

# 論文

## ドリフト修正機能を有する動画像からの身体動作推定法

山本 正信<sup>†a)</sup>

### A Simple and Robust Approach to Drift Reduction in Motion Estimation of Human Body

Masanobu YAMAMOTO<sup>†a)</sup>

あらまし 連続する画像から運動パラメータを推定し、それを追跡開始フレームでの姿勢に累積することによって動作を追跡する手法は、推定誤差の累積によるドリフトが問題となっている。本論文では、簡単で強力なドリフト修正法を提案する。最終フレームでも姿勢を与え、動作追跡により得られた姿勢とのドリフトを求める。ドリフトからフレーム間の平均ドリフトを求める。フレーム順に、得られた身体姿勢から平均ドリフトを差し引くことにより姿勢の修正を行う。動作追跡と姿勢修正を交互に繰り返すことにより、中間フレームでのモデルを身体に一致させる。姿勢の修正は、三元一次連立方程式を解くだけでよく、数回の繰返しでモデルは身体に一致する。いくつかの実映像に対し、身体の動作を測定し提案手法の実効性を確認した。

キーワード モーションキャプチャ、ビデオ映像、ドリフト、位置・姿勢の修正、隠れの補完

#### 1. ま え が き

ビデオ映像からの動作測定法では、身体の構造を表すモデルを利用することが一般的となっている[1]。モデルを画面上で身体像に当てはめたとき、モデルから身体の位置・姿勢を知ることができる[2]~[10]。当てはめのためには、画像から身体像を抽出しておく必要がある。この抽出は必ずしも安定に行えるとは限らない。

これに対し、位置・姿勢の変位分は連続する画像間の差分値を使って容易に推定できる[11]~[13]。したがって、動画像の最初のフレームでモデルを当てはめ、初期位置・姿勢を得た後、これに姿勢の変位を累積すれば身体の位置・姿勢を知ることができる。ところが、変位の推定誤差も累積されるため、累積から得た姿勢は実際の姿勢から時間とともに大きくずれてしまう。このずれをドリフトと呼んでいる。

ドリフトを解消するために、動画像の最後のフレームでもモデルを当てはめ、得られた最終姿勢に対し、累積により得た姿勢が一致するように姿勢変位の修正

を行う手法[14]が提案されている。しかし、この修正法は弛緩法による姿勢伝搬を行っているため収束するまで多くの計算量を必要とする。

Tomasiら[15]は、中間フレームでの位置姿勢をロボットの動作計画を使って補間した。この手法は、姿勢変位の推定を行う必要はないが、初期フレームと最終フレームとの間隔が広がった場合には、補間動作は必ずしも実際の動作を正確に表しているとは限らない。

Rahimiら[16]は、最終フレームでの姿勢が正しい姿勢となるように、すべてのフレームにおいて位置と姿勢を再計算した。ただし、彼らの手法は、姿勢変位の推定に他の物体の運動が含まれていないことを前提としている。実際は、画像中には複数の物体が含まれ異なる運動をしている。モデルが対象からずれると、測定される運動には異なる運動が含まれ、それがますますドリフトを加速させる。したがって、中間フレームで複数の運動の影響を排除する必要がある。

本論文では、簡単で強力なドリフト修正法[21]を提案する。動作追跡により得られた最終フレームでのドリフトからフレーム間の平均ドリフトを求める。フレーム順に、得られた身体姿勢から平均ドリフトを差し引くことにより姿勢の修正を行う。動作追跡と姿勢修正を交互に繰り返すことにより、中間フレームでのモデルを身体に収束させる。姿勢の修正は、三元一次

<sup>†</sup> 新潟大学工学部情報工学科, 新潟市

Department of Information Engineering, Niigata University,  
Niigata-shi, 950-2181 Japan

a) E-mail: yamamoto@ie.niigata-u.ac.jp

の線形連立方程式を解くだけでよく、数回の繰返しでモデルは身体に収束する。

長時間の動作追跡では、位置・姿勢を開始・終了フレームで与えるだけでは不十分で、いくつかの中間フレームでも、位置・姿勢を与えることになる。位置・姿勢が与えられたフレームをキーフレームと呼ぶ。

次章では、準備として身体モデルを導入し、動作追跡法の概要と隠れを伴った運動推定法を述べる。3.では、姿勢及び位置ドリフトの修正方法について述べる。4.では、キーフレームでの位置・姿勢を与えるためのGUIについて述べる。5.では、提案手法により実際の動画像から動作を測定した例を示す。6.では、提案手法と従来法との比較を行う。

## 2. 多関節モデルによる動作追跡

### 2.1 身体が多関節モデル

身体の構造を多関節モデルで表す。図1左に多関節モデルの例が示されている。このモデルは全部で16個の部位から構成され、Waistを木構造の最上位に置き根部位と呼ぶ。図1右は、各部位の接続関係を木構造で表している。矢印の向きは親から子への向きを示す。

各部位は固有の座標系をもっている。この座標系を部位座標系  $\Sigma_i$  とする。ここで、添字  $i$  は部位の番号で、根部位を1とし、親子順に付けられているとする。部位座標系は通常原点を関節位置（腰部では重心位置）に置き、座標軸の一つを部位の体軸に一致させておく。

身体の位置・姿勢・運動は部位座標系の位置・姿勢・運動で表すことができる。身体各部位の姿勢は、その親の座標系を基準に表す。すなわち、部位  $i$  の座標系

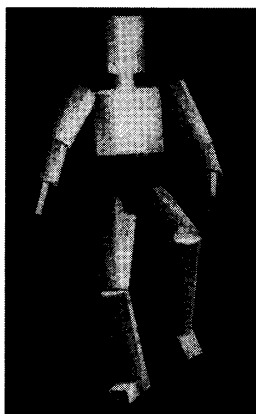


図1 身体が多関節モデル

Fig. 1 An articulated model for human body.

$\Sigma_i$  は、その親である部位  $i-1$  の座標系  $\Sigma_{i-1}$  を  $x, y, z$  軸方向に  $t_x, t_y, t_z$  移動させ、 $z, y, x$  軸の順に  $\theta_{z_i}, \theta_{y_i}, \theta_{x_i}$  回転させて得る。このとき、座標系  $\Sigma_i$  から  $\Sigma_{i-1}$  への変換を同次変換行列

$$T_i = \begin{pmatrix} T_i(\theta) & \mathbf{t}_i \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

で表す。ここで、 $T_i(\theta)$  は回転行列で、 $z, y, x$  軸回りの回転行列を、それぞれ  $Q_z(\theta_z), Q_y(\theta_y), Q_x(\theta_x)$  としたとき、

$$T_i(\theta) = Q_z(\theta_{z_i})Q_y(\theta_{y_i})Q_x(\theta_{x_i})$$

である。また、 $\mathbf{t}_i$  は並進ベクトル

$$\mathbf{t}_i = \begin{pmatrix} t_{x_i} \\ t_{y_i} \\ t_{z_i} \end{pmatrix}$$

である。なお、根部位の親は簡単のためにカメラ座標系  $\Sigma_c$  としておく。

部位の位置と姿勢を表すパラメータを根部位から順に並べたものをベクトル  $\mathbf{q}$  で表す。すなわち、

$$\mathbf{q} = (t_{x_1}, t_{y_1}, t_{z_1}, \theta_{x_1}, \theta_{y_1}, \theta_{z_1}, \dots, t_{x_i}, t_{y_i}, t_{z_i}, \theta_{x_i}, \theta_{y_i}, \theta_{z_i})^\top$$

とし、この  $\mathbf{q}$  を身体の位置・姿勢パラメータと呼ぶ。

カメラ座標系で表された点  $\mathbf{p}$  が、部位  $i$  の座標系  $\Sigma_i$  では  $\mathbf{p}_i$  で表されているとする。この点を同次座標系表示で

$$\tilde{\mathbf{p}}_i = \begin{pmatrix} \mathbf{p}_i \\ 1 \end{pmatrix}$$

と表す。また、 $\mathbf{p}$  の同次座標系表示を  $\tilde{\mathbf{p}}$  とする。このとき部位座標系からカメラ座標系への変換は、

$$\tilde{\mathbf{p}} = T_1 T_2 \cdots T_i \tilde{\mathbf{p}}_i \quad (1)$$

となる。

部位  $i$  が更に自身の座標系を基準に  $x, y, z$  軸方向に  $r_{x_i}, r_{y_i}, r_{z_i}$  移動し、 $z, y, x$  軸の順に  $\phi_{z_i}, \phi_{y_i}, \phi_{x_i}$  回転したとする。この運動を同次変換行列で

$$R_i = \begin{pmatrix} R_i(\phi) & \mathbf{r}_i \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

と表す。ここで、 $R_i(\phi)$  は回転行列のオイラー分解

$$R_i(\phi) = Q_z(\phi_{z_i})Q_y(\phi_{y_i})Q_x(\phi_{x_i})$$

であり、 $\mathbf{r}_i$  は並進ベクトル

$$\mathbf{r}_i = \begin{pmatrix} r_{x_i} \\ r_{y_i} \\ r_{z_i} \end{pmatrix}$$

である。

根部位の自由度は並進と回転の6自由度であるが、それ以外の部位は親とリンク結合しているため、自由度は回転のみの3自由度である。部位の並進と回転を表す変数を運動パラメータと呼び、パラメータを根部位から順に並べたものをベクトル  $\mathbf{x}$  で表す。

$$\mathbf{x} = (r_{x_1}, r_{y_1}, r_{z_1}, \phi_{x_1}, \phi_{y_1}, \phi_{z_1}, \phi_{x_2}, \phi_{y_2}, \phi_{z_2}, \dots, \phi_{x_i}, \phi_{y_i}, \phi_{z_i})^T$$

位置・姿勢パラメータ  $\mathbf{q}$  とはパラメータの数が異なることに注意。

モデルが運動したとき、位置  $\tilde{\mathbf{p}}_i$  はカメラ座標系では  $\tilde{\mathbf{p}}'$  に移動したとする。この変換は、

$$\tilde{\mathbf{p}}' = T_1 R_1 T_2 R_2 \cdots T_i R_i \tilde{\mathbf{p}}_i \quad (2)$$

で与えられる。 $\tilde{\mathbf{p}}'$  は運動パラメータ  $\mathbf{x}$  の関数となる。この関数は一般に非線形関数であるが、運動パラメータが微小なときには線形化することができる。すなわち、

$$\tilde{\mathbf{p}}' - \tilde{\mathbf{p}} = \mathbf{J}\mathbf{x} \quad (3)$$

と表せる。ここで、 $\mathbf{J} = \frac{\partial \tilde{\mathbf{p}}'}{\partial \mathbf{x}}$  はヤコビ行列と呼ばれ、その計算法は付録1.に示されている。

## 2.2 ドリフト修正機能を有する動作追跡法の概要

追跡開始フレーム（第0フレーム）と追跡終了フレーム（第  $n$  フレーム）で、モデルを身体に当てはめ身体的位置と姿勢が得られているとする。このとき中間フレームでの位置と姿勢の決定問題は、ロボットの軌道計画問題として解くことができる。市販のアニメーション作成ツールのいくつかは、キーフレームを与えたときキーフレーム間の姿勢を補間する機能も持っている。

本論文では、最も簡単な方法として線形補間を用いる。姿勢の線形補間は、開始姿勢と終了姿勢の差を軸回りの回転で表し、回転角度をフレーム数で案分し、中間姿勢を得る。位置の線形補間は、最上位の部位のみで行う。開始位置と終了位置の差をフレーム数で案

分して中間位置を得る。

得られた位置・姿勢列はパラメータ空間内に軌道を描く。図2左に簡単のためにパラメータ空間を二次元として、生成された軌道が描かれている。この軌道を初期軌道とする。同時に実際の身体の動作軌道も描かれている。初期軌道は必ずしも動作軌道に一致しない。両者が一致するように初期軌道を修正する。

初期軌道から各フレームでのモデルの位置・姿勢が与えられたとき、次節で述べる手法で運動パラメータを推定する。得られた運動パラメータを追跡開始フレームでの位置・姿勢に累積した結果を図2右に示す。累積軌道上で動作軌道からのドリフトが発生している。このドリフトは、モデルが身体とずれているために、異なる動きが混入していることが主な原因である。

ドリフトの修正方法は3.で示す。図2右に修正された軌道が示されているが、一度の修正で動作軌道に一致するとは限らない。修正軌道の評価は次のように行う。

モデルが身体に一致しているとすれば、身体像上の点  $(X, Y)$  の投影元  $\mathbf{p}_i$  を身体モデルから求めることができる。次のフレームでの点  $\mathbf{p}_i$  の移動先を式(2)から計算し、その投影先  $(X', Y')$  を求める。もし、軌道が正しければ、濃淡値  $E(X, Y, j)$  と  $E(X', Y', j+1)$  の誤差は0に近づくはずである。そこで、身体像全体及び全フレームでの誤差の総和を次式で求める。

$$\sum_{j=0}^n \sum_{X, Y} (E(X', Y', j+1) - E(X, Y, j))^2 \quad (4)$$

軌道の修正を行うたびに誤差の総和を求め、最も誤差の総和が小さな軌道を動作軌道とする。

ところで、正しい動作軌道上で計算した運動パラ

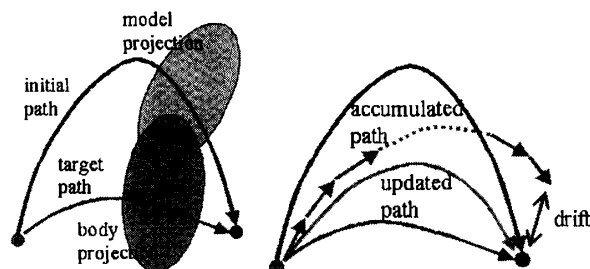


図2 初期軌道、動作軌道、累積軌道、ドリフト及び軌道の修正

Fig. 2 Pose trajectories in configuration space. Black curve denotes the initial trajectory, light gray one accumulation on the initial poses, and dark gray one final trajectory.

メータを累積しても、画像には様々なノイズが含まれているためドリフトは起こる [16]。この場合でも、ドリフトの修正により正しい軌道に戻す。

### 2.3 隠れを考慮した身体動作の推定

モデルが身体に一致しているとき、身体部位の動作パラメータ  $\mathbf{x}$  は線形連立方程式

$$A\mathbf{x} = \mathbf{b} \quad (5)$$

を解くことによって得られる。この連立方程式の係数行列及び定数ベクトルの導出は付録 2. を参照。

ただし、方程式 (5) の根拠である移動ベクトルの束縛式 (A.7) は濃淡分布の一次近似に基づいているため、動作パラメータを正しく得るためには式 (5) の反復適用が必要である [17], [18]。

身体の部位が一部隠された場合には、その動作パラメータを決定することができない。この状態は線形推定式 (5) が一次従属であることを意味している。したがって、問題は線形推定式の一次従属性の判定と一次従属と判定されたときの解の確定法である。

一次従属性を判定するために、線形連立方程式 (5) の係数行列  $A$  を特異値分解する。

$$A = U\Sigma V^T \quad (6)$$

ここで、 $A$  の次元を  $m \times n (m \gg n)$  とする。 $U$ 、 $V$  はそれぞれ  $m \times m$ 、 $n \times n$  の直交行列である。 $A$  の特異値を大きい順に、 $\sigma_1, \sigma_2, \dots, \sigma_n$  とする。 $\Sigma$  は特異値を大きさの順に対角成分に並べた  $m \times n$  の対角行列である。

最大特異値と各特異値との比  $\sigma_i/\sigma_1$  があるしきい値以下になったとき、その特異値を 0 とみなす。すなわち、このしきい値以下の特異値  $\sigma_{k+1}, \sigma_{k+2}, \dots, \sigma_n$  としたとき、それらをすべて 0 とし、改めて対角行列  $\Sigma$  を定義する。ここで、 $\text{rank}(A) = k$  である。このとき、式 (5) の残差を最小とする解は

$$A^+\mathbf{b} + W\mathbf{z} \quad (7)$$

として与えられる。ここで、 $A^+$  はムーアペンローズ一般化逆行列で、 $A^+ = V\Sigma^{-1}U^T$  と与えられる。ただし、 $\Sigma^{-1}$  は特異値の逆数  $1/\sigma_i, (i = 1, \dots, k)$  を順に対角成分に並べた  $n \times m$  の対角行列である。また、 $W$  は行列  $V$  の第  $k+1$  番目以降の列ベクトルを順に並べた行列である。 $\mathbf{z}$  は  $n-k$  次元の未定ベクトルである。

未定ベクトル  $\mathbf{z}$  は、隠れにより発生する未知数であ

る。もし、動作パラメータの予測値が与えられているならば、隠された部位の動きがその予測値となるように未定ベクトル  $\mathbf{z}$  の値を決定する。

予測値は修正前の軌道から計算されるフレーム間動作パラメータをそのままあてる。動作パラメータの予測値を  $\mathbf{y}$  とする。解が  $\mathbf{y}$  に近づくように未定ベクトルを決定する。それには、連立方程式

$$\mathbf{y} = A^+\mathbf{b} + W\mathbf{z} \quad (8)$$

を  $\mathbf{z}$  について解けばよい。この式の総数は未知数の数より多いので、方程式の左辺と右辺の差が最小となる解を求める。最小二乗解は方程式の正規表現から得られる。正規表現は  $W^T$  を左から掛けて得られる。 $W$  の列ベクトルは互いに直交するので、結局

$$\mathbf{z} = W^T(\mathbf{y} - A^+\mathbf{b}) \quad (9)$$

となる。これを式 (7) に代入すれば、最終解

$$\mathbf{x} = A^+\mathbf{b} + WW^T(\mathbf{y} - A^+\mathbf{b}) \quad (10)$$

が得られる。

## 3. ドリフトの修正

### 3.1 姿勢の修正

最初に姿勢を修正する。身体部位の姿勢及び運動を直交行列で表す。部位  $i$  の親部位に対する  $j$  フレーム目の姿勢は  $T_{i,j}(\theta)$  と表すが、 $T_j$  と略す。この部位の初期姿勢を  $T_0$ 、最終姿勢を  $T_n$  とする。部位  $i$  の  $j-1$  フレームから  $j$  フレームまでの回転運動を  $R_{i,j}(\phi)$  と表す。これを  $R_j$  と略す。初期フレームから最終フレームまでの部位の回転運動を  $R_1, R_2, \dots, R_n$  とする。これらはすべてモデル照合や動作推定により得られているものとする。

最終姿勢は、初期姿勢に回転運動を累積することによって、

$$T_0 R_1 R_2 \cdots R_n$$

と計算することができる。回転運動に誤差が含まれると、計算された姿勢は次第に身体から離れドリフトが発生する。

計算により得られた姿勢が与えられた最終姿勢  $T_n$  と等しくなるためには、補正回転運動  $X_i$  を

$$T_0 R_1 X_1 R_2 X_2 \cdots R_n X_n = T_n \quad (11)$$

となるように与えればよい。

$X_i$  は各フレームで発生したドリフトの補正である。ドリフトはフレームごとに異なるが、時間とともに全ドリフトが大きくなるのは、フレームごとのドリフトが平均的に偏っていると考えてよい。この平均ドリフトの補正を  $X$  とする。すなわち、

$$X_i = X \quad (12)$$

とする。更に、補正量は微小であるとして、姿勢補正行列を

$$X = \begin{pmatrix} 1 & -a & b \\ a & 1 & -c \\ -b & c & 1 \end{pmatrix} \quad (13)$$

と表す。式 (11) に  $T_0$  の逆行列を左から掛けて、 $T_0^{-1}T_n = T$  とし、

$$R_1XR_2X \cdots R_nX = T \quad (14)$$

を得る。方程式 (14) を  $X$  について解けば補正行列が得られる。この方程式は  $X$  に関する非線形方程式であるので、ニュートン法による逐次近似により解を求めることにする。

この方程式の左辺を展開し両辺を比較すれば、左辺の要素は  $X$  の要素の  $n$  次多項式となる。 $n$  次方程式を代数的に解くのは難しい。

そこで、左辺を  $X$  の要素の一次式で近似し、一次方程式を解く。 $X$  は

$$X = I + Y \quad (15)$$

と分解できる。ただし、 $I$  は大きさ 3 の単位行列、

$$Y = \begin{pmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{pmatrix}$$

とすれば、式 (14) の左辺は

$$\begin{aligned} & R_1(I+Y)R_2(I+Y) \cdots R_n(I+Y) \\ &= R_1R_2 \cdots R_n + R_1YR_2 \cdots R_n + R_1R_2YR_3 \cdots R_n \\ & \quad + \cdots + R_1R_2 \cdots R_nY + O(2) \end{aligned}$$

となる。ただし、行列  $O(2)$  の要素は、 $a, b, c$  に関する二次以上の積項からなる多項式となる。 $a, b, c$  を微小量とすれば、二次以上の積項は無視することができ、行列  $O(2)$  も 0 行列とみなすことができる。

ここで、

$$F_j = R_1 \cdots R_j$$

$$G_i = R_j \cdots R_n$$

ただし、 $G_{n+1} = I$  とすれば、式 (14) の左辺は

$$F_n + F_1YG_2 + F_2YG_3 + \cdots + F_nYG_{n+1} \quad (16)$$

と近似される。これが式 (14) の右辺と一致するとすれば、 $Y$  に関する一次方程式が得られる<sup>(注1)</sup>。

$$F_1YG_2 + F_2YG_3 + \cdots + F_nYG_{n+1} = S \quad (17)$$

ただし、 $S = T - F_n$  である。

さて、式 (17) の一般項を  $H_j$  とすれば、 $H_j = F_jYG_{j+1}$  であるが、添字  $j$  を省略して

$$H = FYG$$

としたとき、この一般項の行列の要素を計算する。行列  $H, F, G$  を

$$H = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix}$$

$$F = \begin{pmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{pmatrix}$$

$$G = \begin{pmatrix} g_1 & g_2 & g_3 \\ g_4 & g_5 & g_6 \\ g_7 & g_8 & g_9 \end{pmatrix}$$

としたとき、

$$\begin{aligned} h_1 &= (f_2g_1 - f_1g_4)a + (f_1g_7 - f_3g_1)b \\ & \quad + (f_3g_4 - f_2g_7)c \end{aligned}$$

$$\begin{aligned} h_2 &= (f_2g_2 - f_1g_5)a + (f_1g_8 - f_3g_2)b \\ & \quad + (f_3g_5 - f_2g_8)c \end{aligned}$$

$$\begin{aligned} h_3 &= (f_2g_3 - f_1g_6)a + (f_1g_9 - f_3g_3)b \\ & \quad + (f_3g_6 - f_2g_9)c \end{aligned}$$

(注1) :  $F(\delta\mathbf{x}) = R_1XR_2X \cdots R_nX$ , ただし、 $\delta\mathbf{x} = (a, b, c)^T$  とおき、 $F(\delta\mathbf{x})$  を  $\delta\mathbf{x} = \mathbf{0}$  でテイラー展開し、二次以上の微小項を無視すれば、 $F(\delta\mathbf{x}) \approx F(\mathbf{0}) + \left( \frac{\partial F(\delta\mathbf{x})}{\partial \delta\mathbf{x}} \Big|_{\delta\mathbf{x}=\mathbf{0}} \right)^T \delta\mathbf{x} = R_1R_2 \cdots R_n + R_1YR_2 \cdots R_n + R_1R_2YR_3 \cdots R_n + \cdots + R_1R_2 \cdots R_nY$  を得る。この式が  $T$  に等しいとしても、方程式 (17) が得られる。

$$\begin{aligned}
h_4 &= (f_{5g1} - f_{4g4})a + (f_{4g7} - f_{6g1})b \\
&\quad + (f_{6g4} - f_{5g7})c \\
h_5 &= (f_{5g2} - f_{4g5})a + (f_{4g8} - f_{6g2})b \\
&\quad + (f_{6g5} - f_{5g8})c \\
h_6 &= (f_{5g3} - f_{4g6})a + (f_{4g9} - f_{6g3})b \\
&\quad + (f_{6g6} - f_{5g9})c \\
h_7 &= (f_{8g1} - f_{7g4})a + (f_{7g7} - f_{9g1})b \\
&\quad + (f_{9g4} - f_{8g7})c \\
h_8 &= (f_{8g2} - f_{7g5})a + (f_{7g8} - f_{9g2})b \\
&\quad + (f_{9g5} - f_{8g8})c \\
h_9 &= (f_{8g3} - f_{7g6})a + (f_{7g9} - f_{9g3})b \\
&\quad + (f_{9g6} - f_{8g9})c
\end{aligned}$$

を得る。すなわち、行列  $H$  の要素は、定数項を含まない  $a, b, c$  の一次式である。これを  $j$  フレーム目の回転運動に対応する一般項であることを強調して、

$$h_{k,j} = p_{k,j}a + q_{k,j}b + r_{k,j}c \quad (18)$$

と表す。ただし、 $(k = 1, \dots, 9, j = 1, \dots, n)$  とする。式 (17) の各要素はすべての時系列についての総和であり、これが右辺  $S$  の各要素に等しいとすれば、九つの方程式が得られる。すなわち、

$$\alpha_k a + \beta_k b + \gamma_k c = s_i \quad (k = 1, \dots, 9) \quad (19)$$

ただし、

$$\alpha_k = \sum_{j=1}^n p_{k,j} \quad (20)$$

$$\beta_k = \sum_{j=1}^n q_{k,j} \quad (21)$$

$$\gamma_k = \sum_{j=1}^n r_{k,j} \quad (22)$$

$$S = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_4 & s_5 & s_6 \\ s_7 & s_8 & s_9 \end{pmatrix}$$

とする。

式 (19) は  $a, b, c$  に関する線形連立方程式であり、行列の形で書き表せば、

$$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \vdots & \vdots & \vdots \\ \alpha_9 & \beta_9 & \gamma_9 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_9 \end{pmatrix} \quad (23)$$

である。方程式の数が未知数の数を上回るので、最小二乗法により解を得る。

このようにして得られた解は、近似方程式 (17) の解であるので、もとの方程式 (14) を満たすとは限らない。そこで、運動  $R_j$  に補正  $X$  を掛けた行列を新たな運動  $R_j$  として、再度、補正行列を計算する。方程式 (14) の左辺と右辺の差のノルムが小さな値  $\epsilon (> 0)$  より小さくなるまで、この計算を繰り返す。実験によれば、 $\epsilon = 10^{-10}$  のとき、繰返しの回数は 5 回程度であった。

### 3.2 位置の修正

位置の修正は最上位の部位（腰部）についてのみ行う。根部位 1 の座標系  $\Sigma_1$  について、0 フレームと  $n$  フレーム目の原点位置をカメラ座標系で表し、それぞれ  $\mathbf{t}_{1,0}$ ,  $\mathbf{t}_{1,n}$  とする。これらは与えられているものとする。一方、 $j-1$  フレームから  $j$  フレーム間の並進移動量を  $\mathbf{r}_{1,j}$  と表す。初期フレームから最終フレームまでの並進移動量  $\mathbf{r}_{1,1}, \mathbf{r}_{1,2}, \dots, \mathbf{r}_{1,n}$  は運動パラメータの推定により得られているものとする。

原点の最終位置は、その初期位置に並進移動量を累積することによって得られる。ただし、計算された最終位置は正しい最終位置  $\mathbf{t}_{1,n}$  とはずれているので、 $\mathbf{t}'_{1,n}$  とする。

$$\mathbf{t}'_{1,n} = \mathbf{t}_{1,0} + T_{1,0} \mathbf{r}_{1,1} + \dots + T_{1,0} \prod_{j=1}^{n-1} R_{1,j} \mathbf{r}_{1,n} \quad (24)$$

ここで、 $R_{1,j}$  は修正前の回転行列である。また、 $T_{1,0}$  は 0 フレームにおける根部位座標系からカメラ座標系への変換である。

この累積位置が、与えられた位置  $\mathbf{t}_{1,n}$  に等しくなるためには、差  $\mathbf{t}_{1,n} - \mathbf{t}'_{1,n}$  を  $n$  等分し、得た値  $\delta \mathbf{t}$  を使って、原点位置を第 1 フレームでは  $\delta \mathbf{t}$  移動させ、第 2 フレームでは  $2\delta \mathbf{t}$  移動させ、 $\dots$ 、第  $n$  フレームでは  $n\delta \mathbf{t}$  移動させればよい。

### 4. GUI による姿勢設定

キーフレームでのモデルの当てはめは自動的に行えることが望ましい。しかし、画像から身体像の抽出、身体部位の隠れ、正確でない身体モデル、キーフレー

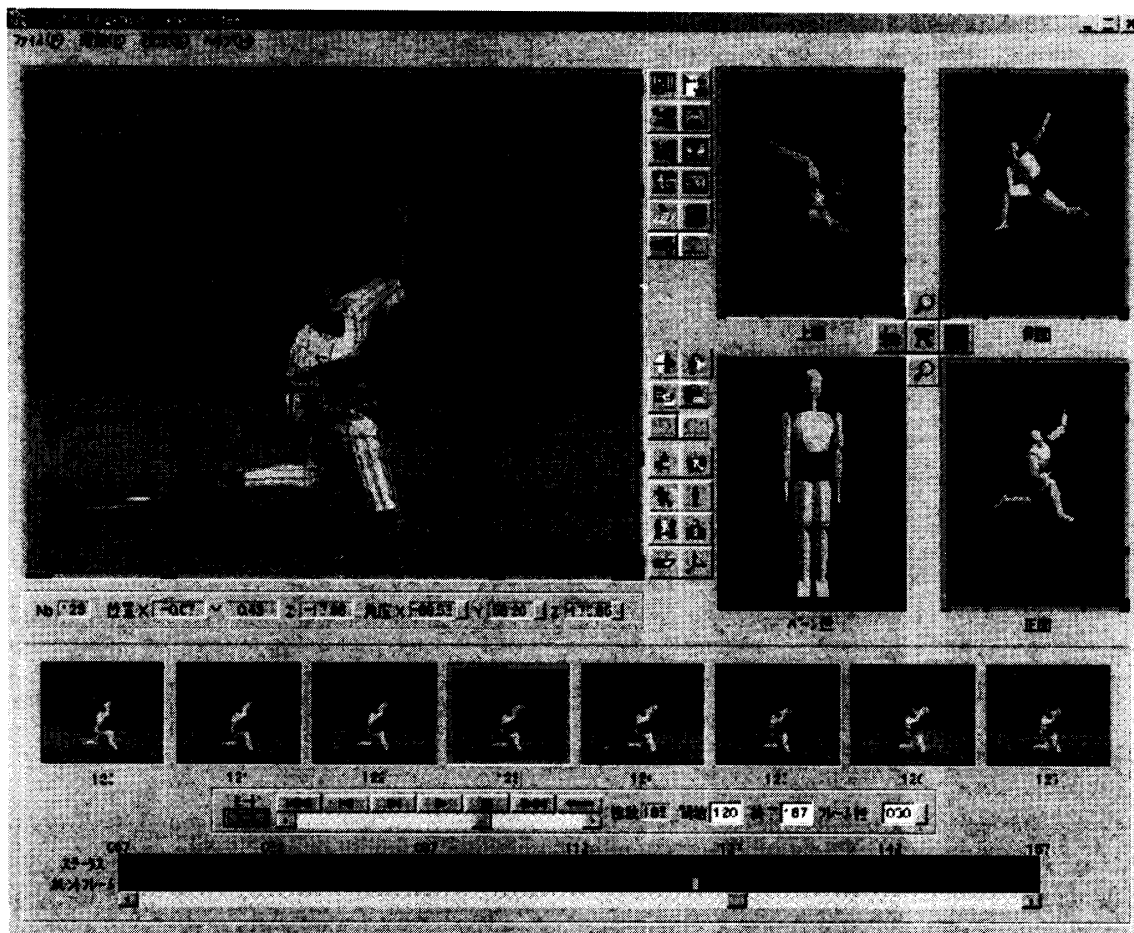


図 3 位置・姿勢を設定するための GUI

Fig. 3 A graphic user interface to set a body pose.

ムを選択など様々な問題があり、現在のところ自動化は難しい。そこで、現実的な方法として、手動でキーフレームを選択しモデルを修正しながら身体に当てはめることのできる簡単な GUI を使用した。

提案システムの GUI を図 3 に示す。左側の最も大きな画面をメイン画面と呼び、追跡画像が示されている。身体モデルは、ワイヤフレーム表示あるいはソリッド表示で、このメイン画面に重ねて描かれる。右に小さな四つの画面がありモデルのソリッド表示が示されている。このうち、上面図、側面図、正面図には、互いに直交するカメラから見たモデルの投影像が示されている。残りのパーツ選択画面は、位置・姿勢を設定あるいはサイズを変更すべき部位を指定するために使用する。

各部位の位置・姿勢の設定は、すべてマウス操作により行う。モデルと身体像の一致は、メイン画像上でワイヤフレームモデルの輪郭が身体像の輪郭にほぼ重なることにより評価する。画面に対する手足の向きや、

隠された部位の姿勢は、3 面図を見ながら調整する。身体全体のモデル照合は、1 キーフレーム当たり 10~30 秒くらいで実行することができる。

図 1 左に示す身体モデルでは、腕、足、首は円筒、腰部、胸部、頭部は錐台、手や足先は直方体で近似している。モデルは、身体データベース [19] を参考に作成した。成人男女、老年男女の 4 種類を用意しているが、身体の大きさには個人差があるため、対象ごとに部位のサイズと部位間の接続位置を変更する必要がある。これらのモデルパラメータは、モデルが身体像に一致するように変更する。

モデルの変更と姿勢の設定は、画面上では互いに関連しているため、どちらを優先させるかは難しい問題である。例えば、腕のモデルを照合させるとき、モデルの長さや姿勢を同時に設定することは難しい。しかし、もし腕の体軸が画面に平行になっているならば、モデルの長さは画面から比較的正確に得られる。このように、モデルのサイズは、対象部位のサイズが正確

表 1 身体部位の運動パラメータの自由度

Table 1 Setting d.o.f for parts.

Parts name	d.o.f	Parts name	d.o.f
Waist	6	Chest	3
Left thigh	2	Right thigh	2
Left shin	1	Right shin	1
Left upper arm	2	Right upper arm	2
Left forearm	1	Right forearm	1
Head	2	Total d.o.f	23

に画面に現れやすいキーフレームを選んで変更する。

身体モデルの運動パラメータは、根部位が6自由度、その他の部位は3自由度であるが、小さな部位や表面に十分な模様のない部位の運動を測定するのは難しい。そのため、身体の動きの自由度に制限を加えている。表1に本論文で設定した各部位の自由度を示す。手や足先、首は部位が小さいため自由度を0とした。なお、位置・姿勢を設定するときの自由度は、根部位が6自由度、その他の部位は3自由度である。

## 5. 追跡実験結果

提案した動作推定法の性能を実際の動画像を使って評価した。まず、位置と姿勢の修正機能を剛体と多関節物体の追跡により評価する。次に、隠れや複雑な動作を含む運動の追跡を評価する。また、これらの追跡にかかる計算コストを示す。

### 5.1 剛体

直方体の箱を対象に40フレームの運動を追跡した。第0と40フレームをキーフレームとしてモデル照合を行った。図4の行は、上から順に初期姿勢列、2回目、3回目及び5回目の修正結果による姿勢列画像である。列は、左から順に0, 8, 16, 25, 40フレームにおける画像にモデルを重ねて示している。動作軌道の修正を5回繰り返して最小誤差となる姿勢が得られた。剛体6自由度の修正が可能であることを示している。なお、この例では最終フレームが40フレームを超えると、追跡のための繰返し演算は発散した。モデルが物体像に重なる割合が大きければ、追跡軌道は物体の真の軌道に収束する。しかし重なる割合が小さければ、運動パラメータは物体の動きを表さなくなるため、動作追跡の繰返しのたびにモデルは物体像からますます離れ、ついには発散する。

### 5.2 身体

身体を対象に腰、胸、頭部、右腕、右上脚から構成されるモデルを使って、演説をしている人物の180フレームの動作を追跡した。キーフレームは、第0フ

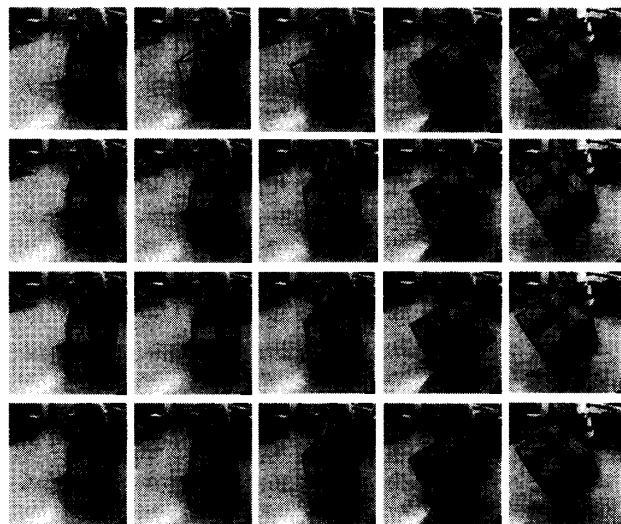


図4 箱の追跡. 各行は繰返しごとの結果. 各列はフレームに対応.

Fig. 4 Box tracking. The 1st, 2nd, 3rd and 4th rows denote poses at initial, 1st, 2nd and 5th (final) iterations, respectively. The 1st, 2nd, 3rd, 4th and 5th columns denote poses at 0, 8th, 16th, 25th and 40th frames, respectively.

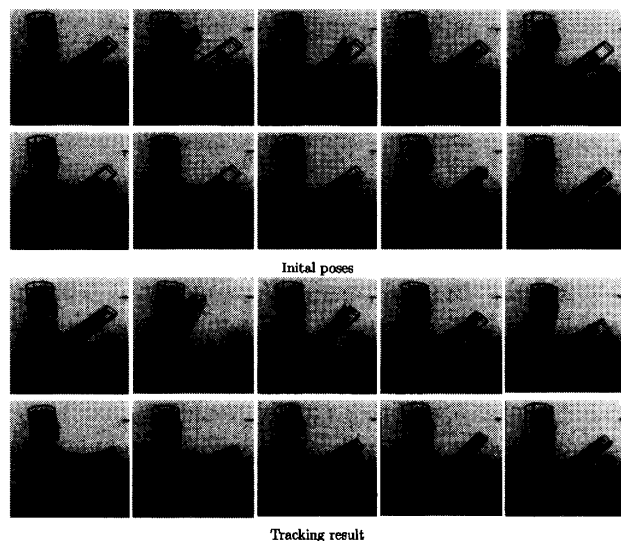


図5 演説人物の動作追跡結果(下)と初期軌道における姿勢列(上)

Fig. 5 Human body tracking.

レームと第180フレームのみである。図5の上行に初期姿勢列、下行に最終姿勢列を、左上から右下にかけて0から180フレームまで20フレームおきに表示している。また、図6はモデルの初期姿勢が身体像と最も離れているフレームでの修正過程を示している。各行は上から18, 87, 156番目のフレームを、各列は左から0, 1, 2, 3, 5回目の修正結果を示している。5回の繰返しで誤差が最小となる姿勢が得られた。最初



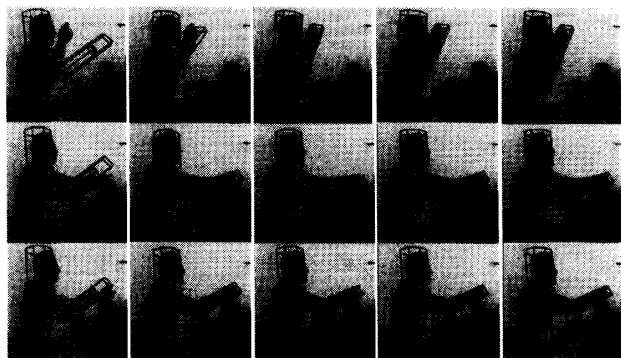


図6 姿勢の修正過程. 各行は上から18, 87, 156フレームを示し, 各列は左から初期姿勢, 1, 2, 3回目の修正及び最終姿勢.

Fig.6 Correction process. The rows correspond to 18th, 87th and 156th frames, and the columns 0, 1, 2, 3 and 5 iterations, respectively.

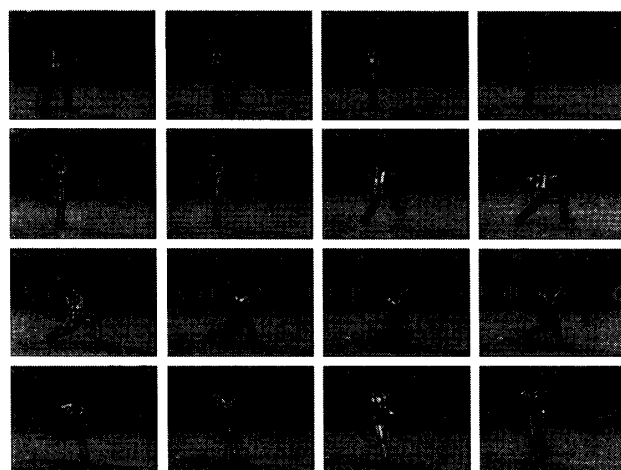


図7 投球動作の追跡結果

Fig.7 Tracking result of pitching in the baseball game.

の1, 2回の繰返しでほぼ最終姿勢が得られている.

### 5.3 様々な動作の追跡

動作が複雑になれば, 追跡開始と終了フレームのモデル照合だけでは十分ではない. 図7は野球の投球動作を追跡した結果である. 全188フレームを20フレームおきに表示している. また, 図8は, 投球動作をバッターの視点から復元したCG映像である.

全フレーム数は, 演説(180フレーム)とはほぼ同じであるが, 16枚のキーフレームを使用した. その理由は主に二つある. まず, 複雑な動作では線形補間から得られた初期姿勢と実際の姿勢との隔たりが大きくなる. 例として, 101から127フレームまでの区間でその両端をキーフレームとして追跡した結果を示す. 図9左は, 107フレームでの初期姿勢, 同図右は姿勢

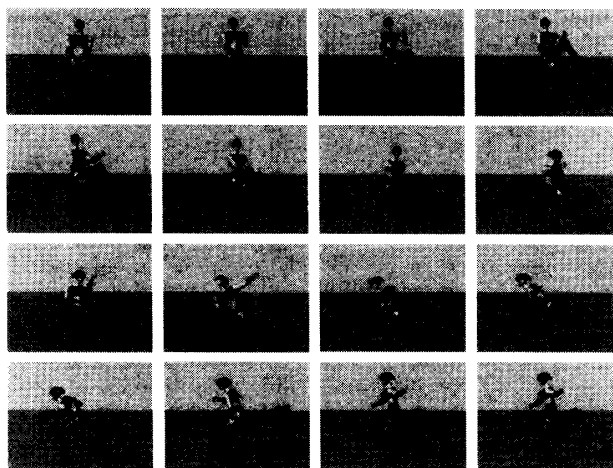


図8 投球動作のCG映像. 追跡結果からの再構成.

Fig.8 Reconstruction of the pitching from tracking result.



図9 初期姿勢(左), 修正姿勢(中), キーフレーム追加後の修正姿勢(右)

Fig.9 Initial pose (left), final pose (middle) and final pose after adding key-frames (right).

の修正結果である. 足部の追跡は良好であるが, 複雑な動きをする腕の姿勢修正は難しい. もう一つは, 動きが速くなると画質が低下し動作推定の前提条件が崩れることである. 図3に示す123フレーム目の画像では投手の左腕がぶれている. 以上の理由から, この区間には112と120フレームをキーフレームに加えた. 図9右に, キーフレームを加えた後での姿勢の修正結果(107フレーム)を示す.

単眼視である限り部位の隠れは起こる. 本論文では, 部位の隠れは動作推定式(5)の一次従属性により判定され, 隠された部位の姿勢は初期姿勢が当てはめられる.

図10は, 投球動作の追跡について, 動作推定式(5)の係数行列の階数と部位の隠れの関係を示している. 上図は横軸がフレーム数, 縦軸が階数である. 階数は, 最大特異値と各特異値の比があるしきい値以上である特異値の個数とした. 運動パラメータは23個(表1)であるが, カメラモデルを望遠と仮定しているため, カメラ光軸方向の運動は測定できない. したがって, 最大階数は22である. 見えている部位も表面模様が

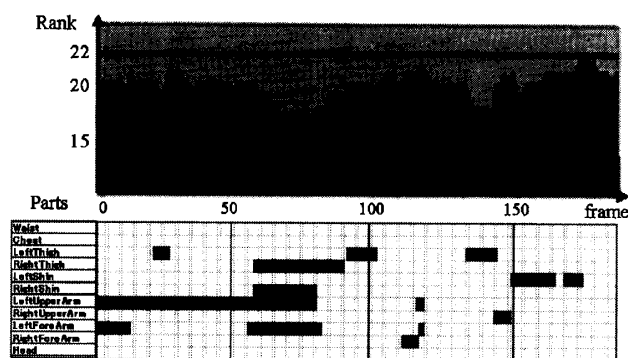


図 10 動作推定式の階数と部位の隠れ

Fig. 10 Rank of the system of motion equations vs. occlusion.

表 2 追跡実験結果

Table 2 Statistics in various tracking results.

追跡動作	総フレーム数	キーフレーム数	平均区間	標準偏差
茶道点前	351	13	28.4	1.2
野球・バッティング	151	9	18.8	11.3
気功	196	11	18.5	12.1
ゴルフ・バッティング	71	6	14.0	8.1
野球・ピッチング	188	16	12.5	5.0
歩行	240	21	12.0	7.0
ノルディックスキー	60	7	9.8	3.1
フィギュアスケート	355	40	9.1	4.3
体操・床運動	91	13	7.5	1.2
マラソン	60	10	6.6	2.1

顕著でなかったり、見かけが小さい場合には隠れとみなされ、最大階数には達していない。下図は、各部位ごとに隠れが起きたフレームを黒で塗りつぶしている。上下の図を比較すると隠れがほぼ正確にとらえられていることが分かる。

様々な動作について、追跡を行った結果を表 2 に示す。この表で、平均区間とはキーフレームに挟まれた追跡区間の平均フレーム長であり、その標準偏差をその右欄に示している。動作は上から平均区間の長さ順に並べられている。茶道の点前のように比較的ゆっくりとした日常動作では、ほぼ 30 フレームおきにキーフレームを挿入すればよい。しかし、スポーツのように速くて複雑な動作には頻りにキーフレームを必要とする。

### 5.4 修正計算のコスト

本手法は、位置と姿勢を修正するための繰返し計算にコストを必要とする。式 (4) で計算される誤差の総和が、最も小さくなる時の位置・姿勢列を求めている。図 11 は、ある追跡実験で得られた、繰返し回数と誤差の総和である。最小誤差解への収束は単調では

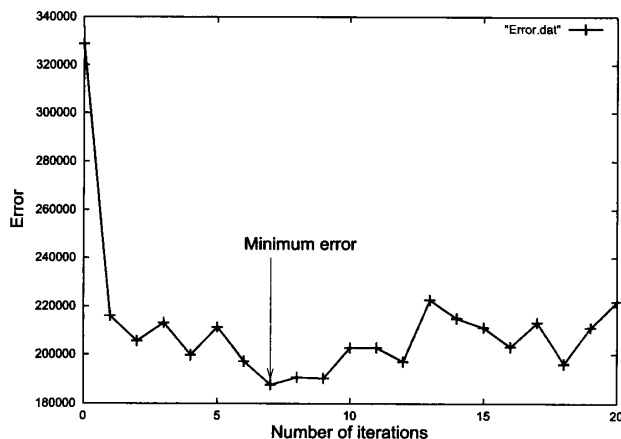


図 11 追跡の繰返し回数と誤差の総和

Fig. 11 Iterative reduction of residual error for model fitting.

なく、繰返し回数を増やしても最適解が得られるとは限らない。また、初期位置や姿勢が、正しい位置・姿勢と大きくずれている場合は、解が発散することもある。計算速度は、Pentium4, 2GHz のプロセッサで、画像サイズをハーフサイズ、繰返し回数 10 回、キーフレーム間の区間を 10 フレームとしたとき、約 10 秒であった。

## 6. 考 察

本手法と既に提案されているドリフト修正法との違いを考察する。

Tomasi ら [15] は、キーフレーム間の姿勢をロボットの動作計画によって補間している。キーフレームの間隔が短い場合には有効であるが、長くなると人間特有の動作が現れ、ロボットの動作計画だけでは生成できない動作を含む。実際、線形補間による動作計画から生成された図 5 の初期姿勢列には、モデルが身体からかなりずれたフレームもある。他の動作計画法を使用したところで、身振り手振りを交えながら演説する動作の生成は困難であると思われる。本手法では、動作計画による軌道が誤っていても逐次修正により正しい動作軌道に収束させている。ただし、Tomasi ら [15] の手法の主眼は、モデル部位の衝突回避にある。単眼視からの身体追跡で、奥行方向でモデルの部位同士の衝突を回避することができよう。

Rahimi ら [16] は、画像の二次元並進移動と頭部の三次元回転運動の例について追跡ドリフトの修正を行っている。二次元並進移動では画像全体を移動させて動画像を作り、三次元回転運動では距離情報により

頭部画像を背景から切り離している。いずれの場合にも、測定された運動に背景や他の物体の運動が含まれないように配慮されている。これに対して、本手法は対象を背景や他の物体と切り離さずに運動を測定しているため、運動の測定に他の運動が含まれる可能性がある。それにもかかわらず、動作追跡の繰返しにより正しい動作軌跡に収束できている。

次に、ドリフトの修正精度を比較する。文献[16]の例によると、二次元並進移動におけるドリフト量は初期位置からの移動量にほぼ比例している。これは、最終フレームでのドリフト量をフレーム数で案分し、その案分量をフレーム順に差し引くことにより各フレームでのドリフトをある程度解消できることを意味する。ドリフト量が比例関係から外れている個所では、本手法は外れた分だけ誤差になるが、Rahimiらの手法は精度良く追従している。これは、本手法はドリフト量をフレーム間で一定としているのに対し、Rahimiらは異なるとしているためである。

一方、計算量は、本手法では、二次元ベクトル量の案分と加算のみであるが、Rahimiらの手法では未知数の数が全フレーム数  $\times$  二元の線形連立方程式を解く必要がある。したがって、計算量は本手法の方が少ないといえる。なお、Rahimiらの手法でキーフレームの姿勢を中間フレームに伝搬させるという考えは大田ら[14]と同じであるが、離れたフレームへも直接伝搬させることによって修正精度を向上させている。

## 7. む す び

連続する画像から運動パラメータを推定し、それを累積することによって動作を測定する手法は、推定誤差の累積によるドリフトが問題となっていた。本論文は、このドリフトを修正するための簡単で強力な手法を提案した。いくつかの実映像に対し、身体の動作を測定し提案手法の実効性を確認した。

今後の課題は自動化である。既に、画像から身体像の抽出、身体部位の特定、身体のサイズと姿勢の決定などの技術が開発されている。これらの技術と本提案手法を組み合わせることで動作測定の自動化を進めたい。

**謝辞** 本システムの GUI の開発及び手法の評価に御尽力頂いた、伊吹彰洋氏 ((株) Cube Inn) 及び土居真人氏 ((株) ASE) に感謝致します。また、野球映像の使用を許可頂きました、日本テレビ放送網株式会社及び読売巨人軍に感謝致します。

## 文 献

- [1] T. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol.81, pp.231–268, 2001.
- [2] J. Rehg and T. Kanade, "Visual tracking of high dof articulated structures: An application to human hand tracking," *Proc. European Conference on Computer Vision*, vol.2, pp.35–46, 1994.
- [3] J. Ohya and F. Kishino, "Human pose estimation from multiple images using genetic algorithm," *Proc. Intl. Conference on Pattern Recognition*, vol.1, pp.750–753, 1994.
- [4] D. Gavrila and L. Davis, "3-d model-based tracking of humans in action: A multi-view approach," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.73–80, 1996.
- [5] I. Kakadiaris and D. Metaxas, "Model-based estimation of 3d human motion with occlusion based on active multi-viewpoint selection," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.81–87, 1996.
- [6] 天谷賢治, 原 裕二, 青木 繁, "逆解析手法による 3 次元人体運動の再構成," *機械学会論文集 (C 編)*, vol.163, no.608, pp.1167–1171, 1997.
- [7] I. Haritaoglu, D. Harwood, and L.S. Davis, " $w^4$ :who? when? where? what? a realtime system for detecting and tracking people," *Proc. Intl. Conference on Automatic Face and Gesture Recognition*, pp.222–227, 1998.
- [8] C. Barron and I.A. Kakadiaris, "Estimating anthropometry and pose from a single image," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.669–676, 2000.
- [9] C. Taylor, "Reconstruction of articulated objects from point correspondences in a single uncalibrated image," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.255–261, 2000.
- [10] C. Sminchisescu and B. Triggs, "Kinematic jump processes for monocular 3d human tracking," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, vol.1, pp.69–76, 2003.
- [11] M. Yamamoto and K. Koshikawa, "Human motion analysis based on a robot arm model," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.664–665, 1991.
- [12] M. Yamamoto, A. Sato, S. Kawada, T. Kondo, and Y. Osaki, "Incremental tracking of human actions from multiple views," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.2–7, 1998.
- [13] C. Bregler and J. Malik, "Tracking people with twists and exponential maps," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.8–15, 1998.
- [14] 大田佳人, 山際貴志, 山本正信, "キーフレーム拘束を利

用した単眼動画画像からの人間動作の追跡,” 信学論 (D-II), vol.J81-D-II, no.9, pp.2208-2018, Sept. 1998.

[15] C. Tomasi, S. Patov, and A. Sastry, “3d tracking = classification + interpolation,” Proc. Intl. Conference on Computer Vision, pp.1441-1448, 2003.

[16] A. Rahimi, L.-P. Morency, and T. Darrell, “Reducing drift in parametric motion tracking,” Proc. Intl. Conference on Computer Vision, pp.315-322, 2001.

[17] B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” Proc. 7th Intl. Joint Conference on Artificial Intelligence, pp.674-679, 1981.

[18] 山本正信, “勾配法による動画画像と距離情報からの複数動物体の3次元運動解析,” 信学論 (D), vol.J69-D, no.5, pp.785-793, May 1986.

[19] 河内まさ子, 持丸正明, 岩澤 洋, 三谷誠二, “日本人人体寸法データベース 1997-98,” 通産省工技院くらしと JIS センター, 2000.

[20] B. Horn and B. Schunck, “Determining optical flow,” Artif. Intell., vol.17, pp.185-203, 1981.

[21] 山本正信, 特願 2003-182413.

付 録

1. ヤコビ行列の導出

$\frac{\partial \tilde{\mathbf{p}}'}{\partial \mathbf{x}}$  はヤコビ行列と呼ばれ, 次のように計算される.

$$\begin{aligned} \frac{\partial \tilde{\mathbf{p}}'}{\partial \mathbf{x}} &= \left( \frac{\partial \tilde{\mathbf{p}}'}{\partial r_{x_1}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial r_{y_1}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial r_{z_1}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{x_1}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{y_1}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{z_1}}, \right. \\ &\quad \left. \dots, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{x_i}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{y_i}}, \frac{\partial \tilde{\mathbf{p}}'}{\partial \phi_{z_i}} \right) \\ &= \left( T_1 \frac{\partial R_1}{\partial r_{x_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, T_1 \frac{\partial R_1}{\partial r_{y_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, \right. \\ &\quad T_1 \frac{\partial R_1}{\partial r_{z_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, T_1 \frac{\partial R_1}{\partial \phi_{x_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, \\ &\quad T_1 \frac{\partial R_1}{\partial \phi_{y_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, T_1 \frac{\partial R_1}{\partial \phi_{z_1}} T_2 \dots T_i \tilde{\mathbf{p}}_i, \\ &\quad \vdots \\ &\quad T_1 \dots T_i \frac{\partial R_i}{\partial \phi_{x_1}} \tilde{\mathbf{p}}_i, T_1 \dots T_i \frac{\partial R_i}{\partial \phi_{y_1}} \tilde{\mathbf{p}}_i, \\ &\quad \left. T_1 \dots T_i \frac{\partial R_i}{\partial \phi_{z_1}} \tilde{\mathbf{p}}_i \right) \end{aligned}$$

ここで,  $K_x, K_y, K_z, N_x, N_y, N_z$  を次のように定義する.

$$K_x = \frac{\partial R_i}{\partial r_{x_i}} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$K_y = \frac{\partial R_i}{\partial r_{y_i}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$K_z = \frac{\partial R_i}{\partial r_{z_i}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$N_x = \frac{\partial R_i}{\partial \phi_{x_i}} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$N_y = \frac{\partial R_i}{\partial \phi_{y_i}} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$N_z = \frac{\partial R_i}{\partial \phi_{z_i}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

また,  $T_i$  から  $T_1$  までの座標変換を

$$H_i = T_1 T_2 \dots T_i \tag{A.1}$$

とすれば,  $H_i$  は物体座標系  $\Sigma_i$  からカメラ座標系  $\Sigma_c$  への変換である.

$$\begin{aligned} H_i &= T_1 \dots T_j T_{j+1} \dots T_i \\ &= H_j T_{j+1} \dots T_i \end{aligned}$$

より,

$$H_j^{-1} H_i = T_{j+1} \dots T_i$$

なる関係を用い, 更に, 位置  $\tilde{\mathbf{p}}_i$  を式 (1) よりカメラ座標での位置  $\tilde{\mathbf{p}}$  で表せば, ヤコビ行列は

$$\begin{aligned} \frac{\partial \tilde{\mathbf{p}}'}{\partial \mathbf{x}} &= (H_1 K_x H_1^{-1} \tilde{\mathbf{p}}, H_1 K_y H_1^{-1} \tilde{\mathbf{p}}, H_1 K_z H_1^{-1} \tilde{\mathbf{p}}, \\ &\quad H_1 N_x H_1^{-1} \tilde{\mathbf{p}}, H_1 N_y H_1^{-1} \tilde{\mathbf{p}}, H_1 N_z H_1^{-1} \tilde{\mathbf{p}}, \\ &\quad \vdots \\ &\quad H_i N_x H_i^{-1} \tilde{\mathbf{p}}, H_i N_y H_i^{-1} \tilde{\mathbf{p}}, H_i N_z H_i^{-1} \tilde{\mathbf{p}}) \end{aligned} \tag{A.2}$$

と表せる.

## 2. 運動パラメータの推定

身体モデルを身体に一致させたとき、身体像はモデルに貼り付けられたテクスチャが画像上に投影されたものとみなすことができる。このとき、モデルの運動に応じて移動先の身体像を予測することができる。この予測画像が次の画像と一致するようにモデルの動きを求めれば、それが身体の運動となる。身体の運動パラメータ  $\mathbf{x}$  は画像間の差分を使って次のように得られる。

身体上の位置  $\mathbf{p}$  が  $\mathbf{p}'$  へ移動したとき、三次元の移動ベクトルを  $\Delta\mathbf{p} = \mathbf{p}' - \mathbf{p} = (\Delta x, \Delta y, \Delta z)^\top$  とする。同次座標系で表された位置  $\tilde{\mathbf{p}}$  とその移動先  $\tilde{\mathbf{p}}'$  から移動ベクトルへの変換は、

$$\Delta\mathbf{p} = \mathbf{p}' - \mathbf{p} = P_n(\tilde{\mathbf{p}}' - \tilde{\mathbf{p}}) \quad (\text{A.3})$$

である。ただし、

$$P_n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

カメラモデルを透視変換モデルとする。このとき、点  $\mathbf{p}$  は画面上の点  $(X, Y)$  へ、 $f$  をカメラの焦点距離として、

$$\begin{cases} X &= f \frac{x}{z} \\ Y &= f \frac{y}{z} \end{cases} \quad (\text{A.4})$$

で投影される。点  $\mathbf{p}'$  も同様に点  $(X', Y')$  へ投影されるとする。

投影面上の二次元移動ベクトルを  $\Delta\mathbf{P} = (\Delta X, \Delta Y)^\top = (X' - X, Y' - Y)^\top$  としたとし、三次元移動ベクトル  $\Delta\mathbf{p}$  と二次元移動ベクトル  $\Delta\mathbf{P}$  との関係は、変位量が小さいとき次式で関係づけられる。

$$\Delta\mathbf{P} = \frac{1}{z} P_f \Delta\mathbf{p} \quad (\text{A.5})$$

ただし、

$$P_f = \begin{pmatrix} f & 0 & -X \\ 0 & f & -Y \end{pmatrix} \quad (\text{A.6})$$

一方、投影面上で観測される動画像の濃淡値を  $E(X, Y, t)$  とする。 $t$  は時刻を表す。移動先でも濃淡値が変化しないとすれば、画面上の移動ベクトルは次の一次方程式に束縛されることが知られている [20]。

$$E_X \Delta X + E_Y \Delta Y + E_t = 0 \quad (\text{A.7})$$

ここで、 $E_X, E_Y$  はそれぞれ濃淡値  $E$  の  $X, Y$  方向の変化率（空間こう配）であり、 $E_t$  は位置  $(X, Y)$  における時間変化率（時間こう配）である。それぞれ、動画像から差分演算などで容易に計算することができる。

このとき、式 (3), (A.3), (A.5), (A.7) から運動パラメータ  $\mathbf{x}$  を拘束する次の方程式が得られる。

$$\mathbf{G}\mathbf{J}\mathbf{x} + E_t = 0 \quad (\text{A.8})$$

ここで、 $\mathbf{G} = (fE_X/z, fE_Y/z, -(XE_X + YE_Y)/z, 0)^\top$  とする。

モデルは身体に一致しているので、身体上の点の三次元座標値  $(x, y, z)$  はモデルから得られる。また、その点が属している部位からヤコビ行列を計算することができる。したがって、得られた式 (A.8) は運動パラメータ  $\mathbf{x}$  を求めるための線形推定式となる。モデル（身体）上のいくつかの点から導かれるこの推定式を連立させれば、次の線形連立方程式が得られる。

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (\text{A.9})$$

ここで、

$$\mathbf{A} = \begin{pmatrix} \mathbf{G}_1 \mathbf{J}_1 \\ \mathbf{G}_2 \mathbf{J}_2 \\ \vdots \\ \mathbf{G}_m \mathbf{J}_m \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -E_{t_1} \\ -E_{t_2} \\ \vdots \\ -E_{t_m} \end{pmatrix}$$

ただし、添字  $1, 2, \dots, m$  は測定点を表す。

(平成 16 年 7 月 28 日受付, 17 年 1 月 7 日再受付)



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。平元～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。現在、新潟大工学部情報工学科教授。工博、情報処理学会、IEEE CS 各会員。