

Smoothed Local Generalized Cones による パーツ構造の復元

佐藤 嘉伸 大谷 淳

NTT ヒューマンインタフェース研究所

あらまし: 本研究の目的は、視点に依存しない 3D 形状の分割記述を 2.5D 画像から復元することである。本稿では、*Smoothed Local Generalized Cones* (SLGCs) と呼ぶ新しい 3D 形状表現を提案する。この表現法は、Brady の提案による *Smoothed Local Symmetries* (SLSs) の 3 次元版と考えられる。まず、3D 形状自体によって定まる局所基準座標系を推定する方法を示す。2D 形状の局所座標系は局所対称により一意に定めることができるが、3D 表面形状の局所対称からは、局所座標系の位置と方向に関する拘束条件が得られるのみである。そこで、さらに、局所一般化円筒 (Local Generalized Cones: LGCs) の当てはめを行うことによって、局所座標系の決定を行う。SLGC プリミティブは、滑らかに接続する LGC の系列を抽出することによって復元される。次に、抽出された SLGC プリミティブを用いて、3D 形状のパーツ構造を復元する方法を述べる。パーツ構造復元のためには、SLGC プリミティブの選択・統合を行う必要があり、その基準を論じ、組み合わせ最適化問題として定式化する。本表現法の利点は、復元形状の多様性と復元の安定性である。局所的な規則性モデルを利用することにより、Superquadrics 等のパラメトリックモデルや大局的一般化円筒に比べ、柔軟にデータに当てはめることができるので、多様な形状の復元が可能である。安定性の観点からも、数値的に不安定な大局的モデルの直接的な当てはめに比べ有利である。得られた記述は、3D 物体認識において、形状の可変性を含む物体モデルの利用、大規模モデルベースの利用、および、物体モデル記述の学習を行う場合、特に有効である。

Recovering Part Structure by Smoothed Local Generalized Cones

Yoshinobu Sato and Jun Ohya

NTT Human Interface Laboratories

E-mail: sato%nttcvg.ntt.jp@relay.cs.net

1-2356, Take, Yokosuka-shi, Kanagawa-ken, 238-03, JAPAN

Abstract: The goal of this work is to recover viewpoint-independent descriptions of 3D shapes from 2.5D images. In this paper, we propose a novel 3D shape representation called the *Smoothed Local Generalized Cones* (SLGCs), which can be considered as a three dimensional extension of Brady's *Smoothed Local Symmetries* (SLSs). First, we present a method for estimating local canonical frames determined by the 3D shape itself. While local symmetry of 2D shapes determines its local frame, local (mirror) symmetry of 3D shapes only provides constraints on position and orientation of its local frame. The 3D local frames are estimated by finding frames where the Local Generalized Cones (LGCs) best fit to sensory data. The SLGC primitives can be recovered by extracting long sequences of LGCs connected smoothly. Next, we present a method for recovering a part structure of a 3D object using the SLGC primitives. The criteria for grouping and selecting the SLGC primitives to recover a part structure are presented and the problem is formulated as a combinatorial optimization. The advantages of the SLGCs are their versatility and stability through the use of local models. The local models can fit to data flexibly and do not suffer from unstable localization, as compared with using parametric models such as superquadrics or global Generalized Cones. Obtained descriptions are well-suited for object recognition with a large model database and learning object models.

1 まえがき

高度な3D物体認識システムには、物体モデル形状の可変性、大規模モデルベース、および、モデル記述の学習等を扱う能力が必要である。現在のシステムでは(例えば、[1])、これらの能力に重大な限界があり、その限界はモデルとの照合を行う際の記述レベルに起因している。この限界を克服するために、最近の研究では物体に固有の“パーツ構造(part structure)”の復元に関して、大きな注目が払われている[2],[3],[4],[5],[6]。

しかし、従来研究では以下のような問題点がある。まず、パーツの位置決め、および、分割は、非線型最適化に過度に依存しており、位置決めは数値的に不安定で、かつ、安定な分割を得ることは難しい。また、復元される形状の適用範囲にも制約が大きい。さらに、パーツ分割の問題は、本質的に曖昧性を含むものであるにもかかわらず、ほとんどの手法は、すべてのもってもらい分けを得ることなく、唯一の分割を得るのみである。これは、Marrの*least commitment principle*[3]に反する。曖昧さの解消には、対象に依存した知識が必要な場合も多く、パーツ復元の処理過程では、時期尚早な決定は避けるべきである。本研究の目的は、広範囲の形状に適用できる明確な基準に基づいたパーツの位置決め法、即ち、3D形状自体に基づくパーツの基準座標系(canonical frame)の決定法、および、すべてのもってもらい分けを安定に得るための方法を開発することである。近年、精密な距離画像入力装置が開発されている[7],[8]。そこで、本研究では、このような装置から獲得される密な表面情報が利用できることを前提とする。

パーツ構造の復元に関する従来研究は、3つのアプローチに分類できる。第1のアプローチは、パーツの表現にパラメトリックモデルを利用する[4],[5],[6],[9]。Superquadricsはその典型例である。Superquadricsは少数のパラメータで表現できるにもかかわらず、パーツ記述子としての特性をかなりの程度満足する点で魅力的である。しかし、適用範囲を拡大するためには、曲げ(bending)や先細め(tapering)などを導入する必要があり、非線形最小2乗法を必要とする位置決め処理は数値的に不安定なものとなる[10]。第2のアプローチは、一般化円筒(Generalized Cones: GCs)を用いる[2],[11],[12]。詳細に研究されているGCはStraight Homogeneous Generalized Cones (SHGCs)とそのサブクラスに限られ、したがって、復元形状の範囲は限定される。上記の2つのアプローチが大局的モデル(global model)を用いるのに対して、第3のアプローチは、局所的な規則性モデル(local regularity model)を用いる。BradyのSmoothed Local Symmetries (SLSs)は、2次元形状の記述に関するものではあるが、局所モデルを用いたパーツ構造復元のアプローチである[13],[14],[15]。SLSは、局所モデルとして、局所対称(local symmetry)を用いる。SLSでは、まず、局所対称を用いて、2D形状の局所基準座標系(local canonical frame)を決定し、次に、局所座標系を統合してパーツの候補を形成する。SLSは小さなノッチに過敏であると言われていたが、図1に示すように、選択と統合を行う後処理によって主構造とノッチの両方を復元できる。

本研究のアプローチは2D-SLSの3Dへの拡張である。局所モデルは、Superquadrics等のパラメトリックモデルや一般化円筒に比べ柔軟にデータに当てはめることができ、多様な形状復元が可能である。安定性の観点からも、数値的に不安定な大局的モデルの直接的当てはめに比べ有利である。SLSの3D拡張において、第1の問題は、“3D形状の局所基準座標系が、形状自体からどのように推定できるのか?”という問題である。この推定は、センサデータ中の規則性を綿密に解析することにより行う。センサデータから得られる局所座標系は、パーツの存在に関する有力な証言となる。2D-SLSと同様に、パーツ候補は局所座標系の統合によって形成される。第2の問題は、“どのようにパーツ候補を選択・統合してパーツ構造を復元するか?”という問題である。この問題は、一種の組合せ最適化問題と考えることができ、パーツ候補の選択・統合を行う基準を論じ、最適化問題として定式化する。

本稿の構成を以下に示す。2.では、局所基準座標系の推定について述べる。局所規則性モデルとして、“局所”一般化円筒(Local Generalized Cone: LGC)モデルを用いる。まず、局所面対称により局所座標系の自由度を拘束し、次に、LGCモデルを当てはめることによって局所座標系を決定する。3.では、局所面対称とLGCの抽出のインプリメンテーションの詳細と実距離画像に対する実験結果を示す。4.では、パーツ構造の復元を行うために、LGCの当てはめによって得られたパーツ候補をどのように選択・統合するかについて述べる。パーツ構造は、パーツの凸性(convexity)、最小パーツ数(minimal-number of parts)、データの最大被覆(maximal-covering)、および、パーツ間の最小重複(minimal-overlapping)の4つの基準に基づく目的関数を最小化するパーツ分割を探索することによって行う。5.では、本手法のまとめを行い、問題点と今後の課題について述べる。

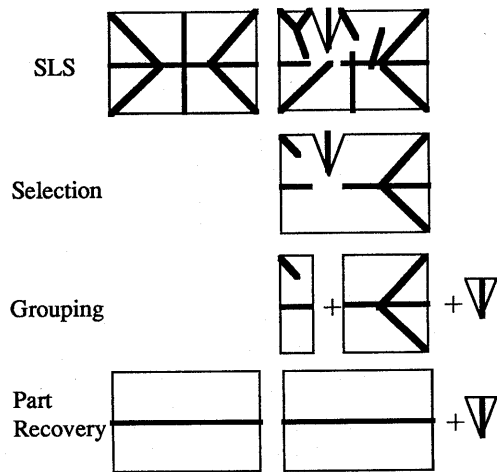


図1 Smoothed Local Symmetriesによるパーツ構造の復元

2 3D形状の局所基準座標系の推定

2.1 局所面対称による拘束

局所対称によって、座標軸の可能な2つの向きを除いて、2D形状の局所基準座標系を決定することができる(図2)。3D形状の局所面対称から、どのような拘束が与えられるだろうか？

x_1, x_2 を表面上の点の3D位置、 n_1, n_2 を、それぞれ、 x_1, x_2 における表面法線とする(図3)。これら2点の面対称条件は次式で表される:

$$(n_1 + n_2) \cdot (x_1 - x_2) = 0, \quad (1)$$

$$(n_1 - n_2) = \lambda(x_1 - x_2). \quad (2)$$

これら2点が面対称条件を満足し、かつ、2点間の断面形状が対称であるとみなされる場合、局所面対称が存在すると判断でき、その対称点 x_s 、および、局所対称面の法線 n_s は、次式で表される:

$$x_s = \frac{x_1 + x_2}{2}, \quad (3)$$

$$n_s = \frac{x_1 - x_2}{|x_1 - x_2|}. \quad (4)$$

局所対称面が定まったとして、局所座標系の自由度を限定する拘束を考えたとき、局所対称面が局所座標系の1つの座標面に一致すると仮定するのはきわめて妥当である。そこで、3D局所面対称により、局所座標系の位置・方向に対して以下の拘束を強いる(図4)。

1. 局所対称面は、局所座標系の yz -平面に対応する。

2. 局所対称点は、局所座標系の y -軸上である。

2番目の拘束は、局所座標系を対称点に関係づけるために必要である。以上の拘束のもとで、対称面内での z -軸の方向 θ 、および、対称点から z -軸までの距離 ρ の2自由度が残る。

2.2 局所一般化円筒の当てはめ

局所座標系の z -軸を決定するために3D形状自体の自然軸を検出し、それを z -軸に一致させることを考える。自然軸の検出において、なじみのあるモデルとして一般化円筒がある。そこで、“局所”一般化円筒(Local GC: LGC)を当てはめ、当てはまり具合のよい軸を選択するという方法をとる。即ち、 xy -平面で切断される切り口をGCの断面とし、 z -軸方向を掃引方向と考え、GCモデルを当てはめる。GCモデルとしては、できる限り一般性のあるものが要求される。そこで、本研究では、Shaferの分類[16]によるところのStraight Bilinear GC(SBGC)を用いる。通常用いられるHomogeneous GCは、軸に沿って一様な拡大・縮小のみが許されるが、Bilinear GCでは、断面上の2つの直交方向に対して異なる拡大・縮小を行うことができる。よって、Bilinear GCの適用範囲は、Homogeneous GCに比べ、大幅に拡大される。また、ここで考えているGCは、局所的な(即ち、軸が微小な長さしか持たない)GCであるので、軸が直線(straight)であり、掃引関数が線形(linear)であるというGCの線形近似の仮定は適用範囲を狭めることはない。

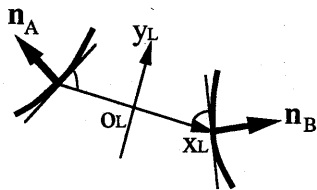


図2 2D局所対称によって定まる局所基準座標系

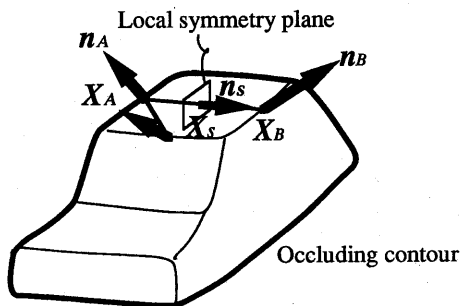


図3 局所面対称

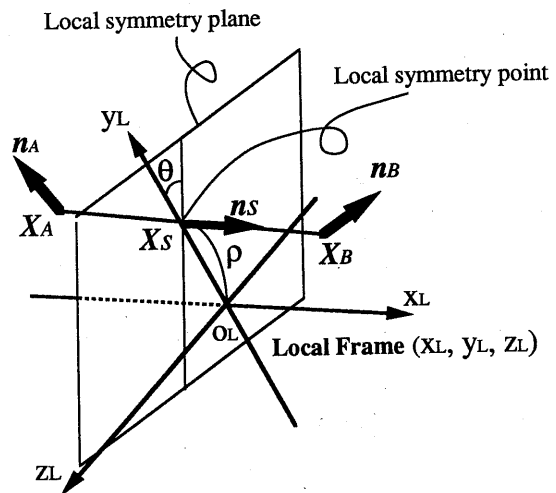


図4 局所対称面による局所基準座標系の拘束

精度よく局所対称面を検出するためには、十分な local support に基づく必要がある。そこで、遮蔽輪郭上の 1 点に対応する局所対称面を抽出するため、その点を中心とした遮蔽輪郭に沿った局所窓を考える。そして、局所窓内の各点と画像中の全点に対して対称条件を調べ、対称条件を満たす点の対から定まる対称面の平面パラメータを平面を表すパラメータ空間に投票する (Hough 変換)。次に、パラメータ空間中でピーク値を持つパラメータ値を初期値として、IRLS (Iterative Reweighted Least Squares) 法 [17] によって、精密化を施す。一般に、遮蔽輪郭の各点に対応する局所対称面は複数個存在する。IRLS 法による局所対称面の推定では、精密、かつ、非極大抑制 (推定された局所対称面に非常に近い“偽”の対称面の抑制) が行われるので、SLMS セグメントの抽出には、動的計画法や発見的探索等の追跡法は不要で、単に、遮蔽輪郭点を順にたどり、各点に対応づけられた局所対称面の中で、滑らかに結合可能な局所対称面を順に接続することによって抽出できる。

図 7 と図 8 に SLMS セグメントの抽出結果を示す。(a) はスリット光投影型のレンジスキャナ [7] によって獲得した 2 種類の電話受話器の距離画像の陰影表示である。(b) は遮蔽輪郭画像である。(c) は抽出された SLMS セグメントである。SLMS のセグメント上の各点は、局所対称面の法線方向を持っている。

3.2 SLGC プリミティブの抽出

今度は、SLMS セグメントの各局所対称面に対して、LGC の当てはめを行う。一般に、各局所対称面に対応する LGC は複数個存在する。SLGC プリミティブの抽出は、SLMS セグメントの抽出と同様の方法で、SLMS セグメントを順にたどってゆき、各局所対称面に対応づけられ

た LGC の中で、滑らかに結合可能な LGC を順に接続することによって抽出できる。現在のインプリメンテーションでは、滑らかに接続された各 LGC の z -軸に直線当てはめを行っているので、直線近似可能な軸をもつ SLGC プリミティブのみ抽出できる (湾曲した軸の抽出に対しても容易に拡張可能である)。図 7 と図 8 に対する SLGC プリミティブの抽出結果を図 9 と図 10 に示す。これらの例では、16 方向の θ に対して、切り口を生成し LGC の当てはめを行った。抽出された LGC の断面関数は B-スプラインで近似しており、不可視部分は対称性の拘束を用いて復元している。

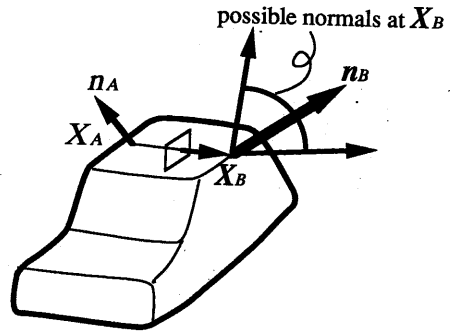


図 6 法線方向が急変する点での法線の割り当て

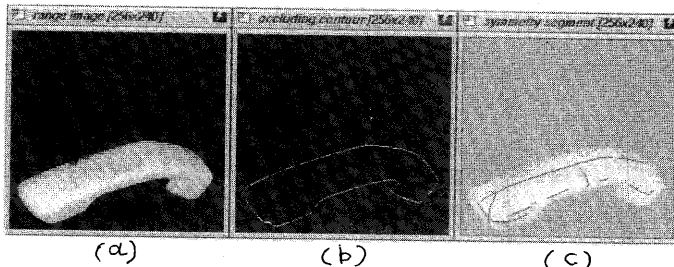


図 7 SLMS セグメント抽出結果 1

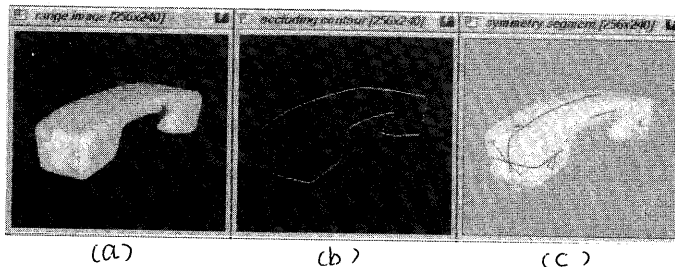


図 8 SLMS セグメント抽出結果 2

4 パーツ構造の復元

4.1 SLGC プリミティブの選択・統合の基準

3.2に示したSLGCプリミティブはパーツの存在に関する強力な証言を与えるが、併合すべきプリミティブや取り除くべきプリミティブが存在する。そのような選択・統合を行うための基準として、Bradyらは、滑らかな延長(smooth extension)、包括(subsumption)、非重複(non-overlapping)、および、連結(connected)などの基準を提案している[13], [15]。本研究では、それらの基準を一般化し、最適化問題として定式化する。

本研究での基準を以下に示す。

1. **パーツ形状の凸性 (Convexity):** 2つのSLGCプリミティブを併合したとき、どちらかのSLGCのz軸に関して凸形状となるなら、それらは1つのパーツに併合できる(現在のインプリメンテーションでは、直線軸のみを考えているので、この基準は、単に“凸形状”である)。この基準は、滑らかな延長と包括の2つの基準を含むものである。
2. **最小パーツ数 (Minimal-number of parts):** パーツの数は、できるだけ少ないほうがよい。即ち、単純な記述を優先する。

3. **データの最大被覆 (Maximal-covering):** 少なくとも1つのパーツによって被われるデータの数はできるだけ多いほうがよい。即ち、復元されたパーツによって、データは余すところなく説明できることが望ましい。

4. **パーツ間の最小重複 (Minimal-overlapping):** 2つ以上のパーツによって被われるデータの数はできるだけ少ないほうがよい。即ち、異なるパーツ間の重なりはできるだけ少ないほうが望ましい。

P をSLGCプリミティブのリストとする。 C を併合によって凸形状を構成することができるSLGCの対のリストであるとする。即ち、このリストに含まれる対は1つのパーツに併合可能である。 D を可能な分割を表すSLGCのクラスタのリストであるとする。ここで、分割 D の各クラスタ内のすべてのSLGC対は C に含まれていなければならない。いま、9個のSLGCからなるリスト $P_1 = (a, b, c, d, e, f, g, h, i)$ を考える(図11)。併合可能なSLGC対のリスト C_1 は $((a, b), (a, c), (a, d), (b, c), (b, d), (c, d), (f, g), (f, h), (f, i), (g, h), (g, i), (h, i))$ で与えられる。このとき、 $D_1 = ((a, b, c, d), (g, h, i))$ は可能な分割の1つである。分割 D_1 は、2つのパーツ (a, b, c, d) と (g, h, i) からなっている。

あるパーツ分割 D に対する目的関数 $F(D)$ は、パーツの数 $n_1(D)$ 、少なくとも1つのパーツによって被われるデータ数 $n_2(D)$ 、および、2つ以上のパーツによって被われるデータ数 $n_3(D)$ の線型結合として表現できる:

$$F(D) = -\epsilon_1 n_1(D) + \epsilon_2 n_2(D) - \epsilon_3 n_3(D). \quad (13)$$

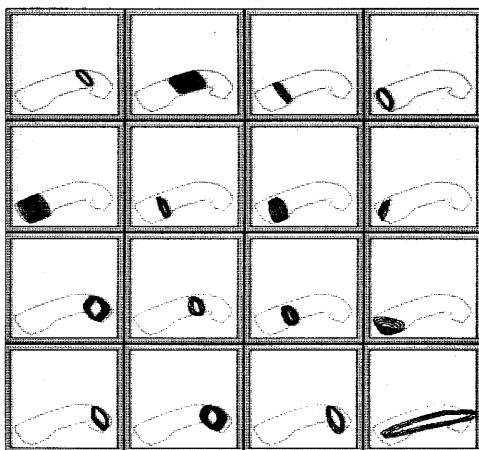


図9 SLGCプリミティブ抽出結果1

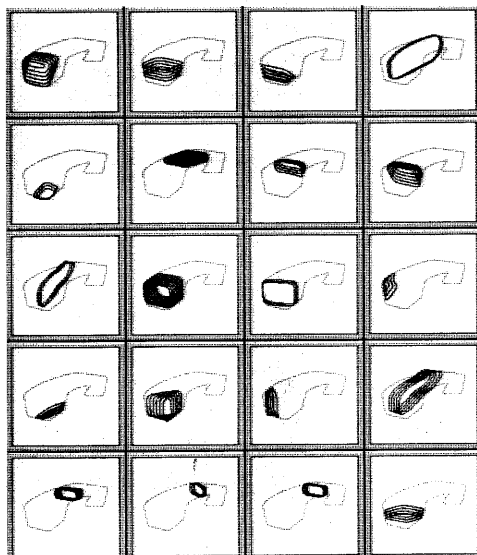


図10 SLGCプリミティブ抽出結果2

したがって、パーツ分割問題は、以下のように定式化できる：“SLGC プリミティブのリスト P 、および、併合可能な SLGC 対のリスト C が与えられ、分割 D の各クラスタ内のすべての SLGC 対が C に含まれるという制約のもとで、目的関数 $F(D)$ を最大化する SLGC のクラスタのリスト D を求めよ”。

重みパラメータ e_1 、 e_2 、および、 e_3 は、所望の感度に基づいて決定される。 e_1 が大きければ、小さいパーツは取り除かれる。これらのパラメータの調節は、認識処理等の後処理によって制御される。

4.2 可能なパーツ構造の探索

このパーツ分割問題を解決するために、組み合わせ最適化技法が必要となる。同様の最適化問題が、Hopfield ネットによって解かれている [5]。しかし、Hopfield ネットはコスト関数の極小値を検出できるのみであり、その極小値に対応する 1 つの分割が得られるのみである。即ち、可能なすべての分割は得られない。初期値を変えれば、異なる極小値を検出できる可能性があるが、それを組織的に行うことは難しい。さらに、本研究での問題は、SLGC プリミティブを取捨選択するだけでなく、SLGC の併合をも伴うので、0-1 計画問題のみを解く Hopfield ネットに埋め込むことは容易ではない。

ここでは、分枝限定法 (branch-and-bound method) による解決を考える。まず、問題を、SLGC リストにおける 1 番目の SLGC を選択する場合、および、そうでない場合の 2 つの場合に分割する。さらに、1 番目の SLGC を選択した場合は、2 番目の SLGC を選択し 1 番目の SLGC に併合する場合、2 番目の SLGC を選択するが 1 番目の SLGC に併合しない場合、および、2 番目の SLGC を選択しない場合の 3 つの場合に分割される。1 番目の SLGC を選択しなかった場合は、2 番目 SLGC を選択する場合としない場合の 2 つの場合に分割される。このように木探索が行われる (図 12)。ここで、併合可能な SLGC 対のリスト C によって、探索木の枝刈りが行える。また、探索木の各ノードにおいて、目的関数の値が、親ノードでの値より小さいならば、現在のノードにおける分割 D の部分リストを含む子孫ノードを探索する必要はない。被覆するデータの数の多い (即ち、大きい) SLGC プリミティブ

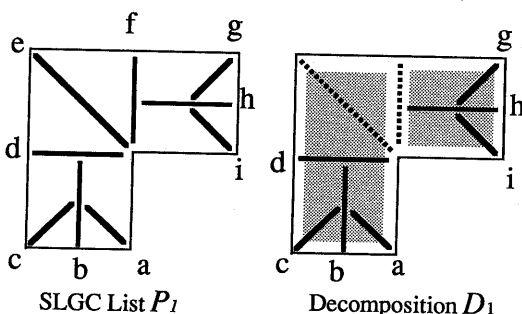


図11 SLGCプリミティブの選択・統合

ブから順に探索して行けば、これによる探索範囲の削減効果は非常に大きいと考えられる。よって、分枝限定法によって、十分に効率的な探索が可能である。

5 むすび

本稿では、距離画像からの 3D 形状の視点に依存しない分割記述の復元について述べ、暫定的な実験結果を示した。まず、局所一般化円筒 (LGC) の当てはめにより、3D 形状自体により定まる局所基準座標系を決定し、これらの局所座標系をグルーピングすることによって、体積的プリミティブ (SLGC) を形成する。以上の手法はインプリメントされ、実距離画像に対する実験により、その有効性を確認した。パーツ構造復元のためには、これらのプリミティブの選択・統合を行う必要があり、その基準を論じ、組み合わせ最適化問題として定式化した。そのインプリメントは現在行われている。

本手法の利点は、明確な基準に基づいて、パーツの位置決めを行い、組織的に可能な分割を探索するという点にある。従来、大局的パーツモデルを非線型最小 2 乗法を利用して強引に当てはめる方法と異なり、局所一般化円筒モデルを基に、信頼できる局所情報を集積していくことにより、もっともらしいパーツ構造の組織的な探索を可能にした。

本手法の問題点としては、インプリメンテーションの速度が上げられる。入力画像から、図 9 の出力まで、Silicon Graphics ワークステーション (20MIPS) において約 1 時間かかっている。しかし、いたるところで行われる局所対称面の抽出、および、局所一般化円筒の当てはめ処理はすべて独立な処理であるので、並列化が極めて容易であり、問題はさほど大きくないと考えている。

今後の課題としては、まず、復元形状の適用範囲の拡大があげられる。現在の方法では、2.2 におけるセンサデータの切り口は 1 つのパーツの断面であると仮定しているが、それが、2 つ以上のパーツからなる場合には、局所一般化円筒を当てはめる際、前もって切り口をセグメンテーションしておく必要がある。1 つのパーツの断面は凸形状であるといえるので、分割点として凹の極大点を用い

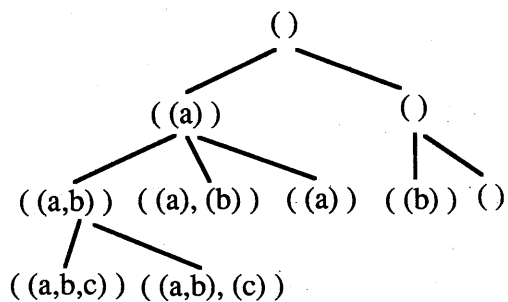


図12 最適分割の探索木

る方法が考えられる。さらに、復元された個々のパーツから形状特徴を表す有効なパラメータの抽出法も課題である。大規模モデルベースからの認識を考えたい場合に、特徴パラメータを用いたインデクシングにより効率よく物体モデルの候補を取り出すことができる。

謝辞：本研究をサポートしていただいた視覚情報研究部小林幸雄部長、および、石井健一郎グループリーダー、有益な助言を賜った尺長健主幹研究員、および、間瀬健二主任研究員、日頃計算機環境の面でお世話になっている渡部保日児主任研究員、および、距離画像入力に協力いただいた安野貴之社員に感謝致します。

参考文献

- [1] Grimson, W.E.L. and Lozano-Perez, T.: Model-based recognition and localization from sparse range and tactile data, *International Journal of Robotics Research*, 3, 3, 3-35 (1984).
- [2] Nevatia, R and Binford, T.O.: Description and Recognition of Curved Objects, *Artificial Intelligence*, 8, 1, 77-98 (1977).
- [3] Marr, D, *Vision*, San Francisco, Freeman (1982).
- [4] Pentland, A.P.: Recognition by Parts, *Proc. of ICCV '87* (1987)
- [5] Pentland, A.P.: Automatic Extraction of Deformable Part Models, *International Journal of Computer Vision*, 4, 107-126 (1990).
- [6] Ferrie, F.P., Deriving Coarse 3D Models of Objects, *Proc. of CVPR '88* (1988).
- [7] Sato, K. and Inokuchi, S.: Range-Imaging System Utilizing Nematic Liquid Crystal Mask, *Proc. of ICCV '87* (1987).
- [8] Zuk, D. and Dell'eva, M.: 3D Sensor for The Autonomous Land Vehicle, *Technical Report, Environmental Research Institute of Michigan* (1985)
- [9] Solina, F. and Bajcsy, R.: Recovery of Parametric Models from Range Images: The Case for Superquadrics with Global Deformations, *IEEE Trans., PAMI-12*, 2, 131-146 (1990).
- [10] Boulton, T.E. and Gross, A.D.: Recovery of Superquadrics from Depth Information, *Proc. of Workshop on Spatial Reasoning and Multi-Sensor Fusion* (1987).
- [11] Rao, K and Nevatia, R.: Computing Volume Descriptions from Sparse 3-D Data, *International Journal of Computer Vision*, 2, 33-50 (1988).
- [12] Ponce, J., Chelberg, D. and Mann, W.B.: Invariant Properties of Straight Homogeneous Generalized Cylinders and Their Contours, *IEEE Trans., PAMI-11*, 9, 951-966 (1990).
- [13] Brady, M.: Criteria for Representations of Shape, In Rosenfeld, A. and Beck, J. (Eds.), *Human and Machine Vision* Academic Press (1983).
- [14] Brady, M and Asada, H.: Smoothed Local Symmetries and Their Implementation, *International Journal of Robotics Research*, 3, 3, 36-60 (1984).
- [15] Cornell, J.H.: Learning Shape Description: Generating and Generalizing Models of Visual Objects, *MIT Technical Report 853* (1985)
- [16] Shafer, S.A.: *Shadows and Silhouettes in Computer Vision*, Kluwer Academic Publishers.
- [17] Besl, P.J., Birch, J.B. and Watson, L.T.: Robust Window Operator, *Proc. of ICCV '88* (1988)