

拡張ガウス像による3次元被写体の観測方向の決定

池内克史
(電子技術総合研究所)

1. はじめに

3次元被写体の2次元画像を解析する場合、中間結果として被写体の各点における面素の方向を得る手法がこれまで多く研究されてきた。^[1~5]この中間結果は、Z-1/ZD表現^[6]と呼ばれる。Z-1/ZD表現とは、被写体の各面素上における、法線単位ベクトルを画像上の対応する点に平行移動した単位ベクトルの集合である。たとえば、図1の(a)に示すようなゴルフボール上の小円形の中がみと、もとの小図形が円であるという仮定から図1の(b)に示されるようなZ-1/ZD表現が得られる。^[2]同様のアルゴリズムを用いて、2次元濃淡画像とその被写体の反射率地図を用いてZ-1/ZD表現を得ることも可能である。^[5]本論文ではこの種のZ-1/ZD表現が得られたとして、この情報をどのように利用するかを考える。

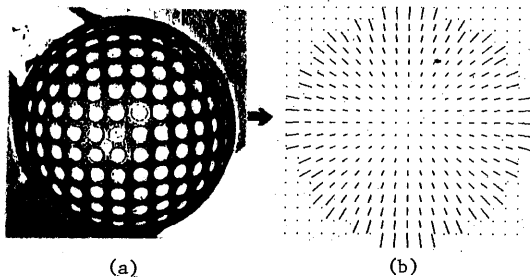


図1

図1の(b)に示された表現形は、観測者中心座標系により表示されている。すなわち、観測者の位置を中心に、その視線をZ軸にとり、それに直交する平面に被写体像を投影している。面素の法線方向も、視線方向を基準として表示される。これに反して、通常モデルは、そのモデルに固有の座標系——

たとえば、円筒ならばの軸方向と半径方向、楕円体ならば長軸、短軸方向といった座標系——によりモデル化されるのが自然である。この2つの座標系は互いに独立である。従って、この2つの座標系を関係づける手法が必用となる。

2. 拡張ガウス像 (EGI)

上記2つの座標を関係づけるものとして、拡張ガウス像という概念を導入する。本章では、まず拡張ガウス像の作り方をのべ、ついでその特性についてふれる。

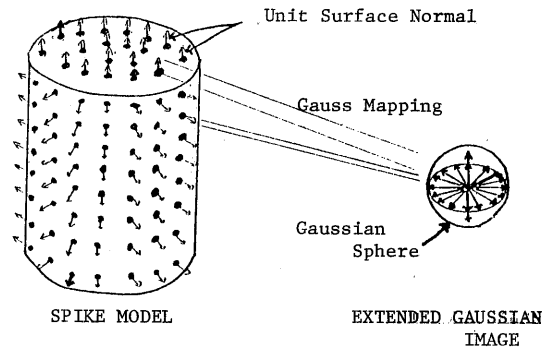


図2 Spike model と EGI

2. 1 拡張ガウス像の原理

モデル化しようとしている3次元物体の表面は、単位面積をもつ面素でおおわれていると仮定する。この面素上において面素の法線単位ベクトルが定義できる。この単位ベクトルの始点を3次元空間の原点へ平行移動する(この写像をガウス写像と呼ぶ)。結果として、それらのベクトルの終点は単位半径をもつ球(ガウス球)上にのることになる。曲率半径の大きい部分に対

応するガウス球上の領域では、ベクトルの存在密度が高く、曲率半径の小さい部分に対応する領域ではベクトルの存在密度が低くなる。このベクトルの終点がおのおの微小質量をもっていると仮定すれば、3次元物体の表面形状に対応してガウス球上にその物体に特有の質量分布が得られることになる。これをその物体の拡張ガウス像^[7,8]と呼ぶ。図2参照。

2.2 拡張ガウス像の実装

拡張ガウス像を計算機に実装するためにはまず、球面を一様に分割することが必要となる。すなわち、面素の法線方向が連続的に変化していくような面、たとえば、楕円面に対応するガウス球上の領域では、質量は連続的に分布している。従って、これをデジタルサイズするためには、球面を分割しなければならぬ。この分割は一様であることが望ましい。すなわち、各分割されたメッシュは、球面上に同一面積をもち、かつどの方向にも均一に存在して球面をおお、まなければならぬ。これは、その物体をどの方向から観測しても、モデルが同一の精度を保っているためである。

この条件を満たす一つの方法として、球に内接する正多面体の稜を球の中心から球面上に投影し、結果として得られる球面上の円弧によって球面を分割する方法がある。この方法から得られる分割された球面のことを測地ドーム (geodesic dome) と言う。正20面体以上の高次の正多面体は存在しないので、正多面体の各面をなるべく同一の面積をもちようとして3角形分割したのから得られる測地ドームを使用する。測地ドームについては、[9]に詳しい。本論文では、正12面体から第2次分割まで行なつたもの (2 frequency dodecahedron) を基本とし、球面を240の小球面3角

形に分割した。図3にこれを示す。なお、この構造内には各3角形間の球面上の距離も保持されている。

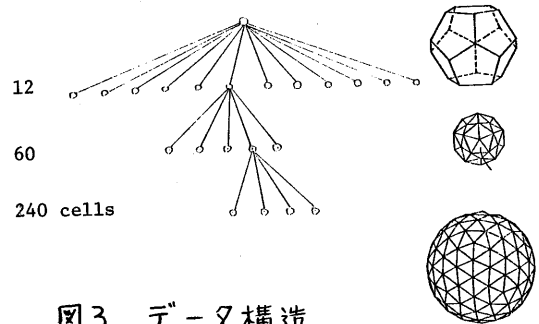


図3 データ構造

次に、測地ドームを利用して拡張ガウス像を実装する。実際、この測地ドームは、2つの目的に使用される。一つは、本来の目的、すなわち拡張ガウス像を保持するためである。この時、一つのメッシュ内の質量の総和が、メッシュの中心方向に割り当てられる。第2の目的は、モデル化しようとしている3次元物体を均一にサンプルするのに使用される。なぜなら、測地ドーム上の各メッシュは、中心からみても均一に分布している。従って各メッシュの方向をサンプルの方向とすることができる。

サンプル用の測地ドームの中心は、モデル化されている物体の重心と同一位置にあると、測地ドームの各メッシュの中心方向を延長して物体表面に達したとする。この時中心から物体表面までの距離を r 、中心方向と物体表面における法線(\vec{p})のなす角を γ とすれば $r/\cos\gamma$ が、拡張ガウス像を保持する測地ドームの \vec{p} に対応するメッシュに加えられる。図4参照。このようにして、全方向についてサンプルした後、拡張ガウス像を保持する測地ドーム上の値は、総和が1となるように正規化される。この正規化された測地ドーム上の質量分布がその物体の拡張ガウス像となる。

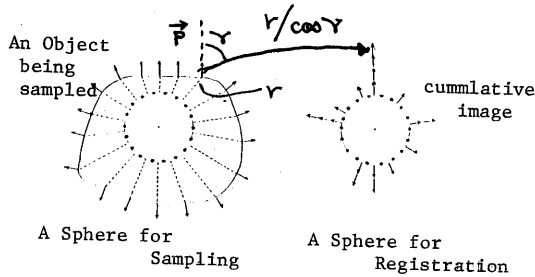


図4 サンプル方法

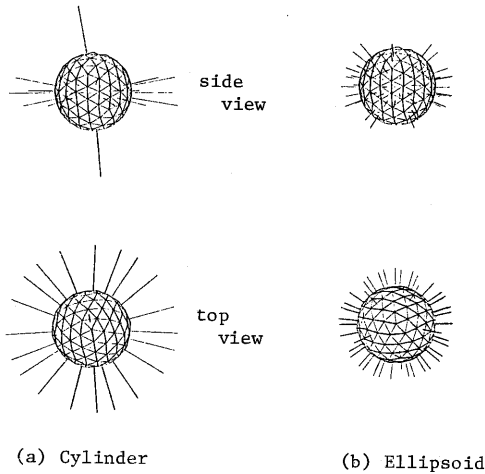


図5 EGIの例

2.3 拡張ガウス像の特性

本論文で、拡張ガウス像を利用して3次元物体を認識することを提案する理由は、(1)拡張ガウス像が座標系に多くの点で独立であるということ、(2)拡張ガウス像の一致が2つの凸体の合同条件である、という二点である。まず(1)の理由の利点について考える。入力画像から得られる2-1/2D記述は、観測者中心座標系に基づいており、被写体のモデルは対象物中心座標系に基づいている。この両者は、全く関係がない。しかし、両者の情報をいったん座標系に独立な拡張ガウス像とし、モデルと画像のマッチングを取れば2座標系の差異を意識しなくてよい。座標系は、①原点の位置、②座標軸のスケール、③座標軸の方向の3特性で記述できる。この内、拡張ガウス像は前二者

に独立である。まず、拡張ガウス像が座標系の原点の位置に独立な理由は、被写体の3次元モデルから拡張ガウス像をつくる際、面素上の法線ベクトルは、その位置に無関係にガウス球の中心に平行移動される。従って、モデルが記述されている座標系の原点がどこにあっても常に同一の像が得られる。次に拡張ガウス像が座標系のスケールに独立である理由を述べる。座標軸のスケールの大小に応じて、被写体の3次元モデルの表面積は変化する。これにともなう、ガウス球上の質量の総和も変化するようになる。しかし、拡張ガウス像をつくる際には、総和が1になるよう正規化してあるため、座標軸のスケールには独立となる。たとえば、一辺1cmの立方体も2cmの立方体も常に、北極、南極、および赤道上の4点にそれぞれ1/6の質量をもつ拡張ガウス像として投影される。この特性は、物体認識上非常に有利である。一般にTVカメラと被写体との距離は、状況に応じて変化し、被写体の大きさも変化するにも関わらず、それから導出される拡張ガウス像は同一物体に関しても、常に同一だからである。第③の特性、すなわち対象物座標系の方向が変化すれば拡張ガウス像もガウス球上でそれにつれて回転する。しかし、両者の回転様式は一致しており、拡張ガウス像の相対的位置関係は変化しない。丁度、被写体は拡張ガウスをもつガウス球でおおわれており、被写体の回転に応じて、そのガウス球も回転すると考えよう。従って、拡張ガウス像を用いて観測者中心座標系上の記述と対象物中心座標系上の記述とのマッチングを取る際、この両者の回転方向の一致を取る必要がある。換言すれば、被写体をどの方向から観測したかを知る必要がある。これについては次章で展開する。

次に、(2)について述べる。2つの拡張がウス像が同一であることは、2つの凸体が合同であるための必要十分条件^[10]であり、一般の物体に関しては、必要条件である。従って、凸体に関しては、モデルの拡張がウス像と画像からの拡張がウス像との一致をみるだけで、観測者の観測方向が定まるだけでなく、被写体などのモデルに対応するかということも定まってしまう。一般の物体に関しても、多くの場合候補者内の物体の拡張がウス像が異なっているため、物体の認識が可能である。たとえ、候補者内の2つの物体が全く同一の拡張がウス像をもつ場合でも、本論文に提案する方法を使用して、観測者の観測方向が決定できる。この場合いったん観測方向が定まってしまうとモデルの他の情報——例えば、拡張がウス像が2つの座標系の差を埋めるために放棄した位置情報——を使用して最終決定する場合も、はるかに比較が容易となる。

3. 観測者の観測方向に関する拘束

一般に、観測者が写体をながめる方向にはる自由度がある。まず空間上のある地点に被写体を置いたとする。この被写体にはがウス球が付随している。どの方向からこの被写体をながめるか、すなわち視線方向ががウス球上のどの地点と直交するかについて、2自由度

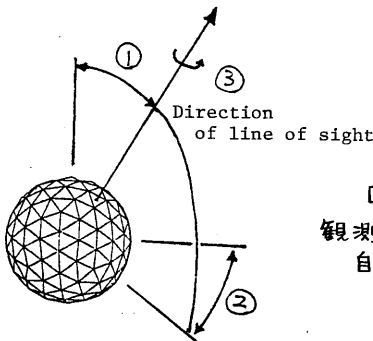


図6
観測方向の
自由度

ある。次に観測者はこの視線を軸として回転できる。この自由度は1ある。図6参照。視線方向の自由度、視線を軸とする回転の自由度の各々について、自由度を減少させる拘束を考える。

3. 1 視線方向に関する拘束

表面積に対する断面積の比率は、視線方向に対する拘束になる。例えば、楕円を長軸の方向からながめた場合、この比率は他の方向からながめた場合より、小さいものとなる。

この比率は、拡張がウス像において、

$$A_v = \sum_i z_i m_i / \sum_i m_i, \quad (1)$$

と書ける。 m_i は、がウス球上の i 番目のメッシュ内の質量、 z_i は、観測者の観測方向(\vec{N}_v)とメッシュ i があらわす方向(\vec{N}_i)との余弦である。すなわち、 m_i がメッシュのあらわす方向 N_i の方向をもつ面素の面積をあらわすから、 $m_i z_i$ はその面素の視線方向に直交する面(画像面)への投影面積を表わす。総和は、視線方向 N_v から可視の半球についてとられる。従って、(1)式の分子が断面積、分母が表面積に対応することになる。図7参照。

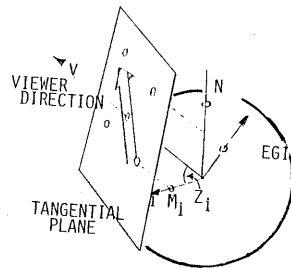


図7 EG I と画像面

一方、観測画像においては、断面積は面素の方向が得られた画素数を n とすると、 n 。表面積は各画素 i における得られた面素の方向と画面の法線方向とのなす角を $\cos \alpha_i$ とすると、

$$\sum_{i=1}^n 1/\cos \alpha_i$$

で求められる。ゆえに上記の比率は、

$$A_v = n / \sum_{i=1}^n 1 / \cos \alpha_i \quad (2)$$
 とする。この比率 A_v は、丁度反射率地図に基づいた時の濃淡情報^[2]のようにガウス球上における観測者の視線に関する自由度を1つ減らす。

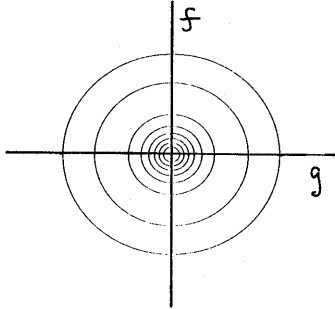


図8 投影率地図(極投影平面)

図8は例として楕円体におけるこの比率をガウス球上で求め、それを極投影法で平面に投影したものである。図8の各軌跡は、同一の比率をもつ点を表わす。この楕円体は、Z軸方向が長軸、X軸とY軸方向が同一長さの短軸となるものをとったので、極投影平面では、等比率の軌跡は同心円状となり、原点に近い点すなわち長軸方向からながめると断面積/表面積の比は小さく、原点から離れるにつれ、大きくなることを示している。

3. 2 視線を軸とする回転に関する拘束

視線を軸とする回転の自由度を拘束することを考える。拡張ガウス像の慣性モーメントの方向を利用する。まず、ガウス球上の各セル i 内の質量を m_i 、このセルを視線に直交する平面上に投影した点の平面における座標を (x_i, y_i) とする。

$$\begin{aligned}
 I_{xx} &= \sum_{i=1}^n m_i x_i^2 \\
 I_{yy} &= \sum_{i=1}^n m_i y_i^2 \\
 I_{xy} &= \sum_{i=1}^n m_i x_i y_i \quad (3)
 \end{aligned}$$

とし、総和は視線方向を極とする可視の半球上にあるセル i についてのみとられる。これらの I_{xx} , I_{xy} , I_{yy}

から

$$\alpha = \left\{ \tan^{-1} 2 I_{xy} / (I_{yy} - I_{xx}) \right\} / 2$$

なる α は、視線に直交する面上における2次慣性モーメント最小の方向となる。観測された画像面においても、この方向が求められる。従って、両者が一致するように、ガウス球を回転させればよいこととなる。図9は各セルが視線方向になつた場合の2次慣性モーメント最小の方向を、針の方向で示している。

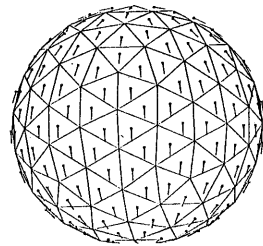


図9 2次慣性モーメント最小の方向

たいていの被写体の大部分の視線方向に関しては、この方向は一意に定まるので、視線を軸とする回転の自由度は、完全に拘束されることとなる。

3. 3 ガウス球上の情報

最終的に、ガウス球上の各セルは、3種類の情報を含んでいることとなる。まず第1は、各セルにおける拡張ガウス像の質量、この質量分布によって最終決定が行われる。第2に、そのセルが視線方向になつた場合の断面積/表面積の比率。これによって、視線方向の自由度が拘束される。第3に、そのセルが視線方向になつた場合の2次慣性モーメント最小の方向。これによって視線を軸とする回転が完全に決定できる。

上記の第2、第3の拘束は、被写体が他の物体にさえぎられているような場合にも適用可能ないように変形するこ

とができる。式(1)、(3)を用いてモデルとなる拡張ガウス像から、投影率、2次慣性モーメント最小の方向を求める際、視線方向を極とする半球の上で総和をとった。これを、部分観測にも成立するように変形するには、総和がとられる範囲をもう少し小さくすればよい。上記の投影率、2次慣性モーメント最小の方向は、視線方向に対応する点とそれを極とするガウス球上の領域——この領域を示す量として視線方向と領域周辺部分に対応する方向のなす角をみこみ角と呼び、 ω と記す——との関数になっている。例えば、投影率の場合この領域として、半球全体($\omega = \frac{\pi}{2}$)をとったものが、式(1)の拘束条件となり、1つの点に縮退した場合($\omega \rightarrow 0$)、ある方向のセル内の質量の有無を判別する条件になる。この中間の各種のみこみ ω により、各種の拘束条件の定式化が可能となる。図10に、みこみ角 ω による投影率の変化の例を挙げる。

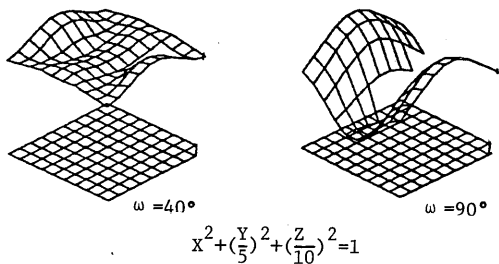


図10 みこみ角による投影率の変化の例

各セルについて、それが視線方向と合った時の投影率、2次慣性モーメントの方向以外に代表的のみこみ角、例えば 80° 、 70° 等の領域における投影率、2次慣性モーメントを計算して、記憶しておく。一方、観測された面積の方向をガウス球面にプロットすることにより、点の分布具合から視線方向を中心にとれただけのみこみ角に対応する領域が観測されたのかかわかるので、そ

のみこみ角について有効な拘束を使用すればよい。ただし、重ねている部分が平面の時や凹領域の場合は、この部分領域による拘束は使用できない。

3. 4 判別関数

観測された面素の方向の分布とモデルとの比較は次の判別関数を用いて行われる。

```

if  $M_i^{Proto} = 0.0$ 
  do nothing
if  $M_i^{Proto} \neq 0$ 
   $\Delta = M_i^{Proto} - M_{i+e}^{observe}$ 
   $d = \text{distance}(i, e)$ 
  if  $d < d^0$  and  $\Delta < \Delta^0$ 
    add  $e^{-\lambda_1 d - \lambda_2 \Delta}$  to
    total point
  
```

M_i^{Proto} はモデルの*i*番目のセルの質量、 $M_{i+e}^{observe}$ は観測された面素の方向をガウス球上に投影した時の*i*番目のセルの質量、 λ_1 、 λ_2 は係数である。上式の意味は、モデルの各セルを順々にサーチして行き、あるセルにおいて、質量 M_i^{Proto} があったとすると、その近傍*i+e*をサーチし、 M_i^{Proto} と同じような質量をもつセルが存在するかどうかを順々にチェックする。同じような質量 $M_{i+e}^{observe}$ が存在した場合には、その類似度*d*と距離 Δ に応じてその観測方向に対する得点に $e^{-\lambda_1 d - \lambda_2 \Delta}$ が加算される。この判別関数は、各モデルの拘束条件を満たした観測方向についてのみ適用される。この判別関数の最も高い点を得たモデルの観測方向が最も確からしいモデルの観測方向ということになる。

4. 実験例

拡張ガウス像によるマッチングの有効性をみるために、次のような実験を行った。 $x^2 + y^2 + \left(\frac{z}{5}\right)^2 = 1$ で表わされる楕円体がZ軸から 10° 傾いた方向の

ら観測されたとする。まずこのシステムに入力することによる $2-1/2$ 次元記述を生成した。各面素の方向を針で表わしたものを図11に示す。

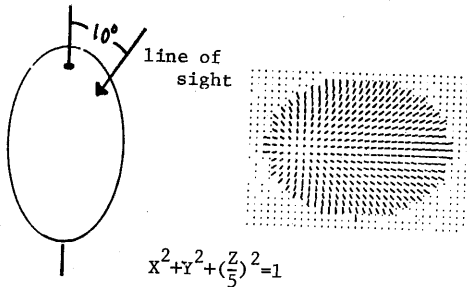


図11 楕円体と $2-1/2$ 次元記述

第一の仕事は、与えられた面素の方向分布から拡張ガウス像を求めることである。各面素は、十分に小さいと仮定する。まず面素の方向からそれがガウス球上のどのセルに対応するかが決定される。次に、面素の微小表面積がそのセルに加算される。この微小表面積は、 $(1/\cos\alpha)$ に比例する。従って、この値がセルに加算される。ただし、 α は面素の方向と Z 軸のなす角である。全画素についてこの操作をほどこれた後、ガウス球上の質量は総和として、 $1/2$ の値をもつように正規化される。図12に、このようにして図11から得た拡張ガウス像を示す。

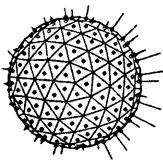


図12 観測された EGI

拡張ガウス像を求める際に、同時に投影率、2次慣性モーメント最小の方向も求めておく。領域周辺では、ノイズの影響が大きいのでみこみ角 W が 80° 以内となる画素上の点についてのみ上記の投影率、2次慣性モーメントを計

算する際対象とした。3.3節でも述べたように、投影率、2次慣性モーメントの方向は任意の、みこみ角について成立するので、これが可能である。投影率として、0.27、2次慣性モーメントの最小の方向として、Z 軸方向を得た。

今回は、候補となるモデルとして、上記の楕円体と円筒を用意した。円筒の半径は 1.0、高さは 3.0 である。両者の拡張ガウス像を図13に示す。

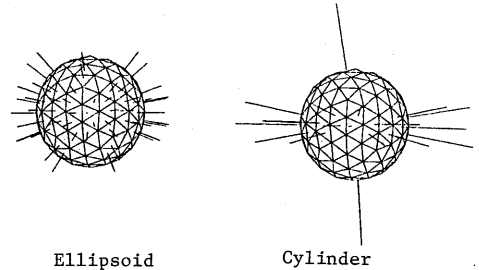


図13 モデルの EGI

投影率 0.27 ± 0.1 の拘束条件より、240 の可能な視線方向のうち楕円体の場合 20 方向に、円筒の場合 42 方向に拘束された。両者の物体とも、上記の各セルにおいて、2次慣性モーメント最小の方向を一意に持ち、ているので、視線を軸とする回転の自由度は完全に拘束される。

楕円体の場合、セル 16 (そして、16 の反対側にある 220) において最も高い点数 6.59 を与え、セル 0、セル 8 などが次点で、5.2 であった。一方、円筒の場合、セル 0 が最大で、1.6 だった。従って、楕円体のセル 16 の方向が、観測方向として、最も確からしい、ということになる。このセル 16 は、 $(0.1426, -0.1036, 0.9843)$ の方向に対応するので、頂角 $10^\circ, 15^\circ$ の方向から観測されたと言う、ほぼ正しい解が得られたことになる。

5. まとめ

生データから得られる $2-1/2D$ 表現を解析する手段として、拡張ガウス像を使用することを提案した。 $2-1/2D$ 表現は観測者中心座標系に基づいているのに対し、物体のモデルはこれと独立な対象物中心座標系で表現される。この両座標系間の差を乗り越えるために、座標系に独立な拡張ガウス像に一度投影し直して観測者の観測方向を見つけ出そうとするものである。まず各候補者について最も確からしい観測方向を見つける。次に各候補者間で最大の確からしさを得た候補者の観測方向が、解として採用される。

観測者の観測方向は、3自由度ある。視線方向に2自由度、視線を軸とした回転に1自由度である。画像面から得られる断面積/表面積なる比率から視線方向の自由度を1つ減らすことができる。また、画素面における拡張ガウス像の2次慣性モーメント最小方向から、視線を軸とする回転の自由度を完全に拘束することができる。可能な観測者の観測方向を上記の拘束を用いて減小させた後に、観測方向の最終決定は、モデルと実際に得られた拡張ガウス像の分布の類似度を評価して行われる。

候補者内の2つの凹領域を含む物体が、たまたま同一の拡張ガウス像をもつ場合がある。この場合も、本手法を使用して、観測者の観測方向を決定した後に、拡張ガウス像が2つの座標系の差を埋めるために放棄した位置情報を使用し、最終決定をすればよい。この場合、観測者の観測方向が限定されてしまっているのだから、拡張ガウス像を用いにくい場合より、はるかに有利となるのである。

《謝辞》

熱心に御検討頂いた、電子技術総合研究所、白井視覚情報研究室長、同研

究室員の方々、並びにMITのHorn教授に、深く感謝致します。

REFERENCES

- [1] Ikeuchi, K. and Horn, B.K.P. (1981), "Numerical Shape from Shading and Occluding Boundaries," Artificial Intelligence, Vol. 17, 141-184.
- [2] Ikeuchi, K. (1980), "Shape from Regular Patterns: An Example of Constraint Propagation in Vision," Proc. ICFR-80, Miami Beach, Dec. 1980, 1032-1039.
- [3] Ota, Y., Maenobu, K., and Sakai, T., (1981) "Obtaining Surface Orientation from Textels under Perspective Projection," Proc. IJCAI-81, Vancouver, Canada, Aug., 1981, 746-751.
- [4] Witkin, A.P. (1981), "Recovering Surface Shape and Orientation from Texture," Artificial Intelligence, Vol. 17, 17-45.
- [5] Koshikawa, K. (1979), "A Polarimetric Approach to Shape Understanding," Proc. IJCAI-79, Tokyo, Aug., 1979, 493-495.
- [6] Marr, D. and Nishihara, K.H. (1978), "Representation and Recognition of the Spatial Organization of Three-dimensional Shape," Proc. R. Soc. Lond., B:200.
- [7] Horn, B.K.P. (1981), "Sequills & Quills - Representations for Surface Topography," in: Bajcsy, R. (Ed.) Representation of 3-Dimensional Objects, Springer, Berlin.
- [8] Smith, D. (1979), "Using Enhanced Spherical Images for Object Representation," MIT-AI Memo, No. 530, MIT-AI Lab, Cambridge, Mass.
- [9] Wenninger, M.J. (1979), Spherical Models, Cambridge University Press, Cambridge, UK.
- [10] Lysternik, L.A. (1963) Convex Figures and Polyhedra, Dover, New York

