

# ビデオ映像とCGの合成によるヴァーチャルファッションの実現

星野 准<sup>†,††</sup> 斉藤 啓史<sup>†††</sup>

本稿では、仮想的な鏡の前でCGによる衣服の試着や髪型の変更を行うことができる、仮想ファッションシステムを提案する。本システムでは、テレビカメラにより利用者の映像を取り込み、動作に合わせた衣服や髪型を合成してディスプレイ上に表示する。利用者の3次元動作に連動して、CGによる衣服や髪型も動くことにより、仮想的な鏡の前で試着を行っている効果を実現する。本稿では、人物の大まかな3次元モデルと時空間勾配法を利用して、人物の3次元動作を推定する手法を述べる。また、推定した3次元動作に合わせて、CG画像を生成することにより、実写映像との高精度な重ね合わせを行う。具体例として、人物動作に合わせた衣服と髪型を生成して、ビデオ映像への重ね合わせを行った結果を示す。

## Building Virtual Fashion Simulator by Merging CG and Humans in Video Sequences

JUNICHI HOSHINO<sup>†,††</sup> and HIROFUMI SAITO<sup>†††</sup>

This paper presents a virtual fashion simulator by merging video and CG. A person can wear any virtual cloth before the virtual mirror. The synthesized virtual cloth and hair style can be matched with the use's movement as if the user can ware them before the virtual mirror. In this paper, we describe a new method for merging virtual objects onto the human video sequence. First, we track current 3D pose of human body by using the spatio-temporal analysis. Then we generate CG objects and merge it with the human body in video. We demonstrate examples of merging virtual cloth and virtual hairstyle with the video captured images.

### 1. ま え が き

服装やヘアスタイルなどのファッションを変更することは、自己表現の手段や、変身願望を満たすための手段としても広く利用されている。我々が新しいファッションを試すときには、店頭飾られている衣服を試着したり、雑誌やカタログで新しい衣服を見て自分に似合うかどうかを想像したりする。このような活動を支援するために、近年のコンピュータグラフィクスや仮想現実の分野では、多様なファッションを計算機で生成して、利用者に分かりやすく見せる試みが行われている<sup>1)</sup>。

従来のシステムでは主に、CGによる人物キャラクターに衣服を着せて歩かせて、仮想的なファッション

ショーを生成することが行われてきた<sup>2),3)</sup>。これらの研究では、主に3次元的人物モデルと衣服の衝突計算や、衣服の動的なシミュレーション<sup>5)~9)</sup>などの問題を扱っている。ところが、個々の利用者が自分で衣服を着たときに、どのように見えるかについては、重点が置かれていない。

個々の利用者がファッションを変えた姿をシミュレートする方法としては、利用者の姿をビデオカメラで取り込んで、新しい衣服やヘアスタイルを合成することが考えられる。従来のシステムでは、静止画を対象とした合成手法の検討が行われている。たとえば、静止した人物の写真に衣服画像を合成したり<sup>4)</sup>、人物の顔写真にヘアスタイルを合成する手法<sup>10),11)</sup>が提案されている。ところが、従来のシステムでは静止画を対象としているため、利用者の動きに合わせて衣服や髪型が変化の様子を見ることができない。

本稿では、人物の動作に合わせた衣服やヘアスタイルを合成することができる、仮想ファッションシステムを提案する。本システムでは、利用者の3次元動作に連動して、CGによる衣服や髪型も動くことにより、

† 筑波大学機能工学系

College of Engineering Systems, University of Tsukuba

†† 科学技術振興事業団さきがけ研究 21

PRESTO, JST

††† 新潟大学

Niigata University

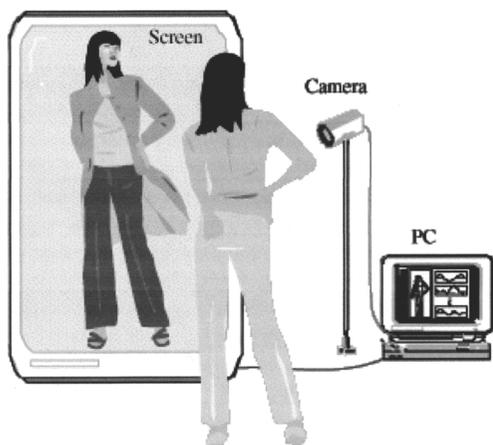


図1 仮想ファッションシステム

Fig. 1 Concept of the virtual fashion simulator.

仮想的な鏡の前で試着を行っている効果を実現する。本稿では、まず人物の大まかな3次元モデルと時空間勾配法を利用して、人物の3次元動作を推定する手法を述べる。次に、推定した3次元動作に合わせて、CG画像を生成することにより、実写映像との高精度な重ね合わせを行う手法を述べる。また、人体の初期位置を推定するために、拡張カルマンフィルタを用いてシルエット画像と人体モデルの位置合わせを行う。

本稿では、2章でシステムの概要を述べた後、3章でフレーム間の人体の動きを推定する手法を述べる。4章で人体の初期位置を推定する手法を述べる。5章で推定した人体の姿勢を利用して、衣服やヘアスタイルなどのCGを合成する手法と、実写画像とCGの重ね合わせを行う手法を述べる。6章で人物動作に合わせた衣服と髪型を生成して、ビデオ映像への重ね合わせを行った実験結果を示す。

## 2. システムの概要

図1に仮想ファッションシステムの概念図を示す。本システムでは、テレビカメラにより利用者の映像を取り込み、利用者の動作に合わせた衣服や髪型を合成してディスプレイ上に表示する。利用者の3次元動作に連動して、CGによる衣服や髪型も動くことにより、仮想的な鏡の前で試着を行っている効果を実現する。図2に本手法による衣服CG画像の合成例を示す。

実写映像とCG画像の動きの整合性をとる技術は、マッチムーブとも呼ばれている<sup>11)</sup>。従来手法では、情景とカメラの相対的な運動を推定して、CG画像を合成することが行われている<sup>13)~19)</sup>。また、人物の頭部に3次元モデルをあてはめて、照明と連動した影を生成したり、頭部に帽子などのCGの物体を合成する手



(a) 入力画像

(b) 合成結果

図2 衣服CG画像の合成例

Fig. 2 Example of merging CG cloth.

法が提案されている<sup>12)</sup>。これらの手法では、情景が多面体で近似できる場合や、顔のように単一の物体で表される場合を扱っている。そのため、人物のような複雑な被写体画像にCG画像を合成するときには、作業者が手で位置合わせをすることが必要となる。

本稿では、ビデオ映像から人物の3次元動作を推定して、実写映像とCG画像の合成を自動化することに重点を置く。本手法では、次の手順で実写画像とCG画像の合成を行う(図3参照)。

- (1) 動画像中の人物の3次元姿勢を推定する。
- (2) 推定した3次元姿勢を利用して衣服やヘアスタイルなどのCG画像を合成する。
- (3) 生成したCG画像と動画像中の人物を合成する。
- (4) 入力画像の各フレームで(1)~(3)の処理を繰り返す。

人体の動作を推定する手法としては、時空間勾配法などの運動解析を利用することが考えられる。ところが、運動解析に基づく手法では、フレーム間の変化量しか推定できないため、初期フレームの人体姿勢を利用者が正確に与えることが必要である。そのため、入力画像と人体モデルの位置合わせをするための労力が大きいという問題がある。本稿では、シルエット画像により初期位置を求めて、フレーム間の運動を時空間勾配法で推定する手法を述べる。

## 3. 人物動作の推定

本章では、人体の3次元モデルと動画像の時空間勾配法を利用して、3次元動作を推定する手法を述べる。時空間勾配法により人物の動作を推定する際には、関節数が増えた場合に運動モデルが複雑化するため、数式の導出が困難になる。本稿では、剛体変換を利用して簡便に定式化するとともに、関節数が増えた場合の運動モデルの導出を規則化する。これにより、Nリン

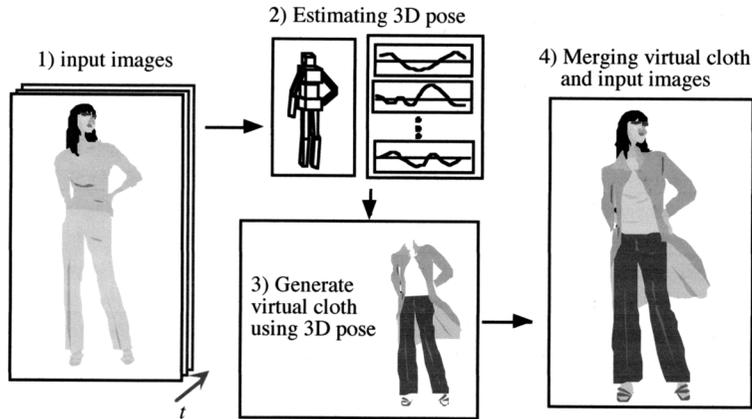


図3 システムの概要

Fig. 3 Overview of the proposed system.

クの物体の運動推定を容易に行うことができる。

### 3.1 時空間勾配法

本稿では、図4のような人体モデルを利用してフレーム間の人物動作を推定する。これらのパーツは胴体を頂点として、腕や頭部といった末端のパーツに向かって親子関係が結ばれている。親パーツが動くと、その子にあたるパーツは親パーツの座標系で一緒に動くことになる。図5にカメラ座標系およびワールド座標系と人体モデルの関係を示す。人体モデル上の点  $p = (x, y, z)^T$  と、画面上の点  $(X, Y)$ 、および輝度値の関係は、時空間勾配法により次のように表すことができる(付録A)。

$$G\dot{p} = -E_t \quad (1)$$

$$G = \left( f \frac{E_X}{z}, f \frac{E_Y}{z}, \frac{XE_X + YE_Y}{z} \right)$$

次に、人体モデルの全関節角の位置と姿勢を表すベクトルを  $q = (q_1, q_2, \dots, q_n)$ 、人体モデル上の任意の点を  $p = (x, y, z)^T$  とすると、 $q$  と  $p$  の関係は非線形関数  $p = f(q)$  で表すことができる。 $f$  は非線形であるが、 $q$  の時間変化率  $\dot{q}$  と速度ベクトル  $\dot{p}$  の関係は、ヤコビ行列  $J(q) = \partial f / \partial q$  により、線形式で関係付けることができる。

$$\dot{p} = J(q)\dot{q} \quad (2)$$

よって、人体モデルの全関節  $q$  の時間変化率  $\dot{q}$  と、画面上の輝度値の関係式を導くことができる。

$$GJ(q)\dot{q} = -E_t \quad (3)$$

式(3)を人体モデル上の複数の点から求めて、最小2乗法で解くことにより、人体の動作パラメータ  $\dot{q}$  を求める。初期姿勢に動作パラメータ  $\dot{q}$  を加算することで、各フレームの人体モデルの姿勢を得る。

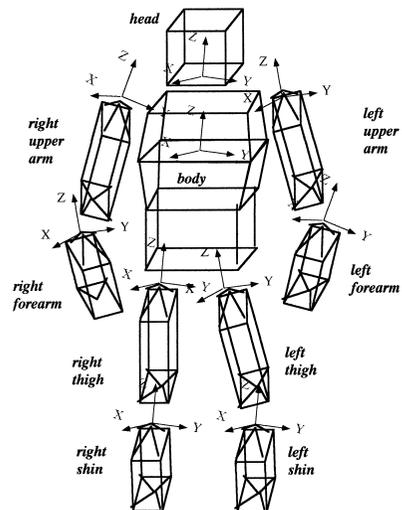


図4 人体モデル

Fig. 4 Human body model.

### 3.2 胴体のヤコビ行列の導出

3.1節では姿勢パラメータ  $q$  と人体モデル上の任意の点  $p$  の関係を  $p = f(q)$  で表した。この関係は人体の幾何学構造から導出することができる。本稿では、人体モデルの幾何学的な関係を記述する際に、同次座標系を利用する。同次座標系では、 $(x, y, z, w)$  の3次元空間内の座標値は  $(x/w, y/w, z/w)$  となる。

人体モデルの運動は、物体座標系における剛体変換  $E_i$  で表すことができる。剛体変換  $E_i$  は、 $3 \times 3$  の回転行列  $Q$  および、3次元の並行移動ベクトル  $T$  で次のように書くことができる。

$$E_i = \begin{bmatrix} Q_i & T_i \\ 0^t & 1 \end{bmatrix} \quad (4)$$

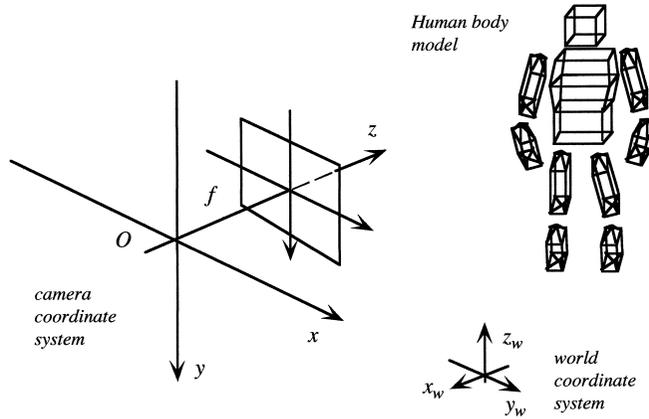


図5 カメラ座標系と人体モデルの関係

Fig. 5 Relationship of camera coordinate system and body model.

ただし、 $0^t$  はゼロの転置行列を表す。また、物体座標の点  $P_{oi} = (x_{oi}, y_{oi}, z_{oi}, 1)^T$  に剛体変換  $E_i$  を施す場合は、次式で表すことができる。

$$P'_{oi} = E_i P_{oi} \quad (5)$$

ヤコビ行列を求めるためには、人体の各パーツが運動したときに、カメラ座標系で見た人体モデル上の点  $P_s$  が、どのように変化するかを求めることが必要となる。この関係式は、次の3つのステップで求める。

(1) まず、カメラ座標系で見た胴体モデル上の点  $P_s$  を、ワールド座標系に変換する操作を、剛体変換  $E_s$  で表す。

$$P_w = E_s P_s \quad (6)$$

(2) 次に、胴体の運動を適用すると、胴体モデル上の点  $P_w$  が  $P'_w$  に移動する。 $P'_w$  を求めるためには、ワールド座標系の点  $P_w$  を物体座標系に変換して、胴体の運動  $E_i$  を適用した後に、ワールド座標系に戻す操作を行う。この変換規則を  $F_i = E_{oi} E_i E_{oi}^{-1}$  とおくと、胴体モデル上の点  $P'_w$  は次式で求めることができる。

$$P'_w = F_i P_w = F_i E_s P_s \quad (7)$$

(3) 最後に、ワールド座標系の点  $P'_w$  をカメラ座標系に変換する。

$$P'_s = E_s^{-1} P'_w = E_s^{-1} F_i E_s P_s \quad (8)$$

式(8)で得られた  $P'_s$  を時間微分することで、ヤコビ行列が得られる。

$$\frac{dP'_s}{dt} = J(q)\dot{q} \quad (9)$$

### 3.3 その他の関節のヤコビ行列の導出

3.2節では胴体の運動を表すヤコビ行列を導出したが、他の人体パーツの場合は胴体を親としているため、運動に依存関係が生じる。以下では、依存関係が1段

階の場合(上脚, 上腕, 頭)と、依存関係が2段階の場合(下脚, 下腕)の場合について、関係式を導出する。

(1) 上脚, 上腕, 頭の運動

胴体だけを親とする, 左右上脚, 左右上腕, 頭の5つの物体に関して定式化を行う。まず,  $P_s$  をワールド座標系に変換する。

$$P_w = E_{sj} P_s \quad (10)$$

この場合は、子の移動にともなう剛体変換と座標変換  $F_i$  を適用する。そして、親(この場合は胴体 ( $i=0$ ))の移動による変換  $F_1 = [E_{o1} E_1 E_{o1}^{-1}]$  を適用する。

$$P''_{sj} = E_{sj}^{-1} F_1 F_i E_{sj} P_s \quad (11)$$

式(9)と同様に  $P'_s$  を時間微分することで、ヤコビ行列を求める。

(2) 下脚, 下腕の運動

右左下脚, 右左下腕の場合は、上脚, 上腕を親に持ち、さらに胴体を親として持っている。この場合も、式(11)と同様な変換規則を適用して次式が得られる。

$$P'''_{sj} = E_{sj}^{-1} F_1 F_{i-1} F_i E_{sj} P_s \quad (12)$$

式(9)と同様に  $P'_s$  を時間微分することで、ヤコビ行列を求める。

(3)  $N$  個の物体が接続された場合

物体の数が増加した場合も、(2)と同様の規則で求めることができる。

$$P^n_{sj} = E_{sj}^{-1} F_1 \left[ \prod_{m=0}^{n-2} F_{i-m} \right] E_{sj} P_s \quad (13)$$

以上のヤコビ行列を利用して、式(3)を最小2乗法で解くことにより、人体の動作パラメータ  $\dot{q}$  を求める。

## 4. 初期姿勢の推定

3章で述べた動作推定法ではフレーム間の動きしか

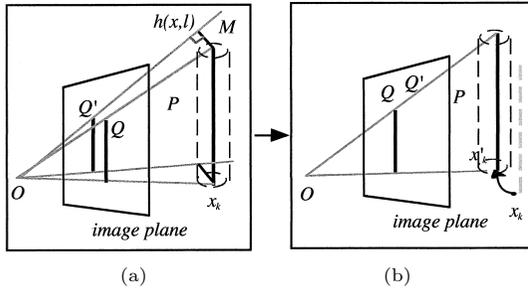


図 6 画面上の線分と胴体パーツの関係

Fig. 6 A relationship of the 2D center line and a corresponding body part.

推定することができないため、人体モデルの初期位置を推定することが必要となる。本稿では、拡張カルマンフィルタにより、シルエット画像から求めた中心線と、人体モデルの体軸を一致させる手法を述べる。なお、本稿の実験例では文献 25) の手法によりシルエット画像を抽出しているが、以下の手法はシルエット画像の抽出法に依存しない。

#### 4.1 マッチング誤差の定式化

グローバル座標系での人体の各パーツの姿勢  $x = [r, t]^T$  を位置  $t$  と姿勢  $r$  で表す。 $t$  は原点からの移動成分を表すベクトルである。また、 $r$  は回転軸にそったベクトルであり、ベクトルの大きさは回転角に等しい。図 6 において、2D 線分  $Q' = (q'_1, q'_2)$  を 2次元画面上で始点および終点を表すベクトル  $q'_1$  および  $q'_2$  で表す。3D 線分  $P = (p_1, p_2)$  をグローバル座標系での線分の始点および終点を表すベクトル  $p_1$  および  $p_2$  で表す。また、人体姿勢の初期位置で観測された 2D 線分  $Q'$  と、3D 線分  $P$  の組を、 $l = (P, Q')$  で表す。

図 6 (b) に示すように、人体姿勢  $x = [r, t]^T$  が正確な場合は、観測画像から得られた 2D 線分  $Q'$  とカメラの原点が作る平面  $M$  の上に、3D 線分  $P = (p_1, p_2)$  が含まれる。ところが図 6 (a) のように、人体姿勢  $x = [r, t]^T$  に誤差がある場合は、平面  $M$  と 3D 線分  $P$  の間に距離が生じる。本稿では、3D 線分  $P = (p_1, p_2)$  の始点および終点の平面  $M$  への最短距離を、予測誤差  $h(x, l)$  として利用する。

$$h(x, l) = \begin{bmatrix} (Rp_1 + t) \cdot n \\ (Rp_2 + t) \cdot n \end{bmatrix} \quad (14)$$

$h(x, l)$  は  $2 \times 1$  のベクトルである。 $R$  は  $r$  から導出される  $3 \times 3$  の回転行列である。また、 $n$  は面  $M$  の単位法線ベクトル

$$n = \frac{q'_1 \times q'_2}{\|q'_1 \times q'_2\|} \quad (15)$$

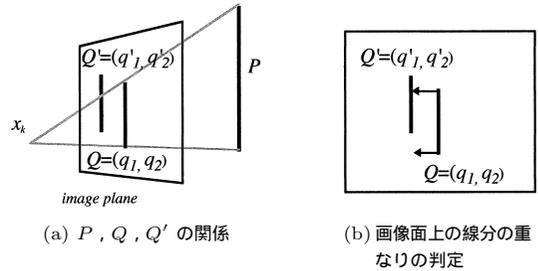


図 7 2D 画面上での対応付け

Fig. 7 Line correspondence on the 2D image plane.

である。 $x$  に誤差がない場合は、 $h(x, l) = 0$  となる。 $h(x, l)$  は非線形なので、観測値  $l = \hat{l}_i$  と推定値  $x = \hat{x}_{i-1}$  のまわりに 1 次テーラー展開をして線形近似する。

$$h(x_i, l_i) = \frac{\partial h}{\partial x}(x_i, l_i)(x - x_i) + \frac{\partial h}{\partial l}(x_i, l_i)(l - l_i) \quad (16)$$

ただし  $\frac{\partial h}{\partial x}$ 、 $\frac{\partial h}{\partial l}$  は偏微分を表す(付録 B 参照)。これらの偏微分係数は、3.4 節でカルマンフィルタを適用して  $x = [r, t]^T$  の最適値を求める際に利用する。

#### 4.2 2D, 3D 線分の対応付け

$h(x, l)$  を求めるためには 3D 線分  $P$  と、それに対応する画面上での 2D 観測線分  $Q'$  の組  $l = (P, Q')$  が必要である。しかし撮影環境内に多数の線分がある場合、1 つの 3D 線分に対して複数の 2D 線分の候補があることが多い。本稿では、以下の 2 つの手法を併用して対応付けを行う。まず 2D 画面上で見たときに重なっている線分を候補として選択する。次に、推定値の誤差分布と、2D, 3D 線分の誤差分布を考慮に入れて、統計的に一致すると判定できる線分を選択する。

##### (1) 2次元画面上での対応付け

図 7 に 2次元画面上で対応付けを行う手法を示す。まず、図 4 (a) のように、3D 線分  $P$  を画面上に投影して、2D 予測線分  $Q = (q_1, q_2)$  を求める。また、観測画像から 2D 線分  $Q' = (q'_1, q'_2)$  を得る。図 7 (b) のように、2D 線分  $Q' = (q'_1, q'_2)$  と 2D 予測線分  $Q$  が、画面上で重複領域を持つ場合は、対応の候補として選択する。

重複領域の判定方法としては、 $(q_1, q_2)$  を  $(q'_1, q'_2)$  に射影したときに、 $q_1$  あるいは  $q_2$  のどちらかの点が  $(q'_1, q'_2)$  の内部にあれば、2 つの線分は重なっていると考えることができる。まず、 $q_1$  を  $(q'_1, q'_2)$  に射影した場合の座標  $s$  は、 $(q'_1, q'_2)$  に沿う単位方向ベクトルを  $u = N[q'_2 - q'_1]$  ( $N[\cdot]$  はベクトルの正規化を表す) とすると、

$$S = q'_1 + u \cdot (q_1 - q'_1)u \quad (17)$$

である。\$s\$ が \$(q'\_1, q'\_2)\$ の内部に含まれる場合は、

$$(s - q'_1) \cdot (s - q'_2) < 0 \quad (18)$$

となる。この操作を \$q\_1, q\_2, q'\_1, q'\_2\$ に関して行い、2つの線分が画像面上で重なっているかを判定し、重なっていない線分は棄却する。

## (2) 誤差分布を考慮に入れた対応付け

人体姿勢の誤差分布 \$\sigma\_{x\_i}\$ と、2D、3D 線分の誤差分布 \$\sigma\_{l\_i}\$ を使って \$\chi^2\$ 検定を行い、統計的に一致しないと考えられる線分を棄却する。ここでは、2D 線分と 3D 線分が一致していると仮定して、仮説検定を行う。2D 線分と 3D 線分が一致している場合は、式 (3) の \$h(x, l\_i)\$ の線形近似式がゼロとなる。人体姿勢 \$x = [r, t]^T\$ と観測値 \$l = (P, Q')\$ の誤差分布を、正規分布 \$\sigma\_{x\_i} \sim N(0, P\_0)\$, \$\sigma\_{l\_i} \sim N(0, L\_0)\$ と仮定する。\$h(\hat{x}\_{i-1}, \hat{l}\_i)\$ の 1 次、2 次統計量は、

$$\begin{aligned} E[h(\hat{x}_{i-1}, \hat{l}_i)] &= 0 \quad (19) \\ E[h(\hat{x}_{i-1}, \hat{l}_i)] &= S_i \\ &= \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial x} P_{i-1} \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)^T}{\partial x} \\ &\quad + \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial l} L_{i-1} \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)^T}{\partial l} \quad (20) \end{aligned}$$

となる。これよりマハラノビス距離 \$d\$ を次のように求めることができる。

$$d(\hat{x}_{i-1}, \hat{l}_i) = h(\hat{x}_{i-1}, \hat{l}_i)^T S_i^{-1} h(\hat{x}_{i-1}, \hat{l}_i) \quad (21)$$

\$h(\hat{x}\_{i-1}, \hat{l}\_i)\$ の誤差分布は \$h(\hat{x}\_{i-1}, \hat{l}\_i) \sim N(0, S\_i)\$ であるため、\$d(\hat{x}\_{i-1}, \hat{l}\_i) \sim \chi^2(s)\$ となる (ただし \$s = \text{Rank}(S\_i)\$)。\$\chi^2(s)\$ 分布から、確率に基づくしきい値を設置することができる。たとえば \$d(\hat{x}\_{i-1}, \hat{l}\_i) < 9.21\$ は、\$S\_i\$ がランク 2 (\$s = 2\$) のときに 99% の線分が受け付けられることを意味している。しきい値以上の線分は、統計的に一致すると考えられないため、棄却することができる。複数の線分が候補として残った場合は、\$d(\hat{x}\_{i-1}, \hat{l}\_i)\$ の距離値が最小のものを選択する。

以上の 2 つの手法を併用して、3D 線分 \$P\$ と画面上で対応する 2D 観測線分 \$Q'\$ の組 \$l = (P, Q')\$ を求める。

## 4.3 人体姿勢の推定

対応付けがついた 2D 観測線分 \$Q'\$ と 3D 線分 \$P\$ の誤差 \$h(x, l)\$ が最小になるように、人体姿勢 \$x\$ を補正し、位置の誤差分布 \$\sigma\_{x\_i}\$ を減少させる。式 (3) の \$h(x, l\_i)\$ の線形近似式をゼロとおいて書き直すと、次のような観測方程式が得られる。

$$z_i = H_i x + v_i \quad (22)$$

ただし、

$$z_i = \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial x} \hat{x}_{i-1} - h(\hat{x}_{i-1}, \hat{l}_i) \quad (23)$$

$$H_i = \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial x} \quad (24)$$

$$v_i = \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial l} (l_i - \hat{l}_i) \quad (25)$$

である。ここで \$v\_i\$ の分散 \$R\_i\$ は次のように求める。

$$E[v_i v_i^T] = R_i = \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)}{\partial l} L_i \frac{\partial h(\hat{x}_{i-1}, \hat{l}_i)^T}{\partial l} \quad (26)$$

以上の観測方程式より、カメラ位置 \$x\_i\$ の 2 乗誤差を最小にする推定値は、カルマンフィルタ<sup>21), 22)</sup>で与えられる。

$$\hat{x}_i = \hat{x}_{i-1} + K_i [z_i - H_i \hat{x}_{i-1}] \quad (27)$$

$$K_i = \sigma_{i-1} H_i (H_i \sigma_{i-1} H_i^T + R)^{-1} \quad (28)$$

$$\sigma_i = \sigma_{i-1} - K_i H_i \sigma_{i-1} \quad (29)$$

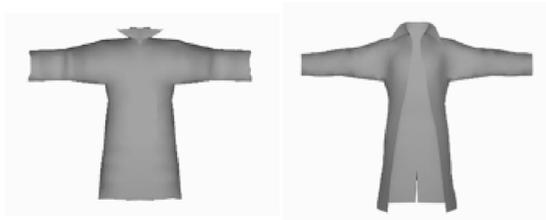
人体姿勢 \$x\_i\$ と誤差分布 \$\sigma\_i\$ は、新しい線分の対応付けが行われるたびに更新される。以上の姿勢推定を人体の各部位に適用することで、人体モデルの初期姿勢を推定する。本稿では、胴体の姿勢を最初に求めて、上腕、上脚、下腕、下脚と胴体に近い順番に中心線とのマッチングを行い、姿勢の推定を行った。

## 4.4 中心線の推定

シルエット画像から中心線を推定する際には、手や足などの部位が胴体に隠れた場合、姿勢を推定できない。ところが、本稿ではフレーム間の移動量を推定する手法と組み合わせて利用するため、処理に適したフレームだけを選択できればよい。本稿では、次の手順で中心線の推定を行う。

- (1) 中心線が抽出しやすい姿勢のシルエット画像をあらかじめ選択して記憶する。このときに、シルエット画像に対応した人体モデルの姿勢パラメータも与えておく。姿勢パラメータは大まかな値で構わない。
- (2) 低解像度の入力画像系列で、記憶したシルエット画像とのマッチングを行い、姿勢が類似しているフレームを選択する。
- (3) 選択したシルエット画像では、人体モデルの大まかな姿勢が分かっているため、画面上に各パーツの領域を投影することができる。それぞれの領域内のシルエットを切り出して、主軸を求めて中心線とする。

抽出した中心線を利用して、4.1~4.3 節で述べた手法を適用することで、人体モデルの正確な初期位置の推定を行う。



(a) 衣服を作成したところ (b) 人体モデルに衣服をフィッティングしたところ



(c) 歩行動作に合わせて衣服の動的シミュレーションをしているところ

図 8 衣服シミュレーションの例

Fig. 8 Example of cloth simulation.

## 5. 衣服とヘアスタイルの合成手法

本章では、洋服とヘアスタイルを具体例として、CGの生成手法と、ビデオ映像への合成手法を説明する。

### 5.1 衣服画像の生成

本稿では、人体の3次元モデルと姿勢パラメータが得られているため、モデルと衣服画像の衝突計算を行うことで、人体モデルの形状に合わせた衣服画像を生成する。また、フレーム間の人体の動作の変化に合わせて、衣服のダイナミックな動きをシミュレートする。衣服と人体の衝突計算を行う際に、図4の人体の近似モデルを利用すると、衣服の形状が不自然になる。そのため、本実験では、表面が滑らかな人体モデルに推定した関節角を与えて衣服CG画像を生成した。各パーツの寸法は、動作を推定した人体モデルと同一のものを利用した。

衣服画像の生成法としては、さまざまな手法が提案されている<sup>2)~9)</sup>。本稿では、実写映像と衣服CG画像の位置合わせを行うことに重点を置いているため、市販のソフトウェア(MAYA Cloth™)を利用した。その他の動的シミュレーションが可能な衣服生成法も同様に利用可能である。また、衣服と人体モデルの衝突計算もCGソフトの機能を利用したが、従来から提案されている手法も同様に利用することができる<sup>2),5)</sup>。図8に衣服シミュレーションを行っているところを示



(a) 頭部の3次元モデルに頭髮を張り付けたところ



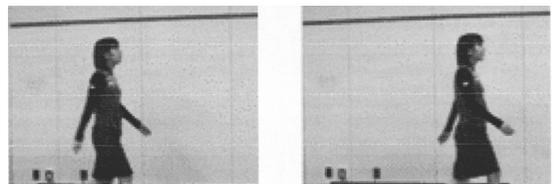
(b) 頭部の形状にフィッティングしたところ



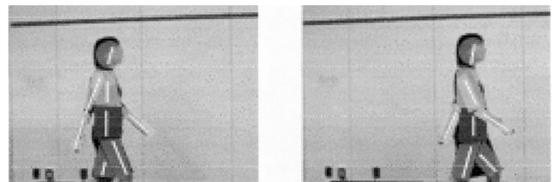
(c) 人物の移動に合わせてシミュレーションを行ったところ

図 9 ヘアスタイル生成の例

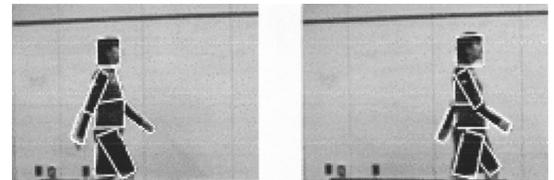
Fig. 9 Example of hair style simulation.



(a) 入力画像



(b) 中心線を抽出したところ



(c) 人体モデルの姿勢を推定したところ

図 10 初期位置の推定結果

Fig. 10 Result of initial position estimation.

す。(a)は衣服を生成したところ、(b)は人物の3次元形状にフィッティングしたところである。(c)は歩行動作に合わせて衣服シミュレーションを行っているところである。

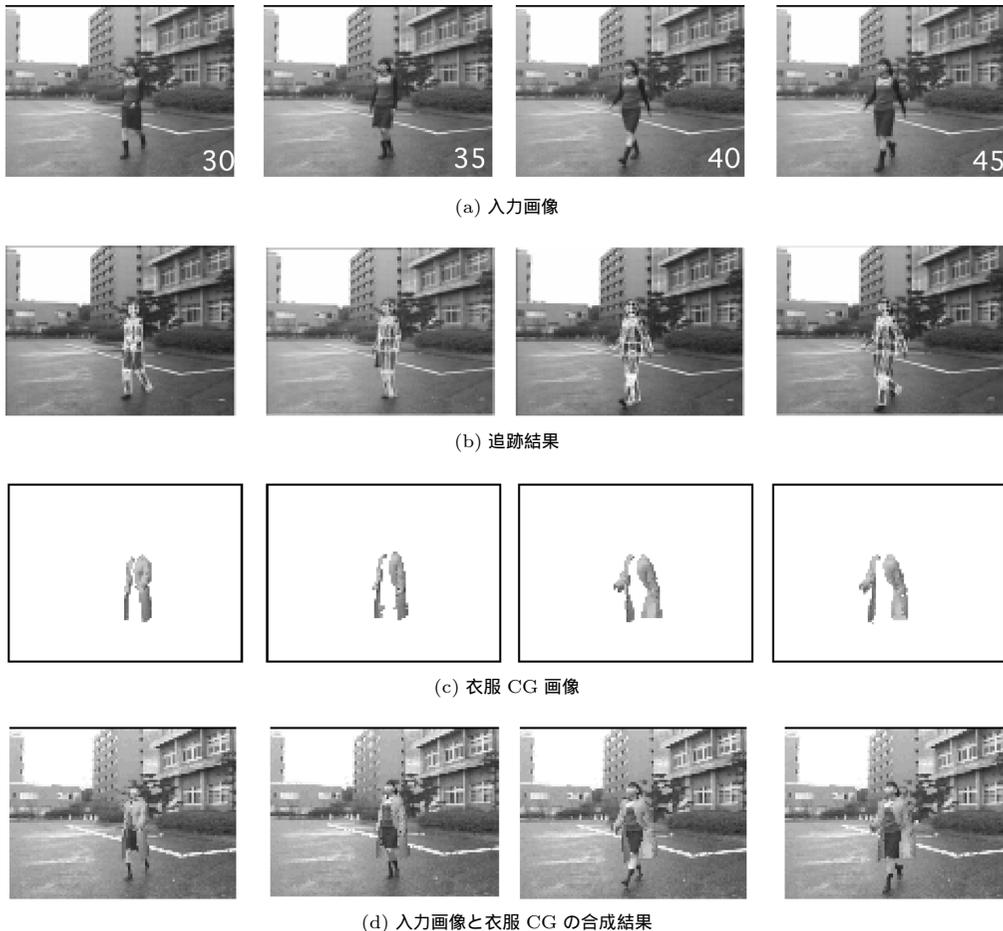


図 11 動画像系列に重ね合わせた例

Fig. 11 Example of merging virtual cloth.

## 5.2 ヘアスタイルの生成

ヘアスタイルを生成する場合は、頭部の 3 次元モデルに変形可能な頭髪を張り付けて、人物の移動に合わせて動的シミュレーションを行う。髪型のモデルは、頭部の 3 次元モデルに短冊状に細い髪を張り付けて、髪の様をテクスチャマッピングすることで生成する。本実験では、5.1 節と同様にヘアスタイルの動的シミュレーションに MAYA Cloth™ を利用した。図 9 にヘアスタイルの動的シミュレーションを行った例を示す。(a) は頭部に髪を張り付けたところである。(b) は重力を働かせて髪を頭部の形状にフィッティングしたところである。(c) は人物の移動に合わせて頭髪の動きをシミュレートしたところである。

## 5.3 実写画像との合成

衣服画像を合成する際には、人体部位による隠れを考慮に入れることが必要となる。たとえば、衣服の前に腕がある場合は、合成後も腕が見えていなくては

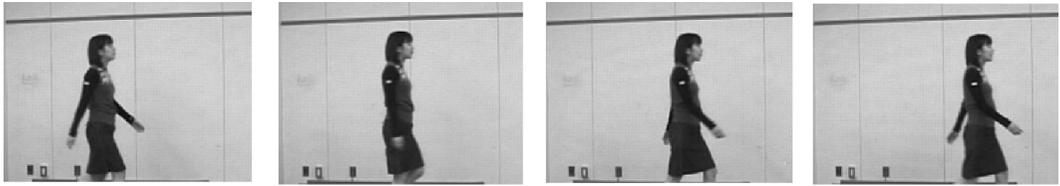
いけないので、単純に衣服画像を上から重ね合わせることができない。人体部位による隠れを判断するためには、衣服との前後関係を判断することが必要となる。本稿では、運動推定や衣服画像の生成に人体モデルを利用しているため、合成の段階でも利用することにした。本手法では、次のステップで人体画像と CG の合成を行う。

- (1) 入力画像を人体モデル上にテクスチャマッピングする。
- (2) 人体モデルに衣服画像を重ね合わせる。
- (3) 人体モデルと衣服モデルを同時にレンダリングする。

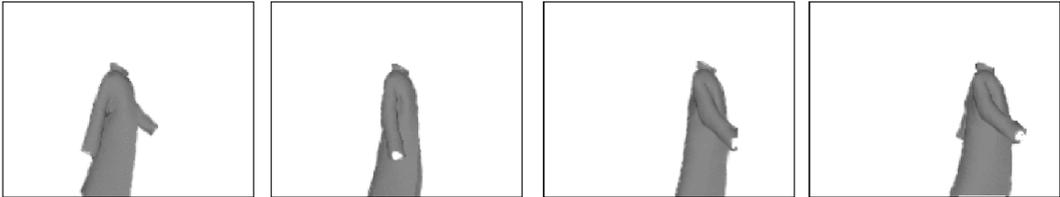
以上の処理により、推定した人体姿勢を利用して、CG の生成と実写画像への合成を行うことができる。

## 6. 実験結果

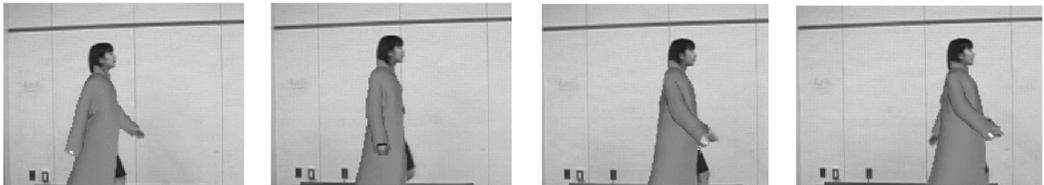
本手法の有効性を確認するために、ビデオ映像が



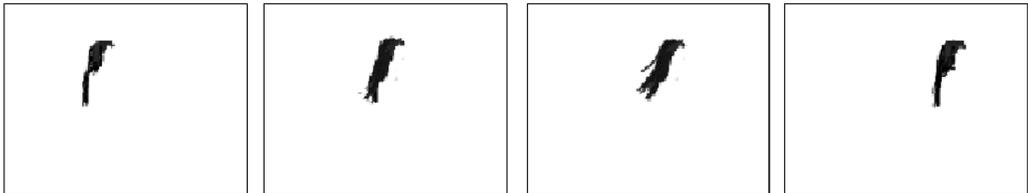
(a) 入力画像



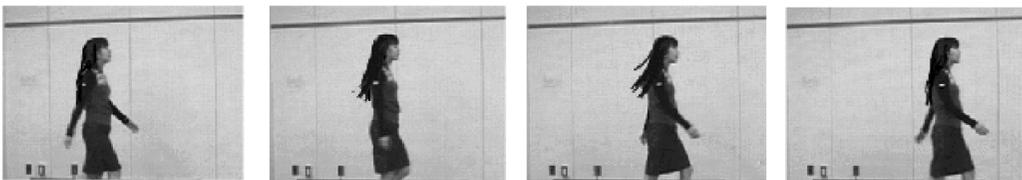
(b) 衣服 CG の生成結果



(c) 衣服 CG の合成結果



(d) ヘアスタイルの生成結果



(e) ヘアスタイルの合成結果

図 12 衣服とヘアスタイル合成結果

Fig. 12 Experimental results of merging CG cloth and hairstyle.

ら人物動作の推定を行い、CGによる衣服とヘアスタイルの合成を行った。本実験では、人物動作の推定はPC( PentiumIII 700M ,RAM192M ,Linux )で行い、衣服画像の生成と人物画像への合成はSGI Onyxで行った。

図 10 に人体モデルの初期位置を推定した結果を示す。(a)は入力画像、(b)は中心線を推定したところである。(c)は拡張カルマンフィルタにより中心線と人体モデルを一致させたところである。

図 11 に歩いている人物に衣服 CG 画像を重ね合わせたところを示す。(a)は入力画像、(b)は3次元動作を推定したところである。(c)は人物動作に合わせて衣服CGを生成したところである。(d)は(c)の衣服画像を(a)の入力画像に重ね合わせたところである。人物の移動に合わせた自然な衣服CGの合成が可能になっていることが確認できる。

図 12 に衣服とヘアスタイルの重ね合わせを行った例を示す。(a)は入力画像、(b)は人物動作に合わせて

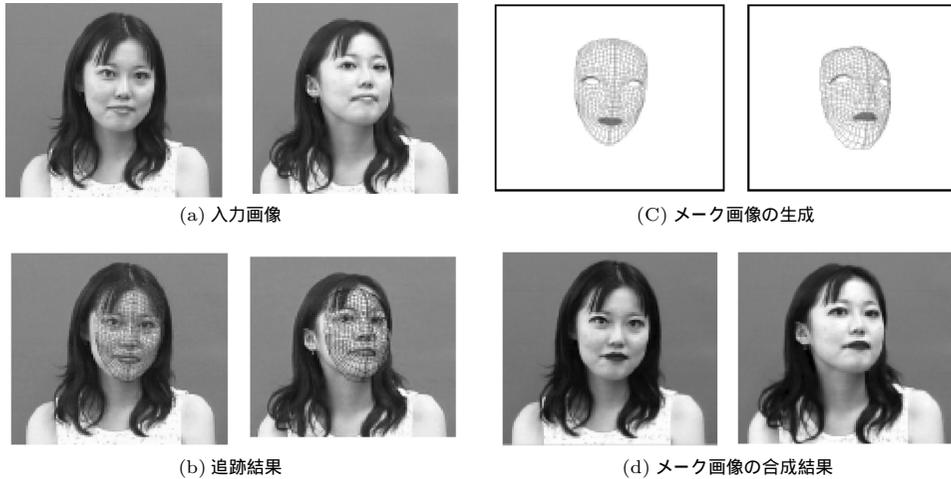


図 13 メーク画像を合成したところ

Fig. 13 Experimental results of merging virtual make-up.

衣服CGを生成したところである。(c)は(b)の衣服画像を(a)の入力画像に重ね合わせたところである。(d)は人物動作に合わせてヘアスタイルを生成したところである。(e)は(d)のヘアスタイルを合成したところである。図13は同様の手法により、顔画像にメーク画像を重ね合わせた実験例である。これらの実験から、人物の移動に合わせて衣服とヘアスタイルの合成が行われていることが確認できる。

上記の処理に必要な計算時間は、各フレームで位置合わせに2~3秒、衣服や髪型の生成とレンダリングに2~10秒程度である。衣服や髪型の生成に関しては、モデルの複雑さに依存して処理時間が変化する。本稿では実時間処理のための高速化を行っていないため、今後の検討により改善することが可能であると考えられる。

## 7. む す び

本稿では、仮想的な鏡の前で衣服やヘアスタイルの変更を行うことができる仮想ファッションシステムを提案した。本システムでは、従来のCG作品を作ることを重視したシステムに対して、利用者が自分の新しい姿を見ることに重点を置いている。また、利用者の動作に追従して、衣服やヘアスタイルも動的に動くことにより、現実感の向上を狙っている。本稿では、ビデオ映像からの3次元的な人物動作を推定するとともに、実写映像を利用して合成実験を行い、システムの実現が可能であることを示した。

今後の課題としては、本手法の高速化により、対話的に仮想的なファッションの変更を行うことができるシステムを構築することが考えられる。

謝辞 本稿の人物イラストレーションを作成していただきました星野裕美さんに感謝いたします。また、6章のメーク画像の合成実験を手伝っていただきました新潟大学の里見智一さんに感謝いたします。

## 参 考 文 献

- 1) Gray, S.: Virtual Fashion, *IEEE Spectrum*, pp.18-25 (1992).
- 2) Carignan, M., Yang, Y., Thalman, N. and Thalman, D.: Dressing Animated Synthetic Actors with Complex Deformable Clothes, *ACM SIGGRAPH'92*, pp.99-104 (1992).
- 3) Volino, P., Courchesne, M. and Thalmann, N.M.: Versatile and Efficient Techniques for Simulating Cloth and Other Deformable Objects, *Computer Graphics Proceedings*, pp.137-144 (1995).
- 4) 坂口嘉之, 美濃導彦, 池田克夫: 自分に合った衣服をオーダーする: 仮想服飾オーダーメイドシステム, *電子情報通信学会誌*, Vol.82, No.4, pp.404-411 (1999).
- 5) 坂口嘉之, 美濃導彦, 池田克夫: 仮想服飾環境PARTY—衣服形状計算における, 衣服と人体との衝突計算方法, *電子通信学会論文誌*, Vol.J78-D-II, No.3, pp.483-491 (1995).
- 6) Okabe, H., Imaoka, H., Tomita, T. and Niwaya, H.: Three Dimensional Apparel CAD System, *ACM SIGGRAPH'92*, pp.105-110 (1992).
- 7) Baraff, D. and Witkin, A.: Large Steps in Cloth Simulation, *Proc. SIGGRAPH98 Conference*, pp.43-54 (1998).
- 8) Eberhardt, B., Weber, A. and Strasser, W.: A fast, flexible, particle-system model for cloth

- draping, *IEEE Computer Graphics and Applications*, Vol.16, pp.52-59 (1996).
- 9) Ng, H.N. and Grimsdale, R.L.: Computer graphics techniques for modeling cloth, *IEEE Computer Graphics and Applications*, Vol.16, pp.28-41 (1996).
- 10) Barron, C.: Matte Painting in the Digital Age, Animation Sketches, *SIGGRAPH'98* (1998).
- 11) Blanz, V. and Vetter, T.: A Morphable Model for the Synthesis of 3D Faces, *Proc. SIGGRAPH 99*, pp.187-194 (1999).
- 12) Bajura, M., Fuchs, H. and Ohbuchi, R.: Merging virtual objects with the real world: Seeing ultrasound imagery within the patient, *SIGGRAPH '92*, pp.203-210 (1992).
- 13) Bajura, M. and Neumann, U.: Dynamic Registration Correction in Video-Based Augmented Reality Systems, *IEEE Computer Graphics and Applications*, Vol.15, No.5, pp.52-60 (1995).
- 14) Bajura, M. and Neumann, U.: Dynamic Registration Correction in Augmented Reality Systems, *Proc. IEEE VRAIS 1995*, pp.189-196 (1995).
- 15) Edwards, E.K., Rolland, J.P. and Keller, K.P.: Video See-Through Design for Merging of Real and Virtual Environments, *Proc. IEEE VRAIS 1993*, pp.222-233 (1993).
- 16) Gottschalk, S. and Hughes, J.: Autocalibration for Virtual Environments Tracking Hardware, *SIGGRAPH'93*, pp.65-72 (1993).
- 17) Azuma, R. and Bishop, G.: Improving Static and Dynamic Registration in an Optical See-Through HMD, *SIGGRAPH '94*, pp.197-204 (1994).
- 18) Janin, A.L., Mizwell, D.W. and Caudell, T.P.: Calibration of Head-Mounted Display for Augmented Reality Applications, *Proc. IEEE VRAIS 1993*, pp.246-255 (1993).
- 19) 大谷 淳, 海老原一之, 棚沢 順: 仮想変身システム—自分の姿を好みの姿に変えられるシステムを目指して, *ATR Journal*, No.26, pp.6-7 (1996).
- 20) 宮田一乗, 黒川雅人: ヘアスタイル・シュミレーション・システムの試作, 情報処理学会グラフィックスとCAD 74-9, pp.47-52 (1995).
- 21) 宮田一乗: 画像のワーピング処理を用いたヘアスタイルの合成, 画像電子学会論文誌, Vol.25, No.5 (1996).
- 22) 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物動作の3次元画像追跡, 信学論(D-II), Vol.79-D-II, No.1, pp.71-83 (1996).
- 23) Gelb, A., et al.: *Applied Optimal Estimation*,

MIT Press, Cambridge, MA (1974).

- 24) Sorenson, H.W.: *Kalman Filtering: Theory and Application*, IEEE Press, New York (1985).
- 25) 星野准一, 増田 功: 照明変動にロバストな移動物体抽出法, 画像電子学会論文誌, Vol.28, No.1, pp.41-47 (1999).

## 付 録 A

入力画像から人物の姿勢データを推定する手法を述べる。まず、人体モデル上の任意の点を  $p = (x, y, z)^T$  とすると、カメラ座標系に投影した画面上の点  $(X, Y)$  は、

$$X = fx/z, \quad Y = fy/z \quad (A1)$$

となる。同様に、人体モデル上の点  $p$  の3次元速度ベクトル  $\dot{p} = (u, v, w)^T$  と、画面上の速度ベクトル  $(U, V)$  の関係は、

$$U = (f_u - X_w)/z, \quad V = (f_v - Y_w)/z \quad (A2)$$

である。画面上の点  $(X, Y)$  における時刻  $t$  での明るさを  $E(X, Y, t)$  とする。また  $(E_X, E_Y)$  を輝度微分,  $E_t$  を時間微分とすると、画面上の速度ベクトル  $(U, V)$  は、次の時空間勾配の拘束条件式を満たす。

$$E_X U + E_Y V + E_t = 0 \quad (A3)$$

式(A2)を式(A3)に代入すると、次式が得られる。

$$G\dot{p} = -E_t V \quad (A4)$$

$$G = \left( f \frac{E_X}{z}, f \frac{E_Y}{z}, \frac{X E_X + Y E_Y}{z} \right)$$

式(A4)は、人体モデル上の点  $p = (x, y, z)^T$  と、画面上の点  $(X, Y)$ 、および輝度値の関係を示している。

## 付 録 B

4.2節式(16)の偏微分係数  $\frac{\partial h}{\partial x}$ ,  $\frac{\partial h}{\partial l}$  を求める。

(1)  $\frac{\partial h}{\partial x}$  の算出

以下、(a)  $\frac{\partial h}{\partial r}$ , (b)  $\frac{\partial h}{\partial t}$  を求める。

(a)  $\frac{\partial h}{\partial r}$

まず、回転行列  $R$  は次のように書くことができる。

$$R = e^H = 1 + H/1! + H^2/2! + \dots + H^n/n!$$

$$H = \begin{pmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_y \\ -r_y & r_x & 0 \end{pmatrix}$$

ただし  $r = (r_x, r_y, r_z)$  である。この関係を使うと  $\frac{\partial R_p}{\partial r}$  は、

$$\frac{\partial R_p}{\partial r} = \frac{\partial e_p^H}{\partial r} = \left[ \frac{\partial e_p^H}{\partial r_x}, \frac{\partial e_p^H}{\partial r_y}, \frac{\partial e_p^H}{\partial r_z} \right]$$

と書き直すことができる． $\frac{\partial e_p^H}{\partial r_x}$  は，

$$\frac{\partial e_p^H}{\partial r_x} = e^H \frac{\partial H}{\partial r_x} = e^H \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = e^H [i \times I]$$

である．同様に  $\frac{\partial e_p^H}{\partial r_y}$  および  $\frac{\partial e_p^H}{\partial r_z}$  は，

$$\frac{\partial e_p^H}{\partial r_x} = e^H [j \times I], \quad \frac{\partial e_p^H}{\partial r_z} = e^H [k \times I]$$

したがって，

$$\frac{\partial R_p}{\partial r} = e^H [p \times I] = R[p \times I]$$

となる．よって  $\frac{\partial h}{\partial r}$  は次のように求められる．

$$\frac{\partial h}{\partial r} = \begin{bmatrix} n^T \cdot R[p_1 \times I] \\ n^T \cdot R[p_2 \times I] \end{bmatrix}$$

(b)  $\frac{\partial h}{\partial t}$

$h(x, l)$  を  $t$  で微分する．

$$\frac{\partial h}{\partial t} = \begin{bmatrix} n^T \\ n^T \end{bmatrix}$$

(2)  $\frac{\partial h}{\partial l}$  の算出

$$\frac{\partial h}{\partial l} = \left[ \frac{\partial h}{\partial p_1}, \frac{\partial h}{\partial p_2}, \frac{\partial h}{\partial q'_1}, \frac{\partial h}{\partial q'_2} \right]$$

以下， $\frac{\partial h}{\partial p_1}$ ， $\frac{\partial h}{\partial p_2}$ ， $\frac{\partial h}{\partial p'_1}$ ， $\frac{\partial h}{\partial p'_2}$  をそれぞれ導出する．

(a)  $\frac{\partial h}{\partial p_1}$ ， $\frac{\partial h}{\partial p_2}$

$h(x, l)$  を  $p_1, p_2$  で微分する．

$$\frac{\partial h}{\partial p_1} = \begin{bmatrix} n^T R \\ 0 \end{bmatrix}, \quad \frac{\partial h}{\partial p_2} = \begin{bmatrix} 0 \\ n^T R \end{bmatrix}$$

(b)  $\frac{\partial h}{\partial q'_1}$ ， $\frac{\partial h}{\partial q'_2}$

まず  $\frac{\partial n}{\partial q'_1}$  を計算する．

$$\begin{aligned} \frac{\partial n}{\partial q'_1} &= \frac{\partial}{\partial q'_1} (q'_1 \times q'_2) \frac{1}{\|q'_1 \times q'_2\|} \\ &\quad + (q'_1 \times q'_2) \left( \frac{\partial}{\partial q'_1} \frac{1}{\|q'_1 \times q'_2\|} \right) \\ &= \frac{1}{\|q'_1 \times q'_2\|} [-q'_2 \times I] + nn^T [q'_2 \times I] \end{aligned}$$

が得られる．したがって，

$$\begin{aligned} \frac{\partial n}{\partial q'_1} &= \left[ \frac{1}{\|q'_1 \times q'_2\|} [Rp_1 + t]^T [-q'_2 \times I] + nn^T [q'_2 \times I] \right] \\ &\quad \left[ \frac{1}{\|q'_1 \times q'_2\|} [Rp_2 + t]^T [-q'_2 \times I] + nn^T [q'_2 \times I] \right] \end{aligned}$$

同様に  $\frac{\partial n}{\partial q'_2}$  は，

$$\begin{aligned} \frac{\partial n}{\partial q'_2} &= \left[ \frac{1}{\|q'_1 \times q'_2\|} [Rp_1 + t]^T [-q'_1 \times I] + nn^T [q'_1 \times I] \right] \\ &\quad \left[ \frac{1}{\|q'_1 \times q'_2\|} [Rp_2 + t]^T [-q'_1 \times I] + nn^T [q'_1 \times I] \right] \end{aligned}$$

となる．

(平成 12 年 9 月 28 日受付)

(平成 13 年 1 月 11 日採録)



星野 准一 (正会員)

平成元年，早稲田大学理工学部電子通信学科卒業．同年，セコム IS 研究所入所．平成 3 年から平成 5 年 MIT メディアラボ客員研究員．平成 11 年から新潟大学自然科学研究科助手．平成 12 年から筑波大学機能工学系講師．同年より科学技術振興事業団さきがけ研究 21「情報と知」領域研究員を兼務．工学博士．人物動作の推定，実写画像と CG の合成，デジタルアクター等の研究に従事．IEEE，ACM，電子情報通信学会，画像電子学会各会員．



斉藤 啓史

平成 11 年，新潟大学工学部卒業．現在，新潟大学大学院自然科学研究科在籍．実写画像と CG の合成，ヴァーチャルファッション等の研究に従事．