

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2020-9316

(P2020-9316A)

(43) 公開日 令和2年1月16日(2020.1.16)

(51) Int.Cl. F 1 テーマコード (参考)  
 G 0 6 T 7 / 7 0 (2017.01) G 0 6 T 7 / 7 0 Z 5 L 0 9 6

審査請求 未請求 請求項の数 9 O L (全 14 頁)

(21) 出願番号 特願2018-131662 (P2018-131662)  
 (22) 出願日 平成30年7月11日 (2018.7.11)

(71) 出願人 000003207  
 トヨタ自動車株式会社  
 愛知県豊田市トヨタ町1番地  
 (74) 代理人 100113608  
 弁理士 平川 明  
 (74) 代理人 100123319  
 弁理士 関根 武彦  
 (74) 代理人 100123098  
 弁理士 今堀 克彦  
 (74) 代理人 100143797  
 弁理士 宮下 文徳  
 (74) 代理人 100176201  
 弁理士 小久保 篤史  
 (74) 代理人 100138357  
 弁理士 矢澤 広伸

最終頁に続く

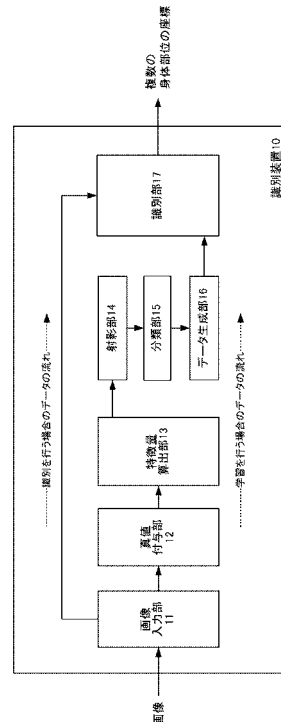
(54) 【発明の名称】 学習方法および学習装置

(57) 【要約】

【課題】 人の姿勢を精度良く認識する識別器を得る。

【解決手段】 対象者が有する複数の身体部位の位置を識別する識別器を学習させる学習方法。前記対象者を含む画像を取得する画像取得ステップと、前記画像に含まれる前記対象者が有する複数の身体部位の位置に関する情報である第一の部位情報を取得する真値取得ステップと、前記第一の部位情報を特徴量に変換する変換ステップと、複数の前記特徴量をクラスタリングする分類ステップと、生成された複数のクラスタのそれぞれについて、対応する前記第一の部位情報に類似する第二の部位情報を生成する生成ステップと、前記第二の部位情報を用いて、前記識別器を学習させる学習ステップと、を含む。

【選択図】 図 1



## 【特許請求の範囲】

## 【請求項 1】

対象者が有する複数の身体部位の位置を識別する識別器を学習させる学習方法であって、

前記対象者を含む画像を取得する画像取得ステップと、  
 前記画像に含まれる前記対象者が有する複数の身体部位の位置に関する情報である第一の部位情報を取得する真値取得ステップと、  
 前記第一の部位情報を特徴量に変換する変換ステップと、  
 複数の前記特徴量をクラスタリングする分類ステップと、  
 生成された複数のクラスタのそれぞれについて、対応する前記第一の部位情報に類似する第二の部位情報を生成する生成ステップと、  
 前記第二の部位情報を用いて、前記識別器を学習させる学習ステップと、  
 を含む、学習方法。

10

## 【請求項 2】

前記生成ステップでは、前記複数のクラスタにそれぞれ含まれる特徴量の数に基づいて、クラスタごとに生成する前記第二の部位情報の数を決定する、  
 請求項 1 に記載の学習方法。

## 【請求項 3】

前記特徴量は、前記複数の身体部位の位置関係に基づいた特徴量である、  
 請求項 1 または 2 に記載の学習方法。

20

## 【請求項 4】

前記特徴量は、前記複数の身体部位間の距離の順位に基づいた特徴量である、  
 請求項 3 に記載の学習方法。

## 【請求項 5】

前記生成ステップでは、前記第一の部位情報に対応する複数の身体部位の位置を、それぞれ所定の半径を有する領域内において移動させることで、前記第二の部位情報を生成する、  
 請求項 3 または 4 に記載の学習方法。

## 【請求項 6】

前記生成ステップでは、前記領域に含まれる複数の点について、位置に基づいた重みを付与し、前記重みに基づいて移動先を決定する、  
 請求項 5 に記載の学習方法。

30

## 【請求項 7】

前記分類ステップでは、前記特徴量の次元数を削減した結果をクラスタリングする、  
 請求項 1 から 6 のいずれかに記載の学習方法。

## 【請求項 8】

請求項 1 から 7 のいずれかに記載の学習方法の各ステップをコンピュータに実行させるためのプログラム。

## 【請求項 9】

対象者が有する複数の身体部位の位置を識別する識別器を学習させる学習装置であって、

前記対象者を含む画像を取得する画像取得手段と、  
 前記画像に含まれる前記対象者が有する複数の身体部位の位置に関する情報である第一の部位情報を取得する真値取得手段と、  
 前記第一の部位情報を特徴量に変換する変換手段と、  
 複数の前記特徴量をクラスタリングする分類手段と、  
 生成された複数のクラスタのそれぞれについて、対応する前記第一の部位情報に類似する第二の部位情報を生成する生成手段と、  
 前記第二の部位情報を用いて、前記識別器を学習させる学習手段と、  
 を有する、学習装置。

40

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は画像処理に関し、特に人の姿勢を推定する技術に関する。

## 【背景技術】

## 【0002】

画像中の人物の姿勢を推定する方法として、例えば、非特許文献1に記載された方法が提案されている。当該方法では、まず、CG（コンピュータ・グラフィックス）を用いて様々な体格・姿勢の人物に対する奥行画像を大量に生成する。また、これらの奥行画像のそれぞれに対して、頭部、顔面、首、上腕などといった、身体部位に対応する領域ごとに色分け（腑分け）を行う。さらに、古典的機械学習の一種であるランダム・フォレストを用いて、奥行画像中の任意の一点が、複数の身体部位に対応する複数の領域のうちのどれに属するのかを推定するための識別器を学習する。最終的に、得られた識別器を用いて、奥行画像中のすべての点について、複数の領域のうちのどれに属するのかを推定する。

10

## 【0003】

また、画像中の人物の姿勢を推定する方法として、例えば、非特許文献2に記載された方法が提案されている。当該方法では、まず、様々な体格・姿勢の人物に対応する画像を大量に用意する。また、これらの画像のそれぞれに対して、頭部、顔面、首、肩、肘などといった、人体の部位の位置を付与する。さらに、深層学習の一種で、VGG-16と呼ばれる畳み込みニューラル・ネットワーク（CNN）を用いて、画像中の人物の人体の部位の位置を推定するための識別器を学習する。最終的に、得られた識別器を用いて画像中の人物の人体の部位の位置を推定する。

20

## 【先行技術文献】

## 【非特許文献】

## 【0004】

【非特許文献1】J. Shotton et al., Real-Time Human Pose Recognition in Parts from Single Depth Images, CVPR2011

【非特許文献2】K. Simonyan et al., "Very Deep Convolutional Networks for Large-Scale Image Recognition", International Conference on Learning Representations, 2015.

30

【非特許文献3】田口善弘・大野克嗣・横山和成(2001):非計量多次元尺度構成法への期待と新しい視点、統計数理、第49巻第1号133-153

【非特許文献4】M. Belkin et al., Laplacian Eigenmaps for Dimensionality Reduction and Data Representation, Neural Computation, Vol. 15, No.6, pp. 1373-1396 (2003)

【非特許文献5】Normalized cuts and image segmentation, 2000 Jianbo Shi, Jitendra Malik

【非特許文献6】M. Sugiyama et al., On information maximization clustering: Tuning parameter selection and analytic solution. In L. Getoor and T. Scheffer, editors, Proceedings of 28th International Conference on Machine Learning (ICML2011), pages 65-72, Bellevue, Washington, USA, Jun. 28-Jul. 2 2011.

40

【非特許文献7】A. Toshev et al., Pose Estimation via Deep Neural Networks, CVPR 2015

【非特許文献8】M. Schwarz et al., RGB-D Object Recognition and Pose Estimation Based on Pre-Trained Convolutional Neural Network Features ICRA2015

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0005】

従来技術に係る識別方法では、識別器の推定精度が、学習データの質と量（具体的には、画像中の人物の体格・服装・姿勢、カメラの解像度・向き・角度などのばらつき）に強

50

く依存するといった問題がある。

すなわち、従来技術においては、識別器を学習させるために、十分な質と量を満たす学習データを用意しなければならないという点において課題があった。

【0006】

本発明は上記の問題点を考慮してなされたものであり、人の姿勢を精度良く識別するための識別器を得ることを目的とする。

【課題を解決するための手段】

【0007】

本発明の第一の形態に係る学習方法は、

対象者が有する複数の身体部位の位置を識別する識別器を学習させる学習方法であって、前記対象者を含む画像を取得する画像取得ステップと、前記画像に含まれる前記対象者が有する複数の身体部位の位置に関する情報である第一の部位情報を取得する真値取得ステップと、前記第一の部位情報を特徴量に変換する変換ステップと、複数の前記特徴量をクラスタリングする分類ステップと、生成された複数のクラスタのそれぞれについて、対応する前記第一の部位情報に類似する第二の部位情報を生成する生成ステップと、前記第二の部位情報を用いて、前記識別器を学習させる学習ステップと、を含むことを特徴とする。

10

【0008】

本発明の第二の形態に係る学習装置は、

対象者が有する複数の身体部位の位置を識別する識別器を学習させる学習装置であって、前記対象者を含む画像を取得する画像取得手段と、前記画像に含まれる前記対象者が有する複数の身体部位の位置に関する情報である第一の部位情報を取得する真値取得手段と、前記第一の部位情報を特徴量に変換する変換手段と、複数の前記特徴量をクラスタリングする分類手段と、生成された複数のクラスタのそれぞれについて、対応する前記第一の部位情報に類似する第二の部位情報を生成する生成手段と、前記第二の部位情報を用いて、前記識別器を学習させる学習手段と、を有することを特徴とする。

20

【発明の効果】

【0009】

本発明によれば、人の姿勢を精度良く識別するための識別器を得ることができる。

【図面の簡単な説明】

30

【0010】

【図1】第一の実施形態に係る識別装置10の機能構成図。

【図2】識別装置10が行う学習処理のフローチャート。

【図3】画像入力部11に入力される画像の例。

【図4】真値付与部12によって付与された座標の例。

【図5】部位間の距離の大きさの順位に基づく順位特徴量の説明図。

【図6】分類部15によってクラスタリングされた特徴量を説明する図。

【図7】第二の部位情報を生成する方法を説明する図。

【図8】第二の実施形態に係る識別装置10の機能構成図。

【発明を実施するための形態】

40

【0011】

本発明に係る学習方法は、人が有する複数の身体部位の位置を、入力された画像に基づいて識別する識別器を学習させる方法である。

識別器を学習させるためには、一般的に、大量の学習データが必要になる。そこで、本発明では、既知の学習データに基づいて、数の不足が予測される学習データを推定して生成し、生成した学習データを併用して識別器の学習を行う。かかる構成によると、識別器の汎化性能を高めることができる。

【0012】

本発明に係る学習方法は、対象者を含む画像を取得し、当該対象者が有する身体部位の位置に関する情報である第一の部位情報を取得する。画像は、例えば、可視光画像や赤外

50

画像である。また、可視光画像や赤外画像と、距離画像とを組み合わせ用いてもよい。

また、第一の部位情報は、真の値を表すデータ（正解データ）である。第一の部位情報には、例えば、頭、首、肩、肘、手首、掌、胸、股関節、膝、足首などの位置（例えば、各部位の中心位置）が含まれる。また、第一の部位情報には、頭部領域、顔の向き、視線方向、手（手指）領域、手指形状が含まれてもよい。

これらの情報は、特徴量に変換される。

#### 【0013】

本発明に係る学習方法は、さらに、複数の画像にそれぞれ対応する複数の特徴量をクラスタリングし、生成された複数のクラスタのそれぞれについて、対応する第一の部位情報に類似する第二の部位情報を生成する。第二の部位情報は、新たに生成された学習データ

10

#### 【0014】

かかる形態によれば、正解データに基づいて学習データの数を増やせるため、人の姿勢をより精度よく識別する識別器を得ることができる。

さらに、特徴量をクラスタリングして得られたクラスタごとに第二の部位情報を生成するため、学習データの充足度を判定することが可能になる。

#### 【0015】

また、前記生成ステップでは、前記複数のクラスタにそれぞれ含まれる特徴量の数に基づいて、クラスタごとに生成する前記第二の部位情報の数を決定することを特徴としてもよい。

20

かかる構成によると、含まれる特徴量の数が相対的に少ないクラスタがある場合に、当該クラスタに対応する第二の部位情報を相対的に多く生成するといった対応が可能になる。

#### 【0016】

なお、前記特徴量は、前記複数の身体部位の位置関係に基づいた特徴量であることを特徴としてもよい。また、前記特徴量は、前記複数の身体部位間の距離の順位に基づいた特徴量であることを特徴としてもよい。

#### 【0017】

このような特徴量は、スケール変換・回転・平行移動に対して不変であり、また、微小変動に対して頑強である。したがって、このような特徴量を採用することで、どのような学習データが不足しているのかを精度よく推定することができる。

30

#### 【0018】

また、前記生成ステップでは、前記第一の部位情報に対応する複数の身体部位の位置を、それぞれ所定の半径を有する領域内において移動させることで、前記第二の部位情報を生成することを特徴としてもよい。

このように、複数の身体部位の位置に微小変動を加えることで、学習データを増やし、位置の推定誤差が少ない識別器を生成できるようになる。

#### 【0019】

また、前記生成ステップでは、前記領域に含まれる複数の点について、位置に基づいた重みを付与し、前記重みに基づいて移動先を決定することを特徴としてもよい。

40

例えば、第一の部位情報に対応する座標に近いほど重みを大きくするようにしてもよい。重みは、例えば、正規分布などを用いて付与してもよい。

#### 【0020】

また、前記分類ステップでは、前記特徴量の次元数を削減した結果をクラスタリングしてもよい。かかる構成によると、処理効率を向上させることができる。

#### 【0021】

（第一の実施形態）

本発明の第一の実施形態に関して、図面を参照して説明する。図1は、第一の実施形態に係る識別装置10の機能構成を示すブロック図である。

本実施形態に係る識別装置10は、入力された画像に基づいて、当該画像に含まれる人

50

物が有する複数の身体部位の座標を識別し出力する装置である。識別装置 10 は、識別器を有しており、当該識別器を学習させる学習モードと、学習結果に基づいて、識別を行う識別モードに切り替え可能に構成される。以降の説明において、識別器を学習させる処理を学習処理と称し、学習した識別器を用いて識別を行う処理を識別処理と称する。

#### 【0022】

本実施形態に係る識別装置 10 は、半導体集積回路 (LSI) を用いて実現できる。識別装置 10 は、画像入力部 11、真値付与部 12、特徴量算出部 13、射影部 14、分類部 15、データ生成部 16、識別部 17 を有して構成される。これらの構成要素はソフトウェアモジュールによって実現され、識別装置 10 が果たす機能にそれぞれ対応している。

10

#### 【0023】

これらの各機能部については、以下に説明する学習処理および識別処理の中で合わせて説明する。

#### 【0024】

次に、識別装置 10 が行う学習処理について説明する。図 2 は、学習処理の流れを示すフローチャートである。

#### 【0025】

まず、ステップ S10 において、画像の入力インターフェースである画像入力部 11 が、対象者の体が含まれる画像を取得する。

画像入力部 11 は、識別装置 10 に接続されたカメラや、記憶装置、ネットワーク等を介して画像を取得する。例えば、車両内の乗員を対象として処理を行う場合、車両内に設置されたカメラから画像を取得してもよい。なお、本実施形態では、画像入力部 11 が赤外画像を取得しているが、可視光画像を取得してもよい。ここで取得される画像は、当該画像に含まれる対象者の姿勢が既知である画像である。

20

#### 【0026】

画像入力部 11 は、図 3 のような、対象者が含まれた画像を複数取得し、時刻  $t$  ( $t = 1, 2, \dots, T$ ) に対応する画像  $I(t)$  を真値付与部 12 に出力する。なお、本実施形態では、学習を行うための画像の集合として、所定の間隔で取得された  $T$  枚の画像を用いるが、画像を取得する間隔や枚数は特に限定されない。

#### 【0027】

次に、ステップ S11 において、真値付与部 12 が、画像入力部 11 が取得した画像  $I(t)$  のそれぞれについて、対象者の体に存在する  $M$  個の身体部位の、画像中における二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) を取得する。

30

図 4 は、 $M = 10$  である場合の例を示した図である。ここでは、真値付与部 12 が、頭、首、肩 (左右)、肘 (左右)、掌 (左右)、股関節 (左右) の各部位の中心座標を取得するものとする。すなわち、10 組の二次元座標が取得される。

ここで、 $x_m(t)$  は、時刻  $t$  における、 $m$  番目の部位の、画像  $I(t)$  中における水平方向座標を表す。また、 $y_m(t)$  は、時刻  $t$  における、 $m$  番目の部位の、画像  $I(t)$  中における垂直方向座標を表す。

#### 【0028】

ここで取得する座標の組が、本発明における第一の部位情報である。第一の部位情報は、真の値を表す情報である。第一の部位情報は、装置のユーザによって与えられてもよいし、座標データが画像に付属している場合、当該データを取得してもよい。

40

#### 【0029】

次に、ステップ S12 において、特徴量算出部 13 が、真値付与部 12 で得られた、時刻  $t$  における対象者の  $M$  個の部位の二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) に基づいて、特徴量  $F(t)$  を算出する。具体的には、特徴量  $F(t)$  は、次の式 (1) を用いて算出される。

#### 【数 1】

$$F(t) = (R(D(1,2)), R(D(1,3)), \dots, R(D(8,9)), R(D(9,10))) \quad \dots (1)$$

50

## 【 0 0 3 0 】

ここで、 $D(m,n)$  は、 $m$  番目の部位と  $n$  番目の部位との画像空間上におけるユークリッド距離を表す。また、 $R(D(m,n))$  は、 $D(1,2), D(1,3), \dots, D(8,9), D(9,10)$  を降順にソートした際における、 $D(m,n)$  の順位を表す。

## 【 0 0 3 1 】

図 5 ( A ) は、対象者の身体部位を便宜上 4 つとした場合の例である。ここでは、すべての部位のペア間の距離  $D(t)$  が、以下のように与えられたとする。

$$D(t) = (D(1,2), D(1,3), D(1,4), D(2,3), D(2,4), D(3,4)) \\ = (5.5, 2.6, 2.8, 3.5, 4.3, 4.0)$$

この場合、時刻  $t$  における特徴量  $F(t)$  は、

$$F(t) = (1, 6, 5, 4, 2, 3)$$

のように算出される。

10

## 【 0 0 3 2 】

特徴量  $F(t)$  は、複数の身体部位の間の距離の大きさの順位に基づく特徴量である。より詳細には、複数の身体部位間の距離の、全ての身体部位の組み合わせ中における順位を列挙したものである。

本明細書では、このような特徴量を「順位特徴量（または単に特徴量）」と称する。

## 【 0 0 3 3 】

順位特徴量  $F(t)$  は、スケール変換や微小変動に対して頑強であるという特徴を持つ。例えば、ある画像から得た身体部位が、図 5 ( A ) に示したような位置関係にある場合、当該画像を拡大すると、各身体部位の位置は図 5 ( B ) のようになる。これらの図からも分かるように、身体部位のスケールが変動しても、順位特徴量は変わらない。

20

## 【 0 0 3 4 】

図 5 ( C ) は、図 5 ( A ) に示した身体部位の位置が微小に変動した場合の例である。図 5 ( A ) および図 5 ( C ) から分かるように、身体部位の微小変動に対して順位特徴量は頑健である。順位特徴量  $F(t)$  は、身体部位間の距離の順位のみ依存するため、拡大・縮小といったスケール変動に影響されない。また、身体部位の位置が微小に変動しても、その変動がそれらの順位に影響しなければ、順位特徴量  $F(t)$  は変化しない。

このような特性により、対象者の姿勢を推定する際に生じる様々な変動、例えば、対象者の水平移動、対象者の体格差、カメラの位置や向き、深層学習による人体の部位の位置の推定誤差などによる影響を抑制することができる。

30

なお、順位特徴量  $F(t)$  の次元数  $D$  は、身体部位の個数が  $M$  個であるとき、その組み合わせの個数  ${}_M C_2$  によって与えられる。例えば、身体部位の個数が 10 個である場合、順位特徴量は 45 次元のベクトルとなる。

## 【 0 0 3 5 】

以上に説明した処理により、画像  $I(t)$  ごとに一つの順位特徴量  $F(t)$  が取得される。すなわち、画像が  $T$  枚ある場合、 $T$  個の順位特徴量が取得される。

## 【 0 0 3 6 】

次に、ステップ S 1 3 において、射影部 1 4 が、特徴量算出部 1 3 が取得した  $T$  個の順位特徴量  $F(t)$  ( $t=1, \dots, T$ ) を、 $d (< D)$  次元の低次元空間（例えば、 $d=2$  の平面）に射影し、 $T$  個の特徴量  $f(t)$  ( $t=1, \dots, T$ ) を生成する。以降、低次元空間に射影された特徴量を、低次元特徴量と称する。

40

## 【 0 0 3 7 】

特徴量を低次元空間に射影するための手法として、例えば、古典的な統計解析手法の一種である主成分分析 (PCA) がある。あるいは、例えば、非特許文献 3 に記載されたような、多次元尺度構成法 (MDS) を用いてもよい。あるいは、例えば、非特許文献 4 に記載されたような、ラプラシアン固有値マップ (LLE) を用いてもよい。

本ステップにより、例えば、順位特徴量の次元数が、45 次元から 2 次元に削減される。

## 【 0 0 3 8 】

50

次に、ステップ S 1 4 において、分類部 1 5 が、低次元特徴量  $f(t)$  をクラスタリングし、複数のクラスタ（低次元特徴量  $f(t)$  ( $t=1, \dots, T$ ) の部分集合）を生成する。

クラスタを生成するための手法として、例えば、古典的なクラスタリング手法の一種である  $k$ -means を用いることができる。あるいは、例えば、非特許文献 5 に記載されたような、スペクトラル・クラスタリングを用いてもよい。あるいは、例えば、非特許文献 6 に記載されたような、二乗損失相互情報量を用いた情報量最大化クラスタリングを用いてもよい。

【 0 0 3 9 】

図 6 ( A ) は、二次元平面に射影された複数の低次元特徴量を示した図であり、図 6 ( B ) は、複数の低次元特徴量をクラスタリングした結果を示した図である。ここでは、4 つのクラスタが生成されたものとする。

10

【 0 0 4 0 】

ここで生成されたクラスタは、順位特徴量に基づいたものである。すなわち、異なるクラスタに分類された低次元特徴量は、異なる姿勢に対応するものである可能性が高い。すなわち、複数のクラスタに含まれる低次元特徴量の数に基づいて、学習データが相対的に不足している姿勢を特定することができる。本実施形態に係る識別装置 1 0 は、複数のクラスタに含まれる低次元特徴量の数に基づいて、学習データが相対的に不足している姿勢を特定し、さらなる学習データを自動的に生成する。

【 0 0 4 1 】

ステップ S 1 5 では、データ生成部 1 6 が、ステップ S 1 5 で得られたクラスタリングの結果を参照して、第一の部位情報から、新たな部位情報（第二の部位情報）を生成する。

20

【 0 0 4 2 】

具体的には、順位特徴量  $F(t)$  から変換された低次元特徴量  $f(t)$  ( $t=1, \dots, T$ ) をクラスタリングした結果に基づいて、対象者の  $M$  個の身体部位に対する二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) に対して、以下で説明するデータ拡張を適用し、新たな二次元座標  $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, T'$ ) を生成する。

【 0 0 4 3 】

なお、データ拡張とは、対象者の  $M$  個の身体部位に対する二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) に微小な変動を加えることで、元の二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) とは似て異なる新たな二次元座標  $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, T'$ ) を生成することを意味する。

30

【 0 0 4 4 】

微小な変動は、例えば、以下のようにして加えることができるが、これらに限られない。

- ( 1 )  $M$  個の身体部位のうちの一つ以上を、それぞれ所定の範囲内で別個に移動させる
- ( 2 )  $M$  個の身体部位の組を拡大または縮小させる
- ( 3 )  $M$  個の身体部位の組を平行移動させる
- ( 4 )  $M$  個の身体部位の組を回転させる
- ( 5 )  $M$  個の身体部位の組を所定のルールによって変形させる

40

【 0 0 4 5 】

ここで、 $T$  個の低次元特徴量  $f(t)$  ( $t=1, \dots, T$ ) から、 $T'$  個の新たな二次元座標  $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, T'$ ) を生成する場合を考える。

$k$  番目のクラスタに、低次元特徴量  $f(t_k)$  ( $t_k=1, \dots, L_k$ ) が含まれるとした場合、当該クラスタに対応する二次元座標  $(x_m(t), y_m(t))$  ( $m=1, 2, \dots, M$ ) に対してデータ拡張を適用し、 $L'_k$  ( $=w_k \times c$ ) 個の二次元座標  $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, L'_k$ ) を生成する。ここで、係数  $w_k$  は、 $k$  番目のクラスタに対する重みであり、式 ( 2 ) で与えられる。

## 【数 2】

$$w_k = \frac{T}{L_k} \quad \dots (2)$$

## 【0046】

また、係数 $c$ は、データ拡張を適用して生成する低次元特徴量 $f'(t)$  ( $t=1, \dots, T'$ )の個数 $T'$ の関数であり、式(3)で与えられる。

## 【数 3】

$$c = \frac{T'}{T \cdot \sum_k \frac{1}{L_k}} \quad \dots (3)$$

10

## 【0047】

すなわち、 $k$ 番目のクラスに含まれる低次元特徴量 $f(t_k)$  ( $t_k=1, \dots, L_k$ )の個数 $L_k$ が相対的に少ない場合、データ拡張により、相対的に多くの二次元座標 $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, L'_k$ )が生成される。

一方、 $k$ 番目のクラスに含まれる低次元特徴量 $f(t_k)$  ( $t_k=1, \dots, L_k$ )の個数 $L_k$ が相対的に多い場合、データ拡張により、相対的に少ない二次元座標 $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, L'_k$ )が生成される。

なお、データ拡張によって生成されるデータの総数は、試行錯誤的に決定すればよい。

20

## 【0048】

微小な変動は、例えば、複数の身体部位を表す座標を含んだ所定の範囲を生成し、当該範囲内において行うことができる。例えば、図7に示したように、複数の身体部位の位置を中心とした半径 $r$ の円を生成し、当該円内において座標を移動させてもよい。移動は、ランダムに行ってもよいし、座標の移動先がより円の中心側に偏るように重みを付したうえで行ってもよい。また、半径 $r$ の値は、対応する人物の属性(身長、体格、体幹の長さ、顔の大きさ等)に応じて試行錯誤的に決定すればよい。

以上の処理により、第二の部位情報が生成される。

## 【0049】

次に、ステップS16で、第一の部位情報と、新たに生成した第二の部位情報を用いて、対象者が有する身体部位の位置を推定するための識別器(識別部17)を学習させる。

30

具体的には、第二の部位情報に対応する画像を新たに生成し、以下のデータを、学習データとして識別部17が有する識別器に入力し、学習させる。

(1) 第一の部位情報に対応する画像と、第一の部位情報の組み合わせ

(2) 第二の部位情報に対応する画像と、第二の部位情報の組み合わせ

なお、第二の部位情報に対応する画像は、元の画像に変形を加えることで生成することができる。

本実施形態では、識別部17は、識別器を含んで構成される。具体的には、識別部17は、取得した画像に基づいて、予め定められた複数の身体部位の座標を推定して出力する識別器を含む。

40

ここで、識別器は、例えば、非特許文献7の深層学習を用いて学習させることができる。あるいは、例えば、非特許文献8の深層学習を用いて学習させてもよい。

## 【0050】

これにより、ある時刻 $t_0$ における画像に基づいて、当該画像に含まれる対象者の $M$ 個の身体部位の二次元座標 $(X_m(t_0), Y_m(t_0))$  ( $m=1, 2, \dots, M$ )を出力する識別器を得ることができる。すなわち、画像中の人物の身体部位の座標を取得する装置を得ることができる。

また、身体部位の座標を、身体部位を活用するより上位の装置に送信することで、人物がとっている姿勢や行動を特定することが可能になる。例えば、車両内乗員によるスマートフォンの操作や飲食といった危険姿勢を識別し、車両の走行状態と照らし合わせることで適応的に注意を促すことなどが可能になる。

50

## 【0051】

以上説明したように、第一の実施形態によると、複数の順位特徴量をクラスタリングした結果を用いることで、数が相対的に不足している学習データを特定し、追加の学習データを自動生成することができる。かかる形態によると、身体部位の位置を識別する識別器をより精度よく学習させることが可能になる。

## 【0052】

(第二の実施形態)

第一の実施形態では、識別部17が、対象者の複数の身体部位の座標を出力した。これに対し、第二の実施形態は、識別部17が、対象者の複数の身体部位の座標に加え、対象者の姿勢を示す姿勢ラベルを出力する実施形態である。図8は、第二の実施形態における識別装置10の機能構成を示すブロック図である。

10

## 【0053】

第二の実施形態では、真値付与部12が、身体部位の二次元座標に加え、対象者が取っている姿勢を表す姿勢ラベル $V(t)$ を取得する。姿勢ラベルは、姿勢を表すものであればどのように定義されていてもよい。例えば、車両の乗員を対象とする場合、ハンドル操作、バックミラーの調整、コントロール・パネルの調整、シートベルトの着脱、スマートフォン操作、飲食といった姿勢ラベルを定義してもよい。本例では、説明を簡単にするため、「正常姿勢」と「異常姿勢」の二種類のラベルを付与するものとする。例えば、正常姿勢の場合、 $V(t)=0$ という値を与え、異常姿勢の場合、 $V(t)=1$ という値を与える。

## 【0054】

第二の実施形態では、射影部14が生成した低次元特徴量 $f(t)$  ( $t=1, \dots, T$ )が、分類部15の他に識別部17へも出力される。

20

また、第二の実施形態では、データ生成部16が生成した第二の部位情報(すなわち、対象者の $M$ 個の身体部位に対する新たな二次元座標 $(x'_m(t'), y'_m(t'))$  ( $m=1, 2, \dots, M, t'=1, \dots, T'$ ))が、識別部17の他に特徴量算出部13へも出力される。

## 【0055】

第二の実施形態では、識別部17は、二種類の識別器を有して構成される。一つ目(第一の識別器)は、第一の実施形態で述べたような、ある時刻 $t_0$ における画像に含まれる対象者の $M$ 個の身体部位の座標を推定して出力する識別器である。二つ目(第二の識別器)は、ある時刻 $t_0$ における画像に含まれる対象者の姿勢ラベル $V(t_0)$ を推定して出力する識別器である。このような識別器は、例えば、古典的機械学習の一種であるサポート・ベクターマシンを用いて学習することができる。

30

## 【0056】

第二の識別器は、順位特徴量を用いて学習される。すなわち、データ生成部16によって生成された第二の部位情報が、特徴量算出部13によって順位特徴量に変換され、射影部14によって低次元特徴量に変換されたのち、識別部17に入力される。

かかる構成によると、対象者の身体部位の位置を推定する識別器に加え、対象者の姿勢そのものを識別する識別器を得ることができるようになる。すなわち、対象者の姿勢に基づいた様々な機能を提供できるようになる。

## 【0057】

(変形例)

本発明は、上記手段の少なくとも一部を備える識別装置または学習装置として捉えることができる。また、本発明は、上記処理の少なくとも一部を実行する識別方法または学習方法として捉えることもできる。また、本発明は、これら方法をコンピュータに実行させるためのコンピュータプログラム、あるいはこのコンピュータプログラムを非一時的に記憶したコンピュータ可読記憶媒体として捉えることもできる。上記手段および処理の各々は可能な限り互いに組み合わせて本発明を構成することができる。

40

## 【0058】

なお、学習処理(ステップS11)において用いられる身体部位の位置は、どのようにして求めても構わない。これは部位検出のアルゴリズムが特定のものに限定されないこと

50

を意味するだけでなく、部位検出を人手で行ってもよいことを意味する。

【0059】

また、実施形態の説明では、二次元の画像空間上における距離を用いたが、人体の部位が三次元座標によって得られている場合、三次元空間内における距離を採用することもできる。例えば、赤外画像や可視光画像に加え、距離画像を取得可能である場合、処理系を三次元に拡張することもできる。

【0060】

本発明における識別装置10は、半導体集積回路(LSI)による実装に限定されず、汎用的なマイクロプロセッサやメモリを有するコンピュータがプログラムを実行することによって実現されても構わない。

10

【0061】

また、実施形態の説明では、同一の装置が学習処理と識別処理の双方を行ったが、学習処理を行う装置と、識別処理を行う装置が分かれていてもよい。さらに、識別装置10から、学習処理に関連した手段を省いてもよい。例えば、例示した識別装置10から、真値付与部12、特徴量算出部13、射影部14、分類部15、データ生成部16を省略し、識別処理のみを行う装置(識別装置)として実施してもよい。

この場合、学習済みの識別部17(または、識別部17が有する識別器)を装置から論理的に切り離し、別の装置に組み込み可能に構成してもよい。例えば、実施形態に係る識別装置10によって学習された識別部17を、識別処理のみを行う装置に組み込めるようにしてもよい。

20

また、反対に、学習処理のみを行う装置として発明を実施してもよい。

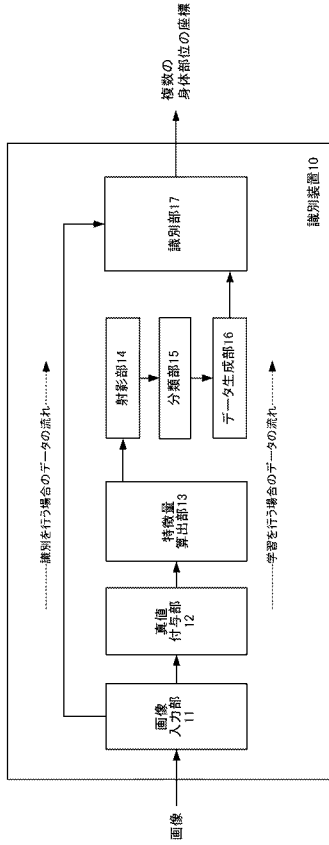
【符号の説明】

【0062】

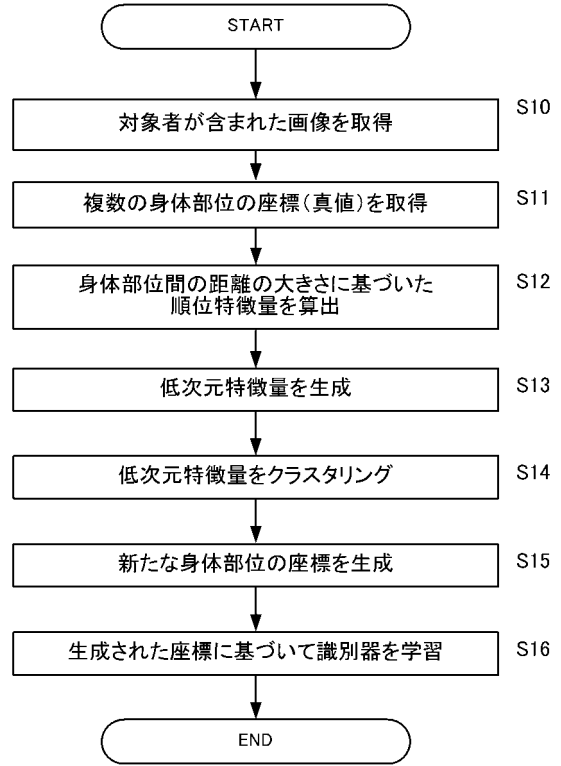
- 10・・・識別装置
- 11・・・画像入力部
- 12・・・真値付与部
- 13・・・特徴量算出部
- 14・・・射影部
- 15・・・分類部
- 16・・・データ生成部
- 17・・・識別部

30

【 図 1 】



【 図 2 】

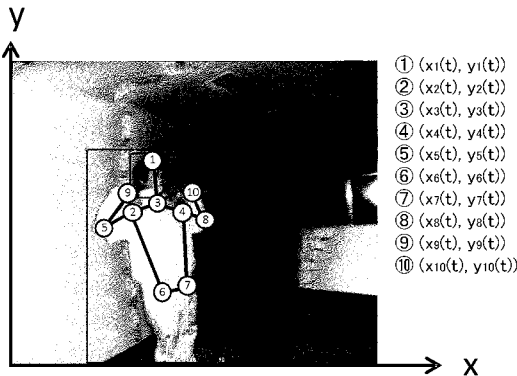


【 図 3 】

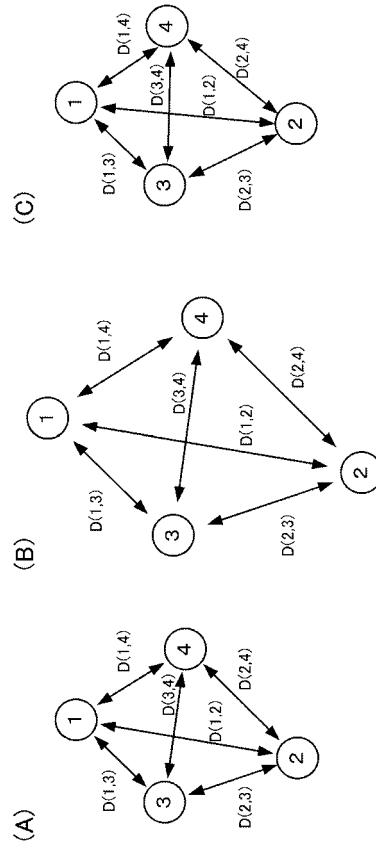


時刻tに対応する画像I(t)

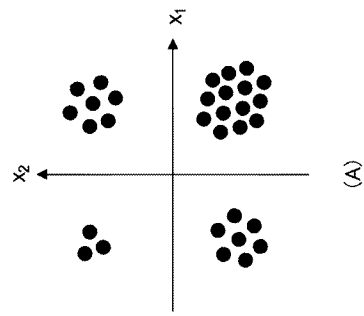
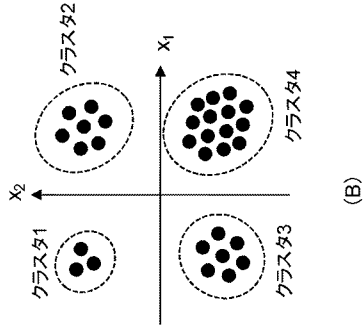
【 図 4 】



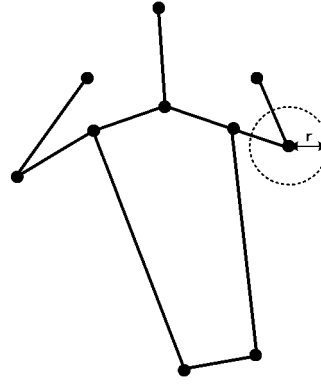
【 図 5 】



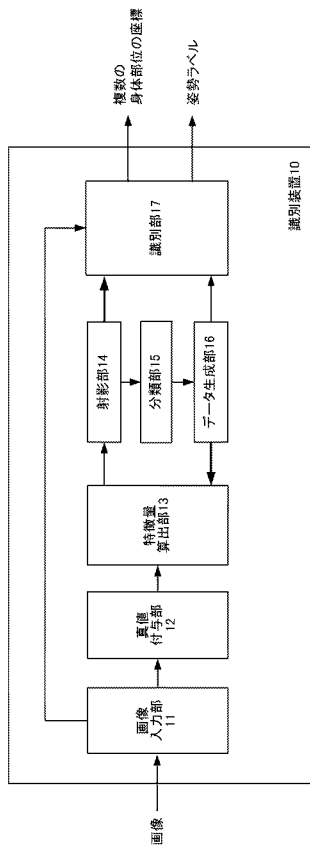
【 図 6 】



【 図 7 】



【 図 8 】



フロントページの続き

(72)発明者 山中 正雄

東京都港区赤坂6丁目6番20号 株式会社トヨタIT開発センター内

Fターム(参考) 5L096 BA20 EA39 FA52 FA66 FA79 JA11 KA04 MA07