

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-181761

(P2017-181761A)

(43) 公開日 平成29年10月5日(2017.10.5)

(51) Int.Cl.	F I	テーマコード (参考)
G10L 21/0232 (2013.01)	G10L 21/0232	5D220
H04R 3/00 (2006.01)	H04R 3/00 320	
G10L 21/0272 (2013.01)	G10L 21/0272 100A	
G10L 21/028 (2013.01)	G10L 21/0272 100Z	
	G10L 21/028 B	

審査請求 未請求 請求項の数 11 OL (全 14 頁)

(21) 出願番号 特願2016-68405 (P2016-68405)  
 (22) 出願日 平成28年3月30日 (2016. 3. 30)

(71) 出願人 000000295  
 沖電気工業株式会社  
 東京都港区虎ノ門一丁目7番12号  
 (74) 代理人 100180275  
 弁理士 吉田 倫太郎  
 (74) 代理人 100161861  
 弁理士 若林 裕介  
 (74) 代理人 100090620  
 弁理士 工藤 宣幸  
 (72) 発明者 藤枝 大  
 東京都港区虎ノ門一丁目7番12号 沖電  
 気工業株式会社内  
 Fターム(参考) 5D220 BA06

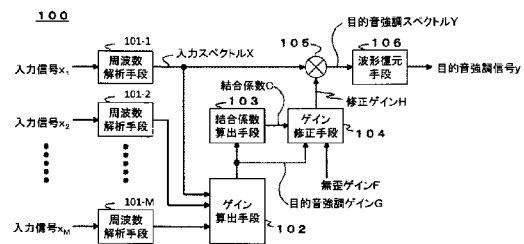
(54) 【発明の名称】 信号処理装置及びプログラム、並びに、ゲイン処理装置及びプログラム

(57) 【要約】 (修正有)

【課題】 入力信号に対して目的音を強調する処理を行う際の音質をより向上させる。

【解決手段】 音響信号処理装置100は、複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調する所定の帯域ごとの第1のゲインを得るゲイン算出手段102と、第1のゲインに基づいて帯域ごとの結合係数を得る結合係数算出手段103と、結合係数に基づいて第1のゲインと帯域ごとに所定の定数値で与えられる第2のゲインとを凸結合することで第1のゲインを修正した帯域ごとの第3のゲインを得るゲイン修正手段104と、複数の入力スペクトルの1つに帯域ごとに第3のゲインを乗じて目的音強調スペクトルを得る目的音強調手段105とを有することを特徴とする。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調する所定の帯域ごとの第 1 のゲインを得るゲイン算出手段と、

前記第 1 のゲインに基づいて前記帯域ごとの結合係数を得る結合係数算出手段と、

前記結合係数に基づいて前記第 1 のゲインと前記帯域ごとに所定の定数値で与えられる第 2 のゲインとを凸結合することで前記第 1 のゲインを修正した前記帯域ごとの第 3 のゲインを得るゲイン修正手段と、

前記複数の入力スペクトルの一部又は全部に前記帯域ごとに前記第 3 のゲインを乗じて目的音強調スペクトルを得る目的音強調手段と

を有することを特徴とする音響信号処理装置。

10

**【請求項 2】**

前記結合係数算出手段は、前記第 1 のゲインを周波数方向と時間方向とに平滑化して、前記結合係数を得ることを特徴とする請求項 1 に記載の音響信号処理装置。

**【請求項 3】**

前記結合係数算出手段は、前記第 1 のゲインを周波数方向に移動平均法によって平滑化し、さらに時間方向に時定数フィルタによって平滑化して前記結合係数を得ることを特徴とする請求項 2 に記載の音響信号処理装置。

**【請求項 4】**

前記結合係数算出手段は、前記第 1 のゲインを周波数方向に平滑化する際に、平均する範囲を臨界帯域幅に比例させることを特徴とする請求項 3 に記載の音響信号処理装置。

20

**【請求項 5】**

前記第 1 のゲインを周波数方向と時間方向とに平滑化して前記第 2 のゲインを取得する平滑化手段をさらに備え、

前記ゲイン修正手段は、前記平滑化手段が取得した前記第 2 のゲインを用いて前記第 3 のゲインを得る

ことを特徴とする請求項 1 ~ 4 のいずれかに記載の音響信号処理装置。

**【請求項 6】**

前記平滑化手段は、前記第 1 のゲインを周波数方向に移動平均法によって平滑化し、さらに時間方向に時定数フィルタによって平滑化することで、前記第 2 のゲインを得ることを特徴とする請求項 5 に記載の音響信号処理装置。

30

**【請求項 7】**

前記平滑化手段は、前記第 1 のゲインを周波数方向に平滑化する際に、平均する範囲を臨界帯域幅に比例させることを特徴とする請求項 6 に記載された音響信号処理装置。

**【請求項 8】**

前記目的音強調手段は、前記複数の入力スペクトルのうちの少なくとも 2 つについて、それぞれ前記帯域ごとに前記第 3 のゲインを乗じて複数の目的音強調スペクトルを得ることを特徴とする請求項 1 ~ 7 のいずれかに記載の音響信号処理装置。

**【請求項 9】**

複数の入力音響信号に基づく複数の入力スペクトルに基づいて目的音成分を強調した所定の帯域ごとの第 1 のゲインについて、前記帯域ごとの結合係数を得る結合係数算出手段と、

40

前記結合係数に基づいて前記第 1 のゲインと前記帯域ごとに所定の定数値で与えられる第 2 のゲインとを凸結合することで、前記第 1 のゲインを修正した前記帯域ごとの第 3 のゲインを得るゲイン修正手段と

を有することを特徴とするゲイン処理装置。

**【請求項 10】**

コンピュータを、

複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調する所定の帯域ごとの第 1 のゲインを得るゲイン算出手段と、

50

前記第 1 のゲインに基づいて前記帯域ごとの結合係数を得る結合係数算出手段と、  
前記結合係数に基づいて前記第 1 のゲインと前記帯域ごとに所定の定数値で与えられる  
第 2 のゲインとを凸結合することで前記第 1 のゲインを修正した前記帯域ごとの第 3 のゲ  
インを得るゲイン修正手段と、

前記複数の入力スペクトルの一部又は全部に前記帯域ごとに前記第 3 のゲインを乗じて  
目的音強調スペクトルを得る目的音強調手段と

して機能させることを特徴とする音響信号処理プログラム。

【請求項 11】

コンピュータを、

複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調した  
所定の帯域ごとの第 1 のゲインについて、前記帯域ごとの結合係数を得る結合係数算出  
手段と、

前記結合係数に基づいて前記第 1 のゲインと前記帯域ごとに所定の定数値で与えられる  
第 2 のゲインとを凸結合することで前記第 1 のゲインを修正した前記帯域ごとの第 3 のゲ  
インを得るゲイン修正手段と

して機能させることを特徴とするゲイン処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音響信号処理装置及びプログラム、並びに、ゲイン処理装置及びプログラム  
に関し、例えば、複数の音源が存在する環境下で、特定の音源のみ強調し収録する音声処  
理装置（音響信号処理装置）に適用し得る。

【背景技術】

【0002】

複数の音源が存在する環境下において、ある目的の音源を抽出する技術として、複数の  
マイクロホンを用いた音源分離や、マイクロホンを直線や平面、球面上等に配置したマイ  
クロホンアレイを用いたビームフォーマや、ヌルフォーマ等がある。特に、目的の音源以  
外の音源が非定常である場合や、複数ある場合には、単一のマイクロホンを用いたノイズ  
サプレッサによる目的音源の抽出は難しく、2つ以上のマイクロホンを用いることが必須  
となる。

【0003】

上述したマイクロホンアレイを用いたビームフォーマとは、ある特定の方向の音のみ強  
調し収録する技術である。ビームフォーマとは、各マイクロホンに到達する信号の時間差  
を利用して指向性を形成する技術である。

【0004】

ビームフォーマには、加算型と減算型という2つの種類がある。加算型ビームフォーマ  
に比べて、減算型ビームフォーマはより少ないマイクロホン数で鋭い指向性を形成でき  
るという利点がある。

【0005】

図 4 は、マイクロホン数が 2 個の場合の減算型ビームフォーマを適用した従来の收音装  
置 P S の構成について示したブロック図である。

【0006】

減算型ビームフォーマを適用した收音装置 P S では、まず遅延器により目的とする方向  
に存在する目的音が各マイクロホン M 1 - M 2 に到来する信号の時間差を算出し、遅延を  
加えることにより目的音の位相を合わせる。時間差は ( 1 ) 式により算出される。( 1 )  
式において、d はマイクロホン間の距離、c は音速、 $\theta$  は遅延量をそれぞれ表している  
。また、( 1 ) 式において、 $\theta$  は、2つのマイクロホン M 1 - M 2 の位置を結んだ直線  
に対する垂直方向から目的方向への角度を表している。

【0007】

收音装置 P S の遅延器は、例えば、死角がマイクロホン M 1 とマイクロホン M 2 の中心

10

20

30

40

50

に対し、マイクロホン M 1 の方向に存在する場合、マイクロホン M 1 の入力信号  $x_1(t)$  に対し遅延処理を行う。その後、收音装置 P S の減算器では、(2)式に従い処理が行われる。

【0008】

收音装置 P S の減算器では、周波数領域でも減算処理を同様に行うことができ、その場合(2)式は(3)式のように変更される。

【数1】

$$\tau_L = \frac{d \sin \theta_L}{c} \quad \dots (1)$$

10

$$a(t) = x_2(t) - x_1(t - \tau_L) \quad \dots (2)$$

$$A(\omega) = X_2(\omega) - e^{-j\omega\tau_L} X_1(\omega) \quad \dots (3)$$

【0009】

例えば、 $\theta_L = \pm \pi/2$  の場合、(3)式により形成される指向性は、図5に示すようなカーゴイド型の単一指向性となる。また、例えば、 $\theta_L = 0$  の場合、(3)式により形成される指向性は、図6のような8の字型の双指向性となる。以下では、入力信号から単一指向性を形成するフィルタを「単一指向性フィルタ」と呼び、双指向性を形成するフィルタを「双指向性フィルタ」と呼ぶものとする。

20

【0010】

また、従来の收音装置では、スペクトル減算法 (Spectral Subtraction: SS) を用いることで、双指向性の死角方向に強い指向性を形成することができる。従来の收音装置において、SSによる指向性の形成は、例えば、以下の(4)式を用いて行うことができる。(4)式では、マイクロホン M 1 の入力信号  $X_1(\omega)$  を用いているが、マイクロホン M 2 の入力信号  $X_2(\omega)$  でも同様の効果を得ることができる。また、(4)式において、 $\alpha$  は SS の強度を調整するための係数であり、 $\beta$  は最大抑圧量である。

30

【数2】

$$|Y(\omega)| = \max \left[ |X_1(\omega)| - \alpha |A(\omega)|, \beta |X_1(\omega)| \right] \quad \dots (4)$$

【0011】

非特許文献1では、単一指向性及び双指向性の2種類の指向性を利用することで、目的方向に強い指向性を形成する手法を提案している。非特許文献1の手法では、まず目的方向に対して左右(垂直方向)に死角を向ける2つの単一指向性と、目的方向と同じ方向に死角を向ける双指向性を形成する。その後、非特許文献1の手法では、2つの単一指向性フィルタ出力のうち、パワーの小さい方を選択し、選択した単一指向性フィルタ出力から双指向性フィルタ出力をSSすることにより、非目的音を抑圧し、目的音を強調する。非特許文献1の手法では、周囲に存在する非目的音を抑圧し、目的音のみを強調できる。特に、SSを用いる場合、使用するマイクロホン数が少なく、且つ低演算量で鋭い指向性を形成できる。

40

【先行技術文献】

【非特許文献】

【0012】

【非特許文献1】矢頭隆、森戸誠、山田圭、小川哲司、“正方形マイクロホンアレイによる音源分離技術”、情報処理、Vol. 51、No. 11、2010

50

**【発明の概要】****【発明が解決しようとする課題】****【0013】**

しかしながら、従来の目的音強調処理において、SSが適用される場合、通常フロアリング処理を行うが、このとき（時間周波数領域における）周囲の周波数成分に比べてパワーの大きい孤立周波数成分ができてしまうと、不快なミュージカルノイズが発生する原因となってしまう。従来の目的音強調処理では、ビームフォーマに限らず、音源分離やヌルフォーマでも、除去しきれずに歪んだ状態で非目的音が残留したり、特に周波数領域で処理した場合には上述のSSと同様にミュージカルノイズが発生したりすることが多い。

**【0014】**

以上のような問題に鑑みて、入力信号に対して目的音を強調する処理を行う際の音質をより向上（例えば、非目的音の残留音やミュージカルノイズのような不快な音を軽減）させることができる音響信号処理装置及びゲイン処理装置が望まれている。

**【課題を解決するための手段】****【0015】**

第1の本発明の音響信号処理装置は、（1）複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調する所定の帯域ごとの第1のゲインを得るゲイン算出手段と、（2）前記第1のゲインに基づいて前記帯域ごとの結合係数を得る結合係数算出手段と、（3）前記結合係数に基づいて前記第1のゲインと前記帯域ごとに所定の定数値で与えられる第2のゲインとを凸結合することで前記第1のゲインを修正した前記帯域ごとの第3のゲインを得るゲイン修正手段と、（4）前記複数の入力スペクトルの一部又は全部に前記帯域ごとに前記第3のゲインを乗じて目的音強調スペクトルを得る目的音強調手段とを有することを特徴とする。

**【0016】**

第2の本発明のゲイン処理装置は、（1）複数の入力音響信号に基づく複数の入力スペクトルに基づいて目的音成分を強調した所定の帯域ごとの第1のゲインについて、前記帯域ごとの結合係数を得る結合係数算出手段と、（2）前記結合係数に基づいて前記第1のゲインと前記帯域ごとに所定の定数値で与えられる第2のゲインとを凸結合することで、前記第1のゲインを修正した前記帯域ごとの第3のゲインを得るゲイン修正手段とを有することを特徴とする。

**【0017】**

第3の本発明の音響信号処理プログラムは、コンピュータを、（1）複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調する所定の帯域ごとの第1のゲインを得るゲイン算出手段と、（2）前記第1のゲインに基づいて前記帯域ごとの結合係数を得る結合係数算出手段と、（3）前記結合係数に基づいて前記第1のゲインと前記帯域ごとに所定の定数値で与えられる第2のゲインとを凸結合することで前記第1のゲインを修正した前記帯域ごとの第3のゲインを得るゲイン修正手段と、（4）前記複数の入力スペクトルの一部又は全部に前記帯域ごとに前記第3のゲインを乗じて目的音強調スペクトルを得る目的音強調手段として機能させることを特徴とする。

**【0018】**

第4の本発明のゲイン処理プログラムは、コンピュータを、（1）複数の入力音響信号に基づく複数の入力スペクトルに基づいて、目的音成分を強調した所定の帯域ごとの第1のゲインについて、前記帯域ごとの結合係数を得る結合係数算出手段と、（2）前記結合係数に基づいて前記第1のゲインと前記帯域ごとに所定の定数値で与えられる第2のゲインとを凸結合することで前記第1のゲインを修正した前記帯域ごとの第3のゲインを得るゲイン修正手段として機能させることを特徴とする。

**【発明の効果】****【0019】**

本発明によれば、入力信号に対して目的音を強調する処理を行う際の音質をより向上（例えば、信号の歪みやミュージカルノイズの抑制）させることができる。

10

20

30

40

50

## 【図面の簡単な説明】

【0020】

【図1】第1の実施形態に係る音声処理装置の機能的構成について示したブロック図である。

【図2】第2の実施形態に係る音声処理装置の機能的構成について示したブロック図である。

【図3】第3の実施形態に係る音声処理装置の機能的構成について示したブロック図である。

【図4】従来の收音装置の構成について示した説明図である。

【図5】従来の単一指向性フィルタにより形成される指向特性の一例を説明する説明図である。

【図6】従来の双指向性フィルタにより形成される指向特性の一例を説明する説明図である。

## 【発明を実施するための形態】

【0021】

(A)第1の実施形態

以下、本発明による音響信号処理装置及びプログラム、並びに、ゲイン処理装置及びプログラムの第1の実施形態を、図面を参照しながら詳述する。以下では、本発明の音響信号処理装置及びゲイン処理装置を音声処理装置に適用した例について説明する。

【0022】

(A-1)第1の実施形態の構成

図1は、この実施形態の音声処理装置100の全体構成を示すブロック図である。

【0023】

音声処理装置100は、目的音源を含むM個の入力信号 $x_1 \sim x_M$ （時間領域の入力信号）を用いて、所定の入力信号について目的音強調処理を行い、目的音強調信号 $y$ （時間領域の出力信号）を生成するものである。音声処理装置100としては、例えば、通信端末（例えば、携帯電話端末やスマートホン等の音声通話が可能な通信端末）、オーディオ機器、音声認識装置等の目的音を收音して種々の音響信号処理を行う装置に適用することができる。

【0024】

なお、入力信号の数は限定されないものであるが、例えば、音声処理装置100が2つのマイクロホンにより構成されるマイクロホンアレイから入力される入力信号を処理する場合、 $M=2$ となる。

【0025】

音声処理装置100は、周波数解析手段101-1~101-Mと、ゲイン算出手段102と、結合係数算出手段103と、ゲイン修正手段104と、乗算手段105と、波形復元手段106とを有する。

【0026】

音声処理装置100は、例えば、プロセッサ及びメモリを有するコンピュータにプログラム（実施形態に係る音響信号処理プログラムを含む）をインストールすることにより実現するようにしてもよい。

【0027】

第1の実施形態では、結合係数算出手段103、及びゲイン修正手段104により、本発明のゲイン処理装置が構成されている。本発明のゲイン処理装置についても、例えば、プロセッサ及びメモリを有するコンピュータにプログラム（実施形態に係るゲイン処理プログラム）をインストールすることにより実現するようにしてもよい。

【0028】

第1の実施形態の音声処理装置100は、周波数領域で目的音を強調する際に、歪みやミュージカルノイズ等を低減するものである。第1の実施形態の音声処理装置100において、目的音強調に適用する手法（方式）としては、例えば、(4)式のビームフォーマ

10

20

30

40

50

手法（以下、「SSBF」と呼ぶ）や、非特許文献1に記載のビームフォーマ手法（以下、「CSSBF」と呼ぶ）や、独立成分分析や独立ベクトル分析を用いて音源を分離する方法（以下、「ICABBS」と呼ぶ）や、マイクロホン間の振幅比や位相差に注目して時間周波数点ごとに0又は1のゲインを乗じて音源を分離する手法（以下「MaskBSS」と呼ぶ）等（以下、これらの手法を総称して「従来手法」と呼ぶ）を適用することができる。

【0029】

ゲイン算出手段102では、従来手法を用いて、特定の音源、又は特定の方向を強調する帯域ごとのゲイン（第1のゲイン）を算出する。第1のゲインを入力スペクトルの1つに帯域ごとに乗じると、従来手法によるスペクトルが得られるが、上述した通り、歪んだ非目的音が残留したり、ミュージカルノイズが発生したりしてしまう場合がある。

10

【0030】

結合係数算出手段103は、第1のゲインを周波数方向と時間方向に平滑化して、帯域ごとの結合係数（範囲0.0～1.0）を算出する。第1のゲインは、目的音が存在すると判断した帯域では1.0に近い値を取り、目的音が存在しないと判断した帯域では0.0に近い値を取る。すなわち、第1のゲインは目的音らしさの指標となっており、これを平滑化した結合係数も目的音らしさを表している。平滑化する理由は後述する。

【0031】

第1の実施形態においては、ゲイン修正手段104は、歪みやミュージカルノイズを生じる第1のゲインと、まったく歪みを生じない第2のゲインとを、結合係数によってファジーに切り替えることで、不快な雑音を生じない第3のゲインを算出する。第2のゲインは、帯域ごとに所定の定数で与えられる。当該所定の定数は、ある程度小さく（例えば、0.1程度）、且つ周波数方向に十分滑らかとするのが好適である（例えば、隣り合う帯域の比が±1dB以内）。局所的（例えば、区間100ミリ秒且つ帯域200Hz程度）に見れば、第2のゲインは「音の大きさ」を変化させているだけなので、もし第2のゲインを入力スペクトルの1つに帯域ごとに乗じると、まったく歪みを生じないが、目的音もろとも抑圧されたスペクトルが得られる。ある帯域において、第1のゲインをG（目的音強調ゲイン）、第2のゲインをF（無歪ゲイン）、結合係数をC、第3のゲインをH（修正ゲイン）とすると、ゲイン修正手段104は、(5)式を用いて第3のゲインHを算出することができる。すなわち、Cが目的音らしさを意味しているため、第3のゲインHを適用することは、目的音が優勢な帯域については第1のゲインGを適用し、目的音が優勢でない帯域については第2のゲインFを適用することになる。第3のゲインHにおいて、第1のゲインGと第2のゲインFの切り替えが急激に行われると、不快な歪みを生じてしまうが、Cは周波数方向と時間方向に平滑化されているので、切り替えは緩やかに（ファジーに）行われ、不快な歪みが抑制される。

20

30

【数3】

$$H = C \cdot G + (1 - C) \cdot F \quad \dots (5)$$

【0032】

乗算手段105（目的音強調手段）は、以上のようにして得られた第3のゲインを、入力スペクトルの1つに帯域ごとに乗じて、不快な歪みを生じない目的音強調スペクトルを得る。

40

【0033】

（A-2）第1の実施形態の動作

次に、上述した構成を有する第1の実施形態の音声処理装置100の動作について、図1を参照しながら説明する。

【0034】

周波数解析手段101-1～101-Mは、フーリエ変換に代表される任意の周波数解析手法、またはフィルタバンクに代表される任意の帯域分割手段によって、入力信号 $x_1$

50

$\sim x_M$  をそれぞれ  $K$  個の帯域に分割する。入力信号  $x_m$  から  $k$  番目の帯域に分割された信号を  $X_{k; m}$  と表し、 $K$  個の帯域すべてをまとめて入力スペクトル  $X_m$  と表す。そして、周波数解析手段 101-1 ~ 101-M は、得られた入力スペクトル  $X_1 \sim X_M$  をゲイン算出手段 102 に与えるとともに、所定の入力信号に基づく入力スペクトル（この実施形態では、 $X_1$  とする）を乗算手段 105 に与える。なお、乗算手段 105 に与えられる入力スペクトルは 1 番目の入力スペクトル  $X_1$  としたが、これに限定されるものではなく、1 ~  $M$  番目のいずれの入力スペクトルを乗算手段 105 に与えても良い。

【0035】

ゲイン算出手段 102 は、周波数解析手段 101-1 ~ 101-M から与えられた入力スペクトル  $X_1 \sim X_M$  に基づいて、所定の目的音強調方法によって入力スペクトル  $X_1 \sim X_M$  に含まれる目的音成分又は目的音方向の推定を行い、得られた目的音強調ゲイン  $G$ （第 1 のゲイン）を結合係数算出手段 103 及びゲイン修正手段 104 に与える。

10

【0036】

目的音強調ゲイン  $G$  には、それぞれの帯域について、個別の値を有している。以下では、目的音強調ゲイン  $G$  における、任意の帯域  $k$  の値を  $G_k$  と表す。 $G_k$  は  $0.0 \sim 1.0$  の範囲の実数値となる。

【0037】

ゲイン算出手段 102 における目的音強調方法としては、例えば、SSBF、CSSBF、ICABBS、MaskBBS 等に代表される任意の処理（例えば、ビームフォーマ、ヌルフォーマ、音源分離等）を適用することができる。以下では、ゲイン算出手段 102 は、一例として、SSBF を用いた目的音強調ゲイン  $G$  の算出を行うものとして説明する。

20

【0038】

以下では、算出手段 102 では、SSBF のパラメータ  $M = 2$  とし、目的音源は正面方向（2つのマイクロホンを結んだ直線の鉛直方向）に存在しているものとする。ゲイン算出手段 102 は、まず双指向性フィルタ、すなわち（6）式によって、目的音を除去した妨害音スペクトル  $D$  を得る。そして、ゲイン算出手段 102 は、（7）式によって目的音強調ゲイン  $G$  を得る。ところで、（7）式の両辺に  $|X_{k; 1}|$  を乗じると、（4）式と同じ形となる。そこで、ゲイン算出手段 102 は、（7）式で得られたゲインを  $X_1$  の帯域ごとに乗じれば  $SS$  を実行したことと同じ結果になることが確認できる。

30

【数 4】

$$D_k = X_{k,2} - X_{k,1} \quad \dots (6)$$

$$G_k = \max \left[ 1 - \frac{\alpha |D_k|}{|X_{k,1}|}, \beta \right] \quad \dots (7)$$

40

【0039】

結合係数算出手段 103 は、ゲイン算出手段 102 から与えられた目的音強調ゲイン  $G$  に基づいて結合係数  $C$  を算出し、得られた結合係数  $C$  をゲイン修正手段 104 に与える。

【0040】

結合係数  $C$  は、目的音強調ゲイン  $G$  と同様に、帯域それぞれに個別の値を有している。以下では、結合係数  $C$  の任意の帯域  $k$  の値を  $C_k$  と表す。 $C_k$  は  $0.0 \sim 1.0$  の範囲の実数値となる。 $C_k$  は、大きいほど目的音らしさが高いことを意味する。

【0041】

結合係数算出手段 103 は、目的音強調ゲイン  $G$  を周波数方向と時間方向に平滑化した

50



値を結合係数  $C$  として算出する。結合係数算出手段 103 が行う平滑化の方法は任意であり、例えば画像処理で用いられるような 2 次元の移動平均法や、2 次元ガウシアンフィルタなどを適用できる。結合係数算出手段 103 が行う平滑化方式として、より好適には、まず周波数方向に移動平均法を用いて平滑化し、次に時間方向に時定数フィルタ（リーク積分ともいう）を用いて平滑化する方式がある。

【0042】

また、結合係数算出手段 103 は、周波数方向に移動平均法によって平滑化する際には、単なる平均を取っても良いし、加重平均を取るようにしても良い。また、結合係数算出手段 103 は、周波数によって平均又は加重平均する範囲を変更しても良く、例えば臨界帯域幅に比例する範囲を平均又は加重平均するようにしてもよい。

10

【0043】

臨界帯域幅とは、人間の聴覚において、ある純音に対する聴覚マスキングが有効な帯域幅のことである。臨界帯域幅は、純音の周波数に対して非線形で、500 Hz 以下の純音に対しては常に 100 Hz であり、500 Hz より上の周波数では純音の周波数に対して  $1/3 \sim 1/4$  オクターブの割合で増加していく特性を持つ。臨界帯域幅（単位は Bark）の近似式がいくつか提案されている。結合係数算出手段 103 は、例えば、以下の(8)式を用いて臨界帯域幅  $b$  [Bark] を求めることができる。なお、(8)式において、 $f$  は周波数（単位は Hz）である。

【数5】

$$b = 13 \tan^{-1}(0.00076 f) + 3.5 \tan^{-1} \left( \left( \frac{f}{7500} \right)^2 \right) \dots (8)$$

20

【0044】

ゲイン修正手段 104 は、目的音強調ゲイン  $G$  と所定の定数値で与えられた無歪ゲイン  $H$ （第2のゲイン）とを、結合係数  $C$  に基づいて帯域ごとに凸結合し、得られた修正ゲイン  $H$ （第3のゲイン）を乗算手段 105 に与える。

【0045】

無歪ゲイン  $F$  及び修正ゲイン  $H$  は、目的音強調ゲイン  $G$  と同様に、帯域それぞれに個別の値を有している。以下では、無歪ゲイン  $F$  の任意の帯域  $k$  の値を  $F_k$  と表す。また、以下では、修正ゲイン  $H$  の任意の帯域  $k$  の値を  $H_k$  と表す。 $F_k$ 、 $H_k$  はいずれも  $0.0 \sim 1.0$  の範囲の実数値となる。

30

【0046】

無歪ゲイン  $F$  は、帯域ごとに所定の定数となる。無歪ゲイン  $F$  の帯域ごとに設定される所定の定数は、ある程度小さく（例えば、 $0.1$  程度）、且つ周波数方向に十分滑らかとする（例えば、隣り合う帯域の比を  $\pm 1$  dB 以内とする）ことが望ましい。なお、無歪ゲイン  $F$  では、全帯域に渡って同じ値を設定するようにしてもよい。

【0047】

ゲイン修正手段 104 は、例えば、(9)式を用いて、帯域ごとの修正ゲイン  $H_k$  を得ることができる。

40

【数6】

$$H_k = C_k \cdot G_k + (1 - C_k) \cdot F_k \dots (9)$$

【0048】

乗算手段 105 は、入力スペクトル  $X_1$  に帯域ごとに修正ゲイン  $H$  を乗じ、得られた目的音強調スペクトル  $Y$  を波形復元手段 106 に与える。

【0049】

50

目的音強調スペクトル $Y$ は、入力スペクトル $X_1$ と同様に、帯域それぞれに個別の値を有している。以下では、目的音強調スペクトル $Y$ の任意の帯域 $k$ の値を $Y_k$ と表す。

【0050】

波形復元手段106は、周波数解析手段101-1~101-Mで用いた周波数解析手法または帯域分割手法に対応する波形復元手法を用いて、乗算手段105から与えられた目的音強調スペクトル $Y$ に基づいて時間波形を再構成し、得られた目的音強調信号 $y$ を出力する。

【0051】

なお、音声処理装置100は、目的音強調スペクトル $Y$ の波形を復元して目的音強調信号 $y$ を出力するものとして説明したが、波形を復元せずに目的音強調スペクトル $Y$ を出力しても良い。また、目的音強調スペクトル $Y$ と目的音強調信号 $y$ の両方を出力するようにしても良い。

10

【0052】

(A-3)第1の実施形態の効果

第1の実施形態によれば、以下のような効果を奏することができる。

【0053】

第1の実施形態の音声処理装置100は、目的音強調ゲイン $G$ (目的音源又は目的音方向を強調できるが歪みを生じる場合のあるゲイン)と、所定の定数値で与えられた無歪ゲイン $F$ (歪みはないが音声成分を強調できないゲイン)とを、目的音らしさを意味し且つ周波数方向と時間方向とに滑らかに変化する結合係数 $C$ に基づいて凸結合することによって修正ゲイン $H$ を得る。そして、第1の実施形態の音声処理装置100は、修正ゲイン $H$ を用いて入力スペクトルについて目的音成分を強調することで、従来の目的音強調方法と同等の目的音強調性能を維持しながら歪みを生じない信号処理を行うことができる。

20

【0054】

(B)第2の実施形態

以下、本発明による音響信号処理装置及びプログラム、並びに、ゲイン処理装置及びプログラムの第2の実施形態を、図面を参照しながら詳述する。以下では、本発明の音響信号処理装置及びゲイン処理装置を音声処理装置に適用した例について説明する。

【0055】

(B-1)第2の実施形態の構成

図2は、第2の実施形態の音声処理装置200の構成を示すブロック図である。図2では、上述の図1と同一部分又は対応部分には、同一符号又は対応符号を付している。

30

【0056】

以下では、第2の実施形態について第1の実施形態との差異を説明する。

【0057】

図2に示すように、第2の実施形態の音声処理装置200には、平滑化手段207が追加されている点で第1の実施形態と異なっている。

【0058】

第2の実施形態では、結合係数算出手段103、ゲイン修正手段104、及び平滑化手段207により、本発明のゲイン処理装置が構成されている。

40

【0059】

(B-2)第2の実施形態の動作

次に、上述した構成を有する第2の実施形態の音声処理装置200の動作について、図2を参照しながら説明する。

【0060】

以下では、第2の実施形態の音声処理装置200の動作について、第1の実施形態との差異部分のみを説明する。

【0061】

第2の実施形態の音声処理装置200では、ゲイン算出手段102によって算出された目的音強調ゲイン $G$ が平滑化手段207にも与えられる点と、平滑化手段207により、

50

目的音強調ゲイン  $G$  に応じた無歪ゲイン  $F$  が生成され、ゲイン修正手段 104 に供給される点で、第 1 の実施形態と異なっている。

【0062】

平滑化手段 207 は、ゲイン算出手段 102 から与えられた目的音強調ゲイン  $G$  に基づいて無歪ゲイン  $F$  を算出し、得られた無歪ゲイン  $F$  をゲイン修正手段 104 に与える。

【0063】

平滑化手段 207 は、目的音強調ゲイン  $G$  を周波数方向と時間方向に平滑化した値を無歪ゲイン  $F$  として算出する。平滑化手段 207 が行う平滑化の方法は任意であり、例えば、画像処理で用いられるような 2 次元の移動平均法や 2 次元ガウシアンフィルタなどを適用できる。平滑化手段 207 が行う平滑化方式として、より好適には、まず周波数方向に移動平均法を用いて平滑化し、次に時間方向に時定数フィルタ（リーク積分ともいう）を用いて平滑化する方式がある。

10

【0064】

また、平滑化手段 207 は、周波数方向に移動平均法によって平滑化する際には、単なる平均を取っても良いし、加重平均を取るようにしても良い。また、平滑化手段 207 は、周波数によって平均又は加重平均する範囲を変更しても良く、例えば臨界帯域幅に比例する範囲を平均又は加重平均するようにしてもよい。また、平滑化手段 207 では、結合係数算出手段 103 における平滑化方法と同一としても良いし、異なる方法やパラメータを用いても良いが、無歪ゲイン  $F$  は歪みをまったく生じない特性を持つ必要がある。したがって、平滑化手段 207 は、周波数方向の平滑化は臨界帯域幅で平滑化した場合と同等以上に滑らかにし、時間方向の平滑化は結合係数算出手段 103 よりも滑らかにすることが好ましい。

20

【0065】

(B-3) 第 2 の実施形態の効果

第 2 の実施形態によれば、以下のような効果を奏することができる。

【0066】

第 2 の実施形態では、平滑化手段 207 が、目的音強調ゲイン  $G$  を周波数方向と時間方向に平滑化した値を無歪ゲイン  $F$  として算出する。そして、ゲイン修正手段 104 は、平滑化手段 207 から供給された無歪ゲイン  $F$  を用いて、目的音強調ゲイン  $G$  を修正する。これにより、第 2 の実施形態では、入力信号の特性によって変化する目的音強調ゲイン  $G$  に対して適応した無歪ゲイン  $F$  を適用することができ、目的音強調ゲイン  $G$  が優勢な帯域及び時刻と無歪ゲイン  $F$  が優勢な帯域及び時刻との差異が適応的に調整される。これにより、第 2 の実施形態の音声処理装置 100 では、音声強調性能を維持しながら、入力信号の特性に適応して歪み等を低減させることができる。

30

【0067】

(C) 第 3 の実施形態

以下、本発明による音響信号処理装置及びプログラム、並びに、ゲイン処理装置及びプログラムの第 3 の実施形態を、図面を参照しながら詳述する。以下では、本発明の音響信号処理装置及びゲイン処理装置を音声処理装置に適用した例について説明する。

【0068】

(C-1) 第 3 の実施形態の構成

次に、上述した構成を有する第 2 の実施形態の音声処理装置 300 の動作について、図 3 を参照しながら説明する。

【0069】

以下では、第 3 の実施形態の音声処理装置 300 の動作について、第 2 の実施形態との差異部分のみを説明する。

【0070】

第 3 の実施形態の音声処理装置 300 では、乗算手段 105 と波形復元手段 106 が入力信号  $x_1 \sim x_M$  ごとに配置されている点で第 1 の実施形態と異なっている。すなわち、第 3 の実施形態の音声処理装置 300 では、乗算手段 105 と波形復元手段 106 が周波

40

50

数解析手段 101 と同じ数 (M 個) 配置されている。第 3 の実施形態に係る音声処理装置 300 は、M 個の乗算手段 105 (105-1 ~ 105-M) と、M 個の波形復元手段 106 (106-1 ~ 106-M) とを有している。

【0071】

(C-2) 第 3 の実施形態の動作

次に、上述した構成を有する第 2 の実施形態の音声処理装置 300 の動作について、図 3 を参照しながら説明する。

【0072】

以下では、第 3 の実施形態の音声処理装置 300 の動作について、第 1 の実施形態との差異部分のみを説明する。

10

【0073】

第 1 の実施形態では、入力信号  $x_1$  のみが乗算手段 105 に供給されていた。これに対して、第 2 の実施形態では、入力信号  $x_1 \sim x_M$  が、それぞれ乗算手段 105-1 ~ 105-M に供給される。

【0074】

また、第 1 の実施形態の音声処理装置 100 では、ゲイン修正手段 104 が、得られた修正ゲイン H を、ただ 1 つの乗算手段 105 に与えていた。これに対して、第 2 の実施形態のゲイン修正手段 104 は、得られた修正ゲイン H を、M 個の乗算手段 105-1 ~ 105-M に与える。

【0075】

そして、第 2 の実施形態では、乗算手段 105-1 ~ 105-M により、それぞれ入力信号  $x_1 \sim x_M$  と修正ゲイン H とが乗算される。以下では、乗算手段 105-1 ~ 105-M の乗算処理により得られる目的音強調スペクトルを  $Y_1 \sim Y_m$  と表す。

20

【0076】

そして、目的音強調スペクトル  $Y_1 \sim Y_m$  は、それぞれ波形復元手段 106-1 ~ 106-M に供給される。そして、波形復元手段 106-1 ~ 106-M は、それぞれ目的音強調スペクトル  $Y_1 \sim Y_m$  に基づいて時間波形を再構成し、得られた目的音強調信号を出力する。以下では、波形復元手段 106-1 ~ 106-M が出力する目的音強調信号を、 $y_1 \sim y_M$  と表す。

【0077】

音声処理装置 300 は、例えば、複数の入力信号 (マイクロホンの入力) を用いる任意の装置の前処理として適用できる。音声処理装置 300 から出力される目的音強調信号  $y_1 \sim y_M$  は、例えば、所定の方向に音源を定位する装置 (音像定位装置) に供給するようにしてもよい。

30

【0078】

(C-3) 第 3 の実施形態の効果

第 3 の実施形態によれば、以下のような効果を奏することができる。

【0079】

第 3 の実施形態の音声処理装置 100 では、すべての入力信号 (マイクロホンの入力) に対して同じ修正ゲイン H を用いて目的音強調を行うので、複数の入力信号 (マイクロホンの入力) を用いる任意の装置の前処理として好適な目的音強調処理 (歪みやミュージカルノイズの少ない目的音強調処理) を提供できる。

40

【0080】

(D) 他の実施形態

本発明は、上記の実施形態に限定されるものではなく、以下に例示するような変形実施形態も挙げることができる。

【0081】

(D-1) 上記の各実施形態では、強調する目的音を音声としているが、音声以外の音 (例えば、機械音や操作音等の音) を目的音 (強調対象となる音響信号) としてもよい。

【0082】

50

(D-2) 第3の実施形態では、音声処理装置300は、すべての目的音強調スペクトルの波形を復元して目的音強調信号を出力するものとして説明したが、波形を復元せずすべての目的音強調スペクトルを出力するようにしても良いし、一部の目的音強調スペクトルを波形に復元して目的音強調信号と目的音強調スペクトルの両方を出力するようにしても良い。また、第3の実施形態では、音声処理装置300は、すべてのマイクロホンの入力に関して目的音強調信号を出力するものとして説明したが、少なくとも2つ以上の任意の入力信号(図示しないマイクロホンからの入力信号)を選択して目的音強調信号を生成して出力するようにしても良い。

【0083】

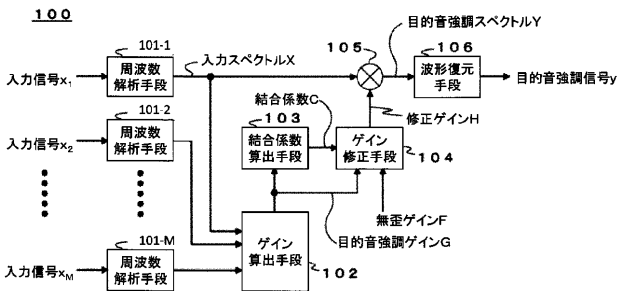
(D-3) 上記の各実施形態では、音声処理装置に時間領域の入力信号を供給しているが、音声処理装置に予め周波数領域に変換した入力スペクトルを供給するようにしても良い。

【符号の説明】

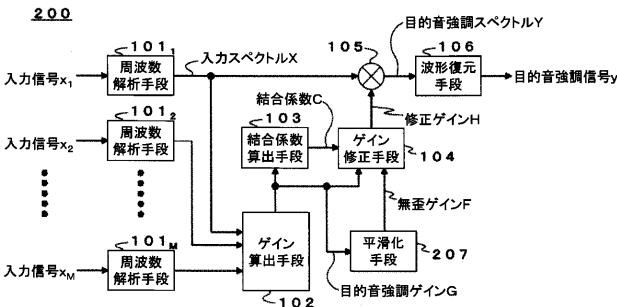
【0084】

100... 音声処理装置、101、101-1~101-M... 周波数解析手段、102... ゲイン算出手段、103... 結合係数算出手段、104... ゲイン修正手段、105... 乗算手段、106... 波形復元手段。

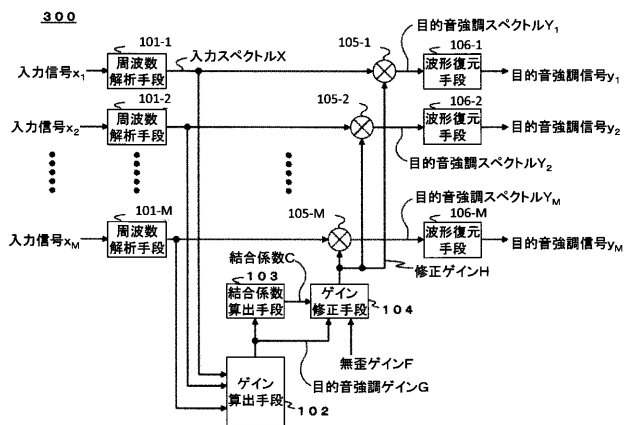
【図1】



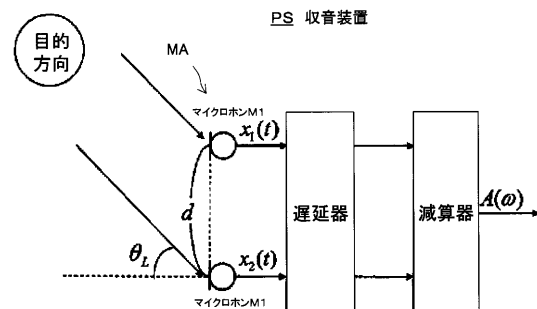
【図2】



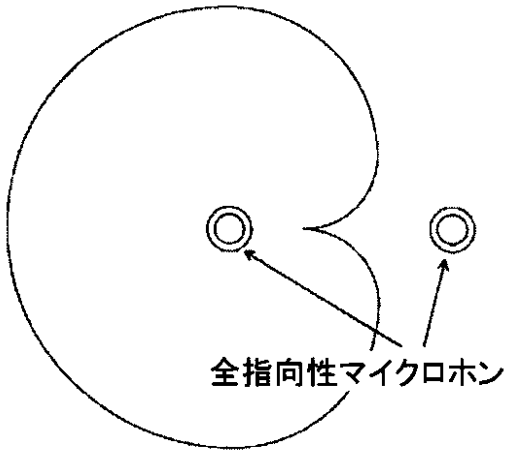
【図3】



【図4】



【 図 5 】



【 図 6 】

