

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7229925号

(P7229925)

(45)発行日 令和5年2月28日(2023.2.28)

(24)登録日 令和5年2月17日(2023.2.17)

(51)国際特許分類

H 0 4 S 7/00 (2006.01)

F I

H 0 4 S 7/00 3 2 0

請求項の数 22 (全32頁)

(21)出願番号	特願2019-537888(P2019-537888)	(73)特許権者	315002955
(86)(22)出願日	平成29年9月14日(2017.9.14)		ノキア テクノロジーズ オーユー
(65)公表番号	特表2019-535216(P2019-535216 A)		フィンランド共和国 0 2 6 1 0 エスポーカラカーリ 7
(43)公表日	令和1年12月5日(2019.12.5)	(74)代理人	100127188
(86)国際出願番号	PCT/FI2017/050655		弁理士 川守田 光紀
(87)国際公開番号	WO2018/060549	(72)発明者	マキネン ヨーマ
(87)国際公開日	平成30年4月5日(2018.4.5)		フィンランド共和国 3 3 5 6 0 タンペレ ローツンカツ 4
審査請求日	令和1年5月9日(2019.5.9)	(72)発明者	タンミ ミッコ
審判番号	不服2022-2503(P2022-2503/J1)		フィンランド共和国 3 3 3 1 0 タンペレ コハマンカーリ 1 C 1 3
審判請求日	令和4年2月21日(2022.2.21)	(72)発明者	ライティネン ミッコ・ヴィッレ
(31)優先権主張番号	1616479.0		フィンランド共和国 0 0 2 1 0 ヘルシンキ メルコンカツ 9 A 6
(32)優先日	平成28年9月28日(2016.9.28)		最終頁に続く
(33)優先権主張国・地域又は機関	英国(GB)		

(54)【発明の名称】 空間オーディオシステムにおける利得制御

(57)【特許請求の範囲】

【請求項1】

基準位置に対して少なくとも1つの音源の位置及び／又は向きを設ける手段と、

前記少なくとも1つの音源に対応付けられた少なくとも1つの入力オーディオ信号を分析し、前記少なくとも1つの入力オーディオ信号のための少なくとも1つの最大入力レベルを推定する手段と、

ユーザの位置及び／又は向きを決定する手段と、

前記少なくとも1つの音源の位置及び／又は向きと、前記ユーザの位置及び／又は向きとの間の差を決定する手段と、

前記少なくとも1つの最大入力レベルおよび前記差に基づいて、少なくとも1つの利得値を決定する手段であって、前記少なくとも1つの利得値は、前記少なくとも1つの入力オーディオ信号に対応付けられた第1周波数帯域のための、第1周波数帯域用レベル制御利得を含む、前記決定する手段と、

前記少なくとも1つの入力オーディオ信号と、方向伝達関数対と、前記少なくとも1つの利得値と、前記少なくとも1つの音源の位置及び／又は向きに少なくとも部分的に基づいて、少なくとも2つの出力チャネルを合成する手段と、
を備える装置。

【請求項2】

前記設けることは、

少なくとも2つのマイクオーディオ信号を分析し、前記少なくとも1つの音源の位置及

10

20

び／又は向きを決定することと、

前記少なくとも1つの入力オーディオ信号に対応付けられたメタデータから、前記少なくとも1つの音源の位置及び／又は向きを取得することと、
の少なくともいずれかを含み、前記メタデータは、ある周波数帯域に対する方向パラメータを含む、

請求項1に記載の装置。

【請求項3】

前記基準位置に対する追跡部の位置および／または向き
を決定するようにさらに構成され、前記ユーザの位置及び／又は向きを決定することは、
前記追跡部の位置及び／又は向きを決定することを含む、請求項1に記載の装置。

10

【請求項4】

前記少なくとも1つの利得値を、前記少なくとも2つの出力チャンネルのそれぞれの最大出力レベルに基づいて決定するように構成される、請求項1に記載の装置。

【請求項5】

前記少なくとも1つの入力オーディオ信号と、前記方向伝達関数対と、前記第1周波数帯域用レベル制御利得と、前記差に基づいて、前記少なくとも2つの出力チャンネルを合成するように構成された、請求項4に記載の装置。

【請求項6】

デジタルコンパスからの方向、

ジャイロスコープからの方向、

ビーコン測位システムからの方向、

前記位置が未来の実際の追跡部位置を推定したものとなるように慣性および質量でモデル化された方向、

の内の少なくとも1つを受信するように構成される、請求項3に記載の装置。

20

【請求項7】

前記入力オーディオ信号に対応付けられた第2周波数帯域に対する、第2周波数帯域用レベル制御利得；

第1出力チャンネルに対する、第1出力チャンネル用レベル制御利得；

第2出力チャンネルに対する、第2出力チャンネル用レベル制御利得；

の少なくともいずれかを、前記差に基づいて決定するように構成された、請求項1に記載の装置。

30

【請求項8】

前記追跡部は、前記追跡部に対応付けられた方向を生成するように構成されたヘッドバンドを備え、前記方向はユーザの聴き取り方向に対応付けられている、請求項3に記載の装置。

【請求項9】

前記ユーザの位置及び／又は向きは、前記装置の少なくとも一部の位置及び／又は向きを含み、

前記少なくとも2つの出力チャンネルの出力信号値の少なくともいずれかを制限し、それによって、該出力信号値が定義されたパルス符号変調方式のダイナミックレンジに含まれるようにするようにさらに構成された、請求項1に記載の装置。

40

【請求項10】

前記基準位置は、

少なくとも1つの空間的に取得される音源に対する原点位置および／または向き、

前記少なくとも1つの入力オーディオ信号をキャプチャするマイクアレイの位置及び／又は向き、

カメラに対する原点位置および／または向き、

前記少なくとも1つの音源が存在する音場面に対する原点位置および／または向き、
の内の少なくとも1つである、請求項1に記載の装置。

【請求項11】

50

前記少なくとも2つの出力チャンネルを合成することは、

前記差を用いる第1頭部伝達関数と、モデル化された左耳位置の組合せに、左レベル制御利得値を乗算することに基づいて、左出力チャンネルオーディオ信号を生成することと、

前記差を用いる第2頭部伝達関数と、モデル化された右耳位置の組合せに、右レベル制御利得値を乗算することに基づいて、右出力チャンネルオーディオ信号を生成することと、を含む、請求項5に記載の装置。

【請求項12】

前記少なくとも1つの利得値は、前記少なくとも1つの入力オーディオ信号が、少なくとも1つの推定最大入力レベルに達するレベルを有している場合に、前記少なくとも2つの出力チャンネルのそれぞれのレベルが、該少なくとも2つの出力チャンネルのそれぞれの最大出力レベルを超えないように制御するように構成される、請求項4に記載の装置。

10

【請求項13】

基準位置に対して少なくとも1つの音源の位置及び／又は向きを設けることと、

前記少なくとも1つの音源に対応付けられた少なくとも1つの入力オーディオ信号を分析し、前記少なくとも1つの入力オーディオ信号のための少なくとも1つの最大入力レベルを推定することと、

ユーザの位置及び／又は向きを決定することと、

前記少なくとも1つの音源の位置及び／又は向きと、前記ユーザの位置及び／又は向きとの間の差を決定することと、

前記少なくとも1つの最大入力レベルおよび前記差に基づいて、少なくとも1つの利得値を決定することであって、前記少なくとも1つの利得値は、前記少なくとも1つの入力オーディオ信号に対応付けられた第1周波数帯域のための、第1周波数帯域用レベル制御利得を含む、前記決定することと、

20

前記少なくとも1つの入力オーディオ信号と、方向伝達関数対と、前記少なくとも1つの利得値と、前記少なくとも1つの音源の位置及び／又は向きに少なくとも部分的に基づいて、少なくとも2つの出力チャンネルを合成することと、を含む方法。

【請求項14】

前記設けることは、

少なくとも2つのマイクオーディオ信号を分析し、前記少なくとも1つの音源の位置及び／又は向きを決定することと、

前記少なくとも1つの入力オーディオ信号に対応付けられたメタデータから、前記少なくとも1つの音源の位置及び／又は向きを取得することと、

の少なくともいずれかを含み、前記メタデータは、ある周波数帯域に対する方向パラメータを含む、

請求項13に記載の方法。

30

【請求項15】

前記基準位置に対する追跡部の位置および／または向きを決定することを更に含み、前記ユーザの位置及び／又は向きを決定することは、前記追跡部の位置および／または向きを決定することを含む、請求項13に記載の方法。

40

【請求項16】

前記少なくとも1つの利得値を、前記少なくとも2つの出力チャンネルのそれぞれの最大出力レベルに基づいて決定することを更に含む、請求項13に記載の方法。

【請求項17】

前記少なくとも1つの入力オーディオ信号と、前記方向伝達関数対と、前記第1周波数帯域用レベル制御利得と、前記差に基づいて、前記少なくとも2つの出力チャンネルを合成することを更に含む、請求項16に記載の方法。

【請求項18】

前記入力オーディオ信号に対応付けられた第2周波数帯域に対する、第2周波数帯域用レベル制御利得；

50

第1出力チャンネルに対する、第1出力チャンネル用レベル制御利得；
第2出力チャンネルに対する、第2出力チャンネル用レベル制御利得；
の少なくともいずれかを、前記差に基づいて決定することを更に含む、請求項16に記載の方法。

【請求項19】

前記少なくとも2つの出力チャンネルを合成することは、

前記差を用いる第1頭部伝達関数と、モデル化された左耳位置の組合せに、左レベル制御利得値を乗算することに基づいて、左出力チャンネルオーディオ信号を生成することと、

前記差を用いる第2頭部伝達関数と、モデル化された右耳位置の組合せに、右レベル制御利得値を乗算することに基づいて、右出力チャンネルオーディオ信号を生成することと、を含む、請求項17に記載の方法。

10

【請求項20】

前記少なくとも1つの利得値は、前記少なくとも1つの入力オーディオ信号が、少なくとも1つの推定最大入力レベルに達するレベルを有している場合に、前記少なくとも2つの出力チャンネルのそれぞれのレベルが、該少なくとも2つの出力チャンネルのそれぞれの最大出力レベルを超えないように制御するように構成される、請求項17に記載の方法。

【請求項21】

処理手段及び記憶手段を備える装置であって、前記記憶手段はプログラム命令を格納し、前記プログラム命令は、前記処理手段に実行されると、前記装置に、請求項13から20のいずれかに記載の方法を遂行させるように構成される、装置。

20

【請求項22】

装置の処理手段に実行されると、前記装置に、請求項13から20のいずれかに記載の方法を遂行させるように構成されるプログラム命令を備える、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、頭部追跡入力を利用したオーディオシステムにおいて、利得制御を実現する装置および方法に関する。本発明はさらに、オーディオ信号の空間的再生を実現可能とすべく、頭部追跡入力を利用した空間オーディオシステムにおいて、利得制御を実現する装置および方法に関するが、これに限定されるものではない。

30

【背景】

【0002】

空間場において、多数の信号源からオーディオ信号を取得し、提供するには、莫大な処理労力が伴う。例えば、劇場や講堂等の音響環境における話し手や演者のようなオーディオ信号源からのキャプチャ、ミキシングを実行して、聴き手に届け、効果的な音響空間を作り出すのに必要な処理は並大抵ではない。

【0003】

一般的に用いられているシステムでは、近接または外部マイクが利用される。例えば、ラベリアマイクをユーザに装着したり、オーディオ信号を取得するため、ブームポールに取り付けたマイクを話し手、または空間オーディオ信号外のその他の信号源に近付けたりすることが行われている。これら信号は、マイクアレイから取得された空間（あるいは環境または音響場）オーディオ信号とミキシングされて、決定されたユーザからの方向入力に基づいて、ユーザに提供されうる。

40

【0004】

録音状況に応じて、取得されるオーディオコンテンツのレベルやダイナミクスは大いに変動する。典型的には、遠くまたは音の小さな音源からの聴き取りを容易にするためには、レンダリングまたはミキシングされたオーディオ信号ヘッドルーム、すなわちオーディオ信号が、クリッピングにより再生システムまたはオーディオ信号を損傷することなく一時的に「最大レベル」を超越しうる量を最小限に抑えるべきである。これは、音を聴く環境で大きい騒音や、ヘッドホンやスピーカ出力に関する音圧制限を伴う、携帯型プレーヤ

50

に関して特に重要である。

【0005】

典型的には、頭部追跡システムや、ヘッドホンを利用して音を聴いているユーザは、その頭部の動きに応じて動くミキシングされたオーディオ信号または音景が、バイノーラル (binaural) 形式にレンダリングされる。すなわち、音を聴くユーザの位置が、レンダリングされた (バイノーラル) オーディオ信号のレベルに影響を及ぼす (一方で、典型的には向きのみが考慮される)。例えば、耳が直接音源に向いているため高周波数が増幅されるか、耳が音源を向いていないため高周波数が減衰するかにより、出力レベルが変化しうる。オーディオ信号におけるこの向きまたは位置の変化に対応可能として、クリッピングを防止するためには、オーディオコンテンツにさらなるヘッドルームが必要となるが、これは望ましくない。

10

【0006】

さらに、オーディオ信号の「音量」を調整するための従来の自動利得制御 (Automatic Gain Control: AGC) およびダイナミックレンジ圧縮 (Dynamic Range Compression: DRC) アルゴリズムでは、オーディオ信号を処理、出力する前に、遅延をかける必要がある。これは頭部追跡入力を利用した音景のレンダリングには不都合である。音景の動きは、時間遅延を感じさせずに頭部の動きに準じるのが理想である。さもないと、空間オーディオ品質が極めて低く知覚されてしまうのである。したがって、レンダリングされたオーディオ信号に、知覚可能な遅延をもたらすようなレベル制御関数は適用できない。

【0007】

したがって、オーディオシステムにおいて、入力の複雑性と、さらなる遅延をもたらさないことの両方を考慮して利得制御を適用可能にする手法の開発が望まれている。

20

【摘要】

【0008】

本願は、聴き手の位置が、合成されたオーディオのレベルに影響しうる、空間オーディオ再生システムにおける利得制御に関する。利得制御は、利得制御関数に遅延を伴うことなく合成されたオーディオが、位置の変化に反応できるようにするものである。これは、音源と聴き手の位置についての (推定された、または既知の) 情報と、入力の特徴、さらに、空間オーディオ再生システムを利用してレンダリングされたバイノーラル出力信号のレベルを推定し、利得を適用することで実現された。

30

【0009】

第1の態様によると装置が提供され、該装置は、基準位置に対して少なくとも1つの音源の位置を設け、前記少なくとも1つの音源に対応付けられた少なくとも1つの入力オーディオ信号を分析し、前記少なくとも1つの入力オーディオ信号および前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも1つの利得値を決定し、前記少なくとも1つの入力オーディオ信号と、方向伝達関数対と、前記少なくとも1つの利得値と、前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも2つの出力チャネルを合成するように構成されたプロセッサを備える。

【0010】

前記基準位置に対して前記少なくとも1つの音源の位置を設けるように構成された前記プロセッサは、少なくとも2つのマイクオーディオ信号を分析し、前記基準位置に対する前記音源の位置を決定するように構成されてもよい。

40

【0011】

前記基準位置に対して前記少なくとも1つの音源の位置を設けるように構成された前記プロセッサは、前記少なくとも1つの入力オーディオ信号に対応付けられたメタデータから、前記基準位置に対する前記音源の位置を取得するように構成されてもよい。

【0012】

前記少なくとも1つの入力信号に対応付けられた前記メタデータは、ある周波数帯域に対する前記メタデータの方向パラメータを含んでもよい。

【0013】

50

前記少なくとも1つの音源に対応付けられた前記少なくとも1つの入力オーディオ信号を分析するように構成された前記プロセッサは、少なくとも1つの入力オーディオ信号レベルを決定するように構成されてもよく、前記少なくとも1つの利得値は、前記少なくとも1つの入力オーディオ信号レベルおよび前記基準位置に対する前記少なくとも1つの音源の位置に基づいて決定されてもよい。

【0014】

前記プロセッサは、前記基準位置に対する追跡部の位置および／または向きを決定するようにさらに構成されてもよい。

【0015】

前記プロセッサは、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、レベル制御利得を決定するようにさらに構成されてもよい。

10

【0016】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、レベル制御利得を決定するように構成される前記プロセッサは、前記少なくとも1つの入力オーディオ信号レベルに基づいて、少なくとも1つのレベル制御利得関数を決定し、前記少なくとも1つのレベル制御利得関数に適用される、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定してもよい。

【0017】

前記プロセッサは、前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成するように構成されてもよい。

20

【0018】

前記追跡部の前記位置および／または向きを決定するように構成された前記プロセッサは、デジタルコンパスからの方向、ジャイロスコープからの方向、ビーコン測位システムからの方向、前記位置が未来の実際の追跡部位置を推定したものとなるように慣性および質量でモデル化された方向、の内の少なくとも1つを受信するように構成されてもよい。

【0019】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定するように構成されたプロセッサは、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第1の周波数帯域に対する、第1の周波数帯域用レベル制御利得を決定し、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第2の周波数帯域に対する、第2の周波数帯域用レベル制御利得を決定するように構成されてもよい。

30

【0020】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定するように構成されたプロセッサは、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第1の出力チャンネルに対する、第1の出力チャンネル用レベル制御利得を決定し、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第2の出力チャンネルに対する、第2の出力チャンネル用レベル制御利得を決定するように構成されてもよい。

40

【0021】

前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成するように構成された前記プロセッサは、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第1の頭部伝達関数と、モデル化された左耳位置の組合せに、左レベル制御利得値をかけることで、前記少なくとも2つの出力チャンネルの左

50

出力チャネルオーディオ信号を生成し、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第2の頭部伝達関数と、モデル化された右耳位置の組合せに、右レベル制御利得値をかけることで、前記少なくとも2つの出力チャネルの右出力チャネルオーディオ信号を生成するように構成されてもよい。

【0022】

前記追跡部は、前記追跡部に対応付けられた方向を生成するように構成されたヘッドバンドを備え、前記方向はユーザの聴き取り方向に対応付けられてもよい。

【0023】

前記プロセッサは、出力信号値の範囲が定義されたパルス符号変調方式のダイナミックレンジに含まれるように、前記少なくとも2つの出力チャネルの出力信号値を制限するようにさらに構成されてもよい。

【0024】

前記少なくとも2つの出力チャネルは、オーバーヘッドが最小限に抑えられていてもよい。

【0025】

前記基準位置は、少なくとも1つの空間的に取得される音源に対する原点位置および／または向きと、カメラに対する原点位置および／または向きと、前記少なくとも1つの音源が存在する音場面に対する原点位置および／または向きと、の内の少なくとも1つであってもよい。

【0026】

第2の態様によると方法が提供され、該方法は、基準位置に対して少なくとも1つの音源の位置を設けることと、前記少なくとも1つの音源に対応付けられた少なくとも1つの入力オーディオ信号を分析し、前記少なくとも1つの入力オーディオ信号および前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも1つの利得値を決定することと、前記少なくとも1つの入力オーディオ信号と、方向伝達関数対と、前記少なくとも1つの利得値と、前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも2つの出力チャネルを合成することと、を含む。

【0027】

前記基準位置に対して前記少なくとも1つの音源の位置を設けることは、少なくとも2つのマイクオーディオ信号を分析し、前記基準位置に対する前記音源の位置を決定することを含んでもよい。

【0028】

前記基準位置に対して前記少なくとも1つの音源の位置を設けることは、前記少なくとも1つの入力オーディオ信号に対応付けられたメタデータから、前記基準位置に対する前記音源の位置を取得することを含んでもよい。

【0029】

前記少なくとも1つの入力信号に対応付けられた前記メタデータは、ある周波数帯域に対する前記メタデータの方向パラメータを含んでもよい。

【0030】

前記少なくとも1つの音源に対応付けられた前記少なくとも1つの入力オーディオ信号を分析することは、少なくとも1つの入力オーディオ信号レベルを決定することを含んでもよく、前記少なくとも1つの利得値を決定することは、前記少なくとも1つの入力オーディオ信号レベルおよび前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、前記少なくとも1つの利得値を決定することを含んでもよい。

【0031】

前記方法は、前記基準位置に対する追跡部の位置および／または向きを決定することをさらに含んでもよい。

【0032】

前記方法は、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、レベル制御利得を決定することをさらに含んでもよい。

10

20

30

40

50

【0033】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定することは、前記少なくとも1つの入力オーディオ信号レベルに基づいて、少なくとも1つのレベル制御利得関数を決定することと、前記少なくとも1つのレベル制御利得関数に適用される、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定することと、を含んでもよい。

【0034】

前記方法は、前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成することをさらに含んでもよい。

10

【0035】

前記追跡部の位置および／または向きを決定することは、デジタルコンパスからの方向、ジャイロスコープからの方向、ビーコン測位システムからの方向、前記位置が未来の実際の追跡部位置を推定したものとなるように慣性および質量でモデル化された方向、の内の少なくとも1つを受信することを含んでもよい。

【0036】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定することは、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第1の周波数帯域に対する、第1の周波数帯域用レベル制御利得を決定することと、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第2の周波数帯域に対する、第2の周波数帯域用レベル制御利得を決定すること、を含んでもよい。

20

【0037】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定することは、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第1の出力チャンネルに対する、第1の出力チャンネル用レベル制御利得を決定することと、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第2の出力チャンネルに対する、第2の出力チャンネル用レベル制御利得を決定することとを含んでもよい。

30

【0038】

前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成することは、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第1の頭部伝達関数と、モデル化された左耳位置の組合せに、左レベル制御利得値をかけることで、前記少なくとも2つの出力チャンネルの左出力チャンネルオーディオ信号を生成することと、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第2の頭部伝達関数と、モデル化された右耳位置の組合せに、右レベル制御利得値をかけることで、前記少なくとも2つの出力チャンネルの右出力チャンネルオーディオ信号を生成することとを含んでもよい。

40

【0039】

前記追跡部は、前記追跡部に対応付けられた方向を生成するように構成されたヘッドバンドを備え、前記方向はユーザの聴き取り方向に対応付けられてもよい。

【0040】

前記方法は、出力信号値の範囲が定義されたパルス符号変調方式のダイナミックレンジに含まれるように、前記少なくとも2つの出力チャンネルの出力信号値を制限することをさらに含んでもよい。

50

【0041】

前記少なくとも2つの出力チャンネルは、オーバーヘッドが最小限に抑えられてもよい。

【0042】

前記基準位置は、少なくとも1つの空間的に取得される音源に対する原点位置および／または向きと、カメラに対する原点位置および／または向きと、前記少なくとも1つの音源が存在する音場面に対する原点位置および／または向きと、の内の少なくとも1つであってもよい

【0043】

第3の態様によると装置が提供され、該装置は、基準位置に対して少なくとも1つの音源の位置を設ける手段と、前記少なくとも1つの音源に対応付けられた少なくとも1つの入力オーディオ信号を分析し、前記少なくとも1つの入力オーディオ信号および前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも1つの利得値を決定する手段と、前記少なくとも1つの入力オーディオ信号と、方向伝達関数対と、前記少なくとも1つの利得値と、前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、少なくとも2つの出力チャンネルを合成する手段を備える。

10

【0044】

前記基準位置に対して前記少なくとも1つの音源の位置を設ける前記手段は、少なくとも2つのマイクオーディオ信号を分析し、前記基準位置に対する前記音源の位置を決定する手段を備えてもよい。

【0045】

前記基準位置に対して前記少なくとも1つの音源の位置を設ける前記手段は、前記少なくとも1つの入力オーディオ信号に対応付けられたメタデータから、前記基準位置に対する前記音源の位置を取得する手段を備えてもよい。

20

【0046】

前記少なくとも1つの入力信号に対応付けられた前記メタデータは、ある周波数帯域に対する前記メタデータの方向パラメータを含んでもよい。

【0047】

前記少なくとも1つの音源に対応付けられた前記少なくとも1つの入力オーディオ信号を分析する手段は、少なくとも1つの入力オーディオ信号レベルを決定する手段を備えてもよく、前記少なくとも1つの利得値を決定する手段は、前記少なくとも1つの入力オーディオ信号レベルおよび前記基準位置に対する前記少なくとも1つの音源の位置に基づいて、前記少なくとも1つの利得値を決定する手段を備えてもよい。

30

【0048】

前記装置は、前記基準位置に対する追跡部の位置および／または向きを決定する手段をさらに備えてもよい。

【0049】

前記装置は、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、レベル制御利得を決定する手段をさらに備えてもよい。

【0050】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、レベル制御利得を決定する手段は、前記少なくとも1つの入力オーディオ信号レベルに基づいて、少なくとも1つのレベル制御利得関数を決定する手段と、前記少なくとも1つのレベル制御利得関数に適用される、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定する手段とを備えてもよい。

40

【0051】

前記装置は、前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成する手段をさらに備えてもよい。

50

【0052】

前記追跡部の位置および／または向きを決定する手段は、デジタルコンパスからの方向、ジャイロ스코プからの方向、ビーコン測位システムからの方向、前記位置が未来の実際の追跡部位置を推定したものとなるように慣性および質量でモデル化された方向、の内の少なくとも1つを受信する手段を備えてもよい。

【0053】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定する手段は、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第1の周波数帯域に対する、第1の周波数帯域用レベル制御利得を決定する手段と、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記入力オーディオ信号に対応付けられた第2の周波数帯域に対する、第2の周波数帯域用レベル制御利得を決定する手段と、を備えてもよい。

10

【0054】

前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記レベル制御利得を決定する手段は、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第1の出力チャンネルに対する、第1の出力チャンネル用レベル制御利得を決定する手段と、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、第2の出力チャンネルに対する、第2の出力チャンネル用レベル制御利得を決定する手段とを備えてもよい。

20

【0055】

前記入力オーディオ信号と、前記方向伝達関数対と、前記少なくとも1つのレベル制御利得値と、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づいて、前記少なくとも2つの出力チャンネルを合成する前記手段は、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第1の頭部伝達関数と、モデル化された左耳位置の組合せに、左レベル制御利得値をかけることで、前記少なくとも2つの出力チャンネルの左出力チャンネルオーディオ信号を生成する手段と、前記入力オーディオ信号への、前記少なくとも1つの音源と前記追跡部との間の位置および／または向きの差に基づく第2の頭部伝達関数と、モデル化された右耳位置の組合せに、右レベル制御利得値をかけることで、前記少なくとも2つの出力チャンネルの右出力チャンネルオーディオ信号を生成する手段とを備えてもよい。

30

【0056】

前記追跡部は、前記追跡部に対応付けられた方向を生成するように構成されたヘッドバンドを備え、前記方向はユーザの聴き取り方向に対応付けられてもよい。

【0057】

前記装置は、出力信号値の範囲が定義されたパルス符号変調方式のダイナミックレンジに含まれるように、前記少なくとも2つの出力チャンネルの出力信号値を制限する手段をさらに備えてもよい。

【0058】

前記少なくとも2つの出力チャンネルは、オーバーヘッドが最小限に抑えられてもよい。

40

【0059】

前記基準位置は、少なくとも1つの空間的に取得される音源に対する原点位置および／または向きと、カメラに対する原点位置および／または向きと、前記少なくとも1つの音源が存在する音場面に対する原点位置および／または向きと、の内の少なくとも1つであってもよい。媒体に記憶されたコンピュータプログラム製品により、装置に本明細書に記載の方法を実行させてもよい。

【0060】

電子デバイスは、本明細書に記載の装置を含んでもよい。

【0061】

チップセットは、本明細書に記載の装置を含んでもよい。

50

【0062】

本願の実施形態は、現状の技術に関する問題を解決することを目的とする。

【図面の簡単な説明】

【0063】

本願をよりよく理解できるように、以下の添付の図を例示的に参照する。

【図1a】図1aは、いくつかの実施形態の実施に適した、分散型オーディオキャプチャおよび処理システムおよび装置を概略的に示す。

【図1b】図1bは、いくつかの実施形態の実施に適した、分散型オーディオキャプチャおよび処理システムおよび装置を概略的に示す。

【図2】図2は、マイクアレイに対する音源の例示的向きを概略的に示す。

10

【図3】図3は、頭部追跡部第1の向き、すなわち基準向きに対する音源の例示的向きを概略的に示す。

【図4】図4は、頭部追跡部第2の向き、すなわち回転向きに対する音源の例示的向きを概略的に示す。

【図5】図5は、例示的オーディオ信号に対する、第1のおよび第2の向きの間の頭部追跡部の向きの差に応じた出力レベル分布の例を概略的に示す。

【図6】図6は、例示的オーディオ信号に対する、第1のおよび第2の向きの間の頭部追跡部の向きの差に応じた出力レベル周波数分布の例を概略的に示す。

【図7】図7は、図1bに示すレベルコントローラをさらに詳細に、概略的に示す。

【図8】図8は、図7に示すレベルコントローラの効果を概略的に示す。

20

【図9】図9は、実施形態の実施に適した電子デバイスの例を示す。

【発明を実施するための形態】

【0064】

以下に、空間的に処理されたオーディオ信号と、特に追跡部からの入力に基づいてレンダリングされたオーディオ信号における、オーディオ信号の効果的な利得制御を実現するのに適した装置や、利用できる機構についてより詳細に説明する。例えば、追跡部は聴き手の頭部の向きを監視する頭部追跡部であってもよい。以下の例では、オーディオ信号とオーディオキャプチャ信号が説明される。ただし、ある実施形態において、本装置は、オーディオ信号を取得する、またはオーディオ信号とその他の情報信号を受信するように構成された、任意の適切な電子デバイスまたは装置の一部であってもよいことが理解されよう。

30

【0065】

上述したように、利得制御に対する従来の手法は、遅延線、利得（またはダイナミックレンジ）コントローラおよび乗算器により、音量制御を実現しうるものである。利得またはダイナミックレンジコントローラは、入力信号の振幅を決定し、この入力信号が遅延線を通過する際に、信号に適用される利得を計算し、その後その利得を乗算器により適用して、出力信号の振幅を調整してもよい。入力信号の振幅が所定の閾値を超えると、利得コントローラは、出力信号の振幅を制限して、所定の閾値未満に抑える。マルチバンド利得またはダイナミックレンジコントローラシステムの実現により、これを向上できうる。該システムは、それぞれ既定の周波数帯に対する利得を制御する複数のコントローラを用いて、閾値および利得を制御することにより、出力信号の全体的な音量を上げることができる。周波数に依存した制御により、1つの周波数帯域に入力電力が集中すると、出力信号の音量が上がりうる。ただし、これにより通常、フィルタリングまたは周波数分割工程が必要となり、これがシステムの処理遅延および計算負荷を増加させる。

40

【0066】

以下に詳細に記載される概念は、例えば空間オーディオ再生システムの一部として実施されてもよい。このシステムは、外部（スピーカ、楽器、またはその他の音源）オーディオ信号と、空間（音響場）オーディオ信号の両方を取得するように構成されたキャプチャシステムの一部であってもよい。したがって、本明細書に記載の実施形態では、利得制御システムが空間オーディオ再生システムに統合されて、ヘッドルームの必要性を最小限に

50

抑え、計算の複雑さと時間遅延とを最小限に抑えるようにする。

【0067】

したがって、本明細書に記載の実施形態は、バイノーラル空間オーディオ再生方法を実現しうる。この方法は、任意の音響源の到来方向（Direction of Arrival：DOA）と、直接要素（音響源等）と周囲要素（すなわち背景）の割合を推定し、この情報と聴き手の頭部追跡部入力の向きを利用して、バイノーラル出力信号を合成するものである。

【0068】

したがって、統合された利得制御方法および装置は、音源と、（例えば頭部追跡装置からの）出力信号を聴いているユーザの、（推定された、または既知の）位置についての情報を利用して、レンダリングされたバイノーラル出力信号のレベルを推定する。このようにして、該方法および装置は、利得制御機能を、空間オーディオ処理およびレンダリングシステムに統合しうる。

【0069】

これによりもたらされる効果は主に2つ考えられる。まず、空間オーディオ再生システムのレンダリング機能が利用でき、滑らかな利得変化や、周波数依存利得が、遅延や処理負荷の増加を伴わず適用できうる。さらに、理想的なヘッドルーム構成に「到達」できる。

【0070】

分散型オーディオキャプチャシステムの一部として実現される、利得制御に適切な方法および装置を以下に説明する。以下の例は、ライブコンテンツのキャプチャおよび処理について説明される。ただし、ライブおよび記録コンテンツ（または記録コンテンツのみ）の処理に適切なさらなる例を採用できる。

【0071】

以下にさらに詳細に説明する方法は、頭部追跡装置からの出力を利用して、オーディオ信号に対応付けられた利得／処理パラメータを変更することからなる。以下の例において、キャプチャ、ミキサ、出力システムはそれぞれ個別であるものと示されるが、単一の装置で実現されてもよいし、物理的に分離しているが、互いに通信可能な一連の装置で分散されてもよい。

【0072】

さらに、以下の装置の少なくとも一部の要素は、いわゆる「クラウド」等の、分散型演算システム内で実現されうる。

【0073】

図1aに、いくつかの実施形態に係る利得制御の実現に適した、分散型オーディオキャプチャおよび処理システムの一例を示す。

【0074】

以下の例では、3つの外部音源が示されるが、キャプチャ対象の音源の数を3つよりも少なくして、または多くして、以下の装置および方法を適用することもできる。例えば、システム99は第1の音源マイク111を備える。第1の音源マイク111は、ラベリアマイクまたは「近接」音源キャプチャ装置のその他の例であってもよい。例えばある実施形態において、第1の音源マイクは、ブームマイクまたは同様の接近型マイクキャプチャシステムであってもよい。ある実施形態において、第1の音源マイク111はマイクアレイであってもよい。図1aに示す第1の音源マイク111は、ボーカル等の第1の音響源に対応付けられたマイクであってもよい。

【0075】

システム99は、第2の音源マイク113をさらに備えてもよい。第2の音源マイク113は、楽器等の第2の音響源に対応付けられたマイクまたはマイクアレイであってもよい。例えば、第2の音源マイクは、エレキギターに内蔵された内部マイクシステム（ピックアップマイク）である。

【0076】

システム99は、第3の音源マイク115をさらに備えてもよい。第3の音源マイク115は、第3の音響源に対応付けられたマイクまたはマイクアレイであってもよい。例え

10

20

30

40

50

ば、第3の音源マイクは、ドラムに対応付けられたマイクである。

【0077】

音源マイク111、113、115は、いずれも対応付けられた音源と共にダイナミックに動くもので、かつ、音源のダイナミックオーディオ出力を反映するものであってもよいことが理解されたい。

【0078】

音源マイク111、113、115は、それぞれが取得したオーディオ信号を、プロセッサ兼レンダラ100に出力するように構成されてもよい。例えば、第1の音源マイク111は、プロセッサ兼レンダラ100の受信部（図示せず）へとオーディオ信号を無線送信する送信部（図示せず）を備えてもよい。

【0079】

システム99は、空間オーディオキャプチャ（Spatial Audio Capture：SPAC）デバイス105をさらに備えてもよい。空間オーディオキャプチャデバイス105は、「音響場」キャプチャ装置の一例であって、ある実施形態においては、本明細書に記載の音源により表される周囲音場面やその他の周囲音源に対応付けられたオーディオ信号を取得するように構成された指向性マイクアレイまたは全方向性マイクアレイであってもよい。空間オーディオキャプチャデバイス105は、取得したオーディオ信号を、プロセッサ兼レンダラ100に出力するように構成されてもよい。

【0080】

ある実施形態において、空間オーディオキャプチャデバイス105はモバイルデバイス内に実現される。この場合、空間オーディオキャプチャデバイスは、空間オーディオを取得するように構成されており、当該空間オーディオは、聴き手に届けられると、聴き手が実際に空間オーディオキャプチャデバイスの所在地に存在するかのよう感じられる音場を体験可能とするものである。

【0081】

空間オーディオキャプチャデバイス105は、位置決定部または位置追跡部（図1aに図示せず）をさらに備えてもよい。位置決定部または位置追跡部は、空間オーディオキャプチャデバイス105に対する、音源マイク111、113、115（したがって、それらに対応付けられた音源）の位置を特定するのに適した出力を生成し、この情報をプロセッサ兼レンダラ100に出力するように構成されてもよい。図1aにおいて音源マイク111、113、115から延びるように示された点線が、位置追跡部または決定部から測位システム151への出力伝送を示すものであってもよい。以下の例においては、空間オーディオキャプチャデバイスは基準位置の一例とされるが、任意の物体位置を基準としてもよい。位置決定／追跡は任意の適切な手段または方法で実現されてもよい。

【0082】

例えばある実施形態において、測位システムは衛星測位信号（またはその他のビーコン信号）を利用して、音源と空間オーディオキャプチャデバイス（ある実施形態においてはミキサ）の位置を推定してもよい。このように推定された位置は、位置決定部または位置追跡部に送信されてもよい。位置決定部または位置追跡部は、これらの位置を受信し、ある実施形態においては空間キャプチャデバイス（および／またはミキサまたはミキシングコントローラ）に対する音源の位置を決定するように構成される。

【0083】

位置決定は、ローカル無線周波数信号推定により実行されてもよい。すなわち、例えば音源マイクは位置タグを備えてもよい。この位置タグは、無線周波数信号を生成および送信するように構成されてもよい。当該信号は、適切な受信部により受信可能であって、受信部または位置決定部に、受信部に対する音源マイクの位置または場所を特定するのに適切な情報を決定可能とするものでありうる。したがって、空間オーディオキャプチャデバイス105は、位置タグ受信部をさらに備えてもよい。この位置タグ受信部は、音源に対応付けられた位置タグからの無線周波数信号を受信するように構成される。空間オーディオキャプチャデバイス105は、空間オーディオキャプチャデバイスに対するタグの推定

10

20

30

40

50

位置を決定するように構成された位置決定部をさらに備えてもよい。人に装着される／楽器に設けられたマイクは、音響空間内を自由に移動できるので、位置決定部は、マイク位置の連続検知に対応可能である必要がある。

【0084】

ある実施形態において、受信部および／または位置追跡部または決定部は、空間オーディオキャプチャデバイス105外に設けられてもよい。そのような実施形態では、空間オーディオキャプチャデバイス自体が、位置タグを備える（または位置タグに対応付けられている）。この位置タグは、受信部および／または位置決定部が、位置決定部に対する空間オーディオキャプチャデバイスの位置を推定可能とする無線周波数信号を生成する。それに応じて、位置決定部は音源（または音源に対応付けられた位置タグ）と、空間オーディオキャプチャデバイス（およびデバイスに対応付けられた位置タグ）の（1つまたは複数の）相対位置を決定してもよい。

10

【0085】

ある実施形態において、高精度屋内測位（High Accuracy Indoor Positioning：HAIP）システム、またはその他の適切な屋内測位技術を利用して、位置タグおよび位置決定部が実現される。したがって、位置タグはHAIPタグであって、位置受信部および／または位置追跡部は、タグの位置を追跡するように構成されたHAIP探知部である。

【0086】

ある実施形態において、無線周波数に基づく位置推定に加えて、位置決定部および／または位置追跡部はビデオコンテンツ分析および／または音源位置特定を利用してもよい。

20

【0087】

すなわち、ある実施形態において、空間オーディオキャプチャデバイスの位置または配置箇所が決定される。空間オーディオキャプチャデバイスの配置箇所は、（0時点において）以下のように表される。

$$(x_s(0), y_s(0))$$

【0088】

ある実施形態において、第1の音源（s1）マイクが、HAIP探知部から特定の距離範囲で、SPACアレイの前方に配置される、較正工程または動作が実現されてもよい（言い換えると、0時点を定義する）。この第1の音源マイクの位置は、以下のように表される。

30

$$(x_{s1}(0), y_{s1}(0))$$

【0089】

さらに、ある実施形態において、この較正工程は、HAIP座標系における空間オーディオキャプチャデバイスの「前方向」を決定できる。これは、まず次のベクトルからアレイ前方向を定義することで実行されうる。

$$(x_{s1}(0) - x_s(0), y_{s1}(0) - y_s(0))$$

40

【0090】

このベクトルにより、位置追跡部はアレイに対する方位角 α_{s1} および距離 d_{s1} を決定可能としてもよい。

【0091】

例えば、時点 t での第1の音源マイク位置を次のようにする。

$$(x_{s1}(t), y_{s1}(t))$$

50

【0092】

アレイに対する方向は次のベクトルで定義される。

$$(x_{s1}(t) - x_s(0), y_{s1}(t) - y_s(0))$$

【0093】

ここから方位 α は以下のように決定される。

$$\alpha = \text{atan2}(y_{s1}(t) - y_s(0), x_{s1}(t) - x_s(0)) - \text{atan2}(y_{s1}(0) - y_s(0), x_{s1}(0) - x_s(0))$$

10

式中、 $\text{atan2}(y, x)$ は、正の x 軸と点 (x, y) との間の角度が得られる「四象限逆正接」である。したがって、第1の項により、正の x 軸 $(x_s(0), y_s(0))$ を原点とすると、点 $(x_{s1}(t), y_{s1}(t))$ との間の角度が求められ、第2の項により、 x 軸と、初期位置 $(x_{s1}(0), y_{s1}(0))$ との間の角度が求められる。方位角は第1の角度を第2の角度から引くことで得られてもよい。

【0094】

距離 d は次のように求められる。

$$\sqrt{(x_{s1}(t) - x_s(0))^2 + (y_{s1}(t) - y_s(0))^2}$$

20

【0095】

ある実施形態において、H A I P 位置データは多分にノイズを含みうるので、数秒（例えば30秒）の期間にわたって、オーディオキャプチャデバイスと第1の音源とのH A I P タグの位置を記録し、記録された位置の平均として位置 $(x_{s1}(0), y_{s1}(0))$ および $(x_s(0), y_s(0))$ を求め、上述の式の入力を得るようにしてもよい。

【0096】

ある実施形態において、較正工程はS P A C デバイス（例えばモバイルデバイス）により開始されてもよい。この場合、ユーザ（複数可）が、30秒間にアレイの前方から動かないように音声またはその指示を送り、当該期間の終了時に音で知らせるように構成される。

30

【0097】

上述の例では、二次元における位置情報が位置決定部および／または位置追跡部により生成されているが、この情報は三次元情報にまで拡大されてもよいことが理解されよう。その場合、位置決定部および／または位置追跡部は、方位角および距離に加えて、仰角も決定しうる。

【0098】

ある実施形態において、その他の位置決定／追跡手段を利用して、移動する音源の位置特定および追跡を実行してもよい。その他の追跡手段の例としては、慣性センサ、レーダー、超音波センシング、L I D A R またはレーザー測距計、視覚的分析、オーディオ分析等が挙げられる。

40

【0099】

例えば視覚的分析により、人や楽器のようなあらかじめ決められた音源の位置特定および追跡を実行してもよい。視覚的分析は、空間オーディオと共に撮影されたパノラマ映像に対して適用されてもよい。すなわち、この分析では、音源マイクを持つ人を視覚的に位置特定することに基づいて、人の位置を特定および追跡するものであってもよい。視覚的追跡の利点として、音源が音を発しておらず、音響に基づいた追跡が困難であるような状況でも利用可能である。視覚的追跡は、各パノラマ映像フレームについて、適切なデータ

50

群（歩行者を含む画像のデータ群等）で訓練した検出器を実行または動作することに基づいてもよい。別の実施形態において、動画フレーム間の、正しい人の軌跡を得るために、カルマンフィルタリングまたは粒子フィルタリング等の追跡技術を利用してもよい。ここで、空間オーディオキャプチャデバイスの前方向に一致する、パノラマ映像の前方向に対する人の位置を、該当する音源の到来方向として利用できる。ある実施形態において、音源マイクの外觀に基づく視覚的マーカまたは検出器を利用して、視覚的追跡方法の精度向上に寄与してもよい。

【0100】

ある実施形態において、視覚的分析は、音源（すなわち、パノラマ映像フレーム内の座標）の二次元位置についての情報のみを提供するものではなく、検出音源のサイズに比例する、距離についての情報も提供できる。ただし、後者は音源クラスの「標準的な」サイズがわかっていることが前提となる。例えば、「あらゆる」人の距離は、平均身長に基づいて推定できる。あるいは、システムにおいて特定の音源のサイズがわかっているならば、より正確な距離の推定も可能である。例えば、追跡が必要な各人の身長をシステムに伝えるまたはそれで訓練するようにしてもよい。

10

【0101】

ある実施形態において、奥行き検知装置を使用することで、三次元または距離情報を実現してもよい。例えば、「Kinect」システム、飛行時間型カメラ、ステレオカメラまたはカメラアレイを利用して、分析対象の画像を生成してもよく、多数の画像からの画像差分から、奥行きまたは三次元視覚的シーンが生成されてもよい。これら画像はカメラにより生成されてもよい。

20

【0102】

ある実施形態において、音響源位置決定および追跡を使用して音源を追跡しうる。例えば、到達時間差（Time Difference of Arrival：TDOA）法により、この音源の方向を推定しうる。音源位置決定は、ある実施形態において、ステアドビームフォーマを粒子フィルタ型追跡アルゴリズムと共に使用し実現してもよい。

【0103】

ある実施形態において、音源の追跡に、音響自己位置特定を利用してもよい。

【0104】

無線技術と、接続性に対するソリューションにおいて、さらにデバイス間の高精度同期に対応可能な技術が存在する。これは、音響相関解析における時間のずれの不確定性を取り除くことで、距離測定を簡潔化できるものである。これらの技術は、マルチチャネル音響再生システムの、将来的なWi-Fi規格化のために提案されている。

30

【0105】

ある実施形態において、屋内測位、視覚的分析、音響源位置特定のそれぞれによる位置推定を共に利用してもよい。例えば、各決定部または推定部からの推定を平均化して、位置決定および追跡精度を向上してもよい。さらに、視覚的分析の計算負荷（通常、オーディオ信号またはHAIP信号の分析よりもかなり「重い」）を最小限に抑えるように、パノラマフレーム全体の、音源が存在すると音響および／またはHAIP分析サブシステムが予測した空間的位置に対応する一部のみに視覚的分析を実施してもよい。

40

【0106】

ある実施形態において、位置推定は、多数の音源からの情報を組み合わせてもよい。多数の推定を組み合わせることにより、提供されたシステムに対して、正確な位置情報を提供する可能性がある。ただし、システムは低い解像度であっても、位置推定を実現する位置検出技術のサブセットを利用するように構成されることが有利である。

【0107】

システム99は、音源マイク111、113、115および空間オーディオキャプチャデバイス105の出力を受信するように構成されたプロセッサ兼レンダラ100をさらに備えてもよい。さらに、ある実施形態において、プロセッサ兼レンダラ100は、位置追跡部から音源位置（および追跡情報）を受信するように構成されてもよい。

50

【0108】

プロセッサ兼レンダラ100は、音源マイク111、113、115および空間オーディオキャプチャデバイス105からの出力を処理（例えばミキシング）して、このように処理された信号をヘッドホン121に出力するように構成されてもよい。本明細書で示すプロセッサ兼レンダラ100は、入力コントローラ103および頭部追跡部（または追跡部）101からの入力を受信し、この入力に基づいて音響源／空間オーディオキャプチャデバイス（または装置）からのオーディオ信号を処理するように構成されてもよい。これについては詳細に後述する。

【0109】

システム99は、入力コントローラ103をさらに備えてもよい。入力コントローラ103は、プロセッサ兼レンダラ100の音響出力を（例えばヘッドホン121を介して）聴いているユーザに、出力の処理を制御することを可能としてもよい。したがってある実施形態において、入力コントローラ103は、適切な選択を実現し、プロセッサ兼レンダラ100への出力を制御することで、音響源と、プロセッサ兼レンダラ100が生成した空間オーディオキャプチャデバイスオーディオ信号の処理（例えばミキシング）を制御してもよい。入力コントローラ103は、物理的コントローラ要素を備えてもよい。例えば、スライダ、ダイヤル、ボタンでもよいし、あるいはタッチスクリーンにおいて物理的コントローラ要素を仮想的に表したものであってもよい。ある実施形態において物理的コントローラ要素は、動的に再設定可能である。すなわち、同一のコントローラ要素で、第1の期間においては第1のパラメータまたは機能が制御され、第2の期間においては第2のパラメータまたは機能が制御されてもよい。

【0110】

システム99は、頭部追跡部（音源追跡部または単純に追跡部とも呼ばれるものである）101をさらに備えてもよい。これは、出力を聴くユーザの頭部の向きを追跡する任意の適切な慣性センサであってもよい。言い換えると、頭部追跡部101は、方位および／または仰角を測定するための任意の適切な手段であってもよい。例えば、頭部追跡部はヘッドホン121に取り付けられ、出力を聴いているユーザの、定義されたまたは基準向き（マイクアレイ「前」方向）に対する向きおよび／または方位を監視し、プロセッサ兼レンダラ100に出力可能な値または入力を提供するように構成されたセンサであってもよい。頭部追跡部101は、ある実施形態において、少なくとも1つのジャイロスコープおよび／またはデジタルコンパスにより実現されてもよい。動作を簡単にするため、頭部追跡部および空間オーディオキャプチャデバイスマイクアレイは、磁気コンパスを備えて、専用の手動校正ステップを必要なくしてもよい。磁気コンパスがない場合、頭部追跡部を装着したユーザは、オーディオキャプチャシステムの基準軸と同一の方向を向くことで、システムの校正が行われる。

【0111】

図1aに示すように、システム99は、一対のヘッドホン121をさらに備える。これは聴き手の頭部に装着可能で、適切なバンドを介して接続された一対のトランスデューサを備える。バンドに頭部追跡部101が装着される。ある実施形態において、ヘッドホン121は、電気信号をユーザへの音響出力に変換するのに適したイヤピース、イヤホン等であってもよい。

【0112】

図1bは、プロセッサ兼レンダラ100の例をさらに詳細に示す。プロセッサ兼レンダラ100は、オーディオ信号／ストリームを受信するように構成される。例えば、プロセッサ兼レンダラ100は、（空間オーディオキャプチャデバイス105内の）マイクアレイ141、あるいは第1の音源マイク111、第2の音源マイク113、または第3の音源マイク115からオーディオ信号を受信するように構成されてもよい。したがって入力、マイクアレイの2以上のマイク信号、あるいは既知の方向に配置されたラベリアマイクまたは楽器用マイクからの1以上の音響物体信号を含んでもよい（特定の実施形態では、オーディオ信号／ストリームとして、対応付けられた位置値を有するその他の任意のオ

10

20

30

40

50

オーディオ信号が使用されてもよい)。ある実施形態において、入力は「記録された」または記憶されたオーディオ信号であってもよい。例えば入力は、マルチチャネルラウドスピーカ信号、すなわちラウドスピーカチャネル構成に配置されたオーディオ信号を含んでもよい。ある実施形態において、音響入力は、サンプリングされたオーディオ信号と、聴き手に対する、音響源または物体の方向または位置を表すメタデータを含んでもよい。ある実施形態において、音響入力信号は、利得値等のその他の任意のパラメータ、またはオーディオ信号にかけられる等化フィルタを含んでもよい。

【0113】

ある実施形態において、プロセッサ兼レンダラ100はアナライザ131を有する。アナライザ131は、音響源に対応付けられた向きまたは位置を決定するために、音響入力を分析するように構成される。アナライザ131は任意の適切な方向または位置解析を実行してもよい。例えばある実施形態において、アナライザ131は、空間オーディオキャプチャデバイス105からの、測位システム151の出力を受信し、この情報から音源の向きまたは位置を決定するように構成される。

10

【0114】

ある実施形態において、アナライザ131は、空間オーディオキャプチャデバイス105からオーディオ信号を受信し、それを分析して音景内の音響源または物体に対応付けられた方向または位置情報を決定するように構成されてもよい。同様に、上述のように音源マイクはマイクアレイを含みうるため、その場合、アナライザは音源マイクからのオーディオ信号を分析し、音景内の音響源／物体を決定する。

20

【0115】

アナライザ131は、音源の到来方向(DOA)を決定するように構成されてもよい。さらに、アナライザ131は、直接成分と周囲成分の割合、またはその他の音景を表現および定義するのに寄与するその他のパラメータを決定するように構成されてもよい。

【0116】

図2は、例えば単一の音源201、およびM個のマイクを含むマイクアレイ200(例えば空間オーディオキャプチャデバイス105のマイクアレイ141でありうる)が存在する例を示す。マイクアレイ200は、マイク信号 $x(m, n')$ を提供するように構成されうる。ここで m と n' はそれぞれマイクと時間のインデックスである。音源201は、球面座標系において、方位角 $\phi(k, n)$ と、仰角 $\theta(k, n)$ とを有するように示されている。なお、 k 、 m 、 n はそれぞれ、周波数、マイク、時間のインデックスである。

30

【0117】

アナライザ131は、短時間フーリエ変換により、入力信号を周波数領域に変換してもよい。

$$X(k, m, n) = F(x(m, n'))$$

式中、 X は変換された信号を示す。

【0118】

音源の到来方向を決定するための方法が数多く知られている。ここでは、方向情報を決定するための1つの方法が説明される。有用性が確認された方法である。この方法はあくまで例示であって、別の方法を利用してもよい。この方法は、音響入力信号チャネル間の相関に基づく。この方法では、到来音の方向は、B周波数領域サブ帯域について個別に推定される。全サブ帯域について、最も突出して知覚される音源の方向を確認するという概念である。最も突出して知覚される音源の方向を全サブ帯域について決定すると、複数の音源がどこにあるか、複数の音源がそれぞれ異なるサブ帯域において突出している場合には、それらに対応付けられた方向を決定できる。

40

【0119】

上述のように、各入力信号 m はフーリエ変換により周波数領域に変換されるが、フーリエ変換は離散フーリエ変換(Discrete Fourier Transform: DFT)またはその他の適切に実施される変換であってもよい。例示的实施形態において、変換窓かけには、重複50%、有効長20ミリ秒(ms)の正弦波窓を利用する。ある実施形態において、フーリ

50

エ変換実施前に、窓の終端に $D_{tot} = D_{max} + D_{HRTF}$ ゼロが加えられる。 D_{max} は、マイク間のサンプルの最大遅延に対応する。最大遅延は以下のとおりに求められる。

$$D_{max} = \frac{dF_s}{v}$$

式中、 F_s は信号のサンプリングレートであって、

$$v$$

は空気中の音の速度である。 D_{HRTF} の値は、シンセサイザ兼 HRTF レンダラ 135 内で実施されうる任意の頭部伝達関数 (Head Related Transfer Function: HRTF) 処理により信号に生じた最大遅延である。

【0120】

DFT 変換後、全マイク m に対する周波数領域表現 $X(k, m, n)$ は $k = 0, \dots, N - 1$ となる。 N は、正弦波窓 (長さ N_s) と、追加された

$$D_{tot}$$

ゼロを考慮した窓の全長である。

【0121】

周波数領域表現は、 B 個のサブ帯域に分割される。

$$X_m^b(k) = X_m(k_b + i), i = 0, \dots, i_{b+1} - i_b - 1, b = 0, \dots, B - 1, \quad (2)$$

式中、

$$k_b$$

は第 b サブ帯域の第 1 インデックスである。サブ帯域の幅は、例えば等価矩形帯域幅 (Equivalent Rectangular Bandwidth: ERB) という尺度に基づいてもよい。

【0122】

全てのサブ帯域について、以下のように方向解析が実行される。

1. サブ帯域が選択される。
2. 当該サブ帯域内の信号に対して方向解析が実行される。方向解析は、(例えば突出した) 音源の方向 α_b を決定するものである。
3. 全てのサブ帯域が選択されたか判定する。まだ全て選択済みではなければ、新たなサブ帯域を選択してステップ 2 を繰り返す。選択されていれば、決定されたサブ帯域方向を処理して、突出した音源 (およびその他の音源) の方向を決定する。

【0123】

上記ステップ 2 は、以下のサブステップまたは動作にさらに分割されてもよい。

【0124】

最初に、2 つのマイク信号 (この例ではマイク 2 および 3) により、方向が推定される。2 つの入力マイクオーディオ信号について、これらチャンネルの周波数領域信号間の時間

差が除去される。目的は、サブ帯域 b についての 2 つのチャネル間の相関が最大となる遅延 τ_b を見つけることである。

【 0 1 2 5 】

例えば

$$X_m^b(k)$$

の周波数領域表現は、以下を使用して τ_b 時間領域サンプルずらすことができる。

$$X_{m,\tau_b}^b(k) = X_m^b(k) e^{-j \frac{2\pi k \tau_b}{N}} \quad 10$$

【 0 1 2 6 】

最適遅延が以下のとおりに求められる。

$$\max_{\tau_b} \operatorname{Re} \left(\sum_{k=0}^{n_{b+1}-n_b-1} (X_{2,\tau_b}^b(k)^* X_3^b(k)) \right), \tau_b \in [-D_{\max}, D_{\max}]$$

20

式中、 Re は、結果の実数部で、 $*$ は複素共役である。

$$X_{2,\tau_b}^b$$

および

$$X_3^b$$

は、

$$k_{b+1} - k_b \quad 30$$

サンプルの長さを持つベクトルと考えられる。1 サンプルの解像度は、概して上述の遅延を見つけるのに適したものである。また、相関とは異なる、その他の同様の知覚に基づく測定も利用可能である。遅延情報により、以下の論理から和信号が生成される。

$$X_{sum}^b = \begin{cases} (X_{2,\tau_b}^b + X_3^b)/2 & \tau_b \leq 0 \\ (X_2^b + X_{3,-\tau_b}^b)/2 & \tau_b > 0 \end{cases}$$

40

式中、 τ_b は前述のように定義された理想的な遅延である。

【 0 1 2 7 】

この和信号によると、イベントが生じるマイクのコンテンツ（すなわち、周波数領域信号）がまずそのまま加えられ、その後イベントが生じるマイクのコンテンツ（すなわち周波数領域信号）はシフトされて、最高の組合せが得られる。

【 0 1 2 8 】

したがって、音源は、例えばアレイの第 2 のマイクのようなマイクで受信される例示的時間領域関数で示されるイベントを生成するものと考えられる。同じイベントが第 3 のマイクで受信される。理想的な状況としては、アレイの第 2 のマイクで受信される例示的時間領域関数が、単純に第 3 のマイクで受信される関数を時間シフトしたものとなる。この

50

状況が理想的である理由は、現実では2つのマイクが、例えば、それらによるイベントの記録が、強め合うまたは弱め合う干渉や、イベントからの音を遮蔽または増幅する要素により影響されるような、異なる様々な環境に置かれることが多いためである。

【0129】

シフト値 τ_b は、音源が第3のマイクよりも第2のマイクにどれ程近いかを示す (τ_b が正であれば、音源は第3のマイクよりも第2のマイクに近い)。実際の距離の差は以下のように計算される。

$$\Delta_{23} = \frac{v\tau_b}{F_s} \quad 10$$

【0130】

基本的な幾何学を利用して、到来音の角度が以下に等しいと判定できる。

$$\dot{\alpha}_b = \pm \cos^{-1} \left(\frac{\Delta_{23}^2 + 2b\Delta_{23} - d^2}{2db} \right)$$

式中、 d はマイク間の距離であり、 b は音源と最近傍のマイクとの推定距離である。典型的には、 b は固定値に設定できる。例えば $b = 2$ メートルとすることで、安定した結果が得られることが確認されている。

20

【0131】

なお、2つのマイクだけでは正確な方向が判定できず、音の到来方向については2つの可能性が存在する。

【0132】

そこで、例えば3つのマイクのアレイにおける第1のマイクのようなさらなるマイクを利用して、いずれの符号 (+ または -) が正しいかを定義できる。

第1のマイクと、2つの推定された音源との距離は以下のとおりに定義できる。

30

$$\delta_b^+ = \sqrt{(h + b \sin(\dot{\alpha}_b))^2 + (d/2 + b \cos(\dot{\alpha}_b))^2}$$

$$\delta_b^- = \sqrt{(h - b \sin(\dot{\alpha}_b))^2 + (d/2 + b \cos(\dot{\alpha}_b))^2}$$

式中、 h は、第1、第2、および第3のマイクの例示的配置 (マイクが三角形の頂点にそれぞれ配置される) で定義される二等辺三角形の高さを表す。この例では、 h は以下のとおりに定義される。

$$h = \frac{\sqrt{3}}{2} d \quad 40$$

距離は、サンプル内の遅延に等しい。

50

$$\tau_b^+ = \frac{\delta^{+-b}}{v} F_s$$

$$\tau_b^- = \frac{\delta^{-b}}{v} F_s$$

10

これら2つの遅延の内、和信号とのよりよい相関を実現する一方が選択される。相関は以下のように求められる。

$$c_b^+ = \text{Re} \left(\sum_{k=0}^{k_{b+1}-k_b-1} (X_{sum, \tau_b^+}^b(k)^* X_1^b(k)) \right)$$

$$c_b^- = \text{Re} \left(\sum_{k=0}^{k_{b+1}-k_b-1} (X_{sum, \tau_b^-}^b(k)^* X_1^b(k)) \right)$$

この上で、サブ帯域 b について顕著な音源に対する方向が以下のように求められる。

20

$$\alpha_b = \begin{cases} \dot{\alpha}_b & c_b^+ \geq c_b^- \\ -\dot{\alpha}_b & c_b^+ < c_b^- \end{cases}$$

【0133】

同じ推定が、各サブ帯域に対して繰り返される。

【0134】

この方法は、「水平」および「垂直」に変位があるようなマイクアレイにも適用できる。これによって、方位と仰角とが決定できる。

30

【0135】

したがって、音源の方向は、 b をサブ帯域（すなわち、 k 個のインデックス群）として、球面座標系内の方位 $\phi(b, n)$ と、仰角 $\theta(b, n)$ により表される。したがって、アナライザ131は入力処理して、方向情報と、音場を表す別のメタデータまたはパラメータを提供できる。この位置メタデータは、シンセサイザ／レンダラ135およびレベルコントローラ133に送ってもよい。

【0136】

さらに、図1bに示すように、プロセッサ兼レンダラ100はシンセサイザ／レンダラ135を備えてもよい。シンセサイザ／レンダラ135は、アナライザ131からの方向メタデータおよび入力信号 $X(k, n, m)$ と、レベルコントローラ133からの利得またはレベル制御と、頭部追跡部の向きを利用して、バイノーラル出力オーディオ信号を合成するように構成されてもよい。

40

【0137】

シンセサイザ／レンダラ135は、入力オーディオ信号に対して頭部伝達関数(HRTF)を適用することで、左および右チャンネル出力オーディオ信号を生成するまたは合成するように構成されてもよい。HRTFは、人の耳、頭部、胴等が、特定の方向からの音にどのように影響するかを表すものである。したがって、単純に複数の音響源のDOAと、複数の頭部追跡部の向きに応じて、異なる複数のHRTFが適用される。

【0138】

図3に、頭部追跡部第1の（すなわち基準）向きに対する音源の向きの一例を示す。こ

50

の例では、マイクアレイ 200 の基準向き 300 に対して、音源 201 は固定の仰角 ($\theta = 0$) と、方位角 ϕ 203 に配置されているように示されている。したがって、頭部追跡部が、マイクアレイと同じ「基準」向きまたは方向に向けられると、例示的音源 201 は聴き手の左側に位置することになる。

【0139】

そのような例では、シンセサイザ 135 は、第 1 の「左チャンネル」HRTF を入力オーディオ信号に適用して、左チャンネルバイノーラル出力オーディオ信号を生成し、さらに第 2 の「右チャンネル」HRTF を入力オーディオ信号に適用して、右チャンネルバイノーラル出力オーディオ信号を出力するように構成される。

【0140】

なお、上述の例では、聴き手の頭部が、音響源から右耳への直接経路を遮り、対応する HRTF の振幅反応は、ローパスフィルタ形状となることが理解されよう。一方で、音響源から左耳オーディオ出力への経路は頭部により遮られることはないので、対応する HRTF の振幅反応は上述のようなローパス的特性を帯びず、高周波数の割合が高まる。シンセサイザ／レンダラ 135 は、聴き手が頭部追跡部を利用する際に、その頭部の動きに合わせてレンダリングされた音景も動くようにさらに構成される。

【0141】

例えば、図 4 は、第 1 の、すなわち基準向きに対して、第 2 の、すなわち回転向きに頭部追跡部が配置された例を示す。すなわち、例えば頭部追跡部の向き 400 が、マイクアレイ基準向き 300 に対して向き角度 ϕ_{head} 401 となっている。聴き手に対する音源 201 の向き（向き 400 により定義される）は、 $\phi - \phi_{head}$ 403 と定義されうる。この例では、シンセサイザ 135 はここで、元の方位角 ϕ ではなく、新たな方位角 $\phi - \phi_{head}$ の音源入力向きに基づく HRTF を利用して、オーディオ信号をレンダリングする。頭部追跡部の可変の向きに基づく可変 HRTF の提供の効果として、周波数応答に応じて左および右チャンネル出力オーディオ信号を変化させるだけでなく、出力レベルを変化させることが挙げられる。

【0142】

図 5 に、方位角変化に応じた出力レベルの変化を示す。図 5 では、音源および頭部追跡部の向き $\phi - \phi_{head}$ の変化のための、左チャンネルオーディオ出力の出力レベル分布の例を示す。角度 $\phi_0 = \phi - \phi_{head}$ は、任意の固定角度である。最大出力レベルは、 ϕ_{Max} 502 で聴こえうる。 ϕ_0 503 では数デシベル低くなる。図 8 に示す理想的曲線を、所望の最大出力レベルを超えないように選択できる。

【0143】

本明細書に記載のように、より高い周波数帯で、出力レベルがより激しく変化する。例えば図 6 は、例示的なオーディオ信号に対する、頭部追跡部の向きの差 $\phi - \phi_{head} = \pi / 2$ 601 および $\phi - \phi_{head} = -\pi / 2$ 603 に基づく出力レベル周波数分布の例を概略的に示す。本例に示すように、左耳が例示的音源 $\phi - \phi_{head} = \pi / 2$ 601 に対する到来方向に「向いている」場合、出力オーディオ信号の周波数応答は、スペクトルの広範囲で高くなる。左耳が例示的音源 $\phi - \phi_{head} = -\pi / 2$ 603 に対する到来方向に「向いていない」場合、頭部の遮蔽効果により、スペクトルの出力レベルが低減する。この低減は高い周波数（例えば $> 3,500\text{ Hz}$ ）で顕著である。

【0144】

この出力レベルの変化は、利得制御およびダイナミックレンジ制御動作において考慮すべきものである。

【0145】

図 7 は、図 1 b に示すレベルコントローラ 133 をより詳細に示すものである。図 1 b に示すレベルコントローラは、アナライザ 13 から入力オーディオ信号 $X(k, n, m)$ と、この入力オーディオ信号に対応付けられたその他のメタデータ（音源方向 $\phi(k, n)$ 等）を受信するように構成される。さらに、レベルコントローラ 133 は、頭部追跡部 $\phi_{HEAD}(k, n)$ から頭部の向き値 701 を受信するように構成されうる。

10

20

30

40

50

【0146】

レベルコントローラ133は、入力オーディオ信号と、音源の方向と、頭部追跡部の向き値に基づいて、利得値 $G(k, n)$ を生成し、その利得値 $G(k, n)$ 761をシンセサイザ／レンダラ135に出力するように構成される。

【0147】

シンセサイザ／レンダラ135は、これらの利得値を受信し、入力オーディオ信号 $X(k, n, m)$ または出力オーディオ信号のそれぞれ的一方に適用するように構成されてもよい。

【0148】

ある実施形態において、レベルコントローラ133は、 M 個のマイク入力 $m = [1, M]$ に対する入力レベル予測 $X'(k, n) = \max(\text{abs}(X(k, m, n)))$ と、 ϕ_{head} により定義された方位角に基づく、入力レベルと出力レベルとの所望の実現する出力利得を計算するように構成される。ある実施形態において、利得関係は、一連の関数（または曲線）として表されてもよく、これを角度 $\phi(k) - \phi_{\text{head}}(k)$ において所望の最大出力レベルを超えないように選択される。ある実施形態において、関数または曲線は、図5に示すような出力レベル推定に基づいて生成されてもよい。入力レベルおよび出力レベルの関係は、異なる様々な方位角 $\phi - \phi_{\text{head}}$ と、周波数帯 k について測定され、所望の実現する利得が決定される。一例示的实施形態では、レベル制御アルゴリズムがテーブルルックアップ機能により、所望の入力レベルおよび出力レベル関係を実現する曲線を選択する。

【0149】

オーバーフローを防ぐため、利得曲線は、例えば一定の利得を低い入力レベルで実現し、より高い入力レベルではより低い利得が実現されるように、2つ以上のセクションからなるものであってもよい。ヘッドルームを最小限に抑えるため、最大予測入力レベルで使用される利得 $g_{\text{MIN}}(k, \phi)$ を決定する必要がある。システム閾値 $Y_{\text{THMAX}}(k)$ が左および右出力の最大レベルを定義し、符号 $Y_{\text{MAX}}(k, \phi)$ が最大予測入力レベルでの出力レベルを示すものとして、利得は、 $g_{\text{MIN}}(k, \phi) = Y_{\text{THMAX}}(k) / Y_{\text{MAX}}(k, \phi)$ となる。例えば一定利得値のようなその他の利得曲線パラメータは、使用されるシステムに依存するものであって、多くの場合主観的に判断される。線形性を仮定することで、出力レベル $Y_{\text{EST}}(k, \phi, n) = C * X'(k, n) * Th(k, \phi(n))$ に対する推定を導ける。項 $Th(k, \phi(n))$ により、出力レベル推定に方位角への依存性が加わる。ここで C はスカラー定数である。測定データは、 C および Th 値の決定に利用できる。

【0150】

例えば図8は、単一の周波数帯と、一対の方位角に対する、利得関数曲線を示す、一連の関数の例を示す。したがって、例えば既定の入力レベルまでは、利得関数800は一定であり、方位角に依存しない。既定の入力レベルを超えると、最大出力レベル利得関数803が、基準出力レベル利得関数801よりも低くなるように、利得関数が方位角に応じて変化する。このように位置と周波数に依存した曲線を利用することで、レベルコントローラは必要なヘッドルームを最小限に抑えられる。

【0151】

ある実施形態において、既知の方向に近接場マイクを設け、レベルコントローラ133は M 個のマイク入力 $m = [1, M]$ に対して推定出力レベル $Y_{\text{EST}}(k, n) = \sqrt{\sum(\text{abs}(X(k, m, n))^2 * Th(k, \phi(m, n) - \phi_{\text{head}}(n)))}$ を計算するように構成されてもよい。

【0152】

項 $Th(k, \phi(m, n) - \phi_{\text{head}}(n))$ は、各近接場マイク m に対する入力レベルおよび出力レベル関係の方位角依存性を補償するものである。

【0153】

したがって、レベル制御アルゴリズムは、各周波数帯 k に対する推定出力レベル Y_{EST}

に基づく所望の出力レベルを定義する固定曲線を利用してもよい。

【0154】

上述のように、シンセサイザ／レンダラ135は、利得入力G(k, n)を受信し、乗算器を利用して利得を信号経路に適用するように構成されうる。ある実施形態において、利得はアナライザ131と、シンセサイザ／レンダラ135との間の乗算器ブロックにより適用される。

【0155】

上述したような実施形態では、システムは、自身に対する遅延や計算負荷を最低限に抑えうる。これは、レベル制御関数の適用に個別のバッファリング、周波数領域変換、またはフィルタバンクが必要ないためである。

【0156】

さらに、ある実施形態において、プロセッサ兼レンダラ100はリミッタ137を備えてもよい。リミッタ137は、シンセサイザ／レンダラ135からの出力を受信して、所望の振幅範囲内にパルス符号変調(Pulse Code Modulated: PCM)信号が収まることを保証し、それにより、若干小さくなったヘッドルームを利用可能とする。システム遅延増大を防ぐため、リミッタ137はシンセサイザ／レンダラ135の出力バッファリングを利用できる。

【0157】

ある実施形態において、本明細書に記載のレベルコントローラを変形して、個別直列レベルコントローラを特定の入力(例えばパーカッション楽器またはベースギター)に実現可能となってもよい。ある実施形態において、ドラム等の近接マイク入力は、出力信号にミキシングされる前に、個別レベル制御関数を有していてもよい。

【0158】

ある実施形態において、位置追跡部から受信したユーザの頭部の向き／位置は、質量および慣性でモデル化可能な連続した関数として扱うことができる。したがって、現在の頭部位置／向きから、並進および／または回転速度により未来の頭部位置／向きを予測することができる。ある実施形態において、レベルコントローラは利得推定を、現在の頭部位置だけでなく、予測頭部位置についても判定可能に構成されてもよい。例えば聴き手が耳を音源に向けていれば、レベル制御によってより滑らかな制御曲線が実現できうる。

【0159】

上述のように従来技術の利得制御方法では、頭部追跡バイノーラルレンダリングでは極めて望ましくない余分な遅延が発生するが、詳述した上記実施形態では、利得制御を空間一音処理に組み込むことで、これを防止できる。

【0160】

従来技術の利得制御方法では、独立した処理ブロックを要するため、計算が過度に複雑化してしまう。本明細書に提示の実施形態は、レベル(利得)制御を空間一音レンダリングに組み込むことで、これを防止できうる。したがって、余分な処理を最小限に抑えながら、レベル(利得)制御が実行できる。

【0161】

従来技術の利得制御方法では、頭部追跡情報もDOA情報も考慮されない。したがって、クリッピングの防止のために、余分なヘッドルームまたは余分の遅延が必要となる。本明細書に提示の発明は、上述の情報を利用するので、余分なヘッドルームも余分の遅延も不要となる。

【0162】

シンセサイザ／レンダラの後段に追加のリミッタを使用することで、本明細書に記載の実施形態は、従来技術の手段よりも優れた出力オーディオ品質を実現することが期待される。通常、ソフトリミット機能が必要となることはあまりないためである。

【0163】

図9に、プロセッサ兼レンダラ100の少なくとも一部、またはシステム99の一部として利用可能な電子デバイスの例を示す。このデバイスは、任意の適切なデバイスまたは

10

20

30

40

50

装置であってもよい。例えばある実施形態において、デバイス1200は仮想または拡張現実キャプチャデバイス、モバイルデバイス、ユーザ機器、タブレットコンピュータ、コンピュータ、オーディオ再生装置等である。

【0164】

デバイス1200は、マイクアレイ1201を備えてもよい。マイクアレイ1201は、複数（例えばM個）のマイクを備えてもよい。ただし、マイクは任意の適切な構成であってもよいし、任意の適切な数のマイクを利用してもよい。ある実施形態において、マイクアレイ1201は装置から分離しており、オーディオ信号は有線または無線接続を介して装置に送られる。マイクアレイ1201は、ある実施形態において、図1aに示すようなSPACマイクアレイ144であってもよい。

10

【0165】

マイクは、音波を適切な電氣的オーディオ信号に変換するように構成されたトランスデューサであってもよい。ある実施形態において、マイクはソリッドステートマイクであってもよい。言い換えると、マイクはオーディオ信号を取得して、適切なデジタル形式の信号を出力可能であってもよい。別の実施形態において、マイクまたはマイクアレイ1201は任意の適切なマイクまたはオーディオキャプチャ手段を含み入る。例としては、コンデンサマイク、キャパシタマイク、静電マイク、エレクトレットコンデンサマイク、ダイナミックマイク、リボンマイク、カーボンマイク、圧電マイク、または微小電気機械システム（Microelectrical-Mechanical System：MEMS）マイクが挙げられる。このマイクはある実施形態において、オーディオキャプチャされた信号をA/Dコンバータ（ADC）1203に出力可能である。

20

【0166】

デバイス1200は、A/Dコンバータ1203をさらに備えてもよい。A/Dコンバータ1203は、マイクアレイ1201における各マイクからのオーディオ信号を受信し、それを処理に適した形式に変換するようにさらに構成されてもよい。ある実施形態において、マイクが一体型マイクであれば、A/Dコンバータは不要である。A/Dコンバータ1203は、任意の適切なA/D変換または処理手段であってもよい。A/Dコンバータ1203は、オーディオ信号のデジタル表現を、プロセッサ1207またはメモリ1211に出力するように構成されてもよい。

【0167】

ある実施形態において、デバイス1200は、少なくとも1つのプロセッサまたはCPU1207を備える。プロセッサ1207は様々なプログラムコードを実行するように構成できる。実行されるプログラムコードは、例えば本明細書に記載のSPAC制御、レベル制御、分析、合成／レンダリングを含んでもよい。

30

【0168】

ある実施形態において、デバイス1200はメモリ1211を備える。ある実施形態において、少なくとも1つのプロセッサ1207がメモリ1211に接続される。メモリ1211は、任意の適切な記憶手段であってもよい。ある実施形態において、メモリ1211は、プロセッサ1207で実行可能なプログラムコードを記憶するためのプログラムコード部を備える。さらに、ある実施形態において、メモリ1211は、データ（例えば、本明細書に記載の実施形態のとおり処理されたまたは処理されるデータ）を記憶するための記憶データ部をさらに備えてもよい。プログラムコード部内に記憶された、実行されたプログラムコードと、記憶データ部内に記憶されたデータとは、必要に応じてプロセッサ1207がメモリープロセッサ接続を介して取得できる。

40

【0169】

ある実施形態において、デバイス1200はユーザインタフェース1205を備える。ある実施形態において、ユーザインタフェース1205は、プロセッサ1207に接続可能である。ある実施形態において、プロセッサ1207は、ユーザインタフェース1205の動作を制御して、ユーザインタフェース1205からの入力を受信可能である。ある実施形態において、ユーザインタフェース1205はユーザに、例えばキーパッドを介し

50

てデバイス1200に対してコマンドを入力可能とするものである。ある実施形態において、ユーザインタフェース1205はユーザに、デバイス1200から情報を取得可能とする。例えば、ユーザインタフェース1205は、デバイス1200からの情報をユーザに表示するように構成されたディスプレイを備えてもよい。ある実施形態において、ユーザインタフェース1205はタッチスクリーンまたはタッチインタフェースを備えてもよい。これらはデバイス1200への情報の入力と、デバイス1200のユーザへの情報の表示の両方を実現可能である。

【0170】

実施形態によっては、デバイス1200は送受信部1209を備える。当該実施形態において、送受信部1209はプロセッサ1207に接続されて、別の装置または電子デバイスと、例えば無線通信ネットワークを介して通信可能とするように構成される。送受信部1209または任意の適切な送受信部または送信および／または受信手段は、ある実施形態において、有線または無線接続を介して別の電子デバイスまたは装置を通信するように構成されてもよい。

10

【0171】

例えば図9に示すように、送受信部1209は頭部追跡部101、SPACデバイス105、またはヘッドホン121と通信するように構成されてもよい。

【0172】

送受信部1209はさらなる装置と、任意の適切な公知の通信プロトコルにより通信可能である。例えばある実施形態において、送受信部1209または送受信手段は、好適なユニバーサル移動体通信システム（Universal Mobile Telecommunications System：UMTS）プロトコル、例えばIEEE802.X等の無線ローカルエリアネットワーク（Wireless Local Area Network：WLAN）プロトコル、Bluetooth等の好適な短距離無線周波数通信プロトコル、または赤外線通信経路（Infrared Data communication pathway：IRDA）を用いる。

20

【0173】

ある実施形態において、デバイス1200は、ミキサおよび／またはレンダリング装置としてさらに利用されてもよい。この場合、送受信部1209は、オーディオ信号および位置情報を受信し、適切なコードを実行するプロセッサ1207を利用して、適切なオーディオ信号レンダリングを生成するように構成されてもよい。デバイス1200は、D/Aコンバータ1213を備えてもよい。D/Aコンバータ1213は、プロセッサ1207および／またはメモリ1211に接続されて、（例えば本明細書に記載のオーディオ信号のオーディオレンダリング後、プロセッサ1207からの）オーディオ信号のデジタル表現を、オーディオサブシステム出力を介して提供するのに適した、適切なアナログ形式に変換するように構成されてもよい。D/Aコンバータ（DAC）1213または信号処理手段は、ある実施形態において任意の適切なDAC技術を利用してもよい。

30

【0174】

さらに、デバイス1200は、ある実施形態において、オーディオサブシステム出力1215を備えてもよい。図9に示すような一例として、オーディオサブシステム出力1215はヘッドホン121への接続を可能とするように構成された出力ソケットであってもよい。ただし、オーディオサブシステム出力1215は、任意の適切なオーディオ出力またはオーディオ出力への接続部でありうる。例えば、オーディオサブシステム出力1215は、マルチチャネルスピーカシステムへの接続部であってもよい。

40

【0175】

ある実施形態において、D/Aコンバータ1213と、オーディオサブシステム1215とは、物理的に別々の出力デバイスで実施されてもよい。例えば、DAC1213およびオーディオサブシステム1215が、送受信部1209を介してデバイス1200と通信するコードレスイヤホンとして実施されてもよい。

【0176】

図示のデバイス1200は、オーディオキャプチャおよびオーディオレンダリング要素

50

を両方備えているが、ある実施形態においては、デバイス 1200 はオーディオキャプチャまたはオーディオレンダリング装置要素のみを備えてもよい。

【0177】

一般に、本発明の様々な実施形態が、ハードウェアまたは特定用途向け回路、ソフトウェア、ロジック、またはそれらの組合せで実装されてもよい。例えば、ある態様はハードウェアで実装されてもよく、別の態様はコントローラやマイクロプロセッサ等のコンピュータデバイスによって実行されるファームウェアやソフトウェアで実装されてもよいが、本発明はこれらに限定されない。本発明の種々の態様はブロック図、フローチャート、または他の図的記述を使用して記述ないし図示される。本明細書に記載されるこれらのブロック、装置、システム、技術、方法は、非限定的な例として、ハードウェア、ソフトウェア、ファームウェア、特定用途向け回路やロジック、汎用ハードウェア、コントローラや他のコンピュータデバイス、またはそれらの組合せで実装されてもよいと理解されるべきである。

【0178】

本発明の実施形態は、プロセッサエンティティ内にあるようなモバイルデバイスのデータプロセッサによって実行可能なコンピュータソフトウェアによって、またはハードウェアによって、またはソフトウェアとハードウェアの組合せによって実装されてもよい。またこの点に関して、添付する図面に示される論理フローの任意のブロックが、プログラムのステップや、相互接続された論理回路・ブロック・機能、またはプログラムのステップ、論理回路・ブロック・機能の組合せを表現してもよいことに留意されたい。ソフトウェアは、メモリチップ等の物理媒体、プロセッサ内に実装されるメモリブロック、ハードディスクやフレキシブルディスク等の磁気媒体、DVD、そのデータ異形態であるCD等の光学媒体に格納されてもよい。

【0179】

メモリは、ローカルな技術環境に適したあらゆるタイプのものであってよく、例えば、半導体ベースのメモリデバイス、磁気メモリデバイス・システム、光学式メモリデバイス・システム、固定式メモリ、移動式メモリ等の様々な適合するデータ格納技術を用いて実装されてもよい。データプロセッサは、ローカルな技術環境に適したあらゆるタイプのものであってよく、非限定的な例として、1つ以上の汎用コンピュータ、特定用途向けコンピュータ、マイクロプロセッサ、デジタル信号プロセッサ (Digital Signal Processor : DSP)、特定用途向け集積回路 (Application Specific Integrated Circuit : ASIC)、ゲートレベル回路、マルチコアプロセッサ・アーキテクチャに基づくプロセッサを含んでもよい。

【0180】

本発明の実施形態は、集積回路モジュールのような、様々な要素で実施されることもできる。集積回路の設計は多くは自動化されたプロセスである。論理レベルの設計を、半導体基板上にエッチング・形成するための半導体回路設計に変換する複雑で強力なソフトウェアツールが利用可能である。

【0181】

カリフォルニア州マウンテンビューの Synopsys, Inc や、カリフォルニア州サンノゼの Cadence Design のような業者が提供するプログラムは、定評のある設計ルールと実績のある設計モジュールのライブラリに基づいて、半導体チップ上に導電経路や要素を自動的に配する。半導体回路の設計が完了すると、その設計は製造のために、Opus や GDSII 等の標準的な電子フォーマットの形で半導体製造設備、いわゆる fab に送られる。

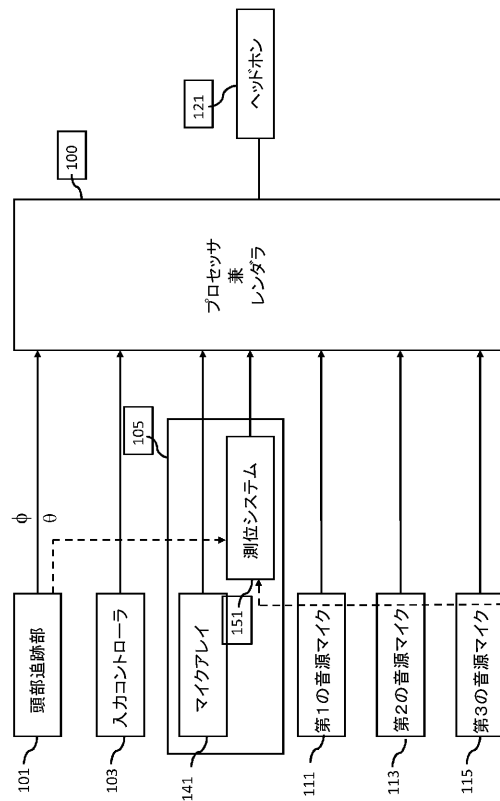
【0182】

前述の説明は、本発明の例示的で非限定的な実施例を十分かつ詳細に記述している。しかし、こうした前述の説明を、添付する図面および特許請求の範囲と併せて考慮すれば、種々の修正および変更が可能であることは、本願に関連する技術分野の当業者には明らかであろう。さらに、本発明が教示するこうした事項の全ておよび類似する変形は、添付の

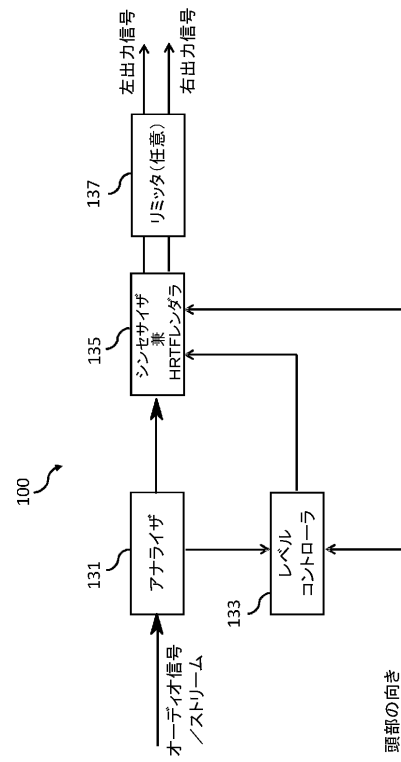
特許請求の範囲で定義されるように、その全てが本発明の範囲内にある。

【図面】

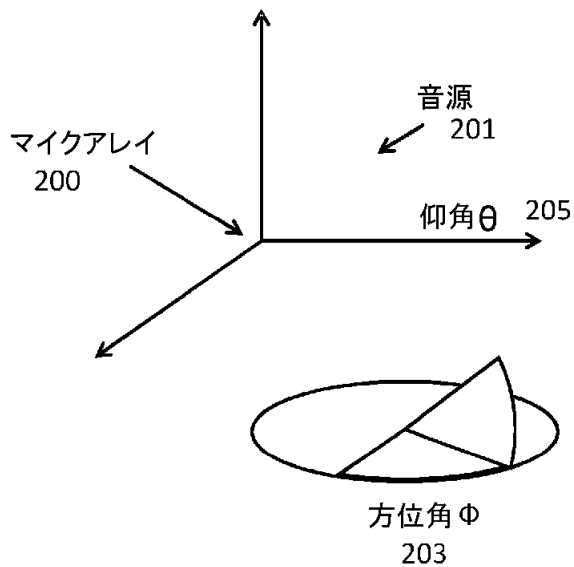
【図 1 a】



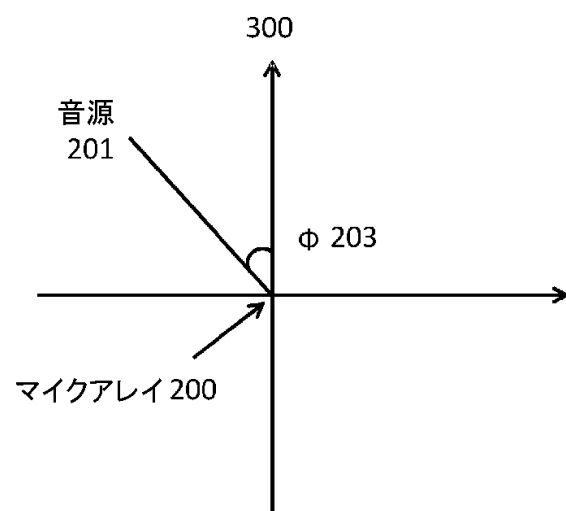
【図 1 b】



【図 2】



【図 3】



10

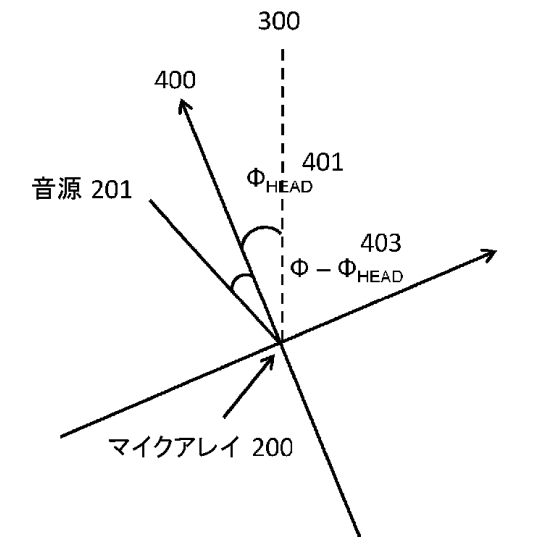
20

30

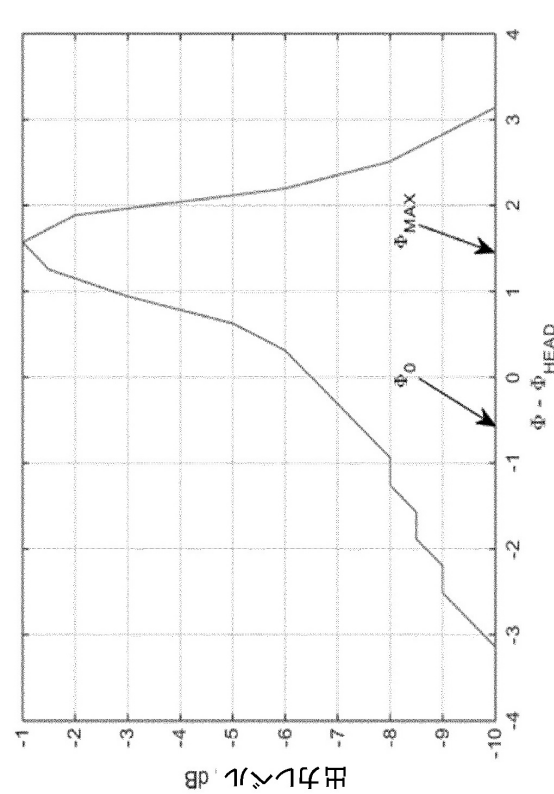
40

50

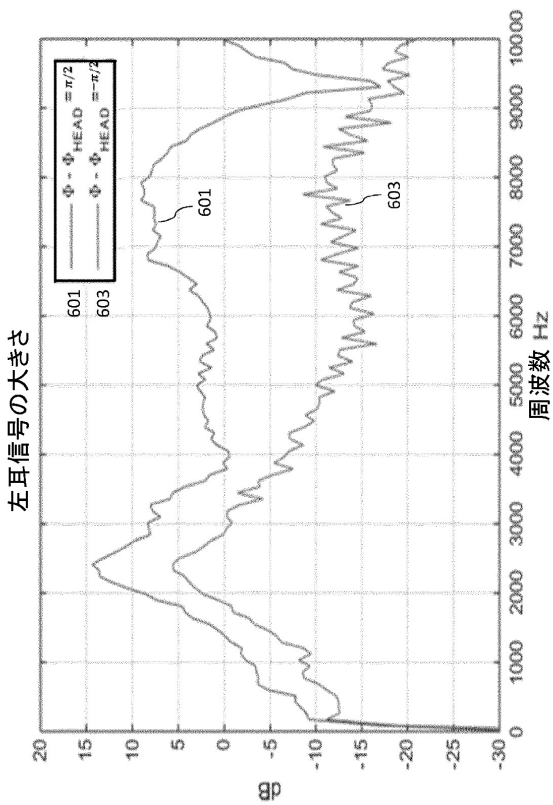
【図 4】



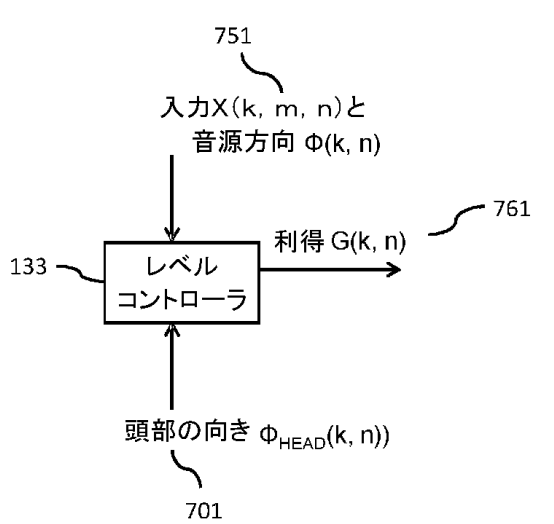
【図 5】



【図 6】



【図 7】



10

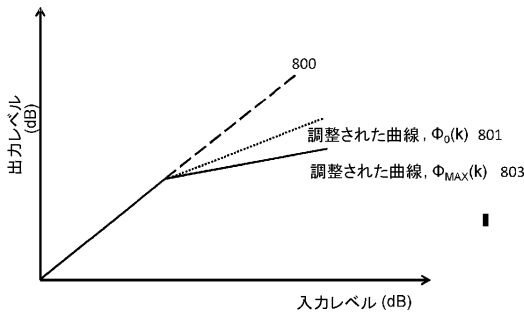
20

30

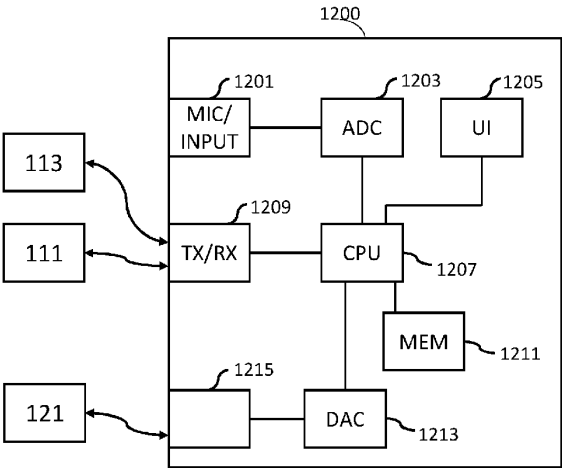
40

50

【図 8】



【図 9】



10

20

30

40

50

フロントページの続き

(72)発明者 ヴィロライネン ユッシ
 フィンランド共和国 02210 エスプー コロナカツ 5 C 20
合議体
 審判長 五十嵐 努
 審判官 梶本 剛
 審判官 田中 啓介
(56)参考文献 特開2015-233252 (JP, A)
 国際公開第2007/119330 (WO, A1)
 米国特許出願公開第2005/0080616 (US, A1)
(58)調査した分野 (Int.Cl., DB名)
 H04S 1/00 - 7/00