

マーカレスモーションキャプチャによる仮想ファッションショー

岡田 隆三[†] シュテンガービヨン[†] 池 司[†] 近藤 伸宏^{††}

† (株)東芝 研究開発センター
〒 211-0016 神奈川県川崎市幸区小向東芝町 1
†† (株)東芝 セミコンダクター社
〒 212-8520 川崎市幸区堀川町 580-1
E-mail: †ryuzo.okada@toshiba.co.jp

あらまし 2台のカメラによって、人物の追跡と姿勢推定をリアルタイムに行うシステムを試作した。まず、事前に追跡対象人物の三次元形状モデルと、様々な姿勢標本から構成される姿勢辞書を取得しておく。姿勢推定は、形状モデルを変形させて姿勢辞書中の様々な姿勢のシルエット画像を生成し、画像から背景差分によって得られる観測シルエットと最も一致する姿勢を選択することによって行う。このシステムを用いて、舞台上を歩く女性モデルの姿勢を推定し、CGによって衣服を着せて表示する仮想ファッションショーシステムを試作した。このシステムは、ファッションショーで使用できる高品質な動きを生成することが特徴である。

キーワード モーションキャプチャ、姿勢推定、関節物体追跡

Virtual fashion show using marker-less motion capture

Ryuzo OKADA[†], Bjorn STENGER[†], Tsukasa IKE[†], and Nobuhiro KONDOH^{††}

† Corporate Research & Development Center, Toshiba Corporation
1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 212-8582
†† Semiconductor Company, Toshiba Corporation
580-1, Horikawa-cho, Saiwai-ku, Kawasaki-shi 212-8520
E-mail: †ryuzo.okada@toshiba.co.jp

Abstract We have developed a motion capture system using two cameras, which is capable of tracking and estimating postures of a person in real-time. We, beforehand, obtain a 3-D shape model of the person to be tracked and a motion dictionary consisting of many posture exemplars. We estimate the postures by selecting a posture in the dictionary whose silhouette projected onto an image is the most similar to that observed from a current image. Based on this system, we have developed a virtual fashion show system also. The system generates a high-quality motion that is capable of being used for a fashion show.

Key words motion capture, posture estimation, articulated object tracking

1. はじめに

人間の姿勢を獲得する技術であるモーションキャプチャは、コンピュータグラフィック(CG)キャラクタの動作生成や、ジェスチャー認識等人間の動作のコンピュータへの入力などに応用が期待されている。現在実用化されているモーションキャプチャシステムには、以下のようなものがある。

- 体の様々な位置に目印(マーカー)をつけて、それを多数のカメラで撮影して各関節の角度を推定する[1]。現在のところ、この手法が普及しているが、装置が高価であることや、マーカーを着用しなければならないなど被計測者の負担が大き

いこと、専用のスタジオを必要とすることなど、制約が大きい。

- 体の各関節部分にポテンシオメータやジャイロを装着して、関節角度を推定する[2]。この場合も装置が高価で、被験者への負担が大きい。

• 画像から手作業で体の関節位置を抽出し、人体モデルを当てはめる手動式[3]、もしくは、動作の開始および終了の姿勢を手動で画像を用いて当てはめ、その間の動きを画像処理[4]によって推定する半自動式[6]がある。この方法は、ソフトウェア的に実現できるので、PCとカメラで安価に装置を構成できるが、姿勢を手動で与えなければならない。

このように、現在実用化されているシステムは、グラフィッ

クリエイタなどのプロフェッショナル向けで、一般的なユーザがジェスチャ入力に使用するなど手軽に使えるものではない。マーカーやセンサーを取り付けることは、一般ユーザが使用する上で大きな障害となるため、画像による自動のモーションキャプチャシステムが一般ユーザ向けには最も有力と考えられる。

画像によるモーションキャプチャは、人体の三次元モデルを使用しない手法と、使用する方法に大別される。前者は、Shape-from-Silhouette [7], [8] と呼ばれるアプローチで、被計測者を取り囲むように多数のカメラを配置し、各カメラから被計測者のシルエットを生成する。得られた全てのシルエットを三次元空間に逆投影し、その共通領域を求める。それが被計測者の形状である。

後者は、三次元モデルから様々な姿勢を生成し、その画像への投影像が画像から得られる被計測者のシルエット、速度情報、エッジ情報などの特徴量と最も良く一致する姿勢を選択することにより姿勢推定を行う。この方法はカメラの台数が少なくてすむが、汎用性や安定性、精度に問題が多く残っている。大別すると、(1) 一般的な環境で画像から安定な被計測者の特徴量抽出、(2) 抽出した特徴と三次元モデルとの類似度の定義、(3) 姿勢パラメータが張る高次元パラメータ空間内の最適姿勢の探索手法という課題がある。(1), (2) に関しては、背景差分 [9]、色情報 [10]、エッジ [11]、動き (オブティカルフロー) [12] といった特徴量を用いて、様々な類似度が定義されている。(3) は、類似度が最も高くなる姿勢を高次元の姿勢パラメータ空間の中で探索する問題であるが、類似度と姿勢パラメータの関係は非線形であり、多数の極大値、つまり偽の最適姿勢が存在する。また、画像特徴から姿勢が一意に定まらない特異点問題も存在する。このような問題に対して、姿勢の確率密度分布を、サンプリングした追跡候補 (particle) の確率密度分布で近似し、フィルタリングによって確率密度分布の時間的な伝播をおこなう Particle filtering [13] が注目されているが、人体の姿勢推定のような複雑な確率密度分布を近似するためには、多数の Particle が必要で、実時間実行が難しい。また、探索コストを軽減するため、逆運動学 (Inverse Kinematics; IK) を利用した手法も提案 [5] されている。この手法は、手足の位置など人体の各パーツの末端を検出する必要がある。このような三次元形状モデルに基づく方法は、形状モデルをもっているため、運動を修正したり、服を着せ替えたりといった加工が、前者に比べて容易に出来るという利点もある。

本報告では、後者のアプローチで実時間動作するモーションキャプチャシステムについて報告する。肌色領域などを用いて手足の位置など人体の各パーツの末端を検出するアプローチは取らず、より汎用性の高いシルエットマッチングに基づく姿勢推定手法を採用する。また、モーションキャプチャシステムの結果を用いて実際のモデルとは異なる衣服を CG で表現する、仮想ファッショントーショーシステムを試作した。このようなアプリケーションでは、高い品質、つまり自然な動きを生成する必要があるため、あらかじめ取得しておいた動きデータを使用して、自然な動きの生成を行う。

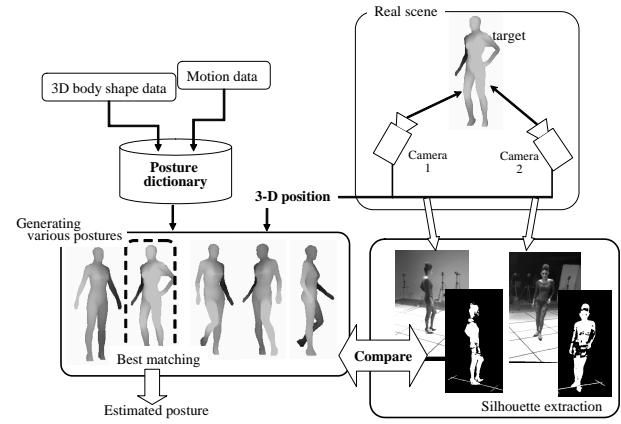


図 1 モーションキャプチャシステムの構成
Fig. 1 Organization of our motion capture system

2. 姿勢推定手法

試作したモーションキャプチャシステムの構成を図 1 に示す。追跡対象の人物をあらかじめ三次元スキャナで計測し、また、市販のモーションキャプチャシステムによって、推定に必要な姿勢を計測しておく。これらの追跡対象に関するデータを用いて、姿勢推定に必要な姿勢辞書を事前に作成しておく。

本システムでは、2 台以上の高速度カメラを用いて追跡対象人物を撮影する。フレームレートは 100fps、解像度 640×480 画素である。各カメラの画像に対して、高フレームレート画像を用いた追跡アルゴリズム [16] を適用して、画像内で人物を追跡する。カメラはあらかじめ Tsai の手法 [14], [15] によってキャリブレーションを行い、各カメラの画像内の追跡結果を用いて追跡対象の実世界中での三次元位置を求める。

推定した追跡対象の三次元位置に、姿勢辞書を用いて様々な姿勢の候補を生成し、画像に投影し、姿勢候補のシルエット画像（候補シルエットと呼ぶ）を生成する。一方、背景差分によって、撮影した画像から現在の追跡対象人物のシルエット画像（観測シルエットと呼ぶ）を抽出する。観測シルエット画像と、候補シルエットを比較し、観測シルエットに最も近い候補シルエットの姿勢を現在の姿勢として推定する。このようなシルエットによる比較を、4 フレーム毎に行う (25fps) ことにより、リアルタイムモーションキャプチャを実現している。

試作したモーションキャプチャシステムは、4. 章で述べるデジタルファッショントーショー向けのシステムである。そのため、推定を行う姿勢は、ファッショントーショーを行うのに必要な歩行、ポージング、ターンに限定している。

2.1 人体モデルと座標系

人体は、各部位は関節によって接合されており、関節物体とみなすことができる。以下では、関節と関節を繋ぐ線分をボーンと呼ぶ。各ボーンには固有の座標系 (局所座標系) を定義する。また、シーンに固定な座標系をワールド座標系、各カメラに固定の座標系をカメラ座標系とする。

関節物体は木構造で表すのが一般的で、人体の局所座標系とワールド座標系、カメラ座標系を含めて、シーン全体の座標系

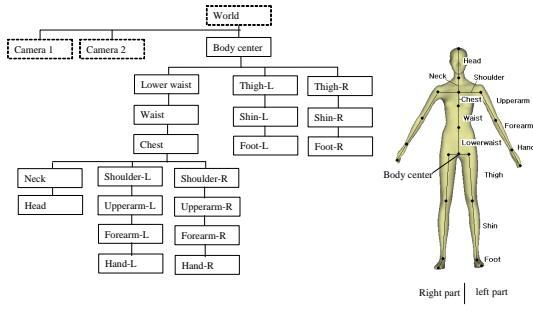


図 2 人体モデルの構造

Fig. 2 Structure of our body model

のつながりを木構造で表現すると、図 2 のようになる。点線の四角形で囲われたノードは、実体としてのボーンは持たない。ワールド座標系には、カメラ座標系と人体モデルのトップノードである Body center ノードが接続されている。Body center ノード以下の構造が、人体モデルの木構造である。

2.2 姿勢辞書

姿勢辞書は、市販されているモーションキャプチャシステム、および三次元人体形状スキャナを用いて作成する。

三次元形状スキャナからは、総頂点数 200 万点程度の人体表面形状のポリゴンデータが得られるが、欠落している部分があるので、手作業で穴埋めし、さらに頂点数を 2000 点程度に間引いて使用する。

使用したモーションキャプチャシステムは、赤外線に反射するマーカーを全身に取り付け赤外線を照射し、12 台程度の赤外カメラで捕らえたマーカーの動きから姿勢を計算する。得られるデータは、人体モデルの各ボーンのローカル座標系を、その親ボーンのローカル座標系に変換する回転、平行移動のパラメータである。ただし、人体モデルの親ノードである Body center 以外は関節なので回転成分しか存在せず、平行移動のパラメータは一定である。以下では、ある姿勢 p についての回転のパラメータを姿勢データと呼び、 $r_p = (r_{p1}, \dots, r_{p(3N_b)})$ と表す。ここで、 N_b は人体モデルの関節数である。このマーカーを用いたモーションキャプチャは、画像処理による推定を行う全ての姿勢について行い、類似した冗長な姿勢データを削除する。姿勢 a と姿勢 b の差 d_{1ab} を次のように定義し、 d_{1ab} が一定の閾値内にあるとき、姿勢 a と姿勢 b を似ていると定義する。

$$d_{1ab} = \max_{i=1, \dots, 3N_b} |r_{ai} - r_{bi}| \quad (1)$$

このようにして得られた姿勢データの集合を姿勢辞書とする。

2.3 三次元位置検出

2 台の高速度カメラで撮影した画像は、それぞれ、高速度画像を用いた追跡アルゴリズム [16] によって処理され、画面内に入ってきた人間を追跡する。この追跡アルゴリズムは、画像から抽出した特徴点（コーナー点）を、フレーム間の動きが小さいという高速度画像の特性を利用して、安定に追跡することができる。特徴点の位置の重心を追跡対象の位置として出力するため、位置の検出精度はそれほど高くない（図 3 参照）。

次に、各カメラ c から得られる画像上の重心位置を用いて、追跡対象の三次元位置を計算する方法について述べる。姿勢推

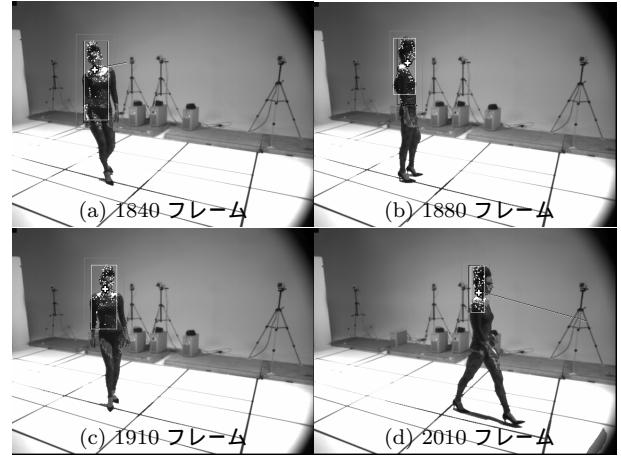


図 3 特徴点追跡の結果（原画像に追跡結果をオーバーレイ表示している。白い四角形は各画像での追跡ウインドウ、白い小さい点は追跡に用いる特徴点（コーナー点）、右方向に伸びている白い線分の長さと方向によって、追跡対象の画像上の運動方向と速度を表している。この線分の左端が重心位置で、白い大きい十字マークで示されている。）

Fig. 3 Result of feature point tracking

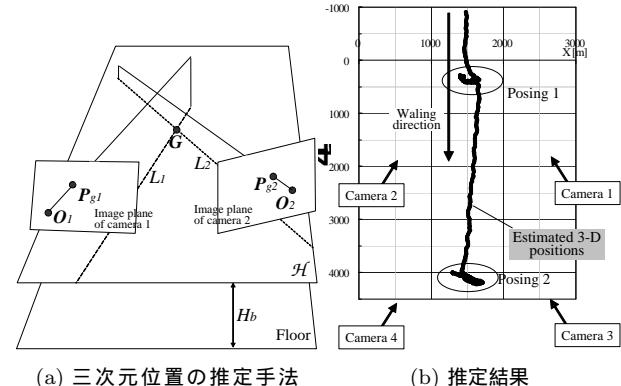


図 4 三次元位置の推定

Fig. 4 Estimation of 3-D position

定を行うのは、歩行、ポージング、ターン等のほぼ直立に近い姿勢であるので、鉛直方向の位置はほぼ一定値 H_b とする。 H_b は、人体モデルの Body center ノードの原点の高さである。床面に平行で床面からの高さが H_b である平面を \mathcal{H} に、各カメラのカメラ中心 O_c と画像平面上の追跡対象の重心 P_{gc} を通る直線を投影し、 L_c とする。直線 L_c の交点を追跡対象の三次元位置とする。ただし、カメラが 3 台以上の場合は、各直線 L_c からの距離の二乗和が最も小さくなる点とする。

図 3 の画像列について、カメラ 4 台で三次元位置の推定を行った結果を図 4(b) に示す。鉛直方向 (Y 方向) は一定なので、 $X-Z$ 平面に結果をプロットしている。約 $X = 1500$ の位置を Z 軸の正の方向に向かって歩いている途中、約 $Z = 500$ 、 $Z = 4000$ の位置で左右に体を振ってポージングを行っているため、 X 座標が左右に変化している。図 3(b)、(c)、(d) は 2 回めのポージングに当たる。

2.4 姿勢推定

あらかじめ作成しておいた姿勢辞書から、前のフレームの姿

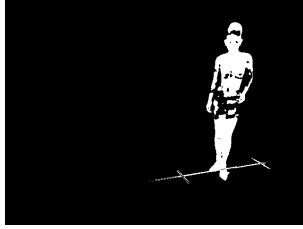


図 5 背景差分による観測シルエット抽出結果

Fig. 5 Silhouette extraction using background subtraction

勢に近い姿勢を姿勢候補として抽出し、候補シルエットを作成する。また、現在の画像とあらかじめ記憶しておいた背景画像を用いて、背景差分によってシルエット画像（観測シルエット）を生成する。次に、候補シルエットと観測シルエットの類似度を計算し、類似度が最も高い姿勢候補を推定姿勢とする。最後に、表示の質を高めるため、推定結果から滑らかな動きを生成する。以下では、各処理の詳細を説明する。

2.4.1 姿勢候補

類似度の計算コストは高く、姿勢辞書にある全ての姿勢について類似度計算を伴う探索を行うのはリアルタイムでは現在のところ不可能であった。そのため、探索の空間を限定して、類似度の計算回数を減らすために、前フレームの姿勢 p に類似している姿勢のみを姿勢候補として抽出する。

姿勢の差には、各関節角の差の最大値である $d1_{ab}$ （式（1））と、各関節位置を画像平面に並行投影した位置の差の最大値である $d2_{ab}$ を用いる。

$$d2_{ab} = \max_{i=1, \dots, N_b} |\mathbf{p}_{ai} - \mathbf{p}_{bi}| \quad (2)$$

ここで、 \mathbf{p}_{ai} 、 \mathbf{p}_{bi} は、それぞれ姿勢 a 、 b の i 番目のボーンのローカル座標系原点を画像平面に並行投影した画像上の点である。前フレームの姿勢 p と姿勢辞書に登録されている姿勢について、これらの差を計算し、共にある閾値より小さい姿勢を、姿勢候補の集合 M とする。

$$M = \{m | d1_{pm} < T_{d1}, d2_{pm} < T_{d2}\} \quad (3)$$

2.4.2 シルエット抽出

候補シルエットは、次のように生成する。（1）姿勢候補集合 M から、シルエットを抽出したい姿勢 m を選択し、（2）人体形状モデルを、姿勢 m の関節角データを用いて変形させ、（3）変形した人体形状モデルの Body center ノードの原点が、観測された三次元位置 G と一致するように人体形状モデルを移動し、（4）人体形状モデルのポリゴンデータを画像平面に投影して、人体モデルが投影された画素値を 1、それ以外を 0 とする。

次に、観測シルエットの抽出は、最も単純な背景差分法を用いて行う。背景となる画像をあらかじめ撮影しておき、現在の画像との差分を計算し、差分が閾値より大きい点の集合をシルエットとして抽出する。シルエット画像は、差分が閾値より大きい点の画素値は 1、それ以外の画素値は 0 である。

2.4.3 類似度

2.3 章で述べたように、三次元位置の検出精度は高くないた

め、位置に関する局所最適化と、姿勢に関する最適化を同時に使う必要がある。カメラ c 、姿勢候補 m について、候補シルエット画像の画素値が 1 である画素の集合を $S_1(c, m)$ 、 $S_1(c, m)$ を (x, y) だけ並行移動した画素集合を $S_1(c, m, x, y)$ 、観測シルエット画像の画素値が 1 である画素の集合を $S_2(c)$ とする。また、姿勢候補集合 M に含まれる全てのシルエットを包含する四角形内部の点の集合を $R(c)$ とする。このとき、類似度 $s(m)$ を次式のように定義する。

$$s(m) = \sum_c \max_{(x, y \in D)} \left(\frac{A(S_1(c, m, x, y) \cap S_2(c))}{A(S_1(c, m, x, y))} - \frac{A(\overline{S_1(c, m, x, y)} \cap R(c) \cap S_2(c))}{A(S_1(c, m, x, y) \cap R(c))} \right) \quad (4)$$

ここで、 $A(\cdot)$ は面積、 D は近傍の画素の集合である。上式では、候補シルエットに含まれる観測シルエットの密度と、候補シルエット外部領域に含まれる観測シルエット密度の差を、1 つのカメラから得られる画像に関する類似度と定義している。

2.4.4 最適姿勢探索

まず、姿勢候補集合 M を用いて、次のように 2 層の探索木を構成する。姿勢候補集合 M の要素は、前のフレームの姿勢に近い順に整列しているので、先頭から順に、間引き間隔 t で第一層目の姿勢を選択する。第二層目は、残りの姿勢で、最も近い第一層目の姿勢に連結する。

最適な姿勢の探索は、この探索木を用いて次のように行う。まず、第一層目の姿勢について、式（4）によって類似度を計算し、類似度の上位 k 個の姿勢を選択する。次に、これら k 個の姿勢に連結されている二層目の類似度を計算する。最後に、類似度計算を行った全ての姿勢の中で、類似度の最も高い姿勢を推定結果とする。

リアルタイムで探索を実行するためには、計算量を一定以下に保つ必要がある。そこで、 M の要素のうち T_n 番目以降の候補姿勢は計算量が一定値以上増えないように計算を行わない。

現在の実装では、 $t = 3$ 、 $k = 3$ 、 $T_n = 40$ である。

2.4.5 運動生成

推定姿勢として出力されるのは、辞書に登録されている姿勢のみなので、推定精度が粗く、また振動もしている。このような推定結果を用いても、4. 章で述べるような映像や動きの品質が重視されるアプリケーションでは、鑑賞に耐えられる品質の映像にはならない。

ステージ上の女性モデルの動きは、姿勢辞書作成に用いたマーカーによるモーションキャプチャの動き $\{p(t)\}$ と、ポージングのタイミングや細かい動作は異なるが、おおまかな動作は同じである。そこで、推定した姿勢に応じて $\{p(t)\}$ の再生スピードを変化させて CG 描画モジュールに送出することによって滑らかな動きを生成した。具体的には、次のような手順で生成した。

（1）CG 描画をはじめる初期姿勢 $p(t_o)$ をあらかじめ決めておき、推定された姿勢 p_e と初期姿勢 $p(t_o)$ の差 $d1$ を式（1）により計算する。

（2） $d1$ が閾値より大きい場合は、出力姿勢 p_o を初期姿勢

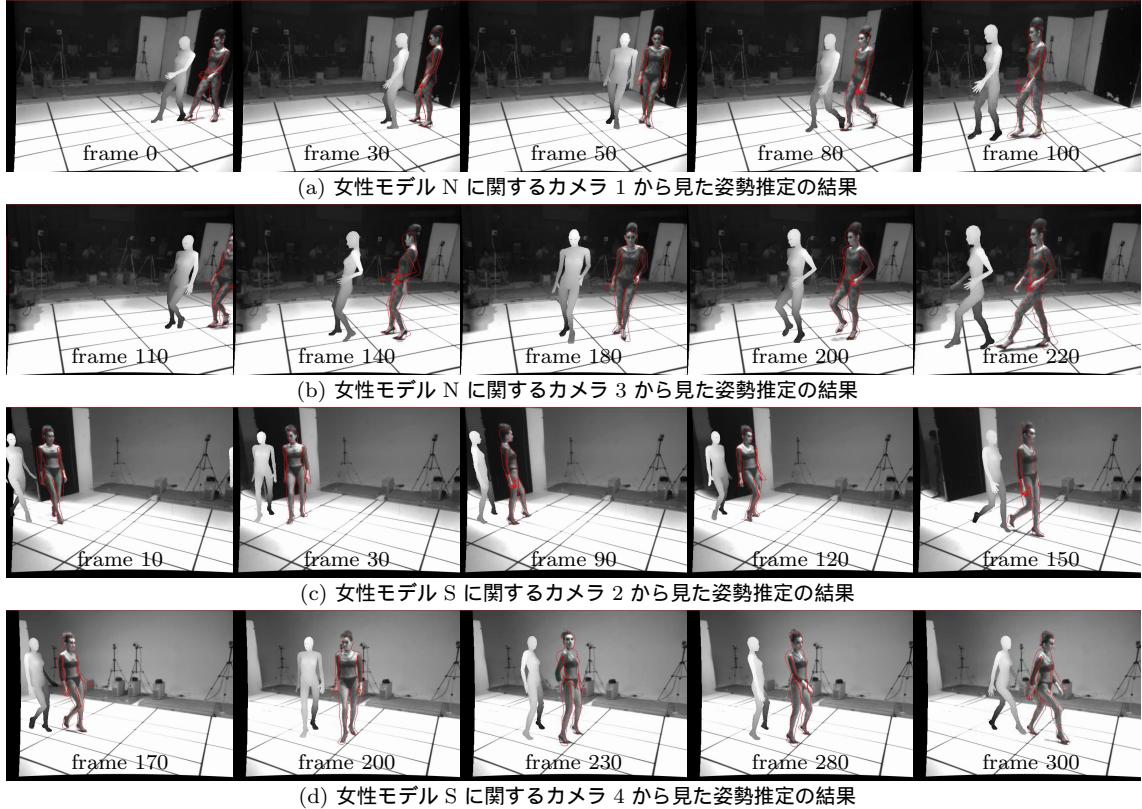


図 6 4 台のカメラを用いた姿勢推定の結果 1(姿勢推定の結果は、赤色の輪郭線で描かれている。
CG 描画用の運動生成を行った結果は、濃淡によって擬似 3D モデルで表示されている。)

$p(t_o)$ として、1 に戻る。 d_1 が閾値より小さくなったら 3 へ。
(3) $\{p(t)\}$ の中で、前の時刻の出力姿勢 p_o と時間的に連続している周辺の姿勢を選択し、そのなかで最も現在の推定姿勢 p_e と最も差が小さい姿勢を現在の時刻の出力姿勢 p_o として出力する。

3. 実験結果

図 6 に姿勢推定の結果を示す。いずれの場合も、カメラの配置は図 4(b) と同様で、4 台の高速度カメラを、女性モデルのウォーキング経路の両脇に 2 台づつ配置して姿勢推定を行っている。同時に使用するのは 2 台のみで、カメラ 1 とカメラ 2 の組、およびカメラ 3 とカメラ 4 のように、2 台一組のカメラを 2 組用いて広い領域をカバーしている。各カメラ組が姿勢推定を行う担当領域は、あらかじめカメラの画角に応じて決定しており、検出された三次元位置が担当領域に入っているカメラ組の画像を使用する。図 6(a)(b) と図 6(c)(d) は、別の女性モデルが異なる動きをしており、前者はウォーキング方向左側の 2 台のカメラの画像、後者は右側の 2 台のカメラの画像である。

姿勢推定は正しく行われているが、姿勢推定の結果を原画像に上書き表示した、人体モデルの輪郭を見ると、腕の角度、足の角度などが若干異なっている場合がある。これは、姿勢辞書の中に現在の姿勢にフィットする最適な姿勢が存在しないことが原因である。

濃淡による擬似 3D 表示のモデルは、前節で述べた CG 描画用の運動生成を行っているので、動きはスムーズである。しか

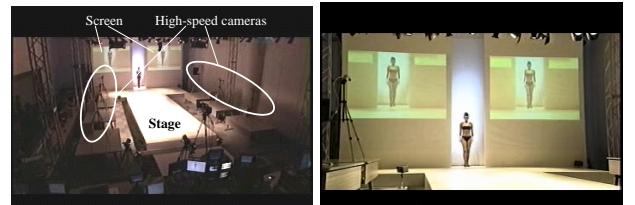


図 7 仮想ファッションショーの概観
Fig. 7 Overview of virtual fashion show

し、姿勢推定そのものに上記のような誤差がある上、運動生成も行っているため、実際の女性モデルの姿勢と重ねるとやはり誤差がある。

4. 仮想ファッションショー

2. 章で述べたモーションキャプチャシステムの実証アプリケーションとして、デジタルファッションショーシステムを試作した。女性のモデルが、ステージ上でウォーキング、ポージングを行い、その動きをモーションキャプチャシステムによって推定する。推定した動きと三次元人体モデルを用い、衣服シミュレーションと CG の技術によって、実際の女性モデルとは異なる衣服を描画して大型スクリーンに投影する。ショーの概観を図 7 に示す。

図 8 にシステム構成とステージ配置を、図 9 に 4 種類のシリオを示す。女性モデルは、黒いタイツを着用してステージをウォーキングし、ステージ先端まで来る。途中 2 度のポージン

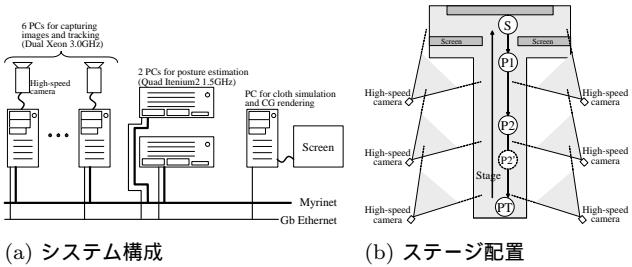


図 8 仮想ファッションショーのシステム構成とステージ配置

Fig. 8 System organization and structure of the stage for virtual fashion show



図 9 仮想ファッションショーで用いた 4 種類のシナリオ (各画像の左側の映像は、モーションキャプチャの結果から生成した衣服をまとった CG 映像である。)

Fig. 9 Four types of scenarios used in our virtual fashion show

グを行うが、ポージングの位置はシナリオによって異なる。シナリオ 1 と 2 は、P1 と P2 の位置でポージングを行い、シナリオ 3 と 4 は P1 と P2' の位置でポージングを行う (図 9 参照)。

6 台の高速度カメラを、P2' のポージングまでモーションキャプチャできるように、ステージ両脇に配置した。高速度カメラは、それぞれ Dual Xeon 3.0GHz 搭載の PC に接続され、この PC は画像入力と同時に 2.3 章の前半で述べた画像上で女性モデルの追跡を行う。撮影された画像と、女性モデルの画像上の位置は、データ通信用の高速ネットワーク Myrinet を通して、Quad Itanium2 1.6GHz 搭載の姿勢推定用マシンに送信される。姿勢推定用の 2 台のマシンでは、2.3 章の後半で述べた三次元位置検出を行い、さらに 2.4 章で述べた姿勢推定を実行する。姿勢推定の結果は、Gb Ethernet を通して CG レンダリング用 PC に送信され、衣服シミュレーションおよびレンダリングを行い、大型スクリーンに CG 映像を表示する。

5. おわりに

2 台のカメラ組を複数用いて、実時間でモーションキャプチャを行うシステムについて述べた。また、実証アプリケーションとして、仮想ファッションショーシステムを試作し、実際に動作することを確かめた。

現在のシステムにはまだ多くの制限があり、実用的には不十分な点がある。今回の実験のように、照明や女性モデルの服装などの条件をコントロールできる環境であれば、背景差分のような単純な手法でもそれほど問題にはならないが、家庭やオフィス等一般的の環境で使用するためには、汎用的な画像特徴と形状モデルの類似度を定義する必要がある。また、現状のシステムでは、女性モデルの歩行と数種類のポージングだけに動きが限定されているが、様々な動きに対応するためには、姿勢の探索空間も大きくなるので、最適姿勢を探索する探索手法も改善の必要がある。さらに、3. 章で述べたように、姿勢推定の誤差はまだ大きく、推定精度の向上も重要な課題の一つである。

文 献

- [1] 株式会社スパイク STT
http://www.mocap.jp/stt/stt_1.html
- [2] 株式会社スパイク GYPSE
http://www.mocap.jp/gypsy/gypsy_1.html
- [3] 株式会社ジースポーツ KROPS
<http://www.gsport.co.jp/products/krops.html>
- [4] 山本正信, “コンピュータビジョン技術による次世代モーションキャプチャ,” 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol.43, No.SIG4(CVIM4), pp.1-9, 2002
- [5] N. Date, H. Yoshimoto, D. Arita, and R. Taniguchi, “Real-time Human Motion Sensing based on Vision-based Inverse Kinematics for Interactive Applications,” Proc. of ICPR, No. 3, pp. 318-321, 2004
- [6] 株式会社スリーディー MY-Motion
[http://www.ddd.co.jp/tech_info/tech_my-motion.htm](http://www.ddd.co.jp/tech_info/tech_mymotion.htm)
- [7] 小川和田俊和, 東海彰吾, 松山隆司, “平面間透視投影を用いた並列視体積交差法,” 情報処理学会 CVIM 研究会論文誌, Vol.42, No.SIG 6(CVIM 2), pp.33-43, 2001.6
- [8] G.K.M. Cheung, S. Baker and T. Kanade, “Shape-From-Silhouette of Articulated Objects and its Use for Human Body Kinematics Estimation and Motion Capture,” Proc. of CVPR, Vol. 1, pages 77-84, 2003
- [9] M. Brand, “Shadow puppetry,” Proc. of ICCV, pp.1237–1244, 1999
- [10] C. Wren, A. Azarbayejani, T. Darrel, and A. Pentland, “Pfinder:Real-time tracking of the human body,” PAMI, 19(7), pp. 780–785, 1997
- [11] D. M. Gavrila and L. S. Davis, “3-D model-based tracking of humans in action: A multi-view approach,” CVPR, pp.73–80, 1996
- [12] Hedvig Sidenbladh and Michael J. Black and David J. Fleet, “Stochastic Tracking of 3D Human Figures Using 2D Image Motion,” Proc. of ECCV, pp. 702-718, 2000
- [13] Michael Isard, Andrew Blake: ICONDENSATION: Unifying Low-Level and High-Level Tracking in a Stochastic Framework. ECCV (1), pp. 893–908, 1998
- [14] Roger Y. Tsai, ”An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision”, Proc. of CVPR, pages 364-374, 1986
- [15] Roger Y. Tsai, ”A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses”, IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, pp. 323-344, 1987
- [16] R. Okada, et al., ”High-speed Object Tracking in Ordinary Surroundings Based on Temporary Evaluated Optical Flow,” Proc. of IROS, pp. 242 – 247, 2003