

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7697602号
(P7697602)

(45)発行日 令和7年6月24日(2025.6.24)

(24)登録日 令和7年6月16日(2025.6.16)

(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

請求項の数 9 (全24頁)

(21)出願番号	特願2024-541918(P2024-541918)	(73)特許権者	000004237
(86)(22)出願日	令和4年1月21日(2022.1.21)		日本電気株式会社
(65)公表番号	特表2024-547215(P2024-547215 A)	(74)代理人	東京都港区芝五丁目7番1号 100103894
(43)公表日	令和6年12月26日(2024.12.26)		弁理士 家入 健
(86)国際出願番号	PCT/JP2022/002221	(72)発明者	パン ヤドン
(87)国際公開番号	WO2023/139757		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5年7月27日(2023.7.27)		
審査請求日	令和6年7月11日(2024.7.11)	審査官	藤原 敬利

最終頁に続く

(54)【発明の名称】 姿勢推定装置、姿勢推定方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

一人以上の人物が撮像された対象画像を取得する取得手段と、
前記対象画像からキーポイントを検出するキーポイント検出手段と、
前記キーポイントに基づいて1つ以上の選択ファクタを算出し、前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、前記キーポイントに対して適用するアルゴリズムを選択するアルゴリズム選択手段と、
前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するキーポイント関連付け手段と、
前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定する推定手段と、を有し、
前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、姿勢推定装置。

【請求項2】

前記キーポイントの種類は、右肩と左肩とを含み、
前記アルゴリズム選択手段は、前記選択ファクタとして、前記対象画像内の人物の密集度を算出し、
前記密集度の前記算出は、

各前記左肩の前記キーポイントを、最も近い右肩の前記キーポイントとリンクすることと、
複数の前記左肩の前記キーポイントにリンクされた右肩の前記キーポイントごとに、前記
右肩の前記キーポイントを有する最長のリンクの長さが、前記右肩の前記キーポイントを
有する最短のリンクの長さの所定倍を超えると、前記最長のリンクを削除することと、
複数の前記左肩の前記キーポイントとリンクされた前記右肩の前記キーポイントの数に
基づいて、前記対象画像内の人物の前記密集度を算出することと、を含む、請求項 1 に記
載の姿勢推定装置。

【請求項 3】

前記キーポイントの種類は、右肩と左肩とを含み、
前記アルゴリズム選択手段は、前記選択ファクタとして、前記対象画像内の人物の解像
度を算出し、
前記解像度の前記算出は、

各前記左肩の前記キーポイントを、最も近い右肩の前記キーポイントとリンクすることと、
複数の前記左肩の前記キーポイントにリンクされた右肩の前記キーポイントごとに、前記
右肩の前記キーポイントを有する最長のリンクの長さが、前記右肩の前記キーポイントを
有する最短のリンクの長さの所定倍を超えると、前記最長のリンクを削除することと、
前記削除後に残っている前記リンクの数と、前記対象画像の寸法のうちの 1 つに基づ
いて定義された閾値未満の長さである前記リンクの数と、に基づいて、前記対象画像内の
人物の前記解像度を算出することと、を含む、請求項 1 又は 2 に記載の姿勢推定装置。

【請求項 4】

前記事前定義されたアルゴリズムは、ミッドポイントアルゴリズム、方向マップアルゴ
リズム、位置マップアルゴリズム、又はそれらのうちの 2 つ若しくは 3 つを含む、請求項
1 から 3 のいずれか一項に記載の姿勢推定装置。

【請求項 5】

前記アルゴリズムの選択は、

前記選択ファクタが、前記選択ファクタの閾値未満であるか否かを判定することと、
前記選択ファクタが、前記選択ファクタの前記閾値未満であるとき、第 1 の前記アル
ゴリズムを選択することと、

前記選択ファクタが、前記選択ファクタの前記閾値以上であるとき、第 2 の前記アル
ゴリズムを選択することと、を含む、請求項 1 から 4 のいずれか一項に記載の姿勢推定装
置。

【請求項 6】

前記アルゴリズムの選択は、

前記対象画像内の人物の前記解像度が、前記解像度の閾値未満であるか否かを判定す
ることと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値未満であるとき、ミッ
ドポイントアルゴリズムを選択することと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値以上であるとき、方向
マップアルゴリズム又は位置マップアルゴリズムを選択することと、を含む、請求項 5 に
記載の姿勢推定装置。

【請求項 7】

前記アルゴリズムの選択は、

前記対象画像内の人物の前記密集度が、前記密集度の閾値未満であるか否かを判定す
ることと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値未満であるとき、ミッ
ドポイントアルゴリズム又は方向マップアルゴリズムを選択することと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値以上であるとき、位置
マップアルゴリズムを選択することと、を含む、請求項 5 又は 6 に記載の姿勢推定装置。

【請求項 8】

1 つ以上のコンピュータによって実行される姿勢推定方法であって、

10

20

30

40

50

一人以上の人物が撮像された対象画像を取得するステップと、
前記対象画像からキーポイントを検出するステップと、
前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップと、
前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、前記キーポイントに対して適用するアルゴリズムを選択するステップと、

前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するステップと、

前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を含み、

前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、姿勢推定方法。

【請求項9】

一人以上の人物が撮像された対象画像を取得するステップと、

前記対象画像からキーポイントを検出するステップと、

前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップと、

前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、前記キーポイントに対して適用するアルゴリズムを選択するステップと、

前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するステップと、

前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を1つ以上のコンピュータに実行させ、

前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、全体として、画像を分析して、画像に撮像された人物の姿勢を推定する技術に関する。

【背景技術】

【0002】

一人以上の人物が撮像された画像に対して実行される分析には様々な種類がある。これらの分析のうちの1つは、画像に撮像された各人物の姿勢を推定する姿勢推定である。人物の姿勢は、画像から検出された身体の間節など、キーポイントに基づいて推定される。

【0003】

画像内に複数の人物が撮像されている場合、姿勢推定は、各グループが互いに同一人物に属するキーポイントを含むように、キーポイントをグループに分割する「キーポイント関連付け (key-point association)」と呼ぶ処理を含む。特許文献1は、キーポイント関連付けのためのアルゴリズムのうちの1つを開示している。

【先行技術文献】

【特許文献】

【0004】

【文献】国際公開第2021/014530号

【発明の概要】

【発明が解決しようとする課題】

【0005】

キーポイント関連付けには複数のアルゴリズムがあり、あらゆる状況に対して最適なアルゴリズムはない。本開示の目的は、画像に撮像された人物の姿勢を正確に推定する新規な技術を提供することである。

【課題を解決するための手段】

【0006】

本開示によって提供される姿勢推定装置は、命令を記憶するように構成される少なくとも1つのメモリと、少なくとも1つのプロセッサを有する。前記プロセッサは、一人以上の人物が撮像された対象画像を取得することと、前記対象画像からキーポイントを検出することと、前記キーポイントに基づいて1つ以上の選択ファクタを算出することと、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ことと、前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択することと、前記キーポイントと、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行することと、前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定することと、を実行する前記命令を実行するように構成される。

【0007】

本開示のよって提供される姿勢推定方法は、1つ以上のコンピュータによって実行される。前記方法は、一人以上の人物が撮像された対象画像を取得するステップと、前記対象画像からキーポイントを検出するステップと、前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップと、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ステップと、前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択するステップと、前記キーポイントと、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するステップと、前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を含む。

【0008】

本開示によって提供される非一時的コンピュータ可読記憶媒体は、プログラムを格納する。前記プログラムは、1つ以上のコンピュータに、一人以上の人物が撮像された対象画像を取得するステップと、前記対象画像からキーポイントを検出するステップと、前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップと、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ステップと、前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択するステップと、前記キーポイントと、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するステップと、前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を実行させる。

【発明の効果】

【0009】

本開示によれば、画像内の人物の姿勢を正確に推定する新規の技術が提供される。

【図面の簡単な説明】

【0010】

【図 1】実施形態 1 の姿勢推定装置の概要を示す図である。

【図 2】実施形態 1 の姿勢推定装置の機能構成の一例を示すブロック図である。

【図 3】実施形態 1 の姿勢推定装置のハードウェア構成の一例を示すブロック図である。

【図 4】実施形態 1 の姿勢推定装置によって実行される処理の例示的な流れを示すフローチャートである。

【図 5】対象画像における人物の密集度を算出する方法の一例を示すフローチャートである。

【図 6】ステップ 202 から 206 の場合の例を示す図である。

【図 7】ミッドポイントモデルと方向マップモデルの訓練を示す図である。

【図 8】方向マップモデルと位置マップモデルの訓練を示す図である。

【図 9】ミッドポイントモデル、方向マップモデル、及び位置マップモデルの訓練を示す図である。

【図 10】キーポイント関連付けアルゴリズムを選択する処理の第 1 の例示的な流れを示すフローチャートである。

【図 11】キーポイント関連付けアルゴリズムを選択する処理の第 2 の例示的な流れを示すフローチャートである。

【図 12】キーポイント関連付けアルゴリズムを選択する処理の第 3 の例示的な流れ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】

【0011】

本開示による実施の形態について、図面を参照しつつ以下に説明する。各図面において同一の要素には同一の符号を付し、必要に応じて重複する説明を省略する。また、所定の情報（例えば、所定の値又は所定の閾値）は、特に断らない限り、その情報を使用するコンピュータがアクセス可能な記憶装置に予め記憶されている。

【0012】

実施形態 1

<概要>

図 1 は、実施形態 1 の姿勢推定装置 2000 の概要を示す。なお、図 1 に示す概要は、姿勢推定装置 2000 を理解しやすくするために、姿勢推定装置 2000 の動作の一例を示したものであり、姿勢推定装置 2000 が取りうる動作の範囲を限定したり、狭めたりするものではない。

【0013】

姿勢推定装置 2000 は、一人以上の人物が撮像された対象画像 10 を取得し、それぞれの人物の姿勢を推定する。そのために、姿勢推定装置 2000 は、対象画像 10 からキーポイントを検出し、検出したキーポイントに対してキーポイント関連付けを実行する。キーポイントは、関節などの人物の身体の特徴点を示しうる。キーポイント関連付けは、各グループが互いに同一人物に属するキーポイントを含むように、キーポイントをグループに分割する処理である。キーポイント関連付けによってその人物に属すると特定されたキーポイントに基づいて、各人物の姿勢を推定することができる。

【0014】

キーポイント関連付けには複数のアルゴリズムがあり、どのアルゴリズムがその画像に撮像された人物の姿勢を推定するために適しているかは、分析される画像に依存する。以下、キーポイント関連付けのためのアルゴリズムを「キーポイント関連付けアルゴリズム」と呼ぶ。したがって、姿勢推定装置 2000 は、対象画像 10 に関連するファクタを算出して、事前定義された複数のキーポイント関連付けアルゴリズムから、対象画像 10 のためのキーポイント関連付けアルゴリズムを選択する。以下、このファクタを「選択ファクタ」と呼ぶ。選択ファクタは、対象画像 10 内の人物の密集度、解像度、又はその両方を含んでもよい。

【0015】

姿勢推定装置 2000 は、対象画像 10 から検出されたキーポイントに対して、選択し

10

20

30

40

50

たキーポイント関連付けアルゴリズムを実行することにより、キーポイントのグループ（以下、キーポイントグループ）を取得し、各キーポイントグループは、互いに同一人物に属すると推定されるキーポイントを含む。そして、キーポイントごとに、姿勢推定装置 2000 は、キーポイントグループに含まれるキーポイントに基づいて、キーポイントグループに対応する人物の姿勢を特定する。

【0016】

＜作用効果の例＞

様々なキーポイント関連付けアルゴリズムが存在し、あらゆる状況に対して最適なアルゴリズムは存在しない。姿勢推定装置 2000 によれば、対象画像 10 に適用するキーポイント関連付けアルゴリズムは固定ではなく、選択ファクタに基づいて、事前定義された複数のキーポイント関連付けアルゴリズムから選択される。選択ファクタは、対象画像 10 に撮像された人物の密集度、解像度、又はその両方を含みうる。このように、対象画像 10 に適用するキーポイント関連付けアルゴリズムは、対象画像 10 に撮像された人物の密集度、解像度、又はその両方に基づいて適宜選択される。したがって、対象画像 10 内の人物の姿勢を正確に推定することができる。

【0017】

以下、姿勢推定装置 2000 の詳細な説明を述べる。

【0018】

＜機能構成の例＞

図 2 は、実施形態 1 の姿勢推定装置 2000 の機能構成の一例を示すブロック図である。姿勢推定装置 2000 は、取得部 2020 と、キーポイント検出部 2040 と、アルゴリズム選択部 2060 と、キーポイント関連付け部 2080 と、推定部 2100 と、を備える。取得部 2020 は、対象画像 10 を取得する。キーポイント検出部 2040 は、対象画像 10 からキーポイントを検出する。アルゴリズム選択部 2060 は、1 つ以上の選択ファクタを算出し、算出した選択ファクタに基づいて、事前定義されたものの中からキーポイント関連付けアルゴリズムを選択する。キーポイント関連付け部 2080 は、検出したキーポイントに対して、選択したキーポイント関連付けアルゴリズムを実行し、それによってキーポイントグループを生成する。キーポイントグループごとに、推定部 2100 は、キーポイントグループにおけるキーポイントに基づいて、キーポイントグループに対応する人物の姿勢を推定する。

【0019】

＜ハードウェア構成の例＞

姿勢推定装置 2000 は、1 つ以上のコンピュータによって実現されてもよい。1 つ以上のコンピュータの各々は、姿勢推定装置 2000 を実現するために製造された専用のコンピュータであってもよいし、パーソナルコンピュータ(PC: personal computer)、サーバマシン、モバイルデバイスといった汎用のコンピュータであってもよい。

【0020】

姿勢推定装置 2000 は、コンピュータにアプリケーションをインストールすることで実現されてもよい。アプリケーションは、コンピュータを姿勢推定装置 2000 として機能させるためのプログラムによって実現される。すなわち、プログラムは、姿勢推定装置 2000 の機能部を実装するものである。

【0021】

図 3 は、実施形態 1 の姿勢推定装置 2000 を実現するコンピュータ 1000 のハードウェア構成の一例を示すブロック図である。図 3 において、コンピュータ 1000 は、バス 1020 と、プロセッサ 1040 と、メモリ 1060 と、記憶装置 1080 と、入出力 (Input/Output) インターフェース 1100 と、ネットワークインターフェース 1120 と、を有する。

【0022】

バス 1020 は、プロセッサ 1040、メモリ 1060、記憶装置 1080、入出力インターフェース 1100 及びネットワークインターフェース 1120 が相互にデータを送

10

20

30

40

50

受信するためのデータ伝送路である。プロセッサ 1040 は、CPU (Central Processing Unit)、GPU (Graphics Processing Unit)、FPGA (Field-Programmable Gate Array) などのプロセッサである。メモリ 1060 は、RAM (Random Access Memory) 又は ROM (Read Only Memory) などの主記憶要素である。記憶装置 1080 は、ハードディスク、SSD (Solid State Drive)、メモリカードなどの補助記憶要素である。入出力インターフェース 1100 は、コンピュータ 1000 と、キーボード、マウス、ディスプレイデバイスなどの周辺機器と、の間のインターフェースである。ネットワークインターフェース 1120 は、コンピュータ 1000 とネットワークとの間のインターフェースである。ネットワークは、LAN (Local Area Network) であってもよく、又は WAN (Wide Area Network) であってもよい。

10

【0023】

コンピュータ 1000 のハードウェア構成は、図 3 に示すものに限定しない。例えば、上述したように、姿勢推定装置 2000 は、複数のコンピュータによって実現されてもよい。この場合、それらのコンピュータは、ネットワークを介して互いに接続されてもよい。

【0024】

<処理の流れ>

図 4 は、実施形態 1 の姿勢推定装置 2000 によって実行される処理の例示的な流れを示すフローチャートである。取得部 2020 は、対象画像 10 を取得する (S102)。キーポイント検出部 2040 は、対象画像 10 からキーポイントを検出する (S104)。アルゴリズム選択部 2060 は、1 つ以上の選択ファクタを算出する (S106)。アルゴリズム選択部 2060 は、算出した選択ファクタに基づいて、事前定義されたものからキーポイント関連付けアルゴリズムを選択する (S108)。キーポイント関連付け部 2080 は、キーポイントグループを生成するために、検出したキーポイントに対して、選択したアルゴリズムを実行する (S110)。推定部 2100 は、キーポイントグループごとに、人物の姿勢を推定する (S112)。

20

【0025】

<対象画像 10 の取得：S102>

取得部 2020 は、対象画像 10 を取得する。対象画像 10 の取得方法は様々である。いくつかの態様では、対象画像 10 は、姿勢推定装置 2000 が取得することができるように、記憶装置に予め記憶される。この場合、取得部 2020 は、対象画像を取得するために、記憶装置にアクセスしてもよい。他の実施形態では、対象画像 10 は、対象画像 10 を生成するカメラなど、別のコンピュータによって送信されてもよい。この場合、取得部 2020 は、対象画像 10 を受信することにより取得してもよい。

30

【0026】

いくつかの実施形態では、対象画像 10 は、ビデオを形成するビデオフレームなど、連続する画像のうちの 1 つであってもよい。この場合、姿勢推定装置 2000 は、連続する画像の全部又は一部を対象画像 10 として取得し、対象画像 10 ごとにそれぞれの人物の姿勢を推定しうる。

【0027】

<キーポイントの検出：S104>

キーポイント検出部 2040 は、対象画像 10 からキーポイントを検出する (S104)。画像から人物のキーポイントを検出する様々な方法があり、キーポイント検出部 2040 は、対象画像 10 からキーポイントを検出するために、これらの方法のうちの 1 つを使用しうる。人物のキーポイントは、首、肩、肘、手首、腰、膝、足首などにおける関節など、人物の身体の特徴であってもよい。

40

【0028】

いくつかの態様において、キーポイント検出部 2040 は、入力として画像を取得するように構成され、更に画像の入力に応じて入力画像からキーポイントを検出するように予め訓練された、機械学習ベースのモデル (例えば、ニューラルネットワーク) を有する。以下、このモデルを「キーポイント検出モデル」と呼ぶ。

50

【0029】

キーポイント検出モデルは、入力として対象画像10を取得し、対象画像10から特徴を抽出し、抽出した特徴に基づいて対象画像10から1つ以上のキーポイントを検出し、抽出した特徴に基づいて各キーポイント（例えば、首、右肩など）のクラスを特定しうる。この場合、キーポイント検出モデルは、対象画像10から特徴を抽出するように予め訓練された第1のモデルと、対象画像10から抽出した特徴に基づいてキーポイントを検出して分類するように事前に訓練された第2のモデルと、を含みうる。第1のモデル及び第2のモデルの各々は、ニューラルネットワークなどの機械学習ベースのモデルとして構成されうる。なお、入力画像からキーポイントを検出でき、更にそれらを分類できる機械学習モデルには様々な種類があり、キーポイント検出モデルを、そのようなモデルのうちの1つとして構成することができる。

10

【0030】

<選択ファクタの算出：S106>

対象画像10に適したキーポイント関連付けアルゴリズムを選択するために、アルゴリズム選択部2060は、検出したキーポイントに基づいて選択ファクタを算出する（S106）。上述したように、選択ファクタは、対象画像10内の人物の密集度、対象画像10内の人物の解像度、又はその両方を含んでもよい。以下、これらのファクタの算出方法の一例について説明する。

【0031】

《人物の密集度》

人物の密集度を選択ファクタとして使用する場合、アルゴリズム選択部2060は、対象画像10から検出したキーポイントに基づいて、対象画像10内の人物の密集度を算出する。人物の密集度は、右肩及び左肩のキーポイントを使用して測ることができる。

20

【0032】

図5は、対象画像10における人物の密集度の算出方法の一例を示すフローチャートである。アルゴリズム選択部2060は、検出したすべてのキーポイントから、左肩又は右肩を表すキーポイントを抽出する（S202）。そして、左肩のキーポイントごとに、アルゴリズム選択部2060は、最も近い右肩のキーポイントを探索し、それらをリンクする（S204）。このステップにより、右肩のいくつかのキーポイントを、左肩の複数のキーポイントにリンクしうる。

30

【0033】

MPポイントごとに、アルゴリズム選択部2060は、そのMPポイントを一端とする最長リンク及び最短リンクを抽出し、最長リンクの長さが最短リンクの長さの所定倍（例えば、2倍）を超えると、その最長リンクを除去する（S206）。このステップにより、一部のMPポイントは、最長リンクを除去した結果として、非MPポイント（すなわち、単一の左肩のキーポイントのみにリンクされた右肩のキーポイント）になりうる。

【0034】

図6は、ステップ202からステップ206の場合の一例を示している。この例では、ステップS202において、3つの右肩22-1から22-3及び5つの左肩24-1から24-5が検出されている。次に、ステップS204において、左肩24-1から24-5を、右肩22-1、22-2、22-2、22-3、及び22-3に、それぞれリンクする。この場合、右肩22-2及び22-3は、MPポイントである。

40

【0035】

ステップS206では、右肩22-3について、右肩22-3と左肩24-5との間の最長リンクの長さが、右肩22-3と左肩24-4との間の最短リンクの長さの所定倍を超えると判定する。これにより、右肩22-3と左肩24-5との間のリンクが除去される。この除去の結果、右肩22-3は、非MPポイントとなる。

【0036】

ステップS206は、MPポイントの数が収束する（例えば、一定になる）まで繰り返される。以下、このMPポイントの数を「NMP」と呼ぶ。具体的には、アルゴリズム選

50

択部 2060 は、NMP が収束したか否かを判定する (S208)。NMP が収束していない場合 (S208:NO)、再度ステップ S206 が実行される。一方、NMP が収束した場合 (S208:YES)、アルゴリズム選択部 2060 は、対象画像 10 内の人物の密集度として NMP を設定する。概念的には、NMP が大きいほど、対象画像 10 内に人物が密集している。

【0037】

他の態様において、アルゴリズム選択部 2060 は、NMP に基づいて、対象画像 10 内の人物の密集度を算出してもよい。例えば、密集度の算出のために、入力値に比例した値を出力する関数が予め定義されている。この場合、アルゴリズム選択部 2060 は、この関数に NMP を入力し、NMP に比例した出力値を取得し、この出力値を対象画像 10 内の人物の密集度として使用しうる。

10

【0038】

《人物の解像度》

人物の解像度を選択ファクタとして使用する場合、アルゴリズム選択部 2060 は、上述した NMP に基づいて人物の解像度を算出してもよい。具体的には、アルゴリズム選択部 2060 は、図 5 のステップ S208 の後に残った MP ポイントから、対象画像の寸法のうちの 1 つに基づいて定義されうる閾値 (例えば、対象画像 10 の幅の 1/25) 未満の長さのリンクを抽出する。ここで抽出されたリンクを、「SL リンク」と呼ぶ。

【0039】

アルゴリズム選択部 2060 は、以下のように RSL と呼ぶ値を算出する。

20

【数 1】

$$RSL = \begin{cases} \frac{NSL}{NAL} & (if\ NAL > 0) \\ 0 & (if\ NAL = 0) \end{cases} \quad (1)$$

式 (1) において、NSL は、SL リンクの数を表し、NAL は、MP ポイントを一端とする、すべてのリンクの数を表す。

【0040】

30

概念的には、RSL が大きいほど、対象画像 10 内の人物の解像度は低くなる。したがって、いくつかの態様において、アルゴリズム選択部 2060 は、RSL が小さくなるにつれて大きくなる値として、対象画像 10 内の人物の解像度を算出する。例えば、解像度の算出のために、予め、入力値の逆数に比例した値を出力する関数が予め定義されている。この場合、アルゴリズム選択部 2060 は、この関数に RSL を入力して、1/RSL に比例した出力値を取得し、この出力値を対象画像 10 内の人物の解像度として使用してもよい。なお、この関数は、入力として 0 が与えられたとき、最大値を出力するように定義される。

【0041】

他の態様において、アルゴリズム選択部 2060 は、対象画像 10 内の人物の解像度を表す選択ファクタとして RSL を使用しうる。この場合、後述するように、対象画像 10 内の人物の解像度が閾値未満であるか否かを判定するとき、アルゴリズム選択部 2060 は、RSL が閾値よりも大きい場合に、対象画像 10 内の人物の解像度が閾値未満であると判定する。

40

【0042】

<キーポイント関連付けアルゴリズム>

事前定義されたキーポイント関連付けアルゴリズムは、1) ミッドポイントアルゴリズム、2) 方向マップアルゴリズム、及び 3) 位置マップアルゴリズムのうちの 2 つ以上を含んでもよい。以下、各アルゴリズムについて説明する。

【0043】

50

<<ミッドポイントアルゴリズム>>

ミッドポイントアルゴリズムは、キーポイント関連付けを実行するために、対象画像 10 からミッドポイント (mid-point) を検出する。ミッドポイントは、2 つのキーポイントの中間に位置するポイントである。ミッドポイントアルゴリズムの詳細は、特許文献 1 に開示されている。

【0044】

ミッドポイントアルゴリズムは、ニューラルネットワークなどの機械学習モデルを使用して実現されうる。以下、この機械学習モデルを「ミッドポイントモデル」と呼ぶ。ミッドポイントモデルは、対象画像 10 を入力として取得するように構成され、入力データの入力に応じて、対象画像 10 内のミッドポイントを出力するように予め訓練されうる。

【0045】

ミッドポイントアルゴリズムは、対象画像 10 内のミッドポイントを取得するために、対象画像 10 をミッドポイントモデルに入力しうる。次いで、ミッドポイントアルゴリズムは、ミッドポイントに基づいて、キーポイントをキーポイントグループに分割する。

【0046】

<<方向マップアルゴリズム>>

方向マップアルゴリズムは、対象画像 10 の方向マップを生成し、キーポイント関連付けのための方向マップを使用して、キーポイントをキーポイントグループに分割する。方向マップは、対象画像 10 から抽出した特徴マップであり、対象画像 10 と同じサイズである。方向マップは、対象画像 10 における人物の領域（以下、人物領域）の画素ごとの単位ベクトルを示す。人物領域内の画素に対応する単位ベクトルは、その画素から、その人物領域の事前定義された基準ポイント (reference point) までを指す。人物領域の基準ポイントは、人物領域に対応する人物の特定のキーポイント（例えば、首のキーポイント）であってもよい。

【0047】

より具体的には、方向マップは、H 方向マップ及び V 方向マップと呼ぶ 2 つの特徴マップのセットを含んでもよい。H 方向マップにおいて、人物領域の画素は、それに対応する単位ベクトルの水平成分（すなわち、x 成分）を示す。一方、V 方向マップにおいて、人物領域の画素は、それに対応する単位ベクトルの垂直成分（y 成分）を示す。

【0048】

方向マップアルゴリズムは、ニューラルネットワークなどの機械学習モデルを使用して実現されうる。以下、この機械学習モデルを「方向マップモデル」と呼ぶ。方向マップモデルは、対象画像 10 を入力として取得するように構成され、入力データの入力に応じて、対象画像 10 の方向マップを出力するように予め訓練されうる。方向マップアルゴリズムは、方向マップを取得するために、対象画像 10 を方向マップモデルに入力しうる。次いで、方向マップアルゴリズムは、方向マップを使用してキーポイントをキーポイントグループに分割する。

【0049】

方向マップアルゴリズムは、方向マップを使用して、対象画像 10 から検出したキーポイントと人物のペアごとにスコアを算出し、算出したスコアに基づいて、どのキーポイントがどの人物に属するかを特定する。次いで、キーポイントごとに、方向マップアルゴリズムは、そのキーポイントが、そのキーポイントに関するスコアのうちの最大スコアに対応する人物に属すると特定する。

【0050】

キーポイント K1 と人物 P1 のペアがスコア S1、キーポイント K1 と人物 P2 のペアがスコア S2、キーポイント K1 と人物 P3 のペアがスコア S3 であるとする。また、S1、S2、S3 の最大スコアは S2 である。この場合、キーポイント K1 と人物 P2 とのペアのスコアが最大であるため、方向マップアルゴリズムは、キーポイント K1 が人物 P2 に属すると特定する。

【0051】

10

20

30

40

50

キーポイントと人物とのペアのスコアは、3つのファクタ OB 、 RoD 、及び D の積として算出することができる（すなわち、 $S=OB*RoD*D$ ）。これら3つのファクタを、以下のように算出する。方向マップアルゴリズムは、キーポイントと、人物の基準ポイントとの間の線上に1つ以上の中間ポイント（intermediate point）を生成する。方向マップアルゴリズムは、中間ポイントごとに、中間ポイントが人物領域に位置するか否かを判定する。方向マップアルゴリズムは、ファクタ OB として、中間ポイントの何パーセントが人物領域内に位置するかを算出する。例えば、3つの中間ポイントのうち2つが人物領域に位置する場合、ファクタ OB は、 $2/3$ となる。

【0052】

そして、人物領域に位置する中間ポイントごとに、方向マップアルゴリズムは、方向マップからその中間ポイントに対応する単位ベクトルを取得する。対象画像10上の中間ポイントの座標を $(x1, y1)$ とする。この場合、中間ポイントに対応する単位ベクトルは、方向マップの $(x1, y1)$ に位置するものとなる。方向マップアルゴリズムはまた、キーポイントに対応する単位ベクトルを取得する。

10

【0053】

方向マップアルゴリズムは、ファクタ RoD として、取得された単位ベクトルの方向の変動（variation）を算出する。単位ベクトルの方向の変動は、それらの単位ベクトルのうちの2つの間の最大差を表す。

【0054】

ファクタ D として、方向マップアルゴリズムは、人物のキーポイントと基準ポイントとの間の距離を算出する。

20

【0055】

<<位置マップアルゴリズム>>

位置マップアルゴリズムは、対象画像10の位置マップを算出し、キーポイント関連付けのために位置マップを使用して、キーポイントをキーポイントグループに分割する。位置マップは、対象画像10から抽出した特徴マップであり、対象画像10と同じサイズである。位置マップにおいて、人物領域の画素は、人物領域の相対位置を示す。人物領域の相対位置は、対象画像10のベース位置（base position）に対する基準ポイント（例えば、首のキーポイント）の位置である。対象画像10のベース位置は、その原点（例えば、左上隅）でありうる。

30

【0056】

より具体的には、位置マップは、 H 位置マップ及び V 位置マップと呼ぶ2つの特徴マップを含んでもよい。 H 位置マップにおいて、人物領域内の画素は、ベース位置（例えば、対象画像10の左端）に対する、その人物領域の基準ポイントの水平位置を示す。一方、 V 位置マップにおいて、人物領域の画素は、ベース位置（例えば、対象画像10の上端）に対する、その人物領域の基準ポイントの垂直位置を示す。

【0057】

対象画像10の幅を W 、高さを H とし、対象画像10における人物領域の基準ポイントの絶対座標を $(x1, y1)$ とする。この場合、人物領域の相対位置は、 $(x1/W, y1/W)$ となる。したがって、この人物領域の画素は、 H 位置マップでは $x1/W$ を示し、 V 位置マップでは $y1/H$ を示す。

40

【0058】

位置マップアルゴリズムは、ニューラルネットワークなどの機械学習モデルを使用して実現されうる。以下、この機械学習モデルを「位置マップモデル」と呼ぶ。位置マップモデルは、対象画像10を入力として取得するように構成され、入力データの入力に応じて、対象画像10の位置マップを出力するように予め訓練されうる。位置マップアルゴリズムは、位置マップを取得するために、対象画像10を、位置マップモデルに入力しうる。次いで、位置マップアルゴリズムは、位置マップを使用して、キーポイントをキーポイントグループに分割する。

【0059】

50

対象画像 10 から検出されたキーポイントごとに、位置マップアルゴリズムは、対象画像 10 から検出された各人物からの距離を算出する。距離は、位置マップから取得したキーポイントと人物との相対位置を使用して算出される。

【0060】

具体的には、位置マップアルゴリズムは、位置マップ上の座標が対象画像 10 上のキーポイントの座標と同じである位置マップの画素から画素値を取得し、取得した値をキーポイントの相対位置として使用する。対象画像 10 において、キーポイント K1 の座標を (x_1, y_1) とする。また、H 位置マップの (x_1, y_1) の画素は x_2 を示し、V 位置マップの (x_1, y_1) の画素は y_2 を示す。この場合、位置マップアルゴリズムは、H 位置マップにおける (x_1, y_1) の画素から、画素値 x_2 を、キーポイント K1 の相対位置の x 座標として取得する。また、位置マップアルゴリズムは、V 位置マップにおける (x_1, y_1) の画素の画素値 y_2 を、キーポイント K1 の相対位置の y 座標として取得する。これにより、キーポイント K1 の相対位置は、 (x_2, y_2) に特定される。

10

【0061】

人物の相対位置に関して、位置マップアルゴリズムは、その人物に対応する人物領域内の画素から画素値を取得し、取得した値を人物の相対位置として使用する。人物 P1 の人物領域の画素が、H 位置マップで値 x_3 を示し、人物 P1 の人物領域の画素が、V 位置マップで値 y_3 を示すものとする。この場合、人物 P1 の相対位置は、 (x_3, y_3) に特定される。上述のように、キーポイントの相対位置が (x_2, y_2) で示されるとき、位置マップアルゴリズムは、 (x_2, y_2) と (x_3, y_3) との間の距離を、キーポイントと人物 P1 との間の距離として算出する。

20

【0062】

キーポイントに対する各人物までの距離を算出した後、位置マップアルゴリズムは、キーポイントからの距離が最も短い人物を特定する。次いで、位置マップは、キーポイントが特定された人物に属すると特定する。

【0063】

<モデルの構成>

<<方向マップモデルについて>>

方向マップは、基準ポイント（首のキーポイントなど）と、人物の全身体領域内の任意の他の画素との間の幾何学的関係を記述する特徴マップである。方向マップの品質を良好にするために、方向マップモデルは、人物の身体状況（すなわち、異なる身体部分間の接続方法）を十分に理解していることが好ましい。2つのペアにしたキーポイントの中間に定義されるミッドポイントは、方向マップモデルによる異なる身体部分間の接続の好適な理解を助け、したがって、方向マップの品質の向上を助けることができる。

30

【0064】

上記の洞察に基づいて、方向マップは、対象画像 10 から検出されたキーポイント及びミッドポイントを使用して、方向マップを生成するように構成されることが好ましい。したがって、ミッドポイントアルゴリズム及び方向マップモデルアルゴリズムが、事前定義されたキーポイント関連付けアルゴリズムとして採用されるとき、方向マップモデルは、対象画像 10 だけでなく、ミッドポイントモデルの出力も入力として取得するように構成することができる。この場合、ミッドポイントモデル及び方向マップモデルを、互いに同じ訓練データを使用して、一括して訓練することができる。なお、ミッドポイントモデルは、対象画像 10 からキーポイントを検出するために、キーポイント検出モデルとしての機能を有してもよい。

40

【0065】

図 7 は、ミッドポイントモデル及び方向マップモデルの訓練を示している。この例では、ミッドポイントモデルは、対象画像 10 を入力として取得し、対象画像 10 からキーポイント及びミッドポイントを検出する。方向マップモデルが、ミッドポイントの出力（すなわち、キーポイント及びミッドポイント）を入力として取得しうるように、方向マップモデルは、ミッドポイントモデルに接続される。

50

【0066】

図7に示すモデルの訓練のための訓練データは、一人以上の人物が撮像された入力画像と、正解データとのセットを含む。正解データは、対応する入力画像から検出されるキーポイント及びミッドポイント、並びに対応する入力画像から生成される方向マップを示す。モデルは、それらの出力（すなわち、ミッドポイントによって検出されたキーポイント及びミッドポイント、並びに方向マップモデルによって生成された方向マップ）と、正解データとの間の相違度を表す損失を算出し、更に、算出した損失に基づいてモデルの訓練可能なパラメータを更新することによって、訓練される。

【0067】

<<位置マップモデルについて>>

基準ポイントを除く人物の身体領域内における任意の画素について、人物の2方向マップ（それぞれ v_x 及び v_y で示される）におけるその値は、画素から基準ポイントまでの単位ベクトルの X 成分及び Y 成分である。したがって、 v_x 及び v_y は、「 $v_x^2 + v_y^2 = 1$ 」の条件を満たし、これは、基準ポイントを除く人物の身体領域内のすべての画素について、 v_x と v_y との二乗和が一定値であることを意味する。一方、位置マップは、画像内の人物の位置に対応する一定の値を用いて、人物の身体領域を埋めるものとして定義される。したがって、方向マップの二乗は、人物の身体領域のすべての画素の値を一定値に収束させることによって、位置マップの品質の向上を助けることができる。なお、方向マップの二乗とは、 H 方向マップにおいて、各画素が、対応する画素によって示される値の二乗を示す、 H 方向マップの二乗と、 V 方向マップにおいて、各画素が、対応する画素によって示される値の二乗を示す、 V 方向マップの二乗と、を含む。

【0068】

上記の洞察に基づいて、方向マップの二乗を使用して位置マップを生成するように位置マップモデルを構成することが好ましい。したがって、方向マップモデルアルゴリズム及び位置マップアルゴリズムが、事前定義されたキーポイント関連付けアルゴリズムとして採用されるとき、位置マップモデルは、対象画像10だけでなく、方向マップモデルの出力の二乗も入力として取得するように構成することができる。この場合、方向マップモデル及び位置マップを、互いに同じ訓練データを使用して、一括して訓練することができる。

【0069】

図8は、方向マップモデル及び位置マップモデルの訓練を示している。この例では、方向マップモデルと位置マップモデルとの間に、方向マップモデルの出力の二乗を算出する構成部が存在する。この構成部は、方向マップモデルの出力を取得し、この出力の二乗を算出し、算出の結果を位置マップモデルに供給するように構成される。

【0070】

図8に示すモデルの訓練のための訓練データは、一人以上の人物が撮像された入力画像と、正解データと、のセットを含む。正解データは、対応する入力画像から生成される方向マップ及び位置マップを示す。モデルは、それらの出力（すなわち、方向マップモデルによって生成された方向マップ、及び位置マップモデルによって生成された位置マップ）と、正解データとの間の相違度を表す損失を算出し、更に、算出した損失に基づいてモデルの訓練可能なパラメータを更新することによって、訓練される。

【0071】

図7及び図8に示す構成を組み合わせることによって、ミッドポイントモデル、方向マップモデル、及び位置マップモデルは、それらのモデルが事前定義されたキーポイント関連付けアルゴリズムとして採用されるときに、一括して訓練することができる。図9は、ミッドポイントモデル、方向マップモデル、位置マップモデルの訓練を示している。

【0072】

図9に示すモデルの訓練のための訓練データは、一人以上の人物が撮像された入力画像と、正解データと、のセットを含む。正解データは、対応する入力画像から検出されるキーポイント及びミッドポイント、並びに対応する入力画像から生成される方向マップ及び位置マップを示す。モデルは、それらの出力（すなわち、ミッドポイントによって検出さ

10

20

30

40

50

れたキーポイント及びミッドポイント、方向マップモデルによって生成された方向マップ、及び位置マップモデルによって生成された位置マップ）と、正解データとの間の相違度を表す損失を算出し、更に、算出された損失に基づいてモデルの訓練可能なパラメータを更新することによって、訓練される。

【0073】

＜キーポイント関連付けアルゴリズムの選択：S108＞

アルゴリズム選択部2060は、選択ファクタに基づいて、対象画像10に適したキーポイント関連付けアルゴリズムを選択する（S108）。いくつかの態様において、アルゴリズム選択部2060は、選択ファクタが事前定義された閾値よりも大きいかなにかに基づいて、キーポイント関連付けアルゴリズムを決定してもよい。

10

【0074】

図10は、キーポイント関連付けアルゴリズムを選択する処理の第1の例示的な流れを示すフローチャートである。この例では、事前定義されたキーポイントアルゴリズムは、ミッドポイントアルゴリズムと、方向マップアルゴリズムと、位置マップアルゴリズムと、を含む。また、この例では、選択ファクタは、対象画像10内の人物の密集度と、対象画像10内の人物の解像度と、を含む。

【0075】

具体的には、アルゴリズム選択部2060は、対象画像10内の人物の解像度が閾値 ThR よりも小さいかなかを判定する（S302）。解像度が閾値 ThR よりも小さいとき（S302：YES）、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、ミッドポイントアルゴリズムを選択する（S304）。一方、解像度が閾値 ThR 以上であるとき（S302：NO）、アルゴリズム選択部2060は、対象画像10内の人物の密集度が閾値 ThD よりも小さいかなかを判定する（S306）。

20

【0076】

密集度が閾値 ThD よりも小さいとき（S306：YES）、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、方向マップアルゴリズムを選択する（S308）。一方、密集度が閾値 ThD 以上であるとき（S306：NO）、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、位置マップアルゴリズムを選択する（S310）。

30

【0077】

図11は、キーポイント関連付けアルゴリズムを選択する処理の第2の例示的な流れを示すフローチャートである。この例では、事前定義されたキーポイントアルゴリズムは、ミッドポイントアルゴリズムと、方向マップアルゴリズムと、を含む。また、この例では、対象画像10内の人物の解像度が、選択ファクタとして使用される。

【0078】

具体的には、アルゴリズム選択部2060は、対象画像10内の人物の解像度が閾値 ThR よりも小さいかなかを判定する（S402）。解像度が閾値 ThR よりも小さいとき（S402：YES）、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、ミッドポイントアルゴリズムを選択する（S404）。一方、解像度が閾値 ThR 以上であるとき（S402：NO）、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、方向マップアルゴリズムを選択する（S406）。

40

【0079】

なお、図11に示す例では、方向マップアルゴリズムに代えて、位置マップアルゴリズムを、事前定義されたキーポイント関連付けアルゴリズムのうちの1つとして採用してもよい。この場合、解像度が閾値 ThR 以上であるとき、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、位置マップアルゴリズムを選択する。

【0080】

50

図12は、キーポイント関連付けアルゴリズムを選択する処理の第3の例示的な流れを示すフローチャートである。この例では、事前定義されたキーポイントアルゴリズムは、方向マップアルゴリズムと、位置アルゴリズムと、を含む。また、この例では、対象画像10内の人物の密集度が、選択ファクタとして使用される。

【0081】

具体的には、アルゴリズム選択部2060は、対象画像10内の人物の密集度が閾値 ThD よりも小さいか否かを判定する (S502)。密集度が閾値 ThD よりも小さいとき (S502: YES)、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、方向マップアルゴリズムを選択する (S504)。一方、密集度が閾値 ThD 以上であるとき (S502: NO)、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、位置マップアルゴリズムを選択する (S506)。

【0082】

なお、図12に示す例では、方向マップアルゴリズムに代えて、ミッドポイントアルゴリズムを、事前定義されたキーポイント関連付けアルゴリズムのうちの1つとして採用してもよい。この場合、密集度が閾値 ThD 以上であるとき、アルゴリズム選択部2060は、対象画像10に適用するキーポイント関連付けアルゴリズムとして、ミッドポイントアルゴリズムを選択する。

【0083】

<姿勢推定装置2000からの出力>

姿勢推定装置2000は、姿勢推定の結果を示す情報（出力情報と呼ぶ）を出力するように構成されうる。例えば、出力情報は、対象画像10の識別子（例えば、フレーム番号）と、キーポイントグループごとに、キーポイントグループの推定姿勢の識別子、及びキーポイントグループ内の各キーポイントのキーポイント情報のセットと、を含んでもよい。推定姿勢の識別子は、キーポイントグループに対応する人物がどのような種類の姿勢を取っているかを示す。キーポイント情報は、キーポイントの種類（例えば、首、右肩など）及び位置（例えば、座標）を示す。

【0084】

出力情報を出力するには、様々な方法がある。いくつかの態様において、出力情報は、記憶装置に格納され、ディスプレイデバイスに表示され、又は姿勢推定装置2000のユーザのPC若しくはスマートフォンなど、別のコンピュータに送信されてもよい。

【0085】

プログラムは、様々なタイプの非一時的なコンピュータ可読媒体 (non-transitory computer readable medium) を用いて格納され、コンピュータに提供することができる。非一時的なコンピュータ可読媒体は、様々なタイプの実体のある記録媒体 (tangible storage medium) を含む。非一時的なコンピュータ可読媒体の例は、磁気記録媒体（例えば、フレキシブルディスク、磁気テープ、ハードディスクドライブ）、光磁気記録媒体（例えば、光磁気ディスク）、CD-ROM、CD-R、CD-R/W、半導体メモリ（例えば、マスクROM、PROM (Programmable ROM)、EPROM (Erasable PROM)、フラッシュROM、RAM)を含む。また、プログラムは、様々なタイプの一時的なコンピュータ可読媒体 (transitory computer readable medium) によってコンピュータに提供されてもよい。一時的なコンピュータ可読媒体の例は、電気信号、光信号、及び電磁波を含む。一時的なコンピュータ可読媒体は、電線及び光ファイバ等の有線通信路、又は無線通信路を介して、プログラムをコンピュータに供給できる。

【0086】

実施の形態を参照して、上記に本開示を説明したが、本開示は、上述の実施の形態に限定されるものではない。本発明の範囲内で、当業者が理解しうる様々な変更を本開示の構成及び詳細について行うことができる。

【0087】

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

10

20

30

40

50

られない。

<付記>

(付記 1)

姿勢推定装置であって、

命令を記憶するように構成される少なくとも1つのメモリと、

一人以上の人物が撮像された対象画像を取得することと、

前記対象画像からキーポイントを検出することと、

前記キーポイントに基づいて1つ以上の選択ファクタを算出することであって、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ことと、

前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択することと、

前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行することと、

前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定することと、
を実行する前記命令を実行するように構成される少なくとも1つのプロセッサと、を有する、姿勢推定装置。

(付記 2)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記密集度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の前記キーポイントにリンクすることと、

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除することと、

左肩の複数のキーポイントとリンクされた右肩の前記キーポイントの数に基づいて、前記対象画像内の人物の前記密集度を算出することと、を含む、付記1に記載の姿勢推定装置。

(付記 3)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記解像度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の前記キーポイントにリンクすることと、

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除することと、

前記削除後に残っている前記リンクの数と、前記対象画像の寸法のうちの1つに基づいて定義された閾値未満の長さである前記リンクの数と、に基づいて、前記対象画像内の人物の前記解像度を算出することと、を含む、付記1又は2に記載の姿勢推定装置。

(付記 4)

前記事前定義されたアルゴリズムは、ミッドポイントアルゴリズム、方向マップアルゴリズム、位置マップアルゴリズム、又はそれらのうちの2つ若しくは3つを含む、付記1から3のいずれか一項に記載の姿勢推定装置。

(付記 5)

キーポイント関連付けのための選択アルゴリズムは、

前記選択ファクタが、前記選択ファクタの閾値未満であるか否かを判定することと、

前記選択ファクタが、前記選択ファクタの前記閾値未満であるとき、キーポイント関連付けのための第1のアルゴリズムを選択することと、

前記選択ファクタが、前記選択ファクタの前記閾値以上であるとき、キーポイント関連付けのための第2のアルゴリズムを選択することと、を含む、付記1から4のいずれか一項に記載の姿勢推定装置。

(付記6)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記解像度が、前記解像度の閾値未満であるか否かを判定することと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値未満であるとき、ミッドポイントアルゴリズムを選択することと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値以上であるとき、方向マップアルゴリズム又は位置マップアルゴリズムを選択することと、を含む、付記5に記載の姿勢推定装置。

10

(付記7)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記密集度が、前記密集度の閾値未満であるか否かを判定することと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値未満であるとき、ミッドポイントアルゴリズム又は方向マップアルゴリズムを選択することと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値以上であるとき、位置マップアルゴリズムを選択することと、を含む、付記5又は6に記載の姿勢推定装置。

20

(付記8)

1つ以上のコンピュータによって実行される姿勢推定方法であって、

一人以上の人物が撮像された対象画像を取得するステップと、

前記対象画像からキーポイントを検出するステップと、

前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップであって、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ステップと、

前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択するステップと、

前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポイントに対してキーポイント関連付けを実行するステップと、

30

前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を含む、姿勢推定方法。

(付記9)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記密集度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の前記キーポイントにリンクするステップと、

40

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除するステップと、

左肩の複数のキーポイントとリンクされた右肩の前記キーポイントの数に基づいて、前記対象画像内の人物の前記密集度を算出するステップと、を含む、付記8に記載の姿勢推定方法。

(付記10)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記解像度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の

50

前記キーポイントにリンクするステップと、

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除するステップと、

前記削除後に残っている前記リンクの数と、前記対象画像の寸法のうちの1つに基づいて定義された閾値未満の長さである前記リンクの数と、に基づいて、前記対象画像内の人物の前記解像度を算出するステップと、を含む、付記8又は9に記載の姿勢推定方法。

(付記11)

前記事前定義されたアルゴリズムは、ミッドポイントアルゴリズム、方向マップアルゴリズム、位置マップアルゴリズム、又はそれらのうちの2つ若しくは3つを含む、付記8から10のいずれか一項に記載の姿勢推定方法。

(付記12)

キーポイント関連付けのための選択アルゴリズムは、

前記選択ファクタが、前記選択ファクタの閾値未満であるか否かを判定するステップと、

前記選択ファクタが、前記選択ファクタの前記閾値未満であるとき、キーポイント関連付けのための第1のアルゴリズムを選択するステップと、

前記選択ファクタが、前記選択ファクタの前記閾値以上であるとき、キーポイント関連付けのための第2のアルゴリズムを選択するステップと、を含む、付記8から11のいずれか一項に記載の姿勢推定方法。

(付記13)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記解像度が、前記解像度の閾値未満であるか否かを判定するステップと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値未満であるとき、ミッドポイントアルゴリズムを選択するステップと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値以上であるとき、方向マップアルゴリズム又は位置マップアルゴリズムを選択するステップと、を含む、付記12に記載の姿勢推定方法。

(付記14)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記密集度が、前記密集度の閾値未満であるか否かを判定するステップと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値未満であるとき、ミッドポイントアルゴリズム又は方向マップアルゴリズムを選択するステップと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値以上であるとき、位置マップアルゴリズムを選択するステップと、を含む、付記12又は13に記載の姿勢推定方法。

(付記15)

プログラムを記憶する非一時的コンピュータ可読記憶媒体であって、前記プログラムが、1つ以上のコンピュータに、

一人以上の人物が撮像された対象画像を取得するステップと、

前記対象画像からキーポイントを検出するステップと、

前記キーポイントに基づいて1つ以上の選択ファクタを算出するステップであって、前記1つ以上の選択ファクタは、前記対象画像内の人物の密集度、解像度、又はその両方を含む、ステップと、

前記1つ以上の選択ファクタに基づいて、キーポイント関連付けのための事前定義されたアルゴリズムから、キーポイント関連付けのためのアルゴリズムを選択するステップと、

前記キーポイントを、各々が互いに同一人物の前記キーポイントを含む1つ以上のキーポイントグループに分割するために、前記選択されたアルゴリズムを使用して前記キーポ

10

20

30

40

50

イントに対してキーポイント関連付けを実行するステップと、

前記キーポイントグループごとに、前記キーポイントグループに含まれる前記キーポイントに基づいて、前記キーポイントグループに対応する前記人物の姿勢を推定するステップと、を実行させる、非一時的コンピュータ可読記憶媒体。

(付記 16)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記密集度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の前記キーポイントにリンクすることと、

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除することと、

左肩の複数のキーポイントとリンクされた右肩の前記キーポイントの数に基づいて、前記対象画像内の人物の前記密集度を算出することと、を含む、付記 15 に記載の記憶媒体。

(付記 17)

前記キーポイントの種類は、右肩と左肩とを含み、

前記対象画像内の人物の前記解像度の前記算出は、

左肩の前記キーポイントごとに、左肩の前記キーポイントを、それに最も近い右肩の前記キーポイントにリンクすることと、

左肩の複数のキーポイントにリンクされた右肩の前記キーポイントごとに、右肩の前記キーポイントを有する最短のリンクの長さの所定倍を最長のリンクの長さが超えるとき、右肩の前記キーポイントを有する前記最長のリンクを削除することと、

前記削除後に残っている前記リンクの数と、前記対象画像の寸法のうちの 1 つに基づいて定義された閾値未満の長さである前記リンクの数と、に基づいて、前記対象画像内の人物の前記解像度を算出することと、を含む、付記 15 又は 16 に記載の記憶媒体。

(付記 18)

前記事前定義されたアルゴリズムは、ミッドポイントアルゴリズム、方向マップアルゴリズム、位置マップアルゴリズム、又はそれらのうちの 2 つ若しくは 3 つを含む、付記 15 から 17 のいずれか一項に記載の記憶媒体。

(付記 19)

キーポイント関連付けのための選択アルゴリズムは、

前記選択ファクタが、前記選択ファクタの閾値未満であるか否かを判定するステップと、

前記選択ファクタが、前記選択ファクタの前記閾値未満であるとき、キーポイント関連付けのための第 1 のアルゴリズムを選択するステップと、

前記選択ファクタが、前記選択ファクタの前記閾値以上であるとき、キーポイント関連付けのための第 2 のアルゴリズムを選択するステップと、を含む、付記 15 から 18 のいずれか一項に記載の記憶媒体。

(付記 20)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記解像度が、前記解像度の閾値未満であるか否かを判定するステップと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値未満であるとき、ミッドポイントアルゴリズムを選択するステップと、

前記対象画像内の人物の前記解像度が、前記解像度の前記閾値以上であるとき、方向マップアルゴリズム又は位置マップアルゴリズムを選択するステップと、を含む、付記 19 に記載の記憶媒体。

(付記 21)

キーポイント関連付けのための選択アルゴリズムは、

前記対象画像内の人物の前記密集度が、前記密集度の閾値未満であるか否かを判定するステップと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値未満であるとき、ミッドポイントアルゴリズム又は方向マップアルゴリズムを選択するステップと、

前記対象画像内の人物の前記密集度が、前記密集度の前記閾値以上であるとき、位置マップアルゴリズムを選択するステップと、を含む、付記 19 又は 20 に記載の記憶媒体。

【符号の説明】

【0088】

10	対象画像	
22	右肩のキーポイント	10
24	左肩のキーポイント	
1000	コンピュータ	
1020	バス	
1040	プロセッサ	
1060	メモリ	
1080	記憶装置	
1100	入出力インターフェース	
1120	ネットワークインターフェース	
2000	姿勢推定装置	
2020	取得部	20
2040	キーポイント検出部	
2060	アルゴリズム選択部	
2080	キーポイント関連付け部	
2100	推定部	

30

40

50

【図面】
【図 1】

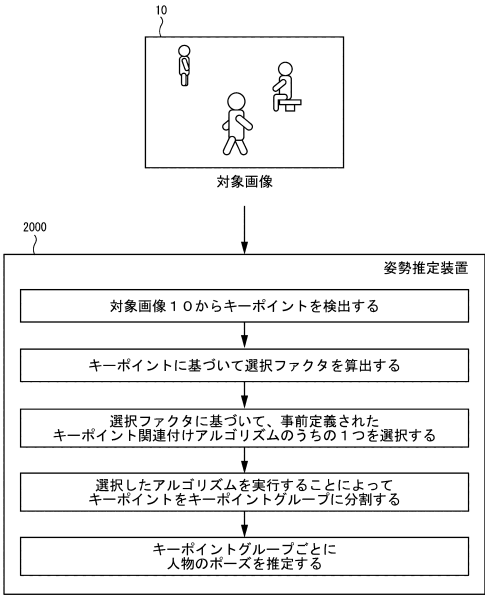


Fig. 1

【図 2】

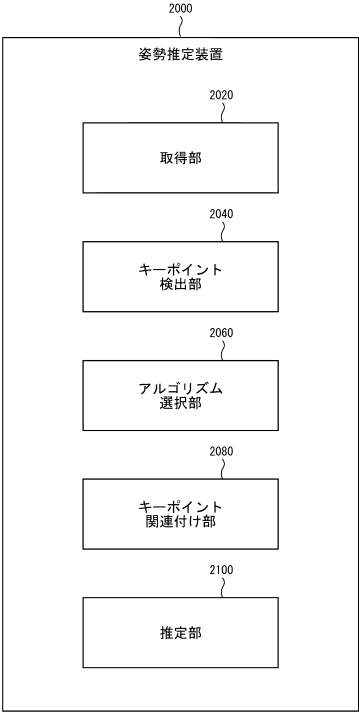


Fig. 2

【図 3】

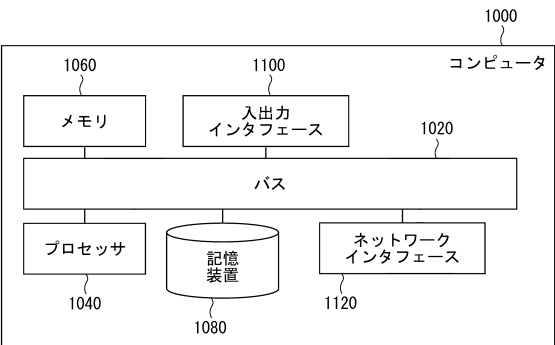


Fig. 3

【図 4】

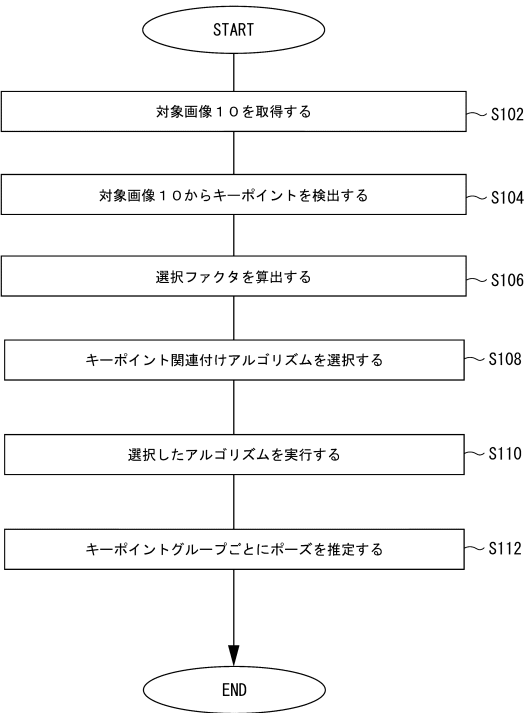


Fig. 4

10

20

30

40

50

【図 5】

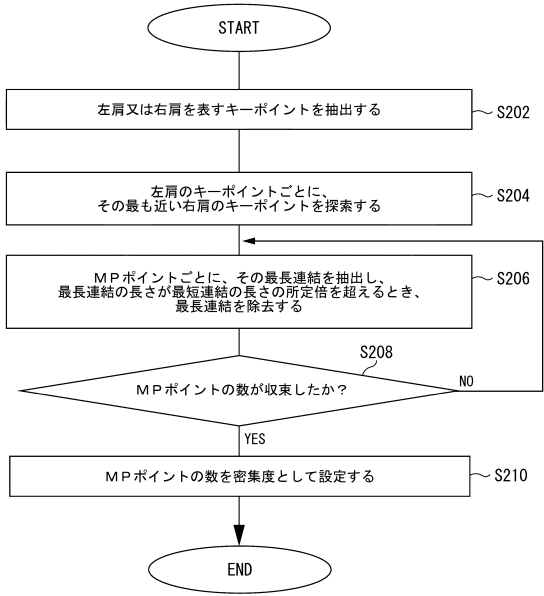


Fig. 5

【図 6】

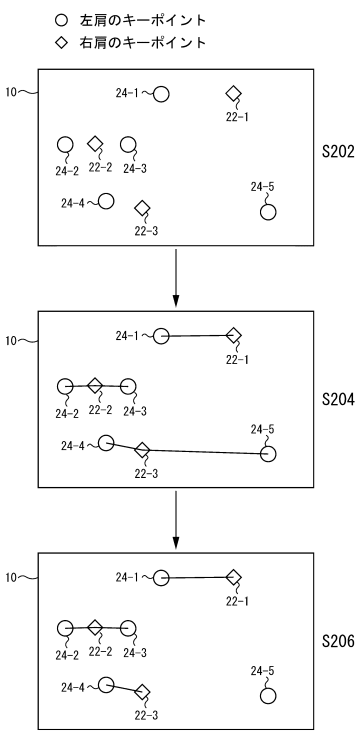


Fig. 6

【図 7】

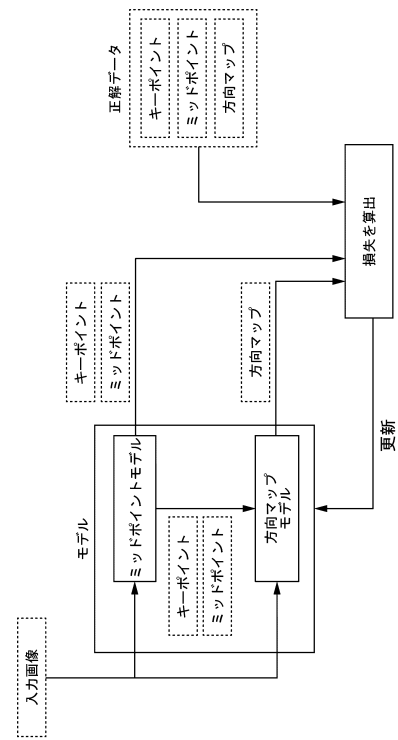


Fig. 7

【図 8】

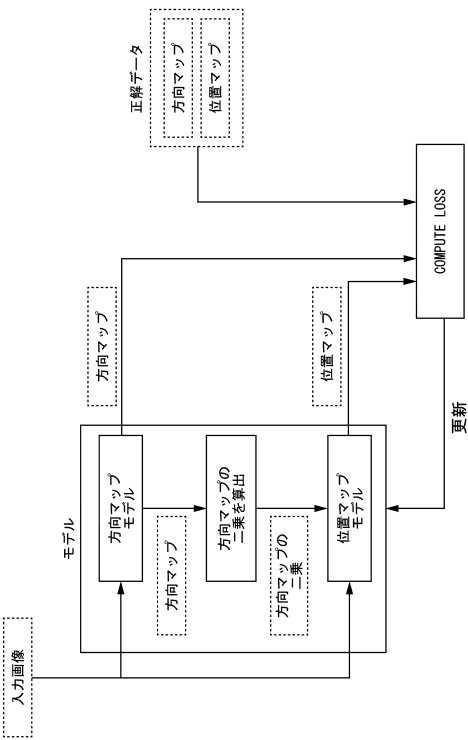
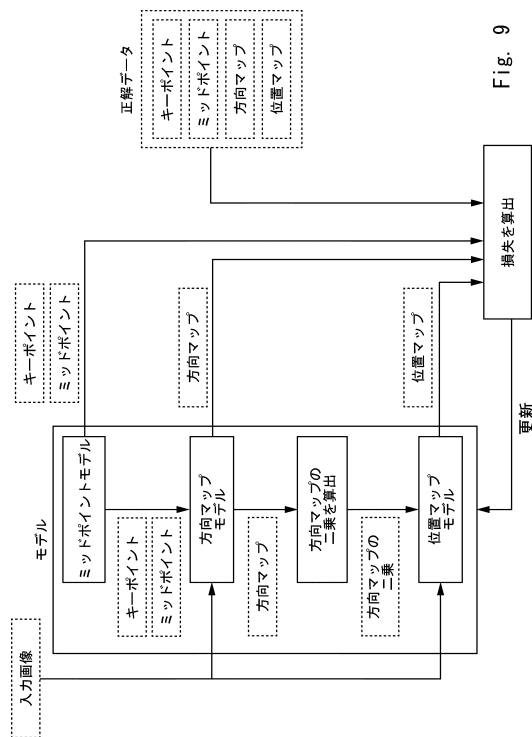
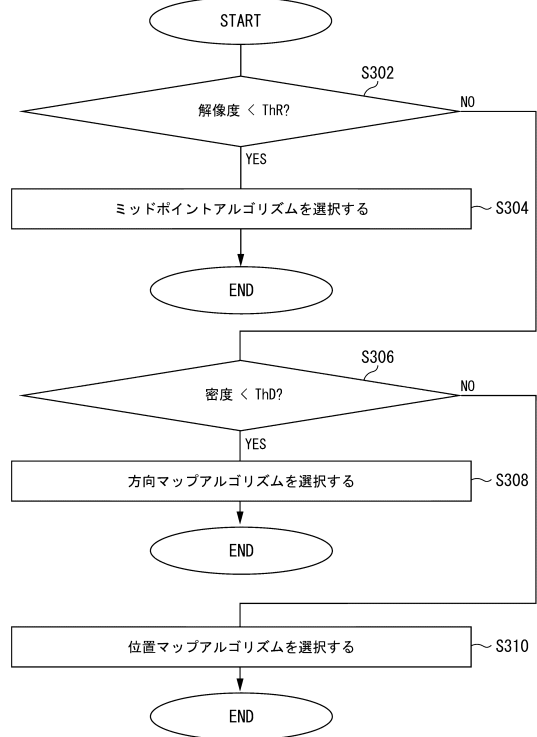


Fig. 8

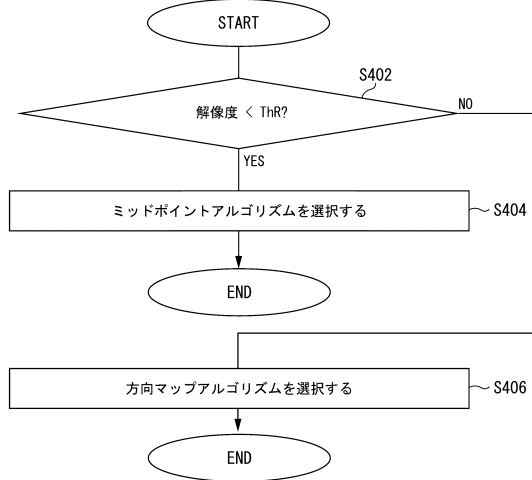
【図 9】



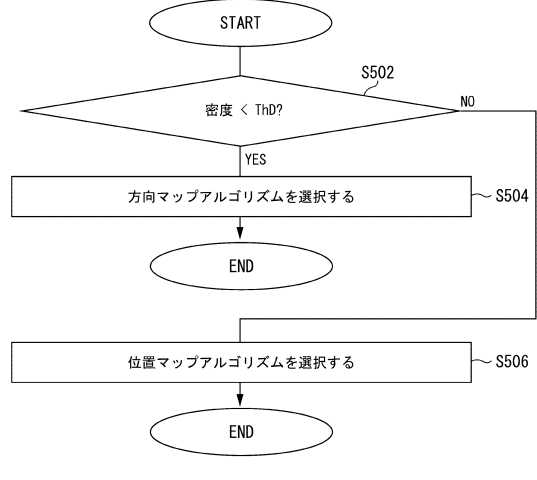
【図 10】



【図 11】



【図 12】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2021-157835 (JP, A)
 米国特許出願公開第 2019/0019030 (US, A1)
 中国特許出願公開第 111368594 (CN, A)
 中国特許出願公開第 111310625 (CN, A)
 中国特許出願公開第 110909655 (CN, A)
- (58)調査した分野 (Int.Cl., DB名)
 G 0 6 T 7 / 0 0
 G 0 6 T 7 / 7 0