

照度差ステレオとプリズムステレオによる ビンピッキングタスク中のグリップ姿勢決定

正員 池内 克史[†] 正員 長田 茂美^{††}
非会員 ホーン B.^{†††} 非会員 ニシハラ K.^{†††}

Determining Gripper Configuration in Bin Picking Tasks
Using Photometric Stereo System and PRISM Stereo System

Katsushi IKEUCHI[†], Shigemi NAGATA^{††}, Members, Berthold HORN^{†††}
and Keith NISHIHARA^{††††}, Nonmembers

あらまし ビンピッキングのための目を構成した。照度差ステレオを用いて、シーンのいがり表現像を得る。得られたいがり表現像中に、部品の重なりを示す。影、輪郭線、面素の不連続部分に対応する場所を検出し、領域分割する。各連結領域中最適な領域を一定の評価関数に従って決定する。最適領域上での面素分布より、拡張ガウス像を得る。計算機中に蓄えられた部品の各姿勢における拡張ガウス像と得られた拡張ガウス像を比較することにより部品の姿勢を決定する。部品の幾何モデルを利用して決定された姿勢下での安定把握姿勢を決定する。一方、プリズムステレオにより深さ分布図を得る。これから、各安定把握姿勢下で、グリップが他の部品と接触しないかをチェックし、他の部品から最も遠い姿勢を最適把握姿勢とし、この姿勢を手にする。このシステムを二台のLispマシン上に実装し、ビンから部品をつかみ上げることができた。

1. ま え が き

知能ロボットにとって、感覚器の保持は一つの不可欠な能力である。外界の情報収集がなければ、制御はオープンループとなり、作業は予め決められたものしか実行できない。感覚器を持たないシステムは外乱に弱く、フレキシブルでもない。

各種の感覚器の内、視覚器は最も重要な機能の一つであろう。視覚は非接触で情報を収集できるため、環境への攪乱が少ない。また、環境の状況を大づかみにできるという点に特色がある。

一般に、知能ロボットのタスクは次の三要素に細分

できる。

- (1) 対象物を探し出し、つかみ点を決定する。
- (2) つかみ上げた物体を目的地まで他の物体との衝突を回避しながら運ぶ。
- (3) 運ばれてきた物体を使って、組み立て作業を行なう。

各サブタスクにおいては、知能ロボットと外界との拘り合いは少しずつ異なっている。従って、視覚のはたすべき役割りも少しずつ異なっていく。本論文では、(1)のサブタスクにおいて視覚のはたすべき役割りについて考えてみたい。

(1)のサブタスクにおいて視覚がない場合、ロボットは一定位置に一定姿勢で置かれたものしかつかめないことになる。ここでは、視覚を用いてこの欠点を解決することを考える。視覚を用いた欠点解決における未知要素は、被写体の位置、種類、姿勢の三つである。典型的な例として、部品の山からのつかみ上げが挙げられる。部品の山のことをビンと呼ぶ。ビンからのつかみ上げの場合、特に位置と姿勢の決定が重要になる。従来ビンピッキングは、明暗情報の変化する線(エッ

[†]電子技術総合研究所制御部, 茨城県
Electrotechnical Laboratory, Ibaraki-ken, 305 Japan

^{††}(株)富士通研究所, 川崎市
FUJITSU LABORATORIES LTD., Kawasaki-shi, 211 Japan

^{†††}マサチューセッツ工科大, USA
Massachusetts Institute of Technology, Cambridge, MA,
02139 USA

^{††††}フェアチャイルド人工知能研究所, USA
Fairchild Laboratory for AI Research, Palo Alto, CA,
94304 USA

ジ)を抽出し、このエッジを処理することで進められてきた^{(1),(2)}。エッジを利用してシーンを小領域に分割する。各小領域の中に既知の形を探し出し、各小領域を意味づける。この情報を基に、位置姿勢を決定するものである。

エッジに基づく手法では、次の二点で困難を伴う。

(1)重なり合った三次元物体間は、明暗情報の変化が少ない。故に、単純なフィルターではエッジが抽出しにくい。(2)三次元物体は観測する方向に応じてその形が変化する。この各方向における形を蓄えるのでは大変である。また、ピンピッキングの場合、三次元物体が重なっているため、元の形を復元するのが難しい。

Kellyらは、この二つの問題を避けるため金属円柱状の被写体に限定して、表面に発生する正反射の帯(ハイライト)を利用した⁽³⁾。このハイライトから領域分割を行なったのである。しかし、被写体の姿勢は完全には決定できない。そこで、この欠点をグリップの工夫により回避している。

近年、三次元情報(2-1/2D)をシーンから直接抽出する方法が数多く開発されてきた⁽⁴⁾。濃淡情報を利用するもの⁽⁵⁾、形のみかけのひずみを利用するもの⁽⁶⁾、偏光を利用するもの⁽⁷⁾、レンジファインダー^{(8),(9)}、ステレオ⁽¹⁰⁾などである。本稿では、近年開発された濃淡情報を利用する手法と2眼ステレオ法を利用して、ピンピッキングの目を構成することを考える。

2. 問題の設定

本ピンピッキングシステムは次のような仮定を置く。

- (1) 全て同一の部品が山積みされている。
- (2) 部品の材質が既知である。
- (3) 部品の形が既知である。

これらは、工業部品供給の場合には、容易に満足される条件であり、この仮定によりシステムの一般性は失われない。

3. システムの概要

本システムは、図1に示されるように8段階の処理よりなる。

- (1) 照度差ステレオを用いて、シーンのいがぐり表現像を得る。
- (2) 得られたいがぐり表現像により、領域分割する。
- (3) 一定の評価関数に従って、最適な領域を得る。
- (4) 最適領域内の面素分布を照度差ステレオでより精密に再計算する。

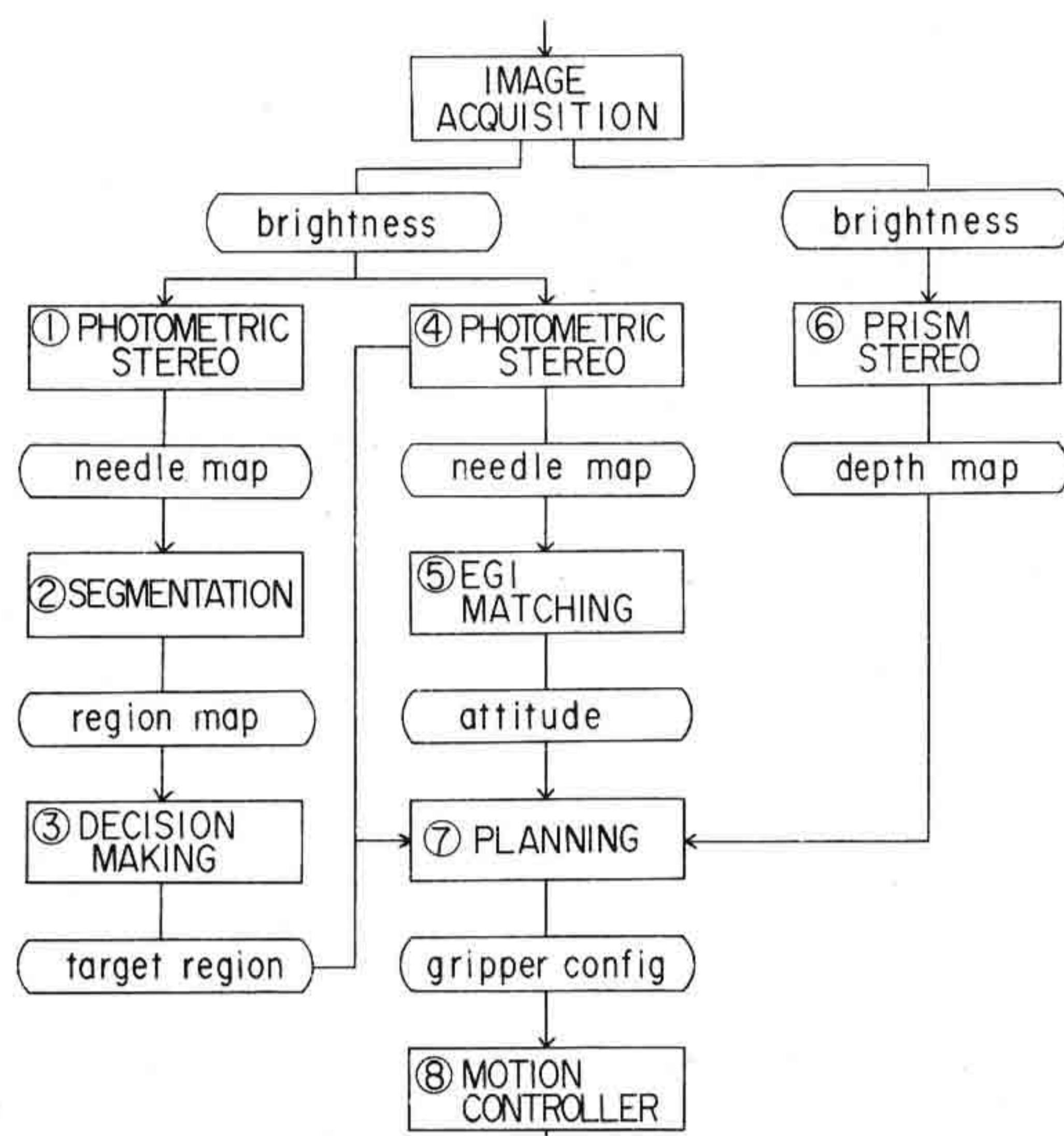


図1 システムの概要

Fig.1 Out line of the system.

- (5) 得られた画素分布から拡張ガウス像を求め、これと計算機内の拡張ガウス像を比較することにより、部品の空間中における姿勢を決定する。
- (6) プリズムステレオの出力から対象部品の周辺の状態を得る。
- (7) (5), (6)の情報から部品の安定把握姿勢を決定する。
- (8) 得られた位置姿勢情報をもとに部品をつかみ上げる。

3.1 照明度ステレオ

同一シーンを光源の方向のみを変化させ、同一位置より写した複数枚の画像から、物体各点の面素の方向を得る手法を照度差ステレオと呼ぶ^{(1)~(5)}。ここでは、3方向の光源下における画像から面素の方向を得る。

照度差ステレオは、各画像間に視差が存在しない。これは、各画像の同一位置の画素どうしが対応するからである。このためマッチングが非常に高速に行なえる。さらに、予め明るさの3つ組と面素の関係を計算し、3つ組から面素の方向が参照できる表を計算機に蓄えておく。この参照表法を使用すれば、実時間処理は表の読み出しのみとなる。

参照表は理論的に求めることもできるが、ここでは実験的に求める。まず、被写体と同一の反射特性を持ち形状が既知(通常は球形)の物体を光源下に置く。3光源と補助光源を利用して、画像から球の輪郭線を抽出する。この輪郭線と物体の形状から、画素に対応する面素の方向が計算できる。

次に、光源を1つずつ点灯し、3枚の画像を取り込む。球上の各画素について、明るさの3つ組に対応する参照表の位置に、その画素に対応する面素の方向を記入する。最終的に各要素について平均を取る。このようにして3次元参照表が得られる。

3.2 領域分割と評価関数

3.2.1 領域分割のための知識

多数の部品が重なった場所から、1つの部品を切り出すためには、シーンを領域に分割する必要がある。本稿では、3次元物体の重なりと照度差ステレオの出力の関係から導出される知識を利用して領域分割する。

(1) 影

3次元物体が重なり合った場所に光を投げ掛けた場合、上に存在する物体は、下に存在する物体に影を投げ掛ける。従って、多数の方向から光を投げ掛けると影部によって各物体が分離される。照度差ステレオは、3方向から光を投げ掛けるため、影の部分を検出すれば良い。すなわち、照度差ステレオへの入力を一定の域値で切ればよい。

(2) 輪郭線

物体が重なり合った場合、各物体の境界は輪郭線で区切られる。輪郭線上では、面素の方向は視線方向と直交する。今、視線方向と面素の方向とのなす角を見込み角と呼ぶ。見込み角が 90° に近い所を検出すれば、各物体の分離が可能となる。特にこの手法は、曲面体に有効である。照度差ステレオの参照表において、ある程度以上の見込み角を持つ面素が格納されている場所を面素の求まらない場所に変換することでこの知識が実装可能となる。

(3) 面素の不連続性

一般に、2つの物体が重なった線の両側では、面素の方向は不連続となる。従って、面素が不連続になる場所を検出すれば、2物体を分離することができる。不連続の測度としては、なめらか拘束を使用した。

(1), (2), (3)を利用して、領域分割を行なう。今後の処理では、画像上の(1), (2), (3)の領域を背景とし、0の値を持たせる。その他の領域、すなわち照度差ステレオにより面素の求まった画素では、1の値を持たせた2値画像を構成する。(1), (2)の知識は、照度差ステレオ参照表に埋め込むことができる。(1)と(2)の領域に対応する面素を参照表から間引くのである。この参照表を用いて照度差ステレオを実行し、面素の求まった場所では1、求まらなかった場所では0と置く操作を行えば、求める2値画像が、いがり表現像と同時

に得られる。

次に、この2値画像の連結領域を調べてラベル付ける。すなわち、各連結領域ごとに、異なった番号を割り振るのである。これは、 2×2 ウインドウについて、通常のラベル操作を画面右上から順次行なって行く。ただし、このラベリングに際して(3)の知識も併用する。すなわち、不連続性が大となる面素を発見すると、その面素を0に変換する。

3.2.2 対象についての知識

複数の連結領域から特定の領域を選ぶためには、部品に対して固有な知識を利用するのが有利である。このため、対象を設定する必要がある。本稿では、対象としてドーナツを使用する。

一般に、横になったドーナツほどつかみやすい。横になるほど観測される領域が広がる。同時に、ドーナツの穴も見えるようになる。下層にあるドーナツは、上層にあるドーナツに隠されて観測される領域が減少する。従って、上層のドーナツに対応する領域を、穴の有る領域で最大のものとし、これを切り出す。

領域の面積は、単に同一ラベルを振られている画素の数を数えれば良い。また、領域に穴があるか否かは、オイラー数を調べればよい。ただこのオイラー数はノイズに影響されやすいので、オイラー数を調べる前に、各連結領域ごとに膨張処理を行なう。

最適領域は下記の手順で選ばれた。

(ステップ1)

* 最大から順に5つの大きな連結領域を選ぶ。

(ステップ2)

* 2つ以上の領域がオイラー数0の場合、その内最大のものを選ぶ。

* オイラー数0の領域が唯一の場合、それを選ぶ。

* オイラー数0の領域がない場合、最大の領域が選ばれ、拡張ガウス像マッチングシステムに穴なしの情報を送られる。

3.3 プリズムステレオ

照度差ステレオとプリズムステレオは互いに相補な関係にある。照度差ステレオは高速に、精密に画素の方向が決まる。しかし、絶対深さは定まらない。一方プリズムステレオは、高速に絶対深さが定まる。しかし、従さの求まる密度が粗い。

照度差ステレオの出力を用いて、領域分割、目標領域の切り出し、姿勢決定を行なうと共に、プリズムステレオを用いて、他の物体と衝突しない目標点の決定と目標点の深さ情報の決定を行なう。

プリズムステレオは Nishihara らによって実装された、2眼ステレオの一種である Marr-Poggio ステレオのハードウェア版である。まず左右の画像にガウシアン分布の2次微分フィルタ(これをDOGと呼ぶ)をかけ、濃淡画像の変化の大きい所の検出を行なう。これは、ゼロクロッシング線(ZC線)と呼ばれる。ZC線の抽出までは、Lispマシンに実装された専用ハードウェアで行なう。

左右の画像上のZC線の対応づけにより、視差を計算する。正確には、ノイズの影響を避けるためフィルタ出力の同一符号領域の比較を行なう。DOGフィルタの幅とZC線の出現間隔の関係を利用して、マッチングの探策域を限定する。幅の広いフィルタの処理から始め、狭いフィルタの処理へと移り、深さ情報を精密化して行く。

Marr-Poggioのステレオは、ZC線上でしか深さ情報が定まらない。Nishiharaらは、画像上のZC線の数を増加させるために、ランダムドットを投げ掛ける。このあと、画像を小区画に分割し、その内部に存在するZC線上の深さ情報の平均を取り、これをその小区画の深さとしている。従って、出力は2次元の棒グラフのようになる。これが、プリズムステレオと呼ばれる所以である。

3.4 拡張ガウス像

ある物体の拡張ガウス像とは、その物体表面の面素の空間ヒストグラムである⁽¹³⁾。まず、物体表面が同一面積の微小面で覆われていると仮定しよう。この面素上において、面素の単位法線ベクトルが定義できる。各法線ベクトルの視点を座標原点に平行移動する。各ベクトルの視点はガウス球上に存在する。各終点を質点と見なせば、ガウス球上に1つの質量分布が定義できる。これをその物体の拡張ガウス像と呼ぶ。図2参照。

物体の画面上での見掛けの形は、その3次元移動

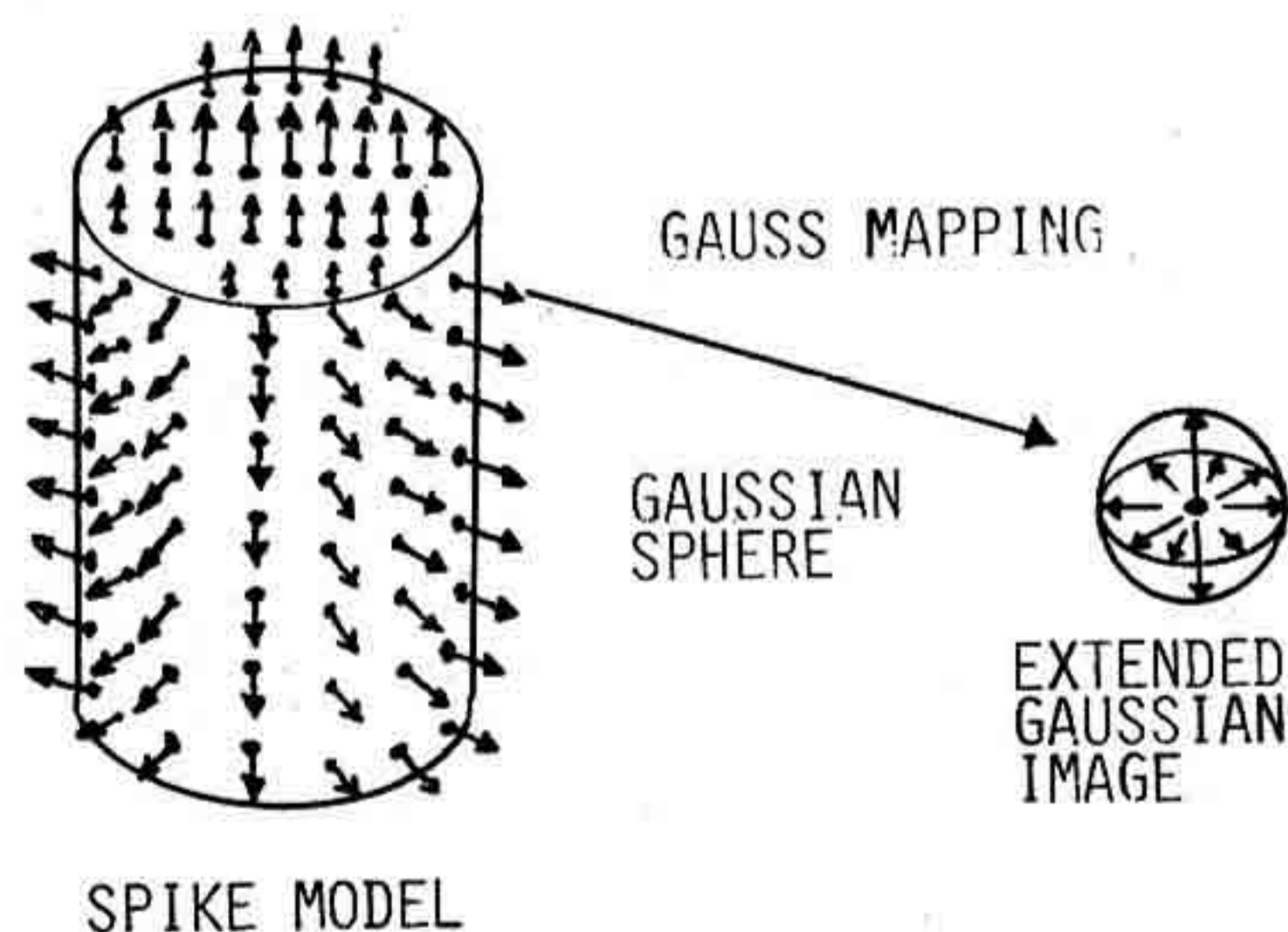


図2 拡張ガウス像の原理

Fig.2 Definition of extended Gaussian image.

により変化を受ける。この内、平行移動は拡張ガウス像に変化を及ぼさない。一方、被写体が回転すると、ガウス球上の拡張ガウス像も回転する。しかし両者の回転様式は一致している。従って、各姿勢での可視半球上での拡張ガウス像と、観測された拡張ガウス像を比較することにより、被写体の姿勢を決定することができる。

拡張ガウス像のマッチングは、物体の回転の自由度に対応して3自由度有る。可視半球がガウス球上のどの部分に対応するかで2自由度。可視半球の視線まわりの回転で1自由度である。この3自由度のマッチングを始める前に、可視半球上での拡張ガウス像の重心位置、慣性モーメントの方向を用いて自由度を減少させた後、マッチングを取ることにする。

拡張ガウス像を実装するためには、球面を分割する必要が有る。この分割は一様であることが望ましい。すなわち、分割された各メッシュの面積がなるべく均一で、かつ近傍のメッシュへの距離がなるべく均一であるような分割が必要である。

ここでは、正12面体の各面を2次分割して得られる多面体に基づく測地ドームを使用する。この測地ドームは、240の小区画を持つ。

非凸物体においては、視線とのなす角が90°以内の面素(可視半球に乗る面素)であっても、物体の他の部分にさえぎられて観測できない場合がある。言い換えると、ガウス球上の拡張ガウス像の回転によって、可視半球上の拡張ガウス像を得ることはできない。従って、各視線方向について、可視半球の拡張ガウス像を定義し直す。

理論的には各視線方向における可視半球上の拡張ガウス像全体を蓄えるためには、(視線方向の自由度) × (面素方向の自由度)次元の参照表が必要となる。これを、デジタイズすることにより、2次元の参照表に蓄えることができる。

ガウス球面は、前述の測地ドームを用いて小区画に分割できる。この小区画に番号を割り振り、その番号順に並べることで、観測方向を1次元配列とできる。

次に、各視線下における可視半球も同様の分割で小区画となる。同様に、これに番号を割り振ることで1次元配列となる。この両者を組み合わせて、各行が視線方向、各列が面素方向を示す2次元配列を得る。この2次元配列上に拡張ガウス像を表現する。これを拡張ガウス像参照表と呼ぶ。

マッチングに便利なように次のような工夫をする。

まず、各視線方向について、対称物中心座標系と一定の関係を持つ仮想的画像面を考える。この仮想的画像面の座標軸を基準として、可視半球をメッシュ化し1次元配列とする。次に、この仮想的画像面上で拡張ガウス像の慣性モーメント最小軸の方向を求める。この軸と、仮想的画像面上のX軸が一致するように物体を回転させた後、物体の拡張ガウス像を先の1次元配列に蓄える。次に総和が1となるように正規化する。一方、1次元配列に2個の新しい要素を付加し、これに拡張ガウス像の重心と先の回転角を蓄える。この操作を各視線方向について行ない、参照表を完成する。

観測されたいがぐり表現は、その拡張ガウス像モーメント最小軸が画像のX軸と一致するように回転される。次いで、回転されたいがぐり表現像から、画像面の座標系に基づいて作られた測地ドームに対応する1次元配列に拡張ガウス像を得る。正規化し、重心も得る。

まず、重心位置から可能性のない視線方向を消去する。生き残った視線方向について、観測拡張ガウス像との類似度が計算される。ある視線方向 v の類似度は下記の評価関数を用いて計算する。

「参照表の v 行について、各 i 要素について、

* $EGIM^{observe}(i) = 0.0$ ならならにもせず

* $EGIM^{observe}(i) \neq 0.0$

$$\Delta = \frac{|EGIM^{observe}(i) - EGIM^{model}(i)|}{EGIM^{observe}(i)}$$

$$d = distance(i, i+e)$$

$$d < d^{standard}, \Delta < \Delta^{standard}$$

ならば、 $EGIM^{observe}(i)(1-\Delta)\cos d$ を v 行の得点 $S(v)$ に加える。

$EGIM^{observe}(i)$ は、観測拡張ガウス像の第 i 番目のセルの質量、 $distance(i, i+e)$ は、 i 番目のセルと $i+e$ 番目のセルの代表角のなす角度である。

上記の操作の後、 $S(v)$ が最大となった視線方向が、その時の視線方向とされる。次に、参照表に登録された回転角と、観測拡張ガウス像の回転角から、視線まわりの回転角が決定され、観測方向と物体中心座標系の関係が決定される。

3.5 安全な物体把握

安全に物体を把握するため、つぎの2条件が満たされなければならない。

(1) グリッパが安定に物体を把握すること。

これは、つかもうとしている物体の形状とグリッパの形状との関係により決定される(安定把握)。

(2) グリッパがつかもうとしている時に、他の物体

と衝突しないこと。

これは、グリッパの形状、姿勢と他の物体の存在状況により決定される(衝突回避)。

前述の2つの条件ともグリッパの形状に左右されるので、使用するグリッパを決定する。2個の直方体により構成される2本指である。

安定把握については、花房ら^[4]の3本指についての研究、Bradyの2本指^[5]についての研究がある。両者とも物体の重心位置との関係から把握系のポテンシャルが最小となる姿勢で物体を把握することを提案している。

ビンピッキングの場合、物体の周囲に他の物体が存在し、上記のような理想的な姿勢で物体を把握できるのは極めてまれである。従って、もう少しゆるい条件で安定把握姿勢を定義する。下記の2条件が満たされれば、安定把握とする。

(1) グリッパが把握するために閉じる際、物体が平行移動しないこと。

(2) グリッパが把握するために閉じる際、物体が回転移動をしないこと。

2本指のグリッパでは、2点で物体に力を及ぼす。

(1)、(2)の条件が満たされるためには、この2力が、互いに反対方向で絶対値が等しければよい。

一般に、物体の摩擦角が予想できない場合には、2つの接触点で、そこでの面素の方向と把握線の方向が一致するような点のペアのみを採用すればよい。こういったペアは、幾何モデルを利用して抽出できる。

安定姿勢の内、他の物体と衝突しない姿勢を選ぶ必要がある。この衝突回避姿勢は、グリッパの形状と物体の形状により決定される。

一般に、衝突回避問題は、C拘束法^[6]を用いて解析されることが多い。しかし、ビンピッキング問題では、障害物が数多く互いに重なっているため、C拘束面の導出に手間がかかる。従って下記に述べるような直接法を取る。

前述の手法により導出された安定把握姿勢についてそれが他の物体と衝突しないかどうかを調べる。グリッパが物体を把握しようとしている時、その下面は位置方向が一つの安定把握姿勢により決定される2つの長方形領域上を動く。従って、この長方形領域が他の物体と交わるか否かを調べればよい。

3.6 カメラモデル

カメラ座標系とマニピュレータの座標系の関係を定める必要が有る。これを解析的に行なう方法も提案

されているが、カメラ座標系の測定に誤差を伴う。従って、次に述べるような実験手法を取った。

(u, v) をカメラ画像面の座標、 (x, y, z) をマンピューレータの座標系とする。マーカーをマンピューレータの z 軸が一定な平面で動かし、その時の (u, v) と (x, y, z) を記録する。この場合、

$$x = au + bv + e, \quad y = cu + dv + f$$

なる関係がある。よって、この係数を記録した値を用いて、最小二乗法で決定する。

これらの係数は、当然 z の値によって変化する。従って、物体が存在するであろう範囲の上底と下底において、上記の手法で2組の係数を決定する。中間の z の値に対しては、2組の係数を内挿して使用する。

4. 実験

図3にハードウェア構成を示す。照度差ステレオ、領域分割、拡張ガウス像マッチングを行なうメインシステムが一台のLispマシンの上に実装された。一方プリズムステレオは、もう一台のLispマシンに実装

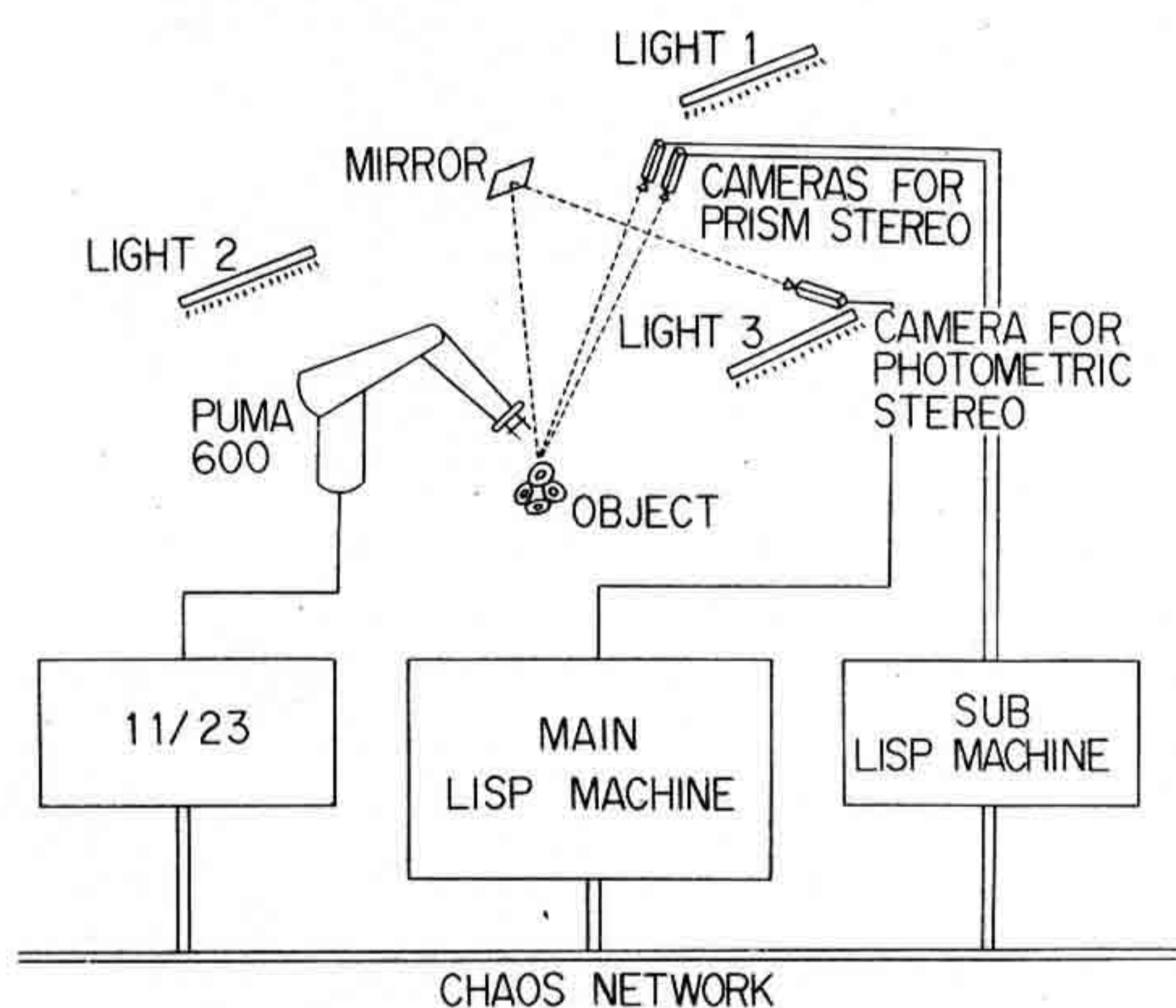


図3 ハードウェア構成
Fig.3 Hardware configuration.

され、二台のLispマシンは並列処理を行なう。マンピューレータのコントロールは11/23が受けもつ。二台のLispマシン及び11/23は、Chaosネットワークを介して結ばれ、並列処理の開始、プリズムステレオの出力などの転送が行なわれる。照度差ステレオ入力から、拡張ガウス像マッチングまでのメインプロセスの処理は、中央のメインLispマシンで行なわれ、プリズムステレオの処理はサブLispマシンで並列に進む。

メインLispマシンが照度差ステレオ用の3光源及びプリズムステレオ用のランダムドット投影スライドプロジェクターをコントロールする。メインLispマシンは照度差ステレオ用の3枚の画像を取り込み、サブのLispマシンは、2台のテレビカメラをコントロールして、ステレオペアの画像を取り込む。この時点で照度差ステレオとプリズムステレオが同時に実行を開始する。

本実験では、照度差ステレオを2度動かす。第1回目は、いがぐり表現像を粗く生成して、領域分割を行なうため。2回目は、いがぐり表現像を精密に生成して、拡張ガウス像を生成するためである。

これは、照度差ステレオが使用する参照表の特性を変化させることにより実現できる。照度差ステレオ1が使用する参照表は、精度が低く、かつ見込み角が40°までしか求まらない。一方、照度差ステレオ2が用いる参照表は、精度が高く、見込み角が65°まで求められる。図4に3画像、図5に照度差ステレオ1により得られた各面素の方向を針で示した。

領域分割は、3.2で述べた知識を用いて行なわれ、図6に示すような連結領域が得られた。

目標領域は、オイラー数と面積より中央部やや右よりの領域と決定された。

照度差ステレオ2は、見込み角が60°以内の面素を決定できる。照度差ステレオ2の参照表には、各面素の拡張ガウス像質量も登録されている。図8に照度差

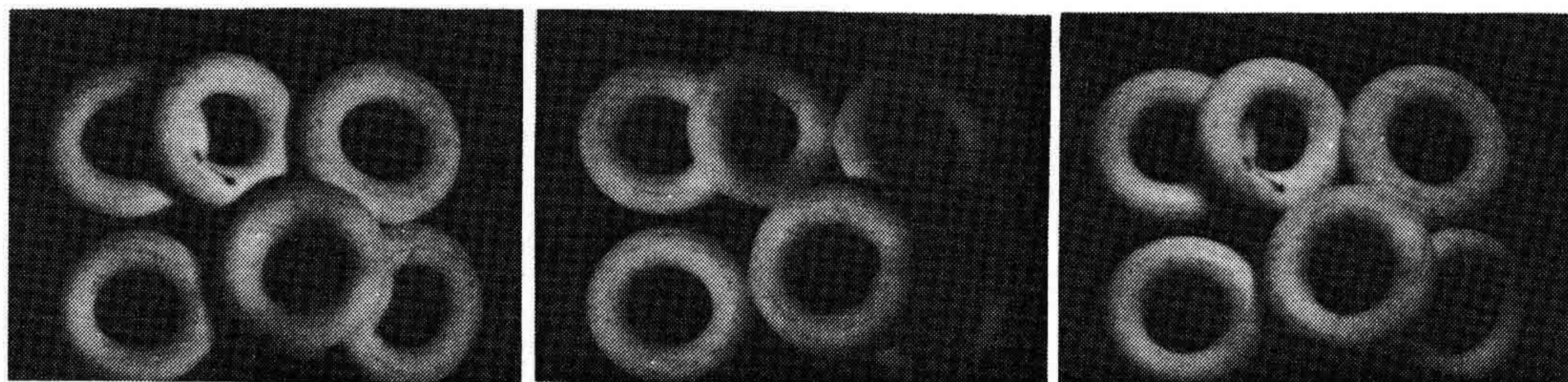


図4 照度差ステレオ入力画像
Fig.4 Three brightness arrays for photometric stereo 1.

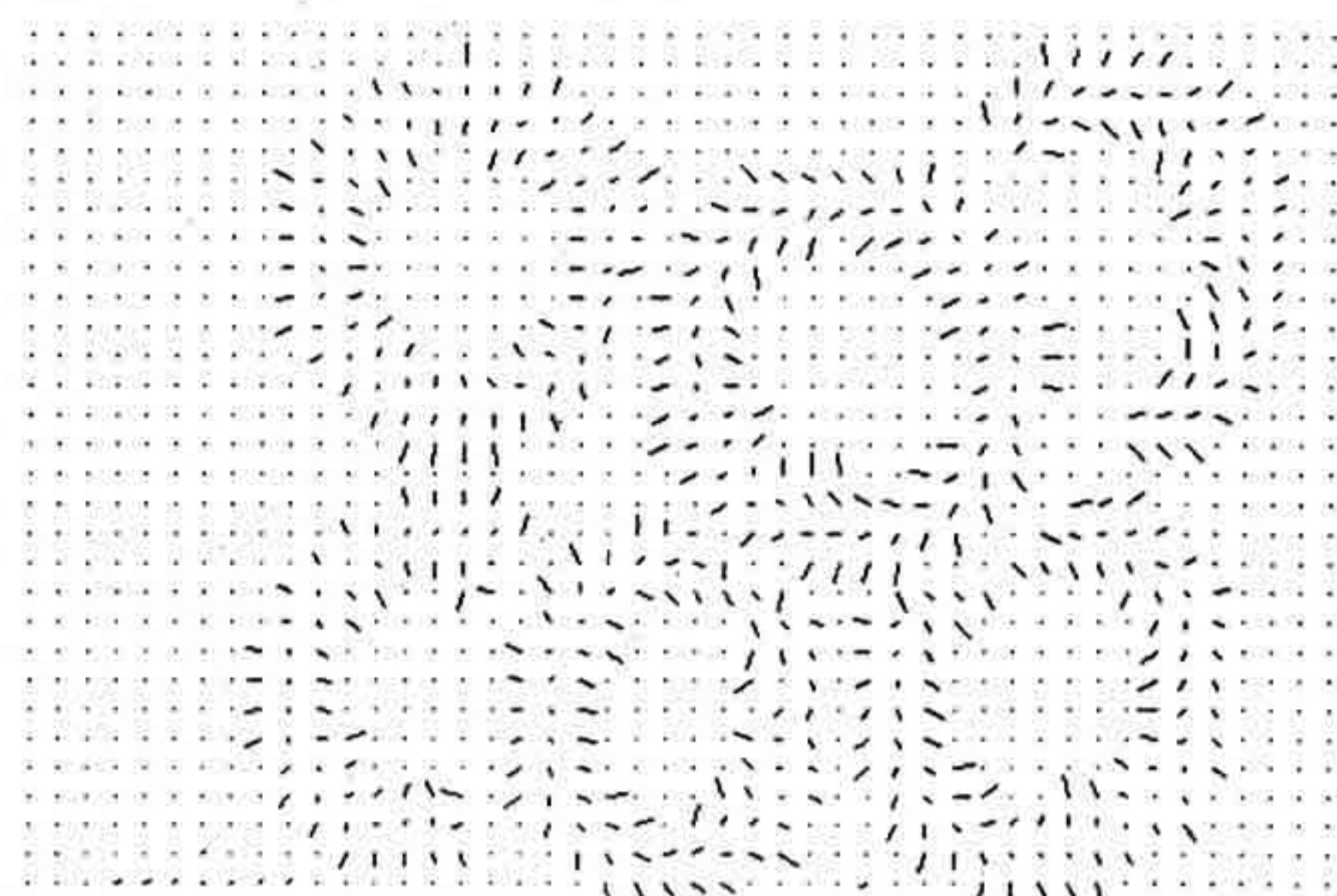


図5 いがぐり表現像
Fig.5 Obtained needle map.

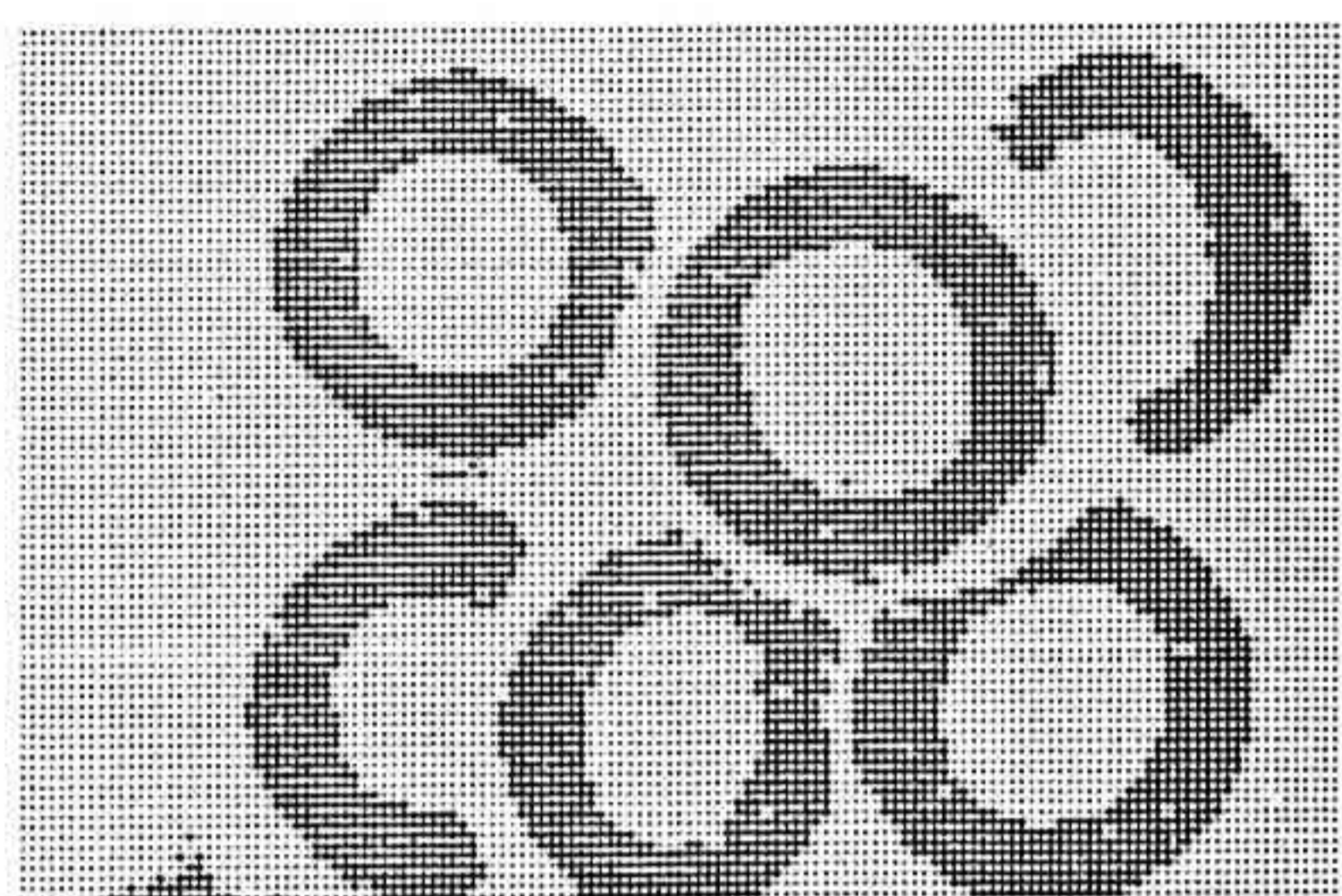


図6 領域分割結果
Fig.6 Segmented region.

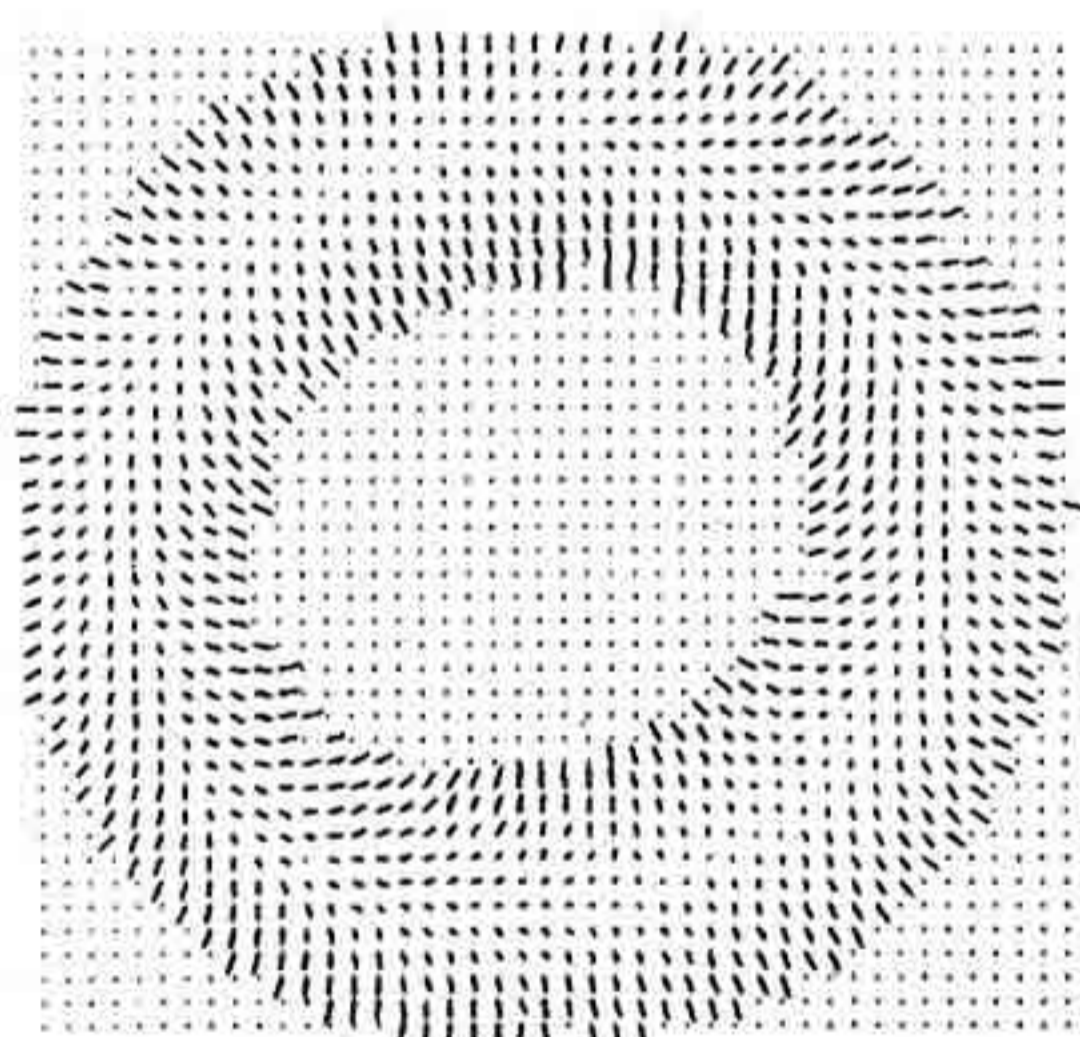


図7 目標領域のいがぐり表現像
Fig.7 Needle map on the target region.

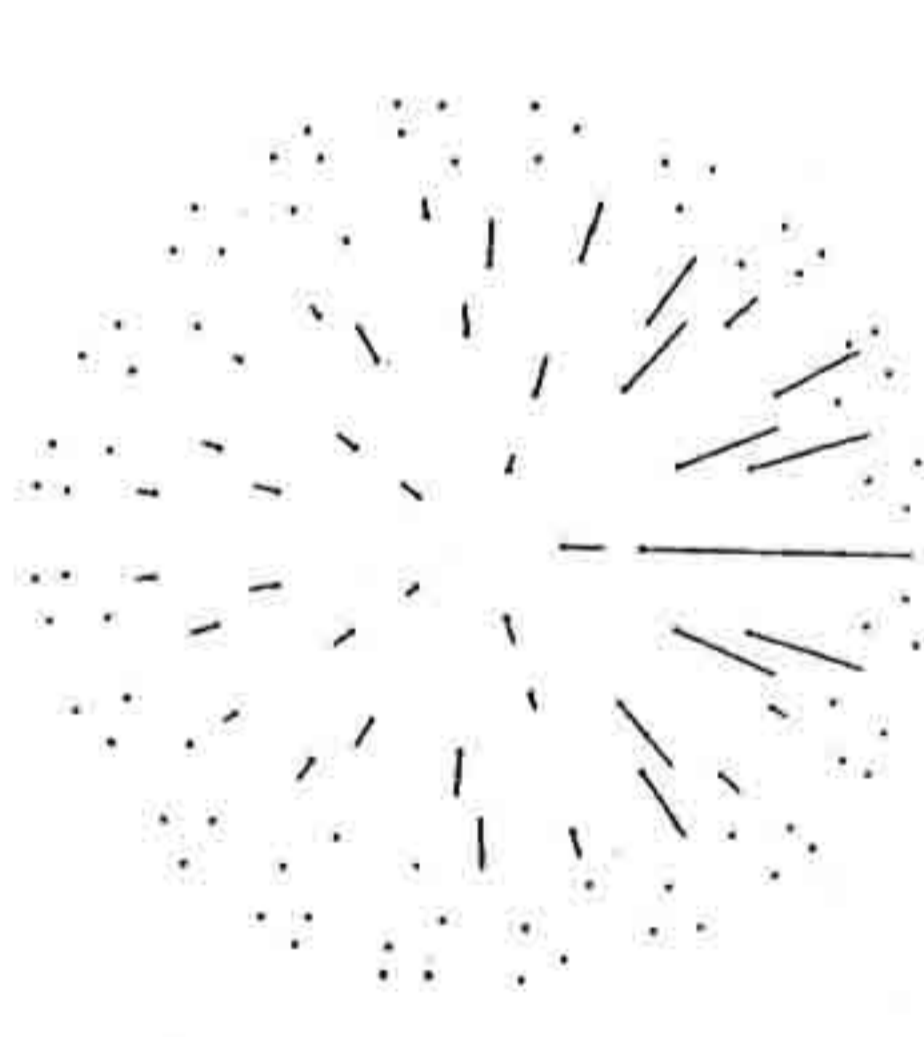


図8 目標領域の拡張
ガウス像
Fig.8 EGI on the target region

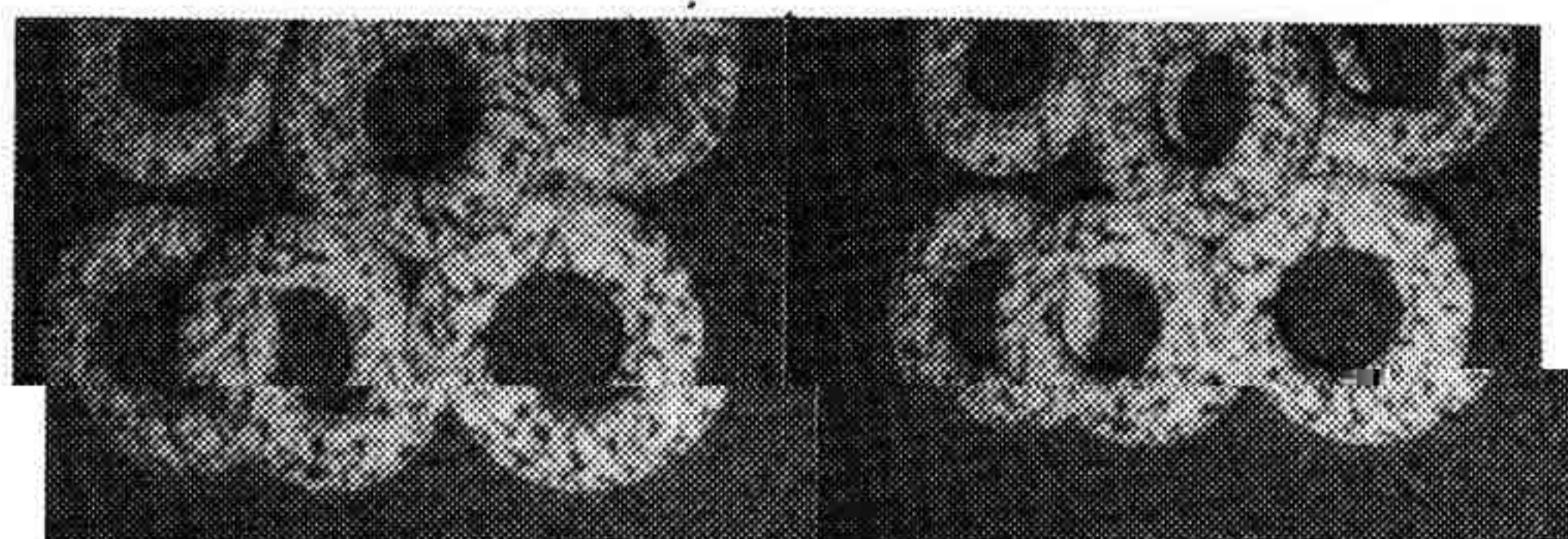


図9 入力ステレオペア
Fig.9 Input images for PRISM stereo.

ステレオ2によるいがぐり表現像を示す。

拡張ガウス像の重心, 慣性モーメント方向, 正規化像が計算される。図9は, 正規化拡張ガウス像であり,

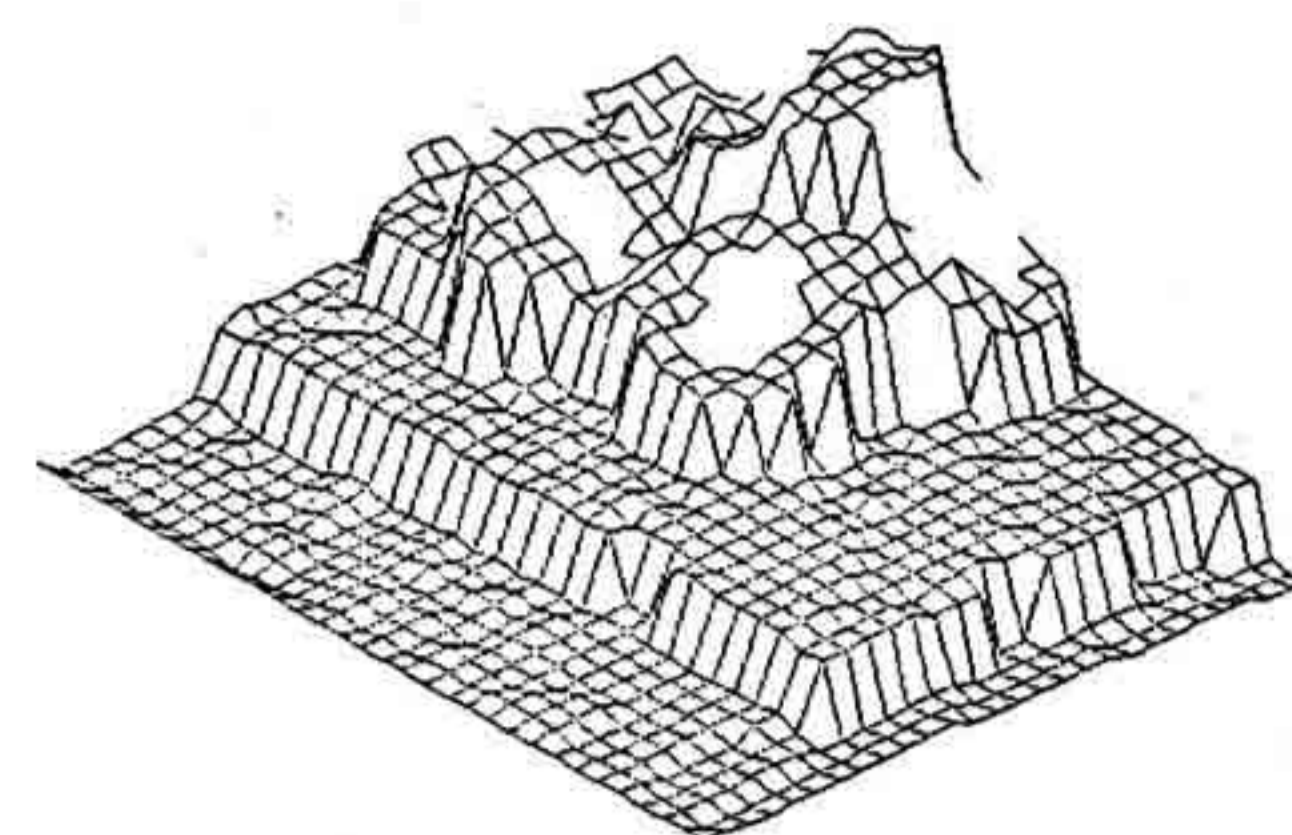


図10 得られた深さ情報
Fig.10 Depth map obtained by PRISM stereo.

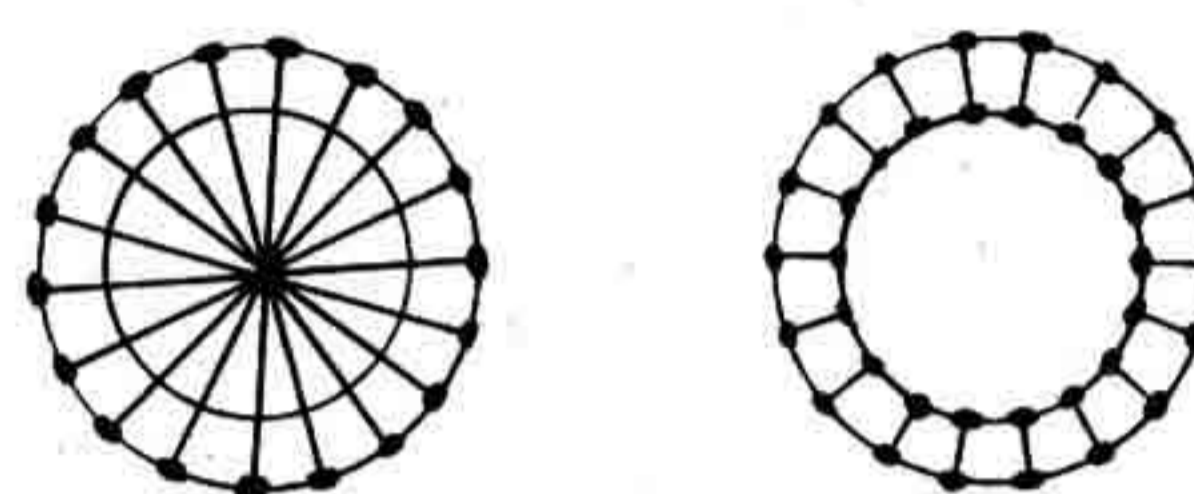


図11 候補つかみ点
Fig.11 Candidate grasping points.

慣性モーメント最小軸が, 紙面水平方向に重なるように回転している。

ドーナツは回転対象なので, 全ての方向からの拡張ガウス像を参照表に登録する必要はない。ドーナツ軸を含む平面とガウス球交わりである大円の4分の1の円弧上の方向で充分である。この区間 10° ごとの方向からの拡張ガウス像を参照表に登録した。モーメント最小軸が正確にX軸に一致した正規化像の他に観測誤差を吸収するため少しX軸とずれた像も登録した。これにより, ドーナツの姿勢が決定された。

この時点でプリズムステレオへ, 深さ情報の伝送が要求される。図10にステレオペア, 図11に得られた深さ情報を示す。なお, 照度差ステレオのTVカメラは鏡を通し, プリズムステレオのTVカメラは直接ピンを眺めているので, 両者の画像の上下が逆転している。

まずドーナツの安定把握位置を導出する。ここでは, 簡単のため, ドーナツを2次元のリングとして取り扱う。ドーナツに対する安定把握姿勢のオリエンテーション方向は, 図12に示す2種類である。図12(a)の方がポテンシャルは低い, グリップの幅の制限から, 実装できない。従って, 図12(b)に対応する点を画像中に探せばよい。一方, アプローチ方向の自由度を考慮しても, 衝突回避の可能性はそれほど増大しない。よって, 一律にドーナツ面に直交する方向とする。目標ドーナツの輪郭線は, 領域分割によって抽出されている。重心の水平位置も決定でき, さらに, ドーナツの姿勢もわかっているため, 輪郭線上の1点が片方の接触点となった時の, グリップの姿勢が決定できる。おのお

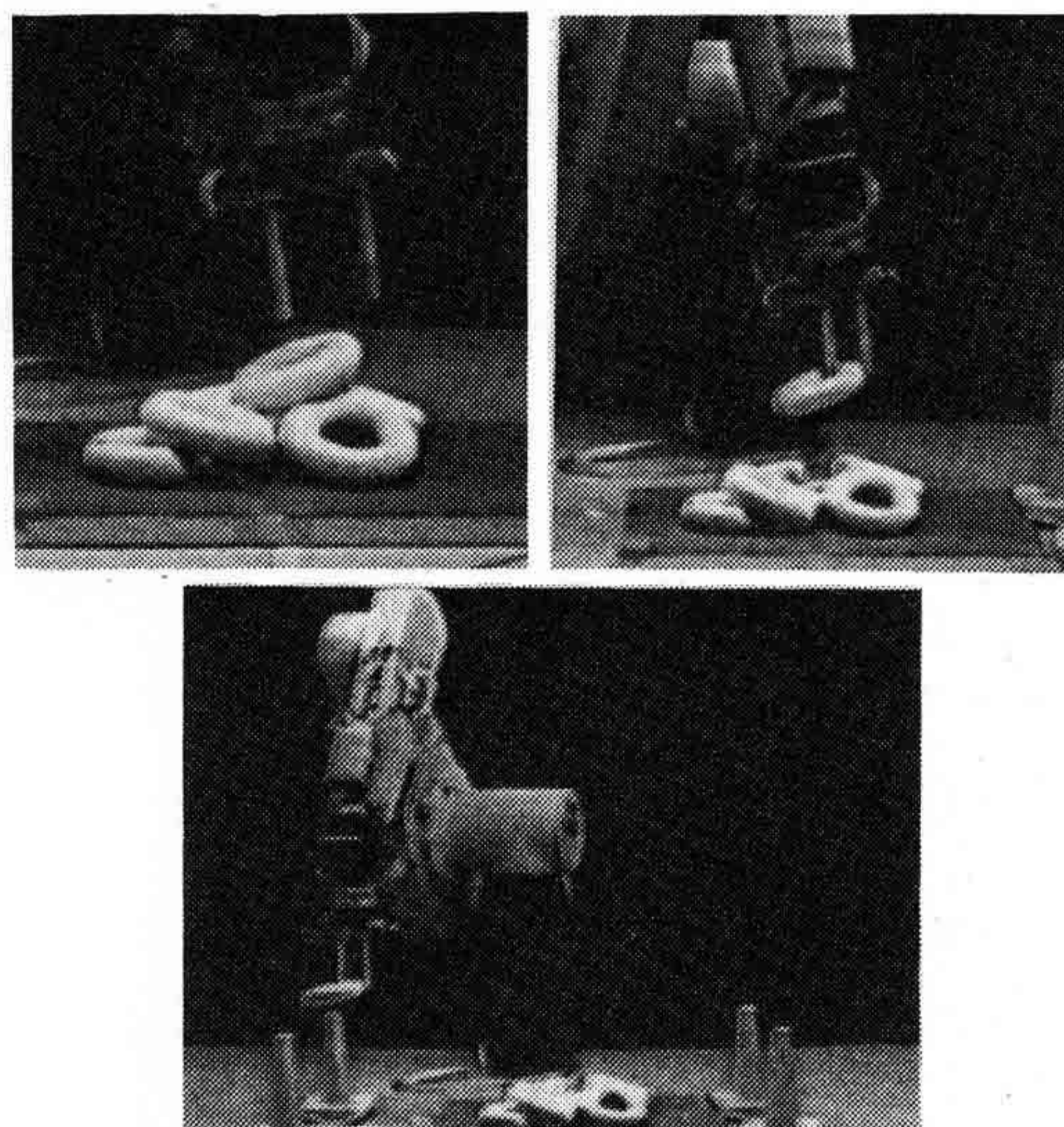


図12 手による実行
Fig.12 Picking up an object.

のの位置で、プリズムステレオの出力とグリップの低面との相関を調べて行く。これから、図13の実行例に示された安全把握姿勢が決定された。

5. む す び

ビンピッキングを行なうためのシステムを示した。システムは、照度差ステレオ、拡張ガウス像で姿勢を決定し、プリズムステレオで安全把握姿勢を決定した。

画像取り込みからグリップへの位置姿勢命令送らまでの実行時間は、40-50秒であった。全システムはZeta-lispを使用してコードした。

本稿の目的は、近年開発されてきた技術を用いて、コンピュータビジョンシステムの出発点であったハンドアイシステムを構成してみることにあった。これについては、ある程度の見通しが得られたと思う。今後は、積極的にモデルを利用した手法の開発が待たれる。

謝辞 本研究は、池内、長田がMIT人工知能研究所滞在中にHorn, Nishiharaらと共同研究したものである。この研究の機会を与えて下さった電総研・石

井前部長、若松部長、白井室長、富士通研・棚橋部長、吉田課長、MIT・Winston所長に感謝致します。

文 献

- (1) S. Tsuji and A. Nakamura: "Recognition of an Object in a Stack of Industrial Parts", Proc. 4th-IJCAI, pp.811-818 (1975).
- (2) W. A. Perkins: "Model-based Vision System for Scene Containing Multiple Parts", Proc. 5th-IJCAI, pp.678-684 (1977).
- (3) R. B. Kelley, et al.: "A Robot System Which Acquires Cylindrical Workpieces from Bins", IEEE Trans. SMC, SMC-12, 2, pp.151-160 (1982).
- (4) D. H. Ballard and C. M. Brown: Computer Vision, Prentice-Hall (1982).
- (5) 池内克史: "反射率地図に基づき、二次元濃淡画像より三次元形状を再構成する2手法", 信学論(D), J65-D, 7, pp.842-849 (1982).
- (6) 池内克史: "被写体表面上の小図形の幾何学ひずみから三次元形状を再構成する1手法", 信学論(D), J65-D, 7, pp.850-857 (1982).
- (7) 越川和忠: "反射偏光を利用する面方向検出法", 計測制御論, 18, 10, pp.1031-1033 (1982).
- (8) 杉原厚吉: "頂点辞書を利用した距離画像解析", 情理論, 20, 1, pp.39-44 (1979).
- (9) 大島, 白井: "3次元情報を用いた物体認識", 信学論(D), J65-D, 5, pp.629-636 (1982).
- (10) W. E. L. Grimson: From Images to Surface, MIT Press (1981).
- (11) R. J. Woodham: "Reflectance Map Techniques for Analyzing Surface Defects in Metal Casting", AI-TR-457, MIT (1978).
- (12) K. H. Nishihara: "PRISM: Practical Realtime Image Stereo Matcher", Proc. SPIE, 449, pp.134-142 (1983).
- (13) 池内克史: "拡張ガウス像に基づく被写体いがぐり表現像から観測方向決定のための一手法", 信学論(D), J66-D, 5, pp.463-470 (1983).
- (14) 花房, 浅田: "弾性指を有するロボットハンドによる物体の安定な把握", 計測制御論, 13, 4, pp.370-377 (1977).
- (15) M. Brady: "Parts Description and Acquisition using Vision", Proc. SPIE, 336, pp.20-28 (1982).
- (16) T. Lozano-Perez: "Spatial Planning: A Configuration Space Approach", AI. Memo No.605 (1980).

(昭和59年8月22日受付)