

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号
特開2024-98820
(P2024-98820A)

(43)公開日 令和6年7月24日(2024.7.24)

(51)国際特許分類	F I	テーマコード (参考)
G 0 3 B 13/36 (2021.01)	G 0 3 B 13/36	2 H 0 0 2
G 0 3 B 15/00 (2021.01)	G 0 3 B 15/00	Q 2 H 0 1 1
G 0 2 B 7/28 (2021.01)	G 0 2 B 7/28	Z 2 H 1 5 1
G 0 3 B 7/091(2021.01)	G 0 3 B 7/091	5 C 1 2 2
H 0 4 N 23/67 (2023.01)	H 0 4 N 23/67	1 0 0
審査請求 未請求 請求項の数 16 O L (全33頁)		

(21)出願番号	特願2023-2558(P2023-2558)	(71)出願人	000001007
(22)出願日	令和5年1月11日(2023.1.11)		キヤノン株式会社
			東京都大田区下丸子3丁目30番2号
		(74)代理人	110003281
			弁理士法人大塚国際特許事務所
		(72)発明者	八島 浩
			東京都大田区下丸子3丁目30番2号
			キヤノン株式会社内
		(72)発明者	小椋 秀樹
			東京都大田区下丸子3丁目30番2号
			キヤノン株式会社内
		(72)発明者	杉谷 邦明
			東京都大田区下丸子3丁目30番2号
			キヤノン株式会社内
		(72)発明者	松井 遥平
			最終頁に続く

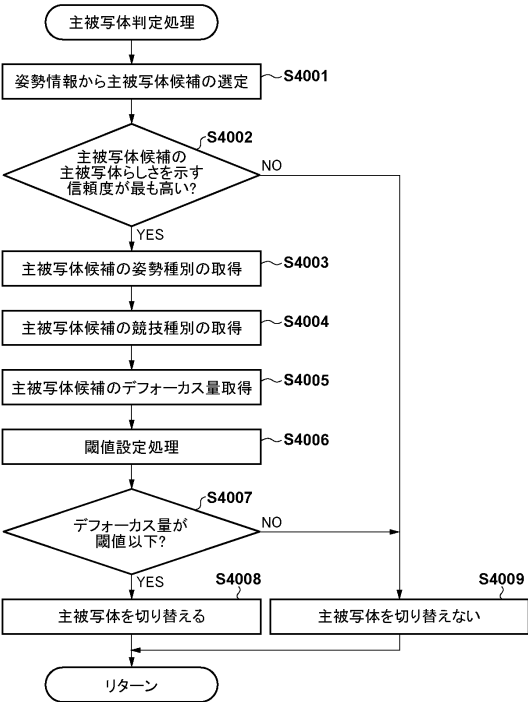
(54)【発明の名称】 撮像装置及びその制御方法、プログラム、記憶媒体

(57)【要約】

【課題】複数の被写体が存在する場合に、焦点を合わせるべき被写体を適切に選択することができる撮像装置を提供する。

【解決手段】被写体を検出する被写体検出手段と、被写体の姿勢を検出する姿勢検出手段と、被写体の焦点状態を検出する焦点検出手段と、被写体の姿勢と焦点状態とに基づいて、被写体を主被写体として選択するか否かを判定する閾値を設定する設定手段と、を備える。

【選択図】図15



【特許請求の範囲】**【請求項 1】**

被写体を検出する被写体検出手段と、
前記被写体の姿勢を検出する姿勢検出手段と、
前記被写体の焦点状態を検出する焦点検出手段と、
前記被写体の姿勢と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定する設定手段と、
を備えることを特徴とする撮像装置。

【請求項 2】

前記主被写体として選択された被写体に焦点を合わせる焦点調節手段をさらに備えることを特徴とする請求項 1 に記載の撮像装置。

10

【請求項 3】

前記被写体の姿勢に基づいて、前記被写体が主被写体である信頼性を取得する取得手段をさらに備えることを特徴とする請求項 1 に記載の撮像装置。

【請求項 4】

前記被写体の姿勢に基づいて、前記被写体の状態を判定する状態判定手段をさらに備え、前記設定手段は、前記被写体の状態と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定することを特徴とする請求項 1 に記載の撮像装置。

【請求項 5】

前記状態判定手段は、前記被写体の姿勢に基づいて前記被写体がどのような動作を行っているかの状態を判定することを特徴とする請求項 4 に記載の撮像装置。

20

【請求項 6】

前記被写体の姿勢に基づいて、前記被写体が行っている競技の種別を判定する競技判定手段をさらに備えることを特徴とする請求項 1 に記載の撮像装置。

【請求項 7】

前記設定手段は、前記競技の種別にさらに基づいて、前記閾値を設定することを特徴とする請求項 6 に記載の撮像装置。

【請求項 8】

前記設定手段は、前記被写体の姿勢継続時間にさらに基づいて、前記閾値を設定することを特徴とする請求項 1 に記載の撮像装置。

30

【請求項 9】

前記閾値は、前記被写体のデフォーカス量に対する閾値であることを特徴とする請求項 8 に記載の撮像装置。

【請求項 10】

前記被写体のデフォーカス量が前記閾値以下である場合に、前記被写体を前記主被写体として選択する選択手段をさらに備えることを特徴とする請求項 9 に記載の撮像装置。

【請求項 11】

前記設定手段は、前記姿勢継続時間が長いほど、前記閾値を大きく設定することを特徴とする請求項 8 に記載の撮像装置。

40

【請求項 12】

前記撮像装置のパンニングまたはチルティングを検出する検出手段をさらに備えることを特徴とする請求項 1 に記載の撮像装置。

【請求項 13】

前記パンニングまたはチルティングにより画面の中央に近づいた被写体を主被写体として選択する第 2 の選択手段をさらに備えることを特徴とする請求項 12 に記載の撮像装置。

【請求項 14】

被写体を検出する被写体検出工程と、
前記被写体の姿勢を検出する姿勢検出工程と、

50

前記被写体の焦点状態を検出する焦点検出工程と、

前記被写体の姿勢と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定する設定工程と、
を有することを特徴とする撮像装置の制御方法。

【請求項 15】

請求項 14 に記載の制御方法の各工程をコンピュータに実行させるためのプログラム。

【請求項 16】

請求項 14 に記載の制御方法の各工程をコンピュータに実行させるためのプログラムを記憶したコンピュータが読み取り可能な記憶媒体。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、撮像装置に関するものである。

【背景技術】

【0002】

複数の撮影を連続して行う連続撮影において、撮像装置により検出された動きのある複数の被写体を撮影する場合に、複数の被写体から主被写体を判定し、主被写体に対して焦点を合わせ続ける必要がある。主被写体を判定する方法として、2つの異なる種類の被写体を検出して主被写体を判定し、焦点調節する方法が特許文献 1 に開示されている。

【先行技術文献】

20

【特許文献】

【0003】

【特許文献 1】 特開 2018-66889 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、特許文献 1 に開示されているように、2つの異なる種類の被写体の位置関係から、主被写体を判定する場合には、撮影者の意思とは異なる被写体を主被写体と判定し、焦点を合わせ続けてしまう場合がある。また、主被写体が入れ替わることで焦点位置が急変化してしまう場合もある。

30

【0005】

本発明は上述した課題に鑑みてなされたものであり、その目的は、複数の被写体が存在する場合に、焦点を合わせるべき被写体を適切に選択することができる撮像装置を提供することである。

【課題を解決するための手段】

【0006】

本発明に係わる撮像装置は、被写体を検出する被写体検出手段と、前記被写体の姿勢を検出する姿勢検出手段と、前記被写体の焦点状態を検出する焦点検出手段と、前記被写体の姿勢と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定する設定手段と、を備えることを特徴とする。

40

【発明の効果】

【0007】

本発明によれば、複数の被写体が存在する場合に、焦点を合わせるべき被写体を適切に選択することが可能となる。

【図面の簡単な説明】

【0008】

【図 1】 本発明の実施形態であるカメラの構成を示すブロック図。

【図 2】 実施形態のカメラにおける画素配列を示す図。

【図 3】 実施形態における画素の平面図と断面図。

【図 4】 実施形態における画素構造の説明図。

50

- 【図 5】実施形態における瞳分割の説明図。
【図 6】実施形態におけるデフォーカス量と像ずれ量の関係を示す図。
【図 7】実施形態における焦点検出領域を示す図。
【図 8】実施形態における A F ・撮像処理を説明するフローチャート。
【図 9】実施形態における撮影処理を説明するフローチャート。
【図 10】実施形態における被写体追従 A F 処理を説明するフローチャート。
【図 11】実施形態における姿勢情報の説明図
【図 12 A】実施形態における主被写体判定処理を説明する概念図。
【図 12 B】実施形態における主被写体判定処理を説明する概念図。
【図 12 C】実施形態における主被写体判定処理を説明する概念図。
【図 13 A】実施形態における主被写体判定処理を説明する概念図。
【図 13 B】実施形態における主被写体判定処理を説明する概念図。
【図 13 C】実施形態における主被写体判定処理を説明する概念図。
【図 14】実施形態における被写体検出・追尾処理を説明するフローチャート。
【図 15】実施形態における主被写体判定処理を説明するフローチャート。
【図 16】実施形態におけるニューラルネットワークの構造の一例を示す図。
【図 17 A】実施形態における主被写体判定処理を説明する概念図。
【図 17 B】実施形態における主被写体判定処理を説明する概念図。
【図 18 A】実施形態における主被写体判定処理を説明する概念図。
【図 18 B】実施形態における主被写体判定処理を説明する概念図。
【図 19】第 2 の実施形態における主被写体判定処理を説明するフローチャート。
【図 20 A】第 2 の実施形態における主被写体判定処理を説明する概念図。
【図 20 B】第 2 の実施形態における主被写体判定処理を説明する概念図。
【図 20 C】第 2 の実施形態における主被写体判定処理を説明する概念図。
【発明を実施するための形態】

【0009】

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。

【0010】

（第 1 の実施形態）

図 1 は、本発明の撮像装置の第 1 の実施形態であるデジタルカメラ（以下、カメラ）100 の構成を示す図である。

【0011】

図 1 において、第 1 レンズ群 101 は、結像光学系としての撮像光学系のうち最も被写体側（前側）に配置され、光軸方向に移動可能に保持されている。絞り 102 は、その開口径を調節することで光量調節を行う。第 2 レンズ群 103 は、絞り 102 と一体となって光軸方向に移動し、光軸方向に移動する第 1 レンズ群 101 とともに変倍（ズーム）を行う。

【0012】

第 3 レンズ群（フォーカスレンズ）105 は、光軸方向に移動して焦点調節を行う。光学ローパスフィルタ 106 は、撮像画像の偽色やモアレを軽減するための光学素子である。第 1 レンズ群 101、絞り 102、第 2 レンズ群 103、第 3 レンズ群 105 および光学ローパスフィルタ 106 により撮像光学系が構成される。

【0013】

ズームアクチュエータ 111 は、不図示のカム筒を光軸回りで回転させることで、カム筒に設けられたカムによって第 1 レンズ群 101 および第 2 レンズ群 103 を光軸方向に移動させて変倍を行わせる。絞りアクチュエータ 112 は、絞り 102 の光量調節動作の

ために不図示の複数の遮光羽根を開閉方向に駆動する。フォーカスアクチュエータ 114 は、第 3 レンズ群 105 を光軸方向に移動させて焦点調節を行わせる。

【0014】

焦点調節手段としてのフォーカス駆動回路 126 は、カメラ CPU 121 からのフォーカス駆動指令に応じてフォーカスアクチュエータ 114 を駆動し、第 3 レンズ群 105 を光軸方向に移動させる。絞り駆動回路 128 は、カメラ CPU 121 からの絞り駆動指令に応じて絞りアクチュエータ 112 を駆動する。ズーム駆動回路 129 は、ユーザのズーム操作に応じてズームアクチュエータ 111 を駆動する。

【0015】

なお、本実施形態では、撮像光学系、アクチュエータ 111, 112, 114 および駆動回路 126, 128, 129 が撮像素子 107 を含むカメラ本体と一体に設けられている場合について説明する。しかし、撮像光学系、アクチュエータ 111, 112, 114 および駆動回路 126, 128, 129 を有する交換レンズがカメラ本体に着脱可能であってもよい。

【0016】

電子フラッシュ 115 は、キセノン管や LED 等の発光素子を有し、被写体を照明する光を発する。AF (オートフォーカス) 補助光発光部 116 は、LED 等の発光素子を有し、所定の開口パターンを有するマスクの像を投光レンズを介して被写体に投射することで、暗いまたは低コントラストの被写体に対する焦点検出性能を向上させる。電子フラッシュ制御回路 122 は、撮像動作に同期して電子フラッシュ 115 を点灯させるように制御する。補助光駆動回路 123 は、焦点検出動作に同期して AF 補助光発光部 116 を点灯させるように制御する。

【0017】

カメラ CPU 121 は、カメラ 100 における各種制御を司る。カメラ CPU 121 は、演算部、ROM、RAM、A/D コンバータ、D/A コンバータおよび通信インターフェイス回路等を有する。カメラ CPU 121 は、ROM に記憶されたコンピュータプログラムに従ってカメラ 100 内の各種回路を駆動したり、AF、撮像、画像処理および記録等の一連の動作を制御したりする。カメラ CPU 121 は、画像処理装置としても機能する。

【0018】

撮像素子 107 は、複数の画素を含む 2 次元 CMOS フォトセンサとその周辺回路からなり、撮像光学系の結像面に配置されている。撮像素子 107 は、撮像光学系により形成される被写体像を光電変換する。撮像素子駆動回路 124 は、撮像素子 107 の動作を制御するとともに、光電変換により生成したアナログ信号を A/D 変換して、デジタル信号をカメラ CPU 121 に送信する。

【0019】

シャッター 108 は、フォーカルプレーンシャッターの構成を有しており、カメラ CPU 121 からの指示に基づき、シャッター 108 に内蔵されているシャッター駆動回路からの指令で、フォーカルプレーンシャッターの駆動を行う。撮像素子 107 の信号の読み出し中には、撮像素子 107 を遮光する。また、露光が行われているときには、フォーカルプレーンシャッターが開かれ、撮像素子 107 に撮影光束が導かれる。

【0020】

画像処理回路 125 は、カメラ CPU 121 内の RAM に蓄積された画像データに対して予め定められた画像処理を適用する。画像処理回路 125 が適用する画像処理には、ホワイトバランス調整処理、色補間 (デモザイク) 処理、ガンマ補正処理といった所謂現像処理のほか、信号形式変換処理、スケーリング処理などがあるが、これらに限定されない。さらに画像処理回路 125 は、処理した画像データや各被写体の関節位置、固有物体の位置や大きさ情報、被写体の重心、顔や瞳の位置情報などをカメラ CPU 121 内の RAM に保存する。判定処理の結果を他の画像処理 (例えばホワイトバランス調整処理) に利用してもよい。

10

20

30

40

50

【0021】

表示器（表示手段）131は、LCD等の表示素子を備え、カメラ100の撮像モードに関する情報、撮像前のプレビュー画像、撮像後の確認用画像、焦点検出領域の指標および合焦画像等を表示する。操作スイッチ群132は、メイン（電源）スイッチ、リリース（撮影トリガ）スイッチ、ズーム操作スイッチ、撮影モード選択スイッチ等を含み、ユーザにより操作される。フラッシュメモリ133は、撮像画像を記録する。フラッシュメモリ133は、カメラ100に対して着脱が可能である。

【0022】

被写体検出手段としての被写体検出部140は、機械学習により生成される辞書データに基づいて、被写体検出を行う。本実施形態では、被写体検出部140は、複数種類の被写体を検出するために、被写体ごとの辞書データを用いる。各辞書データは、例えば、対応する被写体の特徴が登録されたデータである。被写体検出部140は、被写体ごとの辞書データを順次切り替えながら被写体検出を行う。本実施形態では、被写体ごとの辞書データは辞書データ記憶部141に記憶される。従って、辞書データ記憶部141には、複数の辞書データが記憶される。カメラCPU121は、複数の辞書データの中から何れの辞書データを用いて被写体検出を行うかを、事前に設定された被写体の優先度や撮像装置の設定に基づいて決定する。被写体検出としては人物の検出と、人物の顔、瞳、胴体等の器官検出を行う。

【0023】

さらに人以外のボール等の物体検出も行う。被写体検出部140では人物被写体と人物とは異なる物体物の被写体（例：ボール、ゴールリング、ネット）の被写体検出を行う。

【0024】

姿勢検出手段としての姿勢取得部142は被写体検出部140で検出された複数の被写体それぞれに対して姿勢推定を行い、姿勢情報を取得する。取得すべき姿勢情報の内容については、被写体の種類に応じて決定される。ここでは、被写体が人物であるので、姿勢取得部142は、被写体として人物の複数の関節の位置を取得するものとする。なお、姿勢取得部142は人物の関節の位置を取得できるため、人物が特定の動きを伴う動作を行っているか否かを判定するアクション認識を行うこともできる。

【0025】

なお、姿勢推定の方法には、如何なる方法を用いてもよく、例えば文献1に記載の方法を用いることができる。姿勢情報の取得の詳細は後述する。

（文献1）

Cao, Zhe, et al., "Realtime multi-person 2d pose estimation using part affinity fields.", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

辞書データ記憶部141には、被写体ごとの辞書データが記憶されている。被写体検出部140は、撮像された画像データと辞書データとに基づいて、画像中の被写体の位置を推定する。被写体検出部140は、被写体の位置やサイズ、信頼度等を推定して、推定した情報を出力してもよい。被写体検出部140は、他の情報を出力してもよい。

【0026】

被写体検出用の辞書データとしては、例えば、被写体として「人物」を検出するための辞書データや「動物」を検出するための辞書データ、「乗物」を検出するための辞書データ、ボールを検出するための辞書データ等がある。また、「人物の全体」を検出するための辞書データと「人物の顔」等の器官を検出するための辞書データとが別個に辞書データ記憶部141に記憶されていてもよい。

【0027】

本実施形態では、被写体検出部140は、機械学習されたCNNにより構成され、画像データに含まれる被写体の位置等を推定する。本実施形態では、被写体検出部140は、それぞれ異なるCNN（コンボリュショナル・ニューラル・ネットワーク）により構成

される。被写体検出部 140 は、GPU（グラフィックス・プロセッシング・ユニット）や CNN による推定処理に特化した回路で実現されてもよい。

【0028】

CNN の機械学習は、任意の手法で行われ得る。例えば、サーバ等の所定のコンピュータが、CNN の機械学習を行い、カメラ 100 は、学習された CNN を、所定のコンピュータから取得してもよい。例えば、所定のコンピュータが、学習用の画像データを入力とし、学習用の画像データに対応する被写体の位置等を教師データとした教師あり学習を行うことで、被写体検出部 140 の CNN の学習が行われてもよい。以上により、学習済みの CNN が生成される。CNN の学習は、カメラ 100 または上述した画像処理装置で行われてもよい。

10

【0029】

競技検出手段としての競技種別取得部 143 は、被写体検出部 140 と辞書データ記憶部 141 と姿勢取得部 142 の情報から人物被写体が行っている競技種別を特定する。なお、競技種別取得部 143 は、予め撮影者の意思によって競技種別を特定（設定）する機能を有してもよい。

【0030】

パン・チルト検出部 144 は、ジャイロセンサ等を用いて構成され、カメラのパンニング動作及びチルティング動作を検出する。本体姿勢判定部 145 は、パン・チルト検出部 144 の出力に基づいて、パンニング又はチルティングの動作開始から動作終了までの間に撮影者が行う一連のカメラ操作を本体姿勢として判定する。判定結果は撮影者意思推定部 146 に通知される。

20

【0031】

撮影者意思推定部 146 は、撮影者の撮影意図を推定する。撮影者意思推定部 146 は、被写体検出部 140 からの被写体情報及び姿勢取得部 142 からの姿勢種別（前述したアクション認識）の情報と、本体姿勢判定部 145 からの撮影者によるカメラ操作（パンニング又はチルティングの動作）の履歴情報に基づいて、撮影者が意図的に主被写体を切り替えようとしているか否かを推定する。推定結果はカメラ CPU 121 に送られた後、カメラ CPU 121 からの各駆動命令によって主被写体へのフォーカス駆動制御に用いられる。

【0032】

次に、図 2 を用いて撮像素子 107 の画像配列について説明する。図 2 は、撮像素子 107 のうち 4 画素列×4 画素行の範囲の画素配列を光軸方向（z 方向）から見て示している。

30

【0033】

1 つの画素ユニット 200 は 2 行×2 列に配置された 4 つの撮像素素を含む。撮像素子 107 上に多数の画素ユニット 200 が配列されることで、2 次元の被写体像の光電変換を行うことができる。1 つの画素ユニット 200 のうち左上には R（赤）の分光感度を有する撮像素素（以下、R 画素という）200R が配置され、右上と左下には G（緑）の分光感度を有する撮像素素（以下、G 画素という）200G が配置されている。さらに、右下には B（青）の分光感度を有する撮像素素（以下、B 画素という）200B が配置されている。また、各撮像素素は、水平方向（x 方向）に分割された第 1 焦点検出画素 201 と第 2 焦点検出画素 202 を含む。

40

【0034】

本実施形態の撮像素子 107 において、撮像素素の画素ピッチ P は $4\mu\text{m}$ 、撮像素素数 N は水平（x）5575 列×垂直（y）3725 行＝約 2075 万画素である。また、焦点検出画素の画素ピッチ PAF は $2\mu\text{m}$ 、焦点検出画素数 NAF は水平 11150 列×垂直 3725 行＝約 4150 万画素である。

【0035】

本実施形態では、各撮像素素が水平方向に 2 分割されている場合について説明するが、垂直方向に分割されていてもよい。また、本実施形態の撮像素子 107 は、それぞれ第 1

50

および第2焦点検出画素を含む撮像素子を複数有するが、撮像素子と第1および第2焦点検出画素とを別々の画素として設けてもよい。例えば、複数の撮像素子の中に第1および第2焦点検出画素を離散的に配置してもよい。

【0036】

図3(a)は、撮像素子107の受光面側(+z方向)から見た1つの撮像素子(200R, 200G, 200B)を示している。図3(b)は、図3(a)の撮像素子のa-a断面を-y方向から見て示している。図3(b)に示すように、1つの撮像素子には、入射光を集光するための1つのマイクロレンズ305が設けられている。

【0037】

また、撮像素子には、x方向にN分割(本実施形態では2分割)された光電変換部301, 302が設けられている。光電変換部301, 302はそれぞれ、第1焦点検出画素201および第2焦点検出画素202に相当する。光電変換部301, 302の重心はそれぞれ、マイクロレンズ305の光軸に対して、-x側および+x側に偏心している。

【0038】

各撮像素子におけるマイクロレンズ305と光電変換部301, 302との間には、R, GまたはBのカラーフィルタ306が設けられている。なお、光電変換部ごとにカラーフィルタの分光透過率を変えてもよいし、カラーフィルタを省略してもよい。

【0039】

撮像素子から撮像素子に入射した光は、マイクロレンズ305により集光され、カラーフィルタ306で分光された後、光電変換部301, 302で受光され、ここで光電変換される。

【0040】

次に、図4を用いて、図3に示した画素構造と瞳分割との関係を説明する。図4は、図3(a)に示した撮像素子のa-a断面を+y側から見て示すとともに、撮像素子の射出瞳を示す。図4では、射出瞳の座標軸との対応を取るために、撮像素子のx方向とy方向とを図3(b)に対して反転させている。

【0041】

射出瞳のうち+x側に重心が偏心した第1瞳領域501は、マイクロレンズ305によって、撮像素子のうち-x側の光電変換部301の受光面と概ね共役な関係とされた領域である。第1瞳領域501を通過した光束は、光電変換部301、すなわち第1焦点検出画素201により受光される。また、射出瞳のうち-x側に重心が偏心した第2瞳領域502は、マイクロレンズ305によって、撮像素子のうち+x側の光電変換部302の受光面と概ね共役な関係とされた領域である。第2瞳領域502を通過した光束は、光電変換部302、すなわち第2焦点検出画素202により受光される。瞳領域500は、光電変換部301, 302(第1および第2焦点検出画素201, 202)を全て合わせた撮像素子全体で受光可能な瞳領域を示している。

【0042】

図5は、撮像素子107による瞳分割を示している。第1瞳領域501と第2瞳領域502をそれぞれ通過した一対の光束は、撮像素子107の各画素にそれぞれ異なる角度で入射し、2分割された第1および第2焦点検出画素201, 202により受光される。本実施形態では、撮像素子107の複数の第1焦点検出画素201からの出力信号を集めて第1焦点検出信号を生成し、複数の第2焦点検出画素202からの出力信号を集めて第2焦点検出信号を生成する。また、複数の撮像素子の第1焦点検出画素201からの出力信号と第2焦点検出画素202からの出力信号とを加算して撮像素子信号を生成する。そして、複数の撮像素子からの撮像素子信号を合成して、有効画素数Nに相当する解像度の画像を生成するための撮像素子信号を生成する。

次に、撮像素子のデフォーカス量と撮像素子107から取得される第1焦点検出信号と第2焦点検出信号の位相差(像ずれ量)との関係について、図6を用いて説明する。図中の撮像素子600には撮像素子107が配置されており、図4および図5を参照して説明したように、撮像素子の射出瞳は第1瞳領域501と第2瞳領域502とに2分割され

10

20

30

40

50

ている。デフォーカス量 d は、被写体 (801, 802) からの光束の結像位置 C から撮像面 600 までの距離 (大きさ) を $|d|$ として、結像位置 C が撮像面 600 より被写体側にある前ピン状態を負符号 ($d < 0$) で表し、結像位置 C が撮像面 600 より被写体とは反対側にある後ピン状態を正符号 ($d > 0$) で表すように定義される。結像位置 C が撮像面 600 にある合焦状態では、 $d = 0$ である。撮像光学系は、被写体 801 に対して合焦状態 ($d = 0$) にあり、被写体 802 に対して前ピン状態 ($d < 0$) にある。前ピン状態 ($d < 0$) と後ピン状態 ($d > 0$) を合わせてデフォーカス状態 ($|d| > 0$) という。

【0043】

前ピン状態 ($d < 0$) では、被写体 802 からの光束のうち第1瞳領域 501 (第2瞳領域 502) を通過した光束は、一旦集光した後、光束の重心位置 $G1$ ($G2$) を中心として幅 $\Gamma1$ ($\Gamma2$) に広がり、撮像面 600 上でボケ像を形成する。このボケ像は、撮像素子 107 上の各第1焦点検出画素 201 (各第2焦点検出画素 202) により受光され、第1焦点検出信号 (第2焦点検出信号) が生成される。つまり、第1焦点検出信号 (第2焦点検出信号) は、撮像面 600 上での光束の重心位置 $G1$ ($G2$) にて被写体 802 がボケ幅 $\Gamma1$ ($\Gamma2$) だけボケた被写体像を表す信号となる。

【0044】

被写体像のボケ幅 $\Gamma1$ ($\Gamma2$) は、デフォーカス量 d の大きさ $|d|$ の増加に概ね比例して増加する。同様に、第1焦点検出信号と第2焦点検出信号間の像ずれ量 p (= 光束の重心位置の差 $G1 - G2$) の大きさ $|p|$ も、デフォーカス量 d の大きさ $|d|$ の増加に概ね比例して増加する。後ピン状態 ($d > 0$) でも、第1焦点検出信号と第2焦点検出信号間の像ずれ方向は前ピン状態とは反対となるが、同様である。

【0045】

このように、デフォーカス量の大きさが増加するのに伴って第1および第2焦点検出信号間の像ずれ量の大きさが増加する。本実施形態では、撮像素子 107 を用いて得られた第1および第2焦点検出信号間の像ずれ量からデフォーカス量を算出する撮像面位相差検出方式焦点検出を行う。

【0046】

次に、撮像素子 107 のうち第1および第2焦点検出信号を取得する焦点検出領域について、図7を用いて説明する。図7において、 $A(n, m)$ は、撮像素子 107 の有効画素領域 1000 に設定された複数 (x 方向および y 方向に3つずつの計9つ) の焦点検出領域のうち x 方向での n 番目、 y 方向での m 番目の焦点検出領域を示す。焦点検出領域 $A(n, m)$ に含まれる複数の第1および第2焦点検出画素 201, 202 からの出力信号から第1および第2焦点検出信号が生成される。 $I(n, m)$ は、表示器 131 において焦点検出領域 $A(n, m)$ の位置を表示する指標を示す。

【0047】

なお、図7に示した9つの焦点検出領域は例に過ぎず、焦点検出領域の数、位置およびサイズは限定されない。例えば、ユーザが指定した位置や被写体検出器により検出された被写体位置を中心とした所定の範囲に一つもしくは複数の領域を焦点検出領域として設定してもよい。本実施形態では、後述するデフォーカスマップ取得に際して、より高分解能に焦点検出結果が得られるよう焦点検出領域を配置する。例えば、撮像素子上に、水平120分割、垂直80分割で計9600点の焦点検出領域を配置する。

【0048】

図8のフローチャートは、本実施形態のカメラ100にAF (オートフォーカス) 動作および撮像動作を行わせるAF・撮像処理 (画像処理方法) を示す。具体的には、カメラ100に、表示器131にライブビュー画像を表示する撮像前から静止画撮像を行うまでの動作を行わせる処理を示す。コンピュータであるカメラCPU121は、コンピュータプログラムに従って本処理を実行する。以下の説明において、 S はステップを意味する。

【0049】

まず $S1$ では、カメラCPU121は、撮像素子駆動回路124に撮像素子107を駆

10

20

30

40

50

動させ、撮像素子107から撮像データを取得する。その後、カメラCPU121は、得られた撮像データのうち、図7に示した焦点検出領域のそれぞれに含まれる複数の第1および第2焦点検出画素から第1および第2焦点検出信号を取得する。また、カメラCPU121は、撮像素子107の全有効画素の第1および第2焦点検出信号を加算して撮像信号を生成し、画像処理回路125に撮像信号（撮像データ）に対する画像処理を行わせて画像データを取得する。なお、撮像画素と第1および第2焦点検出画素とを別々に設けた場合は、カメラCPU121は焦点検出用画素に対する補完処理を行って画像データを取得する。

【0050】

次にS2では、カメラCPU121は、画像処理回路125にS1で得られた画像データからライブビュー画像を生成させ、これを表示器131に表示させる。なお、ライブビュー画像は、表示器131の解像度に合わせた縮小画像であり、ユーザはこれを見ながら撮像構図や露出条件等を調整することができる。そのため、カメラCPU121は、画像データから得られた測光値に基づき露出調整を行い、表示器131に表示する。露出調整は、露光時間、撮影レンズの絞り開口の開閉、撮像素子出力に対するゲイン調整を適宜行うことにより実現される。

【0051】

次にS3では、カメラCPU121は、操作スイッチ群132に含まれるリリーススイッチの半押し操作によって、撮像準備動作の開始を指示するスイッチSw1がオンされたか否かを判定する。カメラCPU121は、Sw1がオンされていない場合は、Sw1がオンになるタイミングを監視するため、S3の判定を繰り返す。一方、Sw1がオンされた場合は、カメラCPU121はS400に進み、被写体追従オートフォーカス（AF）処理を行う。ここでは、得られた撮像信号や、焦点検出信号からの被写体領域の検出、焦点検出領域の設定、焦点検出処理と記録画像の撮像処理までのタイムラグの影響を抑制するための予測AF処理などを行う。詳細は、後述する。

【0052】

そして、カメラCPU121はS5に進み、リリーススイッチの全押し操作によって、撮像動作の開始を指示するスイッチSw2がオンされたか否かを判定する。カメラCPU121は、Sw2がオンされていない場合はS3に戻る。一方、Sw2がオンされた場合はS300に進み、撮像サブルーチンを実行する。撮像サブルーチンの詳細については後述する。撮像サブルーチンが終了すると、S7に進む。

【0053】

S7では、カメラCPU121は、操作スイッチ群132に含まれるメインスイッチがオフされたか否かを判別する。カメラCPU121は、メインスイッチがオフされた場合は本処理を終了し、メインスイッチがオフされていない場合はS3に戻る。

【0054】

本実施形態では、S3でSw1のオンが検出された後、被写体検出処理やAF処理を行うように構成したが、これらの処理を行うタイミングはこれに限らない。Sw1がオンされる前の状態で、S400で行う被写体追従AF処理を行うことにより、撮影者の撮影前の予備動作を不要とすることができる。

【0055】

次に、図9に示すフローチャートを用いて、図8のS300でカメラCPU121が実行する撮像サブルーチンについて説明する。

【0056】

S301において、カメラCPU121は、露出制御処理を行い、撮像条件（シャッター速度、絞り値、撮像感度等）を決定する。この露出制御処理は、ライブビュー画像の画像データから取得した輝度情報を用いて行うことができる。

【0057】

そして、カメラCPU121は、決定した絞り値を絞り駆動回路128に送信して、絞り102を駆動させる。また、カメラCPU121は、決定したシャッター速度をシャッ

10

20

30

40

50

ター１０８に送信して、フォーカルプレーンシャッターを開く動作を行う。さらにカメラＣＰＵ１２１は、撮像素子駆動回路１２４を通じて露光期間の間、撮像素子１０７に電荷蓄積を行わせる。

【００５８】

露光制御処理を行ったカメラＣＰＵ１２１は、Ｓ３０２において、撮像素子駆動回路１２４に、撮像素子１０７からの静止画撮像のための撮像信号の全画素読み出しを行わせる。また、カメラＣＰＵ１２１は、撮像素子駆動回路１２４に撮像素子１０７内の焦点検出領域（合焦目標領域）からの第１および第２焦点検出信号のうち一方の読み出しを行わせる。このときに読み出される第１または第２焦点検出信号は、後述する画像再生時に画像のピント状態を検出するために用いられる。第１および第２焦点検出信号のうち一方の焦点検出信号を撮像信号から差し引くことで他方の焦点検出信号を取得することができる。

10

【００５９】

次にＳ３０３では、カメラＣＰＵ１２１は、画像処理回路１２５に、Ｓ３０２で読み出されてＡ／Ｄ変換された撮像データに対して欠陥画素補正処理を行わせる。

【００６０】

さらにＳ３０４では、カメラＣＰＵ１２１は、画像処理回路１２５に、欠陥画素補正処理後の撮像データに対して、デモザイク（色補間）処理、ホワイトバランス処理、γ補正（階調補正）処理、色変換処理およびエッジ強調処理等の画像処理や符号化処理を行わせる。

【００６１】

そしてＳ３０５では、カメラＣＰＵ１２１は、Ｓ３０４での画像処理や符号化処理が行われることで得られた画像データとしての静止画データと、Ｓ３０２で読み出された一方の焦点検出信号とを、画像データファイルとしてフラッシュメモリ１３３に記録する。

20

【００６２】

次にＳ３０６では、カメラＣＰＵ１２１は、Ｓ３０５で記録した静止画データに対応付けて、カメラ１００の特性情報としてのカメラ特性情報をフラッシュメモリ１３３とカメラＣＰＵ１２１内のメモリに記録する。カメラ特性情報は、例えば以下の情報を含む。

- ・撮像条件（絞り値、シャッタ速度、撮像感度等）
- ・画像処理回路１２５で行った画像処理に関する情報
- ・撮像素子１０７の撮像画素および焦点検出画素の受光感度分布に関する情報
- ・カメラ１００内での撮像光束のケラレに関する情報
- ・カメラ１００における撮像光学系の取付け面から撮像素子１０７までの距離の情報
- ・カメラ１００の製造誤差に関する情報

30

撮像画素および焦点検出画素の受光感度分布に関する情報（以下、単に受光感度分布情報という）は、撮像素子１０７から光軸上の距離（位置）に応じた撮像素子１０７の感度の情報である。この受光感度分布情報は、マイクロレンズ３０５と光電変換部３０１，３０２に依存するため、これらに関する情報であってもよい。また、受光感度分布情報は、光の入射角度に対する感度の変化の情報であってもよい。

【００６３】

次にＳ３０７では、カメラＣＰＵ１２１は、Ｓ３０５で記録した静止画データに対応付けて、撮像光学系の特性情報としてのレンズ特性情報をフラッシュメモリ１３３とカメラＣＰＵ１２１内のメモリに記録する。レンズ特性情報は、例えば、射出瞳に関する情報、光束をける鏡筒等の枠に関する情報、撮像時の焦点距離やＦナンバーの情報、撮像光学系の収差に関する情報、撮像光学系の製造誤差に関する情報および撮像時のフォーカスレンズ１０５の位置（被写体距離）の情報を含む。

40

【００６４】

次にＳ３０８では、カメラＣＰＵ１２１は、静止画データに関する情報としての画像関連情報をフラッシュメモリ１３３とカメラＣＰＵ１２１内のメモリに記録する。画像関連情報は、例えば、撮像前の焦点検出動作に関する情報、被写体の移動に関する情報、焦点検出精度に関する情報を含む。

50

【0065】

次にS309では、カメラCPU121は、表示器131に、撮像画像のプレビュー表示を行う。これにより、ユーザは撮像画像を簡易的に確認することができる。

【0066】

S309の処理が終わると、カメラCPU121は本撮像サブルーチンを終了して、図8のS7に進む。

【0067】

次に、図10に示すフローチャートを用いて、図8のS400でカメラCPU121が実行する被写体追従AF処理サブルーチンについて説明する。

【0068】

S401において、カメラCPU121は、図8のS1で取得した複数の焦点検出領域のそれぞれで得られた第1および第2焦点検出信号間の像ずれ量を算出し、その像ずれ量から焦点検出領域ごとのデフォーカス量を算出する。上述の通り、本実施形態では、撮像素子上に、水平120分割、垂直80分割の計9600点配置された焦点検出領域から得られた焦点検出結果の群を、デフォーカスマップと称する。

【0069】

次に、S402では、カメラCPU121は、被写体検出、および、追尾処理を行う。被写体検出処理は、上述の被写体検出部140が行う。被写体の検出は、得られた画像の状態によっては、検出不可能な場合があるため、そのような場合には、テンプレートマッチングなど他の手段を用いた追尾処理を行い、被写体の位置を推定する。詳細は後述する。

【0070】

次にS403では、カメラCPU121は、被写体検出部140で検出された複数の被写体の各関節位置から姿勢情報を取得する。

【0071】

図11は、姿勢取得部142が取得する情報の概念図である。図11(a)は、処理対象の画像を表しており、被写体901は、ボール903をキャッチしている。被写体901は、撮影シーン中の重要な被写体である。本実施形態では、被写体の姿勢情報を用いることで、撮影者がピントを合わせたいと意図している可能性が高い被写体を判定する。一方、被写体902は非主被写体である。

【0072】

図11(b)は、被写体901及び902の姿勢情報、並びにボール903の位置及び大きさの例を表す図である。関節911は、被写体901の各関節を表しており、関節912は、被写体902の各関節を表している。図11(b)では、関節として、頭頂部、首、肩、肘、手首、腰、膝、足首の位置を取得する例を示しているが、関節位置はこれらの一部でもよいし、別の位置を取得してもよい。また、関節位置だけではなく、関節同士を結ぶ軸などの情報を用いてもよく、被写体の姿勢を表す情報であれば、任意の情報を姿勢情報として用いることができる。以下では、姿勢情報として、関節位置を取得する場合について説明する。

【0073】

姿勢取得部142は、関節911及び関節912の画像中における2次元座標(x, y)を取得する。ここで、(x, y)の単位はピクセルである。重心位置913は、ボール903の重心位置を表し、矢印914は、ボール903の画像中での大きさを表している。被写体検出部140は、ボール903の重心位置の画像中における2次元座標(x, y)、及び、画像中におけるボール903の幅を示すピクセル数を取得する。

【0074】

次にS404では、カメラCPU121は、S402での被写体検出結果、S403での被写体の姿勢情報に基づいて姿勢種別判定処理を行う。ここで姿勢種別判定について説明する。

【0075】

図 1 2 A ~ 図 1 2 C は、バスケットボール競技でのプレーヤーの位置を示す模式図である。図 1 2 A ~ 図 1 2 C では、プレーヤー 9 2 4 とプレーヤー 9 2 5 は、姿勢取得部 1 4 2 からの情報に基づいて、アクション姿勢をとっている被写体と認識された主被写体候補である。ここで、アクション姿勢とは、プレーヤーがシュートを狙っているなどの、ユーザが撮影を望むと予測される被写体の姿勢を意味する。

【 0 0 7 6 】

図 1 2 A では、プレーヤー 9 2 5 を主被写体として、プレーヤー 9 2 5 に A F 枠 1 9 0 0 が設定されている。図 1 2 B では、プレーヤー 9 2 4 がボール 9 0 3 を持ち、ゴールリング 9 4 0 に向けてシュートを狙う姿勢をとっている。カメラ C P U 1 2 1 は、姿勢取得部 1 4 2 のニューラルネットワークを使用した公知の方法でプレーヤーの関節を検出する。そして、関節の位置情報と辞書データ記憶部 1 4 1 の情報とから、プレーヤー 9 2 4 がシュートアクション姿勢をとっていると推定し、かつシュートアクション姿勢時の平均的な姿勢継続時間の情報を取得する。なお、ここでの姿勢推定には、例えば、特開 2 0 2 2 - 1 3 5 5 5 2 号に開示されているように、画像に対して深度検出と、関節検出もしくは器官検出を行い、姿勢を推定する方法などを用いることができる。図 1 2 C は、プレーヤー 9 2 4 がシュートしてボール 9 0 3 がゴールリング 9 4 0 に向かっているシーンを示している。

【 0 0 7 7 】

図 1 3 A ~ 図 1 3 C は、バスケットボール競技での別の状態のプレーヤーの位置を示す模式図である。図 1 3 A は、プレーヤー 9 2 6 とプレーヤー 9 2 7 が存在するシーンを示している。図 1 3 B は、プレーヤー 9 2 6 がボール 9 0 3 を持っており、プレーヤー 9 2 7 がプレーヤー 9 2 6 からのパスを待っているシーンを示している。図 1 3 C は、プレーヤー 9 2 6 がプレーヤー 9 2 7 へパスしたシーンを示している。図 1 3 B、図 1 3 C においては、姿勢取得部 1 4 2 で得られたプレーヤー 9 2 6 とプレーヤー 9 2 7 の姿勢情報に基づいて、カメラ C P U 1 2 1 は、プレーヤー 9 2 6 とプレーヤー 9 2 7 をパスアクション姿勢と認識する。また、パスアクション姿勢時の平均的な姿勢継続時間の情報を取得する。このときに取得した姿勢の種別判定と姿勢の継続時間の情報は、後述する主被写体選定のための主被写体らしさ信頼度の判定要素に用いられ、姿勢継続時間の長さとは信頼度の高さは比例する。

【 0 0 7 8 】

図 1 0 の説明に戻り、S 4 0 5 では、カメラ C P U 1 2 1 は、競技種別判定処理を行う。S 4 0 2 での被写体検出結果、S 4 0 3 での姿勢情報、S 4 0 4 での姿勢種別情報、検出されている複数の被写体の動作から、競技種別を判定する処理を行う。競技種別については後述する。なお、競技種別の判定ではなく、予め競技種別を選択する機能を有していてもよい。

【 0 0 7 9 】

S 4 0 5 での競技種別の判定方法について説明する。例えば、図 1 2 B の画像について、カメラ C P U 1 2 1 は、姿勢取得部 1 4 2 からの姿勢推定による関節接続情報と、競技種別取得部 1 4 3 からの人物及び人物以外の被写体の動きベクトル情報と、辞書データ記憶部 1 4 1 の情報とを取得する。そして、これらの情報から、2 人のプレーヤー 9 2 4 と 9 2 5 が行っている競技種別をバスケットボールと判定する。このとき、動体物以外のゴールリングやコート情報を加味して競技の種類を判定してもよい。

【 0 0 8 0 】

次に S 4 0 0 0 では、カメラ C P U 1 2 1 は、主被写体判定処理を行う。S 4 0 1 で得られたデフォーカスマップ、S 4 0 2 で得られた被写体検出結果、S 4 0 3 で得られた姿勢情報、S 4 0 4 で得られた姿勢種別情報、S 4 0 5 で得られた競技種別情報を用いて、主被写体を判定する。詳細は図 1 5 のサブフローチャートを用いて後述する。

【 0 0 8 1 】

次に、S 4 0 7 では、カメラ C P U 1 2 1 は、S 4 0 1 で取得した焦点検出結果と、過去の焦点検出を行ったタイミングの時系列データである複数のデフォーカス量を用いて、

予測 A F 処理を行う。

【 0 0 8 2 】

これは、焦点検出を行ったタイミングと、撮影画像の露光を行うタイミングにタイムラグがある場合に必要となる処理である。この処理では、焦点検出を行ったタイミングに対して所定時間後である撮影画像の露光を行うタイミングでの光軸方向の被写体の位置を予測して、A F 制御を行う。被写体の像面位置の予測では、過去の被写体の像面位置と時刻の履歴データを用いて、多変量解析（例えば最小 2 乗法）を行い、予測曲線の式を求める。求めた予測曲線の式に撮影画像の露光を行うタイミングの時刻を代入することで、被写体の像面予測位置を算出できる。

【 0 0 8 3 】

さらに光軸方向だけでなく、3次元の位置を予測してもよい。画面上を X Y 面とし、光軸方向を Z 方向とした X Y Z 方向のベクトルを求める。具体的には、S 4 0 2 の被写体検出・追尾処理で得られた被写体の X Y 位置と、S 4 0 5 で得られたデフォーカス量に基づく Z 方向位置の時系列データとから、撮影画像の露光を行うタイミングでの被写体の位置を予測する。さらに被写体である人物の関節位置の時系列データから予測してもよい。なお、予測対象は主被写体とその他の複数の人物と人物以外の動体物を含む。

【 0 0 8 4 】

次に、S 4 0 8 では、カメラ C P U 1 2 1 は、S 4 0 0 0 での主被写体判定処理の結果と、S 4 0 1 で得られたデフォーカス量と、S 4 0 7 での予測 A F 処理の結果を用いて、フォーカスレンズの駆動量を計算する。そして、カメラ C P U 1 2 1 は、この駆動量に基づいてフォーカスアクチュエータ 1 1 4 を駆動し、第 3 レンズ群 1 0 5 を光軸方向に移動させることで焦点調節処理を行う。

【 0 0 8 5 】

なお、焦点調節処理では、S 4 0 0 0 での主被写体判定処理により判定された主被写体に対し、急な加減速ピント移動を避け、滑らかなピント遷移となるような焦点調節を行う。また、撮像装置の撮影シーケンスや焦点調節を行う撮影レンズの制御及び駆動性能に応じて焦点調節を行ってもよい。例えば、撮影シーケンスとして、レンズ駆動時間が長い場合にはフォーカス駆動時間が確保できるので、後述のデフォーカス量の閾値を大きくしてレンズ駆動しやすくする。逆に、レンズ駆動時間が短いシーケンスの場合には、フォーカス駆動時間が確保できないため、後述のデフォーカス量の閾値を小さくして、レンズ駆動しにくくしてもよい。その他フォーカスレンズの駆動源の違いにより、単位時間あたりのフォーカス駆動可能な駆動量が異なるため、フォーカスレンズの駆動源の違いに応じて、後述のデフォーカス量の閾値を変更してもよい。

【 0 0 8 6 】

S 4 0 8 の処理が終わると、カメラ C P U 1 2 1 は、被写体追従 A F 処理のサブルーチンを終了して、図 8 の S 5 に進む。

【 0 0 8 7 】

次に、図 1 4 に示すフローチャートを用いて、図 1 0 の S 4 0 2 でカメラ C P U 1 2 1 が実行する被写体検出・追尾処理のサブルーチンについて説明する。

【 0 0 8 8 】

S 2 0 0 0 において、カメラ C P U 1 2 1 は、図 8 の S 1 で取得した画像データから検出したデータに基づいて、検出したい被写体の種別に応じた辞書データの設定を行う。事前に設定された被写体の優先度や撮像装置の設定に基づき、辞書データ記憶部 1 4 1 に記憶された複数の辞書データから、本処理で用いる辞書データを選択する。例えば、複数の辞書データとして、「人物」「乗物」「動物」など被写体を分類した複数種類の辞書データが記憶されている。本実施形態では、選択する辞書データは一つでもよいし、複数でもよい。一つの場合は、一つの辞書データが検出可能な被写体を、高頻度で繰り返し検出を行うことが可能になる。一方で、複数の辞書データを選択する場合は、検出被写体としての優先度に応じて、シーケンシャルに辞書データを設定することで、順次、被写体を検出することができる。

10

20

30

40

50

【0089】

次に、S2001では、被写体検出部140は、ステップS1で読み出した画像データを入力画像として、ステップS2000で設定された辞書データを用いて、人物または人物以外の被写体検出を行う。このとき、被写体検出部140は、検出した被写体の位置やサイズ、信頼度等の情報を出力する。このとき、カメラCPU121は、被写体検出部140が出力した上記の情報を表示器131に表示させてもよい。

【0090】

S2001では、被写体検出部140は、画像データから、第1の種類の被写体である人物について、階層的に複数の領域を検出する。例えば、辞書データとして「人物」が設定されていた場合は、「全身」の領域、「顔」の領域、「目」の領域といった複数の器官を検出する。人物の目や顔など局所領域は、被写体としてピントや露出状態を合わせたい領域である一方、周囲の障害物や顔の向きによって検出できない場合がある。そのような場合でも、全身の検出を行うことで、被写体をロバストに検出し続けることができるため、階層的に被写体を検出するように構成している。

【0091】

次にS2002では、被写体検出部140は、S2001での第1の種類の被写体とは異なる第2の種類の被写体である人物または人物以外の物体検出を行う。例えば、辞書データ記憶部141に記憶された複数の辞書データから競技に関わる人物を検出する辞書データを選択する。そして、人物が被写体検出された後に検出物体の辞書データに変更し、物体全体の領域検出と、物体中心位置やサイズの検出を行う。なお、検出物体を予め特定して検出してもよい。

【0092】

物体検出には如何なる方法を用いてもよく、例えば、以下の文献2に記載の方法を用いることができる。本実施形態では第2の種類の被写体をボールとするが、ラケット等の別の固有物体でもよい。

(文献2)

Redmon, Joseph, et al., "You only look once: Unified, real-time object detection.", Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

次にS2003では、カメラCPU121は、S2001で得られた被写体検出領域をテンプレートとして、公知のテンプレートマッチング処理を行う。図8のS1で得られた複数枚の画像を用いて、過去の画像で得られた被写体検出領域をテンプレートとして、直前に得られた画像内に、類似の領域を検索する。テンプレートマッチングに用いる情報としては、周知の通り、輝度情報、色のヒストグラム情報、コーナーやエッジなどの特徴点情報など、いずれを用いてもよい。マッチング方法やテンプレート更新方法も種々の方法が考えられるが、いずれの方法を用いてもよい。S2003で行う追尾処理は、S2001で被写体を検出されなかった場合に、過去の被写体検出データと類似の領域を直前に得られた画像データから検出することにより、安定した被写体検出・追尾処理を実現するために行う。

【0093】

S2003の処理が終わると、カメラCPU121は、本被写体検出・追尾処理のサブルーチンを終了して、図10のS403に進む。

【0094】

次に図15に示すフローチャートを用いて、図10のS4000で行う主被写体判定処理のサブルーチンについて説明する。

【0095】

S4001において、カメラCPU121は、図10のS401で取得したデフォーカスマップ（焦点状態）とS403の姿勢情報に基づいて複数の被写体の中から主被写体らしい被写体候補を選定する。主被写体らしさとは、姿勢情報に基づいて算出される、図10のS405までに説明したシュートやパスのアクション姿勢をとっている被写体の主被

写体である信頼度（確率）である。以下では、主被写体らしさを表す信頼度（被写体が処理対象画像の主被写体である可能性の度合い）として、被写体が処理対象画像の主被写体である確率を採用する場合について説明するが、確率以外の値を用いてもよい。例えば、信頼度として、被写体の重心位置と固有物体の重心位置との間の距離の逆数を用いることができる。

【0096】

＜主被写体である確率の算出方法＞

以下、各関節の座標及び大きさに基づいて主被写体らしさを表す確率を算出する方法について説明する。以下では、機械学習の一手法であるニューラルネットワークを用いる場合について説明する。

10

【0097】

図16は、ニューラルネットワークの構造の一例を示す図である。図16において、1001は入力層、1002は中間層、1003は出力層、1004はニューロン、1005はニューロン1004同士の接続関係を表す線である。ここでは、図示の都合上、代表的なニューロンと接続線にのみ番号を付与している。入力層1001のニューロン1004の数は、入力するデータの次元に等しく、出力層1003のニューロンの数は2であるとする。これは、主被写体らしいか否かを判定する2クラス分類の問題であることに対応している。

【0098】

入力層1001のi番目のニューロン1004と、中間層1002のj番目のニューロン1004とを接続する線1005には、重み w_{ji} が与えられており、中間層1002におけるj番目のニューロン1004が出力する値 z_j は、以下の式で与えられる。

20

【0099】

【数1】

$$z_j = h(b_j + \sum_i w_{ji} x_i) \quad \cdots (1)$$

$$h(z) = \max(z, 0) \quad \cdots (2)$$

30

【0100】

式(1)において、 x_i は入力層1001のi番目のニューロン1004に入力される値を表す。和は、j番目のニューロンと接続されている、入力層1001の全てのニューロン1004に対してとるものとする。 b_j はバイアスと呼ばれ、j番目のニューロン1004の発火のしやすさをコントロールするパラメータである。また、式(2)で定義される関数 h は、ReLU (Rectified Linear Unit) と呼ばれる活性化関数である。活性化関数としては、シグモイド関数など別の関数を用いることも可能である。

【0101】

また、出力層1003のk番目のニューロン504が出力する値 y_k は、以下の式で与えられる。

40

【0102】

【数2】

$$y_k = f(b_k + \sum_j w_{kj} z_j) \quad \cdots (3)$$

$$f(y_k) = \frac{\exp(y_k)}{\sum_i \exp(y_i)} \quad \cdots (4)$$

【0103】

50

式（３）において、 z_j は中間層１００２の j 番目のニューロン１００４が出力する値を表し、 $i, k = 0, 1$ である。０は非主被写体、１は主被写体らしさに対応する。和は、 k 番目のニューロンと接続されている、中間層１００２の全てのニューロンに対してとるものとする。また、式（４）で定義される関数 f は、ソフトマックス関数と呼ばれ、 k 番目のクラスに属する確率値を出力する。本実施形態では、 $f(y_1)$ を主被写体らしさを表す確率として用いる。

【０１０４】

学習時は、人物の関節の座標とボールの座標及び大きさとを入力する。そして、出力される確率と正解ラベルとを用いる損失関数を最小化するように、全ての重みとバイアスの最適化を行う。ここで、正解ラベルとは、主被写体の場合には“１”、非主被写体の場合には“０”の２値をとるものとする。損失関数 L としては、下記で表されるような二値クロスエントロピーを用いることができる。

【０１０５】

【数３】

$$L(y, t) = - \sum_m t_m \log y_m - \sum_m (1 - t_m) \log (1 - y_m) \quad \cdots (5)$$

【０１０６】

式（５）において、添え字の m は学習対象の被写体のインデックスを表している。 y_m は、出力層１００３における $k = 1$ のニューロンから出力される確率値であり、 t_m は正解ラベルである。損失関数は、式（５）の他にも、平均二乗誤差など、正解ラベルとの一致度を測れる関数であれば、何でもよい。式（５）に基づいて最適化することにより、正解ラベルと出力される確率値とが近づくように重みやバイアスを決定することができる。学習済みの重みやバイアス値は、予めフラッシュメモリ１３３に保存しておき、必要に応じてカメラＣＰＵ１２１内のＲＡＭに格納する。シーンに応じて複数種類の重みやバイアス値を用意しても構わない。学習済みの重みやバイアス（事前に行われた機械学習の結果）を用いて、式（１）～式（４）に基づいて確率値 $f(y_1)$ を出力する。

【０１０７】

なお、学習する際には、重要な行動に移る前の状態を主被写体らしさの状態として学習できる。例えば、ボールを投げる場合であれば、ボールを投げるときに、手を前に伸ばした状態を主被写体らしさの状態の１つとして学習することができる。この構成を採用する理由は実際に主被写体らしさが重要な行動を起こした際に、的確に撮像装置の制御が実行されている必要があるためである。例えば、主被写体らしさに対応する信頼度（確率値）が予め設定した第１の所定値を超えた場合に、自動で画像や映像を記録する制御（記録制御）を開始することにより、撮影者は重要な瞬間を逃さず撮影することが可能になる。この際、学習対象の状態から、重要な行動までの典型的な時間の情報を撮像装置の制御に使用しても構わない。

【０１０８】

以上、ニューラルネットワークを用いた確率の算出方法について説明したが、主被写体らしいか否かのクラス分類ができれば、サポートベクタマシンや決定木など別の機械学習の手法を用いても構わない。また、機械学習に限らず、あるモデルに基づいて信頼度又は確率値を出力する関数を構築してもよい。人物とボールとの間の距離が近いほど、主被写体らしさである信頼性が大きいと仮定し、人物とボールとの間の距離に対する単調減少関数の値を用いることも可能である。

【０１０９】

なお、ボールの情報も用いて主被写体らしさの判定を行ったが、被写体の姿勢情報のみを用いて主被写体らしさの判定を行うことも可能である。被写体の姿勢情報の種別（例えば、パス、シュートなど）に応じてボールの情報も用いた方が良い場合と良くない場合がある。例えばシュートの場合には人物とボールとの間の距離が遠くなるが、シュートを打

10

20

30

40

50

った被写体を撮影者が主被写体らしいとしたい場合もあるため、ボールによらず、被写体である人物の姿勢情報のみで主被写体らしさを判定してもよいし、被写体の姿勢情報からの種別に応じて、ボールの情報も用いて主被写体らしさを判定してもよい。

【0110】

その他にも、各関節の座標やボールの座標や大きさに対して、線形変換など、所定の変換を施したデータを入力データとして用いてもよい。またデフォーカス差がある2つの被写体間で主被写体らしさが頻繁に切り替わると撮影者の意図とは異なる場合が多い。そのため、各被写体の信頼度の時系列データから頻繁に切り替わっていることを検出し、2つの被写体間であれば、どちらかの被写体（例えば至近側の被写体）の信頼度を高くすることで切り替わりを防止してもよい。また、2つの被写体を含んだ領域を主被写体らしさを表す領域としてもよい。

10

【0111】

さらに別の方法として、人物の姿勢情報、人物とボールの位置、各被写体のデフォーカス量、主被写体らしさを示す信頼度の時系列データを入力データとして用いてもよい。また、前述の予測処理を行い、撮影画像の露光を行うタイミングの時刻の人物の関節の座標やボールの座標や大きさを予測したデータを入力データとして用いて、信頼度を算出してもよい。被写体の像面移動速度や各関節の座標の時系列変化量に応じて、予測処理を行ったデータを用いるか否かを切り替えてもよい。そうすることで、被写体の姿勢変化が少ない場合には主被写体らしさを示す信頼度の精度を維持でき、被写体の姿勢変化が大きい場合には予測処理での結果を用いることで、主被写体らしさを示す被写体をより早い時刻に検出できる。以上の方法により、複数の第1の種類の被写体の信頼度を算出する。

20

【0112】

次にS4002では、カメラCPU121は、主被写体らしさの信頼度を判定する。撮影者がAF枠1900で狙っている被写体の他に、S4001で主被写体候補として判定した複数の被写体の中で姿勢情報に基づく主被写体らしさを示す信頼度が最も高い（より高い）被写体が存在する場合にはS4003へ進む。そのような被写体が存在しない、つまり現在のAF枠1900に対応する主被写体の信頼度が最も高い場合にはS4009へ進む。

【0113】

次にS4003では、カメラCPU121は、S4001で判定した信頼度が最も高い主被写体候補の姿勢種別情報を取得する（状態判定）。姿勢種別情報とは前述したように姿勢取得部142において取得された、被写体動作がシュート、パス、ドリブル、スパイク、ブロック、レシーブなど、どのような動作なのかの判定情報とその姿勢種別の姿勢継続時間に関する情報である。

30

【0114】

次にS4004では、カメラCPU121は、S4003で判定した主被写体候補に対して競技種別取得部143から競技種別情報を取得する（競技判定）。競技種別とは、前述したように姿勢取得部142の情報も加味して、バスケットボール、バレーボール、サッカーなど、どのような競技をしているのかの情報で、競技毎に姿勢継続時間を設定する。このとき、被写体人物の姿勢動作情報以外に、ボールなどの動体物情報とゴールリング、ネット、ゴールネットなどの固定物情報を競技種別判定の付加情報として使用してもよく、また予め競技種別を設定できるようにしてもよい。

40

【0115】

次にS4005では、カメラCPU121は、S4003で判定した主被写体候補のデフォーカス量を取得する。

【0116】

次にS4006では、カメラCPU121は、S4003、S4004、S4005で判定した主被写体候補の姿勢種別と競技種別の情報から推定した姿勢継続時間とデフォーカス量（焦点状態）とに基づいて、急激なピント変化をさせずに主被写体候補毎にピント合わせするためのデフォーカス量の閾値を設定する。デフォーカス量の閾値設定の詳細に

50

ついて以下に説明する。

【0117】

図12A～12Cを使用し、競技種別がバスケットボール、姿勢種別がシュートアクション（シュート姿勢）の場合の閾値設定について説明する。

【0118】

バスケットボールにおいて、競技の状況が、図12Aに示す状況から、図12Bに示す状況に変化したとする。図12Aは、2人のプレイヤー924, 925が存在し、AF枠1900がプレイヤー925に設定されている状況を示している。図12Bは、プレイヤー924がボール903を持ってゴールリング940へのシュート姿勢をとっている状況を示している。

10

【0119】

ここで、図12AのようにAF枠1900がプレイヤー925に位置する状態から、図12Bのように、プレイヤー924に位置する状態に切り替えるか否かを判定する制御を図15のフローチャートに当てはめると、以下のようになる。

【0120】

プレーの状況が図12Aの状況から図12Bの状況に変化すると、図15のフローチャートにおいて、S4002では、現在AF枠1900が位置している被写体よりも、例えばボールに近い主被写体の候補であるプレイヤー924の方が主被写体らしさの信頼度が高くなる。そのため、処理は、S4003に進む。S4003では、図12Bのプレイヤー924の姿勢種別がシュート姿勢であると判定される。S4004では、現在の競技種別がバスケットボールであると判定される。S4005では、主被写体候補であるプレイヤー924のデフォーカス量が取得される。S4006では、主被写体を切り替えるためのデフォーカス量の閾値、言い換えると、AF枠1900をプレイヤー925からプレイヤー924へ切り替えるためのプレイヤー924のデフォーカス量の閾値を設定する。

20

【0121】

ここで、バスケットボールの場合、シュートするプレイヤー924の姿勢継続時間が比較的長く、予め定めた所定時間より長い。そして、動作が頻繁に切り替わることがないため、主被写体らしさが維持される。そのため、AF枠1900を、プレイヤー925からプレイヤー924に切り替えたとしても、フォーカス駆動する時間が確保でき、主被写体候補であるプレイヤー924にピント移動できる可能性が高い。

30

【0122】

このようなことから、S4006では、主被写体候補であるプレイヤー924のデフォーカス量の閾値を大きい値、例えば90Fδ（Fδとは撮影レンズの絞り値F、許容錯乱円δでベストピント位置を0とした場合のデフォーカス量を表す）に設定する。デフォーカス量の閾値を大きく設定すれば、プレイヤー924のデフォーカス量が多少大きくても、S4007において閾値以下と判断され、AF枠1900を移動させる判断がなされる。そのため、AF枠が移動しやすくなり、S4008において、AF枠1900がプレイヤー925からプレイヤー924に移動される。このようにして、ピントを合わせるべき主被写体候補であるプレイヤー924に、速やかにAF枠が移動される。

【0123】

次に、競技の状況が、図12Bに示す状況から、図12Cに示す状況に変化したとする。図12Cは、プレイヤー924がシュートしてボール903がゴールリング940に向かっていている状況を示している。この場合、図12Cにおいて、プレイヤー924以外に主被写体らしさの信頼度が高いプレイヤーが存在しない。そのため、図15のS4002からS4009に進み、プレイヤー924のAF枠1900は継続される。

40

【0124】

次に、図13A～13Cを使用し、競技種別がバスケットボール、姿勢種別がパスアクション（パス姿勢）の場合の閾値設定について説明する。

【0125】

バスケットボールにおいて、競技の状況が、図13Aに示す状況から、図13B、図1

50

3 Cに示す状況に変化したとする。図13 Aは、2人のプレーヤー926とプレーヤー927が存在し、AF枠1900がプレーヤー927に設定されている状況を示している。図13 Bは、プレーヤー926がボール903を受け取った状況を示している。図13 Cは、プレーヤー926がプレーヤー927へボール903をパスした状況を示している。

【0126】

ここで、図13 AのようにAF枠1900がプレーヤー927に位置する状態から、プレーヤー926に位置する状態に切り替えるか否かを判定する制御を図15のフローチャートに当てはめると、以下のようになる。

【0127】

プレーの状況が図13 Aの状況から図13 Bの状況に変化すると、図15のフローチャートにおいて、S4002では、現在AF枠1900が位置している被写体よりも、例えばボールに近い主被写体の候補であるプレーヤー926の方が主被写体らしさの信頼度が高くなる。そのため、処理は、S4003に進む。S4003では、図13 Bのプレーヤー926の姿勢種別がパス姿勢であると判定される。S4004では、現在の競技種別がバスケットボールであると判定される。S4005では、主被写体候補であるプレーヤー926のデフォーカス量が取得される。S4006では、主被写体を切り替えるためのデフォーカス量の閾値、言い換えると、AF枠1900をプレーヤー927からプレーヤー926へ切り替えるためのプレーヤー926のデフォーカス量の閾値を設定する。

【0128】

ここで、バスケットボールの場合、パスするプレーヤー924の姿勢継続時間は、シュートをする動作などの姿勢継続時間よりも短く、予め定めた所定時間よりも短いことが予想される。そして、動作が頻繁に切り替わることがあるため、主被写体らしさが維持されにくい。そのため、AF枠1900を、プレーヤー927からプレーヤー926に切り替えたとしても、切り替え後に主被写体が切り替わる可能性がある。

【0129】

このようなことから、S4006では、主被写体候補であるプレーヤー926のデフォーカス量の閾値を小さい値、例えば20Fδ（Fδとは撮影レンズの絞り値F、許容錯乱円δでベストピント位置を0とした場合のデフォーカス量を表す）に設定する。デフォーカス量の閾値を小さく設定すれば、プレーヤー926のデフォーカス量が比較的小さくても、S4007において閾値より大きいと判断され、AF枠1900を移動させない判断がなされる。そのため、AF枠が移動しにくくなり、S4008において、AF枠1900はプレーヤー927からプレーヤー926に移動されない。このようにして、主被写体が頻繁に切り替わる場合に、フォーカス駆動が行ったり来たりハンチングすることを防止し、フォーカス駆動の安定性を向上させることができる。

【0130】

次に、競技の状況が、図13 Bに示す状況から、図13 Cに示す状況に変化したとする。この場合、図13 Cにおいて、プレーヤー927がボール903を持っているため、プレーヤー927以外に主被写体らしさの信頼度が高いプレーヤーが存在しない。そのため、図15のS4002からS4009に進み、プレーヤー927のAF枠1900は継続される。

【0131】

次に、図17 A、図17 Bを使用し、競技種別がバレーボール、姿勢種別がスパイクアクション（スパイク姿勢）の場合の閾値設定について説明する。

【0132】

バレーボールにおいて、競技の状況が、図17 Aに示す状況から、図17 Bに示す状況に変化したとする。図17 Aは、プレーヤー932がプレーヤー934のボール940のトスのタイミングに合わせてジャンプし、ネット905越しにプレーヤー933が存在する状況を示している。そして、AF枠1900はプレーヤー932に設定されている。図13 Bは、プレーヤー932がスパイクを打った直後で、プレーヤー933がスパイクをブロックしようとジャンプした状況を示している。

【0133】

ここで、図17AのようにAF枠1900がプレイヤー932に位置する状態から、プレイヤー933に位置する状態に切り替えるか否かを判定する制御を図15のフローチャートに当てはめると、以下のようになる。

【0134】

プレーの状況が図17Aの状況から図17Bの状況に変化すると、図15のフローチャートにおいて、S4002では、現在AF枠1900が位置している被写体よりも、例えばボールに近い主被写体の候補であるプレイヤー933の方が主被写体らしさの信頼度が高くなる場合がある。その場合、処理は、S4003に進む。S4003では、図17Bのプレイヤー932の姿勢種別がスパイク姿勢であると判定される。S4004では、現在の競技種別がバレーボールであると判定される。S4005では、主被写体候補であるプレイヤー933のデフォーカス量が取得される。S4006では、主被写体を切り替えるためのデフォーカス量の閾値、言い換えると、AF枠1900をプレイヤー932からプレイヤー933へ切り替えるためのプレイヤー933のデフォーカス量の閾値を設定する。

10

【0135】

ここで、バレーボールの場合、スパイクするプレイヤー932の姿勢継続時間が非常に短い。そして、動作が頻繁に切り替わることがあるため、主被写体らしさが維持されにくい。そのため、AF枠1900を、プレイヤー932からプレイヤー933に切り替えたとしても、切り替え後に主被写体が切り替わる可能性がある。

20

【0136】

このようなことから、S4006では、主被写体候補であるプレイヤー932のデフォーカス量の閾値を小さい値、例えば15Fδ（Fδとは撮影レンズの絞り値F、許容錯乱円δでベストピント位置を0とした場合のデフォーカス量を表す）に設定する。デフォーカス量の閾値を小さく設定すれば、プレイヤー933のデフォーカス量が比較的小さくても、S4007において閾値より大きいと判断され、AF枠1900を移動させない判断がなされる。そのため、AF枠が移動しにくくなり、S4008において、AF枠1900はプレイヤー932からプレイヤー933に移動されない。このようにして、主被写体が頻繁に切り替わる場合に、フォーカス駆動が行ったり来たりハンティングすることを防止し、フォーカス駆動の安定性を向上させることができる。

30

【0137】

図18A、図18Bを使用し、競技種別がサッカー、姿勢種別がシュートアクション（シュート姿勢）の場合の閾値設定について説明する。

【0138】

サッカーにおいて、競技の状況が、図18Aに示す状況から、図18Bに示す状況に変化したとする。図18Aは、プレイヤー936がボール960を持ち、プレイヤー935がボール960を待ち受けており、プレイヤー937がゴールキーパーとして待ち構えている状況を示している。そして、AF枠1900はプレイヤー936に設定されている。図18Bは、プレイヤー936からパスされたプレイヤー935がサッカーゴール961に向けてシュートした状況を示している。

40

【0139】

ここで、図18AのようにAF枠1900がプレイヤー936に位置する状態から、図18Bのように、プレイヤー935に位置する状態に切り替えるか否かを判定する制御を図15のフローチャートに当てはめると、以下のようになる。

【0140】

プレーの状況が図18Aの状況から図18Bの状況に変化すると、図15のフローチャートにおいて、S4002では、現在AF枠1900が位置している被写体よりも、例えばボールに近い主被写体の候補であるプレイヤー935の方が主被写体らしさの信頼度が高くなる。そのため、処理は、S4003に進む。S4003では、図18Bのプレイヤー935の姿勢種別がシュート姿勢であると判定される。S4004では、現在の競技種

50

別がサッカーであると判定される。S 4 0 0 5では、主被写体候補であるプレーヤー 9 3 5のデフォーカス量が取得される。S 4 0 0 6では、主被写体を切り替えるためのデフォーカス量の閾値、言い換えると、A F 枠 1 9 0 0をプレーヤー 9 3 6からプレーヤー 9 3 5へ切り替えるためのプレーヤー 9 3 5のデフォーカス量の閾値を設定する。

【0 1 4 1】

ここで、サッカーの場合、シュートするプレーヤー 9 3 5の姿勢継続時間が比較的長く、予め定めた所定時間よりも長い。そして、動作が頻繁に切り替わることがないため、主被写体らしさが維持される。そのため、A F 枠 1 9 0 0を、プレーヤー 9 3 6からプレーヤー 9 3 5に切り替えたとしても、フォーカス駆動する時間が確保でき、主被写体候補であるプレーヤー 9 3 5にピント移動できる可能性が高い。

【0 1 4 2】

このようなことから、S 4 0 0 6では、主被写体候補であるプレーヤー 9 3 5のデフォーカス量の閾値を大きい値、例えば8 0 F δ（F δとは撮影レンズの絞り値F、許容錯乱円δでベストピント位置を0とした場合のデフォーカス量を表す）に設定する。デフォーカス量の閾値を大きく設定すれば、プレーヤー 9 3 5のデフォーカス量が多少大きくても、S 4 0 0 7において閾値以下と判断され、A F 枠 1 9 0 0を移動させる判断がなされる。そのため、A F 枠が移動しやすくなり、S 4 0 0 8において、A F 枠 1 9 0 0がプレーヤー 9 3 6からプレーヤー 9 3 5に移動される。このようにして、ピントを合わせるべき主被写体候補であるプレーヤー 9 3 5に、速やかにA F 枠が移動される。

【0 1 4 3】

以上説明したように、本実施形態では、被写体の動作が頻繁に切り替わることがなく、主被写体として維持されやすい競技種別かつ姿勢種別の場合には、デフォーカス量の閾値を大きくして、主被写体が切り替わりやすく（A F 枠が移動しやすく）する。一方、被写体の動作が頻繁に切り替わることがあり、主被写体として維持されにくい競技種別かつ姿勢種別の場合には、デフォーカス量の閾値を小さくして、主被写体が切り替わりにくく（A F 枠が移動しにくく）する。これにより、競技種別、姿勢種別かつデフォーカス量に応じて、適切にフォーカス制御することが可能となる。

【0 1 4 4】

なお、S 4 0 0 6で設定されるデフォーカス量の閾値は、競技に関わらず、姿勢種別（アクション認識）のみに基づいて設定されてもよい。また、撮像装置側の撮影シーケンスや撮影光学系の駆動処理時間などの違いにより変更してもよい。

【0 1 4 5】

（第2の実施形態）

次に、本発明の第2実施形態について説明する。本実施形態における撮像装置の構成は、図1に示す第1の実施形態の場合と同様であり、以下では、第1の実施形態と同じ部分には第1の実施形態と同じ符号を付して説明を省略し、第1の実施形態と相違する点についてのみ説明する。

【0 1 4 6】

図19のフローチャートを用いて、主に撮影者意思推定部146からの情報に基づいてカメラCPU121が行う主被写体の切り替えについて説明する。なお、S 4 0 0 1～S 4 0 0 5、S 4 0 0 6～S 4 0 0 9のステップに示す処理は、図15の場合と同じであり、以下、図15と相違するS 5 0 0 6～S 5 0 0 8について説明する。

【0 1 4 7】

S 5 0 0 6では、カメラCPU121は、パン・チルト検出部144により、撮影者がパンニング／チルティング動作を行っているかを検出し、その動作を検出した場合はS 5 0 0 7へ進む。また、その動作を検出しなかった場合はS 4 0 0 6へ進む。

【0 1 4 8】

ここで、図20A～図20Cを用いて、パンニング／チルティング動作を検出した場合の主被写体の切り替え動作の例について説明する。

【0 1 4 9】

図 20 A は、撮影画角 1901 の中央像高 1902 内にプレーヤー 941 が存在し、周辺像高 1903 にプレーヤー 942 が存在する状況を示している。AF 枠 1900 は、プレーヤー 941 に設定されており、プレーヤー 942 は主被写体候補として姿勢が検出されている。

【0150】

図 20 B は、撮影者が、パンニング／チルティング動作により、故意にプレーヤー 942 を撮影画角 1901 の中央像高 1902 内に移動させた状態を示している。この時、パン・チルト検出部 144 は、パンニング／チルティング動作を検出し、カメラ CPU 121 は、プレーヤー 942 を主被写体候補から主被写体に切り替え、AF 枠 1900 をプレーヤー 942 に設定する。

【0151】

図 20 C は、図 20 B で主被写体として切り替えられたプレーヤー 942 のシュートシーンを示している。プレーヤー 942 は引き続き中央像高 1902 内の主被写体であり、周辺像高 1903 内のプレーヤー 941 は主被写体候補として姿勢が検出される。なお、中央像高 1902 の範囲は撮影条件や被写体条件により変更してもよく、また主被写体の切り替えは、主被写体候補の姿勢種別によって優先付けしてもよい。

【0152】

図 20 の S5007 では、本体姿勢判定部 145 からのカメラの移動・位置情報に基づいて、撮影者意思推定部 146 は撮影者が意図的に主被写体を切り替えようとしているかの意思を判定する。ユーザの意思により主被写体を切り替えたとは判定された場合は、S5008 へ進み、ユーザの意思により主被写体を切り替えるのではないと判定された場合は、S4009 へ進む。

【0153】

なお、S5008 で主被写体を切り替える場合には、主被写体のデフォーカス量に関係なくピント合わせを行うが、パンニング／チルティング動作に対する応答性を考慮し、主被写体へのピント遷移（フォーカス駆動制御等）を変更してもよい。

【0154】

本明細書の開示は、以下の撮像装置、方法、プログラムおよび記憶媒体を含む。

【0155】

（項目 1）

被写体を検出する被写体検出手段と、
前記被写体の姿勢を検出する姿勢検出手段と、
前記被写体の焦点状態を検出する焦点検出手段と、
前記被写体の姿勢と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定する設定手段と、
を備えることを特徴とする撮像装置。

【0156】

（項目 2）

前記主被写体として選択された被写体に焦点を合わせる焦点調節手段をさらに備えることを特徴とする項目 1 に記載の撮像装置。

【0157】

（項目 3）

前記被写体の姿勢に基づいて、前記被写体が主被写体である信頼性を取得する取得手段をさらに備えることを特徴とする項目 1 または 2 に記載の撮像装置。

【0158】

（項目 4）

前記被写体の姿勢に基づいて、前記被写体の状態を判定する状態判定手段をさらに備え、前記設定手段は、前記被写体の状態と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定することを特徴とする項目 1 乃至 3 のいずれか 1 項目に記載の撮像装置。

10

20

30

40

50

【 0 1 5 9 】

(項 目 5)

前記状態判定手段は、前記被写体の姿勢に基づいて前記被写体がどのような動作を行っているかの状態を判定することを特徴とする項目 4 に記載の撮像装置。

【 0 1 6 0 】

(項 目 6)

前記被写体の姿勢に基づいて、前記被写体が行っている競技の種別を判定する競技判定手段をさらに備えることを特徴とする項目 1 乃至 5 のいずれか 1 項目に記載の撮像装置。

【 0 1 6 1 】

(項 目 7)

前記設定手段は、前記競技の種別にさらに基づいて、前記閾値を設定することを特徴とする項目 6 に記載の撮像装置。

【 0 1 6 2 】

(項 目 8)

前記設定手段は、前記被写体の姿勢継続時間にさらに基づいて、前記閾値を設定することを特徴とする項目 1 乃至 7 のいずれか 1 項目に記載の撮像装置。

【 0 1 6 3 】

(項 目 9)

前記閾値は、前記被写体のデフォーカス量に対する閾値であることを特徴とする項目 8 に記載の撮像装置。

【 0 1 6 4 】

(項 目 1 0)

前記被写体のデフォーカス量が前記閾値以下である場合に、前記被写体を前記主被写体として選択する選択手段をさらに備えることを特徴とする項目 9 に記載の撮像装置。

【 0 1 6 5 】

(項 目 1 1)

前記設定手段は、前記姿勢継続時間が長いほど、前記閾値を大きく設定することを特徴とする項目 8 乃至 1 0 のいずれか 1 項目に記載の撮像装置。

【 0 1 6 6 】

(項 目 1 2)

前記撮像装置のパンニングまたはチルティングを検出する検出手段をさらに備えることを特徴とする項目 1 乃至 1 1 のいずれか 1 項目に記載の撮像装置。

【 0 1 6 7 】

(項 目 1 3)

前記パンニングまたはチルティングにより画面の中央に近づいた被写体を主被写体として選択する第 2 の選択手段をさらに備えることを特徴とする項目 1 2 に記載の撮像装置。

【 0 1 6 8 】

(項 目 1 4)

被写体を検出する被写体検出工程と、

前記被写体の姿勢を検出する姿勢検出工程と、

前記被写体の焦点状態を検出する焦点検出工程と、

前記被写体の姿勢と前記焦点状態とに基づいて、前記被写体を主被写体として選択するか否かを判定する閾値を設定する設定工程と、
を有することを特徴とする撮像装置の制御方法。

【 0 1 6 9 】

(項 目 1 5)

項目 1 4 に記載の制御方法の各工程をコンピュータに実行させるためのプログラム。

【 0 1 7 0 】

(項 目 1 6)

項目 1 4 に記載の制御方法の各工程をコンピュータに実行させるためのプログラムを記

10

20

30

40

50

憶したコンピュータが読み取り可能な記憶媒体。

【0171】

(他の実施形態)

また本発明は、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できる。また、1以上の機能を実現する回路(例えば、ASIC)によっても実現できる。

【0172】

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。

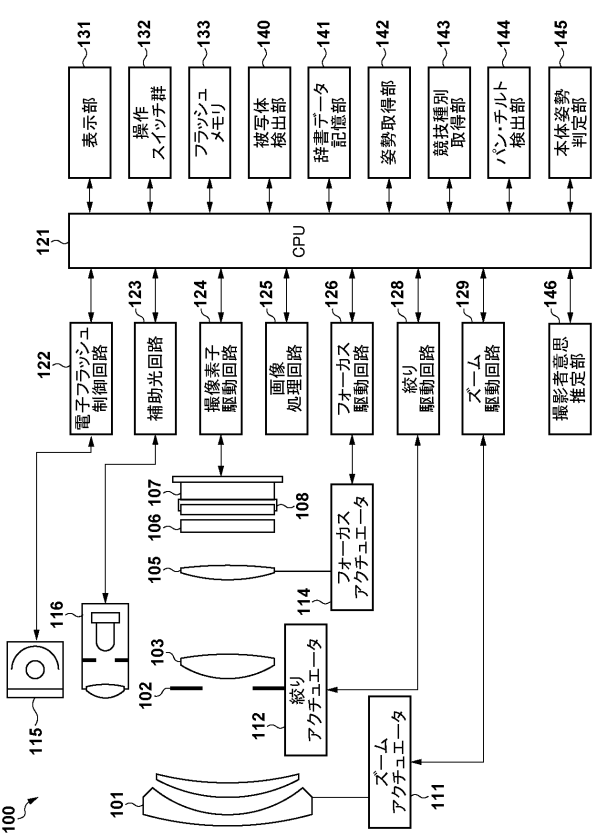
【符号の説明】

【0173】

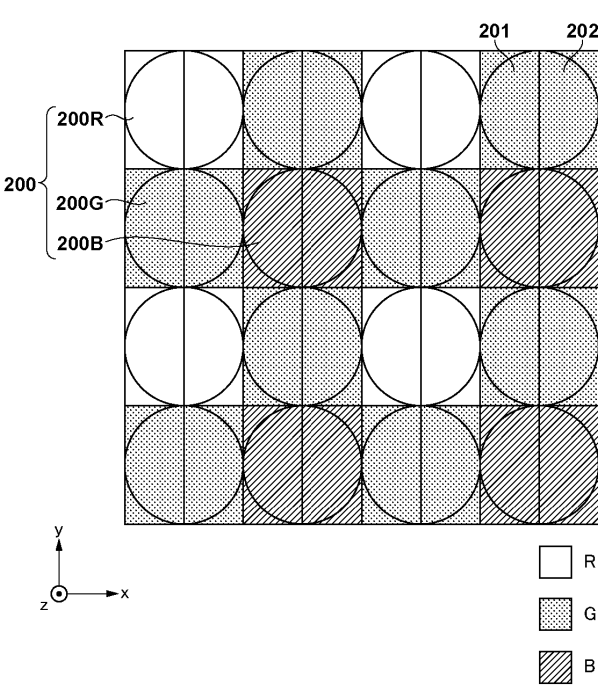
100：カメラ、107：撮像素子、121：カメラCPU、126：フォーカス駆動回路、140：被写体検出部、142：姿勢取得部、143：競技種別取得部、144：パン・チルト検出部、145：本体姿勢判定部、146：撮影者意思推定部

【図面】

【図1】



【図2】



10

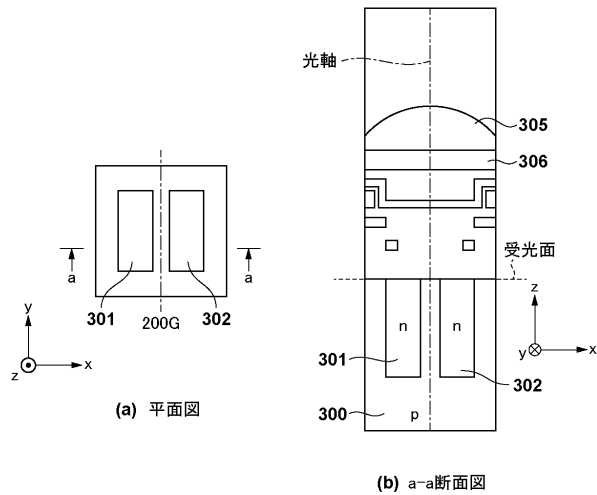
20

30

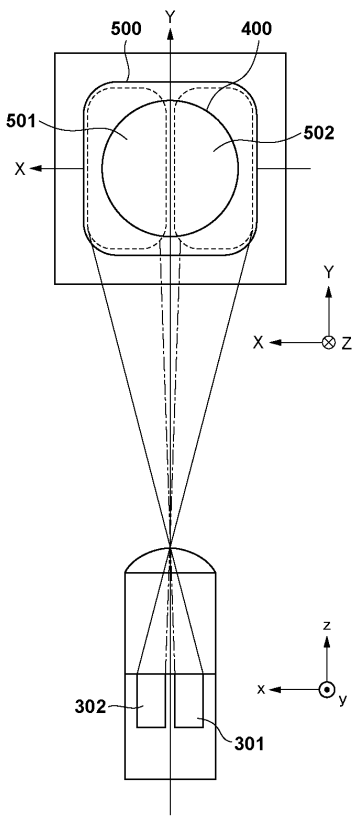
40

50

【 図 3 】



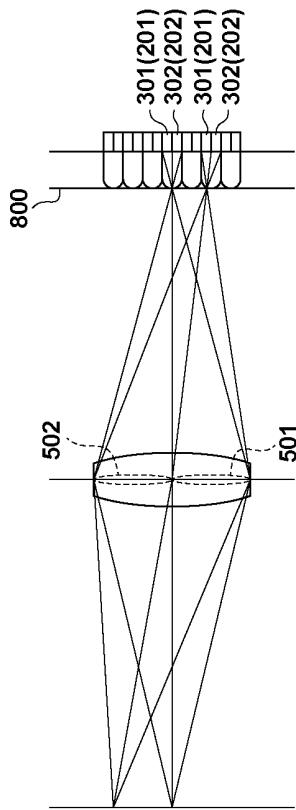
【 図 4 】



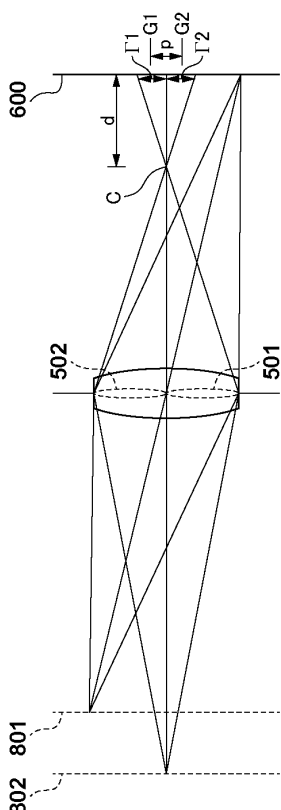
10

20

【 図 5 】



【 図 6 】

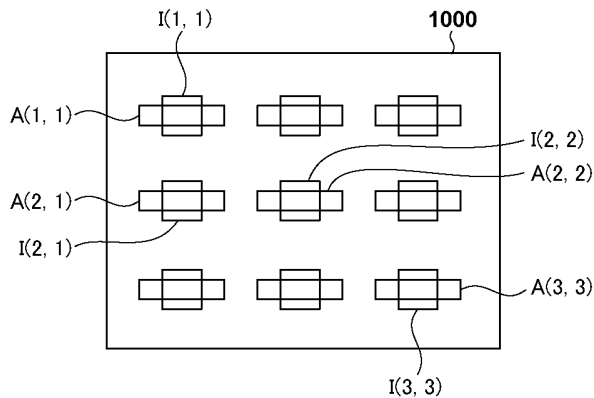


30

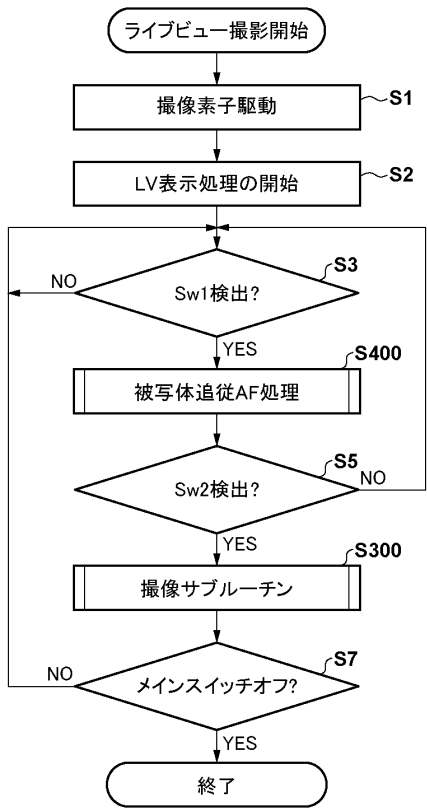
40

50

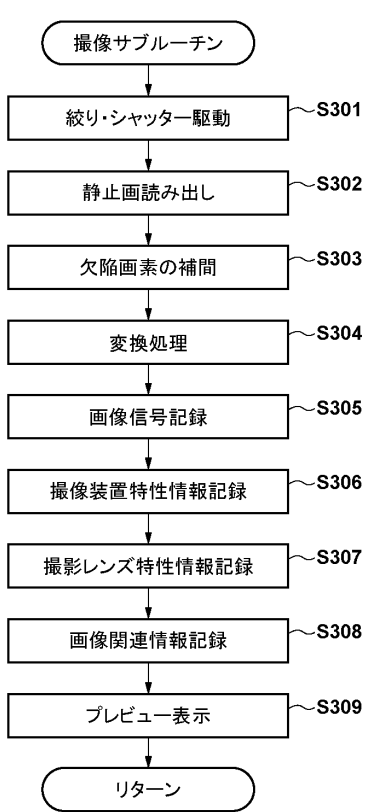
【図 7】



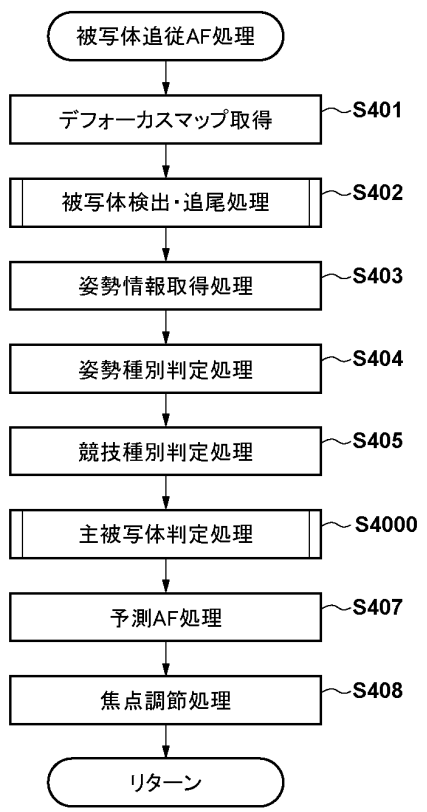
【図 8】



【図 9】



【図 10】



10

20

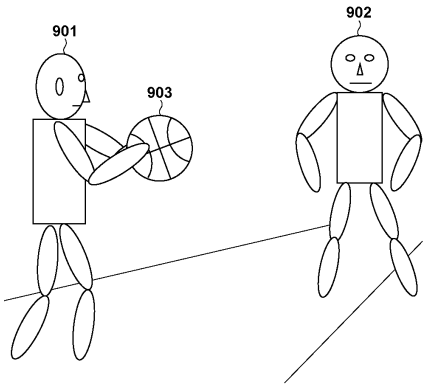
30

40

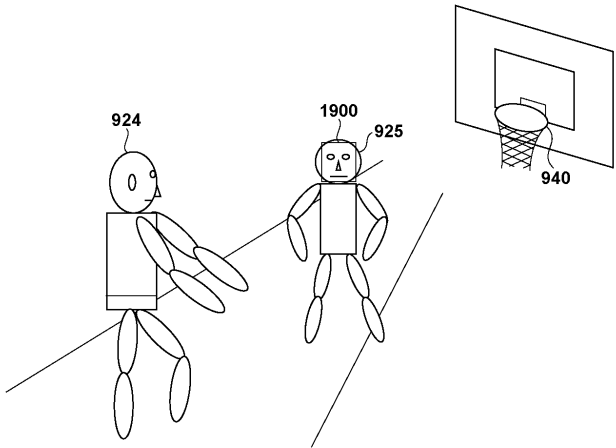
50

【図 1 1】

(a)

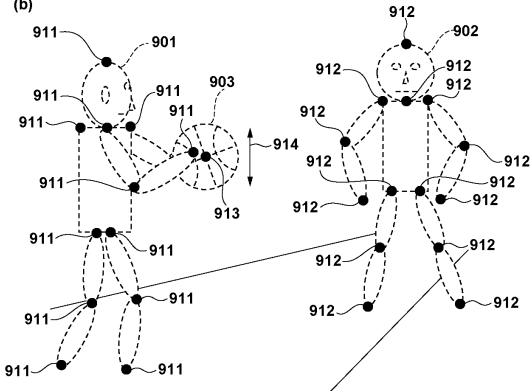


【図 1 2 A】



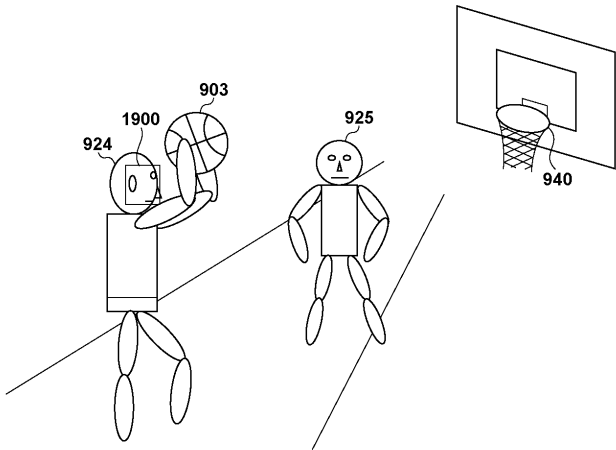
10

(b)

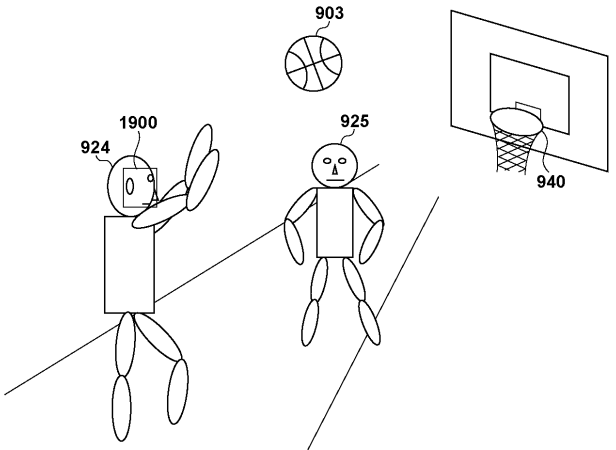


20

【図 1 2 B】



【図 1 2 C】

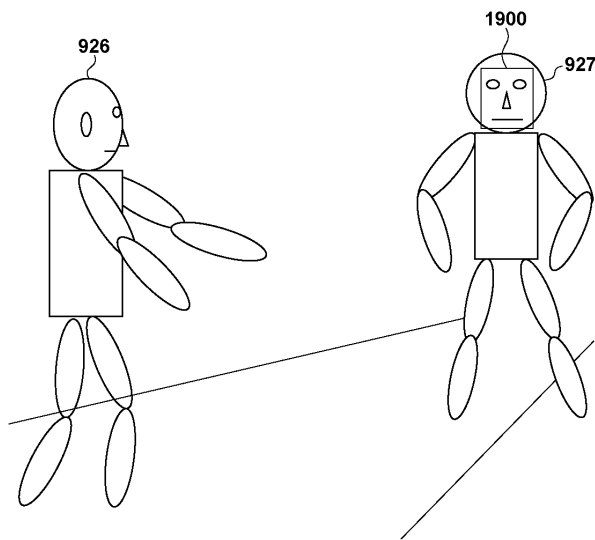


30

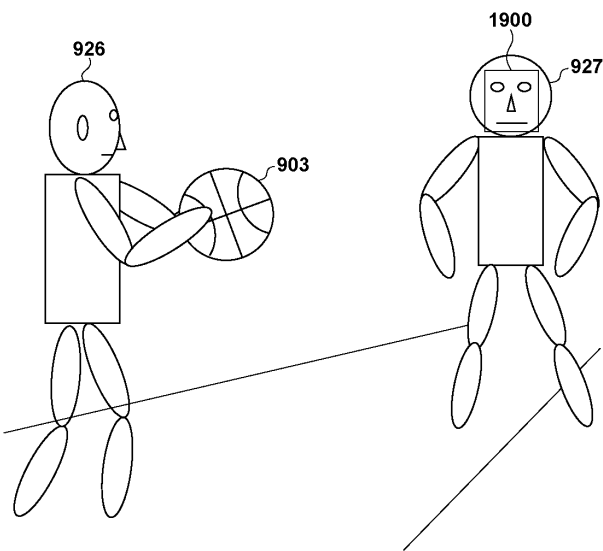
40

50

【図 1 3 A】

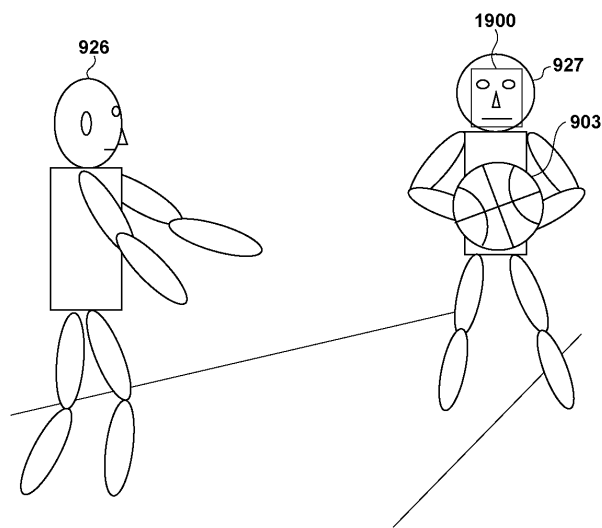


【図 1 3 B】

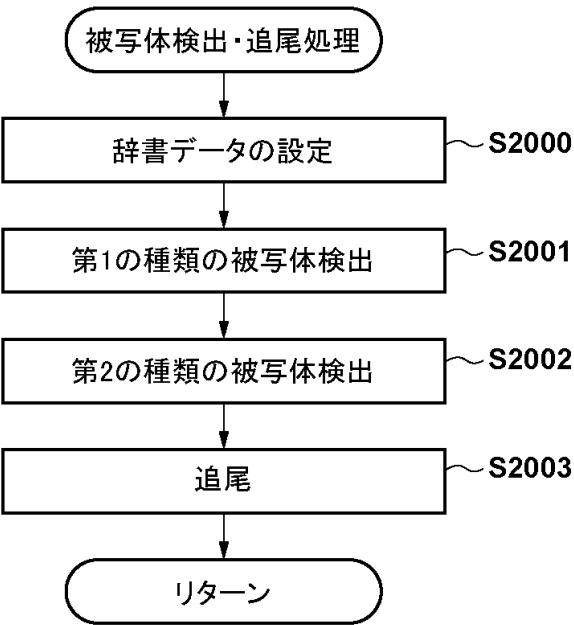


10

【図 1 3 C】



【図 1 4】



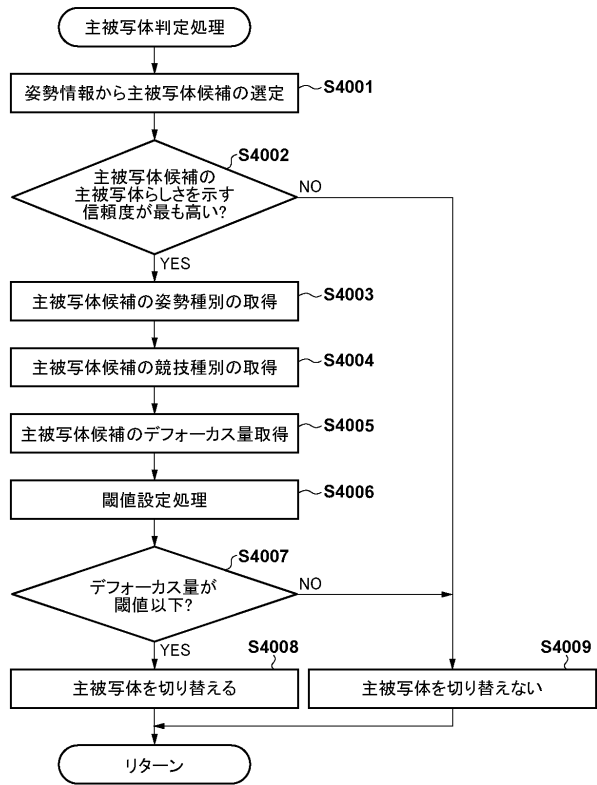
20

30

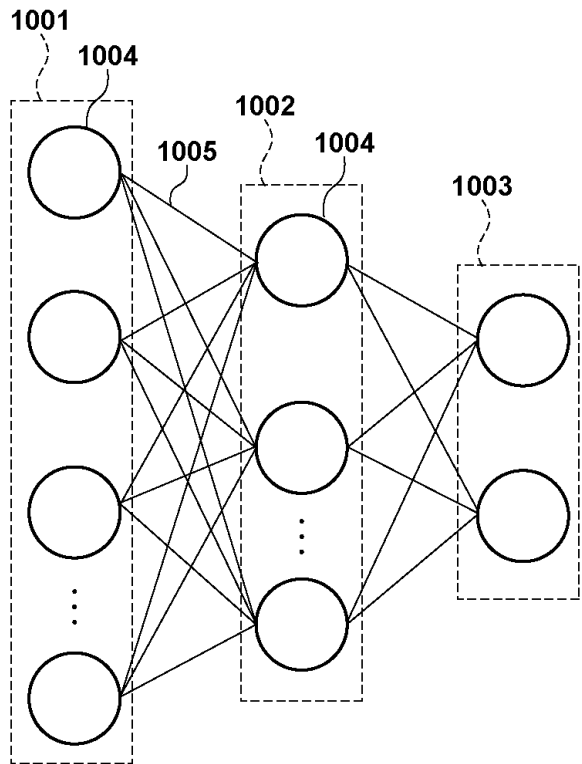
40

50

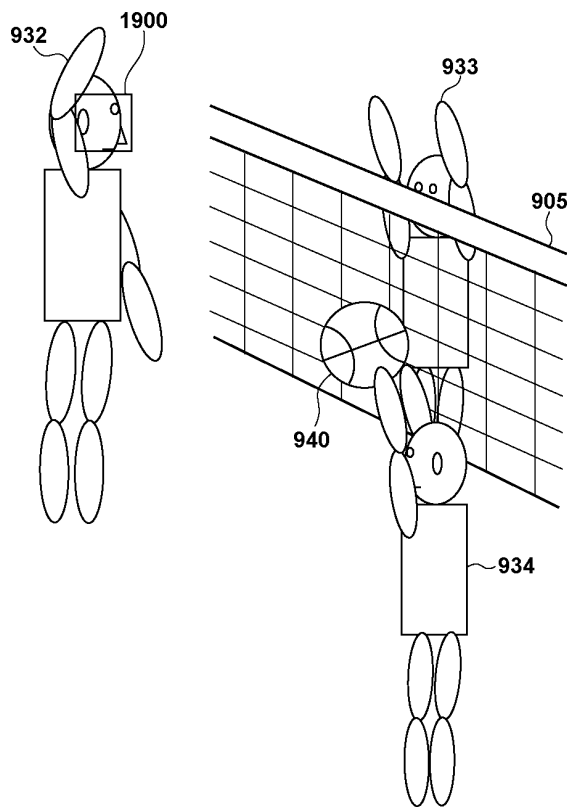
【図 15】



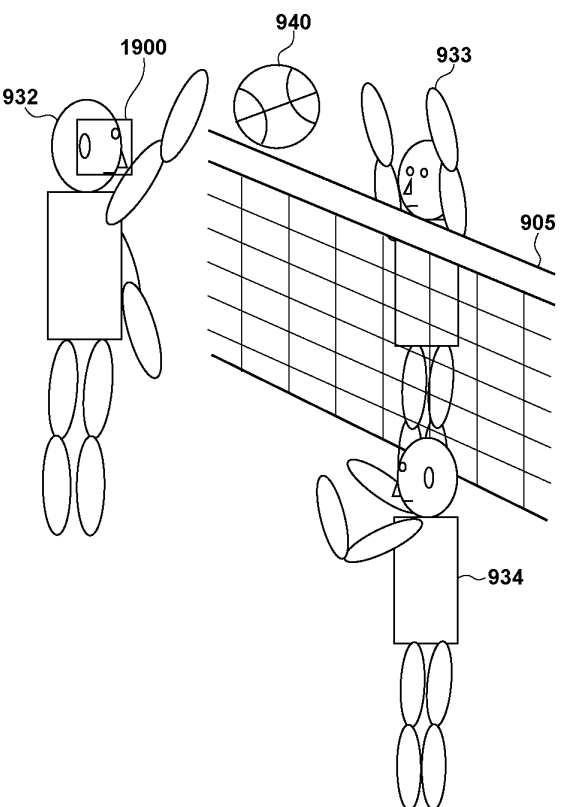
【図 16】



【図 17 A】



【図 17 B】



10

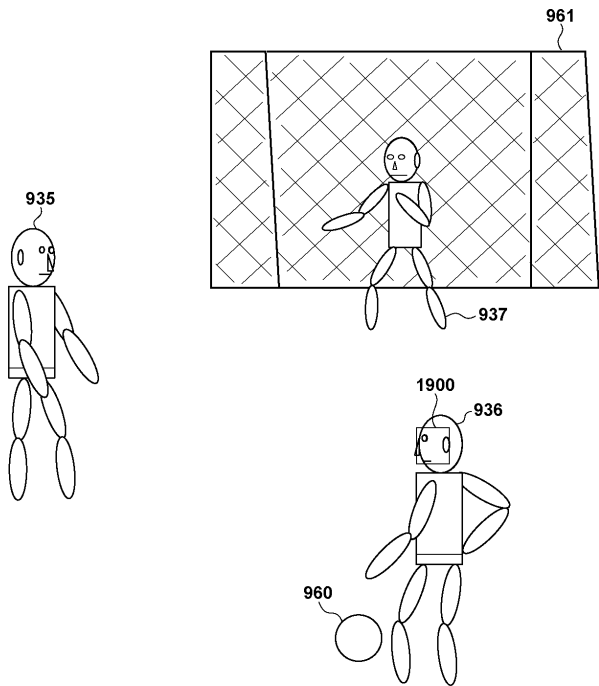
20

30

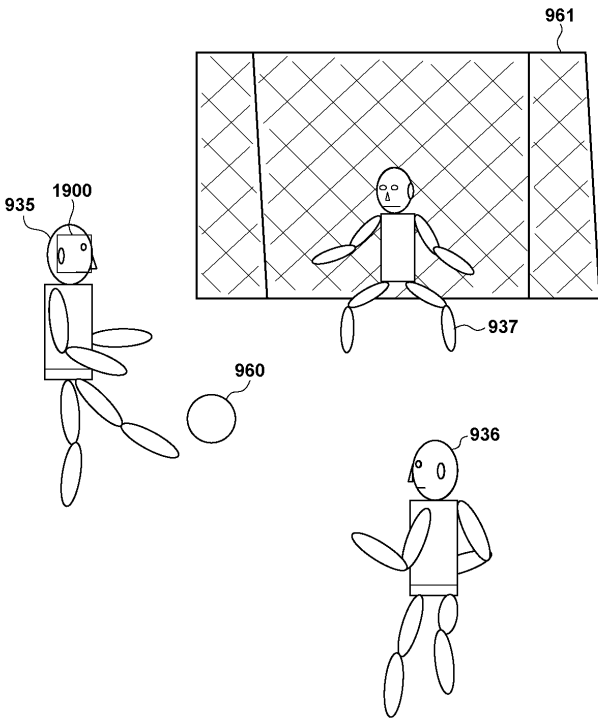
40

50

【図 18 A】



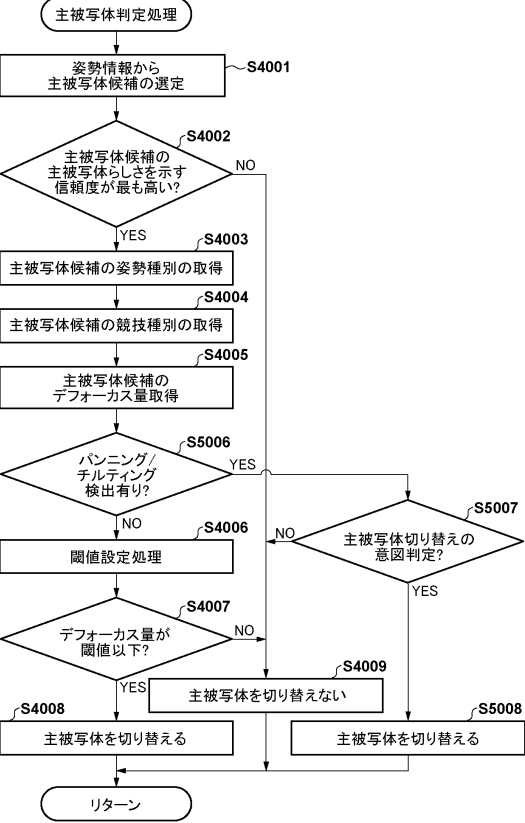
【図 18 B】



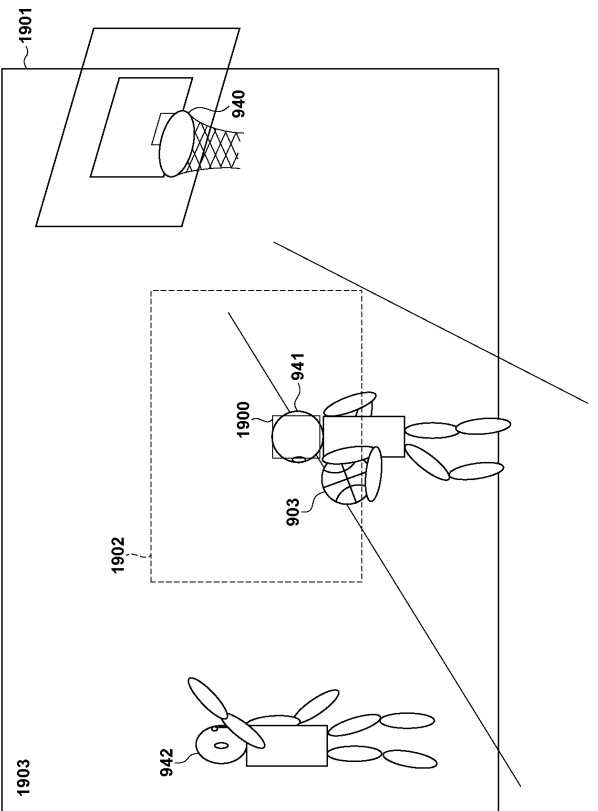
10

20

【図 19】



【図 20 A】

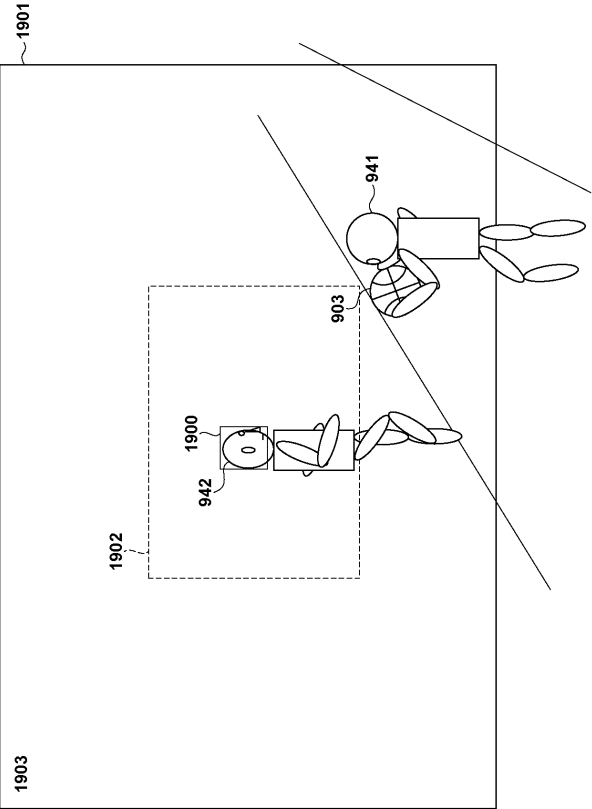


30

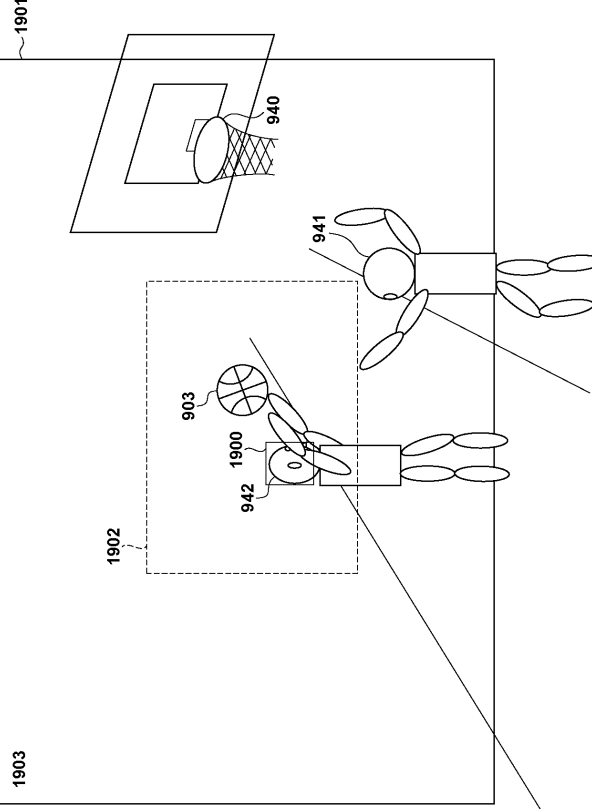
40

50

【図 20 B】



【図 20 C】



10

20

30

40

50

フロントページの続き

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 上田 暁彦
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
Fターム(参考) 2H002 GA16 GA19 GA54 GA66 HA11
2H011 BB03 CA21
2H151 CE33 CE34 DA15 DA18 DA34 DB01
5C122 DA03 DA04 EA65 FB05 FC02 FC06 FD01 FD07 FD13 FH10
FH11 FH14 HA48 HA86 HA88 HB01 HB10