

モデルベースステレオビジョンによるロボットハンドの認識

5D-5

高橋裕信† 寺崎 肇† 田中利晴† 富田文明*
 (†三洋電機筑波研究所) (*電子技術総合研究所)

1. はじめに

予想外の事象にも対処可能なロボット制御を行うため、視覚によってロボットの動作環境と対象物を認識するビジュアルフィードバックについて検討している。ここではロボットハンドの位置検出、特に動作誤差検出のためのステレオ法を適用した例について報告する。

(図1)

ステレオ法を用いている利点は対象物の認識¹⁾と同一の画像を用いることができるため対象物とロボットとの相対距離の評価に適している点と、単眼法に比較して特に奥行き方向で高い精度が期待できる点があげられる。しかし、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は困難である。そのため画像の境界線表現などが提案されている。²⁾

ロボットハンドの認識においては、形状が既知であることを拘束として利用することができる。各画像毎にハンドのモデルを対応付けることにより、モデルを単位としたステレオ対応が可能となる。

またロボットの位置と姿勢は誤差を伴うもののロボットのコントローラから得られるので、各画像中での見え方をあらかじめ予測することができる利点がある。そこでこれらの拘束を利用したモデルベースのステレオについて提案する。

2. アルゴリズム

ハンドのモデル上の各点 $pi(x_i, y_i, z_i)$ の各カメラ画像上での位置 (X_{ni}, Y_{ni}) を求める。各点はロボットの位置と回転の6自由度で変換される。

$$(x'_i, y'_i, z'_i) = \text{Trans}(t_x, t_y, t_z) \text{Rot}(r_x, r_y, r_z) (x_i, y_i, z_i)$$

各パラメータはコントローラの示す値 (tc, rc) と誤差 (dt, dr) で表される。

$$\begin{cases} t_x = tc_x + dt_x \\ t_y = tc_y + dt_y \\ t_z = tc_z + dt_z \end{cases} \quad \begin{cases} r_x = rc_x + dr_x \\ r_y = rc_y + dr_y \\ r_z = rc_z + dr_z \end{cases}$$

次にカメラNの画像上への投影点 (X_{ni}, Y_{ni}) はカメラモデルとしてピンホールカメラを仮定して1次の分数式で表すことができる。

$$\begin{cases} X_{ni} = \frac{m_{xx}x'_i + m_{xy}y'_i + m_{xz}z'_i + m_x}{m_{zx}x'_i + m_{zy}y'_i + m_{zz}z'_i + 1} \\ Y_{ni} = \frac{m_{yx}x'_i + m_{yy}y'_i + m_{yz}z'_i + m_y}{m_{zx}x'_i + m_{zy}y'_i + m_{zz}z'_i + 1} \end{cases}$$

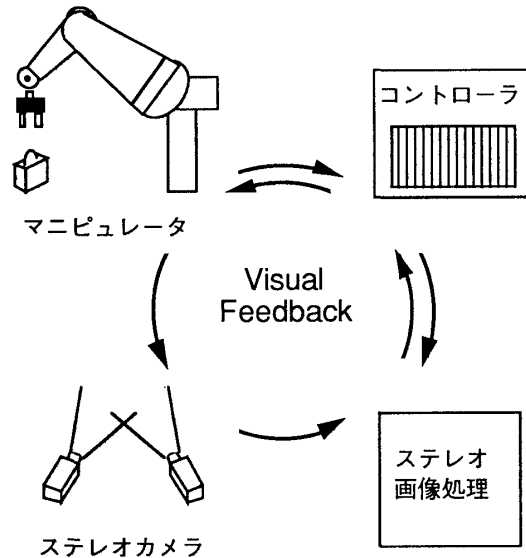


図1 Visual Feedbackによるロボット制御

ここに示された11のパラメータはカメラに固有な値であり、あらかじめ実測により求めることができる。セグメントの可視性は、今回は簡便法として、属する面の位置と法線方向を用いて決定し、面相互のオクルージョン等については考慮していない。

カメラにより観測された画像とモデルから予測される像をセグメント単位で比較し、対応付けを行なう。シーン中の対応するセグメントは近傍 $(< \sigma)$ にあると仮定し、個々の対応の評価値として、図2に示したように、モデル上のセグメントの両端点から対応するセグメントに垂線をおろして囲まれる斜線の部分の面積 S_{ni} を用いる。この面積が小さいほどよく一致していると見ることが

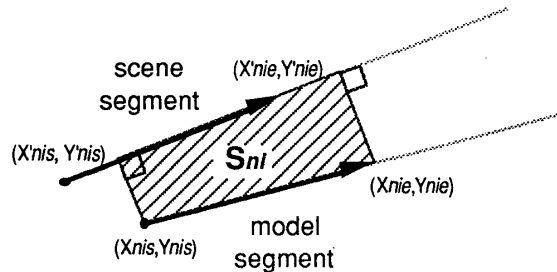


図2 モデルとシーンの対応の評価

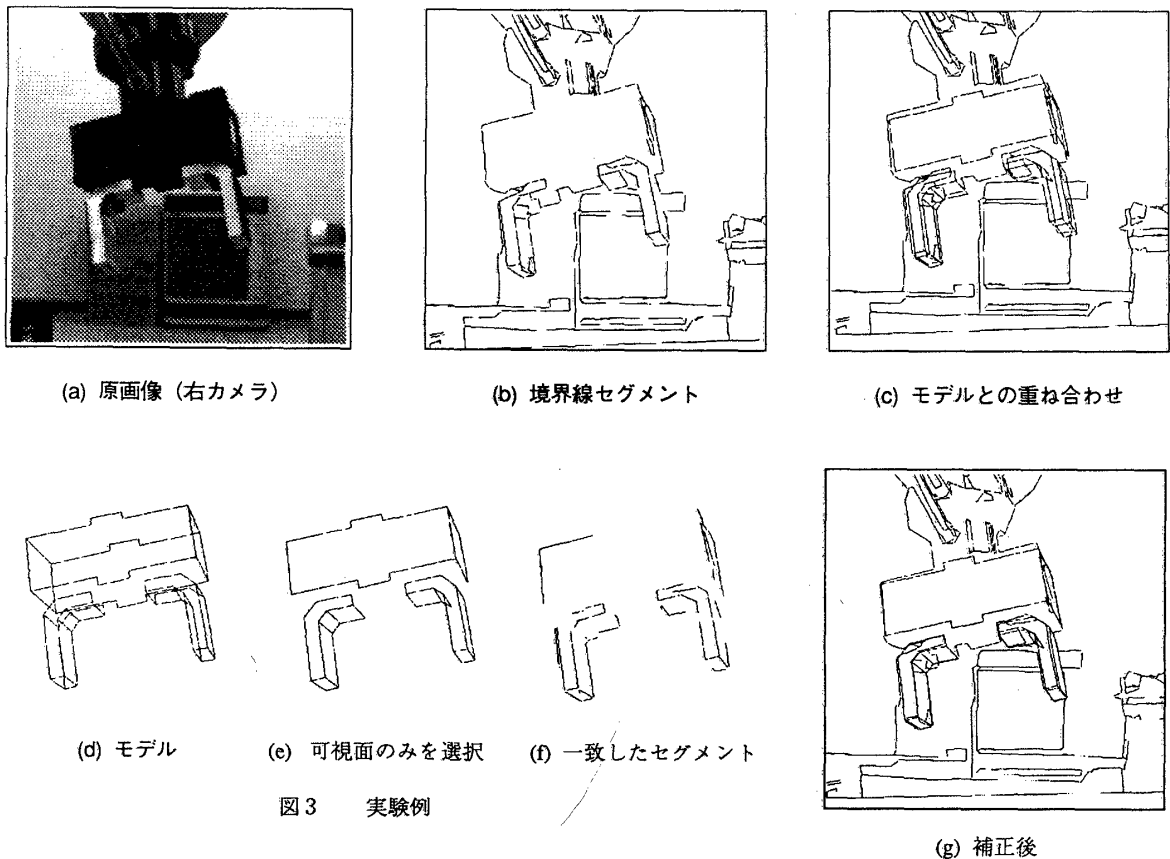


図3 実験例

できる。ここで単純に端点間の距離を評価値として用いない理由は、セグメントの軸方向(1)に対する不要の拘束を避けるためである。各カメラ毎に得られる、対応セグメント全てについて S_{ni} を求め、総和を S とする。

ここで S はロボット制御系の誤差 (dt, dr) の関数とみることができる。 S を最小とするように最小二乗法を設定し、最適値を得ることができる。実際の解法としては各パラメータについて1次偏微分を行ない、繰り返し法により解を求めている。さらに対応の誤りに対処するため、得られた解によりパラメータを補正し、より小さな近傍 ($\sigma < \sigma$) を設定して対応付けを繰り返す。

3. 実験結果

実験には2眼のステレオカメラおよび6自由度垂直多関節型マニピュレータであるPUMA 560を使用した。ロボットを任意の位置、姿勢に移動し、これをテレビカメラで撮像する(図3a)。コントローラよりハンドのパラメータを得て、モデルを変換し(d)、見えるセグメントを予想する(e)。これらを重ねると(c)のように誤差があることがわかる。ロボットの公称繰り返し精度(0.5mm)にロボット座標系のゆがみを考慮して、最大(2.0mm)程度を誤差の許容範囲 σ として与えた。画像、モデル間に対応付けを行ない画像中の対応するセグメントを得る(f)。各対応について最小二乗評価を行なう

ための行列に加算する。さらに他のカメラからの像についても同様の処理を行ない同じ行列に加算する。行列を解いて得られる値をモデルにフィードバックした結果、正確に重なり正しい補正値が得られていることがわかる。(g)

4. 考察

既知のモデルを用いて、ステレオの各画像中での位置を予測することにより、複雑なノイズの多いシーンでも安定して適用することができ、計算コストの点でも有利である。特に左右の画像間で対応するセグメントが得られない場合には、通常のステレオ法での測定はできない。しかしモデルを持つことにより位置および姿勢を決定することができる。

本報告ではロボットハンドの位置、姿勢の予測値が得られる場合について検討した。しかし未知の場合でも、中間表現としてシーンとモデルの表現として境界線表現を用いているため、モデルと画像間の対応付けを行なうことが可能であり、この方法を適用することができる。

参考文献

1. 寺崎他 ステレオによる物体の階層的表現に基づく認識 第39回情報大会
2. 富田、高橋 ステレオ画像の境界線表現に基づく対応 信学論 J71-D p1074(1988)