

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7364077号
(P7364077)

(45)発行日 令和5年10月18日(2023.10.18)

(24)登録日 令和5年10月10日(2023.10.10)

(51)国際特許分類 F I
G 0 6 T 7/00 (2017.01) G 0 6 T 7/00 6 6 0 B
G 0 6 V 10/22 (2022.01) G 0 6 V 10/22

請求項の数 8 (全41頁)

(21)出願番号	特願2022-531159(P2022-531159)	(73)特許権者	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(86)(22)出願日	令和2年6月17日(2020.6.17)	(74)代理人	100110928 弁理士 速水 進治
(86)国際出願番号	PCT/JP2020/023707	(72)発明者	潘 雅冬 東京都港区芝五丁目7番1号 日本電気株式会社内
(87)国際公開番号	WO2021/255846	(72)発明者	吉田 登 東京都港区芝五丁目7番1号 日本電気株式会社内
(87)国際公開日	令和3年12月23日(2021.12.23)	(72)発明者	川合 諒 東京都港区芝五丁目7番1号 日本電気株式会社内
審査請求日	令和4年12月8日(2022.12.8)	(72)発明者	劉 健全

最終頁に続く

(54)【発明の名称】 画像処理装置、画像処理方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

画像を取得する画像取得手段と、
前記画像に含まれる人物の2次元骨格構造を検出する骨格構造検出手段と、
前記2次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段と、
前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段と、
前記補正後画像に含まれる人物の前記2次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段と、
を有する画像処理装置。

【請求項2】

前記指定手段は、ユーザ入力に基づき、又は、画像解析で、前記未検出人物領域を指定する請求項1に記載の画像処理装置。

【請求項3】

前記指定手段は、
前記2次元骨格構造の検出を行った前記画像の中に、前記2次元骨格構造の検出に失敗した人物が存在するか判定し、
前記2次元骨格構造の検出に失敗した人物が存在すると判定した前記画像に対し、前記未検出人物領域を指定する処理を実行する請求項1に記載の画像処理装置。

【請求項 4】

前記指定手段は、ユーザ入力に基づき、又は、画像解析で、前記 2 次元骨格構造の検出を行った前記画像の中に、前記 2 次元骨格構造の検出に失敗した人物が存在するか判定する請求項 3 に記載の画像処理装置。

【請求項 5】

前記指定手段は、

前記画像に含まれる人物を検出し、

検出された人物の数が、前記 2 次元骨格構造の検出に成功した人物の数より多い場合、前記 2 次元骨格構造の検出に失敗した人物が存在すると判定する請求項 4 に記載の画像処理装置。

10

【請求項 6】

検出された前記 2 次元骨格構造の特徴量を算出する特徴量算出手段と、

算出された前記特徴量の類似度に基づき、前記画像に含まれる人物の状態の認識処理を行う認識手段と、

を有する請求項 1 から 5 のいずれか 1 項に記載の画像処理装置。

【請求項 7】

コンピュータが、

画像を取得し、

前記画像に含まれる人物の 2 次元骨格構造を検出し、

前記 2 次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定し、

20

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成し、前記補正後画像に含まれる人物の前記 2 次元骨格構造を検出する処理を実行する画像処理方法。

【請求項 8】

コンピュータを、

画像を取得する画像取得手段、

前記画像に含まれる人物の 2 次元骨格構造を検出する骨格構造検出手段、

前記 2 次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段、

30

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段、

前記補正後画像に含まれる人物の前記 2 次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段、

として機能させるプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、画像処理装置、画像処理方法、及びプログラムに関する。

【背景技術】

40

【0002】

近年、監視システム等において、監視カメラの画像から人物の姿勢や行動等の状態の検出や検索を行う技術が利用されている。関連する技術として、例えば、特許文献 1 乃至 4 が知られている。

【0003】

特許文献 1 には、深さ映像に含まれる人物の頭や手足等のキージョイントに基づいて、類似する人物の姿勢を検索する技術が開示されている。特許文献 2 には、人物の姿勢と関連しないが、画像に付加された傾き等の姿勢情報を利用して類似画像を検索する技術が開示されている。特許文献 3 には、姿勢情報の入力を受け、指定された姿勢を含む画像を検索する技術が開示されている。特許文献 4 には、ユーザから画像内の人物の頭に対応す

50

る一点を指定する入力を受付けると、その点を所定位置とした矩形領域を切り出し、その後、エッジ検出等でその矩形領域の画像を背景領域と人物領域とを分け、人物領域の画像を指定された人物画像として処理する技術が開示されている。なお、その他に、人物の骨格推定に関連する技術として、非特許文献 1 が知られている。

【先行技術文献】

【特許文献】

【0004】

【文献】特表 2014 - 522035 号公報

特開 2006 - 260405 号公報

特開 2019 - 91138 号公報

国際公開第 2016 / 067573 号

【非特許文献】

【0005】

【文献】Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, P. 7291-7299

【発明の概要】

【発明が解決しようとする課題】

【0006】

画像内における人物の位置（端に位置するなど）や周辺の背景画像等に起因して、人物の状態（姿勢や行動）の検出が困難になる場合がある。しかし、このような人物の状態も検出できることが好ましい。本発明の目的は、画像内における人物の位置や周辺の背景画像等に関わらず、人物の状態を精度よく検出できるようにすることである。

【課題を解決するための手段】

【0007】

本発明によれば、
 画像を取得する画像取得手段と、
 前記画像に含まれる人物の 2 次元骨格構造を検出する骨格構造検出手段と、
 前記 2 次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段と、
 前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段と、
 前記補正後画像に含まれる人物の前記 2 次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段と、
 を備える画像処理装置が提供される。

【0008】

また、本発明によれば、
 コンピュータが、
 画像を取得し、
 前記画像に含まれる人物の 2 次元骨格構造を検出し、
 前記 2 次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定し、
 前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成し、
 前記補正後画像に含まれる人物の前記 2 次元骨格構造を検出する処理を実行する画像処理方法が提供される。

【0009】

また、本発明によれば、
 コンピュータを、
 画像を取得する画像取得手段、
 前記画像に含まれる人物の 2 次元骨格構造を検出する骨格構造検出手段、

10

20

30

40

50

前記 2 次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段、

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段、

前記補正後画像に含まれる人物の前記 2 次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段、

として機能させるプログラムが提供される。

【発明の効果】

【 0 0 1 0 】

本発明によれば、画像内における人物の位置や周辺の背景画像等に関わらず、人物の状態を精度よく検出できるようになる。

10

【図面の簡単な説明】

【 0 0 1 1 】

【図 1】実施の形態に係る画像処理装置の概要を示す構成図である。

【図 2】実施の形態 1 に係る画像処理装置の構成を示す構成図である。

【図 3】実施の形態 1 に係る画像処理方法を示すフローチャートである。

【図 4】実施の形態 1 に係る分類方法を示すフローチャートである。

【図 5】実施の形態 1 に係る検索方法を示すフローチャートである。

【図 6】実施の形態 1 に係る骨格構造の検出例を示す図である。

【図 7】実施の形態 1 に係る人体モデルを示す図である。

20

【図 8】実施の形態 1 に係る骨格構造の検出例を示す図である。

【図 9】実施の形態 1 に係る骨格構造の検出例を示す図である。

【図 10】実施の形態 1 に係る骨格構造の検出例を示す図である。

【図 11】実施の形態 1 に係る分類方法の具体例を示すグラフである。

【図 12】実施の形態 1 に係る分類結果の表示例を示す図である。

【図 13】実施の形態 1 に係る検索方法を説明するための図である。

【図 14】実施の形態 1 に係る検索方法を説明するための図である。

【図 15】実施の形態 1 に係る検索方法を説明するための図である。

【図 16】実施の形態 1 に係る検索方法を説明するための図である。

【図 17】実施の形態 1 に係る検索結果の表示例を示す図である。

30

【図 18】実施の形態 2 に係る画像処理装置の構成を示す構成図である。

【図 19】実施の形態 2 に係る画像処理方法を示すフローチャートである。

【図 20】実施の形態 2 に係る身長画素数算出方法の具体例 1 を示すフローチャートである。

【図 21】実施の形態 2 に係る身長画素数算出方法の具体例 2 を示すフローチャートである。

【図 22】実施の形態 2 に係る身長画素数算出方法の具体例 2 を示すフローチャートである。

【図 23】実施の形態 2 に係る正規化方法を示すフローチャートである。

【図 24】実施の形態 2 に係る人体モデルを示す図である。

40

【図 25】実施の形態 2 に係る骨格構造の検出例を示す図である。

【図 26】実施の形態 2 に係る骨格構造の検出例を示す図である。

【図 27】実施の形態 2 に係る骨格構造の検出例を示す図である。

【図 28】実施の形態 2 に係る人体モデルを示す図である。

【図 29】実施の形態 2 に係る骨格構造の検出例を示す図である。

【図 30】実施の形態 2 に係る身長画素数算出方法を説明するためのヒストグラムである。

【図 31】実施の形態 2 に係る骨格構造の検出例を示す図である。

【図 32】実施の形態 2 に係る 3 次元人体モデルを示す図である。

【図 33】実施の形態 2 に係る身長画素数算出方法を説明するための図である。

【図 34】実施の形態 2 に係る身長画素数算出方法を説明するための図である。

50

【図 3 5】実施の形態 2 に係る身長画素数算出方法を説明するための図である。

【図 3 6】実施の形態 2 に係る正規化方法を説明するための図である。

【図 3 7】実施の形態 2 に係る正規化方法を説明するための図である。

【図 3 8】実施の形態 2 に係る正規化方法を説明するための図である。

【図 3 9】画像処理装置が処理する情報の一例を模式的に示す図である。

【図 4 0】画像処理装置が処理する情報の一例を模式的に示す図である。

【図 4 1】画像処理装置が出力する画面の一例を模式的に示す図である。

【図 4 2】画像処理装置が出力する画面の一例を模式的に示す図である。

【図 4 3】画像処理装置のハードウェア構成例を示す図である。

【図 4 4】画像処理装置の処理の流れの一例を示すフローチャートである。

10

【図 4 5】画像処理装置が出力する画面の一例を模式的に示す図である。

【図 4 6】画像処理装置が生成する画像の一例を模式的に示す図である。

【図 4 7】画像処理装置の処理の流れの一例を示すフローチャートである。

【発明を実施するための形態】

【0012】

以下、本発明の実施の形態について、図面を用いて説明する。尚、すべての図面において、同様な構成要素には同様の符号を付し、適宜説明を省略する。

【0013】

(実施の形態に至る検討)

近年、ディープラーニング等の機械学習を活用した画像認識技術が様々なシステムに応用されている。例えば、監視カメラの画像により監視を行う監視システムへの適用が進められている。監視システムに機械学習を活用することで、画像から人物の姿勢や行動等の状態をある程度把握することが可能とされつつある。

20

【0014】

しかしながら、このような関連する技術では、必ずしもオンデマンドにユーザが望む人物の状態を把握できない場合がある。例えば、ユーザが検索し把握したい人物の状態を事前に特定できている場合もあれば、未知の状態のように具体的に特定できていない場合もある。そうすると、場合によっては、ユーザが検索したい人物の状態を詳細に指定することができない。また、人物の体の一部が隠れているような場合には検索等を行うことができない。関連する技術では、特定の検索条件のみからしか人物の状態を検索できないため、所望の人物の状態を柔軟に検索や分類することが困難である。

30

【0015】

発明者らは、オンデマンドに画像からユーザ所望の人物の状態を認識するため、非特許文献 1 などの骨格推定技術を利用する方法を検討した。非特許文献 1 に開示された Open Pose 等のように、関連する骨格推定技術では、様々なパターンの正解付けされた画像データを学習することで、人物の骨格を推定する。以下の実施の形態では、このような骨格推定技術を活用することで、人物の状態を柔軟に認識することを可能とする。

【0016】

なお、Open Pose 等の骨格推定技術により推定される骨格構造は、関節等の特徴的な点である「キーポイント」と、キーポイント間のリンクを示す「ボーン(ボーンリンク)」とから構成される。このため、以下の実施の形態では、骨格構造について「キーポイント」と「ボーン」という用語を用いて説明するが、特に限定されない限り、「キーポイント」は人物の「関節」に対応し、「ボーン」は人物の「骨」に対応している。

40

【0017】

(実施の形態の概要)

図 1 は、実施の形態に係る画像処理装置 10 の概要を示している。図 1 に示すように、画像処理装置 10 は、骨格検出部 11、特徴量算出部 12、及び認識部 13 を備えている。骨格検出部 11 は、カメラ等から取得される 2 次元画像に基づいて、複数の人物の 2 次元骨格構造(以下、単に「骨格構造」という場合がある)を検出する。特徴量算出部 12 は、骨格検出部 11 により検出された複数の 2 次元骨格構造の特徴量を算出する。認識部

50

13は、特徴量算出部12により算出された複数の特徴量の類似度に基づいて、複数の人物の状態の認識処理を行う。認識処理は、人物の状態の分類処理や検索処理等である。

【0018】

このように、実施の形態では、2次元画像から人物の2次元骨格構造を検出し、この2次元骨格構造から算出される特徴量に基づいて人物の状態の分類や検索等の認識処理を行う。

【0019】

(実施の形態1)

以下、図面を参照して実施の形態1について説明する。図2は、本実施の形態に係る画像処理装置100の構成を示している。画像処理装置100は、カメラ200及び記憶手段(データベース(DB)110)、とともに画像処理システム1を構成する。画像処理装置100を含む画像処理システム1は、画像から推定される人物の骨格構造に基づき、人物の姿勢や行動等の状態を分類及び検索するシステムである。

【0020】

カメラ200は、2次元の画像を生成する監視カメラ等の撮像部である。カメラ200は、所定の箇所に設置されて、その設置箇所から撮像領域における人物等を撮像する。カメラ200は、撮像した画像(映像)を画像処理装置100へ出力可能に有線又は無線で直接接続、もしくは任意の通信ネットワーク等を介して接続されている。なお、カメラ200を画像処理装置100の内部に設けてもよい。

【0021】

データベース110は、画像処理装置100の処理に必要な情報(データ)や処理結果等を格納するデータベースである。データベース110は、画像取得部101が取得した画像や、補正部111が生成した補正後画像や、骨格構造検出部102の検出結果、機械学習用のデータ、特徴量算出部103が算出した特徴量、分類部104の分類結果、検索部105の検索結果等を記憶する。データベース110は、画像処理装置100と必要に応じてデータを入出力可能に有線又は無線で直接接続、もしくは任意の通信ネットワーク等を介して接続されている。なお、データベース110をフラッシュメモリなどの不揮発性メモリやハードディスク装置等として、画像処理装置100の内部に設けてもよい。

【0022】

図2に示すように、画像処理装置100は、画像取得部101、骨格構造検出部102、特徴量算出部103、分類部104、検索部105、入力部106、表示部107、指定部109、補正部111及び再検出部112を備えている。なお、各部(ブロック)の構成は一例であり、後述の方法(動作)が可能であれば、その他の各部分で構成されてもよい。また、画像処理装置100は、例えば、プログラムを実行するパーソナルコンピュータやサーバ等のコンピュータ装置で実現されるが、1つの装置で実現してもよいし、ネットワーク上の複数の装置で実現してもよい。例えば、入力部106や表示部107等を外部の装置としてもよい。また、分類部104及び検索部105の両方を備えていてもよいし、いずれか一方のみを備えていてもよい。分類部104及び検索部105の両方、もしくは一方は、人物の状態の認識処理を行う認識部13である。また、骨格構造検出部102、指定部109、補正部111及び再検出部112は、図1の骨格検出部11に該当する。

【0023】

画像処理装置100は、データ蓄積処理と、分類処理と、検索処理とをこの順に実行する。なお、以下で説明するように、画像処理装置100は分類処理を実行しなくてもよい。

【0024】

<データ蓄積処理>

データ蓄積処理は、解析対象の画像(以下、「解析対象画像」)を取得し、複数の解析対象画像各々から人物の2次元骨格構造を検出し、検出された2次元骨格構造の特徴量を算出し、算出した特徴量を各解析対象画像に紐付けてデータベース110に格納する処理である。以下、データ蓄積処理に関わる機能部の構成を説明する。

10

20

30

40

50

【 0 0 2 5 】

画像取得部 1 0 1 は、解析対象画像を取得する。本明細書において、「取得」とは、ユーザ入力に基づき、又は、プログラムの指示に基づき、「自装置が他の装置や記憶媒体に格納されているデータを取りに行くこと（能動的な取得）」、たとえば、他の装置にリクエストまたは問い合わせして受信すること、他の装置や記憶媒体にアクセスして読み出すこと等、および、ユーザ入力に基づき、又は、プログラムの指示に基づき、「自装置に他の装置から出力されるデータを入力すること（受動的な取得）」、たとえば、配信（または、送信、プッシュ通知等）されるデータを受信すること、また、受信したデータまたは情報の中から選択して取得すること、及び、「データを編集（テキスト化、データの並び替え、一部データの抽出、ファイル形式の変更等）などして新たなデータを生成し、当該新たなデータを取得すること」の少なくともいずれか一方を含む。

10

【 0 0 2 6 】

例えば、画像取得部 1 0 1 は、所定の監視期間にカメラ 2 0 0 が撮像した人物を含む 2 次元の画像を解析対象画像として取得する。その他、画像取得部 1 0 1 は、データベース 1 1 0 等の記憶手段に記憶されている人物を含む 2 次元の画像を解析対象画像として取得してもよい。

【 0 0 2 7 】

骨格構造検出部 1 0 2 は、取得された解析対象画像各々から人物の 2 次元骨格構造を検出する。骨格構造検出部 1 0 2 は、解析対象画像の中で認識される全ての人物について、骨格構造を検出することができる。骨格構造検出部 1 0 2 は、機械学習を用いた骨格推定技術を用いて、認識される人物の関節等の特徴に基づき人物の骨格構造を検出する。骨格構造検出部 1 0 2 は、例えば、非特許文献 1 の Open Pose 等の骨格推定技術を用い、関節等の特徴的な点であるキーポイントを検出したりする。

20

【 0 0 2 8 】

指定部 1 0 9 は、2 次元骨格構造の検出を行った解析対象画像の中に、2 次元骨格構造の検出に失敗した人物が存在するか判定する（検出漏れ判定処理）。そして、指定部 1 0 9 は、2 次元骨格構造の検出に失敗した人物が存在すると判定した解析対象画像に対し、未検出人物領域を指定する処理を実行する（未検出人物領域指定処理）。

【 0 0 2 9 】

「2 次元骨格構造の検出に失敗した人物」は、骨格構造検出部 1 0 2 により所定数（1 以上の整数）以上のキーポイントが検出されなかった人物である。「未検出人物領域」は、2 次元骨格構造の検出に失敗した人物が存在する画像内の領域である。

30

【 0 0 3 0 】

まず、検出漏れ判定処理の一例を説明する。以下で説明するように、指定部 1 0 9 は、ユーザ入力に基づき、又は、画像解析で、2 次元骨格構造の検出を行った解析対象画像の中に 2 次元骨格構造の検出に失敗した人物が存在するか判定することができる。

【 0 0 3 1 】

「検出漏れ判定処理例 1」

指定部 1 0 9 は、図 4 5 に示すように、画像取得部 1 0 1 が取得した解析対象画像上に、骨格構造検出部 1 0 2 による 2 次元骨格構造の検出結果を重畳した画像を生成し、ユーザに向けて出力する。当該画像では、図示するように、検出されたキーポイントやボーンを検出された画像上の位置に表示する。結果、各人物から検出されたキーポイントやボーンは、各人物に重ねて表示される。図示する人物 P 1 乃至 P 3 は、キーポイントやボーンが表示されている。一方、人物 P 4 は、キーポイントやボーンが表示されていない。この表示により、人物 P 1 乃至 P 3 は 2 次元骨格構造の検出に成功し、人物 P 4 は 2 次元骨格構造の検出に失敗したことが分かる。

40

【 0 0 3 2 】

上述の画像を表示後、指定部 1 0 9 は、2 次元骨格構造の検出に失敗した人物が画像内に存在するか否かの入力をユーザから受け付ける。例えば、「当該画像の中に 2 次元骨格構造が検出されていない人物はいますか？」のような問い合わせ文を表示し、「Y e s

50

OR No」でその回答を受付けてもよい。ユーザは、キーポイントやボーンが重畳表示されていない人物や、その数が基準より少ない人物が画像内に存在する場合、2次元骨格構造の検出に失敗した人物が画像内に存在する旨を入力する。指定部109は、当該ユーザ入力に基づき、2次元骨格構造の検出に失敗した人物が存在するか判定する。

【0033】

「検出漏れ判定処理例2」

指定部109は、画像取得部101が取得した解析対象画像に対して、人物検出処理を行う。人物検出処理は、骨格構造検出部102による2次元骨格構造の検出と異なるアルゴリズム（例えば、顔を検出するアルゴリズム）で、画像内の人物を検出する処理である。当該人物検出処理で採用するアルゴリズムは特段制限されず、周知のあらゆる技術を採用できる。

10

【0034】

そして、指定部109は、上記人物検出処理で検出された人物の数が、2次元骨格構造の検出に成功した人物の数より多い場合、2次元骨格構造の検出に失敗した人物が存在すると判定する。なお、「2次元骨格構造の検出に成功した人物」は、所定数（1以上の整数）以上のキーポイントが検出された人物である。

【0035】

次に、未検出人物領域指定処理の一例を説明する。以下で説明するように、指定部109は、ユーザ入力に基づき、又は、画像解析で、未検出人物領域を指定することができる。

【0036】

「未検出人物領域指定処理例1」

指定部109は、図45に示すような画像を表示し、当該画像上で未検出人物領域を指定する入力を受付ける。当該入力で、人物の輪郭を指定する入力を受付けてもよい。その他、当該入力で、人物を含む所定形状（矩形）の領域を指定する入力を受付けてもよい。後者の例の場合、指定された領域内に、人物以外の背景画像が含まれる。指定部109は、エッジ検出、予め用意された背景画像との差分算出、セマンテックセグメンテーション、インスタンスセグメンテーション等の任意の手段を用いて、人物領域と背景領域とを切り分け、人物のみが存在する未検出人物領域を得ることができる。

20

【0037】

「未検出人物領域指定処理例2」

指定部109は、画像取得部101が取得した解析対象画像に対して、人物検出処理を行い、人物が存在する領域を検出する。人物検出処理は、骨格構造検出部102による2次元骨格構造の検出と異なるアルゴリズム（例えば、顔を検出するアルゴリズム）で、画像内の人物を検出する処理である。当該人物検出処理で採用するアルゴリズムは特段制限されず、周知のあらゆる技術を採用できる。例えば、セマンテックセグメンテーションやインスタンスセグメンテーション等の任意の手段を用いて、人物のみの領域を検出してもよいし、人物及び背景を含む所定形状（矩形）の領域を検出してもよい。

30

【0038】

次いで、指定部109は、当該人物検出処理で検出された人物が存在する領域の画像内の位置と、骨格構造検出部102により検出された2次元骨格構造各々の画像内の位置とを照合し、人物検出処理で検出された領域のうち、その領域内で2次元骨格構造の検出に成功していない領域を特定する。

40

【0039】

そして、指定部109は、特定した領域に基づき、未検出人物領域を指定する。上記人物検出処理で人物のみの領域が検出されている場合、指定部109は、特定した領域を、未検出人物領域として指定する。一方、上記人物検出処理で人物及び背景を含む所定形状（矩形）の領域が検出されている場合、指定部109は、エッジ検出、予め用意された背景画像との差分算出、セマンテックセグメンテーション、インスタンスセグメンテーション等の任意の手段を用いて、人物領域と背景領域とを切り分け、人物のみが存在する未検出人物領域を得ることができる。

50

【 0 0 4 0 】

補正部 1 1 1 は、解析対象画像に対し未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する。単純化する処理は、例えば、未検出人物領域以外の領域を所定の単色画像（例：白、青等）にする処理であってもよいし、未検出人物領域以外の領域を所定のパターン画像（例：ドットパターン、千鳥格子パターン等）にする処理であってもよいし、その他であってもよい。図 4 6 に、図 4 5 の解析対象画像に対して上記単純化する処理を行うことで生成された補正後画像の一例を示す。図 4 6 に示す例では、人物 P 4 が存在する領域（未検出人物領域）以外の領域が、白の単色画像になっている。未検出人物領域以外の領域の情報を単純化することで、当該単純化を行う前に比べて、未検出人物領域に存在する人物の 2 次元骨格構造が検出されやすくなる。

10

【 0 0 4 1 】

再検出部 1 1 2 は、補正後画像に含まれる人物の 2 次元骨格構造を検出する処理を骨格構造検出部 1 0 2 に実行させる。骨格構造検出部 1 0 2 は、解析対象画像に対する処理と同じアルゴリズムで補正後画像を処理し、補正後画像に含まれる人物の 2 次元骨格構造を検出する。

【 0 0 4 2 】

特徴量算出部 1 0 3 は、検出された 2 次元骨格構造の特徴量を算出し、算出した特徴量を、その 2 次元骨格構造が検出された解析対象画像に紐付けてデータベース 1 1 0 に格納する。骨格構造の特徴量は、人物の骨格の特徴を示しており、人物の骨格に基づいて人物の状態を分類や検索するための要素となる。通常、この特徴量は、複数のパラメータ（例えば後述する分類要素）を含んでいる。特徴量は、骨格構造の全体の特徴量でもよいし、骨格構造の一部の特徴量でもよく、骨格構造の各部のように複数の特徴量を含んでもよい。特徴量の算出方法は、機械学習や正規化等の任意の方法でよく、正規化として最小値や最大値を求めてもよい。一例として、特徴量は、骨格構造を機械学習することで得られた特徴量や、骨格構造の頭部から足部までの画像上の大きさ等である。骨格構造の大きさは、画像上の骨格構造を含む骨格領域の上下方向の高さや面積等である。上下方向（高さ方向または縦方向）は、画像における上下の方向（Y 軸方向）であり、例えば、地面（基準面）に対し垂直な方向である。また、左右方向（横方向）は、画像における左右の方向（X 軸方向）であり、例えば、地面に対し平行な方向である。

20

【 0 0 4 3 】

なお、ユーザが望む分類や検索を行うためには、分類や検索処理に対しロバスト性を有する特徴量を用いることが好ましい。例えば、ユーザが、人物の向きや体型に依存しない分類や検索を望む場合、人物の向きや体型にロバストな特徴量を使用してもよい。同じ姿勢で様々な方向に向いている人物の骨格や同じ姿勢で様々な体型の人物の骨格を学習することや、骨格の上下方向のみの特徴を抽出することで、人物の向きや体型に依存しない特徴量を得ることができる。

30

【 0 0 4 4 】

< 分類処理 >

分類処理は、データ蓄積処理でデータベース 1 1 0 に格納されたデータ（解析対象画像と、各解析対象画像から検出された 2 次元骨格構造の特徴量とを紐付けたデータ）に基づき、解析対象画像から検出された複数の 2 次元骨格構造をその特徴量が類似するもの同士でまとめて分類（グループ分け）する処理である。なお、解析対象画像と、各解析対象画像から検出された 2 次元骨格構造とは互いに紐付けられている。このため、分類処理による複数の 2 次元骨格構造の分類は、複数の解析対象画像の分類にもなる。分類処理により、複数の解析対象画像は、類似する 2 次元骨格構造を含むもの同士でまとめられる。以下、分類処理に関わる機能部の構成を説明する。

40

【 0 0 4 5 】

分類部 1 0 4 は、データベース 1 1 0 に格納された複数の骨格構造を、骨格構造の特徴量の類似度に基づいて分類する（クラスタリングする）。分類部 1 0 4 は、人物の状態の認識処理として、骨格構造の特徴量に基づいて複数の人物の状態を分類しているとも言え

50

る。類似度は、骨格構造の特徴量間の距離である。分類部 104 は、骨格構造の全体の特徴量の類似度により分類してもよいし、骨格構造の一部の特徴量の類似度により分類してもよく、骨格構造の第 1 の部分（例えば両手）及び第 2 の部分（例えば両足）の特徴量の類似度により分類してもよい。なお、各画像における人物の骨格構造の特徴量に基づいて人物の姿勢を分類してもよいし、時系列に連続する複数の画像における人物の骨格構造の特徴量の変化に基づいて人物の行動を分類してもよい。すなわち、分類部 104 は、骨格構造の特徴量に基づいて人物の姿勢や行動を含む人物の状態を分類できる。例えば、分類部 104 は、所定の監視期間に撮像された複数の画像における複数の骨格構造を分類対象とする。分類部 104 は、分類対象の特徴量間の類似度を求め、類似度の高い骨格構造が同じクラス（似た姿勢のグループ）となるように分類する。なお、検索と同様に、分類条件をユーザが指定できるようにしてもよい。分類部 104 は、骨格構造の分類結果をデータベース 110 に格納するとともに、表示部 107 に表示することができる。

10

【0046】

< 検索処理 >

検索処理は、データ蓄積処理でデータベース 110 に格納されたデータ（解析対象画像と、各解析対象画像から検出された 2 次元骨格構造の特徴量とを紐付けたデータ）に基づき、解析対象画像から検出された複数の 2 次元骨格構造の中から所定の骨格構造を検索する処理である。なお、解析対象画像と、各解析対象画像から検出された 2 次元骨格構造とは互いに紐付けられている。このため、上記「解析対象画像から検出された複数の 2 次元骨格構造の中から所定の骨格構造を検索する処理」により、所定の骨格構造を含む解析対象画像を検索することができる。

20

【0047】

本実施形態の検索処理では、画像処理装置 100 は、1 つ又は複数のクエリ画像の候補を取得すると、各々の評価値を算出する。この評価値は、コンピュータ処理により人物の状態の特徴部分がよく検出できる画像であるか否かの指標となる。そして、画像処理装置 100 は、このような評価値に基づきクエリ画像の候補の中からクエリ画像を選択し、選択したクエリ画像に基づき検索を行う。このような画像処理装置 100 によれば、検索用に好ましい画像をクエリ画像として選択することが可能となる。そして、所定の状態の人物を含む画像を高精度に検索できるようになる。以下、検索処理に関わる機能部の構成を説明する。

30

【0048】

画像取得部 101 は、クエリ画像の候補を取得する。画像取得部 101 は、例えば以下の取得例のいずれかにより、クエリ画像の候補を取得することができる。

【0049】

「取得例 1」

当該例では、画像取得部 101 は、解析対象画像の中のいずれかをクエリ画像の候補として取得する。例えば、画像取得部 101 は、上記分類処理で生成した複数のグループ各々から所定のルールで所定数の解析対象画像を選択し、選択した解析対象画像をクエリ画像の候補としてもよい。

【0050】

「取得例 2」

当該例では、画像取得部 101 は、ユーザが用意し、画像処理装置 100 に入力した画像をクエリ画像の候補として取得する。

40

【0051】

「取得例 3」

当該例では、画像取得部 101 は、ユーザが指定したキーワードで検索した画像をクエリ画像の候補として取得する。キーワードは、「座る」、「立つ」等のように、人物の状態（姿勢、行動等）に関する内容が想定される。キーワードの入力は、例えばテキストボックス、ドロップダウン、チェックボックス等の周知の GUI を利用して実現できる。

【0052】

50

例えば、予め、図39に示すように、クエリ画像として利用されるために用意された画像（以下、「クエリ用の画像」と、キーワード（各画像に含まれる人物の状態を示すワード）とを紐付けた情報が、データベース110に登録されてもよい。そして、画像取得部101は、入力されたキーワードが紐付けられたクエリ用の画像を当該情報の中から検索し、検索結果に含まれるクエリ用の画像の一部又は全部をクエリ画像の候補として取得してもよい。

【0053】

その他、図40に示すように、解析対象画像の一部と、キーワード（各画像に含まれる人物の状態を示すワード）とを紐付けた情報が、データベース110に登録されてもよい。そして、画像取得部101は、入力されたキーワードが紐付けられた解析対象画像を当該情報の中から検索し、検索結果に含まれる解析対象画像の一部又は全部をクエリ画像の候補として取得してもよい。

10

【0054】

その他、画像取得部101は、キーワードに関連する画像を検索する検索エンジンに、入力されたキーワードを送信し、当該検索エンジンから検索結果を取得してもよい。そして、画像取得部101は、検索結果に含まれる画像の一部又は全部をクエリ画像の候補として取得してもよい。

【0055】

骨格構造検出部102は、クエリ画像の候補に含まれる人物の2次元骨格構造を検出する。骨格構造検出部102の当該検出処理は、データ蓄積処理で説明した通りである。

20

【0056】

検索部105は、骨格構造検出部102による2次元骨格構造の検出結果に基づき、クエリ画像の候補の評価値を算出する。検索部105は、検出されたキーポイントの数、及び検出されたキーポイント各々の信頼度（検出結果の信頼度）の少なくとも一方に基づき、評価値を算出することができる。また、検索部105は、画像内における人物の大きさに基づき、評価値を算出することができる。

【0057】

評価値算出の詳細なアルゴリズムは特段制限されないが、以下の内容を満たすように設計される。

- ・検出されたキーポイントの数が多いほど、評価値が高い。
- ・検出されたキーポイントの信頼度が高いほど、評価値が高い。
- ・画像内における人物が大きいほど評価値が高い。

30

【0058】

なお、検索部105は、ユーザ入力に基づき設定された複数のキーポイント各々の重み付け値に基づき、評価値を算出してもよい。この場合、評価値算出のアルゴリズムは、上記内容に加えて、さらに以下の内容を満たすように設計される。

- ・より重み付け値が高いキーポイントが検出されているほど、評価値が高い。
- ・より重み付け値が高いキーポイントが検出され、その信頼度が高いほど、評価値が高い。

【0059】

ここで、ユーザ入力に基づき複数のキーポイント各々の重み付け値を設定する処理の一例を説明する。

40

【0060】

まず、画像処理装置100は、以下の3つの処理のいずれかに基づき、複数のキーポイントの中の一部を特定する。

- ・人物の身体を示す画像上で身体の一部を枠で囲むユーザ入力を受付け、当該枠内に含まれるキーポイントを特定する。
- ・人物の身体及び身体のキーポイントを示す画像（人物の身体上に複数のキーポイントを重畳した画像）上で一部のキーポイントを指定するユーザ入力を受付け、指定されたキーポイントを特定する。
- ・人物の身体の一部をその部位の名称（例；右腕、右足、下半身、上半身等）で指定する

50

ユーザ入力を受付け、指定された部位に含まれるキーポイントを特定する。

【 0 0 6 1 】

そして、画像処理装置 1 0 0 は、特定したキーポイントの重み付け値をユーザ入力に基づき設定する。

【 0 0 6 2 】

検索部 1 0 5 は、各クエリ画像の候補の評価値に基づき、クエリ画像の候補の中からクエリ画像を選択する。例えば、検索部 1 0 5 は、図 4 1 や図 4 2 に示すように、複数のクエリ画像の候補の評価値（図中、各画像の下に表示された数値）を比較可能に出力し、複数のクエリ画像の候補の中から 1 つ又は複数のクエリ画像を指定するユーザ入力を受付ける。そして、検索部 1 0 5 は、ユーザ入力で指定された 1 つ又は複数のクエリ画像の候補をクエリ画像として選択する。なお、図 4 1 及び図 4 2 の例では、複数のクエリ画像の候補の評価値を一覧表示することで、それらが比較可能になっているが、画面に表示するクエリ画像の候補を順に切り替えることで、それらを比較可能にしてもよい。そして、検索部 1 0 5 は、評価値が高い順にクエリ画像の候補を表示してもよい。例えば、検索部 1 0 5 は、評価値の順に並べて、クエリ画像の候補及びその評価値を一覧表示してもよい。また、検索部 1 0 5 は、クエリ画像の候補及びその評価値を、評価値の順に順次表示してもよい。なお、図 4 1 や図 4 2 に示す例では複数のクエリ画像の候補の評価値が表示されているが、画像取得部 1 0 1 が 1 つのクエリ画像の候補を取得した場合、その 1 つのクエリ画像の候補の評価値が表示される。

10

【 0 0 6 3 】

図 4 1 は、上述したキーワードでクエリ画像の候補を検索した場合の例であり、検索結果として、キーワードでヒットした画像（クエリ画像の候補）と、各々の評価値（各画像に紐付いた数値）が表示されている。

20

【 0 0 6 4 】

図 4 2 は、解析対象画像の一部をクエリ画像の候補として取得する場合の例である。上記分類処理で生成した複数のグループ各々から所定のルールで選択された所定数の解析対象画像が、クエリ画像の候補として表示されている。複数のクエリ画像の候補は、グループ毎にまとめて表示されている。そして、複数のクエリ画像の候補各々の評価値（各画像に紐付いた数値）が表示されている。

【 0 0 6 5 】

なお、図 4 2 に示すように検出されたキーポイントを表示してもよいし、図 4 1 に示すように検出されたキーポイントを表示しなくてもよい。

30

【 0 0 6 6 】

「選択例 2」

検索部 1 0 5 は、評価値が基準（基準値以上）を満たすクエリ画像の候補を、クエリ画像として選択する。この例の場合、画像処理装置 1 0 0 がクエリ画像の候補の中からクエリ画像を自動選択するので、ユーザによる選択操作が不要となる。

【 0 0 6 7 】

特徴量算出部 1 0 3 は、検索部 1 0 5 が選択したクエリ画像から検出された 2 次元骨格構造の特徴量を算出する。特徴量算出部 1 0 3 の当該算出処理は、データ蓄積処理で説明した通りである。

40

【 0 0 6 8 】

そして、検索部 1 0 5 は、データ蓄積処理でデータベース 1 1 0 に格納された複数の骨格構造の中から、検索クエリ（クエリ状態）の特徴量（クエリ画像から検出された 2 次元骨格構造の特徴量）と類似度の高い骨格構造を検索する。

【 0 0 6 9 】

例えば、検索部 1 0 5 は、検索クエリの特徴量と、複数の解析対象画像各々から検出された骨格構造の特徴量とを照合することで、検索クエリの特徴量と類似度の高い骨格構造を検索してもよい。この構成の場合、上述した分類処理は不要となる。しかし、照合対象が複数の解析対象画像の全てとなるので、照合におけるコンピュータの処理負担が大きく

50

なる。

【0070】

そこで、検索部105は、分類処理で得られたグループ毎に2次元骨格構造の特徴量の代表を任意の手段で決定し、代表と上記検索クエリの特徴量との照合により、検索クエリの特徴量と類似度の高い骨格構造を検索してもよい。この構成の場合、照合対象の数が少なくなるので、照合におけるコンピュータの処理負担が小さくなる。

【0071】

なお、解析対象画像と、各解析対象画像から検出された2次元骨格構造とは互いに紐付けられている。このため、上記「解析対象画像から検出された複数の2次元骨格構造の中から所定の骨格構造を検索する処理」により、所定の骨格構造（検索クエリの特徴量と類似度の高い骨格構造）を含む解析対象画像を検索することができる。すなわち、解析対象画像の中から、クエリ画像に含まれる人物の状態と類似する状態の人物を含む解析対象画像を検索することができる。

10

【0072】

類似度は、骨格構造の特徴量間の距離である。検索部105は、骨格構造の全体の特徴量の類似度により検索してもよいし、骨格構造の一部の特徴量の類似度により検索してもよく、骨格構造の第1の部分（例えば両手）及び第2の部分（例えば両足）の特徴量の類似度により検索してもよい。なお、各画像における人物の骨格構造の特徴量に基づいて人物の姿勢を検索してもよいし、時系列に連続する複数の画像における人物の骨格構造の特徴量の変化に基づいて人物の行動を検索してもよい。すなわち、検索部105は、骨格構造の特徴量に基づいて人物の姿勢や行動を含む人物の状態を検索できる。例えば、検索部105は、所定の監視期間に撮像された複数の解析対象画像における複数の骨格構造の特徴量を検索対象とする。

20

【0073】

入力部106は、画像処理装置100を操作するユーザから入力された情報を取得する入力インターフェイスである。例えば、ユーザは、監視カメラの画像から不審な状態の人物を監視する監視者である。入力部106は、例えば、GUI（Graphical User Interface）であり、キーボード、マウス、タッチパネル、マイク、物理ボタン等の入力装置から、ユーザの操作に応じた情報が入力される。

【0074】

表示部107は、画像処理装置100の動作（処理）の結果等を表示する表示部であり、例えば、液晶ディスプレイや有機EL（Electro Luminescence）ディスプレイ等のディスプレイ装置である。表示部107は、分類部104の分類結果、検索部105の検索結果、上述したクエリ画像の候補の評価値等を表示する。

30

【0075】

次に、画像処理装置100のハードウェア構成の一例を説明する。画像処理装置100の各機能部は、任意のコンピュータのCPU（Central Processing Unit）、メモリ、メモリにロードされるプログラム、そのプログラムを格納するハードディスク等の記憶ユニット（あらかじめ装置を出荷する段階から格納されているプログラムのほか、CD（Compact Disc）等の記憶媒体やインターネット上のサーバ等からダウンロードされたプログラムをも格納できる）、ネットワーク接続用インターフェイスを中心にハードウェアとソフトウェアの任意の組合せによって実現される。そして、その実現方法、装置にはいろいろな変形例があることは、当業者には理解されるところである。

40

【0076】

図43は、画像処理装置100のハードウェア構成を例示するブロック図である。図43に示すように、画像処理装置100は、プロセッサ1A、メモリ2A、入出力インターフェイス3A、周辺回路4A、バス5Aを有する。周辺回路4Aには、様々なモジュールが含まれる。画像処理装置100は周辺回路4Aを有さなくてもよい。なお、画像処理装置100は物理的及び/又は論理的に分かれた複数の装置で構成されてもよいし、物理的及び/又は論理的に一体となった1つの装置で構成されてもよい。画像処理装置100が

50

物理的及び/又は論理的に分かれた複数の装置で構成される場合、複数の装置各々が上記ハードウェア構成を備えることができる。

【0077】

バス5Aは、プロセッサ1A、メモリ2A、周辺回路4A及び入出力インターフェイス3Aが相互にデータを送受信するためのデータ伝送路である。プロセッサ1Aは、例えばCPU、GPU(Graphics Processing Unit)などの演算処理装置である。メモリ2Aは、例えばRAM(Random Access Memory)やROM(Read Only Memory)などのメモリである。入出力インターフェイス3Aは、入力装置、外部装置、外部サーバ、外部センサー、カメラ等から情報を取得するためのインターフェイスや、出力装置、外部装置、外部サーバ等に情報を出力するためのインターフェイスなどを含む。入力装置は、例えばキーボード、マウス、マイク、物理ボタン、タッチパネル等である。出力装置は、例えばディスプレイ、スピーカ、プリンター、メーラ等である。プロセッサ1Aは、各モジュールに指令を出し、それらの演算結果をもとに演算を行うことができる。

10

【0078】

図3~図5、図44、図47は、本実施の形態に係る画像処理装置100の処理の流れを示している。図3は、画像処理装置100における画像取得から検索処理までの流れを示し、図4は、図3の分類処理(S104)の流れを示し、図5、図44は、図3の検索処理(S105)の流れを示している。図47は、図3のS102及び図44のS322の骨格構造検出処理の流れの一例を示している。

【0079】

図3に示すように、画像取得部101は、複数の解析対象画像を取得する(S101)。続いて、骨格構造検出部102は、取得された複数の解析対象画像各々から人物の2次元骨格構造を検出する(S102)。図6は、骨格構造の検出例を示している。図6に示すように、解析対象画像には複数の人物が含まれている場合がある。この場合、骨格構造検出部102は、解析対象画像に含まれる各人物について骨格構造を検出する。

20

【0080】

図7は、このとき検出する人体モデル300の骨格構造を示しており、図8~図10は、骨格構造の検出例を示している。骨格構造検出部102は、OpenPose等の骨格推定技術を用いて、2次元の画像から図7のような人体モデル(2次元骨格モデル)300の骨格構造を検出する。人体モデル300は、人物の関節等のキーポイントと、各キーポイントを結ぶボーンから構成された2次元モデルである。

30

【0081】

骨格構造検出部102は、例えば、画像の中からキーポイントとなり得る特徴点を検出し、キーポイントの画像を機械学習した情報を参照して、人物の各キーポイントを検出する。図7の例では、人物のキーポイントとして、頭A1、首A2、右肩A31、左肩A32、右肘A41、左肘A42、右手A51、左手A52、右腰A61、左腰A62、右膝A71、左膝A72、右足A81、左足A82を検出する。さらに、これらのキーポイントを連結した人物の骨として、頭A1と首A2を結ぶボーンB1、首A2と右肩A31及び左肩A32をそれぞれ結ぶボーンB21及びボーンB22、右肩A31及び左肩A32と右肘A41及び左肘A42をそれぞれ結ぶボーンB31及びボーンB32、右肘A41及び左肘A42と右手A51及び左手A52をそれぞれ結ぶボーンB41及びボーンB42、首A2と右腰A61及び左腰A62をそれぞれ結ぶボーンB51及びボーンB52、右腰A61及び左腰A62と右膝A71及び左膝A72をそれぞれ結ぶボーンB61及びボーンB62、右膝A71及び左膝A72と右足A81及び左足A82をそれぞれ結ぶボーンB71及びボーンB72を検出する。骨格構造検出部102は、検出した人物の骨格構造をデータベース110に格納する。

40

【0082】

図8は、起立した状態の人物を検出する例である。図8では、起立した人物が正面から撮像されており、正面から見たボーンB1、ボーンB51及びボーンB52、ボーンB61及びボーンB62、ボーンB71及びボーンB72がそれぞれ重ならずに出され、右

50

足のボーン B 6 1 及びボーン B 7 1 は左足のボーン B 6 2 及びボーン B 7 2 よりも多少折れ曲がっている。

【 0 0 8 3 】

図 9 は、しゃがみ込んでいる状態の人物を検出する例である。図 9 では、しゃがみ込んでいる人物が右側から撮像されており、右側から見たボーン B 1、ボーン B 5 1 及びボーン B 5 2、ボーン B 6 1 及びボーン B 6 2、ボーン B 7 1 及びボーン B 7 2 がそれぞれ検出され、右足のボーン B 6 1 及びボーン B 7 1 と左足のボーン B 6 2 及びボーン B 7 2 は大きく折れ曲がり、かつ、重なっている。

【 0 0 8 4 】

図 10 は、寝ている状態の人物を検出する例である。図 10 では、寝ている人物が左斜め前から撮像されており、左斜め前から見たボーン B 1、ボーン B 5 1 及びボーン B 5 2、ボーン B 6 1 及びボーン B 6 2、ボーン B 7 1 及びボーン B 7 2 がそれぞれ検出され、右足のボーン B 6 1 及びボーン B 7 1 と左足のボーン B 6 2 及びボーン B 7 2 は折れ曲がり、かつ、重なっている。

【 0 0 8 5 】

続いて、図 3 に示すように、特徴量算出部 103 は、検出された骨格構造の特徴量を算出する (S 103)。例えば、骨格領域の高さや面積を特徴量とする場合、特徴量算出部 103 は、骨格構造を含む領域を検出し、その領域の高さ (画素数) や面積 (画素面積) を求める。骨格領域の高さや面積は、検出される骨格領域の端部の座標や端部のキーポイントの座標から求められる。特徴量算出部 103 は、求めた骨格構造の特徴量をデータベース 110 に格納する。なお、この骨格構造の特徴量は、人物の状態を示す情報としても用いられる。

【 0 0 8 6 】

図 8 の例では、起立した人物の骨格構造から全てのボーンを含む骨格領域を検出する。この場合、骨格領域の上端は頭部のキーポイント A 1、骨格領域の下端は左足のキーポイント A 8 2、骨格領域の左端は右肘のキーポイント A 4 1、骨格領域の右端は左手のキーポイント A 5 2 となる。このため、キーポイント A 1 とキーポイント A 8 2 の Y 座標の差分から骨格領域の高さを求める。また、キーポイント A 4 1 とキーポイント A 5 2 の X 座標の差分から骨格領域の幅を求め、骨格領域の高さと幅から面積を求める。

【 0 0 8 7 】

図 9 の例では、しゃがみ込んだ人物の骨格構造から全てのボーンを含む骨格領域を検出する。この場合、骨格領域の上端は頭部のキーポイント A 1、骨格領域の下端は右足のキーポイント A 8 1、骨格領域の左端は右腰のキーポイント A 6 1、骨格領域の右端は右手のキーポイント A 5 1 となる。このため、キーポイント A 1 とキーポイント A 8 1 の Y 座標の差分から骨格領域の高さを求める。また、キーポイント A 6 1 とキーポイント A 5 1 の X 座標の差分から骨格領域の幅を求め、骨格領域の高さと幅から面積を求める。

【 0 0 8 8 】

図 10 の例では、寝ている人物の骨格構造から全てのボーンを含む骨格領域を検出する。この場合、骨格領域の上端は左肩のキーポイント A 3 2、骨格領域の下端は左手のキーポイント A 5 2、骨格領域の左端は右手のキーポイント A 5 1、骨格領域の右端は左足のキーポイント A 8 2 となる。このため、キーポイント A 3 2 とキーポイント A 5 2 の Y 座標の差分から骨格領域の高さを求める。また、キーポイント A 5 1 とキーポイント A 8 2 の X 座標の差分から骨格領域の幅を求め、骨格領域の高さと幅から面積を求める。

【 0 0 8 9 】

続いて、図 3 に示すように、分類部 104 は、分類処理を行う (S 104)。分類処理では、図 4 に示すように、分類部 104 は、算出された骨格構造の特徴量の類似度を算出し (S 111)、算出された特徴量に基づいて骨格構造を分類する (S 112)。分類部 104 は、分類対象であるデータベース 110 に格納されている全ての骨格構造間の特徴量の類似度を求め、最も類似度が高い骨格構造 (姿勢) を同じクラスに分類する (クラスタリングする)。さらに、分類したクラス間の特徴度を求めて分類し、所定の数のク

10

20

30

40

50

ラスタとなるまで分類を繰り返す。図 1 1 は、骨格構造の特徴量の分類結果のイメージを示している。図 1 1 は、2次元の分類要素によるクラスタ分析のイメージであり、2つ分類要素は、例えば、骨格領域の高さと骨格領域の面積等である。図 1 1 では、分類の結果、複数の骨格構造の特徴量が3つのクラスタC 1 ~ C 3 に分類されている。クラスタC 1 ~ C 3 は、例えば、立っている姿勢、座っている姿勢、寝ている姿勢のように各姿勢に対応し、似ている姿勢ごとに骨格構造（人物）が分類される。

【 0 0 9 0 】

本実施の形態では、人物の骨格構造の特徴量に基づいて分類することにより、多様な分類方法を用いることができる。なお、分類方法は、予め設定されていてもよいし、ユーザが任意に設定できるようにしてもよい。また、後述する検索方法と同じ方法により分類を行ってもよい。つまり、検索条件と同様の分類条件により分類してもよい。例えば、分類部 1 0 4 は、次の分類方法により分類を行う。いずれかの分類方法を用いてもよいし、任意に選択された分類方法を組み合わせてもよい。

10

【 0 0 9 1 】

< 分類方法 1 >

複数の階層による分類全身の骨格構造による分類や、上半身や下半身の骨格構造による分類、腕や脚の骨格構造による分類等を階層的に組み合わせて分類する。すなわち、骨格構造の第 1 の部分や第 2 の部分の特徴量に基づいて分類し、さらに、第 1 の部分や第 2 の部分の特徴量に重みづけを行って分類してもよい。

【 0 0 9 2 】

< 分類方法 2 >

時系列に沿った複数枚の画像による分類時系列に連続する複数の画像における骨格構造の特徴量に基づいて分類する。例えば、時系列方向に特徴量を積み重ねて、累積値に基づいて分類してもよい。さらに、連続する複数の画像における骨格構造の特徴量の変化（変化量）に基づいて分類してもよい。

20

【 0 0 9 3 】

< 分類方法 3 >

骨格構造の左右を無視した分類人物の右側と左側が反対の骨格構造を同じ骨格構造として分類する。

【 0 0 9 4 】

さらに、分類部 1 0 4 は、骨格構造の分類結果を表示する（S 1 1 3）。分類部 1 0 4 は、データベース 1 1 0 から必要な骨格構造や人物の画像を取得し、分類結果として似ている姿勢（クラスタ）ごとに骨格構造及び人物を表示部 1 0 7 に表示する。図 1 2 は、姿勢を3つに分類した場合の表示例を示している。例えば、図 1 2 に示すように、表示ウィンドウ W 1 に、姿勢ごとの姿勢領域 W A 1 ~ W A 3 を表示し、姿勢領域 W A 1 ~ W A 3 にそれぞれ該当する姿勢の骨格構造及び人物（イメージ）を表示する。姿勢領域 W A 1 は、例えば立っている姿勢の表示領域であり、クラスタ C 1 に分類された、立っている姿勢に似た骨格構造及び人物を表示する。姿勢領域 W A 2 は、例えば座っている姿勢の表示領域であり、クラスタ C 2 に分類された、座っている姿勢に似た骨格構造及び人物を表示する。姿勢領域 W A 3 は、例えば寝ている姿勢の表示領域であり、クラスタ C 2 に分類された、寝ている姿勢に似た骨格構造及び人物を表示する。

30

【 0 0 9 5 】

続いて、図 3 に示すように、画像処理装置 1 0 0 は、検索処理を行う（S 1 0 5）。検索処理では、図 5 に示すように、画像処理装置 1 0 0 は、検索条件の入力を受け付ける（S 1 2 1）。

【 0 0 9 6 】

S 1 2 1 の処理の一例を、図 4 4 を用いて説明する。画像取得部 1 0 1 は上述した取得例 1 乃至 3 のいずれかに基づき、1つ又は複数のクエリ画像の候補を取得する（S 3 2 1）。次いで、骨格構造検出部 1 0 2 は、各クエリ画像の候補に含まれる人物の2次元骨格構造を検出する（S 3 2 2）。次いで、検索部 1 0 5 は、S 3 2 2 における2次元骨格構

40

50

造の検出結果に基づき、各クエリ画像の候補の評価値を算出する（S323）。

【0097】

次いで、検索部105は、S323で算出された評価値に基づき、クエリ画像の候補の中からクエリ画像を選択する（S324）。次いで、特徴量算出部103は、S324で選択されたクエリ画像から検出された2次元骨格構造の特徴量を算出する（S325）。クエリ画像から検出された2次元骨格構造の特徴量が、検索クエリとなる。

【0098】

図5に戻り、その後、検索部105は、検索条件に基づいて骨格構造を検索する（S122）。検索部105は、ユーザにより指定された骨格構造を検索クエリとして、検索対象であるデータベース110に格納されている骨格構造の中から特徴量の類似度が高い骨格構造を検索する。検索部105は、検索クエリの骨格構造の特徴量と検索対象の骨格構造の特徴量（解析対象画像から検出された骨格構造の特徴量）との類似度を算出し、算出した類似度が所定の閾値よりも高い骨格構造を検出する。検索クエリの骨格構造の特徴量は、予め算出された特徴量を使用してもよいし、検索時に求めた特徴量を使用してもよい。なお、検索クエリは、ユーザの操作に応じて骨格構造の各部を動かすことで入力してもよいし、ユーザがカメラの前で実演した姿勢を検索クエリとしてもよい。

【0099】

本実施の形態では、分類方法と同様に、人物の骨格構造の特徴量に基づいて検索することにより、多様な検索方法を用いることができる。なお、検索方法は、予め設定されていてもよいし、ユーザが任意に設定できるようにしてもよい。例えば、検索部105は、次の検索方法により検索を行う。いずれかの検索方法を用いてもよいし、任意に選択された検索方法を組み合わせてもよい。複数の検索方法（検索条件）を論理式（例えばAND（論理積）、OR（論理和）、NOT（否定））により組み合わせて検索してもよい。例えば、検索条件を「（右手を挙げている姿勢）AND（左足を挙げている姿勢）」として検索してもよい。

【0100】

<検索方法1>

高さ方向の特徴量のみによる検索人物の高さ方向の特徴量のみを用いて検索することで、人物の横方向の変化の影響を抑えることができ、人物の向きや人物の体型の変化に対しロバスト性が向上する。例えば、図13の骨格構造501～503のように、人物の向きや体型が異なる場合でも、高さ方向の特徴量は大きく変化しない。このため、骨格構造501～503では、検索時（分類時）に同じ姿勢であると判断することができる。

【0101】

<検索方法2>

部分検索画像において人物の体の一部が隠れている場合、認識可能な部分の情報のみを用いて検索する。例えば、図14の骨格構造511及び512のように、左足が隠れていることにより、左足のキーポイントが検出できない場合でも、検出されている他のキーポイントの特徴量を使用して検索できる。このため、骨格構造511及び512では、検索時（分類時）に同じ姿勢であると判断することができる。つまり、全てのキーポイントではなく、一部のキーポイントの特徴量を用いて、分類や検索を行うことができる。図15の骨格構造521及び522の例では、両足の向きが異なっているものの、上半身のキーポイント（A1、A2、A31、A32、A41、A42、A51、A52）の特徴量を検索クエリとすることで、同じ姿勢であると判断することができる。また、検索したい部分（特徴点）に対して、重みを付けて検索してもよいし、類似度判定の閾値を変化させてもよい。体の一部が隠れている場合、隠れた部分を無視して検索してもよいし、隠れた部分を加味して検索してもよい。隠れた部分も含めて検索することで、同じ部位が隠れているような姿勢を検索することができる。

【0102】

<検索方法3>

骨格構造の左右を無視した検索人物の右側と左側が反対の骨格構造を同じ骨格構造とし

10

20

30

40

50

て検索する。例えば、図 16 の骨格構造 5 3 1 及び 5 3 2 のように、右手を挙げている姿勢と、左手を挙げている姿勢を同じ姿勢として検索（分類）できる。図 16 の例では、骨格構造 5 3 1 と骨格構造 5 3 2 は、右手のキーポイント A 5 1、右肘のキーポイント A 4 1、左手のキーポイント A 5 2、左肘のキーポイント A 4 2 の位置が異なるものの、その他のキーポイントの位置は同じである。骨格構造 5 3 1 の右手のキーポイント A 5 1 及び右肘のキーポイント A 4 1 と骨格構造 5 3 2 の左手のキーポイント A 5 2 及び左肘のキーポイント A 4 2 のうち、一方の骨格構造のキーポイントを左右反転させると、他方の骨格構造のキーポイントと同じ位置となり、また、骨格構造 5 3 1 の左手のキーポイント A 5 2 及び左肘のキーポイント A 4 2 と骨格構造 5 3 2 の右手のキーポイント A 5 1 及び右肘のキーポイント A 4 1 のうち、一方の骨格構造のキーポイントを左右反転させると、他方の骨格構造のキーポイントと同じ位置となるため、同じ姿勢と判断する。

10

【 0 1 0 3 】

< 検索方法 4 >

縦方向と横方向の特徴量による検索人物の縦方向（Y 軸方向）の特徴量のみで検索を行った後、得られた結果をさらに人物の横方向（X 軸方向）の特徴量を用いて検索する。

【 0 1 0 4 】

< 検索方法 5 >

時系列に沿った複数枚の画像による検索時系列に連続する複数の画像における骨格構造の特徴量に基づいて検索する。例えば、時系列方向に特徴量を積み重ねて、累積値に基づいて検索してもよい。さらに、連続する複数の画像における骨格構造の特徴量の変化（変化量）に基づいて検索してもよい。

20

【 0 1 0 5 】

さらに、検索部 1 0 5 は、骨格構造の検索結果を表示する（S 1 2 3）。検索部 1 0 5 は、データベース 1 1 0 から必要な骨格構造や人物の画像を取得し、検索結果として得られた骨格構造及び人物を表示部 1 0 7 に表示する。例えば、検索クエリ（検索条件）が複数指定されている場合、検索クエリごとに検索結果を表示する。図 17 は、3 つの検索クエリ（姿勢）により検索した場合の表示例を示している。例えば、図 17 に示すように、表示ウィンドウ W 2 において、左端部に指定された検索クエリ Q 1 0、Q 2 0、Q 3 0 の骨格構造及び人物を表示し、検索クエリ Q 1 0、Q 2 0、Q 3 0 の右側に各検索クエリの検索結果 Q 1 1、Q 2 1、Q 3 1 の骨格構造及び人物を並べて表示する。

30

【 0 1 0 6 】

検索結果を検索クエリの隣から並べて表示する順番は、該当する骨格構造が見つかった順でもよいし、類似度が高い順でもよい。部分検索の部分（特徴点）に重みを付けて検索した場合に、重み付けて計算した類似度順に表示してもよい。ユーザが選択した部分（特徴点）のみから計算した類似度順に表示してもよい。また、検索結果の画像（フレーム）を中心に、時系列の前後の画像（フレーム）を一定時間分切り出して表示してもよい。

【 0 1 0 7 】

次に、図 47 を用いて、図 3 の S 1 0 2 及び図 44 の S 3 2 2 の骨格構造検出処理の流れの一例を説明する。

【 0 1 0 8 】

まず、骨格構造検出部 1 0 2 は、処理対象の画像（解析対象画像、クエリ画像の候補等）に含まれる人物の 2 次元骨格構造を検出する（S 4 2 1）。画像内に複数の人物が存在する場合、骨格構造検出部 1 0 2 は、複数の人物各々の 2 次元骨格構造を検出する。

40

【 0 1 0 9 】

次いで、指定部 1 0 9 は、例えば上述した検出漏れ判定処理例 1 又は 2 に基づき、S 4 2 1 で処理された画像の中に、2 次元骨格構造の検出に失敗した人物が存在するか判定する（S 4 2 2）。2 次元骨格構造の検出に失敗した人物が存在しない場合（S 4 2 2 の No）、S 4 2 1 での 2 次元骨格構造の検出結果が出力される（S 4 2 6）。

【 0 1 1 0 】

一方、2 次元骨格構造の検出に失敗した人物が存在する場合（S 4 2 2 の Yes）、指

50

定部 109 は、例えば上述した未検出人物領域指定処理例 1 又は 2 に基づき、2次元骨格構造の検出に失敗した人物が存在する画像内の領域である未検出人物領域を指定する (S423)。

【0111】

次いで、補正部 111 は、処理対象の画像に対し、記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する (S424)。その後、再検出部 112 は、補正後画像に含まれる人物の 2次元骨格構造を検出する処理を骨格構造検出部 102 に実行させる (S425)。そして、S421 及び S425 での 2次元骨格構造の検出結果が出力される (S426)。

【0112】

以上のように、本実施の形態では、2次元画像から人物の骨格構造を検出し、検出した骨格構造の特徴量に基づいて分類や検索を行うことを可能とした。これにより、類似度が高い似た姿勢ごとに分類することができ、また、検索クエリ (検索キー) と類似度が高い似た姿勢を検索することができる。画像から似ている姿勢を分類し表示することで、ユーザが姿勢等を指定することなく、画像中の人物の姿勢を把握することができる。分類結果の中からユーザが検索クエリの姿勢を指定できるため、予めユーザが検索したい姿勢を詳細に把握していない場合でも、所望の姿勢を検索することができる。例えば、人物の骨格構造の全体や一部等を条件として分類や検索を行うことができるため、柔軟な分類や検索が可能となる。

【0113】

また、本実施の形態では、2次元骨格構造の検出に失敗した人物が画像内に存在する場合、その人物が存在する領域 (未検出人物領域) 以外の領域の情報を単純化する補正を行い、当該補正後画像に基づき、再度 2次元骨格構造を検出する処理を行う。このような処理を行うことで、補正前の画像では 2次元骨格構造を検出できなかった人物の 2次元骨格構造を検出することが可能となる。本発明者は、当該処理は、画像内における人物の位置 (端に位置するなど) や周辺の背景画像等に起因して人物の状態を検出できない場合の改善策として有効であることを確認している。

【0114】

また、本実施の形態では、2次元骨格構造の検出に失敗した人物が存在する場合のみ、上記補正後画像の生成、及び、補正後画像に基づく、再度の 2次元骨格構造の検出を行うことができる。すなわち、補正後画像でなく補正前の画像から次元骨格構造を検出できる人物に対しては当該処理を回避できる。このため、不要に多く当該処理を実行することによるコンピュータ負担を軽減できる。

【0115】

また、本実施の形態では、コンピュータによる画像解析で、2次元骨格構造の検出に失敗した人物が存在するか否かの判定や、未検出人物領域の検出を行うことができる。このため、ユーザの負担を軽減できる。また、上述のような特徴的な手法で当該検出や判定を行うことができるので、精度よく当該検出や判定を行うことができる。

【0116】

また、本実施の形態では、画像処理装置 100 は、クエリ画像の候補を取得すると、各々の評価値を算出する。この評価値は、コンピュータ処理により人物の状態の特徴部分がよく検出できる画像であるか否かの指標となる。そして、画像処理装置 100 は、このような評価値に基づきクエリ画像の候補の中からクエリ画像を選択し、選択したクエリ画像に基づき検索を行う。このような画像処理装置 100 によれば、検索用に好ましい画像をクエリ画像として選択することが可能となる。そして、所定の状態の人物を含む画像を高精度に検索できるようになる。

【0117】

また、本実施の形態では、画像処理装置 100 は、検出されたキーポイントの数、及び検出されたキーポイント各々の信頼度の少なくとも一方に基づき、評価値を算出することができる。また、画像処理装置 100 は、画像内における人物の大きさに基づき、評価値

10

20

30

40

50

を算出することができる。このような画像処理装置 100 によれば、コンピュータ処理により人物の状態の特徴部分がよく検出できる画像であるか否かをよく表した評価値を算出することができる。

【0118】

また、本実施の形態では、画像処理装置 100 は、ユーザ入力に基づき設定された複数のキーポイント各々の重み付け値に基づき、評価値を算出することができる。このような画像処理装置 100 によれば、検索したい人物の状態の中における特に特徴的な部分が、コンピュータ処理によりよく検出できる画像であるか否かを精度よく評価することが可能となる。例えば、右手を挙げている人物を検索したい場合、右手部分が特に特徴的な部分となる。そして、右手部分に含まれるキーポイントの重み付け値が高めに設定される。

10

【0119】

また、本実施の形態では、画像処理装置 100 は、「人物の身体を示す画像上で身体の一部を枠で囲むユーザ入力を受付け、当該枠内に含まれるキーポイントを特定する処理」、「人物の身体及び身体のキーポイントを示す画像上で一部のキーポイントを指定するユーザ入力を受付け、指定されたキーポイントを特定する処理」、又は「人物の身体の一部をその部位の名称で指定するユーザ入力を受付け、指定された部位に含まれるキーポイントを特定する処理」、により一部のキーポイントを特定し、特定したキーポイントの重み付け値をユーザ入力に基づき設定することができる。このような画像処理装置 100 によれば、ユーザは、所望のキーポイントを容易に指定することが可能となる。すなわち、操作性の高いユーザフレンドリーな仕組みが実現される。

20

【0120】

また、本実施形態では、画像処理装置 100 は、クエリ画像の候補を取得すると、評価値を算出し、算出した評価値をユーザに提示することができる。そして、画像処理装置は、その後クエリ画像の候補の中からクエリ画像を選択するユーザ入力を受付け、選択されたクエリ画像に基づく検索を行うことができる。このような画像処理装置 100 によれば、ユーザは、評価値に基づき、クエリ画像の候補の中から所望のクエリ画像を容易に選択することが可能となる。特に、クエリ画像の候補が多数ある場合に、特に有益である。

【0121】

(実施の形態 2)

以下、図面を参照して実施の形態 2 について説明する。本実施の形態では、実施の形態 1 における特徴量算出の具体例について説明する。本実施の形態では、人物の身長を用いて正規化することで特徴量を求める。その他については、実施の形態 1 と同様である。

30

【0122】

図 18 は、本実施の形態に係る画像処理装置 100 の構成を示している。図 18 に示すように、画像処理装置 100 は、実施の形態 1 の構成に加えて、さらに身長算出部 108 を備える。なお、特徴量算出部 103 と身長算出部 108 を一つの処理部としてもよい。

【0123】

身長算出部(身長推定部) 108 は、骨格構造検出部 102 により検出された 2 次元骨格構造に基づき、2 次元の画像内の人物の起立時の高さ(身長画素数という)を算出(推定)する。身長画素数は、2 次元の画像における人物の身長(2 次元画像空間上の人物の全身の長さ)であるとも言える。身長算出部 108 は、検出された骨格構造の各ボーンの長さ(2 次元画像空間上の長さ)から身長画素数(ピクセル数)を求める。

40

【0124】

以下の例では、身長画素数を求める方法として具体例 1 ~ 3 を用いる。なお、具体例 1 ~ 3 のいずれかの方法を用いてもよいし、任意に選択される複数の方法を組み合わせて用いてもよい。具体例 1 では、骨格構造の各ボーンのうち、頭部から足部までのボーンの長さを合計することで、身長画素数を求める。骨格構造検出部 102 (骨格推定技術) が頭頂と足元を出力しない場合は、必要に応じて定数を乗じて補正することもできる。具体例 2 では、各ボーンの長さ(2 次元画像空間上の身長)との関係を示す人体モデルを用いて、身長画素数を算出する。具体例 3 では、3 次元人体モデルを 2 次元骨格構

50

造にフィッティング（あてはめる）することで、身長画素数を算出する。

【 0 1 2 5 】

本実施の形態の特徴量算出部 1 0 3 は、算出された人物の身長画素数に基づいて、人物の骨格構造（骨格情報）を正規化する正規化部である。特徴量算出部 1 0 3 は、正規化した骨格構造の特徴量（正規化値）をデータベース 1 1 0 に格納する。特徴量算出部 1 0 3 は、骨格構造に含まれる各キーポイント（特徴点）の画像上での高さを、身長画素数で正規化する。本実施の形態では、例えば、高さ方向は、画像の 2 次元座標（X - Y 座標）空間における上下の方向（Y 軸方向）である。この場合、キーポイントの高さは、キーポイントの Y 座標の値（画素数）から求めることができる。あるいは、高さ方向は、実世界の 3 次元座標空間における地面（基準面）に対し垂直な鉛直軸の方向を、2 次元座標空間に投影した鉛直投影軸の方向（鉛直投影方向）でもよい。この場合、キーポイントの高さは、実世界における地面に対し垂直な軸を、カメラパラメータに基づいて 2 次元座標空間に投影した鉛直投影軸を求め、この鉛直投影軸に沿った値（画素数）から求めることができる。なお、カメラパラメータは、画像の撮像パラメータであり、例えば、カメラパラメータは、カメラ 2 0 0 の姿勢、位置、撮像角度、焦点距離等である。カメラ 2 0 0 により、予め長さや位置が分かっている物体を撮像し、その画像からカメラパラメータを求めることができる。撮像された画像の両端ではひずみが発生し、実世界の鉛直方向と画像の上下方向が合わない場合がある。これに対し、画像を撮影したカメラのパラメータを使用することで、実世界の鉛直方向が画像中でどの程度傾いているのかが分かる。このため、カメラパラメータに基づいて画像中に投影した鉛直投影軸に沿ったキーポイントの値を身長で正規化することで、実世界と画像のずれを考慮してキーポイントの特徴量化することができる。なお、左右方向（横方向）は、画像の 2 次元座標（X - Y 座標）空間における左右の方向（X 軸方向）であり、または、実世界の 3 次元座標空間における地面に対し平行な方向を、2 次元座標空間に投影した方向である。

【 0 1 2 6 】

図 1 9 ~ 図 2 3 は、本実施の形態に係る画像処理装置 1 0 0 の処理の流れを示している。図 1 9 は、画像処理装置 1 0 0 における画像取得から検索処理までの流れを示し、図 2 0 ~ 図 2 2 は、図 1 9 の身長画素数算出処理（S 2 0 1）の具体例 1 ~ 3 の流れを示し、図 2 3 は、図 1 9 の正規化処理（S 2 0 2）の流れを示している。

【 0 1 2 7 】

図 1 9 に示すように、本実施の形態では、実施の形態 1 における特徴量算出処理（S 1 0 3）として、身長画素数算出処理（S 2 0 1）及び正規化処理（S 2 0 2）を行う。その他については実施の形態 1 と同様である。

【 0 1 2 8 】

画像処理装置 1 0 0 は、画像取得（S 1 0 1）及び骨格構造検出（S 1 0 2）に続いて、検出された骨格構造に基づいて身長画素数算出処理を行う（S 2 0 1）。この例では、図 2 4 に示すように、画像における直立時の人物の骨格構造の高さを身長画素数（ h ）とし、画像の人物の状態における骨格構造の各キーポイントの高さをキーポイント高さ（ y_i ）とする。以下、身長画素数算出処理の具体例 1 ~ 3 について説明する。

【 0 1 2 9 】

< 具体例 1 >

具体例 1 では、頭部から足部までのボーンの長さを用いて身長画素数を求める。具体例 1 では、図 2 0 に示すように、身長算出部 1 0 8 は、各ボーンの長さを取得し（S 2 1 1）、取得した各ボーンの長さを合計する（S 2 1 2）。

【 0 1 3 0 】

身長算出部 1 0 8 は、人物の頭部から足部の 2 次元の画像上のボーンの長さを取得し、身長画素数を求める。すなわち、骨格構造を検出した画像から、図 2 4 のボーンのうち、ボーン B 1（長さ L_1 ）、ボーン B 5 1（長さ L_{21} ）、ボーン B 6 1（長さ L_{31} ）及びボーン B 7 1（長さ L_{41} ）、もしくは、ボーン B 1（長さ L_1 ）、ボーン B 5 2（長さ L_{22} ）、ボーン B 6 2（長さ L_{32} ）及びボーン B 7 2（長さ L_{42} ）の各長さ（画

10

20

30

40

50

素数)を取得する。各ボーンの長さは、2次元の画像における各キーポイントの座標から求めることができる。これらを合計した、 $L1 + L21 + L31 + L41$ 、もしくは、 $L1 + L22 + L32 + L42$ に補正定数を乗じた値を身長画素数(h)として算出する。両方の値を算出できる場合、例えば、長い方の値を身長画素数とする。すなわち、各ボーンは正面から撮像された場合が画像中での長さが最も長くなり、カメラに対して奥行き方向に傾くと短く表示される。従って、長いボーンの方が正面から撮像されている可能性が高く、真実の値に近いと考えられる。このため、長い方の値を選択することが好ましい。

【0131】

図25の例では、ボーンB1、ボーンB51及びボーンB52、ボーンB61及びボーンB62、ボーンB71及びボーンB72がそれぞれ重ならずに出検されている。これらのボーンの合計である、 $L1 + L21 + L31 + L41$ 、及び、 $L1 + L22 + L32 + L42$ を求め、例えば、検出されたボーンの長さが長い左足側の $L1 + L22 + L32 + L42$ に補正定数を乗じた値を身長画素数とする。

10

【0132】

図26の例では、ボーンB1、ボーンB51及びボーンB52、ボーンB61及びボーンB62、ボーンB71及びボーンB72がそれぞれ検出され、右足のボーンB61及びボーンB71と左足のボーンB62及びボーンB72が重なっている。これらのボーンの合計である、 $L1 + L21 + L31 + L41$ 、及び、 $L1 + L22 + L32 + L42$ を求め、例えば、検出されたボーンの長さが長い右足側の $L1 + L21 + L31 + L41$ に補正定数を乗じた値を身長画素数とする。

20

【0133】

図27の例では、ボーンB1、ボーンB51及びボーンB52、ボーンB61及びボーンB62、ボーンB71及びボーンB72がそれぞれ検出され、右足のボーンB61及びボーンB71と左足のボーンB62及びボーンB72が重なっている。これらのボーンの合計である、 $L1 + L21 + L31 + L41$ 、及び、 $L1 + L22 + L32 + L42$ を求め、例えば、検出されたボーンの長さが長い左足側の $L1 + L22 + L32 + L42$ に補正定数を乗じた値を身長画素数とする。

【0134】

具体例1では、頭から足までのボーンの長さを合計することで身長を求めることができるため、簡易な方法で身長画素数を求めることができる。また、機械学習を用いた骨格推定技術により、少なくとも頭から足までの骨格を検出できればよいため、しゃがみ込んでいる状態など、必ずしも人物の全体が画像に写っていない場合でも精度よく身長画素数を推定することができる。

30

【0135】

<具体例2>

具体例2では、2次元骨格構造に含まれる骨の長さ2次元画像空間上の人物の全身の長さとの関係を示す2次元骨格モデルを用いて身長画素数を求める。

【0136】

図28は、具体例2で用いる、2次元画像空間上の各ボーンの長さ2次元画像空間上の全身の長さとの関係を示す人体モデル(2次元骨格モデル)301である。図28に示すように、平均的な人物の各ボーンの長さ2次元画像空間上の全身の長さとの関係(全身の長さに対する各ボーンの長さの割合)を、人体モデル301の各ボーンに対応付ける。例えば、頭のボーンB1の長さは全身の長さ $\times 0.2$ (20%)であり、右手のボーンB41の長さは全身の長さ $\times 0.15$ (15%)であり、右足のボーンB71の長さは全身の長さ $\times 0.25$ (25%)である。このような人体モデル301の情報をデータベース110に記憶しておくことで、各ボーンの長さから平均的な全身の長さを求めることができる。平均的な人物の人体モデルの他に、年代、性別、国籍等の人物の属性ごとに人体モデルを用意してもよい。これにより、人物の属性に応じて適切に全身の長さ(身長)を求めることができる。

40

【0137】

具体例2では、図21に示すように、身長算出部108は、各ボーンの長さ取得する

50

(S221)。身長算出部108は、検出された骨格構造において、全てのボーンの長さ(2次元画像空間上の長さ)を取得する。図29は、しゃがみ込んでいる状態の人物を右斜め後ろから撮像し、骨格構造を検出した例である。この例では、人物の顔や左側面が写っていないことから、頭のボーンと左腕及び左手のボーンが検出できていない。このため、検出されているボーンB21、B22、B31、B41、B51、B52、B61、B62、B71、B72の各長さを取得する。

【0138】

続いて、身長算出部108は、図21に示すように、人体モデルに基づき、各ボーンの長さから身長画素数を算出する(S222)。身長算出部108は、図28のような、各ボーンと全身の長さとの関係を示す人体モデル301を参照し、各ボーンの長さから身長画素数を求める。例えば、右手のボーンB41の長さが全身の長さ $\times 0.15$ であるため、ボーンB41の長さ $/ 0.15$ によりボーンB41に基づいた身長画素数を求める。また、右足のボーンB71の長さが全身の長さ $\times 0.25$ であるため、ボーンB71の長さ $/ 0.25$ によりボーンB71に基づいた身長画素数を求める。

10

【0139】

このとき参照する人体モデルは、例えば、平均的な人物の人体モデルであるが、年代、性別、国籍等の人物の属性に応じて人体モデルを選択してもよい。例えば、撮像した画像に人物の顔が写っている場合、顔に基づいて人物の属性を識別し、識別した属性に対応する人体モデルを参照する。属性ごとの顔を機械学習した情報を参照し、画像の顔の特徴から人物の属性を認識することができる。また、画像から人物の属性が識別できない場合に、平均的な人物の人体モデルを用いてもよい。

20

【0140】

また、ボーンの長さから算出した身長画素数をカメラパラメータにより補正してもよい。例えばカメラを高い位置において、人物を見下ろすように撮影した場合、二次元骨格構造において肩幅のボーン等の横の長さはカメラの俯角の影響を受けないが、首-腰のボーン等の縦の長さは、カメラの俯角が大きくなる程小さくなる。そうすると、肩幅のボーン等の横の長さから算出した身長画素数が実際より大きくなる傾向がある。そこで、カメラパラメータを活用すると、人物がどの程度の角度でカメラに見下ろされているかがわかるため、この俯角の情報を使って正面から撮影したような二次元骨格構造に補正することができる。これによって、より正確に身長画素数を算出できる。

30

【0141】

続いて、身長算出部108は、図21に示すように、身長画素数の最適値を算出する(S223)。身長算出部108は、ボーンごとに求めた身長画素数から身長画素数の最適値を算出する。例えば、図30に示すような、ボーンごとに求めた身長画素数のヒストグラムを生成し、その中で大きい身長画素数を選択する。つまり、複数のボーンに基づいて求められた複数の身長画素数の中で他よりも長い身長画素数を選択する。例えば、上位30%を有効な値とし、図30ではボーンB71、B61、B51による身長画素数を選択する。選択した身長画素数の平均を最適値として求めてもよいし、最も大きい身長画素数を最適値としてもよい。2次元画像のボーンの長さから身長を求めるため、ボーンを正面からできていない場合、すなわち、ボーンがカメラから見て奥行き方向に傾いて撮像された場合、ボーンの長さが正面から撮像した場合よりも短くなる。そうすると、身長画素数が大きい値は、身長画素数が小さい値よりも、正面から撮像された可能性が高く、より尤もらしい値となることから、より大きい値を最適値とする。

40

【0142】

具体例2では、2次元画像空間上のボーンと全身の長さとの関係を示す人体モデルを用いて、検出した骨格構造のボーンに基づき身長画素数を求めるため、頭から足までの全ての骨格が得られない場合でも、一部のボーンから身長画素数を求めることができる。特に、複数のボーンから求められた値のうち、より大きい値を採用することで、精度よく身長画素数を推定することができる。

【0143】

50

< 具体例 3 >

具体例 3 では、2次元骨格構造を3次元人体モデル（3次元骨格モデル）にフィッティングさせて、フィッティングした3次元人体モデルの身長画素数を用いて全身の骨格ベクトルを求める。

【 0 1 4 4 】

具体例 3 では、図 2 2 に示すように、身長算出部 1 0 8 は、まず、カメラ 2 0 0 の撮像した画像に基づき、カメラパラメータを算出する（S 2 3 1）。身長算出部 1 0 8 は、カメラ 2 0 0 が撮像した複数の画像の中から、予め長さが分かっている物体を検出し、検出した物体の大きさ（画素数）からカメラパラメータを求める。なお、カメラパラメータを予め求めておき、求めておいたカメラパラメータを必要に応じて取得してもよい。

10

【 0 1 4 5 】

続いて、身長算出部 1 0 8 は、3次元人体モデルの配置及び高さを調整する（S 2 3 2）。身長算出部 1 0 8 は、検出された2次元骨格構造に対し、身長画素数算出用の3次元人体モデルを用意し、カメラパラメータに基づいて、同じ2次元画像内に配置する。具体的には、カメラパラメータと、2次元骨格構造から、「実世界におけるカメラと人物の相対的な位置関係」を特定する。例えば、仮にカメラの位置を座標（0, 0, 0）としたときに、人物が立っている（または座っている）位置の座標（x, y, z）を特定する。そして、特定した人物と同じ位置（x, y, z）に3次元人体モデルを配置して撮像した場合の画像を想定することで、2次元骨格構造と3次元人体モデルを重ね合わせる。

【 0 1 4 6 】

図 3 1 は、しゃがみ込んでいる人物を左斜め前から撮像し、2次元骨格構造 4 0 1 を検出した例である。2次元骨格構造 4 0 1 は、2次元の座標情報を有する。なお、全てのボーンを検出していることが好ましいが、一部のボーンが検出されていなくてもよい。この2次元骨格構造 4 0 1 に対し、図 3 2 のような、3次元人体モデル 4 0 2 を用意する。3次元人体モデル（3次元骨格モデル） 4 0 2 は、3次元の座標情報を有し、2次元骨格構造 4 0 1 と同じ形状の骨格のモデルである。そして、図 3 3 のように、検出した2次元骨格構造 4 0 1 に対し、用意した3次元人体モデル 4 0 2 を配置し重ね合わせる。また、重ね合わせるとともに、3次元人体モデル 4 0 2 の高さを2次元骨格構造 4 0 1 に合うように調整する。

20

【 0 1 4 7 】

なお、このとき用意する3次元人体モデル 4 0 2 は、図 3 3 のように、2次元骨格構造 4 0 1 の姿勢に近い状態のモデルでもよいし、直立した状態のモデルでもよい。例えば、機械学習を用いて2次元画像から3次元空間の姿勢を推定する技術を用いて、推定した姿勢の3次元人体モデル 4 0 2 を生成してもよい。2次元画像の関節と3次元空間の関節の情報を学習することで、2次元画像から3次元の姿勢を推定することができる。

30

【 0 1 4 8 】

続いて、身長算出部 1 0 8 は、図 2 2 に示すように、3次元人体モデルを2次元骨格構造にフィッティングする（S 2 3 3）。身長算出部 1 0 8 は、図 3 4 のように、3次元人体モデル 4 0 2 を2次元骨格構造 4 0 1 に重ね合わせた状態で、3次元人体モデル 4 0 2 と2次元骨格構造 4 0 1 の姿勢が一致するように、3次元人体モデル 4 0 2 を変形させる。すなわち、3次元人体モデル 4 0 2 の身長、体の向き、関節の角度を調整し、2次元骨格構造 4 0 1 との差異がなくなるように最適化する。例えば、3次元人体モデル 4 0 2 の関節を、人の可動範囲で回転させていき、また、3次元人体モデル 4 0 2 の全体を回転させたり、全体のサイズを調整する。なお、3次元人体モデルと2次元骨格構造のフィッティング（あてはめ）は、2次元空間（2次元座標）上で行う。すなわち、2次元空間に3次元人体モデルを写像し、変形させた3次元人体モデルが2次元空間（画像）でどのように変化するかを考慮して、3次元人体モデルを2次元骨格構造に最適化する。

40

【 0 1 4 9 】

続いて、身長算出部 1 0 8 は、図 2 2 に示すように、フィッティングさせた3次元人体モデルの身長画素数を算出する（S 2 3 4）。身長算出部 1 0 8 は、図 3 5 のように、3

50

次元人体モデル 4 0 2 と 2 次元骨格構造 4 0 1 の差異がなくなり、姿勢が一致すると、その状態の 3 次元人体モデル 4 0 2 の身長画素数を求める。最適化された 3 次元人体モデル 4 0 2 を直立させた状態として、カメラパラメータに基づき、2 次元空間上の全身の長さを求める。例えば、3 次元人体モデル 4 0 2 を直立させた場合の頭から足までのボーンの長さ（画素数）により身長画素数を算出する。具体例 1 と同様に、3 次元人体モデル 4 0 2 の頭部から足部までのボーンの長さを合計してもよい。

【 0 1 5 0 】

具体例 3 では、カメラパラメータに基づいて 3 次元人体モデルを 2 次元骨格構造にフィッティングさせて、その 3 次元人体モデルに基づいて身長画素数を求めることで、全てのボーンが正面に写っていない場合、すなわち、全てのボーンが斜めに映っているため誤差

10

【 0 1 5 1 】

< 正規化処理 >

図 1 9 に示すように、画像処理装置 1 0 0 は、身長画素数算出処理に続いて、正規化処理 (S 2 0 2) を行う。正規化処理では、図 2 3 に示すように、特徴量算出部 1 0 3 は、キーポイント高さを算出する (S 2 4 1)。特徴量算出部 1 0 3 は、検出された骨格構造に含まれる全てのキーポイントのキーポイント高さ（画素数）を算出する。キーポイント高さは、骨格構造の最下端（例えばいずれかの足のキーポイント）からそのキーポイントまでの高さ方向の長さ（画素数）である。ここでは、一例として、キーポイント高さを、画像におけるキーポイントの Y 座標から求める。なお、上記のように、キーポイント高さは、カメラパラメータに基づいた鉛直投影軸に沿った方向の長さから求めてもよい。例えば、図 2 4 の例で、首のキーポイント A 2 の高さ (y_i) は、キーポイント A 2 の Y 座標から右足のキーポイント A 8 1 または左足のキーポイント A 8 2 の Y 座標を引いた値である。

20

【 0 1 5 2 】

続いて、特徴量算出部 1 0 3 は、正規化のための基準点を特定する (S 2 4 2)。基準点は、キーポイントの相対的な高さを表すための基準となる点である。基準点は、予め設定されていてもよいし、ユーザが選択できるようにしてもよい。基準点は、骨格構造の中心もしくは中心よりも高い（画像の上下方向における上である）ことが好ましく、例えば、首のキーポイントの座標を基準点とする。なお、首に限らず頭やその他のキーポイントの座標を基準点としてもよい。キーポイントに限らず、任意の座標（例えば骨格構造の中心座標等）を基準点としてもよい。

30

【 0 1 5 3 】

続いて、特徴量算出部 1 0 3 は、キーポイント高さ (y_i) を身長画素数で正規化する (S 2 4 3)。特徴量算出部 1 0 3 は、各キーポイントのキーポイント高さ、基準点、身長画素数を用いて、各キーポイントを正規化する。具体的には、特徴量算出部 1 0 3 は、基準点に対するキーポイントの相対的な高さを身長画素数により正規化する。ここでは、高さ方向のみに着目する例として、Y 座標のみを検出し、また、基準点を首のキーポイントとして正規化を行う。具体的には、基準点（首のキーポイント）の Y 座標を (y_c) として、次の式 (1) を用いて、特徴量（正規化値）を求める。なお、カメラパラメータに基づいた鉛直投影軸を用いる場合は、(y_i) 及び (y_c) を鉛直投影軸に沿った方向の値に変換する。

40

【 数 1 】

$$f_i = (y_i - y_c) / h \quad \dots (1)$$

【 0 1 5 4 】

例えば、キーポイントが 1 8 個の場合、各キーポイントの 1 8 点の座標 (x_0 、 y_0)、(x_1 、 y_1)、 \dots (x_{17} 、 y_{17}) を、上記式 (1) を用いて、次のように 1

50

8次元の特徴量に変換する。

【数 2】

$$\begin{aligned} f_0 &= (y_0 - y_c) / h \\ f_1 &= (y_1 - y_c) / h \\ &\vdots \\ f_{17} &= (y_{17} - y_c) / h \end{aligned} \quad \dots (2)$$

10

【0155】

図36は、特徴量算出部103が求めた各キーポイントの特徴量の例を示している。この例では、首のキーポイントA2を基準点とするため、キーポイントA2の特徴量は0.0となり、首と同じ高さの右肩のキーポイントA31及び左肩のキーポイントA32の特徴量も0.0である。首よりも高い頭のキーポイントA1の特徴量は-0.2である。首よりも低い右手のキーポイントA51及び左手のキーポイントA52の特徴量は0.4であり、右足のキーポイントA81及び左足のキーポイントA82の特徴量は0.9である。この状態から人物が左手を挙げると、図37のように左手が基準点よりも高くなるため、左手のキーポイントA52の特徴量は-0.4となる。一方で、Y軸の座標のみを用いて正規化を行っているため、図38のように、図36に比べて骨格構造の幅が変わっても特徴量は変わらない。すなわち、本実施の形態の特徴量(正規化値)は、骨格構造(キーポイント)の高さ方向(Y方向)の特徴を示しており、骨格構造の横方向(X方向)の変化に影響を受けない。

20

【0156】

以上のように、本実施の形態では、2次元画像から人物の骨格構造を検出し、検出した骨格構造から求めた身長画素数(2次元画像空間上の直立時の高さ)を用いて、骨格構造の各キーポイントを正規化する。この正規化された特徴量を用いることで、分類や検索等を行った場合のロバスト性を向上することができる。すなわち、本実施の形態の特徴量は、上記のように人物の横方向の変化に影響を受けないため、人物の向きや人物の体型の変化に対しロバスト性が高い。

30

【0157】

さらに、本実施の形態では、OpenPose等の骨格推定技術を用いて人物の骨格構造を検出することで実現できるため、人物の姿勢等を学習する学習データを用意する必要がない。また、骨格構造のキーポイントを正規化し、データベースに格納しておくことで、人物の姿勢等の分類や検索が可能となるため、未知な姿勢に対しても分類や検索を行うことができる。また、骨格構造のキーポイントを正規化することで、明確でわかりやすい特徴量を得ることができるため、機械学習のようにブラックボックス型のアルゴリズムと異なり、処理結果に対するユーザの納得性が高い。

【0158】

以上、図面を参照して本発明の実施形態について述べたが、これらは本発明の例示であり、上記以外の様々な構成を採用することもできる。

40

【0159】

また、上述の説明で用いた複数のフローチャートでは、複数の工程(処理)が順番に記載されているが、各実施形態で実行される工程の実行順序は、その記載の順番に制限されない。各実施形態では、図示される工程の順番を内容的に支障のない範囲で変更することができる。また、上述の各実施形態は、内容が相反しない範囲で組み合わせることができる。

【0160】

上記の実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限定されない。

50

1. 画像を取得する画像取得手段と、

前記画像に含まれる人物の2次元骨格構造を検出する骨格構造検出手段と、

前記2次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段と、

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段と、

前記補正後画像に含まれる人物の前記2次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段と、

を有する画像処理装置。

2. 前記指定手段は、ユーザ入力に基づき、又は、画像解析で、前記未検出人物領域を指定する1に記載の画像処理装置。

10

3. 前記指定手段は、

前記2次元骨格構造の検出を行った前記画像の中に、前記2次元骨格構造の検出に失敗した人物が存在するか判定し、

前記2次元骨格構造の検出に失敗した人物が存在すると判定した前記画像に対し、前記未検出人物領域を指定する処理を実行する1に記載の画像処理装置。

4. 前記指定手段は、ユーザ入力に基づき、又は、画像解析で、前記2次元骨格構造の検出を行った前記画像の中に、前記2次元骨格構造の検出に失敗した人物が存在するか判定する3に記載の画像処理装置。

5. 前記指定手段は、

20

前記画像に含まれる人物を検出し、

検出された人物の数が、前記2次元骨格構造の検出に成功した人物の数より多い場合、前記2次元骨格構造の検出に失敗した人物が存在すると判定する4に記載の画像処理装置。

6. 検出された前記2次元骨格構造の特徴量を算出する特徴量算出手段と、

算出された前記特徴量の類似度に基づき、前記画像に含まれる人物の状態の認識処理を行う認識手段と、

を有する1から5のいずれかに記載の画像処理装置。

7. コンピュータが、

画像を取得し、

30

前記画像に含まれる人物の2次元骨格構造を検出し、

前記2次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定し、

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成し、

前記補正後画像に含まれる人物の前記2次元骨格構造を検出する処理を実行する画像処理方法。

8. コンピュータを、

画像を取得する画像取得手段、

前記画像に含まれる人物の2次元骨格構造を検出する骨格構造検出手段、

前記2次元骨格構造の検出に失敗した人物が存在する前記画像内の領域である未検出人物領域を指定する指定手段、

40

前記未検出人物領域以外の領域の情報を単純化する処理を行い、補正後画像を生成する補正手段、

前記補正後画像に含まれる人物の前記2次元骨格構造を検出する処理を前記骨格構造検出手段に実行させる再検出手段、

として機能させるプログラム。

【符号の説明】

【0161】

1 画像処理システム

10 画像処理装置

50

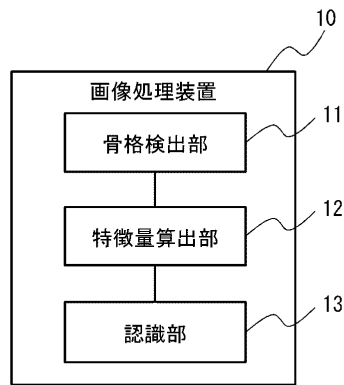
- 1 1 骨格検出部
- 1 2 特徴量算出部
- 1 3 認識部
- 1 0 0 画像処理装置
- 1 0 1 画像取得部
- 1 0 2 骨格構造検出部
- 1 0 3 特徴量算出部
- 1 0 4 分類部
- 1 0 5 検索部
- 1 0 6 入力部
- 1 0 7 表示部
- 1 0 8 身長算出部
- 1 0 9 指定部
- 1 1 0 データベース
- 1 1 1 補正部
- 1 1 2 再検出部
- 2 0 0 カメラ
- 3 0 0、3 0 1 人体モデル
- 4 0 1 2次元骨格構造
- 4 0 2 3次元人体モデル

10

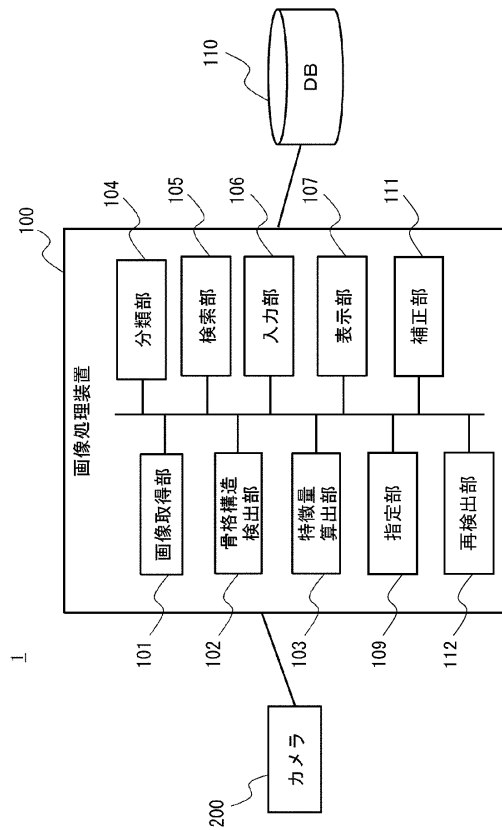
20

【図面】

【図 1】



【図 2】

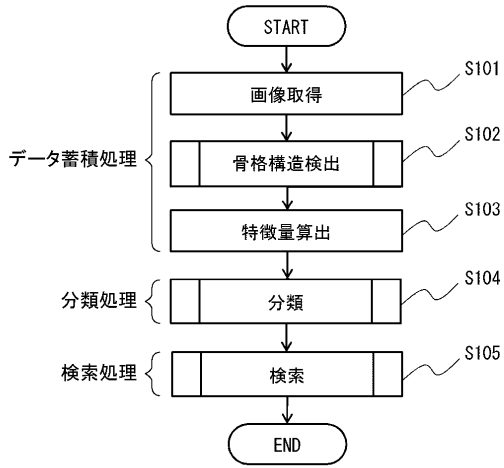


30

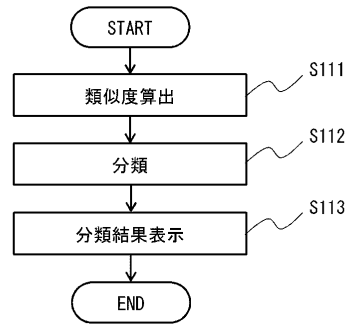
40

50

【 図 3 】

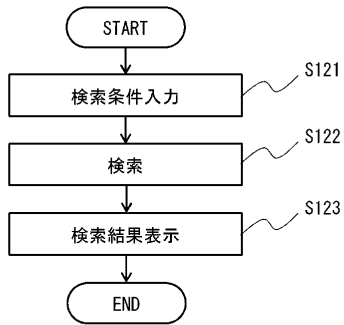


【 図 4 】

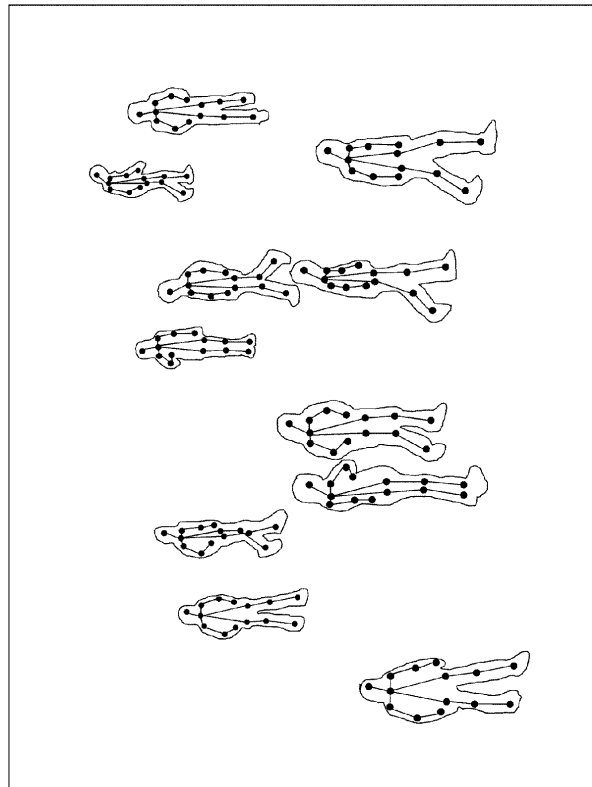


10

【 図 5 】



【 図 6 】



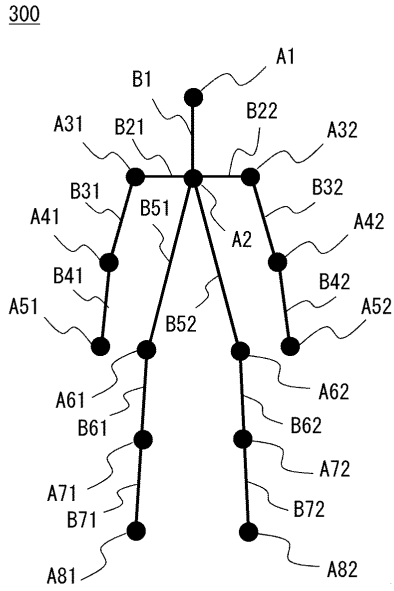
20

30

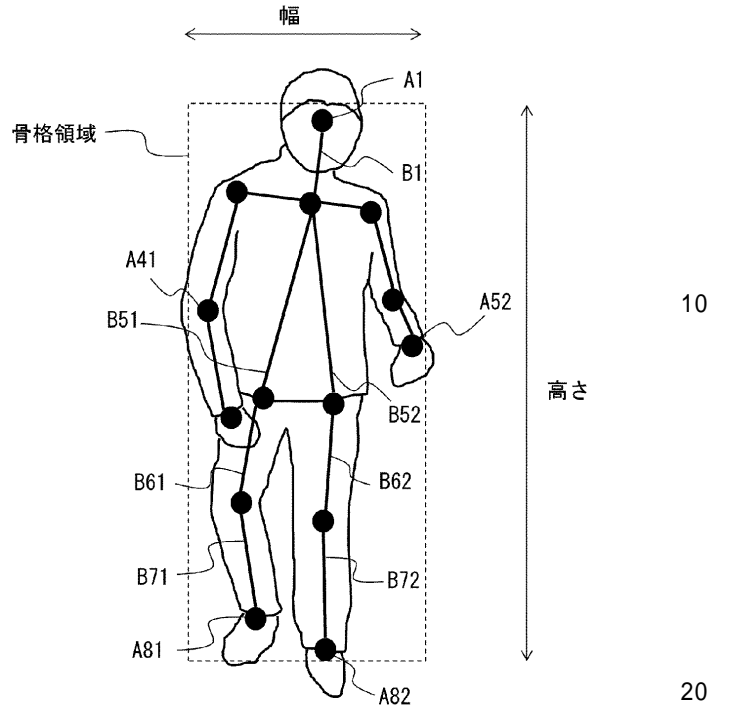
40

50

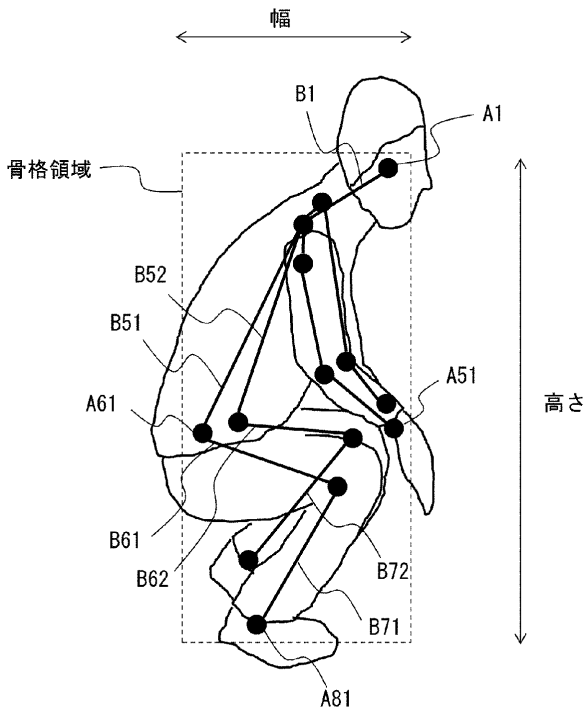
【 図 7 】



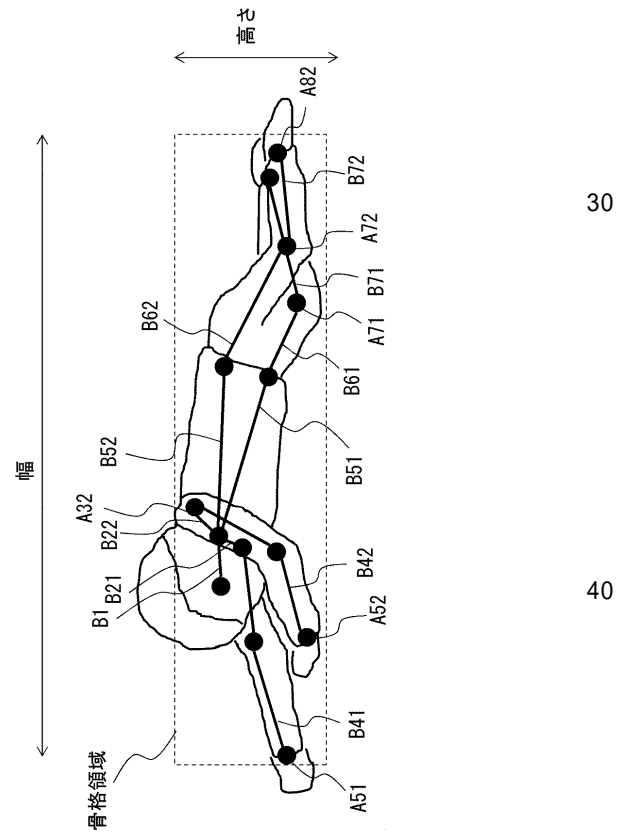
【 図 8 】



【 図 9 】



【 図 10 】



10

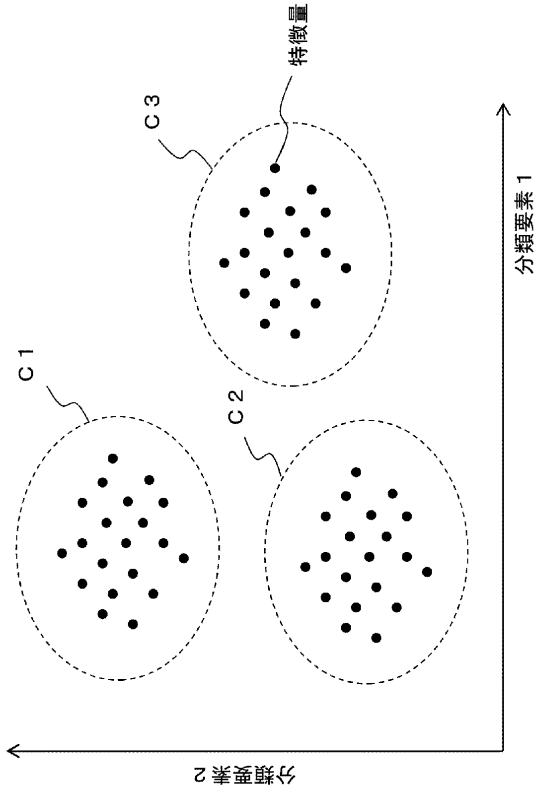
20

30

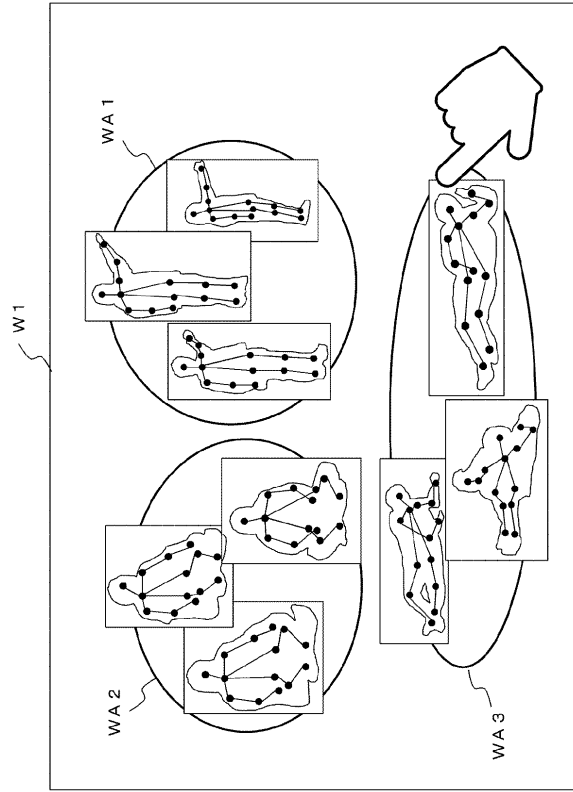
40

50

【図 1 1】



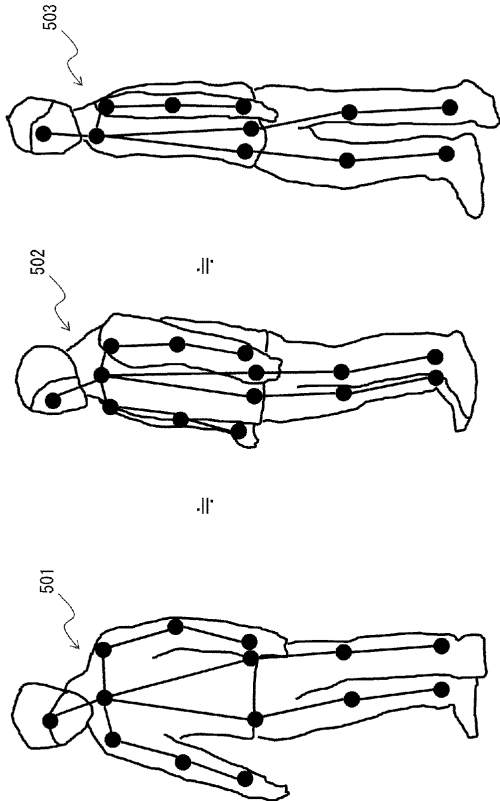
【図 1 2】



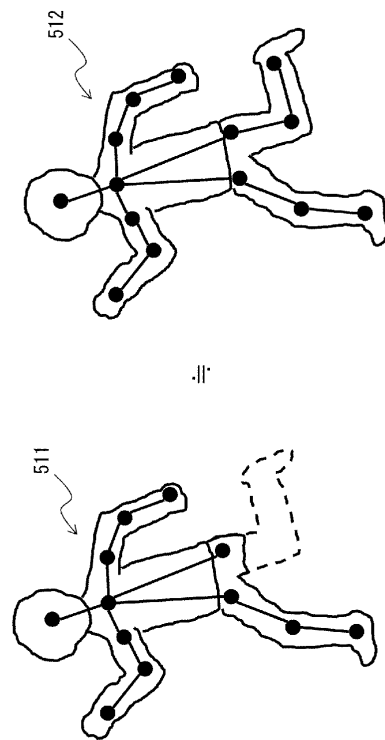
10

20

【図 1 3】



【図 1 4】

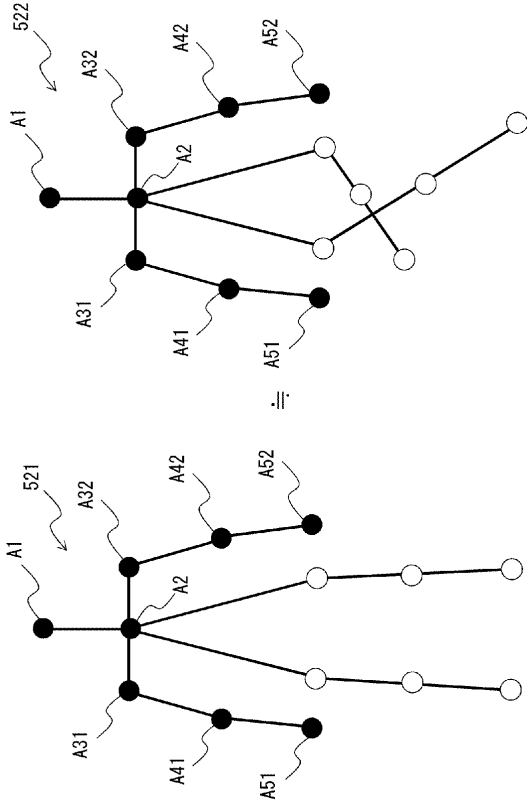


30

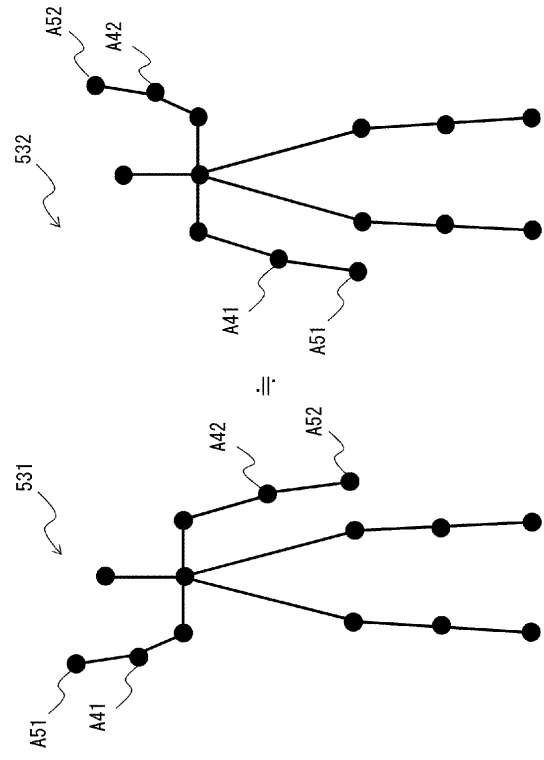
40

50

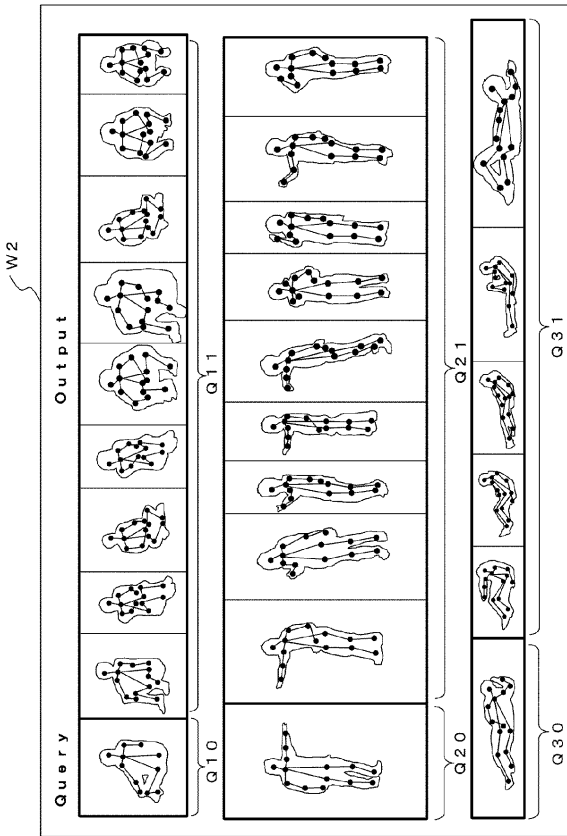
【図 15】



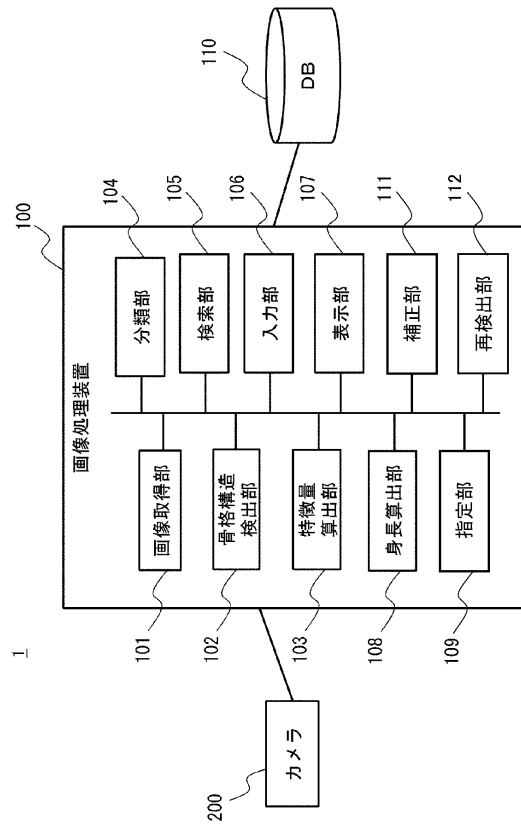
【図 16】



【図 17】



【図 18】



10

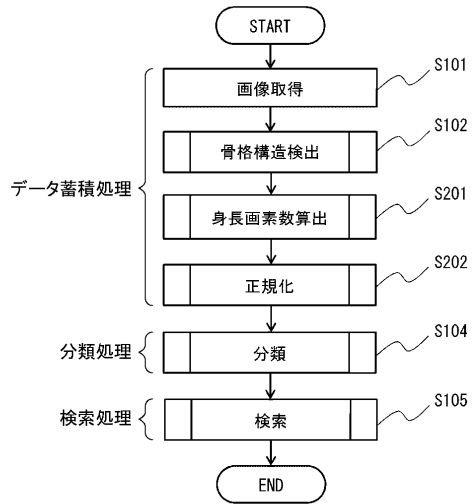
20

30

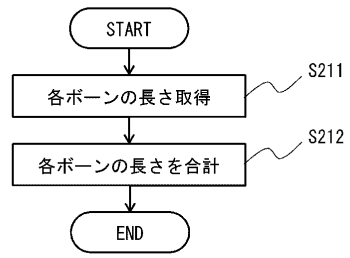
40

50

【図 19】

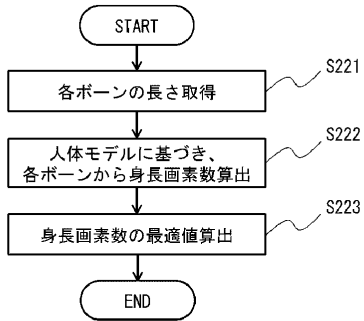


【図 20】

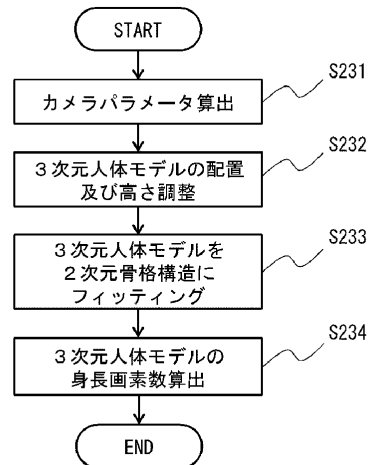


10

【図 21】



【図 22】



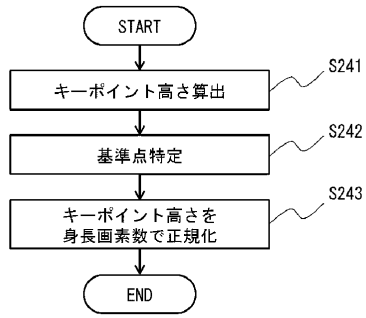
20

30

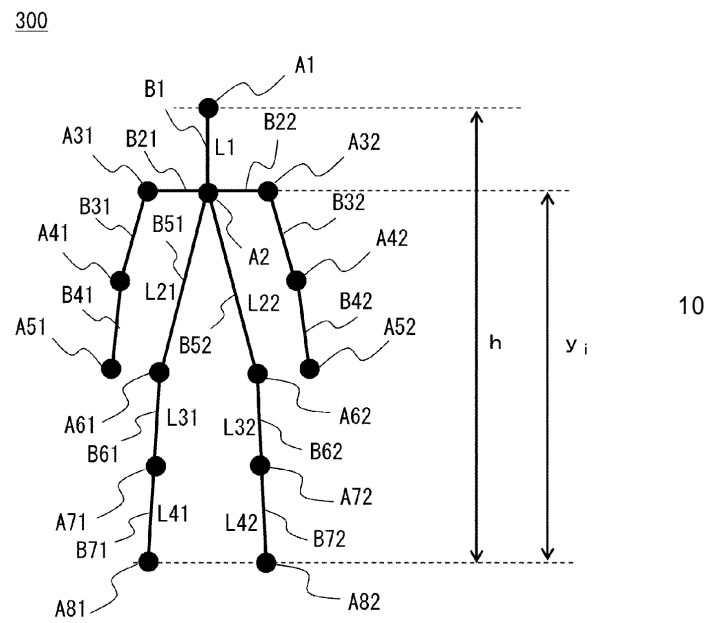
40

50

【 図 2 3 】



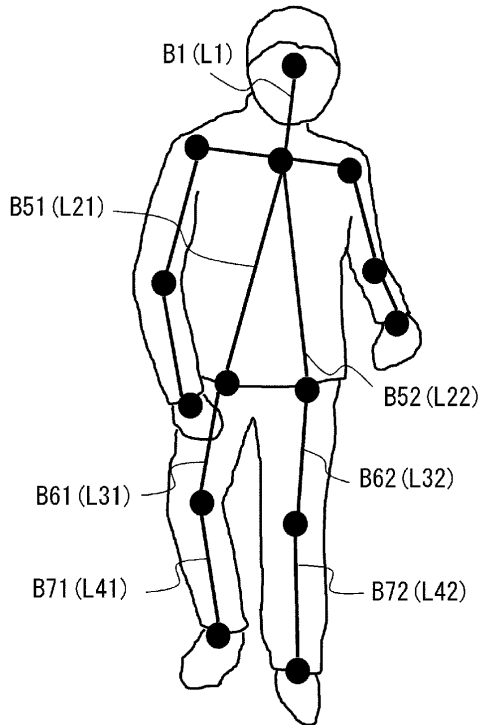
【 図 2 4 】



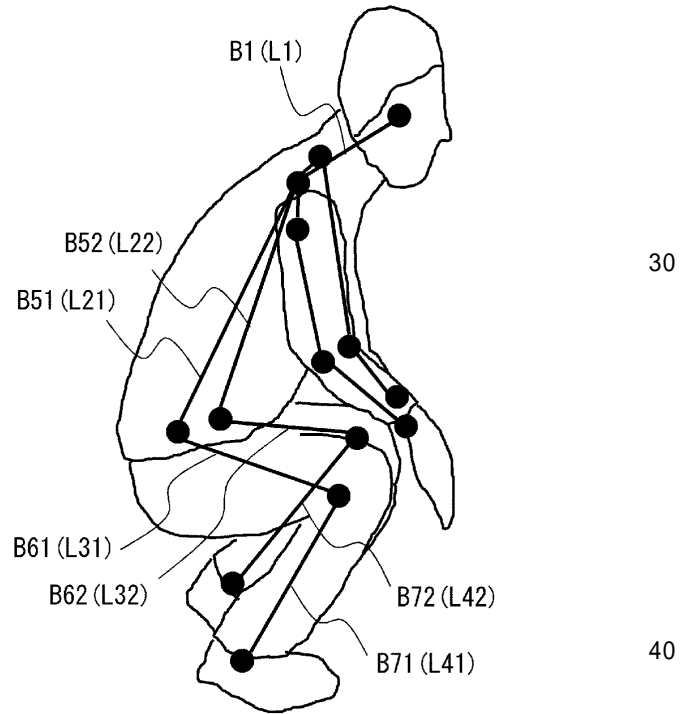
10

20

【 図 2 5 】



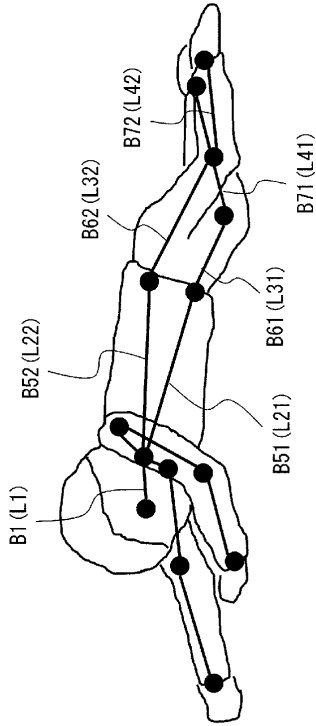
【 図 2 6 】



30

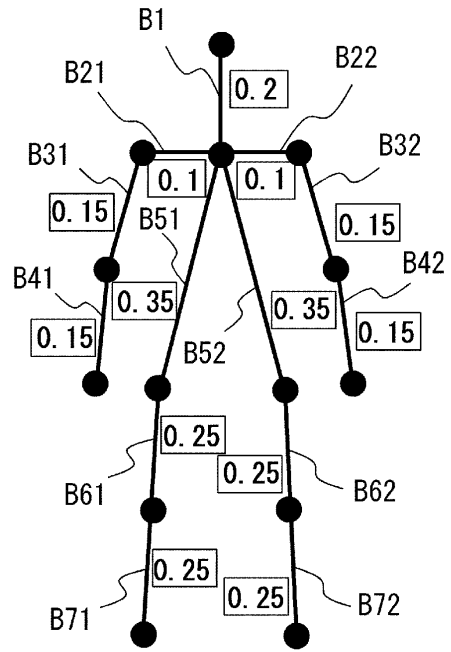
40

【図 27】



【図 28】

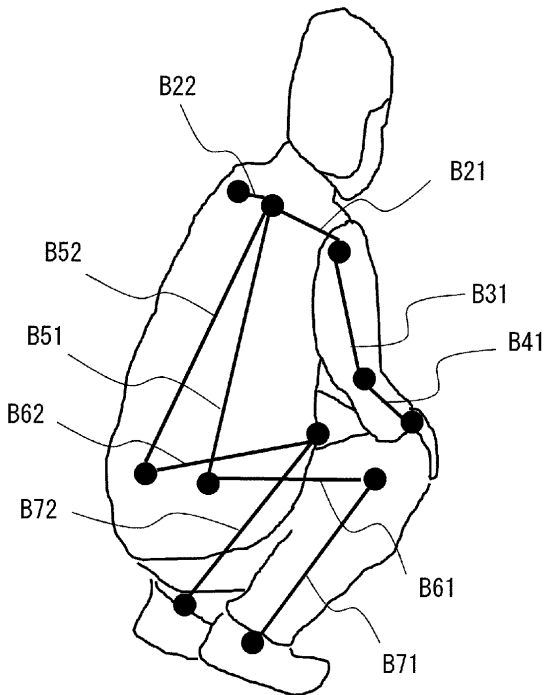
301



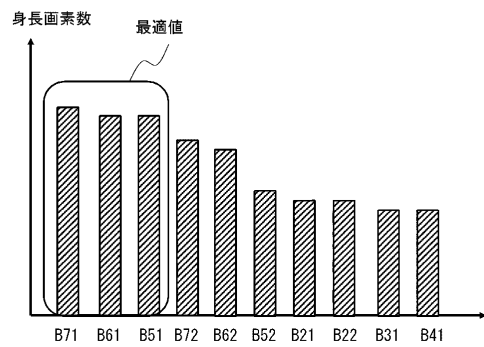
10

20

【図 29】



【図 30】

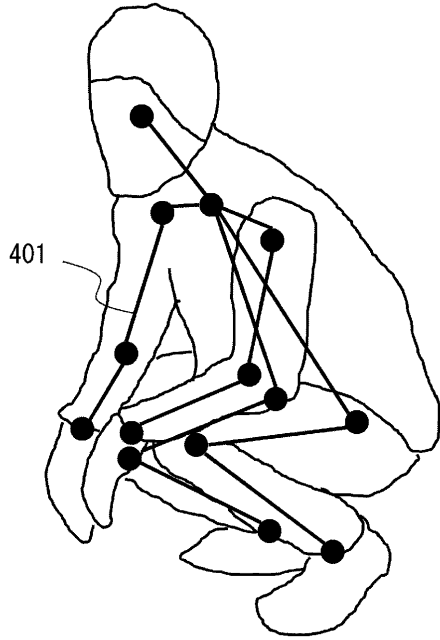


30

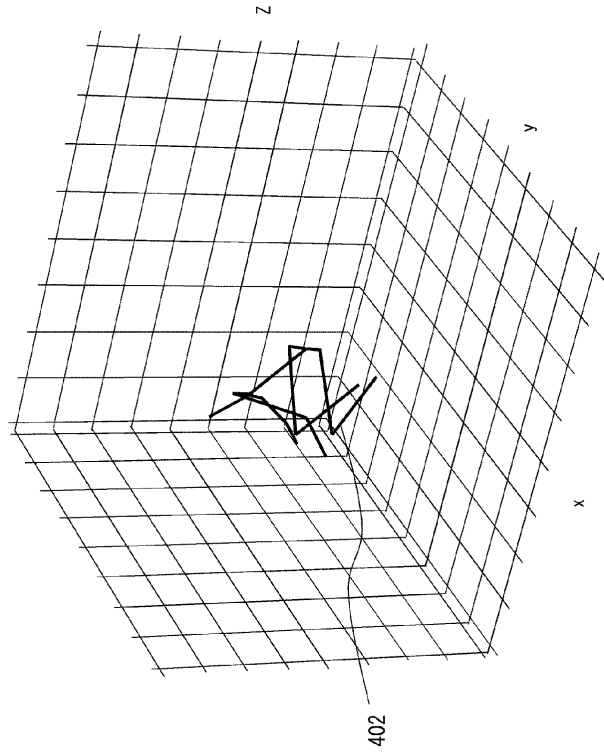
40

50

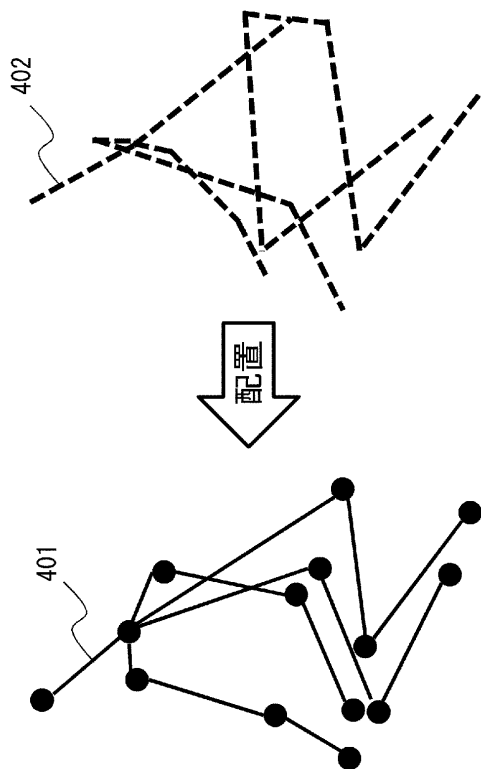
【図 3 1】



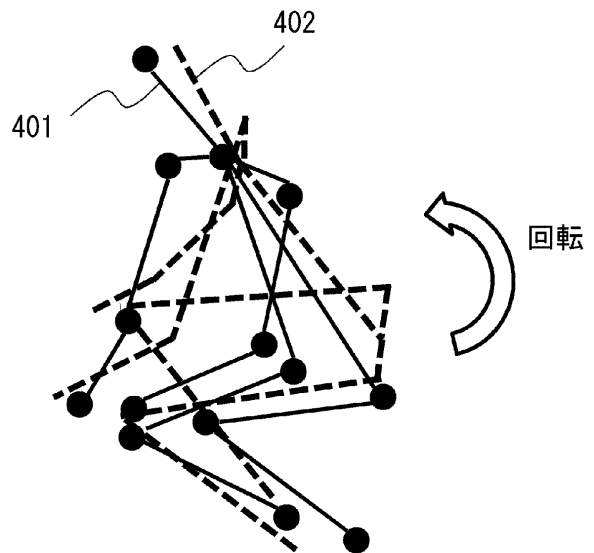
【図 3 2】



【図 3 3】



【図 3 4】



10

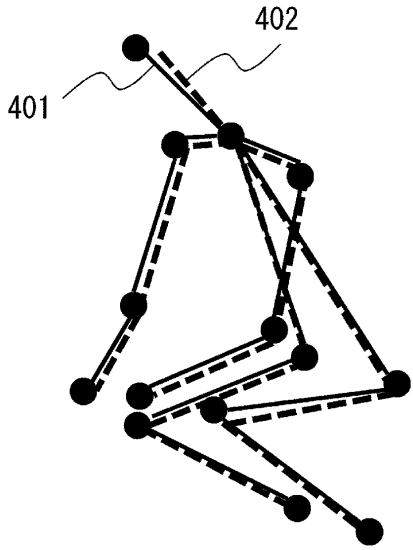
20

30

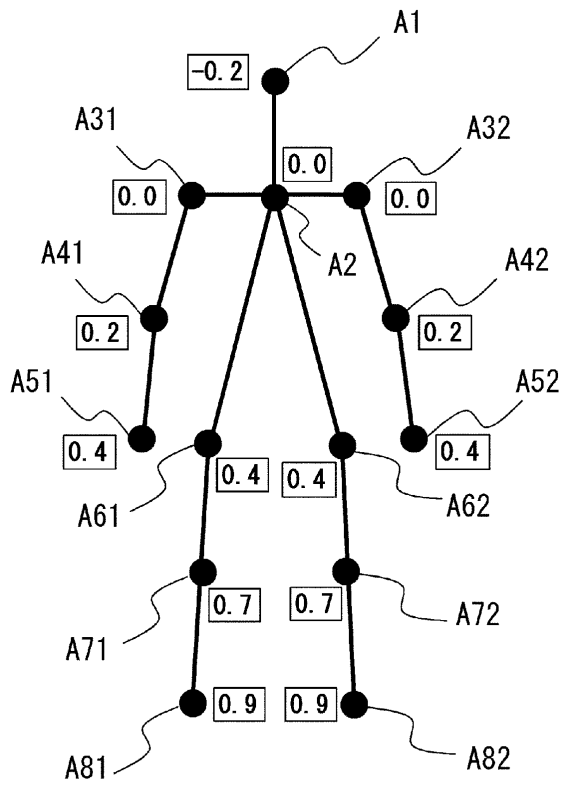
40

50

【図 3 5】



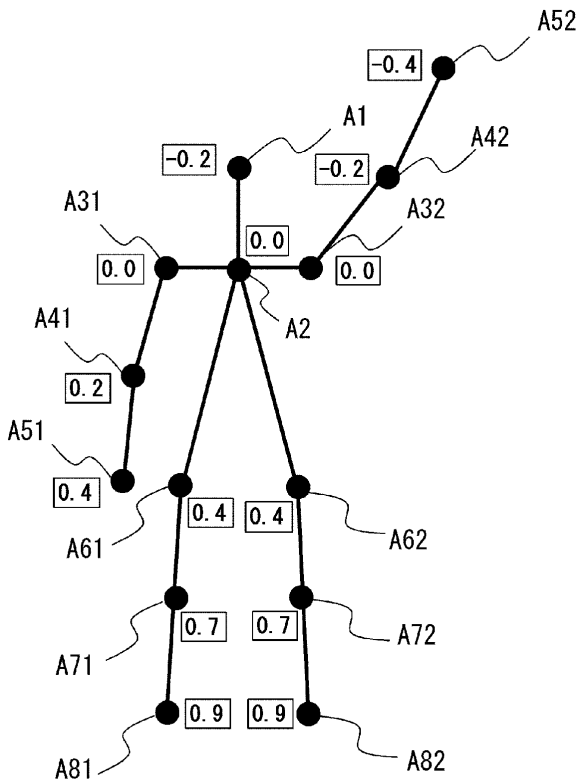
【図 3 6】



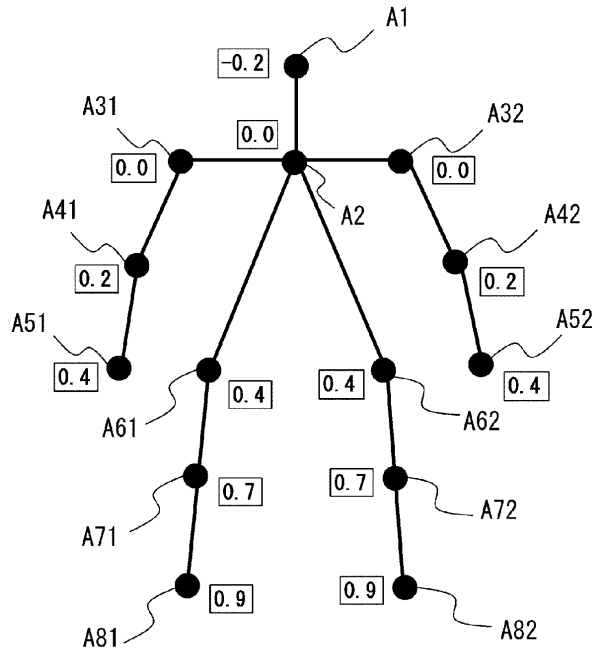
10

20

【図 3 7】



【図 3 8】



30

40

50

【 図 3 9 】

クエリ用の画像

画像識別情報	キーワード
C0001	座り込み、正座
C0002	座り込み、あぐら
.	.
.	.
.	.

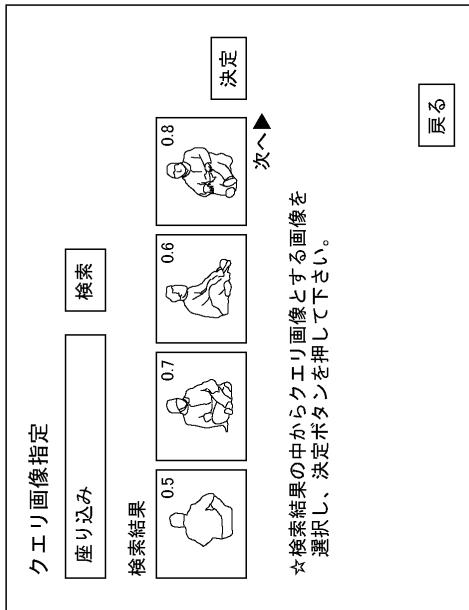
【 図 4 0 】

解析対象画像

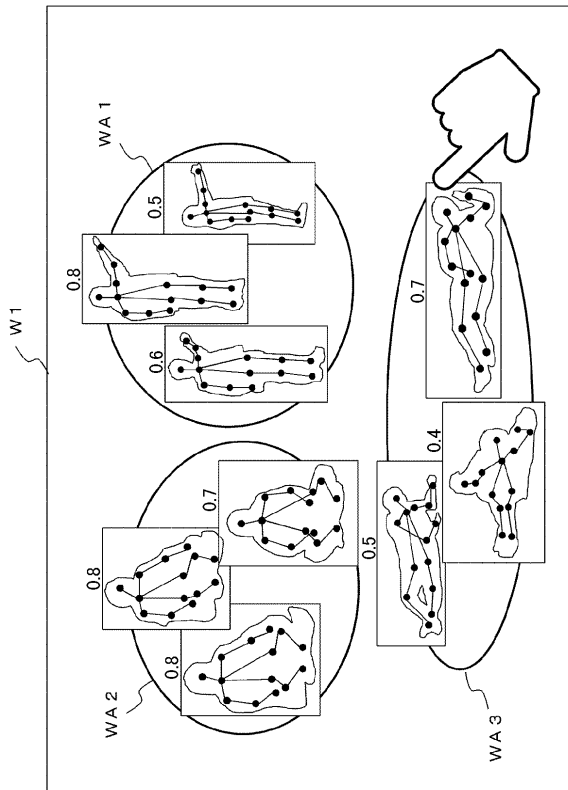
画像識別情報	キーワード	...
000001	座り込み、正座	...
000002
000003
000004
.	.	.
.	.	.
.	.	.

10

【 図 4 1 】



【 図 4 2 】



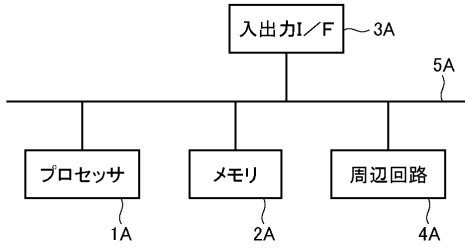
20

30

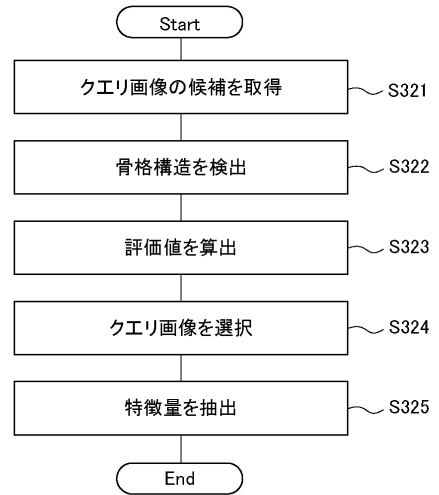
40

50

【図 4 3】

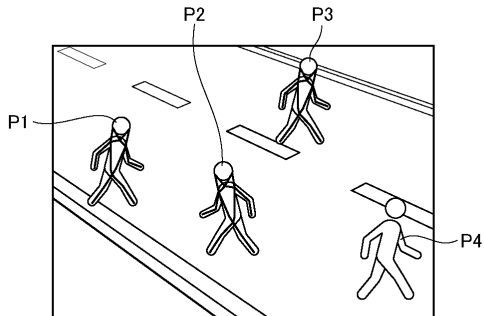


【図 4 4】

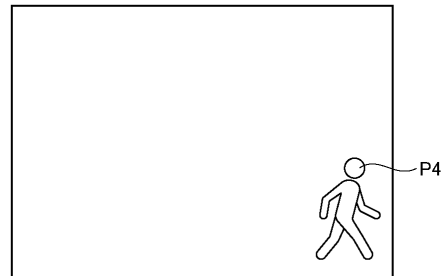


10

【図 4 5】



【図 4 6】



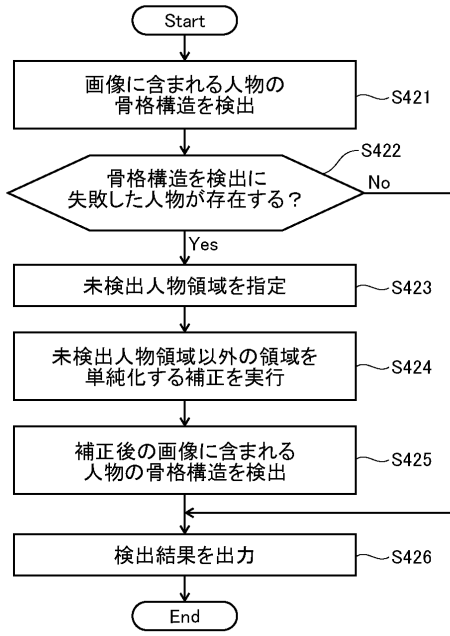
20

30

40

50

【 図 4 7 】



10

20

30

40

50

フロントページの続き

- 東京都港区芝五丁目7番1号 日本電気株式会社内
(72)発明者 西村 祥治
東京都港区芝五丁目7番1号 日本電気株式会社内
審査官 山田 辰美
(56)参考文献 特開2012-120647(JP,A)
特開2017-199303(JP,A)
特許第6593949(JP,B1)
(58)調査した分野 (Int.Cl., DB名)
G06T 7/00-7/90
G06V 10/22