

## 2値画像による 多面体の位置・姿勢計測

久野敦司, 山口芳徳, 政木俊道, 坂和彦, 中塚信雄,  
加藤充孝 (立石電機)

### 1. まえがき

物体の3次元形状を認識したり、その3次元的な位置、姿勢を測定できる視覚機能は知能ロボットの実現にとって極めて重要である。3次元的な形状認識、位置決めを行なおうとすると、物体の各点までの奥行き情報を信頼性良く入力できることが基本となるため、多くの研究がなされている。その有力な手法として投光法がある。しかし投光法には、1) 違い物体の測定が困難、2) 明るい所での使用が困難等の欠点がある。また、他の有力な手法として、ステレオ法がある。このステレオ法では、上に述べた欠点は解決されるが、2つの画面の間での対応部分を決定する、いわゆる対応づけの問題が未解決である。本研究では、安定に抽出できる輪郭線に着目することと、もう1台のカメラを導入して3眼にする事によって対応づけの問題を解決した。<sup>(1)</sup>

また、物体の3次元位置、姿勢を求めるためには物体のモデルのどの部分が、現に観測している物体のどの部分に対応するかを決定する必要があるが、本研究では、Maximal Clique法の採用により高速に、これを実行した。本手法は従来の3次元物体計測方法に比較して、次の特徴を有する。

- (1) 安定・高速に抽出できる輪郭線に着目した。
- (2) 3眼を用いるので、広い視野が得られる。
- (3) 3眼立体視法により高速立体視が可能

### 2. システム構成

図1に本報告のシステム構成を示す。3台の同軸照明(TVカメラの光軸と照明軸を一致させた照明)付きTVカメラ、画像処理装置及び汎用コンピュータから構成される。

#### (1) 観測系

撮像装置：3台の2次元CCDテレビカメラ

照明装置：3台の同軸照明装置

背景：シート状再帰性反射物(住友3M製)

これにより高コントラストの画像が得られる。

#### (2) 画像処理装置

Deanza IP8500 Image Processor  
(512×512画素、256階調の画像メモリー付き)

#### (3) 汎用コンピュータ

Dec社 VAX11/780 (VMS)

#### (4) プログラミング言語

C言語+FORTRAN

### 3. 原理および処理の流れ

本システムは2個のプログラムから構成されている。各プログラムの処理原理は、次のとおりである。

#### (1) モデル登録プログラム

本プログラムは、カメラモデル、形状モデル、特徴点間距離モデル、フレーム定義の各ファイルを生成する。

カメラモデルはピンホールカメラを表現するものを採用しており、3台のTVカメラそれぞれについてインタラクティブに生成する。

形状モデルはワイヤーフレームモデルを、採用している。

特徴点間距離モデルは、形状モデルから自動生成される。

フレーム定義ファイルは、ロボットが物体を把握する部分を示すフレームを定義するものであり、インタラクティブに生成される。

#### (2) 位置・姿勢計測プログラム

本プログラムは、3台のTVカメラからの画像を線画化し、線画の特徴点を用いて物体の頂点を立体視した後に、物体の3次元構造の記述を生成する。

この3次元構造の記述は、形状モデル及び特徴点間距離モデルと比較照合され、観測物体の頂点と形状モデルの頂点との対応付けが行なわれる。

頂点間の対応付けデータ及び各頂点の3次元座標を用いて、観測物体の位置・姿勢を算出する。

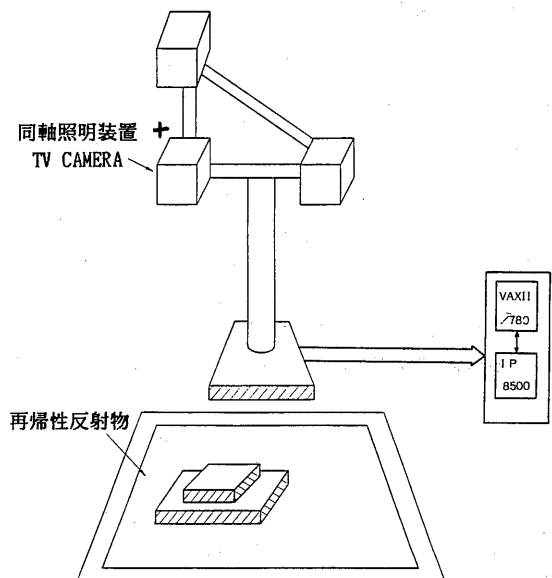


図1 システム構成図

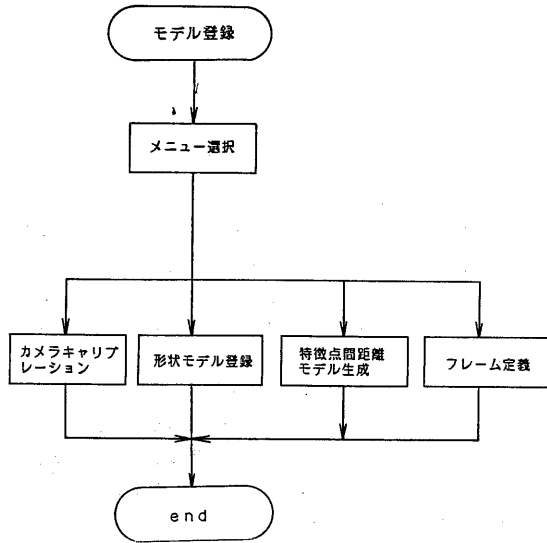


図2 モデル登録

### 3.1 モデル登録

#### 3.1.1 カメラキャリブレーション

##### (1) カメラモデル

カメラモデルは、TVカメラの位置・姿勢、レンズの結像距離、ビデオ信号のサンプリングと画像メモリーへの設定過程を表現するものである。

図3においてカメラモデルは、3次元空間上の点 $P_i$ を画像上の点 $S_i$ に写像する。

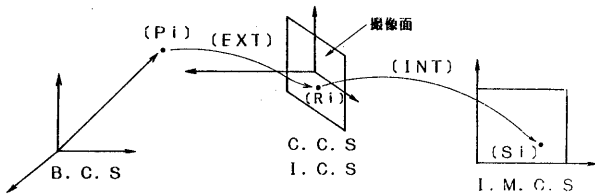


図3 カメラモデル

カメラモデルには多くの座標系が関与しており、各座標系を用いて座標変換の過程を記述できる。

#### 座標

#### 基準座標系

- $(P_i) = (x_i, y_i, z_i, 1)^t$  : B.C.S.(Base Coordinate System)
- $(Q_i) = (e_i, f_i, g_i, 1)^t$  : C.C.S.(Camera Coordinate System)
- $(R_i)^* = (p_i^*, q_i^*, 0, h_i^*)^t$  : I.C.S.(Image Coordinate System)
- $(R_i) = (p_i, q_i, 1)^t$  : I.C.S.
- $(S_i) = (i_i, j_i, 1)^t$  : I.M.C.S.(Image Memory Coordinate System)

$$(P_i) \xrightarrow{(FRAME)^{-1}} (Q_i) \xrightarrow{(LENZ)} (R_i)^* \xrightarrow{(SAMP)} (S_i)$$

図4 各種座標間の関係

図4に示す座標を用いて、この過程を示すと次のようになる。

$$(Q_i) = (FRAME)^{-1} \cdot (P_i) \quad \text{--- (1)}$$

$$(R_i)^* = (LENZ) \cdot (Q_i) \quad \text{--- (2)}$$

$$p_i = p_i^* / h_i^* \quad \text{--- (3)}$$

$$q_i = q_i^* / h_i^* \quad \text{--- (4)}$$

$$(S_i) = (SAMP) \cdot (R_i) \quad \text{--- (5)}$$

(a)  $(FRAME)^{-1}$ について

$(FRAME)$ はB. C. SからみたTVカメラの位置・姿勢をあらわす4行4列の行列である。

また、 $(FRAME)^{-1}$ はB. C. S表現の座標をC. C. S表現の座標に変換する行列である。

$$(FRAME)^{-1} \equiv \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{--- (6)}$$

(b)  $(LENZ)$ について

これは結像変換をあらわす行列であり、C. C. S表現の座標をI. C. S表現の同次座標に変換する。

$$(LENZ) \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-1}{lenb} & 1 \end{bmatrix} \quad \text{--- (7)}$$

ここで、 $lenb$ はレンズ中心と撮像面との距離である。

(C)  $(SAMP)$ について

これは撮像面上の像をサンプリングして、画像メモリー上の画像に変換するものである。

すなわち、I. C. S表現の座標をI. M. C. S表現のものに変換する座標変換行列である。

$$(SAMP) \equiv \begin{bmatrix} \frac{1}{PICHX} & 0 & \frac{-ORGX}{PICHX} \\ 0 & \frac{-1}{PICHY} & \frac{ORGX}{PICHY} \\ 0 & 0 & 1 \end{bmatrix} \quad \text{--- (8)}$$

ここで (EXT)  $\equiv$  (LENZ)  $\cdot$  (FRAME)<sup>-1</sup> をカメラの外部モデル、(SAMP) を内部モデルと呼ぶ。

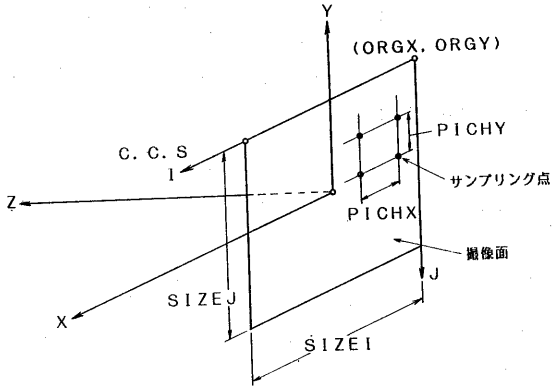


図5 サンプルング

(2) カメラモデルのキャリブレーション

カメラの内部モデルは、TVカメラの移動やレンズ調整の影響を受けない。したがって、ここでいうカメラモデルのキャリブレーションは、カメラの外部モデルのキャリブレーションである。カメラの外部モデルのキャリブレーションにおいては、(1)、……、(5)式で座標(Pi)、(Si)を入力とし、(LENZ)、(FRAME)<sup>-1</sup>を出力する。

[2]

この過程は、解析的手法で初期値を算出した後に山登り法によって最適なカメラの外部モデルを得るという手法によって実現している。

3.1.2 形状モデル登録

形状モデルとしては、ワイヤフレームモデルを用いる。ワイヤフレームモデルは、物体の3次元形状を頂点和辺によって記述する。

3.1.3 特徴点間距離モデル登録

特徴点間距離モデルは、上記の形状モデルから自動生成される。特徴点間距離モデルは、特徴点の属性として次に示す距離ランクベクトルを用い、特徴点間の位置関係の表現として、特徴点間距離を用いる。

特徴点Qiの距離ランクベクトルA(Qi)は次式で与えられる。

$$A(Q_i) = (l_{i1}, \dots, l_{ik}, \dots, l_{in})$$

$$l_{ik} = \begin{cases} 1: & rk - \alpha k \leq |Q_i Q_j| < rk + \alpha k \\ & \text{を満たす} Q_j \text{が、} Q_i \text{の近傍} N(Q_i) \text{内に存在する} \\ 0: & \text{その他} \end{cases}$$

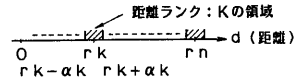


図6 距離ランク

3.2 位置・姿勢計測

位置・姿勢の計測は次図に示す処理によって実行する。

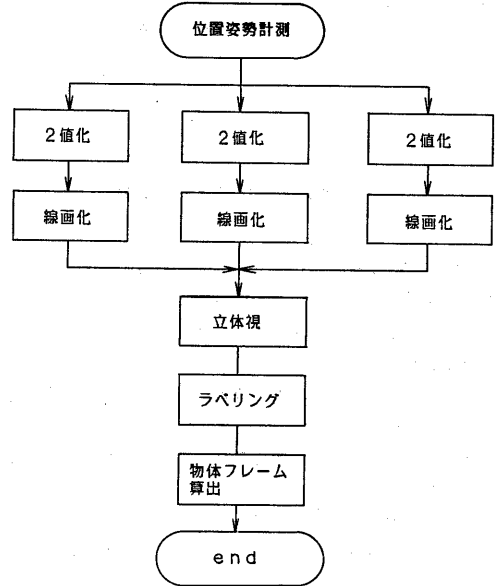


図7 位置・姿勢計測

3.2.1 画像の線画化

2値画像の輪郭を追跡し、その輪郭に直線をあてはめて、線画を得る。

(1) 輪郭追跡

2値画像の輪郭線を次の方法によって追跡する。画像上を3x3マスクをスキャンさせて、まず追跡の開始点を発見する。その開始点から常に黒地を右に見るようにマスクを次の輪郭点上に移動させる。この移動中において、マスクの中心点が通過する軌跡が、輪郭線である。

(2) 直線のあてはめ

輪郭追跡によって得た画素列を、直線部分に分割して、各部分Dに最小自乗法により直線をあてはめる。[4]

$$D = \{ (x_1, y_1), \dots, (x_N, y_N) \}$$

N: 画素数、(xi, yi): 座標

### 3.2.2 立体視

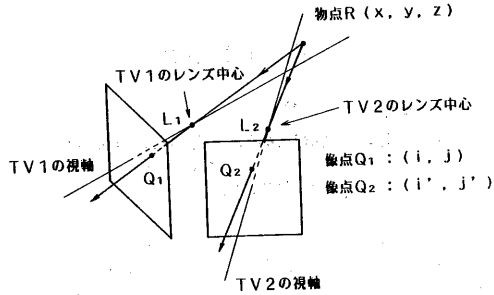


図8 物点と像点の関係

物点Rの3次元座標(x, y, z)を、2台のTVカメラ上の対応する像点Q<sub>1</sub>、Q<sub>2</sub>とカメラモデルにより算出する。TV1、TV2のレンズ中心点の3次元座標は、それぞれのカメラモデルを用いて算出する。

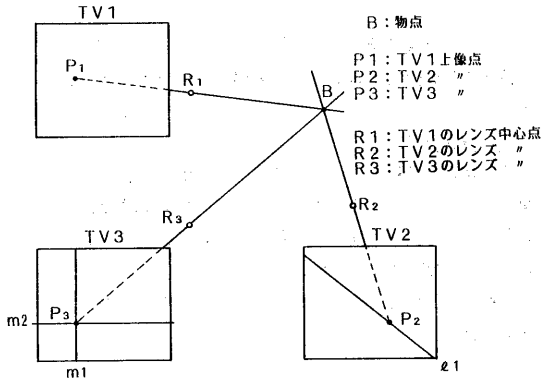
また、像点Q<sub>1</sub>、Q<sub>2</sub>の3次元座標はカメラモデルを参照する事によって算出する。

物点R(x, y, z)は、直線Q<sub>1</sub>L<sub>1</sub>と直線Q<sub>2</sub>L<sub>2</sub>の交点として得られる。

#### (1) エピポーララインを用いた像点の対応付け

立体視においては、複数のTVカメラ上の像点を対応付ける必要がある。

像点の対応付けには、3眼立体視法を用いる。



l1は直線P<sub>1</sub>R<sub>1</sub>がTV2上に生成するエピポーラライン  
m1は直線P<sub>1</sub>R<sub>1</sub>がTV3上に生成するエピポーラライン  
m2は直線P<sub>2</sub>R<sub>2</sub>がTV3上に生成するエピポーラライン

図9 3眼立体視法

TV1上の像点P<sub>1</sub>に対応するTV2上の像点は、l1上に存在する。TV1上の像点P<sub>1</sub>に対応するTV3上の像点は、m1上に存在する。

ここでl1上の像点P<sub>2</sub>がTV3上に生成するエピポーララインm2上に、像点P<sub>3</sub>が存在し、かつP<sub>3</sub>がm1上にあるならば、P<sub>1</sub>、P<sub>2</sub>、P<sub>3</sub>を対応点とする。

#### (2) 視野範囲の拡大

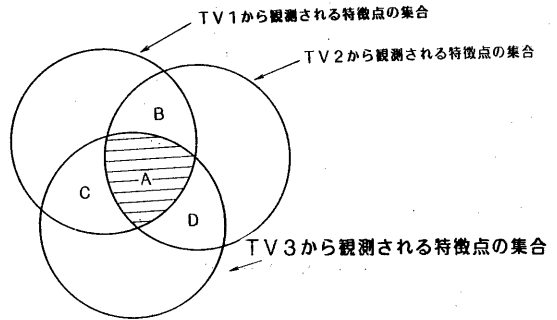


図10 視野範囲の拡大

3眼立体視のみでは、上図Aの部分の特徴点しか立体視できない。これに両眼立体視により3次元座標の得られた特徴点の集合B、C、Dを加える事によって、広い視野を実現した。B、C、Dを加える段階においてAの部分の抽出に用いられた像点は、対応候補から除かれている。そのため、B、C、Dの特徴点の両眼視を行なうときには対応づけのあいまい性は非常に減少しており、信頼性の高い対応づけが行なえる。

#### (3) 特徴点の接続関係の検出

立体視した各特徴点間に辺が存在するか否かを、次のようにして検出した。

立体視できた特徴点をB<sub>1</sub>、B<sub>2</sub>とする。

B<sub>1</sub>、B<sub>2</sub>に対応するTV1、TV2、TV3上の像点を、それぞれP<sub>1</sub>、Q<sub>1</sub>、R<sub>1</sub>およびP<sub>2</sub>、Q<sub>2</sub>、R<sub>2</sub>とする。このとき、次の条件がすべて成立しているときに、特徴点B<sub>1</sub>、B<sub>2</sub>の間に辺が存在するとみなす。

(条件1) P<sub>1</sub>とP<sub>2</sub>が辺で結合している。

(条件2) Q<sub>1</sub>とQ<sub>2</sub>が辺で結合している。

(条件3) R<sub>1</sub>とR<sub>2</sub>が辺で結合している。

このようにして、物体の頂点の3次元座標と頂点間の接続関係から成る3次元物体の記述が作成される。

### 3.2.3 ラベリング

立体視の過程で作成された3次元物体の記述は、モデル登録時に生成された形状モデル及び特徴点間距離モデルと比較・照合される。そして、物体の特徴点とモデルの特徴点の対応付けを実行する。

**(1) 辺の長さを用いた対応候補の検出**

モデルの特徴点Qと観測物体の特徴点Pの比較は、次のようにして行なう。Pから出る辺の集合をEP、Qから出る辺の集合をEQとする。

$$EP = \{L1, \dots, Li, \dots, LN\}$$

$$EQ = \{l1, \dots, lj, \dots, lM\}$$

ただし、Li、ljは辺の長さをも表わす。

このとき、次の量を求める。

$$di = \min \{ \text{abs}(Li - lj) \} \quad \text{--- (9)}$$

$$EQ = EQ - \{lj\} \quad \text{--- (10)}$$

$$\text{maxd} = \max \{ di \} \quad \text{--- (11)}$$

$$Li \in EP$$

ここで、maxd < th1ならば(P, Q)を対応候補とする。また、maxd ≥ th1ならばPとQは対応しないと判定する。ただし、th1はしきい値である。

**(2) 距離ランクベクトルを用いた対応候補の検出**

観測物体の特徴点Piとモデルの特徴点Qjの対応度Cijは、次のようにして算出する。

Pi、Qjの距離ランクベクトルをA(Pi)、A(Qj)とする。

$$A(Pi) = (l_{i1}, \dots, l_{ik}, \dots, l_{in})$$

$$A(Qj) = (l_{j1}, \dots, l_{jk}, \dots, l_{jn})$$

$$l_{ik} = \begin{cases} 1 : rk - \alpha k \leq |Pi - Pm| < rk + \alpha k \text{ を} \\ \quad \text{満たす特徴点Pmが、Piの近傍N(Pi)} \\ \quad \text{内に存在する。} \\ 0 : \text{その他} \end{cases} \quad \text{--- (12)}$$

$$Cij = n - \sum_{k=1}^n (l_{ik} + l_{jk}) \quad \text{--- (13)}$$

ここで、Cij > th2ならば(Pi, Qj)を対応候補とする。th2はしきい値である。

**(3) 対応組間の矛盾の検出**

前記の両基準ともに対応候補と判定した対応組を、ノードとするノードの集合をVとする。

$$V = \{nij \in (Pi, Qj) \mid nij \text{は対応候補}\}$$

ここで、相互に矛盾のないノード間をエッジで結合した無向グラフ G = (V, E) を生成する。

$$E = \{eijkl \in nij \text{ と } nk \text{ の適合度}\}$$

$$eijkl = \begin{cases} 0 \text{ (矛盾)} : (i = k \text{ OR } j = l) \\ \quad \text{OR } (PiPk \neq QjQl) \\ 1 \text{ (適合)} : \text{その他} \end{cases}$$

**(4) 信頼性の高い対応組集合の抽出**

対応候補間の適合関係をあらわすグラフGから、信頼性の高い対応組集合を抽出する。

グラフG中のMaximal Clique<sup>[3]</sup>は、相互に適合する最大のノード(対応組)集合であり、高い信頼性を有する。

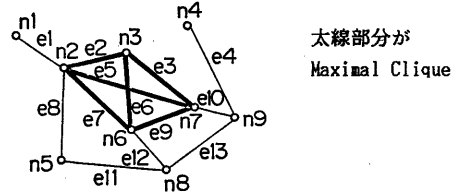


図11 グラフ中のMaximal Clique

**3. 2. 4 位置・姿勢行列の算出**

モデルの特徴点と観測物体の特徴点の3次元座標を、対応するものが同じ列になるように行列化したものを、(A)、

(B)とする。

観測物体の位置・姿勢をあらわす行列を(W)とする。

$$(A) = \begin{bmatrix} x_1 & x_i & x_N \\ y_1 & y_i & y_N \\ z_1 & z_i & z_N \\ 1 & 1 & 1 \end{bmatrix} \quad \text{--- (14)}$$

$$(B) = \begin{bmatrix} x'_1 & x'_i & x'_N \\ y'_1 & y'_i & y'_N \\ z'_1 & z'_i & z'_N \\ 1 & 1 & 1 \end{bmatrix} \quad \text{--- (15)}$$

このとき4行4列の行列(W)は、次式により求める。

$$(W) = (A) \cdot (B)^t \cdot ((B) \cdot (B)^t)^{-1} \quad \text{--- (16)}$$

**4. 実験結果**

物体の3次元物体の位置、姿勢の計測結果を下表に示す。対象物体の2値画像および、計測した位置、姿勢をもとに作成したワイヤフレームの像を写真1、2に示す。

**(1) 誤差平均値**

X座標誤差 (mm)	Y座標誤差 (mm)	Z座標誤差 (mm)
-0.34	0.31	-0.84

X軸方向誤差 (degree)	Y軸方向誤差 (degree)	Z軸方向誤差 (degree)
5.7	2.7	4.6

(2) 95%信頼区間の中

X座標誤差 (mm)	Y座標誤差 (mm)	Z座標誤差 (mm)	X軸方向誤差 (degree)	Y軸方向誤差 (degree)	Z軸方向誤差 (degree)
0.21	0.19	0.57	2.5	1.39	2.06

5. あとがき

本システムは2値画像に対して3眼立体視を適用して多面体の位置・姿勢を計測するものである。本システムは、立体視における対応付けに3眼立体視法を用い、観測物体とモデルとの対応付けにMaximal Clique法を用いることにより、高信頼度の対応づけを高速に抽出するというソフトウェア上の特長を有する。また、位置・姿勢の計測精度は良好な値を示しておりロボットによる組立作業への応用も可能である。

最後に、本研究をすすめるにあたり多くの討論と助言をいただいた大阪大学基礎工学部の谷内田助教授に感謝します。

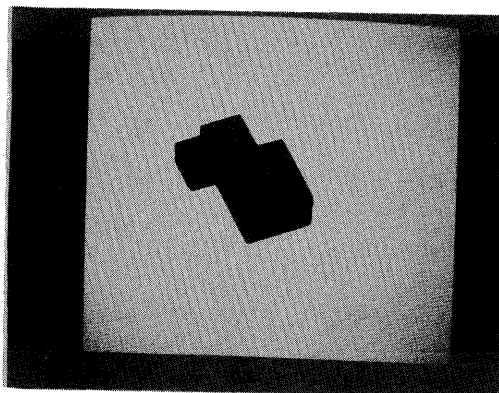


写真1

6. 参考文献

- [1] 谷内田正彦 「多重画像を用いて3次元物体を認識、配線・ハンダ付けの自動化に成功」  
日経メカニカル 1984. 1. 2
- [2] COMPUTER VISION 481p~483p  
DANA H. BALLARD. CHRISTOPHER M. BROWN  
PRENTICE-HALL
- [3] R.C. BOLLES "ROBUST FEATURE MATCHING THROUGH MAXIMAL CLIQUE" PROC. SPIE'S TECHNIC. SYMP. IMAGE APPLIC. WASHINGTON D.C. APR. 1979 P140
- [4] 村上, 岡田, 土屋: "線図形の入力法に関する一考察" 信学論 (D), J59-D, PP117-124  
( '76/2)

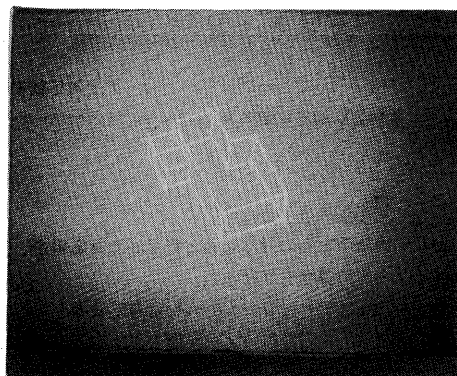


写真2