

3次元モデルによる2次元画像検索

3E-5

矢野尾 一男[†] 坂内 正夫[†][†]東京大学生産技術研究所

1 はじめに

今までに、絵画の検索などを目的とした簡略画によるインタラクティブな画像検索などの研究が行なわれている。この簡略画による画像検索を拡張してユーザが直接3Dモデルを与えれば、視点の位置に依存しないより柔軟な検索を行なうことができると考えられる。本稿では、現在構築中のシステムにおける、3Dモデルと2次元画像との照合による検索機構について述べる。

2 対応関係を利用したマッチング

3Dモデルを用いる利点として、検索結果が視点の位置に依存しないことが挙げられる。その反面、視点の位置に依存しないマッチング機構が必要となる。対象として2D画像を想定しているため、アスペクト法などの距離画像や法線ベクトル図などの3次元的数据を必要とする手法を用いることはできない。

ユーザが生成したモデルから回りから見た投影像群を生成し、それぞれの投影像についてマッチングを行なう方法も考えられるが、どの程度の数の投影像を生成する必要があるか分からず、またマッチングには回転、拡大縮小、平行移動に不変な手法を用いる必要があるという欠点がある。

そこで、対象となる2D画像中の物体を表す領域と3Dモデルとの対応が領域の色などの属性によって領域とモデルとの対応を割り出し、そのモデルと領域の対応から視点の位置を決定してマッチングを行なう手法を提案する(図1)。そのプロセスは以下のように表すことができる。

1. 予め2次元画像中を領域分割しておく。
2. ユーザが記述した3Dシーンから、3Dモ

デルの位置、色を調べる。

3. オブジェクト同士の対応を調べる。
4. それぞれのオブジェクトの位置を調べて、矛盾がないかどうか検査し、視線方向を割り出す。
5. 矛盾がないとき、得られた視線方向から見たラスタイメージと、2次元画像との2次元画像同士のマッチングを行なう。

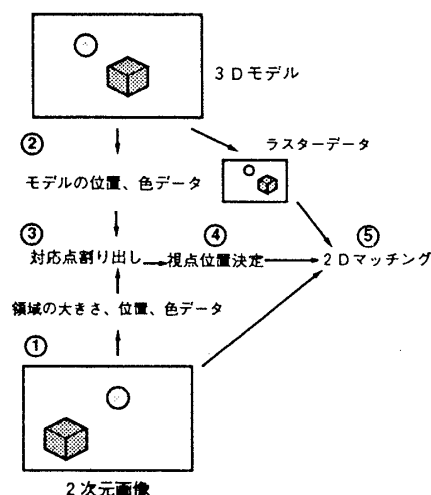


図1: 対応関係を利用したマッチング

3 視点決定

数点の対応関係が分かっているならば、3次元位置と2次元投影位置の対応によって視点位置を算出することができる[1]。本研究では、ユーザが厳密に正確なモデルを与えることは期待できないため、できるだけロバストな手法を採用する必要がある。そこで、非線形性が比較的少ない正投影を仮定した。

3次元モデルの世界座標系内での位置を (x_i, y_i, z_i) 、それに対応する画像中の領域の位置を (u_i, v_i) 回転を表す 3×3 の正規直行行

列を $M = \{m_{ij}\}$ 、平行移動を表すベクトルを (s, t) 、スケールファクタを f とおくと、

$$\begin{aligned} u_i &= f(m_{11}x_i + m_{21}y_i + m_{31}z_i + s) \\ v_i &= f(m_{12}x_i + m_{22}y_i + m_{32}z_i + t) \end{aligned} \quad (1)$$

という関係が成り立つ。 s を独立に求める必要はないので、4点以上与えられれば、線形連立方程式(1)は解くことができる。

上述の線形な解法は高速であるが、実際は3つの独立変数 (α, β, γ) で表される M のそれぞれの成分を独立変数として扱っているため、雑音に弱いという欠点がある。そこで、式(1)を

$$u_i = \phi_i(\alpha, \beta, \gamma, s, f) \quad (2)$$

$$v_i = \psi_i(\alpha, \beta, \gamma, t, f) \quad (3)$$

と表し、最小自乗法を用いて解を求めることにする。微小変化 $(\Delta\alpha, \Delta\beta, \Delta\gamma, \Delta s, \Delta t)$ を Δv 、式(2)の右辺のヤコビアン行列を Φ とおくと、以下の式が成り立つ。

$$\Phi \Delta v = \begin{pmatrix} u_1 - \phi_1(\alpha, \beta, \gamma, s, f) \\ \dots\dots\dots \\ u_n - \phi_n(\alpha, \beta, \gamma, s, f) \end{pmatrix} \quad (4)$$

式(3)についても同様の式が得られ、これらをまとめて最小自乗法の計算を Δv が十分小さくなるまで繰り返すことによって解が得られる。初期値としては、式(1)によって得られる近似解を用いる。

4 実験結果

前節で述べた照合手法を用い、4つの球、4つの楕円体、5つの異なった物体からなる3種類のシーンについて、その3次元モデルとそのシーンを周囲60方向からみた投影画像との照合を行い、表1のような結果を得た。楕円体のようにいびつな形になると隠蔽による誤差が影響し、また隠蔽によって計算に必要な4組のペアが生成できずに視点計算が失敗しているものもある。しかし、相関が0.5以上だと主観的にはほぼ一致しているように見えることを考慮すると、全体的に見て隠蔽に対しロバストであるといえるだろう。

また、4つの球の位置にノイズを加え、視点位置計算の安定性を検証した(図2)。縦軸

は算出した視点位置から見た投影画像と対象画像との相関、横軸はノイズの合計である。

相関が0となっているのが、解が収束しなかった場合である。これより、ノイズが小さい場合、収束性は問題ないことがわかる。

表1: 投影画像と3Dモデルの照合実験結果

照合結果	4つの球	4楕円体	5物体
失敗(相関0-0.5)	0	0	8
ほぼ一致(0.5-0.7)	0	14	8
一致(0.8-1)	60	46	44

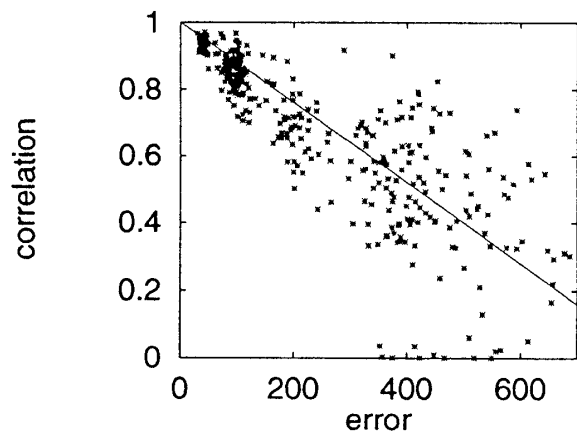


図2: モデル誤差と投影画像間の相関との関係

5 おわりに

本稿で述べた方法は、領域とモデルとの対応をとる初期マッチ段階で組合せ爆発を起こる可能性があるという欠点があるが、モデルと画像との照合は高速であり、モデルの不正確さ、隠蔽によるノイズにも耐え得ることが実験によって明らかになった。今後の課題として、初期マッチの効率化、主観を考慮した評価手法などが考えられる。

文献

- [1] D. G. Lowe, Three-Dimensional Object Recognition from Single Two-Dimensional Images, Artificial Intelligence 31 (1987) 355-395