(i)



(e)



(j)

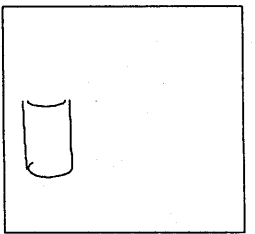


図7 ティシュペーパーの箱とジュースの缶あるシーン

である。すなわち、(1)により同じレベルのプリミティブ間でしか照合しない。(2)により同じ種類のプリミティブ間でしか照合しない。そして、(3)では、簡単のために左右画像の各プリミティブの外接長方形を考え、その頂点の座標をそれぞれ $(x_L(i), y_L(i))$, $(x_R(i), y_R(i))$ ($i = 1, \dots, 4$) とすると、カメラが平行である (または、画像がそのように変換されている) ことを前提として、次の条件を満足しなければならない。

$$\begin{aligned} x_L(i) &> x_R(i) \\ y_L(i) &= y_R(i) \end{aligned}$$

上位から下位への対応の伝播の結果、最下位のレベルですべての辺の間で厳密なエビポーラ拘束が満足されると、対応が完了する。

4.2 実験結果

図7(a)は、ティッシュペーパーの箱とジュースの缶のステレオ画像であり、(b)はそれから得られた一次プリミティブである。同図(c)は対応する直方体、(d)と(e)は対応する円柱、(f)~(h)は対応する平行四辺形、(i)は対応する楕円であり、(j)は対応する円柱側面である。このように、画像からできるだけ上位のプリミティブが発見できれば、対応が容易になることがわかる。

上記の処理は、ステレオ画像の対応法であるが、結果として、3次元プリミティブの発見法の1つでもある (他に、ステレオにより3次元データが得られた後に、3次元プリミティブを発見する方法もある)。

5. あとがき

画像から得られる物体の不完全な線画から、物体を構成する共通的なプリミティブを並列的に発見する方法について述べた。この方法はまた、物体認識に限らず、ステレオ、モーションなどコンピュータビジョンにおけるすべての対応問題に対して有効な方法となろう。さらに、基本となるすべてのプリミティブとプリミティブ間の関係 (adjacentなど) を定義することができれば、Winstonのアーチの例のように⁽¹⁰⁾、新しい物体をサンプルとして与えた画像から、プリミティブの階層を自動生成し、上位の新しいプリミティブ (概念) を学習することが可能となる。しかし、まだ一般的にどのようなプリミティブと関係を利用すべきかは明らかではない。また、複雑な物体を扱う場合には、特定のプリミティブをトップダウン的に発見する処理も必要になるであろう。そして、

このようなシステムを実現するための並列協調アーキテクチャの開発も重要な課題である。

参考文献

- (1) Roberts, L.: Machine perception of three-dimensional information, in Tippet, J. (Ed.), Optical and Electro-Optical Information Processing, pp.159-197, MIT Press, Cambridge, MA., 1965.
- (2) Brooks, R. A., Greiner, R., and Binford, T. O.: The ACRONYM model-based vision system, Proc. 6th IJCAI, 1979, pp.105-113.
- (3) Oshima, M. and Shirai, Y.: Object recognition using three-dimensional information, IEEE Trans., PAMI-5, 4, pp.353-361, 1983.
- (4) Lowe, D. G.: Perceptual Organization and Visual Recognition, Kluwer, Boston, 1985.
- (5) 富田: 3次元物体の境界線に基づく記述の生成と認識, 信学論, J71-D, 2, pp.343-349, 1988.
- (6) Fishler, R. B.: Model invocation for three dimensional scene understanding, Proc. 10th IJCAI, 1987, pp.805-807.
- (7) Burns, J. B. and Kitchen, L. J.: Recognition in 2D images of 3D objects from large model bases using prediction hierarchies, Proc. 10th IJCAI, 1987, pp.763-766.
- (8) 小泉, 富田: ソリッドモデルと画像の定性的及び定量的照合, 情報処理学会, CV54-5, 1988.
- (9) 富田, 高橋: 画像のB-REPのためのアルゴリズム, 信学会, PRU86-87, 1987.
- (10) Winston, P. H.: Learning structural descriptions from examples, Winston, P. H. (Ed.), The Psychology of Computer Vision, pp.157-210, McGraw-Hill, New York, 1975.