

仮説検証モジュールの階層的構成による シーン理解

荒川賢一 尺長 健 金子 博

NTTヒューマンインタフェース研究所

シーン理解とは、シーンに存在する物体を画像情報により認識することである。この認識過程を物体の同定、及び物体に関する諸情報の獲得の2段階から成るとする。まず、我々は、シーン理解に用いるシーンの知識を記述する枠組みとして、シーンの構成要素を階層的に記述する景観記述モデルを導入する。景観記述モデルでは、各物体は物体木と呼ぶ記述単位に分解して記述されており、その選択が物体同定を意味する。また、物体の概形が選択に用いられる情報として、物体木中に記述される。物体同定過程は、画像情報から得られる特徴と物体に関する知識とを対応付ける過程である。本報告では、画像情報から得られた特徴と物体概形の部分形状について階層的に記述された知識との間で、逐次、照合をおこなっていく仮説検証モジュールによる物体同定過程の実現法を提案する。物体同定過程は、仮説検証モジュールの利用により、仮説単位形成のための統合ルールと仮説する形状に関する知識が用意されることで、実現可能となる。その一例として、室内シーンについて距離画像から概形を獲得する過程をシミュレーション実験し、その結果について考察した。

An Approach to Scene Analysis

Using Hypothesis Generation and verification Modules

Ken-ichi ARAKAWA, Takeshi SHAKUNAGA and Hiroshi KANEKO

NTT Human Interface Laboratories

1-2356 Take Yokosuka-shi

Kanagawa 238-03 Japan

Scene analysis is regarded as the recognition process of objects which exist in scenes. The recognition process consists of the object identification step and the configuration map acquisition step. We propose the Scene Representation Model as an adequate framework of knowledge representation for the steps. In the model, an object is represented as a processing element named the Object Tree. The selection of an Object Tree means the object identification. Descriptions in a level of the Object Tree called the Coarse Structure Level are suitable for the key of the selection.

Furthermore, we propose the Hypothesis Generation and Verification Module, which is a processing module with a simple control flow. The object identification step can be developed easily using the modules, because the knowledge necessary for the processing of each module is definite; rules of feature merging and frames as candidates of hypotheses. We also report the simulation of the object identification step (coarse structure detection) and discuss its performance.

1. はじめに

シーン理解は、対象シーン内に存在するそれぞれの物体について、画像情報をもとに物体認識をおこなうことである。ここで、物体認識とは、物体の同定、及び、物体に関する諸情報（例えば、配置情報、色彩情報など）の獲得の2段階からなるとみなせる。本報告では、対象世界を人工物からなる室内シーンに限定し、世界マップ（物体の配置マップ）の獲得を目的として、シーン理解問題を論ずる。

シーン理解の研究において、いかに処理を進めるかという問題と同様に、いかにシーンを記述するかがひとつの問題として扱われてきた¹⁾²⁾。これらの問題は、本来切り離して考えることはできないものである。本報告では、まず、後者の問題に対して知識の記述の枠組みを導入し、その後、前者の問題を考察する。

我々は、目的である配置情報の獲得に必要な知識を記述する枠組みとして、シーンの構成要素を階層的に記述する景観記述モデルを提案する。景観記述モデルでは、物体を物体木という単位で階層的に記述し、物体木の集合でシーンを記述する。物体の同定は、物体木の選択により実現される。

一方、処理の流れの観点から見ると、知識の存在を仮定したmodel-based visionの研究が多くおこなわれている³⁾が、その制御の流れは、一般に複雑であるため、目的遂行のための知識の利用が場当り的になるという問題がある。本報告では、物体同定過程（ボトムアップ過程）に注目し、仮説検証モジュールの階層的構成による実現法を提案する。これは、形状の知識と画像情報中の特徴を局所的なものから大局的なものへと順次、対応付けしていくことで物体同定をおこなうことに相当する。ここで、仮説検証モジュールは、ある特徴に対応する仮説の獲得（生成）と検証をおこなう処理単位で、その内部での制御が非常に単純であるという特徴を持つ。

以下、2章では、シーンを記述するの枠組みとして、景観記述モデルを導入する。3章においては、仮説検証モジュールの基本構成と、その連続から成るシーン理解のボトムアップ過程について述べる。また、4章では、以上の考えに基づいたシーン理解実験について報告し、5章において、今後の課題について考察する。

2. 景観記述モデル

シーンに関する知識を記述する枠組みとして、シーンの構成要素である物体を階層的に記述する景観記述モデルを導入する。景観記述モデルは、次の2つの条件を満足する必要がある。

- ①物体の同定に有効な情報を記述し、しかもそれが使いやすい形式で含まれている。
- ②対象シーンに存在する物体に関して、配置情報獲得などの目的遂行に必要な情報を記述しうる。

条件①は、物体に対応する知識を選択するのに効果的で、かつ、画像情報からも獲得しやすい情報をモデルが持つことを

要請する。このような情報として物体の構造を残した大まかな形状（概形）に注目する。一方、条件②に関しては、本報告では、シーン中の物体の配置マップ作成をシーン理解の目的として想定している。このため、景観記述モデルにより、配置情報を獲得するのに必要な物体の詳細な形状やサイズ等の情報、さらに、ある時点の物体の配置情報までもひとつの枠組みの中に記述できるようにする。

2.1 物体木

景観記述モデルは、シーンを物体（人工物）の集合として記述する。その物体の記述単位にあたるのが、物体木である。物体木の例を図1に示す。物体木には、物体に関する情報を次の4階層で記述する。

- (1) 概念レベル：物体木のルートであり、物体の概念に対応する。物体の概念とは、ここでは、例えば、いす、机といったような言語的に一般名詞で表現される単位に一般化されたものとする。すなわち、物体木は物体を表現する一般名詞に対応して存在する。
- (2) 概形記述レベル：限られた基本形状の組合せにより、物体の基本的な構造を大まかに表現する概形記述に対応する階層である。概形は、形状の中でも画像情報から比較的獲得しやすいものであり、物体の同定（物体木の検索で実現される）において、キーとして使われる。詳しくは、2.2節で述べる。
- (3) 詳細記述レベル：人工物は、ある設計図面に従い、大量に生産されることが多い。そのスペック情報に対応する属性を含むのが、この階層である。形状について言えば、CADで作成された計算機上での記述に含まれる情報と等価な情報が記述される階層である。また、色や（詳細な）大きさも、この階層の属性として扱われる。
- (4) インスタンスレベル：ある時点での物体のインスタンス（実在）を同定するのに必要な属性のうち、詳細記述レベルまでで扱われないもの記述される。本報告では、物体の配置情報（世界座標系内での絶対座標、向き）をこの階層の属性として扱う。

なお、インスタンスに対応する情報は、属性の継承（インヘリタンス）の考え方を適用して参照することで、獲得可能である。また、インスタンスレベルの情報の集合と、シーン内の物体の配置マップは等価である。

ところで、シーン理解において、物体の配置に関する拘束が知識として役立つ場合がある。例えば、同じ中空の円筒構造を持つコップとくずいれを考えたとき、コップは机などの物体の上に置かれるが、くずいれは床に直接置かれることが多い。このような情報を記述するため、物体間の配置に関する拘束関係を物体木間の拘束グラフとして表現することが可能である。また、例えばオフィスに存在する物体について物体木を集合化することにより、環境に合わせた知識ベースの選択が可能のように管理することもできる。

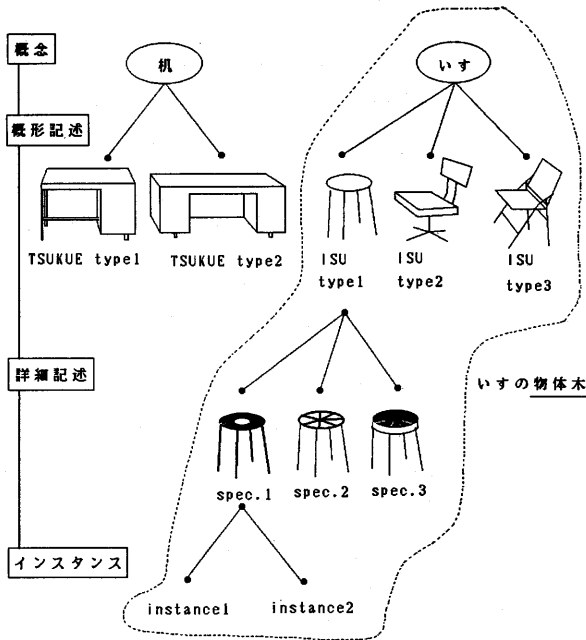


図1 景観記述モデルの物体木

2.2 概形記述

概形記述と同様に物体の大きな形状を階層的に記述する考えは従来からあり⁴⁾、ACRONYM⁵⁾のobject graphにおいても利用されている。object graphは物体の大きな形状と詳細な形状の階層的な関係を示すが、その各ノードは、一般化円筒の集合として記述した形状を参照する。本報告でも、物体の形状を分解し、限定された基本形状に近似して記述することで概形記述を実現する。すなわち、概形記述は、基本形状に対応するノードと、接続関係を示すアークにより構成されるグラフ構造で表現される。その例を図2に示す。なお、一般的に物体の部分形状への分解は一意ではない。しかし、高々2次曲面からなる形状に限定すれば、変換アルゴリズムを統一することで、一意に分解できる可能性がある⁶⁾。この事実については、更に検討せねばならないが、概形を高々2次曲面までの基本形状の接続したものと限定することで分解の一意性について可能性を残しておく。従って、本報告では、基本形状として{多面体、1表面が2次曲面であるボリューム、補助形状}を考える(図3参照)。補助形状とは、棒、中空の円筒形などの概形記述に必要である小体積要素をいう。

概形記述は、物体に関する知識の単位である物体木を検索するための手がかりとして利用できる。我々は、シーン理解過程のうち、物体同定過程を実現する際、物体木の検索において、キーとして概形記述を用いる。

3. 仮説検証モジュールによる物体同定過程の構成

シーン理解は、画像情報から得られる情報により物体を同定する過程(物体同定過程)と、個々の物体に関する知識により物体に関する諸情報(配置・サイズ・色彩など)を抽出する過程(物体具現化過程)の組合せにより構成されるとみ

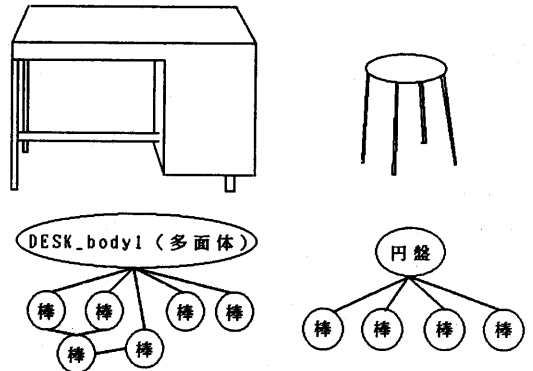
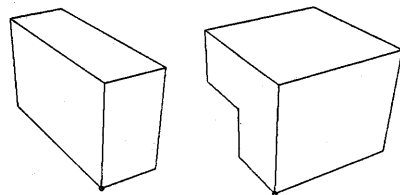


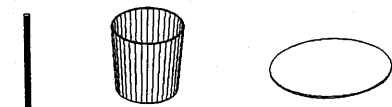
図2 概形記述の例(TSUKUE type1, ISU type1)



・多面体の例



・2次曲面を含むボリュームの例



・補助形状の例(棒, 中空の円筒, 円盤)

図3 基本形状の例

なすことができる。ここで、一般には、両者は独立ではなく、相補的な関係にあると考えられ⁷⁾、両者を包括的に制御する機構が必要であると考えられる。

本報告では、物体同定過程をボトムアップ過程としてまず実行し、その後、トップダウン過程として物体具現化過程を実行することでシーン理解を簡単化して実現することを考える。以下、その前半としての物体同定過程の構成について論じる。すなわち、複数の物体を含むシーンにおいて、各物体に関する知識ベースが与えられている状態で、物体同定過程を連続した仮説検証過程により実現する方法を提案する。

3.1節では、物体同定過程の大局的な構成を示す。また、3.2節では、これを実現するための処理要素である仮説検証モジュールについて、その構造を示すと同時に、具備すべき条件を論じる。

3.1 物体同定過程の階層的構成

物体の同定を画像情報からおこなうとき、その結果として物体に対応する2次元の領域に物体の種類ラベルが付加される。すなわち、画像の面から処理を考えると、画像情報への物体に関する知識に対応するラベル付けと考えることができる。画像情報上での意味のあるラベルが付けられる構造的な単位を画像特徴単位と呼ぶ。例えば、直線のエッジ、一樣領域などは画像特徴単位である。物体同定過程では、以下の理由により、局所的な画像特徴単位から順に大局的なものへ知識との対応を表すラベルを付けていくべきであると考えられる。

- ①画像上では局所的な画像特徴単位の方が検出しやすい。
- ②局所的な画像特徴単位から順に大局的な画像特徴単位に統合しつつラベル付けすれば、物体に対応する単位への切り出しが同時におこなわれることになる。

画像情報上の領域にラベル付けをする問題は従来から、多く論じられている⁸⁾⁹⁾が、これらにおいては、3次元構造との対応を明確に考えてはいない。本報告では、物体の同定は、物体の概形によっておこなわれる。物体の概形獲得は、その物体の概形に関する知識との対応を示すラベルを画像単位特徴に付加することでおこなう。その過程は、画像特徴単位と概形の3次元的部分形状の知識との対応付けを順次おこなっていくことにより実現される。すなわち、3次元の形状の(部分的な)見えが画像特徴単位であるとみなしながら、ラベル付けがおこなう。

一般に、ある画像特徴単位に対して形状に関するラベルを付ける処理は仮説検証とみなせる。なぜなら、画像特徴単位は形状の見えにしかすぎず、対応する3次元形状を画像特徴単位から一意に定めることはできない。そこで、画像特徴単位に仮に形状を対応付けた後、なんらかの検証が必要になるわけである。すなわち、この考えに従うと、画像上での各画像特徴単位に対応する形状の階層を考察し、その階層に合わせて仮説検証過程を構成することにより物体同定過程を実現することが可能となる。このとき、各階層で仮説検証によるラベル付けのメカニズムと、局所的な画像特徴単位から大局的な画像特徴単位を作り出す統合のメカニズムが必要である。これらは、具体的には、次節で提案する仮説検証モジュールにより実現される。

3.2 仮説検証モジュール

前節の考えに従うと、物体同定過程にあたる形状獲得は、図4に示すように画像特徴単位に対して統合、及び、仮説獲得・検証による知識との対応付け(ラベル付け)を連続的におこなうことで一般的に実現できる。ここでは、その各ステップを実現する仮説検証モジュールと呼ぶ処理要素を提案する。

仮説検証モジュールの構成を図5に示す。仮説検証モジュールは、以下の基本的な処理により構成される。

- (1)仮説に対応する画像特徴単位の抽出
- (2)仮説獲得
- (3)仮説検証

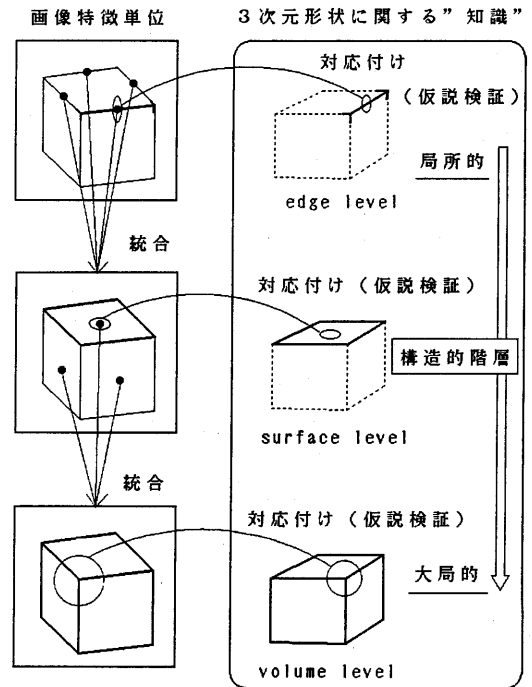


図4 階層的な知識と画像特徴単位の対応付けの例

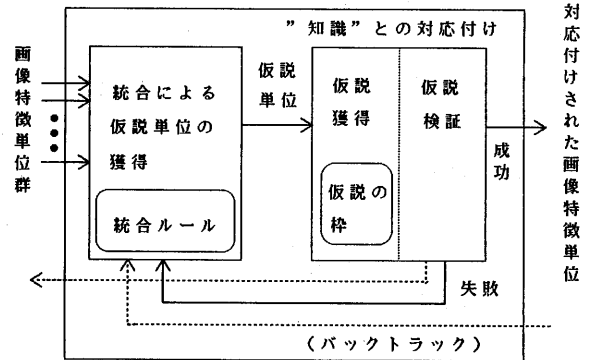


図5 仮説検証モジュールの基本構成

仮説検証モジュールでは、ある3次元の形状に関する画像情報上での見え(a view)にあたる画像特徴単位を抽出し、その形状としての仮説(形状知識との対応付けの仮説)を形成し、さらに、対応する知識から画像特徴単位を解析することで検証する。画像特徴単位は局所的なものから順に大局的なものへとラベル付けしていくので、前段階のモジュールで仮説検証された画像特徴単位を統合することにより、そのモジュールで注目する画像特徴単位が抽出される。

ここで、仮説検証モジュールを用いる際に、以下のことを考察する必要がある。

- ①画像特徴単位として画像情報中に現れる形状の階層の設定
- ②各仮説検証モジュールでの画像特徴単位の統合ルール
- ③各仮説検証モジュールでの仮説検証法

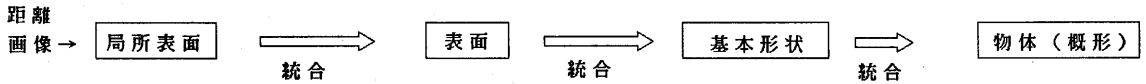
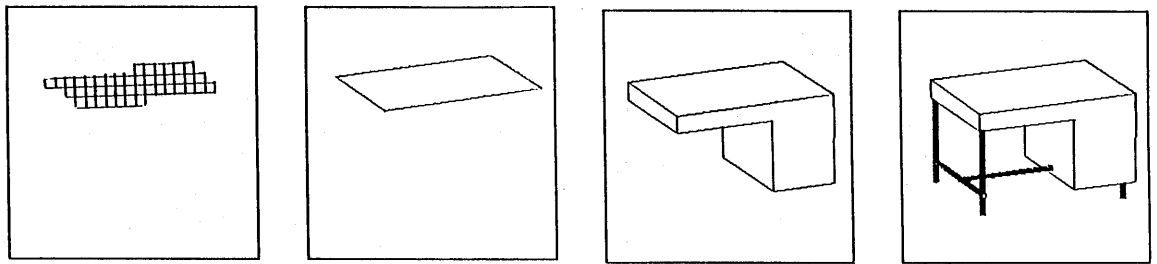


図6 距離画像からの物体同定過程

なお、フレーム¹⁰⁾のスロットを適当な値で埋めていくことで仮説検証を実行するメカニズムを利用すれば、仮説に対応するフレームと統合ルールを用意することで物体同定過程を構成することができる。この観点から、距離画像からの物体同定過程を例にとり、次章でその構成について述べる。

4. 距離画像からの物体同定過程の実現

距離画像から物体の概形を求め、物体を同定する過程を仮説検証モジュールにより構成した。距離画像からの3次元形状獲得については、従来からいくつかの研究がある³⁾。例えば、大島らの方法¹¹⁾では、ボトムアップ過程を経て得られた核と呼ばれる表面によりモデルを選択し、核に隣接する表面からmodel-drivenの照合（トップダウン過程）をしていくことで形状を獲得している。これに対し、本報告では、表面に注目して概形を獲得していく過程を仮説検証モジュールを用いることで実現する。ここで、便宜上、距離画像について平行投影の仮定をおく。

距離画像から物体の概形を求めるボトムアップ過程を仮説検証モジュールにより構成するとき、まず、各モジュールで獲得する画像特徴単位を定める必要がある。高々2次曲面からなる概形を求める過程を図6に示すように、画素→局所表面→表面→基本形状→概形の階層的な形状にあたる距離画像上での画像特徴単位を順に獲得することで実現する。なお、以降では、各画像特徴単位を対応する形状にあわせて、表面特徴などと呼ぶ。また、仮説を与える画像特徴単位（群）を仮説単位と呼ぶ。各モジュールは、以下の処理を仮説検証によりおこなう。

- ・仮説検証モジュール1：隣接画素群を統合して、局所表面特徴を求める。局所表面特徴は、高々2次の小表面に対応する。
- ・仮説検証モジュール2：局所表面特徴群を統合して、表面特徴を求める。表面特徴は、高々2次曲面の表面に対応する。
- ・仮説検証モジュール3：隣接する表面特徴群を統合して、基本形状特徴を求める。
- ・仮説検証モジュール4：隣接する基本形状特徴群を統合し、概形特徴を求める。

この過程によって、概形を検索キーとし、物体木群の概形記述レベルを部分グラフの照合により検索することで、物体が同定される。これらの仮説検証モジュールで用いられる知識を表1に簡潔に示す。以下に、それぞれの仮説検証モジュールでの処理について述べる。

表1 仮説検証モジュールで用いる知識

	画素→局所表面特徴	局所表面特徴→表面特徴
統合ルール	距離値変化の連続性	距離値変化の連続性
仮説候補	平面、円筒面、球面、双曲面	平面、円筒面、球面

	表面特徴→基本形状特徴	基本形状特徴→概形特徴
統合ルール	隣接関係、境界形状	隣接関係
仮説候補	多面体、2次曲面を含むボリューム、補助形状等	(物体木の概形記述)

4.1 距離画像からの局所表面の獲得

距離画像を $z(x, y)$ で表す。距離画像上の間隔 s の格子点 (x_0, y_0) を局所表面特徴の核画素と呼ぶ。このとき、局所表面特徴 $LS(x_0, y_0) = \{z(x, y) | (x-x_0)^2 + (y-y_0)^2 < r^2\}$ は、以下の式で表される2次曲面に対応する円形の小領域（画素群）である。

$$z(x, y) = A \times x^2 + B \times y^2 + C \times x \times y + D \times x + E \times y + F \quad [4-1]$$

上記の小領域に平面、球面、円筒面、及び、双曲面の局所表面の種類を持つ仮説を与え、検証して、局所表面特徴を得る。
[仮説の獲得]

核画素から距離 r 以内にある画素群を局所表面の仮説を与える領域（仮説単位）とする。仮説の獲得法を以下に示す。

注*) 本報告では、基本形状として2次曲面を含むボリュームまで考えているが、実現については多面体の抽出までしかおこなっていない。

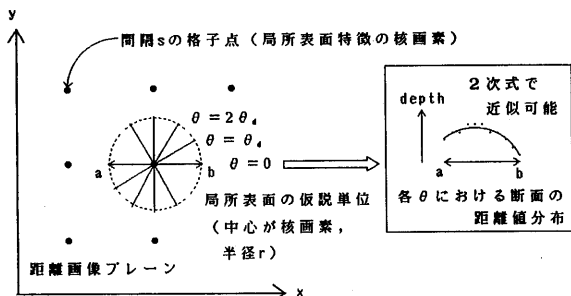


図7 局所表面の仮説単位の獲得

- (1) 核画素を通る断面の距離値分布の近似式を図7に示す各方向について求める。すなわち、各 θ について、

$$Z_{\theta}(i) = A_{\theta} \times i^2 + B_{\theta} \times i + C_{\theta} \quad [4-2]$$

$$\text{ここで、} i = \begin{cases} ((x-x_g)^2 + (y-y_g)^2)^{0.5} & \text{if } y > y_g \\ -((x-x_g)^2 + (y-y_g)^2)^{0.5} & \text{if } y < y_g \end{cases}$$

$-r < i < r$, r : 局所表面特徴の半径

$$\theta = 0, \theta_d, 2\theta_d, \dots < 180$$

を $y = \tan \theta \times (x - x_g) + y_g$ 上で、核画素からの距離が r 以下の画素の距離値より最小自乗近似で求める。

- (2) $Z_{\theta}(i)$ のうち、近似誤差の平均が大きなもの(近似ができなかった方向)が半数以上ある時、その仮説単位については局所表面仮説が得られないとする。

- (3) 近似誤差の平均が小さい $Z_{\theta}(i)$ の係数 A_{θ} 以下の基準を適用し、各種類の局所表面仮説が得る。

すべての A_{θ} の絶対値が十分小さい \rightarrow 平面

すべての A_{θ} の符号が同じで、絶対値が大きい \rightarrow 球面

すべての A_{θ} の符号が同じで、絶対値が変動する \rightarrow 円筒面
 A_{θ} の符号が正のものも負のものもあり、

絶対値が変動する \rightarrow 双曲面

仮説の枠はフレームにより表2のように記述され、局所表面の種類に対応したsubclassの選択がおこなわれる。

[仮説の検証]

核画素から半径 r 以内の領域(仮説単位)について、平面の仮説については[4-3]式、その他の仮説については[4-1]式の形の距離値の近似式を最小自乗近似により求め、その近似誤差が小さければ、検証されたとする。

$$z(x, y) = D \times x + E \times y + F \quad [4-3]$$

表2 仮説検証モジュール1のフレーム

(\$で始まる文字列は関数名を表す。下線付き文字列はスロット名を表す。)

(LOCAL-SURFACE

(circular-region ; top frame of local-surface

(subclass (val (local-surface)))

(center) : (xg, yg)

(radius) : r

(local-surface

(a-kind-of (val (circular-region)))

(subclass (candidate (planar cylindrical spherical hyperbolic

Sselect-local-surface)) ; selection of the kind

(ld-coefficients (if-needed \$local-surface-detector))

(coefficients (if-needed \$approximation-2d))

(error (if-needed \$calculate-approximate-error-2d))

...)

4.2 局所表面からの表面の獲得

仮説検証モジュール2では、隣接する局所表面特徴を統合し、仮説検証をおこない高々2次曲面の表面に対応する表面特徴を求める。対象シーンに含まれる物体の表面が高々2次曲面であると仮定すると、同一の表面に含まれる局所表面ペアについて、その間の断面の距離値の変化は高々2次である。この考えに従ったルールにより局所表面特徴を統合し、表面の仮説単位を求める(図8参照)。表面仮説については、平面、円筒面、球面の種類がある。また、最小自乗近似による2次曲面のフィッティングにより仮説を検証する。

[仮説の獲得]

以下の過程で局所表面を統合する。

- (1) 核画素が (x_{g1}, y_{g1}) , (x_{g2}, y_{g2}) である局所表面特徴ペアについて、それを結ぶ線分上での距離値分布の近似式 $Z_{g1-g2}(i)$ を最小自乗近似で求め、その近似誤差を計算する。

$$Z_{g1-g2}(i) = A \times i^2 + B \times i + C \quad [4-4]$$

ここで、上式は $(x-x_{g2})(y_{g1}-y_{g2}) = (y-y_{g2})(x_{g1}-x_{g2})$ の直線上で、 (x_{g1}, y_{g1}) , (x_{g2}, y_{g2}) 間にある各画素について求められるが、 i はその (x_{g1}, y_{g1}) からの距離である。局所表面特徴をノードとし、そのペア間の近似誤差の逆数をアークの重みとしてもつ重み付きグラフである局所表面結合グラフを作る(図8参照)。

- (2) その近似誤差が大きい(重みが小さい)アークを切り、局所表面結合グラフの最大連結部分グラフで表される局所表面特徴群を表面の仮説単位とする。

- (3) 仮説の面種については、平面の局所表面特徴が多く含まれる表面仮説を平面とし、また、曲率変化を局所表面と同様の方法で調べることで、円筒面・球面を判断する。

この過程で用いる表面仮説のフレームを表3に示す。

[仮説の検証]

表面仮説を形成する局所表面特徴群について、次の条件を満たせば、仮説を検証したものとする。

・[4-1], [4-3]式への最小自乗近似をおこない、その近似誤差が小さい。

・表面仮説の境界を以下に示す距離画像上での処理により獲得したとき、その境界が局所表面群をすべて含む。

- (1) 表面の近似式と距離画像の差を求め、これをあるしきい値で2値化する。

- (2) 表面仮説に含まれる局所表面を選び、その核画素の座標から最も近い境界点を探索することにより境界を抽出する。

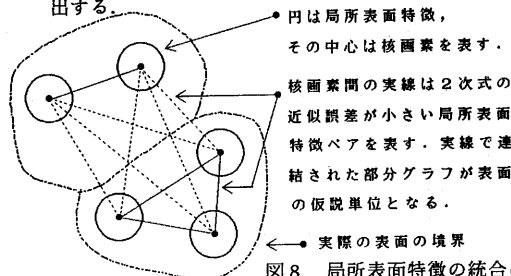


図8 局所表面特徴の統合による表面の仮説単位の獲得

表3 仮説検証モジュール2のフレーム

```
(SURFACE
  (region
    (a-set-of (val (circular-region)))
    (subclass (val (surface))))
  (surface
    (a-kind-of (val (region)))
    (subclass (candidate (planar cylindrical spherical hyperbolic)
      $select-surface)) : selection of the kind
    (coefficients (if-needed $surface-fitting))
    (error (if-needed $surface-fitting-error))
    (boundary (if-needed $calculate-boundary))
    (weight-point (if-needed $calculate-weight-point)))
  ...)
```

4.3 表面からの基本形状の獲得

表面特徴を統合して基本形状の仮説単位を形成し、仮説検証により基本形状特徴を得る過程について述べる。基本形状の種類については2.2節で述べたが、表面特徴群からは棒の仮説は得られないので、ここでは考えない。表面特徴群はその境界線の種類(ステップ、凹凸)により統合され、仮説単位が求められる。また、その基本形状仮説に対応するsubclassが選択される(すべてのスロットが適切な値で埋められる)ことで検証が進められる。

[仮説の獲得]

表面特徴について、距離画像上での隣接関係を調べる。隣接する表面特徴ペアについて、その間の境界の3次元的な形状を表面の近似式から求める。境界の3次元形状の種類として、ステップ状、凸状、凹状、及び縁状がある。各表面特徴をノードとし、隣接する表面特徴間のみアークを持つ表面隣接グラフを作り、以下のルールを適用して仮説単位となる表面特徴群を形成する。

- ・ステップ状、縁状の境界を挟む表面特徴ペア間のアークは切る。(オクルージョンエッジの可能性が高い)
- ・凸状の境界を挟む表面特徴ペア間のアークは接続したままにする。
- ・凹状の境界を挟む表面特徴ペア間のアークを順に切り、それぞれの時点で連結している部分グラフに含まれる表面特徴群を仮説単位とする。

表4 仮説検証モジュール3のフレーム

```
(PRIMITIVE
  (primitive
    (a-set-of (val (region)))
    (subclass (candidate (polyhedron cylinder sphere supplement)
      $select-primitive-by-surface-kind)))
  (polyhedron
    (a-kind-of (val (primitive)))
    (subclass (candidate (desk-body rectangular-parallelepiped ...)
      $select-polyhedron)))
  (desk-body
    (a-kind-of (val polyhedron))
    (a-set-of (val (surface
      (planar 4 (90 90 90 90) (1.0 0.7 1.0 0.7))
      (planar -6 (90 90 90 -90 90 90)
        (1.0 0.6 0.4 0.5 0.6 1.0))
      ...)))
    : attributes of surfaces :
    : (kind vertex# angle-list length-of-side-list)
  (connection-of-surfaces
    (val ((0 1 90) (0 2 90) (0 3 90) (0 4 90)
      ...)))) : surface connections
  ...)
```

[仮説の検証]

基本形状仮説のフレームを表4に示す。統合された表面特徴群を基本形状仮説の単位としてフレームの最上位に組み込む(基本形状仮説の獲得)。そののち、順にsubclassを表面群の部分グラフの照合により選択していく。そして、選択可能であった仮説のみ検証されたとする。例えば、多面体の検証にあたる照合法を以下に示す。

- (1) 表面特徴毎にその境界線に直線フィッティングし、多角形表現に変換する。
- (2) 表面特徴毎に時計周りの内角の列の対応を調べる。
- (3) 対応のついた表面の接続関係を表面特徴群と照合して、フレーム内に記述された表面群の部分グラフをなすか調べる。

4.4 基本形状からの物体の獲得

仮説検証モジュール4では、基本形状特徴を統合し、概形記述の仮説を獲得する。その仮説を物体木の概形記述レベルと照合し、対応するものがあれば検証されたとする。概形が得られることにより、物体木が選択され、物体の同定過程(ボトムアップ過程)が終了する。

[仮説の獲得]

概形(物体)の仮説単位を基本形状特徴から形成するのに、一般的に効果的な統合ルールはない。本報告の段階では、画像上で隣接する基本形状特徴について全ての組を考えて、仮説単位として扱うことにする。

[仮説の検証]

概形仮説のフレームの一部(最上位、及び、机のフレーム)を表5に示す。景観記述モデルに従い記述された物体に関する知識から切り出してきたフレームは、表5に示すフレームのsubclassとして扱われる。検証は、まず、仮説単位を最上位フレームに組み込み、順にsubclassを基本形状群の部分グラフの照合により選択していくことでおこなわれる。

表5 仮説検証モジュール4のフレーム

```
(OBJECT
  (object
    (a-set-of (val (primitive)))
    (subclass (candidate (desk1 chair1 dustpot ...) $select-object)))
  (desk1
    (a-set-of (val (desk-body %lf-foot %rf-foot %lb-foot %sidebar %bar
      %rb-foot))) : primitives (parts) : % means a stick.
    (position-of-parts
      (val (desk-body (0.3 -0.5 -0.4)
        (%lf-foot (-0.8 0 -0.4) (0.0 -1.0 0.0)
          ...)))
    (connection-of-parts
      (val ((desk-body %lf-foot (-0.8 0 -0.4)
        (desk-body %rf-foot (0.2 -0.5 -0.4)
          ...))))))
  ...)
```

4.5 シミュレーション実験

計算機上で生成した室内シーンの距離画像から物体を同定するシミュレーション実験をおこなった。図9-aに示す机、いす、くずいれから成るシーンの距離画像(512×512画素、256階調)について局所表面、及び、表面を獲得した結果をそれぞれ図9-b、cに示す。局所表面を画像上で比較的粗く検出したため、小さな表面が検出できなかった。画像上で十分な面



図9 実験結果

(a:原画像, b:局所表面特徴の獲得結果, 円形の濃淡の明→暗の順に平面, 円筒面, 球面, 双曲面を示す, c:表面特徴の獲得結果, d:多面体の基本形状の獲得結果(desk-body の各表面のエッジを表示))

a	b	c
d		

謝辞

日頃から熱心に御討議, 御討論頂くNTTヒューマンインタフェース研究所 視覚情報研究部 小森和昭部長, 末永康仁グループリーダーをはじめ, 視覚情報研究部の皆様に, 心から感謝致します。

参考文献

- 1)Marr, D. : VISION, 乾敏郎, 安藤広志 (訳), 産業図書 (1987)
- 2)Ballard, D. H., Brown, C. M. :COMPUTER VISION, 福村晃夫他 (訳), 日本コンピュータ協会 (1987)
- 3)Besl, J. B., Jain, R. C. :Three-Dimensional Object Recognition, Computing Surveys, Vol. 17, No. 1, (1985), pp. 75-145
- 4)Barr, A., Feigenbaum, E. A. (ed.): the Handbook of Artificial Intelligence, 田中幸吉, 淵一博 (監訳), 共立出版 (1983), 13章
- 5)Brooks, R. A. :Model-based Computer Vision, UMI Research Press (1984)
- 6)Xie, S., Calvert, T. W. :CSG-EESI:A New Solid Representation Scheme and a Conversion Expert System, IEEE Trans. PAMI, Vol. 10, No. 2 (1988), pp. 221-234
- 7)Lee, H. C., Fu, K. S. :Generating Object Descriptions for Model Retrieval, IEEE Trans. PAMI, Vol. 5, No. 5, (1983), pp. 462-471
- 8)Riseman, E. M., Hanson, A. R. :A Methodology for the Development of General Knowledge-based Vision System, in Vision, Brain, and Cooperative Computation (Arbib, M. A., Hanson, A. R. (ed.)), the MIT Press (1987), pp. 285-328
- 9)Ohta, Y. :A Region-oriented Image-analysis System by Computer, Ph. D. thesis, Kyoto Univ., (1980)
- 10)Winston, P. H. (ed.): Psychology of Computer Vision, 白井良明, 杉原厚吉 (訳), 産業図書 (1979)
- 11)大島正毅, 白井良明: 3次元情報を用いた物体認識, 信学論, Vol. J65-D, No. 5 (1982), pp. 629-636
- 12)Hwang, V. S-S., Davis, L. s., Matsuyama, T. :Hypothesis Integration in Image Understanding System, Computer Vision, Graphics, and Image Processing, Vol. 36 (1986), pp. 321-371

積を持つ表面群に対し, 多面体の基本形状を検出する仮説検証モジュールを実行した。机を構成する基本形状を検出した結果を図9-dに示す。なお, 法線方向が視線方向と直角に近い角を成す多面体については境界線形状が正しく求めにくい。内角の対応付けにおいて, 許容範囲を大きくする対処をする必要があった。

5. むすび

シーン理解過程は物体の同定, 及び, 物体間の配置情報等の獲得の2段階から成る。このうち, 本報告では, 物体同定過程に重点を置き, 仮説検証モジュールの導入により, その具体化を図った。そして, 仮説検証モジュールの階層的構成による実現例として, 距離画像からの物体同定実験をおこない, 多面体を概形とするものには良好な処理結果を得ることができた。これは, 予備実験ではあるが, 今後のより複雑な物体の同定に対しても提案手法の有効性を示すものである。

しかし, 残された問題も少なくない。2次曲面を含むボリュームや補助形状の獲得については今後インプリメントしていく予定であるが, 物体に関する知識の増加に伴う知識選択のメカニズム (実現においては, 計算量の爆発が起きない照合法など) や, 仮説集積 (hypothesis integration)¹²⁾ のメカニズム, 仮説検証モジュールにおけるフィードバック系などについては比較的簡単な対象を設定したために検討が不十分である。

シーン理解は, 最終的には, ボトムアップな処理とトップダウンな処理の融合により実現されると考えられる。本報告では, 主としてボトムアップな処理としての物体同定過程について報告した。今後は上記の問題の解決に加え, トップダウンアプローチの有効な利用法も合わせ, より汎用性があり, かつ, 柔軟なシーン理解系を実現していきたい。