

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4564096号
(P4564096)

(45) 発行日 平成22年10月20日 (2010.10.20)

(24) 登録日 平成22年8月6日 (2010.8.6)

(51) Int.Cl. F I
G O 6 T 7/20 (2006.01) G O 6 T 7/20 3 0 0 A

請求項の数 21 (全 17 頁)

(21) 出願番号	特願2009-506714 (P2009-506714)	(73) 特許権者	000005326
(86) (22) 出願日	平成19年4月17日 (2007.4.17)		本田技研工業株式会社
(65) 公表番号	特表2009-533784 (P2009-533784A)		東京都港区南青山二丁目1番1号
(43) 公表日	平成21年9月17日 (2009.9.17)	(74) 代理人	100064414
(86) 国際出願番号	PCT/US2007/066769		弁理士 磯野 道造
(87) 国際公開番号	W02007/121431	(74) 代理人	100111545
(87) 国際公開日	平成19年10月25日 (2007.10.25)		弁理士 多田 悦夫
審査請求日	平成21年9月30日 (2009.9.30)	(72) 発明者	グブタ, ラケシュ
(31) 優先権主張番号	60/792, 917		アメリカ合衆国、オハイオ州 93067
(32) 優先日	平成18年4月17日 (2006.4.17)		-9705、レイモンド、ステート ルー
(33) 優先権主張国	米国 (US)		ト 739 21001、ホンダ アール
(31) 優先権主張番号	11/735, 829		アンド ディ アメリカズ、リーガル
(32) 優先日	平成19年4月16日 (2007.4.16)		デパートメント気付
(33) 優先権主張国	米国 (US)		
早期審査対象出願		最終頁に続く	

(54) 【発明の名称】 物体との相互作用を含む複合動作の分類

(57) 【特許請求の範囲】

【請求項 1】

単純な動作の組み合わせを含む、観察した複合動作を分類するための、コンピュータが実行する方法であって、

前記観察した複合動作を表す画像データを受信するステップと、

前記画像データから特徴ベクトルを生成するステップと、

コンピュータの記憶装置に記憶されている、2つ以上の単純な動作を表すデータを組み合わせることにより、複数のベースラインベクトルを生成するステップと、

前記特徴ベクトルを前処理して、前記特徴ベクトルを、前記複数のベースラインベクトルのうちの一つのベースラインベクトルに近づけるように修正し、修正特徴ベクトルを生成するステップと、

前記複数のベースラインベクトルの一つ一つと前記修正特徴ベクトルとを比較するステップと、

前記観察した複合動作を構成する前記単純な動作の組み合わせを、前記複数のベースラインベクトルのうちの、前記修正特徴ベクトルと最も相関性の高いベースラインベクトルを構成する、記憶されている複数の単純な動作を表すデータとして分類するステップと、を有することを特徴とする方法。

【請求項 2】

前記特徴ベクトルを生成する前記ステップは、

前記複合動作における第一の動きを表す平行移動行列を算出するステップと、

10

20

前記複合動作における第二の動きを表すスケール値を算出するステップと、
前記複合動作における手と物体の接触を表す肌検出値を算出するステップと、
前記平行移動行列、前記スケール値、前記肌検出値を組み合わせて、前記特徴ベクトルを生成するステップと、
を有することを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記平行移動行列は、水平移動値及び垂直移動値を含むことを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記生成された特徴ベクトルを前処理する前記ステップは、
前記特徴ベクトルの長さを均一化するステップと、
前記特徴ベクトルに非線形時間伸縮アルゴリズムを適用するステップと、
を含むことを特徴とする請求項 1 に記載の方法。 10

【請求項 5】

前記特徴ベクトルの前記長さを均一化する前記ステップは、
前記特徴ベクトルが指定された長さを有するように、前記特徴ベクトルにデータを付け加えるステップを含むことを特徴とする請求項 4 に記載の方法。

【請求項 6】

前記非線形時間伸縮アルゴリズムは、相関最適化伸縮法もしくは動的時間伸縮法を含むことを特徴とする請求項 4 に記載の方法。 20

【請求項 7】

前記観察した複合動作を分類する前記ステップは、特異値分解または主成分分析を前記特徴ベクトルに適用するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 8】

基準となる単純な動作を表す単純な動作の画像データを受信するステップと、
前記受信した単純な動作の画像データから基準ベクトルを生成するステップと、
前記基準ベクトルを前記コンピュータの記憶装置に記憶するステップと、
をさらに有することを特徴とする請求項 1 に記載の方法。

【請求項 9】

前記複数のベースラインベクトルを生成する前記ステップは、
前記単純な動作の水平方向の動きを表す水平移動値を算出するステップと、
前記単純な動作の垂直方向の動きを表す垂直移動値を算出するステップと、
前記単純な動作の第 3 の動きを表すスケール値を算出するステップと、
手と物体の接触を表す肌検出値を算出するステップと、
前記水平移動値、前記垂直移動値、前記スケール値、前記肌検出値を組み合わせるステップと、
を有することを特徴とする請求項 8 に記載の方法。 30

【請求項 10】

前記ベースラインベクトルから動きのないデータを取り除くステップをさらに有することを特徴とする請求項 9 に記載の方法。 40

【請求項 11】

前記複数の単純な動作のそれぞれを分類するステップは、
前記修正特徴ベクトルを正規化して、平均 0、分散 1 のベクトルにするステップと、
特異値分解もしくは主成分分析を前記正規化されたベクトルに適用するステップと、
を含むことを特徴とする請求項 1 に記載の方法。

【請求項 12】

前記ベースラインベクトルのそれぞれと前記修正特徴ベクトルを比較する前記ステップは、それぞれのベースラインベクトルと前記修正特徴ベクトルの間の距離ベクトルを算出することを特徴とする請求項 1 に記載の方法。

【請求項 13】

前記修正特徴ベクトルと最も相関性の高い前記ベースラインベクトルは、前記修正特徴ベクトルから最小の距離ベクトルを有するベースラインベクトルを含むことを特徴とする請求項 1 2 に記載の方法。

【請求項 1 4】

単純な動作の組み合わせを含む複合動作を分類するためのコンピュータが実行可能なプログラムであって、

前記プログラムは、

前記観察した複合動作を表す画像データを受信するステップと、

前記画像データから特徴ベクトルを生成するステップと、

2 つ以上の単純な動作を表すデータを組み合わせることにより、複数のベースラインベクトルを生成するステップと、

前記特徴ベクトルを前処理して、前記特徴ベクトルを前記複数のベースラインベクトルのうちの一つのベースラインベクトルに近づけるように修正して、修正特徴ベクトルを形成するステップと、

前記複数のベースラインベクトルの一つ一つと前記修正特徴ベクトルを比較するステップと、

前記観察した複合動作を構成する前記単純な動作の組み合わせを、前記複数のベースラインベクトルのうちの、前記修正特徴ベクトルと最も相関性の高いベースラインベクトルを構成する、記憶されている複数の単純な動作を表すデータとして分類するステップとを含むステップを前記コンピュータに実行させることを特徴とするプログラム。

【請求項 1 5】

特徴ベクトルを生成する前記ステップは、

前記複合動作における第一の動きを表す平行移動行列を算出するステップと、

前記複合動作の第二の動きを表すスケール値を算出するステップと、

前記複合動作における手と物体の接触を表す肌検出値を算出するステップと、

前記平行移動行列、前記スケール値、前記肌検出値を組み合わせ、前記特徴ベクトルを生成するステップと、

を含むことを特徴とする請求項 1 4 に記載のプログラム。

【請求項 1 6】

前記生成された特徴ベクトルを前処理する前記ステップは、

前記特徴ベクトルの長さを均一化するステップと、

前記特徴ベクトルに非線形時間伸縮アルゴリズムを適用するステップと、

を含むことを特徴とする請求項 1 4 に記載のプログラム。

【請求項 1 7】

基準となる単純な動作を表す単純な動作の画像データを受信するステップと、

前記受信した単純な動作の画像データから基準データを生成するステップと、

前記基準データを前記コンピュータの記憶装置に記憶するステップと、

をさらに有することを特徴とする請求項 1 4 に記載のプログラム。

【請求項 1 8】

前記観察した複合動作を分類する前記ステップは、

前記修正特徴ベクトルを平均 0、分散 1 のベクトルに正規化するステップと、

前記正規化されたベクトルに特異値分解もしくは主成分分析を適用するステップと、

を含むことを特徴とする請求項 1 4 に記載のプログラム。

【請求項 1 9】

前記非線形時間短縮アルゴリズムは、相関最適化伸縮法もしくは動的時間短縮を含むことを特徴とする請求項 1 6 に記載のプログラム。

【請求項 2 0】

前記ベースラインベクトルのそれぞれと前記修正特徴ベクトルを比較する前記ステップは、それぞれのベースラインベクトルと前記修正特徴ベクトルの間の距離ベクトルを算出することを特徴とする請求項 1 4 に記載のプログラム。

【請求項 2 1】

前記修正特徴ベクトルと最も相関性の高い前記ベースラインベクトルは、前記修正特徴ベクトルから最小の距離ベクトルを有するベースラインベクトルを含むことを特徴とする請求項 2 0 に記載のプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、米国仮出願番号 60 / 792、917 (出願日：2006 年 4 月 17 日) 及び米国特許出願番号 11 / 735、829 (出願日：2007 年 4 月 16 日) に基づき優先権を主張するものであり、前記出願の全ての記載をここに引用するものである。

10

【0002】

本発明は主にマシン・ラーニングに関するものであり、より詳しくは、観察した動作を特定し分類するシステム及び方法に関するものである。

【背景技術】

【0003】

観察した動作の特定及び分類は、ロボットに動作を実行させるようにトレーニングするための効率の良いメカニズムを提供するため、マシン・ラーニングの用途において関心を集めている。例えば、ロボットは、ある物体と人間の相互作用を観察し、その相互作用を含む動作を特定して記憶することが出来る。ロボットは、後にその記憶されたデータを使用して、その観察した相互作用を複製することが出来る。一般的に複合動作として記述されるいくつかの相互作用は、連続して行われる単純な動作の組み合わせから成り立っている。しかし、これらの複合動作は、特定したり分類したりすることが難しく、ロボットが後に動作実行することが難しい。特に、複合動作のモデリングと認識は、多くの有りうる複合動作の一つ一つに対してデータを収集して記憶する必要があるため、とても手間のかかるものである。

20

【0004】

ロボットは、一般的に、観察した相互作用に関する情報を記憶したデータベースを使用して、観察した相互作用を後に実行する。例えば、データベースは、「手に取る」、「持ち上げる」、「投げる」、「壊す」、「与える」などの、ある物体に対して行われる動作に関するデータを記憶している。その様なデータベースに複合動作に関する情報を記憶するには、広範なトレーニングデータと記憶容量が必要となる。なぜなら、複合動作は連続して行われる多数の単純な動作を含んでいるからである。一方、「押す」、「引く」、「持ち上げる」、「落とす」、「置く」といったような、単純な動作のデータは容易に収集、記憶することが出来る。単純な動作の数は限られており、単純な動作を正確に記述するために必要なトレーニングデータの量も多くはないからである。したがって、単純な動作を記述するデータを記憶し、そして、観察した相互作用に対応する、記憶された単純な動作の組み合わせを判定するほうがより効率的である。この方法は、複合動作を構成する単純な動作を連続して行うことで、複合動作を再現することを可能にする。

30

【発明の開示】

【発明が解決しようとする課題】

40

【0005】

動作認識のための既存の手法としては、アルゴリズムをシステムにハード・コード化し、そのコード化されたアルゴリズムを動画像列に適用して物体間の支持関係、接触関係、接続関係を判定するものがある (Fern, A., et al. 著 「Specific- to-general learning for temporal events with application to learning event definitions from video」、Journal of Artificial Intelligence Research, vol. 17, P. 379-449 参照。なお、この文献に記載される全ての内容をここに引用する)。しかしながら、ハード・コード化されたアルゴリズムを使用することは、認識可能な動作のタイプを制限することになるため、この手法はその適用範囲が限られている。別の手法としては、手の軌道の速さと方向を表すことにより、複合動作を認識するものがある (Rao, C 著 「View-inv

50

ariant representation and recognition of actions」 International Journal of Computer Vision, 50 (2) : P. 203-226 参照。なお、この文献に記載される全ての内容をここに引用する)。しかし、この手法は、ごくわずかな手の動きしか有さない動作(例「落とす」)などの、いくつかのタイプの動作を認識することが出来ず、さらに、物体と手の相互関係をモデルすることができない。さらに別の手法として、人の手と物体に取り付けた無線周波数特定(RFID)タグにより動作を認識するものがある(Philipose, M., et al. 著 「Inferring activities from interactions with objects」 IEEE Pervasive Computing, p. 50-57 参照。なお、この文献に記載される全ての内容をここに引用する)。しかし、RFIDを使用する手法は手間がかかるものであり、観察対象の人と物体に発信機とセンサを搭載する必要がある。これらの既存の手法は、限られた数の複合動作しか認識できず、追加の複合動作を特定するように修正することが容易ではない。

10

【0006】

単純な動作に対応した記憶データを使用して、観察した動作の特定及び分類をするためのシステム及び方法が必要である。

【課題を解決するための手段】

【0007】

本発明は、単純な動作に対応した記憶データを使用して、単純な動作の組み合わせを含む観察された複合動作を特定して分類する方法を提供するものである。一実施形態においては、前記複合動作を表す画像データを受信して、その画像データを前記複合動作に対応する特徴ベクトルを生成するために使用する。次に、単純な動作を表す記憶された基準データが、複数のベースライン基準ベクトルを生成するために使われる。前記複数のベースライン基準ベクトルの1つ1つは、前記単純な動作の組み合わせを表している。例えば、前記特徴ベクトルは、水平移動値、垂直移動値、スケール値、肌検出値を連結したデータを含むことが出来る。前記生成された特徴ベクトルは、その後、前処理されて、基準データに近づくように修正される。前記ベースライン基準ベクトルの1つ1つは、前記修正特徴ベクトルと比較されて、前記観察された複合動作を表す最適な基準ベクトルと単純な動作を判定する。上述の処理が、単純な動作の基準データを使用して、前記複合動作を構成する単純な動作を特定することを可能にしている。

20

【0008】

一実施形態においては、前記特徴ベクトルが指定された長さを有するように、前記特徴ベクトルに動きのないデータを付け加えるなどすることにより、前記特徴ベクトルの長さを均一化することによって、前記特徴ベクトルは前処理される。次に、相関最適化伸縮法もしくは動的時間伸縮法などの非線形時間伸縮アルゴリズムが前記特徴ベクトルに適用される。前記非線形時間伸縮アルゴリズムは、前記特徴ベクトルと単純な動作に対応する基準ベクトルとの間の時間的なずれを補償する。

30

【0009】

本明細書に記載された特徴及び利点は、全てを包括したものではなく、詳細には、図面、明細書及び特許請求の範囲を見た当業者にとって、多くの追加的な特徴及び利点が自明である。さらに、本明細書で用いられた文言は原則として読みやすさ及び説明目的で選択されたものであり、本発明の構成要件を限定したり線引きしたりするために選択されたものではないことに留意すべきである。

40

【発明を実施するための最良の形態】

【0010】

次に、本発明の好ましい実施形態について、図面を参照しながら説明する。なお、図面内の同一の参照番号は、同一の要素もしくは機能的に類似した要素を表すものとする。また、それぞれの参照番号の左端の数字は、その参照番号が最初に使われた図面の番号を表すものとする。

【0011】

明細書中の「一実施形態において」という言葉は、実施形態と関連して説明されるある

50

特定の機能、構造、特徴が本発明の実施形態の少なくとも一つに含まれる、という意味であり、常に同一の実施形態を表すものではない。

【 0 0 1 2 】

以下の詳細な説明には、コンピュータ・メモリ内のデータビット操作のアルゴリズム及び記号表現によって表された部分がある。このようなアルゴリズム記述及び表現は、データ処理技術分野において、当業者が研究内容を他の当業者に最も効率的に伝えるために用いる手段である。あるアルゴリズムがあり、それが概して期待した結果に至る筋の通ったステップの系列だと理解されたとする。それらのステップは、物理量の物理的处理を要するステップである。必ずという訳ではないが、これらの物理量は記憶、転送、結合、比較、その他処理が可能な電氣的または磁氣的信号の形をとるのが普通である。これらの信号をビット、値、要素、記号、文字、術語、番号などで表わすのが、主に慣用上の理由から便利な場合があることが分かっている。さらに、一般性を損なうわけではないが、物理量の物理的处理を要するステップのある特定の配列を、モジュールもしくはコードデバイスなどと表すのが便利な場合があることが分かっている。

10

【 0 0 1 3 】

しかしながら、このような術語や同様の用語はすべて適切な物理量に関連付けられるべきであり、また、それら物理量に付けた便宜上のラベルに過ぎないという ことに留意すべきである。以下の説明より明らかなように、特に断わらない限り、"処理"、"演算"、"計算"、"判定"、"表示"などの術語によって論じられることは、コンピュータシステムのレジスタ及びメモリの内部の物理的（電子的）な量として表現されたデータを処理して、コンピュータシステムのメモリやレジスタ、その他同様の情報記憶装置、情報伝送装置又は表示装置の内部の同様に物理量として表現された他のデータへ変換する、コンピュータシステムや同様の電子的演算装置の作用及びプロセスを意味する。

20

【 0 0 1 4 】

本発明のある側面は、アルゴリズムの形で記述されたプロセス・ステップ及び命令を含んでいる。本発明におけるそれらのプロセス・ステップ及び命令は、ソフトウェア、ファームウェア、もしくはハードウェアのいずれでも実現することが可能であり、ソフトウェアによって実現する場合、ダウンロードして様々なオペレーティングシステム上の各種プラットフォームに駐在させて動作させることも出来る。

【 0 0 1 5 】

本発明は、本明細書に述べた処理を実行するための装置にも関係するものである。このような装置は、所要目的のために専用に作られてもよいし、汎用コンピュータを内蔵のコンピュータ・プログラムによって選択駆動もしくは再構成したものでもよい。そのようなコンピュータ・プログラムは、コンピュータが読み取り可能な記憶媒体、限定するわけではないが例えば、フロッピー（登録商標）ディスク、光ディスク、CD-ROM、光磁気ディスクなどの任意の種類のディスク、リードオンリーメモリ（ROM）やランダムアクセスメモリ（RAM）、EPROM、EEPROM、磁気カード又は光カードなど、コンピュータのシステムバスに接続された電子的命令の記憶に適した任意種類の媒体に格納することができる。さらに、明細書中でコンピュータと呼ばれるものは、1つのプロセッサのみを含むものでもよいし、演算能力の向上のために設計されたマルチプロセッサ構成を実装したものでもよい。

30

40

【 0 0 1 6 】

本明細書で提示したアルゴリズム及び表示は、本質的に、いかなる特定のコンピュータ、その他の装置とも関わりがない。様々な汎用マシンを、本明細書に述べる内容に従ったプログラムで使用し得るが、所要の手順のステップの実行のために、より特化した装置を作るほうが好都合であるかもしれない。そのような多様なシステムに必要とされる構造は以下の説明から明らかになる。さらに、どのような特定のプログラミング言語とも関連付けることなく本発明を説明する。本明細書に述べる本発明の内容を実現するために様々なプログラミング言語を使用し得ることを理解されよう。本明細書における、特定のプログラミング言語に関する言及は、本発明の実施可能性及び最良の実施形態を開示するた

50

めである。

【 0 0 1 7 】

さらに、本明細書における用語は原則的に可読性及び説明の便宜上選択されたものであり、発明の主題を特定、制限するために選択されたものではない。したがって、本発明の開示は説明のためのものであり、請求の範囲に記載された発明の範囲を制限するものではない。

【 0 0 1 8 】

図 1 は、本発明の一実施形態が稼動するコンピュータシステム 1 0 0 を示す図である。コンピュータシステム 1 0 0 は計算装置 1 0 7 と画像取得モジュール 1 0 5 を備えている。計算装置 1 0 7 は、プロセッサ 1 1 0、入力装置 1 2 0、出力装置 1 3 0、メモリ 1 5 0 を備えている。入力装置 1 2 0 は、画像取得モジュール 1 0 5 と通信を行う。一実施形態においては、計算装置 1 0 7 は、画像処理部 1 4 0 もしくはその他の画像処理に特化したプロセッサを備えている。他の実施形態においては、計算装置 1 0 7 は、無線通信機やコネクタを備えたコミュニケーションモジュール 1 6 0 をさらに備えている。

【 0 0 1 9 】

画像取得モジュール 1 0 5 は、ビデオカメラ、もしくは電子的に動作を取り込み可能なその他の動画取り込み装置などであり、物体 1 0 2 の動作を記述する画像データを計算装置 1 0 7 に取り込む。一実施形態においては、画像取得モジュール 1 0 5 は、一秒間に多くの枚数のフレーム（例 5、8、10 フレームなど）を取り込み、継続的に物体 1 0 2 を観察することで、物体 1 0 2 の動きを検出する。一実施形態においては、画像取得モジュール 1 0 5 は、皮膚検出技術もしくはその他の適切な方法を使って、人の手がフレーム中にあるかどうか、また、その人の手が物体 1 0 2 と接触しているかどうかを判定するソフトウェアもしくはファームウェアを備えている。代替りの方法として、画像取得モジュール 1 0 5 は、物体 1 0 2 の動きのデータを計算装置 1 0 7 に送信して、計算装置 1 0 7 が、プロセッサ 1 1 0、画像処理部 1 4 0 もしくはその他の画像処理に特化したプロセッサを使用して、物体 1 0 2 の動きや人の手の存在、その他の物体 1 0 2 の動きを記述するデータを判定する方法もある。

【 0 0 2 0 】

プロセッサ 1 1 0 は、データ信号を処理し、複合命令セットコンピュータ・アーキテクチャ（CISC）、縮小命令セットコンピュータ・アーキテクチャ（RISC）または、命令セットの組み合わせを実装したアーキテクチャを含む様々なコンピューティング・アーキテクチャを（含んでいても）備えていてもよい。図 1 では単一のプロセッサのみが表示されているが、複数のプロセッサを備えていてもよい。プロセッサ 1 1 0 は、算術演算処理ユニット、マイクロプロセッサ、汎用コンピュータ、もしくは、メモリ 1 5 0、入力装置 1 2 0、出力装置 1 3 0 またはコミュニケーションモジュール 1 6 0 からの電子データ信号を送信、受信、処理するために実装されたその他の情報処理装置を含む。

【 0 0 2 1 】

入力装置 1 2 0 は、計算装置 1 0 7 にユーザの入力を伝えるように構成されたものであればどのような機器でもよく、カーソル・コントローラや、キーボードでもよい。一実施形態では、入力装置 1 2 0 は、キーボードなどの英数字入力機器（例 QWERTY キーボード）、キーパッド、もしくは、情報やコマンドの選択をプロセッサ 1 1 0 もしくはメモリ 1 5 0 に伝えるようになされた、タッチセンサー式の画面に表示されたキーボードやキーパッドを含んでいてもよい。他の実施形態においては、入力装置 1 2 0 は、ジョイスティック、マウス、トラックボール、スタイラス、ペン、タッチスクリーン、カーソル指示キー、もしくは画像の移動調整を行う他のメカニズムなどの、コマンドの選択だけでなく、位置データもプロセッサ 1 1 0 に伝えるために搭載されたユーザ入力機器である。

【 0 0 2 2 】

出力装置 1 3 0 は、以下に説明するように、電子画像や電子データを表示するように構成されたあらゆる機器を表すものとする。出力装置 1 3 0 は、例えば、有機発光ダイオード（OLED）、液晶ディスプレイ（LCD）、陰極線管（CRT）ディスプレイ、もし

10

20

30

40

50

くは、同様の機能を備えた任意のディスプレイ機器、スクリーン、モニタなどであってもよい。一実施形態においては、出力装置 130 は、出力装置 130 の画面がタッチセンサ式の透明パネルで覆われたタッチスクリーンも備えている。

【0023】

一実施形態において、計算装置 107 は、画像取得モジュール 105 やメモリ 150 からの入力を受け付ける画像処理部 140 を備えている。画像処理部 140 は、映像データの処理を効率よく行うために、映像データに処理を施すように特別に設計されたプロセッサを備えている。画像処理部 140 は、映像データを処理することで、プロセッサ 110 がより迅速にデータ処理操作を行うことを可能にしている。

【0024】

メモリ 150 は、プロセッサ 110 が実行する命令やデータを記憶している。その命令やデータは、本明細書中に記載されたあらゆる / 全ての技術を実行するためのコードを含んでいてもよい。メモリ 150 は、ダイナミック・ランダム・アクセス・メモリ (DRAM)、スタティック・ランダム・アクセス・メモリ (SRAM)、Flash RAM (非揮発性メモリ)、上記の装置を組み合わせたもの、もしくは当技術分野で公知のその他のメモリ装置のいずれであってもよい。メモリ 150 は、データ格納部 152、プリプロセッサ 154、分類部 156 を備えており、プロセッサ 110、入力装置 120、出力装置 130、画像処理部 140、コミュニケーションモジュール 160 と情報授受を行うように構成されている。

【0025】

データ格納部 152 は、単純な動作を表す基準データを記憶している。例えば、データ格納部 152 は、複数の、基本的もしくは初歩的な単純な動作 (例 「押す」、「引く」、「持ち上げる」、「落とす」、「置く」など) を表すデータを含んでいる。一実施形態においては、その単純な動作を表すデータは、物体 102 の位置の変化などの、その単純な動作がどのように物体 102 に影響を与えているかを記述した特徴ベクトルとして記憶されている。他の実施形態では、データ格納部 152 は、物体 102 に対して行われた単純な動作を表した映像データを含んでいる。上述したデータ格納部 152 に関する記載は単に例であり、データ格納部 152 は、単純な動作を記述、特定することができるあらゆる情報を含むことが出来るものとする。

【0026】

プリプロセッサ 154 は、観察した複合動作と記憶されている基準データをより精確に比較するために、観察した複合動作を表すデータを修正するための情報を提供するものである。複合動作は複数の単純な動作を含んでいるため、複合動作において、その構成要素である単純な動作を完了するために必要な時間は、記憶された基準データが表す単純な動作を完了するために必要な時間とは異なっている。したがって、独立に行われる単純な動作を連結して実行するのに必要な時間は、関連する単純な動作の組み合わせを含む複合動作を実行するのに必要な時間とは異なる。プリプロセッサ 154 は、これらの時間差を補償するために、観察した動作を表すデータを修正するための命令を含んでおり、観察した動作を記憶された基準データとより精確に比較することを可能にしている。例えば、プリプロセッサ 154 は、動的時間伸縮法 (DTW)、相関最適化伸縮法 (COW) もしくは 2 つのデータ (例 基準値と観察値) を非線形的に修正するためのその他の任意の方法をどう実行するかを記述したデータを含んでいる。時間的要素の修正に関して述べてきたが、プリプロセッサ 154 は、観察したデータと記憶された基準データとの比較に影響を与える、観察したデータのあらゆる特徴を補償するための命令を含むことが出来る。

【0027】

分類部 156 は、観察されたデータと記憶された基準データとの比較に関する情報を提供するものであり、観察されたデータに最も類似した基準データを決定するものである。例えば、分類部 156 は、単純な動作を表す基準データを組み合わせることにより複合動作基準ベクトルを生成し、その生成された複合動作基準ベクトルと観察したデータとの間の最小ユークリッド距離を判定するなどすることにより、観察されたデータに最も類似し

10

20

30

40

50

た、(生成された)複合動作基準ベクトルを決定するための命令を含んでいる。他の構成としては、分類部156は、観察された複合動作をその構成要素である単純な動作に分解し、記憶された基準データと、観察された複合動作の構成要素である単純な動作との間の最小ユークリッド距離を判定するなどすることにより、その観察された複合動作を構成する単純な動作に最も類似した基準データを判定するための命令を含んでいる。分類部156に関する上記の記載は単に例であり、分類部156は、観察したデータに最も類似した基準データを判定することが可能な任意の情報を含むことが出来る。

【0028】

一実施形態においては、計算装置107は、計算装置107を図示しないネットワークもしくは他の計算装置107に接続するコミュニケーションモジュール160を備えている。そのネットワークは、ローカル・エリア・ネットワーク(LAN)、ワイド・エリア・ネットワーク(WAN)(例 インターネット)、複数の装置が通信を行う、その他のあらゆる相互接続されたデータ経路を含んでいてもよい。一実施形態においては、コミュニケーションモジュール160は、USB、IEEE1394、イーサネット(登録商標)などの、ファイルや情報を分配するための、他の計算装置107との従来の有線接続である。他の実施形態においては、コミュニケーションモジュール160は、赤外線通信、IEEE802.11a/b/g/n(WiFi)通信、BluetoothR通信、3G通信、IEEE802.16(WiMax)通信、無線通信などのための従来の型の無線通信機である。他の構成としては、コミュニケーションモジュール160は、無線通信機と有線接続を組み合わせたものを備えている。

【0029】

本発明の精神と範囲から逸脱することなく、計算装置107が図1に示したコンポーネントよりも多くコンポーネントを含むことが可能であり、また、少ないコンポーネントしか含んでいないこともあることは、当業者にとっては自明である。例えば、計算装置107は、一次キャッシュもしくは二次キャッシュなどの増設メモリを含んでいてもよいし、一つ以上の特定用途向け集積回路(ASICs)を含んでいてもよい。同様に、計算装置107は、増設された入力装置、出力装置を備えていてもよい。本発明のいくつかの実施形態においては、一つ以上の下記のコンポーネント(110、120、130、140、150、152、154、156、160)をお互いに近接して配置することが可能であり、一方、他の実施形態においては、これらのコンポーネントを地理的に離れた場所に配置することが可能である。例えば、メモリ140におけるコンポーネントは、離れた計算装置107に配置された一つ以上のプロセッサ110により実行可能なプログラムであってもよい。

【0030】

図2は、本発明の一実施形態による、観察した動作を分類するための方法を示したものである。一実施形態において、コンピュータシステム100は、まず、単純な動作を特定する基準データを集める(210)。例えば、画像取得モジュール105を使用して、物体102に対して行われた単純な動作を記録し、その記録されたデータをデータ格納部152に記憶しておくことで、一セットの単純な動作(例 「引く」、「押す」、「落とす」、「持ち上げる」、「置く」など)に対する基準データが集められる(210)。一実施形態においては、1つ1つの単純な動作について複数の特徴ベクトルを生成するために、1つ1つの単純な動作について複数の事例が記録、記憶されて使用される。単純な動作について複数の事例のデータを取得することは、基準データが、単純な動作の実行時間のバリエーションを説明することを可能にする。例えば、単純な動作が実行されて、5-10フレーム/秒もしくは他の適切なフレームレートなどで、複数のフレームの画像データが記録される。その記録されたフレームデータを使用して、水平変換、垂直変換、スケール値、肌検出値などのデータが算出され、その算出された値が連結され、その単純な動作を表す基準ベクトルが生成される。一実施形態においては、基準ベクトルの格納容量を減らすために、データがある閾値を超えて変化しなくなった時などの動きのないデータは基準ベクトルからは除かれる。他の構成としては、計算装置107は、入力装置120、コ

コミュニケーションモジュール 160、または他の情報源から、特徴ベクトルなどの物体 102 に対して行われた単純な動作の効果の記述するデータを受け取る。基準データに使用される単純な動作は、ユーザが上述の処理を行うことで定義することも出来るし、予め決めておき、データ格納部 152 に格納することも可能である。

【0031】

コンピュータシステム 100 は、その後、物体 102 と関わる複合動作を観察する (220)。画像取得モジュール 105 は、複数のフレームの画像データを記録するなどして、物体 102 の動きを記録し、計算装置 107 はその複数のフレームの画像データを使用して、観察した複合動作を現す特徴ベクトルを生成する (230)。例えば、計算装置 107 は、物体 102 の動きを表す変換行列を算出し、平行移動データやスケールデータなどの変換行列からの値を連結して、特徴ベクトルを生成する (230)。一実施形態においては、計算装置 107 は、人の手が物体 102 と接触しているかどうかを判定するために、肌検出技術を使用し、その人の手と物体 102 の接点を表すデータを特徴ベクトルに含める。一実施形態においては、その生成された特徴ベクトルは、水平移動データ、垂直移動データ、スケールデータ、手と物体の境界を表すデータを含む。他の実施形態においては、その生成された特徴ベクトルは、物体 102 の回転を表すデータや、物体 102 の動きを表す追加のデータを含む。

【0032】

その後、記憶された基準データを使用して、単純な動作の組み合わせを表すベースライン基準ベクトルが生成される。一実施形態においては、単純な動作が連結されてベースライン基準ベクトルが生成される (240)。異なるベースライン基準ベクトルは、基準データにより特定される単純な動作の異なる組み合わせを表し、複数のベースライン基準ベクトルが、基準データの単純な動作の複数の組み合わせを表すことを可能にしている。例えば、基準データが、「引く」、「押す」、「落とす」、「持ち上げる」、「置く」などの単純な動作を表している場合、「引く 持ち上げる」、「押す 持ち上げる」、「引く 置く」などの複合動作、もしくは基準データが表す上記の単純な動作の他の組み合わせを表すベースライン基準ベクトルが生成される (260)。一実施形態においては、基準データが表す単純な動作の各組み合わせに対してベースライン基準ベクトルが生成される (240)。他の構成としては、基準データが表す単純な動作のサブセットや指定された数の単純な動作に対するベースライン基準ベクトルが生成される。

【0033】

特徴ベクトルは、生成されたベースライン基準ベクトルと観察した動作の間の時間差を補償するために、前処理される (250)。複合動作を構成する、相互に依存する単純な動作が完了する時間は、連結された、独立に実行される単純な動作が完了する時間とは異なるため、前処理は、これらの時間差を補償するために、ベースライン基準ベクトルと生成された特徴ベクトルを別々に修正する。例えば、前処理 250 は、動的時間伸縮法 (DTW) もしくは相関最適化伸縮法 (COW) などの、時間伸縮技術を使用して、ベースライン基準ベクトルと生成された特徴ベクトルを非線形的に修正する。この非線形的な修正は、ベースライン基準ベクトルと特徴ベクトルの間の最適なマッチの算出を可能にしている。一実施形態においては、前処理 250 は、特徴ベクトルに動きのないデータを付け加えることで、特徴ベクトルが特定の長さを持つようにしている。特徴ベクトルにデータを付け加えることは、複数の特徴ベクトルが同一の長さを有することを可能にし、後続の処理を簡単にすることが出来る。

【0034】

次に、特異値分解 (SVD)、主成分分析 (PCA) または、他の分解技術を使用して、観察した複合動作とベースライン基準ベクトルとを比較することにより、観察した複合動作を構成する単純な動作群を判定する。一実施形態においては、特徴ベクトルは、ベースライン基準ベクトルとの比較の前に、平均 0、分散 1 となるように正規化される。他の実施形態においては、生成されたベースライン基準ベクトルも、平均 0、分散 1 となるように正規化される。前処理された特徴ベクトルの構成要素の動作群は、次に、生成された

ベースライン基準ベクトルと比較され、分類される(260)。一実施形態においては、前処理された特徴ベクトルとベースライン基準ベクトルとの間のユークリッド距離を含む距離ベクトルを算出することにより、複合動作を構成する単純な動作群が分類される。次に、(その複合動作の)構成要素の動作が、前処理された特徴ベクトルから最小のユークリッド距離を有するベースライン基準ベクトルとして分類される(260)。ただし、複合動作を分類するために最小距離を使用することは単に一例であり、前処理された特徴ベクトルとベースライン基準ベクトル間の相互関係を判定することが出来る任意の手法を使用可能である。このことは、観察した複合動作を構成する単純な動作群を、その動作群に最も類似したベースライン基準ベクトルを生成するために組み合わせられた単純な動作群として分類することを可能にする。単純な動作群を含む記憶されたトレーニングデータは、観察した複合動作を分類するために使われる(260)。したがって、複合動作は、単純な動作を連続して行うことによって再現可能である。

【0035】

図3は、本発明の一実施形態による、観察した動作を表す特徴ベクトルを生成する(230)ための方法を示したフローチャートである。まず最初に、物体102の動きを表す変換値が算出される(310)。例えば、計算装置107は、物体102が動かされたとき、線形変換、回転変換、スケール変換などの、変換値を算出する(310)。他の構成としては、物体102の動きを表す変換行列を算出する(310)。この変換行列は、水平移動、垂直移動、回転運動、画像取得モジュール105との接近動作又は離間動作及び他の適当な動きに関するデータを含むことが出来る。一実施形態においては、人の手がフレーム内に存在するかどうか、また、その人の手が物体102と接触しているかどうかを表すデータを生成する(320)ために、肌検出技術が使用される。一実施形態においては、生成された肌検出データは、物体102に対する手の位置を表すフラグを含んでいる。そのフラグが、「落とす」、「置く」などの類似した動作を区別することを可能にしている。テーブル1は、肌検出フラグの一つの使用例を示している。

テーブル1. 肌検出データフラグ値の例

フラグ値	手の位置
0	手がフレームの中にあるが、物体とは接触していない。
-1	手がフレームの中に存在しない。
正数>0	手がフレームの中にあり、物体と接触している。

【0036】

次に、肌検出データと変換によるデータは連結されて特徴ベクトルを形成する(330)。一実施形態においては、算出された変換値のデータのサブセットは、お互いのサブセットと肌検出データ(330)に連結されて特徴ベクトルを形成する。例えば、単純な動作が物体102の角度を変化させない場合は、一般的に回転運動のデータは用いられない。したがって、その特徴ベクトルは、物体102の水平及び垂直の動きを表すデータを含むだけとなる。他の構成としては、特徴ベクトルは、変換により算出された全てのデータ値と肌検出データを含んでいる。

【0037】

図4は、本発明の一実施形態による、基準ベクトルとの比較のために特徴ベクトルを前処理する(250)ための方法を示したフローチャートである。単純な動作を完了するのに要する時間は、実行される動作によって異なるため、異なった動作が実行されたときは、異なった量のデータが取り込まれる。例えば、「落とす」などの動作は、「置く」などの動作よりも短い時間で完了するため、「落とす」という動作が実行されたときは、より少ないデータが記録、記憶される。データ処理をより簡単にするために、特徴ベクトルの長さはほぼ一定の値に均一化される(410)。一実施形態においては、多数の特徴ベク

トルがほぼ同一の長さを持つように、連結された変換データと肌検出データに動きのないデータが付け加えられて、特徴ベクトルの長さが均一化される。

【 0 0 3 8 】

複合動作を構成する、相互に関連する単純な動作を完了させるのに要する時間は、独立に実行される単純な動作を完了するのに要する時間とは異なっている。この時間差は、観察した複合動作とベースライン基準ベクトルとの比較に影響を与え、比較結果を誤らせる可能性がある。したがって、正しく整列されていない時間値を補償するために、その特徴ベクトルと基準ベクトルは伸縮される(420)。一実施形態においては、特徴ベクトルと基準データは、その2つのデータの比較を最適化するために、非線形的に伸縮される(420)。この特徴ベクトルの伸縮処理(420)はまた、特徴ベクトルを均一化するために付け加えられた動きのないデータによって引き起こされるゆがみを最小化する。例えば、動的時間伸縮法(DTW)もしくは相関最適化伸縮法(COW)が、基準ベクトルと特徴ベクトルの時間的要素を別々に拡大・縮小するために使われ、特徴ベクトルと基準ベクトル間の比較の精度を上げている。

【 0 0 3 9 】

図5は、本発明の一実施形態による、相関最適化伸縮法(COW)を使用して観察した複合動作と複合動作基準ベクトルを比較する処理の例を示したものである。COWは、ベースライン基準ベクトルと特徴ベクトルの時間的要素を別々に拡大・縮小することにより、その特徴ベクトルとベースライン基準ベクトルの比較の精度を向上させるものである。

【 0 0 4 0 】

COWは、基準ベクトル510及び特徴ベクトル520などのデータを、一つ以上の時間点を有するセグメント540に分解する。一実施形態においては、セグメント540は同一の長さを有している。他の実施形態においては、セグメント540は、異なった長さ、もしくは、ユーザによって定義された長さを有している。伸縮されるセグメントの柔軟性、つまり、セグメント540の長さの増加量、減少量を特定する「緩み」パラメータも定義される。一実施形態において、「緩み」パラメータは、線形補間を使用して算出され、セグメント540の長さの最大増加値と最大縮小値を指定する。

【 0 0 4 1 】

基準ベクトル510のセグメント540は、特徴ベクトル520のセグメント540と比較される。比較対象である基準ベクトル510のセグメント540と特徴ベクトル520のセグメント540が、お互いに異なる数の時間点を有する場合、線形補間と「緩み」パラメータ、もしくは他の適切な補間手法を使用して、特徴ベクトル520のセグメント540が補間され、伸縮特徴ベクトル530のセグメントが基準ベクトル510のセグメントと同じもしくはほぼ同じ数の時間点を有するように、伸縮特徴ベクトル530が生成される。上記の手法は、特徴ベクトル520のセグメント540の長さを変えることにより、比較されるセグメントを整列させることを可能にしている。

【 0 0 4 2 】

一実施形態においては、特徴ベクトル520のセグメントの長さは多数の「緩み」パラメータによって修正され、それにより複数の異なった伸縮特徴ベクトル530が生成される。伸縮特徴ベクトル530のセグメントは、その後、基準ベクトル520に最も類似する伸縮特徴ベクトル530を判定するために基準ベクトル510のセグメントの長さと相互に関連付けられる。他の構成としては、その相互関係を最適化するために追加のデータも使われる。一実施形態においては、基準ベクトル510と特徴ベクトル520もしくは伸縮特徴ベクトル530との相互関係は、次の式を使用して算出される。

【 0 0 4 3 】

【数1】

$$p(n) = \frac{\text{cov}[r(n), s\{n\}]}{\sqrt{\text{var}[r(n)]\text{var}(s\{n\})}}$$

10

20

30

40

50

上の式で、 r は基準ベクトル 5 1 0、 s は特徴ベクトル 5 2 0、 n はセグメントを示す指標である。

【0044】

上述のように、COWの適用は、セグメント 5 4 0の数と「緩み」パラメータを使用し、比較処理に使用される伸縮特徴ベクトル 5 3 0を生成することで、基準ベクトル 5 1 0と特徴ベクトル 5 2 0の比較を最適化する。一実施形態においては、基準ベクトル 5 1 0及び特徴ベクトル 5 2 0の始点と終点を合わせるために、基準ベクトル 5 1 0及び特徴ベクトル 5 2 0の最初の境界と最後の境界は固定される。これにより、COWの適用が簡単になる。

【0045】

図6は、本発明の一実施形態による、動的時間伸縮法(DTW)を使用して観察した複合動作を基準ベクトルと比較する処理の一例を示したものである。動的時間伸縮法は、基準ベクトル 6 1 0や特徴ベクトル 6 2 0などの2つのデータ値を非線形的に伸縮して、その2つのデータ値の最適なマッチを見つける方法である。

【0046】

一実施形態においては、基準ベクトル 6 1 0は、 N 個のセグメントを含んでおり、特徴ベクトル 6 2 0は M 個のセグメントを含んでいる。基準ベクトル 6 1 0と特徴ベクトル 6 2 0を最適にマッチさせるために、(i 行、 j 列)の要素が、特徴ベクトル 6 2 0の i 番目のセグメントと基準ベクトル 6 1 0の j 番目のセグメントの間のユークリッド距離を含む N 行 M 列の行列が生成される。他の実施形態においては、その行列の(i 行、 j 列)の要素は、特徴ベクトル 6 2 0の i 番目のセグメントと基準ベクトル 6 1 0の j 番目のセグメントとの間の比較を表した任意の値を含むものとする。

【0047】

その行列を使用して、基準ベクトル 6 1 0との比較のために特徴ベクトル 6 2 0を修正する際に使われるワーピングパス 6 3 0が生成される。一実施形態においては、ワーピングパス 6 3 0は、その行列の対角線の一つ上と一つ下のデータを調べることにより生成される。このセットのデータで、特徴ベクトル 6 2 0と基準ベクトル 6 1 0の最も近いマッチ(例 ベクトル間の最小距離、ベクトル間の最大相関値など)を示すデータが、ワーピングパス 6 3 0に含まれる。これらの値が連結されて、伸縮特徴ベクトルを生成するために使用されるワーピングパス 6 3 0が形成される。伸縮特徴ベクトルは、特徴ベクトル 6 2 0にワーピングパス 6 3 0を掛けることで生成される。伸縮特徴ベクトル 6 2 0は、次に、基準ベクトル 6 1 0と比較される。上述のように、ワーピングパス 6 3 0は、特徴ベクトル 6 2 0の時間的要素を拡大もしくは縮小して、特徴ベクトル 6 2 0と基準ベクトル 6 1 0のデータ量の違いを補償するために使用される。

【0048】

特定の実施形態および用途を例示して本発明を説明したが、本発明は、ここに開示された厳密な構成および構成要素のみに限定されるのではなく、本発明の付属の請求項に定義される技術的な思想と範囲を逸脱しない限りにおいて、各種の修正、変更あるいは変形を、本発明の構成、運用、方法および装置の詳細について行うことが可能である。 HYPERLINK "http://www.wipo.int/news/en/newsfeed.jsp"

【図面の簡単な説明】

【0049】

【図1】本発明の一実施形態によるコンピュータシステムを示す図である。

【図2】本発明の一実施形態による、観察した動作を分類するための方法を示す図である。

【図3】本発明の一実施形態による、観察した動作を表す特徴ベクトルを生成するための方法を示したフローチャートである。

【図4】本発明の一実施形態による、基準ベクトルとの比較のために特徴ベクトルを前処理する(250)ための方法を示したフローチャートである。

【図5】本発明の一実施形態による、相関最適化伸縮法(COW)を使用して観察した複

10

20

30

40

50

合動作と 複合動作基準ベクトルを比較する処理の例を示す図である。

【図6】本発明の一実施形態による、動的時間伸縮法(DTW)して観察した複合動作を基準ベクトルと比較する処理の一例を示す図である。

【符号の説明】

【0050】

100 コンピュータシステム

102 物体

105 画像取得モジュール

107 計算装置

110 プロセッサ

10

120 入力装置

130 出力装置

140 画像処理部

150 メモリ

152 データ格納部

154 プリプロセッサ

156 分類部

160 コミュニケーションモジュール

510 基準ベクトル

520 特徴ベクトル

20

530 伸縮特徴ベクトル

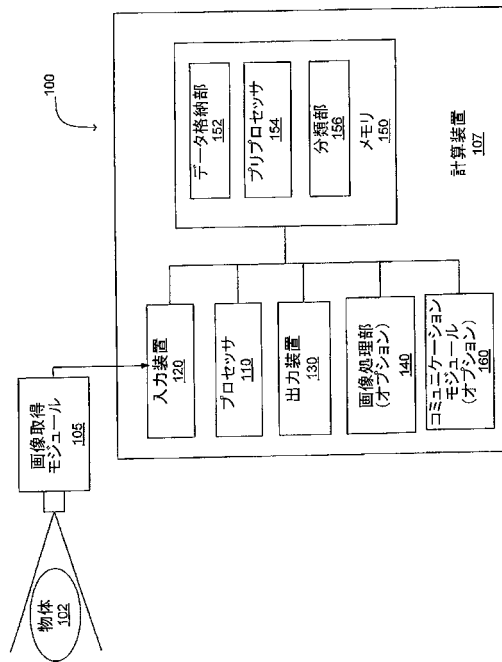
540 セグメント

610 基準ベクトル

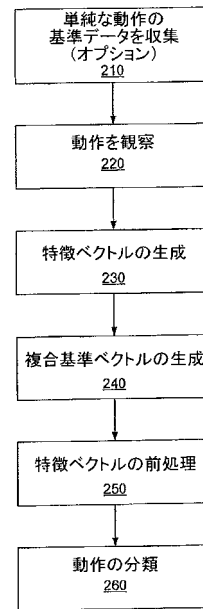
620 特徴ベクトル

630 ワーピングパス

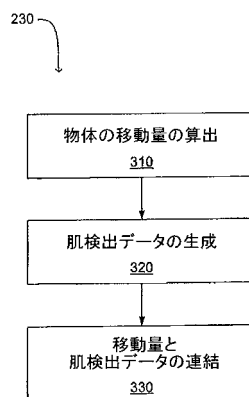
【図 1】



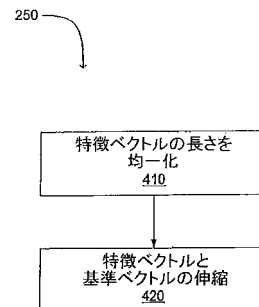
【図 2】



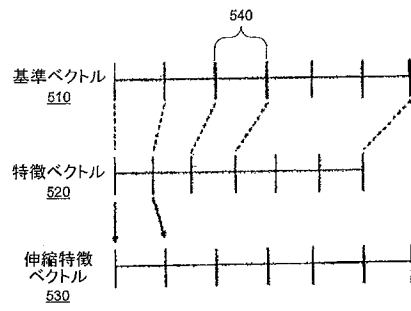
【図 3】



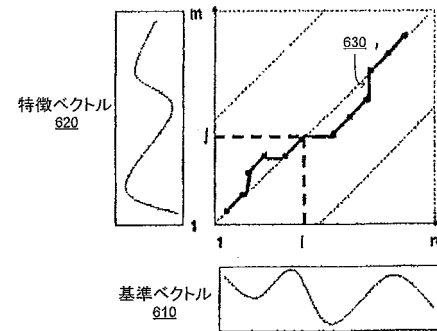
【図 4】



【図 5】



【図 6】



フロントページの続き

(72)発明者 ナヤク, トルプティ, デヴダス

アメリカ合衆国、オハイオ州 93067-9705、レイモンド、ステート ルート 739
21001、ホンダ アール アンド ディ アメリカズ、リーガル デパートメント気付

審査官 松尾 俊介

(56)参考文献 特開2000-181481(JP, A)

特開2004-303251(JP, A)

特開2004-310606(JP, A)

山田博司, 外2名, 連続DPマッチングによる手話の認識, 電子情報通信学会技術研究報告PRMU, 日本, 社団法人電子情報通信学会, 1999年 2月18日, Vol. 98, No. 608, p. 65~72

内田誠一, 外5名, 動作の早期認識およびその予測への応用に関する検討, 電子情報通信学会技術研究報告PRMU, 日本, 社団法人電子情報通信学会, 2004年11月11日, Vol. 104 No. 449, p. 7~12

大崎竜太, 外2名, 速度に基づく切り出しとクラスタリングによる基本動作の抽出, 人工知能学会誌, 日本, 社団法人人工知能学会, 2000年 9月 1日, 第15巻, 第5号, p. 878~886

(58)調査した分野(Int.Cl., DB名)

G06T7/00~7/60