

## ハンドアイロボットのためのステレオビジョン

高橋 裕信<sup>†</sup> 寺崎 肇<sup>†</sup> 杉本 和英<sup>†</sup> 富田 文明<sup>‡</sup>

<sup>†</sup>三洋電機筑波研究所 <sup>‡</sup>電子技術総合研究所

柔軟な制御が可能なロボットシステムについて検討をすすめている。計測や動作の誤差や予測できない変化があっても、エラーリカバリーを行ない確実に動作するロボットを目指している。

知覚系としてステレオ視を用いたビジュアルフィードバックを基本とした。このシステムでは動作計画を階層化している。視覚機構も階層化され動作計画の各レベルに適したものが選択される。各レベルで認識対象について得られる情報（対象物体のモデル、予測姿勢など）を利用してステレオ法の対応問題の軽減を図り、速度と安定性を高めている。積木の Pick and Place を例題とし、その実験例について報告する。

## Stereo Vision for Hand-Eye Robot

Hironobu TAKAHASHI<sup>†</sup> Hajime TERASAKI<sup>†</sup> Kazuhide SUGIMOTO<sup>†</sup> Fumiaki TOMITA<sup>‡</sup>

<sup>†</sup> Tsukuba Research Center, Sanyo Electric Company, Ltd.

2-1 Koyadai Tsukuba Ibaraki 305 Japan

<sup>‡</sup> Electrotechnical Laboratory

1-1-4 Umezono Tsukuba Ibaraki 305 Japan

We are developing a flexible robot system which can control its behavior correctly even if there are errors in actions and unpredictable changes in the environment. We use stereo vision for the multiple levels of action plans. Since the proper scheme for visual perception is prepared for each level, the use of information about objects given on each level makes stereo correspondence fast and stable. We apply the system to the problem of Pick and Place of blocks in the experiments.

## 1. はじめに

柔軟な制御が可能なロボットシステムについて検討をすすめている。動作誤差や予測できない変化があっても、確実に動作することを目指している。

実際に試作しているシステムでは知覚系としてステレオ視を基本的を使用している。誤差補正ばかりではなく、エラーリカバリーまでを含めた広い範囲のビジュアルフィードバックが必要である。

動作計画のレベルごとに必要とされるステレオ視覚（環境認識と物体認識、誤差補正、動作の検証など）の計算量と確実性を、動作系から得られる情報（対象モデルや予測位置・姿勢）によって高めることができる。

ロボットハンドやすでにロボットが把握した物体などでは、そのモデルが既知でその位置と姿勢がほぼ予測できる。この場合ステレオの各画像中での見え方を予測し、モデルを当てはめることができる。画像との一致を評価すれば、その動作の確認に利用される。また画像とのずれから、動作誤差も求められる。

環境中から対象物体を見つけるような場合には、その位置は予測できない。しかし、対象物体のモデルを用いてステレオの対応問題を軽減することができる。

## 2. 階層構造をもったロボットシステム

検討しているシステムは図1のように、計画系、知覚系、検索系から構成されている。計画系では与えられたタスクと条件に従って、障害物回避・把握計画・作業手順の生成などを行ない、接続されたロボットに動作を指示する。知覚系にはセンサーが接続され、必要に応じて環境、対象物、ロボット自身の動作などを監視する。検索系には環境や対象物、動作計画に必要な知識がある。ユーザーから与えられる指令に対してシステムは現在の状況を知覚系によって知り、目標と照らし合わせて計画系が行動計画を立て実行する。その際に、知覚には対象物のモデルなどを、計画・実行時には行動モデルなどを検索系から求めて利用する。

このシステムの特徴は計画系と知覚系が階層構造を持ち、レベル間の縦のつながりも存在することである[1]。図1に示したように、このハンドア

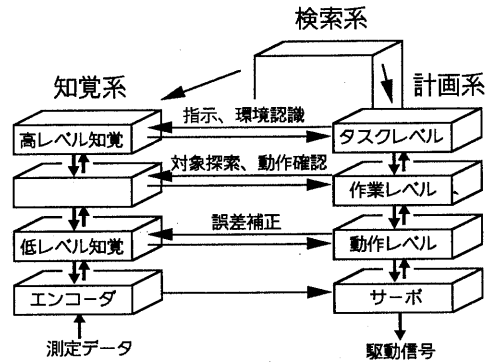


図1 ロボットシステムの構成

イの場合には、レベルとして、タスクレベル、作業レベル、動作レベルの3つの階層を設定した。

具体的な pick and place のシミュレーションを示す。条件としてロボットにはあらかじめ対象物のモデルを与える。オペレータは作業の目標状態を見せる。ロボットは現在の状態と目標状態を画像により認識し、動作の計画と実行を行なう。

### [タスクレベル]

最上位のタスクレベルでは目標状態を認識し、作業手順を作成する。個々の手順はプリミティブタスクで表現される。例えば図2のような場合を考える。初期状態と目標状態でともに積木AにB

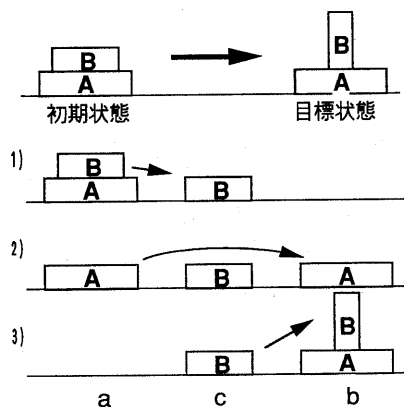


図2 サブタスク

が重なっている。このため、いったん置き換える必要があり、次のような3つのプリミティブタスクが生成される。

- 1) aにある積木Bを一時、場所cに移動。
- 2) aにある積木Aを目標bに移動。
- 3) cにある積木Bを目標bの積木Aに重ねる。

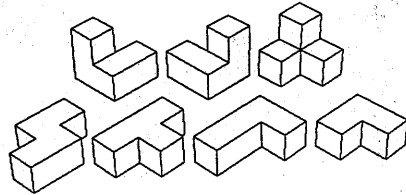


図3 対象物体のモデル

タスクレベルではプリミティブタスクごとに下位の作業レベルでの計画、実行、確認を待つ。下位のレベルで障害物などのために計画できない、もしくは実行したが失敗し対処できないといった場合には、このタスクレベルで環境認識と計画をやり直す。

#### 【作業レベル】

プリミティブタスクを解決するための、作業計画を行なう。作業は「接近」、「把握」、「姿勢変更」などの基本作業と、基本作業を組み合わせた pick and place などの複合作業がある。

プリミティブタスクで姿勢の変更が必要なければ基本作業の「押す」などが可能になる。姿勢を変更する必要がある場合は pick and place を選択する。1回の pick and place でできなければ複数回の pick and place を組み合わせる[2]。

各作業ではその最初に必要があれば対象物を環境中から探索する。精度が要求される作業では知覚系による誤差の補正を行なう。実行後には終了状態を予測し確認する。

タスクレベルと同様に、計画段階で作業不能、もしくは作業後に失敗したことが確認された場合には、対象物の探索と計画をやり直す。このレベルで対処できない場合にはタスクレベルに通知して終了する。

#### 【動作レベル】

移動や把握などの基本作業をロボット動作言語に展開する。実験で使用したPUMAでは、VAL言語の move, open, close などの命令に変換し実行する。

### 3. ステレオ視覚

このシステムでは知覚系としてステレオ法を使用した。これは受動的方法であることから汎用性が期待できる。そのためロボット自身や環境の記述、対象物の認識などに利用できる。また同一の

画像で対象物やロボットを認識できるので相対距離の評価で有利である。他の受動的な方法である単眼法と比較すると、計算コストは大きいものの、高い精度が期待できる。

ステレオ法では対応問題が避けられず、高い安定性を得ることが難しい。そのため画像を境界線表現などの高次のデータに変換し対応付ける方法などが提案されている[3]。

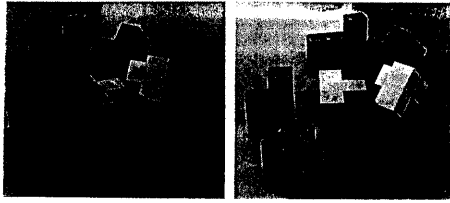
ハンドアイシステムでは適用環境が限定されており、その目的や条件によってステレオの対応問題をさらに軽減することが可能となる。このシステムで知覚系は次の4つの目的で使用される。

- a. 周辺環境の記述、目標状態の認識
- b. 対象物の探索
- c. 動作確認
- d. 動作誤差、認識誤差の補正

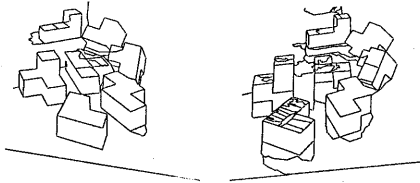
対象についての知識によって認識条件が次のように分類される。

1. 対象についての知識がない。
2. 認識対象のモデルが与えられる。
3. 認識対象のモデルと予測位置が与えられる。

1のモデルを持たない机や障害物について、ステレオによって記述することが可能である[4]。しかしここでは机やロボットなどの環境モデルはあらかじめ与えられているものとした。モデルとしては認識対象(図3)とマニピュレータを与えている。目的のうちa、bは認識対象のモデルが与えられる2に該当する。c、dの動作確認と誤差補正ではおおよその位置と姿勢も与えられるので3に該当する。



(a)



(b)

図4 画像の境界線表現

物体認識は3つの過程にわけて考えられる。(1)初期照合(部分的な照合による座標変換の候補の選択)、(2)検証(全体的な照合による適否の決定)(3)微調整(誤差の測定と改良)である。

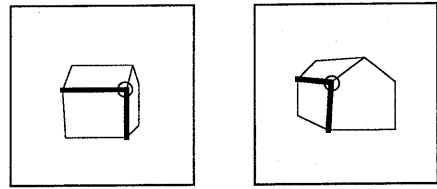
動作確認と誤差補正では、位置と姿勢が予測されているので初期照合を必要としない。動作確認では検証のみを行なえばよく、誤差補正では微調整のみでよい。

【初期照合】

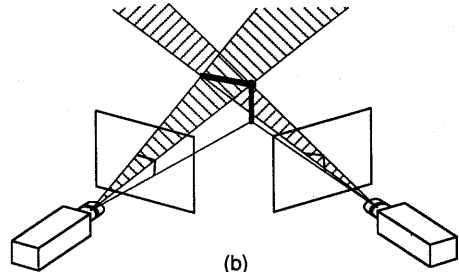
固定した大きさを持ったモデルがあれば単眼法でも対象の初期照合を行なうことができ、最低3点の対応部分があれば空間中の位置を定めることができる(5)。しかし画像中の3点とモデルの3点、もしくは3線どうしを選択しなければならず、その組み合わせ数は大きい。

このシステムでは白黒カメラで撮影して得られたステレオ画像(図4a)を、2次元の境界線表現に変換して使用した(図4b)。

それぞれのステレオ画像中で連結した線分の組を選択し、その連結点がエビポーラ条件を満たすような組合せを求める。例えば図5(a)の画像での○で示した点と2本の線分が条件を満たしている。この2本の線分が3次元空間中で同一の線を表していると仮定すると、空間中の頂点でつながった2本の半直線が求められる(図5b)。この頂点とモデルの頂点間で初期照合する。つぎにそれぞれの頂点のなす角を比較し、一致するもの



(a)



(b)

図5 頂点の対応付け

を選択する。

エビポーラ条件と頂点のなす角によってモデルとシーンの対応が限定されるので単眼法に比べて組み合わせ数を減らすことができる。また半直線と比較しているため、頂点以外の端点がオクルージョンなどによって切れていても影響されないという利点がある。

モデルに回転対称性がある場合には、同一の姿勢にもかかわらず、初期照合と評価を繰り返すことになる。たとえば図6のモデルでは3つの回転

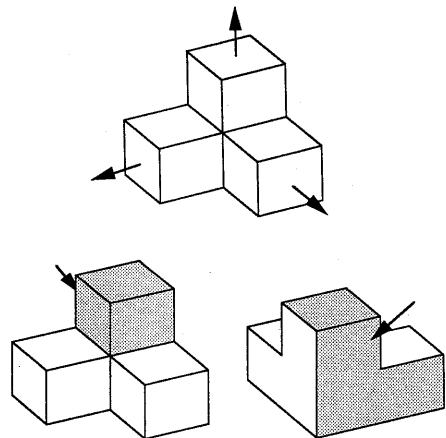


図6 回転対称性

対称性が存在する。そこであらかじめCADデータから対称性を導き、対応する各面のうちのひとつのみを初期照合の候補とする。このモデルの例では網かけで示した部分のみが初期照合を行なう面として照合データに記録される。

【検証】

初期照合での各対応付けの評価と動作確認のために検証を行なう。CADモデルとカメラパラメータおよび対象物の予測位置から、各画像中でのエッジの位置を予測する。これを各画像と比較し、方向が等しく、量子化誤差範囲内で近くにあるものを探す。探索にかかる計算量を減らすため、各画像中の線分はあらかじめHough空間に対応したテーブルにプロットされ、検索できるようになっている。

評価値を規格化するために、画像中で一致した部分をモデル中に投影している。例えば、図7のようにCADモデルから見えることが予測された線分 $M_j$  (a) と実際にステレオの各画像のいずれかで見えた線分 $S_i$  (b) が得られる。それぞれの長さの総和により評価値を定義する。

$$\text{評価値} = \frac{\sum_{i=1}^m S_i}{\sum_{j=1}^n M_j}$$

ステレオ画像間の対応付けが誤っているものや、モデルのあてはめが誤っているものに対しては低い評価値が得られる。

タスクレベルでの目標状態の記述と作業レベルでの対象物の探索では評価値の使い方が異なる。

目標状態の記述では、画像中での個々の線分について、もっともよい評価値を与えたモデルとその位置、姿勢を記憶する。画像中のノイズなどに対しても当てはめることがあるため、しきい値適当に定めて(0.5以上)選択している。

対象物の探索ではそのモデルがもっともよく一致した位置と姿勢を結果とする。

作業の確認では、得られた評価値としきい値を比較し完了または失敗を判断する。

【微調整】

検証と同じように、予測姿勢を用いて画像との比較を行なう。照合の評価とは異なり、画像中での誤差を見込んで対応可能な範囲を大きく取って

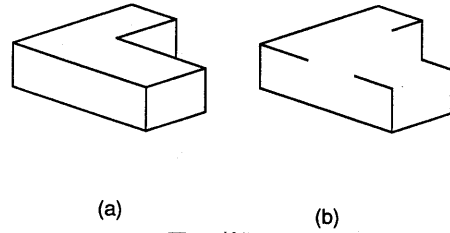


図7 検証

いる。各モデル上の線分 $L_i$  ( $i = 1, 2, 3, \dots, m$ ) に対して、各カメラ $c$ に関するカメラ座標系への変換 $M_c$  ( $c = 1, 2$ ) を行ない透視投影 $P$ を行なう。得られた予測線分 $E_{ic}$  と画像上の線分 $S_{ic}$  と比較する。

動作誤差は移動及び回転の6自由度で与えられるので、この変換を $M_\epsilon$ とする。誤差の総和を以下の式で定義する。

$$\text{総和} = \sum_{c=1}^2 \sum_{i=1}^m f_{\text{dist}}(P M_c M_\epsilon L_i, S_{ic})$$

$f_{\text{dist}}$ は線分間の距離を与える関数で、図8に示したような線分間のなす台形の面積である。 $S$ は $M_\epsilon$ をパラメータとする関数と見ることができるので、最小二乗法により $S$ を最小とする $M_\epsilon$ を求めることができる。これを予測との誤差とした。解法としては6つの各パラメータについて総和を1次偏微分し、繰り返し法により解を求める。

単純に対応する線分の端点間の距離を評価値として用いない理由は、線分の軸方向に対する不要の拘束を避けるためである。図9に示したように端点間で対応付けた場合には、最適解として $b$ しか得られない。このため再度対応付けを繰り返す必要がある。本方法では最適解として $c$ が得られるため、対応付けを繰り返す必要がない。

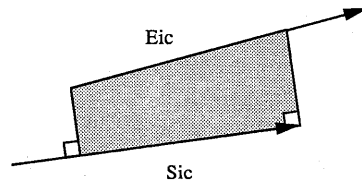


図8 線分間の距離

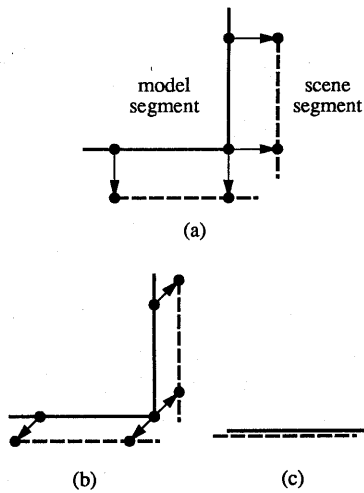


図9 線分の対応付け

#### 4. 実験結果

実験では3台のコンピュータ (SUN3×2、DEC5820×1) をイーサネットで接続し機能を分担した (図10)。C1はシステム全体のコントロールと動作系を担当している。マニピュレータとしてPUMA-560が接続され、そのドライブをする。PUMAには空気圧並行2指ハンドがついている。

C2およびC3は知覚系を担当し、C2は2台の白黒カメラからの画像の入力と前処理、カメラ系のキャリブレーションを行なう。C3は得られたステレオ画像を処理する。C1から送られたモデルや姿勢の予測値に従って、認識と微調整を行なう。

##### 【行動計画】

あらかじめ知識として机の位置とロボットの姿勢が環境条件として入力されている。対象物のCADモデルとロボットハンドのモデルにより、動作開始前に個々の対象物に対するハンドの把握可能な位置と姿勢を求める。また行動計画と認識に用いられるモデルの対称性も求め、これらを知識として検索系に記録する。

オペレータが示した目標状態と現在の状態を認識しpick and placeに展開する。対象物を基準として共通の把握姿勢を求める。この時、対象物の対称性の知識を用いて対称な姿勢は共通のものとする。それぞれの状態での机などとの干渉を調べ不

可能なものを排除する。一つの積木のpick and placeの例を図11に示した。

##### 【初期照合】

図4の画像をステレオ対応した結果、頂点が44組得られた。このうち誤対応によるものは30組あった。しかし誤対応のほとんどの場合に線分のなす角が極端な狭角または広角となり、モデルの各頂点に一致しなかった。そのため実際には正しく対応して得られた頂点にのみモデルが当てはめられた。

##### 【検証】

ステレオ対応から得られた頂点 (図12a○) の部分に対するあてはめを見てみる。例えば、誤った姿勢でモデルを当てはめた例 (a)、正しく当てはめた例 (b) に対して評価値はそれぞれ、0.526, 0.750でありこの評価方法で正しいものが選択できる。

対象物の全てのモデルを与えて状態の記述を行なうと図12(c)の結果が得られた。また特定の対象物の探索では(d)の結果が得られた。

比較のため同一の画像に対して単眼による処理を行なった。単眼法では精度初期照合での精度が劣るため認識率が落ち、計算量の増大が予想される。同一の画像に対する処理の結果、7個のブロックのうち5個の対象物は発見できず、初期照合以降の処理にかかった計算時間は約1.5倍であった。

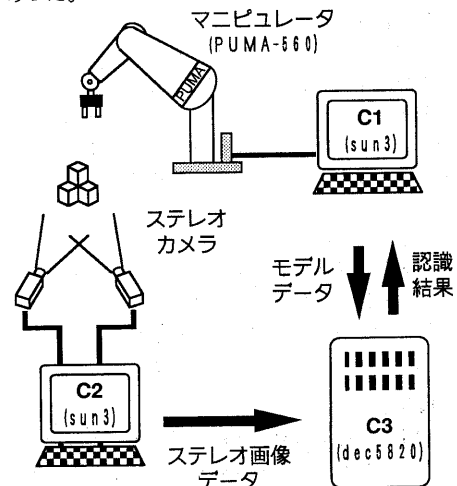


図10 システム構成

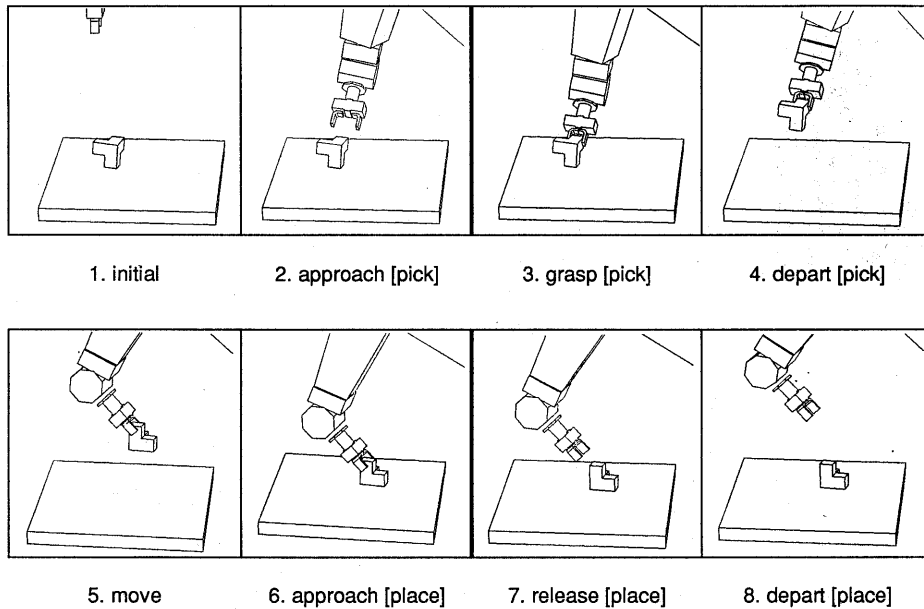


図 1 1 行動計画

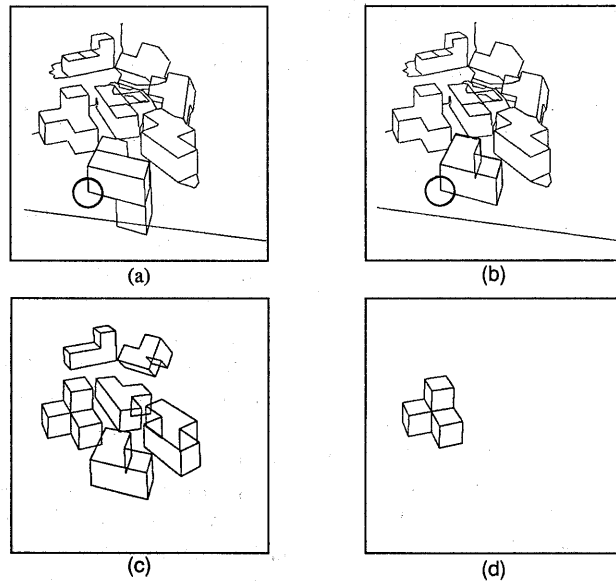


図 1 2 モデルのあてはめと検証

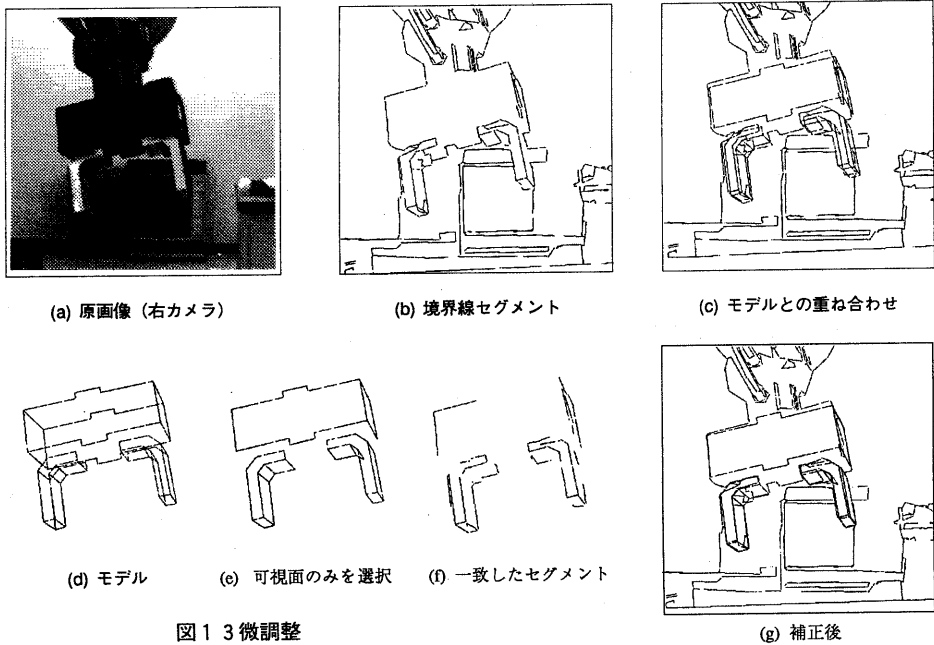


図 1 3 微調整

〔微調整、誤差補正〕

図 1 3 に把握時に接近したハンドの位置補正の例を示す。PUMAでは繰り返し精度は0.5mmであるが、絶対誤差は2～3mmある。

図(a)は接近時のステレオ画像である。ハンドの予測画像を重ねると(c)、3mm程度指定した位置よりもハンドがずれている。そこで対応する部分を抽出する。回転、移動のパラメータを求め、ハンドの位置を補正する。この像を重ねると図13(g)が得られ、1mm程度の精度が得られていることがわかる。

5. まとめ

本システムではステレオ画像をベースとして、目的に従ったステレオビジョンを提供している。動作系の階層ごとにモデルや予測位置の情報が得られるので、これを利用してステレオ法の対応問題の軽減をはかることができた。今後は組み立てなどの複雑な作業に利用していく予定である。

謝辞

このシステムの画像処理プログラムと対象物のモデルを作成していただいた図書館情報大学の福田郁美さんと星野由紀子さんに感謝します。

参考文献

- 1) Albus, J.S., Brains, Behavior, and Robotics, McGraw-Hill, New York, (1981)
- 2) 寺崎、長谷川、高橋 Pick and Place 作業における持ち変え計画 ロボット学会第8回学術講演会予稿集 p105-108 (1990)
- 3) 富田、高橋、ステレオ画像の境界線表現に基づく対応、信学会誌、J71-D,6,p1074-1082(1988)
- 4) Sugimoto, K., Takahashi, H., and Tomita, F., Scene interpretation based on boundary representations of stereo images, Proc. 9th Int. Conf. Pattern Recognition, 155-159, (1988)
- 5) Fischler, M.A., Bolles, R.B., Random sample consensus: A Paradigm for model fitting with applications to image analysis and automated cartography, Comm. ACM, 24-6, p381-395 (1981)