

# 関節構造を持つ柔軟変形モデルを用いた人体運動の推定

小川原光一<sup>†</sup> 李 曉路<sup>†</sup> 池内 克史<sup>†</sup>

<sup>†</sup> 東京大学 〒153-8505 東京都目黒区駒場 4-6-1 東京大学生産技術研究所

E-mail: †{ogawara,gyouro,ki}@cvl.iis.u-tokyo.ac.jp

あらまし 本論文では、マーカレスモーションキャプチャの一手法として、多視点からの時系列ビデオ画像から輪郭情報を利用して対象の三次元復元を行い、これと人体の三次元形状モデルの位置合わせを行うことによって、人間の全身運動を連続推定する手法を提案する。本手法の特徴は、(1) 関節構造を持った人体の柔軟変形モデルを用いることによって肩の複雑な動きや腰のひねりなど繊細な動きに追従することができる点と、(2) ICP とロバスト推定を組み合わせ対応点探索を工夫することによって腕と胴の相互干渉やノイズに強い推定を可能にする点である。また、既存手法に対する優位性の評価を行い、さらに実環境において激しい運動の追跡が可能であることを示した。

キーワード マーカレスモーションキャプチャ、人体運動推定、多視点カメラ、柔軟変形モデル

## Human Motion Tracking by Articulated Deformable Model

Koichi OGAWARA<sup>†</sup>, Xiaolu LI<sup>†</sup>, and Katsushi IKEUCHI<sup>†</sup>

<sup>†</sup> The University of Tokyo Institute of Industrial Science, the University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

E-mail: †{ogawara,gyouro,ki}@cvl.iis.u-tokyo.ac.jp

**Abstract** This paper presents a novel whole body motion tracking method by aligning the 3D model of human body and reconstructed 3D shape of the target obtained from multiple video streams. The advantage of the proposed method is two fold: (1) delicate deformation around shoulder and waist can be accurately tracked by using an articulated deformable model, (2) combination of ICP and robust estimator with an improved correspondence search algorithm make it possible to estimate motion robustly against noise and interference between limb and torso. Also the advantage over the conventional method is evaluated and the power to track challenging whole body motion in real environment is presented.

**Key words** marker-less motion capture, human motion tracking, multi-view geometry, deformable model

### 1. はじめに

人体の全身運動を計測するいわゆるモーションキャプチャの技術は、映画やゲーム業界におけるキャラクターアニメーションの作成、スポーツ選手の運動解析、伝統舞踊に代表される無形文化財の電子保存など様々な分野で利用されている。しかし、現在よく用いられている磁気式・光学式などの方式では、いずれも体表面に多数の磁気センサもしくは反射マーカを取りつける必要があり、対象者に準備や運動の制限を強いるという問題がある。

本論文では、装着型センサやマーカを用いないマーカレスモーションキャプチャの一手法として、多視点からの時系列ビデオ画像から輪郭情報を利用して対象の三次元復元を行い、これと関節構造を持った人体の三次元柔軟変形モデルとの位置合わせを行うことによって、人間の全身運動を連続推定する手法を提案する。

#### 1.1 関連研究

広域監視やユーザインターフェースなどの用途でより生活に密着した全身運動計測のニーズが高まっており、身体への拘束がなくまた簡便に用いることが可能なマーカレスモーションキャプチャの研究開発が盛んに行われている [1]。

マーカレスモーションキャプチャでは、カメラからの入力画像をコンピュータビジョンの手法によって処理し運動推定を行う。このとき、単眼カメラを用いる方法 [2] ~ [5] と多視点カメラを用いる方法 [6] ~ [10] があるが、単眼では隠蔽や曖昧性を解決することが難しいことから、精度の高い運動情報を求めるために多視点カメラを用いた運動推定手法が近年多数研究されている。

Gavrila らは、超 2 次曲面 (super quadrics) を用いて人体の関節モデルを定義し、あらかじめ様々な姿勢パラメータとそのときの 2 次元投影画像の組をデータベースに蓄えておき、4 台のカメラ入力とデータベースとの比較から chamber distance

を指標にして最も類似した姿勢パラメータを推定する手法を提案した [6]。この方法は、初期姿勢の推定や運動追跡が失敗したときの回復に有効であるが、一般に運動追跡がうまく行っている場合には、一つ前の時刻における推定結果を初期値として用いその近傍を探索することによって、より頑健かつ高速な運動推定が可能になる。この考えに基づき、Delamarre らは円柱や直方体などで構成された人体の 3 次元関節モデルを定義し、カメラ画像に投影したモデル輪郭の周囲で画像エッジとの対応を取ることによって、姿勢パラメータの勾配を計算し運動推定を行った [7]。

Cheung らは、画像に投影された対象の輪郭情報から体表面上の 3 次元点である CSP (Colored Surface Point) を求め、異なる剛体運動を行う CSP 群を分離することによって関節位置と関節角度を計算し、人体の関節モデルと運動を同時に推定する手法を提案した [8]。この手法は、関節モデルを生成するために複雑な手順を要し、また肩や胴回りの微妙な変形を検出することは困難である。

人体は内部の関節構造とそれによって滑らかに変形する皮膚から構成されており、上で述べた剛体の関節モデルでは正確な運動推定は難しい。Ilic らは、陰関数で表現された体表面モデルを定義し、FFD (Free Form Deformation) を用いて人体モデルを滑らかに変形させ、上半身の形状推定を行った [9]。しかし、この方法は関節構造を陽に用いないため、運動推定には不向きである。

Kehl らは関節構造を持つ柔軟変形モデルを用いた運動推定手法を提案し、モデルの各頂点とその最近傍点の組から勾配法によって関節角・姿勢の推定を行った [10]。局所解に陥ることを防ぐために確率的に選択したモデル頂点群を使用し、また表面色の情報を最近傍探索に利用して腕と胴体の癒着を回避しているが、逆に癒着しない色の服装を強いるという欠点がある。

### 1.2 提案手法

本論文では、運動の連続性を仮定し、時刻  $t-1$  の姿勢が分かっているとして時刻  $t$  の姿勢を求める問題を扱う。初期姿勢については、例えば [5] の方法でおおまかな姿勢を推定した後本手法を適用するなどして求めることが可能である。

我々の提案手法の特徴は以下の 2 点である。

(1) 関節構造を持った人体の柔軟変形モデルを用いることによって肩の複雑な動きや腰のひねりなど繊細な動きに追従することを可能にする。また関節を持たない構造用のリンクを導入しモデルに局所的な剛性を持たせることによって、胸部や頭部においてより人体に近い変形を実現する。

(2) Iterative Closest Point (ICP) 法 [11] とロバスト推定を組み合わせ、法線付き Kd-Tree を対応点探索に用いることによって、腕と胴の相互干渉やノイズに強い推定を可能にする。Kehl らの手法と異なり表面色の情報を用いずに癒着の問題に対処するため、対象者の服装の色に影響されないという利点がある。

以降では、まず第 2 章で関節構造を持った柔軟変形モデルについて述べ、第 3 章でこのモデルを用いた運動推定手法について述べる。つぎに第 4 章で運動推定の評価と実環境における追跡例を紹介し、第 5 章で結論を述べる。

## 2. 関節構造を持つ柔軟変形モデル

### 2.1 人体モデルの定義

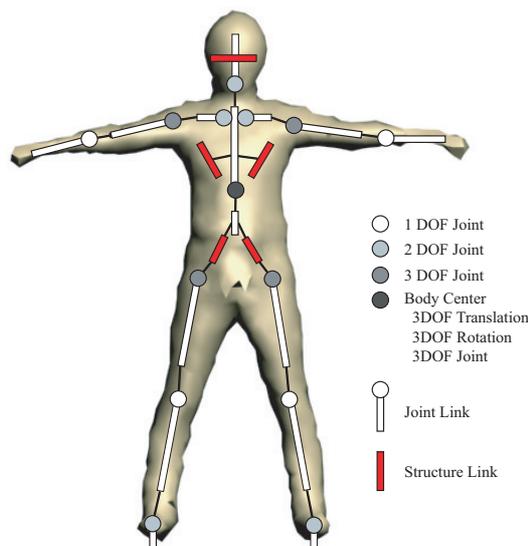


図 1 人体モデル

本研究で使用する人体モデルは、図 1 に示すように関節構造を表すリンクモデルと体表面を表すメッシュモデルによって構成され、関節角の変位に応じてメッシュは近傍のリンクに追従して滑らかに変形する。

リンクモデルは、関節角 2 9、並進 3、回転 3 の自由度を持っており、肩回りの複雑な変形を表現するために、鎖骨に相当する自由度が入っている点の特徴である。また、メッシュの各頂点は近傍のリンクの影響を受けるため、不自然な変位が生じないように、関節構造を表すリンク (Joint Link) 以外に肋骨などに相当する構造を維持するためのリンク (Structure Link) を導入している。

またメッシュモデルは、多視点カメラの入力から第 4.2.1 章で述べる方法によって作成する。

### 2.2 モデルの変形

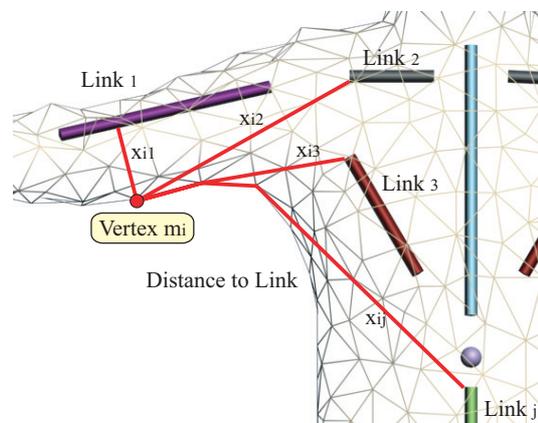


図 2 メッシュ頂点とリンクまでの距離の定義

Kehl らのモデル [10] で用いている変形手法 [12] は、リンク構造に沿ってメッシュの制御点を配置し、関節角の変形に合わ

せて制御点を移動もしくは縮退させ、これに基づいてメッシュの変形を実現している。しかし、図2の肩のように多数のリンクからの影響を考慮しなければならない場合には制御点の計算が複雑になる。

そこで本研究では、体中心座標系におけるメッシュの  $i$  番目の頂点の位置を、関節角  $\theta$  の関数  $m'_i(\theta)$  として次のように決定する。

まず、図2に示すように、頂点  $i$  からリンク  $j$  までの距離  $x_{ij}$  を、頂点  $i$  からリンク  $j$  までのメッシュを貫通しない最短距離として定義する。 $x_{ij}$  を求めるには、まずメッシュの頂点とリンク  $j$  をノードとし、メッシュのエッジを弧としたグラフを作成する。次に、各頂点とリンク  $j$  を結ぶ線分がメッシュを貫通しない場合には、新たにその頂点からリンク  $j$  まで弧を設定する。弧の重みをノード間のユークリッド距離とし、各頂点からリンクまでの最短距離を Dijkstra 法によって求めることで  $x_{ij}$  が計算される。

次に、ある頂点  $i$  に対するリンク  $j$  からの重みを式(1)に示す  $x_{ij}$  の関数  $w(x_{ij})$  によって定義し、式(2)に示すように重み付き和の形式で  $m'_i(\theta)$  を定義する。なお、式(1)の  $a$  と  $b$  は実験的に決定する定数である。

$$w(x_{ij}) = a \cdot e^{-bx_{ij}} \quad \sum_j w(x_{ij}) = 1 \quad (1)$$

$$\begin{pmatrix} m'_i(\theta) \\ 1 \end{pmatrix} = \sum_{j=1}^L w(x_{ij}) \cdot T_j(\theta) \cdot T_j^{-1}(0) \cdot \begin{pmatrix} m_i \\ 1 \end{pmatrix} \quad (2)$$

ここで、 $T_j(\theta)$  は関節角  $\theta$  が与えられたときのリンク  $j$  座標系から体中心座標系への変換行列を表す。 $m_i$  は、全ての関節角が0度のとき、すなわち図1に示す初期姿勢の状態での体中心座標系における頂点  $i$  の座標である。

### 3. 人体運動の推定

#### 3.1 定式化

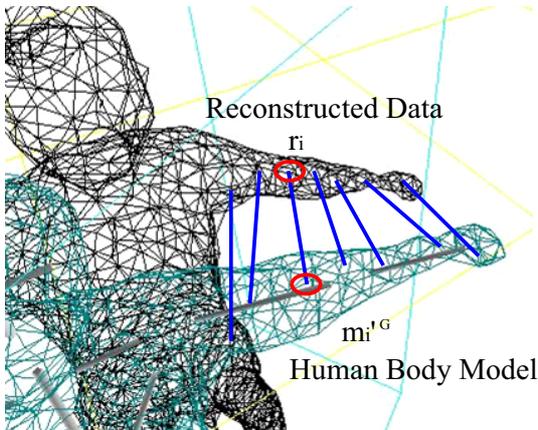


図3 対応点探索に基づく姿勢の推定

多視点からの時系列ビデオ画像を用い、第4.2.1章で述べる方法によって時刻  $t$  における追跡対象の3次元形状を復元する

ことができる。時刻  $t-1$  における運動パラメータ、すなわち関節角  $\theta$ 、位置  $t$  姿勢  $R$  が分かっている場合には、運動推定の問題はこのパラメータを初期値として式(3)を最小化する問題として定式化できる。

$$E = \min \sum_i \| m_i^{tG} - r_i \|^2 \quad (3)$$

ここで、図3に示すように、 $m_i^{tG}$  は  $i$  番目のモデルの頂点を指し、 $r_i$  は復元された3次元形状における  $m_i^{tG}$  の対応点を指す。この定式化は、剛体同士の位置合わせアルゴリズムであるICP[11]を関節構造に沿って滑らかに変形する柔軟変形モデルの位置合わせに拡張したものである。

式(3)の対応点間の誤差がガウス分布に従う場合には、最小二乗法によって式(3)を最小化し運動パラメータを求めることができる。しかし、一般には正しい対応点がない場合や計測誤差の影響があり、これらは最小化を不安定にする要因となる。そこで、本研究ではロバスト推定であるM推定法を導入し、代わりに式(4)を最小化する。

$$E = \min \sum_i \rho(\| R \cdot m'_i(\theta) + t - r_i \|^2) \quad (4)$$

ここで、 $\rho(z)$  は誤差  $z$  の任意の関数である。

式(4)を最小化する  $p = (\theta, t, R)$  は、式(5)を解くことによって求められる。

$$\frac{\delta E}{\delta p} = \sum_i \frac{\delta \rho(z_i)}{\delta z_i} \frac{\delta z_i}{\delta p} \quad (5)$$

ただし、 $z_i = \| R \cdot m'_i(\theta) + t - r_i \|^2$  である。

ここで、式(6)に示すように誤差項を表す重み関数  $e(z)$  を導入する。

$$e(z) = \frac{1}{z} \frac{\delta \rho}{\delta z} \quad (6)$$

すると、式(5)は式(7)のように書き換えることができる。ここで、 $e(z)$  が  $z$  の関数であることを無視すると、これは重み付き最小二乗法の形となる。

$$\frac{\delta E}{\delta p} = \sum_i e(z_i) z_i \frac{\delta z_i}{\delta p} \quad (7)$$

本研究では、外れ値を除外する確率分布としてローレンツ分布を用い、重み関数  $e(z)$  を式(8)のように定義した。

$$e(z) = \left( 1 + \frac{1}{2} \left( \frac{z}{\sigma} \right)^2 \right)^{-1} \quad (8)$$

式(4)を最小にする姿勢パラメータ  $p$  は、式(7)から共役勾配法によって求めることが可能である。

#### 3.2 法線付き Kd-Tree による対応点探索

式(4)の  $r_i$  を求めるためには、復元された3次元データの中から  $m'_i$  の対応点を求める必要がある。このとき、単純に最近傍点を求めると、本手法で扱う多関節物体では特に図4に示すような誤対応が生じやすい。

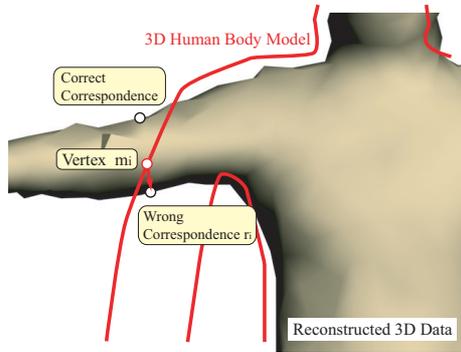


図 4 最近傍点を求めた場合の誤対応

そのため、点の法線の類似性も考慮し、探索時間を削減するために法線付の Kd-Tree によって対応点探索を行う。

Kd-Tree とは、図 5(a) に示すように  $x, y, z$  座標の分散が最大の方向に点群を再帰的に分割してバイナリ木を作成し、探索時間の計算量を  $O(n \log(n))$  に抑えるアルゴリズムである。この場合、図 5(b) の点  $p_1$  の対応点は  $p_2$  になる。

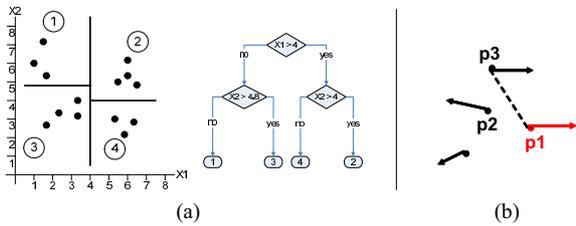


図 5 Kd-tree

これに対し、図 5(b) の点  $p_1$  の対応点が  $p_3$  となるように、点の法線の分散を同時に計算し、法線の重みつき分散と座標の分散を比較して大きい方を分割しバイナリ木を作成する。重みは、法線の偏差の取りうる値域 ( $0 \sim \pi$ ) と座標の偏差の取りうる値域 (物体を囲むバウンディングボックスの大きさ) とがバランスするように決定する。

### 3.3 階層的フィッティング

式 (4) を最小化する際に、運動パラメータ  $\theta, t, R$  を同時に求めると、関節構造の末端のリンクが局所解に陥ってしまうことが多い。そこで、図 6 に示すとおり、体中心の  $(t, R)$  から始めて、関節構造の上位の階層の関節グループ  $\theta_1$  から下位の階層の関節グループ  $\theta_n$  へと順に関節角  $\theta$  のパラメータを求める方法を用いる。

図 7 に、階層的フィッティングによって人体モデルが段階的に収束する様子を示す。

## 4. 実験結果

### 4.1 シミュレーションにおける評価

#### 4.1.1 姿勢推定の評価

まず、位置・姿勢・関節角の真値が分かっている時系列データを作成し、柔軟変形モデルの剛体リンクモデルに体する優位性の評価を行った。なお、剛体リンクの関節構造は柔軟変形モデルと同じものを使用した。

repeat

1. 運動パラメータを保存

$$(t, R) = (t', R')$$

$$(\theta_1, \theta_2, \dots, \theta_N) = (\theta_1', \theta_2', \dots, \theta_N')$$

2. 並進回転パラメータを推定

$$\text{勾配 } \frac{\partial E(t, R, \theta_1, \theta_2, \dots, \theta_N)}{\partial (t, R)}$$

を計算し、 $E(t', R', \theta_1, \theta_2, \dots, \theta_N)$  を最小化する  $(t', R')$  を求める

3. 1 番目の階層の関節角を推定

$$\text{勾配 } \frac{\partial E(t', R', \theta_1, \theta_2, \dots, \theta_N)}{\partial \theta_1}$$

を計算し、 $E(t', R', \theta_1', \theta_2, \dots, \theta_N)$  を最小化する  $\theta_1'$  を求める

4. 2 番目の階層の関節角を推定

$$\text{勾配 } \frac{\partial E(t', R', \theta_1', \theta_2, \dots, \theta_N)}{\partial \theta_2}$$

を計算し、 $E(t', R', \theta_1', \theta_2', \dots, \theta_N)$  を最小化する  $\theta_2'$  を求める

...

5. N 番目の階層の関節角を推定

$$\text{勾配 } \frac{\partial E(t', R', \theta_1', \theta_2', \dots, \theta_N)}{\partial \theta_N}$$

を計算し、 $E(t', R', \theta_1', \theta_2', \dots, \theta_N')$  を最小化する  $\theta_N'$  を求める

until  $|E(t', R', \theta_1', \theta_2', \dots, \theta_N') - E(t, R, \theta_1, \theta_2, \dots, \theta_N)| < \epsilon$

図 6 階層的フィッティングの手順

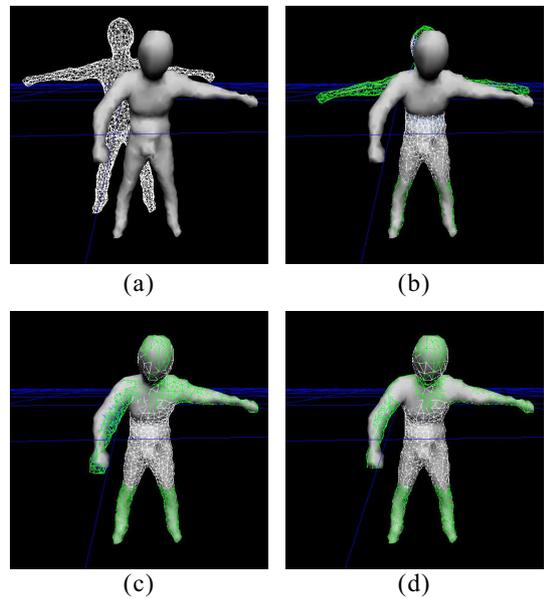
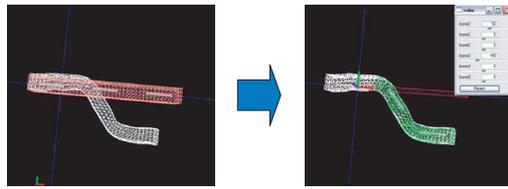


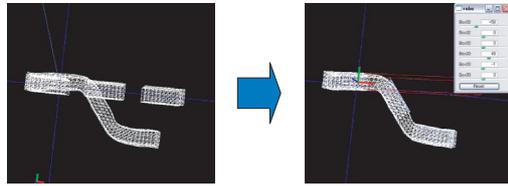
図 7 階層的フィッティング: (a) 初期状態 (b) 1 回目の並進回転パラメータ推定結果 (c) 1 回目の関節角推定結果 (d) 最終の運動パラメータ推定結果

まず、図 8 に示す単純な 2 関節物体における収束結果の比較を行った。形状推定を行う対象は、柔軟変形モデルの関節角を適当に設定して変形を行い、変形後の各頂点にガウスノイズを加えて生成した。図 9 の横軸はガウスノイズの標準偏差を、縦軸は平均誤差を示す。この結果から、剛体リンクモデルでは関節回りの変形に追従できず、柔軟変形モデルに対して 2 倍程度の角度誤差があることが分かる。また、図 9(b) の対応点間の誤差は当然単調増加となるが、一方図 9(a) の関節角度の誤差は発散しないことも確かめられた。

また、図 10 の例では、剛体リンクモデルの場合胴回りの柔軟な変形にモデルが対応できず、結果として誤差が伝播し、関

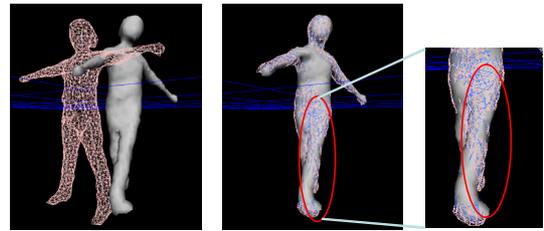


(a) Deformable Model

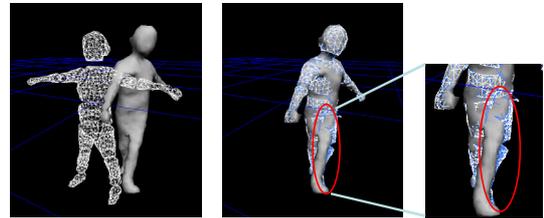


(b) Segment Model

図 8 柔軟変形モデルと剛体モデルの比較

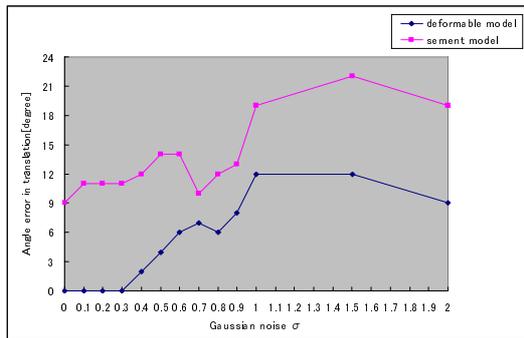


(a) Deformable Model

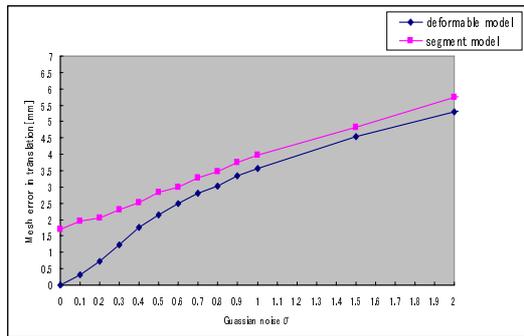


(b) Segment Model

図 10 柔軟変形モデルの定性的評価



(a) 関節角の平均誤差



(b) 対応点間距離の平均誤差

図 9 平均誤差の比較

節構造で下層になる足関節の推定に悪影響が出ていることが分かる。一方柔軟変形モデルでは、この悪影響が小さい。

#### 4.1.2 階層的フィッティングの評価

図 11 に、全運動パラメータを同時に求めた場合と、階層的フィッティングによって順次求めた場合との比較を典型的な例を挙げて示す。図 11(b) にあるように、全運動パラメータを同時に推定すると、局所解に陥り収束が途中で止まってしまうことが分かる。

#### 4.2 実環境における運動推定結果

##### 4.2.1 人体の3次元復元

本研究では、天井にカメラ (Sony DXC-9000) を 8 台設置した環境を用意し、次の手順で複数のビデオ画像から時系列三次元復元データを得る。

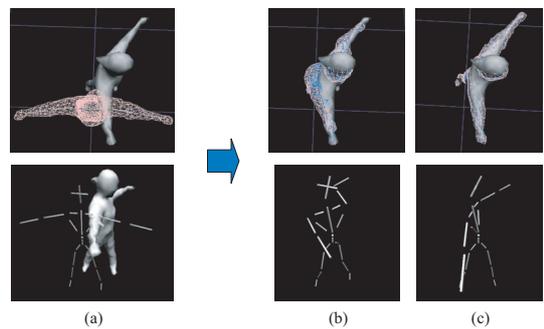


図 11 階層的フィッティングの比較: (a) 初期状態 (b) 全運動パラメータ同時推定結果 (c) 階層的フィッティング結果

- (1) 図 12(a) に示すように、8 台のカメラから毎秒 30 枚の周期で同期のとれた画像を取得し蓄積する
- (2) 背景差分法 [13] を使用し、図 12(b) に示すように輪廓を抽出する
- (3) 視体積交差法 [14] によって三次元復元を行い、Marching Cubes 法 [15] によって図 12(c) に示すメッシュ表現を得る

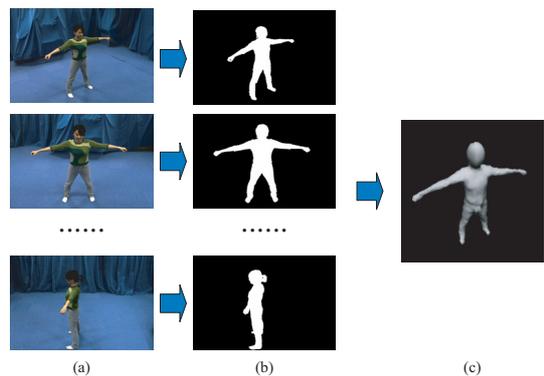


図 12 人体の3次元復元

##### 4.2.2 運動推定結果

図 13 は、全身を使った非常にダイナミックな跳躍動作を追跡した例である。

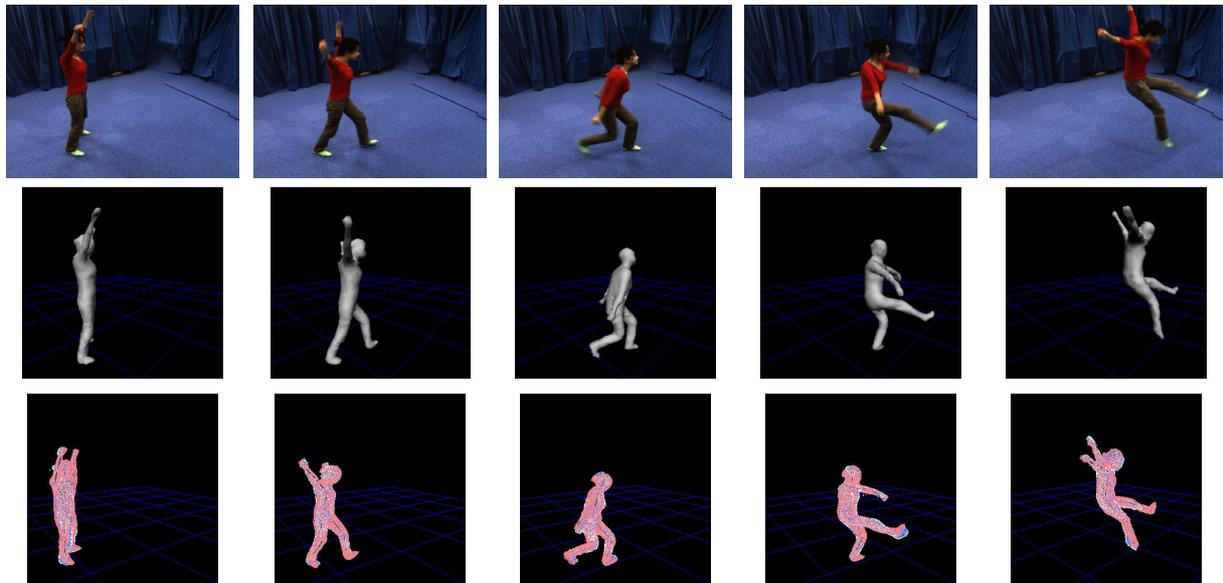


図 13 運動推定例 2 : 跳躍動作  
 上段 : 取得画像, 中段 : 三次元復元結果, 下段 : 人体モデルのフィッティング結果

## 5. ま と め

本論文では, 運動追跡精度の高いマーカレスモーションキャプチャ手法を提案した. 関節構造を持った人体の柔軟変形モデルを用いることで体幹の複雑な運動の追跡を可能にし, また ICP とロバスト推定を組み合わせ法線付 Kd-Tree を対応点探索に用いることにより, 誤対応に強い運動パラメータ推定を可能にした.

今後の課題としては, 初期姿勢の計算や追跡失敗時の対処方法の検討, 人体モデルのスケールを自動調整することによるモデルの汎化, 得られた運動データを利用したコンテンツ生成などが挙げられる.

なお本研究は, 科学技術振興機構戦略的創造研究推進事業 (CREST) デジタルメディア作品の制作を支援する基盤技術の支援を受けて行われた.

## 文 献

- [1] T. B. Moeslund and E. Granum: "A survey of computer vision-based human motion capture", *Computer Vision and Image Understanding: CVIU*, **81**, 3, pp. 231–268 (2001).
- [2] J. M. Rehg and T. Kanade: "Model-based tracking of self-occluding articulated objects", *IEEE International Conference on Computer Vision: ICCV*, pp. 612–617 (1995).
- [3] 亀田, 美濃, 池田: "シルエット画像からの関節物体の姿勢推定法", *電子情報通信学会論文誌 D-II*, **J79-D2**, 1, pp. 26–35 (1996).
- [4] 島田, 白井, 久野, 三浦: "緩やかな制約知識を利用した単眼視動画像からの関節物体の形状と姿勢の同時推定", *電子情報通信学会論文誌 D-II*, **J81-D2**, 1, pp. 45–53 (1998).
- [5] V. A. Romer Rosales and S. Sclaroff: "3d hand pose reconstruction using specialized mappings", *IEEE International Conference on Computer Vision: ICCV*, pp. 378–385 (2001).
- [6] D. Gavrila and L. Davis: "Tracking humans in action: A 3d model-based approach", *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp. 73–80 (1996).
- [7] Q. Delamarre and O. D. Faugeras: "3d articulated models and multi-view tracking with silhouettes", *IEEE International Conference on Computer Vision: ICCV*, pp. 716–721 (1999).
- [8] T. K. G. Cheung, S. Baker: "Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture", *IEEE Conference on Computer Vision and Pattern Recognition: CVPR* (2003).
- [9] S. Ilic and P. Fua: "Generic deformable implicit mesh models for automated reconstruction", *ICCV workshop on Higher-Level Knowledge in 3D Modelling and Motion Analysis*, pp. 29–38 (2003).
- [10] M. B. R. Kehl and L. V. Gool: "Full body tracking from multiple views using stochastic sampling", *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp. 129–136 (2005).
- [11] P. J. Besl and N. D. McKay: "A method for registration of 3-D shapes", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **14**, 2, pp. 239–256 (1992).
- [12] K. Komatsu: "Human skin model capable of natural shape variation", *The Visual Computer*, **4**, 3, pp. 265–271 (1988).
- [13] D. H. Thanarat Horprasert and L. S. Davis: "A statistical approach for real-time robust background subtraction and shadow detection", *ICCV Framerate Workshop*, pp. 1–19 (1999).
- [14] A. Laurentini: "The visual hull concept for silhouette-based image understanding", *PAMI*, **16**, 2, pp. 150–162 (1994).
- [15] C. Montani, R. Scateni and R. Scopigno: "Discretized marching cubes", *Visualization '94 Proceedings IEEE Computer Society*, IEEE Computer Society Press, pp. 281–287 (1994).