

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6613078号
(P6613078)

(45) 発行日 令和1年11月27日(2019.11.27)

(24) 登録日 令和1年11月8日(2019.11.8)

(51) Int.Cl.	F 1
HO4R 3/00	(2006.01)
HO4R 1/40	(2006.01)
HO4S 7/00	(2006.01)
	HO4R 3/00 320
	HO4R 1/40 320A
	HO4S 7/00 330
	HO4R 3/00 310

請求項の数 17 (全 17 頁)

(21) 出願番号 特願2015-169731 (P2015-169731)
 (22) 出願日 平成27年8月28日 (2015.8.28)
 (65) 公開番号 特開2017-46322 (P2017-46322A)
 (43) 公開日 平成29年3月2日 (2017.3.2)
 審査請求日 平成30年8月27日 (2018.8.27)

(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 100076428
 弁理士 大塚 康徳
 (74) 代理人 100115071
 弁理士 大塚 康弘
 (74) 代理人 100112508
 弁理士 高柳 司郎
 (74) 代理人 100116894
 弁理士 木村 秀二
 (74) 代理人 100130409
 弁理士 下山 治
 (74) 代理人 100134175
 弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】信号処理装置及びその制御方法

(57) 【特許請求の範囲】

【請求項 1】

複数の収音部により収音された音響信号に基づいて特定の方向に対応する音響信号である方向音信号を取得する取得手段であって、第1周波数帯域と前記第1周波数帯域より低い第2周波数帯域とを少なくとも含む複数の周波数帯域それぞれについて1以上の方向に対応する1以上の方向音信号を取得する取得手段と、

前記取得手段により取得された前記第1周波数帯域に対応する方向音信号と、前記取得手段により取得された前記第2周波数帯域に対応する方向音信号とを用いて、再生用の音響信号を生成する生成手段とを有し、

前記生成手段により前記再生用の音響信号の生成に用いられる前記第1周波数帯域に対応する方向音信号の数は、前記生成手段により前記再生用の音響信号の生成に用いられる前記第2周波数帯域に対応する方向音信号の数よりも多いことを特徴とする信号処理装置。

【請求項 2】

前記取得手段は、前記複数の収音部により収音された音響信号に対して特定の方向に対応する指向性形成フィルタを適用することで、当該特定の方向に対応する方向音信号を取得することを特徴とする請求項1に記載の信号処理装置。

【請求項 3】

前記複数の収音部は、それぞれ異なる方向に指向性を有する複数の指向性マイクロホンであり、

前記取得手段は、前記複数の指向性マイクロホンのうち特定の方向に指向性を有する指向性マイクロホンにより収音された音響信号を、当該特定の方向に対応する方向音信号として取得することを特徴とする請求項1に記載の信号処理装置。

【請求項4】

前記取得手段により取得される同一の周波数帯域に対応する複数の方向音信号はそれぞれ異なる方向に対応することを特徴とする請求項1乃至3のいずれか1項に記載の信号処理装置。

【請求項5】

前記生成手段により用いられる同一の周波数帯域に対応する複数の方向音信号に対応するビームパターンを合成した合成ビームパターンの凹み量が閾値以下となるように、前記生成手段により用いられる方向音信号の数を決定する決定手段を有することを特徴とする請求項1乃至4のいずれか1項に記載の信号処理装置。

10

【請求項6】

前記生成手段により用いられる同一の周波数帯域に対応する複数の方向音信号に対応するビームパターンの重なり量が所定値以下となるように、前記生成手段により用いられる方向音信号の数を決定する決定手段を有することを特徴とする請求項1乃至4のいずれか1項に記載の信号処理装置。

【請求項7】

前記決定手段は、前記複数の方向音信号に対応するビームパターンの特定の方向の値について、最大値とそれ以外の値との比が閾値以上となるように、前記生成手段により用いられる方向音信号の数を決定することを特徴とする請求項6に記載の信号処理装置。

20

【請求項8】

前記生成手段により生成された再生用の音響信号を再生する再生手段を更に有することを特徴とする請求項1乃至7のいずれか1項に記載の信号処理装置。

【請求項9】

前記生成手段は、前記取得手段により取得された複数の方向音信号に、各方向音信号に対応する方向の頭部伝達関数を適用することで、前記再生用の音響信号を生成し、

前記再生手段は、前記生成手段により生成された再生用の音響信号をユーザの両耳近傍で再生することを特徴とする請求項8に記載の信号処理装置。

【請求項10】

30

前記再生手段は、前記生成手段により生成された再生用の音響信号をユーザの周囲に配置される複数のスピーカにより再生することを特徴とする請求項8に記載の信号処理装置。

【請求項11】

前記生成手段により前記再生用の音響信号の生成に用いられる前記第1周波数帯域に対応する方向音信号の数及び前記第2周波数帯域に対応する方向音信号の数は、頭部伝達関数の周波数別の方向感度に応じて決まることを特徴とする請求項8乃至10のいずれか1項に記載の信号処理装置。

【請求項12】

前記方向感度は、前記頭部伝達関数の両耳間レベル差の方向に応じた変化量であることを特徴とする請求項11に記載の信号処理装置。

40

【請求項13】

前記生成手段により前記再生用の音響信号の生成に用いられる前記第1周波数帯域に対応する方向音信号の数及び前記第2周波数帯域に対応する方向音信号の数は、選択可能な指向方向数の中から、所定の指向方向数との差が小さくなるように決定されることを特徴とする請求項10に記載の信号処理装置。

【請求項14】

前記選択可能な指向方向数は、前記複数のスピーカが再生可能な周波数帯域に応じて決定されることを特徴とする請求項13に記載の信号処理装置。

【請求項15】

50

複数の収音部により収音された音響信号に基づいて特定の方向に対応する音響信号である方向音信号を取得する取得工程であって、第1周波数帯域と前記第1周波数帯域より低い第2周波数帯域とを少なくとも含む複数の周波数帯域それぞれについて1以上の方に対応する1以上の方の方向音信号を取得する取得工程と、

前記取得工程において取得された前記第1周波数帯域に対応する方向音信号と、前記取得工程において取得された前記第2周波数帯域に対応する方向音信号とを用いて、再生用の音響信号を生成する生成工程とを有し、

前記生成工程において前記再生用の音響信号の生成に用いられる前記第1周波数帯域に對応する方向音信号の数は、前記生成工程において前記再生用の音響信号の生成に用いられる前記第2周波数帯域に對応する方向音信号の数よりも多いことを特徴とする信号処理方法。

【請求項16】

前記取得工程において取得される同一の周波数帯域に對応する複数の方向音信号はそれぞれ異なる方向に対応することを特徴とする請求項15に記載の信号処理方法。

【請求項17】

コンピュータを請求項1乃至14のいずれか1項に記載の信号処理装置が有する各手段として機能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、音響信号処理技術に関する。

【背景技術】

【0002】

複数のマイクロホン素子（マイクロホンアレイ）で録音した複数チャネルの音響信号から、方向ごとの音（以下「方向音」という。）を取得する技術が知られている。この技術を用いて、全方位の方向音が各方向から再生されているようにユーザに提示できれば、あたかも録音現場にいるかのような高臨場感を実現することができる。

【0003】

特許文献1には、指向性約45°の8本の指向性マイクロホンを放射状に並べた指向性マイクロホンアレイで指向方向ごとの方向音を取得し、各指向方向に45°間隔で配置した8台の周囲スピーカで再生を行う全方位録音・再生システムが開示されている。

【0004】

方向音を取得する方法としては、指向性マイクロホンアレイを用いる方法の他に、フィルタリングに基づく方法がある。すなわち、（無指向性）マイクロホンアレイで録音した複数チャネルの音響信号に、所望の指向方向に對応する指向性形成のフィルタ係数を適用して加算することで、任意の指向方向の方向音を生成することができる。特許文献2では、8本のマイクロホンで構成されるマイクロホンアレイで録音した8ch音響信号をフィルタリング（遅延制御）することで、ユーザが要求する指向性マイクロホンと同じになるよう指向性を形成し、ユーザが要求する数の方向音を生成している。

【0005】

また、全方位の方向音が各方向から再生されているようにユーザに提示する方法としては、ユーザの周囲にスピーカを配置する方法の他に、ヘッドホンで立体音響再生を行う方法がある。すなわち、各方向音に各指向方向に對応する方向の左右耳の頭部伝達関数を適用し、左右それぞれ加算してヘッドホンから再生することで、仮想スピーカをユーザの周囲に配置するような効果が得られる。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特許第2515101号公報

【特許文献2】特開平9-55925号公報

10

20

30

40

50

【発明の概要】

【発明が解決しようとする課題】

【0007】

一般に方向音の取得において、指向性マイクロホンアレイを用いる場合もフィルタリングで指向性を形成する場合も、形成できる指向性のビームパターンは、低周波数域では鈍くなり、逆に高周波数域では鋭くなる傾向がある。このとき、全方位録音・再生を行うために、所定の指向方向数で均等配置した各指向方向の方向音を取得し、例えばヘッドホンで立体音響再生を行うと、以下のような課題が生じる。

【0008】

すなわち、低周波数域では各指向性のビームパターンの重なりが増えるため、(点)音源の方向感が不明瞭になり、音量も過多になる傾向がある。また高周波数域では、各指向性のビームパターンの重なりが減り、各ビームパターンを合成した合成ビームパターンにおいて、各指向方向の間で凹みが生じる。このため、音源間(例えば全方位に配置した楽器間)の音量バランスが崩れ、全方位の雰囲気音(拡散音源)も方向ごとに音量感が変わってしまう。

【0009】

特許文献1,2には、このような周波数ごとの指向性の違いに起因する課題を解決できる手法は開示されていない。

【0010】

本発明は、音源の方向感の明瞭化、及び、各方向の音量バランスの均一化に有利な技術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】

【0011】

本発明の一側面によれば、複数の収音部により収音された音響信号に基づいて特定の方向に対応する音響信号である方向音信号を取得する取得手段であって、第1周波数帯域と前記第1周波数帯域より低い第2周波数帯域とを少なくとも含む複数の周波数帯域それぞれについて1以上の方向に対応する1以上の方向音信号を取得する取得手段と、前記取得手段により取得された前記第1周波数帯域に対応する方向音信号と、前記取得手段により取得された前記第2周波数帯域に対応する方向音信号とを用いて、再生用の音響信号を生成する生成手段とを有し、前記生成手段により前記再生用の音響信号の生成に用いられる前記第1周波数帯域に対応する方向音信号の数は、前記生成手段により前記再生用の音響信号の生成に用いられる前記第2周波数帯域に対応する方向音信号の数よりも多いことを特徴とする信号処理装置が提供される。

【発明の効果】

【0012】

本発明によれば、音源の方向感の明瞭化、及び、各方向の音量バランスの均一化に有利な技術を提供することができる。

【図面の簡単な説明】

【0013】

【図1】第1実施形態に係る信号処理装置のブロック図。

【図2】第1実施形態に係る信号処理を示すフローチャート。

【図3】指向方向数が5の場合のビームパターンの例を示す図。

【図4】指向方向数が9の場合のビームパターンの例を示す図。

【図5】指向方向数が17の場合のビームパターンの例を示す図。

【図6】周波数ごとの指向方向数の説明図。

【図7】頭部伝達関数の周波数別方向感度の説明図。

【図8】第2実施形態に係る信号処理装置のブロック図。

【図9】第2実施形態に係る信号処理を示すフローチャート。

【発明を実施するための形態】

【0014】

10

20

30

40

50

以下、図面を参照して本発明の実施形態について詳細に説明する。なお、本発明は以下の実施形態に限定されるものではなく、また、実施形態の中で説明されている特徴の組み合わせの全てが本発明の課題解決のために必須のものであるとは限らない。なお、同一の構成については、同じ符号を付して説明する。

【0015】

<第1実施形態>

図1は、第1実施形態に係る信号処理装置100の構成を示すブロック図である。信号処理装置100は、各構成要素を統括的に制御するシステム制御部101、各種データを記憶しておく記憶部102、信号の解析処理を行う信号解析処理部103を備える。記憶部102は、複数のマイクロホン素子(収音部)を有するマイクロホンアレイ106で収音された音響信号を保持している。音響信号入力部107は、マイクロホンアレイ106からの音響信号を入力する。

【0016】

信号処理装置100は、ユーザの周囲に方向音の音像である方向音像を生成する再生系を有する。本実施形態において、再生系は、音響信号出力部104、ヘッドホン105を含む。この再生系は、後述するように、各方向音に各指向方向に対応する方向の頭部伝達関数(HRTF(Head-Related Transfer Function))を適用して、ユーザの両耳近傍で再生することができる。信号解析処理部103は、後述する信号解析処理によって、ヘッドホン105から再生するヘッドホン再生信号を生成する。音響信号出力部104は、ヘッドホン再生信号にD/A変換および増幅を施した信号を、ヘッドホン105に出力する。

【0017】

以下、本実施形態における信号処理について、図2のフローチャートに沿って説明する。なお、図2のフローチャートに対応するプログラムは例えば記憶部102に保持され、特に別記しない限り信号解析処理部103によって実行されるものとする。

【0018】

S201では、記憶部102が保持している、M個のマイクロホン素子(Mチャネルマイクロホンアレイ)で録音したMチャネル音響信号を取得し、チャネルごとにフーリエ変換することで周波数領域のデータ(フーリエ係数)であるz(f)を得る。ここで、各周波数のz(f)はM個の要素を持つベクトルである。

【0019】

S202～S216は周波数ごとの処理であり、周波数ループの中で行う。

【0020】

S202では、現在の周波数ループの周波数における指向方向数D(f)をD(f)=1のように初期化する。S203では、指向方向数D(f)を用いて各指向性の指向方向 $\theta_d(f)$ [d=1～D(f)]を算出する。ここでは複数の指向性で水平全方位をカバーするため、音響信号を録音したマイクロホンアレイの座標系における正面0°を基準方向として、水平の指向方向(方位角)を、 $\theta_d(f)=(d-1) \times 360^\circ / D(f)$ のように算出する。なお、180°を超える指向方向については $\theta_d(f) = \theta_d(f) - 360^\circ$ のように表現するものとする。

【0021】

S204～S205は、S203で指向方向を算出した指向性ごとの処理であり、指向性ループの中で行う。

【0022】

S204では、現在の指向性ループで対象としている指向性を形成するための指向性形成フィルタのフィルタ係数を取得する。ここでは、記憶部102があらかじめ保持している指向性形成フィルタのフィルタ係数から、指向方向 $\theta_d(f)$ に対応する $w_d(f)$ を取得する。ここで、フィルタ係数(ベクトル) $w_d(f)$ は周波数領域のデータ(フーリエ係数)であり、M個の要素で構成される。なお、マイクロホンアレイの構成が異なればフィルタ係数も異なるため、録音に用いたマイクロホンアレイの種別IDを音響信号の付加情報として録音時に記録しておき、そのマイクロホンアレイに対応するフィルタ係数を本ステップで

10

20

30

40

50

用いるようにしてもよい。

【0023】

指向性形成フィルタのフィルタ係数の算出には、各方向（方位角 θ ）の音源と各マイクロホン素子間の伝達関数であるアレイ・マニフォールド・ベクトル $a(f, \theta)$ が一般に用いられる。ここで、 $a(f, \theta)$ は周波数領域のデータ（フーリエ係数）であり、M個の要素で構成される。指向方向 $\theta_d(f)$ に指向性のメインローブを向ける方法として、例えば遅延和法なら $\theta_d(f)$ 方向のアレイ・マニフォールド・ベクトル $a_d(f)$ を用いて、 $w_d(f) = a_d(f) / (a_d^H(f)a_d(f))$ のようにフィルタ係数が得られる。

【0024】

S205では、S204で取得した指向性形成フィルタのフィルタ係数 $w_d(f)$ と、アレイ・マニフォールド・ベクトル $a(f, \theta)$ を用いて指向性のビームパターンを算出する。ビームパターンの方位角 θ 方向の値 $b_d(f, \theta)$ は、式(1)で得られる。

【0025】

【数1】

$$b_d(f, \theta) = w_d^H(f) a(f, \theta) \quad (1)$$

【0026】

$a(f, \theta)$ の θ を、例えば -180° から 180° まで 1° 刻みで変えながら $b_d(f, \theta)$ を計算することで、水平全方位のビームパターンが得られる。なお、音響信号の録音に用いたマイクロホンアレイの構造によっては、自由空間や剛球等の理論式により、任意の解像度でアレイ・マニフォールド・ベクトル $a(f, \theta)$ を算出できる。なお、円状等間隔マイクロホンアレイなどマイクロホン素子が等方的に配置されている場合は、指向方向が正面 0° の場合のビームパターン $b_1(f, \theta)$ を回転させることで、他の指向性のビームパターン $b_d(f, \theta)$ [$d=2 \sim D(f)$]を得ることもできる。

【0027】

S206では、S205で算出した各指向性のビームパターン $b_d(f, \theta)$ [$d=1 \sim D(f)$]を合成することで、式(2)のように合成ビームパターン $b_{sum}(f, \theta)$ を算出する。

【0028】

【数2】

$$b_{sum}(f, \theta) = \sqrt{\sum_{d=1}^{D(f)} b_d^2(f, \theta)} \quad (2)$$

【0029】

ここで、現在の周波数で形成される指向性に対して指向方向数 $D(f)$ が不足していると、図3[D(f)=5の例]のように、指向方向301～305にそれぞれメインローブを向けた各指向性のビームパターン311～315の重なりが減る。この結果、各ビームパターンを合成した合成ビームパターン316において、各指向方向301～305の間で凹みが生じるため、音源間の音量バランスが崩れ、全方位の雰囲気音も方向ごとに音量感が変わってしまう。

【0030】

そこでS207では、S206で算出した合成ビームパターン $b_{sum}(f, \theta)$ の凹み量の目安として標準偏差 $b_{sum}(f)$ を算出し、その値が閾値以下であるか否かを判定する。この閾値を τ_1 とする。算出された標準偏差 $b_{sum}(f)$ が閾値 τ_1 以下でなければ、指向方向数 $D(f)$ が不足しているとして処理はS208へ進み、 τ_1 以下であれば、処理はS209へ進む。なお、標準偏差 $b_{sum}(f)$ は例えばdB表現の $b_{sum}(f, \theta)$ から算出するものとする。なお、 $b_{sum}(f, \theta)$ の最大値と最小値の差(図3の例では双方向矢印317)を凹み量の目安とし、その値を閾値 τ_2 と比較するようにしてもよい。このとき、 $b_{sum}(f, \theta)$ は各指向

10

20

30

40

50

方向で最大値をとり、隣接する指向方向の真ん中で最小値をとる。

【0031】

S208へ進んだ場合は、指向方向数D(f)をD(f) D(f)+1のようにインクリメントし、S203へ戻る。

【0032】

S209へ進んだ場合は、指向方向数が適切な範囲に入ったとして、このときの指向方向数D(f)を現周波数の指向方向数の下限値である下限指向方向数D_{min}(f)と決定する。

【0033】

現在の周波数で形成される指向性に対して指向方向数D(f)が適切となると、図4 [D(f)=9の例] のように、合成ビームパターン334の凹みが消え略円形となる。

【0034】

次に、現在の周波数で形成される指向性に対して指向方向数D(f)が過剰となると、図5 [D(f)=17の例] のように、各指向性のビームパターンの重なりが増えるため、音源の指向感が不明瞭になり、音量も過多になる傾向がある。ただし、指向方向数が過剰な場合については、指向方向数が不足の場合のような合成ビームパターンの乱れは生じず、図5の合成ビームパターン366のように略円形となるため、別の評価法を考える必要がある。ただし、ビームパターンの形状（面積）は描画における表示範囲の設定（図3の場合 - 30 dBから+10 dB）に依存するため、評価指標として各ビームパターンのオーバーラップ部分の面積比などは適さない。

【0035】

そこで、所定の方向における各ビームパターンの値の比を評価指標に用いることを考える。式(3)のように、各方向でビームパターンの値が最大となる指向性のインデックスをd_{max}(f, θ)とする。

【0036】

【数3】

$$d_{max}(f, \theta) = \arg \max_d b_d(f, \theta) \quad (3)$$

【0037】

そして、各方向で最大のビームパターンの値をb_{dmax}(f, θ)とすると、各方向におけるビームパターンの最大値とそれ以外の値との比r(f, θ)は、例えば式(4)のように表される。

【0038】

【数4】

$$r(f, \theta) = \sqrt{\frac{b_{dmax}^2(f, \theta)}{\sum_{d=1}^{D(f)} b_d^2(f, \theta) - b_{dmax}^2(f, \theta)}} \quad (4)$$

【0039】

ここで、図4のように指向方向数が適切な場合は、例えば指向方向321に音源があるとすると、この指向方向θ₁(f)=0°におけるr(f, θ₁)は、例えば8 dBといった正の値をとる。すなわち、指向方向321にメインローブを向けたビームパターン331が捉える音のエネルギー341が、指向方向322～323にメインローブを向けたビームパターン332～333が捉える音のエネルギー342～343の和よりも大きい。つまり、ある方向に音源がある場合に、その方向にメインローブを向けた指向性が捉える音のエネルギーが、他の方向にメインローブを向けた指向性が捉える音のエネルギーの和より大きいため、適切な状態であるといえる。

【0040】

10

20

30

40

50

一方、図5のように指向方向数が過剰な場合は、例えば指向方向351に音源があるとすると、この指向方向 $\alpha_1(f)=0^\circ$ における $r(f, \alpha_1)$ は、例えば0dB未満の小さな値をとる。すなわち、指向方向351にメインロープを向けたビームパターン361が捉える音のエネルギー371より、指向方向352～355にメインロープを向けたビームパターン362～365が捉える音のエネルギー372～375の和の方が大きい。つまり、ある方向に音源がある場合に、その方向にメインロープを向けた指向性が捉える音のエネルギーより、他の方向にメインロープを向けた指向性が捉えるエネルギーの和の方が大きいため、不適切な状態であるといえる。

【0041】

以上を踏まえ、S210では、指向方向 $\alpha_d(f)$ におけるビームパターンの最大値とそれ以外の値との比 $r(f, \alpha_d(f))$ を算出し、その値が閾値以上であるか否かを判定する。ここではその閾値を α_3 とする。算出された比の値が閾値 α_3 （例えば0dB）以上であれば、指向方向数D(f)はまだ適切な範囲内であるとして処理はS208へ進み、閾値 α_3 未満であれば処理はS211へ進む。なお、指向方向 $\alpha_d(f)$ 以外の方向における $r(f, \alpha)$ を閾値 α_4 と比較するようにしてもよいが、 $r(f, \alpha)$ は指向方向 $\alpha_d(f)$ で最大となるため、本実施形態では、例えば、 $\alpha_4 < \alpha_3$ とする。

【0042】

なお、各指向性のビームパターンの重なりが増えると、図5のように合成ビームパターン366の値が大きくなるため、音量も過多になる傾向がある。そこで、合成ビームパターンの最大値 $b_{sum}(f, \alpha_d(f))$ と、各ビームパターンの最大値 $b_d(f, \alpha_d(f))$ [正規化されれば0dB]との差（図5の例では双方向矢印367）を、閾値 α_5 と比較するようにしてもよい。すなわち、 $b_{sum}(f, \alpha_d(f)) - b_d(f, \alpha_d(f))$ が α_5 以下であれば、指向方向数D(f)はまだ適切な範囲内であるとしてS208へ、 α_5 より大きければS211へ進むようにしてもよい。

【0043】

S208へ進んだ場合は、指向方向数D(f)をD(f)→D(f)+1のようにインクリメントし、S203へ戻る。ただし、指向方向数の下限値D_{min}(f)は既に決定しているため、以降はS207とS209をスキップするものとする。

【0044】

S211へ進んだ場合は、指向方向数が適切な範囲から外れたとして、このときの指向方向数D(f)から1を引いたD(f)-1を現在の周波数の指向方向数の上限値である上限指向方向数D_{max}(f)と決定する。

【0045】

一般に、低周波数域では形成できる指向性のビームパターンが鈍くなり、逆に高周波数域では鋭くなる傾向がある。このため、上記S207、210のように周波数ごとにビームパターンを評価すると、図6(a)に模式的に示すように、低周波数域より高周波数域で下限指向方向数D_{min}(f)および上限指向方向数D_{max}(f)が大きくなる。そこで、周波数ごとの指向方向数を例えば次式を用いてD(f)=D_{mean}(f)のように決定する。

【0046】

【数5】

$$D_{mean}(f) = \text{round}\left(\frac{D_{min}(f) + D_{max}(f)}{2}\right) \quad (5)$$

【0047】

このようにすれば、低周波数域より高周波数域の指向方向数が多くなり、全ての周波数で指向方向数が適切な範囲内となるため、音源の方向感が明瞭となり、各方向の音量バランスも均一となる。

【0048】

ここでさらに、人の音源方向に対する周波数ごとの感度特性を考慮して、各周波数の指

10

20

30

40

50

向方向数 $D(f)$ を $D_{min}(f) \sim D_{max}(f)$ の範囲内で適切に決定することを考える。

【0049】

図7(a)は、HRTFから算出した周波数ごとの両耳間レベル差(ILD)について、音源方向が0°から180°の場合まで1°ごとに、計181本のグラフで描画したものである。なお、音源方向が0°から-180°の場合のグラフは、概ね図7(a)を符号反転(上下反転)したものとなる。また、図7(a)のグラフの周波数ごとの標準偏差 $ILD(f)$ を図7(b)に示している。

【0050】

ここで、人の音源方向に対する感度(方向感度)は、HRTFの両耳間レベル差の方向に対する変化量に対応する。例えば、 $ILD(f)$ が大きい周波数、すなわち方向によるILDの変化が大きい周波数は、人の音源方向に対する感度(方向感度)が高い周波数である。そこで、点線501のように $ILD(f)$ が大きい周波数では、人が方向ごとの差が分かりやすいと考えられるため、指向方向数を $D_{max}(f)$ 寄りの値とする。一方、図7(b)の点線502のように $ILD(f)$ が小さい周波数では、人が方向ごとの差が分かりにくいと考えられるため、指向方向数を $D_{min}(f)$ 寄りの値とする。

【0051】

具体的には、例えば図7(b)のように $ILD(f)$ が概ね0~15dBの値をとる場合は、15で割ることで正規化し、0~1の値をとるHRTFの周波数別の方角感度 $s(f)$ と定義する。そして、例えば式(6)を用いることで、人の周波数別の方角感度を考慮した指向方向数を、 $D(f)=D_{sens}(f)$ のように $D_{min}(f) \sim D_{max}(f)$ の範囲内で決定することができる。

【0052】

【数6】

$$D_{sens}(f) = \text{round}\left(D_{min}(f) + s(f)(D_{max}(f) - D_{min}(f))\right) \quad (6)$$

【0053】

なお、音源方向0°~180°のHRTFから $s(f)$ を算出しているため、これは全方位に対する平均的な方向感度であると解釈できる。これは特に、後述するヘッドホン再生信号の生成において、ユーザの頭部運動に応じたHRTFの切り替え(ヘッドトラッキング処理)を行う場合は、全方位のHRTFが使用され得るため妥当であるといえる。

【0054】

なお、人が音を知覚しにくい例えば15kHz以上においては、HRTFから算出した $s(f)$ に適当な減衰カーブを適用することで、 $D_{sens}(f)$ を小さくするようにしてもよい。図6(a)に $D_{sens}(f)$ の例を曲線で模式的に示す。なお、指向方向数に対応する図6(a)の4本のグラフは、整数値をとるため実際には階段状のグラフとなる。

【0055】

以上を踏まえ、S212では、 $D_{min}(f) \sim D_{max}(f)$ の範囲内で各周波数の指向方向数を $D(f)=D_{mean}(f)$ [式(5)] や $D(f)=D_{sens}(f)$ [式(6)] のように決定する。なお、式(6)の $s(f)$ は、あらかじめHRTFから算出して記憶部102が保持しているものを取得して用いればよい。

【0056】

S213では、S212で決定した指向方向数 $D(f)$ を用いて、S203と同様に各指向性の指向方向 $d(f)=(d-1) \times 360^\circ / D(f)$ [$d=1 \sim D(f)$] を算出する。なお、180°を超える指向方向については $d(f) = d(f) - 360^\circ$ のように表現するものとする。

【0057】

S214~S216は、S213で指向方向を算出した指向性ごとの処理であり、指向性ループの中で行う。

【0058】

10

20

30

40

50

S 2 1 4 では、S 2 0 4 と同様に、現在の指向性ループで対象としている指向性を形成するためのフィルタ係数を取得する。すなわち、記憶部 1 0 2 があらかじめ保持している指向性形成フィルタのフィルタ係数から、指向方向 $\theta_d(f)$ に対応する $w_d(f)$ を取得する。

【 0 0 5 9 】

S 2 1 5 では、S 2 0 1 で取得したMチャネル音響信号のフーリエ係数 $z(f)$ に、S 2 1 4 で取得した指向性形成フィルタのフィルタ係数 $w_d(f)$ を適用する。これにより、現在の指向性ループに対応する指向方向 $\theta_d(f)$ の方向音 $Y_d(f)$ を式 (7) のように生成する。ここで、 $Y_d(f)$ は周波数領域のデータ (フーリエ係数) である。

【 0 0 6 0 】

【数 7】

10

$$Y_d(f) = w_d^H(f)z(f) \quad (7)$$

【 0 0 6 1 】

S 2 1 6 では、S 2 1 5 で取得した指向方向 $\theta_d(f)$ の方向音のフーリエ係数 $Y_d(f)$ に、指向方向 $\theta_d(f)$ と同じ方向の左右耳の H R T F [$H_L(f, \theta_d(f)), H_R(f, \theta_d(f))$] を適用する。そして、式 (8) のように左右それぞれのヘッドホン再生信号 $X_L(f), X_R(f)$ に加算して行く。ここで、 $X_L(f), X_R(f)$ は周波数領域のデータ (フーリエ係数) である。なお、H R T F はあらかじめ記憶部 1 0 2 が保持しているものを取得して用いればよい。

【 0 0 6 2 】

【数 8】

20

$$\begin{cases} X_L(f) \leftarrow X_L(f) + H_L(f, \theta_d(f))Y_d(f) \\ X_R(f) \leftarrow X_R(f) + H_R(f, \theta_d(f))Y_d(f) \end{cases} \quad (8)$$

【 0 0 6 3 】

指向性ループの中で本ステップの処理を行うことで、各指向方向の方向音を再生する仮想スピーカをユーザの周囲に順次配置していくことになる。さらに周波数ループの中で本ステップの処理を行うことで、S 2 1 2 で決定した指向方向数 $D(f)$ に対応して、仮想スピーカの数も周波数ごとに制御することになる。すなわち、低周波数域より高周波数域の仮想スピーカ数が多くなり、全ての周波数で仮想スピーカ数が適切な範囲内となるため、音源の方向感が明瞭となり、各方向の音量バランスも均一となる。

30

【 0 0 6 4 】

なお、周波数ごとに指向方向数 $D(f)$ を適切に制御することで、各周波数の合成ビームパターンのレベルは概ね揃うようになるが、より厳密に、全周波数で合成ビームパターンのレベルを一定値にするような周波数ごとのゲイン調整を行ってもよい。

【 0 0 6 5 】

なお、ユーザの頭部運動を検出可能なセンサを、例えばヘッドホン 1 0 5 が備えていてもよい。そして、例えば音響信号の所定の時間フレーム長 (音響フレーム) ごとに、頭部運動に応じて使用する H R T F を切り替えて行くヘッドトラッキング処理を行ってもよい。

40

【 0 0 6 6 】

S 2 1 7 では、S 2 1 6 で生成したヘッドホン再生信号のフーリエ係数 $X_L(f), X_R(f)$ をそれぞれ逆フーリエ変換することで、時間波形であるヘッドホン再生信号 $x_L(t), x_R(t)$ を取得する。

【 0 0 6 7 】

S 2 1 8 では、音響信号出力部 1 0 4 が、S 2 1 7 で取得したヘッドホン再生信号 $x_L(t), x_R(t)$ に D A 変換および增幅を施し、ヘッドホン 1 0 5 から再生する。

【 0 0 6 8 】

50

なお、例えばS202～S213の周波数ごとの各指向方向の決定までは、あらかじめ処理して結果を記憶部102に保持しておいてもよい。そして、S201の音響信号の取得に合わせて、例えば音響フレームごとにS214～S218の音響レンダリング・再生処理のみリアルタイムに行うようにしてもよい。

【0069】

なお、例えばシステム制御部101と相互に結ばれた不図示のG.U.I部を介して、指向方向数D(f)を例えば低域・中域・高域ごとにユーザが制御できるようにしてもよい。

【0070】

なお、第1実施形態では、S215で指向方向_d(f)の方向音のみを生成し、S216で指向方向_d(f)と同じ方向に、生成した方向音と同じ数の仮想スピーカを配置していた。しかし、S215で指向方向_d(f)の方向音だけでなく、例えば1°間隔で水平全方位にメインロープを向けた360°の方向音を生成しておいてもよい。そして、S216でそのうち指向方向_d(f)の方向音のみを選択的に用いて、指向方向_d(f)と同じ方向のみに仮想スピーカを配置するようにしてもよい。

10

【0071】

<第2実施形態>

上記した第1実施形態では、(無指向性)マイクロホンアレイでの指向性形成フィルタリングによる方向音生成と、ヘッドホンでの立体音響再生の組み合わせで、周波数ごとに指向方向数および仮想スピーカ数を制御した。第2実施形態では、指向性マイクロホンアレイによる方向音取得と、周囲スピーカ再生の組み合わせにおいて、周波数ごとに指向方向数および使用スピーカ数を制御する。

20

【0072】

図8は、本実施形態に係る信号処理装置600の構成を示すブロック図である。信号処理装置600は、各構成要素を統括的に制御するシステム制御部101、各種データを記憶しておく記憶部102、信号の解析処理を行う信号解析処理部103を備える。また、信号処理装置600は、ユーザの周囲に方向音の音像である方向音像を生成する生成手段としての再生系を有する。本実施形態において、再生系は例えば、音響信号出力部604、ユーザの周囲(水平周囲)に配置された複数チャネル(例えば12チャネル)を構成する複数のスピーカ611～622を含む。記憶部102は、スピーカ611～622の配置数・方向に合わせて12本の指向性マイクロホンを放射状に並べた、12チャネルの指向性マイクロホンアレイ605で音響信号入力部107を介して録音された12チャネル音響信号を保持している。なお、本発明は、特定のスピーカの数に限定されるものではない。なお、録音に用いた指向性マイクロホンの配置数・方向に合わせて、周囲スピーカを配置するようにしてもよい。

30

【0073】

信号解析処理部103は、後述する信号解析処理によって、スピーカ611～622から再生するスピーカ再生信号を生成する。音響信号出力部104は、生成されたスピーカ再生信号にD/A変換および増幅を施し、スピーカ611～622から再生する。

【0074】

以下、本実施形態の信号解析処理について、図9のフローチャートに沿って説明する。なお、図9のフローチャートに対応するプログラムは例えば記憶部102に保持され、特に別記しない限り信号解析処理部103によって実行されるものとする。

40

【0075】

S701では、記憶部102があらかじめ保持しているスピーカ611～622の配置と再生可能帯域を取得し、そこから各周波数で全方位再生に使用可能なスピーカ数の組み合わせを決定して、後のステップで選択可能な指向方向数D_{sp}(f)とする。なお、周囲スピーカの配置や再生可能帯域は、ユーザの位置であるリスニングポイントに備えたマイクロホンで音響測定を行い算出してもよい。

【0076】

選択可能な指向方向数D_{sp}(f)は、複数のスピーカのそれぞれの再生可能帯域に応じて決

50

定されうる。図8において、大型スピーカ611, 614, 617, 620は低域から高域まで、中型スピーカ613, 615, 619, 621は中域から高域まで、小型スピーカ612, 616, 618, 622は高域のみが再生可能とする。これより、均等配置が可能で全方位再生に使用可能な周波数ごとのスピーカ数の組み合わせ、すなわち後のステップで選択可能な指向方向数 $D_{sp}(f)$ は、以下のようになる。

$$D_{sp}(f)=\{1, 2, 4\} [f < f_M],$$

$$D_{sp}(f)=\{1, 2, 3, 4, 6\} [f_M \leq f < f_H],$$

$$D_{sp}(f)=\{1, 2, 3, 4, 6, 12\} [f_H \leq f]$$

ここで、 f_M は低域と中域の境界周波数、 f_H は中域と高域の境界周波数とする。

【0077】

10

S702の処理は、第1実施形態のS201と同様であるため説明を省略する。

【0078】

S703～S715は周波数ごとの処理であり、周波数ループの中で行う。

【0079】

S703～S704の処理は、第1実施形態のS202～S203と同じであるため説明を省略する。

【0080】

S705は、S704で指向方向を算出した指向性ごとの処理であり、指向性ループの中で行う。

【0081】

20

S705では、現在の指向性ループで対象としている指向性のビームパターンを取得する。すなわち、記憶部102があらかじめ保持している、指向性マイクロホンを指向方向 $d(f)$ に向けた場合のビームパターン $b_d(f, \theta)$ を取得する。なお、指向性マイクロホンのビームパターンは測定やシミュレーション等により得られる。なお、指向性マイクロホンの種別によってビームパターンは異なる。そのため、録音に用いた指向性マイクロホンの種別IDを音響信号の附加情報として録音時に記録しておき、その指向性マイクロホンに対応するビームパターンを本ステップで取得するようにしてもよい。なお、指向性マイクロホンを正面0°に向けた場合のビームパターン $b_0(f, \theta)$ を回転させることで、指向性マイクロホンを他の指向方向 $d(f)$ に向けた場合のビームパターン $b_d(f, \theta) [d=2 \sim 6]$ を得ることができる。

30

【0082】

S706～S711の処理は、第1実施形態のS206～S211と同じであるため説明を省略する。

【0083】

S712では、第1実施形態のS212と同様に、各周波数の指向方向数を $D_{mean}(f)$ [式(5)] や $D_{sens}(f)$ [式(6)] のように決定する。ここで決定した指向方向数を「所定の指向方向数」という。

【0084】

S713では、S712で決定した所定の指向方向数との差が小さくなるように(例えば最小となるように)、S701で決定した選択可能な指向方向数 $D_{sp}(f)$ の中から、各周波数の指向方向数 $D(f)$ を決定する。例えば所定の指向方向数が $D_{mean}(f)$ の場合、図6(b)の太横線で示すように、 $D(f)=4 [f < f_M]$ 、 $D(f)=6 [f_M \leq f < f_D]$ 、 $D(f)=12 [f_D \leq f]$ となる。ここで、 f_D は $D_{mean}=(6+12)/2=9$ となるような周波数である。また、所定の指向方向数が $D_{sens}(f)$ の場合、指向方向数が同じ値となる周波数は必ずしも連続せず、飛び飛びになり得る。

40

【0085】

S714の処理は、第1実施形態のS213と同じであるため説明を省略する。

【0086】

S715では、S702で取得した音響信号から指向方向 $d(f)$ の方向音を取得して、対応するスピーカ再生信号に割り当てる。本実施形態では音響信号が指向性マイクロホン

50

アレイで録音されており、指向方向 $d(f)$ に対応するチャネルの音響信号がそのまま方向音となる。そのため、それを対応するチャネルのスピーカ再生信号に割り当てればよい。

【0087】

12チャネル音響信号のフーリエ係数（ベクトル） $z(f)$ について、そのm番目の要素を $z_m(f)$ [$m=1 \sim 12$] とする。また、12チャネルのスピーカ 611 ~ 622について、各スピーカ再生信号のフーリエ係数を $X_s(f)$ [$s=1 \sim 12$] とする。ここで、指向方向数 $D(f)=4$ で、各指向方向で次のようになる周波数を考える。

$$\begin{aligned}z_1(f) &= 0^\circ, \\z_2(f) &= 90^\circ, \\z_3(f) &= 180^\circ, \\z_4(f) &= -90^\circ\end{aligned}$$

この場合は、

$$\begin{aligned}X_i(f) &= z_i(f) \quad [i=1, 4, 7, 10], \\X_j(f) &= 0 \quad [j=2, 3, 5, 6, 8, 9, 11, 12]\end{aligned}$$

となる。

【0088】

また、指向方向数 $D(f)=6$ で、各指向方向が次のようになる周波数を考える。

$$\begin{aligned}z_1(f) &= 0^\circ, \\z_2(f) &= 60^\circ, \\z_3(f) &= 120^\circ, \\z_4(f) &= 180^\circ, \\z_5(f) &= -120^\circ, \\z_6(f) &= -60^\circ\end{aligned}$$

この場合は、

$$\begin{aligned}X_i(f) &= z_i(f) \quad [i=1, 3, 5, 7, 9, 11], \\X_j(f) &= 0 \quad [j=2, 4, 6, 8, 10, 12]\end{aligned}$$

となる。

【0089】

また、指向方向数 $D(f)=12$ で、各指向方向が次のようになる周波数を考える。

$$\begin{aligned}z_1(f) &= 0^\circ, \\z_2(f) &= 30^\circ, \\z_3(f) &= 60^\circ, \\z_4(f) &= 90^\circ, \\z_5(f) &= 120^\circ, \\z_6(f) &= 150^\circ, \\z_7(f) &= 180^\circ, \\z_8(f) &= -150^\circ, \\z_9(f) &= -120^\circ, \\z_{10}(f) &= -90^\circ, \\z_{11}(f) &= -60^\circ, \\z_{12}(f) &= -30^\circ\end{aligned}$$

この場合は、

$$X_i(f) = z_i(f) \quad [i=1 \sim 12]$$

となる。

【0090】

図6(b)の太横線で示すように、 $D(f)=4$ [$f < f_M$] 、 $D(f)=6$ [$f_M \leq f < f_D$] 、 $D(f)=12$ [$f_D \leq f$] の場合、周波数 f_M 未満の方向音は、4つのスピーカ 611, 614, 617, 620 から再生される。また、周波数 f_M 以上 f_D 未満の方向音は、6つのスピーカ 611, 613, 615, 617, 619, 621 から再生される。また、周波数 f_D 以上の方向音は、全12のスピーカ 611 ~ 622 から再生される。これは、高域ほど使用するスピーカ

10

20

30

40

50

力の数が多くなる、新しいタイプのサラウンド構成である。

【0091】

S716では、S715で生成したスピーカ再生信号のフーリエ係数 $X_s(f)$ をそれぞれ逆フーリエ変換することで、時間波形であるスピーカ再生信号 $x_s(t)$ [s=1~12]を取得する。

【0092】

S717では、音響信号出力部104が、S716で取得したスピーカ再生信号 $x_s(t)$ にD/A変換および増幅を施し、スピーカ611~622から再生する。

【0093】

以上説明した実施形態によれば、周波数ごとに指向方向数を制御することで、音源の方向感が明瞭となり、各方向の音量バランスも均一となる。

10

【0094】

なお、上記実施形態において記憶部102があらかじめ保持しているとした各種データは、システム制御部101と相互に結ばれた不図示のデータ入出力部を介して、外部から入力するようにしてもよい。

【0095】

また、上記第1実施形態と第2実施形態を適宜組み合わせれば、以下のような実施形態が構成可能であり、これらの実施形態も本発明の範囲に含まれる。すなわち、(無指向性)マイクロホンアレイでの指向性形成フィルタリングによる方向音生成と、周囲スピーカ再生の組み合わせにおいて、周波数ごとに指向方向数および使用スピーカ数を制御する実施形態が構成可能である。また、指向性マイクロホンアレイによる方向音取得と、ヘッドホンでの立体音響再生の組み合わせにおいて、周波数ごとに指向方向数および仮想スピーカ数を制御する実施形態が構成可能である。

20

【0096】

なお、信号処理装置100が再生(ヘッドホン、スピーカ)の機能に加えて、録音(マイクロホンアレイ)や撮影(カメラ)・表示(ディスプレイ)の機能を備えていてもよい。このとき、撮影・録音系と表示・再生系がそれぞれ遠隔地で同期的に動作すれば、遠隔ライブシステムを実現することができる。

【0097】

なお、上記実施形態では、水平全方位で音源の方向感が明瞭となり、音量バランスも均一となるようにしていたが、対象とする方向範囲を任意に設定してもよい。例えば、水平方向だけでなく、仰角方向も含めた全方位を対象方向範囲としてもよいし、水平前方半面や、撮影した映像信号の画角範囲などに限定してもよい。この場合、例えば合成ビームパターンの凹み量の目安である標準偏差は、水平全方位ではなく対象方向範囲の合成ビームパターンから算出する。

30

【0098】

(他の実施形態)

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ー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。

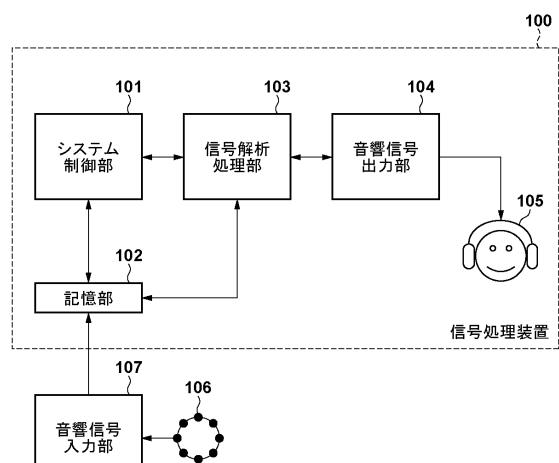
40

【符号の説明】

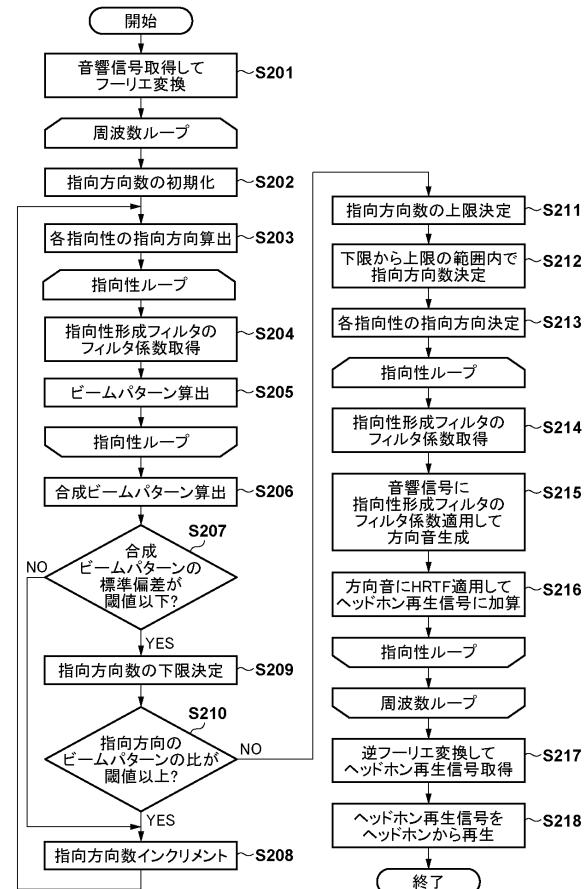
【0099】

100:信号処理装置、101:システム制御部、102:記憶部、103:信号解析処理部、104:音響信号出力部、105:ヘッドホン

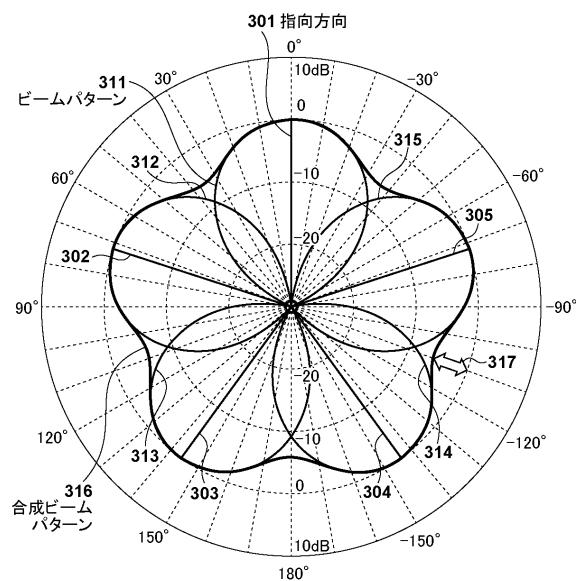
【図1】



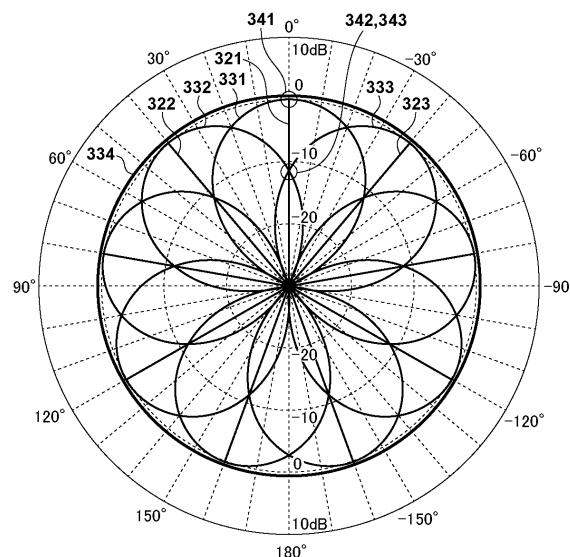
【図2】



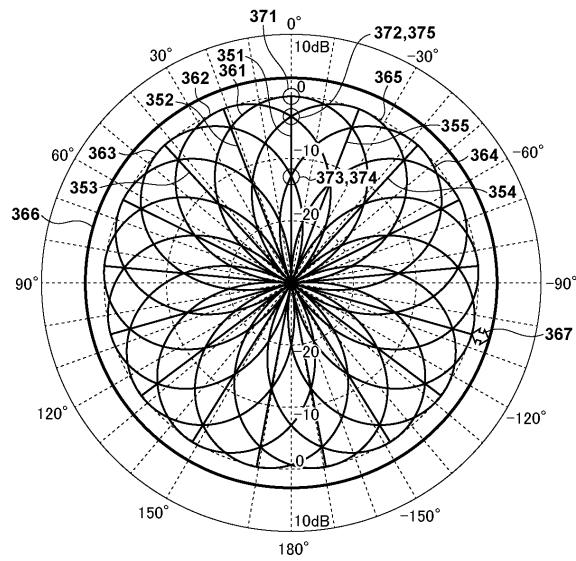
【図3】



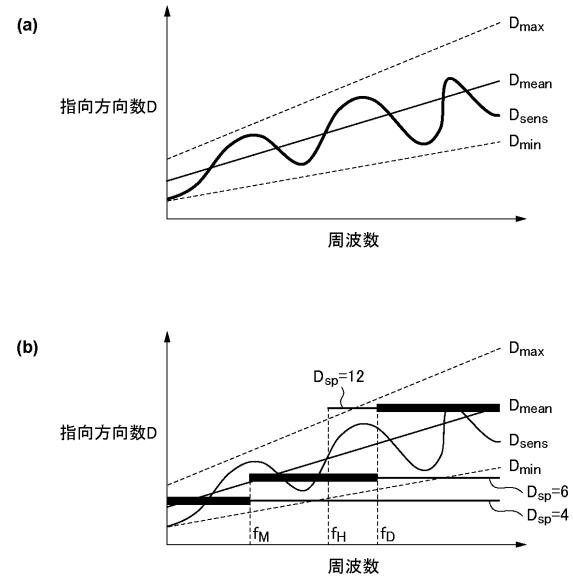
【図4】



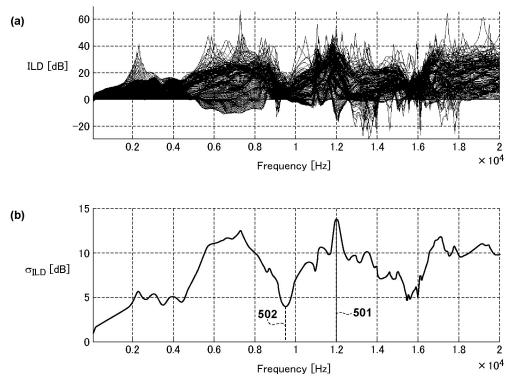
【図5】



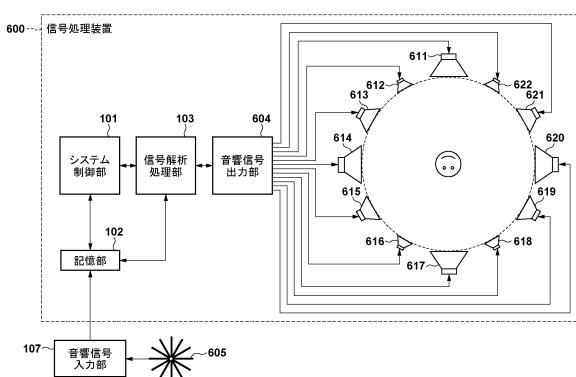
【図6】



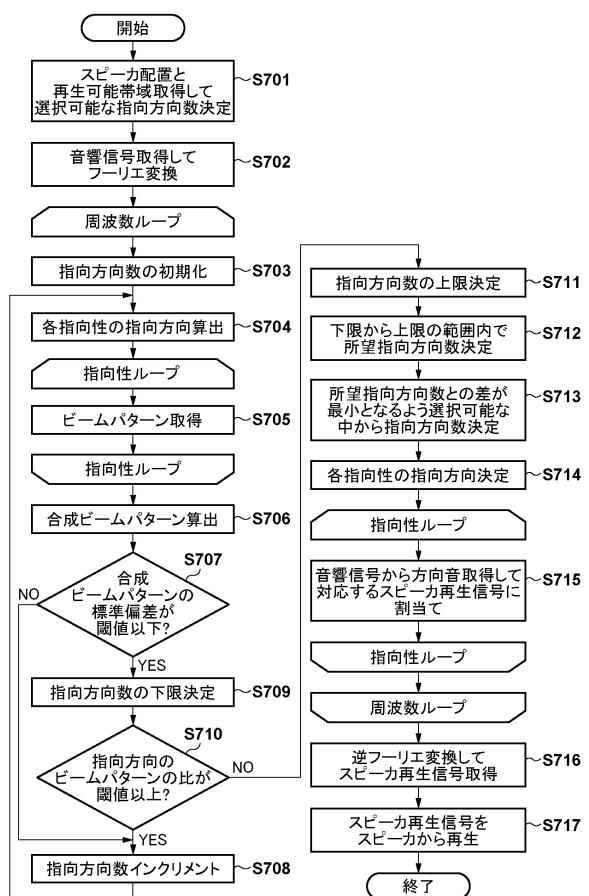
【図7】



【図8】



【図9】



フロントページの続き

(72)発明者 多和田 典朗
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

審査官 大石 剛

(56)参考文献 特開2014-017645(JP,A)
特開昭63-008641(JP,A)
特開2002-044796(JP,A)

(58)調査した分野(Int.Cl., DB名)

H 04 R	1 / 40
H 04 R	3 / 00
H 04 S	7 / 00