

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5801026号  
(P5801026)

(45) 発行日 平成27年10月28日(2015.10.28)

(24) 登録日 平成27年9月4日(2015.9.4)

|              |       |           |      |       |     |
|--------------|-------|-----------|------|-------|-----|
| (51) Int.Cl. |       | F I       |      |       |     |
| HO4S         | 5/02  | (2006.01) | HO4S | 5/02  | K   |
| HO4R         | 3/00  | (2006.01) | HO4R | 3/00  | 320 |
| HO4N         | 5/225 | (2006.01) | HO4N | 5/225 | F   |
|              |       |           | HO4S | 5/02  | Z   |

請求項の数 4 (全 21 頁)

|           |                               |           |                     |
|-----------|-------------------------------|-----------|---------------------|
| (21) 出願番号 | 特願2009-128793 (P2009-128793)  | (73) 特許権者 | 313003417           |
| (22) 出願日  | 平成21年5月28日(2009.5.28)         |           | 株式会社ザクティ            |
| (65) 公開番号 | 特開2010-278725 (P2010-278725A) |           | 大阪府大阪市北区大淀中一丁目1番88号 |
| (43) 公開日  | 平成22年12月9日(2010.12.9)         | (74) 代理人  | 100085501           |
| 審査請求日     | 平成24年4月20日(2012.4.20)         |           | 弁理士 佐野 静夫           |
|           |                               | (74) 代理人  | 100124132           |
|           |                               |           | 弁理士 渋谷 和俊           |
|           |                               | (74) 代理人  | 100128842           |
|           |                               |           | 弁理士 井上 温            |
|           |                               | (74) 代理人  | 100129562           |
|           |                               |           | 弁理士 山本 昌則           |
|           |                               | (74) 代理人  | 100141092           |
|           |                               |           | 弁理士 山本 英生           |

最終頁に続く

(54) 【発明の名称】 画像音響処理装置及び撮像装置

(57) 【特許請求の範囲】

【請求項1】

対になる入力画像信号及び入力音響信号を処理する画像音響処理装置において、前記入力画像信号が示す入力画像を解析する画像解析部と、前記画像解析部によって解析された入力画像中の注目被写体の大きさに応じて前記入力音響信号の指向特性を制御するための目標指向性特性を設定する目標指向特性設定部と、前記目標指向性特性に基づいて出力音響信号を生成する指向特性制御部と、前記目標指向性特性を現す画像を含む表示画像を生成する表示画像生成部と、を備え、前記指向特性は一方の方向から往来する音の集音レベルが他の方向から到来する音の集音レベルよりも相対的に大きくなる方向を示す強調方向及び集音レベルが相対的に大きくなる方向の範囲を示す強調幅を含むことを特徴とする画像音響処理装置。

10

【請求項2】

前記指向特性制御部は、入力画像中の注目被写体の大きさが、入力画像中の画角の第1所定サイズより大きい場合は、該注目被写体を含む第1範囲から到来する音を強調し、入力画像中の注目被写体の大きさが入力画像中の画角の第2所定サイズよりも小さい場合は、該注目被写体を含む前記第1範囲よりもせまい第2範囲から到来する音を強調して出力音声信号を生成することを特徴とする、請求項1記載の画像音響処理装置。

【請求項3】

対になる入力画像信号及び入力音響信号を処理する画像音響処理装置において、前記入力画像信号が示す入力画像を解析する画像解析部と、

20

前記画像解析部によって解析された入力画像中の注目被写体の装置本体への音声発生状況に応じて前記入力音響信号の指向特性を制御し、出力音響信号を生成する指向特性制御部と、

前記出力音響信号の状態を示す画像を含む表示画像を生成する表示画像生成部と、を備え、

前記指向特性は一方の方向から往来する音の集音レベルが他の方向から到来する音の集音レベルよりも相対的に大きくなる方向を示す強調方向及び集音レベルが相対的に大きくなる方向の範囲を示す強調幅を含むことを特徴とする画像音響処理装置。

【請求項 4】

対になる入力画像信号及び入力音響信号を処理する画像音響処理装置において、

10

前記入力画像信号が示す入力画像を解析する画像解析部と、

前記画像解析部によって解析された入力画像中の注目被写体の数に応じて前記入力音響信号の指向特性を制御し、出力音響信号を生成する指向特性制御部と、

前記出力音響信号の状態を示す画像を含む表示画像を生成する表示画像生成部と、を備え

前記指向特性は一方の方向から往来する音の集音レベルが他の方向から到来する音の集音レベルよりも相対的に大きくなる方向を示す強調方向及び集音レベルが相対的に大きくなる方向の範囲を示す強調幅を含むことを特徴とする画像音響処理装置。

【発明の詳細な説明】

20

【技術分野】

【0001】

本発明は、入力される画像信号と、当該画像信号と対になる音響信号と、に所定の処理を施して出力する画像音響処理装置や、当該画像音響処理装置を備えた撮像装置に関する。

【背景技術】

【0002】

撮像及び集音により画像信号及び音響信号を生成して記録するデジタルビデオカメラに代表される撮像装置が、広く普及している。このような撮像装置の中には、所定の方向から到来する音を強調した（指向特性を制御した）音響信号を生成して記録するものがある。

30

【0003】

例えば特許文献 1 では、マイクの指向特性を表現する画像をモニタに表示する撮像装置が提案されている。また、特許文献 2 では、音量や音響信号の指向特性を表現したパターンを撮像中の画像に重畳させてモニタに表示する撮像装置が提案されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開平 6 - 2 2 5 3 8 2 号公報

【特許文献 2】特開 2 0 0 6 - 2 6 1 9 0 0 号公報

40

【発明の概要】

【発明が解決しようとする課題】

【0005】

特許文献 1 及び特許文献 2 で提案される撮像装置では、マイクまたは音響信号の指向特性や音響信号の音量がモニタなどに表示されるため、撮影者はこれらの表示を確認することによって音響信号の指向特性や音量を認識することが可能となる。しかしながら、これらの表示によって撮影者が音響信号の指向特性を認識できたとしても、意図する音響信号を得るための指向特性の制御方法の設定や調整が困難となったり、そのための操作が煩雑となったりするため、問題となる。

【0006】

50

また、特許文献2で提案される撮像装置では、画角内の物体が発する音の音量を表示することは可能である。しかしながら、撮影者などの画角外の物体が発する音の音量は表示することができないため、意図する音響信号を得るためにどのように対応すべきかを撮影者が判断することが困難となり、問題となる。

【0007】

そこで本発明は、ユーザが意図する音響信号を容易かつ精度良く生成することを可能とする画像音響処理装置及び撮像装置を提供することを目的とする。

【課題を解決するための手段】

【0008】

上記目的を達成するために本発明の画像音響処理装置は、対になる入力画像信号及び入力音響信号を処理する画像音響処理装置において、前記入力画像信号が示す入力画像を解析する画像解析部と、前記画像解析部の解析結果に基づいて前記入力音響信号の指向特性を制御し、出力音響信号を生成する指向特性制御部と、前記出力音響信号の状態を示す画像を含む表示画像を生成する表示画像生成部と、を備えることを特徴とする。

10

【0009】

なお、出力音響信号には、以下の実施形態において説明する音量検出用出力音響信号が含まれ得る。同様に、指向特性制御部には、音量検出用指向特性制御部が含まれ得る。

【0010】

また、上記構成の画像音響処理装置において、前記画像解析部が、前記入力画像中から注目被写体を検出し、前記指向特性制御部が、前記画像解析部による前記注目被写体の検出結果に基づいて前記入力音響信号の指向特性を制御して前記出力音響信号を生成し、前記表示画像生成部が、前記出力音響信号の指向特性を示す画像を前記入力画像に重畳させた前記表示画像を生成することとしても構わない。

20

【0011】

このように構成すると、注目被写体に対応した制御方法によって入力音響信号の指向特性を制御することが可能となる。そのため、入力音響信号中の注目被写体が発する音を、容易にユーザが意図するものに制御することが可能となる。なお、以下の実施形態では、出力音響信号の指向特性の一例として、入力音響信号の制御後の指向特性である目標指向特性を挙げて説明している。

【0012】

また、上記構成の画像音響処理装置において、前記出力音響信号の強調された音の到来方向を示す画像を入力画像に重畳して表示画像を生成することとしても構わない。

30

【0013】

このように構成すると、出力音響信号の具体的な指向特性を示す画像を、入力画像に対比させて示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、出力音響信号の指向特性を具体的かつ容易に認識することが可能となる。なお、以下の実施形態では、出力音響信号の強調された音の到来方向を示す画像の一例として軸(図6及び図8参照)を挙げて説明している。

【0014】

また、上記構成の画像音響処理装置において、前記出力音響信号の音量を検出する音量検出部をさらに備え、前記画像解析部が、前記入力画像中から注目被写体を検出し、前記指向特性制御部が、前記入力音響信号中の前記注目被写体が存在する方向以外から到来する音を抑制して前記出力音響信号を生成し、前記表示画像生成部が、前記音量検出部により検出される前記出力音響信号の音量を示す画像を、前記入力画像に重畳させた前記表示画像を生成することとしても構わない。

40

【0015】

このように構成すると、注目被写体が発する音の音量を示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、注目被写体が発する音の状態を具体的に認識することが可能となる。

【0016】

50

また、上記構成の画像音響処理装置において、前記表示画像生成部が、前記出力音響信号の音量を示す画像を、前記入力画像中の前記注目被写体の近傍となる位置に重畳させた前記表示画像を生成することとしても構わない。

【0017】

このように構成すると、音量を注目被写体に関連付けて示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、表示される音量がどの注目被写体が発した音の音量であるかを容易に認識することが可能となる。

【0018】

また、上記構成の画像音響処理装置において、前記表示画像生成部が、前記入力画像中の前記注目被写体の位置を示す画像を前記入力画像に重畳させて前記表示画像を生成することとしても構わない。

10

【0019】

このように構成すると、注目被写体の入力画像中の位置を明確に示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、注目被写体の場所を容易に認識することが可能となる。さらに、注目被写体と出力音響信号の状態との関係を容易に認識することが可能となる。

【0020】

また、上記構成の画像音響処理装置において、前記出力音響信号の音量を検出する音量検出部と、前記入力画像の画角外の画角外音源が存在する方向を検出する音源方向検出部と、をさらに備え、前記指向特性制御部が、前記入力音響信号中の前記画角外音源が存在する方向以外から到来する音を抑制して前記出力音響信号を生成し、前記表示画像生成部が、前記音量検出部により検出される前記出力音響信号の音量を示す画像を、前記入力画像に重畳させた前記表示画像を生成することとしても構わない。

20

【0021】

このように構成すると、画角外音源が発する音の音量を示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、画角外音源が発する音の状態を具体的に認識することが可能となる。なお、音源方向検出部が、例えば入力画像や入力音響信号を解析することによって、画角外音源が存在する方向を検出することとしても構わない。

【0022】

30

また、上記構成の画像音響処理装置において、前記表示画像生成部が、前記出力音響信号の音量を示す画像を、前記入力画像の端部に重畳させた前記表示画像を生成することとしても構わない。

【0023】

このように構成すると、音量を示す画像を最も画角外に近い位置に表示した表示画像を生成することが可能となる。そのため、表示画像を確認したユーザが、表示される音量が画角外音源の発した音の音量であることを容易に認識することが可能となる。

【0024】

また、本発明の撮像装置は、上記のいずれかに記載の画像音響処理装置と、撮像により前記入力画像信号を生成する撮像部と、集音により前記入力音響信号を生成する集音部と、前記表示画像を表示する表示部と、備えることとしても構わない。

40

【0025】

また、上記構成の撮像装置において、前記指向特性制御部による前記入力音響信号の指向特性の制御方法の変更を指示する信号を入力する操作部をさらに備えることとしても構わない。

【0026】

このように構成すると、表示画像を確認した撮影者が、容易に入力音響信号の指向特性の制御方法を変更して意図する出力音響信号を生成することが可能となる。

【発明の効果】

【0027】

50

本発明によると、入力画像の解析結果に基づいた制御方法で入力音響信号の指向特性の制御を行うとともに、得られる出力音響信号の状態を示した画像を含む表示画像が生成されることとなる。即ち、入力画像に連動した入力音響信号の指向特性の制御を行うことにより、入力音響信号の指向特性の制御を容易に行うことが可能となる。さらに、得られる出力音響信号の状態が表示画像に含まれるため、表示画像を確認したユーザが、意図した出力音響信号が得られているか否かを確認することが可能となる。そして、出力音響信号の状態が意図するものと異なる場合には、入力音響信号の指向特性の制御方法を変更するなどの対応を講じることが可能となる。したがって、ユーザが意図する出力音響信号を容易かつ精度良く生成することが可能となる。

【図面の簡単な説明】

10

【0028】

【図1】は、本発明の実施形態における撮像装置の構成について示すブロック図である。

【図2】は、第1実施例の画像音響処理部の構成について示すブロック図である。

【図3】は、第1実施例の画像音響処理部における指向特性制御部の構成例について示すブロック図である。

【図4】は、第1実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。

【図5】は、指向特性画像の種々の例について示す図である。

【図6】は、第1実施例の画像音響処理部における表示画像生成部によって生成される表示画像の別例について示す図である。

20

【図7】は、第2実施例の画像音響処理部の構成について示すブロック図である。

【図8】は、第2実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。

【図9】は、第3実施例の画像音響処理部の構成について示すブロック図である。

【図10】は、第3実施例の画像音響処理部における音量検出用指向特性制御部の構成例について示すブロック図である。

【図11】は、第3実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。

【図12】は、音量検出結果画像の別例について示す図である。

【図13】は、第3実施例の画像音響処理部における表示画像生成部によって生成される表示画像の別例について示す図である。

30

【発明を実施するための形態】

【0029】

本発明の実施形態について、以下図面を参照して説明する。最初に、本発明における撮像装置の一例について説明する。

【0030】

<< 撮像装置 >>

まず、撮像装置の構成について、図1を参照して説明する。図1は、本発明の実施形態における撮像装置の構成について示すブロック図である。

【0031】

40

図1に示すように、撮像装置1は、入射される光学像を電気信号に変換するCCD (Charge Coupled Device) またはCMOS (Complimentary Metal Oxide Semiconductor) センサなどの固体撮像素子から成るイメージセンサ2と、被写体の光学像をイメージセンサ2に結像させるとともに光量などの調整を行うレンズ部3と、を備える。レンズ部3とイメージセンサ2とで撮像部が構成され、この撮像部によって画像信号が生成される。なお、レンズ部3は、ズームレンズやフォーカスレンズなどの各種レンズ(不図示)や、イメージセンサ2に入力される光量を調整する絞り(不図示)などを備える。

【0032】

さらに、撮像装置1は、イメージセンサ2から出力されるアナログ信号である画像信号をデジタル信号に変換するとともにゲインの調整を行うAFE (Analog Front End) 4と

50

、入力される音を電気信号に変換する集音部 5 と、集音部 5 から出力されるアナログ信号である音響信号をデジタル信号に変換する A D C (Analog to Digital Converter) 6 と、A D C 6 から出力される音響信号に対して各種音響処理を施して出力する音響処理部 7 と、A F E 4 から出力される画像信号に対して各種画像処理を施して出力する画像処理部 8 と、画像処理部 8 から出力される画像信号と音響処理部 7 から出力される音響信号とに対して M P E G (Moving Picture Experts Group) 圧縮方式などの動画用の圧縮符号化処理を施す圧縮処理部 9 と、圧縮処理部 9 で圧縮符号化された圧縮符号化信号を記録する外部メモリ 11 と、画像信号を外部メモリ 11 に記録したり読み出したりするドライバ部 10 と、ドライバ部 10 において外部メモリ 11 から読み出した圧縮符号化信号を伸長して復号する伸長処理部 12 と、を備える。

10

**【 0 0 3 3 】**

また、撮像装置 1 は、伸長処理部 12 で復号された画像信号をモニタなどの表示装置 (不図示) で表示可能な形式の信号に変換する画像信号出力回路部 13 と、伸長処理部 12 で復号された音響信号をスピーカなどの出力装置 (不図示) で出力可能な形式の信号に変換する音響信号出力回路部 14 と、を備える。

**【 0 0 3 4 】**

また、撮像装置 1 は、撮像装置 1 内全体の動作を制御する C P U (Central Processing Unit) 15 と、各処理を行うための各プログラムを記憶するとともにプログラム実行時の信号の一時保管を行うメモリ 16 と、撮像を開始するボタンや各種設定の決定を行うボタンなどの撮影者からの指示が入力される操作部 17 と、各部の動作タイミングを一致させるためのタイミング制御信号を出力するタイミングジェネレータ (T G) 部 18 と、C P U 15 と各部との間で信号のやりとりを行うためのバス 19 と、メモリ 16 と各部との間で信号のやりとりを行うためのバス 20 と、を備える。

20

**【 0 0 3 5 】**

なお、外部メモリ 11 は画像信号や音響信号を記録することができればどのようなものでも構わない。例えば、S D (Secure Digital) カードのような半導体メモリ、D V D などの光ディスク、ハードディスクなどの磁気ディスクなどをこの外部メモリ 11 として使用することができる。また、外部メモリ 11 を撮像装置 1 から着脱自在としても構わない。

**【 0 0 3 6 】**

次に、撮像装置 1 の基本動作について図 1 を参照して説明する。まず、撮像装置 1 は、レンズ部 3 より入射される光をイメージセンサ 2 において光電変換することによって、電気信号である画像信号を生成する。イメージセンサ 2 は、T G 部 18 から入力されるタイミング制御信号に同期して、所定のフレーム周期 (例えば、1 / 30 秒) で順次 A F E 4 に画像信号を出力する。そして、A F E 4 によってアナログ信号からデジタル信号へと変換された画像信号は、画像処理部 8 に入力される。画像処理部 8 では、画像信号が Y U V を用いた信号に変換されるとともに、階調補正や輪郭強調等の各種画像処理が施される。また、メモリ 16 はフレームメモリとして動作し、画像処理部 8 が処理を行なう際に画像信号を一時的に保持する。

30

**【 0 0 3 7 】**

また、集音部 5 は、音を集音して電気信号である音響信号に変換し、出力する。集音部 5 から出力される音響信号は A D C 6 に入力されて、アナログ信号からデジタル信号へと変換される。さらに、A D C 6 によってデジタル信号へと変換された音響信号は音響処理部 7 に入力され、ノイズ除去などの各種音響処理が施される。また、音響処理部 7 は、音響信号を処理することによりその指向特性を制御する。なお、指向特性やその制御方法の詳細については、後述する。

40

**【 0 0 3 8 】**

画像処理部 8 から出力される画像信号と、音響処理部 7 から出力される音響信号と、はともに圧縮処理部 9 に入力され、圧縮処理部 9 において所定の圧縮方式で圧縮される。このとき、画像信号と音響信号とは時間的に関連付けられて (対になって) おり、再生時に

50

画像と音とがずれないように構成される。そして、圧縮された画像信号及び音響信号はドライバ部10を介して外部メモリ11に記録される。

【0039】

外部メモリ11に記録された圧縮後の画像信号及び音響信号は、操作部17を介して入力される撮影者の再生指示に基づいて伸長処理部12に読み出される。伸長処理部12は、再生するために読み出される圧縮された画像信号及び音響信号を伸長し、この再生用の画像信号を画像信号出力回路部13、再生用の音響信号を音響信号出力回路部14にそれぞれ出力する。そして、画像信号出力回路部13が、再生用の画像信号を表示装置で表示可能な形式の信号に変換するとともに、音響信号出力回路部14が、再生用の音響信号をスピーカで出力可能な形式の信号に変換して、それぞれ出力する。これにより、再生用の画像が表示装置で表示され、再生用の音がスピーカから出力される。

10

【0040】

また、本実施形態の撮像装置1は、撮像した画像の記録を開始する前や、動画の記録時などに、撮像した画像を表示装置に表示する。このとき、画像処理部8は、表示用の画像信号を生成するとともに、バス20を介して画像信号出力回路部13に出力する。そして、画像信号出力回路部13が、表示用の画像信号を表示装置で表示可能な形式の信号に変換して、出力する。

【0041】

撮影者は、表示装置に表示される画像を確認することで、これから記録するまたは現在記録している画像の画角を認識することができる。さらに、このとき表示装置に表示される画像には、音響処理部7によって制御される音響信号の状態が重畳して表示される。なお、表示装置に表示される画像やその生成方法の詳細については、後述する。

20

【0042】

なお、表示装置やスピーカは、撮像装置1と一体となっているものでも構わないし、別体となっており、撮像装置1に備えられる端子とケーブル等を用いて接続されるようなものでも構わない。ただし、表示用の画像信号を表示する表示装置は、撮像装置1と一体となっているものであると好ましい。以下では、表示装置が、撮像装置1と一体になっているモニタである場合について説明する。

【0043】

また、集音部5が、デジタル信号の音響信号を出力するデジタルマイクを備えるものとして、ADC6を備えない構成としても構わない。

30

【0044】

<画像音響処理部>

以下、上述した表示画像を生成する画像処理部8及び音響処理部7の要部(以下、画像音響処理部とする)の構成及び動作について、図面を参照して説明する。なお、以下の説明では、上述した表示用の画像信号を「表示画像信号」、表示画像信号が示す画像を「表示画像」と呼ぶ。また、撮像して得られる画像信号であり表示用の画像信号の基になる画像信号を「入力画像信号」、入力画像信号が示す画像を「入力画像」と呼ぶ。また、入力画像信号の生成時(入力画像の撮像時)に集音されて得られる音響信号(即ち、入力画像信号と対になる音響信号)を「入力音響信号」と呼び、入力音響信号の指向特性を制御して生成される音響信号を「出力音響信号」と呼ぶ。

40

【0045】

また、指向特性とは、各方向から到来する音の集音レベル(集音して得られる音響信号のレベル)の差異を示すものとし、強調方向や強調幅を用いて表現することができる。強調方向とは、当該方向から到来する音の集音レベルが他の方向から到来する音の集音レベルよりも相対的に大きくなる方向を示すものである。また、強調幅とは、集音レベルが相対的に大きくなる方向の範囲を示すものである。強調幅が広いほど広範囲から到来する音が強調されて集音され、強調幅が狭いほど狭範囲から到来する音のみが限定的に強調されて集音される。なお、強調方向は一つに限られず、同時に複数存在する場合も生じ得る。

【0046】

50

また、ある方向から到来する音を強調するとは、ある方向から到来する音のレベルを絶対的に大きくする場合だけに限られず、ある方向以外から到来する音を抑制することで、ある方向から到来する音のレベルを相対的に大きくする場合をも含むものとする。

【 0 0 4 7 】

[ 第 1 実施例 ]

画像音響処理部の第 1 実施例について図面を参照して説明する。図 2 は、第 1 実施例の画像音響処理部の構成について示すブロック図である。図 2 に示すように、画像音響処理部 3 0 a は、入力画像信号が示す入力画像を解析して画像解析情報を生成する画像解析部 8 1 と、画像解析部 8 1 で生成される画像解析情報に基づいて入力音響信号の指向特性を制御して出力音響信号を生成するとともに入力音響信号の制御後の指向特性（即ち、出力音響信号の指向特性、以下、目標指向特性とする）を設定して目標指向特性情報を生成する指向特性制御部 7 1 と、指向特性制御部 7 1 で生成される目標指向特性情報に基づいた画像を入力画像に重畳させた表示画像となる表示画像信号を生成する表示画像生成部 8 2 と、を備える。また指向特性制御部 7 1 は、表示画像を確認した撮影者から操作部 1 7 を介して入力される指向特性制御指示に基づいて、目標指向特性の設定方法の変更を行う。

10

【 0 0 4 8 】

なお、画像解析部 8 1 及び表示画像生成部 8 2 が、図 1 に示した画像処理部 8 に備えられ、指向特性制御部 7 1 が、図 1 に示した音響処理部 7 に備えられることとしても構わない。

【 0 0 4 9 】

以下、本実施例の画像音響処理部 3 0 a の各部の構成及び動作について説明する。

20

【 0 0 5 0 】

（画像解析部）

画像解析部 8 1 は、例えば、順次入力される入力画像中から注目被写体を順次検出する検出処理（追尾処理）を行い、検出された注目被写体の入力画像中の位置や大きさなどを示す情報を画像解析情報として順次生成し、出力する。検出すべき注目被写体は、検出処理の開始時に撮影者がカーソルキーやタッチパネルなどから成る操作部 1 7 を操作することによって設定されたり、プログラムなどによって自動的に設定されたりする。このとき、例えば設定された注目被写体の形状や色などの特徴が認識され、入力画像内から当該特徴を示す部分が検出されることで、注目被写体の検出が行われる。

30

【 0 0 5 1 】

具体的に例えば、検出すべき注目被写体を不特定の人顔としても構わないし（顔検出）、予め記憶している特定の人顔としても構わない（顔認識）。さらに、検出した顔を有する人の一部（例えば、検出した顔の眉間から口に向かう方向に存在する領域である胴体領域）の色を認識するとともに、入力画像中から当該色の部分を検出することで、注目被写体の検出を行うこととしても構わない。

【 0 0 5 2 】

また、顔検出を行う場合、周知の種々の技術を適用することが可能である。例えば、Adaboost (Yoav Freund, Robert E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", European Conference on Computational Learning Theory, September 20, 1995.) を利用して大量の教師サンプル（顔及び非顔のサンプル画像）から作成した重みテーブルと、入力画像と、を比較することで顔検出を行うことが可能である。

40

【 0 0 5 3 】

以下では説明の具体化のため、画像解析部 8 1 が、人の顔を注目被写体として検出するとともに、入力画像中の注目被写体（人の顔）の位置や大きさを示す情報を含む画像解析情報を生成し出力するものとする。

【 0 0 5 4 】

（指向特性制御部）

指向特性制御部 7 1 は、画像解析部 8 1 から出力される画像解析情報を取得し、注目被

50



写体の位置や大きさ、有無などに基づいて目標指向特性を設定し、当該目標指向特性が実現されるように入力音響信号の指向特性を制御する。また、撮影者から操作部 17 を介して指向特性制御指示が入力されれば、当該指示に基づいて目標指向特性の設定方法を変更する。また、入力音響信号の指向特性の制御は、例えば入力音響信号のレベルを音の到来方向ごとに制御することによって行う。

【 0 0 5 5 】

集音部 5 が、複数の指向性マイク（特定方向から到来する音を強調して集音するマイク）を備えるものである場合、入力音響信号は、強調された方向が異なる複数のチャンネルの信号を備えるものとなる。そのため、それぞれのチャンネルの信号のレベルを制御することにより、指向特性を制御することが可能となる。

10

【 0 0 5 6 】

また、集音部 5 が、複数の無指向性マイク（特定方向から到来する音を強調することなく、一律に集音するマイク）を備えるものである場合、入力音響信号は、強調された方向がない複数のチャンネルの信号を備えるものとなる。この場合、例えばそれぞれのチャンネルの信号の位相差を算出して音の到来方向を求め、音の到来方向に基づいて信号のレベルを制御することで指向特性を制御することが可能となる。なお、この構成の一例について、以下図面を参照して説明する。

【 0 0 5 7 】

図 3 は、第 1 実施例の画像音響処理部における指向特性制御部の構成例について示すブロック図である。なお、説明の具体化のため、図 3 では L c h 及び R c h の二つのチャンネルの信号を備える入力音響信号の指向特性を制御する指向特性制御部 7 1 について示すこととする。

20

【 0 0 5 8 】

図 3 に示すように、指向特性制御部 7 1 は、入力音響信号の L c h の信号を高速フーリエ変換（Fast Fourier Transform、以下 F F T とする）して出力する F F T 部 7 1 1 L と、入力音響信号の R c h の信号を F F T して出力する F F T 部 7 1 1 R と、F F T 部 7 1 1 L , 7 1 1 R から出力される L c h 及び R c h のそれぞれの信号を所定の周波数の帯域毎に比較することで各帯域の位相差を算出して出力する位相差算出部 7 1 2 と、画像解析情報や指向特性制御指示に基づいて目標指向特性を設定して目標指向特性情報を出力する目標指向特性設定部 7 1 3 と、目標指向特性設定部 7 1 3 から出力される目標指向特性情報に示される目標指向特性が実現されるように位相差算出部 7 1 2 から出力される各帯域の位相差に基づいて各チャンネルの各帯域のレベルの制御量を設定する帯域別制御量設定部 7 1 4 と、F F T 部 7 1 1 L から出力される L c h の信号の各帯域のレベルを帯域別制御量設定部 7 1 4 で設定された制御量に応じて制御して出力する帯域別レベル制御部 7 1 5 L と、F F T 部 7 1 1 R から出力される R c h の信号の各帯域のレベルを帯域別制御量設定部 7 1 4 で設定された制御量に応じて制御して出力する帯域別レベル制御部 7 1 5 R と、帯域別レベル制御部 7 1 5 L から出力される L c h の信号を逆高速フーリエ変換（Inverse Fast Fourier Transform、以下 I F F T とする）して L c h の出力音響信号として出力する I F F T 部 7 1 6 L と、帯域別レベル制御部 7 1 5 R から出力される R c h の信号を I F F T して R c h の出力音響信号として出力する I F F T 部 7 1 6 R と、を備える。

30

40

【 0 0 5 9 】

F F T 部 7 1 1 L , 7 1 1 R は、入力音響信号の L c h 及び R c h の信号をそれぞれ F F T し、時間軸の信号から周波数軸の信号へとそれぞれ変換する。位相差算出部 7 1 2 は、F F T 部 7 1 1 L , 7 1 1 R から出力される L c h 及び R c h の信号を、それぞれの周波数の帯域毎に比較する（例えば、帯域毎に L c h 及び R c h の信号の相関を求めるなど）。これにより、L c h 及び R c h の信号の位相差（音源と複数の無指向性マイクのそれぞれとの距離差、到達までの時間差とも解釈し得る）を算出する。

【 0 0 6 0 】

目標指向特性設定部 7 1 3 は、画像解析情報に基づいて目標指向特性の設定を行うとと

50

もに、指向特性制御指示が入力されればこれに基づいて目標指向特性の設定方法の変更を行う。具体的に例えば、画像解析情報が示す注目被写体が存在する方向を強調方向とし、強調幅をその注目被写体の大きさに対応した広さとする設定方法によって、目標指向特性を設定する。

【 0 0 6 1 】

また、この設定方法で設定された目標指向特性が撮影者の意図するものと異なる場合、撮影者は操作部 1 7 を介して指向特性制御指示を入力することで、目標指向特性の設定方法を変更することができる。具体的に例えば、複数の注目被写体が検出されている場合に、特定の注目被写体以外の注目被写体が存在する方向が強調方向とならないようにしたり、強調幅を広くまたは狭くしたりするなどのように、目標指向特性の設定方法を変更することができる。そして、指向特性設定部 7 1 3 は、以上のようにして設定される目標指向特性を目標指向特性情報として出力する。

10

【 0 0 6 2 】

帯域別制御量設定部 7 1 4 は、位相差算出部 7 1 2 から出力される位相差に基づいて帯域毎に音の到来方向を確認するとともに、目標指向特性設定部 7 1 3 から出力される目標指向特性情報に基づいて目標指向特性の強調方向を確認する。そして、音の到来方向が強調方向に含まれる帯域のレベルが増大される、及び（または）音の到来方向が強調方向に含まれない帯域のレベルが抑制されるように、各帯域の制御量を設定する。

【 0 0 6 3 】

また、帯域別レベル制御部 7 1 5 L , 7 1 5 R は、帯域別制御量設定部 7 1 4 で設定された制御量に基づいて、L c h 及び R c h のそれぞれの信号のレベルを帯域毎に制御することで、入力音響信号の指向特性の制御を行う。そして、I F F T 部 7 1 6 L , 7 1 6 R が、帯域別レベル制御部 7 1 5 L , 7 1 5 R から出力される L c h 及び R c h のそれぞれの周波数軸の信号を I F F T して時間軸の信号に変換し、出力音響信号の L c h 及び R c h の信号をそれぞれ生成して出力する。

20

【 0 0 6 4 】

なお、上記の指向特性制御部 7 1 の構成は一例に過ぎず、他の構成を採用しても構わない。例えば、入力音響信号の R c h の信号をある時間遅延させて入力音響信号の L c h の信号に合成（例えば、加算または減算）することで出力音響信号の L c h の信号を生成し、入力音響信号の L c h の信号をある時間遅延させて入力音響信号の R c h の信号に合成することで出力音響信号の R c h の信号を生成する構成としても構わない。また、遅延させる時間を画像解析情報に基づいた可変の時間としても構わない。

30

【 0 0 6 5 】

（表示画像生成部）

表示画像生成部 8 2 は、入力される目標指向特性情報に示される目標指向特性を表現した画像を入力画像に重畳させることで、目標指向特性を視覚的に表現した表示画像を生成する。この表示画像の一例を、図 4 に示す。図 4 は、第 1 実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。

【 0 0 6 6 】

図 4 に示すように、表示画像 P 1 は、目標指向特性を模式的に表現した指向特性画像 S 1 が入力画像の隅（例えば右下隅）に重畳表示されて成る。また、本例の指向特性画像 S 1 は、マイクの模式図 S 1 1 と、設定された目標指向特性の状態を示す複数の弧 S 1 2 と、で構成される。

40

【 0 0 6 7 】

また、表示画像 P 1 は、注目被写体 T（人の顔）が画像解析部 8 1 において入力画像中から検出され、指向特性制御部 7 1 が、注目被写体 T が存在する方向から到来する音を強調する制御を行う場合について示している。この場合、例えば指向特性画像 S 1 が、マイクの模式図 S 1 1 の上部にのみ長い弧 S 1 2 が付された構成にすることで、被写体方向の広い範囲から到来する音を強調する（強調方向が被写体方向、強調幅が広い）目標指向特性が設定されたことを表現する。

50

## 【 0 0 6 8 】

上記の図 4 と同様の方法で目標指向特性を表現した指向特性画像の種々の例について、図 5 を参照して説明する。図 5 は、指向特性画像の種々の例について示す図である。

## 【 0 0 6 9 】

図 5 ( a ) は、図 4 に示した指向特性画像 S 1 と同様のものであり、被写体方向の広い範囲から到来する音を強調するように制御することを表現したものである。図 5 ( b ) は、マイクの模式図の上部にのみ短い弧が付された構成であり、被写体方向の狭い範囲から到来する音を強調するように制御する（強調方向が被写体方向、強調幅が狭い目標指向特性となる）ことを表現したものである。図 5 ( c ) は、マイクの模式図の左右に長い弧が付された構成であり、特定方向から到来する音を強調せず無指向にする（強調方向がない目標指向特性とする）場合を表現したものである。図 5 ( d ) は、マイクの模式図の上下に短い弧が付された構成であり、被写体方向及び撮影者方向から到来する音を強調するように制御する（強調方向が被写体方向及び撮影者方向である目標指向特性とする）ことを表現したものである。

10

## 【 0 0 7 0 】

例えば、入力画像中から検出された注目被写体 T の画角に占める割合が大きい場合に、図 5 ( a ) の指向特性画像が示すような被写体方向の広い範囲から到来する音が強調される目標指向特性が設定され、注目被写体 T の画角に占める割合が小さい場合に、図 5 ( b ) の指向特性画像が示すような被写体方向の狭い範囲から到来する音が強調される目標指向特性が設定されることとしても構わない。また例えば、入力画像中から注目被写体 T が検出されなかった場合に、図 5 ( c ) の指向特性画像が示すような無指向となる目標指向特性が設定されることとしても構わない。さらに例えば、入力画像中から検出された注目被写体 T が撮影者に向かって喋っていることが確認される場合（例えば、注目被写体 T の目線が撮影者方向であることや口が動いていることが確認される場合や、人の声が入力音響信号中に含まれていることが確認された場合など）に、注目被写体 T と撮影者とが会話をしていると推測し、図 5 ( d ) の指向特性画像が示すような被写体方向及び撮影者方向から到来する音が強調される目標指向特性が設定されることとしても構わない。

20

## 【 0 0 7 1 】

撮影者は、モニタに表示される表示画像 P 1 に含まれる指向特性画像 S 1 を確認することで、設定された目標指向特性を認識する。そして、撮影者が意図する目標指向特性と異なることを認識した場合、操作部 1 7 を介して指向特性制御指示を入力することで、目標指向特性の設定方法を変更する。

30

## 【 0 0 7 2 】

このように、入力画像の状態に応じて目標指向特性を設定することにより、撮影者が意図する出力音響信号を生成するための目標指向特性を、容易に設定することが可能となる。さらに、表示画像 P 1 中に指向特性画像 S 1 を表示して、設定された目標指向特性が意図するものであるか否かを撮影者に認識させるとともに、目標指向特性の設定方法を撮影者が変更し得る構成とすることにより、設定される目標指向特性を精度よく撮影者の意図するものにすることが可能となる。したがって、撮影者が意図する出力音響信号を容易かつ精度良く生成することが可能となる。

40

## 【 0 0 7 3 】

なお、目標指向特性を抽象的に表現する指向特性画像 S 1 を表示画像 P 1 中に表示する場合について説明したが、具体的に表現する指向特性画像を表示することとしても構わない。この指向特性画像について、図面を参照して説明する。図 6 は、第 1 実施例の画像音響処理部における表示画像生成部によって生成される表示画像の別例について示す図である。また、図 6 ( a ) 及び ( b ) は、撮影者から指向特性制御指示が入力される前後の表示画像 P 2 1 , P 2 2 を示したものであり、図 5 と同様に注目被写体 T が入力画像内から検出された場合について示すものである。

## 【 0 0 7 4 】

図 6 に示すように、本例の指向特性画像 S 2 は、マイクの模式図 S 2 1 と、強調方向及

50

び強調幅を示す軸 S 2 2 L , S 2 2 R とから成り、軸 S 2 2 L , S 2 2 R に挟まれる領域によって強調方向及び強調幅を表現するものである。図 6 ( a ) に示す表示画像 P 2 1 では、注目被写体 T を中心とした強調方向であり十分広い強調幅となる目標指向特性が設定された場合の指向特性画像 S 2 が表示されており、表示画像 P 2 1 を確認した撮影者が、強調幅を狭くしたいと考えた場合について説明する。

【 0 0 7 5 】

このような場合、上述のように撮影者は操作部 1 7 を介して指向特性制御指示を入力することにより、目標指向特性の設定方法を変更する。例えば、操作部 1 7 がモニタに設けられるタッチパネルなどで構成されている場合、撮影者は、モニタに表示される図 6 ( a ) の軸 S 2 2 L , S 2 2 R の少なくとも一方を選択して動かすなどして、軸 S 2 2 L , S 2 2 R の間隔を狭くする。これにより、強調幅を狭くする指向特性制御指示が、指向特性制御部 7 1 に入力されたことになる。

10

【 0 0 7 6 】

指向特性制御部 7 1 は、入力される指向特性制御指示に基づいて目標指向特性の設定方法を変更し、変更後の設定方法によって目標指向特性の設定を行う。図 6 ( b ) に示す表示画像 P 2 2 は、変更後の設定方法によって目標指向特性が設定された場合の指向特性画像 S 2 を示すものである。図 6 ( b ) の表示画像 P 2 2 では、軸 S 2 2 L , S 2 2 R の間隔が図 6 ( a ) の表示画像 P 2 1 よりも狭いものとなっている。

【 0 0 7 7 】

撮影者は、図 6 ( b ) の表示画像 P 2 2 中の指向特性画像 S 2 を確認することで、意図する目標指向特性が設定されているか否かを認識する。意図する目標指向特性が設定されていないならば、撮影者からさらなる指向特性制御指示が入力される。一方、意図する目標指向特性が設定されていれば、図 6 ( b ) に示す後も同様の設定方法で目標指向特性が設定される。即ち、注目被写体 T を中心とした強調方向であり、狭い強調幅となる目標指向特性が、その後の入力画像信号及び入力音響信号に対して順次設定されることとなる。

20

【 0 0 7 8 】

このように、目標指向特性を具体的に表現する指向特性画像 S 2 を表示画像 P 2 1 , P 2 2 中に表示することにより、設定されている目標指向特性や、指向特性制御指示の入力による目標指向特性の変化を、撮影者が具体的に認識することが可能となる。そのため、容易に目標指向特性を設定することが可能となる。また、この指向特性画像 S 2 を利用することで、撮影者が具体的な指向特性制御指示を入力することが可能となる。

30

【 0 0 7 9 】

[ 第 2 実施例 ]

画像音響処理部の第 2 実施例について図面を参照して説明する。図 7 は、第 2 実施例の画像音響処理部の構成について示すブロック図であり、第 1 実施例の構成について示した図 2 に相当するものである。なお、図 7 において図 2 と同様の構成となる部分については同様の符号を付し、その詳細な説明については省略する。

【 0 0 8 0 】

図 7 に示すように、画像音響処理部 3 0 b は、画像解析部 8 1 と、指向特性制御部 7 1 と、画像解析部 8 1 から出力される画像解析情報と指向特性制御部 7 1 から出力される目標指向特性情報とのそれぞれに基づいた画像を入力画像に重畳させた表示画像を生成して表示画像信号を出力する表示画像生成部 8 2 b と、を備える。

40

【 0 0 8 1 】

本実施例の表示画像生成部 8 2 b は、目標指向特性情報に基づいた画像（即ち、指向特性画像）だけでなく、画像解析情報に基づいた画像（以下、画像解析結果画像とする）をも入力画像に重畳させて表示画像を生成する点において、第 1 実施例と異なる。

【 0 0 8 2 】

本実施例の表示画像生成部 8 2 b が生成する表示画像の一例について、図面を参照して説明する。図 8 は、第 2 実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。なお、説明の具体化のため、本実施例の表示

50

画像生成部 8 2 b は、図 6 に示した指向特性画像（マイクの模式図及び軸を備える画像）と同様の指向特性画像を生成するものとする。また、入力画像中から二つの注目被写体 T 1 , T 2 が検出され、強調方向が注目被写体 T 1 , T 2 が存在するそれぞれの方向であり、強調幅が注目被写体 T 1 , T 2 のそれぞれの大きさに対応した広さとなるように目標指向特性が設定される場合を例に挙げ、以下説明する。

【 0 0 8 3 】

図 8 に示す表示画像 P 3 では、マイクの模式図 S 3 1 と、注目被写体 T 1 が存在する方向の強調方向及び強調幅を示す軸 S 3 2 L , S 3 2 R と、注目被写体 T 2 が存在する方向の強調方向及び強調幅を示す軸 S 3 3 L , S 3 3 R と、が指向特性画像 S 3 として表示されている。さらに、注目被写体 T 1 である人の顔を囲むように表示される顔枠画像 A 1 と、注目被写体 T 2 である人の顔を囲むように表示される顔枠画像 A 2 と、が画像解析結果画像として表示されている。

10

【 0 0 8 4 】

このように、表示画像 P 3 中に指向特性画像 S 3 だけでなく、画像解析結果画像を併せて表示することにより、表示画像 P 3 を確認した撮影者が、設定されている目標指向特性を容易に認識することが可能となる。特に、撮影者が、注目被写体 T 1 , T 2 と設定される目標指向特性との関係（即ち、目標指向特性の設定方法）を容易に認識することが可能となる。

【 0 0 8 5 】

なお、指向特性画像として、図 6 に示したような目標指向特性を具体的に表現するものを適用する場合を例示したが、抽象的に表示する指向特性画像を適用しても構わない。ただし、目標指向特性を具体的に表現する指向特性画像を適用すると、撮影者が、注目被写体と目標指向特性との関係や、目標指向特性の設定方法を容易に認識することが可能となるため、好ましい。

20

【 0 0 8 6 】

[ 第 3 実施例 ]

画像音響処理部の第 3 実施例について図面を参照して説明する。図 9 は、第 3 実施例の画像音響処理部の構成について示すブロック図であり、第 1 実施例の構成について示した図 2 に相当するものである。なお、図 9 において図 2 と同様の構成となる部分については同様の符号を付し、その詳細な説明については省略する。

30

【 0 0 8 7 】

図 9 に示すように、画像音響処理部 3 0 c は、画像解析部 8 1 と、画像解析情報及び指向特性制御指示に基づいて入力音響信号の指向特性を制御し音量検出用出力音響信号を生成する音量検出用指向特性制御部 7 1 c と、音量検出用指向特性制御部 7 1 c から出力される音量検出用出力音響信号の音量を検出して音量検出情報を出力する音量検出部 7 2 と、画像解析部 8 1 から出力される画像解析情報と音量検出部 7 2 から出力される音量検出情報とのそれぞれに基づいた画像を入力画像に重畳させた表示画像を生成して表示画像信号を出力する表示画像生成部 8 2 c と、指向特性制御部 7 1 と、入力される音量指定指示（詳細は後述）を指向特性制御指示に変換して指向特性制御部 7 1 に出力する指向特性制御指示変換部 7 3 と、を備える。

40

【 0 0 8 8 】

本実施例の画像音響処理部 3 0 c は、音量検出用指向特性制御部 7 1 c と、音量検出部 7 2 と、指向特性制御指示変換部 7 3 とを備える点において第 1 実施例と異なる。また、表示画像生成部 8 2 c による表示画像の生成方法についても、第 1 実施例と異なる。以下、音量検出用指向特性制御部 7 1 c 、音量検出部 7 2 、表示画像生成部 8 2 c 及び指向特性制御指示変換部 7 3 について図面を参照してそれぞれ説明する。

【 0 0 8 9 】

（音量検出用指向特性制御部）

図 1 0 は、第 3 実施例の画像音響処理部における音量検出用指向特性制御部の構成例について示すブロック図である。音量検出用指向特性制御部 7 1 c は、指向特性制御部 7 1

50

と同様に、入力音響信号の指向特性を制御して音量検出用出力音響信号を生成するものである。なお、音量検出用出力音響信号は出力音響信号の一種として解釈され得るものであり、音量検出用指向特性制御部 7 1 c は指向特性制御部 7 1 の一種として解釈され得るものである。また、以下では説明の具体化及び簡略化のため、図 1 0 に示す音量検出用指向特性制御部 7 1 c の構成が、図 3 に示した指向特性制御部 7 1 の構成と同様のものであることとし、同様の構成となる部分については同様の符号を付し、その詳細な説明については省略する。

#### 【 0 0 9 0 】

図 1 0 に示すように、本例の音量検出用指向特性制御部 7 1 c は、FFT部 7 1 1 L , 7 1 1 R と、位相差算出部 7 1 2 と、画像解析情報に基づいて音量検出方向を設定するとともに当該音量検出方向から到来する音を抽出するための音量検出用目標指向特性を設定して音量検出用目標指向特性を出力する音量検出用目標指向特性設定部 7 1 3 c と、帯域別制御量設定部 7 1 4 と、帯域別レベル制御部 7 1 5 L , 7 1 5 R と、L c h 及び R c h の音量検出用出力音響信号を出力するIFFT部 7 1 6 L , 7 1 6 R と、を備える。なお、音量検出用目標指向特性設定部 7 1 3 c 及び音量検出用目標指向特性情報は、図 3 の指向特性制御部 7 1 における目標指向特性設定部 7 1 3 及び目標指向特性情報にそれぞれ相当するものであり、これらの一種として解釈され得るものである。

10

#### 【 0 0 9 1 】

音量検出方向とは、例えば、画像解析情報によって示される注目被写体が存在する方向などであり、音源が存在し得る方向である。なお、音量検出方向は入力画像の画角内に限られるものではなく、画角外の方向（例えば、撮影者方向）も音量検出方向に含まれ得る。また、音量検出用の目標指向特性とは、音量検出方向以外から到来する音のレベルを抑制（例えば、略 0 となるまで抑制）するものである。

20

#### 【 0 0 9 2 】

音量検出用目標指向特性設定部 7 1 3 c は、設定した音量検出方向に対応する音量検出用目標指向特性を設定する。音量検出方向が複数設定される場合は、それぞれの音量検出方向に対応する音量検出用目標指向特性を、順次切り替えて設定する。

#### 【 0 0 9 3 】

なお、音量検出用出力音響信号及び出力音響信号においてそれぞれの音量検出方向から到来する音のレベルが略等しくなるように、音量検出用目標指向特性が目標指向特性に関連して設定されるように構成しても構わない。このように構成すると、後述する音量検出部 7 2 において検出される音量が、出力音響信号中の音量検出方向から到来する音の音量を示すものとなるため、好ましい。

30

#### 【 0 0 9 4 】

具体的には、図 9 に示すように、指向特性制御部 7 1 と音量検出用指向特性制御部 7 1 c とのそれぞれに、指向特性制御指示変換部 7 3 （詳細は後述）から出力される指向特性制御指示が入力されるように構成して、目標指向特性と音量検出用目標指向特性とが連動して制御されるように構成しても構わない。この場合、音量検出用目標指向特性設定部 7 1 3 c は、目標指向特性設定部 7 1 3 と同様に入力される指向特性制御指示に基づいて目標指向特性の設定方法を変更するが、上述のように音量検出方向以外から到来する音のレベルは抑制する。したがって、仮に出力音響信号の指向特性を変化させたとしても、音量検出用出力音響信号の指向特性もこれに追随して変化するため、出力音響信号の音量検出方向から到来する音の音量を示す音量検出用出力音響信号が継続して出力されることとなる。

40

#### 【 0 0 9 5 】

また、撮影者が操作部 1 7 を介して音量検出用指向特性制御部 7 1 c （特に音量検出用指向特性設定部 7 1 3 c ）に指示を入力することにより、音量検出方向の調整（音量検出方向の追加または削除、強調方向や強調幅の調整など）が行われることとしても構わない。

#### 【 0 0 9 6 】

50

(音量検出部)

音量検出部 7 2 は、指向特性制御部 7 1 c から出力される音量検出用出力音響信号の音量を検出することで、音量検出方向から到来する音の音量(レベル)を検出する。検出されて得られる音量は、音量検出情報として音量検出部 7 2 から出力され、表示画像生成部 8 2 c に入力される。

【0097】

なお、音量検出用指向特性制御部 7 1 c において、複数の音源に対応する複数の音量検出用目標指向特性が順次設定される場合、表示画像生成部 8 2 c は、入力される音量検出情報がどの音源に対応するものであるかを識別可能であるものとする。

【0098】

(表示画像生成部)

表示画像生成部 8 2 c は、上述した画像解析結果画像と、入力される音量検出情報が示す音量を表現した画像(以下、音量検出結果画像とする)と、を入力画像に重畳させて表示画像を生成する。生成される表示画像の一例を、図 1 1 に示す。図 1 1 は、第 3 実施例の画像音響処理部における表示画像生成部によって生成される表示画像の一例について示す図である。

【0099】

図 1 1 に示すように、表示画像 P 4 は、図 8 と同様の注目被写体 T 1 , T 2 を示す画像解析結果画像(顔枠画像 A 1 , A 2 )と、音量検出結果画像(数値画像 V 1 , V 2 )と、が入力画像に重畳表示されて成る。また、数値画像 V 1 は、注目被写体 T 1 の近傍に表示され、数値画像 V 2 は、注目被写体 T 2 の近傍に表示される。

【0100】

数値画像 V 1 は、注目被写体 T 1 が存在する方向を音量検出方向とした場合の音量検出用出力音響信号から検出される音量の値を表示したものである。また、数値画像 V 2 は、注目被写体 T 2 が存在する方向を音量検出方向とした場合の音量検出用出力音響信号から検出される音量の値を表現したものである。

【0101】

上述の第 1 実施例及び第 2 実施例と同様に、表示画像 P 4 を確認して出力音響信号の状態を認識した撮影者は、必要に応じて指向特性制御部 7 1 における目標指向特性の設定方法を変更することで、意図する出力音響信号を得る。このとき、所定の音源(例えば、注目被写体 T 1 , T 2 )の出力音響信号中の音量(例えば大小、目標値など)を指定する音量指定指示を入力可能な構成とすると、容易に出力音響信号を制御することができるため、好ましい。ただしこの場合、図 9 に示すように、音量指定指示を指向特性制御指示に変換する指向特性制御指示変換部 7 3 を備えることとする。指向特性制御指示変換部 7 3 から出力される指向特性制御指示は、指向特性制御部 7 1 だけでなく、上述のように音量検出用指向特性制御部 7 1 c にも入力される。なお、第 1 及び第 2 実施例と同様に、撮影者が指向特性制御指示を指向特性制御部 7 1 及び音量検出用指向特性制御部 7 1 c に直接的に入力可能な構成としても構わない。

【0102】

また、本例では音源が発する音の音量を確認することができるため、所定の音源(例えば注目被写体 T 1 , T 2 )に働きかけたり集音環境を変更したりすることも可能である。このような方法で、入力音響信号自体を変更することにより出力音響信号の状態を変更することも可能である。

【0103】

このように、入力画像から検出された注目被写体 T 1 , T 2 が発する音の音量を表現する数値画像 V 1 , V 2 を表示画像 P 4 中に表示することで、撮影者が注目被写体 T 1 , T 2 が発するそれぞれの音の状態(音量)を具体的に認識することが可能となる。そのため、撮影者が意図する出力音響信号が得られているか否かを容易に判断して、対応を講じることが可能となる。したがって、撮影者が意図する出力音響信号を容易かつ精度良く生成することが可能となる。

10

20

30

40

50

## 【 0 1 0 4 】

また、数値画像 V 1 , V 2 のそれぞれを、対応する画顔枠画像 A 1 , A 2 の近傍に表示することで、表示される音量がどの注目被写体 T 1 , T 2 が発した音の音量であるかを容易に認識することが可能となる。そのため、撮影者が注目被写体 T 1 , T 2 の一方が発する音を他方のものとして誤認することを抑制することが可能となる。

## 【 0 1 0 5 】

なお、第 1 実施例及び第 2 実施例と本実施例とを組み合わせても構わない。例えば、指向特性制御部から出力される目標指向特性情報が表示画像生成部 8 2 c に入力されるとともに、表示画像中に指向特性画像が表示されることとしても構わない（図 4 ~ 6 及び図 8 参照）。このように構成すると、表示画像を確認した撮影者が目標指向特性と音量とを同時

10

## 【 0 1 0 6 】

また、上記の図 1 1 と異なる方法で音量を表現した音量検出結果画像を用いても構わない。音量検出結果画像の別例について、図 1 2 を参照して説明する。図 1 2 は、音量検出結果画像の別例について示す図である。

## 【 0 1 0 7 】

図 1 2 ( a ) は、上下の長さ（ブロックの数）によって音量の大きさを表現する、いわゆるレベルメータによって音量を表現した音量検出結果画像の例を示したものである。なお、図 1 2 ( a ) では、上下方向に伸縮（増減）するものを示しているが、左右方向に伸縮（増減）するものとしても構わない。図 1 2 ( b ) は、放射線の数及び長さによって音量の値を表現した音量検出結果画像の例を示したものである。なお、図 1 2 ( b ) では、左右方向に伸縮（増減）するものを示しているが、上下方向に伸縮（増減）するものとしても構わない。

20

## 【 0 1 0 8 】

このように、音量を抽象的に表現した音量検出結果画像を用いることとすると、撮影者が音量の大きさを直感的かつ即座に認識することが可能となる。

## 【 0 1 0 9 】

また、上述のように音量検出方向は入力画像の画角外であっても構わない。例えば、撮影者方向を音量検出方向にすることも可能である。撮影者方向が音量検出方向となる場合の表示画像の例について、図 1 3 を参照して説明する。図 1 3 は、第 3 実施例の画像音響処理部における表示画像生成部によって生成される表示画像の別例について示す図である。

30

## 【 0 1 1 0 】

図 1 3 に示す表示画像 P 5 は、図 1 1 と同様の注目被写体 T 1 が検出され、顔枠画像 A 1、数値画像 V 1 が表示されたものとなっている。さらに表示画像 P 5 の端部（本例では下端）に、数値画像 V 3 が表示されている。数値画像 V 3 は、撮影者方向を音量検出方向とした場合の音量検出出力音響信号から検出される音量の値を表現したものである。

## 【 0 1 1 1 】

このように、入力画像の画角外の方向、特に撮影者方向から到来する音の音量を表示可能とすると、画角外の撮影者が発する音の音量までも撮影者が認識することが可能となる。そのため、さらに精度よく撮影者が意図する出力音響信号を生成することが可能となる。

40

## 【 0 1 1 2 】

また、画像解析部 8 1 が入力画像を解析することによって入力画像の画角外に存在する音源を検出して、当該音源の方向を音量検出方向として設定しても構わない。具体的に例えば、図 5 ( d ) について述べたように、入力画像を解析結果より注目被写体と撮影者とが会話していると推測される場合に撮影者を音源の一つとみなし、撮影者方向を音量検出方向として設定しても構わない。また、撮影者の指示によって画角外の音源を検出しても構わないし、図 1 0 の位相差算出部で得られる入力音響信号の位相差に基づいて画角外の

50



音源を検出しては構わない。

【0113】

<<その他変形例>>

上述の第1～第3実施例の画像音響処理部30a～30cによる表示画像及び出力音響信号の生成は、動画の記録時など出力音響信号を記録する際にのみ行われ得るものではなく、記録前のプレビュー時にも行われ得るものである。プレビュー時に表示画像及び出力音響信号の生成を行うことにより、出力音響信号の状態（指向特性や音量）を予め撮影者の意図したものにしておくことが可能となる。なお、プレビュー時において、出力音響信号を画像音響処理部30a～30cから出力しないこととしても構わない。

【0114】

また、上述の例は、動画を記録する撮像装置1に本発明の画像音響処理部（画像音響処理装置）を備える場合について説明したものであるが、再生装置に画像音響処理部を備えさせ、再生時に音響信号の指向特性を制御することも可能である。例えばこの場合、入力画像信号及び入力音響信号が、記録媒体に記録されたものや外部から入力されるものとして、表示画像信号が、テレビなどの表示装置で再生されるものとしても構わない。ただし、表示画像中の指向特性画像や画像解析結果画像、音量検出結果画像の表示/非表示を、ユーザの指示によって切り替え可能な構成とすると好ましい。

【0115】

また、本発明の実施形態における撮像装置1について、画像音響処理部30a～30cの動作を、マイコンなどの制御装置が行うこととしても構わない。さらに、このような制御装置によって実現される機能の全部または一部をプログラムとして記述し、当該プログラムをプログラム実行装置（例えばコンピュータ）上で実行することによって、その機能の全部または一部を実現するようにしても構わない。

【0116】

また、上述した場合に限らず、図2、図7及び図9の画像音響処理部30a～30cは、ハードウェア、あるいは、ハードウェア及びソフトウェアの組み合わせによって実現可能である。また、ソフトウェアを用いて画像音響処理部30a～30cを構成する場合、ソフトウェアによって実現される部位についてのブロック図は、その部位の機能ブロック図を表すこととする。

【0117】

以上、本発明の実施形態についてそれぞれ説明したが、本発明の範囲はこれに限定されるものではなく、発明の主旨を逸脱しない範囲で種々の変更を加えて実行することができる。

【産業上の利用可能性】

【0118】

本発明は、入力される画像信号と、当該画像信号と対になる音響信号と、に所定の処理を施して出力する画像音響処理装置や、当該画像音響処理装置を備えたデジタルビデオカメラに代表される撮像装置に関する。

【符号の説明】

【0119】

- 30a～30c 画像音響処理部
- 7 音響処理部
- 71 指向特性制御部
- 71c 音量検出用指向特性制御部
- 711L, 711R FFT部
- 712 位相差算出部
- 713 目標指向特性設定部
- 713c 音量検出用目標指向特性設定部
- 714 帯域別制御量設定部
- 715L, 715R 帯域別レベル制御部

10

20

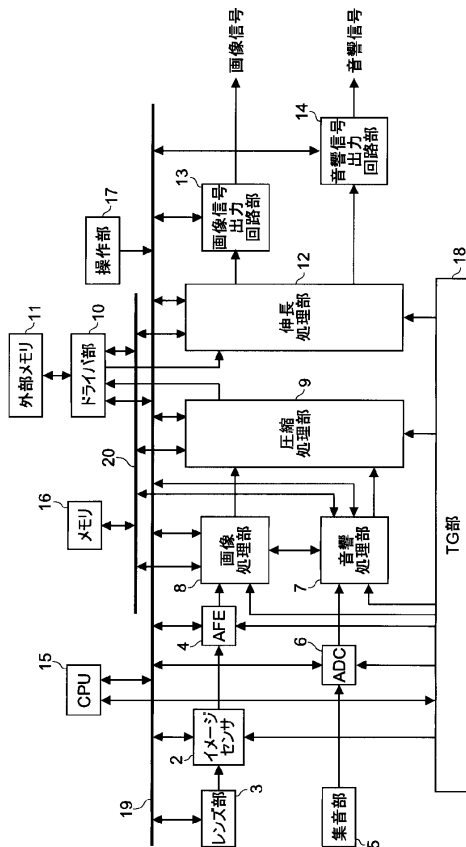
30

40

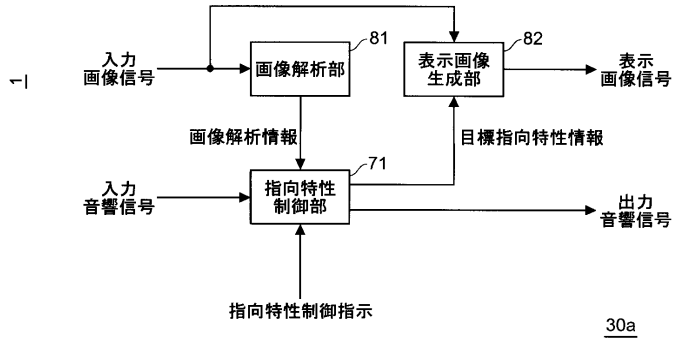
50

- 7 1 6 L , 7 1 6 R I F F T 部
- 7 2 音量検出部
- 7 3 指向特性制御指示変換部
- 8 画像処理部
- 8 1 画像解析部
- 8 2 , 8 2 b , 8 2 c 表示画像生成部

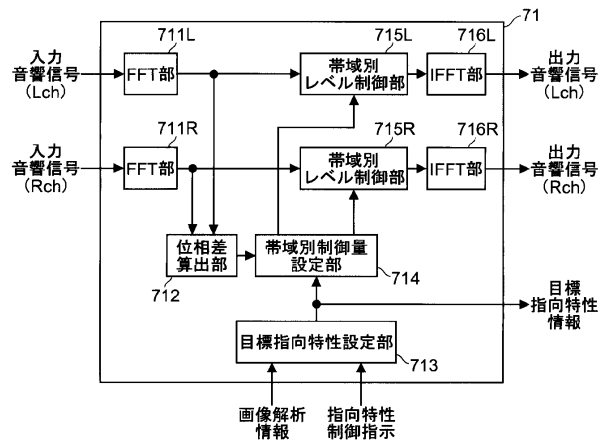
【 図 1 】



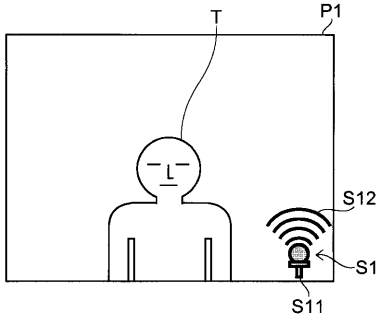
【 図 2 】



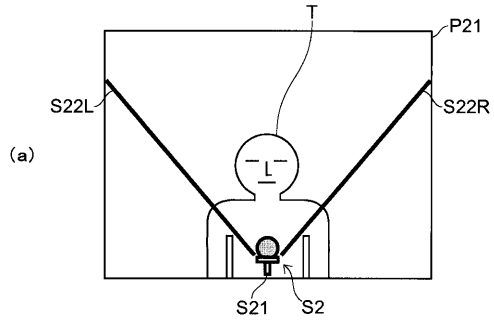
【 図 3 】



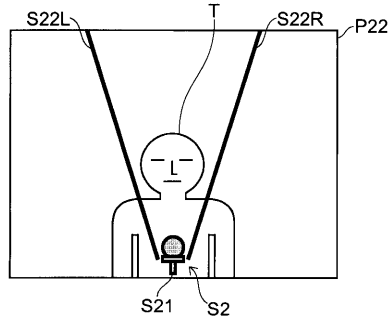
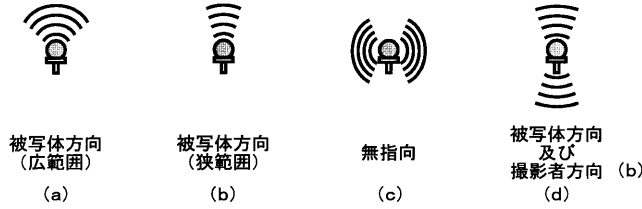
【図4】



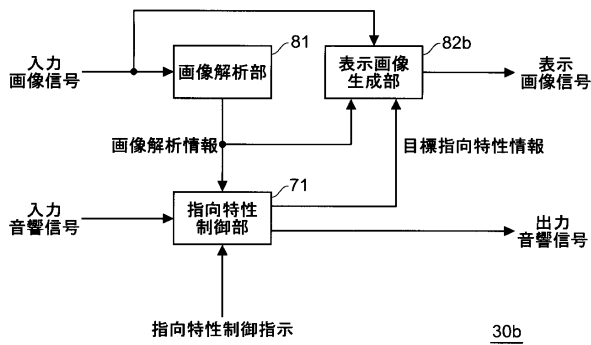
【図6】



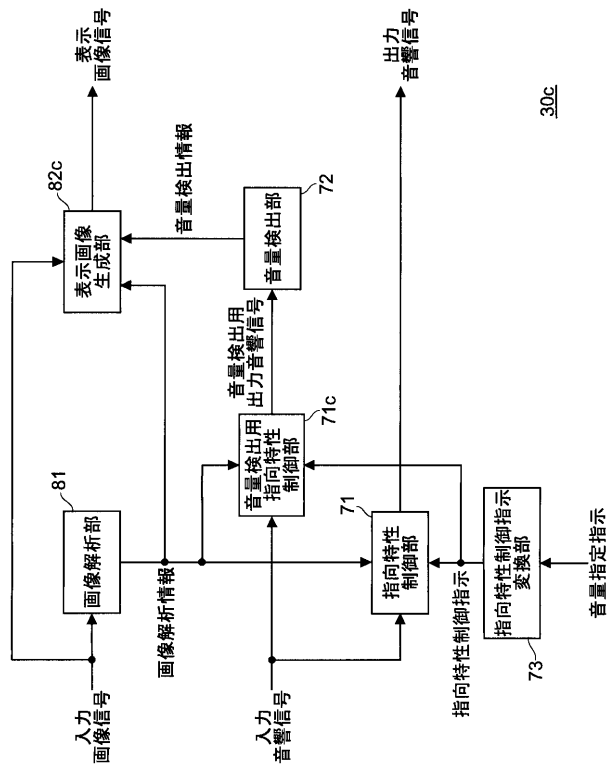
【図5】



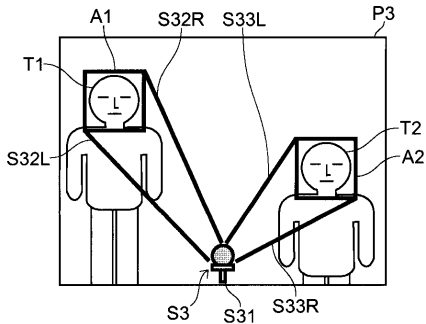
【図7】



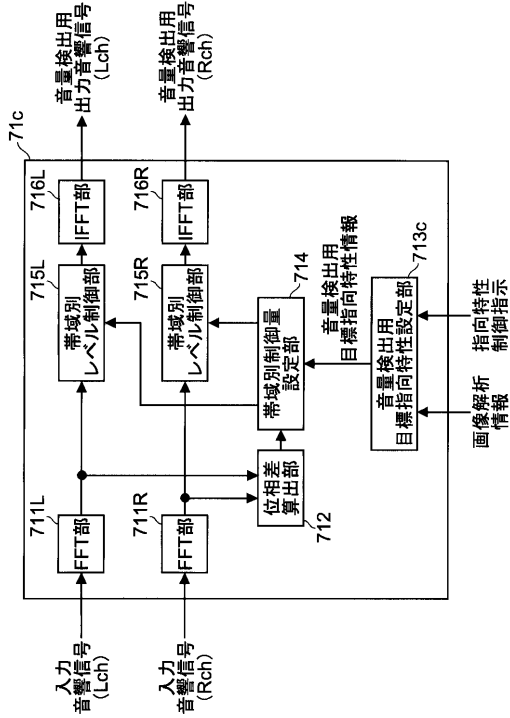
【図9】



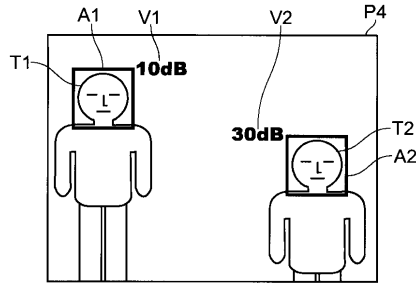
【図8】



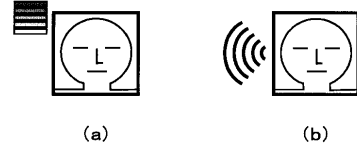
【図10】



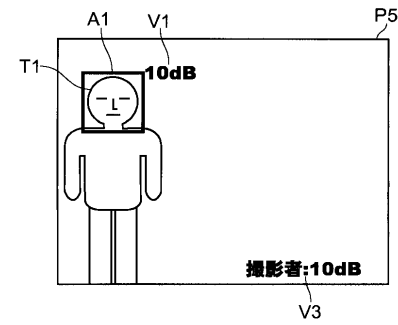
【図11】



【図12】



【図13】



## フロントページの続き

- (72)発明者 奥 智岐  
大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内
- (72)発明者 吉田 昌弘  
大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内
- (72)発明者 山中 誠  
大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内

審査官 吉 澤 雅博

- (56)参考文献 特開2008-236644(JP,A)  
特開2008-278433(JP,A)  
特開2008-193196(JP,A)  
特開2006-222618(JP,A)  
特開2007-251355(JP,A)  
特開平08-088853(JP,A)  
特開2001-245187(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04N 5/225  
H04R 3/00  
H04S 5/02