

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6463749号
(P6463749)

(45) 発行日 平成31年2月6日(2019.2.6)

(24) 登録日 平成31年1月11日(2019.1.11)

(51) Int.Cl.

H04S 7/00 (2006.01)

F 1

H04S 7/00 300

請求項の数 10 (全 21 頁)

(21) 出願番号 特願2016-525578 (P2016-525578)
 (86) (22) 出願日 平成26年10月20日 (2014.10.20)
 (65) 公表番号 特表2016-539554 (P2016-539554A)
 (43) 公表日 平成28年12月15日 (2016.12.15)
 (86) 國際出願番号 PCT/EP2014/072411
 (87) 國際公開番号 WO2015/059081
 (87) 國際公開日 平成27年4月30日 (2015.4.30)
 審査請求日 平成29年6月20日 (2017.6.20)
 (31) 優先権主張番号 13290255.2
 (32) 優先日 平成25年10月23日 (2013.10.23)
 (33) 優先権主張国 歐州特許庁 (EP)

(73) 特許権者 510185767
 ドルビー・インターナショナル・アーベー
 オランダ王国, セーエン アムステルダム
 ズイドースト 1101, ヘリケルベル
 グウェグ 1-35, アポロ ビルティン
 グ 3エー
 (74) 代理人 100107766
 弁理士 伊東 忠重
 (74) 代理人 100070150
 弁理士 伊東 忠彦
 (74) 代理人 100091214
 弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 2 Dセットアップを使用したオーディオ再生のためのアンビソニックス・オーディオ音場表現を復号する方法および装置

(57) 【特許請求の範囲】

【請求項 1】

L 個のスピーカのために、符号化されたアンビソニックス形式オーディオ信号を復号する方法であって、

少なくとも 1 つの仮想スピーカの少なくとも 1 つの仮想位置を前記 L 個のスピーカの位置に追加するステップと、

前記 L 個のスピーカの位置および前記少なくとも 1 つの仮想位置に基づいて第一の行列を決定するステップであって、前記第一の行列が前記 L 個のスピーカの位置および前記少なくとも 1 つの仮想位置に対する係数を有する、ステップと、

第二の行列を決定するステップであって、前記仮想位置に対する係数が重み付けされ、前記 L 個のスピーカの位置に関する係数に分配され、前記 L 個のスピーカの位置に対する係数を有する前記第二の行列が取得される、ステップとを含み、

前記仮想位置に対する前記係数が重み係数 = 1 / L に基づいて重み付けされ、L はスピーカの数である、

方法。

【請求項 2】

前記第二の行列に基づいて前記符号化されたアンビソニックス形式オーディオ信号を復号して、複数の復号されたスピーカ信号を得ることをさらに含む、請求項 1 記載の方法。

【請求項 3】

フロベニウス・ノルムに基づいて前記第二の行列を正規化して、正規化された行列を決

10

20

定するステップをさらに含む、請求項 1 に記載の方法。

【請求項 4】

前記 L 個のスピーカの位置および音場信号の係数の次数 N を特定するステップと、

前記位置に基づいて前記 L 個のスピーカが実質的に 2 D 平面にあると特定するステップと、

仮想スピーカの前記少なくとも 1 つの仮想位置を生成するステップと、
をさらに含む、請求項 1 に記載の方法。

【請求項 5】

前記符号化されたオーディオ信号を帯域通過フィルタを使用して複数の周波数帯域に分離するステップをさらに含み、各周波数帯域に対して 1 つの、複数の別個の 3 D 復号行列が生成され、各 3 D 復号行列はダウンミキシングされ、任意的には別個に正規化される、
請求項 1 に記載の方法。 10

【請求項 6】

L 個のスピーカのために、符号化されたアンビソニックス形式オーディオ信号を復号する装置であって、

少なくとも 1 つの仮想スピーカの少なくとも 1 つの仮想位置を前記 L 個のスピーカの位置に追加する追加部と、

前記 L 個のスピーカの位置および前記少なくとも 1 つの仮想位置に基づいて第一の行列を決定する第一ユニットであって、前記第一の行列が前記 L 個のスピーカの位置および前記少なくとも 1 つの仮想位置に対する係数を有する、第一ユニットと、 20

第二の行列を決定する第二ユニットであって、前記仮想位置に対する係数が重み付けされ、前記 L 個のスピーカの位置に関連する係数に分配され、前記 L 個のスピーカの位置に対する係数を有する前記第二の行列が取得される、第二ユニットとを備える、

前記仮想位置に対する前記係数が重み係数 = $1/L$ に基づいて重み付けされ、 L はスピーカの数である、
装置。

【請求項 7】

前記第二の行列に基づいて前記符号化されたオーディオ信号を復号して、複数の復号されたスピーカ信号を取得する復号部、
を備える、請求項 6 記載の装置。 30

【請求項 8】

フロベニウス・ノルムに基づいて前記第二の行列を正規化して、正規化された行列を取得する正規化部をさらに有する、
請求項 6 記載の装置。

【請求項 9】

前記 L 個のスピーカの位置および音場信号の係数の次数 N を特定する第 1 の特定部と、

前記位置に基づいて前記 L 個のスピーカが実質的に 2 D 平面にあると特定する第 2 の特定部と、

仮想スピーカの前記少なくとも 1 つの仮想位置を生成する仮想スピーカ位置生成部と、
をさらに含む、請求項 6 に記載の装置。 40

【請求項 10】

帯域通過フィルタを使って前記符号化されたオーディオ信号を複数の周波数帯域に分離するユニットをさらに含み、各周波数帯域に対して 1 つ、複数の別個の 3 D 復号行列が生成され、各 3 D 復号行列は、ダウンミキシングされ、任意的には別個に正規化される、請求項 6 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、2 D セットアップまたはnear - 2 D セットアップを使用したオーディオ再生のためのアンビソニックス・オーディオ音場表現、特に、アンビソニックス形式のオ
50

—ディオ表現を復号する方法および装置に関する。

【背景技術】

【0002】

正確な定位は、どのような空間的なオーディオ再生システムにとっても主要な目標である。このような再生システムは、会議システム、ゲーム、または、3Dサウンドの利点を享受する他の仮想環境にとってきわめて実用的である。3Dにおけるサウンド・シーンは、自然な音場として合成または捕捉することができる。例えば、アンビソニックスのような音場信号は、所望の音場の表現を担持する。音場表現から個々のスピーカ信号を取得するには、復号処理が必要である。アンビソニックス形式の信号の復号は、「レンダリング」とも称する。オーディオ・シーンを合成するには、所与の音源の空間的な定位を取得するために空間的なスピーカ配置を参照するパン関数が必要である。自然な音場を記録するためには、空間的な情報の捕捉にマイクロフォン・アレイが必要である。アンビソニックス手法は、これを成し遂げるために大変適したツールである。アンビソニックス形式の信号は、音場の球面調和分解に基づいて、所望の音場の表現を担持する。基本的なアンビソニックス形式またはB形式は、次数0および1の球面調和関数を使用するが、いわゆる高次アンビソニックス(HOA)は、少なくとも2次のさらなる球面調和関数も使用する。スピーカの空間的な配置は、スピーカ・セットアップと称する。復号処理のためには、復号行列(レンダリング行列とも称する)が必要であり、この行列は、所与のスピーカ・セットアップに特化したものであり、既知のスピーカの位置を使用して生成される。

10

【0003】

20

一般的に使用されているスピーカ・セットアップは、2つのスピーカを使用するステレオ・セットアップ、5つのスピーカを使用する標準サラウンド・セットアップ、5つより多くのスピーカを使用するサラウンド・セットアップの拡張である。しかしながら、これらのセットアップはよく知られているが、2次元(2D)に制約され、例えば、高さ情報は再現されない。高さ情報を再現することができる既知のスピーカ・セットアップに対するレンダリングは、音の定位および音色において欠点を有する。これらの欠点は、空間的に垂直なパンが極めて不均一なラウドネスで知覚されるか、スピーカ信号が強いサイドロープを有する点であり、これは、特に、中心から外れた位置で聴き取る際の欠点となる。したがって、スピーカに対するHOA音場の記述をレンダリングする際には、いわゆるエネルギー保存性を有するレンダリング設計が好ましい。これは、単一の音源をレンダリングする結果として、音源の方向とは独立して、一定のエネルギーのスピーカ信号が発生することを意味する。還元すれば、アンビソニックス表現によって保持される入力エネルギーは、スピーカ・レンダラーによって保存される。本発明者による国際公開特許公報第2014/012945号[文献1]は、3Dスピーカ・セットアップに対する良好なエネルギー保存性および定位の特性を有するHOAレンダラー設計について記載している。しかしながら、この手法は、全ての方向をカバーする3Dスピーカ・セットアップに対しては極めて良好に動作するものの、音源の方向の中には、(例えば、5.1サラウンドのような)2Dスピーカ・セットアップでは減衰するものがある。このことは、特に、スピーカが配置されてない、例えば、トップからの方向に当てはまる。

30

【0004】

40

F.ZotterおよびM.Frank著「All-Round Ambisonic Panning and Decoding(オールラウンドなアンビソニック・パンニングおよび復号処理)」[文献2]では、スピーカによって構築される凸包に穴が存在する場合には、「架空の」スピーカが追加される。しかしながら、その架空のスピーカに対する結果として得られる信号は、実際のスピーカでの再生が省略される。したがって、その方向(すなわち、実際のスピーカが配置されていない方向)からの音源信号が依然として減衰されることとなる。さらに、本文献は、V B A P(ベクトル・ベースの振幅パンニング)と共に使用される架空のスピーカの使用を開示するのみである。

【発明の概要】

【0005】

50

したがって、残っている課題は、スピーカが配置されていない方向からの音源の減衰がより少ないか、全く減衰しないようにする、2D(2次元)スピーカ・セットアップに対するエネルギー保存性を有するアンビソニックス・レンダラーを設計することにある。2Dスピーカ・セットアップは、スピーカの仰角が所定の小さな範囲(例えば、10°未満(<10°))で、水平面に近くなるものとして分類することができる。

【0006】

本明細書は、規則的または非規則的な、空間的なスピーカ配置に対するアンビソニックス形式の音場表現をレンダリング/復号処理するための解決法について記載し、そのレンダリング/復号処理は、高度に改善された定位特性および音色特性をもたらし、エネルギー保存性を有し、スピーカを利用可能でない方向からの音をもレンダリングする。スピーカを利用可能でない方向からの音は、スピーカが各方向で利用可能であると仮定した場合と概ね同様のエネルギーおよび知覚されるラウドネスでレンダリングされることは有利である。もちろん、その方向ではスピーカが利用可能でないため、これらの音源の正確な定位は可能ではない。

【0007】

特に、少なくとも幾つかの記載した実施形態は、HOA形式の音場データを復号するための復号行列を取得する新しい方法を提供する。少なくともHOA形式は、スピーカの位置とは直接関連していない音場を記述し、取得されるスピーカの信号は、必ずチャンネル・ベースのオーディオ形式であるため、HOA信号の復号は、常に、オーディオ信号のレンダリングに密接に関連している。原理的には、同じことが、他のオーディオの音場形式にも当てはまる。したがって、本開示内容は、音場に関連するオーディオ形式の復号およびレンダリングの両方に関連する。復号行列およびレンダリング行列の用語は、同意語として使用されている。

【0008】

良好なエネルギー保存特性を有する所与のセットアップに対する復号行列を取得するために、1つ以上の仮想のスピーカがスピーカを利用可能でない場所に追加される。例えば、2Dセットアップに対する改良された復号行列を取得するために、2つの仮想のスピーカがトップおよびボトムに追加される(トップおよびボトムは、概ね仰角0°で設置された2Dスピーカでは+90°および-90°の仰角に対応する。)。この仮想的な3Dスピーカ・セットアップのために、エネルギー保存特性を満たす復号行列が設計される。最後に、仮想のスピーカに対する復号行列からの重み係数は、2Dセットアップの実際のスピーカに対する一定利得とミキシングされる。

【0009】

一実施形態によれば、所与の組のスピーカに対するアンビソニックス形式のオーディオ信号をレンダリングまたは復号する復号行列(またはレンダリング行列)を生成し、その生成は、従来の方法を使用して、変更されたスピーカの位置を使用して、第1の予備復号行列を生成するステップであって、変更されたスピーカの位置が所与の組のスピーカのスピーカ位置および少なくとも1つの追加的な仮想のスピーカ位置を含む、上記生成するステップと、第1の予備復号行列をダウンミキシングするステップであって、上記少なくとも1つの追加的な仮想のスピーカに関連する係数が除かれ、所与の組のスピーカの、スピーカに関連する係数に分配される、上記ダウンミキシングするステップと、によって行われる。一実施形態においては、続いて、復号行列を正規化する後続するステップが行われる。結果として得られる復号行列は、所与の組のスピーカのためのアンビソニックス信号をレンダリングまたは復号するのに適しており、スピーカが存在しない位置からの音でさえも、正確な信号エネルギーで再生される。これは、改良された復号行列の構築によるものである。好ましくは、第1の予備復号行列はエネルギー保存性を有する。

【0010】

一実施形態においては、復号行列はL個の行およびO_{3D}個の列を有する。行の数は2Dスピーカ・セットアップにおけるスピーカの数に対応し、列の数はO_{3D}=(N+1)²に従ったHOA次数Nに依存するアンビソニックス係数O_{3D}の数に対応する。2Dスピーカ

10

20

30

40

50

カ・セットアップに対する復号行列の係数の各々は、少なくとも第1の中間係数および第2の中間係数の合計である。第1の中間係数は、2Dスピーカ・セットアップの現在のスピーカの位置に対するエネルギー保存性を有する3D行列設計方法によって取得され、このエネルギー保存性を有する3D行列設計方法は、少なくとも1つの仮想のスピーカの位置を使用する。第2の中間係数は、少なくとも1つの仮想のスピーカの位置に対する上記エネルギー保存性を有する3D行列設計方法から取得された、重み係数 g を乗算した係数によって取得される。一実施形態においては、重み係数 g は

$$g = \frac{1}{\sqrt{L}}$$

に従って算出され、ここで、 L は2Dスピーカ・セットアップにおけるスピーカの数である。 10

【0011】

一実施形態においては、本発明は、上述した、または、請求の範囲に記載されたステップを含む方法をコンピュータに行わせるための実行可能な命令を記憶したコンピュータ読取可能な媒体に関する。この方法を利用する装置は、請求項9に開示されている。

【0012】

従属請求項、以下の説明および図面には、有利な実施形態が開示されている。

【0013】

本発明の例示的な実施形態が添付図面を参照して説明されている。

【図面の簡単な説明】

20

【0014】

【図1】一実施形態に係る方法のフローチャートである。

【図2】ダウンミキシング済みのHOA復号行列の例示的な構成を示した図である。

【図3】スピーカの位置を取得、変更するためのフローチャートである。

【図4】一実施形態に係る装置を示すブロック図である。

【図5】従来の復号行列から結果的に生じるエネルギー分布を示した図である。

【図6】実施形態に係る復号行列から結果的に生じるエネルギー分布を示した図である。

【図7】複数の異なる周波数帯域に対する別個に最適化された復号行列の使用を示した図である。

【発明を実施するための形態】

30

【0015】

図1は、本発明の一実施形態に係るオーディオ信号、特に、音場信号を復号する方法のフローチャートを示している。音場信号の復号は、一般的には、オーディオ信号がレンダリングされるスピーカの位置を必要とする。 L 個のスピーカに対するこのようなスピーカの位置

$$\hat{\Omega}_1 \dots \hat{\Omega}_L$$

が本処理に入力される(i10)。なお、位置について言及する場合は、本明細書において、実際には、空間的な方向を意味する。すなわち、スピーカの位置は、その傾斜角 γ および方位角 ϕ によって定義され、これらの傾斜角 γ および方位角 ϕ を組み合わせてベクトル

$$\hat{\Omega}_l = [\theta_l, \phi_l]^T$$

40

とする。そして、ステップ10において仮想のスピーカの少なくとも1つの位置を追加する。一実施形態においては、処理i10で入力される全てのスピーカの位置は2Dセットアップを構成するように概ね同一平面にあり、追加される少なくとも1つの仮想のスピーカはこの平面の外にある。一つの特に有利な実施形態においては、処理i10で入力される全てのスピーカの位置は概ね同一平面にあり、ステップ10において2つの仮想のスピーカの位置を追加する。2つの仮想のスピーカの有利な位置について以下に記載する。一実施形態においては、後述する式(6)に従って追加が行われる。追加するステップ10を行った結果として、一組のスピーカの角度

50

$\widehat{\Omega}'_1 \dots \widehat{\Omega}'_{L+L_{\text{virt}}}$

が変更される (q 1 0)。 L_{virt} は仮想のスピーカの数である。変更された一組のスピーカの角度は、3D復号行列設計ステップ11で使用される。さらに、HOAの次数N（一般的には音場信号の係数の次数）はステップ11に供給される必要がある (i 1 1)。

【0016】

3D復号行列設計ステップ11は、3D復号行列を生成するための任意の既知の方法を実行する。好ましくは、3D復号行列は、エネルギー保存タイプの復号／レンダリングに適している。例えば、国際特許出願第E P 2 0 1 3 / 0 6 5 0 3 4号明細書に記載された方法を使用することができる。3D復号行列設計ステップ11の結果として、 $L' = L + L_{\text{virt}}$ 個のスピーカ信号のレンダリングに適した復号行列またはレンダリング行列 D' が得られる。ここで、 L_{virt} は、「仮想のスピーカの位置を追加する」ステップ10で追加された仮想のスピーカの位置の数である。

【0017】

L 個のスピーカのみが物理的に利用可能であるため、3D復号行列設計ステップ11から結果的に生成される復号行列 D' は、ダウンミキシングするステップ12において L 個のスピーカに適応するようにする必要がある。ステップ12では、復号行列 D' のダウンミキシングを行い、ここで、仮想のスピーカに関連する係数が重み付けされ、既存のスピーカに関連する係数に分配される。好ましくは、任意の特定のHOA次数の係数（すなわち、復号行列 D' の列）が重み付けされ、同一のHOA次数の係数（すなわち、復号行列 D' の同一の列）に加算される。一例は、後述する式(8)に従ったダウンミキシングである。ダウンミキシングするステップ12の結果として、 L 個の行を有する、すなわち、復号行列 D' よりも行の数が少ないが、復号行列 D' と列の数が同じダウンミキシング済みの3D復号行列

\tilde{D}

が生成される。換言すれば、復号行列 D' の次元は、 $(L + L_{\text{virt}}) \times 0_{3D}$ であり、ダウンミキシング済みの3D復号行列

\tilde{D}

の次元は、 $L \times 0_{3D}$ である。

【0018】

図2は、HOA復号行列 D' からのダウンミキシング済みのHOA復号行列

\tilde{D}

の例示的な構成を示している。HOA復号行列 D' は、 $L + 2$ 個の行を有し、これは、2つの仮想のスピーカの位置が L 個の利用可能なスピーカの位置に追加されたものである。また、HOA復号行列 D' は、 0_{3D} 個の列を有する。ここで、 0_{3D} は、 $= (N + 1)^2$ であり、 N は、HOAの次数である。ダウンミキシングするステップ12において、HOA復号行列 D' の行 $L + 1$ および $L + 2$ の係数が重み付けされ、各々の列の係数に分配され、行 $L + 1$ および $L + 2$ が除かれる。例えば、行 $L + 1$ および $L + 2$ の各々の第1の係数 $d'_{L+1,1}$ 、および $d'_{L+2,1}$ が重み付けされ、 $d'_{1,1}$ などの各残りの行の第1の係数に追加される。ダウンミキシング済みのHOA復号行列

\tilde{D}

から結果的に得られる係数

$\tilde{d}_{1,1}$

は、 $d'_{1,1}$ 、 $d'_{L+1,1}$ 、 $d'_{L+2,1}$ および重み係数 g の関数である。同様に、例えば、ダウンミキシング済みのHOA復号行列

\tilde{D}

から結果的に得られる係数

$\tilde{d}_{2,1}$

は、 $d'_{2,1}$ 、 $d'_{L+1,1}$ 、 $d'_{L+2,1}$ および重み係数 g の関数であり、ダウンミキシング済みのHOA復号行列

10

20

30

40

50

\tilde{D}

の結果として得られる係数

 $\tilde{d}_{1,2}$ は、 $d'_{1,2}$ 、 $d'_{L+1,2}$ 、 $d'_{L+2,2}$ および重み付け係数 g の関数である。

【0019】

通常、ダウンミキシング済みのHOA復号行列

 \tilde{D}

は、正規化ステップ13において正規化される。しかしながら、このステップ13は、音場信号の復号に非正規化された復号行列を使用することができるため、必要に応じて行われるものである。一実施形態においては、ダウンミキシング済みのHOA復号行列

10

 \tilde{D} は、後述する式(9)に従って正規化される。正規化ステップ13の結果として、正規化されたダウンミキシング済みのHOA行列 D が生成され、このHOA復号行列 D は、ダウンミキシング済みのHOA復号行列 \tilde{D} と同じ次元 $L \times O_{3D}$ を有する。

【0020】

次いで、正規化されたダウンミキシング済みのHOA復号行列 D は、音場復号ステップ14で使用され、ここで、入力音場信号 i_{14} が復号されて L 個のスピーカ信号 q_{14} となる。通常、スピーカ・セットアップが変更されるまでは、正規化されたダウンミキシング済みのHOA復号行列 D は変更される必要はない。したがって、一実施形態においては、正規化されたダウンミキシング済みのHOA復号行列 D は、復号行列ストレージに記憶される。

20

【0021】

図3は、一実施形態において、どのようにスピーカの位置が取得され、変更されるかの詳細を示している。本実施形態は、 L 個のスピーカの位置 $\hat{\Omega}_1 \dots \hat{\Omega}_L$ および音場信号の係数の次数 N を特定するステップ101と、この L 個のスピーカの位置から L 個のスピーカが実質的に2D平面上にあると特定するステップ102と、仮想のスピーカの少なくとも1つの仮想の位置

30

 $\hat{\Omega}'_{L+1}$

を生成するステップ103と、を含む。

【0022】

一実施形態においては、この少なくとも1つの仮想の位置

 $\hat{\Omega}'_{L+1}$

は、

$$\hat{\Omega}'_{L+1} = [0, 0]^T$$

および

$$\hat{\Omega}'_{L+1} = [\pi, 0]^T$$

40

のうちの一方である。

【0023】

一実施形態においては、ステップ103において、2つの仮想のスピーカに対応する2つの仮想の位置

 $\hat{\Omega}'_{L+1}$

および

 $\hat{\Omega}'_{L+2}$

を生成する。ここで、

$$\hat{\Omega}'_{L+1} = [0, 0]^T$$

および

$$\hat{\Omega}'_{L+2} = [\pi, 0]^T$$

である。

【 0 0 2 4 】

一実施形態によれば、既知の位置にある L 個のスピーカに対する符号化されたオーディオ信号を復号する方法は、この L 個のスピーカの位置

$$\hat{\Omega}_1 \dots \hat{\Omega}_L$$

10

および音場信号の係数の次数 N を特定するステップ 101 と、こ L 個のスピーカの位置から L 個のスピーカが実質的に 2D 平面にあると特定するステップ 102 と、仮想のスピーカの少なくとも 1 つの仮想の位置

$$\hat{\Omega}'_{L+1}$$

を生成するステップ 103 と、3D 復号行列 D' を生成するステップ 11 であって、その L 個のスピーカの特定された位置

$$\hat{\Omega}_1 \dots \hat{\Omega}_L$$

および少なくとも 1 つの仮想の位置

$$\hat{\Omega}'_{L+1}$$

20

が使用され、3D 復号行列 D' は、上記特定されたスピーカの位置および仮想のスピーカの位置に対する係数を有する、上記生成するステップ 11 と、3D 復号行列 D' をダウンミキシングするステップ 12 であって、仮想のスピーカの位置に対する係数が重み付けされ、その特定されたスピーカの位置に関する係数に分配され、その特定されたスピーカの位置に対する係数を有するダウンスケーリングされた 3D 復号行列

$$\tilde{D}$$

が取得される、上記ダウンミキシングするステップ 12 と、そのダウンスケーリングされた 3D 復号行列

$$\tilde{D}$$

を使用して符号化されたオーディオ信号 i_{14} を復号するステップ 14 であって、複数の復号されたスピーカ信号 q_{14} が取得される、上記復号するステップ 14 と、を含む。

30

【 0 0 2 5 】

一実施形態においては、符号化されたオーディオ信号は音場信号であり、例えば、HOA 形式の音場信号である。

【 0 0 2 6 】

一実施形態においては、上記の仮想のスピーカの少なくとも 1 つの仮想の位置

$$\hat{\Omega}'_{L+1}$$

は、

$$\hat{\Omega}'_{L+1} = [0, 0]^T$$

および

$$\hat{\Omega}'_{L+1} = [\pi, 0]^T$$

40

のうちの一方である。

【 0 0 2 7 】

一実施形態においては、上記の仮想のスピーカの位置に対する係数が重み係数

$$g = \frac{1}{\sqrt{L}}$$

を用いて重み付けされる。

【 0 0 2 8 】

一実施形態においては、この方法は、ダウンスケーリング済みの 3D 復号行列

50

\tilde{D}

を正規化するステップをさらに含み、正規化されたダウンスケーリング済みの3D復号行列Dが取得され、符号化されたオーディオ信号i14を復号する上記のステップ14は、正規化されたダウンスケーリング済みの3D復号行列Dを使用する。一実施形態においては、この方法は、ダウンスケーリング済みの3D復号行列

 \tilde{D}

または正規化されたダウンミキシング済みのHOA復号行列Dを復号行列ストレージに記憶するステップをさらに含む。

【0029】

一実施形態によれば、所与の組のスピーカに対する音場信号をレンダリングまたは復号する復号行列を生成する。この生成は、従来の方法を使用して、変更されたスピーカの位置を使用して、第1の予備復号行列を生成するステップであって、変更されたスピーカの位置が所与の組のスピーカのスピーカ位置および少なくとも1つの追加的な仮想のスピーカのスピーカ位置を含む、上記生成するステップと、第1の予備復号行列をダウンミキシングするステップであって、少なくとも1つの追加的な仮想のスピーカに関連する係数は除かれ、所与の組のスピーカのスピーカに関連する係数に分配される、上記ダウンミキシングするステップと、によって行われる。一実施形態においては、続いて、復号行列を正規化する以下のステップが行われる。結果として得られる復号行列は、所与の組のスピーカに対する音場信号をレンダリングまたは復号するのに適しており、スピーカが存在しない位置からの音でさえも、正確な信号エネルギーで再生される。これは、改良された復号行列の構成によるものである。好ましくは、第1の予備復号行列はエネルギー保存性を有する。

【0030】

図4a)は、一実施形態に係る装置のブロック図を示している。既知の位置にあるL個のスピーカに対する音場形式の符号化されたオーディオ信号を復号する装置400は、少なくとも1つの仮想のスピーカの少なくとも1つの位置をL個のスピーカの位置に追加する追加部410と、3D復号行列D'を生成する復号行列生成部411であって、そのL個のスピーカの位置

 $\hat{\Omega}_1 \dots \hat{\Omega}_L$

および少なくとも1つの仮想の位置

 $\hat{\Omega}'_{L+1}$

が使用され、3D復号行列D'が上記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有し、3D復号行列D'をダウンミキシングする行列ダウンミキシング部412であって、仮想のスピーカに対する係数が重み付けされ、特定されたスピーカの位置に関連する係数に分配され、特定されたスピーカの位置に対する係数を有するダウンスケーリングされた3D復号行列

 \tilde{D}

が取得される、上記行列ダウンミキシング部412と、ダウンスケーリングされた3D復号行列

 \tilde{D}

を使用して符号化されたオーディオ信号を復号する復号部414であって、複数の復号されたスピーカ信号が取得される、上記復号部414と、を含む。

【0031】

一実施形態においては、装置は、ダウンスケーリングされた3D復号行列

 \tilde{D}

を正規化する正規化部413をさらに含み、正規化されたダウンスケーリング済みの3D復号行列Dが取得され、復号部414は、正規化されたダウンスケーリング済みの3D復号行列を使用する。

【0032】

図4b)に示された一実施形態においては、装置は、L個のスピーカの位置($\Omega_1 \dots \Omega_L$)お

10

20

30

40

50

および音場信号の係数の次数Nを特定する第1の特定部4101と、このL個のスピーカの位置からL個のスピーカが実質的に2D平面にあると特定する第2の特定部4102と、仮想のスピーカの少なくとも1つの仮想の位置

$(\hat{\Omega}_{L+1}')$

を生成する仮想スピーカ位置生成部4103と、を含む。

【0033】

一実施形態においては、装置は、符号化されたオーディオ信号を複数の周波数帯域に分離する帯域通過フィルタ715bを含み、711bで複数の分離された3D復号行列Db'（各周波数帯域に対して1つの分離された3D復号行列Db'）が生成され、712bで各3D復号行列Db'はダウンミキシングされ、さらに別個に正規化されてもよく、復号部714bは各周波数帯域毎に別個に復号する。本実施形態においては、装置は、各スピーカに対して1つ、複数の加算部716bをさらに含む。各加算部は、各々のスピーカに関連する周波数帯域を合計する。10

【0034】

追加部410、復号行列生成部411、行列ダウンミキシング部412、正規化部413、復号部414、第1の特定部4101、第2の特定部4102、および仮想スピーカ位置生成部4103の各々の機能は、1つ以上のプロセッサによって実施され、これらの部の各々は、これらのうちの他の部、または、これらの部ではない他の部と同一のプロセッサを共有することがある。20

【0035】

図7は、入力信号の複数の異なる周波数帯域に対して別個に最適化された復号行列を使用する実施形態を示している。本実施形態においては、復号方法は、符号化されたオーディオ信号を帯域通過フィルタを使用して複数の周波数帯域に分離するステップを含む。711bで複数の分離された3D復号行列Db'（各周波数帯域に対して1つの分離された3D復号行列Db'）が生成され、712bで各3D復号行列Db'は、ダウンミキシングされる。さらに別個に正規化されてもよい。714bで各周波数帯域に対して符号化されたオーディオ信号の復号が別個に行われる。これにより、人間の知覚における周波数依存差が考慮されるという利点が得られ、異なる周波数帯域に対して異なる復号行列が得られることとなる。一実施形態においては、1つのみ、あるいは複数の（全てではないが）復号行列を、上述したように、仮想のスピーカの位置を追加し、次いで、仮想のスピーカの位置の各々の係数を重み付けし、既存のスピーカの位置に対する係数に分配することによって、生成する。別の実施形態においては、各々の符号化行列を、上述したように、仮想のスピーカの位置を追加し、次いで、仮想のスピーカの位置の各々の係数を重み付けし、既存のスピーカの位置に対する係数に分配することによって、生成する。最後に、周波数帯域分割と逆の処理で、1つの周波数帯域加算部716bで同一のスピーカに関連する全ての周波数帯域を、スピーカ毎に、合計する。30

【0036】

追加部410、復号行列生成部711b、行列ダウンミキシング部712b、正規化部713b、復号部714b、周波数帯域加算部716b、および帯域通過フィルタ部715bの各々は、1つ以上のプロセッサによって実施され、これらの機能部の各々は、これらのうちの他の機能部、または、これらの機能部ではない他の機能部と同一のプロセッサを共有することがある。40

【0037】

本開示内容の一態様は、良好なエネルギー保存特性を有する2Dセットアップに対するレンダリング行列を取得するものである。一実施形態においては、2つのスピーカがトップおよびボトム（概ね仰角0°で設置された2Dスピーカでは+90°および-90°の仰角）に追加される。この仮想的な3Dスピーカ・セットアップに対して、エネルギー保存特性を満たすレンダリング行列が設計される。最後に、仮想のスピーカに対するレンダリング行列からの重み係数が2Dセットアップの実際のスピーカに対する一定（コンスタ50

ント) の利得とミキシングされる。

【 0 0 3 8 】

以下において、アンビソニックス(特に、H O A)のレンダリングについて説明する。

【 0 0 3 9 】

アンビソニックス・レンダリングは、アンビソニックス音場の記述からスピーカ信号を算出する処理である。これは、時には、アンビソニックス復号とも呼ばれる。次数Nの3Dアンビソニックス音場表現が考慮され、ここで、係数の数は、以下の式(1)の通りである。

$$O_{3D} = (N + 1)^2 \quad (1)$$

【 0 0 4 0 】

この時間サンプルtの係数は、 O_{3D} 個の要素を有するベクトル
 $\mathbf{b}(t) \in \mathbb{C}^{O_{3D} \times 1}$

によって表される。レンダリング行列

$$\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$$

を用いて、時間サンプルtに対するスピーカ信号は、以下の式(2)によって算出される。

$$\mathbf{w}(t) = \mathbf{D} \mathbf{b}(t) \quad (2)$$

ここで、

$$\mathbf{D} \in \mathbb{C}^{L \times O_{3D}}$$

および

$$\mathbf{w} \in \mathbb{R}^{L \times 1}$$

であり、Lはスピーカの数である。

【 0 0 4 1 】

スピーカの位置は、各々の傾斜角 θ_l および方位角 ϕ_l によって定義され、これらの傾斜角 θ_l および方位角 ϕ_l を組み合わせてベクトル

$$\hat{\Omega}_l = [\theta_l, \phi_l]^T \text{ for } l = 1, \dots, L$$

とする。聴取位置からの相異なるスピーカの距離は、スピーカ・チャンネルに対するそれぞれの遅延を使用することで補償される。

【 0 0 4 2 】

H O A領域における信号エネルギーは、以下の式(3)によって与えられる。

$$E = \mathbf{b}^H \mathbf{b} \quad (3)$$

ここで、 \mathbf{b}^H は、複素共役転置を表している。スピーカ信号の対応するエネルギーは、以下の式(4)によって算出される。

【 数 1 】

$$\hat{E} = \mathbf{w}^H \mathbf{w} = \mathbf{b}^H \mathbf{D}^H \mathbf{D} \mathbf{b}. \quad (4)$$

【 0 0 4 3 】

エネルギー保存性のある復号/レンダリングを成し遂げるために、エネルギー保存性のある復号/レンダリング行列の比

$$\hat{E}/E$$

は一定(コンスタント)であるべきである。

【 0 0 4 4 】

原理的には、改良された2Dレンダリングのための以下の拡張が提案される。2Dスピーカ・セットアップに対するレンダリング行列の設計のために、1つ以上の仮想のスピーカを追加する。2Dセットアップは、スピーカの仰角が所定の小さな範囲内にあり、水平面に近くなるものと考えられる。これは、以下の式(5)のように表現することができる。

10

20

30

40

【数2】

$$\left| \theta_l - \frac{\pi}{2} \right| \leq \theta_{thres2d}; l = 1, \dots, L \quad (5)$$

【0045】

通常、閾値 $\theta_{thres2d}$ は、一実施形態においては、5° ~ 10° の範囲にある値に対応するように選定される。

【0046】

レンダリング設計については、変更された組のスピーカ角度

$$\hat{\Omega}'_l$$

10

が定義される。最後の（この例においては、2つ）のスピーカの位置は、極座標系の南極および北極（垂直方向の、すなわち、トップおよびボトム）の2つの仮想のスピーカのものである。

【数3】

$$\begin{aligned} \hat{\Omega}'_l &= \hat{\Omega}_l; l = 1, \dots, L \\ \hat{\Omega}'_{L+1} &= [0, 0]^T \\ \hat{\Omega}'_{L+2} &= [\pi, 0]^T \end{aligned} \quad (6)$$

【0047】

20

そして、レンダリング設計のために使用されるスピーカの新しい数は、 $L' = L + 2$ である。これらの変更されたスピーカの位置から、エネルギー保存手法を用いてレンダリング行列

$$D' \in \mathbb{C}^{(L+2) \times O_{3D}}$$

が設計される。例えば、[文献1]に記載された設計方法が使用される。次に、元のスピーカ・セットアップに対する最終的なレンダリング行列が D' から導出される。1つの考え方とは、行列 D' に定義されている仮想のスピーカの重み係数を実際のスピーカに対してミキシングすることである。固定された利得係数が使用され、これは、以下の式(7)のように選定される。

【数4】

30

$$g = \frac{1}{\sqrt{L}}. \quad (7)$$

【0048】

中間行列

$$\tilde{D} \in \mathbb{C}^{L \times O_{3D}}$$

の係数（本明細書では、ダウンスケーリングされた3D復号行列とも呼ばれる）は、以下の式(8)のように定義される。

【数5】

$$\tilde{d}_{l,q} = d'_{l,q} + g \cdot d'_{L+1,q} + g \cdot d'_{L+2,q} \quad \text{for } l = 1, \dots, L \text{ and } q = 1, \dots, O_{3D} \quad (8)$$

40

ここで、

$$\tilde{d}_{l,q}$$

は、1番目の行およびq番目の列における

$$\tilde{D}$$

の行列要素である。必要に応じて最後のステップにおいては、中間行列（ダウンスケーリングされた3D復号行列）がフロベニウス・ノルムを使用して正規化してもよい。

【数6】

$$D = \frac{\tilde{D}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{O_{3D}} |\tilde{d}_{l,q}|^2}} \quad (9)$$

【0049】

図5および図6は、5.0サラウンド・スピーカ・セットアップに対するエネルギー分布を示している。両方の図において、エネルギーの値は、グレースケールとして示されており、丸印は、スピーカの位置を示している。開示されている方法を用いて、特に、トップ(ここでは示されていないが、さらに、ボトム)での減衰が減少しているのは明らかである。

【0050】

図5は、従来の復号行列から結果的に得られるエネルギー分布を示している。z=0平面の周りの小さな円は、スピーカの位置を表している。[-3.9, ..., 2.1]デジベル(dB)のエネルギー範囲がカバーされ、この結果として、エネルギー差が6dBとなることが分かる。さらに、単位球面のトップからの信号(さらに、図示されていないが、ボトム上の信号)は、ここではスピーカが利用可能でないため、極めて低エネルギーで再生され、すなわち、聞き取りができない。

【0051】

図6は、1つ以上の実施形態に係る復号行列から生ずるエネルギー分布を示している。図5の場合と同じ位置に同じ数のスピーカが存在する。少なくとも以下の利点がもたらされる。第1に、[-1.6, ..., 0.8]デジベル(dB)のより小さなエネルギー範囲がカバーされ、この結果として、エネルギー差がより小さくなり、2.4dBのみとなる。第2に、単位球面の全ての方向からの信号は、ここにスピーカが存在しない場合であっても、それぞれの正確なエネルギーを用いて再生される。これらの信号は、利用可能なスピーカを通じて再生されるため、それぞれの定位は正確ではない。しかしながら、信号は、正しいラウドネスで聞き取り可能である。この例において、トップからの信号およびボトム上の信号(図示せず)は、改良された復号行列を用いた復号によって聞き取りできるようになる。

【0052】

一実施形態においては、既知の位置にあるL個のスピーカのためのアンビソニックス形式の符号化されたオーディオ信号を復号する方法は、少なくとも1つの仮想のスピーカの少なくとも1つの位置をL個のスピーカの位置に追加するステップと、3D復号行列D'を生成するステップであって、そのL個のスピーカの位置

$\hat{\Omega}_1 \dots \hat{\Omega}_L$

および少なくとも1つの仮想の位置

$\hat{\Omega}'_{L+1}$

が使用され、その3D復号行列D'が上記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、上記生成するステップと、3D復号行列D'をダウンミキシングするステップであって、仮想のスピーカの位置に対する係数が重み付けされ、特定されたスピーカの位置に関する係数に分配され、特定されたスピーカの位置に対する係数を有するダウンスケーリングされた3D復号行列

\tilde{D}

が取得される、上記ダウンミキシングするステップと、ダウンスケーリングされた3D復号行列

\tilde{D}

を使用して符号化されたオーディオ信号を復号するステップであって、複数の復号されたスピーカ信号が取得される、上記復号するステップと、を含む。

【0053】

別の実施形態においては、既知の位置にあるL個のスピーカのためのアンビソニックス

10

20

30

40

50

形式の符号化されたオーディオ信号を復号する装置は、少なくとも 1 つの仮想のスピーカの少なくとも 1 つの位置を L 個のスピーカの位置に追加する追加部 410 と、3D 復号行列 D' を生成する復号行列生成部 411 であって、 L 個のスピーカの位置

$\hat{\Omega}_1 \dots \hat{\Omega}_L$

および少なくとも 1 つの仮想の位置

$\hat{\Omega}'_{L+1}$

が使用され、その 3D 復号行列 D' が上記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、上記復号行列生成部 411 と、3D 復号行列 D' をダウンミキシングするダウンミキシング部 412 であって、仮想のスピーカの位置に対する係数が重み付けされ、特定されたスピーカの位置に関連する係数に分配され、特定されたスピーカの位置に対する係数を有するダウンスケーリングされた 3D 復号行列

\tilde{D}

が取得される、上記ダウンミキシング部 412 と、ダウンスケーリングされた 3D 復号行列

\tilde{D}

を使用して符号化されたオーディオ信号を復号する復号部 414 であって、複数の復号されたスピーカ信号が取得される、上記復号部 414 と、を含む。

【0054】

さらに別の実施形態においては、既知の位置にある L 個のスピーカのためのアンビソニックス形式の符号化されたオーディオ信号を復号する装置は、少なくとも 1 つのプロセッサおよび少なくとも 1 つのメモリを含み、そのメモリは命令を記憶し、その命令がプロセッサ上で実行されると、プロセッサは、少なくとも 1 つの仮想のスピーカの少なくとも 1 つの位置を L 個のスピーカの位置に追加する追加部 410 と、3D 復号行列 D' を生成する復号行列生成部 411 であって、 L 個のスピーカの位置

$\hat{\Omega}_1 \dots \hat{\Omega}_L$

および少なくとも 1 つの仮想の位置

$\hat{\Omega}'_{L+1}$

が使用され、3D 復号行列 D' が上記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、上記復号行列生成部 411 と、3D 復号行列 D' をダウンミキシングする行列ダウンミキシング部 412 であって、仮想のスピーカの位置に対する係数が重み付けされ、特定されたスピーカの位置に関連する係数に分配され、特定されたスピーカの位置に対する係数を有するダウンスケーリングされた 3D 復号行列

\tilde{D}

が取得される、上記行列ダウンミキシング部 412 と、ダウンスケーリングされた 3D 復号行列

\tilde{D}

を使用して符号化されたオーディオ信号を復号する復号部 414 であって、複数の復号されたスピーカ信号が取得される、上記復号部 414 と、の機能を実現する。

【0055】

さらに別の実施形態においては、コンピュータ読取可能な記憶媒体は、既知の位置にある L 個のスピーカのためのアンビソニックス形式の符号化されたオーディオ信号を復号する方法をコンピュータに実行させるための実行可能な命令を記憶し、この方法は、少なくとも 1 つの仮想のスピーカの少なくとも 1 つの位置を L 個のスピーカの位置に追加するステップと、3D 復号行列 D' を生成するステップであって、 L 個のスピーカの位置

$\hat{\Omega}_1 \dots \hat{\Omega}_L$

および少なくとも 1 つの仮想の位置

$\hat{\Omega}'_{L+1}$

が使用され、その 3D 復号行列 D' が上記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、上記生成するステップと、その 3D 復号行列 D' をダウンミキシングするステップであって、仮想のスピーカの位置に対する係数が重み付けされ、特定

10

20

30

40

50

されたスピーカの位置に関連する係数に分配され、特定されたスピーカの位置に対する係数を有するダウンスケーリングされた3D復号行列

\tilde{D}

が取得される、上記ダウンミキシングするステップと、ダウンスケーリングされた3D復号行列

\tilde{D}

を使用して符号化されたオーディオ信号を復号するステップであって、複数の復号されたスピーカ信号が取得される、上記復号するステップと、を含む。コンピュータ読取可能な記憶媒体のさらなる実施形態は、上述した特徴事項、特に、請求項1に從属する従属請求項に開示された特徴事項を任意に含むことができる。

10

【0056】

本発明は、純粋に、例示的な目的で説明されているが、本発明の範囲を逸脱することなく、詳細な事項を変更することが可能である。例えば、HOAに関してのみ説明しているが、本発明は、他の音場オーディオ形式にも適用することができる。

【0057】

明細書、(該当する場合には)請求項、および図面に開示された各構成要素は、独立して設けてもよく、任意に適切に組み合わせて設けてもよい。構成要素は、適宜、ハードウェア、ソフトウェア、または、ハードウェアおよびソフトウェアの両方を組み合わせて実施することができる。請求項に存在する参照符号は例示的な目的のみで記載されており、請求項に係る範囲に限定的な影響を与えるものではない。

20

【0058】

引用した参考文献は、以下の通りである。

[文献1] 国際特許公開公報第2014/012945号(PD120032)

[文献2] F. ZottlerおよびM. Frank著「All-Round Ambisonic Panning and Decoding(オールラウンドなアンビソニック・パンニングおよび復号処理)」、オーディオ技術者協会ジャーナル、2012年、第60巻、807-820頁

いくつかの態様を記載しておく。

〔態様1〕

既知の位置にあるL個のスピーカに対するアンビソニックス形式の符号化されたオーディオ信号を復号する方法であって、

30

-少なくとも1つの仮想のスピーカの少なくとも1つの位置を前記L個のスピーカの位置に追加するステップ(10)と、

-3D復号行列(D')を生成するステップ(11)であって、前記L個のスピーカの位置

$(\hat{\Omega}_1, \dots, \hat{\Omega}_L)$

および前記少なくとも1つの仮想の位置

$(\hat{\Omega}'_{L+1})$

40

が使用され、前記3D復号行列(D')が前記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、前記生成するステップ(11)と、

-前記3D復号行列(D')をダウンミキシングするステップ(12)であって、前記仮想のスピーカの位置に対する係数が重み付けされ、前記特定されたスピーカの位置に関する係数に分配され、前記特定されたスピーカの位置に対する係数を有するダウンスケーリングされた3D復号行列

(\tilde{D})

が取得される、前記ダウンミキシングするステップ(12)と、
-前記ダウンスケーリングされた3D復号行列

(\tilde{D})

を使用して前記符号化されたオーディオ信号(i14)を復号するステップ(14)であ
って、複数の復号されたスピーカ信号(q14)が取得される、前記復号するステップ(

10

14)と、

を含む、前記方法。

[態様2]

前記仮想のスピーカの位置に対する前記係数が重み係数

$$g = \frac{1}{\sqrt{L}}$$

を用いて重み付けされ、Lはスピーカの数である、態様1に記載の方法。

[態様3]

仮想のスピーカの前記少なくとも1つの仮想の位置

20

($\hat{\Omega}'_{L+1}$)

は、

$$\hat{\Omega}'_{L+1} = [0, 0]^T$$

および

$$\hat{\Omega}'_{L+1} = [\pi, 0]^T$$

30

のうちの一方である、態様1または2に記載の方法。

[態様4]

フロベニウス・ノルムを使用して前記ダウンスケーリング済みの3D復号行列

(\tilde{D})

を正規化するステップ(13)をさらに含み、正規化されたダウンスケーリング済みの3D復号行列(D)が取得され、前記符号化されたオーディオ信号を復号するステップ(14)は、前記正規化されたダウンスケーリング済みの3D復号行列(D)を使用する、態様1~3のいずれか1項に記載の方法。

40

[態様5]

前記正規化が

$$D = \frac{\tilde{D}}{\sqrt{\sum_{l=1}^L \sum_{q=1}^{0_{3D}} |\tilde{d}_{l,q}|^2}}.$$

に従って行われる、態様4に記載の方法。

[態様6]

50

- 前記 L 個のスピーカの位置

$(\hat{\Omega}_1, \dots, \hat{\Omega}_L)$

および音場信号の係数の次数 N を特定するステップ (101) と、

- 前記位置から前記 L 個のスピーカが実質的に 2D 平面にあると特定するステップ (102) と、

- 仮想のスピーカの少なくとも 1 つの仮想の位置

$(\hat{\Omega}'_{L+1})$

10

を生成するステップ (103) と、

をさらに含む、態様 1 ~ 5 のいずれか 1 項に記載の方法。

[態様 7]

前記符号化されたオーディオ信号を帯域通過フィルタを使用して複数の周波数帯域に分離するステップをさらに含み、各周波数帯域に対して 1 つの、複数の別個の 3D 復号行列 (D_b') が生成され (711b)、各 3D 復号行列 (D_b') はダウンミキシングされ (712b)、必要に応じて別個に正規化され (713b)、前記符号化されたオーディオ信号 (i14) を復号するステップ (714b) は各周波数帯域に対して別個に行われる、態様 1 ~ 6 のいずれか 1 項に記載の方法。

20

[態様 8]

前記既知の L 個のスピーカの位置は、概ね 1 つの 2D 平面内にあり、仰角が 10° 以下である、態様 1 ~ 7 のいずれか 1 項に記載の方法。

[態様 9]

既知の位置にある L 個のスピーカのためのアンビソニックス形式の符号化されたオーディオ信号を復号する装置であって、

- 少なくとも 1 つの仮想のスピーカの少なくとも 1 つの位置を前記 L 個のスピーカの位置に追加する追加部 (410) と、

- 3D 復号行列 (D') を生成する復号行列生成部 (411) であって、前記 L 個のスピーカの位置

30

$(\hat{\Omega}_1, \dots, \hat{\Omega}_L)$

および前記少なくとも 1 つの仮想の位置

$(\hat{\Omega}'_{L+1})$

が使用され、前記 3D 復号行列 (D') が前記特定されたスピーカおよび仮想のスピーカの位置に対する係数を有する、前記復号行列生成部 (411) と、

- 前記 3D 復号行列 (D') をダウンミキシングする行列ダウンミキシング部 (412) であって、前記仮想のスピーカの位置に対する係数が重み付けされ、前記特定されたスピーカの位置に関する係数に分配され、前記特定されたスピーカの位置に対する係数を有するダウンスケーリングされた 3D 復号行列

40

(\tilde{D})

が取得される、前記行列ダウンミキシング部 (412) と、

- 前記ダウンスケーリングされた 3D 復号行列

【 数 16 】

(\tilde{D})

50

を使用して前記符号化されたオーディオ信号(i 1 4)を復号する復号部(4 1 4)であつて、複数の復号されたスピーカ信号(q 1 4)が取得される、前記復号部(4 1 4)と、
を備える、前記装置。

[態様 1 0]

フロベニウス・ノルムを使用して前記ダウンスケーリングされた 3 D 復号行列

(\tilde{D})

を正規化する正規化部(4 1 3)をさらに含み、

10

正規化されたダウンスケーリング済みの 3 D 復号行列(D)が取得され、前記復号部(4 1 4)は、前記正規化されたダウンスケーリング済みの 3 D 復号行列(D)を使用する、
態様 9 に記載の装置。

[態様 1 1]

- 前記 L 個のスピーカの位置

($\hat{\Omega}_1, \dots, \hat{\Omega}_L$)

および音場信号の係数の次数 N を特定する第 1 の特定部(1 0 1)と、

- 前記位置から前記 L 個のスピーカが概ね 2 D 平面にあると特定する第 2 の特定部(1 0 2)と、

20

- 仮想のスピーカの少なくとも 1 つの仮想の位置

($\hat{\Omega}'_{L+1}$)

を生成する仮想スピーカ位置生成部(1 0 3)と、

をさらに含む、態様 9 または 1 0 に記載の装置。

[態様 1 2]

前記符号化されたオーディオ信号を複数の周波数帯域に分離する複数の帯域通過フィルタ(7 1 5 b)をさらに含み、各周波数帯域に対して 1 つ、複数の別個の 3 D 復号行列(D b')が生成され(7 1 1 b)、各 3 D 復号行列(D b')は、ダウンミキシングされ(7 1 2 b)、必要に応じて別個に正規化され(7 1 3 b)、前記符号化されたオーディオ信号(i 1 4)を復号する部(7 1 4 b)は、各周波数帯域を別個に復号する、
態様 9 ~ 1 1 のいずれか 1 項に記載の装置。

30

[態様 1 3]

既知の位置にある L 個のスピーカのためのアンビソニックス形式の符号化されたオーディオ信号を復号する方法をコンピュータに行わせるための実行可能な命令を記憶したコンピュータ読取可能な記憶媒体であって、前記方法は、

- 少なくとも 1 つの仮想のスピーカの少なくとも 1 つの位置を前記 L 個のスピーカの位置に追加するステップ(1 0)と、

40

- 3 D 復号行列(D')を生成するステップ(1 1)であって、前記 L 個のスピーカの位置

($\hat{\Omega}_1, \dots, \hat{\Omega}_L$)

および前記少なくとも 1 つの仮想の位置

($\hat{\Omega}'_{L+1}$)

が使用され、前記 3 D 復号行列(D')が前記特定されたスピーカおよび仮想のスピーカ

50

の位置に対する係数を有する、前記生成するステップ(11)と、

- 前記3D復号行列(D')をダウンミキシングするステップ(12)であって、前記仮想のスピーカの位置に対する係数が重み付けされ、前記特定されたスピーカの位置に関する係数に分配され、前記特定されたスピーカの位置に対する係数を有するダウンスケーリングされた3D復号行列

(\tilde{D})

が取得される、前記ダウンミキシングするステップ(12)と、

- 前記ダウンスケーリングされた3D復号行列

10

(\tilde{D})

を使用して前記符号化されたオーディオ信号(i14)を復号するステップ(14)であって、複数の復号されたスピーカ信号(q14)が取得される、前記復号するステップ(14)と、

を含む、前記コンピュータ読取可能な記憶媒体。

[態様14]

前記仮想のスピーカの位置に対する前記係数が重み係数

$$g = \frac{1}{\sqrt{L}}$$

20

を用いて重み付けされ、Lは、スピーカの数である、態様13に記載のコンピュータ読取可能な記憶媒体。

[態様15]

仮想のスピーカの前記少なくとも1つの仮想の位置

($\widehat{\Omega}'_{L+1}$)

は、

30

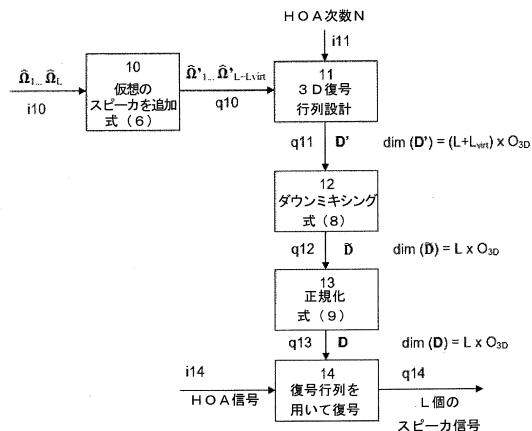
$$\widehat{\Omega}'_{L+1} = [0, 0]^T$$

および

$$\widehat{\Omega}'_{L+1} = [\pi, 0]^T$$

のうちの一方である、態様13または14に記載のコンピュータ読取可能な記憶媒体。

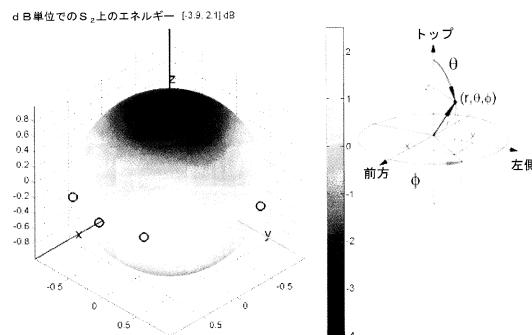
【 四 1 】



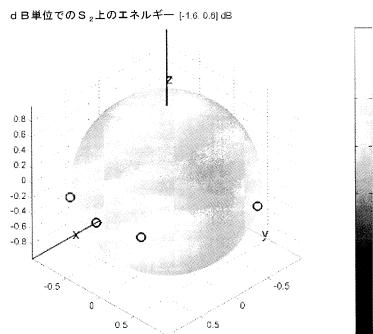
【 図 2 】

$$D' = \begin{bmatrix} d'_{1,1} & d'_{1,2} & \dots & d'_{1,03D} \\ \dots & \dots & \dots & \dots \\ d'_{L,1} & d'_{L,2} & \dots & d'_{L,03D} \\ d'_{L+1,1} & d'_{L+1,2} & \dots & d'_{L+1,03D} \\ d'_{L+2,1} & d'_{L+2,2} & \dots & d'_{L+2,03D} \end{bmatrix} \rightarrow \bar{D} = \begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \dots & \bar{d}_{1,03D} \\ \dots & \dots & \dots & \dots \\ \bar{d}_{L,1} & \bar{d}_{L,2} & \dots & \bar{d}_{L,03D} \end{bmatrix}$$

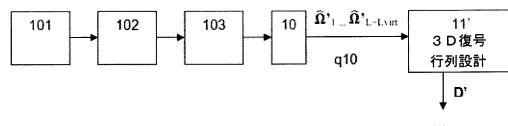
【圖 5】



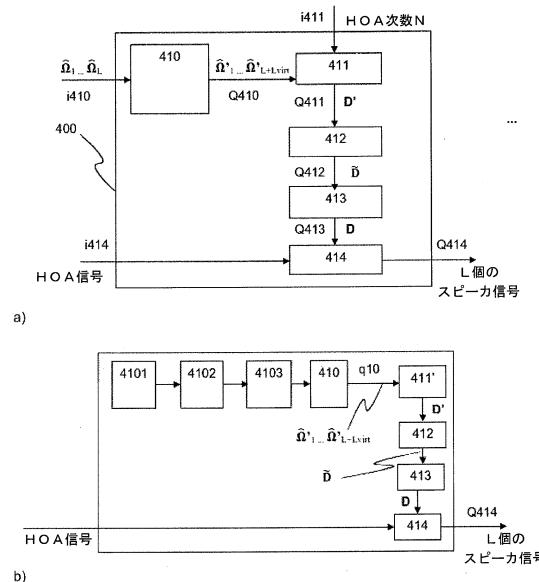
【圖 6】



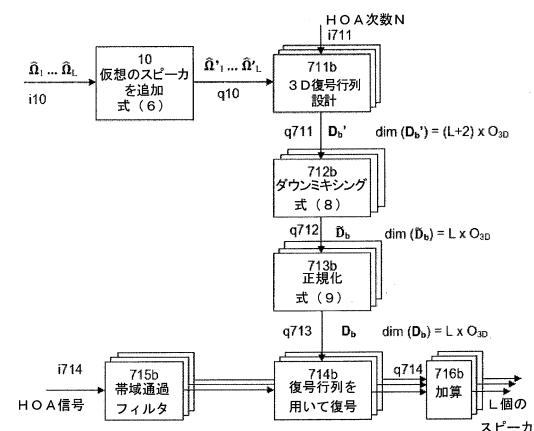
【圖 3】



(义 4)



〔 四 7 〕



フロントページの続き

(72)発明者 ケイラー , フロリアン

ドイツ国 30625 ハノーファー カール - ヴィーヘルト - アレー 74 ドイチエ・トムソン・オーハーゲー リサーチ・アンド・イノベーション

(72)発明者 ベーム , ヨハネス

ドイツ国 30625 ハノーファー カール - ヴィーヘルト - アレー 74 ドイチエ・トムソン・オーハーゲー リサーチ・アンド・イノベーション

審査官 大石 剛

(56)参考文献 特表2006-506918 (JP, A)

国際公開第2011/117399 (WO, A1)

欧州特許出願公開第02645748 (EP, A1)

米国特許第05594800 (US, A)

(58)調査した分野(Int.Cl., DB名)

H04S 7/00