シルエットからの関節物体三次元姿勢推定法

A 3-D Pose Estimation of an Articulated Object from its Silhouette Image

亀田 能成　　美濃 導彦　　池田 克夫

Yoshinari Kameda　Michihiko Minoh　Katsuo Ikeda

京都大学工学部 †

Faculty of Engineering, Kyoto University

1　はじめに

画像認識の研究において，対象物が剛体でない分野にも適用できる汎用的モデルマッチング技術が求められるようになりつつある．その適用分野として，人間の姿勢の認識，行動解析，またマンマシンインターフェースとしての手の形状認識などが挙げられる．これらの分野では，対象は関節を複数もつ関節物体である．

対象物体を具体的に手指や人体に限った場合については，様々な研究が行われている[1, 2, 3]．これらの研究では，計算機側の持つモデルのデータ構造が対象物体に合わせて作られている．

それに対し，本研究では関節物体一般に適用できるモデルのデータ構造を与え，それに合わせたモデルマッチング方法を提案する．

モデルマッチングをどのような特徴空間で行うかは大きな問題である．最も簡単に入力できる画像は輝度画像であり，それに対してよく用いられる特徴量はエッジ情報とシルエット情報である．エッジ情報が画像の撮影状況や抽出方法に依存するのに対して，シルエット情報はカメラパラメータにのみ依存する．そこで，本稿のモデルマッチング方法では，カメラパラメータを既知とし，シルエット情報を利用することにする．

†京都大学工学部情報工学科

〒606 京都府京都市左京区吉田本町

2　関節物体のモデル記述

関節物体は，複数の剛体部品が自由度を持つ関節によって連結されているものである．多くの関節物体において，剛体部品は木状に連結されている．この点に注目して，本稿では関節物体のモデル記述方法として，部品位置決定木というモデル記述を用いる[4]．部品位置決定木においては，各ノードが剛体部品に相当し，各アークが関節に相当する．ノードiのデータ構造を図1に示す．ノードは，それに対応する剛体部品の立体形状をパッチモデルとして保持している．これにより，従来の円筒形の部品表現[1, 2, 3]に比べて，高いモデル記述能力を有している．

部品位置決定木中の根ノードに相当する剛体部品の空間的位置が決定され，かつ各アークにおける関節の角度がすべて決定されれば，モデルの三次元姿勢が一意に定まる．

3　マッチング処理の順序

本稿で提案するマッチング手法では，部品位置決定木中の各ノードを一つ一つ逐次的に処理していく．

まず，部品位置決定木中の全てのノードを三つのノード集合のいずれかに属させる．初期状態では，既定ノード集合Fは空集合とし，先端ノード集合Nには根ノードのみが含まれ，残りの全てのノードは未定ノード集合Rに含まれる

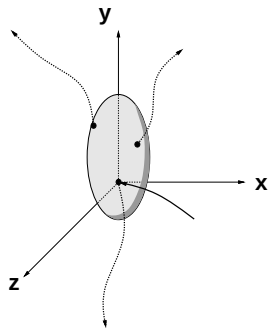


図 1: ノードのデータ構造

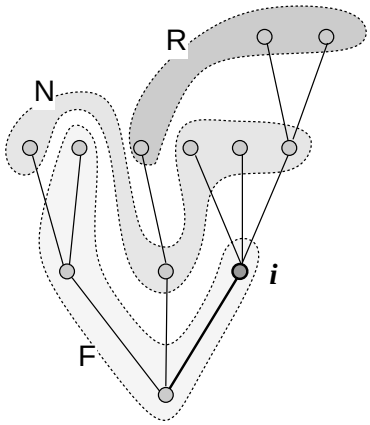


図 2: マッチング処理の実行順序

ものとする．マッチング処理は次のような順序で行われる．

1. まず，部品位置決定木中の根ノードに相当する剛体部品の位置を決定する．これは，シルエット画像撮影時の制約条件から推定されるものとする．この根ノードの番号を i とする．
2. ノード i を先端ノード集合 N から既定ノード集合 F に移し，ノード i に隣接する未定ノード集合 R 中の全てのノードを先端ノード集合 N に移す．
3. 先端ノード集合 N から任意のノード i を選び，シルエットを用いてその親ノードとの間の関節角度を後述の方法 (4節参照) を用いて推定する．もし先端ノード集合 N が空集合ならマッチング終了である．そうでなければ 2. に戻る。

処理中の部品位置決定木の様子を図 2に示す．

4　シルエットを用いたマッチング方法

シルエットを画像からの情報としてモデルマッチングを行う場合，画像平面上でのシルエット領域 S とモデルの投影によってできる領域 M との排他的論理和 $E = S \oplus M$ を最小にすることが目標となる [5]．

実際には，E を最小にするような関節角度を全て同時に解くことは困難である．そこで，E の最小化問題を各剛体部品とシルエットとの関係に関する問題で近似することを考える．いま，あるノード i に対応する剛体部品の画像平面上への投影領域を m_i とする．このとき，次式を満たすようなノード i とその親ノードとの間の関節角度を求める．

$$\begin{aligned} &\min S \oplus (m_i \cap \overline{I(i)}) \\ &= \min\{(S \cap \overline{m_i}) \cup (\overline{S} \cap m_i \cap \overline{I(i)})\} \end{aligned} \tag{1}$$

ただし，$I(i)$ はノード i の親ノードから根ノードへ至る経路上に存在する剛体部品全ての画像平面上への投影領域の論理和を意味する．前節で述べたマッチング処理順序により，この経路上の剛体部品の位置は全て事前に決定されている．上式による近似は経路上にない剛体部品の位置推定結果を必要としないので，前節のマッチング手順 2. において，先端ノード集合 N 中の全てのノードを同時に並列処理することが可能となる．式 (1) の意味を図式化すると，図 3のようになる．

もし何らかの要因によって部品位置決定木の各ノードのマッチング処理順序が予め定められている場合には，式 (1) の代わりに次式を用いて近似すればよい．

$$\min S \oplus \{m_i \cap (\overline{\bigcup_{k \in F} m_k})\} \tag{2}$$

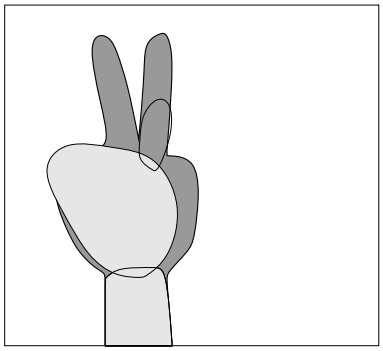


図 3: 式 (1) の図式化

5　実験と考察

対象物体として，典型的な関節物体である人間の手指と人体を取り上げる．本稿で提案するモデルマッチング方法を，式 (1) 及び式 (2) の両方について実験した．

5.1　手指

ある人物の右手指を対象とし，21 枚のサンプルに対して実験を行った．手指はいずれも腕の中心軸がカメラ軸と直交する状態で撮影されている．画像の大きさは横 600 画素で縦 420 画素である．また，カメラパラメータは既知である．

手指モデルは，17 個の剛体部品から構成されている．腕の部分に相当する剛体部品が，部品位置決定木の根ノードに対応している．関節の自由度は 1 ないし 2 であり，モデル全体で 22 の自由度がある．マッチング処理の際の根ノードでの剛体部品位置決定は，別の方法 [4] によって自動的に行った．

実験の結果得られた E の平均値を表 1に示す．式 (2) の方法におけるマッチング処理順序は適当に順序列を与えた．今回の実験においては，式 (1) の方法が式 (2) の方法よりよい結果を出している．これは，式 (2) の方法に与えた順序列があまりよくなかったことを示している．現在のところ，式 (1) の方法より常によい結果を与える順序列をシステマティックに見つけることは難しい．式 (1) の方法はマッチング処理順序に関係なく常に同一シルエットに対して同一姿勢を推定する．

実験結果の一例を図 4に示す．

表 1: 手指に対する実験結果

	$E[\mathrm{mm}^2]$	E[画素]	E/S
式 (1)	2364	9456	0.143
式 (2)	3044	12176	0.196

原画像　　シルエット画像

式 (1) の方法 $\frac{E}{S} = 0.092573$　　式 (2) の方法 $\frac{E}{S} = 0.103474$

図 4: 手指に対する実験例

5.2　人体

ある人物の全身を対象とし，17 枚のサンプルに対して実験を行った．人体はいずれも首の中心軸がカメラ軸と直交する状態で撮影されているものとしている．画像の大きさは全て横 320 画素で縦 360 画素である．また，カメラパラメータは既知である．

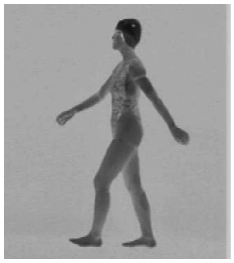
人体モデルは，17 個の剛体部品から構成されている．首の部分に相当する剛体部品が，部品位置決定木の根ノードに対応している．関節の自由度は 1 ないし 3 であり，モデル全体で 40 の自由度がある．マッチング処理の際の根ノードでの剛体部品位置決定は，手指の場合と同様に別法によって自動的に行った．

実験の結果得られた E の平均値を表 2に示す．式 (2) の方法においては，マッチング処理順序には適当な順序列を与えている．実験結果の一例を図 5に示す．

人体に対する実験では，式 (1) の方法と式 (2) の方法はほぼ同等の結果である．しかしながら，このようなマッチング処理順序を何らかの基準で決定する方法については，現在のところ解明されていない．それだけに，式 (1) の方法が安定した姿勢推定結果を得るためには望ましいと結論できる．

表 2: 人体に対する実験結果

	$E[\mathrm{cm}^2]$	E[画素]	E/S
式 (1)	801.4	4093	0.241
式 (2)	767.4	3920	0.232



原画像

シルエット画像

式 (1) の方法
$\frac{E}{S} = 0.190729$

式 (2) の方法
$\frac{E}{S} = 0.219281$

図 5: 人体に対する実験例

6 結論

シルエットを入力画像から得て，関節物体の三次元姿勢をモデルマッチングによって推定する方法を提案し，実験によりその能力を確かめた．

現在の方法ではシルエットとモデルの投影領域との重なりの大きさをマッチング評価の基準としているので，E の値がかなり小さくても三次元の形状としては似ていない場合がある．そこで，今後の課題としては，シルエットの輪郭の曲率など幾何形状に関する特徴量を利用してマッチング方法の改良を行うことが考えられる．また，マッチングを行う際の補完的な情報として，エッジ情報を利用することも検討課題である．

参考文献

[1] C. I. Attwood, G. D. Sullivan, and K. Baker: Model-based Recognition of Human Posture Using Single Synthetic Images, *Proceedings of the Fifth Alvey Vision Conference* (1989), 25–30.

[2] A. C. downston and H. Drouet: Model-based Image Analysis for Unconstrained Human Upperbody Motion, *International Conference on Image Processing and its Applications* (1992), 274–277.

[3] 三井章友, 青木正喜: 体操をしている人間の人体各部の追跡, 第 20 回画電学全大 (1992–6), 65–68.

[4] 亀田能成, 美濃導彦, 池田克夫: シルエットを利用した手指の三次元形状推定法, MIRU'92, 情処シンポジウム論文集, **92-3** (1992), 239–246.

[5] 亀田能成, 美濃導彦, 池田克夫: シルエットを利用した手指の三次元形状推定法, 第 45 回情処全大, **2** (1992), 389–390.