

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6877246号
(P6877246)

(45) 発行日 令和3年5月26日(2021.5.26)

(24) 登録日 令和3年4月30日(2021.4.30)

(51) Int.Cl.		F I			
HO4R	3/00	(2006.01)	HO4R	3/00	320
HO4R	1/08	(2006.01)	HO4R	1/08	
HO4R	1/02	(2006.01)	HO4R	1/02	107
G1OL	21/0208	(2013.01)	G1OL	21/0208	100Z
G1OL	21/0232	(2013.01)	G1OL	21/0232	

請求項の数 5 (全 32 頁)

(21) 出願番号 特願2017-111166 (P2017-111166)
 (22) 出願日 平成29年6月5日(2017.6.5)
 (65) 公開番号 特開2018-207317 (P2018-207317A)
 (43) 公開日 平成30年12月27日(2018.12.27)
 審査請求日 令和2年5月25日(2020.5.25)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 110003281
 特許業務法人大塚国際特許事務所
 (72) 発明者 辻本 悠貴
 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
 (72) 発明者 園田 啓太
 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
 (72) 発明者 佐藤 龍介
 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

最終頁に続く

(54) 【発明の名称】 音声処理装置及びその制御方法

(57) 【特許請求の範囲】

【請求項1】

音声処理装置であって、
 駆動部と、
 前記音声処理装置外からの音声を主として取得するための第1のマイクと、
 前記駆動部からの駆動騒音を主として取得する第2のマイクと、
 前記第1のマイクから得られた時系列の音声データを第1の周波数スペクトルデータに変換し、前記第2のマイクから得られた時系列の音声データを第2の周波数スペクトルデータに変換する変換手段と、
 前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、周波数毎の前記駆動騒音の減算量を演算する駆動音演算処理手段と、
 該駆動音演算処理手段で得た前記駆動騒音の減算量の時間軸に対する変動を抑制する第1の抑制手段と、
 前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから周波数毎の風雑音減算量を演算する風雑音演算処理手段と、
 該風雑音演算処理手段で得た前記風雑音減算量の時間軸に対する変動を抑制する第2の抑制手段と、
 前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、右チャンネルの仮のゲインと左チャンネルの仮のゲインを決定するステレオゲイン演算処理手段と、

10

20

該ステレオゲイン演算処理手段で決定した前記右チャンネルの仮のゲインと前記左チャンネルの仮のゲインの時間軸に対する変動を抑制する第3の抑制手段と、

前記第1、第2、第3の抑制手段による変動が抑制された駆動騒音の減算量、風雑音の減算量、右チャンネルの仮のゲインと左チャンネルの仮のゲインに基づき、左右のチャンネルそれぞれのトータルゲインを決定するトータルゲイン決定手段と、

前記変換手段で得た前記第1の周波数スペクトルデータと、前記トータルゲイン決定手段で得た左右のチャンネルそれぞれのトータルゲインに基づき、ステレオの左右のチャンネルの周波数スペクトルデータを生成するステレオ生成手段と、

該ステレオ生成手段で得た左右チャンネルの周波数スペクトルデータそれぞれを、時系列の音声データに逆変換する逆変換手段と

を有することを特徴とする音声処理装置。

【請求項2】

前記風雑音演算処理手段は、更に、前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータそれぞれの予め設定された周波数以下の成分の差分に基づいて風雑音レベルを算出し、

前記ステレオ生成手段は、

前記風雑音レベルに依存して決定される合成の上限周波数を求め、当該上限周波数以下では前記第1の周波数スペクトルデータに前記第2の周波数スペクトルデータを合成し、前記上限周波数を上回る周波数では前記第1の周波数スペクトルデータを選択する合成手段を有し、

当該合成手段から出力された周波数スペクトルデータと、前記トータルゲイン決定手段で得た左右のチャンネルそれぞれのトータルゲインに基づき、ステレオの左右のチャンネルの周波数スペクトルデータを生成する

ことを特徴とする請求項1に記載の音声処理装置。

【請求項3】

前記第1乃至第3の抑制手段は、それぞれに設定された時定数に基づいて時間軸に対する変動を抑制する

ことを特徴とする請求項1又は2に記載の音声処理装置。

【請求項4】

音声処理装置の制御方法であって、

前記音声処理装置は、駆動部と、前記音声処理装置外からの音声を主として取得するための第1のマイクと、前記駆動部からの駆動騒音を主として取得する第2のマイクとを有し、

前記方法は、

前記第1のマイクから得られた時系列の音声データを第1の周波数スペクトルデータに変換し、前記第2のマイクから得られた時系列の音声データを第2の周波数スペクトルデータに変換する変換工程と、

前記変換工程で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、周波数毎の前記駆動騒音の減算量を演算する駆動音演算処理工程と、

該駆動音演算処理工程で得た前記駆動騒音の減算量の時間軸に対する変動を抑制する第1の抑制工程と、

前記変換工程で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから周波数毎の風雑音減算量を演算する風雑音演算処理工程と、

該風雑音演算処理工程で得た前記風雑音減算量の時間軸に対する変動を抑制する第2の抑制工程と、

前記変換工程で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、右チャンネルの仮のゲインと左チャンネルの仮のゲインを決定するステレオゲイン演算処理工程と、

該ステレオゲイン演算処理工程で決定した前記右チャンネルの仮のゲインと前記左チャンネルの仮のゲインの時間軸に対する変動を抑制する第3の抑制工程と、

10

20

30

40

50

前記第1、第2、第3の抑制工程による変動が抑制された駆動騒音の減算量、風雑音の減算量、右チャンネルの仮のゲインと左チャンネルの仮のゲインに基づき、左右のチャンネルそれぞれのトータルゲインを決定するトータルゲイン決定工程と、

前記変換工程で得た前記第1の周波数スペクトルデータと、前記トータルゲイン決定工程で得た左右のチャンネルそれぞれのトータルゲインに基づき、ステレオの左右のチャンネルの周波数スペクトルデータを生成するステレオ生成工程と、

該ステレオ生成工程で得た左右チャンネルの周波数スペクトルデータそれぞれを、時系列の音声データに逆変換する逆変換工程と

を有することを特徴とする音声処理装置の制御方法。

【請求項5】

音声処理装置のプロセッサが読み込み実行するプログラムであって、

前記音声処理装置は、駆動部と、前記音声処理装置外からの音声を主として取得するための第1のマイクと、前記駆動部からの駆動騒音を主として取得する第2のマイクとを有し、

前記プログラムは、前記プロセッサを、請求項4に記載の方法の各工程を実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は音声処理装置、特に複数のマイクいの感度差を補正する技術に関するものである。

【背景技術】

【0002】

従来、デジタルカメラ、デジタルビデオカメラに代表される撮像装置は、撮像して得た動画像データと、被写体の周囲の音声データとを併せて記録する。以降、記録の目的となる被写体の周囲の音声を、以下、「周囲環境音」と称する。

【0003】

また、撮像装置は、光学レンズを移動させることで、撮像中に動被写体をフォーカスしたりズームしたりすることができる。しかしながら、光学レンズの移動する際の駆動音をマイクで拾ってしまうことがあり、このような駆動騒音を低減させることが望まれている。駆動騒音を低減させる方法として、例えば、特許文献1であらわされるような、スペクトルサブトラクション法などがある。しかしながら、特許文献1の方法により雑音低減処理を行った場合、雑音スペクトルの推定誤差が雑音低減処理後の信号に歪みとして残留し、耳障りな雑音として聞こえる事がある。この音は、人工雑音やミュージカルノイズとも呼ばれ、音声信号の品位を悪化させる要因の一つとなっている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2006-279185号公報

【特許文献2】特願2010-536590号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

このような、ミュージカルノイズを低減させる方法として、特許文献2にあらわされるような方法がある。特許文献2によると、入力信号に対し、S/N比に応じた振幅抑圧と、雑音スペクトルに応じた振幅抑圧の二つの処理を行う。周波数ごとに二つの抑圧処理後のスペクトルの大きい方を選択する事で、ミュージカルノイズを軽減し、かつ、入力音声の雑音を抑圧する事が出来る。しかしながら、特許文献2の方法では、音声処理におけるゲイン変動を要因としたミュージカルノイズは考慮していないため、音声信号にミュージカルノイズが残ってしまう可能性がある。

10

20

30

40

50

【0006】

本発明はかかる問題に鑑みなされたものであり、音声処理におけるゲイン変動を要因としたミュージカルノイズを抑制し、より適切にミュージカルノイズを低減させる技術を提供しようとするものである。

【課題を解決するための手段】

【0007】

この課題を解決するため、例えば本発明の音声処理装置は以下の構成を備える。すなわち、

音声処理装置であって、

駆動部と、

前記音声処理装置外からの音声を主として取得するための第1のマイクと、

前記駆動部からの駆動騒音を主として取得する第2のマイクと、

前記第1のマイクから得られた時系列の音声データを第1の周波数スペクトルデータに変換し、前記第2のマイクから得られた時系列の音声データを第2の周波数スペクトルデータに変換する変換手段と、

前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、周波数毎の前記駆動騒音の減算量を演算する駆動音演算処理手段と、

該駆動音演算処理手段で得た前記駆動騒音の減算量の時間軸に対する変動を抑制する第1の抑制手段と、

前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから周波数毎の風雑音減算量を演算する風雑音演算処理手段と、

該風雑音演算処理手段で得た前記風雑音減算量の時間軸に対する変動を抑制する第2の抑制手段と、

前記変換手段で得た前記第1の周波数スペクトルデータと前記第2の周波数スペクトルデータから、右チャンネルの仮のゲインと左チャンネルの仮のゲインを決定するステレオゲイン演算処理手段と、

該ステレオゲイン演算処理手段で決定した前記右チャンネルの仮のゲインと前記左チャンネルの仮のゲインの時間軸に対する変動を抑制する第3の抑制手段と、

前記第1、第2、第3の抑制手段による変動が抑制された駆動騒音の減算量、風雑音の減算量、右チャンネルの仮のゲインと左チャンネルの仮のゲインに基づき、左右のチャンネルそれぞれのトータルゲインを決定するトータルゲイン決定手段と、

前記変換手段で得た前記第1の周波数スペクトルデータと、前記トータルゲイン決定手段で得た左右のチャンネルそれぞれのトータルゲインに基づき、ステレオの左右のチャンネルの周波数スペクトルデータを生成するステレオ生成手段と、

該ステレオ生成手段で得た左右チャンネルの周波数スペクトルデータそれぞれを、時系列の音声データに変換する逆変換手段とを有する。

【発明の効果】

【0008】

本発明によれば、音声処理におけるゲイン変動を要因としたミュージカルノイズを抑制し、より適切にミュージカルノイズを低減させることが可能になる。

【図面の簡単な説明】

【0009】

【図1】実施形態の撮像装置のブロック構成図。

【図2】実施形態の撮像装置の撮像部、音声入力部の詳細なブロック構成図。

【図3】実施形態の撮像装置の音声入力部のメカ構成図。

【図4】実施形態の撮像装置のRECのシーケンスを示すフローチャート。

【図5】実施形態の撮像装置のL/Rch生成部のタイミングチャート。

【図6】実施形態の撮像装置の音声入力部の詳細な構成を示すブロック図。

【図7】周囲環境音の撮像装置への伝搬する系を示す図。

【図8】実施形態の撮像装置のメインマイクaからの周波数スペクトルとサブマイクbか

10

20

30

40

50

らの周波数スペクトルの位相の関係を示す図。

【図 9】実施形態のステレオ感の強調係数と周波数の関係を示す図。

【図 10】実施形態の撮像装置のメインマイク a とサブマイク b 其々の各周波数毎の振幅スペクトルを示す図。

【図 11】実施形態の撮像装置のサブマイク b の周波数 N ポイント目の時系列の振幅スペクトルを示す図。

【図 12】実施形態の撮像装置のメインマイク a とサブマイク b 其々の時系列の位相を示す図である。

【図 13】実施形態の撮像装置の M c h - S c h 演算部の動作タイミングチャート。

【図 14】実施形態の撮像装置の感度差補正部の動作タイミングチャート。

10

【図 15】実施形態の撮像装置の音声入力部のメカ構成図。

【図 16】実施形態の撮像装置のメインマイク a からの周波数スペクトルとサブマイク b からの周波数スペクトルを示す図。

【図 17】実施形態の風雑音レベルに対する風雑音ゲインの周波数関係を示す図。

【図 18】実施形態の撮像装置のメインマイク a からの周波数スペクトルとサブマイク b からの周波数スペクトルの合成される比率と周波数の関係を示す図。

【図 19】実施形態の撮像装置のステレオ抑制部について、駆動騒音検出時と風雑音検出時に応じて、ステレオ効果の強調に用いる強調係数を変更するタイミングチャート。

【図 20】実施形態の風雑音検出時における、合成比率と周波数とステレオ効果の強調に用いる強調係数の関係を示す図。

20

【図 21】実施形態の駆動騒音除去ゲインと風雑音減算量と L c h 生成用ステレオゲインと R c h 生成用ステレオゲインの時定数を示す図。

【発明を実施するための形態】

【0010】

以下図面に従って本発明に係る実施形態を詳細に説明する。本実施形態では、撮像装置に収容される音声処理装置について説明する。

【0011】

図 1 は実施形態の撮像装置 100 の構成を示すブロック図である。撮像装置 100 は、撮像部 101、音声入力部 102、メモリ 103、表示制御部 104、表示部 105 を有する。また、撮像装置 100 は、符号化処理部 106、記録再生部 107、記録媒体 108、制御部 109、操作部 110、音声出力部 111、スピーカ 112、外部出力部 113、並びに、これらを接続するバス 114 を有する。

30

【0012】

撮像部 101 は、撮影光学レンズにより取り込まれた被写体の光学像を撮像素子により画像信号に変換し、アナログデジタル変換、画像調整処理などを行い、画像データを生成する。撮影光学レンズは、内蔵型の光学レンズであっても、着脱式の光学レンズであっても良い。また、撮像素子は、CCD、CMOS 等に代表される光電変換素子であればよい。

【0013】

音声入力部 102 は、内蔵または音声端子を介して接続されたマイクにより、音声処理装置外（実施形態では撮像装置外）からの周辺の音声を集音し、電気信号を生成する。また、音声入力部 102 は、アナログデジタル変換、音声処理などを行い音声データを生成する。マイクは、指向性、無指向性を問わないが、本実施形態では無指向性のマイクを使用するものとする。

40

【0014】

メモリ 103 は、撮像部 101 により得られた画像データや、音声入力部 102 により得られた音声データを一時的に記憶するために利用される。

【0015】

表示制御部 104 は、撮像部 101 により得られた画像データに係る画像や、撮像装置 100 の操作画面、メニュー画面等を表示部 105 や、不図示の映像端子を介して外部の

50

ディスプレイに表示する。表示部 105 の種類は問わないが、例えば液晶表示器である。

【0016】

符号化処理部 106 は、メモリ 103 に一時的に記憶された画像データや音声データを読み出して所定の符号化を行い、圧縮画像データ、圧縮音声データ等を作成する。また、音声データに関しては圧縮しないようにしてもよい。圧縮画像データは、例えば、MPEG2 や H.264 / MPEG4 - AVC など、どのような圧縮方式で圧縮されたものであってもよい。また、圧縮音声データも、AC3 (A) AC、ATRAC、ADPCM などのような圧縮方式で圧縮されたものであってもよい。また、符号化処理部 106 は、上記の符号化データ（圧縮画像データ、圧縮音声データ）の復号処理も行う。

【0017】

記録再生部 107 は、記録媒体 108 に対して、符号化処理部 106 で生成された圧縮画像データ、圧縮音声データまたは音声データ、各種データを記録したり、記録媒体 108 から読出したりする。ここで、記録媒体 108 は、画像データ、音声データ等を記録する不揮発性の記録媒体である。例えば、磁気ディスク、光学式ディスク、半導体メモリなどであり、その種類は問わない。また、記録媒体 108 は、本装置 100 に対して固定であっても、脱着可能であっても構わない。

【0018】

制御部 109 は、バス 114 を介して、撮像装置 100 の各ブロックに制御信号を送信することで撮像装置 100 の各ブロックを制御するものであり、各種制御を実行するための CPU やメモリなどから構成される。制御部 109 で使用するメモリは、各種制御プログラムを格納する ROM、演算処理のためのワークエリアとして利用する RAM 等であり、制御部 109 の外付けのメモリも含む。

【0019】

操作部 110 は、ボタン、ダイヤル、タッチパネル、或いはそれらの組み合わせであり、ユーザの操作に応じて、指示信号を制御部 109 に送信する。操作部 110 は、具体的には、動画記録開始、終了を指示するための撮影ボタン、光学的もしくは電子的に画像に対してズーム動作する指示するためのズームレバー、各種調整をするための十字キー、決定キーなどを有する。

【0020】

音声出力部 111 は、記録再生部 107 により再生された音声データや圧縮音声データ、または制御部 109 により出力される音声データをスピーカ 112 や音声端子などに出力する。外部出力部 113 は、記録再生部 107 により再生された圧縮映像データや圧縮音声データ、音声データなどを外部機器に出力する。データバス 114 は、音声データや画像データ等の各種データ、各種制御信号を撮像装置 100 の各ブロックに供給する。

【0021】

以上が実施形態における撮像装置 100 の構成の説明である。次に、実施形態における撮像装置の通常の動作について説明する。

【0022】

本実施形態の撮像装置 100 は、ユーザが操作部 110 を操作して電源を投入する指示が出されたことに応じて、不図示の電源供給部からの電力が、撮像装置の各ブロックに供給される。

【0023】

電源が供給されると、制御部 109 は、操作部 110 のモード切り換えスイッチが、例えば、撮影モード、再生モード等のどのモードを指定しているかを、操作部 110 からの指示信号により確認する。撮影モードにおける動画記録モードでは、撮像部 101 により得られた画像データと音声入力部 102 により得られた音声データとを 1 つの画像ファイルとして保存する。再生モードでは、記録媒体 108 に記録された画像ファイルを記録再生部 107 により再生して表示部 105 に表示させ、スピーカ 112 より出力することになる。

【0024】

10

20

30

40

50

撮影モードでは、まず、制御部109は、撮影待機状態に移行させるように制御信号を撮像装置100の各ブロックに送信し、以下のような動作をさせる。

【0025】

撮像部101は、撮影光学レンズにより取り込まれた被写体の光学像を撮像素子により動画像信号に変換し、アナログデジタル変換、画像調整処理などを行い、動画像データを生成する。そして、撮像部101は、得られた動画像データを表示処理部104に送信し、表示部105に表示させる。なお、撮像部101は、1フレームが水平1920画素×垂直1080画素、フレームレートが30フレーム/秒の動画像信号を出力する。ユーザはこの様にして表示された画面を見ながら撮影の準備を行う。

【0026】

音声入力部102は、複数のマイクにより得られたアナログ音声信号をデジタル信号に変換し、得られた複数のデジタル音声信号を処理して、マルチチャンネルの音声データを生成する。そして、得られた音声データを音声出力部111に送信し、接続されたスピーカ112や不図示のイヤホンから音声として出力させる。ユーザは、この様にして出力された音声を聞きながら記録音量を決定するためのマニュアルボリュームの調整をすることもできる。

【0027】

次に、ユーザが操作部110の記録ボタンを操作することにより撮影開始の指示信号が制御部109に送信されると、制御部109は、撮像装置100の各ブロックに撮影開始の指示信号を送信し、撮影モードにおける動画像記録モードに移行する。具体的な、制御部109の処理は以下の通りである。

【0028】

撮像部101は、撮影光学レンズにより取り込まれた被写体の光学像を撮像素子により動画像信号に変換し、アナログデジタル変換、画像調整処理などを行い、動画像データを生成する。そして、得られた動画像データを表示処理部104に送信し、表示部105に表示させる。また、撮像部101は、得られた画像データをメモリ103へ送信する。

【0029】

音声入力部102は、複数のマイクにより得られたアナログ音声信号をデジタル信号に変換し、得られた複数のデジタル音声信号を処理して、マルチチャンネルの音声データを生成する。そして、得られた音声データをメモリ103に送信する。また、マイクが一つの場合には、得られたアナログ音声信号をデジタル変換し音声データを生成し、音声データをメモリ103に送信する。

【0030】

符号化処理部106は、メモリ103に一時的に記憶された動画像データや音声データを読み出して所定の符号化を行い、圧縮動画像データ、圧縮音声データ等を生成し、再びメモリ103に格納する。

【0031】

制御部109は、メモリ103に格納された圧縮動画像データ、圧縮音声データを合成し、データストリームを形成し、記録再生部107に出力する。音声データを圧縮しない場合には、制御部109は、メモリ103に格納された音声データと圧縮動画像データとを合成し、データストリームを形成して記録再生部107に出力する。

【0032】

記録再生部107は、UDF、FAT等のファイルシステム管理のもとに、データストリームを一つの動画ファイルとして記録媒体108に書き込んでいく。

【0033】

撮像装置100は、上記の処理を動画記録状態中、継続することになる。そして、ユーザが操作部110の記録ボタンを操作することにより撮影終了の指示信号が制御部109に送信されると、制御部109は、撮像装置100の各ブロックに撮影終了の指示信号を送信し、以下のような動作をさせる。

【0034】

10

20

30

40

50

撮像部 101、音声入力部 102 は、それぞれ動画像データ、音声データの生成を停止する。符号化処理部 106 は、メモリに記憶されている残りの画像データと音声データとを読み出して所定の符号化を行い、圧縮動画像データ、圧縮音声データ等を生成し終わったら動作を停止する。音声データを圧縮しない場合には、当然、圧縮動画像データの生成が終わったら動作を停止する。

【0035】

そして、制御部 109 は、これらの最後の圧縮動画像データと、圧縮音声データまたは音声データとを合成し、データストリームを形成し、記録再生部 107 に出力する。

【0036】

記録再生部 107 は、UDF、FAT等のファイルシステム管理のもとに、データストリームを一つの動画ファイルとして記録媒体 108 に書き込んでいく。そして、データストリームの供給が停止したら、動画ファイルを完成させて、記録動作を停止させる。

10

【0037】

制御部 109 は、記録動作が停止すると、撮影待機状態に移行させるように制御信号を撮像装置 100 の各ブロックに送信して、撮影待機状態に戻る。

【0038】

次に、再生モードについて説明する。ユーザが操作部 110 を操作して再生モードにした場合、制御部 109 は、再生状態に移行させるように制御信号を撮像装置 100 の各ブロックに送信し、以下のような動作をさせる。

【0039】

20

記録媒体 108 に記録された圧縮動画像データと圧縮音声データとからなる動画ファイルを記録再生部 107 が読み出して、読み出された圧縮動画像データ、圧縮音声データを符号化処理部 106 に送る。

【0040】

符号化処理部 106 は、圧縮動画像データ、圧縮音声データを復号し、それぞれを表示制御部 104、音声出力部 111 に送信する。表示制御部 104 は、復号された動画像データを表示部 105 に表示させる。音声出力部 111 は、復号された音声データを内蔵のスピーカ 112、または、取付けられた外部スピーカに出力して、音響として再生させる。

【0041】

30

本実施形態の撮像装置 100 は以上のように、動画像、音声の記録再生を行うことができる。

【0042】

本実施形態では、音声入力部 102 において、音声信号を得る際に、マイクにより得られた音声信号のレベル調整処理等の処理をしている。この処理は、装置が起動してから常に行われてもよいし、撮影モードが選択されてから行われてもよい。或いは、音声の記録に関連するモードが選択されてから行われてもよい。また、音声の記録に関連するモードにおいて、音声の記録が開始したことに応じて上記の処理を行ってもよい。本実施形態では、動画像撮影の開始されたタイミングで上記の処理を行うものとして説明する。

【0043】

40

図 2 は本実施形態の撮像装置 100 の撮像部 101、音声入力部 102 のブロック構成図である。

【0044】

撮像部 101 は、被写体の光学像を取り込む光学レンズ 201、光学レンズ 201 により取り込まれた被写体の光学像を電気信号（画像信号）に変換させる撮像素子 202 を有する。さらに、撮像部 101 は、撮像素子 202 により得られたアナログ画像信号をデジタル画像信号に変換し、画質調整処理をして画像データを形成し、メモリに送信する画像処理部 203 を有している。さらに、撮像部 101 は、光学レンズ 201 を移動させるための位置センサ、モータ等の公知の駆動メカニズムを有する光学レンズ制御部 204 を有している。本実施形態では、撮像部 101 に光学レンズ 201、光学レンズ制御部 204

50

が内蔵されているように記載しているが、光学レンズ201は、レンズマウントを介して撮像装置100に着脱自在な交換レンズであっても良い。また、光学レンズ制御部204は、交換レンズ内に設けられるようにしても良い。

【0045】

ここで、ズーム動作、フォーカス調整などの指示を、ユーザが操作部110を操作して入力すると、制御部109は、光学レンズ制御部204に光学レンズ201を移動させるための制御信号(駆動信号)を送信する。光学レンズ制御部204は、この制御信号に応じて、不図示の位置センサで光学レンズ201の位置を確認し、不図示のモータ等で光学レンズ201の移動を行う。また、画像処理部203により得られた画像や被写体との距離を制御部109が確認し、自動的に調整する場合は、光学レンズを駆動させる制御信号を送信することになる。また、画像のブレを防止する、いわゆる防振機能を備えている場合には、制御部109は、不図示の振動センサにより検出された振動に基づいて、光学レンズ201を移動させるための制御信号を光学レンズ制御部204に送信することになる。

10

【0046】

このときに、光学レンズ201の移動による駆動騒音や光学レンズ201を移動させるためのモータの駆動騒音が発生することになる。制御部109からの光学レンズ201を駆動させる制御信号に応じて、光学レンズ制御部204が光学レンズ201を駆動させる。従って、制御部109は、駆動騒音が発生するタイミングを知る(検出するまたは、決定する)ことができる。

20

【0047】

本実施形態において、光学レンズ201の制御により、例えば最大で50倍、最小で1倍のズームングを光学的に行うことができる。これを本実施形態では光学ズームと言う。勿論、光学ズームの倍率は前記以上でも前記以下でも構わないものとする。光学ズームは、制御部109からの指示で、光学レンズ制御部204が、光学レンズ201の光学レンズを移動させることで、被写体の光学像をズームングさせるものである。また、画像処理部203は、撮像素子202により得られた画像信号の一部をズームインした画像信号を出力する電子ズーム機能を備えている。また、撮像素子202により得る画像の範囲を広くし、画像処理部203で画像サイズをズームアウトした画像信号を出力する電子ズーム機能を備えている。

30

【0048】

以上が実施形態における撮像部101の構成とその動作である。次に、音声入力部102の構成と動作を説明する。

【0049】

実施形態の撮像装置100は、参照符号205a、205bで示す2つのマイクを有する。これらマイク205a、205bは空気(媒体)を伝播する振動を電気信号に変換し、音声信号を出力するものである。マイク205aがメイン(MAIN)マイク、マイク205bがサブ(SUB)マイク205bであって、以降、この名称で表現する。

【0050】

詳細は後述する説明で明らかにするが、メインマイク205aは、ステレオ音声の一方のチャンネルに対応するマイクとして機能し、且つ、音声処理装置外(実施形態では撮像装置100外)から音声を主として取得するためのマイクである。また、サブマイク205bは、ステレオ音声のもう一方のチャンネルに対応するマイクとして機能する位置に配置される。サブマイク205bは、メインマイク205aと比較して、音声処理装置内(撮像装置100)の駆動部からの駆動騒音を主として取得するためのマイクである。

40

【0051】

メインマイク205aはアナログの音声信号をMch(メインチャンネル)、サブマイク205bはアナログの音声信号をSch(サブチャンネル)として出力する。本実施形態において第一の音声入力部をメインマイク205a、第一の音声信号をMchとする。また、第二の音声入力部をサブマイク205b、第二の音声信号をSchとする。本実施形態

50

では2チャンネルで構成されたステレオ方式とするため、メインマイク205aとサブマイク205bの配置位置は、撮像部101の正立に構えた際の水平方向に所定距離隔てた位置に設けられている。なお、実施形態では、マイク数を2としているが、それ以上のマイクを保持する構成でも構わない。

【0052】

メインマイク205a、サブマイク205bにより得られたアナログ音声信号はA/D変換部206に供給され、ここでそれぞれの音声信号がデジタルの音声データに変換される。本実施形態におけるA/D変換部206は、48kHzのサンプリングレートでサンプリングを行い、1サンプリング当たり16bitのデジタルデータを生成するものとする。

10

【0053】

A/D変換部206で得られた、予め設定された音声信号の期間(フレーム)の時系列のデジタルの音声データはFFT部207に供給され、ここで高速フーリエ変換され、周波数毎の周波数スペクトルデータに変換される。本実施形態において、周波数スペクトルは、0Hzから48kHzまでにおいて1024ポイントの周波数スペクトルデータとして変換され、ナイキスト周波数である24kHzまでにおいては512ポイントの周波数スペクトルを持つものとする。メインマイク205aからの周波数スペクトルデータをMain[0]~[511]、サブマイク205bからの周波数スペクトルデータをSub[0]~[511]と表す。また、本実施形態において、第一の音声スペクトルデータをMain[0]~[511]、第二の音声スペクトルデータをSub[0]~[511]と表すものとする。なお、各スペクトルデータの添え字が「0」が最低周波数を、「511」が最大周波数を表すものとする。

20

【0054】

駆動音演算処理部209は、駆動部を駆動させるための、制御部109からの制御信号に応じて、FFT部207により得た周波数スペクトルデータの周波数成分毎の、駆動騒音の減算量を決定する。この駆動騒音は、光学レンズ201が駆動されることにより発生される。なお、本実施形態における駆動部はズーム動作、フォーカス調整により駆動する光学レンズ201を指すものとする。駆動音演算処理部209は、周波数スペクトル毎の減算量を表すNC_Gain[0]~[511]と、駆動騒音検出信号を出力する。

【0055】

詳細は後述する説明から明らかになるが、感度差補正部208は、駆動音演算処理部209からの、1フレーム前の駆動騒音検出信号に応じて、現フレームのMain[0]~[511]に対するSub[0]~[511]の感度を補正し、補正後の周波数スペクトルデータMain[0]~[511]、Sub[0]~[511]を出力する。

30

【0056】

風雑音演算処理部210は、FFT部207からの周波数スペクトルデータから、風雑音を検出し、減算量を決定する。そして、風雑音演算処理部210は、決定した風雑音の周波数スペクトルデータWC_Gain[0]~[511]と、風雑音レベル信号を出力する。

【0057】

ステレオゲイン演算処理部211は、FFT部207からの周波数スペクトルデータに対し、ステレオのLch(左チャンネル)及びRch(右チャンネル)それぞれのゲインを決定する。そして、ステレオゲイン演算処理部211は、各チャンネルの、決定した周波数スペクトルの成分毎のゲインを表すGain_L[0]~[511]とGain_R[0]~[511]を出力する。ここで、左チャンネルのゲインがGain_L[0]~[511]、右チャンネルのゲインがGain_R[0]~[511]である。

40

【0058】

トータルゲイン演算部212は、駆動音演算処理部209、風雑音演算処理部210、および、ステレオゲイン演算処理部211において決定したNC_Gain[0]~[511]、WC_Gain[0]~[511]、Gain_L[0]~[511]、Gai

50

$n_R[0] \sim [511]$ を合算し、 $Total_Gain_L[0] \sim [511]$ 、 $Total_Gain_R[0] \sim [511]$ を出力する。具体的には、次式の通りである。実施形態では、このトータルゲイン演算部212は、トータルゲイン決定部として機能する。

$Total_Gain_R[] = NC_Gain[] + WC_Gain[] + Gain_R[]$

$Total_Gain_L[] = NC_Gain[] + WC_Gain[] + Gain_L[]$

【0059】

L/Rch生成部213は、 $MAIN[0] \sim [511]$ の周波数毎の周波数スペクトルと、トータルゲイン演算部212で決定した $Total_Gain_L[0] \sim [511]$ 、 $Total_Gain_R[0] \sim [511]$ を用いて、LchとRchの周波数スペクトルデータを生成する(詳細後述)。つまり、本実施形態におけるL/Rch生成部213はステレオ生成部として機能する。

10

【0060】

iFFT部214は、L/Rch生成部213で生成された各チャンネルの周波数スペクトルデータに対して逆高速フーリエ変換を行い、それぞれのチャンネルの時系列の音声信号に戻す。

【0061】

音声処理部215は、イコライザ等の処理を実施する。オートレベルコントローラは、時系列の音声信号の振幅を所定のレベルに調整する(以後、ALC部216)。

【0062】

20

以上の構成により、音声入力部102は、音声信号に所定の処理を行い音声データを形成し、メモリ103へ送信することになる。

【0063】

次に、本実施形態の撮像装置100の記録動作について図4を用いて説明する。同図は実施形態の撮像装置100の記録のシーケンスを示すフローチャートである。

【0064】

S401にて、ユーザによる操作部110の操作により記録(REC)開始が指示されることで、本処理が開始される。S402にて、制御部109は音声録音するために音声のパスを接続する。音声パスが確立した後、S403にて、制御部109は、本実施形態で説明する制御を含めた信号処理の初期設定をおこない、処理を開始する。この信号処理の内容に関しては後述する。以降、RECシーケンスが終了するまで、本実施形態で説明する制御を含めた信号処理は実施される。

30

【0065】

記録処理シーケンス中、制御部109は、ユーザによる操作部110への操作を監視する。そして、ユーザにより、操作部110の一部であるズームレバーが操作された場合、S404からS405に処理を進め、制御部109は撮像部101を制御し、ズーム処理を行う。このズーム処理は、S406にて、ユーザがズームレバーの操作を止めたと判定されるまで継続する。ズーム処理中は、先に説明したように、レンズ201の移動による駆動騒音が発生し、その騒音が周囲環境音に重畳して録音されてしまう点に注意されたい。

40

【0066】

そして、制御部109は、ユーザによる操作部110の操作や、記録媒体108の状況によって、記録終了が指示されたと判断した場合、S407からS408に処理を進める。S408にて、制御部109は音声パスを切断し、次いで、S409にて信号処理も終了する。

【0067】

次に、本実施形態の撮像装置100の音声入力部102の詳細を図6を用いて説明する。同図は、本実施形態の音声入力部102の詳細な構成を示すブロック図である。

【0068】

本実施形態における音声入力部102は、前述の通り、空気中を伝播する音声振動を電

50

気信号に変換し、音声信号を出力するメインマイク205aとサブマイク205bを有する。また前述の通り、A/D変換部206は、アナログ音声信号を、48KHz、16bitのサンプリングを行い、アナログ音声信号からデジタル音声データに変換する。

【0069】

感度差補正部208は、メインマイク205aからの周波数スペクトルデータMain[0]~[511]と、サブマイク205bからの周波数スペクトルデータSub[0]~[511]との感度差を補正する。このため、感度差補正部208は、感度補正積分器2081、感度補正検出部2082、補正量演算部2083、感度補正ゲインテーブル2084、感度差補正ゲイン部2085を含む。

【0070】

感度補正積分器2081は、メインマイク205aからの周波数スペクトルデータMain[0]~[511]、及び、サブマイク205bからの周波数スペクトルデータSub[0]~[511]に対し、時間軸方向のレベル変化に時定数を持たせる。

【0071】

感度補正検出部2082は、感度補正積分器2081にて時定数を持たせた周波数スペクトルデータであるMain[0]~[511]とSub[0]~[511]のレベル差『Main[n]-Sub[n]』を、全周波数ポイントについて求める。ここで、差分は正負の符号が発生することに注意されたい。

【0072】

補正量演算部2083は、感度補正検出部2082からの差分レベルが負の場合(Main[n]<Sub[n]の場合に等価)、Main[n]=Sub[n]となるようにするため、Sub[n]の補正量を算出する。

【0073】

なお、感度補正検出部2082からの差分レベルが正の場合(Main[n]>Sub[n]の場合に等価)、Sub[n]を補正する必要が無い。したがって、この場合、補正量演算部2083はSub[n]の補正量として0を出力する。

【0074】

感度補正ゲインテーブル2084は、補正量演算部2083にて算出された各周波数スペクトルSub[0]~[511]の具体的な補正量が格納している。

【0075】

感度差補正ゲイン部2085は、実際に、感度補正ゲインテーブル2084を基に各周波数スペクトルSub[0]~[511]のレベル補正を実行する。

【0076】

ここで上記の時定数については、感度補正の追従を限りなく遅くする事を目的とするので数十秒単位とする。また、感度補正積分器2081は、後述する駆動検出部2095により駆動騒音の検出を表す駆動騒音検出信号を受けた場合、その動作を停止する。これは、光学レンズ201が駆動している不安定な期間における積分を排除する事を意図する。

【0077】

以上が実施形態における感度差補正部208を構成する各処理部の説明である。次に、駆動音演算処理部209について説明する。

【0078】

駆動音演算処理部209は、メインマイク205a、サブマイク205bからの周波数スペクトルデータであるMain[0]~[511]、Sub[0]~[511]から、駆動騒音の減算量NC_Gain[0]~[511]を決定し、駆動騒音を検出した事を示す駆動騒音検出信号を出力する。このため、駆動音演算処理部209は、Mch-Sch演算部2091、駆動騒音除去ゲイン演算部2092、時間毎振幅変動検出部2093、時間每位相変動検出部2094、駆動検出部2095、フレーム間振幅差検出部2096、駆動音減算量積分器2097を有する。

【0079】

Mch-Sch演算部2091は、メインマイク205aからの周波数スペクトルデー

10

20

30

40

50

タMain[0]~[511]から、サブマイク205bからの周波数スペクトルデータSub[0]~[511]を差し引いた値を、駆動騒音の減算量として出力する。

【0080】

ただし、周波数スペクトルnポイント目において、Main[n]>Sub[n]の場合には、減算量[n]は0とする。つまり、Mch-Sch演算部2091は、周波数スペクトルnポイント目において、Main[n]-Sub[n]<0であることを条件に負の値を減算量[n]として出力する。

【0081】

また、Main[n]に対してSub[n]が十分に大きく、Main[n]-Sub[n]が予め設定した閾値(負の値)を下回る場合、Mch-Sch演算部2091は、駆動騒音を検出したことを示す検出信号[n]を出力し、否の場合には検出信号を出力しない。なお、実際には、騒音検出を"1"、非検出を"0"として表しても良い。

【0082】

また、駆動騒音検出の判定は、減算関係を逆にして、Sub[n]-Main[n]と閾値(正の値を持つ)との比較で行っても良い。この場合、Mch-Sch演算部2091は、この演算の結果が閾値を上回った場合に駆動騒音検出を示す信号を出力することになる。

【0083】

駆動検出部2095は、Mch-Sch演算部2091からの1フレーム分の検出信号[0]~[511]を受け、その中に、1以上の検出信号が存在した場合、該当のフレームにおいては駆動騒音を検出した事を表す駆動騒音検出信号を出力する。

【0084】

Mch-Sch演算部2091及び駆動検出部2095による処理は、正の閾値をThと定義したとき、次式を満たす「i」(iは0から511のいずれか)が存在するか否かの判定を行い、その判定結果を駆動騒音検出を示す信号として出力していると、と言える。

$Ma i n [i] + T h < S u b [i]$

【0085】

時間毎振幅変動検出部2093は、メインマイク205aからの周波数スペクトルデータMain[0]~[511]、サブマイク205bからの周波数スペクトルデータSub[0]~[511]に対し、時間方向のフレーム間での振幅変動量の検出を行う。具体的には、時間毎振幅変動検出部2093は、現在のフレームの周波数スペクトルのnポイント目の成分値と、前フレームの周波数スペクトルのnポイント目の成分値との差分値を求め、出力する。そして、nポイント目での変動量が予め設定された閾値を超えた場合、時間毎振幅変動検出部2093は、時間毎振幅変動量[n]を出力し、閾値以下の場合には0を出力する。

【0086】

時間毎位相変動検出部2094は、後述の位相差判定部2111から取得する位相情報に基づき、メインマイク205aからの周波数スペクトルデータMain[0]~[511]、サブマイク205bからの周波数スペクトルデータSub[0]~[511]の位相変動量の検出を行う。例えば周波数スペクトルnポイント目において前記変動量が予め定められた閾値を超えた場合は、時間毎位相変動検出部2094は、時間毎位相変動量[n]を出力する。また、変動量が閾値以下の場合、時間毎位相変動検出部2094は、時間毎位相変動量[n]を出力しない、又は、時間毎位相変動量[n]=0として出力する。

【0087】

フレーム間振幅差検出部2096は、駆動検出部2095からの駆動騒音検出信号に基づき、サブマイク205bからの周波数スペクトルデータであるSub[0]~[511]の時間方向のフレーム間での振幅差の検出を行う。例えば周波数スペクトルnポイント目において、駆動騒音検出信号があり、前フレームと現フレームとの振幅差が予め定めら

10

20

30

40

50

れた閾値を超えた場合は、フレーム間振幅差検出部 2096 は、フレーム間振幅差量 [n] を出力する。また、差が閾値以下の場合、フレーム間振幅差検出部 2096 は、フレーム間振幅差量 [n] を出力しない、もしくは、フレーム間振幅差量 [n] = 0 として出力する。

【0088】

駆動騒音除去ゲイン演算部 2092 は、同一フレームにおいて、前述の M c h - S c h 演算部 2095 からの減算量 [0] ~ [511]、時間毎振幅変動検出部 2093 からの時間毎振幅変動量 [0] ~ [511]、時間毎位相変動検出部 2094 からの時間毎位相変動量 [0] ~ [511]、フレーム間振幅差検出部 2096 からのフレーム間振幅差量 [0] ~ [511] 其々の結果に対し、予め定められた係数を乗算して、加算した駆動騒音除去量 [0] ~ [511] を算出し、出力する。

10

【0089】

駆動音減算量積分器 2097 は、駆動騒音除去ゲイン演算部 2092 から出力された駆動騒音除去量 [0] ~ [511] に対し、時間方向の変動量に時定数を持たせ、駆動騒音除去ゲイン N C _ G a i n [0] ~ [511] (正負の符号付き) を出力する。

【0090】

以上が実施形態の駆動音演算処理部 209 の構成と動作である。次に、風雑音演算処理部 210 について説明する。

【0091】

風雑音演算処理部 210 は、メインマイク 205 a からの周波数スペクトルデータ M a i n [0] ~ [511]、サブマイク 205 b からの周波数スペクトルデータ S u b [0] ~ [511] から風雑音を検出し、減算量を表す W C _ G a i n [0] ~ [511] と、風雑音レベル信号を出力する。風雑音演算処理部 210 は、風検出部 2101、風雑音ゲイン演算部 2102、風雑音減算量積分器 2103 を有する。

20

【0092】

風検出部 2101 は、メインマイク 205 a からの周波数スペクトル M a i n [0] ~ [511]、及び、サブマイク 205 b からの周波数スペクトル S u b [0] ~ [511] のうちから、それぞれの低周波数域の所定数のポイントの相関に応じて風雑音レベルの検出を行う。例えば低域の 10 ポイントにおいて、次式に従い風雑音レベルを求め、出力する。なお、ここでの " n " は、実施形態の場合には 0 乃至 9 であるが、この数は適宜変更しても構わない。

30

$$\text{風雑音レベル} = (\text{M a i n} [n] - \text{S u b} [n]) / (\text{M a i n} [n] + \text{S u b} [n])$$

なお、上式の は、 n = 0 乃至 9 の合算を示している。

【0093】

また、風雑音ゲイン演算部 2102 は、図 17 に示すような特性線分を持つテーブルを有する。図示のように、1つの線分は、或る周波数以下ではゲインが負、その周波数以上ではゲインが 0 となる。そして、ゲインが負から 0 となる周波数の位置が互いに異なる複数の線分を含む。そして、風雑音ゲイン演算部 2102 は、風雑音レベルに従った 1つの線分を用いて、風雑音ゲイン [0] ~ [511] を決定し、出力する。なお、実施形態では、風雑音ゲイン [0] ~ [511] をテーブルを用いて決定するものとしたが、風雑音レベルを引数とする関数を用いて、風雑音ゲイン [0] ~ [511] を決定しても良い。

40

【0094】

風雑音減算量積分器 2103 は、風雑音ゲイン演算部 2102 から出力された風雑音ゲイン [0] ~ [511] に対し、時間方向の変動量に時定数を持たせ、風雑音ゲイン W C _ G a i n [0] ~ [511] (正負の符号付き) を出力する。

【0095】

以上が実施形態における風雑音演算処理部 210 の構成と動作である。次に、実施形態におけるステレオゲイン演算処理部 211 を説明する。

【0096】

50

ステレオゲイン演算処理部 2 1 1 は、メインマイク 2 0 5 a からの周波数スペクトルデータ $Main[0] \sim [511]$ 、サブマイク 2 0 5 b からの周波数スペクトルデータ $Sub[0] \sim [511]$ から、ステレオの Lch のゲイン $Gain_L[0] \sim [511]$ と、 Rch のゲイン $Gain_R[0] \sim [511]$ を生成し、出力する。このために、ステレオゲイン演算処理部 2 1 1 は、位相差判定部 2 1 1 1、ステレオゲイン演算部 2 1 1 2、ステレオ抑制部 2 1 1 3、左ゲイン積分器 2 1 1 4、右ゲイン積分器 2 1 1 5 を有する。

【 0 0 9 7 】

位相差判定部 2 1 1 1 は、周波数スペクトルデータ $Main[0] \sim [511]$ に対する $Sub[0] \sim [511]$ の位相情報を算出する。

10

【 0 0 9 8 】

例えば、周波数スペクトルデータにおける各ポイントの位相ベクトルを $V()$ として表した場合の、周波数ポイント n の位相情報 $[n]$ は次式に従って算出される。

位相情報 $[n] = |V(Main[n]) \times V(Sub[n])| / (|V(Main[n])| \cdot |V(Sub[n])|)$

ここで、右辺の " $|x|$ " はベクトル x の絶対値 (スカラー) を表し、分母の " \cdot " はスカラーどうしの積、分子の " \times " は 2 つのベクトルの正弦である外積を表している。

【 0 0 9 9 】

位相差判定部 2 1 1 1 は上式に従って算出した位相情報 $[0] \sim [511]$ を出力する。

20

【 0 1 0 0 】

ステレオゲイン演算部 2 1 1 2 は、位相差判定部 2 1 1 1 からの位相情報 $[0] \sim [511]$ からステレオゲイン $[0] \sim [511]$ の演算を行う。例えば周波数ポイント n において、次式に従って各チャンネルのゲインを得る。

Lch 生成用のステレオゲイン = $1 + \text{位相情報}[n] \times \text{強調係数}$

Rch 生成用のステレオゲイン = $1 - \text{位相情報}[n] \times \text{強調係数}$

ステレオゲイン演算部 2 1 1 2 は、上式にて算出された Lch 、 Rch のステレオゲイン $[n]$ を出力する。ここで、強調係数は周波数に応じて変更されるものであり、上限を 1、下限を 0 とするものである。

【 0 1 0 1 】

30

ステレオ抑制部 2 1 1 3 は、駆動音演算処理部 2 0 9 内の $Mch - Sch$ 演算部 2 0 9 1 からの駆動騒音を検出したことを示す検出信号を受けた場合に強調係数を 0 にする。また、ステレオ抑制部 2 1 1 3 は、風雑音演算処理部 2 1 0 内の風検出部 2 1 0 1 からの風雑音レベルに応じて強調係数を 0 にする。

【 0 1 0 2 】

左ゲイン積分器 2 1 1 4 は、ステレオゲイン演算部 2 1 1 2 から出力された、 Lch 生成用のステレオゲイン $[0] \sim [511]$ に対し、時間方向の変動量に所定の時定数を持たせ、それをステレオゲイン $GainL[0] \sim [511]$ (正負の符号付き) として出力する。

【 0 1 0 3 】

40

右ゲイン積分器 2 1 1 5 は、ステレオゲイン演算部 2 1 1 2 から出力された、 Rch 生成用のステレオゲイン $[0] \sim [511]$ に対し、時間方向の変動量に所定の時定数を持たせ、それをステレオゲイン $GainR[0] \sim [511]$ (正負の符号付き) として出力する。

【 0 1 0 4 】

以上が実施形態のステレオゲイン演算処理部 2 1 1 の構成と動作である。次に、実施形態におけるトータルゲイン演算部 2 1 2 を説明する。

【 0 1 0 5 】

トータルゲイン演算部 2 1 2 は、駆動音演算処理部 2 0 9、風雑音演算処理部 2 1 0、および、ステレオゲイン演算処理部 2 1 1 において決定した $NC_Gain[0] \sim [5$

50

11]、WC_Gain[0]～[511]、Gain_L[0]～[511]、Gain_R[0]～[511]を合算し、Total_Gain_L[0]～[511]、Total_Gain_R[0]～[511]を出力する。具体的には次式である。

Total_Gain_L[] = NC_Gain[] + WC_Gain[] + Gain_L[]

Total_Gain_R[] = NC_Gain[] + WC_Gain[] + Gain_R[]

【0106】

次に、L/Rch生成部213を説明する。このL/Rch生成部213は、周波数スペクトルデータMAIN[0]～[511]から、トータルゲイン演算部212で決定したTotal_Gain_L[0]～[511]、Total_Gain_R[0]～[511]を用いて、LchとRchの出力用の周波数スペクトルデータを作成する。L/Rch生成部213は、Mch/Sch選択部2131、L/Rchゲイン加算部2132を有する。

10

【0107】

Mch/Sch選択部2131は、風検出部2101による風雑音レベルに応じて、周波数スペクトルデータMain[0]～[511]に合成することになるSub[0]～[511]の周波数ポイントの範囲を選択する。また、Mch/Sch選択部2131は、風雑音レベルに応じて、合成する境界位置を低周波数ポイントから高周波数ポイントへと変化させる。また、風を検出されない場合、Mch/Sch選択部2131は合成を行わず、周波数スペクトルデータMain[0]～[511]をそのまま出力する。

20

【0108】

L/Rchゲイン加算部2132は、Mch/Sch選択部2132から出力された周波数スペクトルデータMain[0]～[511]に対して、トータルゲイン演算部212で決定したTotal_Gain_L[0]～[511]、Total_Gain_R[0]～[511]を用いて、左右チャンネル(LchとRch)の周波数スペクトルデータを作成する。

【0109】

以上が実施形態のL/Rch生成部213の構成と動作である。

【0110】

iFFT部214は、L/Rch生成部213で生成された各チャンネルの周波数スペクトルデータを逆変換(逆FFT変換)し、元の時系列の音声信号に戻す。音声処理部215は、イコライザ等の処理を実施する。ALC(オートレベルコントローラ)216は、時系列の音声信号の振幅を所定のレベルに調整する。

30

【0111】

以上の構成を備え、音声入力部102は、音声信号に所定の処理を行い音声データを形成して、メモリ103へ送信し、格納することになる。

【0112】

ここで、本実施形態の音声入力部102の一部を構成するメカ構成について、図3(a)、3(b)を用いて説明する。

【0113】

図3(a)は、本実施形態の撮像装置の筐体の外観図である。撮影対象に撮像装置が向いた状態で、撮影者から見て右側の所定位置の参照符号"a"がメインマイク205aの入力穴(開口部)、左側の対向する位置の参照符号"b"がサブマイク205bの入力穴となる。図3(b)においての拡大図は、音声入力部102の一部であるメインマイク205aとサブマイク205bのメカ構成部である。図3(b)は、前記メカ構成を示す断面図である。マイク穴を構成する外装部102-1、メインマイク205aを保持するメインマイクブッシュ102-2a、サブマイク205bを保持するサブマイクブッシュ102-2b、其々のマイクブッシュを外装部へ押し付け保持をする押し付け部103により構成される。外装部102-1、押し付け部103についてはPC材等のモールド部材で構成されるが、アルミ、ステンレス等の金属部材であっても問題ない。また、メインマイクブッシュ102-2a、サブマイクブッシュ102-2bについては、エチレンプロピレ

40

50

ンジエンゴム等のゴム材にて構成される。

【0114】

ここで、外装部におけるマイク穴の径について説明する。サブマイク205bへのマイク穴の径（開口している面積）は、メインマイク205aへのマイク穴の径（同面積）に対して小さく、所定の倍率にて縮小された構成をとる。マイク穴形状については円状か楕円状が望ましいが、方形状でも構わない。また、其々の穴形状について、同形状でも別形状でも構わない。前記構成は、撮像装置内部でマイクに空気伝搬して伝わる駆動騒音についてサブマイク205bのマイク穴側から外部へ漏れにくくなる事を目的とする。

【0115】

次に、外装部102-1とマイクブッシュで構成されるマイク前面の空間について説明する。外装部102-1とサブマイクブッシュ102-2bで構成されるサブマイク205bの前面の空間の容積は、外装部102-1とメインマイクブッシュ102-2aで構成されるメインマイク205aの前面の空間のそれより大きく、所定の倍率の容積を確保する構成をとる。この構成は、サブマイク205bの前面の空間において、空間内の気圧変化が大きくなり、駆動騒音が強調される事を目的とする。

【0116】

前述の通り、マイク入力メカ構成におけるサブマイク205b入力は、メインマイク205a入力に対して、駆動騒音の振幅が大きく強調される構成をとる。各マイクへ入力される駆動騒音の音声レベルの関係は、メインマイク205a<サブマイク205bとなる。一方、マイク穴の前面から空気伝搬により各マイクへ入力される、装置外からの音声（本来の集音目的である周辺環境音）のレベル関係は、メインマイク205a サブマイク205bの関係となることに注意されたい。

【0117】

ここで、本実施形態の音声入力部102でのステレオゲイン演算処理部211の動作について、図7から図9を用いて説明する。

【0118】

図7は、撮像装置100に内蔵されたマイクに対する外部からの音声の経路と、内蔵の光学レンズ201の駆動時の音声の経路の一例を示している。この時のマイクは、図2に示すメインマイク205aおよびサブマイク205bが該当する。図7のように周囲環境音の音源と撮像装置100との距離は、メインマイク205aとサブマイク205b間の距離に対して十分に大きい。よって、周囲環境音の音源からのメインマイク205aへの音声の伝播経路と、周囲環境音の音源とサブマイク205bへの音声の伝播経路は殆ど同一と考えて良い。しかし、撮像装置内蔵の光学レンズ201は、メインマイク205aとサブマイク205bに近接している。また、光学レンズ201の移動を行うためのモータからマイクへの距離が均等でなかったり、撮像装置内での音声の経路が異なる可能性もある。故に、光学レンズ駆動系からメインマイク205a、サブマイク205bそれぞれへの音声経路（距離）は大きく異なってしまふ。つまり、周囲環境音と駆動騒音とでは、MchとSchの音声レベルの差分に大きな差が出る事となる。それ故、周囲環境音と光学レンズの駆動騒音は大きく差が出て、これらを容易に区別することができる。

【0119】

一方、本来、周囲環境音は左右のどちら側から発生したかはMchとSchでは大きさでは判断することは難しい。そこで、周囲環境音は音声信号の位相を利用して判断することができる。詳細について説明する。

【0120】

図8(a)~(c)は、或る周波数スペクトルデータMain[n]とSub[n]の関係を示している。

【0121】

ステレオゲイン演算処理部211は、メインマイク205aからの周波数スペクトルデータMain[0]~[511]、サブマイク205bからの周波数スペクトルデータSub[0]~[511]から、ステレオのLchのゲインGain_L[0]~[511]

10

20

30

40

50

、RchのゲインGain_R[0]～[511]を出力する。ステレオゲイン演算処理部211は以下の構成を備えている。

【0122】

位相差判定部2111は、周波数スペクトルデータMain[0]～[511]に対する周波数スペクトルデータSub[0]～[511]の位相情報を算出する。

【0123】

例えば周波数ポイントnの周囲環境音が、メインマイク205a側から発生した場合、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の関係は図8(a)のような関係になる。本実施形態でのマイク配置においても、周波数スペクトルの大きさは変わってしまっても、位相は変わることはない。そこで、位相情報を、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の外積($|V(\text{Main}[n]) \times V(\text{Sub}[n])|$)を用いることで得る。

位相情報[n] = $|V(\text{Main}[n]) \times V(\text{Sub}[n])| / (|V(\text{Main}[n])| \cdot |V(\text{Sub}[n])|)$

位相差判定部2111は、上式にて算出された位相情報[n]を出力する。ここで求められる位相情報[n]は、すなわち、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ のsinであり、周囲環境音がメインマイク205a側(撮像装置100を構えるユーザの右側)から発生した場合は、 $0 < \text{位相情報}[n] < 1$ となる。

【0124】

また、周波数ポイントnの周囲環境音が、サブマイク205b側から発生した場合、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の関係は図8(b)のような周波数スペクトルの関係になる。本実施形態でのマイク配置においても、周波数スペクトルの大きさは変わってしまっても、位相は変わることはない。

【0125】

そこで、位相情報を $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の外積($|V(\text{Main}[n]) \times V(\text{Sub}[n])|$)を用いることで得る。

位相情報[n] = $|V(\text{Main}[n]) \times V(\text{Sub}[n])| / (|V(\text{Main}[n])| \cdot |V(\text{Sub}[n])|)$

位相差判定部2111は、上式にて算出された位相情報[n]を出力する。ここで求められる位相情報[n]は、すなわち、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ のsinであり、周囲環境音がサブマイク205b側からの場合、 $0 > \text{位相情報}[n] > -1$ となる。

【0126】

また周波数ポイントnの周囲環境音がメインマイク205a、サブマイク205bと同じ距離、すなわち光学レンズ201の中心から発生した場合、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の関係は図8(c)のような周波数スペクトルの関係になる。本実施形態でのマイク配置においても、周波数スペクトルの大きさは変わってしまっても、位相は変わることはない。

【0127】

位相情報は、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ の外積($|V(\text{Main}[n]) \times V(\text{Sub}[n])|$)を用いることで得ることができる。

位相情報[n] = $|V(\text{Main}[n]) \times V(\text{Sub}[n])| / (|V(\text{Main}[n])| \cdot |V(\text{Sub}[n])|)$

位相差判定部2111は、上式にて算出された位相情報[n]を出力する。ここで求められる位相情報[n]は、 $V(\text{Main}[n])$ と $V(\text{Sub}[n])$ のsinであり、周囲環境音が、サブマイク205b側からは位相情報[n] = 0となる。

【0128】

ステレオゲイン演算部2112は、上記のようにして決定した位相情報[0]～[511]を用いて、ステレオゲイン[0]～[511]の演算を行っている。例えば周波数ポイントnにおいて、ステレオゲイン演算部2112は次式に従って各チャンネルのゲインを算出する。

10

20

30

40

50

L c h生成用のステレオゲイン = 1 + 位相情報 [n] × 強調係数

R c h生成用のステレオゲイン = 1 - 位相情報 [n] × 強調係数

そして、ステレオゲイン演算部 2 1 1 2 は、上式にて算出された各チャンネルのステレオゲイン [n] を出力する。

【 0 1 2 9 】

図 9 はステレオゲイン演算部 2 1 1 2 で用いられる各周波数ポイントにおける強調係数を示した図である。

【 0 1 3 0 】

横軸を周波数ポイント、縦軸を強調係数とした時、もっとも強調したい周波数の強調係数を最大値の 1 . 0 として、位相差がでにくい低域と位相差が判断できない高域の強調係数は最小値の 0 とする。

10

【 0 1 3 1 】

例えばもっとも強調したい 1 k H z ~ 5 k H z は強調係数を 1 . 0 とし、2 0 0 H z 以下は 0 とする。

【 0 1 3 2 】

位相差が判断できない高域の強調係数は、メインマイク 2 0 5 a とサブマイク 2 0 5 b の距離で決定する。例えば、メインマイク 2 0 5 a とサブマイク 2 0 5 b の距離が 1 5 m m の時、音速を 3 4 0 m / s とすると、1 5 m m の間に半波長が入る 1 1 . 3 k H z 以上になると、正しい位相情報が取れず、左右が反転してしまう可能性がある。また、1 5 m m の間に 1 / 4 波長が入る 5 . 7 k H z 以上は正確性が低い。そこで図 9 に示すような周波数に応じた強調係数のかけ方を行う。

20

【 0 1 3 3 】

ここで、本実施形態の音声入力部 1 0 2 での駆動音演算処理部 2 0 9、トータルゲイン演算部 2 1 2、L / R c h 生成部 2 1 3 の動作について、図 5、図 1 0 から図 1 3 を用いて説明する。

【 0 1 3 4 】

図 1 0 は、メインマイク 2 0 5 a とサブマイク 2 0 5 b それぞれの各周波数の振幅スペクトルデータの例を示している。

【 0 1 3 5 】

F F T 部 2 0 7 により、各チャンネルの音声信号は 0 H z から 4 8 k H z までにおいて 1 0 2 4 ポイントの周波数スペクトルとして変換される。変換後の周波数スペクトルデータは、ナイキスト周波数である 2 4 k H z までにおいては 5 1 2 ポイントの周波数スペクトルを持つものとする。

30

【 0 1 3 6 】

先に図 3 (a) , (b) を用いて説明したように、実施形態の撮像装置 1 0 0 のマイク入力のメカ構成によれば、サブマイク 2 0 5 b は、メインマイク 2 0 5 a に対して、駆動騒音の振幅が大きく強調された信号を生成する。つまり振幅スペクトルにおいて、

周囲環境音レベル：メインマイク 2 0 5 a サブマイク 2 0 5 b

駆動騒音レベル：メインマイク 2 0 5 a < サブマイク 2 0 5 b

との関係となる。

40

【 0 1 3 7 】

図 1 0 に、メインマイク 2 0 5 a からの振幅スペクトルデータ M a i n []、サブマイク 2 0 5 b からの振幅スペクトルデータ S u b [] の一例を示す。また、同図における「M a i n - S u b」は、M c h - S c h 演算部 2 0 9 1 にて演算される、M a i n [] から S u b [] を差し引いた減算量 [0] ~ [5 1 1] を示している。

【 0 1 3 8 】

例えば、S c h における N ポイント目の周辺の振幅スペクトルを着目すると、S c h > M c h であり、つまり駆動騒音が支配的なポイントである事が言える。この時、M a i n - S u b には、N ポイント目周辺にて予め定められたズーム検出閾値を超える (下回る) 減算量が算出され、N ポイント目周辺は「駆動騒音」とされる振幅スペクトルと検出され

50

る。一方、M c hにおけるN 2ポイント目の振幅スペクトルに着目すると、S c h M c hである。つまり周囲環境音が支配的なポイントであることが言える。この時、M a i n - S u bには、ズーム検出閾値を超える減算量は算出されないため、N 2ポイント目周辺の振幅スペクトルは駆動騒音とは検出されることはない。上記演算を[0] ~ [5 1 1]の振幅スペクトル全ての範囲において実行する。

【 0 1 3 9 】

図 1 1 は、サブマイク 2 0 5 b の周波数 N ポイント目の時系列の振幅スペクトルを示す図である。

【 0 1 4 0 】

図示の「S u b c h」は、Nポイント目の振幅スペクトルデータが時系列にて変動する事を示す。

10

【 0 1 4 1 】

$S c h | t_n - t_{(n-1)} |$ は、S c h Nポイント目の振幅スペクトルに対し、時間毎振幅変動検出部 2 0 9 3 により演算される時間方向のフレーム間での振幅変動量を示し、時間毎変動量 [n] として出力される。例えば、 t_1 から t_2 にて S c h の振幅スペクトルに着目すると、時間方向での変動量は大きくなっており、 $S c h | t_n - t_{(n-1)} |$ には、 t_1 から t_2 において、変動量検出閾値を超える時間毎変動量が算出される。この演算を [0] ~ [5 1 1] の振幅スペクトルの全てのポイントにおいて実行する。

【 0 1 4 2 】

図 1 2 (a) , (b) は、メインマイク 2 0 5 a からの振幅スペクトル、サブマイク 2 0 5 b からの振幅スペクトルにおける、周波数 N ポイント目の時系列の位相を示す図である。

20

【 0 1 4 3 】

同図 (a) は複素数平面 I m , R e により、時間方向における「周囲環境音」の位相の変化を示しており、実線部は M c h を、点線部は S c h を表している。 t_0 , t_1 , t_2 , t_3 , t_4 については、時間方向の推移を示す。

【 0 1 4 4 】

同図 (b) は、「駆動騒音」の位相の変化を示している。

【 0 1 4 5 】

ここにおいて、周囲環境音については、M c h と S c h の位相は、 t_0 から t_4 の時間の推移において一定である。駆動騒音については、M c h と S c h の位相は t_0 から t_4 の時間の推移において大きく変動している。其々の時間方向での位相の変動は時間毎位相変動検出部 2 0 9 4 にて検出され、時間毎位相変動量 [n] として出力される。時間毎位相変動検出部 2 0 9 4 は、この演算を [0] ~ [5 1 1] の振幅スペクトルの全て周波数ポイントについて実行する。

30

【 0 1 4 6 】

図 1 3 (a) , (b) は、M c h - S c h 演算部 2 0 9 1 の動作タイミングチャートの一例を表している。

【 0 1 4 7 】

同図 (a) における M a i n [N] , S u b [N] , M a i n [N] - S u b [N] は、それぞれ周波数 N ポイント目の M c h の振幅スペクトルデータ、S c h の振幅スペクトルデータ、M c h 振幅スペクトルから S c h 振幅スペクトルを差し引いた減算量 [N] を示している。M a i n [N] - S u b [N] は、M c h - S c h 演算部 2 0 9 1 にて演算を行われた結果を出力している。

40

【 0 1 4 8 】

ここで、同図 (a) の t_1 から t_2 の期間に着目すると、S u b [N] の振幅スペクトルは、M a i n [N] に対して大きく上回っており、M a i n [N] - S u b [N] の演算結果はズーム閾値を上回る結果となっており、駆動騒音として検出され、減算量 [N] が出力される。

【 0 1 4 9 】

50

図13(b)におけるMain[N2]、Sub[N2]、Main[N2]-Sub[N2]は、それぞれ周波数N2ポイント目のMchの振幅スペクトル、Schの振幅スペクトル、Mch振幅スペクトルからSch振幅スペクトルを差し引いた減算量[n]を示す。ここで、同図(b)のt1からt2の期間を着目すると、Main[N2]とSub[N2]が同レベルで変動しており、Main[N2]-Sub[N2]の演算結果もズーム閾値を上回る結果はない。周波数N2ポイント目において駆動騒音は検出されない結果となる。Mch-Sch演算部2091は上記タイミングチャートで示した演算を[0]~[511]の振幅スペクトル全てにおいて実行する。

【0150】

図5はL/Rch生成部213のタイミングチャートの一例を表す。ズーム駆動動作は、制御部109からの制御を受け、t1からt2のタイミングにおいて、光学レンズ201が駆動動作となる。Mchスペクトルは、図5において抽出した特定の周波数Nポイント目のスペクトルを表す。Lch、Rchについては、トータルゲイン演算部212で決定したTotal_Gain_L、Total_Gain_RをMchに加算することで生成される。同図のタイミングチャートに示されるように、例えば、Mchに対し、Total_Gain_Lを下げ、Total_Gain_Rを上げることで、Rchが強調することができ、1chの入力で2chのステレオ信号を生成する事が可能である。

【0151】

また、t1からt2における光学レンズの駆動動作中においても、Total_Gain_L、Total_Gain_Rを下げることで、Lch、Rchに対し、駆動騒音を除去することが可能である。

【0152】

ここで、ここで、本実施形態の音声入力部102での感度差補正部208の動作について、図14を用いて説明する。

【0153】

図14は、感度差補正部208の動作タイミングチャートの一例を示している。同図において、ズーム検出は駆動検出部2095の駆動騒音の検出結果を示す。入力スペクトルNPointは、周波数Nポイント目のMchの振幅スペクトル、Schの振幅スペクトルを示す。実線部はMchを、点線部はSchを示わしている。

【0154】

入力スペクトル(積分)NPointは、周波数Nポイント目の感度補正積分器2081のMch、Schの積分結果を示す。感度調整出力スペクトルNPointは、周波数Nポイント目の感度差補正ゲイン部2085によりレベル補正されたMchの振幅スペクトル、Schの振幅スペクトルを示す。実線部はMchを、点線部はSchを示す。

【0155】

図14において、t0はREC開始のタイミングであり、t0からt1にかけては数10秒程度の充分長い時間を表わしている。タイミングt2からt3にかけては、ズーム検出がONされており、駆動検出部2095により駆動騒音が発生していることを表す。

【0156】

入力スペクトルNPointは、MchとSchはREC開始時t0においてレベル差が生じている。それに対し、入力スペクトル(積分)NPointは、感度補正積分器2081により、積分されt0からt1にかけてゆっくりとレベル差に追従していく。感度調整出力スペクトルNPointも入力スペクトル(積分)NPointの積分結果に対し、t0からt1にかけて十分に時間を掛けて感度差補正ゲイン部2085にてゲイン補正をしていく。これは、感度差補正部208は、メインマイク205aとサブマイク205bの感度補正を目的としているので、数十秒程度の充分な時間を掛けてのレベル補正で良く、過渡的な応答性を必要としない。

【0157】

また、タイミングt2からt3にかけてのズーム検出ON期間においては、感度補正積分器2081が停止状態となる。よって、駆動騒音が発生することにより、Mchの振幅

10

20

30

40

50

スペクトル、S c hの振幅スペクトルに大きなレベル差が発生するが、感度補正積分器2081が停止状態にあるので、レベル差に追従することなく、値は保持される。前述したが、感度差補正部208は、メインマイク205aとサブマイク205bの感度補正を目的としているので、駆動騒音による過渡的なレベル差分に対する応答は必要としない。感度差補正部208は上記タイミングチャートで示した補正を[0]～[511]の振幅スペクトル全てにおいて実行する。

【0158】

ここで、本実施形態の音声入力部102での風雑音演算処理部210の動作について、図15から図17を用いて説明する。

【0159】

図15は、音声入力部102の一部であるサブマイク205bに対し、風防材102-3を構成したメカ構成を示す断面図である。

【0160】

マイク穴を構成する外装部102-1は、メインマイク205aを保持するメインマイクブッシュ102-2a、サブマイク205bを保持するサブマイクブッシュ102-2b、其々のマイクブッシュを外装部へ押し付け保持をする押し付け部103により構成される。外装部102-1、押し付け部103についてはP C材等のモールド部材で構成されるが、アルミ、ステンレス等の金属部材であっても問題ない。また、メインマイクブッシュ102-2a、サブマイクブッシュ102-2bについては、エチレンプロピレンジエンゴム等のゴム材にて構成される。

【0161】

ここで、外装部102-1におけるマイク穴の穴径について説明する。サブマイク205bへのマイク穴の径は、メインマイク205aへのマイク穴の径に対して小さい。実施形態では、サブマイク205bのマイク孔の径(直径)は、メインマイク205aのマイク穴のその1/3の寸法としている。マイク穴形状については円状、楕円状が望ましいが、方形状でも構わない。また、其々の穴形状について、同形状でも別形状でも構わない。

【0162】

次に、外装部102-1とマイクブッシュ102-2a、102-2bで構成されるマイク前面の空間と、クッション材の配置について説明する。外装部102-1とサブマイクブッシュ102-2bで構成されるサブマイク205bの前面の空間の容積は、外装部102-1とメインマイクブッシュ102-2aで構成されるメインマイク205a前面の空間のそれより大きく、3倍確保する構成をとる。

【0163】

外装部102-1とサブマイクブッシュ102-2bで構成されるサブマイク205bの前面の空間においては、風防材102-3として、風防クッション材やシールマイクを配置する。いずれも風の周波数に対応した0～4kHz程度の低周波帯の信号成分をフィルタする部材として構成される。風防材102-3により低周波帯が支配的である風雑音のサブマイク205bへの空気伝搬の影響を大幅に軽減する事が可能である。

【0164】

図16は、風雑音入力時のメインマイク205aからの周波数スペクトルデータMain[0]～[511]と、サブマイク205bからの周波数スペクトルデータSub[0]～[511]を示している。風雑音入力時、風雑音成分は点線部の低周波帯域において存在している。風検出部2101は、メインマイク205aからの周波数スペクトルMain[0]～[511]、サブマイク205bからの周波数スペクトルSub[0]～[511]のうちから、低域周波数帯の例えば10ポイントの相関をみて風雑音レベルの検出を行っている。風検出部2101は、例えば低域の周波数ポイントnにおいて、次式に従って風雑音レベルを算出し、出力する。

風雑音レベル = (Main[n] - Sub[n]) / (Main[n] + Sub[n])

10

20

30

40

50

なお、上式は、低周波成分の10ポイントとしているので、 n は0乃至9の範囲内である。また、実施形態では、低域周波数帯を10ポイントとしたが、この数は一例である。撮像装置の設計に応じて適宜設定することが望まれる。

【0165】

図17は、風雑音ゲイン演算部2102にて演算される、風検出部2101からの風雑音レベルに対する風雑音ゲイン[0]～[511]の周波数関係を示す。風検出部2101からの風雑音レベルが大きい程、風雑音ゲインはマイナス側へシフトし、点線の示すカットオフ周波数を高周波帯域へシフトする。前記カットオフの周波数により風雑音ゲイン[0]～[511]は決定される。

【0166】

次に、本実施形態の音声入力部102でのMch/Sch選択部213の動作について、図18(a)、(b)を用いて説明する。

【0167】

図18(a)は、Mch/Sch選択部2131にて合成される、メインマイク205aからの周波数スペクトルデータMain[0]～[511](図示のMain ch)とサブマイク205bからの周波数スペクトルデータSub[0]～[511](図示のSub ch)との、風雑音レベルに応じた合成比率と周波数の関係を示している。

【0168】

ここで図3(a)は、図3(b)記載のメインマイク205aとサブマイク205bのメカ構成に対応した実施形態を示す。ここでMch/Sch選択部2131は、図18(a)に示すように、風雑音レベルに基づき、Main chを1.0から0.5の比率で、また、Sub chを0から0.5の比率で合成する。

【0169】

風雑音レベルが大きい程、Main chにおいては1.0から0.5へ合成比率を下げ、Sub chにおいては0から0.5へ合成比率を上げ、Main chとSub chを合成するクロスオーバーの周波数(合成の上限周波数)を上げていく。そして、Mch/Sch選択部2131は、風雑音レベルに依存する上限周波数以下ではMain chとSub chとを図示の比率で合成し、上限周波数を上回る周波数ではMain chを選択して出力する。風雑音レベルが0の場合は、Schの合成比率は0となる。ここで、図3(b)で記載の通り、サブマイク205bへのマイク穴の径は、メインマイク205aへのマイク穴の径に対して小さく、1/3に縮小された構成をとる。よって、サブマイク205bへの風雑音の影響度はメインマイク205aよりも弱い。よって、風検出部2101からの風雑音レベルに応じて、Mchに対しSchを合成する事により、風雑音の軽減に効果を果たす。

【0170】

次に図18(b)は、図15に示すようにサブマイク205bに対し、風防材102-3を構成したメカ構成に対応した実施形態を示す。ここでMch/Sch選択部2131は風雑音レベルから、Mchを1.0から0の比率で、Schを0から1.0の比率で合成する。つまり、風雑音レベルが大きい程、Mchにおいては1.0から0へ合成比率を下げ、Schにおいては0から1.0へ合成比率を上げ、MchとSchを合成するクロスオーバーの周波数を上げていく。風雑音レベルが0の場合は、Schの合成比率は0となる。ここで図3(a)や図15で記載の通り、サブマイク205bへのマイク穴の径は、メインマイク205aへのマイク穴の径に対して小さく、1/3に縮小された構成をとる。かつ、外装部102-1とサブマイクブッシュ102-2bで構成されるサブマイク205b前面の空間には、風防材102-3を設けている。よって、サブマイク205bへの風雑音の影響度はメインマイク205aに対し、更に小さくできる。よって、風検出部2101からの風雑音レベルに応じて、MchからSchに切り替えていく事で、風雑音の軽減に効果を果たす。

【0171】

ここで、本実施形態の音声入力部102でのステレオ抑制部2113の具体的動作につ

10

20

30

40

50

いて図19、図20を用いて説明する。

【0172】

図19は、ステレオ抑制部2113について、駆動騒音検出時と風雑音検出時に応じて、ステレオ効果の強調に用いる強調係数を変更するタイミングチャートを示している。図19において、Main[N]は、周波数Nポイント目のMchの振幅スペクトルデータを示す。また、駆動騒音検出信号は、駆動検出部2095により駆動騒音を検出した事を示す検出信号を示す。また、風雑音検出信号は、風検出部2101により風雑音を検出した事を示す風雑音レベル(予め設定された閾値以上の風雑音レベル)を示す。GainL[N]、GainR[N]は、ステレオゲイン演算処理部2112により決定された周波数Nポイント目のMchの振幅スペクトルに加算するステレオのLch及び、Rchのゲインを示す。

10

【0173】

ステレオ抑制部2113は、Mch-Sch演算部2091からの駆動騒音を検出したことを示す検出信号を受けて、強調係数を0にする。また、風検出部2101からの風雑音を検出した事を示す風雑音レベルを受けて、強調係数を周波数に応じて0にする。

【0174】

ここで、タイミングt1からt2の期間を着目すると、Main[N]の振幅スペクトルは大きく変動しており、Mch-Sch演算部2091からの検出信号は、駆動騒音有りとして検出を示している。この期間、GainL[N]、GainR[N]は0に固定されている。つまりステレオ抑制部2113が強調係数を0にした事を表している。また、タイミングt3からt4の期間を着目すると、Main[N]の振幅スペクトルは大きく変動しており、風検出部2101からの風雑音検出信号は検出を示している。この期間、GainL[N]、GainR[N]は0に固定されている。つまりステレオ抑制部2113が強調係数を0にした事を表している。

20

【0175】

図20は、風検出部2101からの風雑音レベル検出時において、Mch/Sch選択部2131にてメインマイク205aからの周波数スペクトルMain[0]~[511]とサブマイク205bからの周波数スペクトルSub[0]~[511]が合成される比率と、周波数に対し、ステレオ抑制部2113にて強調係数を0にする周波数の関係を示す図である。ここでMch/Sch選択部2131は風雑音レベルから、風雑音レベルが大きい程、Mchにおいては1.0から0.5へ合成比率を下げ、Schにおいては0から0.5へ合成比率を上げ、MchとSchを合成するクロスオーバーの周波数を上げていく。風雑音レベルの場合、前記クロスオーバーの周波数は500Hzである。これに対しステレオ抑制部2113は、前記クロスオーバーの周波数よりも高い周波数750Hzまで強調係数を0に固定する。ステレオ抑制部2113は、風検出部2101からの風雑音レベルが大きい程、強調係数を0に固定する周波数を上げていく。GainL、GainRのステレオゲインによる強調により、風雑音も強調されることを防ぐ。

30

【0176】

ここで、本実施形態の音声入力部102での駆動音減算量積分器2097、風雑音減算量積分器2103、右ゲイン積分器2114、左ゲイン積分器2115の動作について図21を用いて説明する。

40

【0177】

図21は、周波数Nポイント目のMchの振幅スペクトルデータについての、それぞれ決定される駆動騒音除去ゲインNC_GAIN[N]、風雑音減算量WC_GAIN[N]、Lch生成用ステレオゲインL_GAIN[N]、Rch生成用ステレオゲインR_GAIN[N]のそれぞれに対する時定数を示す。これらは、駆動音減算量積分器2097、風雑音減算量積分器2103、左ゲイン積分器2114、右ゲイン積分器2115により決定される。駆動騒音減算量積分器の時定数は、右ゲイン積分器2115、左ゲイン積分器2114の時定数に対して遅く、風雑音減算量積分器の時定数は右ゲイン積分器2115、左ゲイン積分器2114の時定数に対して遅い。駆動騒音と風雑音は、それぞれ

50

駆動騒音成分であり、時系列でのばらつきも大きく、時定数を遅くしてそれぞれ駆動騒音減算と風雑音減算の追従を遅くすることで前記ばらつきを抑える。また、ステレオゲインについては、時定数を早くすることで、発音する被写体の移動に対する追従を早くする。

【0178】

本実施形態においては、2系統の音声が入力される場合について説明したが、それ以上のチャンネル数であっても適用することができる。

【0179】

また、本実施形態においては、撮像装置について説明したが、本実施形態の音声入力部102の音声処理は、外部の音声を記録、または入力するような装置つまり、音声記録装置であればどのような装置であっても適用することができる。例えば、ICレコーダ、携帯電話等に適用しても良い。

10

【0180】

また、実施形態では、図6に示す構成をハードウェアにより実現する例を説明したが、例えば、同図のマイクやAD変換部等を除く処理部の多くを、プロセッサが実行するプロセスやサブルーチン等のプログラムで実現しても構わない。

【0181】

(その他の実施例)

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

20

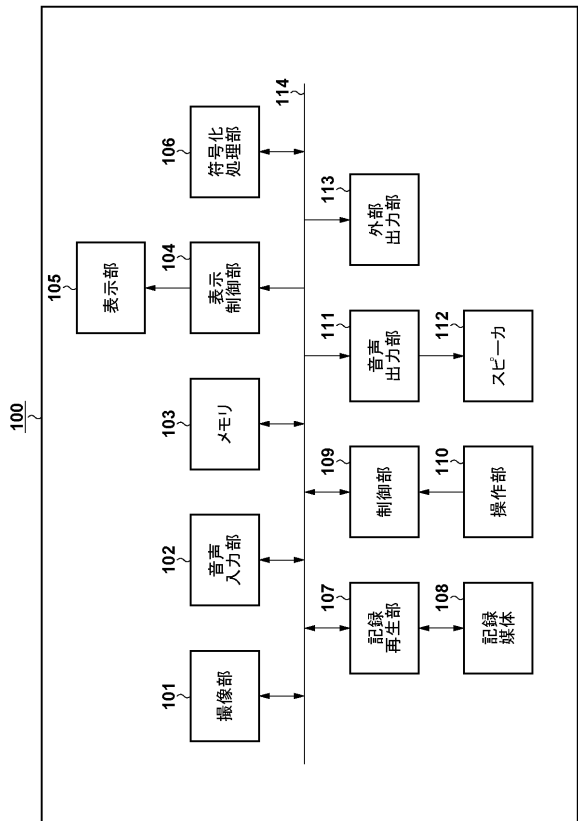
【符号の説明】

【0182】

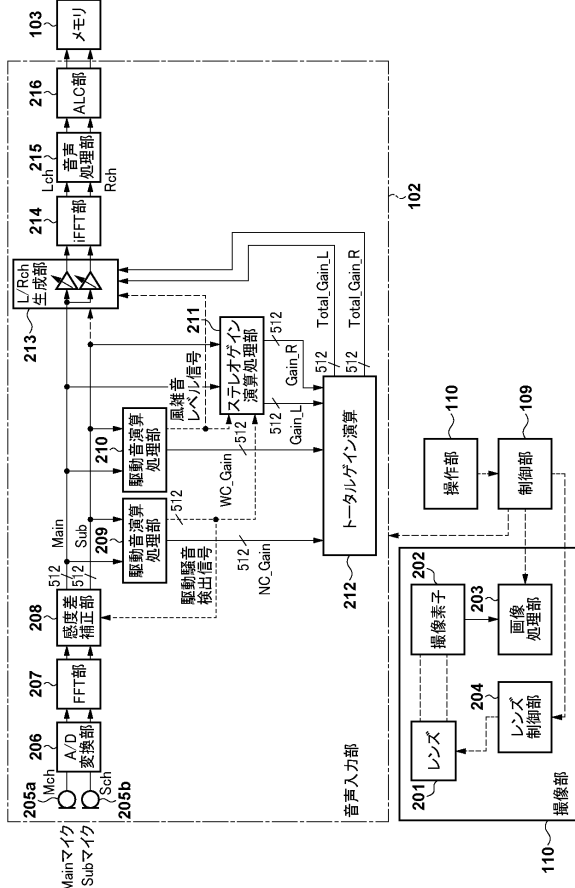
100...撮像装置、101...撮像部、102...音声入力部、103...メモリ、104...表示制御部、105...表示部、106...符号化処理部、107...記録再生部、108...記録媒体、109...制御部、110...操作部、111...音声出力部、112...スピーカ、113...外部出力部、114...データバス、201...光学レンズ、202...撮像素子、203...画像処理部、204...光学レンズ制御部、205...マイク、205a...メインマイク、205b...サブマイク、206...A/D変換部、207...FFT部、208...感度差補正部、209...駆動音演算処理部、210...風雑音演算処理部、211...ステレオゲイン演算処理部、212...トータルゲイン演算部、213...L/Rch生成部、214...IFFT部、215...音声処理部、216...ALC部、102-1...外装部、102-2a...メインマイクブッシュ、102-2b...サブマイクブッシュ、102-3...風防材

30

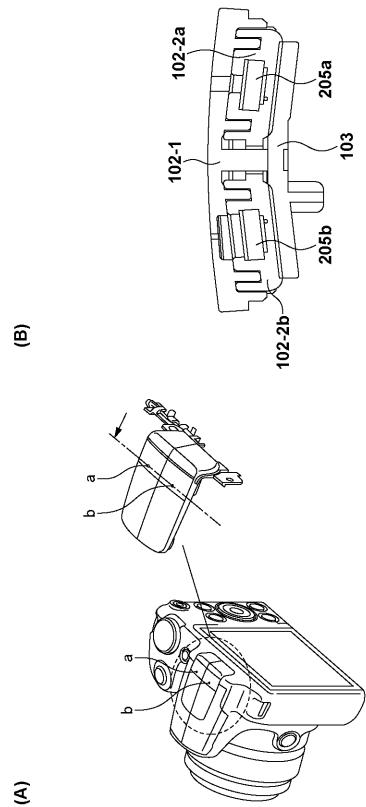
【図1】



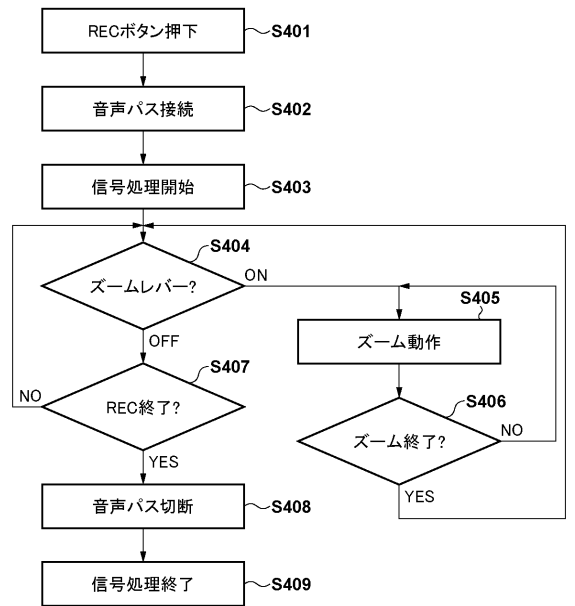
【図2】



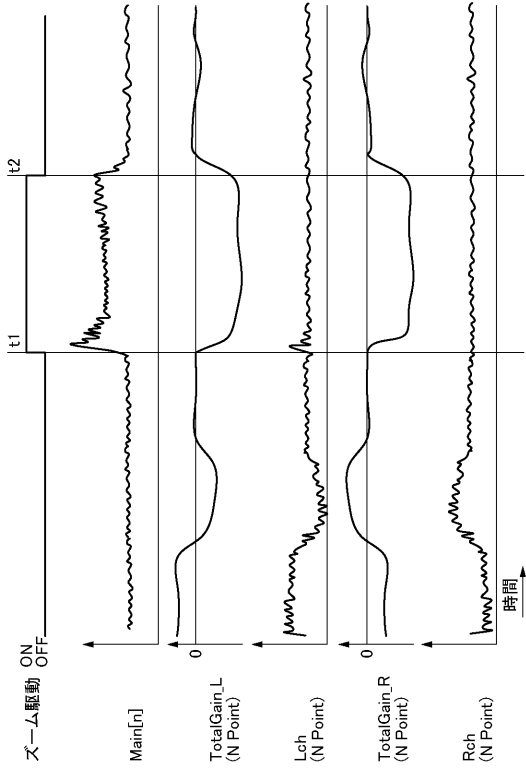
【図3】



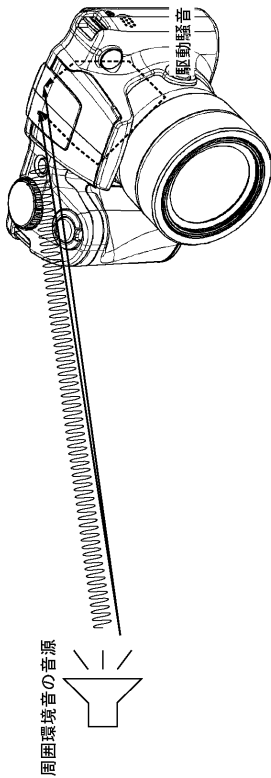
【図4】



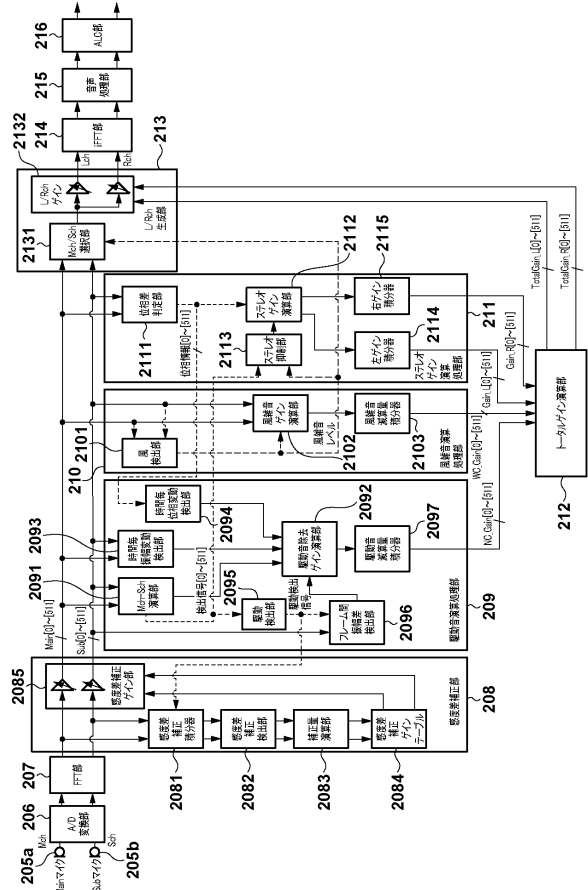
【 図 5 】



【 図 7 】

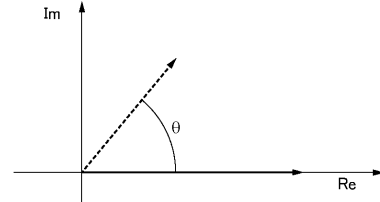


【 図 6 】

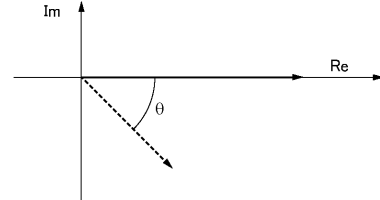


【 図 8 】

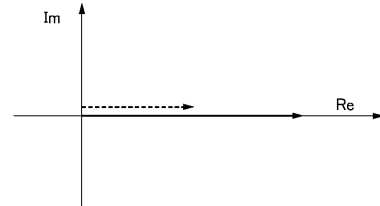
(a) Mainマイク側(右側)から周囲環境音になっている時



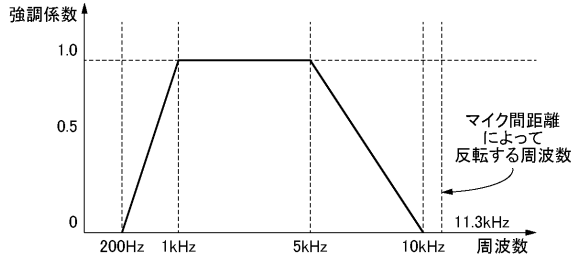
(b) Subマイク側(左側)から周囲環境音になっている時



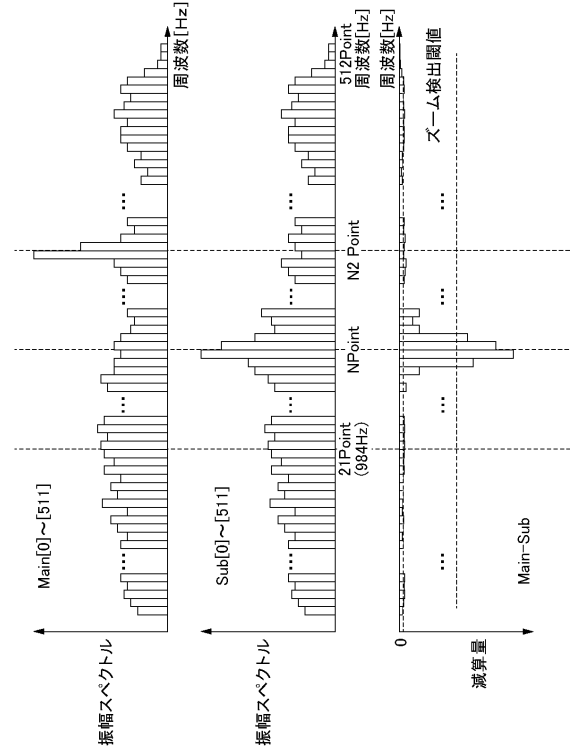
(c) 光軸中心から周囲環境音になっている時



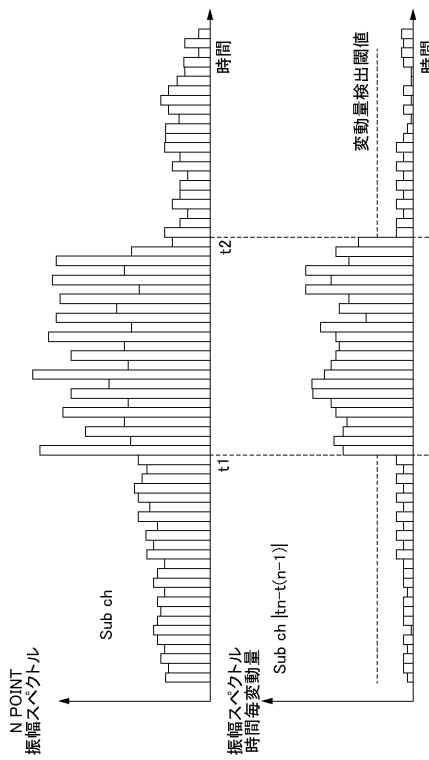
【図9】



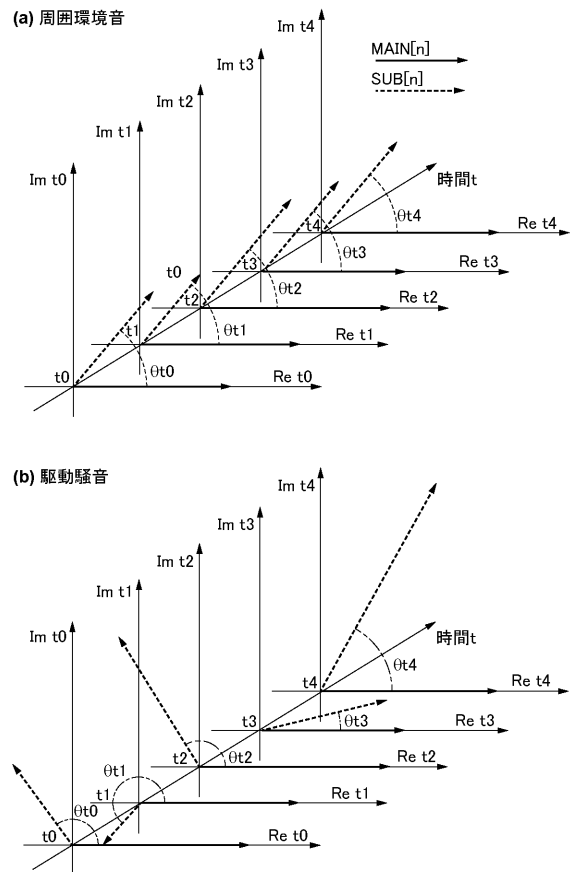
【図10】



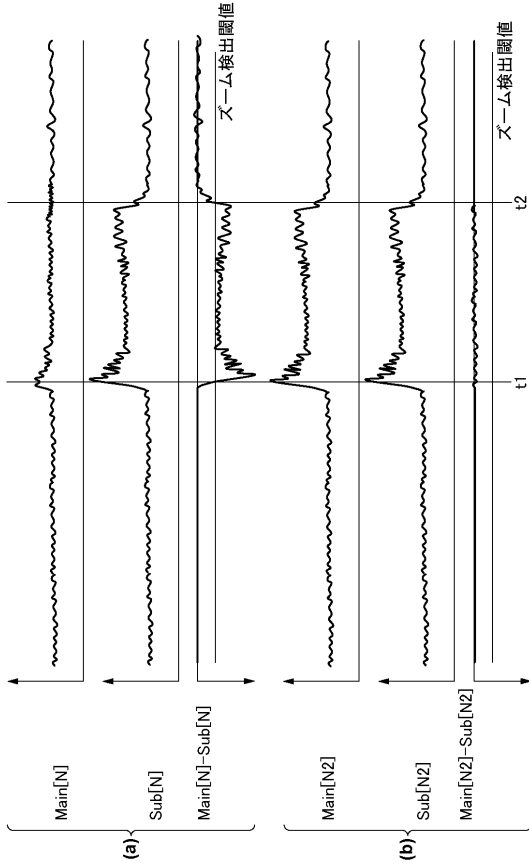
【図11】



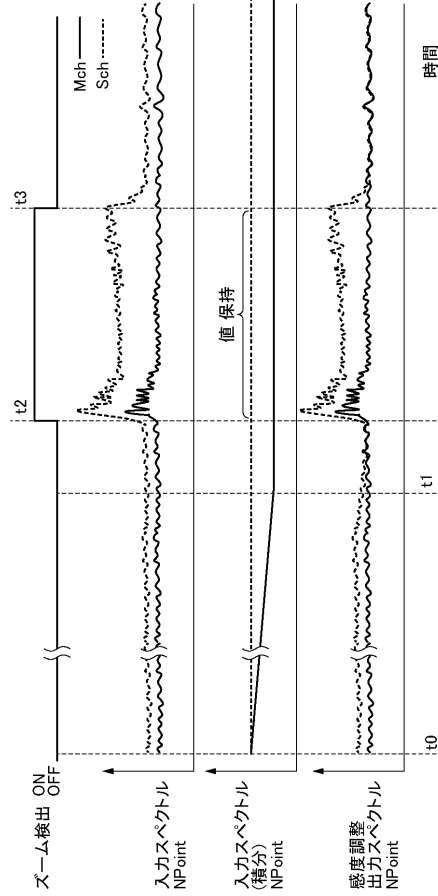
【図12】



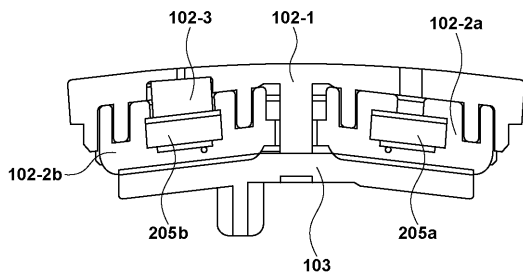
【図13】



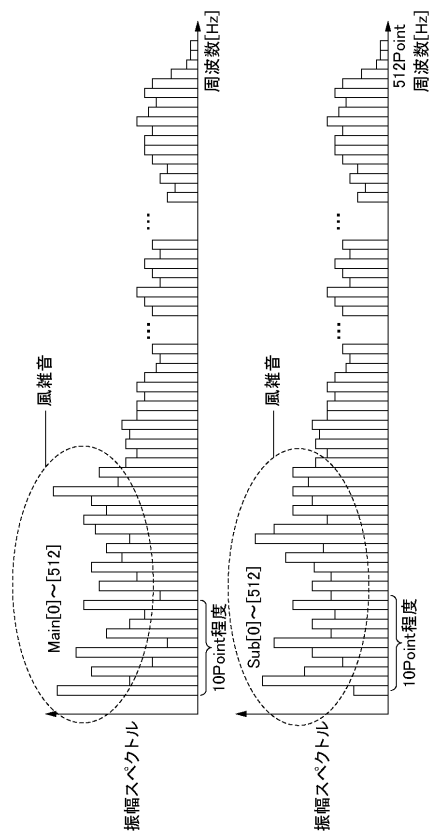
【図14】



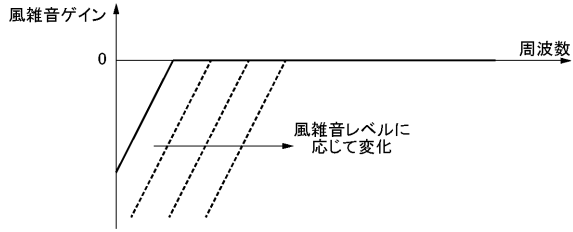
【図15】



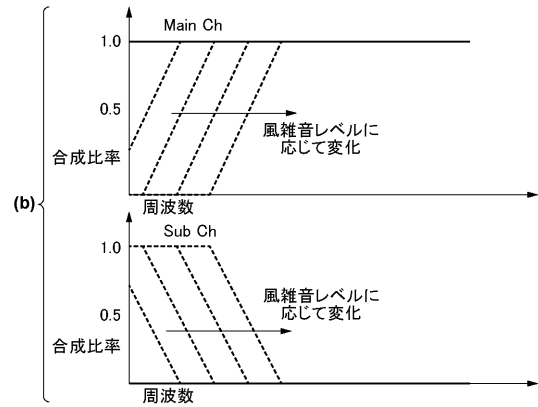
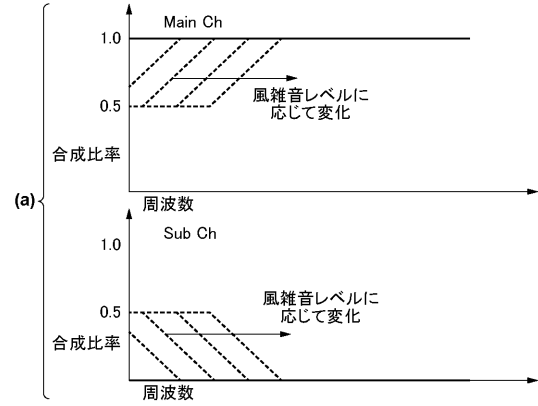
【図16】



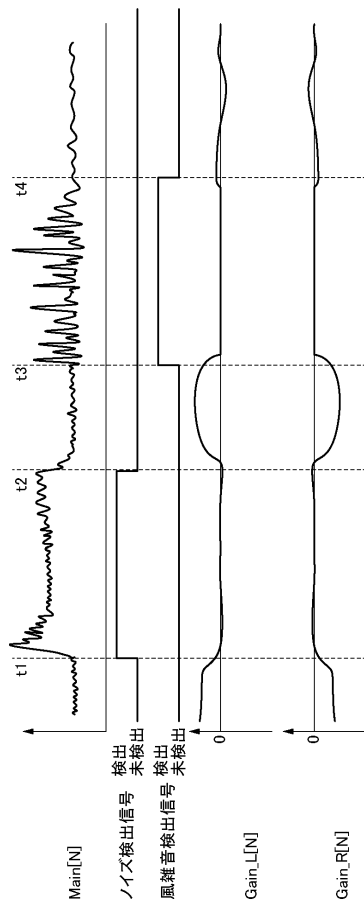
【図 17】



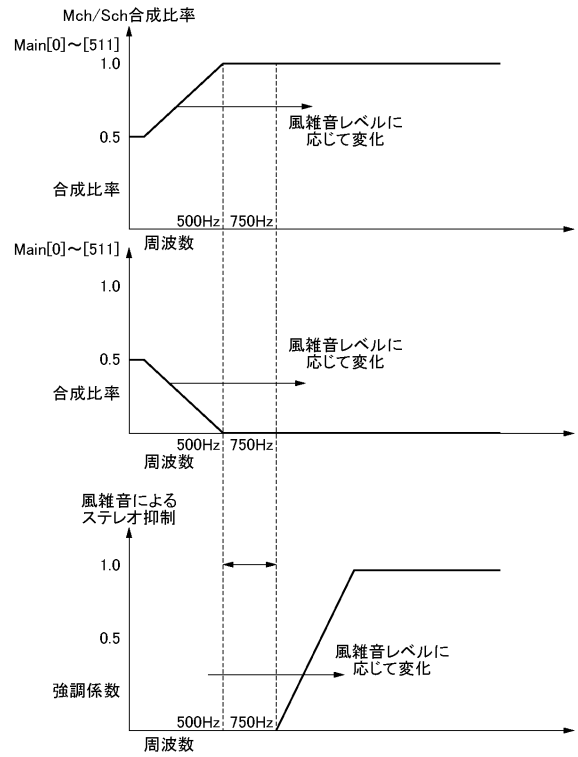
【図 18】



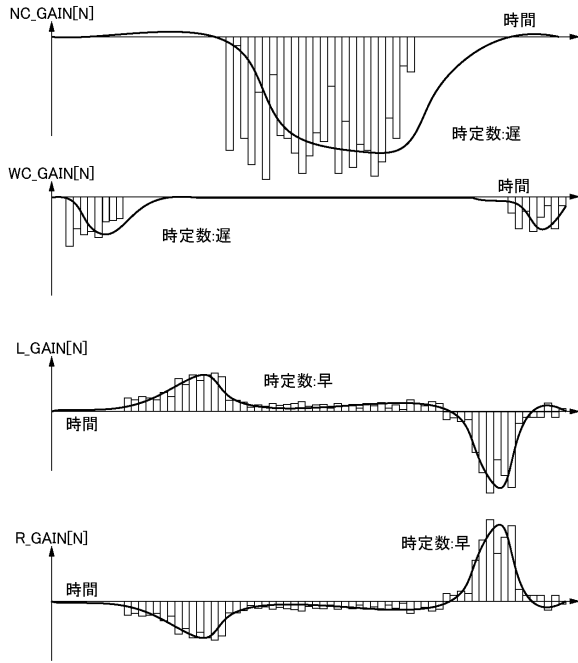
【図 19】



【図 20】



【 図 2 1 】



フロントページの続き

審査官 齊田 寛史

- (56)参考文献 特開2012-3021(JP,A)
特開2015-114444(JP,A)
特開2012-142745(JP,A)
特開2013-110629(JP,A)
特開2016-054462(JP,A)
米国特許出願公開第2012/0154610(US,A1)

(58)調査した分野(Int.Cl., DB名)

H04R 1/02
H04R 1/08
H04R 3/00
H04R 5/027
G10L 21/0208
G10L 21/0232