

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4934810号  
(P4934810)

(45) 発行日 平成24年5月23日(2012.5.23)

(24) 登録日 平成24年3月2日(2012.3.2)

(51) Int.Cl.		F I			
<b>G06T</b>	<b>1/00</b>	<b>(2006.01)</b>	G06T	1/00	315
<b>G06T</b>	<b>7/20</b>	<b>(2006.01)</b>	G06T	1/00	340B
			G06T	7/20	300A

請求項の数 5 (全 13 頁)

(21) 出願番号	特願2006-355681 (P2006-355681)	(73) 特許権者	504174135
(22) 出願日	平成18年12月28日(2006.12.28)		国立大学法人九州工業大学
(65) 公開番号	特開2008-165580 (P2008-165580A)		福岡県北九州市戸畑区仙水町1番1号
(43) 公開日	平成20年7月17日(2008.7.17)	(74) 代理人	100090697
審査請求日	平成21年11月25日(2009.11.25)		弁理士 中前 富士男
		(74) 代理人	100139262
			弁理士 中嶋 和昭
		(72) 発明者	石川 聖二
			福岡県北九州市戸畑区仙水町1-1 九州工業大学内
		(72) 発明者	タン ジュークイ
			福岡県北九州市戸畑区仙水町1-1 九州工業大学内
		審査官	鹿野 博嗣

最終頁に続く

(54) 【発明の名称】 モーションキャプチャ方法

(57) 【特許請求の範囲】

【請求項1】

画像入力手段によって取得した対象動体の2次元的な動作画像から、該対象動体の3次元的な動作を再現するためのモーションキャプチャ方法であって、固有空間データ作成手段により、予め動体の基本動作ごとに、該動体の基本動作の各フレーム画像データAが点で表示される固有空間データAを作成し記憶手段に格納してデータベース化する固有空間データ作成工程と、木構造作成手段により、前記固有空間データ作成手段でデータベース化された前記固有空間データAを、前記動体の基本動作が保有している情報ごとに木構造群に分解し前記記憶手段に格納して構造化する木構造作成工程と、判別手段により、判別しようとする前記対象動体の動作のフレーム画像データBが点で表示された固有空間データBと、前記木構造作成手段で構造化された前記動体の基本動作ごとの固有空間データAとを比較して、前記固有空間データBからの距離が最も近い固有空間データAを選び、前記対象動体の3次元的な動作を特定する判別工程とを有し、前記木構造群では、大小比較が可能な境界を表す値であるキーによって根および節が構築され、しかも、前記固有空間データAは葉のみに格納され、更に、前記判別工程での前記固有空間データBと前記固有空間データAとの比較は前記キーを用いて行うことを特徴とするモーションキャプチャ方法。

【請求項2】

請求項1記載のモーションキャプチャ方法において、前記基本動作ごとの前記各フレーム

画像データAは、連続するまたは間隔を有する2つずつのフレーム画像を重ね合わせ、変化のない部分を削除して得られる複数の差分画像を、それぞれ重ね合わせることで得られることを特徴とするモーションキャプチャ方法。

【請求項3】

請求項1および2のいずれか1項に記載のモーションキャプチャ方法において、前記固有空間データ作成工程での前記動体は疑似人モデルまたは人であり、しかも前記画像入力手段は、前記動体が疑似人モデルの場合は仮想カメラ群、前記動体は人の場合はカメラ群であり、前記動体の基本動作を前記画像入力手段を用いて多方向から撮影し、前記基本動作ごとの複数のフレーム画像データAを得ることを特徴とするモーションキャプチャ方法。

【請求項4】

請求項1～3のいずれか1項に記載のモーションキャプチャ方法において、前記固有空間データAは、前記フレーム画像データAに微分処理を行って作成されることを特徴とするモーションキャプチャ方法。

【請求項5】

請求項1～4のいずれか1項に記載のモーションキャプチャ方法において、前記固有空間データAは、前記フレーム画像データAをカラーネン・レーベ変換して求められた固有値および固有ベクトルから作成される固有空間に投影して得られ、前記固有空間データBは、前記フレーム画像データBを前記固有空間に投影して得られることを特徴とするモーションキャプチャ方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、例えば、ビデオカメラのような画像入力手段によって取得した対象動体の2次元的な動作画像、特に人の2次元的な動作画像から、アニメーション、三次元アバター、またはロボットの3次元的な動作を再現するためのモーションキャプチャ方法に関する。

【背景技術】

【0002】

従来、人の動作を測定し、その立体モデルを造る技術として、モーションキャプチャ方法がある。

このモーションキャプチャ方法としては、例えば、機械式方法、磁気式方法、および光学式方法が主要な方法として挙げられる（例えば、特許文献1～4参照）。

中でも、光学式方法は、例えば、対象動体である人をカメラにより撮影するだけであるため、動作の制限を最も受けにくい方法であり、広く利用されている。例えば、映画、ビデオゲーム、スポーツ、またはダンスのように、運動または動作を表現する3次元メディアに関連する分野で、それらの立体コンテンツの製作に利用されてきた。これらの分野では、たくさんカメラがセットされたスタジオを所有しているため、このスタジオでシナリオ通りの動作を行うことにより、その動作データが得られ、立体モデル化ができるからである。

【0003】

【特許文献1】特開2005-345161号公報

【特許文献2】特開2004-101273号公報

【特許文献3】特開2000-321044号公報

【特許文献4】特開平10-74249号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、前記従来からのモーションキャプチャ方法には、未だ解決すべき以下のような問題があった。

機械式方法は、測定機械を身体に装着する必要があるため、表現可能な動作に制限がある。

10

20

30

40

50

また、磁気式方法は、磁場を発生させた環境でしか使用できないという問題があり、適用可能な場所が制約される。

そして、光学式方法は、広く利用されてはいるが、事前にカメラのキャリブレーション（配置および位置設定）が必要であり、また人の動きを測定するためのマーカも、身体に装着する必要があるため、作業性が悪い。なお、視体積交差法（バックプロジェクト法）を使用すれば、マーカの使用は不要となるが、カメラのキャリブレーションは必要である。一方、因子分解に基づく方法を使用すれば、カメラのキャリブレーションは不要となるが、マーカの使用が必要となる。

【 0 0 0 5 】

本発明はかかる事情に鑑みてなされたもので、画像入力手段によって取得した対象動体の2次元的な動作画像から、観察方向によらずに、自動かつ高速に、3次元的な動作を再現可能なモーションキャプチャ方法を提供することを目的とする。

【課題を解決するための手段】

【 0 0 0 6 】

前記目的に沿う本発明に係るモーションキャプチャ方法は、画像入力手段によって取得した対象動体の2次元的な動作画像から、該対象動体の3次元的な動作を再現するためのモーションキャプチャ方法であって、

固有空間データ作成手段により、予め動体の基本動作ごとに、該動体の基本動作の各フレーム画像データAが点で表示される固有空間データAを作成し記憶手段に格納してデータベース化する固有空間データ作成工程と、

木構造作成手段により、前記固有空間データ作成手段でデータベース化された前記固有空間データAを、前記動体の基本動作が保有している情報ごとに木構造群に分解（木構造内に分配）し前記記憶手段に格納して構造化する木構造作成工程と、

判別手段により、判別しようとする前記対象動体の動作のフレーム画像データBが点で表示された固有空間データBと、前記木構造作成手段で構造化された前記動体の基本動作ごとの固有空間データAとを比較して、前記固有空間データBからの距離が最も近い固有空間データAを選び、前記対象動体の3次元的な動作を特定する判別工程とを有し、

前記木構造群では、大小比較が可能な境界を表す値であるキーによって根および節が構築され、しかも、前記固有空間データAは葉のみに格納され、更に、前記判別工程での前記固有空間データBと前記固有空間データAとの比較は前記キーを用いて行う。

【 0 0 0 7 】

ここで、木構造とは、動体の基本動作が保有している情報、例えば、動体の基本動作をその特徴ごとに区分する方法であり、例えば、B - tree、B \* - tree、またはB + - treeが従来知られている。

また、動体として人を対象とした場合、その基本動作として、例えば、重たい荷物を持ち上げる動作、物を拾う動作、腹痛でおなかを抱えて座り込む動作、頭上から落ちてくる物を避けようと両手で頭を覆う動作、歩く動作、および転倒する動作等がある。また、動体としては、人の他に、動物、車等の乗り物、またはロボット等を適用することができる。

【 0 0 0 8 】

本発明に係るモーションキャプチャ方法において、前記基本動作ごとの前記各フレーム画像データAは、連続するまたは間隔を有する2つずつのフレーム画像を重ね合わせ、変化のない部分を削除して得られる複数の差分画像を、それぞれ重ね合わせることで得られることが好ましい。

ここで、差分画像は、例えば、動体の背景画像を削除することにより得られる画像であり、このような背景画像を削除することで、処理するデータ量を少なくできる。

【 0 0 0 9 】

本発明に係るモーションキャプチャ方法において、前記固有空間データ作成工程での前記動体は疑似人モデルまたは人であり、しかも前記画像入力手段は、前記動体が疑似人モデルの場合は仮想カメラ群、前記動体は人の場合はカメラ群であり、前記動体の基本動作を前記画像入力手段を用いて多方向から撮影し、前記基本動作ごとの複数のフレーム画像デ

10

20

30

40

50

ータ A を得ることが好ましい。

ここで、疑似人モデルとは、人の 3 次元モデルのことであり、一般にアバターといわれる。疑似人モデルを使用することにより、基本動作を行う動体を標準化することができる。

【 0 0 1 0 】

本発明に係るモーションキャプチャ方法において、前記固有空間データ A は、前記フレーム画像データ A に微分処理を行って作成されることが好ましい。

ここで、微分処理とは、例えば、ログ ( L o G ) フィルタ、またはソーベルフィルタ等によって行う方法である。なお、ログフィルタとは、画像データをぼかして微分する方法である。

【 0 0 1 1 】

本発明に係るモーションキャプチャ方法において、前記固有空間データ A は、前記フレーム画像データ A をカルーネン・レーベ変換して求められた固有値および固有ベクトルから作成される固有空間に投影して得られ、前記固有空間データ B は、前記フレーム画像データ B を前記固有空間に投影して得られることが好ましい。

ここで、カルーネン・レーベ変換とは、カルーネン・レーベ展開ともいわれ、高次元の各フレーム画像データを低次元に変換する方法である。

【 発明の効果 】

【 0 0 1 2 】

請求項 1 ~ 5 記載のモーションキャプチャ方法は、フレーム画像データ A、B を点で表示した固有空間データ A、B を用いて固有空間を構築するので、処理するデータの量を少なくでき、対象動体の動作を高速に処理できる。また、固有空間データ A を木構造群に分解するので、固有空間データ A と固有空間データ B との全てを比較することなく、固有空間データ A を選んで対象動体の 3 次元的な動作を特定でき、処理速度の更なる高速化を図ることができる。

従って、動体の多数の基本動作を記憶手段に格納することで、例えば、任意の場所における不特定の人物の動作を、観察方向によらず、ビデオカメラのような画像入力手段で撮影し、立体的に再現することができる。

【 0 0 1 3 】

特に、請求項 2 記載のモーションキャプチャ方法においては、基本動作ごとの各フレーム画像データ A が、複数の差分画像をそれぞれ重ね合わせることで得られるので、処理するデータ量を少なくでき、動作認識の処理時間を更に短くできる。

請求項 3 記載のモーションキャプチャ方法においては、基本動作を疑似人モデルに行わせ複数のフレーム画像データ A を得た場合、標準化した人のデータとすることができ、体型の違いを無くすることができる。更に、疑似人モデルまたは人を多方向から観察した動作画像を用いるので、対象物体をどの方向から観察した場合でも動作を判別することができる。

【 0 0 1 4 】

請求項 4 記載のモーションキャプチャ方法においては、フレーム画像データ A に微分処理を行って、固有空間データ A を作成するので、例えば、服装の違いによる誤差 ( ノイズ ) を減少させることができ、例えば、疑似人モデルをより標準化することができる。

請求項 5 記載のモーションキャプチャ方法においては、固有空間データ A が、フレーム画像データ A をカルーネン・レーベ変換して求められた固有値および固有ベクトルから作成される固有空間に投影して得られ、固有空間データ B が、フレーム画像データ B を固有空間に投影して得られるので、次元を低くでき、動作認識の処理時間を短くできる。

【 発明を実施するための最良の形態 】

【 0 0 1 5 】

続いて、添付した図面を参照しつつ、本発明を具体化した実施の形態につき説明し、本発明の理解に供する。

本発明の一実施の形態に係るモーションキャプチャ方法は、1 台のビデオカメラ ( 画像入力手段の一例 ) によって取得した対象動体の一例である対象人の 2 次元的な動作画像を、

10

20

30

40

50

予め登録された動体の一例である人が行った複数の基本動作と比較して、対象人の3次元的な動作を再現するための方法である。以下、詳しく説明する。

【0016】

まず、人が行う基本動作の固有空間データAを作成する固有空間データ作成工程について説明する。

人を中心にして等距離で、しかも等角度に、複数台（例えば、4台）のビデオカメラ（動画が撮影可能であればよい）を配置し、人が行う各基本動作（例えば、重たい荷物を持ち上げる動作、物を拾う動作、腹痛でおなかを抱えて座り込む動作、頭上から落ちてくる物を避けようと両手で頭を覆う動作、歩く動作、および転倒する動作等）を撮影する。なお、ビデオカメラとしては、例えば、CCDカメラ、高速度カメラ、ハンディタイプカメラ、デジタルVTR、またはデジタルビデオカメラを使用してもよい。

10

次に、それぞれの基本動作を撮影した映像をコンピュータに取り込む。なお、以下の作業は、コンピュータ内で計算して行われ、コンピュータ内のプログラムにより処理される。

【0017】

コンピュータ内に取り込まれた画像のうち、各ビデオカメラごとに、例えば、1秒間に1コマ以上50コマ以下の間隔で得られる連続する複数のフレーム画像は、コンピュータ内の前処理手段により重ね合わせられる。このとき、変化のない部分、例えば、人の周囲に存在する背景画像（例えば、壁、床、および空）を削除するが、人の画像についても、動きが無い部分（僅かに動く部分を含んでもよく、また含まなくてもよい）を削除してもよい。なお、複数のフレーム画像は、例えば、2枚ごと、または3枚ごとのように、複数枚ごとに間隔を有するものでもよい。

20

これにより、基本動作の一連の動作が残像として示される1枚の圧縮された画像を、基本動作の複数のフレーム画像データAとして、コンピュータ内の記憶手段に格納できる。

【0018】

このとき、前記した複数のフレーム画像を、連続する2つのフレーム画像ごとに重ね合わせ、変化のない部分、例えば、人の周囲に存在する背景画像を引き算して削除した後、得られる複数の差分画像をそれぞれ重ね合わせることで、1枚の圧縮画像を得てもよい。

なお、上記した人の行う基本動作の3次元データは、人に実際に動作をしてもらうことで得たが、例えば、コンピュータグラフィクスを用いた疑似人モデルにより作成してもよく、また予め他のモーションキャプチャ法で獲得したデータを用いて疑似人モデルにより作成してもよい。

30

この場合、人の基本動作を疑似人モデルに行わせ、この疑似人モデルを中心として、水平方向、上方向、および下方向のいずれか1または2以上に、等間隔で等角度に配置される多数（例えば、6台以上）の仮想ビデオカメラからなる仮想カメラ群により、疑似人モデルを撮影して、複数のフレーム画像データAを得る。

【0019】

ここで、各フレーム画像データAは、基本動作を撮影した動画中の画像の集合であって、1枚の画像データが、例えば、縦が256ピクセル、横が256ピクセルで構成されている場合、総画素数が65536画素、つまり65536(N)次元のデータが得られる。また、例えば、1秒間に15コマで2秒間撮影した場合、1方向から30(P)枚のフレーム画像が得られるが、前記したように、画像の前処理を行うことで、その動作を表す複数のフレーム画像は、1画像で圧縮表現される。

40

次に、コンピュータ内の固有空間データ作成手段により、予め人の基本動作ごとに、人の基本動作の各フレーム画像データAが点で表示される固有空間データAを作成する。なお、この固有空間データAの作成は、特願2005-237785の方法と同様の手法で実施できる。

【0020】

得られた1つの基本動作のフレーム画像データA（以下、単に画像ともいう）に対し、正規化を行い、従来公知のTVラスタースキャンと同様の方法で走査して、(1)式に示すベクトルを得る。

50

$$x_p = (x_1, x_2, \dots, x_N)^T \dots (1)$$

ここで、ベクトルの各要素は、スキャンした順番で並んでいる画素数である。なお、Nはピクセル数を示し、Tは転置を示し、また $x_p$ は $x_p = 1$ となるように正規化されている。

次に、N行P列の行列Xを、(2)式のように定義する。

$$X = (x_1 - c, x_2 - c, \dots, x_P - c) \dots (2)$$

ここで、cは画像の平均値であり、(3)式で計算される。

【0021】

【数1】

$$c = \sum_{p=1}^P x_p / P \dots (3)$$

10

【0022】

また、共分散行列Qは、行列Xより(4)式で定義される。

$$Q = X X^T \dots (4)$$

【0023】

カルーネン・レーベ変換により、(5)式を用いて共分散行列Qの固有値 $\lambda_1, \lambda_2, \dots, \lambda_N$ を求める。但し、 $\lambda_1 > \lambda_2 > \dots > \lambda_N$ である。

20

$$Q u = \lambda u \dots (5)$$

ここで、uはN個の成分を持つベクトルである。

得られた固有値 $\lambda_1, \lambda_2, \dots, \lambda_N$ から、固有ベクトル $e_1, e_2, \dots, e_N$ が求められる。

【0024】

ここで、固有ベクトルのk個の最大固有値 $\lambda_1, \lambda_2, \dots, \lambda_k$ 、および、それに対応する固有ベクトル $e_1, e_2, \dots, e_k$ を選択し、k個の固有ベクトルの張る空間、即ち、(6)式に示すk次元の固有空間ESを作成する。

$$ES = (e_1, e_2, \dots, e_k) \dots (6)$$

なお、 $k \leq N$ であり、固有空間ES上に画像データを写像する変換行列Eは、(7)式で示される。例えば、kを100とした場合には、N次元からk次元、すなわち、65536次元から100次元に次元を下げるができる。

30

$$E = (e_1, e_2, \dots, e_k) \dots (7)$$

【0025】

ここで、(8)式により、各フレーム画像データAを固有空間ES上に投影して、固有空間データAとして点の集合 $g_p$ を得る。

$$g_p = (e_1, e_2, \dots, e_k)^T x_p \dots (9)$$

このようにして、人の姿勢は、固有空間上で単なる点として登録される。

この得られた点の集合 $g_p$ を、記憶手段に格納してデータベース化する。

なお、固有空間データAの作成に際しては、事前に、コンピュータ内に取り込まれた画像の各フレーム画像データAのそれぞれの画像データを、従来公知のログフィルタにかけ、各フレーム画像データAをぼかして微分処理してもよい。

40

【0026】

以下、同様に他の方向から撮影した基本動作の各フレーム画像データAからそれぞれ固有空間データAを作成し、得られた点の集合を、記憶手段に格納してデータベース化する。また、複数の基本動作の全てのフレーム画像データAから、同様にして固有空間データAを作成し、記憶手段に格納してデータベース化する。

次に、以上に示した固有空間データ作成手段でデータベース化された固有空間データAを、コンピュータ内の木構造作成手段により、人の基本動作が保有している情報ごとに木構造群に分解する木構造作成工程について説明する。なお、木構造としては、例えば、B -

50

$tree$ 、 $B^* - tree$ 、または $B^+ - tree$ が従来知られている。

【0027】

固有空間に対して、 $B - tree$ を適用するという考えは、固有空間を複数に分割し、点として表現された姿勢をそれぞれ格納するピン（貯蔵箱：人の基本動作が所有している情報ごとに分解され構成される木構造群）を作り、入力された未知の姿勢と似た画像の格納されたピンを高速に探し出すことにある。

$B - tree$ 構造を固有空間に導入して、固有空間の構造化を行うことにより、圧縮画像が点として表現された固有空間は、複数のピンに分けられ、ピンは $B - tree$ 構造で表現される。

なお、人の動作を表すこの固有空間を、動作データベースと呼ぶ。

10

【0028】

ここで、 $B - tree$ について説明する。

以下の条件を満たすものを、 $(m, H)$ に属する $B - tree$  という。ここで、 $m$ は、根（ルート）または節（ノード）が持つことのできる子供の数である。また、 $H$ は木の高さを表し、検索速度に関係する。

1．根は葉であるか、または $2 \sim m$ 個の子を持つ。

2．根、葉以外の節は、 $[m/2] \sim m$ 個の子を持つ。ただし、 $[x]$ は $x$ 以下の最大の整数を表す。

3．根からすべての葉までの経路の長さは等しい。

$B - tree$ では、格納するデータから造られる「境界を表す値」、即ちキーが重要な意味を持ち、このキーによって根や節が構築される。このキーは、大小比較することが可能なスカラー値である。また、データは、葉のみに格納される。

20

【0029】

この $B - tree$ を固有空間に適用するときは、それぞれの固有空間上の座標値 $e_k$  ( $k = 1, 2, \dots, K$ )を、 $R$ 個のある幅 $L$ を持ったセクションに分割し、木構造を作成する。

ここで、画像 $I_p$ が式(9)によって固有空間の点 $g = (g_1, g_2, \dots, g_k)$ に投影されると、 $g_k$  ( $k = 1, 2, \dots, K$ )は、いずれかのセクションに含まれるから、そのセクションの固有の番号 $S_{k,r}$  ( $r = 0, 1, \dots, R - 1$ )が与えられる。

30

この結果 $g$ は、式(10)によって、 $K$ 桁 $R$ 進数である $S_p$ に変換される。

$$S_p = S_{1,r_1} S_{2,r_2} S_{3,r_3} \dots S_{K,r_K} \dots \quad (10)$$

これにより、画像は、 $S_p$ をキーとして、木構造である $B - tree$ に分配されて格納されるので、これを、記憶手段に格納して、構造化する。

以上の方法により、人の各基本動作がデータベース化される。

【0030】

次に、前記したコンピュータ内の固有空間データ作成手段により、判別しようとする対象人のフレーム画像データ $B$ が点で表示された固有空間データ $B$ を作成する。

まず、対象人の動作を1台のビデオカメラで撮影する。

動作画像をコンピュータに取り込み、(11)式に示す各フレーム画像データ $B$ の集合 $y$ を得る。

40

$$y = (y_1, y_2, \dots, y_p) \dots \quad (11)$$

そして、前記した固有空間データ $A$ と前処理と同様の方法により、その動作を表す連続フレームを圧縮表現して1画像とすることで、フレーム画像データ $B$ が作成される。

【0031】

更に、前記した固有値および固有ベクトルから作成される固有空間 $ES$ に、(12)式を用いて、フレーム画像データ $B$  ( $y'$ と表記)を投影し、固有空間データ $B$ である点 $h$ を得る。

$$h = E^T y' = (e_1, e_2, \dots, e_k)^T y' \dots \quad (12)$$

そして、コンピュータ内の判別手段により、固有空間データ $B$ と、木構造作成手段で構造

50

化された人の基本動作ごとの固有空間データ A とを比較する判別工程について説明する。

【0032】

人の姿勢認識では、未知の姿勢を持つ画像  $I_p$  を固有空間に投影し、(10)式によってセクション番号  $S_p$  を得る。次に、 $S_p$  を検索キーとして  $g$  を検索し、候補姿勢  $g_{p_r}$  ( $r = 1, 2, \dots, R$ ) を得る。

最後に、(13)式を適用すれば、固有空間データ B を示す点  $h$  からの距離が最も近い(距離が最小)固有空間データ A を示す点の集合  $g_i$  が選ばれ、最も近い姿勢  $p = p^*$  が得られる。

$$d_{p^*} = \min g_{p_r} - g_p \dots (13)$$

ここでは、 $R = P$  となることが期待されるため、検索速度は大幅に改善される。

10

ただし、 $R$  動作データベースに登録されている全基本動作の数である。

【0033】

このように、人の動作を任意方向からビデオ撮影し、その動作に最も近い動作を動作データベースの検索により探し、見つければそれを3次元的な動作と特定することで、例えば、アニメーション、アバター、またはロボットのような3次元媒体で再現できる。

これにより、固有空間データ B に最も近い圧縮画像が検索されるが、この圧縮画像は、もとの動作情報(即ち、動体の基本動作)を持っているため、これを参照することで、3次元動作を再現できる。

なお、未知の動作がカメラで撮影される場合、画像の前処理により、その動作を表す連続するフレーム画像は、1画像  $I$  で圧縮表現されるため、画像  $I$  に最も近い圧縮画像が、コンピュータ内の基本動作のデータベースから検索される。このデータベースは、前記したように、*B-tree* 構造を持つため、検索は高速に行われる。従って、画像  $I$  と最も距離の短い画像が検索され、この距離がある閾値より小さければ、未知動作は、その動作として判断される。

20

以上の方法により、動作データベース検索によるモーションキャプチャが実現する。

【実施例】

【0034】

次に、本発明の作用効果を確認するために行った実施例について説明する。

ここでは、本願発明のモーションキャプチャ方法を適用し、ビデオカメラによって取得した対象人の2次元的な動作画像から、対象人の3次元的な動作を再現する方法について説明する。

30

まず、図1(A)、(B)、図2(A)、(B)、図3(A)、(B)に示すように、人に、重たい荷物を持ち上げる動作、物を拾う動作、腹痛でおなかを抱えて座り込む動作、頭上から落ちてくる物を避けようと両手で頭を覆う動作、歩く動作、および転倒する動作の各基本動作を行ってもらい、この動作画像をビデオカメラで連続的に撮影して、コンピュータに入力する。なお、ここでは、説明の便宜上、一方向からのみ撮影した映像を示す。

【0035】

次に、前記した前処理について説明する。

ここでは、図1~図3の各基本動作を、その動作ごとに重ね合わせ、図4(A)の(a)~(f)に示す1枚の画像データを得る。なお、図1(A)は図4(A)の(a)、図1(B)は図4(A)の(b)、図2(A)は図4(A)の(c)、図2(B)は図4(A)の(d)、図3(A)は図4(A)の(e)、および図3(B)は図4(A)の(f)に、それぞれ対応する。

40

そして、図4(A)の(a)~(f)の背景画像を削除することで、図4(B)に示す抽出画像を得た後、前記した正規化を行うことで、図5に示す画像が得られる。

【0036】

また、前処理は、以下の方法で行うこともできる。

ここでは、図1~図3の各基本動作を、連続する2つのフレーム画像ごとに重ね合わせ、変化のない部分、即ち人の周囲に存在する背景画像(例えば、壁、床、および空)を引き

50



算して削除する。そして、この差分画像をそれぞれ重ね合わせることで、図6(A)に示すように、基本動作の一連の動作が残像として示される1枚の圧縮画像を作成できる。なお、図1(A)は図6(A)の(a)、図1(B)は図6(A)の(b)、図2(A)は図6(A)の(c)、図2(B)は図6(A)の(d)、図3(A)は図6(A)の(e)、および図3(B)は図6(A)の(f)に、それぞれ対応する。

そして、図6(A)の(a)~(f)に残存する背景画像を削除することで、図6(B)に示す抽出画像を得た後、前記した正規化を行うことで、図7に示す画像が得られる。

#### 【0037】

これにより、各圧縮画像には、対応する3次元の動作を再現するのに必要な情報を持たせることができる。

この各フレーム画像データAからカラーネン・レーベ変換により、固有値および固有ベクトルを計算し、固有空間データAを作成し、複数の基本動作の固有空間データAを、記憶手段に格納してデータベース化する。

次に、このデータベース化された固有空間データAを、木構造に分解(動体の基本動作が保有している情報、例えば、その画像特徴ごとに区分)し、記憶手段に格納してデータベース化する。これにより、固有空間データAは、例えば、類似画像(同じピンに異なる基本動作が入る場合もある)ごとに分類される。

#### 【0038】

一方、3次元的な動作を得るための元データとなる人の動作については、1台のビデオカメラで対象人を撮影し、動作画像をコンピュータに取り込み、各フレーム画像データBを得る。そして、前記した各フレーム画像データAと同様の方法で、1画像に圧縮されたフレーム画像データBを、前記した固有値および固有ベクトルから作成される固有空間に投影して固有空間データBを得る。

そして、固有空間データBと基本動作ごとの固有空間データAとを比較して、固有空間データBからの距離が最も近い固有空間データAを選び、これを対象人の3次元的な動作と特定する。

これにより、例えば、アニメーション、アバター、またはロボットのような3次元媒体で再現できる。

#### 【産業上の利用可能性】

#### 【0039】

本発明のモーションキャプチャ方法によって、例えば、従来のもーションキャプチャ方法では不可能であった予期せぬ突発的な事象(広場で突然ストリートダンスが始まる場合のように、身体にマーカをつけず、自然な状態で動作してもらえらる事象)などの、立体モデル化が可能となり、また容易となる。これは、画像入力手段により、現地で対象動体の動作を撮影するだけでよいからである。

また、従来のように、事前のカメラキャリブレーションおよびマーカも必要なく、対象となる動体の動作のみを撮影するだけでよいので、立体モデル化できる動作数が増え、また操作が簡単なモーションキャプチャ方法を提供できる。これにより、低価格で使い易いモーションキャプチャシステムを提供できるので、モーションキャプチャ技術の他分野への普及が期待できる。

#### 【0040】

以上、本発明を、実施の形態を参照して説明してきたが、本発明は何ら上記した実施の形態に記載の構成に限定されるものではなく、特許請求の範囲に記載されている事項の範囲内で考えられるその他の実施の形態や変形例も含むものである。例えば、前記したそれぞれの実施の形態や変形例の一部または全部を組合せて本発明のモーションキャプチャ方法を構成する場合も本発明の権利範囲に含まれる。

また、前記実施の形態においては、本願発明である動体として人を適用した場合について説明したが、これに限定されるものではなく、動体として、例えば、人以外の動物、車等の乗り物、またはロボットでもよい。

#### 【図面の簡単な説明】

10

20

30

40

50

【 0 0 4 1 】

【図1】(A)は人が重い荷物を持ち上げる基本動作の連続するフレーム画像の説明図、(B)は人が物を拾う基本動作の連続するフレーム画像の説明図である。

【図2】(A)は人が腹痛でおなかを抱えて座り込む基本動作の連続するフレーム画像の説明図、(B)は人が頭上から落ちてくる物を避けようと両手で頭を覆う基本動作の連続するフレーム画像の説明図である。

【図3】(A)は人が歩く基本動作の連続するフレーム画像の説明図、(B)は人が転倒する基本動作の連続するフレーム画像の説明図である。

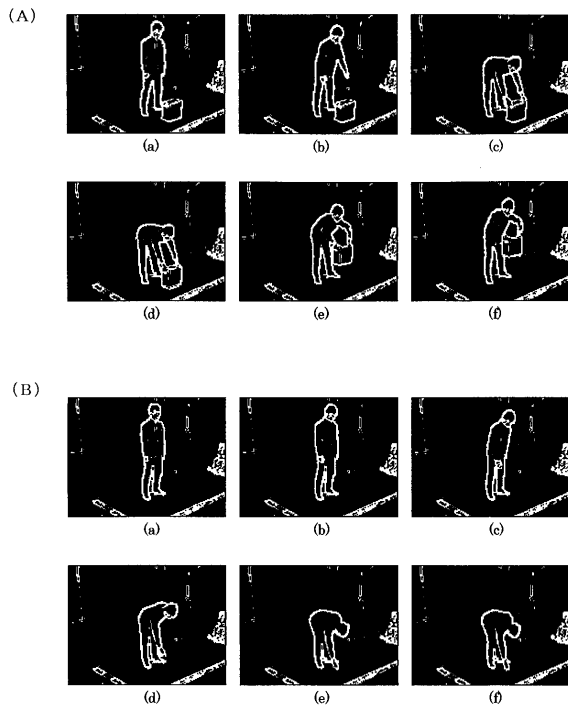
【図4】(A)は図1～図3に示す各基本動作ごとの画像の説明図、(B)は(A)の背景画像を削除した後の画像の説明図である。

【図5】図4(B)の画像を正規化した画像の説明図である。

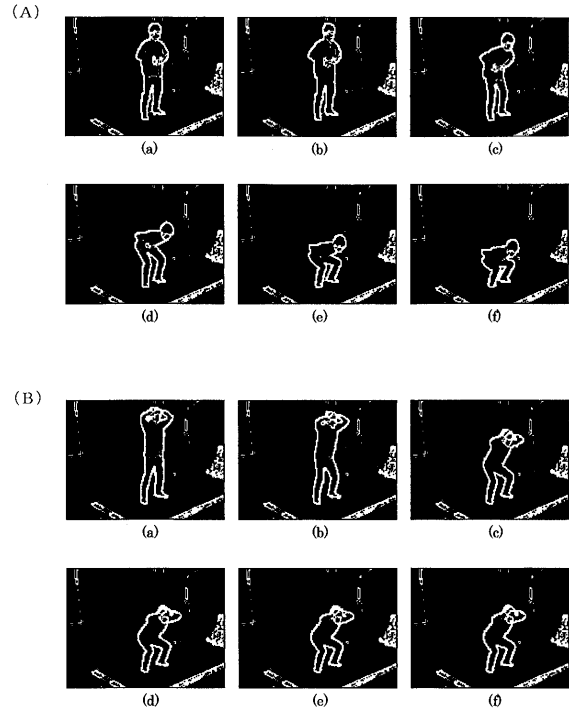
【図6】(A)は図1～図3に示す各基本動作ごとの連続する画像を重ね合わせ変化のない部分を削除した後得られる圧縮画像の説明図、(B)は(A)の背景画像を削除した後の圧縮画像の説明図である。

【図7】図6(B)の画像を正規化した画像の説明図である。

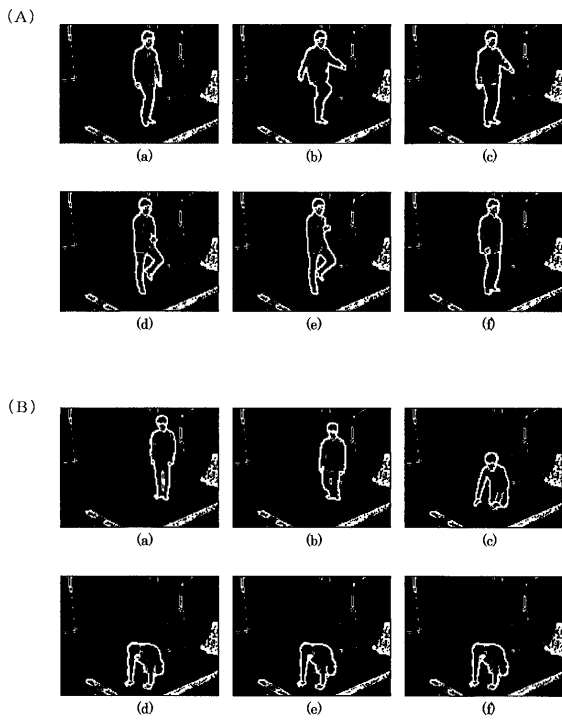
【図1】



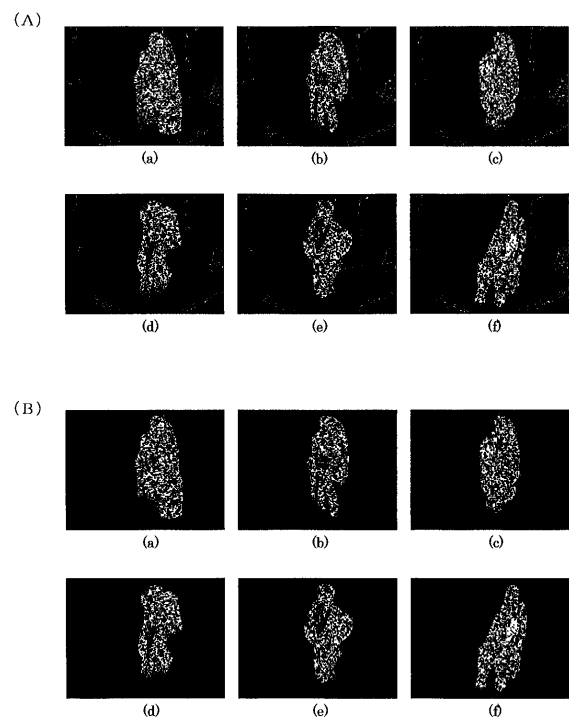
【図2】



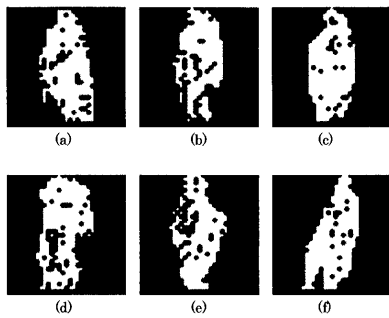
【 図 3 】



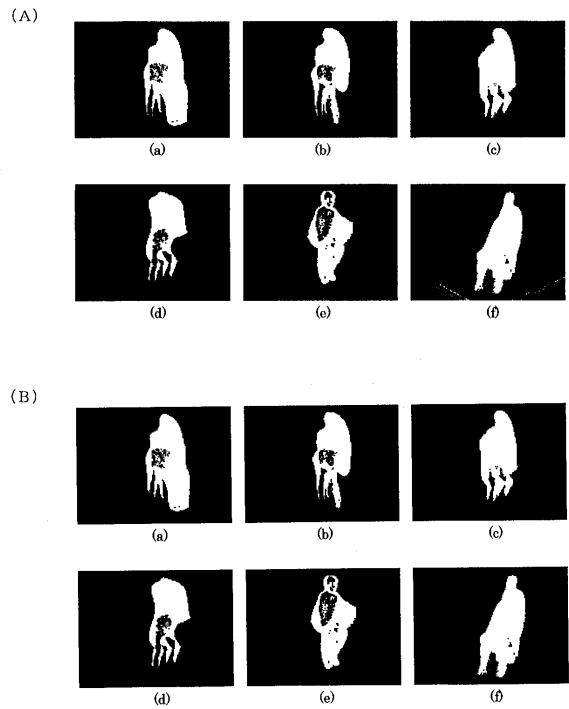
【 図 4 】



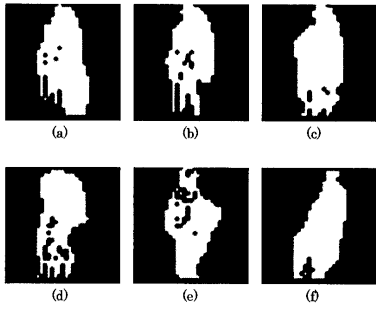
【 図 5 】



【 図 6 】



【 図 7 】



---

フロントページの続き

- (56)参考文献 特開平10-003544(JP,A)  
国際公開第2005/104010(WO,A1)  
特開平11-185055(JP,A)  
特開2004-145818(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00  
G06T 7/20