論　文

シルエット画像からの関節物体の姿勢推定法

Nonmember 亀田能成 †　　*Member* 美濃導彦†　　*Member* 池田克夫†

A Pose Estimation Method for an Articulated Object from its Silhouette Image

Yoshinari KAMEDA, *Nonmember*, Michihiko MINOH, Katsuo IKEDA, *Members*

あらまし　一枚のシルエット画像から，その画像中に写された関節物体の姿勢を，モデルマッチングによって推定する方法を提案する．対象となる関節物体モデルをあらかじめ計算機に構築しておく．関節物体モデルが対象物体を正確に反映するものならば，特徴点に基づく逆運動学的な解法やマッチング処理の内部にヒューリスティックな知識を埋め込まなくとも，本方法により姿勢推定が可能であることを示す．本稿で取り扱う関節物体は，部品が関節によって繋がった木構造をしているものと仮定する．その姿勢は関節のすべての角度を決定することで定まる．CG 合成画像に対して本方法を適用し，その能力と特徴を評価した後，実際の人体についても実験を行い有効性を確認した．

キーワード　シルエット画像, 関節物体, 姿勢推定, モデルマッチング

1. まえがき

計算機の応用分野で，関節物体の姿勢推定に対する要求が高まりつつある．ここでいう姿勢とは，対象物体の関節の曲がり具合がどのような状態にあるかを規定するものである．仮想世界の構築を行うためには人体全体に対する姿勢推定が必要であるし，機械の遠隔操作を想定するなら，手指の姿勢推定が重要となる．また，計算機の端末にカメラを設置して計算機利用者の行動を数値化できれば，ヒューマン・インターフェースの一助となる．

特に対象を人体や手指に限定する場合，関節物体の姿勢推定を行う一つのアプローチとして，人体や手指の円筒近似できる部分をリボンとして画像から抽出し，段階的に統合していく方法が考えられる．

倉掛らは明確な人体モデルを持たず，画像中のエッジからリボンを抽出し，そのリボンの接続関係を考察することによって平面内での関節物体の動作を認識している[1]．

また，ユーニパンらはカラー画像を用いて肌色領域を抽出し，領域の形状からどの領域が腕であるかを判別した後，画像平面内での領域間の関係をもとに腕の姿勢推定を行っている[2]．

しかし，これらの手法では立体的な対象物モデルを持たないので，三次元的に姿勢を再現することが難しい．これに対して，関節物体の空間的な形状，姿勢，動作を直接認識するため，さまざまな研究が行われてきた．実用に近い方法では，データグローブの装着や指先の着色などが挙げられる[3]．これらの方法の特徴は，対象となる人体に特別なマーカーや装置を付帯させることにある．しかしながら，こうした装着型のデバイスは認識対象である利用者に負担を強いる．また，遠隔地の対象物体の姿勢推定を行うような状況にあっては，そうした工夫自体が無理な場合があり，特別な装具に頼らない姿勢推定法が求められる．

そこで，装着型のデバイスを用いるかわりに，人体に対応する立体的で明確なモデルを構築し，モデルマッチング手法により問題を解決しようとする研究が行われている．この場合，人体に負担は強いられない．しかし，従来の研究では，モデルマッチング処理の実現に際して，モデルに関する先見的な知識が多くの場合あらかじめ組み込まれている．すなわち，画像特徴とモデルの特定の部分との対応づけがヒューリスティックな知識によって行われている．最もよく見られるのが，人体モデルや手指モデルを用いる場合に，人体や手指の円筒近似できる部分をリボンないし長方形領域という形の画像特徴に対応させる方法である[4],[5]．

例えば，木本らは，人体をスティックモデルで表現

† 京都大学工学部, 京都市
Faculty of Engineering, Kyoto University, Kyoto-shi, 606–01, Japan

し，対象のシルエット領域を心線化してモデルとの対応づけを行い，姿勢の解析を行っている[6]．一方，ユーニバンらはMRFを用い，モデルの構造を利用して長方形領域の抽出を効率よく行っているが，モデル自身は二次元内で動くことしか考慮されていない[7]．

画像特徴に領域の形状を用いない研究としては，石井らが人体の動作パターンの抽出を目的として，肌色領域の位置のみに注目し，ステレオ画像を使用する研究が挙げられる[8]．しかし，この研究では肌色領域がモデルのどこに対応するかがやはりあらかじめ考察により解決されている．島田らは手指に対する姿勢推定方法を提案している[9]が，彼らの方法は画像上の突起領域が指先に対応するという対象物体に関する知見をマッチング処理に採り入れている．

このような画像の特徴とモデルとの対応づけをマッチング処理内部にあらかじめ組み込んでおく方法はマッチング処理の高速化や効率化には有利であるが，その一方でアルゴリズムの汎用性が犠牲となる．

そこで，本稿ではモデルを正確に表現することでマッチング処理を対象物体に依存しないようにし，できるだけその一般的な定式化と解法を目指す．モデルと対象物体との合致度を高くすれば，画像からの特徴量はそれほど必要ないとの観点に立ち，シルエット画像を前提にしたマッチング処理方法について提案する．

以下，2章では対象とする関節物体について述べ，それに対するモデルを定義する．3章では入力情報として用いるシルエット画像などについて説明し，4章では本研究における姿勢推定問題の定式化について述べ，5章で今回提案する方法について述べる．さらに6章で本方法による実験とその結果への考察を行う．最後に，7章で結論を述べる．

2. 関節物体とそのモデル

本稿でいう関節物体とは，関節角度によって，かつそれによってのみその姿勢が一意に決定される物体をいう．

関節物体をモデル化するに際して，本稿では以下の条件を設ける．これらは多くの関節物体に見られる共通点であり，本研究の一般性を失うものではない．

- 一つの関節は最大三軸までの関節回転軸の組合せで実現される．
- 関節物体を構成する体節は関節によって互いに接続される．
- 各体節は剛体である．
- 体節間の関節部分の形状は，その関節角度によってのみ決定される．
- 各体節は木構造状に連結されている．

関節物体モデルは，木構造のデータ構造で表現する．このデータ構造を姿勢決定木と呼ぶ．姿勢決定木の各ノードは関節物体の一つの体節に相当し，その体節の幾何形状ならびに親ノードとの間の関節に関する情報を持つ．表1にその属性の一覧を示す．幾何形状はさらに剛体部と関節角度依存部に分けられ，後者は関節角度の値に応じて変形する．ノードiのローカル座標系は，その原点が親ノード座標系内の関節接続点に一致するように親ノードと接続する．関節の回転は表1中の$k = 0, 1, 2$の順に行われる．関節角度$a_i(k)$は関節可動範囲を逸脱しない．

表1 姿勢決定木のノードのデータ構造

ノード番号	$i, i \geqq 0$
親体節番号	p_i
幾何形状	ノードiのサーフェスをパッチで表現 ($\text{ノード}\ i\ \text{のローカル座標系で記述}$)
関節接続点	親ノードの局所座標系内の一点
関節回転軸	$A_i(k), 0 \leqq k < n_i$ (n_iは軸数)
関節角度	$a_i(k), 0 \leqq k < n_i$ (n_iは軸数)
関節可動範囲	$r_low_i(k) \leqq a_i(k) \leqq r_high_i(k)$

なお，姿勢決定木中の根ノードrにおいて，関節接続点は三次元世界座標系内での根ノードの位置を表す．つまり，本稿では，関節物体の存在位置は根ノードの関節接続点によって規定される．また，その関節角度$\mathbf{a}_r = \{a_r(k) | 0 \leqq k < n_r\}$は三次元世界座標系に対する回転を表す．

従来の研究[6],[8]では各体節の幾何形状を円筒ないしそれに類するもので近似していた．これに対して，本研究では各体節の幾何形状を表現するのにパッチを用いているので，複雑な幾何形状の対象物体であっても精度のよい近似表現が可能である．

人体を対象にした関節物体モデルの幾何形状の表現例を図1,2に示す．図1は人体の胸部に対応するノードである．淡色表示されたパッチが体節の剛体部に相当する．濃色表示されたパッチは親ノードとの間の関節角度依存部に対応し，関節角度の変化に従って変形する．

3. 入力情報

関節物体の姿勢推定において，一般には画像中に関節物体が存在するかどうかを調べ，存在する場合にはその位置を抽出した上で姿勢の推定を行う．しかし，

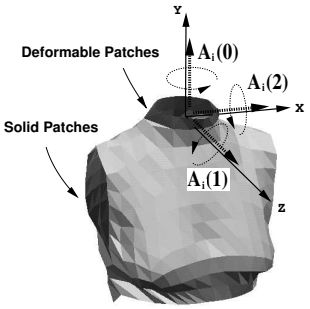

図 1　胸部に対応するノード
Fig. 1　A node corresponding to chest

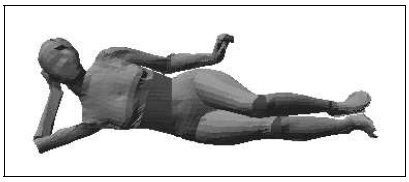

図 2　女性モデル
Fig. 2　A female model

本研究は姿勢推定を主眼とするので，対象とする関節物体の種類 ・構造が既知であるとする．関節物体の存在位置を認識することは本稿の目的外である．そこで，以後本研究では画像中の関節物体の位置および撮影時のカメラパラメータについては既知であるとする．つまり，マッチング処理に先だって，関節物体モデルの根ノードの関節接続点があらかじめ与えられているものとする．本稿の関節物体モデルは木構造であるため，任意のノードを根ノードにすることが可能である．そこで，最も位置を正確に抽出できる体節の位置を求め，その体節を根ノードとする関節物体モデルを用意することができる．

なお，対象となる関節物体以外の物体によるシルエットは画像中に存在しない．

4. 姿勢推定問題

本研究では，対象物体の姿勢に関する手がかりとして二次元情報であるシルエット画像のみを用いているから，姿勢推定の評価もまた二次元上で行う．

いま，あるノード i の関節角度を $\mathbf{a}_i = \{a_i(k) | 0 \leqq k < n_i\}$ で表すと，関節物体モデルの姿勢は $\{\mathbf{a}_i\}$ で表現される．また，与えられた画像中のシルエット領域を Sとし，ノード i の幾何形状を画像平面上へ投影して得られる領域を $P_i(\mathbf{a}_i)$ とする．ここで，$f(A)$ を画像平面上の領域 A の面積を求める写像とすると，姿勢推定問題は，

$$J[\{\mathbf{a}_i\}] = f\left(\left(\bigcup_i P_i(\mathbf{a}_i)\right) \oplus S\right) \qquad (1)$$

を最小化する問題と考えることができる．なお，ここで⊕は領域どうしの排他的論理和を求める演算子である．

対象物体と関節物体モデルの形状および関節の構造が同一である場合，$J[\{\mathbf{a}_i\}]$ の最小値は 0 である．一般的には，これを満たす姿勢のみが入力画像中の姿勢と一致する．これを満たす姿勢 $\{\mathbf{a}_i\}$ は，一意である場合と一意でない場合がある．一意でない場合，それはその姿勢についてある $\mathbf{a}_i$ が変化しても $\bigcup_i P_i(\mathbf{a}_i)$ が変化しないことを意味する．この場合，原因はノードどうしの隠蔽などが考えられる．

実際には対象物体と関節物体モデルの形状および関節の構造が全く同一であるという仮定は現実的ではないので，$J[\{\mathbf{a}_i\}]$ を最小にする姿勢を求めることをもって，本研究では姿勢推定とする．

5. マッチング処理

本研究では入力画像としてシルエット画像を用いている．この場合，モデルとシルエットとのマッチング評価の基準として，面積ないしシルエットの輪郭の形状に注目することが考えられる．

輪郭の形状に注目してマッチング評価する方法については，対象とモデルの幾何形状がよく一致する限りにおいては良好な姿勢推定が行えることが予測される．しかしながら，この方法は輪郭の微妙なずれが存在するだけで姿勢推定がうまく行かないことがある．そこで，本稿では面積をマッチング評価基準に用いる．

本マッチング処理では，関節物体モデルの木構造を利用して各ノードごとにサブマッチング処理を行うことにより，姿勢全体の関節角度変数を同時に扱う方法に比べてマッチング処理量を削減している．処理を分割しているので，局所的最適解に陥る可能性があるため，処理の二段階目で局所的最適解の回避を行っている．なお，本マッチング処理の収束性については，各ノードの処理におけるシルエット領域と当該ノードとの関係評価を工夫することで $J[\{\mathbf{a}_i\}]$ が最小となることを狙っているが，完全な収束性の保証はされていない．

本稿で提案するマッチング処理は大きく二つの部分に分けられる．一つは姿勢が定まっていない姿勢決定木中のノードの関節角度を順次決定していく処理である．もう一つは，いったん得られた姿勢推定結果からさらによい結果を求めるために，再処理すべきノードを決定する処理である．以下に，この二つの処理について説明を行う．

5.1 姿勢推定処理

姿勢推定処理は姿勢決定木を根ノードから葉ノードに向けて辿りながら進める．いま，$\mathcal{F}$を処理済みのノード集合とし，$\mathcal{F}$中のノードを親に持つ未処理ノードの集合を $\mathcal{N}$，残りのノードの集合を $\mathcal{R}$ とする．最初，$\mathcal{N}$には根ノードのみが存在し，他のノードはすべて $\mathcal{R}$ に含まれる．

本処理では，すべてのノードについて関節角度を順次決定していく．最終的には，すべてのノードが $\mathcal{F}$に含まれるようになるまで以下の姿勢推定処理を行う．図 3に処理の流れを示す．

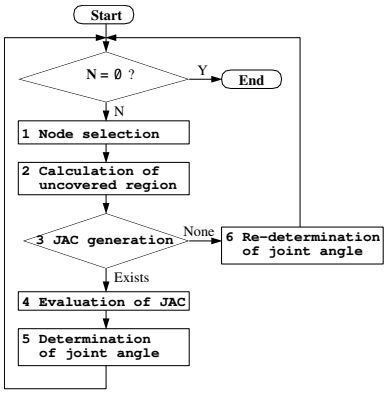

図 3 姿勢推定処理
Fig. 3 Pose estimation procedure

(1) ノード選択
$\mathcal{N}$の中で，もっとも根ノードまでの段数の少ないノード i を選択する．もし $\mathcal{N} = \emptyset$ なら，姿勢推定結果が得られたことになるので，処理終了．

(2) 未被覆領域の計算
$\mathcal{F}$に属すすべてのノードを投影して，画像上に生成した領域を既被覆領域 Cと呼び，式 (2) で表す．さらに，シルエット領域 Sから既被覆領域 Cに相当する部分を削除したものを，未被覆領域 S'とする．これは式 (3) のように表される．

$$C = \bigcup_{i \in \mathcal{F}} P_i \tag{2}$$

$$S' = S \cap \overline{C} \tag{3}$$

(3) 関節角度候補の生成
ノード i の関節角度を決定するには，P_iと Sとの関係を調べなくてはならない．P_iは関節角度によって変化するので，一般的には，P_iを関節角度の関数として表現しそれを解析的に処理することが考えられる．しかし，本研究では各ノードの幾何形状に制限がないので，ノード i の関節可動範囲を一定間隔でサンプリングする方法を採る．すなわち，サンプリングされた関節角度ごとに P_iを生成し，Sの外側にはみでる領域が形状誤差限界面積 r 以下であるようなサンプリング関節角度のみを関節角度候補 $\mathbf{JAC}_i(t)$ とする．なお，t はこれがノード i の t 番目の関節角度候補であることを意味する．$\mathbf{JAC}_i(t)$ は関節角度 $a_i(k),\ 0 \leqq k < n_i$を要素とするベクトルである．

$$\mathbf{JAC}_i(t) \ \ where \ \ f\left(P_i\left(\mathbf{JAC}_i(t)\right) \cap \overline{S}\right) \leqq r \tag{4}$$

もし関節角度候補が一つも生成されなければ，処理 6. へ進む．

(4) 関節角度候補の評価
各関節角度候補 $\mathbf{JAC}_i(t)$ について，式 (5) に示されるような評価値 $e_i(t)$ を計算する (図 4参照)．

$$e_i(t) = f\left(S' \cap P_i\left(\mathbf{JAC}_i(t)\right)\right) \tag{5}$$

$e_i(t)$ は，ノード i の投影領域 P_iが未被覆領域 S'をより広く被覆するほど，よい評価値を与える．ここで留意すべきことは，P_iと既被覆領域 C とがどれだけ重畳しているかが評価に影響を与えないことである．式 (5) は，P_iが既被覆領域 Cといくら重畳しても構わないが未被覆領域 S'をカバーするべきである，との考えに基づく面積を用いた評価基準である．
処理 4. における姿勢決定木内の姿勢推定処理の進行の様子を，図 5に示す．

(5) 関節角度の決定
$e_i(t)$ の値の大きいものから順に，$\mathbf{JAC}_i(t)$ を並べ，リストを作成する．このとき関節角度候補の順番 t を並べた順に付け替える．最初の関節角度候補 $\mathbf{JAC}_i(1)$ が，ノード i で採用される関節角度となる．ノード i

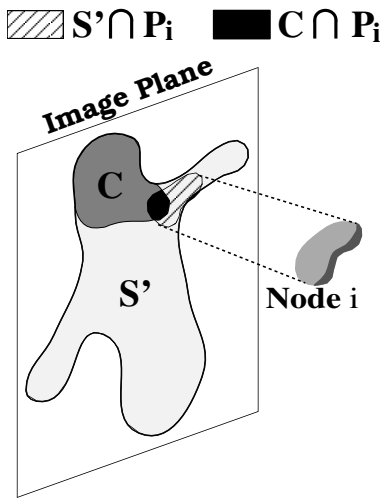

図 4　関節角度候補の画像上での評価
Fig. 4　Evaluation of a Joint Angle Candidate(JAC) on an image plane

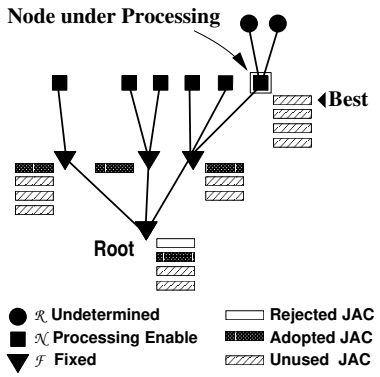

図 5　姿勢推定処理中の姿勢決定木の状態
Fig. 5　A status of a pose decision tree under pose estimation processing

は $\mathcal{F}$に移され，その子ノードは $\mathcal{N}$に移される．処理 1. に戻る．

(6) 関節角度の再決定

関節角度候補がノード i において一つも生成されないのは，それより根側のノードのどれかで誤った関節角度が採用されているからである．そこで，ノード i から根ノードへ向けて木を辿り，未採用の関節角度候補を有する最初のノード jを探す．このノード jにおいて，u_jは現在採用されている関節角度候補のリスト中の順位を表す．ノード jで採用されている関節角度候補 $\mathbf{JAC}_j(u_j)$ を棄却し，その次の関節角度候補 $\mathbf{JAC}_j(u_j+1)$ を新たに採用する．ノード jの子ノードを $\mathcal{N}$に移し，その子ノードに子孫ノードがあれば，すべて $\mathcal{R}$ に帰属させる．処理 1. へ戻る．

5.2　姿勢再推定処理

前節の処理 1. において $\mathcal{N}=\emptyset$ となれば，ある姿勢推定結果が得られたことになる．理想的には，$J[\{\mathbf{a}_i\}]$ を最小にするすべての姿勢が得られればよいことになるが，そのためのマッチング処理量は膨大なものとなる．例えば，図 2，表 2で示されるモデルを 20 度間隔でサンプリングした場合，すべてを探索するためには最悪の場合 5.7×10^8回だけ処理 3. を実行する必要があることになる．そこで，本節ではいったん得られた姿勢推定結果に対してよりよい $J[\{\mathbf{a}_i\}]$ の値を得られるよう，姿勢を漸近的に求める方法について述べる．

5.2.1　葉ノード集合の組合せ

葉ノードに未採用の関節角度候補が複数残っている場合，それらの関節角度候補の組合せを考える．それぞれの姿勢について $J[\{\mathbf{a}_i\}]$ を計算し，よりよい評価値が求められた場合は，その姿勢を推定結果とする．本処理による結果の改善は，複数の葉ノードの投影領域が互いに重なりあう場合に有効である．

5.2.2　最大露出ノード

葉ノード以外のノードで関節角度が変化すると，その子孫ノードすべての関節角度について再推定処理をしなくてはならない．このため，関節角度候補の変更をどのノードで行うかは，マッチング処理量に大きく影響する．

本稿では，他のノードとの隠蔽関係が少ないノードは他のノードより先に関節角度を決定すべきであると考え，以下のような処理を行う．

まず，葉ノード以外で，未採用の関節角度候補を持つノードの集合 $\mathcal{A}$ を求める．$\mathcal{A}$ の中から，ノード i を選ぶとし，式 (6) で表される非隠蔽領域 R_iを定義する．これは，ノード i 以外のモデル中の全ノードを投影し，その投影領域を Sから除去した領域に相当する．

この非隠蔽領域 R_i上へノード i を関節角度候補 $\mathbf{JAC}_i(t)$(但し $t\geq u_i$) の値に従って投影し，それぞれの被覆面積 $cover_i(t)$ を式 (7) に従って計算する．これを $\mathcal{A}$ 中のすべてのノードについて計算し，最も値の大きいノード m とその時の関節角度候補 $\mathbf{JAC}_m(t_m)$ を求める．このノード m を最大露出ノードと呼び，その関節角度候補を最大露出関節角度候補と呼ぶ．

表 2　人体モデルの関節角度可動範囲 (単位 $[degree]$)

ノード	軸	最小値	最大値	ノード	軸	最小値	最大値	ノード	軸	最小値	最大値	ノード	軸	最小値	最大値
頭部	Y	-90	90	右手	Y	-90	90	腰部	Y	-50	50	左大腿	Y	-20	50
	Z	-30	30		Z	-45	40		Z	-30	30		Z	0	40
	X	-30	30		X	-10	30		X	-20	20		X	-100	10
胸部	Y	-60	60	左上腕	X	-180	20	右大腿	Y	-50	20	左脛	X	0	120
	Z	-30	30		Z	-20	90		Z	-40	0	左足首	X	-40	40
	X	-30	30	左前腕	Y	-90	20		X	-100	10		Y	0	40
右上腕	X	-180	20		X	-160	0	右脛	X	0	120		Z	-50	20
	Z	-90	20	左手	Y	-90	90	右足首	X	-40	40				
右前腕	Y	-20	90		Z	-40	45		Y	-40	0				
	X	-160	0		X	-10	30		Z	-20	50				

$$R_i = \left(\bigcup_{k \neq i} P_k \left(\mathbf{JAC}_k(u_k) \right) \right) \cap S \quad (6)$$

$$cover_i(t) = f\left(R_i \cap P_i \left(\mathbf{JAC}_i(t) \right) \right) \quad (7)$$

$$\mathbf{JAC}_m(t_m) \quad where \quad \max_{m \in \mathcal{A}} \left(\max_{t_m \geqq u_m} \left(cover_m(t_m) \right) \right) \quad (8)$$

もし $t_m = u_m$ ならば，ノード m を $\mathcal{A}$ から抹消したのち，最大露出ノードと最大露出関節角度候補を再度求める．もし $t_m > u_m$ ならば，最大露出ノード m の u_m の値を t_m とし，その関節角度を変更する．

関節角度の変更によって，最大露出ノード m の子孫ノードの関節角度は再決定されなければならない．このため，最大露出ノード m の子ノードを $\mathcal{N}$ に帰属させ，さらにそれらの子孫ノードがもし存在すれば $\mathcal{R}$ に帰属させる．この後，前節 5.1 に述べた方法を再度実行する．

なお，もし $\mathcal{A}$ に一つもノードが存在しなくなれば，すべての姿勢推定処理の終了となる．

6. 実験と評価

本稿で提案する姿勢推定方法について，まず CG で合成したシルエット画像に対して行った実験について結果を示し，考察を行う．続いて，実在の物体に適用した実験例を示す．

6.1 CG 合成画像に対する実験と評価

CG 合成画像に対する実験では，女性人体を対象とした．まず，姿勢推定処理に用いる女性人体モデルを利用してシルエット画像を CG で 219 枚合成する．関節角度は 20 度刻みでランダムに発生させた．これらの画像それぞれに対して，根ノードの関節接続点の位置が既知であるとして本方法を適用した．画像の大きさは縦横とも 700 画素である．カメラ条件としては並行投影を仮定し，実世界の $1cm$ が 2.26 画素分の長さに相当するものとする．

実験に用いた女性人体モデルの関節角度の可動範囲を表 2に示す．図 6はすべての関節角度が 0 度である時のモデルの姿勢を示す．各ノードのローカル座標軸はすべて，画像平面内にモデルをカメラに対して図 6 のように直立させたときの水平方向を X 軸，垂直方向を Y 軸，カメラ方向を Z 軸としている．各ノードの各軸の表記の順は上から下へそのまま回転の順番を示す．また，根ノードには頭部が相当する．これは，本実験のように人体が自由な姿勢をとる場合，位置の抽出は頭部が一番行いやすいのではないかと考えたからである．

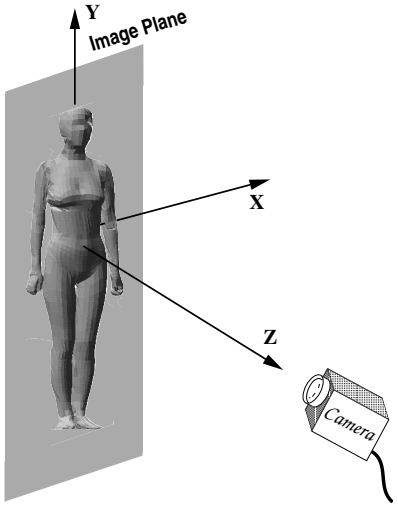

図 6　直立したモデルと座標軸との関係
Fig. 6　Relationship between a uplight model and axes of coordinates

姿勢推定処理を行うにあたって，本実験では関節角度候補の生成のためのサンプリング間隔を 20 度に設

定した．また，ノード一つについて生成される関節角度候補を最大 5 個に制限する．すなわち，関節角度候補の評価において，評価値 $e_i(t)$ の上位 5 候補のみでリストを作成する．

CG 合成画像で実験を行う場合，形状誤差限界面積を $r = 0$ とすれば各ノードの投影領域が完全にシルエット領域に包含されることになるので，姿勢推定の結果がよくなることが予想される．しかし，これは対象物体とモデルとが全く等しい幾何形状を持つという仮定をおいているのと同じである．実際には，モデルに対象物体と完全に等しい幾何形状を与えることはほぼ不可能であるから，この仮定は現実的ではない．そこで，今回の実験では形状誤差限界面積 rを 200 画素に定めた．これはカメラ条件から換算すると $39cm^2$ で，ノードの平均投影面積の約 12.6%に相当する．

本研究では，姿勢推定を $J[\{\mathbf{a}_i\}]$ を最小にすることで行っている．ここでは，各姿勢による $f(S)$ の大きさの違いを正規化して評価するため，式 (9) で定義される E_{total}を用いる．

$$E_{total} = \frac{J[\{\mathbf{a}_i\}]}{f(S)} \qquad (9)$$

219 例の平均評価値は 2.27%であった．Sと $P = \bigcup_i (P_i(\mathbf{a}_i))$ との完全一致は 135 例であった．これは，全例の 61.6%において本方法が $J[\{\mathbf{a}_i\}] = 0$ を満たす姿勢を求めたことを意味する．このうち，101 例についてはすべての関節角度が一致した．図 7にその数例を挙げる．表示は濃淡画像であるが，姿勢推定処理時にはシルエット画像がシステムに入力されている．図を見ても分かるように，これらはあまり隠蔽等が生じていない姿勢である．一方，残りの 34 例については Sと Pとが一致しているにも関わらず，姿勢が一致しない．これは，4 章でも述べたように，$J[\{\mathbf{a}_i\}] = 0$ を満たす姿勢が必ずしも一意ではないからである．具体的には，これには二つの原因がある．

一つは隠蔽であり，これは単眼視画像を用いている限り解決されない．このような例が 14 例存在した．図 8に二例挙げる．上段が原画像であり，中段は元の姿勢を横から観測した画像，下段は推定結果を横から観測した画像である．

もう一つの原因はシルエット画像を用いたことによるエッジ情報の欠如である．これは 22 例見られた．図 9に例示するが，左が元の姿勢，右が推定結果である．シルエット領域内のエッジ情報を用いれば，これらの例は正確な姿勢推定が可能になるものと思われる．

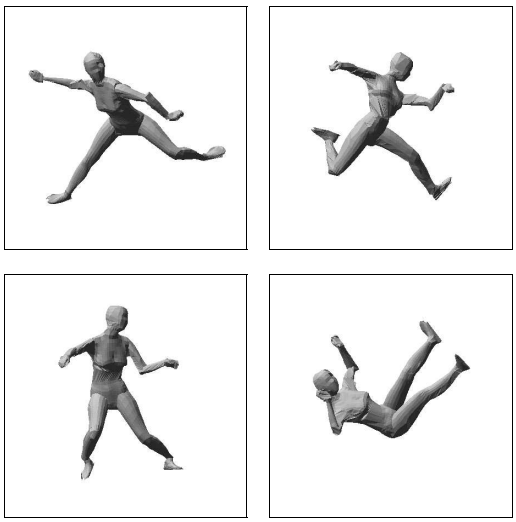

図 7　姿勢一致した成功例
Fig. 7　Succeeded examples of coincident poses

これらの原因に基づく姿勢の一致しない推定結果の場合，いくつかのノードでは比較的高い評価値 $e_i(t)$ を持つ関節角度候補 $\mathbf{JAC}_i(t)$ が未採用のまま残されている．そこで，本方法の後処理としてエッジ情報や他の視点方向からの観測情報を導入すれば，それらの関節角度候補のうちから正しいものを選択することが可能になる．このように，曖昧な姿勢に対しても本方法は容易に対応可能である．

実験中，75 例については最初に求めた姿勢推定結果より姿勢再推定処理 (5. 2節) 後で得た推定結果のほうが E_{total}の値を改善した．これは，姿勢再推定処理による局所的最適解の回避が行われたことを示している．図 10に示した例では，下段の姿勢に対するシルエットが与えられた時，姿勢再推定処理前では上段のような誤った結果となったが，姿勢再推定を行うことで元の姿勢と一致する結果に改善されている．マッチング処理量の観点から考察すると，平均処理ノード数すなわち処理 3. の実行回数は平均で 26.2 回である．実行回数の最小は 15 回であるから，マッチング処理量は最小実行回数に近い数字であることが分かる．処理ノード数の分布を図 11に示す．処理ノード数の多い上位 7 例を除くと，残りの平均は 20.1 となる．

ところで，本実験において Sと Pが一致しない例が存在する最大の理由は，ノード一つについて生成される関節角度候補を 5 つに制限していることである．図

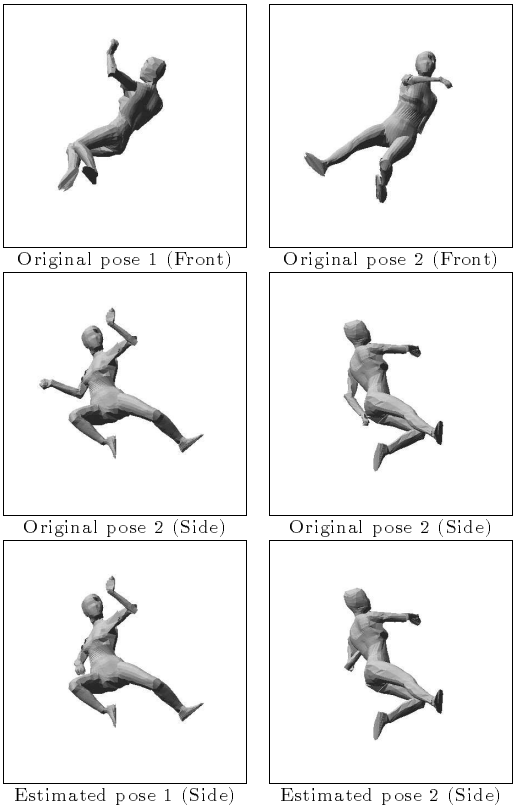

図 8　隠蔽下でのシルエット一致例
Fig. 8　Occluded examples fitted to silhouettes

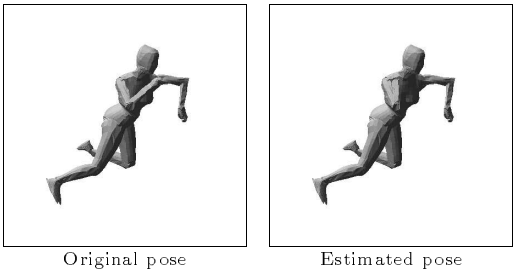

図 9　エッジ情報欠如下でのシルエット一致例
Fig. 9　An example fitted to a silhouette under lack of edge information

12に一例を示す．この例では，胸部と右前腕とが接する形で配置されているので，胸部のノードにおいて関節角度候補が多数発生し，正しい関節角度候補が上位5つに入れなかったため，シルエットの一致が得られなかった．生成される関節角度候補の生成数を増やせばシルエットが一致する率を向上させることは可能であるが，生成数に対してマッチング処理量が指数的に増大するため，そのトレードオフが問題となる．

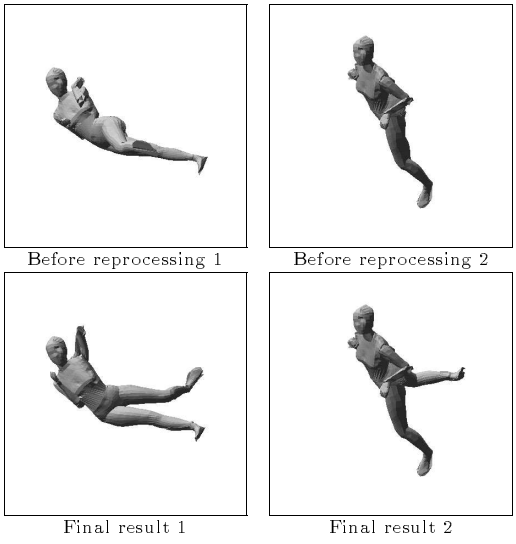

図 10　改善された例
Fig. 10　Improved examples

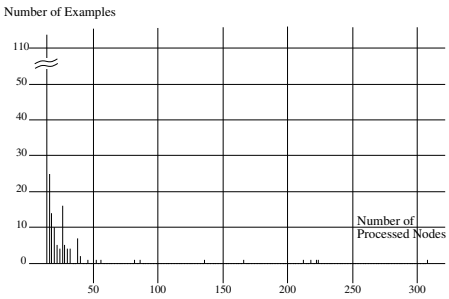

図 11　処理ノード数
Fig. 11　Number of processed nodes

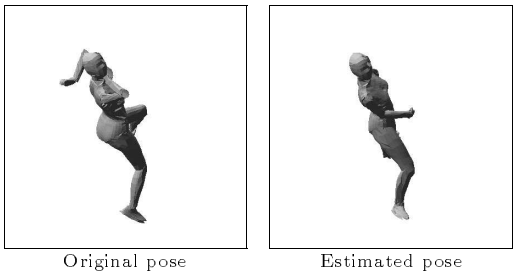

図 12　シルエットが一致しなかった例
Fig. 12　Examples misfitted to silhouettes

6.2 実在の物体に対する実験と評価

女性人体モデルを使用して，実際の人体をビデオカメラで撮影した画像に対して姿勢推定した例を図 13 に示す．但し，根ノードは右足首であり，その関節接続点の位置が既知で，関節角度のサンプリング間隔を 10 度刻みとしている．右足首を根ノードとするのは，それが地面に接しているので抽出が容易であろうと考えられる上に，シルエット画像では判別不能な足の左右の区別がつけられるからである．使用した女性人体モデルは対象である女性から作成したものである(10)．モデル作成時の関節接続点の設定や各ノード内の幾何形状の剛体部と関節角度依存部との境界の設定が正確に行えていないため，CG 画像の実験結果に比べると姿勢推定結果の精度はいくらか落ちる．画像をビデオカメラで撮影しているため，その大きさは横 320 画素，縦 360 画素と CG 画像の実験より粗い．同様に 6 例について実験を行い，全部について E_{total}の平均を求めると 22.3%であった．対象物体の姿勢推定結果から，本方法が有効であることが分かる．

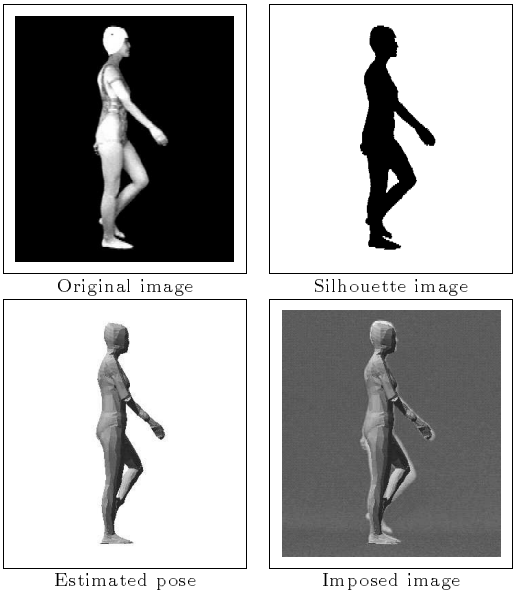

図 13 女性の姿勢推定例
Fig. 13 A pose estimation example of a female

また，他の関節物体に本方法を適用した例として，人間の左手首を対象にした実験を 3 例行った．その 1 例を図 14に示す．手を対象にする場合，前腕の位置が最も抽出しやすいと考えられるので，モデルの根ノードを左手前腕とする．実験にあたって，その位置は既知とする．画像の大きさは横 448 画素，縦 400 画素である．この実験ではモデルと対象物体との幾何形状が異なるので，指の長さ等は若干一致しない．3 例による E_{total}の平均は 17.2%であった．

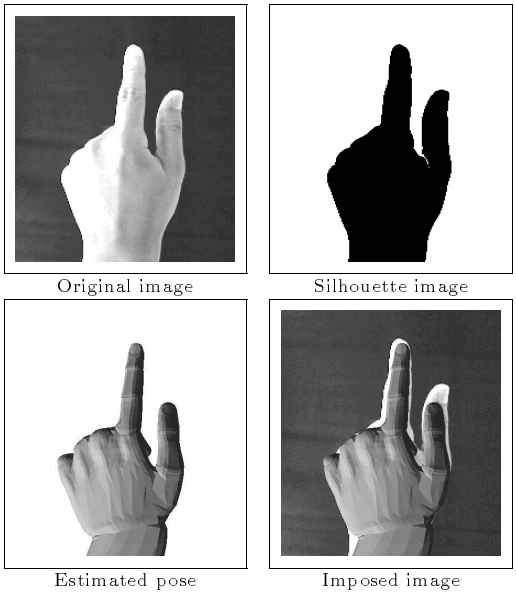

図 14 左手首の姿勢推定例
Fig. 14 A pose estimation example of a hand

これらの実験から，実在の物体に対しても本方法によって関節物体の姿勢を推定できることが実証された．実在の物体に対する実験においても，CG による実験の考察で挙げたような隠蔽の問題などは残されるが，これについては前節で述べたように将来的に対応が可能である．

7. むすび

本稿では，一枚のシルエット画像からその画像中に写された関節物体の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た．本方法は逆運動学的な解法やマッチング処理の内部でのヒューリスティックな知識を使用していないので，関節物体一般の姿勢推定に適用可能である．CG 合成画像による実験を通して本方法の姿勢推定能力が明らかになり，また一枚のシルエット画像のみでは推定不可能な部分に対する対応方法についても考察した．さらに，実際に人体と手指に適用することで，本方法の有効性を検証した．

今後の課題としては，後処理としてのエッジ情報の利用や，複数方向からの画像の利用，動画像への応用などが考えられる．

謝辞 関節物体モデル作成に御協力して頂きました東洋紡けいはんな研究室の坂口嘉之氏に感謝致します．

文　献

(1) 倉掛正治, ラマカント ネヴァティア, “関節のある動物体の記述と追跡,” 信学論 (D-II), **J76-D-II**, 4, pp. 854–862 (1993).
(2) クンラポン ユーニハン, 木下宏揚, 酒井善則, “視覚言語処理システムにおける手の認識法,” 信学論 (D-II), **J75-D-II**, 9, pp. 1489–1497 (1992).
(3) J. Lee and T. L. Kunii, “Constraint-based hand animation,” Models and Techniques in Computer Animation (Computer Animation Series), pp. 110–127 (1993).
(4) M. K. Leung and Y. Yang, “A region based approarch for human body motion analysis,” Pattern Recognition, **20**, 3, pp. 321–339 (1987).
(5) W. Long and Y. Yang, “Log-tracker: An attribute-based approarch to tracking human body motion,” IJPRAI, **5**, 3, pp. 439–458 (1991).
(6) 木本伊彦, 梶谷昭彦, 安田靖彦, “スティックモデルに基づく単眼視動画像からの人体歩行運動の解析の一手法,” 信学論 (D-II), **J74-D-II**, 3, pp. 376–387 (1991).
(7) クンラポン ユーニハン, 木下宏揚, 酒井善則, “スティックモデルを用いた手振りの認識,” 信学論 (D-II), **J77-D-II**, 1, pp. 51–60 (1994).
(8) 石井浩史, 望月研二, 岸野文郎, “人物像合成のためのステレオ画像からの動作認識法,” 信学論 (D-II), **J76-D-II**, 8, pp. 1805–1812 (1993).
(9) 島田信敬, 白井良明, 久野善徳, “三次元モデルを用いた二次元動画像からの手指姿勢の推定,” 信学技報, **PRU94**, 94-50, pp. 25–32 (1994).
(10) 坂口嘉之, 美濃導彦, 池田克夫, “仮想服飾環境 party — 人体と型紙のための幾何学的制約充足型格子形成法 —,” 信学論 (D-II), **J77-D-II**, 11, pp. 2210–2219 (1994).

亀田 能成

平 3 京大・工・情報卒．平 5 同大大学院修士課程終了．現在，同大博士後期課程在学中．モデルを用いた三次元物体認識の研究に従事．情報処理学会会員．

美濃 導彦

昭和５３年京都大学工学部情報工学科卒業．昭和５８年同大学院博士課程終了．同年，工学部助手．平成元年助教授．画像処理，人工知能，知的コミュニケーション関係の研究に従事．工学博士．電子情報通信学会，画像電子学会，ロボット学会各会員．

池田 克夫

昭和１２年生．昭和３５年京都大学工学部電子工学科卒業．昭和３７年同大学院修士課程終了．昭和４０年同博士課程学修退学．同年京都大学助手．昭和４６年同助教授，工学部．昭和４６年９月より１年間文部省在外研究員として，米国ユタ大学および MIT に留学．昭和５３年筑波大学教授，電子・情報工学系．昭和６３年８月京都大学教授，工学部情報工学教室．コンピュータ組織法,LAN，画像処理 / 理解に興味を持つ．著書に「オペレーティングシステム論」(電子情報通信学会) などがある．工学博士．電子情報通信学会，人工知能学会,IEEE,ACM 各会員.IPL editor．