シルエット画像からの関節物体の姿勢推定能力について

亀田 能成　美濃 導彦　池田 克夫

京都大学工学部

1　はじめに

関節物体が写された一枚の画像について，そのシルエットに対象物モデルの投影領域が一致するように全ての関節の角度を推定可能かどうかを検討する．関節物体とは，剛体部品が関節で連結された物体をいう．具体的例としては，人体全体やその一部が挙げられる．

これまでの筆者らの方法では，各部品の画像に対するマッチング評価基準として，画像中のシルエットと部品の画像平面への投影領域との重畳関係を重なり領域を基に評価してきた [1, 2]．この方法では，モデルと対象物体との幾何的特徴情報を利用していないため，関節角度の推定誤差が生じることや，処理が逐次実行されるのでいったんある部品において局所解に陥ると正しい推定結果を得られない，などの問題があった．

そこで本稿では，部品の投影領域と画像中のシルエットとの重畳関係を輪郭も利用して評価する方法を提案する．また，人体の CG 画像について実験を行い，シルエットを入力情報とする方法の限界についても考察を行う．

2　モデル

人体などの関節物体は，剛体部品が木構造状に連結したと見なせることから，本稿では関節物体モデルを部品位置決定木 [2] で表現する．関節は 1 から 3 の自由度を持ち，回転角度の範囲が定められている．各々の剛体部品はパッチ表現で幾何情報を有している．システムは，対象物体中の関節角度以外の情報を予め保持している．

3　マッチング方法

マッチング処理は部品位置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られる．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．また，与えられた画像のシルエット領域を S で表す．いま，ノード $i \in \mathcal{N}$ の剛体部品をマッチング処理することを考えると，以下のような手順となる．

1. ノード i の関節の回転範囲の大きさに応じた複数の関節角度候補 $\mathbf{v}_{(i,k)}$ を作成し，各々について，ノード i の剛体部品の画像平面上への投影領域の画素集合を $P_i(\mathbf{v}_{(i,k)})$ とする．
2. モデルの投影領域はシルエットから逸脱してはならないという知識から，面積に基づくマッチング評価値 $e_{rgn}(\mathbf{v}_{(i,k)})$ を求める．関数 $pix(A)$ は指定された領域 A の画素数を計算する．

$$e_{rgn}(\mathbf{v}_{(i,k)}) = pix\left(\overline{S} \wedge P_i(\mathbf{v}_{(i,k)})\right) \tag{1}$$

3. もし $e_{rgn}(\mathbf{v}_{(i,k)}) = 0$ を満たす k が一つもなければ，5. に処理を進める．
4. $e_{rgn}(\mathbf{v}_{(i,l)}) = 0$ を満たす関節角度候補についてだけ，輪郭に基づくマッチング評価値 $e_{cntr}(\mathbf{v}_{(i,l)})$ を求める．また，関数 $cntr(A)$ は領域 A の輪郭を構成する画素集合を求める関数である．その中で，e_{cntr} を最大にする関節角度候補 $\mathbf{v}_{(i,n_i)}$ をノード i の関節角度として採用し，ノード i を $\mathcal{N}$ から $\mathcal{F}$ に移動させる．ノード i の子ノードは $\mathcal{N}$ に移る．6. へ進む．

$$e_{cntr}(\mathbf{v}_{(i,l)}) = pix\left(cntr(S) \wedge cntr(P_i(\mathbf{v}_{(i,l)}))\right) \tag{2}$$

5. ノード i から根ノードに向けて木構造を後戻りして，$e_{cntr}(\mathbf{v}_{(j,l)})$（ただし $l \neq n_j$）が存在する最初のノード $j \in \mathcal{F}$ において，$\mathbf{v}_{(j,n_j)}$ を破棄し，新しく $\mathbf{v}_{(j,l)}$ をノード j の関節角度として採用する．ノード j の子ノードは全て $\mathcal{N}$ に，それ以降の子孫ノードは $\mathcal{R}$ に移される．
6. $\mathcal{N}$ から一つを取りだし，1. へ戻る．もし $\mathcal{N} = \emptyset$ なら推定終了となる．

4　実験と考察

実験は人体を対象として行った．モデルは 17 個の剛体部品からなり，根ノードには頭部分が相当する．関節の自由度は合計で 40 あり，総パッチ数は 7672 である．

Pose Estimation Ability for an Articulated Object from its Silhouette

Yoshinari KAMEDA, Michihiko MINOH, Katsuo IKEDA

Faculty of Engineering, Kyoto University

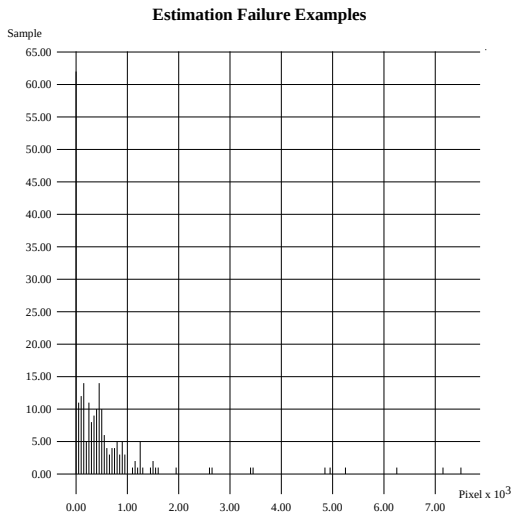

図 1: 推定不成功時の E の様子

今回の実験では，CG によって合成した人体モデルの画像を入力とする．この際，関節角度は 20 度間隔の乱数で発生させた．

今回はマッチング評価基準にシルエット情報のみを利用しているので，姿勢推定が成功したかどうかの判定は，シルエットと推定結果の投影領域との排他的論理和 E が 0 かどうかで行う．

701 例に対して実験を行った結果，$E = 0$ になったのは 472 例 (67.3%) であった．姿勢推定に失敗した 229 例について図 1に傾向を示す．横軸は排他的論理和を画素で表し，縦軸は処理例の数を示している．姿勢推定失敗は，あるノードがシルエットの一部分をその投影によって構成しているにも関わらず，その部分以外のシルエットの部分に対応付けられてしまう場合に生ずる．これについては，e_{cntr} の定義および終了判定条件を変更すれば解決可能であるが，計算量が膨大になる可能性がある．

成功例の一つを図 2に示す．成功例でのマッチング処理に要した処理ノード数の平均は 17.38，最大で 79 であった．なお，実験においては，各ノード i の関節角度候補 $\mathbf{v}_{(i,k)}$ のサンプリングを 20 度間隔で行った．

関節物体の三次元的な姿勢推定という観点から見ると，$E = 0$ は必ずしも対象の姿勢と推定結果が同一であることを保証しない．推定終了時の各ノード i において，関節角度候補 $\mathbf{v}_{(i,l)}, (l \neq n_i)$ が一つ以上残されている場合，推定結果が対象画像に対して複数解を有する可能性がある．本実験では，472 個の成功例のうち，95 例がこの可能性を有していた．図 3にその一例を挙げる．このようにその関節角度に任意性が残されたノードについては，少なくともそのノードが他のノードよりカメラ側に近い場合，輝度画像のエッジ情報などを利用すれば関節角度の推定が可能である．

図 2: 推定成功例

図 3: 三次元的な曖昧さを残した推定成功例

5 おわりに

本稿では，シルエットの面積と輪郭の両方の情報及び処理の後戻りによって，シルエット領域に過不足なく一致するような関節物体の姿勢推定が可能であることを示した．また，シルエットを入力とするモデルマッチングの限界にも触れ，考察を行った．

参考文献

[1] 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, pp. 255-256, 1992.

[2] 亀田能成, 美濃導彦, 池田克夫: シルエットからの関節物体三次元姿勢推定法, 第 21 回画電学大 Visual Computing'93, pp. 133-136, 1993.