

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7101829号
(P7101829)

(45)発行日 令和4年7月15日(2022.7.15)

(24)登録日 令和4年7月7日(2022.7.7)

(51)国際特許分類	F I
G 0 6 T 7/00 (2017.01)	G 0 6 T 7/00 3 0 0 F
	G 0 6 T 7/00 6 6 0 Z

請求項の数 18 (全53頁)

(21)出願番号	特願2020-572391(P2020-572391)	(73)特許権者	321006888
(86)(22)出願日	令和2年4月29日(2020.4.29)		ベイジン・センスタイム・テクノロジー
(65)公表番号	特表2022-503426(P2022-503426 A)		・デベロップメント・カンパニー・リミテッド
(43)公表日	令和4年1月12日(2022.1.12)		中華人民共和国・1 0 0 0 8 0 ・ベイジン・ハイディエン・ノースウェスト・フォース・リング・ロード・ナンバー・5 8 ・イレブンス・フロア・ルーム・1 1 0 1 - 1 1 1 7
(86)国際出願番号	PCT/CN2020/087826	(74)代理人	100078282
(87)国際公開番号	WO2021/057027		弁理士 山本 秀策
(87)国際公開日	令和3年4月1日(2021.4.1)	(74)代理人	100113413
審査請求日	令和2年12月24日(2020.12.24)		弁理士 森下 夏樹
(31)優先権主張番号	201910926373.4	(74)代理人	100181674
(32)優先日	令和1年9月27日(2019.9.27)		弁理士 飯田 貴敏
(33)優先権主張国・地域又は機関	中国(CN)		

最終頁に続く

(54)【発明の名称】 人体検出方法、装置、コンピュータ機器及び記憶媒体

(57)【特許請求の範囲】

【請求項1】

人体検出方法であって、前記人体検出方法は、
検出対象画像を取得することと、
前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報および人体輪郭を示すための輪郭キーポイントの位置情報を決定することと、
前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成することと
を含み、
前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報および人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、
前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴および輪郭特徴を取得し、取得された骨格特徴および輪郭特徴に対して特徴融合を行うことと、
前記特徴融合の結果に基づいて、前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報を決定することと
を含む、人体検出方法。

【請求項2】

前記輪郭キーポイントは、主輪郭キーポイントおよび補助輪郭キーポイントを含み、2つの隣接する前記主輪郭キーポイントの間に、少なくとも1つの補助輪郭キーポイントが存在する、請求項1に記載の人体検出方法。

【請求項 3】

前記検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、

前記検出対象画像に基づいて、前記主輪郭キーポイントの位置情報を決定することと、

前記主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定することと、

決定された前記人体輪郭情報に基づいて、複数の前記補助輪郭キーポイントの位置情報を決定することと

を含む、請求項 2 に記載の人体検出方法。

【請求項 4】

前記人体検出結果は、

骨格キーポイントマークおよび輪郭キーポイントマークが付加された検出対象画像、前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報を含むデータ群のうちの 1 つまたは複数を含む、請求項 1 ～ 3 のいずれか一項に記載の人体検出方法。

【請求項 5】

前記人体検出方法は、

前記人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集、人体画像の貼り付けのうちの 1 つまたは複数を実行することを更に含む、請求項 4 に記載の人体検出方法。

【請求項 6】

前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴および輪郭特徴を取得し、取得された骨格特徴および輪郭特徴に対して特徴融合を行うことは、

前記検出対象画像に基づいて、少なくとも 1 回の特徴抽出を行い、特徴抽出を行うたびに取得された骨格特徴および輪郭特徴に対して特徴融合を行うことであって、複数回の特徴抽出を行う場合、 i 回目の特徴融合の結果に基づいて $i + 1$ 回目の特徴抽出を行い、 i は正整数である、ことを含み、

前記特徴融合の結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報および人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、

最終回の特徴融合の結果に基づいて、前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報を決定することを含む、請求項 1 に記載の人体検出方法。

【請求項 7】

前記人体検出方法は、人体検出モデルによって実現され、前記人体検出モデルは、第 1 特徴抽出ネットワークおよび／または特徴融合ニューラルネットワークを含む、請求項 1 ～ 6 のいずれか一項に記載の人体検出方法。

【請求項 8】

前記検出対象画像に基づいて、少なくとも 1 回の特徴抽出を行うことは、

1 回目の特徴抽出において、事前訓練された前記第 1 特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第 1 ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第 1 ターゲット輪郭特徴行列を抽出することと、

$i + 1$ 回の特徴抽出において、事前訓練された第 2 特徴抽出ネットワークを用いて、 i 回目の特徴融合の結果から、前記第 1 ターゲット骨格特徴行列および前記第 1 ターゲット輪郭特徴行列を抽出することと

を含み、

前記第 1 特徴抽出ネットワークのネットワークパラメータは、前記第 2 特徴抽出ネットワークのネットワークパラメータとは異なり、かつ、特徴抽出のたびに用いられる前記第 2 特徴抽出ネットワークのネットワークパラメータは異なる、請求項 7 に記載の人体検出方法。

【請求項 9】

抽出して得られた骨格特徴および輪郭特徴に対して特徴融合を行うことは、

事前訓練された前記特徴融合ニューラルネットワークを用いて、前記第 1 ターゲット骨格

10

20

30

40

50

特徴行列および前記第 1 ターゲット輪郭特徴行列に対して特徴融合を行い、第 2 ターゲット骨格特徴行列および第 2 ターゲット輪郭特徴行列を得ることを含み、前記第 2 ターゲット骨格特徴行列は、三次元骨格特徴行列であり、前記三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、前記二次元骨格特徴行列における各要素の値は、該要素に対応する画素点が、対応する骨格キーポイントに属する確率を表し、

前記第 2 ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、前記三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、前記二次元輪郭特徴行列における各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表し、

特徴融合のたびに用いられる前記特徴融合ニューラルネットワークのネットワークパラメータは異なる、請求項 8 に記載の人体検出方法。

【請求項 10】

前記第 1 特徴抽出ネットワークは、共通特徴抽出ネットワークと、第 1 骨格特徴抽出ネットワークと、第 1 輪郭特徴抽出ネットワークとを含み、

前記第 1 特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第 1 ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第 1 ターゲット輪郭特徴行列を抽出することは、

前記共通特徴抽出ネットワークを用いて、前記検出対象画像に対して畳み込み処理を行い、骨格特徴および輪郭特徴を含む基礎特徴行列を得ることと、

前記第 1 骨格特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第 1 骨格特徴行列を得て、前記第 1 骨格特徴抽出ネットワークにおける第 1 ターゲット畳み込み層から、第 2 骨格特徴行列を取得し、前記第 1 骨格特徴行列および前記第 2 骨格特徴行列に基づいて、前記第 1 ターゲット骨格特徴行列を得ることであって、前記第 1 ターゲット畳み込み層は、前記第 1 骨格特徴抽出ネットワークにおける、最後の 1 つの畳み込み層以外のいずれか 1 つの畳み込み層である、ことと、

前記第 1 輪郭特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第 1 輪郭特徴行列を得て、前記第 1 輪郭特徴抽出ネットワークにおける第 2 ターゲット畳み込み層から、第 2 輪郭特徴行列を取得し、前記第 1 輪郭特徴行列および前記第 2 輪郭特徴行列に基づいて、前記第 1 ターゲット輪郭特徴行列を得ることであって、前記第 2 ターゲット畳み込み層は、前記第 1 輪郭特徴抽出ネットワークにおける、最後の 1 つの畳み込み層以外のいずれか 1 つの畳み込み層である、ことと

を含む、請求項 8 に記載の人体検出方法。

【請求項 11】

第 1 骨格特徴行列および前記第 2 骨格特徴行列に基づいて、前記第 1 ターゲット骨格特徴行列を得ることは、

前記第 1 骨格特徴行列および前記第 2 骨格特徴行列に対して連結処理を行い、第 1 連結骨格特徴行列を得ることと、前記第 1 連結骨格特徴行列を次元変換処理し、前記第 1 ターゲット骨格特徴行列を得ることとを含み、

前記第 1 輪郭特徴行列および前記第 2 輪郭特徴行列に基づいて、前記第 1 ターゲット輪郭特徴行列を得ることは、

前記第 1 輪郭特徴行列および前記第 2 輪郭特徴行列に対して連結処理を行い、第 1 連結輪郭特徴行列を得ることと、前記第 1 連結輪郭特徴行列を次元変換処理し、前記ターゲット輪郭特徴行列を得ることとを含み、

前記第 1 ターゲット骨格特徴行列の次元は、前記第 1 ターゲット輪郭特徴行列の次元と同じであり、かつ、前記第 1 ターゲット骨格特徴行列と前記第 1 ターゲット輪郭特徴行列は、同一の次元における次元数が同じである、請求項 10 に記載の人体検出方法。

【請求項 12】

前記特徴融合ニューラルネットワークは、第 1 畳み込みニューラルネットワークと、第 2 畳み込みニューラルネットワークと、第 1 変換ニューラルネットワークと、第 2 変換ニュー

10

20

30

40

50

ーラルネットワークとを含み、

前記特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列および第2ターゲット輪郭特徴行列を得ることは、

前記第1畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得て、前記第2畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得ることと、

前記第1中間輪郭特徴行列および前記第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、前記第1変換ニューラルネットワークを用いて、前記第1連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第1中間骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、前記第2変換ニューラルネットワークを用いて、前記第2連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることとを含む、請求項9に記載の人体検出方法。

【請求項13】

前記特徴融合ニューラルネットワークは、第1定向畳み込みニューラルネットワークと、第2定向畳み込みニューラルネットワークと、第3畳み込みニューラルネットワークと、第4畳み込みニューラルネットワークと、第3変換ニューラルネットワークと、第4変換ニューラルネットワークとを含み、

前記特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列および第2ターゲット輪郭特徴行列を得ることは、

前記第1定向畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して、定向畳み込み処理を行い、第1定向骨格特徴行列を得て、前記第3畳み込みニューラルネットワークを用いて、前記第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得ることと、

前記第2定向畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して、定向畳み込み処理を行い、第1定向輪郭特徴行列を得て、前記第4畳み込みニューラルネットワークを用いて、前記第1定向輪郭特徴に対して、畳み込み処理を行い、第2中間輪郭特徴行列を得ることと、

前記第2中間輪郭特徴行列および前記第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、前記第3変換ニューラルネットワークを用いて、前記第3連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第2中間骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得て、前記第4変換ニューラルネットワークを用いて、前記第4連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることとを含む、請求項9に記載の人体検出方法。

【請求項14】

前記特徴融合ニューラルネットワークは、変位推定ニューラルネットワークと、第5変換ニューラルネットワークとを含み、

前記特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列および第2ターゲット輪郭特徴行列を得ることは、

前記第1ターゲット骨格特徴行列および前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第5連結特徴行列を得ることと、

前記第5連結特徴行列を前記変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する変位情報を得て、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとし、前記現在のキ

10

20

30

40

50

ーポイントのペアであるもう1つのキーポイントに対応する三次元特徴行列から、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列を得ることと、
前記ペアであるもう1つのキーポイントから前記現在のキーポイントまでの変位情報に基づいて、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、前記現在のキーポイントに対応する変位特徴行列を得ることと、
各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と、該骨格キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット骨格特徴行列を生成することと、
各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と、該輪郭キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット輪郭特徴行列を生成することと
を含む、請求項9に記載の人体検出方法。

10

【請求項15】

前記人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、前記サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報および人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされている、請求項7～14のいずれか一項に記載の人体検出方法。

20

【請求項16】

人体検出装置であって、前記人体検出装置は、
検出対象画像を取得するように構成されている取得モジュールと、
前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報および人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成されている検出モジュールと、
前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成するように構成されている生成モジュールと
を備え、
前記検出モジュールは、
前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴および輪郭特徴を取得し、取得された骨格特徴および輪郭特徴に対して特徴融合を行うことと、
前記特徴融合の結果に基づいて、前記骨格キーポイントの位置情報および前記輪郭キーポイントの位置情報を決定することと
を行うように構成されている、人体検出装置。

30

【請求項17】

コンピュータ機器であって、
前記コンピュータ機器は、プロセッサと非一時的な記憶媒体とバスとを備え、前記非一時的な記憶媒体には、前記プロセッサによって実行可能な機器読み取り可能な命令が記憶されており、前記コンピュータ機器が実行される場合、前記プロセッサおよび前記非一時的な記憶媒体は、前記バスを介して通信し、前記プロセッサは、前記機器読み取り可能な命令を実行することにより、請求項1～15のいずれか一項に記載の人体検出方法を実行するように構成されている、コンピュータ機器。

40

【請求項18】

コンピュータ読み取り可能な記憶媒体であって、前記コンピュータ読み取り可能な記憶媒体には、コンピュータプログラムが記憶されており、前記コンピュータプログラムは、プロセッサによって実行されると、請求項1～15のいずれか一項に記載の人体検出方法を

50

実行することを前記プロセッサに行わせる、コンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、画像処理技術分野に関し、具体的には、人体検出方法、装置、コンピュータ機器並びに記憶媒体に関する。

【背景技術】

【0002】

ニューラルネットワークが画像、映像、音声、テキストなどの分野に適用されることに伴い、ユーザによりニューラルネットワークに基づいた様々なモデルの精度に求められる要求がますます高くなる。ニューラルネットワークの重要な適用シーンとして、画像において人体検出を行うことであり、人体検出の精細度及び演算データ量に対する要求が高い。

【発明の概要】

【発明が解決しようとする課題】

【0003】

本願の実施例の目的は、人体検出方法、装置、コンピュータ機器並びに記憶媒体を提供することである。

【課題を解決するための手段】

【0004】

第1態様によれば、本願の実施例は、人体検出方法を提供する。前記方法は、検出対象画像を取得することと、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することと、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成することと、を含む。

【0005】

本願の実施例は、検出対象画像から、人体骨格構造を示すための骨格キーポイントの位置情報、人体輪郭を示すための輪郭キーポイントの位置情報を決定し、骨格キーポイントの位置情報、輪郭キーポイントの位置情報に基づいて、人体検出結果を生成し、表現精細度の向上と、演算データ量と、を両立させることができる。

【0006】

また、本願の実施形態において、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を利用して、人体検出結果を得るため、人体を表すための情報は、より豊富で、画像編集、人体体型調整などのような、より広くシーンに適用する。

【0007】

任意選択的な実施形態において、前記輪郭キーポイントは、主輪郭キーポイント及び補助輪郭キーポイントを含み、ここで、2つの隣接する前記主輪郭キーポイントの間に、少なくとも1つの補助輪郭キーポイントが存在する。

【0008】

該実施形態において、主輪郭キーポイントの位置情報及び補助輪郭キーポイントの位置情報により、人体輪郭を表すことで、人体輪郭の識別の精度を一層高くし、情報量をより豊富にする。

【0009】

任意選択的な実施形態において、前記検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、前記検出対象画像に基づいて、前記主輪郭キーポイントの位置情報を決定することと、前記主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定することと、決定された前記人体輪郭情報に基づいて、複数の前記補助輪郭キーポイントの位置情報を決定することと、を含む。

【0010】

該実施形態において、主輪郭キーポイントの位置情報、及び補助輪郭キーポイントの位置

10

20

30

40

50

情報をより正確に位置決めすることができる。

【0011】

任意選択的な実施形態において、前記人体検出結果は、骨格キーポイントマーク及び輪郭キーポイントマークが付加された検出対象画像、前記骨格キーポイントの位置情報及び前記輪郭キーポイントの位置情報を含むデータ群のうちの1つ又は複数を含む。

【0012】

該実施形態において、骨格キーポイントマーク及び輪郭キーポイントマークを含む検出対象画像は、より直感的な視覚的イメージを人々に与えることができる。骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を含むデータ群は、後続の処理をより容易にする。

【0013】

任意選択的な実施形態において、該方法は、前記人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集及び人体画像の貼り付けのうちの1つ又は複数を実行することを更に含む。

【0014】

該実施形態において、表現精細度がより高く、演算データ量がより少ない人体検出結果に基づいて、より高い精度及びより速い速度でより多くの操作を実現させることができる。

【0015】

任意選択的な実施形態において、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことと、特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定することと、を含む。

【0016】

該実施形態において、検出対象画像に対して特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、それによって、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を得ることができる。該方法で得られた人体検出結果によれば、より少ないデータ量で人体を表すことができるだけでなく、人体を表すための人体の骨格特徴及び輪郭特徴を抽出することもでき、表現精細度の向上も両立させる。

【0017】

任意選択的な実施形態において、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことは、前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行い、特徴抽出を行うたびに得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことであって、複数回の特徴抽出を行う場合、 i 回目の特徴融合の特徴融合結果に基づいて $i+1$ 回目の特徴抽出を行い、 i は正整数である、ことを含み、特徴融合結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定することを含む。

【0018】

該実施形態において、検出対象画像に対して少なくとも1回の特徴抽出を行い、特徴抽出を行うたびに得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことで、位置関連関係を有する骨格特徴点及び輪郭特徴点を相互補正することができる。最終的に得られた骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報に、より高い精度を持たせることができる。

【0019】

任意選択的な実施形態において、前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行うことは、1回目の特徴抽出において、事前訓練された第1特徴抽出ネットワーク

10

20

30

40

50

を用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出することと、 $i + 1$ 回の特徴抽出において、事前訓練された第2特徴抽出ネットワークを用いて、 i 回目の特徴融合の特徴融合結果から、前記第1ターゲット骨格特徴行列、及び前記第1ターゲット輪郭特徴行列を抽出することと、を含み、第1特徴抽出ネットワークのネットワークパラメータは、第2特徴抽出ネットワークのネットワークパラメータと異なり、且つ特徴抽出のたびに用いられる第2特徴抽出ネットワークのネットワークパラメータは異なる。

【0020】

該実施例において、骨格特徴と輪郭特徴に対して少なくとも1回の抽出及び少なくとも1回の融合を行うことで、最終的に得られた骨格キーポイントの位置情報及び輪郭キーポイントの位置情報は、より高い精度を有する。

【0021】

任意選択的な実施形態において、抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことは、事前訓練された特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることであって、前記第2ターゲット骨格特徴行列は、三次元骨格特徴行列であり、該三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、前記二次元骨格特徴行列における各要素の値は、該要素に対応する画素点が、対応する骨格キーポイントに属する確率を表し、前記第2ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、該三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、前記二次元輪郭特徴行列における各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表し、特徴融合のたびに用いられる特徴融合ニューラルネットワークのネットワークパラメータは異なる、ことを含む。

【0022】

該実施形態において、事前訓練された特徴融合ネットワークに基づいて骨格特徴及び輪郭特徴を融合することで、より良好な特徴融合結果を得ることができる。最終的に得られた骨格キーポイントの位置情報及び輪郭キーポイントの位置情報に、より高い精度を持たせることができる。

【0023】

任意選択的な実施形態において、最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定することは、最終回の特徴融合で得られた第2ターゲット骨格特徴行列に基づいて、前記骨格キーポイントの位置情報を決定することと、最終回の特徴融合で得られた第2ターゲット輪郭特徴行列に基づいて、前記輪郭キーポイントの位置情報を決定することと、を含む。

【0024】

該実施形態において、少なくとも1回の特徴抽出及び特徴融合により、最終的に得られた骨格キーポイントの位置情報及び輪郭キーポイントの位置情報に、より高い精度を持たせる。

【0025】

任意選択的な実施形態において、前記第1特徴抽出ネットワークは、共通特徴抽出ネットワーク、第1骨格特徴抽出ネットワーク及び第1輪郭特徴抽出ネットワークを含み、第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出することは、前記共通特徴抽出ネットワークを用いて、前記検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得ることと、前記第1骨格特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1骨格特徴行列を得て、前記第1骨格特徴抽出ネットワークにおける第1ターゲット畳み込み層から、第2骨格特徴行列を取得し、前記第1

10

20

30

40

50

骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得ることであって、前記第1ターゲット畳み込み層は、前記第1骨格特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、前記第1輪郭特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1輪郭特徴行列を得て、前記第1輪郭特徴抽出ネットワークにおける第2ターゲット畳み込み層から、第2輪郭特徴行列を取得し、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることであって、前記第2ターゲット畳み込み層は、前記第1輪郭特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、を含む。

【0026】

該実施形態において、共通特徴抽出ネットワークを用いて、骨格特徴及び輪郭特徴を抽出し、検出対象画像における、骨格特徴及び輪郭特徴以下のほかの特徴を除去した後、第1骨格特徴抽出ネットワークを用いて、骨格特徴を意図的に抽出し、第1輪郭特徴抽出ネットワークを用いて、輪郭特徴を意図的に抽出する。従って、必要な演算量がより少ない。

【0027】

任意選択的な実施形態において、第1骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得ることは、前記第1骨格特徴行列及び前記第2骨格特徴行列に対して連結処理を行い、第1連結骨格特徴行列を得ることと、前記第1連結骨格特徴行列を次元変換処理し、前記第1ターゲット骨格特徴行列を得ることと、を含み、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることは、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に対して連結処理を行い、第1連結輪郭特徴行列を得ることと、前記第1連結輪郭特徴行列を次元変換処理し、前記ターゲット輪郭特徴行列を得ることと、を含み、前記第1ターゲット骨格特徴行列の次元は、前記第1ターゲット輪郭特徴行列の次元と同じであり、且つ前記第1ターゲット骨格特徴行列と前記第1ターゲット輪郭特徴行列は、同一の次元における次元数が同じである。

【0028】

該実施形態において、第1骨格特徴行列及び前記第2骨格特徴行列に対して連結処理を行うことで、第1ターゲット骨格特徴行列に、より豊富な骨格特徴情報を持たせる。それと同時に、第1輪郭特徴行列及び前記第2輪郭特徴行列に対して連結処理を行い、第1ターゲット輪郭特徴行列に、より豊富な骨格特徴情報を持たせる。後続の特徴融合過程において、より高い精度で、骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を抽出することができる。

【0029】

任意選択的な実施形態において、前記特徴融合ニューラルネットワークは、第1畳み込みニューラルネットワーク、第2畳み込みニューラルネットワーク、第1変換ニューラルネットワーク及び第2変換ニューラルネットワークを含み、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、前記第1畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得て、前記第2畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得ることと、前記第1中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、前記第1変換ニューラルネットワークを用いて、前記第1連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第1中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、前記第2変換ニューラルネットワークを用いて、前記第2連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含む。

【0030】

該実施形態において、前記第1中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、連結処理結果に基づいて、第2ターゲット骨格特徴行列を得ることで、骨格特徴及び輪郭特徴を融合し、輪郭特徴を用いて、抽出された骨格特徴を補正することを実現させる。また、前記第1中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、連結処理結果に基づいて、第2ターゲット輪郭特徴行列を得ることで、骨格特徴及び輪郭特徴を融合し、骨格特徴を用いて、抽出された輪郭特徴を補正することを実現させる。更に、より高い精度で、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報を抽出することができる。

【0031】

任意選択的な実施形態において、前記特徴融合ニューラルネットワークは、第1定向畳み込みニューラルネットワーク、第2定向畳み込みニューラルネットワーク、第3畳み込みニューラルネットワーク、第4畳み込みニューラルネットワーク、第3変換ニューラルネットワーク、及び第4変換ニューラルネットワークを含み、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、前記第1定向畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して、定向畳み込み処理を行い、第1定向骨格特徴行列を得て、第3畳み込みニューラルネットワークを用いて、前記第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得ることと、前記第2定向畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して、定向畳み込み処理を行い、第1定向輪郭特徴行列を得て、第4畳み込みニューラルネットワークを用いて、前記第1定向輪郭特徴に対して、畳み込み処理を行い、第2中間輪郭特徴行列を得ることと、前記第2中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、第3変換ニューラルネットワークを用いて、前記第3連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第2中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得て、第4変換ニューラルネットワークを用いて、前記第4連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含む。

【0032】

該実施形態において、定向畳み込みの方式により、特徴を融合処理することで、より高い精度で、骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を抽出することができる。

【0033】

任意選択的な実施形態において、前記特徴融合ニューラルネットワークは、変位推定ニューラルネットワーク、第5変換ニューラルネットワークを含み、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第5連結特徴行列を得ることと、前記第5連結特徴行列を前記変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する変位情報を得て、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとし、該現在のキーポイントのペアであるもう1つのキーポイントに対応する三次元特徴行列から、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列を得ることと、前記ペアであるもう1つのキーポイントから前記現在のキーポイントまでの変位情報に基づいて、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応する変位特徴行列を得ることと、各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と、該骨格キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポ

10

20

30

40

50

イントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット骨格特徴行列を生成することと、各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と、該輪郭キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット輪郭特徴行列を生成することと、を含む。

【0034】

該実施形態において、骨格キーポイント及び輪郭キーポイントに対して変位変換を行うことで特徴融合を実現させる。従って、より高い精度で、骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を抽出することができる。

【0035】

任意選択的な実施形態において、前記人体検出方法は、人体検出モデルにより実現され、前記人体検出モデルは、前記第1特徴抽出ネットワーク及び／又は前記特徴融合ニューラルネットワークを含み、前記人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、前記サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報、及び人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされている。

【0036】

該実施形態において、該訓練方法で得られた人体検出モデルは、より高い検出精度を持ち、該人体検出モデルによれば、表現精細度と演算データ量を両立させる人体検出結果を得ることができる。

【0037】

第2態様によれば、本願の実施例は、人体検出装置を更に提供する。前記装置は、検出対象画像を取得するように構成される取得モジュールと、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成される検出モジュールと、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成するように構成される生成モジュールと、を備える。

【0038】

第3態様によれば、本願の実施例は、コンピュータ機器を更に提供する。前記コンピュータ機器は、プロセッサと、非一時的記憶媒体と、バスと、を備え、前記非一時的記憶媒体に、前記プロセッサによる実行可能な機器可読命令が記憶されており、コンピュータ機器が運転される場合、前記プロセッサと前記記憶媒体とは、バスを介して通信し、前記機器可読命令が前記プロセッサにより実行される場合、上記第1態様又は第1態様におけるいずれか1つの可能な実施形態における工程を実行する。

【0039】

第4態様によれば、本願の実施例は、コンピュータ可読記憶媒体を更に提供する。該コンピュータ可読記憶媒体に、コンピュータプログラムが記憶されており、該コンピュータプログラムがプロセッサにより実行される場合、上記第1態様又は第1態様におけるいずれか1つの可能な実施形態における工程を実行する。

例えば、本願は以下の項目を提供する。

(項目1)

人体検出方法であって、前記方法は、

検出対象画像を取得することと、

前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報

、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することと、

前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、

10

20

30

40

50

人体検出結果を生成することと、を含むことを特徴とする、人体検出方法。

(項目2)

前記輪郭キーポイントは、主輪郭キーポイント及び補助輪郭キーポイントを含み、ここで、2つの隣接する前記主輪郭キーポイントの間に、少なくとも1つの補助輪郭キーポイントが存在することを特徴とする

項目1に記載の人体検出方法。

(項目3)

前記検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、

前記検出対象画像に基づいて、前記主輪郭キーポイントの位置情報を決定することと、
前記主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定することと、
決定された前記人体輪郭情報に基づいて、複数の前記補助輪郭キーポイントの位置情報を決定することと、を含むことを特徴とする

項目2に記載の人体検出方法。

(項目4)

前記人体検出結果は、
骨格キーポイントマーク及び輪郭キーポイントマークが付加された検出対象画像、前記骨格キーポイントの位置情報及び前記輪郭キーポイントの位置情報を含むデータ群のうちの1つ又は複数を含むことを特徴とする

項目1から3のいずれか一項に記載の人体検出方法。

(項目5)

前記方法は、
前記人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集及び人体画像の貼り付けのうちの1つ又は複数を実行することを更に含むことを特徴とする

項目4に記載の人体検出方法。

(項目6)

前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、

前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことと、特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定することと、を含むことを特徴とする

項目1から5のいずれか一項に記載の人体検出方法。

(項目7)

前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことは、

前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行い、特徴抽出を行うたびに得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことであって、複数回の特徴抽出を行う場合、 i 回目の特徴融合の特徴融合結果に基づいて $i+1$ 回目の特徴抽出を行い、 i は正整数である、ことを含み、

特徴融合結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定することは、

最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定することを含むことを特徴とする

項目6に記載の人体検出方法。

(項目8)

前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行うことは、

1回目の特徴抽出において、事前訓練された第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行

10

20

30

40

50

列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出することと、

$i + 1$ 回の特徴抽出において、事前訓練された第2特徴抽出ネットワークを用いて、 i 回目の特徴融合の特徴融合結果から、前記第1ターゲット骨格特徴行列、及び前記第1ターゲット輪郭特徴行列を抽出することと、を含み、

第1特徴抽出ネットワークのネットワークパラメータは、第2特徴抽出ネットワークのネットワークパラメータと異なり、且つ特徴抽出のたびに用いられる第2特徴抽出ネットワークのネットワークパラメータは異なることを特徴とする

項目7に記載の人体検出方法。

(項目9)

抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行うことは、

事前訓練された特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることを含み、

前記第2ターゲット骨格特徴行列は、三次元骨格特徴行列であり、該三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、前記二次元骨格特徴行列における各要素の値は、該要素に対応する画素点が、対応する骨格キーポイントに属する確率を表し、

前記第2ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、該三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、前記二次元輪郭特徴行列における各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表し、

特徴融合のたびに用いられる特徴融合ニューラルネットワークのネットワークパラメータは異なる、ことを含むことを特徴とする

項目8に記載の人体検出方法。

(項目10)

前記第1特徴抽出ネットワークは、共通特徴抽出ネットワーク、第1骨格特徴抽出ネットワーク及び第1輪郭特徴抽出ネットワークを含み、

第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出することは、

前記共通特徴抽出ネットワークを用いて、前記検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得ることと、

前記第1骨格特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1骨格特徴行列を得て、前記第1骨格特徴抽出ネットワークにおける第1ターゲット畳み込み層から、第2骨格特徴行列を取得し、前記第1骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得ることであって、前記第1ターゲット畳み込み層は、前記第1骨格特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、

前記第1輪郭特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1輪郭特徴行列を得て、前記第1輪郭特徴抽出ネットワークにおける第2ターゲット畳み込み層から、第2輪郭特徴行列を取得し、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることであって、前記第2ターゲット畳み込み層は、前記第1輪郭特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、を含むことを特徴とする

項目8に記載の人体検出方法。

(項目11)

第1骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得ることは、

前記第1骨格特徴行列及び前記第2骨格特徴行列に対して連結処理を行い、第1連結骨

10

20

30

40

50

格特徴行列を得ることと、前記第1連結骨格特徴行列を次元変換処理し、前記第1ターゲット骨格特徴行列を得ることと、を含み、

前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることは、

前記第1輪郭特徴行列及び前記第2輪郭特徴行列に対して連結処理を行い、第1連結輪郭特徴行列を得ることと、前記第1連結輪郭特徴行列を次元変換処理し、前記ターゲット輪郭特徴行列を得ることと、を含み、

前記第1ターゲット骨格特徴行列の次元は、前記第1ターゲット輪郭特徴行列の次元と同じであり、且つ前記第1ターゲット骨格特徴行列と前記第1ターゲット輪郭特徴行列は、同一の次元における次元数が同じであることを特徴とする

項目10に記載の人体検出方法。

(項目12)

前記特徴融合ニューラルネットワークは、第1畳み込みニューラルネットワーク、第2畳み込みニューラルネットワーク、第1変換ニューラルネットワーク及び第2変換ニューラルネットワークを含み、

特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、

前記第1畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得て、前記第2畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得ることと、

前記第1中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、前記第1変換ニューラルネットワークを用いて、前記第1連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、

前記第1中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、前記第2変換ニューラルネットワークを用いて、前記第2連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含むことを特徴とする

項目9に記載の人体検出方法。

(項目13)

前記特徴融合ニューラルネットワークは、第1定向畳み込みニューラルネットワーク、第2定向畳み込みニューラルネットワーク、第3畳み込みニューラルネットワーク、第4畳み込みニューラルネットワーク、第3変換ニューラルネットワーク、及び第4変換ニューラルネットワークを含み、

特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、

前記第1定向畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して、定向畳み込み処理を行い、第1定向骨格特徴行列を得て、第3畳み込みニューラルネットワークを用いて、前記第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得ることと、

前記第2定向畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して、定向畳み込み処理を行い、第1定向輪郭特徴行列を得て、第4畳み込みニューラルネットワークを用いて、前記第1定向輪郭特徴に対して、畳み込み処理を行い、第2中間輪郭特徴行列を得ることと、

前記第2中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、第3変換ニューラルネットワークを用いて、前記第3連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、

前記第2中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い

10

20

30

40

50

、第4連結特徴行列を得て、第4変換ニューラルネットワークを用いて、前記第4連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含むことを特徴とする

項目9に記載の人体検出方法。

(項目14)

前記特徴融合ニューラルネットワークは、変位推定ニューラルネットワーク、第5変換ニューラルネットワークを含み、

特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることは、

前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第5連結特徴行列を得ることと、

前記第5連結特徴行列を前記変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する変位情報を得て、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとし、該現在のキーポイントのペアであるもう1つのキーポイントに対応する三次元特徴行列から、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列を得ることと、

前記ペアであるもう1つのキーポイントから前記現在のキーポイントまでの変位情報に基づいて、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応する変位特徴行列を得ることと、

各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と、該骨格キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット骨格特徴行列を生成することと、

各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と、該輪郭キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット輪郭特徴行列を生成することと、を含むことを特徴とする

項目9に記載の人体検出方法。

(項目15)

前記人体検出方法は、人体検出モデルにより実現され、前記人体検出モデルは、前記第1特徴抽出ネットワーク及び／又は前記特徴融合ニューラルネットワークを含み、

前記人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、前記サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報、及び人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされていることを特徴とする

項目1から14のいずれか一項に記載の人体検出方法。

(項目16)

人体検出装置であって、前記装置は、

検出対象画像を取得するように構成される取得モジュールと、

前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成される検出モジュールと、

前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成するように構成される生成モジュールと、を備えることを特徴とする

、人体検出装置。

(項目 1 7)

前記輪郭キーポイントは、主輪郭キーポイント及び補助輪郭キーポイントを含み、ここで、2つの隣接する前記主輪郭キーポイントの間に、少なくとも1つの補助輪郭キーポイントが存在することを特徴とする

項目 1 6 に記載の人体検出装置。

(項目 1 8)

前記検出モジュールは、前記検出対象画像に基づいて、前記主輪郭キーポイントの位置情報を決定し、前記主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定し、決定された前記人体輪郭情報に基づいて、複数の前記補助輪郭キーポイントの位置情報を決定するという方式で、前記検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成されることを特徴とする

項目 1 7 に記載の人体検出装置。

(項目 1 9)

前記人体検出結果は、

骨格キーポイントマーク及び輪郭キーポイントマークが付加された検出対象画像、前記骨格キーポイントの位置情報及び前記輪郭キーポイントの位置情報を含むデータ群のうちの1つ又は複数を含むことを特徴とする

項目 1 6 から 1 8 のいずれか一項に記載の人体検出装置。

(項目 2 0)

前記人体検出装置は、

前記人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集及び人体画像の貼り付けのうちの1つ又は複数を実行するように構成される実行モジュールを更に備えることを特徴とする

項目 1 9 に記載の人体検出装置。

(項目 2 1)

前記検出モジュールは、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定するという方式で、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成されることを特徴とする

項目 1 6 から項目 2 0 のいずれか一項に記載の人体検出装置。

(項目 2 2)

前記検出モジュールは、前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行い、特徴抽出を行うたびに得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、複数回の特徴抽出を行う場合、 i 回目の特徴融合の特徴融合結果に基づいて $i + 1$ 回目の特徴抽出を行い、 i は正整数であるという方式で、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うように構成され、

前記検出モジュールは、最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定するという方式で、特徴融合結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成されることを特徴とする

項目 2 1 に記載の人体検出装置。

(項目 2 3)

前記検出モジュールは、1 回目の特徴抽出において、事前訓練された第 1 特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第 1 ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第 1

10

20

30

40

50

ターゲット輪郭特徴行列を抽出し、 $i + 1$ 回の特徴抽出において、事前訓練された第 2 特徴抽出ネットワークを用いて、 i 回目の特徴融合の特徴融合結果から、前記第 1 ターゲット骨格特徴行列、及び前記第 1 ターゲット輪郭特徴行列を抽出するという方式で、前記検出対象画像に基づいて、少なくとも 1 回の特徴抽出を行うように構成され、

第 1 特徴抽出ネットワークのネットワークパラメータは、第 2 特徴抽出ネットワークのネットワークパラメータと異なり、且つ特徴抽出のたびに用いられる第 2 特徴抽出ネットワークのネットワークパラメータは異なることを特徴とする

項目 2 2 に記載の人体検出装置。

(項目 2 4)

前記検出モジュールは、事前訓練された特徴融合ニューラルネットワークを用いて、前記第 1 ターゲット骨格特徴行列及び前記第 1 ターゲット輪郭特徴行列に対して特徴融合を行い、第 2 ターゲット骨格特徴行列及び第 2 ターゲット輪郭特徴行列を得るという方式で、抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行うように構成され、

前記第 2 ターゲット骨格特徴行列は、三次元骨格特徴行列であり、該三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、前記二次元骨格特徴行列における各要素の値は、該要素に対応する画素点が、対応する骨格キーポイントに属する確率を表し、

前記第 2 ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、該三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、前記二次元輪郭特徴行列における各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表し、

特徴融合のたびに用いられる特徴融合ニューラルネットワークのネットワークパラメータは異なることを特徴とする

項目 2 3 に記載の人体検出装置。

(項目 2 5)

第 1 特徴抽出ネットワークは、共通特徴抽出ネットワーク、第 1 骨格特徴抽出ネットワーク及び第 1 輪郭特徴抽出ネットワークを含み、

前記検出モジュールは、

前記共通特徴抽出ネットワークを用いて、前記検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得ることと、

前記第 1 骨格特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第 1 骨格特徴行列を得て、前記第 1 骨格特徴抽出ネットワークにおける第 1 ターゲット畳み込み層から、第 2 骨格特徴行列を取得し、前記第 1 骨格特徴行列及び前記第 2 骨格特徴行列に基づいて、前記第 1 ターゲット骨格特徴行列を得ることであって、前記第 1 ターゲット畳み込み層は、前記第 1 骨格特徴抽出ネットワークにおける、最後の 1 つの畳み込み層以外のいずれか 1 つの畳み込み層であることと、

前記第 1 輪郭特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第 1 輪郭特徴行列を得て、前記第 1 輪郭特徴抽出ネットワークにおける第 2 ターゲット畳み込み層から、第 2 輪郭特徴行列を取得し、前記第 1 輪郭特徴行列及び前記第 2 輪郭特徴行列に基づいて、前記第 1 ターゲット輪郭特徴行列を得ることであって、前記第 2 ターゲット畳み込み層は、前記第 1 輪郭特徴抽出ネットワークにおける、最後の 1 つの畳み込み層以外のいずれか 1 つの畳み込み層であることと、を実行する、

という方式で、第 1 特徴抽出ネットワークを用いて、検出対象画像から、前記第 1 ターゲット骨格特徴行列及び前記第 1 ターゲット輪郭特徴行列を抽出するように構成されることを特徴とする

項目 2 3 に記載の人体検出装置。

(項目 2 6)

前記検出モジュールは、前記第 1 骨格特徴行列及び前記第 2 骨格特徴行列に対して連結処理を行い、第 1 連結骨格特徴行列を得て、前記第 1 連結骨格特徴行列を次元変換処理し、前記第 1 ターゲット骨格特徴行列を得るという方式で、前記第 1 骨格特徴行列及び前記

10

20

30

40

50

第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得るように構成され、
前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることは、

前記第1輪郭特徴行列及び前記第2輪郭特徴行列に対して連結処理を行い、第1連結輪郭特徴行列を得ることと、

前記第1連結輪郭特徴行列を次元変換処理し、前記ターゲット輪郭特徴行列を得ることと、を含み、

前記第1ターゲット骨格特徴行列の次元は、前記第1ターゲット輪郭特徴行列の次元と同じであり、且つ前記第1ターゲット骨格特徴行列と前記第1ターゲット輪郭特徴行列は、同一の次元における次元数が同じであることを特徴とする

項目25に記載の人体検出装置。

(項目27)

前記特徴融合ニューラルネットワークは、第1畳み込みニューラルネットワーク、第2畳み込みニューラルネットワーク、第1変換ニューラルネットワーク及び第2変換ニューラルネットワークを含み、

前記検出モジュールは、

前記第1畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得て、前記第2畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得ることと、

前記第1中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、前記第1変換ニューラルネットワークを用いて、前記第1連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、

前記第1中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、前記第2変換ニューラルネットワークを用いて、前記第2連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を実行する、

という方式で、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るように構成されることを特徴とする

項目24に記載の人体検出装置。

(項目28)

前記特徴融合ニューラルネットワークは、第1定向畳み込みニューラルネットワーク、第2定向畳み込みニューラルネットワーク、第3畳み込みニューラルネットワーク、第4畳み込みニューラルネットワーク、第3変換ニューラルネットワーク、及び第4変換ニューラルネットワークを含み、

前記検出モジュールは、

前記第1定向畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して、定向畳み込み処理を行い、第1定向骨格特徴行列を得て、第3畳み込みニューラルネットワークを用いて、前記第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得ることと、

前記第2定向畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して、定向畳み込み処理を行い、第1定向輪郭特徴行列を得て、第4畳み込みニューラルネットワークを用いて、前記第1定向輪郭特徴行列に対して、畳み込み処理を行い、第2中間輪郭特徴行列を得ることと、

前記第2中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、第3変換ニューラルネットワークを用いて、前記第3連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、

前記第2中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得て、第4変換ニューラルネットワークを用いて、前記第4連結特

10

20

30

40

50

徴行列に対して次元変換を行い、前記第 2 ターゲット輪郭特徴行列を得ることと、を実行する、

という方式で、特徴融合ニューラルネットワークを用いて、前記第 1 ターゲット骨格特徴行列及び前記第 1 ターゲット輪郭特徴行列に対して特徴融合を行い、第 2 ターゲット骨格特徴行列及び第 2 ターゲット輪郭特徴行列を得るように構成されることを特徴とする

項目 2 4 に記載の人体検出装置。

(項目 2 9)

前記特徴融合ニューラルネットワークは、変位推定ニューラルネットワーク、第 5 変換ニューラルネットワークを含み、

前記検出モジュールは、

前記第 1 ターゲット骨格特徴行列及び前記第 1 ターゲット輪郭特徴行列に対して連結処理を行い、第 5 連結特徴行列を得ることと、

前記第 5 連結特徴行列を前記変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの 1 つのキーポイントがもう 1 つのキーポイントまでに移動する変位情報を得て、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとし、該現在のキーポイントのペアであるもう 1 つのキーポイントに対応する三次元特徴行列から、前記ペアであるもう 1 つのキーポイントに対応する二次元特徴行列を得ることと、

前記ペアであるもう 1 つのキーポイントから前記現在のキーポイントまでの変位情報に基づいて、前記ペアであるもう 1 つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応する変位特徴行列を得ることと、

各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と、該骨格キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポイントの連結二次元特徴行列を前記第 5 変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第 2 ターゲット骨格特徴行列を生成することと、

各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と、該輪郭キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を前記第 5 変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第 2 ターゲット輪郭特徴行列を生成することと、を実行する、

という方式で、特徴融合ニューラルネットワークを用いて、前記第 1 ターゲット骨格特徴行列及び前記第 1 ターゲット輪郭特徴行列に対して特徴融合を行い、第 2 ターゲット骨格特徴行列及び第 2 ターゲット輪郭特徴行列を得るように構成されることを特徴とする

項目 2 4 に記載の人体検出装置。

(項目 3 0)

前記人体検出装置の人体検出機能は、人体検出モデルにより実現され、前記人体検出モデルは、前記第 1 特徴抽出ネットワーク及び／又は前記特徴融合ニューラルネットワークを含み、

前記人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、前記サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報、及び人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされていることを特徴とする

項目 1 6 から 2 9 のいずれか一項に記載の人体検出装置。

(項目 3 1)

コンピュータ機器であって、プロセッサと、非一時的記憶媒体と、バスと、を備え、前記非一時的記憶媒体に、前記プロセッサによる実行可能な機器可読命令が記憶されており、コンピュータ機器が実行される場合、前記プロセッサと前記非一時的記憶媒体は、バス

10

20

30

40

50

を介して通信し、前記機器可読命令は、前記プロセッサにより実行され、項目 1 から 1 5 のいずれか一項に記載の方法のステップを実行することを特徴とする、コンピュータ機器。

(項目 3 2)

コンピュータ可読記憶媒体であって、該コンピュータ可読記憶媒体に、コンピュータプログラムが記憶されており、該コンピュータプログラムがプロセッサにより実行され、項目 1 から 1 5 のいずれか一項に記載の方法のステップを実行することを特徴とする、コンピュータ可読記憶媒体。

【発明の効果】

【0040】

本願の実施例は、検出対象画像から、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定し、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報に基づいて、人体検出結果を生成する。表現精細度の向上と演算データ量を両立させる。

【0041】

本願の上記目的、特徴及び利点を明確かつ分かりやすくするために、以下、好適な実施例を挙げて、添付図面を参照しながら、以下のように、詳しく説明する。

【図面の簡単な説明】

【0042】

【図 1】本願の実施例による人体検出方法を示すフローチャートである。

【図 2 a】本願の実施例による輪郭キーポイント及び骨格キーポイントの位置の例を示す。

【図 2 b】本願の実施例による主輪郭キーポイント及び補助輪郭キーポイントの位置の例を示す。

【図 2 c】本願の実施例によるもう 1 つの主輪郭キーポイント及び補助輪郭キーポイントの位置の例を示す。

【図 2 d】本願の実施例によるもう 1 つの主輪郭キーポイント及び補助輪郭キーポイントの位置の例を示す。

【図 3】本願の実施例による第 1 特徴抽出ネットワークの構造を示す概略図である。

【図 4】本願の実施例による特徴抽出方法を示すフローチャートである。

【図 5】本願の実施例による特徴融合ネットワークの構造を示す概略図である。

【図 6】本願の実施例による特徴融合方法を示すフローチャートである。

【図 7】本願の実施例によるもう 1 つの特徴融合ネットワークの構造を示す概略図である。

【図 8】本願の実施例によるもう 1 つの特徴融合方法を示すフローチャートである。

【図 9 a】本願の実施例による散乱畳み込み演算子を用いて反復更新を行う過程を示す概略図である。

【図 9 b】本願の実施例による集約畳み込み演算子を用いて反復更新を行う過程を示す概略図である。

【図 10】本願の実施例によるもう 1 つの特徴融合ネットワークの構造を示す概略図である。

【図 11】本願の実施例によるもう 1 つの特徴融合方法を示すフローチャートである。

【図 12】本願の実施例による骨格キーポイント及び輪郭キーポイントの例を示す。

【図 13】本願の実施例による二次元特徴行列における要素に対する変位変換の具体的な例を示す。

【図 14】本願の実施例による第 2 特徴抽出ネットワークの構造を示す概略図である。

【図 15】本願の実施例による人体検出装置を示す概略図である。

【図 16】本願の実施例によるコンピュータ機器を示す概略図である。

【0043】

本願の実施例の技術的解決手段をより明確に説明するために、以下、実施例に必要な図面を簡単に説明する。下記図面は、説明のみの目的で本願の幾つかの実施形態を図示しており、本願を制限するものではないことが理解されるべきである。当業者は、創造的労働を行わない前提で、これらの図面に基づいてその他の関連する図面を得ることもできる。図

10

20

30

40

50

面において、同一又は類似した符号は、同一の要素又は均等な要素を表す。1つの符号が1つの図面において定義されていると、後続の図面において、これに対して定義及び説明を行う必要がない。

【発明を実施するための形態】

【0044】

本願の実施例の目的、技術的解決手段及び利点をより明確にするために、以下、本願の実施例における図面を参照しながら、本願の実施例における具体的な技術的解決手段をを明瞭かつ完全に説明する。勿論、記述される実施例は、全ての実施例ではなく、ただ本出願の一部の実施例である。一般的には、図面に説明されて示されている本願の実施例のユニットは、多岐にわたる異なる構成で配置及び設計することができる。従って、図面を参照しながら提供される本願の実施例の下記詳細な説明は、特許請求されている本願の範囲を限定するものではなく、本願の実施例を示すためのものに過ぎない。本願の実施例に基づいて、当業者が創造的な労力なしに得られる他の実施例の全ては、本願の保護の範囲に含まれる。

【0045】

検討によれば、人体検出を行う時に、一般的には、骨格キーポイント検出法及びセマンティックセグメンテーション法という2つの方式がある。

【0046】

骨格キーポイント検出法について、該方法において、ニューラルネットワークモデルにより、画像から、人体の骨格キーポイントを抽出し、骨格キーポイントに基づいて、対応する人体検出結果を得る。該人体検出方法において、簡略な人体表現方法を用いるため、データ量がより少ない。従って、該方法で得られた人体検出結果に基づいて他の後続処理を行う場合、必要な演算量も少ない。例えば、挙動検出、人体姿勢に基づいたマンマシンインタラクションなどのような人体姿勢、動作認識などの分野に広く適用される。しかし、該方法は、人体の輪郭情報を抽出できないため、得られた人体検出結果の表現精細度が低い。

【0047】

セマンティックセグメンテーション法について、該方法において、セマンティックセグメンテーションモデルにより、画像における各画素点が人体に属する確率を識別し、画像における各画素点が人体に属する確率に基づいて、人体検出結果を得ることができるが、該人体検出方法によれば、人体の輪郭情報を完全に得ることができるが、人体認識結果に含まれる演算データ量が大きい。

【0048】

従って、表現精細度と演算データ量を両立させることができる早急に解決しなければならない課題である。

【0049】

上記検討によれば、本願は、人体検出方法、装置、コンピュータ機器及び記憶媒体を提供し、検出対象画像に対して特徴抽出を行い、人体の骨格特徴及び輪郭特徴を得て、抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、それによって、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を得ることができる。該方法で得られた人体検出結果は、データ量がより少なく、人体の骨格特徴及び輪郭特徴を反映すると共に、表現精細度の向上も実現させる。

【0050】

また、本願の実施例において、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を用いて、人体検出結果を得るため、人体を表す情報がより豊富であり、適用シーンがより広い。

【0051】

従来の人検出方式における欠点に対して、繰り返した実践及び鋭意検討を行う必要がある。従って、既存の課題のマイニング及び本願で提供される解決手段は、いずれも本願の範囲内に含まれる。

【0052】

以下、本願の実施例による人体検出方法を詳しく説明する。該人体検出方法は、例えば、コンピュータのような、データ処理能力を有するいかなる装置に適用可能である。

【0053】

図1は、本願の実施例による人体検出方法を示すフローチャートである。ここで、S101において、検出対象画像を取得する。

【0054】

S102において、検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定する。

【0055】

S103において、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報に基づいて、人体検出結果を生成する。

【0056】

以下、上記S101～S103についてそれぞれ説明する。

【0057】

I：上記S101において、検出対象画像は、例えば、ターゲット位置に取り付けられたカメラにより撮られた検出対象画像、ほかのコンピュータ機器から送信された検出対象画像、ローカルデータベースから読み出された事前記憶した検出対象画像等であってもよい。検出対象画像には、人体画像が含まれてもよく、人体画像が含まれなくてもよい。検出対象画像に人体画像が含まれる場合、本願の実施例で提供される人体検出方法により、最終的な人体検出結果を得ることができる。検出対象画像に人体画像が含まれない場合、得られた人体検出結果は、例えば、殻である。

【0058】

II：上記S102において、図2aに示すように、骨格キーポイントは、人体の骨格特徴を表すために用いられる。該骨格特徴は、人体の関節部位の特徴を含む。関節は、例えば、肘関節、手首関節、肩関節、頸関節、股関節、膝関節、足首関節等である。例示的には、人体の頭部に骨格キーポイントを設けてもよい。

【0059】

輪郭キーポイントは、人体の輪郭特徴を表すために用いられる。それは、図2aに示すように、主輪郭キーポイントを含んでもよく、又は図2b～図2dに示すように、主輪郭キーポイント及び補助輪郭キーポイントを含んでもよい。ここで、図2b～図2dは、図2aにおける枠内の部位の部分図である。

【0060】

ここで、主輪郭キーポイントは、人体関節部位輪郭を示すための輪郭キーポイントであり、図2aに示すように、例えば、肘関節の輪郭、手首関節の輪郭、肩関節の輪郭、頸関節の輪郭、股関節の輪郭、膝関節の輪郭、足首関節の輪郭などが挙げられる。それは、一般的には、対応する関節部位の骨格キーポイントに対応する。

【0061】

補助輪郭キーポイントは、人体関節部位間の輪郭を示すための輪郭キーポイントである。2つの隣接する主輪郭キーポイント間の補助輪郭キーポイントは、少なくとも1つがある。例えば、図2bに示す例において、2つの主輪郭キーポイント間の補助輪郭キーポイントは、1つである。図2cに示す例において、2つの主輪郭キーポイント間の補助輪郭キーポイントは、2つがある。図2dに示す例において、2つの主輪郭キーポイント間の補助輪郭キーポイントは、3つがある。

【0062】

上記図面及び記述的説明に係る骨格キーポイント及び輪郭キーポイントは、本願を理解しやすくするための例だけである。実際の適用において、実際のシーンに応じて骨格キーポイント及び輪郭キーポイントの数及び位置を適当に調整することができる。本願は、これを限定するものではない。

【0063】

輪郭キーポイントが主輪郭キーポイント及び補助輪郭キーポイントを含む場合について、下記方式で、検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定することができる。

【0064】

検出対象画像に基づいて、主輪郭キーポイントの位置情報を決定し、主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定し、決定された人体輪郭情報に基づいて、複数の補助輪郭キーポイントの位置情報を決定する。

【0065】

輪郭キーポイントが主輪郭キーポイントを含む場合、検出対象画像に基づいて、主輪郭キーポイントの位置情報を直接的に決定すればよい。

【0066】

本願の実施例は、検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定する具体的な方法を提供する。

【0067】

検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、特徴融合結果に基づいて、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報を決定する。

【0068】

検出対象画像に基づいて、骨格特徴及び輪郭特徴を抽出する時に、下記A又はBのうちのいずれか1つを用いてもよいが、これらに限定されない。

【0069】

Aにおいて、検出対象画像に対して、一回の特徴抽出を行い、該特徴抽出で得られた骨格特徴及び輪郭特徴に対して特徴融合を行う。

【0070】

Bにおいて、検出対象画像に対して、複数回の特徴抽出を行い、毎回の特徴抽出を行った後、該特徴抽出で得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、最終回の特徴融合の特徴融合結果に基づいて、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報を決定する。

【0071】

以下、まずAを具体的に説明する。

【0072】

Aの場合、該特徴融合の特徴融合結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報及び人体輪郭を示すための輪郭キーポイントの位置情報を決定する。

【0073】

以下、a1及びa2において、特徴抽出過程及び特徴融合過程をそれぞれ説明する。

【0074】

a1：特徴抽出過程：

事前訓練された第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出することができる。

【0075】

具体的には、図3に示すように、本願の実施例は、第1特徴抽出ネットワークの構造を示す概略図を提供する。第1特徴抽出ネットワークは、共通特徴抽出ネットワーク、第1骨格特徴抽出ネットワーク及び第1輪郭特徴抽出ネットワークを含む。

【0076】

図4に示すように、本願の実施例は、図3で提供される第1特徴抽出ネットワークにより、検出対象画像から、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列を抽出する具体的な過程を更に提供する。それは、下記ステップを含む。

【0077】

10

20

30

40

50

S 4 0 1において、共通特徴抽出ネットワークを用いて、検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得る。

【0078】

具体的な実施において、検出対象画像は、画像行列として表すことが可能である。検出対象画像は、例えば、グレースケール画像のような単一色チャンネル画像である場合、二次元画像行列として表れることが可能である。二次元画像行列における各要素は、検出対象画像の画素点に一対一に対応する。二次元画像行列における各要素の値は、各要素に対応する画素点の画素値である。検出対象画像は、例えば、RGB形式の画像のような多色チャンネル画像である場合、三次元画像行列として表れることが可能である。三次元画像行列において、異なる色（例えば、R、G、B）チャンネルに一対一に対応する3つの二次元画像行列が含まれる。いずれか1つの二次元画像行列における各要素の値は、対応する色チャンネルでの、各要素に対応する画素点の画素値である。

10

【0079】

共通特徴抽出ネットワークは、少なくとも1つの畳み込み層を含む。検出対象画像の画像行列を共通特徴抽出ネットワークに入力した後、共通特徴抽出ネットワークを用いて検出対象画像の画像行列に対して畳み込み処理を行い、検出対象画像における特徴を抽出する。この場合において、抽出された特徴は、骨格特徴を含むだけでなく、輪郭特徴も含む。

【0080】

S 4 0 2において、第1骨格特徴抽出ネットワークを用いて、基礎特徴行列に対して畳み込み処理を行い、第1骨格特徴行列を得て、前記第1骨格特徴抽出ネットワークにおける第1ターゲット畳み込み層から、第2骨格特徴行列を取得し、第1骨格特徴行列及び第2骨格特徴行列に基づいて、第1ターゲット骨格特徴行列を得て、第1ターゲット畳み込み層は、第1骨格特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である。

20

【0081】

具体的な実施において、第1骨格特徴抽出ネットワークは、複数の畳み込み層を含む。複数の畳み込み層は、順に接続され、下位層の畳み込み層の入力は、上位層の畳み込み層の出力である。このような構造を有する第1骨格特徴抽出ネットワークは、基礎特徴行列に対して複数回の畳み込み処理を行い、最後の畳み込み層から、第1骨格特徴行列を得ることができる。ここで、第1骨格特徴行列は、三次元特徴行列であり、該三次元特徴行列において、複数の二次元特徴行列が含まれ、且つ、各二次元特徴行列は、事前決定された複数の骨格キーポイントに一対一に対応する。1つの骨格キーポイントに対応する二次元特徴行列における要素の値は、該要素に対応する画素点が該骨格キーポイントに属する確率を表し、1つの要素に対応する画素点は、一般的には複数である。

30

【0082】

また、複数の畳み込み層により、基礎特徴行列に対して複数回の畳み込み処理を行うことで、基礎特徴行列から、人体の骨格特徴を抽出できるが、畳み込み回数の増加に伴い、検出対象画像におけるいくつかの情報が損失される。これらの情報には、人体の骨格特徴に関わる情報が含まれる可能性がある。検出対象画像における損失された情報が多すぎると、最終的に得られた、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列は、十分に正確でないことがある。従って、本願の実施例において、また、第1骨格特徴抽出ネットワークの第1ターゲット畳み込み層から、第2骨格特徴行列を取得し、第1骨格特徴行列及び第2骨格特徴行列に基づいて、第1ターゲット骨格特徴行列を得る。

40

【0083】

ここで、第1ターゲット畳み込み層は、第1骨格特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である。図3の例において、第1骨格特徴抽出ネットワークにおける最後から二番目の畳み込み層は、第1ターゲット畳み込み層として選ばれる。

【0084】

50

例えば、下記方式で、第1骨格特徴行列及び第2骨格特徴行列に基づいて、第1ターゲット骨格特徴行列を得ることができる。

【0085】

第1骨格特徴行列及び第2骨格特徴行列に対して連結処理を行い、第1連結骨格特徴行列を得て、第1連結骨格特徴行列を次元変換処理し、第1ターゲット骨格特徴行列を得る。

【0086】

ここで、第1連結骨格特徴行列に対して次元変換処理を行う場合、それを次元変換ニューラルネットワークに入力し、該次元変換ニューラルネットワークを用いて、第1連結骨格特徴行列に対して少なくとも1回の畳み込み処理を行い、第1ターゲット骨格特徴行列を得る。

【0087】

ここで、次元変換ニューラルネットワークは、第1骨格特徴行列及び第2骨格特徴行列における特徴情報を融合し、それによって、得られた第1ターゲット骨格特徴行列に、より豊富な情報を包含させることができる。

【0088】

S403において、第1輪郭特徴抽出ネットワークを用いて、基礎特徴行列に対して畳み込み処理を行い、第1輪郭特徴行列を得て、第1輪郭特徴抽出ネットワークにおける第2ターゲット畳み込み層から、第2輪郭特徴行列を取得し、第1輪郭特徴行列及び第2輪郭特徴行列に基づいて、第1ターゲット輪郭特徴行列を得て、第2ターゲット畳み込み層は、第1輪郭特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である。図3の例において、第1輪郭特徴抽出ネットワークにおける最後から二番目の畳み込み層は、第2ターゲット畳み込み層として選ばれる。

【0089】

具体的な実施において、第1輪郭特徴抽出ネットワークも複数の畳み込み層を含む。複数の畳み込み層は、順に接続され、下位層の畳み込み層の入力は、上位層の畳み込み層の出力である。このような構造を有する第1輪郭特徴抽出ネットワークは、基礎特徴行列に対して複数回の畳み込み処理を行い、最後の畳み込み層から、第1輪郭特徴行列を得ることができる。ここで、第1輪郭特徴行列は、三次元特徴行列である。該三次元特徴行列において、複数の二次元特徴行列が含まれ、且つ、各二次元特徴行列は、事前決定された複数の輪郭キーポイントに一对一に対応する。1つの輪郭キーポイントに対応する二次元特徴行列における要素の値は、該要素に対応する画素点が該輪郭キーポイントに属する確率を表し、1つの要素に対応する画素点は、一般的には複数である。

【0090】

ここで、輪郭キーポイントの数は、一般的には、骨格キーポイントの数と異なる。従って、得られた第1輪郭特徴行列に含まれる二次元特徴行列の数は、第1骨格特徴行列に含まれる二次元特徴行列の数と異なってもよいことに留意されたい。

【0091】

例えば、骨格キーポイントの数が14個であり、輪郭キーポイントの数が25個である場合、第1輪郭特徴行列に含まれる二次元特徴行列の数は、25個であり、第1骨格特徴行列に含まれる二次元特徴行列の数は、14個である。

【0092】

また、第1ターゲット輪郭特徴行列にもより豊富な情報を包含させるために、上記S402と類似した方式で、第1輪郭特徴抽出ネットワークにおける第2ターゲット畳み込み層から第2輪郭特徴行列を取得し、続いて、第1輪郭特徴行列及び第2輪郭特徴行列に基づいて、第1ターゲット輪郭特徴行列を得ることができる。

【0093】

ここで、第1輪郭特徴行列及び第2輪郭特徴行列に基づいて、第1ターゲット輪郭特徴行列を得る方式は、例えば、

第1輪郭特徴行列及び第2輪郭特徴行列に対して連結処理を行い、第1連結輪郭特徴行列を得て、第1連結輪郭特徴行列に対して次元変換処理を行い、第1ターゲット輪郭特徴行

10

20

30

40

50

列を得ることを含む。

【0094】

上記S402及びS403において、第1ターゲット骨格特徴行列の次元は、第1ターゲット輪郭特徴行列の次元と同じであり、且つ同一の次元での、第1ターゲット骨格特徴行列と第1ターゲット輪郭特徴行列の次元数は同じであり、それによって、後続で第1ターゲット骨格特徴行列と第1ターゲット輪郭特徴行列に基づいて特徴融合処理を行うことを容易にすることに留意されたい。

【0095】

例えば、第1ターゲット骨格特徴行列の次元が3であり、且つ各次元の次元数は、それぞれ64、32及び14である場合、該第1ターゲット骨格特徴行列の次元数は、 $64 \times 32 \times 14$ として表れてもよく、第1ターゲット輪郭特徴行列の次元数も $64 \times 32 \times 14$ として表れてもよい。

10

【0096】

また、もう1つの実施例において、下記方式で、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列を得ることもできる。

【0097】

共通特徴抽出ネットワークを用いて、検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得て、

第1骨格特徴抽出ネットワークを用いて、基礎特徴行列に対して畳み込み処理を行い、第1骨格特徴行列を得て、第1骨格特徴行列に対して次元変換処理を行い、第1ターゲット骨格特徴行列を得て、

20

第1輪郭特徴抽出ネットワークを用いて、基礎特徴行列に対して畳み込み処理を行い、第1輪郭特徴行列を得て、第1輪郭特徴行列に対して次元変換処理を行い、第1ターゲット輪郭特徴行列を得る。

【0098】

該方式において、高い精度で、人体の骨格特徴及び輪郭特徴を検出対象画像から抽出することもできる。

【0099】

また、本願の実施例で提供される第1特徴抽出ネットワークは、事前訓練されたものである。

30

【0100】

ここで、本願の実施例で提供される人体検出方法は、人体検出モデルにより実現され、人体検出モデルは、第1特徴抽出ネットワーク及び／又は特徴融合ニューラルネットワークを含み、

人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報、及び人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされている。

【0101】

具体的には、人体検出モデルに第1特徴抽出ネットワークが含まれる場合について、第1特徴抽出ネットワークは、単独で訓練されてもよく、特徴融合ニューラルネットワークと共に共同訓練されてもよく、単独訓練と共同訓練を組み合わせてもよい。

40

【0102】

訓練により第1特徴抽出ネットワークを得る過程は、下記(1)及び(2)を含むが、これらに限定されない。

【0103】

(1) 第1特徴抽出ネットワークに対して単独訓練を行うことは、例えば、以下を含む。

【0104】

ステップ1. 1において、複数枚のサンプル画像、及び各サンプル画像のラベル付きデータを取得し、ラベル付きデータは、人体骨格構造を示すための骨格キーポイントの実際の位置情報、及び人体輪郭を示すための輪郭キーポイントの実際の位置情報を含む。

50

【0105】

ステップ1. 2において、複数枚のサンプル画像を第1基礎特徴抽出ネットワークに入力し、第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列を得る。

【0106】

ステップ1. 3において、第1サンプルターゲット骨格特徴行列に基づいて、骨格キーポイントの第1予測位置情報を決定し、第1サンプルターゲット輪郭特徴行列に基づいて、輪郭キーポイントの第1予測位置情報を決定する。

【0107】

ステップ1. 4において、骨格キーポイントの実際の位置情報及び骨格キーポイントの第1予測位置情報に基づいて、第1損失を決定し、輪郭キーポイントの実際の位置情報及び輪郭キーポイントの第1予測位置情報に基づいて、第2損失を決定する。

【0108】

ステップ1. 5において、第1損失及び第2損失に基づいて、第1基礎特徴抽出ネットワークに対して今回の訓練を行う。

【0109】

第1基礎特徴抽出ネットワークに対して複数回の訓練を行うことで、第1特徴抽出ネットワークを得る。

【0110】

図3に示すように、第1損失は、図3におけるLS1であり、第2損失は、図3におけるLC1である。第1損失及び第2損失に基づいて、第1基礎特徴抽出ネットワークの訓練を行い、精度の高い第1特徴抽出ネットワークを得る。

【0111】

(2) 第1特徴抽出ネットワーク及び特徴融合ニューラルネットワークを共同訓練することは、例えば、以下を含む。

【0112】

ステップ2. 1において、複数枚のサンプル画像及び各サンプル画像のラベル付きデータを得て、ラベル付きデータは、人体骨格構造を示すための骨格キーポイントの実際の位置情報及び人体輪郭を示すための輪郭キーポイントの実際の位置情報を含む。

【0113】

ステップ2. 2において、複数枚のサンプル画像を第1基礎特徴抽出ネットワークに入力し、第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列を得る。

【0114】

ステップ2. 3において、基礎特徴融合ニューラルネットワークを用いて、第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列に対して特徴融合を行い、第2サンプルターゲット骨格特徴行列及び第2サンプルターゲット輪郭特徴行列を得る。

【0115】

ステップ2. 4において、第2サンプルターゲット骨格特徴行列に基づいて、骨格キーポイントの第2予測位置情報を決定し、第2サンプルターゲット輪郭特徴行列に基づいて、輪郭キーポイントの第2予測位置情報を決定する。

【0116】

ステップ2. 5において、骨格キーポイントの実際の位置情報及び骨格キーポイントの第2予測位置情報に基づいて、第3損失を決定し、輪郭キーポイントの実際の位置情報及び輪郭キーポイントの第2予測位置情報に基づいて、第4損失を決定する。

【0117】

ステップ2. 6において、第3損失及び第4損失に基づいて、第1基礎特徴抽出ネットワーク及び基礎特徴融合ニューラルネットワークに対して、今回の訓練を行う。

【0118】

10

20

30

40

50

第1基礎量み込みニューラルネットワーク及び基礎特徴融合ニューラルネットワークに対して複数回の訓練を行うことで、第1特徴抽出ネットワーク及び特徴融合ニューラルネットワークを得る。

【0119】

(3) 単独訓練と共同訓練を組み合わせて第1特徴抽出ネットワークを得る過程において、上記(1)及び(2)におけるプロセスにより同期して訓練を行うことができる。

【0120】

又は、まず、(1)におけるプロセスにより、第1特徴抽出ネットワークを事前訓練し、事前訓練された第1特徴抽出ネットワークと特徴融合ニューラルネットワークに対して、上記(2)における共同訓練を行うこともできる。

【0121】

第1特徴抽出ネットワークに対して単独訓練と共同訓練を行う場合に用いられる画像は、同じであっても異なってもよいことに留意されたい。

【0122】

第1特徴抽出ネットワーク及び特徴融合ニューラルネットワークを共同訓練する前に、まず、特徴融合ニューラルネットワークを事前訓練し、事前訓練された特徴融合ニューラルネットワークを第1特徴抽出ネットワークと共に共同訓練することもできる。

【0123】

特徴融合ニューラルネットワークに対して単独訓練を行う詳細な過程は、下記a2に示す実施例に関する説明を参照することができる。

【0124】

a2：特徴融合過程：

人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列及び人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を得た後、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列に基づいて、特徴融合処理を行うことができる。

【0125】

具体的には、検出対象画像に基づいて、骨格特徴及び輪郭特徴を抽出する過程において、用いられる基礎行列が同一であるが、第1骨格特徴抽出ネットワークは、基礎特徴行列から骨格特徴を抽出し、それに対して、第1輪郭特徴抽出ネットワークは、基礎特徴行列から輪郭特徴を抽出する。2つの過程は相互独立して存在する。しかしながら、同一の人体は、輪郭特徴と骨格特徴が、相互関連付けられている。輪郭特徴と骨格特徴を融合させる目的は、骨格特徴と輪郭特徴との相互作用関係を利用することである。例えば、輪郭特徴に基づいて、最終的に抽出された骨格キーポイントの位置情報を修正し、骨格特徴に基づいて、最終的に抽出された輪郭キーポイントの位置情報を修正することで、更に、より正確な骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を得て、精度のより高い人体検出結果を得る。

【0126】

本願の実施例は、抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行うための具体的な方法を提供する。前記方法は、事前訓練された特徴融合ニューラルネットワークを用いて、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得ることを含む。

【0127】

ここで、第2ターゲット骨格特徴行列は、三次元骨格特徴行列であり、該三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、二次元骨格特徴行列における各要素の値は、該要素に対応する画素点に対応する骨格キーポイント（即ち、該二次元骨格特徴行列に対応する骨格キーポイント）に属する確率を表す。第2ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、該三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、二次元輪郭特徴行列にお

10

20

30

40

50

る各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表す。

【0128】

本願の実施例で提供される特徴融合ニューラルネットワークは、単独訓練されてもよく、第1特徴抽出ネットワークと共に共同訓練されてもよく、単独訓練と共同訓練を組み合わせてもよい。

【0129】

特徴融合ニューラルネットワークと第1特徴抽出ネットワークを共同訓練する過程は、上記(2)を参照することができる。ここで、詳細な説明を省略する。

【0130】

構造の異なる特徴融合ニューラルネットワークに対して単独訓練を行う場合、用いられる訓練方法も異なる。構造の異なる特徴融合ニューラルネットワークの訓練方法は、下記M1～M3を参照することができる。

【0131】

骨格特徴及び輪郭特徴に対して特徴融合を行う過程は、下記M1～M3のうちの少なくとも1つを含んでもよく、これらに限定されない。

【0132】

M1:

図5に示すように、本願の実施例は、特徴融合ニューラルネットワークの具体的な構造を提供する。該構造は、第1畳み込みニューラルネットワーク、第2畳み込みニューラルネットワーク、第1変換ニューラルネットワーク、及び第2変換ニューラルネットワークを含む。

【0133】

図6に示すように、本願の実施例は、図5で提供される特徴融合ニューラルネットワークに基づいて、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るための具体的な方法を更に提供する。前記方法は、下記ステップを含む。

【0134】

S601において、第1畳み込みニューラルネットワークを用いて、第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得る。S603を実行する。

【0135】

ここで、第1畳み込みニューラルネットワークは、少なくとも1つの畳み込み層を含む。第1畳み込みニューラルネットワークは、複数層を含む場合、複数の畳み込み層は、順次接続される。該層の畳み込み層の入力は、上位層の畳み込み層の出力である。第1ターゲット骨格特徴行列を第1畳み込みニューラルネットワークに入力し、各畳み込み層を用いて、第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得る。

【0136】

該過程は、骨格特徴を第1ターゲット骨格特徴行列から更に抽出するためのものである。

【0137】

S602において、第2畳み込みニューラルネットワークを用いて、第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得る。S604を実行する。

【0138】

ここで、該処理過程は、上記S601と同様であり、ここで、詳細な説明を省略する。

【0139】

S601とS602の実行には、順序がないことに留意されたい。同期して実行してもよく、非同期で実行してもよい。

【0140】

10

20

30

40

50

S 6 0 3において、第1中間輪郭特徴行列と第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、第1変換ニューラルネットワークを用いて、第1連結特徴行列に対して次元変換を行い、第2ターゲット骨格特徴行列を得る。

【0141】

ここで、第1中間輪郭特徴行列と第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得る。得られた第1連結特徴行列に、輪郭特徴が含まれるだけでなく、骨格特徴も含まれる。

【0142】

第1変換ニューラルネットワークを用いて、第1連結行列に対して更に次元変換を行うことは、実際に、第1変換ニューラルネットワークを用いて、再び、第1連結特徴行列から骨格特徴を抽出することである。第1連結特徴行列を得る過程において、検出対象画像における骨格特徴及び輪郭特徴以外の他の特徴が除去されており、骨格特徴及び輪郭特徴のみが含まれるため、第1連結特徴行列に基づいて得られた第2ターゲット骨格特徴行列に含まれる骨格特徴は、輪郭特徴による影響を受けており、それによって骨格特徴と輪郭特徴を関連付けることができ、骨格特徴と輪郭特徴の融合を実現させることができる。

【0143】

S 6 0 4において、第1中間骨格特徴行列と第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、第2変換ニューラルネットワークを用いて、第2連結特徴行列に対して次元変換を行い、第2ターゲット輪郭特徴行列を得る。

【0144】

ここで、第1中間骨格特徴行列と第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得る過程は、上記S 6 0 2における第1連結特徴行列を得る過程と同様であり、ここで、詳細な説明を省略する。

【0145】

同様に、第2ターゲット輪郭特徴行列に含まれる輪郭特徴は、骨格特徴による影響を受けており、それによって骨格特徴と輪郭特徴を関連付け、骨格特徴と輪郭特徴の融合を実現させる。

【0146】

もう1つの実施例において、下記方式で、特徴融合ニューラルネットワークを単独訓練することができる。

【0147】

ステップ3. 1において、複数枚のサンプル画像の第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列を取得する。

【0148】

取得方式は、上記実施例における第1ターゲット骨格特徴行列、第1ターゲット輪郭特徴行列を取得する方式と同様であり、ここで、詳細な説明を省略する。第1特徴抽出ネットワークと共に共同訓練を行う場合に取得してもよく、事前訓練された第1特徴抽出ネットワークにより取得してもよい。

【0149】

ステップ3. 2において、第1基礎量み込みニューラルネットワークを用いて、第1サンプルターゲット骨格特徴行列に対して量み込み処理を行い、第1サンプル中間骨格特徴行列を得る。

【0150】

ステップ3. 3において、第2基礎量み込みニューラルネットワークを用いて、第1サンプルターゲット輪郭特徴行列に対して量み込み処理を行い、第1サンプル中間輪郭特徴行列を得る。

【0151】

ステップ3. 4において、第1サンプル中間輪郭特徴行列と第1サンプルターゲット骨格特徴行列に対して連結処理を行い、第1サンプル連結特徴行列を得て、第1基礎変換ニューラルネットワークを用いて、第1サンプル連結特徴行列に対して次元変換を行い、第2

10

20

30

40

50

サンプルターゲット骨格特徴行列を得る。

【0152】

ステップ3. 5において、第1サンプル中間骨格特徴行列と第1サンプルターゲット輪郭特徴行列に対して連結処理を行い、第2サンプル連結特徴行列を得て、第2基礎変換ニューラルネットワークを用いて、第2サンプル連結特徴行列に対して、次元変換を行い、第2サンプルターゲット輪郭特徴行列を得る。

【0153】

ステップ3. 6において、第2サンプルターゲット骨格特徴行列に基づいて、骨格キーポイントの第3予測位置情報を決定し、第2サンプルターゲット輪郭特徴行列に基づいて、輪郭キーポイントの第3予測位置情報を決定する。

【0154】

ステップ3. 7において、骨格キーポイントの実際の位置情報及び骨格キーポイントの第3予測位置情報に基づいて、第5損失を決定し、輪郭キーポイントの実際の位置情報及び輪郭キーポイントの第3予測位置情報に基づいて、第6損失を決定する。

【0155】

ステップ3. 8において、第5損失及び第6損失に基づいて、第1基礎畳み込みニューラルネットワーク、第2基礎畳み込みニューラルネットワーク、第1基礎変換ニューラルネットワーク、及び第2基礎変換ニューラルネットワークに対して今回の訓練を行う。

【0156】

第1基礎畳み込みニューラルネットワーク、第2基礎畳み込みニューラルネットワーク、第1基礎変換ニューラルネットワーク、及び第2基礎変換ニューラルネットワークに対して複数回の訓練を行うことで、特徴融合ニューラルネットワークを得る。

【0157】

ここで、第5損失は、図5におけるLS2であり、第6損失は、図5におけるLC2である。

【0158】

M2:

図7に示すように、本願の実施例で提供されるもう1つの特徴融合ニューラルネットワークの具体的な構造は、第1定向畳み込みニューラルネットワーク、第2定向畳み込みニューラルネットワーク、第3畳み込みニューラルネットワーク、第4畳み込みニューラルネットワーク、第3変換ニューラルネットワーク、及び第4変換ニューラルネットワークを含む。

【0159】

図8に示すように、本願の実施例は、図7で提供される特徴融合ニューラルネットワークに基づいて、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るための具体的な方法を更に提供する。前記方法は、下記ステップを含む。

【0160】

S801において、第1定向畳み込みニューラルネットワークを用いて、第1ターゲット骨格特徴行列に対して定向畳み込み処理を行い、第1定向骨格特徴行列を得る。第3畳み込みニューラルネットワークを用いて、第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得る。S804を実行する。

【0161】

S802において、第2定向畳み込みニューラルネットワークを用いて、第1ターゲット輪郭特徴行列に対して定向畳み込み処理を行い、第1定向輪郭特徴行列を得る。第4畳み込みニューラルネットワークを用いて、第1定向輪郭特徴行列に対して畳み込み処理を行い、第2中間輪郭特徴行列を得る。S803を実行する。

【0162】

S803において、第2中間輪郭特徴行列と第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、第3変換ニューラルネットワークを用いて、第3連

10

20

30

40

50

結特徴行列に対して次元変換を行い、第2ターゲット骨格特徴行列を得る。

【0163】

S804において、第2中間骨格特徴行列と第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得て、第4変換ニューラルネットワークを用いて、第4連結特徴行列に対して次元変換を行い、第2ターゲット輪郭特徴行列を得る。

【0164】

具体的に実施する場合、骨格特徴と輪郭特徴に対して特徴融合を行う過程において、骨格キーポイントは、一般的に、人体の骨格に集中しており、輪郭キーポイントは、人体の輪郭に集中しており、つまり、骨格の周囲に分布する。従って、骨格特徴及び輪郭特徴に対してそれぞれ局所的な空間変換を行う必要がある。例えば、骨格特徴を輪郭特徴の輪郭特徴行列における位置に変換し、輪郭特徴を骨格特徴の骨格特徴行列における位置に変換することで、骨格特徴及び輪郭特徴をより好適に抽出し、骨格特徴及び輪郭特徴の融合を実現させる。

【0165】

該目的を実現させるために、本願の実施例において、まず、第1定向畳み込みニューラルネットワークを用いて、第1ターゲット骨格特徴行列に対して定向畳み込み処理を行う。該定向畳み込みは、特徴面において骨格特徴の定向空間変換を効果的に実現させることができる。続いて、第3畳み込みニューラルネットワークを用いて、得られた第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得る。この場合、第1定向畳み込み層により骨格特徴に対して定向空間変換を行ったため、骨格特徴は、実際に、輪郭特徴方向へ移動した。続いて、第2中間骨格特徴行列と第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得る。第4連結特徴行列は、輪郭特徴を含むと共に、定向空間変換が行われた骨格特徴を更に含む。続いて、第4変換ニューラルネットワークを用いて、第4連結特徴行列に対して次元変換を行い、つまり、第4連結特徴行列から、再び輪郭特徴を抽出する。このような方式で得られた第2ターゲット輪郭特徴行列は、骨格特徴による影響を受けており、骨格特徴と輪郭特徴との融合を実現させる。

【0166】

同様に、本願の実施例において、まず、第2定向畳み込みニューラルネットワークを用いて、第1ターゲット輪郭特徴行列に対して定向畳み込み処理を行う。該定向畳み込みは、特徴面において輪郭特徴の定向空間変換を効果的に実現させることができる。続いて、第4畳み込みニューラルネットワークを用いて、得られた第1定向輪郭特徴行列に対して畳み込み処理を行い、第2中間輪郭特徴行列を得る。この場合、第2定向畳み込み層により輪郭特徴に対して定向空間変換を行ったため、輪郭特徴は、実際に、骨格特徴方向へ移動した。続いて、第2中間輪郭特徴行列と第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得る。第3連結特徴行列は、骨格特徴を含むと共に、定向空間変換が行われた輪郭特徴を更に含む。続いて、第3変換ニューラルネットワークを用いて、第3連結特徴行列に対して次元変換を行い、つまり、第3連結特徴行列から、再び骨格特徴を抽出する。このような方式で得られた第2ターゲット骨格特徴行列は、輪郭特徴による影響を受けており、骨格特徴と輪郭特徴との融合を実現させる。

【0167】

具体的には、定向畳み込みは、複数回の反復畳み込みステップからなる。効果的な定向畳み込みは下記要求を満たす。

【0168】

(1) 毎回の反復畳み込みのステップにおいて、特徴行列における一組の要素の要素値のみを更新する。

【0169】

(2) 最終回の反復畳み込みのステップを行った後、全ての要素の要素値は、一回のみ更新されるべきである。

【0170】

第1ターゲット骨格特徴行列に対して定向畳み込みを行うことを例として、定向畳み込み

10

20

30

40

50

過程を実現させるために、要素の更新順序を制御するための特徴関数シーケンス

【0171】

【化1】

$$F = \{F_k\}_{k=1}^K$$

を定義することができる。ここで、関数 F_k の入力は、第1ターゲット骨格特徴行列における各要素の位置であり、関数 F_k の出力は、 k 回目の反復における要素を更新するかどうかを表す。該出力は、1又は0であってもよい。1は、更新を表し、0は、更新しないことを表す。具体的には、 k 回目の反復過程において、 $F_k = 1$ の領域における要素の要素値のみを更新し、他の領域における要素の要素値をそのまま維持する。 i 回目の反復の更新は、

10

【0172】

【化2】

$$T_i(X) = F_i \cdot (W \times T_{i-1}(X) + b) + (1 - F_i) \cdot T_{i-1}(X)$$

で表されてもよい。

【0173】

ここで、 $T_0(X) = X$ であり、 X は、定向畳み込みの入力を表し、即ち、第1ターゲット骨格特徴行列である。 W 及び b はそれぞれ複数回の反復過程における共通重み及び偏差を表す。

20

【0174】

骨格特徴と輪郭特徴の融合を実現させるために、一対の対称な定向畳み込み演算子を設定してもよい。つまり、上記特徴関数シーケンス

【0175】

【化3】

$$F = \{F_k\}_{k=1}^K$$

30

は、それぞれ、散乱畳み込み演算子

【0176】

【化4】

$$F_i^S$$

及び集約畳み込み演算子

【0177】

【化5】

$$F_i^G$$

40

である。ここで、散乱畳み込み演算子は、内から外へ特徴行列における要素を順次更新する。集約畳み込み演算子は、外から内へ特徴行列における要素を順次更新する。

【0178】

第1定向畳み込みニューラルネットワークを用いて、第1ターゲット骨格特徴行列に対して定向畳み込み処理を行う場合、骨格特徴元素を該要素周囲の位置（輪郭特徴に更に関わる位置）に定向空間変換する必要があるため、散乱畳み込み演算子

【0179】

50

【化 6】

$$F_i^S$$

を用いる。第 2 定向畳み込みニューラルネットワークを用いて、第 1 ターゲット輪郭特徴行列に対して定向畳み込み処理を行う場合、輪郭特徴元素を輪郭特徴行列の中間位置（骨格特徴に更に関わる位置）に定向空間変換する必要があるため、集約畳み込み演算子

【0 1 8 0】

【化 7】

$$F_i^G$$

10

を用いる。

【0 1 8 1】

具体的には、第 1 定向畳み込みニューラルネットワークにより、第 1 ターゲット骨格特徴行列に対して定向畳み込み処理を行う過程は以下のとおりである。

【0 1 8 2】

第 1 ターゲット骨格特徴行列を複数のサブ行列に分ける。各サブ行列は、1 つのグリッドと呼ばれる。ここで、第 1 ターゲット骨格特徴が三次元行列である場合、3 つの次元の次元数は、それぞれ m 、 n 、 s である。従って、第 1 ターゲット骨格特徴行列の次元数は、 $m * n * s$ で表される。グリッドのサイズが 5 である場合、各グリッドの次元数はいずれも $5 * 5 * s$ で表される。

20

【0 1 8 3】

続いて、各グリッドに対して、散乱畳み込み演算子

【0 1 8 4】

【化 8】

$$F_i^S$$

を用いて複数回の反復畳み込みを行い、ターゲットサブ行列を得る。図 9 a に示すように、散乱畳み込み演算子

【0 1 8 5】

30

【化 9】

$$F_i^S$$

を用いて、グリッドサイズが 5 であるサブ行列における要素の要素値に対して 2 回の反復更新を行う過程を提供する。ここで、図 9 a における a は、初期サブ行列を表し、b は、一回の反復を行った後に得られたサブ行列を表し、c は、二回の反復を行った後に得られたサブ行列を表し、つまり、ターゲットサブ行列を表す。

【0 1 8 6】

各グリッドに対応するターゲットサブ行列を連結し、第 1 定向骨格特徴行列を得る。

40

【0 1 8 7】

同様に、第 2 定向畳み込みニューラルネットワークにより、第 1 ターゲット輪郭特徴行列に対して定向畳み込み処理を行う過程は以下のとおりである。

【0 1 8 8】

第 1 ターゲット輪郭特徴行列を複数のサブ行列に分ける。各サブ行列は、1 つのグリッドと呼ばれる。ここで、第 1 ターゲット輪郭特徴行列は三次元行列である場合、3 つの次元の次元数は、それぞれ m 、 n 、 s である。その場合、第 1 ターゲット輪郭特徴行列の次元数は、 $m * n * s$ で表される。グリッドのサイズが 5 である場合、各グリッドの次元数はいずれも $5 * 5 * s$ で表される。

【0 1 8 9】

50

続いて、各グリッドに対して、集約畳み込み演算子

【0 1 9 0】

【化1 0】

$$F_i^G$$

を用いて複数回の反復畳み込みを行い、ターゲットサブ行列を得る。

【0 1 9 1】

図9 bに示すように、集約畳み込み演算子

【0 1 9 2】

【化1 1】

$$F_i^G$$

10

を用いて、グリッドサイズが5であるサブ行列における要素の要素値に対して2回の反復更新を行う過程を提供する。ここで、図9 bにおけるaは、初期サブ行列を表し、bは、一回の反復を行った後に得られたサブ行列を表し、cは、二回の反復を行った後に得られたサブ行列を表し、つまり、ターゲットサブ行列を表す。

【0 1 9 3】

各グリッドに対応するターゲットサブ行列を連結し、第1定向輪郭特徴行列を得る。

【0 1 9 4】

ここで、各サブ行列の反復畳み込みを並行して実行することができることに留意されたい。

【0 1 9 5】

図9 a及び図9 bにおける例は、散乱畳み込み演算子

【0 1 9 6】

【化1 2】

$$F_i^S$$

20

及び集約畳み込み演算子

【0 1 9 7】

【化1 3】

$$F_i^G$$

30

を用いてサブ行列における要素の要素値を反復更新する例だけである。

【0 1 9 8】

もう1つの実施例において、下記方式で、特徴融合ニューラルネットワークを単独訓練することができる。

【0 1 9 9】

ステップ4. 1において、複数枚のサンプル画像の第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列を取得する。

【0 2 0 0】

取得方式は、上記実施例における第1ターゲット骨格特徴行列、第1ターゲット輪郭特徴行列の取得方式と同様であり、ここで詳細な説明を省略する。第1特徴抽出ネットワークと共に共同訓練する場合に取得してもよく、事前訓練された第1特徴抽出ネットワークにより取得してもよい。

【0 2 0 1】

ステップ4. 2において、第1基礎定向畳み込みニューラルネットワークを用いて、第1サンプルターゲット骨格特徴行列に対して定向畳み込み処理を行い、第1サンプル定向骨

40

50

格特徴行列を得て、第1サンプル定向骨格特徴行列及び輪郭キーポイントの実際の位置情報を用いて、第7損失を得る。第7損失に基づいて、第1基礎定向畳み込みニューラルネットワークに対して今回の訓練を行う。

【0202】

ここで、第7損失は、図7におけるLC3である。

【0203】

ここで、第1基礎定向畳み込みニューラルネットワークを用いて、第1サンプルターゲット骨格特徴行列に対して定向畳み込み処理を行う。つまり、第1サンプルターゲット骨格特徴行列に対して定向空間変換を行う。この場合、得られた第1サンプル定向骨格特徴行列で表されるキーポイントの位置情報を輪郭キーポイントの位置情報と可能な限り一致させる必要がある。従って、第1サンプル定向骨格特徴行列及び輪郭キーポイントの実際の位置情報に基づいて、第7損失を得る。第7損失を用いて、第1基礎定向畳み込みニューラルネットワークに対する訓練を行う。

10

【0204】

ステップ4.3において、第2基礎定向畳み込みニューラルネットワークを用いて、第1サンプルターゲット輪郭特徴行列に対して定向畳み込み処理を行い、第1サンプル定向輪郭特徴行列を得て、第1サンプル定向輪郭特徴行列及び骨格キーポイントの実際の位置情報を用いて、第8損失を得る。第8損失に基づいて、第2基礎定向畳み込みニューラルネットワークに対して今回の訓練を行う。

【0205】

20

ここで、第8損失は、図7におけるLS3である。

【0206】

ステップ4.4において、第4基礎畳み込みニューラルネットワークを用いて、第1サンプル定向輪郭特徴行列に対して畳み込み処理を行い、第2サンプル中間輪郭特徴行列を得て、得られた第2サンプル中間輪郭特徴行列と第1サンプルターゲット骨格特徴行列に対して連結処理を行い、第3サンプル連結特徴行列を得て、第3基礎変換ニューラルネットワークを用いて、第3サンプル連結特徴行列に対して次元変換を行い、第2サンプルターゲット骨格特徴行列を得る。

【0207】

ステップ4.5において、第2サンプルターゲット骨格特徴行列に基づいて、骨格キーポイントの第4予測位置情報を決定し、骨格キーポイントの実際の位置情報及び骨格キーポイントの第4予測位置情報に基づいて、第9損失を決定する。

30

【0208】

ここで、第9損失は、図7におけるLS4である。

【0209】

ステップ4.6において、第3基礎畳み込みニューラルネットワークを用いて、第1サンプル定向骨格特徴行列に対して畳み込み処理を行い、第2サンプル中間骨格特徴行列を得て、得られた第2サンプル中間骨格特徴行列と第1サンプルターゲット輪郭特徴行列に対して連結処理を行い、第4サンプル連結特徴行列を得て、第4基礎変換ニューラルネットワークを用いて、第4サンプル連結特徴行列に対して次元変換を行い、第2サンプルターゲット輪郭特徴行列を得る。

40

【0210】

ステップ4.7において、第2サンプルターゲット輪郭特徴行列に基づいて、輪郭キーポイントの第4予測位置情報を決定し、輪郭キーポイントの実際の位置情報及び輪郭キーポイントの第4予測位置情報に基づいて、第10損失を決定する。

【0211】

ここで、第10損失は、図7におけるLC4である。

【0212】

ステップ4.8において、第9損失及び第10損失に基づいて、第3基礎畳み込みニューラルネットワーク、第4基礎畳み込みニューラルネットワーク、第3基礎変換ニューラル

50

ネットワーク、及び第4基礎変換ニューラルネットワークに対して今回の訓練を行う。

【0213】

第1基礎定向量み込みニューラルネットワーク、第2基礎定向量み込みニューラルネットワーク、第3基礎量み込みニューラルネットワーク、第4基礎量み込みニューラルネットワーク、第3基礎変換ニューラルネットワーク、及び第4基礎変換ニューラルネットワークに対して複数回の訓練を行うことで、訓練された特徴融合ニューラルネットワークを得る。

【0214】

M3:

図10に示すように、本願の実施例で提供されるもう1つの特徴融合ニューラルネットワークの具体的な構造は、変位推定ニューラルネットワーク、第5変換ニューラルネットワークを含む。

10

図11に示すように、本願の実施例は、図10で提供される特徴融合ニューラルネットワークに基づいて、第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るための具体的な方法を更に提供する。前記方法は、下記ステップを含む。

【0215】

S1101において、第1ターゲット骨格特徴行列と第1ターゲット輪郭特徴行列に対して連結処理を行い、第5連結特徴行列を得る。

【0216】

20

S1102において、第5連結特徴行列を変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する変位情報を得て、ここで、各キーポイントペアのうちの2つのキーポイントの位置は、隣接し、該2つのキーポイントは、1つの骨格キーポイント及び1つの輪郭キーポイントを含むか、又は2つの骨格キーポイントを含むか、又は2つの輪郭キーポイントを含む。

【0217】

具体的な実施において、人体のために、複数の骨格キーポイント及び複数の輪郭キーポイントを事前決定する。図12に示すように、人体のために事前決定された複数の骨格キーポイント及び輪郭キーポイントの例を提供する。該例において、骨格キーポイントは、14個であり、図12における大きなドットによりそれぞれ、頭頂、頸、両肩、両肘、両手首、両股、両膝、及び両足首を表す。輪郭キーポイントは、26個であり、図12における小さなドットで表される。人体の頭頂を示す骨格キーポイント以外、他の各骨格キーポイントは、2つの輪郭キーポイントに対応する。ここで、両股の骨格キーポイントは、同一の輪郭キーポイントに対応する。

30

【0218】

位置が隣接する2つのキーポイントは、1つのキーポイントペアを構成する。図12に示すように、線分により直接的に接続した2つのキーポイントは、1つのキーポイントペアを構成する。つまり、キーポイントペアは、(骨格キーポイント、骨格キーポイント)、(輪郭キーポイント、輪郭キーポイント)、又は(骨格キーポイント、輪郭キーポイント)から構成される可能性がある。

40

【0219】

変位推定ニューラルネットワークは、複数の量み込み層を含む。複数の量み込み層は、順次接続され、第5連結特徴行列における骨格特徴及び輪郭特徴に対して特徴学習を行い、各キーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する変位情報を得るために用いられる。各キーポイントに対応する変位情報は2組である。

【0220】

例えば、キーポイントペアが(P、Q)であり、P及びQはそれぞれ1つのキーポイントを表すと、該キーポイントペアの変位情報は、PからQまでに移動する変位情報、及びQからPまでに移動する変位情報を含む。

50

【0221】

各組の変位情報は、いずれも移動方向及び移動距離を含む。

【0222】

S1103において、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとして、現在のキーポイントのペアであるもう1つのキーポイントに対応する三次元特徴行列から、ペアであるもう1つのキーポイントに対応する二次元特徴行列を取得し、ペアであるもう1つのキーポイントは、骨格キーポイントである場合、該骨格キーポイントに対応する三次元特徴行列は、第1骨格特徴行列であり、ペアであるもう1つのキーポイントは、輪郭キーポイントである場合、該輪郭キーポイントに対応する三次元特徴行列は、第1輪郭特徴行列である。

10

【0223】

S1104において、ペアであるもう1つのキーポイントから現在のキーポイントまでの変位情報に基づいて、ペアであるもう1つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応する変位特徴行列を得る。

【0224】

ここで、依然としてキーポイントペア(P、Q)を例として、まず、Pを現在のキーポイントとして、Qに対応する三次元特徴行列から、Qに対応する二次元特徴行列を取得する。

【0225】

ここで、Qが骨格キーポイントである場合、Qに対応する三次元特徴行列は、第1骨格特徴行列(上記S402を参照する)である。Qが輪郭キーポイントである場合、Qに対応する三次元特徴行列は、第1輪郭特徴行列(上記S403を参照する)である。

20

【0226】

ここで、Qが骨格キーポイントである場合、第1骨格特徴行列をQの三次元特徴行列として、第1骨格特徴行列から、Qの二次元特徴行列を得る。第1骨格特徴行列に骨格特徴のみが含まれるため、後続の処理過程で学習された骨格特徴に、更なる指向性を持たせる。同様に、Qが輪郭キーポイントである場合、第1輪郭特徴行列をQの三次元特徴行列として、第1輪郭特徴行列から、Qの二次元特徴行列を得る。第1輪郭特徴行列に輪郭特徴のみが含まれるため、後続の処理過程で学習された輪郭特徴に、更なる指向性を持たせる。

【0227】

Qの二次元特徴行列を得た後、QからPまでに移動する変位情報に基づいて、Qの二次元特徴行列における要素に対して位置変換を行い、Pに対応する変位特徴行列を得る。

30

【0228】

例えば、図13に示すように、QからPまでに移動する変位情報は、(2, 3)である。ここで、2は、第1次元で移動した距離が2であることを表す。3は、第2次元で移動した距離が3であることを表す。従って、Qの二次元特徴行列は、図13におけるaに示すとおりである。Qの二次元特徴行列における要素に対して位置変換を行った後に得られたPに対応する変位特徴行列は、図13におけるbに示すとおりである。ここで、数字により変位情報を相対的に表す。実際の実施において、具体的な解決手段を参照しながら、変位情報を理解すべきである。例えば、変位情報「2」は、2つの要素、2つのセルなどを指してもよい。

40

【0229】

続いて、Qを現在のキーポイントとして、Pに対応する三次元特徴行列から、Pに対応する二次元特徴行列を取得する。続いて、PからQまでに移動する変位情報に基づいて、Pの二次元特徴行列における要素に対して位置変換を行い、Qに対応する変位特徴行列Qを得る。

【0230】

これにより、各骨格キーポイントに対応する変位特徴行列及び各輪郭キーポイントに対応する変位特徴行列を得ることができる。

【0231】

ここで、各骨格キーポイントは、複数のキーポイントとそれぞれペアリングされる可能性

50

があるため、得られた各骨格キーポイントの変位特徴行列も複数である可能性があり、各輪郭キーポイントは、複数のキーポイントとそれぞれペアリングされる可能性があるため、得られた各輪郭キーポイントの変位特徴行列も複数である可能性があることに留意されたい。また、異なる輪郭キーポイントに対応する変位特徴行列の数も異なる可能性があり、異なる骨格キーポイントに対応する変位特徴行列の数も異なる可能性がある。

【0232】

S1105において、各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と該骨格キーポイントに対応する各変位特徴行列に対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポイントの連結二次元特徴行列を第5変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、第2ターゲット骨格特徴行列を生成する。

10

【0233】

S1106において、各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と該輪郭キーポイントに対応する各変位特徴行列に対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を第5変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、第2ターゲット輪郭特徴行列を生成する。

【0234】

例えば、Pが骨格キーポイントであり、且つPに対応する二次元特徴行列がP'であり、Pが3つのキーポイントペアに位置すると、上記過程により、Pの3つの変位特徴行列を得て、それぞれP1'、P2'、及びP3'である。P'、P1'、P2'及びP3'を連結し、Pの連結二次元特徴行列を得る。この場合、Pの3つの変位特徴行列のうち、骨格キーポイントに対応する二次元特徴行列における要素を位置変換することで得られた変位特徴行列があるだけでなく、輪郭キーポイントに対応する二次元特徴行列における要素を位置変換することで得られた変位特徴行列もある。従って、P'、P1'、P2'及びP3'を連結し、位置がPの位置に隣接する各キーポイントの特徴を融合させる。また、第5変換ニューラルネットワークを用いて、Pの連結二次元特徴行列に対して畳み込み処理を行うことで、得られたPのターゲット二次元特徴行列に、骨格特徴を包含させるだけでなく、輪郭特徴も包含させ、骨格特徴と輪郭特徴の融合を実現させる。

20

30

【0235】

同様に、Pが輪郭キーポイントである場合にも、上記過程により、骨格特徴と輪郭特徴の融合を実現させることもできる。

【0236】

もう1つの実施例において、下記方式で特徴融合ニューラルネットワークを単独訓練することができる。

【0237】

ステップ5.1において、複数枚のサンプル画像の第1サンプルターゲット骨格特徴行列及び第1サンプルターゲット輪郭特徴行列を取得する。

40

【0238】

取得方式は、上記実施例における第1ターゲット骨格特徴行列、第1ターゲット輪郭特徴行列を取得する方式と同様であり、ここで、詳細な説明を省略する。第1特徴抽出ネットワークと共に共同訓練を行う場合に取得してもよく、事前訓練された第1特徴抽出ネットワークにより取得してもよい。

【0239】

ステップ5.2において、第1サンプルターゲット骨格特徴行列と第1サンプルターゲット輪郭特徴行列に対して連結処理を行い、第5サンプル連結特徴行列を得る。

【0240】

ステップ5.3において、第5サンプル連結特徴行列を基礎変位推定ニューラルネットワ

50

ークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイントがもう1つのキーポイントまでに移動する予測変位情報を得て、各キーポイントペアのうちの2つのキーポイントの位置は隣接し、該2つのキーポイントは、1つの骨格キーポイント及び1つの輪郭キーポイントを含むか、又は2つの骨格キーポイントを含むか、又は2つの輪郭キーポイントを含む。

【0241】

ステップ5.4において、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとして、該現在のキーポイントのペアであるもう1つのキーポイントに対応するサンプル三次元特徴行列から、ペアであるもう1つのキーポイントに対応するサンプル二次元特徴行列を取得する。

【0242】

ステップ5.5において、ペアであるもう1つのキーポイントから現在のキーポイントまでの予測変位情報に基づいて、ペアであるもう1つのキーポイントに対応するサンプル二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応するサンプル変位特徴行列を得る。

【0243】

ステップ5.6において、現在のキーポイントに対応するサンプル変位特徴行列及び現在のキーポイントに対応するサンプル二次元特徴行列に基づいて、変位損失を決定する。

【0244】

ステップ5.7において、変位損失に基づいて、変位推定ニューラルネットワークに対して今回の訓練を行う。

【0245】

ステップ5.8において、各骨格キーポイントについて、該骨格キーポイントに対応するサンプル二次元特徴行列と該骨格キーポイントに対応する各サンプル変位特徴行列に対して連結処理を行い、該骨格キーポイントのサンプル連結二次元特徴行列を得て、該骨格キーポイントのサンプル連結二次元特徴行列を第5基礎変換ニューラルネットワークに入力し、該骨格キーポイントに対応するサンプルターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するサンプルターゲット二次元特徴行列に基づいて、第2サンプルターゲット骨格特徴行列を生成する。

【0246】

ステップ5.9において、各輪郭キーポイントについて、該輪郭キーポイントに対応するサンプル二次元特徴行列と該輪郭キーポイントに対応する各サンプル変位特徴行列に対して連結処理を行い、該輪郭キーポイントのサンプル連結二次元特徴行列を得て、該輪郭キーポイントのサンプル連結二次元特徴行列を第5基礎変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するサンプルターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するサンプルターゲット二次元特徴行列に基づいて、第2サンプルターゲット輪郭特徴行列を生成する。

【0247】

ステップ5.10において、第2サンプルターゲット骨格特徴行列、第2サンプルターゲット輪郭特徴行列、骨格キーポイントの実際の位置情報及び輪郭キーポイントの実際の位置情報に基づいて、変換損失を決定する。例えば、第2サンプルターゲット骨格特徴行列に基づいて、骨格キーポイントの予測位置情報を決定し、第2サンプルターゲット輪郭特徴行列に基づいて、輪郭キーポイントの予測位置情報を決定することができる。骨格キーポイントの予測位置情報、実際の位置情報及び輪郭キーポイントの予測位置情報、実際の位置情報に基づいて、変換損失を決定する。

【0248】

ステップ5.11において、変換損失に基づいて、第5基礎変換ニューラルネットワークに対して今回の訓練を行う。

【0249】

ステップ5.12において、基礎変位推定ニューラルネットワーク、第5基礎変換ニュー

10

20

30

40

50

ラルネットワークに対して複数回の訓練を行うことで、特徴融合ニューラルネットワークを得る。

【0250】

Bにおいて、検出対象画像に対して、複数回の特徴抽出を行い、毎回の特徴抽出を行った後、該特徴抽出で得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、最終回の特徴融合の特徴融合結果に基づいて、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報を決定する。

【0251】

複数回の特徴抽出を行う場合、 i 回回の特徴融合の特徴融合結果に基づいて、 $i + 1$ 回回の特徴抽出を行い、 i は、正整数である。

【0252】

Bにおいて、1 回回の特徴抽出を行う過程は、上記 A における検出対象画像に対して骨格特徴及び輪郭特徴を抽出する過程と一致し、ここで、詳細な説明を省略する。

【0253】

Bにおいて、1 回回の特徴抽出以外の他の特徴抽出を行う具体的な過程は以下のとおりである。

【0254】

第2特徴抽出ネットワークを用いて、前回の特徴融合の特徴融合結果から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出する。

【0255】

ここで、第1特徴抽出ネットワークのネットワークパラメータは、第2特徴抽出ネットワークのネットワークパラメータと異なり、且つ異なる特徴抽出に用いられる第2特徴抽出ネットワークのネットワークパラメータは異なる。

【0256】

ここで、第1特徴抽出ネットワーク及び第2特徴抽出ネットワークはいずれも複数の畳み込み層を含む。第1特徴抽出ネットワーク及び第2特徴抽出ネットワークのネットワークパラメータは、例えば、畳み込み層の数、各畳み込み層に用いられる畳み込みカーネルのサイズ、各畳み込み層に用いられる畳み込みカーネルの数などを含むが、これらに限定されない。

【0257】

図14に示すように、本願の実施例は、第2特徴抽出ネットワークの構造を示す概略図を提供する。第2特徴抽出ネットワークは、第2骨格特徴抽出ネットワーク、及び第2輪郭特徴抽出ネットワークを含む。

【0258】

該第2特徴抽出ネットワークを用いて今回の特徴提取を行うための前回の特徴融合の特徴融合結果は、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を含む。第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るための具体的な過程は、上記 A を参照することができ、ここで詳細な説明を省略する。

【0259】

該第2特徴抽出ネットワークを用いて、前回の特徴融合の特徴融合結果から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出するための具体的な過程は例えば以下のとおりである。

【0260】

第2骨格特徴抽出ネットワークを用いて、前回の特徴融合で得られた第2ターゲット骨格特徴行列に対して畳み込み処理を行い、第3骨格特徴行列を得て、第2骨格特徴抽出ネットワークにおける第3ターゲット畳み込み層から、第4骨格特徴行列を得て、第3骨格特徴行列及び第4骨格特徴行列に基づいて、第5ターゲット骨格特徴行列を得る。ここで、第3ターゲット畳み込み層は、第2骨格特徴抽出ネットワークにおける最後の一つの畳み

10

20

30

40

50

込み層以外のいずれか1つの畳み込み層である。

【0261】

第2輪郭特徴抽出ネットワークを用いて、前回の特徴融合で得られた第2ターゲット輪郭特徴行列に対して畳み込み処理を行い、第3輪郭特徴行列を得て、第2輪郭特徴抽出ネットワークにおける第4ターゲット畳み込み層から、第4輪郭特徴行列を得て、第3輪郭特徴行列及び第4輪郭特徴行列に基づいて、第6ターゲット輪郭特徴行列を得る。第4ターゲット畳み込み層は、第2輪郭特徴抽出ネットワークにおける最後の一つの畳み込み層以外のいずれか1つの畳み込み層である。

【0262】

具体的な処理方式は、上記Aにおける第1骨格特徴抽出ネットワークを用いて検出対象画像から第1ターゲット骨格特徴行列及び第1ターゲット輪郭特徴行列を抽出する具体的な過程と同様であり、ここで詳細な説明を省略する。

【0263】

上記実施例において、上記IIにおける骨格キーポイント及び輪郭キーポイントの位置情報を決定する方式を説明した。

【0264】

IIIにおいて、上記IIに基づいて骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を得た後、各骨格キーポイントの位置及び輪郭キーポイントの位置を検出対象画像から決定することができる。続いて、人体検出結果を生成することができる。

【0265】

人体検出結果は、骨格キーポイントマーク及び輪郭キーポイントマークを含む検出対象画像、骨格キーポイントの位置情報及び輪郭キーポイントの位置情報を含むデータ群のうちの1つ又は複数を含む。

【0266】

続いて、人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集及び人体画像の貼り付けのうちの1つ又は複数を実行することもできる。

【0267】

ここで、動作認識は、例えば、殴り合い、ランニングなどのような現在の動作を認識することを指す。人体姿勢認識は、例えば、伏倒、指定した動作を行っているかどうかなどのような人体の現在の姿勢を認識することを指す。人体輪郭調整は、例えば、人体の体型、身長などを調整するを指す。人体画像編集は、例えば、人体画像に対して拡大縮小、回転、トリミングなどを行うことを指す。人体画像の貼り付けは、例えば、画像Aにおける人体を検出した後、対応する人体画像を画像Bに貼り付けることを指す。

【0268】

本願の実施例は、検出対象画像から、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定し、骨格キーポイントの位置情報、及び輪郭キーポイントの位置情報に基づいて、人体検出結果を生成する。表現精細度の向上と演算データ量を両立させる。

【0269】

また、本願の実施形態において、人体骨格構造を示すための骨格キーポイントの位置情報及び人体輪郭を示すための輪郭キーポイントの位置情報を用いて人体検出結果を得るため、人体を表すための情報はより豊富であり、適用シーンは、より広く、例えば、画像編集、人体体型調整などに適用可能である。

【0270】

同一の技術的思想によれば、本願の実施例は、人体検出方法に対応する人体検出装置を更に提供する。本願の実施例における装置による課題を解決するための原理は、本願の実施例における上記人体検出方法と類似するため、装置の実施は、方法の実施を参照することができ、重複の説明を省略する。

【0271】

図15は、本願の実施例による人体検出装置を示す概略図である。前記装置は、取得モジ

10

20

30

40

50

ジュール151、検出モジュール152及び生成モジュール153を含み、取得モジュール151は、検出対象画像を取得するように構成され、検出モジュール152は、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成され、生成モジュール153は、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成するように構成される。

【0272】

可能な実施形態において、前記輪郭キーポイントは、主輪郭キーポイント及び補助輪郭キーポイントを含み、ここで、2つの隣接する前記主輪郭キーポイントの間に、少なくとも1つの補助輪郭キーポイントが存在する。

10

【0273】

可能な実施形態において、前記検出モジュール152は、前記検出対象画像に基づいて、前記主輪郭キーポイントの位置情報を決定し、前記主輪郭キーポイントの位置情報に基づいて、人体輪郭情報を決定し、決定された前記人体輪郭情報に基づいて、複数の前記補助輪郭キーポイントの位置情報を決定するという方式で、前記検出対象画像に基づいて、人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成される。

【0274】

可能な実施形態において、前記人体検出結果は、骨格キーポイントマーク及び輪郭キーポイントマークが付加された検出対象画像、前記骨格キーポイントの位置情報及び前記輪郭キーポイントの位置情報を含むデータ群のうちの1つ又は複数を含む。

20

【0275】

可能な実施形態において、該人体検出装置は、前記人体検出結果に基づいて、人体動作認識、人体姿勢検出、人体輪郭調整、人体画像編集及び人体画像の貼り付けのうちの1つ又は複数を実行するように構成される実行モジュール154を更に備える。

【0276】

可能な実施形態において、前記検出モジュール152は、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定するという方式で、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成される。

30

【0277】

可能な実施形態において、前記検出モジュール152は、前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行い、特徴抽出を行うたびに得られた骨格特徴及び輪郭特徴に対して特徴融合を行い、複数回の特徴抽出を行う場合、 i 回目の特徴融合の特徴融合結果に基づいて $i+1$ 回目の特徴抽出を行い、 i は正整数であるという方式で、前記検出対象画像に基づいて、特徴抽出を行い、骨格特徴及び輪郭特徴を得て、得られた骨格特徴及び輪郭特徴に対して特徴融合を行うように構成され、前記検出モジュール152は、最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定するという方式で、特徴融合結果に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定するように構成される。

40

【0278】

可能な実施形態において、前記検出モジュール152は、1回目の特徴抽出において、事前訓練された第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出し、 $i+1$ 回の特徴抽出において、事前訓練された第2特徴抽出ネットワークを用いて、 i 回目の特徴融合の特徴融合結果から、前記第1ターゲット骨格特徴行列、及び前記第1ターゲット輪郭特徴行列を抽出するという方式で、前記検出対象画像に基づいて、少なくとも1回の特徴抽出を行う

50

ように構成され、第1特徴抽出ネットワークのネットワークパラメータは、第2特徴抽出ネットワークのネットワークパラメータと異なり、且つ特徴抽出のたびに用いられる第2特徴抽出ネットワークのネットワークパラメータは異なる。

【0279】

考えられる実現形態において、前記検出モジュール152は、事前訓練された特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るという方式で、抽出して得られた骨格特徴及び輪郭特徴に対して特徴融合を行うように構成され、前記第2ターゲット骨格特徴行列は、三次元骨格特徴行列であり、該三次元骨格特徴行列は、各骨格キーポイントにそれぞれ対応する二次元骨格特徴行列を含み、前記二次元骨格特徴行列における各要素の値は、該要素に対応する画素点が、対応する骨格キーポイントに属する確率を表し、前記第2ターゲット輪郭特徴行列は、三次元輪郭特徴行列であり、該三次元輪郭特徴行列は、各輪郭キーポイントにそれぞれ対応する二次元輪郭特徴行列を含み、前記二次元輪郭特徴行列における各要素の値は、該要素に対応する画素点が、対応する輪郭キーポイントに属する確率を表し、特徴融合のたびに用いられる特徴融合ニューラルネットワークのネットワークパラメータは異なる。

10

【0280】

考えられる実現形態において、前記検出モジュール152は、最終回の特徴融合で得られた第2ターゲット骨格特徴行列に基づいて、前記骨格キーポイントの位置情報を決定し、最終回の特徴融合で得られた第2ターゲット輪郭特徴行列に基づいて、前記輪郭キーポイントの位置情報を決定するという方式で、最終回の特徴融合の特徴融合結果に基づいて、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報を決定するように構成される。

20

【0281】

考えられる実現形態において、第1特徴抽出ネットワークは、共通特徴抽出ネットワーク、第1骨格特徴抽出ネットワーク及び第1輪郭特徴抽出ネットワークを含み、前記検出モジュール152は、前記共通特徴抽出ネットワークを用いて、前記検出対象画像に対して畳み込み処理を行い、骨格特徴及び輪郭特徴を含む基礎特徴行列を得ることと、前記第1骨格特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1骨格特徴行列を得て、前記第1骨格特徴抽出ネットワークにおける第1ターゲット畳み込み層から、第2骨格特徴行列を取得し、前記第1骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得ることとあって、前記第1ターゲット畳み込み層は、前記第1骨格特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、前記第1輪郭特徴抽出ネットワークを用いて、前記基礎特徴行列に対して畳み込み処理を行い、第1輪郭特徴行列を得て、前記第1輪郭特徴抽出ネットワークにおける第2ターゲット畳み込み層から、第2輪郭特徴行列を取得し、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることとあって、前記第2ターゲット畳み込み層は、前記第1輪郭特徴抽出ネットワークにおける、最後の1つの畳み込み層以外のいずれか1つの畳み込み層である、ことと、を含む方式で、第1特徴抽出ネットワークを用いて、検出対象画像から、人体骨格特徴を示すための骨格キーポイントの第1ターゲット骨格特徴行列を抽出し、人体輪郭特徴を示すための輪郭キーポイントの第1ターゲット輪郭特徴行列を抽出するように構成される。

30

40

【0282】

考えられる実現形態において、前記検出モジュール152は、前記第1骨格特徴行列及び前記第2骨格特徴行列に対して連結処理を行い、第1連結骨格特徴行列を得て、前記第1連結骨格特徴行列を次元変換処理し、前記第1ターゲット骨格特徴行列を得るという方式で、前記第1骨格特徴行列及び前記第2骨格特徴行列に基づいて、前記第1ターゲット骨格特徴行列を得るように構成され、前記第1輪郭特徴行列及び前記第2輪郭特徴行列に基づいて、前記第1ターゲット輪郭特徴行列を得ることは、前記第1輪郭特徴行列及び前記

50

第2輪郭特徴行列に対して連結処理を行い、第1連結輪郭特徴行列を得ることと、前記第1連結輪郭特徴行列を次元変換処理し、前記ターゲット輪郭特徴行列を得ることと、を含み、前記第1ターゲット骨格特徴行列の次元は、前記第1ターゲット輪郭特徴行列の次元と同じであり、且つ前記第1ターゲット骨格特徴行列と前記第1ターゲット輪郭特徴行列は、同一の次元における次元数が同じである。

【0283】

可能な実施形態において、前記特徴融合ニューラルネットワークは、第1畳み込みニューラルネットワーク、第2畳み込みニューラルネットワーク、第1変換ニューラルネットワーク及び第2変換ニューラルネットワークを含み、

前記検出モジュール152は、前記第1畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して畳み込み処理を行い、第1中間骨格特徴行列を得て、前記第2畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して畳み込み処理を行い、第1中間輪郭特徴行列を得ることと、前記第1中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第1連結特徴行列を得て、前記第1変換ニューラルネットワークを用いて、前記第1連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第1中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第2連結特徴行列を得て、前記第2変換ニューラルネットワークを用いて、前記第2連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含む方式で、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るように構成される。

【0284】

可能な実施形態において、前記特徴融合ニューラルネットワークは、第1定向畳み込みニューラルネットワーク、第2定向畳み込みニューラルネットワーク、第3畳み込みニューラルネットワーク、第4畳み込みニューラルネットワーク、第3変換ニューラルネットワーク、及び第4変換ニューラルネットワークを含み、

前記検出モジュール152は、前記第1定向畳み込みニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列に対して、定向畳み込み処理を行い、第1定向骨格特徴行列を得て、第3畳み込みニューラルネットワークを用いて、前記第1定向骨格特徴行列に対して畳み込み処理を行い、第2中間骨格特徴行列を得ることと、前記第2定向畳み込みニューラルネットワークを用いて、前記第1ターゲット輪郭特徴行列に対して、定向畳み込み処理を行い、第1定向輪郭特徴行列を得て、第4畳み込みニューラルネットワークを用いて、前記第1定向輪郭特徴行列に対して、畳み込み処理を行い、第2中間輪郭特徴行列を得ることと、前記第2中間輪郭特徴行列と前記第1ターゲット骨格特徴行列に対して連結処理を行い、第3連結特徴行列を得て、第3変換ニューラルネットワークを用いて、前記第3連結特徴行列に対して次元変換を行い、前記第2ターゲット骨格特徴行列を得ることと、前記第2中間骨格特徴行列と前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第4連結特徴行列を得て、第4変換ニューラルネットワークを用いて、前記第4連結特徴行列に対して次元変換を行い、前記第2ターゲット輪郭特徴行列を得ることと、を含む方式で、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るように構成される。

【0285】

可能な実施形態において、前記特徴融合ニューラルネットワークは、変位推定ニューラルネットワーク、第5変換ニューラルネットワークを含み、

前記検出モジュール152は、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して連結処理を行い、第5連結特徴行列を得ることと、前記第5連結特徴行列を前記変位推定ニューラルネットワークに入力し、事前決定された複数組のキーポイントペアに対して変位推定を行い、各組のキーポイントペアのうちの1つのキーポイン

10

20

30

40

50

トがもう1つのキーポイントまでに移動する変位情報を得て、各組のキーポイントペアのうちの各キーポイントをそれぞれ現在のキーポイントとし、該現在のキーポイントのペアであるもう1つのキーポイントに対応する三次元特徴行列から、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列を得ることと、前記ペアであるもう1つのキーポイントから前記現在のキーポイントまでの変位情報に基づいて、前記ペアであるもう1つのキーポイントに対応する二次元特徴行列における要素に対して位置変換を行い、該現在のキーポイントに対応する変位特徴行列を得ることと、各骨格キーポイントについて、該骨格キーポイントに対応する二次元特徴行列と、該骨格キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該骨格キーポイントの連結二次元特徴行列を得て、該骨格キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該骨格キーポイントに対応するターゲット二次元特徴行列を得て、各骨格キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット骨格特徴行列を生成することと、各輪郭キーポイントについて、該輪郭キーポイントに対応する二次元特徴行列と、該輪郭キーポイントに対応する各変位特徴行列とに対して連結処理を行い、該輪郭キーポイントの連結二次元特徴行列を得て、該輪郭キーポイントの連結二次元特徴行列を前記第5変換ニューラルネットワークに入力し、該輪郭キーポイントに対応するターゲット二次元特徴行列を得て、各輪郭キーポイントにそれぞれ対応するターゲット二次元特徴行列に基づいて、前記第2ターゲット輪郭特徴行列を生成することと、を含む方式で、特徴融合ニューラルネットワークを用いて、前記第1ターゲット骨格特徴行列及び前記第1ターゲット輪郭特徴行列に対して特徴融合を行い、第2ターゲット骨格特徴行列及び第2ターゲット輪郭特徴行列を得るように構成される。

10

20

【0286】

可能な実施形態において、前記人体検出方法は、人体検出モデルにより実現され、前記人体検出モデルは、前記第1特徴抽出ネットワーク及び／又は前記特徴融合ニューラルネットワークを含み、前記人体検出モデルは、訓練サンプル集合におけるサンプル画像を利用して訓練されたものであり、前記サンプル画像に、人体骨格構造の骨格キーポイントの実際の位置情報、及び人体輪郭の輪郭キーポイントの実際の位置情報がラベル付けされている。

【0287】

装置における各モジュールの処理フロー及び各モジュール間のインタラクションフローは、上記方法実施例における関連説明を参照することができ、ここで、詳細な説明を省略する。

30

【0288】

本願の実施例は、コンピュータ機器を更に提供する。図16は、本願の実施例によるコンピュータ機器の構造を示す概略図である。前記コンピュータ機器は、プロセッサ11と、記憶媒体12と、バス13と、を備え、記憶媒体12は、実行可能な命令を記憶するためのものであり、メモリ121及び外部メモリ122を含み、このメモリ121は、内部メモリとも呼ばれ、プロセッサ11における処理データ及びハードディスク等の外部メモリ122と交換されるデータを一時的に記憶するためのものであり、プロセッサ11は、メモリ121を介して外部メモリ122とデータ交換を行う。前記コンピュータ機器100が実行される場合、前記プロセッサ11と前記記憶媒体12は、バス13を介して通信し、前記プロセッサ11に、検出対象画像を取得し、前記検出対象画像に基づいて、人体骨格構造を示すための骨格キーポイントの位置情報、及び人体輪郭を示すための輪郭キーポイントの位置情報を決定し、前記骨格キーポイントの位置情報、及び前記輪郭キーポイントの位置情報に基づいて、人体検出結果を生成するという命令を実行させる。

40

【0289】

本願の実施例は、コンピュータ可読記憶媒体を更に提供する。該コンピュータ可読記憶媒体にコンピュータプログラムが記憶されており、該コンピュータプログラムがプロセッサにより実行される場合、上記方法の実施例に記載の人体検出方法のステップを実行する。

50

【0290】

本願の実施例で提供される人体検出方法のコンピュータプログラム製品は、プログラムコードを記憶したコンピュータ可読記憶媒体を含む。前記プログラムコードに含まれる命令は、上記方法実施例における前記人体検出方法のステップを実行するために用いられる。具体的には、上記方法実施例を参照することができ、ここで詳細な説明を省略する。

【0291】

説明上の便宜及び簡素化を図るために、上記説明されたシステム、及び装置の具体的な作動過程は、前記方法の実施例における対応した過程を参照することができるから、ここで詳しく説明しないようにすることは、当業者にはっきり理解されるべきである。本発明で提供する幾つかの実施例で開示したシステム、装置及び方法は、他の方式によって実現できることを理解すべきである。例えば、以上に記載した装置の実施例はただ例示的なもので、例えば、前記ユニットの分割はただロジック機能の分割で、実際に実現する時は他の分割方式によってもよい。例えば、複数のユニット又は組立体を組み合わせてもよいし、別のシステムに組み込んでよい。又は若干の特徴を無視してもよいし、実行しなくてもよい。また、示したか或いは検討した相互間の結合又は直接的な結合又は通信接続は、幾つかのインタフェース、装置又はユニットによる間接的な結合又は通信接続であってもよく、電氣的、機械的または他の形態であってもよい。

【0292】

分離部材として説明した該ユニットは、物理的に別個のものであってもよいし、そうでなくてもよい。ユニットとして示された部材は、物理的ユニットであってもよいし、そうでなくてもよい。即ち、同一の位置に位置してもよいし、複数のネットワークに分布してもよい。実際の需要に応じてそのうちの一部又は全てのユニットにより本実施例の方策の目的を実現することができる。

【0293】

また、本願の各実施例における各機能ユニットは一つの処理ユニットに集積されてもよいし、各ユニットが物理的に別個のものとして存在してもよいし、2つ以上のユニットが一つのユニットに集積されてもよい。

【0294】

前記機能はソフトウェア機能ユニットの形で実現され、かつ独立した製品として販売または使用されるとき、コンピュータにより読み取り可能な記憶媒体内に記憶されてもよい。このような理解のもと、本発明の技術的解決手段は、本質的に、又は、従来技術に対して貢献をもたらした部分又は該技術的解決手段の一部は、ソフトウェア製品の形式で具現することができる。このようなコンピュータソフトウェア製品は、記憶媒体に記憶しても良く、また、コンピュータ設備（パソコン、サーバ、又はネットワーク装置など）に、本発明の各実施例に記載の方法の全部又は一部のステップを実行させるための若干の命令を含む。前記の記憶媒体は、Uディスク、リムーバブルハードディスク、読み出し専用メモリ（Read-only Memory：ROM）、ランダムアクセスメモリ（Random Access Memory：RAM）、磁気ディスク、又は光ディスクなど、プログラムコードを記憶可能な各種の媒体を含む。

【0295】

なお、上記実施例は本願の具体的な実施形態に過ぎず、本願の技術的解決手段を説明するためのものであり、これを限定するものではなく、本願の保護範囲はこれに制限されるものではなく、前記実施例を参照しながら、本願を詳細に説明したが、本技術分野を周知するいかなる当業者であれば、本願で開示された技術範囲内で、前記実施例に記載の技術的解決手段に対して変化または代替を容易に思いつくことができ、又は一部の技術的特徴に対して均等物による置換を行うこともでき、これらの修正、変化又は置換は、対応する技術的解決手段の本質を本願の実施例の技術的解決手段の製品及び範囲を離脱させるものではなく、本願の保護範囲内に含まれるものとするとは、当業者であれば、理解すべきである。従って、本願の保護範囲は特許請求の範囲の保護範囲を基準とするべきである。

10

20

30

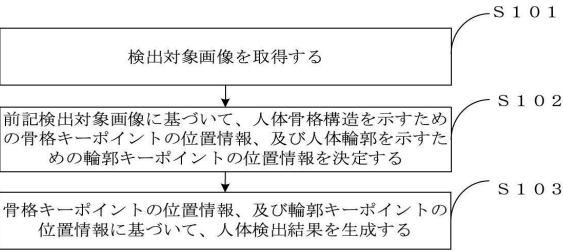
40

50

【図面】

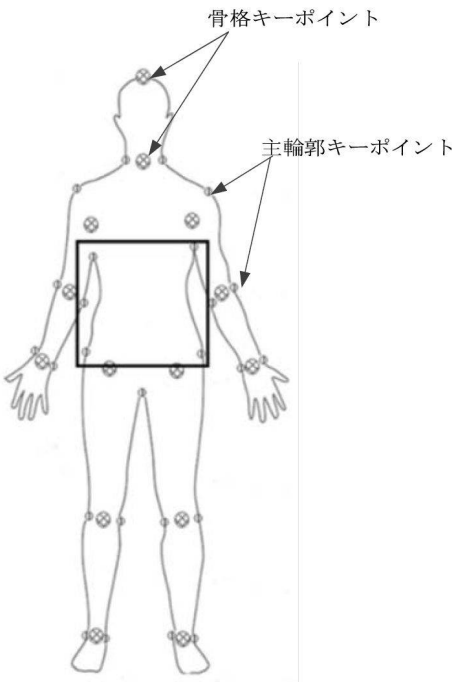
【図 1】

【図 1】



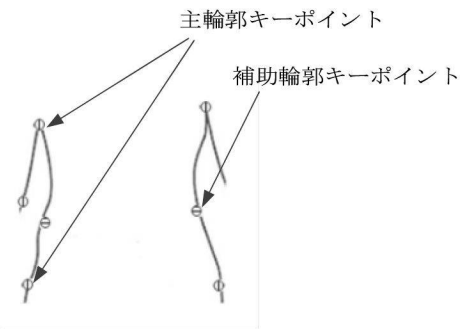
【図 2 a】

【図 2 a】



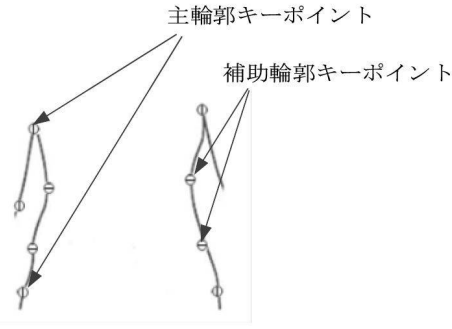
【図 2 b】

【図 2 b】



【図 2 c】

【図 2 c】



10

20

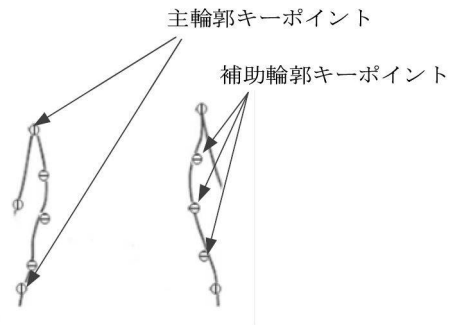
30

40

50

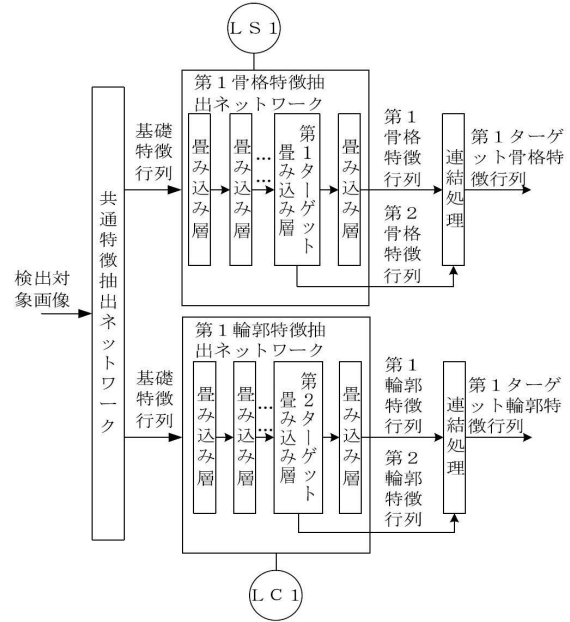
【図 2 d】

【図 2 d】



【図 3】

【図 3】

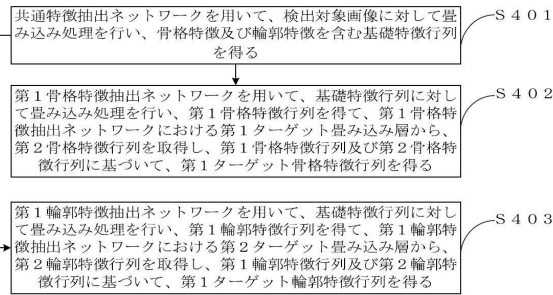


10

20

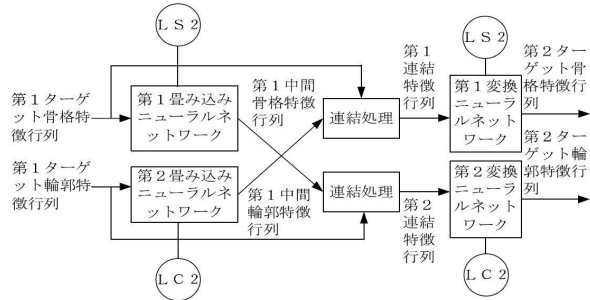
【図 4】

【図 4】



【図 5】

【図 5】



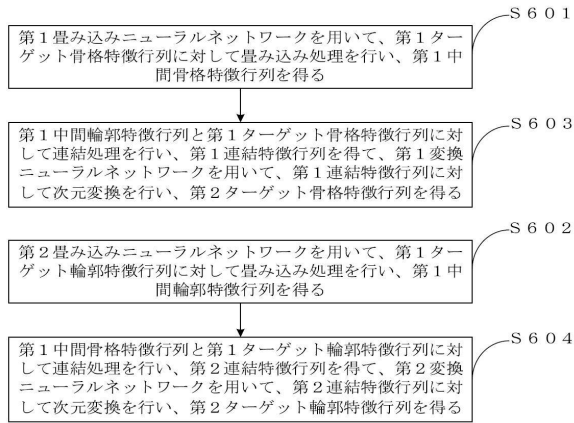
30

40

50

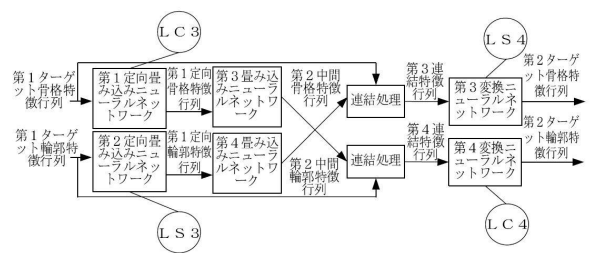
【図 6】

【図 6】



【図 7】

【図 7】



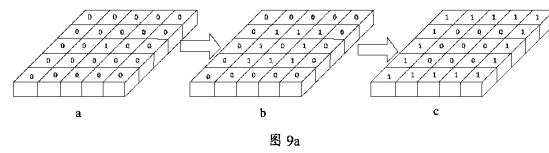
10

【図 8】

【図 8】



【図 9 a】



20

30

40

50

【図 9 b】

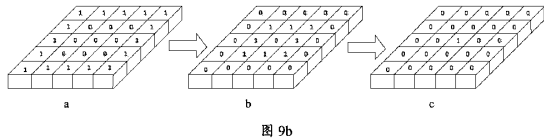
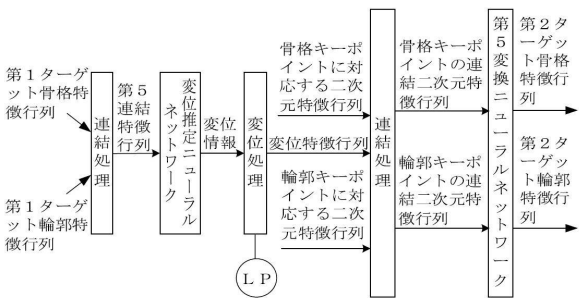


図 9b

【図 10】

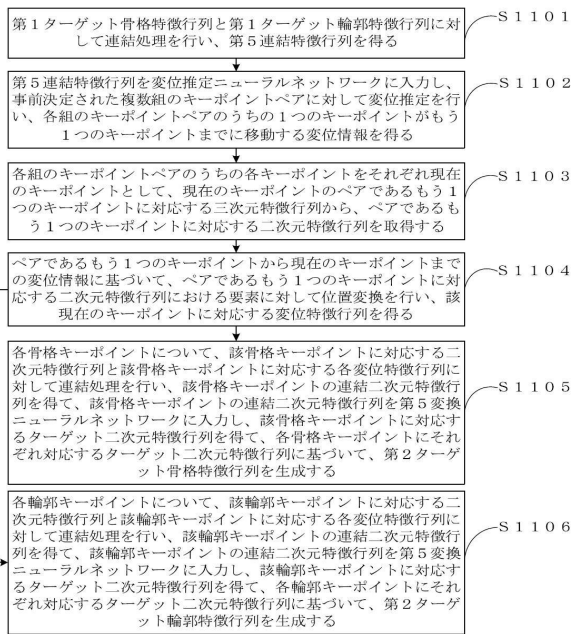
【図 10】



10

【図 11】

【図 11】



【図 12】

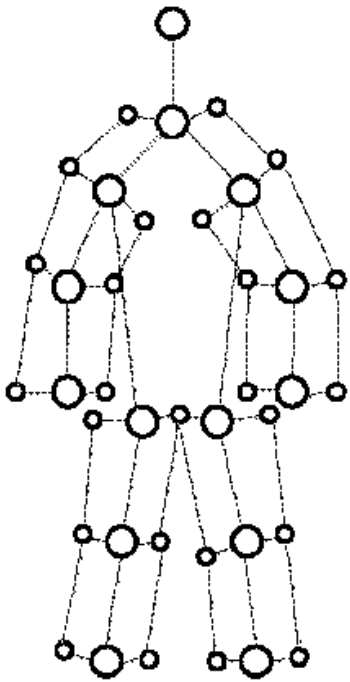


図 12

20

30

40

50

【図 1 3】

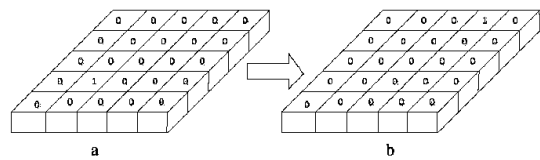
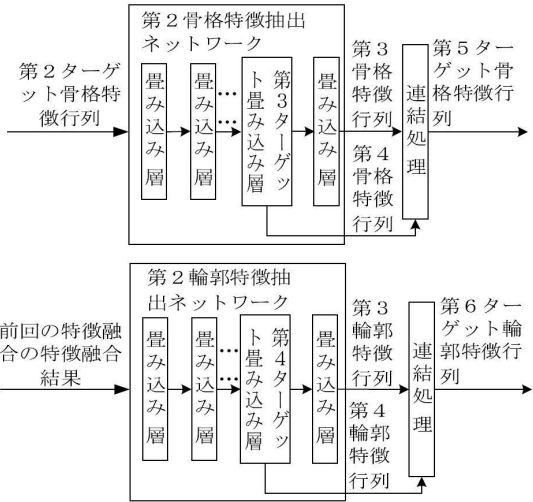


図 13

【図 1 4】

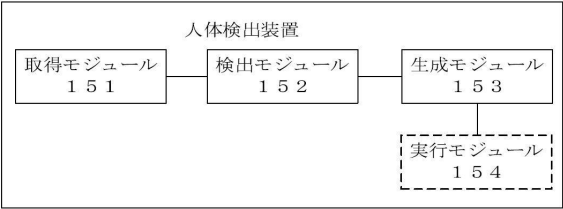
【図 1 4】



10

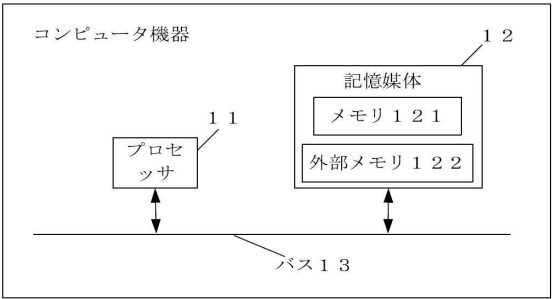
【図 1 5】

【図 1 5】



【図 1 6】

【図 1 6】



20

30

40

50

フロントページの続き

- (74)代理人 100181641
弁理士 石川 大輔
- (74)代理人 230113332
弁護士 山本 健策
- (72)発明者 段浩▲東▼
中華人民共和国 ベイジン 100080, ハイディアン, ナンバー 58 ノースウェスト 4
ティーエイチ リング ロード, フロア 11ティーエイチ, ルーム 1101-1117
- (72)発明者 ▲劉▼文▲韜▼
中華人民共和国 ベイジン 100080, ハイディアン, ナンバー 58 ノースウェスト 4
ティーエイチ リング ロード, フロア 11ティーエイチ, ルーム 1101-1117
- 審査官 間野 裕一
- (56)参考文献 中国特許出願公開第109242868 (CN, A)
特開2007-164641 (JP, A)
特表2014-522058 (JP, A)
特開2014-89665 (JP, A)
米国特許出願公開第2019/0197299 (US, A1)
国際公開第2019/141104 (WO, A1)
国際公開第2019/100888 (WO, A1)
国際公開第2019/001481 (WO, A1)
中国特許出願公開第109508625 (CN, A)
- (58)調査した分野 (Int.Cl., DB名)
G06T 7/00