

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4695275号
(P4695275)

(45) 発行日 平成23年6月8日(2011.6.8)

(24) 登録日 平成23年3月4日(2011.3.4)

(51) Int.Cl. F I
G06T 13/00 (2011.01) G O 6 T 13/00 B
G06T 13/80 (2011.01) G O 6 T 15/00 I O O A
G06T 15/00 (2011.01)

請求項の数 6 (全 20 頁)

<p>(21) 出願番号 特願2001-63854 (P2001-63854) (22) 出願日 平成13年3月7日(2001.3.7) (65) 公開番号 特開2002-269580 (P2002-269580A) (43) 公開日 平成14年9月20日(2002.9.20) 審査請求日 平成19年11月27日(2007.11.27)</p> <p>前置審査</p>	<p>(73) 特許権者 503360115 独立行政法人科学技術振興機構 埼玉県川口市本町四丁目1番8号 (74) 代理人 100105371 弁理士 加古 進 (72) 発明者 星野 准一 茨城県つくば市春日一丁目6-2-501</p> <p>審査官 田中 幸雄</p>
---	---

最終頁に続く

(54) 【発明の名称】 動画像生成システム

(57) 【特許請求の範囲】

【請求項1】

動画像生成システムであって、

人体の3次元モデルのシルエット画像と、該シルエット画像と対応して人体の3次元モデルの姿勢パラメータとを複数記憶している記憶手段と、

2次元動画像を取得する2次元動画像取得手段と、

該2次元動画像取得手段からの2次元動画像内の人物と該記憶手段に記憶している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シルエット画像を選択し、該シルエット画像に対応する3次元モデルの姿勢パラメータから各パーツの中心線を推定することにより、3次元人体モデルの初期姿勢の姿勢パラメータを推定する初期姿勢推定手段と

10

、
 前記取得した2次元動画像内の人物から、人体の3次元モデルを構成する各パーツ(オブジェクト)の動作パラメータを得る動作推定手段と、

前記2次元動画像内の人物より3次元モデルの各パーツと対応するテクスチャデータを抽出するテクスチャ抽出手段と、

前記3次元モデルの各パーツの動作パラメータを用いて、初期姿勢から3次元モデルを生成する3次元モデル生成手段と、

前記生成した3次元モデルから、視点を設定して2次元の動画像を生成し、前記生成した新たな3次元モデルにおいて、前記抽出したテクスチャデータでは不足する2次元動画像部分は、前記抽出したテクスチャデータより補間して得る動画像生成手段と

20

を備えることを特徴とする動画像生成システム。

【請求項 2】

請求項 1 に記載の動画像生成システムにおいて、

前記 3 次元モデル生成手段は、前記動作推定手段で得た動作パラメータを変更する動作変更手段をさらに備え、取得した 2 次元動画像の人物とは異なる動きの動画像を生成することを特徴とする動画像生成システム。

【請求項 3】

請求項 1 又は 2 に記載の動画像生成システムにおいて、

前記 3 次元モデル生成手段は、さらに、前記生成した 3 次元モデルに、他のオブジェクトを合成することを特徴とする動画像生成システム。

10

【請求項 4】

請求項 3 に記載の動画像生成システムにおいて、

前記 3 次元モデル生成手段の前記他のオブジェクトは、衣装モデルから前記 3 次元モデルの動作パラメータを用いてフィッティングされた衣装であり、

前記動画像生成手段は、前記 3 次元モデルと前記衣装とを同時にレンダリングして合成画像を得る

ことを特徴とする動画像生成システム。

【請求項 5】

請求項 1 ~ 4 のいずれかに記載の動画像生成システムをコンピュータ・システムに構成させるコンピュータ・プログラム。

20

【請求項 6】

請求項 1 ~ 4 のいずれかに記載の動画像生成システムをコンピュータ・システムに構成させるコンピュータ・プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、2次元のビデオ映像から、リアルな人物 3次元CGを容易に生成することができるシステムであり、特にビデオ映像から生成した人物の3次元モデルを用いて、リアルな動画像を生成できる人物動画像生成システムに関する。

【0002】

30

【背景技術】

今日、映画の特殊効果の生成や、ビデオゲームの開発では、リアルな人物CGの生成が重要な課題である。ところが、人物のように複雑な被写体のCGを生成する作業は、ほとんどがアニメーターの手作業で行われていて、非常にコストが高いという問題がある。

近年では、実際に存在する俳優のCGが必要とされることも多い。例えば、映画の制作現場では、危険なスタント・シーンを撮影するときに、代役で撮影した映像を加工して、主演している俳優が演技をしているように見せることが必要となる。また、商業フィルム(CM)では、過去の有名な人物を登場させて、他の俳優と競演する場面が増えている。ところが、従来手法では、既に撮影したフィルムから手作業で映像を切り出してつなぎ合わせる程度しかできないため、制約が大きいだけでなく、作業に必要な労力も膨大である。

40

また、ビデオ編集においても、より自由度の高い人物画像の編集が望まれている。例えば、既に撮影した映像中の人物を、異なるカメラ・アングルから見た場合に変更したり、衣服や髪型の変更、姿勢の変更を行うことができれば、ポスト・プロダクションの作業をより効果的に行うことができる。

【0003】

これらのようにリアルな映像を容易に生成する手法としては、CGの分野でイメージ・ベース・レンダリング手法が提案されている。この手法では、実写画像に3次元モデルを当てはめることで、自由な視点位置から見た映像を生成する。ところが、従来手法では、屋外の建築物などの単純な形状を持つ物体や、顔のように単一の物体を対象としている。そ

50

のため、人物のような複雑な動作を行う被写体へ適用することができないという問題がある。

自由な視点位置から画像を生成する手法については、コンピュータ・ビジョンの分野でも、Virtualized Reality（仮想化された現実）とも呼ばれて盛んに研究が行われている。これらの手法では、撮影環境内に多数のカメラを設置して、ステレオ法などの手段により3次元モデルを生成する。ところが、正確な3次元モデルを推定するためには、多数のカメラが必要となり、コストが高くなるという問題がある。また、オクルージョンによって見えない部分については、3次元形状を復元することができないという問題がある。

【0004】

【発明が解決しようとする課題】

本発明の目的は、ビデオ画像より復元した人体の3次元モデルを利用した動画像生成処理の適用により、リアルな人物CGを容易に生成することである。

【0005】

【課題を解決するための手段】

上記の目的を達成するために、本発明は、動画像生成システムであって、人体の3次元モデルのシルエット画像と、該シルエット画像と対応して人体の3次元モデルの姿勢パラメータとを複数記憶している記憶手段と、2次元動画像を取得する2次元動画像取得手段と、該2次元動画像取得手段からの2次元動画像内の人物と該記憶手段に記憶している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シルエット画像を選択し、該シルエット画像に対応する3次元モデルの姿勢パラメータから各パーツの中心線を推定することにより、3次元人体モデルの初期姿勢の姿勢パラメータを推定する初期姿勢推定手段と、前記取得した2次元動画像内の人物から、人体の3次元モデルを構成する各パーツ（オブジェクト）の動作パラメータを得る動作推定手段と、前記2次元動画像内の人物より3次元モデルの各パーツに対応するテクスチャ・データを抽出するテクスチャ抽出手段と、前記3次元モデルの各パーツの動作パラメータを用いて、初期姿勢から3次元モデルを生成する3次元モデル生成手段と、前記生成した3次元モデルから、視点を設定して2次元の動画像を生成し、前記生成した新たな3次元モデルにおいて、前記抽出したテクスチャ・データでは不足する2次元動画像部分は、前記抽出したテクスチャ・データより補間して得る動画像生成手段とを備える。これにより、生成した人体モデルとテクスチャを利用して、新たな視点位置を設定することにより別の視点位置から見た人物画像も生成することができる。

【0006】

前記3次元モデル生成手段は、前記動作推定手段で得た動作パラメータを変更する動作変更手段をさらに備え、取得した2次元動画像の人物とは異なる動きの動画像を生成することもできる。

前記3次元モデル生成手段は、前記生成した3次元モデルに他のオブジェクトを合成することができる。

例えば、前記3次元モデル生成手段の前記他のオブジェクトは、衣装モデルから前記3次元モデルの動作パラメータを用いてフィッティングされた衣装であり、前記動画像生成手段は、前記3次元モデルと前記衣装とを同時にレンダリングして合成画像を得ることができる。

本動画像生成システムをコンピュータ・システムに構成させるコンピュータ・プログラムおよびコンピュータ・プログラムを記録した記録媒体も本発明である。

【0007】

【発明の実施の形態】

<本発明の概要>

本発明は人体の3次元モデルを利用した動画像処理の適用により、リアルな人物CGを容易に生成する人物画像生成システムである。

図1は本発明を用いた実施形態の一例である、モデルに基づく人物動画像生成システムの機能構成図を示す。図1に示すように、本システムは、テレビカメラにより人物の動画像

10

20

30

40

50

を連続的に取得する画像取得部 10、人物の各関節のパラメータを推定したり、顔や衣服などの人物テクスチャを抽出する動作推定部 20、後述する人体モデルを利用した隠蔽箇所の復元や、異なる視点位置から見た画像の生成を行ったり、衣服 CG や髪型 CG の合成や、姿勢の変更などを行う画像生成部 40、そして生成した結果を表示する画像表示部 50 で構成している。

【0008】

このような構成の本システムでは、画像取得部 10 で取得した人物動画像に、あらかじめシステムに用意してある人体の 3 次元モデルを当てはめて、動作推定部 20 において、各関節の 3 次元動作を推定すると共に、顔や衣服などの輝度情報を人体モデルのテクスチャとして蓄積する。また、画像生成部 40 において、撮影した視点位置から見たときに隠蔽されている箇所を、人体モデルを利用した CG の生成で補うことで、全身像を復元する。このようにして得た画像を画像表示部 50 から表示している。

図 2 ~ 図 4 は、本発明における処理の概要を分かりやすく説明するために、ビデオ映像に上述の処理を適用した例を示している。図 2 (a) は画像取得部 10 で取得した入力画像の例である。図 2 (b) は動作推定部 20 で人体モデルの姿勢を推定した例である。図 3 (c) は、画像生成部 40 で人体の 3 次元モデルにテクスチャ・マッピングをした人体画像を、自由な視点位置から見た例である。そしてこのとき、画像生成部 40 は撮影した視点位置から見たときに隠蔽されている箇所を、人体モデルを利用した CG の生成で補っている。図 3 (d) は前述の CG の生成で補った画像の例を示す図である。

【0009】

このように、本システムにおいては、生成した人体モデルとテクスチャを利用して、自由な視点位置から見た人物画像を生成したり、入力画像とは異なる動作の人物画像を生成したり、人物のような複雑な被写体に、CG による別の衣服や髪型の合成を行うことができる。図 4 (e) は衣服の CG を合成した例、図 4 (f) は異なる姿勢へ変更した例である。

また、多様な姿勢を持つ人物画像を、関節角の変更データを用いて生成できるため、動画画像圧縮の手段としても利用することができる。

以上のような処理を行うシステムを以下に詳細に説明する。

【0010】

<モデルに基づく人物動作の推定>

動作推定部 20 におけるモデルに基づく人物動作を、動画像を構成する連続した画像 (フレーム) 間から推定する処理を以下に詳細に説明する。

(移動量の推定)

動作推定部 20 においては、人体の 3 次元モデルを当てはめて、各関節の 3 次元動作を推定している。図 5 に、動作推定部 20 で用いている人体モデル 300 を示す。図 5 のように、人体モデル 300 は、各パーツをオブジェクトで表し、それらを接続することで全体を表している。この人体モデル 300 は、頭部 310、胴体 315、右上腕 322、左上腕 324、右前腕 332、左前腕 334、右腿 342、左腿 344、右膝下 352、左膝下 354 の各パーツ (各オブジェクト) で構成している。動作推定部 20 では、このような人体モデル 300 を利用してフレーム間の人物動作を推定する。人体モデル 300 の各パーツ (各オブジェクト) の接続関係は、胴体 315 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の接続として定めることができる。親パーツが動くと、その子に当たるパーツは親パーツの座標系で規定される接続関係で一緒に動くことになる。この人体モデル 300 を用いて、以下で動作について説明を行う。

なお、この人体モデルは標準形をシステムで用意して、ユーザが対象となる人物に合わせて調整することで得ることができる。

図 6 は、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図である。この図において、動画像はカメラ座標系で記述される画像面 410 に表示される。人体モデル 300 上の任意の点を、ワールド座標系を用いて、

【数 1】

$$p = (x, y, z)^T$$

とする（上式で括弧の肩に付けられている T は転置行列を表す）と、カメラ座標系に投影した画像面 4 1 0 上の点 (X, Y) は、

【数 2】

$$X = f x / z$$

$$Y = f y / z \quad (f \text{ は画像面 } 4 \ 1 \ 0 \text{ を生成するカメラの焦点距離})$$

となる。同様に、人体モデル 3 0 0 上にある任意の点 p に関する 3 次元速度ベクトルを、

【数 3】

$$\dot{p} = (u, v, w)^T$$

10

とすると、このベクトルはフレーム間での点 p の移動速度を表している。画像面 4 1 0 上の速度ベクトル (U, V) とすると、両者の関係は、

【数 4】

$$U = (f u - X w) / z$$

$$V = (f v - Y w) / z$$

となる。

【0 0 1 1】

また、画像面 4 1 0 上の点 (X, Y) における時刻 t での輝度を E (X, Y, t) とする。そして、(E_x, E_y) を輝度微分、E_t を時間微分とすると、画像面 4 1 0 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。

20

【数 5】

$$E_x U + E_y V + E_t = 0$$

この式に上述の U, V を代入すると、次式のように人体モデル 3 0 0 上の任意の点 p と、画像面 4 1 0 上の点 (X, Y) および輝度値 E_t の関係を示すことができる。

【数 6】

$$G \dot{p} = -E_t$$

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

30

【0 0 1 2】

次に、人体モデル 3 0 0 の全関節角の位置と姿勢を表すベクトル s を

【数 7】

$$s = (s_1, s_2, \dots, s_n)$$

として、s と p の関係を便宜的に非線形関数 g を用いて

【数 8】

$$p = g(s)$$

で表す。関数 g は非線形であるが、s の時間変化率と速度ベクトルの関係は、ヤコビ行列 J (s) = g / s により、線形式で関係付けることができる。

40

【数 9】

$$\dot{p} = J(s) \dot{s}$$

よって、人体モデル 3 0 0 の全関節 s の時間変化率と、画像面 4 1 0 上の点の輝度値の関係式を導くことができる。

【数 10】

$$G J(s) \dot{s} = -E_t$$

上述では姿勢パラメータ s と人体モデル上の任意の点 p の関係を p = g (s) で表した。

50

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。以下では、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。

【0013】

(胴体のヤコビ行列の導出)

人体モデルの運動は、物体座標系における剛体変換 ε_i で表すことができる。3×3の回転行列をC、3次元の並行移動ベクトルをDとすると、剛体変換 ε_i は、次式のように表すことができる。

【数11】

$$\varepsilon_i = \begin{bmatrix} C_i & D_i \\ 0^t & 1 \end{bmatrix}$$

10

ただし、 0^t はゼロの転置行列を表す。また、物体座標の点

【数12】

$o_i = (x_{o_i}, y_{o_i}, z_{o_i}, 1)^T$
に剛体変換 ε_i を施す場合は、次式で表すことができる。

【数13】

$$o_i' = \varepsilon_i o_i$$

【0014】

20

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 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 s を、ワールド座標系に変換する操作を、剛体変換 ε_s で表す。

【数14】

$$w = \varepsilon_s s$$

2) 次に、胴体の運動を適用すると、胴体モデル上の点 w が w' に移動する。 w' を求めるためには、ワールド座標系の点 w を物体座標系に変換して、胴体の運動 ε_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

30

【数15】

$F_i = \varepsilon_i^{-1} \varepsilon_i^{-1} o_i^{-1}$
とおくと、胴体モデル上の点 w' は次式で求めることができる。

【数16】

$$w' = F_i w = F_i \varepsilon_s s$$

3) 最後に、ワールド座標系の点 w' をカメラ座標系に変換する。

【数17】

$s' = \varepsilon_s^{-1} w' = \varepsilon_s^{-1} F_i \varepsilon_s s$
上記の P_s' を時間微分することで、下記のヤコビ行列が得られる。

【数18】

40

$$\frac{d\rho_s'}{dt} = J(s)\dot{s}$$

【0015】

(その他の関節のヤコビ行列の導出)

上記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。

50

1) 上脚, 上腕, 頭の運動

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

【数19】

$w = s_j s$
この場合は、子の移動に伴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 $F_i = [\quad o_i \quad i \quad o_i^{-1}]$ を適用する。

【数20】

$s_j'' = s_j^{-1} F_1 F_i s_j s$
そして胴体のヤコビ行列の導出と同様に s' を時間微分することで、ヤコビ行列を求める。

10

【0016】

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、更に胴体を親として持っている。この場合も、胴体だけを親とする、依存関係が1段階の場合（上脚、上腕、頭）と同様な変換規則を適用することで次式が得られる。

【数21】

$s_j''' = s_j^{-1} F_1 F_{i1} F_i s_j s$
そしてこれまでのヤコビ行列の導出と同様に s' を時間微分することで、ヤコビ行列を求める。

20

【0017】

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。

【数22】

$$\rho_{sj}^n = \varepsilon_{sj}^{-1} F_1 \left[\prod_{m=0}^{n-2} F_{i-m} \right] \varepsilon_{sj} \rho_s$$

以上のヤコビ行列を利用して、上述の画像面410上の点の輝度値の関係式を人体モデル300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 s を求める。そして以下で説明する初期姿勢に、上記で求めた動作パラメータを加算することで、各フレームの人体モデル300の姿勢を得る。

30

【0018】

<初期姿勢の推定>

上述の動作推定法ではフレーム間の動きしか推定することができないため、人体モデル300の初期姿勢を推定することが必要となる。以下では、拡張カルマン・フィルタにより、シルエット画像から抽出した中心線と、人体モデルの体軸を一致させる手法を述べる。ワールド座標系での人体モデル300の各パーツの姿勢 a を表した式を以下に示す。

【数23】

$$a = [r, d]^T$$

40

d はワールド座標の原点からの移動成分を表すベクトルである。また、 r はワールド座標の各座標軸である回転軸に沿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。

【0019】

図7は、画面上の線分の関係と、それに対応する人体モデル300中のパーツとの関係を示した図である。3次元(3D)線分 P の投影された画像面410上の2次元(2D)予測線分が Q である。図7における2次元(2D)観測線分 Q' と2D予測線分 Q を、2次元画面上410での始点および終点を表わすワールド座標系のベクトル q_1', q_2', q_1 および q_2 で表わした式を以下に示す。

【数24】

$$Q' = (q_1', q_2')$$

50

$$Q = (q_1, q_2)$$

また、3D線分Pをワールド座標系での線分の始点および終点を表わすベクトル p_1 および p_2 で表わした式を以下に示す。

【数25】

$$P = (p_1, p_2)$$

そして、初期位置で観測された2D観測線分 Q' と3D線分Pの組を l として、以下の式に示す。

【数26】

$$l = (P, Q')$$

この3D線分Pと、それに対応する画像面410上での2D観測線分 Q' の組 l を求める手法を以下に説明する。図8は入力した動画像より人体モデルの初期位置を推定するまでの処理の過程を示す図である。この図8を参照しながら以下に説明をする。

10

【0020】

(中心線の推定)

画像取得部10より取得した画像の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、本発明のシステムでは、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良い。以下の手順で中心線の推定を行う。

【0021】

1) 中心線が抽出しやすい姿勢のシルエット画像を予め選択してシステムに記憶しておく。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。姿勢パラメータは後述する手法により姿勢を微調整することができるので、大まかな値で構わない。

20

2) 図8(a)(i)および(ii)のような、画像取得部10より取得した低解像度の入力画像系列で、あらかじめシステムに記憶してある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、画像面410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。図8(b)(i)および(ii)は入力画像に人体モデルを当てはめ、各パーツの中心線を抽出し、表示した画像を示す。

30

【0022】

上記の手法より、抽出した中心線を利用して、3D線分Pと、それに対応する画像面410上での2D観測線分 Q' の組 l を求めることができ、後述する手法により人体モデルの正確な初期位置の推定を行うことができる。

【0023】

(初期姿勢の調整)

各パーツの姿勢 a が正確な場合は、観測画像から得られた画像面410上の2D観測線分 Q' とカメラの原点Oとが作る平面Mの上に3D線分Pが含まれる。ところが各パーツの姿勢 a に誤差がある場合は、図7(a)のように、画像面410上に表示される線分 Q' とカメラの原点Oとが作る平面Mと3D線分Pの間に距離が生じる。この距離の最短距離を、予測誤差 $h(a, l)$ とし、以下の式のように表す。

40

【数27】

$$h(a, l) = \left[\frac{(Rp_1 + d) \cdot n}{(Rp_2 + d) \cdot n} \right]$$

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

【数28】

$$n = \frac{q'_1 \times q'_2}{\|q'_1 \times q'_3\|}$$

である。a に誤差がない場合は、 $h(a, l) = 0$ となり、図 7 (b) のように画像面 410 上の 2D 予測線分 Q が 2D 観測線分 Q' と重なる。そして 2D 観測線分 Q' とカメラの原点が作る平面の上に 3D 線分 P が含まれるようになる。このときの 3D 線分 P の姿勢 a が画像面 410 上に表示されている 2D 観測線分 Q' に対応しているということになる。 $h(a, l)$ は非線形なので、観測値 $l = \hat{l}_i$ と推定値 $a = \hat{a}_{i-1}$ のまわりに 1 次テーラ-展開をして線形近似する。

10

【数 29】

$$h(a_i, l_i) = \frac{\partial h}{\partial a}(a_i, l_i)(a - a_i) + \frac{\partial h}{\partial l}(a_i, l_i)(l - l_i)$$

ただし h/a 、 h/l は偏微分を表す。これらの偏微分係数は、以下のカルマン・フィルタを適用して人体モデルの各パーツの姿勢 $a = [r, d]^T$ の最適値を求める際に利用する。

【0024】

対応付けがついた 2D 観測線分 Q' と 3D 線分 P の誤差 $h(a, l)$ が最小になるように、人体姿勢 a を補正し、位置の誤差分布 x_i を減少させる。直前に示した $h(a, l_i)$ 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。

20

【数 30】

$$z_i = H_i a + v_i$$

ただし、

【数 31】

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

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

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

30

である。ここで v_i の分散 B_i は次のように求める。

【数 32】

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

40

【0025】

以上の観測方程式より、人体パーツの姿勢 x_i の 2 乗誤差を最小にする推定値は、カルマン・フィルタで与えられる。

【数 33】

$$\hat{a}_i = \hat{a}_{i-1} + K_i [z_i - H_i \hat{a}_{i-1}]$$

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

$$\sigma_i = \sigma_{i-1} - K_i H_i \sigma_{i-1}$$

人体姿勢 a_i と誤差分布 σ_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。本発明のシステムでは、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、姿勢の推定を行う。

10

このようにして、図8(a)(i)および(ii)の入力画像に初期姿勢を推定し、図8(c)(i)および(ii)のように人体モデルを重ね合わせた画像を得ることができる。

【0026】

(2D、3D線分の対応付け)

上述の手法では、シルエット画像のマッチングにより、抽出した中心線を利用して、3D線分Pと対応する2D観測線分Q'の組1が求まったとして説明した。しかし、撮影環境内に中心線として抽出可能な複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となる。そこで以下に示す2つの手法を併用して対応付けを行う。まず2D画像面410上で見たときに、2D観測線分Q'に射影したときに射影像が2D観測線分Q'と重な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。

20

【0027】

1) 2次元画面上での対応付け

図7を用いて2次元画像面410上で対応付けを行う手法を示す。まず、図7(a)のように、3D線分Pを画面上に投影して、2D予測線分Q=(q₁, q₂)を求める。また、観測画像から2D観測線分Q'=(q₁', q₂')を得る。2D観測線分Q'に対して射影した2D予測線分Qの射影像が重複領域をもつ場合は、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、2D予測線分Qを2D観測線分Q'に射影したときに、q₁あるいはq₂のどちらかの点がQ'の内部にあれば、2D予測線分Qの射影像が2D観測線分Q'に重なっていると考えることができる。まず、q₁をQ'に射影した場合の座標cは、Q'に沿う単位方向ベクトルを

30

【数34】

$m = N [q_2' - q_1']$ (N[]はベクトルの正規化を表わす)

とすると、

【数35】

$c = q_1' + m \cdot (q_1 - q_1')$

である。そして、cがQ'の内部に含まれる条件は、

【数36】

$(c - q_1') \cdot (c - q_2') < 0$

40

となる。この操作をq₁, q₂, q₁', q₂'に関して行い、上記の条件を満たさない線分は棄却する。

【0028】

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 σ_{x_i} と、2D、3D線分の誤差分布 σ_{l_i} を使って χ^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-3D線分が一致していると仮定して、仮説検定を行う。人体姿勢 $a = [r, d]^T$ と観測値 $l = (P, Q')$ の誤差分布を、正規分布 $a_i \sim N(0, P_0)$, $l_i \sim N(0, L_0)$ と仮定する。h(\hat{a}_{i-1} , \hat{l}_i)の1次、2次統計量は、

【数37】

50

$$E \left[h \left(\hat{a}_{i-1}, \hat{l}_i \right) \right] = 0$$

【数 3 8】

$$E \left[h \left(\hat{a}_{i-1}, \hat{l}_i \right) \right] = Z_i = \frac{\partial h \left(\hat{a}_{i-1}, \hat{l}_i \right)}{\partial a} P_{i-1} \frac{\partial h \left(\hat{a}_{i-1}, \hat{l}_i \right)^T}{\partial a} + \frac{\partial h \left(\hat{a}_{i-1}, \hat{l}_i \right)}{\partial l} L_{i-1} \frac{\partial h \left(\hat{a}_{i-1}, \hat{l}_i \right)^T}{\partial l}$$

となる。これよりマハラノビス距離 k を次のように求めることができる。

【数 3 9】

$$k \left(\hat{a}_{i-1}, \hat{l}_i \right) = h \left(\hat{a}_{i-1}, \hat{l}_i \right)^T Z_i^{-1} h \left(\hat{a}_{i-1}, \hat{l}_i \right)$$

10

$h \left(\hat{a}_{i-1}, \hat{l}_i \right)$ の誤差分布は $h \left(\hat{a}_{i-1}, \hat{l}_i \right) \sim N \left(0, Z_i \right)$ であるため、 $k \left(\hat{a}_{i-1}, \hat{l}_i \right) \sim \chi^2 \left(\text{Rank} \left(Z_i \right) \right)$ となる（ただし $\text{Rank} \left(Z_i \right) = \text{Rank} \left(Z_i \right)$ ）。 $\chi^2 \left(\text{Rank} \left(Z_i \right) \right)$ 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。例えば $k \left(\hat{a}_{i-1}, \hat{l}_i \right) < 9.21$ は、 Z_i がランク 2（ $\text{Rank} \left(Z_i \right) = 2$ ）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、 $k \left(\hat{a}_{i-1}, \hat{l}_i \right)$ 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、2D 予測線分 Q が 2D 観測線分 Q' に一番近い 3D 線分 P と画像面 410 上で対応する 2D 観測線分 Q' の組 $l = (P, Q')$ を求めることができる。

20

なお、上述の人体モデルの初期姿勢は、ユーザーが人体モデルの姿勢を推定して画像面に投影し、入力画像に重ね合わせて当てはめることにより求めてもよい。

【0029】

< 隠蔽箇所の復元 >

画像生成部 40 で行っている隠蔽箇所の復元処理を、以下に詳細に説明する。

図 9 は上述の手法で、人体モデルにテクスチャ・マッピングを行い、撮影位置とは異なる視点位置から見た例を示す図である。図 9 (a) の 2 次元の画像より、人物の姿勢を推定することができる。この姿勢における人体パーツの 3 次元モデル上の輝度分布を求めることができる。顔、皮膚、衣服等の輝度情報を、3 次元モデル上のテクスチャとして蓄積する。このように人物テクスチャをマッピングした 3 次元モデルから、従来の CG モデルをレンダリングするのと同じ手法で、自由な視点位置から見た画像を生成することができる。

30

【0030】

しかし、図 9 (b) から分かるように、1 台のテレビカメラで撮影した画像からでは、人物の全身像の輝度情報を得ることができないという問題が起こる。例えば、図 9 (a) のようにカメラに対して前に向いている人物を撮影した場合は、人物の背後が隠蔽されていて撮影できない。そのため、図 9 (b) のように異なる方向から見た場合に、全身像を表示することができない。

この問題を解決するため、画像生成部 40 では、人体モデルを利用した復元手法を用いている。隠蔽部分を復元する手法としてはいくつかあるが、例として 2 つの手法を説明する。

40

【0031】

1 入力画像から得られた輝度情報の利用による復元

入力画像から得られた輝度情報を利用する方法としては、様々なものがあるが、ここでは腕や足などの人体パーツの対称性を利用した方法を述べる。

図 10 は、入力された画像中の人体パーツの対称性から輝度情報を補間する手法を説明するための図である。図 10 に示すように、体軸を垂直に切ってみたとき（図 10 (a) 参照）に、画像から得られた輝度値で点対称の位置にある隠蔽部分の輝度値を埋める（図 10 (b) 参照）ことで補間して全体の輝度値を得る（図 10 (c) 参照）。

50

このときに、点対称の補間では求まらず、隙間になっている箇所や、部分的に空白になっているところは、近傍の輝度値を参照して、スプライン補間などの手法で輝度値を補間する。

なお、より高度な手法として、入力画像から洋服の模様（テクスチャ）を分析して、隠蔽部分に張り付ける等の補間とすることもできる。

【 0 0 3 2 】

2 CGの生成による復元

頭髪のように形状が予測できるものに関してはCGで補間することができる。頭髪の色は、画像中の輝度値を参照して決定する。図11はCGで頭髪を生成した例を示す図である。図11(a)は、人体モデルに頭髪を張り付けたところである。この頭髪に重力をかけるという計算をすることで、図11(b)のように頭髪を頭部にフィッティングすることができる。図11(c)は人物が歩いた場合の頭髪の動的な変化をシミュレートした例を示したものである。

10

【 0 0 3 3 】

<人物画像とCG画像の合成>

画像生成部40で行っている、人物とCG画像の合成処理について、以下に詳細に説明する。画像生成部40では、推定した人物の3次元姿勢を利用してCG画像を生成することができるため、入力画像とは異なる衣服CGや頭髪CGを人物と合成することも可能である。

図12はCGによる衣服の例を示す図である。以下に、図12に示すような衣服CGを人物と合成する例で説明する。

20

図12(a)は衣服を生成したところである。この衣服を人体の3次元モデルに衣服モデルを着せる(重ね合わせる)。まず、モデルと衣服画像の衝突計算を行うことで、図12(a)で示すような生成した衣服画像から、人体モデルの形状に合わせた衣服画像を生成する。衣服と人体モデルの衝突計算は、従来から提案されている手法を利用する。その結果、図12(b)のように衣服画像が人物の3次元形状にフィッティングする。次にフレーム間の人体の動作に合わせて、衣服のダイナミックな動きをシミュレートする。人体の動作情報は、上述の手法で推定した値を利用する。また、衣服画像のシミュレーション手法は、従来から提案されている手法を利用する。図12(c)は、歩行動作に合わせて衣服シミュレーションを行った例である。

30

このようにして得られた衣服モデルと人体モデルを同時にレンダリングすることで、合成した映像が得られる。図13は衣服CGを合成した例を示す図である。図13(b)は、上記の処理を行って図13(a)のような入力画像の人物に衣服CGを合成した画像を示す図である。

【 0 0 3 4 】

<モデルに基づく人物動画像の編集>

これまで説明したように、画像生成部40では、人物の動作推定と隠蔽箇所の復元をする手法により、人体の3次元モデルと、顔、衣服などの人物テクスチャを得ることができる。

その上、画像生成部40では、人体モデルの各関節の角度を変更するだけで、様々な姿勢を持った人物画像を生成することができる。これについて、以下に詳しく説明する。

40

【 0 0 3 5 】

(異なる姿勢の生成)

図14は人体モデルを動かして異なる姿勢に変更した例を示す図である。図14(a)は、立っている人物画像から得られる人体モデルである。このような入力画像から得られる人体モデルの各関節の角度を変更することにより、例えば、左手を前に突き出したようにした図14(b)や、左手をあげて膝をついているようにした図14(c)の画像を生成することができる。図15は入力画像から得られた人物の人体モデルの姿勢を変形したものを示す図である。図15(a)のような人物画像を入力し、この画像から得た人体モデルの関節の角度を変える。関節の角度を変えた人体モデルにテクスチャをマッピングし、

50

最終的に図 15 (b) のように異なる姿勢の人物画像に変更できる。

【 0 0 3 6 】

(動作の物理特性の変更)

人体モデルに物理特性を変更した動きを与えることで、入力画像とは異なる動きをしている人物動画像を生成することができる。動作の物理特性の変更は、人体モデルの質量や慣性モーメントを変更して、物理シミュレーションを行うことで得られる。例えば、映像中の人物が歩行動作をしている場合、歩いている速さを変えたり、重い荷物を持っているような動作に変更することに利用する。

【 0 0 3 7 】

(異なる動き情報の当てはめ)

人体モデルの関節角を動かすときに、異なる人物の動きを計測した結果を当てはめることもできる。例えば、あらかじめマリリン・モンローの歩き方を映画の中から計測しておく。そして、その計測した歩き方を図 15 に示すような入力した女性の映像に対して適用することにより、例えばマリリン・モンローの歩き方で動かすことができる。

【 0 0 3 8 】

(部分的な人物テクスチャの張り替え)

本手法では、顔、衣服などの人物テクスチャとの対応が取れているため、顔だけを異なる人物に差し替えるなどの、画像の構造を利用したビデオ編集を行うことができる。

以上のように本システムでは、ビデオ映像から人物画像を生成するため、人物CGを生成するときに一般的に必要な複雑な作業を省略して、リアルな人物動画像を容易に生成することができる。そして、実在する俳優のビデオ映像から人物CGを生成できるため、例えば、映画の制作現場において、危険なスタントシーンを撮影するときに、主演している俳優が演技をしているように見せることができる。また、コマーシャルフィルム (CM) の撮影において、過去の有名な人物を登場させて、他の俳優と競演させることができる。

【 0 0 3 9 】

< その他の実施形態 >

(動画像データベース)

図 16 は本発明を動画像データベースに応用した場合の機能構成を示す図である。図 1 の構成と同様であるが、動作推定部 20 で人物動画像を人体の姿勢角と人物テクスチャに分解したデータを格納しておく蓄積部 36 と、蓄積部 36 より与えられた条件に一致するデータを検索して取出し、画像生成部 40 にデータを渡す検索部 38 を備えている。

ビデオ編集においては、より自由度の高い人物画像の編集が望まれている。例えば、既に撮影した映像中の人物を、異なるカメラアングルから見た場合に変更したり、衣服や髪型の変更、姿勢の変更を行うことができれば、ポスト・プロダクションの作業をより効果的に行うことができる。そこで本発明を応用して、動画像データベースに映像素材を蓄積しておき、簡便に検索することで映像素材を再利用することができる。

【 0 0 4 0 】

具体的な実現手段として、本システムの画像取得部 10 からとりこんだ映像を動作推定部 20 によって動作パラメータと人物テクスチャを求め、それらを蓄積部 36 に蓄積する。一方、検索部 38 は「人物が歩いているシーン」「座っているシーン」等の人物の動作を利用した検索条件を与え、検索を行い、検索条件に該当する動作パラメータを取出すことができる。また、本システムでは、人物テクスチャを分離して蓄積しているため、「赤い洋服を着た人物が写っているシーン」のように、画像の構造を検索条件として利用した検索も可能である。もちろん検索条件に該当する人物テクスチャも取出すことができる。動画像データベースである蓄積部 36 から検索部 38 が取出した動作パラメータと人物テクスチャは、画像生成部 40 において、カメラアングルを変更したり、衣服や髪型の変更、姿勢の変更を行うことができる。

【 0 0 4 1 】

上述のようにビデオ編集において、既に撮影した映像中の人物を、異なる視点から見た場

10

20

30

40

50

合に変更したり、衣服や髪型の変更、姿勢の変更を行うことができる。そのため、ポスト・プロダクションの作業をより効果的に行うことができる。

なお、これまでに述べたシステムを実施するためのコンピュータ・プログラムを格納した記録媒体から読み取ることにより、パソコンやワークステーション、複数のコンピュータ・システム上に本発明のシステムを実装することができる。また、プログラムをインターネットのような通信回線を介して、取得することにより、本システムをコンピュータ・システムに実装することもできる。

【 0 0 4 2 】

【発明の効果】

人体の3次元モデルを利用した動画像処理の適用により、リアルな人物CGを容易に生成することができ、ビデオ映像から人物の動作情報と輝度情報を取得することで、リアルな人物CGを容易に生成することができる。

そのため、ビデオ映像中の人物画像に対して、自由な視点位置から見たり、異なる動作に変更するなどの高度な操作を行うことが可能になる。また、人物のような複雑な被写体に対して、衣服や髪型などCGを合成することができる。

さらに、多様な姿勢を持つ人物画像を、関節角の変更で生成できるため、動画像圧縮の手段としても利用することができる。

【図面の簡単な説明】

【図1】 モデルに基づく人物動画像生成システムの機能構成を示す図である。

【図2】 ビデオ映像に本発明の処理を適用した例を示す図である。

【図3】 ビデオ映像に本発明の処理を適用した例を示す図である。

【図4】 ビデオ映像に本発明の処理を適用した例を示す図である。

【図5】 人体モデルを示す図である。

【図6】 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図である。

【図7】 画面上の線分の関係と人体パーツとの関係を示した図である。

【図8】 人体モデルの初期位置を推定するまでの処理の過程を示す図である。

【図9】 撮影位置とは異なる視点位置から画像を見た例を示す図である。

【図10】 入力された画像中の人体パーツから輝度情報を補間する手法を示す図である。

。

【図11】 CGで頭髮を生成した例を示す図である。

【図12】 CGの衣服シミュレーションの例を示す図である。

【図13】 衣服CGを合成した実験例を示す図である。

【図14】 人体モデルを動かして異なる姿勢に変更した例を示す図である。

【図15】 入力画像から得られた人物の人体モデルの姿勢を変形したものを示す図である。

【図16】 本発明を動画像データベースに応用した場合の機能構成を示す図である。

【符号の説明】

1 0	画像取得部
2 0	動作推定部
3 6	蓄積部
3 8	検索部
4 0	画像生成部
5 0	画像表示部
3 0 0	人体モデル
3 1 0	頭部
3 1 5	胴体
3 2 2	右上腕
3 2 4	左上腕
3 3 2	右前腕
3 3 4	左前腕

10

20

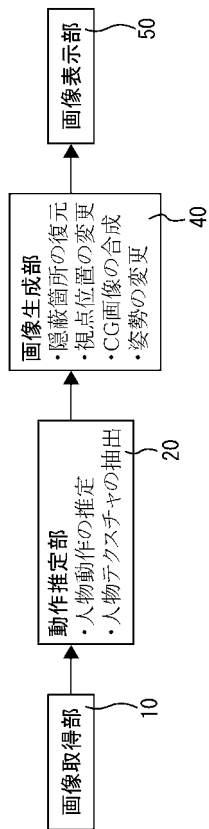
30

40

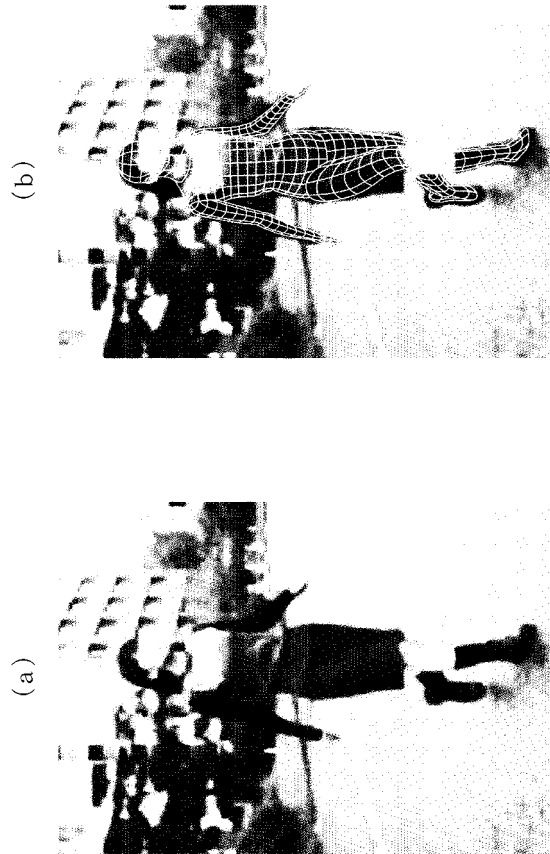
50

- 3 4 2 右腿
- 3 4 4 左腿
- 3 5 2 右膝下
- 3 5 4 左膝下
- 4 1 0 画面面

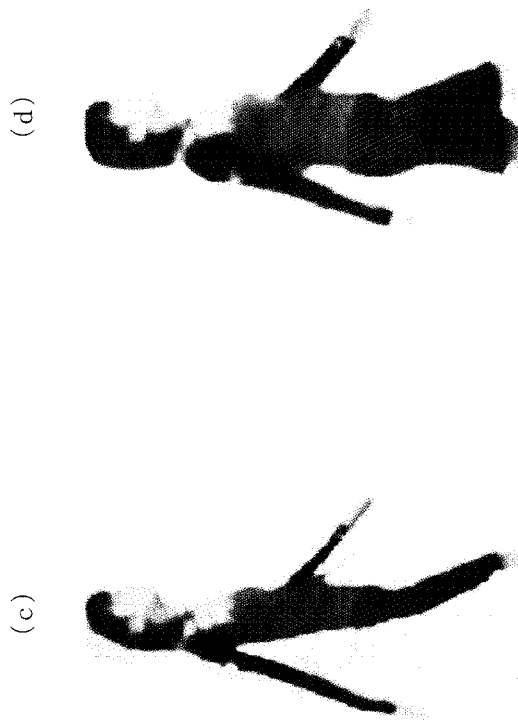
【図1】



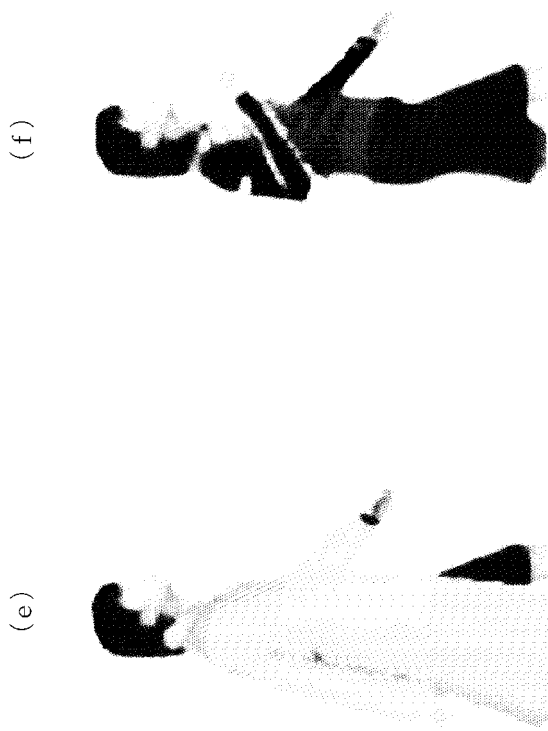
【図2】



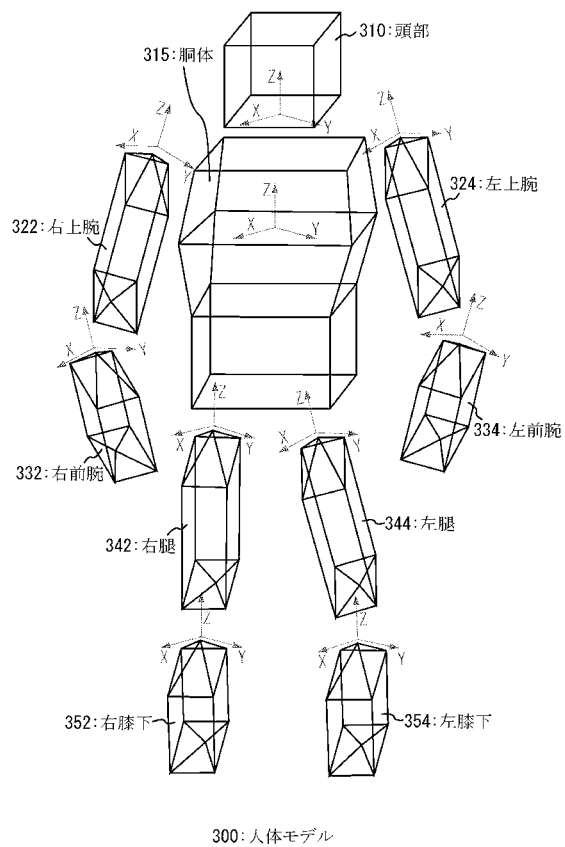
【図3】



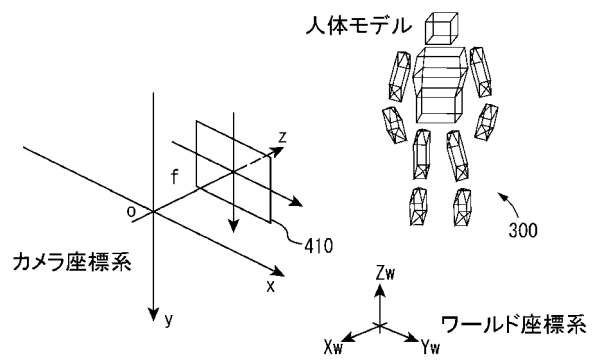
【図4】



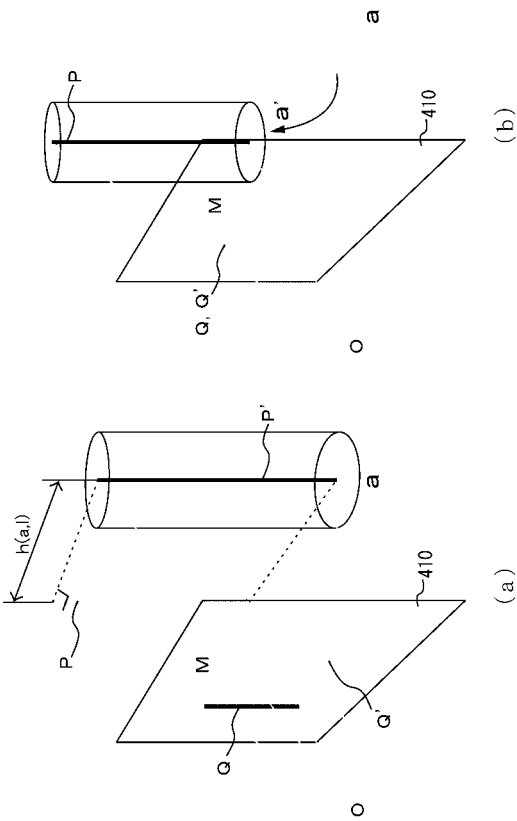
【図5】



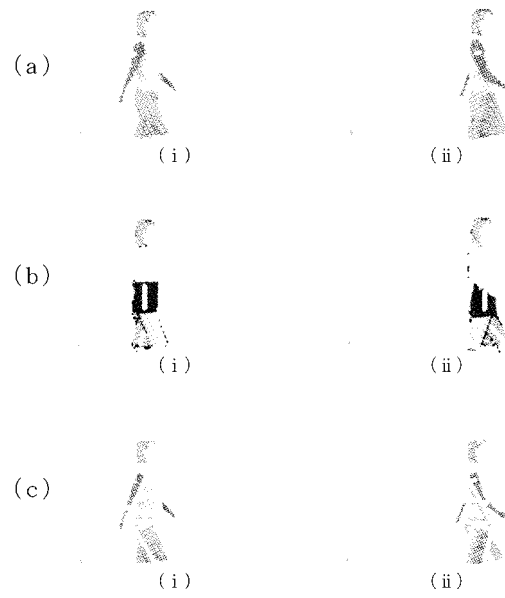
【図6】



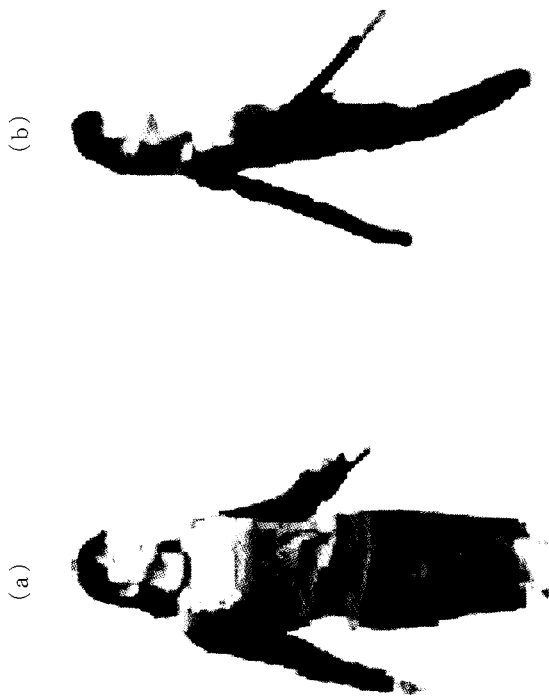
【図7】



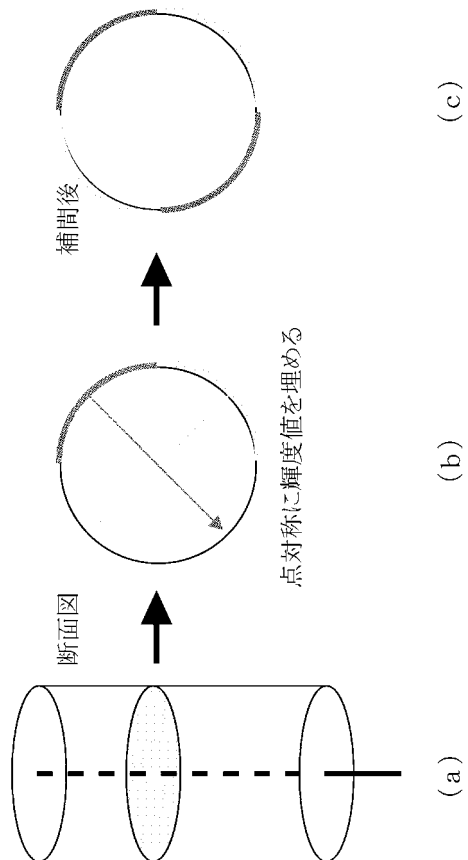
【図8】



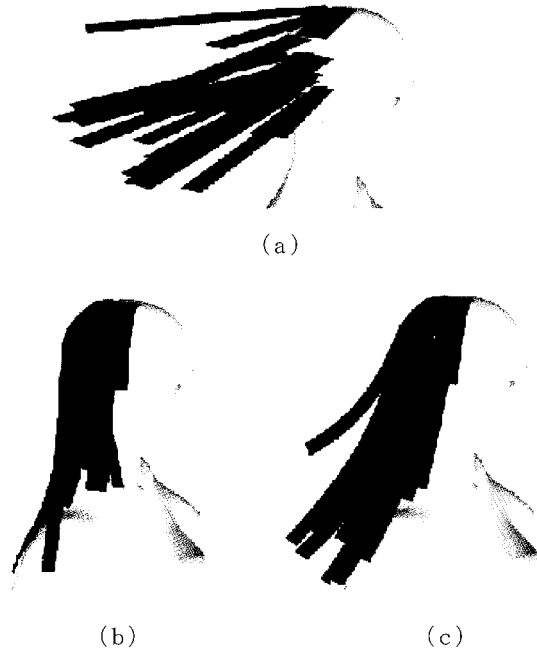
【図9】



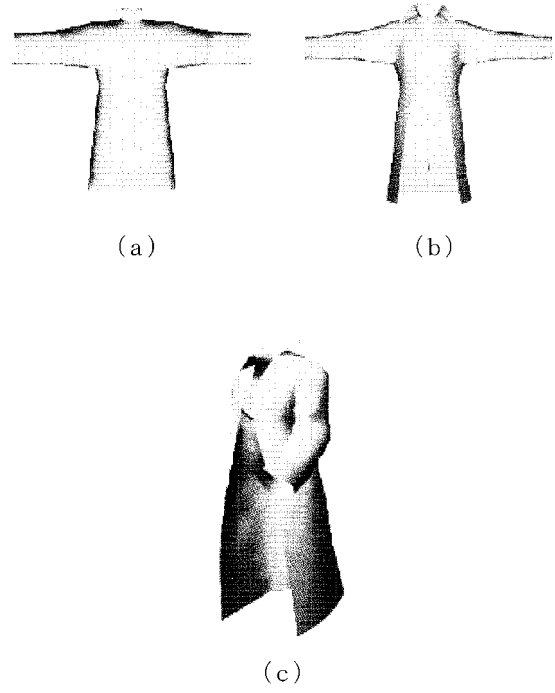
【図10】



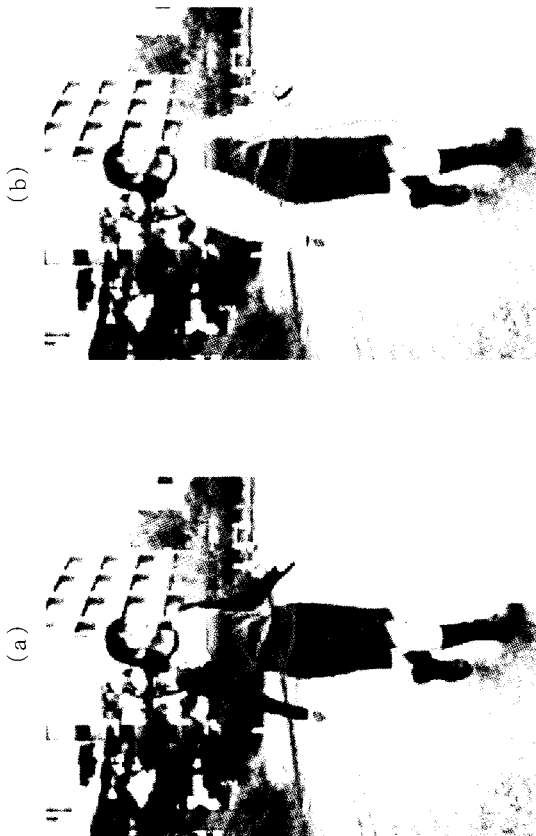
【図 1 1】



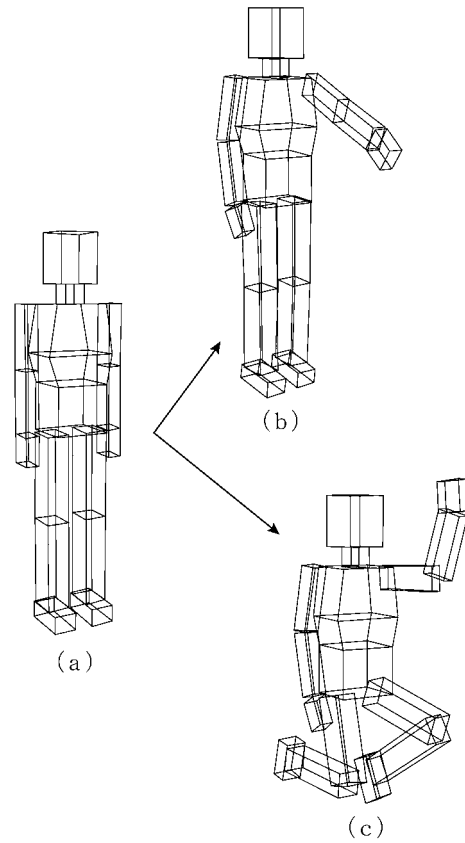
【図 1 2】



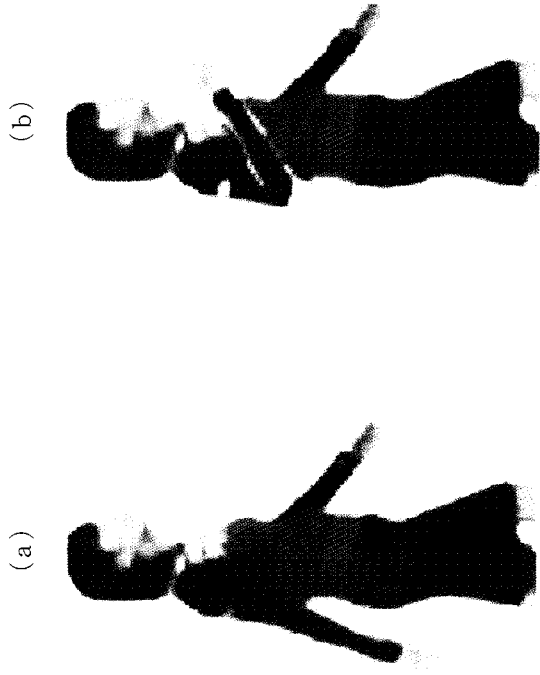
【図 1 3】



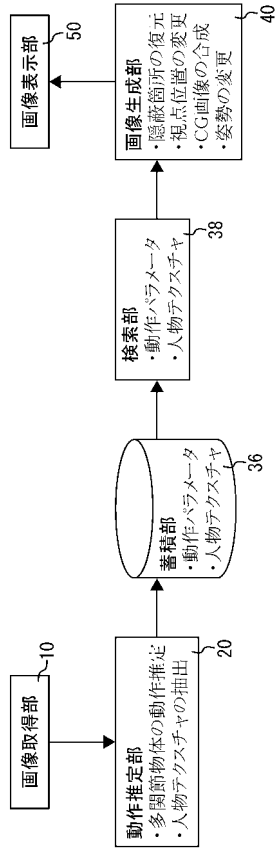
【図 1 4】



【図 15】



【図 16】



フロントページの続き

(56)参考文献 特開平08-147494(JP,A)

特開平09-035082(JP,A)

特開2001-005973(JP,A)

星昌人ほか, デジタルクローン俳優のための動画解析と動作生成, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2000年 6月22日, Vol.100 No.134, 1-8頁

斉藤啓史ほか, 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, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2000年 6月22日, Vol.100 No.134, 15-22頁

(58)調査した分野(Int.Cl., DB名)

G06T 13/00

G06T 15/00