(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6345771号
(P6345771)
(45)発行日　平成30年6月20日(2018.6.20)　(24)登録日　平成30年6月1日(2018.6.1)

(51)Int.Cl.　F I
G10L 19/008 (2013.01)　G10L 19/008

請求項の数 15　(全 282 頁)

(21)出願番号	特願2016-516823 (P2016-516823)
(86)(22)出願日	平成26年5月29日 (2014.5.29)
(65)公表番号	特表2016-526189 (P2016-526189A)
(43)公表日	平成28年9月1日 (2016.9.1)
(86)国際出願番号	PCT/US2014/040041
(87)国際公開番号	WO2014/194105
(87)国際公開日	平成26年12月4日 (2014.12.4)
審査請求日	平成29年5月9日 (2017.5.9)
(31)優先権主張番号	61/828,445
(32)優先日	平成25年5月29日 (2013.5.29)
(33)優先権主張国	米国 (US)
(31)優先権主張番号	61/828,615
(32)優先日	平成25年5月29日 (2013.5.29)
(33)優先権主張国	米国 (US)

(73)特許権者　595020643
クゥアルコム・インコーポレイテッド
ＱＵＡＬＣＯＭＭ　ＩＮＣＯＲＰＯＲＡＴＥＤ
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
(74)代理人　100108855
弁理士　蔵田　昌俊
(74)代理人　100109830
弁理士　福原　淑弘
(74)代理人　100158805
弁理士　井関　守三
(74)代理人　100194814
弁理士　奥村　元宏

最終頁に続く

(54)【発明の名称】音場の空間成分の圧縮のための量子化ステップサイズ

(57)【特許請求の範囲】
【請求項１】
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを備え、前記空間成分は、球面調和領域において定義され、複数の球面調和係数に関して線形分解を実行することによって生成される、
方法。
【請求項２】
前記量子化ステップサイズを決定することは、目標ビットレートに基づいて前記量子化ステップサイズを決定することを備える、
請求項１に記載の方法。
【請求項３】
前記量子化ステップサイズを決定することは、
前記空間成分を表すために使用されるビットの数の推定値を決定することと、
前記推定値と目標ビットレートとの間の差に基づいて前記量子化ステップサイズを決定することと
を備える、請求項１に記載の方法。
【請求項４】
前記量子化ステップサイズを決定することは
前記空間成分を表すために使用されるビットの数の推定値を決定することと、
前記推定値と目標ビットレートとの間の差を決定することと、

前記差を前記目標ビットレートに加算することによって前記量子化ステップサイズを決定することと
を備える、請求項１に記載の方法。
【請求項５】
前記ビットの数の前記推定値を決定することは、
前記目標ビットレートに対応するコードブックを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定値を計算すること、または、
前記空間成分を圧縮するときに使用されるコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定値を計算すること
のいずれかを備える、請求項３に記載の方法。
【請求項６】
前記ビットの数の前記推定値を決定することは、
前記空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第１の推定値を計算することと、
前記空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第２の推定値を計算することと、
前記第１の推定値と前記第２の推定値のうちで、前記ビットの数の前記決定された推定値として使用されるべき、最も少ないビットを有するものを選択することと
を備える、請求項３に記載の方法。
【請求項７】
前記ビットの数の前記推定値を決定することは、
前記空間成分が対応するカテゴリを識別するカテゴリ識別子を識別することと、
前記カテゴリに対応する前記空間成分を圧縮するときに生じうる、前記空間成分の残差値のビット長を識別することと、
少なくとも一部、前記カテゴリ識別子を表すために使用されるビットの数を前記残差値の前記ビット長に加算することによって、前記ビットの数の前記推定値を決定することと
を備える、請求項３に記載の方法。
【請求項８】
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを行うように構成される１つまたは複数のプロセッサを備え、前記空間成分は、球面調和領域において定義され、複数の球面調和係数に関して線形分解を実行することによって生成される、
デバイス。
【請求項９】
前記１つまたは複数のプロセッサは、目標ビットレートに基づいて前記量子化ステップサイズを決定することを行うように構成される、
請求項８に記載のデバイス。
【請求項１０】
前記１つまたは複数のプロセッサは、前記空間成分を表すために使用されるビットの数の推定値を決定することと、前記推定値と目標ビットレートとの間の差に基づいて前記量子化ステップサイズを決定することとを行うように構成される、
請求項８に記載のデバイス。
【請求項１１】
前記１つまたは複数のプロセッサは、前記空間成分を表すために使用されるビットの数の推定値を決定することと、前記推定値と目標ビットレートとの間の差を決定することと、前記差を前記目標ビットレートに加算することによって前記量子化ステップサイズを決定することとを行うように構成される、
請求項８に記載のデバイス。
【請求項１２】
前記１つまたは複数のプロセッサは、

前記目標ビットレートに対応するコードブックを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定値を計算すること、または、
前記空間成分を圧縮するときに使用されるコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定値を計算すること
のいずれかを行うように構成される、請求項１０に記載のデバイス。
【請求項１３】
前記１つまたは複数のプロセッサは、前記空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第１の推定値を計算することと、前記空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第２の推定値を計算することと、前記第１の推定値と前記第２の推定値のうちで、前記ビットの数の前記決定された推定値として使用されるべき、最も少ないビットを有するものを選択することを行うように構成される、
請求項１０に記載のデバイス。
【請求項１４】
前記１つまたは複数のプロセッサは、前記空間成分が対応するカテゴリを識別するカテゴリ識別子を識別することと、前記カテゴリに対応する前記空間成分を圧縮するときに生じうる、前記空間成分の残差値のビット長を識別することと、少なくとも一部、前記カテゴリ識別子を表すために使用されるビットの数を前記残差値の前記ビット長に加算することによって、前記ビットの数の前記推定値を決定することとを行うように構成される、
請求項１０に記載のデバイス。
【請求項１５】
実行されると、１つまたは複数のプロセッサに、
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを行わせる命令を記憶しており、前記空間成分は、球面調和領域において定義され、複数の球面調和係数に関して線形分解を実行することによって生成される、
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、２０１３年５月２９日に出願された米国仮出願第６１／８２８，４４５号、２０１３年５月３１日に出願された米国仮出願第６１／８２９，７９１号、２０１３年１１月１日に出願された米国仮出願第６１／８９９，０３４号、２０１３年１１月１日に出願された米国仮出願第６１／８９９，０４１号、２０１３年５月３０日に出願された米国仮出願第６１／８２９，１８２号、２０１３年５月３０日に出願された米国仮出願第６１／８２９，１７４号、２０１３年５月３０日に出願された米国仮出願第６１／８２９，１５５号、２０１４年１月３０日に出願された米国仮出願第６１／９３３，７０６号、２０１３年５月３１日に出願された米国仮出願第６１／８２９，８４６号、２０１３年１０月３日に出願された米国仮出願第６１／８８６，６０５号、２０１３年１０月３日に出願された米国仮出願第６１／８８６，６１７号、２０１４年１月８日に出願された米国仮出願第６１／９２５，１５８号、２０１４年１月３０日に出願された米国仮出願第６１／９３３，７２１号、２０１４年１月８日に出願された米国仮出願第６１／９２５，０７４号、２０１４年１月８日に出願された米国仮出願第６１／９２５，１１２号、２０１４年１月８日に出願された米国仮出願第６１／９２５，１２６号、２０１４年５月２７日に出願された米国仮出願第６２／００３，５１５号、２０１３年５月２９日に出願された米国仮出願第６１／８２８，６１５号の利益を主張し、これらの各々の内容全体が参照によって本明細書に組み込まれる。
【技術分野】
【０００２】
[0002]本開示は、オーディオデータに関し、より具体的には、オーディオデータの圧縮

に関する。

【背景技術】

【０００３】

[0003]高次アンビソニックス（ＨＯＡ）信号（複数の球面調和係数（ＳＨＣ）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、音場の三次元的表現である。このＨＯＡ表現またはＳＨＣ表現は、このＳＨＣ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ーの幾何学的配置とは無関係な方式で、この音場を表すことができる。このＳＨＣ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、このＳＨＣ信号はまた、下位互換性を容易にし得る。したがって、ＳＨＣ表現は、下位互換性にも対応する、音場のより優れた表現を可能にすることができる。

【発明の概要】

【０００４】

[0004]全般に、高次アンビソニックスオーディオデータの圧縮および解凍のための技術が説明される。

【０００５】

[0005]一態様では、方法は、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを取得することを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される。

【０００６】

[0006]別の態様では、デバイスは、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するように構成される１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される。

【０００７】

[0007]別の態様では、デバイスは、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを取得するための手段と、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方は少なくとも、複数の球面調和係数に関して変換を実行することによって生成される、１つまたは複数の第１のベクトルを記憶するための手段とを備える。

【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、実行されると、１つまたは複数のプロセッサに、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを取得させる命令を記憶しており、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される。

【０００９】

[0009]別の態様では、方法は、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の解凍スキームの１つを選択することと、複数の解凍スキームの選択された１つを使用して球面調和係数の圧縮されたバージョンを解凍することとを備える。

【００１０】

[0010]別の態様では、デバイスは、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の解凍スキームの１つを選択し、複数の解凍スキームの選択された１つを使用して球面調

和係数の圧縮されたバージョンを解凍するように構成される１つまたは複数のプロセッサを備える。

【００１１】

[0011]別の態様では、デバイスは、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の解凍スキームの１つを選択するための手段と、複数の解凍スキームの選択された１つを使用して球面調和係数の圧縮されたバージョンを解凍するための手段とを備える。

【００１２】

[0012]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、統合された復号デバイスの１つまたは複数のプロセッサに、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の解凍スキームの１つを選択させ、複数の解凍スキームの選択された１つを使用して球面調和係数の圧縮されたバージョンを解凍させる命令を記憶している。

【００１３】

[0013]2別の態様では、方法は、音場を表す球面調和係数が合成オーディオオブジェクトから生成されたかどうかのインジケーションを取得することを備える。

【００１４】

[0014]別の態様では、デバイスは、音場を表す球面調和係数が合成オーディオオブジェクトから生成されたかどうかのインジケーションを取得するように構成される１つまたは複数のプロセッサを備える。

【００１５】

[0015]別の態様では、デバイスは、音場を表す球面調和係数を記憶するための手段と、球面調和係数が合成オーディオオブジェクトから生成されるかどうかのインジケーションを取得するための手段とを備える。

【００１６】

[0016]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかのインジケーションを取得させる命令を記憶している。

【００１７】

[0017]別の態様では、方法は、音場の１つまたは複数の成分を表す１つまたは複数の第１のベクトルを量子化することと、１つまたは複数の第１のベクトルの量子化が原因で音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにもたらされた誤差を補償することとを備える。

【００１８】

[0018]別の態様では、デバイスは、音場の１つまたは複数の成分を表す１つまたは複数の第１のベクトルを量子化し、１つまたは複数の第１のベクトルの量子化が原因で音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにもたらされた誤差を補償するように構成される１つまたは複数のプロセッサを備える。

【００１９】

[0019]別の態様では、デバイスは、音場の１つまたは複数の成分を表す１つまたは複数の第１のベクトルを量子化するための手段と、１つまたは複数の第１のベクトルの量子化が原因で音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにもたらされた誤差を補償するための手段とを備える。

【００２０】

[0020]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場の１つまたは複数の成分を表す１つまたは複数の第１のベクトルを量子化させ、１つまたは複数の第１のベクトルの量子化が原因で音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにもたらされた誤差を補償させる命令を記憶している。

【００２１】

[0021]別の態様では、方法は、目標ビットレートに基づいて、低減された球面調和係数または低減された球面調和係数のそれの分解を生成するために、複数の球面調和係数または複数の球面調和係数のそれの分解に関して次数低減を実行することを備え、複数の球面調和係数は音場を表す。

【００２２】

[0022]別の態様では、デバイスは、目標ビットレートに基づいて、低減された球面調和係数または低減された球面調和係数のそれの分解を生成するために、複数の球面調和係数または複数の球面調和係数のそれの分解に関して次数低減を実行するように構成される１つまたは複数のプロセッサを備え、ここにおいて、複数の球面調和係数は音場を表す。

【００２３】

[0023]別の態様では、デバイスは、複数の球面調和係数または複数の球面調和係数のそれの分解を記憶するための手段と、目標ビットレートに基づいて、低減された球面調和係数または低減された球面調和係数のそれの分解を生成するために、複数の球面調和係数または複数の球面調和係数のそれの分解に関して次数低減を実行するための手段とを備え、ここにおいて、複数の球面調和係数は音場を表す。

【００２４】

[0024]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、目標ビットレートに基づいて、低減された球面調和係数または低減された球面調和係数のそれの分解を生成するために、複数の球面調和係数または複数の球面調和係数のそれの分解に関して次数低減を実行させる命令を記憶しており、ここにおいて、複数の球面調和係数は音場を表す。

【００２５】

[0025]別の態様では、方法は、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを取得することを備え、このベクトルは、音場を記述する複数の球面調和係数から分解されたものである。

【００２６】

[0026]別の態様では、デバイスは、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを取得するように構成される１つまたは複数のプロセッサを備え、このベクトルは、音場を記述する複数の球面調和係数から分解されたものである。

【００２７】

[0027]別の態様では、デバイスは、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを取得するための手段と、このベクトルは音場を記述する複数の球面調和係数から分解されたものである、係数の第１の０ではないセットを記憶するための手段とを備える。

【００２８】

[0028]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを決定させる命令を記憶しており、このベクトルは、音場を記述する複数の球面調和係数から分解されたものである。

【００２９】

[0029]別の態様では、方法は、ビットストリームから、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルの少なくとも１つを取得することを備え、ここにおいて、球面調和係数は音場を記述し、バックグラウンド球面調和係数は同じ音場の１つまたは複数のバックグラウンド成分を記述した。

【００３０】

[0030]別の態様では、デバイスは、ビットストリームから、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルの少なくとも１つを決定するように構成される１つまたは複数のプロセッサを備え、ここにおいて、球面調和係数は音場を記述し、バックグラウンド球面調和係数は同じ音場の１つまたは複数のバックグラウンド成分を記述した。

【００３１】
[0031]別の態様では、デバイスは、ビットストリームから、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルの少なくとも１つを取得するための手段を備え、ここにおいて、球面調和係数は音場を記述し、バックグラウンド球面調和係数は同じ音場の１つまたは複数のバックグラウンド成分を記述した。
【００３２】
[0032]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、ビットストリームから、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルの少なくとも１つを取得させる命令を記憶しており、ここにおいて、球面調和係数は音場を記述し、バックグラウンド球面調和係数は同じ音場の１つまたは複数のバックグラウンド成分を記述した。
【００３３】
[0033]別の態様では、方法は、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別することを備える。
【００３４】
[0034]別の態様では、デバイスは、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別するように構成される１つまたは複数のプロセッサを備える。
【００３５】
[0035]別の態様では、デバイスは、１つまたは複数の球面調和係数（ＳＨＣ）を記憶するための手段と、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別するための手段とを備える。
【００３６】
[0036]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別させる命令を記憶している。
【００３７】
[0037]別の態様では、方法は、１つまたは複数のオーディオオブジェクトと対応する指向性情報とを表す複数の球面調和係数の分解された表現を生成するために複数の球面調和係数に関してベクトルベースの合成を実行することと、ここにおいて、球面調和係数はある次数と関連付けられ音場を記述する、指向性情報から明瞭な指向性情報とバックグラウンド指向性情報とを決定することと、変換されたバックグラウンド指向性情報を生成するためにバックグラウンドオーディオオブジェクトと関連付けられる指向性情報の次数を低減することと、音場の全体的なエネルギーを保存するために変換された指向性情報の値を増大させるように補償を適用することとを備える。
【００３８】
[0038]別の態様では、デバイスは、１つまたは複数のオーディオオブジェクトと対応する指向性情報とを表す複数の球面調和係数の分解された表現を生成するために複数の球面調和係数に関してベクトルベースの合成を実行し、ここにおいて、球面調和係数はある次数と関連付けられ音場を記述する、指向性情報から明瞭な指向性情報とバックグラウンド指向性情報とを決定し、変換されたバックグラウンド指向性情報を生成するためにバック

グラウンドオーディオオブジェクトと関連付けられる指向性情報の次数を低減し、音場の全体的なエネルギーを保存するために変換された指向性情報の値を増大させるように補償を適用するように構成される１つまたは複数のプロセッサを備える。

【００３９】

[0039]別の態様では、デバイスは、１つまたは複数のオーディオオブジェクトと対応する指向性情報とを表す複数の球面調和係数の分解された表現を生成するために複数の球面調和係数に関してベクトルベースの合成を実行するための手段と、ここにおいて、球面調和係数はある次数と関連付けられ音場を記述する、指向性情報から明瞭な指向性情報とバックグラウンド指向性情報とを決定するための手段と、変換されたバックグラウンド指向性情報を生成するためにバックグラウンドオーディオオブジェクトと関連付けられる指向性情報の次数を低減するための手段と、音場の全体的なエネルギーを保存するために変換された指向性情報の値を増大させるように補償を適用するための手段とを備える。

【００４０】

[0040]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、１つまたは複数のオーディオオブジェクトと対応する指向性情報とを表す複数の球面調和係数の分解された表現を生成するために複数の球面調和係数に関してベクトルベースの合成を実行させ、ここにおいて、球面調和係数はある次数と関連付けられ音場を記述する、指向性情報から明瞭な指向性情報とバックグラウンド指向性情報とを決定させ、変換されたバックグラウンド指向性情報を生成するためにバックグラウンドオーディオオブジェクトと関連付けられる指向性情報の次数を低減させ、音場の全体的なエネルギーを保存するために変換された指向性情報の値を増大させるように補償を適用させる命令を記憶している。

【００４１】

[0041]別の態様では、方法は、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得することを備える。

【００４２】

[0042]別の態様では、デバイスは、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得するように構成される１つまたは複数のプロセッサを備える。

【００４３】

[0043]別の態様では、デバイスは、第１の複数の球面調和係数と第２の複数の球面調和係数とを記憶するための手段と、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得するための手段とを備える。

【００４４】

[0044]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得させる命令を記憶している。

【００４５】

[0045]別の態様では、方法は、音場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を取得することを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００４６】

[0046]別の態様では、デバイスは、音場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を取得するように構成される１つまたは複数のプロセッサを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成され

る。
【００４７】
[0047]別の態様では、デバイスは、音場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を取得するための手段と、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、ビットストリームを記憶するための手段とを備える。
【００４８】
[0048]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場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を取得させる命令を記憶しており、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００４９】
[0049]別の態様では、方法は、音場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を生成することを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００５０】
[0050]別の態様では、デバイスは、音場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を生成するように構成される１つまたは複数のプロセッサを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００５１】
[0051]別の態様では、デバイスは、音場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を生成するための手段と、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、ビットストリームを記憶するための手段とを備える。
【００５２】
[0052]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場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を生成させる命令を有し、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００５３】
[0053]別の態様では、方法は、複数の圧縮された空間成分のある空間成分の圧縮されたバージョンを複数の圧縮された空間成分の残りに対するその空間成分の圧縮されたバージョンの次数に基づいて解凍するときに使用するハフマンコードブックを識別することを備え、その空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００５４】
[0054]別の態様では、デバイスは、複数の圧縮された空間成分のある空間成分の圧縮されたバージョンを複数の圧縮された空間成分の残りに対するその空間成分の圧縮されたバージョンの次数に基づいて解凍するときに使用するハフマンコードブックを識別するように構成される１つまたは複数のプロセッサを備え、その空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【００５５】
[0055]別の態様では、デバイスは、複数の圧縮された空間成分のある空間成分の圧縮されたバージョンを複数の圧縮された空間成分の残りに対するその空間成分の圧縮されたバージョンの次数に基づいて解凍するときに使用するハフマンコードブックを識別するための手段と、その空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、複数の圧縮された空間成分を記憶するための手段とを備える。
【００５６】
[0056]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまた

は複数のプロセッサに、複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて解凍するときに使用するハフマンコードブックを識別させる命令を記憶しており、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００５７】

[0057]別の態様では、方法は、複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別することを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００５８】

[0058]別の態様では、デバイスは、複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別するように構成される１つまたは複数のプロセッサを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００５９】

[0059]別の態様では、デバイスは、ハフマンコードブックを記憶するための手段と、複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別するための手段とを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００６０】

[0060]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別させる命令を記憶しており、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００６１】

[0061]別の態様では、方法は、音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００６２】

[0062]別の態様では、デバイスは、音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定するように構成される１つまたは複数のプロセッサを備え、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００６３】

[0063]別の態様では、デバイスは、音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定するための手段と、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、量子化ステップサイズを記憶するための手段とを備える。

【００６４】

[0064]別の態様では、非一時的コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定させる命令を記憶しており、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【００６５】

[0065]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。これらの技法の他の特徴、目的、および利点は、説明および図面から、ならびに特許請求の範囲から、明らかになろう。

【図面の簡単な説明】

【００６６】
【図１】様々な次数および位数の球面調和基底関数を示す図。
【図２】様々な次数および位数の球面調和基底関数を示す図。
【図３】本開示において説明される技法の様々な態様を実行し得るシステムを示す図。
【図４】本開示において説明される技法の様々な態様を実行し得る図３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。
【図５】図３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６】本開示において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符号化デバイスのコンテンツ分析ユニットの例示的な動作を示すフローチャート。
【図７】本開示において説明されるベクトルベースの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８】本開示において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９Ａ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｂ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｃ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｄ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｅ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｆ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｇ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｈ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｉ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｊ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｋ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図９Ｌ】図４の例のオーディオ符号化デバイスのある態様をより詳細に示すブロック図。
【図１０Ａ】圧縮された空間成分を指定し得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１０Ｂ】圧縮された空間成分を指定し得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１０Ｃ】圧縮された空間成分を指定し得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１０Ｃ（ｉ）】圧縮された空間成分を指定し得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１０Ｄ】圧縮された空間成分を指定し得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１０Ｄ（ｉ）】圧縮された空間成分を指定し得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１０Ｅ】圧縮された空間成分を指定し得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１０Ｅ（ｉ）】圧縮された空間成分を指定し得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１０Ｆ】圧縮された空間成分を指定し得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１０Ｇ】圧縮された空間成分を指定し得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１０Ｈ（ｉ）】圧縮された空間成分を指定し得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１０Ｈ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｉ（ｉ）】圧縮された空間成分を指定し得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１０Ｉ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｊ（ｉ）】圧縮された空間成分を指定し得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１０Ｊ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｋ（ｉ）】圧縮された空間成分を指定し得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１０Ｋ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｌ（ｉ）】圧縮された空間成分を指定し得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１０Ｌ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｍ（ｉ）】圧縮された空間成分を指定し得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１０Ｍ（ｉｉ）】圧縮された空間成分を指定し得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１０Ｎ（ｉ）】圧縮された空間成分を指定し得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１０Ｎ（ｉｉ）】圧縮された空間成分を指定し得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１００（ｉ）】圧縮された空間成分を指定し得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１００（ｉｉ）】圧縮された空間成分を指定し得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１１Ａ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１１Ｂ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１１Ｃ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１１Ｄ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１１Ｅ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１１Ｆ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示

すブロック図。
【図１１Ｇ】図５の例に示されるオーディオ復号デバイスのあるユニットをより詳細に示すブロック図。
【図１２】本開示において説明される技法の様々な態様を実行し得る例示的なオーディオエコシステムを示す図。
【図１３】図１２のオーディオエコシステムの一例をより詳細に示す図。
【図１４】図１２のオーディオエコシステムの一例をより詳細に示す図。
【図１５Ａ】図１２のオーディオエコシステムの他の例をより詳細に示す図。
【図１５Ｂ】図１２のオーディオエコシステムの他の例をより詳細に示す図。
【図１６】本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ符号化デバイスを示す図。
【図１７】図１６のオーディオ符号化デバイスの一例をより詳細に示す図。
【図１８】本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ復号デバイスを示す図。
【図１９】図１８のオーディオ復号デバイスの一例をより詳細に示す図。
【図２０Ａ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｂ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｃ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｄ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｅ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｆ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２０Ｇ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ取得デバイスを示す図。
【図２１Ａ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ再生デバイスを示す図。
【図２１Ｂ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ再生デバイスを示す図。
【図２１Ｃ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ再生デバイスを示す図。
【図２１Ｄ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ再生デバイスを示す図。
【図２１Ｅ】本開示において説明される技法のある態様を実行し得る例示的なオーディオ再生デバイスを示す図。
【図２２Ａ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｂ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｃ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｄ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｅ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｆ】本開示において説明される１つまたは複数の技法による例示的なオーディオ

再生環境を示す図。
【図２２Ｇ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２２Ｈ】本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図。
【図２３】本開示において説明される１つまたは複数の技法による、ユーザがヘッドフォンを装着しながらスポーツの試合の３Ｄ音場を経験し得る例示的な使用事例を示す図。
【図２４】本開示において説明される１つまたは複数の技法による、３Ｄ音場が録音され得るスポーツスタジアムを示す図。
【図２５】本開示において説明される１つまたは複数の技法による、ローカルのオーディオランドスケープに基づいて３Ｄ音場をレンダリングするための技法を示すフローチャート。
【図２６】本開示において説明される１つまたは複数の技法による例示的なゲームスタジオを示す図。
【図２７】本開示において説明される１つまたは複数の技法によるレンダリングエンジンを含む複数のゲームシステムを示す図。
【図２８】本開示において説明される１つまたは複数の技法による、ヘッドフォンにより模擬され得るスピーカー構成を示す図。
【図２９】本開示において説明される１つまたは複数の技法による、３Ｄ音場を取得および／または編集するために使用され得る複数のモバイルデバイスを示す図。
【図３０】本開示において説明される１つまたは複数の技法に従って処理され得る３Ｄ音場と関連付けられるビデオフレームを示す図。
【図３１Ａ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｂ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｃ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｄ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｅ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｆ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｇ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｈ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｉ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｊ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｋ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｌ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３１Ｍ】本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行のあるシミュレーション結果を示すグラフを示す図。
【図３２】本開示において説明される技法による、高次アンビソニックス係数から分解されたＳ行列からの特異値のグラフを示す図。

【図３３Ａ】本開示において説明される技法による、音場のフォアグラウンド成分を記述するベクトルを符号化するときに再順序付けが有する潜在的な影響を示すグラフを示す図。
【図３３Ｂ】本開示において説明される技法による、音場のフォアグラウンド成分を記述するベクトルを符号化するときに再順序付けが有する潜在的な影響を示すグラフを示す図。
【図３４】本開示による、明瞭なオーディオオブジェクトのエネルギーのみに基づく識別と指向性のみに基づく識別との差を示す概念図。
【図３５】本開示による、明瞭なオーディオオブジェクトのエネルギーのみに基づく識別と指向性のみに基づく識別との差を示す概念図。
【図３６Ａ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｂ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｃ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｄ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｅ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｆ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３６Ｇ】本開示において説明される技法のある態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図。
【図３７】本明細書で説明されるように空間－時間的な補間を得るための技法の表現を示す図。
【図３８】本明細書で説明される技法に従った、多次元信号のためのシーケンシャルＳＶＤブロックのための人工的なＵＳ行列、すなわちUS_1とUS_2とを示すブロック図。
【図３９】本開示において説明される技法による、特異値分解と空間－時間的成分の平滑化とを使用した高次アンビソニックス（ＨＯＡ）信号の後続のフレームの分解を示すブロック図。
【図４０Ａ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｂ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｃ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｄ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｅ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｆ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｇ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示

において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｈ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｉ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４０Ｊ】２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイスを示すブロック図。
【図４１Ａ】２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ復号デバイスを示すブロック図。
【図４１Ｂ】２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ復号デバイスを示すブロック図。
【図４１Ｃ】２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ復号デバイスを示すブロック図。
【図４１Ｄ】２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ復号デバイスを示すブロック図。
【図４２Ａ】図４０Ｂ～図４０Ｊの例に示される次数低減ユニットをより詳細に示すブロック図。
【図４２Ｂ】図４０Ｂ～図４０Ｊの例に示される次数低減ユニットをより詳細に示すブロック図。
【図４２Ｃ】図４０Ｂ～図４０Ｊの例に示される次数低減ユニットをより詳細に示すブロック図。
【図４３】図４０Ｉに示されるＶ圧縮ユニットをより詳細に示す図。
【図４４】本開示において説明される技法の様々な態様による、量子化誤差を補償するためにオーディオ符号化デバイスによって実行される例示的な動作を示す図。
【図４５Ａ】本開示において説明される技法のある態様による、２つのフレームの部分からのサブフレームの補間を示す図。
【図４５Ｂ】本開示において説明される技法のある態様による、２つのフレームの部分からのサブフレームの補間を示す図。
【図４６Ａ】本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図。
【図４６Ｂ】本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図。
【図４６Ｃ】本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図。
【図４６Ｄ】本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図。
【図４６Ｅ】本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図。
【図４７】図４１Ａ～図４１Ｄの例に示されるオーディオ復号デバイスの抽出ユニットをより詳細に示すブロック図。
【図４８】図４１Ａ～図４１Ｄの例に示されるオーディオ復号デバイスのオーディオレンダリングユニットをより詳細に示すブロック図。

【図４９Ａ】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｂ】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｃ】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｄ（ｉ）】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｄ（ｉｉ）】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｅ（ｉ）】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図４９Ｅ（ｉｉ）】本開示において説明される技法のある態様を実装し得るオーディオコーディングシステムを示す図。
【図５０Ａ】本開示において説明される技法による、バックグラウンドコンテンツの次数を潜在的に低減するための２つの異なる手法の１つを示すブロック図。
【図５０Ｂ】本開示において説明される技法による、バックグラウンドコンテンツの次数を潜在的に低減するための２つの異なる手法の１つを示すブロック図。
【図５１】球面調和係数を圧縮するために本開示で説明された技法の様々な態様を実施し得るオーディオ符号化デバイスの明瞭な成分の圧縮経路の例を示すブロック図。
【図５２】球面調和係数（ＳＨＣ）を再構築またはほぼ再構築するために本開示で説明された技法の様々な態様を実施し得るオーディオ復号デバイスの別の例を示すブロック図。
【図５３】本開示において説明される技法の様々な態様を実行し得るオーディオ符号化デバイスの別の例を示すブロック図。
【図５４】図５３の例に示されるオーディオ符号化デバイスの例示的な実装形態をより詳細に示すブロック図。
【図５５Ａ】音場を回転させるために本開示において説明される技法のある態様を実行する例を示す図。
【図５５Ｂ】音場を回転させるために本開示において説明される技法のある態様を実行する例を示す図。
【図５６】第１の基準フレームに従ってキャプチャされ、次いで第２の基準フレームに対して音場を表すために本開示において説明される技法に従って回転される例示的な音場を示す図。
【図５７Ａ】本開示において説明される技法に従って形成されるビットストリームを示す図。
【図５７Ｂ】本開示において説明される技法に従って形成されるビットストリームを示す図。
【図５７Ｃ】本開示において説明される技法に従って形成されるビットストリームを示す図。
【図５７Ｄ】本開示において説明される技法に従って形成されるビットストリームを示す図。
【図５７Ｅ】本開示において説明される技法に従って形成されるビットストリームを示す図。
【図５８】本開示において説明される技法の回転態様を実施する際の図５３の例に示される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５９】本開示において説明される技法の変換態様を実行する際の図５３の例に示される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。
【発明を実施するための形態】
【００６７】
[0123]サラウンド音響の発展は、今日の娯楽のための多くの出力フォーマットを利用可

能にしてきた。そのような消費者向けのサラウンド音響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「チャネル」ベースである。これらは、普及している５．１フォーマット（これは、フロントレフト（ＦＬ）と、フロントライト（ＦＲ）と、センターまたはフロントセンターと、バックレフトまたはサラウンドレフトと、バックライトまたはサラウンドライトと、低周波効果（ＬＦＥ）という、６つのチャネルを含む）、発展中の７．１フォーマット、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、切られた二十面体（truncated icosohedron）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。

【００６８】

[0124]将来のＭＰＥＧ符号化器への入力は、任意選択で、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）従来のチャネルベースオーディオ、（ｉｉ）（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（ｉｉｉ）球面調和基底関数の係数（「球面調和係数」すなわちＳＨＣ、「高次アンビソニックス」すなわちＨＯＡ、および「ＨＯＡ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。この未来のＭＰＥＧ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／ｗ１３４１１．ｚｉｐにおいて利用可能な、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。

【００６９】

[0125]市場には様々な「サラウンド音響」チャ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成することを望み、スピーカー構成ごとにサウンドトラックをリミックス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。

【００７０】

[0126]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。

【００７１】

[0127]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。

【００７２】

【数１】

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

【００７３】

[0128]この式は、時間ｔにおける音場の任意の点｛r_r，θ_r，φ_r｝における圧力p_iが、ＳＨＣ、

【００７４】

【数２】

$$A_n^m(k)$$

【００７５】

によって一意に表され得ることを示す。ここで、

【００７６】

【数３】

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

【００７７】

であり、ｃは、音の速さ（約３４３ｍ／ｓ）であり、｛r_r，θ_r，φ_r｝は、基準の点（または観測点）であり、j_n（・）は、次数ｎの球ベッセル関数であり、

【００７８】

【数４】

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

【００７９】

は、次数ｎおよび位数ｍの球面調和基底関数である。大括弧中の項は、離散フーリエ変換（ＤＦＴ）、離散コサイン変換（ＤＣＴ）、またはウェーブレット変換などの、様々な時間－周波数変換によって近似され得る信号の周波数領域表現（すなわち、Ｓ｛ω_r、r_r、θ_r、φ_r｝）であることが、認識され得る。階層セットの他の例は、ウェーブレット変換係数のセット、および多分解能基底関数の係数の他のセットを含む。

【００８０】

[0129]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。理解され得るように、各次数に対して、図示されているが説明を簡単にするために図１の例では明示的に述べられていない位数ｍという拡張が存在する。

【００８１】

[0130]図２は、０次（ｎ＝０）から４次（ｎ＝４）までの球面調和基底関数を示す別の図である。図２では、球面調和基底関数は、図示された次数と位数の両方を伴う３次元座標空間において示されている。

【００８２】

[0131]

【００８３】

【数５】

$$\mathrm{SHC}\ A_n^m(k)$$

【００８４】
は、様々なマイクロフォンアレイ構成によって物理的に取得（たとえば、記録）されてよく、または代替的に、それらは音場のチャ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進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。

【００８５】
[0132]上で述べられたように、ＳＨＣはマイクロフォン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、ｐｐ．１００４－１０２５において説明されている。

【００８６】
[0133]これらのＳＨＣがどのようにオブジェクトベースの記述から導出され得るかを例示するために、次の式を考える。
個々のオーディオオブジェクトに対応する音場に対する

【００８７】
【数６】

$$A_n^m(k)$$

【００８８】
は、以下のように表わされることが可能である：

【００８９】
【数７】

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

【００９０】
ここで、ｉは、

【００９１】
【数８】

$$\sqrt{-1}$$

【００９２】
であり、

【００９３】
【数９】

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

【００９４】

は、次数ｎの（第２の種類の）球ハンケル関数であり、｛ｒ$_s$，θ$_s$、φ$_s$｝は、オブジェクトのロケーションである。オブジェクトソースエネルギーｇ（ω）を（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の時間周波数分析技法を使用して）周波数の関数として知ることによって、各ＰＣＭオブジェクトとそのロケーションとを

【００９５】
【数１０】

$$\text{SHC } A_n^m(k)$$

【００９６】
へと変換することが可能になる。さらに、各オブジェクトの
【００９７】
【数１１】

$$A_n^m(k)$$

【００９８】
係数は、（上記が線形であり直交分解であるので）加法的であることが示され得る。このようにして、多数のＰＣＭオブジェクトが、
【００９９】
【数１２】

$$A_n^m(k)$$

【０１００】
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$_r$，θ$_r$，φ$_r$｝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。
【０１０１】
[0134]図３は、本開示で説明される技法の様々な態様を実行し得るシステム１０を示す図である。図３の例に示されているように、システム１０は、コンテンツ作成者１２とコンテンツ消費者１４とを含む。コンテンツ作成者１２およびコンテンツ消費者１４の文脈で説明されているが、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てよい。その上、コンテンツ作成者１２は、いくつか例を挙げると、ハンドセット（または携帯電話）、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１４は、いくつか例を挙げると、ハンドセット（または携帯電話）、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。
【０１０２】
[0135]コンテンツ作成者１２は、コンテンツ消費者１４などのコンテンツ消費者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し得る映画撮影所または他のエ

ンティティを表し得る。いくつかの例では、コンテンツ作成者１２は、ＨＯＡ係数１１を圧縮することを望む個々のユーザを表し得る。しばしば、このコンテンツ作成者は、ビデオコンテンツと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１４は、オーディオ再生システムへのアクセス権を所有するまたは有する個人を表し、このオーディオ再生システムは、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。図３の例では、コンテンツ消費者１４は、オーディオ再生システム１６を含む。

【０１０３】

[0136]コンテンツ作成者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１２は、様々なフォーマットのライブ録音７（直接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１２はこれらをオーディオ編集システム１８を使用して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様相を識別しようとしてレンダリングされたスピーカーフィードを聞くことができる。コンテンツ作成者１２は次いで、（場合によっては、上で説明された方式でソース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を利用することができる。音声編集システム１８は、音声データを編集し、１つまたは複数のソース球面調和係数としてこの音声データを出力することが可能な任意のシステムを表す。

【０１０４】

[0137]編集プロセスが完了すると、コンテンツ作成者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を符号化しまたは別様に圧縮するように構成されるデバイスを表す、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。

【０１０５】

[0138]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、ベクトルベースの合成または指向性ベースの合成に基づいてＨＯＡ係数１１を符号化するように構成され得る。ベクトルベースの合成方法を実行するか指向性ベースの合成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２０は、ＨＯＡ係数１１に少なくとも一部基づいて、ＨＯＡ係数１１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、ＰＣＭ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）生成されたかを決定することができる。ＨＯＡ係数１１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２０は、指向性ベースの合成方法を使用してＨＯＡ係数１１を符号化することができる。ＨＯＡ係数１１がたとえば固有マイクを使用してライブをキャプチャしたとき、オーディオ符号化デバイス２０は、ベクトルベースの合成方法に基づいてＨＯＡ係数１１を符号化することができる。上の区別は、ベクトルベースの合成方法または指向性ベースの合成方法が利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、ＨＯＡ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。

【０１０６】

[0139]説明を目的に、ＨＯＡ係数１１がライブをキャプチャした、またはそうで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２０が決定すると

仮定すると、オーディオ符号化デバイス２０は、線形可逆変換（ＬＩＴ）の適用を伴うベクトルベースの合成方法を使用してＨＯＡ係数１１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「ＳＶＤ」）と呼ばれる。この例では、オーディオ符号化デバイス２０は、ＨＯＡ係数１１の分解されたバージョンを決定するためにＳＶＤをＨＯＡ係数１１に適用することができる。オーディオ符号化デバイス２０は次いで、様々なパラメータを識別するためにＨＯＡ係数１１の分解されたバージョンを分析することができ、このことは、ＨＯＡ係数１１の分解されたバージョンの再順序付けを容易にし得る。オーディオ符号化デバイス２０は次いで、識別されたパラメータに基づいてＨＯＡ係数１１の分解されたバージョンを再順序付けることができ、そのような再順序付けは、以下でさらに詳細に説明されるように、変換がＨＯＡ係数のフレームにわたってＨＯＡ係数を再順序付けることができるとすると（フレームが一般にＨＯＡ係数１１のＭ個のサンプルを含み、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。ＨＯＡ係数１１の分解されたバージョンを再順序付けた後で、オーディオ符号化デバイス２０は、ＨＯＡ係数１１の分解されたバージョンのうちで、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、または目立つ）成分を表すもの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、フォアグラウンド成分を表すＨＯＡ係数１１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。

【０１０７】

[0140]オーディオ符号化デバイス２０はまた、ＨＯＡ係数１１のうちで、少なくとも一部、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）成分を表すものを識別するために、ＨＯＡ係数１１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２０は、いくつかの例では、バックグラウンド成分がＨＯＡ係数１１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ものではなく、０次および１次の球面基底関数に対応するものなど）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２０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、ＨＯＡ係数１１の残りのバックグラウンドＨＯＡ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。

【０１０８】

[0141]オーディオ符号化デバイス２０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ＨＯＡ係数１１の各々に関して、ある形態の音響心理学的符号化（ＭＰＥＧサラウンド、ＭＰＥＧ－ＡＡＣ、ＭＰＥＧ－ＵＳＡＣ、または他の既知の形態の音響心理学的符号化）を実行することができる。オーディオ符号化デバイス２０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２０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。いくつかの例では、この量子化はスカラー／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２０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２１を形成することができる。オーディオ符号化デバイス２０は次いで、ビットストリーム２１をコンテンツ消費者１４に送信し、または別様に出力することができる。

【０１０９】

[0142]図３ではコンテンツ消費者１４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１２は、コンテンツ作成者１２とコンテンツ消費者１４との間に配置された中間デバイスにビットストリーム２１を出力することができる。この中間デバイス

は、ビットストリーム２１を要求し得るコンテンツ消費者１４に後で配信するために、このビットストリーム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、携帯電話、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２１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。この中間デバイスは、ビットストリーム２１を要求するコンテンツ消費者１４などの加入者にビットストリーム２１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に存在してもよい。

【０１１０】

[0143]代替的に、コンテンツ作成者１２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２１を格納することができ、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、これらの媒体に格納されたコンテンツが送信されるチャネルを指し得る（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３の例に限定されるべきではない。

【０１１１】

[0144]図３の例にさらに示されるように、コンテンツ消費者１４は、オーディオ再生システム１６を含む。オーディオ再生システム１６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１６は、いくつかの異なるレンダラ２２を含み得る。レンダラ２２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ニング（ＶＢＡＰ：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。

【０１１２】

[0145]オーディオ再生システム１６はさらに、オーディオ復号デバイス２４を含み得る。オーディオ復号デバイス２４は、ビットストリーム２１からＨＯＡ係数１１’を復号するように構成されるデバイスを表すことができ、ＨＯＡ係数１１’は、ＨＯＡ係数１１と類似し得るが、有損失の演算（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ＨＯＡ係数に関して音響心理学的復号を実行する。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ＨＯＡ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ＨＯＡ係数およびバックグラウンド成分を表す復号されたＨＯＡ係数に基づいて、ＨＯＡ係数１１’を決定することができる。

【０１１３】

[0146]オーディオ再生システム１６は、ＨＯＡ係数１１’を取得するためにビットストリーム２１を復号した後で、ラウドスピーカーフィード２５を出力するためにＨＯＡ係数１１’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３の例には示されていない）を駆動することができる。

【０１１４】

[0147]適切なレンダラを選択するために、またはいくつかの例では、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。いくつかの例では、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式でラウドスピーカーを駆動することができる。他の例では、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１６を入力するようにユーザに促すことができる。

【０１１５】

[0148]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１つを選択することができる。いくつかの例では、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閾値内にないとき、オーディオ再生システム１６は、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、いくつかの例では、オーディオレンダラ２２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１つを生成することができる。

【０１１６】

[0149]図４は、本開示において説明される技法の様々な態様を実行し得る図３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、ベクトルベース合成方法ユニット２７と、指向性ベース合成方法ユニット２８とを含む。

【０１１７】

[0150]コンテンツ分析ユニット２６は、ＨＯＡ係数１１がライブ録音または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どうかを識別するために、ＨＯＡ係数１１のコンテンツを分析するように構成されるユニットを表す。コンテンツ分析ユニット２６は、ＨＯＡ係数１１が実際の音場の記録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。コンテンツ分析ユニット２６は、この決定を様々な方法で行うことができる。たとえば、コンテンツ分析ユニット２６は、$（Ｎ＋１）^2－１$個のチャネルをコーディングし、最後の残りのチャネル（ベクトルとして表され得る）を予測することができる。コンテンツ分析ユニット２６は、最後の残りのチャネルを決定するために、$（Ｎ＋１）^2－１$個のチャネルのうち少なくともいくつかにスカラーを適用し、結果として得られる値を加算することができる。さらに、この例では、コンテンツ分析ユニット２６は、予測されたチャネルの精度を決定することができる。この例では、予測されたチャネルの精度が比較的高い（たとえば、精度が特定閾値を超える）場合、ＨＯＡ係数１１は、合成オーディオオブジェクトから生成される可能性が高い。対照的に、予測されたチャネルの精度が比較的低い（たとえば、精度が特定の閾値を下回る）場合、ＨＯＡ係数１１は、記録された音場を表す可能性が高い。たとえば、この例では、予測されたチャネルの信号対雑音比（ＳＮＲ）が１００デシベル（ｄｂ）を超える場合、ＨＯＡ係数１１は、合成オーディオオブジェクトから生成された音場を表す可能性が高い。対照的に、ｅｉｇｅｎマイクロフォンを使用して記録された音場のＳＮＲは５～２０ｄｂであり得る。したがって、実際の直接的な録音から生成されたＨＯＡ係数１１によって表される音場と合成オーディオオブジェクトから生成されたＨＯＡ係数１１によって表される音場の間には、ＳＮＲ比における明らかな境界が存在し得る。

【０１１８】

[0151]より具体的には、コンテンツ分析ユニット２６は、音場を表すＨＯＡ係数１１が合成オーディオオブジェクトから生成されたかどうかを決定するとき、ＨＯＡ係数のフレームを取得することができ、これは、４次の表現（すなわち、Ｎ＝４）では２５×１０２４のサイズであり得る。フレーム分割されたＨＯＡ係数（本明細書ではフレーム分割され

たＳＨＣ行列１１とも呼ばれることがあり、後続のフレーム分割されたＳＨＣ行列はフレーム分割されたＳＨＣ行列２７Ｂ、２７Ｃなどと示され得る）を取得した後で。コンテンツ分析ユニット２６は次いで、低減されたフレーム分割されたＨＯＡ係数を生成するために、フレーム分割されたＨＯＡ係数１１の第１のベクトルを除外することができる。いくつかの例では、フレーム分割されたＨＯＡ係数１１から除外されるこの第１のベクトルは、ＨＯＡ係数１１のうちで、０次、０位の球面調和基底関数と関連付けられるものに対応し得る。

【０１１９】

[0152]コンテンツ分析ユニット２６は次いで、低減されたフレーム分割されたＨＯＡ係数の残りのベクトルから、低減されたフレーム分割されたＨＯＡ係数の第１の０ではないベクトルを予測することができる。第１の０ではないベクトルは、０以外の値を有する１次（次数に依存する位数の各々を考慮する）から４次（次数に依存する位数の各々を考慮する）までの第１のベクトルを指し得る。いくつかの例では、低減されたフレーム分割されたＨＯＡ係数の第１の０ではないベクトルは、ＨＯＡ係数１１のうちで、１次、０位の球面調和基底関数と関連付けられるものを指す。第１の０ではないベクトルに関して説明されるが、本技法は、低減されたフレーム分割されたＨＯＡ係数の残りのベクトルから、低減されたフレーム分割されたＨＯＡ係数の他のベクトルを予測することができる。たとえば、コンテンツ分析ユニット２６は、低減されたフレーム分割されたＨＯＡ係数のうちで、１次、１位の球面調和基底関数、または１次、－１次の球面調和基底関数と関連付けられるものを予測することができる。さらに他の例として、コンテンツ分析ユニット２６は、低減されたフレーム分割されたＨＯＡ係数のうちで、２次、０次の球面調和基底関数と関連付けられるものを予測することができる。

【０１２０】

[0153]第１の０ではないベクトルを予測するために、コンテンツ分析ユニット２６は次の式に従って演算することができる：

【０１２１】

【数１３】

$$\sum_i (\alpha_i v_i)$$

【０１２２】

ここで、ｉは、１から（Ｎ＋１）^{2}－２であり、これは４次の表現では２３であり、α_iはｉ番目のベクトルの何らかの定数を示し、v_iはｉ番目のベクトルを指す。第１の０ではないベクトルを予測した後で、コンテンツ分析ユニット２６は、予測された第１の０ではないベクトルおよび実際の０ではないベクトルに基づいて、誤差を取得することができる。いくつかの例では、コンテンツ分析ユニット２６は、誤差を導出するために実際の第１の０ではないベクトルから予測された第１の０ではないベクトルを差し引く。コンテンツ分析ユニット２６は、予測された第１の０ではないベクトルの各成分と実際の第１の０ではないベクトルの各成分との差の絶対値の合計として、誤差を計算することができる。

【０１２３】

[0154]誤差が得られると、コンテンツ分析ユニット２６は、実際の第１の０ではないベクトルのエネルギーおよび誤差に基づいて、比を計算することができる。コンテンツ分析ユニット２６は、第１の０ではないベクトルの各成分を２乗して、２乗された成分を互いに加算することによって、このエネルギーを決定することができる。コンテンツ分析ユニット２６は次いで、この比を閾値と比較することができる。比が閾値を超えないとき、コンテンツ分析ユニット２６は、フレーム分割されたＨＯＡ係数１１が録音から生成されたと決定し、ＨＯＡ係数１１の対応するコーディングされた表現が録音から生成されたことをビットストリームにおいて示すことができる。比が閾値を超えるとき、コンテンツ分析ユニット２６は、フレーム分割されたＨＯＡ係数１１が合成オーディオオブジェクトから

生成されたと決定し、フレーム分割されたＨＯＡ係数１１の対応するコーディングされた表現が合成オーディオオブジェクトから生成されたことをビットストリームにおいて示すことができる。

【０１２４】

[0155]フレーム分割されたＨＯＡ係数１１が録音から生成されたかまたは合成オーディオオブジェクトから生成されたかのインジケーションは、各フレームに対する単一のビットを備え得る。この単一のビットは、対応するフレームを符号化する際の様々な方法を実質的に切り替える各フレームのために、様々な符号化が使用されたことを示し得る。いくつかの例では、フレーム分割されたＨＯＡ係数１１が録音から生成されたとき、コンテンツ分析ユニット２６は、ＨＯＡ係数１１をベクトルベースの合成ユニット２７に渡す。いくつかの例では、フレーム分割されたＨＯＡ係数１１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２６は、ＨＯＡ係数１１を指向性ベースの合成ユニット２８に渡す。指向性ベースの合成ユニット２８は、指向性ベースのビットストリーム２１を生成するためにＨＯＡ係数１１の指向性ベースの合成を実行するように構成されるユニットを表し得る。

【０１２５】

[0156]言い換えれば、本技法は、フロントエンドの分類器を使用してＨＯＡ係数をコーディングすることに基づく。分類器は、次のように作動することができる：

（フレーム分割されたＨＯＡ係数またはＨＯＡ係数とも呼ばれ得る、たとえば４次の、１０２４というフレームサイズの）フレーム分割されたＳＨ行列で開始する－ここで２５×１０２４というサイズの行列が得られる。

第１のベクトル（０次のＳＨ）を除外する－したがって２４×１０２４というサイズの行列がある。

行列中のベクトルの残り（１×１０２４というサイズの２３個のベクトル）から行列中の第１の０ではないベクトル（１×１０２４というサイズのベクトル）を予測する。

予測は次の通りである：予測ベクトル＝ｓｕｍ－ｏｖｅｒ－ｉ［ａｌｐｈａ－ｉ×ｖｅｃｔｏｒ－Ｉ］（ここでＩにわたる合計は、ｉ＝１．．．２３という２３個のインデックスにわたって行われる）。

次いで誤差を確認する：実際の誤差－予測される誤差＝誤差。

ベクトル／誤差のエネルギーの比が大きい（すなわち、誤差が小さい）場合、背後の音場（そのフレームにおける）は希薄／合成的である。それ以外の場合、背後の音場は（たとえばマイクアレイを使用して）録音された音場である。

録音されたものか、合成かの決定に応じて、異なる方法で符号化／復号（帯域幅圧縮を指し得る）を実行する。この決定は、各フレームに対してビットストリームを通じて送信される１ビットの決定である。

【０１２６】

[0157]図４の例に示されるように、ベクトルベースの合成ユニット２７は、線形可逆変換（ＬＩＴ）ユニット３０と、パラメータ計算ユニット３２と、再順序付けユニット３４と、フォアグラウンド選択ユニット３６と、エネルギー補償ユニット３８と、音響心理学的オーディオコーダユニット４０と、ビットストリーム生成ユニット４２と、音場分析ユニット４４と、係数低減ユニット４６と、バックグラウンド（ＢＧ）選択ユニット４８と、空間－時間的補間ユニット５０と、量子化ユニット５２とを含み得る。

【０１２７】

[0158]線形可逆変換（ＬＩＴ）ユニット３０は、ＨＯＡチャネルの形態でＨＯＡ係数１１を受信し、各チャネルは、球面基底関数の所与の次数、位数と関連付けられる係数（ＨＯＡ［ｋ］と示され得る、ここでｋは現在のフレームまたはサンプルのブロックを示し得る）のブロックまたはフレームを表す。ＨＯＡ係数１１の行列は、次元Ｄ：Ｍ×（Ｎ＋１）^{2}、を有し得る。

【０１２８】

[0159]すなわち、ＬＩＴユニット３０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るユニットを表し得る。ＳＶＤ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、別段特別に述べられない限り０ではない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。

【０１２９】

[0160]代替的な変換は主成分分析を備えてよく、これは「ＰＣＡ」と呼ばれることが多い。ＰＣＡ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、直交変換を利用する数学的プロシージャを指す。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）をもたない変数を表す。これらの主成分は、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、この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大の分散を有するような方法で、定義される。ＰＣＡは、ＨＯＡ係数１１に関してＨＯＡ係数１１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、ＰＣＡ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（ＰＯＤ）、および固有値分解（ＥＶＤ）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、多チャネルオーディオデータの「エネルギー圧縮」および「脱相関」である。

【０１３０】

[0161]いずれにしても、ＬＩＴユニット３０は、ＨＯＡ係数１１を変換されたＨＯＡ係数の２つ以上のセットに変換するために、特異値分解（やはり「ＳＶＤ」と呼ばれ得る）を実行する。変換されたＨＯＡ係数のこれらの「セット」は、変換されたＨＯＡ係数のベクトルを含み得る。図４の例では、ＬＩＴユニット３０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、ＨＯＡ係数１１に関してＳＶＤを実行することができる。ＳＶＤは、線形代数学では、ｙ×ｚの実行列または複素行列Ｘ（ここで、Ｘは、ＨＯＡ係数１１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すことができる。

$$X = USV^*$$

Ｕは、ｙ×ｙの実ユニタリ行列または複素ユニタリ行列を表すことができ、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一般に知られる。Ｓは、対角線上に非負実数をもつｙ×ｚの矩形対角行列を表すことができ、ここで、Ｓの対角線上の値は、マルチチャネルオーディオデータの特異値として一般に知られる。V^*（Ｖの共役転置行列を示し得る）はｚ×ｚの実ユニタリ行列または複素ユニタリ行列を表すことができ、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一般に知られる。

【０１３１】

[0162]本開示では、ＨＯＡ係数１１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２０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。

【０１３２】

[0163]いくつかの例では、上で参照されたＳＶＤ数式中のＶ*行列は、複素数を備える行列にＳＶＤが適用され得ることを反映するために、Ｖ行列の共役転置行列として示される。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見なされてよい。以下では、説明を簡単にするために、ＨＯＡ係数１１が実数を備え、その結果、Ｖ*行列ではなくＶ行列がＳＶＤ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ＨＯＡ係数１１に適用されてよく、ここで、ＳＶＤ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ＳＶＤ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ＨＯＡ係数１１へのＳＶＤの適用を含んでよい。

【０１３３】

[0164]いずれにしても、ＬＩＴユニット３０は、高次アンビソニックス（ＨＯＡ）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ＳＶＤを実行することができる（ここで、このアンビソニックスオーディオデータは、ＨＯＡ係数１１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）。上で述べられたように、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１０２４個のオーディオサンプルを含むとき、Ｍは１０２４に等しい。Ｍのこの典型的な値に関して説明されるが、本開示の技法は、Ｍのこの典型的な値に限定されるべきではない。ＬＩＴユニット３０はしたがって、Ｍ×（Ｎ＋１）^{2}のＨＯＡ係数を有するブロックＨＯＡ係数１１に関してブロックごとのＳＶＤを実行することができ、ここでＮはやはりＨＯＡオーディオデータの次数を示す。ＬＩＴユニット３０は、このＳＶＤの実行を通じて、Ｖマトリックスと、Ｓマトリックスと、Ｕマトリックスとを生成することができ、マトリックスの各々は、上で説明されたそれぞれのＶマトリックスと、Ｓマトリックスと、Ｕマトリックス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３０は、次元Ｄ：Ｍ×（Ｎ＋１）^{2}を有するＵＳ［ｋ］ベクトル３３（ＳベクトルとＵベクトルの組み合わされたバージョンを表し得る）と、次元Ｄ：（Ｎ＋１）^{2}×（Ｎ＋１）^{2}を有するＶ［ｋ］ベクトル３５とを出力するために、ＨＯＡ係数１１に関してＳＶＤを実行することができる。ＵＳ［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。

【０１３４】

[0165]Ｕ行列、Ｓ行列、およびＶ行列の分析は、これらの行列が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の中の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的な特性（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すことができる。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。Ｕ行列中のベクトルとＶ行列中のベクトルの両方が、それらの二乗平均平方根のエネルギーが１に等しくなるように正規化される。したがって、Ｕの中のオーディオ信号のエネルギーは、Ｓの中の対角線上の要素によって表される。ＵとＳを乗算してＵＳ［ｋ］（個々のベクトル要素X_{ps}（ｋ）を有する）を形成することで、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）それらの空間的特性とを切り離すＳＶＤ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ＨＯＡ［ｋ］係数ＸをＵＳ［ｋ］とＶ［ｋ］のベクトル乗算によって合成するこのモデルは、本文書全体で使用される、「ベクトルベースの合成方法」という用語を生じさせる。

【０１３５】

[0166]ＨＯＡ係数１１に関して直接実行されるものとして説明されるが、ＬＩＴユニット３０は、線形可逆変換をＨＯＡ係数１１の派生物に適用することができる。たとえば、ＬＩＴユニット３０は、ＨＯＡ係数１１から導出された電力スペクトル密度行列に関してＳＶＤを適用することができる。電力スペクトル密度行列は、ＰＳＤ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、ＨＯＡ係数１１のフレームを指す。

【０１３６】

[0167]ＬＩＴユニット３０は、ＳＶＤ（ｓｖｄ）をＰＳＤに適用した後で、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、すなわちＬＩＴ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。ＬＩＴユニット３０は、いくつかの例では、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。ＬＩＴユニット３０は、ＳＶ［ｋ］’行列を得るために、Ｓ［ｋ］行列を量子化されたＶ［ｋ］’行列とまず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。ＬＩＴユニット３０は次に、ＳＶ［ｋ］’行列の擬似逆行列（ｐｉｎｖ）を取得することができ、次いで、Ｕ［ｋ］行列を取得するためにＳＶ［ｋ］’行列の擬似逆行列とＨＯＡ係数１１を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');
```

【０１３７】

[0168]ＨＯＡ係数自体ではなくＨＯＡ係数の電力スペクトル密度（ＰＳＤ）に関してＳＶＤを実行することによって、ＬＩＴユニット３０は場合によっては、プロセッササイクルと記憶空間の１つまたは複数に関してＳＶＤを実行することの計算的な複雑さを低減しつつ、ＳＶＤがＨＯＡ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の符号化効率を達成することができる。すなわち、上で説明されたＰＳＤタイプのＳＶＤは場合によっては、ＳＶＤがＦ＊Ｆの行列（ＦはＨＯＡ係数の数）に対して行われるので、計算的により負荷が軽いことがある。Ｍはフレーム長であるを伴うＭ＊Ｆの行列と比較されると、すなわち１０２４個以上のサンプル。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＊Ｌ＾２）と比べて、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の表記である）。

【０１３８】

[0169]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るユニットを表す。現在のフレームに対するこれらのパラメータの各々は、Ｒ［ｋ］、θ［ｋ］、φ［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、これらのパラメータを識別するために、ＵＳ［ｋ］ベクトル３３に関してエネルギー分析および／または相関付け（またはいわゆる相互相関付け）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、以前のフレームに対してこれら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］、θ［ｋ－１］、φ［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とを再順序付けユニット３４に出力することができる。

【０１３９】

[0170]すなわち、パラメータ計算ユニット３２は、第１の時間に対応するＬ個の第１のＵＳ［ｋ］ベクトル３３の各々および第２の時間に対応する第２のＵＳ［ｋ－１］ベクトル３３の各々に関してエネルギー分析を実行し、第１のオーディオフレームの少なくとも一部分（しかししばしば全体）および第２のオーディオフレームの一部分（しかししばしば全体）の二乗平均平方根エネルギーを計算し、これによって、第１のオーディオフレームのＬ個の第１のＵＳ［ｋ］ベクトル３３の各々に対して１つ、および第２のオーディオフレームの第２のＵＳ［ｋ－１］ベクトル３３の各々に対して１つの、２Ｌ個のエネルギーを生成することができる。

【０１４０】

[0171]他の例では、パラメータ計算ユニット３２は、第１のＵＳ［ｋ］ベクトル３３の各々および第２のＵＳ［ｋ－１］ベクトル３３の各々のサンプルのセット（そうでなければ全体）のある部分の間の相互相関付けを実行することができる。相互相関付けは、信号処理技術において理解されるような相互相関付けを指し得る。言い換えれば、相互相関付けは、２つの波形（この場合はＭ個のサンプルの個別のセットとして定義される）の１つに加えられる時間遅れの関数としての、２つの波形の間の類似性の尺度を指し得る。いくつかの例では、相互相関付けを実行するために、パラメータ計算ユニット３２は、第１のＵＳ［ｋ］ベクトル２７の各々の最後のＬ個のサンプルを、第２のＵＳ［ｋ－１］ベクトル３３の残りの各々の最初のＬ個のサンプルと順番ごとに比較して、相関パラメータを決定する。本明細書で使用される場合、「順番方向（turn-wise）」の演算は、要素の第１のセットおよび要素の第２のセットに関して行われる要素対要素の演算を指し、ここで、この演算は、セットの順序に従って「順番に」要素の第１のセットおよび第２のセットの各々から１つの要素を引き抜く。

【０１４１】

[0172]パラメータ計算ユニット３２はまた、指向性特性パラメータを決定するために、Ｖ［ｋ］および／またはＶ［ｋ－１］ベクトル３５を分析することができる。これらの指向性特性パラメータは、対応するＵＳ［ｋ］および／またはＵＳ［ｋ－１］ベクトル３３によって表されるオーディオオブジェクトの動きおよびロケーションのインジケーションを提供することができる。パラメータ計算ユニット３２は、前述の現在のパラメータ３７（ＵＳ［ｋ］ベクトル３３および／またはＶ［ｋ］ベクトル３５に関して決定された）の任意の組合せと、以前のパラメータ３９（ＵＳ［ｋ－１］ベクトル３３および／またはＶ［ｋ－１］ベクトル３５に関して決定された）の任意の組合せとを、再順序付けユニット３４に提供することができる。

【０１４２】

[0173]ＳＶＤ分解は、ＵＳ［ｋ－１］［ｐ］ベクトル（または代替的にはＸｐｓ$^{(p)}$（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、やはりＵＳ［ｋ］［ｐ］ベクトル３３（または代替的にはＸｐｓ$^{(p)}$（ｋ））によって示され得る、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もたらすようにオーディオオブジェクトを再順序付けるために、再順序付けユニット３４によって使用され得る。

【０１４３】

[0174]すなわち、再順序付けユニット３４は次いで、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。再順序付けユニット３４は、再順序付けられたＵＳ［ｋ］行列３３’（数学的には

【０１４４】

【数１４】

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

【０１４５】
として示され得る）と、再順序付けられたＶ［ｋ］行列３５’（数学的には
【０１４６】
【数１５】

$$\overline{\mathrm{V}}[k]$$

【０１４７】
として示され得る）とをフォアグラウンド音声（または支配的音声－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）再順序付けることができる。
【０１４８】
[0175]言い換えれば、再順序付けユニット３４は、再順序付けられたＵＳ［ｋ］行列３３’を生成するために、ＵＳ［ｋ］行列３３内のベクトルを再順序付けるように構成されるユニットを表し得る。再順序付けユニット３４は、ＵＳ［ｋ］ベクトル３３（ここでやはり、$Xps^{(p)}(k)$としてやはり代替的に示され得るＵＳ［ｋ］ベクトル３３の各ベクトルは、１つまたは複数の明瞭な（または言い換えれば支配的な）音場の中に存在するモノラルオーディオオブジェクトを表し得る）がオーディオデータの部分から変化し得るので、ＵＳ［ｋ］行列３３を再順序付けることができる。すなわち、いくつかの例においてオーディオ符号化デバイス１２がオーディオフレームと一般に呼ばれるオーディオデータのこれらの部分に対して動作するとすると、導出されるＵＳ［ｋ］行列３３において表されるようなこれらの明瞭なモノラルオーディオオブジェクトに対応するベクトルの位置は、フレームへのＳＶＤの適用およびフレームツーフレームからの各オーディオオブジェクトの変化する特徴が原因で、オーディオフレームごとに変化し得る。
【０１４９】
[0176]オーディオフレームごとにＵＳ［ｋ］行列３３内のベクトルを再順序付けることなく音響心理学的オーディオコーダユニット４０に直接ＵＳ［ｋ］行列３３内のベクトルを渡すことは、モノラルオーディオオブジェクトがオーディオフレームにまたがって連続的である（チャネル的である、この例では互いに対するＵＳ［ｋ］行列３３内のベクトルの位置的な順序によって定義される）ときにより良好に機能する従来の圧縮スキームなどのいくつかの圧縮スキームについて、達成可能な圧縮の程度を下げることがある。その上、再順序付けられないとき、ＵＳ［ｋ］行列３３内のベクトルの符号化は、復号されるときにオーディオデータの品質を下げることがある。たとえば、音響心理学的オーディオコーダユニット４０によって図３の例において表され得るＡＡＣ符号化器は、フレームごとにＵＳ［ｋ］行列３３内のベクトルを直接符号化するときに達成される圧縮と比較して、フレームごとにＵＳ［ｋ］行列３３’内の再順序付けられた１つまたは複数のベクトルをより効率的に圧縮することができる。ＡＡＣ符号化器に関して上で説明されるが、本技法は、モノラルオーディオオブジェクトが複数のフレームにわたって特定の順序または位置に指定される（チャネル的である）ときにより良好な圧縮をもたらす任意の符号化器に関して実行され得る。
【０１５０】
[0177]本技法の様々な態様は、このようにして、オーディオ符号化デバイス１２が、１つまたは複数のベクトルを再順序付ける（たとえば、ＵＳ［ｋ］行列３３内のベクトルを再順序付けて、再順序付けられたＵＳ［ｋ］行列３３’内の再順序付けられた１つまたは

複数のベクトルを生成し、これによって、音響心理学的オーディオコーダユニット４０などの従来のオーディオ符号化器によるＵＳ［ｋ］行列３３内のベクトルの圧縮を容易にする）ことを可能にし得る。

【０１５１】

[0178]たとえば、再順序付けユニット３４は、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－１］行列３３内の１つまたは複数の第２のベクトルが対応する第２のフレームよりも時間的に後の、第１のオーディオフレームからのＵＳ［ｋ］行列３３内の１つまたは複数のベクトルを再順序付けることができる。第１のオーディオフレームが第２のオーディオフレームよりも時間的に後にあるという文脈で説明されるが、第１のオーディオフレームは第２のオーディオフレームよりも時間的に先行してよい。したがって、本技法は、本開示で説明される例に限定されるべきではない。

【０１５２】

[0179]説明するために、ＵＳ［ｋ］行列３３内のｐ個のベクトルの各々がＵＳ［ｋ］［ｐ］として示され、対応するベクトルがｋ番目のフレームからのものか以前の（ｋ－１）番目のフレームからのものかをｋが示し、ｐが同じオーディオフレームのベクトルに対するベクトルの行を示す（ここでＵＳ［ｋ］行列は（Ｎ＋１）^{2}個のそのようなベクトルを有する）、以下の表１を考える。上で述べられたように、Ｎは１であると決定され、ｐはベクトル１から４を示し得ると仮定する。

【０１５３】

【表１】

表1

検討対象のエネルギー	比較対象
US[*k*-1][1]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][2]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][3]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][1], US[*k*][2], US[*k*][3], US[*k*][4]

【０１５４】

[0180]上の表１において、再順序付けユニット３４は、ＵＳ［ｋ－１］［１］に対して計算されたエネルギーをＵＳ［ｋ］［１］、ＵＳ［ｋ］［２］、ＵＳ［ｋ］［３］、ＵＳ［ｋ］［４］の各々に対して計算されるエネルギーと比較し、ＵＳ［ｋ－１］［２］に対して計算されたエネルギーをＵＳ［ｋ］［１］、ＵＳ［ｋ］［２］、ＵＳ［ｋ］［３］、ＵＳ［ｋ］［４］の各々に対して計算されるエネルギーと比較し、以下同様である。再順序付けユニット３４は次いで、第２の（時間的に）先行するオーディオフレームの第２のＵＳ［ｋ－１］ベクトル３３の１つまたは複数を廃棄することができる。説明のために、残りの第２のＵＳ［ｋ－１］ベクトル３３を示す以下の表２を考える。

【０１５５】

【表２】

表2

検討対象のベクトル	検討対象として残存
US[*k*-1][1]	US[*k*][1], US[*k*][2]
US[*k*-1][2]	US[*k*][1], US[*k*][2]
US[*k*-1][3]	US[*k*][3], US[*k*][4]
US[*k*-1][4]	US[*k*][3], US[*k*][4]

【０１５６】

[0181]上の表２において、再順序付けユニット３４は、ＵＳ［ｋ－１］［１］に対して計算されるエネルギーがＵＳ［ｋ］［１］およびＵＳ［ｋ］［２］の各々に対して計算されるエネルギーと類似している、ＵＳ［ｋ－１］［２］に対して計算されるエネルギーがＵＳ［ｋ］［１］およびＵＳ［ｋ］［２］の各々に対して計算されるエネルギーと類似している、ＵＳ［ｋ－１］［３］に対して計算されるエネルギーがＵＳ［ｋ］［３］およびＵＳ［ｋ］［４］の各々に対して計算されるエネルギーと類似している、およびＵＳ［ｋ－１］［４］に対して計算されるエネルギーがＵＳ［ｋ］［３］およびＵＳ［ｋ］［４］の各々に対して計算されるエネルギーと類似しているということを、エネルギー比較に基づいて決定することができる。いくつかの例では、再順序付けユニット３４は、ＵＳ［ｋ］行列３３の第１のベクトルの各々とＵＳ［ｋ－１］行列３３の第２のベクトルの各々との間の類似性を識別するために、さらなるエネルギー分析を実行することができる。

【０１５７】

[0182]他の例では、再順序付けユニット３２は、相互相関付けに関する現在のパラメータ３７および以前のパラメータ３９に基づいてベクトルを再順序付けることができる。これらの例では、上の表２に戻ると、再順序付けユニット３４は、これらの相互相関パラメータに基づいて表３において表される以下の例示的な相関付けを決定することができる。

【０１５８】

【表３】

表3

検討対象のベクトル	相関対象
US[*k*-1][1]	US[*k*][2]
US[*k*-1][2]	US[*k*][1]
US[*k*-1][3]	US[*k*][3]
US[*k*-1][4]	US[*k*][4]

【０１５９】

[0183]上の表３から、再順序付けユニット３４は、一例として、ＵＳ［ｋ－１］［１］ベクトルが異なるように配置されたＵＳ［ｋ］［２］ベクトルに相関する、ＵＳ［ｋ－１］［２］ベクトルが異なるように配置されたＵＳ［ｋ］［１］ベクトルに相関する、ＵＳ［ｋ－１］［３］ベクトルが同様に配置されたＵＳ［ｋ］［３］ベクトルに相関する、およびＵＳ［ｋ－１］［４］ベクトルが同様に配置されたＵＳ［ｋ］［４］ベクトルに相関

すると決定する。言い換えれば、再順序付けユニット３４は、ＵＳ［ｋ］［２］ベクトルがＵＳ［ｋ］行列３３の第１のベクトルの第１の行に再配置され、ＵＳ［ｋ］［１］ベクトルが第１のＵＳ［ｋ］ベクトル３３の第２の行に再配置されるように、ＵＳ［ｋ］行列３３の第１のベクトルをどのように再順序付けるべきかを記述する再順序付け情報と呼ばれ得るものを決定する。再順序付けユニット３４は次いで、再順序付けられたＵＳ［ｋ］行列３３’を生成するために、この再順序付け情報に基づいてＵＳ［ｋ］行列３３の第１のベクトルを再順序付けることができる。

【０１６０】

[0184]加えて、再順序付けユニット３４は、図４の例に示されていないが、この再順序付け情報をビットストリーム生成デバイス４２に提供することができ、ビットストリーム生成デバイス４２は、図３および図５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ＵＳ［ｋ］行列３３のベクトルを復元するためにＵＳ［ｋ］行列３３’の再順序付けられたベクトルをどのように再順序付けるべきかを決定できるように、この再順序付け情報を含むようにビットストリーム２１を生成することができる。

【０１６１】

[0185]まずエネルギー固有のパラメータに、次いで相互相関パラメータに基づく分析を伴う２段階のプロセスを実行するものとして上で説明されるが、再順序付けユニット３２は、再順序付け情報を決定するためにエネルギーパラメータに関してのみこの分析を実行するだけでよく、再順序付け情報を決定するために相互相関パラメータに関してのみこの分析を実行してよく、または、上で説明された方式でエネルギーパラメータと相互相関パラメータの両方に関して分析を実行してよい。加えて、本技法は、エネルギー比較および／または相互相関の１つまたは両方を実行することを伴わない、相関付けを実行するための他のタイプのプロセスを利用することができる。したがって、本技法は、この点で、上に記載されている例に限定されるべきではない。その上、パラメータ計算ユニット３２から得られる他のパラメータ（Ｖベクトルから、またはＶ［ｋ］およびＶ［ｋ－１］の中のベクトルの相関から導出される空間位置パラメータなどの）も、ＵＳ中のベクトルの正しい順序を決定するために、ＵＳ［ｋ］およびＵＳ［ｋ－１］から得られるエネルギーパラメータおよび相互相関パラメータとともに（同時に／一緒に、またはシーケンシャルに）使用され得る。

【０１６２】

[0186]Ｖ行列中のベクトルの相関を使用することの一例として、パラメータ計算ユニット３４は、Ｖ［ｋ］行列３５のベクトルが以下の表４において指定されるように相関付けられると決定することができる。

【０１６３】

【表４】

表４

検討対象のベクトル	相関対象
V[*k*-1][1]	V[*k*][2]
V[*k*-1][2]	V[*k*][1]
V[*k*-1][3]	V[*k*][3]
V[*k*-1][4]	V[*k*][4]

【０１６４】

上の表４から、再順序付けユニット３４は、一例として、Ｖ［ｋ－１］［１］ベクトルが異なるように配置されたＶ［ｋ］［２］ベクトルに相関する、Ｖ［ｋ－１］［２］ベクトルが異なるように配置されたＶ［ｋ］［１］ベクトルに相関する、Ｖ［ｋ－１］［３］ベ

クトルが同様に配置されたＶ［ｋ］［３］ベクトルに相関する、およびＶ［ｋ－１］［４］ベクトルが同様に配置されたＶ［ｋ］［４］ベクトルに相関すると決定する。再順序付けユニット３４は、再順序付けられたＶ［ｋ］行列３５’として、Ｖ［ｋ］行列３５のベクトルの再順序付けられたバージョンを出力することができる。

【０１６５】

[0187]いくつかの例では、ＵＳ行列中のベクトルに適用される同じ再順序付けは、Ｖ行列中のベクトルにも適用される。言い換えれば、Ｖベクトルを再順序付ける際に使用されるあらゆる分析が、ＵＳベクトルを再順序付けるために使用されるあらゆる分析とともに使用され得る。再順序付け情報が単にＵＳ［ｋ］ベクトル３５に関するエネルギーパラメータおよび／または相互相関パラメータに関して決定されない例を説明するために、再順序付けユニット３４はまた、Ｖ［ｋ］ベクトル３５に関して上で説明されたのと同様の方式で、相互相関パラメータおよびエネルギーパラメータに基づいてＶ［ｋ］ベクトル３５に関してこの分析を実行することができる。その上、ＵＳ［ｋ］ベクトル３３は何ら指向性特性を有しないが、Ｖ［ｋ］ベクトル３５は対応するＵＳ［ｋ］ベクトル３３の指向性に関する情報を提供することができる。この意味で、再順序付けユニット３４は、対応する指向性特性パラメータの分析に基づいて、Ｖ［ｋ］ベクトル３５とＶ［ｋ－１］ベクトル３５との相関を識別することができる。すなわち、いくつかの例では、オーディオオブジェクトは、動いているときに連続的な方式で音場の中を動き、または比較的安定したロケーションにとどまる。したがって、再順序付けユニット３４は、Ｖ［ｋ］行列３５およびＶ［ｋ－１］行列３５のベクトルのうちで、何らかの既知の物理的に現実的な動きを示す、または音場の中で静止したままであるものを、相関付けられるものとして識別し、ＵＳ［ｋ］ベクトル３３とＶ［ｋ］ベクトル３５とをこの指向性特性の相関に基づいて再順序付けることができる。いずれにしても、再順序付けユニット３４は、再順序付けられたＵＳ［ｋ］ベクトル３３’と再順序付けられたＶ［ｋ］ベクトル３５’とをフォアグラウンド選択ユニット３６に出力することができる。

【０１６６】

[0188]加えて、本技法は、エネルギー比較および／または相互相関の１つまたは両方を実行することを伴わない、正しい順序を決定するための他のタイプのプロセスを利用することができる。したがって、本技法は、この点で、上に記載されている例に限定されるべきではない。

【０１６７】

[0189]ＵＳ行列のベクトルの再順序付けを反映するためにＶ行列のベクトルを再順序付けるものとして上で説明されたが、いくつかの例では、ＶベクトルはＵＳベクトルとは異なるように再順序付けられてよく、ここで、別個のシンタックス要素がＵＳベクトルの再順序付けとＶベクトルの再順序付けとを示すために生成され得る。いくつかの例では、Ｖベクトルが音響心理学的に符号化され得ないとすると、Ｖベクトルは再順序付けられなくてよく、ＵＳベクトルのみが再順序付けられてよい。

【０１６８】

[0190]Ｖ行列のベクトルおよびＵＳ行列のベクトルの再順序付けが異なる実施形態は、オーディオオブジェクトを空間中で交換する意図があるとき、すなわち、オーディオオブジェクトを元の録音された位置から動かす（背後の音場が自然な録音であったとき）、または人工的に意図される位置から動かす（背後の音場がオブジェクトの人工的なミキシングであるとき）ときである。例として、２つのオーディオソースＡおよびＢがあり、Ａが音場の「左」部分から出る猫の「ニャー」という音であり、Ｂが音場の「右」部分から出る犬の「ウー」という音であり得るとする。ＶおよびＵＳの再順序付けが異なるとき、２つの音のソースの位置が交換される。交換の後で、Ａ（「ニャー」）は音場の右部分から出て、Ｂ（「ウー」）は音場の左部分から出る。

【０１６９】

[0191]音場分析ユニット４４は、目標ビットレート４１を達成可能にするために、ＨＯＡ係数１１に関して音場分析を実行するように構成されるユニットを表し得る。音場分析

ユニット４４は、この分析および／または受信された目標ビットレート４１に基づいて、音響心理学的なコーダの具体物の総数（これは環境またはバックグラウンドチャネルの総数（BG_{TOT}）とフォアグラウンドチャネルの数、または言い換えれば支配的なチャネルの数との関数であり得るを決定することができる。音響心理学的なコーダの具体物の総数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓとして示され得る。音場分析ユニット４４はまた、やはり目標ビットレート４１を達成可能にするために、フォアグラウンドチャネルの総数（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ネルの対応する数（ｎＢＧａ＝（ＭｉｎＡｍｂＨｏａＯｒｄｅｒ＋１）^{2}）と、送信すべき追加のＢＧ　ＨＯＡチャネルのインデックス（ｉ）（これは、図４の例ではバックグラウンドチャネル情報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ネル情報４２は、環境チャネル情報４３とも呼ばれ得る。ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓ－ｎＢＧａで残るチャネルの各々は、「追加のバックグラウンド／環境チャネル」、「アクティブなベクトルベースの支配的なチャネル」、「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。一実施形態では、これらのチャネルタイプは、２ビットによるシンタックス要素（「ＣｈａｎｎｅｌＴｙｐｅ」として）示され得る（たとえば、００：追加のバックグラウンドチャネル、０１：ベクトルベースの支配的な信号、１０：非アクティブな信号、１１：指向性ベースの信号）。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}＋（上の例の）インデックス００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ネルタイプとして現れる階数として、与えられ得る。

【０１７０】

[0192]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ネルおよび／またはフォアグラウンドチャネルを選択することができる。一実施形態では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓは８に設定され得るが、ＭｉｎＡｍｂＨｏａＯｒｄｅｒはビットストリームのヘッダセクションにおいて１に設定され得る（これは図１０～図１００（ｉｉ）に関してより詳細に説明される）。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つのチャネルが確保され得るが、他の４つのチャネルは、フレームごとに、チャ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／環境チャネルと、フォアグラウンド／支配的なチャ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、上で説明されたように、ベクトルベースの信号と指向性ベースの信号の１つであり得る。

【０１７１】

[0193]いくつかの例では、フレームに対するベクトルベースの支配的な信号の総数は、上の例では、そのフレームのビットストリームにおいてＣｈａｎｎｅｌＴｙｐｅインデックスが０１である回数によって与えられ得る。上の実施形態では、各々の追加のバックグラウンド／環境チャネル（たとえば、００というＣｈａｎｎｅｌＴｙｐｅ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ネルにおいて表され得るかの対応する情報。この情報は、４次のＨＯＡコンテンツでは、５と２５の間を示すためのインデックスであり得る（ｍｉｎＡｍｂＨｏａＯｒｄｅｒが１に設定されているときは最初の４つである１～４が常に送信され得るので、５と２５の間の１つを示すことだけが必要である）。この情報はしたがって、「ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ」として示され得る、５ビットのシンタックス要素（４次のコンテンツのための）を使用して送信され得る。

【０１７２】

[0194]第２の実施形態では、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の実施形態では、フォアグラウンド／支配的な信号の総数は、ｎＦＧ＝ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓ－［（ＭｉｎＡｍｂＨｏａＯｒｄｅｒ＋１）^{2}＋インデックス００の回数］として与えられ得る。

【０１７３】

[0195]音場分析ユニット４４は、バックグラウンドチャネル情報４３とＨＯＡ係数１１とをバックグラウンド（ＢＧ）選択ユニット４６に、バックグラウンドチャネル情報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。

【０１７４】

[0196]いくつかの例では、音場分析ユニット４４は、ＵＳ［ｋ］行列３３のベクトルおよび目標ビットレート４１の分析に基づいて、最大の値を有するこれらの成分の変数ｎＦＧの数を選択することができる。言い換えれば、音場分析ユニット４４は、Ｓ［ｋ］行列３３のベクトルの降順の対角線上の値によって作成される曲線の傾きを分析することによって、２つのサブ空間を分離する変数Ａの値（N_{BG}と同様である、または実質的に同様であり得る）を決定することができ、ここで、大きい特異値はフォアグラウンド音声または明瞭な音声を表し、小さい特異値は音場のバックグラウンド成分を表す。すなわち、変数Ａは、フォアグラウンドサブ空間およびバックグラウンドサブ空間へと全体の音場を区分することができる。

【０１７５】

[0197]いくつかの例では、音場分析ユニット４４は、特異値曲線の一次導関数と二次導関数とを使用することができる。音場分析ユニット４４はまた、１と５の間となるように変数Ａの値を制限することができる。別の例として、音場分析ユニット４４は、１と（Ｎ＋１）^{2}の間となるように変数Ａの値を制限することができる。代替的に、音場分析ユニット４４は、４という値などへと変数Ａの値を事前に定めることができる。いずれにしても、Ａの値に基づいて、音場分析ユニット４４は、フォアグラウンドチャネルの総数（ｎＦＧ）４５と、バックグラウンド音場の次数（N_{BG}）と、送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）およびインデックス（ｉ）とを決定する。

【０１７６】

[0198]さらに、音場分析ユニット４４は、ベクトルごとにＶ［ｋ］行列３５のエネルギーを決定することができる。音場分析ユニット４４は、Ｖ［ｋ］行列３５中のベクトルの各々のエネルギーを決定し、高いエネルギーを有するベクトルをフォアグラウンド成分として識別することができる。

【０１７７】

[0199]その上、音場分析ユニット４４は、空間エネルギー分析、空間マスキング分析、拡散分析、または他の形態の聴覚的分析を含む、様々な他の分析をＨＯＡ係数１１に関して実行することができる。音場分析ユニット４４は、空間領域へのＨＯＡ係数１１の変換を通じて空間エネルギー分析を実行することができ、保存されるべき音場の指向性成分を表す高エネルギーの領域を識別する。音場分析ユニット４４は、音場分析ユニット４４が空間的に近接しているより高エネルギーの音によってマスキングされる空間領域を識別できることを除き、空間エネルギー分析と同様の方式で知覚的空間マスキング分析を実行することができる。音場分析ユニット４４は次いで、知覚的にマスキングされた領域に基づいて、いくつかの例ではより少数のフォアグラウンド成分を識別することができる。音場分析ユニット４４はさらに、音場のバックグラウンド成分を表し得る拡散エネルギーの領域を識別するために、ＨＯＡ係数１１に関して拡散分析を実行することができる。

【０１７８】

[0200]音場分析ユニット４４はまた、オーディオデータと関連付けられる指向性ベースの情報を使用して、音場を表すオーディオデータの顕著さ、明瞭性、または優位性を決定するように構成されるユニットを表し得る。エネルギーベースの決定は、音場の明瞭なオーディオ成分を識別するためにＳＶＤによって分解された音場のレンダリングを改善し得

るが、バックグラウンドオーディオ成分が高いエネルギーレベルを示す場合は、エネルギーベースの決定はまた、デバイスに、明瞭なオーディオ成分としてバックグラウンドオーディオ成分を誤って識別させることがある。すなわち、エネルギーのみに基づいて明瞭なオーディオ成分とバックグラウンドオーディオ成分を分離することは、高エネルギーの（たとえば、より大きい）バックグラウンドオーディオ成分が明瞭なオーディオ成分として誤って識別されることがあるので、安定的ではないことがある。音場の明瞭なオーディオ成分とバックグラウンドオーディオ成分をより安定的に区別するために、本開示で説明される技法の様々な態様は、音場分析ユニット４４に、ＨＯＡ係数１１の分解されたバージョンからフォアグラウンドオーディオ成分と環境オーディオ成分とを分離するためにＨＯＡ係数１１の指向性ベースの分析を実行させることができる。

【０１７９】

[0201]この点において、音場分析ユニット４４は、ＵＳ［ｋ］行列３３中のベクトルおよびＶ［ｋ］行列３５中のベクトルの１つまたは複数に含まれるバックグラウンド要素から明瞭な（またはフォアグラウンド）要素を識別するように構成される、または別様に動作可能である、ユニットを表し得る。いくつかのＳＶＤベースの技法に従って、最も高エネルギーの成分（たとえば、ＵＳ［ｋ］行列３３およびＶ［ｋ］行列３５またはそれらから導出されるベクトルの１つまたは複数の最初のいくつかのベクトル）は、明瞭な成分として扱われ得る。しかしながら、ＵＳ［ｋ］行列３３中のベクトルおよびＶ［ｋ］行列３５中のベクトルの１つまたは複数の最も高エネルギーの成分（ベクトルによって表される）は、すべての状況において、最も指向性のある成分／信号を表すとは限らない。

【０１８０】

[0202]音場分析ユニット４４は、ＵＳ［ｋ］行列３３中のベクトルおよびＶ［ｋ］行列３５中のベクトルまたはそれらから導出されたベクトルの１つまたは複数のベクトルの指向性に基づいて、フォアグラウンド／直接の／支配的な要素を識別するために、本明細書で説明される技法の１つまたは複数の態様を実施することができる。いくつかの例では、音場分析ユニット４４は、ベクトルのエネルギーと指向性の両方に基づいて、１つまたは複数のベクトルを明瞭なオーディオ成分として識別または選択することができる（ここで、成分は「オブジェクト」とも呼ばれ得る）。たとえば、音場分析ユニット４４は、ＵＳ［ｋ］行列３３中のベクトルおよびＶ［ｋ］行列３５中のベクトル（またはそれらから導出されたベクトル）の１つまたは複数のベクトルのうち、高いエネルギーと（指向性指数として表される）高い指向性の両方を示すものを、明瞭なオーディオ成分として識別することができる。結果として、特定のベクトルがＵＳ［ｋ］行列３３中のベクトルおよびＶ［ｋ］行列３５中のベクトルの１つまたは複数の他のベクトルと比較されると比較的指向性が小さいと音場分析ユニット４４が決定する場合、特定のベクトルと関連付けられるエネルギーレベルとは無関係に、音場分析ユニット４４は、その特定のベクトルがＨＯＡ係数１１によって表される音場のバックグラウンド（または環境）オーディオ成分を表すと決定することができる。

【０１８１】

[0203]いくつかの例では、音場分析ユニット４４は、以下の演算を実行することによって、指向性に基づいて明瞭なオーディオオブジェクト（上で述べられたように、「成分」とも呼ばれ得る）を識別することができる。音場分析ユニット４４は、Ｓ［ｋ］行列中のベクトル（ＵＳ［ｋ］ベクトル３３から導出され得る、または図４の例には示されないがＬＩＴユニット３０によって別個に出力される）を、（たとえば、１つまたは複数の行列乗算処理を使用して）Ｖ［ｋ］行列３５中のベクトルと乗算することができる。Ｖ［ｋ］行列３５とＳ［ｋ］行列とを乗算することによって、音場分析ユニット４４はＶＳ［ｋ］行列を得ることができる。加えて、音場分析ユニット４４は、ＶＳ［ｋ］行列中のベクトルの各々の成分の少なくともいくつかを二乗する（すなわち、２という冪によって冪乗する）ことができる。いくつかの例では、音場分析ユニット４４は、１より大きい次数と関連付けられる各ベクトルの二乗された成分を加算することができる。

【０１８２】

[0204]一例として、ＶＳ［ｋ］行列の各ベクトルが２５個の成分を含む場合、音場分析ユニット４４は、各ベクトルに関して、５番目の成分から始まり２５番目の成分で終わるように各ベクトルの成分を二乗し、二乗された成分を加算して指向性指数（または指向性インジケータ）を決定することができる。各加算演算は、対応するベクトルの指向性指数をもたらし得る。この例では、音場分析ユニット４４は、１以下の次数と関連付けられる各行の成分、すなわち１番目から４番目の成分が、一般的にエネルギーの量をより対象としており、成分の指向性はあまり対象としていないと、決定することができる。すなわち、０または１という次数と関連付けられる低次アンビソニックスは、図１および図２に示されるように、圧力波の方向に関して多くのことを提供せずむしろ何らかの大きさ（これはエネルギーを表す）を提供する、球面基底関数に対応する。

【０１８３】

[0205]上の例で説明される演算はまた、以下の擬似コードに従って表され得る。以下の擬似コードは、（引用符を伴わない）文字列「/*」および「*/」の連続する実例の中に含まれるコメント記述の形式の、注釈を含む。

```
[U,S,V] = svd(audioframe,'ecom');
VS = V*S;
/*次の行は、各行を独立に分析することと、対応するベクトルの指向性指数または指向性尺度を決定するために５番目の成分から２５番目の成分まで（一例として）第１の行の中の値を加算することとを対象とする。加算の前に成分を二乗する。１より大きい次数と関連付けられる各行の中の成分は、高次アンビソニックスと関連付けられるので、指向性がある可能性がより高い。*/
sumVS = sum(VS(5:end,:).^2,1);
/*次の行は、生成されたＶＳ行列の二乗の合計をソートすることと、最大の値のセット（たとえば、最大の値の３つまたは４つ）を選択することとを対象とする*/
[~,idxVS] = sort(sumVS,'descend');
U = U(:,idxVS);
V = V(:,idxVS);
S = S(idxVS,idxVS);
```

【０１８４】

[0206]言い換えれば、上の擬似コードに従って、音場分析ユニット４４は、ＨＯＡ係数１１のうちで１より大きい次数を有する球面基底関数に対応するものから分解されたＶＳ［ｋ］行列の各ベクトルの成分を選択することができる。音場分析ユニット４４は次いで、ＶＳ［ｋ］行列の各ベクトルのこれらの成分を二乗し、二乗された成分を加算して、ＶＳ［ｋ］行列の各ベクトルの指向性尺度または指向性指数を識別し、計算し、または別様に決定することができる。次に、音場分析ユニット４４は、ベクトルの各々のそれぞれの指向性尺度に基づいて、ＶＳ［ｋ］行列のベクトルをソートすることができる。音場分析ユニット４４は、最高の対応する指向性と関連付けられるベクトルが最初になり、最低の対応する指向性と関連付けられるベクトルが最後となるように、指向性尺度の降順でこれらのベクトルをソートすることができる。音場分析ユニット４４は次いで、最高の相対的な指向性尺度を有するベクトルの０ではないサブセットを選択することができる。

【０１８５】

[0207]音場分析ユニット４４は、音響心理学的なコーダの具体物の総数（これは環境またはバックグラウンドチャネルの総数（$ＢＧ_{TOT}$）の関数であり得るとフォアグラウンドチャネルの数とを決定するために、前述の分析の任意の組合せを実行することができる。音場分析ユニット４４は、前述の分析の任意の組合せに基づいて、フォアグラウンドチャネルの総数（ｎＦＧ）４５と、バックグラウンド音場の次数（N_{BG}）と、送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）およびインデックス（ｉ）（図４の例ではバックグラウンドチャネル情報４３としてまとめて示され得る）とを決定することができる。

【０１８６】

[0208]いくつかの例では、音場分析ユニット４４はこの分析を各々のＭ個のサンプルに

ついて実行することができ、これはフレームごとに再表示され得る。この点で、Ａの値はフレームごとに変化し得る。決定が各々のＭ個のサンプルについて行われるビットストリームの例が、図１０～図１００（ｉｉ）に示される。他の例では、音場分析ユニット４４は、フレームごとに２回以上この分析を実行し、フレームの２つ以上の部分を分析することができる。したがって、本技法は、この点で、本開示で説明されている例に限定されるべきではない。

【０１８７】

[0209]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ネル情報（たとえば、バックグラウンド音場（N_{BG}）と、送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）およびインデックス（ｉ））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る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１つによって識別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、ここで、ｎＢＧａは、図３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ＢＧ　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ことになる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×［（$N_{BG}+1$）^{2}＋ｎＢＧａ］を有し得る。

【０１８８】

[0210]フォアグラウンド選択ユニット３６は、再順序付けられたＵＳ［ｋ］行列３３’および再順序付けられたＶ［ｋ］行列３５’のうちで音場のフォアグラウンド成分または明瞭な成分を表すものを、ｎＦＧ４５（これらのフォアグラウンドベクトルを識別する１つまたは複数のインデックスを表し得る）に基づいて選択するように構成される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、ｎＦＧ信号４９（再順序付けられたＵＳ［ｋ］$_{1,...,nFG}$４９、$FG_{1,...,nFG}$［ｋ］４９、または

【０１８９】

【数１６】

$$X_{PS}^{(1..nFG)}(k)\ 49$$

【０１９０】

として示され得る）を音響心理学的オーディオコーダユニット４０に出力することができ、ここで、ｎＦＧ信号４９は次元Ｄ：Ｍ×ｎＦＧ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再順序付けられたＶ［ｋ］行列３５’（または$v^{(1...nFG)}$（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、ここで、再順序付けられたＶ［ｋ］行列３５’のうちでフォアグラウンド成分に対応するものは、次元Ｄ：（Ｎ＋１）^{2}×ｎＦＧ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、以下の式として数学的に示され得る）。

【０１９１】

【数１７】

$$\overline{V}_{1,...,nFG}[k]$$

【０１９２】

[0211]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ＨＯＡチャネルの様々なチャ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ＨＯＡ

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るユニットを表し得る。エネルギー補償ユニット３８は、再順序付けられたＵＳ［ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および環境ＨＯＡ係数４７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ＨＯＡ係数４７’を生成するために、こ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ＨＯＡ係数４７’を音響心理学的オーディオコーダユニット４０に出力することができる。

【０１９３】

[0212]実質的に、エネルギー補償ユニット３８は、次数低減された環境ＨＯＡ係数４７（これは、いくつかの例では、［$(N_{BG}+1)^2+nBGa$］）という次数／位数を有する球面基底関数に対応する含まれる係数のみに関してＮ未満の次数を有する）を生成するためにＨＯＡ係数１１によって記述される音場の環境成分の次数を低減することによって引き起こされる、音場のバックグラウンド音声成分の全体的なエネルギーの起こり得る減少を補償するために使用され得る。いくつかの例では、エネルギー補償ユニット３８は、環境ＨＯＡ係数４７を音響心理学的オーディオコーダユニット４０に出力する前に、ＨＯＡ係数４７の二乗平均平方根（ＲＭＳ）エネルギーを（再順序付けられたＵＳ［ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数の総計的なエネルギー分析を通じて決定される）ＨＯＡ係数１１のＲＭＳに等しく、または少なくともそれにより近くなるように増大させるために、環境ＨＯＡ係数４７の［$(N_{BG}+1)^2+nBGa$］個の列の各々に適用すべき増幅値の形態で補償利得を決定することによって、このエネルギーの損失を補償する。

【０１９４】

[0213]いくつかの例では、エネルギー補償ユニット３８は、再順序付けられたＵＳ［ｋ］行列３３’と再順序付けられたＶ［ｋ］行列３５’の１つまたは複数の各行および／または各列のＲＭＳを識別することができる。エネルギー補償ユニット３８はまた、選択されたフォアグラウンドチャネルの１つまたは複数の各行および／または各列のＲＭＳを識別することができ、選択されたフォアグラウンドチャネルは、ｎＦＧ信号４９と、フォアグラウンドＶ［ｋ］ベクトル51_kと、次数低減された環境ＨＯＡ係数４７とを含み得る。再順序付けられたＵＳ［ｋ］行列３３’および再順序付けられたＶ［ｋ］行列３５’の１つまたは複数の各行および／または各列のＲＭＳは、RMS_{FULL}と示されるベクトルに記憶され得るが、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数の各行および／または各列のＲＭＳは、$RMS_{REDUCED}$と示されるベクトルに記憶され得る。エネルギー補償ユニット３８は次いで、$Z=RMS_{FULL}/RMS_{REDUCED}$という式に従って、増幅値ベクトルＺを計算することができる。エネルギー補償ユニット３８は次いで、この増幅値ベクトルＺまたはその様々な部分を、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数に適用することができる。いくつかの例では、増幅値ベクトルＺは、次の式$HOA_{BG-RED}'=HOA_{BG-RED}Z^T$によって、次数低減された環境ＨＯＡ係数４７のみに適用され、ここでHOA_{BG-RED}は次数低減された環境ＨＯＡ係数４７を示し、HOA_{BG-RED}'はエネルギー補償され低減された環境ＨＯＡ係数４７’を示し、Z^TはＺベクトルの転置を示す。

【０１９５】

[0214]いくつかの例では、再順序付けられたＵＳ［ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数のそれぞれの行および／または列の各ＲＭＳを決定するために、エネルギー補償ユニット３８はまず、基準球面調和係数（ＳＨＣ）レンダラを列に適用することができる。エネルギー補償ユニット３８による基準ＳＨＣレンダラの適用は、以下でより詳細に説明されるように、再順序付けられたＵＳ［

ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数の行および／または列によって表されるフレームの各行および／または各列によって記述される音場全体のエネルギーを決定するための、ＳＨＣ領域におけるＲＭＳの決定を可能にする。

【０１９６】

[0215]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るユニットを表し得る。空間－時間的補間ユニット５０は、再順序付けられたフォアグラウンドＨＯＡ係数を復元するために、ｎＦＧ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ｎＦＧ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、再順序付けられたフォアグラウンドＨＯＡ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、フォアグラウンドＶ［ｋ］ベクトル51_kのうちで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ものを出力することができる。フォアグラウンドＶ［ｋ］ベクトル51_kのうちで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ものは、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）使用されることを確実にするために、これらの量子化された／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。

【０１９７】

[0216]この点において、空間－時間的補間ユニット５０は、第１のオーディオフレームおよび第２の時間的に後続のまたは先行するオーディオフレームのいくつかの他の部分から第１のオーディオフレームの第１の部分を補間するユニットを表し得る。いくつかの例では、それらの部分はサブフレームとして示されることがあり、サブフレームに関して実行されるような補間は、図４５～図４６Ｅに関してより詳細に説明される。他の例では、空間－時間的補間ユニット５０は、図３７～図３９に関してより詳細に説明されるように、以前のフレームの最後のある数のサンプル、および後続のフレームの最初のある数のサンプルに関して動作することができる。空間－時間的補間ユニット５０は、この補間を実行する際に、ビットストリーム２１において指定されることが要求されるフォアグラウンドＶ［ｋ］ベクトル51_kのサンプルの数を減らすことができ、それは、フォアグラウンドＶ［ｋ］ベクトル51_kのうちで補間されたＶ［ｋ］ベクトルを生成するために使用されるもののみが、フォアグラウンドＶ［ｋ］ベクトル51_kのサブセットを表すからである。すなわち、（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Ｖ［ｋ］ベクトル51_kの数を減らすことによって）ＨＯＡ係数１１の圧縮を潜在的により効率的に行えるようにするために、本開示で説明される技法の様々な態様は、第１のオーディオフレームの１つまたは複数の部分の補間を提供することができ、それらの部分の各々はＨＯＡ係数１１の分解されたバージョンを表し得る。

【０１９８】

[0217]空間－時間的補間は、いくつかの利益をもたらし得る。まず、ＳＶＤまたは他のＬＩＴが実行される際のブロックごとの性質により、ｎＦＧ信号４９がフレームとフレームの間で連続的ではないことがある。言い換えれば、ＬＩＴユニット３０がフレームごとにＳＶＤを適用するとすると、たとえばＵＳ［ｋ］行列３３およびＶ［ｋ］行列３５の再順序付けられていない性質により証拠として、得られた変換されたＨＯＡ係数に何らかの不連続性が存在し得る。この補間を実行することによって、フレーム境界（または言い換

えれば、ＨＯＡ係数１１のフレームへの区分）が原因でもたらされるあらゆるアーティファクトを低減する可能性のある平滑化効果を補間が有し得るとすると、不連続性は低減され得る。この補間を実行するためにフォアグラウンドＶ［ｋ］ベクトル５１$_k$を使用し、次いで、補間されたフォアグラウンドＶ［ｋ］ベクトル５１$_k$に基づいて復元された再順序付けられたＨＯＡ係数から補間されたｎＦＧ信号４９’を生成することで、フレームごとの演算が原因の、さらにはｎＦＧ信号４９の再順序付けが原因の少なくともいくつかの影響を平滑化することができる。

【０１９９】

[0218]演算において、空間－時間的補間ユニット５０は、第１のフレームに含まれる第１の複数のＨＯＡ係数１１の一部分の第１の分解、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のフレームに含まれる第２の複数のＨＯＡ係数１１の一部分の第２の分解、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して、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ことができる。

【０２００】

[0219]いくつかの例では、第１の分解は、ＨＯＡ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の分解は、ＨＯＡ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。

【０２０１】

[0220]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（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この技法は、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、ＵＳ［ｋ］行列中の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、この方式で扱われると、これらのベクトル（ＵＳ［ｋ］行列中の）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の間で不連続である。これらの不連続性は、成分が変換オーディオコーダ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。

【０２０２】

[0221]本開示で説明される技法は、この不連続性に対処し得る。すなわち、本技法は、球面調和関数領域における直交する空間軸としてＶ行列が解釈され得るという考察に基づき得る。Ｕ［ｋ］行列は、基底関数によって球面調和関数（ＨＯＡ）データの射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したが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の同様の分解とは異なる。これらの点で、ＳＶＤは、ｍａｔｃｈｉｎｇ　ｐｕｒｓｕｉｔアルゴリズムであると見なされ得る。本開示で説明される技法は、空間－時間的補間ユニット５０が、フレームとフレームの間で基底関数（Ｖ［ｋ］）の連続性を、フレーム間を補間することによって保つことを可能にし得る。

【０２０３】

[0222]上で述べられたように、この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上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であり得る。

【０２０４】

【数１８】

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

【０２０５】
この上の式において、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、一実施形態では、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
【０２０６】
【数１９】

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

【０２０７】
が必要とされるサンプルの長さであり、このプロセスの出力がこれらのベクトルのｌを生成することも示す）。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「ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ」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の間で線形に、および単調に変化し得る。他の例では、ｗ（ｌ）は、非線形に、しかしｌの関数として非単調な（二乗余弦の４分の１周期などの）方式で０と１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、「ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄ」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（ｌ）が０に近い値であるとき、出力
【０２０８】
【数２０】

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

【０２０９】
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（ｌ）が１に近い値であるとき、そのことは、出力
【０２１０】
【数２１】

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

【０２１１】
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確実にする。
【０２１２】
[0223]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る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}－ｎＢＧａ］×ｎＦＧを有し得る。

【０２１３】

[0224]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の係数の数を減らすように構成される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、（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］ベクトルの係数のうちで指向性情報をほとんどまたはまったく有しないものを除去するように構成されるユニットを表し得る。上で説明されたように、いくつかの例では、明瞭な、または言い換えればフォアグラウンドＶ［ｋ］ベクトルの係数のうちで、１次または０次の基底関数に対応するもの（N_{BG}と示され得る）は、指向性情報をほとんど提供しないので、フォアグラウンドＶベクトルから（「係数低減」と呼ばれ得るプロセスを通じて）除去され得る。この例では、N_{BG}に対応するこれらの係数を識別するだけではなく、追加のＨＯＡチャネル（変数ＴｏｔａｌＯｆＡｄｄＡｍｂＨＯＡＣｈａｎによって示され得る）を［（N_{BG}＋１）^{2}＋１，（Ｎ＋１）^{2}］のセットから識別するために、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を識別することができ、これらはバックグラウンドチャネル情報４３とまとめて呼ばれ得る。係数低減ユニット４６は次いで、（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から除去して、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}））×ｎＦＧであるより低次元のＶ［ｋ］行列５５を生成することができる。

【０２１４】

[0225]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式の量子化を実行し、これらの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る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１つまたは複数を圧縮するように構成される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ＨＯＡ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、ｎは音場のＨＯＡ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式の量子化を実行することができる。

【０２１５】

[0226]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スキームを実行することができる。この圧縮スキームは、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スキーム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、ある例として、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の整数表現の一様量子化、および、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とコーディングの１つまたは複数を含む、圧縮スキームを実行することができる。

【０２１６】

[0227]いくつかの例では、この圧縮スキームの１つまたは複数のプロセスの様々なものが、一例として、得られるビットストリーム２１の目標ビットレートを達成するために、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と、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。

【０２１７】

[0228]いずれにしても、上で述べられたように、このコーディングスキームはまず、低減されたフォアグラウンドＶ［ｋ］ベクトル５５の各々の各要素の浮動小数点表現（これは、いくつかの例では３２ビットの浮動小数点数である）を１６ビットの整数表現に変換することを伴い得る。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の所与の１つの各要素を2^{15}と乗算することによって、この浮動小数点から整数への変換を実行することができ、これはいくつかの例では、１５だけ右シフトすることによって実行される。

【０２１８】

[0229]量子化ユニット５２は次いで、低減されたフォアグラウンドＶ［ｋ］ベクトル５５の所与の１つの要素のすべてに関して一様量子化を実行することができる。量子化ユニット５２は、ｎｂｉｔｓパラメータ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識別することができる。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ｎｂｉｔｓパラメータを動的に決定することができる。量子化ユニット５２は、このｎｂｉｔｓパラメータ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、$2^{16-nbits}$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、ｎｂｉｔｓが６に等しい場合、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／Δ］に等しく、$-2^{nbits-1} < v_q < 2^{nbits-1}$である。

【０２１９】

[0230]量子化ユニット５２は次いで、量子化されたベクトル素子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を（カテゴリ識別子ｃｉｄを決定することによって）、次の式を使用して識別することができる：

【０２２０】

【数２２】

$$cid = \begin{cases} 0, & v_q = 0 \text{ の場合} \\ \lfloor log_2|v_q| \rfloor + 1, & v_q \neq 0 \text{ の場合} \end{cases}$$

【０２２１】

量子化ユニット５２は次いで、このカテゴリインデックスｃｉｄをハフマンコーディングし、一方で、v_qが正の値から負の値かを示す符号ビットを識別することもできる。量子化ユニット５２は次に、このカテゴリにおける残差を識別することができる。一例として、量子化ユニット５２は、次の式にしたがって残差を決定することができる：

$$残差 = |v_q| - 2^{cid-1}$$

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。

【０２２２】

[0231]次の例は、この分類および残差コーディングのプロセスの簡略化された例を示す

。まず、ｎｂｉｔｓが６に等しいのでv_q∈［－３１，３１］であると仮定する。次に、以下のことを仮定する。

【０２２３】

【表５】

cid	vq	cidに対するハフマンコード
0	0	‘1’
1	-1, 1	‘01’
2	-3,-2, 2,3	‘000’
3	-7,-6,-5,-4, 4,5,6,7	‘0010’
4	-15,-14,…,-8, 8,…,14,15	‘00110’
5	-31,-30,…,-16, 16,…,30,31	‘00111’

【０２２４】

また、以下のことを仮定する。

【０２２５】

【表６】

cid	残差に対するブロックコード
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110…,0000, 1000,…,1110,1111
5	01111, … ,00000, 10000, … ,11111

【０２２６】
したがって、v_q＝［６，－１７，０，０，３］に対して、次のことが決定され得る：

```
>>  3,5,0,0,2
>>  符号 = 1,0,x,x,1
>>  残差 = 2,1,x,x,1
>>  ６に対するビット = ‘0010’ + ’1’ + ’10’
>>  －１７に対するビット = ‘00111’ + ’0’ + ‘0001’
>>  ０に対するビット = ‘0’
>>  ０に対するビット = ‘0’
>>  ３に対するビット = ‘000’ + ‘1’ + ‘1’
>>  総ビット = 7+10+1+1+5 = 24
>>  平均ビット = 24/5 = 4.8
```

【０２２７】
[0232]前述の簡略化された例には示されないが、量子化ユニット５２は、ｃｉｄをコーディングするとき、ｎｂｉｔｓ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、ｎｂｉｔｓ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ｎｂｉｔｓ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ｃｉｄ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。

【０２２８】
[0233]説明するために、量子化ユニット５２は、ｎｂｉｔｓ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１つが低減されたフォアグラウンドＶ［ｋ］ベクトル５５の対応する時間的に後続の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は定義されたもの）の空間情報を表さないときに、使用され得る。量子化ユニット５２は追加で、ｎｂｉｔｓ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１つが低減されたフォアグラウンドＶ［ｋ］ベクトル５５の対応する時間的に後続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のその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、ｎｂｉｔｓ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のその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、構築され得る。

【０２２９】
[0234]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。

【０２３０】

【表７】

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

【０２３１】
前述の表において、予測モード（「予測モード（Ｐｒｅｄ　ｍｏｄｅ）」）は、現在のベクトルに対して予測が実行されたかどうかを示し、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。
【０２３２】
[0235]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。
【０２３３】
【表８】

	録音	合成
予測なし	HT{1,2,3}	HT5
予測あり	HT4	HT5

【０２３４】
前述の表において、「録音」列は、録音されたオーディオオブジェクトをベクトルが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。「予測なし」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「予測あり」行は、予測がベクトル要素に関して実行されるとき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。

【０２３５】

[0236]この点において、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を圧縮するために上で述べられたスカラー量子化および／またはハフマン符号化を実行し、サイドチャネル情報５７と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。このサイドチャネル情報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。量子化ユニット５２は、図１０Ｂおよび図１０Ｃの例に示されるものと同様の方式で、サイドチャネル情報５７を出力することができる。

【０２３６】

[0237]上で述べられたように、量子化ユニット５２は、サイドチャネル情報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、量子化ユニット５２は、複数の構成モードのいずれが選択されたかを示す、アクセスユニット（１つまたは複数のフレームを含み得る）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、量子化ユニット５２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）この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、このシンタックス要素は、この明瞭な成分の指向性の様相を表すために、４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、「ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈ」として示され得る。このようにして、量子化ユニット５２は、４つの構成モードのいずれが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、または別様に指定することができる。４つの構成モードに関して説明されるが、本技法は４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構成モードを含む、任意の数の構成モードに限定されるべきである。スカラー／エントロピー量子化ユニット５３はまた、サイドチャネル情報５７の中の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。

【０２３７】

[0238]オーディオ符号化デバイス２０に含まれる音響心理学的オーディオコーダユニット４０は、音響心理学的オーディオコーダの複数の具体物を表すことができ、これらの各々は、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’の各々の様々なオーディオオブジェクトまたはＨＯＡチャネルを符号化して符号化された環境ＨＯＡ係数５９と符号化されたｎＦＧ信号６１とを生成するために使用される。音響心理学的オーディオコーダユニット４０は、符号化された環境ＨＯＡ係数５９と符号化されたｎＦＧ信号６１とをビットストリーム生成ユニット４２を出力することができる。

【０２３８】

[0239]いくつかの例では、この音響心理学的オーディオコーダユニット４０は、先進的オーディオコーディング（ＡＡＣ）符号化ユニットの１つまたは複数の具体物を表し得る。音響心理学的オーディオコーダユニット４０は、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’の各列または各行を符号化することができる。しばしば、音響心理学的オーディオコーダユニット４０は、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’に残存している次数／位数の各々に対して、ＡＡＣ符号化ユニットの具体物を呼び出すことができる。バックグラウンド球面調和係数３１がどのようにＡＡＣ符号化ユニットを使用して符号化され得るかに関するより多くの情報が、第１２４回会議、２００８年５月１７～２０日において発表され、http://ro.uow.

edu.au/cgi/viewcontent.cgi?article=8025&context=engpapersにおいて利用可能な、Ｅｒｉｃ Ｈｅｌｌｅｒｕｄ他、「Ｅｎｃｏｄｉｎｇ Ｈｉｇｈｅｒ Ｏｒｄｅｒ Ａｍｂｉｓｏｎｉｃｓ ｗｉｔｈ ＡＡＣ」という表題の会議文書において見出され得る。いくつかの例では、オーディオ符号化ユニット１４は、補間されたｎＦＧ信号４９’を符号化するために使用されるものよりも低い目標ビットレートを使用して、エネルギー補償された環境ＨＯＡ係数４７’をオーディオ符号化し、これによって、補間されたｎＦＧ信号４９’と比較して、エネルギー補償された環境ＨＯＡ係数４７’を場合によってはより圧縮することができる。

【０２３９】

[0240]オーディオ符号化デバイス２０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（復号デバイスにより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するユニットを表し、こ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。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ＨＯＡ係数５９と、符号化されたｎＦＧ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ＨＯＡ係数５９と、符号化されたｎＦＧ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。ビットストリーム２１は、基本のまたは主要な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。

【０２４０】

[0241]図４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の合成を使用して符号化されることになるかベクトルベースの合成を使用して符号化されることになるかに基づいて、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。このビットストリーム出力ユニットは、（ＨＯＡ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）指向性ベースの合成が実行されたか、または（ＨＯＡ係数が記録されたことを検出した結果として）ベクトルベースの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、この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のそれぞれの１つとともに現在のフレームのために使用されるこの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。

【０２４１】

[0242]いくつかの例では、本技法の様々な態様はまた、ＨＯＡ係数１１が合成オーディオオブジェクトから生成されるかどうかをオーディオ符号化デバイス２０が決定することを可能にし得る。本技法のこれらの態様は、オーディオ符号化デバイス２０が、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかのインジケーションを取得するように構成されることを可能にし得る。

【０２４２】

[0243]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するように構成される。

【０２４３】

[0244]これらの例および他の例において、オーディオ符号化デバイス２０は、低減されたフレーム分割された球面調和係数行列を取得するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外するように構成される。

【０２４４】
[0245]これらの例および他の例において、オーディオ符号化デバイス２０は、低減されたフレーム分割された球面調和係数行列を取得して、低減されたフレーム分割された球面調和係数行列の残りのベクトルに基づいて低減されたフレーム分割された球面調和係数行列のベクトルを予測するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外するように構成される。
【０２４５】
[0246]これらの例および他の例において、オーディオ符号化デバイス２０は、低減されたフレーム分割された球面調和係数行列を取得して、低減されたフレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて低減されたフレーム分割された球面調和係数行列のベクトルを予測するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外するように構成される。
【０２４６】
[0247]これらの例および他の例において、オーディオ符号化デバイス２０は、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測するように構成される。
【０２４７】
[0248]これらの例および他の例において、オーディオ符号化デバイス２０は、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予想し、予測されたベクトルに基づいて誤差を計算するようにさらに構成される。
【０２４８】
[0249]これらの例および他の例において、オーディオ符号化デバイス２０は、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予想し、フレーム分割された球面調和係数行列の予測されたベクトルと対応するベクトルとに基づいて誤差を計算するようにさらに構成される。
【０２４９】
[0250]これらの例および他の例において、オーディオ符号化デバイス２０は、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測し、フレーム分割された球面調和係数行列の予測されたベクトルと対応するベクトルとの差の絶対値の合計として誤差を計算するように構成される。
【０２５０】
[0251]これらの例および他の例において、オーディオ符号化デバイス２０は、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測し、フレーム分割された球面調和係数行列の予測されるベクトルと対応するベクトルとに基づいて誤差を計算し、フレーム分割された球面調和係数行列の対応するベクトルのエネルギーと誤差とに基づいて比を計算し、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するために比と閾値を比較するように構成される。
【０２５１】
[0252]これらの例および他の例において、オーディオ符号化デバイス２０は、球面調和係数の圧縮されたバージョンを記憶するビットストリーム２１中のインジケーションをコンフィギャードツー指定するように構成される。
【０２５２】
[0253]いくつかの例において、様々な技法は、オーディオ符号化デバイス２０が、ＨＯ

Ａ係数１１に関して変換を実行することを可能にし得る。これらの例および他の例において、オーディオ符号化デバイス２０は、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを取得するように構成されてよく、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数１１に関して変換を実行することによって生成される。

【０２５３】

[0254]これらの例および他の例において、オーディオ符号化デバイス２０、ここで変換は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数１１の右特異ベクトルを表すＶ行列とを生成する、特異値分解を備える。

【０２５４】

[0255]これらの例および他の例において、オーディオ符号化デバイス２０、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルを備え、Ｕ行列およびＳ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成される。

【０２５５】

[0256]これらの例および他の例において、オーディオ符号化デバイス２０、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列およびＳ行列およびＶ行列は、少なくとも複数の球面調和係数１１に関して特異値分解を実行することによって生成される。

【０２５６】

[0257]これらの例および他の例において、オーディオ符号化デバイス２０、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ符号化デバイス２０はさらに、１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルと１つまたは複数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄを取得するように構成される。

【０２５７】

[0258]これらの例および他の例において、オーディオ符号化デバイス２０、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ符号化デバイス２０はさらに、１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルと１つまたは複数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄをオーディオフレームごとに取得するように構成される。

【０２５８】

[0259]これらの例および他の例において、オーディオ符号化デバイス２０、ここで変換は、音場の明瞭な成分と音場のバックグラウンド成分とを識別するための主要成分分析を備える。

【０２５９】
[0260]本開示で説明される技法の様々な態様は、量子化誤差を補償するように構成されるオーディオ符号化デバイス２０を提供し得る。
【０２６０】
[0261]いくつかの例では、オーディオ符号化デバイス２０は、音場の１つまたは複数の成分を表す１つまたは複数の第１のベクトルを量子化し、１つまたは複数の第１のベクトルの量子化が原因で音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにもたらされた誤差を補償するように構成され得る。
【０２６１】
[0262]これらの例および他の例において、オーディオ符号化デバイスは、音場を記述する複数の球面調和係数に関して特異値分解を実行することによって少なくとも一部生成されるＶ行列の転置から１つまたは複数のベクトルを量子化するように構成される。
【０２６２】
[0263]これらの例および他の例において、オーディオ符号化デバイスはさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成される。
【０２６３】
[0264]これらの例および他の例において、オーディオ符号化デバイスはさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成され、Ｕ行列の１つまたは複数のＵベクトルをＳ行列の１つまたは複数のＳベクトルと乗算することによって計算される１つまたは複数のＵ＊Ｓベクトルにおいて量子化が原因でもたらされる誤差を補償するように構成される。
【０２６４】
[0265]これらの例および他の例において、オーディオ符号化デバイスはさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成し、音場の明瞭な成分に各々が対応するＵ行列の１つまたは複数のU_{DIST}ベクトルを決定し、音場の同じ明瞭な成分に各々が対応するＳ行列の１つまたは複数のS_{DIST}ベクトルを決定し、音場の同じ明瞭な成分に各々が対応するＶ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルを決定するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは複数のV^{T}_{DIST}ベクトルを量子化するように構成され、１つまたは複数の誤差補償されたU_{DIST}＊S_{DIST}ベクトルを生成するために、Ｕ行列の１つまたは複数のU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによって計算される１つまたは複数のU_{DIST}＊S_{DIST}ベクトルにおいて量子化が原因でもたらされる誤差を補償するように構成される。
【０２６５】
[0266]これらの例および他の例において、オーディオ符号化デバイスは、１つまたは複数のU_{DIST}ベクトル、１つまたは複数のS_{DIST}ベクトル、および１つまたは複数のV^{T}_{DIST}ベクトルに基づいて明瞭な球面調和係数を決定し、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルによって明瞭な球面調和係数を分割するために$V^{T}_{Q_DIST}$ベクトルに関して擬似逆変換を実行し、それによって、V^{T}_{DIST}ベクトルの量子化を通じてもたらされる誤差を少なくとも一部補償する誤差補償された１つまたは複数のU_{C_DIST}＊S_{C_DIST}ベクトルを生成するように構成される。
【０２６６】
[0267]これらの例および他の例において、オーディオ符号化デバイスはさらに、複数の

球面調和係数の左特異ベクトルを表すU行列と、複数の球面調和係数の特異値を表すS行列と、複数の球面調和係数の右特異ベクトルを表すV行列とを生成し、音場の１つまたは複数のバックグラウンド成分を記述するU行列の１つまたは複数のU_{BG}ベクトルと音場の１つまたは複数の明瞭な成分を記述するU行列の１つまたは複数のU_{DIST}ベクトルとを決定し、音場の１つまたは複数のバックグラウンド成分を記述するS行列の１つまたは複数のS_{BG}ベクトルと音場の１つまたは複数の明瞭な成分を記述するS行列の１つまたは複数のS_{DIST}ベクトルとを決定し、V行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}ベクトルとを決定するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、ここにおいて、V^{T}_{DIST}ベクトルは音場の１つまたは複数の明瞭な成分を記述し、V^{T}_{BG}は音場の１つまたは複数のバックグラウンド成分を記述する、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは複数のV^{T}_{DIST}ベクトルを量子化するように構成され、誤差補償されたバックグラウンド球面調和係数を生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算することによって形成されるバックグラウンド球面調和係数において量子化が原因でもたらされる誤差を補償するように構成される。

【０２６７】

[0268]これらの例および他の例において、オーディオ符号化デバイスは、V^{T}_{DIST}ベクトルと、U_{DIST}ベクトルをS_{DIST}ベクトルと乗算することによって形成される１つまたは複数の$U_{DIST} * S_{DIST}$ベクトルとに基づいて誤差を決定し、誤差補償されたバックグラウンド球面調和係数を生成するために決定された誤差をバックグラウンド球面調和係数に追加するように構成される。

【０２６８】

[0269]これらの例および他の例において、オーディオ符号化デバイスは、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分をやはり表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償するように構成され、１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するようにさらに構成される。

【０２６９】

[0270]これらの例および他の例において、オーディオ符号化デバイスは、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分をやはり表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償するように構成され、１つまたは複数の誤差補償された第２のベクトルをオーディオ符号化し、オーディオ符号化された１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するようにさらに構成される。

【０２７０】

[0271]本技法の様々な態様はさらに、オーディオ符号化デバイス２０が低減された球面調和係数またはその分解を生成することを可能にし得る。いくつかの例では、オーディオ符号化デバイス２０は、目標ビットレートに基づいて、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解に関する次数低減を実行するように構成されてよく、複数の球面調和係数は音場を表す。

【０２７１】

[0272]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、次数低減を実行する前に、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を識別する１つまたは複数の第２のベクトルとを識別するために、複数の球面調和係数に関して特異値分解を実行するように構成され、１つまたは複数の第１のベクトル、１つまたは複数の第２のベクトル、または１つまたは複数の第１

のベクトルと１つまたは複数の第２のベクトルの両方に関して次数低減を実行するように構成される。

【０２７２】

[0273]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数の球面調和係数またはその分解に関してコンテンツ分析を実行するように構成され、目標ビットレートおよびコンテンツ分析に基づいて、低減された球面調和係数または球面調和係数の低減されたその分解を生成するために、複数の球面調和係数またはその分解に関して次数低減を実行するように構成される。

【０２７３】

[0274]これらの例および他の例において、オーディオ符号化デバイス２０は、複数の球面調和係数またはそれの分解に関して空間分析を実行するように構成される。

【０２７４】

[0275]これらの例および他の例において、オーディオ符号化デバイス２０は、複数の球面調和係数またはそれの分解に関して拡散分析を実行するように構成される。

【０２７５】

[0276]これらの例および他の例において、オーディオ符号化デバイス２０は１つまたは複数のプロセッサであるは、複数の球面調和係数またはそれの分解に関して空間分析と拡散分析とを実行するように構成される。

【０２７６】

[0277]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、低減された球面調和係数または球面調和係数の低減されたそれの分解を含むビットストリームにおいて、低減された球面調和係数または球面調和係数の低減されたそれの分解のそれらが対応する球面基底関数の１つまたは複数の次数および／または１つまたは複数の位数を指定するように構成される。

【０２７７】

[0278]これらの例および他の例において、低減された球面調和係数または球面調和係数の低減されたそれの分解は、複数の球面調和係数または球面調和係数のそれの分解よりも小さい値を有する。

【０２７８】

[0279]これらの例および他の例において、オーディオ符号化デバイス２０は、低減された球面調和係数または球面調和係数の低減されたそれの分解を生成するために、複数の球面調和係数または球面調和係数のそれの分解のベクトルうちで指定された次数および／または位数を有するものを除去するように構成される。

【０２７９】

[0280]これらの例および他の例において、オーディオ符号化デバイス２０は、低減された球面調和係数または球面調和係数の低減されたそれの分解を生成するために、複数の球面調和係数または球面調和係数のそれの分解のベクトルうちで指定された次数および／または位数を有するものを０にするように構成される。

【０２８０】

[0281]本技法の様々な態様はまた、オーディオ符号化デバイス２０が、音場の明瞭な成分を表すように構成されることを可能にし得る。これらの例および他の例において、オーディオ符号化デバイス２０は、音場の明瞭な成分を表すために使用されるべきベクトルの係数の第１の０ではないセットを取得するように構成され、このベクトルは、音場を記述する複数の球面調和係数から分解される。

【０２８１】

[0282]これらの例および他の例において、オーディオ符号化デバイス２０は、係数のすべてを含むようにベクトルの係数の第１の０ではないセットを決定するように構成される。

【０２８２】

[0283]これらの例および他の例において、オーディオ符号化デバイス２０は、係数の第

１の０ではないセットを、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものとして、決定するように構成される。

【０２８３】

[0284]これらの例および他の例において、オーディオ符号化デバイス２０は、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含め、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含めないように、係数の第１の０ではないセットを決定するように構成される。

【０２８４】

[0285]これらの例および他の例において、オーディオ符号化デバイス２０は、係数の第１の０ではないセットを、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを除いて係数のすべてを含むように、決定するように構成される。

【０２８５】

[0286]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、サイドチャネル情報中のベクトルの係数の第１の０ではないセットを指定するように構成される。

【０２８６】

[0287]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、ベクトルの係数の第１の０ではないセットをオーディオ符号化することなく、サイドチャネル情報中のベクトルの係数の第１の０ではないセットを指定するように構成される。

【０２８７】

[0288]これらの例および他の例において、ベクトルは、ベクトルベースの合成を使用して複数の球面調和係数から分解されるベクトルを備える。

【０２８８】

[0289]これらの例および他の例において、ベクトルベースの合成は特異値分解を備える。

【０２８９】

[0290]これらの例および他の例において、ベクトルは、特異値分解を使用して複数の球面調和係数から分解されるＶベクトルを備える。

【０２９０】

[0291]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、ベクトルの係数の０ではないセットをそれによって指定する複数の構成モードの１つを選択し、複数の構成モードの選択された１つに基づいてベクトルの係数の０ではないセットを指定するように構成される。

【０２９１】

[0292]これらの例および他の例において、複数の構成モードの１つは、係数の０ではないセットが係数のすべてを含むことを示す。

【０２９２】

[0293]これらの例および他の例において、複数の構成モードの１つは、係数の第１の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含むことを示す。

【０２９３】

[0294]これらの例および他の例において、複数の構成モードの１つは、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含み、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含まないことを示す。

【０２９４】

[0295]これらの例および他の例において、複数の構成モードの１つは、係数の０ではな

いセットが係数の少なくとも１つを除き係数のすべてを含むことを示す。

【０２９５】

[0296]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数の構成モードの選択された１つをビットストリームにおいて指定するように構成される。

【０２９６】

[0297]本開示で説明される技法の様々な態様はまた、オーディオ符号化デバイス２０が、音場の明瞭な成分を様々な方法で表すように構成されることを可能にし得る。これらの例および他の例において、オーディオ符号化デバイス２０は、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを取得するように構成され、このベクトルは、音場を記述する複数の球面調和係数から分解されたものである。

【０２９７】

[0298]これらの例および他の例において、係数の第１の０ではないセットはベクトルの係数のすべてを含む。

【０２９８】

[0299]これらの例および他の例において、係数の第１の０ではないセットは、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含む。

【０２９９】

[0300]これらの例および他の例において、係数の第１の０ではないセットは、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含み、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含まない。

【０３００】

[0301]これらの例および他の例において、係数の第１の０ではないセットは、十分な指向性情報を有しないものとして識別される係数の少なくとも１つを除き、係数のすべてを含む。

【０３０１】

[0302]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、ベクトルの第１の部分として係数の第１の０ではないセットを抽出するように構成される。

【０３０２】

[0303]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、サイドチャネル情報からベクトルの第１の０ではないセットを抽出し、ベクトルの係数の第１の０ではないセットに基づいて複数の球面調和係数の再合成されたバージョンを取得するように構成される。

【０３０３】

[0304]これらの例および他の例において、ベクトルは、ベクトルベースの合成を使用して複数の球面調和係数から分解されるベクトルを備える。

【０３０４】

[0305]これらの例および他の例において、ベクトルベースの合成は特異値分解を備える。

【０３０５】

[0306]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数の構成モードの１つに従って、ベクトルの係数の０ではないセットをそれによって抽出する複数の構成モードの１つを決定し、複数の構成モードの取得された１つに基づいてベクトルの係数の０ではないセットを抽出するように構成される。

【０３０６】

[0307]これらの例および他の例において、複数の構成モードの１つは、係数の０ではないセットが係数のすべてを含むことを示す。

【０３０７】

[0308]これらの例および他の例において、複数の構成モードの１つは、係数の第１の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含むことを示す。

【０３０８】

[0309]これらの例および他の例において、複数の構成モードの１つは、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含み、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含まないことを示す。

【０３０９】

[0310]これらの例および他の例において、複数の構成モードの１つは、係数の０ではないセットが係数の少なくとも１つを除き係数のすべてを含むことを示す。

【０３１０】

[0311]これらの例および他の例において、オーディオ符号化デバイス２０は、ビットストリームにおいてシグナリングされる値に基づいて複数の構成モードの１つを決定するように構成される。

【０３１１】

[0312]本技法の様々な態様はまた、いくつかの例では、オーディオ符号化デバイス２０が、１つまたは複数の明瞭なオーディオオブジェクト（または言い換えれば、支配的なオーディオオブジェクト）を識別することを可能にし得る。いくつかの例では、オーディオ符号化デバイス２０は、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別するように構成され得る。

【０３１２】

[0313]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、オーディオオブジェクトと関連付けられる球面調和係数に基づいて１つまたは複数のオーディオオブジェクトの指向性を決定するように構成される。

【０３１３】

[0314]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、球面調和係数に関して特異値分解を実行し、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すように構成され、オーディオ符号化デバイス２０は、１つまたは複数のオーディオオブジェクトのそれぞれの指向性が少なくとも一部Ｖ行列に基づくと決定するように構成される。

【０３１４】

[0315]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、再順序付けられたＶ行列においてより大きい指向性指数を有するベクトルがより小さい指向性指数を有するベクトルよりも上に配置されるように、Ｖ行列の１つまたは複数のベクトルを再順序付けるように構成される。

【０３１５】

[0316]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、より大きい指向性指数を有するベクトルがより小さい指向性指数を有するベクトルよりも多くの指向性情報を含むと決定するように構成される。

【０３１６】

[0317]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、ＶＳ行列を生成するためにＶ行列をＳ行列と乗算するように構成され、ＶＳ行列は１つまたは複数のベクトルを含む。

【０３１７】

[0318]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、１４よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択し、対応する二乗された成分を形成するために選択された成分の各々を二乗し、ＶＳ行列の各行に対して、対応するベクトルの指向性指数を決定するために二乗された成分のすべてを加算するように構成される。

【０３１８】

[0319]これらの例および他の例において、オーディオ符号化デバイス２０は、１４よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することが、ＶＳ行列の各行の１８番目の成分において開始しＶＳ行列の各行の３８番目の成分において終了するすべての成分を選択することを備えるように構成される。

【０３１９】

[0320]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、明瞭なオーディオオブジェクトを表すためにＶＳ行列のベクトルのサブセットを選択するように構成される。これらの例および他の例において、オーディオ符号化デバイス２０はＶＳ行列の４つのベクトルを選択するように構成され、選択された４つのベクトルはＶＳ行列のベクトルのすべての４つの最大の指向性指数を有する。

【０３２０】

[0321]これらの例および他の例において、オーディオ符号化デバイス２０は、ベクトルの選択されたサブセットが、明瞭なオーディオオブジェクトが各ベクトルの指向性とエネルギーの両方に基づくことを表すと、決定するように構成される。

【０３２１】

[0322]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、再順序付けられた１つまたは複数の第１のベクトルを決定するために、明瞭なオーディオオブジェクトを表す１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとのエネルギー比較を実行するように構成され、１つまたは複数の第１のベクトルがオーディオデータの第１の部分中の明瞭なオーディオオブジェクトを記述し、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の明瞭なオーディオオブジェクトを記述する。

【０３２２】

[0323]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、再順序付けられた１つまたは複数の第１のベクトルを決定するために、明瞭なオーディオオブジェクトを表す１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとの相互相関付けを実行するように構成され、１つまたは複数の第１のベクトルがオーディオデータの第１の部分中の明瞭なオーディオオブジェクトを記述し、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の明瞭なオーディオオブジェクトを記述する。

【０３２３】

[0324]本技法の様々な態様はまた、いくつかの例では、オーディオ符号化デバイス２０が、ＨＯＡ係数１１の分解に関してエネルギー補償を実行するように構成されることを可能にし得る。これらの例および他の例において、オーディオ符号化デバイス２０は、１つまたは複数のオーディオオブジェクトと対応する指向性情報とを表す複数の球面調和係数の分解された表現を生成するために複数の球面調和係数に関してベクトルベースの合成を実行し、ここにおいて、球面調和係数はある次数と関連付けられ音場を記述する、指向性情報から明瞭な指向性情報とバックグラウンド指向性情報とを決定し、変換されたバックグラウンド指向性情報を生成するためにバックグラウンドオーディオオブジェクトと関連付けられる指向性情報の次数を低減し、音場の全体的なエネルギーを保存するために変換された指向性情報の値を増大させるように補償を適用するように構成され得る。

【０３２４】

[0325]これらの例および他の例において、オーディオ符号化デバイス２０は、オーディオオブジェクトを表すＵ行列およびＳ行列と指向性情報を表すＶ行列とを生成するために

複数の球面調和係数に関して特異値分解を実行し、Ｖ行列の明瞭な列ベクトルとＶ行列のバックグラウンド列ベクトルとを決定し、Ｖ行列の変換されたバックグラウンド列ベクトルを生成するためにＶ行列のバックグラウンド列ベクトルの次数を低減し、音場の全体的なエネルギーを保存するようにＶ行列の変換されたバックグラウンド列ベクトルの値を増大させるために補償を適用するように構成され得る。

【０３２５】

[0326]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、Ｓ行列の顕著な特異値の数を決定するように構成され、Ｖ行列の明瞭な列ベクトルの数はＳ行列の顕著な特異値の数である。

【０３２６】

[0327]これらの例および他の例において、オーディオ符号化デバイス２０は、球面調和係数に対する低減された次数を決定し、低減された次数よりも大きい次数と関連付けられるＶ行列のバックグラウンド列ベクトルの行に対して０という値を決定するように構成される。

【０３２７】

[0328]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、修正された球面調和係数を生成するために、Ｕ行列のバックグラウンド列と、Ｓ行列のバックグラウンド列と、Ｖ行列の変換されたバックグラウンド列の転置とを組み合わせるように構成される。

【０３２８】

[0329]これらの例および他の例において、修正された球面調和係数は、音場の１つまたは複数のバックグラウンド成分を記述する。

【０３２９】

[0330]これらの例および他の例において、オーディオ符号化デバイス２０は、Ｖ行列のバックグラウンド列ベクトルのあるベクトルの第１のエネルギーとＶ行列の変換されたバックグラウンド列ベクトルのあるベクトルの第２のエネルギーとを決定し、Ｖ行列の変換されたバックグラウンド列ベクトルのそのベクトルの各要素に増幅値を適用するように構成され、増幅値は第２のエネルギーに対する第１のエネルギーの比を備える。

【０３３０】

[0331]これらの例および他の例において、オーディオ符号化デバイス２０は、Ｖ行列のバックグラウンド列ベクトルのあるベクトルの第１の二乗平均平方根エネルギーとＶ行列の変換されたバックグラウンド列ベクトルのあるベクトルの第２の二乗平均平方根エネルギーとを決定し、Ｖ行列の変換されたバックグラウンド列ベクトルのそのベクトルの各要素に増幅値を適用するように構成され、増幅値は第２のエネルギーに対する第１のエネルギーの比を備える。

【０３３１】

[0332]本開示で説明される技法の様々な態様はまた、オーディオ符号化デバイス２０が、ＨＯＡ係数１１の分解されたバージョンに関して補間を実行することを可能にし得る。いくつかの例では、オーディオ符号化デバイス２０は、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得するように構成され得る。

【０３３２】

[0333]これらの例および他の例において、第１の分解は、第１の複数の球面調和係数の右特異ベクトルを表す第１のＶ行列を備える。

【０３３３】

[0334]これらの事例および他の事例において、第２の分解は、第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える。

【０３３４】

[0335]これらの例および他の例において、第１の分解は、第１の複数の球面調和係数の

右特異ベクトルを表す第１のＶ行列を備え、第２の分解は、第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える。

【０３３５】

[0336]これらの例および他の例において、時間セグメントはオーディオフレームのサブフレームを備える。

【０３３６】

[0337]これらの例および他の例において、時間セグメントはオーディオフレームの時間サンプルを備える。

【０３３７】

[0338]これらの例および他の例において、オーディオ符号化デバイス２０は、第１の複数の球面調和係数のある球面調和係数に対する第１の分解および第２の分解の補間された分解を取得するように構成される。

【０３３８】

[0339]これらの例および他の例において、オーディオ符号化デバイス２０は、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分に対する第１の分解および第２のフレームに含まれる第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を取得するように構成され、オーディオ符号化デバイス２０はさらに、第１の複数の球面調和係数の第１の人工時間成分を生成するために、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分の第１の時間成分に補間された分解を適用し、第２の複数の球面調和係数の第２の人工時間成分を生成するために、第２のフレームに含まれる第２の複数の球面調和係数の第２の部分の第２の時間成分にそれぞれの補間された分解を適用するように構成される。

【０３３９】

[0340]これらの例および他の例において、第１の時間成分は、第１の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０３４０】

[0341]これらの例および他の例において、第２の時間成分は、第２の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０３４１】

[0342]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、第１の人工時間成分と第２の人工時間成分とを受信し、第１の複数の球面調和係数の第１の部分に対する第１の分解および第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を計算し、第１の時間成分を復元するために補間された分解の逆数を第１の人工時間成分を適用し、第２の時間成分を復元するために補間された分解の逆数を第２の人工時間成分を適用するように構成される。

【０３４２】

[0343]これらの例および他の例において、オーディオ符号化デバイス２０は、第１の複数の球面調和係数の第１の空間成分と第２の複数の球面調和係数の第２の空間成分とを補間するように構成される。

【０３４３】

[0344]これらの例および他の例において、第１の空間成分は、第１の複数の球面調和係数の左特異ベクトルを表す第１のＵ行列を備える。

【０３４４】

[0345]これらの例および他の例において、第２の空間成分は、第２の複数の球面調和係数の左特異ベクトルを表す第２のＵ行列を備える。

【０３４５】

[0346]これらの例および他の例において、第１の空間成分は第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分は第２の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表す。

【０３４６】

[0347]これらの例および他の例において、第１の空間成分は第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分は第２の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、オーディオ符号化デバイス２０は第１の空間成分の最後のＮ個の要素と第２の空間成分の最初のＮ個の要素とを補間するように構成される。

【０３４７】

[0348]これらの例および他の例において、第２の複数の球面調和係数は、時間領域において第１の複数の球面調和係数に後続する。

【０３４８】

[0349]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、第１の複数の球面調和係数の第１の分解を生成するために、第１の複数の球面調和係数を分解するように構成される。

【０３４９】

[0350]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、第２の複数の球面調和係数の第２の分解を生成するために、第２の複数の球面調和係数を分解するように構成される。

【０３５０】

[0351]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、第１の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第１の複数の球面調和係数の特異値を表すＳ行列と、第１の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第１の複数の球面調和係数に関して特異値分解を実行するように構成される。

【０３５１】

[0352]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、第２の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第２の複数の球面調和係数の特異値を表すＳ行列と、第２の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第２の複数の球面調和係数に関して特異値分解を実行するように構成される。

【０３５２】

[0353]これらの例および他の例において、第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、音場の平面波表現を表す。

【０３５３】

[0355]これらの例および他の例において、第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、一緒にミキシングされた１つまたは複数のモノラルオーディオオブジェクトを表す。

【０３５４】

[0356]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、３次元音場を表すそれぞれの第１の球面調和係数および第２の球面調和係数を備える。

【０３５５】

[0357]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、１よりも大きい次数を有する少なくとも１つの球面基底関数と関連付けられる。

【０３５６】

[0358]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、４に等しい次数を有する少なくとも１つの球面基底関数と関連付けられる。

【０３５７】

[0359]これらの例および他の例において、補間は第１の分解および第２の分解の重み付けられた補間であり、第１の分解に適用された重み付けられた補間の重みは、第１の分解

および第２の分解のベクトルによって表される時間に反比例し、第２の分解に適用される重み付けられた補間の重みは、第１の分解および第２の分解のベクトルによって表される時間に比例する。

【０３５８】

[0360]これらの例および他の例において、分解された補間された球面調和係数は、第１の複数の球面調和係数および第２の複数の球面調和係数の空間成分と時間成分の少なくとも１つを平滑化する。

【０３５９】

[0361]これらの例および他の例において、オーディオ符号化デバイス２０は、スカラーを取得するためにＵｓ［ｎ］＝ＨＯＡ（ｎ）＊（Ｖ＿ｖｅｃ［ｎ］）－１を計算するように構成される。

【０３６０】

[0362]これらの例および他の例において、補間は線形補間を備える。これらの例および他の例において、補間は非線形補間を備える。これらの例および他の例において、補間はコサイン補間を備える。これらの例および他の例において、補間は重みづけられたコサイン補間を備える。これらの例および他の例において、補間はキュービック補間を備える。これらの例および他の例において、補間は適応スプライン補間を備える。これらの例および他の例において、補間は最小曲率補間を備える。

【０３６１】

[0363]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、時間セグメントの分解された補間された球面調和係数の表現と補間のタイプのインジケーションとを含むビットストリームを生成するように構成される。

【０３６２】

[0364]これらの例および他の例において、このインジケーションは、補間のタイプと対応付けられる１つまたは複数のビットを備える。

【０３６３】

[0365]このようにして、本開示で説明される技法の様々な態様は、オーディオ符号化デバイス２０が、時間セグメントの分解された補間された球面調和係数の表現と補間のタイプのインジケーションとを含むビットストリームを取得するように構成されることを可能にし得る。

【０３６４】

[0366]これらの例および他の例において、このインジケーションは、補間のタイプと対応付けられる１つまたは複数のビットを備える。

【０３６５】

[0367]この点において、オーディオ符号化デバイス２０は、オーディオ符号化デバイス２０がいくつかの例では音場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を生成するように構成され得るという点で、本技法の一実施形態を表すことができ、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０３６６】

[0368]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分を圧縮するときに使用される予測モードを指定するフィールドを含むようにビットストリームを生成するように構成される。

【０３６７】

[0369]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を含むようにビットストリームを生成するように構成される。

【０３６８】

[0370]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィー

ルドを含むようにビットストリームを生成するように構成される。
【０３６９】
[0371]これらの例および他の例において、値はｎｂｉｔｓ値を備える。
【０３７０】
[0372]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを含むようにビットストリームを生成するように構成され、値は複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す。
【０３７１】
[0373]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためにハフマンコードを含むようにビットストリームを生成するように構成される。
【０３７２】
[0374]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分が正の値か負の値かを識別する符号ビットを含むようにビットストリームを生成するように構成される。
【０３７３】
[0375]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分の残差値を表すためにハフマンコードを含むようにビットストリームを生成するように構成される。
【０３７４】
[0376]これらの例および他の例において、ベクトルベースの合成は特異値分解を備える。
【０３７５】
[0377]この点において、オーディオ符号化デバイス２０はさらに、オーディオ符号化デバイス２０がいくつかの例では複数の空間成分のある空間成分を複数の空間成分の残りに対するその空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別するように構成され得るという点で、本技法の様々な態様を実装することができ、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【０３７６】
[0378]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を圧縮するときに使用される予測モードに基づいてハフマンコードブックを識別するように構成される。
【０３７７】
[0379]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、ハフマンコードブックを識別するハフマンテーブル情報を使用して、ビットストリームにおいて表される。
【０３７８】
[0380]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィールドを使用して、ビットストリームにおいて表される。
【０３７９】
[0381]これらの例および他の例において、値はｎｂｉｔｓ値を備える。
【０３８０】
[0382]これらの例および他の例において、ビットストリームは、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを備え、値は複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す。
【０３８１】
[0383]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なく

とも一部、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すために識別されたハフマンコードブックから選択されたハフマンコードを使用して、ビットストリームにおいて表される。

【０３８２】

[0384]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分が正の値か負の値かを識別する符号ビットを使用して、ビットストリームにおいて表される。

【０３８３】

[0385]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分の残差値を表すために識別されたハフマンコードブックから選択されたハフマンコードを使用して、ビットストリームにおいて表される。

【０３８４】

[0386]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分の圧縮されたバージョンを生成するために、識別されたハフマンコードブックに基づいて空間成分を圧縮し、空間成分の圧縮されたバージョンを含むようにビットストリームを生成するように構成される。

【０３８５】

[0387]その上、オーディオ符号化デバイス２０は、いくつかの例では、オーディオ符号化デバイス２０が音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定するように構成され得るという点で、本技法の様々な態様を実装することができ、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０３８６】

[0388]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、目標ビットレートに基づいて量子化ステップサイズを決定するように構成される。

【０３８７】

[0389]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を表すために使用されるビットの数の推定値を決定し、その推定値と目標ビットレートとの間の差に基づいて量子化ステップサイズを決定するように構成される。

【０３８８】

[0390]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を表すために使用されるビットの数の推定値を決定し、その推定値と目標ビットレートとの間の差を決定し、その差を目標ビットレートに加算することによって量子化ステップサイズを決定するように構成される。

【０３８９】

[0391]これらの例および他の例において、オーディオ符号化デバイス２０は、目標ビットレートに対応するコードブックを仮定して、空間成分に対して生成されるべきビットの数の推定値を計算するように構成される。

【０３９０】

[0392]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を圧縮するときに使用されるコーディングモードを仮定して、空間成分に対して生成されるべきビットの数のエスティメイテッドを計算するように構成される。

【０３９１】

[0393]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して空間成分に対して生成されるべきビットの数の第１の推定値を計算し、空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して空間成分に対して生成されるべきビットの数の第２の推定値を計算し、ビットの数の決定された推定値として使用されることになる、第１の推定値と第２の推定値のうちでビットの数が少ない方を選択するように構成される。

【０３９２】
[0394]これらの例および他の例において、オーディオ符号化デバイス２０は、空間成分が対応するカテゴリを識別するカテゴリ識別子を識別し、カテゴリに対応する空間成分を圧縮するときに生じるであろう空間成分の残差値のビット長を識別し、カテゴリ識別子を表すために使用されるビットの数を残差値のビット長に加算することによって少なくとも一部、ビットの数の推定値を決定するように構成される。
【０３９３】
[0395]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分を圧縮するときに使用されるべき、複数のコードブックのうちの１つを選択するように構成される。
【０３９４】
[0396]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数のコードブックの各々を使用して空間成分を表すために使用されるビットの数の推定値を決定し、決定された推定値をもたらした複数のコードブックのうちでビットの数が最小のものを選択するように構成される。
【０３９５】
[0397]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、複数のコードブックの１つまたは複数を使用して空間成分を表すために使用されるビットの数の推定値を決定し、複数のコードブックの１つまたは複数は、空間成分の他の要素に対する、圧縮されるべき空間成分の要素の次数に基づいて選択されるように構成される。
【０３９６】
[0398]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分が後続の空間成分から予測されないときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するように構成される。
【０３９７】
[0399]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分が後続の空間成分から予測されるときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するように構成される。
【０３９８】
[0400]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分が音場の中の合成オーディオオブジェクトを表すときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するように構成される。
【０３９９】
[0401]これらの例および他の例において、合成オーディオオブジェクトは、パルス符号変調（ＰＣＭ）されたオーディオオブジェクトを備える。
【０４００】
[0402]これらの例および他の例において、オーディオ符号化デバイス２０はさらに、空間成分が音場の中の録音されたオーディオオブジェクトを表すときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するように構成される。
【０４０１】
[0403]上で説明された様々な例の各々において、オーディオ符号化デバイス２０は、方法を実行することができ、またはそうでなければ、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。いくつかの例では、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。いくつかの例では、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が実行され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。

【０４０２】

[0404]図５は、図３のオーディオ復号デバイス２４をより詳細に示すブロック図である。図５の例に示されているように、オーディオ復号デバイス２４は、抽出ユニット７２と、指向性ベース再構築ユニット９０と、ベクトルベース再構築ユニット９２とを含み得る。

【０４０３】

[0405]抽出ユニット７２は、ビットストリーム２１を受信してＨＯＡ係数１１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るユニットを表し得る。抽出ユニット７２は、ＨＯＡ係数１１が様々なバージョンを介して符号化されたかどうかを、上で述べられたシンタックス要素（たとえば、図１０Ｅおよび図１０Ｈ（ｉ）～図１００（ｉｉ）の例に示されるＣｈａｎｎｅｌＴｙｐｅシンタックス要素）から決定することができる。指向性ベースの符号化が実行されたとき、抽出ユニット７２は、ＨＯＡ係数１１の指向性ベースのバージョンと、この符号化されたバージョンと関連付けられるシンタックス要素（図５の例では指向性ベースの情報９１として示される）とを抽出し、この指向性ベースの情報９１を指向性ベース再構築ユニット９０に渡すことができる。この指向性ベース再構築ユニット９０は、指向性ベースの情報９１に基づいてＨＯＡ係数１１’の形態でＨＯＡ係数を再構築するように構成される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１０～図１００（ｉｉ）および図１１の例に関してより詳細に説明される。

【０４０４】

[0406]ＨＯＡ係数１１がベクトルベースの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と、符号化された環境ＨＯＡ係数５９と、符号化されたｎＦＧ信号５９とを抽出することができる。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を量子化ユニット７４に渡し、符号化された環境ＨＯＡ係数５９を符号化されたｎＦＧ信号６１とともに音響心理学的復号ユニット８０に渡すことができる。

【０４０５】

[0407]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ＨＯＡ係数５９と、符号化されたｎＦＧ信号５９とを抽出するために、抽出ユニット７２は、ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈと示されるシンタックス要素を含むサイドチャネル情報５７を取得することができる。抽出ユニット７２は、サイドチャネル情報５７からのｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈを解析することができる。抽出ユニット７２は、ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素に基づいて、上で説明された構成モードの任意の１つで動作するように構成され得る。

【０４０６】

[0408]抽出ユニット７２は次いで、サイドチャネル情報５７から圧縮された形態の低減されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}を解析するために、構成モードの任意の１つに従って動作する。抽出ユニット７２は、ＶＶｅｃｔｏｒＤａｔａのための以下のシンタックステーブルにおいて提示されるシンタックスを伴う以下の擬似コードにおいて提示される、スイッチステートメントに従って動作することができる。

【０４０７】

【数２３】

```
switch CodedVVecLength{
      case 0:
            VVecLength = NumOfHoaCoeffs;
            for (m=0; m<VVecLength; ++m){
                  VVecCoeffId[m] = m;
            }
            break;
      case 1:
            VVecLength = NumOfHoaCoeffs-MinNumOfCoeffsForAmbHOA -
      NumOfContAddHoaChans;
      n = 0;
      for(m=MinNumOfCoeffsForAmbHOA; m<NumOfHoaCoeffs; ++m){
                  CoeffIdx = m+1;
            if(CoeffIdx isNotMemberOf ContAddHoaCoeff){
                  VVecCoeffId[n] = CoeffIdx-1;
                  n++;
            }
       }
            break;
case 2:
            VVecLength = NumOfHoaCoeffs - MinNumOfCoeffsForAmbHOA;
            for (m=0; m< VVecLength; ++m){
                  VVecCoeffId[m] = m + MinNumOfCoeffsForAmbHOA;
            }
     break;
     case 3:
     VVecLength = NumOfHoaCoeffs - NumOfContAddHoaChans;
     n = 0;
     for(m=0; m<NumOfHoaCoeffs; ++m){
                  c = m+1;
           if(c isNotMemberOf ContAddHoaCoeff){
                  VVecCoeffId[n] = c-1;
                  n++;
            }
     }
}
```

【０４０８】

【表９】

シンタックス	ビットの数	ニーモニック
VVectorData(i)		
{		
if (NbitsQ(k)[i] == 5){		
for (m=0; m< VVecLength; ++m){		
VVec[i][VVecCoeffId[m]](k) = (**VecVal** / 128.0) – 1.0;	**8**	**uimsbf**
}		
elseif(NbitsQ(k)[i] >= 6){		
for (m=0; m< VVecLength; ++m){		
huffIdx = *huffSelect*(VVecCoeffId[m], PFlag[i], CbFlag[i]);		
cid = *huffDecode*(NbitsQ[i], huffIdx, **huffVal**);	**動的**	**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**	**uimsbf**
}		
}		
VVec[i][VVecCoeffId[m]](k) = aVal[i][m]*(2^(16 – NbitsQ(k)[i])*aVal[i][m])/2^15;		
if (PFlag(k)[i] == 1) {		
VVec[i][VVecCoeffId[m]](k)+= VVec[i][VVecCoeffId[m]](k-1)		
}		
}		
}		

【０４０９】

[0409]前述のシンタックステーブルにおいて、４つのケース（ケース０～３）を伴う第１のスイッチステートメントは、V^{T}_{DIST}ベクトルの長さを係数の数（ＶＶｅｃＬｅｎｇｔｈ）およびインデックス（ＶＶｅｃＣｏｅｆｆＩｄ）に関して決定する方法を提供する。第１のケースであるケース０は、V^{T}_{DIST}ベクトルの係数のすべて（ＮｕｍＯｆＨｏａＣｏｅｆｆｓ）が指定されることを示す。第２のケースであるケース１は、V^{T}_{DIST}ベクトルの係数のうちでＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡより大きい数に対応するもののみが指定されることを示し、これは上で$(N_{DIST}+1)^2-(N_{BG}+1)^2$と呼ばれるものを示し得る。さらに、ＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎにおいて識別されるＮｕｍＯｆＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎ係数が差し引かれる。リストＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎは、次数ＭｉｎＡｍｂＨｏａＯｒｄｅｒを超える次数に対応する追加のチャネル（ここで「チャネル」はある次数と位数の組合せに対応する特定の係数を指す）を指定する。第３のケースであるケース２は、V^{T}_{DIST}ベクトルの係数のうちでＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡより大きい数に対応するものが指定されることを示し、これは上で$(N_{DIST}+1)^2-(N_{BG}+1)^2$と呼ばれるものを示し

得る。第４のケースであるケース３は、V^{T}_{DIST}ベクトルの係数のうちでＮｕｍＯｆＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎによって識別される係数を除去した後で残るものが指定されることを示す。ＶＶｅｃＬｅｎｇｔｈとＶＶｅｃＣｏｅｆｆＩｄリストの両方が、ＨＯＡＦｒａｍｅ内のすべてのＶＶｅｃｔｏｒに対して有効である。

【０４１０】

[0410]このスイッチステートメントの後で、一様逆量子化を実行するかどうかの決定は、ＮｂｉｔｓＱ（または上で示されたようにｎｂｉｔｓ）によって制御されてよく、ＮｂｉｔｓＱが５に等しい場合、一様な８ビットの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。上で論じられた予測モードは上のシンタックステーブルではＰＦｌａｇとして示されるが、ＨＴ情報ビットは上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、復号が上で説明されたものと実質的に同様の方式でどのように行われるかを指定する。上で述べられた様々なケースの各々に適合するビットストリーム２１の様々な例は、図１０Ｈ（ｉ）～図１００（ｉｉ）に関して以下でより詳細に説明される。

【０４１１】

[0411]ベクトルベースの再構築ユニット９２は、ＨＯＡ係数１１’を再構築するために、ベクトルベースの合成ユニット２７に関して上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るユニットを表す。ベクトルベースの再構築ユニット９２は、量子化ユニット７４と、空間－時間的補間ユニット７６と、フォアグラウンド定式化ユニット７８と、音響心理学的復号ユニット８０と、ＨＯＡ係数定式化ユニット８２と、再順序付けユニット８４とを含み得る。

【０４１２】

[0412]量子化ユニット７４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を逆量子化してそれによって低減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、図４の例に示される量子化ユニット５２とは逆の方式で動作するように構成されるユニットを表し得る。逆量子化ユニット７４は、いくつかの例では、量子化ユニット５２に関して上で説明されたものとは逆の方式で、ある形態のエントロピー復号とスカラー逆量子化とを実行することができる。逆量子化ユニット７４は、低減されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}を再順序付けユニット８４に転送することができる。

【０４１３】

[0413]音響心理学的復号ユニット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４の例に示される音響心理学的オーディオコーディングユニット４０とは逆の方式で動作することができる。音響心理学的復号ユニット８０は、エネルギー補償された環境ＨＯＡ係数４７’をＨＯＡ係数定式化ユニット８２に渡し、ｎＦＧ信号４９’を再順序付け８４に渡すことができる。

【０４１４】

[0414]再順序付けユニット８４は、再順序付けユニット３４に関して上で説明されたものとは逆の方式で動作するように構成されるユニットを表し得る。再順序付けユニット８４は、ＨＯＡ係数１１のフォアグラウンド成分の元の次数を示すシンタックス要素を受信することができる。再順序付けユニット８４は、これらの再順序付けシンタックス要素に基づいて、再順序付けられたｎＦＧ信号４９’’と再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’とを生成するために、補間されたｎＦＧ信号４９’と低減されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}とを再順序付けることができる。再順序付けユニット８４は、再順序付けられたｎＦＧ信号４９’’をフォアグラウンド定式化ユニット７８に出力し、再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’を空間－時間的補間ユニット７６に出力することができる。

【０４１５】
[0415]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に関して上で説明されたものと同様の方式で動作することができる。空間－時間的補間ユニット７６は、再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’を受信し、補間されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’’を生成するために再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’および再順序付けら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}’’をフォアグラウンド定式化ユニット７８に転送することができる。
【０４１６】
[0416]フォアグラウンド定式化ユニット７８は、フォアグラウンドＨＯＡ係数６５を生成するために、補間されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’’および再順序付けられたｎＦＧ信号４９’’に関して行列乗算を実行するように構成されるユニットを表し得る。フォアグラウンド定式化ユニット７８は、補間されたフォアグラウンドＶ［ｋ］ベクトル５５_{k}’’による再順序付けられたｎＦＧ信号４９’’の行列乗算を実行することができる。
【０４１７】
[0417]ＨＯＡ係数定式化ユニット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を環境ＨＯＡチャネル４７’に加算するように構成されるユニットを表すことができ、プライム記号の表記は、これらのＨＯＡ係数１１’がＨＯＡ係数１１と類似し得るが同一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’の差は、有損失の送信媒体を通じた送信、量子化、または他の有損失の演算が原因の損失に起因し得る。
【０４１８】
[0418]このようにして、本技法は、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、ビットストリームから、量子化された指向性情報と、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、符号化された環境高次アンビソニックス（ＨＯＡ）係数とを決定することを可能とし、ここにおいて、量子化された指向性情報および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は音場のフォアグラウンド成分を記述するフォアグラウンドＨＯＡ係数を表し、符号化された環境ＨＯＡ係数は音場の環境成分を記述する、指向性情報を生成するために量子化された指向性情報を逆量子化し、補間された指向性情報を生成するために指向性情報に関して空間－時間的補間を実行し、フォアグラウンドオーディオオブジェクトを生成するために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をオーディオ復号し環境ＨＯＡ係数を生成するために符号化された環境ＨＯＡ係数をオーディオ復号し、補間された指向性情報およびフォアグラウンドオーディオオブジェクトの関数としてフォアグラウンドＨＯＡ係数を決定し、フォアグラウンドＨＯＡ係数および環境ＨＯＡ係数の関数としてＨＯＡ係数を決定する。
【０４１９】
[0419]このようにして、本技法の様々な態様は、統一されたオーディオ復号デバイス２４が２つの異なる圧縮スキームを切り替えることを可能にし得る。いくつかの例では、オーディオ復号デバイス２４は、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の圧縮スキームの１つを選択し、複数の圧縮スキームの選択された１つを使用して球面調和係数の圧縮されたバージョンを解凍するように構成され得る。これらの例および他の例において、オーディオ復号デバイス２４は統合された復号器を備える。
【０４２０】
[0420]いくつかの例は、オーディオ復号デバイス２４は、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかのインジケーションを取得するように構成され得る。
【０４２１】

[0421]これらの例および他の例において、オーディオ復号デバイス２４は、球面調和係数の圧縮されたバージョンを記憶するビットストリームからインジケーションを取得するように構成される。

【０４２２】

[0422]このようにして、本技法の様々な態様は、オーディオ復号デバイス２４が、音場の明瞭な成分とバックグラウンド成分とを記述するベクトルを取得することを可能にし得る。いくつかの例において、オーディオ復号デバイス２４は、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するように構成されてよく、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される。

【０４２３】

[0423]これらの例および他の例において、オーディオ復号デバイス２４、ここで変換は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成する、特異値分解を備える。

【０４２４】

[0424]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルを備え、Ｕ行列およびＳ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成される。

【０４２５】

[0425]これらの例および他の例において、オーディオ復号デバイス２４はさらに、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルのオーディオ復号されたバージョンを生成するために、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルをオーディオ復号するように構成される。

【０４２６】

[0426]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列およびＳ行列およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成される。

【０４２７】

[0427]これらの例および他の例において、オーディオ復号デバイス２４はさらに、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルのオーディオ復号されたバージョンを生成するために、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルをオーディオ復号するように構成される。

【０４２８】

[0428]これらの例および他の例において、複数の球面調和関数のうちで音場の明瞭な成分を表すものを復元するために、$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{DIST}ベクトルと乗算するようにさらに構成されるオーディオ復号デバイス２４。

【０４２９】

[0429]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第２のベクトルは、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{BG}*$

S_{BG}＊V^{T}_{BG}ベクトルを備え、Ｓ行列、Ｕ行列、およびＶ行列は各々、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成される。

【０４３０】

[0430]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第２のベクトルは、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成され、１つまたは複数のオーディオ符号化されたU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルを備え、Ｓ行列、Ｕ行列、およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ復号デバイス２４はさらに、１つまたは複数のオーディオ復号されたU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルを生成するために、１つまたは複数のオーディオ符号化されたU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルをオーディオ復号するように構成される。

【０４３１】

[0431]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化されたU_{DIST}＊S_{DIST}ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ復号デバイス２４はさらに、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを生成するために１つまたは複数のオーディオ符号化されたU_{DIST}＊S_{DIST}ベクトルをオーディオ復号し、複数の球面調和係数のうちで音場の明瞭な成分を記述するものを復元するためにU_{DIST}＊S_{DIST}ベクトルをV^{T}_{DIST}ベクトルと乗算するように構成され、１つまたは複数の第２のベクトルは、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化されたU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルを備え、オーディオ復号デバイス２４はさらに、音場のバックグラウンド成分を記述する複数の球面調和係数の少なくともある部分を復元するために１つまたは複数のオーディオ符号化されたU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルを復号し、複数の球面調和係数の再構築されたバージョンを生成するために、音場の明瞭な成分を記述する複数の球面調和係数を音場のバックグラウンド成分を記述する複数の球面調和係数の少なくともその部分に加算するように構成される。

【０４３２】

[0432]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列およびＳ行列およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ復号デバイス２０はさらに、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと１つまたは複数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄを取得するように構成される。

【０４３３】

[0433]これらの例および他の例において、オーディオ復号デバイス２４、ここで１つまたは複数の第１のベクトルは、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列は、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、オーディオ復号デバイス２４はさらに、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと１つまたは複

数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄをオーディオフレームごとに取得するように構成される。

【０４３４】

[0434]これらの例および他の例において、オーディオ復号デバイス２４、ここで変換は、音場の明瞭な成分と音場のバックグラウンド成分とを識別するための主要成分分析を備える。

【０４３５】

[0435]本開示で説明される技法の様々な態様はまた、オーディオ符号化デバイス２４が、ＨＯＡ係数の分解されたバージョンに関して補間を実行することを可能にし得る。いくつかの例では、オーディオ復号デバイス２４は、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得するように構成され得る。

【０４３６】

[0436]これらの例および他の例において、第１の分解は、第１の複数の球面調和係数の右特異ベクトルを表す第１のＶ行列を備える。

【０４３７】

[0437]これらの事例および他の事例において、第２の分解は、第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える。

【０４３８】

[0438]これらの例および他の例において、第１の分解は、第１の複数の球面調和係数の右特異ベクトルを表す第１のＶ行列を備え、第２の分解は、第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える。

【０４３９】

[0439]これらの例および他の例において、時間セグメントはオーディオフレームのサブフレームを備える。

【０４４０】

[0440]これらの例および他の例において、時間セグメントはオーディオフレームの時間サンプルを備える。

【０４４１】

[0441]これらの例および他の例において、オーディオ復号デバイス２４は、第１の複数の球面調和係数のある球面調和係数に対する第１の分解および第２の分解の補間された分解を取得するように構成される。

【０４４２】

[0442]これらの例および他の例において、オーディオ復号デバイス２４は、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分に対する第１の分解および第２のフレームに含まれる第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を取得するように構成され、オーディオ復号デバイス２４はさらに、第１の複数の球面調和係数の第１の人工時間成分を生成するために、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分の第１の時間成分に補間された分解を適用し、第２の複数の球面調和係数の第２の人工時間成分を生成するために、第２のフレームに含まれる第２の複数の球面調和係数の第２の部分の第２の時間成分にそれぞれの補間された分解を適用するように構成される。

【０４４３】

[0443]これらの例および他の例において、第１の時間成分は、第１の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０４４４】

[0444]これらの例および他の例において、第２の時間成分は、第２の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０４４５】

[0445]これらの例および他の例において、オーディオ復号デバイス２４はさらに、第１の人工時間成分と第２の人工時間成分とを受信し、第１の複数の球面調和係数の第１の部分に対する第１の分解および第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を計算し、第１の時間成分を復元するために補間された分解の逆数を第１の人工時間成分を適用し、第２の時間成分を復元するために補間された分解の逆数を第２の人工時間成分を適用するように構成される。

【０４４６】

[0446]これらの例および他の例において、オーディオ復号デバイス２４は、第１の複数の球面調和係数の第１の空間成分と第２の複数の球面調和係数の第２の空間成分とを補間するように構成される。

【０４４７】

[0447]これらの例および他の例において、第１の空間成分は、第１の複数の球面調和係数の左特異ベクトルを表す第１のＵ行列を備える。

【０４４８】

[0448]これらの例および他の例において、第２の空間成分は、第２の複数の球面調和係数の左特異ベクトルを表す第２のＵ行列を備える。

【０４４９】

[0449]これらの例および他の例において、第１の空間成分は第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分は第２の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表す。

【０４５０】

[0450]これらの例および他の例において、第１の空間成分は第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分は第２の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、オーディオ復号デバイス２４は第１の空間成分の最後のＮ個の要素と第２の空間成分の最初のＮ個の要素とを補間するように構成される。

【０４５１】

[0451]これらの例および他の例において、第２の複数の球面調和係数は、時間領域において第１の複数の球面調和係数に後続する。

【０４５２】

[0452]これらの例および他の例において、オーディオ復号デバイス２４はさらに、第１の複数の球面調和係数の第１の分解を生成するために、第１の複数の球面調和係数を分解するように構成される。

【０４５３】

[0453]これらの例および他の例において、オーディオ復号デバイス２４はさらに、第２の複数の球面調和係数の第２の分解を生成するために、第２の複数の球面調和係数を分解するように構成される。

【０４５４】

[0454]これらの例および他の例において、オーディオ復号デバイス２４はさらに、第１の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第１の複数の球面調和係数の特異値を表すＳ行列と、第１の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第１の複数の球面調和係数に関して特異値分解を実行するように構成される。

【０４５５】

[0455]これらの例および他の例において、オーディオ復号デバイス２４はさらに、第２の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第２の複数の球面調和係数の特異値を表すＳ行列と、第２の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第２の複数の球面調和係数に関して特異値分解を実行するように構成される。

【０４５６】

[0456]これらの例および他の例において、第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、音場の平面波表現を表す。

【０４５７】

[0457]これらの例および他の例において、第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、一緒にミキシングされた１つまたは複数のモノラルオーディオオブジェクトを表す。

【０４５８】

[0458]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、３次元音場を表すそれぞれの第１の球面調和係数および第２の球面調和係数を備える。

【０４５９】

[0459]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、１よりも大きい次数を有する少なくとも１つの球面基底関数と関連付けられる。

【０４６０】

[0460]これらの例および他の例において、第１の複数の球面調和係数および第２の複数の球面調和係数は各々、４に等しい次数を有する少なくとも１つの球面基底関数と関連付けられる。

【０４６１】

[0461]これらの例および他の例において、補間は第１の分解および第２の分解の重み付けられた補間であり、第１の分解に適用された重み付けられた補間の重みは、第１の分解および第２の分解のベクトルによって表される時間に反比例し、第２の分解に適用される重み付けられた補間の重みは、第１の分解および第２の分解のベクトルによって表される時間に比例する。

【０４６２】

[0462]これらの例および他の例において、分解された補間された球面調和係数は、第１の複数の球面調和係数および第２の複数の球面調和係数の空間成分と時間成分の少なくとも１つを平滑化する。

【０４６３】

[0463]これらの例および他の例において、オーディオ復号デバイス２４は、スカラーを取得するためにＵｓ［ｎ］＝ＨＯＡ（ｎ）＊（Ｖ＿ｖｅｃ［ｎ］）－１を計算するように構成される。

【０４６４】

[0464]これらの例および他の例において、補間は線形補間を備える。これらの例および他の例において、補間は非線形補間を備える。これらの例および他の例において、補間はコサイン補間を備える。これらの例および他の例において、補間は重みづけられたコサイン補間を備える。これらの例および他の例において、補間はキュービック補間を備える。これらの例および他の例において、補間は適応スプライン補間を備える。これらの例および他の例において、補間は最小曲率補間を備える。

【０４６５】

[0465]これらの例および他の例において、オーディオ復号デバイス２４はさらに、時間セグメントの分解された補間された球面調和係数の表現と補間のタイプのインジケーションとを含むビットストリームを生成するように構成される。

【０４６６】

[0466]これらの例および他の例において、このインジケーションは、補間のタイプと対応付けられる１つまたは複数のビットを備える。

【０４６７】

[0467]これらの例および他の例において、オーディオ復号デバイス２４はさらに、時間セグメントの分解された補間された球面調和係数の表現と補間のタイプのインジケーションとを含むビットストリームを取得するように構成される。

【０４６８】
[0468]これらの例および他の例において、このインジケーションは、補間のタイプと対応付けられる１つまたは複数のビットを備える。
【０４６９】
[0469]本技法の様々な態様は、いくつかの例では、オーディオ復号デバイス２４が、音場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を取得するように構成されることをさらに可能にでき、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【０４７０】
[0470]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用される予測モードを指定するフィールドを使用して、ビットストリームにおいて表される。
【０４７１】
[0471]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を使用して、ビットストリームにおいて表される。
【０４７２】
[0472]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィールドを使用して、ビットストリームにおいて表される。
【０４７３】
[0473]これらの例および他の例において、値はｎｂｉｔｓ値を備える。
【０４７４】
[0474]これらの例および他の例において、ビットストリームは、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを備え、値は複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す。
【０４７５】
[0475]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためのハフマンコードを使用して、ビットストリームにおいて表される。
【０４７６】
[0476]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分が正の値か負の値かを識別する符号ビットを使用して、ビットストリームにおいて表される。
【０４７７】
[0477]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分の残差値を表すためのハフマンコードを使用して、ビットストリームにおいて表される。
【０４７８】
[0478]これらの例および他の例において、デバイスはオーディオ復号デバイスを備える。
【０４７９】
[0479]技法の様々な態様はまた、オーディオ復号デバイス２４が、複数の圧縮された空間成分のある空間成分の圧縮されたバージョンを複数の圧縮された空間成分の残りに対するその空間成分の圧縮されたバージョンの次数に基づいて解凍するときに使用するハフマンコードブックを識別することを可能にでき、空間成分は複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。
【０４８０】
[0480]これらの例および他の例において、オーディオ復号デバイス２４は、音場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を取得し、空間成分を取得するた

めに、少なくとも一部、識別されたハフマンコードブックを使用して、空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８１】

[0481]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用される予測モードを指定するフィールドを使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、空間成分を取得するために、予測モードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８２】

[0482]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、ハフマンテーブル情報に少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８３】

[0483]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィールドを使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、その値に少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８４】

[0484]これらの例および他の例において、値はｎｂｉｔｓ値を備える。

【０４８５】

[0485]これらの例および他の例において、ビットストリームは、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを備え、値は、複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表し、オーディオ復号デバイス２４は、その値に少なくとも一部基づいて空間成分の複数の圧縮されたバージョンを解凍するように構成される。

【０４８６】

[0486]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためのハフマンコードを使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、ハフマンコードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８７】

[0487]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分が正の値か負の値かを識別する符号ビットを使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、符号ビットに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８８】

[0488]これらの例および他の例において、空間成分の圧縮されたバージョンは、少なくとも一部、空間成分の残差値を表すためのハフマンコードを使用して、ビットストリームにおいて表され、オーディオ復号デバイス２４は、識別されたハフマンコードブックに含まれるハフマンコードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するように構成される。

【０４８９】

[0489]上で説明された様々な例の各々において、オーディオ復号デバイス２４は、ある方法を実行し、またはそうでなければ、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。いくつかの例では、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。いくつかの例では

、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が実行され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。

【０４９０】

[0490]図６は、本開示において説明される技法の様々な態様を実行する際の、図４の例に示されるコンテンツ分析ユニット２６などのオーディオ符号化デバイスのコンテンツ分析ユニットの例示的な動作を示すフローチャートである。

【０４９１】

[0491]コンテンツ分析ユニット２６は、音場を表すＨＯＡ係数１１が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割されたＨＯＡ係数を取得することができ（９３）、これは、４次の表現（すなわち、Ｎ＝４）では２５×１０２４のサイズであり得る。フレーム分割されたＨＯＡ係数（これは本明細書ではフレーム分割されたＳＨＣ行列１１として示されることもあり、後続のフレーム分割されたＳＨＣ行列はフレーム分割されたＳＨＣ行列２７Ｂ、２７Ｃなどとして示され得る、など）を取得した後で、コンテンツ分析ユニット２６は次いで、低減されたフレーム分割されたＨＯＡ係数を生成するために、フレーム分割されたＨＯＡ係数１１の第１のベクトルを除外することができる（９４）。

【０４９２】

[0492]コンテンツ分析ユニット２６は次いで、低減されたフレーム分割されたＨＯＡ係数の残りのベクトルから、低減されたフレーム分割されたＨＯＡ係数の第１の０ではないベクトルを予測することができる（９５）。第１の０ではないベクトルを予測した後で、コンテンツ分析ユニット２６は、予測された第１の０ではないベクトルおよび実際の０ではないベクトルに基づいて、誤差を得ることができる（９６）。誤差が得られると、コンテンツ分析ユニット２６は、実際の第１の０ではないベクトルのエネルギーおよび誤差に基づいて、比を計算することができる（９７）。コンテンツ分析ユニット２６は次いで、この比を閾値と比較することができる（９８）。比が閾値を超えないとき（「いいえ」９８）、コンテンツ分析ユニット２６は、フレーム分割されたＳＨＣ行列１１が録音から生成されたと決定し、ＳＨＣ行列１１の対応するコーディングされた表現が録音から生成されたことをビットストリームにおいて示すことができる（１００、１０１）。比が閾値を超えるとき（「はい」９８）、コンテンツ分析ユニット２６は、フレーム分割されたＳＨＣ行列１１が合成オーディオオブジェクトから生成されると決定し、ＳＨＣ行列１１の対応するコーディングされた表現が合成オーディオオブジェクトから生成されたことをビットストリームにおいて示すことができる（１０２、１０３）。いくつかの例では、フレーム分割されたＳＨＣ行列１１が録音から生成されたとき、コンテンツ分析ユニット２６は、フレーム分割されたＳＨＣ行列１１をベクトルベースの合成ユニット２７に渡す（１０１）。いくつかの例では、フレーム分割されたＳＨＣ行列１１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２６は、フレーム分割されたＳＨＣ行列１１を指向性ベースの合成ユニット２８に渡す（１０４）。

【０４９３】

[0493]図７は、本開示において説明されるベクトルベースの合成技法の様々な態様を実行する際の、図４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が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ＨＯＡ係数（たとえば、ＳＶＤの場合、変換されたＨＯＡ係数はＵＳ［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ＨＯＡ係数に関してＬＩＴを適用することができる（１０７）。

【０４９４】

[0494]オーディオ符号化デバイス２０は次に、上で説明された方式で様々なパラメータを識別するために、ＵＳ［ｋ］ベクトル３３、ＵＳ［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ＨＯＡ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）。

【０４９５】

[0495]オーディオ符号化デバイス２０は次いで再順序付けユニット３４を呼び出すことができ、再順序付けユニット３４は、上で説明されたように、再順序付けられた変換されたＨＯＡ係数３３’／３５’（または言い換えれば、ＵＳ［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ＨＯＡ係数（これはやはり、ＳＶＤの文脈では、ＵＳ［ｋ］ベクトル３３とＶ［ｋ］ベクトル３５とを指し得る）を再順序付けることができる（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、上で説明されたように、フォアグラウンドチャネルの総数（ｎＦＧ）４５と、バックグラウンド音場の次数（N_{BG}）と、送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）およびインデックス（ｉ）（図４の例ではバックグラウンドチャネル情報４３としてまとめて示され得る）とを決定するために、ＨＯＡ係数１１および／または変換されたＨＯＡ係数３３／３５に関して音場分析を実行することができる（１０９）。

【０４９６】

[0496]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ネル情報４３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、再順序付けられたＵＳ［ｋ］ベクトル３３’および再順序付けられたＶ［ｋ］ベクトル３５’のうちで音場のフォアグラウンド成分または明瞭な成分を表すものを、ｎＦＧ４５（これらのフォアグラウンドベクトルを識別する１つまたは複数のインデックスを表し得る）に基づいて選択することができる（１１２）。

【０４９７】

[0497]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ＨＯＡチャネルの様々なチャ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ＨＯＡ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ＨＯＡ係数４７’を生成することができる。

【０４９８】

[0498]オーディオ符号化デバイス２０はまた、次いで、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、再順序付けられた変換されたＨＯＡ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、バックグラウンドチャネル情報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ことができる（１１８）。

【０４９９】

[0499]オーディオ符号化デバイス２０は次いで、上で説明された方式で、低減されたフォアグラウンドＶ［ｋ］ベクトル５５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５７を生成するために、量子化ユニット５２を呼び出すことができる（１２０）。

【０５００】

[0500]オーディオ符号化デバイス２０はまた、音響心理学的オーディオコーダユニット４０を呼び出すことができる。音響心理学的オーディオコーダユニット４０は、符号化された環境ＨＯＡ係数５９と符号化されたｎＦＧ信号６１とを生成するために、エネルギー補償された環境ＨＯＡ係数４７’および補間されたｎＦＧ信号４９’の各ベクトルを音響心理学的に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４２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４２は、コーディングされたフォアグラウンド指向性情報５７と、コーディングされた環境ＨＯＡ係数５９と、コーディングされたｎＦＧ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。

【０５０１】

[0501]図８は、本開示において説明される技法の様々な態様を実行する際の、図５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がビットストリーム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。議論を目的に、ベクトルベースの再構築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、上で述べられた情報を取り出すためにこのビットストリームを解析し、この情報をベクトルベースの再構築ユニット９２に渡すことができる。

【０５０２】

[0502]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（やはり、コーディングされたフォアグラウンドＶ［ｋ］ベクトル５７とも呼ばれ得る）と、コーディングされた環境ＨＯＡ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ｎＦＧ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とも呼ばれ得る）とを、上で説明された方式でビットストリーム２１から抽出することができる（１３２）。

【０５０３】

[0503]オーディオ復号デバイス２４はさらに、量子化ユニット７４を呼び出すことができる。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して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、音響心理学的復号ユニット８０を呼び出すことができる。音響心理学的オーディオコーディングユニット８０は、エネルギー補償された環境ＨＯＡ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ＨＯＡ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）。音響心理学的復号ユニット８０は、エネルギー補償された環境ＨＯＡ係数４７’をＨＯＡ係数定式化ユニット８２に渡し、ｎＦＧ信号４９’を再順序付けユニット８４に渡すことができる。

【０５０４】

[0504]再順序付けユニット８４は、ＨＯＡ係数１１のフォアグラウンド成分の元の次数を示すシンタックス要素を受信することができる。再順序付けユニット８４は、これらの再順序付けシンタックス要素に基づいて、再順序付けられたｎＦＧ信号４９’’と再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_k’とを生成するために、補間されたｎＦＧ信号４９’と低減されたフォアグラウンドＶ［ｋ］ベクトル５５_kとを再順序付けることができる（１４０）。再順序付けユニット８４は、再順序付けられたｎＦＧ信号４９’’をフォアグラウンド定式化ユニット７８に出力し、再順序付けられたフォアグラウンドＶ［ｋ］ベクトル５５_k’を空間－時間的補間ユニット７６に出力することができる。

【０５０５】

[0505]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、再順序付けられたフォアグラウンド指向性情報55_k’を受信し、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（１４２）。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’’をフォアグラウンド定式化ユニット７１８に転送することができる。

【０５０６】

[0506]オーディオ復号デバイス２４は、フォアグラウンド定式化ユニット７８を呼び出すことができる。フォアグラウンド定式化ユニット７８は、フォアグラウンドＨＯＡ係数６５を取得するために、補間されたフォアグラウンド指向性情報５５ｋ’’と補間されたフォアグラウンド信号４９’’との行列乗算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、ＨＯＡ係数定式化ユニット８２を呼び出すことができる。ＨＯＡ係数定式化ユニット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を環境ＨＯＡチャネル４７’に加算することができる（１４６）。

【０５０７】

[0507]図９Ａ～図９Ｌは、図４の例のオーディオ符号化デバイス２０の様々な態様をより詳細に示すブロック図である。図９Ａは、オーディオ符号化デバイス２０のＬＩＴユニット３０をより詳細に示すブロック図である。図９Ａの例に示されるように、ＬＩＴユニット３０は、複数の様々な線形可逆変換２００～２００Ｎを含み得る。ＬＩＴユニット３０は、いくつか例を挙げると、特異値分解（ＳＶＤ）変換２００Ａ（「ＳＶＤ２００Ａ」）、主要成分分析（ＰＣＡ：principle component analysis）変換２００Ｂ（「ＰＣＡ２００Ｂ」）、カルーネンレーベ変換（ＫＬＴ）２００Ｃ（「ＫＬＴ２００Ｃ」）、高速フーリエ変換（ＦＦＴ）２００Ｄ（「ＦＦＴ２００Ｄ」）、および離散コサイン変換（ＤＣＴ）２００Ｎ（「ＤＣＴ２００Ｎ」）を含み得る。ＬＩＴユニット３００は、ＨＯＡ係数１１に関してそれぞれの変換を適用し、それぞれの変換されたＨＯＡ係数３３／３５を生成するために、これらの線形可逆変換２００のいずれか１つを呼び出すことができる。

【０５０８】

[0508]ＨＯＡ係数１１に関して直接実行されるものとして説明されるが、ＬＩＴユニット３０は、線形可逆変換２００をＨＯＡ係数１１の派生物に適用することができる。たとえば、ＬＩＴユニット３００は、ＨＯＡ係数１１から導出された電力スペクトル密度行列に関してＳＶＤ２００を適用することができる。電力スペクトル密度行列は、ＰＳＤ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、ＨＯＡ係数１１のフレームを指す。

【０５０９】

[0509]ＬＩＴユニット３０は、ＳＶＤ２００（ｓｖｄ）をＰＳＤに適用した後で、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、すなわち、ＬＩＴユニット３０（または代替的に、一例としてＳＶＤユニット２００）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。ＳＶＤ２００は、いくつかの例では、量子化されたＶ［ｋ］行列（Ｖ［ｋ］’行列として示され得る）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。ＬＩＴユニット３０は、ＳＶ［ｋ］’行列を得るために、Ｓ［ｋ］行列を量子化されたＶ［ｋ］’行列とまず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。ＬＩＴユニット３０は次に、ＳＶ［ｋ］’行列の擬似逆行列（ｐｉｎｖ）を取得することができ、次いで、Ｕ［ｋ］行列を取得するためにＳＶ［ｋ］’行列の擬似逆行列とＨＯＡ係数１１を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');
```

【０５１０】

[0510]ＨＯＡ係数自体ではなくＨＯＡ係数の電力スペクトル密度（ＰＳＤ）に関してＳＶＤを実行することによって、ＬＩＴユニット３０は場合によっては、プロセッササイクルと記憶空間の１つまたは複数に関してＳＶＤを実行することの計算的な複雑さを低減しつつ、ＳＶＤがＨＯＡ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の符号化効率を達成することができる。すなわち、上で説明されたＰＳＤタイプのＳＶＤは場合によっては、ＳＶＤがＦ＊Ｆの行列（ＦはＨＯＡ係数の数）に対して行われるので、計算的により負荷が軽いことがある。Ｍはフレーム長であるを伴うＭ＊Ｆの行列と比較されると、すなわち１０２４個以上のサンプル。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＊Ｌ＾２）と比べて、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の表記である）。

【０５１１】

[0511]図９Ｂは、オーディオ符号化デバイス０のパラメータ計算ユニット３２をより詳細に示すブロック図である。パラメータ計算ユニット３２は、エネルギー分析ユニット２０２と相互相関付けユニット２０４とを含み得る。エネルギー分析ユニット２０２は、現在のフレーム（ｋ）または以前のフレーム（ｋ－１）の１つまたは複数のための、相関付けパラメータ（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）の１つまたは複数を生成するために、ＵＳ［ｋ］ベクトル３３およびＶ［ｋ］ベクトル３５の１つまたは複数に関して上で説明されたエネルギー分析を実行することができる。同様に、相互相関付けユニット２０４は、現在のフレーム（ｋ）または以前のフレーム（ｋ－１）の１つまたは複数のための、相関付けパラメータ（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）の１つまたは複数を生成するために、ＵＳ［ｋ］ベクトル３３およびＶ［ｋ］ベクトル３５の１つまたは複数に関して上で説明された相互相関付けを実行することができる。パラメータ計算ユニット３２は、現在のフレームパラメータ３７と以前のフレームパラメータ３９とを出力することができる。

【０５１２】

[0512]図９Ｃは、オーディオ符号化デバイス２０の再順序付けユニット３４をより詳細に示すブロック図である。再順序付けユニット３４は、パラメータ評価ユニット２０６とベクトル再順序付けユニット２０８とを含む。パラメータ評価ユニット２０６は、再順序付けインデックス２０５を生成するために、上で説明された方式で以前のフレームパラメータ３９と現在のフレームパラメータ３７とを評価するように構成されるユニットを表す。再順序付けインデックス２０５は、ＵＳ［ｋ］ベクトル３３のベクトルおよびＶ［ｋ］ベクトル３５のベクトルがどのようにレンダリングされるべきかを（たとえば、現在のベクトルロケーションのインデックスを識別するペアの第１のインデックスとベクトルの再順序付けられたロケーションを識別するペアの第２のインデックスとを伴うインデックスペアによって）識別する、インデックスを含む。ベクトル再順序付けユニット２０８は、再順序付けインデックス２０５に従って、ＵＳ［ｋ］ベクトル３３とＶ［ｋ］ベクトル３５とを再順序付けるように構成されるユニットを表す。再順序付けユニット３４は、再順序付けられたＵＳ［ｋ］ベクトル３３’と再順序付けられたＶ［ｋ］ベクトル３５’とを出力しながら、再順序付けインデックス２０５を１つまたは複数のシンタックス要素としてビットストリーム生成ユニット４２に渡すこともできる。

【０５１３】

[0513]図９Ｄは、オーディオ符号化デバイス２０の音場分析ユニット４４をより詳細に示すブロック図である。図９Ｄの例に示されるように、音場分析ユニット４４は、特異値分析ユニット２１０Ａと、エネルギー分析ユニット２１０Ｂと、空間分析ユニット２１０Ｃと、空間マスキング分析ユニット２１０Ｄと、拡散分析ユニット２１０Ｅと、指向性分析ユニット２１０Ｆとを含み得る。特異値分析ユニット２１０Ａは、（ＵＳ［ｋ］ベクト

ル３３の一部を形成する）Ｓベクトルの降順の対角線上の値によって作り出される曲線の傾きを分析するように構成されるユニットを表すことができ、ここで、上で説明されたように、大きい特異値はフォアグラウンド音声または明瞭な音声を表し、小さい特異値は音場のバックグラウンド成分を表す。エネルギー分析ユニット２１０Ｂは、ベクトルごとにＶ［ｋ］ベクトル３５のエネルギーを決定するように構成されるユニットを表し得る。

【０５１４】

[0514]空間分析ユニット２１０Ｃは、空間領域へのＨＯＡ係数１１の変換および保存されるべき音場の指向性成分を表す高エネルギーの領域の識別を通じて、上で説明された空間エネルギー分析を実行するように構成されるユニットを表し得る。空間マスキング分析ユニット２１０Ｄは、空間マスキング分析ユニット２１０Ｄが空間的に近接するより高エネルギーの音によってマスキングされる空間領域を識別できることを除き、空間エネルギー分析と同様の方式で空間マスキング分析を実行するように構成されるユニットを表し得る。拡散分析ユニット２１０Ｅは、音場のバックグラウンド成分を表し得る拡散エネルギーの領域を識別するために、ＨＯＡ係数１１に関して上で説明された拡散分析を実行するように構成されるユニットを表し得る。指向性分析ユニット２１０Ｆは、ＶＳ［ｋ］ベクトルを計算することと、指向性指数を識別するためにこれらのＶＳ［ｋ］ベクトルの各々の各成分を二乗して加算することとを伴う、上で述べられた指向性分析を実行するように構成されるユニットを表し得る。指向性分析ユニット２１０Ｆは、ＶＳ［ｋ］ベクトルの各々のこの指向性指数を、バックグラウンド／フォアグラウンド（ＢＧ／ＦＧ）識別（ＩＤ）ユニット２１２に提供することができる。

【０５１５】

[0515]音場分析ユニット４４はまた、ＢＧ／ＦＧ　ＩＤユニット２１２を含んでよく、これは、フォアグラウンドチャネルの総数（ｎＦＧ）４５、バックグラウンド音場の次数（N_{BG}）、および送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）とインデックス（ｉ）（これらは図４の例ではバックグラウンドチャネル情報４３としてまとめて示され得る）を、分析ユニット２１０～２１０Ｆの任意の組合せによって出力される分析の任意の組合せに基づいて決定するように構成されるユニットを表し得る。ＢＧ／ＦＧ　ＩＤユニット２１２は、目標ビットレート４１を達成するために、ｎＦＧ４５とバックグラウンドチャネル情報４３とを決定することができる。

【０５１６】

[0516]図９Ｅは、オーディオ符号化デバイス２０のフォアグラウンド選択ユニット３６をより詳細に示すブロック図である。フォアグラウンド選択ユニット３６は、再順序付けられたＵＳ［ｋ］ベクトル３３’および再順序付けられたＶ［ｋ］ベクトル３５’からのｎＦＧシンタックス要素４５によって識別されるフォアグラウンドＵＳ［ｋ］ベクトル４９とフォアグラウンドＶ［ｋ］ベクトル51_kとを解析し、または別様に抽出することができる、ベクトル解析ユニット２１４を含む。ベクトル解析ユニット２１４は、音場分析ユニット４４によって識別されｎＦＧシンタックス要素４５（フォアグラウンドチャネル情報４５とも呼ばれ得る）によって指定される音場のフォアグラウンド成分を表す様々なベクトルを解析することができる。図９Ｅの例に示されるように、ベクトル解析ユニット２１４は、いくつかの例では、音場のフォアグラウンド成分を表すために、フォアグラウンドＵＳ［ｋ］ベクトル４９およびフォアグラウンドＶ［ｋ］ベクトル51_kの中の非連続的なベクトルを選択することができる。その上、ベクトル解析ユニット２１４は、いくつかの例では、音場のフォアグラウンド成分を表すために、フォアグラウンドＵＳ［ｋ］ベクトル４９およびフォアグラウンドＶ［ｋ］ベクトル51_kの同じベクトル（位置ごとの）を選択することができる。

【０５１７】

[0517]図９Ｆは、オーディオ符号化デバイス２０のバックグラウンド選択ユニット４８をより詳細に示すブロック図である。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ネル情報（たとえば、バックグラウンド音場（N_{BG}）および送信すべき追加のＢＧ　ＨＯＡチャネルの数（ｎＢＧａ）とインデックス（ｉ））に基づいて、バックグラ

ウンドまたは環境ＨＯＡ係数４７を決定することができる。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１つによって識別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、ここで、ｎＢＧａは、図５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ＢＧ　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ことになる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：$M\times[(N_{BG}+1)^2+nBGa]$を有し得る。

【０５１８】

[0518]図９Ｇは、オーディオ符号化デバイス２０のエネルギー補償ユニット３８をより詳細に示すブロック図である。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ＨＯＡチャネルの様々なチャ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ＨＯＡ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るユニットを表し得る。エネルギー補償ユニット３８は、エネルギー決定ユニット２１８と、エネルギー分析ユニット２２０と、エネルギー増幅ユニット２２２とを含み得る。

【０５１９】

[0519]エネルギー決定ユニット２１８は、再順序付けられたＵＳ［ｋ］行列３３’と再順序付けられたＶ［ｋ］ベクトル３５’の１つまたは複数の各行および／または各列のＲＭＳを識別するように構成されるユニットを表し得る。エネルギー決定ユニット３８はまた、選択されたフォアグラウンドチャネルの１つまたは複数の各行および／または各列のＲＭＳを識別することができ、これは、ｎＦＧ信号４９と、フォアグラウンドＶ［ｋ］ベクトル51_kと、次数低減された環境ＨＯＡ係数４７とを含み得る。再順序付けられたＵＳ［ｋ］行列３３’および再順序付けられたＶ［ｋ］行列３５’の１つまたは複数の各行および／または各列のＲＭＳは、RMS_{FULL}と示されるベクトルに記憶され得るが、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数の各行および／または各列のＲＭＳは、$RMS_{REDUCED}$と示されるベクトルに記憶され得る。

【０５２０】

[0520]いくつかの例では、再順序付けられたＵＳ［ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数のそれぞれの行および／または列の各ＲＭＳを決定するために、エネルギー決定ユニット２１８はまず、基準球面調和係数（ＳＨＣ）レンダラを列に適用することができる。エネルギー決定ユニット２１８による基準ＳＨＣレンダラの適用は、再順序付けられたＵＳ［ｋ］行列３３’、再順序付けられたＶ［ｋ］行列３５’、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数の行および／または列によって表されるフレームの各行および／または各列によって記述される音場全体のエネルギーを決定するための、ＳＨＣ領域におけるＲＭＳの決定を可能にする。エネルギー決定ユニット３８は、このRMS_{FULL}ベクトルと$RMS_{REDUCED}$ベクトルとをエネルギー分析ユニット２２０に渡すことができる。

【０５２１】

[0521]エネルギー分析ユニット２２０は、次の式$Z=RMS_{FULL}/RMS_{REDUCED}$に従って、増幅値ベクトルＺを計算するように構成されるユニットを表し得る。エネルギー分析ユニット２２０は次いで、この増幅値ベクトルＺをエネルギー増幅ユニット２２２に渡すことができる。エネルギー増幅ユニット２２２は、ｎＦＧ信号４９、フォアグラウンドＶ［ｋ］ベクトル51_k、および次数低減された環境ＨＯＡ係数４７の１つまたは複数に、この増幅値ベクトルＺまたはその様々な部分を適用するように構成されるユニットを表

し得る。いくつかの例では、増幅値ベクトルＺは、次の式$HOA_{BG-RED}' = HOA_{BG-RED} Z^{T}$によって、次数低減された環境ＨＯＡ係数４７のみに適用され、ここでHOA_{BG-RED}は次数低減された環境ＨＯＡ係数４７を示し、HOA_{BG-RED}'はエネルギー補償され低減された環境ＨＯＡ係数４７’を示し、Z^{T}はＺベクトルの転置を示す。

【０５２２】

[0522]図９Ｈは、図４の例に示されるオーディオ符号化デバイス２０の空間－時間的補間ユニット５０をより詳細に示すブロック図である。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るユニットを表し得る。空間－時間的補間ユニット５０は、Ｖ補間ユニット２２４とフォアグラウンド適応ユニット２２６とを含み得る。

【０５２３】

[0523]Ｖ補間ユニット２２４は、現在のフォアグラウンドＶ［ｋ］ベクトル51_{k}および以前のフォアグラウンドＶ［ｋ－１］ベクトル51_{k-1}の残りの部分に基づいて、補間すべき現在のフォアグラウンドＶ［ｋ］ベクトル51_{k}の部分を選択することができる。Ｖ補間ユニット２２４は、上で述べられたサブフレームの１つまたは複数、またはフレームごとに変化し得る１つだけの定義されていない部分として、その部分を選択することができる。Ｖ補間ユニット２２４は、いくつかの例では、補間すべき現在のフォアグラウンドＶ［ｋ］ベクトル51_{k}の１０２４個のサンプルのうちの単一の１２８サンプルの部分を選択することができる。Ｖ補間ユニット２２４は次いで、ベクトルを球面上へ（Ｔ設計行列などの射影行列を使用して）射影することによって、現在のフォアグラウンドＶ［ｋ］ベクトル51_{k}と以前のフォアグラウンドＶ［ｋ－１］ベクトル51_{k-1}の中のベクトルの各々を、別々の空間的な写像に変換することができる。Ｖ補間ユニット２２４は次いで、Ｖの中のベクトルを球面上の形状として解釈することができる。２５６サンプルの部分についてＶ行列を補間するために、Ｖ補間ユニット２２４は次いで、これらの空間的な形状を補間し、次いで、射影行列の逆行列を介してそれらを球面調和関数領域のベクトルに変換し戻すことができる。本開示の技法は、このようにして、Ｖ行列間での滑らかな移行を実現することができる。Ｖ補間ユニット２２４は次いで、フォアグラウンドＶ［ｋ］ベクトル51_{k}の補間された部分を除去するように修正された後で、フォアグラウンドＶ［ｋ］ベクトル51_{k}を表す残りのＶ［ｋ］ベクトル５３を生成することができる。Ｖ補間ユニット２２４は次いで、補間されたフォアグラウンドＶ［ｋ］ベクトル$51_{k}'$をｎＦＧ適応ユニット２２６に渡すことができる。

【０５２４】

[0524]補間するために単一の部分を選択するとき、Ｖ補間ユニット２２４は、補間の継続時間、言い換えれば時間を（サンプルの数の単位で）識別する、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４と示されるシンタックス要素を生成することができる。サブフレーム補間を実行するために単一の部分を選択するとき、Ｖ補間ユニット２２４はまた、実行される補間のタイプ（またはいくつかの例では、補間が実行されたかされなかったか）を識別し得る、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄ２５５と示される別のシンタックス要素を生成することができる。空間－時間的補間ユニット５０は、これらのシンタックス要素２５４および２５５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。

【０５２５】

[0525]ｎＦＧ適応ユニット２２６は、適応させられたｎＦＧ信号４９’を生成するように構成されるユニットを表し得る。ｎＦＧ適応ユニット２２６は、フォアグラウンドＶ［ｋ］ベクトル51_{k}によるｎＦＧ信号４９の乗算を通じてフォアグラウンドＨＯＡ係数をまず取得することによって、適応させられたｎＦＧ信号４９’を生成することができる。フォアグラウンドＨＯＡ係数を取得した後で、ｎＦＧ適応ユニット２２６は、適応させられたｎＦＧ信号４９’（これらの信号は、補間されたフォアグラウンドＶ［ｋ］ベクトル

５１ｋ’から導出されるとすると、補間されたｎＦＧ信号４９’とも呼ばれ得る）を取得するために、補間されたフォアグラウンドＶ［ｋ］ベクトル５３によってフォアグラウンドＨＯＡ係数を除算することができる。

【０５２６】

[0526]図９Ｉは、図４の例に示されるオーディオ符号化デバイス２０の係数低減ユニット４６をより詳細に示すブロック図である。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る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}－ｎＢＧａ］×ｎＦＧを有し得る。

【０５２７】

[0527]係数低減ユニット４６は係数最小化ユニット２２８を含んでよく、これは、（バックグラウンドチャネル情報４３によって識別されるような）バックグラウンドＨＯＡ係数４７の中の説明がつく任意の係数を除去することによって、残りのフォアグラウンドＶ［ｋ］ベクトル５３の各々のサイズを低減し、または別様に最小化するように構成されるユニットを表し得る。係数最小化ユニット２２８は、低減されたフォアグラウンドＶ［ｋ］ベクトル５５を取得するために、バックグラウンドチャネル情報４３によって識別される係数を除去することができる。

【０５２８】

[0528]図９Ｊは、図４の例に示されるオーディオ符号化デバイス２０の音響心理学的オーディオコーダユニット４０をより詳細に示すブロック図である。音響心理学的オーディオコーダユニット４０は、エネルギー補償されたバックグラウンドＨＯＡ係数４７’および補間されたｎＦＧ信号４９’に関して音響心理学的な符号化を実行するように構成されるユニットを表し得る。図９Ｈの例に示されるように、音響心理学的オーディオコーダユニット４０は、エネルギー補償されたバックグラウンドＨＯＡ係数４７’のチャネルの各々（ここで、この文脈におけるチャネルは、特定の次数／位数の球面基底関数に対応するフレーム中のサンプルのすべてに対する係数を指す）と、補間されたｎＦＧ信号４９’の各信号とをオーディオ符号化するために、音響心理学的オーディオ符号化器４０Ａ～４０Ｎの複数の具体物を呼び出すことができる。いくつかの例では、音響心理学的オーディオコーダユニット４０は、エネルギー補償されたバックグラウンドＨＯＡ係数４７’（またはｎＢＧａとインデックスの総数（ｉ）とを足したもの）の各チャネルと、ｎＢＧａと追加の環境ＨＯＡチャネルのインデックスの総数（ｉ）とｎＦＧとの合計に対する補間されたｎＦＧ信号４９’（またはｎＦＧ）の各信号とを別々に符号化するのに十分な数の、オーディオ符号化器４０Ａ～４０Ｎを実体化し、またはそうでなければ（ハードウェアで実装されるときは）含む。オーディオ符号化器４０Ａ～４０Ｎは、符号化されたバックグラウンドＨＯＡ係数５９と符号化されたｎＦＧ信号６１とを出力することができる。

【０５２９】

[0529]図９Ｋは、図４の例に示されるオーディオ符号化デバイス２０の量子化ユニット５２をより詳細に示すブロック図である。図９Ｋの例では、量子化ユニット５２は、一様量子化ユニット２３０と、ｎｂｉｔｓユニット２３２と、予測ユニット２３４と、予測モードユニット２３６（「予測モードユニット（Ｐｒｅｄ　Ｍｏｄｅ　Ｕｎｉｔ）２３６」）と、カテゴリおよび残差コーディングユニット２３８と、ハフマンテーブル選択ユニット２４０とを含む。一様量子化ユニット２３０は、空間成分の１つ（低減されたフォアグラウンドＶ［ｋ］ベクトル５５の任意の１つを表し得る）に関して上で説明された一様量子化を実行するように構成されるユニットを表す。ｎｂｉｔｓユニット２３２は、ｎｂｉｔｓパラメータまたはｎｂｉｔｓ値を決定するように構成されるユニットを表す。

【０５３０】

[0530]予測ユニット２３４は、量子化された空間成分に関して予測を実行するように構成されるユニットを表す。予測ユニット２３４は、低減されたフォアグラウンドＶ［ｋ］

ベクトル５５の時間的に後の対応する１つ（低減されたフォアグラウンドＶ［ｋ－１］ベクトル５５として示され得る）により、低減されたフォアグラウンドＶ［ｋ］ベクトル５５の現在のベクトルの要素ごとの減算を実行することによって、予測を実行することができる。この予測の結果は、予測された空間成分と呼ばれ得る。

【０５３１】

[0531]予測モードユニット２３６は、予測モードを選択するように構成されるユニットを表し得る。ハフマンテーブル選択ユニット２４０は、ｃｉｄのコーディングに適切なハフマンテーブルを選択するように構成されるユニットを表し得る。予測モードユニット２３６およびハフマンテーブル選択ユニット２４０は、一例として、次の擬似コードに従って動作することができる：

```
For a given nbits, retrieve all the Huffman Tables having nbits
B00 = 0; B01 = 0; B10 = 0; B11 = 0; // コーディングモードごとに予想されるビットを計算するために初期化する
  for m = 1:(# elements in the vector)
  // ベクトル要素ｖ（ｍ）のための期待されるビットの数を計算する
  // 予測なしでハフマンテーブル５を使用する
  B00 = B00 + calculate_bits(v(m), HT5);
    // 予測なしでハフマンテーブル｛１，２，３｝を使用する
  B01 = B01 + calculate_bits(v(m), HTq); q in {1,2,3}

  // 予測残差ｅ（ｍ）のための期待されるビットの数を計算する
  e(m) = v(m) - vp(m); // vp(m): previous frame vector element
  // 予測ありでハフマンテーブル４を使用する
  B10 = B10 + calculate_bits(e(m), HT4);
  // 予測ありでハフマンテーブル５を使用する
  B11 = B11 + calculate_bits(e(m), HT5);
end

// 最小ビットを生じる最良の予測モードとハフマンテーブルとを発見する
// 最良の予測モードおよびハフマンテーブルに、それぞれｐｆｌａｇおよびＨｔｆｌａｇによってフラグが付けられる
[Be, id] = min( [B00 B01 B10 B11] );
Switch id
  case 1: pflag = 0; HTflag = 0;
  case 2: pflag = 0; HTflag = 1;
  case 3: pflag = 1; HTflag = 0;
  case 4: pflag = 1; HTflag = 1;
end
```

【０５３２】

[0532]カテゴリおよび残差コーディングユニット２３８は、上でより詳細に説明された方式で、予測された空間成分または（予測が無効にされているときは）量子化された空間成分の分類と残差コーディングとを実行するように構成されるユニットを表し得る。

【０５３３】

[0533]図９Ｋの例に示されるように、量子化ユニット５２は、ビットストリーム２１とサイド情報（これ自体がビットストリーム２１とは別のビットストリームであり得る）のいずれかに含めるために、様々なパラメータまたは値を出力することができる。情報がサイドチャネル情報において指定されると仮定すると、スカラー／エントロピー量子化ユニット５０は、ｎｂｉｔｓ値２３３としてｎｂｉｔｓ値を、予測モード２３７として予測モードを、ハフマンテーブル情報２４１としてハフマンテーブル情報を、空間成分の圧縮されたバージョン（図４の例においてコーディングされたフォアグラウンドＶ［ｋ］ベクト

ル５７として示される）とともに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、空間成分の圧縮されたバージョンはこの例では、ｃｉｄと、符号ビットと、ブロックコーディングされた残差とを符号化するために選択されるハフマンコードを指し得る。ｎｂｉｔｓ値は、コーディングされたフォアグラウンドＶ［ｋ］ベクトル５７のすべてに対するサイドチャネル情報において一度指定され得るが、予測モードおよびハフマンテーブル情報は、コーディングされたフォアグラウンドＶ［ｋ］ベクトル５７の各々１つに対して指定され得る。空間成分の圧縮されたバージョンを指定するビットストリームの部分は、図１０Ｂおよび／または図１０Ｃの例においてさらに示される。

【０５３４】

[0534]図９Ｌは、図４の例に示されるオーディオ符号化デバイス２０のビットストリーム生成ユニット４２をより詳細に示すブロック図である。ビットストリーム生成ユニット４２は、メインチャネル情報生成ユニット２４２とサイドチャネル情報生成ユニット２４４とを含み得る。メインチャネル情報生成ユニット２４２は、再順序付けインデックス２０５、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅシンタックス要素２５４、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５、符号化されたバックグラウンドＨＯＡ係数５９、および符号化されたｎＦＧ信号６１のすべてではないとしても１つまたは複数を含む、メインビットストリーム２１を生成することができる。サイドチャネル情報生成ユニット２４４は、ｎｂｉｔｓ値２３３、予測モード２３７、ハフマンテーブル情報２４１、およびコーディングされたフォアグラウンドＶ［ｋ］ベクトル５７のすべてではないとしても１つまたは複数を含み得る、サイドチャネルビットストリーム２１Ｂを生成するように構成されるユニットを表し得る。ビットストリーム２１および２１Ｂは、ビットストリーム２１と総称的に呼ばれ得る。いくつかの文脈では、ビットストリーム２１はメインチャネルビットストリーム２１のみを指し得るが、ビットストリーム２１Ｂはサイドチャネル情報２１Ｂと呼ばれ得る。

【０５３５】

[0535]図１０Ａ～図１００（ｉｉ）は、圧縮された空間成分をより詳細に指定し得る、ビットストリームの部分またはサイドチャネル情報を示す図である。図１０Ａの例において、部分２５０は、レンダラ識別子（「レンダラＩＤ」）フィールド２５１とＨＯＡＤｅｃｏｄｅｒＣｏｎｆｉｇフィールド２５２とを含む。レンダラＩＤフィールド２５１は、ＨＯＡコンテンツのミキシングのために使用されたことのあるレンダラのＩＤを記憶するフィールドを表し得る。ＨＯＡＤｅｃｏｄｅｒＣｏｎｆｉｇフィールド２５２は、ＨＯＡ空間復号器を初期化するための情報を記憶するように構成されるフィールドを表し得る。

【０５３６】

[0536]ＨＯＡＤｅｃｏｄｅｒＣｏｎｆｉｇフィールド２５２はさらに、指向性情報（「ｄｉｒｅｃｔｉｏｎ　ｉｎｆｏ」）フィールド２５３と、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅフィールド２５４と、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄフィールド２５５と、ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈフィールド２５６と、利得情報フィールド２５７とを含む。指向性情報フィールド２５３は、指向性ベース合成復号器を構成するための情報を記憶するフィールドを表し得る。ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅフィールド２５４は、ベクトルベースの信号の空間－時間的な補間の時間を記憶するフィールドを表し得る。ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄフィールド２５５は、ベクトルベースの信号の空間－時間的な補間の間に適用される補間タイプのインジケーションを記憶するフィールドを表し得る。ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈフィールド２５６は、ベクトルベースの信号を合成するために使用される送信されたデータベクトルの長さを記憶するフィールドを表し得る。利得情報フィールド２５７は、信号に適用される利得修正を示す情報を記憶するフィールドを表す。

【０５３７】

[0537]図１０Ｂの例では、部分２５８Ａはサイド情報チャネルの一部分を表し、ここで部分２５８Ａはバイト数フィールド２６０とｎｂｉｔｓフィールド２６１とを含むフレー

ムヘッダ２５９を含む。バイト数フィールド２６０は、バイトアラインメント０フィールド２６４を含む空間成分ｖ１からｖｎを指定するためにフレームに含まれるバイトの数を表すためのフィールドを表し得る。ｎｂｉｔｓフィールド２６１は、空間成分ｖ１～ｖｎを解凍する際に使用するために識別されるｎｂｉｔｓ値を指定し得るフィールドを表す。

【０５３８】

[0538]図１０Ｂの例にさらに示されるように、部分２５８Ａはｖ１～ｖｎのためのサブビットストリームを含んでよく、サブビットストリームの各々が、予測モードフィールド２６２と、ハフマンテーブル情報フィールド２６３と、圧縮された空間成分ｖ１～ｖｎの対応する１つとを含み得る。予測モードフィールド２６２は、圧縮された空間成分ｖ１～ｖｎの対応する１つに関して予測が実行されたかどうかのインジケーションを記憶するためのフィールドを表し得る。ハフマンテーブル情報フィールド２６３は、どのハフマンテーブルが圧縮された空間成分ｖ１～ｖｎの対応する１つの様々な態様を復号するために使用されるべきかを少なくとも一部示すためのフィールドを表す。

【０５３９】

[0539]この点において、本技法は、オーディオ符号化デバイス２０が、音場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を取得することを可能にでき、空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０５４０】

[0540]図１０Ｃは、圧縮された空間成分を指定し得るサイドチャネル情報の部分２５８Ｂの代替的な例をより詳細に示す図である。図１０Ｃの例において、部分２５８Ｂは、Ｎｂｉｔｓフィールド２６１を含むフレームヘッダ２５９を含む。Ｎｂｉｔｓフィールド２６１は、空間成分ｖ１～ｖｎを解凍する際に使用するために識別されるｎｂｉｔｓ値を指定し得るフィールドを表す。

【０５４１】

[0541]図１０Ｃの例にさらに示されるように、部分２５８Ｂはｖ１～ｖｎのサブビットストリームを含んでよく、サブビットストリームの各々が、予測モードフィールド２６２と、ハフマンテーブル情報フィールド２６３と、圧縮された空間成分ｖ１～ｖｎの対応する１つとを含む。予測モードフィールド２６２は、圧縮された空間成分ｖ１～ｖｎの対応する１つに関して予測が実行されたかどうかのインジケーションを記憶するためのフィールドを表し得る。ハフマンテーブル情報フィールド２６３は、どのハフマンテーブルが圧縮された空間成分ｖ１～ｖｎの対応する１つの様々な態様を復号するために使用されるべきかを少なくとも一部示すためのフィールドを表す。

【０５４２】

[0542]示される例におけるＮｂｉｔｓフィールド２６１は、サブフィールドＡ２６５と、サブフィールドＢ２６６と、サブフィールドＣ２６７とを含む。この例では、Ａ２６５およびＢ２６６は各々１ビットのサブフィールドであるが、Ｃ２６７は２ビットのサブフィールドである。他の例は、異なるサイズのサブフィールド２６５と、２６６と、２６７とを含み得る。Ａフィールド２６５およびＢフィールド２６６は、Ｎｂｉｔｓフィールド２６１の最上位ビットと２番目の上位ビットとを記憶するフィールドを表し得るが、Ｃフィールド２６７は、Ｎｂｉｔｓフィールド２６１の最下位ビットを記憶するフィールドを表し得る。

【０５４３】

[0543]部分２５８Ｂはまた、ＡｄｄＡｍｂＨｏａＩｎｆｏＣｈａｎｎｅｌフィールド２６８を含み得る。ＡｄｄＡｍｂＨｏａＩｎｆｏＣｈａｎｎｅｌフィールド２６８は、追加の環境ＨＯＡ係数のための情報を記憶するフィールドを表し得る。図１０Ｃの例に示されるように、ＡｄｄＡｍｂＨｏａＩｎｆｏＣｈａｎｎｅｌ２６８は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘフィールド２４６と、ＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎフィールド２４７とを含む。ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘフィールド２４６は、追加の環境ＨＯＡ係数のインデックスを記憶するフィールドを表し得る。ＡｍｂＣｏｅｆｆＩｄｘ

Ｔｒａｎｓｉｔｉｏｎフィールド２４７は、フレームの中で追加の環境ＨＯＡ係数がフェードインされるかフェードアウトされるかを示すデータを記憶するように構成されるフィールドを表し得る。

【０５４４】

[0544]図１０Ｃ（ｉ）は、圧縮された空間成分を指定し得るサイドチャネル情報の部分２５８Ｂ’の代替的な例をより詳細に示す図である。図１０Ｃ（ｉ）の例では、部分２５８Ｂ’は、Ｎｂｉｔｓフィールド２６１を含むフレームヘッダ２５９を含む。Ｎｂｉｔｓフィールド２６１は、空間成分ｖ１～ｖｎを解凍する際の使用について識別されるｎｂｉｔｓ値を指定することができる。

【０５４５】

[0545]図１０Ｃ（ｉ）の例にさらに示されるように、部分２５８Ｂ’はｖ１～ｖｎのサブビットストリームを含んでよく、サブビットストリームの各々が、予測モードフィールド２６２を含まずに、ハフマンテーブル情報フィールド２６３と、圧縮された指向性成分ｖ１～ｖｎの対応する１つとを含む。すべての他の点において、部分２５８Ｂ’は部分２５８Ｂと同様であり得る。

【０５４６】

[0546]図１０Ｄは、ビットストリーム２１の部分２５８Ｃをより詳細に示す図である。部分２５８Ｃは、図１０Ｄの例に示されるように、フレームヘッダ２５９および０バイトアラインメント２６４が除去され、一方でＮｂｉｔｓ２６１フィールドがｖ１～ｖｎのビットストリームの各々の前に追加されたことを除き、部分２５８と同様である。

【０５４７】

[0547]図１０Ｄ（ｉ）は、ビットストリーム２１の部分２５８Ｃ’をより詳細に示す図である。部分２５８Ｃ’は、部分２５８Ｃ’がＶベクトルｖ１～ｖｎの各々のための予測モードフィールド２６２を含まないことを除き、部分２５８Ｃと同様である。

【０５４８】

[0548]図１０Ｅは、ビットストリーム２１の部分２５８Ｄをより詳細に示す図である。部分２５８Ｄは、図１０Ｅの例に示されるように、フレームヘッダ２５９および０バイトアラインメント２６４が除去され、一方でＮｂｉｔｓ２６１フィールドがｖ１～ｖｎのビットストリームの各々の前に追加されたことを除き、部分２５８と同様である。

【０５４９】

[0549]図１０Ｅ（ｉ）は、ビットストリーム２１の部分２５８Ｄ’をより詳細に示す図である。部分２５８Ｄ’は、部分２５８Ｄ’がＶベクトルｖ１～ｖｎの各々のための予測モードフィールド２６２を含まないことを除き、部分２５８Ｄと同様である。この点において、オーディオ符号化デバイス２０は、図１０Ｃ（ｉ）、図１０Ｄ（ｉ）、および図１０Ｅ（ｉ）の例に関して例証されたように、各々の圧縮されたＶベクトルのための予測モードフィールド２６２を含まないビットストリーム２１を生成することができる。

【０５５０】

[0550]図１０Ｆは、図１０Ａの例に示されるビットストリーム２１の部分２５０を異なる方式で示す図である。図１０Ｄの例に示される部分２５０は、ＨＯＡＯｒｄｅｒフィールド（説明を簡単にするために図１０Ｆの例では示されなかった）と、ＭｉｎＡｍｂＨｏａＯｒｄｅｒフィールド（やはり説明を簡単にするために図１０の例では示されなかった）と、指向性情報フィールド２５３と、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅフィールド２５４と、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄフィールド２５５と、ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈフィールド２５６と、利得情報フィールド２５７とを含む。図１０Ｆの例に示されるように、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅフィールド２５４は３ビットのフィールドを備えてよく、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄフィールド２５５は１ビットのフィールドを備えてよく、ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈフィールド２５６は２ビットのフィールドを備えてよい。

【０５５１】

[0551]図１０Ｇは、ビットストリーム２１の部分２４８をより詳細に示す図である。部分２４８は、ＨＯＡｆｒａｍｅフィールド２４９（サイドバンド情報、サイドチャネル情報、またはサイドチャネルビットストリームとしても示され得る）を含む、統合された音声／オーディオコーダ（ＵＳＡＣ：unified speech/audio coder）３次元（３Ｄ）ペイロードを表す。図１０Ｅの例に示されるように、ＨＯＡＦｒａｍｅフィールド２４９の拡大図は、図１０Ｃの例に示されるビットストリーム２１の部分２５８Ｂと同様であり得る。「ＣｈａｎｎｅｌＳｉｄｅＩｎｆｏＤａｔａ」は、説明を簡単にするために図１０Ｃの例には示されなかったＣｈａｎｎｅｌＴｙｐｅフィールド２６９と、図１０Ｅの例において「ｂａ」として示されるＡフィールド２６５と、図１０Ｅの例において「ｂｂ」として示されるＢフィールド２６６と、図１０Ｅの例において「ｕｎｉｔＣ」として示されるＣフィールド２６７とを含む。ＣｈａｎｎｅｌＴｙｐｅフィールドは、チャネルが指向性ベースの信号か、ベクトルベースの信号か、追加の環境ＨＯＡ係数かを示す。異なるＣｈａｎｎｅｌＳｉｄｅＩｎｆｏＤａｔａの間には、灰色で示される異なるＶベクトルビットストリーム（たとえば、「ｖ１のためのビットストリーム」および「ｖ２のためのビットストリーム」）を伴うＡｄｄＡｍｂＨｏａＩｎｆｏＣｈａｎｎｅｌフィールド２６８がある。

【０５５２】

[0552]図１０Ｈ～図１０Ｏ（ｉｉ）は、付随するＨＯＡｃｏｎｆｉｇ部分２５０Ｈ～２５０Ｏとともにビットストリーム２１の別の様々な例示的な部分２４８Ｈ～２４８Ｏをより詳細に示す図である。図１０Ｈ（ｉ）および図１０Ｈ（ｉｉ）は、上の擬似コードにおいてケース０に対応するように生成された、第１の例示的なビットストリーム２４８Ｈと付随するＨＯＡ　ｃｏｎｆｉｇ部分２５０Ｈとを示す。図１０Ｈ（ｉ）の例において、ＨＯＡｃｏｎｆｉｇ部分２５０Ｈは、Ｖベクトルのすべての要素、たとえばすべての１６個のＶベクトルの要素がコーディングされることを示すように設定された、ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｈはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｈはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｈはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、$(1+1)^2$すなわち４に等しいものとしてＭｉｎＮｕｍｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｈは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、$(N+1)^2$すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５５３】

[0553]図１０Ｈ（ｉ）の例にさらに示されるように、部分２４８Ｈは、スペクトル帯域複製（ＳＢＲ：spectral band replication）が有効にされるときに２つのオーディオフレームが１つの統合された音声およびオーディオコーディング（ＵＳＡＣ）３次元（ＵＳＡＣ－３Ｄ）フレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ａおよび２４９ＢがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓ

ＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５５４】

[0554]図１０Ｈ（ｉｉ）は、フレーム２４９Ａと２４９Ｂとをより詳細に示す。図１０Ｈ（ｉｉ）の例に示されるように、フレーム２４９Ａは、ＣｈａｎｎｅｌＳｉｄｅＩｎｆｏＤａｔａ（ＣＳＩＤ）フィールド１５４～１５４Ｃと、ＨＯＡＧａｉｎＣｏｒｒｅｃｔｉｏｎＤａｔａ（ＨＯＡＧＣＤ）フィールドと、ＶＶｅｃｔｏｒＤａｔａフィールド１５６および１５６Ｂと、ＨＯＡＰｒｅｄｉｃｔｉｏｎＩｎｆｏフィールドとを含む。ＣＳＩＤフィールド１５４は、ＣｈａｎｎｅｌＴｙｐｅ２６９とともにｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とを含み、これらの各々は、図１０Ｈ（ｉ）の例に示される対応する値０１、１、０、および０１に設定される。ＣＳＩＤフィールド１５４Ｂは、ＣｈａｎｎｅｌＴｙｐｅ２６９とともにｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とを含み、これらの各々は、図１０Ｈ（ｉｉ）の例に示される対応する値０１、１、０、および０１に設定される。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。ＣＳＩＤフィールド１５４～１５４Ｃの各々は、トランスポートチャネル１、２、および３のそれぞれ１つに対応する。事実上、各ＣＳＩＤフィールド１５４～１５４Ｃは、対応するペイロード１５６および１５６Ｂが指向性ベースの信号か（対応するＣｈａｎｎｅｌＴｙｐｅが０に等しいとき）、ベクトルベースの信号か（対応するＣｈａｎｎｅｌＴｙｐｅが１に等しいとき）、追加の環境ＨＯＡ係数か（対応するＣｈａｎｎｅｌＴｙｐｅが２に等しいとき）、空か（ＣｈａｎｎｅｌＴｙｐｅが３に等しいとき）を示す。

【０５５５】

[0555]図１０Ｈ（ｉｉ）の例において、フレーム２４９Ａは、（ＣＳＩＤフィールド１５４および１５４ＢにおいてＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）２つのベクトルベースの信号と（ＣＳＩＤフィールド１５４ＣにおいてＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｈを仮定すると、オーディオ復号デバイス２４は、すべての１６個のＶベクトルの要素が符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６および１５６Ｂは各々、各々が８ビットによって一様に量子化されるすべての１６個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、単一のアスタリスク（*）によって示されるように、コーディングスキームは、対応するトランスポートチャネルのためのＣＳＩＤフィールドにおいて、ＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５５６】

[0556]フレーム２４９Ｂにおいて、ＣＳＩＤフィールド１５４および１５４Ｂはフレーム２４９の中のそれらと同じであるが、フレーム２４９ＢのＣＳＩＤフィールド１５４Ｃは１というＣｈａｎｎｅｌＴｙｐｅに切り替えられる。フレーム２４９ＢのＣＳＩＤフィールド１５４Ｃはしたがって、ＣｂＦｌａｇ２６７と、Ｐｆｌａｇ２６７（ハフマン符号化を示す）と、Ｎｂｉｔｓ２６１（１２に等しい）とを含む。結果として、フレーム２４９Ｂは、各々が１２ビットによって一様に量子化されハフマンコーディングされた、１６個のＶベクトル要素を含む第３のＶＶｅｃｔｏｒＤａｔａフィールド１５６Ｃを含む。上で述べられたように、コーディングされたＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定されるが、ハフマンコーディングスキームは、この特定のトランスポートチャネル（たとえば、トランスポートチャネル第３）のためのＣＳＩＤフィールド１５４Ｃの中のＮｂｉｔｓＱ＝１２、ＣｂＦｌａｇ＝０、およびＰｆｌａｇ＝０によってシグナリングされる。

【０５５７】

[0557]図１０Ｉ（ｉ）および図１０Ｉ（ｉｉ）の例は、上の擬似コードにおいてケース０に対応するように生成された、第２の例示的なビットストリーム２４８Ｉと付随するＨＯＡ　ｃｏｎｆｉｇ部分２５０Ｉとを示す。図１０Ｉ（ｉ）の例では、ＨＯＡｃｏｎｆｉ

ｇ部分２５０Ｉは、Ｖベクトルのすべての要素、たとえばすべての１６個のＶベクトル要素がコーディングされることを示すように設定された、ＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｉはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｉはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。

【０５５８】

[0558]ＨＯＡｃｏｎｆｉｇ部分２５０Ｉはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、$(1+1)^2$すなわち４に等しいものとしてＭｉｎＮｕｍｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。オーディオ復号デバイス２４はまた、ＮｕｍＯｆＨｏａＣｏｅｆｆｓシンタックス要素とＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡとの間の差に設定されるものとしてＭａｘＮｏｏｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓシンタックス要素を導出することができ、これはこの例では１６－４すなわち１２に等しいと仮定される。オーディオ復号デバイス２４はまた、ｃｅｉｌ（ｌｏｇ２（ＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓ））＝ｃｅｉｌ（ｌｏｇ２（１２））＝４に設定されるものとしてＡｍｂＡｓｉｇｎｍＢｉｔｓシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｈは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、$(N+1)^2$すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５５９】

[0559]図１０Ｉ（ｉ）の例にさらに示されるように、部分２４８Ｈは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｃおよび２４９ＤがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５６０】

[0560]図１０Ｉ（ｉｉ）は、フレーム２４９Ｃと２４９Ｄとをより詳細に示す。図１０Ｉ（ｉｉ）の例に示されるように、フレーム２４９Ｃは、ＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６とを含む。ＣＳＩＤフィールド１５４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６と、ＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７（ここで、ダブルアスタリスク（**）は、フレキシブルトランスポートチャネル第１に対して、復号器の内部状態がここではＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎＳｔａｔｅ＝２であると仮定されることを示し、このことは、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘビットフィールドがシグナリングされる、またはビットストリームにおいて別様に指定されるをもたらす）と、ＣｈａｎｎｅｌＴｙｐｅ２６９（これは２に等しく、対応するペイロードが追加の環境ＨＯＡ係数であることをシグナリングする）とを含む。オーディオ復号デバイス２４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ＋１＋ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡすなわちこの例では５に等しいものとしてＡｍｂＣｏｅｆｆＩｄｘを導出することができる。ＣＳＩＤフィールド１５４Ｂは、図１０Ｉ

（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。

【０５６１】

[0561]図１０Ｉ（ｉｉ）の例では、フレーム２４９Ｃは、（ＣＳＩＤフィールド１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）単一のベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｉを仮定すると、オーディオ復号デバイス２４は、すべての１６個のＶベクトル要素が符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６は、各々が８ビットによって一様に量子化される、すべての１６個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、脚注２によって示されるように、コーディングスキームは、対応するトランスポートチャネルのＣＳＩＤフィールドにおいてＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５６２】

[0562]フレーム２４９Ｄにおいて、ＣＳＩＤフィールド１５４は、移行が起こらなかったことを示すＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７を含むので、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６は以前のフレームから推測されることが可能であり、再びシグナリングされること、または別様に指定されることは必要ではない。フレーム２４９ＤのＣＳＩＤフィールド１５４Ｂおよび１５４Ｃは、フレーム２４９Ｃのそれらと同じであるので、フレーム２４９Ｃのように、フレーム２４９Ｄは単一のＶＶｅｃｔｏｒＤａｔａフィールド１５６を含み、これは、各々が８ビットによって一様に量子化されるすべての１６個のベクトル要素を含む。

【０５６３】

[0563]図１０Ｊ（ｉ）および図１０Ｊ（ｉｉ）は、上の擬似コードにおけるケース１に対応するように生成された、第１の例示的なビットストリーム２４８Ｊと付随するＨＯＡｃｏｎｆｉｇ部分２５０Ｊとを示す。図１０Ｊ（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５０Ｊは、要素１からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素まで、およびＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素において指定される要素（この例では０であると仮定される）を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｊはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｊはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｊはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、$(1+1)^2$すなわち４に等しいものとしてＭｉｎＮｕｍｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｊは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、$(N+1)^2$すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５６４】

[0564]図１０Ｊ（ｉ）の例にさらに示されるように、部分２４８Ｊは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｅおよび２４９ＦがＵＳＡ

Ｃ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５６５】

[0565]図１０Ｊ（ｉｉ）は、フレーム２４９Ｅと２４９Ｆとをより詳細に示す。図１０Ｊ（ｉｉ）の例に示されるように、フレーム２４９Ｅは、ＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６および１５６Ｂとを含む。ＣＳＩＤフィールド１５４は、図１０Ｊ（ｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｂは、図１０Ｊ（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。ＣＳＩＤフィールド１５４～１５４Ｃの各々は、トランスポートチャネル１、２、および３のそれぞれ１つに対応する。

【０５６６】

[0566]図１０Ｊ（ｉｉ）の例では、フレーム２４９Ｅは、（ＣＳＩＤフィールド１５４および１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）２つのベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｈを仮定すると、オーディオ復号デバイス２４は、すべての１２個のＶベクトル要素が符号化されると決定することができる（ここで、１２は、（ＨＯＡＯｒｄｅｒ＋１）^{2}－（ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡ）－（ＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎ）＝１６－４－０＝１２として導出される）。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６および１５６Ｂは各々、各々が８ビットによって一様に量子化されるすべての１２個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、単一のアスタリスク（*）によって示されるように、コーディングスキームは、対応するトランスポートチャネルのためのＣＳＩＤフィールドにおいて、ＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５６７】

[0567]フレーム２４９Ｆにおいて、ＣＳＩＤフィールド１５４および１５４Ｂはフレーム２４９Ｅの中のそれらと同じであるが、フレーム２４９ＦのＣＳＩＤフィールド１５４Ｃは１というＣｈａｎｎｅｌＴｙｐｅに切り替えられる。フレーム２４９ＢのＣＳＩＤフィールド１５４Ｃはしたがって、ＣｂＦｌａｇ２６７と、Ｐｆｌａｇ２６７（ハフマン符号化を示す）と、Ｎｂｉｔｓ２６１（１２に等しい）とを含む。結果として、フレーム２４９Ｆは、各々が１２ビットによって一様に量子化されハフマンコーディングされる、１２個のＶベクトル要素を含む第３のＶＶｅｃｔｏｒＤａｔａフィールド１５６Ｃを含む。上で述べられたように、コーディングされたＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定されるが、ハフマンコーディングスキームは、この特定のトランスポートチャネル（たとえば、トランスポートチャネル第３）のためのＣＳＩＤフィールド１５４Ｃの中のＮｂｉｔｓＱ＝１２、ＣｂＦｌａｇ＝０、およびＰｆｌａｇ＝０によってシグナリングされる。

【０５６８】

[0568]図１０Ｋ（ｉ）および図１０Ｋ（ｉｉ）の例は、上の擬似コードにおけるケース１に対応するように生成された、第２の例示的なビットストリーム２４８Ｋと付随するＨＯＡ ｃｏｎｆｉｇ部分２５０Ｋとを示す。図１０Ｋ（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５０Ｋは、要素１からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素まで、およびＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素において指定される要素（この例では０であると仮定される）を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｋはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｋはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。

【０５６９】

[0569]ＨＯＡｃｏｎｆｉｇ部分２５０Ｋはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、$(1+1)^2$すなわち４に等しいものとしてＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。オーディオ復号デバイス２４はまた、ＮｕｍＯｆＨｏａＣｏｅｆｆｓシンタックス要素とＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡとの間の差に設定されるものとしてＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓシンタックス要素を導出することができ、これはこの例では１６－４すなわち１２に等しいと仮定される。オーディオ復号デバイス２４はまた、ｃｅｉｌ（ｌｏｇ２（ＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓ））＝ｃｅｉｌ（ｌｏｇ２（１２））＝４に設定されるものとしてＡｍｂＡｓｉｇｎｍＢｉｔｓシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｋは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、$(N+1)^2$すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５７０】

[0570]図１０Ｋ（ｉ）の例にさらに示されるように、部分２４８Ｋは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｇおよび２４９ＨがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５７１】

[0571]図１０Ｋ（ｉｉ）は、フレーム２４９Ｇと２４９Ｈとをより詳細に示す。図１０Ｋ（ｉｉ）の例に示されるように、フレーム２４９ＧはＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６とを含む。ＣＳＩＤフィールド１５４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６と、ＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７（ここで、ダブルアスタリスク（**）は、フレキシブルトランスポートチャネル第１に対して、復号器の内部状態がここではＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎＳｔａｔｅ＝２であると仮定されることを示し、このことは、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘビットフィールドがシグナリングされる、またはビットストリームにおい

て別様に指定されるをもたらす）と、ＣｈａｎｎｅｌＴｙｐｅ２６９（これは2に等しく、対応するペイロードが追加の環境ＨＯＡ係数であることをシグナリングする）とを含む。オーディオ復号デバイス２４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ＋１＋ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡすなわちこの例では５に等しいものとしてＡｍｂＣｏｅｆｆＩｄｘを導出することができる。ＣＳＩＤフィールド１５４Ｂは、図１０Ｋ（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。

【０５７２】

[0572]図１０Ｋ（ｉｉ）の例では、フレーム２４９Ｇは、（ＣＳＩＤフィールド１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）単一のベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｋを仮定すると、オーディオ復号デバイス２４は、１１個のＶベクトル要素が符号化されると決定することができる（ここで、１２は、（ＨＯＡＯｒｄｅｒ＋１）^{2}－（ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡ）－（ＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎ）＝１６－４－１＝１１として導出される）。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６は、各々が８ビットによって一様に量子化される、すべての１１個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、脚注２によって示されるように、コーディングスキームは、対応するトランスポートチャネルのＣＳＩＤフィールドにおいてＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５７３】

[0573]フレーム２４９Ｈにおいて、ＣＳＩＤフィールド１５４は、移行が起こらなかったことを示すＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７を含むので、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６は以前のフレームから推測されることが可能であり、再びシグナリングされること、または別様に指定されることは必要ではない。フレーム２４９ＨのＣＳＩＤフィールド１５４Ｂおよび１５４Ｃは、フレーム２４９Ｇのそれらと同じであるので、フレーム２４９Ｇのように、フレーム２４９Ｈは単一のＶＶｅｃｔｏｒＤａｔａフィールド１５６を含み、これは、各々が８ビットによって一様に量子化される１１個のベクトル要素を含む。

【０５７４】

[0574]図１０Ｌ（ｉ）および図１０Ｌ（ｉｉ）は、上の擬似コードにおけるケース２に対応するように生成された、第１の例示的なビットストリーム２４８Ｌと付随するＨＯＡｃｏｎｆｉｇ部分２５０Ｌとを示す。図１０Ｌ（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５０Ｌは、０次からＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０によって指定される次数（これは、この例では（ＨｏａＯｒｄｅｒ＋１）^{2}－（ＭｉｎＡｍｂＨｏａＯｒｄｅｒ＋１）^{2}＝１６－４＝１２に等しい）までの要素を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｌはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｌはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｌはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、（１＋１）^{2}すなわち４に等しいものとしてＭｉｎＮｕｍＯｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｌは、３に等しいものとして

コンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、（Ｎ＋１）^{2}すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５７５】

[0575]図１０Ｌ（ｉ）の例にさらに示されるように、部分２４８Ｌは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｉおよび２４９ＪがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５７６】

[0576]図１０Ｌ（ｉｉ）は、フレーム２４９Ｉと２４９Ｊとをより詳細に示す。図１０Ｌ（ｉｉ）の例に示されるように、フレーム２４９ＩはＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６および１５６Ｂとを含む。ＣＳＩＤフィールド１５４は、図１０Ｊ（ｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｂは、図１０Ｌ（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。ＣＳＩＤフィールド１５４～１５４Ｃの各々は、トランスポートチャネル１、２、および３のそれぞれ１つに対応する。

【０５７７】

[0577]図１０Ｌ（ｉｉ）の例では、フレーム２４９Ｉは、（ＣＳＩＤフィールド１５４および１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）２つのベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｈを仮定すると、オーディオ復号デバイス２４は、１２個のＶベクトル要素が符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６および１５６Ｂは各々、各々が８ビットによって一様に量子化される１２個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、単一のアスタリスク（*）によって示されるように、コーディングスキームは、対応するトランスポートチャネルのためのＣＳＩＤフィールドにおいて、ＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５７８】

[0578]フレーム２４９Ｊにおいて、ＣＳＩＤフィールド１５４および１５４Ｂはフレーム２４９Ｉの中のそれらと同じであるが、フレーム２４９ＦのＣＳＩＤフィールド１５４Ｃは１というＣｈａｎｎｅｌＴｙｐｅに切り替えられる。フレーム２４９ＢのＣＳＩＤフィールド１５４Ｃはしたがって、ＣｂＦｌａｇ２６７と、Ｐｆｌａｇ２６７（ハフマン符号化を示す）と、Ｎｂｉｔｓ２６１（１２に等しい）とを含む。結果として、フレーム２４９Ｆは、各々が１２ビットによって一様に量子化されハフマンコーディングされる、１２個のＶベクトル要素を含む第３のＶＶｅｃｔｏｒＤａｔａフィールド１５６Ｃを含む。

上で述べられたように、コーディングされたＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定されるが、ハフマンコーディングスキームは、この特定のトランスポートチャネル（たとえば、トランスポートチャネル第３）のためのＣＳＩＤフィールド１５４Ｃの中のＮｂｉｔｓＱ＝１２、ＣｂＦｌａｇ＝０、およびＰｆｌａｇ＝０によってシグナリングされる。

【０５７９】

[0579]図１０Ｍ（ｉ）および図１０Ｍ（ｉｉ）の例は、上の擬似コードにおけるケース２に対応するように生成された、第２の例示的なビットストリーム２４８Ｍと付随するＨＯＡ　ｃｏｎｆｉｇ部分２５０Ｍとを示す。図１０Ｍ（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５０Ｍは、０次からＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０によって指定される次数（これは、この例では（ＨｏａＯｒｄｅｒ＋１）^{2}－（ＭｉｎＡｍｂＨｏａＯｒｄｅｒ＋１）^{2}＝１６－４＝１２に等しい）までの要素を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｍはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｍはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。

【０５８０】

[0580]ＨＯＡｃｏｎｆｉｇ部分２５０Ｍはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、（１＋１）^{2}すなわち４に等しいものとしてＭｉｎＮｕｍｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。オーディオ復号デバイス２４はまた、ＮｕｍＯｆＨｏａＣｏｅｆｆｓシンタックス要素とＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡとの間の差に設定されるものとしてＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓシンタックス要素を導出することができ、これはこの例では１６－４すなわち１２に等しいと仮定される。オーディオ復号デバイス２４はまた、ｃｅｉｌ（ｌｏｇ２（ＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓ））＝ｃｅｉｌ（ｌｏｇ２（１２））＝４に設定されるものとしてＡｍｂＡｓｉｇｎｍＢｉｔｓシンタックス要素導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｍは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、（Ｎ＋１）^{2}すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５８１】

[0581]図１０Ｍ（ｉ）の例にさらに示されるように、部分２４８Ｍは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｋおよび２４９ＬがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５８２】

[0582]図１０Ｍ（ｉｉ）は、フレーム２４９Ｋと２４９Ｌとをより詳細に示す。図１０

Ｍ（ｉｉ）の例に示されるように、フレーム２４９ＫはＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６とを含む。ＣＳＩＤフィールド１５４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６と、ＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７（ここで、ダブルアスタリスク（**）は、フレキシブルトランスポートチャネル第１に対して、復号器の内部状態がここではＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎＳｔａｔｅ＝２であると仮定されることを示し、このことは、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘビットフィールドがシグナリングされる、またはビットストリームにおいて別様に指定されるをもたらす）と、ＣｈａｎｎｅｌＴｙｐｅ２６９（これは２に等しく、対応するペイロードが追加の環境ＨＯＡ係数であることをシグナリングする）とを含む。オーディオ復号デバイス２４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ＋１＋ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡすなわちこの例では５に等しいものとしてＡｍｂＣｏｅｆｆＩｄｘを導出することができる。ＣＳＩＤフィールド１５４Ｂは、図１０Ｍ（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。

【０５８３】

[0583]図１０Ｍ（ｉｉ）の例では、フレーム２４９Ｋは、（ＣＳＩＤフィールド１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）単一のベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｍを仮定すると、オーディオ復号デバイス２４は、１２個のＶベクトル要素が符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６は、各々が８ビットによって一様に量子化される１２個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、脚注２によって示されるように、コーディングスキームは、対応するトランスポートチャネルのＣＳＩＤフィールドにおいてＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５８４】

[0584]フレーム２４９Ｌにおいて、ＣＳＩＤフィールド１５４は、移行が起こらなかったことを示すＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７を含むので、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６は以前のフレームから推測されることが可能であり、再びシグナリングされること、または別様に指定されることは必要ではない。フレーム２４９ＬのＣＳＩＤフィールド１５４Ｂおよび１５４Ｃは、フレーム２４９Ｋのそれらと同じであるので、フレーム２４９Ｋのように、フレーム２４９Ｌは単一のＶＶｅｃｔｏｒＤａｔａフィールド１５６を含み、これは、各々が８ビットによって一様に量子化される１２個のベクトル要素を含む。

【０５８５】

[0585]図１０Ｎ（ｉ）および図１０Ｎ（ｉｉ）は、上の擬似コードにおけるケース３に対応するように生成された、第１の例示的なビットストリーム２４８Ｎと付随するＨＯＡｃｏｎｆｉｇ部分２５０Ｎとを示す。図１０Ｎ（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５０Ｎは、ＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素（これは、この例では０であると仮定される）において指定される要素を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｎはまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｎはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。ＨＯＡｃｏｎｆｉｇ部分２５０Ｋはさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が

１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、（１＋１）^{2}すなわち４に等しいものとしてＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５０Ｎは、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、（Ｎ＋１）^{2}すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５８６】

[0586]図１０Ｎ（ｉ）の例にさらに示されるように、部分２４８Ｎは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｍおよび２４９ＮがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５８７】

[0587]図１０Ｎ（ｉｉ）は、フレーム２４９Ｍと２４９Ｎとをより詳細に示す。図１０Ｎ（ｉｉ）の例に示されるように、フレーム２４９ＭはＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６および１５６Ｂとを含む。ＣＳＩＤフィールド１５４は、図１０Ｊ（ｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｂは、図１０Ｎ（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。ＣＳＩＤフィールド１５４～１５４Ｃの各々は、トランスポートチャネル１、２、および３のそれぞれ１つに対応する。

【０５８８】

[0588]図１０Ｎ（ｉｉ）の例では、フレーム２４９Ｍは、（ＣＳＩＤフィールド１５４および１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）２つのベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５０Ｍを仮定すると、オーディオ復号デバイス２４は、１６個のＶベクトル要素が符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６および１５６Ｂは各々、各々が８ビットによって一様に量子化される１６個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、単一のアスタリスク（*）によって示されるように、コーディングスキームは、対応するトランスポートチャネルのためのＣＳＩＤフィールドにおいて、ＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５８９】

[0589]フレーム２４９Ｎにおいて、ＣＳＩＤフィールド１５４および１５４Ｂはフレーム２４９Ｍの中のそれらと同じであるが、フレーム２４９ＦのＣＳＩＤフィールド１５４Ｃは１というＣｈａｎｎｅｌＴｙｐｅに切り替えられる。フレーム２４９ＢのＣＳＩＤフィールド１５４Ｃはしたがって、ＣｂＦｌａｇ２６７と、Ｐｆｌａｇ２６７（ハフマン符

号化を示す）と、Ｎｂｉｔｓ２６１（１２に等しい）とを含む。結果として、フレーム２４９Ｆは、各々が１２ビットによって一様に量子化されハフマンコーディングされる１６個のＶベクトル要素を含む、第３のＶＶｅｃｔｏｒＤａｔａフィールド１５６Ｃを含む。上で述べられたように、コーディングされたＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定されるが、ハフマンコーディングスキームは、この特定のトランスポートチャネル（たとえば、トランスポートチャネル第３）のためのＣＳＩＤフィールド１５４Ｃの中のＮｂｉｔｓＱ＝１２、ＣｂＦｌａｇ＝０、およびＰｆｌａｇ＝０によってシグナリングされる。

【０５９０】

[0590]図１００（ｉ）および図１００（ｉｉ）の例は、上の擬似コードにおけるケース３に対応するように生成された、第２の例示的なビットストリーム２４８Ｏと付随するＨＯＡ　ｃｏｎｆｉｇ部分２５００とを示す。図１００（ｉ）の例では、ＨＯＡｃｏｎｆｉｇ部分２５００は、ＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素（これは、この例では１であると仮定される）において指定される要素を除き、Ｖベクトルのすべての要素がコーディングされることを示すように設定されるＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈシンタックス要素２５６を含む。ＨＯＡｃｏｎｆｉｇ部分２５００はまた、空間－時間的な補間の補間関数が二乗余弦であることを示すように設定された、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５を含む。ＨＯＡｃｏｎｆｉｇ部分２５００Ｏはその上、２５６という補間されたサンプル継続時間を示すように設定された、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅ２５４を含む。

【０５９１】

[0591]ＨＯＡｃｏｎｆｉｇ部分２５００はさらに、環境ＨＯＡコンテンツのＭｉｎｉｍｕｍＨＯＡ次数が１であることを示すように設定された、ＭｉｎＡｍｂＨｏａＯｒｄｅｒシンタックス要素１５０を含み、ここでオーディオ復号デバイス２４は、$(1+1)^2$すなわち４に等しいものとしてＭｉｎＮｕｍｏｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を導出することができる。オーディオ復号デバイス２４はまた、ＮｕｍＯｆＨｏａＣｏｅｆｆｓシンタックス要素とＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡとの間の差に設定されるものとしてＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓシンタックス要素を導出することができ、これはこの例では１６－４すなわち１２に等しいと仮定される。オーディオ復号デバイス２４はまた、ｃｅｉｌ（ｌｏｇ２（ＭａｘＮｏＯｆＡｄｄＡｃｔｉｖｅＡｍｂＣｏｅｆｆｓ））＝ｃｅｉｌ（ｌｏｇ２（１２））＝４に設定されるものとしてＡｍｂＡｓｉｇｎｍＢｉｔｓをシンタックス要素導出することができる。ＨＯＡｃｏｎｆｉｇ部分２５００は、３に等しいものとしてコンテンツのＨＯＡ次数を示すように設定されたＨｏａＯｒｄｅｒシンタックス要素１５２を含み（または言い換えればＮ＝３）、ここでオーディオ復号デバイス２４は、$(N+1)^2$すなわち１６に等しいものとしてＮｕｍＯｆＨｏａＣｏｅｆｆｓを導出することができる。

【０５９２】

[0592]図１００（ｉ）の例にさらに示されるように、部分２４８Ｏは、スペクトル帯域複製（ＳＢＲ）が有効にされるときに２つのオーディオフレームが１つのＵＳＡＣ－３Ｄフレームに記憶されるとすると、２つのＨＯＡフレーム２４９Ｏおよび２４９ＰがＵＳＡＣ拡張ペイロードに記憶される、ＵＳＡＣ－３Ｄオーディオフレームを含む。オーディオ復号デバイス２４は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素およびＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素の関数として、フレキシブルトランスポートチャネルの数を導出することができる。次の例では、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素が７に等しく、ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素が４に等しいことが仮定され、ここで、フレキシブルトランスポートチャネルの数は、ｎｕｍＨＯＡＴｒａｎｓｐｏｒｔＣｈａｎｎｅｌｓシンタックス要素からＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡシンタックス要素を引いたもの（すなわち３）に等しい。

【０５９３】

[0593]図１００（ｉｉ）は、フレーム２４９Ｏと２４９Ｐとをより詳細に示す。図１００（ｉｉ）の例に示されるように、フレーム２４９ＯはＣＳＩＤフィールド１５４～１５４ＣとＶＶｅｃｔｏｒＤａｔａフィールド１５６とを含む。ＣＳＩＤフィールド１５４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６と、ＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７（ここで、ダブルアスタリスク（**）は、フレキシブルトランスポートチャネル第１に対して、復号器の内部状態がここではＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎＳｔａｔｅ＝２であると仮定されることを示し、このことは、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘビットフィールドがシグナリングされる、またはビットストリームにおいて別様に指定されるをもたらす）と、ＣｈａｎｎｅｌＴｙｐｅ２６９（これは２に等しく、対応するペイロードが追加の環境ＨＯＡ係数であることをシグナリングする）とを含む。オーディオ復号デバイス２４は、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ＋１＋ＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡすなわちこの例では５に等しいものとしてＡｍｂＣｏｅｆｆＩｄｘを導出することができる。ＣＳＩＤフィールド１５４Ｂは、図１００（ｉｉ）の例に示される各々が対応する値０１、１、０、および０１に設定される、ｕｎｉｔＣ２６７と、ｂｂ２６６と、ｂａ２６５とともにＣｈａｎｎｅｌＴｙｐｅ２６９を含む。ＣＳＩＤフィールド１５４Ｃは、３という値を有するＣｈａｎｎｅｌＴｙｐｅフィールド２６９を含む。

【０５９４】

[0594]図１００（ｉｉ）の例では、フレーム２４９Ｏは、（ＣＳＩＤフィールド１５４Ｂの中のＣｈａｎｎｅｌＴｙｐｅ２６９が１に等しいとすると）単一のベクトルベースの信号と（ＣＳＩＤフィールド１５４Ｃの中のＣｈａｎｎｅｌＴｙｐｅ２６９が３に等しいとすると）空とを含む。前述のＨＯＡｃｏｎｆｉｇ部分２５００を仮定すると、オーディオ復号デバイス２４は、１６からＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素に（たとえば、６というインデックスと関連付けられるベクトル要素がＣｏｎｔＡｄｄＡｍｂＨｏａＣｈａｎシンタックス要素として指定される）よって指定されるものを引いたもの、すなわち１５個のＶベクトルが符号化されると決定することができる。したがって、ＶＶｅｃｔｏｒＤａｔａ１５６は、各々が８ビットによって一様に量子化される１５個のベクトル要素を含む。脚注１によって示されるように、コーディングされるＶＶｅｃｔｏｒＤａｔａ要素の数およびインデックスは、パラメータＣｏｄｅｄＶＶｅｃＬｅｎｇｔｈ＝０によって指定される。その上、脚注２によって示されるように、コーディングスキームは、対応するトランスポートチャネルのＣＳＩＤフィールドにおいてＮｂｉｔｓＱ＝５によってシグナリングされる。

【０５９５】

[0595]フレーム２４９Ｐにおいて、ＣＳＩＤフィールド１５４は、移行が起こらなかったことを示すＡｍｂＣｏｅｆｆＩｄｘＴｒａｎｓｉｔｉｏｎ２４７を含むので、ＣｏｄｅｄＡｍｂＣｏｅｆｆＩｄｘ２４６は以前のフレームから推測されることが可能であり、再びシグナリングされること、または別様に指定されることは必要ではない。フレーム２４９ＰのＣＳＩＤフィールド１５４Ｂおよび１５４Ｃは、フレーム２４９Ｏのそれらと同じであるので、フレーム２４９Ｏのように、フレーム２４９Ｐは単一のＶＶｅｃｔｏｒＤａｔａフィールド１５６を含み、これは、各々が８ビットによって一様に量子化される１５個のベクトル要素を含む。

【０５９６】

[0596]図１１Ａ～図１１Ｇは、図５の例に示されるオーディオ復号デバイス２４の様々なユニットをより詳細に示すブロック図である。図１１Ａは、オーディオ復号デバイス２４の抽出ユニット７２をより詳細に示すブロック図である。図１１Ａの例に示されるように、抽出ユニット７２は、モード解析ユニット２７０と、モード構成ユニット２７２（「モード構成ユニット２７２（ｍｏｄｅ　ｃｏｎｆｉｇ　ｕｎｉｔ　２７２）」）と、構成可能抽出ユニット２７４とを含み得る。

【０５９７】

[0597]モード解析ユニット２７０は、ビットストリーム２１を形成するためにＨＯＡ係

数１１を符号化するために使用されるコーディングモードを示す上で述べられたシンタックス要素（たとえば、図１０Ｅの例に示されるＣｈａｎｎｅｌＴｙｐｅシンタックス要素）を解析するように構成されるユニットを表し得る。モード解析ユニット２７０は、決定されたシンタックス要素をモード構成ユニット２７２に渡すことができる。モード構成ユニット２７２は、解析されたシンタックス要素に基づいて構成可能抽出ユニット２７４を構成するように構成されるユニットを表し得る。モード構成ユニット２７２は、ビットストリーム２１からＨＯＡ係数１１の指向性ベースのコーディングされた表現を抽出し、または解析されたシンタックス要素に基づいてビットストリーム２１からＨＯＡ係数１１のベクトルベースのコーディングされた表現を抽出するように、構成可能抽出ユニット２７４を構成することができる。

【０５９８】

[0598]指向性ベースの符号化が実行されたとき、構成可能抽出ユニット２７４は、ＨＯＡ係数１１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と関連付けられるシンタックス要素（図１１Ａの例では指向性ベースの情報９１として示される）を抽出することができる。指向性ベースの情報９１は、０に等しいＣｈａｎｎｅｌＴｙｐｅによって定義されるように、図１０Ｄの例に示される指向性情報２５３と図１０Ｅの例に示される指向性ベースのＳｉｄｅＣｈａｎｎｅｌＩｎｆｏＤａｔａとを含み得る。

【０５９９】

[0599]ＨＯＡ係数１１がベクトルベースの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（たとえば、ＣｈａｎｎｅｌＴｙｐｅシンタックス要素が１に等しいとき）、構成可能抽出ユニット２７４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ＨＯＡ係数５９と、符号化されたｎＦＧ信号５９とを抽出することができる。構成可能抽出ユニット２７４はまた、ＨＯＡ係数１１がベクトルベースの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決定すると、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅシンタックス要素２５４とＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５とをビットストリーム２１から抽出し、これらのシンタックス要素２５４および２５５を空間－時間的補間ユニット７６に渡すことができる。

【０６００】

[0600]図１１Ｂは、図５の例に示されるオーディオ復号デバイス２４の量子化ユニット７４をより詳細に示すブロック図である。量子化ユニット７４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をエントロピー復号して逆量子化し、それによって低減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、図４の例に示される量子化ユニット５２とは逆の方式で動作するように構成されるユニットを表し得る。スカラー／エントロピー逆量子化ユニット９８４は、カテゴリ／残差復号ユニット２７６と、予測ユニット２７８と、一様逆量子化ユニット２８０とを含み得る。

【０６０１】

[0601]カテゴリ／残差復号ユニット２７６は、ハフマンテーブル情報２４１によって識別されるハフマンテーブル（これは、上で述べられたように、ビットストリーム２１中のシンタックス要素として表現される）を使用して、コーディングされたフォアグラウンドＶ［ｋ］ベクトル５７に関してハフマン復号を実行するように構成されるユニットを表し得る。カテゴリ／残差復号ユニット２７６は、量子化されたフォアグラウンドＶ［ｋ］ベクトルを予測ユニット２７８に出力することができる。予測ユニット２７８は、予測モード２３７に基づいて量子化されたフォアグラウンドＶ［ｋ］ベクトルに関して予測を実行し、補強された量子化されたフォアグラウンドＶ［ｋ］ベクトルを一様逆量子化ユニット２８０に出力するように構成されるユニットを表し得る。一様逆量子化ユニット２８０は、ｎｂｉｔｓ値２３３に基づいて、補強された量子化されたフォアグラウンドＶ［ｋ］ベクトルに関して逆量子化を実行し、低減されたフォアグラウンドＶ［ｋ］ベクトル５５_kを出力するように構成されるユニットを表し得る。

【０６０２】

[0602]図１１Ｃは、図５の例に示されるオーディオ復号デバイス２４の音響心理学的復号ユニット８０をより詳細に示すブロック図である。上で述べられたように、音響心理学的復号ユニット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４の例に示される音響心理学的オーディオコーディングユニット４０とは逆の方式で動作することができる。音響心理学的復号ユニット８０は、エネルギー補償された環境ＨＯＡ係数４７’をＨＯＡ係数定式化ユニット８２に渡し、ｎＦＧ信号４９’を再順序付け８４に渡すことができる。音響心理学的復号ユニット８０は、音響心理学的オーディオコーディングユニット４０と同様の複数のオーディオ復号器８０～８０Ｎを含み得る。オーディオ復号器８０～８０Ｎは、上で述べられたように、バックグラウンドＨＯＡ係数４７’の各チャネルおよびｎＦＧ信号４９’の各信号の同時の復号をサポートするのに十分な量で、音響心理学的オーディオコーディングユニット４０の中で実体化され、または別様に含まれ得る。

【０６０３】

[0603]図１１Ｄは、図５の例に示されるオーディオ復号デバイス２４の再順序付けユニット８４をより詳細に示すブロック図である。再順序付けユニット８４は、再順序付けユニット３４に関して上で説明されたものとは逆の方式で動作するように構成されるユニットを表し得る。再順序付けユニット８４は、ＨＯＡ係数１１のフォアグラウンド成分のオリジナルの次数を示すシンタックス要素２０５を受信するように構成されるユニットを表し得る、ベクトル再順序付けユニット２８２を含み得る。抽出ユニット７２は、ビットストリーム２１からのこれらのシンタックス要素を解析し、シンタックス要素２０５を再順序付けユニット８４に渡すことができる。ベクトル再順序付けユニット２８２は、これらの再順序付けシンタックス要素２０５に基づいて、再順序付けられたｎＦＧ信号４９’’と再順序付けられたフォアグラウンドＶ［ｋ］ベクトル55_k’とを生成するために、補間されたｎＦＧ信号４９’と低減されたフォアグラウンドＶ［ｋ］ベクトル55_kとを再順序付けることができる。再順序付けユニット８４は、再順序付けられたｎＦＧ信号４９’’をフォアグラウンド定式化ユニット７８に出力し、再順序付けられたフォアグラウンドＶ［ｋ］ベクトル55_k’を空間－時間的補間ユニット７６に出力することができる。

【０６０４】

[0604]図１１Ｅは、図５の例に示されるオーディオ復号デバイス２４の空間－時間的補間ユニット７６をより詳細に示すブロック図である。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に関して上で説明されたものと同様の方式で動作することができる。空間－時間的補間ユニット７６は、Ｖ補間ユニット２８４を含んでよく、これは、再順序付けられたフォアグラウンドＶ［ｋ］ベクトル55_k’を受信し、補間されたフォアグラウンドＶ［ｋ］ベクトル55_k’’を生成するために再順序付けられたフォアグラウンドＶ［ｋ］ベクトル55_k’および再順序付けら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－時間的な補間を実行するように構成されるユニットを表し得る。Ｖ補間ユニット２８４は、ＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅシンタックス要素２５４およびＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５に基づいて補間を実行することができる。いくつかの例では、Ｖ補間ユニット２８５は、ＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＭｅｔｈｏｄシンタックス要素２５５によって識別された補間のタイプを使用してＣｏｄｅｄＳｐａｔｉａｌＩｎｔｅｒｐｏｌａｔｉｏｎＴｉｍｅシンタックス要素２５４によって指定される継続時間にわたり、Ｖベクトルを補間することができる。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’’をフォアグラウンド定式化ユニット７８に転送することができる。

【０６０５】

[0605]図１１Ｆは、図５の例に示されるオーディオ復号デバイス２４のフォアグラウンド定式化ユニット７８をより詳細に示すブロック図である。フォアグラウンド定式化ユニ

ット７８は乗算ユニット２８６を含んでよく、これは、フォアグラウンドＨＯＡ係数６５を生成するために補間されたフォアグラウンドＶ［ｋ］ベクトル５５$_k$’’および再順序付けられたｎＦＧ信号４９’’に関して行列乗算を実行するように構成されるユニットを表し得る。

【０６０６】

[0606]図１１Ｇは、図５の例に示されるオーディオ復号デバイス２４のＨＯＡ係数定式化ユニット８２をより詳細に示すブロック図である。ＨＯＡ係数定式化ユニット８２は加算ユニット２８８を含んでよく、これは、ＨＯＡ係数１１’を取得するために、フォアグラウンドＨＯＡ係数６５を環境ＨＯＡチャネル４７’に加算するように構成されるユニットを表し得る。

【０６０７】

[0607]図１２は、本開示において説明される技法の様々な態様を実行し得る例示的なオーディオエコシステムを示す図である。図１２に示されるように、オーディオエコシステム３００は、取得３０１と、編集３０２と、コーディング３０３と、送信３０４と、再生３０５とを含み得る。

【０６０８】

[0608]取得３０１は、オーディオコンテンツが取得されるオーディオエコシステム３００の技法を表し得る。取得３０１の例は、限定はされないが、音声（たとえばライブの音声）を記録すること、オーディオの生成（たとえば、オーディオオブジェクト、フォーリー作成物、音声合成、シミュレーション）などを含む。いくつかの例では、音声は、コンサートにおいて、スポーツイベントにおいて、および監視を行うときに記録され得る。いくつかの例では、オーディオは、シミュレーション、およびオーサード／ミキシング（たとえば、映画、ゲーム）を実行するときに生成され得る。オーディオオブジェクトは、ハリウッド（たとえば、ＩＭＡＸスタジオ）において使用され得る。いくつかの例では、取得３０１は、図３のコンテンツ作成者１２などのコンテンツ作成者によって実行され得る。

【０６０９】

[0609]編集３０２は、オーディオコンテンツが編集および／または修正されるオーディオエコシステム３００の技法を表し得る。一例として、オーディオコンテンツは、オーディオコンテンツの複数のユニットをオーディオコンテンツの単一のユニットへと組み合わせることによって編集され得る。別の例として、オーディオコンテンツは、実際のオーディオコンテンツを調整する（たとえば、オーディオコンテンツの１つまたは複数の周波数成分のレベルを調整する）ことによって編集され得る。いくつかの例では、編集３０２は、図３のオーディオ編集システム１８などのオーディオ編集システムによって実行され得る。いくつかの例では、編集３０２は、図２９に示されるモバイルデバイスの１つまたは複数などのモバイルデバイス上で実行され得る。

【０６１０】

[0610]コーディング３０３は、オーディオコンテンツがオーディオコンテンツの表現へとコーディングされるオーディオエコシステム３００の技法を表し得る。いくつかの例では、オーディオコンテンツの表現は、図３のビットストリーム２１などのビットストリームであり得る。いくつかの例では、コーディング３０２は、図３の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によって実行され得る。

【０６１１】

[0611]送信３０４は、オーディオコンテンツがコンテンツ作成者からコンテンツ消費者に運ばれるオーディオエコシステム３００の要素を表し得る。いくつかの例では、オーディオコンテンツは、リアルタイムで、またはほぼリアルタイムで運ばれ得る。たとえば、オーディオコンテンツはコンテンツ消費者にストリーミングされ得る。いくつかの例では、オーディオコンテンツは、コンピュータ可読記憶媒体などの媒体上にオーディオコンテンツをコーディングすることによって運ばれ得る。たとえば、オーディオコンテンツは、ディスク、ドライブなど（たとえば、ブルーレイ（登録商標）ディスク、メモリカード、

ハードドライブなど）に記憶され得る。

【０６１２】

[0612]再生３０５は、オーディオコンテンツがレンダリングされコンテンツ消費者に向かって再生されるオーディオエコシステム３００の技法を表し得る。いくつかの例では、再生３０５は、再生環境の１つまたは複数の態様に基づいて３Ｄ音場をレンダリングすることを含み得る。言い換えれば、再生３０５は、ローカルの音響ランドスケープに基づき得る。

【０６１３】

[0613]図１３は、図１２のオーディオエコシステムの一例をより詳細に示す図である。図１３に示されるように、オーディオエコシステム３００は、オーディオコンテンツ３０８と、映画撮影所３１０と、音楽スタジオ３１１と、ゲーミングオーディオスタジオ３１２と、チャネルベースのオーディオコンテンツ３１３と、コーディングエンジン３１４と、ゲームオーディオステム３１５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３１６と、配信システム３１７とを含み得る。例示的なゲーミングオーディオスタジオ３１２が図２６に示されている。いくつかの例示的なゲームオーディオコーディング／レンダリングエンジン３１６が図２７に示されている。

【０６１４】

[0614]図１３によって示されるように、映画撮影所３１０、音楽スタジオ３１１、およびゲーミングオーディオスタジオ３１２はオーディオコンテンツ３０８を受け取り得る。いくつかの例では、オーディオコンテンツ３０８は、図１２の取得３０１の出力を表し得る。映画撮影所３１０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、チャネルベースのオーディオコンテンツ３１３（たとえば、２．０、５．０、および７．１の）を出力することができる。音楽スタジオ３１０は、ＤＡＷを出力することなどによって、チャネルベースのオーディオコンテンツ３１３（たとえば、２．０、および５．１の）を出力することができる。いずれの場合でも、コーディングエンジン３１４は、配信システム３１７による出力のために、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）に基づいて、チャネルベースのオーディオコンテンツ３１３を受信し符号化することができる。このようにして、コーディングエンジン３１４は、図１２のコーディング３０３の例であり得る。ゲーミングオーディオスタジオ３１２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ステム３１５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３１６は、配信システム３１７による出力のために、オーディオステム３１５をチャネルベースの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。いくつかの例では、映画撮影所３１０、音楽スタジオ３１１、およびゲーミングオーディオスタジオ３１２の出力は、図１２の編集３０２の出力を表し得る。いくつかの例では、コーディングエンジン３１４および／またはゲームオーディオコーディング／レンダリングエンジン３１６の出力は、図１２の送信３０４の技法を介して配信システム３１７に運ばれ得る。

【０６１５】

[0615]図１４は、図１２のオーディオエコシステムの別の例をより詳細に示す図である。図１４に示されるように、オーディオエコシステム３００Ｂは、ブロードキャストされる録音オーディオオブジェクト３１９と、プロフェッショナルオーディオシステム３２０と、消費者向けオンデバイスキャプチャ３２２と、ＨＯＡオーディオフォーマット３２３と、オンデバイスレンダリング３２４と、消費者向けオーディオ、ＴＶ、およびアクセサリ３２５と、カーオーディオシステム３２６とを含み得る。

【０６１６】

[0616]図１４に示されるように、ブロードキャストされる録音オーディオオブジェクト３１９、プロフェッショナルオーディオシステム３２０、および消費者向けオンデバイスキャプチャ３２２はすべて、ＨＯＡオーディオフォーマット３２３を使用して出力をコー

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３２４と、消費者向けオーディオ、ＴＶ、およびアクセサリ３２５と、カーオーディオシステム３２６とを使用して再生され得る単一の表現へと、ＨＯＡオーディオフォーマット３２３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、汎用的なオーディオ再生システム（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な）において再生され得る。

【０６１７】

[0617]図１５Ａおよび図１５Ｂは、図１２のオーディオエコシステムの他の例をより詳細に示す図である。図１５Ａに示されるように、オーディオエコシステム３００Ｃは、取得要素３３１と再生要素３３６とを含み得る。取得要素３３１は、有線および／またはワイヤレス取得デバイス３３２（たとえば、Ｅｉｇｅｎマイクロフォン）と、オンデバイスサラウンドサウンドキャプチャ３３４と、モバイルデバイス３３５（たとえば、スマートフォンおよびタブレット）とを含み得る。いくつかの例では、有線および／またはワイヤレス取得デバイス３３２は、有線および／またはワイヤレス通信チャネル３３３を介してモバイルデバイス３３５に結合され得る。

【０６１８】

[0618]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３３５が音場を取得するために使用され得る。たとえば、モバイルデバイス３３５は、有線および／もしくはワイヤレス取得デバイス３３２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３３４（たとえば、モバイルデバイス３３５へと統合される複数のマイクロフォン）を介して、音場を取得することができる。モバイルデバイス３３５は次いで、再生要素３３６の１つまたは複数による再生のために、取得された音場をＨＯＡ３３７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３３５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（その音場を取得し）、録音をＨＯＡへとコーディングすることができる。

【０６１９】

[0619]モバイルデバイス３３５はまた、ＨＯＡコーディングされる音場を再生するために再生要素３３６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３３５は、ＨＯＡコーディングされた音場を復号し、信号を再生要素３３６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３３６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３３５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３３８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３３５は、１つまたは複数のドッキングステーションおよび／または１つまた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３３９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３３５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３４０を利用することができる。

【０６２０】

[0620]いくつかの例では、特定のモバイルデバイス３３５は、３Ｄ音場を取得することと、より後の時間に同じ３Ｄ音場を再生すること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３３５は、３Ｄ音場を取得し、３Ｄ音場をＨＯＡへと符号化し、符号化された３Ｄ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。

【０６２１】

[0621]図１５Ｂに示されるように、オーディオエコシステム３００Ｄは、オーディオコンテンツ３４３と、ゲームスタジオ３４４と、コーディングされるオーディオコンテンツ３４５と、レンダリングエンジン３４６と、配信システム３４７とを含み得る。いくつかの例では、ゲームスタジオ３４４は、ＨＯＡ信号の編集をサポートし得る１つまたは複数

のＤＡＷを含み得る。たとえば、１つまたは複数のＤＡＷ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、ＨＯＡ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３４４は、ＨＯＡ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。いずれの場合でも、ゲームスタジオ３４４は、配信システム３４７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３４６に、コーディングされたオーディオコンテンツ３４５を出力することができる。

【０６２２】

[0622]図１６は、本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ符号化デバイスを示す図である。図１６に示されるように、オーディオエコシステム３００Ｅは、オリジナルの３Ｄオーディオコンテンツ３５１と、符号化器３５２と、ビットストリーム３５３と、復号器３５４と、レンダラ３５５と、再生要素３５６とを含み得る。図１６によってさらに示されるように、符号化器３５２は、音場分析および分解３５７と、バックグラウンド抽出３５８と、バックグラウンドの顕著さの決定３５９と、オーディオコーディング３６０と、フォアグラウンド／明瞭オーディオ抽出３６１と、オーディオコーディング３６２とを含み得る。いくつかの例では、符号化器３５２は、図３および図４のオーディオ符号化デバイス２０と同様の動作を実行するように構成され得る。いくつかの例では、音場分析および分解３５７は、図４の音場分析ユニット４４と同様の動作を実行するように構成され得る。いくつかの例では、バックグラウンド抽出３５８およびバックグラウンドの顕著さの決定３５９は、図４のＢＧ選択ユニット４８と同様の動作を実行するように構成され得る。いくつかの例では、オーディオコーディング３６０およびオーディオコーディング３６２は、図４の音響心理学的オーディオコーダユニット４０と同様の動作を実行するように構成され得る。いくつかの例では、フォアグラウンド／明瞭オーディオ抽出３６１は、図４のフォアグラウンド選択ユニット３６と同様の動作を実行するように構成され得る。

【０６２３】

[0623]いくつかの例では、フォアグラウンド／明瞭オーディオ抽出３６１は、図３３のビデオフレーム３９０に対応するオーディオコンテンツを分析することができる。たとえば、フォアグラウンド／明瞭オーディオ抽出３６１は、領域３９１Ａ～３９１Ｃに対応するオーディオコンテンツがフォアグラウンドオーディオであると決定することができる。

【０６２４】

[0624]図１６に示されるように、符号化器３５２は、２５～７５Ｍｂｐｓというビットレートを有し得るオリジナルのコンテンツ３５１を、２５６ｋｂｐｓ～１．２Ｍｂｐｓというビットレートを有し得るビットストリーム３５３へと符号化するように構成され得る。図１７は、図１６のオーディオ符号化デバイスの一例をより詳細に示す図である。

【０６２５】

[0625]図１８は、本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ復号デバイスを示す図である。図１８に示されるように、オーディオエコシステム３００Ｅは、オリジナルの３Ｄオーディオコンテンツ３５１と、符号化器３５２と、ビットストリーム３５３と、復号器３５４と、レンダラ３５５と、再生要素３５６とを含み得る。図１６によってさらに示されるように、復号器３５４は、オーディオ復号器３６３と、オーディオ復号器３６４と、フォアグラウンド再構築３６５と、ミキシング３６６とを含み得る。いくつかの例では、復号器３５４は、図３および図５のオーディオ復号デバイス２４と同様の動作を実行するように構成され得る。いくつかの例では、オーディオ復号器３６３、オーディオ復号器３６４は、図５の音響心理学的復号ユニット８０と同様の動作を実行するように構成され得る。いくつかの例では、フォアグラウンド再構築３６５は、図５のフォアグラウンド定式化ユニット７８と同様の動作を実行するように構成され得る。

【０６２６】

[0626]図１６に示されるように、復号器３５４は、ビットストリーム３５３を受信して

復号し、得られた再構築された３Ｄ音場をレンダラ３５５に出力するように構成されてよく、レンダラ３５５は次いで、再生要素３５６の１つまたは複数に、オリジナルの３Ｄコンテンツ３５１の表現を出力させ得る。図１９は、図１８のオーディオ復号デバイスの一例をより詳細に示す図である。

【０６２７】

[0627]図２０Ａ～図２０Ｇは、本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ取得デバイスを示す図である。図２０Ａは、３Ｄ音場を記録するように集団的に構成される複数のマイクロフォンを含み得る、Ｅｉｇｅｎマイクロフォン３７０を示す。いくつかの例では、Ｅｉｇｅｎマイクロフォン３７０の複数のマイクロフォンは、約４ｃｍ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３７０から直接ビットストリーム１７を出力するために、Ｅｉｇｅｎマイクロフォンに統合され得る。

【０６２８】

[0628]図２０Ｂは、１つまたは複数のＥｉｇｅｎマイクロフォン３７０などの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３７２を示す。製作トラック３７２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。

【０６２９】

[0629]図２０Ｃ～図２０Ｅは、３Ｄ音場を記録するように集団的に構成される複数のマイクロフォンを含み得るモバイルデバイス３７４を示す。言い換えれば、複数のマイクロフォンは、Ｘ、Ｙ、Ｚのダイバーシチを有し得る。いくつかの例では、モバイルデバイス３７４は、モバイルデバイス３７４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のダイバーシチを提供するように回転され得るマイクロフォン３７６を含み得る。モバイルデバイス３７４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。

【０６３０】

[0630]図２０Ｆは、３Ｄ音場を記録するように構成され得る、耐衝撃性のビデオキャプチャデバイス３７８を示す。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３７８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３７８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３７８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３Ｄ音場をキャプチャすることができる。

【０６３１】

[0631]図２０Ｇは、３Ｄ音場を記録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３８０を示す。いくつかの例では、モバイルデバイス３８０は、図１５のモバイルデバイス３３５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、Ｅｉｇｅｎマイクロフォンは、アクセサリで増強されたモバイルデバイス３８０を形成するために、図１５のモバイルデバイス３３５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３８０は、アクセサリで増強されたモバイルデバイス３８０と一体のサウンドキャプチャコンポーネントをただ使用したときよりも高品質なバージョンの３Ｄ音場をキャプチャすることができる。

【０６３２】

[0632]図２１Ａ～図２１Ｅは、本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ再生デバイスを示す図である。図２１Ａおよび図２１Ｂは、複数のスピーカー３８２とサウンドバー３８４とを示す。本開示の１つまたは複数の技法によれば、スピーカー３８２および／またはサウンドバー３８４は、あらゆる任意の構成で配置されることが可能であり、それでも３Ｄ音場を再生する。図２１Ｃ～図２１Ｅは、複数のヘッドフォン再生デバイス３８６～３８６Ｃを示す。ヘッドフォン再生デバイス３８６～３８６Ｃは、有線接続とワイヤレス接続のいずれかを介して復号器に結合され得る。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３８２、サウンドバー３８４、およびヘッドフォン再生デバイス３８６～３８６Ｃの任意の組合せで音場をレンダリングするために利用され得る。

【０６３３】

[0633]図２２Ａ～図２２Ｈは、本開示において説明される１つまたは複数の技法による例示的なオーディオ再生環境を示す図である。たとえば、図２２Ａは５．１スピーカー再生環境を示し、図２２Ｂは２．０（たとえば、ステレオ）スピーカー再生環境を示し、図２２Ｃはフルハイトの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を示し、図２２Ｄおよび図２２Ｅは各々２２．２スピーカー再生環境を示し、図２２Ｆは１６．０スピーカー再生環境を示し、図２２Ｇは自動車向けのスピーカー再生環境を示し、図２２Ｈはイヤーバッドの再生環境を伴うモバイルデバイスを示す。

【０６３４】

[0634]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、図２２Ａ～図２２Ｈに示される再生環境のいずれかにおいても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、レンダードが、図２２Ａ～図２２Ｈに示される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。

【０６３５】

[0635]図２３に示されるように、ユーザはヘッドフォン３８６を装着しながらスポーツの試合を見ることがある。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音場が取得されることが可能であり（たとえば、１つまたは複数のＥｉｇｅｎマイクロフォンが図２４に示される野球場の中および／または周りに配置されることが可能であり）、３Ｄ音場に対応するＨＯＡ係数が取得され復号器に送信されることが可能であり、復号器がＨＯＡ係数に基づいて３Ｄ音場を再構築するために決定し、再構築されたザ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）についてのインジケーションを取得し、再構築された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。いくつかの例では、レンダラは、図２５の技法に従って再生環境のタイプについてのインジケーションを取得することができる。このようにして、レンダラは、様々なスピーカーロケーション、タイプの数、サイズに「適応」し、また、ローカルの環境に対して最適にイコライズすることができる。

【０６３６】

[0636]図２８は、本開示において説明される１つまたは複数の技法による、ヘッドフォンにより模擬され得るスピーカー構成を示す図である。図２８によって示されるように、本開示の技法は、ヘッドフォン３８９を装着しているユーザが、音場がスピーカー３８８によって再生されているかのように音場を体験することを可能にし得る。このようにして、ユーザは、広い領域に音が出力されることなく、３Ｄ音場を聞くことができる。

【０６３７】

[0637]図３０は、本開示において説明される１つまたは複数の技法に従って処理され得る３Ｄ音場と関連付けられるビデオフレームを示す図である。

【０６３８】

[0638]図３１Ａ～図３１Ｍは、本開示において説明される技法の様々な態様による、音場の合成または記録された分類の実行の様々なシミュレーション結果を示すグラフ４００Ａ～４００Ｍを示す図である。図３１Ａ～図３１Ｍの例では、グラフ４００Ａ～４００Ｍの各々は、点線によって示される閾値４０２と、破線によって示されるそれぞれのオーディオオブジェクト４０４Ａ～４０４Ｍ（総称的に「オーディオオブジェクト４０４」）とを含む。

【０６３９】
[0639]コンテンツ分析ユニット２６に関して上で説明された分析を経てオーディオオブジェクト４０４が閾値４０２の下にあると決定されるとき、コンテンツ分析ユニット２６は、オーディオオブジェクト４０４の対応する１つが、録音されたオーディオオブジェクトを表すと決定する。図３１Ｂ、図３１Ｄ～図３１Ｈ、および図３１Ｊ～図３１Ｌの例に示されるように、コンテンツ分析ユニット２６は、オーディオオブジェクト４０４Ｂ、４０４Ｄ～４０４Ｈ、４０４Ｊ～４０４Ｌが（少なくとも時間の９０％、しばしば時間の１００％の間）閾値４０２の下にあり、したがって録音されたオーディオオブジェクトを表すと決定する。図３１Ａ、図３１Ｃ、および図３１Ｉの例に示されるように、コンテンツ分析ユニット２６は、オーディオオブジェクト４０４Ａ、４０４Ｃ、および４０４Ｉが閾値４０２を超え、したがって合成オーディオオブジェクトを表すと決定する。

【０６４０】
[0640]図３１Ｍの例では、オーディオオブジェクト４０４Ｍは、いくつかの合成部分（たとえば、閾値４０２を上回る）といくつかの合成部分（たとえば、閾値４０２を下回る）とを有する、混合された合成オーディオオブジェクト／録音されたオーディオオブジェクトを表す。この例でのコンテンツ分析ユニット２６は、オーディオオブジェクト４０４Ｍの合成部分と録音部分とを識別し、その結果、オーディオ符号化デバイス２０は、指向性ベースの符号化されたオーディオデータとベクトルベースの符号化されたオーディオデータの両方を含むようにビットストリーム２１を生成する。

【０６４１】
[0641]図３２は、本開示において説明される技法による、高次アンビソニックス係数から分解されたＳ行列からの特異値のグラフ４０６を示す図である。図３２に示されるように、大きい値を有する０ではない特異値はわずかである。図４の音場分析ユニット４４は、これらの特異値を分析して、再順序付けられたＵＳ［ｋ］ベクトル３３’および再順序付けられたＶ［ｋ］ベクトル３５’のｎＦＧフォアグラウンド（または言い換えれば支配的な）成分（しばしばベクトルによって表される）を決定することができる。

【０６４２】
[0642]図３３Ａおよび図３３Ｂは、本開示において説明された技法による、音場のフォアグラウンド成分を記述するベクトルを符号化するときに再順序付けが有する潜在的な影響を示すそれぞれのグラフ４１０Ａおよび４１０Ｂを示す図である。グラフ４１０Ａは、再順序付けられていない（または言い換えれば、オリジナルの）ＵＳ［ｋ］ベクトル３３の少なくともいくつかを符号化した結果を示し、グラフ４１０Ｂは、再順序付けられたＵＳ［ｋ］ベクトル３３’の対応する１つを符号化した結果を示す。グラフ４１０Ａおよび４１０Ｂの各々の一番上のプロットは符号化の誤差を示し、グラフ４１０Ｂにはフレーム境界のみにおいて目立つ誤差がある可能性が高い。したがって、本開示で説明される再順序付け技法は、従来のオーディオコーダを使用したモノラルオーディオオブジェクトのコーディングを容易にし、または別様に促進し得る。

【０６４３】
[0643]図３４および図３５は、本開示による、明瞭なオーディオオブジェクトのエネルギーのみに基づく識別と指向性のみに基づく識別との差を示す概念図である。図３４の例では、より大きいエネルギーを示すベクトルは、指向性とは無関係に明瞭なオーディオオブジェクトであるものとして識別される。図３４に示されるように、（ｙ軸上にプロットされる）より高いエネルギー値に従って配置されるオーディオオブジェクトは、（たとえば、ｘ軸上にプロットされる指向性指数によって表される）指向性とは無関係に、「フォアグラウンドの中にある」と決定される。

【０６４４】
[0644]図３５は、図４の音場分析ユニット４４によって実施される技法などによる、指向性とエネルギーの両方に基づく明瞭なオーディオオブジェクトの識別を示す。図３５に示されるように、より大きい指向性指数はｘ軸の左側に向かってプロットされ、より大きいエネルギーレベルはｙ軸の上側に向かってプロットされる。この例では、音場分析ユニ

ット４４は、明瞭なオーディオオブジェクト（たとえば、「フォアグラウンドの中」にある）が、相対的にグラフの左上に向かってプロットされたベクトルデータと関連付けられると決定することができる。一例として、音場分析ユニット４４は、グラフの左上の象限にプロットされたベクトルが明瞭なオーディオオブジェクトと関連付けられると決定することができる。

【０６４５】

[0645]図３６Ａ～図３６Ｆは、本開示において説明される技法の様々な態様による、補間を実行するための球面調和係数の分解されたバージョンの少なくとも一部分の空間領域への射影を示す図である。図３６Ａは、球面４１２へのＶ［ｋ］ベクトル３５の１つまたは複数の射影を示す図である。図３６Ａの例において、各々の数字は、（場合によってはＶ行列１９’の１つの行および／または列と関連付けられる）球面に射影される異なる球面調和係数を識別する。異なる色は明瞭なオーディオ成分の方向を示唆し、より明るい（および徐々により暗くなる）色は明瞭な成分の主要な方向を示す。図４の例に示されるオーディオ符号化デバイス２０の時間－空間的補間ユニット５０は、図３６Ａの例に示される球面を生成するために、赤い点の各々の間の空間－時間的な補間を実行することができる。

【０６４６】

[0646]図３６Ｂは、Ｖ［ｋ］ベクトル３５の１つまたは複数のビームへの射影を示す図である。空間－時間的補間ユニット５０は、図３６Ｂの例に示されるビーム４１４を生成するために、Ｖ［ｋ］ベクトル３５の１つの行および／もしくは列、またはＶ［ｋ］ベクトル３５の複数の行および／もしくは列を射影することができる。

【０６４７】

[0647]図３６Ｃは、図３６の例に示される球面４１２などの球面への、Ｖ［ｋ］ベクトル３５の１つまたは複数の１つまたは複数のベクトルの射影の断面を示す図である。

【０６４８】

[0648]図３６Ｄ～図３６Ｇに示されるのは、異なる音源（蜂、ヘリコプター、電子音楽、および競技場の中の人々）が３次元空間において示され得るときの、（約２０ミリ秒である１フレームにわたる）時間のスナップショットの例である。

【０６４９】

[0649]本開示において説明される技法は、これらの異なる音源の表現が、単一のＵＳ［ｋ］ベクトルと単一のＶ［ｋ］ベクトルとを使用して識別され表されることを可能にする。音源の時間的な変動性はＵＳ［ｋ］ベクトルにおいて表されるが、各音源の空間的な分布は単一のＶ［ｋ］ベクトルによって表される。１つのＶ［ｋ］ベクトルは、音源の幅と、ロケーションと、サイズとを表し得る。その上、単一のＶ［ｋ］ベクトルは、球面調和基底関数の線形結合として表され得る。図３６Ｄ～図３６Ｇのプロットにおいて、音源の表現は、空間座標系への単一のＶベクトルの変換に基づく。音源を示す同様の方法が、図３６～図３６Ｃにおいて使用される。

【０６５０】

[0650]図３７は、本明細書において説明されるような空間－時間的な補間を得るための技法の表現を示す。図４の例に示されるオーディオ符号化デバイス２０の空間－時間的補間ユニット５０は、以下でより詳細に説明される空間－時間的な補間を実行することができる。空間－時間的な補間は、空間次元と時間次元の両方において、より高分解能の空間成分を取得することを含み得る。空間成分は、高次アンビソニックス（ＨＯＡ）係数（または、ＨＯＡ係数は「球面調和係数」とも呼ばれ得るので）からなる多次元信号の直交方向の分解に基づき得る。

【０６５１】

[0651]示されるグラフにおいて、ベクトルV_1およびV_2は、多次元信号の２つの異なる空間成分の対応するベクトルを表す。空間成分は、多次元信号のブロックごとの分解によって取得され得る。いくつかの例では、空間成分は、高次アンビソニックス（ＨＯＡ）オーディオデータ（このアンビソニックスオーディオデータは、ブロック、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）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ＳＶＤを実行することにより得られる。サンプル中のオーディオフレームの長さを示すために、変数Ｍが使用され得る。

【０６５２】

[0652]したがって、V_1およびV_2は、ＨＯＡ係数１１のシーケンシャルなブロックのフォアグラウンドＶ［ｋ］ベクトル51_kとフォアグラウンドＶ［ｋ－１］ベクトル51_{k-1}の対応するベクトルを表し得る。V_1は、たとえば、第１のフレーム（ｋ－１）に対するフォアグラウンドＶ［ｋ－１］ベクトル51_{k-1}の第１のベクトルを表すが、V_2は、第２の後続のフレーム（ｋ）に対するフォアグラウンドＶ［ｋ］ベクトル51_kの第１のベクトルを表し得る。V_1およびV_2は、多次元信号に含まれる単一のオーディオオブジェクトの空間成分を表し得る。

【０６５３】

[0653]各々のｘに対する補間されたベクトルV_xが、時間セグメントまたは「時間サンプル」の番号ｘに従ってV_1とV_2とを重み付けることによって、時間成分（およびしたがって、いくつかの場合には空間成分）を平滑化するために補間されたベクトルV_xが適用され得る多次元信号の時間成分に対して取得される。上で説明されたようなＳＶＤの分解を仮定すると、ｎＦＧ信号４９の平滑化は、対応する補間されたV_xによる各時間サンプルベクトル（たとえば、ＨＯＡ係数１１のサンプル）のベクトル分割を行うことによって取得され得る。すなわち、ＵＳ［ｎ］＝ＨＯＡ［ｎ］×V_x［ｎ］$^{-1}$であり、これは、列ベクトルによって乗算される行ベクトルを表すので、ＵＳのスカラー要素を生み出す。V_x［ｎ］$^{-1}$は、V_x［ｎ］の擬似逆ベクトルとして取得され得る。

【０６５４】

[0654]V_1およびV_2の重み付けに関して、V_1は、V_2がV_1よりも時間的に後に発生することが原因で、時間次元に沿って比例的により低く重み付けられる。すなわち、フォアグラウンドＶ［ｋ－１］ベクトル51_{k-1}は分解の空間成分であるが、時間的にシーケンシャルなフォアグラウンドＶ［ｋ］ベクトル51_kは時間にわたる空間成分の異なる値を表す。したがって、ｘがｔに沿って増大するにつれて、V_1の重みは低下するが、V_2の重みは上昇する。ここで、d_1およびd_2は重みを表す。

【０６５５】

[0655]図３８は、本明細書で説明される技法による、多次元信号のためのシーケンシャルなＳＶＤブロックのための人工的なＵＳ行列、すなわちUS_1とUS_2とを示すブロック図である。補間されたＶベクトルが、オリジナルの多次元信号を復元するために、人工的なＵＳベクトルの行ベクトルに適用され得る。より具体的には、空間－時間的補間ユニット５０は、Ｋ／２個の補間されたサンプルを取得するために、ｎＦＧ信号４９とフォアグラウンドＶ［ｋ］ベクトル51_k（フォアグラウンドＨＯＡ係数とも示され得る）を乗算した結果に、補間されたフォアグラウンドＶ［ｋ］ベクトル５３の擬似逆ベクトルを乗算することができ、このＫ／２個の補間されたサンプルは、U_2行列の図３８の例に示されるような最初のＫ／２個のサンプルとして、ｎＦＧ信号のＫ／２個のサンプルの代わりに使用され得る。

【０６５６】

[0656]図３９は、本開示において説明される技法による、特異値分解と空間－時間的成分の平滑化とを使用した高次アンビソニックス（ＨＯＡ）信号の後続のフレームの分解を示すブロック図である。フレームｎ－１およびフレームｎ（フレームｎおよびフレームｎ＋１としても示され得る）は、時間的に後続のフレームを表し、各フレームは１０２４個の時間セグメントを備えて４というＨＯＡ次数を有し、（４＋１）^{2}＝２５個の係数を与える。フレームｎ－１およびフレームｎにおいて人工的に平滑化されるＵ行列であるＵＳ行列は、示されるような補間されたＶベクトルの適用によって取得され得る。各々の灰色の行ベクトルまたは列ベクトルは、１つのオーディオオブジェクトを表す。

【０６５７】

[0657]アクティブなベクトルベースの信号のＨＯＡ表現を計算する

[0658]ＸＶＥＣｋで表されるベクトルベースの信号の各々を取り出し、それを対応する（逆量子化された）空間ベクトルＶＶＥＣｋと乗算することによって、瞬時的なＣＶＥＣｋが作り出される。各ＶＶＥＣｋはＭＶＥＣｋで表される。したがって、次数ＬのＨＯＡ信号およびＭ個のベクトルベースの信号に対して、各々がフレーム長Ｐによって与えられる次元を有する、Ｍ個のベクトルベースの信号がある。これらの信号はしたがって、ＸＶＥＣｋｍｎ、ｎ＝０，．．．Ｐ－１；ｍ＝０，．．．Ｍ－１として表され得る。これに対応して、Ｍ個の空間ベクトルと、次元（Ｌ＋１）^{2}のＶＶＥＣｋがある。これらは、ＭＶＥＣｋｍｌとして表されることが可能であり、ｌ＝０，．．．，（Ｌ＋１）$^{2-1}$；ｍ＝０，．．．，Ｍ－１である。各ベクトルベースの信号のＨＯＡ表現ＣＶＥＣｋｍは、以下の式によって与えられる行列ベクトル乗算である：

【０６５８】

【数２４】

$$CVECkm=(XVECkm(MVECkm)T)T$$

【０６５９】

これは、（Ｌ＋１）^{2}対Ｐの行列を生み出す。完全なＨＯＡ表現は、次のように各ベクトルベースの信号の寄与を加算することによって与えられる：

【０６６０】

【数２５】

$$CVECk=m=0M-1CVECk[m]$$

【０６６１】

[0659]Ｖベクトルの空間－時間的な補間

しかしながら、滑らかな空間－時間的な継続性を維持するために、上の計算は、フレーム長の一部であるＰ－Ｂのみに対して実行される。ＨＯＡ行列の最初のＢ個のサンプルは代わりに、ｍ＝０，．．．，Ｍ－１；ｌ＝０，．．．，（Ｌ＋１）^{2}として、現在のＭＶＥＣｋｍおよび以前の値ＭＶＥＣｋ－１ｍから導出される、ＭＶＥＣｋｍｌの補間されたセットを使用して実行される。これは、次のように各時間サンプルｐに対するベクトルを導出すると、より高い時間密度の空間ベクトルをもたらす：

【０６６２】

【数２６】

$$MVECkmp=pB-1MVECkm+B-1-pB-1MVECk-1m,\ p=0,..,B-1$$

【０６６３】

各時間サンプルｐに対して、（Ｌ＋１）２次元の新たなＨＯＡベクトルが次のように計算される：

【０６６４】

【数２７】

$$CVECkp=(XVECkmp)MVECkmp,\ p=0,..,B-1$$

【０６６５】

これらの最初のＢ個のサンプルは、以前のセクションのＰ－Ｂ個のサンプルによって補強され、ｍ番目のベクトルベースの信号の完全なＨＯＡ表現であるＣＶＥＣｋｍをもたらす。

【０６６６】
[0660]復号器（たとえば、図５の例に示されるオーディオ復号デバイス２４）において、ある明瞭な、フォアグラウンドの、またはベクトルベースの支配的な音声に対して、以前のフレームからのＶベクトルおよび現在のフレームからのＶベクトルは、特定の時間セグメントにわたってより高分解能の（時間的に）補間されたＶベクトルを生成するために、線形（または非線形）補間を使用して補間され得る。空間－時間的補間ユニット７６はこの補間を実行することができ、ここで空間－時間的補間ユニット７６は次いで、その特定の時間セグメントにわたってＨＯＡ行列を生成するために、現在のフレームの中のＵＳベクトルをより高分解能の補間されたＶベクトルと乗算することができる。
【０６６７】
[0661]あるいは、空間－時間的補間ユニット７６は、第１のＨＯＡ行列を作成するために、ＵＳベクトルを現在のフレームのＶベクトルと乗算することができる。復号器は追加で、第２のＨＯＡ行列を作成するために、ＵＳベクトルを以前のフレームからのＶベクトルと乗算することができる。空間－時間的補間ユニット７６は次いで、ある特定の時間セグメントにわたって、線形（または非線形）補間を第１および第２のＨＯＡ行列に適用することができる。この補間の出力は、共通の入力行列／ベクトルを仮定すると、補間されたＶベクトルとのＵＳベクトルの乗算の出力と一致し得る。
【０６６８】
[0662]この点において、本技法は、オーディオ符号化デバイス２０および／またはオーディオ復号デバイス２４が、以下の条項に従って動作するように構成されることを可能にし得る。
【０６６９】
[0663]項１３５０５４－１Ｃ。空間と時間の両方において複数のより高分解能の空間成分を取得するように構成される１つまたは複数のプロセッサを備え、空間成分が球面調和係数からなる多次元信号の直交方向の分解に基づく、オーディオ符号化デバイス２０またはオーディオ復号デバイス２４などのデバイス。
【０６７０】
[0664]項１３５０５４－１Ｄ。第１の複数の球面調和係数および第２の複数の球面調和係数の空間成分と時間成分の少なくとも１つを平滑化するように構成される１つまたは複数のプロセッサを備える、オーディオ符号化デバイス２０またはオーディオ復号デバイス２４などのデバイス。
【０６７１】
[0665]項１３５０５４－１Ｅ。空間と時間の両方において複数のより高分解能の空間成分を取得するように構成される１つまたは複数のプロセッサを備え、空間成分が球面調和係数からなる多次元信号の直交方向の分解に基づく、オーディオ符号化デバイス２０またはオーディオ復号デバイス２４などのデバイス。
【０６７２】
[0666]項１３５０５４－１Ｇ。少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して分解能を増大させることによって、時間セグメントのための分解され解像度の増大した球面調和係数を取得するように構成される１つまたは複数のプロセッサを備える、オーディオ符号化デバイス２０またはオーディオ復号デバイス２４などのデバイス。
【０６７３】
[0667]項１３５０５４－２Ｇ。第１の分解が、第１の複数の球面調和係数の右特異ベクトルを表す第１のＶ行列を備える、項１３５０５４－１Ｇのデバイス。
【０６７４】
[0668]項１３５０５４－３Ｇ。第２の分解が、第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える、項１３５０５４－１Ｇのデバイス。
【０６７５】
[0669]項１３５０５４－４Ｇ。第１の分解が第１の複数の球面調和係数の右特異ベクト

ルを表す第１のＶ行列を備え、第２の分解が第２の複数の球面調和係数の右特異ベクトルを表す第２のＶ行列を備える、項１３５０５４－１Ｇのデバイス。

【０６７６】

[0670]項１３５０５４－５Ｇ。時間セグメントがオーディオフレームのサブフレームを備える、項１３５０５４－１Ｇのデバイス。

【０６７７】

[0671]項１３５０５４－６Ｇ。時間セグメントがオーディオフレームの時間サンプルを備える、項１３５０５４－１Ｇのデバイス。

【０６７８】

[0672]項１３５０５４－７Ｇ。１つまたは複数のプロセッサが、第１の複数の球面調和係数の球面調和係数に対する第１の分解および第２の分解の補間された分解を取得するように構成される、項１３５０５４－１Ｇのデバイス。

【０６７９】

[0673]項１３５０５４－８Ｇ。１つまたは複数のプロセッサが、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分に対する第１の分解および第２のフレームに含まれる第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を取得するように構成され、１つまたは複数のプロセッサがさらに、第１の複数の球面調和係数の第１の人工時間成分を生成するために、第１のフレームに含まれる第１の複数の球面調和係数の第１の部分の第１の時間成分に補間された分解を適用し、第２の複数の球面調和係数の第２の人工時間成分を生成するために、第２のフレームに含まれる第２の複数の球面調和係数の第２の部分の第２の時間成分にそれぞれの補間された分解を適用するように構成される、項１３５０５４－１Ｇのデバイス。

【０６８０】

[0674]項１３５０５４－９Ｇ。第１の時間成分が、第１の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、項１３５０５４－８Ｇのデバイス。

【０６８１】

[0675]項１３５０５４－１０Ｇ。第２の時間成分が、第２の複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、項１３５０５４－８Ｇのデバイス。

【０６８２】

[0676]項１３５０５４－１１Ｇ。１つまたは複数のプロセッサがさらに、第１の人工時間成分と第２の人工時間成分とを受信し、第１の複数の球面調和係数の第１の部分に対する第１の分解および第２の複数の球面調和係数の第２の部分に対する第２の分解の補間された分解を計算し、第１の時間成分を復元するために補間された分解の逆数を第１の人工時間成分を適用し、第２の時間成分を復元するために補間された分解の逆数を第２の人工時間成分を適用するように構成される、項１３５０５４－８Ｇのデバイス。

【０６８３】

[0677]項１３５０５４－１２Ｇ。１つまたは複数のプロセッサが、第１の複数の球面調和係数の第１の空間成分と第２の複数の球面調和係数の第２の空間成分とを補間するように構成される、項１３５０５４－１Ｇのデバイス。

【０６８４】

[0678]項１３５０５４－１３Ｇ。第１の空間成分が、第１の複数の球面調和係数の左特異ベクトルを表す第１のＵ行列を備える、項１３５０５４－１２Ｇのデバイス。

【０６８５】

[0679]項１３５０５４－１４Ｇ。第２の空間成分が、第２の複数の球面調和係数の左特異ベクトルを表す第２のＵ行列を備える、項１３５０５４－１２Ｇのデバイス。

【０６８６】

[0680]項１３５０５４－１５Ｇ。第１の空間成分が第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分が第２の複数の球面調和係

数に対する球面調和係数のＭ個の時間セグメントを表す、項１３５０５４－１２Ｇのデバイス。

【０６８７】

[0681]項１３５０５４－１６Ｇ。第１の空間成分が第１の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、第２の空間成分が第２の複数の球面調和係数に対する球面調和係数のＭ個の時間セグメントを表し、１つまたは複数のプロセッサが時間セグメントに対する分解された補間された球面調和係数を取得するように構成されるが第１の空間成分の最後のＮ個の要素と第２の空間成分の最初のＮ個の要素とを補間することを備える、項１３５０５４－１２Ｇのデバイス。

【０６８８】

[0682]項１３５０５４－１７Ｇ。第２の複数の球面調和係数が、時間領域において第１の複数の球面調和係数に後続する、項１３５０５４－１Ｇのデバイス。

【０６８９】

[0683]項１３５０５４－１８Ｇ。１つまたは複数のプロセッサがさらに、第１の複数の球面調和係数の第１の分解を生成するために、第１の複数の球面調和係数を分解するように構成される、項１３５０５４－１Ｇのデバイス。

【０６９０】

[0684]項１３５０５４－１９Ｇ。１つまたは複数のプロセッサがさらに、第２の複数の球面調和係数の第２の分解を生成するために、第２の複数の球面調和係数を分解するように構成される、項１３５０５４－１Ｇのデバイス。

【０６９１】

[0685]項１３５０５４－２０Ｇ。１つまたは複数のプロセッサがさらに、第１の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第１の複数の球面調和係数の特異値を表すＳ行列と、第１の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第１の複数の球面調和係数に関して特異値分解を実行するように構成される、項１３５０５４－１Ｇのデバイス。

【０６９２】

[0686]項１３５０５４－２１Ｇ。１つまたは複数のプロセッサがさらに、第２の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第２の複数の球面調和係数の特異値を表すＳ行列と、第２の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第２の複数の球面調和係数に関して特異値分解を実行するように構成される、項１３５０５４－１Ｇのデバイス。

【０６９３】

[0687]項１３５０５４－２２Ｇ。第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、音場の平面波表現を表す、項１３５０５４－１Ｇのデバイス。

【０６９４】

[0688]項１３５０５４－２３Ｇ。第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、一緒にミキシングされた１つまたは複数のモノラルオーディオオブジェクトを表す、項１３５０５４－１Ｇのデバイス。

【０６９５】

[0689]項１３５０５４－２４Ｇ。第１の複数の球面調和係数と第２の複数の球面調和係数の各々が、３次元音場を表すそれぞれの第１の球面調和係数および第２の球面調和係数を備える、項１３５０５４－１Ｇのデバイス。

【０６９６】

[0690]項１３５０５４－２５Ｇ。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、１よりも大きい次数を有する少なくとも１つの球面基底関数と関連付けられる、項１３５０５４－１Ｇのデバイス。

【０６９７】

[0691]項１３５０５４－２６Ｇ。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、４に等しい次数を有する少なくとも１つの球面基底関数と関連付けられる

、項１３５０５４－１Ｇのデバイス。
【０６９８】
[0692]項１３５０５４－２７Ｇ。補間が第１の分解および第２の分解の重み付けられた補間であり、第１の分解に適用された重み付けられた補間の重みが、第１の分解および第２の分解のベクトルによって表される時間に反比例し、第２の分解に適用される重み付けられた補間の重みが、第１の分解および第２の分解のベクトルによって表される時間に比例する、項１３５０５４－１Ｇのデバイス。
【０６９９】
[0693]項１３５０５４－２８Ｇ。分解された補間された球面調和係数が、第１の複数の球面調和係数および第２の複数の球面調和係数の空間成分と時間成分の少なくとも１つを平滑化する、項１３５０５４－１Ｇのデバイス。
【０７００】
[0695]図４０Ａ～図４０Ｊは各々、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法の様々な態様を実行し得る、例示的なオーディオ符号化デバイス５１０Ａ～５１０Ｊを示すブロック図である。図４０Ａ～図４０Ｊの例の各々において、オーディオ符号化デバイス５１０Ａおよび５１０Ｂは各々、いくつかの例では、デスクトップコンピュータ、ラップトップコンピュータ、ワークステーション、タブレットまたはスレートコンピュータ、専用オーディオ記録デバイス、携帯電話（いわゆる「スマートフォン」を含む）、個人用メディアプレーヤデバイス、個人用ゲーミングデバイス、またはオーディオデータを符号化することが可能な任意の他のタイプのデバイスを表す。
【０７０１】
[0696]単一のデバイス、すなわち図４０Ａ～図４０Ｊの例におけるデバイス５１０Ａ～５１０Ｊとして示されているが、デバイス５１０Ａ～５１０Ｊに含まれるものとして以下で参照される様々なコンポーネントまたはユニットは、実際にはデバイス５１０Ａ～５１０Ｊの外部の別個のデバイスを形成し得る。言い換えれば、単一のデバイス、すなわち図４０Ａ～図４０Ｊの例におけるデバイス５１０Ａ～５１０Ｊによって実行されるものとして本開示では説明されているが、本技法は、複数のデバイスを備えるシステムによって実施され、または別様に実行されてよく、これらのデバイスの各々は、以下でより詳細に説明される様々なコンポーネントまたはユニットの１つまたは複数を各々含み得る。したがって、本技法は図４０Ａ～図４０Ｊの例に限定されるべきではない。
【０７０２】
[0697]いくつかの例では、オーディオ符号化デバイス５１０Ａ～５１０Ｊは、図３および図４の例に関して上で説明されたものに対する代替的なオーディオ符号化デバイスを表す。オーディオ符号化デバイス５１０Ａ～５１０Ｊの以下の議論全体で、動作に関する様々な類似性が、図４に関して上で説明されたオーディオ符号化デバイス２０の様々なユニット３０～５２に関して言及される。多くの点で、オーディオ符号化デバイス５１０Ａ～５１０Ｊは、以下で説明されるように、わずかな誘導または修正を伴っても、オーディオ符号化デバイス２０と実質的に同様の方式で動作することができる。
【０７０３】
[0698]図４０Ａの例に示されるように、オーディオ符号化デバイス５１０Ａは、オーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを備える。オーディオ圧縮ユニット５１２は、球面調和係数（ＳＨＣ）５１１（「ＳＨＣ５１１」）を圧縮するユニットを表すことができ、ＳＨＣ５１１は高次アンビソニックス（ＨＯＡ）係数５１１としても示され得る。オーディオ圧縮ユニット５１２メイいくつかの例では、オーディオ圧縮ユニット５１２は、ＳＨＣ５１１に関して無損失で圧縮することができる、または有損失の圧縮を実行できるユニットを表す。ＳＨＣ５１１は複数のＳＨＣを表すことができ、複数のＳＨＣの少なくとも１つが、上でより詳細に説明されるように、１より大きい次数を有する球面基底関数に対応する（この種類のＳＨＣは、１つの例がいわゆる「Ｂフォーマット」である低次アンビソニックスと区別す

るために、高次アンビソニックス（ＨＯＡ）と呼ばれる）。オーディオ圧縮ユニット５１２はＳＨＣ５１１を無損失で圧縮することができるが、いくつかの例では、オーディオ圧縮ユニット５１２は、ＳＨＣ５１１のうちで、再生成されるときに音場を記述する際に（一部が人の聴覚系によって聞かれることが可能ではないという点で）顕著ではないまたは関連のないものを除去する。この点で、この圧縮の有損失の性質は、ＳＨＣ５１１の圧縮されたバージョンから再生成されるときに音場の知覚される品質に過剰に影響を与えないことがある。

【０７０４】

[0699]図４０Ａの例において、オーディオ圧縮ユニットは、分解ユニット５１８と音場成分抽出ユニット５２０とを含む。分解ユニット５１８は、オーディオ符号化デバイス２０の線形可逆変換ユニット３０と同様であり得る。すなわち、分解ユニット５１８は、特異値分解と呼ばれる形態の分析を実行するように構成されるユニットを表し得る。ＳＶＤに関して説明されているが、技法は、線形的に無相関なデータ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、別段特別に述べられない限り「０ではない」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。

【０７０５】

[0700]いずれにしても、分解ユニット５１８は、球面調和係数５１１を変換された球面調和係数の２つ以上のセットへと変換するために、特異値分解（やはり頭字語「ＳＶＤ」によって示され得る）を実行する。図４０の例では、分解ユニット５１８は、いわゆるＶ行列５１９と、Ｓ行列５１９Ｂと、Ｕ行列５１９Ｃとを生成するために、ＳＨＣ５１１に関してＳＶＤを実行することができる。図４０の例では、分解ユニット５１８は、線形可逆変換ユニット３０に関して上で論じられたような組み合わされた形態でＵＳ［ｋ］ベクトルを出力するのではなく、行列の各々を別々に出力する。

【０７０６】

[0701]上で述べられたように、上で参照されたＳＶＤの数学的表現の中のＶ*行列は、ＳＶＤが複素数を備える行列に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行列に等しいと見なされてよい。説明を簡単にするために、ＳＨＣ５１１が実数を備え、その結果Ｖ行列がＶ*行列ではなくＳＶＤを通じて出力されることが以下で仮定される。Ｖ行列であると仮定されるが、本技法は、複素係数を有するＳＨＣ５１１に同様の方式で適用されてよく、このときＳＶＤの出力はＶ*行列である。したがって、本技法はこの点で、Ｖ行列を生成するためにＳＶＤ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ＳＨＣ５１１へＳＶＤを適用することを含み得る。

【０７０７】

[0702]いずれにしても、分解ユニット５１８は、高次アンビソニックス（ＨＯＡ）オーディオデータ（このアンビソニックスオーディオデータは、ＳＨＣ５１１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）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ＳＶＤを実行することができる。サンプル中のオーディオフレームの長さを示すために、変数Ｍが使用され得る。たとえば、オーディオフレームが１０２４個のオーディオサンプルを含むとき、Ｍは１０２４に等しい。分解ユニット５１８はしたがって、Ｍ対$(N+1)^2$のＳＨＣを有するブロックＳＨＣ５１１に関してブロックごとのＳＶＤを実行することができ、ＮはやはりＨＯＡオーディオデータの次数を示す。分解ユニット５１８は、このＳＶＤの実行を通じて、Ｖ行列５１９と、Ｓ行列５１９Ｂと、Ｕ行列５１９Ｃとを生成することができ、行列５１９～５１９Ｃの各々（「行列５１９」）は、上でより詳細に説明されたそれぞれのＶ行列、Ｓ行列、およびＵ行列を表し得る。分解ユニット５１８は、これらの行列５１９Ａを音場成分抽出ユニット５２０に渡し、または出力することができる。Ｖ行列５１９Ａは$(N+1)^2$対$(N+1)^2$のサイズであってよく、Ｓ行列５１９Ｂは$(N+1)^2$対（Ｎ＋１

）^{2}のサイズであってよく、Ｕ行列はＭ対（Ｎ＋１）^{2}のサイズであってよく、Ｍはオーディオフレーム中のサンプルの数を指す。Ｍの典型的な値は１０２４であるが、本開示の技法はＭのこの典型的な値に限定されるべきではない。

【０７０８】

[0703]音場成分抽出ユニット５２０は、音場の明瞭な成分と音場のバックグラウンド成分とを決定し、次いで抽出して、音場のバックグラウンド成分から音場の明瞭な成分を実質的に分離するように構成されるユニットを表し得る。この点において、音場成分抽出ユニット５２０は、図４の例に示されるオーディオ符号化デバイス２０の音場分析ユニット４４、バックグラウンド選択ユニット４８、およびフォアグラウンド選択ユニット３６に関して上で説明された動作の多くを実行することができる。いくつかの例において、音場の明瞭な成分が、これらの成分の明瞭な性質を正確に表すために、（音場のバックグラウンド成分に対して相対的に）高次の基底関数（したがってより多くのＳＨＣ）を必要とするとすると、バックグラウンド成分から明瞭な成分を分離することで、より多くのビットが明瞭な成分に割り振られることと、（相対的に言って）より少数のビットがバックグラウンド成分に割り振られることとが可能になり得る。したがって、（ＳＶＤの形態で、またはＰＣＡを含む任意の他の形態の変換で）この変換の適用を通じて、本開示で説明される技法は、様々なＳＨＣへのビットの割振りを容易にすることができ、これによってＳＨＣ１１の圧縮を容易にすることができる。

【０７０９】

[0704]その上、本技法はまた、図４０Ｂに関して以下でより詳細に説明されるように、いくつかの例において、音場のバックグラウンド成分の拡散的な性質またはバックグラウンド的な性質を仮定した場合に、音場のバックグラウンド部分を表すために高次の基底関数が必要とされないとすると、音場のバックグラウンド成分の次数低減を可能にし得る。本技法はしたがって、音場の拡散的な側面またはバックグラウンド的な側面の圧縮を可能にしながら、ＳＨＣ５１１へのＳＶＤの適用を通じて音場の顕著な明瞭な成分または側面を保存することができる。

【０７１０】

[0705]図４０の例にさらに示されるように、音場成分抽出ユニット５２０は、転置ユニット５２２と、顕著成分分析ユニット５２４と、算出ユニット５２６とを含む。転置ユニット５２２は、「V^T行列５２３」として示されるＶ行列５１９の転置を生成するためにＶ行列５１９Ａを転置するように構成されるユニットを表す。転置ユニット５２２は、このV^T行列５２３を算出ユニット５２６に出力することができる。V^T行列５２３は、（Ｎ＋１）^{2}対（Ｎ＋１）^{2}のサイズであり得る。

【０７１１】

[0706]顕著成分分析ユニット５２４は、Ｓ行列５１９Ｂに関して顕著さの分析を実行するように構成されるユニットを表す。顕著成分分析ユニット５２４は、この点において、図４の例に示されるオーディオ符号化デバイス２０の音場分析ユニット４４に関して上で説明されたものと同様の動作を実行することができる。顕著成分分析ユニット５２４は、Ｓ行列５１９Ｂの対角線上の値を分析して、最大の値を有するこれらの成分の数である変数Ｄを選択することができる。言い換えれば、顕著成分分析ユニット５２４は、Ｓの降順の対角線上の値によって作り出される曲線の傾きを分析することによって、２つのサブ空間（たとえば、フォアグラウンドサブ空間または支配的なサブ空間およびバックグラウンドサブ空間または環境サブ空間）を分離する値Ｄを決定することができ、ここで大きい特異値はフォアグラウンド音声または明瞭な音声を表し、小さい特異値は音場のバックグラウンド成分を表す。いくつかの例では、顕著成分分析ユニット５２４は、特異値曲線の一次導関数と二次導関数とを使用することができる。顕著成分分析ユニット５２４はまた、１と５の間となるように数Ｄを制限することができる。別の例として、顕著成分分析ユニット５２４は、１と（Ｎ＋１）^{2}の間となるように数Ｄを制限することができる。あるいは、顕著成分分析ユニット５２４は、４という値などに、値Ｄを事前に定めることができる。いずれにしても、数Ｄが推定されると、顕著成分分析ユニット２４は、行列Ｕ、Ｖ、

およびＳからフォアグラウンドサブ空間とバックグラウンドサブ空間とを抽出する。

【０７１２】

[0707]いくつかの例では、顕著成分分析ユニット５２４は、この分析をＭ個のサンプルごとに実行することができ、これはフレームごとと言い換えられ得る。この点において、Ｄはフレームごとに変化し得る。他の例では、顕著成分分析ユニット２４は、この分析をフレームごとに２回以上実行して、フレームの２つ以上の部分を分析することができる。したがって、本技法は、この点で、本開示で説明されている例に限定されるべきではない。

【０７１３】

[0708]つまり、顕著成分分析ユニット５２４は、図４０の例においてＳ行列５１９Ｂとして示される対角行列の特異値を分析し、対角Ｓ行列５１９Ｂの他の値よりも相対的に大きい値を有する値を識別することができる。顕著成分分析ユニット５２４は、Ｄの値を識別し、S_{DIST}行列５２５ＡとS_{BG}行列５２５Ｂとを生成するためにこれらの値を抽出することができる。S_{DIST}行列５２５Ａは、オリジナルのＳ行列５１９Ｂの$(N+1)^2$を有するＤ個の列を備える対角行列を表し得る。いくつかの例では、S_{BG}行列５２５Ｂは、各々がオリジナルのＳ行列５１９Ｂの$(N+1)^2$個の変換された球面調和係数を含む$(N+1)^2-D$個の列を有する行列を表し得る。オリジナルのＳ行列５１９Ｂの$(N+1)^2$個の値を有するＤ個の列を備える行列を表すS_{DIST}行列として説明されるが、Ｓ行列５１９Ｂが対角行列であり、Ｄ個の列の$(N+1)^2$個の値のうちで各列の中のＤ番目の値より後のものが０という値であることが多いとすると、顕著成分分析ユニット５２４は、この行列を切り取って、オリジナルのＳ行列５１９ＢのＤ個の値を有するＤ個の列を有するS_{DIST}行列を生成することができる。完全なS_{DIST}行列５２５Ａおよび完全なS_{BG}行列５２５Ｂに関して説明されたが、本技法は、これらのS_{DIST}行列５２５Ａの切り取られたバージョンおよびこのS_{BG}行列５２５Ｂの切り取られたバージョンに関して実施され得る。したがって、本開示の技法はこの点において限定されるべきではない。

【０７１４】

[0709]言い換えれば、S_{DIST}行列５２５ＡはＤ対$(N+1)^2$のサイズであり得るが、S_{BG}行列５２５Ｂは$(N+1)^2-D$対$(N+1)^2$のサイズであり得る。S_{DIST}行列５２５Ａは、音場の明瞭な（ＤＩＳＴ）オーディオ成分であるということに関して顕著であると決定される主要成分、または言い換えれば特異値を含み得るが、S_{BG}行列５２５Ｂは、音場のバックグラウンド（ＢＧ）成分、または言い換えれば環境成分または不明瞭なオーディオ成分であると決定される特異値を含み得る。図４０の例では別々の行列５２５Ａおよび５２５Ｂとして示されるが、行列５２５Ａおよび５２５Ｂは、変数Ｄを使用して単一の行列として指定され、S_{DIST}行列５２５を表すこの単一の行列の（左から右への）列の数を示すことができる。いくつかの例では、変数Ｄは４に設定され得る。

【０７１５】

[0710]顕著成分分析ユニット５２４はまた、U_{DIST}行列５２５ＣとU_{BG}行列５２５Ｄとを生成するためにＵ行列５１９Ｃを分析することができる。しばしば、顕著成分分析ユニット５２４は、変数Ｄを識別するためにＳ行列５１９Ｂを分析して、変数Ｄに基づいてU_{DIST}行列５２５ＣとU_{BG}行列５２５Ｂとを生成することができる。すなわち、顕著であるＳ行列５１９ＢのＤ個の列を識別した後で、顕著成分分析ユニット５２４は、この決定された変数Ｄに基づいてＵ行列５１９Ｃを分割することができる。この例において、顕著成分分析ユニット５２４は、Ｕ行列５１９Ｃの$(N+1)^2$個の変換された球面調和係数の（左から右への）Ｄ個の列を含むようにU_{DIST}行列５２５Ｃを生成し、オリジナルのＵ行列５１９Ｃの$(N+1)^2$個の変換された球面調和係数の残りの$(N+1)^2-D$個の列を含むようにU_{BG}行列５２５Ｄを生成することができる。U_{DIST}行列５２５ＣはＭ対Ｄのサイズであり得るが、U_{BG}行列５２５ＤはＭ対$(N+1)^2-D$のサイズであり得る。図４０の例では別々の行列５２５Ｃおよび５２５Ｄであるものとして示されるが、行列５２５Ｃおよび５２５Ｄは、変数Ｄを使用して単一の行列として指定され、U_{DIST}行列５２５Ｂを表すこの単一の行列の（左から右への）列の数を示すことができる。

【０７１６】

[0711]顕著成分分析ユニット５２４はまた、V^T_{DIST}行列５２５ＥとV^T_{BG}行列５２５Ｆとを生成するためにV^T行列５２３を分析することができる。しばしば、顕著成分分析ユニット５２４は、変数Ｄを識別するためにＳ行列５１９Ｂを分析して、変数Ｄに基づいてV^T_{DIST}行列５２５ＥとV_{BG}行列５２５Ｆとを生成することができる。すなわち、顕著であるＳ行列５１９ＢのＤ個の列を識別した後で、顕著成分分析ユニット２５４は、この決定された変数Ｄに基づいてＶ行列５１９Ａを分割することができる。この例では、顕著成分分析ユニット５２４は、オリジナルのV^T行列５２３のＤ個の値の（上から下への）$(N+1)^2$個の行を含むようにV^T_{DIST}行列５２５Ｅを生成し、オリジナルのV^T行列５２３の$(N+1)^2-D$個の値の残りの$(N+1)^2$個の行を含むようにV^T_{BG}行列５２５Ｆを生成することができる。V^T_{DIST}行列５２５Ｅは$(N+1)^2$対Ｄのサイズであり得るが、V^T_{BG}行列５２５Ｄは$(N+1)^2$対$(N+1)^2-D$のサイズであり得る。図４０の例では別々の行列５２５Ｅおよび５２５Ｆであるものとして示されるが、行列５２５Ｅおよび５２５Ｆは、変数Ｄを使用して単一の行列として指定され、V_{DIST}行列５２５Ｅを表すこの単一の行列の（左から右への）列の数を示すことができる。顕著成分分析ユニット５２４は、S_{DIST}行列５２５と、S_{BG}行列５２５Ｂと、U_{DIST}行列５２５Ｃと、U_{BG}行列５２５Ｄと、V^T_{BG}行列５２５Ｆとを算出ユニット５２６に出力しながら、V^T_{DIST}行列５２５Ｅをビットストリーム生成ユニット５１６に出力することもできる。

【０７１７】

[0712]算出ユニット５２６は、行列乗算、または、１つまたは複数の行列（またはベクトル）に関して実行されることが可能な任意の他の数学的演算を実行するように構成されるユニットを表し得る。より具体的には、図４０の例に示されるように、算出ユニット５２６は、U_{DIST}行列５２５ＣをS_{DIST}行列５２５Ａと乗算してＭ対ＤのサイズのU_{DIST}ベクトル＊S_{DIST}ベクトル５２７を生成するために行列乗算を実行するように構成されるユニットを表し得る。行列算出ユニット５２６はまた、U_{BG}行列５２５ＤをS_{BG}行列５２５Ｂと乗算し、次いでV^T_{BG}行列５２５Ｆと乗算して、$U_{BG}＊S_{BG}＊V^T_{BG}$行列５２５Ｆを生成し、サイズＭ対$(N+1)^2$のバックグラウンド球面調和係数５３１（これは、球面調和係数５１１のうちで音場のバックグラウンド成分を表すものを表し得る）を生成するために、行列乗算を実行するように構成されるユニットを表し得る。算出ユニット５２６は、U_{DIST}ベクトル＊S_{DIST}ベクトル５２７とバックグラウンド球面調和係数５３１とをオーディオ符号化ユニット５１４に出力することができる。

【０７１８】

[0713]オーディオ符号化デバイス５１０はしたがって、符号化プロセスの終わりにおける行列乗算を通じてU_{DIST}ベクトル＊S_{DIST}ベクトル５２７とバックグラウンド球面調和係数５３１とを生成するように構成されるこの算出ユニット５２６を含むという点で、オーディオ符号化デバイス２０とは異なる。オーディオ符号化デバイス２０の線形可逆変換ユニット３０は、符号化プロセスの比較的初めにおいてＵ行列とＳ行列の乗算を実行してＵＳ［ｋ］ベクトル３３を出力し、このことは、図４０の例に示されない再順序付けなどのより後の動作を容易にし得る。その上、オーディオ符号化デバイス２０は、符号化プロセスの終わりにおいてバックグラウンドＳＨＣ５３１を復元するのではなく、ＨＯＡ係数１１から直接バックグラウンドＨＯＡ係数４７を選択し、これによって、バックグラウンドＳＨＣ５３１を復元するための行列演算を場合によっては回避する。

【０７１９】

[0714]オーディオ符号化ユニット５１４は、U_{DIST}ベクトル＊S_{DIST}ベクトル５２７とバックグラウンド球面調和係数５３１とをさらに圧縮するために、ある形態の符号化を実行するユニットを表し得る。オーディオ符号化ユニット５１４は、図４の例に示されるオーディオ符号化デバイス２０の音響心理学的オーディオコーダユニット４０と実質的に同様の方式で動作することができる。いくつかの例では、このオーディオ符号化ユニット５１４は、先進的オーディオコーディング（ＡＡＣ）符号化ユニットの１つまたは複数の具体物を表し得る。オーディオ符号化ユニット５１４は、U_{DIST}ベクトル＊S_{DIST}ベクトル

５２７の各列または各行を符号化することができる。しばしば、オーディオ符号化ユニット５１４は、バックグラウンド球面調和係数５３１に残っている次数／位数の組合せの各々のために、ＡＡＣ符号化ユニットの具体物を呼び出すことができる。バックグラウンド球面調和係数５３１がどのようにＡＡＣ符号化ユニットを使用して符号化され得るかに関するより多くの情報が、第１２４回会議、２００８年5月１７～２０日において発表され、http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapersにおいて利用可能な、Ｅｒｉｃ Ｈｅｌｌｅｒｕｄ他、「Ｅｎｃｏｄｉｎｇ Ｈｉｇｈｅｒ Ｏｒｄｅｒ Ａｍｂｉｓｏｎｉｃｓ ｗｉｔｈ ＡＡＣ」という表題の会議文書において見出され得る。オーディオ符号化ユニット１４は、U_{DIST}ベクトル＊S_{DIST}ベクトル５２７の符号化されたバージョン（「符号化されたU_{DIST}ベクトル＊S_{DIST}ベクトル５１５」と示される）とバックグラウンド球面調和係数５３１の符号化されたバージョン（「符号化されたバックグラウンド球面調和係数５１５Ｂ」と示される）とをビットストリーム生成ユニット５１６に出力することができる。いくつかの例では、オーディオ符号化ユニット５１４は、U_{DIST}ベクトル＊S_{DIST}ベクトル５２７を符号化するために使用されるものよりも低い目標ビットレートを使用してバックグラウンド球面調和係数５３１をオーディオ符号化し、これによって、場合によっては、U_{DIST}＊S_{DIST}ベクトル５２７と比較してより大きくバックグラウンド球面調和係数５３１を圧縮することができる。

【０７２０】

[0715]ビットストリーム生成ユニット５１６は、既知のフォーマット（復号デバイスにより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、それによってビットストリーム５１７を生成するユニットを表す。ビットストリーム生成ユニット４２は、図４の例に示されるオーディオ符号化デバイス２４のビットストリーム生成ユニット４２に関して上で説明されたものと同様の方式で動作することができる。ビットストリーム生成ユニット５１６は、符号化されたU_{DIST}ベクトル＊S_{DIST}ベクトル５１５と、符号化されたバックグラウンド球面調和係数５１５６ＢおよびV^{T}_{DIST}行列５２５Ｅとを多重化するマルチプレクサを含み得る。

【０７２１】

[0716]図４０Ｂは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法の様々な態様を実行し得る、例示的なオーディオ符号化デバイス５１０Ｂを示すブロック図である。オーディオ符号化デバイス５１０Ｂは、オーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含むという点で、オーディオ符号化デバイス５１０と同様であり得る。その上、オーディオ符号化デバイス５１０Ｂのオーディオ圧縮ユニット５１２は、分解ユニット５１８を含むという点で、オーディオ符号化デバイス５１０のオーディオ圧縮ユニットと同様であり得る。オーディオ符号化デバイス５１０Ｂのオーディオ圧縮ユニット５１２は、音場成分抽出ユニット５２０が次数低減ユニット５２８Ａ（「次数低減ユニット５２８（ｏｒｄｅｒ ｒｅｄｕｃｔ ｕｎｉｔ ５２８）」）として示される追加のユニットを含むという点で、オーディオ符号化デバイス５１０のオーディオ圧縮ユニット５１２とは異なり得る。この理由で、オーディオ符号化デバイス５１０Ｂの音場成分抽出ユニット５２０は、「音場成分抽出ユニット５２０Ｂ」として示される。

【０７２２】

[0717]次数低減ユニット５２８Ａは、バックグラウンド球面調和係数５３１の追加の次数低減を実行するように構成されるユニットを表す。いくつかの例では、次数低減ユニット５２８Ａは、音場を表すのに必要なバックグラウンド球面調和係数５３１の数を減らすために、表されたバックグラウンド球面調和係数５３１によって音場を回転することができる。いくつかの例では、バックグラウンド球面調和係数５３１が音場のバックグラウンド成分を表すとすると、次数低減ユニット５２８Ａは、バックグラウンド球面調和係数５３１のうちで高次の球面基底関数に対応するものを取り除き、除去し、または（しばしば０にすることによって）別様に削除することができる。この点において、次数低減ユニット５２８Ａは、図４の例に示されるオーディオ符号化デバイス２０のバックグラウンド選

択ユニット４８と同様の動作を実行することができる。次数低減ユニット５２８Ａは、バックグラウンド球面調和係数５３１の低減されたバージョン（「低減されたバックグラウンド球面調和係数５２９」として示される）をオーディオ符号化ユニット５１４に出力することができ、オーディオ符号化ユニット５１４は、低減されたバックグラウンド球面調和係数５２９を符号化するために上で説明された方式でオーディオ符号化を実行し、それによって、符号化された低減されたバックグラウンド球面調和係数５１５Ｂを生成することができる。

【０７２３】

[0718]以下で列挙される様々な条項は、本開示で説明された技法の様々な態様を提示し得る。

【０７２４】

[0719]項１３２５６７－１。複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すように構成される１つまたは複数のプロセッサを備える、オーディオ符号化デバイス５１０またはオーディオ符号化デバイス５１０Ｂなどのデバイス。

【０７２５】

[0720]項１３２５６７－２。１つまたは複数のプロセッサがさらに、それらの組合せまたは派生物を含むＵ行列、Ｓ行列、およびＶ行列の１つまたは複数のベクトルとしての複数の球面調和係数の表現を含むようにビットストリームを生成するように構成される、項１３２５６７－１のデバイス。

【０７２６】

[0721]項１３２５６７－３。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分を記述するＵ行列に含まれる１つまたは複数のU_{DIST}ベクトルを決定するように構成される、項１３２５６７－１のデバイス。

【０７２７】

[0722]項１３２５６７－４。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分を記述するＵ行列に含まれる１つまたは複数のU_{DIST}ベクトルを決定し、音場の明瞭な成分を同様に記述するＳ行列に含まれる１つまたは複数のS_{DIST}ベクトルを決定し、U_{DIST}＊S_{DIST}ベクトルを生成するために１つまたは複数のU_{DIST}ベクトルと１つまたは複数のS_{DIST}ベクトルを乗算するように構成される、項１３２５６７－１のデバイス。

【０７２８】

[0723]項１３２５６７－５。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分を記述するＵ行列に含まれる１つまたは複数のU_{DIST}ベクトルを決定し、音場の明瞭な成分を同様に記述するＳ行列に含まれる１つまたは複数のS_{DIST}ベクトルを決定し、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを生成するために１つまたは複数のU_{DIST}ベクトルと１つまたは複数のS_{DIST}ベクトルを乗算するように構成され、１つまたは複数のプロセッサがさらに、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルのオーディオ符号化されたバージョンを生成するために１つまたは複数のU_{DIST}＊S_{DIST}ベクトルをオーディオ符号化するように構成される、項１３２５６７－１のデバイス。

【０７２９】

[0724]項１３２５６７－６。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、Ｕ行列に含まれる１つまたは複数のU_{BG}ベクトルを決定するように構成される、項１３２５６７－１のデバイス。

【０７３０】

[0725]項１３２５６７－７。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分とバックグラウンド成分とを識別するためにＳ行列を分析するように構成される、項１３２５６７－１のデバイス。

【０７３１】
[0726]項１３２５６７－８。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分とバックグラウンド成分とを識別するためにＳ行列を分析し、Ｓ行列の分析に基づいて、音場の明瞭な成分を記述するＵ行列の１つまたは複数のU_{DIST}ベクトルと音場のバックグラウンド成分を記述するＵ行列の１つまたは複数のU_{BG}ベクトルとを決定するように構成される、項１３２５６７－１のデバイス。
【０７３２】
[0727]項１３２５６７－９。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分とバックグラウンド成分とを識別するためにオーディオフレームごとにＳ行列を分析し、オーディオフレームごとのＳ行列の分析に基づいて、音場の明瞭な成分を記述するＵ行列の１つまたは複数のU_{DIST}ベクトルと音場のバックグラウンド成分を記述するＵ行列の１つまたは複数のU_{BG}ベクトルとを決定するように構成される、項１３２５６７－１のデバイス。
【０７３３】
[0728]項１３２５６７－１０。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、音場の明瞭な成分とバックグラウンド成分とを識別するためにＳ行列を分析し、Ｓ行列の分析に基づいて、音場の明瞭な成分を記述するＵ行列の１つまたは複数のU_{DIST}ベクトルと音場のバックグラウンド成分を記述するＵ行列の１つまたは複数のU_{BG}ベクトルとを決定し、Ｓ行列の分析に基づいて、１つまたは複数のU_{DIST}ベクトルおよび１つまたは複数のU_{BG}ベクトルに対応するＳ行列の１つまたは複数のS_{DIST}ベクトルと１つまたは複数のS_{BG}ベクトルとを決定すること、Ｓ行列の分析に基づいて、１つまたは複数のU_{DIST}ベクトルおよび１つまたは複数のU_{BG}ベクトルに対応するＶ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}ベクトルとを決定するように構成される、項１３２５６７－１のデバイス。
【０７３４】
[0729]項１３２５６７－１１。１つまたは複数のプロセッサがさらに、複数の球面調和係数をさらに表すときに、１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算するように構成され、１つまたは複数のプロセッサがさらに、$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルのオーディオ符号化されたバージョンを生成するために$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルをオーディオ符号化するように構成される、項１３２５６７－１０のデバイス。
【０７３５】
[0730]項１３２５６７－１２。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算し、１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルの係数のうちで球面調和基底関数の１つまたは複数の次数と関連付けられるものを除去し、それによって１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルの次数低減されたバージョンを生成するために、次数低減処理を実行するように構成される、項１３２５６７－１０のデバイス。
【０７３６】
[0731]項１３２５６７－１３。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算し、１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルの係数のうちで球面調和基底関数の１つまたは複数の次数と関連付けられるものを除去し、それによって１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルの次数低減されたバージョンを生成するために、次数低減処理を実行するように構成され、１つまたは複数のプロセッサがさらに、次数低減された１つまたは複数の$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルのオーディオ符号化

バージョンを生成するために、１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルの次数低減されたバージョンをオーディオ符号化するように構成される、項１３２５６７－１０のデバイス。

【０７３７】

[0732]項１３２５６７－１４。１つまたは複数のプロセッサがさらに、複数の球面調和係数を表すときに、１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルを生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算し、１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルの係数のうちで球面調和基底関数の１よりも大きい１つまたは複数の次数と関連付けられるものを除去し、それによって１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルの次数低減されたバージョンを生成するために、次数低減処理を実行し、次数低減された１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルのオーディオ符号化されたバージョンを生成するために、１つまたは複数のU_{BG}＊S_{BG}＊V^{T}_{BG}ベクトルの次数低減されたバージョンをオーディオ符号化するように構成される、項１３２５６７－１０のデバイス。

【０７３８】

[0733]項１３２５６７－１５。１つまたは複数のプロセッサがさらに、１つまたは複数のV^{T}_{DIST}ベクトルを含むようにビットストリームを生成するように構成される、項１３２５６７－１０のデバイス。

【０７３９】

[0734]項１３２５６７－１６。１つまたは複数のプロセッサがさらに、１つまたは複数のV^{T}_{DIST}ベクトルをオーディオ符号化することなく１つまたは複数のV^{T}_{DIST}ベクトルを含むようにビットストリームを生成するように構成される、項１３２５６７－１０のデバイス。

【０７４０】

[0735]項１３２５６７－１Ｆ。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関して特異値分解を実行し、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数としてマルチチャネルオーディオデータを表すための、１つまたは複数のプロセッサを備える、オーディオ符号化デバイス５１０または５１０Ｂなどのデバイス。

【０７４１】

[0736]項１３２５６７－２Ｆ。マルチチャネルオーディオデータが、複数の球面調和係数を備える、項１３２５６７－１Ｆのデバイス。

【０７４２】

[0737]項１３２５６７－３Ｆ。１つまたは複数のプロセッサがさらに、項１３２５６７－２から項１３２５６７－１６の任意の組合せによって記載されるように実行するように構成される、項１３２５６７－２Ｆのデバイス。

【０７４３】

[0738]上で説明された様々な項の各々から、オーディオ符号化デバイス５１０Ａ～５１０Ｊのいずれもが、オーディオ符号化デバイス５１０Ａ～５１０Ｊが実行するように構成される方法を実行することができ、またはその方法の各ステップを実行するための手段を別様に備え得ることを理解されたい。いくつかの例では、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。いくつかの例では、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５１０Ａ～５１０Ｊ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。

【０７４４】

[0739]たとえば、項１３２５６７－１７は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行することと、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すこととを備える方法として、前述の項１３２５６７－１から導出され得る。

【０７４５】

[0740]別の例として、項１３２５６７－１８は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するための手段と、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すための手段とを備える、オーディオ符号化デバイス５１０Ｂなどのデバイスとして、前述の項１３２５６７－１から導出され得る。

【０７４６】

[0741]さらに別の例として、項１３２５６７－１８は、実行されると、１つまたは複数のプロセッサに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行させ、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表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として、前述の項１３２５６７－１から導出され得る。

【０７４７】

[0742]様々な項が同様に、上で例示されたように導出される様々なデバイス、方法、および非一時的コンピュータ可読記憶媒体について、項１３２５６７－２～１３２５６７－１６から導出され得る。本開示全体で列挙される様々な他の項について、同じことが実行され得る。

【０７４８】

[0743]図４０Ｃは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法の様々な態様を実行し得る、例示的なオーディオ符号化デバイス５１０Ｃを示すブロック図である。オーディオ符号化デバイス５１０Ｃは、オーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含むという点で、オーディオ符号化デバイス５１０Ｂと同様であり得る。その上、オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニット５１２は、分解ユニット５１８を含むという点でオーディオ符号化デバイス５１０Ｂのオーディオ圧縮ユニットと同様であり得る。

【０７４９】

[0744]オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニット５１２は、しかしながら、音場成分抽出ユニット５２０がベクトル再順序付けユニット５３２として示される追加のユニットを含むという点で、オーディオ符号化デバイス５１０Ｂのオーディオ圧縮ユニット５１２とは異なり得る。この理由で、オーディオ符号化デバイス５１０Ｃの音場成分抽出ユニット５２０は、「音場成分抽出ユニット５２０Ｃ」として示される。

【０７５０】

[0745]ベクトル再順序付けユニット５３２は、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３を生成するためにU_{DIST}＊S_{DIST}ベクトル５２７を再順序付けるように構成されるユニットを表し得る。この点で、ベクトル再順序付けユニット５３２は、図４の例に示されるオーディオ符号化デバイス２０の再順序付けユニット３４に関して上で説明されたものと同様の方式で動作することができる。音場成分抽出ユニット５２０Ｃは、U_{DIST}＊S_{DIST}ベクトル５２７（ここで、U_{DIST}＊S_{DIST}ベクトル５２７の各ベクトルは音場の中に存在する１つまたは複数の明瞭なモノラルオーディオオブジェクトを

表し得る）の次数が上で述べられた理由でオーディオデータの部分から変化し得るので、U_{DIST}＊S_{DIST}ベクトル５２７を再順序付けるためにベクトル再順序付けユニット５３２を呼び出すことができる。すなわち、オーディオ圧縮ユニット５１２がいくつかの例ではオーディオフレーム（球面調和係数５１１のM個のサンプルを有してよく、ここでMはいくつかの例では１０２４に設定される）と一般に呼ばれるオーディオデータのこれらの部分に対して動作するとすると、U_{DIST}＊S_{DIST}ベクトル５２７がそれから導出されるU行列５１９Ｃにおいて表されるようなこれらの明瞭なモノラルオーディオオブジェクトに対応するベクトルの部分は、オーディオフレームごとに変化し得る。

【０７５１】

[0746]これらのU_{DIST}＊S_{DIST}ベクトル５２７をオーディオフレームごとに再順序付けることなくこれらのU_{DIST}＊S_{DIST}ベクトル５２７を直接オーディオ符号化ユニット５１４に渡すことは、モノラルオーディオオブジェクトがオーディオフレームにわたって相関がある（この例では互いに対するU_{DIST}＊S_{DIST}ベクトル５２７の次数によって定義される、チャネル－ワイズ（channel-wise））ときにより良好に機能する従来の圧縮スキームなどのいくつかの圧縮スキームについて達成可能な圧縮の程度を下げることがある。その上、再順序付けられないとき、U_{DIST}＊S_{DIST}ベクトル５２７の符号化は、復元されるときのオーディオデータの品質を下げることがある。たとえば、オーディオ符号化ユニット５１４によって図４０Ｃの例において表され得るＡＡＣ符号化器は、U_{DIST}＊S_{DIST}ベクトル５２７をフレームごとに直接符号化するときに達成される圧縮と比較して、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３をフレームごとにより効率的に圧縮することができる。ＡＡＣ符号化器に関して上で説明されるが、本技法は、モノラルオーディオオブジェクトが複数のフレームにわたって特定の順序または位置に指定される（チャネル的である）ときにより良好な圧縮をもたらす任意の符号化器に関して実行され得る。

【０７５２】

[0747]以下でより詳細に説明されるように、本技法は、オーディオ符号化デバイス５１０Ｃが、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３を生成し、それによってオーディオ符号化ユニット５１４などの従来のオーディオ符号化器によるU_{DIST}＊S_{DIST}ベクトル５２７の圧縮を容易にするために、１つまたは複数のベクトル（すなわち、U_{DIST}＊S_{DIST}ベクトル５２７を再順序付けることを可能にし得る。オーディオ符号化デバイス５１０Ｃはさらに、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３の符号化されたバージョン５１５Ａを生成するために、オーディオ符号化ユニット５１４を使用して、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３をオーディオ符号化するための本開示で説明された技法を実行することができる。

【０７５３】

[0748]たとえば、音場成分抽出ユニット５２０Ｃは、１つまたは複数の第２のU_{DIST}＊S_{DIST}ベクトル５２７が対応する第２のフレームに時間的に後続する第１のオーディオフレームから、１つまたは複数の第１のU_{DIST}＊S_{DIST}ベクトル５２７を再順序付けるために、ベクトル再順序付けユニット５３２を呼び出すことができる。第１のオーディオフレームが第２のオーディオフレームよりも時間的に後にあるという文脈で説明されるが、第１のオーディオフレームは第２のオーディオフレームよりも時間的に先行してよい。したがって、本技法は、本開示で説明される例に限定されるべきではない。

【０７５４】

[0749]ベクトル再順序付けユニット５３２はまず、第１のU_{DIST}＊S_{DIST}ベクトル５２７と第２のU_{DIST}＊S_{DIST}ベクトル５２７の各々に関してエネルギー分析を実行し、第１のオーディオフレームの少なくとも一部分（しかししばしば全体）および第２のオーディオフレームの少なくとも一部分（しかししばしば全体）の二乗平均平方根エネルギーを計算して、それによって、第１のオーディオフレームの第１のU_{DIST}＊S_{DIST}ベクトル５２７の各々に対して１つ、第２のオーディオフレームの第２のU_{DIST}＊S_{DIST}ベクトル５２７の各々に対して１つの、（Ｄを４であると仮定して）８つのエネルギーを生成すること

ができる。ベクトル再順序付けユニット５３２は次いで、表１～４に関して上で説明された第２のU_{DIST}＊S_{DIST}ベクトル５２７の各々に対して、第１のU_{DIST}＊S_{DIST}ベクトル５２７からの各エネルギーを順番ごとに比較することができる。

【０７５５】

[0750]言い換えれば、ＨｏＡ信号に対するフレームベースのＳＶＤ（またはＫＬＴおよびＰＣＡなどの関連する方法）の分解を使用するとき、フレームごとのベクトルの次数は一貫していることが保証されないことがある。たとえば、背後の音場に２つのオブジェクトがある場合、分解（適切に実行されるときは「理想的な分解」と呼ばれ得る）は、１つのベクトルがＵ行列中の１つのオブジェクトを表すように、２つのオブジェクトの分離をもたらし得る。しかしながら、分解が「理想的な分解」として示され得るときでも、ベクトルは、Ｕ行列の中の（およびそれに対応してＳ行列およびＶ行列の中の）位置がフレームごとに変化し得る。さらに、位相の差もあることがあり、ここで、ベクトル再順序付けユニット５３２は、位相反転を使用して（反転されたベクトルの各要素をマイナス１すなわち－１によってドット乗算することによって）位相を反転することができる。これらのベクトルを供給するために、フレームごとに同じ「ＡＡＣ／オーディオコーディングエンジン」へのフレームごとは、次数が識別されること（または言い換えれば、信号が照合されること）と、位相が調整されることと、フレーム境界における注意深い補間が適用されることとを必要とし得る。これがなければ、背後のオーディオコーデックは、「テンポラルスミアリング（temporal smearing）」または「プリエコー（pre-echo）」として知られているものを含む極めて激しいアーティファクトを生成し得る。

【０７５６】

[0751]本開示で説明される技法の様々な態様によれば、オーディオ符号化デバイス５１０Ｃは、ベクトルのフレーム境界におけるエネルギーと相互相関付けとを使用して、ベクトルを識別／照合するために複数の方法を適用することができる。オーディオ符号化デバイス５１０Ｃはまた、フレーム境界において現れることの多い１８０度の位相変化が修正されることを確実にし得る。ベクトル再順序付けユニット５３２は、フレーム間の滑らかな移行を確実にするために、ベクトル間にある形態のフェードイン／フェードアウト補間ウィンドウを適用することができる。

【０７５７】

[0752]このようにして、オーディオ符号化デバイス５３０Ｃは、再順序付けられた１つまたは複数の第１のベクトルを生成するために１つまたは複数のベクトルを再順序付け、それによって従来のオーディオ符号化器による符号化を支援し、ここにおいて、１つまたは複数のベクトルは音場の明瞭な成分を記述する表す、再順序付けられた１つまたは複数のベクトルの符号化されたバージョンを生成するために、従来のオーディオ符号化器を使用して、再順序付けられた１つまたは複数のベクトルをオーディオ符号化することができる。

【０７５８】

[0753]本開示で説明される技法の様々な態様は、オーディオ符号化デバイス５１０Ｃが以下の項に従って動作することを可能にし得る。

【０７５９】

[0754]項１３３１４３－１Ａ。再順序付けられた１つまたは複数の第１のベクトルを決定し、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方の抽出を容易にするために、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとの間のエネルギー比較を実行するように構成される、１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルがオーディオデータの第１の部分の中の音場の明瞭な成分を記述し、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の音場の明瞭な成分を記述する、オーディオ符号化デバイス５１０Ｃなどのデバイス。

【０７６０】

[0755]項１３３１４３－２Ａ。１つまたは複数の第１のベクトルがオーディオデータの

第１の部分の中の音場のバックグラウンド成分を表さず、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の音場のバックグラウンド成分を表さない、項１３３１４３－１Ａのデバイス。

【０７６１】

[0756]項１３３１４３－３Ａ。１つまたは複数のプロセッサがさらに、エネルギー比較を実行した後で、１つまたは複数の第２のベクトルと相関した１つまたは複数の第１のベクトルを識別するために、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとの間の相互相関付けを実行するように構成される、項１３３１４３－１Ａのデバイス。

【０７６２】

[0757]項１３３１４３－４Ａ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄し、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行し、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付けるように構成される、項１３３１４３－１Ａのデバイス。

【０７６３】

[0758]項１３３１４３－５Ａ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄し、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行し、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付け、再順序付けられた１つまたは複数の第１のベクトルのオーディオ符号化されたバージョンを生成するために、再順序付けられた１つまたは複数の第１のベクトルを符号化するように構成される、項１３３１４３－１Ａのデバイス。

【０７６４】

[0759]項１３３１４３－６Ａ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄し、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行し、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付け、再順序付けられた１つまたは複数の第１のベクトルのオーディオ符号化されたバージョンを生成するために、再順序付けられた１つまたは複数の第１のベクトルを符号化し、再順序付けられた１つまたは複数の第１のベクトルの符号化されたバージョンを含むようにビットストリームを生成するように構成される、項１３３１４３－１Ａのデバイス。

【０７６５】

[0760]項１３３１４３－７Ａ。オーディオデータの第１の部分がＭ個のサンプルを有する第１のオーディオフレームを備え、オーディオデータの第２の部分が同数のＭ個のサンプルを有する第２のオーディオフレームを備え、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは

複数の第１のベクトルの少なくとも１つの最後のＭ－Ｚ個の値および低減された１つまたは複数の第２のベクトルの各々の最初のＭ－Ｚ個の値に関して相互相関付けを実行するように構成され、ＺがＭより小さい、請求項３Ａから６Ａのデバイス。

【０７６６】

[0761]項１３３１４３－８Ａ。オーディオデータの第１の部分がＭ個のサンプルを有する第１のオーディオフレームを備え、オーディオデータの第２の部分が同数のＭ個のサンプルを有する第２のオーディオフレームを備え、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つの最後のＭ－Ｙ個の値および低減された１つまたは複数の第２のベクトルの各々の最初のＭ－Ｚ個の値に関して相互相関付けを実行するように構成され、ＺとＹの両方がＭより小さい、請求項３Ａから６Ａのデバイス。

【０７６７】

[0762]項１３３１４３－９Ａ。１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの少なくとも１つを反転するように構成される、請求項３Ａから６Ａのデバイス。

【０７６８】

[0763]項１３３１４３－１０Ａ。１つまたは複数のプロセッサがさらに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成される、項１３３１４３－１Ａのデバイス。

【０７６９】

[0764]項１３３１４３－１１Ａ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の関数として１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを生成するように構成される、項１３３１４３－１Ａのデバイス。

【０７７０】

[0765]項１３３１４３－１２Ａ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、Ｕ行列の１つまたは複数のＵＤＩＳＴベクトルとＳ行列の１つまたは複数のＳＤＩＳＴ行列とを識別するために、Ｓ行列に関して顕著さの分析を実行し、１つまたは複数のＵＤＩＳＴベクトルを１つまたは複数のＳＤＩＳＴベクトルと乗算することによって少なくとも一部、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを決定するように構成される、項１３３１４３－１Ａのデバイス。

【０７７１】

[0766]項１３３１４３－１３Ａ。オーディオデータの第１の部分が、オーディオデータの第２の部分よりも時間的に前に現れる、項１３３１４３－１Ａのデバイス。

【０７７２】

[0767]項１３３１４３－１４Ａ。オーディオデータの第１の部分が、オーディオデータの第２の部分よりも時間的に後に現れる、項１３３１４３－１Ａのデバイス。

【０７７３】

[0768]項１３３１４３－１５Ａ。１つまたは複数のプロセッサがさらに、エネルギー比較を実行するときに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの各々の二乗平均平方根エネルギーを計算し、１つまたは複数の第１のベクトルの少なくとも１つに対して計算される二乗平均平方根エネルギーを１つまたは複数の第２のベクトルの各々に対して計算される二乗平均平方根エネルギーと比較するように構成される、

項１３３１４３－１Ａのデバイス。
【０７７４】
[0769]項１３３１４３－１６Ａ。１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、エネルギー比較に基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付けるように構成され、１つまたは複数のプロセッサがさらに、第１のベクトルを再順序付けるときに、再順序付けられた１つまたは複数の第１のベクトルを生成するときの滑らかな移行を確実にするために、１つまたは複数の第１のベクトルの間でフェードイン／フェードアウト補間ウィンドウを適用するように構成される、項１３３１４３－１Ａのデバイス。
【０７７５】
[0770]項１３３１４３－１７Ａ。１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、エネルギー比較に少なくとも基づいて１つまたは複数の第１のベクトルを再順序付け、再順序付けられた１つまたは複数の第１のベクトルまたは再順序付けられた１つまたは複数の第１のベクトルの符号化されたバージョンを含むようにビットストリームを生成し、１つまたは複数の第１のベクトルがどのように再順序付けられたかを記述する再順序付け情報をビットストリームにおいて指定するように構成される、項１３３１４３－１Ａのデバイス。
【０７７６】
[0771]項１３３１４３－１８Ａ。エネルギー比較が、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方のオーディオ符号化を促進するために、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方の抽出を容易にする、項１３３１４３－１Ａのデバイス。
【０７７７】
[0772]項１３３１４３－１Ｂ。再順序付けられた１つまたは複数の第１のベクトルを決定するために１つまたは複数の第１のベクトルおよび１つまたは複数の第２のベクトルに関して相互相関付けを実行して、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方の抽出を容易にするように構成される、１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルがオーディオデータの第１の部分の中の音場の明瞭な成分を記述し、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の音場の明瞭な成分を記述する、オーディオ符号化デバイス５１０Ｃなどのデバイス。
【０７７８】
[0773]項１３３１４３－２Ｂ。１つまたは複数の第１のベクトルがオーディオデータの第１の部分の中の音場のバックグラウンド成分を表さず、１つまたは複数の第２のベクトルがオーディオデータの第２の部分の中の音場のバックグラウンド成分を表さない、項１３３１４３－１Ｂのデバイス。
【０７７９】
[0774]項１３３１４３－３Ｂ。１つまたは複数のプロセッサがさらに、相互相関付けを実行する前に、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとの間のエネルギー比較を実行するように構成され、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方のオーディオ符号化を容易にするために、１つまたは複数の第１のベクトルと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行するように構成される、項１３３１４３－１Ｂのデバイス。
【０７８０】
[0775]項１３３１４３－４Ｂ。１つまたは複数のプロセッサがさらに、エネルギー比較を実行するときに、１つまたは複数の第１のベクトルおよび１つまたは複数の第２のベクトルの各々の二乗平均平方根エネルギーを計算し、１つまたは複数の第１のベクトルの少

なくとも１つに対して計算される二乗平均平方根エネルギーを１つまたは複数の第２のベクトルの各々に対して計算される二乗平均平方根エネルギーと比較するように構成される、項１３３１４３－３Ｂのデバイス。

【０７８１】

[0776]項１３３１４３－５Ｂ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄するように構成され、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行するように構成され、１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付けるように構成される、項１３３１４３－３Ｂのデバイス。

【０７８２】

[0777]項１３３１４３－６Ｂ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄するように構成され、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行するように構成され、１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付け、再順序付けられた１つまたは複数の第１のベクトルのオーディオ符号化されたバージョンを生成するために、再順序付けられた１つまたは複数の第１のベクトルを符号化するように構成される、項１３３１４３－３Ｂのデバイス。

【０７８３】

[0778]項１３３１４３－７Ｂ。１つまたは複数のプロセッサがさらに、１つまたは複数の第２のベクトルよりも少数のベクトルを有する低減された１つまたは複数の第２のベクトルを生成するために、エネルギー比較に基づいて第２のベクトルの１つまたは複数を廃棄するように構成され、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つと低減された１つまたは複数の第２のベクトルとの間の相互相関付けを実行するように構成され、１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付け、再順序付けられた１つまたは複数の第１のベクトルのオーディオ符号化されたバージョンを生成するために、再順序付けられた１つまたは複数の第１のベクトルを符号化し、再順序付けられた１つまたは複数の第１のベクトルの符号化されたバージョンを含むようにビットストリームを生成するように構成される、項１３３１４３－３Ｂのデバイス。

【０７８４】

[0779]項１３３１４３－８Ｂ。オーディオデータの第１の部分がＭ個のサンプルを有する第１のオーディオフレームを備え、オーディオデータの第２の部分が同数のＭ個のサンプルを有する第２のオーディオフレームを備え、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは

複数の第１のベクトルの少なくとも１つの最後のＭ－Ｚ個の値および低減された１つまたは複数の第２のベクトルの各々の最初のＭ－Ｚ個の値に関して相互相関付けを実行するように構成され、ＺがＭより小さい、請求項３Ｂから７Ｂのデバイス。

【０７８５】

[0780]項１３３１４３－９Ｂ。オーディオデータの第１の部分がＭ個のサンプルを有する第１のオーディオフレームを備え、オーディオデータの第２の部分が同数のＭ個のサンプルを有する第２のオーディオフレームを備え、１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルの少なくとも１つに相関する低減された１つまたは複数の第２のベクトルの１つを識別するために、１つまたは複数の第１のベクトルの少なくとも１つの最後のＭ－Ｙ個の値および低減された１つまたは複数の第２のベクトルの各々の最初のＭ－Ｚ個の値に関して相互相関付けを実行するように構成され、ＺとＹの両方がＭより小さい、請求項３Ｂから７Ｂのデバイス。

【０７８６】

[0781]項１３３１４３－１０Ｂ。１つまたは複数のプロセッサがさらに、相互相関付けを実行するときに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの少なくとも１つを反転するように構成される、請求項１Ｂのデバイス。

【０７８７】

[0782]項１３３１４３－１１Ｂ。１つまたは複数のプロセッサがさらに、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成される、項１３３１４３－１Ｂのデバイス。

【０７８８】

[0783]項１３３１４３－１２Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の関数として１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを生成するように構成される、項１３３１４３－１Ｂのデバイス。

【０７８９】

[0784]項１３３１４３－１３Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、Ｕ行列の１つまたは複数のU_{DIST}ベクトルとＳ行列の１つまたは複数のS_{DIST}行列とを識別するために、Ｓ行列に関して顕著さの分析を実行し、１つまたは複数のU_{DIST}ベクトルを１つまたは複数のS_{DIST}ベクトルと乗算することによって少なくとも一部、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを決定するように構成される、項１３３１４３－１Ｂのデバイス。

【０７９０】

[0785]項１３３１４３－１４Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを決定するときに、１つまたは複数の第１のベクトルおよび１つまたは複数の第２のベクトルの少なくとも１つとしてＶ行列の１つまたは複数のＶＤＩＳＴベクトルを識別するために、Ｓ行列に関して顕著さの分析を実行するように構成される、項１３３１４３－１Ｂのデバイス。

【０７９１】

[0786]項１３３１４３－１５Ｂ。オーディオデータの第１の部分がオーディオデータの第２の部分よりも時間的に前に現れる、項１３３１４３－１Ｂのデバイス。

【０７９２】

[0787]項１３３１４３－１６Ｂ。オーディオデータの第１の部分がオーディオデータの第２の部分よりも時間的に後に現れる、項１３３１４３－１Ｂのデバイス。

【０７９３】

[0788]項１３３１４３－１７Ｂ。１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに基づいて１つまたは複数の第１のベクトルの少なくとも１つを再順序付け、第１のベクトルを再順序付けるときに、再順序付けられた１つまたは複数の第１のベクトルを生成するときの滑らかな移行を確実にするために、１つまたは複数の第１のベクトルの間でフェードイン／フェードアウト補間ウィンドウを適用するように構成される、項１３３１４３－１Ｂのデバイス。

【０７９４】

[0789]項１３３１４３－１８Ｂ。１つまたは複数のプロセッサがさらに、再順序付けられた１つまたは複数の第１のベクトルを生成するために、相互相関付けに少なくとも基づいて１つまたは複数の第１のベクトルを再順序付け、再順序付けられた１つまたは複数の第１のベクトルまたは再順序付けられた１つまたは複数の第１のベクトルの符号化されたバージョンを含むようにビットストリームを生成し、１つまたは複数の第１のベクトルがどのように再順序付けられたかをビットストリームにおいて指定するように構成される、項１３３１４３－１Ｂのデバイス。

【０７９５】

[0790]項１３３１４３－１９Ｂ。相互相関付けが、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方のオーディオ符号化を促進するために、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの一方または両方の抽出を容易にする、項１３３１４３－１Ｂのデバイス。

【０７９６】

[0791]図４０Ｄは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法の様々な態様を実行し得る、例示的なオーディオ符号化デバイス５１０Ｄを示すブロック図である。オーディオ符号化デバイス５１０Ｄは、オーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含むという点で、オーディオ符号化デバイス５１０Ｃと同様であり得る。その上、オーディオ符号化デバイス５１０Ｄのオーディオ圧縮ユニット５１２は、分解ユニット５１８を含むという点で、オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニットと同様であり得る。

【０７９７】

[0792]オーディオ符号化デバイス５１０Ｄのオーディオ圧縮ユニット５１２は、しかしながら、音場成分抽出ユニット５２０が量子化ユニット５３４（「量子化ユニット５３４（ｑｕａｎｔ　ｕｎｉｔ　５３４）」）として示される追加のユニットを含むという点で、オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニット５１２とは異なり得る。この理由で、オーディオ符号化デバイス５１０Ｄの音場成分抽出ユニット５２０は、「音場成分抽出ユニット５２０Ｄ」として示される。

【０７９８】

[0793]量子化ユニット５３４は、対応する１つもしくは複数の$V^T_{Q_DIST}$ベクトル５２５Ｇおよび／または１つもしくは複数の$V^T_{Q_BG}$ベクトル５２５Ｈを生成するために、１つもしくは複数のV^T_{DIST}ベクトル５２５Ｅおよび／または１つもしくは複数のV^T_{BG}ベクトル５２５Ｆを量子化するように構成されるユニットを表す。量子化ユニット５３４は、ビットストリーム５１７において１つまたは複数のV^T_{DIST}ベクトル５２５Ｅを表すために使用されるビットの数を減らすために、１つまたは複数のV^T_{DIST}ベクトル５２５Ｅを量子化する（これは、値を表すために使用されるビットの削除を通じた数学的な丸めのための信号処理用語である）ことができる。いくつかの例では、量子化ユニット５３４は、１つまたは複数のV^T_{DIST}ベクトル５２５Ｅの３２ビット値を量子化して、これらの３２ビット値を丸められた１６ビット値で置き換えて１つまたは複数の$V^T_{Q_DIST}$ベクトル５

２５Ｇを生成することができる。この点で、量子化ユニット５３４は、図４の例に示されるオーディオ符号化デバイス２０の量子化ユニット５２に関して上で説明されたものと同様の方式で動作することができる。

【０７９９】

[0794]この性質の量子化は、量子化の粗さに応じて変化する誤差を音場の表現へともたらし得る。言い換えれば、１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅを表すために使用されるビットがより多いと、量子化誤差はより小さくなり得る。V^{T}_{DIST}ベクトル５２５Ｅの量子化による量子化誤差（「E_{DIST}」と示され得る）は、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇから１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅを差し引くことによって決定され得る。

【０８００】

[0795]本開示で説明される技法によれば、オーディオ符号化デバイス５１０Ｄは、$U_{DIST}*S_{DIST}$ベクトル５２７の１つまたは複数に、あるいは、１つまたは複数のU_{BG}ベクトル５２５Ｄを１つまたは複数のS_{BG}ベクトル５２５Ｂと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトル５２５Ｆと乗算することによって生成されるバックグラウンド球面調和係数５３１に、E_{DIST}誤差を射影することによって、またはそれらを別様に修正することによって、E_{DIST}量子化誤差の１つまたは複数を補償することができる。いくつかの例では、オーディオ符号化デバイス５１０Ｄは、$U_{DIST}*S_{DIST}$ベクトル５２７の中のE_{DIST}誤差のみを補償することができる。他の例では、オーディオ符号化デバイス５１０Ｄは、バックグラウンド球面調和係数の中のE_{BG}誤差のみを補償することができる。さらに他の例では、オーディオ符号化デバイス５１０Ｄは、$U_{DIST}*S_{DIST}$ベクトル５２７とバックグラウンド球面調和係数の両方の中のE_{DIST}誤差を補償することができる。

【０８０１】

[0796]動作において、顕著成分分析ユニット５２４は、１つまたは複数のS_{DIST}ベクトル５２５と、１つまたは複数のS_{BG}ベクトル５２５Ｂと、１つまたは複数のU_{DIST}ベクトル５２５Ｃと、１つまたは複数のU_{BG}ベクトル５２５Ｄと、１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅと、１つまたは複数のV^{T}_{BG}ベクトル５２５Ｆとを、算出ユニット５２６に出力するように構成され得る。顕著成分分析ユニット５２４はまた、１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅを量子化ユニット５３４に出力することができる。量子化ユニット５３４は、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇを生成するために、１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅを量子化することができる。量子化ユニット５３４は、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇを算出ユニット５２６に提供しながら、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇを（上で説明されたように）ベクトル再順序付けユニット５３２にも提供することができる。ベクトル再順序付けユニット５３２は、V^{T}_{DIST}ベクトル５２５Ｅに関して上で説明されたものと同様の方式で、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇに関して動作することができる。

【０８０２】

[0797]これらのベクトル５２５～５２５Ｇ（「ベクトル５２５」）を受け取ると、算出ユニット５２６はまず、音場の明瞭な成分を記述する明瞭な球面調和係数と、音場のバックグラウンド成分を記述したバックグラウンド球面調和係数とを決定することができる。行列算出ユニット５２６は、１つまたは複数のU_{DIST}５２５Ｃベクトルを１つまたは複数のS_{DIST}ベクトル５２５Ａと乗算し、次いで１つまたは複数のV^{T}_{DIST}ベクトル５２５Ｅと乗算することによって、明瞭な球面調和係数を決定するように構成され得る。算出ユニット５２６は、１つまたは複数のU_{BG}５２５Ｄベクトルを１つまたは複数のS_{BG}ベクトル５２５Ａと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトル５２５Ｅと乗算することによって、バックグラウンド球面調和係数を決定するように構成され得る。

【０８０３】

[0798]算出ユニット５２６は次いで、１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇに関して擬似逆演算を実行し、次いで明瞭な球面調和関数を１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５２５Ｇの擬似逆ベクトルと乗算することによって、１つまたは複数の補償された

U_{DIST}＊S_{DIST}ベクトル５２７’（これらのベクトルがE_{DIST}誤差を補償するための値を含むことを除き、U_{DIST}＊S_{DIST}ベクトル５２７と同様であり得る）を決定することができる。ベクトル再順序付けユニット５３２は、再順序付けられたベクトル５２７’を生成するために上で説明された方式で動作することができ、再順序付けられたベクトル５２７’は、上でやはり説明されたように、オーディオ符号化された再順序付けられたベクトル５１５’を生成するために、オーディオ符号化ユニット５１５Ａによってオーディオ符号化される。

【０８０４】

[0799]算出ユニット５２６は次に、バックグラウンド球面調和係数へとE_{DIST}誤差を射影することができる。算出ユニット５２６は、この射影を実行するために、明瞭な球面調和係数をバックグラウンド球面調和係数に加算することによって、オリジナルの球面調和係数５１１を決定し、または別様に復元することができる。算出ユニット５２６は次いで、V^T_{DIST}ベクトル５１９の量子化が原因の残りの誤差を決定するために、量子化された明瞭な球面調和係数（U_{DIST}ベクトル５２５ＣをS_{DIST}ベクトル５２５Ａと乗算し、次いで$V^T_{Q_DIST}$ベクトル５２５Ｇと乗算することによって生成され得る）およびバックグラウンド球面調和係数を球面調和係数５１１から差し引くことができる。算出ユニット５２６は次いで、補償された量子化されたバックグラウンド球面調和係数５３１’を生成するために、量子化されたバックグラウンド球面調和係数にこの誤差を加算することができる。

【０８０５】

[0800]いずれにしても、次数低減ユニット５２８Ａは、補償された量子化されたバックグラウンド球面調和係数５３１’を低減されたバックグラウンド球面調和係数５２９’へと低減するために、上で説明されたように実行することができ、低減されたバックグラウンド球面調和係数５２９’は、オーディオ符号化された低減されたバックグラウンド球面調和係数５１５Ｂ’を生成するために、上で説明された方式でオーディオ符号化ユニット５１４によってオーディオ符号化され得る。

【０８０６】

[0801]このようにして、本技法は、オーディオ符号化デバイス５１０Ｄが、音場の１つまたは複数の成分を表す、V^T_{DIST}ベクトル５２５Ｅなどの１つまたは複数の第１のベクトルを量子化することと、音場の同じ１つまたは複数の成分を同様に表す、U_{DIST}＊S_{DIST}ベクトル５２７および／またはバックグラウンド球面調和係数５３１のベクトルなどの１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされた誤差を補償することとを可能にし得る。

【０８０７】

[0802]その上、本技法は、以下の項に従ってこの量子化誤差の補償を提供することができる。

【０８０８】

[0803]項１３３１４６－１Ｂ。音場の１つまたは複数の明瞭な成分を表す１つまたは複数の第１のベクトルを量子化し、音場の同じ１つまたは複数の明瞭な成分を同様に表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償するように構成される、１つまたは複数のプロセッサを備える、オーディオ符号化デバイス５１０Ｄなどのデバイス。

【０８０９】

[0804]項１３３１４６－２Ｂ。１つまたは複数のプロセッサが、少なくとも一部、音場を記述する複数の球面調和係数に関して特異値分解を実行することによって生成される、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成される、項１３３１４６－１Ｂのデバイス。

【０８１０】

[0805]項１３３１４６－３Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複

数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、１つまたは複数のプロセッサが、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成される、項１３３１４６－１Ｂのデバイス。

【０８１１】

[0806]項１３３１４６－４Ｂ。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、１つまたは複数のプロセッサが、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成され、１つまたは複数のプロセッサが、Ｕ行列の１つまたは複数のＵベクトルをＳ行列の１つまたは複数のＳベクトルと乗算することによって計算される１つまたは複数のＵ＊Ｓベクトルにおいて量子化が原因でもたらされる誤差を補償するように構成される、項１３３１４６－１Ｂのデバイス。

【０８１２】

[0807]項１３３１４６－５Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、各々が音場の明瞭な成分の１つに対応するＵ行列の１つまたは複数のU_{DIST}ベクトルを決定し、各々が音場の明瞭な成分の同じ１つに対応するＳ行列の１つまたは複数のS_{DIST}ベクトルを決定し、各々が音場の明瞭な成分の同じ１つに対応するＶ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルを決定するように構成され、

【０８１３】

[0808]１つまたは複数のプロセッサが、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは複数のV^{T}_{DIST}ベクトルを量子化するように構成され、１つまたは複数のプロセッサが、１つまたは複数の誤差補償されたU_{DIST}＊S_{DIST}ベクトルを生成するために、Ｕ行列の１つまたは複数のU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによって計算される１つまたは複数のU_{DIST}＊S_{DIST}ベクトルにおいて量子化が原因でもたらされる誤差を補償するように構成される、項１３３１４６－１Ｂのデバイス。

【０８１４】

[0809]項１３３１４６－６Ｂ。１つまたは複数のプロセッサが、１つまたは複数のU_{DIST}ベクトルと、１つまたは複数のS_{DIST}ベクトルと、１つまたは複数のV^{T}_{DIST}ベクトルとに基づいて、明瞭な球面調和係数を決定し、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルによって明瞭な球面調和係数を除算し、それによって、V^{T}_{DIST}ベクトルの量子化を通じてもたらされた誤差を少なくとも一部補償する誤差補償された１つまたは複数のU_{C_DIST}＊S_{C_DIST}ベクトルを生成するために、$V^{T}_{Q_DIST}$ベクトルに関して擬似逆変換を実行するように構成される、項１３３１４６－５Ｂのデバイス。

【０８１５】

[0810]項１３３１４６－７Ｂ１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償されたU_{DIST}＊S_{DIST}ベクトルをオーディオ符号化するように構成される、項１３３１４６－５Ｂのデバイス。

【０８１６】

[0811]項１３３１４６－８Ｂ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、音場の１つまたは複数のバックグラウンド成分を記述するＵ行列の１つまたは複数のU_{BG}ベクトルと、音場の１つまたは複数の明瞭な成分を記述するＵ行列の１つまたは複数のU_{DIST}ベクトルとを決定し、音場の１つまたは

複数のバックグラウンド成分を記述するＳ行列の１つまたは複数のS_{BG}ベクトルと、音場の１つまたは複数の明瞭な成分を記述するＳ行列の１つまたは複数のS_{DIST}ベクトルとを決定し、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}ベクトルとを決定するように構成され、V^{T}_{DIST}ベクトルが音場の１つまたは複数の明瞭な成分を記述し、V^{T}_{DIST}ベクトルが音場の１つまたは複数の明瞭な成分を記述し、V^{T}_{BG}が音場の１つまたは複数のバックグラウンド成分を記述し、１つまたは複数のプロセッサが、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは複数のV^{T}_{DIST}ベクトルを量子化するように構成され、１つまたは複数のプロセッサがさらに、誤差補償されたバックグラウンド球面調和係数を生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算することによって形成されるバックグラウンド球面調和係数において量子化が原因でもたらされる誤差の少なくとも一部分を補償するように構成される、項１３３１４６－１Ｂのデバイス。

【０８１７】

[0812]項１３３１４６－９Ｂ。１つまたは複数のプロセッサが、V^{T}_{DIST}ベクトルと、U_{DIST}ベクトルをS_{DIST}ベクトルと乗算することによって形成される１つまたは複数の$U_{DIST} * S_{DIST}$ベクトルとに基づいて誤差を決定し、誤差補償されたバックグラウンド球面調和係数を生成するために、バックグラウンド球面調和係数に決定された誤差を加算するように構成される、項１３３１４６－８Ｂのデバイス。

【０８１８】

[0813]項１３３１４６－１０Ｂ。１つまたは複数のプロセッサがさらに、誤差補償されたバックグラウンド球面調和係数をオーディオ符号化するように構成される、項１３３１４６－８Ｂのデバイス。

【０８１９】

[0814]項１３３１４６－１１Ｂ。

【０８２０】

[0815]１つまたは複数のプロセッサが、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償するように構成され、１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するように構成される、項１３３１４６－１Ｂのデバイス。

【０８２１】

[0816]項１３３１４６－１２Ｂ。１つまたは複数のプロセッサが、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償するように構成され、１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償された第２のベクトルをオーディオ符号化し、オーディオ符号化された１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するように構成される、項１３３１４６－１Ｂのデバイス。

【０８２２】

[0817]項１３３１４６－１Ｃ。音場の１つまたは複数の明瞭な成分を表す１つまたは複数の第１のベクトルを量子化し、音場の１つまたは複数のバックグラウンド成分を表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされた誤差を補償するように構成される、１つまたは複数のプロセッサを備える、オーディオ符号化デバイス５１０Ｄなどのデバイス。

【０８２３】

[0818]項１３３１４６－２Ｃ。１つまたは複数のプロセッサが、少なくとも一部、音場

を記述する複数の球面調和係数に関して特異値分解を実行することによって生成される、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成される、項１３３１４６－１Ｃのデバイス。

【０８２４】

[0819]項１３３１４６－３Ｃ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行するように構成され、１つまたは複数のプロセッサが、Ｖ行列の転置から１つまたは複数のベクトルを量子化するように構成される、項１３３１４６－１Ｃのデバイス。

【０８２５】

[0820]項１３３１４６－４Ｃ。１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、各々が音場の明瞭な成分の１つに対応するＵ行列の１つまたは複数のU_{DIST}ベクトルを決定し、各々が音場の明瞭な成分の同じ１つに対応するＳ行列の１つまたは複数のS_{DIST}ベクトルを決定し、各々が音場の明瞭な成分の同じ１つに対応するＶ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルを決定するように構成され、１つまたは複数のプロセッサが、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは複数のV^{T}_{DIST}ベクトルを量子化するように構成され、１つまたは複数の誤差補償された$U_{DIST} * S_{DIST}$ベクトルを生成するために、Ｕ行列の１つまたは複数のU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによって計算される１つまたは複数の$U_{DIST} * S_{DIST}$ベクトルにおいて量子化が原因でもたらされる誤差の少なくとも一部分を補償するように構成される、項１３３１４６－１Ｃのデバイス。

【０８２６】

[0821]項１３３１４６－５Ｃ。１つまたは複数のプロセッサが、１つまたは複数のU_{DIST}ベクトルと、１つまたは複数のS_{DIST}ベクトルと、１つまたは複数のV^{T}_{DIST}ベクトルとに基づいて、明瞭な球面調和係数を決定し、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルによって明瞭な球面調和係数を除算し、それによって、V^{T}_{DIST}ベクトルの量子化を通じてもたらされた誤差を少なくとも一部補償する１つまたは複数の$U_{C_DIST} * S_{C_DIST}$ベクトルを生成するために、$V^{T}_{Q_DIST}$ベクトルに関して擬似逆変換を実行するように構成される、項１３３１４６－４Ｃのデバイス。

【０８２７】

[0822]項１３３１４６－６Ｃ。１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償された$U_{DIST} * S_{DIST}$ベクトルをオーディオ符号化するように構成される、項１３３１４６－４Ｃのデバイス。

【０８２８】

[0823]項１３３１４６－７Ｃ１つまたは複数のプロセッサがさらに、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す複数の球面調和係数に関して特異値分解を実行し、音場の１つまたは複数のバックグラウンド成分を記述するＵ行列の１つまたは複数のU_{BG}ベクトルと、音場の１つまたは複数の明瞭な成分を記述するＵ行列の１つまたは複数のU_{DIST}ベクトルとを決定し、音場の１つまたは複数のバックグラウンド成分を記述するＳ行列の１つまたは複数のS_{BG}ベクトルと、音場の１つまたは複数の明瞭な成分を記述するＳ行列の１つまたは複数のS_{DIST}ベクトルとを決定し、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}ベクトルとを決定するように構成され、V^{T}_{DIST}ベクトルが音場の１つまたは複数の明瞭な成分を記述し、V^{T}_{BG}が音場の１つまたは複数のバックグラウンド成分を記述し、１つまたは複数のプロセッサが、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために１つまたは

複数のV^{T}_{DIST}ベクトルを量子化するように構成され、１つまたは複数のプロセッサが、誤差補償されたバックグラウンド球面調和係数を生成するために、１つまたは複数のU_{BG}ベクトルを１つまたは複数のS_{BG}ベクトルと乗算し、次いで１つまたは複数のV^{T}_{BG}ベクトルと乗算することによって形成されるバックグラウンド球面調和係数において量子化が原因でもたらされる誤差を補償するように構成される、項１３３１４６－１Ｃのデバイス。

【０８２９】

[0824]項１３３１４６－８Ｃ。１つまたは複数のプロセッサが、V^{T}_{DIST}ベクトルと、U_{DIST}ベクトルをS_{DIST}ベクトルと乗算することによって形成される１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルとに基づいて誤差を決定し、誤差補償されたバックグラウンド球面調和係数を生成するために、バックグラウンド球面調和係数に決定された誤差を加算するように構成される、項１３３１４６－７Ｃのデバイス。

【０８３０】

[0825]項１３３１４６－９Ｃ。１つまたは複数のプロセッサがさらに、誤差補償されたバックグラウンド球面調和係数をオーディオ符号化するように構成される、項１３３１４６－７Ｃのデバイス。

【０８３１】

[0826]項１３３１４６－１０Ｃ。１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償し、１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するように構成される、項１３３１４６－１Ｃのデバイス。

【０８３２】

[0827]項１３３１４６－１１Ｃ。１つまたは複数のプロセッサがさらに、１つまたは複数の誤差補償された第２のベクトルを生成するために、音場の同じ１つまたは複数の成分を同様に表す１つまたは複数の第２のベクトルにおいて１つまたは複数の第１のベクトルの量子化が原因でもたらされる誤差を補償し、１つまたは複数の誤差補償された第２のベクトルをオーディオ符号化し、オーディオ符号化された１つまたは複数の誤差補償された第２のベクトルと量子化された１つまたは複数の第１のベクトルとを含むようにビットストリームを生成するように構成される、項１３３１４６－１Ｃのデバイス。

【０８３３】

[0828]言い換えれば、帯域幅低減の目的でＨｏＡ信号に対してフレームベースのＳＶＤ（またはＫＬＴおよびＰＣＡなどの関連する方法）の分解を使用するとき、本開示で説明される技法は、オーディオ符号化デバイス１０Ｄが、Ｕ行列の最初のいくつかのベクトル（Ｓ行列の対応する特異値によって乗算された）と、さらにはＶベクトルの対応するベクトルとを量子化することを可能にし得る。これは、音場の「フォアグラウンド」成分または「明瞭な」成分を備える。本技法は次いで、オーディオ符号化デバイス５１０ＤがＡＡＣ符号化器などの「ブラックボックス」オーディオコーディングエンジンを使用してＵ＊Ｓベクトルをコーディングすることを可能にし得る。Ｖベクトルは、スカラー量子化されるか、ベクトル量子化されるかのいずれかであり得る。

【０８３４】

[0829]加えて、Ｕ行列中の残りのベクトルのいくつかは、Ｓ行列およびＶ行列の対応する特異値によって乗算され、また、「ブラックボックス」オーディオコーディングエンジンを使用してコーディングされ得る。これらは、音場の「バックグラウンド」成分を備える。Ｖベクトルの単純な１６ビットのスカラー量子化は、４次（２５個の係数）では約８０ｋｂｐｓのオーバーヘッドを、６次（４９個の係数）では約１６０ｋｂｐｓのオーバーヘッドをもたらし得る。より粗い量子化は、より大きい量子化誤差をもたらし得る。本開示で説明される技法は、Ｖベクトルの量子化誤差をフォアグラウンド成分およびバックグラウンド成分に「射影する」ことによって、Ｖベクトルの量子化誤差を補償することがで

きる。

【０８３５】

[0830]本開示の技法は、実際のＶベクトルの量子化されたバージョンを計算することを含み得る。この量子化されたＶベクトルは、Ｖ’（ここでＶ’＝Ｖ＋ｅである）と呼ばれ得る。フォアグラウンド成分に対して本技法が再作成することを試みている背後のＨｏＡ信号は、Ｈ＿ｆ＝ＵＳＶによって与えられ、Ｕ、Ｓ、およびＶはフォアグラウンド要素のみを含む。この議論では、ＵＳはベクトルＵの単一のセットによって置き換えられる。したがって、Ｈ＿ｆ＝ＵＶである。Ｖ’に誤差が多いとすると、本技法はＨ＿ｆを可能な限り厳密に再作成することを試みている。したがって、本技法は、オーディオ符号化デバイス１０Ｄが、Ｈ＿ｆ＝Ｕ’Ｖ’となるような Ｕ’を見つけることを可能にし得る。オーディオ符号化デバイス１０Ｄは、Ｕ’＝Ｈ＿ｆ［Ｖ’］＾（－１）を可能にする、擬似逆変換方法を使用することができる。Ｕ’をコーディングするためにいわゆる「ブラックボックス」オーディオコーディングエンジンを使用して、本技法は、誤差の多いＶ’ベクトルと呼ばれ得るものによって引き起こされる、Ｈの誤差を最小にすることができる。

【０８３６】

[0831]同様の方法で、本技法はまた、オーディオ符号化デバイスが、Ｖを量子化することが原因の誤差をバックグラウンド要素へと射影することを可能にし得る。オーディオ符号化デバイス５１０Ｄは、フォアグラウンドＨｏＡ信号とバックグラウンドＨｏＡ信号の組合せ、すなわちＨ＝Ｈ＿ｆ＋Ｈ＿ｂである全体のＨｏＡ信号を再作成するように構成され得る。これは、Ｖ’の量子化誤差が原因で、Ｈ＝Ｈ＿ｆ＋ｅ＋Ｈ＿ｂとしてモデル化され得る。このようにして、「ブラックボックスオーディオコーダ」にＨ＿ｂを通す代わりに、（ｅ＋Ｈ＿ｂ）をオーディオコーダに通し、実質的にＶ’の誤差を補償する。実際には、これは、最大でも、バックグラウンド要素のために送信すべきであるとオーディオ符号化デバイス５１０Ｄによって決定された次数までしか補償しない。

【０８３７】

[0832]図４０Ｅは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｅを示すブロック図である。オーディオ符号化デバイス５１０Ｅは、オーディオ符号化デバイス５１０Ｅがオーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含む点において、オーディオ符号化デバイス５１０Ｄと類似であり得る。その上、オーディオ符号化デバイス５１０Ｅのオーディオ圧縮ユニット５１２は、オーディオ圧縮ユニット５１２が分解ユニット５１８を含む点において、オーディオ符号化デバイス５１０Ｄのものと類似であり得る。

【０８３８】

[0833]ただし、オーディオ符号化デバイス５１０Ｅのオーディオ圧縮ユニット５１２は、音場成分抽出ユニット５２０の数学ユニット５２６が、ビットストリーム５１７にＶ行列５１９Ａの転置の低減されたバージョンを含める前に、Ｖ行列５１９Ａをさらに低減するために、本開示で説明される技法の追加の態様を実行する点において、オーディオ符号化デバイス５１０Ｄのオーディオ圧縮ユニット５１２とは異なり得る。この理由で、オーディオ符号化デバイス５１０Ｅの音場成分抽出ユニット５２０は、「音場成分抽出ユニット５２０Ｅ」として示される。

【０８３９】

[0834]図４０Ｅの例では、次数低減ユニット５２８は、低減されたバックグラウンド球面調和係数５２９’をオーディオ符号化ユニット５１４に転送するのではなく、低減されたバックグラウンド球面調和係数５２９’を数学ユニット５２６に返す。上で述べられたように、これらの低減されたバックグラウンド球面調和係数５２９’は、球面基底関数に対応する係数のうちで、１つまたは複数の識別された次数および／または位数を有するものを除去することによって、低減されていてよい。低減されたバックグラウンド球面調和係数５２９’の低減された次数は、変数N_{BG}によって示され得る。

【０８４０】

[0835]音場成分抽出ユニット５２０Ｅが、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトル５３３’に関して次数低減を実行しなくてよいとすると、音場の明瞭な成分を記述する球面調和係数のこの分解の次数（変数N_{DIST}によって示され得る）は、バックグラウンド次数N_{BG}よりも大きくなり得る。言い換えれば、N_{BG}は、一般にN_{DIST}よりも小さくなり得る。N_{BG}がN_{DIST}よりも小さくなり得る１つの可能な理由は、高次球面基底関数が必要とされないように、バックグラウンド成分があまり指向性を有しておらず、それによって次数低減を可能にし、N_{BG}がN_{DIST}よりも小さくなることをもたらすと仮定されることである。

【０８４１】

[0836]図４０Ａ～図４０Ｄの例に示されるように、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５３９が、ビットストリーム５１７においてこれらのベクトル５３９をオーディオ符号化することなく、以前にオープンに送られたとすると、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５３９は、かなりの帯域幅を消費し得る。一例として、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５３９の各々は、１６ビットスカラー値に量子化されるとき、４次アンビソニックスオーディオデータではほぼ２０Ｋｂｐｓ（各ベクトルが２５個の係数を有する場合）と、６次アンビソニックスオーディオデータでは４０Ｋｂｐｓ（各ベクトルが４９個の係数を有する場合）とを消費し得る。

【０８４２】

[0837]本開示で説明される技法の様々な態様によれば、音場成分抽出ユニット５２０Ｅは、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトル５３９など、球面調和係数またはその分解のために指定される必要があるビットの量を低減することができる。いくつかの例では、数学ユニット５２６は、次数低減された球面調和係数５２９’に基づいて、再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９のうちで、除去されて、次数低減された球面調和係数５２９’と再び組み合わせられるべきであるものと、再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９のうちで、V^{T}_{SMALL}ベクトル５２１を形成するべきであるものとを決定することができる。すなわち、数学ユニット５２６は、次数低減された球面調和係数５２９’の次数を決定することができ、ここで、この次数はN_{BG}と示され得る。再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９は、変数N_{DIST}によって示された次数のものであってよく、ここで、N_{DIST}は次数N_{BG}よりも大きい。

【０８４３】

[0838]数学ユニット５２６は次いで、再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９の第１のN_{BG}次数を解析し、球面基底関数に対応する、分解された球面調和係数を指定するベクトルのうちで、N_{BG}以下の次数を有するものを除去することができる。これらの除去された再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９は、次いで、中間の明瞭な球面調和係数を形成するために、球面基底関数に対応する球面調和係数５１１の分解されたバージョンを表す再順序付けられたU_{DIST}＊S_{DIST}ベクトル５３３’のうちで、N_{BG}以下の次数を有するものを、除去された再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９と乗算することによって、中間の球面調和係数を形成するために使用され得る。数学ユニット５２６は次いで、中間の明瞭な球面調和係数を、次数低減された球面調和係数５２９’に加算することによって、修正されたバックグラウンド球面調和係数５３７を生成することができる。数学ユニット５２６は次いで、この修正されたバックグラウンド球面調和係数５３７をオーディオ符号化ユニット５１４に渡すことができ、オーディオ符号化ユニット５１４は、オーディオ符号化された修正されたバックグラウンド球面調和係数５１５Ｂ’を形成するために、これらの係数５３７をオーディオ符号化する。

【０８４４】

[0839]数学ユニット５２６は次いで、１つまたは複数のV^{T}_{SMALL}ベクトル５２１を渡すことができ、V^{T}_{SMALL}ベクトル５２１は、球面基底関数に対応する球面調和係数５１１の分解された形態を表すベクトル５３９のうちで、N_{BG}よりも大きくN_{DIST}以下の次数を有するものを表し得る。この点で、数学ユニット５２６は、図４の例に示されるオーディオ符号化デバイス２０の係数低減ユニット４６に類似した動作を実行することができる。数

学ユニット５２６は、１つまたは複数のV^T_{SMALL}ベクトル５２１をビットストリーム生成ユニット５１６に渡すことができ、ビットストリーム生成ユニット５１６は、V^T_{SMALL}ベクトル５２１をしばしばそれらのオリジナルのオーディオ符号化されていない形態において含むように、ビットストリーム５１７を生成することができる。V^T_{SMALL}ベクトル５２１が、再順序付けられた$V^T_{Q_DIST}$ベクトル５３９よりも少ないベクトルを含むとすると、本技法は、ビットストリーム５１７中でV^T_{SMALL}ベクトル５２１のみを指定することによって、再順序付けられた$V^T_{Q_DIST}$ベクトル５３９へのより少ないビットの割振りを容易にすることができる。

【０８４５】

[0840]量子化されていないように示されるが、いくつかの例では、オーディオ符号化デバイス５１０Ｅは、V^T_{BG}ベクトル５２５Ｆを量子化することができる。いくつかの例では、オーディオ符号化ユニット５１４がバックグラウンド球面調和係数を圧縮するために使用されないときなど、オーディオ符号化デバイス５１０Ｅは、V^T_{BG}ベクトル５２５Ｆを量子化することができる。

【０８４６】

[0841]このようにして、本技法は、オーディオ符号化デバイス５１０Ｅが、ビットストリーム中の１つまたは複数のベクトルに割り振られることが必要とされるビットの量を低減するために、バックグラウンド球面調和係数と再び組み合わせられるべき、球面調和係数から分解された１つまたは複数のベクトルのうちの少なくとも１つを決定することを可能にすることができ、球面調和係数が音場を記述し、バックグラウンド球面調和係数が、同じ音場の１つまたは複数のバックグラウンド成分を記述した。

【０８４７】

[0842]すなわち、本技法は、オーディオ符号化デバイス５１０Ｅが、次の項によって示された方式で構成されることを可能にすることができる。

【０８４８】

[0843]１３３１４９－１Ａ項。オーディオ符号化デバイス５１０Ｅなどのデバイスであって、ビットストリーム中の１つまたは複数のベクトルに割り振られることが必要とされるビットの量を低減するために、バックグラウンド球面調和係数と再び組み合わせられるべき、球面調和係数から分解された１つまたは複数のベクトルのうちの少なくとも１つを決定するように構成される１つまたは複数のプロセッサを備え、球面調和係数が音場を記述し、バックグラウンド球面調和係数が、同じ音場の１つまたは複数のバックグラウンド成分を記述した、デバイス。

【０８４９】

[0844]１３３１４９－２Ａ項。１つまたは複数のプロセッサが、１つまたは複数のベクトルから、１つまたは複数のベクトルのうちの決定された少なくとも１つを除去することによって、１つまたは複数のベクトルの低減されたセットを生成するようにさらに構成される、１３３１４９－１Ａ項のデバイス。

【０８５０】

[0845]１３３１４９－３Ａ項。１つまたは複数のプロセッサが、１つまたは複数のベクトルから、１つまたは複数のベクトルのうちの決定された少なくとも１つを除去することによって、１つまたは複数のベクトルの低減されたセットを生成することと、修正されたバックグラウンド球面調和係数を生成するために、１つまたは複数のベクトルのうちの除去された少なくとも１つを、バックグラウンド球面調和係数と再び組み合わせることと、１つまたは複数のベクトルの低減されたセットと、修正されたバックグラウンド球面調和係数とを含むように、ビットストリームを生成することとを行うようにさらに構成される、１３３１４９－１Ａ項のデバイス。

【０８５１】

[0846]１３３１４９－４Ａ項。１つまたは複数のベクトルの低減されたセットが、最初にオーディオ符号化されることなく、ビットストリームに含まれる、１３３１４９－３Ａ項のデバイス。

【０８５２】
[0847]１３３１４９－５Ａ項。１つまたは複数のプロセッサが、１つまたは複数のベクトルから、１つまたは複数のベクトルのうちの決定された少なくとも１つを除去することによって、１つまたは複数のベクトルの低減されたセットを生成することと、修正されたバックグラウンド球面調和係数を生成するために、１つまたは複数のベクトルのうちの除去された少なくとも１つを、バックグラウンド球面調和係数と再び組み合わせることと、修正されたバックグラウンド球面調和係数をオーディオ符号化することと、１つまたは複数のベクトルの低減されたセットと、オーディオ符号化された修正されたバックグラウンド球面調和係数とを含むように、ビットストリームを生成することとを行うようにさらに構成される、１３３１４９－１Ａ項のデバイス。

【０８５３】
[0848]１３３１４９－６Ａ項。１つまたは複数のベクトルが、音場の１つまたは複数の明瞭な成分の少なくともある態様を表すベクトルを備える、１３３１４９－１Ａ項のデバイス。

【０８５４】
[0849]１３３１４９－７Ａ項。１つまたは複数のベクトルが、音場を記述する複数の球面調和係数に関して特異値分解を実行することによって少なくとも一部、生成されるＶ行列の転置からの１つまたは複数のベクトルを備える、１３３１４９－１Ａ項のデバイス。

【０８５５】
[0850]１３３１４９－８Ａ項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、複数の球面調和係数に関して特異値分解を実行するようにさらに構成され、１つまたは複数のベクトルが、Ｖ行列の転置からの１つまたは複数のベクトルを備える、１３３１４９－１Ａ項のデバイス。

【０８５６】
[0851]１３３１４９－９Ａ項。１つまたは複数のプロセッサが、球面基底関数に対応するバックグラウンド球面調和係数のうちで、識別された次数および／または位数を有するものを除去するために、バックグラウンド球面調和係数に関して次数低減を実行するようにさらに構成され、バックグラウンド球面調和係数が次数N_{BG}に対応する、１３３１４９－１Ａ項のデバイス。

【０８５７】
[0852]１３３１４９－１０Ａ項。１つまたは複数のプロセッサが、球面基底関数に対応するバックグラウンド球面調和係数のうちで、識別された次数および／または位数を有するものを除去するために、バックグラウンド球面調和係数に関して次数低減を実行するようにさらに構成され、バックグラウンド球面調和係数が、明瞭な球面調和係数の次数N_{DIST}未満である次数N_{BG}に対応し、明瞭な球面調和係数が音場の明瞭な成分を表す、１３３１４９－１Ａ項のデバイス。

【０８５８】
[0853]１３３１４９－１１Ａ項。１つまたは複数のプロセッサが、球面基底関数に対応するバックグラウンド球面調和係数のうちで、識別された次数および／または位数を有するものを除去するために、バックグラウンド球面調和係数に関して次数低減を実行するようにさらに構成され、バックグラウンド球面調和係数が、明瞭な球面調和係数の次数N_{DIST}未満である次数N_{BG}に対応し、明瞭な球面調和係数が音場の明瞭な成分を表し、次数低減を受けない、１３３１４９－１Ａ項のデバイス。

【０８５９】
[0854]１３３１４９－１２Ａ項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、複数の球面調和係数に関して特異値分解を実行することと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}とを決定することとを行うようにさらに構成され、１つまたは

複数のV^{T}_{DIST}ベクトルが、音場の１つまたは複数の明瞭な成分を記述し、１つまたは複数のV^{T}_{BG}ベクトルが、音場の１つまたは複数のバックグラウンド成分を記述し、１つまたは複数のベクトルが、１つまたは複数のV^{T}_{DIST}ベクトルを含む、１３３１４９－１Ａ項のデバイス。

【０８６０】

[0855]１３３１４９－１３Ａ項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、複数の球面調和係数に関して特異値分解を実行することと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}とを決定することと、１つまたは複数のV_{DIST}ベクトルが、音場の１つまたは複数の明瞭な成分を記述し、１つまたは複数のV_{BG}ベクトルが、音場の１つまたは複数のバックグラウンド成分を記述する、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために、１つまたは複数のV^{T}_{DIST}ベクトルを量子化することとを行うようにさらに構成され、１つまたは複数のベクトルが、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを含む、１３３１４９－１Ａ項のデバイス。

【０８６１】

[0856]１３３１４９－１４Ａ項。１つまたは複数のプロセッサが、Ｕ行列の１つまたは複数のU_{DIST}ベクトルと１つまたは複数のU_{BG}ベクトルとを決定することと、１つまたは複数のU_{DIST}ベクトルが、音場の１つまたは複数の明瞭な成分を記述し、１つまたは複数のU_{BG}ベクトルが、音場の１つまたは複数のバックグラウンド成分を記述する、Ｓ行列の１つまたは複数のS_{DIST}ベクトルと１つまたは複数のS_{BG}ベクトルとを決定することと、１つまたは複数のS_{DIST}ベクトルが、音場の１つまたは複数の明瞭な成分を記述し、１つまたは複数のS_{BG}ベクトルが、音場の１つまたは複数のバックグラウンド成分を記述する、を行うようにさらに構成される、１３３１４９－１２Ａ項または１３３１４９－１３Ａ項のいずれかのデバイス。

【０８６２】

[0857]１３３１４９－１５Ａ項。１つまたは複数のプロセッサが、１つまたは複数のU_{BG}ベクトル、１つまたは複数のS_{BG}ベクトル、および１つまたは複数のV^{T}_{BG}の関数として、バックグラウンド球面調和係数を決定することと、N_{BG}に等しい次数を有する低減されたバックグラウンド球面調和係数を生成するために、バックグラウンド球面調和係数に関して次数低減を実行することと、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを生成するために、１つまたは複数のU_{DIST}を１つまたは複数のS_{DIST}ベクトルと乗算することと、１つまたは複数のベクトルの低減されたセットを生成するために、１つまたは複数のベクトルから、１つまたは複数のベクトルのうちの決定された少なくとも１つを除去することと、中間の明瞭な球面調和係数を生成するために、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを、１つもしくは複数のV^{T}_{DIST}ベクトルまたは１つもしくは複数の$V^{T}_{Q_DIST}$ベクトルのうちの除去された少なくとも１つと乗算することと、１つもしくは複数のV^{T}_{DIST}ベクトルまたは１つもしくは複数の$V^{T}_{Q_DIST}$ベクトルのうちの除去された少なくとも１つを、バックグラウンド球面調和係数と再び組み合わせるために、中間の明瞭な球面調和係数をバックグラウンド球面調和係数に加算することとを行うようにさらに構成される、１３３１４９－１４Ａ項のデバイス。

【０８６３】

[0858]１３３１４９－１６Ａ項。１つまたは複数のプロセッサが、１つまたは複数のU_{BG}ベクトル、１つまたは複数のS_{BG}ベクトル、および１つまたは複数のV^{T}_{BG}の関数として、バックグラウンド球面調和係数を決定することと、N_{BG}に等しい次数を有する低減されたバックグラウンド球面調和係数を生成するために、バックグラウンド球面調和係数に関して次数低減を実行することと、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを生成するために、１つまたは複数のU_{DIST}を１つまたは複数のS_{DIST}ベクトルと乗算することと、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを生成するために、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを再順序付けることと、１つまたは複数のベクトルの低

減されたセットを生成するために、１つまたは複数のベクトルから、１つまたは複数のベクトルのうちの決定された少なくとも１つを除去することと、中間の明瞭な球面調和係数を生成するために、再順序付けられた１つまたは複数のU_{DIST}＊S_{DIST}ベクトルを、１つもしくは複数のV^{T}_{DIST}ベクトルまたは１つもしくは複数の$V^{T}_{Q_DIST}$ベクトルのうちの除去された少なくとも１つと乗算することと、１つもしくは複数のV^{T}_{DIST}ベクトルまたは１つもしくは複数の$V^{T}_{Q_DIST}$ベクトルのうちの除去された少なくとも１つを、バックグラウンド球面調和係数と再び組み合わせるために、中間の明瞭な球面調和係数をバックグラウンド球面調和係数に加算することとを行うようにさらに構成される、１３３１４９－１４Ａ項のデバイス。

【０８６４】

[0859]１３３１４９－１７Ａ項。１つまたは複数のプロセッサが、中間の明瞭な球面調和係数をバックグラウンド球面調和係数に加算した後、バックグラウンド球面調和係数をオーディオ符号化することと、オーディオ符号化されたバックグラウンド球面調和係数を含むように、ビットストリームを生成することとを行うようにさらに構成される、１３３１４９－１５Ａ項または１３３１４９－１６Ａ項のいずれかのデバイス。

【０８６５】

[0860]１３３１４９－１８Ａ項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、複数の球面調和係数に関して特異値分解を実行することと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルと１つまたは複数のV^{T}_{BG}とを決定することと、１つまたは複数のV_{DIST}ベクトルが、音場の１つまたは複数の明瞭な成分を記述し、１つまたは複数のV_{BG}ベクトルが、音場の１つまたは複数のバックグラウンド成分を記述し、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために、１つまたは複数のV^{T}_{DIST}ベクトルを量子化することと、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを生成するために、１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを再順序付けることとを行うようにさらに構成され、１つまたは複数のベクトルが、再順序付けられた１つまたは複数の$V^{T}_{Q_DIST}$ベクトルを含む、１３３１４９－１Ａ項のデバイス。

【０８６６】

[0861]図４０Ｆは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｆを示すブロック図である。オーディオ符号化デバイス５１０Ｆは、オーディオ符号化デバイス５１０Ｆがオーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含む点において、オーディオ符号化デバイス５１０Ｃと類似であり得る。その上、オーディオ符号化デバイス５１０Ｆのオーディオ圧縮ユニット５１２は、オーディオ圧縮ユニット５１２が、オーディオ符号化デバイス５１０Ｃの同様のユニットに類似して動作し得る、分解ユニット５１８とベクトル再順序付けユニット５３２とを含む点において、オーディオ符号化デバイス５１０Ｃのものと類似であり得る。いくつかの例では、オーディオ符号化デバイス５１０Ｆは、U_{DIST}ベクトル５２５Ｃ、U_{BG}ベクトル５２５Ｄ、V^{T}_{DIST}ベクトル５２５Ｅ、およびV^{T}_{BG}ベクトル５２５Ｊのうちのいずれかの１つまたは複数のベクトルを量子化するために、図４０Ｄおよび図４０Ｅに関して説明されたような、量子化ユニット５３４を含み得る。

【０８６７】

[0862]ただし、オーディオ符号化デバイス５１０Ｆのオーディオ圧縮ユニット５１２は、音場成分抽出ユニット５２０の顕著成分分析ユニット５２４が、図４０Ａ～図４０Ｊの状況でＤとして示される、フォアグラウンド成分の数を選択するために、コンテンツ分析を実行することができる点において、オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニット５１２とは異なり得る。言い換えれば、顕著成分分析ユニット５２４は、球面調和係数の分解されたバージョンが合成オーディオオブジェクトから生成されたか、マイクロフォンを用いた自然な録音から生成されたかを識別するために、上で説明された方式

でＵ、ＳおよびＶ行列５１９に関して動作することができる。顕著成分分析ユニット５２４は次いで、この合成決定に基づいてＤを決定することができる。

【０８６８】

[0863]その上、オーディオ符号化デバイス５１０Ｆのオーディオ圧縮ユニット５１２は、音場成分抽出ユニット５２０が、追加のユニット、次数低減およびエネルギー保存ユニット５２８Ｆ（「次数ｒｅｄ．およびエネルギーｐｒｓｖ．ユニット５２８Ｆ」と示される）を含み得る点において、オーディオ符号化デバイス５１０Ｃのオーディオ圧縮ユニット５１２とは異なり得る。これらの理由で、オーディオ符号化デバイス５１０Ｆの音場成分抽出ユニット５２０は、「音場成分抽出ユニット５２０Ｆ」として示される。

【０８６９】

[0864]次数低減およびエネルギー保存ユニット５２８Ｆは、完全V_{BG}行列５２５Ｈによって一部記述される音場の全体的なエネルギー（と、付随する音圧と）を保存しながら、複数の球面調和係数５１１の右特異ベクトルを表すV_{BG}行列５２５Ｈのバックグラウンド成分の次数低減を実行するように構成されるユニットを表す。この点で、次数低減およびエネルギー保存ユニット５２８Ｆは、図４の例に示されるオーディオ符号化デバイス２０のバックグラウンド選択ユニット４８およびエネルギー補償ユニット３８に関して上で説明されたものと類似した動作を実行することができる。

【０８７０】

[0865]完全V_{BG}行列５２５Ｈは、次元$(N+1)^2 \times (N+1)^2 - D$を有し、ただし、Ｄは、主成分、または言い換えれば、音場の明瞭なオーディオ成分であることに関して顕著であると決定される特異値の数を表す。すなわち、完全V_{BG}行列５２５Ｈは、音場のバックグラウンド（ＢＧ）、または言い換えれば、環境的もしくは不明瞭なオーディオ成分であると決定されるそれらの特異値を含む。

【０８７１】

[0866]たとえば、図４０Ｂ～図４０Ｅの次数低減ユニット５２４に関して上で説明されたように、次数低減およびエネルギー保存ユニット５２８Ｆは、V_{BG}行列５２５Ｈのバックグラウンド特異値のうちで、高次球面基底関数に対応するものを（しばしばゼロにすることによって）除去、排除、または別様に削除することができる。次数低減およびエネルギー保存ユニット５２８Ｆは、転置ユニット５２２に、V_{BG}行列５２５Ｈの低減されたバージョン（「V_{BG}’行列５２５Ｉ」と示され、以下で「低減されたV_{BG}’行列５２５Ｉ」と呼ばれる）を出力することができる。低減されたV_{BG}’行列５２５Ｉは、次元

【０８７２】

【数２８】

$$(\tilde{\eta}+1)^2 \mathrm{x}\ (N+1)^2 - D$$

【０８７３】

を有することができ、ただし

【０８７４】

【数２９】

$$\tilde{\eta} < N$$

【０８７５】

である。転置ユニット５２２は、転置された低減されたV^T_{BG}’行列５２５Ｊを生成し、数学ユニット５２６に出力するために、低減されたV_{BG}’行列５２５Ｉに転置演算を適用し、数学ユニット５２６は、U_{BG}行列５２５Ｄと、S_{BG}行列５２５Ｂと、転置された低減されたV^T_{BG}’行列５２５Ｊとを使用して、$U_{BG} * S_{BG} * V^T_{BG}$を計算することによって、音場のバックグラウンド音声成分を再構築するように動作することができる。

【０８７６】

[0867]本明細書で説明される技法によれば、次数低減およびエネルギー保存ユニット５２８Ｆは、低減されたＶ$_{BG}$’行列５２５Ｉを生成するために、完全Ｖ$_{BG}$行列５２５Ｈの次数を低減することによって引き起こされる、音場のバックグラウンド音声成分の全体的なエネルギーの起こり得る低減を補償するようにさらに構成される。いくつかの例では、次数低減およびエネルギー保存ユニット５２８Ｆは、低減されたＶ$_{BG}$’行列５２５Ｉを転置ユニット５２２に出力する前に、低減されたＶ$_{BG}$’行列５２５Ｉの二乗平均平方根（ＲＭＳ）エネルギーを完全Ｖ$_{BG}$行列５２５ＨのＲＭＳに等しく、または少なくともそれにより近くなるように増大させるために、低減されたＶ$_{BG}$’行列５２５Ｉの（Ｎ＋１）^{2}－Ｄ個の列の各々に適用すべき増幅値の形態で補償利得を決定することによって補償する。

【０８７７】

[0868]いくつかの例では、次数低減およびエネルギー保存ユニット５２８Ｆは、完全Ｖ$_{BG}$行列５２５Ｈの各列のＲＭＳエネルギーと、低減されたＶ$_{BG}$’行列５２５Ｉの各列のＲＭＳエネルギーとを決定し、次いで、次の式に示されるように、前者と後者との比としてその列のための増幅値を決定することができ、

$$\propto = v_{BG} / v_{BG}'$$

【０８７８】

[0869]ただし、∝は列のための増幅値であり、ｖ$_{BG}$はＶ$_{BG}$行列５２５Ｈの単一の列を表し、ｖ$_{BG}$’はＶ$_{BG}$’行列５２５Ｉの対応する単一の列を表す。これは、次のような行列表記で表されてよく、

【０８７９】

【数３０】

$$A = V_{BG}{}^{RMS} / V_{BG}{}'^{RMS},$$

$$A = [\propto_1 \ldots \propto_{(N+1)^2 - D}]$$

【０８８０】

ただし、Ｖ$_{BG}{}^{RMS}$は、Ｖ$_{BG}$行列５２５Ｈの各列のＲＭＳを示す要素を有するＲＭＳベクトルであり、Ｖ$_{BG}$’^{RMS}は、低減されたＶ$_{BG}$’行列５２５Ｉの各列のＲＭＳを示す要素を有するＲＭＳベクトルであり、Ａは、Ｖ$_{BG}$行列５２５Ｈの各列のための要素を有する増幅値ベクトルである。次数低減およびエネルギー保存ユニット５２８Ｆは、対応する増幅値α、またはベクトル形態で、下の式を使用して、低減されたＶ$_{BG}$行列５２５Ｉの各列にスカラー乗算を適用し、

$$V_{BG}'' = V_{BG}' A^{T}$$

[0870]ただし、Ｖ$_{BG}$’’は、エネルギー補償を含む低減されたＶ$_{BG}$’行列５２５Ｉを表す。次数低減およびエネルギー保存ユニット５２８Ｆは、低減されたＶ$_{BG}$’行列５２５ＩのＲＭＳを、完全Ｖ$_{BG}$行列５２５ＨのＲＭＳに等しくする（またはほぼ等しくする）ために、エネルギー補償を含む低減されたＶ$_{BG}$’行列５２５Ｉを転置ユニット５２２に出力することができる。エネルギー補償を含む低減されたＶ$_{BG}$’行列５２５Ｉの出力次元は、下の式であり得る。

【０８８１】

【数３１】

$$(\tilde{\eta}+1)^2 \times (N+1)^2 - D$$

【０８８２】

[0871]いくつかの例では、低減されたＶ$_{BG}$’行列５２５Ｉおよび完全Ｖ$_{BG}$行列５２５Ｈのそれぞれの列の各ＲＭＳを決定するために、次数低減およびエネルギー保存ユニット５２８Ｆはまず、基準球面調和係数（ＳＨＣ）レンダラを列に適用することができる。次数

低減およびエネルギー保存ユニット５２８Ｆによる基準ＳＨＣレンダラの適用は、低減されたＶ$_{BG}$’行列５２５Ｉおよび完全Ｖ$_{BG}$行列５２５Ｈによって表されるフレームの各列によって記述される音場全体のエネルギーを決定するために、ＳＨＣ領域におけるＲＭＳの決定を可能にする。したがって、そのような例では、次数低減およびエネルギー保存ユニット５２８Ｆは、完全Ｖ$_{BG}$行列５２５Ｈの各列に、および、低減されたＶ$_{BG}$’行列５２５Ｉの各低減された列に基準ＳＨＣレンダラを適用し、列および低減された列のためのそれぞれのＲＭＳ値を決定し、列のためのＲＭＳ値と低減された列へのＲＭＳ値との比として、列のための増幅値を決定することができる。いくつかの例では、低減されたＶ$_{BG}$’行列５２５Ｉへの次数低減は、エネルギー保存と同時発生的に列方向に進む。これは、次のように擬似コードで表され得る：

```
R = ReferenceRenderer;
for m = numDist+1 : numChannels
  fullV = V(:,m); // Vの１列を取る => fullV
  reducedV =[fullV(1:numBG); zeros(numChannels-numBG,1)];
  alpha=sqrt( sum((fullV'*R).^2)/sum((reducedV'*R).^2) );
  if isnan(alpha) || isinf(alpha), alpha = 1; end;
  V_out(:,m) = reducedV * alpha;
End
```

【０８８３】

[0872]上の擬似コードでは、ｎｕｍＣｈａｎｎｅｌｓは、（Ｎ＋１）^{2}－Ｄを表すことができ、ｎｕｍＢＧは、下の式を表すことができ、

【０８８４】

【数３２】

$$(\tilde{\eta}+1)^2$$

【０８８５】

Ｖは、Ｖ$_{BG}$行列５２５Ｈを表すことができ、Ｖ＿ｏｕｔは、低減されたＶ$_{BG}$’行列５２５Ｉを表すことができ、Ｒは、次数低減およびエネルギー保存ユニット５２８Ｆの基準ＳＨＣレンダラを表すことができる。Ｖの次元は、（Ｎ＋１）^{2}×（Ｎ＋１）^{2}－Ｄであり得、Ｖ＿ｏｕｔの次元は、下の式であり得る。

【０８８６】

【数３３】

$$(\tilde{\eta}+1)^2 \times (N+1)^2 - D$$

【０８８７】

[0873]結果として、オーディオ符号化デバイス５１０Ｆは、複数の球面調和係数５１１を表すとき、次数低減プロセスの結果として損失され得るエネルギーの補償を含む、次数低減されたＶ$_{BG}$’行列５２５Ｉを使用して、バックグラウンド音声成分を再構築することができる。

【０８８８】

[0874]図４０Ｇは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｇを示すブロック図である。図４０Ｇの例では、オーディオ符号化デバイス５１０Ｇは、音場成分抽出ユニット５２０Ｆを含む。今度は、音場成分抽出ユニット５２０Ｆは、顕著成分分析ユニット５２４Ｇを含む。

【０８８９】

[0875]ただし、オーディオ符号化デバイス５１０Ｇのオーディオ圧縮ユニット５１２は

、オーディオ符号化デバイス５１０Ｇのオーディオ圧縮ユニット５１２が顕著成分分析ユニット５２４Ｇを含む点において、オーディオ符号化デバイス１０Ｆのオーディオ圧縮ユニット５１２とは異なり得る。顕著成分分析ユニット５２４Ｇは、オーディオデータと関連付けられる指向性ベースの情報を使用して、音場を表すオーディオデータの顕著さまたは明瞭性を決定するように構成されるユニットを表し得る。

【０８９０】

[0876]エネルギーベースの決定は、音場の明瞭なオーディオ成分を識別するためにＳＶＤによって分解された音場のレンダリングを改善し得るが、バックグラウンドオーディオ成分が高いエネルギーレベルを示す場合は、エネルギーベースの決定はまた、デバイスに、明瞭なオーディオ成分としてバックグラウンドオーディオ成分を誤って識別させることがある。すなわち、エネルギーのみに基づいて明瞭なオーディオ成分とバックグラウンドオーディオ成分とを分離することは、高エネルギーの（たとえば、より大きい）バックグラウンドオーディオ成分が明瞭なオーディオ成分として誤って識別されることがあるので、安定的ではないことがある。音場の明瞭なオーディオ成分とバックグラウンドオーディオ成分とをより安定的に区別するために、本開示で説明される技法の様々な態様は、顕著成分分析ユニット５２４Ｇに、ＳＨＣ５１１の分解されたバージョンから明瞭なオーディオ成分とバックグラウンドオーディオ成分とを分離するためにＳＨＣ５１１の指向性ベースの分析を実行させることができる。

【０８９１】

[0877]顕著成分分析ユニット５２４Ｇは、図４０Ｈの例では、以前に説明されたオーディオ符号化デバイス５１０～５１０Ｆの顕著成分分析ユニット５２４に類似して、Ｖ行列５１９、Ｓ行列５１９Ｂ、およびＵ行列５１９Ｃのうちの１つまたは複数に含まれるバックグラウンド要素から、明瞭な（またはフォアグラウンド）要素を分離するように構成されるか、またはさもなければ動作可能なユニットを表す。いくつかのＳＶＤベースの技法に従って、最も高エネルギーの成分（たとえば、Ｖ、ＳおよびＵ行列５１９～５１９Ｃまたはそれらから導出された行列のうちの１つまたは複数の最初のいくつかのベクトル）は、明瞭な成分として扱われ得る。しかしながら、行列５１９～５１９Ｃのうちの１つまたは複数の最も高エネルギーの成分（ベクトルによって表される）は、すべての状況において、最も指向性のある成分／信号を表すとは限らない。

【０８９２】

[0878]以前に説明された顕著成分分析ユニット５２４とは異なり、顕著成分分析ユニット５２４Ｇは、行列５１９～５１９Ｃまたはそれらから導出された行列のうちの１つまたは複数のベクトルの指向性に基づいて、フォアグラウンド要素を識別するために、本明細書で説明される技法の１つまたは複数の態様を実施することができる。いくつかの例では、顕著成分分析ユニット５２４Ｇは、ベクトルのエネルギーと指向性の両方に基づいて、１つまたは複数のベクトルを明瞭なオーディオ成分として識別または選択することができる（ここで、成分は「オブジェクト」とも呼ばれ得る）。たとえば、顕著成分分析ユニット５２４Ｇは、行列５１９～５１９Ｃ（またはそれらから導出された行列）の１つまたは複数のベクトルのうち、高いエネルギーと（たとえば、指向性指数として表される）高い指向性の両方を示すものを、明瞭なオーディオ成分として識別することができる。結果として、特定のベクトルが行列５１９～５１９Ｃ（またはそれらから導出された行列）のうちの１つまたは複数の他のベクトルと比較されると比較的指向性が小さいと顕著成分分析ユニット５２４Ｇが決定する場合、特定のベクトルと関連付けられるエネルギーレベルとは無関係に、顕著成分分析ユニット５２４Ｇは、その特定のベクトルがＳＨＣ５１１によって表される音場のバックグラウンド（または環境）オーディオ成分を表すと決定することができる。この点で、顕著成分分析ユニット５２４Ｇは、図４の例に示されるオーディオ符号化デバイス２０の音場分析ユニット４４に関して上で説明されたものに類似した動作を実行することができる。

【０８９３】

[0879]いくつかの実装形態では、顕著成分分析ユニット５２４Ｇは、以下の演算を実行

することによって、指向性に基づいて明瞭なオーディオオブジェクト（上で述べられたように、「成分」とも呼ばれ得る）を識別することができる。顕著成分分析ユニット５２４Ｇは、（たとえば、１つまたは複数の行列乗算プロセスを使用して）Ｖ行列５１９ＡをＳ行列５１９Ｂと乗算することができる。Ｖ行列５１９ＡをＳ行列５１９Ｂと乗算することによって、顕著成分分析ユニット５２４Ｇは、ＶＳ行列を得ることができる。加えて、顕著成分分析ユニット５２４Ｇは、ＶＳ行列の（行であり得る）ベクトルの各々の成分の少なくともいくつかを二乗する（すなわち、２という冪によって冪乗する）ことができる。いくつかの例では、顕著成分分析ユニット５２４Ｇは、１よりも大きい次数と関連付けられる各ベクトルの二乗された成分を加算することができる。一例として、行列の各ベクトルが２５個の成分を含む場合、顕著成分分析ユニット５２４Ｇは、各ベクトルに関して、５番目の成分から始まり２５番目の成分で終わるように各ベクトルの成分を二乗することができ、二乗された成分を加算して指向性指数（または指向性インジケータ）を決定する。各加算演算は、対応するベクトルの指向性指数をもたらし得る。この例では、顕著成分分析ユニット５２４Ｇは、１以下の次数と関連付けられる各行の成分、すなわち１番目から４番目の成分が、一般的にエネルギーの量をより対象としており、それらの成分の指向性はあまり対象としていないと、決定することができる。すなわち、０または１という次数と関連付けられる低次アンビソニックスは、図１および図２に示されるように、圧力波の方向に関して多くのことを提供せず、むしろ何らかの大きさ（これはエネルギーを表す）を提供する、球面基底関数に対応する。

【０８９４】

[0880]上の例で説明される演算はまた、以下の擬似コードに従って表され得る。以下の擬似コードは、（引用符を伴わない）文字列「/*」および「*/」の連続する実例の中に含まれるコメント記述の形式の、注釈を含む。

```
        [U,S,V] = svd(audioframe,'ecom');
        VS = V*S;
/*次の行は、各行を独立に分析することと、対応するベクトルの指向性指数または指向性尺度を決定するために５番目の成分から２５番目の成分まで（一例として）第１の行の中の値を加算することとを対象とする。加算の前に成分を二乗する。１よりも大きい次数と関連付けられる各行の中の成分は、高次アンビソニックスと関連付けられるので、指向性がある可能性がより高い。*/
        sumVS = sum(VS(5:end,:).^2,1);

/*次の行は、生成されたＶＳ行列の二乗の合計をソートすることと、最大の値のセット（たとえば、最大の値の３つまたは４つ）を選択することとを対象とする*/
        [~,idxVS] = sort(sumVS,'descend');
        U = U(:,idxVS);
        V = V(:,idxVS);
        S = S(idxVS,idxVS);
```

【０８９５】

[0881]言い換えれば、上の擬似コードに従って、顕著成分分析ユニット５２４Ｇは、ＳＨＣ５１１のうちで１よりも大きい次数を有する球面基底関数に対応するものから分解されたＶＳ行列の各ベクトルの成分を選択することができる。顕著成分分析ユニット５２４Ｇは次いで、ＶＳ行列の各ベクトルのこれらの成分を二乗することができ、二乗された成分を加算して、ＶＳ行列の各ベクトルの指向性尺度または指向性指数を識別し、計算し、または別様に決定する。次に、顕著成分分析ユニット５２４Ｇは、ベクトルの各々のそれぞれの指向性尺度に基づいて、ＶＳ行列のベクトルをソートすることができる。顕著成分分析ユニット５２４Ｇは、最高の対応する指向性をもつベクトルが最初になり、最低の対応する指向性をもつベクトルが最後となるように、指向性尺度の降順でこれらのベクトルをソートすることができる。顕著成分分析ユニット５２４Ｇは次いで、最高の相対的な指向性尺度を有するベクトルの０ではないサブセットを選択することができる。

【０８９６】
[0882]本明細書で説明される技法のいくつかの態様によれば、オーディオ符号化デバイス５１０Ｇ、またはその１つもしくは複数の構成要素は、明瞭なオーディオ成分として、ＶＳ行列のあらかじめ決定された数のベクトルを識別し、または別様に使用することができる。たとえば、ＶＳ行列の各行の成分５から２５を選択し、各それぞれのベクトルのための相対的な指向性尺度を決定するために、選択された成分を二乗および加算した後、顕著成分分析ユニット５２４Ｇは、明瞭なオーディオ成分を表すベクトルを識別するために、ベクトルの中でさらなる選択を実施することができる。いくつかの例では、顕著成分分析ユニット５２４Ｇは、ベクトルの指向性指数を比較することによって、ＶＳ行列のあらかじめ決定された数のベクトルを選択することができる。一例として、顕著成分分析ユニット５２４Ｇは、は、４つの最高の指向性指数を有する（および、ソートされたＶＳ行列の最初の４つのベクトルである）、ＶＳ行列において表された４つのベクトルを選択することができる。今度は、顕著成分分析ユニット５２４Ｇは、４つの選択されたベクトルが、音場の対応するＳＨＣ表現と関連付けられる４つの最も明瞭なオーディオオブジェクトを表すと決定することができる。
【０８９７】
[0883]いくつかの例では、顕著成分分析ユニット５２４Ｇは、上で説明されたように、４つの選択されたベクトルの明瞭性を反映するために、ＶＳ行列から導出されたベクトルを再順序付けることができる。一例では、顕著成分分析ユニット５２４Ｇは、４つの選択された成分がＶＳ行列の最上位に再配置されるように、ベクトルを再順序付けることができる。たとえば、顕著成分分析ユニット５２４Ｇは、４つの選択された成分のすべてが、得られる再順序付けられたＶＳ行列の最初（または最上）の行に配置されるように、ＶＳ行列を修正することができる。顕著成分分析ユニット５２４Ｇに関して本明細書で説明されるが、様々な実装形態では、ベクトル再順序付けユニット５３２など、オーディオ符号化デバイス５１０Ｇの他の構成要素が再順序付けを実行することができる。
【０８９８】
[0884]顕著成分分析ユニット５２４Ｇは、得られる行列（すなわち、場合によって、再順序付けられるか、またはそうでないＶＳ行列）を、ビットストリーム生成ユニット５１６に通信することができる。今度は、ビットストリーム生成ユニット５１６は、ビットストリーム５１７を生成するためにＶＳ行列５２５Ｋを使用することができる。たとえば、顕著成分分析ユニット５２４ＧがＶＳ行列５２５Ｋを再順序付けた場合、ビットストリーム生成ユニット５１６は、ＶＳ行列５２５Ｋの再順序付けられたバージョンの残りのベクトルを量子化または廃棄することによってなど、ＶＳ行列５２５Ｋの再順序付けられたバージョンの最上行を明瞭なオーディオオブジェクトとして使用することができる。ＶＳ行列５２５Ｋの再順序付けられたバージョンの残りのベクトルを量子化することによって、ビットストリーム生成ユニット１６は、残りのベクトルを環境またはバックグラウンドオーディオデータとして扱うことができる。
【０８９９】
[0885]顕著成分分析ユニット５２４ＧがＶＳ行列５２５Ｋを再順序付けていない例では、ビットストリーム生成ユニット５１６は、顕著成分分析ユニット５２４Ｇによって選択されるような、ＶＳ行列５２５Ｋの各行の特定の成分（たとえば、５番目から２５番目の成分）に基づいて、明瞭なオーディオデータをバックグラウンドオーディオデータと区別することができる。たとえば、ビットストリーム生成ユニット５１６は、ＶＳ行列５２５Ｋの各行の最初の４つの成分を量子化または廃棄することによって、ビットストリーム５１７を生成することができる。
【０９００】
[0886]このようにして、オーディオ符号化デバイス５１０Ｇ、および／または、顕著成分分析ユニット５２４Ｇなど、その構成要素は、音場を表す明瞭なオーディオオブジェクトとバックグラウンドオーディオデータとの間で区別するために、オーディオデータのより高い係数のエネルギーとより低い係数のエネルギーとの比を決定し、または別様に利用

するために、本開示の技法を実施することができる。たとえば、説明されたように、顕著成分分析ユニット５２４Ｇは、顕著成分分析ユニット５２４Ｈによって生成されたＶＳ行列５２５Ｋの様々な成分の値に基づいて、エネルギー比を利用することができる。Ｖ行列５１９ＡとＳ行列５１９Ｂとによって提供されたデータを組み合わせることによって、顕著成分分析ユニット５２４Ｇは、ベクトルおよび関係するデータ（たとえば、指向性指数）の形態で、オーディオデータの様々な成分の指向性と全体的なエネルギーの両方の情報を提供するために、ＶＳ行列５２５Ｋを生成することができる。より具体的には、Ｖ行列５１９Ａは、指向性決定に関係する情報を提供することができるが、Ｓ行列５１９Ｂは、オーディオデータの成分のための全体的なエネルギー決定に関係する情報を提供することができる。

【０９０１】

[0887]他の例では、顕著成分分析ユニット５２４Ｇは、再順序付けられたV^{T}_{DIST}ベクトル５３９を使用して、ＶＳ行列５２５Ｋを生成することができる。これらの例では、顕著成分分析ユニット５２４Ｇは、Ｓ行列５１９Ｂに基づくいかなる修正の前にも、Ｖ行列５１９に基づいて明瞭性を決定することができる。言い換えれば、これらの例によれば、顕著成分分析ユニット５２４Ｇは、ＶＳ行列５２５Ｋを生成するステップを実行することなく、Ｖ行列５１９のみを使用して、指向性を決定することができる。より具体的には、Ｖ行列５１９Ａは、オーディオデータの成分（たとえば、Ｖ行列５１９のベクトル）がミキシングされる方式の情報と、潜在的に、ベクトルによって伝達されるデータの様々な相乗効果の情報とを提供することができる。たとえば、Ｖ行列５１９Ａは、ＥｉｇｅｎＭｉｋｅ（登録商標）によってオーディオ符号化デバイス５１０Ｇに中継されるような、各オーディオ成分の到来方向など、ベクトルによって表される様々なオーディオ成分の「到来方向」の情報を提供することができる。本明細書で使用される場合、「オーディオデータの成分」という用語は、行列５１９またはそれらから導出された任意の行列のいずれかの「成分」と互換的に使用され得る。

【０９０２】

[0888]本開示の技法のいくつかの実装形態によれば、顕著成分分析ユニット５２４Ｇは、本明細書で説明される様々な決定を行うために、外部情報でＳＨＣ表現を補足または補強することができる。一例として、顕著成分分析ユニット５２４Ｇは、行列５１９～５１９Ｃに表された様々なオーディオ成分の顕著さを決定するために、外部情報でＳＨＣを補強することができる。別の例として、顕著成分分析ユニット５２４Ｇおよび／またはベクトル再順序付けユニット５３２は、明瞭なオーディオオブジェクトとバックグラウンドオーディオデータとの間で区別するために、外部データでＨＯＡを補強することができる。

【０９０３】

[0889]いくつかの例では、顕著成分分析ユニット５２４Ｇは、オーディオデータの部分（たとえば、明瞭なオーディオオブジェクト）がケインズ的エネルギー（Keynesian energy）を示すことを検出することができる。そのような明瞭なオブジェクトの一例は、変調する人間の音声と関連付けられ得る。変調する音声ベースのオーディオデータの場合、顕著成分分析ユニット５２４Ｇは、変調するデータのエネルギーが、残りの成分のエネルギーの比として、ほぼ一定（たとえば、閾値範囲内で一定）のままであるか、または経時的にほぼ静止したままであると決定することができる。従来は、ケインズ的エネルギーをもつ明瞭なオーディオ成分のエネルギー特性（たとえば、変調する音声と関連付けられるもの）が、あるオーディオフレームから別のオーディオフレームへと変化する場合、デバイスは、一連のオーディオ成分を単一の信号として識別することができないことがある。しかしながら、顕著成分分析ユニット５２４Ｇは、様々な行列においてベクトルとして表された明瞭なオブジェクトの指向性または開口を決定するために、本開示の技法を実施することができる。

【０９０４】

[0890]より具体的には、顕著成分分析ユニット５２４Ｇは、指向性および／または開口などの特性がオーディオフレームにわたって実質的に変化する可能性が低いと決定するこ

とができる。本明細書で使用される場合、開口は、オーディオデータ内の、高次係数と低次係数との比を表す。Ｖ行列５１９Ａの各行は、特定のＳＨＣに対応するベクトルを含み得る。顕著成分分析ユニット５２４Ｇは、低次ＳＨＣ（たとえば、１以下の次数と関連付けられる）が環境データを表す傾向があるが、高次成分が明瞭なデータを表す傾向があると決定することができる。加えて、顕著成分分析ユニット５２４Ｇは、多くの例では、高次ＳＨＣ（たとえば、１よりも大きい次数と関連付けられる）がより大きいエネルギーを示すことと、高次ＳＨＣと低次ＳＨＣとのエネルギー比がオーディオフレームとオーディオフレームの間で実質的に類似した（または、ほぼ一定の）ままであることとを決定することができる。

【０９０５】

[0891]顕著成分分析ユニット５２４Ｇの１つまたは複数の構成要素は、Ｖ行列５１９を使用して、指向性および開口など、オーディオデータの特性を決定することができる。このようにして、顕著成分分析ユニット５２４Ｇなど、オーディオ符号化デバイス５１０Ｇの構成要素は、指向性ベースの情報を使用して、顕著さを決定し、および／または明瞭なオーディオオブジェクトをバックグラウンドオーディオと区別するために、本明細書で説明される技法を実施することができる。顕著さおよび／または明瞭性を決定するために指向性を使用することによって、顕著成分分析ユニット５２４Ｇは、エネルギーベースのデータのみを使用して顕著さおよび／または明瞭性を決定するように構成されるデバイスの場合よりも、安定的な決定に達することができる。顕著さおよび／または明瞭性の指向性ベースの決定に関して上で説明されたが、顕著成分分析ユニット５２４Ｇは、行列５１９～５１９Ｃ（または、それらから導出された任意の行列）のうちの１つまたは複数のベクトルによって表されるような、オーディオデータの特定の成分の顕著さおよび／または明瞭性を決定するために、エネルギーなどの他の特性に加えて指向性を使用するために、本開示の技法を実施することができる。

【０９０６】

[0892]いくつかの例では、方法は、オーディオオブジェクトの１つまたは複数に対して決定された指向性に基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数（ＳＨＣ）から１つまたは複数の明瞭なオーディオオブジェクトを識別することを含む。一例では、方法は、オーディオオブジェクトと関連付けられる球面調和係数に基づいて、１つまたは複数のオーディオオブジェクトの指向性を決定することをさらに含む。いくつかの例では、方法は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、球面調和係数に関して特異値分解を実行することと、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すこととをさらに含み、１つまたは複数のオーディオオブジェクトのそれぞれの指向性を決定することが、少なくとも一部Ｖ行列に基づく。

【０９０７】

[0893]一例では、方法は、再順序付けられたＶ行列においてより大きい指向性指数を有するベクトルがより小さい指向性指数を有するベクトルよりも上に配置されるように、Ｖ行列の１つまたは複数のベクトルを再順序付けることをさらに含む。一例では、方法は、より大きい指向性指数を有するベクトルがより小さい指向性指数を有するベクトルよりも多くの指向性情報を含むと決定することをさらに含む。一例では、方法は、ＶＳ行列を生成するためにＶ行列をＳ行列と乗算することをさらに含み、ＶＳ行列が１つまたは複数のベクトルを含む。一例では、方法は、１よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することと、対応する二乗された成分を形成するために、選択された成分の各々を二乗することと、ＶＳ行列の各行に対して、対応するベクトルの指向性指数を決定するために、二乗された成分のすべてを加算することとをさらに含む。

【０９０８】

[0894]いくつかの例では、ＶＳ行列の各行は、２５個の成分を含む。一例では、１よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することは、ＶＳ行列の各行

の５番目の成分において開始しＶＳ行列の各行の２５番目の成分において終了するすべての成分を選択することを含む。一例では、方法は、明瞭なオーディオオブジェクトを表すためにＶＳ行列のベクトルのサブセットを選択することをさらに含む。いくつかの例では、サブセットを選択することは、ＶＳ行列の４つのベクトルを選択することを含み、選択された４つのベクトルは、ＶＳ行列のベクトルのすべての４つの最大の指向性指数を有する。一例では、ベクトルの選択されたサブセットが明瞭なオーディオオブジェクトを表すと決定することは、各ベクトルの指向性とエネルギーの両方に基づく。

【０９０９】

[0895]いくつかの例では、方法は、オーディオオブジェクトの１つまたは複数に対して決定された指向性とエネルギーとに基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数から１つまたは複数の明瞭なオーディオオブジェクトを識別することを含む。一例では、方法は、オーディオオブジェクトと関連付けられる球面調和係数に基づいて、１つまたは複数のオーディオオブジェクトの指向性とエネルギーの一方または両方を決定することをさらに含む。いくつかの例では、方法は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す球面調和係数に関して特異値分解を実行することと、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すこととをさらに含み、１つまたは複数のオーディオオブジェクトのそれぞれの指向性を決定することが、少なくとも一部Ｖ行列に基づき、１つまたは複数のオーディオオブジェクトのそれぞれのエネルギーを決定することが、少なくとも一部Ｓ行列に基づく。

【０９１０】

[0896]一例では、方法は、ＶＳ行列を生成するためにＶ行列をＳ行列と乗算することをさらに含み、ＶＳ行列が１つまたは複数のベクトルを含む。いくつかの例では、方法は、１よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することと、対応する二乗された成分を形成するために選択された成分の各々を二乗することと、ＶＳ行列の各行に対して、ＶＳ行列の対応するベクトルの指向性指数を生成するために、二乗された成分のすべてを加算することとをさらに含む。いくつかの例では、ＶＳ行列の各行は、２５個の成分を含む。一例では、１よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することは、ＶＳ行列の各行の５番目の成分において開始しＶＳ行列の各行の２５番目の成分において終了するすべての成分を選択することを備える。いくつかの例では、方法は、明瞭なオーディオオブジェクトを表すためにベクトルのサブセットを選択することをさらに含む。一例では、サブセットを選択することは、ＶＳ行列の４つのベクトルを選択することを備え、選択された４つのベクトルは、ＶＳ行列のベクトルのすべての４つの最大の指向性指数を有する。いくつかの例では、ベクトルの選択されたサブセットが明瞭なオーディオオブジェクトを表すと決定することは、各ベクトルの指向性とエネルギーの両方に基づく。

【０９１１】

[0897]いくつかの例では、方法は、指向性ベースの情報を使用して、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定することを含み、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される。一例では、変換は、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成する、特異値分解を備える。一例では、変換は、音場の明瞭な成分と音場のバックグラウンド成分とを識別するための主成分分析を備える。

【０９１２】

[0898]いくつかの例では、デバイスは、本明細書で説明される技法のいずれか、またはそれらの技法の任意の組合せを実行するように構成されるか、またはさもなければ動作可

能である。いくつかの例では、コンピュータ可読記憶媒体は、実行されると、１つまたは複数のプロセッサに、本明細書で説明される技法のいずれか、またはそれらの技法の任意の組合せを実行させる命令とともに符号化される。いくつかの例では、デバイスは、本明細書で説明される技法のいずれか、またはそれらの技法の任意の組合せを実行するための手段を含む。

【０９１３】

[0899]すなわち、本技法の上記の態様は、オーディオ符号化デバイス５１０Ｇが、次の項に従って動作するように構成されることを可能にすることができる。

【０９１４】

[0900]１３４９５４－１Ｂ項。オーディオ符号化デバイス５１０Ｇなどのデバイスであって、オーディオオブジェクトの１つまたは複数に対して決定された指向性とエネルギーとに基づいて、オーディオオブジェクトと関連付けられる１つまたは複数の球面調和係数から１つまたは複数の明瞭なオーディオオブジェクトを識別するように構成される１つまたは複数のプロセッサを備えるデバイス。

【０９１５】

[0901]１３４９５４－２Ｂ項。１つまたは複数のプロセッサが、オーディオオブジェクトと関連付けられる球面調和係数に基づいて、１つまたは複数のオーディオオブジェクトの指向性とエネルギーの一方または両方を決定するようにさらに構成される、１３４９５４－１Ｂ項のデバイス。

【０９１６】

[0902]１３４９５４－３Ｂ項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、音場を表す球面調和係数に関して特異値分解を実行することと、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数の少なくとも一部分の関数として複数の球面調和係数を表すこととを行うようにさらに構成され、１つまたは複数のプロセッサが、少なくとも一部Ｖ行列に基づいて、１つまたは複数のオーディオオブジェクトのそれぞれの指向性を決定するように構成され、１つまたは複数のプロセッサが、少なくとも一部Ｓ行列に基づいて、１つまたは複数のオーディオオブジェクトのそれぞれのエネルギーを決定するように構成される、請求項１ＢまたはＢまたはそれらの組合せのいずれかのデバイス。

【０９１７】

[0903]１３４９５４－４Ｂ項。１つまたは複数のプロセッサが、ＶＳ行列を生成するためにＶ行列をＳ行列と乗算するようにさらに構成され、ＶＳ行列が１つまたは複数のベクトルを含む、１３４９５４－３Ｂ項のデバイス。

【０９１８】

[0904]１３４９５４－５Ｂ項。１つまたは複数のプロセッサが、１よりも大きい次数と関連付けられるＶＳ行列の各行の成分を選択することと、対応する二乗された成分を形成するために、選択された成分の各々を二乗することと、ＶＳ行列の各行に対して、ＶＳ行列の対応するベクトルの指向性指数を生成するために、二乗された成分のすべてを加算することとを行うようにさらに構成される、１３４９５４－４Ｂ項のデバイス。

【０９１９】

[0905]１３４９５４－６Ｂ項。ＶＳ行列の各行が２５個の成分を含む、請求項５ＢおよびＢまたはそれらの組合せのいずれかのデバイス。

【０９２０】

[0906]１３４９５４－７Ｂ項。１つまたは複数のプロセッサが、ＶＳ行列の各行の５番目の成分において開始しＶＳ行列の各行の２５番目の成分において終了するすべての成分を選択するように構成される、１３４９５４－６Ｂ項のデバイス。

【０９２１】

[0907]１３４９５４－８Ｂ項。１つまたは複数のプロセッサが、明瞭なオーディオオブジェクトを表すためにベクトルのサブセットを選択するようにさらに構成される、１３４

９５４－６Ｂ項および１３４９５４－７Ｂ項またはそれらの組合せのいずれかのデバイス。

【０９２２】

[0908]１３４９５４－９Ｂ項。１つまたは複数のプロセッサが、ＶＳ行列の４つのベクトルを選択するように構成され、選択された４つのベクトルは、ＶＳ行列のベクトルのすべての４つの最大の指向性指数を有する、１３４９５４－８Ｂ項のデバイス。

【０９２３】

[0909]１３４９５４－１０Ｂ項。１つまたは複数のプロセッサが、ベクトルの選択されたサブセットが、各ベクトルの指向性とエネルギーの両方に基づいて明瞭なオーディオオブジェクトを表すと決定するようにさらに構成される、１３４９５４－８Ｂ項および１３４９５４－９Ｂ項またはそれらの組合せのいずれかのデバイス。

【０９２４】

[0910]１３４９５４－１Ｃ項。オーディオ符号化デバイス５１０Ｇなどのデバイスであって、指向性ベースの情報を使用して、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するように構成される１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して変換を実行することによって生成される、デバイス。

【０９２５】

[0911]１３４９５４－２Ｃ項。変換が、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成する、特異値分解を備える、１３４９５４－１Ｃ項の方法。

【０９２６】

[0912]１３４９５４－３Ｃ項。１３４９５４－１Ａ項から１３４９５４－１２Ａ項、および１３４９５４－１Ｂ項から１３４９５４－９Ｂ項の任意の組合せによって記載される動作をさらに備える、１３４９５４－２Ｃ項の方法。

【０９２７】

[0913]１３４９５４－４Ｃ項。変換が、音場の明瞭な成分と音場のバックグラウンド成分とを識別するための主成分分析を備える、１３４９５４－１Ｃ項の方法。

【０９２８】

[0914]図４０Ｈは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｈを示すブロック図である。オーディオ符号化デバイス５１０Ｈは、オーディオ符号化デバイス５１０Ｈがオーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含む点において、オーディオ符号化デバイス５１０Ｇと類似であり得る。その上、オーディオ符号化デバイス５１０Ｈのオーディオ圧縮ユニット５１２は、オーディオ圧縮ユニット５１２が、オーディオ符号化デバイス５１０Ｇの同様のユニットに類似して動作し得る、分解ユニット５１８と音場成分抽出ユニット５２０Ｇとを含む点において、オーディオ符号化デバイス５１０Ｇのものと類似であり得る。いくつかの例では、オーディオ符号化デバイス５１０Ｈは、図４０Ｄ～図４０Ｅに関して説明されるように、U_{DIST}ベクトル５２５Ｃ、U_{BG}ベクトル５２５Ｄ、V^{T}_{DIST}ベクトル５２５Ｅ、およびV^{T}_{BG}ベクトル５２５Ｊのうちのいずれかの１つまたは複数のベクトルを量子化するために、量子化ユニット５３４を含み得る。

【０９２９】

[0915]ただし、オーディオ符号化デバイス５１０Ｈのオーディオ圧縮ユニット５１２は、オーディオ符号化デバイス５１０Ｈのオーディオ圧縮ユニット５１２が補間ユニット５５０として示される追加のユニットを含む点において、オーディオ符号化デバイス５１０Ｇのオーディオ圧縮ユニット５１２とは異なり得る。補間ユニット５５０は、図４５および図４５Ｂに関して以下でより詳細に説明されるように、第１のオーディオフレームおよび第２の時間的に後続のまたは先行するオーディオフレームのサブフレームから、第１の

オーディオフレームのサブフレームを補間するユニットを表し得る。補間ユニット５５０は、この補間を実行することにおいて、分解ユニット５１８がＳＨＣ５１１を分解するために必要とされる程度を潜在的に低減することによって、（処理サイクルおよび／またはメモリ消費に関して）計算的な複雑さを低減することができる。この点で、補間ユニット５５０は、図４の例に示されるオーディオ符号化デバイス２４の空間－時間的補間ユニット５０に関して上で説明されたものに類似した動作を実行することができる。

【０９３０】

[0916]すなわち、分解ユニット５１８によって実行される特異値分解は、潜在的に極めてプロセッサおよび／またはメモリ集約的であるが、また、いくつかの例では、特にＳＨＣ５１１の次数が増大するにつれて、ＳＨＣ５１１を分解するために大量の時間を要する。（処理サイクルおよび／またはメモリ消費に関して）時間の量を低減し、ＳＨＣ５１１の圧縮をより効率的にするために、本開示で説明される技法は、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の補間を提供することができ、ここで、サブフレームの各々は、ＳＨＣ５１１の分解されたバージョンを表し得る。フレーム全体に関してＳＶＤを実行するのではなく、本技法は、分解ユニット５１８が第１のオーディオフレームの第１のサブフレームを分解し、Ｖ行列５１９’を生成することを可能にすることができる。

【０９３１】

[0917]分解ユニット５１８はまた、第２のオーディオフレームの第２のサブフレームを分解することもでき、ここで、この第２のオーディオフレームは、第１のオーディオフレームに時間的に後続し、または時間的に先行し得る。分解ユニット５１８は、第２のオーディオフレームのこのサブフレームのためのＶ行列５１９’を出力することができる。補間ユニット５５０は次いで、第１のサブフレームおよび第２のサブフレームから分解されたＶ行列５１９’に基づいて、第１のオーディオフレームの残りのサブフレームを補間し、Ｖ行列５１９と、Ｓ行列５１９Ｂと、Ｕ行列５１９Ｃとを出力することができ、ここで、残りのサブフレームのための分解は、ＳＨＣ５１１と、第１のオーディオフレームのためのＶ行列５１９Ａと、第１のオーディオフレームの残りのサブフレームのための補間されたＶ行列５１９とに基づいて計算され得る。したがって、補間は、第１のオーディオフレームの残りのサブフレームのための分解の計算を回避することができる。

【０９３２】

[0918]その上、上で述べられたように、Ｕ行列５１９Ｃは、フレームとフレームの間で連続的でなくてもよく、ここで、ＳＨＣ５１１の第１のオーディオフレームから分解されたＵ行列５１９Ｃの明瞭な成分は、ＳＨＣ５１１の第２のオーディオフレームから分解されたＵ行列５１９Ｃにおけるものとは異なる行および／または列で指定され得る。この補間を実行することによって、フレーム境界（または言い換えれば、ＳＨＣ５１１のフレームへの区分）が原因でもたらされるあらゆるアーティファクトを低減し得る平滑化効果を線形補間が有し得るとすると、不連続性は低減され得る。この補間を実行するためにＶ行列５１９’を使用し、次いで、ＳＨＣ５１１からの補間されたＶ行列５１９’に基づいて、Ｕ行列５１９Ｃを復元することは、Ｕ行列５１９Ｃを再順序付けることからのいかなる影響をも平滑化することができる。

【０９３３】

[0919]演算において、補間ユニット５５０は、１つまたは複数のサブフレームのための分解された補間された球面調和係数を生成するために、第１のフレームに含まれる第１の複数の球面調和係数５１１の一部分の第１の分解、たとえばＶ行列５１９’と、第２のフレームに含まれる第２の複数の球面調和係数５１１の一部分の第２の分解、たとえばＶ行列５１９’とから、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。

【０９３４】

[0920]いくつかの例では、第１の分解は、第１の複数の球面調和係数５１１の部分の右特異ベクトルを表す第１のＶ行列５１９’を備える。同様に、いくつかの例では、第２の

分解は、第２の複数の球面調和係数の部分の右特異ベクトルを表す第２のＶ行列５１９’を備える。

【０９３５】

[0921]補間ユニット５５０は、第１のＶ行列５１９’および第２のＶ行列５１９’に基づいて、１つまたは複数のサブフレームに関して時間的補間を実行することができる。すなわち、補間ユニット５５０は、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’と、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’とに基づいて、たとえば、第１のオーディオフレームのための４つの全サブフレームからの第２のサブフレームと、第３のサブフレームと、第４のサブフレームとを時間的に補間することができる。いくつかの例では、この時間的補間は、線形時間的補間であり、ここで、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、第１のオーディオフレームの第４のサブフレームを補間するときよりも、第１のオーディオフレームの第２のサブフレームを補間するときにより重く重み付けられる。第３のサブフレームを補間するとき、Ｖ行列５１９’は均等に重み付けられ得る。第４のサブフレームを補間するとき、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’よりも重く重み付けられ得る。

【０９３６】

[0922]言い換えれば、線形時間的補間は、補間されるべき第１のオーディオフレームのサブフレームのうちの１つの近接を仮定して、Ｖ行列５１９’に重み付けることができる。補間されるべき第２のサブフレームについて、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、補間されるべき第２のサブフレームへのその近接を仮定して、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’よりも重く重み付けられる。重みは、Ｖ行列５１９’に基づいて第３のサブフレームを補間するとき、この理由のために等価であり得る。補間されるべき第４のサブフレームが、第１のオーディオフレームの第１のサブフレームよりも第２のオーディオフレームの第１のサブフレームにより近接しているとすると、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’に適用される重みは、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’に適用されるものよりも大きくなり得る。

【０９３７】

[0923]いくつかの例では、各オーディオフレームの第１のサブフレームのみが、補間を実行するために使用されるが、第１の複数の球面調和係数の部分は、第１の複数の球面調和係数５１１の４つのサブフレームのうちの２つを備え得る。これらの例および他の例において、第２の複数の球面調和係数５１１の部分は、第２の複数の球面調和係数５１１の４つのサブフレームのうちの２つを備える。

【０９３８】

[0924]上で述べられたように、単一のデバイス、たとえばオーディオ符号化デバイス５１０Ｈは、補間を実行し、一方でまた、第１の複数の球面調和係数の部分の第１の分解を生成するために、第１の複数の球面調和係数の部分を分解することもできる。これらの例および他の例において、分解ユニット５１８は、第２の複数の球面調和係数の部分の第２の分解を生成するために、第２の複数の球面調和係数の部分を分解することができる。単一のデバイスに関して説明されるが、２つ以上のデバイスが本開示で説明される技法を実行することができ、ここで、本開示で説明される技法に従って、２つのデバイスのうちの１つは分解を実行し、それらのデバイスのうちの別のものは補間を実行する。

【０９３９】

[0925]言い換えれば、球面調和ベースの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（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この技法は、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中のベクトルのいくつかを、背後にある音場の指向性成分として扱うことができる。しかしながら、この方式で扱われると、（Ｕ中の）これらのベクトルは、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の間で不連続である。これらの不連続性は、成分が変換オーディオコーダ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。

【０９４０】

[0926]本開示で説明される技法は、この不連続性に対処し得る。すなわち、本技法は、球面調和関数領域における直交する空間軸としてＶ行列が解釈され得るという考察に基づき得る。Ｕ行列は、それらの基底関数によって球面調和関数（ＨＯＡ）データの射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したがってそれら自体が不連続である基底関数（Ｖ）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の同様の分解とは異なる。これらの点で、ＳＶＤは、マッチング追跡アルゴリズムであると見なされ得る。本開示で説明される技法は、補間ユニット５５０が、フレームとフレームの間で基底関数（Ｖ）間の連続性を、それらの間で補間することによって保つことを可能にし得る。

【０９４１】

[0927]いくつかの例では、本技法は、補間ユニット５５０が、上で説明され、図４５および図４５Ｂに関して以下でさらに説明されるように、ＳＨデータのフレームを４つのサブフレームに分割することを可能にする。補間ユニット５５０は次いで、第１のサブフレームのためのＳＶＤを計算することができる。同様に、我々は、第２のフレームの第１のサブフレームのためのＳＶＤを計算する。第１のフレームおよび第２のフレームの各々について、補間ユニット５５０は、（Ｔ設計行列などの射影行列を使用して）ベクトルを球面上に射影することによって、Ｖ中のベクトルを空間マップに変換することができる。補間ユニット５５０は次いで、Ｖ中のベクトルを球面上の形状として解釈することができる。第１のフレームの第１のサブフレームと次のフレームの第１のサブフレームとの間の３つのサブフレームのためのＶ行列を補間するために、補間ユニット５５０は、次いで、これらの空間的な形状を補間し、および次いで、射影行列の逆行列を介して、それらをＳＨベクトルに戻すように変換することができる。本開示の技法は、このようにして、Ｖ行列間の平滑な遷移を提供することができる。

【０９４２】

[0928]このようにして、オーディオ符号化デバイス５１０Ｈは、次の項に関して以下で記載される技法の様々な態様を実行するように構成され得る。

【０９４３】

[0929]１３５０５４－１Ａ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、１つまたは複数のサブフレームのための分解された補間された球面調和係数を生成するために、第１のフレームに含まれる第１の複数の球面調和係数の一部分の第１の分解と、第２のフレームに含まれる第２の複数の球面調和係数の一部分の第２の分解とから、第１のフレームの１つまたは複数のサブフレームを補間するように構成される、１つまたは複数のプロセッサを備えるデバイス。

【０９４４】

[0930]１３５０５４－２Ａ項。第１の分解が、第１の複数の球面調和係数の部分の右特異ベクトルを表す第１のＶ行列を備える、１３５０５４－１Ａ項のデバイス。

【０９４５】

[0931]１３５０５４－３Ａ項。第２の分解が、第２の複数の球面調和係数の部分の右特異ベクトルを表す第２のＶ行列を備える、１３５０５４－１Ａ項のデバイス。

【０９４６】

[0932]１３５０５４－４Ａ項。第１の分解が、第１の複数の球面調和係数の部分の右特

異ベクトルを表す第１のＶ行列を備え、第２の分解が、第２の複数の球面調和係数の部分の右特異ベクトルを表す第２のＶ行列を備える、１３５０５４－１Ａ項のデバイス。

【０９４７】

[0933]１３５０５４－５Ａ項。１つまたは複数のプロセッサが、１つまたは複数のサブフレームを補間するとき、第１の分解および第２の分解に基づいて、１つまたは複数のサブフレームを時間的に補間するようにさらに構成される、１３５０５４－１Ａ項のデバイス。

【０９４８】

[0934]１３５０５４－６Ａ項。１つまたは複数のプロセッサが、１つまたは複数のサブフレームを補間するとき、第１の射影された分解を生成するために、第１の分解を空間領域に射影することと、第２の射影された分解を生成するために、第２の分解を空間領域に射影することと、第１の空間的に補間された射影された分解と第２の空間的に補間された射影された分解とを生成するために、第１の射影された分解と第２の射影された分解とを空間的に補間することと、第１の空間的に補間された射影された分解と第２の空間的に補間された射影された分解とに基づいて、１つまたは複数のサブフレームを時間的に補間することとを行うようにさらに構成される、１３５０５４－１Ａ項のデバイス。

【０９４９】

[0935]１３５０５４－７Ａ項。１つまたは複数のプロセッサが、１つまたは複数のサブフレームを補間することから生じる時間的に補間された球面調和係数を、球面調和関数領域に戻すように射影するようにさらに構成される、１３５０５４－６Ａ項のデバイス。

【０９５０】

[0936]１３５０５４－８Ａ項。第１の複数の球面調和係数の部分が、第１の複数の球面調和係数の単一のサブフレームを備える、１３５０５４－１Ａ項のデバイス。

【０９５１】

[0937]１３５０５４－９Ａ項。第２の複数の球面調和係数の部分が、第２の複数の球面調和係数の単一のサブフレームを備える、１３５０５４－１Ａ項のデバイス。

【０９５２】

[0938]１３５０５４－１０Ａ項。

【０９５３】

第１のフレームが４つのサブフレームに分割され、
第１の複数の球面調和係数の部分が、第１の複数の球面調和係数の第１のサブフレームのみを備える、１３５０５４－１Ａ項のデバイス。

【０９５４】

[0941]１３５０５４－１１Ａ項。

【０９５５】

第２のフレームが４つのサブフレームに分割され、
第２の複数の球面調和係数の部分が、第２の複数の球面調和係数の第１のサブフレームのみを備える、１３５０５４－１Ａ項のデバイス。

【０９５６】

[0944]１３５０５４－１２Ａ項。第１の複数の球面調和係数の部分が、第１の複数の球面調和係数の４つのサブフレームのうちの２つを備える、１３５０５４－１Ａ項のデバイス。

【０９５７】

[0945]１３５０５４－１３Ａ項。第２の複数の球面調和係数の部分が、第２の複数の球面調和係数の４つのサブフレームのうちの２つを備える、１３５０５４－１Ａ項のデバイス。

【０９５８】

[0946]１３５０５４－１４Ａ項。１つまたは複数のプロセッサが、第１の複数の球面調和係数の部分の第１の分解を生成するために、第１の複数の球面調和係数の部分を分解するようにさらに構成される、１３５０５４－１Ａ項のデバイス。

【０９５９】
[0947]１３５０５４－１５Ａ項。１つまたは複数のプロセッサが、第２の複数の球面調和係数の部分の第２の分解を生成するために、第２の複数の球面調和係数の部分を分解するようにさらに構成される、１３５０５４－１Ａ項のデバイス。
【０９６０】
[0948]１３５０５４－１６Ａ項。１つまたは複数のプロセッサが、第１の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第１の複数の球面調和係数の特異値を表すＳ行列と、第１の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第１の複数の球面調和係数の部分に関して特異値分解を実行するようにさらに構成される、１３５０５４－１Ａ項のデバイス。
【０９６１】
[0949]１３５０５４－１７Ａ項。１つまたは複数のプロセッサが、第２の複数の球面調和係数の左特異ベクトルを表すＵ行列と、第２の複数の球面調和係数の特異値を表すＳ行列と、第２の複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、第２の複数の球面調和係数の部分に関して特異値分解を実行するようにさらに構成される、１３５０５４－１Ａ項のデバイス。
【０９６２】
[0950]１３５０５４－１８Ａ項。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、音場の平面波表現を表す、１３５０５４－１Ａ項のデバイス。
【０９６３】
[0951]１３５０５４－１９Ａ項。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、一緒にミキシングされた１つまたは複数のモノラルオーディオオブジェクトを表す、１３５０５４－１Ａ項のデバイス。
【０９６４】
[0952]１３５０５４－２０Ａ項。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、３次元音場を表すそれぞれの第１の球面調和係数および第２の球面調和係数を備える、１３５０５４－１Ａ項のデバイス。
【０９６５】
[0953]１３５０５４－２１Ａ項。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、１よりも大きい次数を有する少なくとも１つの球面基底関数と関連付けられる、１３５０５４－１Ａ項のデバイス。
【０９６６】
[0954]１３５０５４－２２Ａ項。第１の複数の球面調和係数および第２の複数の球面調和係数が各々、４に等しい次数を有する少なくとも１つの球面基底関数と関連付けられる、１３５０５４－１Ａ項のデバイス。
【０９６７】
[0955]オーディオ符号化デバイス５１０Ｈによって実行されるものとして上で説明されたが、様々なオーディオ複号デバイス２４および５４０もまた、１３５０５４－１Ａ項から１３５０５４－２２Ａ項に関して上で記載された技法の様々な態様のいずれかを実行することができる。
【０９６８】
[0956]図４０Ｉは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｉを示すブロック図である。オーディオ符号化デバイス５１０Ｉは、オーディオ符号化デバイス５１０Ｉがオーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含む点において、オーディオ符号化デバイス５１０Ｈと類似であり得る。その上、オーディオ符号化デバイス５１０Ｉのオーディオ圧縮ユニット５１２は、オーディオ圧縮ユニット５１２が、オーディオ符号化デバイス５１０Ｈの同様のユニットに類似して動作し得る、分解ユニット５１８と音場成分抽出ユニット５２０とを含む点において、オーディオ符号化デバイス５１０Ｈのものと類

似であり得る。いくつかの例では、オーディオ符号化デバイス１０Ｉは、図３Ｄ～図３Ｅに関して説明されるように、U_{DIST}２５Ｃ、U_{BG}２５Ｄ、V^{T}_{DIST}２５Ｅ、およびV^{T}_{BG}２５Ｊのうちのいずれかの１つまたは複数のベクトルを量子化するために、量子化ユニット３４を含み得る。

【０９６９】

[0957]ただし、オーディオ符号化デバイス５１０Ｉのオーディオ圧縮ユニット５１２と、オーディオ符号化デバイス１０Ｈのオーディオ圧縮ユニット５１２の両方は、音場成分抽出ユニットを含むが、オーディオ符号化デバイス５１０Ｉの音場成分抽出ユニット５２０Ｉは、Ｖ圧縮ユニット５５２と呼ばれる追加のモジュールを含み得る。Ｖ圧縮ユニット５５２は、音場の空間成分、すなわちこの例ではV^{T}_{DIST}ベクトル５３９のうちの１つまたは複数を圧縮するように構成されるユニットを表し得る。すなわち、ＳＨＣに関して実行される特異値分解は、ＳＨＣ（音場を表す）を、Ｓ行列のベクトルによって表されるエネルギー成分と、Ｕ行列によって表される時間成分と、Ｖ行列によって表される空間成分とに分解することができる。Ｖ圧縮ユニット５５２は、量子化ユニット５２に関して上で説明されたものと同様の動作を実行することができる。

【０９７０】

[0958]例示の目的で、V^{T}_{DIST}ベクトル５３９は、各々２５個の要素を有する（これは音場の４次のＨＯＡ表現を示唆する）２つの行ベクトルを備えると仮定される。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V^{T}_{DIST}ベクトル５３９に含まれてよく、ただし、ｎは音場のＨＯＡ表現の次数を示す。

【０９７１】

[0959]Ｖ圧縮ユニット５５２は、V^{T}_{DIST}ベクトル５３９を受信し、圧縮されたV^{T}_{DIST}ベクトル表現５３９’を生成するために、圧縮スキームを実行することができる。この圧縮スキームは、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スキーム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。

【０９７２】

[0960]Ｖ圧縮ユニット５５２は、ある例として、V^{T}_{DIST}ベクトル５３９の各要素の浮動小数点表現をV^{T}_{DIST}ベクトル５３９の各要素の整数表現へと変換すること、V^{T}_{DIST}ベクトル５３９の整数表現の一様量子化、および、V^{T}_{DIST}ベクトル５３９の量子化された整数表現の分類とコーディングの１つまたは複数を含む、圧縮スキームを実行することができる。この圧縮スキームの１つまたは複数のプロセスの様々なものが、一例として、得られるビットストリーム５１７の目標ビットレートを達成するために、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。

【０９７３】

[0961]V^{T}_{DIST}ベクトル５３９の各々が互いに正規直交であるとすると、V^{T}_{DIST}ベクトル５３９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V^{T}_{DIST}ベクトル５３９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。

【０９７４】

[0962]いずれにしても、上で述べられたように、このコーディングスキームはまず、V^{T}_{DIST}ベクトル５３９の各々の各要素の浮動小数点表現（これは、いくつかの例では３２ビットの浮動小数点数である）を１６ビットの整数表現に変換することを伴い得る。Ｖ圧縮ユニット５５２は、V^{T}_{DIST}ベクトル５３９の所与の１つの各要素を2^{15}と乗算することによって、この浮動小数点から整数への変換を実行することができ、これはいくつかの例では、１５だけ右シフトすることによって実行される。

【０９７５】

[0963]Ｖ圧縮ユニット５５２は次いで、V^{T}_{DIST}ベクトル５３９の所与の１つの要素のすべてに関して一様量子化を実行することができる。Ｖ圧縮ユニット５５２は、ｎｂｉｔｓパラメータ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識別することができる。Ｖ圧縮ユニット５５２は、目標ビットレートに基づいて、このｎｂｉｔｓパラメー

タを動的に決定することができる。Ｖ圧縮ユニット５５２は、このｎｂｉｔｓパラメータの関数として、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、Ｖ圧縮ユニット５５２は、$2^{16-nbits}$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、ｎｂｉｔｓが６に等しい場合、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／Δ］に等しく、$-2^{nbits-1} < v_q < 2^{nbits-1}$である。

【０９７６】

[0964]Ｖ圧縮ユニット５５２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、Ｖ圧縮ユニット５５２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を（カテゴリ識別子ｃｉｄを決定することによって）、次の式を使用して識別することができる：

【０９７７】

【数３４】

$$cid = \begin{cases} 0, & v_q = 0 \text{ の場合} \\ \lfloor log_2|v_q| \rfloor + 1, & v_q \neq 0 \text{ の場合} \end{cases}$$

【０９７８】

Ｖ圧縮ユニット５５２は次いで、このカテゴリインデックスｃｉｄをハフマンコーディングし、一方で、v_qが正の値か負の値かを示す符号ビットを識別することもできる。Ｖ圧縮ユニット５５２は次に、このカテゴリにおける残差を識別することができる。一例として、Ｖ圧縮ユニット５５２は、次の式

$$残差 = |v_q| - 2^{cid-1}$$

に従って、この残差を決定することができる。

Ｖ圧縮ユニット５５２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。

【０９７９】

[0965]次の例は、この分類および残差コーディングプロセスの簡略化された例を示す。まず、ｎｂｉｔｓが６に等しいので$v_q \in$［－３１，３１］であると仮定する。次に、次のことを仮定する。

【０９８０】

【表１０】

cid	vq	cidのための ハフマンコード
0	0	‘1’
1	-1, 1	‘01’
2	-3,-2, 2,3	‘000’
3	-7,-6,-5,-4, 4,5,6,7	‘0010’
4	-15,-14,…,-8, 8,…,14,15	‘00110’
5	-31,-30,…,-16, 16,…,30,31	‘00111’

【０９８１】
また、次のことを仮定する。
【０９８２】
【表１１】

cid	残差のためのブロックコード
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110…,0000, 1000,…,1110,1111
5	01111, … ,00000, 10000, … ,11111

【０９８３】
したがって、v_q＝［６，－１７，０，０，３］に対して、次のことが決定され得る。

```
>>    cid = 3,5,0,0,2
```

>> 符号 = 1,0,x,x,1
>> 残差 = 2,1,x,x,1
>> ６に対するビット = ‘0010’ + ’1’ + ’10’
>> －１７に対するビット = ‘00111’ + ’0’ + ‘0001’
>> ０に対するビット = ‘0’
>> ０に対するビット = ‘0’
>> ３に対するビット = ‘000’ + ‘1’ + ‘1’
>> 総ビット = 7+10+1+1+5 = 24
>> 平均ビット = 24/5 = 4.8

【０９８４】

[0966]前述の簡略化された例には示されないが、Ｖ圧縮ユニット５５２は、ｃｉｄをコーディングするとき、ｎｂｉｔｓ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、Ｖ圧縮ユニット５５２は、ｎｂｉｔｓ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、Ｖ圧縮ユニット５５２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ｎｂｉｔｓ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、Ｖ圧縮ユニット５５２は、いくつかの異なる統計的な状況においてｃｉｄ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。

【０９８５】

[0967]説明するために、Ｖ圧縮ユニット５５２は、ｎｂｉｔｓ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V^{T}_{DIST}ベクトル５３９の１つがV^{T}_{DIST}ベクトル５３９の対応する時間的に後続の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は定義されたもの）の空間情報を表さないときに、使用され得る。Ｖ圧縮ユニット５５２は追加で、ｎｂｉｔｓ値の各々に対して、V^{T}_{DIST}ベクトル５３９の１つがV^{T}_{DIST}ベクトル５３９の対応する時間的に後続するベクトルから予測されるとき、V^{T}_{DIST}ベクトル５３９のその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。Ｖ圧縮ユニット５５２はまた、ｎｂｉｔｓ値の各々に対して、V^{T}_{DIST}ベクトル５３９の１つが合成オーディオオブジェクトを表すとき、V^{T}_{DIST}ベクトル５３９のその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、構築され得る。

【０９８６】

[0968]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。

【０９８７】

【表１２】

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

【０９８８】
前述の表において、予測モード（「Ｐｒｅｄモード」）は、現在のベクトルに対して予測が実行されたかどうかを示し、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。
【０９８９】
[0969]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。
【０９９０】
【表１３】

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

【０９９１】
前述の表において、「録音」列は、録音されたオーディオオブジェクトをベクトルが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のコーディング状況を示す。この表に示されるように、Ｖ圧縮ユニット５５２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。Ｖ

圧縮ユニット５５２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。Ｖ圧縮ユニット５５２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。Ｖ圧縮ユニット５５２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。

【０９９２】

[0970]このようにして、本技法は、オーディオ圧縮デバイスが、音場の空間成分を圧縮することを可能にすることができ、ここで、空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【０９９３】

[0971]図４３は、図４０Ｉに示されるＶ圧縮ユニット５５２をより詳細に示す図である。図４３の例では、Ｖ圧縮ユニット５５２は、一様量子化ユニット６００と、ｎｂｉｔｓユニット６０２と、予測ユニット６０４と、予測モードユニット６０６（「Ｐｒｅｄモードユニット６０６」）と、カテゴリおよび残差コーディングユニット６０８と、ハフマンテーブル選択ユニット６１０とを含む。一様量子化ユニット６００は、図４３の例においてｖとして示される空間成分のうちの１つ（V^T_{DIST}ベクトル５３９のうちの任意の１つを表し得る）に関して上で説明された一様量子化を実行するように構成されるユニットを表す。ｎｂｉｔｓユニット６０２は、ｎｂｉｔｓパラメータまたは値を決定するように構成されるユニットを表す。

【０９９４】

[0972]予測ユニット６０４は、図４３の例においてv_qとして示される量子化された空間成分に関して予測を実行するように構成されるユニットを表す。予測ユニット６０４は、V^T_{DIST}ベクトル５３９の対応する時間的に後続するベクトルによる、V^T_{DIST}ベクトル５３９の現在のベクトルの要素ごとの減算を実行することによって、予測を実行することができる。この予測の結果は、予測された空間成分と呼ばれ得る。

【０９９５】

[0973]予測モードユニット６０６は、予測モードを選択するように構成されるユニットを表し得る。ハフマンテーブル選択ユニット６１０は、ｃｉｄのコーディングのための適切なハフマンテーブルを選択するように構成されるユニットを表し得る。予測モードユニット６０６およびハフマンテーブル選択ユニット６１０は、一例として、次の擬似コードに従って動作することができる。

```
For a given nbits, retrieve all the Huffman Tables having nbits
B00 = 0; B01 = 0; B10 = 0; B11 = 0; // コーディングモードごとに期待されるビットを計算するために初期化する
for m = 1:(# elements in the vector)
  // ベクトル要素ｖ（ｍ）のための期待されるビットの数を計算する
  // 予測なしでハフマンテーブル５を使用する
  B00 = B00 + calculate_bits(v(m), HT5);
  // 予測なしでハフマンテーブル｛１，２，３｝を使用する
  B01 = B01 + calculate_bits(v(m), HTq); q in {1,2,3}

  // 予測残差ｅ（ｍ）のための期待されるビットの数を計算する
  e(m) = v(m) - vp(m); // vp(m): previous frame vector element
  // 予測ありでハフマンテーブル４を使用する
  B10 = B10 + calculate_bits(e(m), HT4);
  // 予測ありでハフマンテーブル５を使用する
  B11 = B11 + calculate_bits(e(m), HT5);
end
```

```
    // 最小ビットを生じる最良の予測モードとハフマンテーブルとを発見する
    // 最良の予測モードおよびハフマンテーブルに、それぞれｐｆｌａｇおよびＨｔ
ｆｌａｇによってフラグが付けられる
    [Be, id] = min( [B00 B01 B10 B11] );
    Switch id
      case 1: pflag = 0; HTflag = 0;
      case 2: pflag = 0; HTflag = 1;
      case 3: pflag = 1; HTflag = 0;
      case 4: pflag = 1; HTflag = 1;
end
```

【０９９６】

[0974]カテゴリおよび残差コーディングユニット６０８は、上でより詳細に説明された方式で、予測された空間成分または量子化された空間成分（予測が無効化されるとき）の分類と残差コーディングとを実行するように構成されるユニットを表し得る。

【０９９７】

[0975]図４３の例に示されるように、Ｖ圧縮ユニット５５２は、ビットストリーム５１７またはサイド情報（それ自体がビットストリーム５１７とは別個のビットストリームであり得る）のいずれかに含めるために様々なパラメータまたは値を出力することができる。情報がビットストリーム５１７において指定されると仮定すると、Ｖ圧縮ユニット５５２は、この例では、ｃｉｄを符号化するために選択されるハフマンコードと、符号ビットと、ブロックコーディングされた残差とを指し得る、空間成分の圧縮されたバージョン（図４０Ｉの例では、圧縮された空間成分５３９’として示される）とともに、ｎｂｉｔｓ値と、予測モードと、ハフマンテーブル情報とを、ビットストリーム生成ユニット５１６に出力することができる。ｎｂｉｔｓ値は、V^{T}_{DIST}ベクトル５３９のすべてについて、ビットストリーム５１７において一度指定され得るが、予測モードおよびハフマンテーブル情報は、V^{T}_{DIST}ベクトル５３９の各１つずつについて指定され得る。空間成分の圧縮されたバージョンを指定するビットストリームの部分は、図１０Ｂおよび図１０Ｃの例に示されている。

【０９９８】

[0976]このようにして、オーディオ符号化デバイス５１０Ｈは、次の項に関して以下で記載される技法の様々な態様を実行することができる。

【０９９９】

[0977]１４１５４１－１Ａ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、音場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を取得するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。

【１０００】

[0978]１４１５４１－２Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用される予測モードを指定するフィールドを使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項のデバイス。

【１００１】

[0979]１４１５４１－３Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項および１４１５４１－２Ａ項の任意の組合せのデバイス。

【１００２】

[0980]１４１５４１－４Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィールドを使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項から１４１５４１－３Ａ項の任意の組合せのデバイス。

【１００３】
[0981]１４１５４１－５Ａ項。値がｎｂｉｔｓ値を備える、１４１５４１－４Ａ項のデバイス。
【１００４】
[0982]１４１５４１－６Ａ項。ビットストリームが、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを備え、値が、複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す、１４１５４１－４Ａ項および１４１５４１－５Ａ項の任意の組合せのデバイス。
【１００５】
[0983]１４１５４１－７Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためのハフマンコードを使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項から１４１５４１－６Ａ項の任意の組合せのデバイス。
【１００６】
[0984]１４１５４１－８Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分が正の値か負の値かを識別する符号ビットを使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項から１４１５４１－７Ａ項の任意の組合せのデバイス。
【１００７】
[0985]１４１５４１－９Ａ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分の残差値を表すためのハフマンコードを使用して、ビットストリームにおいて表される、１４１５４１－１Ａ項から１４１５４１－８Ａ項の任意の組合せのデバイス。
【１００８】
[0986]１４１５４１－１０Ａ項。デバイスが、オーディオ符号化デバイス、ビットストリーム生成デバイスを備える、１４１５４１－１Ａ項から１４１５４１－９Ａ項の任意の組合せのデバイス。
【１００９】
[0987]１４１５４１－１２Ａ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ａ項から１４１５４１－１１Ａ項の任意の組合せのデバイス。
【１０１０】
[0988]オーディオ符号化デバイス５１０Ｈによって実行されるとして説明されるが、本技法はまた、オーディオ複号デバイス２４および／または５４０のいずれかによっても実行され得る。
【１０１１】
[0989]このようにして、オーディオ符号化デバイス５１０Ｈは、追加として、次の項に関して以下で記載される技法の様々な態様を実行することができる。
【１０１２】
[0990]１４１５４１－１Ｄ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、音場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を生成するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。
【１０１３】
[0991]１４１５４１－２Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分を圧縮するときに使用される予測モードを指定するフィールドを含むように、ビットストリームを生成するようにさらに構成される、１４１５４１－１Ｄ項のデバイス。
【１０１４】
[0992]１４１５４１－３Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を含むように、ビットストリームを生成するようにさらに構成される、１４

１５４１－１Ｄ項および１４１５４１－２Ｄ項の任意の組合せのデバイス。
【１０１５】
[0993]１４１５４１－４Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を含むフィールドを含むように、ビットストリームを生成するようにさらに構成される、１４１５４１－１Ｄ項から１４１５４１－３Ｄ項の任意の組合せのデバイス。
【１０１６】
[0994]１４１５４１－５Ｄ項。値がｎｂｉｔｓ値を備える、１４１５４１－４Ｄ項のデバイス。
【１０１７】
[0995]１４１５４１－６Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを含むように、ビットストリームを生成するようにさらに構成され、値が、複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す、１４１５４１－４Ｄ項および１４１５４１－５Ｄ項の任意の組合せのデバイス。
【１０１８】
[0996]１４１５４１－７Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためのハフマンコードを含むように、ビットストリームを生成するようにさらに構成される、１４１５４１－１Ｄ項から１４１５４１－６Ｄ項の任意の組合せのデバイス。
【１０１９】
[0997]１４１５４１－８Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分が正の値か負の値かを識別する符号ビットを含むように、ビットストリームを生成するようにさらに構成される、１４１５４１－１Ｄ項から１４１５４１－７Ｄ項の任意の組合せのデバイス。
【１０２０】
[0998]１４１５４１－９Ｄ項。１つまたは複数のプロセッサが、ビットストリームを生成するとき、空間成分の残差値を表すためのハフマンコードを含むように、ビットストリームを生成するようにさらに構成される、１４１５４１－１Ｄ項から１４１５４１－８Ｄ項の任意の組合せのデバイス。
【１０２１】
[0999]１４１５４１－１０Ｄ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ｄ項から１４１５４１－１０Ｄ項の任意の組合せのデバイス。
【１０２２】
[1000]オーディオ符号化デバイス５１０Ｈは、次の項に記載されるような技法の様々な態様を実施するようにさらに構成され得る。
【１０２３】
[1001]１４１５４１－１Ｅ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、音場の空間成分を圧縮するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。
【１０２４】
[1002]１４１５４１－２Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、浮動小数点表現から整数表現へと空間成分を変換するようにさらに構成される、１４１５４１－１Ｅ項のデバイス。
【１０２５】
[1003]１４１５４１－３Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、量子化ステップサイズを示す値を動的に決定することと、量子化された空間成分を生成するために、その値に基づいて空間成分を量子化することとを行うようにさらに構成される、１４１５４１－１Ｅ項および１４１５４１－２Ｅ項の任意の組合せのデバイス。

【１０２６】
[1004]１４１５４１－４Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、空間成分が対応するカテゴリを識別するようにさらに構成される、請求項１Ｅ～３Ｅの任意の組合せのデバイス。
【１０２７】
[1005]１４１５４１－５Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、空間成分のための残差値を識別するようにさらに構成される、１４１５４１－１Ｅ項から１４１５４１－４Ｅ項の任意の組合せのデバイス。
【１０２８】
[1006]１４１５４１－６Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、予測された空間成分を生成するために、空間成分および後続の空間成分に関して予測を実行するようにさらに構成される、１４１５４１－１Ｅ項から１４１５４１－５Ｅ項の任意の組合せのデバイス。
【１０２９】
[1007]１４１５４１－７Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、浮動小数点表現から整数表現へと空間成分を変換することと、量子化ステップサイズを示す値を動的に決定することと、量子化された空間成分を生成するために、その値に基づいて空間成分の整数表現を量子化することと、カテゴリ識別子を生成するために、量子化された空間成分に基づいて、空間成分が対応するカテゴリを識別することと、空間成分の符号を決定することと、量子化された空間成分およびカテゴリ識別子に基づいて、空間成分のための残差値を識別することと、カテゴリ識別子、符号および残差値に基づいて、空間成分の圧縮されたバージョンを生成することとを行うようにさらに構成される、１４１５４１－１Ｅ項の任意の組合せのデバイス。
【１０３０】
[1008]１４１５４１－８Ｅ項。１つまたは複数のプロセッサが、空間成分を圧縮するとき、浮動小数点表現から整数表現へと空間成分を変換することと、量子化ステップサイズを示す値を動的に決定することと、量子化された空間成分を生成するために、その値に基づいて空間成分の整数表現を量子化することと、予測された空間成分を生成するために、空間成分および後続の空間成分に関して予測を実行することと、カテゴリ識別子を生成するために、量子化された空間成分に基づいて、予測された空間成分が対応するカテゴリを識別することと、空間成分の符号を決定することと、量子化された空間成分およびカテゴリ識別子に基づいて、空間成分のための残差値を識別することと、カテゴリ識別子、符号および残差値に基づいて、空間成分の圧縮されたバージョンを生成することとを行うようにさらに構成される、１４１５４１－１Ｅ項の任意の組合せのデバイス。
【１０３１】
[1009]１４１５４１－９Ｅ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ｅ項から１４１５４１－８Ｅ項の任意の組合せのデバイス。
【１０３２】
[1010]本技法の様々な態様は、オーディオ符号化デバイス５１０Ｈが、次の項に記載されるように動作するように構成されることをさらに可能にすることができる。
【１０３３】
[1011]１４１５４１－１Ｆ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、複数の空間成分のうちの現在の空間成分を複数の空間成分の残りに対する現在の空間成分の次数に基づいて圧縮するときに使用するハフマンコードブックを識別するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。
【１０３４】
[1012]１４１５４１－２Ｆ項。１つまたは複数のプロセッサが、１４１５４１－１Ａ項から１４１５４１－１２Ａ項、１４１５４１－１Ｂ項から１４１５４１－１０Ｂ項、および１４１５４１－１Ｃ項から１４１５４１－９Ｃ項に記載されたステップの任意の組合せ

を実行するようにさらに構成される、１４１５４１－３Ｆ項のデバイス。
【１０３５】
[1013]本技法の様々な態様は、オーディオ符号化デバイス５１０Ｈが、次の項に記載されるように動作するように構成されることをさらに可能にすることができる。
【１０３６】
[1014]１４１５４１－１Ｈ項。オーディオ符号化デバイス５１０Ｈなどのデバイスであって、音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。
【１０３７】
[1015]１４１５４１－２Ｈ項。１つまたは複数のプロセッサが、量子化ステップサイズを決定するとき、目標ビットレートに基づいて量子化ステップサイズを決定するようにさらに構成される、１４１５４１－１Ｈ項のデバイス。
【１０３８】
[1016]１４１５４１－３Ｈ項。１つまたは複数のプロセッサが、複数の量子化ステップサイズのうちの１つを選択するとき、空間成分を表すために使用されるビットの数の推定値を決定することと、推定値と目標ビットレートとの間の差に基づいて量子化ステップサイズを決定することとを行うようにさらに構成される、１４１５４１－１Ｈ項のデバイス。
【１０３９】
[1017]１４１５４１－４Ｈ項。１つまたは複数のプロセッサが、複数の量子化ステップサイズのうちの１つを選択するとき、空間成分を表すために使用されるビットの数の推定値を決定することと、推定値と目標ビットレートとの間の差を決定することと、その差を目標ビットレートに加算することによって、量子化ステップサイズを決定することとを行うようにさらに構成される、１４１５４１－１Ｈ項のデバイス。
【１０４０】
[1018]１４１５４１－５Ｈ項。１つまたは複数のプロセッサが、ビットの数の推定値を決定するとき、目標ビットレートに対応するコードブックを仮定して、空間成分に対して生成されるべきビットの数のエスティメイテッドを計算するようにさらに構成される、１４１５４１－３Ｈ項または１４１５４１－４Ｈ項のデバイス。
【１０４１】
[1019]１４１５４１－６Ｈ項。１つまたは複数のプロセッサが、ビットの数の推定値を決定するとき、空間成分を圧縮するときに使用されるコーディングモードを仮定して、空間成分に対して生成されるべきビットの数のエスティメイテッドを計算するようにさらに構成される、１４１５４１－３Ｈ項または１４１５４１－４Ｈ項のデバイス。
【１０４２】
[1020]１４１５４１－７Ｈ項。１つまたは複数のプロセッサが、ビットの数の推定値を決定するとき、空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して、空間成分に対して生成されるべきビットの数の第１の推定値を計算することと、空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して、空間成分に対して生成されるべきビットの数の第２の推定値を計算することと、ビットの数の決定された推定値として使用されることになる、第１の推定値と第２の推定値のうちでビットの数が少ない方を選択することとを行うようにさらに構成される、１４１５４１－３Ｈ項または１４１５４１－４Ｈ項のデバイス。
【１０４３】
[1021]１４１５４１－８Ｈ項。１つまたは複数のプロセッサが、ビットの数の推定値を決定するとき、空間成分が対応するカテゴリを識別するカテゴリ識別子を識別することと、カテゴリに対応する空間成分を圧縮するときに生じるであろう空間成分の残差値のビット長を識別することと、カテゴリ識別子を表すために使用されるビットの数を残差値のビット長に加算することによって少なくとも一部、ビットの数の推定値を決定することとを

行うようにさらに構成される、１４１５４１－３Ｈ項または１４１５４１－４Ｈ項のデバイス。

【１０４４】

[1022]１４１５４１－９Ｈ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ｈ項から１４１５４１－８Ｈ項の任意の組合せのデバイス。

【１０４５】

[1023]オーディオ符号化デバイス５１０Ｈによって実行されるとして説明されるが、上記の項、１４１５４１－１Ｈ項から１４１５４１－９Ｈ項に記載された本技法はまた、オーディオ複号デバイス５４０Ｄによっても実行され得る。

【１０４６】

[1024]加えて、本技法の様々な態様は、オーディオ符号化デバイス５１０Ｈが、次の項に記載されるように動作するように構成されることを可能にすることができる。

【１０４７】

[1025]１４１５４１－１Ｊ項。オーディオ符号化デバイス５１０Ｊなどのデバイスであって、音場の空間成分を圧縮するときに使用されるべき複数のコードブックのうちの１つを選択するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。

【１０４８】

[1026]１４１５４１－２Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、複数のコードブックの各々を使用して空間成分を表すために使用されるビットの数の推定値を決定することと、決定された推定値をもたらした複数のコードブックのうちでビットの数が最小のものを選択することとを行うようにさらに構成される、１４１５４１－１Ｊ項のデバイス。

【１０４９】

[1027]１４１５４１－３Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、複数のコードブックの１つまたは複数を使用して空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成され、複数のコードブックの１つまたは複数が、空間成分の他の要素に対する、圧縮されるべき空間成分の要素の次数に基づいて選択される、１４１５４１－１Ｊ項のデバイス。

【１０５０】

[1028]１４１５４１－４Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、空間成分が後続の空間成分から予測されないときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成される、１４１５４１－１Ｊ項のデバイス。

【１０５１】

[1029]１４１５４１－５Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、空間成分が後続の空間成分から予測されるときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成される、１４１５４１－１Ｊ項のデバイス。

【１０５２】

[1030]１４１５４１－６Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、空間成分が音場の中の合成オーディオオブジェクトを表すときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成される、１４１５４１－１Ｊ項のデバイス。

【１０５３】

[1031]１４１５４１－７Ｊ項。合成オーディオオブジェクトが、パルス符号変調（ＰＣ

M）されたオーディオオブジェクトを備える、１４１５４１－１Ｊ項のデバイス。

【１０５４】

[1032]１４１５４１－８Ｊ項。１つまたは複数のプロセッサが、複数のコードブックのうちの１つを選択するとき、空間成分が音場の中の録音されたオーディオオブジェクトを表すときに使用されるように設計される複数のコードブックの１つを使用して、空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成される、１４１５４１－１Ｊ項のデバイス。

【１０５５】

[1033]１４１５４１－９Ｊ項。ベクトルベースの合成が特異値分解を備える、請求項１Ｊ～８Ｊの任意の組合せのデバイス。

【１０５６】

[1034]上で説明された様々な例の各々において、オーディオ符号化デバイス５１０は、ある方法を実行し、またはそうでなければ、オーディオ符号化デバイス５１０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。いくつかの例では、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。いくつかの例では、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５１０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。

【１０５７】

[1035]図４０Ｊは、２次元または３次元の音場を記述する球面調和係数を圧縮するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ符号化デバイス５１０Ｊを示すブロック図である。オーディオ符号化デバイス５１０Ｊは、オーディオ符号化デバイス５１０Ｊがオーディオ圧縮ユニット５１２と、オーディオ符号化ユニット５１４と、ビットストリーム生成ユニット５１６とを含む点において、オーディオ符号化デバイス５１０Ｇと類似であり得る。その上、オーディオ符号化デバイス５１０Ｊのオーディオ圧縮ユニット５１２は、オーディオ圧縮ユニット５１２が、オーディオ符号化デバイス５１０Ｉの同様のユニットに類似して動作し得る、分解ユニット５１８と音場成分抽出ユニット５２０とを含む点において、オーディオ符号化デバイス５１０Ｇのものと類似であり得る。いくつかの例では、オーディオ符号化デバイス５１０Ｊは、図４０Ｄ～図４０Ｅに関して説明されるように、U_{DIST}ベクトル５２５Ｃ、U_{BG}ベクトル５２５Ｄ、V^{T}_{DIST}ベクトル５２５Ｅ、およびV^{T}_{BG}ベクトル５２５Ｊのうちのいずれかの１つまたは複数のベクトルを量子化するために、量子化ユニット５３４を含み得る。

【１０５８】

[1036]ただし、オーディオ符号化デバイス５１０Ｊのオーディオ圧縮ユニット５１２は、オーディオ符号化デバイス５１０Ｊのオーディオ圧縮ユニット５１２が補間ユニット５５０として示される追加のユニットを含む点において、オーディオ符号化デバイス５１０Ｇのオーディオ圧縮ユニット５１２とは異なり得る。補間ユニット５５０は、図４５および図４５Ｂに関して以下でより詳細に説明されるように、第１のオーディオフレームおよび第２の時間的に後続のまたは先行するオーディオフレームのサブフレームから、第１のオーディオフレームのサブフレームを補間するユニットを表し得る。補間ユニット５５０は、この補間を実行することにおいて、分解ユニット５１８がＳＨＣ５１１を分解するために必要とされる程度を潜在的に低減することによって、（処理サイクルおよび／またはメモリ消費に関して）計算的な複雑さを低減することができる。補間ユニット５５０は、図４０Ｈおよび図４０Ｉの例に示されるオーディオ符号化デバイス５１０Ｈおよび５１０Ｉの補間ユニット５５０に関して上で説明されたものと類似した方式で動作することができる。

【１０５９】

[1037]演算において、補間ユニット２００は、１つまたは複数のサブフレームのための

分解された補間された球面調和係数を生成するために、第１のフレームに含まれる第１の複数の球面調和係数１１の一部分の第１の分解、たとえばＶ行列１９’と、第２のフレームに含まれる第２の複数の球面調和係数１１の一部分の第２の分解、たとえばＶ行列１９’とから、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。

【１０６０】

[1038]補間ユニット５５０は、少なくとも一部、第１の複数の球面調和係数の第１の分解および第２の複数の球面調和係数の第２の分解に関して補間を実行することによって、時間セグメントのための分解された補間された球面調和係数を取得することができる。平滑化ユニット５５４は、第１の複数の球面調和係数および第２の複数の球面調和係数の空間成分と時間成分のうちの少なくとも１つを平滑化するために、分解された補間された球面調和係数を適用することができる。平滑化ユニット５５４は、図３７～図３９に関して上で説明されたように、平滑化されたU_{DIST}行列５２５Ｃ’を生成することができる。第１の分解および第２の分解は、図４０ＪのＶ$_1^T$５５６、Ｖ$_2^T$５５６Ｂを指し得る。

【１０６１】

[1039]いくつかの場合には、V^Tまたは他のＶベクトルまたはＶ行列は、補間のために量子化されたバージョンで出力され得る。このようにして、補間のためのＶベクトルは、復号器におけるＶベクトルと同一であってよく、復号器もまた、たとえば、多次元信号を復元するために、Ｖベクトル補間を実行する。

【１０６２】

[1040]いくつかの例では、第１の分解は、第１の複数の球面調和係数５１１の部分の右特異ベクトルを表す第１のＶ行列５１９’を備える。同様に、いくつかの例では、第２の分解は、第２の複数の球面調和係数の部分の右特異ベクトルを表す第２のＶ行列５１９’を備える。

【１０６３】

[1041]補間ユニット５５０は、第１のＶ行列５１９’および第２のＶ行列１９’に基づいて、１つまたは複数のサブフレームに関して時間的補間を実行することができる。すなわち、補間ユニット５５０は、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’と、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’とに基づいて、たとえば、第１のオーディオフレームのための４つの全サブフレームからの第２のサブフレームと、第３のサブフレームと、第４のサブフレームとを時間的に補間することができる。いくつかの例では、この時間的補間は、線形時間的補間であり、ここで、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、第１のオーディオフレームの第４のサブフレームを補間するときよりも、第１のオーディオフレームの第２のサブフレームを補間するときにより重く重み付けられる。第３のサブフレームを補間するとき、Ｖ行列５１９’は均等に重み付けられ得る。第４のサブフレームを補間するとき、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’よりも重く重み付けられ得る。

【１０６４】

[1042]言い換えれば、線形時間的補間は、補間されるべき第１のオーディオフレームのサブフレームのうちの１つの近接を仮定して、Ｖ行列５１９’に重み付けることができる。補間されるべき第２のサブフレームについて、第１のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’は、補間されるべき第２のサブフレームへのその近接を仮定して、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’よりも重く重み付けられる。重みは、Ｖ行列５１９’に基づいて第３のサブフレームを補間するとき、この理由のために等価であり得る。補間されるべき第４のサブフレームが、第１のオーディオフレームの第１のサブフレームよりも第２のオーディオフレームの第１のサブフレームにより近接しているとすると、第２のオーディオフレームの第１のサブフレームから分解されたＶ行列５１９’に適用される重みは、第１のオーディオフ

レームの第１のサブフレームから分解されたＶ行列５１９’に適用されるものよりも大きくなり得る。

【１０６５】

[1043]いくつかの例では、補間ユニット５５０は、第１の射影された分解を生成するために、第１のオーディオフレームの第１のサブフレームから分解された第１のＶ行列５１９’を、空間領域に射影することができる。いくつかの例では、この射影は、（たとえば、Ｔ設計行列などの射影行列を使用する）球面への射影を含む。補間ユニット５５０は次いで、第２の射影された分解を生成するために、第２のオーディオフレームの第１のサブフレームから分解された第２のＶ行列５１９’を、空間領域に射影することができる。補間ユニット５５０は次いで、第１の空間的に補間された射影された分解と第２の空間的に補間された射影された分解とを生成するために、第１の射影された分解と第２の射影された分解とを空間的に補間する（この場合も線形補間であり得る）ことができる。補間ユニット５５０は次いで、第１の空間的に補間された射影された分解と第２の空間的に補間された射影された分解とに基づいて、１つまたは複数のサブフレームを時間的に補間することができる。

【１０６６】

[1044]補間ユニット５５０がＶ行列５１９’を空間的に、次いで時間的に射影するそれらの例では、補間ユニット５５０は、１つまたは複数のサブフレームを補間することから生じる時間的に補間された球面調和係数を、球面調和関数領域に戻すように射影し、それによって、Ｖ行列５１９と、Ｓ行列５１９Ｂと、Ｕ行列５１９Ｃとを生成することができる。

【１０６７】

[1045]いくつかの例では、第１の複数の球面調和係数の部分は、第１の複数の球面調和係数５１１の単一のサブフレームを備える。いくつかの例では、第２の複数の球面調和係数の部分は、第２の複数の球面調和係数５１１の単一のサブフレームを備える。いくつかの例では、Ｖ行列１９’がそれから分解されるこの単一のサブフレームは、第１のサブフレームである。

【１０６８】

[1046]いくつかの例では、第１のフレームは４つのサブフレームに分割される。これらの例および他の例において、第１の複数の球面調和係数の部分は、複数の球面調和係数５１１の第１のサブフレームのみを備える。これらの例および他の例において、第２のフレームは４つのサブフレームに分割され、第２の複数の球面調和係数５１１の部分は、第２の複数の球面調和係数５１１の第１のサブフレームのみを備える。

【１０６９】

[1047]いくつかの例では、各オーディオフレームの第１のサブフレームのみが、補間を実行するために使用されるが、第１の複数の球面調和係数の部分は、第１の複数の球面調和係数５１１の４つのサブフレームのうちの２つを備え得る。これらの例および他の例において、第２の複数の球面調和係数５１１の部分は、第２の複数の球面調和係数５１１の４つのサブフレームのうちの２つを備える。

【１０７０】

[1048]上で述べられたように、単一のデバイス、たとえばオーディオ符号化デバイス５１０Ｊは、補間を実行し、一方でまた、第１の複数の球面調和係数の部分の第１の分解を生成するために、第１の複数の球面調和係数の部分を分解することもできる。これらの例および他の例において、分解ユニット５１８は、第２の複数の球面調和係数の部分の第２の分解を生成するために、第２の複数の球面調和係数の部分を分解することができる。単一のデバイスに関して説明されるが、２つ以上のデバイスが本開示で説明される技法を実行することができ、ここで、本開示で説明される技法に従って、２つのデバイスのうちの１つは分解を実行し、それらのデバイスのうちの別のものは補間を実行する。

【１０７１】

[1049]いくつかの例では、分解ユニット５１８は、第１の複数の球面調和係数５１１の

右特異ベクトルを表すＶ行列５１９’（ならびに、説明を簡単にするために図示されない、Ｓ行列５１９Ｂ’およびＵ行列５１９Ｃ’）を生成するために、第１の複数の球面調和係数５１１の部分に関して特異値分解を実行することができる。これらの例および他の例において、分解ユニット５１８は、第２の複数の球面調和係数の右特異ベクトルを表すＶ行列５１９’（ならびに、説明を簡単にするために図示されない、Ｓ行列５１９Ｂ’およびＵ行列５１９Ｃ’）を生成するために、第２の複数の球面調和係数５１１の部分に関して特異値分解を実行することができる。

【１０７２】

[1050]いくつかの例では、上で述べられたように、第１の複数の球面調和係数と第２の複数の球面調和係数は各々、音場の平面波表現を表す。これらの例および他の例において、第１の複数の球面調和係数５１１と第２の複数の球面調和係数５１１は各々、一緒にミキシングされた１つまたは複数のモノラルオーディオオブジェクトを表す。

【１０７３】

[1051]言い換えれば、球面調和ベースの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（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この技法は、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中のベクトルのいくつかを、背後にある音場の指向性成分として扱うことができる。しかしながら、この方式で扱われると、（Ｕ中の）これらのベクトルは、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の間で不連続である。これらの不連続性は、成分が変換オーディオコーダ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。

【１０７４】

[1052]本開示で説明される技法は、この不連続性に対処し得る。すなわち、本技法は、球面調和関数領域における直交する空間軸としてＶ行列が解釈され得るという考察に基づき得る。Ｕ行列は、それらの基底関数によって球面調和関数（ＨＯＡ）データの射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したがってそれら自体が不連続である基底関数（Ｖ）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の同様の分解とは異なる。これらの点で、ＳＶＤは、マッチング追跡アルゴリズムであると見なされ得る。本開示で説明される技法は、補間ユニット５５０が、フレームとフレームの間で基底関数（Ｖ）間の連続性を、それらの間で補間することによって保つことを可能にし得る。

【１０７５】

[1053]いくつかの例では、本技法は、補間ユニット５５０が、上で説明され、図４５および図４５Ｂに関して以下でさらに説明されるように、ＳＨデータのフレームを４つのサブフレームに分割することを可能にする。補間ユニット５５０は次いで、第１のサブフレームのためのＳＶＤを計算することができる。同様に、我々は、第２のフレームの第１のサブフレームのためのＳＶＤを計算する。第１のフレームおよび第２のフレームの各々について、補間ユニット５５０は、（Ｔ設計行列などの射影行列を使用して）ベクトルを球面上に射影することによって、Ｖ中のベクトルを空間マップに変換することができる。補間ユニット５５０は次いで、Ｖ中のベクトルを球面上の形状として解釈することができる。第１のフレームの第１のサブフレームと次のフレームの第１のサブフレームとの間の３つのサブフレームのためのＶ行列を補間するために、補間ユニット５５０は、次いで、これらの空間的な形状を補間し、および次いで、射影行列の逆行列を介して、それらをＳＨベクトルに戻すように変換することができる。本開示の技法は、このようにして、Ｖ行列間の平滑な遷移を提供することができる。

【１０７６】

[1054]図４１～図４１Ｄは各々、２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法の様々な態様を実行し得る例示的なオーディオ復号デバイス５４０Ａ～５４０Ｄを示すブロック図である。オーディオ復号デバイス５４０Ａは、デスクトップコンピュータ、ラップトップコンピュータ、ワークステーション、タブレットもしくはスレートコンピュータ、専用オーディオ録音デバイス、携帯電話（いわゆる「スマートフォン」を含む）、パーソナルメディアプレーヤデバイス、パーソナルゲームデバイス、または、オーディオデータを復号することが可能な任意の他のタイプのデバイスなど、オーディオデータを復号することが可能な任意のデバイスを表し得る。

【１０７７】

[1055]いくつかの例では、オーディオ復号デバイス５４０Ａは、いくつかの例では、外部の無関係なデータの除去を容易にするためにオーディオ符号化デバイス５１０Ｂ～５１０Ｊによって使用される、（図４０Ｂ～図４０Ｊの例に関して上で説明されたような）次数低減を実行することを除いて、オーディオ符号化デバイス５１０または５１０Ｂのいずれかによって実行されるオーディオ符号化プロセスとは逆であるオーディオ復号プロセスを実行する。

【１０７８】

[1056]図４１の例では、単一のデバイス、すなわち、デバイス５４０Ａとして示されるが、デバイス５４０Ａ内に含まれるものとして以下で言及される様々な構成要素またはユニットは、デバイス５４０から外部である別個のデバイスを形成することができる。言い換えれば、図４１の例では、単一のデバイス、すなわち、デバイス５４０Ａによって実行されるものとして本開示で説明されるが、本技法は、複数のデバイスを備えるシステムによって実施され、またはさもなければ実行されてよく、ここで、これらのデバイスの各々は、以下でより詳細に説明される様々な構成要素またはユニットのうちの１つまたは複数をそれぞれ含み得る。したがって、本技法は、この点で、図４１の例に限定されるべきではない。

【１０７９】

[1057]図４１の例に示されているように、オーディオ復号デバイス５４０Ａは、抽出ユニット５４２と、オーディオ復号ユニット５４４と、数学ユニット５４６と、オーディオレンダリングユニット５４８とを備える。抽出ユニット５４２は、ビットストリーム５１７から、符号化された低減されたバックグラウンド球面調和係数５１５Ｂと、符号化されたU_{DIST}＊S_{DIST}ベクトル５１５Ａと、V^{T}_{DIST}ベクトル５２５Ｅとを抽出するように構成されるユニットを表す。抽出ユニット５４２は、符号化された低減されたバックグラウンド球面調和係数５１５Ｂと、符号化されたU_{DIST}＊S_{DIST}ベクトル５１５Ａとをオーディオ復号ユニット５４４に出力し、一方でまた、V^{T}_{DIST}行列５２５Ｅを数学ユニット５４６に出力する。この点で、抽出ユニット５４２は、図５の例に示されるオーディオ複号デバイス２４の抽出ユニット７２に類似した方式で動作することができる。

【１０８０】

[1058]オーディオ復号ユニット５４４は、U_{DIST}＊S_{DIST}ベクトル５２７と低減されたバックグラウンド球面調和係数５２９とを復元するように、（しばしば、ＡＡＣ復号スキームなど、逆のオーディオ複号スキームに従って）符号化されたオーディオデータを復号するためのユニットを表す。オーディオ復号ユニット５４４は、U_{DIST}＊S_{DIST}ベクトル５２７と低減されたバックグラウンド球面調和係数５２９とを、数学ユニット５４６に出力する。この点で、オーディオ復号ユニット５４４は、図５の例に示されるオーディオ複号デバイス２４の音響心理学的復号ユニット８０に類似した方式で動作することができる。

【１０８１】

[1059]数学ユニット５４６は、行列乗算および加算（ならびに、いくつかの例では、任意の他の行列数学演算）を実行するように構成されるユニットを表し得る。数学ユニット５４６はまず、U_{DIST}＊S_{DIST}ベクトル５２７のV^{T}_{DIST}行列５２５Ｅとの行列乗算を実

行することができる。数学ユニット５４６は次いで、復元された球面調和係数５４７として示される、オリジナルの球面調和係数１１の低減されたバージョンを生成するために、U_{DIST}＊S_{DIST}ベクトル５２７のV^{T}_{DIST}行列５２５Ｅと、低減されたバックグラウンド球面調和係数５２９との乗算の結果（この場合も、U_{BG}行列５２５ＤのS_{BG}行列５２５Ｂとの、次いでV^{T}_{BG}行列５２５Ｆとの乗算の結果を指し得る）を、U_{DIST}＊S_{DIST}ベクトル５２７のV^{T}_{DIST}行列５２５Ｅとの行列乗算の結果に加算することができる。数学ユニット５４６は、復元された球面調和係数５４７をオーディオレンダリングユニット５４８に出力することができる。この点で、数学ユニット５４６は、図５の例に示されるオーディオ複号デバイス２４のフォアグラウンド定式化ユニット７８およびＨＯＡ係数定式化ユニット８２に類似した方式で動作することができる。

【１０８２】

[1060]オーディオレンダリングユニット５４８は、チャネル５４９Ａ～５４９Ｎ（「マルチチャネルオーディオデータ５４９」または「ラウドスピーカーフィード５４９」と一般に呼ばれることもある、「チャネル５４９」）をレンダリングするように構成されるユニットを表す。オーディオレンダリングユニット５４８は、復元された球面調和係数５４７に変換（しばしば、行列の形式で表される）を適用することができる。復元された球面調和係数５４７は、３次元で音場を記述するので、復元された球面調和係数５４７は、大部分の復号器ローカルスピーカーの幾何学的配置（マルチチャネルオーディオデータ５４９を再生するようになるスピーカーの幾何学的配置を指し得る）に対応することが可能である方式で、マルチチャネルオーディオデータ５４９Ａのレンダリングを容易にするオーディオフォーマットを表す。マルチチャネルオーディオデータ５４９Ａのレンダリングに関する、より多くの情報は、図４８に関して上で説明される。

【１０８３】

[1061]マルチチャネルオーディオデータ５４９Ａがサラウンド音響マルチチャネルオーディオデータ５４９である状況において説明されるが、オーディオレンダリングユニット４８はまた、復元された球面調和係数５４９Ａをバイノーラル化し（binauralize）、それによって２つのバイノーラルにレンダリングされたチャネル５４９を生成するために、ある形態のバイノーラル化（binauralization）を実行することもできる。したがって、本技法は、マルチチャネルオーディオデータのサラウンド音響形式に限定されるべきではなく、バイノーラル化されたマルチチャネルオーディオデータを含み得る。

【１０８４】

[1062]以下に記載される様々な項は、本開示で説明される技法の様々な態様を提示することができる。

【１０８５】

[1063]１３２５６７－１Ｂ項。オーディオ複号デバイス５４０などのデバイスであって、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するように構成される１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して特異値分解を実行することによって生成される、デバイス。

【１０８６】

[1064]１３２５６７－２Ｂ項。１つまたは複数の第１のベクトルが、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化されたU_{DIST}＊S_{DIST}ベクトルを備え、Ｕ行列およびＳ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数のオーディオ符号化されたU_{DIST}＊S_{DIST}ベクトルのオーディオ複号されたバージョンを生成するために、１つまたは複数のオーディオ符号化されたU_{DIST}＊S_{DIST}ベクトルをオーディオ復号するようにさらに構成される、１３２５６７－１Ｂ項のデバイス。

【１０８７】
[1065]１３２５６７－３Ｂ項。１つまたは複数の第１のベクトルが、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列およびＳ行列およびＶ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルのオーディオ複号されたバージョンを生成するために、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルをオーディオ復号するようにさらに構成される、１３２５６７－１Ｂ項のデバイス。

【１０８８】
[1066]１３２５６７－４Ｂ項。１つまたは複数のプロセッサが、複数の球面調和関数のうちで音場の明瞭な成分を表すものを復元するために、$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{DIST}ベクトルと乗算するようにさらに構成される、１３２５６７－３Ｂ項のデバイス。

【１０８９】
[1067]１３２５６７－５Ｂ項。１つまたは複数の第２のベクトルが、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを備え、Ｓ行列、Ｕ行列、およびＶ行列が各々、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成される、１３２５６７－１Ｂ項のデバイス。

【１０９０】
[1068]１３２５６７－６Ｂ項。１つまたは複数の第２のベクトルが、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを備え、Ｓ行列、Ｕ行列、およびＶ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数のオーディオ復号された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを生成するために、１つまたは複数のオーディオ符号化された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルをオーディオ復号するようにさらに構成される、１３２５６７－１Ｂ項のデバイス。

【１０９１】
[1069]１３２５６７－７Ｂ項。１つまたは複数の第１のベクトルが、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルを生成するために１つまたは複数のオーディオ符号化された$U_{DIST}*S_{DIST}$ベクトルをオーディオ復号し、複数の球面調和係数のうちで音場の明瞭な成分を記述するものを復元するために$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{DIST}ベクトルと乗算するようにさらに構成され、１つまたは複数の第２のベクトルが、Ｕ行列に含まれるU_{BG}ベクトルをＳ行列に含まれるS_{BG}ベクトルと乗算し、次いでＶ行列の転置に含まれるV^{T}_{BG}ベクトルと乗算することによって、オーディオ符号化の前に生成された、１つまたは複数のオーディオ符号化された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルを備え、１つまたは複数のプロセッサが、音場のバックグラウンド成分を記述する複数の球面調和係数の少なくともある部分を復元するために、１つまたは複数のオーディオ符号化された$U_{BG}*S_{BG}*V^{T}_{BG}$ベクトルをオーディオ復号し、複数の球面調和係数の再構築されたバージョンを生成するために、音場の明瞭な成分を記述する複数の球面調和係数を音場のバックグ

ラウンド成分を記述する複数の球面調和係数の少なくともその部分に加算するようにさらに構成される、１３２５６７－１Ｂ項のデバイス。

【１０９２】

[1070]１３２５６７－８Ｂ項。１つまたは複数の第１のベクトルが、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと１つまたは複数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄを決定するようにさらに構成される、１３２５６７－１Ｂ項のデバイス。

【１０９３】

[1071]１３２５６７－９Ｂ項。１つまたは複数の第１のベクトルが、Ｕ行列の１つまたは複数のオーディオ符号化されたU_{DIST}ベクトルをＳ行列の１つまたは複数のS_{DIST}ベクトルと乗算することによってオーディオ符号化の前に生成された、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと、Ｖ行列の転置の１つまたは複数のV^{T}_{DIST}ベクトルとを備え、Ｕ行列、Ｓ行列、およびＶ行列が、少なくとも複数の球面調和係数に関して特異値分解を実行することによって生成され、１つまたは複数のプロセッサが、１つまたは複数のU_{DIST}＊S_{DIST}ベクトルと１つまたは複数のV^{T}_{DIST}ベクトルとを形成するためにビットストリームから抽出されるべきベクトルの数を示す値Ｄをオーディオフレームごとに決定するようにさらに構成される、１３２５６７－１０Ｂ項のデバイス。

【１０９４】

[1072]１３２５６７－１Ｇ項。オーディオ複号デバイス５４０などのデバイスであって、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するように構成される１つまたは複数のプロセッサを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、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関して特異値分解を実行することによって生成される、デバイス。

【１０９５】

[1073]１３２５６７－２Ｇ項。マルチチャネルオーディオデータが、複数の球面調和係数を備える、１３２５６７－１Ｇ項のデバイス。

【１０９６】

[1074]１３２５６７－３Ｇ項。１つまたは複数のプロセッサが、１３２５６７－２Ｂ項から１３２５６７－９Ｂ項の任意の組合せを実行するようにさらに構成される、１３２５６７－２Ｇ項のデバイス。

【１０９７】

[1075]上で説明された様々な項の各々から、オーディオ復号デバイス５４０Ａ～５４０Ｄは、ある方法を実行し、またはそうでなければ、オーディオ復号デバイス５４０Ａ～５４０Ｄ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。いくつかの例では、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。いくつかの例では、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５４０Ａ～５４０Ｄ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。

【１０９８】

[1076]たとえば、１３２５６７－１０Ｂ項は、方法が音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数

の第２のベクトルとを決定することを備え、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、音場を表す複数の球面調和係数に関して特異値分解を実行することによって生成される、ように前述の１３２５６７－１Ｂ項から導出され得る。

【１０９９】

[1077]別の例として、１３２５６７－１１Ｂ項は、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定するための手段と、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、複数の球面調和係数に関して特異値分解を実行することによって生成される、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルとを記憶するための手段とを備える、オーディオ復号デバイス５４０などのデバイスになるように、前述の１３２５６７－１Ｂ項から導出され得る。

【１１００】

[1078]また別の例として、１３２５６７－１２Ｂ項は、実行されると、１つまたは複数のプロセッサに、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を記述する１つまたは複数の第２のベクトルとを決定させ、１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方が少なくとも、音場を記述する高次アンビソニックスオーディオデータに含まれる複数の球面調和係数に関して特異値分解を実行することによって生成される、命令を記憶した、非一時的コンピュータ可読記憶媒体になるように、前述の１３２５６７－１Ｂ項から導出され得る。

【１１０１】

[1079]様々な項が、上で例示されたように導出される様々なデバイス、方法および非一時的コンピュータ可読記憶媒体について、１３２５６７－２Ｂ項ないし１３２５６７－９Ｂ項から同様に導出され得る。同じことが、本開示全体にわたって記載される様々な他の項について実行され得る。

【１１０２】

[1080]図４１Ｂは、２次元または３次元の音場を記述する球面調和係数を復号するために本開示において説明される技法のある態様を実行し得る例示的なオーディオ復号デバイス５４０Ｂを示すブロック図である。オーディオ復号デバイス５４０Ｂは、いくつかの例では、抽出ユニット５４２がV^{T}_{DIST}ベクトル５２５Ｅではなく、再順序付けられたV^{T}_{DIST}ベクトル５３９を抽出することができることを除いて、オーディオ復号デバイス５４０と類似であり得る。他の例では、抽出ユニット５４２は、V^{T}_{DIST}ベクトル５２５Ｅを抽出し、次いで、再順序付けられたV^{T}_{DIST}ベクトル５３９を決定するために、ビットストリーム中で指定されるか、または（他のベクトルの分析を通じて）推定される再順序付け情報に基づいて、これらのV^{T}_{DIST}ベクトル５２５Ｅを再順序付けることができる。この点で、抽出ユニット５４２は、図５の例に示されるオーディオ復号デバイス２４の抽出ユニット７２に類似した方式で動作することができる。いずれにしても、抽出ユニット５４２は、再順序付けられたV^{T}_{DIST}ベクトル５３９を数学ユニット５４６に出力することができ、ここで、球面調和係数を復元することに関して上で説明されたプロセスが、これらの再順序付けられたV^{T}_{DIST}ベクトル５３９に関して実行され得る。

【１１０３】

[1081]このようにして、本技法は、オーディオ復号デバイス５４０Ｂが、音場の明瞭な成分を表す再順序付けられた１つまたは複数のベクトルをオーディオ復号することを可能にすることができ、再順序付けられた１つまたは複数のベクトルは、１つまたは複数のベクトルを圧縮することを容易にするために再順序付けられている。これらの例および他の例において、オーディオ復号デバイス５４０Ｂは、音場の明瞭な成分を表す球面調和係数を復元するために、再順序付けられた１つまたは複数のベクトルを、再順序付けられた１つまたは複数の追加のベクトルと再び組み合わせることができる。これらの例および他の例において、オーディオ復号デバイス５４０Ｂは次いで、音場の明瞭な成分を表す球面調和係数と、音場のバックグラウンド成分を表す球面調和係数とに基づいて、複数の球面調

和係数を復元することができる。

【１１０４】

[1082]すなわち、本技法の様々な態様は、次の項に従って、再順序付けられた１つまたは複数のベクトルを復号するように構成されるべきオーディオ復号デバイス５４０Ｂを提供することができる。

【１１０５】

[1083]１３３１４６－１Ｆ項。オーディオ符号化デバイス５４０Ｂなどのデバイスであって、音場における成分に対応するベクトルの数を決定するように構成される１つまたは複数のプロセッサを備えるデバイス。

【１１０６】

[1084]１３３１４６－２Ｆ項。１つまたは複数のプロセッサが、上で説明された例の任意の組合せに従って、次数低減を実行した後、ベクトルの数を決定するように構成される、１３３１４６－１Ｆ項のデバイス。

【１１０７】

[1085]１３３１４６－３Ｆ項。１つまたは複数のプロセッサが、上で説明された例の任意の組合せに従って、次数低減を実行するようにさらに構成される、１３３１４６－１Ｆ項のデバイス。

【１１０８】

[1086]１３３１４６－４Ｆ項。１つまたは複数のプロセッサが、ビットストリームにおいて指定された値からベクトルの数を決定するように構成され、１つまたは複数のプロセッサが、音場の明瞭な成分を表すビットストリーム中の１つまたは複数のベクトルを識別するために、決定されたベクトルの数に基づいてビットストリームを解析するようにさらに構成される、１３３１４６－１Ｆ項のデバイス。

【１１０９】

[1087]１３３１４６－５Ｆ項。１つまたは複数のプロセッサが、ビットストリームにおいて指定された値からベクトルの数を決定するように構成され、１つまたは複数のプロセッサが、音場のバックグラウンド成分を表すビットストリーム中の１つまたは複数のベクトルを識別するために、決定されたベクトルの数に基づいてビットストリームを解析するようにさらに構成される、１３３１４６－１Ｆ項のデバイス。

【１１１０】

[1088]１３３１４３－１Ｃ項。オーディオ復号デバイス５４０Ｂなどのデバイスであって、音場の明瞭な成分を表す再順序付けられた１つまたは複数のベクトルを再順序付けるように構成される１つまたは複数のプロセッサを備えるデバイス。

【１１１１】

[1089]１３３１４３－２Ｃ項。１つまたは複数のプロセッサが、再順序付けられた１つまたは複数のベクトルを決定することと、再順序付けられた１つまたは複数のベクトルがどのように再順序付けられたかを記述する再順序付け情報を決定することとを行うようにさらに構成され、１つまたは複数のプロセッサが、再順序付けられた１つまたは複数のベクトルを再順序付けるとき、決定された再順序付け情報に基づいて、再順序付けられた１つまたは複数のベクトルを再順序付けるようにさらに構成される、１３３１４３－１Ｃ項のデバイス。

【１１１２】

[1090]１３３１４３－３Ｃ項。再順序付けられた１つまたは複数のベクトルが、請求項１Ａから１８Ａの任意の組合せ、または請求項１Ｂから１９Ｂの任意の組合せによって記載された、１つまたは複数の再順序付けられた第１のベクトルを備え、１つまたは複数の第１のベクトルが、請求項１Ａから１８Ａの任意の組合せ、または請求項１Ｂから１９Ｂの任意の組合せによって記載された方法に従って決定される、１Ｃのデバイス。

【１１１３】

[1091]１３３１４３－４Ｄ項。オーディオ復号デバイス５４０Ｂなどのデバイスであって、音場の明瞭な成分を表す再順序付けられた１つまたは複数のベクトルをオーディオ復

号するように構成される１つまたは複数のプロセッサを備え、再順序付けられた１つまたは複数のベクトルが、１つまたは複数のベクトルを圧縮することを容易にするために再順序付けられている、デバイス。

【１１１４】

[1092]１３３１４３－５Ｄ項。１つまたは複数のプロセッサが、音場の明瞭な成分を表す球面調和係数を復元するために、再順序付けられた１つまたは複数のベクトルを、再順序付けられた１つまたは複数の追加のベクトルと再び組み合わせるようにさらに構成される、１３３１４３－４Ｄ項のデバイス。

【１１１５】

[1093]１３３１４３－６Ｄ項。１つまたは複数のプロセッサが、音場の明瞭な成分を表す球面調和係数と、音場のバックグラウンド成分を表す球面調和係数とに基づいて、複数の球面調和係数を復元するようにさらに構成される、１３３１４３－５Ｄ項のデバイス。

【１１１６】

[1094]１３３１４３－１Ｅ項。オーディオ復号デバイス５４０Ｂなどのデバイスであって、再順序付けられた１つまたは複数の第１のベクトルを生成するために、１つまたは複数のベクトルを再順序付け、それによって、従来のオーディオ符号化器による符号化を容易にすることと、１つまたは複数のベクトルが、音場の明瞭な成分を記述する表す、再順序付けられた１つまたは複数のベクトルの符号化されたバージョンを生成するために、従来のオーディオ符号化器を使用して、再順序付けられた１つまたは複数のベクトルをオーディオ符号化することとを行うように構成される１つまたは複数のプロセッサを備えるデバイス。

【１１１７】

[1095]１３３１４３－２Ｅ項。再順序付けられた１つまたは複数のベクトルが、請求項１Ａから１８Ａの任意の組合せ、または請求項１Ｂから１９Ｂの任意の組合せによって記載された、１つまたは複数の再順序付けられた第１のベクトルを備え、１つまたは複数の第１のベクトルが、請求項１Ａから１８Ａの任意の組合せ、または請求項１Ｂから１９Ｂの任意の組合せによって記載された方法に従って決定される、１Ｅのデバイス。

【１１１８】

[1096]図４１Ｃは、別の例示的なオーディオ符号化デバイス５４０Ｃを示すブロック図である。オーディオ復号デバイス５４０Ｃは、デスクトップコンピュータ、ラップトップコンピュータ、ワークステーション、タブレットもしくはスレートコンピュータ、専用オーディオ録音デバイス、携帯電話（いわゆる「スマートフォン」を含む）、パーソナルメディアプレーヤデバイス、パーソナルゲームデバイス、または、オーディオデータを復号することが可能な任意の他のタイプのデバイスなど、オーディオデータを復号することが可能な任意のデバイスを表し得る。

【１１１９】

[1097]図４１Ｃの例では、オーディオ復号デバイス５４０Ｃは、いくつかの例では、外部の無関係なデータの除去を容易にするためにオーディオ符号化デバイス５１０Ｂ～５１０Ｊによって使用される、（図４０Ｂ～図４０Ｊの例に関して上で説明されたような）次数低減を実行することを除いて、オーディオ符号化デバイス５１０Ｂ～５１０Ｅのいずれかによって実行されるオーディオ符号化プロセスとは逆であるオーディオ復号プロセスを実行する。

【１１２０】

[1098]図４１Ｃの例では、単一のデバイス、すなわち、デバイス５４０Ｃとして示されるが、デバイス５４０Ｃ内に含まれるものとして以下で言及される様々な構成要素またはユニットは、デバイス５４０Ｃから外部である別個のデバイスを形成することができる。言い換えれば、図４１Ｃの例では、単一のデバイス、すなわち、デバイス５４０Ｃによって実行されるものとして本開示で説明されるが、本技法は、複数のデバイスを備えるシステムによって実施され、またはさもなければ実行されてよく、ここで、これらのデバイスの各々は、以下でより詳細に説明される様々な構成要素またはユニットのうちの１つまた

は複数をそれぞれ含み得る。したがって、本技法は、この点で、図４１Ｃの例に限定されるべきではない。

【１１２１】

[1099]その上、オーディオ符号化デバイス５４０Ｃは、オーディオ符号化デバイス５４０Ｂと類似であり得る。ただし、抽出ユニット５４２は、再順序付けられた$V^{T}_{Q_DIST}$ベクトル５３９または（図４０のオーディオ符号化デバイス５１０に関して説明された場合のように）V^{T}_{DIST}ベクトル５２５Ｅではなく、ビットストリーム５１７から１つまたは複数のV^{T}_{SMALL}ベクトル５２１を決定することができる。結果として、抽出ユニット５４２は、V^{T}_{SMALL}ベクトル５２１を数学ユニット５４６に渡すことができる。

【１１２２】

[1100]加えて、抽出ユニット５４２は、ビットストリーム５１７からオーディオ符号化された修正されたバックグラウンド球面調和係数５１５Ｂ’を決定し、これらの係数５１５Ｂ’をオーディオ復号ユニット５４４に渡すことができ、オーディオ復号ユニット５４４は、修正されたバックグラウンド球面調和係数５３７を復元するために、符号化された修正されたバックグラウンド球面調和係数５１５Ｂをオーディオ復号することができる。オーディオ復号ユニット５４４は、これらの修正されたバックグラウンド球面調和係数５３７を数学ユニット５４６に渡すことができる。

【１１２３】

[1101]数学ユニット５４６は次いで、高次の明瞭な球面調和係数を復元するために、オーディオ復号された（および、場合によっては再順序付けられていない）$U_{DIST}*S_{DIST}$ベクトル５２７’を、１つまたは複数のV^{T}_{SMALL}ベクトル５２１と乗算することができる。数学ユニット５４６は次いで、複数の球面調和係数５１１またはそのある派生物（符号化器ユニット５１０Ｅにおいて実行される次数低減による派生物であり得る）を復元するために、高次の明瞭な球面調和係数を、修正されたバックグラウンド球面調和係数５３７に加算することができる。

【１１２４】

[1102]このようにして、本技法は、オーディオ複号デバイス５４０Ｃが、ビットストリームから、ビットストリーム中の１つまたは複数のベクトルに割り振られることが必要とされるビットの量を低減するために、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルのうちの少なくとも１つを決定することを可能にすることができ、球面調和係数は音場を記述し、バックグラウンド球面調和係数は、同じ音場の１つまたは複数のバックグラウンド成分を記述した。

【１１２５】

[1103]本技法の様々な態様は、この点で、オーディオ復号デバイス５４０Ｃが、いくつかの例では、ビットストリームから、バックグラウンド球面調和係数と再び組み合わせられた球面調和係数から分解された１つまたは複数のベクトルの少なくとも１つを決定するように構成されることを可能にすることができ、球面調和係数は音場を記述し、バックグラウンド球面調和係数は同じ音場の１つまたは複数のバックグラウンド成分を記述した。

【１１２６】

[1104]これらの例および他の例において、オーディオ復号デバイス５４０Ｃは、ビットストリームから、N_{BG}に等しい次数を有する第１の球面調和係数の一部を取得するように構成される。

【１１２７】

[1105]これらの例および他の例において、オーディオ復号デバイス５４０Ｃは、ビットストリームから、N_{BG}に等しい次数を有する第１のオーディオ符号化された部分球面調和係数を取得し、球面調和係数の第１の部分を生成するために、球面調和係数のオーディオ符号化された第１の部分をオーディオ復号するようにさらに構成される。

【１１２８】

[1106]これらの例および他の例において、１つまたは複数のベクトルの少なくとも１つは、１つまたは複数のV^{T}_{SMALL}ベクトルを備え、１つまたは複数のV^{T}_{SMALL}ベクトルは、

複数の球面調和係数に関して特異値分解を実行することによって生成されたＶ行列の転置から決定されている。

【１１２９】

[1107]これらの例および他の例において、１つまたは複数のベクトルの少なくとも１つは、１つまたは複数のV^{T}_{SMALL}ベクトルを備え、１つまたは複数のV^{T}_{SMALL}ベクトルは、複数の球面調和係数に関して特異値分解を実行することによって生成されたＶ行列の転置から決定されており、オーディオ復号デバイス５４０Ｃは、ビットストリームから、その両方が複数の球面調和係数に関して特異値分解を実行することによって生成された、Ｕ行列およびＳ行列から導出された１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルを取得し、$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{SMALL}ベクトルと乗算するようにさらに構成される。

【１１３０】

[1108]これらの例および他の例において、１つまたは複数のベクトルの少なくとも１つは、１つまたは複数のV^{T}_{SMALL}ベクトルを備え、１つまたは複数のV^{T}_{SMALL}ベクトルは、複数の球面調和係数に関して特異値分解を実行することによって生成されたＶ行列の転置から決定されており、オーディオ復号デバイス５４０Ｃは、ビットストリームから、その両方が複数の球面調和係数に関して特異値分解を実行することによって生成された、Ｕ行列およびＳ行列から導出された１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルを取得し、高次の明瞭なバックグラウンド球面調和係数を復元するために、$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{SMALL}ベクトルと乗算し、少なくとも一部、複数の球面調和係数を復元するために、低次の明瞭なバックグラウンド球面調和係数を含むバックグラウンド球面調和係数を、高次の明瞭なバックグラウンド球面調和係数に加算するようにさらに構成される。

【１１３１】

[1109]これらの例および他の例において、１つまたは複数のベクトルの少なくとも１つは、１つまたは複数のV^{T}_{SMALL}ベクトルを備え、１つまたは複数のV^{T}_{SMALL}ベクトルは、複数の球面調和係数に関して特異値分解を実行することによって生成されたＶ行列の転置から決定されており、オーディオ復号デバイス５４０Ｃは、ビットストリームから、その両方が複数の球面調和係数に関して特異値分解を実行することによって生成された、Ｕ行列およびＳ行列から導出された１つまたは複数の$U_{DIST}*S_{DIST}$ベクトルを取得し、高次の明瞭なバックグラウンド球面調和係数を復元するために、$U_{DIST}*S_{DIST}$ベクトルをV^{T}_{SMALL}ベクトルと乗算し、少なくとも一部、複数の球面調和係数を復元するために、低次の明瞭なバックグラウンド球面調和係数を含むバックグラウンド球面調和係数を、高次の明瞭なバックグラウンド球面調和係数に加算し、復元された複数の球面調和係数をレンダリングするようにさらに構成される。

【１１３２】

[1110]図４１Ｄは、別の例示的なオーディオ符号化デバイス５４０Ｄを示すブロック図である。オーディオ復号デバイス５４０Ｄは、デスクトップコンピュータ、ラップトップコンピュータ、ワークステーション、タブレットもしくはスレートコンピュータ、専用オーディオ録音デバイス、携帯電話（いわゆる「スマートフォン」を含む）、パーソナルメディアプレーヤデバイス、パーソナルゲームデバイス、または、オーディオデータを復号することが可能な任意の他のタイプのデバイスなど、オーディオデータを復号することが可能な任意のデバイスを表し得る。

【１１３３】

[1111]図４１Ｄの例では、オーディオ復号デバイス５４０Ｄは、いくつかの例では、外部の無関係なデータの除去を容易にするためにオーディオ符号化デバイス５１０Ｂ～５１０Ｊによって使用される、（図４０Ｂ～図４０Ｊの例に関して上で説明されたような）次数低減を実行することを除いて、オーディオ符号化デバイス５１０Ｂ～５１０Ｊのいずれかによって実行されるオーディオ符号化プロセスとは逆であるオーディオ復号プロセスを実行する。

【１１３４】

[1112]図４１Ｄの例では、単一のデバイス、すなわち、デバイス５４０Ｄとして示され

るが、デバイス５４０Ｄ内に含まれるものとして以下で言及される様々な構成要素またはユニットは、デバイス５４０Ｄから外部である別個のデバイスを形成することができる。言い換えれば、図４１Ｄの例では、単一のデバイス、すなわち、デバイス５４０Ｄによって実行されるものとして本開示で説明されるが、本技法は、複数のデバイスを備えるシステムによって実施され、またはさもなければ実行されてよく、ここで、これらのデバイスの各々は、以下でより詳細に説明される様々な構成要素またはユニットのうちの１つまたは複数をそれぞれ含み得る。したがって、本技法は、この点で、図４１Ｄの例に限定されるべきではない。

【１１３５】

[1113]その上、オーディオ復号デバイス５４０Ｄは、オーディオ復号デバイス５４０Ｄが、図４０Ｉに関して上で説明されたＶ圧縮ユニット５５２によって実行された圧縮とは全般に逆である追加のＶ解凍を実行することを除いて、オーディオ復号デバイス５４０Ｂと類似であり得る。図４１Ｄの例では、抽出ユニット５４２は、ビットストリーム５１７に含まれる（および、図１０Ｂおよび図１０Ｃのうちの１つに示された例に従って一般に指定される）圧縮された空間成分５３９’のこのＶ解凍を実行する、Ｖ解凍ユニット５５５を含む。Ｖ解凍ユニット５５５は、次の式に基づいて、V^{T}_{DIST}ベクトル５３９を解凍することができる。

【１１３６】

【数３５】

$$\hat{v}_q = \begin{cases} 0, & cid = 0 \text{ の場合} \\ sgn * (2^{cid-1} + \text{残差}), & cid \neq 0 \text{ の場合} \end{cases}$$

【１１３７】

言い換えれば、Ｖ解凍ユニット５５５は、まずビットストリーム５１７からｎｂｉｔｓ値を解析し、ｃｉｄを表すハフマンコードを復号するときに使用するために、５つのハフマンコードテーブルの適切なセットを識別することができる。ビットストリーム５１７において指定された予測モードおよびハフマンコーディング情報と、場合によっては空間成分の他の要素に対する空間成分のその要素の次数とに基づいて、Ｖ解凍ユニット５５５は、解析されたｎｂｉｔｓ値について定義された５つのハフマンテーブルのうちの正しい１つを識別することができる。このハフマンテーブルを使用して、Ｖ解凍ユニット５５５は、ハフマンコードからｃｉｄ値を復号することができる。Ｖ解凍ユニット５５５は次いで、符号ビットと残差ブロックコードとを解析し、残差を識別するために、残差ブロックコードを復号することができる。上の式に従って、Ｖ解凍ユニット５５５は、V^{T}_{DIST}ベクトル５３９のうちの１つを復号することができる。

【１１３８】

[1114]上記は、次のシンタックステーブルにおいて要約され得る。

【１１３９】

【表１４－１】

テーブル－復号されたベクトル

シンタックス	ビットの数	ニーモニック

```
decodeVVec(i)
{
 switch codedVVecLength {
    case 0: ／／完全ベクトル
       VVecLength = NumOfHoaCoeffs;
       for (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
       break;
    case 1: ／／より低い次数が除去される
       VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA;

       for (m=0; m< VVecLength; ++m){
        VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA +
1;
        }
       break;
    case 2:
       VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
       n = 0;
       for(m=0;m<NumOfHoaCoeffs–
MinNumOfCoeffsForAmbHOA; ++m){
        c = m + MinNumOfCoeffsForAmbHOA + 1;
        if ( ismember(c, AmbCoeffIdx) == 0){
           VecCoeff[n] = c;
           n++;
         }
       }
       break;
    case 3:
       VVecLength = NumOfHoaCoeffs –
NumOfAddAmbHoaChan;
      n = 0;
      for(m=0; m<NumOfHoaCoeffs; ++m){
        c = m + 1;
```

【表１４－２】

if (*ismember*(c, AmbCoeffIdx) == 0){
VecCoeff[n] = c;
n++;
}
}
}
}
 	
if (NbitsQ[i] == 5) {／*一様量子化器*／
for (m=0; m< VVecLength; ++m){
VVec(k)[i][m] = (**VecValue** / 128.0) – 1.0;
	8	**uimsbf**

}
}
else { ／*ハフマン復号*／
for (m=0; m< VVecLength; ++m){
Idx = 5;
If (CbFlag[i] == 1) {
idx = (min(3, max(1, ceil(sqrt(VecCoeff[m]) -
1)));
}
else if (PFlag[i] == 1) {idx = 4;}
cid =
	動的	**huffDecode**

huffDecode(huffmannTable[NbitsQ].codebook[idx]; **huffVal**);
if (cid > 0) {
aVal = sgn = (**sgnVal** * 2) - 1;
	1	**bslbf**

if (cid > 1) { aVal = sgn * (2.0^(cid -1) + **intAddVal**);	**cid-1**	**uimsbf**
} } else {aVal = 0.0;} } } }		

注:
－ 一様量子化器のためのエンコーダ関数はmin(255, round((x+1.0)*128.0))
－ ニーモニックhuffDecodeのためのビットの数は動的である

【１１４０】

[1115]前述のシンタックステーブルにおいて、４つのケース（ケース０～３）を伴う第１のスイッチステートメントは、V^{T}_{DIST}ベクトルの長さを係数の数に関して決定する方法を提供する。第１のケースであるケース０は、V^{T}_{DIST}ベクトルの係数のすべてが指定されることを示す。第２のケースであるケース１は、V^{T}_{DIST}ベクトルの係数のうちでＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡよりも大きい次数に対応するもののみが指定されることを示し、これは上で（N_{DIST}＋１）－（N_{BG}＋１）と呼ばれるものを示し得る。第３のケースであるケース２は、第２のケースに類似しているが、ＮｕｍＯｆＡｄｄＡｍｂＨｏａＣｈａｎによって識別される係数をさらに差し引き、これは次数N_{BG}を超

える次数に対応する追加のチャネル（ここで「チャネル」は、ある次数と位数の組合せに対応する特定の係数を指す）を指定するための変数を示す。第４のケースであるケース３は、V^{T}_{DIST}ベクトルの係数のうちでＮｕｍＯｆＡｄｄＡｍｂＨｏａＣｈａｎによって識別される係数を除去した後で残るもののみが指定されることを示す。

【１１４１】

[1116]このスイッチステートメントの後で、一様逆量子化を実行するかどうかの決定は、ＮｂｉｔｓＱ（または上で示されたようにｎｂｉｔｓ）によって制御され、ＮｂｉｔｓＱが５に等しくない場合、ハフマン復号の適用となる。上で言及されるｃｉｄ値は、ＮｂｉｔｓＱ値の下位２ビットに等しい。上で論じられた予測モードは上のシンタックステーブルではＰＦｌａｇとして示されるが、ＨＴ情報ビットは上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、復号が上で説明されたものと実質的に同様の方式でどのように行われるかを指定する。

【１１４２】

[1117]このようにして、本開示の技法は、オーディオ復号デバイス５４０Ｄが、音場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を取得することと、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、空間成分を取得するために空間成分の圧縮されたバージョンを解凍することとを可能にすることができる。

【１１４３】

[1118]その上、本技法は、オーディオ複号デバイス５４０Ｄが、音場の空間成分の圧縮されたバージョンを解凍することを可能にすることができ、空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される。

【１１４４】

このようにして、オーディオ符号化デバイス５４０Ｄは、次の項に関して以下で記載される技法の様々な態様を実行することができる。

【１１４５】

[1119]１４１５４１－１Ｂ項。

【１１４６】

[1120]音場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を取得することと、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、空間成分を取得するために空間成分の圧縮されたバージョンを解凍することとを行うように構成される１つまたは複数のプロセッサを備えるデバイス。

【１１４７】

[1121]１４１５４１－２Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用される予測モードを指定するフィールドを使用して、ビットストリームにおいて表され、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、空間成分を取得するために、予測モードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項のデバイス。

【１１４８】

[1122]１４１５４１－３Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用されるハフマンテーブルを指定するハフマンテーブル情報を使用して、ビットストリームにおいて表され、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、ハフマンテーブル情報に少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項および１４１５４１－２Ｂ項の任意の組合せのデバイス。

【１１４９】

[1123]１４１５４１－４Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表す値を示すフィールドを使用して、ビットストリームにおいて表され、１つまたは複数のプロセッ

サが、空間成分の圧縮されたバージョンを解凍するとき、その値に少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項から１４１５４１－３Ｂ項の任意の組合せのデバイス。

【１１５０】

[1124]１４１５４１－５Ｂ項。値がｎｂｉｔｓ値を備える、１４１５４１－４Ｂ項のデバイス。

【１１５１】

[1125]１４１５４１－６Ｂ項。ビットストリームが、空間成分の圧縮されたバージョンが含まれる音場の複数の空間成分の圧縮されたバージョンを備え、値が、複数の空間成分を圧縮するときに使用される量子化ステップサイズまたはその変数を表し、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、その値に少なくとも一部基づいて空間成分の複数の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－４Ｂ項および１４１５４１－５Ｂ項の任意の組合せのデバイス。

【１１５２】

[1126]１４１５４１－７Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分が対応する圧縮カテゴリを識別するカテゴリ識別子を表すためのハフマンコードを使用して、ビットストリームにおいて表され、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、ハフマンコードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項から１４１５４１－６Ｂ項の任意の組合せのデバイス。

【１１５３】

[1127]１４１５４１－８Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分が正の値か負の値かを識別する符号ビットを使用して、ビットストリームにおいて表され、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、符号ビットに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項から１４１５４１－７Ｂ項の任意の組合せのデバイス。

【１１５４】

[1128]１４１５４１－９Ｂ項。空間成分の圧縮されたバージョンが、少なくとも一部、空間成分の残差値を表すためのハフマンコードを使用して、ビットストリームにおいて表され、１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、ハフマンコードに少なくとも一部基づいて空間成分の圧縮されたバージョンを解凍するようにさらに構成される、１４１５４１－１Ｂ項から１４１５４１－８Ｂ項の任意の組合せのデバイス。

【１１５５】

[1129]１４１５４１－１０Ｂ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ｂ項から１４１５４１－１０Ｂ項の任意の組合せのデバイス。

【１１５６】

[1130]さらに、オーディオ複号デバイス５４０Ｄは、次の項に関して以下で記載される技法の様々な態様を実行するように構成され得る。

【１１５７】

[1131]１４１５４１－１Ｃ項。オーディオ複号デバイス５４０Ｄなどのデバイスであって、音場の空間成分の圧縮されたバージョンを解凍するように構成される１つまたは複数のプロセッサを備え、空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。

【１１５８】

[1132]１４１５４１－２Ｃ項。１つまたは複数のプロセッサが、空間成分の圧縮されたバージョンを解凍するとき、空間成分が圧縮されたときに分類されたカテゴリを識別するカテゴリ識別子を取得することと、空間成分が正の値か負の値かを識別する符号を取得することと、空間成分の圧縮されたバージョンと関連付けられる残差値を取得することと、

カテゴリ識別子と、符号と、残差値とに基づいて空間成分の圧縮されたバージョンを解凍することとを行うようにさらに構成される、１４１５４１－１Ｃ項および１４１５４１－２Ｃ項の任意の組合せのデバイス。

【１１５９】

[1133]１４１５４１－３Ｃ項。１つまたは複数のプロセッサが、カテゴリ識別子を取得するとき、カテゴリ識別子を表すハフマンコードを取得することと、カテゴリ識別子を取得するためにハフマンコードを複号することとを行うようにさらに構成される、１４１５４１－２Ｃ項のデバイス。

【１１６０】

[1134]１４１５４１－４Ｃ項。１つまたは複数のプロセッサが、ハフマンコードを複号するとき、少なくとも一部、複数の空間成分を指定するベクトルにおける空間成分の相対位置に基づいて、ハフマンコードを複号するために使用されるハフマンテーブルを識別するようにさらに構成される、１４１５４１－３Ｃ項のデバイス。

【１１６１】

[1135]１４１５４１－５Ｃ項。１つまたは複数のプロセッサが、ハフマンコードを複号するとき、少なくとも一部、空間成分を圧縮するときに使用される予想モードに基づいて、ハフマンコードを複号するために使用されるハフマンテーブルを識別するようにさらに構成される、１４１５４１－３Ｃ項および１４１５４１－４Ｃ項の任意の組合せのデバイス。

【１１６２】

[1136]１４１５４１－６Ｃ項。１つまたは複数のプロセッサが、ハフマンコードを複号するとき、少なくとも一部、空間成分の圧縮されたバージョンと関連付けられるハフマンテーブル情報に基づいて、ハフマンコードを複号するために使用されるハフマンテーブルを識別するようにさらに構成される、１４１５４１－３Ｃ項から１４１５４１－５Ｃ項の任意の組合せのデバイス。

【１１６３】

[1137]１４１５４１－７Ｃ項。１つまたは複数のプロセッサが、ハフマンコードを複号するとき、少なくとも一部、複数の空間成分を指定するベクトルにおける空間成分の相対位置と、空間成分を圧縮するときに使用される予想モードと、空間成分の圧縮されたバージョンと関連付けられるハフマンテーブル情報とに基づいて、ハフマンコードを複号するために使用されるハフマンテーブルを識別するようにさらに構成される、１４１５４１－３Ｃ項のデバイス。

【１１６４】

[1138]１４１５４１－８Ｃ項。１つまたは複数のプロセッサが、残差値を取得するとき、残差値を取得するために、残差値を表すブロックコードを復号するようにさらに構成される、１４１５４１－２Ｃ項のデバイス。

【１１６５】

[1139]１４１５４１－９Ｃ項。ベクトルベースの合成が特異値分解を備える、１４１５４１－１Ｃ項から１４１５４１－８Ｃ項の任意の組合せのデバイス。

【１１６６】

[1140]さらに、オーディオ複号デバイス５４０Ｄは、次の項に関して以下で記載される技法の様々な態様を実行するように構成され得る。

【１１６７】

[1141]１４１５４１－１Ｇ項。オーディオ複号デバイス５４０Ｄなどのデバイスであって、複数の圧縮された空間成分の現在の空間成分の圧縮されたバージョンを複数の圧縮された空間成分の残りに対する現在の空間成分の圧縮されたバージョンの次数に基づいて解凍するときに使用するハフマンコードブックを識別するように構成される１つまたは複数のプロセッサを備え、その空間成分が、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、デバイス。

【１１６８】

[1142]１４１５４１－２Ｇ項。１つまたは複数のプロセッサが、１４１５４１－１Ｄ項から１４１５４１－１０Ｄ項、および１４１５４１－１Ｅ項から１４１５４１－９Ｅ項に記載されたステップの任意の組合せを実行するようにさらに構成される、１４１５４１－１Ｇ項のデバイス。

【１１６９】

[1143]図４２～図４２Ｃは各々、図４０Ｂ～図４０Ｊの例に示される次数低減ユニット５２８Ａをより詳細に示すブロック図である。図４２は、図４０Ｂ～図４０Ｊの次数低減ユニット５２８Ａの一例を表し得る、次数低減ユニット５２８を示すブロック図である。次数低減ユニット５２８Ａは、目標ビットレート５３５を受信するか、または別様に決定し、この目標ビットレート５３５のみに基づいて、バックグラウンド球面調和係数５３１に関して次数低減を実行することができる。いくつかの例では、次数低減ユニット５２８Ａは、低減されたバックグラウンド球面調和係数５２９を生成するために、次数および／または位数のうちで、バックグラウンド球面調和係数５３１から除去されるべきであるものを識別するために、目標ビットレート５３５を使用して、テーブルまたは他のデータ構造にアクセスすることができる。

【１１７０】

[1144]このようにして、本技法は、オーディオ符号化デバイス５１０Ｂ～４１０Ｊなどのオーディオ符号化デバイスが、目標ビットレート５３５に基づいて、低減された球面調和係数５２９または低減された球面調和係数５２９の分解を生成するために、バックグラウンド球面調和係数５３１など、複数の球面調和係数または複数の球面調和係数の分解に関する次数低減を実行することを可能にすることができ、複数の球面調和係数は音場を表す。

【１１７１】

[1145]上で説明された様々な例の各々において、オーディオ復号デバイス５４０は、ある方法を実行し、またはそうでなければ、オーディオ復号デバイス５４０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。いくつかの例では、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。いくつかの例では、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５４０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。

【１１７２】

[1146]図４２Ｂは、図４０Ｂ～図４０Ｊの次数低減ユニット５２８Ａの一例を表し得る、次数低減ユニット５２８Ｂを示すブロック図である。図４２Ｂの例では、目標ビットレート５３５のみに基づいて次数低減を実行するのではなく、次数低減ユニット５２８Ｂは、バックグラウンド球面調和係数５３１のコンテンツ分析に基づいて、次数低減を実行することができる。次数低減ユニット５２８Ｂは、このコンテンツ分析を実行するコンテンツ分析ユニット５３６Ａを含み得る。

【１１７３】

[1147]いくつかの例では、コンテンツ分析ユニット５３６Ａは、空間分析と呼ばれるコンテンツ分析の形態を実行する空間分析ユニット５３６Ａを含み得る。空間分析は、音場のバックグラウンド成分の形状または他の空間特性を記述する空間情報を識別するために、バックグラウンド球面調和係数５３１を分析することを伴い得る。この空間情報に基づいて、次数低減ユニット５２８Ｂは、低減されたバックグラウンド球面調和係数５２９を生成するために、次数および／または位数のうちで、バックグラウンド球面調和係数５３１から除去されるべきであるものを識別することができる。

【１１７４】

[1148]いくつかの例では、コンテンツ分析ユニット５３６Ａは、拡散分析と呼ばれるコンテンツ分析の形態を実行する拡散分析ユニット５３６Ｂを含み得る。拡散分析は、音場

のバックグラウンド成分の拡散率を記述する拡散情報を識別するために、バックグラウンド球面調和係数５３１を分析することを伴い得る。この拡散情報に基づいて、次数低減ユニット５２８Ｂは、低減された球面調和係数５２９を生成するために、次数および／または位数のうちで、バックグラウンド球面調和係数５３１から除去されるべきであるものを識別することができる。

【１１７５】

[1149]空間分析ユニット５３６Ａと拡散分析ユニット３６Ｂの両方を含むものとして示されるが、コンテンツ分析ユニット５３６Ａは、空間分析ユニット５３６のみ、拡散分析ユニット５３６Ｂのみ、または空間分析ユニット５３６Ａと拡散分析ユニット５３６Ｂの両方を含み得る。いくつかの例では、コンテンツ分析ユニット５３６Ａは、空間分析と拡散分析の一方または両方に加えて、またはその代替として、他の形態のコンテンツ分析を実行することができる。したがって、本開示で説明される技法は、この点について限定されるべきではない。

【１１７６】

[1150]このようにして、本技法は、オーディオ符号化デバイス５１０Ｂ～５１０Ｊなどのオーディオ符号化デバイスが、音場を記述する複数の球面調和係数または複数の球面調和係数の分解のコンテンツ分析に基づいて、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解に関する次数低減を実行することを可能にすることができる。

【１１７７】

[1151]言い換えれば、本技法は、オーディオ符号化デバイス５１０Ｂ～５１０Ｊなどのデバイスが、次の項に従って構成されることを可能にすることができる。

【１１７８】

[1152]１３３１４６－１Ｅ項。オーディオ符号化デバイス５１０Ｂ～５１０Ｊのいずれかなどのデバイスであって、音場を記述する複数の球面調和係数または複数の球面調和係数の分解のコンテンツ分析に基づいて、低減された球面調和係数または低減され球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解に関する次数低減を実行するように構成される１つまたは複数のプロセッサを備えるデバイス。

【１１７９】

[1153]１３３１４６－２Ｅ項。１つまたは複数のプロセッサが、次数低減を実行する前に、音場の明瞭な成分を記述する１つまたは複数の第１のベクトルと、音場のバックグラウンド成分を識別する１つまたは複数の第２のベクトルとを識別するために、複数の球面調和係数に関して特異値分解を実行するようにさらに構成され、１つまたは複数のプロセッサが、１つまたは複数の第１のベクトル、１つまたは複数の第２のベクトル、あるいは１つまたは複数の第１のベクトルと１つまたは複数の第２のベクトルの両方に関して次数低減を実行するように構成される、１３３１４６－１Ｅ項のデバイス。

【１１８０】

[1154]１３３１４６－３Ｅ項。１つまたは複数のプロセッサが、複数の球面調和係数または複数の球面調和係数の分解に関してコンテンツ分析を実行するようにさらに構成される、１３３１４６－１Ｅ項のデバイス。

【１１８１】

[1155]１３３１４６－４Ｅ項。１つまたは複数のプロセッサが、複数の球面調和係数または複数の球面調和係数の分解に関して空間分析を実行するように構成される、１３３１４６－３Ｅ項のデバイス。

【１１８２】

[1156]１３３１４６－５Ｅ項。コンテンツ分析を実行することが、複数の球面調和係数または複数の球面調和係数の分解に関して拡散分析を実行することを備える、１３３１４６－３Ｅ項のデバイス。

【１１８３】

[1157]１３３１４６－６Ｅ項。１つまたは複数のプロセッサが、複数の球面調和係数または複数の球面調和係数の分解に関して空間分析と拡散分析とを実行するように構成される、１３３１４６－３Ｅ項のデバイス。

【１１８４】

[1158]１３３１４６－７Ｅ項。１つまたは複数のプロセッサが、複数の球面調和係数または複数の球面調和係数の分解のコンテンツ分析と、目標ビットレートとに基づいて、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解に関する次数低減を実行するように構成される、請求項１のデバイス。

【１１８５】

[1159]１３３１４６－８Ｅ項。１つまたは複数のプロセッサが、低減された球面調和係数または低減された球面調和係数の分解をオーディオ符号化するようにさらに構成される、１３３１４６－１Ｅ項のデバイス。

【１１８６】

[1160]１３３１４６－９Ｅ項。１つまたは複数のプロセッサが、低減された球面調和係数または低減された球面調和係数の分解をオーディオ符号化することと、低減された球面調和係数または低減された球面調和係数の分解を含むようにビットストリームを生成することとを行うようにさらに構成される、１３３１４６－１Ｅ項のデバイス。

【１１８７】

[1161]１３３１４６－１０Ｅ項。１つまたは複数のプロセッサが、低減された球面調和係数または低減された球面調和係数の分解を含むビットストリームにおいて、低減された球面調和係数または低減された球面調和係数の分解のそれらが対応する球面基底関数の１つまたは複数の次数および／あるいは１つまたは複数の位数を指定するようにさらに構成される、１３３１４６－１Ｅ項のデバイス。

【１１８８】

[1162]１３３１４６－１１Ｅ項。低減された球面調和係数または低減された球面調和係数の分解が、複数の球面調和係数または複数の球面調和係数の分解よりも小さい値を有する、１３３１４６－１Ｅ項のデバイス。

【１１８９】

[1163]１３３１４６－１２Ｅ項。１つまたは複数のプロセッサが、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解のベクトルのうちで指定された次数および／または位数を有するものを除去するようにさらに構成される、１３３１４６－１Ｅ項のデバイス。

【１１９０】

[1164]１３３１４６－１３Ｅ項。１つまたは複数のプロセッサが、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解のベクトルのうちで指定された次数および／または位数を有するものを０にするように構成される、１３３１４６－１Ｅ項のデバイス。

【１１９１】

[1165]図４２Ｃは、図４０Ｂ～図４０Ｊの次数低減ユニット５２８Ａの一例を表し得る、次数低減ユニット５２８Ｃを示すブロック図である。図４２Ｂの次数低減ユニット５２８Ｃは、次数低減ユニット５２８Ｂと実質的に同じであるが、図４２の次数低減ユニット５２８Ａに関して上で説明された方式で目標ビットレート５３５を受信するか、または別様に決定し、一方でまた、図４２Ｂの次数低減ユニット５２８Ｂに関して上で説明された方式でコンテンツ分析を実行することもできる。次数低減ユニット５２８Ｃは次いで、この目標ビットレート５３５およびコンテンツ分析に基づいて、バックグラウンド球面調和係数５３１に関して次数低減を実行することができる。

【１１９２】

[1166]このようにして、本技法は、オーディオ符号化デバイス５１０Ｂ～５１０Ｊなどのオーディオ符号化デバイスが、複数の球面調和係数または複数の球面調和係数の分解に

関してコンテンツ分析を実行することを可能にすることができる。次数低減を実行するとき、オーディオ符号化デバイス５１０Ｂ～５１０Ｊは、目標ビットレート５３５およびコンテンツ分析に基づいて、低減された球面調和係数または低減された球面調和係数の分解を生成するために、複数の球面調和係数または複数の球面調和係数の分解に関して次数低減を実行することができる。

【１１９３】

[1167]１つまたは複数のベクトルが除去されるとすると、オーディオ符号化デバイス５１０Ｂ～５１０Ｊは、ビットストリーム中のベクトルの数を制御データとして指定することができる。オーディオ符号化デバイス５１０Ｂ～５１０Ｊは、オーディオ復号デバイスによるビットストリームからのベクトルの抽出を容易にするために、このビットストリーム中のベクトルの数を指定することができる。

【１１９４】

[1168]図４４は、本開示において説明される技法の様々な態様による、量子化誤差を補償するためにオーディオ符号化デバイス４１０Ｄによって実行される例示的な動作を示す図である。図４４の例では、オーディオ符号化デバイス５１０Ｄの数学ユニット５２６は、数学演算がオーディオ複号デバイス５１０Ｄの数学ユニット５２６によって実行され得ることを示すために、破線ブロックとして示される。

【１１９５】

[1169]図４４の例に示されるように、数学ユニット５２６はまず、明瞭な球面調和係数（「H_{DIST}ベクトル６３０」として示される）を生成するために、U_{DIST}＊S_{DIST}ベクトル５２７をV^{T}_{DIST}ベクトル５２５Ｅと乗算することができる。数学ユニット５２６は次いで、H_{DIST}ベクトル６３０を、V^{T}_{DIST}ベクトル５２５Ｅの量子化されたバージョン（この場合も「$V^{T}_{Q_DIST}$ベクトル５２５Ｇ」として示される）で除算することができる。数学ユニット５２６は、$V^{T}_{Q_DIST}$ベクトル５２５Ｇの擬似逆ベクトルを決定し、次いで、H_{DIST}ベクトルを$V^{T}_{Q_DIST}$ベクトル５２５Ｇの擬似逆ベクトルと乗算し、U_{DIST}＊S_{DIST}（「US_{DIST}」または「US_{DIST}ベクトル」として省略され得る）の誤差補償されたバージョンを出力することによって、この除算を実行することができる。US_{DIST}の誤差補償されたバージョンは、図４４の例では、US^{*}_{DIST}ベクトル５２７’として示され得る。このようにして、本技法は、少なくとも一部、US_{DIST}ベクトル５２７に量子化誤差を実質的に射影し、US^{*}_{DIST}ベクトル５２７’を生成することができる。

【１１９６】

[1170]数学ユニット５２６は次いで、US_{ERR}ベクトル６３４（U_{DIST}＊S_{DIST}ベクトル５２７に射影された量子化が原因の誤差の少なくとも一部分を表し得る）を決定するために、U_{DIST}＊S_{DIST}ベクトル５２７からUS^{*}_{DIST}ベクトル５２７’を差し引くことができる。数学ユニット５２６は次いで、H_{ERR}ベクトル６３６を決定するために、US_{ERR}ベクトル６３４を$V^{T}_{Q_DIST}$ベクトル５２５Ｇと乗算することができる。数学的に、H_{ERR}ベクトル６３６は、US_{DIST}ベクトル５２７－US^{*}_{DIST}ベクトル５２７’に等価であり得、その結果が次いで、V^{T}_{DIST}ベクトル５２５Ｅと乗算される。数学ユニット５２６は次いで、U_{BG}ベクトル５２５ＤをS_{BG}ベクトル５２５Ｂと、次いでV^{T}_{BG}ベクトル５２５Ｆと乗算することによって計算された、バックグラウンド球面調和係数５３１（図４４の例では、H_{BG}ベクトル５３１として示される）に、H_{ERR}ベクトル６３６を加算することができる。数学ユニット５２６は、補償されたH_{BG}ベクトル５３１’を生成するために、H_{ERR}ベクトル６３６をH_{BG}ベクトル５３１に加算し、H_{BG}ベクトル５３１に量子化誤差の少なくとも一部分を実質的に射影することができる。このようにして、本技法は、H_{BG}ベクトル５３１に量子化誤差の少なくとも一部分を射影することができる。

【１１９７】

[1171]図４５および図４５Ｂは、本開示において説明される技法の様々な態様による、２つのフレームの部分からのサブフレームの補間を示す図である。図４５の例では、第１のフレーム６５０および第２のフレーム６５２が示される。第１のフレーム６５０は、Ｕ［１］行列、Ｓ［１］行列およびＶ’［１］行列に分解され得る球面調和係数（「ＳＨ［

１］」）を含み得る。第２のフレーム６５２は、球面調和係数（「ＳＨ［２］」）を含み得る。これらのＳＨ［１］およびＳＨ［２］は、上で説明されたＳＨＣ５１１の異なるフレームを識別することができる。

【１１９８】

[1172]図４５Ｂの例では、図４０Ｈの例に示されるオーディオ符号化デバイス５１０Ｈの分解ユニット５１８は、フレーム６５０および６５２の各々を４つのそれぞれのサブフレーム６５１Ａ～６５１Ｄおよび６５３Ａ～６５３Ｄに分離することができる。分解ユニット５１８は次いで、フレーム６５０の第１のサブフレーム６５１Ａ（「ＳＨ［１，１］」として示される）を、Ｕ［１，１］行列、Ｓ［１，１］行列およびＶ［１，１］行列に分解し、Ｖ［１，１］行列５１９’を補間ユニット５５０に出力することができる。分解ユニット５１８は次いで、フレーム６５２の第２のサブフレーム６５３Ａ（「ＳＨ［２，１］」として示される）を、Ｕ［１，１］行列、Ｓ［１，１］行列およびＶ［１，１］行列に分解し、Ｖ［２，１］行列５１９’を補間ユニット５５０に出力することができる。分解ユニット５１８はまた、ＳＨＣ１１のＳＨ［１，１］、ＳＨ［１，２］、ＳＨ［１，３］およびＳＨ［１，４］と、ＳＨＣ５１１のＳＨ［２，１］、ＳＨ［２，２］、ＳＨ［２，３］およびＳＨ［２，４］とを補間ユニット５５０に出力することもできる。

【１１９９】

[1173]補間ユニット５５０は次いで、図４５Ｂの例に示される例示の下部で識別される補間を実行することができる。すなわち、補間ユニット５５０は、Ｖ’［１，１］およびＶ’［２，１］に基づいてＶ’［１，２］を補間することができる。補間ユニット５５０はまた、Ｖ’［１，１］およびＶ’［２，１］に基づいてＶ’［１，３］を補間することもできる。さらに、補間ユニット５５０はまた、Ｖ’［１，１］およびＶ’［２，１］に基づいてＶ’［１，４］を補間することもできる。これらの補間は、図４６～図４６Ｅの例に示されるように、空間領域へのＶ’［１，１］およびＶ’［２，１］の射影と、後続する時間的補間および次いで球面調和関数領域に戻る射影とを伴い得る。

【１２００】

[1174]補間ユニット５５０は次に、ＳＨ［１，２］を（Ｖ’［１，２］）$^{-1}$と乗算することによってＵ［１，２］Ｓ［１，２］と、ＳＨ［１，３］を（Ｖ’［１，３］）$^{-1}$と乗算することによってＵ［１，３］Ｓ［１，３］と、ＳＨ［１，４］を（Ｖ’［１，４］）$^{-1}$と乗算することによってＵ［１，４］Ｓ［１，４］とを導出することができる。補間ユニット５５０は次いで、Ｖ行列５１９と、Ｓ行列５１９Ｂと、Ｕ行列５１９Ｃとを出力する、分解された形態においてフレームを形成し直すことができる。

【１２０１】

[1175]図４６Ａ～図４６Ｅは、本開示において説明される技法に従って補間された、複数の球面調和係数の分解されたバージョンの１つまたは複数のベクトルの射影の断面を示す図である。図４６Ａは、ＳＶＤプロセスを通じて第１のフレームからの第１のサブフレームのＳＨＣ５１１から分解された第１のＶ行列１９’の１つまたは複数の第１のベクトルの射影の断面を示す。図４６Ｂは、ＳＶＤプロセスを通じて第２のフレームからの第１のサブフレームのＳＨＣ５１１から分解された第２のＶ行列５１９’の１つまたは複数の第２のベクトルの射影の断面を示す。

【１２０２】

[1176]図４６Ｃは、第１のフレームからの第２のサブフレームを表すＶ行列５１９Ａのための１つまたは複数の補間されたベクトルの射影の断面を示し、これらのベクトルは、ＳＨＣ５１１の第１のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されており（すなわち、この例では、図４６の例に示されるＶ行列５１９’の１つまたは複数のベクトル）、および、ＳＨＣ５１１の第２のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されている（すなわち、この例では、図４６Ｂの例に示されるＶ行列５１９’の１つまたは複数のベクトル）。

【１２０３】

[1177]図４６Ｄは、第１のフレームからの第３のサブフレームを表すＶ行列５１９Ａのための１つまたは複数の補間されたベクトルの射影の断面を示し、これらのベクトルは、ＳＨＣ５１１の第１のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されており（すなわち、この例では、図４６の例に示されるＶ行列５１９’の１つまたは複数のベクトル）、および、ＳＨＣ５１１の第２のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されている（すなわち、この例では、図４６Ｂの例に示されるＶ行列５１９’の１つまたは複数のベクトル）。

【１２０４】

[1178]図４６Ｅは、第１のフレームからの第４のサブフレームを表すＶ行列５１９Ａのための１つまたは複数の補間されたベクトルの射影の断面を示し、これらのベクトルは、ＳＨＣ５１１の第１のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されており（すなわち、この例では、図４６の例に示されるＶ行列５１９’の１つまたは複数のベクトル）、および、ＳＨＣ５１１の第２のフレームの第１のサブフレームから分解されたＶ行列５１９’から、本開示で説明される技法に従って補間されている（すなわち、この例では、図４６Ｂの例に示されるＶ行列５１９’の１つまたは複数のベクトル）。

【１２０５】

[1179]図４７は、図４１～図４１Ｄの例に示されるオーディオ復号デバイス５４０Ａ～５４０Ｄの抽出ユニット５４２をより詳細に示すブロック図である。いくつかの例では、抽出ユニット５４２は、２つ以上の復号スキームを実行することができる、「統合された復号器」と呼ばれ得るものへのフロントエンドを表し得る（ここで、これらの２つ以上のスキームを実行することによって、復号器は、２つ以上のスキームを「統合する」と見なされ得る）。図４４の例に示されるように、抽出ユニット５４２は、マルチプレクサ６２０と、抽出サブユニット６２２Ａおよび６２２Ｂ（「抽出サブユニット６２２」）とを含む。マルチプレクサ６２０は、関連付けられる符号化されたフレーム分割されたＳＨＣ行列５４７～５４７Ｎが合成オーディオオブジェクトから生成されるか、録音から生成されるかの対応するインジケーションに基づいて、符号化されたフレーム分割されたＳＨＣ行列５４７～５４７Ｎのうちで、抽出サブユニット６２２Ａおよび抽出サブユニット６２２Ｂに送られるべきものを識別する。抽出サブユニット６２２Ａの各々は、いくつかの例では、合成オーディオオブジェクトから生成されたＳＨＣまたは録音から生成されたＳＨＣのいずれかに適合させられる、異なる復号（「解凍」と呼ばれ得る）スキームを実行することができる。抽出サブユニット６２２Ａの各々は、ＳＨＣ５４７に出力される、ＳＨＣ５４７のフレームを生成するために、これらの解凍スキームのうちのそれぞれの１つを実行することができる。

【１２０６】

[1180]たとえば、抽出ユニット６２２Ａは、次の公式

ＨＯＡ＝ＤｉｒＶ　ｘ　ＰＳ

を使用して、支配的な信号（ＰＳ）からＳＡを再構築するために、解凍スキームを実行することができ、

ただし、ＤｉｒＶは、サイドチャネルを通じて送信され得る、方向ベクトル（様々な方向と幅とを表す）である。抽出ユニット６２２Ｂは、この例では、次の公式

ＨＯＡ＝ｓｑｒｔ（４π）＊Ｙｎｍ（シータ，ファイ）＊ＰＳ

を使用して、ＰＳからＨＯＡ行列を再構築する解凍スキームを実行することができ、

ただし、Ｙｎｍは球面調和関数であり、シータおよびファイ情報は、サイドチャネルを通じて送られ得る。

【１２０７】

[1181]この点で、本技法は、抽出ユニット５３８が、音場を表す球面調和係数の圧縮されたバージョンが合成オーディオオブジェクトから生成されるかどうかのインジケーションに基づいて複数の解凍スキームの１つを選択し、複数の解凍スキームの選択された１つ

を使用して球面調和係数の圧縮されたバージョンを解凍することを可能にする。いくつかの例では、デバイスは、統合された復号器を備える。

【１２０８】

[1182]図４８は、図４１Ａ～図４１Ｄの例に示されるオーディオ復号デバイス５４０Ａ～５４０Ｄのオーディオレンダリングユニット４８をより詳細に示すブロック図である。図４８は、復元された球面調和係数５４７から、復号器ローカルスピーカーの幾何学的配置と互換性があるマルチチャネルオーディオデータ５４９Ａへの変換を示す。（この場合も、復号器におけるスピーカーの幾何学的配置を指すことがある）いくつかのローカルスピーカーの幾何学的配置では、可逆性を保証するいくつかの変換が、望ましいとは言えない音像品質を生じることがある。すなわち、音再現は、キャプチャされているオーディオと比較されたとき、常に音の正しい定位を生じるとは限らない。この望ましいとは言えない像品質を補正するために、本技法は、「仮想スピーカー」と呼ばれ得る概念を導入するようにさらに補強され得る。

【１２０９】

[1183]１つまたは複数のラウドスピーカーが、上で述べられたＩＴＵ－Ｒ　ＢＳ．７７５－１などの規格によって指定されたいくつかの角度トレランスを有する特定のまたは定義された空間領域中で再配置または配置されることを必要とするのではなく、上記のフレームワークは、ベクトルベースの振幅パンニング（ＶＢＡＰ）、距離ベースの振幅パンニング、または他の形態のパンニングなど、何らかの形態のパンニングを含むように修正され得る。説明のためにＶＢＡＰに焦点を当てると、ＶＢＡＰは、「仮想スピーカー」として特徴付けられ得るものを実質的に導入し得る。ＶＢＡＰは、１つまたは複数のラウドスピーカーが、仮想スピーカーをサポートする１つまたは複数のラウドスピーカーのロケーションおよび／または角度のうちの少なくとも１つと異なるロケーションおよび角度のうちの１つまたは複数において仮想スピーカーから発生するように思われる音を実質的に出力するように、これらの１つまたは複数のラウドスピーカーへのフィードを修正し得る。

【１２１０】

[1184]説明するために、ＳＨＣに関してラウドスピーカーフィードを決定するための次の式は、次のようになり得る。

【１２１１】

【数３６】

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

【１２１２】

[1185]上の式において、ＶＢＡＰ行列はＭ行×Ｎ列のサイズであり、ただし、Ｍはスピーカーの数を示し（上の式では５に等しくなり）、Ｎは仮想スピーカーの数を示す。ＶＢＡＰ行列は、聴取者の定義されたロケーションからスピーカーの位置の各々へのベクトルと、聴取者の定義されたロケーションから仮想スピーカーの位置の各々へのベクトルとの関数として計算され得る。上の式中のＤ行列はＮ行×（ｏｒｄｅｒ＋１）^{2}列のサイズであり得、ただし、ｏｒｄｅｒはＳＨ関数の次数を指し得る。Ｄ行列は次のものを表し得る。

【１２１３】

【数３７】

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

【１２１４】

[1186]ｇ行列（または、単一の列のみがあるとすると、ベクトル）は、復号器ローカル幾何学的配置において構成されたスピーカーのためのスピーカーフィードについての利得を表し得る。この式では、ｇ行列はＭのサイズである。Ａ行列（または、単一の列のみがあるとすると、ベクトル）は、ＳＨＣ５２０を示すことができ、（Ｏｒｄｅｒ＋１）^{2}とも示され得る、（Ｏｒｄｅｒ＋１）（Ｏｒｄｅｒ＋１）のサイズである。

【１２１５】

[1187]事実上、ＶＢＡＰ行列は、スピーカーのロケーションと仮想スピーカーの位置とを考慮する「利得調整」と呼ばれ得るものを提供するＭ×Ｎ行列である。このようにしてパンニングを導入することによって、ローカルスピーカーの幾何学的配置によって再現されたとき、より良質の像を生じるマルチチャネルオーディオのより良い再現がもたらされ得る。その上、この式にＶＢＡＰを組み込むことによって、本技法は、様々な規格において指定されているものとはアラインしないプアなスピーカーの幾何学的配置を克服し得る。

【１２１６】

[1188]実際には、この式は、反転させられ、ＳＨＣを、この場合も本開示で復号器ローカル幾何学的配置と呼ばれることがあるラウドスピーカーの特定の幾何学的配置または構成のためのマルチチャネルフィードに変換し戻すために利用され得る。すなわち、この式は、ｇ行列について解くために反転させられ得る。反転させられた式は次のようになり得る。

【１２１７】

【数３８】

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

【１２１８】

[1189]ｇ行列は、この例では、５．１スピーカー構成における５つのラウドスピーカーの各々についてのスピーカー利得を表し得る。この構成において使用される仮想スピーカーロケーションは、５．１マルチチャネルフォーマット仕様または規格において定義されているロケーションに対応し得る。これらの仮想スピーカーの各々をサポートし得るラウドスピーカーのロケーションは、任意の数の知られているオーディオ定位技法を使用して決定され得、それらの多くは、（オーディオ／ビデオ受信機（Ａ／Ｖ受信機）、テレビジョン、ゲームシステム、デジタルビデオディスクシステム、または他のタイプのヘッドエ

ンドシステムなどの）ヘッドエンドユニットに対して各ラウドスピーカーのロケーションを決定するために特定の周波数を有するトーンを再生することを伴う。代替的に、ヘッドエンドユニットのユーザが、ラウドスピーカーの各々のロケーションを手動で指定し得る。いずれの場合も、これらの知られているロケーションと考えられる角度とを鑑みて、ヘッドエンドユニットは、利得について解き、ＶＢＡＰを介して仮想ラウドスピーカーの理想的な構成を仮定し得る。

【１２１９】

[1190]この点で、デバイスまたは装置は、異なるローカル幾何学的配置において構成された仮想スピーカーから発生するように思われる音を発するために、復号器ローカル幾何学的配置におけるスピーカーを駆動する複数のチャネルを生成するために、複数の仮想チャネルにおいて、ベクトルベースの振幅パンニングまたは他の形態のパンニングを実行することができる。したがって、本技法は、オーディオ復号デバイス４０が、複数のチャネルを生成するために、復元された球面調和係数４７など、複数の球面調和係数において変換を実行することを可能にすることができる。複数のチャネルの各々は、対応する異なる空間領域と関連付けられ得る。その上、複数のチャネルの各々は、複数の仮想チャネルを備えてよく、ここで、複数の仮想チャネルは、対応する異なる空間領域と関連付けられ得る。したがって、デバイスは、マルチチャネルオーディオデータ４９の複数のチャネルを生成するために、仮想チャネルにおいてベクトルベースの振幅パンニングを実行することができる。

【１２２０】

[1191]図４９Ａ～図４９Ｅ（ｉｉ）は、本開示において説明される技法の様々な態様を実装し得る、それぞれのオーディオコーディングシステム５６０Ａ～５６０Ｃ、５６７Ｄ、５６９Ｄ、５７１Ｅおよび５７３Ｅを示す図である。図４９Ａの例に示されるように、オーディオコーディングシステム５６０Ａは、オーディオ符号化デバイス５６２とオーディオ復号デバイス５６４とを含み得る。オーディオ符号化デバイス５６２は、それぞれ図４および図４０Ａ～図４０Ｄの例に示されるオーディオ符号化デバイス２０および５１０Ａ～５１０Ｄのうちの任意のものと類似であり得る。オーディオ複号デバイス５６４は、図５および図４１の例に示されるオーディオ複号デバイス２４および４０と類似であり得る。

【１２２１】

[1192]上で説明されたように、高次アンビソニックス（ＨＯＡ）は、空間フーリエ変換に基づいて音場のすべての指向性情報を記述するための方法である。いくつかの例では、アンビソニックス次数Ｎが高いほど、空間分解能は高く、球面調和（ＳＨ）係数の数（Ｎ＋１）^{2}は大きくなる。したがって、アンビソニックス次数Ｎが高いほど、いくつかの例では、係数を送信し、記憶するためにより大きい帯域幅要件をもたらす。ＨＯＡの帯域幅要件は、たとえば、５．１または７．１サラウンド音響オーディオデータと比較して、むしろ高いので、帯域幅低減が、多くの適用形態にとって望まれ得る。

【１２２２】

[1193]本開示で説明される技法によれば、オーディオコーディングシステム５６０Ａは、空間音響シーンにおける不明瞭な（バックグラウンドまたは環境）要素から明瞭な（フォアグラウンド）要素を分離することに基づく方法を実行することができる。この分離は、オーディオコーディングシステム５６０Ａが、フォアグラウンド要素とバックグラウンド要素とを互いとは無関係に処理することを可能にし得る。この例では、オーディオコーディングシステム５６０Ａは、フォアグラウンド要素が（聴取者による）より多くの注意を引くことができ、バックグラウンド要素と比較して（この場合も聴取者によって）定位することがより容易であり得るという、特性を活用する。結果として、オーディオコーディングシステム５６０Ａは、より効率的にＨＯＡコンテンツを記憶または送信することができる。

【１２２３】

[1194]いくつかの例では、オーディオコーディングシステム５６０Ａは、特異値分解（

ＳＶＤ）プロセスを利用することによって、この分離を達成することができる。ＳＶＤプロセスは、ＨＯＡ係数のフレームを３つの行列（Ｕ、Ｓ、Ｖ）に分離することができる。行列Ｕは左特異ベクトルを含み、Ｖ行列は右特異ベクトルを含む。対角行列Ｓは、その対角線において非負のソートされた特異値を含む。ＨＯＡ係数の一般に良い（または、いくつかの例では、ＨＯＡ係数を表すことにおいて無限の精度を仮定すると、完全な）再構築は、Ｕ＊Ｓ＊Ｖ’によって与えられるようになる。Ｄ個の最大の特異値Ｕ（：，１：Ｄ）＊Ｓ（１：Ｄ，：）＊Ｖ’をもつサブ空間を再構築することのみによって、オーディオコーディングシステム５６０Ａは、このＨＯＡフレームから最も顕著な空間情報、すなわち、フォアグラウンド音声要素（と、場合によっては、いくつかの強い初期の室内反射と）を抽出することができる。残りのＵ（：，Ｄ＋１：終了）＊Ｓ（Ｄ＋１：終了，：）＊Ｖ’は、コンテンツからバックグラウンド要素と残響とを再構築することができる。

【１２２４】

[1195]オーディオコーディングシステム５６０Ａは、Ｓの降順の対角線上の値によって作成される曲線の傾きを分析することによって、２つのサブ空間を分離する値Ｄを決定することができ、大きい特異値はフォアグラウンド音声を表し、小さい特異値はバックグラウンド値を表す。オーディオコーディングシステム５６０Ａは、特異値曲線の一次導関数と二次導関数とを使用することができる。オーディオコーディングシステム５６０Ａはまた、数Ｄを１と５との間になるように限定することもできる。代替的に、オーディオコーディングシステム５６０Ａは、４という値などへと数Ｄを事前に定めることができる。いずれにしても、数Ｄが推定されると、オーディオコーディングシステム５６０Ａは、行列ＵおよびＳからフォアグラウンドサブ空間とバックグラウンドサブ空間とを抽出する。

【１２２５】

[1196]オーディオコーディングシステム５６０Ａは次いで、Ｕ（：，Ｄ＋１：終了）＊Ｓ（Ｄ＋１：終了，：）＊Ｖ’を介してバックグラウンドシーンのＨＯＡ係数を再構築し、ＨＯＡ係数の$(N+1)^2$チャネルをもたらすことができる。バックグラウンド要素は、いくつかの例では、フォアグラウンド要素と比較してそれほど顕著ではなく、それほど定位可能ではないことが知られているので、オーディオコーディングシステム５６０Ａは、ＨＯＡチャネルの次数を切り捨てることができる。さらに、オーディオコーディングシステム５６０Ａは、ＡＡＣなど、有損失または無損失のオーディオコーデックを用いて、または任意選択で、顕著なフォアグラウンド要素を圧縮するために使用されるものと比較して、よりアグレッシブなオーディオコーデックを用いて、これらのチャネルを圧縮することができる。いくつかの例では、帯域幅を節約するために、オーディオコーディングシステム５６０Ａは、フォアグラウンド要素を異なるように送信することができる。すなわち、オーディオコーディングシステムは、有損失または無損失のオーディオコーデック（ＡＡＣなど）を用いて圧縮された後に左特異ベクトルＵ（：，１：Ｄ）を送信し、再構築行列Ｒ＝Ｓ（１：Ｄ，：）＊Ｖ’とともにこれらの圧縮された左特異値を送信することができる。Ｒは、フレームにわたって異なり得る、Ｄ×$(N+1)^2$行列を表し得る。

【１２２６】

[1197]オーディオコーディングシステム５６０の受信機側で、オーディオコーディングシステムは、$(N+1)^2$ＨＯＡチャネルのフレームを再構築するために、これらの２つの行列を乗算することができる。バックグラウンドＨＯＡチャネルおよびフォアグラウンドＨＯＡチャネルが一緒に加算されると、オーディオコーディングシステム５６０Ａは、任意の適切なアンビソニックスレンダラを使用して、任意のラウドスピーカーセットアップにレンダリングすることができる。本技法はバックグラウンド要素からのフォアグラウンド要素（直接または明瞭な音声）の分離を提供するので、聴覚障害者は、了解度を増すために、フォアグラウンドからバックグラウンド要素のミキシングを制御することができる。また、たとえば、フォアグラウンド要素のみにおけるダイナミックコンプレッサーなど、他のオーディオ効果もまた適用可能であり得る。

【１２２７】

[1198]図４９Ｂは、オーディオ符号化システム５６０Ｂをより詳細に示すブロック図で

ある。図４９Ｂの例に示されるように、オーディオコーディングシステム５６０Ｂは、オーディオ符号化デバイス５６６とオーディオ復号デバイス５６８とを含み得る。オーディオ符号化デバイス５６６は、図４および図４０Ｅの例に示されるオーディオ符号化デバイス２４および５１０Ｅと類似であり得る。オーディオ複号デバイス５６８は、図５および図４１Ｂの例に示されるオーディオ複号デバイス２４および５４０Ｂと類似であり得る。

【１２２８】

[1199]本開示で説明される技法に従って、ＨｏＡ信号におけるフレームベースのＳＶＤ（または、ＫＬＴ＆ＰＣＡなどの関連方法）分解を使用するとき、帯域幅低減の目的で、オーディオ符号化デバイス６６は、（Ｓ行列の対応する特異値と乗算された）Ｕ行列の最初のいくつかのベクトル、ならびに、V^Tベクトルの対応するベクトルを量子化することができる。これは、音場の「フォアグラウンド」成分を備えるようになる。本技法は、オーディオ符号化デバイス５６６が、「ブラックボックス」オーディオコーディングエンジンを使用して、U_{DIST}＊S_{DIST}ベクトルをコーディングすることを可能にすることができる。Ｖベクトルには、スカラー量子化またはベクトル量子化のいずれかが行われ得る。加えて、Ｕ行列における残りのベクトルの一部または全部は、Ｓ行列およびＶ行列の対応する特異値と乗算され、また「ブラックボックス」オーディオコーディングエンジンを使用してコーディングされ得る。これらは、音場の「バックグラウンド」成分を備えるようになる。

【１２２９】

[1200]最大の聴覚的成分が「フォアグラウンド成分」に分解されるので、オーディオ符号化デバイス５６６は、バックグラウンドが重要な定位可能なコンテンツを含まない（と我々が仮定する）ので、「ブラックボックス」オーディオコーディングエンジンを使用する前に「バックグラウンド」成分のアンビソニックス次数を低減することができる。フォアグラウンド成分のアンビソニックス次数に応じて、オーディオ符号化ユニット５６６は、むしろ大きくなり得る、対応するＶベクトル（複数可）を送信することができる。たとえば、Ｖベクトルの単純な１６ビットスカラー量子化は、フォアグラウンド成分ごとに、４次（２５個の係数）ではほぼ２０ｋｂｐｓオーバーヘッドと、６次（４９個の係数）では４０ｋｂｐｓとをもたらすようになる。本開示で説明される技法は、Ｖベクトルのこのオーバーヘッドを低減するための方法を提供することができる。

【１２３０】

[1201]説明するために、上で説明されたように、フォアグラウンド要素のアンビソニックス次数がN_{DIST}であることと、バックグラウンド要素のアンビソニックス次数N_{BG}とを仮定する。オーディオ符号化デバイス５６６は、上で説明されたように、バックグラウンド要素のアンビソニックス次数を低減することができるので、N_{BG}はN_{DIST}未満であり得る。受信機側でフォアグラウンド要素を再構築するために送信される必要があるフォアグラウンドＶベクトルの長さは、フォアグラウンド要素ごとに$(N_{DIST}+1)^2$の長さを有するのに対して、最初の$((N_{DIST}+1)^2)-((N_{BG}+1)^2)$個の係数は、次数$N_{BG}$までのフォアグラウンドまたは明瞭な成分を再構築するために使用され得る。本開示で説明される技法を使用して、オーディオ符号化デバイス５６６は、次数N_{BG}までのフォアグラウンドを再構築し、得られた$(N_{BG}+1)^2$個のチャネルをバックグラウンドチャネルとマージし、次数N_{BG}までの完全な音場をもたらすことができる。オーディオ符号化デバイス５６６は次いで、Ｖベクトルを、送信のために、$(N_{BG}+1)^2$よりも高いインデックスをもつそれらの係数まで低減することができる（ここで、これらのベクトルは「V^T_{SMALL}」と呼ばれ得る）。受信機側で、オーディオ復号ユニット５６８は、フォアグラウンド要素をV^T_{SMALL}ベクトルと乗算することによって、N_{BG}よりも大きいアンビソニックス次数のためのフォアグラウンドオーディオチャネルを再構築することができる。

【１２３１】

[1202]図４９Ｃは、オーディオ符号化システム５６０Ｃをより詳細に示すブロック図である。図４９Ｃの例に示されるように、オーディオコーディングシステム５６０Ｂは、オーディオ符号化デバイス５６７とオーディオ復号デバイス５６９とを含み得る。オーディ

オ符号化デバイス５６７は、図４および図４０Ｆの例に示されるオーディオ符号化デバイス２０および５１０Ｆと類似であり得る。オーディオ複号デバイス５６９は、図５および図４１Ｂの例に示されるオーディオ複号デバイス２４および５４０Ｂと類似であり得る。

【１２３２】

[1203]本開示で説明される技法に従って、ＨｏＡ信号におけるフレームベースのＳＶＤ（または、ＫＬＴ＆ＰＣＡなどの関連方法）分解を使用するとき、帯域幅低減の目的で、オーディオ符号化デバイス５６７は、（Ｓ行列の対応する特異値と乗算された）Ｕ行列の最初のいくつかのベクトル、ならびに、V^Tベクトルの対応するベクトルを量子化することができる。これは、音場の「フォアグラウンド」成分を備えるようになる。本技法は、オーディオ符号化デバイス５６７が、「ブラックボックス」オーディオコーディングエンジンを使用して、$U_{DIST} * S_{DIST}$ベクトルをコーディングすることを可能にすることができる。Ｖベクトルには、スカラー量子化またはベクトル量子化のいずれかが行われ得る。加えて、Ｕ行列における残りのベクトルの一部または全部は、Ｓ行列およびＶ行列の対応する特異値と乗算され、また「ブラックボックス」オーディオコーディングエンジンを使用してコーディングされ得る。これらは、音場の「バックグラウンド」成分を備えるようになる。

【１２３３】

[1204]最大の聴覚的成分が「フォアグラウンド成分」に分解されるので、オーディオ符号化デバイス５６７は、バックグラウンドが重要な定位可能なコンテンツを含まない（と我々が仮定する）ので、「ブラックボックス」オーディオコーディングエンジンを使用する前に「バックグラウンド」成分のアンビソニックス次数を低減することができる。オーディオ符号化デバイス５６７は、本明細書で説明される技法に従って、音場の全体的なエネルギーを保存するような方法で、次数を低減することができる。フォアグラウンド成分のアンビソニックス次数に応じて、オーディオ符号化ユニット５６７は、むしろ大きくなり得る、対応するＶベクトル（複数可）を送信することができる。たとえば、Ｖベクトルの単純な１６ビットスカラー量子化は、フォアグラウンド成分ごとに、４次（２５個の係数）ではほぼ２０ｋｂｐｓオーバーヘッドと、６次（４９個の係数）では４０ｋｂｐｓとをもたらすようになる。本開示で説明される技法は、Ｖベクトル（複数可）のこのオーバーヘッドを低減するための方法を提供することができる。

【１２３４】

[1205]説明するために、フォアグラウンド要素およびバックグラウンド要素のアンビソニックス次数がＮであると仮定する。オーディオ符号化デバイス５６７は、Ｖベクトル（複数可）のバックグラウンド要素のアンビソニックス次数をＮから

【１２３５】

【数３９】

$$\tilde{\eta}$$

【１２３６】

まで低減し、

【１２３７】

【数４０】

$$\tilde{\eta} < N$$

【１２３８】

であるようにすることができる。オーディオ符号化デバイス６７は、ＳＨＣによって記述される音場の全体的なエネルギーを保存するために、Ｖベクトル（複数可）のバックグラウンド要素の値を増すために、補償をさらに適用する。補償を適用するための例示的な技法は、図４０Ｆに関して上で説明されている。受信機側で、オーディオ復号ユニット５６

９は、アンビソニックス次数のためのバックグラウンドオーディオチャネルを再構築することができる。

【１２３９】

[1206]図４９Ｄ（ｉ）および図４９Ｄ（ｉｉ）は、それぞれオーディオ符号化デバイス５６７Ｄおよびオーディオ復号デバイス５６９Ｄを示す。オーディオ符号化デバイス５６７Ｄおよびオーディオ復号デバイス５６９Ｄは、本開示の態様に従って、１つまたは複数の指向性ベースの明瞭性決定を実行するように構成され得る。高次アンビソニックス（ＨＯＡ）は、空間フーリエ変換に基づいて音場のすべての指向性情報を記述するための方法である。アンビソニックス次数Ｎが高いほど、空間分解能は高くなり、球面調和（ＳＨ）係数の数（Ｎ＋１）＾２は大きくなり、データを送信し、記憶するために必要とされる帯域幅はより大きくなる。ＨＯＡの帯域幅要件がむしろ高いので、多くの適用形態では、帯域幅低減が望まれる。

【１２４０】

[1207]以前の説明は、ＳＶＤ（特異値分解）または関連プロセスが空間オーディオ圧縮のためにどのように使用され得るかについて説明した。本明細書で説明される技法は、フォアグラウンド要素とも呼ばれる顕著な要素を選択するための改善されたアルゴリズムを提示する。ＨＯＡオーディオフレームの、そのＵ、Ｓ、およびＶ行列へのＳＶＤベースの分解の後、本技法は、Ｋ個の顕著な要素の選択を、Ｕ行列の最初のＫ個のチャネル［Ｕ（：，１：Ｋ）＊Ｓ（１：Ｋ，１：Ｋ）］のみに基づかせる。これは、最高のエネルギーをもつオーディオ要素を選択することをもたらす。ただし、それらの要素もまた指向性があることは保証されない。したがって、本技法は、高いエネルギーを有し、指向性もある音声要素を発見することを対象とする。これは、潜在的には、Ｓ行列を用いてＶ行列を重み付けることによって達成される。次いで、この得られた行列の行ごとに、より高くインデックス付けられた要素（高次ＨＯＡ係数と関連付けられる）が二乗および加算され、１行につき１つの値が得られる［図４０Ｈに関して説明された擬似コードにおけるｓｕｍＶＳ］。擬似コードにおいて表されたワークフローに従って、５番目のインデックスで開始する高次アンビソニックス係数が考えられる。これらの値は、それらのサイズに従ってソートされ、ソートインデックスは、オリジナルのＵ、Ｓ、およびＶ行列をそれに応じて再構成するために使用される。本開示で以前に説明されたＳＶＤベースの圧縮アルゴリズムが次いで、さらなる修正なしに適用され得る。

【１２４１】

[1208]図４９Ｅ（ｉ）および図４９Ｅ（ｉｉ）は、それぞれオーディオ符号化デバイス５７１Ｅおよびオーディオ復号デバイス５７３Ｅを示すブロック図である。オーディオ符号化デバイス５７１Ｅおよびオーディオ復号デバイス５７３Ｅは、オーディオ符号化デバイス５７１Ｅが、S^2行列とＶ行列とを生成するために、ＨＯＡ係数の電力スペクトル密度行列（ＰＤＳ）に関して特異値分解を実行することができることを除いて、図４９～図４９Ｄ（ｉｉ）の例に関して上で説明された技法の様々な態様を実行することができる。S^2行列は、二乗されたＳ行列を示すことができ、すなわちS^2行列は、Ｓ行列を取得するために平方根演算を受けることができる。オーディオ符号化デバイス５７１Ｅは、いくつかの例では、量子化されたＶ行列（Ｖ’行列と示され得る）を取得するために、Ｖ行列に関して量子化を実行することができる。

【１２４２】

[1209]オーディオ符号化デバイス５７１Ｅは、ＳＶ’行列を生成するために、Ｓ行列を量子化されたＶ’行列とまず乗算することによって、Ｕ行列を取得することができる。オーディオ符号化デバイス５７１Ｅは次に、ＳＶ’行列の擬似逆行列を取得することができ、次いで、Ｕ行列を取得するためにＳＶ’行列の擬似逆行列とＨＯＡ係数を乗算することができる。ＨＯＡ係数自体ではなくＨＯＡ係数の電力スペクトル密度に関してＳＶＤを実行することによって、オーディオ符号化デバイス５７１Ｅは場合によっては、プロセッササイクルと記憶空間の１つまたは複数に関してＳＶＤを実行することの計算的な複雑さを低減しつつ、ＳＶＤがＨＯＡ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の符

号化効率を達成することができる。

【１２４３】

[1210]オーディオ複号デバイス５７３Ｅは、オーディオ複号デバイス５７３Ｅが、直接的にＨＯＡ係数ではなくＨＯＡ係数の電力スペクトル密度へのＳＶＤの適用を通じて達成されたＨＯＡ係数の分解からＨＯＡ係数を再構築することができることを除いて、上で説明されたそれらのオーディオ復号デバイスと同様であり得る。

【１２４４】

[1211]図５０Ａおよび図５０Ｂは、本開示において説明される技法による、バックグラウンドコンテンツの次数を潜在的に低減するための２つの異なる手法の１つを各々が示すブロック図である。図５０の例に示されるように、第１の手法は、次数をＮから

【１２４５】

【数４１】

$$\tilde{\eta}$$

【１２４６】

まで低減するために、$U_{BG}*S_{BG}*V^{T}$ベクトルに関して次数低減を利用することができ、ただし、

【１２４７】

【数４２】

$$\tilde{\eta}$$

【１２４８】

は（＜）Ｎ未満である。すなわち、図４０Ｂ～図４０Ｊの例に示される次数低減ユニット５２８Ａは、$U_{BG}*S_{BG}*V^{T}$ベクトルの次数Ｎを

【１２４９】

【数４３】

$$\tilde{\eta}$$

【１２５０】

に切り捨てるかまたは別様に低減するために、次数低減を実行することができ、ただし、

【１２５１】

【数４４】

$$\tilde{\eta}$$

【１２５２】

は（＜）Ｎ未満である。

【１２５３】

[1212]代替手法として、次数低減ユニット５２８Ａは、図５０Ｂの例に示されるように、V^{T}に関してこの切り捨てを実行し、行を

【１２５４】

【数４５】

$$(\tilde{\eta}+1)^{2}$$

【１２５５】

になるように排除することができ、
これは、説明を簡単にするために、図４０Ｂの例に示されない。言い換えれば、次数低減ユニット５２８Ａは、V_{BG}行列を実質的に生成するために、V^T行列の１つまたは複数の次数を除去することができる。このV_{BG}行列のサイズは、
【１２５６】
【数４６】

$$(\tilde{\eta}+1)^2 \times (N+1)^2 - D$$

【１２５７】
であり、ここで、このV_{BG}行列は次いで、U_{BG}＊S_{BG}＊V^Tベクトルを生成するときにV^T行列の代わりに使用され、
【１２５８】
【数４７】

$$M \times (\tilde{\eta}+1)^2$$

【１２５９】
のサイズのU_{BG}＊S_{BG}＊V^Tベクトルを生成するために、切り捨てが実質的に実行される。
【１２６０】
[1213]図５１は、球面調和係数７０１を圧縮するために本開示で説明された技法の様々な態様を実施し得るオーディオ符号化デバイス７００Ａの明瞭な成分の圧縮経路の例を示すブロック図である。図５１の例では、明瞭な成分の圧縮経路は、ＳＨＣ７０１によって表される音場の明瞭な成分を圧縮する、オーディオ符号化デバイス７００Ａの処理経路を指し得る。バックグラウンド成分圧縮経路と呼ばれ得る、別の経路は、ＳＨＣ７０１のバックグラウンド成分を圧縮する、オーディオ符号化デバイス７００Ａの処理経路を表し得る。
【１２６１】
[1214]説明を簡単にするために示されていないが、バックグラウンド成分圧縮経路は、ＳＨＣ７０１の分解ではなく、直接的にＳＨＣ７０１に関して動作することができる。これは、U_{BG}行列、S_{BG}行列およびV_{BG}行列からバックグラウンド成分を再合成し、次いでこれらの再合成されたバックグラウンド成分のある形態の（たとえば、ＡＡＣ符号化器を使用する）音響心理学的符号化を実行するのではなく、バックグラウンド成分処理経路が、（図４の例に示されたオーディオ符号化デバイス２０に関して上で説明されたように）直接的にＳＨＣ７０１に関して動作し、音響心理学的符号化器を使用して、これらのバックグラウンド成分を圧縮することができることを除いて、図４９～図４９Ｃに関して上で説明されたものと類似である。直接的にＳＨＣ７０１に関して音響心理学的符号化を実行することによって、不連続性が低減され、一方でまた、再合成されたバックグラウンド成分に関して音響心理学的符号化を実行することと比較して、（バックグラウンド成分を圧縮するために必要とされる動作に関する）計算的な複雑さを低減することもできる。明瞭およびバックグラウンドに関して言及されるが、本開示では「目立つ」という用語が「明瞭」の代わりに使用されることがあり、「環境的」という用語が「バックグラウンド」の代わりに使用されることがある。
【１２６２】
[1215]いずれにしても、球面調和係数７０１（「ＳＨＣ７０１」）は、Ｍ×（Ｎ＋１）^{2}のサイズを有する係数の行列を備えることができ、ここで、Ｍは、オーディオフレームにおけるサンプルの数を示し（および、いくつかの例では１０２４であり）、Ｎは、係数が対応する基底関数の最高次数を示す。上で述べられたように、Ｎは一般に、全体で１０２４×２５個の係数では４に設定される。特定の次数と位数の組合せに対応するＳＨＣ７

０１の各々は、チャネルと呼ばれ得る。たとえば、１次数、０位数の基底関数に対応するＭ個のサンプル係数のすべては、あるチャネルを表し得るが、０次数、０位数の基底関数に対応する係数は、別のチャネルを表し得る、などとなる。ＳＨＣ７０１はまた、本開示では、高次アンビソニックス（ＨＯＡ）コンテンツ７０１またはＳＨ信号７０１と呼ばれることもある。

【１２６３】

[1216]図５１の例に示されるように、オーディオ符号化デバイス７００Ａは、分析ユニット７０２と、ベクトルベースの合成ユニット７０４と、ベクトル低減ユニット７０６と、音響心理学的符号化ユニット７０８と、係数低減ユニット７１０と、圧縮ユニット７１２（「ｃｏｍｐｒユニット７１２」）とを含む。分析ユニット７０２は、音場の明瞭な成分（Ｄ）７０３と、バックグラウンド成分の総数（$ＢＧ_{TOT}$）７０５とを識別するように、ＳＨＣ７０１に関して分析を実行するように構成されるユニットを表し得る。上で説明されたオーディオ符号化デバイスと比較して、オーディオ符号化デバイス７００Ａは、ＳＨＣ７０１の分解に関してこの決定を実行するのではなく、ＳＨＣ７０１に関して直接的に実行する。

【１２６４】

[1217]ベクトルベースの合成ユニット７０４は、ＳＶＤの例では、Ｍ×（Ｎ＋１）^{2}のサイズを有する［ＵＳ］行列７０７と、（Ｎ＋１）^{2}×（Ｎ＋１）^{2}のサイズを有する［Ｖ］行列７０９とを生成するために、ＳＶＤ、ＫＬＴ、ＰＣＡまたは任意の他のベクトルベースの合成など、ＳＨＣ７０１に関してある形態のベクトルベースの合成を実行するように構成されるユニットを表す。［ＵＳ］行列７０７は、ＳＨＣ７０１へのＳＶＤの適用を通じて生成された［Ｕ］行列および［Ｓ］行列の行列乗算から得られる行列を表し得る。

【１２６５】

[1218]ベクトル低減ユニット７０６は、［ＵＳ}行列７０７および［Ｖ］行列７０９の残りのベクトルの各々が音場の明瞭または目立つ成分を識別するように、［ＵＳ］行列７０７および［Ｖ］行列７０９のベクトルの数を低減するように構成されるユニットを表し得る。ベクトル低減ユニット７０６は、明瞭な成分の数Ｄ７０３に基づいて、この低減を実行することができる。明瞭な成分の数Ｄ７０３は、実際には、数のアレイを表すことができ、ここで、各数は、行列７０７および７０９の異なる明瞭なベクトルを識別する。ベクトル低減ユニット７０６は、Ｍ×Ｄのサイズの低減された［ＵＳ］行列７１１と、（Ｎ＋１）^{2}×Ｄのサイズの低減された［Ｖ］行列７１３とを出力することができる。

【１２６６】

[1219]説明を簡単にするために示されていないが、［Ｖ］行列７０９の補間は、上でより詳細に説明されたものに類似した方式で、［Ｖ］行列７０９の低減の前に行われ得る。その上、説明を簡単にするために示されていないが、上でより詳細に説明された方式で、低減された［ＵＳ］行列７１１および／または低減された［Ｖ］行列７１２の再順序付けである。したがって、本技法は、これらおよび他の点（誤差射影、または上で説明されているが図５１の例に示されていない前述の技法の任意の他の態様など）において限定されるべきではない。

【１２６７】

[1220]音響心理学的符号化ユニット７０８は、ビットストリーム７１５を生成するために、［ＵＳ］行列７１１に関して音響心理学的符号化を実行するように構成されるユニットを表す。係数低減ユニット７１０は、低減された［Ｖ］行列７１３のチャネルの数を低減するように構成されるユニットを表し得る。言い換えれば、係数低減ユニット７１０は、（低減された［Ｖ］行列７１３を形成する）明瞭なＶベクトルの係数のうちで指向性情報をほとんどまたはまったく有しないものを排除するように構成されるユニットを表し得る。上で説明されたように、いくつかの例では、明瞭なＶベクトルの係数のうちで、１次または０次の基底関数に対応するもの（上で$Ｎ_{BG}$と示される）は、指向性情報をほとんど提供せず、したがって、明瞭なＶベクトルから（上で「次数低減」と呼ばれるものを通じて）除去され得る。この例では、対応するこれらの係数$Ｎ_{BG}$を識別するだけではなく、追

加のＨＯＡチャネル（変数ＴｏｔａｌＯｆＡｄｄＡｍｂＨＯＡＣｈａｎによって示され得る）を［（N_{BG}＋１）^{2}＋１，（Ｎ＋１）^{2}］のセットから識別するために、より大きい柔軟性が与えられ得る。分析ユニット７０２は、（N_{BG}＋１）^{2}のみではなく、ＴｏｔａｌＯｆＡｄｄＡｍｂＨＯＡＣｈａｎをも識別し得る、BG_{TOT}を決定するために、ＳＨＣ７０１を分析することができる。係数低減ユニット７１０は次いで、（（Ｎ＋１）^{2}－（BG_{TOT}）×Ｄのサイズの小さい［Ｖ］行列７１７を生成するために、低減された［Ｖ］行列７１３から、係数のうちで（N_{BG}＋１）^{2}およびＴｏｔａｌＯｆＡｄｄＡｍｂＨＯＡＣｈａｎに対応するものを除去することができる。

【１２６８】

[1221]圧縮ユニット７１２は次いで、小さい［Ｖ］行列７１７を圧縮するために、上で述べられたスカラー量子化および／またはハフマン符号化を実行し、サイドチャネル情報７１９（「サイドチャネルｉｎｆｏ７１９」）として、圧縮された小さい［Ｖ］行列７１７を出力することができる。圧縮ユニット７１２は、図１０～図１００（ｉｉ）の例に示されるものに類似した方式で、サイドチャネル情報７１９を出力することができる。いくつかの例では、上で説明されたものに類似したビットストリーム生成ユニットは、サイドチャネル情報７１９をビットストリーム７１５に組み込むことができる。その上、ビットストリーム７１５と呼ばれるが、オーディオ符号化デバイス７００Ａは、上で述べられたように、別のビットストリームをもたらすバックグラウンド成分処理経路を含むことができ、ここで、上で説明されたものに類似したビットストリーム生成ユニットは、ビットストリーム７１５と、バックグラウンド成分処理経路によって出力されたビットストリームとを含む、上で説明されたビットストリーム１７に類似したビットストリームを生成することができる。

【１２６９】

[1222]本開示で説明される技法に従って、分析ユニット７０２は、ベクトル、すなわち、この例では、音場の明瞭な成分を表すために使用されるべき、低減された［Ｖ］行列７１３のベクトルの、係数の第１の０ではないセットを決定するように構成され得る。いくつかの例では、分析ユニット７０２は、低減された［Ｖ］行列７１３を形成するあらゆるベクトルの係数のすべてがサイドチャネル情報７１９に含まれるべきであると決定することができる。したがって、分析ユニット７０２は、BG_{TOT}を０に等しく設定することができる。

【１２７０】

[1223]したがって、オーディオ符号化デバイス７００Ａは実質的に、「複号されたベクトル」として示された表に関して上で説明されたものとは逆の方式で作用することができる。加えて、オーディオ符号化デバイス７００Ａは、アクセスユニット（１つまたは複数のフレームを含み得る）のヘッダ中のシンタックス要素、複数の構成モードのいずれが選択されたか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、分析ユニット７０２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）この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、このシンタックス要素は、この明瞭な成分の指向性の様相を表すために、４つの構成モードのいずれが低減された［Ｖ］行列７１３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、「ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈ」として示され得る。このようにして、オーディオ符号化デバイス７００Ａは、４つの構成モードのいずれが小さい［Ｖ］行列７１７をビットストリーム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、または別様に指定することができる。４つの構成モードに関して説明されるが、本技法は４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構成モードを含む、任意の数の構成モードに限定されるべきである。

【１２７１】

[1224]したがって、本技法の様々な態様は、オーディオ符号化デバイス７００Ａが、次の項に従って動作するように構成されることを可能にすることができる。

【１２７２】
[1225]１３３１４９－１Ｆ項。ベクトルの係数の０ではないセットをそれによって指定する複数の構成モードのうちの１つを選択することと、ベクトルが、音場を記述する複数の球面調和係数から分解されており、音場の明瞭な成分を表す、複数の構成モードのうちの選択された１つに基づいて、ベクトルの係数の０ではないセットを指定することとを行うように構成される、１つまたは複数のプロセッサを備えるデバイス。
【１２７３】
[1226]１３３１４９－２Ｆ項。複数の構成モードのうちの１つが、係数の０ではないセットが係数のすべてを含むことを示す、１３３１４９－１Ｆ項のデバイス。
【１２７４】
[1227]１３３１４９－３Ｆ項。複数の構成モードのうちの１つが、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含むことを示す、１３３１４９－１Ｆ項のデバイス。
【１２７５】
[1228]１３３１４９－４Ｆ項。複数の構成モードのうちの１つが、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含め、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含めないことを示す、１３３１４９－１Ｆ項のデバイス。
【１２７６】
[1229]１３３１４９－５Ｆ項。複数の構成モードのうちの１つが、係数の０ではないセットが係数の少なくとも１つを除き係数のすべてを含むことを示す、１３３１４９－１Ｆ項のデバイス。
【１２７７】
[1230]１３３１４９－６Ｆ項。１つまたは複数のプロセッサが、ビットストリームにおいて、複数の構成モードのうちの選択された１つを指定するようにさらに構成される、１３３１４９－１Ｆ項のデバイス。
【１２７８】
[1231]１３３１４９－１Ｇ項。複数の構成モードのうちの１つに従って、ベクトルの係数の０ではないセットをそれによって抽出する複数の構成モードのうちの１つを決定することと、ベクトルが、音場を記述する複数の球面調和係数から分解されており、音場の明瞭な成分を表す、複数の構成モードのうちの取得された１つに基づいて、ベクトルの係数の０ではないセットを抽出することとを行うように構成される、１つまたは複数のプロセッサを備えるデバイス。
【１２７９】
[1232]１３３１４９－２Ｇ項。複数の構成モードのうちの１つが、係数の０ではないセットが係数のすべてを含むことを示す、１３３１４９－１Ｇ項のデバイス。
【１２８０】
[1233]１３３１４９－３Ｇ項。複数の構成モードのうちの１つが、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含むことを示す、１３３１４９－１Ｇ項のデバイス。
【１２８１】
[1234]１３３１４９－４Ｇ項。複数の構成モードのうちの１つが、係数の０ではないセットが、係数のうちで、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応するものを含め、複数の球面調和係数の１つまたは複数が対応する基底関数の次数よりも大きい次数に対応する係数の少なくとも１つを含めないことを示す、１３３１４９－１Ｇ項のデバイス。
【１２８２】
[1235]１３３１４９－５Ｇ項。複数の構成モードのうちの１つが、係数の０ではないセットが係数の少なくとも１つを除き係数のすべてを含むことを示す、１３３１４９－１Ｇ

項のデバイス。

【１２８３】

[1236]１３３１４９－６Ｇ項。１つまたは複数のプロセッサが、複数の構成モードのうちの１つを決定するとき、ビットストリームにおいてシグナリングされた値に基づいて、複数の構成モードのうちの１つを決定するようにさらに構成される、１３３１４９－１Ｇ項のデバイス。

【１２８４】

[1237]図５２は、ＳＨＣ７０１を再構築またはほぼ再構築するために本開示で説明される技法の様々な態様を実施し得るオーディオ復号デバイス７５０Ａの別の例を示すブロック図である。図５２の例では、オーディオ復号デバイス７５０Ａは、抽出ユニット５４２がビットストリーム７１５’（ビットストリーム７１５’がSHC_{BG}７５２のオーディオ符号化されたバージョンをも含むことを除いて、図５１の例に関して上で説明されたビットストリーム７１５と類似である）と、サイドチャネル情報７１９とを受信することを除いて、図４１Ｄの例に示されるオーディオ復号デバイス５４０Ｄと類似である。この理由で、抽出ユニットは「抽出ユニット５４２’」として示される。

【１２８５】

[1238]その上、抽出ユニット５４２’は、抽出ユニット５４２’がＶ解凍ユニット５５５の修正された形態（図５２の例では「Ｖ解凍ユニット５５５’」と示される）を含む点において、抽出ユニット５４２とは異なる。Ｖ解凍ユニット５５５’は、サイドチャネル情報７１９と、ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈ７５４と示されるシンタックス要素とを受信する。抽出ユニット５４２’は、ビットストリーム７１５’から（および、一例では、ビットストリーム７１５’内に含まれたアクセスユニットヘッダから）ｃｏｄｅｄＶＶｅｃＬｅｎｇｔｈ７５４を解析する。Ｖ解凍ユニット５５５’は、上記の説明された構成モード７６０のうちの任意の１つに従って動作するように構成可能な、モード構成ユニット７５６（「モードｃｏｎｆｉｇユニット７５６」）と解析ユニット７５８とを含む。

【１２８６】

[1239]モード構成ユニット７５６は、シンタックス要素７５４を受信し、構成モード７６０のうちの１つを選択する。モード構成ユニット７５６は次いで、構成モード７６０のうちの選択された１つを用いて、解析ユニット７５８を構成する。解析ユニット７５８は、サイドチャネル情報７１９から圧縮された形態の小さい［Ｖ］ベクトル７１７を解析するために、構成モード７６０のうちの任意の１つに従って動作するように構成されるユニットを表す。解析ユニット７５８は、次の表に提示されたスイッチステートメントに従って動作することができる。

【１２８７】

【表１５－１】

テーブル－復号されたベクトル

シンタックス	ビットの数　ニーモニック

decodeVVec(i)
{
switch codedVVecLength {
case 0: ／／完全ベクトル
VVecLength = NumOfHoaCoeffs;
for (m=0; m< VVecLength; ++m){ VecCoeff[m] = m+1; }
break;
case 1: ／／より低い次数が除去される
VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA;
 	
for (m=0; m< VVecLength; ++m){
VecCoeff[m] = m + MinNumOfCoeffsForAmbHOA +
1;
}
break;
case 2:
VVecLength = NumOfHoaCoeffs –
MinNumOfCoeffsForAmbHOA - NumOfAddAmbHoaChan;
n = 0;
for(m=0;m<NumOfHoaCoeffs–
MinNumOfCoeffsForAmbHOA; ++m){
c = m + MinNumOfCoeffsForAmbHOA + 1;
if (ismember(c, AmbCoeffIdx) == 0){
VecCoeff[n] = c;
n++;
}
}	

【表１５－２】

break;
case 3:
VVecLength = NumOfHoaCoeffs – NumOfAddAmbHoaChan;
n = 0;
for(m=0; m<NumOfHoaCoeffs; ++m){
c = m + 1;
if (*ismember*(c, AmbCoeffIdx) == 0){
VecCoeff[n] = c;
n++;
}
}
}
 	
if (NbitsQ[i] == 5) { ／*一様量子化器*／
for (m=0; m< VVecLength; ++m){
VVec(k)[i][m] = (**VecValue** / 128.0) – 1.0;
	8	**uimsbf**

}
}
else { ／*ハフマン復号*／
for (m=0; m< VVecLength; ++m){
Idx = 5;
If (CbFlag[i] == 1) {
idx = (min(3, max(1, ceil(sqrt(VecCoeff[m]) - 1)));
}
else if (PFlag[i] == 1) {idx = 4;}
cid = *huffDecode*(huffmannTable[NbitsQ].codebook[idx]; **huffVal**);
	動的	**huffDecode**

if (cid > 0) { aVal = sgn = (**sgnVal** * 2) - 1;	**1**	**bslbf**
if (cid > 1) { aVal = sgn * (2.0^(cid -1) + **intAddVal**);	**cid-1**	**uimsbf**
} } else {aVal = 0.0;} } } }		

注：
－　一様量子化器のためのエンコーダ関数はmin(255, round((x+1. 0)*128. 0))
－　ニーモニックhuffDecodeのためのビットの数は動的である

【１２８８】

[1240]前述のシンタックステーブルにおいて、４つのケース（ケース０～３）を伴う第１のスイッチステートメントは、小さい［Ｖ］行列７１７の各ベクトルの長さを係数の数に関して決定する方法を提供する。第１のケースであるケース０は、V^{T}_{DIST}ベクトルの係数のすべてが指定されることを示す。第２のケースであるケース１は、V^{T}_{DIST}ベクト

ルの係数のうちでＭｉｎＮｕｍＯｆＣｏｅｆｆｓＦｏｒＡｍｂＨＯＡよりも大きい次数に対応するもののみが指定されることを示し、これは上で（N_{DIST}＋１）－（N_{BG}＋１）と呼ばれるものを示し得る。第３のケースであるケース２は、第２のケースに類似しているが、ＮｕｍＯｆＡｄｄＡｍｂＨｏａＣｈａｎによって識別される係数をさらに差し引き、これは次数N_{BG}を超える次数に対応する追加のチャネル（ここで「チャネル」は、ある次数と位数の組合せに対応する特定の係数を指す）を指定するための変数を示す。第４のケースであるケース３は、V^{T}_{DIST}ベクトルの係数のうちでＮｕｍＯｆＡｄｄＡｍｂＨｏａＣｈａｎによって識別される係数を除去した後で残るもののみが指定されることを示す。

【１２８９】

[1241]この点について、オーディオ復号デバイス７５０Ａは、音場の明瞭な成分を表すベクトルの係数の第１の０ではないセットを決定するために、本開示で説明される技法に従って動作することができ、このベクトルは、音場を記述する複数の球面調和係数から分解されたものである。

【１２９０】

[1242]その上、オーディオ復号デバイス７５０Ａは、複数の構成モードのうちの１つに従って、ベクトルの係数の０ではないセットをそれによって抽出する複数の構成モードのうちの１つを決定することと、ベクトルは、音場を記述する複数の球面調和係数から分解されており、音場の明瞭な成分を表す、複数の構成モードのうちの取得された１つに基づいて、ベクトルの係数の０ではないセットを抽出することとを行うために、本開示で説明される技法に従って動作するように構成され得る。

【１２９１】

[1243]図５３は、本開示において説明される技法の様々な態様を実行し得るオーディオ符号化デバイス５７０の別の例を示すブロック図である。図５３の例では、オーディオ符号化デバイス５７０は、オーディオ符号化デバイス５１０Ａ～５１０Ｊのうちの１つまたは複数と類似であり得る（ここで、次数低減ユニット５２８Ａは、音場成分抽出ユニット２０内に含まれると仮定されるが、説明を簡単にするために示されない）。ただし、オーディオ符号化デバイス５７０は、いくつかの例では、分解ユニット５１８を備え得る、より一般的な変換ユニット５７２を含み得る。

【１２９２】

[1244]図５４は、図５３の例に示されるオーディオ符号化デバイス５７０の例示的な実装形態をより詳細に示すブロック図である。図５４の例に示されるように、オーディオ符号化デバイス５７０の変換ユニット５７２は、回転ユニット６５４を含む。オーディオ符号化デバイス５７０の音場成分抽出ユニット５２０は、空間分析ユニット６５０と、コンテンツ特性分析ユニット６５２と、コヒーレント成分抽出ユニット６５６と、拡散成分抽出ユニット６５８とを含む。オーディオ符号化デバイス５７０のオーディオ符号化ユニット５１４は、ＡＡＣコーディングエンジン６６０と、ＡＡＣコーディングエンジン１６２とを含む。オーディオ符号化デバイス５７０のビットストリーム生成ユニット５１６は、マルチプレクサ（ＭＵＸ）１６４を含む。

【１２９３】

[1245]ＳＨＣの形態の３Ｄオーディオデータを表すために必要とされる帯域幅－ビット／秒に関して－は、消費者の使用に関して禁止とすることがある。たとえば、４８ｋＨｚのサンプリングレートを使用するとき、および３２ビット／同じ分解能を用いて－４次ＳＨＣ表現は、３６Ｍｂｉｔｓ／秒（２５×４８０００×３２ｂｐｓ）の帯域幅を表す。一般に約１００ｋｂｉｔｓ／秒である、ステレオ信号のための最先端のオーディオコーディングと比較すると、これは大きい数字である。図５４の例において実施される技法は、３Ｄオーディオ表現の帯域幅を低減することができる。

【１２９４】

[1246]空間分析ユニット６５０、コンテンツ特性分析ユニット６５２、および回転ユニット６５４は、ＳＨＣ５１１を受信することができる。本開示の他の場所で説明されるように、ＳＨＣ５１１は音場を表すことができる。図５４の例では、空間分析ユニット６５

０、コンテンツ特性分析ユニット６５２、および回転ユニット６５４は、音場の４次（ｎ＝４）表現に対して２５個のＳＨＣを受信することができる。

【１２９５】

[1247]空間分析ユニット６５０は、音場の明瞭な成分と音場の拡散成分とを識別するために、ＳＨＣ５１１によって表される音場を分析することができる。音場の明瞭な成分は、識別可能な方向から来ると知覚されるか、またはさもなければ音場のバックグラウンド成分もしくは拡散成分とは別個である音である。たとえば、個々の楽器によって生成される音は、識別可能な方向から来ると知覚され得る。対照的に、音場の拡散成分またはバックグラウンド成分は、識別可能な方向から来ると知覚されない。たとえば、森を通る風の音は、音場の拡散成分であり得る。

【１２９６】

[1248]空間分析ユニット６５０は、１つまたは複数の明瞭な成分を識別し、明瞭な成分のうちで最も大きいエネルギーを有するものを、（この音場を録音した、推定されるマイクロフォンに対する）垂直軸および／または水平軸とアラインするために、音場を回転させるための最適な角度を識別するように試みることができる。空間分析ユニット６５０は、これらの明瞭な成分が図１および図２の例に示される背後の球面基底関数とよりよくアラインするように音場が回転され得るように、この最適な角度を識別することができる。

【１２９７】

[1249]いくつかの例では、空間分析ユニット６５０は、拡散音（低レベルの方向または低次ＳＨＣを有する音を指すことがあり、ＳＨＣ５１１のうちで１以下の次数を有するものを意味する）を含むＳＨＣ５１１によって表される音場のパーセンテージを識別するために、ある形態の拡散分析を実行するように構成されるユニットを表し得る。一例として、空間分析ユニット６５０は、２００７年６月付けのＪ．　Ａｕｄｉｏ　Ｅｎｇ．　Ｓｏｃ．、第５５巻第６号で公開された「Ｓｐａｔｉａｌ　Ｓｏｕｎｄ　Ｒｅｐｒｏｄｕｃｔｉｏｎ　ｗｉｔｈ　Ｄｉｒｅｃｔｉｏｎａｌ　Ａｕｄｉｏ　Ｃｏｄｉｎｇ」という名称の、Ｖｉｌｌｅ　Ｐｕｌｋｋｉによる論文で説明される方式に類似した方式で拡散分析を実行することができる。いくつかの例では、空間分析ユニット６５０は、拡散パーセンテージを決定するために拡散分析を実行するとき、ＳＨＣ５１１の０次のものおよび１次のものなど、ＨＯＡ係数の０ではないサブセットのみを分析することができる。

【１２９８】

[1250]コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１に少なくとも一部基づいて、ＳＨＣ５１１が音場の自然な録音を介して生成されたか、または一例として、ＰＣＭオブジェクトなどのオーディオオブジェクトから人工的に（すなわち、合成的に）生成されたかを決定することができる。さらに、コンテンツ特性分析ユニット６５２は次いで、ＳＨＣ５１１が音場の実際の録音を介して生成されたか、人工的なオーディオオブジェクトから生成されたかに少なくとも一部基づいて、ビットストリーム５１７に含むべきチャネルの総数を決定することができる。たとえば、コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１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に少なくとも一部基づいて、ビットストリーム５１７が１６個のチャネルを含むべきであると決定することができる。チャネルの各々はモノラルチャネルであり得る。コンテンツ特性分析ユニット６５２は、ビットストリーム５１７の出力ビットレート、たとえば１．２Ｍｂｐｓに基づいて、ビットストリーム５１７に含むべきチャネルの総数の決定をさらに実行することができる。

【１２９９】

[1251]加えて、コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１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に少なくとも一部基づいて、チャネルのうちいくつを音場のコヒーレント成分または言い換えれば明瞭な成分に割り振るべきかと、チャネルのうちいくつを音場の拡散成分または言い換えればバックグラウンド成分に割り振るべきかとを決定することができる。たとえば、ＳＨＣ５１１が、一例としてＥｉｇｅｎｍｉｃを使用して実際の音場の録音から生成されたとき、コ

ンテンツ特性分析ユニット６５２は、チャネルのうちの３つを音場のコヒーレント成分に割り振ることがあり、残りのチャネルを音場の拡散成分に割り振ることがある。この例では、ＳＨＣ５１１が人工的なオーディオオブジェクトから生成されたとき、コンテンツ特性分析ユニット６５２は、チャネルのうちの５つを音場のコヒーレント成分に割り振ることがあり、残りのチャネルを音場の拡散成分に割り振ることがある。このようにして、コンテンツ分析ブロック（すなわち、コンテンツ特性分析ユニット６５２）は、音場のタイプ（たとえば、拡散／指向性など）を決定し、次に抽出するべきコヒーレント／拡散成分の数を決定することができる。

【１３００】

[1252]目標ビットレートは、成分の数と、個々のＡＡＣコーディングエンジン（たとえば、コーディングエンジン６６０、６６２）のビットレートとに影響を及ぼすことができる。言い換えれば、コンテンツ特性分析ユニット６５２は、ビットストリーム５１７の出力ビットレート、たとえば１．２Ｍｂｐｓに基づいて、いくつのチャネルをコヒーレント成分に割り振るべきかと、いくつのチャネルを拡散成分に割り振るべきかとの決定をさらに実行することができる。

【１３０１】

[1253]いくつかの例では、音場のコヒーレント成分に割り振られるチャネルは、音場の拡散成分に割り振られるチャネルよりも大きいビットレートを有し得る。たとえば、ビットストリーム５１７の最大ビットレートは、１．２Ｍｂ／ｓｅｃであり得る。この例では、コヒーレント成分に割り振られる４つのチャネル、および拡散成分に割り振られる１６個のチャネルがあり得る。さらに、この例では、コヒーレント成分に割り振られるチャネルの各々は、６４ｋｂ／ｓｅｃの最大ビットレートを有し得る。この例では、拡散成分に割り振られるチャネルの各々は、４８ｋｂ／ｓｅｃの最大ビットレートを有し得る。

【１３０２】

[1254]上で示されたように、コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１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。コンテンツ特性分析ユニット６５２は、この決定を様々な方法で行うことができる。たとえば、ビットストリーム符号化デバイス５７０は、４次ＳＨＣを使用することができる。この例では、コンテンツ特性分析ユニット６５２は、２４個のチャネルをコーディングし、２５番目のチャネル（ベクトルとして表され得る）を予測することができる。コンテンツ特性分析ユニット６５２は、２５番目のベクトルを決定するために、２４個のチャネルのうちの少なくともいくつかにスカラーを適用し、得られる値を加算することができる。さらに、この例では、コンテンツ特性分析ユニット６５２は、予測された２５番目のチャネルの精度を決定することができる。この例では、予測された２５番目のチャネルの精度が比較的高い（たとえば、精度が特定の閾値を超える）場合、ＳＨＣ５１１は、合成オーディオオブジェクトから生成される可能性が高い。対照的に、予測された２５番目のチャネルの精度が比較的低い（たとえば、精度が特定の閾値を下回る）場合、ＳＨＣ５１１は、録音された音場を表す可能性がより高い。たとえば、この例では、２５番目のチャネルの信号対雑音比（ＳＮＲ）が１００デシベル（ｄｂ）を超える場合、ＳＨＣ５１１は、合成オーディオオブジェクトから生成された音場を表す可能性がより高い。対照的に、ｅｉｇｅｎマイクロフォンを使用して録音された音場のＳＮＲは、５～２０ｄｂであり得る。したがって、実際の直接的な録音から生成されたＳＨＣ５１１によって表される音場と合成オーディオオブジェクトから生成されたＳＨＣ５１１によって表される音場との間には、ＳＮＲ比における明らかな境界が存在し得る。

【１３０３】

[1255]さらに、コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１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に少なくとも一部基づいて、Ｖベクトルを量子化するためのコードブックを選択することができる。言い換えれば、コンテンツ特性分析ユニット６５２は、ＨＯＡ係数によって表される音場が録音されるか合成であるかに応じて、Ｖベクトルを量子化するのに使用するための異なるコ

ードブックを選択することができる。

【１３０４】

[1256]いくつかの例では、コンテンツ特性分析ユニット６５２は、ＳＨＣ５１１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、繰り返し決定することができる。いくつかのそのような例では、この繰返しの基準は、フレームごとであり得る。他の例では、コンテンツ特性分析ユニット６５２は、この決定を１回実行することができる。さらに、コンテンツ特性分析ユニット６５２は、チャネルの総数と、コヒーレント成分チャネルおよび拡散成分チャネルの割振りとを、繰り返し決定することができる。いくつかのそのような例では、この繰返しの基準は、フレームごとであり得る。他の例では、コンテンツ特性分析ユニット６５２は、この決定を１回実行することができる。いくつかの例では、コンテンツ特性分析ユニット６５２は、Ｖベクトルを量子化するのに使用するためのコードブックを繰り返し選択することができる。いくつかのそのような例では、この繰返しの基準は、フレームごとであり得る。他の例では、コンテンツ特性分析ユニット６５２は、この決定を１回実行することができる。

【１３０５】

[1257]回転ユニット６５４は、ＨＯＡ係数の回転演算を実行することができる。本開示の他の場所で（たとえば、図５５および図５５Ｂに関して）説明されるように、回転演算を実行することによって、ＳＨＣ５１１を表すために必要とされるビットの数を低減することができる。いくつかの例では、回転ユニット６５２によって実行される回転分析は、特異値分解（「ＳＶＤ」）分析の一例である。主成分分析（「ＰＣＡ」）、独立成分分析（「ＩＣＡ」）、およびカルーネン－レーベ変換（「ＫＬＴ」）は、適用可能であり得る関連技法である。

【１３０６】

[1258]図５４の例では、コヒーレント成分抽出ユニット６５６は、回転されたＳＨＣ５１１を回転ユニット６５４から受信する。さらに、コヒーレント成分抽出ユニット６５６は、回転されたＳＨＣ５１１から、回転されたＳＨＣ５１１のうちで音場のコヒーレント成分と関連付けられるものを抽出する。

【１３０７】

[1259]加えて、コヒーレント成分抽出ユニット６５６は、１つまたは複数のコヒーレント成分チャネルを生成する。コヒーレント成分チャネルの各々は、音場のコヒーレント係数と関連付けられる回転されたＳＨＣ５１１の異なるサブセットを含み得る。図５４の例では、コヒーレント成分抽出ユニット６５６は、１から１６個のコヒーレント成分チャネルを生成することができる。コヒーレント成分抽出ユニット６５６によって生成されるコヒーレント成分チャネルの数は、コンテンツ特性分析ユニット６５２によって音場のコヒーレント成分に割り振られるチャネルの数によって決定され得る。コヒーレント成分抽出ユニット６５６によって生成されるコヒーレント成分チャネルのビットレートは、コンテンツ特性分析ユニット６５２によって決定され得る。

【１３０８】

[1260]同様に、図５４の例では、拡散成分抽出ユニット６５８は、回転されたＳＨＣ５１１を回転ユニット６５４から受信する。さらに、拡散成分抽出ユニット６５８は、回転されたＳＨＣ５１１から、回転されたＳＨＣ５１１のうちで音場の拡散成分と関連付けられるものを抽出する。

【１３０９】

[1261]加えて、拡散成分抽出ユニット６５８は、１つまたは複数の拡散成分チャネルを生成する。拡散成分チャネルの各々は、音場の拡散係数と関連付けられる回転されたＳＨＣ５１１の異なるサブセットを含み得る。図５４の例では、拡散成分抽出ユニット６５８は、１から９個の拡散成分チャネルを生成することができる。拡散成分抽出ユニット６５８によって生成される拡散成分チャネルの数は、コンテンツ特性分析ユニット６５２によって音場の拡散成分に割り振られるチャネルの数によって決定され得る。拡散成分抽出ユニット６５８によって生成される拡散成分チャネルのビットレートは、コンテンツ特性分

析ユニット６５２によって決定され得る。

【１３１０】

[1262]図５４の例では、ＡＡＣコーディングユニット６６０は、コヒーレント成分抽出ユニット６５６によって生成されるコヒーレント成分チャネルを符号化するために、ＡＡＣコーデックを使用することができる。同様に、ＡＡＣコーディングユニット６６２は、拡散成分抽出ユニット６５８によって生成される拡散成分チャネルを符号化するために、ＡＡＣコーデックを使用することができる。マルチプレクサ６６４（「ＭＵＸ６６４」）は、ビットストリーム５１７を生成するために、サイドデータ（たとえば、空間分析ユニット６５０によって決定される最適な角度）とともに、符号化されたコヒーレント成分チャネルと符号化された拡散成分チャネルとを多重化することができる。

【１３１１】

[1263]このようにして、本技法は、オーディオ符号化デバイス５７０が、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定することを可能にすることができる。

【１３１２】

[1264]いくつかの例では、オーディオ符号化デバイス５７０は、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場の明瞭な成分を表す球面調和係数のサブセットを決定することができる。これらの例および他の例において、オーディオ符号化デバイス５７０は、球面調和係数のサブセットを含むようにビットストリームを生成することができる。オーディオ符号化デバイス５７０は、いくつかの例では、球面調和係数のサブセットをオーディオ符号化し、球面調和係数のオーディオ符号化されたサブセットを含むようにビットストリームを生成することができる。

【１３１３】

[1265]いくつかの例では、オーディオ符号化デバイス５７０は、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場のバックグラウンド成分を表す球面調和係数のサブセットを決定することができる。これらの例および他の例において、オーディオ符号化デバイス５７０は、球面調和係数のサブセットを含むようにビットストリームを生成することができる。これらの例および他の例において、オーディオ符号化デバイス５７０は、球面調和係数のサブセットをオーディオ符号化し、球面調和係数のオーディオ符号化されたサブセットを含むようにビットストリームを生成することができる。

【１３１４】

[1266]いくつかの例では、オーディオ符号化デバイス５７０は、球面調和係数によって表される音場を回転させるための角度を識別するために、球面調和係数に関して空間分析を実行し、回転された球面調和係数を生成するために、識別された角度だけ音場を回転させるために、回転演算を実行することができる。

【１３１５】

[1267]いくつかの例では、オーディオ符号化デバイス５７０は、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場の明瞭な成分を表す球面調和係数の第１のサブセットを決定し、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場のバックグラウンド成分を表す球面調和係数の第２のサブセットを決定することができる。これらの例および他の例において、オーディオ符号化デバイス５７０は、球面調和係数の第２のサブセットをオーディオ符号化するために使用されるものよりも高い目標ビットレートを有する、球面調和係数の第１のサブセットをオーディオ符号化することができる。

【１３１６】

[1268]このようにして、本技法の様々な態様は、オーディオ符号化デバイス５７０が次の項に従って、ＳＣＨ５１１が合成オーディオオブジェクトから生成されるかどうかを決定することを可能にし得る。

【１３１７】

[1269]１３２５１２－１項。オーディオ符号化デバイス５７０などのデバイスであって、１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するようにさらに構成される、を備えるデバイス。

【１３１８】

[1270]１３２５１２－２項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、低減されたフレーム分割された球面調和係数行列を取得するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外するようにさらに構成される、１３２５１２－１項のデバイス。

【１３１９】

[1271]１３２５１２－３項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、低減されたフレーム分割された球面調和係数行列を取得するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外することと、低減されたフレーム分割された球面調和係数行列の残りのベクトルに基づいて、低減されたフレーム分割された球面調和係数行列のベクトルを予測することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２０】

[1272]１３２５１２－４項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、低減されたフレーム分割された球面調和係数行列を取得するために、音場を表す球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列から第１のベクトルを除外することと、低減されたフレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、低減されたフレーム分割された球面調和係数行列のベクトルを予測することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２１】

[1273]１３２５１２－５項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測するようにさらに構成される、１３２５１２－１項のデバイス。

【１３２２】

[1274]１３２５１２－６項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測することと、予測されるベクトルに基づいて誤差を計算することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２３】

[1275]１３２５１２－７項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測することと、フレーム分割された球面調和係数行列の予測されるベクトルと対応するベクトルとに基づいて、誤差を計算することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２４】

[1276]１３２５１２－８項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数

の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測することと、フレーム分割された球面調和係数行列の予測されるベクトルと対応するベクトルとの差の絶対値の合計として、誤差を計算することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２５】

[1277]１３２５１２－９項。１つまたは複数のプロセッサが、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するとき、フレーム分割された球面調和係数行列の残りのベクトルの合計に少なくとも一部基づいて、球面調和係数の少なくとも一部分を記憶するフレーム分割された球面調和係数行列のベクトルを予測することと、フレーム分割された球面調和係数行列の予測されるベクトルと対応するベクトルとに基づいて誤差を計算することと、フレーム分割された球面調和係数行列の対応するベクトルのエネルギーと誤差とに基づいて比を計算することと、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するために比を閾値と比較することとを行うようにさらに構成される、１３２５１２－１項のデバイス。

【１３２６】

[1278]１３２５１２－１０項。１つまたは複数のプロセッサが、ベクトルを予測するとき、球面調和係数の少なくともその部分を記憶するフレーム分割された球面調和係数行列の第１の０ではないベクトルを予測するようにさらに構成される、請求項４乃至９のいずれかのデバイス。

【１３２７】

[1279]１３２５１２－１１項。１つまたは複数のプロセッサが、球面調和係数の圧縮されたバージョンを記憶するビットストリームにおいて、球面調和係数が合成オーディオオブジェクトから生成されるかどうかのインジケーションを指定するようにさらに構成される、請求項１乃至１０のいずれかのデバイス。

【１３２８】

[1280]１３２５１２－１２項。インジケーションが単一のビットである、１３２５１２－１１項のデバイス。

【１３２９】

[1281]１３２５１２－１３項。１つまたは複数のプロセッサが、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場の明瞭な成分を表す球面調和係数のサブセットを決定するようにさらに構成される、１３２５１２－１項のデバイス。

【１３３０】

[1282]１３２５１２－１４項。１つまたは複数のプロセッサが、球面調和係数のサブセットを含むようにビットストリームを生成するようにさらに構成される、１３２５１２－１３項のデバイス。

【１３３１】

[1283]１３２５１２－１５項。１つまたは複数のプロセッサが、球面調和係数のサブセットをオーディオ符号化することと、球面調和係数のオーディオ符号化されたサブセットを含むようにビットストリームを生成することとを行うようにさらに構成される、１３２５１２－１３項のデバイス。

【１３３２】

[1284]１３２５１２－１６項。１つまたは複数のプロセッサが、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場のバックグラウンド成分を表す球面調和係数のサブセットを決定するようにさらに構成される、１３２５１２－１項のデバイス。

【１３３３】

[1285]１３２５１２－１７項。１つまたは複数のプロセッサが、球面調和係数のサブセットを含むようにビットストリームを生成するようにさらに構成される、１３２５１２－１６項のデバイス。

【１３３４】
[1286]１３２５１２－１８項。１つまたは複数のプロセッサが、球面調和係数のサブセットをオーディオ符号化することと、球面調和係数のオーディオ符号化されたサブセットを含むようにビットストリームを生成することとを行うようにさらに構成される、１３２５１２－１５項のデバイス。
【１３３５】
[1287]１３２５１２－１８項。１つまたは複数のプロセッサが、球面調和係数によって表される音場を回転させるための角度を識別するために、球面調和係数に関して空間分析を実行することと、回転された球面調和係数を生成するために、識別された角度だけ音場を回転させるために、回転演算を実行することとを行うようにさらに構成される、１３２５１２－１項のデバイス。
【１３３６】
[1288]１３２５１２－２０項。１つまたは複数のプロセッサが、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場の明瞭な成分を表す球面調和係数の第１のサブセットを決定することと、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、音場のバックグラウンド成分を表す球面調和係数の第２のサブセットを決定することとを行うようにさらに構成される、１３２５１２－１項のデバイス。
【１３３７】
[1289]１３２５１２－２１項。１つまたは複数のプロセッサが、球面調和係数の第２のサブセットをオーディオ符号化するために使用されるものよりも高い目標ビットレートを有する、球面調和係数の第１のサブセットをオーディオ符号化するようにさらに構成される、１３２５１２－２０項のデバイス。
【１３３８】
[1290]１３２５１２－２２項。１つまたは複数のプロセッサが、複数の球面調和係数の左特異ベクトルを表すＵ行列と、複数の球面調和係数の特異値を表すＳ行列と、複数の球面調和係数の右特異ベクトルを表すＶ行列とを生成するために、球面調和係数に関して特異値分解を実行するようにさらに構成される、１３２５１２－１項のデバイス。
【１３３９】
[1291]１３２５１２－２３項。１つまたは複数のプロセッサが、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数のうちで音場の明瞭な成分を表す部分を決定するようにさらに構成される、１３２５１２－２２項のデバイス。
【１３４０】
[1292]１３２５１２－２４項。１つまたは複数のプロセッサが、球面調和係数が合成オーディオオブジェクトから生成されるかどうかに基づいて、Ｕ行列、Ｓ行列、およびＶ行列の１つまたは複数のうちで音場のバックグラウンド成分を表す部分を決定するようにさらに構成される、１３２５１２－２２項のデバイス。
【１３４１】
[1293]１３２５１２－１Ｃ項。オーディオ符号化デバイス５７０などのデバイスであって、少なくとも、球面調和係数のベクトルのエネルギーと、球面調和係数のベクトルの予測されるバージョンおよび球面調和係数のベクトルに基づいて導出される誤差との関数として、計算される比に基づいて、音場を表す球面調和係数が合成オーディオオブジェクトから生成されるかどうかを決定するように構成される１つまたは複数のプロセッサを備えるデバイス。
【１３４２】
[1294]上で説明された様々な例の各々において、オーディオ符号化デバイス５７０は、ある方法を実行し、またはそうでなければ、オーディオ符号化デバイス５７０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。いくつかの例では、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

の例では、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５７０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。

【１３４３】

[1295]図５５および５５Ｂは、音場６４０を回転させるために本開示において説明される技法の様々な態様を実行する例を示す図である。図５５は、本開示で説明される技法の様々な態様による回転の前の音場６４０を示す図である。図５５の例では、音場６４０は、ロケーション６４２Ａおよび６４２Ｂと示される、高圧の２つのロケーションを含む。これらのロケーション６４２Ａおよび６４２Ｂ（「ロケーション６４２」）は、０ではない傾きを有する線６４４に沿って存在する（水平線は０の傾きを有するので、これは水平ではない線を指す別の方法である）。ロケーション６４２がｘ座標およびｙ座標に加えてｚ座標を有するとすると、高次球面基底関数が、この音場６４０を正しく表すために必要とされ得る（これらの高次球面基底関数は、音場の上側部分と下側部分または非水平部分とを記述するので。音場６４０をＳＨＣ５１１に直接的に低減するのではなく、オーディオ符号化デバイス５７０は、ロケーション６４２を接続する線６４４が水平になるまで、音場６４０を回転させることができる。

【１３４４】

[1296]図５５Ｂは、ロケーション６４２を接続する線６４４が水平になるまで回転された後の音場６４０を示す図である。この方式で音場６４０を回転させた結果として、回転された音場６４０がもはやｚ座標との圧力（またはエネルギー）のいかなるロケーションをも有していないとすると、ＳＨＣ５１１は、ＳＨＣ５１１の高次のものが０と指定されるように導出され得る。このようにして、オーディオ符号化デバイス５７０は、０ではない値を有するＳＨＣ５１１の数を低減するために音場６４０を回転させ、トランスレートし、またはより一般的には調整することができる。本技法の様々な他の態様とともに、オーディオ符号化デバイス５７０は次いで、ＳＨＣ５１１のこれらの高次のものが０という値を有することを識別する３２ビット符号付き数をシグナリングするのではなく、ＳＨＣ５１１のこれらの高次のものがシグナリングされないことをビットストリーム５１７のフィールド内でシグナリングすることができる。オーディオ符号化デバイス５７０はまた、しばしば、上で説明された方式で方位角と仰角とを表すことによって、音場６４０がどのように回転されたかを示す、ビットストリーム５１７内の回転情報を指定することもできる。オーディオ符号化デバイスなど、抽出デバイスは次いで、ＳＨＣ５１１のこれらのシグナリングされないものが０という値を有することを暗示し、ＳＨＣ５１１に基づいて音場６４０を再現するとき、図５５の例に示された音場６４０に音場６４０が似ているように音場６４０を回転させるために、回転を実行することができる。このようにして、オーディオ符号化デバイス５７０は、本開示で説明される技法に従って、ビットストリーム５１７において指定されることが必要とされるＳＨＣ５１１の数を低減することができる。

【１３４５】

[1297]「空間圧縮」アルゴリズムは、音場の最適な回転を決定するために使用され得る。一実施形態では、オーディオ符号化デバイス５７０は、可能な方位角と仰角の組合せ（すなわち、上記の例では１０２４×５１２個の組合せ）のすべてを反復し、各組合せに対して音場を回転させ、閾値を上回るＳＨＣ５１１の数を計算するためにアルゴリズムを実行することができる。閾値を上回るＳＨＣ５１１の最小数を生じる方位角／仰角候補の組合せは、「最適な回転」と呼ばれることがあるものと考えられ得る。この回転された形態では、音場は、音場を表すためのＳＨＣ５１１の最小数を必要とすることがあり、次いで、圧縮されると考えられ得る。いくつかの例では、調整は、この最適な回転を備えることがあり、上で説明された調整情報は、この回転（「最適な回転」と呼ばれることがある）情報（方位角角度および仰角角度に関する）を含むことがある。

【１３４６】

[1298]いくつかの例では、方位角角度と仰角角度とを指定するのみではなく、オーディオ符号化デバイス５７０は、一例としてオイラー角の形態の追加の角度を指定することができる。オイラー角は、ｚ軸、前ｘ軸、および前ｚ軸の周りの回転の角度を指定する。本開示では方位角角度と仰角角度の組合せに関して説明されているが、本開示の技法は、方位角角度と仰角角度のみを指定することに限定されるべきではなく、上で述べられた３つのオイラー角を含む任意の数の角度を指定することを含み得る。この意味で、オーディオ符号化デバイス５７０は、音場を記述することに関連する情報を提供し、ビットストリーム内の回転情報としてオイラー角を指定する、複数の階層的な要素の数を低減するために、音場を回転させることができる。オイラー角は、上で述べられたように、音場がどのように回転されたかを記述することができる。オイラー角を使用するとき、ビットストリーム抽出デバイスは、オイラー角を含む回転情報を決定するためにビットストリームを解析し、さらに、複数の階層的な要素のうちで音場を記述することに関連する情報を提供するものに基づいて音場を再現するとき、オイラー角に基づいて音場を回転させることができる。

【１３４７】

[1299]その上、いくつかの例では、これらの角度をビットストリーム５１７内で明示的に指定するのではなく、オーディオ符号化デバイス５７０は、回転を指定する１つまたは複数の角度のあらかじめ定義された組合せと関連付けられるインデックス（「回転インデックス」と呼ばれることがある）を指定することができる。言い換えれば、回転情報は、いくつかの例では、回転インデックスを含むことがある。これらの例では、０の値などの回転インデックスの所与の値は、回転が実行されなかったことを示すことがある。この回転インデックスは、回転テーブルに関連して使用され得る。すなわち、オーディオ符号化デバイス５７０は、方位角角度と仰角角度の組合せの各々に関するエントリを備える回転テーブルを含み得る。

【１３４８】

[1300]代替的に、回転テーブルは、方位角角度と仰角角度の各組合せを表す各行列変換に関するエントリを含み得る。すなわち、オーディオ符号化デバイス５７０は、方位角角度と仰角角度の組合せの各々によって音場を回転させるための各行列変換に関するエントリを有する回転テーブルを記憶することができる。一般に、オーディオ符号化デバイス５７０はＳＨＣ５１１を受信し、回転が実行されるとき、次の式に従ってＳＨＣ５１１’を導出する。

【１３４９】

【数４８】

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

【１３５０】

上の式では、ＳＨＣ５１１’は、第２の基準フレーム（$EncMat_2$）に関して音場を符号化するための符号化行列、第１の基準フレーム（$InvMat_1$）に関してＳＨＣ５１１を音場に戻すための逆行列、およびＳＨＣ５１１の関数として計算される。$EncMat_2$は２５×３２のサイズであり、$InvMat_2$は３２×２５のサイズである。ＳＨＣ５１１’とＳＨＣ５１１の両方は２５のサイズであり、ここで、ＳＨＣ５１１’は、顕著なオーディオ情報を指定しないものの除去により、さらに低減され得る。$EncMat_2$は、方位角角度と仰角角度の各組合せに対して変化し得るが、$InvMat_1$は、方位角角度と仰角角度の各組合せに対して変化しないままであり得る。回転テーブルは、各異なる$EncMat_2$を$InvMat_1$に乗算した結果を記憶するエントリを含み得る。

【１３５１】

[1301]図５６は、第１の基準フレームに従ってキャプチャされ、次いで第２の基準フレ

ームに対して音場を表すために本開示において説明される技法に従って回転される例示的な音場を示す図である。図５６の例では、Ｅｉｇｅｎマイクロフォン６４６を取り囲む音場は、図５６の例ではX_1軸、Y_1軸、およびZ_1軸によって示される第１の基準フレームを仮定してキャプチャされる。ＳＨＣ５１１は、この第１の基準フレームに関して音場を記述する。ＩｎｖＭａｔ$_1$は、ＳＨＣ５１１を変換して音場に戻し、音場が、図５６の例ではX_2軸、Y_2軸、およびZ_2軸によって示される第２の基準フレームに回転されることを可能にする。上で説明されたＥｎｃＭａｔ$_2$は、音場を回転させ、第２の基準フレームに関してこの回転された音場を記述するＳＨＣ５１１’を生成することができる。

【１３５２】

[1302]いずれにしても、上の式は、次のように導出され得る。音場が、正面がｘ軸の方向と見なされるように、特定の座標系を用いて記録されるとすると、Ｅｉｇｅｎマイクロフォンの３２個のマイクロフォン位置（または他のマイクロフォン構成）は、この基準座標系から定義される。次いで、音場の回転は、この基準フレームの回転と見なされ得る。仮定される基準フレームに対して、ＳＨＣ５１１は、次のように計算され得る。

【１３５３】

【数４９】

$$\begin{bmatrix} SHC \\ 27 \end{bmatrix} = \begin{bmatrix} Y_0^0(Pos_1) & Y_0^0(Pos_2) & \cdots & Y_0^0(Pos_{32}) \\ Y_1^{-1}(Pos_1) & \cdots & & Y_1^{-1}(Pos_{32}) \\ \vdots & \ddots & & \\ Y_4^4(Pos_1) & & & Y_4^4(Pos_{32}) \end{bmatrix} \begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix}$$

【１３５４】

上の式では、

【１３５５】

【数５０】

$$Y_n^m$$

【１３５６】

は、ｉ番目のマイクロフォン（ここで、この例では、ｉは１～３２であり得る）の位置（Pos_i）における球面基底関数を表す。mic_iベクトルは、時刻ｔに対するｉ番目のマイクロフォンのためのマイクロフォン信号を示す。位置（Pos_i）は、第１の基準フレーム（すなわち、この例では、回転の前の基準フレーム）におけるマイクロフォンの位置を指す。

【１３５７】

[1303]上の式は、代替的に、上で示された数式に関して

［ＳＨＣ＿２７］＝［E_s（θ，φ）］［m_i（ｔ）］

と表され得る。

【１３５８】

[1304]音場を（または第２の基準フレーム内で）回転させるために、位置（Pos_i）は第２の基準フレーム内で計算されるようになる。オリジナルのマイクロフォン信号が存在する限り、音場は、恣意的に回転されてよい。しかしながら、オリジナルのマイクロフォン信号（mic_i（ｔ））は入手不可能なことが多い。次いで、問題は、マイクロフォン信号（mic_i（ｔ））をＳＨＣ５１１からどのように取り出すかであり得る。（３２マイクロフォンのＥｉｇｅｎマイクロフォンの場合のように）Ｔ字型設計が使用される場合、この問題の解決策は、次の式を解くことによって達成され得る。

【１３５９】

【数５１】

$$\begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix} = [InvMat_1] \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

【１３６０】
このＩｎｖＭａｔ$_1$は、第１の基準フレームに対して指定されたマイクロフォンの位置に従って計算される球面調和基底関数を指定することができる。この式はまた、上で述べられたように、［ｍ$_i$（ｔ）］＝［Ｅ$_s$（θ，φ）］$^{-1}$［ＳＨＣ］として表され得る。
【１３６１】
[1305]マイクロフォン信号（ｍｉｃ$_i$（ｔ））が上の式に従って取り出されると、音場を記述するマイクロフォン信号（ｍｉｃ$_i$（ｔ））は、第２の基準フレームに対応するＳＨＣ５１１’を計算するために回転され、次の式になり得る。
【１３６２】
【数５２】

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

【１３６３】
ＥｎｃＭａｔ$_2$は、回転された位置（Ｐｏｓ$_i$’）から球面調和基底関数を指定する。このようにして、ＥｎｃＭａｔ$_2$は、方位角角度と仰角角度の組合せを効果的に指定することができる。したがって、回転テーブルが、方位角角度と仰角角度の各組合せに対する
【１３６４】
【数５３】

$$\begin{bmatrix} EncMat_2 \\ (25x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x25) \end{bmatrix}$$

【１３６５】
の結果を記憶するとき、回転テーブルは、方位角角度と仰角角度の各組合せを効果的に指定する。上の式はまた、以下のように表され得、

［ＳＨＣ２７′］＝［Ｅ$_s$（θ$_2$，φ$_2$）］［Ｅ$_s$（θ$_1$，φ$_1$）］$^{-1}$［ＳＨＣ２７］

ここで、θ$_2$，φ$_2$は、θ$_1$，φ$_1$によって表される第１の方位角角度および仰角角度と異なる第２の方位角角度および第２の仰角角度を表す。θ$_1$，φ$_1$は第１の基準フレームに対応するが、θ$_2$，φ$_2$は第２の基準フレームに対応する。したがって、ＩｎｖＭａｔ$_1$は、［Ｅ$_s$（θ$_1$，φ$_1$）］$^{-1}$に対応し得るが、ＥｎｃＭａｔ$_2$は、［Ｅ$_s$（θ$_2$，φ$_2$）］に対応し得る。
【１３６６】
[1306]上記は、次数ｎの球ベッセル関数を指すｊ$_n$（・）関数によって周波数領域におけるＳＨＣ５１１の導出を示す様々な式において上で表されるフィルタリング演算を考慮しない計算のより簡略化されたバージョンを表すことができる。時間領域では、このｊ$_n$（・）関数は、特定の次数ｎに固有のフィルタリング演算を表す。フィルタリングにより

、回転は、次数ごとに実行され得る。説明するために、次の式について考える。

【１３６７】

【数５４】

$$a_n^k(t) \square b_n(t) * \left\langle \left[Y_n^m \right] \square \left[m_i(t) \right] \right\rangle$$

$$a_n^k(t) \square \left\langle \left[Y_n^m \right] \square b_n(t) * \left[m_i(t) \right] \right\rangle$$

【１３６８】

[1307]これらの式から、b_n（ｔ）は互いに対して異なるので、次数に対する回転されたＳＨＣ５１１’は個別に行われる。結果として、上の式は、回転されたＳＨＣ５１１’の１次のものを計算するために、次のように変更され得る。

【１３６９】

【数５５】

$$\begin{bmatrix} 1次 \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (3x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x3) \end{bmatrix} \begin{bmatrix} 1次 \\ Order \\ SHC \\ 27 \end{bmatrix}$$

【１３７０】

ＳＨＣ５１１の３つの１次のものが存在するとすると、ＳＨＣ５１１’ベクトルおよびＳＨＣ５１１ベクトルの各々は、上の式では、３のサイズである。同様に、２次の場合、次の式が適用され得る。

【１３７１】

【数５６】

$$\begin{bmatrix} 2次 \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (5x32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32x5) \end{bmatrix} \begin{bmatrix} 2次 \\ Order \\ SHC \\ 27 \end{bmatrix}$$

【１３７２】

この場合も、ＳＨＣ５１１の５つの２次のものが存在するとすると、ＳＨＣ５１１’ベクトルおよびＳＨＣ５１１ベクトルの各々は、上の式では、５のサイズである。他の次数、すなわち３次および４次に対する残りの式は、（$EncMat_2$の行の数、$InvMat_1$の列の数、ならびに３次および４次のＳＨＣ５１１ベクトルおよびＳＨＣ５１１’ベクトルのサイズが３次球面調和基底関数および４次球面調和基底関数の各々の位数の数（ｍ×２＋１）に等しいことにおいて、行列のサイズに関する同じパターンに従って、上で説明されたものと類似であり得る。

【１３７３】

[1308]したがって、オーディオ符号化デバイス５７０は、いわゆる最適な回転を識別しようとして、方位角角度と仰角角度のあらゆる組合せに関して、この回転演算を実行することができる。オーディオ符号化デバイス５７０は、この回転演算を実行した後、閾値を上回るＳＨＣ５１１’の数を計算することができる。いくつかの例では、オーディオ符号

化デバイス５７０は、オーディオフレームなどの持続時間にわたって音場を表す一連のＳＨＣ５１１’を導出するために、この回転を実行することができる。この持続時間にわたって音場を表す一連のＳＨＣ５１１’を導出するためにこの回転を実行することによって、オーディオ符号化デバイス５７０は、フレームまたは他の長さよりも短い持続時間にわたって音場を記述するＳＨＣ５１１の各セットに対してこれを行うために比較すると、実行されなければならない回転演算の数を低減することができる。いずれにしても、オーディオ符号化デバイス５７０は、このプロセス全体を通じて、ＳＨＣ５１１’のうちで、閾値よりも大きいＳＨＣ５１１’の最小数を有するものを保存することができる。

【１３７４】

[1309]しかしながら、方位角角度と仰角角度のあらゆる組合せに関してこの回転演算を実行することは、プロセッサ集約的であり、または時間がかかることがある。結果として、オーディオ符号化デバイス５７０は、回転アルゴリズムのこの「力づくの」実装形態と特徴付けられ得るものを実行しないことがある。代わりに、オーディオ符号化デバイス５７０は、一般に良い圧縮を提供する方位角角度と仰角角度のおそらく既知の（統計学的な）組合せのサブセットに関して回転を実行し、このサブセットのうちでサブセット内の他の組合せと比較してより良い圧縮を提供するものの周りの組合せに関してさらなる回転を実行することができる。

【１３７５】

[1310]別の代替として、オーディオ符号化デバイス５７０は、組合せの既知のサブセットのみに関してこの回転を実行することができる。別の代替として、オーディオ符号化デバイス５７０は、組合せの軌道を（空間的に）たどり、この組合せの軌道に関して回転を実行することができる。別の代替として、オーディオ符号化デバイス５７０は、閾値を上回る０ではない値を有するＳＨＣ５１１’の最大数を定義する圧縮閾値を指定することができる。この圧縮閾値は、オーディオ符号化デバイス５７０が回転を実行し、設定された閾値を上回る値を有するＳＨＣ５１１’の数が圧縮閾値以下である（または、いくつかの例では、未満である）と決定するとき、オーディオ符号化デバイス５７０が、残りの組合せに関していかなる追加の回転演算を実行することも停止するように、探索に対する停止点を効果的に設定することができる。さらに別の代替として、オーディオ符号化デバイス５７０は、組合せの階層的に構成された木（または他のデータ構造）を通り、現在の組合せに関して回転演算を実行し、閾値よりも大きい０ではない値を有するＳＨＣ５１１’の数に応じて木を右または左に（たとえば、二分木の場合）通ることができる。

【１３７６】

[1311]この意味で、これらの代替の各々は、第１の回転演算と第２の回転演算とを実行することと、閾値よりも大きい０ではない値を有するＳＨＣ５１１’の最小数をもたらす第１の回転演算および第２の回転演算のうちの１つを識別するために、第１の回転演算と第２の回転演算とを実行した結果を比較することとを伴う。したがって、オーディオ符号化デバイス５７０は、第１の方位角角度および第１の仰角角度に従って音場を回転させ、音場を記述することに関連する情報を提供する第１の方位角角度および第１の仰角角度に従って回転された音場を表す複数の階層的な要素の第１の数を決定するために、音場に対して第１の回転演算を実行することができる。オーディオ符号化デバイス５７０はまた、第２の方位角角度および第２の仰角角度に従って音場を回転させ、音場を記述することに関連する情報を提供する第２の方位角角度および第２の仰角角度に従って回転された音場を表す複数の階層的な要素の第２の数を決定するために、音場に対して第２の回転演算を実行することもできる。さらに、オーディオ符号化デバイス５７０は、複数の階層的な要素の第１の数と複数の階層的な要素の第２の数との比較に基づいて、第１の回転演算または第２の回転演算を選択することができる。

【１３７７】

[1312]いくつかの例では、回転アルゴリズムは持続時間に対して実行されることがあり、ここで、回転アルゴリズムのその後の呼出しは、回転アルゴリズムの過去の呼出しに基づいて回転演算を実行することができる。言い換えれば、回転アルゴリズムは、以前の持

続時間にわたって音場を回転させたときに決定された過去の回転情報に基づいて適応的であり得る。たとえば、オーディオ符号化デバイス５７０は、第１の持続時間たとえばオーディオフレームにわたってＳＨＣ５１１’を識別するために、この第１の持続時間にわたって音場を回転させることができる。オーディオ符号化デバイス５７０は、上で説明された方法のいずれかにおいて、ビットストリーム５１７内で回転情報とＳＨＣ５１１’とを指定することができる。この回転情報は、第１の持続時間にわたって音場の回転を記述することにおいて、第１の回転情報と呼ばれることがある。次いで、オーディオ符号化デバイス５７０は、第２の持続時間たとえば第２のオーディオフレームにわたってＳＨＣ５１１’を識別するために、この第１の回転情報に基づいて、この第２の持続時間にわたって音場を回転させることができる。オーディオ符号化デバイス５７０は、一例として、方位角角度と仰角角度の「最適な」組合せに対して探索を初期化するために、第２の持続時間にわたって第２の回転演算を実行するとき、この第１の回転情報を利用することができる。次いで、オーディオ符号化デバイス５７０は、ビットストリーム５１７内で第２の持続時間（「第２の回転情報」と呼ばれることがある）に対するＳＨＣ５１１’と対応する回転情報とを指定することができる。

【１３７８】

[1313]処理時間および／または消費を低減するために回転アルゴリズムを実施するいくつかの異なる方法に関して上で説明されているが、本技法は、「最適な回転」と呼ばれ得るものの識別を低減またはさもなければ高速化し得る任意のアルゴリズムに関して実行され得る。その上、本技法は、非最適な回転を識別するが、速度またはプロセッサもしくは他のリソースの利用に関して測定されることが多い、他の態様では実行を改善し得る任意のアルゴリズムに関して実行され得る。

【１３７９】

[1314]図５７～図５７Ｅは各々、本開示で説明される技法に従って形成されるビットストリーム５１７Ａ～５１７Ｅを示す図である。図５７Ａの例では、ビットストリーム５１７Ａは、上で図５３に示されたビットストリーム５１７の一例を表し得る。ビットストリーム５１７Ａは、ＳＨＣ存在フィールド６７０と、ＳＨＣ５１１’を記憶するフィールド（ここで、このフィールドは「ＳＨＣ５１１’」と示される）とを含む。ＳＨＣ存在フィールド６７０は、ＳＨＣ５１１の各々に対応するビットを含み得る。ＳＨＣ５１１’は、ＳＨＣ５１１のうちで、ＳＨＣ５１１の数よりも数が少なくなり得る、ビットストリーム内で指定されるものを表し得る。一般に、ＳＨＣ５１１’の各々は、ＳＨＣ５１１のうちで０ではない値を有するものである。上で述べられたように、任意の所与の音場の４次表現の場合、$(1+4)^2$すなわち２５個のＳＨＣが必要とされる。これらのＳＨＣのうちの１つまたは複数を排除し、これらの０という値が付けられたＳＨＣを単一のビットで置き換えることによって、３１ビットを節約することができ、この３１ビットは、より詳細に音場の他の部分を表すことに割り振られてもよいし、またはさもなければ、効率的な帯域幅利用を容易にするために除去されてもよい。

【１３８０】

[1315]図５７Ｂの例では、ビットストリーム５１７Ｂは、上で図５３に示されたビットストリーム５１７の一例を表し得る。ビットストリーム５１７Ｂは、変換情報フィールド６７２（「変換情報６７２」）と、ＳＨＣ５１１’を記憶するフィールド（ここで、このフィールドは「ＳＨＣ５１１’」と示される）とを含む。変換情報６７２は、上で述べられたように、トランスレーション情報、回転情報、および／または音場への調整を示す任意の他の形態の情報を備え得る。いくつかの例では、変換情報６７２はまた、ビットストリーム５１７Ｂ内でＳＨＣ５１１’と指定されるＳＨＣ５１１の最高次を指定することもできる。すなわち、変換情報６７２は３の次数を示すことができ、抽出デバイスはこれを、ＳＨＣ５１１’が、ＳＨＣ５１１のうちで、ＳＨＣ５１１のうちで３の次数を有するものまでのものを含むことを示すこととして理解することができる。次いで、抽出デバイスは、４以上の次数を有するＳＨＣ５１１を０に設定し、それによって、ビットストリーム内の４以上の次数のＳＨＣ５１１の明示的なシグナリングを潜在的に除去するように構成

され得る。

【１３８１】

[1316]図５７Ｃの例では、ビットストリーム５１７Ｃは、上で図５３に示されたビットストリーム５１７の一例を表し得る。ビットストリーム５１７Ｃは、変換情報フィールド６７２（「変換情報６７２」）と、ＳＨＣ存在フィールド６７０と、ＳＨＣ５１１’を記憶するフィールド（ここで、このフィールドは「ＳＨＣ５１１’」と示される）とを含む。上で図５７Ｂに関して説明されたように、ＳＨＣ５１１のどの次数がシグナリングされないかを理解するように構成されるのではなく、ＳＨＣ存在フィールド６７０は、ＳＨＣ５１１のうちどれがビットストリーム５１７Ｃ内でＳＨＣ５１１’と指定されるかを明示的にシグナリングすることができる。

【１３８２】

[1317]図５７Ｄの例では、ビットストリーム５１７Ｄは、上で図５３に示されたビットストリーム５１７の一例を表し得る。ビットストリーム５１７Ｄは、次数フィールド６７４（「次数６０」）と、ＳＨＣ存在フィールド６７０と、方位角フラグ６７６（「ＡＺＦ６７６」）と、仰角フラグ６７８（「ＥＬＦ６７８」）と、方位角角度フィールド６８０（「方位角６８０」）と、仰角角度フィールド６８２（「仰角６８２」）と、ＳＨＣ５１１’を記憶するフィールド（ここで、この場合も、このフィールドは「ＳＨＣ５１１’」と示される）とを含む。次数フィールド６７４は、ＳＨＣ５１１’の次数、すなわち、音場を表すために使用される球面基底関数の最高次数に対して上記でｎによって示される次数を指定する。次数フィールド６７４は、８ビットフィールドであると示されているが、３（４次を指定するために必要とされるビットの数である）などの他の様々なビットサイズであってよい。ＳＨＣ存在フィールド６７０は、２５ビットフィールドと示されている。この場合も、しかしながら、ＳＨＣ存在フィールド６７０は、他の様々なビットサイズであってよい。ＳＨＣ存在フィールド６７０は、ＳＨＣ存在フィールド６７０が音場の４次表現に対応する球面調和係数の各々のための１ビットを含み得ることを示すために、２５ビットと示される。

【１３８３】

[1318]方位角フラグ６７６は、方位角フィールド６８０がビットストリーム５１７Ｄ内に存在するかどうかを指定する１ビットフラグを表す。方位角フラグ６７６が１に設定されるとき、ＳＨＣ５１１’のための方位角フィールド６８０がビットストリーム５１７Ｄ内に存在する。方位角フラグ６７６が０に設定されるとき、ＳＨＣ５１１’のための方位角フィールド６８０は、ビットストリーム５１７Ｄ内に存在しないかまたはさもなければ指定されない。同様に、仰角フラグ６７８は、仰角フィールド６８２がビットストリーム５１７Ｄ内に存在するかどうかを指定する１ビットフラグを表す。仰角フラグ６７８が１に設定されるとき、ＳＨＣ５１１’のための仰角フィールド６８２がビットストリーム５１７Ｄ内に存在する。仰角フラグ６７８が０に設定されるとき、ＳＨＣ５１１’のための仰角フィールド６８２は、ビットストリーム５１７Ｄ内に存在しないかまたはさもなければ指定されない。１は、対応するフィールドが存在することをシグナリングし、０は、対応するフィールドが存在しないことをシグナリングすると説明されているが、この規則は、０は、対応するフィールドがビットストリーム５１７Ｄ内で指定されていることを指定し、１は、対応するフィールドがビットストリーム５１７Ｄ内で指定されていないことを指定するように、逆にされてよい。したがって、本開示で説明される技法は、この点について限定されるべきではない。

【１３８４】

[1319]方位角フィールド６８０は、ビットストリーム５１７Ｄ内に存在するときに方位角角度を指定する１０ビットフィールドを表す。１０ビットフィールドとして示されているが、方位角フィールド６８０は他のビットサイズであってもよい。仰角フィールド６８２は、ビットストリーム５１７Ｄ内に存在するときに仰角角度を指定する９ビットフィールドを表す。フィールド６８０および６８２で指定される方位角角度および仰角角度はそれぞれ、フラグ６７６および６７８と連動して、上で説明された回転情報を表すことがで

きる。この回転情報は、オリジナルの基準フレームにおけるＳＨＣ５１１を復元するように音場を回転させるために使用され得る。

【１３８５】

[1320]ＳＨＣ５１１’フィールドは、Ｘのサイズである可変フィールドとして示されている。ＳＨＣ５１１’フィールドは、ＳＨＣ存在フィールド６７０によって示されるビットストリーム内で指定されるＳＨＣ５１１’の数により変化し得る。Ｘのサイズは、ＳＨＣ存在フィールド６７０内のＳＨＣ５１１’の数×３２ビット（各ＳＨＣ５１１’のサイズである）の関数として導出され得る。

【１３８６】

[1321]図５７Ｅの例では、ビットストリーム５１７Ｅは、上記で図５３に示されたビットストリーム５１７の別の例を表し得る。ビットストリーム５１７Ｅは、次数フィールド６７４（「次数６０」）と、ＳＨＣ存在フィールド６７０と、回転インデックスフィールド６８４と、ＳＨＣ５１１’を記憶するフィールド（ここで、この場合も、このフィールドは「ＳＨＣ５１１’」と示される）とを含む。次数フィールド６７４、ＳＨＣ存在フィールド６７０、およびＳＨＣ５１１’フィールドは、上で説明されたものに実質的に類似であり得る。回転インデックスフィールド６８４は、仰角角度と方位角角度の１０２４×５１２（すなわち、言い換えれば、５２４２８８）個の組合せのうちの１つを指定するために使用される２０ビットフィールドを表すことができる。いくつかの例では、この回転インデックスフィールド６８４を指定するために１９ビットのみが使用されることがあり、オーディオ符号化デバイス５７０は、回転演算が実行されたかどうか（および、したがって、回転インデックスフィールド６８４がビットストリーム内に存在するかどうか）を示すために、ビットストリーム内で追加のフラグを指定することができる。この回転インデックスフィールド６８４は、上で述べられた回転インデックスを指定し、回転インデックスは、オーディオ符号化デバイス５７０とビットストリーム抽出デバイスの両方に共通する回転テーブル内のエントリを指し得る。この回転テーブルは、いくつかの例では、方位角角度と仰角角度の異なる組合せを記憶することができる。代替的に、回転テーブルは、上で説明された行列を記憶することができ、この行列は、方位角角度と仰角角度の異なる組合せを行列形態で効果的に記憶する。

【１３８７】

[1322]図５８は、本開示において説明される技法の回転態様を実施する際の図５３の例に示されるオーディオ符号化デバイス５７０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５７０は、上で説明された様々な回転アルゴリズムのうちの１つまたは複数に従って方位角角度と仰角角度の組合せを選択することができる（８００）。次いで、オーディオ符号化デバイス５７０は、選択された方位角角度および仰角角度に従って音場を回転させることができる（８０２）。上で説明されたように、オーディオ符号化デバイス５７０は、上で述べられた$InvMat_1$を使用してＳＨＣ５１１から音場を最初に導出することができる。オーディオ符号化デバイス５７０はまた、回転された音場を表すＳＨＣ５１１’を決定することもできる（８０４）。別個のステップまたは動作であると説明されているが、オーディオ符号化デバイス５７０は、方位角角度と仰角角度の組合せの選択を表す変換（$[EncMat_2]$ $[InvMat_1]$の結果を表し得る）を適用し、ＳＨＣ５１１から音場を導出し、音場を回転させ、回転された音場を表すＳＨＣ５１１’を決定することができる。

【１３８８】

[1323]いずれにしても、オーディオ符号化デバイス５７０は次いで、閾値よりも大きい決定されたＳＨＣ５１１’の数を計算し、この数を、以前の方位角角度と仰角角度の組合せに対する以前の反復のために計算された数と比較することができる（８０６、８０８）。第１の方位角角度と仰角角度の組合せに対する第１の反復では、この比較は、あらかじめ定義された以前の数（０に設定し得る）に対するものであり得る。いずれにしても、ＳＨＣ５１１’の決定された数が以前の数よりも小さい場合（「はい」８０８）、オーディオ符号化デバイス５７０は、ＳＨＣ５１１’と、方位角角度と、仰角角度とを記憶し、多

くの場合、回転アルゴリズムの以前の反復から記憶された、以前のＳＨＣ５１１’と、方位角角度と、仰角角度とを置き換える（８１０）。

【１３８９】

[1324]ＳＨＣ５１１’の決定された数が以前の数よりも小さくない場合（「いいえ」８０８）、または以前に記憶されたＳＨＣ５１１’、方位角角度、および仰角角度の代わりにＳＨＣ５１１’と、方位角角度と、仰角角度とを記憶した後、オーディオ符号化デバイス５７０は、回転アルゴリズムが終了したかどうかを決定することができる（８１２）。すなわち、オーディオ符号化デバイス５７０は、一例として、方位角角度と仰角角度のすべての利用可能な組合せが評価されたかどうかを決定することができる。他の例では、オーディオ符号化デバイス５７０は、オーディオ符号化デバイス５７０が回転アルゴリズムを実行することを終了しているように、他の基準が満たされるかどうか（組合せの定義されたサブセットのすべてが実行された、所与の軌道が通られたかどうか、階層木が葉ノードまで通られたかどうかなど）を決定することができる。終了していない場合（「いいえ」８１２）、オーディオ符号化デバイス５７０は、別の選択された組合せに対して上記のプロセスを実行することができる（８００～８１２）。終了した場合（「はい」８１２）、オーディオ符号化デバイス５７０は、上で説明された様々な方法のうちの１つで、ビットストリーム５１７において、記憶されたＳＨＣ５１１’と、方位角角度と、仰角角度とを指定することができる（８１４）。

【１３９０】

[1325]図５９は、本開示において説明される技法の変換態様を実行する際の図５３の例に示されるオーディオ符号化デバイス５７０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５７０は、線形可逆変換を表す行列を選択することができる（８２０）。線形可逆変換を表す行列の一例は、［ＥｎｃＭａｔ$_1$］［ＩｎｃＭａｔ$_1$］の結果である、上で示された行列であり得る。次いで、オーディオ符号化デバイス５７０は、音場を変換するために、音場にこの行列を適用することができる（８２２）。オーディオ符号化デバイス５７０はまた、回転された音場を表すＳＨＣ５１１’を決定することもできる（８２４）。別個のステップまたは動作であると説明されているが、オーディオ符号化デバイス５７０は、変換（［ＥｎｃＭａｔ$_2$］［ＩｎｖＭａｔ$_1$］の結果を表し得る）を適用し、ＳＨＣ５１１から音場を導出し、音場を変換し、変換音場を表すＳＨＣ５１１’を決定することができる。

【１３９１】

[1326]いずれにしても、オーディオ符号化デバイス５７０は次いで、閾値よりも大きい決定されたＳＨＣ５１１’の数を計算し、この数を、変換行列の以前の適用に対する以前の反復のために計算された数と比較することができる（８２６、８２８）。ＳＨＣ５１１’の決定された数が以前の数よりも小さい場合（「はい」８２８）、オーディオ符号化デバイス５７０は、ＳＨＣ５１１’と、行列（または、行列と関連付けられるインデックスなど、その何らかの派生物）とを記憶し、多くの場合、回転アルゴリズムの以前の反復から記憶された、以前のＳＨＣ５１１’と行列（またはその派生物）とを置き換える（８３０）。

【１３９２】

[1327]ＳＨＣ５１１’の決定された数が以前の数よりも小さくない場合（「いいえ」８２８）、または以前に記憶されたＳＨＣ５１１’および行列の代わりに、ＳＨＣ５１１’と行列とを記憶した後、オーディオ符号化デバイス５７０は、変換アルゴリズムが終了したかどうかを決定することができる（８３２）。すなわち、オーディオ符号化デバイス５７０は、一例として、すべての利用可能な変換行列が評価されたかどうかを決定することができる。他の例では、オーディオ符号化デバイス５７０は、オーディオ符号化デバイス５７０が変換アルゴリズムを実行することを終了しているように、他の基準が満たされるかどうか（利用可能な変換行列の定義されたサブセットのすべてが実行された、所与の軌道が通られたかどうか、階層木が葉ノードまで通られたかどうかなど）を決定することができる。終了していない場合（「いいえ」８３２）、オーディオ符号化デバイス５７０は

、別の選択された変換行列に対して上記のプロセスを実行することができる（８２０～８３２）。終了した場合（「はい」８３２）、オーディオ符号化デバイス５７０は、上で説明された様々な方法のうちの１つで、ビットストリーム５１７において、記憶されたＳＨＣ５１１’と行列とを指定することができる（８３４）。

【１３９３】

[1328]いくつかの例では、変換アルゴリズムは、単一の反復を実行し、単一の変換行列を評価することができる。すなわち、変換行列は、線形可逆変換を表す任意の行列を備えることができる。いくつかの例では、線形可逆変換は、音場を空間領域から周波数領域に変換することができる。そのような線形可逆変換の例としては、離散フーリエ変換（ＤＦＴ）があり得る。ＤＦＴの適用は、単一の反復のみを伴うことがあり、したがって、変換アルゴリズムが終了されるかどうかを決定するステップを必ずしも含むとは限らない。したがって、本技法は、図５９の例に限定されるべきではない。

【１３９４】

[1329]言い換えれば、線形可逆変換の一例は離散フーリエ変換（ＤＦＴ）である。２５個のＳＨＣ５１１’は、２５個の複素係数のセットを形成するために、ＤＦＴによって操作され得る。オーディオ符号化デバイス５７０はまた、ＤＦＴのビンサイズの分解能を潜在的に増加させ、たとえば高速フーリエ変換（ＦＦＴ）を適用することを通じてＤＦＴのより効率的な実装形態を潜在的に有するように、２の倍数である整数になるように２５個のＳＨＣ５１１’をゼロパディングすることもできる。いくつかの例では、ＤＦＴの分解能を２５個の点を超えて増加させることは、必ずしも必要とされるとは限らない。変換領域では、オーディオ符号化デバイス５７０は、特定のビンにいずれかのスペクトルエネルギーが存在するかどうかを決定するために、閾値を適用することができる。オーディオ符号化デバイス５７０は、この文脈では、次いで、この閾値を下回るスペクトル係数エネルギーを廃棄またはゼロにすることができ、オーディオ符号化デバイス５７０は、廃棄されたまたはゼロにされたＳＨＣ５１１’のうちの１つまたは複数を有するＳＨＣ５１１’を復元するために逆変換を適用することができる。すなわち、逆変換が適用された後、閾値を下回る係数は存在せず、結果として、より少ないビットが、音場を符号化するために使用され得る。

【１３９５】

[1330]１つまたは複数の例では、説明された機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せで実施され得る。ソフトウェアで実施される場合、これらの機能は、コンピュータ可読媒体上に１つまたは複数の命令またはコードとして記憶または送信され、ハードウェアベースの処理ユニットによって実行され得る。コンピュータ可読媒体は、たとえば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を容易にする任意の媒体を含む、データ記憶媒体または通信媒体などの有形媒体に対応するコンピュータ可読記憶媒体を含み得る。このようにして、コンピュータ可読媒体は、一般に、（１）非一時的である有形のコンピュータ可読記憶媒体、または（２）信号もしくは搬送波などの通信媒体に対応し得る。データ記憶媒体は、本開示で説明される技法の実施のために、１つもしくは複数のコンピュータまたは１つもしくは複数のプロセッサによって、命令、コードおよび／またはデータ構造を取り出すためにアクセスされ得る、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。

【１３９６】

[1331]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記憶装置、磁気ディスク記憶装置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で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得る。また、任意の接続がコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、ウェブサイト、サーバ、または他の遠隔ソースから、同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者回線（

ＤＳＬ）、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して送信される場合、同軸ケーブル、光ファイバケーブル、ツイストペア、ＤＳＬ、または赤外線、無線、マイクロ波などのワイヤレス技術は、媒体の定義に含まれる。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まないが、代わりに非一時的な有形記憶媒体を対象とすることを理解されたい。本明細書において使用されるとき、ディスク（disk）およびディスク（disc）は、コンパクトディスク（ＣＤ）、レーザーディスク（登録商標）、光ディスク、デジタルバーサタイルディスク（ＤＶＤ）、フロッピー（登録商標）ディスクおよびブルーレイディスクを含み、この場合、ディスク（disk）は、通常、磁気的にデータを再生し、ディスク（disc）は、レーザーを用いてデータを光学的に再生する。上記の組合せも、コンピュータ可読媒体の範囲内に含められるべきである。

【１３９７】

[1332]命令は、１つまたは複数のデジタル信号プロセッサ（ＤＳＰ）、汎用マイクロプロセッサ、特定用途向け集積回路（ＡＳＩＣ）、フィールドプログラマブルロジックアレイ（ＦＰＧＡ）、または他の同等の統合された、もしくは個別の論理回路など、１つまたは複数のプロセッサによって実行され得る。したがって、本明細書で使用される「プロセッサ」という用語は、前述の構造または本明細書で説明される技法の実装形態に適した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る機能は、符号化および復号のために構成された専用のハードウェアモジュールおよび／またはソフトウェアモジュール内で提供されるか、あるいは複合コーデックに組み込まれ得る。また、本技法は、１つまたは複数の回路または論理要素において完全に実施され得る。

【１３９８】

[1333]本開示の技法は、ワイヤレスハンドセット、集積回路（ＩＣ）、またはＩＣのセット（たとえば、チップセット）を含む多種多様なデバイスまたは装置において実施され得る。様々な構成要素、モジュール、またはユニットが、開示された技法を実行するように構成されるデバイスの機能的態様を強調するために本開示で説明されているが、必ずしも異なるハードウェアユニットによる実現を必要とするとは限らない。むしろ、上で説明されたように、様々なユニットは、コーデックハードウェアユニットの中で組み合わせられ、または、上で説明された１つもしくは複数のプロセッサを含む、適切なソフトウェアおよび／もしくはファームウェアとともに相互動作可能なハードウェアユニットの集合によって提供され得る。

【１３９９】

[1334]本技法の様々な実施形態が説明された。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、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］

音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを備え、前記空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、

方法。

［Ｃ２］

前記量子化ステップサイズを決定することは、目標ビットレートに基づいて前記量子化ステップサイズを決定することを備える、

Ｃ１に記載の方法。

［Ｃ３］

前記複数の量子化ステップサイズのうちの１つを選択することは、

前記空間成分を表すために使用されるビットの数の推定値を決定することと、

前記推定値と目標ビットレートとの間の差に基づいて前記量子化ステップサイズを決定することと

を備える、Ｃ１に記載の方法。
［Ｃ４］
前記複数の量子化ステップサイズのうちの１つを選択することは、
前記空間成分を表すために使用されるビットの数の推定値を決定することと、
前記推定値と目標ビットレートとの間の差を決定することと、
前記差を前記目標ビットレートに加算することによって前記量子化ステップサイズを決定することと
を備える、Ｃ１に記載の方法。
［Ｃ５］
前記ビットの数の前記推定値を決定することは、前記目標ビットレートに対応するコードブックを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定されたものを計算することを備える、
Ｃ３に記載の方法。
［Ｃ６］
前記ビットの数の前記推定値を決定することは、前記空間成分を圧縮するときに使用されるコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定されたものを計算することを備える、
Ｃ３に記載の方法。
［Ｃ７］
前記ビットの数の前記推定値を決定することは、
前記空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第１の推定値を計算することと、
前記空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の第２の推定値を計算することと、
前記第１の推定値と前記第２の推定値のうちで、前記ビットの数の前記決定された推定値として使用されるべき、より少数のビットを有するものを選択することと
を備える、Ｃ３に記載の方法。
［Ｃ８］
前記ビットの数の前記推定値を決定することは、
前記空間成分が対応するカテゴリを識別するカテゴリ識別子を識別することと、
前記カテゴリに対応する前記空間成分を圧縮するときに生じる、前記空間成分の残差値のビット長を識別することと、
少なくとも一部、前記カテゴリ識別子を表すために使用されるビットの数を前記残差値の前記ビット長に加算することによって、前記ビットの数の前記推定値を決定することと
を備える、Ｃ３に記載の方法。
［Ｃ９］
前記空間成分を圧縮するときに使用されるべき、複数のコードブックのうちの１つを選択することをさらに備える、
Ｃ１に記載の方法。
［Ｃ１０］
前記複数のコードブックのうちの１つを選択することは、
前記複数のコードブックの各々を使用して、前記空間成分を表すために使用されるビットの数の推定値を決定することと、
最小の数のビットを有する、前記決定された推定値をもたらした前記複数のコードブックのうちの前記１つを選択することと
を備える、Ｃ９に記載の方法。
［Ｃ１１］
前記複数のコードブックのうちの１つを選択することは、前記複数のコードブックのうちの１つまたは複数を使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを備え、前記複数のコードブックの前記１つまたは複数は、前記空間成分

の他の要素に対する圧縮されるべき前記空間成分の要素の次数に基づいて選択される、
Ｃ９に記載の方法。
［Ｃ１２］
前記複数のコードブックのうちの１つを選択することは、前記空間成分が後続の空間成分から予測されないときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを備える、
Ｃ９に記載の方法。
［Ｃ１３］
前記複数のコードブックのうちの１つを選択することは、前記空間成分が後続の空間成分から予測されるときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを備える、
Ｃ９に記載の方法。
［Ｃ１４］
前記複数のコードブックのうちの１つを選択することは、前記空間成分が前記音場の中の合成オーディオオブジェクトを表すときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを備える、
Ｃ９に記載の方法。
［Ｃ１５］
前記合成オーディオオブジェクトは、パルスコード変調（ＰＣＭ）オーディオオブジェクトを備える、
Ｃ９に記載の方法。
［Ｃ１６］
前記複数のコードブックのうちの１つを選択することは、前記空間成分が前記音場の中の録音されたオーディオオブジェクトを表すときに使用されるように設計される前記複数のコードブックの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを備える、
Ｃ９に記載の方法。
［Ｃ１７］
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを行うように構成される１つまたは複数のプロセッサを備え、前記空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、
デバイス。
［Ｃ１８］
前記１つまたは複数のプロセッサは、目標ビットレートに基づいて前記量子化ステップサイズを決定することを行うように構成される、
Ｃ１７に記載のデバイス。
［Ｃ１９］
前記１つまたは複数のプロセッサは、前記空間成分を表すために使用されるビットの数の推定値を決定することと、前記推定値と目標ビットレートとの間の差に基づいて前記量子化ステップサイズを決定することとを行うように構成される、
Ｃ１７に記載のデバイス。
［Ｃ２０］
前記１つまたは複数のプロセッサは、前記空間成分を表すために使用されるビットの数の推定値を決定することと、前記推定値と目標ビットレートとの間の差を決定することと、前記差を前記目標ビットレートに加算することによって前記量子化ステップサイズを決定することとを行うように構成される、
Ｃ１７に記載のデバイス。

［Ｃ２１］
前記１つまたは複数のプロセッサは、前記目標ビットレートに対応するコードブックを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定されたものを計算することを行うように構成される、
Ｃ１９に記載のデバイス。
［Ｃ２２］
前記１つまたは複数のプロセッサは、前記空間成分を圧縮するときに使用されるコーディングモードを仮定して前記空間成分のために生成されるべき前記ビットの数の前記推定されたものを計算するように構成される、
Ｃ１９に記載のデバイス。
［Ｃ２３］
前記１つまたは複数のプロセッサは、前記空間成分を圧縮するときに使用されるべき第１のコーディングモードを仮定して前記空間成分のために生成されるべきビットの前記数の第１の推定値を計算することと、前記空間成分を圧縮するときに使用されるべき第２のコーディングモードを仮定して前記空間成分のために生成されるべきビットの前記数の第２の推定値を計算することと、前記第１の推定値と前記第２の推定値のうちで、ビットの前記数の前記決定された推定値として使用されるべき、より少数のビットを有するものを選択することを行うように構成される、
Ｃ１９に記載のデバイス。
［Ｃ２４］
前記１つまたは複数のプロセッサは、前記空間成分が対応するカテゴリを識別するカテゴリ識別子を識別することと、前記カテゴリに対応する前記空間成分を圧縮するときに生じる、前記空間成分の残差値のビット長を識別することと、少なくとも一部、前記カテゴリ識別子を表すために使用されるビットの数を前記残差値の前記ビット長に加算することによって、前記ビットの数の前記推定値を決定することとを行うように構成される、
Ｃ１９に記載のデバイス。
［Ｃ２５］
前記１つまたは複数のプロセッサは、前記空間成分を圧縮するときに使用されるべき、複数のコードブックのうちの１つを選択することを行うようにさらに構成される、
Ｃ１７に記載のデバイス。
［Ｃ２６］
前記１つまたは複数のプロセッサは、前記複数のコードブックの各々を使用して、前記空間成分を表すために使用されるビットの数の推定値を決定することと、最小の数のビットを有する、前記決定された推定値をもたらした前記複数のコードブックの前記１つを選択することとを行うようにさらに構成される、
Ｃ２５に記載のデバイス。
［Ｃ２７］
前記１つまたは複数のプロセッサは、前記複数のコードブックのうちの１つまたは複数を使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを行うようにさらに構成され、前記複数のコードブックの前記１つまたは複数は、前記空間成分の他の要素に対する圧縮されるべき前記空間成分の要素の次数に基づいて選択される、
Ｃ２５に記載のデバイス。
［Ｃ２８］
前記１つまたは複数のプロセッサは、前記空間成分が後続の空間成分から予測されないときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを行うようにさらに構成される、
Ｃ２５に記載のデバイス。
［Ｃ２９］
前記１つまたは複数のプロセッサは、前記空間成分が後続の空間成分から予測されると

きに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを行うようにさらに構成される、
Ｃ２５に記載のデバイス。
［Ｃ３０］
前記１つまたは複数のプロセッサは、前記空間成分が前記音場の中の合成オーディオオブジェクトを表すときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定することを行うようにさらに構成される、
Ｃ２５に記載のデバイス。
［Ｃ３１］
前記合成オーディオオブジェクトは、パルスコード変調（ＰＣＭ）オーディオオブジェクトを備える、
Ｃ２５に記載のデバイス。
［Ｃ３２］
前記１つまたは複数のプロセッサは、前記空間成分が前記音場の中の録音されたオーディオオブジェクトを表すときに使用されるように設計される前記複数のコードブックのうちの１つを使用して前記空間成分を表すために使用されるビットの数の推定値を決定するようにさらに構成される、
Ｃ２５に記載のデバイス。
［Ｃ３３］
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定するための手段、前記空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、と、
前記量子化ステップサイズを記憶するための手段と
を備える、デバイス。
［Ｃ３４］
実行されると、１つまたは複数のプロセッサに、
音場の空間成分を圧縮するときに使用されるべき量子化ステップサイズを決定することを行わせる命令を記憶しており、前記空間成分は、複数の球面調和係数に関してベクトルベースの合成を実行することによって生成される、
非一時的コンピュータ可読記憶媒体。

【図１】

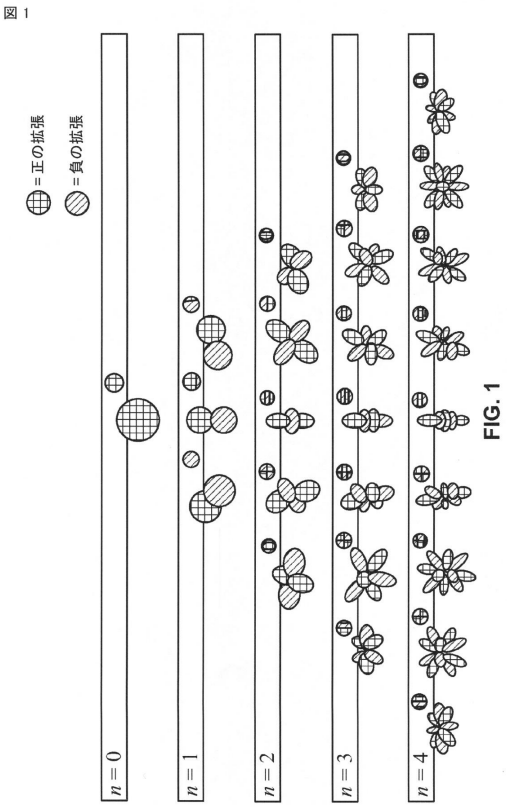


FIG. 1

【図２】

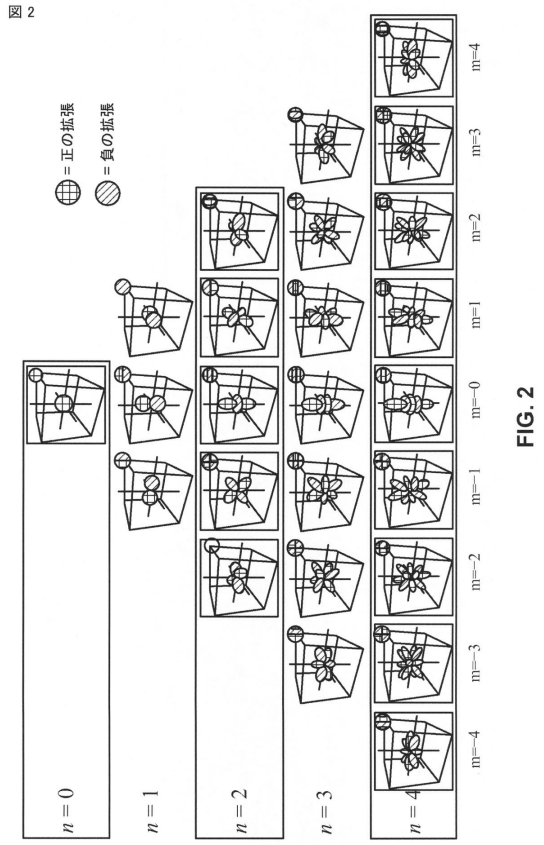


FIG. 2

【図３】

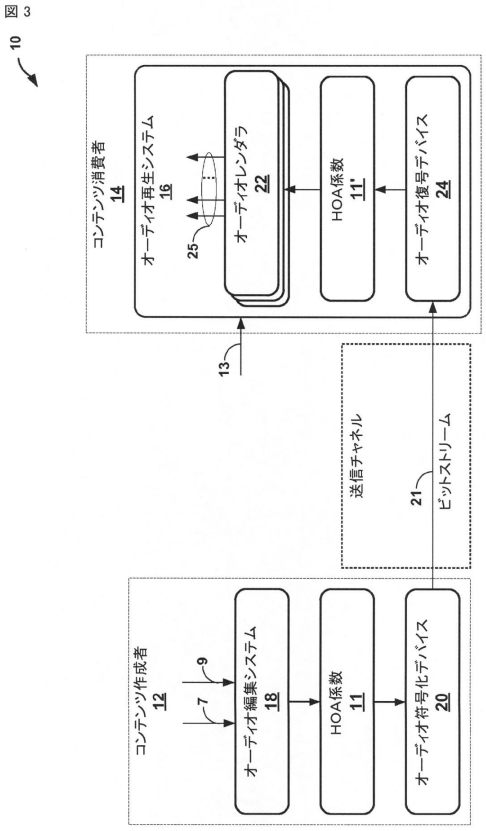


FIG. 3

【図４】

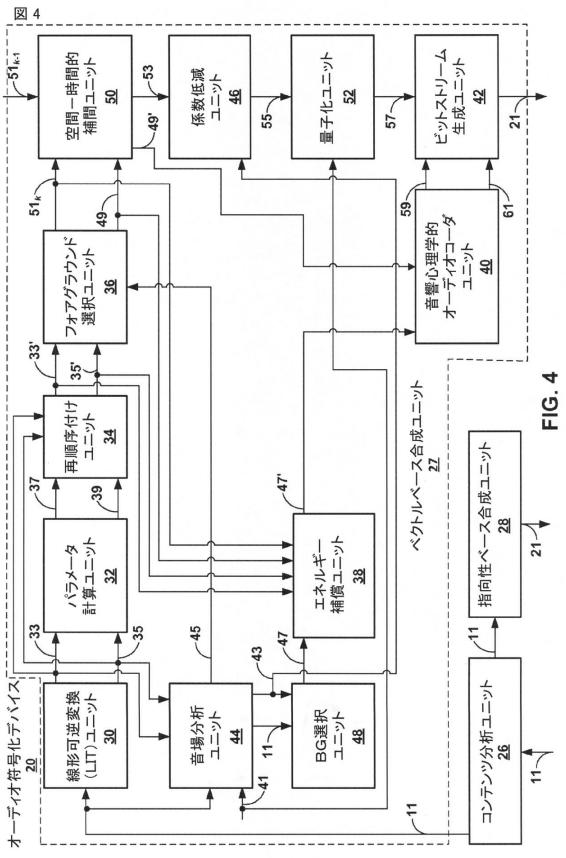


FIG. 4

【図５】

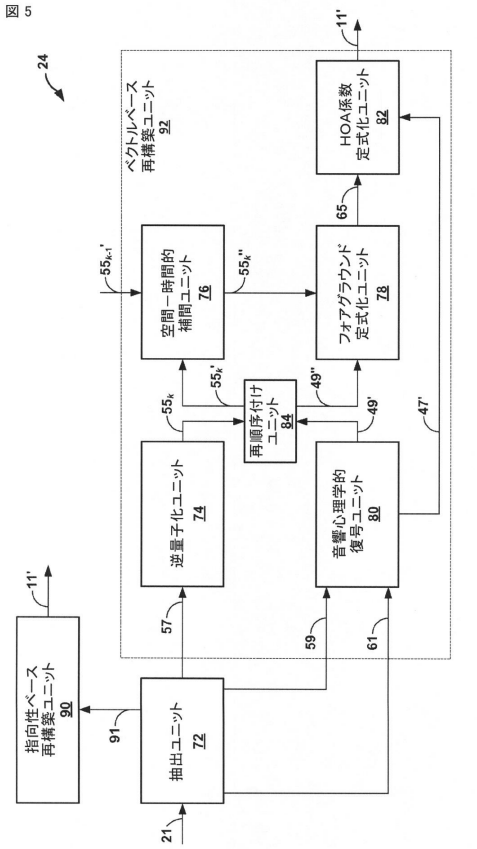


FIG. 5

【図６】

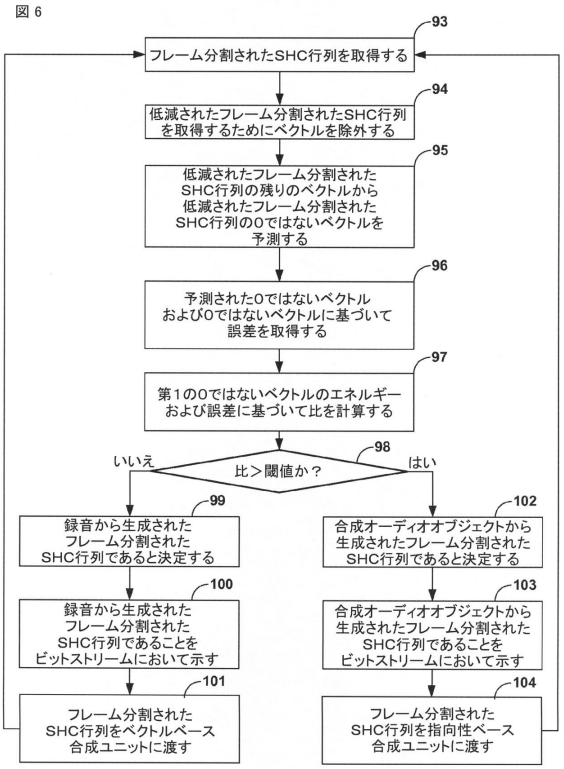


FIG. 6

【図７】

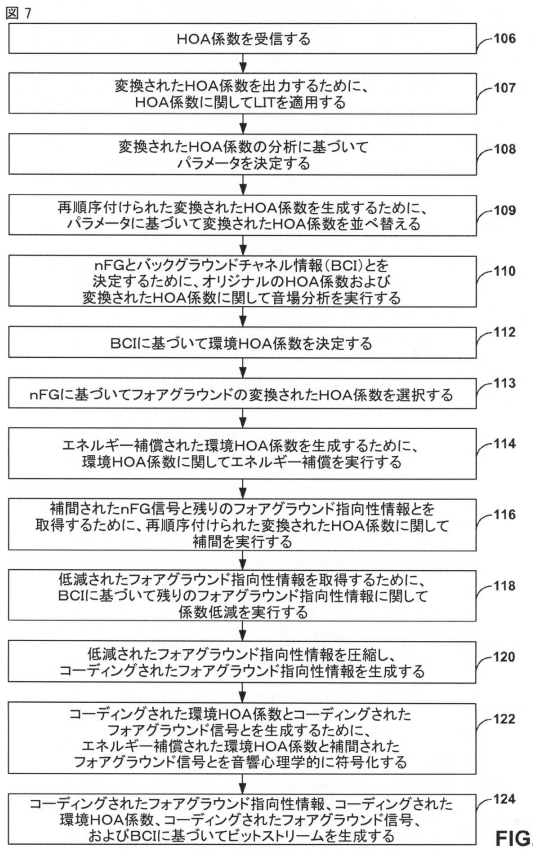


FIG. 7

【図８】

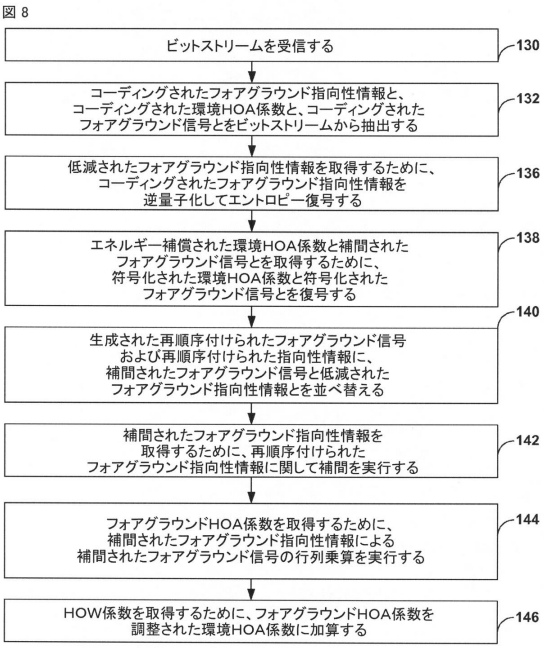


FIG. 8

【図９Ａ】

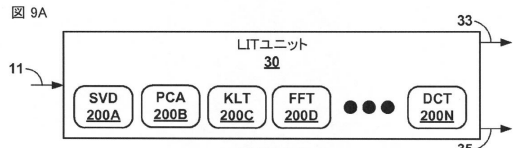


FIG. 9A

【図９Ｂ】

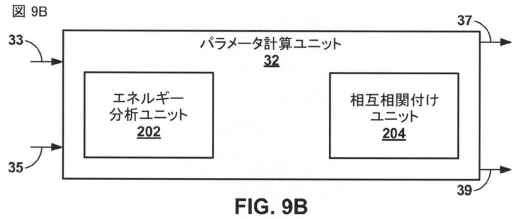


FIG. 9B

【図９Ｃ】

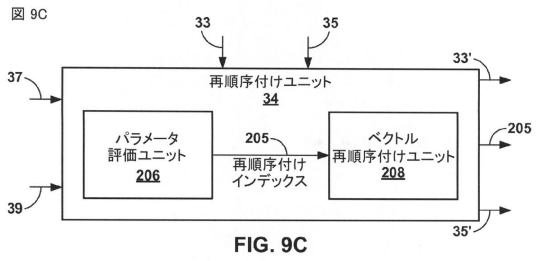


FIG. 9C

【図９Ｄ】

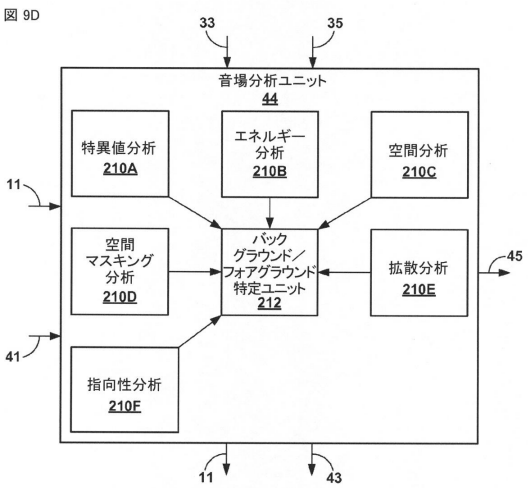


FIG. 9D

【図９Ｅ】

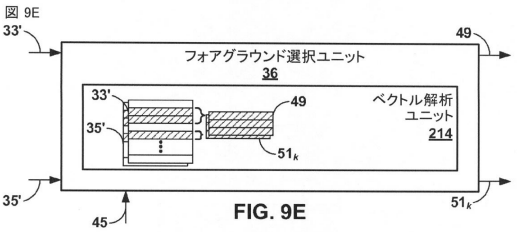


FIG. 9E

【図９Ｆ】

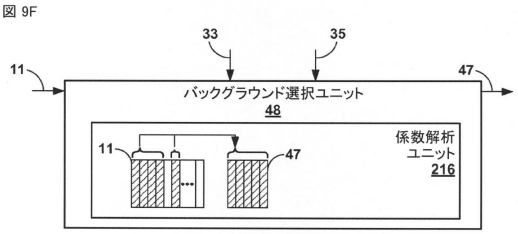


FIG. 9F

【図９Ｇ】

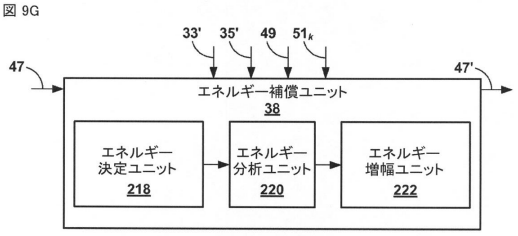


FIG. 9G

【図９Ｈ】

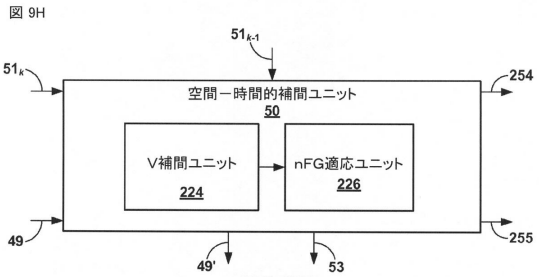


FIG. 9H

【図９Ｉ】

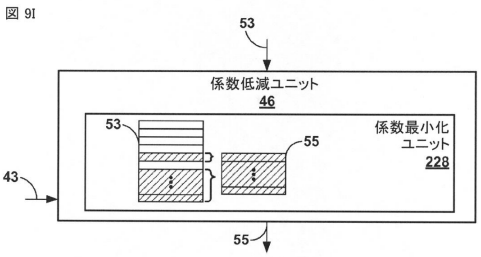


FIG. 9I

【図９Ｊ】

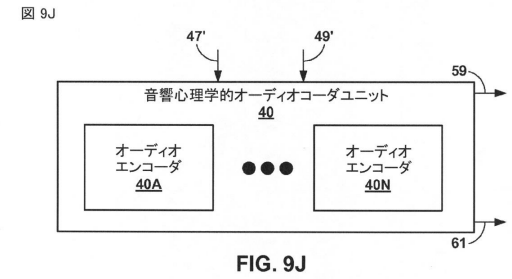


FIG. 9J

【図９Ｋ】

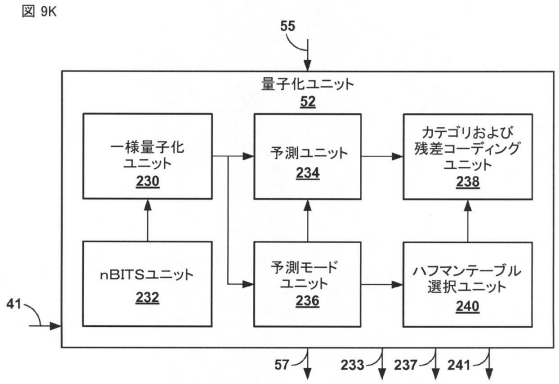


FIG. 9K

【図９Ｌ】

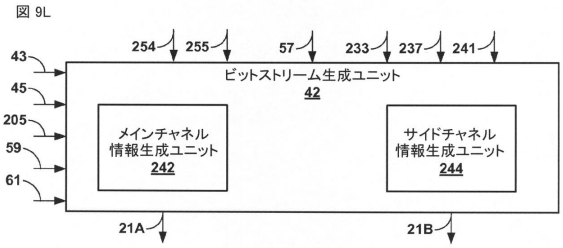


FIG. 9L

【図１０Ａ】

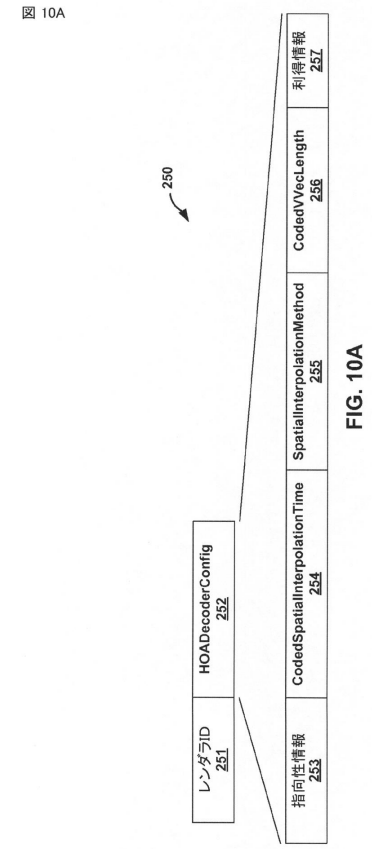


FIG. 10A

【図１０Ｂ】

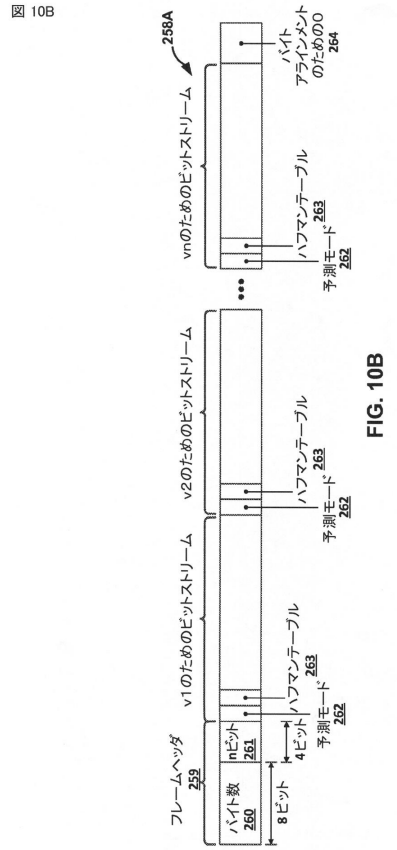


FIG. 10B

【図１０Ｃ】

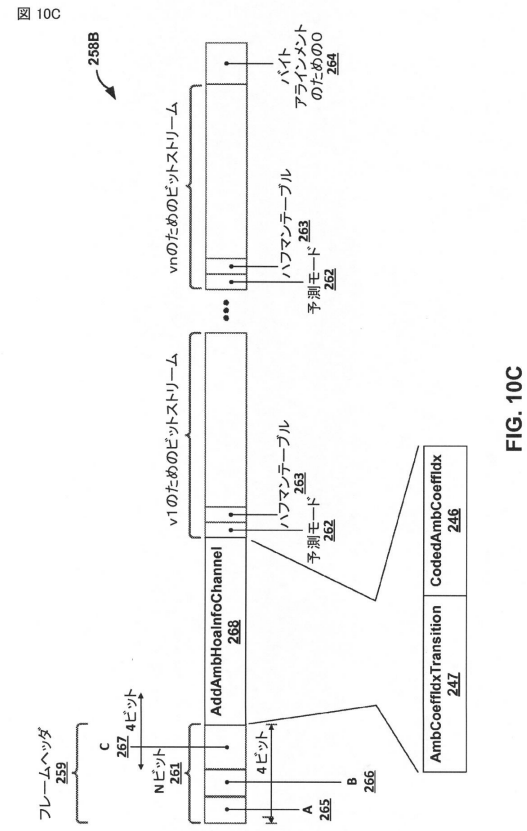


FIG. 10C

【図１０Ｃ（ｉ）】

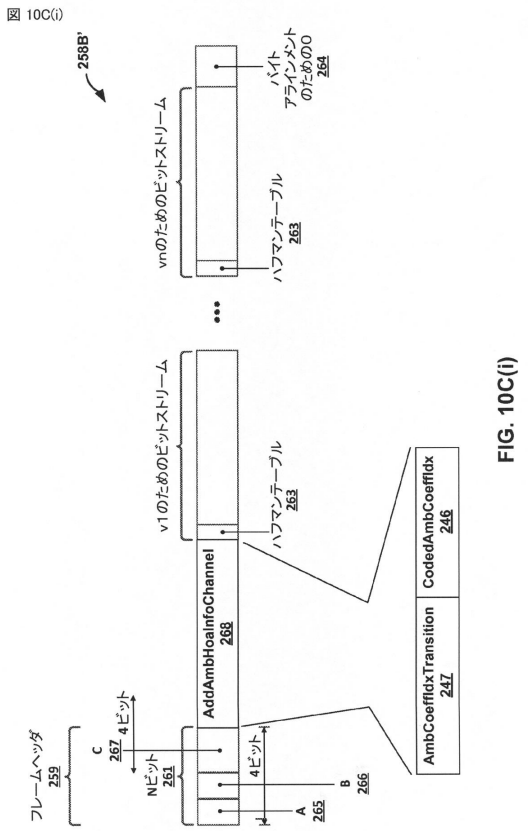


FIG. 10C(i)

【図１０Ｄ】

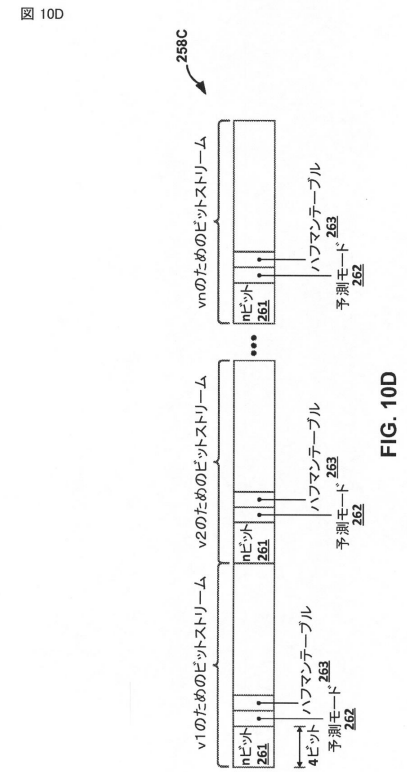


FIG. 10D

【図１０Ｄ（ｉ）】

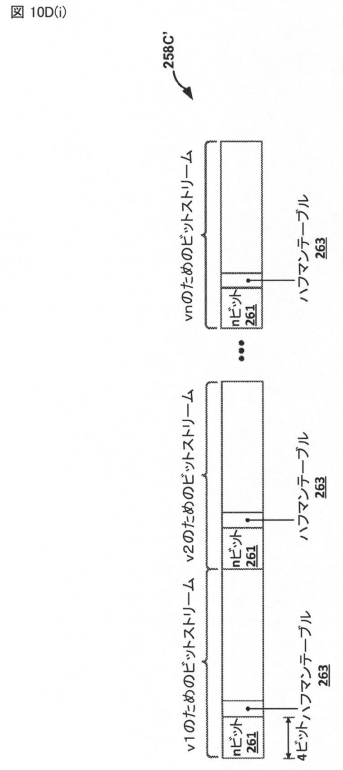


FIG. 10D(i)

【図１０Ｅ】

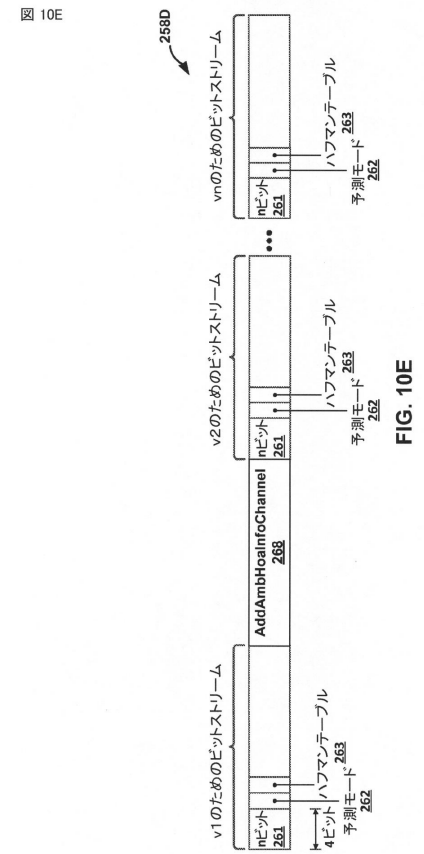


FIG. 10E

【図１０Ｅ（ｉ）】

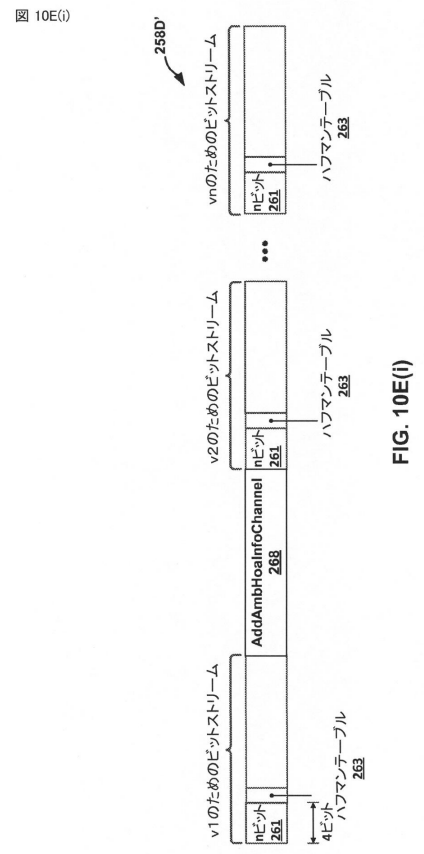


FIG. 10E(i)

【図１０Ｆ】

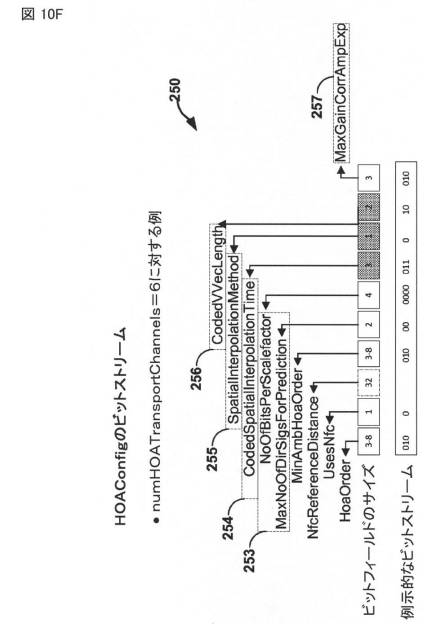


FIG. 10F

【図１０Ｇ】

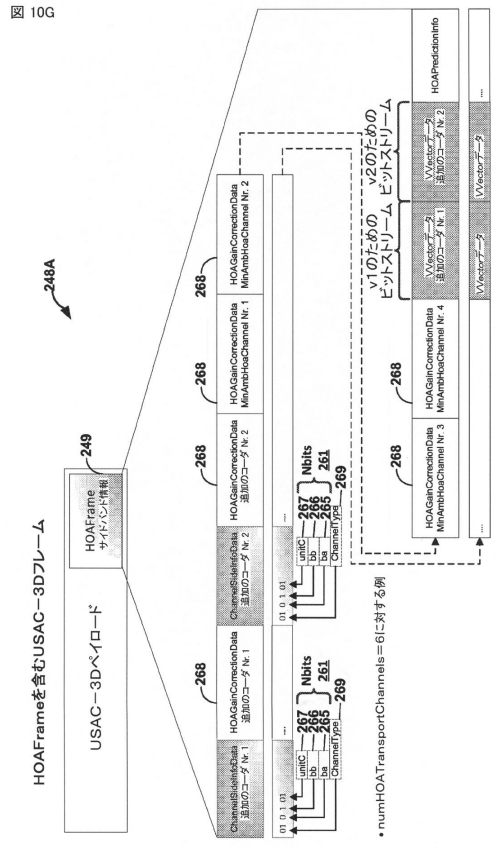


FIG. 10G

【図１０Ｈ（ｉ）】

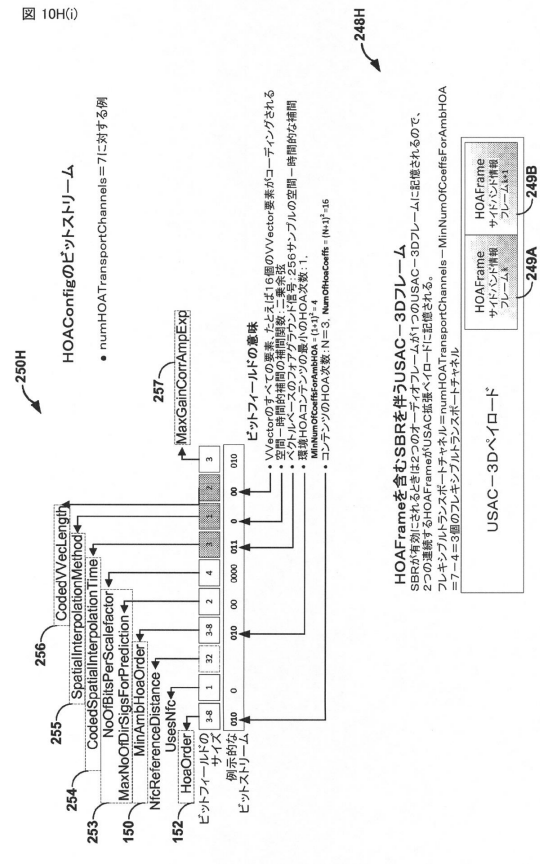


FIG. 10H(i)

【図１０Ｈ（ｉｉ）】

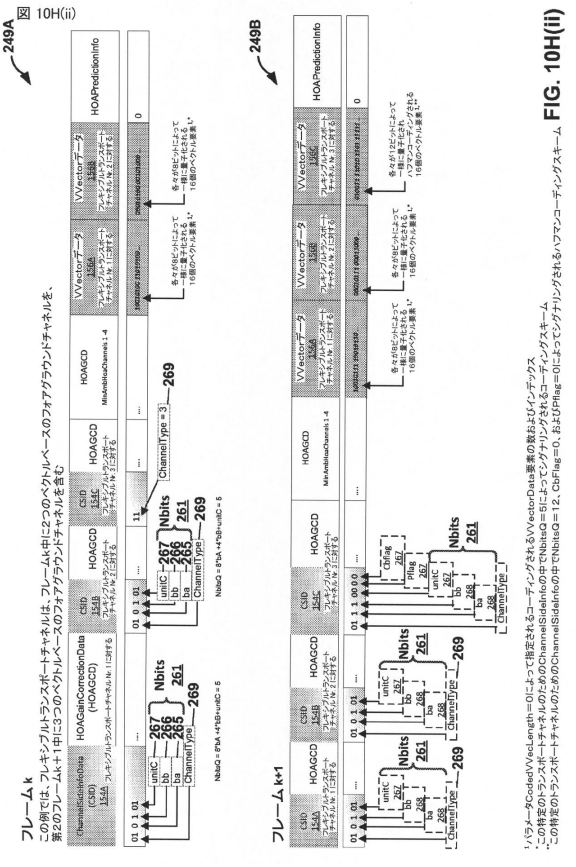


FIG. 10H(ii)

【図１０Ｉ（ｉ）】

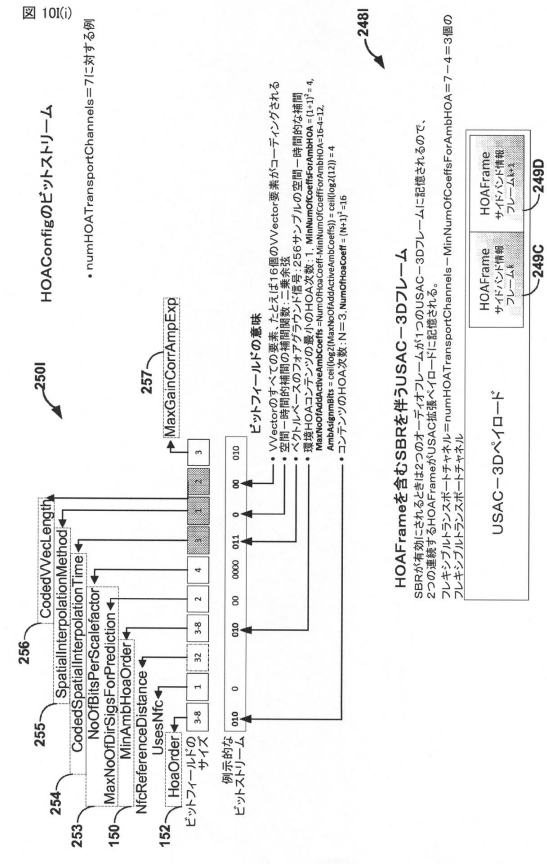


FIG. 10I(i)

【図１０Ｉ（ｉｉ）】

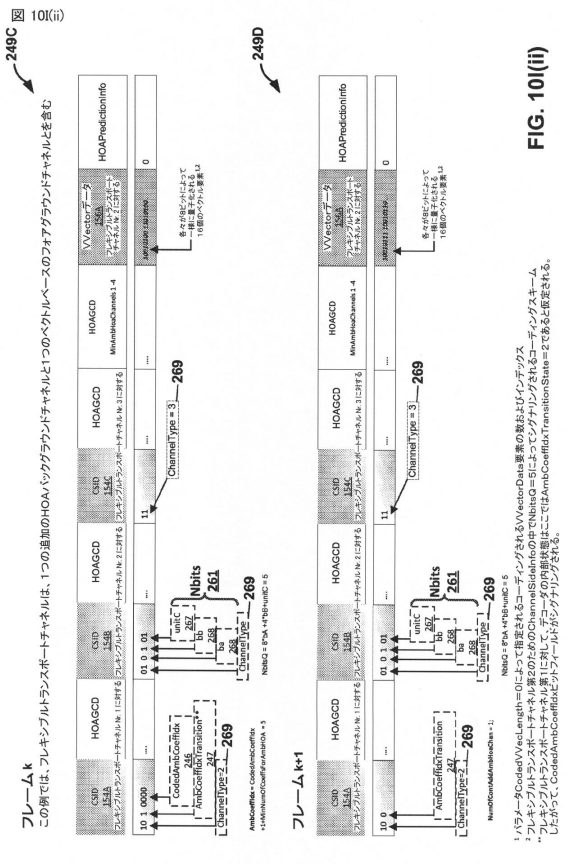


FIG. 10I(ii)

【図１０Ｋ（ｉ）】

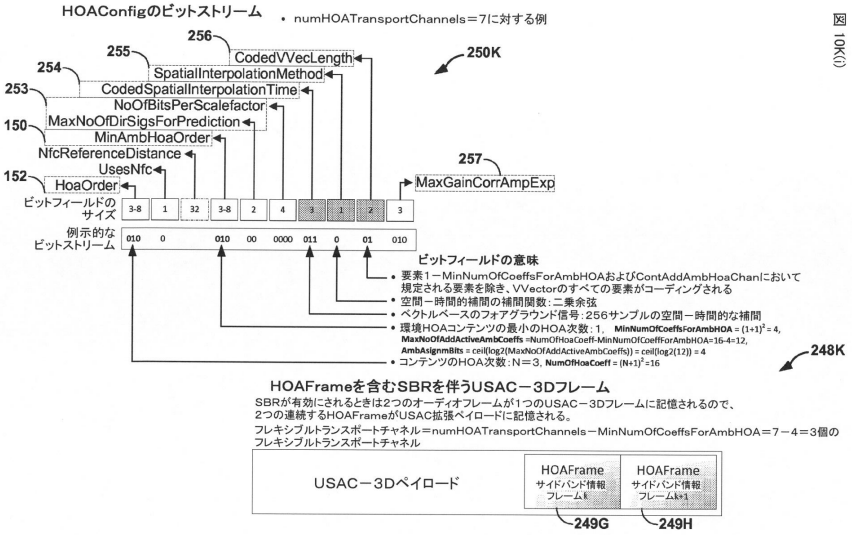
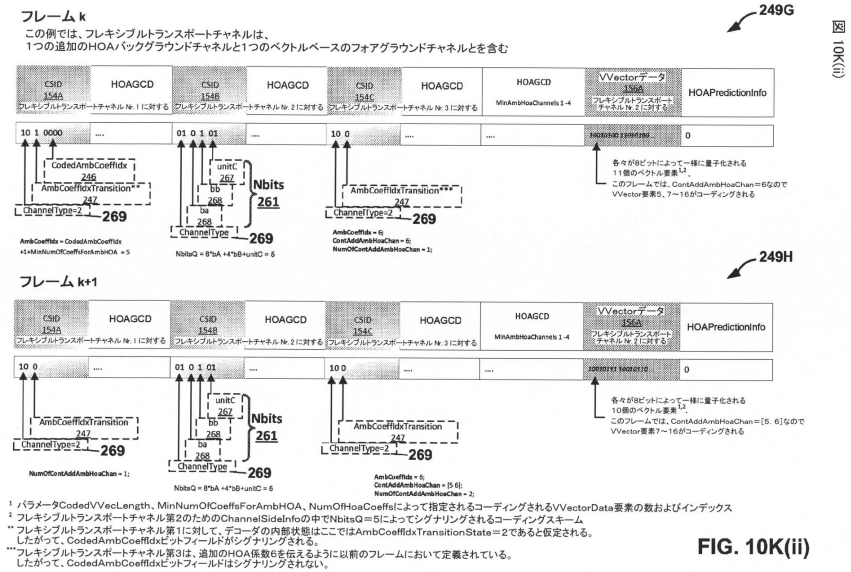


FIG. 10K(i)

【図１０Ｋ（ｉｉ）】



[1] パラメータCodedVVecLength、MinNumOfCoeffsForAmbHOA、NumOfHoaCoeffsによって指定されるコーディングされるVVectorData要素の数およびインデックス
[2] フレキシブルトランスポートチャネル第2のためのChannelSideInfoの中でNbitsQ＝5によってシグナリングされるコーディングスキーム
** フレキシブルトランスポートチャネル第1に対して、デコーダの内部状態はここではAmbCoeffIdxTransitionState＝2であると仮定される。
したがって、CodedAmbCoeffIdxビットフィールドがシグナリングされる。
***フレキシブルトランスポートチャネル第3は、追加のHOA係数6を伝えるように以前のフレームにおいて定義されている。
したがって、CodedAmbCoeffIdxビットフィールドはシグナリングされない。

FIG. 10K(ii)

【図１０Ｊ（ｉ）】

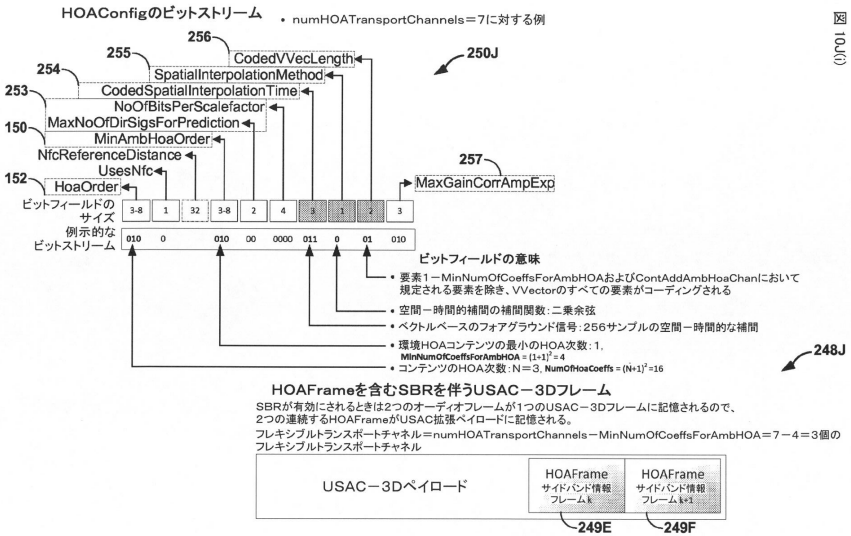
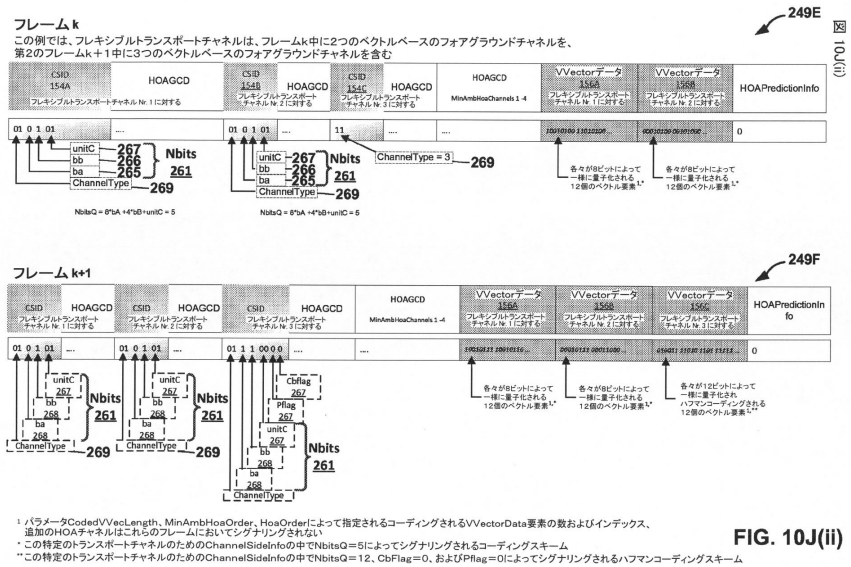


FIG. 10J(i)

【図１０Ｊ（ｉｉ）】



[1] パラメータCodedVVecLength、MinAmbHoaOrder、HoaOrderによって指定されるコーディングされるVVectorData要素の数およびインデックス、
追加のHOAチャネルはこれらのフレームにおいてシグナリングされない
* この特定のトランスポートチャネルのためのChannelSideInfoの中でNbitsQ＝5によってシグナリングされるコーディングスキーム
** この特定のトランスポートチャネルのためのChannelSideInfoの中でNbitsQ＝12、CbFlag＝0、およびPflag＝0によってシグナリングされるハフマンコーディングスキーム

FIG. 10J(ii)

【図１０Ｌ（ｉ）】

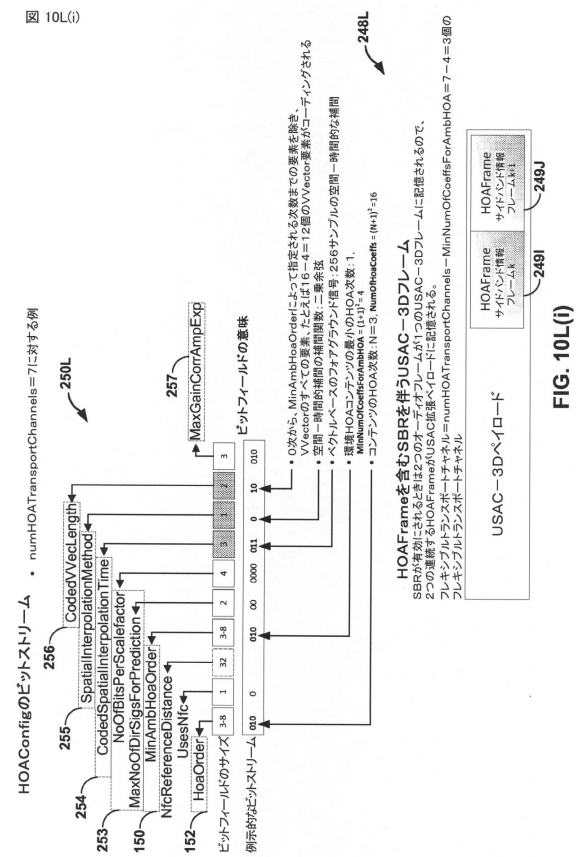


FIG. 10L(i)

【図１０Ｌ（ｉｉ）】

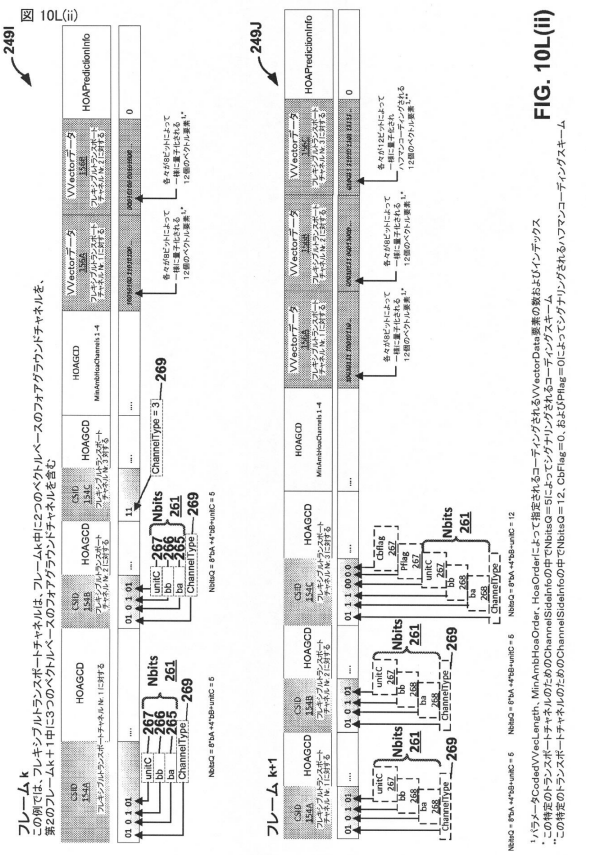


FIG. 10L(ii)

【図１０Ｍ（ｉ）】

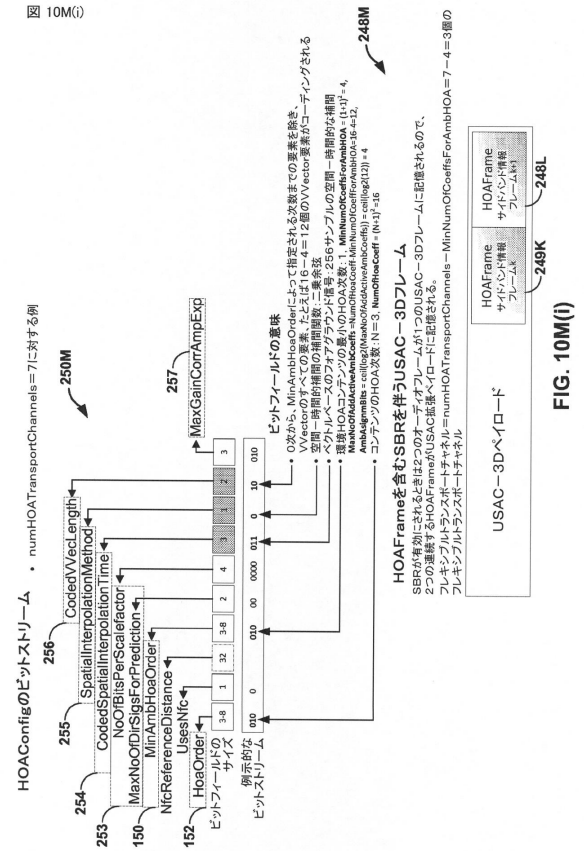


FIG. 10M(i)

【図１０Ｍ（ｉｉ）】

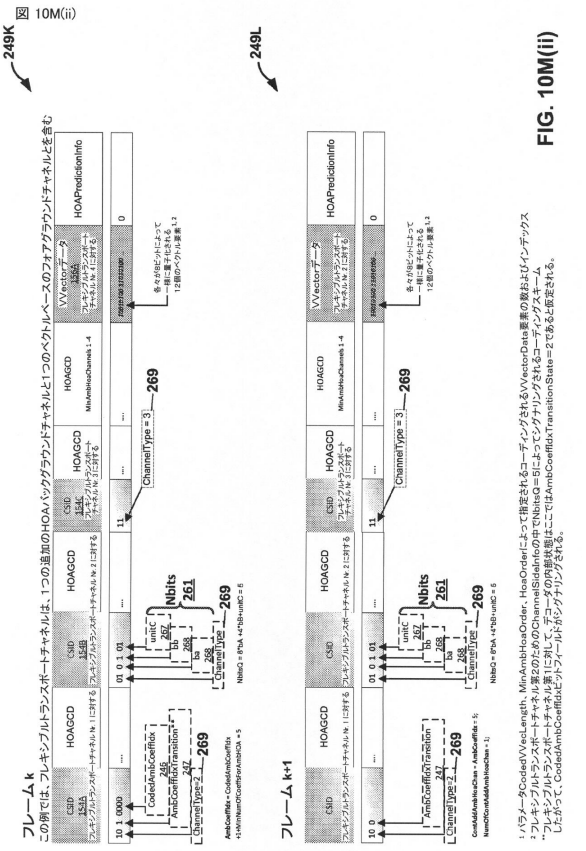


FIG. 10M(ii)

【図１０Ｎ（ｉ）】

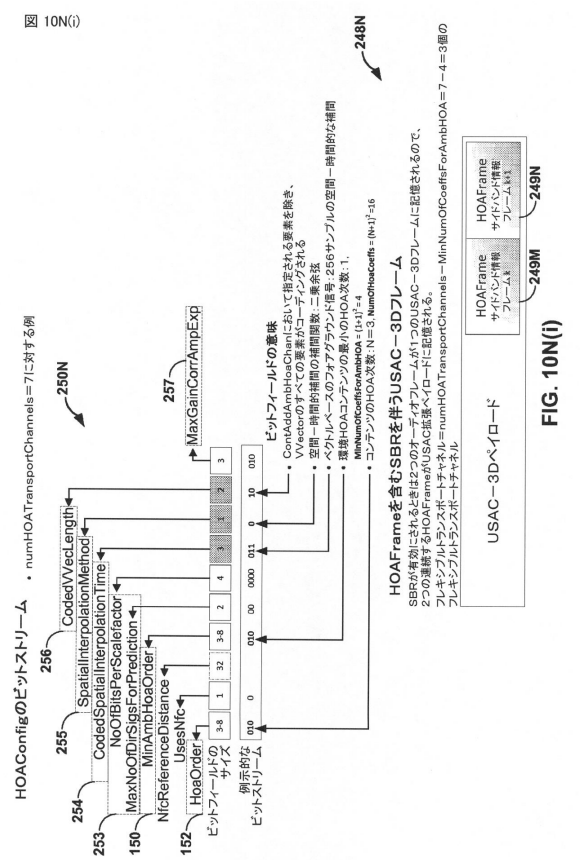
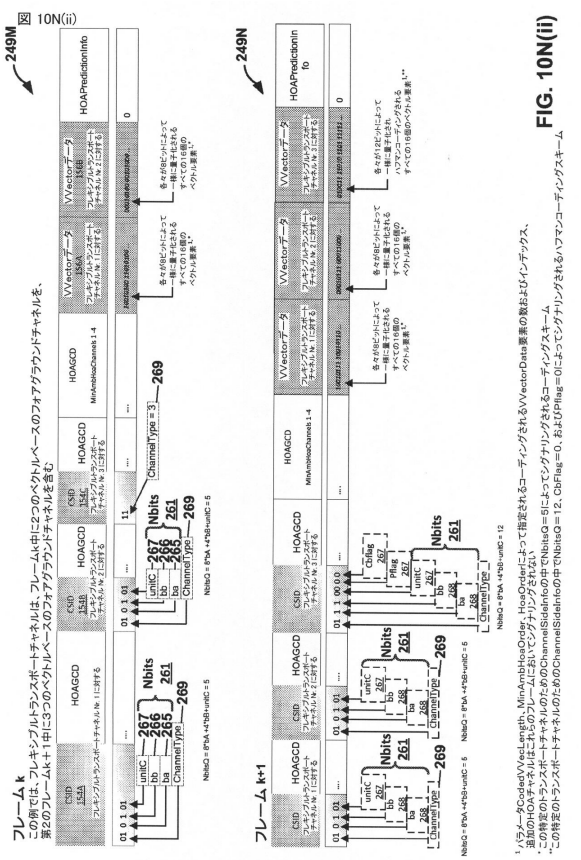


FIG. 10N(i)

【図１０Ｎ（ｉｉ）】



[1] パラメータCodedVVecLength、MinNumOfCoeffsForAmbHOA、NumOfHoaCoeffによって指定されるコーディングされるVVectorData要素の数およびインデックス
[2] フレキシブルトランスポートチャネル第2のためのChannelSideInfoの中でNbitsQ＝5によってシグナリングされるコーディングスキーム
[3] フレキシブルトランスポートチャネル第1に対して、データの内部状態はここではAmbCoeffIdxTransitionState＝2であると仮定される。
したがって、CodedAmbCoeffIdxビットフィールドがシグナリングされる。
[4] フレキシブルトランスポートチャネルは、追加のHOA係数を伝えるように以前のフレームにおいて定義されている。
したがって、CodedAmbCoeffIdxビットフィールドがシグナリングされない。

FIG. 10N(ii)

【図１００（ｉ）】

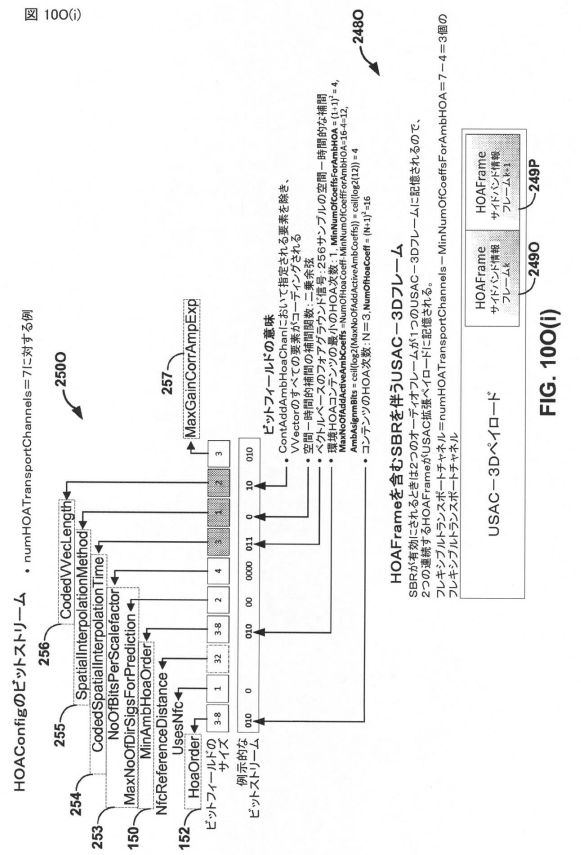
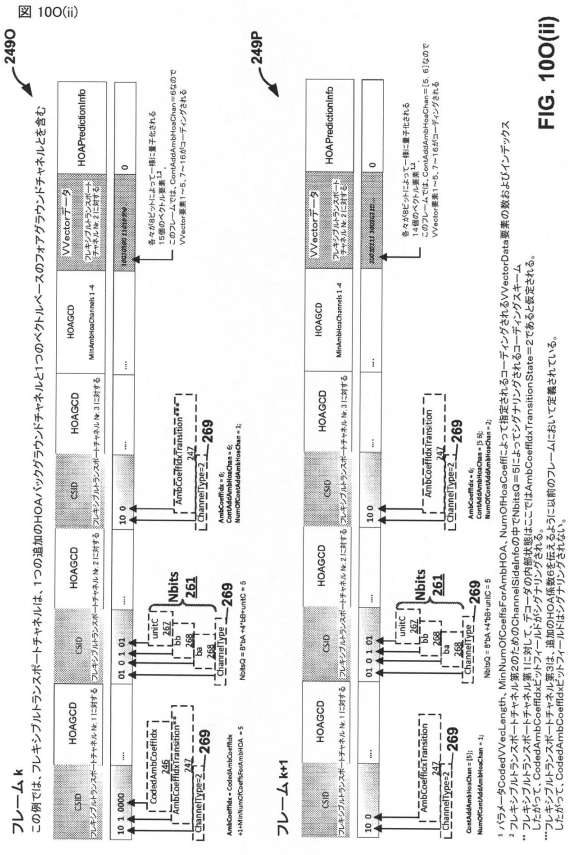


FIG. 10O(i)

【図１００（ｉｉ）】



[1] パラメータCodedVVecLength、MinNumOfCoeffsによって指定されるコーディングされるVVectorData要素の数およびインデックス
[2] フレキシブルトランスポートチャネル第2のためのChannelSideInfoの中でNbitsQ＝5によってシグナリングされるコーディングスキーム
[3] フレキシブルトランスポートチャネル第1に対して、データの内部状態はここではAmbCoeffIdxTransitionState＝2であると仮定される。
したがって、CodedAmbCoeffIdxビットフィールドがシグナリングされる。
[4] フレキシブルトランスポートチャネルは、追加のHOA係数を伝えるように以前のフレームにおいて定義されている。
したがって、CodedAmbCoeffIdxビットフィールドがシグナリングされない。

FIG. 10O(ii)

【図１１Ａ】

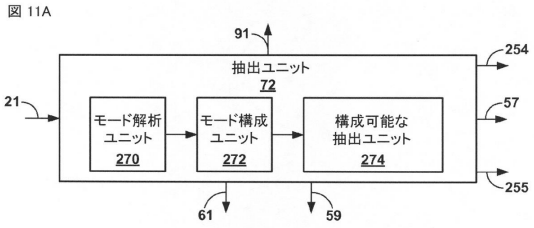


FIG. 11A

【図１１Ｂ】

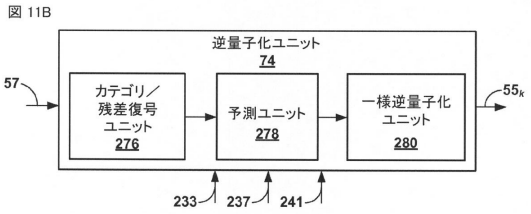


FIG. 11B

【図１１Ｃ】

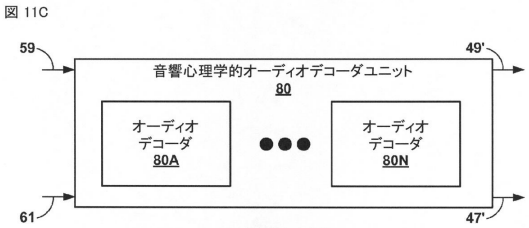


FIG. 11C

【図１１Ｄ】

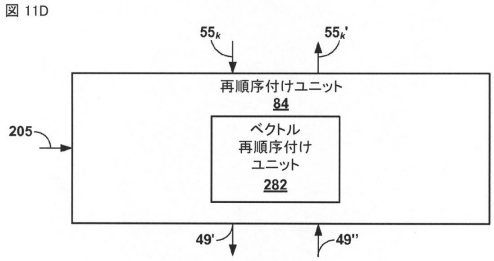


FIG. 11D

【図１１Ｅ】

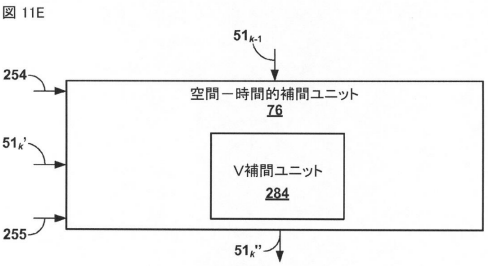


FIG. 11E

【図１１Ｇ】

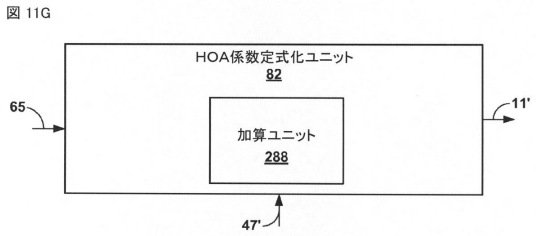


FIG. 11G

【図１１Ｆ】

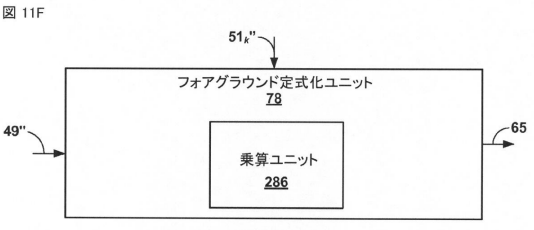


FIG. 11F

【図１２】

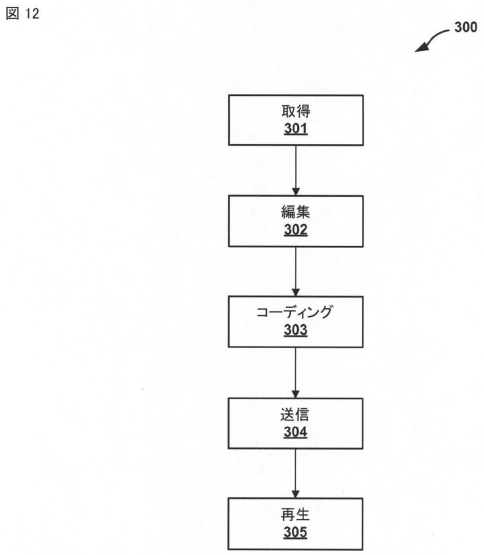


FIG. 12

【図１３】

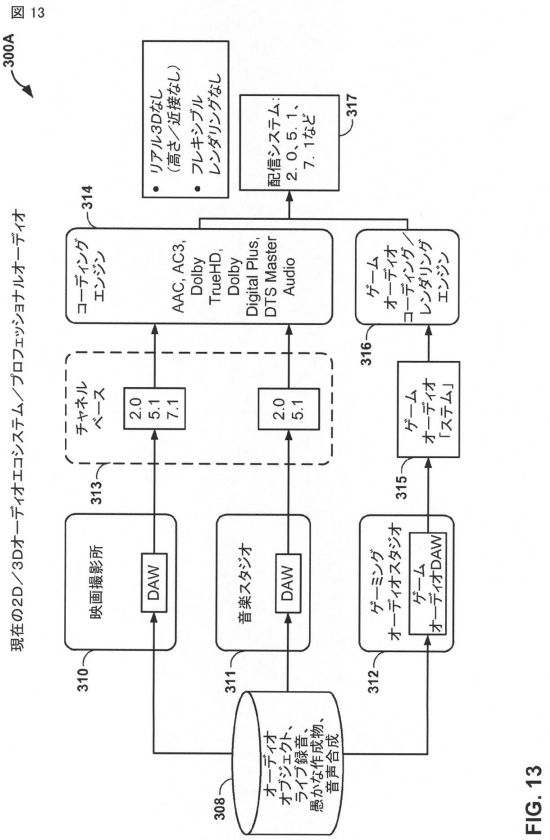


FIG. 13

【図１４】

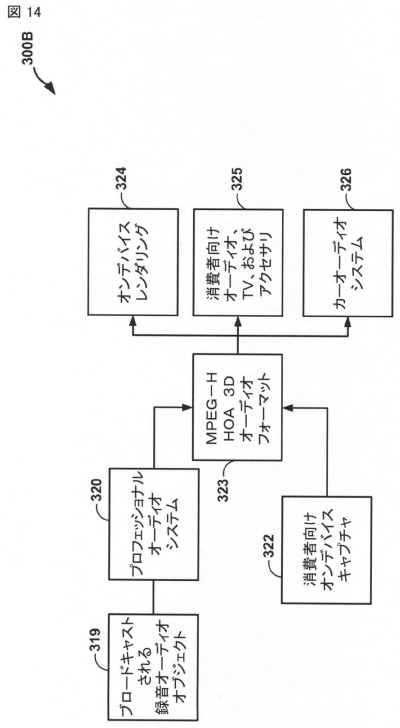


FIG. 14

【図１５Ａ】

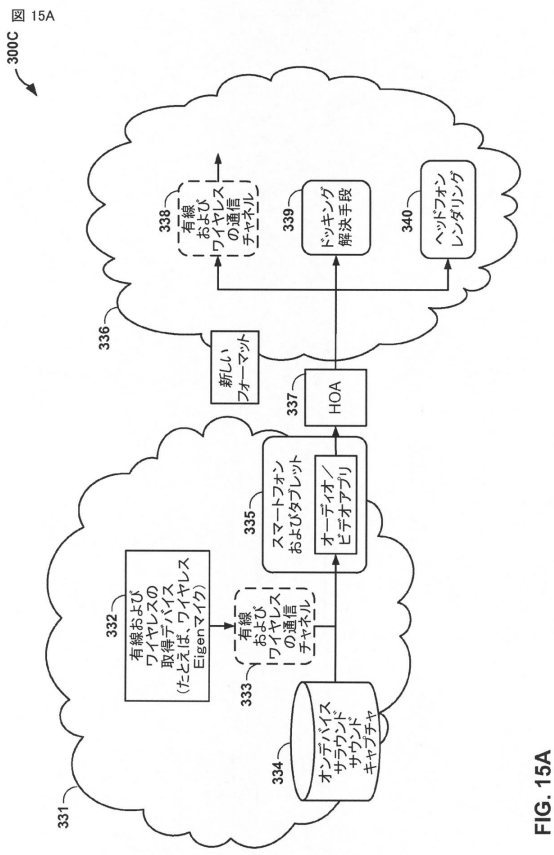


FIG. 15A

【図１５Ｂ】

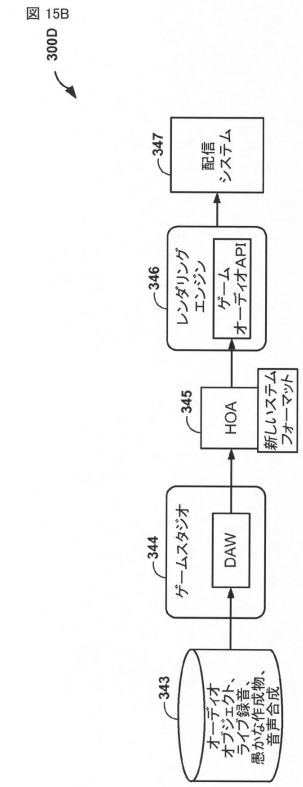


FIG. 15B

【図１６】

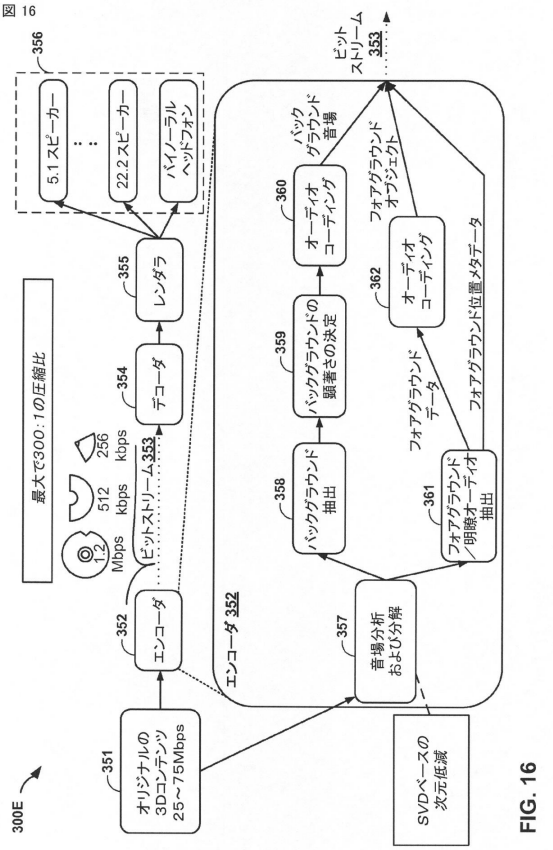


FIG. 16

【図１７】

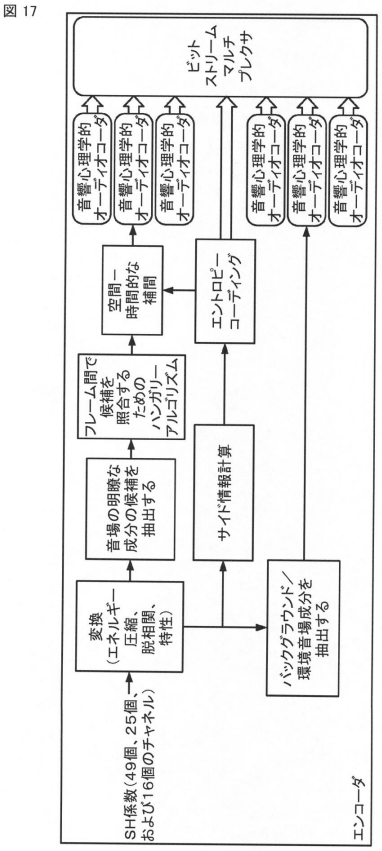


FIG. 17

【図１８】

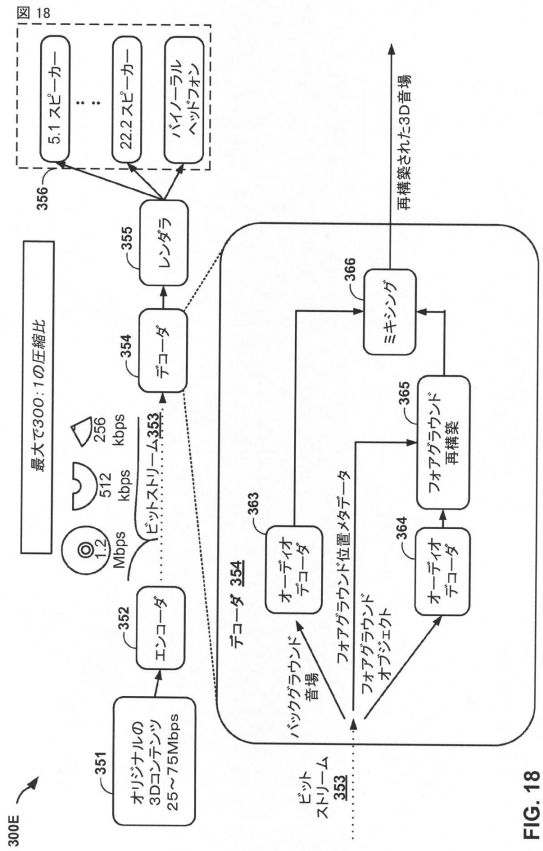


FIG. 18

【図１９】

図 19

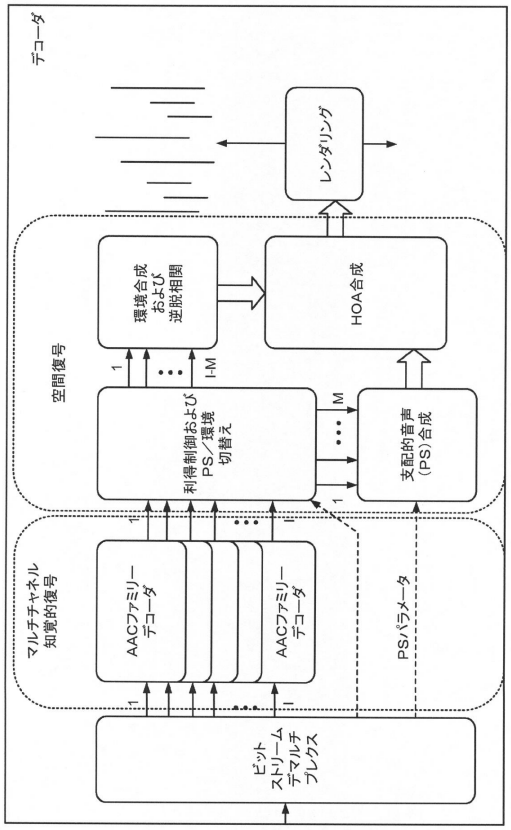


FIG. 19

【図２０Ａ】

図 20A

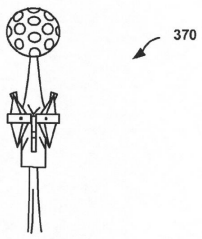


FIG. 20A

【図２０Ｂ】

図 20B

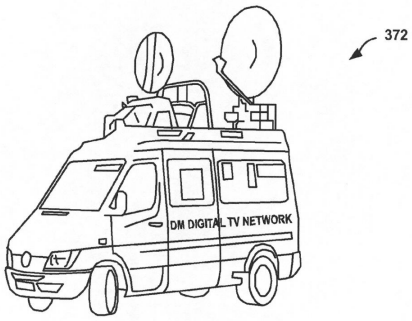


FIG. 20B

【図２０Ｃ】

図 20C

FIG. 20C

【図２０Ｄ】

図 20D

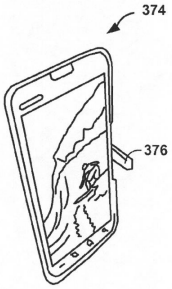


FIG. 20D

【図２０Ｅ】

図 20E

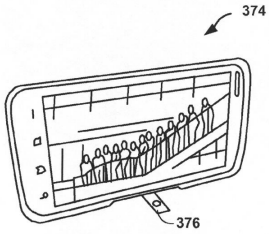


FIG. 20E

【図２０Ｆ】

図 20F

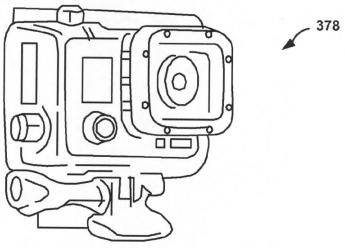


FIG. 20F

【図２０Ｇ】

図 20G

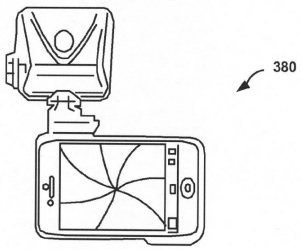


FIG. 20G

【図２１Ａ】

図 21A

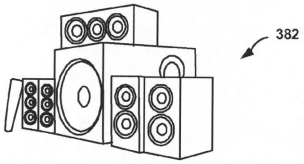


FIG. 21A

【図２１Ｂ】

図 21B

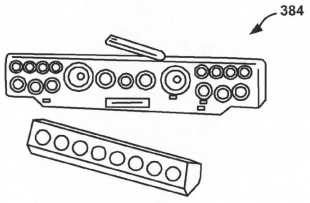


FIG. 21B

【図２１Ｃ】

図 21C

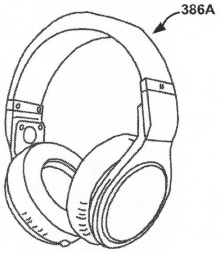


FIG. 21C

【図２１Ｄ】

図 21D

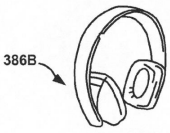


FIG. 21D

【図２１Ｅ】

図 21E

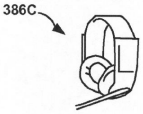


FIG. 21E

【図２２Ａ】

図 22A

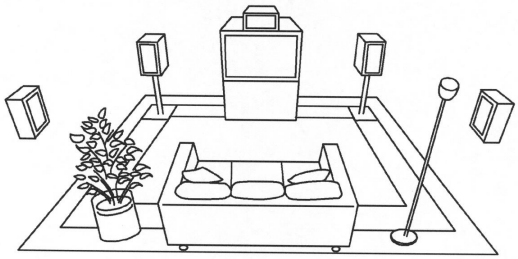


FIG. 22A

【図２２Ｂ】

図 22B

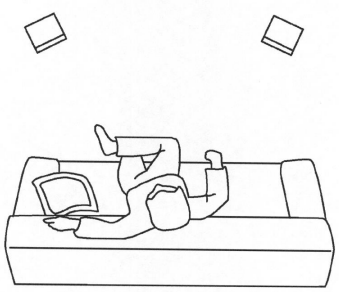
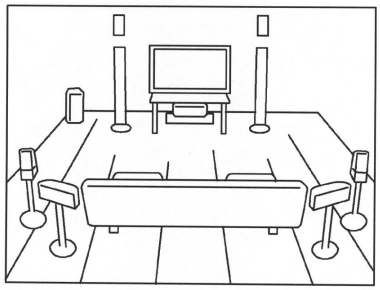


FIG. 22B

【図２２Ｃ】

図 22C



フルハイトを伴う9．1

FIG. 22C

【図２２Ｄ】

図 22D

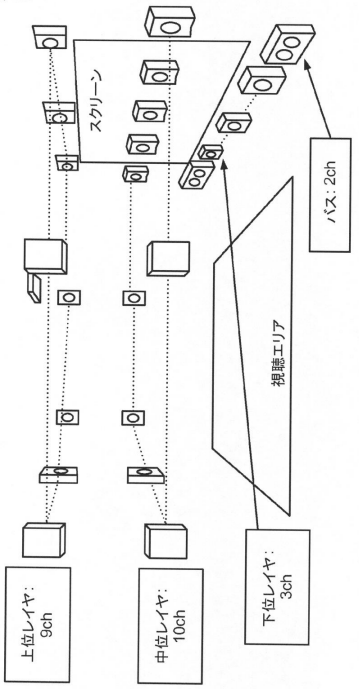


FIG. 22D

【図２２Ｅ】

図 22E

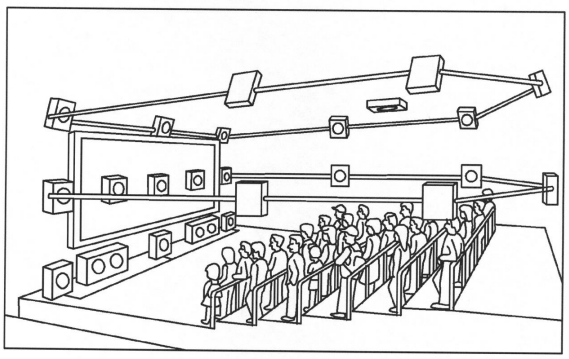


FIG. 22E

【図２２Ｆ】

図 22F

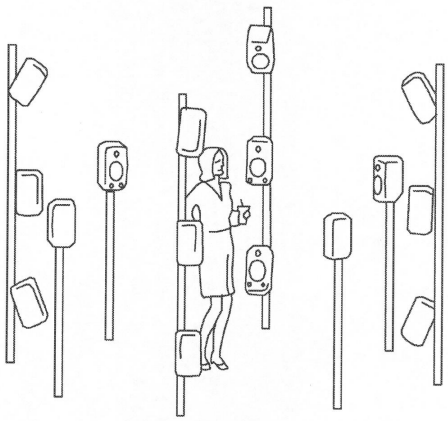


FIG. 22F

【図２２Ｇ】

図 22G

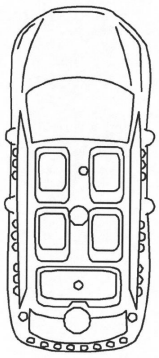


FIG. 22G

【図２２Ｈ】

図 22H

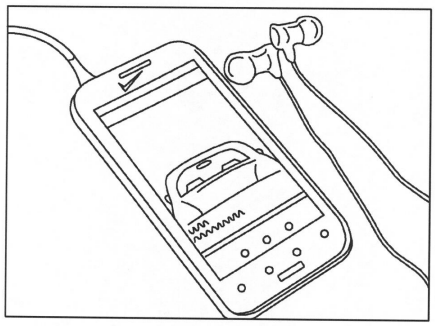


FIG. 22H

【図２３】

図 23

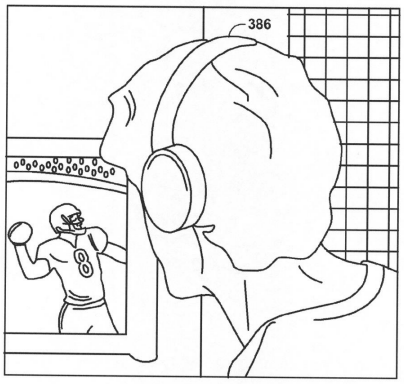


FIG. 23

【図２４】

図 24

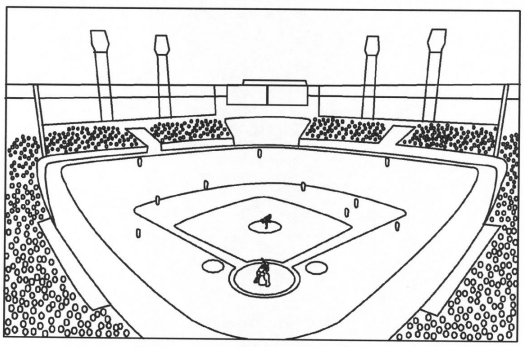


FIG. 24

【図２５】

図 25

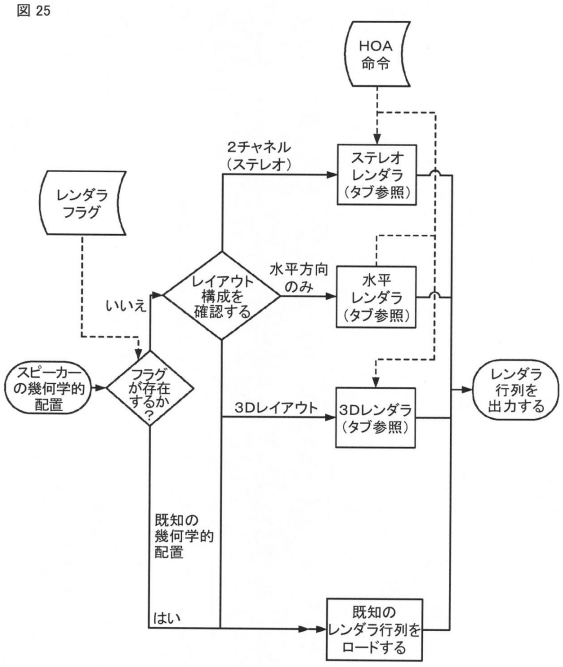


FIG. 25

【図２６】

図 26

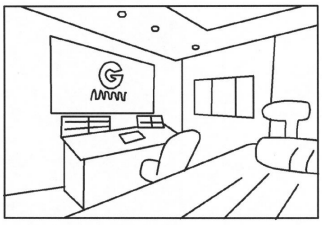


FIG. 26

【図２７】

図 27

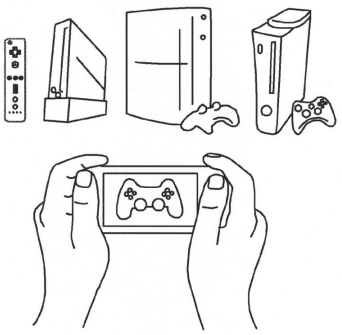


FIG. 27

【図２８】

図 28

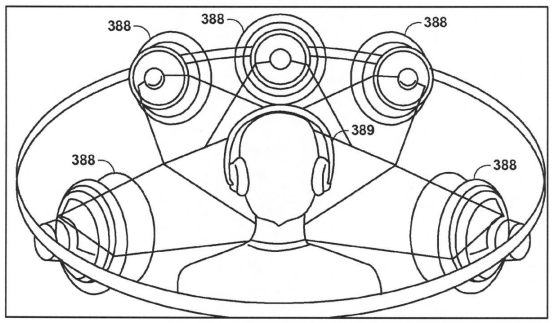


FIG. 28

【図２９】

図 29

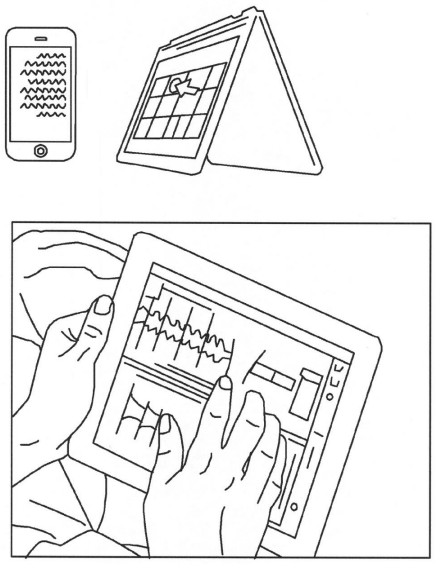


FIG. 29

【図３０】

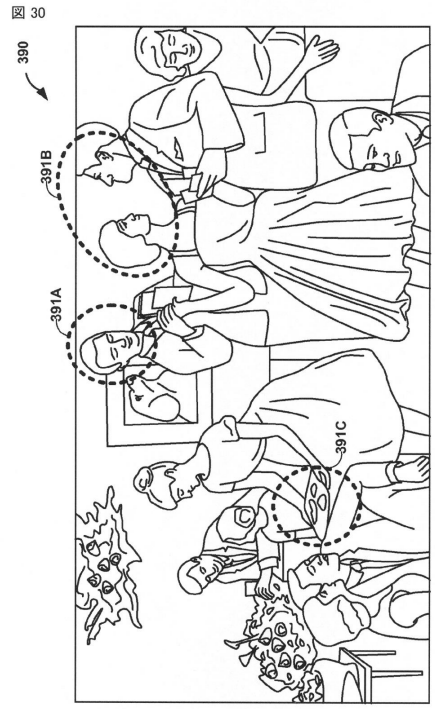


FIG. 30

【図３１Ａ】

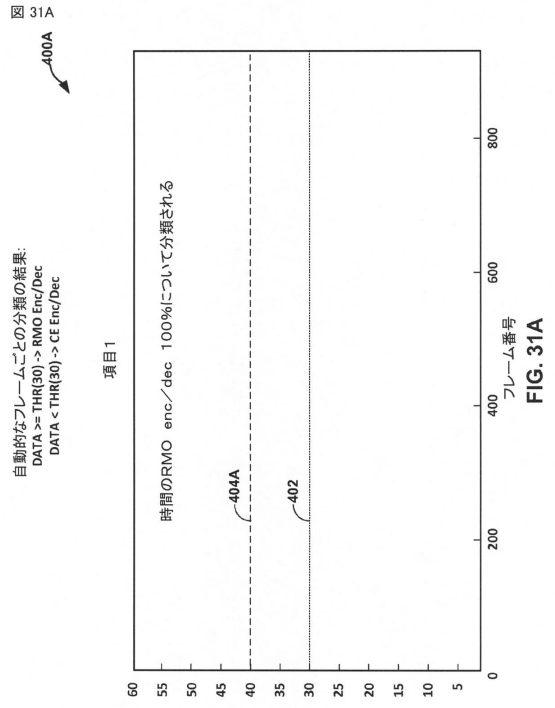


FIG. 31A

【図３１Ｂ】

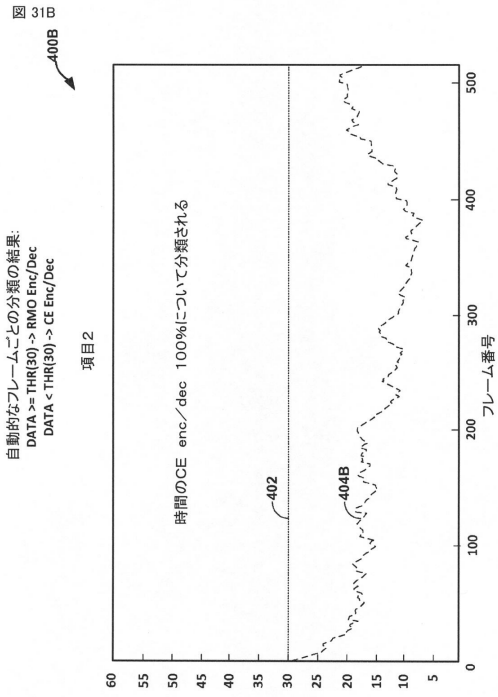


FIG. 31B

【図３１Ｃ】

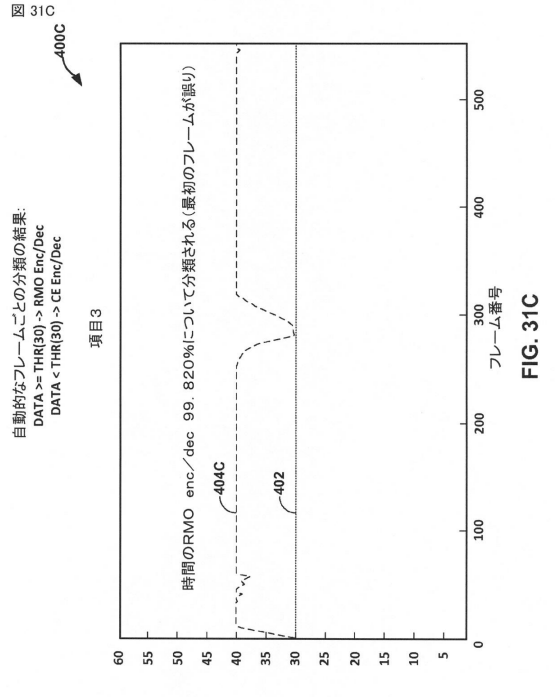


FIG. 31C

【図３１Ｄ】

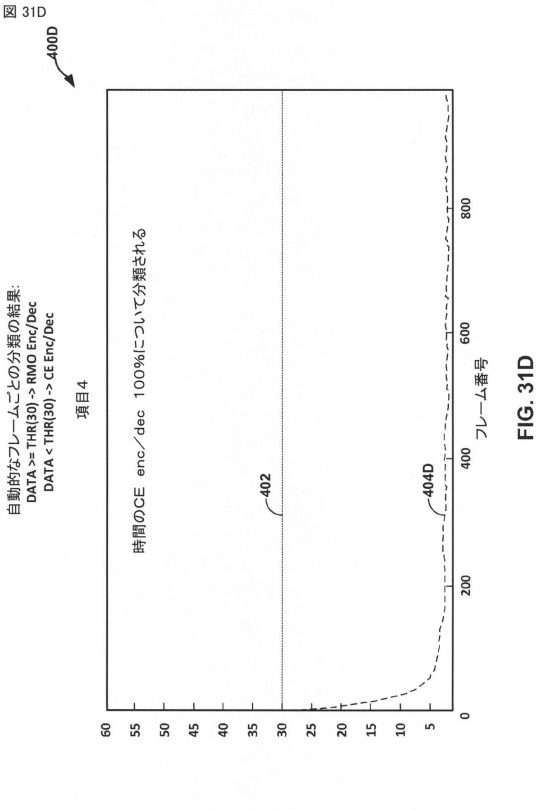


FIG. 31D

【図３１Ｅ】

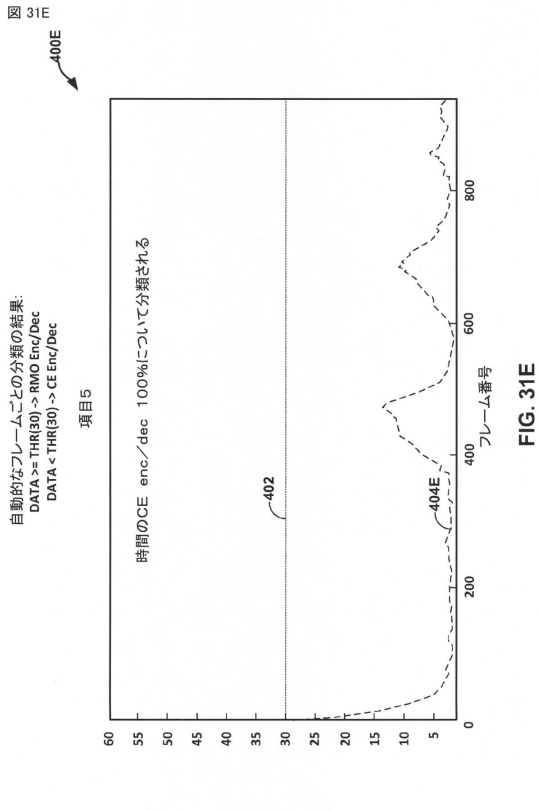


FIG. 31E

【図３１Ｆ】

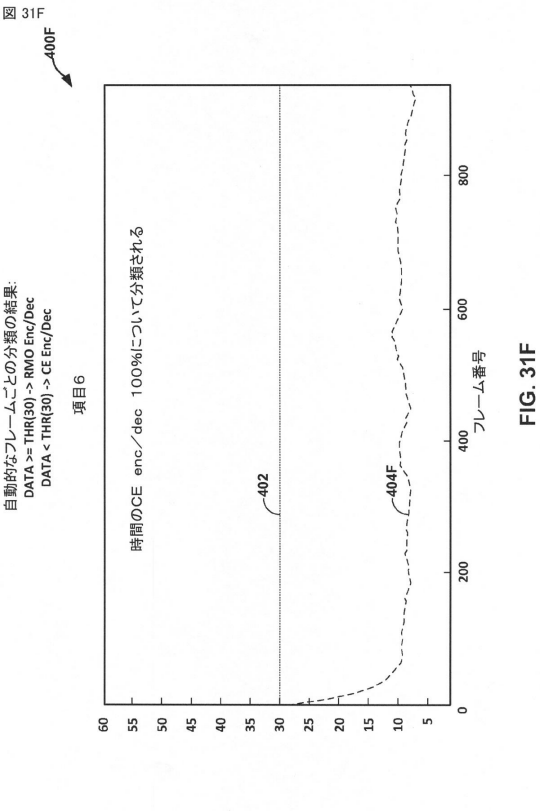


FIG. 31F

【図３１Ｇ】

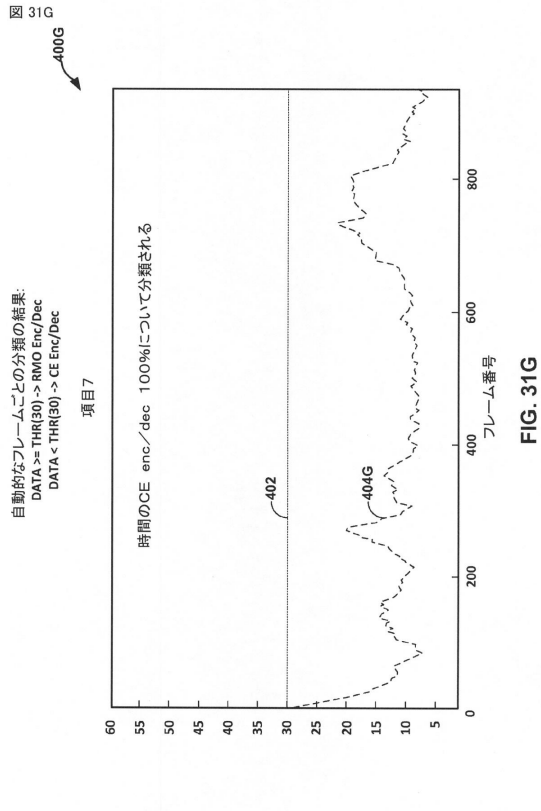


FIG. 31G

【図３１Ｈ】

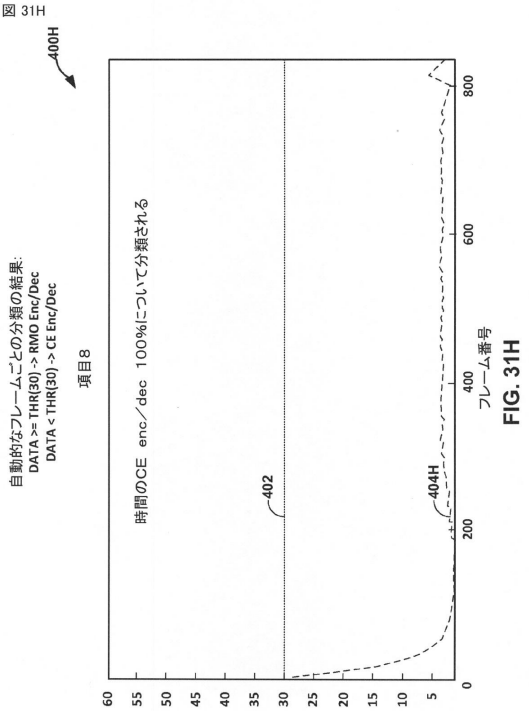


FIG. 31H

【図３１Ｉ】

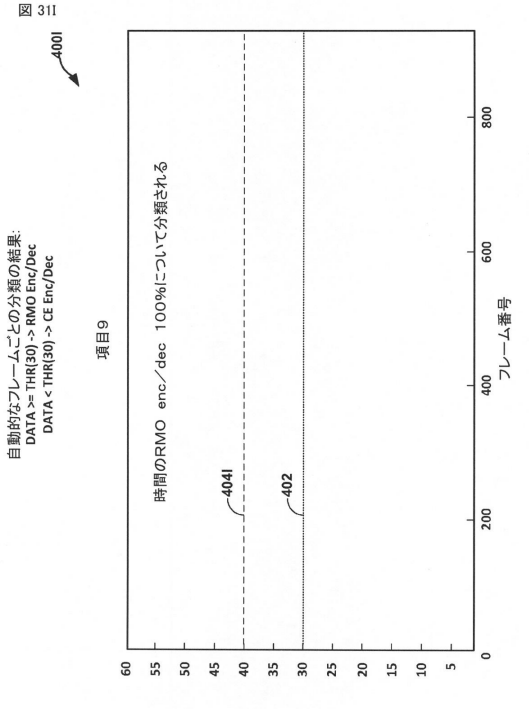


FIG. 31I

【図３１Ｊ】

FIG. 31J

【図３１Ｋ】

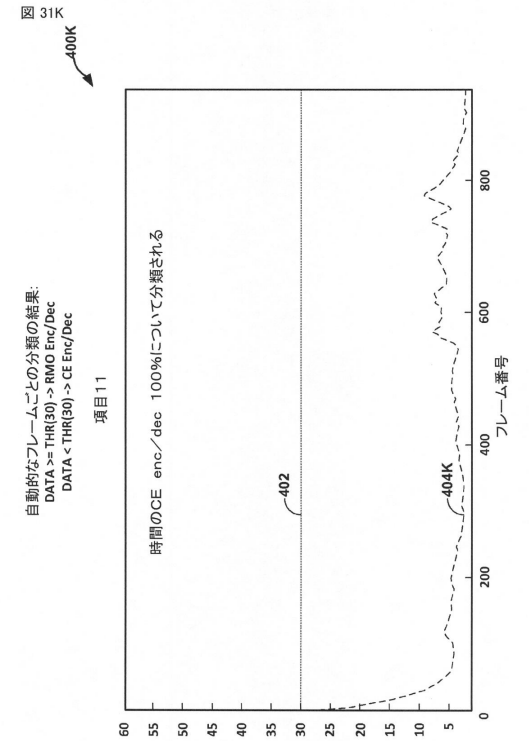


FIG. 31K

【図３１Ｌ】

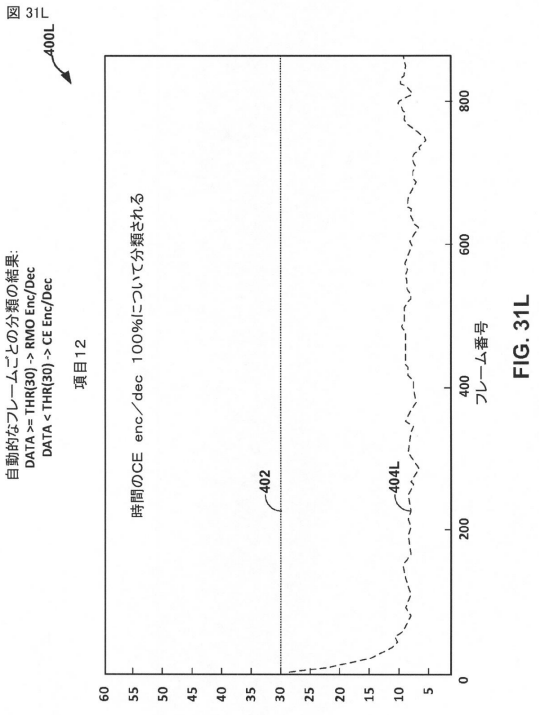


FIG. 31L

【図３１Ｍ】

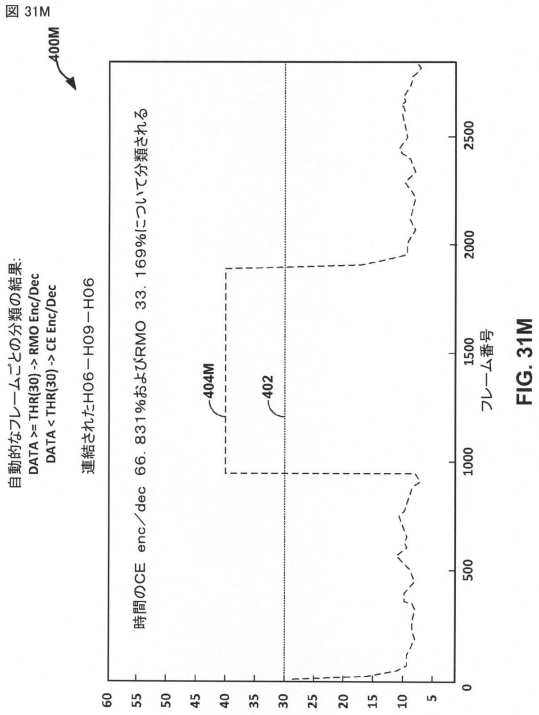


FIG. 31M

【図３２】

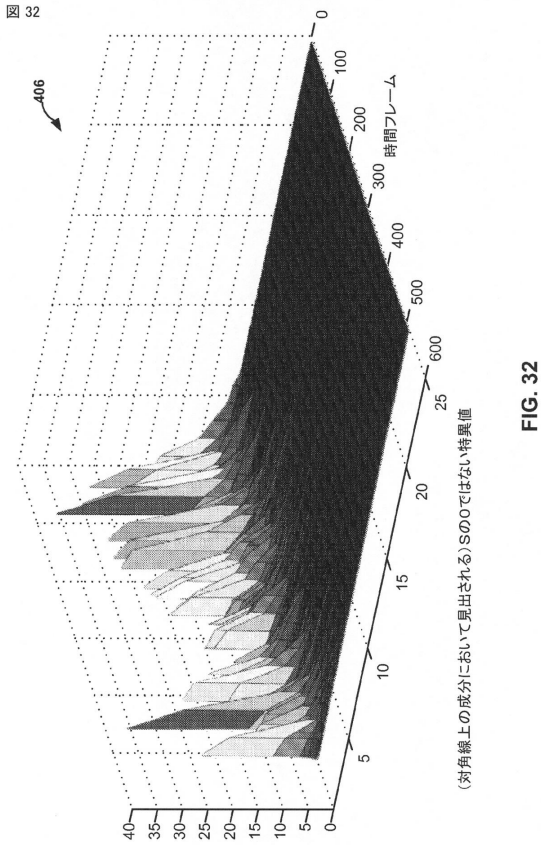


FIG. 32

【図３３Ａ】

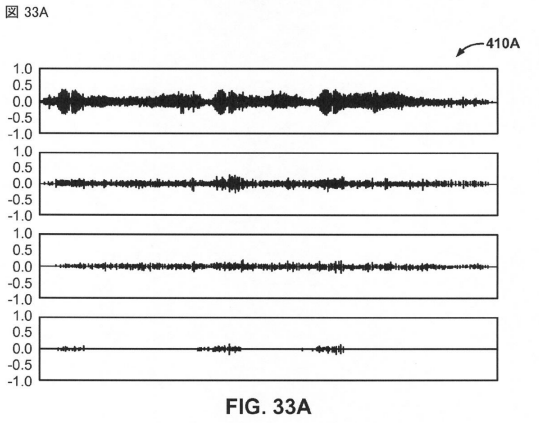


FIG. 33A

【図３３Ｂ】

図 33B

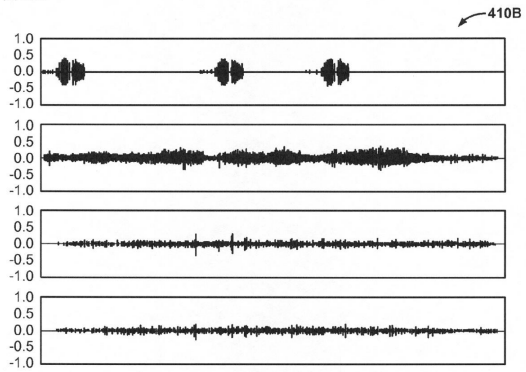


FIG. 33B

【図３４】

図 34

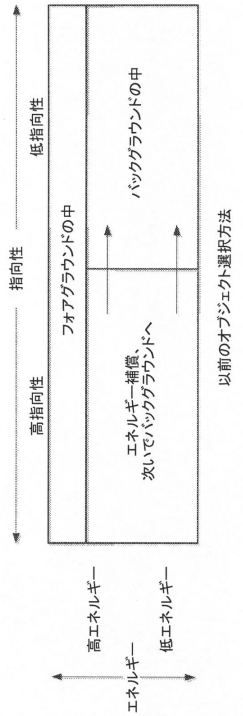


FIG. 34

【図３５】

図 35

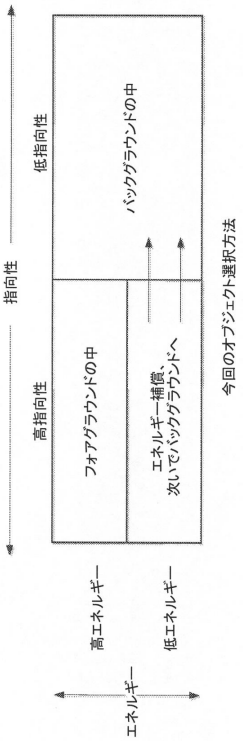


FIG. 35

【図３６Ａ】

図 36A

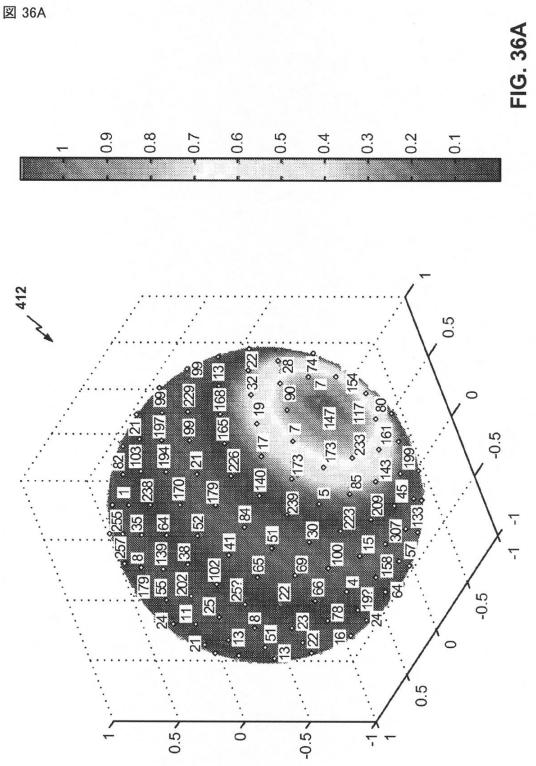


FIG. 36A

【図３６Ｂ】

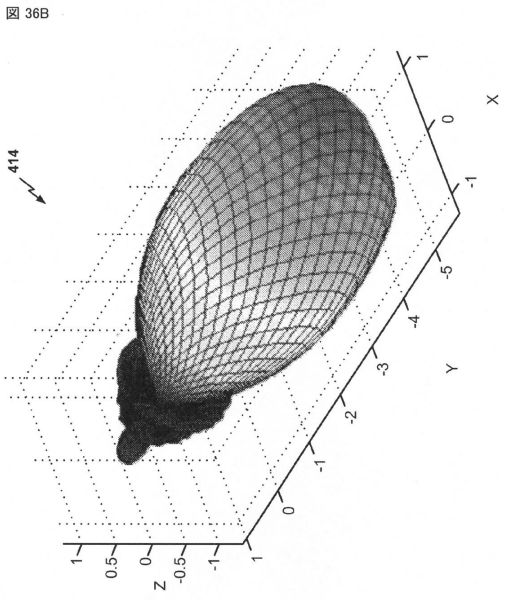


FIG. 36B

【図３６Ｃ】

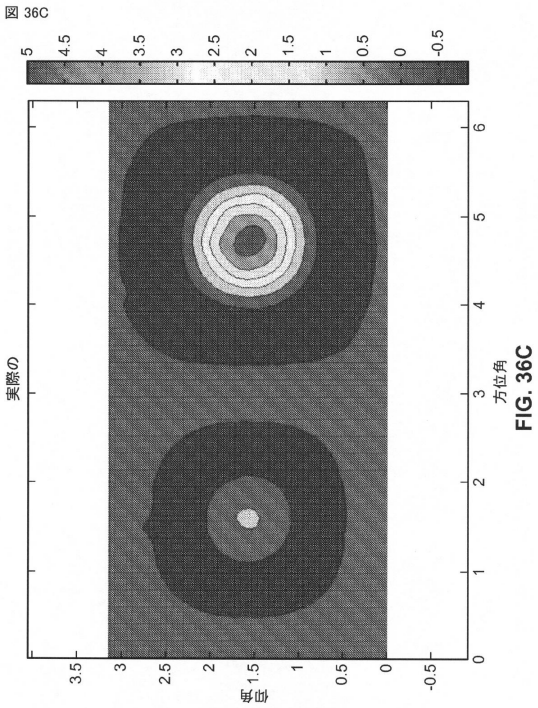


FIG. 36C

【図３６Ｄ】

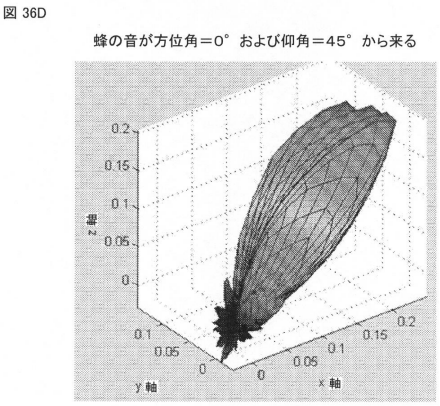


FIG. 36D

【図３６Ｅ】

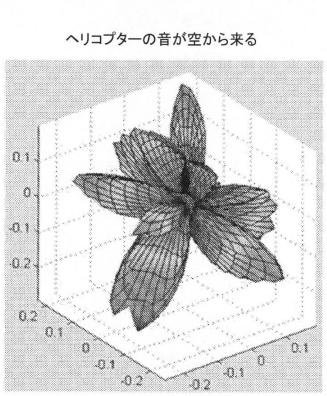


FIG. 36E

【図３６Ｆ】

図 36F

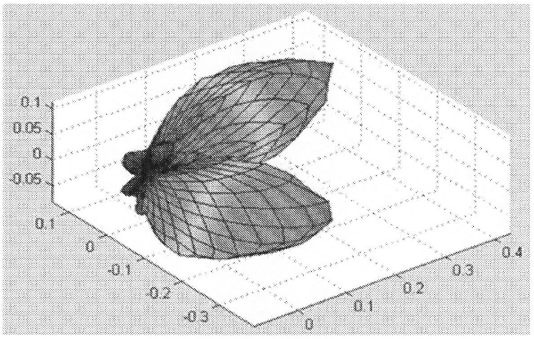


FIG. 36F

【図３６Ｇ】

図 36G

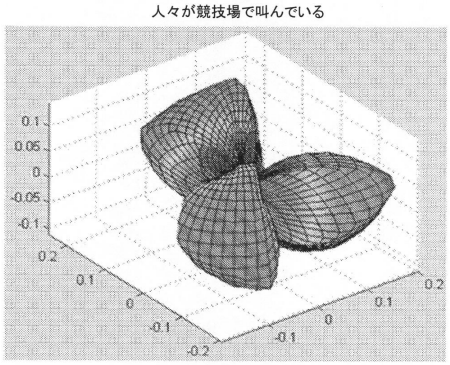


FIG. 36G

【図３７】

図 37

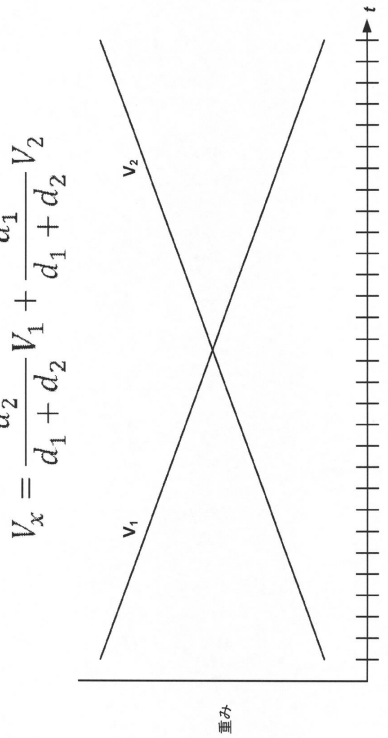


FIG. 37

【図３８】

図 38

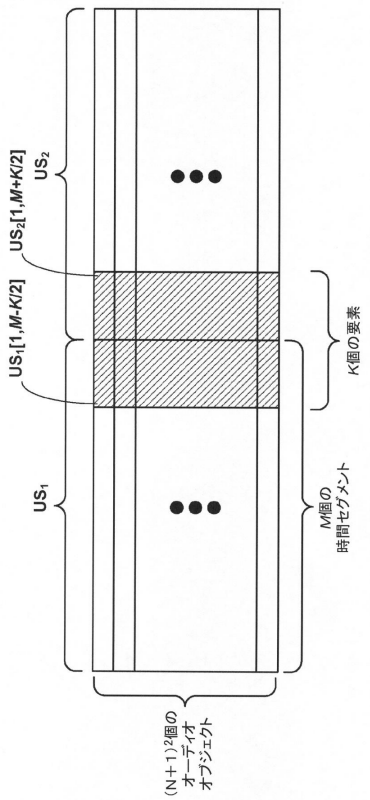


FIG. 38

【図３９】

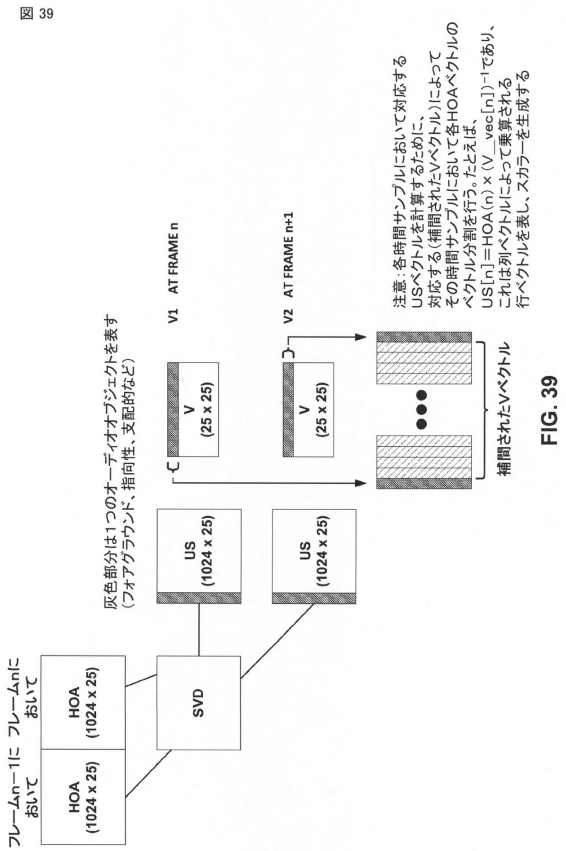


FIG. 39

【図４０Ａ】

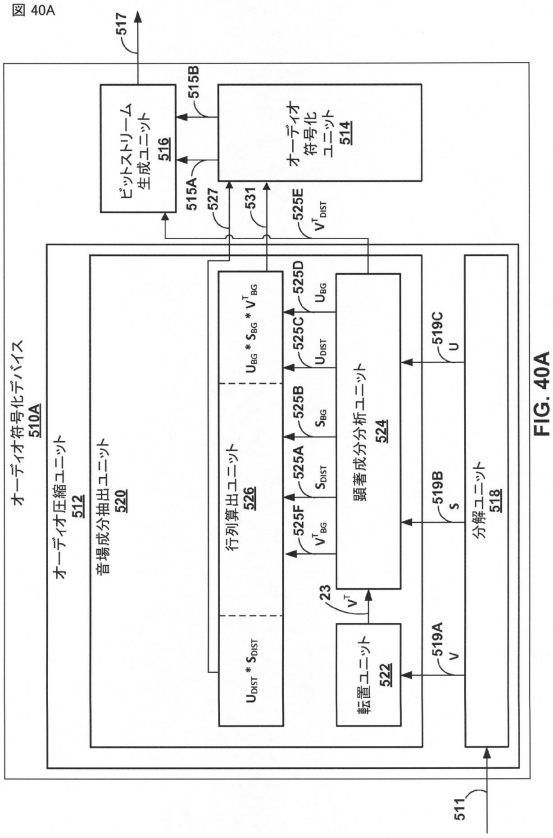


FIG. 40A

【図４０Ｂ】

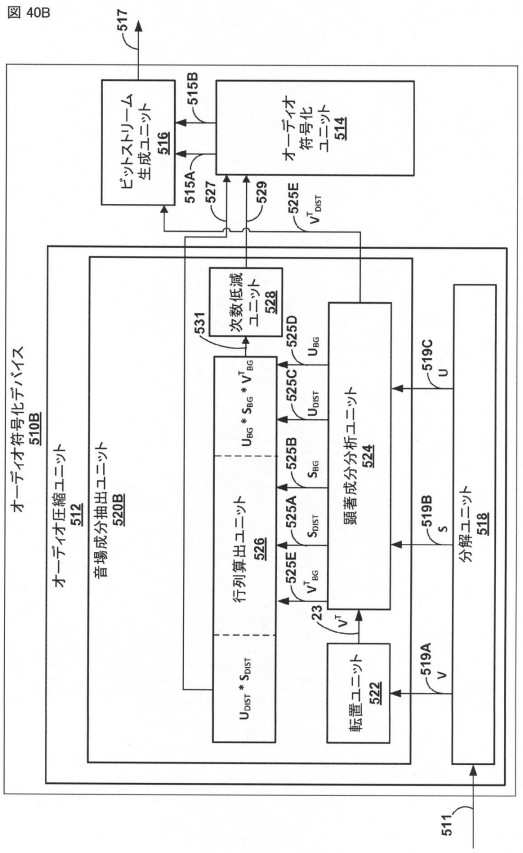


FIG. 40B

【図４０Ｃ】

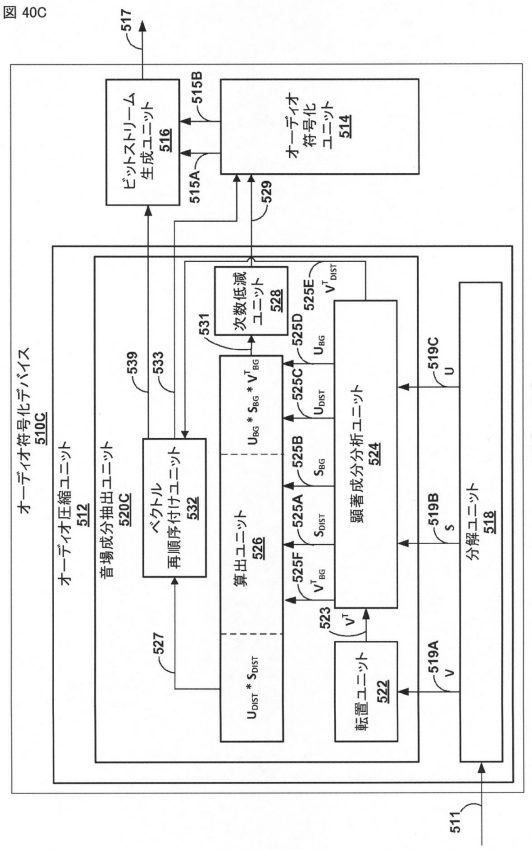


FIG. 40C

【図４０Ｄ】

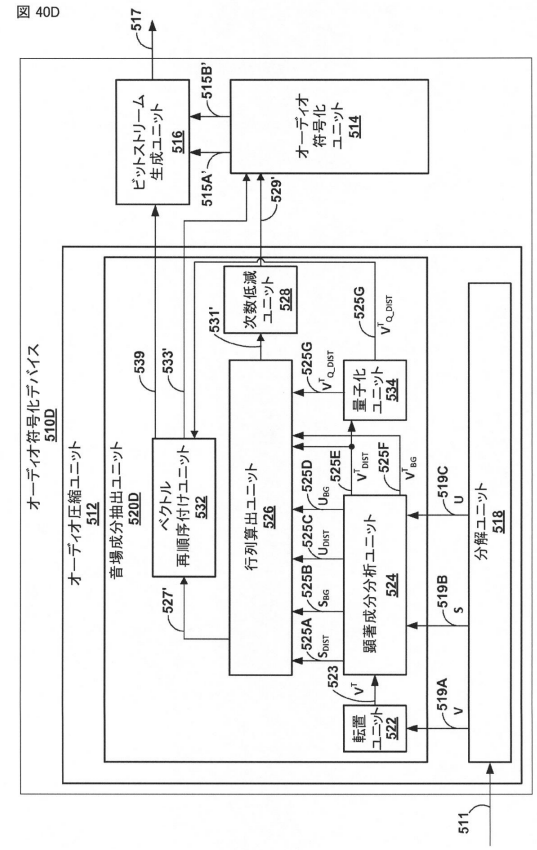


FIG. 40D

【図４０Ｅ】

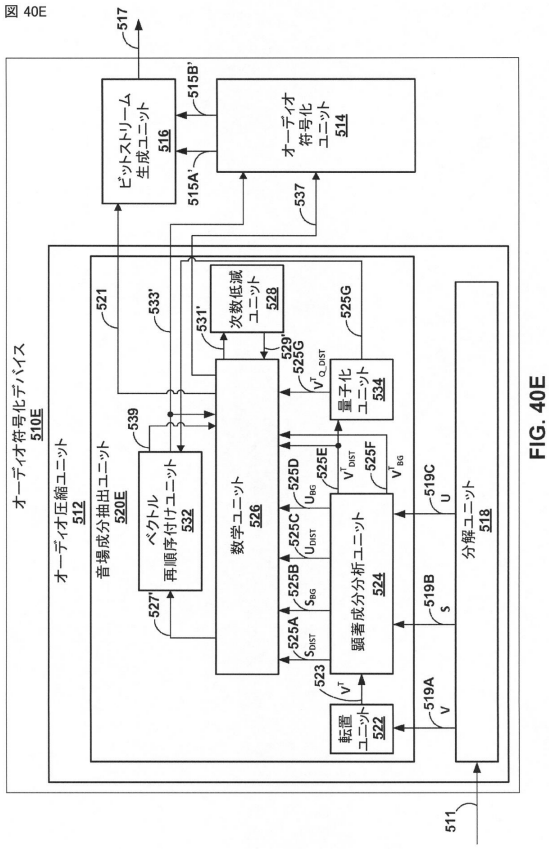


FIG. 40E

【図４０Ｆ】

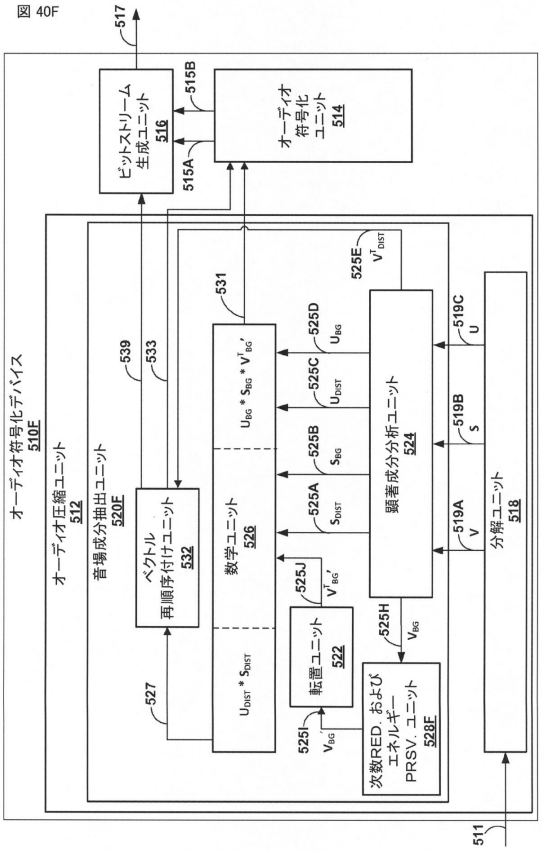


FIG. 40F

【図４０Ｇ】

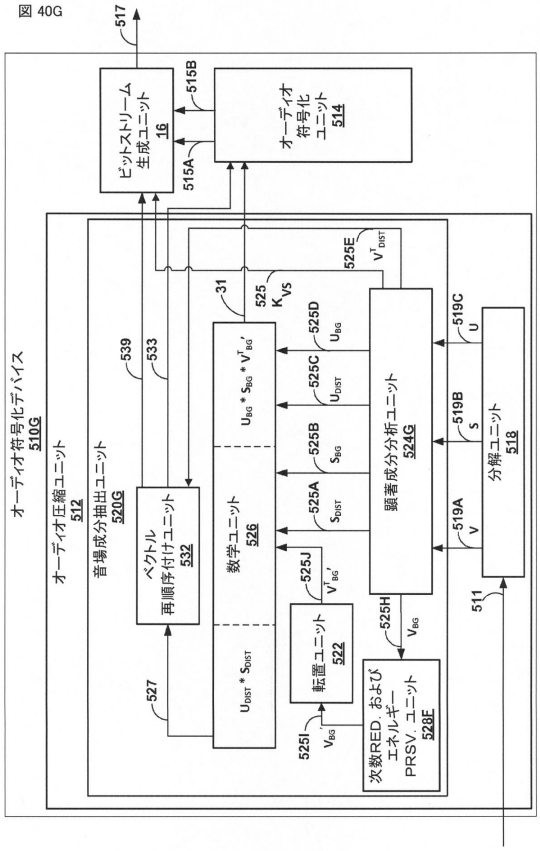


FIG. 40G

【図４０Ｈ】

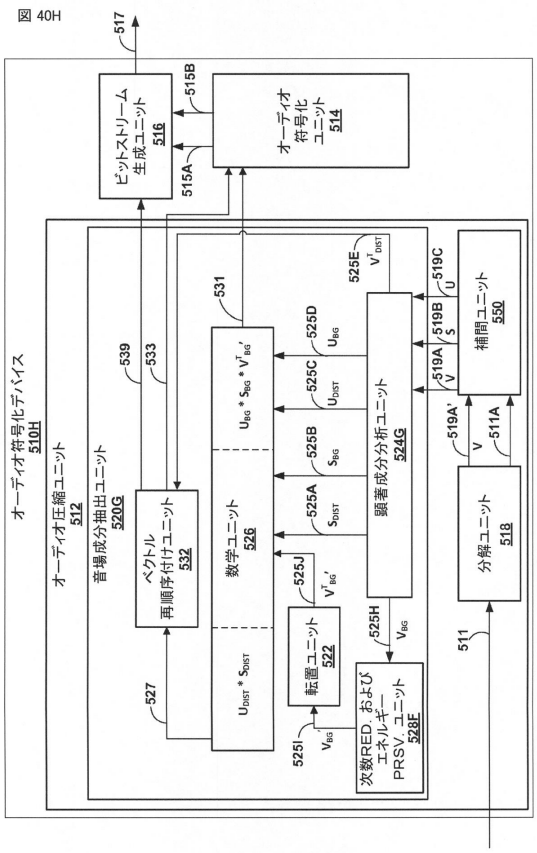


FIG. 40H

【図４０Ｉ】

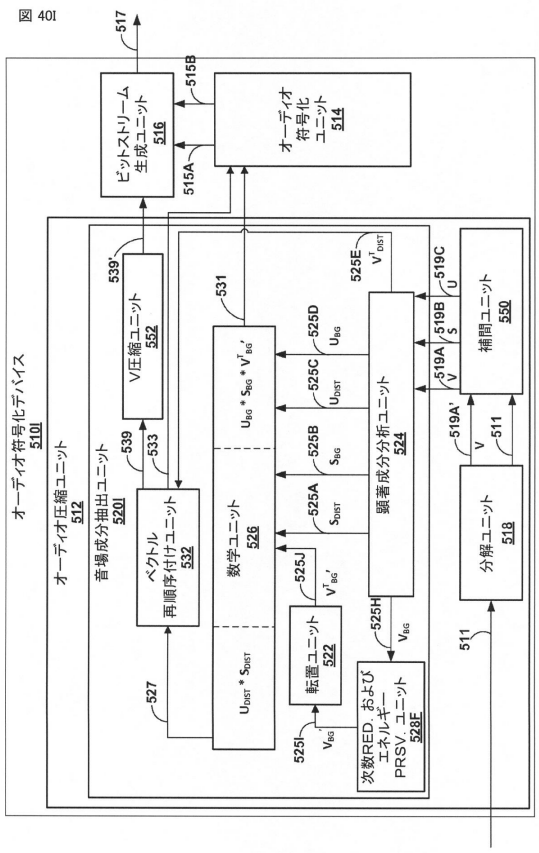


FIG. 40I

【図４０Ｊ】

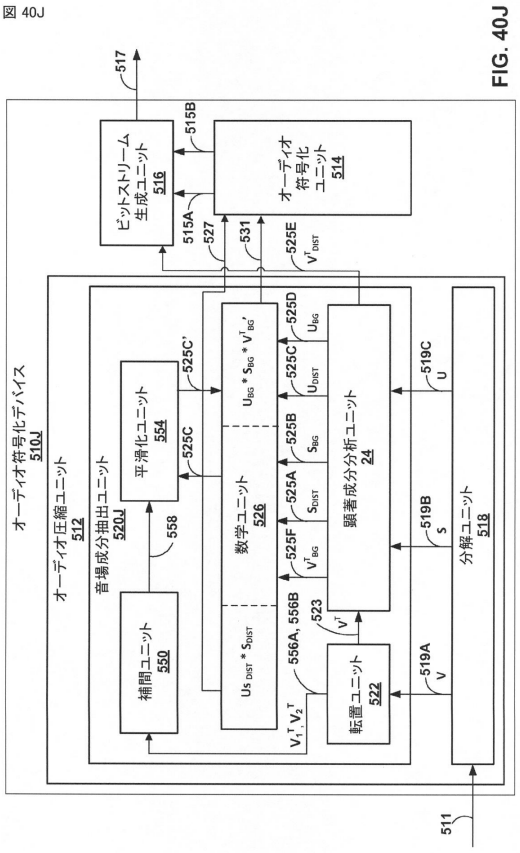


FIG. 40J

【図４１Ａ】

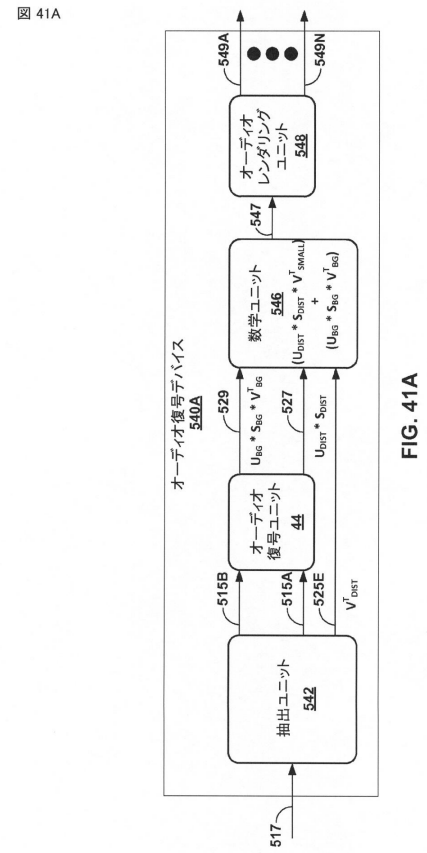


FIG. 41A

【図４１Ｂ】

図 41B

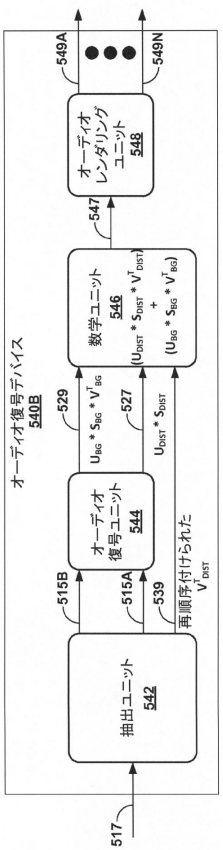


FIG. 41B

【図４１Ｃ】

図 41C

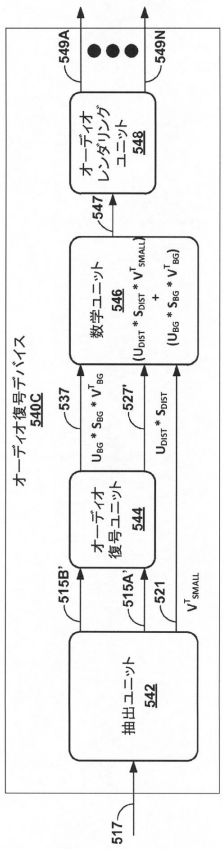


FIG. 41C

【図４１Ｄ】

図 41D

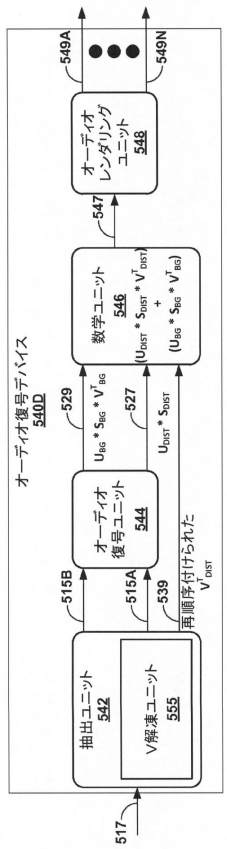


FIG. 41D

【図４２Ａ】

図 42A

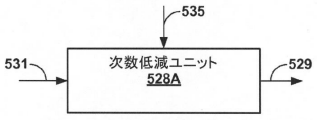


FIG. 42A

【図４２Ｂ】

図 42B

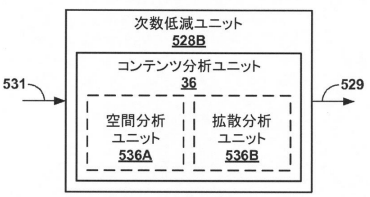


FIG. 42B

【図４２Ｃ】

図 42C

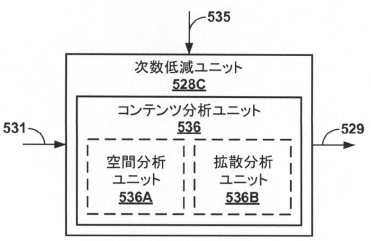


FIG. 42C

【図４３】

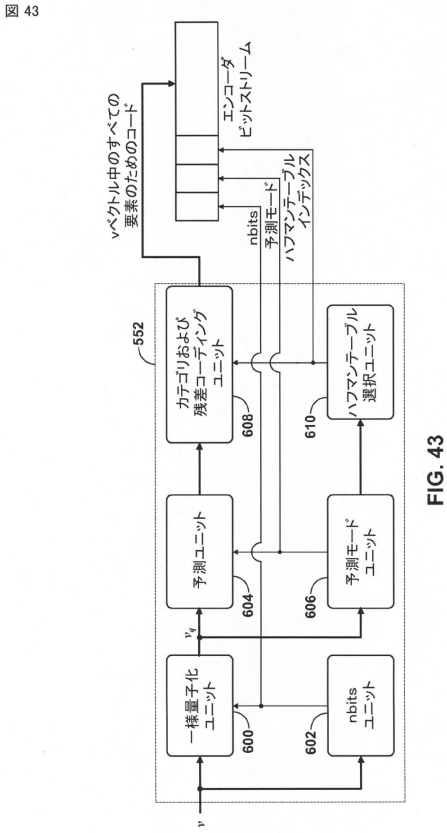


FIG. 43

【図４４】

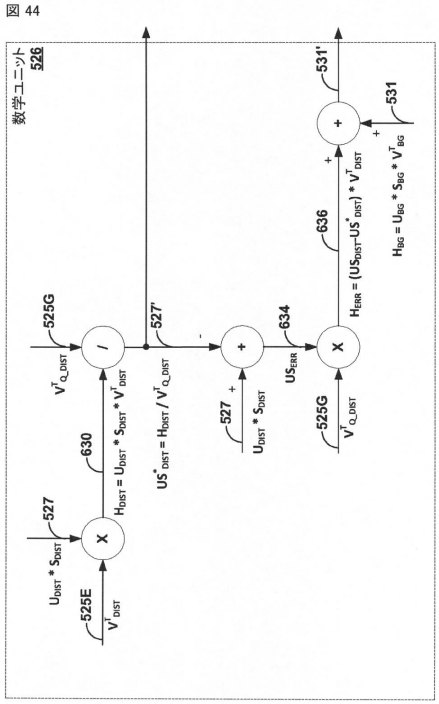


FIG. 44

【図４５Ａ】

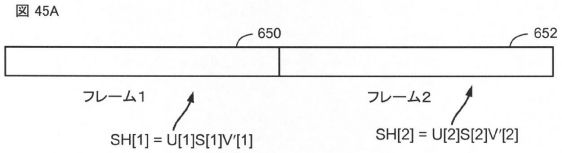
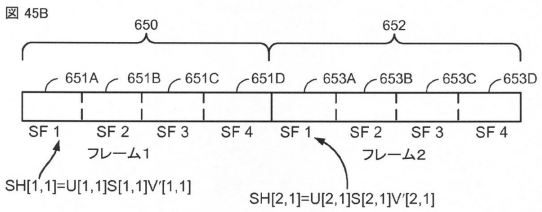


FIG. 45A

【図４５Ｂ】



V'[1,2] = 補間 (V'[1,1], V'[2,1]) → U[1,2]S[1,2] = SH[1,2](V'[1,2])-1
V'[1,3] = 補間 (V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1
V'[1,4] = 補間 (V'[1,1], V'[2,1]) → U[1,3]S[1,3] = SH[1,3](V'[1,3])-1

FIG. 45B

【図４６Ａ】

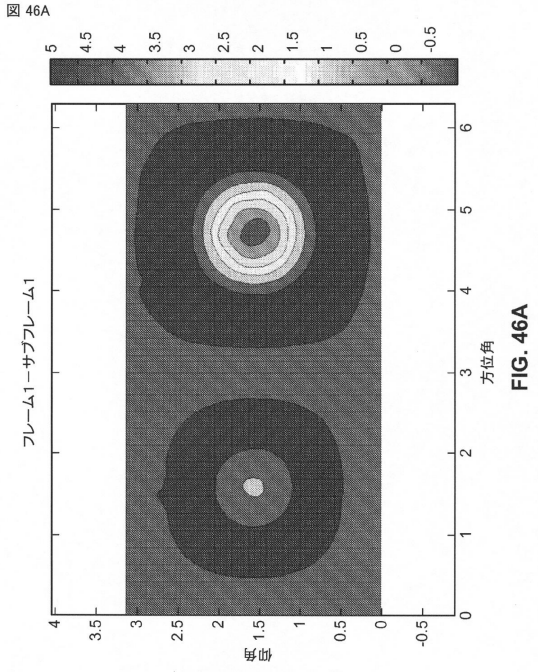


FIG. 46A

【図４６Ｂ】

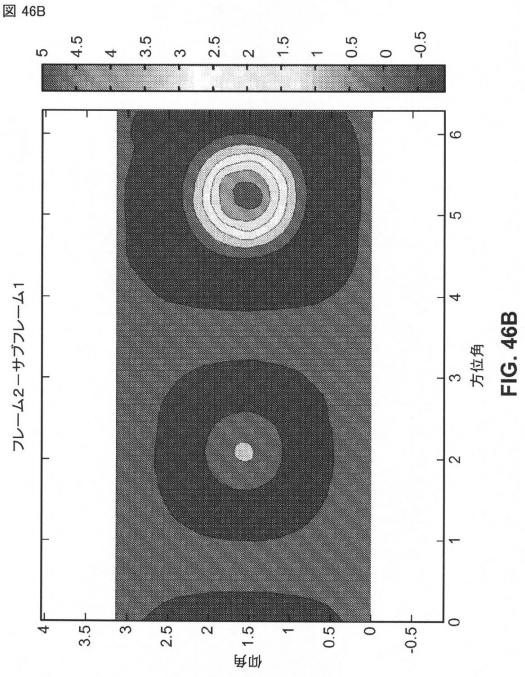


FIG. 46B

【図４６Ｃ】

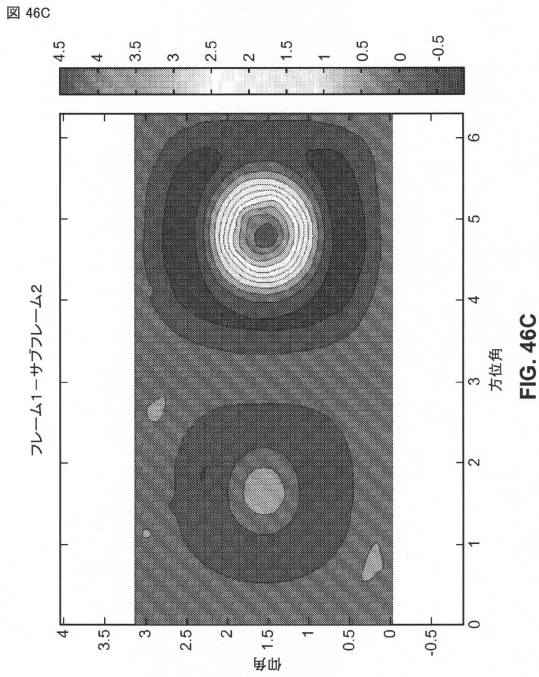


FIG. 46C

【図４６Ｄ】

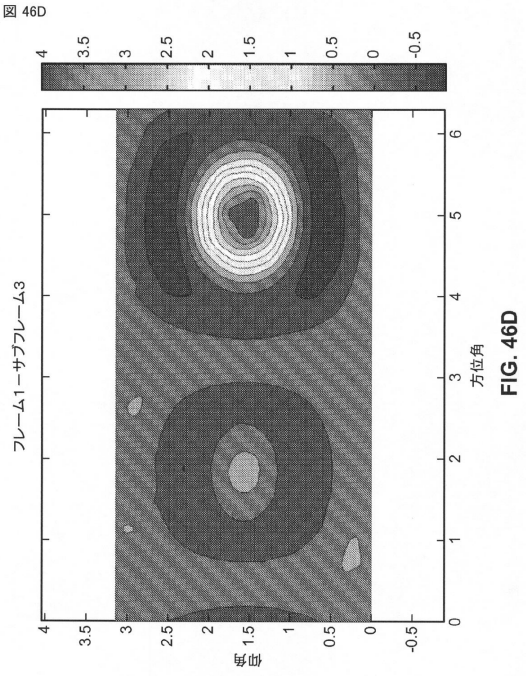


FIG. 46D

【図４６Ｅ】

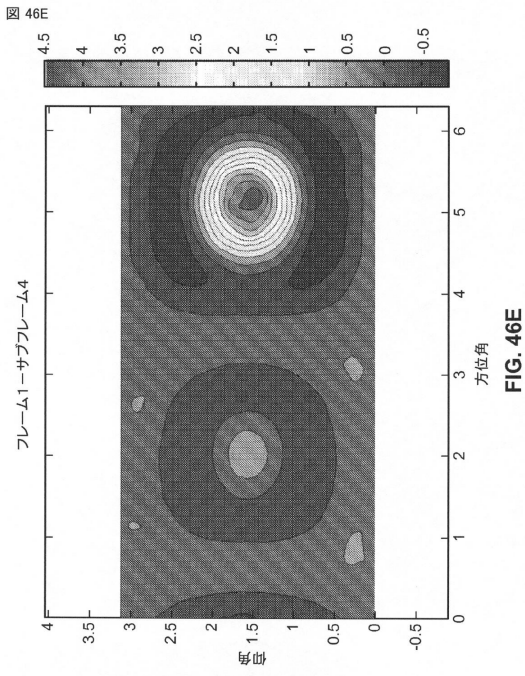


FIG. 46E

【図４７】

図 47

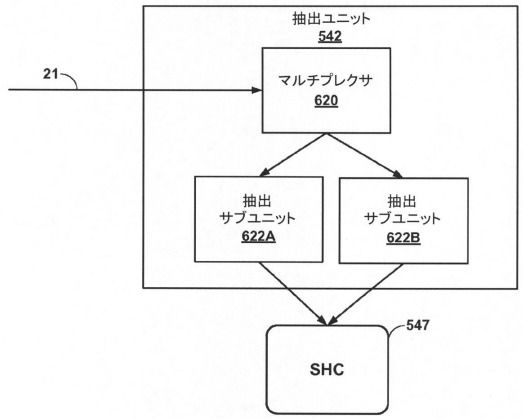


FIG. 47

【図４８】

図 48

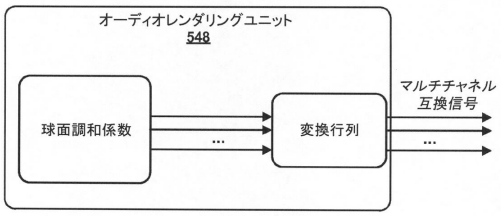


FIG. 48

【図４９Ａ】

図 49A

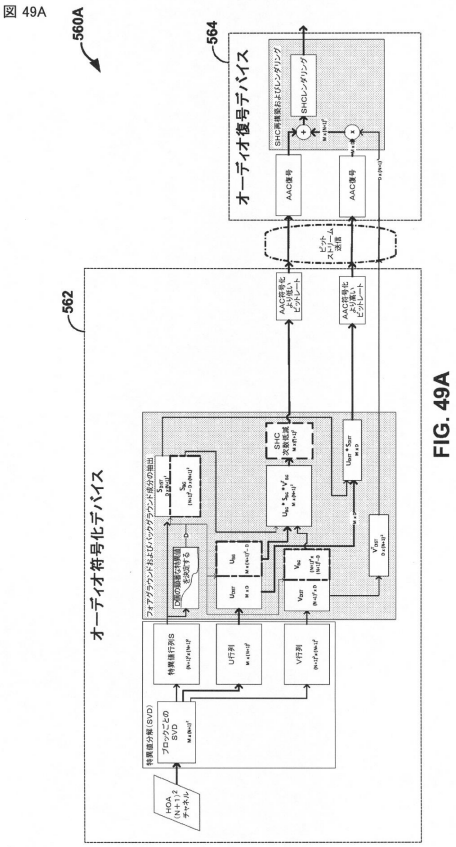


FIG. 49A

【図４９Ｂ】

図 49B

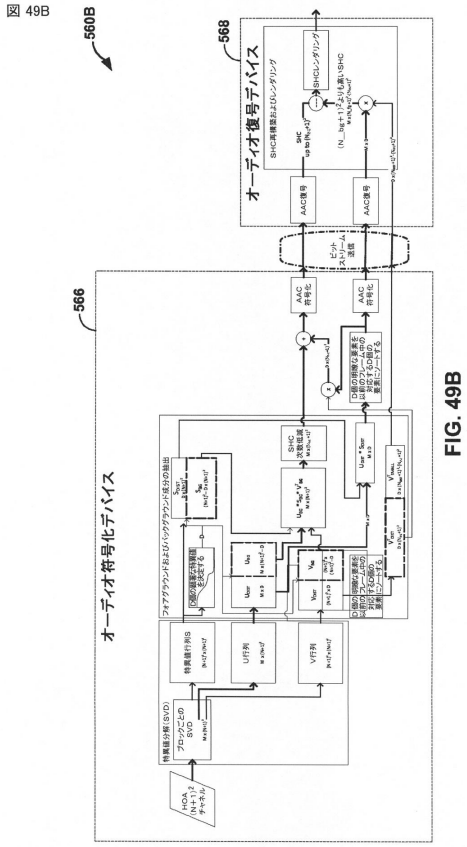


FIG. 49B

【図４９Ｃ】

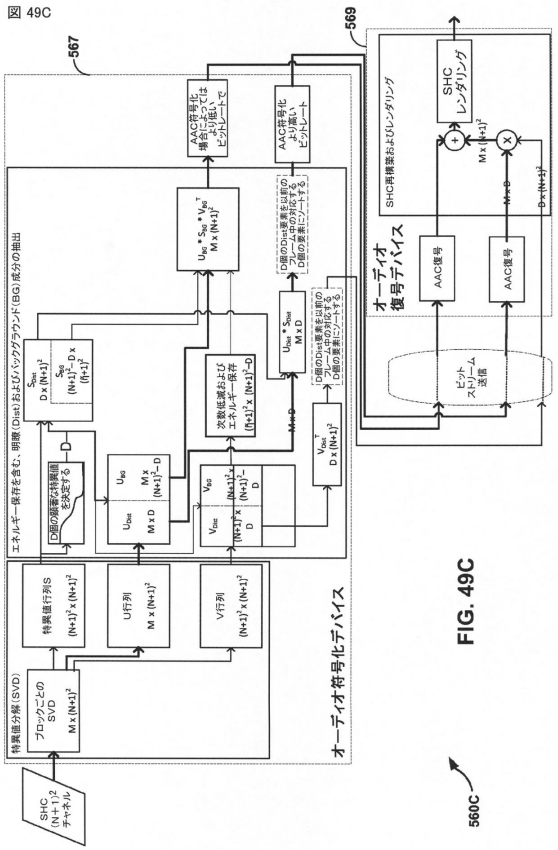


FIG. 49C

【図４９Ｄ（ｉ）】

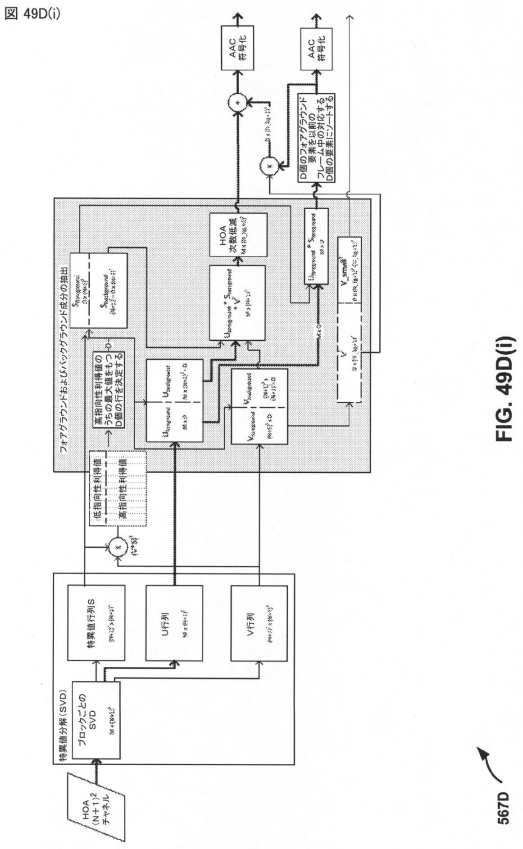


FIG. 49D(i)

【図４９Ｄ（ｉｉ）】

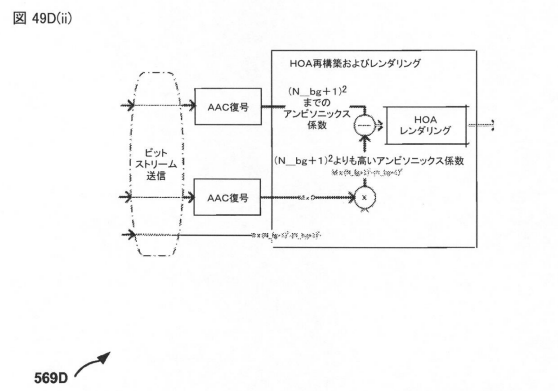


FIG. 49D(ii)

【図４９Ｅ（ｉ）】

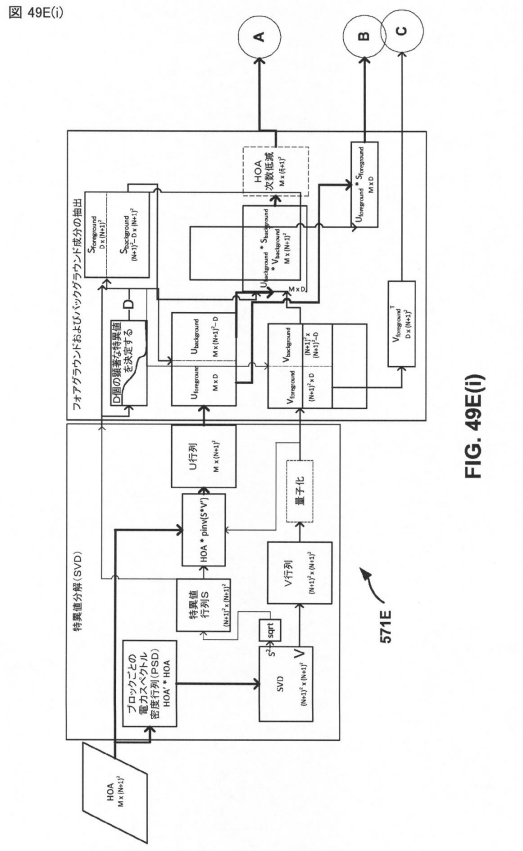


FIG. 49E(i)

【図４９Ｅ（ｉｉ）】

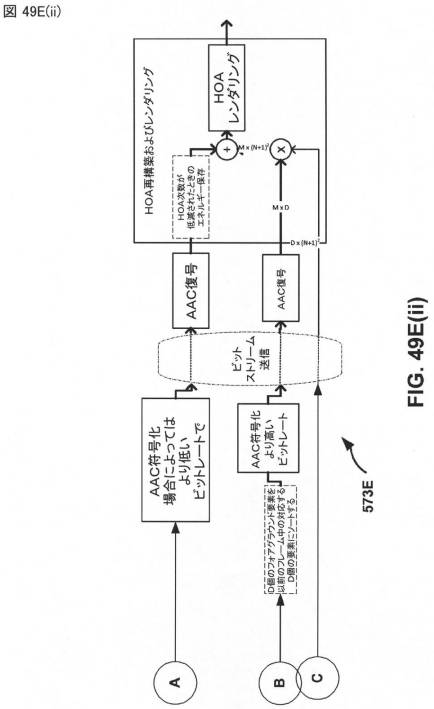


FIG. 49E(ii)

【図５０Ａ】

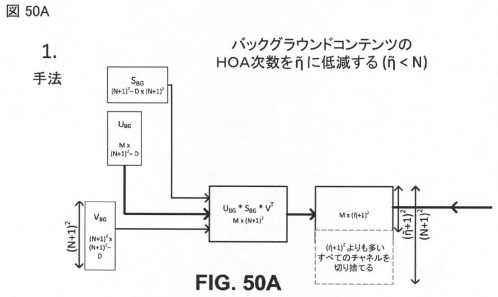


FIG. 50A

【図５０Ｂ】

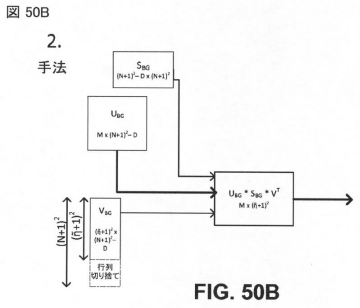


FIG. 50B

【図５１】

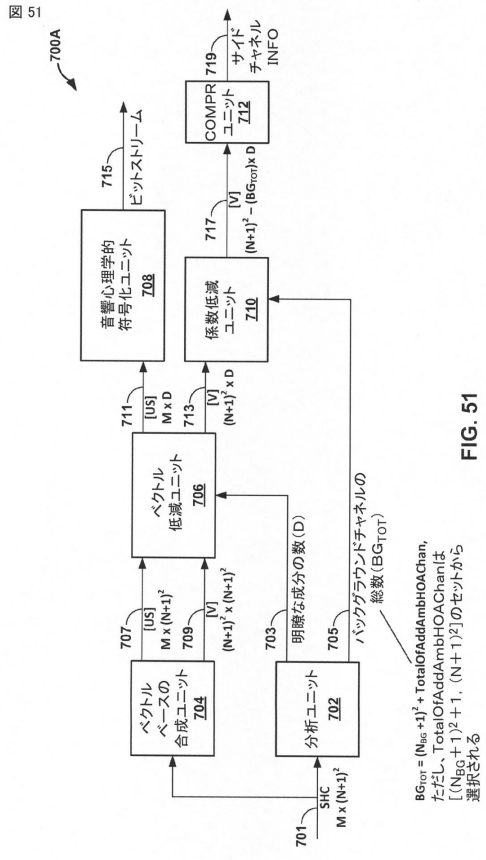


FIG. 51

【図５２】

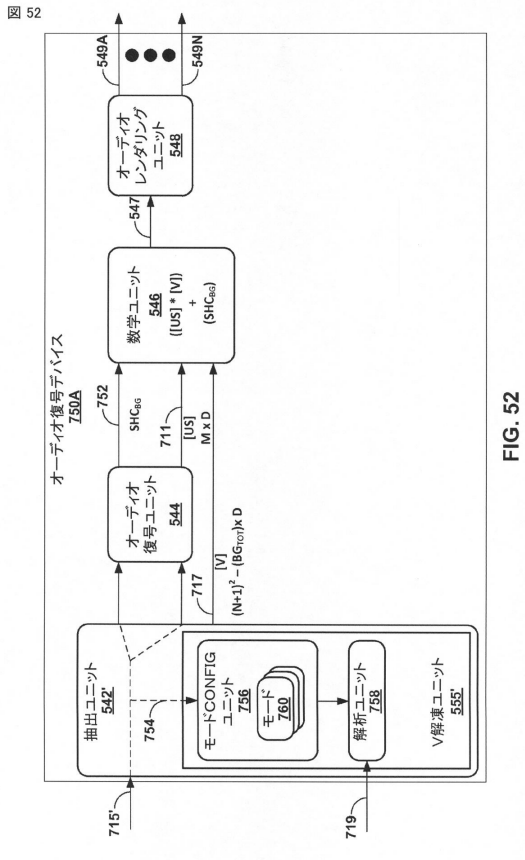


FIG. 52

【図５３】

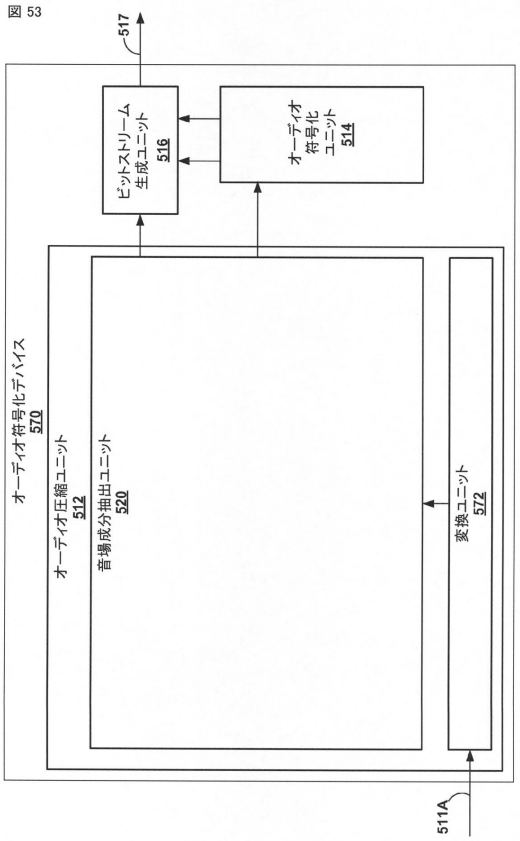


FIG. 53

【図５４】

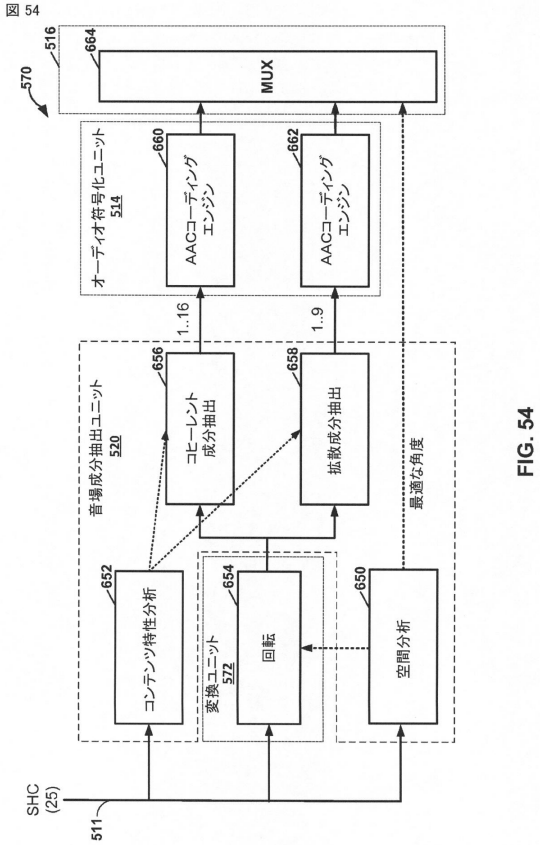


FIG. 54

【図５５Ａ】

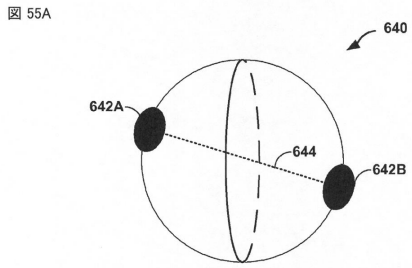


FIG. 55A

【図５６】

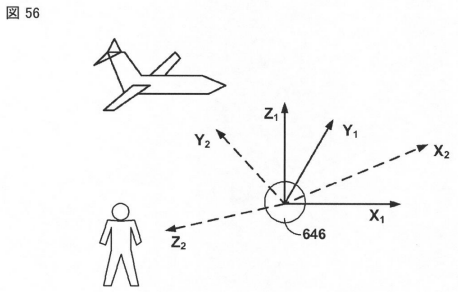


FIG. 56

【図５５Ｂ】

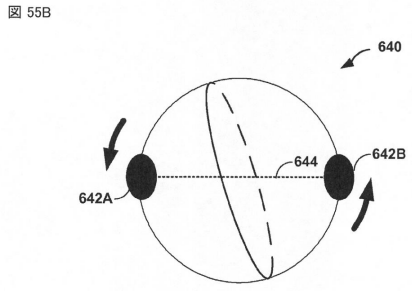


FIG. 55B

【図５７Ａ】

図 57A

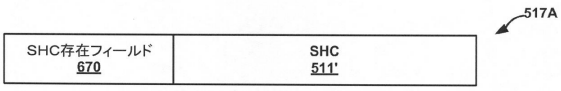


FIG. 57A

【図５７Ｂ】

図 57B

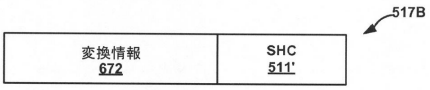


FIG. 57B

【図５７Ｃ】

図 57C

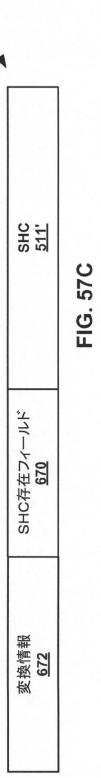


FIG. 57C

【図５７Ｄ】

図 57D

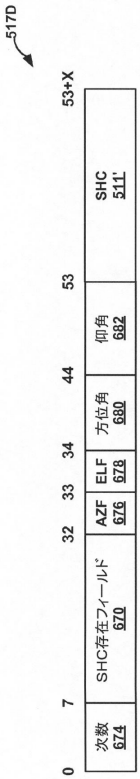


FIG. 57D

【図５７Ｅ】

図 57E

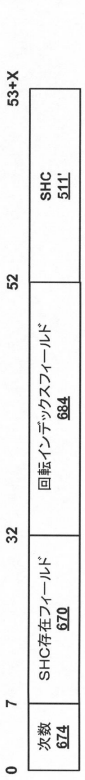


FIG. 57E

【図５８】

図 58

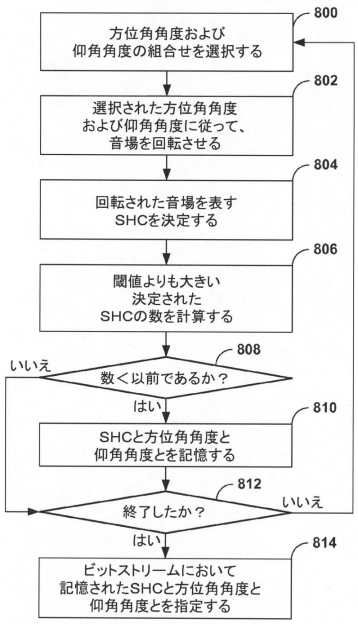


FIG. 58

【図５９】

図 59

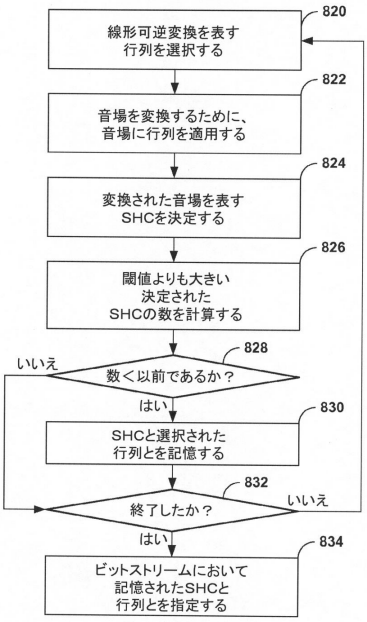


FIG. 59

フロントページの続き

(31)優先権主張番号 61/829,174
(32)優先日 平成25年5月30日(2013.5.30)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/829,155
(32)優先日 平成25年5月30日(2013.5.30)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/829,182
(32)優先日 平成25年5月30日(2013.5.30)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/829,846
(32)優先日 平成25年5月31日(2013.5.31)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/829,791
(32)優先日 平成25年5月31日(2013.5.31)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/886,617
(32)優先日 平成25年10月3日(2013.10.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/886,605
(32)優先日 平成25年10月3日(2013.10.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/899,041
(32)優先日 平成25年11月1日(2013.11.1)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/899,034
(32)優先日 平成25年11月1日(2013.11.1)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/925,074
(32)優先日 平成26年1月8日(2014.1.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/925,112
(32)優先日 平成26年1月8日(2014.1.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/925,126
(32)優先日 平成26年1月8日(2014.1.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/925,158
(32)優先日 平成26年1月8日(2014.1.8)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/933,706
(32)優先日 平成26年1月30日(2014.1.30)
(33)優先権主張国 米国(US)
(31)優先権主張番号 61/933,721
(32)優先日 平成26年1月30日(2014.1.30)
(33)優先権主張国 米国(US)
(31)優先権主張番号 62/003,515
(32)優先日 平成26年5月27日(2014.5.27)
(33)優先権主張国 米国(US)

(31)優先権主張番号　14/289,477
(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)

早期審査対象出願

(72)発明者　セン、ディパンジャン
アメリカ合衆国、カリフォルニア州　９２１２１－１７１４、サン・ディエゴ、モアハウス・ドライブ　５７７５
(72)発明者　リュ、サン－ウク
アメリカ合衆国、カリフォルニア州　９２１２１－１７１４、サン・ディエゴ、モアハウス・ドライブ　５７７５

審査官　菊池　智紀

(56)参考文献　国際公開第２００９／０６７７４１（WO，Ａ１）
安藤彰男，"３次元音響の空間的性質を利用した符号化・伝送"，日本音響学会2011年春季研究発表会講演論文集CD-ROM，２０１１年　3月　2日，pp.751-752

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１９／００－１９／２６
Ｈ０４Ｓ　１／００－　７／００
ＩＥＥＥ　Ｘｐｌｏｒｅ