

# 論 文

## 多視点動画像から的人間動作の追跡と再構成

佐藤 明知<sup>†</sup> 川田 聡<sup>†\*</sup> 大崎 喜彦<sup>†\*\*</sup> 山本 正信<sup>†</sup>

### 3D Model-Based Tracking of Human Actions from Multiple Image Sequences

Akitsu SUGU SATO<sup>†</sup>, Satoshi KAWADA<sup>†\*</sup>, Yoshihiko OSAKI<sup>†\*\*</sup>,  
and Masanobu YAMAMOTO<sup>†</sup>

あらまし 本論文は、多関節構造をもつ人体モデルを用いて、多視点動画像から人体の運動を3次元的に追跡する手法を提案する。本手法は、追跡開始フレームにモデルと人体像の照合を行い、追跡は位置と姿勢の変位を多視点動画像から計算し、逐次累積することにより行われる。変位量の計算は線形連立方程式を解くだけでよい。この計算量はカメラの数が増えてもカメラ1台の場合とほぼ同じである。実際の動画像の解析結果を通して、本手法がオクルージョンに強く信頼度の高い測定法であることを実証する。

キーワード 多視点動画像、動画像処理、コンピュータアニメーション、ロボットモデル、時空間こう配法、人間動作

#### 1. まえがき

人間の動作の測定は、ロボットやアニメーションキャラクターの動作生成、臨場感通信のための動画像の高能率符号化、医療におけるリハビリテーションの評価等、多くの分野で利用されている。人間の動作を測定する装置として、データスーツやデータグローブなどが普及しているが、これらは演技者に大きな負担をかけることが多く、自然な動作を測定することが困難である。

これに対し、ビデオカメラから得られた動画像を利用すれば、人体に接触することなく、自然な状態で動作をとらえることが可能である。動画像を使って動作を解析するとき、人体は複雑な構造をもち、複雑な動作を行うため、多関節構造をもつ人体モデルの利用が一般的となっている[2],[3],[6],[10],[13]~[15]。このモデルと人体像との照合により、人体の姿勢および動作を測定することができる。この測定はカメラが1台あれば可能であるが、(1)オクルージョンが発生したとき、隠された部分の動きを知ることが難しい、(2)カメラの光軸方向への動きは見かけの動きが小さくなるため測定結果の信頼性が低い、(3)人体モデルが一般に近似モデルであるため、正確な3次元照合が難し

い、などの問題点が挙げられる。そこで、異なる位置に置かれた複数台のカメラの利用が考えられている。

RehgとKanade[12]は、ステレオカメラを使って手の動きの3次元追跡を行った。GavrilaとDavis[1]は、人体の周囲に4台のカメラを配置し、ダンスをしている人間全体の追跡を行っている。KakadiarisとMetaxas[4]は、3台のカメラを互いに直交するように配置し、常に信頼度の高い動作の測定結果が得られるようにカメラの自動選択を行っている。

これらの研究は、上記の三つの問題点すべてを同時に解決しているわけではない。また、多関節モデルと多視点カメラからの画像との照合をフレームごとに行っているため、照合のための探索に膨大な計算量を必要とする。Kakadiarisらの研究[4]では、カメラの選択により照合のための計算量は減少すると考えられるが、カメラ選択のための計算が必要である。

本論文では、山本らによって提案された単一カメラによる人体の追跡法[15]を多視点カメラの場合に拡張する<sup>(注1)</sup>。すなわち、モデルと人体像の照合は追跡開始フレームだけで行い、その後は位置と姿勢の変位分を計算し、逐次累積していく。変位分の計算は線形連立方程式を解くだけでよい。この計算量はカメラの数が増えてもカメラ1台の場合とほぼ同じであることを2.で示す。この拡張された手法により、上記の三つの

<sup>†</sup> 新潟大学工学部情報工学科, 新潟市

Faculty of Engineering, Niigata University, Niigata-shi, 950-21 Japan

\* 現在, 伊藤忠テクノサイエンス

\*\* 現在, 安藤電気

(注1): 本論文は、文献[5]を発展させたものである。

問題点すべてが同時に解決できることを 3. で実証する。更に、多視点画像による観測の冗長性を利用して、測定精度の向上が可能であることも併せて示す。

## 2. 多視点動画からの動作推定

人体の動きを表現する運動モデルを定義し、このモデルの運動パラメータを多視点動画から推定する方法を示す。

### 2.1 人体の運動モデル

人間の体を頭部、腹部、胸部、腰部、左右上腕、下腕、上脚、下脚に相当する 12 個の剛体の接続構造で表す。各部を多面体で近似し、これを CAD モデリングシステム [7] で作成する。それぞれの物体には、異なるラベルが付けられている。作成したモデルを図 1 に示す。各物体には物体固有の座標系が付加されておりこの座標系の原点を各関節の位置におく。人体の接続構造は図 2 に示されるような階層構造である。すなわち腹部を根 (root) とし矢印で結んだ方向に親子関係を結ぶことにより表す。

各部分の運動を各物体座標系での原点周りの回転を表す行列  $Q_{b_j}$  と並行移動を表すベクトル  $S_{b_j} = (S_{x_j}, S_{y_j}, S_{z_j})$  によって定義する。ここで、 $j$  は物体の番号である。この物体上の運動前の点の座標値を  $p_j$  とすると、運動後の座標値  $p'_j$  は、

$$p'_j = Q_{b_j} p_j + S_{b_j} \quad (1)$$

となる。

回転行列  $Q_{b_j}$  は、各軸周りの回転量  $\theta_{x_j}, \theta_{y_j}, \theta_{z_j}$  が小さいとすれば、以下のように表される。

$$Q_{b_j} = \begin{pmatrix} 1 & -\theta_{z_j} & \theta_{y_j} \\ \theta_{z_j} & 1 & -\theta_{x_j} \\ -\theta_{y_j} & \theta_{x_j} & 1 \end{pmatrix} \quad (2)$$

各物体の回転量、並行移動量を運動パラメータとし、ベクトル  $\phi$  で表す。すなわち、

$$\phi = (S_{x_1}, S_{y_1}, S_{z_1}, \theta_{x_1}, \theta_{y_1}, \theta_{z_1}, \dots, S_{x_j}, S_{y_j}, S_{z_j}, \theta_{x_j}, \theta_{y_j}, \theta_{z_j}, \dots)^T$$

### 2.2 カメラモデル

各カメラに座標系を設定する。カメラ座標系は、視点を原点とする 3次元座標系  $(x_i, y_i, z_i)$  とする ( $i$  はカメラ番号)。平面  $z_i = 1$  を投影面とし、投影面座標系を  $(X_i, Y_i)$  とする。3次元空間中の点は、視点を

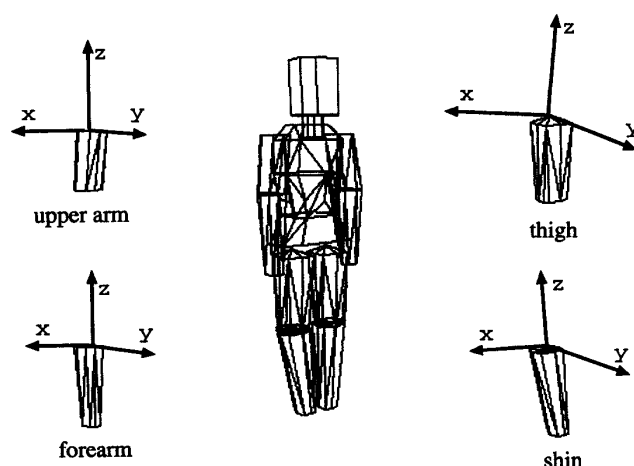


図 1 人体モデル  
Fig.1 Body model.

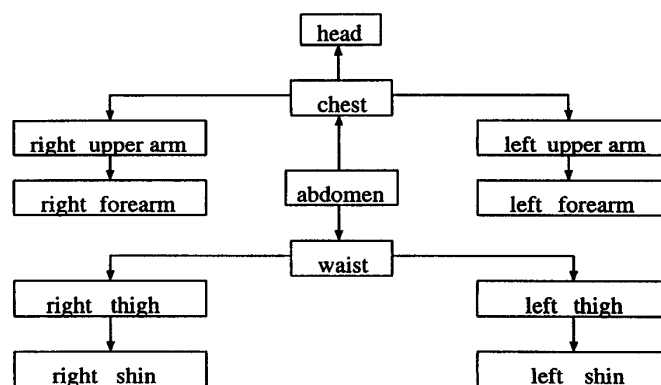


図 2 各物体の相互関係  
Fig.2 Articulated bodies.

中心に投影面上に中心投影されるとする。3次元座標  $(x_i, y_i, z_i)$  と投影面座標  $(X_i, Y_i)$  との関係は、

$$\begin{cases} X_i = f_{x_i} \frac{x_i}{z_i} \\ Y_i = f_{y_i} \frac{y_i}{z_i} \end{cases} \quad (3)$$

と表す。ここで、 $f_{x_i}, f_{y_i}$  は  $x$  方向、 $y$  方向のスケール変換である。

この 3次元座標値が  $(\delta x_i, \delta y_i, \delta z_i)$  変位したとき、投影面上での変位  $(\delta X_i, \delta Y_i)$  は、変位量が小さいとき次式で近似できる。

$$\begin{cases} \delta X_i = (f_{x_i} \delta x_i - X_i \delta z_i) / z_i \\ \delta Y_i = (f_{y_i} \delta y_i - Y_i \delta z_i) / z_i \end{cases} \quad (4)$$

3次元シーンに固定された座標系をワールド座標系としたとき、カメラ座標系で表された点  $p_{c_i}$  は、ワールド座標系では次の座標変換式により  $p_w$  と表される。

$$p_w = R_{c_i} p_{c_i} + T_{c_i} \quad (5)$$

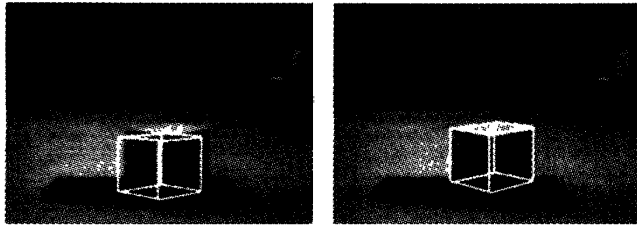


図3 カメラキャリブレーションの結果  
Fig.3 Result of camera calibration.

ここで,  $R_{c_i}$  は, ワールド座標系に対するカメラの姿勢を表し,  $T_{c_i}$  は, カメラの視点の位置  $(T_{x_i}, T_{y_i}, T_{z_i})^T$  を表す.

### 2.3 カメラキャリブレーション

カメラのスケール変換および位置, 姿勢はカメラキャリブレーションを行うことにより求める.

本論文では, 大きさが既知である立方体を画面中央に置き, Loweの方法[9]によりカメラキャリブレーションを行った. この立方体に固有の座標系をワールド座標系とする. 図3にモデルの初期位置(左)および, カメラキャリブレーションの結果(右)として白線で描いた立方体のモデルを画像に重ねて表示する. 図3右のようにモデルのりょう線と立方体のりょう線が一致したときカメラ座標系からワールド座標系への変換  $R_{c_i}, T_{c_i}$  が求まる.

### 2.4 人体モデルの初期フレーム照合

本論文では, 追跡開始フレームにおいて人体のモデルを人体像に次のように手動で照合する. カメラキャリブレーションにより各カメラモデルのパラメータが確定する. 得られたカメラモデルにより, 人体のモデルの投影された像と, 画像上の人体像が一致するようにモデルの位置, 姿勢を手動で調整する. 物体  $j$  の物体座標系とワールド座標系との座標変換は,

$$p_w = R_{b_j} p_{b_j} + T_{b_j} \quad (6)$$

となる. ここで  $p_w, p_{b_j}$  は, それぞれワールド座標系, 物体  $j$  の座標系で表された座標値であり,  $R_{b_j}$  はワールド座標系に対する物体座標系の姿勢を表す変換行列であり,  $T_{b_j}$  は原点の位置を表す変換ベクトルである. 人体モデルが人体像に一致したとき, 変換  $R_{b_j}, T_{b_j}$  が求まる. このとき, モデルの投影像から領域画像(図4左)および, 距離画像(図4右)が得られる. 領域画像は, カメラから見た人体の各部を表すラベル画像である. オクルージョンの有無はこの領域画像から知ることができる.

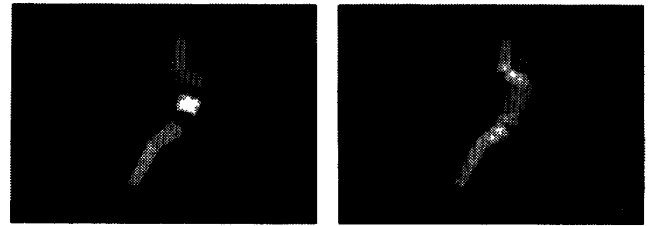


図4 図14のカメラ1から見た領域画像(左)と距離画像(右)  
Fig.4 Region image (left) and depth image (right).

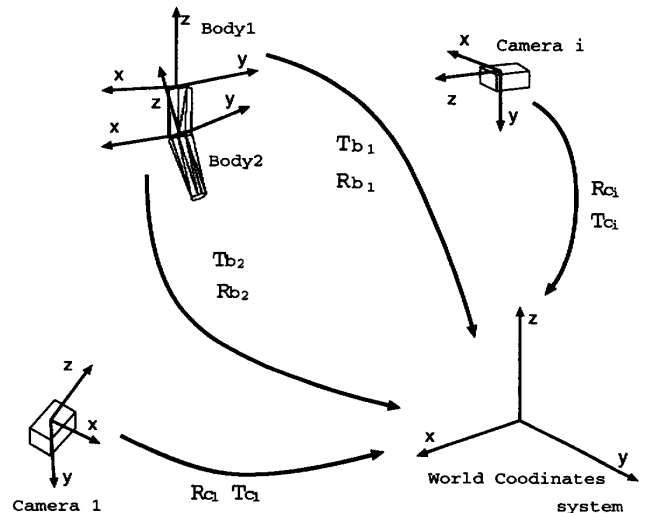


図5 カメラ座標系, 物体座標系, ワールド座標系  
Fig.5 The camera-centered, body, and world coordinates system.

カメラが1台の場合, カメラの奥行方向のモデルの姿勢を決めることは困難であるが, 多視点画像上でのモデルの投影像と人体像を一致させることにより, 人間の位置, 姿勢を正確に知ることができる.

### 2.5 時空間こう配法による運動パラメータの推定式

運動パラメータを推定するための方程式を導く.

まず, 人体上の点の運動後の座標を求める. 図5を例とする. この図において物体1と物体2は, 物体1を親, 物体2を子とする親子関係が結ばれている. 物体2上の点のカメラ  $i$  の座標系での座標を  $p_i$ , 運動後の座標を  $p'_i$  とすると,  $p'_i$  は,  $p_i$  を次の順序で座標変換と運動を行うことにより求めることができる.

- (a) カメラ座標系からワールド座標系へ変換
- (b) ワールド座標系から物体座標系へ変換
- (c) 物体の運動
- (d) ワールド座標系に変換
- (e) 親の物体の物体座標系に変換
- (f) 親の運動

(g) ワールド座標系に変換

(h) カメラ座標系に変換

もし、物体 1 に更に親の物体があるならば、この親の物体の運動を行うため、手順 (e), (f), (g) を繰り返す。

この結果、図 5 の物体 2 では、以下のようになる。

$$\begin{aligned} \mathbf{p}'_i &= R_{c_i} \left( R_{b_1} \left( Q_{b_1} R_{b_1}^{-1} \left( R_{b_2} \left( Q_{b_2} R_{b_2}^{-1} \right. \right. \right. \right. \\ &\quad \left. \left. \left. \left( R_{c_i} \mathbf{p}_i + \mathbf{T}_{c_i} - \mathbf{T}_{b_2} \right) + \mathbf{S}_{b_2} \right) \right. \right. \\ &\quad \left. \left. + \mathbf{T}_{b_2} - \mathbf{T}_{b_1} \right) + \mathbf{S}_{b_1} \right) + \mathbf{T}_{b_1} - \mathbf{T}_{c_i} \end{aligned} \quad (7)$$

ワールド座標系に対して、カメラおよび、物体の位置、姿勢が得られているので、 $\mathbf{p}'_i$  は運動パラメータを変数とする関数である。 $\mathbf{p}'_i$  を運動パラメータに関してテイラー展開して 2 次以上の高次の微小項を無視すれば、3 次元変位  $\delta \mathbf{p}_i = \mathbf{p}'_i - \mathbf{p}_i$  は、運動パラメータの線形結合で

$$\begin{aligned} \delta \mathbf{p}_i &= \mathbf{p}'_i - \mathbf{p}_i \\ &= (\delta x_i, \delta y_i, \delta z_i)^\top \\ &= \sum_{j=1}^2 \left\{ \frac{\partial \mathbf{p}'_i}{\partial S_{x_j}} S_{x_j} + \frac{\partial \mathbf{p}'_i}{\partial S_{y_j}} S_{y_j} + \frac{\partial \mathbf{p}'_i}{\partial S_{z_j}} S_{z_j} \right. \\ &\quad \left. + \frac{\partial \mathbf{p}'_i}{\partial \theta_{x_j}} \theta_{x_j} + \frac{\partial \mathbf{p}'_i}{\partial \theta_{y_j}} \theta_{y_j} + \frac{\partial \mathbf{p}'_i}{\partial \theta_{z_j}} \theta_{z_j} \right\} \\ &= J_i(\phi) \phi \end{aligned} \quad (8)$$

と表される。ここで、 $J_i(\phi)$  は、ヤコビ行列である。

カメラ  $i$  において得られる動画像を  $E(X_i, Y_i, t)$  とする。点  $(X_i, Y_i)$  の単位時間後の移動ベクトル  $(\delta X_i, \delta Y_i)$  は、以下の式により拘束されることが知られている。

$$E_{X_i} \delta X_i + E_{Y_i} \delta Y_i + E_t = 0 \quad (9)$$

ここで、 $(E_{X_i}, E_{Y_i})$  は画像の空間こう配であり、 $E_t$  は時間こう配である。

式 (8) の  $\delta x_i, \delta y_i, \delta z_i$  を式 (4) に代入し、式 (4) の  $\delta X_i, \delta Y_i$  を式 (9) に代入すれば、運動パラメータ  $\phi$  を未知ベクトルとする次の線形推定式が導出される。

$$E_{p_i} J_i(\phi) \phi = E_t \quad (10)$$

但し

$$E_{p_i} = \left[ -\frac{f_{x_i} E_{X_i}}{z_i}, -\frac{f_{y_i} E_{Y_i}}{z_i}, \frac{X_i E_{X_i} + Y_i E_{Y_i}}{z_i} \right]$$

ここで、奥行の距離  $z_i$  は、人体像と投影像を一致させたとき、モデルの距離画像から与えられる。

## 2.6 多視点動画像からの運動パラメータの決定

カメラ  $i$  により撮影された画像上の各点が、人体のどの部分に属するのか、領域画像から決定する。決定された各点に対して、導かれた推定式 (10) を連立させることにより次の線形連立推定式が得られる。

$$A_i \phi = B_i \quad (11)$$

ここで、 $A_i$  は係数行列、 $B_i$  は定数ベクトルである。

カメラの数を  $n$  個とした場合、それぞれのカメラシステムに対し同様の線形連立推定式が導出される。求める  $\phi$  は人体モデル固有の運動パラメータであるので、これらを更に連立させることができる。

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \phi = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \quad (12)$$

運動パラメータは、線形連立推定式 (12) を最小 2 乗法を用いて解くことによって得られる。この推定は、個々のカメラに対する推定式 (11) を逐次解くのではなく、式 (12) を一度解くだけでよい。計算の手間はカメラの数に無関係である。但し、推定精度の向上のためには、画像をあらかじめスムージングしておくこと、および式 (12) の反復適用 [15] を行っている。

運動パラメータの推定と、それに基づくモデルの移動、を繰り返すことにより人間の動きを追跡する。

## 3. 多視点動画像の効果

多視点動画像を用いることにより、オクルージョンが解消できること、光軸方向の動きに対して信頼性の高い結果が得られること、および同時に多方向から観測されることによる冗長性を利用した推定精度の向上が期待される。本章では、これら多視点動画像の効果を実証するための実験を行った。各実験の目的を以下のように定めた。

- (a) オクルージョンの解消および精度の向上
- (b) 光軸方向の動きに対しての信頼性の向上
- (c) 複雑な動作の追跡

以上三つの実験について以下に説明と結果を示す。

### 3.1 動画像入力システム

最大 4 台までの動画像をコンピュータに入力する方法を示す。各カメラは、1 台の同期信号発生器より発生した外部同期信号により駆動されている。従って、

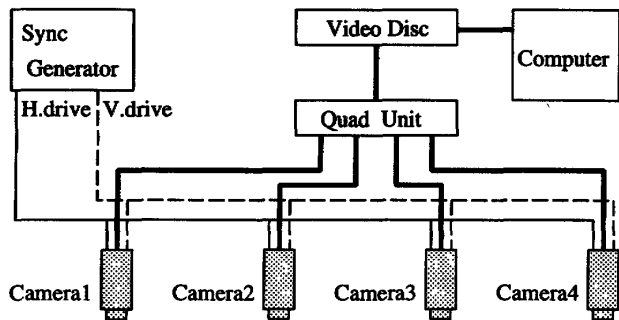


図6 動画画像入力システム

Fig. 6 A system for acquisition of multiple image sequences.

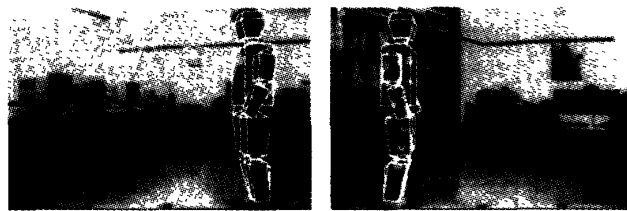


図8 初期状態

Fig. 8 Fitting a model to initial frame.

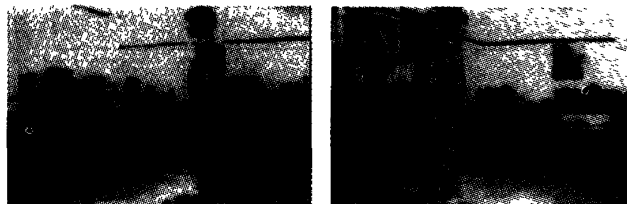


図9 追跡結果

Fig. 9 Tracking result at 50th frame.

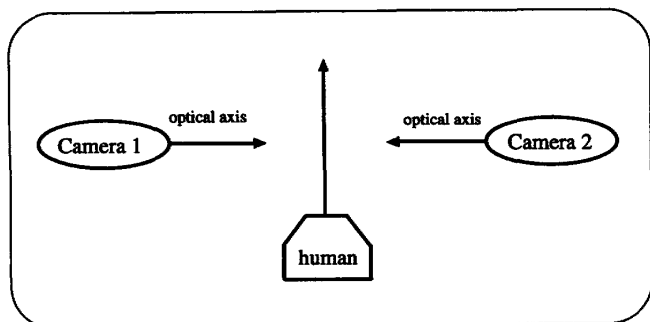


図7 カメラの配置

Fig. 7 An arrangement of two cameras.

各映像信号は完全に同期されている。各カメラからの画像を4画面分割ユニット (SONY, YS-Q400) に入力する。このユニットは、四つの入力画像を縦横 1/2 に縮小し、4分割した画面内にそれぞれ入力し、1枚の画像として出力するものである。この出力された画像をいったんビデオディスク (松下, LQ-4100) に保存しておき、追跡を行うとき計算機に入力する。

### 3.2 オクルージョンの解消および、精度の向上

カメラ1台からの画像では、体の一部がオクルージョンにより見えなくなる場合がある。例えば、図7のように、歩行動作している人間をカメラ1から観測したとき、この人間の右側手足は、隠れにより常時観測することができない。このとき、カメラ1と向かい合うようにカメラ2を配置し、2台のカメラで観測すれば人体全体を追跡することができる。

追跡対象が歩行という簡単な動作であるため、モデルは胴体と頭および左右上腕、下腕、上脚、下脚の九つの物体によって構成した。運動パラメータは、胴体については前後方向、左右方向の並進速度  $S_{x1}$ ,  $S_{y1}$  および、体軸周りの回転角速度  $\theta_{z1}$  とした。また、左右の上腕、下腕、上脚、下脚については、直立した状態で胴体の左右方向の軸周りの回転角速度のみとし、

順に  $\theta_{x2}$ ,  $\theta_{x3}$ , ...,  $\theta_{x7}$ ,  $\theta_{x8}$ ,  $\theta_{x9}$  とした。つまり、運動パラメータ列  $\phi$  は、

$$\phi = (S_{x1}, S_{y1}, \theta_{z1}, \theta_{x2}, \theta_{x3}, \theta_{x4}, \theta_{x5}, \theta_{x6}, \theta_{x7}, \theta_{x8}, \theta_{x9})^T$$

の合計 11 個となる。図8に追跡開始時刻の画像にモデルの輪郭を白線で重ねて描く。

カメラ1より見た画像 (図8左) には、胴体により右上腕が隠されている。そのためカメラ1で観測したときの運動パラメータの線形推定式では、領域画像中に右上腕に対応する領域が存在しないので、右上腕のパラメータ  $\theta_{x6}$  に対応する係数列がすべて 0 となる。このため解  $\theta_{x6}$  を定めることができない。

また、カメラ2からの観測画像 (図8右) には右上腕は見えているが、逆に左上腕が見えていない。そのためカメラ2から得られる運動パラメータの線形推定式では、同様に、左上腕に対応するパラメータ  $\theta_{x2}$  の係数列がすべて 0 となる。従って、この推定式からは  $\theta_{x2}$  を定めることができない。そこで、この二つのカメラからの連立方程式を更に連立させる。得られた連立方程式の係数行列には 0 列がなくなるため運動パラメータを一意に決めることが可能となる。2台のカメラから追跡を行ったときの 50 フレーム目の結果を図9に示す。

図10は、カメラ1 (◇印) またはカメラ2 (+印) からの画像のみでそれぞれ追跡を行った結果と、2台のカメラ (□印) と使って追跡を行った結果を、70 フレームまでの胴体の体軸周りの回転角速度で示してい

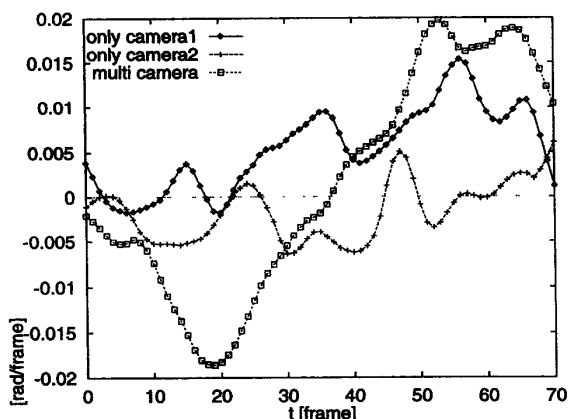


図 10 胴体の回転量  
Fig. 10 Angular velocity of rotation.

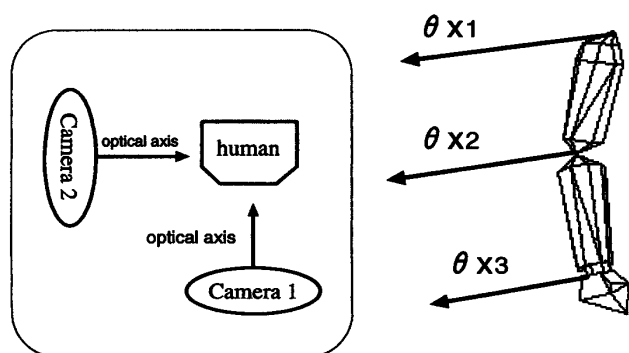


図 11 カメラの配置と CAD モデル  
Fig. 11 An arrangement of two cameras (left) and CAD model (right).

る。70 フレームは歩行のほぼ 1 周期に相当する。歩行動作においては、胴体の体軸周りの回転角速度は歩行の周期に等しい周期の正弦関数で近似できるが、単一のカメラによる追跡では、偏りや小さなぶれが見られる。また、同じ胴体を追跡しているにもかかわらず、カメラ 1 とカメラ 2 の追跡結果が異なっている。このような結果は、画像上の雑音により生じたと考えられる。しかし、2 台のカメラによる追跡では、正弦関数に近い結果が得られており、高い精度の推定結果を得ることができた。

### 3.3 信頼性の向上

人間の動きがカメラの光軸方向に向っているとき、画像上での変化が非常に小さく、光軸方向の動きを正確に復元することができない [2]。実際、図 11 左に示すように人間がカメラ 1 を向き、右腕を真下から前方へ振りあげる動作の追跡を行った。図 11 右にこの実験に用いた人体のモデルを示す。簡単のため、モデルは上腕と下腕と手首のみで構成され、動きは胴体の左

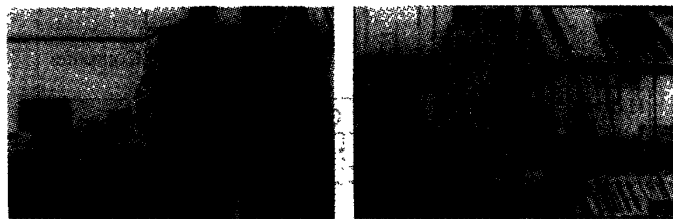


図 12 モノカメラによる実験結果  
Fig. 12 Tracking result from a view of camera 1 alone.



図 13 マルチカメラによる実験結果  
Fig. 13 Tracking result from multiple views of cameras 1 and 2.

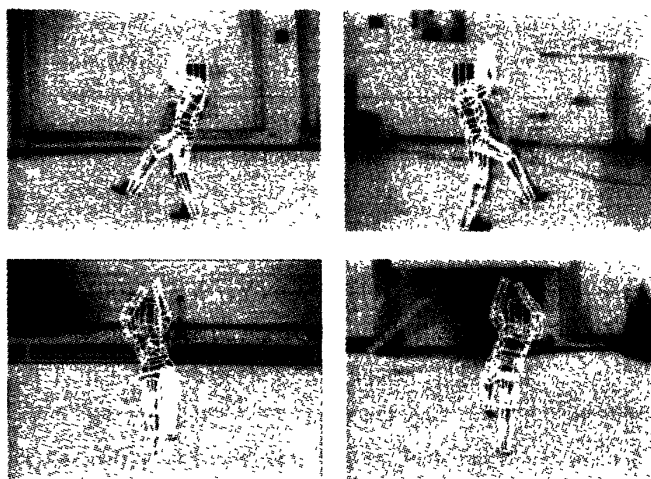


図 14 追跡開始時刻の画像  
Fig. 14 Fitting a model to the body images at initial frames.

右方向を回転軸とする回転のみとした。

図 12 左に、カメラ 1 で追跡を行った結果を示す。右の画像には追跡結果を人体の真横に配置したカメラ 2 (図 11 左) からの画像に重ねて表示している。図 12 左左では追跡しているように見えるが、同図右では、わずかなノイズにも敏感に反応し、大きなずれに至ることが示されている。正確な追跡は、カメラ 2 で行うことができる。しかし、動きの向きが異なる場合には、正確な追跡が可能な方のカメラを選択する必要性が生じてくる [4]。そこで、二つのカメラを併用する。その追跡の結果を図 13 に示す。カメラ 2 のみの場合と画像上で同程度に、どちらのカメラから見ても正確な追跡

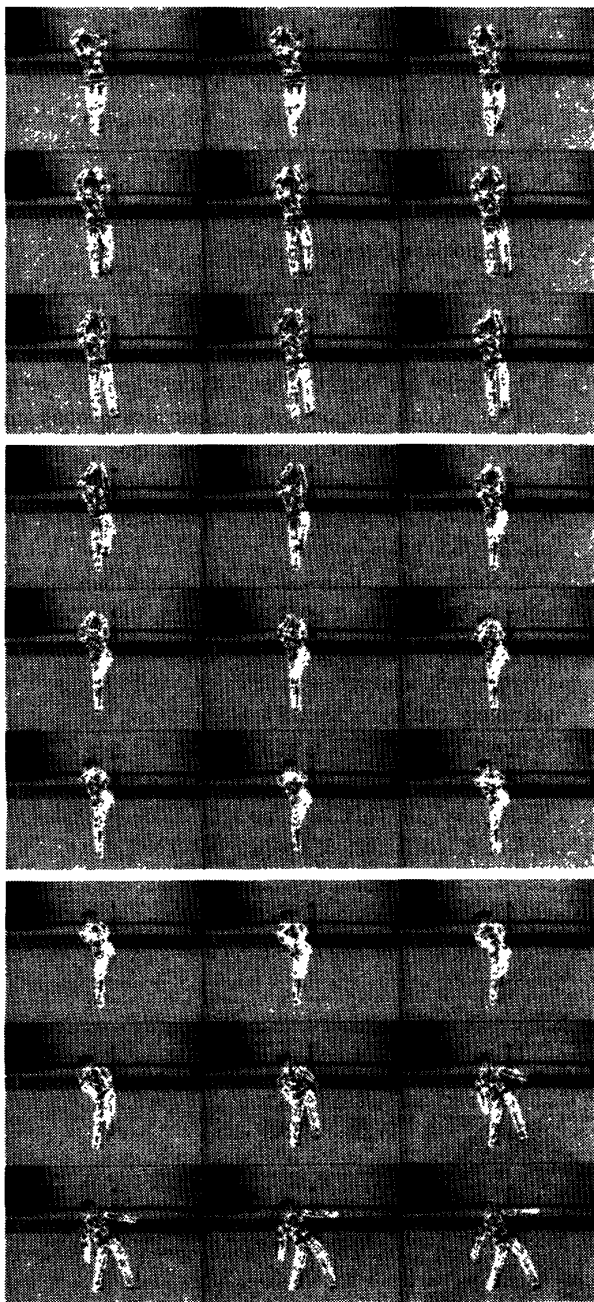


図15 追跡結果

Fig.15 Tracking result of pitcher's motion.

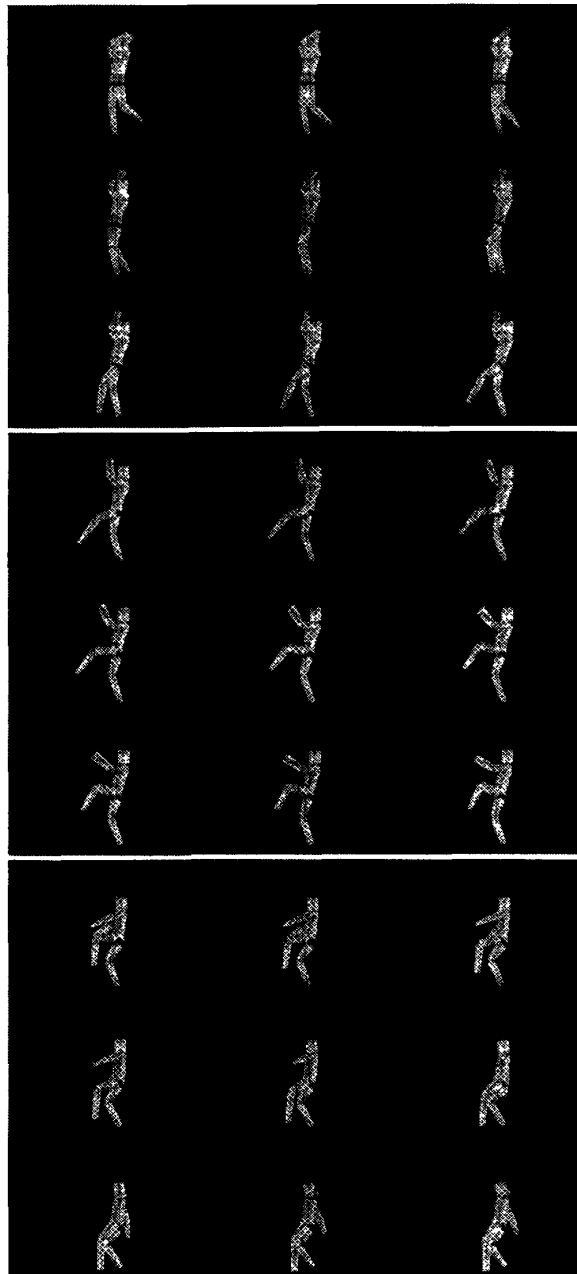


図16 CGアニメーション

Fig.16 CG animation.

が行われた。このことは、提案した手法は、カメラの選択の必要がなく、どのような動きに対しても信頼のおける測定が可能であることを示している。

### 3.4 カメラ4台による実験

複雑な動作の例として、野球の投球動作の追跡を行った。

この実験では、ほぼ人体の前後左右に4台のカメラを配置した。実験で用いたモデルは図1に示す12個の部分からなるモデルである。運動パラメータは各部

の3軸回転、腹部の3軸方向への並進の合計39個とした。

追跡開始フレームでのモデルフィッティングを図14に示す。図15にカメラ3から見た追跡の結果を開始フレームから2フレームおきに示す。

図14では、カメラ1において左腕に、カメラ2において右腕にオクルージョンが起こっている。カメラ3では左脚が光軸方向の動きであること、また、左上脚が徐々に隠れていくため、カメラ3の画像のみの追跡では、信頼性が低く、途中までしか追跡できないこ

とになる。しかし、四つのカメラから得られる推定式を連立させて解くことによりすべての運動パラメータを高い信頼度で求めることが可能となる。

得られた投球動作の追跡データおよび、追跡に用いた人間の幾何モデルをもとに、コンピュータアニメーションを作成し、3次元的に動作が復元できることを確認した。その結果の一部を図16に示す。

#### 4. む す び

本論文では、多視点動画像を用いオクルージョンに強く、精度の高い追跡手法を提案した。この手法により動きの向きによらない信頼度の高い人体の3次元追跡が可能となった。

追跡結果は3次元の動きとして得られている。従ってコンピュータアニメーションにより任意の視点から人間の動作を再現することが可能となった。

今後の課題として、まず、追跡開始画像でのモデルフィッティングの自動化が挙げられる。この方法として、人体モデルと人体像を一致させる手法[1],[4],[11],[12]との融合が考えられる。本手法では、推定値にわずかながらでも誤差<sup>(注2)</sup>があれば誤差が累積し、長時間経つと追跡に失敗する場合がある。誤差の修正の方法として、ステレオカメラシステムによる3次元モデルフィッティングや、誤差が大きくなるたびにそこを追跡開始フレームとしてモデルと人体像を再度照合し追跡を行う方法が考えられる。

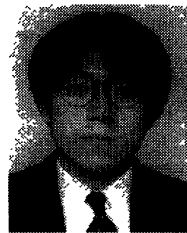
#### 文 献

- [1] D.M. Gavrila and L.S. Davis, "3-D model-based tracking of humans in action: A multi-view approach," Proc. of IEEE CVPR '96, pp.73-80, 1996.
- [2] L. Goncalves, E.D. Bernardo, E. Ursella, and P. Perona, "Monocular tracking of the human arm in 3D," Proc. of 5th ICCV, pp.764-770, 1995.
- [3] D. Hogg, "Model-based vision: a program to see a walking person," Image and Vision Computing, vol.1, no.1, pp.5-20, 1983.
- [4] I.A. Kakadiaris and D. Metaxas, "Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection," Proc. of IEEE CVPR '96, pp.81-87, 1996.
- [5] 川田 聡, 佐藤明知, 大崎喜彦, 山本正信, "人間動作のマルチカメラによる追跡と仮想世界での再現," 信学技報, PRU95-96, 1995.
- [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), vol.J74-D-II, no.3, pp.376-387, March 1991.

- [7] K. Koshikawa and Y. Shirai, "A 3-D modeler vision research," in Proc. of the Int. Conference on Advanced Robotics, pp.185-190, 1985.
- [8] M.K. Leung and Y.H. Yang, "A region based approach for human body motion analysis," Pattern Recognition, vol.20, no.3, pp.321-339, 1987.
- [9] D.G. Lowe, "Fitting parameterized three dimensional models to images," IEEE PAMI, vol.13, no.5, pp.441-450, 1991.
- [10] J.O' Rourke and N.J. Badler, "Model-based image analysis of human motion using constraint propagation," IEEE PAMI, vol.2, no.6, pp.522-536, 1980.
- [11] J. Ohya and F. Kisino, "Human posture estimation from multiple images using genetic algorithm," 12th ICPR, pp.750-753, 1994.
- [12] J.M. Rehg and T. Kanade, "DigitEyes: Vision-based human hand tracking," CMU-CS-93-220, 1993.
- [13] K. Rohr, "Towards model-based recognition of human movements in image sequences," CVGIP: Image Understanding, vol.59, no.1, pp.94-115, 1994.
- [14] 島田伸敬, 白井良明, 久野義徳, "確率に基づく探索と照合を用いた画像からの手指の3次元姿勢推定," 信学論 (D-II), vol.J79-D-II, no.7, pp.1210-1217, July 1996.
- [15] 山本正信, 川田 聡, 近藤拓也, 越川和忠, "ロボットモデルに基づく人間動作の3次元動画追跡," 信学論 (D-II), vol.J79-D-II, no.1, pp.71-83, Jan. 1996.

(平成8年10月21日受付, 9年1月14日再受付)



佐藤 明知

平7新潟大・工・情報工学卒。現在、同大大学院情報・計算機工学専攻博士前期課程在学中。動画像処理に関する研究に従事。



川田 聡

平6新潟大・工・情報工学卒。平8同大大学院情報工学専攻修士課程了。同年、伊藤忠テクノサイエンス(株)に入社。在学中は、コンピュータビジョンに関する研究に従事。

(注2): 誤差要因として、基本拘束式の近似誤差、人体境界付近での異なる動きの混入等が考えられる。追跡フレーム数は、画像の複雑さや、動作の速さに関係するが、現在のところ最長500フレームである。





大崎 喜彦

平6新潟大・工・情報工学卒。平8同大大学院情報工学専攻修士課程了。同年、安藤電気(株)に入社。在学中は、コンピュータビジョンに関する研究に従事。



山本 正信 (正員)

昭48九工大・工・制御卒。昭50東工大大学院修士課程了。同年、電総研入所。動画像処理、コンピュータビジョン等の研究に従事。平1~2カナダ国立研究協議会招聘研究員。昭62情処学会研究賞受賞。平4より新潟大工学部情報工学科教授。工博。情報処理学会、IEEE各会員。