

## R7 スポーツ動作の3次元姿勢推定

原田 篤志†

玉木 徹‡

山本 正信‡

†新潟大学大学院自然科学研究科

‡新潟大学工学部

## 1. はじめに

スポーツ技術を向上させる方法の一は、上手な人の動作を真似ることである。スポーツ時の自分の動作がプロの選手の動作に近づくことができれば、スポーツ技術の向上に繋がるのが期待できる。そのためには、プロ選手に直接指導してもらうことが望ましいがそれは簡単にはできないことではない。一方、テレビやビデオからプロ選手の映像を収集することは比較的容易である。そこで、これらの映像から動作データ(各フレームでの3次元姿勢)を得て、自分の動作データと比較することが可能になれば、プロ選手と直接会うことなく、自分自身で練習し、上達することができる。

本研究ではテレビ中継などの一台のカメラで撮影された映像から動作データを推定することを目的とする。ここでは、同じ動作を異なる人物が行った場合でも、一連の動作の中で近い姿勢データが得られると仮定する。そして、動作データをあらかじめ教示データとして用意しておく、映像中の人物姿勢に最も近い3次元姿勢を推定する。

## 2. 多視点カメラによる教示データの収集

教示データはある動作を行っている映像から推定された動作データである。教示データは単眼カメラの映像中の姿勢を推定するためにあらかじめ用意しておく必要がある。この時、推定したい映像の姿勢を含む動作を撮影することが重要となる。この一連の映像を用い、3次元動作データを得る。

## 3. 中心線と体軸の導出

本手法では、単眼カメラ映像と教示データを比較するために、中心線と体軸を用いる。中心線と体軸の導出方法を以下に述べる。

## 3.1 単眼カメラ映像からの中心線の導出

本手法では推定されるべきはテレビ映像中の姿勢である。2次元の映像と前節で得られた姿勢データが一致しているかどうかを調べるために本手法で用いる特徴は四肢(両上腕、両大腿部)の中心線(計4本の線分)である。中心線の導出方法として、2次元画像上でマウスをクリックして、線分の両端点、左右の肩・ひじ・脚の付け根・ひざの8点の各 $xy$ 座標を取る。そして、それぞれ(肩とひじ、脚の付け根とひざ)の点を結んだ線分を中心線とする(図1)。

## 3.2 教示データからの体軸の導出

教示データからは体軸を導出する。教示データには各フレームでの各関節(肩、ひじ、脚の付け根、ひざ)の3次元空間上における点の座標が記録されている。その座標点を元に四肢の中心を通る3次元上の直線を求め、この直線を体軸とする。

## 4. 教示データと単眼カメラ映像との照合

次に、単眼カメラ映像中の姿勢に最も近い姿勢を見つけるために、教示データとの照合を行う。教示データの各フレーム $n$ の姿勢を人物モデルに当てはめ(図2)、それを単眼カメラ映像に投影し、人物モデルの四肢の体軸

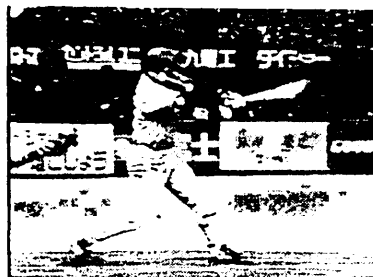


図1: 単眼カメラ映像の中心線。白丸は端点を、白線が中心線を表す。

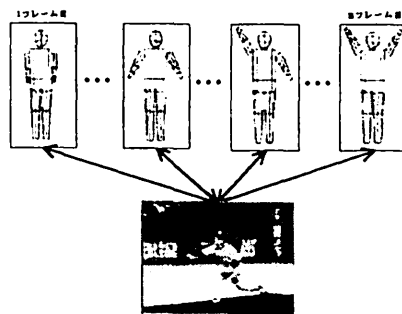


図2: 教示データと単眼カメラ映像の照合

と抽出されている中心線との差を計算する。そして、最も差の小さいフレーム $n^*$ の教示データの姿勢を、単眼カメラ映像中の姿勢であるとみなす。

## 4.1 重複データの削除

教示データ中に重複した姿勢データが存在する場合、同じ作業を繰り返すことになる。照合の効率化を計るため、以下のような方法で、教示データ中の類似したフレームの一方を削除する。まず、各姿勢データの部位毎に3次元空間上の移動距離を求める。この移動距離が一定値以下ならば同一の姿勢データとみなし、削除する。一定値以上の場合、異なる姿勢データとみなし削除はしない。この操作を全フレームに対して繰り返す。

## 4.2 中心線と体軸との距離の定義

中心線と体軸上の点との距離の計算はLoweの方法[1]を用いて行う。図2に示すように、投影面上に姿勢データより得られた体軸と、それに対応する2次元画像より求められた中心線があるとき、体軸と中心線とのずれは体軸上の点から中心線への距離として測定できる。体軸の方程式はヘッセの標準形[2]より、

$$X \cos \theta + Y \sin \theta = d$$

で与えられる。ここで、 $d$ は原点から中心線に下ろした垂線の長さ、 $\theta$ は垂線が $x$ 軸と成す角度であり、投影画像から求めることが出来る。今、図3のように中心線上の2点を $A(X_1, Y_1), B(X_2, Y_2)$ とすれば、 $d \cdot \cos \theta, \sin \theta$ はそれぞれ次式のように算出される。

$$d = \frac{X_1 Y_2 - X_2 Y_1}{L}, \cos \theta = \frac{Y_2 - Y_1}{L}, \sin \theta = \frac{X_1 - X_2}{L}$$

ただし、このときの距離  $L$  は次式とする。

$$L = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

ここで、体軸上の点を  $(X, Y)$  とすれば、この点から中心線までの有効距離  $e$  は、

$$e = X \cos \theta + Y \sin \theta - d$$

となる。このとき、あるパーツ  $k$  (例えば上腕など) について、体軸上の点  $(X_i, Y_i)$  について距離  $e_i$  を定義する。

$$e_i = X_i \cos \theta + Y_i \sin \theta - d$$

これを多数の点について連立し、距離が最小となるようにパーツを移動させるパラメータを反復計算により求める。反復計算終了後のパーツ  $k$  での距離を  $e'_k$  とする。あるフレーム  $n$  において全てのパーツについての距離を  $e_n$  とし、前後フレームで平滑化したものを  $e_w$  とすると、

$$e_n = \sum_i e'_i$$

$$e_w = \frac{1}{3}(e_{n-1} + e_n + e_{n+1})$$

ここで、

$$n^* = \arg \min_n e_w$$

となる第  $n^*$  フレーム目の教示データの姿勢が最も単眼カメラ画像中の姿勢であると仮定する。

## 5. 実験

本手法の有効性を示すため、プロ野球選手のバッティンしている画像(図 5)を対象に実験を行った。また、あらかじめバッティングフォームの教示データとして200フレーム分の3次元姿勢データを得ておいた。本手法による照合結果を図 4 と図 6 に示す。図 4 において白い点が最も中心線と体軸に近い教示データである。図 6 において、線分は画像中の人物の四肢の中心線線分、丸はその端点を表す。また直線は、体軸を通る線である。また、図 7 は照合結果より得られた3次元モデルである。

## 6. おわりに

2次元画像から、3次元の姿勢データを基に照合することにより3次元姿勢を推定する方法を提案した。また、照合作業の改善のために、重複データの削除やフィッティング誤差の平滑化を提案した。

今後の課題として、照合時に用いる体軸を8本に拡張することや、距離計算において各軸の向きを考慮に入れることが挙げられる。

## 参考文献

- [1] D.G.Lowe, "Fitting parameterized three-dimensional models to images", *IEEE Trans. Pattern Anal. Machine Intel.*, Vol.13, No.5, 1991, pp.441-450.
- [2] 矢野健太郎, "立体解析幾何学", 養華房, 1970.

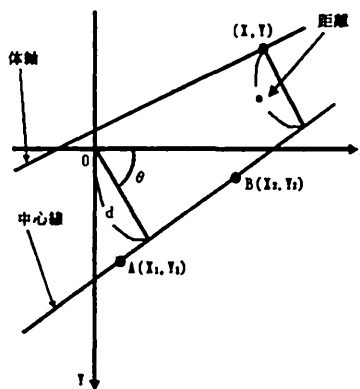


図 3: 投影面上の中心線と体軸

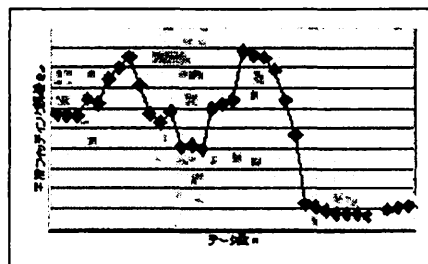


図 4: 単眼カメラ画像と各教示データとの距離



図 5: 原画像



図 6: 照合結果



図 7: 3次元モデルによる推定された姿勢の復元結果