

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6309545号
(P6309545)

(45) 発行日 平成30年4月11日(2018.4.11)

(24) 登録日 平成30年3月23日(2018.3.23)

(51) Int.Cl. F I
HO4S 7/00 (2006.01) HO4S 7/00 380
 HO4S 7/00 300

請求項の数 29 (全 55 頁)

(21) 出願番号	特願2015-557125 (P2015-557125)	(73) 特許権者	595020643
(86) (22) 出願日	平成26年2月7日(2014.2.7)		クアアルコム・インコーポレイテッド
(65) 公表番号	特表2016-509819 (P2016-509819A)		QUALCOMM INCORPORATED
(43) 公表日	平成28年3月31日(2016.3.31)		ED
(86) 国際出願番号	PCT/US2014/015311		アメリカ合衆国、カリフォルニア州 92
(87) 国際公開番号	W02014/124264		121-1714、サン・ディエゴ、モア
(87) 国際公開日	平成26年8月14日(2014.8.14)		ハウス・ドライブ 5775
審査請求日	平成29年1月12日(2017.1.12)	(74) 代理人	100108855
(31) 優先権主張番号	61/762,302		弁理士 蔵田 昌俊
(32) 優先日	平成25年2月7日(2013.2.7)	(74) 代理人	100109830
(33) 優先権主張国	米国 (US)		弁理士 福原 淑弘
(31) 優先権主張番号	61/829,832	(74) 代理人	100158805
(32) 優先日	平成25年5月31日(2013.5.31)		弁理士 井関 守三
(33) 優先権主張国	米国 (US)	(74) 代理人	100194814
			弁理士 奥村 元宏

最終頁に続く

(54) 【発明の名称】 球面調和係数のためのレンダラを決定すること

(57) 【特許請求の範囲】

【請求項1】

デバイスの1つまたは複数のプロセッサが、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカークーラルスピーカークー幾何学的配置を決定することと、

前記1つまたは複数のプロセッサが、前記ローカルスピーカークー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することと、

前記1つまたは複数のプロセッサが、前記決定された2次元レンダラまたは3次元レンダラを使用して前記球面調和係数からマルチチャンネルオーディオデータをレンダリングすることと、前記マルチチャンネルオーディオデータは、空間領域において定義される、を備える、方法。

【請求項2】

前記ローカルスピーカークー幾何学的配置に基づいて前記2次元レンダラまたは3次元レンダラを決定することは、前記ローカルスピーカークー幾何学的配置がステレオスピーカークー幾何学的配置に合致するとき、2次元ステレオレンダラを決定することを備える、請求項1に記載の方法。

【請求項3】

前記ローカルスピーカークー幾何学的配置に基づいて前記2次元レンダラまたは3次元レンダラを決定することは、前記ローカルスピーカークー幾何学的配置が、3つ以上のスピーカークーを有する水平マルチチャンネルスピーカークー幾何学的配置に合致するとき、水平2次元マルチ

チャンネルレングラを決定することを備える、請求項 1 に記載の方法。

【請求項 4】

前記水平 2 次元マルチチャンネルレングラを決定することは、前記決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な水平 2 次元マルチチャンネルレングラを決定することを備える、請求項 3 に記載の方法。

【請求項 5】

前記水平 2 次元マルチチャンネルレングラを決定することは、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な水平 2 次元マルチチャンネルレングラを決定することを備える、請求項 3 に記載の方法。

【請求項 6】

前記ローカルスピーカー幾何学的配置に基づいて前記 2 次元レングラまたは 3 次元レングラを決定することは、前記ローカルスピーカー幾何学的配置が、2 つ以上の水平面上に 3 つ以上のスピーカーを有する 3 次元マルチチャンネルスピーカー幾何学的配置に合致するとき、3 次元マルチチャンネルレングラを決定することを備える、請求項 1 に記載の方法。

【請求項 7】

前記 3 次元マルチチャンネルレングラを決定することは、前記決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な 3 次元マルチチャンネルレングラを決定することを備える、請求項 6 に記載の方法。

【請求項 8】

前記 3 次元マルチチャンネルレングラを決定することは、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な 3 次元マルチチャンネルレングラを決定することを備える、請求項 6 に記載の方法。

【請求項 9】

前記 3 次元マルチチャンネルレングラを決定することは、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な 3 次元マルチチャンネルレングラを決定することを備える、請求項 6 に記載の方法。

【請求項 10】

前記 2 次元レングラまたは 3 次元レングラを決定することは、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカー幾何学的配置を鑑みてレングラリングされることを必要とされる前記球面調和係数のものを識別する、

前記決定された許容次数に基づいて前記 2 次元レングラまたは 3 次元レングラを決定することとを備える、請求項 1 に記載の方法。

【請求項 11】

前記 2 次元レングラまたは 3 次元レングラを決定することは、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカー幾何学的配置を鑑みてレングラリングされることを必要とされる前記球面調和係数のものを識別する、

前記 2 次元レングラまたは 3 次元レングラが、前記決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた前記球面調和係数のもののみをレングラリングするように、前記 2 次元レングラまたは 3 次元レングラを決定することとを備える、請求項 1 に記載の方法。

【請求項 12】

前記 1 つまたは複数のスピーカーの前記ローカルスピーカー幾何学的配置を決定することが、前記ローカルスピーカー幾何学的配置を記述するローカルスピーカー幾何学的配置情報を指定する入力を聴取者から受信することを備える、請求項 1 に記載の方法。

【請求項 13】

前記ローカルスピーカー幾何学的配置に基づいて前記 2 次元レングラまたは 3 次元レングラを決定することは、前記ローカルスピーカー幾何学的配置がモノスピーカー幾何学的配置に合致するとき、モノレングラを決定することを備える、請求項 1 に記載の方法。

10

20

30

40

50

【請求項 14】

デバイスであって、

音場を表す球面調和係数の再生のために使用される 1 つまたは複数のスピーカークールスピーカージェOMETRIC配置を決定することと、

前記ロークールスピーカージェOMETRIC配置に基づいて 2 次元レンダラまたは 3 次元レンダラを決定することと、

前記球面調和係数からマルチチャンネルオーディオデータをレンダリングするために前記決定された 2 次元レンダラまたは 3 次元レンダラに従って動作するように前記デバイスを構成することと、前記マルチチャンネルオーディオデータは、空間領域において定義される、

を行うようにさらに構成された 1 つまたは複数のプロセッサと、

前記 1 つまたは複数のプロセッサに結合され、前記決定された 2 次元レンダラまたは 3 次元レンダラを記憶するように構成されたメモリとを備える、デバイス。

【請求項 15】

前記 1 つまたは複数のプロセッサは、前記ロークールスピーカージェOMETRIC配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ロークールスピーカージェOMETRIC配置がステレオスピーカージェOMETRIC配置に合致するときに 2 次元ステレオレンダラを決定するようにさらに構成された、請求項 14 に記載のデバイス。

【請求項 16】

前記 1 つまたは複数のプロセッサは、前記ロークールスピーカージェOMETRIC配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ロークールスピーカージェOMETRIC配置が、3 つ以上のスピーカークールスピーカージェOMETRIC配置に合致するときに水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 14 に記載のデバイス。

【請求項 17】

前記 1 つまたは複数のプロセッサは、前記水平 2 次元マルチチャンネルレンダラを決定するとき、前記決定されたロークールスピーカージェOMETRIC配置が不規則なスピーカージェOMETRIC配置を示すときに不規則な水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 16 に記載のデバイス。

【請求項 18】

前記 1 つまたは複数のプロセッサは、前記水平 2 次元マルチチャンネルレンダラを決定するとき、前記決定されたロークールスピーカージェOMETRIC配置が規則的なスピーカージェOMETRIC配置を示すときに規則的な水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 16 に記載のデバイス。

【請求項 19】

前記 1 つまたは複数のプロセッサは、前記ロークールスピーカージェOMETRIC配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ロークールスピーカージェOMETRIC配置が、2 つ以上の水平面上に 3 つ以上のスピーカークールスピーカージェOMETRIC配置を合致するとき、3 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 14 に記載のデバイス。

【請求項 20】

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたロークールスピーカージェOMETRIC配置が不規則なスピーカージェOMETRIC配置を示すときに不規則な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 19 に記載のデバイス。

【請求項 21】

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたロークールスピーカージェOMETRIC配置が規則的なスピーカージェOMETRIC配置を示すときに規則的な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 19 に記載のデバイス。

10

20

30

40

50

【請求項 2 2】

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカ幾何学的配置が規則的なスピーカ幾何学的配置を示すときに規則的な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、請求項 1 9 に記載のデバイス。

【請求項 2 3】

前記 1 つまたは複数のプロセッサは、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる前記球面調和係数のものを識別する、前記決定された許容次数に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定することとを行うように構成された、請求項 1 4 に記載のデバイス。

10

【請求項 2 4】

前記 1 つまたは複数のプロセッサは、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる前記球面調和係数のものを識別する、前記 2 次元レンダラまたは 3 次元レンダラが、前記決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた前記球面調和係数のものみをレンダリングするように、前記 2 次元レンダラまたは 3 次元レンダラを決定することとを行うように構成された、請求項 1 4 に記載のデバイス。

【請求項 2 5】

前記 1 つまたは複数のプロセッサが、前記ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を聴取者から受信するように構成された、請求項 1 4 に記載のデバイス。

20

【請求項 2 6】

前記 1 つまたは複数のプロセッサは、前記ローカルスピーカ幾何学的配置がモノスピーカ幾何学的配置に合致するときモノレンダラを決定するように構成された、請求項 1 4 に記載のデバイス。

【請求項 2 7】

実行されたとき、

音場を表す球面調和係数の再生のために使用される 1 つまたは複数のスピーカのローカルスピーカ幾何学的配置を決定することと、

30

前記ローカルスピーカ幾何学的配置に基づいて 2 次元レンダラまたは 3 次元レンダラを決定することと、

マルチチャンネルオーディオデータを生成するために前記決定された 2 次元レンダラまたは 3 次元レンダラを使用して前記球面調和係数をレンダリングすることと、前記マルチチャンネルオーディオデータは、空間領域において定義される、を 1 つまたは複数のプロセッサに行わせる命令を記憶した非一時的コンピュータ可読記憶媒体。

【請求項 2 8】

前記 1 つまたは複数のプロセッサに結合された前記 1 つまたは複数のスピーカが、前記マルチチャンネルオーディオデータに基づいて前記音場を再現することをさらに備える、請求項 1 に記載の方法。

40

【請求項 2 9】

前記 1 つまたは複数のスピーカをさらに備え、前記 1 つまたは複数のスピーカは、前記 1 つまたは複数のプロセッサに結合され、前記マルチチャンネルオーディオデータに基づいて前記音場を再現するように構成された、請求項 1 4 に記載のデバイス。

【発明の詳細な説明】

【技術分野】

【0001】

[0001] 本出願は、2013年5月31日に出願された米国仮出願第61/829,832号および2013年2月7日に出願された米国仮出願第61/762,302号の利

50

益を主張する。

【0002】

[0002] 本開示は、オーディオレンダリング (audio rendering) に関し、より詳細には、球面調和係数 (spherical harmonic coefficient) のレンダリングに関する。

【背景技術】

【0003】

[0003] 高次アンビソニックス (H O A : higher order ambisonics) 信号 (しばしば複数の球面調和係数 (S H C : spherical harmonic coefficient) または他の階層要素 (hierarchical element) によって表される) は、音場の3次元表現である。このH O AまたはS H C表現は、このS H C信号からレンダリングされたマルチチャンネルオーディオ信号を再生するために使用されるローカルスピーカー幾何学的配置 (local speaker geometry) に依存しない様式でこの音場を表し得る。このS H C信号はまた、このS H C信号が、5.1オーディオチャンネルフォーマットまたは7.1オーディオチャンネルフォーマットなど、よく知られ大いに採用されているマルチチャンネルフォーマットにレンダリングされ得るような後方互換性を可能にし得る。S H C表現は、したがって、後方互換性にも適応する音場のより良い表現を可能にする。

【発明の概要】

【0004】

[0004] 概して、特定のローカルスピーカー幾何学的配置に適するオーディオレンダラを決定するための技法について説明される。S H Cはよく知られているマルチチャンネルスピーカーフォーマットに適応し得るが、通常、エンドユーザリスナーは、これらのマルチチャンネルフォーマットによって要求される様式でスピーカーを適切に置いたりまたは配置したりせず、その結果、不規則なスピーカー幾何学的配置 (irregular speaker geometry) が生じる。本開示で説明される技法は、ローカルスピーカー幾何学的配置を決定し、次いで、このローカルスピーカー幾何学的配置に基づいてS H C信号をレンダリングするためのレンダラを決定し得る。レンダリングデバイスは、いくつかの異なるレンダラ、たとえば、モノレンダラ、ステレオレンダラ、水平専用レンダラまたは3次元レンダラの中から選択し、ローカルスピーカー幾何学的配置に基づいてこのレンダラを生成し得る。このレンダラは、不規則なスピーカー幾何学的配置を考慮し、それによって、規則的なスピーカー幾何学的配置 (regular speaker geometry) のために設計された規則的なレンダラと比較して不規則なスピーカー幾何学的配置にもかかわらず、音場のより良い再現を可能にし得る。

【0005】

[0005] その上、本技法は、可逆性を維持しS H Cを復元するように、仮想スピーカー幾何学的配置 (virtual speaker geometry) と呼ばれ得る、均一なスピーカー幾何学的配置にレンダリングし得る。本技法は、その場合、これらの仮想スピーカーを (仮想スピーカーが最初に配置された水平面とは異なる仰角 (elevation) であり得る) 様々な水平面に投射 (project) するための様々な動作を実行し得る。本技法は、デバイスが、これらの投射された仮想スピーカーを、不規則なスピーカー幾何学的配置で配置された様々な物理スピーカーにマッピングするレンダラを生成することを可能にし得る。これらの仮想スピーカーをこのように投射することは、音場のより良い再現を可能にし得る。

【0006】

[0006] 一例では、方法が、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカーのローカルスピーカー幾何学的配置を決定することと、ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することとを備える。

【0007】

[0007] 別の例では、デバイスが、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカーのローカルスピーカー幾何学的配置を決定することと、決定されたローカルスピーカー幾何学的配置に基づいて動作するようにデバイスを構成するこ

10

20

30

40

50

とを行うように構成された1つまたは複数のプロセッサを備える。

【0008】

[0008] 別の例では、デバイスが、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカークーラルスピーカークー幾何学的配置を決定するための手段と、クーラルスピーカークー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定するための手段とを備える。

【0009】

[0009] 別の例では、非一時的コンピュータ可読記憶媒体は、実行されたとき、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカークーラルスピーカークー幾何学的配置を決定することと、クーラルスピーカークー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することとを1つまたは複数のプロセッサに行わせる命令を記憶している。

10

【0010】

[0010] 別の例では、方法が、複数の物理スピーカークーのうちの一つと幾何学的配置で配置された複数の仮想スピーカークーのうちの一つとの間の位置の差を決定することと、位置の決定された差に基づいて、および複数の仮想スピーカークーを複数の物理スピーカークーにマッピングするより前に、幾何学的配置内の複数の仮想スピーカークーのうちの一つの位置を調整することとを備える。

【0011】

[0011] 別の例では、デバイスが、複数の物理スピーカークーのうちの一つと幾何学的配置で配置された複数の仮想スピーカークーのうちの一つとの間の位置の差を決定することと、位置の決定された差に基づいて、および複数の仮想スピーカークーを複数の物理スピーカークーにマッピングするより前に、幾何学的配置内の複数の仮想スピーカークーのうちの一つの位置を調整することとを行うように構成された1つまたは複数のプロセッサを備える。

20

【0012】

[0012] 別の例では、デバイスが、複数の物理スピーカークーのうちの一つと幾何学的配置で配置された複数の仮想スピーカークーのうちの一つとの間の位置の差を決定するための手段と、位置の決定された差に基づいて、および複数の仮想スピーカークーを複数の物理スピーカークーにマッピングするより前に、幾何学的配置内の複数の仮想スピーカークーのうちの一つの位置を調整するための手段とを備える。

30

【0013】

[0013] 別の例では、非一時的コンピュータ可読記憶媒体は、実行されたとき、複数の物理スピーカークーのうちの一つと幾何学的配置で配置された複数の仮想スピーカークーのうちの一つとの間の位置の差を決定することと、位置の決定された差に基づいて、および複数の仮想スピーカークーを複数の物理スピーカークーにマッピングするより前に、幾何学的配置内の複数の仮想スピーカークーのうちの一つの位置を調整することとを1つまたは複数のプロセッサに行わせる命令を記憶している。

【0014】

[0014] 本技法の1つまたは複数の態様の詳細が添付の図面および以下の説明に記載されている。これらの技法の他の特徴、目的、および利点は、その説明および図面から、また特許請求の範囲から明らかになる。

40

【図面の簡単な説明】

【0015】

【図1】 [0015] 様々な次数および副次数の球面調和基底関数を示す図。

【図2】 様々な次数および副次数の球面調和基底関数を示す図。

【図3】 [0016] 本開示で説明される技法の様々な態様を実装し得るシステムを示す図。

【図4】 [0017] 本開示で説明される技法の様々な態様を実装し得るシステムを示す図。

【図5】 [0018] 本開示で説明される技法の様々な態様を実施する際の、図4の例に示されたレンダラ決定ユニットの例示的な動作を示す流れ図。

【図6】 [0019] 図4の例に示されたステレオレンダラ生成ユニットの例示的な動作を示

50

す流れ図。

【図 7】[0020] 図 4 の例に示された水平レンダラ生成ユニットの例示的な動作を示す流れ図。

【図 8 A】[0021] 図 4 の例に示された 3 D レンダラ生成ユニットの例示的な動作を示す流れ図。

【図 8 B】図 4 の例に示された 3 D レンダラ生成ユニットの例示的な動作を示す流れ図。

【図 9】[0022] 不規則な 3 D レンダラを決定するときには下半球処理と上半球処理とを実行する際の、図 4 の例に示された 3 D レンダラ生成ユニットの例示的な動作を示す流れ図。

【図 1 0】[0023] 本開示に記載された技法に従ってどのようにステレオレンダラが生成され得るかを示すユニット空間におけるグラフ 2 9 9 を示す図。 10

【図 1 1】[0024] 本開示に記載された技法に従ってどのように不規則な水平レンダラが生成され得るかを示すユニット空間におけるグラフ 3 0 4 を示す図。

【図 1 2 A】[0025] 本開示で説明される技法に従ってどのように不規則な 3 D レンダラが生成され得るかを示すグラフ 3 0 6 A を示す図。

【図 1 2 B】本開示で説明される技法に従ってどのように不規則な 3 D レンダラが生成され得るかを示すグラフ 3 0 6 B を示す図。

【図 1 3 A】[0026] 本開示で説明される技法の様々な態様に従って形成されるビットストリームを示す図。

【図 1 3 B】本開示で説明される技法の様々な態様に従って形成されるビットストリームを示す図。 20

【図 1 3 C】本開示で説明される技法の様々な態様に従って形成されるビットストリームを示す図。

【図 1 3 D】本開示で説明される技法の様々な態様に従って形成されるビットストリームを示す図。

【図 1 4 A】[0027] 本開示で説明される技法の様々な態様を実装し得る 3 D レンダラ決定ユニットを示す図。

【図 1 4 B】本開示で説明される技法の様々な態様を実装し得る 3 D レンダラ決定ユニットを示す図。

【図 1 5 A】[0028] 2 2 . 2 スピーカー幾何学的配置を示す図。 30

【図 1 5 B】2 2 . 2 スピーカー幾何学的配置を示す図。

【図 1 6 A】[0029] 本開示で説明される技法の様々な態様による、仮想スピーカーのうちの 1 つまたは複数が投射された水平面によってセグメント化される、仮想スピーカーがその上に配置された仮想球体を示す図。

【図 1 6 B】本開示で説明される技法の様々な態様による、仮想スピーカーのうちの 1 つまたは複数が投射された水平面によってセグメント化される、仮想スピーカーがその上に配置された仮想球体を示す図。

【図 1 7】[0030] 本開示で説明される技法の様々な態様による、要素の階層セットに適用され得るウィンドウ処理関数を示す図。

【発明を実施するための形態】 40

【0 0 1 6】

[0031] サラウンド音の発展は、現今では娯楽のための多くの出力フォーマットを利用可能にしている。そのようなサラウンド音フォーマットの例は、普及している 5 . 1 フォーマット（これは、フロントレフト（FL）と、フロントライト（FR）と、センターまたはフロントセンターと、バックレフトまたはサラウンドレフトと、バックライトまたはサラウンドライトと、低周波効果（LFE：low frequency effect）という、6 つのチャネルを含む）、発展中の 7 . 1 フォーマット、および来るべき 2 2 . 2 フォーマット（たとえば、超高精細度テレビジョン規格とともに使用するための）を含む。さらなる例は、球面調和アレイ（spherical harmonic array）のためのフォーマットを含む。

【0 0 1 7】 50

【0032】 (2013年1月付けの、ジュネーブ、スイスにおいて協定で公開された、「Call for Proposals for 3D Audio」と題する、ISO/IEC JTC1/SC29/WG11/N13411文書に回答して概して開発され得る) 将来のMPEGエンコーダへの入力は、随意に以下の3つの可能なフォーマットのうちの1つである。(i) 事前指定された位置にあるラウドスピーカを通して再生されるように意図された、従来のチャンネルベースのオーディオ、(ii) (情報の中でも) 単一オーディオオブジェクトのロケーション座標を含んでいる関連するメタデータ (metadata) とともに単一オーディオオブジェクトのための離散パルスコード変調 (PCM) データを伴う、オブジェクトベースのオーディオ、および (iii) 球面調和基底関数の係数 (「球面調和係数 (spherical harmonic coefficient)」またはSHCとも呼ばれる) を使用して音場を表すことを伴う、シーンベースのオーディオ。

10

【0018】

【0033】 市場には様々な「サラウンド音 (surround-sound)」フォーマットがある。それらは、たとえば、5.1ホームシアターシステム (これは、ステレオを超えたリビングルームへの進出に関して最も成功している) から、NHK (日本放送協会 (Nippon Hoso Kyokai) または日本放送協会 (Japan Broadcasting Corporation)) によって開発された22.2システムにまでわたる。コンテンツ作成者 (たとえば、ハリウッドスタジオ) は、一度に映画のサウンドトラックを作成することを望み、スピーカー構成ごとにサウンドトラックをリミックス (remix) する労力を費やすことを望まない。最近では、規格化委員会が、規格化されたビットストリームへの符号化と、スピーカーの幾何学的配置およびレンダラの位置における音響条件に適応可能でアグノスティック (agnostic) である後続の復号とを提供するための方法を考察している。

20

【0019】

【0034】 コンテンツ作成者にそのようなフレキシビリティを提供するために、要素の階層セットが音場を表すために使用され得る。要素の階層セットは、モデル化された音場の完全な表現をより低次の要素の基本セットが提供するように要素が順序付けられる、要素のセットを指し得る。このセットはより高次の要素を含むように拡張されるので、表現はより詳細なものになる。

【0020】

【0035】 要素の階層セットの一例は、球面調和係数 (SHC) のセットである。次の式は、SHCを使用した音場の記述または表現を示す。

30

【数1】

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

【0021】

この式は、音場の任意の点 $\{r_r, \theta_r, \varphi_r\}$ における圧力 p_i がSHC

【数2】

40

$$A_n^m(k)$$

【0022】

によって一意に表され得ることを示す。ここで、

【数3】

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

【0023】

50

であり、 c は音速(約343 m/s)であり、 $\{r_r, r_r, r_r\}$ は基準の点(または観測点)であり、 $j_n(\cdot)$ は次数 n の球ベッセル関数(spherical Bessel function)であり、

【数4】

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

【0024】

は次数 n と副次数 m との球面調和基底関数である。角括弧の中の項は、離散フーリエ変換(DFT)、離散コサイン変換(DCT)、またはウェーブレット変換などの様々な時間周波数変換によって近似され得る、信号の周波数領域の表現(すなわち、 $S(\cdot, r_r, r_r, r_r)$)であることが認識され得る。階層セットの他の例は、ウェーブレット変換係数のセット、および多分解能基底関数(multiresolution basis function)の係数の他のセットを含む。

10

【0025】

[0036] 図1は、0次($n=0$)から4次($n=4$)までの球面調和基底関数を示す図である。わかるように、各次数について、図示されてはいるが説明を簡単にするために図2の例では明示的に述べられていない副次数 m という拡張がある。

【0026】

[0037] 図2は、0次($n=0$)から4次($n=4$)までの球面調和基底関数を示す別の図である。図2では、球面調和基底関数は、図示された次数と副次数の両方を伴う3次元座標空間において示されている。

20

【0027】

[0038] いずれの場合も、SHC

【数5】

$$A_n^m(k)$$

【0028】

は、様々なマイクロフォンアレイ構成によって物理的に取得(たとえば、記録)され得、あるいは代替的に、それらは音場のチャンネルベースまたはオブジェクトベースの記述から導出され得る。前者は、エンコーダへのシーンベース(scene-based)のオーディオ入力を表す。たとえば、 $1+2^4$ 個の(25個の、したがって4次)係数を伴う4次表現が使用され得る。

30

【0029】

[0039] これらのSHCがどのようにオブジェクトベースの記述から導出され得るかを例示するために、次の式について考える。個々のオーディオオブジェクトに対応する音場の係数

【数6】

$$A_n^m(k)$$

40

【0030】

は、

【数7】

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

50

【 0 0 3 1 】

として表され得、

ただし、 i は

【 数 8 】

$$\sqrt{-1}$$

【 0 0 3 2 】

であり、

【 数 9 】

10

$$h_n^{(2)}(\cdot)$$

【 0 0 3 3 】

は次数 n の (第 2 の種類) の球ハンケル関数 (spherical Hankel function) であり、 $\{ r_s, \theta_s, \phi_s \}$ はオブジェクトのロケーションである。周波数の関数として (たとえば、PCM ストリームに対して高速フーリエ変換を実行するなど、時間周波数分析技法を使用して) ソースエネルギー $g(\cdot)$ を知ることで、各 PCM オブジェクトとそのロケーションとを SHC

【 数 1 0 】

20

$$A_n^m(k)$$

【 0 0 3 4 】

に変換することが可能になる。さらに、各オブジェクトの

【 数 1 1 】

$$A_n^m(k)$$

【 0 0 3 5 】

係数は、(上記が線形および直交分解であるので) 加法的であることが示され得る。このようにして、多数の PCM オブジェクトが

【 数 1 2 】

30

$$A_n^m(k)$$

【 0 0 3 6 】

係数によって (たとえば、個々のオブジェクトについての係数ベクトルの和として) 表され得る。本質的に、これらの係数は、音場に関する情報 (3D 座標の関数としての圧力) を含んでおり、上記は、観測点 $\{ r_r, \theta_r, \phi_r \}$ の近傍における、音場全体の表現への個々のオブジェクトからの変換を表す。残りの図は、オブジェクトベースおよび SHC ベースのオーディオコーディングのコンテキストにおいて以下で説明される。

【 0 0 3 7 】

[0040] 図 3 は、本開示で説明される技法の様々な態様を実行し得るシステム 20 を示す図である。図 3 の例に示されているように、システム 20 は、コンテンツ作成者 22 とコンテンツ消費者 24 とを含む。コンテンツ作成者 22 は、コンテンツ消費者 24 など、コンテンツ消費者による消費のためにマルチチャネルオーディオコンテンツを生成し得る映画スタジオまたは他のエンティティを表し得る。しばしば、このコンテンツ作成者は、ビデオコンテンツと併せてオーディオコンテンツを生成する。コンテンツ消費者 24 は、マルチチャネルオーディオコンテンツを再生することが可能な任意の形態のオーディオ再

40

50

生システムを指し得るオーディオ再生システム 3 2 を所有するかまたはそれへのアクセスを有する個人を表す。図 3 の例では、コンテンツ消費者 2 4 はオーディオ再生システム 3 2 を含む。

【 0 0 3 8 】

[0041] コンテンツ作成者 2 2 は、オーディオレンダラ 2 8 とオーディオ編集システム 3 0 とを含む。オーディオレンダラ 2 6 は、スピーカーフィード（「ラウドスピーカーフィード (loudspeaker feed)」、「スピーカー信号 (speaker signal)」、または「ラウドスピーカー信号 (loudspeaker signal)」とも呼ばれることがある）をレンダリングするかまたはさもなければ生成するオーディオ処理ユニットを表し得る。各スピーカーフィードは、マルチチャンネルオーディオシステムの特定のチャンネルのための音を再現するスピーカーフィードに対応し得る。図 3 の例では、レンダラ 3 8 は、5 . 1、7 . 1 または 2 2 . 2 サラウンド音スピーカーシステム中の 5 つ、7 つまたは 2 2 個のスピーカーの各々のためのスピーカーフィードを生成する、従来の 5 . 1、7 . 1 または 2 2 . 2 サラウンド音フォーマットのためのスピーカーフィードをレンダリングし得る。代替的に、レンダラ 2 8 は、上記で論じられたソース球面調和係数 (source spherical harmonic coefficient) の特性を鑑みて、任意の数のスピーカーを有する任意のスピーカー構成のためにソース球面調和係数からスピーカーフィードをレンダリングするように構成され得る。レンダラ 2 8 は、このようにして、図 3 ではスピーカーフィード 2 9 として示されている、いくつかのスピーカーフィードを生成し得る。

【 0 0 3 9 】

[0042] コンテンツ作成者は、編集プロセス中に、球面調和係数 2 7（「SHC 2 7」）をレンダリングして、高忠実度を有しないかまたは納得のいくサラウンド音エクスペリエンスを提供しない音場の態様を識別しようとする試みにおいて、レンダリングされたスピーカーフィードを聴取し得る。コンテンツ作成者 2 2 は、次いで、（しばしば、上記で説明された様式でソース球面調和係数がそれから導出され得る様々なオブジェクトの操作を通して間接的に）ソース球面調和係数を編集し得る。コンテンツ作成者 2 2 は、球面調和係数 2 7 を編集するためにオーディオ編集システム 3 0 を使用し得る。オーディオ編集システム 3 0 は、オーディオデータを編集し、このオーディオデータを 1 つまたは複数のソース球面調和係数として出力することが可能な任意のシステムを表す。

【 0 0 4 0 】

[0043] 編集プロセスが完了すると、コンテンツ作成者 2 2 は、球面調和係数 2 7 に基づいてビットストリーム 3 1 を生成し得る。すなわち、コンテンツ作成者 2 2 は、ビットストリーム 3 1 を生成することが可能な任意のデバイスを表し得る、ビットストリーム生成デバイス 3 6 を含む。いくつかの事例では、ビットストリーム生成デバイス 3 6 は、球面調和係数 2 7 を（一例として、エンтроピー符号化によって）バンド幅圧縮し、容認されるフォーマットで球面調和係数 2 7 のバンド幅圧縮バージョンを構成してビットストリーム 3 1 を形成するエンコーダを表し得る。他の事例では、ビットストリーム生成デバイス 3 6 は、マルチチャンネルオーディオコンテンツまたはその派生 (derivative) を圧縮するために、一例として、従来のオーディオサラウンド音符号化プロセスのものと同様のプロセスを使用してマルチチャンネルオーディオコンテンツ 2 9 を符号化するオーディオエンコーダ（場合によっては、MPEGサラウンドまたはその派生など、知られているオーディオコーディング規格に準拠するオーディオエンコーダ）を表し得る。圧縮されたマルチチャンネルオーディオコンテンツ 2 9 は、次いで、コンテンツ 2 9 をバンド幅圧縮するためにエンтроピー符号化されるかまたは何らかの他の方法でコーディングされ、合意されたフォーマットに従って構成されて、ビットストリーム 3 1 が形成され得る。直接圧縮されてビットストリーム 3 1 が形成されるか、またはレンダリングされ、次いで圧縮されてビットストリーム 3 1 が形成されるかにかかわらず、コンテンツ作成者 2 2 は、ビットストリーム 3 1 をコンテンツ消費者 2 4 に送信し得る。

【 0 0 4 1 】

[0044] 図 3 ではコンテンツ消費者 2 4 に直接送信されるものとして示されているが、

コンテンツ作成者 2 2 は、コンテンツ作成者 2 2 とコンテンツ消費者 2 4 との間に配置された中間デバイスにビットストリーム 3 1 を出力し得る。この中間デバイスは、このビットストリームを要求し得るコンテンツ消費者 2 4 に後で配信するためにビットストリーム 3 1 を記憶し得る。中間デバイスは、ファイルサーバ、ウェブサーバ、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、スマートフォン、スマートフォンの、またはオーディオデコーダによる後での取出しのためにビットストリーム 3 1 を記憶することが可能な任意の他のデバイスを備え得る。代替的に、コンテンツ作成者 2 2 は、コンパクトディスク、デジタルビデオディスク、高精細度ビデオディスクまたは他の記憶媒体などの記憶媒体にビットストリーム 3 1 を記憶し得、それらの大部分は、コンピュータによって読み取られることが可能であり、したがって、コンピュータ可読記憶媒体と呼ばれることがある。このコンテキストでは、送信チャネルは、これらの媒体に記憶されたコンテンツが送信されるチャネルを指し得る（ならびに、小売店および他の店舗ベースの配信機構を含み得る）。いずれの場合も、本開示の技法は、したがって、この点において図 3 の例に限定されるべきではない。

【 0 0 4 2 】

[0045] 図 3 の例にさらに示されているように、コンテンツ消費者 2 4 はオーディオ再生システム 3 2 を含む。オーディオ再生システム 3 2 は、マルチチャネルオーディオデータを再生することが可能な任意のオーディオ再生システムを表し得る。オーディオ再生システム 3 2 はいくつかの異なるレンダラを含み得る。オーディオ再生システム 3 2 はまた、複数のオーディオレンダラの中からオーディオレンダラ 3 4 を決定するかまたはさもなければ選択するように構成されたユニットを表し得るレンダラ決定ユニット 4 0 を含み得る。いくつかの事例では、レンダラ決定ユニット 4 0 は、いくつかのあらかじめ定義されたレンダラからレンダラ 3 4 を選択し得る。他の事例では、レンダラ決定ユニット 4 0 は、ローカルスピーカ幾何学的配置情報 4 1 に基づいてオーディオレンダラ 3 4 を動的に決定し得る。ローカルスピーカ幾何学的配置情報 4 1 は、オーディオ再生システム 3 2 、聴取者、あるいは任意の他の識別可能な領域またはロケーションに対する、オーディオ再生システム 3 2 に結合された各スピーカのロケーションを指定し得る。しばしば、聴取者は、ローカルスピーカ幾何学的配置情報 4 1 を入力するためにグラフィカルユーザインターフェース (GUI) または他の形態のインターフェースを介してオーディオ再生システム 3 2 とインターフェースし得る。いくつかの事例では、オーディオ再生システム 3 2 は、しばしば、いくつかのトーンを発することと、オーディオ再生システム 3 2 に結合されたマイクロフォンを介してそれらのトーンを測定することとによって自動的に（この例では、いかなる聴取者の介入も必要とせず、を意味する）ローカルスピーカ幾何学的配置情報 4 1 を決定し得る。

【 0 0 4 3 】

[0046] オーディオ再生システム 3 2 はさらに、抽出デバイス 3 8 を含み得る。抽出デバイス 3 8 は、ビットストリーム生成デバイス 3 6 のプロセスとは概して逆であり得るプロセスを通して球面調和係数 2 7 ' (「S H C 2 7 '」、これは球面調和係数 2 7 の修正形態または複製を表し得る) を抽出することが可能な任意のデバイスを表し得る。オーディオ再生システム 3 2 は、球面調和係数 2 7 ' を受信し、S H C 2 7 ' と、指定されたかまたは利用可能な場合はオーディオレンダリング情報 (audio rendering information) 3 9 とを抽出するために抽出デバイス 3 8 を呼び出し得る。

【 0 0 4 4 】

[0047] いずれの場合も、上記のレンダラ 3 4 の各々は、異なる形態のレンダリングを提供し得、ここで、異なる形態のレンダリングは、ベクトルベース振幅パンニング (V B A P : vector-base amplitude panning) を実行する様々な方法のうちの 1 つまたは複数、距離ベースの振幅パンニング (D B A P : distance based amplitude panning) を実行する様々な方法のうちの 1 つまたは複数、単純なパンニングを実行する様々な方法のうちの 1 つまたは複数、ニアフィールド補償 (N F C : near field compensation) フィルタ処理を実行する様々な方法のうちの 1 つまたは複数および/または波動場合成 (wave fie

10

20

30

40

50

ld synthesis) を実行する様々な方法のうちの一つまたは複数を含み得る。選択されたレンダラ 34 は、次いで、球面調和係数 27' をレンダリングして、(説明を簡単にするために図 3 の例には示されていない、オーディオ再生システム 32 に電氣的にまたは場合によってはワイヤレスに結合されたラウドスピーカ-の数に対応する) いくつかのスピーカ-フィールド 35 を生成し得る。

【0045】

[0048] 典型的には、オーディオ再生システム 32 は、複数のオーディオレンダラのうちのいずれかの一つを選択し得、(いくつかの例を挙げると、DVDプレーヤ、Blu-ray (登録商標) プレーヤ、スマートフォン、タブレットコンピュータ、ゲーミングシステム、およびテレビジョンなどの) ビットストリーム 31 が受信されたソースに応じてオーディオレンダラのうちの一つまたは複数を選択するように構成され得る。オーディオレンダラのうちのどの一つでも選択され得るが、しばしば、コンテンツを作成するときに使用されたオーディオレンダラは、コンテンツがオーディオレンダラのうちのこの一つ、すなわち、図 3 の例ではオーディオレンダラ 28 を使用してコンテンツ作成者 22 によって作成されたということに起因して、より良い(および場合によっては最良の)形態のレンダリングを提供する。ローカルスピーカ-幾何学的配置のレンダリング形態と同じであるかまたは少なくともそれに近いレンダリング形態を有するオーディオレンダラ 34 のうちの一つを選択することにより、コンテンツ消費者 24 にとってより良いサラウンド音エクスペリエンスをもたらし得る音場のより良い表現が提供され得る。

【0046】

[0049] ビットストリーム生成デバイスは、オーディオレンダリング情報(audio rendering information) 39 (「オーディオレンダリング情報(audio rendering info) 39」) を含むようにビットストリーム 31 を生成し得る。オーディオレンダリング情報 39 は、マルチチャンネルオーディオコンテンツを生成するときに使用されたオーディオレンダラ、すなわち、図 4 の例ではオーディオレンダラ 28 を識別する信号値を含み得る。いくつかの事例では、信号値は、複数のスピーカ-フィールドへの球面調和係数をレンダリングするために使用される行列を含む。

【0047】

[0050] いくつかの事例では、信号値は、ビットストリームが複数のスピーカ-フィールドへの球面調和係数をレンダリングするために使用される行列を含むことを示すインデックスを定義する 2 つ以上のビットを含む。いくつかの事例では、インデックスが使用されるとき、信号値は、ビットストリーム中に含まれる行列の行の数を定義する 2 つ以上のビットと、ビットストリーム中に含まれる行列の列の数を定義する 2 つ以上のビットとをさらに含む。この情報を使用して、および 2 次元行列の各係数は典型的には 32 ビット浮動小数点数によって定義されることを鑑みて、行列のビットで表されるサイズは、行列の各係数を定義している、行の数と、列の数と、浮動小数点数のサイズ、すなわち、この例では 32 ビットとの関数として計算され得る。

【0048】

[0051] いくつかの事例では、信号値は、複数のスピーカ-フィールドへの球面調和係数をレンダリングするために使用されるレンダリングアルゴリズムを指定する。レンダリングアルゴリズムは、ビットストリーム生成デバイス 36 と抽出デバイス 38 の両方に知られている行列を含み得る。すなわち、レンダリングアルゴリズムは、パンニング(たとえば、VBA P、DBA Pまたは単純なパンニング)あるいはNFCフィルタ処理など、他のレンダリングステップに加えて行列の適用を含み得る。いくつかの事例では、信号値は、複数のスピーカ-フィールドへの球面調和係数をレンダリングするために使用される複数の行列のうちの一つに関連付けられたインデックスを定義する 2 つ以上のビットを含む。この場合も、ビットストリーム生成デバイス 36 と抽出デバイス 38 の両方は、インデックスが複数の行列のうちの特定の 1 つを一意に識別し得るように、複数の行列と複数の行列の次数とを示す情報で構成され得る。代替的に、ビットストリーム生成デバイス 36 は、インデックスが複数の行列のうちの特定の 1 つを一意に識別し得るように、複数の行列

および/または複数の行列の次数を定義するビットストリーム 31 中のデータを指定し得る。

【0049】

[0052] いくつかの事例では、信号値は、複数のスピーカフィールドへの球面調和係数をレンダリングするために使用される複数のレンダリングアルゴリズムのうちの1つに関連付けられたインデックスを定義する2つ以上のビットを含む。この場合も、ビットストリーム生成デバイス 36 と抽出デバイス 38 の両方は、インデックスが複数の行列のうちの特定の1つを一意に識別し得るように、複数のレンダリングアルゴリズムと複数のレンダリングアルゴリズムの次数とを示す情報で構成され得る。代替的に、ビットストリーム生成デバイス 36 は、インデックスが複数の行列のうちの特定の1つを一意に識別し得る

10

【0050】

[0053] いくつかの事例では、ビットストリーム生成デバイス 36 は、ビットストリーム中のオーディオフレームごとにオーディオレンダリング情報 39 を指定する。他の事例では、ビットストリーム生成デバイス 36 は、ビットストリーム中でオーディオレンダリング情報 39 を1回指定する。

【0051】

[0054] 抽出デバイス 38 は、次いで、ビットストリーム中で指定されたオーディオレンダリング情報 39 を決定し得る。オーディオレンダリング情報 39 中に含まれる信号値に基づいて、オーディオ再生システム 32 は、オーディオレンダリング情報 39 に基づく複数のスピーカフィールド 35 をレンダリングし得る。上述されたように、信号値は、いくつかの事例では、複数のスピーカフィールドへの球面調和係数をレンダリングするために使用される行列を含み得る。この場合、オーディオ再生システム 32 は、その行列に基づいてスピーカフィールド 35 をレンダリングするためにオーディオレンダラ 34 のうちの1つを使用して、その行列でオーディオレンダラ 34 のうちのこの1つを構成し得る。

20

【0052】

[0055] いくつかの事例では、信号値は、ビットストリームがスピーカフィールド 35 への球面調和係数 $27'$ をレンダリングするために使用される行列を含むことを示すインデックスを定義する2つ以上のビットを含む。抽出デバイス 38 は、インデックスに

30

【0053】

[0056] いくつかの事例では、信号値は、スピーカフィールド 35 への球面調和係数 $27'$ をレンダリングするために使用されるレンダリングアルゴリズムを指定する。これらの事例では、オーディオレンダラ 34 の一部または全部がこれらのレンダリングアルゴリズムを実行し得る。オーディオ再生デバイス 32 は、次いで、球面調和係数 $27'$ からスピーカフィールド 35 をレンダリングするために、指定されたレンダリングアルゴリズム、たとえば、オーディオレンダラ 34 のうちの1つを利用し得る。

40

【0054】

[0057] 信号値が、スピーカフィールド 35 への球面調和係数 $27'$ をレンダリングするために使用される複数の行列のうち

50

たオーディオレンダラ 3 4 のうちの 1 つを使用して球面調和係数 2 7 ' からスピーカーフ
ィード 3 5 をレンダリングし得る。

【 0 0 5 5 】

[0058] 信号値が、スピーカーフィード 3 5 への球面調和係数 2 7 ' をレンダリングす
るために使用される複数のレンダリングアルゴリズムのうちの 1 つに関連付けられたイン
デックスを定義する 2 つ以上のビットを含むとき、オーディオレンダラ 3 4 の一部または
全部はこれらのレンダリングアルゴリズムを表し得る。したがって、オーディオ再生シス
テム 3 2 は、インデックスに関連付けられたオーディオレンダラ 3 4 のうちの 1 つを使用
して球面調和係数 2 7 ' からスピーカーフィード 3 5 をレンダリングし得る。

【 0 0 5 6 】

[0059] このオーディオレンダリング情報がビットストリーム中で指定される頻度に
応じて、抽出デバイス 3 8 は、オーディオレンダリング情報 3 9 をオーディオフレームごと
にまたは 1 回決定し得る。

【 0 0 5 7 】

[0060] このようにしてオーディオレンダリング情報 3 9 を指定することによって、本
技法は、潜在的に、マルチチャネルオーディオコンテンツ 3 5 が再現されるようにコンテ
ンツ作成者 2 2 が意図した様式に従って、マルチチャネルオーディオコンテンツ 3 5 のよ
り良い再現を生じ得る。その結果、本技法は、より没入できるサラウンド音またはマルチ
チャネルオーディオエクスペリエンスを提供し得る。

【 0 0 5 8 】

[0061] ビットストリーム中でシグナリングされる（かまたはさもなければ指定される
）ものとして説明されたが、オーディオレンダリング情報 3 9 は、ビットストリームとは
別個のメタデータとして、または言い換えれば、ビットストリームとは別個のサイド情報
として指定され得る。ビットストリーム生成デバイス 3 6 は、本開示で説明される技法を
サポートしない抽出デバイスとのビットストリーム互換性を維持するために（およびそれ
によってそれらの抽出デバイスによる成功したパースを可能にするために）、ビットスト
リーム 3 1 とは別個にこのオーディオレンダリング情報 3 9 を生成し得る。したがって、
ビットストリーム中で指定されるものとして説明されたが、本技法は、ビットストリーム
3 1 とは別個にオーディオレンダリング情報 3 9 を指定するための他の方法を可能にし得
る。

【 0 0 5 9 】

[0062] その上、ビットストリーム 3 1 中であるいはビットストリーム 3 1 とは別個の
メタデータまたはサイド情報中でシグナリングされるかまたはさもなければ指定されるも
のとして説明されたが、本技法は、ビットストリーム生成デバイス 3 6 が、ビットストリ
ーム 3 1 中でオーディオレンダリング情報 3 9 の一部分を指定し、ビットストリーム 3 1
とは別個のメタデータとしてオーディオレンダリング情報 3 9 の一部分を指定することを
可能にし得る。たとえば、ビットストリーム生成デバイス 3 6 は、ビットストリーム 3 1
で行列を識別するインデックスを指定し得、ここで、識別される行列を含む複数の行列を
指定するテーブルは、ビットストリームとは別個のメタデータとして指定され得る。オー
ディオ再生システム 3 2 は、次いで、インデックスの形態でビットストリーム 3 1 から、
およびビットストリーム 3 1 とは別個に指定されたメタデータからオーディオレンダリン
グ情報 3 9 を決定し得る。オーディオ再生システム 3 2 は、いくつかの事例では、（オー
ディオ再生システム 3 2 の製造業者または規格化団体によってたいがいホストされる）事
前構成されたかまたは構成されたサーバからテーブルおよび任意の他のメタデータをダウ
ンロードするかまたはさもなければ取り出すように構成され得る。

【 0 0 6 0 】

[0063] しかしながら、よくあることだが、コンテンツ消費者 2 4 は、（典型的にはサ
ラウンド音オーディオフォーマット団体によって）指定された幾何学的配置に従ってスピー
カーを適切に構成しない。しばしば、コンテンツ消費者 2 4 は、固定された高さに、お
よび聴取者に対して正確に指定されたロケーションにスピーカーを置かない。コンテンツ

10

20

30

40

50

消費者 24 は、これらのロケーションにスピーカーを置くことが不可能であるか、または、好適なサラウンド音エクスペリエンスを達成するためにスピーカーをそこに置くべき指定されたロケーションがあることにさえ気づいていないことがある。SHCを使用することは、SHCが2次元または3次元で音場を表すことを鑑みて、スピーカーのよりフレキシブルな配置を可能にし、これは、SHCから、たいていの任意のスピーカー幾何学的配置で構成されたスピーカーによって、音場の容認できる（または、非SHCオーディオシステムのそれと比較して、少なくともより良いサウンディング）再現が提供され得ることを意味する。

【0061】

[0064] たいていの任意のローカルスピーカー幾何学的配置へのSHCのレンダリングを可能にするために、本開示で説明される技法は、レンダラ決定ユニット40が、上記で説明された様式でオーディオレンダリング情報39を使用して標準のレンダラを選択することだけでなく、ローカルスピーカー幾何学的配置情報41に基づいてレンダラを動的に生成することをも可能にし得る。図4～図12Cに関してより詳細に説明されるように、本技法は、ローカルスピーカー幾何学的配置情報41によって指定された特定のローカルスピーカー幾何学的配置に適合されたレンダラ34を生成するための少なくとも4つの例示的な方法を提供し得る。これらの3つの方法は、モノレンダラ34と、ステレオレンダラ34と、水平マルチチャンネルレンダラ34（ここで、たとえば、「水平マルチチャンネル（horizontal multi-channel）」は、3つ以上のスピーカーを有し、それらのスピーカーのすべてが概して同じ水平面上にあるかまたはその近くにある、マルチチャンネルスピーカー構成を指す）と、3次元（3D）レンダラ34（ここで、3次元レンダラは、スピーカーの複数の水平面のためにレンダリングし得る）とを生成するための方法を含み得る。

【0062】

[0065] 動作中、オーディオ決定ユニット40は、オーディオレンダリング情報39またはローカルスピーカー幾何学的配置情報41に基づいてレンダラ34を選択し得る。しばしば、コンテンツ消費者24は、レンダラ決定ユニット40が、オーディオレンダリング情報39に基づいて（これはすべてのビットストリーム中で存在するとは限らないので、存在するとき）レンダラ34を選択し、および存在しないとき、ローカルスピーカー幾何学的配置情報41に基づいてレンダラ34を決定する（またはあらかじめ決定されている場合は選択する）という選好を指定し得る。いくつかの事例では、コンテンツ消費者24は、レンダラ決定ユニット40が、レンダラ34の選択中にオーディオレンダリング情報39を考慮することさえせずに、ローカルスピーカー幾何学的配置情報41に基づいてレンダラ34を決定する（またはあらかじめ決定されている場合は選択する）という選好を指定し得る。ただ2つの代替形態が提供されたが、レンダラ決定ユニット40がどのようにオーディオレンダリング情報39および/またはローカルスピーカー幾何学的配置41に基づいてレンダラ34を選択するかについて任意の数の選好が指定され得る。したがって、本技法は、この点において上記で論じられた2つの例示的な代替形態に限定されるべきではない。

【0063】

[0066] いずれの場合も、レンダラ決定ユニット40がローカルスピーカー幾何学的配置情報41に基づいてレンダラ34を決定すべきであると仮定すると、レンダラ決定ユニット40は、最初に、ローカルスピーカー幾何学的配置を、上記で手短かに述べられた4つのカテゴリーのうちの1つにカテゴリー分類し得る。すなわち、レンダラ決定ユニット40は、最初に、ローカルスピーカー幾何学的配置情報41がモノスピーカー幾何学的配置、ステレオスピーカー幾何学的配置、同じ水平面上に3つ以上のスピーカーを有する水平マルチチャンネルスピーカー幾何学的配置、または、3つ以上のスピーカーを有し、そのうちの2つが（しばしば何らかのしきい値高さによって分離された）異なる水平面上にある3次元マルチチャンネルスピーカー幾何学的配置に概して準拠することをローカルスピーカー幾何学的配置情報41が示すかどうかを決定し得る。このローカルスピーカー幾何学的配置情報41に基づいてローカルスピーカー幾何学的配置をカテゴリー分類すると、レン

ダラ決定ユニット40は、モノレンダラと、ステレオレンダラと、水平マルチチャンネルレンダラと、3次元マルチチャンネルレンダラとのうちの1つを生成し得る。レンダラ決定ユニット40は、次いで、オーディオ再生システム32による使用のためにこのレンダラ34を提供し得、その後、オーディオ再生システム32は、マルチチャンネルオーディオデータ35を生成するために上記で説明された様式でSHC27'をレンダリングし得る。

【0064】

[0067] このようにして、本技法は、オーディオ再生システム32が、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカークーカルスピーカ幾何学的配置を決定することと、ローカルスピーカ幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することとを可能にし得る。

10

【0065】

[0068] いくつかの例では、オーディオ再生システム32は、マルチチャンネルオーディオデータを生成するために、決定されたレンダラを使用して球面調和係数をレンダリングし得る。

【0066】

[0069] いくつかの例では、オーディオ再生システム32は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置がステレオスピーカ幾何学的配置に合致 (conform) するときステレオレンダラを決定し得る。

【0067】

[0070] いくつかの例では、オーディオ再生システム32は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置が、3つ以上のスピーカを有する水平マルチチャンネルスピーカ幾何学的配置に合致するとき水平マルチチャンネルレンダラを決定し得る。

20

【0068】

[0071] いくつかの例では、オーディオ再生システム32は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置が、2つ以上の水平面上に3つ以上のスピーカを有する3次元マルチチャンネルスピーカ幾何学的配置を合致するとき3次元マルチチャンネルレンダラを決定し得る。

【0069】

[0072] いくつかの例では、オーディオ再生システム32は、1つまたは複数のスピーカークーカルスピーカ幾何学的配置を決定するとき、ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を聴取者から受信し得る。

30

【0070】

[0073] いくつかの例では、オーディオ再生システム32は、1つまたは複数のスピーカークーカルスピーカ幾何学的配置を決定するとき、ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を、グラフィカルユーザインターフェースを介して聴取者から受信し得る。

【0071】

[0074] いくつかの例では、オーディオ再生システム32は、1つまたは複数のスピーカークーカルスピーカ幾何学的配置を決定するとき、ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を自動的に決定し得る。

40

【0072】

[0075] 以下は、上記の技法を集約するための1つの方法である。概して、SHC27などの高次アンビソニックス信号は、球面調和基底関数を使用する3次元音場の表現であり、ここで、球面調和基底関数のうちの少なくとも1つは、1よりも大きい次数を有する球面基底関数に関連付けられる。この表現は、エンドユーザスピーカ幾何学的配置に依存しないので理想的な音フォーマットを提供し得、その結果、この表現は、符号化側の事前知識なしに、コンテンツ消費者において任意の幾何学的配置にレンダリングされ得る。

50

次いで、最終スピーカ信号が、その特定のスピーカ方向を向く極パターンを概して表す球面調和係数の線形結合によって導出され得る。5.0/5.1などの共通スピーカレイアウトのための特定のHOAレンダラを設計することについて、また、不規則な2Dおよび3Dスピーカ幾何学的配置のために（通常「オンザフライで（on the fly）」と呼ばれる）リアルタイムまたはほぼリアルタイムでレンダラを生成することについて）調査が行われた。規則的な（t設計（t-design））スピーカ幾何学的配置の「ゴールデン（golden）」ケースは、擬逆元（pseudo-inverse）ベースのレンダリング行列を使用することによってよく知られていることがある。来るべきMPEG-H規格の場合、任意のスピーカ幾何学的配置を取り、当該のスピーカ幾何学的配置にとって最良のレンダリング行列を作り出すための正しい方法を使用することができるシステムが必要とされ得る。

10

【0073】

[0076] 本開示で説明される技法の様々な態様は、HOAまたはSHCレンダラ生成システム/アルゴリズムを提供する。このシステムは、モノ、ステレオ、水平、3次元など、どんなタイプのスピーカ幾何学的配置が使用中であるか、または知られている幾何学的配置/レンダラ行列としてフラグを付けられているかを検出する。

【0074】

[0077] 図4は、図3のレンダラ決定ユニット40をより詳細に示すブロック図である。図4の例に示されているように、レンダラ決定ユニット40は、レンダラ選択ユニット42と、レイアウト決定ユニット44と、レンダラ生成ユニット46とを含み得る。レンダラ選択ユニット42は、レンダリング情報39に基づいてあらかじめ定義されたものを選択するか、またはレンダリング情報39において指定されたレンダリングを選択し、この選択または指定されたレンダラをレンダラ34として出力するように構成されたユニットを表し得る。

20

【0075】

[0078] レイアウト決定ユニット44は、ローカルスピーカ幾何学的配置情報41に基づいてローカルスピーカ幾何学的配置をカテゴリー分類するように構成されたユニットを表し得る。レイアウト決定ユニット44は、ローカルスピーカ幾何学的配置を、1) モノスピーカ幾何学的配置、2) ステレオスピーカ幾何学的配置、3) 水平マルチチャンネルスピーカ幾何学的配置、および4) 3次元マルチチャンネルスピーカ幾何学的配置という、上記で説明された3つのカテゴリーのうちの1つにカテゴリー分類し得る。レイアウト決定ユニット44は、ローカルスピーカ幾何学的配置が3つのカテゴリーのうちのどれに最も合致するかを示すカテゴリー分類情報45をレンダラ生成ユニット46に受け渡し得る。

30

【0076】

[0079] レンダラ生成ユニット46は、カテゴリー分類情報45とローカルスピーカ幾何学的配置情報41とに基づいてレンダラ34を生成するように構成されたユニットを表し得る。レンダラ生成ユニット46は、モノレンダラ生成ユニット48Dと、ステレオレンダラ生成ユニット48Aと、水平レンダラ生成ユニット48Bと、3次元（3D）レンダラ生成ユニット48Cとを含み得る。モノレンダラ生成ユニット48Aは、ローカルスピーカ幾何学的配置情報41に基づいてモノレンダラを生成するように構成されたユニットを表し得る。ステレオレンダラ生成ユニット48Aは、ローカルスピーカ幾何学的配置情報41に基づいてステレオレンダラを生成するように構成されたユニットを表し得る。ステレオレンダラ生成ユニット48Aによって採用されるプロセスについては、図6の例に関して以下でより詳細に説明される。水平レンダラ生成ユニット48Bは、ローカルスピーカ幾何学的配置情報41に基づいて水平マルチチャンネルレンダラを生成するように構成されたユニットを表し得る。水平レンダラ生成ユニット48Bによって採用されるプロセスについては、図7の例に関して以下でより詳細に説明される。3Dレンダラ生成ユニット48Cは、ローカルスピーカ幾何学的配置情報41に基づいて3Dマルチチャンネルレンダラを生成するように構成されたユニットを表し得る。水平レンダラ生成ユ

40

50

ニット 48B によって採用されるプロセスについては、図 8 および図 9 の例に関して以下でより詳細に説明される。

【0077】

[0080] 図 5 は、本開示で説明される技法の様々な態様を実行する際の、図 4 の例に示されたレンダラ決定ユニット 40 の例示的な動作を示す流れ図である。図 5 の流れ図は、概して、いくつかの軽微な表記法の変更を除いて、図 4 に関して上記で説明されたレンダラ決定ユニット 40 によって実行される動作を略述している。図 5 の例では、レンダラフラグは、オーディオレンダリング情報 39 の特定の例を指す。「SHC 次数 (SHC order)」は SHC の最大次数を指す。「ステレオレンダラ (stereo renderer)」はステレオレンダラ生成ユニット 48A を指し得る。「水平レンダラ (horizontal renderer)」は水平レンダラ生成ユニット 48B を指し得る。「3D レンダラ」は 3D レンダラ生成ユニット 48C を指し得る。「レンダラ行列 (Renderer Matrix)」はレンダラ選択ユニット 42 を指し得る。

10

【0078】

[0081] 図 5 の例に示されているように、レンダラ選択ユニット 42 は、レンダリングフラグ 39' として示され得るレンダリングフラグが、ビットストリーム 31 (またはビットストリーム 31 に関連付けられた他のサイドチャンネル情報) 中に存在するかどうかを受信し、決定し得る (60)。レンダラフラグ 39' がビットストリーム 31 中に存在するとき (「YES」60)、レンダラ選択ユニット 42 は、レンダラフラグ 39' に基づいて潜在的な複数のレンダラからレンダラを選択し、選択されたレンダラをレンダラ 34

20

【0079】

[0082] レンダラフラグ 39' がビットストリーム中に存在しないとき (「NO」60)、レンダラ選択ユニット 42 はレンダラ決定ユニット 40 を呼び出し得、レンダラ決定ユニット 40 はローカルスピーカ幾何学的配置情報 41 を決定し得る。ローカルスピーカ幾何学的配置情報 41 に基づいて、レンダラ決定ユニット 40 は、モノレンダラ決定ユニット 48D、スピーカレンダラ決定ユニット 48A、水平レンダラ決定ユニット 48B または 3D レンダラ決定ユニット 48C のうちの 1 つを呼び出し得る。

【0080】

[0083] ローカルスピーカ幾何学的配置情報 41 がモノローカルスピーカ幾何学的配置を示すとき、レンダリング決定ユニット 40 はモノレンダラ決定ユニット 48D を呼び出し得、モノレンダラ決定ユニット 48D は、(SHC 次数に潜在的に基づいて) モノレンダリングを決定し、モノレンダラをレンダラ 34 として出力し得る (66、64)。ローカルスピーカ幾何学的配置情報 41 がステレオローカルスピーカ幾何学的配置を示すとき、レンダリング決定ユニット 40 はステレオレンダラ決定ユニット 48A を呼び出し得、ステレオレンダラ決定ユニット 48A は、(SHC 次数に潜在的に基づいて) ステレオレンダリングを決定し、ステレオレンダラをレンダラ 34 として出力し得る (68、64)。ローカルスピーカ幾何学的配置情報 41 が水平ローカルスピーカ幾何学的配置を示すとき、レンダリング決定ユニット 40 は水平レンダラ決定ユニット 48B を呼び出し得、水平レンダラ決定ユニット 48B は、(SHC 次数に潜在的に基づいて) 水平レンダリングを決定し、水平レンダラをレンダラ 34 として出力し得る (70、64)。ローカルスピーカ幾何学的配置情報 41 がステレオローカルスピーカ幾何学的配置を示すとき、レンダリング決定ユニット 40 は 3D レンダラ決定ユニット 48C を呼び出し得、3D レンダラ決定ユニット 48C は、(SHC 次数に潜在的に基づいて) 3D レンダリングを決定し、3D レンダラをレンダラ 34 として出力し得る (72、64)。

30

40

【0081】

[0084] このようにして、本技法は、レンダラ決定ユニット 40 が、音場を表す球面調和係数の再生のために使用される 1 つまたは複数のスピーカのローカルスピーカ幾何学的配置を決定することと、ローカルスピーカ幾何学的配置に基づいて 2 次元レンダラまたは 3 次元レンダラを決定することとを可能にし得る。

50

【 0 0 8 2 】

[0085] 図6は、図4の例に示されたステレオレンダラ生成ユニット48Aの例示的な動作を示す流れ図である。図6の例では、ステレオレンダラ生成ユニット48Aは、ローカルスピーカー幾何学的配置情報41を受信し(100)、次いで、所与のスピーカー幾何学的配置のための「スイートスポット(sweet spot)」と見なされ得るものにおける聴取者位置に対してスピーカー間の角距離を決定し得る(102)。ステレオレンダラ生成ユニット48Aは、次いで、球面調和係数のHOA/SHC次数によって限定される、最高許容次数を計算し得る(104)。ステレオレンダラ生成ユニット48Aは、次に、決定された許容次数に基づいて等しい離間したアジマス(azimuth)を生成し得る(106)。

10

【 0 0 8 3 】

[0086] ステレオレンダラ生成ユニット48Aは、次いで、2次元(2D)レンダラを形成する仮想または現実のスピーカーのロケーションにおいて球面基底関数をサンプリングし得る。ステレオレンダラ生成ユニット48Aは、次いで、この2Dレンダラの(行列数学のコンテキストで理解される)擬逆元(pseudo-inverse)を実行し得る(108)。数学的に、この2Dレンダラは以下の行列によって表され得る。

【数13】

$$\text{matrix: } \begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2)Y_0^{0*}(\theta_2, \varphi_2) & \dots & \dots \\ h_1^{(2)}(kr_1)Y_1^{1*}(\theta_1, \varphi_1) & \cdot & \dots & \dots \\ \cdot & \cdot & \dots & \dots \\ \cdot & \cdot & \dots & \dots \\ \cdot & \cdot & \dots & \dots \end{bmatrix}$$

20

【 0 0 8 4 】

この行列のサイズはV行×(n+1)²であり得、ただし、Vは仮想スピーカーの数を示し、nはSHC次数を示す。

【数14】

$$h_n^{(2)}(\cdot)$$

30

【 0 0 8 5 】

は次数nの(第二種の)球ハンケル関数である。

【数15】

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

【 0 0 8 6 】

は次数nと副次数mとの球面調和基底関数である。{ θ_r, φ_r }は球面座標に関する基準点(または観測点)である。

40

【 0 0 8 7 】

[0087] ステレオレンダラ生成ユニット48Aは、次いで、アジマスを右位置と左位置とに回転させて2つの異なる2Dレンダラを生成し(110、112)、次いで、それらを2Dレンダラ行列に合成し得る(114)。ステレオレンダラ生成ユニット48Aは、次いで、この2Dレンダラ行列を3Dレンダラ行列に変換し(116)、(図6の例では次数'として示された)許容次数と次数nとの間の差をゼロパディングし得る(120)。ステレオレンダラ生成ユニット48Aは、次いで、3Dレンダラ行列に関してエネルギー保存を実行し(122)、この3Dレンダラ行列を出力し得る(124)。

【 0 0 8 8 】

50

[0088] このようにして、本技法は、ステレオレンダラ生成ユニット48Aが、SHC次数、および左スピーカー位置と右スピーカー位置との間の角距離に基づいてステレオレンダリング行列を生成することを可能にし得る。ステレオレンダラ生成ユニット48Aは、次いで、左スピーカー位置と、次いで右スピーカー位置とに一致するようにレンダリング行列のフロント位置を回転させ、次いで、これらの左行列と右行列とを合成して最終レンダリング行列を形成し得る。

【0089】

[0089] 図7は、図4の例に示された水平レンダラ生成ユニット48Bの例示的な動作を示す流れ図である。図7の例では、水平レンダラ生成ユニット48Bは、ローカルスピーカー幾何学的配置情報41を受信し(130)、次いで、所与のスピーカー幾何学的配置のための「スイートスポット」と見なされ得るものにおける聴取者位置に対してスピーカー間の角距離を見つけ得る(132)。水平レンダラ生成ユニット48Bは、次いで、最小角距離と最大角距離とを計算し、最小角距離を最大角距離と比較し得る(134)。最小角距離が等しい(かまたは何らかの角度しきい値内でほぼ等しい)とき、水平レンダラ生成ユニット48Bは、ローカルスピーカー幾何学的配置が規則的であると決定する。最小角距離が最大角距離に等しくない(かまたは何らかの角度しきい値内でほぼ等しくない)とき、水平レンダラ生成ユニット48Bは、ローカルスピーカー幾何学的配置が不規則であると決定し得る。

【0090】

[0090] ローカルスピーカー幾何学的配置が規則的であると決定されたときについて最初に考えると、水平レンダラ生成ユニット48Bは、上記で説明するように、球面調和係数のHOA/SHC次数によって限定される、最高許容次数を計算し得る(136)。水平レンダラ生成ユニット48Bは、次に、2Dレンダラの擬逆元を生成し(138)、2Dレンダラのこの擬逆元を3Dレンダラに変換し(140)、3Dレンダラをゼロパディングし得る(142)。

【0091】

[0091] ローカルスピーカー幾何学的配置が不規則であると決定されたときについて次に考えると、水平レンダラ生成ユニット48Bは、上記で説明するように、球面調和係数のHOA/SHC次数によって限定される、最高許容次数を計算し得る(144)。水平レンダラ生成ユニット48Bは、次いで、許容次数に基づいて等しい離間したアジマスを生じて(146)、2Dレンダラを生成し得る。水平レンダラ生成ユニット48Bは、2Dレンダラの擬逆元を実行し(148)、随意のウィンドウ処理演算を実行し得る(150)。いくつかの事例では、水平レンダラ生成ユニット48Bはウィンドウ処理演算を実行しないことがある。いずれの場合も、水平レンダラ生成ユニット48Bはまた、等しいアジマスを現実のアジマスにプレースする利得をパンニングし(不規則なスピーカー幾何学的配置について、152)、パンニングされた利得での擬逆元2Dレンダラの行列乗算を実行し得る(154)。数学的に、パンニング利得行列は、ベクトルベース振幅パンニング(VBAP)を実行するサイズ $R \times V$ のVBAP行列を表し得、ただし、 V はここでも仮想スピーカーの数を表し、 R は現実のスピーカーの数を表す。VBAP行列は次のように指定され得る。

【数16】

$$\begin{bmatrix} VBAP \\ MATRIX^{-1} \\ RxV \end{bmatrix}$$

【0092】

乗算は次のように表され得る。

【数 17】

$$\begin{bmatrix} VBAP \\ MATRIX^{-1} \\ RxV \end{bmatrix} \begin{bmatrix} D^{-1} \\ Vx(n+1)^2 \end{bmatrix}$$

【0093】

水平レンダラ生成ユニット48Bは、次いで、2Dレンダラである行列乗算の出力を3Dレンダラに変換し(156)、次いで、ここでも上記で説明されたように、3Dレンダラをゼロパディングし得る(158)。

10

【0094】

[0092] 上記では、仮想スピーカーを現実のスピーカーにマッピングするために特定のタイプのパンニングを実行するものとして説明されたが、本技法は、仮想スピーカーを現実のスピーカーにマッピングするどんな方法に関しても実行され得る。その結果、行列は、 $R \times V$ のサイズを有する「仮想対現実スピーカーマッピング行列(virtual-to-real speaker mapping matrix)」として示され得る。したがって、乗算は、次のようにより一般的に表され得る。

【数 18】

$$\begin{bmatrix} Virtual_to_Real_ \\ Speaker_Mapping_Matrix^{-1} \\ RxV \end{bmatrix} \begin{bmatrix} D^{-1} \\ Vx(n+1)^2 \end{bmatrix}$$

20

【0095】

このVirtual_to_Real_Speaker_Mapping_Matrixは、ベクトルベース振幅パンニング(VBAP)を実行するための行列のうちの1つまたは複数、距離ベースの振幅パンニング(DBAP)を実行するための行列のうちの1つまたは複数、単純なパンニングを実行するための行列のうちの1つまたは複数、ニアフィールド補償(NFC)フィルタ処理を実行するための行列のうちの1つまたは複数および/または波動場合成を実行するための行列のうちの1つまたは複数を含むを含む、仮想スピーカーを現実のスピーカーにマッピングし得る任意のパンニングまたは他の行列を表し得る。

30

【0096】

[0093] 規則的な3Dレンダラが生成されるか不規則な3Dレンダラが生成されるかにかかわらず、水平レンダラ生成ユニット48Bは、規則的な3Dレンダラまたは不規則な3Dレンダラに関してエネルギー保存を実行し得る(160)。すべてではないがいくつかの例では、水平レンダラ生成ユニット48Bは、3Dレンダラの空間特性に基づいて最適化を実行し(162)、この最適化された3Dレンダラまたは最適化されていない3Dレンダラを出力し得る(164)。

【0097】

40

[0094] 水平のサブカテゴリでは、本システムは、したがって、概して、スピーカーの幾何学的配置が規則的に離間しているか不規則であるかを検出し、次いで、擬逆元またはAllRAD手法に基づいてレンダリング行列を作成し得る。AllRAD手法は、2013年3月18~21日、メラノにおけるAIA-DAGA中に提示された、「Comparison of energy-preserving and all-round Ambisonic decoders」と題する、Franz Zotterらによる論文においてより詳細に論じられている。ステレオサブカテゴリでは、HOA次数、および左スピーカー位置と右スピーカー位置との間の角距離に基づいて、規則的な水平のためのレンダラ行列を作成することによってレンダリング行列が生成される。レンダリング行列のフロント位置が、次いで、左スピーカー位置と、次いで右スピーカー位置と

50

に一致するように回転させられ、次いで、合成されて最終レンダリング行列が形成される。

【 0 0 9 8 】

[0095] 図 8 A ~ 図 8 B は、図 4 の例に示された 3 D レンダラ生成ユニット 4 8 C の例示的な動作を示す流れ図である。図 8 A の例では、3 D レンダラ生成ユニット 4 8 C が、ローカルスピーカ幾何学的配置情報 4 1 を受信し (1 7 0)、次いで、第 1 の次数の幾何学的配置と H O A / S H C 次数 n の幾何学的配置とを使用して球面調和基底関数を決定し得る (1 7 2、1 7 4)。3 D レンダラ生成ユニット 4 8 C は、次いで、第 1 の次数以下の基底関数と、1 の次数よりも大きい n 以下の球面基底関数に関連付けられた基底関数との両方のための条件数を決定し得る (1 7 6、1 7 8)。3 D レンダラ生成ユニット 4 8 C は、次いで、いくつかの例では 1 . 0 5 の値を有するしきい値を表し得るいわゆる「規則的値 (regular value)」に条件値の両方を比較する (1 8 0)。

10

【 0 0 9 9 】

[0096] 条件値の両方が規則的値を下回るとき、3 D レンダラ生成ユニット 4 8 C は、ローカルスピーカ幾何学的配置が規則的である (左から右におよび前方から後方に、ある意味で対称的であり、スピーカが等間隔に離間している) と決定し得る。条件値の両方が規則的値を下回らないかまたはそれよりも小さくないとき、3 D レンダラ生成ユニット 4 8 C は、第 1 の次数以下の球面基底関数から計算された条件値を規則的値と比較し得る (1 8 2)。第 1 の次数以下の条件値が規則的値よりも小さいとき (「 Y E S 」 1 8 2)、3 D レンダラ生成ユニット 4 8 C は、ローカルスピーカ幾何学的配置がほぼ規則的 (nearly regular) である (または、図 8 の例に示されているように、「ほぼ規則的 (near regular)」である) と決定する。第 1 の次数以下の条件値が規則的値を下回らないとき (「 N O 」 1 8 2)、3 D レンダラ生成ユニット 4 8 C は、ローカル幾何学的配置が不規則であると決定する。

20

【 0 1 0 0 】

[0097] ローカルスピーカ幾何学的配置が規則的であると決定されたとき、3 D レンダラ生成ユニット 4 8 C は、図 7 の例に関して記載された規則的な 3 D 行列決定に関して上記で説明された様式と同様の様式で 3 D レンダリング行列を決定し、ただし、3 D レンダラ生成ユニット 4 8 C は、スピーカの複数の水平面のためにこの行列を生成することが例外である (1 8 4)。ローカルスピーカ幾何学的配置がほぼ規則的であると決定されたとき、3 D レンダラ生成ユニット 4 8 C は、図 7 の例に関して記載された不規則な 2 D 行列決定に関して上記で説明された様式と同様の様式で 3 D レンダリング行列を決定し、ただし、3 D レンダラ生成ユニット 4 8 C は、スピーカの複数の水平面のためにこの行列を生成することが例外である (1 8 6)。ローカルスピーカ幾何学的配置が不規則であると決定されたとき、3 D レンダラ生成ユニット 4 8 C は、「 P E R F O R M I N G 2 D A N D / O R 3 D P A N N I N G W I T H R E S P E C T T O H E I R A R C H I C A L S E T S O F E L E M E N T S 」と題する米国仮出願第 U . S . 6 1 / 7 6 2 , 3 0 2 号において説明された様式と同様の様式で 3 D レンダリング行列を決定し、ただし、この決定のより一般的な性質に適応するための軽微な変更が例外である (本開示の技法は、この仮出願において例として提供された 2 2 . 2 スピーカ幾何学的配置には限定されないという点において、1 8 8)。

30

40

【 0 1 0 1 】

[0098] 規則的な 3 D レンダリング行列が生成されるか、ほぼ規則的な 3 D レンダリング行列が生成されるか、不規則な 3 D レンダリング行列が生成されるかにかかわらず、3 D レンダラ生成ユニット 4 8 C は、生成された行列に関してエネルギー保存を実行し (1 9 0)、それに続いて、いくつかの事例では、3 D レンダリング行列の空間特性に基づいてこの 3 D レンダリング行列を最適化する (1 9 2)。3 D レンダラ生成ユニット 4 8 C は、次いで、このレンダラをレンダラ 3 4 として出力し得る (1 9 4)。

【 0 1 0 2 】

[0099] 結果として、3次元の場合、本システムは、(擬逆元を使用して) 規則的を、

50

(第1の次数では規則的であるが、H O A次数では規則的でなく、A I I R A D方法を使用する) ほぼ規則的を、または最終的に不規則(これは上記の米国仮出願第U . S . 6 1 / 7 6 2 , 3 0 2号に基づくが、潜在的により一般的な手法として実装される)を検出し得る。3次元不規則プロセス188は、不規則な3次元リスニングのための包囲レンダラを作成するために、適切な場合、スピーカーによってカバーされるエリアのための3D - V B A P三角測量(triangulation)、トップボトムにおける高低パンニングリング、水平バンド、伸長ファクタ(stretch factor)などを生成し得る。上記のオプションのすべては、幾何学的配置間のオンザフライスイッチングが、同じ知覚されるエネルギーを有するように、エネルギー保存を使用し得る。多くの不規則またはほぼ不規則オプションは、随意的の球面調和ウィンドウ処理を使用する。

10

【0103】

[0100] 図8Bは、不規則な3Dローカルスピーカー幾何学的配置を介したオーディオコンテンツの再生のための3Dレンダラを決定する際の、3Dレンダラ決定ユニット48Cの動作を示す流れ図である。図8Bの例に示されているように、3Dレンダラ決定ユニット48Cは、上記で説明するように、球面調和係数のH O A / S H C次数によって限定される、最高許容次数を計算し得る(196)。3Dレンダラ生成ユニット48Cは、次いで、許容次数に基づいて等しい離間したアジマスを生成して(198)、3Dレンダラを生成し得る。3Dレンダラ生成ユニット48Cは、3Dレンダラの擬逆元を実行し(200)、随意的ウィンドウ処理演算を実行し得る(202)。いくつかの事例では、3Dレンダラ生成ユニット48Cはウィンドウ処理演算を実行しないことがある。

20

【0104】

[0101] 3Dレンダラ決定ユニット48Cはまた、図9に関して以下でより詳細に説明されるように、下半球処理と上半球処理とを実行し得る(204、206)。3Dレンダラ決定ユニット48Cは、下半球処理と上半球処理とを実行するとき、現実のスピーカー間の角距離を「伸長(stretch)」すべき量と、パンニングをいくつかのしきい値高さに限定するためのパンニング限界を指定し得る2Dパンニング限界と、スピーカーが同じ水平面において考慮される水平高さバンドを指定し得る水平バンド量とを示す(以下でより詳細に説明される)半球データを生成し得る。

【0105】

[0102] 3Dレンダラ決定ユニット48Cは、いくつかの事例では、下半球処理と上半球処理とのうちの1つまたは複数からの半球データに基づいて、場合によってはローカルスピーカー幾何学的配置を「伸長(stretch)」しながら、3D V B A P三角形を構築するために、3D V B A P演算を実行し得る(208)。3Dレンダラ決定ユニット48Cは、より多くの空間をカバーするために、所与の半球内で現実のスピーカー角距離を伸長し得る。3Dレンダラ決定ユニット48Cはまた、下半球と上半球とのための2Dパンニングデュプレットを識別し得(210、212)、ここで、これらのデュプレットは、それぞれ、下半球と上半球とにおいて仮想スピーカーごとに2つの現実のスピーカーを識別する。3Dレンダラ決定ユニット48Cは、次いで、等間隔に離間した幾何学的配置を生成するときに識別された各規則的な幾何学的配置位置にわたってループし、下半球および上半球仮想スピーカーの2Dパンニングデュプレット(2D panning duplet)と3D V B A P三角形とに基づいて以下の分析を実行し得る(214)。

30

40

【0106】

[0103] 3Dレンダラ決定ユニット48Cは、仮想スピーカーが、下半球と上半球とのための半球データにおいて指定された上側と下側の水平バンド値内にあるかどうかを決定し得る(216)。仮想スピーカーがこれらのバンド値内にあるとき(「YES」216)、3Dレンダラ決定ユニット48Cは、これらの仮想スピーカーの仰角を0に設定する(218)。言い換えれば、3Dレンダラ決定ユニット48Cは、いわゆる「スイートスポット」の周りの球体を二等分する中間水平面に近い下半球と上半球とにおける仮想スピーカーを識別し、この水平面上にこれらの仮想スピーカーのロケーションがあるように設定し得る。これらの仮想スピーカーロケーションを0に設定した後に、または仮想スピー

50

カーが上側と下側の水平バンド値内にないとき(「NO」216)、3Dレンダラ決定ユニット48Cは、中間水平面に沿って仮想スピーカーを現実のスピーカーにマッピングするために使用される3Dレンダラの水平面部分を生成するために、3D VBA Pパンニング(または仮想スピーカーを現実のスピーカーにマッピングする任意の他の形態または方法)(220)を実行し得る。

【0107】

[0104] 3Dレンダラ決定ユニット48Cは、仮想スピーカーの各規則的な幾何学的配置位置にわたってループするとき、下半球仮想スピーカーが、下半球データにおいて指定された下半球仰角限界を下回るかどうかを決定するために、下半球においてそれらの仮想スピーカーを評価し得る(222)。3Dレンダラ決定ユニット48Cは、上半球仮想スピーカーが、上半球データにおいて指定された上半球仰角限界を上回るかどうかを決定するために、これらの上半球仮想スピーカーに関して同様の評価を実行し得る(224)。下半球仮想スピーカーの場合は下回るか、または上半球仮想スピーカーの場合は上回るとき(「YES」226、228)、3Dレンダラ決定ユニット48Cは、それぞれ、識別された下側デュプレットと上側デュプレットとでパンニングを実行して(230、232)、仮想スピーカーの仰角をクリッピングするパンニングリングと呼ばれ得るものを効果的に作成し得、所与の半球の水平バンドの上側に現実のスピーカー間でそれをパンニングする。

【0108】

[0105] 3Dレンダラ決定ユニット(3D renderer determination unit)48Cは、次いで、3D VBA Pパンニング行列を下側デュプレットパンニング行列および上側デュプレットパンニング行列と合成し(234)、合成されたパンニング行列で行列マルチプル3Dレンダラへの行列乗算を実行し得る(236)。3Dレンダラディタメーションユニット(3D renderer determination unit)48Cは、次いで、(図6の例では次数'として示された)許容次数と次数nとの間の差をゼロパディングし(238)、不規則な3Dレンダラを出力し得る。

【0109】

[0106] このようにして、本技法は、レンダラ決定ユニット40が、球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、許容次数が、レンダリングされることを必要とされる球面調和係数のものを識別する、決定された許容次数に基づいてレンダラを決定することとを可能にし得る。

【0110】

[0107] いくつかの例では、レンダラ決定ユニット40許容次数は、球面調和係数の再生のために使用されるスピーカーの決定されたローカルスピーカー幾何学的配置を鑑みてレンダリングされることに必要とされる球面調和係数のものを識別する。

【0111】

[0108] いくつかの例では、レンダラ決定ユニット40は、レンダラを決定するとき、レンダラが、決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた球面調和係数のもののみをレンダリングするように、レンダラを決定し得る。

【0112】

[0109] いくつかの例では、レンダラ決定ユニット40はし得る、許容次数は、球面調和係数が関連付けられた球面基底関数の最大次数Nよりも小さい。

【0113】

[0110] いくつかの例では、レンダラ決定ユニット40は、マルチチャンネルオーディオデータを生成するために、決定されたレンダラを使用して球面調和係数をレンダリングし得る。

【0114】

[0111] いくつかの例では、レンダラ決定ユニット40は、球面調和係数の再生のために使用される1つまたは複数のスピーカーのローカルスピーカー幾何学的配置を決定し得

10

20

30

40

50

る。レンダラを決定するとき、レンダラ決定ユニット40は、決定された許容次数とローカルスピーカ幾何学的配置とに基づいてレンダリングを決定し得る。

【0115】

[0112] いくつかの例では、レンダラ決定ユニット40は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置がステレオスピーカ幾何学的配置に合致するときに許容次数の球面調和係数のものをレンダリングするためにステレオレンダラを決定し得る。

【0116】

[0113] いくつかの例では、レンダラ決定ユニット40は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置が、3つ以上のスピーカを有する水平マルチチャンネルスピーカ幾何学的配置に合致するときに許容次数の球面調和係数のものをレンダリングするために水平マルチチャンネルレンダラを決定し得る。

10

【0117】

[0114] いくつかの例では、レンダラ決定ユニット40は、水平マルチチャンネルレンダラを決定するとき、決定されたローカルスピーカ幾何学的配置が不規則なスピーカ幾何学的配置を示すときに許容次数の球面調和係数のものをレンダリングするために不規則な水平マルチチャンネルレンダラを決定し得る。

【0118】

[0115] いくつかの例では、レンダラ決定ユニット40は、水平マルチチャンネルレンダラを決定するとき、決定されたローカルスピーカ幾何学的配置が規則的なスピーカ幾何学的配置を示すときに許容次数の球面調和係数のものをレンダリングするために規則的な水平マルチチャンネルレンダラを決定し得る。

20

【0119】

[0116] いくつかの例では、レンダラ決定ユニット40は、ローカルスピーカ幾何学的配置に基づいてレンダラを決定するとき、ローカルスピーカ幾何学的配置が、2つ以上の水平面上に3つ以上のスピーカを有する3次元マルチチャンネルスピーカ幾何学的配置を合致するときに許容次数の球面調和係数のものをレンダリングするために3次元マルチチャンネルレンダラを決定し得る。

【0120】

[0117] いくつかの例では、レンダラ決定ユニット40は、3次元マルチチャンネルレンダラを決定するとき、決定されたローカルスピーカ幾何学的配置が不規則なスピーカ幾何学的配置を示すときに許容次数の球面調和係数のものをレンダリングするために不規則な3次元マルチチャンネルレンダラを決定し得る。

30

【0121】

[0118] いくつかの例では、レンダラ決定ユニット40は、3次元マルチチャンネルレンダラを決定するとき、決定されたローカルスピーカ幾何学的配置がほぼ規則的なスピーカ幾何学的配置を示すときに許容次数の球面調和係数のものをレンダリングするためにほぼ規則的な3次元マルチチャンネルレンダラを決定し得る。

【0122】

[0119] いくつかの例では、レンダラ決定ユニット40は、3次元マルチチャンネルレンダラを決定するとき、決定されたローカルスピーカ幾何学的配置が規則的なスピーカ幾何学的配置を示すときに許容次数の球面調和係数のものをレンダリングするために規則的な3次元マルチチャンネルレンダラを決定し得る。

40

【0123】

[0120] いくつかの例では、レンダラ決定ユニット40は、1つまたは複数のスピーカのローカルスピーカ幾何学的配置を決定するとき、ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を聴取者から受信し得る。

【0124】

50

【0121】 いくつかの例では、レンダラ決定ユニット40は、1つまたは複数のスピーカ-のローカルスピーカ-幾何学的配置を決定するとき、ローカルスピーカ-幾何学的配置を記述するローカルスピーカ-幾何学的配置情報を指定する入力を、グラフィカルユーザインターフェースを介して聴取者から受信し得る。

【0125】

【0122】 いくつかの例では、レンダラ決定ユニット40は、1つまたは複数のスピーカ-のローカルスピーカ-幾何学的配置を決定するとき、ローカルスピーカ-幾何学的配置を記述するローカルスピーカ-幾何学的配置情報を自動的に決定し得る。

【0126】

【0123】 図9は、不規則な3Dレンダラを決定するとき下半球処理と上半球処理とを実行する際の、図4の例に示された3Dレンダラ生成ユニット48Cの例示的な動作を示す流れ図である。図9の例に示されたプロセスに関するさらなる情報は、上記の米国仮出願第U.S. 61/762,302号において見つけられ得る。図9の例に示されたプロセスは、図8Bに関して上記で説明された下半球処理または上半球処理を表し得る。

【0127】

【0124】 初めに、3Dレンダラ決定ユニット48Cが、ローカルスピーカ-幾何学的配置情報41を受信し、第1の半球の現実のスピーカ-ロケーションを決定し得る(250、252)。3Dレンダラ決定ユニット48Cは、次いで、第1の半球を反対側の半球上に複製し、HOA次数の幾何学的配置を使用して球面調和を生成し得る(254、256)。3Dレンダラ決定ユニット48Cは、ローカルスピーカ-幾何学的配置の規則性(または均一性)を示し得る条件数を決定し得る(258)。条件数がしきい値数よりも小さいか、または現実のスピーカ-間の最大絶対値仰角差が90度に等しいとき(「YES」260)、3Dレンダラ決定ユニット48Cは、0の伸長値(stretch value)と、sign(90)の2Dパンニング限界値と、0の水平バンド値とを含む半球データを決定し得る(262)。上述されたように、伸長値は、現実のスピーカ-間の角距離を「伸長(stretch)」すべき量と、パンニングをいくつかのしきい値高さに限定するためのパンニング限界を指定し得る2Dパンニング限界と、スピーカ-が同じ水平面において考慮される水平高さバンドを指定し得る水平バンド量とを示す。

【0128】

【0125】 3Dレンダラ決定ユニット48Cはまた、最も高い/最も低い(上半球処理が実行されるのか下半球処理が実行されるのかに応じた)スピーカ-のアジマス(azimuth)の角距離を決定し得る(264)。条件数がしきい値数よりも大きいのか、または現実のスピーカ-間の最大絶対値仰角差が90度に等しくないとき(「YES」260)、3Dレンダラ決定ユニット48Cは、最大絶対値仰角差が0よりも大きいかどうか、および最大角距離がしきい値角距離よりも小さいかどうかを決定し得る(266)。最大絶対値仰角差が0よりも大きく、最大角距離がしきい値角距離よりも小さいとき(「YES」266)、3Dレンダラ決定ユニット48Cは、次いで、仰角の最大絶対値が70よりも大きいかどうかを決定し得る(268)。

【0129】

【0126】 仰角の最大絶対値が70よりも大きいとき(「YES」268)、3Dレンダラ決定ユニット48Cは、0に等しい伸長値と、仰角の絶対値の最大値の符号に等しい2Dパンニング限界と、0に等しい水平バンド値とを含む半球データを決定する(270)。仰角の最大絶対値が70よりも小さいかまたはそれに等しいとき(「NO」268)、3Dレンダラ決定ユニット48Cは、10-仰角の最大絶対値×70×10に等しい伸長値と、仰角の絶対値の最大値の符号付き形式-伸長値に等しい2Dパンニング限界と、仰角の最大絶対値の符号付き形式×0.1に等しい水平バンド値とを含む半球データを決定し得る(272)。

【0130】

【0127】 最大絶対値仰角差が0よりも小さいかまたはそれに等しいか、あるいは最大角距離がしきい値角距離よりも大きいかまたはそれに等しいとき(「NO」266)、3D

10

20

30

40

50

レンダラ決定ユニット48Cは、次いで、仰角の絶対値の最小値が0に等しいかどうかを決定し得る(274)。仰角の絶対値の最小値が0に等しいとき(「YES」274)、3Dレンダラ決定ユニット48Cは、0に等しい伸長値と、0に等しい2Dパンニング限界と、0に等しい水平バンド値と、仰角が0に等しい現実のスピーカーのインデックスを識別する制限半球値とを含む半球データを決定し得る(276)。仰角の絶対値の最小値が0に等しくないとき(「NO」274)、3Dレンダラ決定ユニット48Cは、最低仰角スピーカーのインデックスに等しい制限半球値を決定し得る(278)。3Dレンダラ決定ユニット48Cは、次いで、仰角の最大絶対値が70よりも大きいかどうかを決定し得る(280)。

【0131】

[0128] 仰角の最大絶対値が70よりも大きいとき(「YES」280)、3Dレンダラ決定ユニット48Cは、0に等しい伸長値と、仰角の絶対値の最大値の符号付き形式に等しい2Dパンニング限界と、0に等しい水平バンド値とを含む半球データを決定し得る。仰角の最大絶対値が70よりも小さいかまたはそれに等しいとき(「NO」280)、3Dレンダラ決定ユニット48Cは、10 - 仰角の最大絶対値 × 70 × 10に等しい伸長値と、仰角の絶対値の最大値の符号付き形式 - 伸長値に等しい2Dパンニング限界と、仰角の最大絶対値の符号付き形式 × 0.1に等しい水平バンド値とを含む半球データを決定し得る。

【0132】

[0129] 図10は、本開示に記載された技法に従ってどのようにステレオレンダラが生成され得るかを示すユニット空間におけるグラフ299を示す図である。図10の例に示されているように、仮想スピーカー300A~300Hは、(いわゆる「スイートスポット」を中心とする)ユニット球体を二等分する水平面の円周の周りに均一な幾何学的配置で配置される。物理スピーカー302Aおよび302Bは、仮想スピーカー300Aから測定されて(それぞれ)30度および-30度の角距離に配置される。ステレオレンダラ決定ユニット48Aは、上記でより詳細に説明された様式で仮想スピーカー300Aを物理スピーカー302Aおよび302Bにマッピングするステレオレンダラ34を決定し得る。

【0133】

[0130] 図11は、本開示に記載された技法に従ってどのように不規則な水平レンダラが生成され得るかを示すユニット空間におけるグラフ304を示す図である。図11の例に示されているように、仮想スピーカー300A~300Hは、(いわゆる「スイートスポット」を中心とする)ユニット球体を二等分する水平面の円周の周りに均一な幾何学的配置で配置される。物理スピーカー302A~302D(「物理スピーカー(physical speaker)302」)は、水平面の円周の周りに不規則に配置される。水平レンダラ決定ユニット48Bは、上記でより詳細に説明された様式で仮想スピーカー300A~300H(「仮想スピーカー(virtual speaker)300」)を物理スピーカー302にマッピングする不規則な水平レンダラ34を決定し得る。

【0134】

[0131] 水平レンダラ決定ユニット48Bは、仮想スピーカー300を(最も小さい角距離を有することに関して)仮想スピーカーの各々に最も近い現実のスピーカー302のうち2つにマッピングし得る。マッピングは次の表に記載されている。

【表 1】

仮想スピーカー	現実のスピーカー
300A	302A および 302B
300B	302B および 302C
300C	302B および 302C
300D	302C および 302D
300E	302C および 302D
300F	302C および 302D
300G	302D および 302A
300H	302D および 302A

10

20

【 0 1 3 5 】

[0132] 図 1 2 A および 図 1 2 B は、本開示で説明される技法に従ってどのように不規則な 3 D レンダラが生成され得るかを示すグラフ 3 0 6 A および 3 0 6 B を示す図である。図 1 2 A の例では、グラフ 3 0 6 A は、伸長されたスピーカーロケーション 3 0 8 A ~ 3 0 8 H (「伸長されたスピーカーロケーション (stretched speaker location) 3 0 8」) を含む。3 D レンダラ決定ユニット 4 8 C は、図 9 の例に関して上記で説明された様式で伸長された現実のスピーカーロケーション 3 0 8 を有する半球データを識別し得る。グラフ 3 0 6 A はまた、伸長されたスピーカーロケーション 3 0 8 に対して現実のスピーカーロケーション 3 0 2 A ~ 3 0 2 H (「現実のスピーカーロケーション (real speaker location) 3 0 2」) を示しており、ここで、いくつかの事例では、現実のスピーカーロケーション 3 0 2 は、伸長されたスピーカーロケーション 3 0 8 と同じであり、他の事例では、現実のスピーカーロケーション 3 0 2 は、伸長されたスピーカーロケーション 3 0 8 と同じではない。

30

【 0 1 3 6 】

[0133] グラフ 3 0 6 A はまた、上側 2 D パンニングデュプレットを表す上側 2 D パンニング補間ライン 3 1 0 A と、下側 2 D パンニングデュプレットを表す下側 2 D パンニング補間ライン 3 1 0 B とを含み、それらの各々については、図 8 の例に関して上記でより詳細に説明されている。手短かに言えば、3 D レンダラ決定ユニット 4 8 C は、上側 2 D パンニングデュプレットに基づいて上側 2 D パンニング補間ライン 3 1 0 A を決定し、下側 2 D パンニングデュプレットに基づいて下側 2 D パンニング補間ライン 3 1 0 B を決定し得る。上側 2 D パンニング補間ライン 3 1 0 A は上側 2 D パンニング行列を表し得、一方、下側 2 D パンニング補間ライン 3 1 0 B は下側 2 D パンニング行列を表し得る。上記で説明されたこれらの行列は、次いで、3 D V B A P 行列および規則的な幾何学的配置レンダラと合成されて、不規則な 3 D レンダラ 3 4 が生成され得る。

40

【 0 1 3 7 】

[0134] 図 1 2 B の例では、グラフはグラフ 3 0 6 A に仮想スピーカー 3 0 0 を追加し、ここで、伸長されたスピーカーロケーション 3 0 8 への仮想スピーカー 3 0 0 のマッピングを示すラインとの不要な混同を回避するために、仮想スピーカー 3 0 0 は図 1 2 B の例では正式に示されていない。典型的には、上記で説明されたように、3 D レンダラ決定

50

ユニット48Cは、図11および図12の水平例に示されたものと同様に、仮想スピーカー300の各々を、仮想スピーカーに最も近い角距離を有する伸長されたスピーカーロケーション308のうち2つ以上にマッピングする。したがって、不規則な3Dレンダラは、図12Bの例に示された様式で、仮想スピーカーを、伸長されたスピーカーロケーションにマッピングし得る。

【0138】

[0135] したがって、本技法は、第1の例では、音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカーのローカルスピーカー幾何学的配置を決定するための手段、たとえば、レンダラ決定ユニット40と、ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定するための手段、たとえば、レンダラ決定ユニット40とを備える、オーディオ再生システム32などのデバイスを提供し得る。

10

【0139】

[0136] 第2の例では、第1の例のデバイスは、マルチチャンネルオーディオデータを生成するために、決定された2次元レンダラまたは3次元レンダラを使用して球面調和係数をレンダリングするための手段、たとえば、オーディオレンダラ34をさらに備え得る。

【0140】

[0137] 第3の例では、第1の例のデバイス、ここにおいて、ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定するための手段は、ローカルスピーカー幾何学的配置がステレオスピーカー幾何学的配置に合致するとき、2次元ステレオレンダラを決定するための手段、たとえば、ステレオレンダラ生成ユニット48Aを備え得る。

20

【0141】

[0138] 第4の例では、第1の例のデバイス、ここにおいて、ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定するための手段は、ローカルスピーカー幾何学的配置が、3つ以上のスピーカーを有する水平マルチチャンネルスピーカー幾何学的配置に合致するとき、水平2次元マルチチャンネルレンダラを決定するための手段、たとえば、水平レンダラ生成ユニット48Bを備える。

【0142】

[0139] 第5の例では、第4の例のデバイス、ここにおいて、水平2次元マルチチャンネルレンダラを決定するための手段は、図7の例に関して説明されたように、決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な水平2次元マルチチャンネルレンダラを決定するための手段を備える。

30

【0143】

[0140] 第6の例では、第4の例のデバイス、ここにおいて、水平2次元マルチチャンネルレンダラを決定するための手段は、図7の例に関して説明されたように、決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な水平2次元マルチチャンネルレンダラを決定するための手段を備える。

【0144】

[0141] 第7の例では、第1の例のデバイス、ここにおいて、ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定するための手段は、ローカルスピーカー幾何学的配置が、2つ以上の水平面上に3つ以上のスピーカーを有する3次元マルチチャンネルスピーカー幾何学的配置に合致するとき、3次元マルチチャンネルレンダラを決定するための手段、たとえば、3Dレンダラ生成ユニット48Cを備える。

40

【0145】

[0142] 第8の例では、第7の例のデバイス、ここにおいて、3次元マルチチャンネルレンダラを決定するための手段は、図8Aおよび図8Bの例に関して上記で説明されたように、決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な3次元マルチチャンネルレンダラを決定するための手段を備える。

【0146】

50

【0143】 第9の例では、第7の例のデバイス、ここにおいて、3次元マルチチャンネルレンダラを決定するための手段は、図8Aの例に関して上記で説明されたように、決定されたローカルスピーカ幾何学的配置がほぼ規則的なスピーカ幾何学的配置を示すとき、ほぼ規則的な3次元マルチチャンネルレンダラを決定するための手段を備える。

【0147】

【0144】 第10の例では、第7の例のデバイス、ここにおいて、3次元マルチチャンネルレンダラを決定するための手段は、図8Aの例に関して上記で説明されたように、決定されたローカルスピーカ幾何学的配置が規則的なスピーカ幾何学的配置を示すとき、規則的な3次元マルチチャンネルレンダラを決定するための手段を備える。

【0148】

10

【0145】 第11の例では、第1の例のデバイス、ここにおいて、レンダラを決定するための手段は、図5～図8Bの例に関して上記で説明されたように、球面調和係数が関連付けられた球面基底関数の許容次数を決定するための手段と、許容次数が、決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる球面調和係数のものを識別する、決定された許容次数に基づいてレンダラを決定するための手段とを備える。

【0149】

【0146】 第12の例では、第1の例のデバイス、ここにおいて、2次元レンダラまたは3次元レンダラを決定するための手段は、図5～図8Bの例に関して上記で説明されたように、球面調和係数が関連付けられた球面基底関数の許容次数を決定するための手段と、許容次数が、決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる球面調和係数のものを識別する、2次元レンダラまたは3次元レンダラが、決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた球面調和係数のもののみをレンダリングするように、2次元レンダラまたは3次元レンダラを決定するための手段とを備える。

20

【0150】

【0147】 第13の例では、第1の例のデバイス、ここにおいて、1つまたは複数のスピーカのローカルスピーカ幾何学的配置を決定するための手段は、ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を聴取者から受信するための手段を備える。

30

【0151】

【0148】 第14の例では、第1の例のデバイス、ここにおいて、ローカルスピーカ幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することは、ローカルスピーカ幾何学的配置がモノスピーカ幾何学的配置に合致するとき、モノレンダラを決定すること、たとえば、モノレンダラ決定ユニット48Dを備える。

【0152】

【0149】 図13A～図13Dは、本開示で説明される技法に従って形成されるビットストリーム31A～31Dを示す図である。図13Aの例では、ビットストリーム31Aは、図3の例に示されたビットストリーム31の一例を表し得る。ビットストリーム31Aは、信号値54を定義する1つまたは複数のビットを含むオーディオレンダリング情報39Aを含む。この信号値54は、以下で説明されるタイプの情報の任意の組合せを表し得る。ビットストリーム31Aはまた、オーディオコンテンツの一例を表し得るオーディオコンテンツ58を含む。

40

【0153】

【0150】 図13Bの例では、ビットストリーム31Bはビットストリーム31Aと同様であり得、ここで、信号値54は、インデックス54Aと、シグナリングされる行列の行サイズ54Bを定義する1つまたは複数のビットと、シグナリングされる行列の列サイズ54Cを定義する1つまたは複数のビットと、行列係数54Dとを備える。インデックス54Aは2～5ビットを使用して定義され得るが、行サイズ54Bと列サイズ54Cとの各々は2～16ビットを使用して定義され得る。

50

【 0 1 5 4 】

[0151] 抽出デバイス 3 8 は、インデックス 5 4 A を抽出し、行列がビットストリーム 3 1 B 中に含まれることをそのインデックスがシグナリングするかどうかを決定し得る（ここで、0 0 0 0 または 1 1 1 1 など、いくつかのインデックス値は、行列が明示的にビットストリーム 3 1 B 中で指定されていることをシグナリングし得る）。図 1 3 B の例では、ビットストリーム 3 1 B は、行列が明示的にビットストリーム 3 1 B 中で指定されていることをシグナリングするインデックス 5 4 A を含む。その結果、抽出デバイス 3 8 は、行サイズ 5 4 B と列サイズ 5 4 C とを抽出し得る。抽出デバイス 3 8 は、行サイズ 5 4 B と、列サイズ 5 4 C と、各行列係数の（図 1 3 A に示されていない）シグナリングされるかまたは暗黙的なビットサイズの関数として、行列係数を表すパースすべきビット数を計算するように構成され得る。決定されたビット数を使用して、抽出デバイス 3 8 は行列係数 5 4 D を抽出し得、オーディオ再生デバイス 2 4 は、その行列係数 5 4 D を使用して、上記で説明されたようにオーディオレンダラ 3 4 のうちの 1 つを構成し得る。ビットストリーム 3 1 B 中でオーディオレンダリング情報 3 9 B を 1 回シグナリングするものとして示されているが、オーディオレンダリング情報 3 9 B は、ビットストリーム 3 1 B 中で複数回シグナリングされるか、あるいは少なくとも部分的にまたは完全に別個のアウトオブバンドチャンネル中で（いくつかの事例では随意のデータとして）シグナリングされ得る。

10

【 0 1 5 5 】

[0152] 図 1 3 C の例では、ビットストリーム 3 1 C は、上記の図 3 の例に示されたビットストリーム 3 1 の一例を表し得る。ビットストリーム 3 1 C は、この例ではアルゴリズムインデックス 5 4 E を指定する信号値 5 4 を含むオーディオレンダリング情報 3 9 C を含む。ビットストリーム 3 1 C はオーディオコンテンツ 5 8 をも含む。アルゴリズムインデックス 5 4 E は、上述されたように、2 ~ 5 ビットを使用して定義され得、ここで、このアルゴリズムインデックス 5 4 E は、オーディオコンテンツ 5 8 をレンダリングするときに使用されるべきレンダリングアルゴリズムを識別し得る。

20

【 0 1 5 6 】

[0153] 抽出デバイス 3 8 は、アルゴリズムインデックスを抽出し、行列がビットストリーム 3 1 C 中に含まれることをアルゴリズムインデックス 5 4 E がシグナリングするかどうかを決定し得る（ここで、0 0 0 0 または 1 1 1 1 など、いくつかのインデックス値は、行列が明示的にビットストリーム 3 1 C 中で指定されていることをシグナリングし得る）。図 8 C の例では、ビットストリーム 3 1 C は、行列が明示的にビットストリーム 3 1 C 中で指定されていないことをシグナリングするアルゴリズムインデックス 5 4 E を含む。その結果、抽出デバイス 3 8 は、アルゴリズムインデックス 5 4 E をオーディオ再生デバイスに転送し、オーディオ再生デバイスは、（利用可能な場合は）対応する 1 つ、（図 3 および図 4 の例ではレンダラ 3 4 として示されている）レンダリングアルゴリズムを選択する。図 1 3 C の例では、ビットストリーム 3 1 C 中でオーディオレンダリング情報 3 9 C を 1 回シグナリングするものとして示されているが、オーディオレンダリング情報 3 9 C は、ビットストリーム 3 1 C 中で複数回シグナリングされるか、あるいは少なくとも部分的にまたは完全に別個のアウトオブバンドチャンネル中で（いくつかの事例では随意のデータとして）シグナリングされ得る。

30

40

【 0 1 5 7 】

[0154] 図 1 3 D の例では、ビットストリーム 3 1 C は、上記の図 4、図 5 および図 8 に示されたビットストリーム 3 1 の一例を表し得る。ビットストリーム 3 1 D は、この例では行列インデックス 5 4 F を指定する信号値 5 4 を含むオーディオレンダリング情報 3 9 D を含む。ビットストリーム 3 1 D はオーディオコンテンツ 5 8 をも含む。行列インデックス 5 4 F は、上述されたように、2 ~ 5 ビットを使用して定義され得、ここで、この行列インデックス 5 4 F は、オーディオコンテンツ 5 8 をレンダリングするときに使用されるべきレンダリングアルゴリズムを識別し得る。

【 0 1 5 8 】

50

[0155] 抽出デバイス 38 は、行列インデックス 50F を抽出し、行列がビットストリーム 31D 中に含まれることを行列インデックス 54F がシグナリングするかどうかを決定し得る（ここで、0000 または 1111 など、いくつかのインデックス値は、行列が明示的にビットストリーム 31C 中で指定されていることをシグナリングし得る）。図 13D の例では、ビットストリーム 31D は、行列が明示的にビットストリーム 31D 中で指定されていないことをシグナリングする行列インデックス 54F を含む。その結果、抽出デバイス 38 は、行列インデックス 54F をオーディオ再生デバイスに転送し、オーディオ再生デバイスは、（利用可能な場合は）対応する 1 つ、レンダラ 34 を選択する。図 13D の例では、ビットストリーム 31D 中でオーディオレンダリング情報 39D を 1 回シグナリングするものとして示されているが、オーディオレンダリング情報 39D は、ビットストリーム 31D 中で複数回シグナリングされるか、あるいは少なくとも部分的にまたは完全に別個のアウトオブバンドチャンネル中で（いくつかの事例では随意のデータとして）シグナリングされ得る。

10

【0159】

[0156] 図 14A および図 14B は、本開示で説明される技法の様々な態様を実行し得る 3D レンダラ決定ユニット 48C の別の例である。すなわち、3D レンダラ決定ユニット 48C は、仮想スピーカーが、球体幾何学的配置を二等分する水平面よりも下側の球体幾何学的配置において配置されたとき、仮想スピーカーをその水平面上のロケーションに投射することと、再現される音場が、仮想スピーカーの投射されたロケーションから発生するように思われる少なくとも 1 つの音を含むように、音場を再現する第 1 の複数のラウドスピーカーチャンネル信号を生成するとき、音場を記述する要素の階層セット上で 2 次元パンニングを実行することと、を行うように構成されたユニットを表し得る。

20

【0160】

[0157] 図 14A の例では、3D レンダラ決定ユニット 48C は、SCH27' を受信し、仮想ラウドスピーカー t 設計レンダリングを実行するように構成されたユニットを表し得る仮想スピーカーレンダラ 350 を呼び出し得る。仮想スピーカーレンダラ 350 は、SCH27' をレンダリングし、所与の数（たとえば、22 個または 32 個）の仮想スピーカーのためのラウドスピーカーチャンネル信号を生成し得る。

【0161】

[0158] 3D レンダラ決定ユニット 48C はさらに、球面重み付けユニット 352 と、上半球 3D パンニングユニット 354 と、イヤレベル 2D パンニングユニット 356 と、下半球 2D パンニングユニット 358 とを含む。球面重み付けユニット 352 は、いくつかのチャンネルを重み付けするように構成されたユニットを表し得る。上半球 3D パンニングユニット 354 は、球状に重み付けされた仮想ラウドスピーカーチャンネル信号上で、様々な上半球の物理スピーカー、または、言い換えれば、現実のスピーカーの間でこれらの信号をパンニングするように 3D パンニングを実行するように構成されたユニットを表す。イヤレベル半球 2D パンニングユニット 356 は、球状に重み付けされた仮想ラウドスピーカーチャンネル信号上で、様々なイヤレベルの物理スピーカー、または、言い換えれば、現実のスピーカーの間でこれらの信号をパンニングするように 2D パンニングを実行するように構成されたユニットを表す。下半球 2D パンニングユニット 358 は、球状に重み付けされた仮想ラウドスピーカーチャンネル信号上で、様々な下半球の物理スピーカー、または、言い換えれば、現実のスピーカーの間でこれらの信号をパンニングするように 2D パンニングを実行するように構成されたユニットを表す。

30

40

【0162】

[0159] 図 14B の例では、3D レンダリング決定ユニット 48C' は、3D レンダリング決定ユニット 48C' が球面重み付けを実行しないことがあるかまたはさもなければ球面重み付けユニット 352 を含まないことがあることを除いて、図 14B に示されたものと同様であり得る。

【0163】

[0160] いずれの場合も、典型的には、ラウドスピーカーフィールドは、各ラウドスピー

50

カーが球面波を生成すると仮定することによって計算される。そのようなシナリオでは、1番目のラウドスピーカーによる、ある位置 r 、 θ 、 φ における（周波数の関数としての）音圧は、

【数 19】

$$P_l(\omega, r, \theta, \varphi) = g_l(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi ik) h_n^{(2)}(kr_l) Y_n^{m*}(\theta_l, \varphi_l) Y_n^m(\theta, \varphi)$$

【0164】

によって与えられ、

ただし、 $\{r_l, \theta_l, \varphi_l\}$ は1番目のラウドスピーカーの位置を表し、 $g_l(\omega)$ は、（周波数領域における）1番目のスピーカーのラウドスピーカーフィールドである。すべての5つのスピーカーによる全音圧 P_t は、したがって、

【数 20】

$$P_t(\omega, r, \theta, \varphi) = \sum_{l=1}^5 g_l(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi ik) h_n^{(2)}(kr_l) Y_n^{m*}(\theta_l, \varphi_l) Y_n^m(\theta, \varphi)$$

【0165】

によって与えられる。

【0166】

[0161] 我々はまた、5つのSHCに関する全圧力が次の式によって与えられることを知っている。

【数 21】

$$P_t(\omega, r, \theta, \varphi) = 4\pi \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta, \varphi)$$

【0167】

[0162] 上記の2つの式を同等とすることにより、次のようにSHCに関してラウドスピーカーフィールドを表すための変換行列を使用することが可能になる。

【数 22】

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ A_2^2(\omega) \\ A_2^{-2}(\omega) \end{bmatrix} = -ik \begin{bmatrix} h_0^{(2)}(kr_1) Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2) Y_0^{0*}(\theta_2, \varphi_2) & \dots & \dots \\ h_1^{(2)}(kr_1) Y_1^{1*}(\theta_1, \varphi_1) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ g_4(\omega) \\ g_5(\omega) \end{bmatrix}$$

【0168】

[0163] この式は、5つのラウドスピーカーフィールドと、選定されたSHCとの間に直接的な関係があることを示している。変換行列は、たとえば、サブセット（たとえば、基本セット）中でどのSHCが使用されたか、およびSH基底関数のどの定義が使用されるかに応じて変化し得る。同様に、選択された基本セットから異なるチャンネルフォーマット（たとえば、7.1、2.2）に変換するための変換行列が構築され得る

[0164] 上記の式中的変換行列によってスピーカーフィールドからSHCへの変換が可能

10

20

30

40

50

になるが、S H Cで開始し、我々が5つのチャネルフィールドを作り出すことができ、次いで、デコーダにおいて、(高度なくすなわち、非レガシー)レンダラが存在するとき)我々が場合によってはS H Cに変換し戻すことができるように、我々は、行列が可逆であることを好むことがある。

【0169】

[0165] 行列の可逆性を保証するために上記のフレームワークを操作する様々な方法が活用され得る。これらは、限定はしないが、ラウドスピーカーの位置を変化させること(たとえば、5.1システムの5つのラウドスピーカーのうちの1つまたは複数の位置を、依然としてそれらがITU-R BS.775-1規格によって指定された角度トレランスに従うように調整すること、T設計に従うものなど、トランスデューサの規則的な離間が、典型的には正常に作動する)、正規化技法(たとえば、周波数依存正規化)、ならびにフルランクおよび明確な固有値を保証するようにしばしば動作する様々な他の行列操作技法を含む。最後に、すべての操作の後に、修正された行列が実際に正しいおよび/または許容できるラウドスピーカーフィールドを再現することを保証するために、5.1レンディションを聴覚心理的にテストすることが望ましいことがある。可逆性が保存される限り、S H Cへの正しい復号を保証する逆問題は問題点でない。

【0170】

[0166] (デコーダにおけるスピーカー幾何学的配置を指すことがある)いくつかのローカルスピーカー幾何学的配置では、可逆性を保証するために上記のフレームワークを操作するための上記で概説された方法は、望ましいとは言えない音像を生じることがある。すなわち、音再現は、キャプチャされているオーディオと比較されたとき、常に音の正しい定位を生じるとは限らない。この望ましいとは言えない音像を補正するために、本技法は、「仮想スピーカー(virtual speaker)」と呼ばれ得る概念を導入するようさらに拡張され得る。1つまたは複数のラウドスピーカーが、上述のITU-R BS.775-1などの規格によって指定された、いくつかの角度トレランスを有する特定のまたは定義された空間領域中で再配置または配置されることを要するのではなく、上記のフレームワークは、ベクトルベース振幅パンニング(VBAP: vector base amplitude panning)、距離ベースの振幅パンニング、または他の形態のパンニングなど、何らかの形態のパンニングを含むように修正され得る。説明のためにVBAPに焦点を当てると、VBAPは、「仮想スピーカー」として特徴づけられ得るものを効果的に導入し得る。VBAPは、概して、1つまたは複数のラウドスピーカーが、仮想スピーカーをサポートする1つまたは複数のラウドスピーカーのロケーションおよび/または角度のうちの少なくとも1つと異なるロケーションおよび角度のうちの1つまたは複数において仮想スピーカーから発生するよう思われる音を効果的に出力するように、これらの1つまたは複数のラウドスピーカーへのフィールドを修正し得る。

【0171】

[0167] 例示のために、S H Cに関してラウドスピーカーフィールドを決定するための上記の式は、次のように修正され得る。

【数23】

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX \\ M \times N \end{bmatrix} \begin{bmatrix} D \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix}$$

【0172】

[0168] 上記の式において、VBAP行列はM行×N列のサイズであり、ただし、Mはスピーカーの数を示し(上記の式では5に等しくなるはずであり)、Nは仮想スピーカー

10

20

30

40

50

の数を示す。VBAP行列は、聴取者の定義されたロケーションからスピーカーの位置の各々へのベクトルと、聴取者の定義されたロケーションから仮想スピーカーの位置の各々へのベクトルとの関数として計算され得る。上記の式中のD行列はN行×(order+1)²列のサイズであり得、ただし、orderはSH関数の次数を指し得る。D行列は次の行列を表し得る。

【数24】

$$\text{matrix:} \begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2)Y_0^{0*}(\theta_2, \varphi_2) & \dots & \dots \\ h_1^{(2)}(kr_1)Y_1^{1*}(\theta_1, \varphi_1) & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

10

【0173】

[0169] 事実上、VBAP行列は、スピーカーのロケーションと仮想スピーカーの位置とを考慮する「利得調整 (gain adjustment)」と呼ばれ得るものを提供するM×N行列である。このようにしてパンニングを導入することにより、ローカルスピーカー幾何学的配置によって再現されたとき、より良質の像を生じるマルチチャンネルオーディオのより良い再現がもたらされ得る。その上、この式にVBAPを組み込むことによって、本技法は、様々な規格において指定されているスピーカー幾何学的配置とは整合しない劣悪なスピーカー幾何学的配置を克服し得る。

20

【0174】

[0170] 実際には、この式は、反転させられ、SHCを、以下で幾何学的配置Bと呼ばれることがあるラウドスピーカーの特定の幾何学的配置または構成のためのマルチチャンネルフィールドに変換し戻すために使用され得る。すなわち、この式は、g行列について解くために反転させられ得る。反転させられた式は次のようになり得る。

【数25】

$$\begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX^{-1} \\ M \times N \end{bmatrix} \begin{bmatrix} D^{-1} \\ [N \times (Order + 1)^2] \end{bmatrix} \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix}$$

30

【0175】

[0171] g行列は、この例では、5.1スピーカー構成における5つのラウドスピーカーの各々についてのスピーカー利得を表し得る。この構成において使用される仮想スピーカーロケーションは、5.1マルチチャンネルフォーマット仕様または規格において定義されているロケーションに対応し得る。これらの仮想スピーカーの各々をサポートし得るラウドスピーカーのロケーションは、任意の数の知られているオーディオ定位技法を使用して決定され得、それらの多くは、(オーディオ/ビデオ受信機(A/V受信機)、テレビジョン、ゲーミングシステム、デジタルビデオディスクシステム、または他のタイプのヘッドエンドシステムなどの)ヘッドエンドユニットに対して各ラウドスピーカーのロケーションを決定するために特定の周波数を有するトーンを再生することを伴う。代替的に、ヘッドエンドユニットのユーザが、ラウドスピーカーの各々のロケーションを手動で指定し得る。いずれの場合も、これらの知られているロケーションと考えられる角度とを鑑みて、ヘッドエンドユニットは、利得について解き、VBAPを介して仮想ラウドスピーカーの理想的な構成を仮定し得る。

40

50

【0176】

[0172] この点において、本技法は、デバイスまたは装置が、第1の複数の仮想ラウドスピーカーチャンネル信号を発生するために、第1の複数のラウドスピーカーチャンネル信号上でベクトルベース振幅パンニングまたは他の形態のパンニングを実行することを可能にし得る。これらの仮想ラウドスピーカーチャンネル信号は、ラウドスピーカーが、仮想ラウドスピーカーから発生するように思われる音を発生することを可能にする、これらのラウドスピーカーに提供される信号を表し得る。その結果、第1の複数のラウドスピーカーチャンネル信号上で第1の変換を実行するとき、本技法は、デバイスまたは装置が、音場を記述する要素の階層セットを発生するために、第1の複数の仮想ラウドスピーカーチャンネル信号上で第1の変換を実行することを可能にし得る。

10

【0177】

[0173] その上、本技法は、装置が、第2の複数のラウドスピーカーチャンネル信号を発生するために要素の階層セット上で第2の変換を実行することを可能にし得、ここで、第2の複数のラウドスピーカーチャンネル信号の各々は、対応する異なる空間領域に関連付けられ、ここで、第2の複数のラウドスピーカーチャンネル信号は第2の複数の仮想ラウドスピーカーチャンネルを備え、およびここで、第2の複数の仮想ラウドスピーカーチャンネル信号は、対応する異なる空間領域に関連付けられる。本技法は、いくつかの事例では、デバイスが、第2の複数のラウドスピーカーチャンネル信号を発生するために、第2の複数の仮想ラウドスピーカーチャンネル信号上でベクトルベース振幅パンニングを実行することを可能にし得る。

20

【0178】

[0174] 上記の変換行列は「モード整合 (mode matching)」基準から導出されたが、音圧整合、エネルギー整合など、他の基準からも代替の変換行列が導出され得る。基本セット (たとえば、SHCサブセット) と従来のマルチチャンネルオーディオとの間の変換を可能にする行列が導出され得ることと、また、(マルチチャンネルオーディオの忠実度を低減しない) 操作後に、可逆でもあるわずかに修正された行列も作成され得ることとで、十分である。

【0179】

[0175] いくつかの事例では、パンニングが3次元空間において実行されるという意味で「3Dパンニング (3D panning)」と呼ばれることもある、上記で説明されたパンニングを実行するときに、上記で説明された3Dパンニングは、アーティファクトを導入するか、またはさもなければスピーカーフィールドのより低品質な再生を生じることがある。例として説明するために、上記で説明された3Dパンニングは、図15Aおよび図15Bに示されている2.2スピーカー幾何学的配置に関して採用されることがある。

30

【0180】

[0176] 図15Aおよび図15Bは、同じ2.2スピーカー幾何学的配置を示し、ここで、図15Aに示されたグラフ中の黒点は、(低周波スピーカーを除く)すべてのラウドスピーカー、2.2個のスピーカーのロケーションを示し、図15Bは、これらの同じスピーカーのロケーションを示すが、これらのスピーカーの(影つき半球の背後に位置するスピーカーをブロックする)半球位置性質をさらに定義する。いずれの場合も、実際のスピーカーのいくつか(その数は上記ではMとして示されている)は、実際にその半球において聴取者の耳の背後にあり、聴取者の頭は、図15Aおよび図15Bのグラフ中の(0, 0, 0)の(x, y, z)点の周りの半球中のどこかに配置される。その結果、聴取者の頭の背後にあるスピーカーを仮想化するように3Dパンニングを実行することを試みることは、特に、SHCを生成するとき通常仮定され、仮想スピーカーの位置を伴う図12Bの例に示されている、完全な球体の周りに均一に配置された仮想スピーカーを有する、3.2スピーカー球体(半球ではない)幾何学的配置を仮想化することを試みる時、困難であり得る。

40

【0181】

[0177] 本開示で説明される技法によれば、図14Aの例に示された3Dレンダラ決定

50

ユニット48Cは、仮想スピーカーが、球体幾何学的配置を二等分する水平面よりも下側の球体幾何学的配置において配置されたとき、仮想スピーカーをその水平面上のロケーションに投射することと、再現される音場が、仮想スピーカーの投射されたロケーションから発生するように思われる少なくとも1つの音を含むように、音場を再現する第1の複数のラウドスピーカーチャンネル信号を生成するとき、音場を記述する要素の階層セット上で2次元パンニングを実行することと、を行うようにユニットを表し得る。

【0182】

[0178] 水平面は、いくつかの事例では、球体幾何学的配置を2つの等しい部分に二等分し得る。図16Aは、本開示で説明される技法による、仮想スピーカーが上方にその上に投射される、水平面402によって二等分された球体400を示している。仮想スピーカー300A~300C、ここで、下側の仮想スピーカー300A~300Cは、図14Aおよび図14Bの例に関して上記で概説された方法で2次元プランニングを実行するより前に、上記で具陳された様式で水平面402上に投射される。球体400を等しく二等分する水平面402上に投射されるものとして説明されているが、本技法は、仮想スピーカーを球体400内の任意の水平面(たとえば仰角)に投射し得る。

10

【0183】

[0179] 図16Bは、本開示で説明される技法による、仮想スピーカーが下方にその上に投射される、水平面402によって二等分された球体400を示している。図16Bのこの例では、3Dレンダラ決定ユニット48Cが、仮想スピーカー300A~300Cを水平面402に下に投射し得る。球体400を等しく二等分する水平面402上に投射されるものとして説明されているが、本技法は、仮想スピーカーを球体400内の任意の水平面(たとえば仰角)に投射し得る。

20

【0184】

[0180] このようにして、本技法は、3Dレンダラ決定ユニット48Cが、幾何学的配置で配置された複数の仮想スピーカーのうちの1つの位置に対して複数の物理スピーカーのうちの1つの位置を決定することと、決定された位置に基づいて幾何学的配置内の複数の仮想スピーカーのうちの1つの位置を調整することとを行うことを可能にし得る。

【0185】

[0181] 3Dレンダラ決定ユニット48Cは、第1の複数のラウドスピーカーチャンネル信号を生成するとき、要素の階層セット上で2次元パンニングに加えて第1の変換を実行するようにさらに構成され得、ここにおいて、第1の複数のラウドスピーカーチャンネル信号の各々は、対応する異なる空間領域に関連付けられる。この第1の変換は、上記の式において D^{-1} として反映され得る。

30

【0186】

[0182] 3Dレンダラ決定ユニット48Cは、要素の階層セット上で2次元パンニングを実行するとき、第1の複数のラウドスピーカーチャンネル信号を生成するときに要素の階層セット上で2次元ベクトルベース振幅パンニング(two dimensional vector base amplitude panning)を実行するようにさらに構成され得る。

【0187】

[0183] いくつかの事例では、第1の複数のラウドスピーカーチャンネル信号の各々は、対応する異なる定義された空間領域に関連付けられる。さらに、これらの異なる定義された空間領域は、オーディオフォーマット仕様およびオーディオフォーマット規格のうちの1つまたは複数において定義される。

40

【0188】

[0184] 3Dレンダラ決定ユニット48Cは、同じくまたは代替的に、仮想スピーカーが、球体幾何学的配置においてイヤレベルでまたはその近くで水平面の近くに球体幾何学的配置で配置されたとき、再現される音場が、仮想スピーカーのロケーションから発生するように思われる少なくとも1つの音を含むように、音場を再現する第1の複数のラウドスピーカーチャンネル信号を生成するときに、音場を記述する要素の階層セット上で2次元パンニングを実行するように構成され得る。

50

【 0 1 8 9 】

[0185] このコンテキストでは、3Dレンダラ決定ユニット48Cは、第1の複数のラウドスピーカチャネル信号を生成するとき、要素の階層セット上で2次元パンニングに加えて(同じく上述の D^{-1} 変換を指し得る)第1の変換を実行するようにさらに構成され得、ここで、第1の複数のラウドスピーカチャネル信号の各々は、対応する異なる空間領域に関連付けられる。

【 0 1 9 0 】

[0186] その上、3Dレンダラ決定ユニット48Cは、要素の階層セット上で2次元パンニングを実行するとき、第1の複数のラウドスピーカチャネル信号を生成するとき要素の階層セット上で2次元ベクトルベース振幅パンニングを実行するようにさらに構成され得る。

10

【 0 1 9 1 】

[0187] いくつかの事例では、第1の複数のラウドスピーカチャネル信号の各々は、対応する異なる定義された空間領域に関連付けられる。さらに、これらの異なる定義された空間領域は、オーディオフォーマット仕様およびオーディオフォーマット規格のうちの1つまたは複数において定義され得る。

【 0 1 9 2 】

[0188] 本開示で説明される技法の他の態様のいずれかの代替としてまたはそれと併せて、デバイス10の1つまたは複数のプロセッサは、仮想スピーカが、球体幾何学的配置を二等分する水平面上の上側に球体幾何学的配置で配置されたとき、音場が、仮想スピーカのロケーションから発生するように思われる少なくとも1つの音を含むように、音場を記述する第1の複数のラウドスピーカチャネル信号を生成するとき要素の階層セット上で3次元パンニングを実行するようにさらに構成され得る。

20

【 0 1 9 3 】

[0189] この場合も、このコンテキストでは、3Dレンダラ決定ユニット48Cは、第1の複数のラウドスピーカチャネル信号を生成するとき、要素の階層セット上で3次元パンニングに加えて第1の変換を実行するようにさらに構成され得、ここにおいて、第1の複数のラウドスピーカチャネル信号の各々は、対応する異なる空間領域に関連付けられる。

【 0 1 9 4 】

[0190] その上、3Dレンダラ決定ユニット48Cは、要素の階層セット、第1の複数のラウドスピーカチャネル信号上で3次元パンニングを実行するとき、第1の複数のラウドスピーカチャネル信号を生成するとき要素の階層セット上で3次元ベクトルベース振幅パンニングを実行するようにさらに構成され得る。いくつかの事例では、第1の複数のラウドスピーカチャネル信号の各々は、対応する異なる定義された空間領域に関連付けられる。さらに、これらの異なる定義された空間領域は、オーディオフォーマット仕様およびオーディオフォーマット規格のうちの1つまたは複数において定義され得る。

30

【 0 1 9 5 】

[0191] 本開示で説明される技法の他の態様のいずれかの代替としてまたはそれと併せて、3Dレンダラ決定ユニット48Cは、要素の階層セットからの複数のラウドスピーカチャネル信号の生成において3次元パンニングと2次元パンニングの両方を実行するとき、要素の階層セットの各々の次数に基づいて要素の階層セットに対して重み付けを実行するようにさらに構成され得る。

40

【 0 1 9 6 】

[0192] 3Dレンダラ決定ユニット48Cは、重み付けを実行するとき、要素の階層セットの各々の次数に基づいて要素の階層セットに対してウィンドウ関数を実行するようにさらに構成され得る。このウィンドウ処理関数は図17の例に示され得、ただし、y軸はデシベルを反映し、x軸はSHCの次数を示す。その上、デバイス10の1つまたは複数のプロセッサは、重み付けを実行するとき、要素の階層セットの各々の次数に基づいて要素の階層セットに対して、一例として、カイザーベッスルウィンドウ関数(Kaiser Bessl

50

e window function) を実行するようにさらに構成され得る。

【0197】

[0193] これらの1つまたは複数のプロセッサは、各々、これらの1つまたは複数のプロセッサにあるとされる様々な機能を実行するための手段を表し得る。他の手段は、単独であるいは本開示で説明される技法の組合せで様々な態様を実行し得るソフトウェアを実行することに専用のまたはそれが可能な、専用特定用途向けハードウェア、フィールドプログラマブルゲートアレイ、特定用途向け集積回路または任意の他の形態のハードウェアを含み得る。

【0198】

[0194] 本技法によって識別され、潜在的に解決される問題は、以下のように要約され得る。高次アンビソニックス/球面調和係数サラウンド音素材の忠実に再生にとって、ラウドスピーカーの配置は重要であり得る。理想的には、等距離のラウドスピーカーの3次元球体が望まれ得る。実世界では、現在のラウドスピーカーセットアップは、典型的には、1) 等しく分散されず、2) 聴取者の周りとは上側の上半球にのみ存在し、下側の下半球には存在せず、3) レガシーサポート(たとえば、5.1スピーカーセットアップ)のために、通常は耳の高さにラウドスピーカーのリングを有する。この問題に対処し得る1つの戦略は、(以下で、「t設計(t-design)」と呼ばれる)理想的なラウドスピーカーレイアウトを仮想的に作成すること、および、これらの仮想ラウドスピーカーを、3次元ベクトルベース振幅パンニング(3D-VBAP)方法を介して現実の(非理想的に配置された)ラウドスピーカー上に投射することである。たとえそうでも、下半球からの仮想ラウドスピーカーの投射により、再生の品質を劣化させる強度の定位誤差および他の知覚アーティファクトが生じ得るので、これは問題への最適な解決策を表さないことがある。

【0199】

[0195] 本開示で説明される技法の様々な態様は、上記で概説された戦略の欠陥を克服し得る。本技法は、以下のように、仮想ラウドスピーカー信号の様々な取扱いを提供し得る。本技法の第1の態様は、デバイス10が、2次元パンニング方法を使用して、下半球から水平面上に来て2つの最も近い現実のラウドスピーカー上に投射される仮想ラウドスピーカーを直交してマッピングすることを可能にし得る。その結果、本技法の第1の態様は、間違っ て投射された仮想ラウドスピーカーによって引き起こされる定位誤差を最小化、低減または除去し得る。第2に、耳の高さ(またはその周り)にある上半球中の仮想ラウドスピーカーも、本開示で説明される技法の第2の態様に従って2次元パンニング方法を使用して、2つの最も近いラウドスピーカーに投射され得る。この第2の修正形態の背後にある理由は、人間は、アジマス方向の知覚と比較して、仰角の音源の知覚はそれほど正確でないことがあるからであり得る。VBAPは、仮想音源のアジマス方向の作成は正確であることが概して知られているが、仰角の音の作成は比較的不正確であり、しばしば、知覚される仮想音源は、意図されたよりも高い仰角で知覚される。本技法の第2の態様は、それから恩恵を受けることがなく品質の劣化さえ生じるであろう、空間エリア中で3D-VBAPを使用することを回避する。

【0200】

[0196] 本技法の第3の態様は、イヤレベルの上側の上半球のすべての残りの仮想ラウドスピーカーが、従来の3次元パンニング方法を使用して投射されることである。いくつかの事例では、本技法の第4の態様が実行され得、ここでは、すべての高次アンビソニックス/球面調和係数サラウンド音素材は、素材のより滑らかな空間再現を高めるために、球面調和次数の関数として重み付け関数を使用して重み付けされる。これは、2Dおよび3Dパンニングされた仮想ラウドスピーカーのエネルギーを整合させるために潜在的に有益であることが示されている。

【0201】

[0197] 本開示で説明される技法の各態様を実行するものとして示されているが、3Dレンダラ決定ユニット48Cは、本開示で説明される態様の任意の組合せを実行し、上記4つの態様のうちの1つまたは複数を実行し得る。いくつかの事例では、球面調和係数を

10

20

30

40

50

生成する異なるデバイスが、本技法の様々な態様を相互の様式で実行し得る。冗長性を回避するために詳細に説明されていないが、本開示の技法は、図14Aの例に厳密に限定されるべきではない。

【0202】

[0198] 上記のセクションでは、5.1互換システムのための設計について論じた。詳細は、異なるターゲットフォーマットのために相応に調整され得る。一例として、7.1システムのための互換性を可能にするために、行列が可逆になるように、2つの余分のオーディオコンテンツチャンネルが互換要件に追加され、2つのさらなるSHCが基本セットに追加され得る。7.1システムのための大多数のラウドスピーカー配置（たとえば、ドルビーTrueHD）は依然として水平面上にあるので、SHCの選択は、高さ情報による選択を依然として除外することができる。このようにして、水平面信号レンダリングは、レンダリングシステム中の追加されたラウドスピーカーチャンネルから恩恵を受けることになる。高さダイバーシティをもつラウドスピーカーを含むシステム（たとえば、9.1、11.1および22.2システム）では、高さ情報をもつSHCを基本セット中に含めることが望ましいことがある。ステレオおよびモノのようにより少ないチャンネル数では、既存の5.1ソリューションは、コンテンツ情報を維持するためにダウンミックス（downmix）をカバーするのに十分であり得る。

10

【0203】

[0199] このように、上記のことは、要素の階層セット（たとえば、SHCのセット）と、複数のオーディオチャンネルとの間で変換するためのロスレス機構を表す。マルチチャンネルオーディオ信号がさらなるコーディング雑音を受けない限り、どんな誤差も発生しない。それらがコーディング雑音を受ける場合、SHCへの変換は誤差を発生し得る。しかしながら、それらの影響を低減するために係数の値を監視し、適切なアクションを取ることによって、これらの誤差をなくすことが可能である。これらの方法は、SHC表現における固有の冗長性を含む、SHCの特性を考慮に入れ得る。

20

【0204】

[0200] 本明細書で説明される手法は、音場のSHCベースの表現の使用における潜在的欠点に対する解決策を提供する。この解決策がなければ、何百万ものレガシー再生システムにおいて機能を有することが可能ではないことによって強いられる顕著な欠点により、SHCベースの表現は展開されないであろう。

30

【0205】

[0201] したがって、本技法は、第1の例では、複数の物理スピーカーのうちの1つと幾何学的配置で配置された複数の仮想スピーカーのうちの1つとの間の位置の差を決定するための手段、たとえば、レンダラ決定ユニット40と、位置の決定された差に基づいて、幾何学的配置内の複数の仮想スピーカーのうちの1つの位置を調整するための手段、たとえば、レンダラ決定ユニット40とを備えるデバイスを提供し得る。

【0206】

[0202] 第2の例では、第1の例のデバイス、ここにおいて、位置の差を決定するための手段は、複数の物理スピーカーのうちの1つと複数の仮想スピーカーのうちの1つとの間の仰角の差を決定するための手段、たとえば、3Dレンダラ決定ユニット48Cを備える。

40

【0207】

[0203] 第3の例では、第1の例のデバイス、ここにおいて、図8A～図9および図14A～図16Bの例に関して上記でより詳細に説明されたように、位置の差を決定するための手段は、複数の物理スピーカーのうちの1つと複数の仮想スピーカーのうちの1つとの間の仰角の差を決定するための手段を備え、およびここにおいて、複数の仮想スピーカーのうちの1つの位置を調整するための手段は、仰角の決定された差がしきい値を超えるとき、複数の仮想スピーカーのうちの1つを複数の仮想スピーカーの元の仰角よりも低い仰角に投射するための手段を備える。

【0208】

50

【0204】 第4の例では、第1の例のデバイス、ここにおいて、図8A～図9および図14A～図16Bの例に関して上記でより詳細に説明されたように、位置の差を決定するための手段は、複数の物理スピーカのうちの一つと複数の仮想スピーカのうちの一つとの間の仰角の差を決定するための手段を備え、およびここにおいて、複数の仮想スピーカのうちの一つの位置を調整するための手段は、仰角の決定された差がしきい値を超えるとき、複数の仮想スピーカのうちの一つを複数の仮想スピーカのうちの一つの元の仰角よりも高い仰角に投射するための手段を備える。

【0209】

【0205】 第5の例では、図8Aおよび図8Bの例に関して上記でより詳細に説明されたように、再現される音場が、仮想スピーカの調整されたロケーションから発生するように思われる少なくとも一つの音を含むように、音場を再現するために、複数の物理スピーカを駆動するための複数のラウドスピーカチャネル信号を生成するとき、音場を記述する要素の階層セット上で2次元パンニングを実行するための手段をさらに備える、第1の例のデバイス。

10

【0210】

【0206】 第6の例では、第5の例のデバイス、ここにおいて、要素の階層セットは、複数の球面調和係数を備える。

【0211】

【0207】 第7の例では、第5の例のデバイス、ここにおいて、図8Aおよび図8Bの例に関して上記でより詳細に説明されたように、要素の階層セット上で2次元パンニングを実行するための手段は、複数のラウドスピーカチャネル信号を生成するとき、要素の階層セット上で2次元ベクトルベースの振幅パンニングを実行するための手段を備える。

20

【0212】

【0208】 第8の例では、図8A～図12Bの例に関して上記でより詳細に説明されたように、複数の物理スピーカのうちの一つまたは複数の位置とは異なる一つまたは複数の伸長された物理スピーカ位置を決定するための手段をさらに備える、第1の例のデバイス。

【0213】

【0209】 第9の例では、図8A～図12Bの例に関して上記でより詳細に説明されたように、複数の物理スピーカのうちの一つまたは複数の位置とは異なる一つまたは複数の伸長された物理スピーカ位置を決定するための手段をさらに備える、第1の例のデバイス、ここにおいて、位置の差を決定するための手段は、複数の仮想スピーカのうちの一つの位置に対して、伸長された物理スピーカ位置のうちの一つとの間の差を決定するための手段を備える。

30

【0214】

【0210】 第10の例では、図8A～図12Bおよび図14A～図16Bの例に関して上記でより詳細に説明されたように、複数の物理スピーカのうちの一つまたは複数の位置とは異なる一つまたは複数の伸長された物理スピーカ位置を決定するための手段をさらに備える、第1の例のデバイス、ここにおいて、位置の差を決定するための手段は、伸長された物理スピーカ位置のうちの一つと複数の仮想スピーカのうちの一つの位置との間の仰角の差を決定するための手段を備え、およびここにおいて、複数の仮想スピーカのうちの一つの位置を調整するための手段は、仰角の決定された差がしきい値を超えるとき、複数の仮想スピーカのうちの一つを複数の仮想スピーカの元の仰角よりも低い仰角に投射するための手段を備える。

40

【0215】

【0211】 第11の例では、図8A～図12Bおよび図14A～図16Bの例に関して上記でより詳細に説明されたように、複数の物理スピーカのうちの一つまたは複数の位置とは異なる一つまたは複数の伸長された物理スピーカ位置を決定するための手段をさらに備える、第1の例のデバイス、ここにおいて、位置の差を決定するための手段は、伸長された物理スピーカ位置のうちの一つと複数の仮想スピーカのう

50

ちの1つの位置との間の仰角の差を決定するための手段を備え、およびここにおいて、複数の仮想スピーカのうち1つの位置を調整するための手段は、仰角の決定された差がしきい値を超えると、複数の仮想スピーカのうち1つを複数の仮想スピーカの元の仰角よりも高い仰角に投射するための手段を備える。

【0216】

[0212] 第12の例では、第1の例のデバイス、ここにおいて、図8A～図12Bおよび図14A～図16Bの例に関して上記でより詳細に説明されたように、複数の仮想スピーカは、球面幾何学的配置(spherical geometry)で配置される。

【0217】

[0213] 第13の例では、第1の例のデバイス、ここにおいて、複数の仮想スピーカは、多面体幾何学的配置(polyhedron geometry)で配置される。説明を簡単にするために、本開示の図1～図17によって示される例のいずれにおいても図示されていないが、本技法は、いくつかの例を提供すれば、立方幾何学的配置、12面体幾何学的配置、20・12面体幾何学的配置、菱形30面体幾何学的配置、プリズム幾何学的配置、およびピラミッド幾何学的配置など、任意の形態の多面体幾何学的配置を含む、任意の仮想スピーカ幾何学的配置に関して実行され得る。

10

【0218】

[0214] 第14の例では、第1の例のデバイス、ここにおいて、複数の物理スピーカは、不規則なスピーカ幾何学的配置で配置される。

【0219】

20

[0215] 第15の例では、第1の例のデバイス、ここにおいて、複数の物理スピーカは、複数の異なる水平面上に不規則なスピーカ幾何学的配置で配置される。

【0220】

[0216] 例に応じて、本明細書で説明された方法のいずれかのいくつかの行為またはイベントは、異なる順序で実行され得、互いに付加、統合、または除外され得る(たとえば、すべての説明された行為またはイベントが、方法の実行のために必要であるとは限らない)ことを理解されたい。その上、いくつかの例では、行為またはイベントは、連続的ではなく、たとえば、マルチスレッド処理、割り込み処理、または複数のプロセッサを通じて、同時に実行され得る。さらに、本開示のいくつかの態様は、明快のために単一のデバイス、モジュールまたはユニットによって実行されるものとして説明されているが、本開示の技法は、デバイス、ユニットまたはモジュールの組合せによって実行され得ることを理解されたい。

30

【0221】

[0217] 1つまたは複数の例では、説明された機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せで実装され得る。ソフトウェアで実装される場合、機能は、1つまたは複数の命令またはコードとしてコンピュータ可読媒体上に記憶されるか、あるいはコンピュータ可読媒体を介して送信され、ハードウェアベースの処理ユニットによって実行され得る。コンピュータ可読媒体は、たとえば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を可能にする任意の媒体を含む、データ記憶媒体または通信媒体などの有形媒体に対応するコンピュータ可読記憶媒体を含み得る。

40

【0222】

[0218] このようにして、コンピュータ可読媒体は、概して、(1)非一時的である有形コンピュータ可読記憶媒体、あるいは(2)信号または搬送波などの通信媒体に対応し得る。データ記憶媒体は、本開示で説明された技法の実装のための命令、コードおよび/またはデータ構造を取り出すために1つまたは複数のコンピュータあるいは1つまたは複数のプロセッサによってアクセスされ得る任意の利用可能な媒体であり得る。コンピュータプログラム製品はコンピュータ可読媒体を含み得る。

【0223】

[0219] 限定ではなく例として、そのようなコンピュータ可読記憶媒体は、RAM、R

50

OM、EEPROM（登録商標）、CD-ROMまたは他の光ディスクストレージ、磁気ディスクストレージ、または他の磁気ストレージデバイス、フラッシュメモリ、あるいは、命令またはデータ構造の形態の所望のプログラムコードを記憶するために使用されコンピュータによってアクセスされ得る、任意の他の媒体を備えることができる。また、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、同軸ケーブル、光ファイバーケーブル、ツイストペア、デジタル加入者回線（DSL）、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して、ウェブサイト、サーバ、または他のリモートソースから送信される場合、同軸ケーブル、光ファイバーケーブル、ツイストペア、DSL、または赤外線、無線、およびマイクロ波などのワイヤレス技術は、媒体の定義に含まれる。

10

【0224】

[0220] ただし、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時媒体を含まないが、代わりに非一時的有形記憶媒体を対象とすることを理解されたい。本明細書で使用されるディスク（disk）およびディスク（disc）は、コンパクトディスク（disc）（CD）、レーザーディスク（登録商標）（disc）、光ディスク（disc）、デジタル多用途ディスク（disc）（DVD）、フロッピー（登録商標）ディスク（disk）およびBlu-rayディスク（disc）を含み、ディスク（disk）は、通常、データを磁気的に再生し、ディスク（disc）は、データをレーザーで光学的に再生する。上記の組合せもコンピュータ可読媒体の範囲内に含まれるべきである。

【0225】

20

[0221] 命令は、1つまたは複数のデジタル信号プロセッサ（DSP）、汎用マイクロプロセッサ、特定用途向け集積回路（ASIC）、フィールドプログラマブル論理アレイ（FPGA）、あるいは他の等価な集積回路またはディスクリート論理回路など、1つまたは複数のプロセッサによって実行され得る。したがって、本明細書で使用される「プロセッサ（processor）」という用語は、前述の構造、または本明細書で説明された技法の実装に適した任意の他の構造のいずれかを指し得る。さらに、いくつかの態様では、本明細書で説明された機能は、符号化および復号のために構成された専用のハードウェアおよび/またはソフトウェアモジュール内に提供されるか、あるいは複合コーデックに組み込まれ得る。また、本技法は、1つまたは複数の回路または論理要素において完全に実装され得る。

30

【0226】

[0222] 本開示の技法は、ワイヤレスハンドセット、集積回路（IC）、またはICのセット（たとえば、チップセット）を含む、多種多様なデバイスまたは装置において実装され得る。本開示では、開示された技法を実行するように構成されたデバイスの機能的態様を強調するために様々な構成要素、モジュール、またはユニットについて説明されたが、それらの構成要素、モジュール、またはユニットを、必ずしも異なるハードウェアユニットによって実現する必要があるとは限らない。むしろ、上記で説明されたように、様々なユニットが、好適なソフトウェアおよび/またはファームウェアとともに、上記で説明された1つまたは複数のプロセッサを含めて、コーデックハードウェアユニットにおいて組み合わせられるか、または相互動作ハードウェアユニットの集合によって提供され得る。

40

【0227】

[0223] 本技法の様々な実施形態が説明された。これらおよび他の実施形態は以下の特許請求の範囲内に入る。

以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

[C1]

音場を表す球面調和係数の再生のために使用される1つまたは複数のスピーカーのローカルスピーカー幾何学的配置を決定することと、

前記ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することと

を備える方法。

50

[C 2]

マルチチャンネルオーディオデータを生成するために前記決定された2次元レンダラまたは3次元レンダラを使用して前記球面調和係数をレンダリングすることをさらに備える、C 1に記載の方法。

[C 3]

前記ローカルスピーカー幾何学的配置に基づいて前記2次元レンダラまたは3次元レンダラを決定することは、前記ローカルスピーカー幾何学的配置がステレオスピーカー幾何学的配置に合致するとき、2次元ステレオレンダラを決定することを備える、C 1に記載の方法。

[C 4]

前記ローカルスピーカー幾何学的配置に基づいて前記2次元レンダラまたは3次元レンダラを決定することは、前記ローカルスピーカー幾何学的配置が、3つ以上のスピーカーを有する水平マルチチャンネルスピーカー幾何学的配置に合致するとき、水平2次元マルチチャンネルレンダラを決定することを備える、C 1に記載の方法。

[C 5]

前記水平2次元マルチチャンネルレンダラを決定することは、前記決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な水平2次元マルチチャンネルレンダラを決定することを備える、C 4に記載の方法。

[C 6]

前記水平2次元マルチチャンネルレンダラを決定することは、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な水平2次元マルチチャンネルレンダラを決定することを備える、C 4に記載の方法。

[C 7]

前記ローカルスピーカー幾何学的配置に基づいて前記2次元レンダラまたは3次元レンダラを決定することは、前記ローカルスピーカー幾何学的配置が、2つ以上の水平面上に3つ以上のスピーカーを有する3次元マルチチャンネルスピーカー幾何学的配置に合致するとき、3次元マルチチャンネルレンダラを決定することを備える、C 1に記載の方法。

[C 8]

前記3次元マルチチャンネルレンダラを決定することは、前記決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すとき、不規則な3次元マルチチャンネルレンダラを決定することを備える、C 7に記載の方法。

[C 9]

前記3次元マルチチャンネルレンダラを決定することは、前記決定されたローカルスピーカー幾何学的配置がほぼ規則的なスピーカー幾何学的配置を示すとき、ほぼ規則的な3次元マルチチャンネルレンダラを決定することを備える、C 7に記載の方法。

[C 10]

前記3次元マルチチャンネルレンダラを決定することは、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すとき、規則的な3次元マルチチャンネルレンダラを決定することを備える、C 7に記載の方法。

[C 11]

前記レンダラを決定することは、
前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカー幾何学的配置を鑑みてレンダリングされることを必要とされる前記球面調和係数のものを識別する、
前記決定された許容次数に基づいて前記レンダラを決定することと
を備える、C 1に記載の方法。

[C 12]

前記2次元レンダラまたは3次元レンダラを決定することは、
前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカー幾何学的配置を鑑みてレンダリングされる

10

20

30

40

50

ことを必要とされる前記球面調和係数のものを識別する、

前記 2 次元レンダラまたは 3 次元レンダラが、前記決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた前記球面調和係数のもののみをレンダリングするように、前記 2 次元レンダラまたは 3 次元レンダラを決定することと

を備える、C 1 に記載の方法。

[C 1 3]

前記 1 つまたは複数のスピーカの前記ローカルスピーカー幾何学的配置を決定することが、前記ローカルスピーカー幾何学的配置を記述するローカルスピーカー幾何学的配置情報を指定する入力を聴取者から受信することを備える、C 1 に記載の方法。

10

[C 1 4]

前記ローカルスピーカー幾何学的配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定することは、前記ローカルスピーカー幾何学的配置がモノスピーカー幾何学的配置に合致するとき、モノレンダラを決定することを備える、C 1 に記載の方法。

[C 1 5]

音場を表す球面調和係数の再生のために使用される 1 つまたは複数のスピーカの前記ローカルスピーカー幾何学的配置を決定することと、前記決定されたローカルスピーカー幾何学的配置に基づいて動作するようにデバイスを構成することとを行うように構成された 1 つまたは複数のプロセッサを備えるデバイス。

[C 1 6]

20

前記 1 つまたは複数のプロセッサが、前記決定されたローカルスピーカー幾何学的配置に基づいて動作するように前記デバイスを構成するとき、前記ローカルスピーカー幾何学的配置に基づいて 2 次元レンダラまたは 3 次元レンダラを決定することと、前記決定された 2 次元レンダラまたは 3 次元レンダラに従って動作するように前記デバイスを構成することとを行うようにさらに構成された、C 1 5 に記載のデバイス。

[C 1 7]

前記 1 つまたは複数のプロセッサが、マルチチャンネルオーディオデータを生成するために前記決定された 2 次元レンダラまたは 3 次元レンダラを使用して前記球面調和係数をレンダリングするようにさらに構成された、C 1 6 に記載のデバイス。

[C 1 8]

30

前記 1 つまたは複数のプロセッサは、前記ローカルスピーカー幾何学的配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ローカルスピーカー幾何学的配置がステレオスピーカー幾何学的配置に合致するときに 2 次元ステレオレンダラを決定するようにさらに構成された、C 1 6 に記載のデバイス。

[C 1 9]

前記 1 つまたは複数のプロセッサは、前記ローカルスピーカー幾何学的配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ローカルスピーカー幾何学的配置が、3 つ以上のスピーカを有する水平マルチチャンネルスピーカー幾何学的配置に合致するときに水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 1 6 に記載のデバイス。

40

[C 2 0]

前記 1 つまたは複数のプロセッサは、前記水平 2 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカー幾何学的配置が不規則なスピーカー幾何学的配置を示すときに不規則な水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 1 9 に記載のデバイス。

[C 2 1]

前記 1 つまたは複数のプロセッサは、前記水平 2 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカー幾何学的配置が規則的なスピーカー幾何学的配置を示すときに規則的な水平 2 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 1 9 に記載のデバイス。

50

[C 2 2]

前記 1 つまたは複数のプロセッサは、前記ローカルスピーカ幾何学的配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ローカルスピーカ幾何学的配置が、2 つ以上の水平面上に 3 つ以上のスピーカを有する 3 次元マルチチャンネルスピーカ幾何学的配置を合致するとき、3 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 1 6 に記載のデバイス。

[C 2 3]

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカ幾何学的配置が不規則なスピーカ幾何学的配置を示すときに不規則な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 2 2 に記載のデバイス。

10

[C 2 4]

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカ幾何学的配置がほぼ規則的なスピーカ幾何学的配置を示すときにほぼ規則的な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 2 2 に記載のデバイス。

[C 2 5]

前記 1 つまたは複数のプロセッサは、前記 3 次元マルチチャンネルレンダラを決定するとき、前記決定されたローカルスピーカ幾何学的配置が規則的なスピーカ幾何学的配置を示すときに規則的な 3 次元マルチチャンネルレンダラを決定するようにさらに構成された、C 2 2 に記載のデバイス。

20

[C 2 6]

前記 1 つまたは複数のプロセッサは、前記レンダラを決定するとき、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる前記球面調和係数のものを識別する、前記決定された許容次数に基づいて前記レンダラを決定することを行うようにさらに構成された、C 1 6 に記載のデバイス。

[C 2 7]

前記 1 つまたは複数のプロセッサは、前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記球面調和係数が関連付けられた球面基底関数の許容次数を決定することと、前記許容次数が、前記決定されたローカルスピーカ幾何学的配置を鑑みてレンダリングされることを必要とされる前記球面調和係数のものを識別する、前記 2 次元レンダラまたは 3 次元レンダラが、前記決定された許容次数よりも小さいかまたはそれに等しい次数を有する球面基底関数に関連付けられた前記球面調和係数のもののみをレンダリングするように、前記 2 次元レンダラまたは 3 次元レンダラを決定することとを行うようにさらに構成された、C 1 6 に記載のデバイス。

30

[C 2 8]

前記 1 つまたは複数のプロセッサが、前記 1 つまたは複数のスピーカの前記ローカルスピーカ幾何学的配置を決定するとき、前記ローカルスピーカ幾何学的配置を記述するローカルスピーカ幾何学的配置情報を指定する入力を聴取者から受信するようにさらに構成された、C 1 6 に記載のデバイス。

40

[C 2 9]

前記 1 つまたは複数のプロセッサは、前記ローカルスピーカ幾何学的配置に基づいて前記 2 次元レンダラまたは 3 次元レンダラを決定するとき、前記ローカルスピーカ幾何学的配置がモノスピーカ幾何学的配置に合致するときにモノレンダラを決定するようにさらに構成された、C 1 6 に記載のデバイス。

[C 3 0]

実行されたとき、
音場を表す球面調和係数の再生のために使用される 1 つまたは複数のスピーカのローカルスピーカ幾何学的配置を決定することと、

50

前記ローカルスピーカー幾何学的配置に基づいて2次元レンダラまたは3次元レンダラを決定することと
 を1つまたは複数のプロセッサに行わせる命令を記憶した非一時的コンピュータ可読記憶媒体。

【図1】

図1

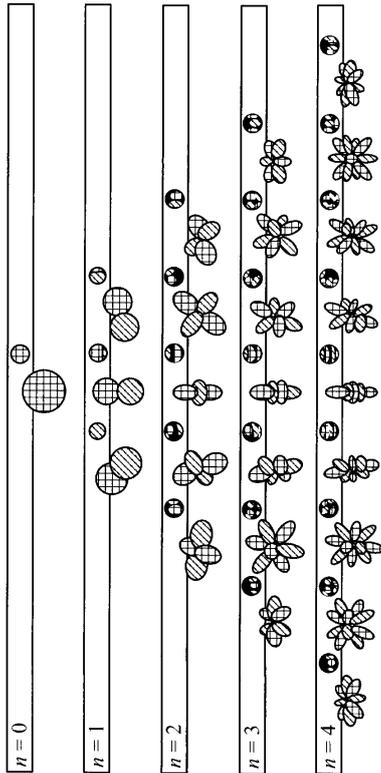


FIG. 1

【図2】

図2

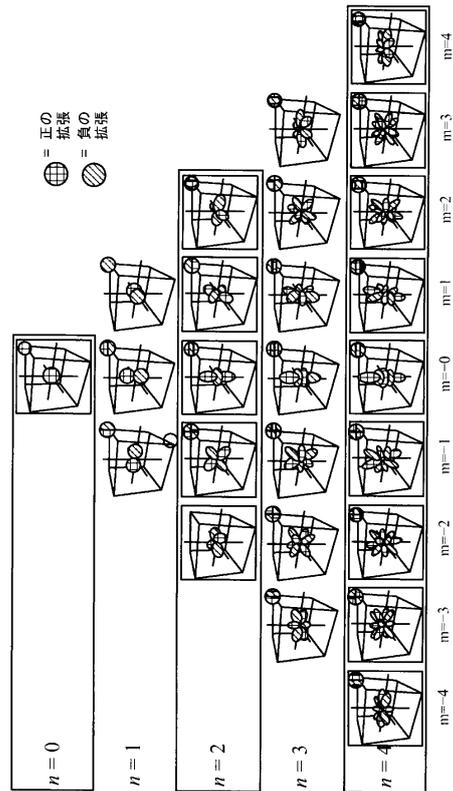


FIG. 2

【 図 3 】

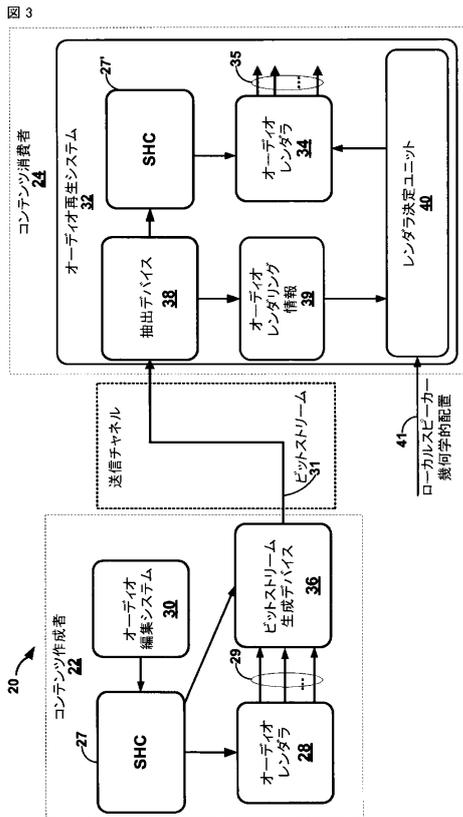


FIG. 3

【 図 4 】

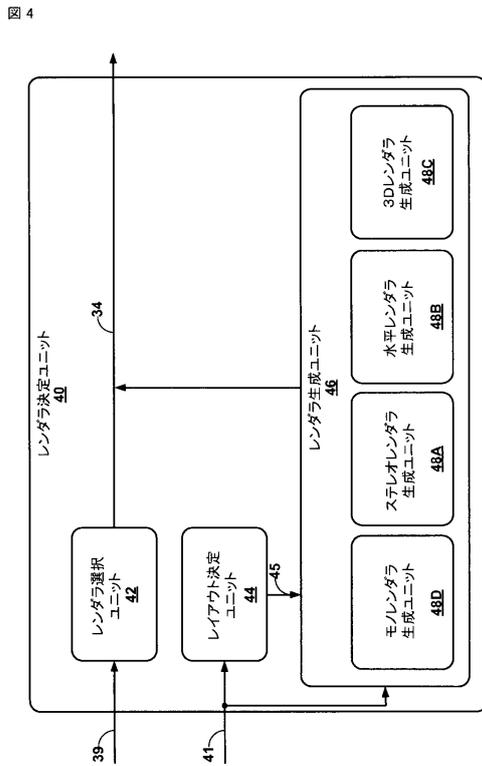


FIG. 4

【 図 5 】

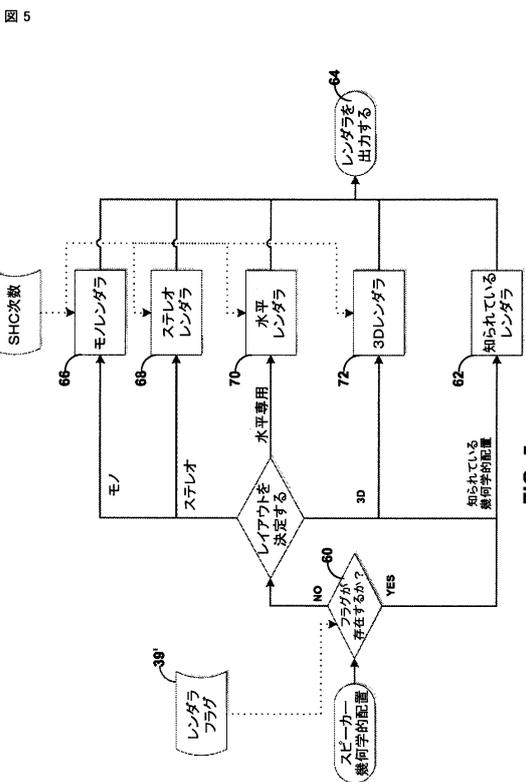


FIG. 5

【 図 6 】

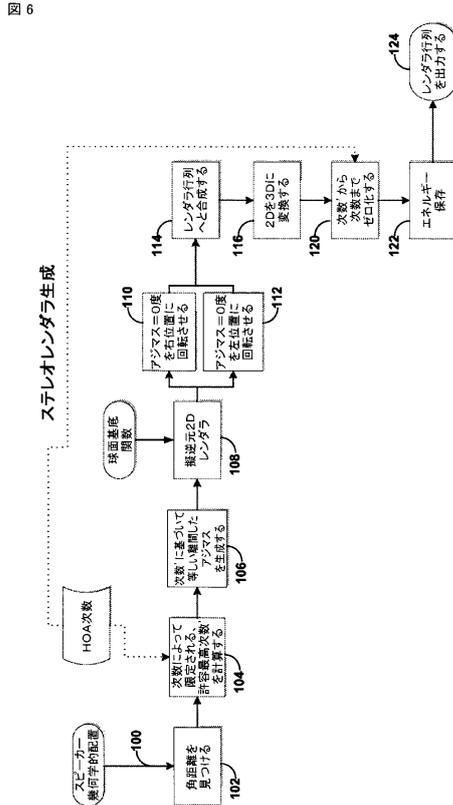


FIG. 6

【図7】

図7

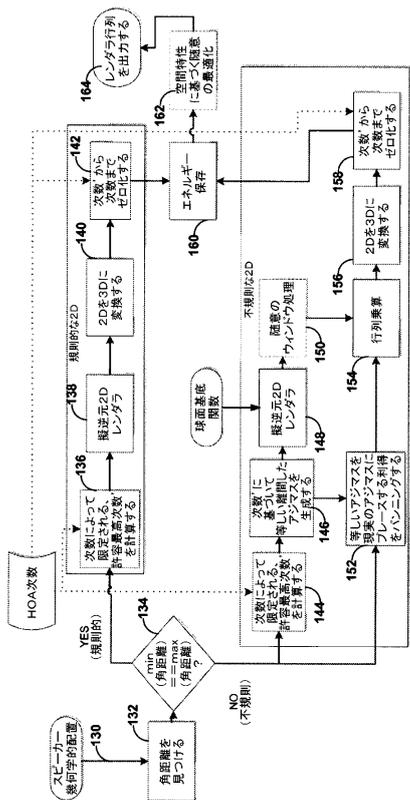


FIG. 7

【図8A】

図8A

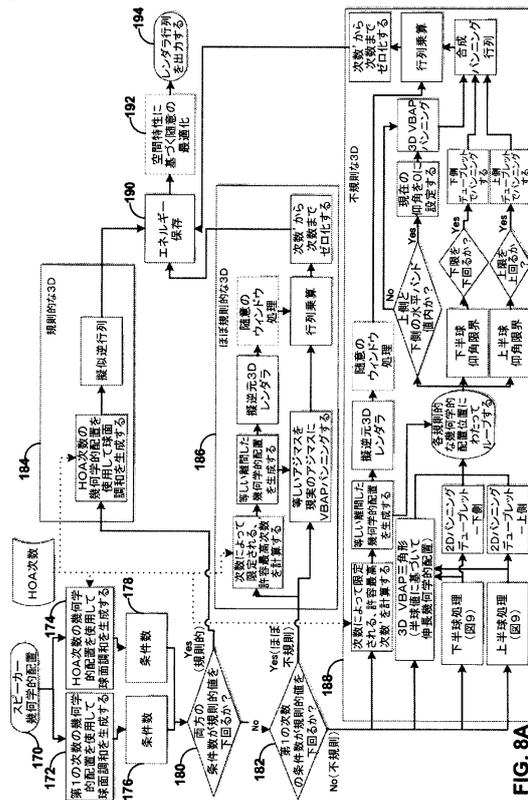


FIG. 8A

【図8B】

図8B

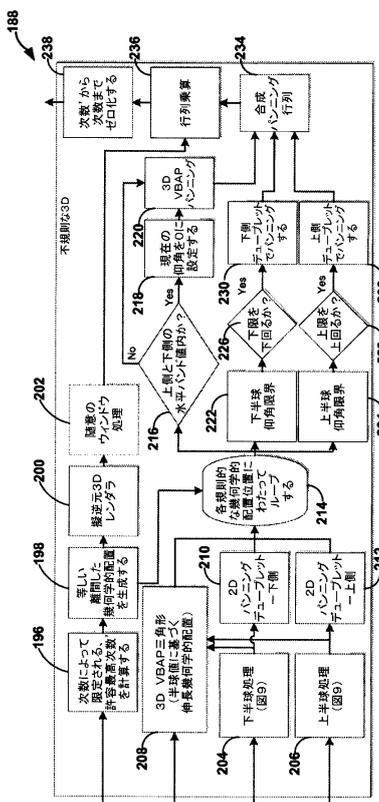


FIG. 8B

【図9】

図9

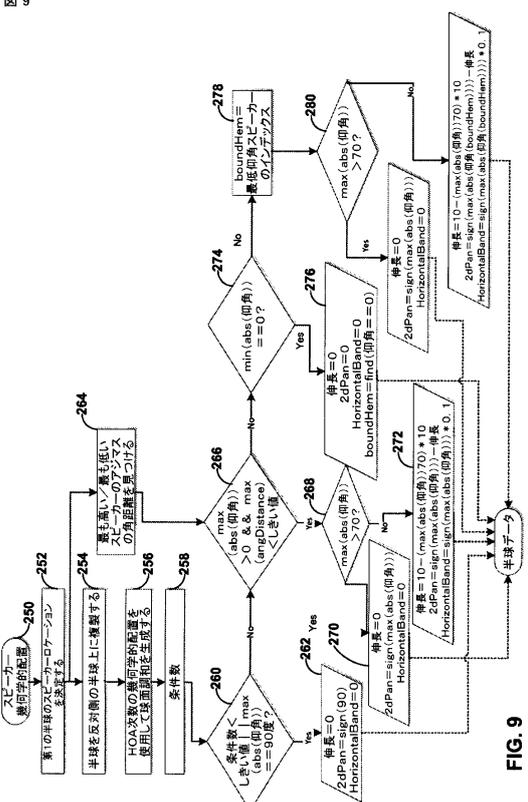


FIG. 9

【図10】

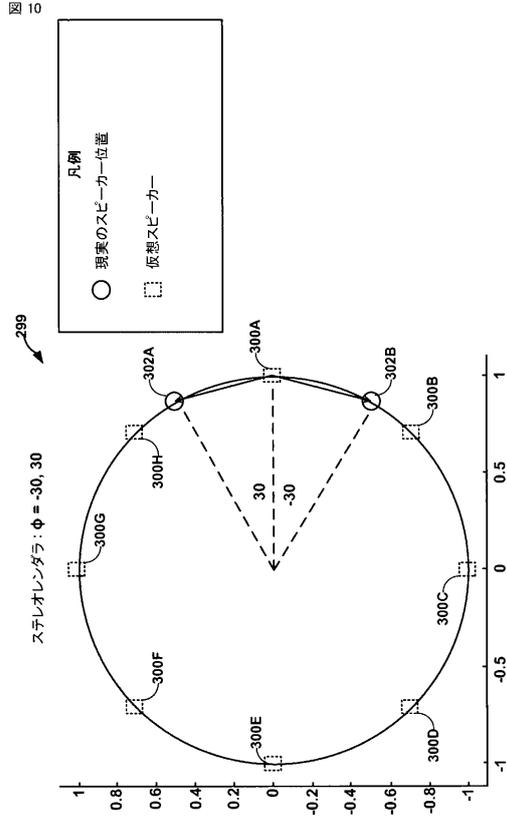


FIG. 10

【図11】

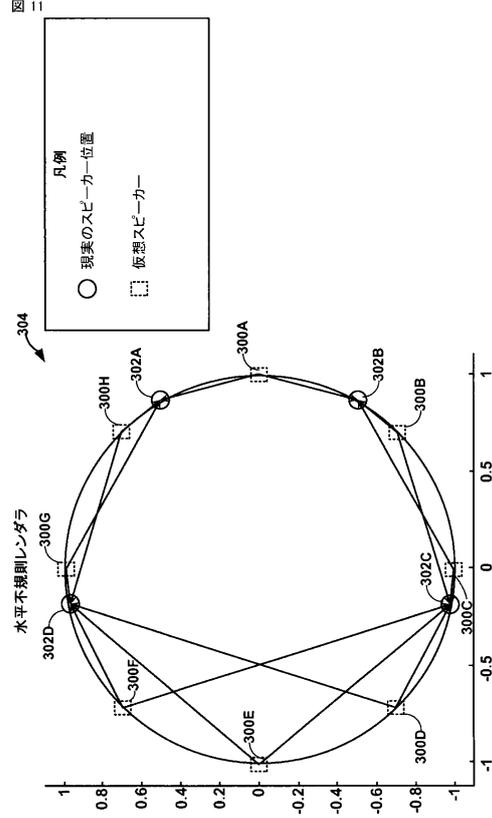


FIG. 11

【図12A】

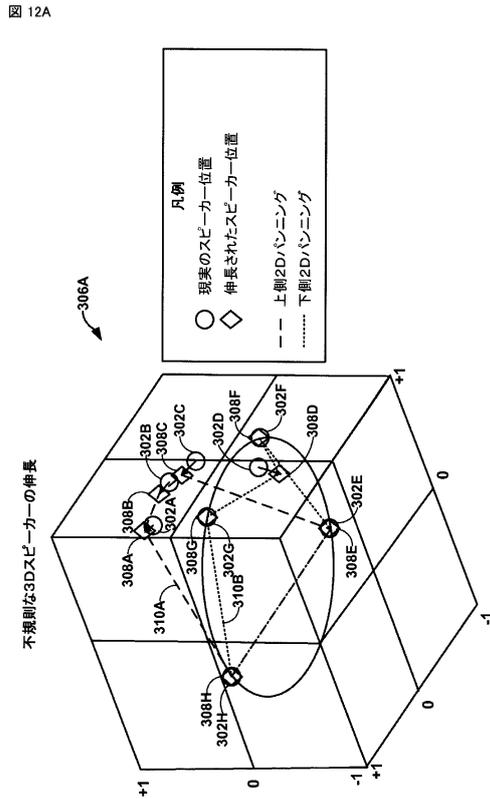


FIG. 12A

【図12B】

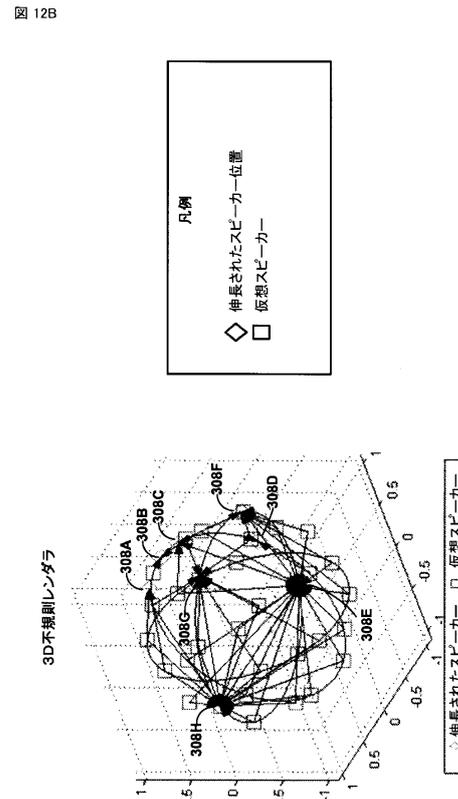


FIG. 12B

【 図 13 A 】

図 13A

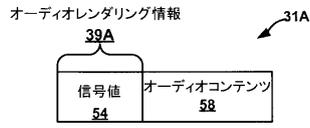


FIG. 13A

【 図 13 B 】

図 13B

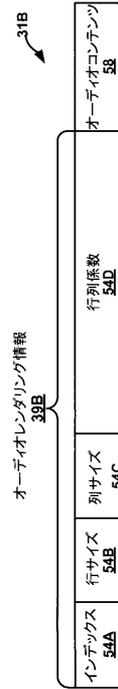


FIG. 13B

【 図 13 C 】

図 13C

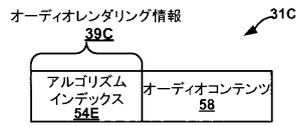


FIG. 13C

【 図 14 A 】

図 14A

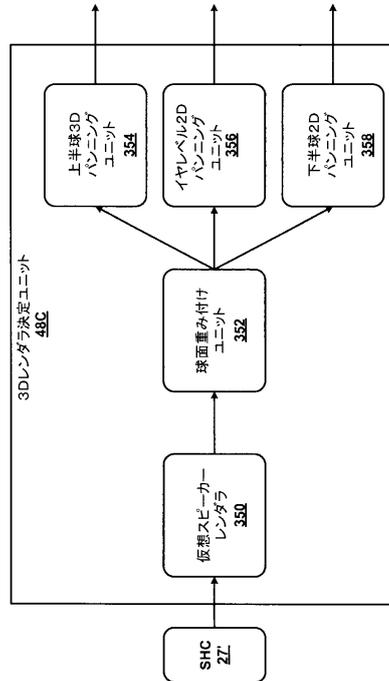


FIG. 14A

【 図 13 D 】

図 13D



FIG. 13D

【 図 14 B 】

図 14B

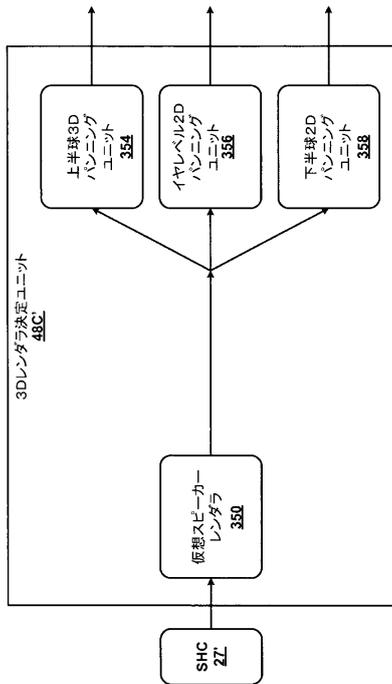


FIG. 14B

【 図 15 A 】

図 15A

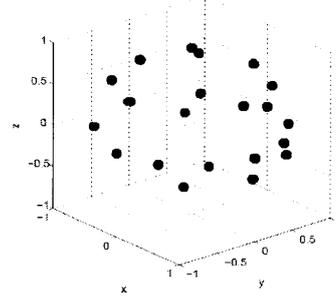


FIG. 15A

【 図 15 B 】

図 15B

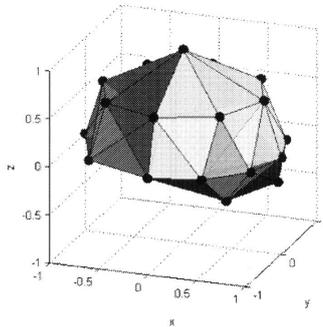


FIG. 15B

【 図 16 A 】

図 16A

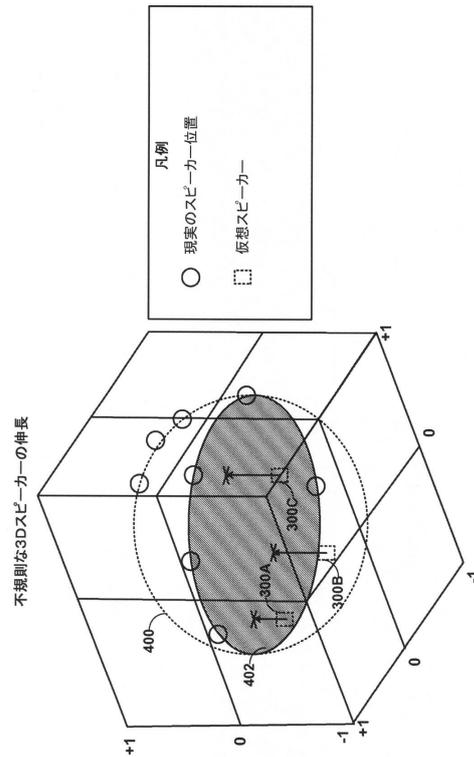


FIG. 16A

【図 16B】

図 16B

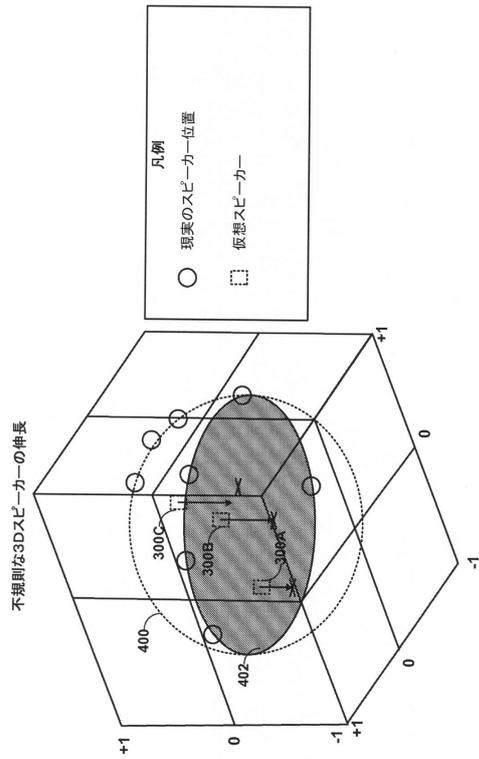


FIG. 16B

【図 17】

図 17

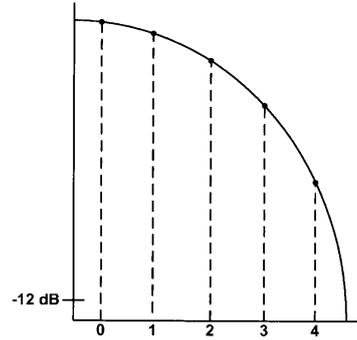


FIG. 17

フロントページの続き

- (31)優先権主張番号 14/174,784
(32)優先日 平成26年2月6日(2014.2.6)
(33)優先権主張国 米国(US)

早期審査対象出願

- (72)発明者 モーレル、マーティン・ジェームズ
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(72)発明者 ペーターズ、ニルス・グンサー
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(72)発明者 セン、ディバンジャン
アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

審査官 下林 義明

- (56)参考文献 特許第4861593(JP, B2)
特表2012-514358(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04S 1/00 - 7/00
H04R 3/00 - 3/14