

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2017-201412

(P2017-201412A)

(43) 公開日 平成29年11月9日 (2017.11.9)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 O L 19/008 (2013.01)	G 1 O L 19/008 2 0 0	
G 1 O L 19/035 (2013.01)	G 1 O L 19/035 Z	
G 1 O L 19/038 (2013.01)	G 1 O L 19/038	

審査請求 有 請求項の数 28 O L (全 82 頁)

(21) 出願番号	特願2017-126158 (P2017-126158)	(71) 出願人	595020643
(22) 出願日	平成29年6月28日 (2017. 6. 28)		クアルコム・インコーポレイテッド
(62) 分割の表示	特願2016-548734 (P2016-548734) の分割		QUALCOMM INCORPORATED
原出願日	平成27年1月30日 (2015. 1. 30)		アメリカ合衆国、カリフォルニア州 92
(31) 優先権主張番号	61/933, 706		121-1714、サン・ディエゴ、モア
(32) 優先日	平成26年1月30日 (2014. 1. 30)		ハウス・ドライブ 5775
(33) 優先権主張国	米国 (US)	(74) 代理人	100108855
(31) 優先権主張番号	61/933, 714		弁理士 蔵田 昌俊
(32) 優先日	平成26年1月30日 (2014. 1. 30)	(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	61/933, 731	(74) 代理人	100158805
(32) 優先日	平成26年1月30日 (2014. 1. 30)		弁理士 井関 守三
(33) 優先権主張国	米国 (US)	(74) 代理人	100112807
			弁理士 岡田 貴志

最終頁に続く

(54) 【発明の名称】 フレームパラメータ再使用可能性を示すこと

(57) 【要約】 (修正有)

【課題】 ベクトルを復号するためのフレームパラメータ再使用可能性を示すための技法を提供する。

【解決手段】 ベクトルを復号するためのフレームパラメータ再使用可能性を示すための技法であって、技法を実行するオーディオ復号デバイス 24 は、プロセッサとメモリとを。プロセッサは、球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得する。ビットストリームは、ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える。メモリは、ビットストリームを記憶する。

【選択図】 図 4

図 4

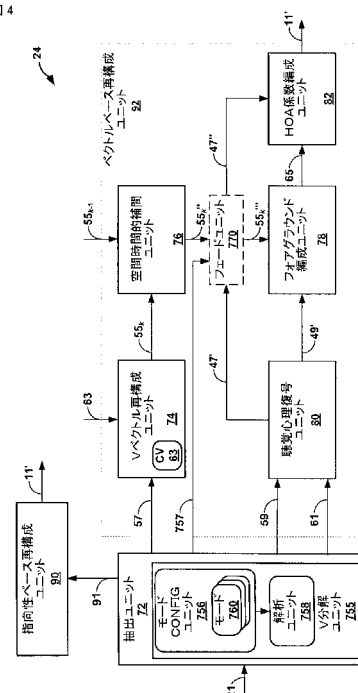


FIG. 4

【特許請求の範囲】**【請求項 1】**

効率的なビット使用の方法であって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを備え、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、

方法。

【請求項 2】

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシンタックス要素の 1 つまたは複数のビットを備える、

請求項 1 に記載の方法。

【請求項 3】

前記シンタックス要素の前記 1 つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも 1 つのシンタックス要素を再使用することを示す、

請求項 2 に記載の方法。

【請求項 4】

前記量子化モードは、ベクトル量子化モードを備える、

請求項 2 に記載の方法。

【請求項 5】

前記量子化モードは、ハフマンコーディングなしのスカラ量子化モードを備える、

請求項 2 に記載の方法。

【請求項 6】

前記量子化モードは、ハフマンコーディングありのスカラ量子化モードを備える、

請求項 2 に記載の方法。

【請求項 7】

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第 2 の最上位ビットとを備える、

請求項 2 に記載の方法。

【請求項 8】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

請求項 1 に記載の方法。

【請求項 9】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

請求項 1 に記載の方法。

【請求項 10】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルが対応する圧縮カテゴリを識別するカテゴリ識別子を示すシンタックス要素を備える、

請求項 1 に記載の方法。

【請求項 11】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

請求項 1 に記載の方法。

【請求項 12】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

請求項 1 に記載の方法。

10

20

30

40

50

【請求項 13】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

請求項 1 に記載の方法。

【請求項 14】

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

請求項 1 に記載の方法。

10

【請求項 15】

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと

、
前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することと

をさらに備える、請求項 1 に記載の方法。

【請求項 16】

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得することと、

高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせることと

をさらに備える、請求項 1 に記載の方法。

20

【請求項 17】

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

請求項 1 に記載の方法。

【請求項 18】

効率的なビット使用を実行するように構成されたデバイスであって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得すること、ここにおいて、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、と、

前記ビットストリームを記憶するように構成されるメモリと

を備える、デバイス。

30

【請求項 19】

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシンタックス要素の 1 つまたは複数のビットを備える、

請求項 18 に記載のデバイス。

【請求項 20】

前記シンタックス要素の前記 1 つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも 1 つのシンタックス要素を再使用することを示す

40

、
請求項 19 に記載のデバイス。

【請求項 21】

前記量子化モードは、ベクトル量子化モードを備える、

請求項 19 に記載のデバイス。

【請求項 22】

前記量子化モードは、ハフマンコーディングなしのスカラー量子化モードを備える、

請求項 19 に記載のデバイス。

【請求項 23】

前記量子化モードは、ハフマンコーディングありのスカラー量子化モードを備える、

50

請求項 19 に記載のデバイス。

【請求項 24】

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第 2 の最上位ビットとを備える、

請求項 19 に記載のデバイス。

【請求項 25】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 26】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 27】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 28】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 29】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 30】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

請求項 18 に記載のデバイス。

【請求項 31】

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

請求項 18 に記載のデバイス。

【請求項 32】

前記 1 つまたは複数のプロセッサは、前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと、前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することとを行うようにさらに構成される、

請求項 18 に記載のデバイス。

【請求項 33】

前記 1 つまたは複数のプロセッサは、前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得することと、高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせることとを行うようにさらに構成される、

請求項 1 に記載の方法。

【請求項 34】

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

請求項 18 に記載のデバイス。

【請求項 35】

効率的なビット使用のデバイスであって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得するための手段、ここにおいて、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、と、

前記インジケータを記憶するための手段と

を備える、デバイス。

【請求項36】

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシンタックス要素の1つまたは複数のビットを備える、

請求項35に記載のデバイス。

【請求項37】

前記シンタックス要素の前記1つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも1つのシンタックス要素を再使用することを示す、

請求項36に記載のデバイス。

【請求項38】

前記量子化モードは、ベクトル量子化モードを備える、

請求項36に記載のデバイス。

【請求項39】

前記量子化モードは、ハフマンコーディングなしのスカラ量子化モードを備える、

請求項36に記載のデバイス。

【請求項40】

前記量子化モードは、ハフマンコーディングありのスカラ量子化モードを備える、

請求項36に記載のデバイス。

【請求項41】

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第2の最上位ビットとを備える、

請求項36に記載のデバイス。

【請求項42】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

請求項35に記載のデバイス。

【請求項43】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

請求項35に記載のデバイス。

【請求項44】

前記以前のフレームからの前記シンタックス要素は、ベクトルが対応する圧縮カテゴリを識別するカテゴリ識別子を示すシンタックス要素を備える、

請求項35に記載のデバイス。

【請求項45】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

請求項35に記載のデバイス。

【請求項46】

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

請求項35に記載のデバイス。

【請求項47】

10

20

30

40

50

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

請求項 35 に記載のデバイス。

【請求項 48】

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

請求項 35 に記載のデバイス。

【請求項 49】

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解するための手段と、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定するための手段と

をさらに備える、請求項 35 に記載のデバイス。

【請求項 50】

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得するための手段と、

高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせるための手段と

をさらに備える、請求項 35 に記載のデバイス。

【請求項 51】

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

請求項 35 に記載のデバイス。

【請求項 52】

実行されると、1つまたは複数のプロセッサに、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを行わせる命令を記憶し、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、

非一時的コンピュータ可読記憶媒体。

【発明の詳細な説明】

【関連出願】

【0001】

[0001] 本出願は、以下の米国仮出願、すなわち、

2014年1月30日に出願された「COMPRESSION OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の米国仮出願第61/933,706号、

2014年1月30日に出願された「COMPRESSION OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の米国仮出願第61/933,714号、

2014年1月30日に出願された「INDICATING FRAM (登録商標) E PARAMETER REUSABILITY FOR DECODING SPATIAL VECTORS」という名称の米国仮出願第61/933,731号、

2014年3月7日に出願された「IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS」という名称の米国仮出願第61/949,591号、

2014年3月7日に出願された「FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の米国仮出願第61/949,583号、

10

20

30

40

50

2014年5月16日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第61/994,794号、

2014年5月28日に出願された「INDICATING FRAME PARAMETER REUSABILITY FOR DECODING SPATIAL VECTORS」という名称の米国仮出願第62/004,147号、

2014年5月28日に出願された「IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS AND FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の米国仮出願第62/004,067号、

10

2014年5月28日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/004,128号、

2014年7月1日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/019,663号、

2014年7月22日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/027,702号、

20

2014年7月23日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/028,282号、

2014年7月25日に出願された「IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS AND FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の米国仮出願第62/029,173号、

2014年8月1日に出願された「CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/032,440号、

30

2014年9月26日に出願された「SWITCHED V-VECTOR QUANTIZATION OF A HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/056,248号、および

2014年9月26日に出願された「PREDICTIVE VECTOR QUANTIZATION OF A DECOMPOSED HIGHER ORDER AMBISONICS (HOA) AUDIO SIGNAL」という名称の米国仮出願第62/056,286号、および

2015年1月12日に出願された「TRANSITIONING OF AMBIENT HIGHER-ORDER AMBISONIC COEFFICIENTS」という名称の米国仮出願第62/102,243号の利益を主張し、上記に記載された米国仮出願の各々は、それらのそれぞれの全体として本明細書に記載されたかのように、参照により組み込まれる。

40

【技術分野】

【0002】

[0002]本開示はオーディオデータに関し、より詳細には、高次アンビソニックオーディオデータのコーディングに関する。

【背景技術】

【0003】

50

[0003]高次アンビソニックス（H O A）信号（複数の球面調和係数（S H C）または他の階層的な要素によって表されることが多い）は、音場の3次元表現である。このH O A表現またはS H C表現は、S H C信号からレンダリングされるマルチチャ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幾何学的配置に依存しない方法で、音場を表し得る。S H C信号は、5. 1オーディオチャネルフォーマットまたは7. 1オーディオチャネルフォーマットなどのよく知られており広く採用されているマルチチャネルフォーマットにレンダリングされ得るので、S H C信号はまた、下位互換性を容易にし得る。したがって、S H C表現は、下位互換性にも対応する、音場のより良い表現を可能にし得る。

【発明の概要】

【0004】

[0004]概して、高次アンビソニックスオーディオデータをコーディングするための技法が説明される。高次アンビソニックスオーディオデータは、1よりも大きい次数を有する球面調和基底関数に対応する少なくとも1つの球面調和係数を備え得る。

【0005】

[0005]一態様では、効率的なビット使用の方法は、球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを備える。ビットストリームは、ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再利用するかどうかのためのインジケータをさらに備える。

【0006】

[0006]別の態様では、効率的なビット使用を実行するように構成されたデバイスは、球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得するように構成される。ビットストリームは、ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再利用するかどうかのためのインジケータをさらに備える。デバイスはさらに、ビットストリームを記憶するように構成されるメモリを備える。

【0007】

[0007]別の態様では、デバイスは、効率的なビット使用を実行するように構成されたデバイスは、球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得するための手段を備える。ビットストリームは、ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再利用するかどうかのためのインジケータをさらに備える。デバイスはさらに、インジケータを記憶するための手段を備える。

【0008】

[0008]別の態様では、非一時的コンピュータ可読記憶媒体は、命令を記憶しており、命令は、実行されると、1つまたは複数のプロセッサに、球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを行わせ、ビットストリームは、ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再利用するかどうかのためのインジケータをさらに備える。

【0009】

[0009]本技法の1つまたは複数の態様の詳細は、添付の図面および以下の説明に記載される。本技法の他の特徴、目的、および利点は、その説明および図面、ならびに特許請求の範囲から明らかになる。

【図面の簡単な説明】

【0010】

【図1】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】本開示で説明される技法の様々な態様を実行することができるシステムを示す図。

【図3】本開示で説明される技法の様々な態様を実行することができる、図2の例に示されるオーディオ符号化デバイスの一例をより詳細に示すブロック図。

10

20

30

40

50

【図 4】図 2 のオーディオ復号デバイスをより詳細に示すブロック図。

【図 5 A】本開示で説明されるベクトルベース合成技法の様々な態様を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャート。

【図 5 B】本開示で説明されるコーディング技法の様々な態様を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャート。

【図 6 A】本開示で説明される技法の様々な態様を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャート。

【図 6 B】本開示で説明されるコーディング技法の様々な態様を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャート。

【図 7】圧縮された空間成分を指定することができるビットストリームまたはサイドチャネル情報の一部分をより詳細に示す図。

【図 8】より詳細に圧縮された空間成分を指定することができるビットストリームの一部分を示す図。

【発明を実施するための形態】

【0011】

[0020] サラウンドサウンドの発展は、現今では娯楽のための多くの出力フォーマットを利用可能にしている。そのような消費者向けのサラウンドサウンドフォーマットの例は、ある幾何学的な座標にあるラウドスピーカーへのフィードを暗黙的に指定するという点で、大半が「チャンネル」ベースである。消費者向けのサラウンドサウンドフォーマットは、普及している 5.1 フォーマット（これは、次の 6 つのチャンネル、すなわち、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）とを含む）、発展中の 7.1 フォーマット、7.1.4 フォーマットおよび 22.2 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）などのハイトスピーカーを含む様々なフォーマットを含む。消費者向けではないフォーマットは、「サラウンドアレイ」と呼ばれることが多い（対称な、および非対称な幾何学的配置の）任意の数のスピーカーに及び得る。そのようなアレイの一例は、切頂二十面体の角の座標に配置される 32 個のラウドスピーカーを含む。

【0012】

[0021] 将来の M P E G 符号化器への入力は、オプションで、次の 3 つの可能なフォーマット、すなわち、（i）あらかじめ指定された位置でラウドスピーカーを通じて再生されることが意図される、（上で論じられたような）従来のチャンネルベースオーディオ、（i i）（情報の中でも）位置座標を含む関連付けられたメタデータを有する単一オーディオオブジェクトのための離散的なパルス符号変調（PCM）データを伴うオブジェクトベースオーディオ、および（i i i）球面調和基底関数の係数（「球面調和係数」すなわち S H C、「高次アンビソニックス」すなわち H O A、および「H O A 係数」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オーディオのうちの 1 つである。将来の M P E G 符号化器は、2013 年 1 月にスイスのジュネーブで発表された、<http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/wl3411.zip> において入手可能な、International Organization for Standardization / International Electrotechnical Commission (ISO) / (IEC) JTC1 / SC29 / WG11 / N13411 による「Call for Proposals for 3D Audio」と題される文書においてより詳細に説明され得る。

【0013】

[0022] 市場には様々な「サラウンドサウンド」チャンネルベースフォーマットがある。これらのフォーマットは、たとえば、5.1 ホームシアターシステム（リビングルームに進出するという点でステレオ以上に最も成功した）から NHK (Nippon Hoso Kyokai すなわち日本放送協会) によって開発された 22.2 システムに及ぶ。コンテンツ作成者（たとえば、ハリウッドスタジオ）は、一度に映画のサウンドトラックを作

10

20

30

40

50

成することを望み、各々のスピーカー構成のためにサウンドトラックをリミックスする努力を行うことを望まない。最近では、規格開発組織が、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（と数）および（レンダラを伴う）再生のロケーションにおける音響条件に適応可能でありそれらに依存しない後続の復号とを提供するための方法を考えている。

【0014】

[0023]コンテンツ作成者にそのような柔軟性を提供するために、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。セットがより高次の要素を含むように拡張されると、表現はより詳細なものになり、分解能は向上する。

10

【0015】

[0024]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を使用する音場の記述または表現を示す。

【0016】

【数1】

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

20

【0017】

[0025]この式は、時間 t における音場の任意の点 $\{r_r, \theta_r, \varphi_r\}$ における圧力 p_i が、SHC、

【0018】

【数2】

$$A_n^m(k)$$

30

【0019】

によって一意に表され得ることを示す。ここで、

【0020】

【数3】

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

【0021】

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

40

【0022】

【数4】

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

【0023】

は次数 n および副次数 m の球面調和基底関数である。角括弧内の項は、離散フーリエ変換

50

(DFT)、離散コサイン変換(DCT)、またはウェーブレット変換などの様々な時間周波数変換によって近似され得る信号の周波数領域表現(すなわち、 $S(\omega, r_r, \theta_r, \phi_r)$)であることが認識できよう。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の係数の他のセットを含む。

【0024】

[0026] 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。理解できるように、各次数に対して、説明を簡単にするために図示されているが図1の例では明示的に示されていない副次数 m の拡張が存在する。

【0025】

[0027] SHC

10

【0026】

【数5】

$$A_n^m(k)$$

【0027】

は、様々なマイクロフォンアレイ構成によって物理的に獲得(たとえば、録音)されてよく、または代替的に、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てよい。SHCはシーンベースのオーディオを表し、ここで、SHCは、より効率的な送信または記憶を促し得る符号化されたSHCを取得するために、オーディオ符号化器に入力され得る。たとえば、 $(1+4)^2$ 個の(25個の、したがって4次の)係数を伴う4次表現が使用され得る。

20

【0028】

[0028] 上述されたように、SHCは、マイクロフォンアレイを使用したマイクロフォン録音から導出され得る。SHCがマイクロフォンアレイからどのように導出され得るかの様々な例は、Poletti, M、「Three-Dimensional Surround Sound Systems Based on Spherical Harmonics」、J. Audio Eng. Soc., Vol. 53, No. 11、2005年11月、1004~1025ページにおいて説明されている。

30

【0029】

[0029] SHCがどのようにオブジェクトベースの記述から導出され得るかを例示するために、次の式を考える。個々のオーディオオブジェクトに対応する音場についての係数

【0030】

【数6】

$$A_n^m(k)$$

【0031】

は、

40

【0032】

【数7】

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

【0033】

と表され得、ただし、 i は

【0034】

50

【数 8】

$$\sqrt{-1}$$

【0 0 3 5】

であり、

【0 0 3 6】

【数 9】

10

$$h_n^{(2)}$$

【0 0 3 7】

は次数 n の（第 2 の種類の）球ハンケル関数であり、 $\{r_s, \theta_s, \phi_s\}$ 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）オブジェクトソースエネルギー $g(\omega)$ を知ることで、各 PCM オブジェクトと対応するロケーションとを SHC

【0 0 3 8】

【数 1 0】

20

$$A_n^m(k)$$

【0 0 3 9】

に変換することが可能となる。さらに、各オブジェクトについての

【0 0 4 0】

【数 1 1】

30

$$A_n^m(k)$$

【0 0 4 1】

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

【0 0 4 2】

【数 1 2】

40

$$A_n^m(k)$$

【0 0 4 3】

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、音場についての情報（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。残りの数字は、以下でオブジェクトベースのオーディオコーディングおよび SHC ベースのオーディオコーディングの文脈で説明される。

【0 0 4 4】

50

[0030]図2は、本開示で説明される技法の様々な態様を実行することができるシステム10を示す図である。図2の例に示されるように、システム10は、コンテンツ作成者デバイス12と、コンテンツ消費者デバイス14とを含む。コンテンツ作成者デバイス12およびコンテンツ消費者デバイス14の文脈で説明されているが、本技法は、オーディオデータを表すビットストリームを形成するために、SHC（HOA係数とも呼ばれ得る）または音場の任意の他の階層的表現が符号化される任意の文脈で実施され得る。その上、コンテンツ作成者デバイス12は、いくつか例を挙げると、ハンドセット（またはセルラーフォン）、タブレットコンピュータ、スマートフォン、またはデスクトップコンピュータを含む、本開示で説明される技法を実施することが可能な任意の形態のコンピューティングデバイスを表し得る。同様に、コンテンツ消費者デバイス14は、いくつか例を挙げると、ハンドセット（またはセルラーフォン）、タブレットコンピュータ、スマートフォン、セットトップボックス、またはデスクトップコンピュータを含む、本開示で説明される技法を実施することが可能な任意の形態のコンピューティングデバイスを表し得る。

10

【0045】

[0031]コンテンツ作成者デバイス12は、コンテンツ消費者デバイス14などのコンテンツ消費者のオペレータによる消費のためのマルチチャネルオーディオコンテンツを生成することができる、映画スタジオまたは他のエンティティによって操作され得る。いくつかの例では、コンテンツ作成者デバイス12は、HOA係数11を圧縮することを望む個人ユーザによって操作され得る。多くの場合、コンテンツ作成者は、ビデオコンテンツとともに、オーディオコンテンツを生成する。コンテンツ消費者デバイス14は、個人によって操作され得る。コンテンツ消費者デバイス14は、マルチチャネルオーディオコンテンツとしての再生のためにSHCをレンダリングすることが可能な任意の形態のオーディオ再生システムを指し得る、オーディオ再生システム16を含み得る。

20

【0046】

[0032]コンテンツ作成者デバイス12は、オーディオ編集システム18を含む。コンテンツ作成者デバイス12は、様々なフォーマットのライブ録音7（HOA係数として直接含む）とオーディオオブジェクト9とを取得し、コンテンツ作成者デバイス12は、オーディオ編集システム18を使用してこれらを編集することができる。コンテンツ作成者は、編集プロセスの間に、オーディオオブジェクト9からのHOA係数11をレンダリングし、さらなる編集を必要とする音場の様々な態様を特定しようとして、レンダリングされたスピーカーフィールドを聞くことができる。コンテンツ作成者デバイス12は次いで、（潜在的に、上記で説明された方法でソースHOA係数がそれから導出され得るオーディオオブジェクト9のうちの様々なオブジェクトの操作を通じて間接的に）HOA係数11を編集することができる。コンテンツ作成者デバイス12は、HOA係数11を生成するためにオーディオ編集システム18を採用することができる。オーディオ編集システム18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。

30

【0047】

[0033]編集プロセスが完了すると、コンテンツ作成者デバイス12は、HOA係数11に基づいてビットストリーム21を生成することができる。すなわち、コンテンツ作成者デバイス12は、ビットストリーム21を生成するために、本開示で説明される技法の様々な態様に従って、HOA係数11を符号化またはさもなければ圧縮するように構成されたデバイスを表す、オーディオ符号化デバイス20を含む。オーディオ符号化デバイス20は、一例として、有線チャネルまたはワイヤレスチャネル、データ記憶デバイスなどであり得る送信チャネルを介した送信のために、ビットストリーム21を生成することができる。ビットストリーム21は、HOA係数11の符号化されたバージョンを表すことができ、主要ビットストリームと、サイドチャネル情報とも呼ばれ得る別のサイドビットストリームとを含み得る。

40

【0048】

[0034]以下でより詳細に説明されるが、オーディオ符号化デバイス20は、ベクトルベ

50

ース合成または指向性ベース合成に基づいてH O A 係数 1 1 を符号化するように構成され得る。ベクトルベース分解方法を実行するか指向性ベース分解方法を実行するかを決定するために、オーディオ符号化デバイス 2 0 は、H O A 係数 1 1 に少なくとも部分的に基づいて、H O A 係数 1 1 が音場の自然な録音（たとえば、ライブ録音 7）を介して生成されたか、または一例として、P C Mオブジェクトなどのオーディオオブジェクト 9 から人工的に（すなわち、合成的に）作成されたかを決定することができる。H O A 係数 1 1 がオーディオオブジェクト 9 から生成されたとき、オーディオ符号化デバイス 2 0 は、指向性ベース分解方法を使用してH O A 係数 1 1 を符号化することができる。H O A 係数 1 1 が、たとえば、e i g e n m i k eを使用してライブでキャプチャされたとき、オーディオ符号化デバイス 2 0 は、ベクトルベース分解方法に基づいてH O A 係数 1 1 を符号化することができる。上の区別は、ベクトルベース分解方法または指向性ベース分解方法が採用され得る一例を表す。これらの一方または両方が、自然な録音、人工的に生成されたコンテンツ、またはこれら 2 つの混合物（ハイブリッドコンテンツ）に対して有用であり得る、他の場合があり得る。さらに、H O A 係数の単一の時間フレームをコーディングするために両方の方法を同時に使用することも可能である。

【0049】

[0035]説明のために、H O A 係数 1 1 がライブでキャプチャされたか、またはさもなければライブ録音 7 などのライブ録音を表すと、オーディオ符号化デバイス 2 0 が決定すると仮定すると、オーディオ符号化デバイス 2 0 は、線形可逆変換（L I T）の適用を伴うベクトルベース分解方法を使用してH O A 係数 1 1 を符号化するように構成され得る。線形可逆変換の一例は、「特異値分解」（または「S V D」）と呼ばれる。この例では、オーディオ符号化デバイス 2 0 は、H O A 係数 1 1 の分解されたバージョンを決定するためにS V DをH O A 係数 1 1 に適用することができる。オーディオ符号化デバイス 2 0 は次いで、様々なパラメータを特定するためにH O A 係数 1 1 の分解されたバージョンを分析することができる、このことは、H O A 係数 1 1 の分解されたバージョンの並べ替えを容易にし得る。オーディオ符号化デバイス 2 0 は次いで、特定されたパラメータに基づいてH O A 係数 1 1 の分解されたバージョンを並べ替えることができ、そのような並べ替えは、以下でさらに詳細に説明されるように、変換がH O A 係数のフレームにわたってH O A 係数を並べ替えることができるとすると（フレームが、H O A 係数 1 1 のM個のサンプルを含み得、Mが、いくつかの例では1024に設定される場合）、コーディング効率を向上させることができる。H O A 係数 1 1 の分解されたバージョンを並べ替えた後、オーディオ符号化デバイス 2 0 は、音場のフォアグラウンド（または言い換えれば、明瞭な、支配的な、もしくは目立つ）成分を表す、H O A 係数 1 1 の分解されたバージョンを選択することができる。オーディオ符号化デバイス 2 0 は、フォアグラウンド成分を表すH O A 係数 1 1 の分解されたバージョンを、オーディオオブジェクトおよび関連付けられる指向性情報として指定することができる。

【0050】

[0036]オーディオ符号化デバイス 2 0 はまた、少なくとも部分的には、音場の1つまたは複数のバックグラウンド（または言い換えれば、環境的な）成分を表すH O A 係数 1 1 を特定するために、H O A 係数 1 1 に関して音場分析を実行することができる。オーディオ符号化デバイス 2 0 は、いくつかの例では、バックグラウンド成分がH O A 係数 1 1 の任意の所与のサンプルのサブセット（たとえば、2 次以上の球面基底関数に対応するH O A 係数 1 1 ではなく、0 次および1 次の球面基底関数に対応するH O A 係数 1 1 など）のみを含み得るとすると、バックグラウンド成分に関してエネルギー補償を実行することができる。言い換えれば、次数低減が実行されるとき、オーディオ符号化デバイス 2 0 は、次数低減を実行したことに起因する全体的なエネルギーの変化を補償するために、H O A 係数 1 1 の残りのバックグラウンドH O A 係数を補強する（たとえば、それにエネルギーを加える／それからエネルギーを差し引く）ことができる。

【0051】

[0037]オーディオ符号化デバイス 2 0 は次に、バックグラウンド成分とフォアグラウン

ドオーディオオブジェクトの各々とを表すH O A係数1 1の各々に関して、ある形態の聴覚心理符号化（M P E Gサラウンド、M P E G－A A C、M P E G－U S A C、または他の既知の形態の聴覚心理符号化など）を実行することができる。オーディオ符号化デバイス2 0は、フォアグラウンド指向性情報に関してある形態の補間を実行し、次いで、次数低減されたフォアグラウンド指向性情報を生成するために、補間されたフォアグラウンド指向性情報に関して次数低減を実行することができる。オーディオ符号化デバイス2 0はさらに、いくつかの例では、次数低減されたフォアグラウンド指向性情報に関して量子化を実行し、コーディングされたフォアグラウンド指向性情報を出力することができる。いくつかの場合には、量子化はスカラー／エントロピー量子化を備え得る。オーディオ符号化デバイス2 0は次いで、符号化されたバックグラウンド成分と、符号化されたフォアグラウンドオーディオオブジェクトと、量子化された指向性情報とを含むように、ビットストリーム2 1を形成することができる。オーディオ符号化デバイス2 0は次いで、ビットストリーム2 1をコンテンツ消費者デバイス1 4に送信またはさもなければ出力することができる。

10

【0052】

[0038]図2では、コンテンツ消費者デバイス1 4に直接的に送信されるものとして示されているが、コンテンツ作成者デバイス1 2は、コンテンツ作成者デバイス1 2とコンテンツ消費者デバイス1 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2 1を出力することができる。中間デバイスは、ビットストリームを要求し得るコンテンツ消費者デバイス1 4に後で配信するために、ビットストリーム2 1を記憶することができる。中間デバイスは、ファイルサーバ、ウェブサーバ、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、モバイルフォン、スマートフォン、または後でのオーディオ復号器による取出しのためにビットストリーム2 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。中間デバイスは、ビットストリーム2 1を要求するコンテンツ消費者デバイス1 4などの加入者にビットストリーム2 1を（場合によっては対応するビデオデータビットストリームを送信するとともに）ストリーミングすることが可能なコンテンツ配信ネットワーク内に存在してもよい。

20

【0053】

[0039]代替的に、コンテンツ作成者デバイス1 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2 1を記憶することができ、記憶媒体の大部分はコンピュータによって読み取り可能であり、したがって、コンピュータ可読記憶媒体または非一時的コンピュータ可読記憶媒体と呼ばれることがある。この文脈において、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すことがある（および、小売店と他の店舗ベースの配信機構とを含み得る）。したがって、いずれにしても、本開示の技法は、この点に関して図2の例に限定されるべきではない。

30

【0054】

[0040]図2の例にさらに示されるように、コンテンツ消費者デバイス1 4は、オーディオ再生システム1 6を含む。オーディオ再生システム1 6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1 6は、いくつかの異なるレンダラ2 2を含み得る。レンダラ2 2は各々、異なる形態のレンダリングを提供することができ、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1つもしくは複数、および／または音場合成を実行する様々な方法の1つもしくは複数を含み得る。本明細書で使用される場合、「Aおよび／またはB」は、「AまたはB」、または「AとB」の両方を意味する。

40

【0055】

[0041]オーディオ再生システム1 6は、オーディオ復号デバイス2 4をさらに含み得る。オーディオ復号デバイス2 4は、ビットストリーム2 1からH O A係数1 1'を復号するように構成されたデバイスを表し得、H O A係数1 1'は、H O A係数1 1と類似し得

50

るが、損失のある演算（たとえば、量子化）および／または送信チャネルを介した送信が原因で異なり得る。すなわち、オーディオ復号デバイス 24 は、ビットストリーム 21 において指定されるフォアグラウンド指向性情報を逆量子化することができ、一方でまた、ビットストリーム 21 において指定されるフォアグラウンドオーディオオブジェクトおよびバックグラウンド成分を表す符号化された H O A 係数に関して聴覚心理復号を実行することができる。オーディオ復号デバイス 24 はさらに、復号されたフォアグラウンド指向性情報に関して補間を実行し、次いで、復号されたフォアグラウンドオーディオオブジェクトおよび補間されたフォアグラウンド指向性情報に基づいて、フォアグラウンド成分を表す H O A 係数を決定することができる。オーディオ復号デバイス 24 は次いで、フォアグラウンド成分を表す決定された H O A 係数およびバックグラウンド成分を表す復号された H O A 係数に基づいて、H O A 係数 1 1 ' を決定することができる。

10

【0056】

[0042] オーディオ再生システム 16 は、H O A 係数 1 1 ' を取得するためにビットストリーム 21 を復号した後、ラウドスピーカーフィールド 25 を出力するために H O A 係数 1 1 ' をレンダリングすることができる。ラウドスピーカーフィールド 25 は、1 つまたは複数のラウドスピーカー（説明を簡単にするために図 2 の例には示されていない）を駆動することができる。

【0057】

[0043] 適切なレンダラを選択するために、またはいくつかの場合には、適切なレンダラを生成するために、オーディオ再生システム 16 は、ラウドスピーカーの数および／またはラウドスピーカーの空間的な幾何学的配置を示すラウドスピーカー情報 13 を取得することができる。いくつかの場合には、オーディオ再生システム 16 は、基準マイクロフォンを使用してラウドスピーカー情報 13 を取得し、ラウドスピーカー情報 13 を動的に決定するような方法でラウドスピーカーを駆動することができる。他の場合には、またはラウドスピーカー情報 13 の動的な決定とともに、オーディオ再生システム 16 は、オーディオ再生システム 16 とインターフェースをとりラウドスピーカー情報 13 を入力するようにユーザに促すことができる。

20

【0058】

[0044] オーディオ再生システム 16 は次いで、ラウドスピーカー情報 13 に基づいて、オーディオレンダラ 22 のうちの 1 つを選択することができる。いくつかの場合には、オーディオ再生システム 16 は、オーディオレンダラ 22 のいずれもがラウドスピーカー情報 13 において指定されたものに対して（ラウドスピーカーの幾何学的配置に関する）何らかの類似性の尺度のしきい値内にないとき、ラウドスピーカー情報 13 に基づいて、オーディオレンダラ 22 のうちの 1 つを生成することができる。オーディオ再生システム 16 は、いくつかの場合には、オーディオレンダラ 22 のうちの既存の 1 つを選択することを最初に試みることなく、ラウドスピーカー情報 13 に基づいて、オーディオレンダラ 22 のうちの 1 つを生成することができる。

30

【0059】

[0045] 図 3 は、本開示で説明される技法の様々な態様を実行することができる、図 2 の例に示されるオーディオ符号化デバイス 20 の一例をより詳細に示すブロック図である。オーディオ符号化デバイス 20 は、コンテンツ分析ユニット 26 と、ベクトルベース分解ユニット 27 と、指向性ベース分解ユニット 28 とを含む。以下で手短かに説明されるが、オーディオ符号化デバイス 20 に関するより多くの情報、および H O A 係数を圧縮またはさもなければ符号化する様々な態様は、2014 年 5 月 29 に出願された「INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の国際特許出願公開第 WO 2014 / 194099 号において入手可能である。

40

【0060】

[0046] コンテンツ分析ユニット 26 は、H O A 係数 1 1 がライブ録音から生成されたコンテンツを表すか、オーディオオブジェクトから生成されたコンテンツを表すかを特定す

50

るために、H O A 係数 1 1 のコンテンツを分析するように構成されたユニットを表す。コンテンツ分析ユニット 2 6 は、H O A 係数 1 1 が実際の音場の録音から生成されたか人工的なオーディオオブジェクトから生成されたかを決定することができる。いくつかの場合には、フレーム化された H O A 係数 1 1 が録音から生成されたとき、コンテンツ分析ユニット 2 6 は、H O A 係数 1 1 をベクトルベース分解ユニット 2 7 に渡す。いくつかの場合には、フレーム化された H O A 係数 1 1 が合成オーディオオブジェクトから生成されたとき、コンテンツ分析ユニット 2 6 は、H O A 係数 1 1 を指向性ベース合成ユニット 2 8 に渡す。指向性ベース合成ユニット 2 8 は、指向性ベースビットストリーム 2 1 を生成するために H O A 係数 1 1 の指向性ベース合成を実行するように構成されたユニットを表し得る。

10

【0061】

[0047] 図 3 の例に示されるように、ベクトルベース分解ユニット 2 7 は、線形可逆変換 (L I T) ユニット 3 0 と、パラメータ計算ユニット 3 2 と、並べ替えユニット 3 4 と、フォアグラウンド選択ユニット 3 6 と、エネルギー補償ユニット 3 8 と、聴覚心理オーディオコーダユニット 4 0 と、ビットストリーム生成ユニット 4 2 と、音場分析ユニット 4 4 と、係数低減ユニット 4 6 と、バックグラウンド (B G) 選択ユニット 4 8 と、空間時間的補間ユニット 5 0 と、量子化ユニット 5 2 とを含み得る。

【0062】

[0048] 線形可逆変換 (L I T) ユニット 3 0 は、H O A チャンネルの形態で H O A 係数 1 1 を受信し、各チャンネルは、球面基底関数の所与の次数、副次数に関連付けられた係数のブロックまたはフレーム (H O A [k]) と示され得、ただし、k はサンプルの現在のフレームまたはブロックを示し得る) を表す。H O A 係数 1 1 の行列は、次元 $D : M \times (N + 1)^2$ を有し得る。

20

【0063】

[0049] すなわち、L I T ユニット 3 0 は、特異値分解と呼ばれるある形態の分析を実行するように構成されたユニットを表し得る。S V D に関して説明されているが、本開示で説明される技法は、線形的に無相関な、エネルギーが圧縮された出力のセットを提供する任意の類似の変換または分解に対して実行されてよい。また、本開示における「セット」への言及は、一般的に、それとは反対に特に明記されていない限り、非 0 のセットを指すことが意図され、いわゆる「空集合」を含む集合の古典的な数学的定義を指すことは意図されない。

30

【0064】

[0050] 代替的な変換は、「P C A」と呼ばれることが多い、主成分分析を備え得る。P C A は、おそらく相関する変数の観測値のセットを、主成分と呼ばれる線形的に無相関な変数のセットに変換するために、直交変換を採用する数学的手順を指す。線形的に無相関な変数とは、互いに対する統計的線形関係 (すなわち依存) を持たない変数を表す。主成分は、互いに対するわずかな統計的相関を有するものとして説明され得る。いずれにしても、いわゆる主成分の数は、元の変数の数以下である。いくつかの例では、変換は、第 1 の主成分が可能な最大の分散を有し (または、言い換えれば、データの変動性をできる限り多く考慮し)、後続の各成分が、連続した成分が先行する成分と直交する (これと無相関と言い換えられ得る) という制約の下で可能な最高分散を有するような方法で、定義される。P C A は、H O A 係数 1 1 に関して H O A 係数 1 1 の圧縮をもたらし得る、ある形態の次数低減を実行することができる。文脈に応じて、P C A は、いくつかの例を挙げれば、離散カルーネンレーベ変換、ホテリング変換、固有直交分解 (P O D)、および固有値分解 (E V D) などのいくつかの異なる名前によって呼ばれることがある。オーディオデータを圧縮するという背後にある目標につながるそのような演算の特性は、マルチチャンネルオーディオデータの「エネルギー圧縮」および「無相関化」である。

40

【0065】

[0051] いずれにしても、L I T ユニット 3 0 が、例として、特異値分解 (やはり「S V D」と呼ばれることがある) を実行すると仮定すると、L I T ユニット 3 0 は、H O A 係

50

数 1 1 を、変換された H O A 係数の 2 つ以上のセットに変換することができる。変換された H O A 係数の「セット」は、変換された H O A 係数のベクトルを含み得る。図 3 の例では、L I T ユニット 3 0 は、いわゆる V 行列と、S 行列と、U 行列とを生成するために、H O A 係数 1 1 に関して S V D を実行することができる。S V D は、線形代数学では、 $y \times z$ の実行列または複素行列 X （ここで、 X は、H O A 係数 1 1 などのマルチチャネルオーディオデータを表し得る）の因数分解を以下の形で表し得る。 $X = U S V^* U$ は $y \times y$ の実ユニタリー行列または複素ユニタリー行列を表し得、ここで、 U の y 個の列は、マルチチャネルオーディオデータの左特異ベクトルとして知られる。 S は、対角線上に非負実数をもつ $y \times z$ の矩形対角行列を表し得、ここで、 S の対角線値は、マルチチャネルオーディオデータの特異値として知られる。 V^* （ V の共役転置を示し得る）は $z \times z$ の実ユニタリー行列または複素ユニタリー行列を表し得、ここで、 V^* の z 個の列は、マルチチャネルオーディオデータの右特異ベクトルとして知られる。

【0066】

[0052] 本開示では、H O A 係数 1 1 を備えるマルチチャネルオーディオデータに適用されるものとして説明されているが、本技法は、任意の形態のマルチチャネルオーディオデータに適用されてよい。このようにして、オーディオ符号化デバイス 2 0 は、マルチチャネルオーディオデータの左特異ベクトルを表す U 行列と、マルチチャネルオーディオデータの特異値を表す S 行列と、マルチチャネルオーディオデータの右特異ベクトルを表す V 行列とを生成するために、音場の少なくとも一部分を表すマルチチャネルオーディオデータに対して特異値分解を実行し、マルチチャネルオーディオデータを U 行列、 S 行列、および V 行列のうちの 1 つまたは複数の少なくとも一部分の関数として表すことができる。

【0067】

[0053] いくつかの例では、上で参照された S V D 数式中の V^* 行列は、複素数を備える行列に S V D が適用され得ることを反映するために、 V 行列の共役転置として示される。実数のみを備える行列に適用されるとき、 V 行列の複素共役（すなわち、言い換えれば、 V^* 行列）は、 V 行列の転置であると思なされてよい。以下では、説明を簡単にするために、H O A 係数 1 1 が実数を備え、その結果、 V^* 行列ではなく V 行列が S V D によって出力されると仮定される。その上、本開示では V 行列として示されるが、 V 行列への言及は、適切な場合には V 行列の転置を指すものとして理解されるべきである。 V 行列であると仮定されているが、本技法は、同様の方式で、複素係数を有する H O A 係数 1 1 に適用されてよく、ここで、S V D の出力は V^* 行列である。したがって、本技法は、この点について、 V 行列を生成するために S V D の適用を提供することのみに限定されるべきではなく、 V^* 行列を生成するために複素成分を有する H O A 係数 1 1 への S V D の適用を含んでよい。

【0068】

[0054] いずれにしても、L I T ユニット 3 0 は、高次アンビソニックス（H O A）オーディオデータの各ブロック（フレームを指し得る）に関して、ブロックごとの形態の S V D を実行することができる（ここで、アンビソニックスオーディオデータは、H O A 係数 1 1 のブロックもしくはサンプル、または任意の他の形態のマルチチャネルオーディオデータを含む）。上述されたように、変数 M は、サンプル中のオーディオフレームの長さを示すために使用され得る。たとえば、オーディオフレームが 1 0 2 4 個のオーディオサンプルを含むとき、 M は 1 0 2 4 に等しい。 M の典型的な値に関して説明されるが、本開示の技法は、 M の典型的な値に限定されるべきではない。L I T ユニット 3 0 はしたがって、 $M \times (N + 1)^2$ の H O A 係数を有するブロック H O A 係数 1 1 に関してブロックごとの S V D を実行することができ、ここで、 N はやはり H O A オーディオデータの次数を示す。L I T ユニット 3 0 は、S V D の実行を通して、 V 行列と、 S 行列と、 U 行列とを生成することができ、ここで、行列の各々は、上記で説明されたそれぞれの V 行列と、 S 行列と、 U 行列とを表すことができる。このようにして、線形可逆変換ユニット 3 0 は、次元 $D : M \times (N + 1)^2$ を有する $U S [k]$ ベクトル 3 3（ S ベクトルと U ベクトルとの組み合わせられたバージョンを表し得る）と、次元 $D : (N + 1)^2 \times (N + 1)^2$ を有する

V[k] ベクトル 35 とを出力するために、H O A 係数 1 1 に関して S V D を実行することができる。U S[k] 行列中の個々のベクトル要素は $X_{ps}(k)$ とも呼ばれることがあり、一方、V[k] 行列の個々のベクトルは $v(k)$ とも呼ばれることがある。

【0069】

[0055] U 行列、S 行列、および V 行列の分析は、それらの行列が X によって上で表される背後の音場の空間的および時間的な特性を伝え、または表すということを明らかにし得る。(M 個のサンプルの長さの) U 中の N 個のベクトルの各々は、(M 個のサンプルによって表される時間期間の間は) 時間の関数として、互いに直交しておりあらゆる空間特性(指向性情報とも呼ばれ得る)とは切り離されている、正規化された分離されたオーディオ信号を表し得る。空間的な形状と位置(r 、 θ 、 ϕ)の幅とを表す空間特性は代わりに、(各々が $(N+1)^2$ の長さの) V 行列の中の個々の i 番目のベクトル、 $v^{(i)}(k)$ によって表され得る。 $v^{(i)}(k)$ ベクトルの各々の個々の要素は、関連付けられたオーディオオブジェクトのための音場の形状と方向とを記述する H O A 係数を表し得る。U 行列中のベクトルと V 行列中のベクトルの両方が、それらの 2 乗平均エネルギーが 1 に等しくなるように正規化される。したがって、U 中のオーディオ信号のエネルギーは、S 中の対角線要素によって表される。したがって、U S[k] (個々のベクトル要素 $X_{ps}(k)$ を有する)を形成するために、U と S とを乗算することは、真のエネルギーを有するオーディオ信号を表す。(U における)オーディオ時間信号と、(S における)それらのエネルギーと、(V における)それらの空間特性とを切り離す S V D 分解の能力は、本開示で説明される技法の様々な態様を支援することができる。さらに、背後の H O A[k] 係数 X を U S[k] と V[k] とのベクトル乗算によって合成するモデルは、本文書全体で使用される、「ベクトルベース分解」という用語を生じさせる。

【0070】

[0056] H O A 係数 1 1 に関して直接実行されるものとして説明されるが、L I T ユニット 30 は、線形可逆変換を H O A 係数 1 1 の派生物に適用することができる。たとえば、L I T ユニット 30 は、H O A 係数 1 1 から導出された電力スペクトル密度行列に関して S V D を適用することができる。電力スペクトル密度行列は、P S D として示され、以下に続く擬似コードにおいて概説されるように、`hoaFrame` の転置の `hoaFrame` への行列乗算を通じて取得され得る。`hoaFrame` という表記は、H O A 係数 1 1 のフレームを指す。

【0071】

[0057] L I T ユニット 30 は、S V D (`svd`) を P S D に適用した後、S[k]² 行列 (`S_squared`) と V[k] 行列とを取得することができる。S[k]² 行列は、二乗された S[k] 行列を示すことができ、すると、L I T ユニット 30 は、S[k] 行列を取得するために平方根演算を S[k]² 行列に適用することができる。L I T ユニット 30 は、いくつかの場合には、量子化された V[k] 行列 ($V[k]'$ 行列と示され得る)を取得するために、V[k] 行列に関して量子化を実行することができる。L I T ユニット 30 は、S V[k]' 行列を取得するために、S[k] 行列を量子化された $V[k]'$ 行列と最初に乗算することによって、U[k] 行列を取得することができる。L I T ユニット 30 は次に、S V[k]' 行列の擬似逆行列 (`pinv`)を取得することができる。次に、U[k] 行列を取得するために、H O A 係数 1 1 を S V[k]' 行列の擬似逆行列と乗算することができる。上記は、以下の擬似コードによって表され得る。

```
PSD = hoaFrame' * hoaFrame;
[V, S_squared] = svd(PSD, 'econ');
S = sqrt(S_squared);
U = hoaFrame * pinv(S*V');
```

[0058] H O A 係数自体ではなく H O A 係数の電力スペクトル密度 (P S D) に関して S V D を実行することによって、L I T ユニット 30 は潜在的に、プロセッササイクルおよび記憶空間のうちの 1 つまたは複数に関して S V D を実行することの計算的な複雑さを低減しつつ、S V D が H O A 係数に直接適用されたかのように同じソースオーディオ符号化

効率を達成することができる。すなわち、上記で説明されたPSDタイプのSVDは、Mがフレーム長さ、すなわち、1024以上のサンプルである、 $M \times F$ 行列と比較して、 $F \times F$ 行列（HOA係数の数Fをもつ）において完了するので、潜在的にそれほど計算的に厳しくないことがある。ここで、SVDの複雑さは、HOA係数11ではなくPSDへの適用を通じて、HOA係数11に適用されたときの $O(M \times L^2)$ と比較して、 $O(L^3)$ 前後であり得る（ここで、 $O(*)$ は、コンピュータサイエンス技術において一般的な計算の複雑さである大文字Oの表記を示す）。

【0072】

[0059]この点において、LITユニット30は、球面調和関数領域における直交空間軸を表すベクトル（例えば、上のV-ベクトル）を取得するために、高次アンビソニックオーディオデータに関して分解を実行するまたはさなければ高次アンビソニックオーディオデータを分解し得る。分解は、SVD、EVD、または分解の任意の他の形式を含み得る。

10

【0073】

[0060]パラメータ計算ユニット32は、相関パラメータ（R）、指向性特性パラメータ（ θ 、 ϕ 、r）、およびエネルギー特性（e）などの様々なパラメータを計算するように構成されたユニットを表す。現在のフレームのためのパラメータの各々は、 $R[k]$ 、 $\theta[k]$ 、 $\phi[k]$ 、 $r[k]$ 、および $e[k]$ として示され得る。パラメータ計算ユニット32は、パラメータを特定するために、 $US[k]$ ベクトル33に関してエネルギー分析および／または相関（もしくはいわゆる相互相関）を実行することができる。パラメータ計算ユニット32はまた、以前のフレームのためのパラメータを決定することができ、ここで、以前のフレームパラメータは、 $US[k-1]$ ベクトルおよび $V[k-1]$ ベクトルの以前のフレームに基づいて、 $R[k-1]$ 、 $\theta[k-1]$ 、 $\phi[k-1]$ 、 $r[k-1]$ 、および $e[k-1]$ として示され得る。パラメータ計算ユニット32は、現在のパラメータ37と以前のパラメータ39とを並べ替えユニット34に出力することができる。

20

【0074】

[0061]SVD分解は、 $US[k-1][p]$ ベクトル（または代替的に、 $X_{ps}^{(p)}(k-1)$ ）として示され得る、 $US[k-1]$ ベクトル33のp番目のベクトルによって表されるオーディオ信号／オブジェクトが、同じく $US[k][p]$ ベクトル33（または代替的に、 $X_{ps}^{(p)}(k)$ ）として示され得る、 $US[k]$ ベクトル33のp番目のベクトルによって表される（時間的に進んだ）同じオーディオ信号／オブジェクトとなることを保証しない。パラメータ計算ユニット32によって計算されるパラメータは、オーディオオブジェクトの自然な評価または時間的な継続性を表すようにオーディオオブジェクトを並べ替えるために、並べ替えユニット34によって使用され得る。

30

【0075】

[0062]すなわち、並べ替えユニット34は、第1の $US[k]$ ベクトル33からのパラメータ37の各々を、第2の $US[k-1]$ ベクトル33のためのパラメータ39の各々に対して順番ごとに比較することができる。並べ替えユニット34は、並べ替えられた $US[k]$ 行列33'（数学的には

【0076】

40

【数13】

$$\overline{US[k]}$$

【0077】

として示され得る）と、並べ替えられた $V[k]$ 行列35'（数学的には

【0078】

【数 1 4】

$$\bar{V}[k]$$

【0079】

として示され得る)とをフォアグラウンドサウンド(または支配的サウンドーPS(predominant sound))選択ユニット36(「フォアグラウンド選択ユニット36」)およびエネルギー補償ユニット38に出力するために、現在のパラメータ37および以前のパラメータ39に基づいて、US[k]行列33およびV[k]行列35内の様々なベクトルを(一例として、ハンガリー法を使用して)並べ替えることができる。

【0080】

[0063]音場分析ユニット44は、目標ビットレート41を潜在的に達成するために、HOA係数11に関して音場分析を実行するように構成されたユニットを表し得る。音場分析ユニット44は、その分析および/または受信された目標ビットレート41に基づいて、聴覚心理コーダのインスタンス化の総数(環境またはバックグラウンドチャンネルの総数(BG_{TOT})とフォアグラウンドチャンネル、または言い換えれば支配的なチャンネルの数との関数であり得る、を決定することができる。聴覚心理コーダのインスタンス化の総数は、numHOATransportChannelsとして示され得る。

【0081】

[0064]音場分析ユニット44はまた、やはり目標ビットレート41を潜在的に達成するために、フォアグラウンドチャンネルの総数(nFG)45と、バックグラウンド(または言い換えれば環境的な)音場の最小次数(N_{BG}、または代替的にはMinAmbHOAorder)と、バックグラウンド音場の最小次数を表す実際のチャンネルの対応する数(nBGa=(MinAmbHOAorder+1)²)と、送るべき追加のBG HOAチャンネルのインデックス(i)(図3の例ではバックグラウンドチャンネル情報43として総称的に示され得る)とを決定することができる。バックグラウンドチャンネル情報42は、環境チャンネル情報43とも呼ばれ得る。numHOATransportChannels-nBGaで残るチャンネルの各々は、「追加のバックグラウンド/環境チャンネル」、「アクティブなベクトルベースの支配的なチャンネル」、「アクティブな指向性ベースの支配的な信号」、または「完全に非アクティブ」のいずれかであり得る。一態様では、チャンネルタイプは、2ビットによって(「ChannelType」として)示されたシンタックス要素であり得る(たとえば、00:指向性ベースの信号、01:ベクトルベースの支配的な信号、10:追加の環境信号、11:非アクティブな信号)。バックグラウンド信号または環境信号の総数、nBGaは、(MinAmbHOAorder+1)²+ (上記の例における)インデックス10がそのフレームのためのビットストリームにおいてチャンネルタイプとして現れる回数によって与えられ得る。

【0082】

[0065]いずれにしても、音場分析ユニット44は、目標ビットレート41に基づいて、バックグラウンド(または言い換えれば環境)チャンネルの数とフォアグラウンド(または言い換えれば支配的な)チャンネルの数とを選択し、目標ビットレート41が比較的高いとき(たとえば、目標ビットレート41が512Kbps以上であるとき)はより多くのバックグラウンドチャンネルおよび/またはフォアグラウンドチャンネルを選択することができる。一態様では、ビットストリームのヘッダセクションにおいて、numHOATransportChannelsは8に設定され得るが、一方で、MinAmbHOAorderは1に設定され得る。このシナリオでは、各フレームにおいて、音場のバックグラウンド部分または環境部分を表すために4つのチャンネルが確保され得るが、一方で、他の4つのチャンネルは、フレームごとに、チャンネルのタイプに応じて変化してよく、たとえば、追加のバックグラウンド/環境チャンネルまたはフォアグラウンド/支配的なチャンネルのい

ずれかとして使用され得る。フォアグラウンド／支配的な信号は、上記で説明されたように、ベクトルベースの信号または指向性ベースの信号のいずれかの1つであり得る。

【0083】

[0066]いくつかの場合には、フレームのためのベクトルベースの支配的な信号の総数は、そのフレームのビットストリームにおいてChannel Type インデックスが01である回数によって与えられ得る。上記の態様では、各々の追加のバックグラウンド／環境チャンネル（たとえば、10というChannel Type に対応する）に対して、（最初の4つ以外の）あり得るHOA係数のいずれがそのチャンネルにおいて表され得るかの対応する情報。その情報は、4次のHOAコンテンツについては、HOA係数5～25を示すためのインデックスであり得る。最初の4つの環境HOA係数1～4は、minAmbHOAorderが1に設定されるときは常に送られ得、したがって、オーディオ符号化デバイスは、5～25のインデックスを有する追加の環境HOA係数のうちの1つを示すことのみが必要であり得る。その情報はしたがって、「CodedAmbCoeffIdx」として示され得る、（4次のコンテンツのための）5ビットのシンタックス要素を使用して送られ得る。

【0084】

[0067]説明のために、minAmbHOAorderが1に設定され、6のインデックスをもつ追加の環境HOA係数が、一例として、ビットストリーム21を介して送られると仮定する。この例では、1のminAmbHOAorderは、環境HOA係数が1、2、3および4のインデックスを有することを示す。環境HOA係数が、 $(\text{minAmbHOAorder} + 1)^2$ 、すなわち、この例では4以下であるインデックスを有するので、オーディオ符号化デバイス20は、それらの環境HOA係数を選択することができる。オーディオ符号化デバイス20は、ビットストリーム21において1、2、3および4のインデックスに関連付けられた環境HOA係数を指定することができる。オーディオ符号化デバイス20はまた、ビットストリーム21において6のインデックスをもつ追加の環境HOA係数を、10のChannel TypeをもつadditionalAmbientHOAchannelとして指定することもできる。オーディオ符号化デバイス20は、CodedAmbCoeffIdxシンタックス要素を使用して、インデックスを指定することができる。実際に、CodedAmbCoeffIdx要素は、1から25のインデックスのすべてを指定することができる。しかしながら、minAmbHOAorderが1に設定されるので、オーディオ符号化デバイス20は、（最初の4つのインデックスが、minAmbHOAorderシンタックス要素を介してビットストリーム21において指定されることが知られているので、）最初の4つのインデックスのいずれをも指定しなくてよい。いずれにしても、オーディオ符号化デバイス20は、（最初の4つについて）minAmbHOAorderと（追加の環境HOA係数について）CodedAmbCoeffIdxとを介して、5つの環境HOA係数を指定するので、オーディオ符号化デバイス20は、1、2、3、4および6のインデックスを有する環境HOA係数に関連付けられた対応するVベクトル要素を指定しなくてよい。結果として、オーディオ符号化デバイス20は、要素[5, 7:25]をもつVベクトルを指定することができる。

【0085】

[0068]第2の態様では、フォアグラウンド／支配的な信号のすべてがベクトルベースの信号である。この第2の態様では、フォアグラウンド／支配的な信号の総数は、 $\text{nFG} = \text{numHOATransportChannels} - [(\text{MinAmbHOAorder} + 1)^2 + \text{additionalAmbientHOAchannel}]$ の各々によって与えられ得る。

【0086】

[0069]音場分析ユニット44は、バックグラウンドチャンネル情報43とHOA係数11とをバックグラウンド（BG）選択ユニット36に、バックグラウンドチャンネル情報43を係数低減ユニット46およびビットストリーム生成ユニット42に、nFG45をフォ

アグラウンド選択ユニット 36 に出力する。

【0087】

[0070]バックグラウンド選択ユニット 48 は、バックグラウンドチャンネル情報（たとえば、バックグラウンド音場（ N_{BG} ）と、送るべき追加の BG HOA チャンネルの数（ n_{BGa} ）およびインデックス（ i ）と）に基づいて、バックグラウンドまたは環境 HOA 係数 47 を決定するように構成されたユニットを表し得る。たとえば、 N_{BG} が 1 に等しいとき、バックグラウンド選択ユニット 48 は、1 以下の次数を有するオーディオフレームの各サンプルの HOA 係数 11 を選択することができる。バックグラウンド選択ユニット 48 は次いで、この例では、インデックス（ i ）のうちの 1 つによって特定されるインデックスを有する HOA 係数 11 を、追加の BG HOA 係数として選択することができ、こ

ここで、 n_{BGa} は、図 2 および図 4 の例に示されるオーディオ復号デバイス 24 などのオーディオ復号デバイスがビットストリーム 21 からバックグラウンド HOA 係数 47 を解析することを可能にするために、ビットストリーム 21 において指定されるために、ビットストリーム生成ユニット 42 に提供される。バックグラウンド選択ユニット 48 は次いで、環境 HOA 係数 47 をエネルギー補償ユニット 38 に出力することができる。環境 HOA 係数 47 は、次元 $D: M \times [(N_{BG} + 1)^2 + n_{BGa}]$ を有し得る。環境 HOA 係数 47 はまた、「環境 HOA 係数 47」と呼ばれることもあり、ここで、環境 HOA 係数 47 の各々は、聴覚心理オーディオコーダユニット 40 によって符号化されるべき別個の環境 HOA チャンネル 47 に対応する。

【0088】

[0071]フォアグラウンド選択ユニット 36 は、（フォアグラウンドベクトルを特定する 1 つまたは複数のインデックスを表し得る） $n_{FG} 45$ に基づいて、音場のフォアグラウンド成分または明瞭な成分を表す、並べ替えられた $US[k]$ 行列 33' と並べ替えられた $V[k]$ 行列 35' とを選択するように構成されたユニットを表し得る。フォアグラウンド選択ユニット 36 は、（並べ替えられた $US[k]_{1, \dots, n_{FG} 49}$ 、 $FG_{1, \dots, n_{FG} 49}$ 、または

【0089】

【数 15】

$$X_{PS}^{(1..n_{FG})}(k) \quad 49$$

【0090】

として示され得る） n_{FG} 信号 49 を、聴覚心理オーディオコーダユニット 40 に出力することができる。ここで、 n_{FG} 信号 49 は次元 $D: M \times n_{FG}$ を有し、モノラルオーディオオブジェクトを各々表し得る。フォアグラウンド選択ユニット 36 はまた、音場のフォアグラウンド成分に対応する並べ替えられた $V[k]$ 行列 35'（または $v^{(1..n_{FG})}(k) 35'$ ）を空間時間的補間ユニット 50 に出力することができる。ここで、フォアグラウンド成分に対応する並べ替えられた $V[k]$ 行列 35' のサブセットは、次元 $D: (N + 1)^2 \times n_{FG}$ を有するフォアグラウンド $V[k]$ 行列 51_k として示され得る（これは、

【0091】

【数 16】

$$\bar{V}_{1..n_{FG}}[k]$$

【0092】

として数学的に示され得る）。

【0093】

10

20

30

40

50

[0072]エネルギー補償ユニット38は、バックグラウンド選択ユニット48によるH O Aチャンネルのうちの様々なチャンネルの除去によるエネルギー損失を補償するために、環境H O A係数47に関してエネルギー補償を実行するように構成されたユニットを表し得る。エネルギー補償ユニット38は、並べ替えられたU S [k] 行列33'、並べ替えられたV [k] 行列35'、n F G信号49、フォアグラウンドV [k] ベクトル51_k、および環境H O A係数47のうちの1つまたは複数に関してエネルギー分析を実行し、次いで、エネルギー補償された環境H O A係数47'を生成するために、そのエネルギー分析に基づいてエネルギー補償を実行することができる。エネルギー補償ユニット38は、エネルギー補償された環境H O A係数47'を聴覚心理オーディオコーデックユニット40に出力することができる。

10

【0094】

[0073]空間時間的補間ユニット50は、k番目のフレームのためのフォアグラウンドV [k] ベクトル51_kと以前のフレームのための(したがってk-1という表記である)フォアグラウンドV [k-1] ベクトル51_{k-1}とを受信し、補間されたフォアグラウンドV [k] ベクトルを生成するために空間時間的補間を実行するように構成されたユニットを表し得る。空間時間的補間ユニット50は、並べ替えられたフォアグラウンドH O A係数を復元するために、n F G信号49をフォアグラウンドV [k] ベクトル51_kと再び組み合わせることができる。空間時間的補間ユニット50は次いで、補間されたn F G信号49'を生成するために、補間されたV [k] ベクトルによって、並べ替えられたフォアグラウンドH O A係数を分割することができる。空間時間的補間ユニット50はまた、オーディオ復号デバイス24などのオーディオ復号デバイスが補間されたフォアグラウンドV [k] ベクトルを生成しそれによってフォアグラウンドV [k] ベクトル51_kを復元できるように、補間されたフォアグラウンドV [k] ベクトルを生成するために使用されたフォアグラウンドV [k] ベクトル51_kを出力することができる。補間されたフォアグラウンドV [k] ベクトルを生成するために使用されたフォアグラウンドV [k] ベクトル51_kは、残りのフォアグラウンドV [k] ベクトル53として示される。同じV [k] およびV [k-1] が符号化器および復号器において(補間されたベクトルV [k] を作成するために)使用されることを保証するために、ベクトルの量子化された/逆量子化されたバージョンが符号化器および復号器において使用され得る。

20

【0095】

[0074]演算において、空間時間的補間ユニット50は、1つまたは複数のサブフレームのための分解され補間された球面調和係数を生成するために、第1のフレーム中に含まれる第1の複数のH O A係数11の一部分の第1の分解物、たとえばフォアグラウンドV [k] ベクトル51_k、および第2のフレーム中に含まれる第2の複数のH O A係数11の一部分の第2の分解物、たとえばフォアグラウンドV [k] ベクトル51_{k-1}から、第1のオーディオフレームの1つまたは複数のサブフレームを補間することができる。

30

【0096】

[0075]いくつかの例では、第1の分解物は、H O A係数11の部分の右特異ベクトルを表す第1のフォアグラウンドV [k] ベクトル51_kを備える。同様に、いくつかの例では、第2の分解物は、H O A係数11の部分の右特異ベクトルを表す第2のフォアグラウンドV [k] ベクトル51_kを備える。

40

【0097】

[0076]言い換えれば、球面調和関数ベースの3Dオーディオは、球面上の直交基底関数による3D圧力場のパラメトリックな表現であり得る。表現の次数Nが高いほど、空間分解能は高くなる可能性があり、(全体で(N+1)²個の係数に対して)球面調和関数(SH)係数の数は大きくなることが多い。多くの適用形態において、係数を効率的に送信し記憶することを可能にするために、係数の帯域幅圧縮が必要とされ得る。本開示において対象とされる本技法は、特異値分解(SVD)を使用した、フレームベースの次元低減プロセスを提供することができる。SVD分析は、係数の各フレームを3つの行列U、S、およびVに分解することができる。いくつかの例では、本技法は、U S [k] 行列中の

50

ベクトルのいくつかを、背後にある音場のフォアグラウンド成分として扱うことができる。しかしながら、この方法で扱われると、(U S [k] 行列中の) ベクトルは、それらが同じ明瞭なオーディオ成分を表すとしても、フレームとフレームとの間で不連続である。不連続性は、成分が変換オーディオコードを通じて供給されるときに、重大なアーティファクトにつながり得る。

【0098】

[0077]いくつかの点で、空間時間的補間は、球面調和関数領域における直交空間軸としてV行列が解釈され得るという観測に依拠し得る。U [k] 行列は、基底関数によって球面調和関数(HOA)データの投影を表すことができ、ここで、不連続性は、フレームごとに変化したことによってそれら自体が不連続である直交空間軸(V [k])に原因を帰すことができる。これは、基底関数がいくつかの例では複数のフレームにわたって一定であるフーリエ変換など、いくつかの他の分解とは異なる。これらの点で、SVDは、マッチング追跡アルゴリズムと見なされ得る。空間時間的補間ユニット50は、フレームとフレームとの間で基底関数(V [k])の連続性を、フレーム間を補間することによって潜在的に維持するために、補間を実行することができる。

10

【0099】

[0078]上述されたように、補間はサンプルに関して実行され得る。この場合は、サブフレームがサンプルの単一のセットを備えるときの上記の説明において一般化される。サンプルにわたる補間とサブフレームにわたる補間の両方の場合において、補間演算は次の式の形態をとり得る。

20

【0100】

【数17】

$$\overline{v(l)} = w(l)v(k) + (1 - w(l))v(k - 1)$$

【0101】

上の式において、補間は、単一のVベクトルv (k - 1) から単一のVベクトルv (k) に関して実行されてよく、このことは、一態様では、隣接するフレームkおよびk - 1 からVベクトルを表し得る。上の式において、lは補間が実行されている分解能を表し、ここで、lは整数のサンプルを示すことができ、l = 1, . . . , Tである(ここで、Tはそれらにわたる補間が実行されており出力される補間されたベクトル

30

【0102】

【数18】

$$\overline{v(l)}$$

【0103】

が必要とされるサンプルの長さであり、そのプロセスの出力がベクトルのlを作成することも示す)。代替的に、lは複数のサンプルからなるサブフレームを示し得る。たとえば、フレームが4つのサブフレームに分割されるとき、lはサブフレームの各々1つに対して、1、2、3、および4という値を備え得る。lの値は、ビットストリームを通じて「Coded Spatial Interpolation Time」という名称のフィールドとしてシグナリングされ得るので、補間演算は復号器において繰り返され得る。w (l) は、補間の重みの値を備え得る。補間が線形であるとき、w (l) は、lの関数として0と1との間で線形に、および単調に変化し得る。他の場合には、w (l) は、lの関数として非線形であるが単調な(二乗余弦の4分の1周期などの)方式で0と1との間で変化し得る。関数w (l) は、同一の補間演算が復号器によって繰り返され得るように、いくつかの異なる関数の可能性の間でインデックスが付けられて、「Spatial Int

40

50

「`interpolationMethod`」という名称のフィールドとしてビットストリームにおいてシグナリングされ得る。 $w(1)$ が 0 に近い値を有するとき、出力

【0104】

【数19】

$$\overline{v(l)}$$

【0105】

は、 $v(k-1)$ によって大きく重み付けられ、またはその影響を受け得る。一方、 $w(1)$ が 1 に近い値を有するとき、そのことは、出力

【0106】

【数20】

$$\overline{v(l)}$$

【0107】

が $v(k-1)$ によって大きく重み付けられ、またはその影響を受けることを保証する。

【0108】

[0079] 係数低減ユニット 46 は、低減されたフォアグラウンド $V[k]$ ベクトル 55 を量子化ユニット 52 に出力するために、バックグラウンドチャンネル情報 43 に基づいて残りのフォアグラウンド $V[k]$ ベクトル 53 に関して係数低減を実行するように構成されたユニットを表し得る。低減されたフォアグラウンド $V[k]$ ベクトル 55 は、次元 D : $[(N+1)^2 - (N_{BG}+1)^2 - BG_{TOT}] \times nFG$ を有し得る。

【0109】

[0080] 係数低減ユニット 46 は、この点において、残りのフォアグラウンド $V[k]$ ベクトル 53 における係数の数を低減するように構成されたユニットを表し得る。言い換えれば、係数低減ユニット 46 は、指向性情報をほとんどまたはまったく有しない（残りのフォアグラウンド $V[k]$ ベクトル 53 を形成する）フォアグラウンド $V[k]$ ベクトルにおける係数を除去するように構成されたユニットを表し得る。上記で説明されたように、いくつかの例では、(N_{BG} と示され得る) 1 次および 0 次の基底関数に対応する、明瞭な、または言い換えればフォアグラウンド $V[k]$ ベクトルの係数は、指向性情報をほとんど提供せず、したがって、「係数低減」と呼ばれ得るプロセスを通じて）フォアグラウンド V ベクトルから除去され得る。この例では、 N_{BG} に対応する係数を特定するだけでなく、追加の HOA チャンネル（変数 $TotalOfAddAmbHOACHan$ によって示され得る）を $[(N_{BG}+1)^2 + 1, (N+1)^2]$ のセットから特定するために、より大きい柔軟性が与えられ得る。音場分析ユニット 44 は、 BG_{TOT} を決定するために HOA 係数 11 を分析することができ、 BG_{TOT} は、($N_{BG}+1)^2$ だけではなく $TotalOfAddAmbHOACHan$ を特定することができ、これらは総称してバックグラウンドチャンネル情報 43 と呼ばれ得る。係数低減ユニット 46 は次いで、低減されたフォアグラウンド $V[k]$ ベクトル 55 とも呼ばれ得る、サイズが $((N+1)^2 - (BG_{TOT})) \times nFG$ であるより低次元の $V[k]$ 行列 55 を生成するために、($N_{BG}+1)^2$ および $TotalOfAddAmbHOACHan$ に対応する係数を残りのフォアグラウンド $V[k]$ ベクトル 53 から除去することができる。

【0110】

[0081] 言い換えれば、公開第 WO 2014/194099 号において示されているように、係数低減ユニット 46 は、サイドチャンネル情報 57 のためのシンタックス要素を生成することができる。たとえば、係数低減ユニット 46 は、複数の構成モードのいずれが選択されたかを示す、(1 つまたは複数のフレームを含み得る) アクセスユニットのヘッダ

中のシンタックス要素を指定することができる。アクセスユニットごとに指定されるものとして説明されるが、係数低減ユニット46は、フレームごとに、または任意の他の周期的な方式で、または非周期的に（ビットストリーム全体で1回など）シンタックス要素を指定することができる。いずれにしても、シンタックス要素は、明瞭な成分の指向性の態様を表すために、3つの構成モードのいずれが低減されたフォアグラウンドV[k]ベクトル55の係数の0ではないセットを指定するために選択されたかを示す2つのビットを備え得る。シンタックス要素は、「C o d e d V V e c L e n g t h」として示され得る。このようにして、係数低減ユニット46は、3つの構成モードのいずれが低減されたフォアグラウンドV[k]ベクトル55をビットストリーム21において指定するために使用されたかを、ビットストリームにおいてシグナリングし、またはさもなければ指定することができる。

10

【0111】

[0082]たとえば、3つの構成モードは、（本文書において後で言及される）V V e c D a t aのためのシンタックステーブルにおいて提示され得る。その例では、構成モードは次のようになる。（モード0）、完全なVベクトル長がV V e c D a t aフィールドにおいて送信される、（モード1）、環境H O A係数のための係数の最小数に関連付けられたVベクトルの要素、および、送信されない追加のH O Aチャンネルを含んだVベクトルの要素のすべて、ならびに（モード2）、環境H O A係数のための係数の最小数に関連付けられたVベクトルの要素が送信されない。V V e c D a t aのシンタックステーブルは、s w i t c hおよびc a s e文とともにモードを示す。3つの構成モードに関して説明されるが、本技法は3つの構成モードに限定されるべきではなく、単一の構成モードまたは複数のモードを含む、任意の数の構成モードを含み得る。公開第W O 2 0 1 4 / 1 9 4 0 9 9号は、4つのモードを有する異なる例を提供している。係数低減ユニット46はまた、サイドチャンネル情報57における別のシンタックス要素としてフラグ63を指定することができる。

20

【0112】

[0083]量子化ユニット52は、コーディングされたフォアグラウンドV[k]ベクトル57を生成するために低減されたフォアグラウンドV[k]ベクトル55を圧縮するための任意の形態の量子化を実行し、コーディングされたフォアグラウンドV[k]ベクトル57をビットストリーム生成ユニット42に出力するように構成されたユニットを表し得る。動作において、量子化ユニット52は、音場の空間成分、すなわちこの例では低減されたフォアグラウンドV[k]ベクトル55のうちの1つまたは複数を圧縮するように構成されたユニットを表し得る。例示の目的で、低減されたフォアグラウンドV[k]ベクトル55は、係数低減の結果として各々25個未満の要素を有する（これは音場の4次のH O A表現を示唆する）2つの行ベクトルを含むと仮定される。2つの行ベクトルに関して説明されるが、最大で $(n+1)^2$ 個までの任意の数のベクトルが低減されたフォアグラウンドV[k]ベクトル55に含まれてよく、ここで、nは音場のH O A表現の次数を示す。その上、スカラー量子化および／またはエントロピー量子化を実行するものとして以下で説明されるが、量子化ユニット52は、低減されたフォアグラウンドV[k]ベクトル55の圧縮をもたらす任意の形態の量子化を実行することができる。

30

40

【0113】

[0084]量子化ユニット52は、低減されたフォアグラウンドV[k]ベクトル55を受信し、コーディングされたフォアグラウンドV[k]ベクトル57を生成するためにある圧縮方式を実行することができる。その圧縮方式は、ベクトルまたはデータの要素を圧縮するための任意の想起可能な圧縮方式を全般に含んでよく、以下でより詳細に説明される例に限定されるべきではない。量子化ユニット52は、一例として、低減されたフォアグラウンドV[k]ベクトル55の各要素の浮動小数点表現を低減されたフォアグラウンドV[k]ベクトル55の各要素の整数表現へと変換すること、低減されたフォアグラウンドV[k]ベクトル55の整数表現の一樣量子化、ならびに、残りのフォアグラウンドV[k]ベクトル55の量子化された整数表現の分類およびコーディングのうちの、1つま

50

たは複数を含む、圧縮方式を実行することができる。

【0114】

[0085]いくつかの例では、圧縮方式の1つまたは複数のプロセスのいくつか、一例として、得られるビットストリーム21のための目標ビットレート41を達成するために、またはほぼ達成するために、パラメータによって動的に制御され得る。低減されたフォアグラウンドV[k]ベクトル55の各々が互いに正規直交であるとする、低減されたフォアグラウンドV[k]ベクトル55の各々は独立にコーディングされ得る。いくつかの例では、以下でより詳細に説明されるように、各々の低減されたフォアグラウンドV[k]ベクトル55の各要素は、(様々なサブモードによって定義される)同じコーディングモードを使用してコーディングされ得る。

10

【0115】

[0086]公開第WO2014/194099号において説明されているように、量子化ユニット52は、低減されたフォアグラウンドV[k]ベクトル55を圧縮するためにスカラー量子化および/またはハフマン符号化を実行し、サイドチャンネル情報57とも呼ばれ得るコーディングされたフォアグラウンドV[k]ベクトル57を出力することができる。サイドチャンネル情報57は、残りのフォアグラウンドV[k]ベクトル55をコーディングするために使用されるシンタックス要素を含み得る。

【0116】

[0087]その上、スカラー量子化の形態に関して説明されるが、量子化ユニット52は、ベクトル量子化または任意の他の形態の量子化を実行することができる。いくつかの場合には、量子化ユニット52は、ベクトル量子化とスカラー量子化との間で切り替えることができる。上記で説明されたスカラー量子化の間に、量子化ユニット52は、(フレームからフレームへのように連続する)2つの連続するVベクトル間の差分を計算し、その差分(または、言い換えれば、残差)をコーディングすることができる。このスカラー量子化は、以前に指定されたベクトルおよび差分信号に基づく、ある形態の予測コーディングを表し得る。ベクトル量子化は、そのような差分コーディングを伴わない。

20

【0117】

[0088]言い換えれば、量子化ユニット52は、入力Vベクトル(たとえば、低減されたフォアグラウンドV[k]ベクトル55のうちの1つ)を受信し、その入力Vベクトルのために使用されるべき量子化のタイプのうちの1つを選択するために、異なるタイプの量子化を実行することができる。量子化ユニット52は、一例として、ベクトル量子化と、ハフマンコーディングなしのスカラー量子化と、ハフマンコーディングありのスカラー量子化とを実行することができる。

30

【0118】

[0089]この例では、量子化ユニット52は、ベクトル量子化されたVベクトルを生成するために、ベクトル量子化モードに従って、入力Vベクトルをベクトル量子化することができる。ベクトル量子化されたVベクトルは、入力Vベクトルを表すベクトル量子化された重み値を含み得る。ベクトル量子化された重み値は、いくつかの例では、量子化コードワードの量子化コードブックにおける量子化コードワード(すなわち、量子化ベクトル)を指す、1つまたは複数の量子化インデックスとして表され得る。量子化ユニット52は、ベクトル量子化を実行するように構成されるとき、コードベクトル63(「CV63」)に基づいて、低減されたフォアグラウンドV[k]ベクトル55の各々をコードベクトルの加重和に分解することができる。量子化ユニット52は、コードベクトル63のうちの選択されたものの各々のための重み値を生成することができる。

40

【0119】

[0090]量子化ユニット52は次に、重み値の選択されたサブセットを生成するために、重み値のサブセットを選択することができる。たとえば、量子化ユニット52は、重み値の選択されたサブセットを生成するために、重み値のセットから、Z個の最大の大きさの重み値を選択することができる。いくつかの例では、量子化ユニット52は、重み値の選択されたサブセットを生成するために、選択された重み値をさらに並べ替えることができ

50

る。たとえば、量子化ユニット 5 2 は、最高の大きさの重み値から開始して、最低の大きさの重み値で終了するように、大きさに基づいて、選択された重み値を並べ替えることができる。

【0 1 2 0】

[0091]ベクトル量子化を実行するとき、量子化ユニット 5 2 は、Z 個の重み値を表すために、量子化コードブックから Z 成分ベクトルを選択することができる。言い換えれば、量子化ユニット 5 2 は、Z 個の重み値を表す Z 成分ベクトルを生成するために、Z 個の重み値をベクトル量子化することができる。いくつかの例では、Z は、単一の V ベクトルを表すために量子化ユニット 5 2 によって選択された重み値の数に対応し得る。量子化ユニット 5 2 は、Z 個の重み値を表すために選択された Z 成分ベクトルを示すデータを生成し、このデータを、コード化された重み 5 7 としてビットストリーム生成ユニット 4 2 に与えることができる。いくつかの例では、量子化コードブックは、インデックス付けされる複数の Z 成分ベクトルを含み得、Z 成分ベクトルを示すデータは、選択されたベクトルを指す量子化コードブックへのインデックス値であり得る。そのような例では、復号器は、インデックス値を復号するために、同様にインデックス付けされた量子化コードブックを含み得る。

10

【0 1 2 1】

[0092]数学的には、低減されたフォアグラウンド V [k] ベクトル 5 5 の各々は、次の式に基づいて表され得る。

【0 1 2 2】

20

【数 2 1】

$$V \approx \sum_{j=1}^J \omega_j \Omega_j \quad (1)$$

【0 1 2 3】

ただし、 Ω_j は、コードベクトルのセット ($\{\Omega_j\}$) における j 番目のコードベクトルを表し、 ω_j は、重みのセット ($\{\omega_j\}$) における j 番目の重みを表し、V は、V ベクトルコーディングユニット 5 2 によって表され、分解および／またはコーディングされている V ベクトルに対応し、J は、V を表すために使用された重みの数とコードベクトルの数とを表す。式 (1) の右辺は、重みのセット ($\{\omega_j\}$) とコードベクトルのセット ($\{\Omega_j\}$) とを含む、コードベクトルの加重和を表し得る。

30

【0 1 2 4】

[0093]いくつかの例では、量子化ユニット 5 2 は、次の式

【0 1 2 5】

【数 2 2】

$$\omega_k = V \Omega_k^T \quad (2)$$

40

【0 1 2 6】

に基づいて、重み値を決定することができ、ただし、

【0 1 2 7】

【数 2 3】

$$\Omega_k^T$$

【0 1 2 8】

は、コードベクトルのセット ($\{\Omega_k\}$) における k 番目のコードベクトルの転置を表し

50

、 V は、量子化ユニット52によって表され、分解および／またはコーディングされている V ベクトルに対応し、 ω_k は、重みのセット（ $\{\omega_k\}$ ）における k 番目の重みを表す。

【0129】

[0094] 25個の重みおよび25個のコードベクトルが、 V ベクトル V_{FG} を表すために使用される一例を検討する。 V_{FG} のそのような分解は、

【0130】

【数24】

$$V_{FG} \approx \sum_{j=1}^{25} \omega_j \Omega_j \quad (3)$$

10

【0131】

として書かれ得、ただし、 Ω_j は、コードベクトルのセット（ $\{\Omega_j\}$ ）における j 番目のコードベクトルを表し、 ω_j は、重みのセット（ $\{\omega_j\}$ ）における j 番目の重みを表し、 V_{FG} は、量子化ユニット52によって表され、分解および／またはコーディングされている V ベクトルに対応する。

【0132】

[0095] コードベクトルのセット（ $\{\Omega_j\}$ ）が正規直交である例では、次の式が適用され得る。

【0133】

【数25】

$$\Omega_j \Omega_k^T = \begin{cases} 1 & j=k \text{ に対して} \\ 0 & j \neq k \text{ に対して} \end{cases} \quad (4)$$

20

【0134】

そのような例では、式（3）の右辺は次のように簡略化し得る。

【0135】

【数26】

$$V_{FG} \Omega_k^T \approx \left(\sum_{j=1}^{25} \omega_j \Omega_j \right) \Omega_k^T = \omega_k \quad (5)$$

30

【0136】

ただし、 ω_k は、コードベクトルの加重和における k 番目の重みに対応する。

【0137】

[0096] 式（3）において使用されたコードベクトルの例示的な加重和では、量子化ユニット52は、（式（2）と同様の）式（5）を使用して、コードベクトルの加重和における重みの各々のための重み値を計算することができ、得られる重みは次のように表され得る。

40

【0138】

【数27】

$$\{\omega_k\}_{k=1, \dots, 25} \quad (6)$$

【0139】

量子化ユニット52が5個の極大重み値（すなわち、最大値または絶対値をもつ重み）を

50

選択する一例を検討する。量子化されるべき重み値のサブセットは、次のように表され得る。

【 0 1 4 0 】

【 数 2 8 】

$$\{\overline{\omega}_k\}_{k=1,\dots,5} \quad (7)$$

【 0 1 4 1 】

重み値のサブセットは、それらの対応するコードベクトルとともに、次の式において示されるように、Vベクトルを推定するコードベクトルの加重和を形成するために使用され得る。 10

【 0 1 4 2 】

【 数 2 9 】

$$\overline{V}_{FG} \approx \sum_{j=1}^5 \overline{\omega}_j \Omega_j \quad (8)$$

【 0 1 4 3 】

ただし、 Ω_j は、コードベクトルのサブセット（ $\{\Omega_j\}$ ）におけるj番目のコードベクトルを表し、 20

【 0 1 4 4 】

【 数 3 0 】

$$\overline{\omega}_j$$

【 0 1 4 5 】

は、重みのサブセット（

【 0 1 4 6 】

【 数 3 1 】

30

$$\{\overline{\omega}_j\}$$

【 0 1 4 7 】

）におけるj番目の重みを表し、

【 0 1 4 8 】

【 数 3 2 】

$$\overline{V}_{FG}$$

40

【 0 1 4 9 】

は、量子化ユニット52によって分解および／またはコーディングされているVベクトルに対応する、推定されたVベクトルに対応する。式（1）の右辺は、重みのセット（

【 0 1 5 0 】

【数 3 3】

$$\{\overline{\omega}_j\}$$

【0 1 5 1】

）とコードベクトルのセット（ $\{\Omega_j\}$ ）とを含む、コードベクトルの加重和を表し得る。

【0 1 5 2】

[0097]量子化ユニット 5 2 は、

10

【0 1 5 3】

【数 3 4】

$$\{\hat{\omega}_k\}_{k=1,\dots,5} \quad (9)$$

【0 1 5 4】

として表され得る量子化された重み値を生成するために、重み値のサブセットを量子化することができる。量子化された重み値は、それらの対応するコードベクトルとともに、次の式において示されるように、推定された V ベクトルの量子化されたバージョンを表すコードベクトルの加重和を形成するために使用され得る。

20

【0 1 5 5】

【数 3 5】

$$\hat{V}_{FG} \approx \sum_{j=1}^5 \hat{\omega}_j \Omega_j \quad (10)$$

【0 1 5 6】

ただし、 Ω_j は、コードベクトルのサブセット（ $\{\Omega_j\}$ ）における j 番目のコードベクトルを表し、

30

【0 1 5 7】

【数 3 6】

$$\hat{\omega}_j$$

【0 1 5 8】

は、重みのサブセット（

【0 1 5 9】

【数 3 7】

40

$$\{\hat{\omega}_j\}$$

【0 1 6 0】

）における j 番目の重みを表し、

【0 1 6 1】

【数 3 8】

$$\hat{V}_{FG}$$

【0 1 6 2】

は、量子化ユニット 5 2 によって分解および／またはコーディングされている V ベクトルに対応する、推定された V ベクトルに対応する。式 (1) の右辺は、重みのセット (

【0 1 6 3】

【数 3 9】

10

$$\{\hat{\omega}_j\}$$

【0 1 6 4】

) とコードベクトルのセット ($\{\Omega_j\}$) とを含む、コードベクトルのサブセットの加重和を表し得る。

【0 1 6 5】

[0098] 上記の代替的な言い換え (大部分は上記で説明されたものと同等である) は、次のようになり得る。V ベクトルは、コードベクトルのあらかじめ定義されたセットに基づいてコーディングされ得る。V ベクトルをコーディングするために、各 V ベクトルは、コードベクトルの加重和に分解される。コードベクトルの加重和は、あらかじめ定義されたコードベクトルと関連付けられた重みとの k 個のペアからなる。

20

【0 1 6 6】

【数 4 0】

$$V \approx \sum_{j=0}^k \omega_j \Omega_j \quad (11)$$

30

【0 1 6 7】

ただし、 Ω_j は、あらかじめ定義されたコードベクトルのセット ($\{\Omega_j\}$) における j 番目のコードベクトルを表し、 ω_j は、あらかじめ定義された重みのセット ($\{\omega_j\}$) における j 番目の実数値の重みを表し、k は、最大 7 であり得る加数のインデックスに対応し、V は、コーディングされている V ベクトルに対応する。k の選定は、符号化器に依存する。符号化器が 2 つ以上のコードベクトルの加重和を選定する場合、符号化器を選定できるあらかじめ定義されたコードベクトルの総数は、 $(N+1)^2$ であり、そのあらかじめ定義されたコードベクトルは、2014 年 7 月 25 日付けの、文書番号 ISO/IEC DIS 23008-3 によって識別される、ISO/IEC JTC1/SC29/WG11 による「Information technology—High efficiency coding and media delivery in heterogeneous environments—Part 3: 3D audio」という名称の 3D Audio 規格の Table F. 3~F. 7 から、HOA 拡張係数として導出される。N が 4 であるとき、32 個のあらかじめ定義された方向を有する、上記の 3D Audio 規格の Annex F. 5 におけるテーブルが使用される。すべての場合において、重み ω の絶対値が、上記の 3D Audio 規格のテーブル F. 12 におけるテーブルの最初の k+1 列において見つかる、あらかじめ定義された重み付け値

40

【0 1 6 8】

【数 4 1】

$$\hat{\omega}$$

【0 1 6 9】

に関してベクトル量子化され、関連付けられた行番号インデックスとともにシグナリングされる。

【0 1 7 0】

[0099]重み ω の数の符号は、

10

【0 1 7 1】

【数 4 2】

$$s_j = \begin{cases} 1, & \omega_j \geq 0 \\ 0, & \omega_j < 0 \end{cases} \quad (12)$$

【0 1 7 2】

として別個にコーディングされる。

【0 1 7 3】

[0100]言い換えれば、値 k をシグナリングした後、 V ベクトルは、 $k+1$ 個のあらかじめ定義されたコードベクトル $\{\Omega_j\}$ を指す $k+1$ 個のインデックスと、あらかじめ定義された重み付けコードブックにおける k 個の量子化された重み

20

【0 1 7 4】

【数 4 3】

$$\{\hat{\omega}_k\}$$

【0 1 7 5】

を指す1つのインデックスと、 $k+1$ 個の数の符号値 s_j とを用いて符号化される。

30

【0 1 7 6】

【数 4 4】

$$\hat{V} = \sum_{j=0}^k (2s_j - 1) \hat{\omega}_j \Omega_j. \quad (13)$$

【0 1 7 7】

符号化器が、1つのコードベクトルの加重和を選択する場合、上記の3D Audio規格のテーブルF. 8から導出されたコードブックが、上記の3D Audio規格のテーブルF. 11のテーブルにおける絶対重み付け値

40

【0 1 7 8】

【数 4 5】

$$\hat{\omega}$$

【0 1 7 9】

と組み合わせて使用され、ここで、これらのテーブルの両方が以下で示される。また、重み付け値 ω の数の符号は、別個にコーディングされ得る。量子化ユニット52は、コードブックインデックスシンタックス要素（以下で「Codebook Index」と示され得る）を

50

使用して、上述されたテーブル F. 3 から F. 12 において記載された上記のコードブックのいずれが入力 V ベクトルをコーディングするために使用されるかをシグナリングすることができる。量子化ユニット 52 はまた、スカラー量子化された V ベクトルをハフマンコーディングすることなく、出力のスカラー量子化された V ベクトルを生成するために、入力 V ベクトルをスカラー量子化することもできる。量子化ユニット 52 は、ハフマンコーディングされスカラー量子化された V ベクトルを生成するために、ハフマンコーディングスカラー量子化モードに従って、入力 V ベクトルをさらにスカラー量子化することができる。たとえば、量子化ユニット 52 は、スカラー量子化された V ベクトルを生成するために、入力 V ベクトルをスカラー量子化し、出力のハフマンコーディングされスカラー量子化された V ベクトルを生成するために、スカラー量子化された V ベクトルをハフマンコーディングすることができる。

10

【0180】

[0101]いくつかの例では、量子化ユニット 52 は、ある形態の予測ベクトル量子化を実行することができる。量子化ユニット 52 は、（量子化モードを示す 1 つまたは複数のビット、たとえば、Nb i t s Q シンタックス要素によって特定されるように）予測がベクトル量子化について実行されるか否かを示すビットストリーム 21 中の 1 つまたは複数のビット（たとえば、P F l a g シンタックス要素）を特定することによって、ベクトル量子化が予測されるか否かを特定することができる。

【0181】

[0102]予測ベクトル量子化について説明するために、量子化ユニット 42 は、ベクトル（たとえば、v ベクトル）のコードベクトルベース分解物に対応する重み値（たとえば、重み値の大きさ）を受信することと、受信された重み値に基づいて、および再構成された重み値（たとえば、1 つまたは複数の以前または後続のオーディオフレームから再構成された重み値）に基づいて、予測重み値を生成することと、予測重み値のセットをベクトル量子化することとを行うように構成され得る。場合によっては、予測重み値のセットにおける各重み値は、単一のベクトルのコードベクトルベース分解物中に含まれる重み値に対応し得る。

20

【0182】

[0103]量子化ユニット 52 は、ベクトルの以前または後続のコーディングから、重み値と重み付きの再構成された重み値とを受信することができる。量子化ユニット 52 は、重み値と重み付きの再構成された重み値とに基づいて、予測重み値を生成することができる。量子化ユニット 42 は、予測重み値を生成するために、重み値から重み付きの再構成された重み値を減算することができる。予測重み値は、代替的に、たとえば、残差、予測残差、残差重み値、重み値差分、誤差、または予測誤差と呼ばれることがある。

30

【0183】

[0104]重み値は、対応する重み値 $w_{i,j}$ の大きさ（または絶対値）である $|w_{i,j}|$ として表され得る。したがって、重み値は代替的に、重み値大きさ、または重み値の大きさと呼ばれることがある。重み値 $w_{i,j}$ は、i 番目のオーディオフレームのための重み値の順序付きサブセットからの j 番目の重み値に対応する。いくつかの例では、重み値の順序付きサブセットは、重み値の大きさに基づいて順序付けされる（たとえば、最大の大きさから最小の大きさへと順序付けされる）ベクトル（たとえば、v ベクトル）のコードベクトルベース分解物中の重み値のサブセットに対応し得る。

40

【0184】

[0105]重み付きの再構成された重み値は、

【0185】

【数 46】

$$|\hat{w}_{i-1,j}|$$

50

【0186】

項を含み得、この項は、対応する再構成された重み値

【0187】

【数47】

$$\hat{w}_{i-1,j}$$

【0188】

の大きさ（または絶対値）に対応する。再構成された重み値

10

【0189】

【数48】

$$\hat{w}_{i-1,j}$$

【0190】

は、 $(i-1)$ 番目のオーディオフレームのための再構成された重み値の順序付きサブセットからの j 番目の再構成された重み値に対応する。いくつかの例では、再構成された重み値の順序付きサブセット（またはセット）は、再構成された重み値に対応する、量子化された予測重み値に基づいて生成され得る。

20

【0191】

[0106] 量子化ユニット42はまた、重み係数 α_j を含む。いくつかの例では、 $\alpha_j = 1$ であり、その場合、重み付きの再構成された重み値は、

【0192】

【数49】

$$|\hat{w}_{i-1,j}|$$

30

【0193】

に低減し得る。他の例では、 $\alpha_j \neq 1$ である。たとえば、 α_j は、次の式に基づいて決定され得る。

【0194】

【数50】

$$\alpha_j = \frac{\sum_{i=1}^I w_{i,j} w_{i-1,j}}{\sum_{i=1}^I w_{i-1,j}^2}$$

40

【0195】

ただし、 I は、 α_j を決定するために使用されたオーディオフレームの数に対応する。前の式において示されたように、重み係数は、いくつかの例では、複数の異なるオーディオフレームからの複数の異なる重み値に基づいて決定され得る。

【0196】

[0107] また、予測ベクトル量子化を実行するように構成されるとき、量子化ユニット52は、次の式に基づいて、予測重み値を生成することができる。

【0197】

【数 5 1】

$$e_{i,j} = \left| w_{i,j} - \alpha_j \hat{w}_{i-1,j} \right|$$

【0 1 9 8】

ただし、 $e_{i,j}$ は、 i 番目のオーディオフレームのための重み値の順序付きサブセットからの j 番目の重み値のための予測重み値に対応する。

【0 1 9 9】

[0108] 量子化ユニット 5 2 は、予測重み値と予測ベクトル量子化 (P V Q) コードブックとに基づいて、量子化された予測重み値を生成する。たとえば、量子化ユニット 5 2 は、量子化された予測重み値を生成するために、コーディングされるべきベクトルのために、またはコーディングされるべきフレームのために生成された、他の予測重み値と組み合わせて、予測重み値をベクトル量子化することができる。 10

【0 2 0 0】

[0109] 量子化ユニット 5 2 は、P V Q コードブックに基づいて、予測重み値 6 2 0 をベクトル量子化することができる。P V Q コードブックは、複数の M 成分候補量子化ベクトルを含み得、量子化ユニット 5 2 は、Z 個の予測重み値を表すために、候補量子化ベクトルのうちの 1 つを選択することができる。いくつかの例では、量子化ユニット 5 2 は、量子化誤差を最小化する（たとえば、最小 2 乗誤差を最小化する）、P V Q コードブックからの候補量子化ベクトルを選択することができる。 20

【0 2 0 1】

[0110] いくつかの例では、P V Q コードブックは、エントリの各々が量子化コードブックインデックスと対応する M 成分候補量子化ベクトルとを含む、複数のエントリを含み得る。量子化コードブックにおけるインデックスの各々は、複数の M 成分候補量子化ベクトルのうちの各々に対応し得る。

【0 2 0 2】

[0111] 量子化ベクトルの各々における成分の数は、単一の v ベクトルを表すために選択される重みの数（すなわち、Z）に依存し得る。概して、Z 成分候補量子化ベクトルをもつコードブックでは、量子化ユニット 5 2 は、単一の量子化ベクトルを生成するために一度に Z 個の予測重み値をベクトル量子化することができる。量子化コードブックにおけるエントリの数は、重み値をベクトル量子化するために使用されるビットレートに依存し得る。 30

【0 2 0 3】

[0112] 量子化ユニット 5 2 が予測重み値をベクトル量子化するとき、量子化ユニット 5 2 は、Z 個の予測重み値を表す予測ベクトルであるように、P V Q コードブックから Z 成分ベクトルを選択することができる。量子化された予測重み値は、

【0 2 0 4】

【数 5 2】

40

$$\hat{e}_{i,j}$$

【0 2 0 5】

として示され得、これは、 i 番目のオーディオフレームのための j 番目の予測重み値のベクトル量子化されたバージョンにさらに対応し得る、 i 番目のオーディオフレームのための Z 成分量子化ベクトルの j 番目の成分に対応し得る。

【0 2 0 6】

[0113] 予測ベクトル量子化を実行するように構成されるとき、量子化ユニット 5 2 はまた、量子化された予測重み値と重み付きの再構成された重み値とに基づいて、再構成され 50

た重み値を生成することもできる。たとえば、量子化ユニット 5 2 は、再構成された重み値を生成するために、量子化された予測重み値に、重み付きの再構成された重み値を加算することができる。その重み付きの再構成された重み値は、上記で説明されている重み付きの再構成された重み値に等しくなり得る。いくつかの例では、重み付きの再構成された重み値は、再構成された重み値の重み付きおよび遅延されたバージョンであり得る。

【0207】

[0114]再構成された重み値は、

【0208】

【数53】

10

$$\left| \hat{w}_{i-1,j} \right|$$

【0209】

として表され得、これは、対応する再構成された重み値

【0210】

【数54】

$$\hat{w}_{i-1,j}^2$$

20

【0211】

の大きさ（または絶対値）に対応する。再構成された重み値

【0212】

【数55】

$$\hat{w}_{i-1,j}$$

【0213】

は、(i-1) 番目のオーディオフレームのための再構成された重み値の順序付きサブセットからの j 番目の再構成された重み値に対応する。いくつかの例では、量子化ユニット 5 2 は、予測コーディングされる重み値の符号を示すデータを別個にコーディングすることができる、復号器は、再構成された重み値の符号を決定するために、この情報を使用することができる。

30

【0214】

[0115]量子化ユニット 5 2 は、次の式に基づいて、再構成された重み値を生成することができる、

【0215】

【数56】

40

$$\left| \hat{w}_{i,j} \right| = \hat{e}_{i,j} + \alpha_j \left| \hat{w}_{i-1,j} \right|$$

【0216】

ただし、

【0217】

【数 5 7】

$$\hat{e}_{i,j}$$

【0 2 1 8】

は、 i 番目のオーディオフレームのための重み値の順序付きサブセットからの j 番目の重み値（たとえば、 M 成分量子化ベクトルの j 番目の成分）のための量子化された予測重み値に対応し、

【0 2 1 9】

【数 5 8】

10

$$|\hat{w}_{i-1,j}|$$

【0 2 2 0】

は、 $(i-1)$ 番目のオーディオフレームのための重み値の順序付きサブセットからの j 番目の重み値のための再構成された重み値の大きさに対応し、 α_j は、重み値の順序付きサブセットからの j 番目の重み値のための重み係数に対応する。

【0 2 2 1】

20

[0116] 量子化ユニット 5 2 は、再構成された重み値に基づいて、遅延され再構成された重み値を生成することができる。たとえば、量子化ユニット 5 2 は、遅延され再構成された重み値を生成するために、1 オーディオフレームだけ、再構成された重み値を遅延させることができる。

【0 2 2 2】

[0117] 量子化ユニット 5 2 はまた、遅延され再構成された重み値と重み係数とに基づいて、重み付きの再構成された重み値を生成することができる。たとえば、量子化ユニット 5 2 は、重み付きの再構成された重み値を生成するために、遅延され再構成された重み値を重み係数と乗算することができる。

【0 2 2 3】

30

[0118] 同様に、量子化ユニット 5 2 は、遅延され再構成された重み値と重み係数とに基づいて、重み付きの再構成された重み値を生成する。たとえば、量子化ユニット 5 2 は、重み付きの再構成された重み値を生成するために、遅延され再構成された重み値を重み係数と乗算することができる。

【0 2 2 4】

[0119] Z 個の予測重み値のための量子化ベクトルであるように、 PVQ コードブックから Z 成分ベクトルを選択することに対応して、量子化ユニット 5 2 は、いくつかの例では、選択された Z 成分ベクトル自体をコーディングするのではなく、選択された Z 成分ベクトルに対応する（ PVQ コードブックからの）インデックスをコーディングすることができる。インデックスは、量子化された予測重み値のセットを示し得る。そのような例では、復号器 2 4 は、 PVQ コードブックと同様のコードブックを含み得、復号器コードブック中の対応する Z 成分ベクトルにインデックスをマッピングすることによって、量子化された予測重み値を示すインデックスを復号することができる。 Z 成分ベクトルにおける成分の各々は、量子化された予測重み値に対応し得る。

40

【0 2 2 5】

[0120] ベクトル（たとえば、 V ベクトル）をスカラー量子化することは、個々に、および／または他の成分とは無関係に、ベクトルの成分の各々を量子化することを伴い得る。たとえば、次の例示的な V ベクトル

$$V = [0.23 \quad 0.31 \quad -0.47 \quad \cdots \quad 0.85]$$

を検討する。この例示的な V ベクトルをスカラー量子化するために、成分の各々が個々に量子化（す

50

なわち、スカラー量子化)され得る。たとえば、量子化ステップが0.1である場合、0.23成分が0.2に量子化され得、0.31成分が0.3に量子化され得るなどとなる。スカラー量子化された成分は、スカラー量子化されたVベクトルを集合的に形成し得る。

【0226】

[0121]言い換えれば、量子化ユニット52は、低減されたフォアグラウンドV[k]ベクトル55のうちの所与の1つの要素のすべてに関して一様スカラー量子化を実行することができる。量子化ユニット52は、Nb i t s Qシンタックス要素として示され得る値に基づいて、量子化ステップサイズを特定することができる。量子化ユニット52は、目標ビットレート41に基づいて、このNb i t s Qシンタックス要素を動的に決定することができる。Nb i t s Qシンタックス要素はまた、以下で再生されるChannel Side Info Dataシンタックステーブルにおいて示されるように、量子化モードを特定し、一方でまた、スカラー量子化のためにステップサイズを特定することもできる。すなわち、量子化ユニット52は、このNb i t s Qシンタックス要素の関数として量子化ステップサイズを決定することができる。一例として、量子化ユニット52は、 $2^{16-N_{bitsQ}}$ に等しいものとして、量子化ステップサイズ(本開示では「デルタ」または「 Δ 」として示される)を決定することができる。この例では、Nb i t s Qシンタックス要素の値が6に等しいとき、デルタは 2^{10} に等しく、 2^6 個の量子化レベルがある。この点において、ベクトル要素vに対して、量子化されたベクトル要素 v_q は $[v/\Delta]$ に等しく、 $-2^{Nb_{bitsQ}-1} < v_q < 2^{Nb_{bitsQ}-1}$ である。

【0227】

[0122]量子化ユニット52は次いで、量子化されたベクトル要素の分類と残差コーディングとを実行することができる。一例として、量子化ユニット52は、所与の量子化されたベクトル要素 v_q に対して、この要素が対応するカテゴリーを(カテゴリー識別子cidを決定することによって)、次の式

【0228】

【数59】

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lceil \log_2 |v_q| \rceil + 1, & \text{if } v_q \neq 0 \end{cases}$$

【0229】

を使用して特定することができる。量子化ユニット52は次いで、このカテゴリーインデックスcidをハフマンコーディングし、一方で、 v_q が正の値であるか負の値であるかを示す符号ビットを特定することもできる。量子化ユニット52は次に、このカテゴリーにおける残差を特定することができる。一例として、量子化ユニット52は、次の式

【0230】

【数60】

$$residual = |v_q| - 2^{cid-1}$$

【0231】

に従って、この残差を決定することができる。量子化ユニット52は次いで、この残差をcid-1ビットによってブロックコーディングすることができる。

【0232】

[0123]量子化ユニット52は、いくつかの例では、cidをコーディングするとき、Nb i t s Qシンタックス要素の異なる値に対して、異なるハフマンコードブックを選択することができる。いくつかの例では、量子化ユニット52は、Nb i t s Qシンタックス

要素値 6, . . . , 15 に対して異なるハフマンコーディングテーブルを提供することができる。その上、量子化ユニット 52 は、全体で 50 個のハフマンコードブックに対して、6, . . . , 15 にわたる異なる N b i t s Q シンタックス要素値の各々に対する 5 個の異なるハフマンコードブックを含み得る。この点において、量子化ユニット 52 は、いくつかの異なる統計的な状況において c i d のコーディングに対処するための、複数の異なるハフマンコードブックを含み得る。

【0233】

[0124]説明するために、量子化ユニット 52 は、N b i t s Q シンタックス要素値の各々に対して、1 から 4 までのベクトル要素をコーディングするための第 1 のハフマンコードブックと、5 から 9 までのベクトル要素をコーディングするための第 2 のハフマンコードブックと、9 以上のベクトル要素をコーディングするための第 3 のハフマンコードブックとを含み得る。これらの最初の 3 つのハフマンコードブックは、圧縮されるべき低減されたフォアグラウンド V [k] ベクトル 55 のうちの 1 つが低減されたフォアグラウンド V [k] ベクトル 55 の時間的に後続の対応するベクトルから予測されず、合成オーディオオブジェクト（たとえば、パルス符号変調（PCM）されたオーディオオブジェクトによって最初に定義されたもの）の空間情報を表さないとき、使用され得る。量子化ユニット 52 は追加で、N b i t s Q シンタックス要素値の各々に対して、低減されたフォアグラウンド V [k] ベクトル 55 のうちの 1 つが低減されたフォアグラウンド V [k] ベクトル 55 の時間的に後続の対応するベクトルから予測されるとき、低減されたフォアグラウンド V [k] ベクトル 55 のうちのこの 1 つをコーディングするための第 4 のハフマンコードブックを含み得る。量子化ユニット 52 はまた、N b i t s Q シンタックス要素値の各々に対して、低減されたフォアグラウンド V [k] ベクトル 55 のうちの 1 つが合成オーディオオブジェクトを表すとき、低減されたフォアグラウンド V [k] ベクトル 55 のうちのこの 1 つをコーディングするための第 5 のハフマンコードブックを含み得る。様々なハフマンコードブックが、これらの異なる統計的な状況の各々に対して、すなわちこの例では、予測されず合成ではない状況、予測される状況、および合成の状況に対して、開発され得る。

【0234】

[0125]以下の表は、ハフマンテーブルの選択と、解凍ユニットが適切なハフマンテーブルを選択することを可能にするためにビットストリーム中で指定されるべきビットとを示す。

【0235】

【表 1】

Pred モード	HT 情報	HTテーブル
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT5

【0236】

上記の表において、予測モード（「P r e d モード」）は、現在のベクトルに対して予測

が実行されたか否かを示し、一方でハフマンテーブル（「HT情報」）は、ハフマンテーブル1から5のうちの1つを選択するために使用される追加のハフマンコードブック（またはテーブル）情報を示す。予測モードはまた、以下で説明されるP F l a gシンタックス要素としても表され得、一方でHT情報は、以下で説明されるC b F l a gシンタックス要素によって表され得る。

【0237】

[0126]以下の表はさらに、様々な統計的な状況またはシナリオが与えられたときのこのハフマンテーブルの選択プロセスを示す。

【0238】

【表2】

	録音	合成
Predなし	HT{1,2,3}	HT5
Predあり	HT4	HT5

【0239】

上記の表において、「録音」列は、ベクトルが録音されたオーディオオブジェクトを表すときのコーディング状況を示し、一方で「合成」列は、ベクトルが合成オーディオオブジェクトを表すときのコーディング状況を示す。「Predなし」行は、予測がベクトル要素に関して実行されないときのコーディング状況を示し、一方で「Predあり」行は、予測がベクトル要素に関して実行されるときにのコーディング状況を示す。この表に示されるように、量子化ユニット52は、ベクトルが録音されたオーディオオブジェクトを表し予測がベクトル要素に関して実行されないとき、HT{1, 2, 3}を選択する。量子化ユニット52は、オーディオオブジェクトが合成オーディオオブジェクトを表し予測がベクトル要素に関して実行されないとき、HT5を選択する。量子化ユニット52は、ベクトルが録音されたオーディオオブジェクトを表し予測がベクトル要素に関して実行されるとき、HT4を選択する。量子化ユニット52は、オーディオオブジェクトが合成オーディオオブジェクトを表し予測がベクトル要素に関して実行されるとき、HT5を選択する。

【0240】

[0127]量子化ユニット52は、本開示で説明される基準の任意の組合せに基づいて、出力切替えされ量子化されたVベクトルとして使用するために、予測されないベクトル量子化されたVベクトル、予測されベクトル量子化されたVベクトル、ハフマンコーディングされないスカラー量子化されたVベクトル、およびハフマンコーディングされスカラー量子化されたVベクトルのうちの1つを選択することができる。いくつかの例では、量子化ユニット52は、ベクトル量子化モードと1つまたは複数のスカラー量子化モードとを含む、量子化モードのセットから量子化モードを選択し、選択されたモードに基づいて（または従って）、入力Vベクトルを量子化することができる。量子化ユニット52は次いで、（たとえば、重み値またはそれを示すビットに関して）予測されないベクトル量子化されたVベクトル、（たとえば、誤差値またはそれを示すビットに関して）予測されベクトル量子化されたVベクトル、ハフマンコーディングされないスカラー量子化されたVベクトル、およびハフマンコーディングされスカラー量子化されたVベクトルのうちの選択されたものを、コーディングされたフォアグラウンドV[k]ベクトル57としてビットストリーム生成ユニット52に与えることができる。量子化ユニット52はまた、量子化モ

ードを示すシンタックス要素（たとえば、N b i t s Qシンタックス要素）と、図4および図7の例に関して以下でより詳細に説明されるように、Vベクトルを逆量子化またはさなければ再構成するために使用される任意の他のシンタックス要素とを与えることができる。

【0241】

[0128]オーディオ符号化デバイス20内に含まれる聴覚心理オーディオコーダユニット40は、聴覚心理オーディオコーダの複数のインスタンスを表し得、これらの各々は、符号化された環境H O A係数59と符号化されたn F G信号61とを生成するために、エネルギー補償された環境H O A係数47'および補間されたn F G信号49'の各々の異なるオーディオオブジェクトまたはH O Aチャネルを符号化するために使用される。聴覚心理オーディオコーダユニット40は、符号化された環境H O A係数59と符号化されたn F G信号61とをビットストリーム生成ユニット42に出力することができる。

【0242】

[0129]オーディオ符号化デバイス20内に含まれるビットストリーム生成ユニット42は、既知のフォーマット（復号デバイスによって知られているフォーマットを指し得る）に適合するようにデータをフォーマットし、それによってベクトルベースのビットストリーム21を生成するユニットを表す。ビットストリーム21は、言い換えれば、上記で説明された方法で符号化されている、符号化されたオーディオデータを表し得る。ビットストリーム生成ユニット42は、いくつかの例ではマルチプレクサを表してよく、マルチプレクサは、コーディングされたフォアグラウンドV[k]ベクトル57と、符号化された環境H O A係数59と、符号化されたn F G信号61と、バックグラウンドチャネル情報43とを受信することができる。ビットストリーム生成ユニット42は次いで、コーディングされたフォアグラウンドV[k]ベクトル57と、符号化された環境H O A係数59と、符号化されたn F G信号61と、バックグラウンドチャネル情報43とに基づいて、ビットストリーム21を生成することができる。このようにして、ビットストリーム生成ユニット42は、図7の例に関してより詳細に以下で説明されるようにビットストリーム21を取得するために、ビットストリーム21におけるベクトル57をそれによって指定し得る。ビットストリーム21は、主要またはメインビットストリームと、1つまたは複数のサイドチャネルビットストリームとを含み得る。

【0243】

[0130]図3の例には示されないが、オーディオ符号化デバイス20はまた、現在のフレームが指向性ベース合成を使用して符号化されるべきであるかベクトルベース合成を使用して符号化されるべきであるかに基づいて、オーディオ符号化デバイス20から出力されるビットストリームを（たとえば、指向性ベースのビットストリーム21とベクトルベースのビットストリーム21との間で）切り替える、ビットストリーム出力ユニットを含み得る。ビットストリーム出力ユニットは、（H O A係数11が合成オーディオオブジェクトから生成されたことを検出した結果として）指向性ベース合成が実行されたか、または（H O A係数が録音されたことを検出した結果として）ベクトルベース合成が実行されたかを示す、コンテンツ分析ユニット26によって出力されるシンタックス要素に基づいて、切替えを実行することができる。ビットストリーム出力ユニットは、ビットストリーム21の各々とともに現在のフレームのために使用される切替えまたは現在の符号化を示すために、正しいヘッダシンタックスを指定することができる。

【0244】

[0131]その上、上述されたように、音場分析ユニット44は、フレームごとに変化し得る、B G_{TOT}環境H O A係数47を特定することができる（が、時々、B G_{TOT}は、2つ以上の（時間的に）隣接するフレームにわたって一定または同じままであり得る）。B G_{TOT}における変化は、低減されたフォアグラウンドV[k]ベクトル55において表された係数への変化を生じ得る。B G_{TOT}における変化は、フレームごとに変化する（「環境H O A係数」と呼ばれることもある）バックグラウンドH O A係数を生じ得る（が、この場合も時々、B G_{TOT}は、2つ以上の（時間的に）隣接するフレームにわたって一定または

同じままであり得る)。この変化は、追加の環境H O A係数の追加または除去と、対応する、低減されたフォアグラウンドV [k] ベクトル5 5からの係数の除去またはそれに対する係数の追加とによって表される、音場の態様のためのエネルギーの変化を生じることが多い。

【0 2 4 5】

[0132]結果として、音場分析ユニット音場分析ユニット4 4は、いつ環境H O A係数がフレームごとに変化するかをさらに決定し、音場の環境成分を表すために使用されることに関して、環境H O A係数への変化を示すフラグまたは他のシンタックス要素を生成することができる(ここで、この変化はまた、環境H O A係数の「遷移」または環境H O A係数の「遷移」と呼ばれることもある)。具体的には、係数低減ユニット4 6は、(Amb Coeff TransitionフラグまたはAmb Coeff Idx Transitionフラグとして示され得る)フラグを生成し、そのフラグが(場合によってはサイドチャンネル情報の一部として)ビットストリーム2 1中に含まれ得るように、そのフラグをビットストリーム生成ユニット4 2に与えることができる。

【0 2 4 6】

[0133]係数低減ユニット4 6は、環境係数遷移フラグを指定することに加えて、低減されたフォアグラウンドV [k] ベクトル5 5が生成される方法を修正することもできる。一例では、環境H O A環境係数のうちの1つが現在のフレームの間に遷移中であると決定すると、係数低減ユニット4 6は、遷移中の環境H O A係数に対応する低減されたフォアグラウンドV [k] ベクトル5 5のVベクトルの各々について、(「ベクトル要素」または「要素」とも呼ばれ得る)ベクトル係数を指定することができる。この場合も、遷移中の環境H O A係数は、BG_{TOT}からバックグラウンド係数の総数を追加または除去し得る。したがって、バックグラウンド係数の総数において生じた変化は、環境H O A係数がビットストリーム中に含まれるか含まれないか、および、Vベクトルの対応する要素が、上記で説明された第2の構成モードおよび第3の構成モードにおいてビットストリーム中で指定されたVベクトルのために含まれるか否かに影響を及ぼす。係数低減ユニット4 6が、エネルギーにおける変化を克服するために、低減されたフォアグラウンドV [k] ベクトル5 5を指定することができる方法に関するより多くの情報は、2015年1月12日に出願された「TRANSITIONING OF AMBIENT HIGHER ORDER AMBISONIC COEFFICIENTS」という名称の米国特許出願第14/594, 533号において提供されている。

【0 2 4 7】

[0134]図4は、図2のオーディオ復号デバイス2 4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、オーディオ復号デバイス2 4は、抽出ユニット7 2と、指向性ベース再構成ユニット9 0と、ベクトルベース再構成ユニット9 2とを含み得る。以下で説明されるが、オーディオ復号デバイス2 4に関するより多くの情報、およびH O A係数を解凍またはさもなければ復号する様々な態様は、2014年5月29日に出願された「INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の国際特許出願公開第WO2014/194099号において入手可能である。

【0 2 4 8】

[0135]抽出ユニット7 2は、ビットストリーム2 1を受信し、H O A係数1 1の様々な符号化されたバージョン(たとえば、指向性ベースの符号化されたバージョンまたはベクトルベースの符号化されたバージョン)を抽出するように構成されたユニットを表し得る。抽出ユニット7 2は、H O A係数1 1が様々な方向ベースのバージョンを介して符号化されたか、ベクトルベースのバージョンを介して符号化されたかを示す、上述されたシンタックス要素から決定することができる。指向性ベース符号化が実行されたとき、抽出ユニット7 2は、H O A係数1 1の指向性ベースのバージョンと、符号化されたバージョンに関連付けられたシンタックス要素(図4の例では指向性ベース情報9 1として示される)とを抽出し、指向性ベース情報9 1を指向性ベース再構成ユニット9 0に渡すことがで

きる。指向性ベース再構成ユニット90は、指向性ベース情報91に基づいてH O A 係数11'の形態でH O A 係数を再構成するように構成されたユニットを表し得る。ビットストリームおよびビットストリーム内のシンタックス要素の構成が、以下で図7A～図7Jの例に関してより詳細に説明される。

【0249】

[0136] H O A 係数11がベクトルベース合成を使用して符号化されたことをシンタックス要素が示すとき、抽出ユニット72は、コーディングされたフォアグラウンドV[k]ベクトル57（コーディングされた重み57および／もしくはインデックス63またはスカラー量子化されたVベクトルを含み得る）と、符号化された環境H O A 係数59と、対応するオーディオオブジェクト61とを抽出することができる。オーディオオブジェクト61各々は、ベクトル57のうちの1つに対応する。抽出ユニット72は、コーディングされたフォアグラウンドV[k]ベクトル57をVベクトル再構成ユニット74に渡し、符号化された環境H O A 係数59を符号化されたn F G 信号61とともに聴覚心理復号ユニット80に渡すことができる。

【0250】

[0137] コーディングされたフォアグラウンドV[k]ベクトル57を抽出するために、抽出ユニット72は、次のChannel Side Info Data (CSID) シンタックステーブルに従って、シンタックス要素を抽出することができる。

【0251】

【表 3 - 1】

テーブル ChannelSideInfoData (i) のシンタックス

シンタックス	ビットの数	ニーモニック
ChannelSideInfoData(i)		
{		
ChannelType[i]	2	uimsbf
switch ChannelType[i]		
{		
case 0:		
ActiveDirIds[i];	NumOfBits PerDirIdx	uimsbf
break;		
case 1:		
bA;	1	bslbf
bB;	1	bslbf
if ((bA + bB) == 0) {		
NbitsQ(k)[i] = NbitsQ(k-1)[i];		
PFlag(k)[i] = PFlag(k-1)[i];		
CbFlag(k)[i] = CbFlag(k-1)[i];		
CodebkIdx(k)[i] = CodebkIdx(k-1)[i];		
NumVecIndices(k)[i] = NumVecIndices[k-1][i];		
}		
else{		
NbitsQ(k)[i] = (8*bA)+(4*bB)+uintC;	2	uimsbf
if (NbitsQ(k)[i] == 4) {		
PFlag(k)[i];	1	bslbf
CodebkIdx(k)[i];	3	uimsbf

【 0 2 5 2 】

【表 3 - 2】

NumVecIndices(k)[i]++;	NumVVec VqElemen tsBits	uimbsf
}		
elseif (NbtsQ(k)[i] >= 6) {		
PFlag(k)[i];	1	bslbf
CbFlag(k)[i];	1	bslbf
}		
}		
}		
break;		
case 2:		
AddAmbHoalInfoChannel(i);		
break;		
default:		
}		
}		

【0253】

[0138] 上記のテーブルのためのセマンティクスは、次のようになる。このペイロードは、*i* 番目のチャンネルのためのサイド情報を保持する。ペイロードのサイズおよびデータは、チャンネルのタイプに依存する。**ChannelType[i]** この要素は、テーブル 95 において定義されている *i* 番目のチャンネルのタイプを記憶する。**ActiveDirIds[i]** この要素は、Annex F. 7 からの 900 のあらかじめ定義された一様に分布した点のインデックスを使用して、アクティブな指向性信号の方向を示す。コードワード 0 は、指向性信号の終了をシグナリングするために使用される。**PFlag[i]** *i* 番目のチャンネルのベクトルベース信号に関連付けられた《スカラー量子化された V ベクトルのハフマン復号のために使用される》予測フラグ（《》内は、取り消し線付である）。**CbFlag[i]** *i* 番目のチャンネルのベクトルベース信号に関連付けられたスカラー量子化された V ベクトルのハフマン復号のために使用されるコードブックフラグ。**CodebkIdx[i]** *i* 番目のチャンネルのベクトルベース信号に関連付けられたベクトル量子化された V ベクトルを逆量子化するために使用される特定のコードブックをシグナリングする。**NbtsQ[i]**

このインデックスは、*i* 番目のチャンネルのベクトルベース信号に関連付けられたデータのハフマン復号のために使用されるハフマンテーブルを決定する。コードワード 5 は、一様 8 ビット逆量子化器の使用を決定する。2 つの MSB00 は、以前のフレーム (*k* - 1) の **NbtsQ[i]** データと、**PFlag[i]** データと、**CbFlag[i]** データとを再使用することを決定する。**bA**、**bB** **NbtsQ[i]** フィールドの **msb(bA)** および第 2 の **msb(bB)**。**uintC** **NbtsQ[i]** フィールドの残りの 2 ビットのコードワード。

NumVecIndices ベクトル量子化された V ベクトルを逆量子化するために使用されるベクトルの数。**AddAmbHoalInfoChannel(i)**

このペイロードは、追加の環境 HOA 係数のための情報を保持する。

【0254】

10

20

30

40

50

[0139] C S I Dシンタックステーブルに従って、抽出ユニット72は、最初に、チャンネルのタイプを示すChannel Typeシンタックス要素を取得することができる（たとえば、ここで、0の値は指向性ベース信号をシグナリングし、1の値はベクトルベース信号をシグナリングし、2の値は追加の環境HOA信号をシグナリングする）。Channel Typeシンタックス要素に基づいて、抽出ユニット72は、この3つのcase間で切り替えることができる。

【0255】

[0140] 本開示で説明される技法の一例について説明するために、case 1に焦点を合わせると、抽出ユニット72は、Nb i t s Qシンタックス要素の最上位ビット（すなわち、上記の例示的なC S I Dシンタックステーブルにおけるb Aシンタックス要素）と、Nb i t s Qシンタックス要素の第2の最上位ビット（すなわち上記の例示的なC S I Dシンタックステーブルにおけるb Bシンタックス要素）とを取得することができる。Nb i t s Q (k) [i] の (k) [i] は、Nb i t s Qシンタックス要素がi番目のトランスポートチャンネルのk番目のフレームについて取得されることを示す。Nb i t s Qシンタックス要素は、HOA係数11によって表される音場の空間成分を量子化するために使用された量子化モードを示す1つまたは複数のビットを表し得る。空間成分はまた、本開示でVベクトルと呼ばれることもあり、またはコーディングされたフォアグラウンドV [k] ベクトル57と呼ばれることもある。

【0256】

[0141] 上記の例示的なC S I Dシンタックステーブルでは、Nb i t s Qシンタックス要素は、対応するV V e c D a t aフィールドにおいて指定されるベクトルを圧縮するために使用される（Nb i t s Qシンタックス要素のための0から3の値が予約済みまたは未使用であるので）12個の量子化モードのうちの1つを示すために4ビットを含み得る。12個の量子化モードは、以下で示された次のものを含む。

0 ~ 3 : 予約済み

4 : ベクトル量子化

5 : ハフマンコーディングなしのスカラー量子化

6 : ハフマンコーディングありの6ビットスカラー量子化

7 : ハフマンコーディングありの7ビットスカラー量子化

8 : ハフマンコーディングありの8ビットスカラー量子化

... ..

16 : ハフマンコーディングありの16ビットスカラー量子化

上記では、6 ~ 16のNb i t s Qシンタックス要素の値は、スカラー量子化がハフマンコーディングとともに実行されるべきであることのみではなく、スカラー量子化の量子化ステップサイズをも示す。この点において、量子化モードは、ベクトル量子化モードと、ハフマンコーディングなしのスカラー量子化モードと、ハフマンコーディングありのスカラー量子化モードとを備えることができる。

【0257】

[0142] 上記の例示的なC S I Dシンタックステーブルに戻ると、抽出ユニット72は、b Aシンタックス要素をb Bシンタックス要素と組み合わせることができ、ここで、この組合せは、上記の例示的なC S I Dシンタックステーブルにおいて示されるような加算であり得る。組み合わされたb A / b Bシンタックス要素は、以前のフレームから、ベクトルを圧縮するとき使用される情報を示す少なくとも1つのシンタックス要素を再利用するかどうかのためのインジケータを表すことができる。抽出ユニット72は次に、組み合わされたb A / b Bシンタックス要素を0の値と比較する。組み合わされたb A / b Bシンタックス要素が0の値を有するとき、抽出ユニット72は、i番目のトランスポートチャンネルの現在のk番目のフレームのための量子化モード情報（すなわち、上記の例示的なC S I Dシンタックステーブルにおける量子化モードを示すNb i t s Qシンタックス要素）が、i番目のトランスポートチャンネルのk - 1番目のフレームの量子化モード情報と同じであると決定することができる。言い換えれば、ゼロ値に設定されるとき、インジケー

10

20

30

40

50

タは、以前のフレームから少なくとも1つのシンタックス要素を再利用するように示す。

【0258】

[0143]抽出ユニット72は、同様に、 i 番目のトランスポートチャネルの現在の k 番目のフレームのための予測情報（すなわち、この例では、予測がベクトル量子化またはスカラー量子化のいずれかの間に実行されるか否かを示すPFlagシンタックス要素）が、 i 番目のトランスポートチャネルの $k-1$ 番目のフレームの予測情報と同じであると決定する。抽出ユニット72はまた、 i 番目のトランスポートチャネルの現在の k 番目のフレームのためのハフマンコードブック情報（すなわち、Vベクトルを再構成するために使用されるハフマンコードブックを示すCbF1agシンタックス要素）が、 i 番目のトランスポートチャネルの $k-1$ 番目のフレームのハフマンコードブック情報と同じであると決定することができる。抽出ユニット72はまた、 i 番目のトランスポートチャネルの現在の k 番目のフレームのためのベクトル量子化情報（すなわち、Vベクトルを再構成するために使用されるベクトル量子化コードブックを示すCodedkIdxシンタックス要素およびVベクトルを再構成するために使用されるコード化ベクトルの数を示すNumVecIndicesシンタックス要素）が、 i 番目のトランスポートチャネルの $k-1$ 番目のフレームのベクトル量子化情報と同じであると決定することができる。

10

【0259】

[0144]組み合わされたbA/bBシンタックス要素が0の値を有していないとき、抽出ユニット72は、 i 番目のトランスポートチャネルの k 番目のフレームのための量子化モード情報と、予測情報と、ハフマンコードブック情報と、ベクトル量子化情報とが、 i 番目のトランスポートチャネルの $k-1$ 番目のフレームの情報と同じではないと決定することができる。結果として、抽出ユニット72は、NbtsQシンタックス要素の最下位ビット（すなわち、上記の例示的なCSIDシンタックステーブルにおけるuintCシンタックス要素）を取得し、NbtsQシンタックス要素を取得するために、bAシンタックス要素とbBシンタックス要素とuintCシンタックス要素とを組み合わせることができる。このNbtsQシンタックス要素に基づいて、抽出ユニット72は、NbtsQシンタックス要素がベクトル量子化をシグナリングするとき、PFlagシンタックス要素、CodedkIdxシンタックス要素、およびNumVecIndicesシンタックス要素と、または、NbtsQシンタックス要素がハフマンコーディングありのスカラー量子化をシグナリングするとき、PFlagシンタックス要素とCbF1agシンタックス要素とのいずれかを取得することができる。このようにして、抽出ユニット72は、Vベクトルを再構成するために使用される上記のシンタックス要素を抽出し、これらのシンタックス要素をベクトルベース再構成ユニット72に渡すことができる。

20

30

【0260】

[0145]抽出ユニット72は次に、 i 番目のトランスポートチャネルの k 番目のフレームからVベクトルを抽出することができる。抽出ユニット72は、CodedVVecLengthと示されたシンタックス要素を含む、HOADecoderConfigコンテナを取得することができる。抽出ユニット72は、HOADecoderConfigコンテナからCodedVVecLengthを解析することができる。抽出ユニット72は、次のVVecDataシンタックステーブルに従って、Vベクトルを取得することができる。

40

【0261】

【表 4 - 1】

シンタックス	ビットの数	ニーモニック
VVectorData(i)		
{		
if (NbitsQ(k)[i] == 4){		
if (NumVecIndices(k)[i] == 1) {		
VecIdx[0] = VecIdx + 1;	10	uimsbf
WeightVal[0] = ((SgnVal*2)-1);	1	uimsbf
} else {		
WeightIdx;	nbitsW	uimsbf
nbitsIdx = ceil(log2(NumOfHoaCoeffs));		
for (j=0; j< NumVecIndices(k)[i]; ++j) {		
VecIdx[j] = VecIdx + 1;	nbitsIdx	uimsbf
if (PFlag[i] == 0) {		
tmpWeightVal(k) [j] = WeightValCdbk[CodebkIdx(k)[i]][WeightIdx][j];		
else {		

10

20

【 0 2 6 2 】

【表 4 - 2】

tmpWeightVal(k) [j] = WeightValPredCdbk[CodebkIdx(k)[i]][WeightIdx[j]] + WeightValAlpha[j] * tmpWeightVal(k-1) [j];		
}		
WeightVal[j] = ((SgnVal*2)-1) * tmpWeightVal(k) [j];	1	uimbsf
}		
}		
else if (NbitsQ(k)[i] == 5) {		
for (m=0; m< VVecLength; ++m)		
aVal[i][m] = (VecVal / 128.0) - 1.0;	8	uimbsf
}		
else if(NbitsQ(k)[i] >= 6) {		
for (m=0; m< VVecLength; ++m){		
huffIdx = huffSelect(VVecCoeffId[m], PFlag[i], CbFlag[i]);		
cid = huffDecode(NbitsQ[i], huffIdx, huffVal);	dynamic	huffDecode
aVal[i][m] = 0.0;		
if (cid > 0) {		
aVal[i][m] = sgn = (SgnVal * 2) - 1;	1	bslbf
if (cid > 1) {		
aVal[i][m] = sgn * (2.0^(cid - 1)) + intAddVal);	cid-1	uimbsf
}		
}		
}		
}		
}		
注: VVecLengthの計算については、 Error! Reference source not found. を参照		

【0263】

VVec(k) [i] これは、i 番目のチャンネルのための k 番目の H O A f r a m e () のための V ベクトルである。VVecLength この変数は、読み出すべきベクトル要素の数を示す。VVecCoeffId このベクトルは、送信された V ベクトル係数のインデックスを含む。VecVal 0 と 255 との間の整数値。aVal V V e c t o r D a t a の復号中に使用される一時的な変数。huffVal ハフマン復号されるべきハフマンコードワード。SgnVal これは、復号中に使用されるコード化された符号値である。intAddVal これは、復号中に使用される追加の整数値である。NumVecIndices ベクトル量子化された V ベクトルを逆量子化するために使用されるベクトルの数。WeightIdx ベクトル量子化された V ベクトルを逆量子化するために

使用される `WeightValCdbk` におけるインデックス。 `nBitsW`

ベクトル量子化された `V` ベクトルを復号するために、`WeightIdx` を読み取るためのフィールドサイズ。 `WeightValCbk` 正の実数値の重み付け係数のベクトルを含むコードブック。 `NumVecIndices` が 1 よりも大きい場合のみ必要。 256 個のエントリをもつ `WeightValCdbk` が与えられる。 `WeightValPredCdbk` 予測重み付け係数のベクトルを含むコードブック。 `NumVecIndices` が 1 よりも大きい場合のみ必要。 256 個のエントリをもつ `WeightValPredCdbk` が与えられる。 `WeightValAlpha` `V` ベクトル量子化の予測コーディングモードのために使用される予測コーディング係数。 `VvecIdx` ベクトル量子化された `V` ベクトルを逆量子化するために使用される、`VecDict` のためのインデックス。 `nbitsIdx` ベクトル量子化された `V` ベクトルを復号するために、`VvecIdx` を読み取るためのフィールドサイズ。

`WeightVal` ベクトル量子化された `V` ベクトルを復号するための実数値の重み付け係数。

【0264】

[0146] 上記のシンタックステーブルでは、抽出ユニット 72 は、`NbitsQ` シンタックス要素の値が 4 に等しい（または、言い換えれば、ベクトル逆量子化が `V` ベクトルを再構成するために使用されることをシグナリングする）か否かを決定することができる。 `NbitsQ` シンタックス要素の値が 4 に等しいとき、抽出ユニット 72 は、`NumVecIndices` シンタックス要素の値を 1 の値と比較することができる。 `NumVecIndices` の値が 1 に等しいとき、抽出ユニット 72 は、`VecIdx` シンタックス要素を取得することができる。 `VecIdx` シンタックス要素は、ベクトル量子化された `V` ベクトルを逆量子化するために使用される `VecDict` のためのインデックスを示す 1 つまたは複数のビットを表し得る。抽出ユニット 72 は、0 番目の要素が `VecIdx` シンタックス要素の値 + 1 に設定された、`VecIdx` アレイをインスタンス化することができる。抽出ユニット 72 はまた、`SgnVal` シンタックス要素を取得することができる。 `SgnVal` シンタックス要素は、`V` ベクトルの復号中に使用されるコーディングされた符号値を示す 1 つまたは複数のビットを表し得る。抽出ユニット 72 は、`WeightVal` アレイをインスタンス化し、`SgnVal` シンタックス要素の値の関数として 0 番目の要素を設定することができる。

【0265】

[0147] `NumVecIndices` シンタックス要素の値が 1 の値に等しくないとき、抽出ユニット 72 は、`WeightIdx` シンタックス要素を取得することができる。 `WeightIdx` シンタックス要素は、ベクトル量子化された `V` ベクトルを逆量子化するために使用される `WeightValCdbk` アレイにおけるインデックスを示す 1 つまたは複数のビットを表し得る。 `WeightValCdbk` アレイは、正の実数値の重み付け係数のベクトルを含むコードブックを表し得る。抽出ユニット 72 は次に、`HOAConfig` コンテナにおいて指定された（一例として、ビットストリーム 21 の開始において指定された）`NumOfHoacoeffs` シンタックス要素の関数として、`nbitsIdx` を決定することができる。抽出ユニット 72 は次いで、`NumVecIndices` 中を反復し、ビットストリーム 21 から `VecIdx` シンタックス要素を取得し、各取得された `VecIdx` シンタックス要素を用いて `VecIdx` アレイ要素を設定することができる。

【0266】

[0148] 抽出ユニット 72 は、ビットストリーム 21 からのシンタックス要素の抽出に無関係である `tmpWeightVal` 変数値を決定することを伴う、次の `PFlag` シンタックス比較を実行しない。したがって、抽出ユニット 72 は次に、`WeightVal` シンタックス要素を決定する際に使用するための `SgnVal` シンタックス要素を取得することができる。

【0267】

[0149] N b i t s Q シンタックス要素の値が 5 に等しい（ハフマン復号なしのスカラ逆量子化が V ベクトルを再構成するために使用されることをシグナリングする）とき、抽出ユニット 7 2 は、0 から V V e c L e n g t h まで反復し、a V a l 変数を、ビットストリーム 2 1 から取得された V e c V a l シンタックス要素に設定する。V e c V a l シンタックス要素は、0 と 2 5 5 との間の整数を示す 1 つまたは複数のビットを表し得る。

【0268】

[0150] N b i t s Q シンタックス要素の値が 6 以上である（ハフマン復号ありの N b i t s Q ビットスカラ逆量子化が V ベクトルを再構成するために使用されることをシグナリングする）とき、抽出ユニット 7 2 は、0 から V V e c L e n g t h まで反復し、h u f f V a l シンタックス要素、S g n V a l シンタックス要素、および i n t A d d V a l シンタックス要素のうちの 1 つまたは複数を取得する。h u f f V a l シンタックス要素は、ハフマンコードワードを示す 1 つまたは複数のビットを表し得る。i n t A d d V a l シンタックス要素は、復号中に使用される追加の整数値を示す 1 つまたは複数のビットを表し得る。抽出ユニット 7 2 は、これらのシンタックス要素をベクトルベース再構成ユニット 9 2 に与えることができる。

10

【0269】

[0151] ベクトルベース再構成ユニット 9 2 は、H O A 係数 1 1 ' を再構成するために、ベクトルベース合成ユニット 2 7 に関して上記で説明されたものとは逆の演算を実行するように構成されたユニットを表し得る。ベクトルベース再構成ユニット 9 2 は、V ベクトル再構成ユニット 7 4 と、空間時間的補間ユニット 7 6 と、フォアグラウンド編成ユニット 7 8 と、聴覚心理復号ユニット 8 0 と、H O A 係数編成ユニット 8 2 と、フェードユニット 7 7 0 と、並べ替えユニット 8 4 とを含み得る。フェードユニット 7 7 0 の破線は、ベクトルベース再構成ユニット 9 2 中に含まれているという観点から見て、フェードユニット 7 7 0 がオプションユニットであり得ることを示す。

20

【0270】

[0152] V ベクトル再構成ユニット 7 4 は、符号化されたフォアグラウンド V [k] ベクトル 5 7 から V ベクトルを再構成するように構成されたユニットを表し得る。V ベクトル再構成ユニット 7 4 は、量子化ユニット 5 2 の動作とは逆の方法で動作することができる。

。

【0271】

[0153] V ベクトル再構成ユニット 7 4 は、言い換えれば、V ベクトルを再構成するために次の擬似コードに従って動作することができる。

30

【0272】

【数 6 1】

```

if (NbitsQ(k)[i] == 4){
    if (NumVvecIndicies == 1){
        for (m=0; m< VVecLength; ++m){
            idx = VVecCoeffID[m];
             $v_{VVecCoeffID[m]}^{(i)}(k) = \text{WeightVal}[0] * \text{VecDict}[900].[\text{VecIdx}[0]][\text{idx}];$ 
        }
    } else {
        cdbLen = 0;
        if (N==4)
            cdbLen = 32;
        if
        for (m=0; m< 0; ++m){
            TmpVVec[m] = 0;
            for (j=0; j< NumVecIndicies; ++j){
                TmpVVec[m] += WeightVal[j] * VecDict[cdbLen].[VecIdx[j]][m];
            }
        }
        FNorm = 0.0;
        for (m=0; m < 0; ++ m) {
            FNorm += TmpVVec[m] * TmpVVec[m];
        }
        FNorm = (N+1)/sqrt(FNorm);
        for (m=0; m< VVecLength; ++m){
            idx = VVecCoeffID[m];
             $v_{VVecCoeffID[m]}^{(i)}(k) = \text{TmpVVec}[\text{idx}] * \text{FNorm};$ 
        }
    }
}
elseif (NbitsQ(k)[i] == 5){
    for (m=0; m< VVecLength; ++m){
         $v_{VVecCoeffID[m]}^{(i)}(k) = (N+1)*aVal[i][m];$ 
    }
}
elseif (NbitsQ(k)[i] >= 6){
    for (m=0; m< VVecLength; ++m){
         $v_{VVecCoeffID[m]}^{(i)}(k) = (N+1)*(2^{(16 - \text{NbitsQ}(k)[i])} * aVal[i][m]) / 2^{15};$ 
        if (PFlag(k)[i] == 1) {
             $v_{VVecCoeffID[m]}^{(i)}(k) += v_{VVecCoeffID[m]}^{(i)}(k - 1);$ 
        }
    }
}

```

【0 2 7 3】

[0154] 上記の擬似コードに従って、Vベクトル再構成ユニット74は、i番目のトラン 50

スポーツチャネルの k 番目のフレームのための N b i t s Q シンタックス要素を取得することができる。N b i t s Q シンタックス要素が 4 に等しい（この場合も、ベクトル量子化が実行されたことをシグナリングする）とき、V ベクトル再構成ユニット 7 4 は、N u m V e c I n d i c i e s シンタックス要素を 1 と比較することができる。N u m V e c I n d i c i e s シンタックス要素は、上記で説明されたように、ベクトル量子化された V ベクトルを逆量子化するために使用されるベクトルの数を示す 1 つまたは複数のビットを表し得る。N u m V e c I n d i c i e s シンタックス要素の値が 1 に等しいとき、V ベクトル再構成ユニット 7 4 は次いで、0 から V V e c L e n g t h シンタックス要素の値まで反復し、i d x 変数を V V e c C o e f f I d に設定し、V V e c C o e f f I d 番目の V ベクトル要素 ($v^{(i)}_{VVecCoeffId[m]}(k)$) を、[9 0 0] [V e c I d x [0]] [i d x] によって特定された V e c D i c t エントリを W e i g h t V a l に乗算したものに設定することができる。言い換えれば、N u m V e c I n d i c i e s の値が 1 に等しいとき、テーブル F. 1 1 において示された 8×1 重み付け値のコードブックとともに、テーブル F. 8 から導出されたベクトルコードブック H O A 拡張係数。

【0 2 7 4】

[0155] N u m V e c I n d i c i e s シンタックス要素の値が 1 に等しくないとき、V ベクトル再構成ユニット 7 4 は、c d b L e n 変数を、ベクトルの数を示す変数である O に設定することができる。c d b L e n シンタックス要素は、コードベクトルの辞書またはコードブックにおけるエントリを示す（ここで、この辞書は、上記の擬似コードにおいて「V e c D i c t」と示され、ベクトル量子化された V ベクトルを復号するために使用される、H O A 拡張係数のベクトルを含む c d b L e n 個のコードブックエントリをもつコードブックを表す）。H O A 係数 1 1 の（「N」によって示される）次数が 4 に等しいとき、V ベクトル再構成ユニット 7 4 は、c d b L e n 変数を 3 2 に設定することができる。V ベクトル再構成ユニット 7 4 は次に、0 から O まで反復し、T m p V V e c アレイを 0 に設定することができる。この反復中に、V ベクトル再構成ユニット 7 4 はまた、0 から N u m V e c I n d e c i e s シンタックス要素の値まで反復し、T e m p V V e c アレイの m 番目のエントリを、V e c D i c t の [c d b L e n] [V e c I d x [j]] [m] エントリを j 番目の W e i g h t V a l に乗算したものに等しくなるように設定することができる。

【0 2 7 5】

[0156] V ベクトル再構成ユニット 7 4 は、次の擬似コードに従って、W e i g h t V a l を導出することができる。

【0 2 7 6】

【数 6 2】

```
for (j=0; j< NumVecIndices(k)[i]; ++j) {
    if (PFlag[i] == 0) {
        tmpWeightVal(k) [j] =
            WeightValCdbk[CodebkIdx(k)[i]][WeightIdx[j]];
    }
    else {
        tmpWeightVal(k) [j] =
            WeightValPredCdbk[CodebkIdx(k)[i]][WeightIdx[j]]
            + WeightValAlpha[j] * tmpWeightVal(k-1) [j];
    }
    WeightVal[j] = ((SgnVal*2)-1) * tmpWeightVal(k) [j];
}
```

【0 2 7 7】

上記の擬似コードでは、Vベクトル再構成ユニット74は、0からNumVecIndicesシンタックス要素の値まで反復し、最初に、PFlagシンタックス要素の値が0に等しいか否かを決定することができる。PFlagシンタックス要素が0に等しいとき、Vベクトル再構成ユニット74は、tmpWeightVal変数を決定し、tmpWeightVal変数を、WeightValCdbkコードブックの[CodebkIdx][WeightIdx]エントリに等しく設定することができる。PFlagシンタックス要素の値が0に等しくないとき、Vベクトル再構成ユニット74は、tmpWeightVal変数を、WeightValPredCdbkコードブックの[CodebkIdx][WeightIdx]エントリ+i番目のトランスポートチャンネルのk-1番目のフレームのtempWeightValをWeightValAlpha変数に
10
乗算したものに等しく設定することができる。WeightValAlpha変数は、オーディオ符号化デバイス20およびオーディオ復号デバイス24において静的に定義され得る、上述された α 値を指し得る。Vベクトル再構成ユニット74は次いで、抽出ユニット72によって取得されたSgnValシンタックス要素とtmpWeightVal変数との関数として、WeightValを取得することができる。

【0278】

[0157] Vベクトル再構成ユニット74は、言い換えれば、重み値コードブック（予測されないベクトル量子化では「WeightValCdbk」と示され、予測ベクトル量子化では「WeightValPredCdbk」と示され、それらの両方は、コードブックインデックス（上記のVVectorData(i)シンタックステーブルにおいて「CodebkIdx」シンタックス要素と示される）および重みインデックス（上記のVVectorData(i)シンタックステーブルにおいて「WeightIdx」シンタックス要素と示される）のうちの1つまたは複数に基づいてインデックス付けされた多次元テーブルを表し得る）に基づいて、Vベクトルを再構成するために使用される各対応するコードベクトルのための重み値を導出することができる。このCodebkIdxシンタックス要素は、以下のChannelSideInfoData(i)シンタックス
20
テーブルにおいて示されるような、サイドチャンネル情報の一部分において定義され得る。

【0279】

[0158] 上記の擬似コードの残りのベクトル量子化部分は、Vベクトルの要素を正規化するためのFNormの計算と、後に続く、TmpVVec[idx]×FNormに等しいものとしてのVベクトル要素($v^{(i)}_{VVecCoeffId[m]}(k)$)の計算とに
30
関係する。Vベクトル再構成ユニット74は、VVecCoeffIDの関数として、idx変数
取得することができる。

【0280】

[0159] NbitsQが5に等しいとき、一様8ビットスカラー逆量子化が実行される。対照的に、6以上のNbitsQの値は、ハフマン復号の適用をもたらし得る。上で言及されるcid値は、NbitsQ値の下位2ビットに等しくてよい。予測モードは、上記のシンタックステーブルではPFlagとして示されるが、一方で、ハフマンテーブル情報ビットは、上記のシンタックステーブルではCbFlagとして示される。残りのシンタックスは、復号が上記で説明されたものと実質的に同様の方法でどのように行われるかを
40
指定する。

【0281】

[0160] 聴覚心理復号ユニット80は、符号化された環境HOA係数59と符号化されたnFG信号61とを復号し、それによってエネルギー補償された環境HOA係数47'と補間されたnFG信号49'（補間されたnFGオーディオオブジェクト49'とも呼ばれ得る）とを生成するために、図3の例に示される聴覚心理オーディオコードユニット40とは逆の方法で動作することができる。聴覚心理復号ユニット80は、エネルギー補償された環境HOA係数47'をフェードユニット770に渡し、nFG信号49'をフォアグラウンド編成ユニット78に渡すことができる。

【0282】

10

20

30

40

50

[0161]空間時間的補間ユニット76は、空間時間的補間ユニット50に関して上記で説明されたものと同様の方法で動作することができる。空間時間的補間ユニット76は、低減されたフォアグラウンドV[k]ベクトル 55_k を受信し、また、補間されたフォアグラウンドV[k]ベクトル $55_k'$ を生成するために、フォアグラウンドV[k]ベクトル 55_k および低減されたフォアグラウンドV[k-1]ベクトル 55_{k-1} に関して空間時間的補間を実行することができる。空間時間的補間ユニット76は、補間されたフォアグラウンドV[k]ベクトル $55_k'$ をフェードユニット770に転送することができる。

【0283】

[0162]抽出ユニット72はまた、いつ環境HOA係数のうちの1つが遷移中であることを示す信号757を、フェードユニット770に出力することもでき、フェードユニット770は次いで、SCH_{BG}47'（ここで、SCH_{BG}47'は、「環境HOAチャンネル47'」または「環境HOA係数47'」とも呼ばれ得る）および補間されたフォアグラウンドV[k]ベクトル $55_k'$ の要素のうちのいずれがフェードインまたはフェードアウトのいずれかを行われるべきであるかを決定することができる。いくつかの例では、フェードユニット770は、環境HOA係数47'および補間されたフォアグラウンドV[k]ベクトル $55_k'$ の要素の各々に関して、反対に動作することができる。すなわち、フェードユニット770は、環境HOA係数47'のうちの対応する1つに関して、フェードインもしくはフェードアウト、またはフェードインもしくはフェードアウトの両方を実行することができ、一方で、補間されたフォアグラウンドV[k]ベクトル $55_k'$ の要素のうちの対応する1つに関して、フェードインもしくはフェードアウト、またはフェードインとフェードアウトの両方を実行することができる。フェードユニット770は、調整された環境HOA係数47'をHOA係数編成ユニット82に出力し、調整されたフォアグラウンドV[k]ベクトル $55_k'$ をフォアグラウンド編成ユニット78に出力することができる。この点において、フェードユニット770は、HOA係数またはその派生物の様々な態様に関して、たとえば、環境HOA係数47'および補間されたフォアグラウンドV[k]ベクトル $55_k'$ の要素の形態で、フェード動作を実行するように構成されたユニットを表す。

【0284】

[0163]フォアグラウンド編成ユニット78は、フォアグラウンドHOA係数65を生成するために、調整されたフォアグラウンドV[k]ベクトル $55_k'$ および補間されたnFG信号49'に関して行列乗算を実行するように構成されたユニットを表し得る。この点において、フォアグラウンド編成ユニット78は、フォアグラウンド、または言い換えればHOA係数11'の支配的な態様を再構成するために、ベクトル $55_k'$ とオーディオオブジェクト49'（それは、補間されたnFG信号49'を示す別の方法である）を組み合わせることができる。フォアグラウンド編成ユニット78は、調整されたフォアグラウンドV[k]ベクトル $55_k'$ による補間されたnFG信号49'の行列乗算を実行することができる。

【0285】

[0164]HOA係数編成ユニット82は、HOA係数11'を取得するために、フォアグラウンドHOA係数65を調整された環境HOA係数47'に組み合わせるように構成されたユニットを表し得る。プライム表記法は、HOA係数11'がHOA係数11と同様であるが同じではないことがあることを反映している。HOA係数11とHOA係数11'との間の差分は、損失のある送信媒体を介した送信、量子化、または他の損失のある演算が原因の損失に起因し得る。

【0286】

[0165]図5Aは、本開示で説明されるベクトルベース合成技法の様々な態様を実行する際の、図3の例に示されるオーディオ符号化デバイス20などのオーディオ符号化デバイスの例示的な動作を示すフローチャートである。最初に、オーディオ符号化デバイス20は、HOA係数11を受信する(106)。オーディオ符号化デバイス20はLITユニ

ット 30 を呼び出すことができ、LIT ユニット 30 は、変換された H O A 係数（たとえば、SVD の場合、変換された H O A 係数は $U S[k]$ ベクトル 33 と $V[k]$ ベクトル 35 とを備え得る）を出力するために H O A 係数に関して LIT を適用することができる（107）。

【0287】

[0166] オーディオ符号化デバイス 20 は次に、上記で説明された方法で様々なパラメータを特定するために、 $U S[k]$ ベクトル 33、 $U S[k-1]$ ベクトル 33、 $V[k]$ ベクトルおよび／または $V[k-1]$ ベクトル 35 の任意の組合せに関して上記で説明された分析を実行するために、パラメータ計算ユニット 32 を呼び出すことができる。すなわち、パラメータ計算ユニット 32 は、変換された H O A 係数 33 / 35 の分析に基づいて少なくとも 1 つのパラメータを決定することができる（108）。

10

【0288】

[0167] オーディオ符号化デバイス 20 は次いで、並べ替えユニット 34 を呼び出すことができ、並べ替えユニット 34 は、上記で説明されたように、並べ替えられた変換された H O A 係数 33' / 35'（または言い換えれば、 $U S[k]$ ベクトル 33' および $V[k]$ ベクトル 35'）を生成するために、パラメータに基づいて、変換された H O A 係数（この場合も、SVD の文脈では、 $U S[k]$ ベクトル 33 と $V[k]$ ベクトル 35 とを指し得る）を並べ替えることができる（109）。オーディオ符号化デバイス 20 は、前述の演算または後続の演算のいずれかの間に、音場分析ユニット 44 を呼び出すこともできる。音場分析ユニット 44 は、上記で説明されたように、フォアグラウンドチャンネルの総数 (nFG) 45 と、バックグラウンド音場の次数 (N_{BG}) と、送るべき追加の B G H O A チャンネルの数 ($nBGa$) およびインデックス (i)（図 3 の例ではバックグラウンドチャンネル情報 43 としてまとめて示され得る）とを決定するために、H O A 係数 11 および／または変換された H O A 係数 33 / 35 に関して音場分析を実行することができる（109）。

20

【0289】

[0168] オーディオ符号化デバイス 20 はまた、バックグラウンド選択ユニット 48 を呼び出すことができる。バックグラウンド選択ユニット 48 は、バックグラウンドチャンネル情報 43 に基づいて、バックグラウンドまたは環境 H O A 係数 47 を決定することができる（110）。オーディオ符号化デバイス 20 はさらに、フォアグラウンド選択ユニット 36 を呼び出すことができ、フォアグラウンド選択ユニット 36 は、 nFG 45（フォアグラウンドベクトルを特定する 1 つまたは複数のインデックスを表し得る）に基づいて、音場のフォアグラウンド成分または明瞭な成分を表す、並べ替えられた $U S[k]$ ベクトル 33' と並べ替えられた $V[k]$ ベクトル 35' とを選択することができる（112）。

30

【0290】

[0169] オーディオ符号化デバイス 20 は、エネルギー補償ユニット 38 を呼び出すことができる。エネルギー補償ユニット 38 は、バックグラウンド選択ユニット 48 による H O A 係数のうちの様々なものの除去によるエネルギー損失を補償するために、環境 H O A 係数 47 に関してエネルギー補償を実行し（114）、それによって、エネルギー補償された環境 H O A 係数 47' を生成することができる。

40

【0291】

[0170] オーディオ符号化デバイス 20 はまた、空間時間的補間ユニット 50 を呼び出すことができる。空間時間的補間ユニット 50 は、補間されたフォアグラウンド信号 49'（「補間された nFG 信号 49'」とも呼ばれ得る）と残りのフォアグラウンド指向性情報 53（「 $V[k]$ ベクトル 53」とも呼ばれ得る）とを取得するために、並べ替えられた変換された H O A 係数 33' / 35' に関して空間時間的補間を実行することができる（116）。オーディオ符号化デバイス 20 は次いで、係数低減ユニット 46 を呼び出すことができる。係数低減ユニット 46 は、低減されたフォアグラウンド指向性情報 55（低減されたフォアグラウンド $V[k]$ ベクトル 55 とも呼ばれ得る）を取得するために、

50

バックグラウンドチャンネル情報 4 3 に基づいて残りのフォアグラウンド V [k] ベクトル 5 3 に関して係数低減を実行することができる (1 1 8) 。

【 0 2 9 2 】

[0171] オーディオ符号化デバイス 2 0 は次いで、上記で説明された方法で、低減されたフォアグラウンド V [k] ベクトル 5 5 を圧縮し、コーディングされたフォアグラウンド V [k] ベクトル 5 7 を生成するために、量子化ユニット 5 2 を呼び出すことができる (1 2 0) 。

【 0 2 9 3 】

[0172] オーディオ符号化デバイス 2 0 はまた、聴覚心理オーディオコーダユニット 4 0 を呼び出すことができる。聴覚心理オーディオコーダユニット 4 0 は、符号化された環境 H O A 係数 5 9 と符号化された n F G 信号 6 1 とを生成するために、エネルギー補償された環境 H O A 係数 4 7 ' および補間された n F G 信号 4 9 ' の各ベクトルを聴覚心理コーディングすることができる。オーディオ符号化デバイスは次いで、ビットストリーム生成ユニット 4 2 を呼び出すことができる。ビットストリーム生成ユニット 4 2 は、コーディングされたフォアグラウンド指向性情報 5 7 と、コーディングされた環境 H O A 係数 5 9 と、コーディングされた n F G 信号 6 1 と、バックグラウンドチャンネル情報 4 3 とに基づいて、ビットストリーム 2 1 を生成することができる。

【 0 2 9 4 】

[0173] 図 5 B は、本開示で説明されるコーディング技法を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャートである。図 3 の例において示されたオーディオ符号化デバイス 2 0 のビットストリーム生成ユニット 4 2 は、本開示で説明される技法を実行するように構成された 1 つの例示的なユニットを表し得る。ビットストリーム生成ユニット 4 2 は、フレームの量子化モードが、(「第 2 のフレーム」として示され得る) 時間的に以前のフレームの量子化モードと同じであるか否かを決定することができる (3 1 4) 。以前のフレームに関して説明されているが、本技法は、時間的に後続のフレームに関して実行され得る。フレームは、1 つまたは複数のトランスポートチャンネルの一部分を含み得る。トランスポートチャンネルの一部分は、あるペイロード (たとえば、図 7 の例では V V e c t o r D a t a フィールド 1 5 6) とともに (C h a n n e l S i d e I n f o D a t a シンタックステーブルに従って形成された) C h a n n e l S i d e I n f o D a t a を含み得る。ペイロードの他の例は、A d d A m b i e n t H O A C o e f f s フィールドを含み得る。

【 0 2 9 5 】

[0174] 量子化モードが同じであるとき (「YES」 3 1 6) 、ビットストリーム生成ユニット 4 2 は、ビットストリーム 2 1 中で量子化モードの一部分を指定することができる (3 1 8) 。量子化モードの一部分は、b A シンタックス要素と b B シンタックス要素とを含み得るが、u i n t C シンタックス要素を含まないことがある。b A シンタックス要素は、N b i t s Q シンタックス要素のビットストリームの最上位ビットを示す、ビットを表し得る。b B シンタックス要素は、N b i t s Q シンタックス要素の第 2 の最上位ビットを示す、ビットを表し得る。ビットストリーム生成ユニット 4 2 は、b A シンタックス要素および b B シンタックス要素の各々の値を 0 に設定し、それによって、ビットストリーム 2 1 中の量子化モードフィールド (すなわち、一例として N b i t s Q フィールド) が u i n t C シンタックス要素を含まないことをシグナリングすることができる。0 値の b A シンタックス要素および b B シンタックス要素のこのシグナリングはまた、以前のフレームからの N b i t s Q 値、P F l a g 値、C b F l a g 値、および C o d e b k I d x 値が、現在のフレームの同じシンタックス要素のための対応する値として使用されるべきであることを示す。

【 0 2 9 6 】

[0175] 量子化モードが同じではないとき (「NO」 3 1 6) 、ビットストリーム生成ユニット 4 2 は、ビットストリーム 2 1 中で全体量子化モードを示す 1 つまたは複数のビットを指定することができる (3 2 0) 。すなわち、ビットストリーム生成ユニット 4 2 は

、ビットストリーム 21 中で b A シンタックス要素と、b B シンタックス要素と、u i n t C シンタックス要素とを指定する。ビットストリーム生成ユニット 42 はまた、量子化モードに基づいて量子化情報を指定することができる (322)。この量子化情報は、ベクトル量子化情報、予測情報、およびハフマンコードブック情報など、量子化に関する任意の情報を含み得る。ベクトル量子化情報は、一例として、C o d e b k I d x シンタックス要素および N u m V e c I n d i c e s シンタックス要素のうち的一方または両方を含み得る。予測情報は、一例として、P F l a g シンタックス要素を含み得る。ハフマンコードブック情報は、一例として、C b F l a g シンタックス要素を含み得る。

【0297】

[0176] この点に関して、技法は、オーディオ符号化デバイス 20 がサウンドフィールドの空間成分の圧縮されたバージョンを備えるビットストリーム 21 を取得するように構成されることができ得る。空間成分は、複数の球面調和関数係数に関してベクトルベース合成を実行することによって生成され得る。ビットストリームは、空間成分を圧縮するときに使用される情報を指定する、以前のフレームからの、ヘッダフィールドの 1 つまたは複数のビットを再使用するかどうかのためのインジケータをさらに備える。

【0298】

[0177] 言い換えれば、技法は、オーディオ符号化デバイス 20 が球面調和関数領域における直交空間軸を表すベクトル 57 を備えるビットストリーム 21 を取得するように構成されることができ得る。ビットストリーム 21 は、ベクトルを圧縮 (たとえば、量子化) するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータ (たとえば、N b i t s Q シンタックス要素の b A / b B シンタックス要素) をさらに備え得る。

【0299】

[0178] 図 6 A 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 に示されるオーディオ復号デバイス 24 などのオーディオ復号デバイスの例示的な動作を示すフローチャートである。最初に、オーディオ復号デバイス 24 は、ビットストリーム 21 を受信することができる (130)。ビットストリームを受信すると、オーディオ復号デバイス 24 は抽出ユニット 72 を呼び出すことができる。説明の目的で、ベクトルベース再構成が実行されるべきであることをビットストリーム 21 が示すと仮定すると、抽出デバイス 72 は、上述された情報を取り出すためにビットストリームを解析し、その情報をベクトルベース再構成ユニット 92 に渡すことができる。

【0300】

[0179] 言い換えれば、抽出ユニット 72 は、コーディングされたフォアグラウンド指向性情報 57 (この場合も、コーディングされたフォアグラウンド V [k] ベクトル 57 と呼ばれ得る) と、コーディングされた環境 H O A 係数 59 と、コーディングされたフォアグラウンド信号 (コーディングされたフォアグラウンド n F G 信号 59 またはコーディングされたフォアグラウンドオーディオオブジェクト 59 と呼ばれ得る) とを、上記で説明された方法でビットストリーム 21 から抽出することができる (132)。

【0301】

[0180] オーディオ復号デバイス 24 はさらに、逆量子化ユニット 74 を呼び出すことができる。逆量子化ユニット 74 は、低減されたフォアグラウンド指向性情報 55_k を取得するために、コーディングされたフォアグラウンド指向性情報 57 をエントロピー復号および逆量子化することができる (136)。オーディオ復号デバイス 24 はまた、聴覚心理復号ユニット 80 を呼び出すことができる。聴覚心理オーディオ復号ユニット 80 は、エネルギー補償された環境 H O A 係数 47' と補間されたフォアグラウンド信号 49' とを取得するために、符号化された環境 H O A 係数 59 と符号化されたフォアグラウンド信号 61 とを復号することができる (138)。聴覚心理復号ユニット 80 は、エネルギー補償された環境 H O A 係数 47' をフェードユニット 770 に渡し、n F G 信号 49' をフォアグラウンド編成ユニット 78 に渡すことができる。

【0302】

[0181] オーディオ復号デバイス 24 は次に、空間時間的補間ユニット 76 を呼び出すことができる。空間時間的補間ユニット 76 は、並べ替えられたフォアグラウンド指向性情報 $55_k'$ を受信し、また、補間されたフォアグラウンド指向性情報 $55_k'$ を生成するために、低減されたフォアグラウンド指向性情報 $55_k/55_{k-1}$ に関して空間時間的補間を実行することができる (140)。空間時間的補間ユニット 76 は、補間されたフォアグラウンド $V[k]$ ベクトル $55_k'$ をフェードユニット 770 に転送することができる。

【0303】

[0182] オーディオ復号デバイス 24 は、フェードユニット 770 を呼び出すことができる。フェードユニット 770 は、エネルギー補償された環境 H O A 係数 $47'$ がいつ遷移中であるかを示すシンタックス要素 (たとえば、Amb Coeff Transition シンタックス要素) を (たとえば、抽出ユニット 72 から) 受信またはさもなければ取得することができる。フェードユニット 770 は、遷移シンタックス要素と維持された遷移状態情報とに基づいて、エネルギー補償された環境 H O A 係数 $47'$ をフェードインまたはフェードアウトし、調整された環境 H O A 係数 $47'$ を H O A 係数編成ユニット 82 に出力することができる。フェードユニット 770 はまた、シンタックス要素と維持された遷移状態情報とに基づいて、および、補間されたフォアグラウンド $V[k]$ ベクトル $55_k'$ の対応する 1 つまたは複数の要素をフェードアウトまたはフェードインし、フォアグラウンド編成ユニット 78 に調整されたフォアグラウンド $V[k]$ ベクトル $55_k'$ を出力することができる (142)。

【0304】

[0183] オーディオ復号デバイス 24 は、フォアグラウンド編成ユニット 78 を呼び出すことができる。フォアグラウンド編成ユニット 78 は、フォアグラウンド H O A 係数 65 を取得するために、調整されたフォアグラウンド指向性情報 $55_k'$ による行列乗算 nFG 信号 $49'$ を実行することができる (144)。オーディオ復号デバイス 24 はまた、H O A 係数編成ユニット 82 を呼び出すことができる。H O A 係数編成ユニット 82 は、H O A 係数 $11'$ を取得するために、フォアグラウンド H O A 係数 65 を調整された環境 H O A 係数 $47'$ に加算することができる (146)。

【0305】

[0184] 図 6 B は、本開示で説明されるコーディング技法を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャートである。図 4 の例において示されたオーディオ符号化デバイス 24 の抽出ユニット 72 は、本開示で説明される技法を実行するように構成された 1 つの例示的なユニットを表し得る。ビットストリーム抽出ユニット 72 は、フレームの量子化モードが、(「第 2 のフレーム」として示され得る) 時間的に以前のフレームの量子化モードと同じであるか否かを示すビットを取得することができる (362)。この場合も、以前のフレームに関して説明されているが、本技法は、時間的に後続のフレームに関して実行され得る。

【0306】

[0185] 量子化モードが同じであるとき (「YES」364)、抽出ユニット 72 は、ビットストリーム 21 から量子化モードの一部分を取得することができる (366)。量子化モードの一部分は、b A シンタックス要素と b B シンタックス要素とを含み得るが、u i n t C シンタックス要素を含まないことがある。抽出ユニット 42 はまた、現在のフレームのための N b i t s Q 値、P F l a g 値、C b F l a g 値、C o d e b k I d x 値、および NumVecIndices 値の値を、以前のフレームのために設定された N b i t s Q 値、P F l a g 値、C b F l a g 値、C o d e b k I d x 値、および NumVecIndices の値と同じように設定することができる (368)。

【0307】

[0186] 量子化モードが同じではないとき (「NO」364)、抽出ユニット 72 は、ビットストリーム 21 から全体量子化モードを示す 1 つまたは複数のビットを取得することができる。すなわち、抽出ユニット 72 は、ビットストリーム 21 から b A シンタックス

要素と、b Bシンタックス要素と、u i n t Cシンタックス要素とを取得する（370）。抽出ユニット72はまた、量子化モードに基づいて、量子化情報を示す1つまたは複数のビットを取得することができる（372）。図5Bに関して上述されたように、量子化情報は、ベクトル量子化情報、予測情報、およびハフマンコードブック情報など、量子化に関する任意の情報を含み得る。ベクトル量子化情報は、一例として、C o d e b k I d xシンタックス要素およびN u m V e c I n d i c e sシンタックス要素のうち的一方または両方を含み得る。予測情報は、一例として、P F l a gシンタックス要素を含み得る。ハフマンコードブック情報は、一例として、C b F l a gシンタックス要素を含み得る。

【0308】

[0187]この点に関して、技法は、オーディオ復号デバイス24がサウンドフィールドの空間成分の圧縮されたバージョンを備えるビットストリーム21を取得するように構成されることができ得る。空間成分は、複数の球面調和関数係数に関してベクトルベース合成を実行することによって生成され得る。ビットストリームは、空間成分を圧縮するときに使用される情報を指定する、以前のフレームからの、ヘッダフィールドの1つまたは複数のビットを再使用するかどうかのためのインジケータをさらに備える。

【0309】

[0188]言い換えれば、技法は、オーディオ復号デバイス24が球面調和関数領域における直交空間軸を表すベクトル57を備えるビットストリーム21を取得するように構成されることができ得る。ビットストリーム21は、ベクトルを圧縮（たとえば、量子化）するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータ（たとえば、N b i t s Qシンタックス要素のb A / b Bシンタックス要素）をさらに備え得る。

【0310】

[0189]図7は、本開示で説明される技法の様々な態様に従って指定された例示的なフレーム249Sおよび249Tを示す図である。図7の例に示されるように、フレーム249Sは、C h a n n e l S i d e I n f o D a t a（C S I D）フィールド154A~154Dと、H O A G a i n C o r r e c t i o n D a t a（H O A G C D）フィールドと、V V e c t o r D a t aフィールド156Aおよび156Bと、H O A P r e d i c t i o n I n f oフィールドとを含む。C S I Dフィールド154Aは、01の値に設定されたC h a n n e l T y p eシンタックス要素（「C h a n n e l T y p e」）269とともに、10の値に設定されたu i n t Cシンタックス要素（「u i n t C」）267と、1の値に設定されたb bシンタックス要素（「b B」）266と、0の値に設定されたb Aシンタックス要素（「b A」）265とを含む。

【0311】

[0190]u i n t Cシンタックス要素267、b Bシンタックス要素266、およびb Aシンタックス要素265は一緒に、N b i t s Qシンタックス要素261を形成し、b Aシンタックス要素265がN b i t s Qシンタックス要素261の最上位ビットを形成し、b Bシンタックス要素266が第2の最上位ビットを形成し、u i n t Cシンタックス要素267が最下位ビットを形成する。N b i t s Qシンタックス要素261は、上述されたように、高次アンビソニックオーディオデータを符号化するために使用された量子化モード（たとえば、ベクトル量子化モード、ハフマンコーディングなしのスカラー量子化モード、およびハフマンコーディングありのスカラー量子化モード）を示す1つまたは複数のビットを表し得る。

【0312】

[0191]C S I Dシンタックス要素154Aはまた、様々なシンタックステーブルにおいて上記で言及されたP F l a gシンタックス要素300とC b F l a gシンタックス要素302とを含む。P F l a gシンタックス要素300は、第1のフレーム249SのH O A係数によって表されるサウンドフィールドの空間成分のコード化要素（ここで、さらに空間成分は、Vベクトルを指し得る）が第2のフレーム（たとえば、この例では以前のフ

10

20

30

40

50

レーム) から予測されるか否かを示す、1つまたは複数のビットを表し得る。C b F l a g シンタックス要素 3 0 2 は、空間成分 (または言い換えれば、V ベクトル要素) を符号化するために使用されたハフマンコードブック (または、言い換えれば、テーブル) のいずれかを特定することができる、ハフマンコードブック情報を示す、1つまたは複数のビットを表し得る。

【0313】

[0192] C S I D フィールド 1 5 4 B は、b B シンタックス要素 2 6 6 と b B シンタックス要素 2 6 5 とを、C h a n n e l T y p e シンタックス要素 2 6 9 とともに含み、その各々が、図 7 の例において対応する値 0 および 0 および 0 1 に設定される。C S I D フィールド 1 5 4 C および 1 5 4 D の各々は、3 (1 1₂) の値を有する C h a n n e l T y p e フィールド 2 6 9 を含む。C S I D フィールド 1 5 4 A ~ 1 5 4 D の各々は、トランスポートチャネル 1、2、3 および 4 の各々に対応する。事実上、各 C S I D フィールド 1 5 4 A ~ 1 5 4 D は、対応するペイロードが指向性ベースの信号か (対応する C h a n n e l T y p e が 0 に等しいとき)、ベクトルベースの信号か (対応する C h a n n e l T y p e が 1 に等しいとき)、追加の環境 H O A 係数か (対応する C h a n n e l T y p e が 2 に等しいとき)、空か (C h a n n e l T y p e が 3 に等しいとき) を示す。

【0314】

[0193] 図 7 の例では、フレーム 2 4 9 S は、(C S I D フィールド 1 5 4 A および 1 5 4 B において 1 に等しい C h a n n e l T y p e シンタックス要素 2 6 9 が与えられる) 2 つのベクトルベース信号と、(C S I D フィールド 1 5 4 C および 1 5 4 D において 3 に等しい C h a n n e l T y p e 2 6 9 が与えられる) 2 つの空とを含む。その上、オーディオ符号化デバイス 2 0 は、P F l a g シンタックス要素 3 0 0 が 1 に設定されることによって示されるような予測を採用した。この場合も、P F l a g シンタックス要素 3 0 0 によって示されるような予測は、圧縮された空間成分 v₁ ~ v_n のうちの対応する 1 つに関して予測が実行されたか否かを示す予測モード指示を指す。P F l a g シンタックス要素 3 0 0 が 1 に設定されるとき、オーディオ符号化デバイス 2 0 は、スカラー量子化では、現在のフレームの対応するベクトル要素との以前のフレームからのベクトル要素の間の差分、または、ベクトル量子化では、現在のフレームの対応する重みとの以前のフレームからの重みの間の差分を取ることに由る予測を採用することができる。

【0315】

[0194] オーディオ符号化デバイス 2 0 はまた、フレーム 2 4 9 S における第 2 のトランスポートチャネルの C S I D フィールド 1 5 4 B のための N b i t s Q シンタックス要素 2 6 1 のための値が、以前のフレーム、たとえば図 7 の例におけるフレーム 2 4 9 T、の第 2 のトランスポートチャネルの C S I D フィールド 1 5 4 B のための N b i t s Q シンタックス要素 2 6 1 の値と同じであると決定した。結果として、オーディオ符号化デバイス 2 0 は、以前のフレーム 2 4 9 T における第 2 のトランスポートチャネルの N b i t s Q シンタックス要素 2 6 1 の値が、フレーム 2 4 9 S における第 2 のトランスポートチャネルの N b i t s Q シンタックス要素 2 6 1 のために再使用されることをシグナリングするために、b A シンタックス要素 2 6 5 および b B シンタックス要素 2 6 6 の各々に対して 0 の値を指定した。結果として、オーディオ符号化デバイス 2 0 は、上で識別された他のシンタックス要素と共にフレーム 2 4 9 S における第 2 のトランスポートチャネルのために u i n t C シンタックス要素 2 6 7 を指定することを回避することができる。

【0316】

[0195] 図 8 は、本明細書で説明される技法による、少なくとも 1 つのビットストリームの 1 つまたは複数のチャネルのための例示的なフレームを示す図である。ビットストリーム 4 5 0 は、1 つまたは複数のチャネルをそれぞれ含み得るフレーム 8 1 0 A ~ 8 1 0 H を含む。ビットストリーム 4 5 0 は、図 7 の例において示されたビットストリーム 2 1 の 1 つの例であり得る。図 8 の例では、オーディオ復号デバイス 2 4 は、状態情報を維持し、どのように現在のフレーム k を復号するかを決定するために、状態情報を更新する。オーディオ復号デバイス 2 4 は、c o n f i g 8 1 4 からの状態情報と、フレーム 8 1 0 B

10

20

30

40

50

～ 8 1 0 Dとを利用することができる。

【0317】

[0196] 言い換えれば、オーディオ符号化デバイス 20 は、ビットストリーム生成ユニット 42 が状態機械 402 に基づいてフレーム 810A～810E の各々のためのシンタックス要素を指定することができる点において、たとえば、ビットストリーム生成ユニット 42 内で、フレーム 810A～810E の各々を符号化するための状態情報を維持する状態機械 402 を含み得る。

【0318】

[0197] オーディオ復号デバイス 24 は、たとえば、ビットストリーム抽出ユニット 72 内で、状態機械 402 に基づいてシンタックス要素（その一部がビットストリーム 21 において明示的に指定されない）を出力する同様の状態機械 402 を同様に含み得る。オーディオ復号デバイス 24 の状態機械 402 は、オーディオ符号化デバイス 20 の状態機械 402 の動作と同様の方法で動作することができる。したがって、オーディオ復号デバイス 24 の状態機械 402 は、状態情報を維持し、`confi g 814`と、図 8 の例では、フレーム 810B～810D の復号とに基づいて、状態情報を更新することができる。状態情報に基づいて、ビットストリーム抽出ユニット 72 は、状態機械 402 によって維持された状態情報に基づいて、フレーム 810E を抽出することができる。状態情報は、オーディオ符号化デバイス 20 がフレーム 810E の様々なトランスポートチャネルを復号するときに利用することができる、いくつかの暗黙的なシンタックス要素を与えることができる。

【0319】

[0198] 上記の技法は、任意の数の異なる状況およびオーディオエコシステムに関して実行され得る。いくつかの例示的な状況が以下で説明されるが、本技法はそれらの例示的な状況に限定されるべきではない。1つの例示的なオーディオエコシステムは、オーディオコンテンツと、映画スタジオと、音楽スタジオと、ゲーミングオーディオスタジオと、チャンネルベースオーディオコンテンツと、コーディングエンジンと、ゲームオーディオシステムと、ゲームオーディオコーディング／レンダリングエンジンと、配信システムとを含み得る。

【0320】

[0199] 映画スタジオ、音楽スタジオ、およびゲーミングオーディオスタジオは、オーディオコンテンツを受信することができる。いくつかの例では、オーディオコンテンツは、獲得物の出力を表し得る。映画スタジオは、デジタルオーディオワークステーション（DAW）を使用することなどによって、（たとえば、2. 0、5. 1、および 7. 1 の）チャンネルベースオーディオコンテンツを出力することができる。音楽スタジオは、DAWを使用することなどによって、（たとえば、2. 0、および 5. 1 の）チャンネルベースオーディオコンテンツを出力することができる。いずれの場合も、コーディングエンジンは、配信システムによる出力のために、チャンネルベースオーディオコンテンツベースの 1 つまたは複数のコーデック（たとえば、AAC、AC3、Dolby True HD、Dolby Digital Plus、および DTS Master Audio）を受信し符号化することができる。ゲーミングオーディオスタジオは、DAWを使用することなどによって、1 つまたは複数のゲームオーディオシステムを出力することができる。ゲームオーディオコーディング／レンダリングエンジンは、配信システムによる出力のために、オーディオシステムをチャンネルベースオーディオコンテンツへとコーディングおよびまたはレンダリングすることができる。本技法が実行され得る別の例示的な状況は、放送録音オーディオオブジェクトと、プロフェッショナルオーディオシステムと、消費者向けオンデバイスキャプチャと、HOA オーディオフォーマットと、オンデバイスレンダリングと、消費者向けオーディオと、TV、およびアクセサリと、カーオーディオシステムとを含み得る、オーディオエコシステムを備える。

【0321】

[0200] 放送録音オーディオオブジェクト、プロフェッショナルオーディオシステム、お

よび消費者向けオンデバイスキャプチャはすべて、H O A オーディオフォーマットを使用して、それらの出力をコーディングすることができる。このようにして、オーディオコンテンツは、オンデバイスレンダリング、消費者向けオーディオ、T V、およびアクセサリ、ならびにカーオーディオシステムを使用して再生され得る単一の表現へと、H O A オーディオフォーマットを使用してコーディングされ得る。言い換えれば、オーディオコンテンツの単一の表現は、オーディオ再生システム 1 6 など、汎用的なオーディオ再生システムにおいて（すなわち、5. 1、7. 1 などの特定の構成を必要とすることとは対照的に）再生され得る。

【0322】

[0201] 本技法が実行され得る状況の他の例には、獲得要素と再生要素とを含み得るオーディオエコシステムがある。獲得要素は、有線および／またはワイヤレス獲得デバイス（たとえば、Eigen マイクロフォン）、オンデバイスサラウンドサウンドキャプチャ、ならびにモバイルデバイス（たとえば、スマートフォンおよびタブレット）を含み得る。いくつかの例では、有線および／またはワイヤレス獲得デバイスは、有線および／またはワイヤレス通信チャネルを介してモバイルデバイスに結合され得る。

【0323】

[0202] 本開示の 1 つまたは複数の技法によれば、モバイルデバイスが音場を獲得するために使用され得る。たとえば、モバイルデバイスは、有線および／もしくはワイヤレス獲得デバイス、ならびに／またはオンデバイスサラウンドサウンドキャプチャ（たとえば、モバイルデバイスに統合された複数のマイクロフォン）を介して、音場を獲得することができる。モバイルデバイスは次いで、再生要素のうちの 1 つまたは複数による再生のために、獲得された音場を H O A 係数へとコーディングすることができる。たとえば、モバイルデバイスのユーザは、ライブイベント（たとえば、会合、会議、劇、コンサートなど）を録音し（その音場を獲得し）、録音を H O A 係数へとコーディングすることができる。

【0324】

[0203] モバイルデバイスはまた、H O A コーディングされた音場を再生するために、再生要素のうちの 1 つまたは複数を利用することができる。たとえば、モバイルデバイスは、H O A コーディングされた音場を復号し、再生要素のうちの 1 つまたは複数に信号を出力することができ、このことは再生要素のうちの 1 つまたは複数に音場を再作成させる。一例として、モバイルデバイスは、1 つまたは複数のスピーカー（たとえば、スピーカーアレイ、サウンドバーなど）に信号を出力するために、ワイヤレスおよび／またはワイヤレス通信チャネルを利用することができる。別の例として、モバイルデバイスは、1 つもしくは複数のドッキングステーションおよび／または 1 つもしくは複数のドッキングされたスピーカー（たとえば、スマート自動車および／またはスマート住宅の中のサウンドシステム）に信号を出力するために、ドッキング解決手段を利用することができる。別の例として、モバイルデバイスは、ヘッドフォンのセットに信号を出力するために、たとえばリアルなバイノーラルサウンドを作成するために、ヘッドフォンレンダリングを利用することができる。

【0325】

[0204] いくつかの例では、特定のモバイルデバイスは、3 D 音場を獲得することと、より後の時間に同じ 3 D 音場を再生することの両方を行うことができる。いくつかの例では、モバイルデバイスは、3 D 音場を獲得し、3 D 音場を H O A へと符号化し、符号化された 3 D 音場を再生のために 1 つまたは複数の他のデバイス（たとえば、他のモバイルデバイスおよび／または他の非モバイルデバイス）に送信することができる。

【0326】

[0205] 本技法が実行され得る Y また別の状況は、オーディオコンテンツと、ゲームスタジオと、コーディングされたオーディオコンテンツと、レンダリングエンジンと、配信システムとを含み得る、オーディオエコシステムを含む。いくつかの例では、ゲームスタジオは、H O A 信号の編集をサポートし得る 1 つまたは複数の D A W を含み得る。たとえば、1 つまたは複数の D A W は、1 つまたは複数のゲームオーディオシステムとともに動作

する（たとえば、機能する）ように構成され得る、H O Aプラグインおよび／またはツールを含み得る。いくつかの例では、ゲームスタジオは、H O Aをサポートする新しいステムフォーマットを出力することができる。いずれの場合も、ゲームスタジオは、配信システムによる再生のために音場をレンダリングすることができるレンダリングエンジンに、コーディングされたオーディオコンテンツを出力することができる。

【0327】

[0206]本技法はまた、例示的なオーディオ獲得デバイスに関して実行され得る。たとえば、本技法は、3 D音場を録音するようにまとめて構成される複数のマイクロフォンを含み得る、E i g e nマイクロフォンに関して実行され得る。いくつかの例では、E i g e nマイクロフォンの複数のマイクロフォンは、約4 c mの半径を伴う実質的に球状の球体の表面に配置され得る。いくつかの例では、オーディオ符号化デバイス20は、マイクロフォンから直接ビットストリーム21を出力するために、E i g e nマイクロフォンに統合され得る。

10

【0328】

[0207]別の例示的なオーディオ獲得状況は、1つまたは複数のE i g e nマイクロフォンなど、1つまたは複数のマイクロフォンから信号を受信するように構成され得る、製作トラックを含み得る。製作トラックはまた、図3のオーディオ符号化器20などのオーディオ符号化器を含み得る。

【0329】

[0208]モバイルデバイスはまた、いくつかの場合には、3 D音場を録音するようにまとめて構成される複数のマイクロフォンを含み得る。言い換えれば、複数のマイクロフォンは、X、Y、Zのダイバーシティを有し得る。いくつかの例では、モバイルデバイスは、モバイルデバイスの1つまたは複数の他のマイクロフォンに関してX、Y、Zのダイバーシティを提供するように回転され得るマイクロフォンを含み得る。モバイルデバイスはまた、図3のオーディオ符号化器20などのオーディオ符号化器を含み得る。

20

【0330】

[0209]耐衝撃性のビデオキャプチャデバイスは、3 D音場を録音するようにさらに構成され得る。いくつかの例では、耐衝撃性のビデオキャプチャデバイスは、ある活動に関与するユーザのヘルメットに取り付けられ得る。たとえば、耐衝撃性のビデオキャプチャデバイスは、急流下りをしているユーザのヘルメットに取り付けられ得る。このようにして、耐衝撃性のビデオキャプチャデバイスは、ユーザの周りのすべての活動（たとえば、ユーザの後ろでくだける水、ユーザの前で話している別の乗員など）を表す3 D音場をキャプチャすることができる。

30

【0331】

[0210]本技法はまた、3 D音場を録音するように構成され得る、アクセサリで増強されたモバイルデバイスに関して実行され得る。いくつかの例では、モバイルデバイスは、上記で説明されたモバイルデバイスと同様であり得るが、1つまたは複数のアクセサリが追加されている。たとえば、E i g e nマイクロフォンが、アクセサリで増強されたモバイルデバイスを形成するために、上述されたモバイルデバイスに取り付けられ得る。このようにして、アクセサリで増強されたモバイルデバイスは、アクセサリで増強されたモバイルデバイスと一体のサウンドキャプチャ構成要素をただ使用するよりも高品質なバージョンの3 D音場をキャプチャすることができる。

40

【0332】

[0211]本開示で説明される本技法の様々な態様を実行することができる例示的なオーディオ再生デバイスが、以下でさらに説明される。本開示の1つまたは複数の技法によれば、スピーカーおよび／またはサウンドバーは、あらゆる任意の構成で配置され得るが、一方で、依然として3 D音場を再生する。その上、いくつかの例では、ヘッドフォン再生デバイスが、有線接続またはワイヤレス接続のいずれかを介して復号器24に結合され得る。本開示の1つまたは複数の技法によれば、音場の単一の汎用的な表現が、スピーカー、サウンドバー、およびヘッドフォン再生デバイスの任意の組合せで音場をレンダリングす

50

るために利用され得る。

【0333】

[0212]いくつかの異なる例示的なオーディオ再生環境はまた、本開示で説明される技法の様々な態様を実行するために好適であり得る。たとえば、5. 1スピーカー再生環境、2. 0（たとえば、ステレオ）スピーカー再生環境、フルハイトフロントラウドスピーカーを伴う9. 1スピーカー再生環境、22. 2スピーカー再生環境、16. 0スピーカー再生環境、自動車スピーカー再生環境、およびイヤバッド再生環境を伴うモバイルデバイスは、本開示で説明される技法の様々な態様を実行するために好適な環境であり得る。

【0334】

[0213]本開示の1つまたは複数の技法によれば、音場の単一の汎用的な表現が、上記の再生環境のいずれかにおいて音場をレンダリングするために利用され得る。加えて、本開示の技法は、レンダーが、上記で説明されたもの以外の再生環境での再生のために、汎用的な表現から音場をレンダリングすることを可能にする。たとえば、設計上の考慮事項が、7. 1スピーカー再生環境に従ったスピーカーの適切な配置を妨げる場合（たとえば、右側のサラウンドスピーカーを配置することが可能ではない場合）、本開示の技法は、再生が6. 1スピーカー再生環境で達成され得るように、レンダーが他の6つのスピーカーとともに補償することを可能にする。

【0335】

[0214]その上、ユーザは、ヘッドフォンを装着しながらスポーツの試合を見ることができ、本開示の1つまたは複数の技法によれば、スポーツの試合の3D音場が獲得され得（たとえば、1つまたは複数のEigenマイクロフォンが野球場の中および／または周りに配置され得）、3D音場に対応するHOA係数が取得され復号器に送信され得、復号器がHOA係数に基づいて3D音場を再構成して、再構成された3D音場をレンダラに出力することができ、レンダラが再生環境のタイプ（たとえば、ヘッドフォン）についての指示を取得し、再構成された3D音場を、ヘッドフォンにスポーツの試合の3D音場の表現を出力させる信号へとレンダリングすることができる。

【0336】

[0215]上記で説明された様々な場合の各々において、オーディオ符号化デバイス20は、ある方法を実行し、またはさもなければ、オーディオ符号化デバイス20が実行するように構成される方法の各ステップを実行するための手段を備え得ることを理解されたい。いくつかの場合には、これらの手段は1つまたは複数のプロセッサを備え得る。いくつかの場合には、1つまたは複数のプロセッサは、非一時的コンピュータ可読記憶媒体に記憶される命令によって構成される、専用のプロセッサを表し得る。言い換えれば、符号化の例のセットの各々における本技法の様々な態様は、実行されると、1つまたは複数のプロセッサに、オーディオ符号化デバイス20が実行するように構成されている方法を実行さ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を提供し得る。

【0337】

[0216]1つまたは複数の例において、前述の機能は、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、コンピュータ可読媒体上の1つまたは複数の命令またはコード上に記憶され、またはこれを介して送信され、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、データ記憶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。データ記憶媒体は、本開示で説明される技法の実装のために命令、コードおよび／またはデータ構造を取り出すために、1つ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。

【0338】

[0217]同様に、上記で説明された様々な場合の各々において、オーディオ復号デバイス24は、ある方法を実行し、またはさもなければ、オーディオ復号デバイス24が実行するように構成される方法の各ステップを実行するための手段を備え得ることを理解された

い。いくつかの場合には、これらの手段は1つまたは複数のプロセッサを備え得る。いくつかの場合には、1つまたは複数のプロセッサは、非一時的コンピュータ可読記憶媒体に記憶される命令によって構成される、専用のプロセッサを表し得る。言い換えれば、符号化の例のセットの各々における本技法の様々な態様は、実行されると、1つまたは複数のプロセッサに、オーディオ復号デバイス24が実行するように構成されている方法を実行さ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を提供し得る。

【0339】

[0218]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、ROM、EEPROM（登録商標）、CD-ROMもしくは他の光ディスクストレージ、磁気ディスクストレージ、もしくは他の磁気記憶デバイス、フラッシュメモリ、または命令もしく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得、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。しかしながら、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時的媒体を含むのではなく、非一時的な有形の記憶媒体を対象とすることを理解されたい。本明細書で使用するディスク(disk)およびディスク(disc)は、コンパクトディスク(disc)(CD)、レーザーディスク(登録商標)(disc)、光ディスク(disc)、デジタル多用途ディスク(disc)(DVD)、フロッピー(登録商標)ディスク(disk)およびBlu-ray(登録商標)ディスク(disc)を含み、ここで、ディスク(disk)は、通常、データを磁氣的に再生し、一方、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の中に含まれるべきである。

【0340】

[0219]命令は、1つもしく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ゲートアレイ(FPGA)、あるいは他の同等の集積回路またはディスクリット論理回路などの1つもしく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」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好適な任意の他の構造のいずれかを指し得る。加えて、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るか、または複合コーデックに組み込まれる、専用のハードウェアモジュールおよび／またはソフトウェアモジュール内で提供され得る。また、本技法は、1つもしくは複数の回路または論理要素で十分に実装され得る。

【0341】

[0220]本開示の技法は、ワイヤレスハンドセット、集積回路(IC)もしくはICのセット(たとえば、チップセット)を含む、多種多様なデバイスまたは装置で実装され得る。本開示では、開示される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、またはユニットが説明されるが、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、上記の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与えられ得る。

【0342】

[0221]本開示の様々な態様が説明された。本技法のこれらおよび他の態様は、以下の特許請求の範囲内に入る。

【図 1】

図 1

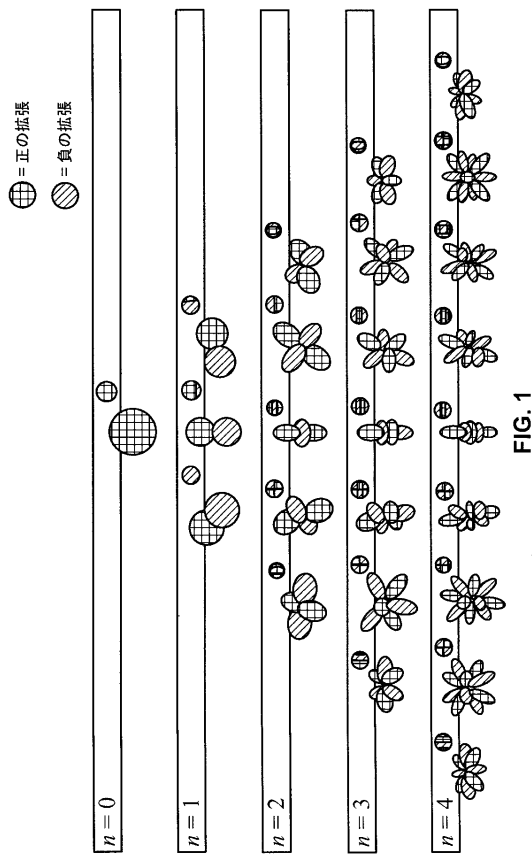


FIG. 1

【図 2】

図 2

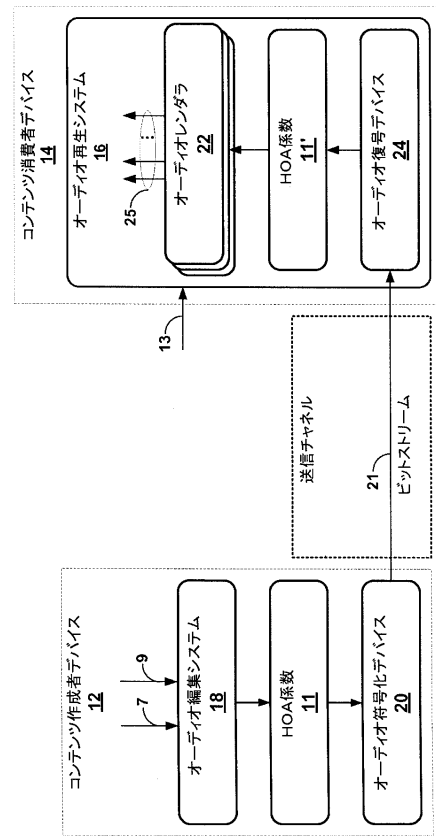


FIG. 2

【図 3】

図 3

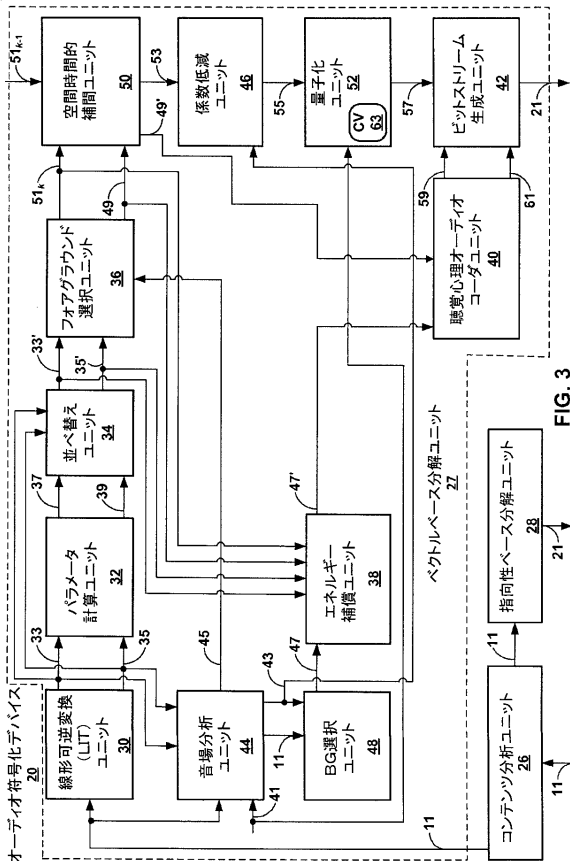


FIG. 3

【図 4】

図 4

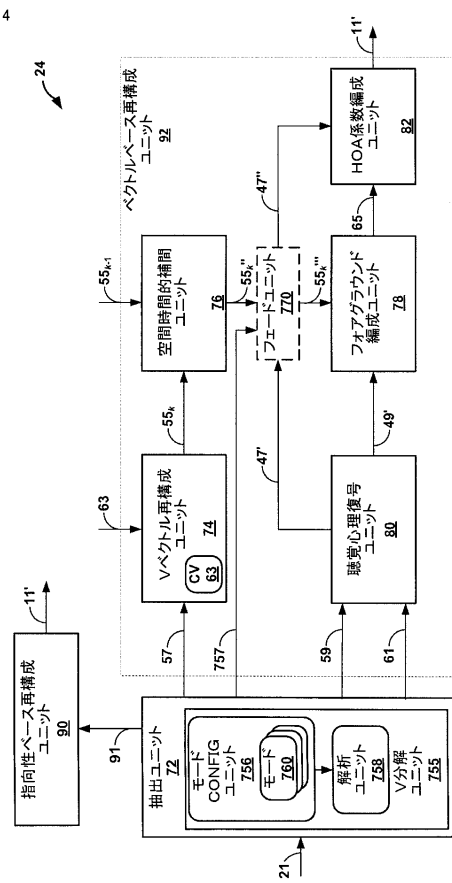
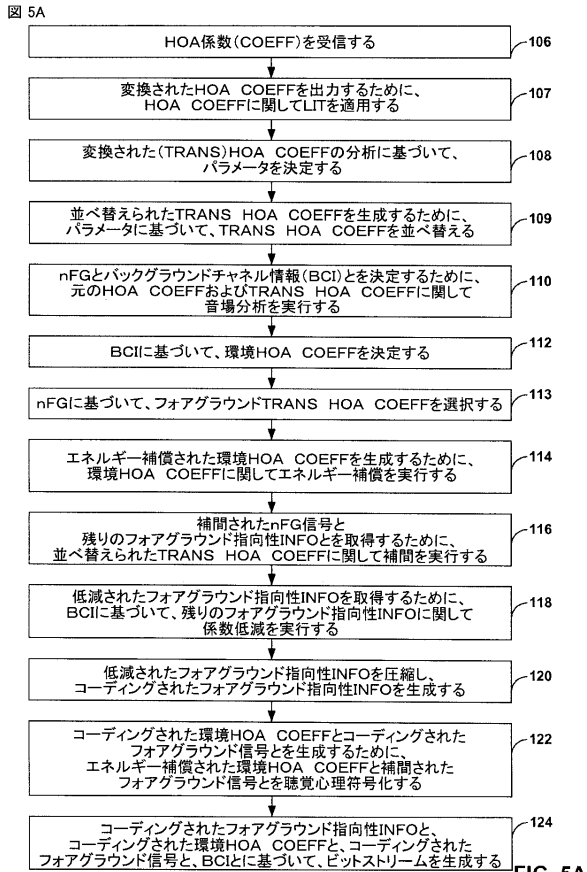
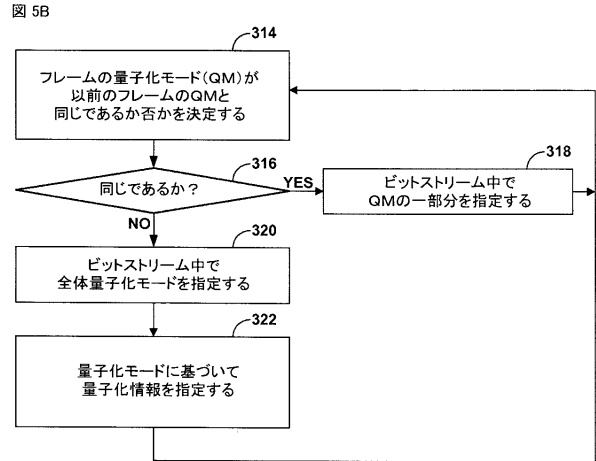


FIG. 4

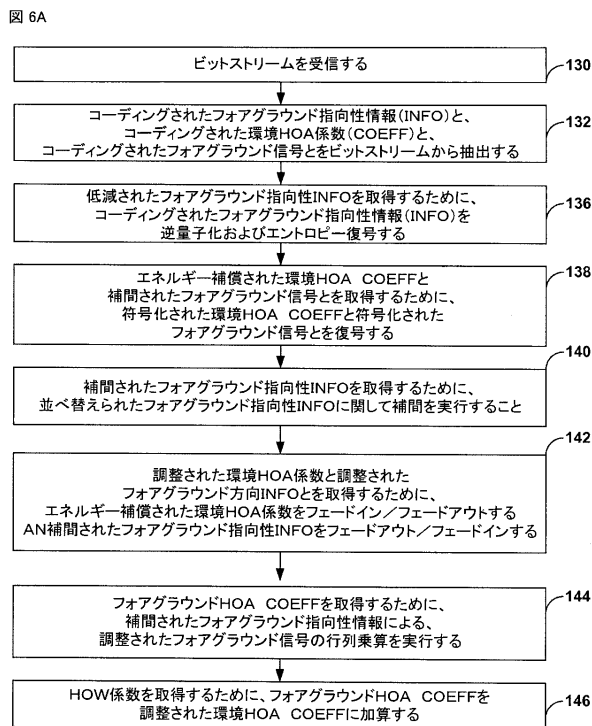
【図 5 A】



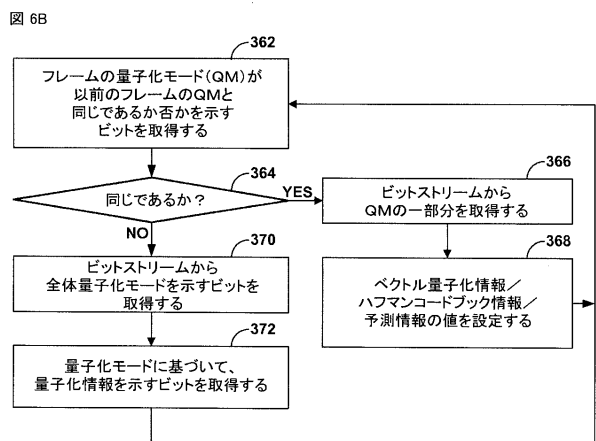
【図 5 B】



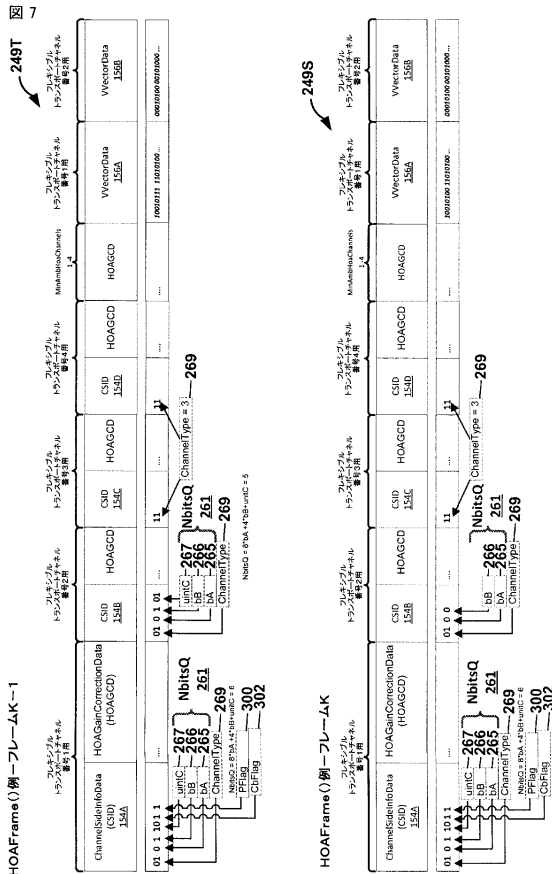
【図 6 A】



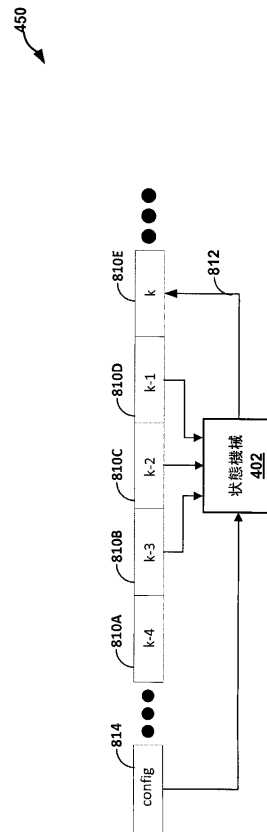
【図 6 B】



【图 7】



【图 8】



【手續補正書】

【提出日】平成29年7月27日(2017.7.27)

【手続補正 1】

【補正対象書類名】 特許請求の範囲

【補正対象項目名】全文

【補正方法】 変更

【補正の内容】

【特許請求の範囲】

【請求項1】

ビットストリームを処理するためのデバイスであって、前記デバイスは、

前記ビットストリームを取得するように構成された1つまたは複数のプロセッサ、前記ビットストリームは、サウンドフィールドの空間成分の圧縮されたバージョンを備え、前記サウンドフィールドの前記空間成分は、球面調和関数領域におけるベクトルによって表され、現在のフレームのためのシンタックス要素の値は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示し、前記ビットストリームは、インジケータをさらに備える、と、

前記インジケータは、前記ビットストリームが前記現在のフレームのための前記シンタックス要素の前記値を含まないこと、および前記現在のフレームのための前記シンタックス要素の前記値が以前のフレームのための前記シンタックス要素の値に等しいことを示す特定の値を有し、

前記 1 つまたは複数のプロセッサに結合されたメモリ、前記メモリは、前記ビットストリームを記憶するように構成される、と
を備える、デバイス。

【請求項2】

前記 1 つまたは複数のプロセッサは、前記ベクトルを再構成するために、前記ベクトル

量子化コードブックを使用するようにさらに構成される、
請求項 1 に記載のデバイス。

【請求項 3】

前記シンタックス要素は、第 1 のシンタックス要素であり、前記インジケータは、前記
現在のフレームのための第 2 のシンタックス要素の値の 1 つまたは複数のビットを備え、
前記現在のフレームのための前記第 2 のシンタックス要素の前記値は、前記ベクトルを圧
縮するときに使用される量子化モードを示す、

請求項 1 に記載のデバイス。

【請求項 4】

前記インジケータは、前記現在のフレームのための第 3 のシンタックス要素の値と、前
記現在のフレームのための第 4 のシンタックス要素の値とを備え、

前記現在のフレームのための前記第 3 のシンタックス要素の前記値プラス前記現在のフ
レームのための前記第 4 のシンタックス要素の前記値がゼロに等しいことは、前記ビット
ストリームが前記現在のフレームのための前記第 1 のシンタックス要素の前記値を含まな
いこと、および前記現在のフレームのための前記第 1 のシンタックス要素の前記値が前記
以前のフレームのための前記第 1 のシンタックス要素の前記値に等しいことを示す、

請求項 3 に記載のデバイス。

【請求項 5】

前記インジケータは、前記現在のフレームのための前記第 2 のシンタックス要素の前記
値の最上位ビットと、前記現在のフレームのための前記第 2 のシンタックス要素の前記値
の第 2 の最上位ビットとを含む、

請求項 3 に記載のデバイス。

【請求項 6】

前記 1 つまたは複数のプロセッサは、

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと

、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを
指定することと

を行うようにさらに構成される、請求項 1 に記載のデバイス。

【請求項 7】

前記 1 つまたは複数のプロセッサは、

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから取得す
ることと、

高次アンビソニック（H O A）オーディオデータを再構成するために、前記ベクトルと
前記オーディオオブジェクトを組み合わせることと

を行うようにさらに構成される、請求項 1 に記載のデバイス。

【請求項 8】

前記 1 つまたは複数のプロセッサは、1 つまたは複数のラウドスピーカーフィールドを出
力するために前記 H O A オーディオデータをレンダリングするように構成され、

前記デバイスは、1 つまたは複数のラウドスピーカーに結合され、前記 1 つまたは複数
のラウドスピーカーフィールドは、前記 1 つまたは複数のラウドスピーカーを駆動する、

請求項 1 に記載のデバイス。

【請求項 9】

前記 1 つまたは複数のプロセッサは、

前記インジケータが前記特定の値を有していないとき、前記ビットストリームから前記
現在のフレームのための前記シンタックス要素の前記値を取得することを行うようにさら
に構成される、

請求項 1 に記載のデバイス。

【請求項 10】

ビットストリームを処理するための方法であって、前記方法は、

前記ビットストリームを取得すること、前記ビットストリームは、サウンドフィールドの空間成分の圧縮されたバージョンを備え、前記サウンドフィールドの前記空間成分は、球面調和関数領域におけるベクトルによって表され、現在のフレームのためのシンタックス要素の値は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示し、前記ビットストリームは、インジケータをさらに備え、

前記インジケータは、前記ビットストリームが前記現在のフレームのための前記シンタックス要素の前記値を含まないこと、および前記現在のフレームのための前記シンタックス要素の前記値が以前のフレームのための前記シンタックス要素の値に等しいことを示す特定の値を有する、と、

前記ビットストリームを記憶することと
を備える、方法。

【請求項 1 1】

前記ベクトルを再構成するために、前記ベクトル量子化コードブックを使用することをさらに備える、

請求項 1 0 に記載の方法。

【請求項 1 2】

前記シンタックス要素は、第 1 のシンタックス要素であり、前記インジケータは、前記現在のフレームのための第 2 のシンタックス要素の値の 1 つまたは複数のビットを備え、前記現在のフレームのための前記第 2 のシンタックス要素の前記値は、前記ベクトルを圧縮するときに使用される量子化モードを示す、

請求項 1 0 に記載の方法。

【請求項 1 3】

前記インジケータは、前記現在のフレームのための第 3 のシンタックス要素の値と、前記現在のフレームのための第 4 のシンタックス要素の値とを備え、

前記現在のフレームのための前記第 3 のシンタックス要素の前記値プラス前記現在のフレームのための前記第 4 のシンタックス要素の前記値がゼロに等しいことは、前記ビットストリームが前記現在のフレームのための前記第 1 のシンタックス要素の前記値を含まないこと、および前記現在のフレームのための前記第 1 のシンタックス要素の前記値が前記以前のフレームのための前記第 1 のシンタックス要素の前記値に等しいことを示す、

請求項 1 2 に記載の方法。

【請求項 1 4】

前記インジケータは、前記現在のフレームのための前記第 2 のシンタックス要素の前記値の最上位ビットと、前記現在のフレームのための前記第 2 のシンタックス要素の前記値の第 2 の最上位ビットとを含む、

請求項 1 2 に記載の方法。

【請求項 1 5】

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することと

をさらに備える、請求項 1 0 に記載の方法。

【請求項 1 6】

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから取得することと、

高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせることと

をさらに備える、請求項 1 0 に記載の方法。

【請求項 1 7】

高次アンビソニックス (HOA) 係数を取得するために前記ビットストリームを復号することと、

1つまたは複数のラウドスピーカーフィードを出力するために前記H O A係数をレンダリングすることと

をさらに備え、

前記1つまたは複数のラウドスピーカーフィードを出力するために前記H O A係数をレンダリングするデバイスは、1つまたは複数のラウドスピーカーに結合され、前記1つまたは複数のラウドスピーカーフィードは、前記1つまたは複数のラウドスピーカーを駆動する、

請求項10に記載の方法。

【請求項18】

前記インジケータが前記特定の値を有していないとき、前記ビットストリームから前記現在のフレームのための前記シンタックス要素の前記値を取得することをさらに備える、
請求項10に記載の方法。

【請求項19】

ビットストリームを処理するためのデバイスであって、前記デバイスは、

前記ビットストリームを取得するための手段、前記ビットストリームは、サウンドフィールドの空間成分の圧縮されたバージョンを備え、前記サウンドフィールドの前記空間成分は、球面調和関数領域におけるベクトルによって表され、現在のフレームのためのシンタックス要素の値は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示し、前記ビットストリームは、インジケータをさらに備え、

前記インジケータは、前記ビットストリームが前記現在のフレームのための前記シンタックス要素の前記値を含まないこと、および前記現在のフレームのための前記シンタックス要素の前記値が以前のフレームのための前記シンタックス要素の値に等しいことを示す特定の値を有する、と、

前記ビットストリームを記憶するための手段と

を備える、デバイス。

【請求項20】

前記ベクトルを再構成するために、前記ベクトル量子化コードブックを使用するための手段をさらに備える、

請求項19に記載のデバイス。

【請求項21】

前記シンタックス要素は、第1のシンタックス要素であり、前記インジケータは、前記現在のフレームのための第2のシンタックス要素の値の1つまたは複数のビットを備え、前記現在のフレームのための前記第2のシンタックス要素の前記値は、前記ベクトルを圧縮するときに使用される量子化モードを示す、

請求項19に記載のデバイス。

【請求項22】

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解するための手段と、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定するための手段と

をさらに備える、請求項19に記載のデバイス。

【請求項23】

前記デバイスは、

前記インジケータが前記特定の値を有していないとき、前記ビットストリームから前記現在のフレームのための前記シンタックス要素の前記値を取得するための手段をさらに備える、

請求項19に記載のデバイス。

【請求項24】

命令を記憶する非一時的コンピュータ可読記憶媒体であって、前記命令は、実行されると、

ビットストリームを取得すること、前記ビットストリームは、サウンドフィールドの空間成分の圧縮されたバージョンを備え、前記サウンドフィールドの前記空間成分は、球面調和関数領域におけるベクトルによって表され、現在のフレームのためのシンタックス要素の値は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示し、前記ビットストリームは、インジケータをさらに備え、

前記インジケータは、前記ビットストリームが前記現在のフレームのための前記シンタックス要素の前記値を含まないこと、および前記現在のフレームのための前記シンタックス要素の前記値が以前のフレームのための前記シンタックス要素の値に等しいことを示す特定の値を有する、と、

前記ビットストリームを記憶することと

を行うようにデバイスを構成する、非一時的コンピュータ可読記憶媒体。

【請求項 25】

前記命令は、実行されると、前記ベクトルを再構成するために、前記ベクトル量子化コードブックを使用するように前記デバイスを構成する、

請求項 24 に記載の非一時的コンピュータ可読記憶媒体。

【請求項 26】

前記シンタックス要素は、第 1 のシンタックス要素であり、前記インジケータは、前記現在のフレームのための第 2 のシンタックス要素の値の 1 つまたは複数のビットを備え、前記現在のフレームのための前記第 2 のシンタックス要素の前記値は、前記ベクトルを圧縮するときに使用される量子化モードを示す、

請求項 24 に記載の非一時的コンピュータ可読記憶媒体。

【請求項 27】

前記命令は、実行されると、前記デバイスに、

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと

、
前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することと

を行わせる、請求項 24 に記載の非一時的コンピュータ可読記憶媒体。

【請求項 28】

前記命令は、実行されると、前記デバイスに、

前記インジケータが前記特定の値を有していないとき、前記ビットストリームから前記現在のフレームのための前記シンタックス要素の前記値を取得すること

を行わせる、請求項 24 に記載の非一時的コンピュータ可読記憶媒体。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0342

【補正方法】変更

【補正の内容】

【0342】

[0221] 本開示の様々な態様が説明された。本技法のこれらおよび他の態様は、以下の特許請求の範囲内に入る。

以下に本願の出願当初の特許請求の範囲に記載された発明を付記する。

[C1]

効率的なビット使用の方法であって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを備え、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、

方法。

[C2]

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシンタックス要素の1つまたは複数のビットを備える、

C 1 に記載の方法。

[C 3]

前記シンタックス要素の前記1つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも1つのシンタックス要素を再使用することを示す、

C 2 に記載の方法。

[C 4]

前記量子化モードは、ベクトル量子化モードを備える、

C 2 に記載の方法。

[C 5]

前記量子化モードは、ハフマンコーディングなしのスカラ量子化モードを備える、

C 2 に記載の方法。

[C 6]

前記量子化モードは、ハフマンコーディングありのスカラ量子化モードを備える、

C 2 に記載の方法。

[C 7]

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第2の最上位ビットとを備える、

C 2 に記載の方法。

[C 8]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

C 1 に記載の方法。

[C 9]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

C 1 に記載の方法。

[C 1 0]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルが対応する圧縮カテゴリを識別するカテゴリ識別子を示すシンタックス要素を備える、

C 1 に記載の方法。

[C 1 1]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

C 1 に記載の方法。

[C 1 2]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

C 1 に記載の方法。

[C 1 3]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

C 1 に記載の方法。

[C 1 4]

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

C 1 に記載の方法。

[C 1 5]

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと

、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することと

をさらに備える、C 1 に記載の方法。

[C 1 6]

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得することと、

高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせることと

をさらに備える、C 1 に記載の方法。

[C 1 7]

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

C 1 に記載の方法。

[C 1 8]

効率的なビット使用を実行するように構成されたデバイスであって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得すること、ここにおいて、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、と、

前記ビットストリームを記憶するように構成されるメモリと

を備える、デバイス。

[C 1 9]

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシンタックス要素の 1 つまたは複数のビットを備える、

C 1 8 に記載のデバイス。

[C 2 0]

前記シンタックス要素の前記 1 つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも 1 つのシンタックス要素を再使用することを示す

、

C 1 9 に記載のデバイス。

[C 2 1]

前記量子化モードは、ベクトル量子化モードを備える、

C 1 9 に記載のデバイス。

[C 2 2]

前記量子化モードは、ハフマンコーディングなしのスカラ量子化モードを備える、

C 1 9 に記載のデバイス。

[C 2 3]

前記量子化モードは、ハフマンコーディングありのスカラ量子化モードを備える、

C 1 9 に記載のデバイス。

[C 2 4]

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第 2 の最上位ビットとを備える、

C 1 9 に記載のデバイス。

[C 2 5]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 2 6]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 2 7]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 2 8]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 2 9]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 3 0]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

C 1 8 に記載のデバイス。

[C 3 1]

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

C 1 8 に記載のデバイス。

[C 3 2]

前記 1 つまたは複数のプロセッサは、前記ベクトルを取得するために高次アンビソニックオーディオデータを分解することと、前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定することとを行うようにさらに構成される、

C 1 8 に記載のデバイス。

[C 3 3]

前記 1 つまたは複数のプロセッサは、前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得することと、高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせることとを行うようにさらに構成される、

C 1 に記載の方法。

[C 3 4]

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

C 1 8 に記載のデバイス。

[C 3 5]

効率的なビット使用のデバイスであって、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得するための手段、ここにおいて、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも 1 つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、と、

前記インジケータを記憶するための手段と

を備える、デバイス。

[C 3 6]

前記インジケータは、前記ベクトルを圧縮するときに使用される量子化モードを示すシ

ンタックス要素の1つまたは複数のビットを備える、
C 3 5に記載のデバイス。

[C 3 7]

前記シンタックス要素の前記1つまたは複数のビットは、ゼロ値に設定されるとき、前記以前のフレームからの前記少なくとも1つのシンタックス要素を再使用することを示す、

C 3 6に記載のデバイス。

[C 3 8]

前記量子化モードは、ベクトル量子化モードを備える、

C 3 6に記載のデバイス。

[C 3 9]

前記量子化モードは、ハフマンコーディングなしのスカラ量子化モードを備える、

C 3 6に記載のデバイス。

[C 4 0]

前記量子化モードは、ハフマンコーディングありのスカラ量子化モードを備える、

C 3 6に記載のデバイス。

[C 4 1]

前記シンタックス要素の一部分は、前記シンタックス要素の最上位ビットと前記シンタックス要素の第2の最上位ビットとを備える、

C 3 6に記載のデバイス。

[C 4 2]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用される以前のモードを示すシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 3]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるハフマンテーブルを示すシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 4]

前記以前のフレームからの前記シンタックス要素は、ベクトルが対応する圧縮カテゴリーを識別するカテゴリー識別子を示すシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 5]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルの要素が正の値であるか負の値であるかを示すシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 6]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるコード化ベクトルの数を示すシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 7]

前記以前のフレームからの前記シンタックス要素は、前記ベクトルを圧縮するときに使用されるベクトル量子化コードブックを示す前記以前のフレームからのシンタックス要素を備える、

C 3 5に記載のデバイス。

[C 4 8]

前記ベクトルの前記圧縮されたバージョンは、前記ベクトルの要素の残差値を表すために、少なくとも部分的に、ハフマンコードを使用して前記ビットストリームにおいて表される、

C 3 5に記載のデバイス。

[C 4 9]

前記ベクトルを取得するために高次アンビソニックオーディオデータを分解するための手段と、

前記ビットストリームを取得するために前記ビットストリームにおける前記ベクトルを指定するための手段と

をさらに備える、C 3 5 に記載のデバイス。

[C 5 0]

前記ベクトルに対応するオーディオオブジェクトを、前記ビットストリームから、取得するための手段と、

高次アンビソニックオーディオデータを再構成するために、前記ベクトルと前記オーディオオブジェクトを組み合わせるための手段と

をさらに備える、C 3 5 に記載のデバイス。

[C 5 1]

前記ベクトルの前記圧縮は、前記ベクトルの量子化を含む、

C 3 5 に記載のデバイス。

[C 5 2]

実行されると、1つまたは複数のプロセッサに、

球面調和関数領域における直交空間軸を表すベクトルを備えるビットストリームを取得することを行わせる命令を記憶し、前記ビットストリームは、前記ベクトルを圧縮するときに使用される情報を示す少なくとも1つのシンタックス要素を、以前のフレームから、再使用するかどうかのためのインジケータをさらに備える、

非一時的コンピュータ可読記憶媒体。

フロントページの続き

- (31)優先権主張番号 61/949,591
(32)優先日 平成26年3月7日(2014.3.7)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/949,583
(32)優先日 平成26年3月7日(2014.3.7)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/994,794
(32)優先日 平成26年5月16日(2014.5.16)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/004,147
(32)優先日 平成26年5月28日(2014.5.28)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/004,067
(32)優先日 平成26年5月28日(2014.5.28)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/004,128
(32)優先日 平成26年5月28日(2014.5.28)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/019,663
(32)優先日 平成26年7月1日(2014.7.1)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/027,702
(32)優先日 平成26年7月22日(2014.7.22)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/028,282
(32)優先日 平成26年7月23日(2014.7.23)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/029,173
(32)優先日 平成26年7月25日(2014.7.25)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/032,440
(32)優先日 平成26年8月1日(2014.8.1)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/056,248
(32)優先日 平成26年9月26日(2014.9.26)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/056,286
(32)優先日 平成26年9月26日(2014.9.26)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/102,243
(32)優先日 平成27年1月12日(2015.1.12)
(33)優先権主張国 米国(US)
(31)優先権主張番号 14/609,190
(32)優先日 平成27年1月29日(2015.1.29)
(33)優先権主張国 米国(US)

(72)発明者 ニルス・ガンザー・ピーターズ

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 ディパンジャン・セン

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