

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4802330号
(P4802330)

(45) 発行日 平成23年10月26日(2011.10.26)

(24) 登録日 平成23年8月19日(2011.8.19)

(51) Int. Cl. F I
G06T 7/20 (2006.01) G06T 7/20 300Z
G06T 7/00 (2006.01) G06T 7/00 300F

請求項の数 7 (全 13 頁)

(21) 出願番号	特願2006-355691 (P2006-355691)	(73) 特許権者	504174135 国立大学法人九州工業大学 福岡県北九州市戸畑区仙水町1番1号
(22) 出願日	平成18年12月28日(2006.12.28)	(74) 代理人	100090697 弁理士 中前 富士男
(65) 公開番号	特開2008-165581 (P2008-165581A)	(74) 代理人	100139262 弁理士 中嶋 和昭
(43) 公開日	平成20年7月17日(2008.7.17)	(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
審査請求日	平成21年11月25日(2009.11.25)	(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】 動体の動作判別方法

(57) 【特許請求の範囲】

【請求項1】

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の複数のフレーム画像データAがそれぞれ点で表示される固有空間データAを、該基本動作ごとに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、判別手段により、判別しようとする対象動体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と前記基本動作ごとの複数の前記固有空間データA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、前記対象動体の動作を判別する判別工程とを有する動体の動作判別方法であって、

前記基本動作ごとの各フレーム画像データAは、前記動体に該基本動作を行わせ、該基本動作を行う前記動体を複数の画像入力手段を用いて多方向から撮影し、前処理手段によって、前記画像入力手段ごとに取得した連続する複数のフレーム画像を重ね合わせることで作成される圧縮画像から得ることを特徴とする固有空間を用いた動体の動作判別方法。

【請求項2】

請求項1記載の動体の動作判別方法において、前記圧縮画像は、前記複数のフレーム画像のうち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動体の動作判別方法。

【請求項3】

請求項1および2のいずれか1項に記載の動体の動作判別方法において、前記固有空間デ

10

20

ータ作成工程と前記判別工程との間には木構造作成工程が設けられ、前記固有空間データ作成工程でデータベース化された前記固有空間データAを、前記木構造作成工程の木構造作成手段により、前記動体の基本動作が保有している情報ごとに木構造内に分配し、前記記憶手段に格納して構造化することを特徴とする動体の動作判別方法。

【請求項4】

請求項1～3のいずれか1項に記載の動体の動作判別方法において、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、しかも前記複数の画像入力手段は、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ることを特徴とする動体の動作判別方法。

【請求項5】

請求項4記載の動体の動作判別方法において、前記動体は前記疑似人モデルであって、該疑似人モデルの行う基本動作は、モーションキャプチャー法によって得られたものであることを特徴とする固有空間を用いた動体の動作判別方法。

【請求項6】

請求項1～5のいずれか1項に記載の動体の動作判別方法において、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動体の動作判別方法。

【請求項7】

請求項1～6のいずれか1項に記載の動体の動作判別方法において、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動体の動作判別方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、任意方向から観察した動体、特に人の動作を認識可能な動体の動作判別方法に関する。

【背景技術】

【0002】

人とロボットが共存する社会では、人のQOL (Quality of Life) を高めるような機能を持つロボットが求められている。具体的には、例えば、重い荷物を持って移動する人、具合が悪くなって座り込む人、または足元のおぼつかない高齢者等に近づく、適切な支援を与えることが可能な機能を持つロボットである。このようなロボットが、基本的に備えるべき機能は、観察方向によらずに、例えば、人の姿勢、動作、または挙動を認識する機能である。

そこで、動作を認識する方法として、3次元復元または3次元モデルフィッティングに基づいて、3次元で処理する方法が知られている(例えば、特許文献1参照)。

【0003】

【特許文献1】特開2003-58906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、人物の動作認識は数多く研究されているが、いずれも人の正面像、あるいは側面像のみを対象とする手法で、任意方向から観察して得られる画像を前提とする手法は提案されていない。

また、動体の動作を3次元で処理して判別するため、データ量が多くなり、処理時間が膨大となるという問題もあった。

【0005】

本発明はかかる事情に鑑みてなされたもので、観察方向によらずに、自動かつ高速に、対

10

20

30

40

50

象動体の動作を認識可能な動体の動作判別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】

【0006】

前記目的に沿う本発明に係る動体の動作判別方法は、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の複数のフレーム画像データAがそれぞれ点で表示される固有空間データAを、該基本動作ごとに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、判別手段により、判別しようとする対象動体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と前記基本動作ごとの複数の前記固有空間データA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、前記対象動体の動作を判別する判別工程とを有する動体の動作判別方法であって、
前記基本動作ごとの各フレーム画像データAは、前記動体に該基本動作を行わせ、該基本動作を行う前記動体を複数の画像入力手段を用いて多方向から撮影し、前処理手段によって、前記画像入力手段ごとに取得した連続する複数のフレーム画像を重ね合わせることで作成される圧縮画像から得る。

10

【0007】

ここで、圧縮画像とは、複数のフレーム画像を重ね合わせることで、動体の一連の動作が残像として表れた画像を意味する。

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。

20

そして、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。

【0008】

本発明に係る動体の動作判別方法において、前記圧縮画像は、前記複数のフレーム画像のうち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、差分画像とは、例えば、2画像間で減算を行い、変化のない画像部分を削除することにより得られる画像であり、このような変化のない画像部分を削除することで、処理するデータ量を少なくできる。

30

【0009】

本発明に係る動体の動作判別方法において、前記固有空間データ作成工程と前記判別工程との間には木構造作成工程が設けられ、前記固有空間データ作成工程でデータベース化された前記固有空間データAを、前記木構造作成工程の木構造作成手段により、前記動体の基本動作が保有している情報ごとに木構造内に分配し（木構造内に記述し）、前記記憶手段に格納して構造化することが好ましい。

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画像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。

【0010】

40

本発明に係る動体の動作判別方法において、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、しかも前記複数の画像入力手段は、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。

【0011】

本発明に係る動体の動作判別方法において、前記動体は前記疑似人モデルであって、該疑似人モデルの行う基本動作は、モーションキャプチャー法によって得られたものであることが好ましい。

ここで、モーションキャプチャー法とは、例えば、光学式、音波（超音波を含む）式、ま

50

たは磁気式等の非接触で計測可能な機器、あるいは機械式のように接触して計測する機器を使用して、人や動物の動きをコンピュータに取り込んで立体化して再現する方法である。

光学式は、動体に、例えば、発光マーカ、反射マーカ、およびパターンマーカ等のいずれかのマーカを取付けて、複数のカメラ（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、およびデジタルビデオカメラのいずれか1または2以上）で動体を撮影し、この映像をコンピュータに入力して、動体の動きを検知することができる。

【0012】

また、音波式では、動体に取付けた超音波の発信器から発信された超音波を、少なくとも3つの超音波受信器で受信し、超音波受信器へ到達する時間の差によって、動体の動きを検知している。音波式では、動体に超音波受信器を取付けてもよい。

更に、磁気式では、動体に磁界を発生する磁界発生器を取付け、磁界中に配置されると電流が流れる磁力計測器によって、誘起された電流の強さから磁界発生器までの距離と方向とを検知して、動体の動きが再現される。

【0013】

本発明に係る動体の動作判別方法において、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、微分処理は、例えば、ログ（LOG）フィルタまたはソーベルフィルタ等によって行うことができる。ログフィルタは、画像データをぼかして、微分することができる。

【0014】

本発明に係る動体の動作判別方法において、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、カルネン・レーベ変換は、カルネン・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ことができる。

【発明の効果】

【0015】

請求項1～7記載の動体の動作判別方法は、連続する複数のフレーム画像を重ね合わせることで作成される圧縮画像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、基本動作として、動体を多方向から撮影した複数のフレーム画像を用いるので、対象動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。

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、任意方向からの人の動作の認識が可能になる。

【0016】

特に、請求項2記載の動体の動作判別方法においては、基本動作ごとの各フレーム画像データAが、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記載の動体の動作判別方法においては、固有空間データAを木構造内に分配するので、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。

【0017】

請求項4記載の動体の動作判別方法においては、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。

請求項5記載の動体の動作判別方法においては、基本動作ごとの複数のフレーム画像データAを得るための疑似人モデルの基本動作が、モーションキャプチャー法によって得られ

10

20

30

40

50

たものであるので、疑似人モデルの動きが滑らかとなり、対象動体の動きとの誤差が少なくなつて判別し易くなる。

【 0 0 1 8 】

請求項 6 記載の動体の動作判別方法においては、各フレーム画像データ A に微分処理を行つて、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 7 記載の動体の動作判別方法においては、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、固有空間データ B も、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。

10

【 発明を実施するための最良の形態 】

【 0 0 1 9 】

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動体の動作判別方法は、対象動体の一例である対象人が行った動作を、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、対象人がどの基本動作（類似している動作も含む）を行ったかを認識可能な方法である。以下、詳しく説明する。

【 0 0 2 0 】

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。

20

人を中心にして複数台（例えば、4 台）のビデオカメラ（画像入力手段の一例：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。ここで、各ビデオカメラは、等距離で等角度に配置されているが、等距離に配置する必要はなく、また等角度に配置する必要もない。

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。

30

【 0 0 2 1 】

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。

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。

40

【 0 0 2 2 】

このとき、前記した複数のフレーム画像を、連続する 2 つず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、また予めモーションキャプチャー法で解析した疑似人モデルにより作成してもよい。

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

50

方向、上方向、および下方向のいずれか1または2以上に配置される多数（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。

【0023】

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枚の画像データは、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）次元のベクトルで表現できる。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。

10

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。

【0024】

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。

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

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。

20

次に、N行P列の行列Xを、（2）式のように定義する。

$$X = (x_{1-c}, x_{2-c}, \dots, x_{p-c}) \dots (2)$$

なお、Pは、1動作の画像の数である。

ここで、cは画像の平均値であり、（3）式で計算される。

【0025】

【数1】

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

30

【0026】

また、共分散行列Qは、行列Xより（4）式で定義される。

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

【0027】

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。

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

ここで、uはN個の成分を持つベクトルである。

40

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。

【0028】

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、それに対応する固定ベクトル 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。

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

なお、 $k < N$ 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。

50

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

【0029】

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。

10

【0030】

以下、同様に他の方向から撮影した基本動作の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内に分配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-tree、B*-tree、またはB+treeが従来知られている。

【0031】

20

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ビン(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ビ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、複数のビンに分けられ、ビンは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。

【0032】

ここで、B-treeについて説明する。

30

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~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。

40

【0033】

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、木構造群を作成する。

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。

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。

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

50

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。

【0034】

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。

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

そして、前記した固有空間データ A の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。

【0035】

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。

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

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、木構造作成手段でデータベース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。

【0036】

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 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。

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。

【0037】

このように、対象人が行った動作から、その動作に最も近い基本動作を動作データベースの検索により探すことで、対象人がどの基本動作(類似している動作も含む)を行ったかを認識できる。

なお、未知の動作を認識する場合、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。

以上の方法により、データベース検索による動作認識が実現する。

【実施例】

【0038】

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、本願発明の動体の動作判別方法を適用し、対象人が行った動作から、対象人がどの基本動作を行ったかを認識する方法について説明する。

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

10

20

30

40

50

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。

【0039】

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。

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。

10

【0040】

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、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)を減算して削除する。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。

20

【0041】

これにより、各圧縮画像に、動作を認識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(例えば、動体の基本動作を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。

30

【0042】

一方、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、複数のフレーム画像を得る。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、対象人がどの基本動作(類似している動作も含む)を行ったかを認識できる。

40

【産業上の利用可能性】

【0043】

本発明の動体の動作判別方法を、例えば、移動ロボットに搭載することにより、移動ロボットが任意方向から人の動作を認識できるようになる。また、認識方向もわかるので、人物への接近の仕方もわかる。従って、例えば、重い荷物を持って移動する人、具合が悪くなって座り込む人、または足元のおぼつかない高齢者等に近づき、適切な支援を与える機能を持つロボットが実現できる。これにより、安全で安心して暮らせる社会環境が実現できる。

【0044】

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

50

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動体の動作判別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、木構造作成手段による木構造作成工程を行った場合について説明したが、行わなくてもよい。

【図面の簡単な説明】

【0045】

【図1】(A)は人が重た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。

10

【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。

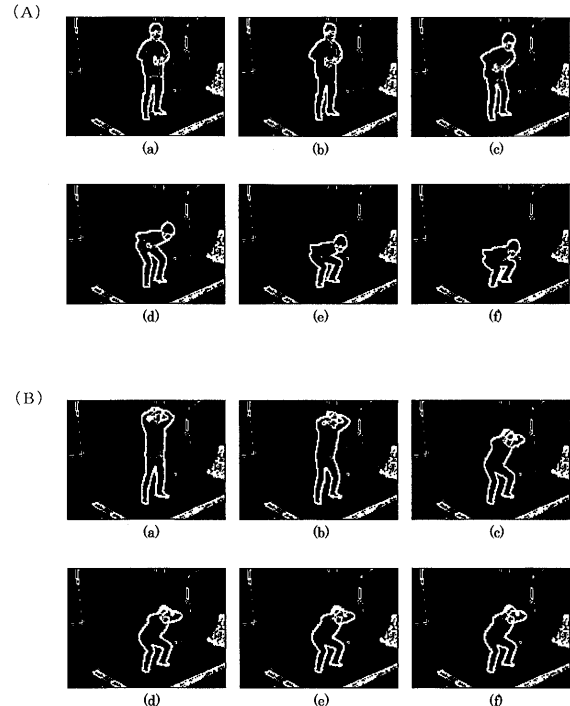
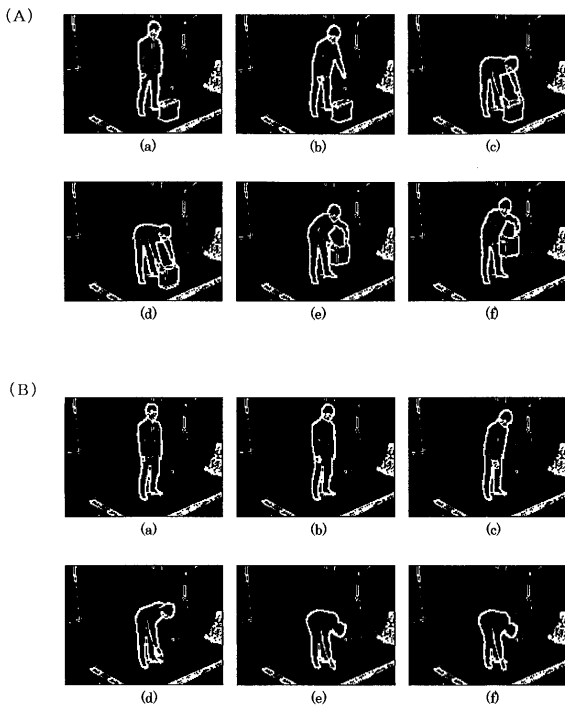
【図6】(A)は図1～図3に示す各基本動作ごとの連続する画像を重ね合わせ変化のない部分を削除した後に得られる圧縮画像の説明図、(B)は(A)の背景画像を削除した後の圧縮画像の説明図である。

20

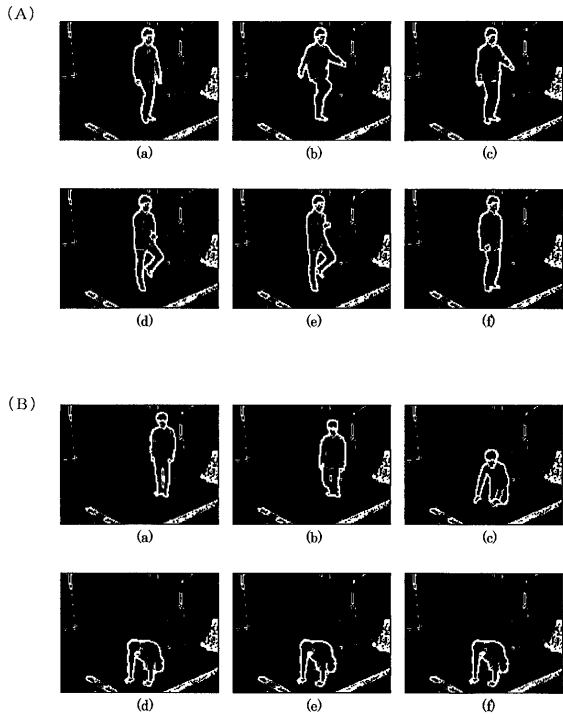
【図7】図6(B)の画像を正規化した画像の説明図である。

【図1】

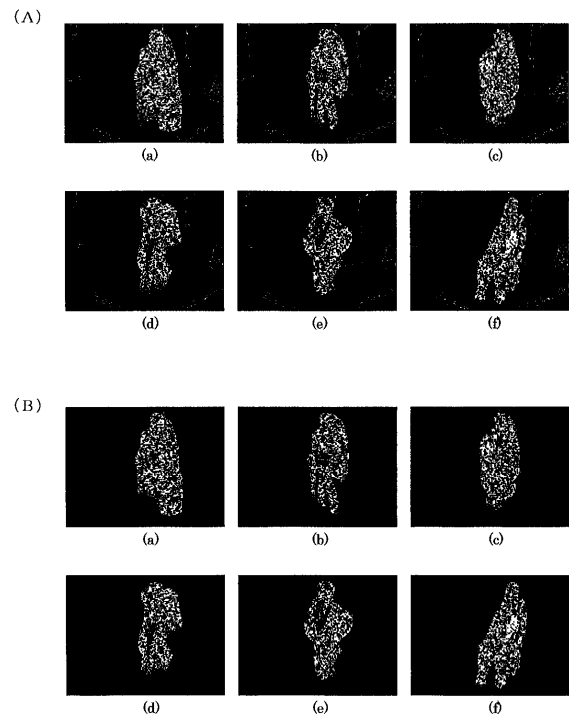
【図2】



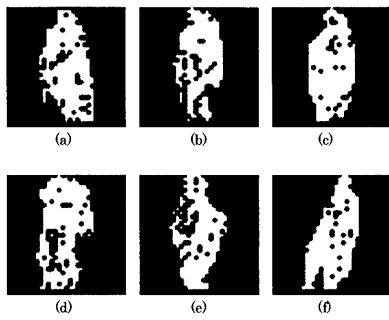
【 図 3 】



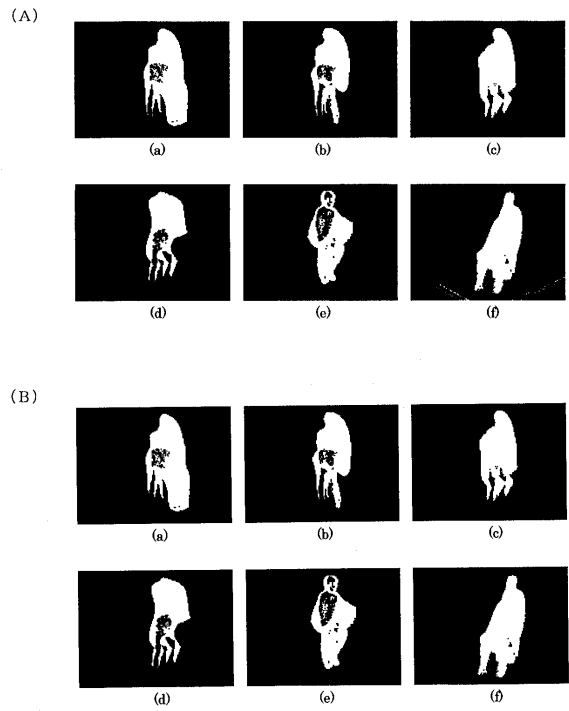
【 図 4 】



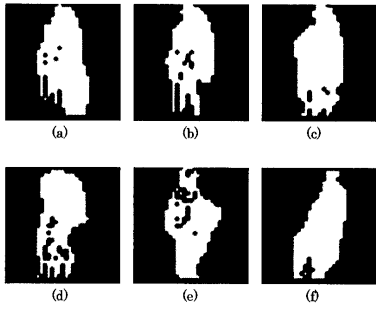
【 図 5 】



【 図 6 】



【 図 7 】



フロントページの続き

- (56)参考文献 特開2005 - 92346 (JP, A)
特開2004 - 220138 (JP, A)
特開2002 - 374520 (JP, A)
特開2004 - 355183 (JP, A)
特開平10 - 334270 (JP, A)
特開2004 - 145818 (JP, A)
特開平10 - 3544 (JP, A)
特開平8 - 329247 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/20
G06T 7/00