

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号
特開2022-3753
(P2022-3753A)

(43) 公開日 令和4年1月11日(2022.1.11)

(51) Int.Cl.	F I	テーマコード (参考)
HO 4 N 5/232 (2006.01)	HO 4 N 5/232	2 H O 5 4
HO 4 N 5/225 (2006.01)	HO 4 N 5/232 1 9 0	5 C 1 2 2
GO 3 B 15/00 (2021.01)	HO 4 N 5/225 4 5 0	
GO 3 B 19/00 (2021.01)	HO 4 N 5/232 9 3 3	
GO 3 B 31/00 (2021.01)	GO 3 B 15/00 Q	

審査請求 未請求 請求項の数 9 O L (全 33 頁) 最終頁に続く

(21) 出願番号	特願2020-108235 (P2020-108235)	(71) 出願人	314012076
(22) 出願日	令和2年6月23日 (2020.6.23)		パナソニックIPマネジメント株式会社
		(74) 代理人	100106518
			弁理士 松谷 道子
		(74) 代理人	100132241
			弁理士 岡部 博史
		(74) 代理人	100199314
			弁理士 竹内 寛
		(72) 発明者	春日井 宏樹
			大阪府門真市大字門真1006番地 パナソニック株式会社内
		(72) 発明者	山本 真一
			大阪府門真市大字門真1006番地 パナソニック株式会社内

最終頁に続く

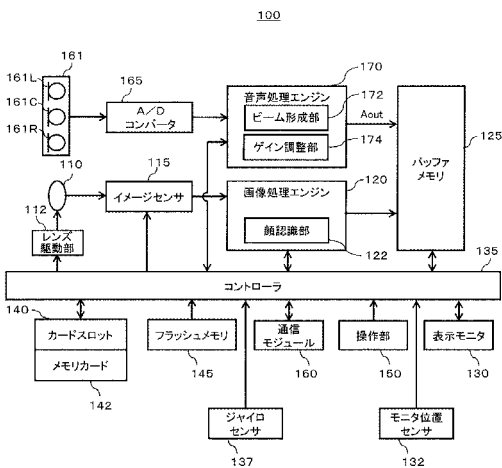
(54) 【発明の名称】 撮像装置

(57) 【要約】

【課題】 音声を取得しながら撮像を行う撮像装置において、ユーザの意図に沿って被写体の音声を收音し易くすることができる撮像装置を提供する。

【解決手段】 撮像装置は、被写体を撮像して画像データを生成する撮像部と、撮像部による撮像中に收音される音声を示す音声信号を取得する音声取得部と、ユーザの指示を受けて、音声取得部の指向性を自動で変える動作モードであるオートモードに自装置を設定する設定部と、音声信号において被写体からの音声を收音する收音エリアを制御する制御部とを備え、制御部は、設定部がオートモードに設定している場合、音声取得部の指向性を自装置の撮影状態に連動して変えることで、被写体を含めるように收音エリアを制御する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

被写体を撮像して画像データを生成する撮像部と、
前記撮像部による撮像中に收音される音声を示す音声信号を取得する音声取得部と、
ユーザの指示を受けて、前記音声取得部の指向性を自動で変える動作モードであるオートモードに自装置を設定する設定部と、
前記音声信号において前記被写体からの音声を收音する收音エリアを制御する制御部とを備え、
前記制御部は、前記設定部が前記オートモードに設定している場合、前記音声取得部の指向性を自装置の撮影状態に連動して変えることで、前記被写体を含めるように前記收音エリアを制御する、撮像装置。

10

【請求項 2】

前記画像データにおいて前記被写体の顔領域を検出する顔検出部を備え、
前記制御部は、前記設定部が前記オートモードに設定している場合、前記顔検出部によって検出された前記顔領域に基づいて前記音声信号における收音対象とする被写体を決定し、前記收音対象に決定した被写体を含めるように、前記收音エリアを制御する、請求項 1 記載の撮像装置。

【請求項 3】

前記制御部は、前記設定部が前記オートモードに設定している場合、自装置が縦向きか又は横向きかで撮影する撮影状態に連動して、前記音声取得部の指向性を変えるように、前記收音エリアを制御する、請求項 1 記載の撮像装置。

20

【請求項 4】

前記制御部は、前記設定部が前記オートモードに設定している場合、撮影者が自身を撮影するか否かの撮影状態に連動して、前記音声取得部の指向性を変えるように、前記收音エリアを制御する、請求項 1 記載の撮像装置。

【請求項 5】

前記被写体の画像を表示する表示面を有し、当該表示面を前記被写体側に変位可能な表示部と、

前記表示部が前記被写体側に前記表示面を変位したか否かを検知する検知部とをさらに備え、

30

前記設定部は、前記検知部が、前記表示部が前記被写体側に前記表示面を変位したことを検知すると、前記オートモードに設定する、請求項 1 記載の撮像装置。

【請求項 6】

前記設定部は、前記ユーザの指示に応じて、前記オートモードのほかに前記音声取得部の指向性が互いに異なる複数の動作モードうちの少なくとも 1 つに設定可能である、請求項 1 記載の撮像装置。

【請求項 7】

前記被写体の画像を表示する表示部と、

前記表示部に表示された被写体を、前記被写体からの音声を收音する收音エリアに設定するユーザ操作を入力する入力部とをさらに備え、

40

前記制御部は、前記收音エリアを設定するユーザ操作が入力されると、前記ユーザ操作に基づいて、前記音声取得部の指向性を変えることで前記被写体を含めるように、前記收音エリアを制御する、請求項 1 記載の撮像装置。

【請求項 8】

被写体を撮像して画像データを生成する撮像部と、

前記撮像部による撮像中に收音される音声を示す音声信号を取得する音声取得部と、

前記被写体の画像を表示する表示部と、

前記表示部に表示された被写体を、前記被写体からの音声を收音する收音エリアに設定するユーザ操作を入力する入力部と、

50

前記音声信号において前記收音エリアを制御する制御部とを備え、

前記制御部は、前記收音エリアを設定するユーザ操作が入力されると、前記ユーザ操作に基づいて、前記音声取得部の指向性を変えることで前記被写体を含めるように、前記收音エリアを制御する、撮像装置。

【請求項 9】

前記表示部は、前記被写体側に表示面を変位可能に構成され、

前記表示部が、前記被写体側に前記表示面が変位された状態において、前記ユーザ操作がなされる、請求項 8 記載の撮像装置。

【発明の詳細な説明】

【技術分野】

10

【0001】

本開示は、音声を取得しながら撮像を行う撮像装置に関する。

【背景技術】

【0002】

特許文献 1 は、顔検出機能を有するビデオカメラを開示している。特許文献 1 のビデオカメラは、ズーム比率及び撮影した画面内の人物の顔の大きさに応じて、マイクの指向角を変化させる。これにより、当該ビデオカメラは、ビデオカメラと被写体映像の距離に、マイクの指向角を関連付けて制御することで、映像と音声の整合をとりつつ、被写体の声をより確実に捉えるようにマイクの指向角を変化させる制御の実現を図っている。この際、当該ビデオカメラは、人物（被写体）の顔の位置及び大きさを検出し、検出した顔部分に枠（顔検出枠）を付けて表示するとともに、顔検出枠の大きさ（顔の大きさ）の情報を利用する。

20

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2010 - 283706 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

本開示は、音声を取得しながら撮像を行う撮像装置において、ユーザの意図に沿って被写体の音声を收音し易くすることができる撮像装置を提供する。

30

【課題を解決するための手段】

【0005】

本開示において、撮像装置は、被写体を撮像して画像データを生成する撮像部と、撮像部による撮像中に收音される音声を示す音声信号を取得する音声取得部と、ユーザの指示を受けて、音声取得部の指向性を自動で変える動作モードであるオートモードに自装置を設定する設定部と、音声信号において被写体からの音声を收音する收音エリアを制御する制御部とを備え、制御部は、設定部がオートモードに設定している場合、音声取得部の指向性を自装置の撮影状態に連動して変えることで、被写体を含めるように收音エリアを制御する。

40

【発明の効果】

【0006】

本開示に係る撮像装置によると、音声を取得しながら撮像を行う撮像装置において、ユーザの意図に沿って被写体の音声を收音し易くすることができる。

【図面の簡単な説明】

【0007】

【図 1】本開示の実施の形態 1 に係るデジタルカメラ 100 の構成を示す図

【図 2】デジタルカメラ 100 の背面を例示する図

【図 3】自分撮り時のデジタルカメラ 100 の状態を例示する図

【図 4】縦撮り時のデジタルカメラ 100 の状態を例示する図

50

- 【図 5】デジタルカメラ 100 におけるビーム形成部 172 の構成を例示する図
- 【図 6】デジタルカメラ 100 において收音エリアを例示する図
- 【図 7】デジタルカメラ 100 における設定メニューの表示例を示す図
- 【図 8】デジタルカメラ 100 において更なる收音エリアを例示する図
- 【図 9】デジタルカメラ 100 のオートモードの動作の概要を説明するための図
- 【図 10】実施の形態 1 に係るデジタルカメラ 100 のフォーカスモードの動作を例示するフローチャート
- 【図 11】デジタルカメラ 100 のフォーカスモードの動作の概要を説明するための図
- 【図 12】実施の形態 1 に係るデジタルカメラ 100 の收音対象の選別処理（図 10 の S3）を例示するフローチャート
- 【図 13】デジタルカメラ 100 における收音対象の選別処理を説明するための図
- 【図 14】デジタルカメラ 100 における收音エリアの決定処理（図 10 の S4）を例示するフローチャート
- 【図 15】デジタルカメラ 100 における收音エリアの決定処理を説明するための図
- 【図 16】デジタルカメラ 100 における顔認識を用いた收音制御（図 10 の S5）を例示するフローチャート
- 【図 17】收音エリアの決定処理によって得られる管理情報を説明するための図
- 【図 18】デジタルカメラ 100 における水平画角および合焦距離からゲインを求める関係を例示する図
- 【図 19】デジタルカメラ 100 における顔認識を用いない收音制御（図 10 の S6）を例示するフローチャート
- 【図 20】実施の形態 1 に係るデジタルカメラ 100 のオートモードの動作を例示するフローチャート
- 【図 21】実施の形態 1 のオートモードにおける横撮り時と縦撮り時の收音制御を説明するための図
- 【図 22】実施の形態 2 に係るデジタルカメラ 100 の表示例を示す図
- 【図 23】実施の形態 2 のデジタルカメラ 100 におけるマニュアル操作を説明するための図
- 【図 24】実施の形態 2 のデジタルカメラ 100 におけるマニュアル操作時の動作を例示するフローチャート
- 【図 25】変形例のデジタルカメラ 100 A におけるマイク 161 A の配置を例示する図
- 【図 26】変形例のオートモードにおける縦撮り時の收音制御を説明するための図
- 【図 27】デジタルカメラ 100 における顔認識に連動した收音制御の動作例を説明するための図

10

20

30

40

【発明を実施するための形態】

【0008】

以下、適宜図面を参照しながら、実施の形態を詳細に説明する。但し、必要以上に詳細な説明は省略する場合がある。例えば、既によく知られた事項の詳細説明や実質的に同一の構成に対する重複説明を省略する場合がある。これは、以下の説明が不必要に冗長になるのを避け、当業者の理解を容易にするためである。なお、発明者（ら）は、当業者が本開示を十分に理解するために添付図面および以下の説明を提供するのであって、これらによって特許請求の範囲に記載の主題を限定することを意図するものではない。

【0009】

（実施の形態 1）

実施の形態 1 では、本開示に係る撮像装置の一例として、画像認識技術に基づいて被写体を検出し、検出した被写体の大きさに応じた收音エリアの制御、及び收音する音声を強調する收音ゲインの制御を行うデジタルカメラについて説明する。

【0010】

〔1-1. 構成〕

【0011】

50

図 1 は、本実施形態に係るデジタルカメラ 100 の構成を示す図である。本実施形態のデジタルカメラ 100 は、イメージセンサ 115 と、画像処理エンジン 120 と、表示モニタ 130 と、コントローラ 135 とを備える。さらに、デジタルカメラ 100 は、バッファメモリ 125 と、カードスロット 140 と、フラッシュメモリ 145 と、操作部 150 と、通信モジュール 160 とを備える。また、デジタルカメラ 100 は、マイク 161 と、マイク用のアナログ/デジタル (A/D) コンバータ 165 と、音声処理エンジン 170 とを備える。また、デジタルカメラ 100 は、例えば光学系 110 及びレンズ駆動部 112 を備える。さらに、デジタルカメラ 100 は、例えば磁気センサ 132 と、加速度センサ 137 とを備える。

【0012】

図 2 は、デジタルカメラ 100 の背面を例示する。図 2 では、デジタルカメラ 100 の 3 軸方向 X, Y, Z と共に、重力方向 G を例示している。X, Y および Z 軸は、それぞれデジタルカメラ 100 の水平画角方向、垂直画角方向および光学系 110 におけるレンズの光軸方向に対応する。図 2 の例では、デジタルカメラ 100 の Y 軸方向が重力方向 G に沿った向き即ち横向きとなっている。

【0013】

本実施形態のデジタルカメラ 100 は、ユーザが自身を撮影する自分撮りをしたり、デジタルカメラ 100 を縦向きに用いる縦撮りをしたりして利用可能である。図 3 は、自分撮り時のデジタルカメラ 100 の状態を例示する。図 4 は、縦撮り時のデジタルカメラ 100 の状態を例示する。

【0014】

図 1 に戻り、光学系 110 は、フォーカスレンズ、ズームレンズ、光学式手ぶれ補正レンズ (OIS)、絞り、シャッタ等を含む。フォーカスレンズは、イメージセンサ 115 上に形成される被写体像のフォーカス状態を変化させるためのレンズである。ズームレンズは、光学系で形成される被写体像の倍率を変化させるためのレンズである。フォーカスレンズ等は、それぞれ 1 枚又は複数枚のレンズで構成される。

【0015】

レンズ駆動部 112 は、光学系 110 におけるフォーカスレンズ等を駆動する。レンズ駆動部 112 はモータを含み、コントローラ 135 の制御に基づいてフォーカスレンズを光学系 110 の光軸に沿って移動させる。レンズ駆動部 112 においてフォーカスレンズを駆動する構成は、DC モータ、ステッピングモータ、サーボモータ、または超音波モータなどで実現できる。

【0016】

イメージセンサ 115 は、光学系 110 を介して形成された被写体像を撮像して、撮像データを生成する。撮像データは、イメージセンサ 115 による撮像画像を示す画像データを構成する。イメージセンサ 115 は、所定のフレームレート (例えば、30 フレーム/秒) で新しいフレームの画像データを生成する。イメージセンサ 115 における、撮像データの生成タイミングおよび電子シャッタ動作は、コントローラ 135 によって制御される。イメージセンサ 115 は、CMOS イメージセンサ、CCD イメージセンサ、または NMOS イメージセンサなど、種々のイメージセンサを用いることができる。

【0017】

イメージセンサ 115 は、動画像、静止画像の撮像動作、スルー画像の撮像動作等を実行する。スルー画像は主に動画像であり、ユーザが例えば静止画像の撮像のための構図を決めるために表示モニタ 130 に表示される。

スルー画像、動画像及び静止画像は、それぞれ本実施形態における撮像画像の一例である。イメージセンサ 115 は、本実施形態における撮像部の一例である。

【0018】

画像処理エンジン 120 は、イメージセンサ 115 から出力された撮像データに対して各種の処理を施して画像データを生成したり、画像データに各種の処理を施して、表示モニタ 130 に表示するための画像を生成したりする。各種処理としては、ホワイト balan

10

20

30

40

50

ス補正、ガンマ補正、Y C 変換処理、電子ズーム処理、圧縮処理、伸張処理等が挙げられるが、これらに限定されない。画像処理エンジン 120 は、ハードワイヤードな電子回路で構成してもよいし、プログラムを用いたマイクロコンピュータ、プロセッサなどで構成してもよい。

【0019】

本実施形態において、画像処理エンジン 120 は、撮像画像の画像認識によって人の顔といった被写体の検出機能を実現する顔認識部 122 を含む。顔認識部 122 は、例えば、ルールベースの画像認識処理によって顔検出を行い、検出情報を出力する。顔検出は、種々の画像認識アルゴリズムによって行われてもよい。検出情報は、被写体の検出結果に対応する位置情報を含む。位置情報は、例えば処理対象の画像 Im 上の水平位置及び垂直位置で規定され、例えば検出された被写体として人の顔を矩形状に囲む領域を示す（図 1 参照）。

10

【0020】

表示モニタ 130 は、種々の情報を表示する表示部の一例である。例えば、表示モニタ 130 は、イメージセンサ 115 で撮像され、画像処理エンジン 120 で画像処理された画像データが示す画像（スルー画像）を表示する。また、表示モニタ 130 は、ユーザがデジタルカメラ 100 に対して種々の設定を行うためのメニュー画面等を表示する。表示モニタ 130 は、例えば、液晶ディスプレイデバイスまたは有機 EL デバイスで構成できる。

【0021】

20

本実施形態のデジタルカメラ 100 は、例えば図 2, 3 に示すように、表示モニタ 130 の位置を変更可能な可動式で構成される。図 2 の例では、表示モニタ 130 が、デジタルカメラ 100 の背面側（-Z 側）に表示面を向けた位置にある。こうした表示モニタ 130 の位置を以下「通常位置」という。図 3 の例では、表示モニタ 130 が、デジタルカメラ 100 の前面側（+Z 側）即ち被写体側に表示面を向けた位置にある。こうした表示モニタ 130 の位置を以下「自分撮り位置」という。

【0022】

磁気センサ 132 は、表示モニタ 130 が通常位置にあるか又は自分撮り位置にあるかを検知する検知部の一例である。磁気センサ 132 は、例えば表示モニタ 130 の位置の検出結果を示す検出信号をコントローラ 135 に出力する。

30

【0023】

可動式の表示モニタ 130 としては、例えばバリエーション式またはチルト式を採用できる。例えば、表示モニタ 130 を回動可能にデジタルカメラ 100 の本体に連結するヒンジ 131 が設けられる。磁気センサ 132 は、例えばヒンジ 131 内部に設けられ、図 2, 3 に対応する 2 状態を有するスイッチ等で構成される。

【0024】

加速度センサ 137 は、例えば 3 軸方向 X, Y, Z の内の 1 つ又は複数の角速度を検出して、検出信号をコントローラ 135 に出力する。加速度センサ 137 は、重力加速度の検出状態に基づき、デジタルカメラ 100 の姿勢が、図 2 に例示したような横向きか、又は図 4 に例示したような縦向きかを検出する姿勢検出部の一例である。

40

【0025】

操作部 150 は、デジタルカメラ 100 の外装に設けられた操作釦や操作レバー等のハードキーの総称であり、使用者による操作を受け付ける。操作部 150 は、例えば、リリース釦、モードダイヤル、タッチパネル、カーソルボタン、ジョイスティックを含む。操作部 150 はユーザによる操作を受け付けると、ユーザ操作に対応した操作信号をコントローラ 135 に送信する。操作部 150 は、例えば図 2 に示すように、リリースボタン 151、選択ボタン 152、決定ボタン 153、機能ボタン 154 およびタッチパネル 155 等を含む。

【0026】

コントローラ 135 は、デジタルカメラ 100 全体の動作を統括制御する。コントロー

50

ラ 1 3 5 は C P U 等を含み、C P U がプログラム（ソフトウェア）を実行することで所定の機能を実現する。コントローラ 1 3 5 は、C P U に代えて、所定の機能を実現するように設計された専用の電子回路で構成されるプロセッサを含んでもよい。すなわち、コントローラ 1 3 5 は、C P U、M P U、G P U、D S U、F P G A、A S I C 等の種々のプロセッサで実現できる。コントローラ 1 3 5 は 1 つまたは複数のプロセッサで構成してもよい。また、コントローラ 1 3 5 は、画像処理エンジン 1 2 0 などと共に 1 つの半導体チップで構成してもよい。

【 0 0 2 7 】

バッファメモリ 1 2 5 は、画像処理エンジン 1 2 0 やコントローラ 1 3 5 のワークメモリとして機能する記録媒体である。バッファメモリ 1 2 5 は、D R A M (D y n a m i c R a n d o m A c c e s s M e m o r y) などにより実現される。フラッシュメモリ 1 4 5 は不揮発性の記録媒体である。また、図示していないが、コントローラ 1 3 5 は各種の内部メモリを有してもよく、例えば R O M を内蔵してもよい。R O M には、コントローラ 1 3 5 が実行する様々なプログラムが記憶されている。また、コントローラ 1 3 5 は、C P U の作業領域として機能する R A M を内蔵してもよい。

【 0 0 2 8 】

カードスロット 1 4 0 は、着脱可能なメモリカード 1 4 2 が挿入される手段である。カードスロット 1 4 0 は、メモリカード 1 4 2 を電氣的及び機械的に接続可能である。メモリカード 1 4 2 は、内部にフラッシュメモリ等の記録素子を備えた外部メモリである。メモリカード 1 4 2 は、画像処理エンジン 1 2 0 で生成される画像データなどのデータを格納できる。

【 0 0 2 9 】

通信モジュール 1 6 0 は、通信規格 I E E E 8 0 2 . 1 1 または W i - F i 規格等に準拠した通信を行う通信モジュール（回路）である。デジタルカメラ 1 0 0 は、通信モジュール 1 6 0 を介して、他の機器と通信することができる。デジタルカメラ 1 0 0 は、通信モジュール 1 6 0 を介して、他の機器と直接通信を行ってもよいし、アクセスポイント経由で通信を行ってもよい。通信モジュール 1 6 0 は、インターネット等の通信ネットワークに接続可能であってもよい。

【 0 0 3 0 】

マイク 1 6 1 は、音を收音する收音部の一例である。マイク 1 6 1 は、收音した音声を電気信号であるアナログ信号に変換して出力する。本実施形態のマイク 1 6 1 は、3 つのマイクロフォン素子 1 6 1 L , 1 6 1 C 及び 1 6 1 R を含む。マイク 1 6 1 は、2 つ又は 4 つ以上のマイクロフォン素子から構成されてもよい。

【 0 0 3 1 】

マイク用の A / D コンバータ 1 6 5 は、マイク 1 6 1 からのアナログ信号をデジタル信号の音声データに変換する。マイク用の A / D コンバータ 1 6 5 は、本実施形態における音声取得部の一例である。なお、マイク 1 6 1 は、デジタルカメラ 1 0 0 の外部にあるマイクロフォン素子を含んでもよい。この場合、デジタルカメラ 1 0 0 は音声取得部として、外部のマイク 1 6 1 に対するインタフェース回路を備える。

【 0 0 3 2 】

音声処理エンジン 1 7 0 は、マイク用の A / D コンバータ 1 6 5 等の音声取得部から出力された音声データを受信して、受信した音声データに対して種々の音声処理を施す。音声処理エンジン 1 7 0 は、本実施形態における音声処理部の一例である。

【 0 0 3 3 】

本実施形態の音声処理エンジン 1 7 0 は、例えば図 1 に示すように、ビーム形成部 1 7 2 と、ゲイン調整部 1 7 4 とを備える。ビーム形成部 1 7 2 は、音声の指向性を制御する機能を実現する。ビーム形成部 1 7 2 の詳細については後述する。ゲイン調整部 1 7 4 は、入力される音声データに、例えばコントローラ 1 3 5 によって設定される收音ゲインを乗じる乗算処理を行って、音声を増幅する。ゲイン調整部 1 7 4 は、入力の音声データに負のゲインを乗じて音声を抑圧する処理を行ってもよい。收音ゲイン調整部 1 4 はさらに

10

20

30

40

50

、入力される音声データの周波数特性及びステレオ特性を変化させる機能を有してもよい。收音ゲインの設定についての詳細は後述する。

【0034】

〔1-1-1. ビーム形成部について〕

本実施形態におけるビーム形成部172の詳細を、以下説明する。

【0035】

ビーム形成部172は、マイク161が收音する音声の指向性を制御するビームフォーミングを行う。本実施形態におけるビーム形成部172の構成例を図5に示す。

【0036】

図5に示すように、ビーム形成部172は、例えばフィルタD1～D3と加算器173を備え、各マイクロフォン素子161L, 161C及び161Rで收音された音声の遅延期間を調整して、その重み付き和を出力する。ビーム形成部172によると、マイク161の收音指向性の方向および範囲を制御して、マイク161が收音する物理的な範囲を設定できる。

【0037】

ビーム形成部172は、図示では1つの加算器173により1チャンネルの出力を行うが、2つ以上の加算器を備え、例えばステレオ出力のような各チャンネルで異なる出力を行う構成であってもよい。また、加算器173の他に減算器を用いて、特に感度が低い方向である死角を特定方向に有する指向性を形成してもよいし、環境に適応して処理を変える適応ビームフォーミングを行ってもよい。また、音声信号の周波数帯域によって異なる処理を適用してもよい。

【0038】

図5では、マイクロフォン素子161L, 161C及び161Rを直線的に配置した例を示しているが、各マイクロフォン素子の配置は、これに限らない。例えば、三角形に配置する場合であっても、フィルタD1～D3の遅延期間及び重みを適宜調整して、マイク161の收音指向性を制御できる。また、ビーム形成部172は、收音指向性の制御に公知の手法を適用してもよい。例えば、OZO Audioといった音声処理技術を用いて、指向性を形成する処理を行い、併せて音声の雑音を抑制する処理等を実行してもよい。

【0039】

上記のようなビーム形成部172により設定可能なデジタルカメラ100の收音エリアについて説明する。

【0040】

〔1-1-2. 收音エリアについて〕

図6は、デジタルカメラ100において定義される收音エリアの例を示す。図6は、收音エリアを、デジタルカメラ100を中心とする円の扇形領域によって例示している。本実施形態のデジタルカメラ100では、水平画角方向は、マイクロフォン素子161R, 161Cおよび161Lが並ぶ方向と一致する。

【0041】

図6(A)は、角度範囲401(例えば70°)において、デジタルカメラ100の前方(すなわち撮影方向)に收音エリアを向ける「前方中心收音エリア」41を示す。図6(B)は、角度範囲401において、デジタルカメラ100の左方に收音エリアを向ける「左半分收音エリア」42を示す。図6(C)は、角度範囲401において、デジタルカメラ100の右方に收音エリアを向ける「右半分收音エリア」43を示す。図6(D)は、角度範囲401より大きい角度範囲402(例えば160°)において、デジタルカメラ100の前方に收音エリアを向ける「前方收音エリア」44を示す。これらの收音エリアは、本実施形態における複数の所定エリアの一例であり、角度範囲401及び402は、第1の角度範囲及び第2の角度範囲の一例である。

【0042】

本実施形態のデジタルカメラ100は、被写体が撮像画像の中心部分に位置するとき、

図 6 (A) の前方中心收音エリア 4 1 を用いる。また、被写体が撮像画像の左半分に位置するとき、図 6 (B) の左半分收音エリア 4 2 を用い、被写体が撮像画像の右半分に位置するとき、図 6 (C) の右半分收音エリア 4 3 を用いる。さらに、被写体が撮像画像の全体に位置するとき、主に図 6 (D) の前方收音エリア 4 4 を用いる。

【 0 0 4 3 】

図 1 1 (B) の例では、收音対象の被写体 R 1 及び R 3 が撮像画像の中心部分に位置するため、前方中心收音エリア 4 1 が用いられる。図 1 1 (C) の例では、收音対象の被写体 R 1 及び R 2 が撮像画像の左半分に位置するため、左半分收音エリア 4 2 が用いられる。

【 0 0 4 4 】

デジタルカメラ 1 0 0 のマイク 1 6 1 といった撮像装置の收音部において、マイクロフォン素子の数および配置は、素子の搭載スペース等の事情により制約を受ける。例えば、ユーザが複数の被写体について音声を記録したい撮影場面において、マイクロフォン素子数の制約により收音指向性を十分に狭められない場合がある。こうした場合でも、本実施形態のデジタルカメラ 1 0 0 は、ユーザの撮影場面を想定して收音エリアを定義しておき、顔認識を用いて收音エリアを決定することで、ユーザの意図に沿った收音エリアを提供することができる。

【 0 0 4 5 】

[1 - 1 - 3 . マイク設定について]

デジタルカメラ 1 0 0 におけるマイク 1 6 1 の收音エリアに関する設定について、図 7 ~ 図 8 を用いて説明する。

【 0 0 4 6 】

図 7 は、デジタルカメラ 1 0 0 における設定メニューの表示例を示す。本実施形態のデジタルカメラ 1 0 0 は、マイク 1 6 1 の收音エリアを制御する動作モード (即ち收音モード) として、例えば図 7 に示すように、「オート」、「サラウンド」、「フロント」、「フォーカス」及び「ナレーション」というモードを有する。

【 0 0 4 7 】

フォーカスモードは、デジタルカメラ 1 0 0 による顔認識及び画角に連動して、自動的にマイク 1 6 1 の指向性を変え、收音エリアを調整する動作モードである。例えば、フォーカスモードは上述した各種收音エリア 4 1 ~ 4 4 を切り替えて用いることにより実現できる。フォーカスモードを 4 つの收音エリア 4 1 ~ 4 4 で敢えて大まかに実現することにより、被写体の僅かな動きによって頻繁に收音指向性が変化するような事態を回避し、ユーザの聴感上の煩わしさを低減することができる。デジタルカメラ 1 0 0 において更なる收音エリアを図 8 に例示する。

【 0 0 4 8 】

図 8 (A) は、サラウンドモードの收音エリア 4 5 を例示する。サラウンドモードは、デジタルカメラ 1 0 0 の左右前後に渡る広い範囲の音を收音するための動作モードである。サラウンドモードの收音エリア 4 5 は、例えば X Z 平面の全周 3 6 0 ° といった角度範囲を有する。

【 0 0 4 9 】

図 8 (B) は、フロントモードの收音エリア 4 6 を例示する。フロントモードは、デジタルカメラ 1 0 0 の前方の音を收音するための動作モードである。フロントモードの收音エリア 4 6 は、デジタルカメラ 1 0 0 から + Z 側に向いており、例えば上述した前方收音エリア 4 4 以上の角度範囲を有する。

【 0 0 5 0 】

図 8 (C) は、ナレーションモードの收音エリア 4 7 を例示する。ナレーションモードは、デジタルカメラ 1 0 0 の後方の音を收音するための動作モードである。ナレーションモードの收音エリア 4 7 は、デジタルカメラ 1 0 0 から - Z 側に向けて形成される。ナレーションモードが設定された状態で、表示モニタ 1 3 0 が自分撮り位置にあることが検出された場合、デジタルカメラ 1 0 0 は、フォーカスモードの動作を実行してもよい。

10

20

30

40

50

【 0 0 5 1 】

オートモードは、デジタルカメラ 1 0 0 の撮影状態に連動して、自動的にマイク 1 6 1 の指向性を変え、收音エリアを調整する動作モードである。デジタルカメラ 1 0 0 の撮影状態は、例えばフォーカスモードで考慮される顔認識等に加えて、自分撮りか否か、及び縦撮りか横撮りかなども含む。

【 0 0 5 2 】

以上のような各種の收音モードを設定するマイク設定は、例えばデジタルカメラ 1 0 0 の設定メニューにおける動画メニューの 1 つとして設けられる。ユーザは、タッチパネル 1 5 5 のタッチ操作あるいは各種ボタン 1 5 2 , 1 5 3 の押下操作により、設定メニューから所望の收音モードを選択可能である。又、マイク設定は、予め機能ボタン 1 5 4 等に割り当てられてもよい。オートモード等の特定の收音モードの設定が、機能ボタン 1 5 4 等に割り当てられてもよい。

【 0 0 5 3 】

〔 1 - 2 . 動作 〕

以上のように構成されるデジタルカメラ 1 0 0 の動作について説明する。以下では、デジタルカメラ 1 0 0 による動画撮影時の動作を説明する。

【 0 0 5 4 】

デジタルカメラ 1 0 0 は順次、光学系 1 1 0 を介して形成された被写体像をイメージセンサ 1 1 5 で撮像して撮像データを生成する。画像処理エンジン 1 2 0 は、イメージセンサ 1 1 5 により生成された撮像データに対して各種処理を施して画像データを生成し、バッファメモリ 1 2 5 に記録する。また、画像処理エンジン 1 2 0 の顔認識部 1 2 2 は、撮像データが示す画像に基づき、被写体の領域を検出して、例えば検出情報をコントローラ 1 3 5 に出力する。

【 0 0 5 5 】

本実施形態のデジタルカメラ 1 0 0 は、顔認識部 1 2 2 に入力された撮像画像において、画像認識処理により顔検出を行い、検出情報に基づいてオートフォーカス (A F) 制御の対象とする被写体を特定する動作モードである顔認識モードを備える。

【 0 0 5 6 】

以上の撮像動作と同時並行で、デジタルカメラ 1 0 0 は、マイク 1 6 1 において收音を行う。マイク用の A / D コンバータ 1 6 5 から收音結果の音声データを音声処理エンジン 1 7 0 にて処理する。音声処理エンジン 1 7 0 は、処理後の音声データ A o u t をバッファメモリ 1 2 5 に記録する。

【 0 0 5 7 】

コントローラ 1 3 5 は、バッファメモリ 1 2 5 を介して、画像処理エンジン 1 2 0 から受け付ける画像データと音声処理エンジン 1 7 0 から受け付ける音声データとの間で、同期を取って動画をメモリカード 1 4 2 に記録する。また、コントローラ 1 3 5 は逐次、表示モニタ 1 3 0 にスルー画像を表示させる。ユーザは、表示モニタ 1 3 0 のスルー画像により随時、撮影の構図等を確認することができる。動画撮影の動作は、操作部 1 5 0 におけるユーザの操作に応じて開始 / 終了される。

【 0 0 5 8 】

以上のようなデジタルカメラ 1 0 0 の動画撮影においては、様々な撮影状態においてユーザが收音したい種々の状況が考えられる。例えば撮影者とその同伴者など、仲間内で会話をしている被写体のグループに注目して行われる場合がある。この場合、音声についても、当該被写体のグループの発声を明瞭に收音したいとのニーズが考えられる。

【 0 0 5 9 】

本実施形態のデジタルカメラ 1 0 0 は、例えば上述したフォーカスモードの動作として、画像処理エンジン 1 2 0 における顔認識部 1 2 2 の検出情報によって被写体を検出し、A F 対象の被写体が決定されたときに、音声処理エンジン 1 7 0 において、当該被写体および撮影する空間で当該被写体の周囲にいる被写体について收音する音声を強調する処理を実行する。このように、画像処理エンジン 1 2 0 の顔認識と音声処理エンジン 1 7 0 の

10

20

30

40

50

音声強調等とを連動させて、上記のような会話をする被写体のグループによる音声を強調した収音を精度良く実現する。

【 0 0 6 0 】

さらに、本実施形態のデジタルカメラ 1 0 0 は、オートモードの動作として、上記のようなフォーカスモードの動作に加えて、さらに様々な撮影状態に応じて適切な収音制御を実現する。図 9 を用いて、オートモードの動作の概要を説明する。

【 0 0 6 1 】

図 9 は、デジタルカメラ 1 0 0 のオートモードと各種収音モードとの対応関係を例示する。オートモードのデジタルカメラ 1 0 0 は、例えば横撮りの状態で顔認識が為されたときは、フォーカスモードと同様の動作を行う。

10

【 0 0 6 2 】

一方、顔認識が為されていないときは、例えば自撮りでない状態（図 2 参照）であれば、デジタルカメラ 1 0 0 は、サラウンドモードと同様の動作を行う、すなわちサラウンドモードの収音エリア 4 5 を採用する。また、顔認識が為されず、自分撮りが行われる状態（図 3 参照）では、デジタルカメラ 1 0 0 は、フロントモードと同様の動作を行う。

【 0 0 6 3 】

また、縦撮りの場合、顔認識が為されていないときの動作は、上述した横撮りの場合と同様である。一方、縦撮りにおいて顔認識が為された場合、本実施形態のデジタルカメラ 1 0 0 は、フォーカスモードの代わりに、フロントモードと同様の動作を行う。

【 0 0 6 4 】

20

以上のようなオートモードの動作によると、図 9 に示すように、種々の撮影状態に応じて各種の収音モードの動作を組み合わせることにより、それぞれの撮影状態において適切な収音制御を実現し易くすることができる。

【 0 0 6 5 】

〔 1 - 2 - 1 . フォーカスモードの動作 〕

図 1 0 及び図 1 1 を用いて、本実施形態に係るデジタルカメラ 1 0 0 のフォーカスモードの動作の概要を説明する。

【 0 0 6 6 】

図 1 0 は、本実施形態に係るデジタルカメラ 1 0 0 のフォーカスモードの動作を例示するフローチャートである。図 1 0 のフローチャートに示す各処理は、例えばデジタルカメラ 1 0 0 がフォーカスモードに設定された状態において所定の周期で繰り返し実行される。所定の周期は、例えば動画のフレーム周期である。図 1 1 は、本実施形態に係るデジタルカメラ 1 0 0 のフォーカスモードの動作の概要を説明するための図である。

30

【 0 0 6 7 】

コントローラ 1 3 5 は、顔認識部 1 2 2 による検出情報に基づいて A F 対象を特定し、A F 制御を実行する（S 1）。A F 対象は、A F 制御の対象とする被写体の画像上の領域を示す。図 1 1（A）は、顔認識部 1 2 2 の検出情報において被写体が検出された領域を示す顔領域 R 1，R 2 及び R 3 を含む撮像画像 I m を例示する。顔領域 R 1，R 2 及び R 3 は、本実施形態における被写体領域の一例である。例えば顔領域 R 1 が A F 対象の顔領域 6 0 として特定される。

40

【 0 0 6 8 】

次にコントローラ 1 3 5 は、A F 対象として特定された顔領域が存在するか否かを判断する（S 2）。具体的には、コントローラ 1 3 5 は、顔領域が検出されていて、かつ A F 対象が顔領域であるか否かを判断する。

【 0 0 6 9 】

A F 対象の顔領域 6 0 がある場合（S 2 で Y E S）、コントローラ 1 3 5 は、検出情報における被写体からマイク 1 6 1 の収音対象を選別する処理を実行する（S 3）。収音対象は、マイク 1 6 1 により音声を強調して収音する対象とする被写体である。A F 対象として特定された顔領域 R 1（6 0）は、収音対象になる。図 1 1（B）は、図 1 1（A）に示す検出情報に基づき、顔領域 R 1 及び R 3 を収音対象に決定し、一方で顔領域 R 2 を

50

收音対象としない例を示す。

【 0 0 7 0 】

本実施形態のデジタルカメラ 1 0 0 は、收音対象の選別処理 (S 3) において、 A F 対象の顔領域 R 1 (6 0) に加え、撮像画像 I m において顔領域 R 1 と同程度の顔の大きさを示す顔 R 3 を、更なる收音対象として決定する。一方で、顔領域 R 1 と異なる大きさの顔領域 R 2 は收音対象から外される。これにより、人物 2 1 と人物 2 3 がデジタルカメラ 1 0 0 から同程度の距離にいて (すなわち、 Z 軸方向の距離の差が小さい)、人物 2 2 は異なる距離にいることを反映して、例えば仲間内で会話する被写体のグループを收音対象とすることができる。收音対象の選別処理 (S 3) についての詳細は後述する。

【 0 0 7 1 】

次に、コントローラ 1 3 5 は、決定した收音対象に基づいて、收音エリアを決定する処理を行う (S 4)。收音エリアの決定処理 (S 4) は、決定した全ての收音対象を含む收音エリアを決定する。図 1 1 (B) の例において、收音エリアを收音対象の顔領域 R 1 及び R 3 を含むように前方中心收音エリア 4 1 (図 6 (A)) に決定されている。收音エリアの決定処理 (S 4) についての詳細は後述する。

【 0 0 7 2 】

次に、コントローラ 1 3 5 は、決定した收音対象及び收音エリアに基づいて、顔認識を用いて收音の制御を行う (S 5)。顔認識を用いた收音制御 (S 5) は、コントローラ 1 3 5 が決定した收音対象、收音エリア及び收音ゲインを含む收音パラメータを、音声処理エンジン 1 7 0 に設定することによって行われる。音声処理エンジン 1 7 0 は、收音パラメータに応じた收音指向性及び收音ゲインを実現する。

【 0 0 7 3 】

一方、例えば顔認識モードの動作中に顔領域が検出されない等、 A F 対象の顔領域 6 0 がない場合 (S 2 で N O)、コントローラ 1 3 5 は、顔認識を用いない收音制御 (S 6) を行う。顔認識を用いた、または用いない收音制御 (S 5 , S 6) についての詳細は後述する。

【 0 0 7 4 】

コントローラ 1 3 5 は、ステップ S 5 または S 6 の收音制御を実行後、ステップ S 1 以降の処理を繰り返す。

【 0 0 7 5 】

以上の処理によると、本実施形態のデジタルカメラ 1 0 0 は、顔認識により検出した被写体から收音対象を選別し、收音対象を全て含む收音エリアを決定して、顔認識を用いた收音制御を行う。これにより、例えば仲間内で会話をする被写体のグループについて、音声を強調して收音することができる。

【 0 0 7 6 】

なお、顔認識による A F 制御 (S 1) において、検出情報に基づく A F 対象の特定は、例えば表示モニタ 1 3 0 に表示させたスルー画像上に顔領域を示す枠表示等を行い、操作部 1 5 0 によりユーザが枠表示を選択する操作を受けて実行することができる。

【 0 0 7 7 】

図 1 1 (C) は、図 1 1 (A) , (B) とは異なる位置に人物 2 1 ~ 2 3 がいる場合の撮像画像 I m の例を示す。デジタルカメラ 1 0 0 は、図 1 1 (B) の例と同様に、まず、例えば顔領域 R 1 を A F 対象の顔領域 6 0 として特定し (S 1)、收音対象に決定する。図 1 1 (C) の例において、收音対象の選別処理 (S 3) は、撮像画像 I m 上で顔領域 R 1 と同程度の顔の大きさである顔領域 R 2 を收音対象に決定し、顔領域 R 3 を收音対象から外す。收音エリアの決定処理 (S 4) は、收音対象として決定された顔領域 R 1 及び R 2 を含む左半分收音エリア 4 2 (図 6 (B)) を收音エリアに決定する。顔認識を用いた收音制御 (S 5) は、左半分收音エリア 4 2 に指向性を制御して人物 2 1 及び 2 2 の音声を明瞭に收音するように、收音パラメータを設定することによって行われる。

【 0 0 7 8 】

[1 - 2 - 2 . 收音対象の選別処理]

10

20

30

40

50

図 10 のステップ S 3 における收音対象の選別処理の詳細を、図 12 ~ 図 13 を用いて説明する。

【 0079 】

図 12 は、デジタルカメラ 100 の收音対象の選別処理 (S 3) を例示するフローチャートである。図 12 に示すフローチャートによる各処理は、図 10 のステップ S 11 で Y E S に進んだとき、例えばデジタルカメラ 100 のコントローラ 135 によって実行される。

【 0080 】

図 13 は、デジタルカメラ 100 における收音対象の選別処理 (S 3) を説明するための図である。以下では、図 11 (A) , (B) の例で收音対象を決定する動作について説明する。

10

【 0081 】

図 12 のフローチャートにおいて、コントローラ 135 は、図 10 のステップ S 1 において特定した A F 対象の顔領域に対応する被写体を收音対象に決定する (S 10)。このとき、コントローラ 135 は、顔認識部 122 から取得した検出情報に基づいて、A F 対象の顔領域の大きさ (即ち顔幅 W) を、他の被写体から收音対象を選別する基準に設定する。

【 0082 】

図 13 (A) は、図 11 (A) , (B) の例において收音対象が選別される場合を例示する。顔幅 W 1 , W 2 , W 3 は、撮像画像 I m における顔領域 R 1 , R 2 , R 3 の大きさを X 軸方向の幅で示す。図 13 (A) の例において、コントローラ 135 は、A F 対象の顔領域 R 1 の顔幅 W 1 を、基準の顔幅 W に設定する (S 10)。設定した顔幅 W は、例えばコントローラ 135 の R A M 等に保持される。

20

【 0083 】

次に、コントローラ 135 は、A F 対象の他に検出された被写体があるか否かを判断する (S 11)。具体的には、コントローラ 135 は、顔認識部 122 の検出情報が A F 対象の顔領域の他に顔領域を含むか否かを判断する。

【 0084 】

A F 対象の他に検出された被写体がある場合 (S 11 で Y E S)、コントローラ 135 は、收音対象の候補である收音候補として、一つの被写体 i を選択する (S 12)。図 13 (A) の例では、検出情報は A F 対象の顔領域 R 1 の他の顔領域 R 2 及び R 3 が、ステップ S 12 ごとに順次、收音候補の被写体 i に対応付けて選択される。

30

【 0085 】

コントローラ 135 は、選択した被写体 i の顔幅 W i と、基準の顔幅 W とを比較する演算を行う (S 13)。具体的には、コントローラ 135 は、基準の顔幅 W に対する被写体 i の顔幅 W i の割合 W_i / W を算出する。図 13 (A) の例において、顔領域 R 2 を收音候補とする選択時 (S 12) には、その顔幅 W 2 についての割合 W_2 / W が算出される (S 13)。

【 0086 】

コントローラ 135 は、收音候補の顔幅 W i と基準の顔幅 W 間の割合 W_i / W が、所定範囲内であるか否かを判断する (S 14)。所定範囲は、例えば收音候補の顔幅 W i が相対的に基準の顔幅 W i と同程度であるとみなす範囲を規定する観点から、「 1 」よりも大きい上限値と、「 1 」よりも小さい下限値で規定される。なお、所定範囲を設定するためのユーザインタフェースが提供されてもよく、例えばユーザが操作部 150 により設定した所定範囲がバッファメモリ 125 等に保持されてもよい。

40

【 0087 】

コントローラ 135 は、顔幅の割合 W_i / W が所定範囲内であると判断すると (S 14 で Y E S)、被写体 i を收音対象とすることを決定する (S 15)。

【 0088 】

一方、コントローラ 135 は、顔幅の割合 W_i / W が所定範囲内でないとは判断すると (

50

S 1 4でNO)、コントローラ135は、被写体*i*を收音対象としないことを決定する(S 1 6)。図13(A)の例において、割合 $W2/W$ は所定範囲の下限値を下回り、顔領域R2を收音対象としないことが決定される。

【0089】

コントローラ135は、被写体*i*を收音対象とするか否かを決定すると(S 1 5またはS 1 6)、例えば被写体*i*について決定した結果をバッファメモリ125に記録する(S 1 7)。次に、コントローラ135は、收音候補として選択済みの被写体とは他の被写体について、ステップS 1 1以降の処理を再び行う。

【0090】

図13(A)の例では、顔領域R2の他に顔領域R3が検出情報に含まれる(S 1 1でYES)。コントローラ135は、顔領域R3に対応する被写体を選択する(S 1 2)と、顔領域R2の場合と同様に、基準の顔幅Wに対する顔幅W3の割合 $W3/W$ を算出する(S 1 3)。図13(A)の例では、割合 $W3/W$ は「1」近傍に算出される。コントローラ135は、算出した顔幅の割合 $W3/W$ が收音対象の所定範囲内であると判断して(S 1 4でYES)、顔領域R3に対応する被写体を收音対象として決定する(S 1 5)。

【0091】

コントローラ135は、收音候補として選択されていない被写体なくなるまで(ステップS 1 1でNO)、ステップS 1 1~S 1 7の処理を繰り返す。その後、コントローラ135は、收音対象の選別処理(S 3)を終了して、図10のステップS 4に進む。

【0092】

以上の処理によると、顔認識により検出した被写体について、AF対象として特定した顔領域R1を基準とする相対的な顔領域R2, R3の大きさ比較が行われる。これにより、相対的な顔領域R3の大きさがAF対象の顔領域R1と同程度である被写体を選別して收音対象に決定することができる。

【0093】

図13(B)は、図11(C)の例において收音対象が選別される場合を例示する。図13(B)の例において、顔領域R1は、図13(A)の例と同様にAF対象として特定されている。このことから、コントローラ135は、顔領域R1を收音対象に決定し、顔幅W1を基準の顔幅Wに設定する(S 1 0)。

【0094】

図13(B)の例では、顔領域R2の顔幅W2は、基準の顔幅W(=W1)と同程度の大きさである。一方、顔領域R3の顔幅W3は、他の顔幅W1及びW2と比較して大きい。本例において、コントローラ135は、割合 $W2/W$ が所定範囲内であると判断して(S 1 4でYES)、顔領域R2の被写体を收音対象として決定する(S 1 5)。一方、割合 $W3/W$ が所定範囲の上限値を上回るため(S 1 4でNO)、顔領域R3の被写体は收音対象としないことが決定される(S 1 6)。よって、本例の收音対象は、顔領域R1及びR2に対応する2つの被写体に決定される(図11(C)参照)。

【0095】

図13(C)は、図11(C)と同様の撮影画像Imにおいて顔領域R3がAF対象の顔領域60として特定された(図10のS 1)場合を例示する。コントローラ135は、顔領域R3を收音対象に決定し、顔幅W3を基準の顔幅Wに設定する(S 1 0)。図13(C)の例において、割合 $W2/W$ 及び $W1/W$ が所定範囲の下限値を下回るため(S 1 4でNO)、顔領域R1及びR2に対応する被写体を收音対象としないことが決定される(S 1 6)。よって、本例の收音対象は、顔領域R3に対応する1つの被写体に決定される。

【0096】

以上のように、本実施形態のデジタルカメラ100は、画像認識により検出した複数の被写体から、AF対象と同程度の大きさである被写体を收音対象として決定することで、後述するユーザの意図に沿った收音エリアの決定に利用することができる。

【0097】

10

20

30

40

50

〔 1 - 2 - 3 . 收音エリアの決定処理 〕

図 10 のステップ S 4 における收音エリアの決定処理の詳細を、図 1 4 ~ 図 1 5 を用いて説明する。

【 0 0 9 8 】

図 1 4 は、本実施形態のデジタルカメラ 1 0 0 における收音エリアの決定処理 (S 4) を例示するフローチャートである。図 1 4 に示すフローチャートによる各処理は、図 1 0 のステップ S 3 を実行した後、例えばデジタルカメラ 1 0 0 のコントローラ 1 3 5 によって実行される。

【 0 0 9 9 】

図 1 5 は、デジタルカメラ 1 0 0 における收音エリアの決定処理 (S 4) を説明するための図である。図 1 5 (A) , (B) は、それぞれ図 1 3 (A) , (B) の例に続いて、收音エリアを決定する場合を例示する。図 1 5 (C) は、図 1 5 (A) , (B) とは更に別の場合を例示する。図 1 5 (A) ~ (C) において、中心位置 x_0 は、X 軸方向における撮像画像 I_m の中心の位置を示し、画像幅 W_h は、X 軸方向における撮像画像 I_m の幅を示す。画像範囲は、撮像画像 I_m 上で中心位置 x_0 を基準に、X 座標 - x_h から x_h の範囲 $x_0 \pm x_h$ と規定される。X 座標 x_h は、 $x_h = W_h / 2 (> 0)$ で定義される。

【 0 1 0 0 】

図 1 4 のフローチャートにおいて、コントローラ 1 3 5 は、全ての收音対象について、顔領域の中心等の位置が撮像画像 I_m の中心範囲にあるか否かを判断する (S 2 0) 。中心範囲は、撮像画像 I_m において前方中心收音エリア 4 1 に対応付けられる範囲である。

【 0 1 0 1 】

中心範囲は、例えば図 1 5 (A) に示すように、撮像画像 I_m 上で中心位置 x_0 を基準に、X 座標 - x_e から x_e の範囲 $x_0 \pm x_e$ として規定される。X 座標 x_e は、所定の画角 e と、画像幅 W_h に対応する水平画角 h とに基づき、例えば $x_e = x_h \times e / h (> 0)$ で定義される。所定の画角 e は、例えば 1 人の人物を含める観点から予め設定され、例えば 30° 等である。コントローラ 1 3 5 は、例えば光学系 1 1 0 のズームレンズのズーム倍率等から現在の水平画角 h を取得して、中心範囲 $x_0 \pm x_e$ を算出する。

【 0 1 0 2 】

水平画角 h が大きい広角撮影においては、X 座標 x_e が小さくなり中心範囲 $x_0 \pm x_e$ は狭い。一方、水平画角 h が小さい望遠撮影においては、X 座標 x_e が大きくなり中心範囲 $x_0 \pm x_e$ は広い。これにより、撮像する物理的な範囲と距離に対応した收音エリアの決定を実現しやすくすることができる。

【 0 1 0 3 】

全收音対象の顔領域の位置が中心範囲内にある場合 (S 2 0 で Y E S) 、コントローラ 1 3 5 は、收音エリアを前方中心收音エリア 4 1 に決定する (S 2 1) 。図 1 5 (A) の例では、收音対象は顔領域 R 1 及び R 3 に対応する。それぞれの顔領域 R 1 及び R 3 の中心の位置 x_1 及び x_3 は、いずれも $x_0 \pm x_e$ の範囲内にある (S 2 0 で Y E S) 。よって、收音エリアが前方中心收音エリア 4 1 に決定される (S 2 1 , 図 1 1 (B) 参照) 。

【 0 1 0 4 】

一方、少なくとも一つ以上の收音対象の顔領域の位置が中心範囲内でない場合 (S 2 0 で N O) 、前方中心收音エリア 4 1 以外の收音エリアが用いられる。この場合、コントローラ 1 3 5 は、全ての收音対象について、例えば顔領域の位置が撮像画像 I_m における左右いずれか半分の範囲のみにあるか否かを判断する (S 2 2) 。左半分の範囲は X 軸方向の中心位置 x_0 よりも X 座標が小さい範囲であり、右半分の範囲は中心位置 x_0 よりも X 座標が大きい範囲である。

【 0 1 0 5 】

全ての收音対象について、顔領域の位置が撮像画像 I_m における左半分または右半分の範囲のみにある場合 (S 2 2 で Y E S) 、コントローラ 1 3 5 は、さらに全收音対象の顔領域の位置が撮像画像 I_m における左半分の範囲内であるか否かを判断する (S 2 3) 。

【 0 1 0 6 】

全收音対象の顔領域の位置が撮像画像 I m における左半分の範囲内である場合 (S 2 3 で Y E S)、コントローラ 1 3 5 は、收音エリアを左半分收音エリア 4 2 に決定する (S 2 4)。図 1 5 (B) の例では、收音対象は顔領域 R 1 及び R 2 に対応する。顔領域 R 1 の位置 x 1 及び顔領域 R 2 の位置 x 2 が、X 軸方向の中心位置 x 0 より左 (すなわち、X 座標が小さい) 側にあるため (S 2 3 で Y E S)、收音エリアは左半分收音エリア 4 2 に決定される (S 2 4 , 図 1 1 (C) 参照)。

【 0 1 0 7 】

一方、全收音対象の顔領域の位置が撮像画像 I m における右半分の範囲内であって、左半分の範囲内でない (S 2 3 で N O)、コントローラ 1 3 5 は、收音エリアを右半分收音エリア 4 3 に決定する (S 2 5)。

10

【 0 1 0 8 】

また、全ての收音対象の顔領域の位置が、撮像画像 I m における左右いずれか半分の範囲のみにはない場合 (S 2 2 で N O)、コントローラ 1 3 5 は、收音エリアを前方收音エリア 4 4 に決定する (S 2 6)。図 6 (D) , (A) に示すように、前方收音エリア 4 4 は、前方中心收音エリア 4 1 の角度範囲 4 0 1 より広い角度範囲 4 0 2 を有する。すなわち、前方收音エリア 4 4 は、撮像画像 I m において X 軸方向に広い範囲に位置する收音対象の被写体を含む。

【 0 1 0 9 】

図 1 5 (C) の例では、收音対象は、顔領域 R 1 , R 2 及び R 3 に対応する。顔領域 R 1 ~ R 3 の中心の位置 x 1 , x 2 及び x 3 は、中心範囲 $x_0 \pm x_e$ 外の位置 x 1 及び x 2 を含み (S 2 0 で N O)、且つ、左半分の範囲内の位置 x 1 と右半分の範囲内の位置 x 2 及び x 3 とを含む (S 2 2 , S 2 3 で N O)。したがって、本例において收音エリアは、前方收音エリア 4 4 に決定される (S 2 6)。

20

【 0 1 1 0 】

コントローラ 1 3 5 は、收音エリアを決定すると (S 2 1 , S 2 4 ~ S 2 6)、決定した收音エリアをバッファメモリ 1 2 5 等に管理情報として記録する (S 2 7)。これにより、收音エリアの決定処理 (S 4) は終了し、図 1 0 のステップ S 5 に進む。

【 0 1 1 1 】

以上の処理によると、收音対象として決定した被写体の撮像画像上での位置に応じて、予め定義した複数の收音エリアから、全ての收音対象を含むように收音エリアが決定される。これにより、動画撮影において、ユーザの意図に沿った收音対象の被写体を含むように、收音エリアを決定することができる。

30

【 0 1 1 2 】

図 1 7 は、收音エリアの決定処理 (S 4) によって得られる管理情報を説明するための図である。図 1 7 (A) は、図 1 3 (A) 及び図 1 5 (A) の例において、收音対象の選別処理 (S 3) 及び收音エリアの決定処理 (S 4) を実行した段階で得られる管理情報を例示する。図 1 7 (B) は、図 1 3 (B) 及び図 1 5 (B) の例における管理情報を例示する。

【 0 1 1 3 】

管理情報は、例えば收音対象の選別処理 (S 3) によって決定される「收音対象」、收音エリアの決定処理 (S 4) によって決定される「收音エリア」、「水平画角」及び「合焦距離」を関連付けて管理する。なお、合焦距離は、例えば顔認識による A F 制御 (S 1) を実行する際に取得される。例えば、コントローラ 1 3 5 は、合焦時における光学系 1 1 0 の各種レンズの位置或いは焦点距離に基づいて、対応する合焦距離を取得してもよい。また、デジタルカメラ 1 0 0 は、D F D (D e p t h f r o m D e f o c u s) 技術または測距センサによる測定により、合焦距離を検出してもよい。

40

【 0 1 1 4 】

なお、本実施形態のデジタルカメラ 1 0 0 は、前方中心收音エリアの判断 (S 2 0) で用いる中心範囲の画角 θ_e を設定可能であり、例えばコントローラ 1 3 5 の R O M 等に記

50

録される。また、画角 θ を設定するためのユーザインタフェースが提供され、例えばユーザが操作部 150 により設定した値がバッファメモリ 125 等に保持されてもよい。

【0115】

〔1-2-4. 收音制御〕

(1) 図10のステップS5について

図10のステップS5における顔認識を用いた收音制御の詳細を、図16～図18を用いて説明する。

【0116】

收音パラメータ設定による收音制御において、本実施形態のデジタルカメラ100は、例えばAF対象の顔領域に対応する被写体について動画音声を強調するように、收音ゲインの設定を行う。收音ゲインは、例えば周波数フィルタ特性およびステレオセパレーション特性を有する。デジタルカメラ100は、例えばデジタルカメラ100が動画の撮影中に、AF対象の顔領域に合焦したときの水平画角及び合焦距離に基づき、收音ゲインを算出する。收音ゲインは、例えば算出される値が大きいほど人の声以外の周波数帯を抑制したりステレオ効果を制御したりして收音ズーム効果を生じさせるように規定される。

10

【0117】

図16は、顔認識を用いた收音制御(S5)を例示するフローチャートである。図16のフローチャートに示す各処理は、図10のステップS4を実行した後、例えばデジタルカメラ100のコントローラ135によって実行される。

【0118】

デジタルカメラ100は、図17に示す管理情報が保持された状態で、ステップS5の処理を開始する。

20

【0119】

コントローラ135は、例えばバッファメモリ125から水平画角を取得して、水平画角に基づくゲイン G_h を算出する(S30)。図18(A)は、水平画角からゲイン G_h を求める関係を例示する。図18(A)の例で、ゲイン G_h は、予め定めたゲインの最大値 G_{max} と最小値 G_{min} の間で、水平画角が小さくなるほど増加する。これにより、ズーム等で水平画角が小さいほど收音時にゲインを大きくして、望遠側で撮影される被写体の音声を強調することができる。

【0120】

コントローラ135は、ステップS30と同様に合焦距離を取得して、合焦距離に基づくゲイン G_d を算出する(S31)。図18(B)は、合焦距離からゲイン G_d を求める関係を例示する。図18(B)の例で、ゲイン G_d は、予め定めたゲインの最大値 G_{max} と最小値 G_{min} の間で、合焦距離が大きくなるほど増加する。これにより、デジタルカメラ100から遠い被写体に合焦するときほど收音時にゲインを大きくして、遠い被写体ほど音声を強調することができる。

30

【0121】

コントローラ135は、算出した水平画角による收音ゲイン G_h と、合焦距離による收音ゲイン G_d と比較し、いずれか大きいゲインを收音ゲイン G とする(S32)。これにより、例えば望遠の水平画角または遠い合焦距離で撮影を行うユーザの意図に沿って被写体の音声を強調するように、收音ゲイン G を算出することができる。

40

【0122】

コントローラ135は、過去の所定回数(例えば5回)にわたり算出された收音ゲイン G 及び決定された收音エリアが、互いに同じであるか否かを判断する(S33)。例えば收音ゲイン G は、図10のステップS1～S5の実行周期における所定回数の範囲内で、算出される毎に上記の管理情報と共に記憶される。コントローラ135は、過去の所定回数の收音ゲイン G および收音エリアが同じであると判断した場合と(S33でYES)、ステップS34に進む。

【0123】

コントローラ135は、ステップS3の收音対象の選別処理により決定した收音対象と

50

、ステップ S 4 の收音エリアの決定処理により決定した收音エリアと、ステップ S 3 2 で算出した收音ゲイン G を、音声処理エンジン 1 7 0 に收音パラメータとして設定する (S 3 4) 。音声処理エンジン 1 7 0 は、ビーム形成部 1 7 2 及びゲイン調整部 1 7 4 により、設定された收音パラメータに応じた收音エリア及び收音ゲインを実現する。

【 0 1 2 4 】

收音パラメータの設定 (S 3 4) 後、コントローラ 1 3 5 は、顔認識を用いた收音制御の処理 (S 5) を終了する。また、コントローラ 1 3 5 は、過去の所定回数の收音ゲイン G および收音エリアが同じでないと判断した場合 (S 3 3 で N O) 、ステップ S 3 4 の処理を行わずに図 1 0 のステップ S 5 の処理を終了する。その後、図 1 0 のステップ S 1 以降の処理が繰り返される。

10

【 0 1 2 5 】

以上の処理によると、算出した收音ゲインと、顔認識に基づいて決定された收音対象及び收音エリアを、收音パラメータに設定して、A F 対象を含む收音対象の被写体の音声を明瞭に收音しやすくする收音エリア及び收音ゲインを実現することができる。

【 0 1 2 6 】

なお、ステップ S 3 0 と S 3 1 の実行順序は、本フローチャートの順に限らず、例えばステップ S 3 1 でゲイン G d を算出してから、ステップ S 3 0 でゲイン G h を算出してもよく、またはステップ S 3 0 と S 3 1 を並列に実行してもよい。

【 0 1 2 7 】

また、以上のステップ S 3 3 によると、收音エリア及び收音ゲイン G が所定回数 (例えば 5 回) 変化しない場合のみ、收音パラメータが設定する処理 (S 3 4) が実行される。これにより、被写体の動きなどにより過度に頻繁に收音エリア及び收音ゲイン G が変更されることを防ぎ、顔認識を用いた收音制御 (S 5) をユーザの意図に沿って精度よく実現することができる。

20

【 0 1 2 8 】

(2) 図 1 0 のステップ S 6 について

図 1 0 のステップ S 6 における顔認識を用いない收音制御 (S 6) の詳細を、図 1 9 を用いて説明する。

【 0 1 2 9 】

図 1 9 は、顔認識を用いない收音制御 (S 6) を例示するフローチャートである。図 1 9 のフローチャートに示す各処理は、顔領域が検出されない等、図 1 0 のステップ S 2 において A F 対象の顔領域がない (S 2 で N O) 場合に、例えばデジタルカメラ 1 0 0 のコントローラ 1 3 5 によって実行される。

30

【 0 1 3 0 】

まず、コントローラ 1 3 5 は、收音エリアを、例えば前方收音エリア 4 4 に決定する (S 4 0) 。

【 0 1 3 1 】

次に、コントローラ 1 3 5 は、水平画角に基づくゲイン G h をステップ S 3 0 と同様に算出して、收音ゲイン G とする (S 4 1) 。さらに、コントローラ 1 3 5 は、ステップ S 3 3 と同様に、過去の所定回数にわたり算出された收音ゲイン G 及び決定された收音エリアが、互いに同じであるか否かを判断する (S 4 2) 。

40

【 0 1 3 2 】

コントローラ 1 3 5 は、過去の所定回数の收音ゲイン G および收音エリアが同じであると判断した場合 (S 4 2 で Y E S) 、收音エリアと收音ゲイン G を收音パラメータに設定し (S 4 3) 、顔認識を用いない收音制御 (S 6) を終了する。また、コントローラ 1 3 5 は、過去の所定回数の收音ゲイン G および收音エリアが同じでないと判断した場合 (S 4 2 で N O) 、ステップ S 4 3 の処理を行わずに図 1 0 のステップ S 6 を終了する。ステップ S 6 の終了後、ステップ S 1 以降の処理が繰り返される。

【 0 1 3 3 】

以上の処理によると、A F 対象の顔領域がない場合でも、デジタルカメラ 1 0 0 の前方

50

における広い範囲の音声を收音するように、また、ズーム等で水平画角が小さいほど收音ゲインを大きくするようにして、撮像される範囲の音声を明瞭に收音しやすくすることができる。

【 0 1 3 4 】

なお、デジタルカメラ 1 0 0 の動作モードに応じて、デジタルカメラ 1 0 0 の周囲 3 6 0 ° の角度範囲を有する全体收音エリアが定義され、ステップ S 4 0 において全体收音エリアに決定されてもよい。このとき、例えば全体收音エリアのみが收音パラメータに設定されてもよい。

【 0 1 3 5 】

〔 1 - 2 - 5 . オートモードの動作 〕

本実施形態に係るデジタルカメラ 1 0 0 のオートモードの動作の詳細を、図 2 0 ~ 図 2 1 を用いて説明する。

【 0 1 3 6 】

図 2 0 は、実施の形態 1 に係るデジタルカメラ 1 0 0 のオートモードの動作を例示するフローチャートである。図 2 0 のフローチャートに示す各処理は、例えば図 1 0 と同様にコントローラ 1 3 5 により、デジタルカメラ 1 0 0 がオートモードに設定された状態において実行される。

【 0 1 3 7 】

図 2 0 に示すように、オートモードのデジタルカメラ 1 0 0 において、コントローラ 1 3 5 は、例えば磁気センサ 1 3 2 の検出信号に基づいて、表示モニタ 1 3 0 が自分撮り位置であるか否かを判断する (S 5 1)。コントローラ 1 3 5 は、表示モニタ 1 3 0 が自分撮り位置でないと判断すると (S 5 1 で N O)、非顔認識時の收音エリアをサラウンドモードの收音エリア 4 5 に設定する (S 5 2)。一方、表示モニタ 1 3 0 が自分撮り位置であると判断すると (S 5 1 で Y E S)、コントローラ 1 3 5 は、非顔認識時の收音エリアをフロントモードの收音エリア 4 6 に設定する (S 5 3)。

【 0 1 3 8 】

コントローラ 1 3 5 は、上述したフォーカスモードと同様に、顔認識の処理を行う (S 1 , S 2)。例えば、A F 対象の顔領域が検出された場合 (S 2 で Y E S)、コントローラ 1 3 5 は、加速度センサ 1 3 7 の検出信号に基づいて、デジタルカメラ 1 0 0 が縦撮りの姿勢か否かを判断する (S 5 4)。コントローラ 1 3 5 は、縦撮りでないと判断すると (S 5 4 で N O)、フォーカスモードと同様のステップ S 3 ~ S 5 の処理を行い、收音制御を実行する。一方、コントローラ 1 3 5 は、縦撮りと判断すると (S 5 4 で Y E S)、フロントモードの收音エリア 4 6 を採用して、收音制御を実行する (S 5 5)。

【 0 1 3 9 】

また、A F 対象の顔領域が検出されなかった場合 (S 2 で N O)、コントローラ 1 3 5 は、ステップ S 5 2 , S 5 3 の設定結果に基づいて、顔認識を用いない收音制御を行う (S 6 A)。ステップ S 6 A の收音制御は、非顔認識時の收音エリアとして設定された收音エリアを用いて、上述したステップ S 6 と同様に行われる。

【 0 1 4 0 】

以上の処理によると、各種の撮影状態に連動してマイク 1 6 1 の指向性を調整するオートモードの動作を実現できる。オートモードにおける横撮り時と縦撮り時の收音制御について、図 2 1 を用いて更に説明する。

【 0 1 4 1 】

図 2 1 (A) は、横撮り時の撮像画像 I m と收音エリア 4 1 ~ 4 3 との関係を例示する。図 2 1 (B) は、縦撮り時にステップ S 5 5 を行わなかった場合を例示する。図 2 1 (C) は、縦撮り時にステップ S 5 5 を行った場合を例示する。

【 0 1 4 2 】

図 2 1 (A) に例示するように、横撮り時には、收音エリアと顔領域 R 1 ~ R 3 と收音エリア 4 1 ~ 4 3 とは同軸上の関係にあり、收音エリア 4 1 ~ 4 3 の切り替えによって所望の顔領域 R 1 ~ R 3 の位置に合わせるような意図通りの收音制御が可能である。

10

20

30

40

50

しかしながら、縦撮り時においては、図 2 1 (B) に示すように、收音エリア 4 1 ~ 4 3 と顔領域 R 1 ~ R 3 との関係が合わなくなる。このことから、所望の顔領域 R 1 ~ R 3 の位置に対して意図通りの收音制御ができず、寧ろ意図に反した收音制御が生じてしまう事態が考えられる。

【 0 1 4 3 】

そこで、本実施形態では、縦撮り時には図 2 1 (C) に示すように、收音エリアをフロントモードの收音エリア 4 6 に固定する。これにより、撮像中の範囲にわたり收音可能な範囲を確保して、意図に反した收音制御が生じるような事態を回避できる。

【 0 1 4 4 】

〔 1 - 3 . 効果等 〕

本実施形態において、デジタルカメラ 1 0 0 は、撮像部の一例のイメージセンサ 1 1 5 と、音声取得部の一例のマイク 1 6 1 と、設定部の一例の操作部 1 5 0 と、制御部の一例のコントローラ 1 3 5 とを備える。イメージセンサ 1 1 5 は、被写体を撮像して画像データを生成する。マイク 1 6 1 は、撮像部による撮像中に收音される音声を示す音声信号を取得する。操作部 1 5 0 は、ユーザの指示を受けて、音声取得部の指向性を自動で変える動作モードであるオートモードに自装置を設定する。コントローラ 1 3 5 は、音声信号において被写体からの音声を收音する收音エリアを制御する。コントローラ 1 3 5 は、オートモードに設定されている場合、マイク 1 6 1 の指向性を自装置の撮影状態に連動して変えることで、被写体を含めるように收音エリアを制御する。これにより、様々な撮影状態に応じて適切な收音制御が実現でき、音声を取得しながら撮像を行う際に、ユーザの意図に沿って被写体の音声を收音し易くすることができる。

【 0 1 4 5 】

本実施形態のデジタルカメラ 1 0 0 は、画像データにおいて被写体の顔領域を検出する顔検出部の一例の顔認識部 1 2 2 を備える。コントローラ 1 3 5 は、オートモードに設定されている場合、顔認識部 1 2 2 によって検出された顔領域に基づいて音声信号における收音対象とする被写体を決定し、收音対象に決定した被写体を含めるように、收音エリアを制御する。これにより、被写体の各種顔認識のような撮影状態に応じて收音制御を行い、ユーザの意図に沿った收音を行い易くできる。

【 0 1 4 6 】

本実施形態において、コントローラ 1 3 5 は、オートモードに設定されている場合、自装置が縦向きか又は横向きかで撮影する撮影状態に連動して、音声取得部の指向性を変えるように、收音エリアを制御する。これにより、縦撮りか横撮りかといった撮影状態に応じて收音制御を行い、ユーザの意図に沿った收音を行い易くできる。

【 0 1 4 7 】

本実施形態において、コントローラ 1 3 5 は、オートモードに設定されている場合、撮影者が自身を撮影するか否かの撮影状態に連動して、音声取得部の指向性を変えるように、收音エリアを制御する。これにより、自分撮りか否かといった撮影状態に応じて收音制御を行い、ユーザの意図に沿った收音を行い易くできる。

【 0 1 4 8 】

本実施形態のデジタルカメラ 1 0 0 は、表示部の一例の表示モニタ 1 3 0 と、検知部の一例の磁気センサ 1 3 2 とをさらに備える。表示モニタ 1 3 0 は、被写体の画像等を表示する表示面を有し、当該表示面を被写体側に変位可能に構成される。磁気センサ 1 3 2 は、表示モニタ 1 3 0 が被写体側に表示面を変位したか否かを検知する。本実施形態の設定部は、磁気センサ 1 3 2 が、表示モニタ 1 3 0 が被写体側に表示面を変位したことを検知すると、オートモードに設定するものであってもよい。例えば、コントローラ 1 3 5 が磁気センサ 1 3 2 からの検出信号に応じて、表示モニタ 1 3 0 が自分撮り位置にあるときに自動的にデジタルカメラ 1 0 0 オートモードに設定してもよい。

【 0 1 4 9 】

本実施形態のデジタルカメラ 1 0 0 において、設定部は、ユーザの指示に応じて、オートモードのほかに音声取得部の指向性が互いに異なる複数の動作モードうちの少なくとも

10

20

30

40

50

１つに設定可能である。例えば、サラウンドモード、フロントモード、或いはナビゲーションモードに設定可能であり、更にフォーカスモードに設定可能であってもよい。

【０１５０】

本実施形態のデジタルカメラ１００において、コントローラ１３５は、設定部がオートモードに設定している場合にイメージセンサ１１５によって撮像が開始されると、表示モニタ１３０に被写体と共にオートモードに設定されていることを示す情報を表示させてもよい。例えば、コントローラ１３５は、表示モニタ１３０にオートモード専用のアイコン等を表示させてもよい。

【０１５１】

（実施の形態２）

以下、図面を用いて実施の形態２を説明する。実施の形態１では、動画撮影時等に收音対象を選別して決定するデジタルカメラ１００について説明した。実施の形態２では、実施の形態１のような動作時に、決定された收音対象に関する情報をユーザに可視化するデジタルカメラ１００について説明する。

【０１５２】

以下、実施の形態１に係るデジタルカメラ１００と同様の構成、動作の説明は適宜省略して、本実施形態に係るデジタルカメラ１００を説明する。

【０１５３】

〔２－１．概要〕

図２２を用いて、本実施形態に係るデジタルカメラ１００が各種情報を表示する動作の概要を説明する。

【０１５４】

図２２は、本実施形態に係るデジタルカメラ１００の表示例を示す。図２２の表示例は、デジタルカメラ１００が図１１（Ｂ）に例示するように收音対象を決定した場合に、表示部１３０においてリアルタイムに表示される一例を示す。本表示例において、デジタルカメラ１００は表示モニタ１３０に、ＡＦ対象の被写体を示すＡＦ枠１１及びＡＦ対象以外の検出された被写体を示す検出枠１３に加えて、收音対象の被写体を示す收音アイコン１２を、撮像画像Ｉｍ上に重畳して表示している。

【０１５５】

本実施形態のデジタルカメラ１００は、收音アイコン１２を、ＡＦ枠１１及び検出枠１３に組み合わせて用いることで、ＡＦ対象のような主要被写体とそれ以外に検出した被写体が、ＡＦ対象及び／または收音対象に決定されているか否かをユーザに可視化する。

【０１５６】

例えば図２２の表示例において、デジタルカメラ１００は、図１１（Ｂ）の例で顔領域Ｒ１（６０）に対応する被写体をＡＦ対象かつ收音対象として決定したことから、人物２１にＡＦ枠１１と收音アイコン１２を表示する。また、デジタルカメラ１００は、図１１（Ｂ）の例で顔領域Ｒ３に対応する被写体をＡＦ対象以外の收音対象として決定したことから、人物２３に検出枠１３と收音アイコン１２を表示する。さらに、デジタルカメラ１００は、收音アイコン１２を伴わない検出枠１３の表示により、図１１（Ｂ）の例で顔領域Ｒ２に対応するＡＦ対象以外の被写体を收音対象としない決定をしたことをユーザに可視化する。

【０１５７】

本実施形態のデジタルカメラ１００において、ユーザは、ＡＦ枠１１または検出枠１３のいずれか一方の表示により、検出した被写体がＡＦ対象か否かを確認できる。ユーザはまた、收音アイコン１２の有無により、收音対象か否かを確認できる。ＡＦ枠１１と收音アイコン１２との組み合わせは、本実施形態における第１の識別情報の一例である。検出枠１３と收音アイコン１２との組み合わせは、本実施形態における第２の識別情報の一例である。検出枠１３は、第３の識別情報の一例である。

【０１５８】

以上のように、本実施形態に係るデジタルカメラ１００は、検出情報が含む被写体から

10

20

30

40

50

決定した收音対象及び A F 対象の被写体を区別する表示を行う。これによりユーザは、デジタルカメラ 100 が検出した被写体のうち收音対象の被写体を把握して、例えば意図に沿った被写体が收音対象として決定されているか否かを確認することができる。

【0159】

図 23 (A), (B) は、本実施形態のデジタルカメラ 100 におけるマニュアル操作を説明するための図である。図 23 (A) は、特定の人物 22 が、顔認識部 122 により検出されていない状態を例示する。例えば、撮影者が人物 22 の音声を收音したくても、人物 22 の顔がデジタルカメラ 100 に対して横向き又は後ろ向きであることにより、顔認識が為されない場合が想定される。また、撮影者が收音したい音源が、人物でないような場合も想定される。こうした場合を考慮して、本実施形態のデジタルカメラ 100 は、收音エリアを手動で設定可能にするマニュアル操作を入力可能に動作する。

10

【0160】

図 23 (B) は、デジタルカメラ 100 における收音エリアのマニュアル操作を例示する。例えば、收音エリアのマニュアル操作は、タッチ操作として実装される。図 24 は、デジタルカメラ 100 におけるマニュアル操作時の動作を例示するフローチャートである。本フローチャートに示す処理は、例えば上述したオートモード又はフォーカスモードの処理とは独立に実行されてもよいし、オートモード又はフォーカスモードの処理時に割り込んで実行されてもよい。

【0161】

まず、デジタルカメラ 100 のコントローラ 135 は、撮影者などのユーザによるマニュアル操作を受け付ける (S61)。コントローラ 135 は、例えば図 23 (B) に示すように、マニュアル操作時において、表示モニタ 130 に收音エリアの指定範囲 48 を表示する。この例はマニュアル操作がオートモード又はフォーカスモードの処理時に割り込んで実施される場合である。

20

【0162】

図 23 (B) の例では、撮影者がタッチ操作により、收音エリアの指定範囲 48 が、顔認識されなかった人物 22 を含むように調整するマニュアル操作を入力している。コントローラ 135 は、入力されたマニュアル操作に基づき、收音エリアの指定範囲 48 を決定する (S62)。マニュアル操作として收音範囲とする始点と終点をタッチ操作で入力すれば、始点と終点を含む所定サイズの收音範囲 48 が表示される。この状態で表示モニタ 130 に表示される確定ボタンをタッチ操作すれば、收音範囲 48 が確定される。

30

【0163】

コントローラ 135 は、決定した收音エリアの指定範囲 48 をマイク 161 の收音制御に反映する (S63)。これにより、指定範囲 48 に対応する收音エリアからの音声を強調するように、收音制御が実行される。

【0164】

〔2-3. 効果等〕

以上のように、本実施形態のデジタルカメラ 100 は、被写体を撮像して画像データを生成するイメージセンサ 115 と、イメージセンサ 115 による撮像中に收音される音声を示す音声信号を取得するマイク 161 と、被写体の画像を表示する表示モニタ 130 とを備える。本実施形態のデジタルカメラ 100 は、表示モニタ 130 に表示された被写体を、被写体からの音声を收音する收音エリアに設定するユーザ操作を入力する操作部 150 等の入力部と、音声信号において收音エリアを制御するコントローラ 135 とを備える。コントローラ 135 は、收音エリアを設定するユーザ操作が入力されると、ユーザ操作に基づいて、マイク 161 の指向性を変えることで被写体を含めるように、收音エリアを制御する。こうしたマニュアル操作により、收音エリアを制御でき、ユーザの意図に沿った收音を行い易くすることができる。

40

【0165】

本実施形態のデジタルカメラ 100 において、表示モニタ 130 は、実施形態 1 と同様に、被写体側に表示面を変位可能に構成されてもよい。表示モニタ 130 が、被写体側に

50

表示面が変位された状態において、ユーザ操作がなされてもよい。即ち、デジタルカメラ 100 において自分撮りを行う際に、上述したマニュアル操作が入力されてもよい。

【0166】

(他の実施の形態)

以上のように、本出願において開示する技術の例示として、上記の各実施の形態を説明した。しかしながら、本開示における技術は、これに限定されず、適宜、変更、置き換え、付加、省略などを行った実施の形態にも適用可能である。また、上記の各実施の形態で説明した各構成要素を組み合わせ、新たな実施の形態とすることも可能である。

【0167】

実施の形態 1 では、マイク 161 に 3 つのマイクロフォン素子 161L, 161C 及び 161R を用いる例を説明した。4 つのマイクロフォン素子を用いる変形例を、図 25 ~ 図 26 を用いて説明する。

【0168】

図 25 は、本変形例のデジタルカメラ 100A におけるマイク 161A の配置を例示する。本変形例において、デジタルカメラ 100A のマイク 161A は、互いに XZ 平面上にある 3 つのマイクロフォン素子 161L, 161C 及び 161R に加えて、4 つ目のマイクロフォン素子 161B を含む。4 つ目のマイクロフォン素子 161B は、Y 方向における位置が他のマイクロフォン素子 161L ~ 161R とは異なるように配置される。

【0169】

図 26 は、本変形例において縦撮り時の撮像画像 Im と收音エリア 41A ~ 43A との関係性を例示する。上記のマイク 161A の構成によると、デジタルカメラ 100 の Y 軸方向においても收音エリアを変えることができる。このため、縦撮り時の收音制御時には、4 つ目のマイクロフォン素子 161B を用いた收音エリアを利用することにより、例えば図 26 に示すように、縦撮り時にも顔認識に追従するような收音制御を実現できる。例えば、本変形例のデジタルカメラ 100A のコントローラ 135 は、図 20 と同様の処理において、ステップ S55 の代わりに、上記のような 4 つ目のマイクロフォン素子 161B を用いた收音制御を行う。これにより、縦撮り時においても、收音エリア 41A ~ 43A を用いて所望の顔領域 R1 ~ R3 の位置に合わせるような意図通りの收音制御が可能となる。

【0170】

また、上記各実施形態において説明した收音制御においては、顔認識と連動して收音エリアを遷移させる期間に緩急を設けることにより、聴感上の違和感をより抑制することができる。この動作例について、図 27 を用いて説明する。

【0171】

図 27 では、デジタルカメラ 100 において顔認識の有無に連動して收音指向性の幅 (即ち收音エリアの角度範囲) を変化させる動作例を示す。本動作例では、時刻 t1 においてデジタルカメラ 100 の顔認識部 122 により被写体の顔が検出されている。すると、誤検出防止制御 (即ちチャタリング) が行われる (図 16 の S33 参照)。

【0172】

例えば顔認識が一瞬だけ為されたり、顔認識した被写体が横を向いてしまったりしたときに收音指向性を変化させると、收音結果を聴くユーザが聴感上の違和感を覚える事態が考えられる。これに対して、上述した誤検出防止制御によると、被写体の顔認識の位置を常に監視し、一定時間收音エリアにいるときに收音指向性を変化させる。これにより、上記のような聴感上の違和感を回避可能になる。

【0173】

また、デジタルカメラ 100 のコントローラ 135 は、時刻 t1 からチャタリング後に、收音指向性を狭めるように收音エリアを遷移させる。收音指向性を狭める際の遷移期間は、例えば比較的短く設定される。これにより、画像認識において被写体の顔が検出された際に、検出された顔に向けて收音指向性がフォーカスされるように速やかな変化を与え、收音結果を聴くユーザに聴感上、良好な印象を与えることができる。また、例えば、前

10

20

30

40

50

方中心收音エリア 4 1 と左半分收音エリア 4 2 など同じ角度範囲の收音エリアの遷移についても、上記と同様に比較的速やかに行われる。

【 0 1 7 4 】

図 2 7 の例においては、例えば被写体が移動したり顔を横向きにしたりする等により、時刻 t_2 において、顔認識部 1 2 2 により被写体の顔が検出されなくなっている。この際も、チャタリング制御の後に、デジタルカメラ 1 0 0 は收音指向性を遷移させる。ここで、收音指向性の範囲を広げるときの遷移期間は、上述した狭めるときの遷移期間よりも長く設定される。これにより、広い範囲からの音が聴こえ出す状況が突然に起こることで生じ得るような聴感上の違和感をユーザに与えることを回避することができる。寧ろ、收音指向性の範囲を広げるときの遷移を遅く実行することにより、ユーザの聴感上の違和感を抑制することができる。

10

【 0 1 7 5 】

さらに、本例では時刻 t_3 において、收音指向性を広げる遷移期間中に、顔認識が再び為されている。この場合、デジタルカメラ 1 0 0 は、割り込み制御として收音指向性を広げきる前に、再び狭める制御に切り替わる。これにより、被写体の顔認識が断続的で有る場合に、顔認識が為された被写体に対して收音指向性を向ける制御が速やかに行われ、ユーザの聴感上の違和感をさらに抑制することができる。

【 0 1 7 6 】

上記の各実施形態においては、收音対象の検出に顔認識部 1 2 2 を用いた。本実施形態において、收音対象の検出は顔認識部 1 2 2 に限らず、例えばこれに代えて、又は加えて、人間の全体又は少なくとも一部を画像認識する人体認識を用いてもよい。また、收音対象は、必ずしも人物でなくてもよく、例えば各種の動物であってもよい。この場合、動物の一部又は全体の画像認識により、收音対象の検出が行われてもよい。

20

【 0 1 7 7 】

また、実施の形態 2 に係る第 1 の識別情報、第 2 の識別情報、及び第 3 の識別情報は、A F 枠 1 1 の有無において主要被写体か否かを識別し、收音アイコン 1 2 の有無において收音対象か否かを識別した。本実施形態において、第 1 ~ 第 3 の識別情報は、特にこれに限らず、例えば 3 種類の枠表示であってもよい。図 2 2 は、本実施形態における 3 種類の枠表示を例示する。図 2 2 の例では、A F 対象かつ收音対象の被写体を示す枠表示 1 1 A、A F 対象以外で收音対象の被写体を示す枠表示 1 3 A 及び收音対象でない被写体を示す枠表示 1 3 B により、A F 対象及びそれ以外の被写体の表示と收音対象の表示とが一体的に為されている。

30

【 0 1 7 8 】

上記の実施形態 1 , 2 では、オートモードの動作とマニュアル操作とをそれぞれ説明したが、これらは組み合わせてもよい。すなわち、本実施形態のデジタルカメラ 1 0 0 は、被写体の画像を表示する表示部としての表示モニタ 1 3 5 と、表示モニタ 1 3 5 に表示された被写体を、被写体からの音声を收音する收音エリアに設定するユーザ操作を入力する入力部としての操作部 1 5 0 とを備えてもよい。コントローラ 1 3 5 は、收音エリアを設定するユーザ操作が入力されると、ユーザ操作に基づいて、マイク 1 6 1 の指向性を変えることで被写体を含めるように、收音エリアを制御してもよい。こうした場合、表示モニタ 1 3 5 は特に可動式でなくてもよく、例えば上述した通常位置などに固定される固定式であってもよい。

40

【 0 1 7 9 】

実施の形態 1 ~ 2 では、図 1 0 のフローチャートにおいて、デジタルカメラ 1 0 0 が内蔵するマイク 1 6 1 について、顔認識を用いた又は用いない收音制御 (S 5 又は S 6) を行う動作例を説明した。本実施形態のデジタルカメラ 1 0 0 は、内蔵のマイク 1 6 1 に代えて、外付けのマイク (以下「マイク 1 6 1 a という」) を備えてもよい。マイク 1 6 1 a は、デジタルカメラ 1 0 0 の外部にあるマイクロフォン素子を含み、3 つ以上のマイクロフォン素子を備える。本実施形態においてコントローラ 1 3 5 は、マイク 1 6 1 a について、予めマイクロフォン素子の配置に関する情報をバッファメモリ 1 2 5 等に保持して

50

おくことにより、実施の形態１と同様にステップＳ５又はＳ６を実行することができる。この場合においても、実施の形態１と同様に決定した收音対象及び／または收音エリアに応じて、被写体の音声を明瞭に得やすくすることができる。

【０１８０】

また、実施の形態１～２では、図１６のフローチャートにおいて、デジタルカメラ１００の撮像範囲に対応する水平画角に基づき、ゲインＧ_hを算出（Ｓ３０）する動作例を説明した。この場合の水平画角は、図１４のフローチャートにおける前方中心收音エリアの判定（Ｓ２０）に用いる水平画角 h と同一である。本実施形態において、ゲインＧ_hの算出に、ステップＳ２０における水平画角 h と異なる用いる水平画角を用いてもよい。例えば、撮像画像上で全ての收音対象の被写体を含むＸ軸方向の幅に対応する角度範囲を、ステップＳ３０における水平画角とする。これにより、收音対象が映る画角に応じて、遠くの被写体の声をより明瞭に收音するように、ゲインＧ_hを算出することができる。

【０１８１】

また、実施の形態１～２では、顔認識部１２２が人の顔を検出した。本実施形態において、顔認識部１２２は、例えば動物の顔を検出してもよい。動物の顔は、動物の種別によって大きさが多様であることが考えられる。この場合でも、例えば收音対象を選別するための所定範囲（Ｓ１４参照）を拡大することにより、実施の形態１と同様に收音対象を選別することができる。さらに、顔認識部１２２が動物の種別ごとに顔を検出し、種別に応じてステップＳ１４における所定範囲を設定してもよい。

【０１８２】

また、実施の形態１～２では、顔認識部１２２を備えるデジタルカメラ１００を説明した。本実施形態において、顔認識部１２２は、外部サーバに設けられてもよい。この場合、デジタルカメラ１００は、通信モジュール１６０を介して、外部サーバに撮像画像の画像データを送信し、外部サーバから顔認識部１２２による処理結果の検出情報を受信してもよい。このようなデジタルカメラ１００においては、通信モジュール１６０が検出部として機能する。

【０１８３】

また、実施の形態１～２では、光学系１１０及びレンズ駆動部１１２を備えるデジタルカメラ１００を例示した。本実施形態の撮像装置は、光学系１１０及びレンズ駆動部１１２を備えなくてもよく、例えば交換レンズ式のカメラであってもよい。

【０１８４】

また、実施の形態１～２では、撮像装置の例としてデジタルカメラを説明したが、これに限定されない。本開示の撮像装置は、画像撮影機能を有する電子機器（例えば、ビデオカメラ、スマートフォン、タブレット端末等）であればよい。

【０１８５】

以上のように、本開示における技術の例示として、実施の形態を説明した。そのために、添付図面および詳細な説明を提供した。

【０１８６】

したがって、添付図面および詳細な説明に記載された構成要素の中には、課題解決のために必須な構成要素だけでなく、上記技術を例示するために、課題解決のためには必須でない構成要素も含まれ得る。そのため、それらの必須ではない構成要素が添付図面や詳細な説明に記載されていることをもって、直ちに、それらの必須ではない構成要素が必須であるとの認定をするべきではない。

【０１８７】

また、上述の実施の形態は、本開示における技術を例示するためのものであるから、特許請求の範囲またはその均等の範囲において種々の変更、置き換え、付加、省略などを行うことができる。

【産業上の利用可能性】

【０１８８】

本開示は、音声を取得しながら撮像を行う撮像装置に適用可能である。

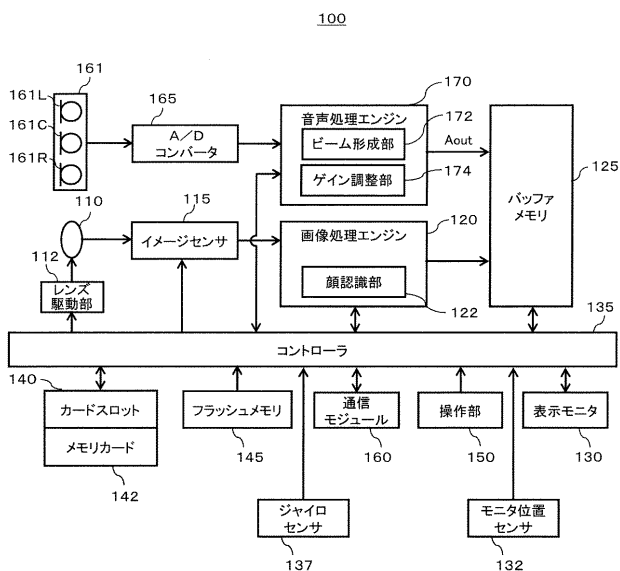
【符号の説明】

【 0 1 8 9 】

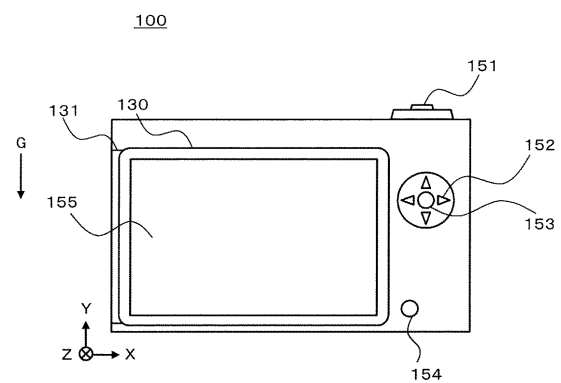
- 1 0 0 デジタルカメラ
- 1 1 5 イメージセンサ
- 1 2 0 画像処理エンジン
- 1 2 2 顔認識部
- 1 2 5 バッファメモリ
- 1 3 0 表示モニタ
- 1 3 2 磁気センサ
- 1 3 5 コントローラ
- 1 3 7 加速度センサ
- 1 4 5 フラッシュメモリ
- 1 5 0 操作部
- 1 1 A F 枠
- 1 2 収音アイコン
- 1 3 検出枠

10

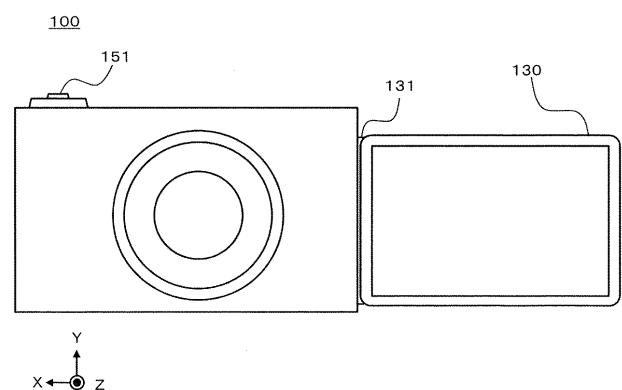
【 図 1 】



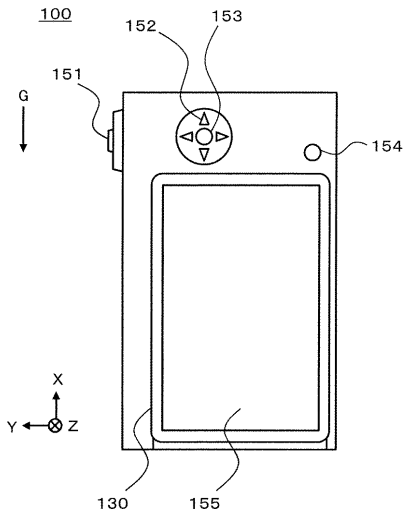
【 図 2 】



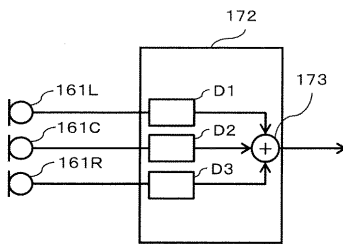
【 図 3 】



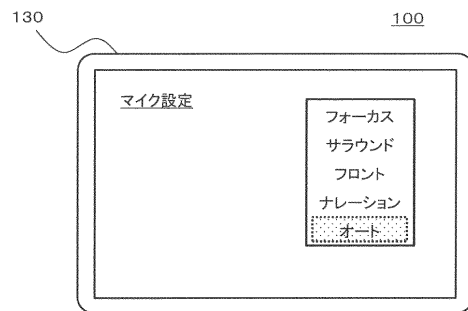
【図 4】



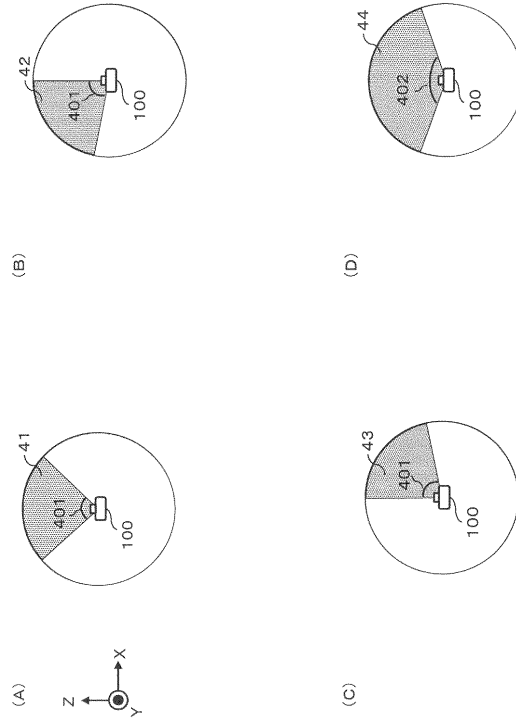
【図 5】



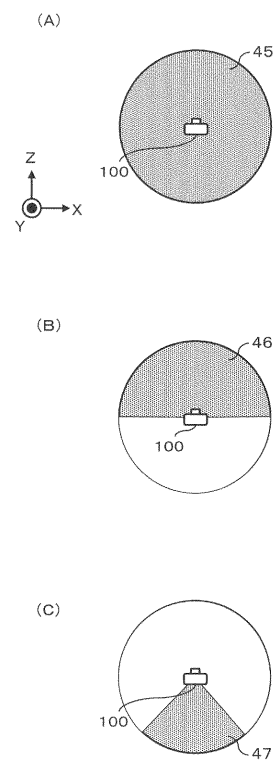
【図 7】



【図 6】



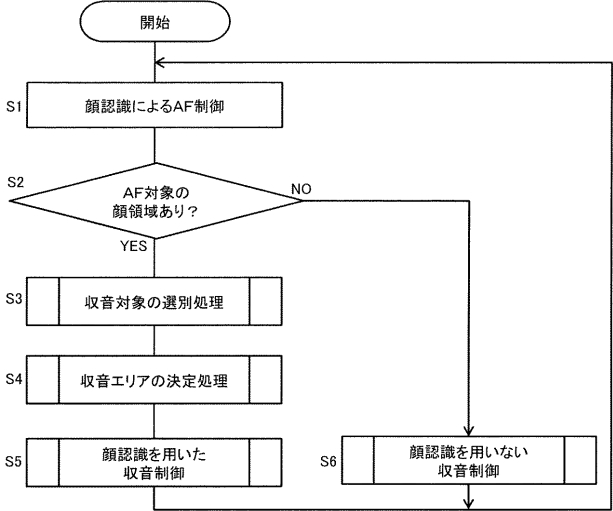
【図 8】



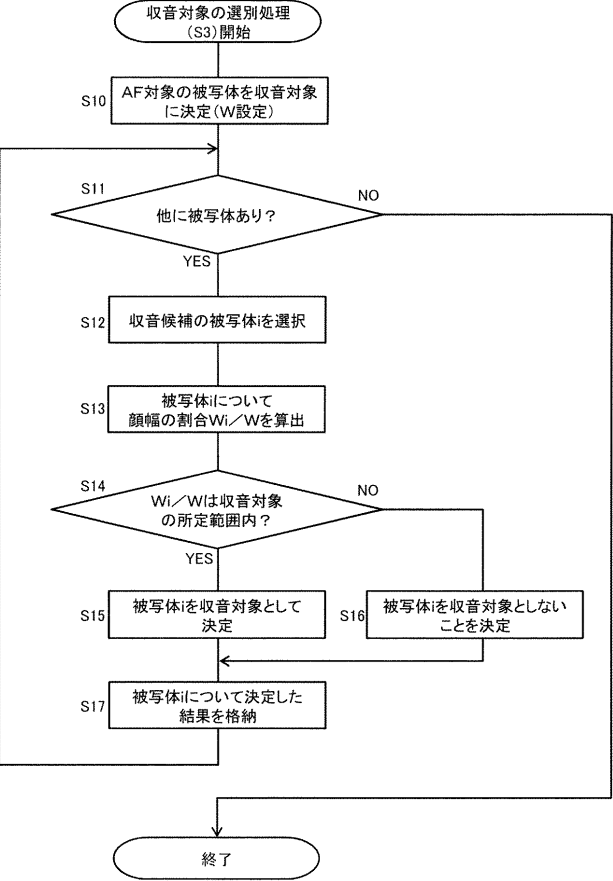
【図 9】

	横撮り		縦撮り	
非自撮り	非顔認識	サラウンド	非顔認識	サラウンド
	顔認識	フォーカス	顔認識	フロント
自撮り	非顔認識	フロント	非顔認識	フロント
	顔認識	フォーカス	顔認識	フロント

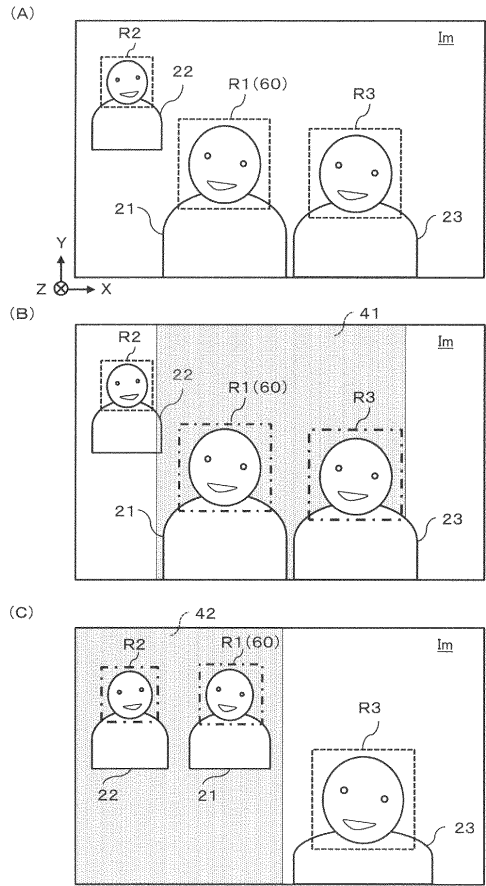
【図 10】



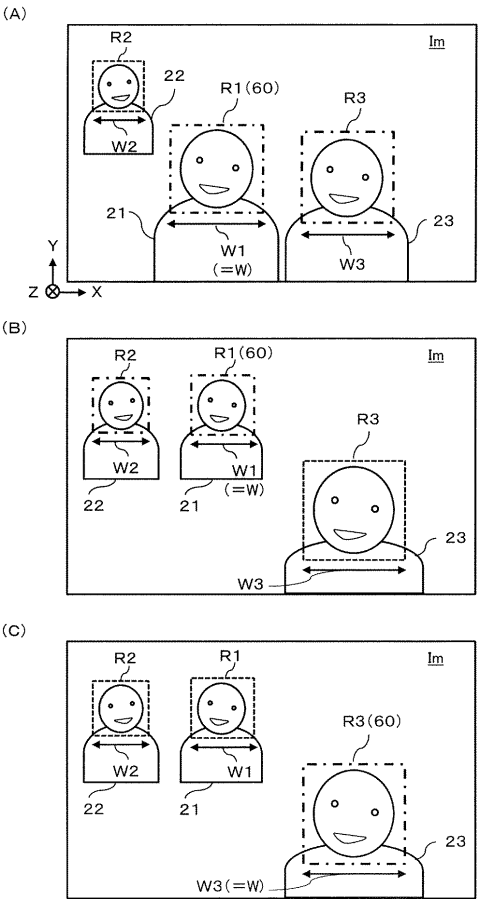
【図 12】



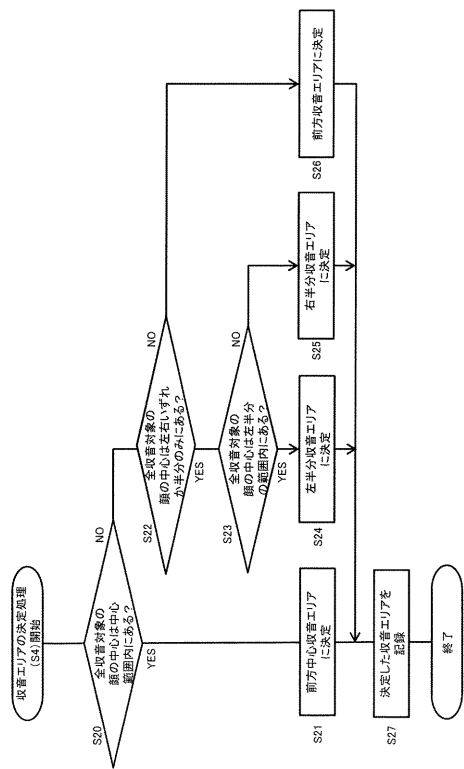
【図 11】



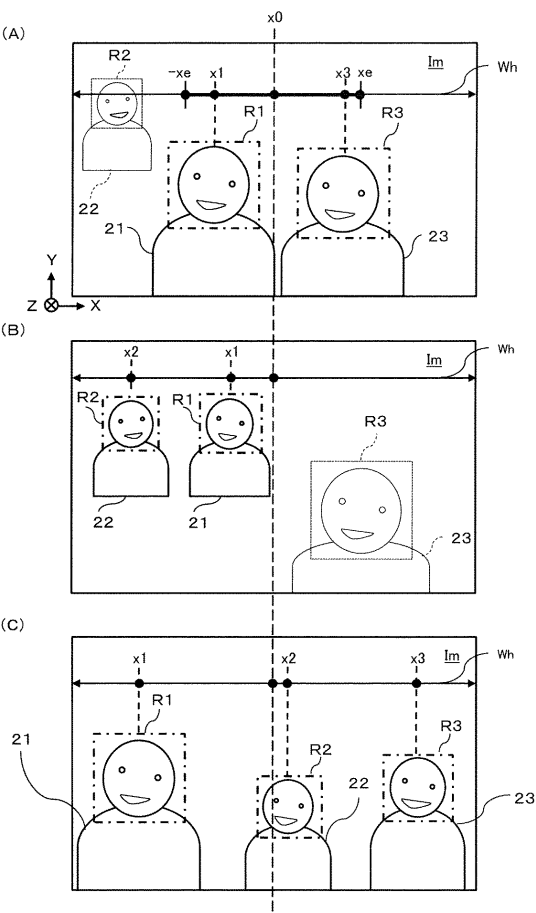
【図 13】



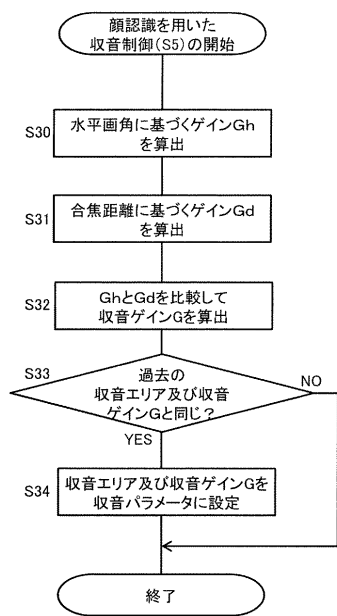
【 図 1 4 】



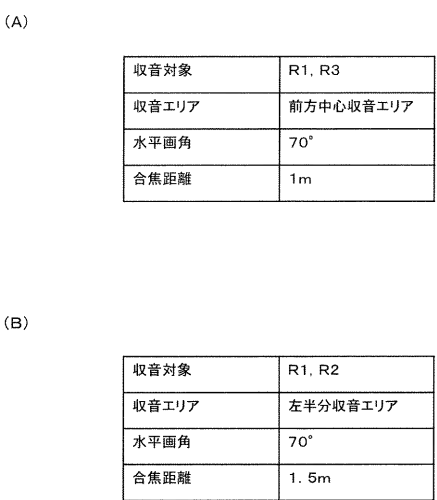
【 図 1 5 】



【 図 1 6 】

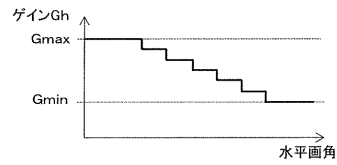


【 図 1 7 】

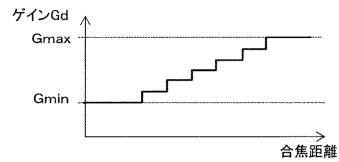


【図 18】

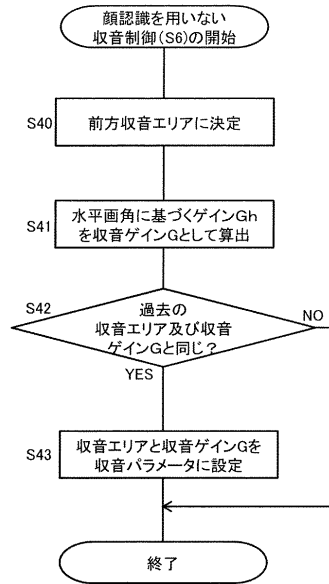
(A)



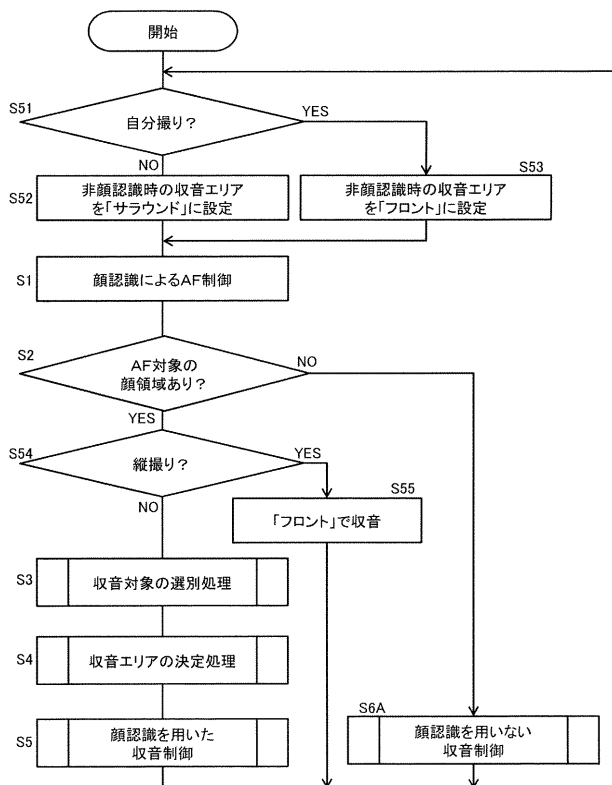
(B)



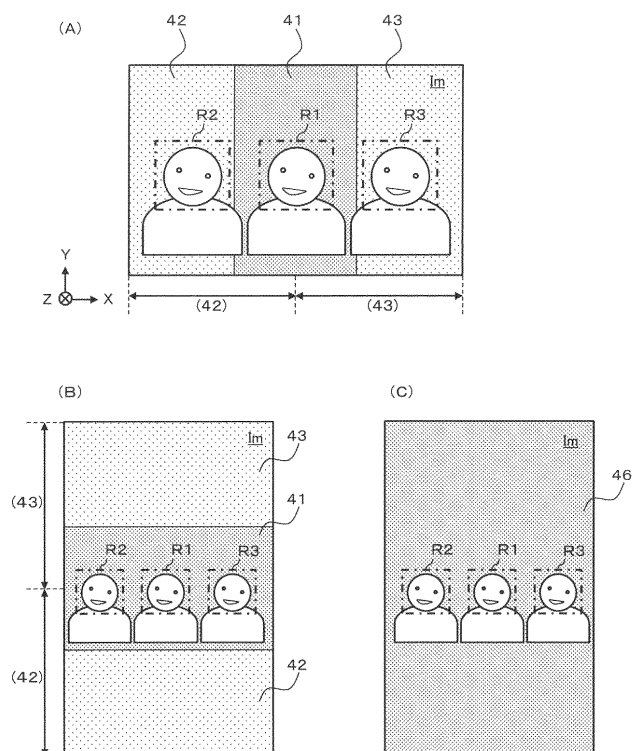
【図 19】



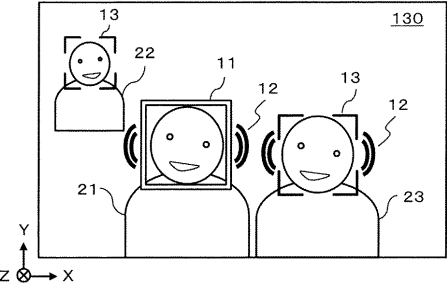
【図 20】



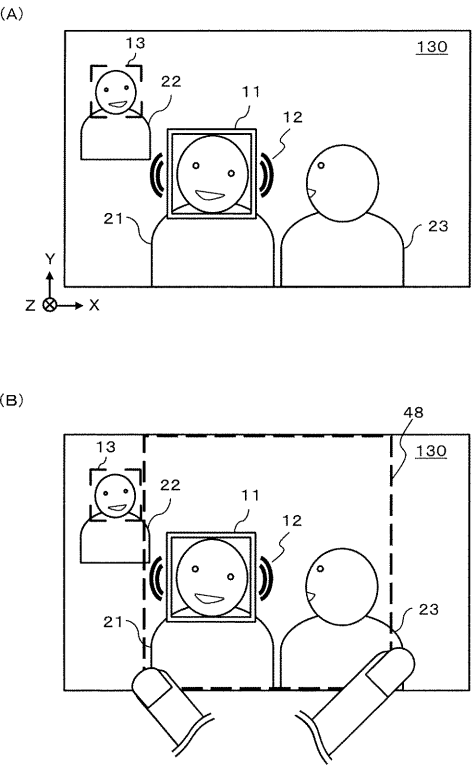
【図 21】



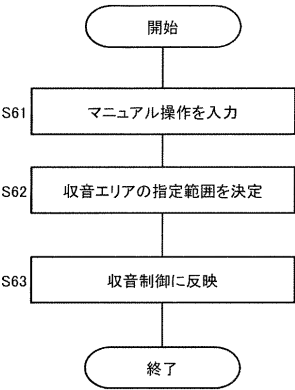
【図 2 2】



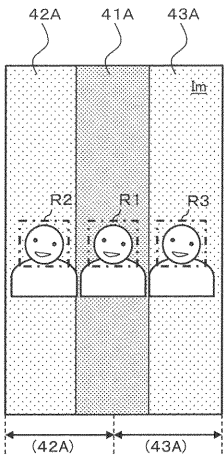
【図 2 3】



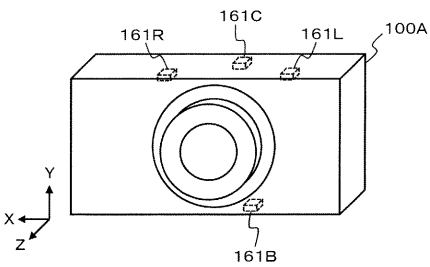
【図 2 4】



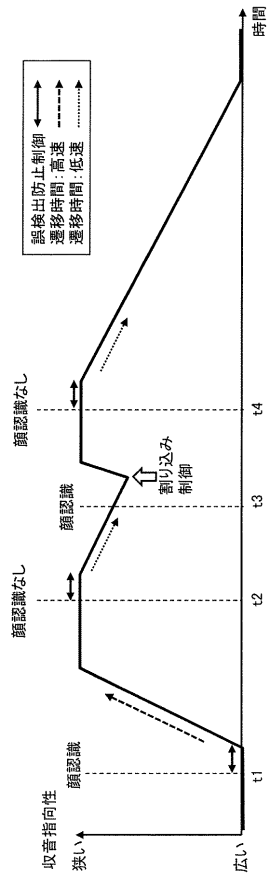
【図 2 6】



【図 2 5】



【図 27】



フロントページの続き

(51)Int.Cl.	F I	テーマコード (参考)
	G 0 3 B 19/00	
	G 0 3 B 31/00	Z

F ターム(参考) 2H054 BB13

5C122	FD13	FH11	FH14	FJ02	FJ04	FJ06	FJ09	FK13	FK29	FK37
FK41	HA75	HB01	HB02	HB05	HB06	HB09				