

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2022-54251

(P2022-54251A)

(43)公開日 令和4年4月6日(2022.4.6)

(51)国際特許分類		F I		テーマコード(参考)	
G 0 6 T	7/55 (2017.01)	G 0 6 T	7/55	5 L 0 9 6	
G 0 6 T	7/00 (2017.01)	G 0 6 T	7/00	3 5 0 B	
G 0 6 T	7/70 (2017.01)	G 0 6 T	7/70	B	
		G 0 6 T	7/00	6 6 0 B	

審査請求 未請求 請求項の数 11 O L (全13頁)

(21)出願番号	特願2020-161329(P2020-161329)	(71)出願人	000001007 キャノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	令和2年9月25日(2020.9.25)	(74)代理人	110003281 特許業務法人大塚国際特許事務所
		(72)発明者	伊藤 秀憲 東京都大田区下丸子3丁目30番2号 キャノン株式会社内
		Fターム(参考)	5L096 AA06 CA05 CA21 DA01 EA43 FA09 FA35 FA38 FA67 FA69 GA08 GA30 GA51 HA11 JA03 KA04

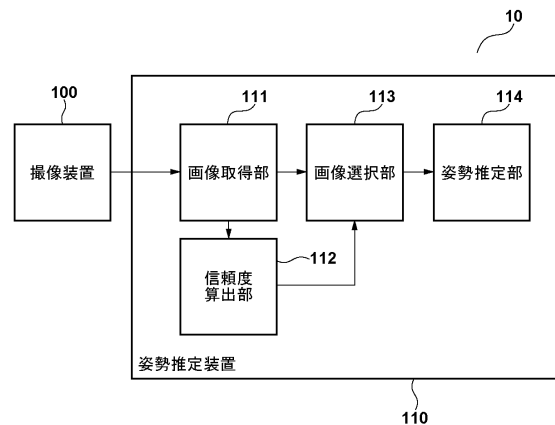
(54)【発明の名称】 生成装置、生成方法、およびプログラム

(57)【要約】

【課題】被写体の姿勢モデルを適切に生成する。

【解決手段】姿勢推定装置(110)は、複数の撮像装置(100)が被写体を異なる方向から撮像することに基づいて得られた複数の撮像画像から、該被写体の領域を抽出した複数の被写体画像を取得し、該複数の被写体画像について、該被写体に対する精度を示す信頼度を取得し、該信頼度に基づいて、該複数の被写体画像から1つ以上の被写体画像を選択し、該選択された1つ以上の被写体画像に基づいて、該被写体の3次元姿勢を示す姿勢モデルを生成する。

【選択図】図1



**【特許請求の範囲】****【請求項 1】**

複数の撮像装置が被写体を異なる方向から撮像することに基づいて得られた複数の撮像画像から、前記被写体の領域を抽出した複数の被写体画像を取得する第 1 の取得手段と、前記複数の被写体画像について、前記被写体に対する精度を示す信頼度を取得する第 2 の取得手段と、前記信頼度に基づいて、前記複数の被写体画像から 1 つ以上の被写体画像を選択する選択手段と、前記選択された 1 つ以上の被写体画像に基づいて、前記被写体の 3 次元姿勢を示す姿勢モデルを生成する生成手段と、  
を有することを特徴とする生成装置。

10

**【請求項 2】**

前記第 1 の取得手段は、前記複数の撮像画像のそれぞれについて、撮像画像における各画素が前記被写体の領域を構成する画素であるか否かを示す確率値を、画素ごとに導出し、当該確率値に基づいて、前記撮像画像における前記被写体画像を取得することを特徴とする請求項 1 に記載の生成装置。

**【請求項 3】**

前記第 1 の取得手段は、前記撮像画像における各画素の前記確率値を所定の閾値で 2 値化することによって、前記被写体画像を取得することを特徴とする請求項 2 に記載の生成装置。

20

**【請求項 4】**

前記第 1 の取得手段は、機械学習による結果に基づいて、前記撮像画像における各画素の前記確率値を導出することを特徴とする請求項 2 または 3 に記載の生成装置。

**【請求項 5】**

前記第 2 の取得手段は、前記撮像画像における各画素の前記確率値の分布に基づいて、前記被写体画像に対する前記信頼度を取得することを特徴とする請求項 2 から 4 のいずれか 1 項に記載の生成装置。

**【請求項 6】**

前記第 2 の手段は、前記確率値の分布において、確率値として取り得る最低値に近い確率値と最高値に近い確率値の割合が大きいほど、値が大きくなるように、前記信頼度を取得することを特徴とする請求項 5 に記載の生成装置。

30

**【請求項 7】**

前記第 1 の取得手段は、前記複数の撮像画像のそれぞれについて、撮像画像の撮影領域に前記被写体が存在しない背景画像の画素値と、前記撮像画像の画素値との差分値を、画素ごとに導出し、当該差分値に基づいて、前記被写体画像を取得することを特徴とする請求項 1 に記載の生成装置。

**【請求項 8】**

前記第 2 の取得手段は、前記撮像画像における前記差分値が大きい値をとるほど、値が大きくなるように、前記信頼度を取得することを特徴とする請求項 7 に記載の生成装置。

**【請求項 9】**

前記選択手段は、前記信頼度と前記複数の撮像装置の配置を示す情報に基づいて、前記 1 つ以上の被写体画像を選択することを特徴とする請求項 1 から 8 のいずれか 1 項に記載の生成装置。

40

**【請求項 10】**

複数の撮像装置が被写体を異なる方向から撮像することに基づいて得られた複数の撮像画像から、前記被写体の領域を抽出した複数の被写体画像を取得する第 1 の取得工程と、前記複数の被写体画像のそれぞれについて、前記被写体に対する精度を示す信頼度を取得する第 2 の取得工程と、前記信頼度に基づいて、前記複数の被写体画像から 1 つ以上の被写体画像を選択する選択工程と、

50

前記選択された1つ以上の被写体画像に基づいて、前記被写体の3次元姿勢を示す姿勢モデルを生成する生成工程と、  
を有することを特徴とする生成方法。

【請求項11】

コンピュータを、請求項1から9のいずれか1項に記載の生成装置として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、人物の姿勢モデル生成技術に関する。

10

【背景技術】

【0002】

複数の異なる位置に設定された複数の撮像装置（多視点カメラ）により得られた画像（多視点カメラ画像）を用いて、被写体（人体）の姿勢モデルを推定する技術（姿勢推定技術）が注目されている。姿勢モデルは、被写体を構成する関節の位置、関節同士の接続関係、被写体の部位間の角度などを表す情報である。姿勢推定技術は、多視点カメラ画像に写る被写体の関節位置を推定することで、肘や膝などの角度などを推定できる。さらに、推定された姿勢モデルを用いることで、スポーツ選手の運動量や疲労度の評価、新旧のフォーム比較のような運動解析が可能になる。特許文献1では、多視点カメラを用いて被写体を撮像し、得られた多視点カメラ画像から被写体領域の画像（被写体画像）を抽出し、該被写体画像から該被写体の3次元関節位置を持つ姿勢モデルを推定している。

20

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2016-126425号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

特許文献1では、被写体一人が写る被写体画像における特徴点を用いて該被写体の形状モデルを推定し、該形状モデルから姿勢モデルを推定している。しかしながら、姿勢モデルを精度よく推定するためには、対象となる被写体を撮像した複数のカメラから、より高精度に被写体部分を抽出したカメラを選択する必要がある。さらに、多数のカメラに同じ被写体が写る場合、これらのカメラからの多視点カメラ画像の全てを用いて姿勢モデルを推定すると処理時間の増加を招く。

30

【0005】

本発明は、上記の課題に鑑みてなされたものであり、被写体の姿勢モデルを適切に生成することを目的とする

【課題を解決するための手段】

【0006】

上記目的を達成するための一手段として、本発明の生成装置は以下の構成を有する。すなわち、複数の撮像装置が被写体を異なる方向から撮像することに基づいて得られた複数の撮像画像から、前記被写体の領域を抽出した複数の被写体画像を取得する第1の取得手段と、前記複数の被写体画像について、前記被写体に対する精度を示す信頼度を取得する第2の取得手段と、前記信頼度に基づいて、前記複数の被写体画像から1つ以上の被写体画像を選択する選択手段と、前記選択された1つ以上の被写体画像に基づいて、前記被写体の3次元姿勢を示す姿勢モデルを生成する生成手段と、を有する。

40

【発明の効果】

【0007】

本発明によれば、被写体の姿勢モデルを適切に生成することができる。

【図面の簡単な説明】

50

## 【 0 0 0 8 】

【図 1】実施形態 1 における画像処理システムの構成例を示す図である。

【図 2】実施形態 1 における標準形状モデルと姿勢モデルを示す模式図である。

【図 3】実施形態 1 における形状推定装置のハードウェア構成例を示す図である。

【図 4】実施形態 1 における撮影対象とカメラ配置の例を示す図である。

【図 5】実施形態 1 における姿勢推定装置により実行される処理のフローチャートである。

。

【図 6】実施形態 1 における前景確率画像と前景確率のヒストグラムの例を示す図である。

。

【図 7】図 6 におけるヒストグラムと 2 クラス分類した結果を示す図である。

10

【図 8】実施形態 1 における複数のカメラに対して設定したカメラグループを示す図である。

【図 9】実施形態 1 における各カメラ（撮像装置）の信頼度、撮影角度およびカメラグループを示した表である。

【発明を実施するための形態】

## 【 0 0 0 9 】

以下、添付図面を参照して実施形態を詳しく説明する。なお、以下の実施形態は特許請求の範囲に係る発明を限定するものではない。実施形態には複数の特徴が記載されているが、これらの複数の特徴の全てが発明に必須のものとは限らず、また、複数の特徴は任意に組み合わせられてもよい。さらに、添付図面においては、同一若しくは同様の構成に同一の参照番号を付し、重複した説明は省略する。

20

## 【 0 0 1 0 】

<実施形態 1>

本実施形態では、姿勢推定に用いる被写体画像を、機械学習の結果（学習済みモデル）を用いて抽出するケースにおいて、被写体抽出時の前景に対する確率値を用いて姿勢推定に用いる画像を選択し、被写体の姿勢モデルを推定する方法について述べる。本実施形態の画像処理システムは、複数の撮像装置が異なる方向から撮像することにより得られた複数の画像に基づいて、被写体の姿勢モデルを推定する。

## 【 0 0 1 1 】

本実施形態における被写体の姿勢推定モデルは、後述するように、一時的な形状モデルを取得した後に、該形状モデルに基づいて推定し生成される。

30

形状モデルは、例えば被写体が人物である場合、被写体のシルエット、輪郭を示し、点群や複数のボクセルで表現されてもよい。また、形状モデルは、複数のポリゴンを含むポリゴンデータとして表現されてもよい。

姿勢モデルは、例えば被写体が人物である場合、その人物の関節位置を表す点と、骨格を表す線とで表現されてもよい。つまり、この場合、姿勢モデルは、複数の点と、2 点間を結ぶ線と、を含んでもよい。姿勢モデルは、これに限定されず、それ以外の表現であってもよく、関節位置を示す点のみで表現されていてもよい。また、すべての関節位置が表現されていなくてもよく、一部の関節位置が表現されていてもよい。また、姿勢モデルが、被写体の 3 次元姿勢を表すものであれば、必ずしも関節位置を点で表現しなくても、いくつか又はすべての関節位置に代えて特徴的な部位を点で表現してもよい。また、人物の顔などの輪郭については、複数の点と直線あるいは曲線で表現してもよいし、球や楕円球で表現してもよい。

40

## 【 0 0 1 2 】

[ 画像処理システムの構成 ]

本実施形態における画像処理システムの構成例を図 1 に示す。本実施形態における画像処理システム 10 は、撮像装置 100 と姿勢推定装置 110 とを含む。なお、図 1 には 1 台の撮像装置 100 を示すが、同様の構成の複数の撮像装置 100 が、無線または有線の接続で姿勢推定装置 110 に接続されているものとする。また、以下の説明において、「撮像装置」は、「カメラ」と同義に用いられるものとする。

50

## 【 0 0 1 3 】

複数の撮像装置 1 0 0 は、多数の異なる方向から撮像領域を撮像する、多視点カメラを構成し、異なる方向から撮像した複数の画像（多視点カメラ画像）を生成・取得する。撮像領域は、例えば、スポーツが行われる競技場の平面と任意の高さで囲まれた領域である。各撮像装置 1 0 0 は、撮像領域を取り囲むようにそれぞれ異なる位置・異なる方向に設置され、同期して撮像を行う。なお、各撮像装置 1 0 0 は撮像領域の全周にわたって設置されなくてもよく、設置場所の制限等によっては撮像領域の一部の方向にのみ設置されてもよい。多視点カメラを構成する撮像装置 1 0 0 の数は限定されず、例えば撮像領域をサッカーやラグビーの競技場とする場合、競技場の周囲に数十～数百台程度の撮像装置 1 0 0 が設置されてもよい。また、望遠カメラと広角カメラなど画角が異なるカメラが撮像装置 1 0 0 として設置されてもよい。撮像装置 1 0 0 は、現実世界の 1 つの時刻情報で同期され、撮像した画像（映像）には毎フレームの画像に撮像時刻情報が付与される。

10

## 【 0 0 1 4 】

さらに、撮像装置 1 0 0 は、自装置の位置、姿勢（向き、撮像方向）、焦点距離、光学中心、歪みなどの状態情報を取得し、管理する。撮像装置 1 0 0 の位置、姿勢（向き、撮像方向）は、撮像装置 1 0 0 自身によって制御されてもよいし、撮像装置 1 0 0 の位置や姿勢を制御する雲台によって制御されてもよい。以下では、撮像装置 1 0 0 の状態情報をカメラパラメータとして説明を行うが、そのパラメータには、雲台等の別の装置により制御されるパラメータ（各種情報）が含まれていてもよい。撮像装置 1 0 0 の位置、姿勢（向き、撮像方向）に関するカメラパラメータは、いわゆる外部パラメータであり、撮像装置 1 0 0 の焦点距離、画像中心、歪みに関するパラメータは、いわゆる内部パラメータである。

20

## 【 0 0 1 5 】

姿勢推定装置 1 1 0 は、複数の撮像装置 1 0 0 から得られた複数の画像から、被写体の姿勢モデルを推定して生成する生成装置として機能する。姿勢推定装置 1 1 0 は、被写体の姿勢モデルを、例えば、次のような方法で推定して生成する。まず、姿勢推定装置 1 1 0 は、複数の撮像装置 1 0 0 から、複数の撮像装置 1 0 0 が異なる方向から撮像することにより得られた複数の撮像画像（多視点カメラ画像）を取得する。次に、姿勢推定装置 1 1 0 は、撮像画像から、人物などの被写体に対応する前景領域を抽出した被写体画像を取得（生成）する。被写体画像とは、撮像装置 1 0 0 により撮像されて取得された撮像画像から、被写体領域（前景領域）を抽出した画像である。前景領域として抽出される被写体とは、一般的に、時系列で同じ方向から撮像を行った場合において動きのある（その位置や形が変化し得る）動的被写体（動体）を指す。被写体は、例えば、ある競技において、それが行われるフィールド内にいる選手や審判などの人物、球技であれば人物に加えボールなども含む。また、コンサートやエンタテインメントにおいては、歌手、演奏者、パフォーマー、司会者などが被写体である。

30

## 【 0 0 1 6 】

本実施形態では、姿勢モデルを推定するために、標準的な人の形を模した 3 次元の標準形状モデルとその姿勢モデル（初期姿勢モデル）が、あらかじめ姿勢推定装置 1 1 0 に入力されるものとする。当該標準形状モデルと初期姿勢モデルは、姿勢推定用パラメータとして所定の 1 つ以上のファイルに記載され、あらかじめ姿勢推定装置 1 1 0 に入力され、ROM 3 1 2 等の記憶手段に格納されうる。

40

## 【 0 0 1 7 】

図 2 に、姿勢推定装置 1 1 0 に入力される標準形状モデル 2 0 0 と初期姿勢モデル 2 1 0 の一例を示す。これらのモデルは CG（Computer Graphic）でも用いられる一般的なモデルで良く、ファイル形式（フォーマット）は FBX 形式など一般的な形式で良い。標準形状モデル 2 0 0 は、例えば、3 次元のメッシュモデルで表現され、頂点座標と三角形もしくは四角形の面を構成する頂点 ID の情報が含まれる。初期姿勢モデル 2 1 0 は、頭部や首、臍、肩、肘、手首、足の付け根、膝、足首のような人体の主要部位や関節部位の位置を表す情報 2 1 1 とそれらの接続関係を示す情報 2 1 2、隣接部位間の

50

角度情報が含まれる。姿勢モデルの部位とメッシュモデルの部位とを対応付けておくことで、姿勢モデルの右腕を回転すれば、対応するメッシュモデルの部位も回転できる。姿勢推定では、このような姿勢モデルに対応したメッシュモデルを、各カメラ画像（各カメラの撮像画像）に射影した画像と各カメラの前景領域とが一致するように変形させ、最も一致した際の姿勢を被写体の姿勢モデルとして推定する。ただし、3次元の姿勢モデルを推定する方法はこれに限定されない。例えば、2次元画像上で2次元の姿勢を推定し、各カメラとの対応に基づいて、3次元の姿勢モデルを推定する方法など、種々の方法を用いてもよい。

#### 【0018】

##### [ 姿勢推定装置の構成 ]

次に、姿勢推定装置110の構成について説明する。まず、姿勢推定装置110の内部構成について、図1を参照して説明する。姿勢推定装置110は、画像取得部111、信頼度算出部112、画像選択部113、姿勢推定部114を有する。

#### 【0019】

画像取得部111は、複数の撮像装置100から、異なる方向から撮像することによって得られた複数の撮像画像を取得する。そして、画像取得部111は、該撮像画像のそれぞれから、撮像画像に含まれる被写体領域（前景画像）を抽出した被写体画像を取得する。画像取得部111は、このようにして、複数の被写体画像を取得する。被写体領域の抽出には、画像内に含まれる特定オブジェクト（サッカーの試合であれば、選手やボール）を前景領域として分離（抽出・検出）する前景背景分離手法を用いることができる。この手法には種々の方法があるが、本実施形態では、機械学習による前景背景分離手法を用いる例について説明する。

#### 【0020】

機械学習を用いた前景背景分離手法では、前景領域を分離したいシーンに類似したシーンの撮影画像と、その中に含まれる前景領域をマーキングしたシルエット画像との組である学習データを大量に準備し、ニューラルネットワークに入力し学習する。学習データとして用いるシルエット画像は、撮像画像内の前景領域を人が判断してマーキングしてもよいし、背景差分法など別の前景背景分離手法を用いて前景領域をマーキングすることで生成してもよい。大量の学習データに含まれる前景領域の特徴を統計的に学習することで、分離対象である撮像画像における各画素が、前景領域に対応するオブジェクトを構成する画素であるか否かを、画素ごとの確率値（前景確率値）として導出する。そして、画像取得部111は、前景確率値を所定の閾値で2値化することで、撮像画像に含まれる前景領域（被写体領域）に対応した被写体画像（マスク画像）を生成する。また、画像取得部111は、撮像画像に含まれる各画素が前景確率値を有する画像である前景確率画像を生成し、さらに、該前景確率画像に基づいて、前景確率の分布状況（例えばヒストグラム）の情報を生成する。該前景確率の分布状況の情報は、信頼度算出部112で使用される。

#### 【0021】

信頼度算出部112は、画像取得部111で生成した、前景確率の分布状況に基づいて、抽出した被写体領域の前景領域に対する信頼度を算出して取得する。この信頼度は、導出した前景確率値に基づく被写体画像がどれほど実際の前景領域に対応するかを示す（すなわち、実際の被写体に対する精度を示す、あるいは被写体の領域の抽出精度（確からしさ）を示す）値（指標）を表す。以下、この信頼度を、単に信頼度と称する。信頼度の算出方法については、動作フローの説明（図5）において後述する。

#### 【0022】

画像選択部113は、信頼度算出部112で算出された信頼度に基づき、姿勢推定に用いる1つ以上の被写体画像を選択する。被写体画像の選択においては、信頼度の他に、撮像装置100の配置を考慮してもよい。姿勢推定に用いるアルゴリズムにより、被写体画像を選択基準が変わる。例えば、なるべくいろいろな方向から撮影した被写体画像を準備した方がよいアルゴリズムの場合であれば、被写体画像の信頼度とともに撮像装置100の配置も考慮して、被写体画像を選択してもよい。

10

20

30

40

50

## 【0023】

姿勢推定部114は、選択した1つ以上の被写体画像と、被写体が写る撮像装置100のカメラパラメータを用いて、被写体の姿勢モデルを推定して生成する。

## 【0024】

次に、姿勢推定装置110のハードウェア構成について説明する。図3に、姿勢推定装置110のハードウェア構成例を示す。姿勢推定装置110は、CPU(Central Processing Unit)311、ROM(Read Only Memory)312、RAM(Random Access Memory)313、補助記憶装置314、表示部315、操作部316、通信I/F(インタフェース)317、及びバス318を有する。CPU311は、ROM312やRAM313に格納されているコンピュータプログラムやデータを用いて姿勢推定装置110の全体を制御することで、図1に示す姿勢推定装置110の各機能を実現する。なお、姿勢推定装置110がCPU311とは異なる1又は複数の専用のハードウェアを有し、CPU311による処理の少なくとも一部を専用のハードウェアが実行してもよい。専用のハードウェアの例としては、ASIC(特定用途向け集積回路)、FPGA(フィールドプログラマブルゲートアレイ)、およびDSP(デジタルシグナルプロセッサ)などがある。ROM312は、変更を必要としないプログラムなどを格納する。RAM313は、補助記憶装置314から供給されるプログラムやデータ、及び通信I/F317を介して外部から供給されるデータなどを一時記憶する。補助記憶装置314は、例えばハードディスクドライブ等で構成され、画像データや音声データなどの種々のデータを記憶する。

10

20

## 【0025】

表示部315は、例えば液晶ディスプレイやLED(Light Emitting Diode)等で構成され、ユーザが姿勢推定装置1を操作するためのGUI(Graphical User Interface)などを表示する。操作部316は、例えばキーボードやマウス、ジョイスティック、タッチパネル等で構成され、ユーザによる操作を受けて各種の指示をCPU311に入力する。CPU311は、表示部315を制御する表示制御部、及び操作部316を制御する操作制御部として動作する。

## 【0026】

通信I/F317は、姿勢推定装置110の外部の装置との通信に用いられる。例えば、姿勢推定装置110が外部の装置と有線で接続される場合には、通信用のケーブルが通信I/F317に接続される。姿勢推定装置110が外部の装置と無線通信する機能を有する場合には、通信I/F317はアンテナ(不図示)を備える。バス318は、姿勢推定装置110の各部をつないで情報を伝達する。

30

## 【0027】

本実施形態では、表示部315と操作部316が姿勢推定装置110の内部に存在するものとするが、表示部315と操作部316との少なくとも一方が姿勢推定装置110の外部に別の装置として存在していてもよい。

## 【0028】

## [動作フロー]

続いて、姿勢推定装置110の動作について説明する。まず、本実施形態における撮影対象とカメラ配置について説明する。図4に、本実施形態における撮影対象とカメラ配置の例を示す。なお、図4並びに後述する図8における参照符号について、番号の後ろに付与したアルファベットのみが異なる符号は、同一機能を持つ装置の別インスタンスを示すものとする。例えば、図4(a)において、カメラ403Aとカメラ403Bは同一の機能を持つ別インスタンスを示している。なお、同一の機能を持つとは、少なくとも特定の機能(撮像機能など)を有することを指すものであり、例えばカメラ403Aとカメラ403Bが有する機能及び性能の一部が異なってもよい。また、以下の説明において、カメラ403は、カメラ403A~403Hの総称を表す。

40

## 【0029】

図4(a)は、撮影対象(被写体の一例)を含む撮影領域とそれを取り囲むように配置し

50

たカメラ403A～403Hを示している。撮影領域401と、該撮影領域401に存在する撮影対象402を取り囲んで撮影するように、カメラ403A～403Hが配置されている。図4(b)は、撮影領域とカメラ配置の位置関係を撮影領域の真上から見下ろした平面上に図示したものである。図4(b)では、撮影対象402を撮影する各カメラの撮影対象402に対する相対的な撮影角度を、カメラ403Bを例として示している。具体的には、カメラ403Bの撮影対象402に対する相対的な撮影角度を、角度基準404と、カメラ403Bと撮影対象を結んだ直線405Bとのなす角度406Bとして示している。このようにしてカメラ403～403Hに対して導出される撮影角度は、カメラの配置を示す情報の一例であり、後述する姿勢推定に用いるカメラを選択する処理において、必要に応じて参照されうる。

10

#### 【0030】

図5は、姿勢推定装置110により実行される処理のフローチャートである。図5に示すフローチャートは、姿勢推定装置110のCPU311がROM312やRAM313に格納されている制御プログラムを実行し、情報の演算および加工並びに各ハードウェアの制御を実行することにより実現されうる。

#### 【0031】

S510において、姿勢推定部114は、姿勢推定用パラメータが記載されたファイルを読み込む。該姿勢推定用パラメータは、図2を用いて前述した姿勢モデルを含む。該姿勢モデルは初期姿勢として姿勢推定(後述するS550)に用いられる。

#### 【0032】

S520において、画像取得部111は被写体画像を取得する。本実施形態では前述の通り、被写体画像(前景画像)の取得方法として機械学習による手法を用いるものとする。また、画像取得部111は、前述の通り、該機械学習を通じて、前景確率画像(各画素が前景領域である確率値(前景確率値)を有する画像)を生成する。前景確率画像は、被写体領域を含む任意のサイズの画像であり得る。さらに、画像取得部111は、生成した前景確率画像に基づいて、前景確率の分布状況の情報として、前景確率のヒストグラムを生成する。

20

#### 【0033】

図6に、本ステップにおいて生成される前景確率画像と前景確率のヒストグラムの例を示す。図6(a)は、前景確率の高低の区別が明確である例を示し、図6(b)は、前景確率の高低の区別が明確でない例を示す。結果として、図6(a)は、被写体領域が比較的上手にできた例であり、図6(b)は比較的上手に検出できなかった例に対応する。

30

#### 【0034】

図6(a)における前景確率画像601は、画素ごと前景確率値を有し、前景確率値の高い画素を白、低い画素を黒で表している。例えば、所定の第1の閾値より高い前景確率値は高いと判定し、所定の第2の閾値より低い前景確率値は、低いと判定することができる。ヒストグラム602は、これらの前景確率の頻度を集計した結果である。図6(a)のヒストグラムは、前景確率値の高い画素の頻度と低い画素の頻度が多く、前景確率値が中間値を取る画素がほとんどない分布となっている。

#### 【0035】

図6(b)における前景確率画像603は、前景確率画像601と同様に、前景確率の高い画素を白、前景確率値の低い画素を黒で表しており、前景確率値が中間値を取る画素を斜線で表している。例えば、前述の所定の第1の閾値と第2の閾値の間にある値を中間値とする。ヒストグラム604は、これらの前景確率値の頻度を集計した結果である。図6(b)のヒストグラムは、前景確率が高い画素、低い画素とともに中間値を取る画素がある程度存在する分布となっている。

40

#### 【0036】

S530において、信頼度算出部112は、画像取得部111で生成した前景確率のヒストグラムを用いて、被写体画像の信頼度を算出して取得する。被写体画像の信頼度の算出方法の一例として、本実施形態では、前景確率のヒストグラムを周知の判別分析法(大津

50

の二値化)により2クラス分類して算出する手法について説明する。大津の二値化は、ある分布に対して、分離度 = (クラス間分散) / (クラス内分散) が最大となるように二値化閾値を決める手法である。分離度は2つのクラスがどれだけ離れた分布になっているかを示す値となっているため、この値が大きいほど、前景確率の中間値が少なく、前景確率の高い画素と低い画素の割合が大きい分布になっていると言える。分離度が大きいということは、機械学習による前景背景分離手法において、前景確率と前景でない確率 (= 背景である確率) がよりはっきりと区別出来ているということの意味し、前景画像がより鮮明に抽出出来ていることを意味する。

#### 【0037】

図6における前景確率のヒストグラム(ヒストグラム602、ヒストグラム604)を大津の二値化により2クラス分類した結果を図7に示す。図7(a)は、ヒストグラム602を2クラス分類した閾値701を、図7(b)は、ヒストグラム604を2クラス分類した閾値702を示している。それぞれの分離度をC1、C2とすると、C1 > C2の大小関係となる。これは、前述のように、前景確率の高い画素と低い画素の割合が大きい分布の場合に、分離度はより大きい値になるからである。本実施形態では、この分離度をそのまま前景画像の信頼度として用いる。よって、図6に示す前景確率のヒストグラム(分布状況)において、確率値として取り得る最低値に近い確率値と最高値に近い確率値の割合が大きいほど、分離度は大きくなり、信頼度は高くなる(大きくなる)。なお、前景画像の信頼度の算出方法はこれに限らず、任意の2クラス化手法と、前景クラスと判定した画素群と背景クラスと判定した画素群の分離度を表す任意の指標を採用することが可能である。

10

20

#### 【0038】

S540において、画像選択部113は姿勢推定に用いる被写体画像を、S530で算出された信頼度などを用いて選択する。信頼度の他に用いる指標は、そのシステムで用いる姿勢推定手法により決定されうる。例えば、姿勢推定対象の被写体をなるべく様々な方向から撮影した画像を用いた方がよい手法を用いる場合、信頼度の他にカメラの配置情報も考慮して被写体画像を選択してもよい。

#### 【0039】

このような、信頼度とカメラの配置情報を考慮した被写体画像の選択方法の一例を、図8と図9を用いて説明する。本例では、姿勢推定用の1つ以上の被写体画像として、カメラ403A~403Hから得られる複数の被写体画像から4つ選択する場合を想定する。カメラの配置のばらつきを持たせるために、図4(b)で説明した撮影角度を用いて各カメラを4つのグループに分類する。図8に、図4(b)のように配置されたカメラに対して設定したカメラグループの例(グループ1~4)を示す。本例では、グループ1~4のそれぞれにおいて、最も信頼度の高い被写体画像を取得できるカメラを選択する。これにより、被写体をなるべく様々な方向から撮影し、かつ信頼度の高い被写体画像を取得できるカメラを選択することが可能となる。各カメラに対する信頼度(本実施形態では、前述した分離度)、撮影角度およびグループを示した表を図9に示す。この表から、同一グループ内で信頼度を比較し、信頼度の高いカメラによる撮像画像から得られる被写体画像を選択することとなる。

30

40

#### 【0040】

S550において、姿勢推定部114は、選択された被写体画像を用いて姿勢を推定する。姿勢推定の方法は多視点カメラ情報と初期姿勢モデルを用いる既知の方法で良い。例えば、次のような方法がある。まず、初期姿勢モデルを変形させて、一時的な形状モデルを取得し、選択された前景画像に射影する。そして、射影された領域と前景画像との類似度を評価する。これらの処理を類似度が一定の閾値を満たすまで姿勢モデルを変えながら繰り返し、閾値を満たしたときの姿勢モデルを、最適な姿勢モデルと推定できる。また、連続的に運動する被写体を撮影した場合、初期姿勢モデルは1フレーム前の姿勢モデルでも良い。

#### 【0041】

50

このように、本実施形態によれば、多視点カメラに写る被写体の姿勢モデルを推定して生成する場合に、被写体部分の検出精度を考慮して、姿勢推定を行う上で最適な1つ以上の被写体画像を選択できる。その結果、姿勢推定の処理時間や誤差を低減できる。

【0042】

なお、本実施形態の説明は、各カメラは1つの同一の被写体を撮影している想定での説明となっているが、撮影領域に複数の被写体が存在する場合についても、本実施形態を適用可能である。その場合、例えば被写体ごとに本実施形態の処理を行えばよい。

【0043】

また、本実施形態では、被写体の撮影方向のばらつきを保証するために、角度に基づきカメラグループを設定する方法を用いたが、他の方法も可能である。例えば、信頼度とカメラ配置のばらつきの度合に基づいて算出する指標が最大化されるようにカメラを選択する方法でもよい。

10

【0044】

<実施形態1の変形例>

実施形態1では、被写体の前景画像を分離手法として、機械学習の結果に基づく手法を用いた場合を説明したが、他の手法を用いてもよい。例えば、前景画像を分離する手法として、予め作成した背景画像と撮影画像の画素値の差分から前景画像を抽出(生成)する背景差分による手法を用いることもできる。この場合、実施形態1で用いた前景確率のヒストグラムの代わりに、抽出した前景画像の周辺領域における撮影画像と背景画像の各画素の差分値のヒストグラムを用いることができる。撮影画像と背景画像の差分値の分離度が大きいということは、前景に対応する画素と背景に対応する画素の色輝度の差分がはっきりしているということであり、これにより前景境界に忠実に前景部分を分離出来たことを意味する。すなわち、抽出した前景画像の周辺領域における撮影画像において、差分値が大きい値をとるほど(もしくは、差分値として取り得る最大値に近い値を取る割合が多いほど)、信頼度は高くなる(大きくなる)。このように前景部分が忠実に分離出来た前景画像を選択することで、より姿勢推定に適した被写体画像を選択することができる。

20

【0045】

<その他の実施形態>

本発明は、上述の実施形態の1以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1以上の機能を実現する回路(例えば、ASIC)によっても実現可能である。

30

【0046】

発明は上記実施形態に制限されるものではなく、発明の精神及び範囲から離脱することなく、様々な変更及び変形が可能である。従って、発明の範囲を公にするために請求項を添付する。

【符号の説明】

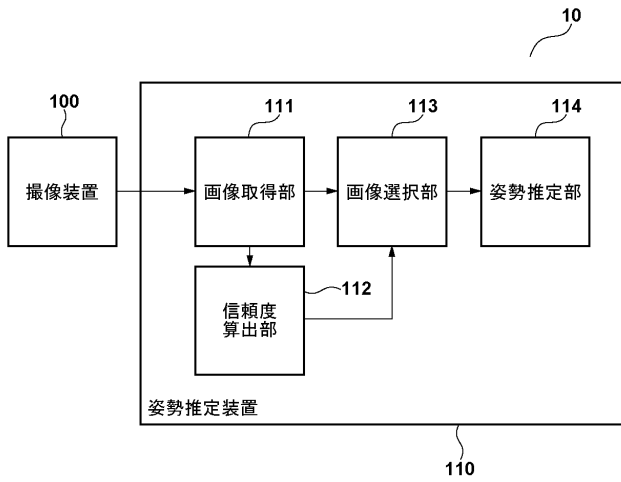
【0047】

100 撮像装置、110 姿勢推定装置、111 画像取得部、112 信頼度算出部、113 画像選択部、114 姿勢推定部

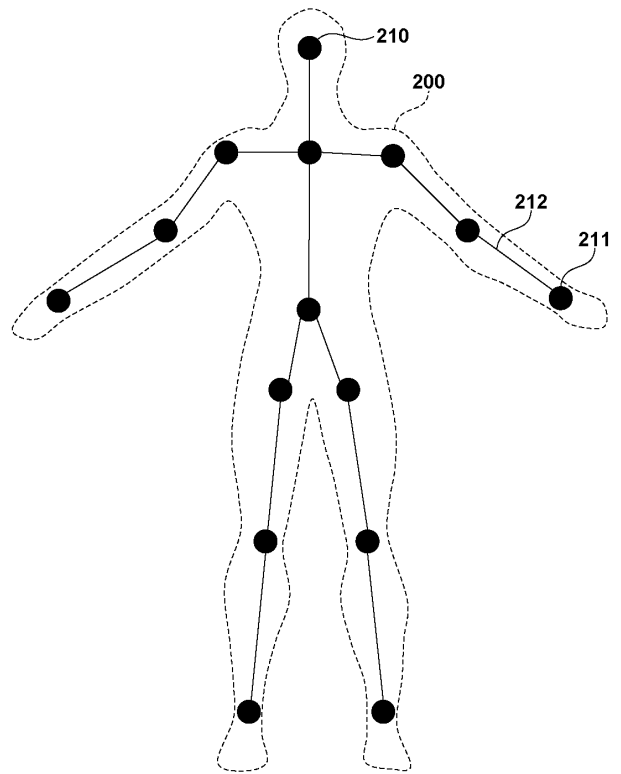
40

【 図面 】

【 図 1 】



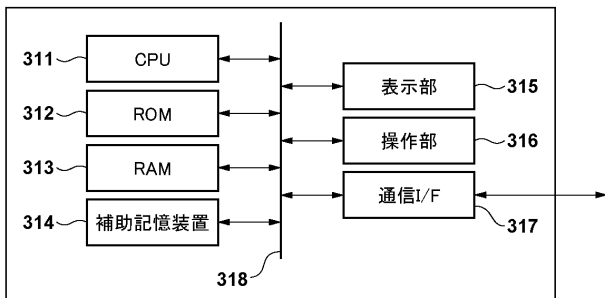
【 図 2 】



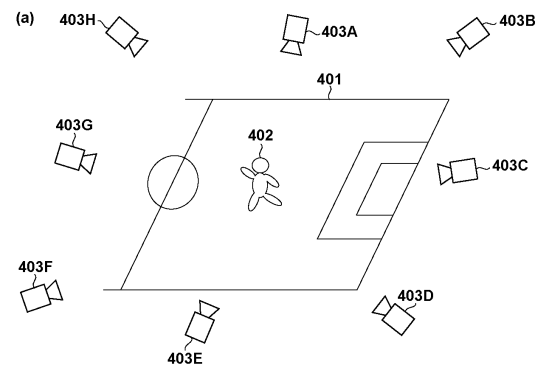
10

20

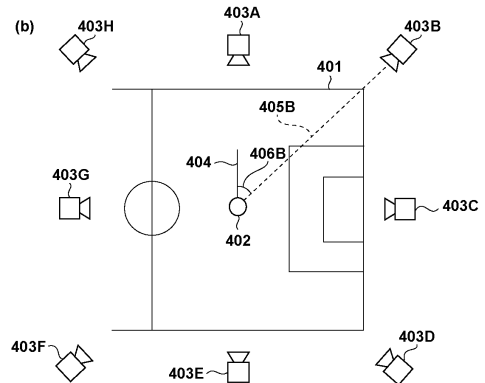
【 図 3 】



【 図 4 】



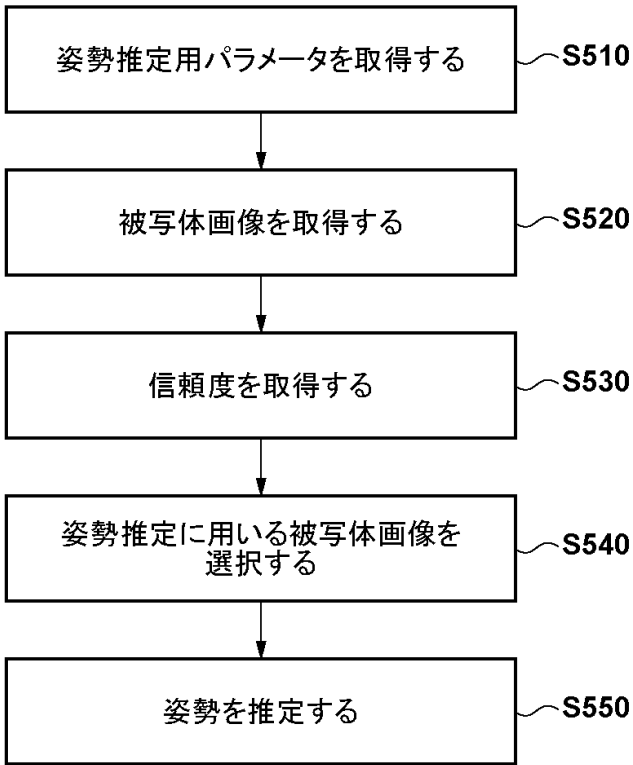
30



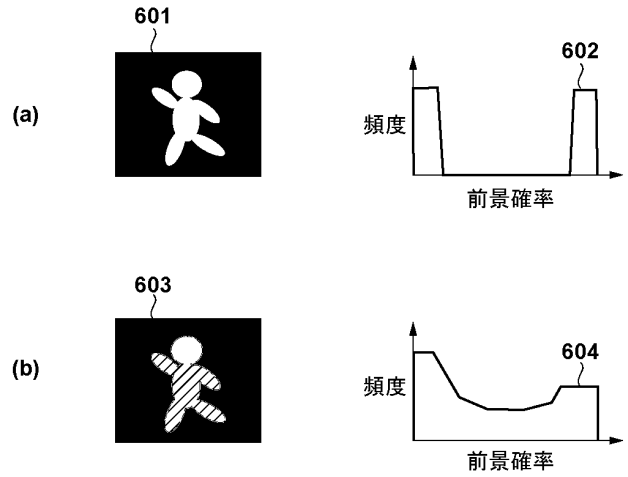
40

50

【図5】



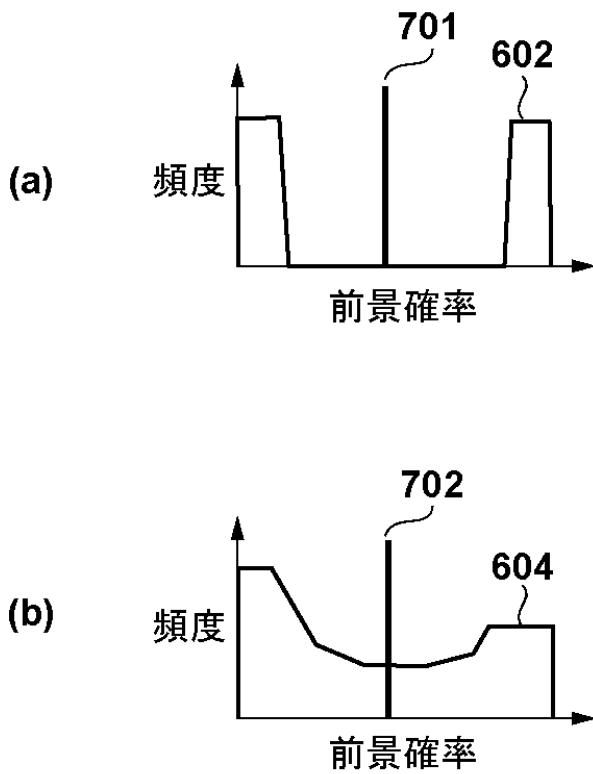
【図6】



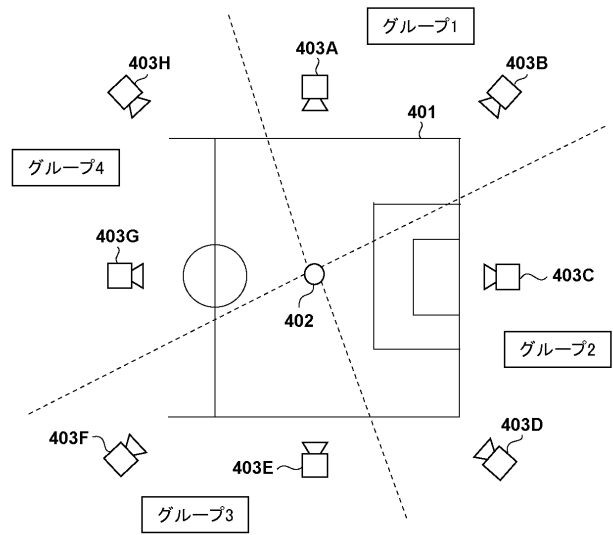
10

20

【図7】



【図8】



30

40

50

## 【 図 9 】

カメラ	信頼度	撮影角度	グループ
403A	Ca	Da	1
403B	Cb	Db	1
403C	Cc	Dc	2
403D	Cd	Dd	2
403E	Ce	De	3
403F	Cf	Df	3
403G	Cg	Dg	4
403H	Ch	Dh	4

10

20

30

40

50