

H-044

# シルエットによる投球動作の3次元姿勢推定

## 3D Pose Estimation of Pitching Motion by Silhouette

子安 大士†  
Hiroshi Koyasu

中村 太†  
Dai Nakamura

前川 仁†  
Hitoshi Maekawa

### 概要

本論文では、野球における投球動作の3次元姿勢推定について述べる。3次元姿勢推定では、マーカを用いたモーションキャプチャシステムや、複数カメラによる手法が提案されている。しかし、このような手法は計測に制限が大きく、また、テレビのライブ画像や過去のアーカイブを利用する上では、単眼カメラでマーカなどを用いずに姿勢の推定を行える必要がある。そこで、本研究ではあらかじめ作成した3次元人体モデルとシルエットの比較による姿勢推定手法を改良して高速化を図った。実際の投球シーンに対して本手法を適用し、手法の有効性を確認した。

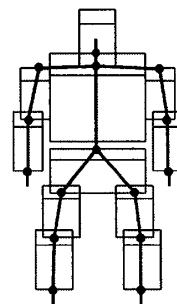


図 1: 15 個の関節による 3 次元人体モデル

### 1 はじめに

3次元姿勢推定はスポーツにおける動作解析、映画やゲームなどのエンターテインメントなど様々な分野で利用されている重要な技術である。野球に関しても、既存のモーションキャプチャシステムを利用した動作解析に関する研究がある [1, 2]。

既存の3次元姿勢推定手法の多くは、被験者にマーカを付ける、複数のカメラを必要とするなど、適用できるシーンに制限がある。一方で、このような制限のない手法として、人体像の輪郭内部を塗りつぶしたシルエット画像を利用した手法 [3] がある。この手法ではあらかじめ3次元人体モデルのシルエット画像による姿勢辞書を作成し、同様にシルエットに変換した入力画像と、もともとシルエット領域の重なりが大きいモデルを選択する事で姿勢を推定する。

シルエットによる姿勢推定手法は、単眼カメラでかつマーカを用いずに姿勢を推定できるため様々な用途に用いることができる。しかしながら、あらかじめ多くの姿勢に対応したシルエット画像を用意しておくため、多くの記憶領域を必要とする。また、大量のモデルの中からの探索をするため、計算コストが大きい。

我々は投手の投球フォームの解析のために、シルエットを用いた3次元姿勢推定の研究を行ってきた [4]。そこで、本論文では既存のシルエットによる3次元姿勢推定手法に対して、ランレングス符号化を用いた画像の圧縮を導入することで記憶領域への制限を緩和するとともに、重複を許容した木構造を用いて探索の効率化を図った。

† 埼玉大学大学院理工学研究科

### 2 シルエットによる3次元姿勢推定

#### 2.1 3次元人体モデル

人体には 200 個以上の関節があり、それぞれの関節の角度により姿勢が決定される。しかし、全ての関節をモデル化することは組み合わせの爆発が生じるため、本研究では特に重要と思われる両肩・両肘・両手首・首・腰・股関節・両膝・両足首の計 15 個の関節のみを考慮したモデルを使用する。図 1 に 15 個の関節を持った人体モデルを図示する。図中、黒丸で示される点がそれぞれの関節の位置を表し、関節をつなぐ実線が関節の連結関係を表す。また、四角は体節を表す。各体節の長さ、太さは自由に設定することが可能である。

#### 2.2 シルエット画像のマッチング

入力された二次元画像と三次元人体モデルの姿勢の一致度の評価は、それぞれを二次元シルエットに変換しその相違を計算することで行う。シルエット間の相違は、最もシンプルな定義ではシルエット部分の画素値を 1、背景部分を 0 とし、両画像の排他的論理和 XOR を用いる。しかし、この定義では着衣などの影響により生じる輪郭のぶれの影響を受けるため、誤った推定結果になりやすい。そこで、シルエット画像に骨格化処理を施し、輪郭に近い所の重みを小さく、中心に近い所の重みを大きくするように重み付けを行った排他的論理和を用いる手法 [3] が提案されており、本研究でもこれを用いる。

#### 2.3 木構造による探索の効率化

膨大な大きさの姿勢辞書の全てに対して上記のマッチング処理を行うことは計算コストの面で不可能である。

そこで、シルエットがよく似た姿勢同士をグループ化した木構造を作成することで、マッチングの回数を減らす。木構造の作成は以下の手順で行う。

1. 姿勢の集合  $S$  の中で最も平均に近いものを代表姿勢として選出する。
2. 選出された代表姿勢からシルエット間距離が最も遠いものを次の代表姿勢として選出する。
3. 選出された各代表姿勢からのシルエット間距離が閾値  $T$  より遠い姿勢が無くなるまで2を繰り返す。
4. 代表姿勢以外の姿勢を、最もシルエット間距離の近い代表姿勢の集合へと振り分ける。
5. 各集合の代表姿勢を選出しないおす。
6. 各集合において1~5を行う。これを、構造木があらかじめ決められた階層数になるまで繰り返す。

### 3 3次元姿勢推定手法の高速化

#### 3.1 ランレングス符号化によるシルエット画像の圧縮

姿勢辞書の要素は2値化されたシルエット画像である。この辞書のサイズが大きくなると様々な問題が発生する。具体的には、シルエット画像ファイルを置くハードディスク上に膨大なスペースが必要となることや、構造化や比較の際にファイルがメモリ上に乗りきらず、ファイルのスワップが頻繁に発生し演算速度が大幅に遅くなることが挙げられる。

そこで、シルエット画像の2値画像であることに着目し、ランレングス符号化により圧縮することで、この問題を軽減する。シルエット画像の展開を実際に画像の比較を行うときのみにすることで、実行時に同時に扱えるモデル数を増やし、プログラムの高速化を図る。

#### 3.2 重複を許容した木構造

2.3節で述べたシルエット間の距離によるグループ化を用いた木構造は、根に近いノードではシルエット間距離のしきい値が大きいため、同一ノード内の姿勢のばらつきが大きい。そのため、集合同士の境界付近に近い姿勢を探索する場合、本来ならば最もマッチングするはずの姿勢が属する集合よりも、存在しない集合の代表姿勢の評価値が高くなってしまった場合、マッチングに失敗する。先行研究 [3] では、これを避けるため根の直下のノードを姿勢のばらつきが十分に小さくなるように細かく分割していた。しかし、これでは木構造のバランスが悪く、計算コストを十分に低減できない。

そこで、各集合で要素の重複のある木構造の作成を提案する。この構造木は、代表姿勢の選出を行い、各姿勢を最近の代表姿勢の集合に割り振るところまでは従来手法と同様である。その後、各姿勢を代表姿勢とのシッ

表 1: 各圧縮形式による演算時間と圧縮効率の比較。

ファイル形式	計算時間	平均ファイルサイズ
BMP	136s	192kB
PNG	306s	2.7kB
ランレングス符号化	138s	3.7kB

ト間距離が閾値  $T_s$  より近い他の集合にも割り振る。ここで、

$$T_s = wd_{\max} \quad (1)$$

である。 $d_{\max}$  は最もシルエット間距離に近い代表姿勢とのシルエット間距離、 $w$  は任意に設定を行う重み係数である。このようにすることで、集合同士の境界付近にある姿勢モデルが両方の集合に属する形になり、探索時に誤ったノードが選択される事を避けられる。これにより、木構造のバランスを効率の良い形に保つことが出来、結果として計算コストを低減することができる。

### 4 姿勢推定実験

#### 4.1 シルエット画像圧縮手法の検証

ランレングス符号化によるファイル圧縮の有効性を検証するため、無圧縮 BMP、PNG 圧縮、ランレングス符号化それぞれの形式で、総数 650 の姿勢集合の構造化を行い計算時間を比較した。計算に使用した PC は、CPU:Core2DUO E6300(1.86GHz)、メモリ:1.0GB である。演算はシングルコアのみで行った。結果を表 1 に示す。PNG 形式では圧縮率は高いが計算コストがかかる。一方提案手法では PNG 形式と遜色ない圧縮率で計算時間も無圧縮とほとんど変わらない。

#### 4.2 木構造の違いによる精度と計算時間

次に提案する木構造の性能を確かめるため、様々な構造木を作成して探索の効率や精度に関して検証を行う。まず、シルエット辞書の要素となる集合として、3426 個の姿勢を用意する。作成した辞書を表 2、表 3 に示す。表 2 の従来手法では葉ノードの要素数が全体姿勢数になっているが、表 3 の提案手法では、階層が進むにつれて全体姿勢数を越えて要素数が増加している。

辞書内の画像 30 枚を抜きだし、それを入力としてそれぞれの辞書に対して探索を行った結果を表 4 に示す。

結果として、正答数は提案手法による辞書 C が最も多かった。従来手法による辞書 A と比較して、計算時間が同程度で探索の精度が向上している。平均比較回数では従来手法による辞書 B が最も少ない回数でマッチングを行っている。しかし、辞書 B は正答数が最も低い。辞書 D は辞書 B よりも平均比較回数が多いが、辞書 A、C と比較するとかなり少ない回数である。各辞書において、正しい画像が得られなかった場合でも、辞書 B 以外は比較的近い画像を得ることが出来た。姿勢の推定では、辞書内に入力画像と完全に一致する画像はまず存在しないと考えられるため、近い画像が得られれば十分である。

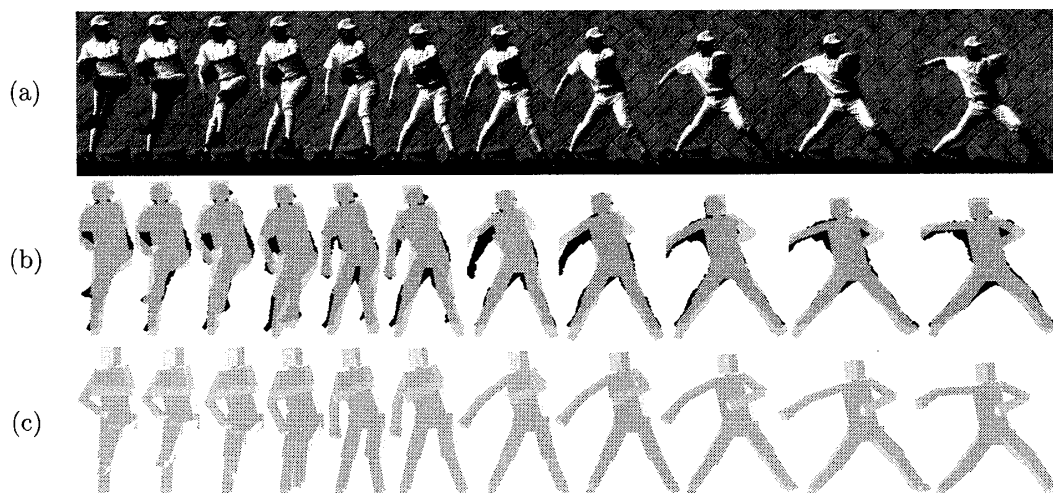


図 2: 3次元姿勢推定結果 (a) は入力実投球画像, (b) は実投球画像と推定されたモデル画像のシルエットを重ね合わせた画像, (c) は推定されたモデル画像.

表 2: 従来型構造木の辞書の各階層のノード数.

シルエット辞書 A			シルエット辞書 B		
階層	要素数	閾値 $T$	階層	要素数	閾値 $T$
1	470	190000	1	6	182000
2	1580	190700	2	11	183900
3	3096	191400	3	21	185800
葉	3426		4	85	187700
			5	514	189600
			葉	3426	

表 4: 各木構造シルエット辞書の比較結果.

	辞書 A	辞書 B	辞書 C	辞書 D
正答数	13	5	21	10
平均比較回数	478.7	30.5	500.6	50.1

表 3: 提案手法による構造木の辞書の各階層のノード数.

シルエット辞書 C			シルエット辞書 D		
階層	要素数	閾値 $T$	階層	要素数	閾値 $T$
1	486	190000	1	7	182000
2	4357	190700	2	19	183900
3	11385	191400	3	62	185800
葉	12353		4	354	187700
			5	3737	189600
			葉	45738	

## 5 まとめと今後の課題

本論文では既存のシルエットによる3次元姿勢推定手法に関する改善と、実投球シーンへの適用結果について述べた。ランレングス符号化を用いた画像の圧縮を導入することで記憶領域への制限を緩和した。また、重複を許容した木構造を用いて精度を保ちつつ効率の良い木構造を構築し、探索の効率化に成功した。今後の課題は時系列の情報を用いて探索の更なる効率化を図ることである。

## 参考文献

- [1] 宮西 智久, 藤井 範久, 阿江 通良, 功力 靖雄, 岡田 守彦, “大学野球選手における速投および遠投動作の3次元比較研究”, 日本体育学会 ISSN:04846710, Vol.40, No.2, pp. 89-103(1995.7).
- [2] 神事 努, 望月 知徳, 湯浅 景元, “野球のピッチング動作における踏み出し脚の運動がボール初速度に与える影響”, 中京大学体育学論叢, Vol.43, No.1, pp. 23-29(2001.11).
- [3] 岡田隆三, シュテンガ ビヨン, “シルエットを用いた *Tree-Based Filtering* による人体の姿勢推定”, 画像の認識, 理解シンポジウム, (2007.7).
- [4] 中村太, 子安大士, 前川仁, “シルエットによる投球・打撃動作の認識” 日本機械学会 [No.08-23] シンポジウム講演論文集, pp.103-107(2008.11.7).

このため、辞書 D は正答数は辞書 A よりも低いが、姿勢の推定に用いることは可能であると考えられる。提案手法では従来型の構造化辞書よりも探索の精度・効率が向上することが確認できた。

### 4.3 実投球シーンの3次元姿勢推定結果

最後に提案手法を用いて、実際の投球シーンを撮影した動画から投手の3次元姿勢を推定する実験を行った。実験に用いた辞書に登録された姿勢の総数は 32763 で、葉ノードを含めて7階層の木構造を構築した。入力動画全体 545 フレームに対して、1 フレームあたりの平均比較回数は 520.1 回、全フレームの推定にかかった時間は 6827 秒であった。実験の結果推定された3次元姿勢の一部を図 2 に示す。