

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7368045号
(P7368045)

(45)発行日 令和5年10月24日(2023.10.24)

(24)登録日 令和5年10月16日(2023.10.16)

(51)国際特許分類	F I
G 0 6 T 7/20 (2017.01)	G 0 6 T 7/20 3 0 0
G 0 6 T 7/00 (2017.01)	G 0 6 T 7/00 3 5 0 C
	G 0 6 T 7/00 6 6 0 B

請求項の数 9 (全18頁)

(21)出願番号	特願2019-140130(P2019-140130)	(73)特許権者	000232092
(22)出願日	令和1年7月30日(2019.7.30)		N E Cソリューションイノベータ株式会
(65)公開番号	特開2021-22323(P2021-22323A)		社
(43)公開日	令和3年2月18日(2021.2.18)		東京都江東区新木場一丁目1 8 番 7 号
審査請求日	令和4年6月6日(2022.6.6)	(74)代理人	110002044
			弁理士法人プライタス
		(72)発明者	谷村 竜太郎
			東京都江東区新木場一丁目1 8 番 7 号
			N E Cソリューションイノベータ株式会
			社内
		(72)発明者	高橋 一徳
			東京都江東区新木場一丁目1 8 番 7 号
			N E Cソリューションイノベータ株式会
			社内
		審査官	片岡 利延

最終頁に続く

(54)【発明の名称】 行動推定装置、行動推定方法およびプログラム

(57)【特許請求の範囲】

【請求項1】

時系列に変化する人物の行動を推定する行動推定装置であって、
時系列の第1画像データから検出された人体部位の時系列の位置データを取得する位置データ取得部と、
前記時系列の位置データを、時系列解析によって複数の位置データ群に分類する分類部と、
前記複数の位置データ群それぞれについて、動作を解析する動作解析部と、
前記動作解析部による解析結果を、時系列に並べて動作シーケンスを生成するシーケンス生成部と、
第2画像データ内で認識された作業に関する作業情報を取得する作業情報取得部と、
前記動作シーケンスに基づいて、前記人物の行動を推定する行動推定部と、
を備え、
前記作業情報取得部は、
前記第2画像データ内の人物の視線とその視線の先にある物体とから、前記人物の行う作業の内容を推定し、推定した前記作業の内容と前記第2画像データ内における前記人物の位置データとを、前記作業情報として取得し、
前記行動推定部は、
前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定する、
ことを特徴とする行動推定装置。

【請求項 2】

請求項 1 に記載の行動推定装置であって、

前記行動推定部は、メモリニューラルネットワーク（Memory Neural Network）を利用して人物の行動を推定する、
ことを特徴とする行動推定装置。

【請求項 3】

請求項 1 に記載の行動推定装置であって、

前記人体部位の時系列の位置データは、前記人物の関節位置を示すデータを含む、
ことを特徴とする行動推定装置。

【請求項 4】

時系列に変化する人物の行動を推定する行動推定方法であって、

時系列の第 1 画像データから検出された人体部位の時系列の位置データを取得し、

前記時系列の位置データを、時系列解析によって複数の位置データ群に分類し

前記複数の位置データ群それぞれについて、動作を解析し

前記動作を解析した結果を、時系列に並べて動作シーケンスを生成し、

第 2 画像データ内で認識された作業に関する作業情報を取得し、

前記動作シーケンスに基づいて、前記人物の行動を推定し、

前記作業情報の取得において、

前記第 2 画像データ内の人物の視線とその視線の先にある物体とから、前記人物の行う作業の内容を推定し、推定した前記作業の内容と前記第 2 画像データ内における前記人物の位置データとを、前記作業情報として取得し、

前記行動の推定において、

前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定する

ことを特徴とする行動推定方法。

【請求項 5】

請求項 4 に記載の行動推定方法であって、

前記人物の行動を推定するときは、メモリニューラルネットワーク（Memory Neural Network）を利用して人物の行動を推定する、

ことを特徴とする行動推定方法。

【請求項 6】

請求項 4 に記載の行動推定方法であって、

前記人体部位の時系列の位置データは、前記人物の関節位置を示すデータを含む、

ことを特徴とする行動推定方法。

【請求項 7】

コンピュータに、時系列に変化する人物の行動を推定させるプログラムであって、

前記コンピュータに、

時系列の第 1 画像データから検出された人体部位の時系列の位置データを取得させ、

前記時系列の位置データを、時系列解析によって複数の位置データ群に分類させ、

前記複数の位置データ群それぞれについて、動作を解析させ、

前記動作を解析した結果を、時系列に並べて動作シーケンスを生成させ、

第 2 画像データ内で認識された作業に関する作業情報として取得させ、

前記動作シーケンスに基づいて、前記人物の行動を推定させる、

命令を含み、

前記作業情報の取得において、

前記第 2 画像データ内の人物の視線とその視線の先にある物体とから、前記人物の行う作業の内容を推定させ、推定した前記作業の内容と前記第 2 画像データ内における前記人物の位置データとを、前記作業情報として取得させ、

前記行動の推定において、

前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定させる、

ことを特徴とするプログラム。

10

20

30

40

50

【請求項 8】

請求項 7 に記載のプログラムであって、

前記人物の行動を推定させるときは、メモリニューラルネットワーク（Memory Neural Network）を利用して人物の行動を推定させる、

ことを特徴とするプログラム。

【請求項 9】

請求項 7 に記載のプログラムであって、

前記人体部位の時系列の位置データは、前記人物の関節位置を示すデータを含む、

ことを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、時系列に変化する人物の行動を推定する行動推定装置および行動推定方法に関し、さらには、これらを実現するためのプログラムに関する。

【背景技術】

【0002】

近年、人物の手を撮像装置で撮像し、時系列の画像データからその人物の手の移動軌跡を検出し、人物が行うジェスチャを認識する技術が提案されている。その技術を利用して、例えば、ジェスチャによる制御装置の操作、または、ジェスチャによるロボットとのコミュニケーションなどを可能としている。

【0003】

しかしながら、手の移動軌跡を検出する場合、短時間の手の位置の変化そのものからどのような動作をしているかを、単純に短時間に区切って判定することは難しい。例えば、手を左右に動かして「バイバイ」する動作では一定期間内での繰り返しが発生するが、単に手を右へ移動させただけの動作もこの動作の中に含まれる。このため、時間の区切り方によっては、検出結果が異なることがある。そこで、特許文献 1 には、周期性の手のジェスチャを認識する際に、ジェスチャの始点と終点とを提示しなくてもジェスチャを認識することができる、ジェスチャ認識装置が開示されている。

【0004】

特許文献 1 に記載のジェスチャ認識装置は、入力された画像データから人間の手の動きの方向を観測信号と捉え、固定長の観測信号列をジェスチャの種類に対応した隠れマルコフモデル（HMM）に入力して尤度値を算出し、尤度値が閾値を越えたものについて対応する HMM からジェスチャの種類を認識するよう構成している。

【先行技術文献】

【特許文献】

【0005】

【文献】特開 2007-087089 号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、特許文献 1 のように、固定長の時間（フレーム数）でジェスチャの種類を認識する場合、その固定長の時間内に、認識したい行動（ジェスチャ）が入らない場合がある。また、同じジェスチャであっても、人によって早い遅いがあるため、固定長の時間で認識する場合、ジェスチャの時間のばらつきに対応できず、十分な認識精度が得られない場合がある。

【0007】

そこで、本発明の目的の一例は、人物の動作の長さに関わらず、その動作状態を解析することができる行動推定装置、行動推定方法およびプログラムを提供することにある。

【課題を解決するための手段】

【0008】

上記目的を達成するため、本発明の一側面における行動推定装置は、
時系列に変化する人物の行動を推定する行動推定装置であって、
時系列の第1画像データから検出された人体部位の時系列の位置データを取得する位置データ取得部と、
前記時系列の位置データを、時系列解析によって複数の位置データ群に分類する分類部と、
前記複数の位置データ群それぞれについて、動作を解析する動作解析部と、
前記動作解析部による解析結果を、時系列に並べて動作シーケンスを生成するシーケンス生成部と、
前記動作シーケンスに基づいて、前記人物の行動を推定する行動推定部と、
を備えることを特徴とする。

10

【0009】

また、上記目的を達成するため、本発明の一側面における行動推定方法は、
時系列に変化する人物の行動を推定する行動推定方法であって、
時系列の第1画像データから検出された人体部位の時系列の位置データを取得し、
前記時系列の位置データを、時系列解析によって複数の位置データ群に分類し
前記複数の位置データ群それぞれについて、動作を解析し
前記動作を解析した結果を、時系列に並べて動作シーケンスを生成し、
前記動作シーケンスに基づいて、前記人物の行動を推定する、
ことを特徴とする。

20

【0010】

更に、上記目的を達成するため、本発明の一側面におけるプログラムは、
コンピュータに、時系列に変化する人物の行動を推定させるプログラムであって、
前記コンピュータに、
時系列の第1画像データから検出された人体部位の時系列の位置データを取得させ、
前記時系列の位置データを、時系列解析によって複数の位置データ群に分類させ、
前記複数の位置データ群それぞれについて、動作を解析させ、
前記動作を解析した結果を、時系列に並べて動作シーケンスを生成させ、
前記動作シーケンスに基づいて、前記人物の行動を推定させる、
命令を含む、ことを特徴とする。

30

【発明の効果】

【0011】

以上のように本発明によれば、人物の動作の長さに関わらず、その動作状態を解析することができる。

【図面の簡単な説明】

【0012】

【図1】図1は、実施形態1の行動推定装置の概略構成を示すブロック図である。
【図2】図2は、実施形態1の行動推定装置の具体的構成を示すブロック図である。
【図3】図3は、行動推定装置による行動の推定を説明するための図である。
【図4】図4は、行動推定装置の動作を示すフロー図である。
【図5】図5は、実施形態2の行動推定装置の具体的構成を示すブロック図である。
【図6】図6は、RGB画像の一例である。
【図7】図7は、動作シーケンスと作業情報とを用いた行動推定を説明するための図である。

40

【図8】図8は、行動推定装置の動作を示すフロー図である。
【図9】図9は、行動推定装置を実現するコンピュータの一例を示すブロック図である。

【発明を実施するための形態】

【0013】

以下、本発明における実施形態の行動推定装置について、図面を参照しながら説明する。

【0014】

50

(実施形態 1)

[装置構成]

図 1 は、実施形態 1 の行動推定装置 10 の概略構成を示すブロック図である。行動推定装置 10 は、時系列に変化する人物の行動を推定する装置である。行動推定装置 10 は、位置データ取得部 1 と、分類部 2 と、動作解析部 3 と、シーケンス生成部 4 と、行動推定部 5 と、を備えている。

【0015】

位置データ取得部 1 は、時系列の第 1 画像データから検出された人体部位の時系列の位置データを取得する。

【0016】

分類部 2 は、位置データ取得部 1 が取得した時系列の位置データを複数に区切り、複数の位置データ群に分類する。

【0017】

動作解析部 3 は、複数の位置データ群それぞれについて、時系列解析によって、動作を解析する。

【0018】

シーケンス生成部 4 は、動作解析部 3 による解析結果を、時系列に並べて動作シーケンスを生成する。

【0019】

行動推定部 5 は、シーケンス生成部 4 が生成した動作シーケンスに基づいて、人物の行動を推定する。

【0020】

この行動推定装置 10 によれば、時系列の動作を短時間毎に区切って動作を解析し、その動作を並べて、行動を推定することで、人物の行動の長さに関わらず、その行動推定を行うことができる。

【0021】

続いて、図 2 を参照して、行動推定装置 10 の構成について具体的説明する。

【0022】

図 2 は、実施形態 1 の行動推定装置 10 の具体的構成を示すブロック図である。図 3 は、行動推定装置 10 による行動の推定を説明するための図である。

【0023】

行動推定装置 10 は撮像装置 50 と接続されている。撮像装置 50 は、例えば、TOF カメラであって、画素毎に深度データが付加された画像データ（以下、TOF 画像と言う）を生成する。なお、撮像装置 50 は、TOF 画像を生成する装置に限定されるものではなく、2 次元画像データを生成する装置であってもよい。

【0024】

撮像装置 50 は人物を撮像する。撮像装置 50 により撮像される人物は、図 3（A）に示すように、動いている。本実施形態では、撮像対象の人物は、立ち止まっている状態から歩きだし、その後、再び立ち止まる動作を行ったものとする。撮像装置 50 は、行動推定装置 10 が備えた装置であってもよいし、行動推定装置 10 とは別装置であってもよい。

【0025】

行動推定装置 10 は、上記した位置データ取得部 1、分類部 2、動作解析部 3、シーケンス生成部 4 および行動推定部 5 に加え、スケルトン認識部 6 を備えている。

【0026】

スケルトン認識部 6 は、撮像装置 50 から TOF 画像を随時取得することで、時系列の第 1 画像データ（以下、時系列データという）を取得する。そして、スケルトン認識部 6 は、時系列データを、時系列のスケルトン画像データに変換する。スケルトン画像データは、例えば、図 3（B）に示すように、関節部分を球体で表し、関節間をつなぐ上腕骨、大腿骨などを棒状体で表して、全体を簡略化した画像データである。

【0027】

10

20

30

40

50

位置データ取得部 1 は、時系列のスケルトン画像データを用いて、人体部位の時系列の位置データを取得する。ここでの人体部位の位置データは、人物の関節位置を示す座標データを含み、例えば、頭部、肩、肘、手首などの座標データである。位置データ取得部 1 により、時系列の位置データを取得することで、人体部位の位置が時系列にどのように変化するかを取得できる。

【0028】

例えば、図 3 (C) に示すように、位置データ取得部 1 は、頭部の位置の時系列変化を取得する。図 3 (C) に示す頭部の位置は、時系列的に最初の頭部の位置を基準とした X 方向の位置である。なお、位置データ取得部 1 は、図 3 (C) において、T O F 画像で設定される座標系の原点を基準とした位置データの変化を取得するようにしてもよい。人体部位の位置データの取得に T O F 画像を用いることで、2 次元画像データを用いた場合と比べて、位置データ取得部 1 は、精度よく人体部位の位置データを取得することができる。

【0029】

なお、スケルトン認識部 6 が行う処理、および、時系列のスケルトン画像データから、人体部位の時系列の位置データを取得する処理は、行動推定装置 10 とは別の装置で行われてもよい。この場合、行動推定装置 10 には、検出された時系列の位置データのみが入力され、位置データ取得部 1 は、入力された時系列の位置データを取得する。

【0030】

分類部 2 は、位置データ取得部 1 が取得した時系列の位置データを、短時間ごとに区切り、複数の位置データ群に分類する。短時間とは、後述する動作解析部 3 が時系列解析を行う際に用いるニューラルネットワークで処理可能な時間である。

【0031】

動作解析部 3 は、教師データ 51 を用いた時系列解析によって、複数の位置データ群それぞれの動作を解析する。つまり、動作解析部 3 は、分類部 2 が区切った短時間ごとに、動作を解析する。時系列解析には、例えば、L S T M (Long short-term memory)、G R U (Gated Recurrent Unit)、R N N (Recurrent Neural Network) などが用いられる。

【0032】

時系列解析の結果には、例えば「移動」、「変化」、「静止」などがある。「移動」は、「一方向への移動」または「ランダムに移動」などを含む。「変化」は、人物が停止した状態で、体の一部を動かした状態をいい、「姿勢変化」、「体幹変化」、「上肢変化」などを含む。また、「姿勢変化」は、「上肢変化」、「体幹変化」、「頭部変化」などを含む。「体幹変化」は、上下変化、角度変化などを含む。「上肢変化」は、手位置変化、肘角度変化などを含む。以下では、短時間ごとに時系列解析が行われた動作を、短時間動作という。

【0033】

なお、教師データ 51 は、行動推定装置 10 が備える記憶装置に記憶されていてもよいし、行動推定装置 10 の外部に設けられた記憶装置に記憶されていてもよい。教師データ 51 は、例えば機械学習により、随時更新される。

【0034】

シーケンス生成部 4 は、短時間動作を、図 3 (D) に示すように、時系列に並べて動作シーケンスを生成する。

【0035】

行動推定部 5 は、シーケンス生成部 4 が生成した動作シーケンスの一または複数の短時間動作から、人物の行動を推定する。行動推定部 5 は、外部メモリ 52 を参照し、例えば、メモリニューラルネットワーク (Memory Neural Network) を利用して行動を推定する。外部メモリ 52 には、推定に用いられる教師データが記憶される。その結果、行動推定部 5 は、図 3 (E) に示すように、「ウロウロしている」、「+5 m 移動した」などの行動を推定する。なお、外部メモリ 52 は、行動推定装置 10 が備えていてもよいし、行動推定装置 10 の外部に設けられていてもよい。

10

20

30

40

50

【0036】

なお、図3の説明では、人物の「頭部」の位置データを用いているが、他の部位の位置データを用いてもよいし、複数の部位の位置データを組み合わせてもよい。

【0037】

[装置動作]

次に、本実施形態における行動推定装置10の動作について図4を用いて説明する。図4は、行動推定装置10の動作を示すフロー図である。以下の説明においては、適宜図1～図3を参照する。また、本実施形態では、行動推定装置10を動作させることによって、行動推定方法が実施される。よって、本実施形態における行動推定方法の説明は、以下の行動推定装置10の動作説明に代える。

【0038】

前提として、撮像装置50によって人物の動作が撮像されているものとする。この前提において、スケルトン認識部6は、撮像装置50からTOF画像を随時取得することで、時系列データを取得する。(S1)。次に、スケルトン認識部6は、取得した時系列データを、時系列のスケルトン画像データへ変換する(S2)。

【0039】

次に、位置データ取得部1は、スケルトン認識部6によるスケルトン認識の結果を用いて、画像データ内の人体部位の位置データを取得する(S3)。そして、分類部2は、位置データ取得部1が取得した時系列の位置データを、短時間ごとに区切り、複数の位置データ群に分類する(S4)。

【0040】

動作解析部3は、RNNなどを利用して、複数の位置データ群それぞれに対し、時系列解析を行う(S5)。シーケンス生成部4は、短時間動作を時系列に並べて動作シーケンスを生成する(S6)。行動推定部5は、S6で生成された動作シーケンスに対して、MemoryNNを利用して、人物の行動を推定する(S7)。

【0041】

[実施形態1の効果]

以上のように、行動推定装置10は、人物の動作を、短時間ごとに、動作を分類している。そして、行動推定装置10は、短時間ごとの動作を時系列に並べて動作シーケンスを生成し、人物の行動を推定する。人物の行動時間が長時間におよぶ場合、行動推定装置10では、短時間の動作が時系列につながられた情報から行動を推定することで、推定が容易となり、推定結果の精度は向上する。つまり、本実施形態によれば、人物の行動時間の長さに関わらず、その行動を推定することができる。

【0042】

(実施形態2)

次に、本発明における実施形態2について説明する。実施形態2は、人物の行動を推定する際、画像データから検出される作業情報を考慮する点で、実施形態1と相違する。以下、相違点について説明する。なお、実施形態1と同じ構成については、同じ符号を付して、その詳細な説明は省略する。

【0043】

[装置構成]

図5は、実施形態2の行動推定装置20の具体的構成を示すブロック図である。行動推定装置20は、撮像装置50および撮像装置53に接続されている。撮像装置53は、各画素がRGBの輝度を含む第2画像データ(以下、RGB画像という)を生成する装置である。なお、撮像装置53は、モノクロ画像データを生成する装置であってもよい。行動推定装置20は、撮像装置50からTOF画像を取得し、撮像装置53からRGB画像を取得する。撮像装置50、53は同じ視野角で、同じ範囲を撮像するように配置されている。撮像装置50は随時撮像し、TOF画像を生成する。一方、撮像装置53は、任意のタイミングで撮像し、RGB画像を生成する。なお、一の撮像装置が、RGB画像とTOF画像とを生成する機能を備えていてもよい。

【0044】

行動推定装置20は、実施形態1で説明した、位置データ取得部1、分類部2、動作解析部3、シーケンス生成部4、行動推定部5およびスケルトン認識部6に加え、作業情報取得部7を備えている。

【0045】

作業情報取得部7は、撮像装置53からRGB画像を取得する。作業情報取得部7は、取得したRGB画像から認識される作業に関する作業情報を取得する。作業情報は、RGB画像内で認識された物体（人物、設備を含む）を特定する情報、または、RGB画像内で認識された人物が行っている作業を特定する情報である。

【0046】

図6は、RGB画像の一例である。例えば、作業情報取得部7は、図6のRGB画像から、人物（A）、人物（B）、パソコン机（C）、パソコン机（D）、扉（E）、およびホワイトボード（F）などの物体を検出し、それらの位置データとともに、作業情報として取得する。また、作業情報取得部7は、検出された人物と設備との距離から、その人物が行っている作業内容を推定し、その作業が行われている画像内における位置データとともに、作業情報として取得する。例えば、図6のRGB画像において、作業情報取得部7は、近傍領域にある人物（B）とパソコン机（C）とから、「パソコン作業中」とする作業内容を推定し、その作業が行われている画像内における位置データとともに、作業情報として取得する。

【0047】

図5に戻る。作業情報取得部7は、撮像装置50からTOF画像を取得する。作業情報取得部7は、取得したTOF画像とRGB画像とを併用して、検出した作業情報の位置データに深度データをさらに付与する。作業情報の位置データに深度データを付与することで、作業内容の推定誤差を低減できる。例えば、二次元の位置データのみで見たときに、パソコンと人物との位置が近い場合であっても、奥行距離が離れている場合がある。このような場合に、「パソコン作業中」とした誤った作業内容を推定することが低減できる。

【0048】

なお、作業情報取得部7が行う作業情報を取得する処理、および、深度データを付与する処理は、行動推定装置20とは別の装置で行われてもよい。この場合、行動推定装置20には処理結果が入力され、作業情報取得部7は、入力された処理結果を取得する。

【0049】

また、撮像装置50と、撮像装置53とでは、視野角が同じでも設置位置が異なるため、RGB画像とTOF画像とでは若干の位置ずれが生じる。このため、作業情報取得部7は、一方の座標系を他方に合わせる変換を行うなど、ずれを修正している。この修正処理は、一の撮像装置が、RGB画像とTOF画像とを生成する機能を備えている場合には、その撮像装置内で行われてもよい。

【0050】

スケルトン認識部6は、撮像装置50からTOF画像を随時取得することで、時系列データを取得し、時系列のスケルトン画像データに変換する。

【0051】

位置データ取得部1は、時系列のスケルトン画像データを用いて、人体部位の時系列の位置データを取得する。

【0052】

分類部2は、位置データ取得部1が取得した時系列の位置データを、短時間ごとに区切り、複数の位置データ群に分類する。

【0053】

動作解析部3は、位置データ群それぞれに対し、教師データ51を用いた時系列解析によって、複数の位置データ群それぞれの動作を分類する。そして、動作解析部3は、教師データ51を用いた時系列解析によって、複数の位置データ群それぞれの動作を解析する。

【0054】

シーケンス生成部 4 は、短時間動作を、時系列に並べて動作シーケンスを生成する。

【0055】

行動推定部 5 は、シーケンス生成部 4 が生成した動作シーケンスと、作業情報取得部 7 が取得した作業情報とに基づいて、人物の行動を推定する。図 7 は、動作シーケンスと作業情報とを用いた行動推定を説明するための図である。図 7 において、動作シーケンスを構成する短時間動作は、「静止」、「移動」、「移動」、「静止」、「頭部変化」、「上肢変化」である。

【0056】

まず、行動推定部 5 は、外部メモリ 52 を参照し、例えば、MemoryNN を利用し、動作シーケンスの三つの短時間動作「移動」、「移動」、「静止」から、動作状態「移動停止」を推定する。また、行動推定部 5 は、二つの短時間動作「頭部変化」および「上肢変化」から、動作状態「着席」を推定する。

【0057】

次に、行動推定部 5 は、動作シーケンス、作業情報、および動作状態から、行動を推定する。例えば、行動推定部 5 は、動作状態「移動停止」と、停止した位置の位置データを含む作業情報「パソコン机」とから、「パソコン前へ移動」という行動を推定する。また、行動推定部 5 は、動作シーケンスの短時間動作「上肢変化」と、動作状態「着席」と、着席時の位置の位置データを含む作業情報「パソコン机」とから、「パソコン操作」という行動を推定する。

【0058】

〔装置動作〕

次に、本実施形態における行動推定装置 20 の動作について図 8 を用いて説明する。図 8 は、行動推定装置 20 の動作を示すフロー図である。本実施形態では、行動推定装置 20 を動作させることによって、行動推定方法が実施される。よって、本実施形態における行動推定方法の説明は、以下の行動推定装置 20 の動作説明に代える。

【0059】

前提として、撮像装置 50、53 が同視野角で撮像しているものとする。この前提において、スケルトン認識部 6 は、撮像装置 50 から TOF 画像を随時取得することで、時系列データを取得する（S11）。次に、スケルトン認識部 6 は、取得した時系列データを、時系列のスケルトン画像データへ変換する（S12）。

【0060】

次に、位置データ取得部 1 は、スケルトン認識部 6 によるスケルトン認識の結果を用いて、画像データ内の人体部位の位置データを取得する（S13）。そして、分類部 2 は、位置データ取得部 1 が取得した時系列の位置データを、短時間ごとに区切り、複数の位置データ群に分類する（S14）。

【0061】

動作解析部 3 は、RNN などを利用して、複数の位置データ群それぞれに対し、時系列解析を行う（S15）。シーケンス生成部 4 は、短時間動作を時系列に並べて動作シーケンスを生成する（S16）。作業情報取得部 7 は、撮像装置 50 から TOF 画像を取得し、撮像装置 53 から RGB 画像を取得する（S17）。作業情報取得部 7 は、RGB 画像および TOF 画像とから作業情報を取得する（S18）。そして、図 7 で説明したように、行動推定部 5 は、S16 で生成した動作シーケンスと、S18 で取得した作業情報とから、人物の行動を推定する（S19）。

【0062】

〔実施形態 2 の効果〕

このように、行動推定装置 20 は、作業情報を併用して行動を推定することで、動作シーケンスのみで行動を推定した場合と比べて、より詳細な行動を推定することができる。また、作業情報の位置データに深度データを付与することで、画像データ内の奥行を考慮した行動推定を行うことができる。例えば、図 7 の説明において、二次元の位置データのみで推定したとき、停止位置と、パソコン机との奥行距離が離れている場合がある。この

10

20

30

40

50

ような場合に、「パソコン前へ移動」という誤った行動を推定することが低減できる。

【0063】

〔実施形態2の変形例〕

なお、作業情報に深度データを付与せずに、行動を推定してもよい。この場合、RGB画像から認識した物体同士の大きさなどを考慮して、物体の奥行距離を推定し、人物の行動を推定することができる。

【0064】

また、作業情報取得部7は、RGB画像内の人物の視線と、その視線先にある物体とから作業情報を取得してもよい。例えば、作業情報取得部7は、人物の視線の先にテレビがある場合、「テレビ視聴中」とする作業内容を推定し、画像内における、その人物の位置データとともに、作業情報として取得する。

【0065】

〔プログラム〕

上記の実施形態におけるプログラムは、コンピュータに、図4または図8に示す処理を実行させるプログラムであればよい。このプログラムをコンピュータにインストールし、実行することによって、上記の実施形態における行動推定装置と行動推定方法とを実現することができる。この場合、コンピュータのプロセッサは、位置データ取得部1、分類部2、動作解析部3、シーケンス生成部4、行動推定部5および作業情報取得部7として機能し、処理を行なう。

【0066】

また、上記の実施形態では、位置データ取得部1、分類部2、動作解析部3、シーケンス生成部4、行動推定部5および作業情報取得部7は、コンピュータに備えられたハードディスク等の記憶装置に、これらを構成するデータファイルを格納することによって、または、このデータファイルが格納された記録媒体をコンピュータと接続された読取装置に搭載することによって実現されている。

【0067】

また、上記の実施形態におけるプログラムは、複数のコンピュータによって構築されたコンピュータシステムによって実行されてもよい。この場合は、例えば、各コンピュータが、それぞれ、位置データ取得部1、分類部2、動作解析部3、シーケンス生成部4、行動推定部5および作業情報取得部7のいずれかとして機能してもよい。

【0068】

〔物理構成〕

ここで、本実施形態におけるプログラムを実行することによって、行動推定装置を実現するコンピュータについて図9を用いて説明する。図9は、行動推定装置を実現するコンピュータの一例を示すブロック図である。

【0069】

図9に示すように、コンピュータ110は、CPU111と、メインメモリ112と、記憶装置113と、入力インターフェイス114と、表示コントローラ115と、データリーダー／ライター116と、通信インターフェイス117とを備える。これらの各部は、バス121を介して、互いにデータ通信可能に接続される。なお、コンピュータ110は、CPU111に加えて、またはCPU111に代えて、GPU (Graphics Processing Unit)、またはFPGA (Field-Programmable Gate Array) を備えていてもよい。

【0070】

CPU111は、記憶装置113に格納された、本実施形態におけるプログラム（コード）をメインメモリ112に展開し、これらを所定順序で実行することにより、各種の演算を実施する。メインメモリ112は、典型的には、DRAM (Dynamic Random Access Memory) 等の揮発性の記憶装置である。また、本実施形態におけるプログラムは、コンピュータ読み取り可能な記録媒体120に格納された状態で提供される。なお、本実施形態におけるプログラムは、通信インターフェイス117を介して接続されたインターネット上で流通するものであってもよい。

10

20

30

40

50

【0071】

また、記憶装置113の具体例としては、ハードディスクドライブの他、フラッシュメモリ等の半導体記憶装置が挙げられる。入力インターフェイス114は、CPU111と、キーボードおよびマウスといった入力機器118との間のデータ伝送を仲介する。表示コントローラ115は、ディスプレイ装置119と接続され、ディスプレイ装置119での表示を制御する。

【0072】

データリーダ／ライタ116は、CPU111と記録媒体120との間のデータ伝送を仲介し、記録媒体120からのプログラムの読み出し、およびコンピュータ110における処理結果の記録媒体120への書き込みを実行する。通信インターフェイス117は、CPU111と、他のコンピュータとの間のデータ伝送を仲介する。

10

【0073】

また、記録媒体120の具体例としては、CF（Compact Flash（登録商標））およびSD（Secure Digital）等の汎用的な半導体記憶デバイス、フレキシブルディスク（Flexible Disk）等の磁気記録媒体、またはCD-ROM（Compact Disk Read Only Memory）などの光学記録媒体が挙げられる。

【0074】

なお、上記の実施形態における行動推定装置は、プログラムがインストールされたコンピュータではなく、各部に対応したハードウェアを用いることによっても実現可能である。更に、行動推定装置は、一部がプログラムで実現され、残りの部分がハードウェアで実現されていてもよい。

20

【0075】

上述した実施形態の一部または全部は、以下に記載する（付記1）～（付記21）によって表現することができるが、以下の記載に限定されるものではない。

【0076】

（付記1）

時系列に変化する人物の行動を推定する行動推定装置であって、
時系列の第1画像データから検出された人体部位の時系列の位置データを取得する位置データ取得部と、
前記時系列の位置データを、時系列解析によって複数の位置データ群に分類する分類部と、
前記複数の位置データ群それぞれについて、動作を解析する動作解析部と、
前記動作解析部による解析結果を、時系列に並べて動作シーケンスを生成するシーケンス生成部と、
前記動作シーケンスに基づいて、前記人物の行動を推定する行動推定部と、
を備えることを特徴とする行動推定装置。

30

【0077】

（付記2）

付記1に記載の行動推定装置であって、
第2画像データ内で認識された作業に関する作業情報を取得する作業情報取得部、
をさらに備え、
前記行動推定部は、
前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定する、
ことを特徴とする行動推定装置。

40

【0078】

（付記3）

付記2に記載の行動推定装置であって、
前記作業情報は、
前記第2画像データ内で認識された物体を特定する情報、および、前記第2画像データ内で認識された人物が行っている作業を特定する情報、の少なくとも一方を含む、

50

ことを特徴とする行動推定装置。

【0079】

(付記4)

付記2または付記3に記載の行動推定装置であって、
前記作業情報は深度データを有する、
ことを特徴とする行動推定装置。

【0080】

(付記5)

付記1から付記4のいずれか一つに記載の行動推定装置であって、
前記行動推定部は、メモリニューラルネットワーク (Memory Neural Network) を利用して人物の行動を推定する、
ことを特徴とする行動推定装置。

10

【0081】

(付記6)

付記1から付記5のいずれか一つに記載の行動推定装置であって、
前記位置データは、前記人物の関節位置を示すデータを含む、
ことを特徴とする行動推定装置。

【0082】

(付記7)

付記1から付記6までのいずれか一つに記載の行動推定装置であって、
前記第1画像データは、画素毎の深度データが付加された画像データを時系列にしたデータである、
ことを特徴とする行動推定装置。

20

【0083】

(付記8)

時系列に変化する人物の行動を推定する行動推定方法であって、
(A) 時系列の第1画像データから検出された人体部位の時系列の位置データを取得するステップと、
(B) 前記時系列の位置データを、時系列解析によって複数の位置データ群に分類するステップと、
(C) 前記複数の位置データ群それぞれについて、動作を解析するステップと、
(D) 前記ステップ(C)での解析結果を、時系列に並べて動作シーケンスを生成するステップと、
(E) 前記動作シーケンスに基づいて、前記人物の行動を推定するステップと、
を備えることを特徴とする行動推定方法。

30

【0084】

(付記9)

付記8に記載の行動推定方法であって、
(F) 第2画像データ内で認識された作業に関する作業情報を取得するステップ、
をさらに備え、
前記ステップ(E)では、
前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定する、
ことを特徴とする行動推定方法。

40

【0085】

(付記10)

付記9に記載の行動推定方法であって、
前記作業情報は、
前記第2画像データ内で認識された物体を特定する情報、および、前記第2画像データ内で認識された人物が行っている作業を特定する情報、の少なくとも一方を含む、
ことを特徴とする行動推定方法。

50

【 0 0 8 6 】

(付記 1 1)

付記 9 または付記 1 0 に記載の行動推定方法であって、
前記作業情報は深度データを有する、
ことを特徴とする行動推定方法。

【 0 0 8 7 】

(付記 1 2)

付記 8 から付記 1 1 のいずれか一つに記載の行動推定方法であって、
前記ステップ (E) では、メモリニューラルネットワーク (Memory Neural Network)
) を利用して人物の行動を推定する、
ことを特徴とする行動推定方法。

10

【 0 0 8 8 】

(付記 1 3)

付記 8 から付記 1 2 のいずれか一つに記載の行動推定方法であって、
前記位置データは、前記人物の関節位置を示すデータを含む、
ことを特徴とする行動推定方法。

【 0 0 8 9 】

(付記 1 4)

付記 8 から付記 1 3 までのいずれか一つに記載の行動推定方法であって、
前記第 1 画像データは、画素毎の深度データが付加された画像データを時系列にしたデ
ータである、
ことを特徴とする行動推定方法。

20

【 0 0 9 0 】

(付記 1 5)

コンピュータに、時系列に変化する人物の行動を推定させるプログラムであって、
前記コンピュータに、
(A) 時系列の第 1 画像データから検出された人体部位の時系列の位置データを取得する
ステップと、
(B) 前記時系列の位置データを、時系列解析によって複数の位置データ群に分類するス
テップと、
(C) 前記複数の位置データ群それぞれについて、動作を解析するステップと、
(D) 前記ステップ (C) での解析結果を、時系列に並べて動作シーケンスを生成するス
テップと、
(E) 前記動作シーケンスに基づいて、前記人物の行動を推定するステップと、
を実行させることを特徴とするプログラム。

30

【 0 0 9 1 】

(付記 1 6)

付記 1 5 に記載のプログラムであって、
前記コンピュータに、
(F) 第 2 画像データ内で認識された作業に関する作業情報を取得するステップ、
をさらに実行させ、
前記ステップ (E) では、
前記動作シーケンスと、前記作業情報とを併用して前記人物の行動を推定する、
ことを特徴とするプログラム。

40

【 0 0 9 2 】

(付記 1 7)

付記 1 6 に記載のプログラムであって、
前記作業情報は、
前記第 2 画像データ内で認識された物体を特定する情報、および、前記第 2 画像デー
タ内で認識された人物が行っている作業を特定する情報、の少なくとも一方を含む、

50

ことを特徴とするプログラム。

【0093】

(付記18)

付記16または付記17に記載のプログラムであって、
前記作業情報は深度データを有する、
ことを特徴とするプログラム。

【0094】

(付記19)

付記15から付記18のいずれか一つに記載のプログラムであって、
前記ステップ(E)では、メモリニューラルネットワーク(Memory Neural Network
)を利用して人物の行動を推定する、
ことを特徴とするプログラム。

10

【0095】

(付記20)

付記15から付記19のいずれか一つに記載のプログラムであって、
前記位置データは、前記人物の関節位置を示すデータを含む、
ことを特徴とするプログラム。

【0096】

(付記21)

付記15から付記20までのいずれか一つに記載のプログラムであって、
前記第1画像データは、画素毎の深度データが付加された画像データを時系列にしたデ
ータである、
ことを特徴とするプログラム。

20

【符号の説明】

【0097】

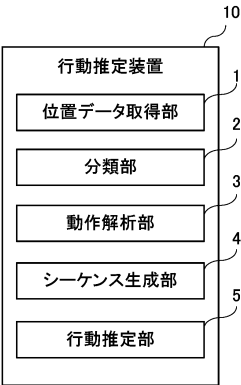
- 1 位置データ取得部
- 2 分類部
- 3 動作解析部
- 4 シーケンス生成部
- 5 行動推定部
- 6 スケルトン認識部
- 7 作業情報取得部
- 10、20 行動推定装置
- 50、53 撮像装置
- 51 教師データ
- 52 外部メモリ

30

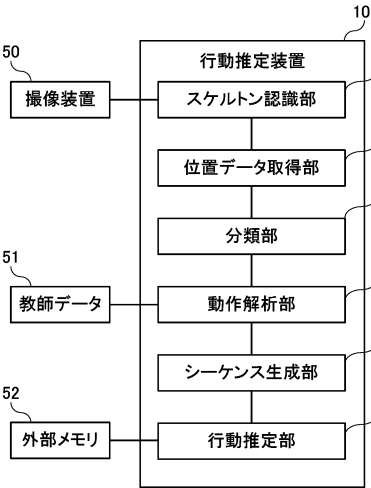
40

50

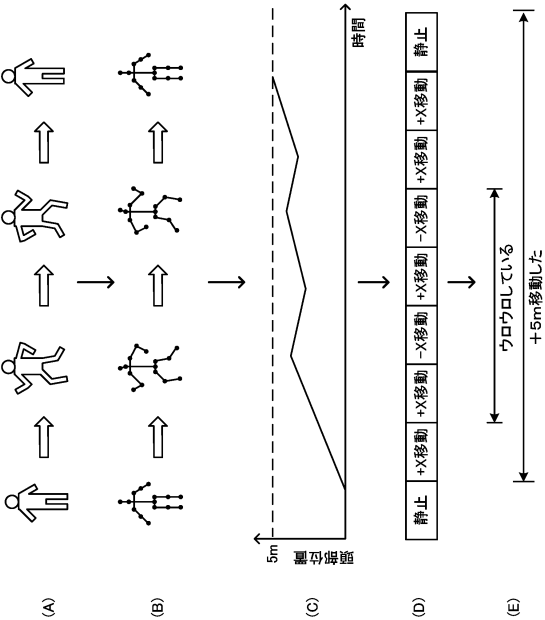
【図面】
【図 1】



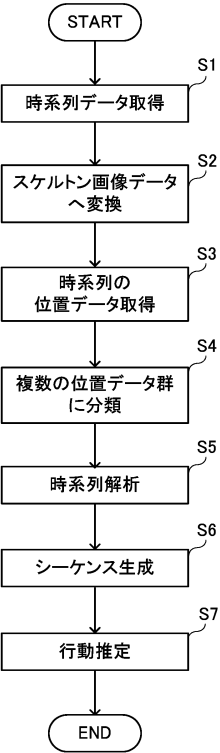
【図 2】



【図 3】



【図 4】



10

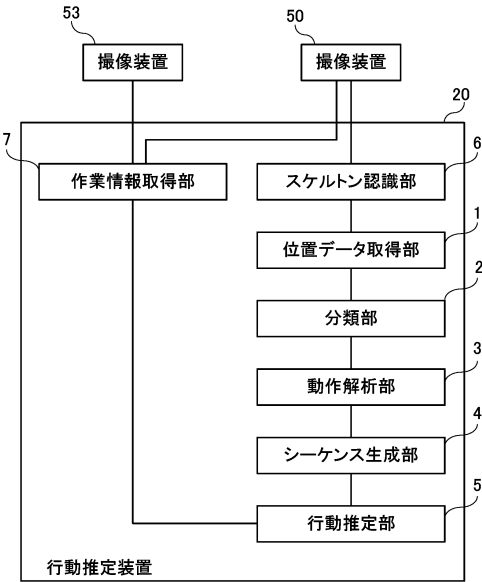
20

30

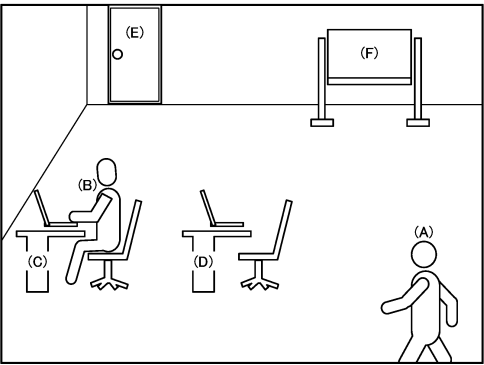
40

50

【図 5】



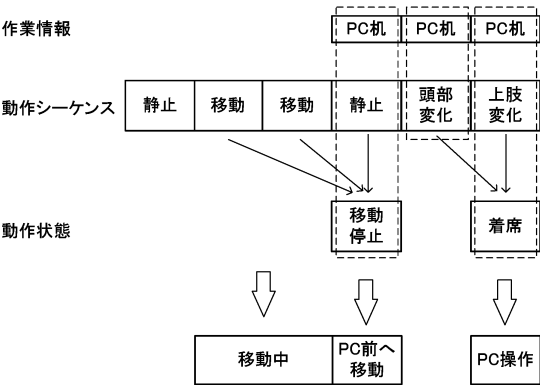
【図 6】



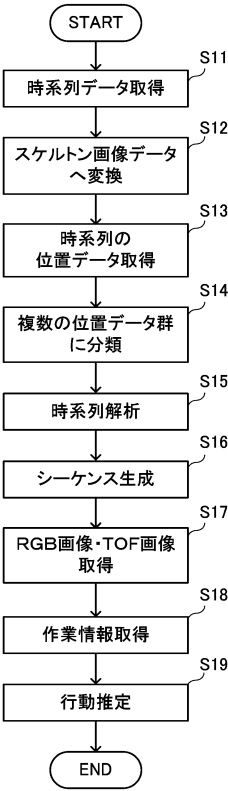
10

20

【図 7】



【図 8】

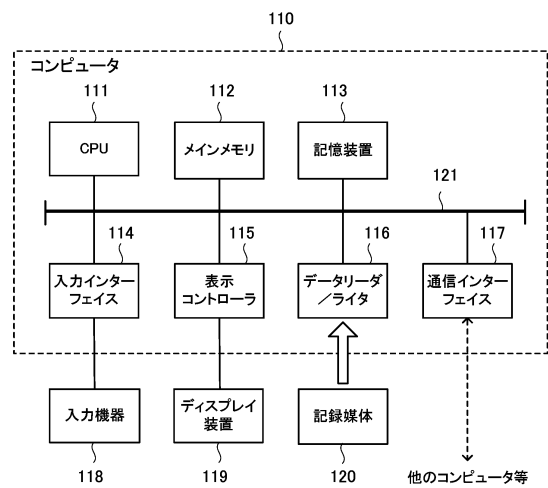


30

40

50

【図 9】



10

20

30

40

50

フロントページの続き

- (56)参考文献 国際公開第2017/150211 (WO, A1)
見持圭一, 塘中哲也, ネットワークロボットによる人の行動認識, 日本ロボット学会誌, 2005年, Vol.23, No.6, pp.687-690
菅原勝也, 外2名, 時系列の画像・深度情報を用いた接触する物体の領域抽出, 信学技報, 2015年, PRMU2014-107, MVE2014-69, pp.245-250
Jason Weston et al., Memory Networks, [online], 2015年11月29日, <https://arxiv.org/abs/1410.3916>
児玉悠, 外2名, 機械学習を用いた一般生活環境下での行動認識手法の検討, 信学技報, 2015年, IBISML2014-95, pp.73-78
- (58)調査した分野 (Int.Cl., DB名)
G06T 7/20
G06T 7/00