

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7023848号
(P7023848)

(45)発行日 令和4年2月22日(2022.2.22)

(24)登録日 令和4年2月14日(2022.2.14)

(51)国際特許分類		F I			
H 0 4 S	1/00 (2006.01)	H 0 4 S	1/00	5 0 0	
H 0 4 S	7/00 (2006.01)	H 0 4 S	7/00	3 0 0	

請求項の数 7 (全26頁)

(21)出願番号	特願2018-539144(P2018-539144)	(73)特許権者	507236292
(86)(22)出願日	平成29年1月26日(2017.1.26)		ドルビー ラボラトリーズ ライセンシン グ コーポレイション
(65)公表番号	特表2019-508947(P2019-508947 A)		アメリカ合衆国 9 4 1 0 3 カリフォル ニア州 サンフランシスコ マーケット ストリート 1 2 7 5
(43)公表日	平成31年3月28日(2019.3.28)	(73)特許権者	510185767
(86)国際出願番号	PCT/US2017/015165		ドルビー・インターナショナル・アー ー
(87)国際公開番号	WO2017/132396		オランダ王国, セーエン アムステルダム ズイドースト 1 1 0 1, ヘリケルベル グウェグ 1 - 3 5, アポロ ビルディン グ 3 エー
(87)国際公開日	平成29年8月3日(2017.8.3)	(74)代理人	100107766
審査請求日	令和2年1月27日(2020.1.27)		弁理士 伊東 忠重
(31)優先権主張番号	16153468.0		
(32)優先日	平成28年1月29日(2016.1.29)		
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		
(31)優先権主張番号	62/288,590		
(32)優先日	平成28年1月29日(2016.1.29)		
	最終頁に続く		最終頁に続く

(54)【発明の名称】 バイノーラル・ダイアログ向上

(57)【特許請求の範囲】

【請求項1】

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法であって、各コンポーネントは空間位置に関連付けられており、当該方法は：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を提供し；

第二のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第二のオーディオ信号呈示を提供し；

前記第一のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用し、前記ダイアログ・コンポーネントのダイアログ呈示を形成し；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせて、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、

前記第一および第二のオーディオ信号呈示のうち一方のみがバイノーラル・オーディオ信号呈示であり、前記ダイアログ推定パラメータは、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう呈示変換をも実行するよう構成されている、

方法。

【請求項 2】

前記第一および第二のオーディオ信号呈示のうち他方がステレオまたはサラウンド・オーディオ信号呈示である、請求項 1 記載の方法。

【請求項 3】

前記第一のオーディオ信号呈示の前記第二のオーディオ信号呈示への変換を可能にするよう構成された呈示変換パラメータの集合を受領し、前記第一のオーディオ信号呈示に呈示変換パラメータの前記集合を適用して前記第二のオーディオ信号呈示を形成することをさらに含む、請求項 1 または 2 記載の方法。

【請求項 4】

前記ダイアログ呈示に因子 G によるレベル修正を適用することをさらに含む、請求項 1 ないし 3 のうちいずれか一項記載の方法。

【請求項 5】

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法であって、各コンポーネントは空間位置に関連付けられており、当該方法は：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を受領し；

前記第一のオーディオ信号呈示の、第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示への変換を可能にするよう構成された呈示変換パラメータの集合を受領し；

前記第二のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し；

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ信号呈示を形成し；

ダイアログ推定パラメータの前記集合を前記第二のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成し；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と加算して、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である、
方法。

【請求項 6】

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダであって、各コンポーネントは空間位置に関連付けられており、当該デコーダは：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、前記第一のオーディオ信号呈示からダイアログ・コンポーネントを推定することを可能にするよう構成されたダイアログ推定パラメータの集合とを受領してデコードするコア・デコーダと；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と；

前記ダイアログ呈示を第二のオーディオ信号呈示と組み合わせて、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する手段とを有し；

前記第一および第二のオーディオ信号呈示のうち一方のみがバイノーラル・オーディオ信号呈示であり、前記ダイアログ推定パラメータは、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう呈示変換をも実行するよう構成されている、
デコーダ。

【請求項 7】

10

20

30

40

50

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダであって、各コンポーネントは空間位置に関連付けられており、当該デコーダは：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合と、前記第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領するコア・デコーダと；

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ再生システムでの再生のために意図された第二のオーディオ信号呈示を形成するよう構成された変換ユニットと；

ダイアログ推定パラメータの前記集合を前記第二のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と加算して、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する加算ブロックとを有し；

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示のうち的一方のみがバイノーラル・オーディオ信号呈示である、
デコーダ。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

本願は2016年1月29日に出願された米国仮特許出願第62/288,590号および2016年1月29日に提出された欧州特許出願第16153468.0号の優先権を主張するものである。両出願の内容はここに参照によってその全体において組み込まれる。

【0002】

発明の分野

本発明は、オーディオ信号処理の分野に関し、特に時に没入的オーディオ・コンテンツと称される立体音響化コンポーネントをもつオーディオ信号について、ダイアログ・コンポーネントの効率的な推定のための方法およびシステムを開示する。

【背景技術】

【0003】

本明細書を通じて背景技術のいかなる議論も、いかなる仕方でも、そのような技術が当分野において広く知られているまたは技術常識の一部をなすことの自認と考えられるべきではない。

【0004】

オーディオのコンテンツ生成、符号化、頒布および再生は、伝統的にはチャンネル・ベースのフォーマットで実行される。すなわち、コンテンツ・エコシステムを通じて一つの特定の目標再生システムが構想されている。そのような目標再生システム・オーディオ・フォーマットの例はモノ、ステレオ、5.1、7.1などであり、これらのフォーマットのことを、もとのコンテンツの異なる呈示 (presentation) と称する。上記の呈示は典型的にはラウドスピーカーを通じて再生されるが、注目すべき例外はステレオ呈示であり、これはヘッドフォンを通じて直接再生されることも多い。

【0005】

一つの具体的な呈示は、典型的にはヘッドフォンでの再生を目標にする、バイノーラル呈示である。バイノーラル呈示の特徴は、二チャンネル信号であって、各信号がそれぞれ左および右の鼓膜でまたは鼓膜近くで知覚されるコンテンツを表わすという点である。バイノーラル呈示は、ラウドスピーカーを通じて直接再生されることができ、好ましくは、

10

20

30

40

50

バイノーラル呈示は、クロストーク打ち消し技法を使ってラウドスピーカーを通じた再生に好適な呈示に変換される。

【0006】

種々の構成、たとえばステレオ、5.1、7.1におけるラウドスピーカーおよびヘッドフォンのような種々のオーディオ再生システムを上記で紹介した。上記の例から、もとのコンテンツの呈示は自然な、意図される、関連するオーディオ再生システムをもつが、もちろん、異なるオーディオ再生システムで再生されることもできることが理解される。

【0007】

コンテンツが意図されたものとは異なる再生システムで再生される場合、ダウンミックス〔下方混合〕またはアップミックス〔上方混合〕・プロセスが適用されることができる。たとえば、5.1コンテンツは、特定のダウンミックスの式を用いることによって、ステレオ再生システムで再生できる。もう一つの例は7.1スピーカー・セットアップでのステレオ・エンコードされたコンテンツの再生であり、これはいわゆるアップミックス・プロセスを含むことがあり、アップミックス・プロセスはステレオ信号に存在している情報によって案内されるまたはされないことができる。アップミックスできる一つのシステムは、ドルビー・ラボラトリーズ社からのDolby Pro Logicである (Roger Dressler, "Dolby Pro Logic Surround Decoder, Principles of Operation", www.Dolby.com)。

【0008】

代替的なオーディオ・フォーマット・システムは、Dolby Atmosシステムによって提供されるようなオーディオ・オブジェクト・フォーマットである。この型のフォーマットでは、オブジェクトまたはコンポーネントが聴取者のまわりの特定の位置をもつものとして定義される。該位置は時間変化してもよい。このフォーマットのオーディオ・コンテンツは時に、没入的オーディオ・コンテンツ (immersive audio content) と称される。本願のコンテキストの範囲内では、オーディオ・オブジェクト・フォーマットは上記のような呈示とは考えられず、むしろ、エンコーダにおいて一つまたは複数の呈示にレンダリングされるもとのコンテンツのフォーマットと考えられることを注意しておく。レンダリング後に、該呈示はエンコードされ、デコーダに伝送される。

【0009】

マルチチャンネルおよびオブジェクト・ベースのコンテンツが上述したようなバイノーラル呈示に変換されるとき、特定の諸位置におけるラウドスピーカーおよびオブジェクトからなる音響シーンは、頭部インパルス応答 (HRIR: head-related impulse response) または両耳室内インパルス応答 (BRIR: binaural room impulse response) によってシミュレートされる。これらは、それぞれ無響のまたは残響のある (シミュレートされた) 環境における各ラウドスピーカー/オブジェクトから鼓膜までの音響経路をシミュレートする。特に、オーディオ信号はHRIRまたはBRIRと畳み込みされて、両耳間レベル差 (ILD)、両耳間時間差 (ITD) およびスペクトル手がかりを復元することができ、これらが聴取者が個々の各ラウドスピーカー/オブジェクトの位置を判別することを許容する。音響環境 (残響) のシミュレーションは、知覚される距離を達成する助けにもなる。図1は、コンテンツ記憶部12から読まれる二つのオブジェクトまたはチャンネル信号 x_{j10} 、 x_{j11} を、四つのHRIR、たとえば14による処理のためにレンダリングするための処理フローの概略的な全体像を示している。HRIR出力はその後、それぞれのチャンネル信号について加算15、16され、ヘッドフォン18を介した聴取者への再生のためのヘッドフォン・スピーカー出力を生成する。HRIRの基本原理はたとえば、非特許文献1で説明されている。

【文献】Wightman, Frederic L., and Doris J. Kistler. "Sound localization." Human psychophysics. Springer New York, 1993. 155-192 HRIR/BRIR畳み込み手法にはいくつかの欠点に伴う。その一つは、ヘッドフォン再生のために必要とされるかなりの量の畳み込み処理である。HRIRまたはBRIR畳み込みは、すべての入力オブジェクトまたはチャンネルについて別個に適用される必要があり、よって計算量は典型的にはチャンネルまたはオブジェクトの数とともに線形に増大する。ヘッドフォンはしばしばバ

10

20

30

40

50

バッテリー駆動のポータブル装置との関連で使われるので、高い計算量は、バッテリー寿命をかなり短くしうるので、望ましくない。さらに、たとえば同時にアクティブな100を超えるオブジェクトを含みうるオブジェクト・ベースのオーディオ・コンテンツの導入で、HRIR畳み込みの計算量は、伝統的なチャンネル・ベースのコンテンツよりも実質的に高くなることがある。

【0010】

この目的のため、2015年8月25日に出願された、同時係属中の未公開の米国仮特許出願第62/209,735号は、ヘッドフォンのための没入的オーディオを効率的に伝送およびデコードするために使用できる呈示変換のためのデュアルエンドの手法を記載している。すべてのオブジェクトのレンダリングをデコーダのみに頼るのではなく、レンダリング・プロセスをエンコーダとデコーダの間で分割することによって、符号化効率およびデコード計算量削減が達成される。

10

【0011】

生成の際に特定の空間位置に関連付けられるコンテンツの一部は、オーディオ・コンポーネントと称される。空間位置は、空間内の点または分散された位置であることができる。オーディオ・コンポーネントは、サウンド・アーティストがサウンドトラック中にミキシングする、すなわち空間的に位置決めする個々のオーディオ源すべてと考えることができる。典型的には、内容的な意味（たとえばダイアログ）が関心対象のコンポーネントに割り当てられ、よって、処理（たとえばダイアログ向上）の目標が定義される。コンテンツ生成の間に生成されるオーディオ・コンポーネントは典型的には、もとのコンテンツから種々の呈示まで、処理チェーンを通じて存在していることを注意しておく。たとえば、オブジェクト・フォーマットでは、関連付けられた空間位置をもつダイアログ・オブジェクトがあることがある。ステレオ呈示では、水平面内に空間的に位置されたダイアログ・コンポーネントがあることがある。

20

【0012】

いくつかの応用では、オーディオ信号中のダイアログ・コンポーネントを抽出することが望ましい。たとえばそのようなコンポーネントを強調または増幅するためである。ダイアログ向上（DE: dialogue enhancement）の目標は、コンテンツのうち、発話と背景オーディオの混合を含む発話部分を修正して、発話がエンドユーザーにとって、より聞き取りやすくなるおよび/またはより疲れにくくなるようにすることであってもよい。DEのもう一つの用途は、たとえばエンドユーザーによってわずらわしいと知覚されるダイアログを減衰させることである。DE方法には、エンコーダ側およびデコーダ側という二つの基本的なクラスがある。デコーダ側DE（シングルエンドと呼ばれる）は、向上されていないオーディオを再構成するデコードされたパラメータおよび信号のみに対して作用する。すなわち、ビットストリームにはDEのための専用のサイド情報は存在しない。エンコーダ側DE（デュアルエンドと呼ばれる）では、デコーダにおいてDEを行なうために使用できる専用のサイド情報がエンコーダにおいて計算されて、ビットストリームに挿入される。

30

【0013】

図2は、通常のステレオ例におけるデュアルエンド・ダイアログ向上の例を示している。ここで、デコーダ24においてデコードされた非向上ステレオ信号23からダイアログ22を抽出できるようにする専用パラメータ21が、エンコーダ20において計算される。抽出されたダイアログは（部分的にはエンドユーザーによって制御される量だけ）レベル修正、たとえばブースト25され、非向上出力23に加えられて、最終的な出力26を形成する。専用パラメータ21は、非向上オーディオ27から盲目的に抽出されることができ、あるいはパラメータ計算において、別個に提供されるダイアログ信号28を活用することができる。

40

【0014】

もう一つの手法は特許文献1に記載されている。ここで、デコーダへのビットストリームは、オブジェクト・ダウンミックス信号（たとえばステレオ呈示）、オーディオ・オブジェクトの再構成を可能にするオブジェクト・パラメータおよび再構成されたオーディオ・

50

オブジェクトの操作を許容するオブジェクト・ベースのメタデータを含んでいる。特許文献1の図10に示されるように、操作は、発話に関係したオブジェクトの増幅を含んでもよい。このように、この手法は、デコーダ側におけるもとのオーディオ・オブジェクトの再構成を必要とするが、これは典型的には計算的に強い要求である。

【文献】米国特許第8,315,396号 バイノーラル・コンテキストでも効率的にダイアログ推定を提供することが一般に望まれている。

【発明の概要】

【発明が解決しようとする課題】

【0015】

バイノーラル・コンテキストにおいて、すなわち、ダイアログ・コンポーネント（単数または複数）が抽出されるもとなるオーディオ呈示または抽出されたダイアログが加えられるオーディオ呈示の少なくとも一方が（残響のあるまたは無響の）バイノーラル表現であるときに、効率的なダイアログ向上を提供することが本発明の目的である。

10

【課題を解決するための手段】

【0016】

本発明の第一の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法が提供される。各コンポーネントは空間位置に関連付けられており、本方法は、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を提供し、第二のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第二のオーディオ信号呈示を提供し、第一のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し、ダイアログ推定パラメータの集合を第一のオーディオ信号呈示に適用し、ダイアログ・コンポーネントのダイアログ呈示を形成し、ダイアログ呈示を第二のオーディオ信号呈示と組み合わせ、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、第一および第二のオーディオ信号呈示の少なくとも一方はバイノーラル・オーディオ信号呈示である。

20

【0017】

本発明の第二の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法が提供される。各コンポーネントは空間位置に関連付けられており、本方法は、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を受領し、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合を受領し、第一のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し、呈示変換パラメータの集合を第一のオーディオ信号呈示に適用して、第二のオーディオ信号呈示を形成し、ダイアログ推定パラメータの集合を第一のオーディオ信号呈示に適用してダイアログ・コンポーネントのダイアログ呈示を形成し、ダイアログ呈示を第二のオーディオ信号呈示と組み合わせ、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、第一のオーディオ信号呈示および第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である。

30

40

【0018】

本発明の第三の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法が提供される。各コンポーネントは空間位置に関連付けられており、本方法は、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を受領し、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合を受領し、第二のオーディオ信号呈示からのダイアログ・コンポーネントの

50

推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し、呈示変換パラメータの集合を第一のオーディオ信号呈示に適用して、第二のオーディオ信号呈示を形成し、ダイアログ推定パラメータの集合を第二のオーディオ信号呈示に適用してダイアログ・コンポーネントのダイアログ呈示を形成し、ダイアログ呈示を第二のオーディオ信号呈示と加算して、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、第一のオーディオ信号呈示および第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である。

【0019】

本発明の第四の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダが提供される。各コンポーネントは空間位置に関連付けられており、本デコーダは、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領してデコードするコア・デコーダと、ダイアログ推定パラメータの集合を第一のオーディオ信号呈示に適用してダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と、ダイアログ呈示を第二のオーディオ信号呈示と組み合わせて、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する手段とを有し、第一および第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である。

【0020】

本発明の第五の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダが提供される。各コンポーネントは空間位置に関連付けられており、本デコーダは、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合と、第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領するコア・デコーダと、呈示変換パラメータの集合を第一のオーディオ信号呈示に適用して、第二のオーディオ再生システムでの再生のために意図された第二のオーディオ信号呈示を形成するよう構成された変換ユニットと、ダイアログ推定パラメータの集合を第一のオーディオ信号呈示に適用してダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と、ダイアログ呈示を第二のオーディオ信号呈示と組み合わせて、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する手段とを有し、第一のオーディオ信号呈示および第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である。

【0021】

本発明の第六の側面によれば、一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダが提供される。各コンポーネントは空間位置に関連付けられており、本デコーダは、第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合と、第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領するコア・デコーダと、呈示変換パラメータの集合を第一のオーディオ信号呈示に適用して、第二のオーディオ再生システムでの再生のために意図された第二のオーディオ信号呈示を形成するよう構成された変換ユニットと、ダイアログ推定パラメータの集合を第二のオーディオ信号呈示に適用してダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と、ダイアログ呈示を第二のオーディオ信号呈示と加算して、第二のオーディオ再生シ

10

20

30

40

50

システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する加算ブロックとを有し、第一のオーディオ信号呈示および第二のオーディオ信号呈示のうちの一つがバイノーラル・オーディオ信号呈示である。

【0022】

本発明は、専用のパラメータ集合が、一つのオーディオ信号呈示からダイアログ呈示を抽出する効率的な仕方を提供しうるという洞察に基づいている。抽出されたダイアログ呈示はその後、別のオーディオ信号呈示と組み合わせられてもよい。ここで、それらの呈示の少なくとも一方はバイノーラル呈示である。本発明によれば、ダイアログを向上させるためにもとのオーディオ・オブジェクトを再構成する必要がない。その代わりに、オーディオ・オブジェクトの呈示、たとえばバイノーラル呈示、ステレオ呈示などに対して直接、専用のパラメータが適用される。本発明概念は、それぞれ個別的な利点をもつ多様な個別の実施形態を可能にする。

10

【0023】

ここでの「ダイアログ向上」という表現は、ダイアログ・コンポーネントを増幅またはブーストすることに制約されず、むしろ、選択されたダイアログ・コンポーネントの減衰にも関係しうることを注意しておく。よって、一般に、「ダイアログ向上」という表現は、オーディオ・コンテンツの一つまたは複数のダイアログに関係したコンポーネントのレベル修正を指す。レベル修正の利得因子Gは、ダイアログを減衰させるために0より小さくても、あるいはダイアログを強調するために0より大きくてもよい。

20

【0024】

いくつかの実施形態では、第一および第二の呈示はいずれも（残響のあるまたは無響の）バイノーラル呈示である。一方のみがバイノーラルである場合、他方の呈示はステレオまたはサラウンド・オーディオ信号呈示であってもよい。

【0025】

異なる呈示の場合、ダイアログ呈示が第二のオーディオ信号呈示に対応するよう、ダイアログ推定パラメータは呈示変換をも実行するよう構成されてもよい。

【0026】

本発明は、有利には、いわゆるサイマルキャスト・システムの特定の型で実装されてもよく、エンコードされたビットストリームは、第一のオーディオ信号呈示を第二のオーディオ信号呈示に変換するのに好適な変換パラメータの集合をも含む。

30

【図面の簡単な説明】

【0027】

本発明の実施形態について、これから単に例として、付属の図面を参照して記述する。

【図1】二つの音源またはオブジェクトについてのHRIR畳み込みプロセスの概略的な全体像を示す図である。各チャンネルまたはオブジェクトは一对のHRIR/BRIRによって処理される。

【図2】ステレオ・コンテキストにおけるダイアログ向上を概略的に示す図である。

【図3】本発明に基づくダイアログ向上の原理を示す概略的なブロック図である。

【図4】本発明のある実施形態に基づく、単一呈示ダイアログ向上の概略的なブロック図である。

40

【図5】本発明のさらなる実施形態に基づく、二呈示ダイアログ向上の概略的なブロック図である。

【図6】本発明のさらなる実施形態に基づく、図5におけるバイノーラル・ダイアログ推定器の概略的なブロック図である。

【図7】本発明のある実施形態に基づく、ダイアログ向上を実装するサイマルキャスト・デコーダの概略的なブロック図である。

【図8】本発明のもう一つの実施形態に基づく、ダイアログ向上を実装するサイマルキャスト・デコーダの概略的なブロック図である。

【図9】a、bは、本発明のさらにもう一つの実施形態に基づく、ダイアログ向上を実装するサイマルキャスト・デコーダの概略的なブロック図である。

50

【図10】本発明のさらにもう一つの実施形態に基づく、ダイアログ向上を実装するサイマルキャスト・デコーダの概略的なブロック図である。

【図11】本発明のさらにもう一つの実施形態に基づく、ダイアログ向上を実装するサイマルキャスト・デコーダの概略的なブロック図である。

【図12】本発明のさらにもう一つの実施形態を示す概略的なブロック図である。

【発明を実施するための形態】

【0028】

下記で開示されるシステムおよび方法は、ソフトウェア、ファームウェア、ハードウェアまたはそれらの組み合わせとして実装されうる。ハードウェア実装では、下記の記述において「段」と称されるタスクの分割は必ずしも物理的なユニットへの分割に対応するものではない。逆に、一つの物理的コンポーネントが複数の機能を有してもよく、一つのタスクが協働するいくつかの物理的コンポーネントによって実行されてもよい。ある種のコンポーネントまたはすべてのコンポーネントは、デジタル信号プロセッサもしくはマイクロプロセッサによって実行されるソフトウェアとして実装されてもよく、あるいはハードウェアとして実装されてもよく、あるいは特定用途向け集積回路として実装されてもよい。そのようなソフトウェアは、コンピュータ記憶媒体（または非一時的な媒体）および通信媒体（または一時的な媒体）を含みうるコンピュータ可読媒体上で頒布されてもよい。当業者にはよく知られているように、コンピュータ記憶媒体という用語は、コンピュータ可読命令、データ構造、プログラム・モジュールまたは他のデータといった情報の記憶のための任意の方法または技術で実装された、揮発性および不揮発性、リムーバブルおよび非リムーバブル媒体を含む。コンピュータ記憶媒体は、これに限られないが、RAM、ROM、EEPROM、フラッシュメモリまたは他のメモリ技術、CD-ROM、デジタル多用途ディスク（DVD）または他の光ディスク記憶、磁気カセット、磁気テープ、磁気ディスク記憶または他の磁気記憶デバイスまたは他の任意の媒体であって所望される情報を記憶するために使用されることができ、コンピュータによってアクセスされることができるものを含む。さらに、当業者には、通信媒体が典型的にはコンピュータ可読命令、データ構造、プログラム・モジュールまたは他のデータを、搬送波または他の転送機構のような変調されたデータ信号において具現し、任意の情報送達媒体を含むことはよく知られている。

【0029】

本発明の実施形態を実装するさまざまな仕方が、図3～図6を参照して論じられる。これらすべての実施形態は概括的には、一つまたは複数のオーディオ・コンポーネントをもつ入力オーディオ信号にダイアログ向上を適用するためのシステムおよび方法に関する。各コンポーネントは空間位置に関連付けられている。示される諸ブロックは典型的にはデコーダにおいて実装される。

【0030】

呈示される実施形態では、入力信号は好ましくは、たとえばフィルタバンク、たとえば直交ミラー・フィルタ（QMF）、離散フーリエ変換（DFT）、離散コサイン変換（DCT）または入力信号を多様な周波数帯域に分割する他の任意の手段によって、時間/周波数タイムに分解される。そのような変換の結果は、インデックス i および離散時間インデックス n をもつ入力についての入力信号 $x_i[n]$ が時間スロット（またはフレーム） k およびサブバンド b についてのサブバンド信号 $x_i[b,k]$ によって表わされる、ということである。たとえば、ステレオ呈示からのバイノーラル・ダイアログ呈示の推定を考える。 $x_j[b,k]$ 、 $j=1,2$ が左および右のステレオ・チャンネルのサブバンド信号を表わし、 \wedge 付きの $d_i[b,k]$ 、 $i=1,2$ が推定された左および右のバイノーラル・ダイアログ信号のサブバンド信号を表わすとする。ダイアログ推定は次のように計算されてもよい。

【0031】

【数1】

10

20

30

40

50

$$\hat{d}_i[b, k] = \sum_{m=0}^{M-1} \sum_{j=1}^J w_{ijm}^{B_p, K} x_j[b, k - m], i = 1, 2, b \in B_p, k \in K, p = 1, \dots, P$$

ここで、 B_p 、 K は所望される時間／周波数タイルに対応する周波数（ b ）および時間（ k ）インデックスの集合であり、 p はパラメータ帯域インデックスであり、 m は畳み込みタップ・インデックスであり、 $w_{ijm}^{B_p, K}$ は入力インデックス j 、パラメータ帯域 B_p 、サンプル範囲もしくは時間スロット K 、出力インデックス i および畳み込みタップ・インデックス m に属する行列係数である。上記の定式化を使うと、ダイアログは（ステレオ信号に関し；このステレオ信号の場合は $J=2$ ）パラメータ w によってパラメータ表現される（parameterized）。集合 K における時間スロットの数は周波数とは独立であり、周波数に対しては定数であり、典型的には時間区間5～40msに対応するよう選ばれる。周波数インデックスの集合の数 P は典型的には1～25の間であり、各集合における周波数インデックスの数は典型的には、聴覚の特性を反映して、周波数が増すとともに増大する（低周波数のほうがパラメータ表現における周波数分解能が高い）。

10

【0032】

ダイアログ・パラメータ w は、エンコーダにおいて計算され、ここに参照によって組み込まれる2015年8月25日に出願された米国仮特許出願第62/209,735号に開示される技法を使ってエンコードされてもよい。これらのパラメータ w は次いでビットストリームにおいて伝送され、デコーダによってデコードされ、その後、上記の式を使って適用される。推定の線形性のため、目標信号（きれいなダイアログまたはきれいなダイアログの推定）が利用可能である場合には、エンコーダ計算は、最小平均平方誤差（MMSE）方法を使って実装されることができる。

20

【0033】

P の選択および K における時間スロット数の選択は品質とビットレートとの間のトレードオフになる。さらに、パラメータ w は、（より低い品質を代償として）ビットレートを下げるために、たとえば $i=j$ のときには $w_{ijm}^{B_p, K} = 0$ と想定し、これらのパラメータは単に伝送しないことによって、制約されることができる。 M の選択も品質／ビットレートのトレードオフである。ここに参照によって組み込まれる2015年8月25日に出願された米国特許出願第62/209,742号参照。信号のバイノーラル化はITD（位相差）を導入するので、パラメータ w は一般に複素数値である。しかしながら、パラメータは、ビットレートを下げるために、実数値であると制約されることができる。さらに、人間が、1.5～2kHzあたりの位相／絶対値カットオフ周波数というある周波数より上では左右の信号の間の位相差および時間差に敏感でないことはよく知られている。よって、その周波数より上では、バイノーラル処理は典型的には、左右のバイノーラル信号の間に位相差が導入されないようになされ、よってパラメータは品質の損失なしに実数値であることができる（非特許文献2参照）。上記の品質／ビットレート・トレードオフは、各時間／周波数タイルにおいて独立に行なうことができる。

30

【文献】Breebaart, J., Nater, F., Kohlrausch, A. (2010). Spectral and spatial parameter resolution requirements for parametric, filter-bank-based HRTF processing. J. Audio Eng. Soc., 58 No 3, p.126-140 一般に、次の形の推定器を使うことが提案される。

40

【0034】

【数2】

$$\hat{y}_i[b, k] = \sum_{m=0}^{M-1} \sum_{j=1}^J w_{ijm}^{B_p, K} x_j[b, k - m], i = 1, \dots, I, b \in B_p, k \in K, p = 1, \dots, P$$

50

ここで、 \hat{y} および x の少なくとも一方がバイノーラル信号である。すなわち、 $l=2$ または $J=2$ または $l=J=2$ である。記法の便宜のため、以下では、ダイアログを推定するために使われる種々のパラメータ集合に言及するときに、しばしば時間/周波数タイルのインデックス B_p 、 K および i, j, m インデックスを省略する。

【0035】

上記の推定器は、行列記法で便利に次のように表現できる（記法の簡単のため時間/周波数タイル・インデックスは省略）。

【0036】

【数3】

$$\hat{Y} = \sum_{m=0}^{M-1} X_m W_m$$

10

ここで、

【0037】

【数4】

$$X_m = [x_1(m) \quad \dots \quad x_J(m)] \text{ および } \hat{Y} = [\hat{y}_1 \quad \dots \quad \hat{y}_I]$$

20

はそれぞれ $x_j[b, k - m]$ および $\hat{y}_i[b, k]$ のベクトル化されたバージョンを列に含んでおり、 W_m は J 行 I 列のパラメータ行列である。推定器の上記の形は、ダイアログ抽出だけを実行するときまたは呈示変換だけを実行するときならびに抽出および呈示変換の両方がパラメータの単一の集合を使ってなされるときに使用されうる。これについては下記の実施形態で詳述される。

【0038】

図3を参照するに、第一のオーディオ信号呈示31は、複数の空間化されたオーディオ・コンポーネントを含む没入型オーディオ信号からレンダリングされている。この第一のオーディオ信号呈示は、一つまたは複数の抽出されたダイアログ・コンポーネントの呈示33を提供するために、ダイアログ推定器32に提供される。ダイアログ推定器32は、ダイアログ推定パラメータ34の専用の集合を提供される。ダイアログ呈示は、利得ブロック35によってレベル修正（たとえばブースト）され、次いで、オーディオ信号の第二の呈示36と組み合わせられて、ダイアログ向上出力37を形成する。のちに論じるように、組み合わせは単純な加算でもよいが、ダイアログ呈示と第一の呈示の加算後に和に変換を適用して、それによりダイアログ向上された第二の呈示を形成することをも含んでいてもよい。

30

【0039】

本発明によれば、呈示の少なくとも一つがバイノーラル呈示（残響ありまたは無響）である。下記でさらに論じるように、第一および第二の呈示は異なってもよく、ダイアログ呈示は第二の呈示に対応してもしなくてもよい。たとえば、第一のオーディオ信号呈示は第一のオーディオ再生システム、たとえば一組のラウドスピーカーでの再生のために意図されていてもよく、一方、第二のオーディオ信号呈示は第二のオーディオ再生システム、たとえばヘッドフォンでの再生のために意図されていてもよい。

40

【0040】

単一の呈示

図4でのデコーダ実施形態では、第一および第二の呈示41、46ならびにダイアログ呈示43はみな（残響のあるまたは無響の）バイノーラル呈示である。よって、（バイノーラル）ダイアログ推定器42 および専用のパラメータ44 はバイノーラル・ダイアログ・コンポーネントを推定し、それが、ブロック45でレベル修正されて第二のオーディオ呈示46に加えられて出力47を形成する。

50

【 0 0 4 1 】

図 4 の実施形態では、パラメータ 4 4 はいかなる呈示変換を実行するようにも構成されない。それでも、最良品質のためには、バイノーラル・ダイアログ推定器 4 2 は、位相 / 絶対値カットオフ周波数までの周波数帯域では複素数値であるべきである。呈示変換がされないときでもなぜ複素数値の推定器が必要とされるかを説明するために、バイノーラル・ダイアログと他のバイノーラル背景コンテンツとの混合であるバイノーラル信号からバイノーラル・ダイアログを推定することを考える。ダイアログの最適な抽出はしばしば、たとえば右のバイノーラル信号の諸部分を左のバイノーラル信号から減算して、背景コンテンツを打ち消すことを含む。バイノーラル処理は、その性質上、左右の信号の間の時間（位相）差を導入するので、何らかの減算ができる前に、それらの位相差が補償される必要がある、そのような補償は複素数値のパラメータを必要とするのである。実際、パラメータの MMSE 計算の結果を調べるとき、パラメータは一般に、実数値であるよう制約されなければ、複素数値として現われる。実際上は、複素数値のパラメータか実数値のパラメータかの選択は、品質とビットレートとの間のトレードオフである。上述したように、パラメータは、高周波数での微細構造波形位相差に敏感でないことを利用して、全く品質損失なしに、周波数位相 / 絶対値カットオフ周波数より上では実数値であることができる。

10

【 0 0 4 2 】

二つの呈示

図 5 のデコーダ実施形態では、第一および第二の呈示が異なっている。図示した例では、第一の呈示 5 1 は非バイノーラル呈示であり（たとえばステレオ 2.0 またはサラウンド 5.1）、一方、第二の呈示 5 6 はバイノーラル呈示である。この場合、ダイアログ推定パラメータ 5 4 の集合は、バイノーラル・ダイアログ推定器 5 2 が、非バイノーラル呈示 5 1 からバイノーラル・ダイアログ呈示 5 3 を推定することを許容するように構成される。呈示は逆にしてもよいことを注意しておく。その場合、バイノーラル・ダイアログ推定器はたとえばバイノーラル・オーディオ呈示からステレオ・ダイアログ呈示を推定することになる。いずれの場合にも、ダイアログ推定器は、ダイアログ・コンポーネントを抽出し、呈示変換を実行する必要がある。バイノーラル・ダイアログ呈示 5 3 はブロック 5 5 によってレベル修正され、第二の呈示 5 6 に加えられる。

20

【 0 0 4 3 】

図 5 に示されるように、バイノーラル・ダイアログ推定器 5 2 は、ダイアログ抽出および呈示変換という二つの動作を実行するために構成された、パラメータ 5 4 の単一の集合を受領する。しかしながら、図 6 に示されるように、（残響のあるまたは無響の）バイノーラル・ダイアログ推定器 6 2 がパラメータの二つの集合 D1、D2 を受領して、一つの集合（D1）がダイアログを抽出するよう構成され（ダイアログ抽出パラメータ）、一つの集合（D2）がダイアログ呈示変換を実行するよう構成される（ダイアログ変換パラメータ）ことも可能である。これは、これらの部分集合 D1、D2 の一方または両方がすでにデコーダにおいて利用可能である実装において有利でありうる。たとえば、ダイアログ抽出パラメータ D1 は、図 2 に示される通常のダイアログ抽出のために利用可能であることがある。さらに、パラメータ変換パラメータ D2 は、のちに論じるように、サイマルキャスト実装において利用可能であることがある。

30

40

【 0 0 4 4 】

図 6 では、ダイアログ抽出（ブロック 6 2 a）が、呈示変換（ブロック 6 2 b）より前に行なわれるものとして示されているが、この順序はむしろ、逆にされてもよい。計算効率の理由で、たとえパラメータが二つの別個の集合 D1、D2 として提供されるとしても、まずパラメータの二つの集合を一つの組み合わせされた行列変換に組み合わせ、その後、この組み合わせられた変換を入力信号 6 1 に適用することが有利であることがある。

【 0 0 4 5 】

さらに、ダイアログ抽出は一次元的であり、抽出されたダイアログがモノ表現となることができることを注意しておく。すると、変換パラメータ D2 は位置メタデータであり、呈示変換はモノ・ダイアログを、その位置に対応する HRTF、HRIR または BRIR を使ってレン

50

ダリングすることを含む。あるいはまた、所望されるレンダリングされたダイアログ呈示がラウドスピーカー再生のために意図されている場合、モノ・ダイアログは、振幅パンまたはベクトル基底振幅パン (VBAP: vector-based amplitude panning) のようなラウドスピーカー・レンダリング技法を使ってレンダリングされることができる。

【 0 0 4 6 】

サイマルキャスト実装

図 7 ~ 図 1 1 は、サイマルキャスト・システム、すなわち一つのオーディオ呈示が一組の変換パラメータと一緒にエンコードされ、デコーダに伝送されるシステムのコンテキストで本発明の実施形態を示している。それらの変換パラメータはデコーダが前記オーディオ呈示を、意図された再生システムに適応した異なる呈示 (たとえば、ヘッドフォンのためのバイノーラル呈示として示される) に変換できるようにする。そのようなシステムのさまざまな側面は、ここに参照によって組み込まれる、2015年8月25日に出願された、同時係属中の未公開の米国仮特許出願第62/209,735号において詳細に記述されている。簡単のため、図 7 ~ 図 1 1 はデコーダ側のみを示している。

10

【 0 0 4 7 】

図 7 に示されるように、コア・デコーダ 7 1 は、オーディオ・コンポーネントの初期のオーディオ信号呈示を含むエンコードされたビットストリーム 7 2 を受領する。図示されている場合、この初期呈示はステレオ呈示 z であるが、他の任意の呈示であってもよい。ビットストリーム 7 2 は、呈示変換パラメータ $w(y)$ の集合をも含んでいる。これらのパラメータは、ステレオ信号 z の行列変換 7 3 を実行して、再構成された無響バイノーラル信号 \hat{y} を生成するために、行列係数として使われる。変換パラメータ $w(y)$ は、米国仮特許出願第 62/209,735号で論じられているように、エンコーダで決定されたものである。図示されている場合、ビットストリーム 7 2 は、ステレオ信号 z の行列変換 7 4 を実行して、無響環境シミュレーション、ここではフィードバック遅延ネットワーク (FDN) 7 5 のための再構成された入力信号 \hat{f} を生成するための行列係数として使われるパラメータ $w(f)$ の集合をも含んでいる。これらのパラメータ $w(f)$ は、呈示変換パラメータ $w(y)$ と同様の仕方で決定されたものである。FDN 7 5 は入力信号 \hat{f} を受領し、音響環境シミュレーション出力 FDN_{out} を提供し、該出力は無響バイノーラル信号 \hat{y} と組み合わせられて、残響のあるバイノーラル信号を提供してもよい。

20

【 0 0 4 8 】

図 7 の実施形態において、ビットストリームはさらに、ステレオ信号 z の行列変換を実行して無響のバイノーラル・ダイアログ呈示 D を生成するためにダイアログ推定器 7 6 において行列係数として使われるダイアログ推定パラメータ $w(D)$ の集合を含んでいる。ダイアログ呈示 D はブロック 7 7 においてレベル修正 (たとえばブースト) され、再構成された無響信号 \hat{y} および無響環境シミュレーション出力 FDN_{out} と、加算ブロック 7 8 において組み合わせられる。

30

【 0 0 4 9 】

図 7 は、本質的には、図 5 の実施形態のサイマルキャスト・コンテキストにおける実装である。

【 0 0 5 0 】

図 8 の実施形態では、ステレオ信号 z 、変換パラメータ $w(y)$ の集合およびパラメータ $w(f)$ のさらなる集合が受領され、図 7 と同じようにデコードされ、要素 7 1、7 3、7 4、7 5、7 8 は図 7 に関して論じたものと等価である。さらに、ここでのビットストリーム 8 2 は、信号 z に対してダイアログ推定器 8 6 によって適用されるダイアログ推定パラメータ $w(D1)$ の集合をも含んでいる。しかしながら、この実施形態では、ダイアログ推定パラメータ $w(D1)$ は、いかなる呈示変換をも提供するよう構成されていない。したがって、ダイアログ推定器 8 6 からのダイアログ呈示出力 D_{stereo} は、初期オーディオ信号呈示、ここではステレオ呈示に対応する。ダイアログ呈示 D_{stereo} はブロック 8 7 においてレベル修正され、次いで加算 8 8 において信号 z に加えられる。ダイアログ向上された信号 ($z + D_{stereo}$) は次いで、変換パラメータ $w(y)$ の集合によって変換される。

40

50

【 0 0 5 1 】

図 8 は、図 6 の実施形態のサイマルキャスト・コンテキストでの実装として見る事ができる。ここで、 $w(D1)$ が $D1$ として使われ、 $w(y)$ が $D2$ として使われる。しかしながら、図 6 ではパラメータの両方の集合がダイアログ推定器 6 2 において適用される一方、図 8 では、抽出されたダイアログ D_{stereo} は信号 z に加えられ、変換 $w(y)$ は組み合わせられた信号 $(z + D)$ に適用される。

【 0 0 5 2 】

パラメータ $w(D1)$ の集合が、サイマルキャスト実装においてステレオ信号のダイアログ向上を提供するために使われるダイアログ向上パラメータと同一であってもよいことを注意しておく。この代替が図 9 の a に示されている。ここでは、ダイアログ抽出 9 6 a がコア・デコーダ 9 1 の一部をなすものとして示されている。さらに、図 9 の a では、パラメータ集合 $w(y)$ を使う呈示変換 9 6 b が利得の前に、信号 z の変換とは別個に実行される。このように、この実施形態は、ダイアログ推定器 6 2 が両方の変換 9 6 a、9 6 b を含んでいて、図 6 に示される場合とさらによく似ている。

10

【 0 0 5 3 】

図 9 の b は、図 9 の a の実施形態の修正版を示している。この場合、呈示変換はパラメータ集合 $w(y)$ を使ってではなく、バイノーラル・ダイアログ推定専用のビットストリーム部分において提供されているパラメータ $w(D2)$ の追加的な集合を用いて実行される。

【 0 0 5 4 】

ある実施形態では、図 9 の b における上述した専用の呈示変換 $w(D2)$ は実数値であり、単一タップ ($M = 1$) であり、フルバンド ($P = 1$) の行列である。

20

【 0 0 5 5 】

図 10 は、図 9 の a ~ b の実施形態の修正版を示している。この場合、ダイアログ抽出器 9 6 a はやはりステレオ・ダイアログ呈示 D_{stereo} を提供し、やはりコア・デコーダ 9 1 の一部をなすものとして示されている。しかしながら、ここでは、ステレオ・ダイアログ呈示 D_{stereo} は、ブロック 9 7 でのレベル修正後に、(FDNからの無響の環境シミュレーションとともに) 無響バイノーラル信号 \hat{y} に直接加えられる。

【 0 0 5 6 】

異なる呈示をもつ信号を組み合わせること、たとえばステレオ・ダイアログ信号を(向上されていないバイノーラル・ダイアログ・コンポーネントを含む)バイノーラル信号に加算することは、当然ながら、空間定位のアーチファクトにつながることを注意しておく。向上されていないバイノーラル・ダイアログ・コンポーネントは、同じコンポーネントのステレオ呈示に比べて空間的に異なっていると知覚されるからである。

30

【 0 0 5 7 】

さらに、異なる呈示をもつ信号を組み合わせることは、ある周波数帯域ではダイアログ・コンポーネントの強め合う加算に、他の周波数帯域では弱め合う加算につながることを注意しておく。その理由は、バイノーラル処理が ITD (位相差) を導入し、ある周波数帯域では同相であり、他の帯域では逆相である信号を加算し、それがダイアログ・コンポーネントにおける音色付け (coloring) アーチファクトにつながるからである (さらに、音色付けは左耳と右耳とで異なることがある)。ある実施形態では、この型のアーチファクトを低減するよう、バイノーラル処理において、位相/絶対値カットオフ周波数より上での位相差は回避される。

40

【 0 0 5 8 】

異なる呈示をもつ信号を組み合わせる場合についての最後の注釈として、一般に、バイノーラル処理はダイアログの了解性を低下させることがあることが認められる。ダイアログ向上の目標が了解性を最大にすることである場合、非バイノーラルであるダイアログ信号を抽出し、レベル修正 (たとえばブースト) することが有利であることがある。より具体的には、たとえ再生のために意図される最終的な呈示がバイノーラルであっても、そのような場合にステレオ・ダイアログ信号を抽出してレベル修正 (たとえばブースト) して、それをバイノーラル呈示と組み合わせることが有利であることがある (了解性向上のため

50

に、上記のように、音色付けアーチファクトと空間定位アーチファクトとをトレードオフ)。

【0059】

図11の実施形態では、ステレオ信号 z 、変換パラメータ $w(y)$ の集合およびパラメータ $w(f)$ のさらなる集合が図7と同じようにして受領され、デコードされる。さらに、図8と同様に、ビットストリームは、いかなる呈示変換も提供しないダイアログ推定パラメータ $w(D1)$ の集合をも含んでいる。しかしながら、この実施形態では、ダイアログ推定パラメータ $w(D1)$ はダイアログ推定器16によって、再構成された無響バイノーラル信号 \hat{y} に対して適用されて、無響バイノーラル・ダイアログ呈示 D を提供する。このダイアログ呈示 D はブロック117によってレベル修正され、加算118において FDN_{out} とともに信号 \hat{y} に加えられる。

10

【0060】

図11は、本質的には、図5の単一呈示の実施形態の、サイマルキャスト・コンテキストでの実装である。しかしながら、それは、図6の実装の $D1$ と $D2$ の順序を逆にしたものと見することもできる。ここで、 $D1$ としてはやはり $w(D1)$ は使われ、 $D2$ として $w(y)$ が使われる。しかしながら、図6ではパラメータの両方の集合がダイアログ推定器において適用されたのに対して、図9では、変換パラメータ $D2$ は、 \hat{y} を得るためにすでに適用されており、ダイアログ推定器16は、残響のあるバイノーラル・ダイアログ呈示 D を得るために、パラメータ $w(D1)$ を信号 \hat{y} に適用する必要があるだけである。

【0061】

いくつかの応用では、ダイアログ・レベル修正因子 G の所望される値に依存して、異なる処理を適用することが望ましいことがある。ある実施形態では、例示的な適切な処理が、因子 G が所与の閾値より大きいか小さいかの判定に基づいて選択される。もちろん、二つ以上の閾値および二つ以上の代替的な処理があってもよい。たとえば、 $th1$ および $th2$ が二つの所与の閾値であるとして、 $G < th1$ のときの第一の処理、 $th1 \leq G < th2$ のときの第二の処理および $G \geq th2$ のときの第三の処理である。

20

【0062】

図12に示される個別的な例では、閾値は0であり、 $G < 0$ (ダイアログの減衰) のときには第一の処理が適用され、 $G > 0$ (ダイアログの強調) のときには第二の処理が適用される。この目的のために、図12の回路は、二つのポジションAおよびBをもつスイッチ121の形の選択論理を含む。スイッチは、ブロック122から利得因子 G の値を提供され、 $G < 0$ のときはポジションAを、 $G > 0$ のときはポジションBを取るよう構成される。

30

【0063】

スイッチがポジションAにあるとき、回路はここでは、行列変換86からの推定されたステレオ・ダイアログをステレオ信号 z と組み合わせ、次いで、組み合わせられた信号に対して行列変換73を実行して、再構成された無響バイノーラル信号を生成するよう構成される。フィードバック遅延ネットワーク75からの出力が次いで、この信号と78において組み合わせられる。この処理は本質的には、上記で論じた図8に対応することを注意しておく。

【0064】

スイッチがポジションBにあるとき、回路はここでは、バイノーラル・ダイアログ推定を提供するために、行列変換86からのステレオ・ダイアログに変換パラメータ $w(D2)$ を適用するよう構成される。次いで、この推定が変換73からの無響のバイノーラル信号およびフィードバック遅延ネットワーク75からの出力に加えられる。この処理は本質的には、上記で論じた図9のbに対応することを注意しておく。

40

【0065】

当業者は、ポジションAおよびBにおけるそれぞれの処理について他の多くの代替を認識するであろう。たとえば、スイッチがポジションBにあるときの処理は上記の代わりに図10のものに対応してもよい。しかしながら、図12の実施形態の主要な貢献は、スイッチ121の導入である。これが利得因子 G の値に依存した代替的な処理を可能にする。

【0066】

50

解釈

本明細書を通じて「一つの実施形態」「いくつかの実施形態」または「ある実施形態」への言及は、その実施形態との関連で記載されている特定の特徴、構造または特性が本発明の少なくとも一つの実施形態に含まれることを意味する。よって、「一つの実施形態において」「いくつかの実施形態において」または「ある実施形態において」という句が本明細書を通じた随所に現われることは、必ずしもみなが同じ実施形態を指しているのではないが、指していてもよい。さらに、特定の特徴、構造または特性は、いかなる好適な仕方

【0067】

本稿での用法では、特に断わりのない限り、共通のオブジェクトを記述する順序形容語「第一」「第二」「第三」などの使用は、単に同様のオブジェクトの異なるインスタンスが言及されていることを示すものであって、そのように記述されるオブジェクトが時間的、空間的、ランキングにおいてまたは他のいかなる仕方においても、所与の序列でなければならぬことを含意することは意図されていない。

【0068】

付属の請求項および本稿の記述において、有する、から構成されるまたは含むという用語の任意のものは、少なくともその後続の要素/特徴を含むが他のものを排除しないことを意味するオープンな用語である。よって、請求項において使われるときの有するの用語は、その後

【0069】

本稿での用法では、用語「例示的」は、性質を示すのではなく、例を挙げる意味で使われる。すなわち、「例示的実施形態」は、必ず例示的な性質の実施形態であるのではなく、例として与えられている実施形態である。

【0070】

本発明の例示的実施形態の上記の記述において、開示の流れをよくし、さまざまな発明側面の一つまたは複数のものの理解を助けるため、本発明のさまざまな特徴が時に単一の実施形態、図面またはその記述にまとめられていることを理解しておくべきである。しかしながら、この開示法は、請求される発明が、各請求項に明示的に記載されているよりも多くの事項を必要とするという意図を反映したものと解釈されるものではない。むしろ、付属の請求項が反映するように、発明の諸側面は、単一の上記の開示される実施形態の全事項よりも少ないものに存する。このように、付属の請求項は、ここに明示的に詳細な説明に組み込まれ、各請求項がそれ自身として本発明の別個の実施形態をなす。

【0071】

さらに、本稿に記載されるいくつかの実施形態が他の実施形態に含まれるいくつかの特徴を含むが他の特徴を含まなくても、異なる実施形態の特徴の組み合わせは本発明の範囲内であり、異なる実施形態をなすことが意図されている。当業者はこれを理解するであろう。たとえば、付属の請求項では、請求される実施形態の任意のものが任意の組み合わせにおいて使用できる。

【0072】

さらに、実施形態のいくつかは、本稿では方法または方法の要素の組み合わせであって、コンピュータ・システムのプロセッサによってまたは該機能を実行する他の手段によって実装されることができるとして記述されている。よって、そのような方法または方法の要素を実行するための必要な命令をもつプロセッサは、前記方法または方法の要素を実行する手段をなす。さらに、装置実施形態の本稿に記載される要素は、本発明を実行するための該要素によって実行される機能を実行する手段の一例である。

10

20

30

40

50

【 0 0 7 3 】

本稿で与えられる記述では、数多くの個別的詳細が記載される。しかしながら、本発明の実施形態がそうした個別的詳細なしでも実施できることは理解される。他方、本記述の理解をかすませないために、よく知られた方法、構造および技法は詳細に示していない。

【 0 0 7 4 】

同様に、請求項において使われるときの用語、結合されたは、直接接続のみに限定されるものと解釈されるべきではない。用語「結合された」および「接続された」ならびにその派生形が使われることがある。これらの用語は互いの同義語として意図されていないことを理解しておくべきである。よって、装置Bに結合された装置Aという表現の範囲は、装置Aの出力が装置Bの入力に直接接続される装置またはシステムに限定されるべきではない。それは、Aの出力とBの入力との間の経路が存在することを意味し、該経路は他の装置または手段を含む経路であってもよい。「結合された」は二つ以上の要素が直接物理的または電氣的に接していること、あるいは二つ以上の要素が互いに直接接触してはいないが、それでも互いと協働または相互作用することを意味する。

10

【 0 0 7 5 】

このように、本発明の個別的実施形態を記述してきたが、当業者は本発明の精神から外れることなく、それに他のおよびさらなる修正がなされてもよいことを認識するであろう。それらすべての変更および修正を本発明の範囲内にはいるものとして請求することが意図されている。たとえば、上記で与えた公式はいずれも単に使用されうる手順の代表である。ブロック図から機能が追加または削除されてもよく、機能ブロックの間で動作が交換されてもよい。本発明の範囲内で記述される方法に段階が追加または削除されてもよい。

20

いくつかの態様を記載しておく。

〔態様1〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法であって、各コンポーネントは空間位置に関連付けられており、当該方法は：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を提供し；

第二のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第二のオーディオ信号呈示を提供し；

30

前記第一のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用し、前記ダイアログ・コンポーネントのダイアログ呈示を形成し；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせて、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、

前記第一および第二のオーディオ信号呈示の少なくとも一方はバイノーラル・オーディオ信号呈示である、

方法。

40

〔態様2〕

前記第一および第二のオーディオ信号呈示がバイノーラル・オーディオ信号呈示である、態様1記載の方法。

〔態様3〕

前記第一および第二のオーディオ信号呈示のうち一方のみがバイノーラル・オーディオ信号呈示である、態様1記載の方法。

〔態様4〕

前記第一および第二のオーディオ信号呈示のうち他方がステレオまたはサラウンド・オーディオ信号呈示である、態様3記載の方法。

〔態様5〕

50

ダイアログ変換パラメータの集合を受領し、ダイアログ推定パラメータの前記集合の適用の前または後にダイアログ変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成することをさらに含む、態様 3 または 4 記載の方法。

〔態様 6〕

前記ダイアログ推定パラメータは、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう呈示変換をも実行するよう構成されている、態様 3 または 4 記載の方法。

〔態様 7〕

前記第一のオーディオ信号呈示を提供することが、初期のオーディオ信号呈示および呈示変換パラメータの集合を受領し、呈示変換パラメータの前記集合を前記初期のオーディオ信号呈示に適用することを含む、態様 2 記載の方法。

10

〔態様 8〕

前記第一のオーディオ信号呈示の前記第二のオーディオ信号呈示への変換を可能にするよう構成された呈示変換パラメータの集合を受領し、前記第一のオーディオ信号呈示に呈示変換パラメータの前記集合を適用して前記第二のオーディオ信号呈示を形成することをさらに含む、態様 1 ないし 7 のうちいずれか一項記載の方法。

〔態様 9〕

ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成することをさらに含む、態様 8 記載の方法。

20

〔態様 10〕

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせることが、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成して、該和に、呈示変換パラメータの前記集合を適用することを含む、態様 8 記載の方法。

〔態様 11〕

前記第一のオーディオ信号呈示がエンコーダから受領される、態様 1 ないし 10 のうちいずれか一項記載の方法。

〔態様 12〕

前記ダイアログ呈示に因子 G によるレベル修正を適用することをさらに含む、態様 1 ないし 11 のうちいずれか一項記載の方法。

30

〔態様 13〕

G が所与の閾値より小さいときは第一の処理が適用され、G が前記閾値より大きいときは第二の処理が適用される、態様 12 記載の方法。

〔態様 14〕

前記閾値が 0 に等しく、 $G < 0$ はダイアログの減衰を表わし、 $G > 0$ はダイアログの強調を表わす、態様 13 記載の方法。

〔態様 15〕

前記第一の処理が、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成して、該和に、呈示変換パラメータの集合を適用することを含む、態様 13 または 14 記載の方法。

40

〔態様 16〕

前記第二の処理が、ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成することを含む、態様 13 ないし 15 のうちいずれか一項記載の方法。

〔態様 17〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法であって、各コンポーネントは空間位置に関連付けられており、当該方法は：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を受領し；

50

前記第一のオーディオ信号呈示の、第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示への変換を可能にするよう構成された呈示変換パラメータの集合を受領し；

前記第一のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し；

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ信号呈示を形成し；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成し；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせて、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み；

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である、

方法。

〔態様 18〕

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせることが、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成して、該和に、呈示変換パラメータの前記集合を適用することを含み、態様 17 記載の方法。

〔態様 19〕

前記ダイアログ推定パラメータは、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう呈示変換をも実行するよう構成されている、態様 17 記載の方法。

〔態様 20〕

ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成することをさらに含む、態様 17 記載の方法。

〔態様 21〕

前記ダイアログ呈示がモノ呈示であり、当該方法がさらに；

前記ダイアログ・コンポーネントに関係する位置データを受領し；

前記第二のオーディオ信号呈示と組み合わせる前に、前記位置データを使って、前記モノ・ダイアログ呈示をレンダリングすることをさらに含む、

態様 17 記載の方法。

〔態様 22〕

前記レンダリングすることが；

前記位置データに基づいてライブラリから頭部伝達関数 (HRTF) を選択し；

選択された HRTF を前記モノ・ダイアログ呈示に適用することを含む、

態様 21 記載の方法。

〔態様 23〕

前記レンダリングすることが、振幅パンを含む、態様 21 記載の方法。

〔態様 24〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するための方法であって、各コンポーネントは空間位置に関連付けられており、当該方法は；

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示を受領し；

前記第一のオーディオ信号呈示の、第二のオーディオ再生システムでの再生のために意図されている前記第二のオーディオ信号呈示への変換を可能にするよう構成された呈示変換パラメータの集合を受領し；

前記第二のオーディオ信号呈示からのダイアログ・コンポーネントの推定を可能にするよう構成されたダイアログ推定パラメータの集合を受領し；

10

20

30

40

50

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ信号呈示を形成し；

ダイアログ推定パラメータの前記集合を前記第二のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成し；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と加算して、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成することを含み、

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である、

方法。

〔態様 25〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダであって、各コンポーネントは空間位置に関連付けられており、当該デコーダは；

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、前記第一のオーディオ信号呈示からダイアログ・コンポーネントを推定することを可能にするよう構成されたダイアログ推定パラメータの集合とを受領してデコードするコア・デコーダと；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と；

前記ダイアログ呈示を第二のオーディオ信号呈示と組み合わせて、第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する手段とを有し；

前記第一および第二のオーディオ信号呈示の一方のみがバイノーラル・オーディオ信号呈示である、

デコーダ。

〔態様 26〕

前記第一および第二のオーディオ信号呈示のうち一方がステレオまたはサラウンド・オーディオ信号呈示である、態様 25 記載のデコーダ。

〔態様 27〕

前記コア・デコーダが、ダイアログ変換パラメータの集合を受領するようさらに構成され、前記ダイアログ推定器が、ダイアログ推定パラメータの前記集合の適用の前または後にダイアログ変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成するようさらに構成されている、態様 25 または 26 記載のデコーダ。

〔態様 28〕

前記ダイアログ推定器は、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう、ダイアログ推定パラメータの前記集合を使って呈示変換をも実行するよう構成されている、態様 25 または 26 記載のデコーダ。

〔態様 29〕

前記コア・デコーダが、呈示変換パラメータの集合を受領するようさらに構成されており、当該デコーダがさらに；

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して前記第二のオーディオ信号呈示を形成するよう構成されている変換ユニットを有する、

態様 25 ないし 28 のうちいずれか一項記載のデコーダ。

〔態様 30〕

前記ダイアログ推定器が、ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成するようさらに構成されている、態様 29 記載のデコーダ。

〔態様 31〕

10

20

30

40

50

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせる手段が、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成する加算ブロックを含み、前記変換ユニットが、該和に、呈示変換パラメータの前記集合を適用するよう構成されている、態様 29 記載のデコーダ。

〔態様 32〕

前記ダイアログ呈示に因子Gによるレベル修正を適用するよう構成されたレベル修正ブロックをさらに有する、態様 25 ないし 31 のうちいずれか一項記載のデコーダ。

〔態様 33〕

Gが所与の閾値より小さいときは前記ダイアログ推定パラメータの第一の適用を選択するよう構成された選択論理をさらに有しており、Gが前記閾値より大きいときは第二の処理が適用される、態様 32 記載のデコーダ。

10

〔態様 34〕

前記閾値が0に等しく、 $G < 0$ はダイアログの減衰を表わし、 $G > 0$ はダイアログの強調を表わす、態様 33 記載のデコーダ。

〔態様 35〕

前記第一の適用が、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成して、該和に、呈示変換パラメータの集合を適用することを含む、態様 33 または 34 記載のデコーダ。

〔態様 36〕

前記第二の適用が、ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成することを含む、態様 33 ないし 35 のうちいずれか一項記載のデコーダ。

20

〔態様 37〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダであって、各コンポーネントは空間位置に関連付けられており、当該デコーダは：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、前記第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合と、前記第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領するコア・デコーダと；

30

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ再生システムでの再生のために意図された第二のオーディオ信号呈示を形成するよう構成された変換ユニットと；

ダイアログ推定パラメータの前記集合を前記第一のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせて、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する手段とを有し；

40

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示のうち一方のみがバイノーラル・オーディオ信号呈示である、

デコーダ。

〔態様 38〕

前記ダイアログ呈示を前記第二のオーディオ信号呈示と組み合わせる手段が、前記ダイアログ呈示と前記第一のオーディオ信号呈示の和を形成する加算ブロックを含み、前記変換ユニットが、該和に、呈示変換パラメータの前記集合を適用するよう構成されている、態様 37 記載のデコーダ。

〔態様 39〕

50

前記ダイアログ推定器は、前記ダイアログ呈示が前記第二のオーディオ信号呈示に対応するよう、ダイアログ推定パラメータの前記集合を使って呈示変換をも実行するよう構成されている、態様 3 7 記載のデコーダ。

〔態様 4 0〕

前記ダイアログ推定器は、ダイアログ推定パラメータの前記集合の適用の前または後に呈示変換パラメータの前記集合を適用して、前記第二のオーディオ信号呈示に対応する変換されたダイアログ呈示を形成するよう構成されている、態様 3 7 記載のデコーダ。

〔態様 4 1〕

前記ダイアログ呈示がモノ呈示であり、前記コア・デコーダが、前記ダイアログ・コンポーネントに係する位置データを受領するようさらに構成されており、当該デコーダがさらに：

前記第二のオーディオ信号呈示と組み合わせる前に、前記位置データを使って、前記モノ・ダイアログ呈示をレンダリングする構成されたレンダラーをさらに有する、態様 3 7 記載のデコーダ。

〔態様 4 2〕

前記レンダラーが：

前記位置データに基づいてライブラリから頭部伝達関数 (HRTF) を選択し；
選択された HRTF を前記モノ・ダイアログ呈示に適用するよう構成されている、態様 4 1 記載のデコーダ。

〔態様 4 3〕

前記レンダラーが振幅パンを適用するよう構成されている、態様 4 1 記載のデコーダ。

〔態様 4 4〕

一つまたは複数のオーディオ・コンポーネントをもつオーディオ・コンテンツをダイアログ向上するためのデコーダであって、各コンポーネントは空間位置に関連付けられており、当該デコーダは：

第一のオーディオ再生システムでの再生のために意図されている前記オーディオ・コンポーネントの第一のオーディオ信号呈示と、第一のオーディオ信号呈示を第二のオーディオ再生システムでの再生のために意図されている第二のオーディオ信号呈示に変換できるようにするよう構成された呈示変換パラメータの集合と、前記第一のオーディオ信号呈示からダイアログ・コンポーネントを推定できるようにするよう構成されたダイアログ推定パラメータの集合とを受領するコア・デコーダと；

呈示変換パラメータの前記集合を前記第一のオーディオ信号呈示に適用して、第二のオーディオ再生システムでの再生のために意図された第二のオーディオ信号呈示を形成するよう構成された変換ユニットと；

ダイアログ推定パラメータの前記集合を前記第二のオーディオ信号呈示に適用して前記ダイアログ・コンポーネントのダイアログ呈示を形成するダイアログ推定器と；

前記ダイアログ呈示を前記第二のオーディオ信号呈示と加算して、前記第二のオーディオ再生システムでの再生のためのダイアログ向上されたオーディオ信号呈示を形成する加算ブロックとを有し；

前記第一のオーディオ信号呈示および前記第二のオーディオ信号呈示のうち的一方のみがバイノーラル・オーディオ信号呈示である、デコーダ。

10

20

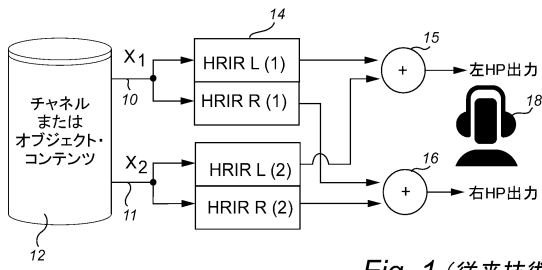
30

40

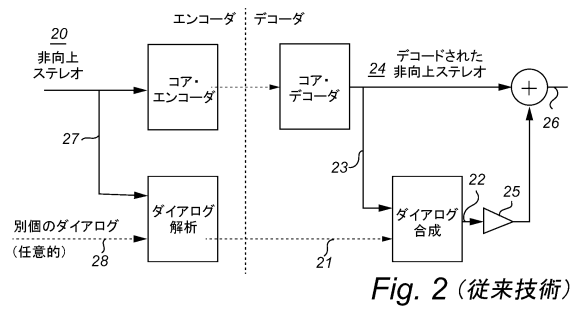
50

【図面】

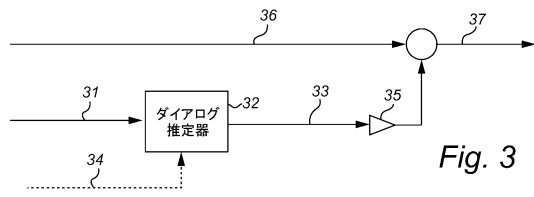
【図 1】



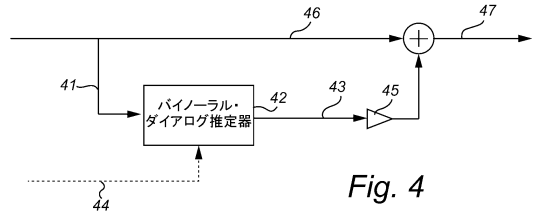
【図 2】



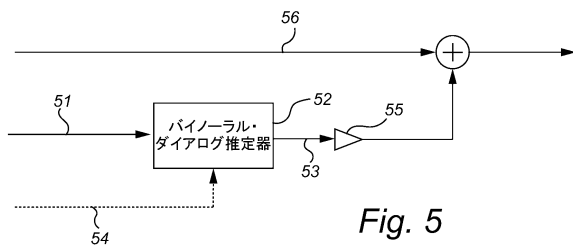
【図 3】



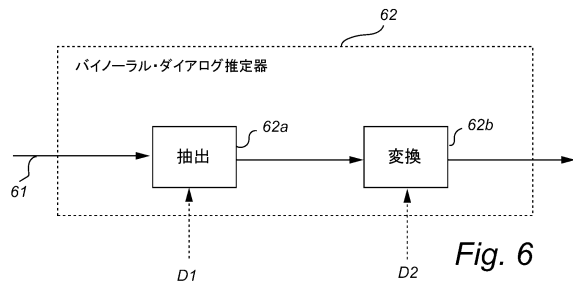
【図 4】



【図 5】



【図 6】



10

20

30

40

50

【図 7】

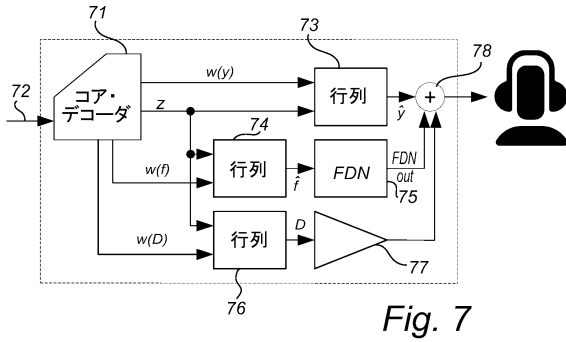


Fig. 7

【図 8】

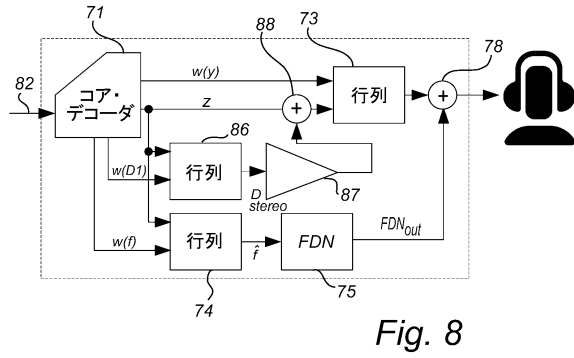


Fig. 8

10

【図 9 a】

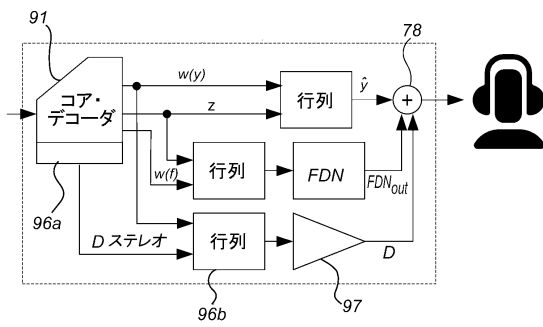


Fig. 9a

【図 9 b】

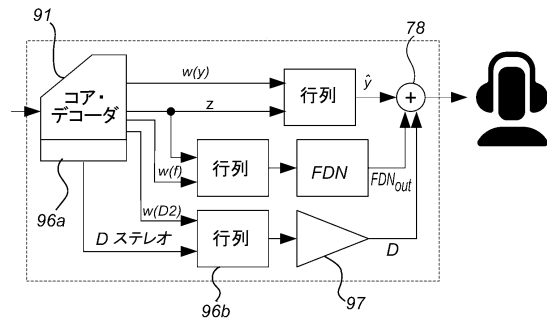


Fig. 9b

20

【図 10】

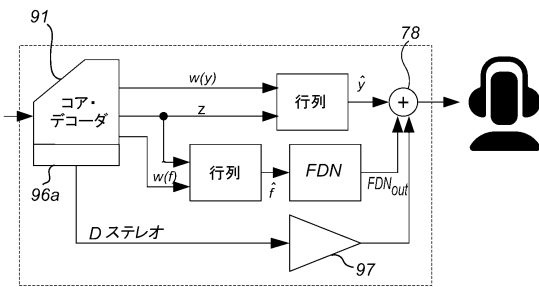


Fig. 10

【図 11】

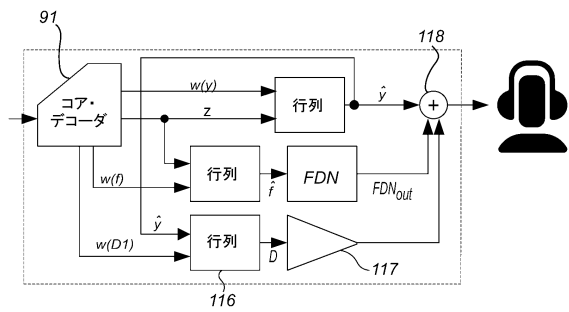


Fig. 11

30

40

50

【 図 1 2 】

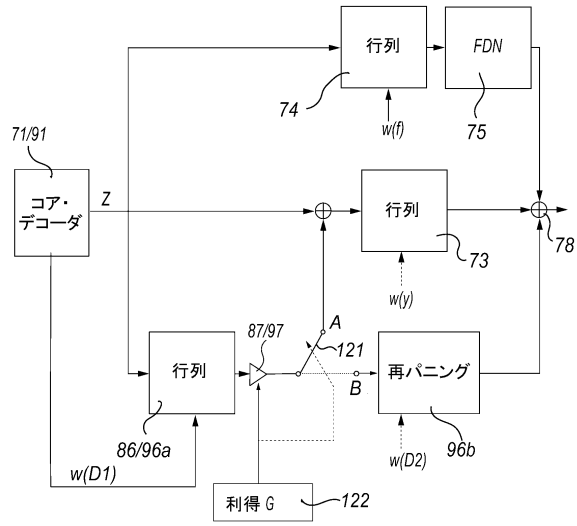


Fig. 12

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関
米国(US)

前置審査

(74)代理人 100070150

弁理士 伊東 忠彦

(74)代理人 100135079

弁理士 宮崎 修

(72)発明者 サミュエルソン, レイフ ジョナス

スウェーデン国, ストックホルム 113 30 イェーヴレガタン 12エー ドルビー スウェーデン アーバー内

(72)発明者 ブリーバート, デイルク ジェローン

オーストラリア国, ニュー・サウス・ウェールズ 2060, シドニー, マクマホンズ・ポイント, ミッチェル・ストリート 35 レベル 3, ドルビー・オーストラリア・プロプライエタリー・リミテッド内

(72)発明者 クーパー, デイヴィッド マシュー

オーストラリア国, ニュー・サウス・ウェールズ 2060, シドニー, マクマホンズ・ポイント, ミッチェル・ストリート 35 レベル 3, ドルビー・オーストラリア・プロプライエタリー・リミテッド内

(72)発明者 コッペンズ, イェルーン

スウェーデン国, ストックホルム 113 30 イェーヴレガタン 12エー ドルビー スウェーデン アーバー内

審査官 堀 洋介

(56)参考文献 国際公開第2015/031505(WO, A1)

特表2003-522472(JP, A)

特表2010-515290(JP, A)

(58)調査した分野 (Int.Cl., DB名)

H04S 1/00

H04S 7/00