

(19) 日本国特許庁(JP)

(12) 特許公報(B1)

(11) 特許番号

特許第6739064号
(P6739064)

(45) 発行日 令和2年8月12日(2020.8.12)

(24) 登録日 令和2年7月27日(2020.7.27)

(51) Int. Cl.			F I		
HO4N	5/232	(2006.01)	HO4N	5/232	930
HO4R	3/00	(2006.01)	HO4N	5/232	290
GO3B	15/00	(2006.01)	HO4N	5/232	190
GO2B	7/28	(2006.01)	HO4N	5/232	127
GO3B	13/36	(2006.01)	HO4R	3/00	320

請求項の数 10 (全 30 頁) 最終頁に続く

(21) 出願番号 特願2020-6969 (P2020-6969)
 (22) 出願日 令和2年1月20日(2020.1.20)
 審査請求日 令和2年3月4日(2020.3.4)

早期審査対象出願

(73) 特許権者 314012076
 パナソニックIPマネジメント株式会社
 大阪府大阪市中央区城見2丁目1番61号
 (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

審査官 ▲徳▼田 賢二

最終頁に続く

(54) 【発明の名称】 撮像装置

(57) 【特許請求の範囲】

【請求項1】

情報を表示する表示部と、
 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、
 前記画像データにおいて前記被写体に対応する被写体領域を検出する検出部と、
 前記検出部によって検出された被写体領域に応じて、前記音声信号における收音対象とする被写体及び主要被写体を決定する制御部と
 を備え、

前記制御部は、

前記検出部によって複数の被写体領域が検出されたとき、前記主要被写体で且つ收音対象とする被写体を示す第1の識別情報と、前記主要被写体とは異なる被写体で且つ前記收音対象を示す第2の識別情報とを区別して、前記表示部に表示するように制御する撮像装置。

【請求項2】

前記制御部は、前記收音対象ではない被写体を示す第3の識別情報を、前記第1の識別情報及び前記第2の識別情報と区別して、前記表示部に表示させる請求項1に記載の撮像装置。

【請求項3】

前記第1の識別情報は、前記主要被写体であるか否かを示す第1の識別表示において前

記主要被写体と識別して、且つ前記收音対象であるか否かを示す第2の識別表示において前記收音対象と識別し、

前記第2の識別情報は、前記第1の識別表示において前記主要被写体でないと識別して、且つ前記第2の識別表示において前記收音対象と識別し、

前記第3の識別情報は、前記第1の識別表示において前記主要被写体でないと識別して、且つ前記第2の識別表示において前記收音対象でないと識別する

請求項2に記載の撮像装置。

【請求項4】

前記表示部に表示された情報に基づいて、前記收音対象から特定の被写体を除外するユーザ操作と、特定の被写体を前記收音対象に追加するユーザ操作との少なくとも一方を入力する操作部をさらに備える、

請求項1から3のいずれか1つに記載の撮像装置。

【請求項5】

前記制御部は、

前記被写体領域の大きさに応じて、前記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

請求項1から4のいずれか1つに記載の撮像装置。

【請求項6】

前記検出部は、前記被写体領域として前記画像データにおける人の顔領域を検出する

請求項1から5のいずれか1つに記載の撮像装置。

【請求項7】

前記制御部は、前記主要被写体に合焦するように、前記撮像部の撮像動作を制御する

請求項1から6のいずれか1つに記載の撮像装置。

【請求項8】

情報を表示する表示部と、

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、

前記画像データにおいて前記被写体に対応する被写体領域を検出する検出部と、

前記検出部によって検出された被写体領域に応じて、前記音声信号における收音対象としない被写体及び主要被写体を決定する制御部と

を備え、

前記制御部は、

前記検出部によって複数の被写体領域が検出されたとき、前記主要被写体で且つ前記收音対象とする被写体を示す第1の識別情報と、前記主要被写体とは異なる被写体で且つ前記收音対象としない被写体を示す第3の識別情報とを区別して、前記表示部に表示するように制御する

撮像装置。

【請求項9】

情報を表示する表示部と、

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、

前記画像データにおいて前記被写体に対応する被写体領域を検出する検出部と、

前記検出部によって検出された被写体領域に応じて、前記音声信号における收音対象とする被写体、前記音声信号における收音対象としない被写体及び主要被写体を決定する制御部と

を備え、

前記制御部は、

前記検出部によって複数の被写体領域が検出されたとき、前記主要被写体で且つ前記收音対象とする第1の被写体と、前記主要被写体とは異なり且つ前記收音対象とする第2の

10

20

30

40

50

被写体と、前記主要被写体とは異なり且つ前記收音対象としない第3の被写体とを夫々区分する情報を前記表示部に表示するように制御する撮像装置。

【請求項10】

前記表示部に表示された情報に基づいて、前記收音対象から特定の被写体を除外するユーザ操作と、特定の被写体を前記收音対象に追加するユーザ操作との少なくとも一方を入力する操作部をさらに備える、

請求項8又は9に記載の撮像装置。

【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

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、ユーザが音声の收音を意図する被写体を確認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、表示部と、撮像部と、音声取得部と、検出部と、制御部とを備える。表示部は、情報を表示する。撮像部は、被写体を撮像して画像データを生成する。音声取得部は、撮像部による撮像中に收音される音声を示す音声データを取得する。検出部は、画像データにおいて被写体に対応する被写体領域を検出する。制御部は、検出部によって検出された被写体領域に応じて、音声信号における收音対象とする被写体及び主要被写体を決定する。制御部は、検出部によって複数の被写体領域が検出されたとき、主要被写体で且つ收音対象とする被写体を示す第1の識別情報と、主要被写体とは異なる被写体で且つ收音対象を示す第2の識別情報とを区別して、表示部に表示するように制御する。

40

【発明の効果】

【0006】

本開示に係る撮像装置によると、画像データにおける検出部の検出結果に基づき、検出した被写体領域に応じて收音対象の被写体を決定し、他の被写体と区別して表示する。これにより、ユーザが音声の收音を意図する被写体を確認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

50

- 【図 2】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 3】デジタルカメラ 100 において收音エリアを例示する図
- 【図 4】実施の形態 1 に係るデジタルカメラ 100 の動作を例示するフローチャート
- 【図 5】デジタルカメラ 100 の動作の概要を説明するための図
- 【図 6】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 4 の S 3）を例示するフローチャート
- 【図 7】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 8】デジタルカメラ 100 における收音エリアの決定処理（図 4 の S 4）を例示するフローチャート
- 【図 9】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 10
- 【図 10】デジタルカメラ 100 における顔認識を用いた收音制御（図 4 の S 5）を例示するフローチャート
- 【図 11】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 12】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 13】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 4 の S 6）を例示するフローチャート
- 【図 14】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 15】実施の形態 2 に係るデジタルカメラ 100 の動作を例示するフローチャート
- 【図 16】実施の形態 2 に係るデジタルカメラ 100 における收音対象の追加操作に応じた動作例を示す図 20
- 【図 17】実施の形態 2 に係るデジタルカメラ 100 における收音対象の除外操作に応じた動作例を示す図
- 【図 18】実施の形態 2 に係るデジタルカメラ 100 における收音対象の選別処理（図 15 の S 3 A）を例示するフローチャート
- 【図 19】実施の形態 2 に係るデジタルカメラ 100 における收音対象の表示処理（図 15 の S 5 1）を例示するフローチャート
- 【図 20】実施の形態 2 の変形例 1 を示す図
- 【図 21】実施の形態 2 の変形例 2 を示す図
- 【図 22】実施の形態 2 の変形例 3 を示す図 30
- 【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） 40
- 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】
- 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 15
- 50

0と、通信モジュール155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。

【0012】

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。

10

【0013】

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。

【0014】

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。

20

【0015】

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。

30

【0016】

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、ホワイトバランス補正、ガンマ補正、YC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。

【0017】

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5参照)。

40

【0018】

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

50

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。

【0019】

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。

【0020】

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。コントローラ135はCPU等を含み、CPU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135は、CPU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135は、CPU、MPU、GPU、DSU、FPGA、ASIC等の種々のプロセッサで実現できる。コントローラ135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135は、画像処理エンジン120などと共に1つの半導体チップで構成してもよい。

【0021】

バッファメモリ125は、画像処理エンジン120やコントローラ135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125は、DRAM(Dynamic Random Access Memory)などにより実現される。フラッシュメモリ145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135は各種の内部メモリを有してもよく、例えばROMを内蔵してもよい。ROMには、コントローラ135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135は、CPUの作業領域として機能するRAMを内蔵してもよい。

【0022】

カードスロット140は、着脱可能なメモリカード142が挿入される手段である。カードスロット140は、メモリカード142を電氣的及び機械的に接続可能である。メモリカード142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142は、画像処理エンジン120で生成される画像データなどのデータを格納できる。

【0023】

通信モジュール155は、通信規格IEEE802.11またはWi-Fi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100は、通信モジュール155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100は、通信モジュール155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155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。

【0024】

マイク161は、音を收音する收音部の一例である。マイク161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161は、3つのマイクロフォン素子161L、161C及び161Rを含む。マイク161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。

【0025】

マイク用のA/Dコンバータ165は、マイク161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165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161は、デジタルカメラ100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100は音声取得部として、外部のマイク161に対するインタフェース回路を備える。

10

20

30

40

50

【0026】

音声処理エンジン170は、マイク用のA/Dコンバータ165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170は、本実施形態における音声処理部の一例である。

【0027】

本実施形態の音声処理エンジン170は、例えば図1に示すように、ビーム形成部172と、ゲイン調整部174とを備える。ビーム形成部172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172の詳細については後述する。ゲイン調整部174は、入力される音声データに、例えばコントローラ135によって設定される収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174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。収音ゲイン調整部14はさらに、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。収音ゲインの設定についての詳細は後述する。

10

【0028】

〔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。

【0029】

ビーム形成部172は、マイク161が収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2に示す。

【0030】

図2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で収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の収音指向性の方向および範囲を制御して、マイク161が収音する物理的な範囲を設定できる。

20

【0031】

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、環境に適應して処理を変える適應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。

30

【0032】

図2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の収音指向性を制御できる。また、ビーム形成部172は、収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。

40

【0033】

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の収音エリアについて説明する。

【0034】

〔1-1-2. 収音エリアについて〕

図3は、デジタルカメラ100において定義される収音エリアの例を示す。図3は、収音エリアを、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おり、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。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161L, 161Cおよび161Rが並ぶ方

50

向と一致する。

【0035】

図3(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3(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3(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3(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。

10

【0036】

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3(A)の前方中心收音エリア41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3(B)の左半分收音エリア42を用い、被写体が撮像画像の右半分に位置するとき、図3(C)の右半分收音エリア43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3(D)の前方收音エリア44を用いる。

【0037】

図5(B)の例では、收音対象の被写体R1及びR3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41が用いられる。図5(C)の例では、收音対象の被写体R1及びR2が撮像画像の左半分に位置するため、左半分收音エリア42が用いられる。

20

【0038】

デジタルカメラ100のマイク161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100は、ユーザの撮影場면을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。

【0039】

30

〔1-2.動作〕

以上のように構成されるデジタルカメラ100の動作について説明する。以下では、デジタルカメラ100による動画撮影時の動作を説明する。

【0040】

デジタルカメラ100は順次、光学系110を介して形成された被写体像をイメージセンサ115で撮像して撮像データを生成する。画像処理エンジン120は、イメージセンサ115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125に記録する。また、画像処理エンジン120の顔認識部122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135に出力する。

40

【0041】

本実施形態のデジタルカメラ100は、顔認識部122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(AF)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。

【0042】

以上の撮像動作と同時並行で、デジタルカメラ100は、マイク161において收音を行う。マイク用のA/Dコンバータ165から收音結果の音声データを音声処理エンジン170にて処理する。音声処理エンジン170は、処理後の音声データAoutをバッファメモリ125に記録する。

コントローラ135は、バッファメモリ125を介して、画像処理エンジン120から

50

受け付ける画像データと音声処理エンジン170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142に記録する。また、コントローラ135は逐次、表示モニタ130にスルー画像を表示させる。ユーザは、表示モニタ130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150におけるユーザの操作に応じて開始/終了される。

【0043】

以上のようなデジタルカメラ100の動画撮影は、例えば撮影者とその同伴者など、仲間内で会話をす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。

【0044】

本実施形態のデジタルカメラ100は、画像処理エンジン120における顔認識部122の検出情報によって被写体を検出し、AF対象の被写体が決定されたときに、音声処理エンジン170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120の顔認識と音声処理エンジン170の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收音を精度良く実現する。

【0045】

〔1-2-1.動作の概要〕

図4及び図5を用いて、本実施形態に係るデジタルカメラ100の動作の概要を説明する。

【0046】

図4は、本実施形態に係るデジタルカメラ100の動作を例示するフローチャートである。図4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100の動作モードが顔認識モードであるとき、動画の撮影中に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5は、本実施形態に係るデジタルカメラ100の動作の概要を説明するための図である。

【0047】

コントローラ135は、顔認識部122による検出情報に基づいてAF対象を特定し、AF制御を実行する(S1)。AF対象は、AF制御の対象とする被写体の画像上の領域を示す。図5(A)は、顔認識部122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R1、R2及びR3を含む撮像画像Imを例示する。顔領域R1、R2及びR3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R1がAF対象の顔領域60として特定される。

【0048】

次にコントローラ135は、AF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(S2)。具体的には、コントローラ135は、顔領域が検出されていて、かつAF対象が顔領域であるか否かを判断する。

【0049】

AF対象の顔領域60がある場合(S2でYES)、コントローラ135は、検出情報における被写体からマイク161の收音対象を選別する処理を実行する(S3)。收音対象は、マイク161により音声を強調して收音する対象とする被写体である。AF対象として特定された顔領域R1(60)は、收音対象になる。図5(B)は、図5(A)に示す検出情報に基づき、顔領域R1及びR3を收音対象に決定し、一方で顔領域R2を收音対象としない例を示す。

【0050】

本実施形態のデジタルカメラ100は、收音対象の選別処理(S3)において、AF対象の顔領域R1(60)に加え、撮像画像Imにおいて顔領域R1と同程度の顔の大きさを示す顔R3を、更なる收音対象として決定する。一方で、顔領域R1と異なる大きさの顔領域R2は收音対象から外される。これにより、人物21と人物23がデジタルカメラ100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22は

10

20

30

40

50

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（S3）についての詳細は後述する。

【0051】

次に、コントローラ135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（S4）。收音エリアの決定処理（S4）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5（B）の例において、收音エリアを收音対象の顔領域R1及びR3を含むように前方中心收音エリア41（図3（A））に決定されている。收音エリアの決定処理（S4）についての詳細は後述する。

【0052】

次に、コントローラ135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（S5）。顔認識を用いた收音制御（S5）は、コントローラ135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170に設定することによって行われる。音声処理エンジン170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。

【0053】

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、AF対象の顔領域60がない場合（S2でNO）、コントローラ135は、顔認識を用いない收音制御（S6）を行う。顔認識を用いた、または用いない收音制御（S5、S6）についての詳細は後述する。

【0054】

コントローラ135は、ステップS5またはS6の收音制御を実行後、ステップS1以降の処理を繰り返す。

【0055】

以上の処理によると、本実施形態のデジタルカメラ100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。

【0056】

なお、顔認識によるAF制御（S1）において、検出情報に基づくAF対象の特定は、例えば表示モニタ130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150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。

【0057】

図5（C）は、図5（A）、（B）とは異なる位置に人物21～23がいる場合の撮像画像Imの例を示す。デジタルカメラ100は、図5（B）の例と同様に、まず、例えば顔領域R1をAF対象の顔領域60として特定し（S1）、收音対象に決定する。図5（C）の例において、收音対象の選別処理（S3）は、撮像画像Im上で顔領域R1と同程度の顔の大きさである顔領域R2を收音対象に決定し、顔領域R3を收音対象から外す。收音エリアの決定処理（S4）は、收音対象として決定された顔領域R1及びR2を含む左半分收音エリア42（図3（B））を收音エリアに決定する。顔認識を用いた收音制御（S5）は、左半分收音エリア42に指向性を制御して人物21及び22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。

【0058】

〔1-2-2. 收音対象の選別処理〕

図4のステップS3における收音対象の選別処理の詳細を、図6～7を用いて説明する。

【0059】

図6は、デジタルカメラ100の收音対象の選別処理（S3）を例示するフローチャートである。図6に示すフローチャートによる各処理は、図4のステップS11でYES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。

【0060】

10

20

30

40

50

図7は、デジタルカメラ100における收音対象の選別処理(S3)を説明するための図である。以下では、図5(A)、(B)の例で收音対象を決定する動作について説明する。

【0061】

図6のフローチャートにおいて、コントローラ135は、図4のステップS1において特定したAF対象の顔領域に対応する被写体を收音対象に決定する(S10)。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、AF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。

【0062】

図7(A)は、図5(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W1、W2、W3は、撮像画像Imにおける顔領域R1、R2、R3の大きさをX軸方向の幅で示す。図7(A)の例において、コントローラ135は、AF対象の顔領域R1の顔幅W1を、基準の顔幅Wに設定する(S10)。設定した顔幅Wは、例えばコントローラ135のRAM等に保持される。

10

【0063】

次に、コントローラ135は、AF対象の他に検出された被写体があるか否かを判断する(S11)。具体的には、コントローラ135は、顔認識部122の検出情報がAF対象の顔領域の他に顔領域を含むか否かを判断する。

【0064】

AF対象の他に検出された被写体がある場合(S11でYES)、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(S12)。図7(A)の例では、検出情報はAF対象の顔領域R1の他の顔領域R2及びR3が、ステップS12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。

20

【0065】

コントローラ135は、選択した被写体iの顔幅Wiと、基準の顔幅Wとを比較する演算を行う(S13)。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Wiの割合Wi/Wを算出する。図7(A)の例において、顔領域R2を收音候補とする選択時(S12)には、その顔幅W2についての割合W2/Wが算出される(S13)。

【0066】

コントローラ135は、收音候補の顔幅Wiと基準の顔幅W間の割合Wi/Wが、所定範囲内であるか否かを判断する(S14)。所定範囲は、例えば收音候補の顔幅Wiが相対的に基準の顔幅Wi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。

30

【0067】

コントローラ135は、顔幅の割合Wi/Wが所定範囲内であると判断すると(S14でYES)、被写体iを收音対象とすることを決定する(S15)。

【0068】

一方、コントローラ135は、顔幅の割合Wi/Wが所定範囲内でないと判断すると(S14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(S16)。図7(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。

40

【0069】

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(S15またはS16)、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(S17)。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S11以降の処理を再び行う。

【0070】

50

図7(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(S11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(S12)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(S13)。図7(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(S14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(S15)。

【0071】

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がなくなるまで(ステップS11でNO)、ステップS11~S17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(S3)を終了して、図4のステップS4に進む。

10

【0072】

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。

【0073】

図7(B)は、図5(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7(B)の例において、顔領域R1は、図7(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(S10)。

20

【0074】

図7(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(S14でYES)、顔領域R2の被写体を收音対象として決定する(S15)。一方、割合 $W3/W$ が所定範囲の上限値を上回るため(S14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(S16)。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5(C)参照)。

【0075】

図7(C)は、図5(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4のS1)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(S10)。図7(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(S14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(S16)。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。

30

【0076】

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。

【0077】

〔1-2-3. 收音エリアの決定処理〕

図4のステップS4における收音エリアの決定処理の詳細を、図8~9を用いて説明する。

40

【0078】

図8は、本実施形態のデジタルカメラ100における收音エリアの決定処理(S4)を例示するフローチャートである。図8に示すフローチャートによる各処理は、図4のステップS3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。

【0079】

図9は、デジタルカメラ100における收音エリアの決定処理(S4)を説明するため

50

の図である。図9(A), (B)は、それぞれ図7(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9(C)は、図9(A), (B)とは更に別の場合を例示する。図9(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。

【0080】

図8のフローチャートにおいて、コントローラ135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(S20)。中心範囲は、撮像画像 I_m において前方中心收音エリア41に対応付けられる範囲である。

10

【0081】

中心範囲は、例えば図9(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135は、例えば光学系110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。

【0082】

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。

20

【0083】

全收音対象の顔領域の位置が中心範囲内にある場合(S20でYES)、コントローラ135は、收音エリアを前方中心收音エリア41に決定する(S21)。図9(A)の例では、收音対象は顔領域R1及びR3に対応する。それぞれの顔領域R1及びR3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(S20でYES)。よって、收音エリアが前方中心收音エリア41に決定される(S21, 図5(B)参照)。

【0084】

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(S20でNO)、前方中心收音エリア41以外の收音エリアが用いられる。この場合、コントローラ135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(S22)。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。

30

【0085】

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(S22でYES)、コントローラ135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(S23)。

【0086】

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(S23でYES)、コントローラ135は、收音エリアを左半分收音エリア42に決定する(S24)。図9(B)の例では、收音対象は顔領域R1及びR2に対応する。顔領域R1の位置 x_1 及び顔領域R2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(S23でYES)、收音エリアは左半分收音エリア42に決定される(S24, 図5(C)参照)。

40

【0087】

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(S23でNO)、コントローラ135は、收音エリアを右半分收音エリア43に決定する(S25)。

50

【 0 0 8 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 3 (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 0 8 9 】

図 9 (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)。

10

【 0 0 9 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 4 のステップ S 5 に進む。

【 0 0 9 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。

20

【 0 0 9 2 】

図 1 1 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 1 (A) は、図 7 (A) 及び図 9 (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 1 (B) は、図 7 (B) 及び図 9 (B) の例における管理情報を例示する。

【 0 0 9 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。

30

【 0 0 9 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 等に記録される。また、画角 θ_e 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 1 5 0 により設定した値がバッファメモリ 1 2 5 等に保持されてもよい。

40

【 0 0 9 5 】

[1 - 2 - 4 . 收音制御]

(1) 図 4 のステップ S 5 について

図 4 のステップ S 5 における顔認識を用いた收音制御の詳細を、図 1 0 ~ 1 2 を用いて説明する。

【 0 0 9 6 】

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 1 0 0 は、例えば A F 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 1 0 0 は、例えばデジタルカメラ 1 0 0 が動画の撮影中

50

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。

【0097】

図10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10のフローチャートに示す各処理は、図4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。

【0098】

デジタルカメラ100は、図11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。

10

【0099】

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12(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12(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。

【0100】

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12(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12(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。

20

【0101】

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。

【0102】

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4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。

30

【0103】

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、ステップS4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S32で算出した收音ゲイン G を、音声処理エンジン170に收音パラメータとして設定する(S34)。音声処理エンジン170は、ビーム形成部172及びゲイン調整部174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。

40

【0104】

收音パラメータの設定(S34)後、コントローラ135は、顔認識を用いた收音制御の処理(S5)を終了する。また、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ないと判断した場合(S33でNO)、ステップS34の処理を行わずに図4のステップS5の処理を終了する。その後、図4のステップS1以降の処理が繰り返される。

【0105】

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

50

び收音エリアを、收音パラメータに設定して、AF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。

【0106】

なお、ステップS30とS31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S31でゲインGdを算出してから、ステップS30でゲインGhを算出してもよく、またはステップS30とS31を並列に実行してもよい。

【0107】

また、以上のステップS33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（S34）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（S5）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。

10

【0108】

（2）図4のステップS6について

図4のステップS6における顔認識を用いない收音制御（S6）の詳細を、図13を用いて説明する。

【0109】

図13は、顔認識を用いない收音制御（S6）を例示するフローチャートである。図13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4のステップS2においてAF対象の顔領域がない（S2でNO）場合に、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。

20

【0110】

まず、コントローラ135は、收音エリアを、例えば前方收音エリア44に決定する（S40）。

【0111】

次に、コントローラ135は、水平画角に基づくゲインGhをステップS30と同様に算出して、收音ゲインGとする（S41）。さらに、コントローラ135は、ステップS33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（S42）。

【0112】

30

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（S42でYES）、收音エリアと收音ゲインGを收音パラメータに設定し（S43）、顔認識を用いない收音制御（S6）を終了する。また、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ないと判断した場合（S42でNO）、ステップS43の処理を行わずに図4のステップS6を終了する。ステップS6の終了後、ステップS1以降の処理が繰り返される。

【0113】

以上の処理によると、AF対象の顔領域がない場合でも、デジタルカメラ100の前方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。

40

【0114】

なお、デジタルカメラ100の動作モードに応じて、デジタルカメラ100の周囲360°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S40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。

【0115】

〔1-3.効果等〕

【0116】

本実施形態において、コントローラ135は、コントローラ135は、顔認識部122

50

によって検出された被写体領域、すなわち顔領域 R 1 ~ R 3 の大きさである顔幅 W 1 , W 2 及び W 3 に応じて、音声信号における收音対象とする被写体を決定し (S 3)、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 (S 4 ~ S 5)。これにより、例えば複数の被写体から、被写体領域の大きさに応じて、收音対象とする被写体と、收音対象としない被写体との選別を実現することができる。

【 0 1 1 7 】

本実施形態において、顔認識部 1 2 2 は、被写体領域として画像データにおける人の顔領域 R 1 ~ R 3 を検出する。なお、被写体領域は、人の顔領域に限らず、例えば動物の顔領域が検出されてもよい。また、被写体領域の大きさは、顔幅に限らず、例えば撮像画像 I m における顔領域 R 1 ~ R 3 の Y 軸方向の幅、または顔領域 R 1 ~ R 3 の面積であって

10

【 0 1 1 8 】

本実施形態において、コントローラ 1 3 5 は、A F 対象の被写体 (主要被写体) に合焦するように、イメージセンサ 1 1 5 (撮像部) の撮像動作を制御する。なお、主要被写体は A F 対象に限定されない。デジタルカメラ 1 0 0 は、主要被写体について合焦の動作に代えて又はこれに加えて、例えば露出制御といった各種の動作を実行してもよい。

【 0 1 1 9 】

(実施の形態 2)

【 0 1 2 0 】

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 1 0 0 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 1 0 0 について説明する。

20

【 0 1 2 1 】

以下、実施の形態 1 に係るデジタルカメラ 1 0 0 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 1 0 0 を説明する。

【 0 1 2 2 】

[2 - 1 . 概要]

図 1 4 を用いて、本実施形態に係るデジタルカメラ 1 0 0 が各種情報を表示する動作の概要を説明する。

30

【 0 1 2 3 】

図 1 4 は、本実施形態に係るデジタルカメラ 1 0 0 の表示例を示す。図 1 4 の表示例は、デジタルカメラ 1 0 0 が図 5 (B) に例示するように收音対象を決定した場合に、表示部 1 3 0 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 1 0 0 は表示モニタ 1 3 0 に、A F 対象の被写体を示す A F 枠 1 1 及び A F 対象以外の検出された被写体を示す検出枠 1 3 に加えて、收音対象の被写体を示す收音アイコン 1 2 を、撮像画像 I m 上に重畳して表示している。

【 0 1 2 4 】

本実施形態のデジタルカメラ 1 0 0 は、收音アイコン 1 2 を、A F 枠 1 1 及び検出枠 1 3 に組み合わせて用いることで、A F 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、A F 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。

40

【 0 1 2 5 】

例えば図 1 4 の表示例において、デジタルカメラ 1 0 0 は、図 5 (B) の例で顔領域 R 1 (6 0) に対応する被写体を A F 対象かつ收音対象として決定したことから、人物 2 1 に A F 枠 1 1 と收音アイコン 1 2 を表示する。また、デジタルカメラ 1 0 0 は、図 5 (B) の例で顔領域 R 3 に対応する被写体を A F 対象以外の收音対象として決定したことから、人物 2 3 に検出枠 1 3 と收音アイコン 1 2 を表示する。さらに、デジタルカメラ 1 0 0 は、收音アイコン 1 2 を伴わない検出枠 1 3 の表示により、図 5 (B) の例で顔領域 R 2 に対応する A F 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。

50

【 0 1 2 6 】

本実施形態のデジタルカメラ 1 0 0 において、ユーザは、A F 枠 1 1 または検出枠 1 3 のいずれか一方の表示により、検出した被写体が A F 対象か否かを確認できる。ユーザはまた、收音アイコン 1 2 の有無により、收音対象か否かを確認できる。A F 枠 1 1 と收音アイコン 1 2 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 1 3 と收音アイコン 1 2 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 1 3 は、第 3 の識別情報の一例である。

【 0 1 2 7 】

以上のように、本実施形態に係るデジタルカメラ 1 0 0 は、検出情報が含む被写体から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 1 0 0 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。さらに、本実施形態のデジタルカメラ 1 0 0 は、ユーザが当該表示に基づき收音対象に被写体を追加する操作をしたり、收音対象から被写体を除外する操作をしたりすることができるように、こうしたユーザ操作を入力可能に動作する。

10

【 0 1 2 8 】

〔 2 - 2 . 動作の詳細 〕

図 1 5 は、本実施形態に係るデジタルカメラ 1 0 0 の動作を例示するフローチャートである。図 1 5 のフローチャートに示す各処理は、実施の形態 1 における図 4 のフローチャートと同様に、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。図 1 5 におけるステップ S 2 , S 4 ~ S 6 の処理は実施の形態 1 (図 4) と同様であり、以下ではその説明を省略する。

20

【 0 1 2 9 】

本実施形態では、例えば顔認識による A F 制御 (S 1 A) において、コントローラ 1 3 5 は、図 4 のステップ S 1 と同様に顔認識部 1 2 2 の検出結果に基づき A F 対象を決定すると、上述した A F 枠 1 1 及び検出枠 1 3 を表示モニタ 1 3 0 に表示させる。

【 0 1 3 0 】

さらに、本実施形態の收音対象の選別処理 (S 3 A) は、図 4 のステップ S 3 と同様の処理において、上記のユーザ操作を反映する処理を実行する。本実施形態における收音対象の選別処理 (S 3 A) についての詳細は後述する。

30

【 0 1 3 1 】

本実施形態においてコントローラ 1 3 5 は、例えば顔認識を用いた收音制御 (S 5) の実行後、收音対象に関する情報を表示する処理を行う (S 5 1) 。図 1 4 の表示例において、コントローラ 1 3 5 は、ステップ S 1 A で表示した A F 枠 1 1 及び検出枠 1 3 に加えて、ステップ S 3 A で決定した收音対象の被写体に收音アイコン 1 2 を表示させる (S 5 1) 。收音対象の表示処理 (S 5 1) についての詳細は後述する。

【 0 1 3 2 】

次に、コントローラ 1 3 5 は、操作部 1 5 0 において、決定した收音対象に被写体を追加するユーザ操作または收音対象から被写体を除外するユーザ操作が入力されたか否かを判断する (S 5 2) 。こうした收音対象の追加または除外操作は、例えば表示モニタ 1 3 0 の表示画面と重畳して配置された操作部 1 5 0 のタッチパネルにおいて、位置を指定するタッチ操作に予め設定される (図 1 6 , 1 7 参照) 。上記のユーザ操作を入力すると、コントローラ 1 3 5 は、指定された位置を含む顔領域等の被写体を示す情報と、收音対象の追加または除外操作の何れかを示す情報とを対応付けて、例えばバッファメモリ 1 2 5 に入力情報として保持する。

40

【 0 1 3 3 】

コントローラ 1 3 5 は、收音対象の追加または除外操作が入力されたと判断した場合 (S 5 2 で Y E S) 、ステップ S 3 A 以降を再び実行する。この際の收音対象の選別処理 (S 3 A) において、ステップ S 5 2 で入力された入力情報が参照される。

【 0 1 3 4 】

50

図16は、收音対象の追加操作に応じた動作例を示す図である。図16(A)は、図14の表示例から、ユーザが收音対象の追加操作を行う例を示す。図16(B)は、図16(A)の追加操作に応じて收音エリアが図5(B)から変更された状態を例示する。図16(C)は、図16(B)の状態に対応する表示モニタ130の表示例を示す。

【0135】

図16の動作例において、ユーザは、收音アイコン12が表示されていない人物22の検出枠13にタッチ操作を行う。当該追加操作に回答して、デジタルカメラ100のコントローラ135は、顔領域R2の被写体を收音対象として決定し(S3A)、顔領域R2を含む前方收音エリア44に決定する(S4)。その後、図16(C)に示すように、コントローラ135は、收音対象に決定した被写体に收音アイコン12を表示する(S51)。

10

【0136】

図17は、收音対象の除外操作に応じた動作例を示す図である。図17(A)は、図16の追加操作が入力された後、ユーザが收音対象の除外操作を行う例を示す。図17(B)は、図17(A)の除外操作に応じて收音エリアが図16(B)から変更された状態を例示する。図17(C)は、図17(B)の状態に対応する表示モニタ130の表示例を示す。

【0137】

図17の動作例において、ユーザは、收音アイコン12が表示されている人物23の検出枠13にタッチ操作を行う。当該除外操作に回答して、コントローラ135は、顔領域R3の被写体を收音対象としないことを決定し(S3A)、顔領域R3を含まない左半分收音エリア42に決定する(S4)。その後、図17(C)に示すように、コントローラ135は、收音対象としないことを決定した被写体には收音アイコン12を表示しない(S51)。

20

【0138】

図15に戻り、收音対象の被写体を追加または除外する操作が特に入力されない場合(S52でNO)、コントローラ135は、ステップS1A以降の処理を繰り返す。

【0139】

以上の処理によると、本実施形態のデジタルカメラ100は、検出した被写体にAF枠11または検出枠13を表示し(S1A)、收音対象を決定して(S3A)、收音対象か否かを区別する表示を行う(S51)。デジタルカメラ100は、これらの表示に基づくユーザによる收音対象の追加または除外操作を受け付ける(S52)。

30

【0140】

このように、本実施形態のデジタルカメラ100は、被写体が收音対象か否かの表示に基づくユーザ操作に応じて、ユーザ指定の被写体を図16に示すように收音対象に追加したり、図17に示すように收音対象から除外したりすることができる。これにより、デジタルカメラ100は、決定した收音対象がユーザの意図とは異なる場合であってもユーザの意図に沿った收音対象に変更可能である。

【0141】

なお、図15のステップS5とS51の処理は、以上の説明とは異なる順序で実行されてもよい。例えば、ステップS51の処理を実行後に、ステップS5の処理が行われてもよい。また、ステップS52におけるユーザ操作の入力は、特にタッチ操作に限らず、例えば操作部150のカーソルキーまたはジョイスティックにより行われてもよい。

40

【0142】

〔2-2-1. 実施の形態2に係る收音対象の選別処理〕

図18は、本実施形態のデジタルカメラ100における收音対象の選別処理(図15のS3A)を例示するフローチャートである。以下、図15のステップS52において收音対象の追加または除外操作が入力された(S52でYES)場合の例について説明する。本実施形態では、実施の形態1における図6のフローチャートに示す処理に加え、ステップS52で入力された操作に応じて收音対象を決定する処理が実行される。図18におけ

50

るステップS 1 0 ~ S 1 2 , S 1 3 ~ S 1 7 の処理は実施の形態 1 (図 6) と同様であり、以下ではその説明を省略する。

【 0 1 4 3 】

コントローラ 1 3 5 は、図 6 のフローチャートと同様に收音候補の被写体 i を選択すると (S 1 2) 、当該被写体 i に関して收音対象の追加操作があったか否かを判断する (S 7 1) 。具体的には、コントローラ 1 3 5 は、図 1 5 のステップ S 5 2 においてバッファメモリ 1 2 5 に格納した入力情報を参照して、被写体 i が追加操作と対応付いて入力情報に記録されているか否かを判断する。

【 0 1 4 4 】

被写体 i の追加操作がない場合 (S 7 1 で NO) 、コントローラ 1 3 5 は、同被写体 i に関して收音対象の除外操作があったか否かを判断する (S 7 2) 。この際、コントローラ 1 3 5 は、ステップ S 7 1 の判断と同様に入力情報を参照して、被写体 i が除外操作と対応付いて入力情報に記録されているか否かを判断する。

10

【 0 1 4 5 】

被写体 i の除外操作がない場合 (S 7 2 で NO) 、コントローラ 1 3 5 は、図 6 のフローチャートと同様にステップ S 1 3 以降の処理を実行する。この場合、実施の形態 1 と同様に、顔領域の大きさの比較に応じて被写体 i を收音対象とするか否かが決定される。

【 0 1 4 6 】

一方、コントローラ 1 3 5 は、被写体 i の追加操作があった判断すると (S 7 1 で YES) 、ステップ S 1 3 ~ S 1 4 の処理を特に実行せずに、被写体 i を收音対象とすることを決定する (S 1 5) 。図 1 6 の例では、人物 2 2 の顔領域 R 2 が收音候補の被写体 i に選択されたときに (S 1 2) 、コントローラ 1 3 5 は、ステップ S 7 1 で YES に進み、顔領域 R 2 の大きさに依らず当該被写体 i を收音対象とすることを決定する。

20

【 0 1 4 7 】

また、コントローラ 1 3 5 は、被写体 i の除外操作があったと判断すると (S 7 2 で YES) 、ステップ S 1 3 ~ S 1 4 の処理を特に実行せずに、被写体 i を收音対象としないことを決定する (S 1 6) 。図 1 7 の例では、人物 2 3 の顔領域 R 3 が收音候補の被写体 i に選択されたときに (S 1 2) 、コントローラ 1 3 5 は、ステップ S 7 2 で YES に進み、顔領域 R 3 の大きさに依らず当該被写体 i を收音対象としないことを決定する (S 1 6) 。

30

【 0 1 4 8 】

なお、上記の例では、收音対象の追加操作及び除外操作が入力されたか否かを判断する (S 7 1 , S 7 2) デジタルカメラ 1 0 0 を説明したが、これに限らない。例えば、デジタルカメラ 1 0 0 は、図 1 5 のステップ S 5 2 において追加操作または除外操作のいずれか一方のみの入力を受け付け、各操作に応じて図 1 8 のステップ S 7 1 または S 7 2 のいずれか一方のみを実行してもよい。

【 0 1 4 9 】

〔 2 - 2 - 2 . 收音対象の表示処理 〕

図 1 5 のステップ S 5 1 における收音対象の表示処理 (S 5 1) の詳細を、図 1 9 を用いて説明する。

40

【 0 1 5 0 】

図 1 9 は、收音対象の表示処理 (S 5 1) を例示するフローチャートである。以下では、図 5 (B) の例と対応する図 1 4 の表示例において、收音対象を表示する動作について説明する。

【 0 1 5 1 】

まず、コントローラ 1 3 5 は、收音対象の選別処理 (図 1 5 の S 3 A) で AF 対象の被写体を收音対象に決定した (図 6 の S 1 0 参照) ことに依じて、表示モニタ 1 3 0 において AF 対象の被写体に收音アイコン 1 2 を表示させる (S 6 0) 。図 5 (B) の例では、表示モニタ 1 3 0 は、図 1 4 に示すように AF 対象の顔領域 6 0 に対応する被写体に、AF 枠 1 1 と共に收音アイコン 1 2 を表示する。

50

【 0 1 5 2 】

次に、コントローラ 1 3 5 は、図 4 のステップ S 4 において決定した收音エリアに含まれる被写体の中で、A F 対象の他に被写体があるか否かを判断する (S 6 1)。具体的には、コントローラ 1 3 5 は、顔認識部 1 2 2 の検出情報における位置情報と、決定した收音エリアの方向及び角度範囲に対応する撮像画像 I m 上の位置とを比較することにより、收音エリアに A F 対象の他に被写体を含むか否かを判断する。

【 0 1 5 3 】

收音エリアに A F 対象の他に含まれる被写体がある場合 (S 6 1 で Y E S)、コントローラ 1 3 5 は、当該被写体から 1 つの被写体 i を選択する (S 6 2)。図 5 (B) の例では、前方中心收音エリア 4 1 に、A F 対象の顔領域 6 0 の他に顔領域 R 3 が含まれることから、コントローラ 1 3 5 は、顔領域 R 3 を選択する。

10

【 0 1 5 4 】

続いてコントローラ 1 3 5 は、選択した被写体 i に收音アイコン 1 2 を表示する (S 6 3)。これにより、図 5 (B) の例では、表示モニタ 1 3 0 は、図 1 4 に示すように顔領域 R 3 に対応する人物 2 3 に、検出枠 1 3 と共に收音アイコン 1 2 を表示している。

【 0 1 5 5 】

その後、コントローラ 1 3 5 は、選択した被写体 i に收音アイコン 1 2 を表示すると (S 6 3)、選択済みの被写体の他に收音エリアに含まれる被写体があるか否かを判断し (S 6 1)、他に被写体があると (S 6 1 で Y E S)、ステップ S 4 2 以降の処理を再び行う。

20

【 0 1 5 6 】

コントローラ 1 3 5 は、收音エリアに他に含まれる被写体がなくなると (S 4 1 で N O)、收音対象の表示処理 (S 5 1) を終了して、図 1 5 のステップ S 5 2 に進む。

【 0 1 5 7 】

以上の処理によると、收音エリアに含まれる被写体に收音アイコン 1 2 を表示することで、ユーザが收音対象の被写体を把握することができる。これによりユーザは、A F 対象といった主要被写体に加えて收音を意図する被写体が、收音対象であるか否かを確認し易い。

【 0 1 5 8 】

なお、図 1 9 に例示するフローチャートにおいて、コントローラ 1 3 5 は、收音エリアに含まれる被写体から被写体 i を選択した (S 6 1 ~ 6 2) が、これに限らない。例えば、被写体 i は、図 1 5 のステップ S 3 A で決定した收音対象から選択されてもよく、検出情報が含む全ての被写体から選択されてもよい。また、例えば図 1 5 のフローチャートにおいてステップ S 5 1 が 2 回目以降に実行される場合、ステップ S 6 1 は省略されてもよく、別の判断基準が用いられてもよい。

30

【 0 1 5 9 】

〔 2 - 3 . 効果等 〕

以上のように、上記の実施の形態のデジタルカメラ 1 0 0 は、表示モニタ 1 3 0 (表示部) と、イメージセンサ 1 1 5 (撮像部) と、マイク用の A / D コンバータ 1 6 5 (音声取得部) と、顔認識部 1 2 2 (検出部) と、コントローラ 1 3 5 (制御部) とを備える。表示モニタ 1 3 0 は、情報を表示する。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。マイク用の A / D コンバータ 1 6 5 は、イメージセンサ 1 1 5 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。顔認識部 1 2 2 は、イメージセンサ 1 1 5 によって生成された画像データにおいて、被写体に対応する被写体領域の一例として顔領域 R 1 , R 2 及び R 3 を検出する。コントローラ 1 3 5 は、顔認識部 1 2 2 によって検出された顔領域 R 1 ~ R 3 に応じて、音声信号における收音対象とする被写体及び A F 対象の被写体 (主要被写体) を決定する。コントローラ 1 3 5 は、顔認識部 1 2 2 によって複数の顔領域 R 1 ~ R 3 が検出されたとき、A F 対象で且つ收音対象とする被写体を示す A F 枠 1 1 及び收音アイコン 1 2 (第 1 の識別情報) と、主要被写体とは異なる被写体で且つ收音対象を示す検出枠 1 3 及び收音アイコン 1 2 (第 2 の識別情報) とを区別し

40

50

て表示モニタ130に表示するように制御する(S1A, S51)。

【0160】

以上のデジタルカメラ100によると、主要被写体で且つ收音対象の被写体と、主要被写体以外で且つ收音対象の被写体とが区別して表示されることから、主要被写体以外にも收音対象となっている被写体があるのがユーザに可視化される。これにより、ユーザが音声の收音を意図する被写体を確認し易くすることができる。

【0161】

実施の形態2において、コントローラ135は、收音対象ではない被写体を示す検出枠13(第3の識別情報)を、第1の識別情報及び第2の識別情報と区別して、表示モニタ130に表示させる(S1A)。これにより、ユーザは表示モニタ130において、検出された被写体のうち收音対象ではない被写体も確認することができる。

10

【0162】

本実施形態において、第1の識別情報は、AF対象の被写体(主要被写体)であるか否かを示すAF枠11の有無(第1の識別表示)において主要被写体と識別して、且つ收音対象であるか否かを示す收音アイコン12の有無(第2の識別表示)において收音対象と識別する。第2の識別情報は、AF枠11の有無においてAF対象の被写体でないと識別して、且つ收音アイコン12の有無において收音対象と識別する。第3の識別情報は、AF枠11の有無においてAF対象でないと識別して、且つ收音アイコン12の有無において收音対象でないと識別する(図14参照)。

【0163】

20

以上のデジタルカメラ100によると、AF枠11においてAF対象か否か、及び收音アイコン12において收音対象か否かの2つの観点から被写体を識別する表示を行うことで、ユーザにとって分かり易い可視化をすることができる。

【0164】

本実施形態において、デジタルカメラ100は、表示モニタ130に表示された情報に基づいて、收音対象から特定の被写体を除外するユーザ操作と、特定の被写体を收音対象に追加するユーザ操作との少なくとも一方を入力する操作部150をさらに備える。これにより、デジタルカメラ100は、決定した收音対象の被写体がユーザの意図とは異なる場合であっても、例えばユーザが所望の被写体を選択する操作を入力することで、ユーザの意図に沿って收音対象に追加したり、收音対象から除外したりすることができる。

30

【0165】

(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。

【0166】

実施の形態2では、デジタルカメラ100は、図14に例示するAF枠11、收音アイコン12及び検出枠13を表示した(図15のS1A, S51)。本実施形態のデジタルカメラ100は、実施の形態2とは異なる表示態様において、決定された收音対象に関する情報をユーザに可視化してもよい。図20、図21及び図22は、それぞれ実施の形態2の変形例1、変形例2及び変形例3を示す図である。

40

【0167】

図20、図21は、それぞれ実施の形態2とは異なる收音アイコン12A及び12Bを表示した表示モニタ130の例を示す。図20は、マイクを模した收音アイコン12Aを例示する。図21は、枠表示による收音アイコン12Bを例示する。收音アイコン12Bは、特に図21の例に限らず、例えば矩形状の枠表示であってもよい。また、本実施形態のデジタルカメラ100は、例えばユーザが收音アイコン12A, 12Bといった各種の收音アイコンから、表示させる收音アイコンを操作部150において選択するユーザ操作を受け付けてもよい。

50

【0168】

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、AF枠11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12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22は、本実施形態における3種類の枠表示を例示する。図22の例では、AF対象かつ收音対象の被写体を示す枠表示11A、AF対象以外で收音対象の被写体を示す枠表示13A及び收音対象でない被写体を示す枠表示13Bにより、AF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。

【0169】

実施の形態1～2では、図4のフローチャートにおいて、デジタルカメラ100が内蔵するマイク161について、顔認識を用いた又は用いない收音制御(S5又はS6)を行う動作例を説明した。本実施形態のデジタルカメラ100は、内蔵のマイク161に代えて、外付けのマイク(以下「マイク161aという」)を備えてもよい。マイク161aは、デジタルカメラ100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135は、マイク161a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125等に保持しておくことにより、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。

【0170】

また、実施の形態1～2では、図10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Gh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8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Gh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Ghを算出することができる。

【0171】

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。

【0172】

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155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155が検出部として機能する。

【0173】

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。

【0174】

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

10

20

30

40

50

カメラ、スマートフォン、タブレット端末等)であればよい。

【0175】

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。

【0176】

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。

10

【0177】

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】

【0178】

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】

【0179】

100 デジタルカメラ

20

115 イメージセンサ

120 画像処理エンジン

122 顔認識部

125 バッファメモリ

130 表示モニタ

135 コントローラ

145 フラッシュメモリ

150 操作部

11 AF枠

12, 12A, 12B 収音アイコン

30

13 検出枠

【要約】

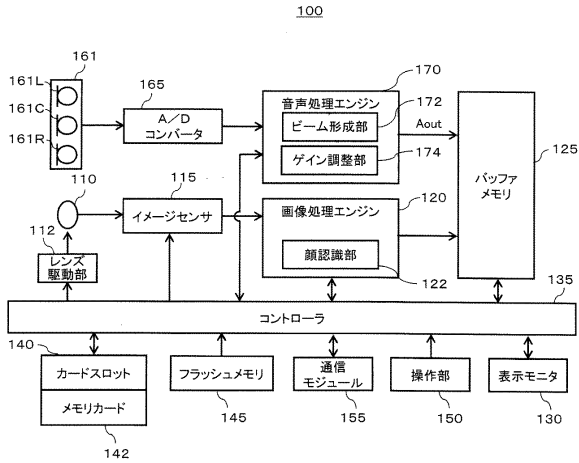
【課題】音声を取得しながら撮像を行う撮像装置において、ユーザが音声の収音を意図する被写体を確認し易くすることができる撮像装置を提供する。

【解決手段】撮像装置(100)は、情報を表示する表示部(130)と、被写体を撮像して画像データを生成する撮像部(115)と、撮像部による撮像中に収音される音声を示す音声データを取得する音声取得部(165)と、検出部は、画像データにおいて被写体に対応する被写体領域を検出する検出部(122)と、検出部によって検出された被写体領域に応じて、音声信号における収音対象とする被写体及び主要被写体を決定する制御部(135)とを備える。制御部は、検出部によって複数の被写体領域が検出されたとき、主要被写体で且つ収音対象とする被写体を示す第1の識別情報と、主要被写体とは異なる被写体で且つ収音対象を示す第2の識別情報とを区別して、表示部に表示するように制御する。

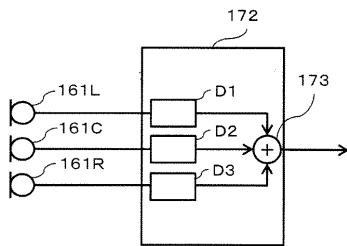
40

【選択図】図14

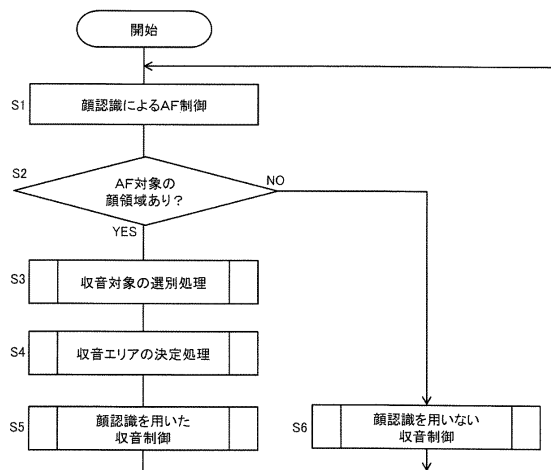
【図1】



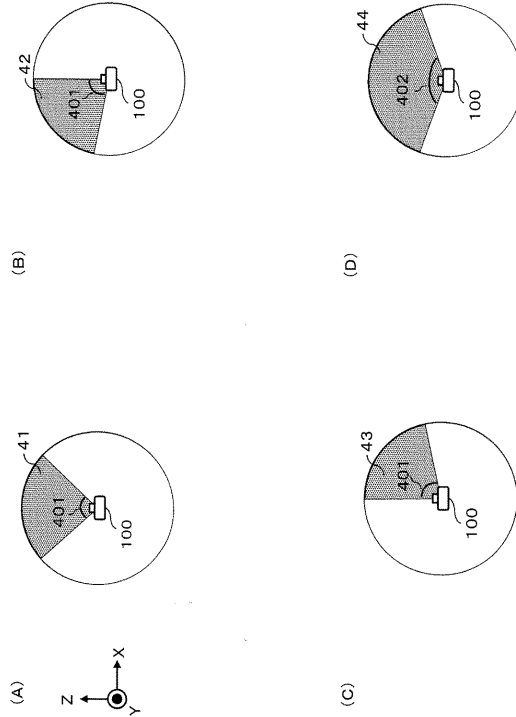
【図2】



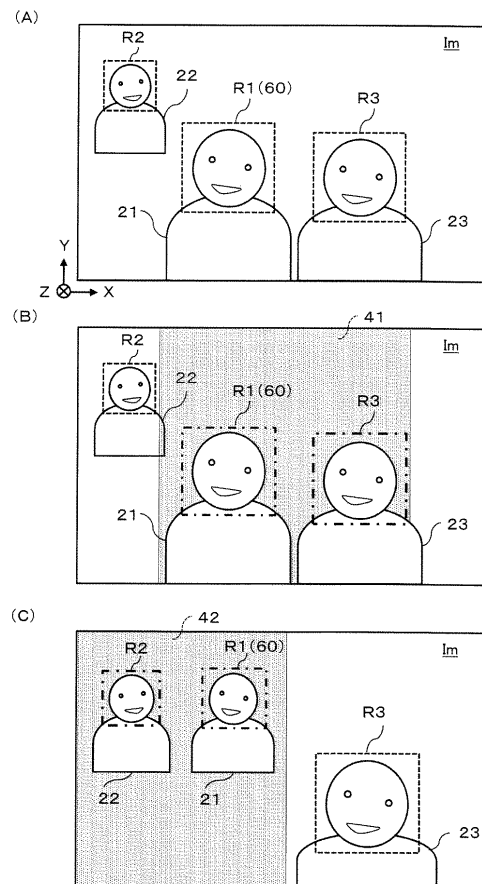
【図4】



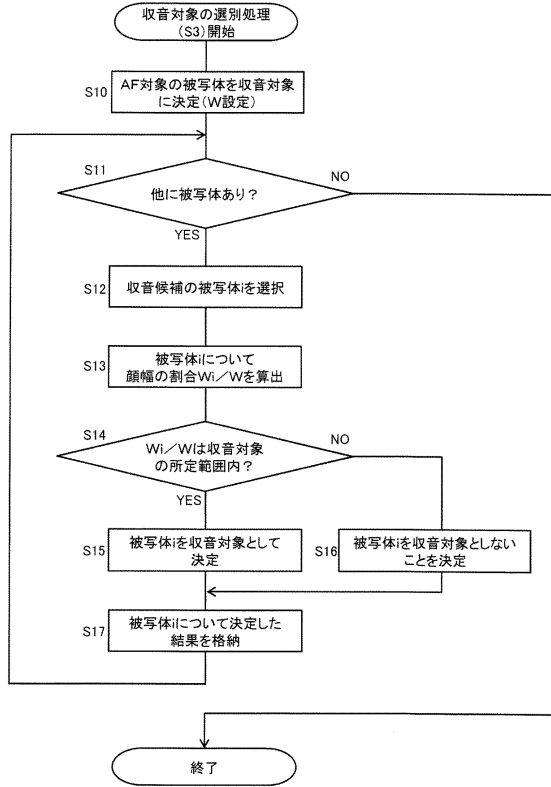
【図3】



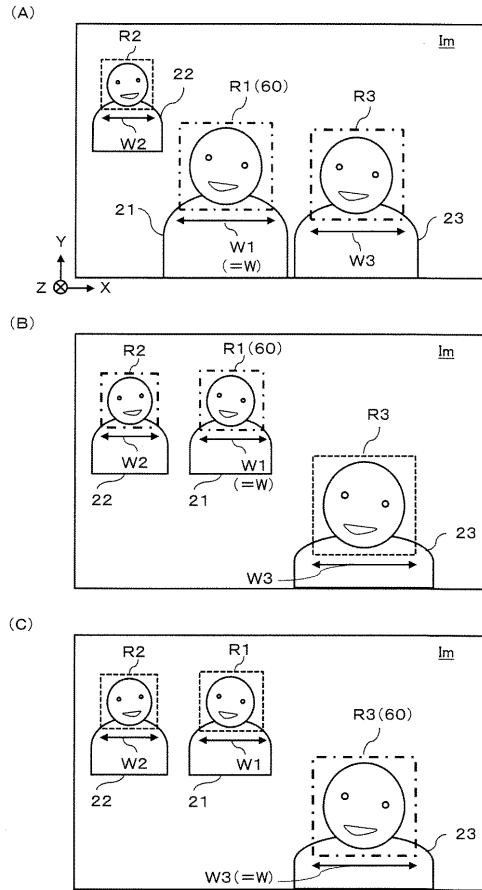
【図5】



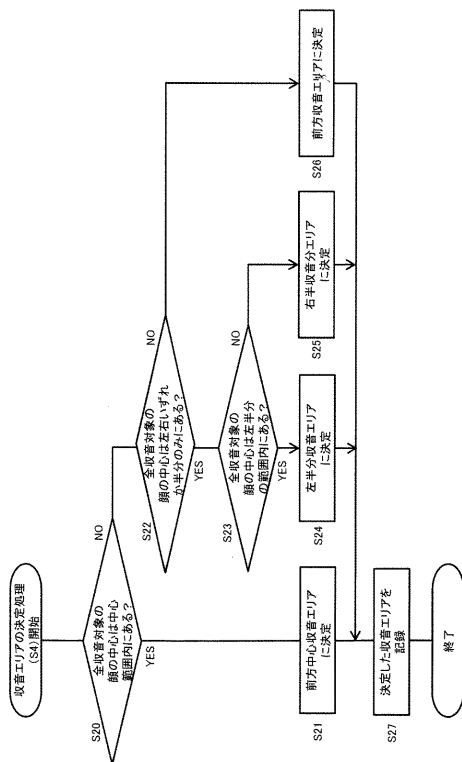
【図6】



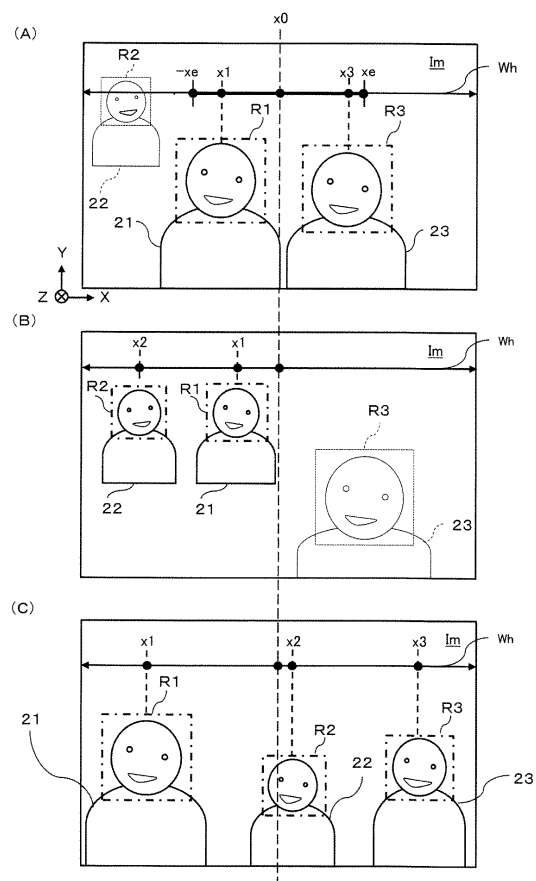
【図7】



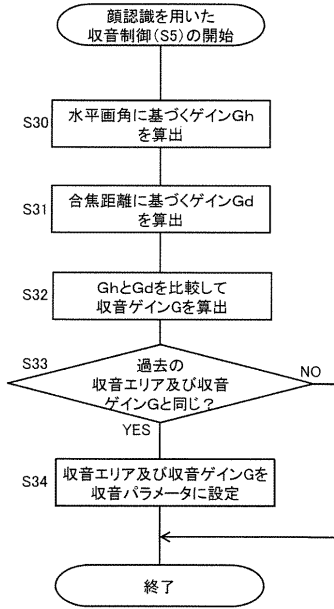
【図8】



【図9】



【図10】



【図11】

(A)

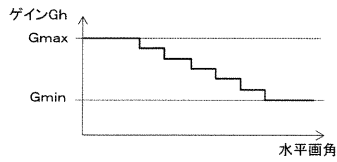
收音対象	R1, R3
收音エリア	前方中心收音エリア
水平面角	70°
合焦距離	1m

(B)

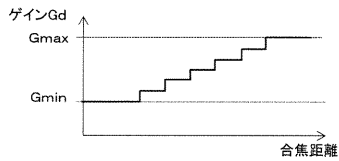
收音対象	R1, R2
收音エリア	左半分收音エリア
水平面角	70°
合焦距離	1.5m

【図12】

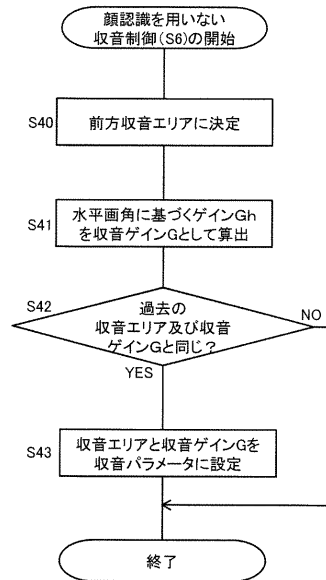
(A)



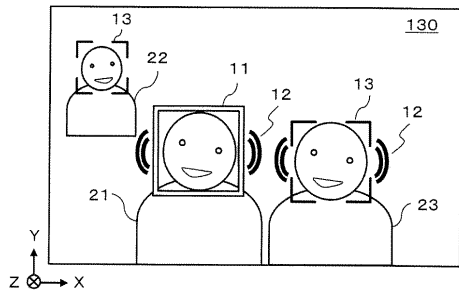
(B)



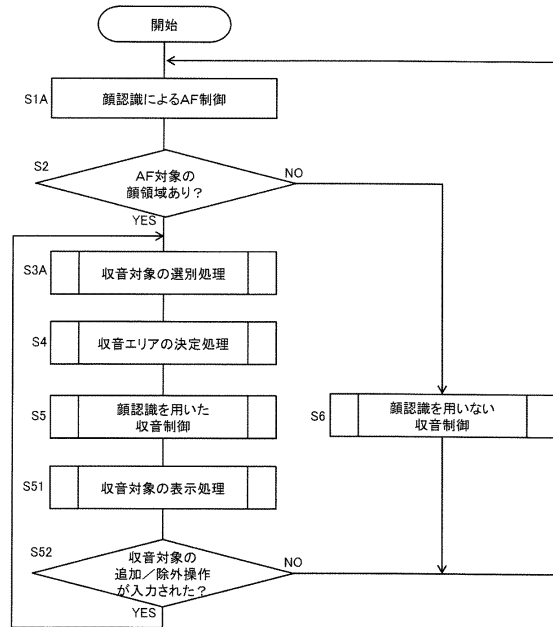
【図13】



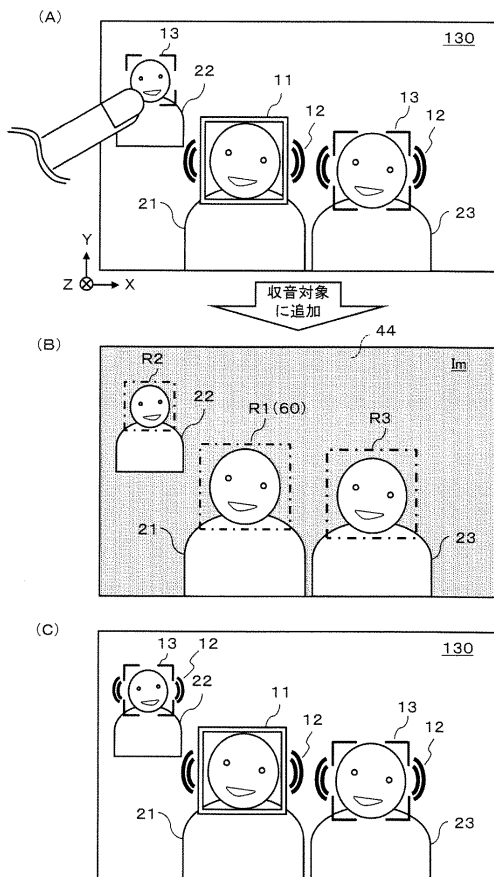
【図14】



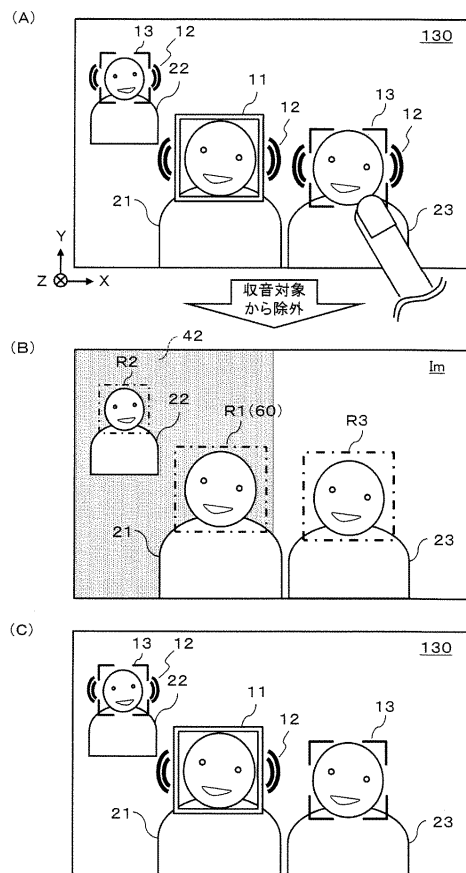
【図15】



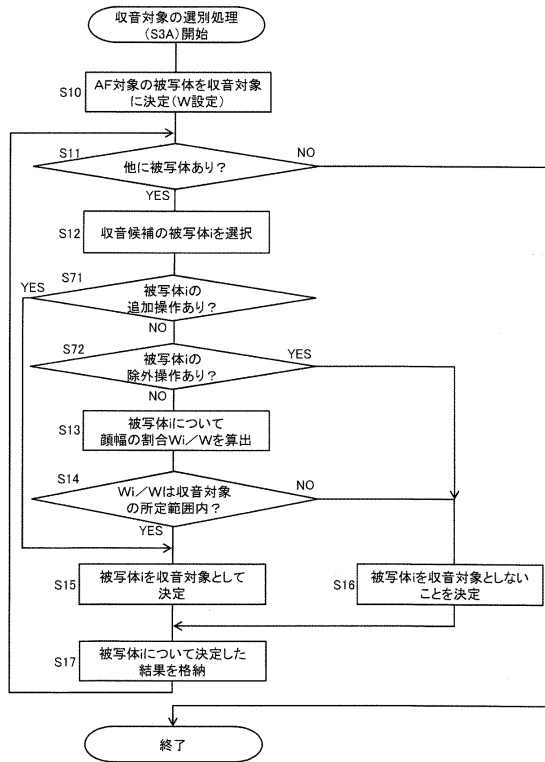
【図16】



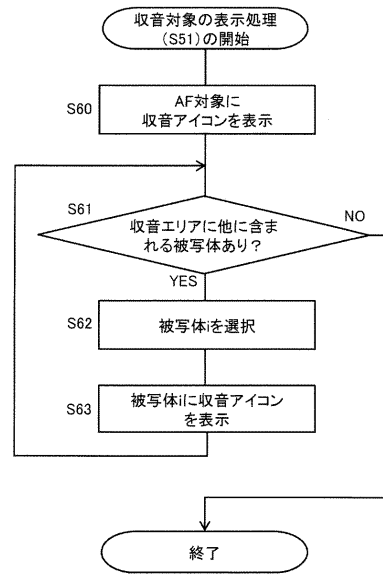
【図17】



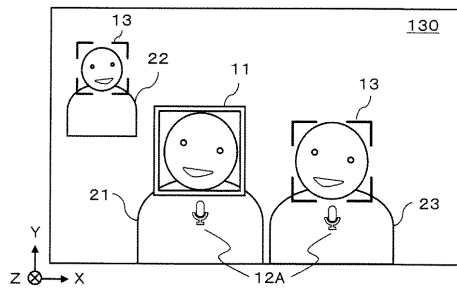
【図18】



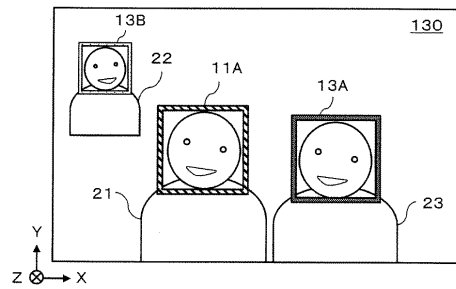
【図19】



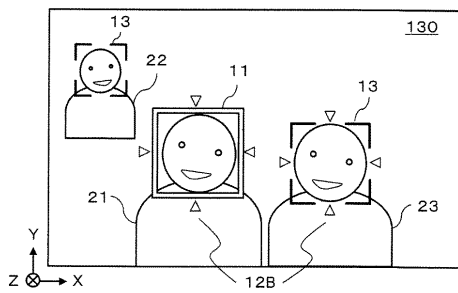
【図20】



【図22】



【図21】



フロントページの続き

(51) Int.Cl. F I
G 0 3 B 17/18 (2006.01) G 0 3 B 15/00 Q
G 0 2 B 7/28 N
G 0 3 B 13/36
G 0 3 B 17/18 Z

(56) 参考文献 特開 2 0 1 0 - 2 7 8 7 2 5 (J P , A)
特開 2 0 0 8 - 2 7 1 0 8 2 (J P , A)
特開 2 0 1 1 - 0 4 1 0 9 6 (J P , A)
特開 2 0 1 2 - 1 4 7 4 2 0 (J P , A)

(58) 調査した分野(Int.Cl. , DB名)
H 0 4 N 5 / 2 3 2
G 0 2 B 7 / 2 8
G 0 3 B 1 3 / 3 6
G 0 3 B 1 5 / 0 0
G 0 3 B 1 7 / 1 8
H 0 4 R 3 / 0 0