

1.はじめに

3次元空間内の物体の運動は、2次元画像上ではエッジ、テクスチャなどの特徴の動きや、オプティカル・フロー(速度場)によって記述される。したがってこの画像の変化は3次元情報を所有しており、画像の変化を解析することによって対象の3次元運動、構造を復元することができる。本研究では複数の画像から速度場を算出し、その速度分布と透視投影関係に基づいて物体の3次元運動、構造を復元することを目的とする。

3次元運動、構造の復元までの手順を3段階に分ける。

- 1)時間的に連続する2枚の濃淡画像から速度場を算出する。
- 2)その画像を被写体のそれぞれに対応する意味のある領域に分割する。
- 3)分割された領域毎に運動パラメータと、距離の情報を算出する。前報¹⁾では第1段階の速度場算出の問題について報告した。本報告では第2段階のセグメンテーション(画像の分割)について述べる。

2.セグメンテーションの概要

分割方法はいくつか提案されている。WohnとWaxman²⁾は物体表面上のテクスチャを利用して画像を分割する方法を提案した。この方法では物体表面上に適当なテクスチャが存在することが仮定されており、分割の分解能も低いようである。一方Adiv³⁾は速度場を利用して連結した速度ベクトルを要素とし、要素間の結合関係を調べて画像を物体に対応する領域に分割する方法を提案している。この方法では物体表面は大まかに平面パッチで近似され、平面パッチの運動の決定問題が、運動を表すパラメータ空間内の最適パラメータの探索問題となるため、膨大なメモリと計算時間が必要となる。本研究では、Adivの方法を参考にしてψ変換を用いたより能率のよい画像の分割方法を提案し、シミュレーションを通してその有効性を検討する。

3.解析方法

3.1 問題の設定

相対運動する互いに独立な複数の物体(剛体)を含むシーンを考える。透視投影を仮定して、図1に示すようにカメラに固定された座標系を(X,Y,Z)、画像座標系を(x,y)とし(x,y軸はX,Y軸に平行)、その原点を(X,Y,Z)=(0,0,1)にとる。今、カメラ座標系においてある物体が並進運動速度T=(Tx,Ty,Tz)、回転運動速度Ω=(Ωx,Ωy,Ωz)で移動したと

する。このとき物体上の点P(X,Y,Z)と画像上での対応点p(x,y)とその速度(α,β)は次式で与えられる。

$$x = \frac{X}{Z}, \quad y = \frac{Y}{Z} \quad (1)$$

$$\alpha = \left\{ x \frac{T_z}{Z} - \frac{T_x}{Z} \right\} + \left\{ xy\Omega_x - (1+x^2)\Omega_y + y\Omega_z \right\}$$

$$\beta = \left\{ y \frac{T_z}{Z} - \frac{T_y}{Z} \right\} + \left\{ (1+y^2)\Omega_x - xy\Omega_y - x\Omega_z \right\}$$

… (2)

式(2)は同じ物体上の全ての点に対応する画像領域内で成り立つので、式(2)を最もよく成立させる運動パラメータT, Ωを求めれば、物体の速度が決まり、これらのパラメータを用いて各点の深さZを求めれば元の3次元構造が決定できる。そのためには、まず画像を独立した物体に対応する領域に分割する必要がある。

3.2 ψ変換

分割問題に対しては、画像をある大きさの正方形に分割し、その正方形に対応する物体表面を平面のパッチと仮定することで、曲面を多面体近似する方法を考える。平面の式を

$$k_x X + k_y Y + k_z Z = 1 \quad (3)$$

とし(k_x, k_y, k_zは0でない実数)、(1)、(3)を使って(2)からZを消去すると次式が得られる。

$$\alpha = a_1 + a_2 x + a_3 y + a_7 x^2 + a_8 xy \quad (4a)$$

$$\beta = a_4 + a_5 x + a_6 y + a_7 xy + a_8 y^2 \quad (4b)$$

ここで、例えばa₁ = Ω_y + k_zT_xである。これはψ変換と呼ばれるもので、a₁…a₈の8つのパラメータで物体の3次元運動を画像上の運動で表す。ある平面の投影結果である画像上の領域内の速度ベクトルは同じψ変換によって表される。したがって同じψ変換を近似的に満たすような平面パッチによって、画像を分割するのである。

パラメータの決定には最小2乗法を用いる。最小化する評価関数は次式で与えられる。

$$E(a_1 \dots a_8) = \sum_{i=1}^n W_i [(\alpha_i - a_1 - a_2 x_i - a_3 y_i - a_7 x_i^2 - a_8 x_i y_i)^2 + (\beta_i - a_4 - a_5 x_i - a_6 y_i - a_7 x_i y_i - a_8 y_i^2)^2] \quad (5)$$

ここで、nは平面パッチに含まれる画素数、W_iは速度

の信頼性の重みである。これを $a_1 \dots a_8$ に関して偏微分して0とおくことによって $a_1 \dots a_8$ に関する正規方程式が得られる。この連立方程式を解くことによって $a_1 \dots a_8$ が得られる。

3.3 セグメンテーション

次に同じ Ψ 変換を与える平面パッチを結合し、画像をその集合で分割する。まず、最初の平面パッチを P_0 とする。次に、 P_0 に隣接した平面パッチ P と P_0 の距離

$$\sqrt{(a_{1P_0} - a_{1P})^2 + \dots + (a_{8P_0} - a_{8P})^2} \quad (6)$$

がしきい値 h 未満の場合は P を P_0 とし、 h 以上の場合は P を P_1 とする。そして次に隣接した平面パッチでも P_0 と(P_1 が存在する場合は P_0 、 P_1 の両方と)の距離を求め、同様の処理をする。これを全ての平面パッチに対して行い、同じ Ψ 変換を与えるパッチの集合によって画像を分割することができる。

4. シミュレーションによる検討

提案した Ψ 変換によるセグメンテーションの有効性を調べるため、シミュレーションを行った。空間内に、
平面： $Z = 50Y + 100$ 、

$$\text{楕円面：}(X-2)^2 + [(Y-2)/4]^2 + (Z-5)^2 = 1$$

が存在する環境の中をカメラが $T=(0,0,0.2,1)$ 、 $\Omega=(0,0,0)$ で運動するとして、 64×64 画素に対応する速度場を計算機で算出した。その結果を図2に示す。画像の下半分の平面と右側の楕円面の運動が表されている。また、背景部分の信頼性の重みを0、それ以外を1とした。

次に速度場を 4×4 画素に区切り、それぞれを1つの平面パッチとした。各パッチの Ψ 変換を求め、画像を分割した。その結果を図3に示す。(a)は $h=0.07$ の結果であり、全体が20種類の領域に分割されている。この例では平面は正確に評価され、楕円面は多面体近似されており、分割が有効に行われていることが分かる。"?"の部分は 4×4 画素の中に背景部分が含まれ、最小2乗法を適用するために必要な数の速度がない領域である。

ここで問題となるのは、しきい値 h の決定である。(b)は $h=0.41$ としたときの結果であり、全体が10種類の領域に分割されている。このようにしきい値を大きくすると粗く分割され、多面体が1平面と評価されてしまうことがわかる。この問題は3次元情報復元問題にも関係するので、適当なしきい値の選択方法を求める必要がある。

5. おわりに

Ψ 変換を利用した画像の分割方法を提案し、シミュレーションを通して、その有効性を調べた。今後はこれらの結果を利用して独立した物体を分離し、運動パラメータ、距離情報の復元を行う予定である。

<参考文献>

- 1) 菅野, 波川, 五十嵐: 1991年精工学会春季大会講演論文集 379-380(1991)
- 2) K. Wahn, A.M. Waxman: Comput. Graphics Image Process. 49, 127-151(1990)
- 3) G. Adiv: IEEE Trans. Pattern Anal. Mach. Intell. PAMI-7, No.4(1985)

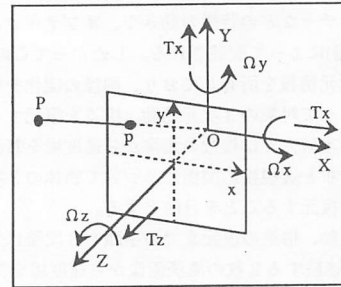


図1 カメラ座標系と画像の関係

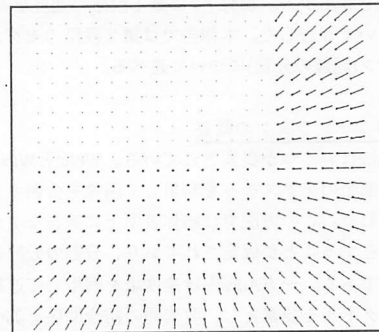
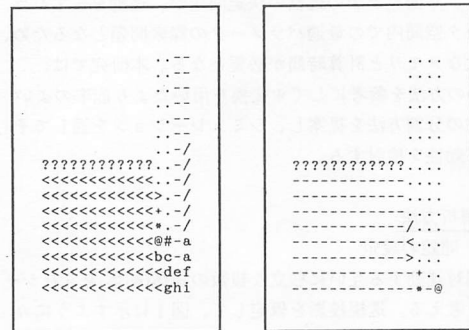


図2 速度場の例



(a) $h=0.07$ での結果 (b) $h=0.41$ での結果

図3 分割結果