

交差角度特徴による3次元物体の位置・姿勢認識

北海道大学工学部 ○矢澤龍太 金子俊一 五十嵐悟

要 旨

2次元単眼画像のみを用いて3次元物体の位置・姿勢を認識するための投票アルゴリズムについて述べる。認識モデルから特徴直線(特徴点を連結して生成)の抽出を行い、その交差角度リストを投票特徴とし、1次元Hash表を用いることにより得た、投票値を累積したスコアの評価をして、モデル及びその位置・姿勢を推定する。

1. はじめに

ロボットの知的制御、いわゆる作業教示や協調作業などを行うシステムの実現にとって、物体ハンドリングの状態を的確に捉える画像理解機能が必要とされている。また実環境において稼動するシステムにおいては、ノイズや不良条件に対してロバストな性能が重要となる。モデル照合による投票型物体認識法である Geometric Hashing 法[1](以下、GH法)及びその改良法[2][3]はそのような特長をもつ3次元物体の識別手法である。GH法は、登録及び認識の2段階を持ち、特に認識段階においては、単純な投票-多数決により判定を行うため実時間性を確保できる。ただし、投票空間である Hash 表が大きくなるにつれて投票値(スコア)の累積に時間を要することになり、できるかぎり小さな投票空間を実現する工夫が望まれている。

2. 特徴点抽出

モデル画像からの特徴点抽出の手順を示す。

- [i]モデル画像に Sobel 微分フィルタをかけ、輝度値変化の大きい画素の抽出を行う。
- [ii]特徴点としてモデルのエッジの折れ曲がり点である部分を抽出する。エッジの折れ曲がり点付近では、エッジ勾配方向が激しく変化していると考えられるため、その勾配方向の変化値を計算し、最大値を取る画素を特徴点とする。
- [iii]細線化効果を持つ極大点抽出処理を行う。一般に細線化処理は計算時間を要し、得られる結果にも特有の歪みを含んでしまう。この処理は通常細線化処理ではなく、次節において説明を行う特徴直線抽出するための前処理として、局所的線分の抽出を行う上で、認識モデルのエッジの存在を知るために行うものである。

3. 特徴直線抽出

ここでは、傾き・切片を用いた直線表現に基づく局所的 Hough 変換を用いた最頻値選択アルゴリズムを提案する。モデルから抽出した特徴点近傍において、認識モデルのエッジが存在すると考えられる直線を局所的に抽出を行う。この処理において、抽出されたエッジを「部分線分」と呼ぶ。その後、各々の特徴点において抽出した部分線分を用いることにより、認識モデルのエッジの存在を知り、部分線分の傾き及び切片の値が一致する特徴点どうしを結ぶことにより特徴直線の抽出を行う。図1に処理の流れを示す。

3.1 部分線分

抽出した特徴点を中心にして x, y 座標系内に存在する

輝度値極大点を、 $y > -x, y \geq x$ であれば $\alpha - \beta$ 空間に、 $y \leq -x, y < x$ であれば $\gamma - \delta$ 空間に、それぞれ用意した2次元空間に投票を行う。ここで3つのパラメータを要する表現($\alpha x_i + \beta y_i = \gamma$)に対応する3次元投票空間の代わりに、異なる2つの投票空間を定義し、計算の複雑さを無くし処理の高速を行う。

それぞれの空間における直線の定義は次のようである。

$$\alpha - \beta \text{ 空間} \quad \alpha x_i + y_i = \beta$$

$$\gamma - \delta \text{ 空間} \quad x_i + \gamma y_i = \delta$$

例えば、極大点が、 $y > -x, y \geq x$ にある場合 x, y 座標においての直線は

$$\alpha x_i + y_i = \beta \tag{1}$$

で表される。 $\alpha - \beta$ 空間への投票は

$$\alpha = (\beta - y_i) / x_i \quad \beta = -L, -L+1, \dots, L \tag{2}$$

であるが、 α 及び β の分解能が異なるため実際の投票は

$$\alpha' = (\beta - y_i) / x_i \times L \tag{3}$$

とスケールして行う。これにより、 $\alpha - \beta$ 空間内において多くの得票値をもつ点 (α, β) を得、部分線分候補を見つけ、その中でも最頻値に対応する傾き α 、切片 β の直線を抽出することにより特徴点近傍において、部分線分を抽出する。 $\gamma - \delta$ 空間についても同様の処理を行う。

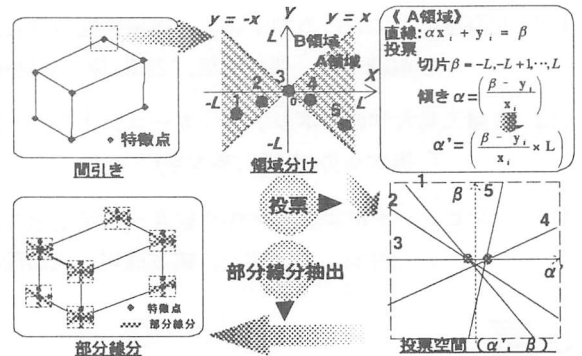


図1 部分線分抽出

3.2 特徴直線(部分線分の連結)

特徴点近傍で抽出された部分線分どうしの連結を行う。任意の特徴点を2個選択し、その間に勾配値(α あるいは γ)の類似する部分線分がある場合には連結する。

4. 交差角度特徴

図2に2次元画像における交差角度を示す。ただし、ここでは対象物の異同に関わらず、異なる見え方を「アスペクト画像」と呼ぶ。交差角度の算出は、あらかじめ用意してお

いた1次元 Hash 表へ、モデルの登録及び認識作業をそれぞれ行うため必要である。各々の作業では、「基準エッジ」(BE)となる特徴直線を、抽出した特徴直線群の中から1本選択する。次にBEと、残りのすべての特徴直線がなす角度値の1次元リストで、あらかじめ用意した1次元 Hash 表に登録又は

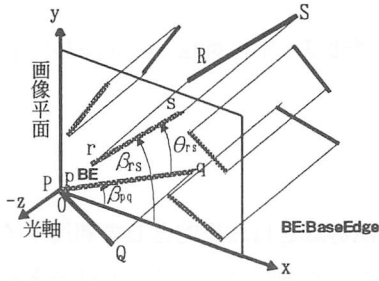


図2 交差角度

投票をする。ここで用いる角度は、登録時及び認識時において、エッジ方向による交差角度の違いを避けるため、常に狭角($\leq \pi/2$)と決める。

第*i*番目のアスペクト画像が N_i 個の特徴直線をもつとき、最大 N_i 個の BE があり得る(あまり短い直線は選ばない)。 N_i-1 個の交差角度値の1次元リストを算出する。これはBEの個数だけ定義される。Hash 表は0度から90度までを幅1度で離散化してある。個々の bin には(モデル画像番号, BE 番号)を登録する。単一の bin には通常複数の項目が多重に登録される。

認識時には、任意の BE に関する交差角度リストによって指示される Hash 表内の bin に登録されている項目を、すべて読みだしソートする。モデル画像番号の最頻値を与えるモデルを認識モデルとし、モデルの姿勢をも推定できる。

5. 位置推定処理

姿勢認識を行った結果に基づき3次元の位置推定を行う処理について述べる。姿勢認識法では、Hash 表へ登録する情報量を、抽出した特徴直線同士のなす交差角度を用いたため、姿勢認識を行った段階では、カメラの光軸回りに対して不変の性質を持つ。そのため、BE とした特徴直線の方向と、その鏡像方向(BE 対称の位置)に対する自由度が残っており、抽出した特徴直線の対応を解くことにより、画像上における認識対象物の位置を推定する。

姿勢認識の結果から、認識画像において BE とした特徴直線と、登録アスペクト画像の特徴直線(be)との対応のみがわかっている。ここで、この2組の端点を合わせることで、位置の推定を行う。

BE の端点座標を $U(x_u, y_u), V(x_v, y_v)$ 、be の端点座標を $u(x_u, y_u), v(x_v, y_v)$ 、登録アスペクト画像における be 以外の特徴直線の端点座標を (x, y) とする。この場合考えられる位置は、図3及び以下で示す4通りである。

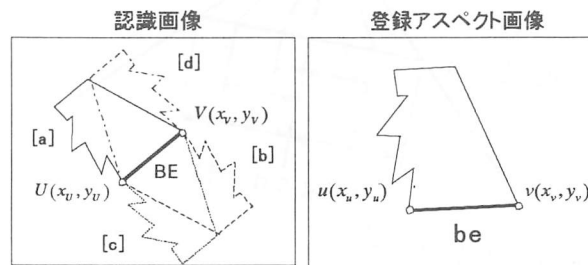


図3 位置の自由度

[a] $u(x_u, y_u)$ が $U(x_u, y_u)$ 、 $v(x_v, y_v)$ が $V(x_v, y_v)$ に対応

[b] $u(x_u, y_u)$ が $V(x_v, y_v)$ 、 $v(x_v, y_v)$ が $U(x_u, y_u)$ に対応

[c] [a] の BE 対称の鏡像反転

[d] [b] の BE 対称の鏡像反転

変換された後の座標を (X'_i, Y'_i) 、回転を表す変換を R 、スケール変換を表す変換行列を s とすると、[a] の場合次式が成り立つ。

[a] $u(x_u, y_u) \rightarrow U(x_u, y_u)$,

$v(x_v, y_v) \rightarrow V(x_v, y_v)$ に対応の場合

$$\begin{pmatrix} X'_i \\ Y'_i \end{pmatrix} = s \cdot R \begin{pmatrix} x_i - x_u \\ y_i - y_u \end{pmatrix} + \begin{pmatrix} x_u \\ y_u \end{pmatrix} \quad (4)$$

ここで、 $R = \begin{pmatrix} \cos(\alpha - \alpha') & -\sin(\alpha - \alpha') \\ \sin(\alpha - \alpha') & \cos(\alpha - \alpha') \end{pmatrix}$

$$\alpha = \tan^{-1} \left(\frac{y_v - y_u}{x_v - x_u} \right) \quad \alpha' = \tan^{-1} \left(\frac{y_v - y_u}{x_v - x_u} \right)$$

$$s = |BE|/|be| = \sqrt{(x_v - x_u)^2 + (y_v - y_u)^2} / \sqrt{(x_v - x_u)^2 + (y_v - y_u)^2}$$

式(4)により、登録アスペクト画像における、be 以外の端点 (x_i, y_i) の変換をし、変換した座標点の近傍に、認識画像の特徴点 (X'_i, Y'_i) が存在すればカウントとする。

カウントがある閾値以上であれば、その位置をモデルの位置と決める。もし、閾値以上のカウントが得られなければ、[b]・[c]・[d] の順番に処理を行い、十分なカウントを得た時点で処理を終了する。

6. 特徴直線抽出実験

図4を用いて実験を行った。図5に部分線分、図6に特徴直線抽出結果を示す。抽出された特徴直線数は16本(誤り6本)であった。また、各々の処理時間は次のようである。Sobel 処理:0.05(秒), 特徴点抽出:0.02(秒), 極大点抽出:0.01(秒), 特徴直線抽出:0.35(秒)。

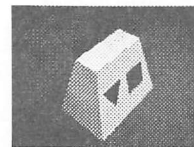


図4 認識モデル



図5 部分線分



図6 特徴直線

7. まとめ

2次元単眼画像を用いた認識対象モデルからの特徴抽出及び特徴直線の抽出を、実時間で行うことができました。この処理は、いわゆる識別と3次元姿勢の推定を含んでいる。今回さらに、姿勢の確定と位置推定処理の定式化を行った。現在、システム開発を行っている。今後の主な課題は、位置誤差の定量的解析と実験による評価である。

参考文献

- [1] Y. Lamdam and H. J. Wolfson: "Geometric Hashing: A general and Efficient Model-based Recognition Scheme" pp.238-249(1988)
- [2] 柴田, 金子, 本多: 画像の認識理解シンポジウム, 2-81/87, (1994)
- [3] 田中, 小野, 他: "空間ポインティングと注目制御型視覚認識", 第10回ヒューマンインタフェースシンポジウム(1994)