

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6100441号
(P6100441)

(45) 発行日 平成29年3月22日(2017.3.22)

(24) 登録日 平成29年3月3日(2017.3.3)

(51) Int.Cl.

F 1

H04S 1/00 (2006.01)
H04S 7/00 (2006.01)H04S 1/00 500
H04S 7/00 350

請求項の数 19 (全 79 頁)

(21) 出願番号 特願2016-516799 (P2016-516799)
 (86) (22) 出願日 平成26年5月28日 (2014.5.28)
 (65) 公表番号 特表2016-523466 (P2016-523466A)
 (43) 公表日 平成28年8月8日 (2016.8.8)
 (86) 國際出願番号 PCT/US2014/039864
 (87) 國際公開番号 WO2014/194005
 (87) 國際公開日 平成26年12月4日 (2014.12.4)
 審査請求日 平成28年7月28日 (2016.7.28)
 (31) 優先権主張番号 61/828,620
 (32) 優先日 平成25年5月29日 (2013.5.29)
 (33) 優先権主張国 米国(US)
 (31) 優先権主張番号 61/847,543
 (32) 優先日 平成25年7月17日 (2013.7.17)
 (33) 優先権主張国 米国(US)

(73) 特許権者 595020643
 クアアルコム・インコーポレイテッド
 QUALCOMM INCORPORATED
 アメリカ合衆国、カリフォルニア州 92
 121-1714、サン・ディエゴ、モア
 ハウス・ドライブ 5775
 (74) 代理人 100108855
 弁理士 蔵田 昌俊
 (74) 代理人 100109830
 弁理士 福原 淑弘
 (74) 代理人 100158805
 弁理士 井関 守三
 (74) 代理人 100194814
 弁理士 奥村 元宏

最終頁に続く

(54) 【発明の名称】コンテンツ解析および重み付けを用いたバイノーラル室内インパルス応答によるフィルタリング

(57) 【特許請求の範囲】

【請求項 1】

音声信号をバイノーラル化する方法であって、前記方法は、左および右の各々に関して下記を備える、

それぞれのバイノーラル音声信号を生成するために、それぞれの複数のバイノーラル室内インパルス応答(B R I R)フィルタを前記音声信号の複数のチャンネルに適用することであって、ここにおいて、前記複数のチャンネル中のチャンネルは、複数のサブグループにグループ化され、サブグループの数は、チャンネルの数よりも少なく、前記それぞれの複数のB R I Rフィルタを適用することは、

前記複数のサブグループの各それぞれのサブグループに関して、

それぞれの複数の適応的に重み付けられたチャンネルを生成すること、ここにおいて、前記それぞれのサブグループに関して前記複数の適応的に重み付けられたチャンネルを生成することは、前記それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルのサンプルに適応的に決定された重みを適用することによって、それぞれの適応的に重み付けられたチャンネルを生成することを備える、と、

結合された信号を生成するために、前記それぞれの複数の適応的に重み付けられたチャンネルを結合することと、

前記それぞれのサブグループに関するフィルタリングされた信号を生成するために、前記結合された信号に反射フィルタを適用することと、

を備える、ことと、

10

20

H R T F フィルタリングされた信号を生成するために、前記複数のチャンネルに頭部伝達関数（H R T F ）を適用することと、

前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と前記H R T F フィルタリングされた信号とを結合すること。

【請求項 2】

左および右の各々に関して、前記それぞれの複数のB R I R フィルタを適用することが、

前記複数のチャンネル中の前記チャンネルのサンプルに、追加の適応的に決定された重みを適用することによって、追加の複数の適応的に重み付けられたチャンネルを生成することと、

追加の結合された信号を生成するために、前記追加の複数の適応的に重み付けられたチャンネルを結合することと、

前記追加の結合された信号にそれぞれの残響フィルタを適用すること、ここにおいて、前記サブグループに関する前記フィルタリングされた信号と前記H R T F フィルタリングされた信号とを結合することは、前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と、前記H R T F フィルタリングされた信号と、前記追加の結合された信号とを結合することを備える、と、

を備える、請求項 1 に記載の方法。

【請求項 3】

前記方法は、左および右の各々に関して、前記残響フィルタを取得することをさらに備え、ここにおいて、前記それぞれの残響フィルタを取得することは、

前記それぞれの残響フィルタを生成するために、前記それぞれの複数のバイノーラル室内インパルス応答フィルタを正規化することなく、前記それぞれの複数のバイノーラル室内インパルス応答フィルタの各々の応答の末尾に対応する残響フィルタの平均を計算すること、

を備える、請求項 2 に記載の方法。

【請求項 4】

前記方法が、左および右の各々に関して、前記それぞれの残響フィルタを取得することをさらに備え、ここにおいて、前記それぞれの残響フィルタを取得することが、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算することと、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記それぞれの残響フィルタを合成することと、

を備える、請求項 2 に記載の方法。

【請求項 5】

前記音声信号の前記複数のチャンネルが、それぞれ、球面調和係数を備える、請求項 1 に記載の方法。

【請求項 6】

前記反射フィルタが第1の反射フィルタであり、前記複数のサブグループの各それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルの前記サンプルに適用された前記それぞれの適応的に決定された重みが、第2のエネルギー値で割った第1のエネルギー値の平方根に等しく、前記第1のエネルギー値が第2の反射フィルタのエネルギーを示し、前記第2のエネルギー値が、前記第1の反射フィルタのエネルギーを示す、

請求項 1 に記載の方法。

【請求項 7】

10

20

30

40

50

1つまたは複数のプロセッサを備えるデバイスであって、前記1つまたは複数のプロセッサは、左および右の各々に関して、

それぞれのバイノーラル音声信号を生成するために、それぞれの複数のバイノーラル室内インパルス応答（B R I R）フィルタを音声信号の複数のチャンネルに適用し、ここにおいて、前記複数のチャンネル中のチャンネルは、複数のサブグループにグループ化され、サブグループの数は、チャンネルの数よりも少なく、ここにおいて、前記1つまたは複数のプロセッサは、前記それぞれの複数のB R I Rフィルタを適用するために、前記1つまたは複数のプロセッサが、

前記複数のサブグループの各それぞれのサブグループに関して、

それぞれの複数の適応的に重み付けられたチャンネルを生成し、ここにおいて、前記1つまたは複数のプロセッサは、前記それぞれのサブグループに関して前記複数の適応的に重み付けられたチャンネルを生成することの一部として、前記1つまたは複数のプロセッサが、前記それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルのサンプルに適応的に決定された重みを適用することによって、それぞれの適応的に重み付けられたチャンネルを生成するように構成され、

結合された信号を生成するために、前記それぞれの複数の適応的に重み付けられたチャンネルを結合し、

前記それぞれのサブグループに関するフィルタリングされた信号を生成するために、前記結合された信号に反射フィルタを適用する、

ように構成され、

H R T F フィルタリングされた信号を生成するために、前記複数のチャンネルに頭部伝達関数（H R T F）を適用し、

前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と前記H R T F フィルタリングされた信号とを結合する、

ように構成される、デバイス。

【請求項8】

左および右の各々に関して、前記1つまたは複数のプロセッサは、前記それぞれの複数のB R I R フィルタを適用することの一部として、前記1つまたは複数のプロセッサが、

前記複数のチャンネル中の前記チャンネルのサンプルに、追加の適応的に決定された重みを適用することによって、追加の複数の適応的に重み付けられたチャンネルを生成し、

追加の結合された信号を生成するために、前記追加の複数の適応的に重み付けられたチャンネルを結合し、

前記追加の結合された信号にそれぞれの残響フィルタを適用し、ここにおいて、前記1つまたは複数のプロセッサは、前記サブグループに関する前記フィルタリングされた信号と前記H R T F フィルタリングされた信号とを結合することの一部として、前記1つまたは複数のプロセッサが、前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と、前記H R T F フィルタリングされた信号と、前記追加の結合された信号とを結合するように構成される、

ように構成される、請求項7に記載のデバイス。

【請求項9】

前記1つまたは複数のプロセッサは、左および右の各々に関して、前記それぞれの残響フィルタを取得するようにさらに構成され、ここにおいて、前記1つまたは複数のプロセッサは、前記それぞれの残響フィルタを取得することの一部として、前記1つまたは複数のプロセッサが、

前記それぞれの残響フィルタを生成するために、前記それぞれの複数のバイノーラル室内インパルス応答フィルタを正規化することなく、前記それぞれの複数のバイノーラル室内インパルス応答フィルタの各々の応答の末尾に対応する残響フィルタの平均を計算する、

ように構成される、請求項8に記載のデバイス。

【請求項10】

10

20

30

40

50

前記 1 つまたは複数のプロセッサは、左および右の各々に関して、前記それぞれの残響フィルタを取得するようにさらに構成され、ここにおいて、前記 1 つまたは複数のプロセッサは、前記それぞれの残響フィルタを取得することの一部として、前記 1 つまたは複数のプロセッサが、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算し、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算し、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記それぞれの残響フィルタを合成する、10

ように構成される、請求項 8 に記載のデバイス。

【請求項 1 1】

前記音声信号の前記複数のチャンネルが、それぞれ、球面調和係数を備える、請求項 7 に記載のデバイス。

【請求項 1 2】

前記反射フィルタが第 1 の反射フィルタであり、前記複数のサブグループの各それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルの前記サンプルに適用された前記それぞれの適応的に決定された重みが、第 2 のエネルギー値で割った第 1 のエネルギー値の平方根に等しく、前記第 1 のエネルギー値が第 2 の反射フィルタのエネルギーを示し、前記第 2 のエネルギー値が、前記第 1 の反射フィルタのエネルギーを示す、20

請求項 7 に記載のデバイス。

【請求項 1 3】

装置であって、

ビットストリームから音声信号の複数のチャンネルを抽出するための手段と、

左および右の各々に関して、

それぞれのバイノーラル音声信号を生成するために、それぞれの複数のバイノーラル室内インパルス応答 (B R I R) フィルタを前記音声信号の前記複数のチャンネルに適用するための手段であって、ここにおいて、前記複数のチャンネル中のチャンネルは、複数のサブグループにグループ化され、サブグループの数は、チャンネルの数よりも少なく、前記それぞれの複数の B R I R フィルタを適用するための前記手段は、30

前記複数のサブグループの各それぞれのサブグループに関して、

それぞれの複数の適応的に重み付けられたチャンネルを生成するための手段、ここにおいて、前記それぞれのサブグループに関して前記複数の適応的に重み付けられたチャンネルを生成するための前記手段は、前記それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルのサンプルに適応的に決定された重みを適用することによって、それぞれの適応的に重み付けられたチャンネルを生成するための手段を備える、と、

結合された信号を生成するために、前記それぞれの複数の適応的に重み付けられたチャンネルを結合するための手段と、40

前記それぞれのサブグループに関するフィルタリングされた信号を生成するために、前記結合された信号に反射フィルタを適用するための手段と、

を備える、手段と、

H R T F フィルタリングされた信号を生成するために、前記複数のチャンネルに頭部伝達関数 (H R T F) を適用するための手段と、

前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と前記 H R T F フィルタリングされた信号とを結合するための手段と、

を備える、装置。50

【請求項 14】

左および右の各々について、前記それぞれの複数の B R I R フィルタを適用するための前記手段が、

前記複数のチャンネル中の前記チャンネルのサンプルに、追加の適応的に決定された重みを適用することによって、追加の複数の適応的に重み付けられたチャンネルを生成するための手段と、

追加の結合された信号を生成するために、前記追加の複数の適応的に重み付けられたチャンネルを結合するための手段と、

前記追加の結合された信号にそれぞれの残響フィルタを適用するための手段、ここにおいて、前記サブグループに関する前記フィルタリングされた信号と前記 H R T F フィルタリングされた信号とを結合するための前記手段は、前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と、前記 H R T F フィルタリングされた信号と、前記追加の結合された信号とを結合するための手段を備える、と、

を備える、請求項 13 に記載の装置。

【請求項 15】

前記装置は、左および右の各々について、前記それぞれの残響フィルタを取得するための手段をさらに備え、ここにおいて、前記それぞれの残響フィルタを取得するための前記手段は、

前記それぞれの残響フィルタを生成するために、前記バイノーラル室内インパルス応答フィルタを正規化することなく、前記バイノーラル室内インパルス応答フィルタの各々の応答の末尾に対応する残響フィルタの平均を計算するための手段、

を備える、請求項 14 に記載の装置。

【請求項 16】

前記装置が、左および右の各々について、前記それぞれの残響フィルタを取得するための手段をさらに備え、ここにおいて、前記それぞれの残響フィルタを取得するための前記手段が、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算するための手段と、

前記それぞれの複数のバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算するための手段と、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記それぞれの残響フィルタを合成するための手段と、

を備える、請求項 14 に記載の装置。

【請求項 17】

前記反射フィルタが第 1 の反射フィルタであり、前記複数のサブグループの各それぞれのサブグループの各それぞれのチャンネルについて、前記それぞれのチャンネルの前記サンプルに適用された前記それぞれの適応的に決定された重みが、第 2 のエネルギー値で割った第 1 のエネルギー値の平方根に等しく、前記第 1 のエネルギー値が第 2 の反射フィルタのエネルギーを示し、前記第 2 のエネルギー値が、前記第 1 の反射フィルタのエネルギーを示す、

請求項 13 に記載の装置。

【請求項 18】

実行されると、1 つまたは複数のプロセッサに、

左および右の各々について、

それぞれのバイノーラル音声信号を生成するために、それぞれの複数のバイノーラル室内インパルス応答 (B R I R) フィルタを音声信号の複数のチャンネルに適用すること、を行わせる命令をその上に記憶した、非一時的コンピュータ可読記憶媒体であって、ここにおいて、前記複数のチャンネル中のチャンネルは、複数のサブグループにグルー

10

20

30

40

50

化され、サブグループの数は、チャンネルの数よりも少なく、前記1つまたは複数のプロセッサに前記それぞれの複数のB R I R フィルタを適用することを行わせることの一部として、前記命令は、前記1つまたは複数のプロセッサに、

前記複数のサブグループの各それぞれのサブグループに関して、

それぞれの複数の適応的に重み付けられたチャンネルを生成することを行わせ、ここにおいて、前記1つまたは複数のプロセッサに前記それぞれのサブグループに関して前記それぞれの複数の適応的に重み付けられたチャンネルを生成することを行わせることの一部として、前記命令は、前記1つまたは複数のプロセッサに、前記それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルのサンプルに適応的に決定された重みを適用することによって、それぞれの適応的に重み付けられたチャンネルを生成することを行わせる、

結合された信号を生成するために、前記それぞれの複数の適応的に重み付けられたチャンネルを結合することを行わせ、

前記それぞれのサブグループに関するフィルタリングされた信号を生成するために、前記結合された信号に反射フィルタを適用することを行わせ、

前記命令は、前記1つまたは複数のプロセッサに、

H R T F フィルタリングされた信号を生成するために、前記複数のチャンネルに頭部伝達関数 (H R T F) を適用することと、

前記それぞれのバイノーラル音声信号を生成するために、前記サブグループに関する前記フィルタリングされた信号と前記H R T F フィルタリングされた信号とを結合することと、

を行わせる、非一時的コンピュータ可読記憶媒体。

【請求項19】

前記反射フィルタが第1の反射フィルタであり、前記複数のサブグループの各それぞれのサブグループの各それぞれのチャンネルに関して、前記それぞれのチャンネルの前記サンプルに適用された前記それぞれの適応的に決定された重みが、第2のエネルギー値で割った第1のエネルギー値の平方根に等しく、前記第1のエネルギー値が第2の反射フィルタのエネルギーを示し、前記第2のエネルギー値が、前記第1の反射フィルタのエネルギーを示す、

請求項18に記載の非一時的コンピュータ可読記憶媒体。

30

【発明の詳細な説明】

【技術分野】

【0001】

優先権主張

[0001]本出願は、2013年5月29日に出願された米国仮特許出願第61/828,620号、2013年7月17日に出願された米国仮特許出願第61/847,543号、2013年10月3日に出願された米国仮出願第61/886,593号、および2013年10月3日に出願された米国仮出願第61/886,620号の利益を主張する。

【0002】

[0002]本開示は、音声レンダリングに関し、より詳細には、音声データのバイノーラルレンダリング (binaural rendering) に関する。

40

【発明の概要】

【0003】

[0003]概して、音声ストリームの源をたどるために、バイノーラル室内インパルス応答 (binaural room impulse response) (B R I R) フィルタの適用によるバイノーラル音声レンダリングに関する技法を説明する。

【0004】

[0004]一例として、音声信号をバイノーラル化する方法は、音声信号の複数の適応的に重み付けられたチャンネルを生成するために適応的に決定された重みを音声信号の複数のチャンネルに適用することと、結合された信号を生成するために音声信号の複数の適応的

50

に重み付けられたチャンネルのうちの少なくとも 2 つを結合することと、バイノーラル音声信号を生成するためにバイノーラル室内インパルス応答フィルタを結合された信号に適用することと、を備える。

【 0 0 0 5 】

[0005]別の例として、デバイスは、音声信号の複数の適応的に重み付けられたチャンネルを生成するために適応的に決定された重みを音声信号の複数のチャンネルに適用することと、結合された信号を生成するために音声信号の複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを結合することと、バイノーラル音声信号を生成するためにバイノーラル室内インパルス応答フィルタを結合された信号に適用することと、を行うように構成された 1 つまたは複数のプロセッサを備える。

10

【 0 0 0 6 】

[0006]別の例として、装置は、音声信号の複数の適応的に重み付けられたチャンネルを生成するために適応的に決定された重みを音声信号の複数のチャンネルに適用するための手段と、結合された信号を生成するために音声信号の複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを結合するための手段と、バイノーラル音声信号を生成するためにバイノーラル室内インパルス応答フィルタを結合された信号に適用するための手段と、を備える。

【 0 0 0 7 】

[0007]別の例として、非一時的コンピュータ可読記憶媒体は、実行されると、1 つまたは複数のプロセッサに、音声信号の複数の適応的に重み付けられたチャンネルを生成するために適応的に決定された重みを音声信号の複数のチャンネルに適用することと、結合された信号を生成するために音声信号の複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを結合することと、バイノーラル音声信号を生成するためにバイノーラル室内インパルス応答フィルタを結合された信号に適用することと、を行わせる命令をその上に記憶している。

20

【 0 0 0 8 】

[0008]技法の 1 つまたは複数の態様の詳細は、添付の図面および以下の説明に記載される。これらの技法の他の特徴、目的、および利点は、説明および図面から、ならびに特許請求の範囲から、明らかになろう。

【 図面の簡単な説明 】

30

【 0 0 0 9 】

【 図 1 】 [0009] 様々な次数および副次数の球面調和基底関数を示す図。

【 図 2 】 様々な次数および副次数の球面調和基底関数を示す図。

【 図 3 】 [0010] 音声信号情報をより効率的にレンダリングするために本開示で説明する技法を実行し得るシステムを示す図。

【 図 4 】 [0011] 例示的なバイノーラル室内インパルス応答 (B R I R) を示すブロック図。

【 図 5 】 [0012] 室内で B R I R を作成するための例示的なシステムモデルを示すブロック図。

【 図 6 】 [0013] 室内で B R I R を作成するためのより詳細なシステムモデルを示すブロック図。

40

【 図 7 】 [0014] 本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイスの一例を示すブロック図。

【 図 8 】 [0015] 本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイスの一例を示すブロック図。

【 図 9 】 [0016] 本開示で説明する技法の様々な態様による、球面調和係数をレンダリングするための、バイノーラルレンダリングデバイスに関する例示的な演算のモードを示すフロー図。

【 図 10 A 】 [0017] 本開示で説明する技法の様々な態様による、図 7 および図 8 の音声再生デバイスによって実施され得る代替の演算のモードを示すフロー図。

50

【図10B】本開示で説明する技法の様々な態様による、図7および図8の音声再生デバイスによって実施され得る代替の演算のモードを示すフロー図。

【図11】[0018]本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイスの一例を示すブロック図。

【図12】[0019]本開示で説明する技法の様々な態様による、図11の音声再生デバイスによって実施され得るプロセスを示すフロー図。

【図13】[0020]例示的なバイノーラル室内インパルス応答フィルタの図。

【図14】[0021]バイノーラル室内インパルス応答をマルチチャンネル音声信号に適用することによって生成されるバイノーラル出力信号の標準的計算のためのシステムを示すブロック図。

10

【図15】[0022]本明細書で説明する技法に従って、バイノーラル室内インパルス応答をマルチチャンネル音声信号に適用することによって生成されるバイノーラル出力信号を計算するためのシステムの機能的構成要素を示すブロック図。

【図16】[0023]複数のバイノーラル室内インパルス応答フィルタの反射セグメントに関する階層的クラスタ分析を示す例示的なプロット。

【図17】[0024]本開示で説明する技法による、音声再生デバイスの動作の例示的なモードを示すフローチャート。

【発明を実施するための形態】

【0010】

[0025]同様の参照符号は、図面およびテキスト全体を通して同じ要素を示す。

20

【0011】

[0026]サラウンドサウンドの発展は、現今では娯楽のための多くの出力フォーマットを利用可能にしている。そのようなサラウンドサウンドフォーマットの例は、一般的な5.1フォーマット（これは、フロントレフト（F L）と、フロントライト（F R）と、センターまたはフロントセンターと、バックレフトまたはサラウンドレフトと、バックライトまたはサラウンドライトと、低周波効果（L F E）という、6つのチャンネルを含む）、発展中の7.1フォーマット、および今後来る22.2フォーマット（たとえば、超高精細テレビ規格で使用するための）を含む。空間音声フォーマットの別の例は、球面調和係数（高次アンビソニックス（Higher Order Ambisonics）としても知られている）である。

30

【0012】

[0027]将来規格化される音声エンコーダ（PCM音声表現をビットストリームに変換するデバイス - 時間サンプルごとに必要なビット数を保存する）への入力は、随意に、3つの可能なフォーマット、（i）あらかじめ指定された位置でラウドスピーカーによって再生されることを意味する、従来のチャンネルベース音声、（ii）（様々な情報の中でも）位置座標を含む関連付けられたメタデータを有する単一音声オブジェクトのための離散的なパルス符号変調（PCM）データを含むオブジェクトベース音声、および（iii）球面調和係数（SHC）を使用して音場を表すことを含むシーンベース音声 - ここで、係数は球面調和基底関数の線形和の「重み」を表す、のうちの1つとすることができます。この文脈では、SHCは、高次アンビソニックス（HOA）モデルによるHOA信号を含み得る。球面調和係数は、代替または追加として、平面モデルと球面モデルとを含み得る。

40

【0013】

[0028]市場には様々な「サラウンドサウンド」フォーマットがある。これらのフォーマットは、たとえば、5.1ホームシアターシステム（リビングルームへの進出を行うという点でステレオ以上に最も成功した）からNHK（Nippon Hoso Kyokaiすなわち日本放送協会）によって開発された22.2システムに及ぶ。コンテンツ作成者（たとえば、ハリウッドスタジオ）は、一度に映画のサウンドトラックを作成することを望み、各々のスピーカー構成のためにサウンドトラックをリミックスする努力を行うことを望まない。最近では、標準化委員会が、標準化されたビットストリームへの符号化と、スピーカーの幾何学的配置およびレンダラの位置における音響条件に適合可能でありそ

50

れらに依存しない後続の復号とを提供するための方法を考えている。

【0014】

[0029]コンテンツ作成者にそのようなフレキシビリティを提供するために、要素の階層的なセットが音場を表すために使用され得る。要素の階層的なセットは、より低次の要素の基本セットがモデル化された音場の完全な表現を提供するように要素が順序付けられている、要素のセットを指し得る。このセットはより高次の要素を含むように拡張されるので、表現はより詳細なものになる。

【0015】

[0030]要素の階層的なセットの一例は、球面調和係数 (SHC) のセットである。次の式は、SHC を使用した音場の記述または表現を示す。

【数1】

10

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t}$$

この式は、任意の点 $\{r_r, \theta_r, \varphi_r\}$ (これは、この例において音場を取り込むマイクロフォンに対する球面座標で表される) における音場の圧力 p_i が、

20

【数2】

$$\text{SHC } A_n^m(k)$$

によって一意に表され得ることを示す。ここで、

【数3】

20

$$k = \frac{\omega}{c}$$

30

、 c は音の速さ (約 343 m/s) であり、 $\{r_r, \theta_r, \varphi_r\}$ は基準の点 (または観測点) であり、 $j_n(\cdot)$ は次数 n の球ベッセル関数であり、および

【数4】

$$Y_n^m(\theta_r, \varphi_r)$$

は次数 n および副次数 m の球面調和基底関数である。角括弧内の項は、離散フーリエ変換 (DFT)、離散コサイン変換 (DCT)、またはウェーブレット変換などの様々な時間周波数変換によって近似され得る信号の周波数領域表現 (すなわち、 $S(\cdot, r_r, \theta_r, \varphi_r)$ である) ことが認識できよう。階層的なセットの他の例は、ウェーブレット変換の係数のセットと、多重解像度の基底関数の係数の他のセットとを含む。

40

【0016】

[0031]図1は、ゼロ次 ($n = 0$) から4次 ($n = 4$) までの球面調和基底関数を示す図である。理解できるように、各次数に対して、説明を簡単にするために図示されているが、図1の例では明示的に示されていない副次数 m の拡張が存在する。

【0017】

[0032]図2は、ゼロ次 ($n = 0$) から第4次 ($n = 4$) までの球面調和基底関数を示す

50

別の図である。図 2 では、球面調和ベースの関数は、示される次数と副次数の両方を伴う 3 次元座標空間において示される。

【0018】

[0033] いずれにしても、

【数 5】

SHC $A_n^m(k)$

10

は、様々なマイクロフォンアレイ構成によって物理的に取得（たとえば、記録）されることが可能であり、または代替的に、音場のチャンネルベースの記述もしくはオブジェクトベースの記述から導出されることが可能である。SHC は、シーンに基づく音声を表す。たとえば、4 次の SHC の表現は、時間サンプルごとに $(1 + 4)^2 = 25$ 個の係数を伴う。

【0019】

[0034] これらの SHC がどのようにオブジェクトベースの記述から導出され得るかを例示するために、次の式を考える。個々の音声オブジェクトに対応する音場に関する係数

【数 6】

20

$A_n^m(k)$

は、

【数 7】

$$A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^{m*}(\theta_s, \varphi_s),$$

30

として表され得、ここで、i は

【数 8】

$$\sqrt{-1},$$

であり、 $h_n^{(2)}(\cdot)$ は次数 n の（第 2 の種類の）球ハンケル関数であり、 $\{r_s, \theta_s, \varphi_s\}$ はオブジェクトの位置である。周波数の関数としての音源のエネルギー $g(\cdot)$ を知ること（たとえば、PCM ストリームに高速フーリエ変換を行うなどの、時間・周波数解析技法を使用して）は、我々が各 PCM オブジェクトとその位置とを

【数 9】

40

SHC $A_n^m(k)$

に変換することを可能にする。さらに、各オブジェクトに関する

【数10】

$$A_n^m(k)$$

係数は、（上式は線形であり直交方向の分解であるので）加法的であることが示され得る。このようにして、多数のPCMオブジェクトが

【数11】

10

$$A_n^m(k)$$

係数によって（たとえば、個々のオブジェクトに関する係数ベクトルの和として）表され得る。本質的に、これらの係数は、音場に関する情報（3D座標の関数としての圧力）を含んでおり、上記は、観測点 $\{r_1, r_2, r_3, r_4\}$ の近傍における、音場全体の表現への個々のオブジェクトからの変換を表す。

【0020】

[0035] SHC はまた、マイクロフォンアレイの記録から次のように導出され得る。

【数12】

20

$$a_n^m(t) = b_n(r_i, t) * \langle Y_n^m(\theta_i, \varphi_i), m_i(t) \rangle$$

ただし、

【数13】

$$a_n^m(t)$$

30

は

【数14】

$$A_n^m(k)$$

(SHC) の時間領域の等価物であり、* は畳み込み演算を表し、< , > は内積を表し、 $b_n(r_i, t)$ は r_i に依存する時間領域のフィルタ関数を表し、 $m_i(t)$ は i 番目のマイクロフォンの信号であり、 i 番目のマイクロフォントランスデューサ (microphone transducer) は、半径 r_i 、仰角 θ_i 、および方位角 φ_i に位置する。したがって、マイクロフォンアレイの中に 32 個のトランスデューサがあり、各マイクロフォンが、 $r_i = a$ が定数となるように球面上に配置される (mathematics の Eigenvector EM32 デバイス上のマイクロフォンのように) 場合、25 個の SHC が、行列演算を使用して次のように導出され得る。

40

【数15】

$$\begin{bmatrix} a_0^0(t) \\ a_1^{-1}(t) \\ \vdots \\ a_4^4(t) \end{bmatrix} = \begin{bmatrix} b_0(a, t) \\ b_1(a, t) \\ \vdots \\ b_4(a, t) \end{bmatrix} * \begin{bmatrix} Y_0^0(\theta_1, \varphi_1) & Y_0^0(\theta_2, \varphi_2) & \dots & Y_0^0(\theta_{32}, \varphi_{32}) \\ Y_1^{-1}(\theta_1, \varphi_1) & Y_1^{-1}(\theta_2, \varphi_2) & \dots & Y_1^{-1}(\theta_{32}, \varphi_{32}) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \varphi_1) & Y_4^4(\theta_2, \varphi_2) & \dots & Y_4^4(\theta_{32}, \varphi_{32}) \end{bmatrix} \begin{bmatrix} m_1(a, t) \\ m_2(a, t) \\ \vdots \\ m_{32}(a, t) \end{bmatrix}$$

上記の式中の行列は、より一般的には $E_s(\cdot, \cdot)$ と呼ばれることがあり、ここで、下付き文字 s は、この行列がある特定の変換器幾何学的配置セット s に関する事を示すことができる。上記の式中の畳み込み (* によって示される) は、行と行に基づき、したがって、たとえば、出力

【数16】

$$a_0^0(t)$$

は $b_0(a, t)$ と、 $E_s(\cdot, \cdot)$ 行列の第1の行とマイクロフォン信号の列（これは時間の関数として変化するベクトル乗算の結果が時系列であるという事実の理由である）とのベクトル乗算から生じる時系列と、の間の畳み込みの結果である。算出は、マイクロフォンアレイの変換器位置が、いわゆる T 字形設計幾何学的配置 (Eigenmike) 変換器幾何学的配置に極めて近い）にあるとき、最も正確であり得る。T字形設計幾何学的配置の1つの特徴は、幾何学的配置から生じる $E_s(\cdot, \cdot)$ 行列は行儀の非常によい (very well behaved) 逆行列（または擬似逆行列）を有すること、さらに、この逆行列は行列 $E_s(\cdot, \cdot)$ の転置によって極めてよく近似され得ることが多いことであり得る。仮に $b_n(a, t)$ を用いたフィルタリング動作が無視される場合、この性質は、SHC からのマイクロフォン信号の復元（すなわち、この例では、 $[m_i(t)] = [E_s(\cdot, \cdot)]^{-1} [SHC]$ ）を可能にする。残りの数字は、以下でオブジェクトベース音声コーディングおよび SHC ベース音声コーディングの文脈で説明される。

【0021】

[0036] 図3は、音声信号情報をより効率的にレンダリングするために本開示で説明する技法を実行し得るシステム20を示す図である。図3の例に示すように、システム20は、コンテンツ作成者22と、コンテンツ消費者24とを含む。コンテンツ作成者22およびコンテンツ消費者24の文脈で説明するが、本技法は、音場の階層的表示を規定するSHC または任意の他の階層要素を利用する任意の文脈において実施され得る。

【0022】

[0037] コンテンツ作成者22は、コンテンツ消費者24などのコンテンツ消費者による消費のためのマルチチャンネル音声コンテンツを生成し得る映画撮影所または他のエンティティを表すことができる。多くの場合、このコンテンツ作成者は、ビデオコンテンツとともに、音声コンテンツを生成する。コンテンツ消費者24は、音声再生システムを所有するまたはそれにアクセスできる個人を表し得、その音声再生システムはマルチチャンネル音声コンテンツを再生する能力がある音声再生システムの任意の形を指し得る。図3の例では、コンテンツ消費者24は、音場の階層的表示を規定する階層要素をレンダリングするための音声再生システム32を所有するかまたはそれへのアクセスを有する。

【0023】

[0038] コンテンツ作成者22は、音声レンダラ28と音声編集システム30とを含む。音声レンダラ28は、スピーカーフィード（「ラウドスピーカーフィード」、「スピーカー信号」、または「ラウドスピーカー信号」と呼ばれることがある）をレンダリングするかまたはさもなければ生成する音声処理ユニットを表し得る。各スピーカーフィードは、

10

20

30

40

50

マルチチャンネル音声システムの特定のチャンネルに対する音を再生するスピーカーフィード、またはスピーカー位置に適合する頭部伝達関数（HRTF）フィルタとの畳み込みについて意図される仮想ラウドスピーカーフィードに対応することができる。各スピーカーフィードは、球面調和係数のチャンネル（ここで、チャンネルは、球面調和係数が対応する関連付けられた球面基底関数の次数および／または副次数によって示され得る）に対応し得、指向性音場を表すためにSHCの多数のチャンネルを使用する。

【0024】

[0039]図3の例では、音声レンダラ28は、従来の5.1、7.1、または22.2のサラウンドサウンドフォーマットのためのスピーカーフィードをレンダリングし、5.1、7.1、または22.2のサラウンドサウンドスピーカーシステムにおいて、5個、7個、または22個のスピーカーの各々に関するスピーカーフィードを生成することができる。代替的に、音声レンダラ28は、上記で検討した音源の球面調和係数の性質が与えられれば、任意の数のスピーカーを有する任意のスピーカー構成のための音源の球面調和係数からスピーカーフィードをレンダリングするように構成され得る。音声レンダラ28は、このようにして、図3ではスピーカーフィード29と示されているいくつかのスピーカーフィードを生成し得る。

【0025】

[0040]コンテンツ作成者は、編集プロセス中に、球面調和係数27（「SHC27」）をレンダリングし、高い忠実度を持たないまたは説得力のあるサラウンドサウンド経験を提供しない音場の様相を識別する試みにおけるレンダリングされたスピーカーフィードをリッスンすることができる。次いで、コンテンツ作成者22は、（多くの場合、上記の様式で音源の球面調和係数が導出され得る異なるオブジェクトの操作を通じて、間接的に）音源の球面調和係数を編集することができる。コンテンツ作成者22は、球面調和係数27を編集するために音声編集システム30を用いることができる。音声編集システム30は、音声データを編集し、この音声データを1つまたは複数の音源の球面調和係数として出力することが可能な任意のシステムを表す。

【0026】

[0041]編集プロセスが完了すると、コンテンツ作成者22は、球面調和係数27に基づいてビットストリーム31を生成することができる。すなわち、コンテンツ作成者22は、ビットストリーム生成デバイス36を含み、それは、ビットストリーム31を生成する能力がある任意のデバイスを表し得る。場合によっては、ビットストリーム生成デバイス36は、球面調和係数27を帯域幅圧縮し（一例として、エントロピー符号化を通じて）、ビットストリーム31を形成するために認められたフォーマットで球面調和係数27のエントロピー符号化バージョンを配置するエンコーダを表し得る。他の例では、ビットストリーム生成デバイス36は、一例としてマルチチャンネル音声コンテンツまたはその派生物を圧縮するために従来の音声サラウンドサウンド符号化プロセスのプロセスに類似したプロセスを使用してマルチチャンネル音声コンテンツ29を符号化する音声エンコーダ（おそらく、MP3サウンドなどの知られている音声コーディング規格またはその派生物に適合する音声エンコーダ）を表すことができる。圧縮されたマルチチャンネル音声コンテンツ29は次いで、コンテンツ29を帯域幅圧縮するためにエントロピー符号化されまたはある他の方法でコーディングされ、ビットストリーム31を形成するために合意したフォーマットに従って配置されてもよい。ビットストリーム31を形成するために直接圧縮されようと、レンダリングされ、次いでビットストリーム31を形成するために圧縮されようと、コンテンツ作成者22は、コンテンツ消費者24にビットストリーム31を送信することができる。

【0027】

[0042]図3ではコンテンツ消費者24に直接送信されるとして示されるが、コンテンツ作成者22は、コンテンツ作成者22とコンテンツ消費者24との間に位置付けられる中間デバイスにビットストリーム31を出力し得る。この中間デバイスは、このビットストリームを要求し得るコンテンツ消費者24への後の配送のためにビットストリーム31を

10

20

30

40

50

記憶し得る。中間デバイスは、ファイルサーバ、ウェブサーバ、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、モバイルフォン、スマートフォン、または音声デコーダによる後の取り出しがためにビットストリーム31を記憶する能力がある任意の他のデバイスを備え得る。この中間デバイスは、ビットストリーム31を要求するコンテンツ消費者24などの加入者にビットストリーム31を（おそらくは対応するビデオデータビットストリームを送信するとともに）ストリーミングすることが可能なコンテンツ配信ネットワークに存在し得る。代替的に、コンテンツ作成者22は、コンパクトディスク、デジタルビデオディスク、高精細度ビデオディスク、または他の記憶媒体などの記憶媒体にビットストリーム31を格納することができ、記憶媒体の大部分はコンピュータによって読み取り可能であり、したがって、コンピュータ可読記憶媒体または非一時的コンピュータ可読記憶媒体と呼ばれ得る。この文脈において、送信チャンネルは、これらの媒体に格納されたコンテンツが送信されるチャンネルを指し得る（および、小売店と他の店舗ベースの配信機構とを含み得る）。したがって、いずれにしても、本開示の技法は、この点に関して図3の例に限定されるべきではない。

【0028】

[0043]図3の例にさらに示すように、コンテンツ消費者24は、音声再生システム32を所有するかまたはそれへのアクセスを有する。音声再生システム32は、マルチチャンネル音声データを再生することが可能な任意の音声再生システムを表すことができる。音声再生システム32は、バイノーラルスピーカーフィード35A～35B（総称して「スピーカーフィード35」）としての出力に関するSHC27'をレンダリングするバイノーラル音声レンダラ34を含む。バイノーラル音声レンダラ34は、ベクトルベース振幅パニング（V B A P : vector-base amplitude panning）を実施する様々な方法のうちの1つまたは複数、および/または音場合成を実施する様々な方法のうちの1つまたは複数など、異なる形態のレンダリングを提供し得る。本明細書で使用するA「および/または」Bは、A、B、またはAとBとの結合を指すことができる。

【0029】

[0044]音声再生システム32は、抽出デバイス38をさらに含むことができる。抽出デバイス38は、一般にビットストリーム生成デバイス36のプロセスに相反し得るプロセスによって球面調和係数27'（球面調和係数27の修正された形態または複製物を表すことができる「SHC27'」）を抽出することが可能な任意のデバイスを表すことができる。いずれにしても、音声再生システム32は、球面調和係数27'を受信し、球面調和係数27'をレンダリングするためにバイノーラル音声レンダラ34を使用し、それによって（音声再生システム32に電気的にまたはおそらくワイヤレスに結合されるラウドスピーカーの数に対応する、このことは例示を容易にするために図3の例には示さない）スピーカーフィード35を生成し得る。スピーカーフィード35の数は2であり得、音声再生システムは、2つの対応するラウドスピーカーを含む一対のヘッドフォンにワイヤレスに結合し得る。しかしながら、様々な例では、バイノーラル音声レンダラ34は、図3に関して図示され、最初に説明されたものより多数または少数のスピーカーフィードを出力することがある。

【0030】

[0045]インパルス位置において生成されたインパルスに関する位置における応答をそれぞれ表す音声再生システムのバイナリ室内インパルス応答（B R I R）フィルタ37。B R I R フィルタ37は、それらがそれぞれ、その位置において人間の耳によって経験されるであろうインパルス応答を表すように生成されるという点において「バイノーラル」である。したがって、インパルスに関するB R I R フィルタは、対のうちの1つの要素が左耳用であり別の要素が右耳用である、対を成すサウンドレンダリングのために生成され、使用されることが多い。図示の例では、バイノーラル音声レンダラ34は、それぞれのバイノーラル音声出力35Aおよび35Bをレンダリングするために、左B R I R フィルタ33Aと右B R I R フィルタ33Bとを使用する。

【0031】

10

20

30

40

50

[0046]たとえば、B R I R フィルタ 3 7 は、音源信号と、インパルス応答 (I R) として測定された頭部伝達関数 (H R T F) とを畳み込むことによって生成され得る。B R I R フィルタ 3 7 の各々に対応するインパルス位置は、仮想空間中の仮想ラウドスピーカーの位置を表し得る。いくつかの例では、バイノーラル音声レンダラ 3 4 は、S H C 2 7 ' と、仮想ラウドスピーカーに対応するB R I R フィルタ 3 7 とを畳み込み、次いで、スピーカーフィード 3 5 としての出力に対してS H C 2 7 ' によって規定される音場をレンダリングするために、得られる畳み込みを累積する（すなわち、合計する）。本明細書で説明するように、バイノーラル音声レンダラ 3 4 は、スピーカーフィード 3 5 としてS H C 2 7 ' をレンダリングしながら、B R I R フィルタ 3 7 を操作することによってレンダリング計算を削減するための技法を適用し得る。

10

【 0 0 3 2 】

[0047]いくつかの例では、本技法は、B R I R フィルタ 3 7 を、室内の一位置における一インパルス応答の異なる段階を表すいくつかのセグメントにセグメント化することを含む。これらのセグメントは、音場の任意の点における圧力（または圧力の欠如）を生成する異なる物理現象に対応する。たとえば、B R I R フィルタ 3 7 の各々はインパルスと同時に計時されるので、第 1 のセグメントまたは「初期」セグメントは、インパルスの位置からの圧力波がインパルス応答が測定される位置に到達するまでの時間を表し得る。タイミング情報を別として、それぞれの初期セグメントに関するB R I R フィルタ 3 7 の値は重要ではなく、音場を記述する階層要素との畳み込みから除外されてよい。同様に、B R I R フィルタ 3 7 の各々は、たとえば、人間の聴覚のダイナミックレンジより低く減衰された、または指定されたしきい値より低く減衰されたインパルス応答信号を含む最終または「末尾」セグメントを含み得る。それぞれの末尾セグメントに関するB R I R フィルタ 3 7 の値もまた重要ではなく、音場を記述する階層要素との畳み込みから除外されてよい。いくつかの例では、本技法は、指定されたしきい値を用いてシュレーダの後方積分 (Schroeder backward integration) を実施すること、および後方積分が指定されたしきい値を超える場合に末尾セグメントから要素を除くことによって末尾セグメントを決定することを含むことがある。いくつかの例では、指定されたしきい値は、残響時間 R T₆₀ について - 6 0 d B である。

20

【 0 0 3 3 】

[0048] B R I R フィルタ 3 7 の各々の追加のセグメントは、室からのエコー効果を含まない、インパルスで生じた圧力波に起因するインパルス応答を表し得る。これらのセグメントは、B R I R フィルタ 3 7 に関する頭部伝達関数 (H R T F) として表され、説明され得、ここで、H R T F は、圧力波が鼓膜まで進むにつれて頭、肩 / 胳、および外耳の周りの圧力波の回折および反射によるインパルス応答を取り込む。H R T F インパルス応答は、線形時不变系 (L T I : linear and time-invariant system) の結果であり、最小位相フィルタとしてモデル化され得る。いくつかの例では、レンダリングの間のH R T F セグメント計算を削減するための技法は、最小位相再構成を含み、元の有限インパルス応答 (F I R) フィルタ（たとえば、H R T F フィルタセグメント）の次数を削減するために、無限インパルス応答 (I I R) フィルタを使用することができる。

30

【 0 0 3 4 】

[0049] I I R フィルタとして実装される最小位相フィルタは、削減されたフィルタ次数を有するB R I R フィルタ 3 7 に関するH R T F フィルタを近似するために使用され得る。次数を削減することは、周波数領域において時間ステップに関する計算の数が付随して削減することをもたらす。加えて、最小位相フィルタの構築に起因する残余 / 余剰フィルタが、音の圧力波が音源から各耳まで進む距離によって引き起こされる時間距離または位相距離を表す両耳間時間差 (I T D : interaural time difference) を推定するために使用され得る。次いで、I T D は、1つまたは複数のB R I R フィルタ 3 7 と、音場を記述する（すなわち、バイノーラル化を決定する）階層要素との畳み込みを計算した後、片耳または両耳に関する音の定位をモデル化するために使用され得る。

40

【 0 0 3 5 】

50

[0050] またさらに、B R I R フィルタ 3 7 の各々のセグメントが H R T F セグメントに後続し、インパルス応答に関する室内の効果を説明し得る。この室内セグメントは、早期エコー（または「早期反射」）セグメントと後期残響セグメントとにさらに分解され得る（すなわち、早期エコーおよび後期残響が、それぞれ、B R I R フィルタ 3 7 の各々の別個のセグメントによって表され得る）。H R T F データが B R I R フィルタ 3 7 に関して利用可能である場合、早期エコーセグメントの開始は、H R T F セグメントを識別するために B R I R フィルタ 3 7 と H R T F との逆畳み込みを行うことによって識別され得る。早期エコーセグメントが、H R T F セグメントに後続する。残余室内応答とは異なり、H R T F セグメントおよび早期エコーセグメントは、対応する仮想スピーカーの位置が重要な点における信号を決定するという点において方向依存性である。

10

【0036】

[0051] いくつかの例では、バイノーラル音声レンダラ 3 4 は、音場を記述する階層要素に関する球面調和領域（ 、 ）または他の領域のために準備された B R I R フィルタ 3 7 を使用する。すなわち、B R I R フィルタ 3 7 は、バイノーラル音声レンダラ 3 4 が、B R I R フィルタ 3 7 の（たとえば、左 / 右の）対称性および S H C 2 7 の対称性を含む、データセットのいくつかの特性を利用しながら高速畳み込みを実施することを可能にするために、球面調和領域（ S H D ）において、変換された B R I R フィルタ 3 7 として規定され得る。そのような例では、変換された B R I R フィルタ 3 7 は、S H C レンダリング行列と元の B R I R フィルタとを乗算する（または時間領域において畳み込みを行う）ことによって生成され得る。数学的に、これは、下式（1）～（5）

20

【数17】

$$B R I R'_{(N+1)^2, L, left} = S H C_{(N+1)^2, L} * B R I R_{L, left} \quad (1)$$

【数18】

$$B R I R'_{(N+1)^2, L, right} = S H C_{(N+1)^2, L} * B R I R_{L, right} \quad (2)$$

30

【数19】

$$B R I R'_{(N+1)^2, L, right} = \begin{bmatrix} Y_0^0(\theta_1, \varphi_1) & Y_0^0(\theta_2, \varphi_2) & \dots & Y_0^0(\theta_L, \varphi_L) \\ Y_1^{-1}(\theta_1, \varphi_1) & Y_1^{-1}(\theta_2, \varphi_2) & \dots & Y_1^{-1}(\theta_L, \varphi_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \varphi_1) & Y_4^4(\theta_2, \varphi_2) & \dots & Y_4^4(\theta_L, \varphi_L) \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_L \end{bmatrix}^T \quad (3)$$

【数20】

40

$$B R I R''_{(N+1)^2, left} = \sum_{k=0}^{L-1} [B R I R'_{(N+1)^2, k, left}] \quad (4)$$

【数21】

$$B R I R''_{(N+1)^2, right} = \sum_{k=0}^{L-1} [B R I R'_{(N+1)^2, k, right}] \quad (5)$$

50

に従って表現され得る。

【0037】

[0052]ここで、(3)は、(1)または(2)のいずれかを、4次の球面調和係数に関する行列形式で示す(これは、4次以下の球面基底関数と関連付けられた球面調和係数の行列形式を表すための代替方法であり得る)。式(3)は、当然ながら、より高次またはより低次の球面調和係数に関して修正され得る。式(4)～式(5)は、合計されたSHC-バイノーラルレンダリング行列(BRIR')'を生成するために、変換された左および右のBRIRフィルタ37をラウドスピーカー次元Lにわたって合計することを示す。相まって、合計されたSHC-バイノーラルレンダリング行列は、次元[(N+1)²、Length、2]を有し、ここで、Lengthは、式(1)～式(5)の任意の結合が適用され得るインパルス応答ベクトルの長さである。式(1)および式(2)のいくつかの例では、レンダリング行列SHCは、式(1)が、 $BRIR'_{(N+1)2,L, left} = SHC_{(N+1)2,L, left} * BRIR_{L, left}$ に修正され、式(2)が、 $BRIR'_{(N+1)2,L, right} = SHC_{(N+1)2,L} * BRIR_{L, right}$ に修正され得るように、バイノーラル化され得る。

【0038】

[0053]上式(1)～(3)において提示される行列をレンダリングするSHC、SHC'は、SHC27'の次数/副次数の結合の各々に関する要素を含み、それは、別個のSHCチャンネルを効率的に規定し、ここで、要素の値は、球面調和領域内のスピーカーLの位置に関するセットである。 $BRIR_{L, left}$ は、左耳、またはスピーカーLに関する位置で生成されたインパルスに関する位置におけるBRIR応答を表し、 $\{i | i \in [0, L]\}$ に関するインパルス応答ベクトル B_i を使用して(3)で表される。 $BRIR'_{(N+1)2,L, left}$ は、「SHC-バイノーラルレンダリング行列」の半分、すなわち、球面調和領域に変換された、左耳またはスピーカーLに関する位置で生成されたインパルスに関する位置におけるSHC-バイノーラルレンダリング行列を表す。 $BRIR'_{(N+1)2,L, right}$ は、SHC-バイノーラルレンダリング行列の他方の半分を表す。

【0039】

[0054]いくつかの例では、本技法は、変換されたBRIRフィルタ37とSHC-バイノーラルレンダリング行列とを生成するために、それぞれの元のBRIRフィルタ37のHRTFおよび早期反射セグメントだけにSHCレンダリング行列を適用することを含み得る。これは、SHC27'との畳み込みの長さを削減し得る。

【0040】

[0055]いくつかの例では、式(4)～(5)に表されるように、球面調和領域における様々なラウドスピーカーを組入れる次元を有するSHC-バイノーラルレンダリング行列は、SHCレンダリングとBRIRレンダリング/ミキシングとを結合する(N+1)²*Length*2のフィルタ行列を生成するように合計され得る。すなわち、L個のラウドスピーカーの各々に関するSHC-バイノーラルレンダリング行列は、たとえば、係数をL次元にわたって合計することによって結合され得る。長さLengthのSHC-バイノーラルレンダリング行列に関して、これは、信号をバイノーラル化するために球面調和係数の音声信号に適用され得る(N+1)²*Length*2の合計された、SHC-バイノーラルレンダリング行列を作成する。Lengthは、本明細書で説明する技法に従ってセグメント化されたBRIRフィルタのセグメントの長さであり得る。

【0041】

[0056]モデル節減のための技法はまた、変更されたレンダリングフィルタに適用され得、それは、SHC27'(たとえば、SHCコンテンツ)が新しいフィルタ行列(合計されたSHC-バイノーラルレンダリング行列)で直接フィルタリングされることを可能にする。次いで、バイノーラル音声レンダラ34は、バイノーラル出力信号35A、35Bを取得するためにフィルタリングされたアレイを合計することによってバイノーラル音声に変換し得る。

【0042】

[0057]いくつかの例では、音声再生システム32のBRIRフィルタ37は、上記で説

10

20

30

40

50

明した技法のうちの任意の1つまたは複数に従って以前に計算された球面調和領域における変換されたB R I R フィルタを表す。いくつかの例では、元のB R I R フィルタ37の変換は、実行時に実施され得る。

【0043】

[0058]いくつかの例では、B R I R フィルタ37は一般的に対称であるので、本技法は、左または右のいずれかの耳に関するS H C - バイノーラルレンダリング行列だけを使用することによって、バイノーラル出力35A、35Bの計算のさらなる節減を促進することができる。フィルタ行列によってフィルタリングされたS H C 27'を合計するとき、バイノーラル音声レンダラ34は、最終出力をレンダリングするとき、第2のチャンネルとしての出力信号35A、35Bのいずれかに関して、条件付き決定を行うことができる。本明細書で説明するように、左または右のいずれかの耳に対して記述された、処理コンテンツまたは修正レンダリング行列に対する言及は、他方の耳に同様に適用可能であるものと理解されるべきである。10

【0044】

[0059]このようにして、本技法は、除外されたB R I R フィルタサンプルと複数のチャンネルとの直接の疊み込みを潜在的に回避するために、B R I R フィルタ37の長さを削減するための複数の手法を提供し得る。その結果、バイノーラル音声レンダラ34は、S H C 27'からのバイノーラル出力信号35A、35Bの効率的なレンダリングを提供し得る。20

【0045】

[0060]図4は、例示的なバイノーラル室内インパルス応答(B R I R)を示すブロック図である。B R I R 40は、5つのセグメント42A～42Eを示す。初期セグメント42Aおよび末尾セグメント42Eは共に、いずれも、重要でなく、レンダリング計算から除外されてよい静止サンプルを含む。頭部伝達関数(H R T F)セグメント42Bは、頭部伝達によるインパルス応答を含み、本明細書で説明する技法を使用して識別され得る。早期エコー(代替として「早期反射」)セグメント42Cおよび後期室内残響セグメント42Dは、H R T F と室内効果とを結合する、すなわち、早期エコーセグメント42Cのインパルス応答は、室内の早期エコーおよび後期残響によってフィルタリングされたB R I R 40に関するH R T F のインパルス応答に匹敵する。しかしながら、早期エコーセグメント42Cは、後期室内残響セグメント42Dと比較して、より離散的なエコーを含むことがある。ミキシング時間は、早期エコーセグメント42Cと後期室内残響セグメント42Dとの間の時間であり、早期エコーが密な残響になる時間を示す。ミキシング時間は、H R T F の中に約 1.5×10^4 サンプルにおいて、またはH R T F セグメント42Bの開始から約 7.0×10^4 サンプルにおいて発生するように図示されている。いくつかの例では、本技法は、統計データと室内容積からの推定とを使用してミキシング時間を計算することを含む。いくつかの例では、50%の内部信頼 t_{mp50} を有する知覚のミキシング時間は約36ミリ秒(ms)であり、95%信頼区間 t_{mp95} を有する知覚のミキシング時間は約80msである。いくつかの例では、B R I R 40に対応するフィルタの後期室内残響セグメント42Dは、コヒーレンス整合された雑音末尾(coherence-matched noise tail)を使用して合成され得る。30

【0046】

[0061]図5は、室内で図4のB R I R 40などのB R I R を作成するための例示的なシステムモデル50を示すブロック図である。このモデルは、ここでは室内52AおよびH R T F 52Bの、カスケード接続されたシステムを含む。H R T F 52Bがインパルスに対して適用された後、インパルス応答は、室内52Aの早期エコーによってフィルタリングされたH R T F のインパルス応答に匹敵する。40

【0047】

[0062]図6は、室内で図4のB R I R 40などのB R I R を作成するための、より詳細なシステムモデル60を示すブロック図である。このモデル60はまた、ここではH R T F 62A、早期エコー62B、および残余室内62C(これはH R T F と室内エコーとを50

結合する)の、カスケード接続されたシステムを含む。モデル 60 は、室内 52A を早期エコー 62B および残余室内 62C に分解することを示し、各システム 62A、62B、62C を線形時不变として取り扱う。

【0048】

[0063] 早期エコー 62B は、残余室内 62C より離散的なエコーを含む。したがって、早期エコー 62B は仮想スピーカーチャンネルごとに変化し得、一方、より長い末尾を有する残余室内 62C は、単一のステレオコピーとして合成され得る。B R I R を取得するために使用されるいくつかの測定用マネキンに関して、H R T F データが、無響室内で測定されるなど入手可能である。早期エコー(「反射」と呼ばれることがある)の位置を識別するために、早期エコー 62B が、B R I R および H R T F のデータを逆疊み込みを行うことによって決定され得る。いくつかの例では、H R T F データはすぐに入手可能ではなく、早期エコー 62B を識別するための技法はプラインド推定を含む。しかしながら、単純な手法は、最初の数ミリ秒(たとえば、最初の 5、10、15、または 20 ms)を、H R T F によってフィルタリングされた直接インパルスと見なすことを含み得る。上記のように、本技法は、統計データと室内容積からの推定とを使用してミキシング時間を計算することを含み得る。

【0049】

[0064] いくつかの例では、本技法は、残余室内 62C に関して 1つまたは複数の B R I R フィルタを合成することを含み得る。ミキシング時間の後、B R I R 残響の末尾(図 6 にシステムの残余室内 62C として表される)は、いくつかの例では、知覚の代償なしに交換され得る。さらに、B R I R 残響の末尾は、エネルギー・ディケイ・リーフ(E D R : Energy Decay Relief)と周波数依存性両耳間コヒーレンス(F D I C : Frequency-Dependent Interaural Coherence)とに適合するガウスノイズで合成され得る。いくつかの例では、共通の合成 B R I R 残響の末尾が、複数の B R I R フィルタに関して生成され得る。いくつかの例では、共通の E D R は、すべてのスピーカーの E D R の平均であり得、または平均エネルギーに匹敵するエネルギーを有するフロントゼロ度 E D R (front zero degree EDR)であり得る。いくつかの例では、F D I C は、すべてのスピーカーにわたる平均 F D I C であり得、または広い空間に関する最大限に相關のない測定に関する、すべてのスピーカーにわたった最小値であってよい。いくつかの例では、残響の末尾はまた、フィードバック遅延ネットワーク(F D N : Feedback Delay Network)による人工的残響を用いてシミュレーションされ得る。

【0050】

[0065] 共通の残響の末尾によって、対応する B R I R フィルタの後ろの部分は、各スピーカーフィードとの個別の疊み込みから除外され得るが、代わりに、一度、すべてのスピーカーフィードのミックスに適用され得る。上記のように、および以下でさらに詳細に説明するように、すべてのスピーカーフィードのミキシングは、球面調和係数信号レンダリングを用いてさらに簡素化され得る。

【0051】

[0066] 図 7 は、本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイスの一例を示すブロック図である。単一のデバイス、すなわち図 7 の例における音声再生デバイス 100 として示されているが、技法は、1つまたは複数のデバイスによって実施され得る。したがって、本技法はこの点において限定され得ない。

【0052】

[0067] 図 7 の例に示すように、音声再生デバイス 100 は、抽出ユニット 104 とバイノーラルレンダリングユニット 102 とを含み得る。抽出ユニット 104 は、ビットストリーム 120 から符号化音声データを抽出するように構成されたユニットを表し得る。抽出ユニット 104 は、球面調和係数(S H C) 122 (これは、S H C 122 が、1 より大きい次数と関連付けられた少なくとも 1 つの係数を含み得るという点において高次アンビソニックス(H O A : higher order ambisonics)と呼ばれることがある)の形態の抽

10

20

30

40

50

出された符号化音声データをバイノーラルレンダリングユニット146に転送し得る。

【0053】

[0068] いくつかの例では、音声再生デバイス100は、SHC122を生成するために符号化音声データを復号するように構成された音声復号ユニットを含む。音声復号ユニットは、いくつかの態様においてSHC122を符号化するために使用される音声符号化プロセスと相反する音声復号プロセスを実施し得る。音声復号ユニットは、符号化音声データのSHCを時間領域から周波数領域に変換するように構成された時間周波数解析ユニットを含み得、それによってSHC122を生成する。すなわち、符号化音声データが、時間領域から周波数領域に変換されていない、SHC122の圧縮形態を表すとき、音声復号ユニットは、SHC122（周波数領域で指定される）を生成するように、SHCを時間領域から周波数領域に変換するために時間周波数解析ユニットを起動し得る。時間周波数解析ユニットは、SHCを時間領域から周波数領域におけるSHC122に変換するために、数例を提示すると、高速フーリエ変換（FFT）と、離散コサイン変換（DCT）と、修正離散コサイン変換（MDCT）と、離散サイン変換（DST）とを含む、フーリエベースの変換の任意の形式を適用し得る。いくつかの例では、SHC122は、すでに、ビットストリーム120において周波数領域内で指定され得る。これらの例では、時間周波数解析ユニットは、変換を適用することなく、またはさもなければ受信されたSHC122を変換することなく、SHC122をバイノーラルレンダリングユニット102に送ることができる。周波数領域で指定されたSHC122に関して説明したが、本技法は、時間領域で指定されたSHC122に関して実施され得る。

【0054】

[0069] バイノーラルレンダリングユニット102は、SHC122をバイノーラル化するように構成されたユニットを表す。言い換えれば、バイノーラルレンダリングユニット102は、SHC122を左および右のチャンネルにレンダリングするように構成されたユニットを表し、そのユニットは、SHC122が記録された室内において、左および右のチャンネルがリスナーによってどのように聞こえうるかのモデル化する空間化の機能を備え得る。バイノーラルレンダリングユニット102は、ヘッドフォンなどのヘッドセットを介する再生に好適な左チャンネル136Aと右チャンネル136B（これらは「チャンネル136」と総称されることがある）とを生成するためにSHC122をレンダリングし得る。図7の例に示すように、バイノーラルレンダリングユニット102は、B R I R フィルタ108と、B R I R 調整ユニット106と、残余室内応答ユニット110と、B R I R SHC - 領域変換ユニット112と、畳み込みユニット114と、結合ユニット116とを含む。

【0055】

[0070] B R I R フィルタ108は、1つまたは複数のB R I R フィルタを含み、図3のB R I R フィルタ37の一例を表し得る。B R I R フィルタ108は、左および右のH R T FがそれぞれのB R I R に与える影響を表す、個別のB R I R フィルタ126A、126Bを含み得る。

【0056】

[0071] B R I R 調整ユニット106は、仮想のラウドスピーカーLの各々ごとの、それぞれ長さNを有するB R I R フィルタ126A、126Bの、L個のインスタンスを受信する。B R I R フィルタ126A、126Bは、すでに、静止サンプルを除去するために調整されていることがある。B R I R 調整ユニット106は、それぞれのH R T Fと、早期反射と、残余室内セグメントとを識別するためにB R I R フィルタ126A、126Bをセグメント化するために、上記の技法を適用し得る。B R I R 調整ユニット106は、B R I R SHC - 領域変換ユニット112にH R T Fと早期反射セグメントとを、サイズ[a, L]の左および右の行列を表す行列129A、129Bとして与え、ここで、aはH R T Fと早期反射セグメントとの連結の長さであり、Lは（仮想または実在の）ラウドスピーカーの数である。B R I R 調整ユニット106は、残余室内応答ユニット110にB R I R フィルタ126A、126Bの残余室内セグメントを、サイズ[b, L]の左

および右の残余室内外行列 $1\ 2\ 8\ A$ 、 $1\ 2\ 8\ B$ として与え、ここで、 b は残余室内外セグメントの長さであり、 L は（仮想または実在の）ラウドスピーカーの数である。

【0057】

[0072] 残余室内外応答ユニット 110 は、SHC122 によって図 7 に表すように、音場を記述する階層要素（たとえば、球面調和係数）の少なくとも幾分かの部分との畳み込みのために、左および右の共通の残余室内外応答セグメントを計算またはさもなければ決定するために、上記の技法を適用し得る。すなわち、残余室内外応答ユニット 110 は、左および右の残余室内外行列 $1\ 2\ 8\ A$ 、 $1\ 2\ 8\ B$ を受信し、左および右の共通の残余室内外応答セグメントを生成するために左および右それぞれの残余室内外行列 $1\ 2\ 8\ A$ 、 $1\ 2\ 8\ B$ を L 個にわたって結合することができる。いくつかの例では、残余室内外応答ユニット 110 は、左および右の残余室内外行列 $1\ 2\ 8\ A$ 、 $1\ 2\ 8\ B$ を L 個にわたって平均化することによって結合を実施し得る。10

【0058】

[0073] 次いで、残余室内外応答ユニット 110 は、左および右の共通の残余室内外応答セグメントと、チャンネル 124B として図 7 に示す SHC122 の少なくとも 1 つのチャンネルとの高速畳み込みを計算し得る。いくつかの例では、左および右の共通の残余室内外応答セグメントは周囲を取り巻く無指向性の音を表すので、チャンネル 124B は、SHC122 の W チャンネル（すなわち、0 次）であり、それは、音場の無指向性部分を符号化する。そのような例では、長さ L_{ength} の W チャンネルサンプルに関して、残余室内外応答ユニット 110 による左および右の共通の残余室内外応答セグメントとの高速畳み込みは、長さ L_{ength} の左および右の出力信号 $1\ 3\ 4\ A$ 、 $1\ 3\ 4\ B$ を生成する。20

【0059】

[0074] 本明細書で使用する「高速畳み込み」および「畳み込み」という用語は、時間領域における畳み込み演算、ならびに周波数領域における点毎の（point-wise）乗算演算を指すことがある。言い換えれば、信号処理の当業者によく知られているように、時間領域における畳み込みは、周波数領域における点毎の乗算と等価であり、ここで時間領域および周波数領域は、互いの変換である。出力変換は、入力変換と伝達関数との点毎の積である。したがって、畳み込みおよび点毎の乗算（または単に「乗算」）は、それぞれの領域（ここでは時間および周波数）に関して行われる概念的に同様の演算を指すことができる。畳み込みユニット 114、214、230；残余室内外応答ユニット 210、354；フィルタ 384 および残響 386 は、代替として、周波数領域における乗算を適用し得、ここでこれらの成分への入力は、時間領域ではなく周波数領域において与えられる。「高速畳み込み」または「畳み込み」として本明細書で説明する他の演算は、同様に、周波数領域における乗算と呼ばれることもあり、ここで、これらの演算への入力は、時間領域ではなく周波数領域で与えられる。30

【0060】

[0075] いくつかの例では、残余室内外応答ユニット 110 は、共通の残余室内外応答セグメントの開始時間に関する値を B R I R 調整ユニット 106 から受信し得る。残余室内外応答ユニット 110 は、B R I R フィルタ 108 に関するより早いセグメントとの結合を見越して、出力信号 $1\ 3\ 4\ A$ 、 $1\ 3\ 4\ B$ をゼロパディングするかまたはさもなければ遅延させ得る。40

【0061】

[0076] B R I R SHC - 領域変換ユニット 112（以後、「領域変換ユニット 112」）は、左および右の B R I R フィルタ $1\ 2\ 6\ A$ 、 $1\ 2\ 6\ B$ を球面調和領域に潜在的に変換し、次いでそのフィルタを L 個にわたって潜在的に合計するために、SHC レンダリング行列を B R I R 行列に適用する。領域変換ユニット 112 は、変換結果を、それぞれ、左および右の SHC - バイノーラルレンダリング行列 $1\ 3\ 0\ A$ 、 $1\ 3\ 0\ B$ として出力する。行列 $1\ 2\ 9\ A$ 、 $1\ 2\ 9\ B$ が $[a, L]$ のサイズである場合、SHC - バイノーラルレンダリング行列 $1\ 3\ 0\ A$ 、 $1\ 3\ 0\ B$ の各々は、フィルタを L 個にわたって合計した後、 $[(N+1)^2, a]$ のサイズになる（たとえば、式（4）～（5）参照）。いくつかの例で50

は、S H C - バイノーラルレンダリング行列 1 3 0 A、1 3 0 B は、実行時または準備時間において計算されるのではなく、音声再生デバイス 1 0 0 の中で構成される。いくつかの例では、S H C - バイノーラルレンダリング行列 1 3 0 A、1 3 0 B の複数のインスタンスは、音声再生デバイス 1 0 0 の中で構成され、音声再生デバイス 1 0 0 は、S H C 1 2 4 A に適用するために、左および右一対の複数のインスタンスを選択する。

【 0 0 6 2 】

[0077] 署み込みユニット 1 1 4 は、左および右のバイノーラルレンダリング行列 1 3 0 A、1 3 0 B と S H C 1 2 4 A とを署み込み、S H C 1 2 4 A は、いくつかの例では、S H C 1 2 2 の次数から次数を削減することができる。周波数（たとえば、S H C ）領域における S H C 1 2 4 A に関して、署み込みユニット 1 1 4 は、S H C 1 2 4 A と左および右のバイノーラルレンダリング行列 1 3 0 A、1 3 0 B とのそれぞれの点毎の乗算を計算し得る。長さ $Length$ の S H C 信号に関して、署み込みは、[$Length$, ($N + 1$)²] のサイズの左および右のフィルタリングされた S H C チャンネル 1 3 2 A、1 3 2 B をもたらし、一般的に、球面調和領域の次数 / 副次数の結合の各々に関して各出力信号行列に関する行が存在する。

【 0 0 6 3 】

[0078] 結合ユニット 1 1 6 は、バイノーラル出力信号 1 3 6 A、1 3 6 B を作成するために、左および右のフィルタリングされた S H C チャンネル 1 3 2 A、1 3 2 B と出力信号 1 3 4 A、1 3 4 B とを結合することができる。次いで、結合ユニット 1 1 6 は、バイノーラル出力信号 1 3 6 A、1 3 6 B を生成するために左および右のバイノーラル出力信号と左および右の出力信号 1 3 4 A、1 3 4 B とを結合する前に、H R T F に関する左および右のバイノーラル出力信号と早期エコー（反射）セグメントとを生成するために、左および右のフィルタリングされた S H C チャンネル 1 3 2 A、1 3 2 B の各々を L 個にわたって別々に合計することができる。

【 0 0 6 4 】

[0079] 図 8 は、本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイスの一例を示すブロック図である。音声再生デバイス 2 0 0 は、音声再生デバイスの例示的な例を表し得、図 7 の 1 0 0 はさらなる詳細である。

【 0 0 6 5 】

[0080] 音声再生デバイス 2 0 0 は、S H C 2 4 2 の次数を削減するために、ビットストリーム 2 4 0 から入ってくる S H C 2 4 2 を処理する随意の S H C 次数削減ユニット 2 0 4 を含み得る。随意の S H C 次数削減は、S H C 2 4 2（たとえば、W チャンネル）の最高次数（たとえば、0 次）のチャンネル 2 6 2 を残余室内応答ユニット 2 1 0 に与え、削減された次数の S H C 2 4 2 を署み込みユニット 2 3 0 に与える。S H C 次数削減ユニット 2 0 4 が S H C 2 4 2 の次数を削減しない例では、署み込みユニット 2 3 0 は、S H C 2 4 2 と同等の S H C 2 7 2 を受信する。いずれにせよ、S H C 2 7 2 は、[$Length$, ($N + 1$)²] の次元を有し、ここで N は S H C 2 7 2 の次数である。

【 0 0 6 6 】

[0081] B R I R 調整ユニット 2 0 6 および B R I R フィルタ 2 0 8 は、図 7 の B R I R 調整ユニット 1 0 6 および B R I R フィルタ 1 0 8 の例示的な例を表し得る。残余応答ユニット 2 1 4 の署み込みユニット 2 1 4 は、上記で説明した技法を使用して B R I R 調整ユニット 2 0 6 によって調整された共通の左および右の残余室内セグメント 2 4 4 A、2 4 4 B を受信し、署み込みユニット 2 1 4 は、左および右の残余室内信号 2 6 2 A、2 6 2 B を生成するために共通の左および右の残余室内セグメント 2 4 4 A、2 4 4 B と最高次数のチャンネル 2 6 2 とを署み込む。遅延ユニット 2 1 6 は、左および右の残余室内出力信号 2 6 8 A、2 6 8 B を生成するために、共通の左および右の残余室内セグメント 2 4 4 A、2 4 4 B に対するサンプルの開始の数で左および右の残余室内信号 2 6 2 A、2 6 2 B をゼロパディングすることができる。

【 0 0 6 7 】

[0082] B R I R S H C - 領域変換ユニット 2 2 0（以後、領域変換ユニット 2 2 0 ）

10

20

30

40

50

は、図7の領域変換ユニット112の例示的な例を表し得る。図示の例では、変換ユニット222は、[a, L]のサイズの左および右の行列を表す行列248A、248Bに($N+1$)²次元のSHCレンダリング行列224を適用し、ここでaはHRTFと早期反射セグメントとの連結の長さであり、Lはラウドスピーカー（たとえば、仮想のラウドスピーカー）の数である。変換ユニット222は、次元[$(N+1)^2$, a, L]を有するSHC-領域における左および右の行列252A、252Bを出力する。合計ユニット226は、次元[$(N+1)^2$, a]を有する左および右の中間SHC-レンダリング行列254A、254Bを作成するために、左および右の行列252A、252Bの各々をL個にわたって合計し得る。削減ユニット228は、最小位相低減など、SHC272にSHC-レンダリング行列を適用すること、および最小位相低減を適用されている中間SHC-レンダリング行列254A、254Bのそれぞれの最小位相部分の周波数応答を近似するようにIIRフィルタを設計するために平衡型モデル打切り法(Balanced Model Truncation method)を使用すること、についての計算の複雑さをさらに削減するために、上記で説明した技法を適用し得る。削減ユニット228は、左および右のSHC-レンダリング行列256A、256Bを出力する。

【0068】

[0083]畳み込みユニット230は、中間信号258A、258Bを生成するためにSHC272の形態のSHCコンテンツをフィルタリングし、合計ユニット232は、左および右の信号260A、260Bを作成するために中間信号258A、258Bを合計する。結合ユニット234は、左および右のバイノーラル出力信号270A、270Bを生成するために左および右の残余室内出力信号268A、268Bと左および右の信号260A、260Bとを結合する。

【0069】

[0084]いくつかの例では、バイノーラルレンダリングユニット202は、変換ユニット222によって生成されたSHC-バイノーラルレンダリング行列252A、252Bのうちの1つだけを使用することによって計算のさらなる削減を実施し得る。その結果、畳み込みユニット230は、左または右の信号の一方だけについて演算し、畳み込み演算を半分に削減することができる。そのような例では、合計ユニット232は、出力260A、260Bをレンダリングするときに、第2のチャンネルに関する条件付き決定を行う。

【0070】

[0085]図9は、本開示で説明する技法による球面調和係数をレンダリングするための、バイノーラルレンダリングデバイスに関する例示的な演算のモードを示すフローチャートである。例示のために、例示的な演算のモードについて、図7の音声再生デバイス200に関して説明する。バイノーラル室内インパルス応答(BRIR)調整ユニット206は、BRIRフィルタ246A、246Bから方向依存性成分/セグメント、特に頭部伝達関数および早期エコーセグメントを抽出することによって、左および右それぞれのBRIRフィルタ246A、246Bを調整する(300)。左および右のBRIRフィルタ126A、126Bの各々は、1つまたは複数の対応するラウドスピーカーに関するBRIRフィルタを含み得る。BRIR調整ユニット106は、抽出された頭部伝達関数と早期エコーセグメントとの連結を、左および右の行列248A、248BとしてBRIRSHC-領域変換ユニット220に与える。

【0071】

[0086]BRIR SHC-領域変換ユニット220は、球面調和(たとえば、HOA)領域内の左および右のフィルタ行列252A、252Bを生成するために、抽出された頭部伝達関数と早期エコーセグメントとを含む左および右のフィルタ行列248A、248Bを変換するためにHOAレンダリング行列224を適用する(302)。いくつかの例では、音声再生デバイス200は、左および右のフィルタ行列252A、252Bを用いて構成され得る。いくつかの例では、音声再生デバイス200は、ビットストリーム240の帯域外または帯域内の信号においてBRIRフィルタ208を受信し、その場合、音声再生デバイス200は、左および右のフィルタ行列252A、252Bを生成する。合

10

20

30

40

50

計ユニット226は、左および右の中間S H C - レンダリング行列254A、254Bを含むS H C領域内のバイノーラルレンダリング行列を生成するために、それぞれの左および右のフィルタ行列252A、252Bをラウドスピーカーの次元にわたって合計する(304)。削減ユニット228は、左および右のS H C - レンダリング行列256A、256Bを生成するために、中間S H C - レンダリング行列254A、254Bをさらに削減し得る。

【0072】

[0087]バイノーラルレンダリングユニット202の畳み込みユニット230は、左および右のフィルタリングされたS H C(たとえば、H O A)チャンネル258A、258Bを作成するために、左および右の中間S H C - レンダリング行列256A、256BをS H Cコンテンツ(球面調和係数272など)に適用する(306)。

10

【0073】

[0088]合計ユニット232は、方向依存性セグメントに関する左および右の信号260A、260Bを作成するために、左および右のフィルタリングされたS H Cチャンネル258A、258Bの各々をS H C次元($N + 1$)²にわたって合計する(308)。次いで、結合ユニット116は、左および右のバイノーラル出力信号270A、270Bを含むバイノーラル出力信号を生成するために、左および右の信号260A、260Bと左および右の残余室内出力信号268A、268Bとを結合し得る。

【0074】

[0089]図10Aは、本開示で説明する技法の様々な態様による、図7および図8の音声再生デバイスによって実施され得る例示的な演算のモード310を示す図である。演算のモード310は、図8の音声再生デバイス200に関して、後で本明細書で説明される。音声再生デバイス200のバイノーラルレンダリングユニット202は、B R I Rデータ312、これはB R I Rフィルタ208の例示的な例であり得ると、H O Aレンダリング行列314、これはH O Aレンダリング行列224の例示的な例であり得る、とを用いて構成され得る。音声再生デバイス200は、帯域内または帯域外のシグナリングチャンネル内のB R I Rデータ312とH O Aレンダリング行列314とをビットストリーム240と相対して受信し得る。この例におけるB R I Rデータ312は、たとえば、L個の実在または仮想のラウドスピーカーを表すL個のフィルタを有し、L個のフィルタの各々は長さKである。L個のフィルタの各々は、左および右の成分を含み得る('x2')。いくつかの場合には、L個のフィルタの各々は、左または右に関する単一の成分を含むことがあり、その成分は、右または左のその相手の成分と対称である。これは、高速畳み込みのコストを削減し得る。

20

【0075】

[0090]音声再生デバイス200のB R I R調整ユニット206は、セグメント化演算と結合演算とを適用することによってB R I Rデータ312を調整し得る。具体的には、例示的な演算のモード310において、B R I R調整ユニット206は、本明細書で説明する技法によるL個のフィルタの各々を、行列315(次元[a, 2, L])を作成するための結合の長さaのH R T Fプラス早期エコーフィルタと、残余行列339(次元[b, 2, L])を作成するための残余室内応答セグメントとにセグメント化する(324)。B R I Rデータ312のL個のフィルタの長さKは、ほぼ、aとbとの合計である。変換ユニット222は、次元[(N + 1)², a, 2, L]の行列317(これは左および右の行列252A、252Bの結合の例示的な例であり得る)を作成するために、(N + 1)²次元のH O A / S H Cレンダリング行列314を作成するためのL個のフィルタに適用し得る。合計ユニット226は、次元[(N + 1)², a, 2]を有する中間S H C - レンダリング行列335を作成するために、左および右の行列252A、252Bの各々をL個にわたって合計し得る(値2を有する第3の次元は左および右の成分を表し、中間S H C - レンダリング行列335は、左および右の両方の中間S H C - レンダリング行列254A、254Bの例示的な例として表すことができる)(326)。いくつかの例では、音声再生デバイス200は、H O Aコンテンツ316(またはその削減されたバ-

30

40

50

ジョン、たとえばHOAコンテンツ321)に適用するための中間SHC-レンダリング行列335を用いて構成され得る。いくつかの例では、削減ユニット228は、行列317の左または右の成分の一方だけを使用することによって、さらなる削減を計算に適用し得る(328)。

【0076】

[0091]音声再生デバイス200は、次数N₁および長さLengthのHOAコンテンツ316を受信し、いくつかの態様では、その中の球面調和係数(SHC)の次数をNに削減するために次数削減演算を適用する(330)。N₁は、入力((I)npot)HOAコンテンツ321の次数を示す。次数削減演算(330)のHOAコンテンツ321は、HOAコンテンツ316と同様に、SHC領域内にある。随意の次数削減演算はまた、最高次数(たとえば、0次)の信号319を生成し、高速畳み込み演算のために残余応答ユニット210に与える(338)。HOA次数削減ユニット204がHOAコンテンツ316の次数を削減しない例では、高速畳み込み適用演算(apply fast convolution operation)(332)は、削減された次数を持たない入力に対して演算する。いずれにしても、高速畳み込み演算(332)に入力されるHOAコンテンツ321は、次元[Length, (N+1)²]を有し、ここでNは次数である。

【0077】

[0092]音声再生デバイス200は、左および右の成分、したがって次元[Length, (N+1)², 2]を有するHOA信号323を作成するために、HOAコンテンツ321と行列335との高速畳み込みを適用し得る(332)。ここでも、高速畳み込みは、周波数領域におけるHOAコンテンツ321と行列335との点毎の乗算、または時間領域における畳み込みを指すことができる。音声再生デバイス200は、次元[Length, 2]を有する合計された信号325を作成するために、HOA信号323を(N+1)²にわたってさらに合計することができる(334)。

【0078】

[0093]次に、残余行列339に戻ると、音声再生デバイス200は、次元「b, 2」を有する共通の残余室内応答行列327を生成するために、本明細書で説明する技法に従つてL個の残余室内応答セグメントを結合することができる(336)。音声再生デバイス200は、次元[Length, 2]を有する室内応答信号329を作成するために、0次のHOA信号319と共に残余室内応答行列327との高速畳み込みを適用し得る(338)。残余行列339のL個の残余応答室内応答セグメントを生成するために、音声再生デバイス200は、BIRデータ312のL個のフィルタのうちの(a+1)番目のサンプルにおいて開始する残余応答室内応答セグメントを取得したので、音声再生デバイス200は、次元[Length, 2]を有する室内応答信号311を生成するためにa個のサンプルを遅延(たとえば、パディング)することによって初期のa個のサンプルを構成する(account for)(340)。

【0079】

[0094]音声再生デバイス200は、次元[Length, 2]を有する出力信号318を作成するために、合計された信号325と室内応答信号311とを、要素を加算することによって結合する(342)。このようにして、音声再生デバイスは、L個の残余室内応答セグメントの各々に関して高速畳み込みを適用することを回避し得る。バイノーラル音声出力信号に変換するために入力される22チャンネルに関して、これは、残余室内応答を生成するための高速畳み込みの数を、22から2に削減し得る。

【0080】

[0095]図10Bは、本開示で説明する技法の様々な態様による、図7および図8の音声再生デバイスによって実施され得る例示的な演算のモード350を示す図である。演算のモード350は、図8の音声再生デバイス200に関して、後で本明細書で説明され、演算のモード310と同様である。しかしながら、演算のモード350は、最初に、HOAコンテンツを、L個の実在または仮想のラウドスピーカーに関して時間領域内のマルチチャンネルスピーカー信号にレンダリングすることと、次いで、本明細書で説明する技法に

10

20

30

40

50

従ってスピーカーフィードの各々に効率的な B R I R フィルタリングを適用することと、を含む。そのために、音声再生デバイス 200 は、H O A コンテンツ 321 を、次元 [L e n g t h , L] を有するマルチチャンネル音声信号 333 に変換する (344)。加えて、音声再生デバイスは、B R I R データ 312 を S H C 領域に変換しない。したがって、音声再生デバイス 200 による削減を信号 314 に適用することは、次元 [a , 2 , L] を有する行列 337 を生成する (328)。

【 0 0 8 1 】

[0096] 次いで、音声再生デバイス 200 は、次元 [L e n g t h , L , 2] (左および右の成分を有する) を有するマルチチャンネル音声信号 341 を作成するために、マルチチャンネル音声信号 333 と行列 337 との高速畳み込み 332 を適用する (348)。次いで、音声再生デバイス 200 は、次元 [L e n g t h , 2] を有する信号 325 を作成するために、L 個のチャンネル / スピーカーによるマルチチャンネル音声信号 341 を合計し得る (346)。

【 0 0 8 2 】

[0097] 図 11 は、本開示で説明するバイノーラル音声レンダリング技法の様々な態様を実施し得る音声再生デバイス 350 の一例を示すプロック図である。単一のデバイス、すなわち図 11 の例における音声再生デバイス 350 として示されているが、本技法は、1 つまたは複数のデバイスによって実施されてよい。したがって、本技法はこの点において限定されるべきではない。

【 0 0 8 3 】

[0098] その上、概して、図 1 ~ 図 10 B の例に関して球面調和領域において適用されるとして上記で説明されているが、本技法はまた、5.1 サラウンドサウンドフォーマット、7.1 サラウンドサウンドフォーマット、および / または 22.2 サラウンドサウンドフォーマットなど、上記のサラウンドサウンドフォーマットに適合するチャンネルベースの信号を含む、任意の形態の音声信号に関して実施され得る。したがって、本技法はまた、球面調和領域内で指定された音声信号に限定されるべきではなく、任意の形態の音声信号に対して適用され得る。

【 0 0 8 4 】

[0099] 図 11 の例に示すように、音声再生デバイス 350 は、図 7 の例に示す音声再生デバイス 100 に類似し得る。しかしながら、音声再生デバイス 350 は、一例として 22.2 サラウンドサウンドフォーマットに適合する一般的なチャンネルベースの音声信号に関する技法を演算またはさもなければ実施することができる。抽出ユニット 104 は、音声チャンネル 352 を抽出し得、ここで音声チャンネル 352 は、一般に「n」チャンネルを含み得、この例では、22.2 サラウンドサウンドフォーマットに適合する 22 チャンネルを含むものと仮定される。これらのチャンネル 352 は、バイノーラルレンダリングユニット 351 の残余室内応答ユニット 354 とチャンネルごとの打切りフィルタユニット 356 の両方に与えられる。

【 0 0 8 5 】

[0100] 上記で説明したように、B R I R フィルタ 108 は、1 つまたは複数の B R I R フィルタを含み、図 3 の B R I R フィルタ 37 の一例を表し得る。B R I R フィルタ 108 は、左および右の H R T F がそれぞれの B R I R に与える影響を表す、個別の B R I R フィルタ 126 A、126 B を含み得る。

【 0 0 8 6 】

[0101] B R I R 調整ユニット 106 は、B R I R フィルタ 126 A、126 B の n 個のインスタンスを受信し、各チャンネル n それぞれに関して、各 B R I R フィルタは長さ N を有する。B R I R フィルタ 126 A、126 B は、すでに、静止サンプルを除去するために調整されていることがある。B R I R 調整ユニット 106 は、それぞれの H R T F と、早期反射と、残余室内セグメントとを識別するために B R I R フィルタ 126 A、126 B をセグメント化するために、上記で説明した技法を適用し得る。B R I R 調整ユニット 106 は、チャンネルごとの打切りフィルタユニット 356 に H R T F と早期反射セグ

10

20

30

40

50

メントとを、サイズ [a , L] の左および右の行列を表す行列 1 2 9 A 、 1 2 9 B として与え、ここで、 a は H R T F と早期反射セグメントとの連結の長さであり、 n は（仮想または実在の）ラウドスピーカーの数である。 B R I R 調整ユニット 1 0 6 は、残余室内応答ユニット 3 5 4 に B R I R フィルタ 1 2 6 A 、 1 2 6 B の残余室内セグメントを、サイズ [b , L] の左および右の残余室内行列 1 2 8 A 、 1 2 8 B として与え、ここで、 b は残余室内セグメントの長さであり、 n は（仮想または実在の）ラウドスピーカーの数である。

【 0 0 8 7 】

[0102] 残余室内応答ユニット 3 5 4 は、音声チャンネル 3 5 2 との畳み込みのための左および右の共通の残余室内応答セグメントを計算またはさもなければ決定するために、上記で説明する技法を適用し得る。すなわち、残余室内応答ユニット 1 1 0 は、左および右の残余室内行列 1 2 8 A 、 1 2 8 B を受信し、左および右の共通の残余室内応答セグメントを生成するために左および右それぞれの残余室内行列 1 2 8 A 、 1 2 8 B を n 個にわたって結合することができる。いくつかの例では、残余室内応答ユニット 3 5 4 は、左および右の残余室内行列 1 2 8 A 、 1 2 8 B を n 個にわたって平均化することによって結合を実施し得る。

【 0 0 8 8 】

[0103] 次いで、残余室内応答ユニット 3 5 4 は、左および右の共通の残余室内応答セグメントと、音声チャンネル 3 5 2 のうちの少なくとも 1 つのチャンネルとの高速畳み込みを計算し得る。いくつかの例では、残余室内応答ユニット 3 5 2 は、共通の残余室内応答セグメントの開始時間に関する値を B R I R 調整ユニット 1 0 6 から受信し得る。残余室内応答ユニット 3 5 4 は、 B R I R フィルタ 1 0 8 に関する、より早いセグメントとの結合を見越して、出力信号 1 3 4 A 、 1 3 4 B をゼロパディングするかまたはさもなければ遅延させ得る。出力信号 1 3 4 A は左音声信号を表す一方で、出力信号 1 3 4 B は右音声信号を表すことができる。

【 0 0 8 9 】

[0104] チャンネルごとの打切りフィルタユニット 3 5 6 （以後、「打切りフィルタユニット 3 5 6 」）は、 H R T F と B R I R フィルタの早期反射セグメントとをチャンネル 3 5 2 に適用し得る。より具体的には、チャンネルごとの打切りフィルタユニット 3 5 6 は、 H R T F と B R I R フィルタの早期反射セグメントとを表す行列 1 2 9 A 、 1 2 9 B をチャンネル 3 5 2 のそれぞれのチャンネルに適用し得る。いくつかの例では、行列 1 2 9 A 、 1 2 9 B は、単一の行列 1 2 9 を形成するように結合され得る。その上、一般的に、 H R T F ならびに早期反射行列 1 2 9 A および 1 2 9 B の各々のうちの左の 1 つと、 H R T F ならびに早期反射行列 1 2 9 A および 1 2 9 B の各々のうちの右の 1 つとが存在する。すなわち、一般的に、左耳および右耳に関する H R T F と早期反射行列とが存在する。チャンネルごとの方向ユニット 3 5 6 は、左および右のフィルタリングされたチャンネル 3 5 8 A および 3 5 8 B を出力するために、左および右の行列 1 2 9 A 、 1 2 9 B の各々を適用し得る。結合ユニット 1 1 6 は、バイノーラル出力信号 1 3 6 A 、 1 3 6 B を作成するために、左のフィルタリングされたチャンネル 3 5 8 A と出力信号 1 3 4 A とを結合する（または、言い換えればミックスする）一方で、右のフィルタリングされたチャンネル 3 5 8 B と出力信号 1 3 4 B とを結合する（または、言い換えればミックスする）ことができる。バイノーラル出力信号 1 3 6 A は左の音声チャンネルに対応し、バイノーラル出力信号 1 3 6 B は右の音声チャンネルに対応することができる。

【 0 0 9 0 】

[0105] いくつかの例では、バイノーラルレンダリングユニット 3 5 1 は、残余室内応答ユニット 3 5 4 が、チャンネルごとの打切りフィルタユニット 3 5 6 の演算と同時に演算するように、残余室内応答ユニット 3 5 4 とチャンネルごとの打切りフィルタユニット 3 5 6 とを互いに同時に起動し得る。すなわち、いくつかの例では、残余室内応答ユニット 3 5 4 は、バイノーラル出力信号 1 3 6 A 、 1 3 6 B が生成され得る速度を改善するために、チャンネルごとの打切りフィルタユニット 3 5 6 と並列に（しかし、同時にないこと

10

20

30

40

50

が多い)演算することが多い。潜在的にカスケード接続方式で演算するよう様々な上記の図において示しているが、本技法は、別段に具体的に規定されていない限り、本開示で説明する説明するユニットまたはモジュールのいずれもの同時演算または並列演算を提供し得る。

【0091】

[0106]図12は、本開示で説明する技法の様々な態様による、図11の音声再生デバイス350によって実施され得るプロセス380を示す図である。プロセス380は、各B R I Rを2つの部分:(a)左フィルタ384A_L~384N_Lおよび右フィルタ384A_R~384N_R(総称して「フィルタ384」)によって表されるH R T Fおよび早期反射の効果を組み込む、より小さい構成要素、および(b)元のB R I Rのすべての末尾の特性から生成され、左残響フィルタ386Lおよび右残響フィルタ386R(総称して「共通のフィルタ386」)によって表される共通の「残響の末尾」に分解することを達成する。プロセス380に示すチャンネルごとのフィルタ384は、上記の部分(a)を表す一方で、プロセス380に示す共通のフィルタ386は、上記の部分(b)を表すことができる。

【0092】

[0107]プロセス380は、不可聴成分を除去し、H R T F / 早期反射を備える成分と後期反射 / 拡散による成分とを決定するためにB R I Rを解析することによってこの分解を実施する。これは、部分(a)に関する、一例として2704タップ(tap)の長さのF I Rフィルタと、部分(b)に関する、別の例として15232タップの長さのF I Rフィルタとをもたらす。プロセス380によれば、音声再生デバイス350は、より短いF I Rフィルタだけを、個別のnチャンネルの各々に適用し得、nは、演算396において例示のために22であると仮定されている。この演算の複雑性は、以下で再生される式(8)における第1の部分の計算(4096点のF F Tを使用する)において表され得る。プロセス380では、音声再生デバイス350は、共通の「残響の末尾」を、22チャンネルの各々にではなく、演算398においてそれらすべての加法的なミックスに適用し得る。この複雑性は、式(8)における複雑性の計算の第2の半分において表され、繰り返すが、これは、添付したアヘンディクス(Appendix)において示される。

【0093】

[0108]この点において、プロセス380は、複数のNチャンネルからの音声コンテンツをミックスすることに基づいて、合成音声信号を生成するバイノーラル音声レンダリングの方法を表し得る。加えて、プロセス380は、さらに、合成音声信号を、遅延によってNチャンネルフィルタの出力と整列させ得、各チャンネルフィルタは、打切りB R I Rフィルタを含む。その上、プロセス380では、音声再生デバイス350は、次いで、演算398において共通の合成残余室内インパルス応答を用いて整列合成音声信号をフィルタリングし、バイノーラル音声出力の左成分388Lおよび右成分388Rのために、演算390Lおよび390Rにおいて、各チャンネルフィルタの出力とフィルタリングされた整列合成音声信号とをミックスすることができる。

【0094】

[0109]いくつかの例では、打切りB R I Rフィルタおよび共通の合成残余インパルス応答は、メモリにプリロードされる。

【0095】

[0110]いくつかの例では、整列合成音声信号のフィルタリングは、時間周波数領域内で実施される。

【0096】

[0111]いくつかの例では、整列合成音声信号のフィルタリングは、畳み込みを介して時間領域内で実施される。

【0097】

[0112]いくつかの例では、打切りB R I Rフィルタおよび共通の合成残余インパルス応答は、分解分析法に基づく。

10

20

30

40

50

【0098】

[0113]いくつかの例では、分解分析法は、N個の室内インパルス応答の各々に対して実施され、N個の打切り室内インパルス応答とN個の残余インパルス応答とをもたらす（ここでNは、nまたはn超として示されることがある）。

【0099】

[0114]いくつかの例では、打切りインパルス応答は、各室内インパルス応答の全長さの40パーセント未満を表す。

【0100】

[0115]いくつかの例では、打切りインパルス応答は、111と17, 830との間のタップ範囲を含む。

10

【0101】

[0116]いくつかの例では、N個の残余インパルス応答の各々は、複雑性を削減する共通の合成残余室内応答内に結合される。

【0102】

[0117]いくつかの例では、各チャンネルフィルタの出力と、フィルタリングされた整列合成音声信号とをミックスすることは、左のスピーカー出力に関するミキシングの第1のセットと右のスピーカー出力に関するミキシングの第2のセットとを含む。

20

【0103】

[0118]様々な例では、上記で説明したプロセス380の様々な例またはそれらの任意の結合の方法は、メモリおよび1つまたは複数のプロセッサを備えるデバイスと、本方法の各ステップを実施するための手段を備えた装置と、非一時的コンピュータ可読記憶媒体上に記憶された命令を実行することによって本方法の各ステップを実施する1つまたは複数のプロセッサとによって実施され得る。

【0104】

[0119]その上、上記で説明した例のいずれにも記載される特定の特徴のいずれも、説明した技法の有益な例の中に組み合わされ得る。すなわち、特定の特徴のいずれも、一般に、本技法のすべての例に適用可能である。本技法の様々な例について説明した。

【0105】

[0120]本開示で説明した技法は、ある場合には、可聴のB R I Rセットにわたってサンプル111～17830だけを識別することができる。例示的な室内の容積からミキシング時間 T_{mp95} を計算し、本技法は、次いで、53.6msの後、すべてのB R I Rに共通の残響の末尾を共有させることができ、15232のサンプル長の共通の残響の末尾と、残留する2704サンプルのH R T F + 反射インパルスとをもたらし、3msのクロスフェードがそれらの間に存在する。計算コスト削減（break down）に関して、以下の項目が到達され得る。

30

【0106】

(a) 共通の残響の末尾： $10 \times 6 \times 10 \log_2 (2 \times 15232 / 10)$ 。

【0107】

(b) 残留するインパルス： $22 \times 6 \times 10 \log_2 (2 \times 4096)$ 、1フレーム内でそれを行うために4096のFFTを使用する。

40

【0108】

(c) 追加の22の加算。

【0109】

[0121]その結果、最終の性能指数は、したがって、ほぼ $C_{mod} = \max (100 \times (C_{conv} - C) / C_{conv}, 0) = 88.0$ に等しく、ここで

【数22】

$$C_{mod} = \max(100 * (C_{conv} - C)/C_{conv}, 0) \quad (6)$$

ここで C_{conv} は最適化されていない実装 (implementation) :

【数23】

10

$$C_{conv} = (22+2)*(10)*(6*\log_2(2*48000/10)) \quad (7)$$

の推定であり、

C は何らかの態様であり、2つの付加的な要素 :

【数24】

20

$$C = 22 * 6 * \log_2(2 * 4096) + 10 * 6 * \log_2(2 * \frac{15232}{10}) \quad (8)$$

によって決定され得る。

【0110】

[0122] したがって、いくつかの態様では、性能指数は、 $C_{mod} = 87.35$ 。

【0111】

[0123] $B_n(z)$ として示される B R I R フィルタは、2つの関数 $B T_n(z)$ と $B R_n(z)$ とに分解され得、それらはそれぞれ、打切り B R I R フィルタと残響 B R I R フィルタとを示す。上記の部分 (a) はこの打切り B R I R フィルタを指す一方で、上記の部分 (b) は残響 B R I R フィルタを指し得る。次いで、 $B_n(z)$ は $B T_n(z) + (z^{-m} * B R_n(z))$ に等しくし得、ここで m は遅延を示す。したがって、出力信号 $Y(z)$ は、

30

【数25】

$$\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} \cdot X_n(z) * BR_n(z)] \quad (9)$$

として計算され得る。

40

【0112】

[0124] プロセス 380 は、共通の合成の残響の末尾のセグメントを導出するために $B R_n(z)$ を解析し得、ここでこの共通の $B R(z)$ は、チャンネル固有の $B R_n(z)$ の代わりに適用され得る。この共通の（またはチャンネル全般の）合成 $B R(z)$ が使用されるとき、 $Y(z)$ は、

【数26】

$$\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} BR_n(z)] \cdot \sum_{n=0}^{N-1} X_n(z) \quad (10)$$

として計算され得る。

【0113】

[0125]図13は、例示的なバイノーラル室内インパルス応答フィルタ(BRIR)400の図である。BRIR400は、5つのセグメント402A～402Cを示す。頭部伝達関数(HRTF)セグメント402Aは、頭部伝達によるインパルス応答を含み、本明細書で説明する技法を使用して識別され得る。HRTFは、無響室内でインパルス応答を測定することと等価である。部屋の第1の反射は、通常、HRTFより長い遅延を有するので、BRIRの最初の部分がHRTFインパルス応答であると仮定される。反射セグメント402Bは、HRTFに室内効果を結合する、すなわち、反射セグメント402Bのインパルス応答は、残響セグメント402Cと比較において早期の離散的なエコーによってフィルタリングされたBRIR400に関するHRTFセグメント402Aのインパルス応答と一致する。ミキシング時間は、反射セグメント402Bと残響セグメント402Cとの間の時間であり、早期エコーが密な残響になる時間を示す。残響セグメント402Cは、ガウス雑音のように拳動し、離散的なエコーは、それ以上分離されない。

【0114】

[0126]近く発表されるMPEG-H規格化において、高解像度と高チャンネルカウントとを有するマルチチャンネル音声が考慮されている。レンダリングをポータブルするために、ヘッドフォン表現が必要である。これは、すべてのスピーカーフィード/チャンネルを1つのステレオヘッドセット内に仮想化することを伴う。ヘッドフォン表現へレンダリングするために、インパルス応答のうちの1つまたは複数のペアのセットが、マルチチャンネル音声に適用され得る。BRIR400は、そのようなインパルス応答のうちの1つのペアを表し得る。標準的なブロック高速フーリエ変換(FFT)を使用してBRIR400フィルタをマルチチャンネル音声のチャンネルに適用することは、計算の負荷が高くなり得る。インパルス応答のペアのセット全体をマルチチャンネル音声の対応するチャンネルに適用することは、いっそう計算の負荷を高くなる。以下で説明する技法は、標準的なフィルタリング(たとえば、ブロックFFT)の結果の品質を著しく犠牲にすることなく、効率的なバイノーラルフィルタリングを提供する。

【0115】

[0127]図14は、バイノーラル室内インパルス応答をマルチチャンネル音声信号に適用することによって生成されるバイノーラル出力信号の計算のためのシステム410を示すブロック図である。入力412A～412Nの各々は、全マルチチャンネル音声信号のうちの単一のチャンネルを表す。BRIR414A～414Nの各々は、左右の構成要素を有するバイノーラルインパルス室内応答フィルタのうちの1つのペアを表す。動作中、計算手順は、単一のチャンネル入力に関するバイノーラル音声信号を、適用されるBRIRによるロケーション表現(represents)においてレンダリングされるように生成するために入力412A～412Nの各々に対して、BRIR414A～414Nのうちの対応する1つのBRIRを単一のチャンネル(モノ)入力に適用する。次いで、N個のバイノーラル音声信号が、ステレオヘッドフォン信号または全バイノーラル音声信号を作成するために、アキュムレータ416によって累積され、出力418としてシステム410によって出力される。

【0116】

[0128]図15は、本明細書で説明する技法に従って、バイノーラル室内インパルス応答をマルチチャンネル音声信号に適用することによって生成されるバイノーラル出力信号を計算するための音声再生デバイス500の構成要素を示すブロック図である。音声再生デ

10

20

30

40

50

バイス 500 は、本開示の様々な計算削減方法を組み合わせて実施するための複数の構成要素を含む。音声再生デバイス 500 のいくつかの態様は、様々な計算削減方法のうちの任意の数の任意の組合せを含み得る。音声再生デバイス 500 は、音声再生システム 32 、音声再生デバイス 100 、音声再生デバイス 200 、および音声再生デバイス 350 のうちのいずれもの一例を表し、本開示の様々な計算削減方法を実施するために上記で列挙されたデバイスのうちのいずれもと同様の構成要素を含むことができる。

【 0117 】

[0129] 計算削減方法は、以下の部分の任意の組合せを含み得る：

[0130] a 部 (H R T F セグメント 402A および H R T F ユニット 504 に対応する) : 通常は位置特定のために数ミリ秒、両耳間遅延 (I T D) および最小位相フィルタに変換することによって計算を削減され得、一例として I I R フィルタを使用してさらに削減され得る。

【 0118 】

[0131] b 部 (反射セグメント 402B および反射ユニット 502 に対応する) : 長さは空間によって変わることがあり、一般的に通常数十ミリ秒続く。各チャンネルについて別々に行われると計算の負荷は高いが、本明細書で説明する技法は、これらのチャンネルのサブグループについて生成されたそれぞれの共通のフィルタを適用することができる。

【 0119 】

[0132] c 部 (残響セグメント 402C および残響ユニット 506 に対応する) : 共通のフィルタが、すべてのチャンネル (たとえば、22.2 フォーマットについて 22 チャンネル) に関して計算される。周波数領域のエネルギー・ディケイ・リーフ (Energy Decay Relief) (E D R) 曲線にわたる直接的な平均に基づいて新しい残響の末尾を再合成するのではなく、残響ユニット 506 は、入力信号コンテンツによって変化する調整する重みによって隨意に改良される平均に、異なる重み付け方式を適用する。

【 0120 】

[0133] 図 14 のシステム 410 と同様の方式で、音声再生デバイス 500 は、マルチチャンネル音声信号の N 個の単一のチャンネル入力 412A ~ 412N (総称して「入力 412 」) を受信し、ステレオヘッドフォン信号または全バイノーラル音声信号を生成して出力するために、バイノーラル室内インパルス応答 (B R I R) フィルタのセグメントを適用する。図 15 に示すように、反射ユニットは、(たとえば、適応的重み付け要素 520A_{1~K} ~ 520M_{1~J} 、 522A ~ 522N を使用して重み付けられた) 重み付けられた合計を使用して離散した入力 412 を異なるグループ内に結合する。共通の残響 (たとえば、図 13 の残響セクション 402C によって示される) に関して、残響ユニット 506 は、入力 412 を、それぞれの適応的重み付け要素 (たとえば、ステレオで、入力ごとに左 / 右に関する異なる重みの 522A ~ 522N) と一緒に組み合わせ、次いで、(遅延 526 を適用した後で) F F T フィルタリングを使用して適用される共通の残響フィルタ 524 (ステレオのインパルス応答フィルタ) を使用して、結合された入力を処理する。

【 0121 】

[0134] 反射ユニット 502 は、共通の残響フィルタ 524 に似た平均反射フィルタ 512A ~ 512M を、適応的重み付け要素 (520A_{1~K} ~ 520M_{1~J}) を用いてサブグループ内に一緒に結合された入力 412 の異なるサブグループに適用する。 H R T F ユニット 504 は、この例示的なデバイスでは、両耳間時間遅延 (I T D) 530A ~ 530N に変換されている頭部伝達関数 (H R T F) フィルタ 414A ~ 414N (総称して「 H R T F フィルタ 414 」) と最小位相フィルタ (これらは多状態無限インパルス応答 (I I R) フィルタを用いてさらに近似され得る) とを適用する。本明細書で使用する「適応的」は、適応的重み付け要素が適用される入力信号の品質に応じた、重み付け要素に関する調整を指す。いくつかの態様では、様々な適応的重み付け要素は、適応的でないことがある。

【 0122 】

10

20

30

40

50

[0135] 入力 4 1 2 の各々に対して B R I R に関するミキシング時間を計算するために、1 つのウインドウの標準偏差の外にあるインパルス応答タップの割合 (fraction of impulse response tap) を 1 0 2 4 のスライディングウインドウにわたって測定するエコー密度プロファイル (Echo Density Profile) が計算される。その値が最初に 1 に到達したとき、これは、インパルス応答がガウス雑音を近似し始め、残響の開始を指示することを示す。個別の H R T F フィルタ 4 1 4 の各々について、異なる計算が存在してよく、測定による最終値 (ミリ秒) は、N チャンネルにわたって平均することによって決定される。

【数 2 7】

10

$T_{mp50} = 36.1$ (50 meaning average perceptual mixing time on regression analysis)

【数 2 8】

$T_{mp95} = 80.7$ (95 meaning transparent on 95% expert listeners, more strict)

【0 1 2 3】

20

[0136] 空間容積に基づくミキシング時間計算のための理論式も存在する。3 0 0 立方メートルの大きさの室内に関して、たとえば、容積による式によれば、

【数 2 9】

$T_{v50} = 31.2$

【数 3 0】

30

$T_{v95} = 53.6$

[0137] 上記のように、H R T F ユニット 5 0 4 は、両耳間時間遅延 (I T D) 5 3 0 A ~ 5 3 0 N と最小位相フィルタとに変換されている頭部伝達関数 (H R T F) フィルタ 4 1 4 を適用する。最小位相フィルタは、元のフィルタのケプストラム (Cepstrum) に窓を適用することによって得られ得、遅延は、位相の 5 0 0 ~ 4 0 0 0 H z の周波数領域での線形回帰によって推定され得、I I R 近似に関して、バランスド・モデル・トランケイション (Balanced Model Truncation) (B M T) 法が、周波数ワープフィルタ (frequency warped filter) 上の振幅応答の最重要成分を抽出するために使用され得る。

40

【0 1 2 4】

[0138] 残響ユニット 5 0 6 に関して、ミキシング時間の後、インパルス応答の末尾 (たとえば、残響セグメント 4 0 2 C) は、理論的に、大した知覚的な差なしに交換可能である。したがって、残響ユニット 5 0 6 は、入力 4 1 2 に対応するそれぞれの B R I R の各応答の末尾を置き換えるために、共通の残響フィルタ 5 2 4 を適用する。音声再生デバイス 5 0 0 の残響ユニット 5 0 6 における適用のために共通の残響フィルタ 5 2 4 を取得するための例示的な方法が存在する：

(1) 各フィルタをそのエネルギー (たとえば、インパルス応答におけるすべてのサンプルの二乗値の合計) によって正規化し、次いで、すべての正規化されたフィルタにわたって平均する。

50

【0125】

(2) すべてのフィルタを直接平均する、たとえば、単純平均を計算する。

【0126】

(3) エネルギーの包絡線 (envelope) およびコヒーレンス制御によって制御された白色雑音を用いて平均化フィルタを再合成する。

【0127】

[0139] 第1の方法(1)は、元のフィルタの各々の特性 / 形状を同等に取る。いくつかのフィルタは、非常に低いエネルギーを有することがある(たとえば、22.2セットアップにおける上中央のチャンネル)けれども、共通のフィルタ524において同等の「票」を有する。

10

【0128】

[0140] 第2の方法(2)は、当然ながら、そのエネルギーレベルに従って各フィルタを重み付けするので、より高エネルギーの、または「より声の大きい」フィルタが、共通のフィルタ524内でより多くの票を得る。この直接の平均はまた、フィルタ間に大した相関がないことを仮定しており、これは、少なくとも、良好なリスニングルーム内で個別に取得されたB R I Rに関して当てはまり得る。

【0129】

[0141] 第3の方法(3)は、周波数依存性両耳間コヒーレンス(FDIC)がB R I Rの残響の末尾を再合成するために使用される技法に基づく。各B R I Rは、最初に、短時間フーリエ変換(STFT: short-term Fourier transform)を完了し、そのFDICは

20

【数31】

$$\Phi(i) = \frac{\sum_{k=k_0}^K \Re(H_L(i, k)H_R(i, k)^*)}{\sqrt{\sum_{k=k_0}^K |H_L(i, k)|^2 \sum_{k=k_0}^K |H_R(i, k)|^2}}$$

として計算される。上式で、iは周波数のインデックスであり、kは時間のインデックスである。R(.)は実部を示す。H_LおよびH_Rは、左右のインパルス応答の短時間フーリエ変換(STFT: Short-time Fourier Transform)である。

30

【0130】

[0142] 一定のFDICおよびEDRを用いて、インパルス応答が、ガウス雑音を使用して

【数32】

$$\tilde{H}_L(i, k) = c(i, k) (a(i, k)N_1(i, k) + b(i, k)N_2(i, k))$$

40

【数33】

$$\tilde{H}_R(i, k) = d(i, k) (a(i, k)N_1(i, k) - b(i, k)N_2(i, k))$$

のように合成され得る。ここで、

50

【数34】

$$a(i, k) = \sqrt{\frac{P_2(i, k)^2(1 + \Phi(i))}{P_1(i, k)^2(1 - \Phi(i)) + P_2(i, k)^2(1 + \Phi(i))}}$$

【数35】

10

$$\begin{aligned} b(i, k) &= \sqrt{1 - a(i, k)^2} \\ &= \sqrt{\frac{P_1(i, k)^2(1 - \Phi(i))}{P_1(i, k)^2(1 - \Phi(i)) + P_2(i, k)^2(1 + \Phi(i))}} \end{aligned}$$

である。

【0131】

20

[0143]ここで、 $H_{\sim L}$ および $H_{\sim R}$ はフィルタの合成されたS T F Tであり、 N_1 および N_2 は単独で生成されたガウス雑音のS T F Tであり、cおよびdは周波数および時間によってインデックスを付けられたE D Rであり、P sは雑音信号の時間平滑化された短時間パワースペクトル推定値である。

【0132】

[0144]平均F D I Cを取得するために、本技法は、

・元のフィルタのF D I Cのうちの1つ、たとえば、前中央のチャンネルを使用する

、

・すべてのF D I Cにわたって直接平均する、

・すべてのF D I Cの最小を使用する、これは、最大限に広範な平均化フィルタを生成することになるが、必ずしも元のフィルタミクスチャに近いとは限らない、

・F D I Cを、E D Rのそれらの相対的なエネルギーで重み付けし、次いで一緒に合計する、

を含み得る。

最後の方法（重み付けられたF D I C）について、各フィルタは、そのエネルギーに応じて共通のF D I Cにおいて「票」を有する。したがって、より声の大きいフィルタは、共通のフィルタ524においてにより多くのそれらのF D I C画像を得る。

【0133】

[0145]さらに、入力信号のレパートリーを検査することによって、追加のパターンが発見され得、コンテンツのエネルギー分布から追加の重みをもたらす。たとえば、22.2セットアップ内の上部チャンネルは、一般的に、低エネルギーB R I Rを有し、コンテンツ作成者は、その位置におけるコンテンツ（たとえば、時折の飛行機の接近飛行）を著することはほとんどない。したがって、共通の残響フィルタ524生成技法は、共通のフィルタ524を合成するときに上部チャンネルに関して正確さをトレードオフする一方で、主たる前中央、左および右のチャンネルが多くの強調を得ることができる。一般式で表現されると、複数の重みを用いて計算された共通または平均のF D I Cは、

30

40

【数36】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

として計算され、上式で、 $FDIC_i$ は i 番目の B R I R チャンネルの $FDIC$ であり、 w_{ji} (> 0) は B R I R チャンネル i に関する規準 j の重み要素である。ここで言及される j 番目の規準のうちの 1 つは B R I R エネルギーであり得、一方で、別のものは信号コンテンツのエネルギーであり得る。分母の合計 (denominator sum) は、結合された重みが最終的に合計 1 になるように正規化する。重みがすべて 1 に等しくなると、式は単純平均に置き換えられる。同様に、共通の E D R (前の式の c および d) は、
【数37】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

として計算され得、ここで、重みは、必ずしも $FDIC$ の重みと同じであるとは限らない。
。

【0134】

[0146] 共通の残響フィルタ 5 2 4 を生成することに關して説明した上記の方法のうちのいずれもまた、反射フィルタ 5 1 2 A ~ 5 1 2 M を合成するために使用され得る。すなわち、チャンネルの反射のサブグループは、同様に合成され得るが、反射によって作成された信号はあまり雑音に似ていないので、エラーは一般的により大きくなる。しかしながら、チャンネルフォーマット (たとえば、2 2 . 2) に従って、中央チャンネルの反射のすべては同様のコヒーレンス評価とエネルギー減衰とを共有し、左横チャンネルの反射のすべては適切な重み付けと結合され得、代替として、左前のチャンネルは 1 つのグループを形成し得、左後ろおよび高さのチャンネルは別のグループを形成し得る、等々。これは、計算を削減するために、反射セグメント (たとえば、反射セグメント 4 0 2 B) をそれぞれ有する N 個のチャンネルを M (たとえば、3 ~ 5) 個のサブグループに削減することができる。同様のコンテンツベースの重み付けは、残響フィルタ 5 2 4 を合成することに關して上記で説明したように、反射結合されたフィルタ 5 1 2 A ~ 5 1 2 M にも適用され得る。反射チャンネルは、任意の組合せでグループ化され得る。インパルス応答の反射セグメントの間の相関を検査することによって、比較的高い相関性があるチャンネルが、サブグループの共通の反射フィルタ 5 1 2 の合成のために一緒にグループ化され得る。
30

【0135】

[0147] 図示の例では、反射ユニット 5 0 2 は、少なくとも、入力 4 1 2 A および 4 1 2 N を 1 つのサブグループにグループ化する。反射フィルタ 5 1 2 A は、このサブグループについて生成された共通のフィルタを表し、反射ユニット 5 0 2 は、反射フィルタ 5 1 2 A を、図示の例では少なくとも入力 4 1 2 A と入力 4 1 2 N とをやはり含むサブグループの入力の結合に適用する。
40

【0136】

[0148] 一例として、B R I R フィルタのセットのそれぞれの反射部分に関する相関行列が検査される。B R I R フィルタのセットは、B R I R フィルタの現在のセットを表すことができる。相関行列は、非類似度行列 (dissimilarity matrix) を取得するために、(1 - corr) / 2 によって調整され、それは、クラスタ分析に関する完全な連結を導くために使用される。
50

【0137】

[0149]図16に示すように、階層的クラスタ分析は、それらの時間の包絡線での相関に従う22.2チャンネルのB R I Rのセットの反射部分で実行され得る。理解されるように、0.6のカットオフスコアを設定することによって、説得力のある類似性を伴う、左チャンネルは4つのサブグループにグループ化され、右チャンネルは3つのサブグループにグループ化され得る。22.2セットアップにおけるスピーカーロケーションを検査することによって、クラスタ分析の結果は、22.2チャンネルのセットアップの常識的な機能性および形状(geometry)と一致する。

【0138】

[0150]次に図15に戻ると、共通のフィルタ(たとえば、反射フィルタ512A~512Mおよび共通の残響フィルタ524)のうちのいずれに関するインパルス応答も、2列ベクトル

【数38】

10

$$\tilde{h} = [h_L \ h_R] = [IFFT(\tilde{H}_L(i, k)) \ IFFT(\tilde{H}_R(i, k))]$$

であり得る。

【0139】

20

[0151]共通のフィルタが計算されると、オンラインの処理において、反射ユニット502および/または残響ユニット506は、最初に入力412をフィルタに関する特定のグループの中にミックスし、次いで共通のフィルタを適用する。たとえば、残響ユニット506は、すべての412を中心にミックスし、次いで共通の残響フィルタ524を適用する。共通のフィルタ合成の前の元のフィルタは変化するエネルギーを有するので、等しくミックスされた入力412は、元の条件に適合しない。フィルタインパルス応答hのエネルギーが、

【数39】

20

$$E(h) = \sum_{n=0}^{N-1} h[n]^2$$

30

として計算され、上式で、nはサンプルのインデックスであり、各h[n]は左/右インパルス応答に関するステレオサンプルである場合)、入力信号に関する初期の重みは、

【数40】

40

$$\hat{w}_i = \sqrt{\frac{E(h_i)}{E(\tilde{h})}}$$

として計算され得、上式で、 h_i は、共通のフィルタ合成前のチャンネル*i*に関する元のフィルタである。

【0140】

50

[0152] 共通のフィルタを使用することによって、
【数41】

$$\sum (in_i \odot h_i)$$

の元のフィルタリングプロセスは
【数42】

10

$$\sum \hat{w}_i in_i \odot \tilde{h}$$

となり、上式で、 in_i は入力信号に関する入力サンプルである。ここで、
【数43】

④

20

は畳み込みを示し、各 h フィルタはステレオインパルス応答であり、したがって、左右の
チャンネルは、これらのプロセスを個別に搬送する。もう少し効率的な処理のために、ス

テレオの重み
【数44】

$$\hat{w}_i$$

のうちのいずれもが、左 / 右の重みを平均することによって単一の値の重みに変換され得
、次いで共通のフィルタの適用時のステレオ入力ミックスが、代わりにモノミックスにな
る。反射ユニット 502 に関する適応的重み要素 520A_{1~K} ~ 520M_{1~J} および残響
ユニット 506 に関する適応的重み要素 522A ~ 522N は、重み
【数45】

30

$$\hat{w}_i$$

のうちのいずれをも表し得る。

【0141】

[0153] 入力信号で

【数46】

$$\hat{w}_i$$

を使用することによって、基本的な仮定は、入力チャンネルは相関されないことであり、

50

したがって、各入力は前と同じエネルギーでフィルタを通過し、合計された信号のエネルギーは、すべての重み付けられた信号のエネルギーの合計とほぼ同じである。実際には、より強い「残響の」音が知覚されることが多く、より高いエネルギーレベルの再合成バージョンが観測される。これは、入力チャンネルが、しばしば相関付けられるという事実に起因する。たとえば、モノ音源をパニングし、それらをあちこち移動させることによって生成されたマルチチャンネルミックスについて、パニングアルゴリズムは、通常、異なるチャンネルにわたって高度に相関付けられた成分を生成する。そして、相関付けられたチャンネルに関して、エネルギーは、初期の重み

【数47】

10

 \hat{w}_i

を使用してより高くなる。

【0142】

[0154]したがって、ミックスされた入力信号を

【数48】

20

$$in_{mix} = \sum \hat{w}_i in_i$$

として計算する代わりに、時間で変わるエネルギー正規化の重みが適用され得、新しい入力信号ミックスは、したがって、

【数49】

$$in_{mix}(n) = w_{norm}(n) \sum \hat{w}_i in_i(n)$$

30

として計算されるべきであり、

上式で、nは離散時間のインデックスであり、正規化 w_{norm} は、信号フレームのうちの1つのセグメントにわたる、重み付けられた信号の合計されたエネルギーと重み付けられた合計された信号のエネルギーとの間のエネルギーの比：

【数50】

$$w_{norm}(n) = \sqrt{\frac{\sum E(\hat{w}_i in_i)}{E(\sum \hat{w}_i in_i)}}$$

40

に従う。式中、信号のインデックスは右辺の中に書かれていない。右辺におけるこの平均エネルギー推定は、合計されたエネルギーのエネルギーおよび合計された信号のエネルギーに対する1次の平滑化フィルタを用いて時間領域において達成され得る。したがって、滑らかなエネルギー曲線が、除算について取得され得る。または、音声再生デバイス500は、すでに、フィルタリングに対してFFTオーバーラップ加算を適用し得るので、各FFTフレームについて、音声再生デバイス500は1つの正規化重みを推定し得、オーバーラップ加算方式は、すでに、時間にわたる平滑化効果を処理することになる。

【0143】

50

[0155] H R T F セグメントと、反射および残響の末尾（または残響）セグメントとの間で、コサインカーブ・クロスフェードが、それらの間を滑らかに移行するために（たとえば、0.2 m s または 10 サンプルの持続時間で）適用される。たとえば、H R T F が 256 サンプル長である場合、反射は 2048 サンプル長であり、残響は 4096 サンプル長であり、レンダラの全等価フィルタ長さは、 $256 + 2048 + 4096 - 2 \times 10 = 6380$ サンプルとなる。

【 0 1 4 4 】

[0156] 結合ステップ 510 は、反射ユニット 502、H R T F ユニット 504、および残響ユニット 506 によって生成されたフィルタリングされた信号のすべてを結合する。いくつかの例では、反射ユニット 502 および残響ユニット 506 のうちの少なくとも一方は、適応的重み要素を適用することを含まない。音声再生デバイス 500 のいくつかの例では、H R T F ユニット 504 は、入力 412 に関する B R I R フィルタの H R T F 部分と反射部分の両方を適用し、すなわち、そのような例における音声再生デバイス 500 は、入力 412 N を、共通の反射フィルタ 512 A ~ 512 M が適用される M 個のサブグループにグループ化しない。

【 0 1 4 5 】

[0157] 図 17 は、本開示で説明する技法による、音声再生デバイスの動作の例示的なモードを示すフローチャートである。例示的な演算のモードについて、図 15 の音声再生デバイス 500 に関して説明する。

【 0 1 4 6 】

[0158] 音声再生デバイス 500 は、単一の入力チャンネルを受信し、適応的に決定された重みをそのチャンネルに適用する（600）。音声再生デバイス 500 は、結合された音声信号を生成するために、これらの適応的に重み付けられたチャンネルを結合する（602）。音声再生デバイス 500 は、さらに、バイノーラル音声信号を生成するために、バイノーラル室内インパルス応答フィルタを結合された音声信号に適用する（604）。バイノーラル室内インパルス応答フィルタは、たとえば、上記で説明した技法のうちのいずれに従って生成された、結合された反射フィルタまたは残響フィルタであってよい。音声再生デバイス 500 は、ステップ 604 において生成されたバイノーラル音声信号から少なくとも部分的に生成された出力音声信号 / 全音声信号を出力する（606）。全音声信号は、1つまたは複数の、結合されたまたはフィルタリングされた反射サブグループ、結合されたまたはフィルタリングされた残響グループ、および音声信号のチャンネルの各々に関するフィルタリングされたそれぞれの H R T F 信号に関する、複数のバイノーラル音声信号の結合であり得る。音声再生デバイス 500 は、全出力バイノーラル音声信号を作成するために結合のための信号を整列させるために、必要に応じて、フィルタリングされた信号に遅延を適用する。

【 0 1 4 7 】

[0159] 上記の追加または代替として、以下の例を説明する。以下の例のうちのいずれにおいて説明する特徴は、本明細書で説明する他の例のうちのいずれとともに利用され得る。

【 0 1 4 8 】

[0160] 一例は、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを取得することと、変換された合計音声信号を生成するために音声信号の複数のチャンネルから決定された合計音声信号に共通のフィルタを適用することと、を備える、音声信号をバイノーラル化する方法を対象とする。

【 0 1 4 9 】

[0161] いくつかの例では、合計音声信号は、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号の複数のチャンネルのサブグループの結合を備える。

【 0 1 5 0 】

[0162] いくつかの例では、方法は、音声信号の複数の変換されたチャンネルを生成する

10

20

30

40

50

ために複数のバイノーラル室内インパルス応答フィルタのそれぞれの頭部伝達関数セグメントを音声信号の複数のチャンネルのうちの対応するチャンネルに適用することと、出力バイノーラル音声信号を生成するために第1の変換された合計音声信号と音声信号の変換されたチャンネルとを結合することと、をさらに備える。

【0151】

[0163]いくつかの例では、共通のフィルタを取得することは、共通のフィルタとして、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を計算することを備える。

【0152】

[0164]いくつかの例では、方法は、合計音声信号を生成するために、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号のチャンネルのサブグループを結合することをさらに備える。 10

【0153】

[0165]いくつかの例では、共通のフィルタは第1の共通のフィルタであり、サブグループは第1のサブグループであり、合計音声信号は第1の合計音声信号であり、変換された合計音声信号は第1の変換された合計音声信号であり、方法は、複数のバイノーラル室内インパルス応答フィルタの第2のサブグループの平均を計算することによって複数のバイノーラル室内インパルス応答フィルタの第2の異なるサブグループに関する第2の共通のフィルタを生成することと、第2の合計音声信号を生成するために複数のバイノーラル室内インパルス応答フィルタの第2のサブグループに対応する音声信号のチャンネルの第2のサブグループを結合することと、第2の変換された合計音声信号を生成するために第2の共通のフィルタを第2の合計音声信号に適用することと、をさらに備え、出力音声信号を生成するために第1の変換された合計音声信号と音声信号の変換されたチャンネルとを結合することは、出力音声信号を生成するために、第1の変換された合計音声信号と、第2の変換された合計音声信号と、音声信号の変換されたチャンネルとを結合することを備える。 20

【0154】

[0166]いくつかの例では、共通のフィルタを取得することは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの加重平均を計算することを備える。 30

【0155】

[0167]いくつかの例では、共通のフィルタを取得することは、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を計算することを備える。

【0156】

[0168]いくつかの例では、共通のフィルタを取得することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの直接平均を計算することを備える。

【0157】

[0169]いくつかの例では、共通のフィルタを取得することは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成することを備える。 40

【0158】

[0170]いくつかの例では、共通のフィルタを取得することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を備える。

【0159】

50

20

30

40

50

[0171] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、直接平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【 0 1 6 0 】

[0172] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【 0 1 6 1 】

[0173] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、エネルギーディケイレリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、を備える。

【 0 1 6 2 】

[0174] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、
【数 5 1】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

10

20

を計算することを備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【 0 1 6 3 】

[0175] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【 0 1 6 4 】

[0176] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することは、

【数 5 2】

30

40

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

を計算することを備え、 $EDR_{average}$ は平均エネルギーディケイレリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイレリーフ値を示し、 w_{ij} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【 0 1 6 5 】

50

[0177]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0166】

[0178]いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0167】

[0179]いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0168】

[0180]いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0169】

[0181]別の例では、方法は、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを生成することを備える。

【0170】

[0182]いくつかの例では、共通のフィルタを生成することは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントの加重平均を計算することを備える。

【0171】

[0183]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタのうちのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの平均を計算することを備える。

【0172】

[0184]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの直接平均を計算することを備える。

【0173】

[0185]いくつかの例では、共通のフィルタを生成することは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成することを備える。

【0174】

[0186]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を備える。

【0175】

[0187]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、直接平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【0176】

[0188]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【0177】

[0189]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、

10

20

30

40

50

を備える。

【0178】

[0190]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、
【数53】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

10

を計算することを備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのうちの1つのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0179】

[0191]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのうちの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

20

【0180】

[0192]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することは、

【数54】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

30

を計算することを備え、 $EDR_{average}$ は平均エネルギーディケイレーフ値であり、 i は音声信号のチャンネルを示し、 EDR_i は音声信号の i 番目のチャンネルに関するエネルギーディケイレーフ値を示し、 w_{ij} は音声信号の i 番目のチャンネルに関する規準 j の重みを示す。

【0181】

[0193]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号の i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

40

【0182】

[0194]別の例では、方法は、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを生成することを備える。

【0183】

[0195]いくつかの例では、共通のフィルタを生成することは、バイノーラル室内インパルス応答フィルタのサブグループのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの加重平均を計算することを備える。

【0184】

[0196]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正

50

規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの平均を計算することを備える。

【0185】

[0197]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの直接平均を計算することを備える。

【0186】

[0198]いくつかの例では、共通のフィルタを生成することは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成することを備える。

10

【0187】

[0199]いくつかの例では、共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を備える。

【0188】

[0200]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、直接平均周波数依存性両耳間コヒーレンス値を計算することを備える。

20

【0189】

[0201]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【0190】

[0202]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、を備える。

30

【0191】

[0203]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算することは、【数55】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

40

を計算することを備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0192】

50

[0204]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0193】

[0205]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することは、

【数56】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

10

を計算することを備え、 $EDR_{average}$ は平均エネルギーディケイリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイリーフ値を示し、 w_{ij} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【0194】

[0206]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

20

【0195】

[0207]別の例では、音声信号をバイノーラル化する方法は、複数のバイノーラル室内インパルス応答フィルタの 1 つまたは複数のセグメントを適用する前に、適応的に決定された重みを音声信号の複数のチャンネルに適用することと、1 つまたは複数のセグメントを複数のバイノーラル室内インパルス応答フィルタに適用することと、を備える。

【0196】

[0208]いくつかの例では、音声信号のチャンネルに関する初期の適応的に決定された重みは、複数のバイノーラル室内インパルス応答フィルタのうちの対応するバイノーラル室内インパルス応答フィルタのエネルギーに従って計算される。

30

【0197】

[0209]いくつかの例では、方法は、複数のバイノーラル室内インパルス応答フィルタに関する共通のフィルタを取得することをさらに備え、 i 番目のチャンネルに関する i 番目の初期の適応的に決定された重み

【数57】

\hat{w}_i

40

は、

【数58】

$$\hat{w}_i = \sqrt{\frac{E(h_i)}{E(\tilde{h})}}$$

50

に従って計算され、 h_i は i 番目のバイノーラル室内インパルス応答フィルタであり、
【数 5 9】

\tilde{h}

は共通のフィルタであり、
【数 6 0】

10

$$E(h) = \sum_{n=0}^{N-1} h[n]^2$$

であり、 n はサンプルのインデックスであり、各 $h[n]$ は n におけるステレオサンプルである。

【0 1 9 8】

[0210] いくつかの例では、方法は、
【数 6 1】

20

$$\sum \hat{w}_i in_i \odot \tilde{h}$$

を計算することによって、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用することをさらに備え、

【数 6 2】

\odot

30

は畳み込み演算を示し、 in_i は音声信号の i 番目のチャンネルを示す。

【0 1 9 9】

[0211] いくつかの例では、それぞれの適応的重み要素をチャンネルに適用することによって合計音声信号を生成するために音声信号のチャンネルを結合することは、
【数 6 3】

$$in_{mix}(n) = w_{norm}(n) \sum \hat{w}_i in_i(n)$$

40

を計算することを備え、 $in_{mix}(n)$ は合計音声信号を示し、 n はサンプルのインデックスであり、

【数64】

$$w_{norm}(n) = \sqrt{\frac{\sum E(\hat{w}_i in_i)}{E(\sum \hat{w}_i in_i)}}$$

であり、 in_i は音声信号の*i*番目のチャンネルを示す。

【0200】

[0212]いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0201】

[0213]いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0202】

[0214]いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0203】

[0215]別の例では、方法は、音声信号の複数の変換されたチャンネルを生成するために複数のバイノーラル室内インパルス応答フィルタのそれぞれの頭部伝達関数セグメントを音声信号の対応するチャンネルに適用することと、複数のバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの加重平均を計算することによって共通のフィルタを生成することと、合計音声信号を生成するために音声信号のチャンネルを結合することと、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用することと、出力音声信号を生成するために変換された合計音声信号と音声信号の変換されたチャンネルとを結合することと、を備える。

【0204】

[0216]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの加重平均を計算することによって共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタのうちのいずれをも正規化することなしに複数のバイノーラル室内インパルス応答フィルタの平均を計算することを備える。

【0205】

[0217]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの加重平均を計算することによって共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタの直接平均を計算することを備える。

【0206】

[0218]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの加重平均を計算することによって共通のフィルタを生成することは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成することを備える。

【0207】

[0219]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの加重平均を計算することによって共通のフィルタを生成することは、複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィ

10

20

30

40

50

ルタを合成することと、を備える。

【0208】

[0220]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することは、直接平均周波数依存性両耳間コヒーレンス値を計算することを備える。

【0209】

[0221]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することは、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算することを備える。

10

【0210】

[0222]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、を備える。

20

【0211】

[0223]いくつかの例では、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することは、

【数65】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

30

を計算することを備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのうちの1つのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0212】

[0224]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号の i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

40

【0213】

[0225]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することは、

【数66】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

を計算することを備え、 $EDR_{average}$ は平均エネルギーディケイリーフ値であり、 i は音声信号のチャンネルを示し、 EDR_i は音声信号の i 番目のチャンネルに関するエネルギーディケイリーフ値を示し、 w_{ij} は音声信号の i 番目のチャンネルに関する規準 j の重みを示す。

【0214】

[0226]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号の i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0215】

[0227]いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0216】

[0228]いくつかの例では、複数の階層要素は、球面調和係数を備える。

20

【0217】

[0229]いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0218】

[0230]別の例では、方法は、音声信号の複数の変換されたチャンネルを生成するために複数のバイノーラル室内インパルス応答フィルタのそれぞれの頭部伝達関数セグメントを音声信号の対応するチャンネルに適用することと、複数のバイノーラル室内インパルス応答フィルタの平均を計算することによって共通のフィルタを生成することと、それぞれの適応的重み要素をチャンネルに適用することによって合計音声信号を生成するために音声信号のチャンネルを結合することと、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用することと、出力音声信号を生成するために変換された合計音声信号と音声信号の変換されたチャンネルとを結合することと、を備える。

30

【0219】

[0231]いくつかの例では、音声信号のチャンネルに関する初期の適応的重み要素は、複数のバイノーラル室内インパルス応答フィルタのうちの対応するバイノーラル室内インパルス応答フィルタのエネルギーに従って計算される。

【0220】

[0232]いくつかの例では、 i 番目のチャンネルに関する i 番目の初期の適応的重み要素

【数67】

は、

\hat{w}_i

40

【数 6 8】

$$\hat{w}_i = \sqrt{\frac{E(h_i)}{E(\tilde{h})}}$$

に従って計算され、 h_i は i 番目のバイノーラル室内インパルス応答フィルタであり、
【数 6 9】

10

\tilde{h}

は共通のフィルタであり、

【数 7 0】

$$E(h) = \sum_{n=0}^{N-1} h[n]^2$$

20

であり、 n はサンプルのインデックスであり、各 $h[n]$ は n におけるステレオサンプルである。

【0 2 2 1】

[0233] いくつかの例では、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用することは、

【数 7 1】

$$\sum \hat{w}_i i n_i \odot \tilde{h}$$

30

を計算することを備え、

【数 7 2】

⊗

40

は畳み込み演算を示し、 $i n_i$ は音声信号の i 番目のチャンネルを示す。

【0 2 2 2】

[0234] いくつかの例では、それぞれの適応的重み要素をチャンネルに適用することによって合計音声信号を生成するために音声信号のチャンネルを結合することは、

【数73】

$$in_{mix}(n) = w_{norm}(n) \sum \hat{w}_i in_i(n)$$

を計算することを備え、 $in_{mix}(n)$ は合計音声信号を示し、 n はサンプルのインデックスであり、

【数74】

$$w_{norm}(n) = \sqrt{\frac{\sum E(\hat{w}_i in_i)}{E(\sum \hat{w}_i in_i)}}$$

であり、 in_i は音声信号の i 番目のチャンネルを示す。

【0223】

[0235] いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0224】

[0236] いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0225】

[0237] いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0226】

[0238] いくつかの例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを記憶するように構成されたメモリと、変換された合計音声信号を生成するために音声信号の複数のチャンネルから決定された合計音声信号に共通のフィルタを適用するように構成されたプロセッサとを備える。

【0227】

[0239] いくつかの例では、合計音声信号は、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号の複数のチャンネルのサブグループの結合を備える。

【0228】

[0240] いくつかの例では、プロセッサは、音声信号の複数の変換されたチャンネルを生成するために複数のバイノーラル室内インパルス応答フィルタのそれぞれの頭部伝達関数セグメントを音声信号の複数のチャンネルのうちの対応するチャンネルに適用することと、出力バイノーラル音声信号を生成するために第 1 の変換された合計音声信号と音声信号の変換されたチャンネルとを結合することと、を行うようにさらに構成される。

【0229】

[0241] いくつかの例では、共通のフィルタは、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を備える。

【0230】

[0242] いくつかの例では、プロセッサは、合計音声信号を生成するために、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号のチャンネルのサブグループを結合するようにさらに構成される。

【0231】

[0243] いくつかの例では、共通のフィルタは第 1 の共通のフィルタであり、サブグループは第 1 のサブグループであり、合計音声信号は第 1 の合計音声信号であり、変換された

10

20

30

40

50

合計音声信号は第1の変換された合計音声信号であり、プロセッサは、複数のバイノーラル室内インパルス応答フィルタの第2のサブグループの平均を計算することによって複数のバイノーラル室内インパルス応答フィルタの第2の異なるサブグループに関する第2の共通のフィルタを生成することと、第2の合計音声信号を生成するために複数のバイノーラル室内インパルス応答フィルタの第2のサブグループに対応する音声信号のチャンネルの第2のサブグループを結合することと、第2の変換された合計音声信号を生成するために第2の共通のフィルタを第2の合計音声信号に適用することと、を行うようにさらに構成され、出力音声信号を生成するために第1の変換された合計音声信号と音声信号の変換されたチャンネルとを結合するために、プロセッサは、出力音声信号を生成するために、第1の変換された合計音声信号と、第2の変換された合計音声信号と、音声信号の変換されたチャンネルとを結合するようにさらに構成される。

【0232】

[0244]いくつかの例では、共通のフィルタは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの加重平均を備える。

【0233】

[0245]いくつかの例では、共通のフィルタは、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を備える。

【0234】

[0246]いくつかの例では、共通のフィルタは、複数のバイノーラル室内インパルス応答フィルタのサブグループの直接平均を備える。

【0235】

[0247]いくつかの例では、共通のフィルタは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して生成され、再合成された共通のフィルタを備える。

【0236】

[0248]いくつかの実施形態では、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を行うようにさらに構成される。

【0237】

[0249]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、直接平均周波数依存性両耳間コヒーレンス値を計算するようにさらに構成される。

【0238】

[0250]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するようにさらに構成される。

【0239】

[0251]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、を行うようにさらに構成される。

10

20

30

40

50

【0240】

[0252] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、
【数75】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

10

を計算するようにさらに構成され、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0241】

[0253] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0242】

[0254] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するために、プロセッサは、

【数76】

20

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

30

を計算するようにさらに構成され、 $EDR_{average}$ は平均エネルギーディケイレリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイレリーフ値を示し、 w_{ij} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【0243】

[0255] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0244】

40

[0256] いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0245】

[0257] いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0246】

[0258] いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0247】

[0259] 別の例では、デバイスは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを生成するように構成されたプロセッサを備える。

50

【0248】

[0260]いくつかの例では、共通のフィルタを生成するために、プロセッサは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントの加重平均を計算するようさら構成される。

【0249】

[0261]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのうちのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの平均を計算するようさら構成される。

10

【0250】

[0262]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの直接平均を計算するようさら構成される。

【0251】

[0263]いくつかの例では、共通のフィルタを生成するために、プロセッサは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成するようさら構成される。

【0252】

[0264]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を行うようさら構成される。

20

【0253】

[0265]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、直接平均周波数依存性両耳間コヒーレンス値を計算するようさら構成される。

30

【0254】

[0266]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するようさら構成される。

【0255】

[0267]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けすることと、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積することと、を行うようさら構成される。

40

【0256】

[0268]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、

【数77】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

を計算するようにさらに構成され、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのうちの1つのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。
10

【0257】

[0269]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのうちの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0258】

[0270]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するために、プロセッサは、
20

【数78】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

を計算するようにさらに構成され、 $EDR_{average}$ は平均エネルギーディケイリーフ値であり、 i は音声信号のチャンネルを示し、 EDR_i は音声信号の i 番目のチャンネルに関するエネルギーディケイリーフ値を示し、 w_{ij} は音声信号の i 番目のチャンネルに関する規準 j の重みを示す。
30

【0259】

[0271]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号の i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0260】

[0272]別の例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを生成するように構成されたプロセッサを備える。

【0261】

[0273]いくつかの例では、共通のフィルタを生成するために、プロセッサは、バイノーラル室内インパルス応答フィルタのサブグループのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの加重平均を計算するようにさらに構成される。
40

【0262】

[0274]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの平均を計算するようにさらに構成される。

【0263】

[0275]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバ
50

イノーラル室内インパルス応答フィルタのサブグループの反射セグメントの直接平均を計算するようにさらに構成される。

【0264】

[0276]いくつかの例では、共通のフィルタを生成するために、プロセッサは、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成するようにさらに構成される。

【0265】

[0277]いくつかの例では、共通のフィルタを生成するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算することと、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成することと、を行うようにさらに構成される。

10

【0266】

[0278]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、直接平均周波数依存性両耳間コヒーレンス値を計算するようにさらに構成される。

【0267】

[0279]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するようにさらに構成される。

20

【0268】

[0280]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けするようにさらに構成され、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積している。

30

【0269】

[0281]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するために、プロセッサは、

【数79】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

40

を計算するようにさらに構成され、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0270】

[0282]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィル

50

タに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0271】

[0283] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するために、プロセッサは、

【数80】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

10

を計算するようにさらに構成され、 $EDR_{average}$ は平均エネルギーディケイリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイリーフ値を示し、 w_{ji} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【0272】

[0284] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

20

【0273】

[0285] いくつかの例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタの 1 つまたは複数のセグメントを適用する前に、適応的に決定された重みを音声信号の複数のチャンネルに適用することと、 1 つまたは複数のセグメントを複数のバイノーラル室内インパルス応答フィルタに適用することと、を行うように構成されたプロセッサを備える。

【0274】

[0286] いくつかの例では、プロセッサは、複数のバイノーラル室内インパルス応答フィルタのうちの対応するバイノーラル室内インパルス応答フィルタのエネルギーに従って、音声信号のチャンネルに関する初期の適応的に決定された重みを計算する。

30

【0275】

[0287] いくつかの例では、プロセッサは、複数のバイノーラル室内インパルス応答フィルタに関する共通のフィルタを取得するようにさらに構成され、 i 番目のチャンネルに関する i 番目の初期の適応的に決定された重み

【数81】

$$\hat{w}_i$$

40

は、

【数82】

$$\hat{w}_i = \sqrt{\frac{E(h_i)}{E(\tilde{h})}}$$

に従って計算され、 h_i は i 番目のバイノーラル室内インパルス応答フィルタであり、

【数 8 3】

 \tilde{h}

は共通のフィルタであり、

【数 8 4】

10

$$E(h) = \sum_{n=0}^{N-1} h[n]^2$$

であり、n はサンプルのインデックスであり、各 $h[n]$ は n におけるステレオサンプルである。

【0 2 7 6】

[0288] いくつかの例では、プロセッサは、

【数 8 5】

20

$$\sum \hat{w}_i in_i \odot \tilde{h}$$

を計算することによって、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用するようにさらに構成され、

【数 8 6】

 \odot

30

は畳み込み演算を示し、 in_i は音声信号の i 番目のチャンネルを示す。

【0 2 7 7】

[0289] いくつかの例では、プロセッサは、

【数 8 7】

$$in_{mix}(n) = w_{norm}(n) \sum \hat{w}_i in_i(n)$$

40

を計算することによってそれぞれの適応的重み要素をチャンネルに適用することによって合計音声信号を生成するために音声信号のチャンネルを結合するようにさらに構成され、 $in_{mix}(n)$ は合計音声信号を示し、n はサンプルのインデックスであり、

【数88】

$$w_{norm}(n) = \sqrt{\frac{\sum E(\hat{w}_i in_i)}{E(\sum \hat{w}_i in_i)}}$$

であり、 in_i は音声信号の*i*番目のチャンネルを示す。

【0278】

[0290]いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0279】

[0291]いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0280】

[0292]いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0281】

[0293]別の例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを取得するための手段と、変換された合計音声信号を生成するために音声信号の複数のチャンネルから決定された合計音声信号に共通のフィルタを適用するための手段と、を備える。

【0282】

[0294]いくつかの例では、合計音声信号は、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号の複数のチャンネルのサブグループの結合を備える。

【0283】

[0295]いくつかの例では、デバイスは、音声信号の複数の変換されたチャンネルを生成するために複数のバイノーラル室内インパルス応答フィルタのそれぞれの頭部伝達関数セグメントを音声信号の複数のチャンネルのうちの対応するチャンネルに適用するための手段と、出力バイノーラル音声信号を生成するために第1の変換された合計音声信号と音声信号の変換されたチャンネルとを結合するための手段と、をさらに備える。

【0284】

[0296]いくつかの例では、共通のフィルタを取得するための手段は、共通のフィルタとして、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を計算するための手段を備える。

【0285】

[0297]いくつかの例では、デバイスは、合計音声信号を生成するために、複数のバイノーラル室内インパルス応答フィルタのサブグループに対応する音声信号のチャンネルのサブグループを結合するための手段をさらに備える。

【0286】

[0298]いくつかの例では、共通のフィルタは第1の共通のフィルタであり、サブグループは第1のサブグループであり、合計音声信号は第1の合計音声信号であり、変換された合計音声信号は第1の変換された合計音声信号であり、デバイスは、複数のバイノーラル室内インパルス応答フィルタの第2のサブグループの平均を計算することによって複数のバイノーラル室内インパルス応答フィルタ第2の異なるサブグループに関する第2の共通のフィルタを生成するための手段と、第2の合計音声信号を生成するために複数のバイノーラル室内インパルス応答フィルタの第2のサブグループに対応する音声信号のチャンネルの第2のサブグループを結合するための手段と、第2の変換された合計音声信号を生成するために第2の共通のフィルタを第2の合計音声信号に適用するための手段と、をさらに備え、出力音声信号を生成するために第1の変換された合計音声信号と音声信号の変換されたチャンネルとを結合するための手段が、出力音声信号を生成するために、第1の変換された合計音声信号と、第2の変換された合計音声信号と、音声信号の変換されたチャンネルとを結合するための手段を備える。

10

20

30

40

50

【0287】

[0299] いくつかの例では、共通のフィルタを取得するための手段は、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの加重平均を計算するための手段を備える。

【0288】

[0300] いくつかの例では、共通のフィルタを取得するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの平均を計算するための手段を備える。

10

【0289】

[0301] いくつかの例では、共通のフィルタを取得するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの直接平均を計算するための手段を備える。

【0290】

[0302] いくつかの例では、共通のフィルタを取得するための手段は、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成するための手段を備える。

【0291】

[0303] いくつかの例では、共通のフィルタを取得するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算するための手段と、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算するための手段と、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段と、を備える。

20

【0292】

[0304] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、直接平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

【0293】

[0305] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

30

【0294】

[0306] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けするための手段と、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積するための手段と、を備える。

40

【0295】

[0307] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、

【数89】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

を計算するための手段を備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス

50

値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $F D I C_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0296】

[0308] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0297】

[0309] いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段は、

【数90】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

を計算するための手段を備え、 $EDR_{average}$ は平均エネルギーディケイレリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイレリーフ値を示し、 w_{ij} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【0298】

[0310] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0299】

[0311] いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

【0300】

[0312] いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0301】

[0313] いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0302】

[0314] 別の例では、デバイスは、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを生成するための手段を備える。

【0303】

[0315] いくつかの例では、共通のフィルタを生成するための手段は、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントの加重平均を計算するための手段を備える。

【0304】

[0316] いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラル室内インパルス応答フィルタのうちのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの平均を計算するための手段を備える。

【0305】

[0317] いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラ

10

20

30

40

50

ル室内インパルス応答フィルタの残響セグメントの直接平均を計算するための手段を備える。

【0306】

[0318]いくつかの例では、共通のフィルタを生成するための手段は、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成するための手段を備える。

【0307】

[0319]いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算するための手段と、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算するための手段と、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段と、を備える。

10

【0308】

[0320]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、直接平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

【0309】

[0321]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

20

【0310】

[0322]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、エネルギーディケイリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタの残響セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けするための手段と、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積するための手段と、を備える。

【0311】

30

[0323]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、

【数91】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

40

を計算するための手段を備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのうちの1つのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0312】

[0324]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのうちの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0313】

50

[0325]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段は、

【数92】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

10

を計算するための手段を備え、 $EDR_{average}$ は平均エネルギーディケイレリーフ値であり、 i は音声信号のチャンネルを示し、 EDR_i は音声信号の*i*番目のチャンネルに関するエネルギーディケイレリーフ値を示し、 w_{ij} は音声信号の*i*番目のチャンネルに関する規準*j*の重みを示す。

【0314】

[0326]いくつかの例では、規準*j*は、*i*番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号の*i*番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0315】

[0327]別の例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを生成するための手段を備える。

20

【0316】

[0328]いくつかの例では、共通のフィルタを生成するための手段は、バイノーラル室内インパルス応答フィルタのサブグループのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの加重平均を計算するための手段を備える。

【0317】

[0329]いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを正規化することなく、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの平均を計算するための手段を備える。

30

【0318】

[0330]いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの直接平均を計算するための手段を備える。

【0319】

[0331]いくつかの例では、共通のフィルタを生成するための手段は、エネルギーの包絡線およびコヒーレンス制御によって制御された白色雑音を使用して共通のフィルタを再合成するための手段を備える。

【0320】

40

[0332]いくつかの例では、共通のフィルタを生成するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算するための手段と、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を使用して平均周波数依存性両耳間コヒーレンス値を計算するための手段と、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段と、を備える。

【0321】

[0333]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、直接平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

50

【0322】

[0334]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値のうちの最小周波数依存性両耳間コヒーレンス値として、平均周波数依存性両耳間コヒーレンス値を計算するための手段を備える。

【0323】

[0335]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、エネルギーディケイレリーフのそれぞれの相対エネルギーによって複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値の各々を重み付けするための手段と、平均周波数依存性両耳間コヒーレンス値を生成するために重み付けられた周波数依存性両耳間コヒーレンス値を累積するための手段と、を備える。10

【0324】

[0336]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を計算するための手段は、

【数93】

$$FDIC_{average} = \frac{\sum_i (\prod_j w_{ji} FDIC_i)}{\sum_i (\prod_j w_{ji})}$$

20

を計算するための手段を備え、 $FDIC_{average}$ は平均周波数依存性両耳間コヒーレンス値であり、 i は複数のバイノーラル室内インパルス応答フィルタのサブグループのバイノーラル室内インパルス応答フィルタを示し、 $FDIC_i$ は i 番目のバイノーラル室内インパルス応答フィルタに関する周波数依存性両耳間コヒーレンス値を示し、 w_{ij} は i 番目のバイノーラル室内インパルス応答フィルタに関する規準 j の重みを示す。

【0325】

30

[0337]いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0326】

[0338]いくつかの例では、平均周波数依存性両耳間コヒーレンス値を使用して共通のフィルタを合成するための手段は、

【数94】

$$EDR_{average} = \frac{\sum_i (\prod_j w_{ji} EDR_i)}{\sum_i (\prod_j w_{ji})}$$

40

を計算するための手段を備え、 $EDR_{average}$ は平均エネルギーディケイレリーフ値であり、 i は音声信号のチャンネルのサブグループのチャンネルを示し、 EDR_i は音声信号のチャンネルのサブグループの i 番目のチャンネルに関するエネルギーディケイレリーフ値を示し、 w_{ij} は音声信号のチャンネルのサブグループの i 番目のチャンネルに関する規準 j の重みを示す。

【0327】

50

[0339] いくつかの例では、規準 j は、 i 番目のバイノーラル室内インパルス応答フィルタに関するエネルギー、または音声信号のチャンネルのサブグループの i 番目のチャンネルに関する信号コンテンツのエネルギーのうちの一方である。

【0328】

[0340] 別の例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタの 1 つまたは複数のセグメントを適用する前に、適応的に決定された重みを音声信号の複数のチャンネルに適用するための手段と、1 つまたは複数のセグメントを複数のバイノーラル室内インパルス応答フィルタに適用するための手段と、を備える。

【0329】

[0341] いくつかの例では、音声信号のチャンネルに関する初期の適応的に決定された重みは、複数のバイノーラル室内インパルス応答フィルタのうちの対応するバイノーラル室内インパルス応答フィルタのエネルギーに従って計算される。 10

【0330】

[0342] いくつかの例では、デバイスは、複数のバイノーラル室内インパルス応答フィルタに関する共通のフィルタを取得するための手段をさらに備え、 i 番目のチャンネルに関する i 番目の初期の適応的に決定された重み

【数95】

\hat{w}_i

20

は、

【数96】

$$\hat{w}_i = \sqrt{\frac{E(h_i)}{E(\tilde{h})}}$$

30

に従って計算され、 h_i は i 番目のバイノーラル室内インパルス応答フィルタであり、

【数97】

\tilde{h}

は共通のフィルタであり、

【数98】

40

$$E(h) = \sum_{n=0}^{N-1} h[n]^2$$

であり、 n はサンプルのインデックスであり、各 $h[n]$ は n におけるステレオサンプルである。

【0331】

[0343] いくつかの例では、デバイスは、

【数 9 9】

$$\sum \hat{w}_i in_i \odot \tilde{h}$$

を計算することによって、変換された合計音声信号を生成するために共通のフィルタを合計音声信号に適用するための手段をさらに備え、

【数 1 0 0】

10

④

は畳み込み演算を示し、 in_i は音声信号の i 番目のチャンネルを示す。

【0 3 3 2】

[0344] いくつかの例では、デバイスは、それぞれの適応的重み要素をチャンネルに適用することによって合計音声信号を生成するために音声信号のチャンネルを結合するための手段をさらに備え、

【数 1 0 1】

20

$$in_{mix}(n) = w_{norm}(n) \sum \hat{w}_i in_i(n)$$

を計算することを備え、 $in_{mix}(n)$ は合計音声信号を示し、 n はサンプルのインデックスであり、

【数 1 0 2】

30

$$w_{norm}(n) = \sqrt{\frac{\sum E(\hat{w}_i in_i)}{E(\sum \hat{w}_i in_i)}}$$

であり、 in_i は音声信号の i 番目のチャンネルを示す。

【0 3 3 3】

[0345] いくつかの例では、音声信号のチャンネルは、複数の階層要素を備える。

40

【0 3 3 4】

[0346] いくつかの例では、複数の階層要素は、球面調和係数を備える。

【0 3 3 5】

[0347] いくつかの例では、複数の階層要素は、高次アンビソニックスを備える。

【0 3 3 6】

[0348] 別の例では、非一時的コンピュータ可読記憶媒体は、実行されると、1つまたは複数のプロセッサに、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを取得することと、変換された合計音声信号を生成するために音声信号の複数のチャンネルから決定された合計音声信号に共通のフィルタを適用することと、を行わせる命令をその上に記憶している。

50

【0337】

[0349]別の例では、非一時的コンピュータ可読記憶媒体は、実行されると、1つまたは複数のプロセッサに、バイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従って重み付けられた複数のバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを生成させる命令をその上に記憶している。

【0338】

[0350]別の例では、非一時的コンピュータ可読記憶媒体は、実行されると、1つまたは複数のプロセッサに、複数のバイノーラル室内インパルス応答フィルタのサブグループの反射セグメントに関する共通のフィルタを生成させる命令をその上に記憶している。

【0339】

[0351]別の例では、非一時的コンピュータ可読記憶媒体は、実行されると、1つまたは複数のプロセッサに、複数のバイノーラル室内インパルス応答フィルタの1つまたは複数のセグメントを適用する前に、適応的に決定された重みを音声信号の複数のチャンネルに適用することと、1つまたは複数のセグメントを複数のバイノーラル室内インパルス応答フィルタに適用することと、を行わせる命令をその上に記憶している。

【0340】

[0352]別の例では、デバイスは、任意の組合せを実施するように構成されたプロセッサを備え、例の任意の組合せの方法が上記を説明した。

【0341】

[0353]別の例では、デバイスは、上記で説明した例のうちの任意の結合結合の方法の各ステップを実施するための手段を備える。

【0342】

[0354]別の例では、非一時的コンピュータ可読記憶媒体は、実行されると、1つまたは複数のプロセッサに、上記で説明した例のうちの任意の結合の方法を実施させる命令をその上に記憶している。

【0343】

[0355]例に応じて、本明細書で説明された方法のいずれのある行為またはイベントは、異なる順序で実行可能であり、追加されてもよいし、マージされてもよいし、全体的に除外されてもよい（たとえば、すべての説明された行為またはイベントが方法の実施に必要とは限らない）ことを理解されたい。その上、ある例では、行為またはイベントは、たとえば、マルチスレッド処理、割込み処理、または複数のプロセッサによって、順次ではなく、同時に実行されることがある。さらに、本開示のある態様は、わかりやすいように、単一のデバイス、モジュール、またはユニットによって実行されると説明されているが、本開示の技法は、デバイス、ユニット、またはモジュールの組合せによって実行されてよいことを理解されたい。

【0344】

[0356]1つまたは複数の例では、説明された機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せで実施されてよい。ソフトウェアで実施される場合、これらの機能は、コンピュータ可読媒体上に1つまたは複数の命令またはコードとして記憶または送信され、ハードウェアベースの処理ユニットによって実行されてもよい。コンピュータ可読媒体は、たとえば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を支援する任意の媒体を含む、データ記憶媒体または通信媒体などの有形媒体に対応するコンピュータ可読記憶媒体を含み得る。

【0345】

[0357]このようにして、コンピュータ可読媒体は、一般に、（1）非一時的である有形のコンピュータ可読記憶媒体または（2）信号もしくはキャリア波などの通信媒体に相当し得る。データ記憶媒体は、本開示で説明する技法の実装のために、命令、コードおよび/またはデータ構造を取り出すために1つもしくは複数のコンピュータまたは1つもしくは複数のプロセッサによってアクセスされ得る、任意の利用可能な媒体であり得る。コンピュータプログラム製品は、コンピュータ可読媒体を含んでもよい。

10

20

30

40

50

【0346】

[0358]例として、それに限定されず、そのようなコンピュータ可読記憶媒体は、RAM、ROM、EEPROM（登録商標）、CD-ROMもしくは他の光ディスク記憶装置、磁気ディスク記憶装置もしくは他の磁気記憶デバイス、フラッシュメモリ、または命令またはデータ構造の形態で所望のプログラムコードを記憶するために使用可能であり、コンピュータによってアクセス可能な他の任意の媒体を備えることができる。さらに、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、ウェブサイト、サーバ、または他の遠隔ソースから、同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者回線（DSL）、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して伝送される場合、同軸ケーブル、光ファイバケーブル、ツイストペア、DSL、または赤外線、無線、マイクロ波などのワイヤレス技術は、媒体の定義に含まれる。

【0347】

[0359]ただし、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時的媒体を含まず、代わりに、非一時的な有形記憶媒体を対象とすることを理解されたい。本明細書で使用するディスク（disk）およびディスク（disc）は、コンパクトディスク（disc）（CD）、レーザーディスク（登録商標）（disc）、光ディスク（disc）、デジタル多用途ディスク（disc）（DVD）、フロッピー（登録商標）ディスク（disk）およびblue-ray（登録商標）ディスク（disc）を含み、ディスク（disk）は、通常、データを磁気的に再生し、ディスク（disc）は、データをレーザーで光学的に再生する。上述の組合せもコンピュータ可読媒体の範囲内に含まれるべきである。

【0348】

[0360]命令は、1つまたは複数のデジタルシグナルプロセッサ（DSP）、汎用マイクロプロセッサ、特定用途向け集積回路（ASIC）、フィールドプログラマブルロジックアレイ（FPGA）、または他の同等の統合された、もしくは個別の論理回路などの、1つまたは複数のプロセッサによって実行され得る。したがって、「プロセッサ」という用語は、本明細書において、前述の構造のうちの任意のものまたは本明細書に記載される技法の実施のために適当な任意の他の構造を参照し得る。加えて、いくつかの態様では、本明細書に記載される機能性は、符号化および復号のために構成され、または組み合わされたコードックに組み込まれる、専用のハードウェア内および/またはソフトウェアモジュール内で提供され得る。また、技法は、1つまたは複数の回路または論理素子内で完全に実施されてよい。

【0349】

[0361]本開示の技法は、ワイヤレスハンドセット、集積回路（IC）またはICのセット（たとえば、チップセット）を含む、多種多様なデバイスまたは装置で実装され得る。様々な構成要素、モジュール、またはユニットは、開示された技法を実行するように構成されるデバイスの機能上の態様を強調するために、本開示に記載されるが、必ずしも異なるハードウェアユニットによる実現を求めるとは限らない。むしろ、上記で説明したように、様々なユニットは、コードックハードウェアユニットの中で組み合わされ、または、上記で説明した1つまたは複数のプロセッサを含む、適切なソフトウェアおよび/またはファームウェアと一緒に相互作用するハードウェアユニットの集合によって提供され得る。

【0350】

[0362]本技法の様々な実施形態が説明された。これらおよび他の実施形態は以下の特許請求の範囲内に入る。

以下に、出願当初の特許請求の範囲に記載された発明を付記する。

[C1]

音声信号をバイノーラル化する方法であって、

前記音声信号の複数の適応的に重み付けられたチャンネルを生成するために、適応的に決定された重みを前記音声信号の複数のチャンネルに適用することと、

10

20

30

40

50

結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも2つを結合することと、
バイノーラル音声信号を生成するために、バイノーラル室内インパルス応答フィルタを前記結合された信号に適用することと、
を備える、方法。

[C 2]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを備える、C 1に記載の方法。

10

[C 3]

前記少なくとも2つのバイノーラル室内インパルス応答フィルタの反射セグメントが、前記少なくとも2つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けられる、C 2に記載の方法。

[C 4]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタの反射セグメントに関する共通のフィルタを備える、C 1に記載の方法。

20

[C 5]

前記少なくとも2つのバイノーラル室内インパルス応答フィルタの前記反射セグメントが、前記少なくとも2つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けられる、C 4に記載の方法。

[C 6]

前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つが、第1のサブグループを備え、

前記結合された信号が、第1の結合された信号を備え、

前記バイノーラル室内インパルス応答フィルタが、第1のバイノーラル室内インパルス応答フィルタを備え、

前記バイノーラル音声信号が、第1のバイノーラル音声信号を備え、前記方法が、

30

第2の結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも2つを備える第2のサブグループを結合することと、

第2のバイノーラル音声信号を生成するために、第2のバイノーラル室内インパルス応答フィルタを前記第2の結合された信号に適用することと、

第3のバイノーラル音声信号を生成するために、前記第1のバイノーラル音声信号と前記第2のバイノーラル音声信号とを結合することと、をさらに備える、C 1に記載の方法。

。

[C 7]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記方法が、

前記共通のフィルタを生成するために、前記少なくとも2つのバイノーラル室内インパルス応答フィルタを正規化することなく、前記少なくとも2つのバイノーラル室内インパルス応答フィルタの平均を計算することをさらに備える、C 1に記載の方法。

[C 8]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記方法が、

前記少なくとも2つのバイノーラル室内インパルス応答フィルタの各々に関するそれぞ

40

50

れの周波数依存性両耳間コヒーレンス値を計算することと、

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算することと、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記共通のフィルタを合成することと、

をさらに備える、C 1 に記載の方法。

[C 9]

前記音声信号の複数のチャンネルに関する初期の適応的に決定された重みが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従つて決定される、C 1 に記載の方法。

10

[C 10]

前記音声信号の前記複数のチャンネルが、それぞれ、球面調和係数を備える、C 1 に記載の方法。

[C 11]

1つまたは複数のプロセッサを備えるデバイスであって、

音声信号の複数の適応的に重み付けられたチャンネルを生成するために、適応的に決定された重みを前記音声信号の複数のチャンネルに適用することと、

結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを結合することと、

20

バイノーラル音声信号を生成するために、バイノーラル室内インパルス応答フィルタを前記結合された信号に適用することと、

を行うように構成される、デバイス。

[C 12]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを備える、C 1 1 に記載のデバイス。

30

[C 13]

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの反射セグメントが、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けられる、C 1 2 に記載のデバイス。

[C 14]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタの反射セグメントに関する共通のフィルタを備える、C 1 1 に記載のデバイス。

[C 15]

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの前記反射セグメントが、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けられる、C 1 4 に記載のデバイス。

40

[C 16]

前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つが、第 1 のサブグループを備え、

前記結合された信号が、第 1 の結合された信号を備え、

前記バイノーラル室内インパルス応答フィルタが、第 1 のバイノーラル室内インパルス応答フィルタを備え、

前記バイノーラル音声信号が、第 1 のバイノーラル音声信号を備え、前記 1 つまたは複数のプロセッサが、

50

第2の結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも2つを備える第2のサブグループを結合することと、

第2のバイノーラル音声信号を生成するために、第2のバイノーラル室内インパルス応答フィルタを前記第2の結合された信号に適用することと、

第3のバイノーラル音声信号を生成するために、前記第1のバイノーラル音声信号と前記第2のバイノーラル音声信号とを結合することと、

を行うようにさらに構成される、C 1 1に記載のデバイス。

[C 1 7]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記1つまたは複数のブロセッサが、

前記共通のフィルタを生成するために、前記少なくとも2つのバイノーラル室内インパルス応答フィルタを正規化することなく、前記少なくとも2つのバイノーラル室内インパルス応答フィルタの平均を計算するようにさらに構成される、C 1 1に記載のデバイス。

[C 1 8]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記1つまたは複数のブロセッサが、

前記少なくとも2つのバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算することと、

前記少なくとも2つのバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算することと、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記共通のフィルタを合成することと、

を行うようにさらに構成される、C 1 1に記載のデバイス。

[C 1 9]

前記音声信号の複数のチャンネルに関する初期の適応的に決定された重みが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従つて決定される、C 1 1に記載のデバイス。

[C 2 0]

前記音声信号の前記複数のチャンネルが、それぞれ、球面調和係数を備える、C 1 1に記載のデバイス。

[C 2 1]

音声信号の複数の適応的に重み付けられたチャンネルを生成するために、適応的に決定された重みを前記音声信号の複数のチャンネルに適用するための手段と、

結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも2つを結合するための手段と、

バイノーラル音声信号を生成するために、バイノーラル室内インパルス応答フィルタを前記結合された信号に適用するための手段と、

を備える、装置。

[C 2 2]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも2つにそれぞれ対応する少なくとも2つのバイノーラル室内インパルス応答フィルタの残響セグメントに関する共通のフィルタを備える、C 2 1に記載の装置。

10

20

30

40

50

[C 2 3]

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの反射セグメントが、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けられる、C 2 2 に記載の装置。

[C 2 4]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタの反射セグメントに関する共通のフィルタを備える、C 2 1 に記載の装置。

[C 2 5]

10

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの前記反射セグメントが、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの少なくとも一部分のそれぞれのエネルギーに従って重み付けされる、C 2 4 に記載の装置。

[C 2 6]

前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つが、第 1 のサブグループを備え、

前記結合された信号が、第 1 の結合された信号を備え、

前記バイノーラル室内インパルス応答フィルタが、第 1 のバイノーラル室内インパルス応答フィルタを備え、

前記バイノーラル音声信号が、第 1 のバイノーラル音声信号を備え、前記装置が、

20

第 2 の結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを備える第 2 のサブグループを結合するための手段と、

第 2 のバイノーラル音声信号を生成するために、第 2 のバイノーラル室内インパルス応答フィルタを前記第 2 の結合された信号に適用するための手段と、

第 3 のバイノーラル音声信号を生成するために、前記第 1 のバイノーラル音声信号と前記第 2 のバイノーラル音声信号とを結合するための手段と、

をさらに備える、C 2 1 に記載の装置。

[C 2 7]

30

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記装置が、

前記共通のフィルタを生成するために、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタを正規化することなく、前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの平均を計算するための手段をさらに備える、C 2 1 に記載の装置。

[C 2 8]

前記バイノーラル室内インパルス応答フィルタが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少なくとも 2 つのバイノーラル室内インパルス応答フィルタに関する共通のフィルタを備え、前記装置が、

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタの各々に関するそれぞれの周波数依存性両耳間コヒーレンス値を計算するための手段と、

40

前記少なくとも 2 つのバイノーラル室内インパルス応答フィルタに関する前記それぞれの周波数依存性両耳間コヒーレンス値の平均周波数依存性両耳間コヒーレンス値を計算するための手段と、

前記平均周波数依存性両耳間コヒーレンス値を使用して前記共通のフィルタを合成するための手段と、

をさらに備える、C 2 1 に記載の装置。

[C 2 9]

前記音声信号の複数のチャンネルに関する初期の適応的に決定された重みが、前記複数の適応的に重み付けられたチャンネルのうちの前記少なくとも 2 つにそれぞれ対応する少

50

なくとも 2 つのバイノーラル室内インパルス応答フィルタのそれぞれのエネルギーに従つて決定される、C 2 1 に記載の装置。

【 C 3 0 】

実行されると、1 つまたは複数のプロセッサに、

音声信号の複数の適応的に重み付けられたチャンネルを生成するために、適応的に決定された重みを前記音声信号の複数のチャンネルに適用することと、

結合された信号を生成するために、前記音声信号の前記複数の適応的に重み付けられたチャンネルのうちの少なくとも 2 つを結合することと、

バイノーラル音声信号を生成するために、バイノーラル室内インパルス応答フィルタを前記結合された信号に適用することと、

を行わせる命令をその上に記憶した、非一時的コンピュータ可読記憶媒体。

10

【図 1】

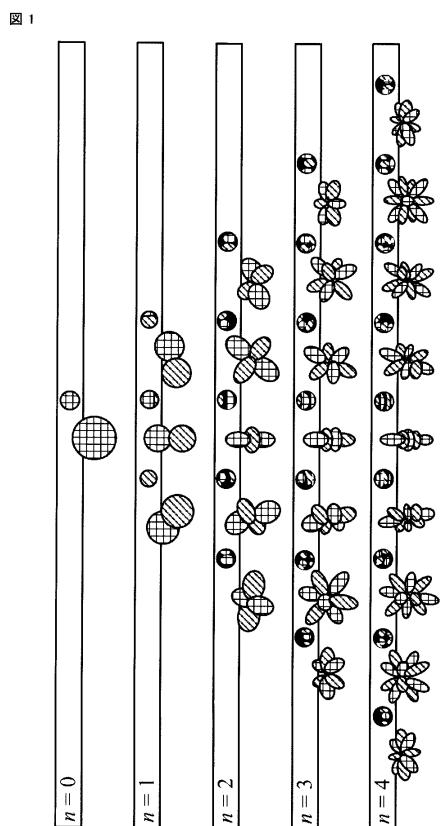


FIG. 1

【図 2】

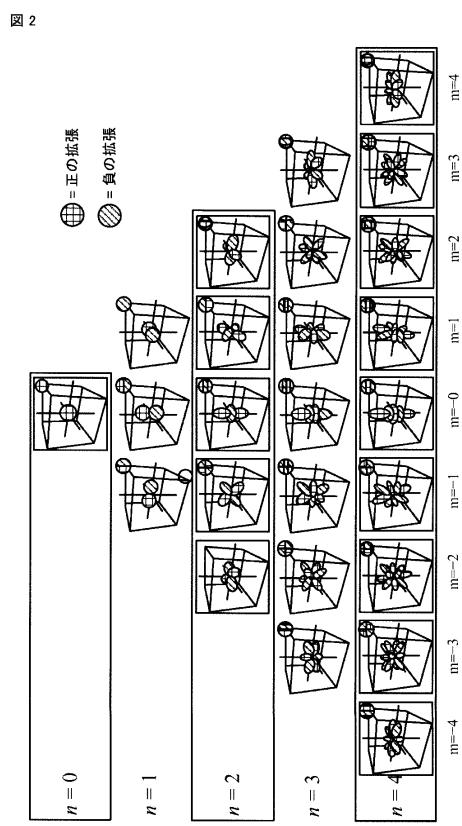


FIG. 2

【図3】

図3

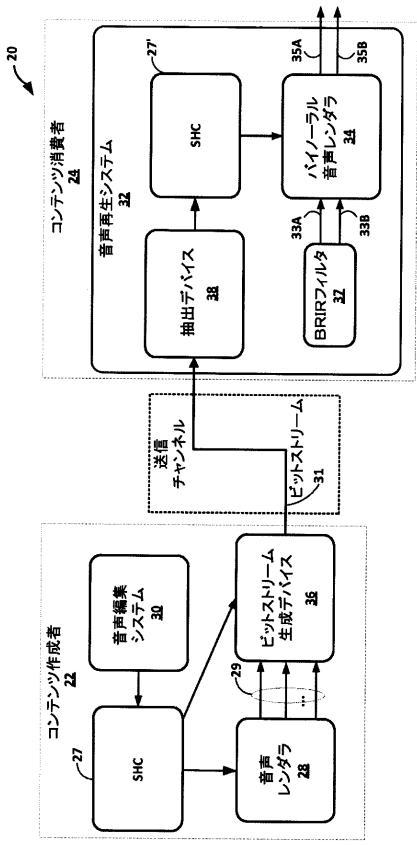


FIG. 3

【図4】

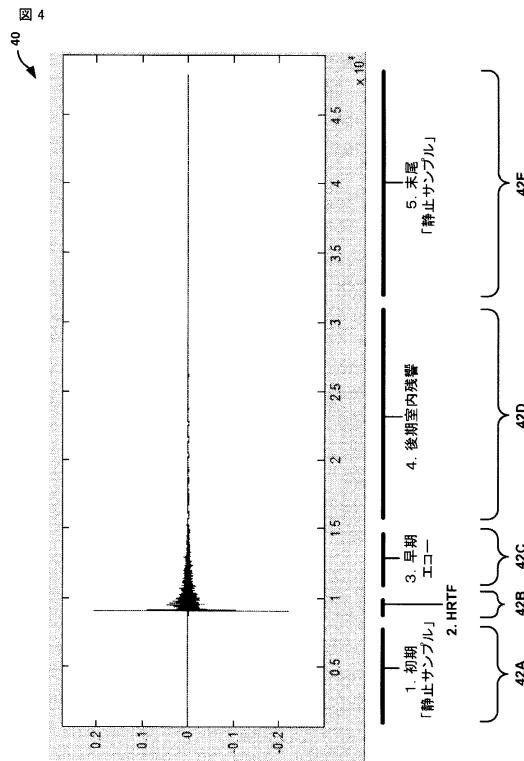


FIG. 4

【図5】

図5

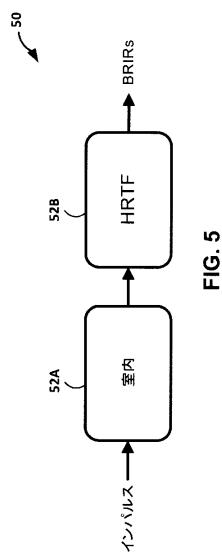


FIG. 5

【図6】

図6

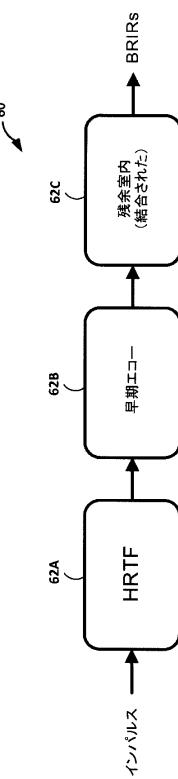
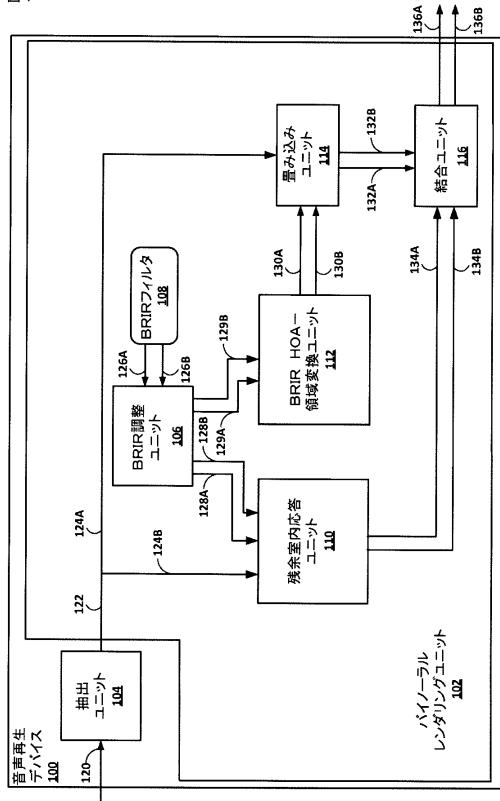


FIG. 6

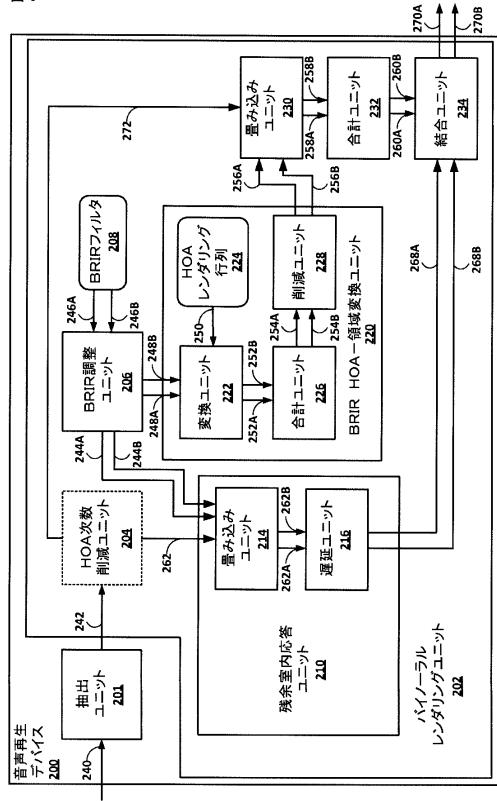
〔 図 7 〕

图 7



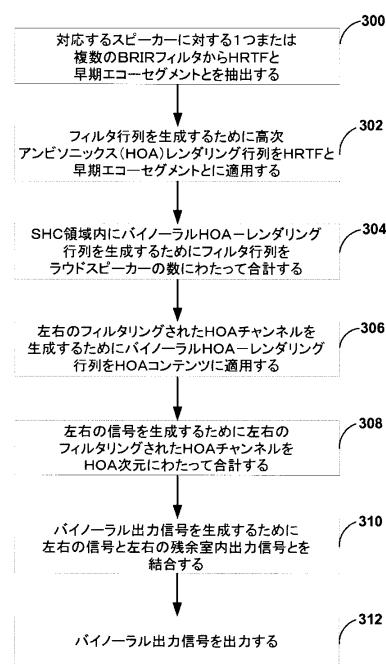
【 図 8 】

图 8



【図9】

图 9



【図10A】

10A

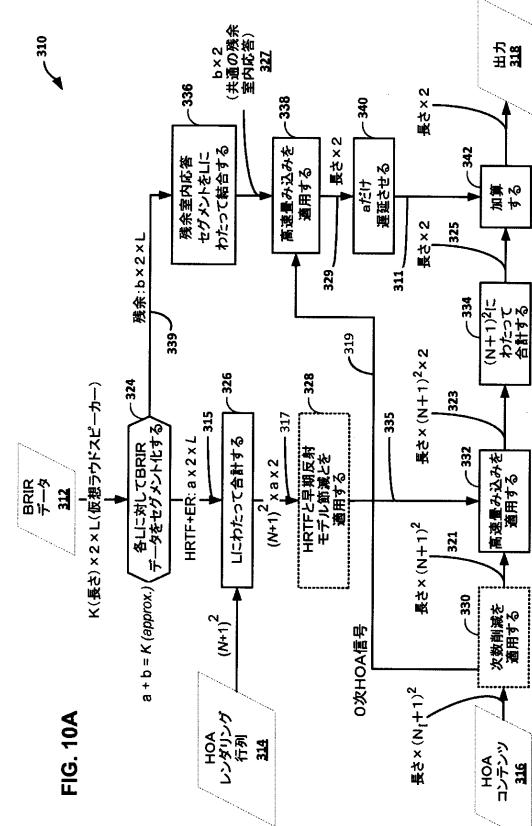
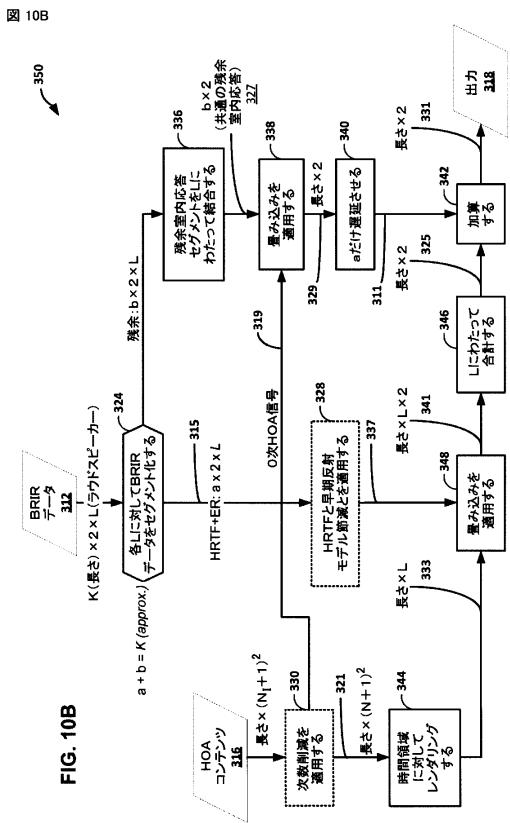


FIG. 9

【図 10B】



【図 11】

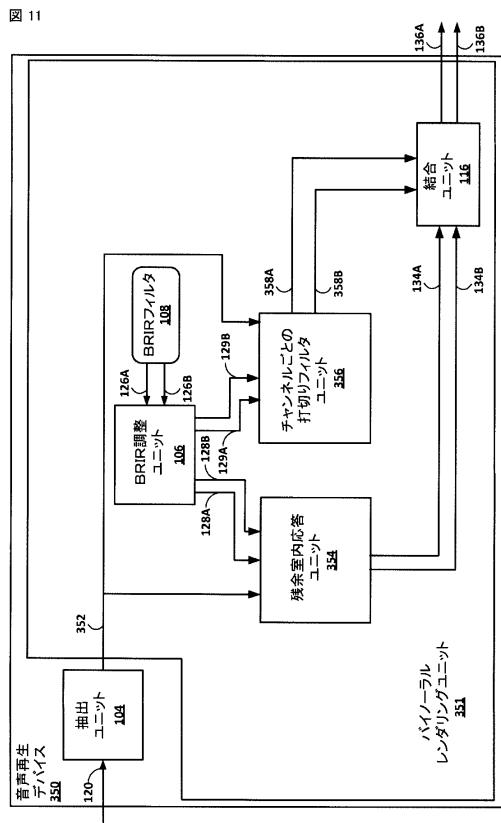


FIG. 11

【図 12】

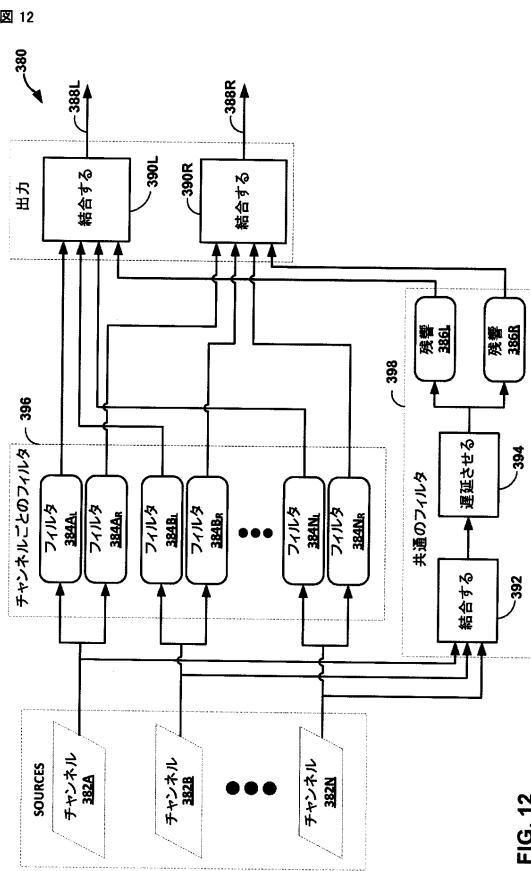


FIG. 12

【図 13】

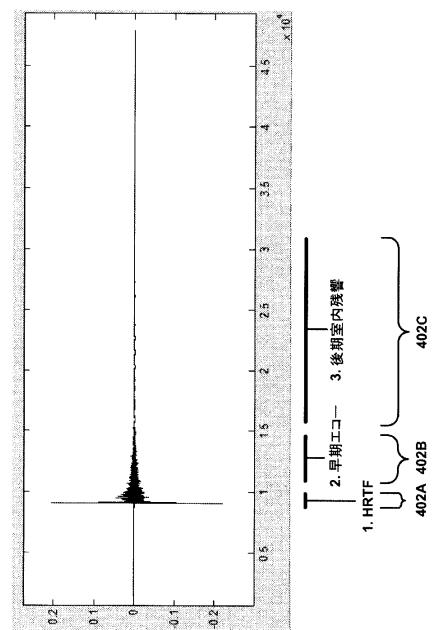
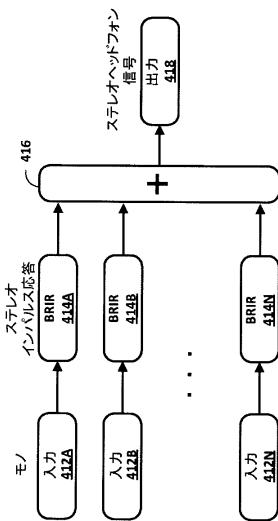


FIG. 13

【図14】

図14

410



【図15】

図15

500

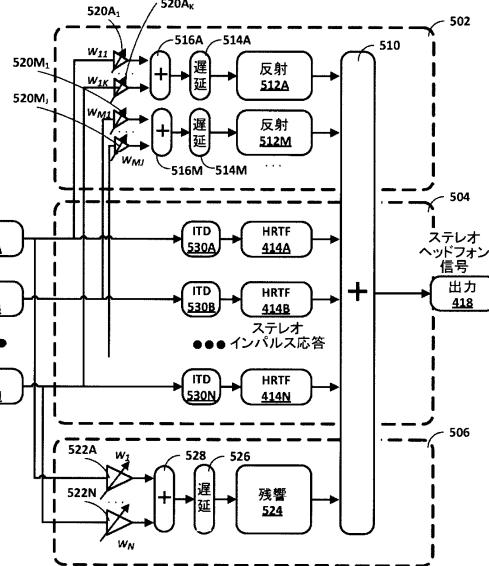
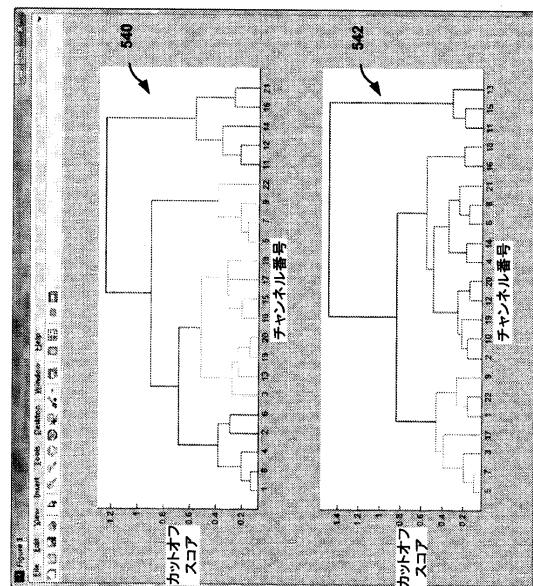


FIG. 15

【図16】

図16



【図17】

図17

FIG. 16

FIG. 17

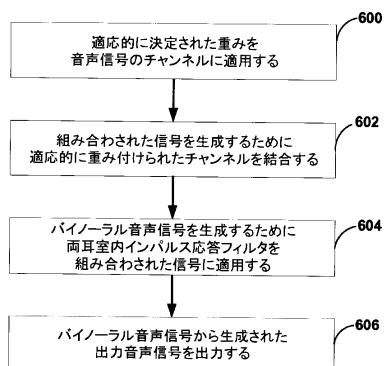


FIG. 17

フロントページの続き

(31)優先権主張番号 61/886,593
(32)優先日 平成25年10月3日(2013.10.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/886,620
(32)優先日 平成25年10月3日(2013.10.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 14/288,277
(32)優先日 平成26年5月27日(2014.5.27)
(33)優先権主張国 米国(US)

早期審査対象出願

(72)発明者 シヤン、ペイ
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(72)発明者 セン、ディパンジャン
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

審査官 大石 剛

(56)参考文献 特開2011-066868(JP, A)
特表2010-508545(JP, A)
米国特許出願公開第2009/0292544(US, A1)

(58)調査した分野(Int.Cl., DB名)

H04S 1/00
H04S 7/00