

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-520785

(P2017-520785A)

(43) 公表日 平成29年7月27日(2017.7.27)

(51) Int.Cl.	F 1			テーマコード (参考)
G 10 L 19/008 (2013.01)	G 10 L	19/008	2 0 0	5 D 0 1 8
G 10 L 19/035 (2013.01)	G 10 L	19/035	Z	5 D 1 6 2
H 04 R 3/00 (2006.01)	H 04 R	3/00	3 2 0	5 D 2 2 0
H 04 R 1/40 (2006.01)	H 04 R	1/40	3 2 0 A	
H 04 S 5/00 (2006.01)	H 04 S	5/00	5 0 0	

審査請求 未請求 予備審査請求 未請求 (全 47 頁)

(21) 出願番号	特願2016-567848 (P2016-567848)	(71) 出願人	595020643 クアアルコム・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
(86) (22) 出願日	平成27年5月15日 (2015.5.15)	(74) 代理人	100108855 弁理士 蔵田 昌俊
(85) 翻訳文提出日	平成29年1月11日 (2017.1.11)	(74) 代理人	100109830 弁理士 福原 淑弘
(86) 國際出願番号	PCT/US2015/031107	(74) 代理人	100158805 弁理士 井関 守三
(87) 國際公開番号	W02015/175953	(74) 代理人	100112807 弁理士 岡田 貴志
(87) 國際公開日	平成27年11月19日 (2015.11.19)		
(31) 優先権主張番号	61/994,493		
(32) 優先日	平成26年5月16日 (2014.5.16)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	61/994,788		
(32) 優先日	平成26年5月16日 (2014.5.16)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	62/004,082		
(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)		

最終頁に続く

(54) 【発明の名称】高次アンビソニック係数の閉ループ量子化

(57) 【要約】

概して、音場の3次元表現を与えるHOA係数の閉ループ量子化のための技法が説明される。オーディオ符号化デバイスは、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、オーディオオブジェクトの閉ループ量子化を実行することができる。オーディオ復号デバイスは、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、閉ループ量子化されたオーディオオブジェクトを取得し、オーディオオブジェクトを逆量子化することができる。

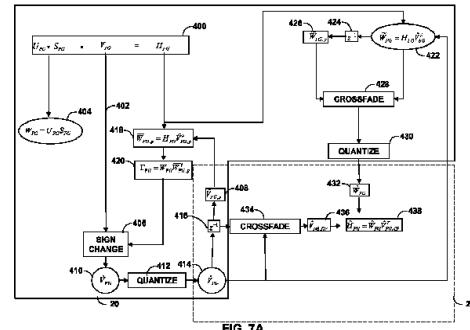


FIG. 7A

【特許請求の範囲】**【請求項 1】**

フォアグラウンドオーディオ信号を量子化するための方法であって、少なくとも1つのプロセッサによって、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、前記オーディオオブジェクトの閉ループ量子化を実行することを備える、方法。

【請求項 2】

前記オーディオオブジェクトの前記閉ループ量子化を実行することがさらに、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することと、

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、前記オーディオオブジェクトの量子化を実行することと、

を備える、請求項1に記載の方法。

【請求項 3】

前記オーディオオブジェクトの量子化を実行することがさらに、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる量子化誤差に少なくとも部分的に基づいて、前記オーディオオブジェクトの量子化を実行することを備える、請求項2に記載の方法。

【請求項 4】

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差に少なくとも部分的に基づいて、前記オーディオオブジェクトの量子化を実行することがさらに、

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差を補償することを備える、請求項3に記載の方法。

【請求項 5】

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差を補償することがさらに、

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果の擬似逆行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定することと、

前記量子化補償されたオーディオオブジェクトの量子化を実行することと、
を備える、請求項4に記載の方法。

【請求項 6】

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果の前記擬似逆行列に少なくとも部分的に基づいて、前記量子化補償されたオーディオオブジェクトを決定することがさらに、

高次アンビソニック(HOA)係数と、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した前記結果の前記擬似逆行列との積として、前記量子化補償されたオーディオオブジェクトを決定することを備える、請求項5に記載の方法。

【請求項 7】

前記オーディオオブジェクトおよび前記指向性情報が、高次アンビソニック係数から分解され、

前記オーディオオブジェクトが、複数の球面調和係数の左特異ベクトルを表すU行列と、前記複数の球面調和係数の特異値を表すS行列との積を備え、

前記オーディオオブジェクトに関連付けられる前記指向性情報が、前記複数の球面調和係数の右特異ベクトルを表すV行列を備える、請求項1に記載の方法。

【請求項 8】

フォアグラウンドオーディオ信号を量子化するためのデバイスであって、
オーディオオブジェクトと、前記オーディオオブジェクトに関連付けられる指向性情報

10

20

30

40

50

とを記憶するように構成されるメモリと、
少なくとも1つのプロセッサであって、

前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、前記オーディオオブジェクトの閉ループ量子化を実行する、ように構成されるプロセッサと、
を備える、デバイス。

【請求項9】

前記少なくとも1つのプロセッサがさらに、
前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行し、
前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、前記オーディオオブジェクトの量子化を実行する、
ように構成される、請求項8に記載のデバイス。
10

【請求項10】

前記オーディオオブジェクトの量子化を実行することがさらに、
前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる量子化誤差に少なくとも部分的に基づいて、前記オーディオオブジェクトの量子化を実行することを備える、請求項9に記載のデバイス。

【請求項11】

前記少なくとも1つのプロセッサがさらに、
前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差を補償するように構成される、請求項10に記載のデバイス。
20

【請求項12】

前記少なくとも1つのプロセッサがさらに、
前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した結果の擬似逆行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定し、
前記量子化補償されたオーディオオブジェクトの量子化を実行する、
ように構成される、請求項11に記載のデバイス。

【請求項13】

前記少なくとも1つのプロセッサがさらに、
高次アンビソニック(HOA)係数と、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した前記結果の前記擬似逆行列との積として、前記オーディオオブジェクトを決定するように構成される、請求項12に記載のデバイス。
30

【請求項14】

前記HOA係数を示すオーディオデータをキャプチャするように構成されるマイクロフォンをさらに備える、請求項13に記載のデバイス。

【請求項15】

オーディオオブジェクトを逆量子化するための方法であって、
少なくとも1つのプロセッサによって、オーディオオブジェクトに関連付けられる指向性情報を量子化を実行した結果に少なくとも部分的に基づいて、閉ループ量子化された前記オーディオオブジェクトを取得することと、
40

前記少なくとも1つのプロセッサによって、前記オーディオオブジェクトを逆量子化することと、
を備える、方法。

【請求項16】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化した結果に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項15に記載の方法。

【請求項17】

10

20

30

40

50

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することから生じる量子化誤差に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 15 に記載の方法。

【請求項 18】

前記オーディオオブジェクトは、

前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差を補償することを含む、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することから生じる前記量子化誤差に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 15 に記載の方法。

10

【請求項 19】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化した結果の逆擬似行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定することと、前記量子化補償されたオーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 15 に記載の方法。

20

【請求項 20】

前記オーディオオブジェクトは、高次アンビソニック (H O A) 係数と、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した前記結果の前記擬似逆行列との積として、前記量子化補償されたオーディオオブジェクトを決定することによって、閉ループ量子化される、請求項 19 に記載の方法。

【請求項 21】

前記オーディオオブジェクトおよび前記指向性情報は、高次アンビソニック係数から分解され、

前記オーディオオブジェクトは、複数の球面調和係数の左特異ベクトルを表す U 行列と、前記複数の球面調和係数の特異値を表す S 行列との積を備え、

前記オーディオオブジェクトに関連付けられる前記指向性情報は、前記複数の球面調和係数の右特異ベクトルを表す V 行列を備える、

30

請求項 15 に記載の方法。

【請求項 22】

ピットストリームを受信することと、

前記閉ループ量子化されたオーディオオブジェクトおよび前記量子化された指向性情報を取得するために、前記ピットストリームを復号することと、

をさらに備える、請求項 15 に記載の方法。

【請求項 23】

フォアグラウンドオーディオ信号を逆量子化するためのデバイスであって、

40

オーディオオブジェクトを記憶するように構成されるメモリと、

少なくとも 1 つのプロセッサであって、

前記オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、閉ループ量子化された前記オーディオオブジェクトを取得し

前記オーディオオブジェクトを逆量子化する、

よう構成されるプロセッサと、

を備えるデバイス。

【請求項 24】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性

50

情報を量子化した結果に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 2 3 に記載のデバイス。

【請求項 2 5】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することから生じる量子化誤差に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 2 3 に記載のデバイス。

【請求項 2 6】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行することから生じる前記量子化誤差を補償することを含む、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することから生じる前記量子化誤差に少なくとも部分的に基づいて前記オーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 2 3 に記載のデバイス。 10

【請求項 2 7】

前記オーディオオブジェクトは、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化することと、前記オーディオオブジェクトに関連付けられる前記指向性情報を量子化した結果の逆擬似行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定することと、前記量子化補償されたオーディオオブジェクトを量子化することによって、閉ループ量子化される、請求項 2 3 に記載のデバイス。 20

【請求項 2 8】

前記オーディオオブジェクトは、高次アンビソニック (H O A) 係数と、前記オーディオオブジェクトに関連付けられる前記指向性情報の量子化を実行した前記結果の前記擬似逆行列との積として、前記量子化補償されたオーディオオブジェクトを決定することによって、閉ループ量子化される、請求項 2 7 に記載のデバイス。

【請求項 2 9】

前記逆量子化されたオーディオオブジェクトからレンダリングされたラウドスピーカーフィードを再生するように構成されるスピーカーをさらに備える、請求項 2 3 に記載のデバイス。 30

【請求項 3 0】

前記少なくとも 1 つのプロセッサがさらに、
ビットストリームを受信し、

前記閉ループ量子化されたオーディオオブジェクトおよび前記量子化された指向性情報を取得するために、前記ビットストリームを復号する、

ように構成される、請求項 2 3 に記載のデバイス。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

[0001]本出願は、以下の米国仮出願、すなわち、
2014年5月16日に出願された「CLOSED LOOP QUANTIZATION OF HIGHER ORDER AMBIISONIC COEFFICIENTS」という名称の米国仮出願第61/994,493号、

2014年5月16日に出願された「CLOSED LOOP QUANTIZATION OF HIGHER ORDER AMBIISONIC COEFFICIENTS」という名称の米国仮出願第61/994,788号、および

2014年5月28日に出願された「CLOSED LOOP QUANTIZATION OF HIGHER ORDER AMBIISONIC COEFFICIENTS」という名称の米国仮出願第62/004,082号

の利益を主張し、上記に記載された米国仮出願の各々は、それらのそれぞれの全体として

40

50

本明細書に記載されたかのように、参照により組み込まれる。

【0002】

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

【背景技術】

【0003】

[0003]高次アンビソニクス(ambisonic)(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しばしば表される)は、音場の3次元表現である。このHOA表現またはSHC表現は、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に依存しない方法で、音場を表し得る。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、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、音場のより良い表現を可能にし得る。

10

【発明の概要】

【0004】

[0004]概して、音場の3次元表現を与えるHOA係数の閉ループ量子化のための技法が説明される。オーディオエンコーダは、閉ループ量子化プロセスを用いて、オーディオオブジェクトおよびオーディオオブジェクトに関連付けられる指向性情報(いずれも、HOA係数から分解され得る)の個別の独立した量子化を実行する代わりに、オーディオオブジェクトに関連付けられる量子化された指向性情報に少なくとも部分的に基づいて、オーディオオブジェクトを量子化し得る。このようにして、量子化されたオーディオオブジェクトは、オーディオオブジェクトに関連付けられる指向性情報を量子化することから生じる量子化誤差を補償し得る。さらに、閉ループ量子化を介して符号化された音場の3次元表現は、開ループ量子化を介して符号化された音場の3次元表現より相対的に少ない量子化誤差でデコーダによって再構成され得る。

20

【0005】

[0005]一態様では、フォアグラウンドオーディオ信号を量子化するための方法が、少なくとも1つのプロセッサによって、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、オーディオオブジェクトの閉ループ量子化を実行することを備える。

30

【0006】

[0006]別の態様では、フォアグラウンドオーディオ信号を量子化するためのデバイスはメモリを含む。そのデバイスはさらに、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、オーディオオブジェクトの閉ループ量子化を実行するように構成される少なくとも1つのプロセッサを含む。

【0007】

[0007]別の態様では、オーディオオブジェクトを逆量子化するための方法が、少なくとも1つプロセッサにおいて、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行する結果に少なくとも部分的に基づいて閉ループ量子化されたオーディオオブジェクトを取得することと、少なくとも1つのプロセッサによって、オーディオオブジェクトを逆量子化することとを備える。

40

【0008】

[0008]別の態様では、オーディオオブジェクトを逆量子化するためのデバイスがメモリを含む。そのデバイスはさらに、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて閉ループ量子化されたオーディオオブジェクトを取得し、オーディオオブジェクトを逆量子化するように構成される少なくとも1つのプロセッサを含む。

【0009】

50

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

【図面の簡単な説明】

【0010】

【図1】[0010]様々な次数および副次数の球面調和基底関数を示す図。

【図2】[0011]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を示す図。

【図3】[0012]本開示で説明される技法の様々な態様を実行し得る、図2の例に示されるオーディオ符号化デバイスの一例をより詳細に示すブロック図。

【図4】[0013]図2のオーディオ復号デバイスをより詳細に示すブロック図。 10

【図5A】[0014]本開示で説明されるベクトルベース合成技法の様々な態様を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャート。

【図5B】[0015]本開示で説明されるコーディング技法の様々な態様を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャート。

【図6A】[0016]本開示で説明される技法の様々な態様を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャート。

【図6B】[0017]本開示で説明されるコーディング技法の様々な態様を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャート。

【図7A】[0018]H O A信号圧縮のための、1つまたは複数のVベクトルの形の指向性情報と、フォアグラウンド信号の形のオーディオオブジェクトとの閉ループ量子化を示すブロック図。 20

【図7B】H O A信号圧縮のための、1つまたは複数のVベクトルの形の指向性情報と、フォアグラウンド信号の形のオーディオオブジェクトとの閉ループ量子化を示すブロック図。

【図8】[0019]図3の例に示されるオーディオ符号化デバイスの量子化ユニットの一例をさらに詳細に示すブロック図。

【図9】[0020]図3の例に示されるオーディオ符号化デバイスの量子化補償ユニットの一例をさらに詳細に示すブロック図。

【発明を実施するための形態】

【0011】

[0021]サラウンドサウンドの発展は、現今では娛樂のための多くの出力フォーマットを利用可能にしている。そのような消費者向けのサラウンドサウンドフォーマットの例は、ある幾何学的な座標にあるラウドスピーカーへのフィードを暗黙的に指定するという点で、大半が「チャネル」ベースである。消費者向けのサラウンドサウンドフォーマットは、普及している5.1フォーマット（これは、次の6つのチャネル、すなわち、フロントレフト（F L）と、フロントライト（F R）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（L F E）とを含む）、発展中の7.1フォーマット、7.1.4フォーマットおよび2.2.2フォーマット（たとえば、超高精細度テレビジョン規格とともに使用するための）などのハイツスピーカーを含む様々なフォーマットを含む。消費者向けではないフォーマットは、「サラウンドアレイ」と呼ばれることが多い（対称な、および非対称な幾何学的配置の）任意の数のスピーカーに及び得る。そのようなアレイの一例は、切頂二十面体の角の座標に配置される32個のラウドスピーカーを含む。 40

【0012】

[0022]将来のM P E Gエンコーダへの入力は、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單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（P C M）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S

10

20

30

40

50

H C、「高次アンビソニックス」すなわち H O A、および「H O A 係数」とも呼ばれる)を使用して音場を表すことを伴うシーンベースオーディオのうちの1つである。将来のMPEGエンコーダは、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>において入手可能な、International Organization for Standardization/International Electrotechnical Commission (ISO)/(IEC) JTC1/SC29/WG11/N13411による「Call for Proposals for 3D Audio」と題される文書においてより詳細に説明され得る。

10

【0013】

[0023]市場には様々な「サラウンドサウンド」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、各々のスピーカー構成のためにサウンドトラックをリミックスする努力を行うことを望まない。最近では、規格開発組織が、規格化されたビットストリームへの符号化と、スピーカーの幾何学的配置(と数)および(レンダラを伴う)再生のロケーションにおける音響条件に適応可能でありそれらに依存しない後続の復号とを提供するための方法を考えている。

20

【0014】

[0024]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。

30

【0015】

[0025]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。

【数1】

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

【0016】

[0026]この式は、時間tにおける音場の任意の点 $\{r_r, \theta_r, \varphi_r\}$ における圧力 p_i が、SHC、 $A_n^m(k)$ によって一意に表され得ることを示す。ここで、

【数2】

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

40

であり、cは音速(約343m/s)であり、 $\{r_r, \theta_r, \varphi_r\}$ は基準点(または観測点)であり、 $j_n(\cdot)$ は次数nの球ベッセル関数であり、 $Y_n^m(\cdot)$ は次数nおよび副次数mの球面調和基底関数である。角括弧内の項は、離散フーリエ変換(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(r_r, θ_r, φ_r))であることが認識できよう。階層セットの他の例は、ウェーブレット変換係数のセット、および多分解能基底関数の係数の他のセットを含む。

【0017】

50

[0027] 図 1 は、0 次 ($n = 0$) から 4 次 ($n = 4$) までの球面調和基底関数を示す図である。理解できるように、各次数に対して、説明を簡単にするために図示されているが図 1 の例では明示的に示されていない副次数 m の拡張が存在する。

【0018】

[0028] S H C $A^m_n(k)$ は、様々なマイクロフォンアレイ構成によって物理的に獲得（たとえば、録音）することができ、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ることができる。S H C はシーンベースのオーディオを表し、ここで、S H C は、より効率的な送信または記憶を促し得る符号化された S H C を取得するために、オーディオエンコーダに入力され得る。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。

10

【0019】

[0029] 上述されたように、S H C は、マイクロフォンアレイを使用したマイクロフォン録音から導出され得る。S H C がマイクロフォンアレイからどのように導出され得るかの様々な例は、Poletti, M、「Three-Dimensional Surround Sound Systems Based on Spherical Harmonics」、J. Audio Eng. Soc., Vol. 53, No. 11, 2005 年 11 月、1004 ~ 1025 ページにおいて説明されている。

【0020】

[0030] S H C がどのようにオブジェクトベースの記述から導出され得るかを例示するために、次の式を考える。個々のオーディオオブジェクトに対応する音場についての係数 $A^m_n(k)$ は、

20

【数 3】

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

と表され得、ただし、 i は、

【数 4】

$$\sqrt{-1}$$

であり、 $h^{(2)}_n(\cdot)$ は次数 n の（第 2 種）球ハンケル関数であり、 $\{r_s, \theta_s, \varphi_s\}$ はオブジェクトのロケーションである。周波数の関数として（たとえば、PCM ストリームに対して高速フーリエ変換を実行するなど、時間周波数分析技法を使用して）オブジェクトソースエネルギー $g(\cdot)$ を知ることで、各 PCM オブジェクトと対応するロケーションとを S H C $A^m_n(k)$ に変換することが可能となる。さらに、各オブジェクトについての $A^m_n(k)$ 係数は、（上式は線形および直交の分解であるので）加法的であることが示され得る。このようにして、多数の PCM オブジェクトが $A^m_n(k)$ 係数によって（たとえば、個々のオブジェクトについての係数ベクトルの和として）表され得る。本質的に、これらの係数は、音場についての情報（3D 座標の関数としての圧力）を含んでおり、上記は、観測点 $\{r_r, \theta_r, \varphi_r\}$ 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 S H C ベースのオーディオコーディングの文脈で説明される。

30

【0021】

[0031] 図 2 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 2 の例に示されるように、システム 10 は、コンテンツ作成者デバイス 12 と、コンテンツ消費者デバイス 14 とを含む。コンテンツ作成者デバイス 12 およびコンテンツ消費者デバイス 14 の文脈で説明されているが、本技法は、オーディオデータを表すビットストリームを形成するために、S H C (H O A 係数とも呼ばれ得る) または音場の任意の他の階層的表現が符号化される任意の文脈で実施され得る。その上、コンテンツ作成者デバイス 12 は、いくつか例を挙げると、ハンドセット（またはセルラーフォン）、タブレットコンピュータ、スマートフォン、またはデスクトップコンピュータを含む、

40

50

本開示で説明される技法を実施することが可能な任意の形態のコンピューティングデバイスを表し得る。同様に、コンテンツ消費者デバイス14は、いくつか例を挙げると、ハンドセット（またはセルラーフォン）、タブレットコンピュータ、スマートフォン、セットトップボックス、またはデスクトップコンピュータを含む、本開示で説明される技法を実施することが可能な任意の形態のコンピューティングデバイスを表し得る。

【0022】

[0032]コンテンツ作成者デバイス12は、コンテンツ消費者デバイス14などのコンテンツ消費者デバイスのオペレータによる消費のためのマルチチャネルオーディオコンテンツを生成し得る、映画スタジオまたは他のエンティティによって操作され得る。いくつかの例では、コンテンツ作成者デバイス12は、HOA係数11を圧縮することを望む個人ユーザによって操作され得る。多くの場合、コンテンツ作成者は、ビデオコンテンツとともに、オーディオコンテンツを生成する。コンテンツ消費者デバイス14は、個人によって操作され得る。コンテンツ消費者デバイス14は、マルチチャネルオーディオコンテンツとしての再生のためにSHCをレンダリングすることが可能な任意の形態のオーディオ再生システムを指し得る、オーディオ再生システム16を含み得る。

10

【0023】

[0033]コンテンツ作成者デバイス12は、オーディオ編集システム18を含む。コンテンツ作成者デバイス12は、様々なフォーマットのライブ録音7（HOA係数として直接含む）とオーディオオブジェクト9とを取得し、コンテンツ作成者デバイス12は、オーディオ編集システム18を使用してこれらを編集し得る。マイクロフォン5はライブ録音7をキャプチャし得る。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特定しようとして、レンダリングされたスピーカーフィードを聞き得る。コンテンツ作成者デバイス12は次いで、（潜在的に、上記で説明された方法でソースHOA係数がそれから導出され得るオーディオオブジェクト9のうちの様々なオブジェクトの操作を通じて間接的に）HOA係数11を編集し得る。コンテンツ作成者デバイス12は、HOA係数11を生成するためにオーディオ編集システム18を採用し得る。オーディオ編集システム18は、オーディオデータを編集し、このオーディオデータを1つまたは複数のソース球面調和係数として出力することが可能な任意のシステムを表す。

20

【0024】

[0034]編集プロセスが完了すると、コンテンツ作成者デバイス12は、HOA係数11に基づいてビットストリーム21を生成し得る。すなわち、コンテンツ作成者デバイス12は、ビットストリーム21を生成するために、本開示で説明される技法の様々な様相に従って、HOA係数11を符号化またはさもなければ圧縮するように構成されたデバイスを表す、オーディオ符号化デバイス20を含む。オーディオ符号化デバイス20は、一例として、有線チャネルまたはワイヤレスチャネル、データ記憶デバイスなどであり得る送信チャネルを介した送信のために、ビットストリーム21を生成し得る。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し得、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。

30

【0025】

[0035]図2では、コンテンツ消費者デバイス14に直接的に送信されるものとして示されているが、コンテンツ作成者デバイス12は、コンテンツ作成者デバイス12とコンテンツ消費者デバイス14との間に配置された中間デバイスにビットストリーム21を出力し得る。中間デバイスは、ビットストリームを要求し得るコンテンツ消費者デバイス14に後で配信するために、ビットストリーム21を記憶し得る。中間デバイスは、ファイルサーバ、ウェブサーバ、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、モバイルフォン、スマートフォン、または後のオーディオデコーダ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他のデバイスを備え得る。中間デバイスは、ビットストリーム21を要求するコンテンツ消費者デバイス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

40

50

ットストリームを送信するとともに、ストリーミングすることが可能なコンテンツ配信ネットワーク内に存在し得る。

【0026】

[0036] 代替的に、コンテンツ作成者デバイス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を記憶し得、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。この文脈において、送信チャネルは、これらの媒体に記憶されたコンテンツが送信されるチャネルを指し得る（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2の例に限定されるべきではない。10

【0027】

[0037] 図2の例にさらに示されるように、コンテンツ消費者デバイス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し得、異なる形態のレンダリングは、ベクトルベース振幅パンニング（V B A P : vector-base amplitude panning）を実行する様々な方法の1つもしくは複数、および／または音場合成を実行する様々な方法の1つもしくは複数を含み得る。本明細書で使用される場合、「Aおよび／またはB」は、「AまたはB」、または「AとB」の両方を意味する。20

【0028】

[0038] オーディオ再生システム16は、オーディオ復号デバイス24をさらに含み得る。オーディオ復号デバイス24は、ビットストリーム21からH O A係数11'を復号するように構成されたデバイスを表し得、H O A係数11'は、H O A係数11と類似し得るが、損失のある演算（たとえば、量子化）および／または送信チャネルを介した送信が原因で異なり得る。オーディオ再生システム16は、H O A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、ラウドスピーカーフィード25を出力するためにH O A係数11'をレンダリングし得る。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2の例には示されていない）を駆動し得る。30

【0029】

[0039] 適切なレンダラを選択するために、またはいくつかの場合には、適切なレンダラを生成するために、オーディオ再生システム16は、ラウドスピーカーの数および／またはラウドスピーカーの空間的な幾何学的配置を示すラウドスピーカー情報13を取得し得る。いくつかの場合には、オーディオ再生システム16は、基準マイクロフォンを使用してラウドスピーカー情報13を取得し、ラウドスピーカー情報13を動的に決定するような方法でラウドスピーカーを駆動し得る。他の場合には、またはラウドスピーカー情報13の動的な決定とともに、オーディオ再生システム16は、オーディオ再生システム16とインターフェースをとりラウドスピーカー情報13を入力するようにユーザに促し得る。40

【0030】

[0040] オーディオ再生システム16は次いで、ラウドスピーカー情報13に基づいて、オーディオレンダラ22のうちの1つを選択し得る。いくつかの場合には、オーディオ再生システム16は、ラウドスピーカー情報13において指定された幾何学的配置に対する何らかの閾値に類似した尺度（ラウドスピーカーの幾何学的配置に関する）内にいずれのオーディオレンダラ22もないとき、ラウドスピーカー情報13に基づいて、オーディオレンダラ22のうちの1つを生成し得る。オーディオ再生システム16は、いくつかの場合には、オーディオレンダラ22のうちの既存の1つを選択することを最初に試みることなく、ラウドスピーカー情報13に基づいて、オーディオレンダラ22のうちの1つを生50

成し得る。その際、1つまたは複数のスピーカー3は、レンダリングされたラウドスピーカーフィード25を再生し得る。

【0031】

[0041]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、図2の例に示されるオーディオ符号化デバイス20の一例をより詳細に示すブロック図である。オーディオ符号化デバイス20は、コンテンツ分析ユニット26と、ベクトルベース分解ユニット27と、指向性ベース分解ユニット28とを含む。以下で手短に説明されるが、オーディオ符号化デバイス20に関するより多くの情報、およびHOA係数を圧縮またはさもなければ符号化する様々な態様は、2014年5月29に出願された「INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の国際特許出願公開第WO2014/194099号において入手可能である。
10

【0032】

[0042]コンテンツ分析ユニット26は、HOA係数11がライブ録音から生成されたコンテンツを表すか、オーディオオブジェクトから生成されたコンテンツを表すかを特定するために、HOA係数11のコンテンツを分析するように構成されたユニットを表す。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し得る。いくつかの場合には、フレーム化された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をベクトルベース分解ユニット27に渡す。いくつかの場合には、フレーム化された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を指向性ベース分解ユニット28に渡す。指向性ベース分解ユニット28は、指向性ベースビットストリーム21を生成するためにHOA係数11の指向性ベース合成を実行するように構成されたユニットを表し得る。
20

【0033】

[0043]図3の例に示されるように、ベクトルベース分解ユニット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、並べ替えユニット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、聴覚心理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。
30

【0034】

[0044]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、副次数に関連付けられた係数のブロックまたはフレーム(HOA[k]と示され得、ただし、kはサンプルの現在のフレームまたはブロックを示し得る)を表す。HOA係数11の行列は、次元D:M × (N+1)²を有し得る。
40

【0035】

[0045]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た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、それとは反対に特に明記されていない限り、非0のセットを指すことが意図され、いわゆる「空集合」を含む集合の古典的な数学的定義を指すことは意図されない。代替的な変換は、「PCA」と呼ばれることが多い、主成分分析を備え得る。文脈に応じて、PCAは、いくつかの例を挙げれば、離散カールーネン・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得る。オーディオデータを圧縮するという背後にある目標につながるそのような演算の特性は、マルチチャネルオーディオデータの「エネルギー圧縮」および「無相関化」である。
50

【0036】

[0046] いざれにしても、LITユニット30が、例として、特異値分解（やはり「SVD」と呼ばれ得る）を実行すると仮定すると、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し得る。変換されたHOA係数の「セット」は、変換されたHOA係数のベクトルを含み得る。図3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を実行し得る。SVDは、線形代数学では、 $y \times z$ の実行列または複素行列X（ここで、Xは、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し得る。

$$X = U S V^*$$

10

Uは $y \times y$ の実ユニタリー行列または複素ユニタリー行列を表し得、ここで、Uのy個の列は、マルチチャネルオーディオデータの左特異ベクトルとして知られる。Sは、対角線上に非負実数をもつ $y \times z$ の矩形対角行列を表し得、ここで、Sの対角要素の値は、マルチチャネルオーディオデータの特異値として知られる。 V^* （Vの共役転置を示し得る）は $z \times z$ の実ユニタリー行列または複素ユニタリー行列を表し得、ここで、 V^* のz個の列は、マルチチャネルオーディオデータの右特異ベクトルとして知られる。

【0037】

[0047] 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得る。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に適用され得、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の適用を含み得る。

20

【0038】

30

[0048] このようにして、LITユニット30は、次元D： $M \times (N + 1)^2$ を有するUS[k]ベクトル33（SベクトルとUベクトルとの組み合わされたバージョンを表し得る）と、次元D： $(N + 1)^2 \times (N + 1)^2$ 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を実行し得る。US[k]行列中の個々のベクトル要素は $X_{ps}(k)$ とも呼ばれ得、一方、V[k]行列の個々のベクトルは $v(k)$ とも呼ばれ得る。

【0039】

40

[0049] U行列、S行列、およびV行列の分析は、それらの行列がXによって上で表される背後の音場の空間的および時間的な特性を伝え、または表すということを明らかにし得る。（M個のサンプルの長さの）U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特性（指向性情報とも呼ばれ得る）とは切り離されている、正規化された分離されたオーディオ信号を表し得る。空間的な形状と位置（r、θ、φ）とを表す空間特性は代わりに、（各々が $(N + 1)^2$ の長さの）V行列の中の個々のi番目のベクトル、 $v^{(i)}(k)$ によって表され得る。 $v^{(i)}(k)$ ベクトルの各々の個々の要素は、関連付けられたオーディオオブジェクトのための音場の形状（幅を含む）と位置とを記述するHOA係数を表し得る。U行列中のベクトルとV行列中のベクトルの両方が、それらの2乗平均エネルギーが1に等しくなるように正規化される。したがって、Uの中のオーディオ信号のエネルギーは、Sの中の対角線要素によって表される。したがって、US[k]（個々のベクトル要素 $X_{ps}(k)$ を有する）を形成するために、UとSとを乗算することは、エネルギーを有

50

するオーディオ信号を表す。(Uにおける)オーディオ時間信号と、(Sにおける)それらのエネルギーと、(Vにおける)それらの空間特性とを切り離すSVD分解の能力は、本開示で説明される技法の様々な様態を支援し得る。さらに、背後のHOA[k]係数XをUS[k]とV[k]とのベクトル乗算によって合成するモデルは、「ベクトルベース分解」という用語を生じさせ、それは本文書全体で使用される。さらに、本明細書全体を通して、エネルギーを伴うオーディオ信号を表すUS[k]は、「オーディオオブジェクト」または「フォアグラウンドオーディオ信号」という用語で呼ばれ得、V[k]は、「オーディオオブジェクトに関連付けられる指向性情報」または「フォアグラウンド信号に関連付けられる指向性情報」と呼ばれ得る。HOA[k]係数は、HOA係数と呼ばれ得、HOA係数 = $US[k]^*V[k]$ であるか、または言い換えると、HOA係数は、オーディオオブジェクト(US[k])とオーディオオブジェクトに関連付けられる指向性情報(V[k])との積である。

【0040】

[0050] 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の派生物に適用し得る。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を適用し得る。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は潜在的に、プロセッササイクルおよび記憶空間のうちの1つまたは複数に関して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符号化効率を達成し得る。

【0041】

[0051] 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(、r)、およびエネルギー特性(e)などの様々なパラメータを計算するように構成されたユニットを表す。現在のフレームのためのパラメータの各々は、R[k]、[k]、[k]、r[k]、およびe[k]として示され得る。パラメータ計算ユニット32は、パラメータを特定するために、US[k]ベクトル33に関してエネルギー分析および/または相関(もしくはいわゆる相互相関)を実行し得る。パラメータ計算ユニット32はまた、以前のフレームのためのパラメータを決定することができ、ここで、以前のフレームパラメータは、US[k-1]ベクトルおよびV[k-1]ベクトルの以前のフレームに基づいて、R[k-1]、[k-1]、[k-1]、r[k-1]、およびe[k-1]と示され得る。パラメータ計算ユニット32は、現在のパラメータ37と以前のパラメータ39とを並べ替えユニット34に出力し得る。

【0042】

[0052] パラメータ計算ユニット32によって計算されるパラメータは、オーディオオブジェクトの自然な評価または時間的な継続性を表すようにオーディオオブジェクトを並べ替えるために、並べ替えユニット34によって使用され得る。並べ替えユニット34は、第1のUS[k]ベクトル33からのパラメータ37の各々を、第2のUS[k-1]ベクトル33のためのパラメータ39の各々に対して順番ごとに比較し得る。並べ替えユニット34は、並べ替えられたUS[k]行列33' (数学的には

【数5】

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

として示され得る)と、並べ替えられたV[k]行列35' (数学的には

【数6】

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

として示され得る)とをフォアグラウンドサウンド(または支配的サウンド - PS (pred

10

20

30

40

50

ominant sound)) 選択ユニット 3 6 (「フォアグラウンド選択ユニット 3 6」) およびエネルギー補償ユニット 3 8 に出力するために、現在のパラメータ 3 7 および以前のパラメータ 3 9 に基づいて、U S [k] 行列 3 3 およびV [k] 行列 3 5 内の様々なベクトルを(一例として、ハンガリー法を使用して)並べ替え得る。

【0043】

[0053]音場分析ユニット 4 4 は、目標ビットレート 4 1 を潜在的に達成するために、H O A 係数 1 1 について音場分析を実行するように構成されたユニットを表し得る。音場分析ユニット 4 4 は、その分析および/または受信された目標ビットレート 4 1 に基づいて、聴覚心理コーダのインスタンス化の総数(環境またはバックグラウンドチャネルの総数($B G_{TOT}$)とフォアグラウンドチャネル、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)を決定し得る。聴覚心理コーダのインスタンス化の総数は、num HOA Transport Channelsとして示されることができる。
10

【0044】

[0054]音場分析ユニット 4 4 はまた、やはり目標ビットレート 4 1 を潜在的に達成するために、フォアグラウンドチャネルの総数($n FG$) 4 5 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Min Amb HOA order)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($n BG_a = (Min Amb HOA order + 1)^2$)と、送るべき追加のBG HOAチャネルのインデックス(i)(図3の例ではバックグラウンドチャネル情報 4 3 として総称的に示され得る)とを決定し得る。バックグラウンドチャネル情報 4 3 は、環境チャネル情報 4 3 とも呼ばれ得る。num HOA Transport Channels - n BG a 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。一態様では、チャネルタイプは、2ビットによって(「Channel Type」として)示されたシンタックス要素であり得る(たとえば、0 0 : 指向性ベースの信号、0 1 : ベクトルベースの支配的な信号、1 0 : 追加の環境信号、1 1 : 非アクティブな信号)。バックグラウンド信号または環境信号の総数、n BG a は、 $(Min Amb HOA order + 1)^2 +$ (上記の例における)インデックス 1 0 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回数によって与えられ得る。
20

【0045】

[0055]音場分析ユニット 4 4 は、目標ビットレート 4 1 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 4 1 が比較的高いとき(たとえば、目標ビットレート 4 1 が 5 1 2 K b p s 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し得る。一態様では、ビットストリームのヘッダセクションにおいて、num HOA Transport Channels は 8 に設定され得るが、一方で、Min Amb HOA order は 1 に設定され得る。このシナリオでは、各フレームにおいて、音場のバックグラウンド部分または環境部分を表すために 4 つのチャネルが確保され得るが、一方で、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/環境チャネルまたはフォアグラウンド/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、上記で説明されたように、ベクトルベースの信号または指向性ベースの信号のいずれかの 1 つであり得る。
30

【0046】

[0056]いくつかの場合には、フレームのためのベクトルベースの支配的な信号の総数は、そのフレームのビットストリームにおいて Channel Type インデックスが 0 1 である回数によって与えられ得る。上記の態様では、各々の追加のバックグラウンド/環境チャネル(たとえば、1 0 という Channel Type に対応する)に対して、(最初の 4 つ以外の)あり得る H O A 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
40

応する情報。その情報は、4次のHOAコンテンツについては、HOA係数5～25を示すためのインデックスであり得る。最初の4つの環境HOA係数1～4は、minAmbHOAorderが1に設定されるときは常に送られ得、したがって、オーディオ符号化デバイスは、5～25のインデックスを有する追加の環境HOA係数のうちの1つを示すことのみが必要であり得る。その情報はしたがって、「Coded AmbCoeffId」として示され得る、(4次のコンテンツのための)5ビットのシンタックス要素を使用して送られ得る。いずれにしても、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36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。

10

【0047】

[0057]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、送るべき追加のBG HOAチャネルの数($nBGa$)およびインデックス(i)と)に基づいて、バックグラウンドまたは環境HOA係数47を決定するように構成されたユニットを表し得る。たとえば、 N_{BG} が1に等しいとき、バックグラウンド選択ユニット48は、1以下の次数を有するオーディオフレームの各サンプルのHOA係数11を選択し得る。バックグラウンド選択ユニット48は次いで、この例では、インデックス(i)のうちの1つによって特定されるインデックスを有するHOA係数11を、追加のBG HOA係数として選択することができ、ここで、 $nBGa$ は、図2および図4の例に示されるオーディオ復号デバイス24などのオーディオ復号デバイスがビットストリーム21からバックグラウンドHOA係数47を解析することを可能にするために、ビットストリーム21において指定するために、ビットストリーム生成ユニット42に提供される。バックグラウンド選択ユニット48は次いで、環境HOA係数47をエネルギー補償ユニット38に出力し得る。環境HOA係数47は、次元D： $M \times [(N_{BG} + 1)^2 + nBGa]$ を有し得る。環境HOA係数47はまた、「環境HOA係数47」と呼ばれ得、ここで、環境HOA係数47の各々は、聴覚心理オーディオコーダユニット40によって符号化されるべき別個の環境HOAチャネルに対応する。

20

【0048】

[0058]フォアグラウンド選択ユニット36は、(フォアグラウンドベクトルを特定する1つまたは複数のインデックスを表し得る) $nFG45$ に基づいて、音場のフォアグラウンド成分または明瞭な成分を表す、並べ替えられた $U_S[k]$ 行列33' と並べ替えられた $V[k]$ 行列35' とを選択するように構成されたユニットを表し得る。フォアグラウンド選択ユニット36は、(並べ替えられた $U_S[k]_1, \dots, nFG49$ 、 $F_G_1, \dots, nFG[k]49$ 、または $X(1 \dots nFG)_{PS}(k)49$ として示され得る) nFG 信号49を、量子化補償ユニット70に出力することができ、ここで、 nFG 信号49は次元D： $M \times nFG$ 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並べ替えられた $V[k]$ 行列35' (または $v^{(1 \dots nFG)}(k)35'$) 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、フォアグラウンド成分に対応する並べ替えられた $V[k]$ 行列35' のサブセットは、次元D： $(N + 1)^2 \times nFG$ 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、

30

【数7】

$$\bar{V}_{1, \dots, nFG}[k]$$

として数学的に示され得る)。

【0049】

[0059]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チャネルのうち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たユニットを表し得る

50

。エネルギー補償ユニット38は、並べ替えられたU S [k] 行列33'、並べ替えられたV [k] 行列35'、n F G信号49、フォアグラウンドV [k] ベクトル51_k、および環境H O A係数47のうちの1つまたは複数に関してエネルギー分析を実行し、次いで、エネルギー補償された環境H O A係数47'を生成するために、そのエネルギー分析に基づいてエネルギー補償を実行し得る。エネルギー補償ユニット38は、エネルギー補償された環境H O A係数47'を聴覚心理オーディオコーダユニット40に出力し得る。

【0050】

[0060]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と以前のフレームに関する（したがってk-1という表記である）フォアグラウンドV [k - 1] ベクトル51_{k-1}とを受信し、補間されたフォアグラウンドV [k] ベクトルを生成するために空間時間的補間を実行するように構成されたユニットを表し得る。空間時間的補間ユニット50は、並べ替えられたフォアグラウンドH O A係数を復元するために、n F G信号49をフォアグラウンドV [k] ベクトル51_kと再び組み合わせ得る。空間時間的補間ユニット50は次いで、補間されたn F G信号49'を生成するために、補間されたV [k] ベクトルによって、並べ替えられたフォアグラウンドH O A係数を分割し得る。空間時間的補間ユニット50はまた、オーディオ復号デバイス24などのオーディオ復号デバイスが補間されたフォアグラウンドV [k] ベクトルを生成しそれによってフォアグラウンドV [k] ベクトル51_kを復元できるように、補間されたフォアグラウンドV [k] ベクトルを生成するために使用されたフォアグラウンドV [k] ベクトル51_kを出力し得る。補間されたフォアグラウンドV [k] ベクトルを生成するために使用されたフォアグラウンドV [k] ベクトル51_kは、残りのフォアグラウンドV [k] ベクトル53として示される。同じV [k] およびV [k - 1] がエンコーダおよびデコーダにおいて（補間されたベクトルV [k] を作成するために）使用されることを保証するために、ベクトルの量子化された／逆量子化されたバージョンがエンコーダおよびデコーダにおいて使用され得る。空間時間的補間ユニット50は、補間されたn F G信号49'を聴覚心理オーディオコーダユニット40に出力し、補間されたフォアグラウンドV [k] ベクトル51_kを係数低減ユニット46に出力し得る。

【0051】

[0061]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たユニットを表し得る。低減されたフォアグラウンドV [k] ベクトル55は、次元D : [(N + 1)² - (N_{BG} + 1)² - B G_{TOT}] × n F Gを有し得る。係数低減ユニット46は、この点において、残りのフォアグラウンドV [k] ベクトル53における係数の数を低減するように構成されたユニットを表し得る。言い換えれば、係数低減ユニット46は、指向性情報をほとんどまたはまったく有しない（残りのフォアグラウンドV [k] ベクトル53を形成する）フォアグラウンドV [k] ベクトルにおける係数を除去するように構成されたユニットを表し得る。いくつかの例では、(N_{BG} と示され得る) 1次および0次の基底関数に対応する、明瞭な、または言い換えればフォアグラウンドV [k] ベクトルの係数は、指向性情報をほとんど提供せず、したがって、（「係数低減」と呼ばれ得るプロセスを通じて）フォアグラウンドVベクトルから除去され得る。この例では、対応する係数N_{BG}を特定するだけではなく、追加のH O Aチャネル（変数T o t a l O f A d d A m b H O A C h a n によって示され得る）を[(N_{BG} + 1)² + 1 , (N + 1)²] のセットから特定するために、より大きい柔軟性が与えられ得る。

【0052】

[0062]量子化ユニット52は、コーディングされたフォアグラウンドV [k] ベクトル57を生成するために低減されたフォアグラウンドV [k] ベクトル55を圧縮するための任意の形態の量子化を実行し、コーディングされたフォアグラウンドV [k] ベクトル57をビットストリーム生成ユニット42に出力するように構成されたユニットを表し得る。動作において、量子化ユニット52は、音場の空間成分、すなわちこの例では低減さ

10

20

30

40

50

れたフォアグラウンド V [k] ベクトル 5 5 の 1 つまたは複数を圧縮するように構成されたユニットを表し得る。量子化ユニット 5 2 は、「 N b i t Q 」で表される量子化モードシンタックス要素によって示されるような、以下の 1 2 の量子化モードのうちのいずれか 1 つを実行し得る。

N b i t Q 値 量子化モードのタイプ

- 0 ~ 3 : 予約済み
- 4 : ベクトル量子化
- 5 : ハフマンコーディングなしのスカラー量子化
- 6 : ハフマンコーディングありの 6 ビットスカラー量子化
- 7 : ハフマンコーディングありの 7 ビットスカラー量子化
- 8 : ハフマンコーディングありの 8 ビットスカラー量子化
- 16 : ハフマンコーディングありの 16 ビットスカラー量子化

量子化ユニット 5 2 は、また、前述のタイプの量子化モードのいずれかの量子化モードの予測されたバージョンを実行し得、ここで、以前のフレームの V ベクトルの要素（またはベクトル量子化が実行されるときの重み）と、現在のフレームの V ベクトルの要素（またはベクトル量子化が実行されるときの重み）との間の差が決定される。量子化ユニット 5 2 は、その際、現在のフレーム自体の V ベクトルの要素の値ではなく、現在のフレームの要素または重みと、以前のフレームの要素または重みとの間の差を量子化し得る。

【 0 0 5 3 】

[0063] 量子化ユニット 5 2 は、低減されたフォアグラウンド V [k] ベクトル 5 5 の複数の符号化されたバージョンを取得するために、低減されたフォアグラウンド V [k] ベクトル 5 5 のそれぞれに対して複数の形の量子化を実行し得る。量子化ユニット 5 2 は、符号化されたフォアグラウンド V [k] ベクトル 5 7 として、低減されたフォアグラウンド V [k] ベクトル 5 5 の符号化されたバージョンのうちの 1 つまたは複数を選択し得る。量子化ユニット 5 2 は、言い換えれば、本開示で説明される基準の任意の組合せに基づいて、出力切替えされ量子化された V ベクトルとして使用するために、予測されないベクトル量子化された V ベクトル、予測されベクトル量子化された V ベクトル、ハフマンコーディングされないスカラー量子化された V ベクトル、およびハフマンコーディングされスカラー量子化された V ベクトルのうちの 1 つを選択し得る。いくつかの例では、量子化ユニット 5 2 は、ベクトル量子化モードと 1 つまたは複数のスカラー量子化モードとを含む、量子化モードのセットから量子化モードを選択し、選択されたモードに基づいて（または従って）、入力 V ベクトルを量子化し得る。量子化ユニット 5 2 は次いで、（たとえば、重み値またはそれを示すビットに関して）予測されないベクトル量子化された V ベクトル、（たとえば、誤差値またはそれを示すビットに関して）予測されベクトル量子化された V ベクトル、ハフマンコーディングされないスカラー量子化された V ベクトル、およびハフマンコーディングされスカラー量子化された V ベクトルのうちの選択されたものを、コーディングされたフォアグラウンド V [k] ベクトル 5 7 としてビットストリーム生成ユニット 4 2 に与え得る。量子化ユニット 5 2 はまた、量子化モードを示すシンタックス要素（たとえば、 N b i t s Q シンタックス要素）と、 V ベクトルを逆量子化またはさもなければ再構成するために使用される任意の他のシンタックス要素とを与える。

【 0 0 5 4 】

[0064] オーディオ符号化デバイス 2 0 に含まれる量子化補償ユニット 7 0 は、補間されたフォアグラウンド V [k] ベクトル 5 3 を量子化することから生じる量子化誤差を補償するために、空間時間的補間ユニット 5 0 から補間された n F G 信号 4 9 ' および補間されたフォアグラウンド V [k] ベクトル 5 3 を、そして、量子化ユニット 5 2 からコーディングされたフォアグラウンド V [k] ベクトル 5 7 を受信し、 n F G 信号 4 9 ' に関する量子化補償を実行するように構成されるユニットを表し得る。量子化補償ユニット 7 0 は、量子化補償された n F G 信号 6 0 を生成し、聴覚心理オーディオコーダユニット 4 0 に出力し得る。

【 0 0 5 5 】

10

20

30

40

50

[0065]量子化補償されたnFG信号60を決定するために、量子化補償ユニット70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は行列であるので、コーディングされたフォアグラウンドV[k]ベクトル57の擬似逆行列を得るために、コーディングされたフォアグラウンドV[k]ベクトル57に関する擬似逆行列関数を実行し得る。擬似逆行列関数は、いくつかの例において、一般化逆行列関数、またはムーア・ペンローズの擬似逆行列関数とし得る。量子化補償ユニット70は、中間的な量子化補償されたnFG信号を決定するために、コーディングされたフォアグラウンドV[k]ベクトル57の擬似逆行列と、フォアグラウンドHOA係数との積を計算し得る。たとえば、量子化補償ユニット70は、補間されたnFG信号49' と補間されたフォアグラウンドV[k]ベクトル57の擬似逆行列と、フォアグラウンドHOA係数との積として、フォアグラウンドHOA係数を決定し得る。中間的な量子化補償されたnFG信号は、補間されたフォアグラウンドV[k]ベクトル53の代わりに、コーディングされたフォアグラウンドV[k]ベクトル57に少なくとも部分的に基づいて計算されるので、コーディングされたフォアグラウンドV[k]ベクトル57の擬似逆行列とフォアグラウンドHOA係数との積の結果として生成されることによって、量子化補償ユニット70によって生成される中間的な量子化補償されたnFG信号は、コーディングされたフォアグラウンドV[k]ベクトル57によって導入される量子化誤差を補償し得る。したがって、オーディオ符号化デバイス20は、それにより、量子化補償されたnFG信号60を量子化する際に、補間されたフォアグラウンドV[k]ベクトル53を量子化する際に導入される任意の誤差を補償し得る。

10

20

30

40

50

【0056】

[0066]量子化補償ユニット70はさらに、量子化補償されたnFG信号60を生成するために、現在のフレームkの中間的な量子化補償されたnFG信号の一部を、以前のフレームk-1の中間的な量子化補償されたnFG信号の一部とクロスフェードさせ得る。たとえば、量子化補償ユニット70は、 1024×2 のサイズの量子化補償されたnFG信号60を生成するために、現在のフレームkの中間的な量子化補償されたnFG信号の最初の256サンプルと、以前のフレームk-1の中間的な量子化補償されたnFG信号の最後の256サンプルとクロスフェードさせ得る。いくつかの例では、量子化補償ユニット70は、現在のフレームkの中間的な量子化補償されたnFG信号を、以前のフレームk-1の中間的な量子化補償されたnFG信号とクロスフェードさせない場合がある。この場合、量子化補償されたnFG信号60は、中間的な量子化補償されたnFG信号と同じであり得る。

【0057】

[0067]オーディオ符号化デバイス20内に含まれる聴覚心理オーディオコーダユニット40は、聴覚心理オーディオコーダの複数のインスタンスを表し得、これらの各々は、符号化された環境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' および量子化補償されたnFG信号60の各々の異なるオーディオオブジェクトまたはHOAチャネルを符号化するために使用される。符号化された環境HOA係数59を生成することは、エネルギー補償された環境HOA係数47' の量子化を実行することを含み得、符号化されたnFG信号61を生成することは、量子化補償されたnFG信号60の量子化を実行することを含み得る。聴覚心理オーディオコーダユニット40は、符号化された環境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に出力し得る。

【0058】

[0068]オーディオ符号化デバイス20内に含まれるビットストリーム生成ユニット42は、既知のフォーマット（復号デバイスによって知られているフォーマットを指し得る）に適合するようにデータをフォーマットし、それによってベクトルベースのビットストリーム21を生成するユニットを表す。ビットストリーム21は、言い換えれば、上記で説明された方法で符号化されている、符号化されたオーディオデータを表し得る。ビットストリーム生成ユニット42は、いくつかの例ではマルチプレクサを表し得、マルチプレクサは、コーディングされたフォアグラウンドV[k]ベクトル57と、符号化された環境

H O A 係数 5 9 と、符号化された n F G 信号 6 1 と、バックグラウンドチャネル情報 4 3 とを受信し得る。ビットストリーム生成ユニット 4 2 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 5 7 と、符号化された環境 H O A 係数 5 9 と、符号化された n F G 信号 6 1 と、バックグラウンドチャネル情報 4 3 とに基づいて、ビットストリーム 2 1 を生成し得る。このようにして、ビットストリーム生成ユニット 4 2 は、それにより、図 7 の例に関して以下により詳細に説明されるように、ビットストリーム 2 1 を取得するために、ビットストリーム 2 1 内のベクトル 5 7 を指定し得る。ビットストリーム 2 1 は、主要またはメインビットストリームと、1 つまたは複数のサイドチャネルビットストリームとを含み得る。

【 0 0 5 9 】

[0069] 図 3 の例には示されないが、オーディオ符号化デバイス 2 0 はまた、現在のフレームが指向性ベース合成を使用して符号化されるべきであるかベクトルベース合成を使用して符号化されるべきであるかに基づいて、オーディオ符号化デバイス 2 0 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 2 1 とベクトルベースのビットストリーム 2 1 との間で）切り替える、ビットストリーム出力ユニットを含み得る。ビットストリーム出力ユニットは、（ H O A 係数 1 1 が合成オーディオオブジェクトから生成されたことを検出した結果として）指向性ベース合成が実行されたか、または（ H O A 係数が録音されたことを検出した結果として）ベクトルベース合成が実行されたかを示す、コンテンツ分析ユニット 2 6 によって出力されるシンタックス要素に基づいて、切替えを実行し得る。ビットストリーム出力ユニットは、ビットストリーム 2 1 の各々とともに現在のフレームのために使用される切替えまたは現在の符号化を示すために、正しいヘッダシンタックスを指定し得る。

【 0 0 6 0 】

[0070] その上、上述されたように、音場分析ユニット 4 4 は、 B G _{TOT} 環境 H O A 係数 4 7 を特定し得、それは、フレームごとに変化し得る（が、時々、 B G _{TOT} は、2 つ以上の（時間的に）隣接するフレームにわたって一定または同じままであり得る）。 B G _{TOT} における変化は、低減されたフォアグラウンド V [k] ベクトル 5 5 において表された係数への変化となり得る。 B G _{TOT} における変化は、フレームごとに変化する（「環境 H O A 係数」と呼ばれることがある）バックグラウンド H O A 係数となり得る（が、この場合も時々、 B G _{TOT} は、2 つ以上の（時間的に）隣接するフレームにわたって一定または同じままであり得る）。この変化は、しばしば、追加の環境 H O A 係数の追加または除去と、対応する、低減されたフォアグラウンド V [k] ベクトル 5 5 からの係数の除去またはそれに対する係数の追加とによって表される、音場の態様のためのエネルギーの変化となる。

【 0 0 6 1 】

[0071] 結果として、音場分析ユニット 4 4 は、いつ環境 H O A 係数がフレームごとに変化するかをさらに決定し、音場の環境成分を表すために使用されることについて、環境 H O A 係数への変化を示すフラグまたは他のシンタックス要素を生成し得る（ここで、この変化はまた、環境 H O A 係数の「遷移」または環境 H O A 係数の「遷移」と呼ばれ得る）。具体的には、係数低減ユニット 4 6 は、（ A m b C o e f f T r a n s i t i o n フラグまたは A m b C o e f f I d x T r a n s i t i o n フラグとして示され得る）フラグを生成し、そのフラグが（場合によってはサイドチャネル情報の一部として）ビットストリーム 2 1 中に含まれ得るように、そのフラグをビットストリーム生成ユニット 4 2 に与え得る。

【 0 0 6 2 】

[0072] 係数低減ユニット 4 6 は、環境係数の遷移のフラグを指定することに加えて、低減されたフォアグラウンド V [k] ベクトル 5 5 が生成される方法を修正し得る。一例では、環境 H O A 環境係数のうちの 1 つが現在のフレームの間に遷移中であると決定すると、係数低減ユニット 4 6 は、遷移中の環境 H O A 係数に対応する低減されたフォアグラウンド V [k] ベクトル 5 5 の V ベクトルの各々について、（「ベクトル要素」または「要

10

20

30

40

50

素」とも呼ばれ得る)ベクトル係数を指定し得る。この場合も、遷移中の環境HOA係数は、 BG_{TOT} からバックグラウンド係数の総数を追加または除去し得る。したがって、バックグラウンド係数の総数において生じた変化は、環境HOA係数がビットストリーム中に含まれるか含まれないか、および、Vベクトルの対応する要素が、上記で説明された第2の構成モードおよび第3の構成モードにおいてビットストリーム中に指定されたVベクトルのために含まれるか否かに影響を及ぼす。係数低減ユニット46が、エネルギーにおける変化を克服するために、低減されたフォアグラウンドV[k]ベクトル55を指定することができる方法に関するより多くの情報は、2015年1月12日に出願された「TRANSITIONING OF AMBIENT HIGHER_ORDER_AMBIENT_COEFFICIENTS」という名称の米国特許出願第14/594,533号において提供されている。10

【0063】

[0073]図4は、図2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4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、指向性ベース再構成ユニット90と、ベクトルベース再構成ユニット92とを含み得る。以下で説明されるが、オーディオ復号デバイス24に関するより多くの情報、およびHOA係数を解凍またはさもなければ復号する様々な態様は、2014年5月29日に出願された「INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD」という名称の国際特許出願公開第WO2014/194099号において入手可能である。20

【0064】

[0074]抽出ユニット72は、ビットストリーム21を受信し、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たユニットを表し得る。抽出ユニット72は、HOA係数11が様々な方向ベースのバージョンを介して符号化されたか、ベクトルベースのバージョンを介して符号化されたかを示す、上述されたシンタックス要素から決定し得る。指向性ベース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、符号化されたバージョンに関連付けられたシンタックス要素(図4の例では指向性ベース情報91として示される)とを抽出し、指向性ベース情報91を指向性ベース再構成ユニット90に渡し得る。指向性ベース再構成ユニット90は、指向性ベース情報91に基づいてHOA係数11'の形態でHOA係数を再構成するように構成されたユニットを表し得る。ビットストリームおよびビットストリーム内のシンタックス要素の構成が、以下で図7A~図7Jの例に関してより詳細に説明される。30

【0065】

[0075]HOA係数11がベクトルベース合成を使用して符号化されたことをシンタックス要素が示すとき、抽出ユニット72は、コーディングされたフォアグラウンドV[k]ベクトル57(コーディングされた重みおよび/もしくはインデックス63またはスカラ-量子化されたVベクトルを含み得る)と、符号化された環境HOA係数59と、対応するオーディオオブジェクト61(符号化されたnFG信号61と呼ばれる場合もある)とを抽出し得る。オーディオオブジェクト61はそれぞれベクトル57のうちの1つに対応する。抽出ユニット72は、コーディングされたフォアグラウンドV[k]ベクトル57をVベクトル再構成ユニット74に渡し、符号化された環境HOA係数59を符号化されたnFG信号61とともに聴覚心理オーディオ復号ユニット80に渡し得る。40

【0066】

[0076]Vベクトル再構成ユニット74(逆量子化ユニットとしても知られる)は、符号化されたフォアグラウンドV[k]ベクトル57から、Vベクトル(たとえば、低減されたフォアグラウンドV[k]ベクトル55_k)を再構成するように構成されるユニットを表し得る。Vベクトル再構成ユニット74は、符号化されたフォアグラウンドV[k]ベクトル57を逆量子化し、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50

るために、量子化ユニット 52 の動作と逆の方法で動作し得る。

【0067】

[0077] いくつかの例では、Vベクトル再構成ユニット74は、クロスフェードされ、量子化されたV[k]ベクトルを生成するために、現在のフレーム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の一部を、以前のフレームのコーディングされたフォアグラウンドV[k-1]ベクトルの一部とクロスフェードさせ得る。たとえば、抽出ユニット72は、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、現在のフレームk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の最初の256サンプルを、以前のフレームk-1の量子化されたフォアグラウンドV[k]ベクトルの最後の256サンプルとクロスフェードさせ、クロスフェードし、量子化されたフォアグラウンドV[k]ベクトルを逆量子化し得る。
10

【0068】

[0078] 聴覚心理オーディオ復号ユニット80は、符号化された環境HOA係数59と符号化されたnFG信号61とを復号し、それによってエネルギー補償された環境HOA係数47'、と補間されたnFG信号49'（補間されたnFGオーディオオブジェクト49'とも呼ばれ得る）とを生成するために、図3の例に示される聴覚心理オーディオコーダユニット40とは逆の方法で動作し得る。聴覚心理オーディオ復号ユニット80は、エネルギー補償された環境HOA係数47'をフェードユニット770に渡し、nFG信号49'をフォアグラウンド編成ユニット78に渡すことができる。
20

【0069】

[0079] 空間時間的補間ユニット76は、空間時間的補間ユニット50に関して上記で説明されたものと同様の方法で動作し得る。空間時間的補間ユニット76は、低減されたフォアグラウンドV[k]ベクトル55_kを受信し、また、補間されたフォアグラウンドV[k]ベクトル55_k''を生成するために、フォアグラウンドV[k]ベクトル55_kおよび低減されたフォアグラウンドV[k-1]ベクトル55_{k-1}に関して空間時間的補間を実行し得る。空間時間的補間ユニット76は、補間されたフォアグラウンドV[k]ベクトル55_k''をフェードユニット770に転送し得る。
20

【0070】

[0080] 抽出ユニット72はまた、いつ環境HOA係数のうちの1つが遷移中であるかを示す信号757を、フェードユニット770に出力し得、フェードユニット770は次いで、SCH_{BG}47'（ここで、SCH_{BG}47'は、「環境HOAチャネル47'」または「環境HOA係数47'」とも呼ばれ得る）および補間されたフォアグラウンドV[k]ベクトル55_k''の要素のうちのいずれがフェードインまたはフェードアウトのいずれかを行われるべきであるかを決定し得る。いくつかの例では、フェードユニット770は、環境HOA係数47'および補間されたフォアグラウンドV[k]ベクトル55_k''の要素の各々に関して、反対に動作し得る。すなわち、フェードユニット770は、環境HOA係数47'のうちの対応する1つに関して、フェードインもしくはフェードアウト、またはフェードインもしくはフェードアウトの両方を実行し得、一方で、補間されたフォアグラウンドV[k]ベクトル55_k''の要素のうちの対応する1つに関して、フェードインもしくはフェードアウト、またはフェードインとフェードアウトの両方を実行し得る。フェードユニット770は、調整された環境HOA係数47'をHOA係数編成ユニット82に出力し、調整されたフォアグラウンドV[k]ベクトル55_k'''をフォアグラウンド編成ユニット78に出力し得る。この点において、フェードユニット770は、HOA係数またはその派生物の様々な態様に関して、たとえば、環境HOA係数47'および補間されたフォアグラウンドV[k]ベクトル55_k'''の要素の形態で、フェード動作を実行するように構成されたユニットを表す。
40

【0071】

[0081] フォアグラウンド編成ユニット78は、フォアグラウンドHOA係数65を生成するために、調整されたフォアグラウンドV[k]ベクトル55_k''''および補間されたnFG信号49'に関して行列乗算を実行するように構成されたユニットを表し得る。
50

この点において、フォアグラウンド編成ユニット78は、フォアグラウンド、または言い換えると、HOA係数11'の支配的態様を再構成するために、オーディオオブジェクト49'（それは、補間されたnFG49'を表す別のある方法である）をベクトル $55_k''$ と組み合わせ得る。フォアグラウンド編成ユニット78は、調整されたフォアグラウンドV[k]ベクトル $55_k''$ による補間されたnFG信号49'の行列乗算を実行し得る。

【0072】

[0082] HOA係数編成ユニット82は、HOA係数11'を取得するために、フォアグラウンドHOA係数65を調整された環境HOA係数47'に組み合わせるように構成されたユニットを表し得る。プライム表記法は、HOA係数11'がHOA係数11と同様であるが同じではないことがあることを反映している。HOA係数11とHOA係数1'との間の差分は、損失のある送信媒体を介した送信、量子化、または他の損失のある演算が原因の損失に起因し得る。

10

【0073】

[0083] 図5Aは、本開示で説明されるベクトルベース合成技法の様々な態様を実行する際の、図3の例に示されるオーディオ符号化デバイス20などのオーディオ符号化デバイスの例示的な動作を示すフローチャートである。最初に、オーディオ符号化デバイス20は、HOA係数11を受信する(106)。オーディオ符号化デバイス20はLITユニット30を呼び出すことができ、LITユニット30は、変換された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を適用し得る(107)。

20

【0074】

[0084] オーディオ符号化デバイス20は次に、上記で説明された方法で様々なパラメータを特定するために、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記で説明された分析を実行するために、パラメータ計算ユニット32を呼び出し得る。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し得る(108)。

30

【0075】

[0085] オーディオ符号化デバイス20は次いで、並べ替えユニット34を呼び出し得、並べ替えユニット34は、上記で説明されたように、並べ替えられた変換された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（この場合も、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）を並べ替え得る(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を呼び出し得る。音場分析ユニット44は、上記で説明されたように、フォアグラウンドチャネルの総数(nFG)45と、バックグラウンド音場の次数(N_{BG})と、送るべき追加のBG HOAチャネルの数(n_{BG}a)およびインデックス(i)（図3の例ではバックグラウンドチャネル情報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し得る(109)。

40

【0076】

[0086] 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を呼び出し得る。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を決定し得る(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を呼び出し得、フォアグラウンド選択ユニット36は、nFG45（フォアグラウンドベクトルを特定する1つまたは複数のインデックスを表し得る）に基づいて、音場のフォアグラウンド成分または明瞭な成分を表す、並べ替えられたUS[k]ベクトル33'と並べ替えられたV

50

[k] ベクトル 3 5 'とを選択し得る(112)。

【0077】

[0087]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を呼び出し得る。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係数のうちの様々なものの除去によるエネルギー損失を補償するために、環境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'を生成し得る。

【0078】

[0088]オーディオ符号化デバイス20はまた、空間時間的補間ユニット50を呼び出し得る。空間時間的補間ユニット50は、補間されたフォアグラウンド信号49'('補間されたnFG信号49')とも呼ばれ得る)と残りのフォアグラウンド指向性情報53('V[k]ベクトル53')とも呼ばれ得る)とを取得するために、並べ替えられた変換されたHOA係数33' / 35'に関して空間時間的補間を実行し得る(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を呼び出し得る。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し得る(118)。

【0079】

[0089]オーディオ符号化デバイス20は次いで、上記で説明された方法で、低減されたフォアグラウンドV[k]ベクトル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を呼び出し得る(120)。

【0080】

[0090]オーディオ符号化デバイス20は、量子化補償ユニット70を呼び出し得る。量子化補償ユニット70は、量子化補償されたnFG信号60を生成するために、コーディングされたフォアグラウンドV[k]ベクトル57の量子化誤差を補償し得る(121)。

【0081】

[0091]オーディオ符号化デバイス20はまた、聴覚心理オーディオコーダユニット40を呼び出し得る。聴覚心理オーディオコーダユニット40は、符号化された環境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'の各ベクトルを聴覚心理コーディングし得る。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を呼び出し得る。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を生成し得る。

【0082】

[0092]図5Bは、本開示で説明されるコーディング技法を実行する際のオーディオ符号化デバイスの例示的な動作を示すフローチャートである。図5Bに示されるように、オーディオ符号化デバイス20のLITユニット30は、HOA係数を、オーディオオブジェクトと、オーディオオブジェクトに関連付けられる指向性情報とに分解し得る(150)。オーディオオブジェクトは、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行列との積を備え得る。オーディオオブジェクトに関連付けられる指向性情報は、複数の球面調和係数の右特異ベクトルを表すV行列を備え得る。

【0083】

[0093]オーディオ符号化デバイス20の聴覚心理オーディオコーダユニット40は、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも

部分的に基づいて、オーディオオブジェクトの閉ループ量子化を実行し得る(152)。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に関連付けられる指向性情報の量子化を実行し、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、オーディオオブジェクトの量子化を実行することによって、オーディオオブジェクトの閉ループ量子化を実行し得る。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に関連付けられる指向性情報の量子化を実行する(すなわち、指向性情報を量子化する)ことから生じる量子化誤差に少なくとも部分的に基づいて、オーディオオブジェクトの量子化を実行する(すなわち、オーディオオブジェクトを量子化する)ことによって、オーディオオブジェクトの量子化を実行(すなわち、オーディオオブジェクトを量子化する)し得る。

10

【0084】

[0094]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に関連付けられる指向性情報の量子化を実行することから生じる量子化誤差を補償することによって、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行することから生じる量子化誤差に少なくとも部分的に基づいて、オーディオオブジェクトの量子化を実行し得る。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に関連付けられる指向性情報の量子化を実行した結果の擬似逆行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定し、量子化補償されたオーディオオブジェクトの量子化を実行することによって、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行することから生じる量子化誤差を補償し得る。

20

【0085】

[0095]オーディオ符号化デバイス20は、高次アンビソニック(HOA)係数と、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果の擬似逆行列との積として量子化補償されたオーディオオブジェクトを決定することによって、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果の擬似逆行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定し得る。

【0086】

[0096]図6Aは、本開示で説明される技法の様々な態様を実行する際の、図4に示されるオーディオ復号デバイス24などのオーディオ復号デバイスの例示的な動作を示すフローチャートである。最初に、オーディオ復号デバイス24は、ビットストリーム21を受信し得る(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を呼び出し得る。説明の目的で、ベクトルベース再構成が実行されるべきであることをビットストリーム21が示すと仮定すると、抽出デバイス72は、上述された情報を取り出すためにビットストリームを解析し、その情報をベクトルベース再構成ユニット92に渡し得る。

30

【0087】

[0097]言い換えれば、抽出ユニット72は、コーディングされたフォアグラウンド指向性情報57(この場合も、コーディングされたフォアグラウンドV[k]ベクトル57とも呼ばれ得る)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61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、上記で説明された方法でビットストリーム21から抽出し得る(132)。

40

【0088】

[0098]オーディオ復号デバイス24はさらに、Vベクトル再構成ユニット74を呼び出し得る。Vベクトル再構成ユニット74は、低減されたフォアグラウンド指向性情報55_kを取得するために、コーディングされたフォアグラウンド指向性情報57をエントロピー復号および逆量子化し得る(136)。オーディオ復号デバイス24はまた、聴覚心理オーディオ復号ユニット80を呼び出し得る。聴覚心理オーディオ復号ユニット80は、エネルギー補償された環境HOA係数47'、と補間されたフォアグラウンド信号49'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

50

号 6 1 とを復号 / 逆量子化し得る (1 3 8)。聴覚心理オーディオ復号ユニット 8 0 は、エネルギー補償された環境 H O A 係数 4 7 ' をフェードユニット 7 7 0 に渡し、n F G 信号 4 9 ' をフォアグラウンド編成ユニット 7 8 に渡し得る。

【 0 0 8 9 】

[0099] オーディオ復号デバイス 2 4 は次に、空間時間的補間ユニット 7 6 を呼び出し得る。空間時間的補間ユニット 7 6 は、並べ替えられたフォアグラウンド指向性情報 5 5_k ' を受信し、また、補間されたフォアグラウンド指向性情報 5 5_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 5 5_k / 5 5_{k-1} に関して空間時間的補間を実行し得る (1 4 0)。空間時間的補間ユニット 7 6 は、補間されたフォアグラウンド V [k] ベクトル 5 5_k ' ' をフェードユニット 7 7 0 に転送し得る。10

【 0 0 9 0 】

[0100] オーディオ復号デバイス 2 4 は、フェードユニット 7 7 0 を呼び出し得る。フェードユニット 7 7 0 は、エネルギー補償された環境 H O A 係数 4 7 ' がいつ遷移中であるかを示すシンタックス要素 (たとえば、A m b C o e f f T r a n s i t i o n シンタックス要素) を (たとえば、抽出ユニット 7 2 から) 受信またはさもなければ取得し得る。フェードユニット 7 7 0 は、遷移シンタックス要素と維持された遷移状態情報とに基づいて、エネルギー補償された環境 H O A 係数 4 7 ' ' を H O A 係数編成ユニット 8 2 に出力し得る。フェードユニット 7 7 0 はまた、シンタックス要素と維持された遷移状態情報とに基づいて、および、補間されたフォアグラウンド V [k] ベクトル 5 5_k ' ' の対応する 1 つまたは複数の要素をフェードアウトまたはフェードインし、フォアグラウンド編成ユニット 7 8 に調整されたフォアグラウンド V [k] ベクトル 5 5_k ' ' を出力し得る (1 4 2)。20

【 0 0 9 1 】

[0101] オーディオ復号デバイス 2 4 は、フォアグラウンド編成ユニット 7 8 を呼び出し得る。フォアグラウンド編成ユニット 7 8 は、フォアグラウンド H O A 係数 6 5 を取得するために、調整されたフォアグラウンド指向性情報 5 5_k ' ' ' による行列乗算 n F G 信号 4 9 ' を実行し得る (1 4 4)。オーディオ復号デバイス 2 4 はまた、H O A 係数編成ユニット 8 2 を呼び出し得る。H O A 係数編成ユニット 8 2 は、H O A 係数 1 1 ' を取得するために、フォアグラウンド H O A 係数 6 5 を調整された環境 H O A 係数 4 7 ' ' に加算し得る (1 4 6)。30

【 0 0 9 2 】

[0102] 図 6 B は、本開示で説明されるコーディング技法を実行する際のオーディオ復号デバイスの例示的な動作を示すフローチャートである。図 6 B に示されるように、オーディオ復号デバイス 2 4 の抽出ユニット 7 2 が、ビットストリームを受信し得る (1 6 0)。オーディオ復号デバイス 2 4 は、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果に少なくとも部分的に基づいて、閉ループ量子化されたオーディオオブジェクトを取得し得る (1 6 2)。たとえば、オーディオ復号デバイス 2 4 の抽出ユニット 7 2 は、閉ループ量子化された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取得するように、ビットストリームを復号し得る。オーディオオブジェクトを取得するのに応答して、オーディオ復号デバイス 2 4 は、オーディオオブジェクトを逆量子化し得る (1 6 4)。40

【 0 0 9 3 】

[0103] いくつかの例では、オーディオオブジェクトは、オーディオオブジェクトに関連付けられる指向性情報を量子化し、オーディオオブジェクトに関連付けられる指向性情報を量子化した結果に少なくとも部分的に基づいてオーディオオブジェクトを量子化することによって、閉ループ量子化される。いくつかの例では、オーディオオブジェクトは、オーディオオブジェクトに関連付けられる指向性情報を量子化し、オーディオオブジェクトに関連付けられる指向性情報を量子化することから生じる量子化誤差に少なくとも部分的に基づいてオーディオオブジェクトを量子化することによって、閉ループ量子化される。

【 0 0 9 4 】

10

20

30

40

50

[0104] いくつかの例では、オーディオオブジェクトは、オーディオオブジェクトに関連付けられる指向性情報を量子化し、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行することから生じる量子化誤差を補償することを含む、オーディオオブジェクトに関連付けられる指向性情報を量子化することから生じる量子化誤差に少なくとも部分的に基づいてオーディオオブジェクトを量子化することによって、閉ループ量子化される。いくつかの例では、オーディオオブジェクトは、オーディオオブジェクトに関連付けられる指向性情報を量子化し、オーディオオブジェクトに関連付けられる指向性情報を量子化した結果の逆擬似行列に少なくとも部分的に基づいて、量子化補償されたオーディオオブジェクトを決定し、量子化補償されたオーディオオブジェクトを量子化することによって、閉ループ量子化される。

10

【0095】

[0105] いくつかの例では、オーディオオブジェクトは、高次アンビソニック（H O A）係数と、オーディオオブジェクトに関連付けられる指向性情報の量子化を実行した結果の擬似逆行列との積として、オーディオオブジェクトを決定することによって、閉ループ量子化される。いくつかの例では、オーディオオブジェクトおよび指向性情報は高次アンビソニック係数から分解され、オーディオオブジェクトは、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行列との積を備え、オーディオオブジェクトに関連付けられる指向性情報は、複数の球面調和係数の右特異ベクトルを表すV行列を備える。

20

【0096】

[0106] 図7Aは、H O A信号圧縮のためのVベクトルおよびフォアグラウンド信号の閉ループ量子化を示すブロック図である。そのような閉ループ量子化は、図3の例に示されるオーディオ符号化デバイス20と、図4の例に示されるオーディオ復号デバイス24によって実行され得る。量子化誤差を低減するために、Vベクトルが量子化され得、そのVベクトルの量子化誤差を補償することによって、USベクトルが量子化され得る。オーディオ符号化デバイス20は、Vベクトルを量子化してQ(V)にし得、Hとpinv(Q(V))との積として新たなターゲット信号T(US)を生成し得、ただし、pinv(A)は、Aの擬似逆行列である。オーディオ符号化デバイス20は、T(US)をQ(T(US))に量子化し得る。オーディオ復号デバイス24は、Q(T(US)).Q(V)'によって生成される量子化されたH O A係数Q(H)に基づいて、H O A係数11'を再構成し得る。このようにして、USベクトルは、Vベクトルの量子化誤差に基づいて量子化され得る。

30

【0097】

[0107] 図7Aに示されるように、H_{FG}は、図7Aの例では、1280×21のサイズを有するフォアグラウンドH O A係数400を表し得る。フォアグラウンドH O A係数400は、U_{FG}と、S_{FG}と、V_{FG}との積に等しくし得、ただし、U_{FG}は1280×2のサイズを有し得、S_{FG}は2×2のサイズを有し得、V_{FG}は21×2のサイズを有し得る。VベクトルV_{FG}402は、図7Aの例では、21×2のサイズを有する、図3の低減されたフォアグラウンドV[k]ベクトル55とし得る。元のターゲットW_{FG}=U_{FG}*S_{FG}404は、図7Aの例では、1280×2のサイズを有する、図3の補間されたnFG信号49'をし得る。

40

【0098】

[0108] オーディオ符号化デバイス20は、後にさらに詳細に論じられるように、以前のフレーム

【数8】

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

408の量子化されたVベクトルに少なくとも部分的に基づいて、VベクトルV_{FG}402に関する符号変更406を実行するか否かを決定し得る。このようにして、オーディオ符

50

号化デバイス 20 は、以前のフレームの遅延した量子化された V ベクトル
【数 9】

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

408 に少なくとも部分的に基づいて、現在のフレームの V ベクトル $V_{FG} 402$ の符号を変更するか否かを決定し得る。オーディオ符号化デバイス 20 は、符号変更されない V ベクトル $V_{FG} 402$ または符号変更された V ベクトル

【数 10】

$$\dot{V}_{FG}$$

10

410 のいずれかが結果として生じるように、V ベクトル $V_{FG} 402$ を符号変更 406 するか否かを決定し得る。オーディオ符号化デバイス 20 は、図 3 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 であり得る、量子化されたフォアグラウンド V ベクトル

【数 11】

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

414 を生成するために、オーディオ符号化デバイス 20 の量子化ユニット 52 を使用することなどによって、符号変更されない V ベクトル $V_{FG} 402$ または符号変更された V ベクトル

20

【数 12】

$$\dot{V}_{FG}$$

410 のいずれかを量子化 412 し得る。オーディオ符号化デバイス 20 が、次のフレームに関する V ベクトル V_{FG} で符号変更を実行するか否かを決定するように、次のフレームにおいて、量子化された V ベクトル

【数 13】

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

30

を利用し得るために、オーディオ符号化デバイスは、遅延された量子化 V ベクトル
【数 14】

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

408 を生成するように 1 フレームだけ量子化されたフォアグラウンド V ベクトル
【数 15】

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

40

414 を遅延させ得る。

【0099】

[0109] オーディオ符号化デバイス 20 は、遅延した量子化された V ベクトル

【数 16】

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

408 に基づいて、

【数17】

$$\bar{W}_{FG,p} = H_{FG} \hat{V}_{FG,p}^{\#}$$

418を決定し、それはフォアグラウンドHOA係数400と、遅延した量子化されたVベクトル

【数18】

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

10

408の擬似逆行列である

【数19】

$$\hat{V}_{FG,p}^{\#}$$

との積である。また、オーディオ符号化デバイス200は、

【数20】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

20

420も決定し得、それは、 $W_{FG} = U_{FG}^* S_{FG} 404$ と、

【数21】

$$\bar{W}_{FG,p} = H_{FG} \hat{V}_{FG,p}^{\#}$$

418との積である。

【数22】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

30

420を決定することに少なくとも部分的に基づいて、オーディオ符号化デバイス20は、Vベクトル $V_{FG} 402$ に関する符号変更406を実行し得る。たとえば、オーディオ符号化デバイス20は、

【数23】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

420の符号が負である場合には、Vベクトル $V_{FG} 402$ に関する符号変更406を実行し得る。

【0100】

40

[0110]オーディオ符号化デバイス20は、量子化されたVベクトル

【数24】

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

408に少なくとも部分的に基づいて、

【数25】

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

の積である新たなターゲット422

50

【数26】

$$\tilde{W}_{FG}$$

を生成し得、それは、量子化されたVベクトル

【数27】

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

408の擬似逆行列と、フォアグラウンドHOA係数400を表し得る H_{FG} との積であり、それにより、新たなターゲット422

【数28】

10

$$\tilde{W}_{FG} = H_{FG} \hat{V}_{FG}^*$$

になる。オーディオ符号化デバイス20は、先行するターゲット426

【数29】

$$\tilde{W}_{FG,p}$$

を生成するように、1フレームだけ新たなターゲット422

【数30】

20

$$\tilde{W}_{FG} = H_{FG} \hat{V}_{FG}^*$$

を遅延させ、図3の量子化補償されたnFG信号60に類似の、量子化補償されたターゲット信号を生成するように、新たなターゲット422

【数31】

$$\tilde{W}_{FG}$$

の、最初の256サンプルのような第1の部分を、先行するターゲット426

【数32】

30

$$\tilde{W}_{FG,p}$$

の、最後の256サンプルのような最後の部分と、クロスフェード428させ得る。オーディオ符号化デバイス20は、オーディオ復号デバイス24に出力される量子化されたフォアグラウンド信号432

【数33】

$$\hat{W}_{FG}$$

を生成するように、図3の聴覚心理オーディオコーダユニット40を用いることなどによって、量子化補償されたターゲット信号を量子化430し得る。

40

【0101】

[0111]オーディオ復号デバイス24は、たとえば、オーディオ符号化デバイス20から、量子化されたフォアグラウンドVベクトル

【数34】

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

414を受信し得、遅延した量子化されたVベクトル

【数 3 5】

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

4 0 8 を生成するように、1 フレームだけ量子化されたフォアグラウンド V ベクトル

【数 3 6】

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

4 1 4 を遅延 4 1 6 させ得る。オーディオ復号デバイス 2 4 は、現在のフレームの量子化 10
されたフォアグラウンド V ベクトル

【数 3 7】

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

4 1 4 の第 1 の数のサンプルを、遅延した量子化された V ベクトル

【数 3 8】

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

4 0 8 の最後の数のサンプルとクロスフェードさせることができ、量子化されたフォアグ 20
ラウンド H O A 係数 4 3 8

【数 3 9】

$$\hat{H}_{FG} = \hat{W}_{FG} \hat{V}_{FG,CF}^T$$

を生成するように、クロスフェードした量子化されたフォアグラウンド V [k] ベクトル

4 3 6

【数 4 0】

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

30

と、量子化されたフォアグラウンド信号 4 3 2

【数 4 1】

$$\hat{W}_{FG}$$

との積を決定し得る。たとえば、オーディオ復号デバイス 2 4 は、現在のフレームの量子
化されたフォアグラウンド V ベクトル

【数 4 2】

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

40

4 1 4 の最初の 2 5 6 サンプルを、遅延した量子化された V ベクトル

【数 4 3】

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

4 0 8 の最後の 2 5 6 サンプルとクロスフェードさせることができ、量子化されたフォア
グラウンド H O A 係数 4 3 8

【数44】

$$\hat{H}_{FG} = \hat{W}_{FG} \hat{V}_{FG,CF}^T$$

を生成するように、クロスフェードした量子化されたフォアグラウンド $V[k]$ ベクトル
4 3 6

【数45】

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

10

と、量子化されたフォアグラウンド信号 4 3 2

【数46】

$$\hat{W}_{FG}$$

との積を決定し得る。スピーカーが、量子化されたフォアグラウンド H O A 係数 4 3 8 から分解された逆量子化されたオーディオオブジェクトからレンダリングされたラウドスピーカーフィードを再生し得るように、オーディオ復号デバイス 2 4 は、フォアグラウンド H O A 係数 4 3 8 を分解および逆量子化し得る。

【0102】

[0112] 図 7 B は、H O A 信号圧縮のための V ベクトルおよびフォアグラウンド信号の閉ループ量子化を示すブロック図である。図 7 B に示されるようなオーディオ復号デバイス 2 4 は、現在のフレームの量子化されたフォアグラウンド V ベクトル

20

【数47】

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

4 1 4 の第 1 の数のサンプルを、遅延した量子化された V ベクトル

【数48】

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

30

4 0 8 の最後の数のサンプルとクロスフェードさせないという点で、図 7 B は図 7 A とは異なる。また、図 7 B に示されるようなオーディオ符号化デバイス 2 0 は、先行するターゲットを生成するために新たなターゲット 4 2 2

【数49】

$$\tilde{W}_{FG} = H_{FG} \hat{V}_{FG}^*$$

を 1 フレームだけ遅延させず、それゆえ、量子化補償されたターゲット信号を生成するために、新たなターゲット 4 2 2

40

【数50】

$$\tilde{W}_{FG}$$

の第 1 の部分を、先行するターゲット 4 の最後の部分とクロスフェードさせないという点で、図 7 B は図 7 A とは異なる。

【0103】

[0113] フォアグラウンド H O A 係数 4 0 0 は、 U_{FG} と、 S_{FG} と、 V_{FG} との積に等しくあり得、ここで、 U_{FG} は 1280×2 のサイズを有し得、 S_{FG} は 2×2 のサイズを有し得、 V_{FG} は 21×2 のサイズを有し得る。 V ベクトル V_{FG} 4 0 2 は、図 7 A の例では、 21×2 のサイズを有する、図 3 の低減されたフォアグラウンド $V[k]$ ベクトル 5 5 であり得

50

る。元のターゲット $W_{FG} = U_{FG}^* S_{FG} 4\ 0\ 4$ は、図 7 A の例では、 $1\ 2\ 8\ 0 \times 2$ のサイズを有する、図 3 の補間された n_{FG} 信号 $4\ 9'$ であり得る。

【数 5 1】

[0114] オーディオ符号化デバイス 20 は、後にさらに詳細に論じられるように、以前のフレーム

【数 5 1】

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

4 0 8 の量子化された V ベクトルに少なくとも部分的に基づいて、V ベクトル $V_{FG} 4\ 0\ 2$ に関する符号変更 $4\ 0\ 6$ を実行するか否かを決定し得る。このようにして、オーディオ符号化デバイス 20 は、以前のフレームの遅延した量子化された V ベクトル

【数 5 2】

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

4 0 8 に少なくとも部分的に基づいて、現在のフレームの V ベクトル $V_{FG} 4\ 0\ 2$ の符号を変更するか否かを決定し得る。オーディオ符号化デバイス 20 は、符号変更されない V ベクトル $V_{FG} 4\ 0\ 2$ または符号変更された V ベクトル

【数 5 3】

20

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

4 1 0 のいずれかが結果として生じるように、V ベクトル $V_{FG} 4\ 0\ 2$ を符号変更 $4\ 0\ 6$ するか否かを決定し得る。オーディオ符号化デバイス 20 は、図 3 のコーディングされたフォアグラウンド V [k] ベクトル 5 7 であり得る、量子化されたフォアグラウンド V ベクトル

【数 5 4】

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

30

4 1 4 を生成するように、オーディオ符号化デバイス 20 の量子化ユニット 5 2 を使用することなどによって、符号変更されない V ベクトル $V_{FG} 4\ 0\ 2$ または符号変更された V ベクトル

【数 5 5】

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

4 1 0 のいずれかを量子化 4 1 2 し得る。オーディオ符号化デバイス 20 は、次のフレームに関する V ベクトル V_{FG} に関する符号変更を実行するか否かを決定するために、次のフレームにおいて、量子化された V ベクトル

【数 5 6】

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

40

を用い得るように、オーディオ符号化デバイスは、遅延した量子化された V ベクトル

【数 5 7】

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

4 0 8 を生成するために、1 フレームだけ量子化されたフォアグラウンド V ベクトル

【数58】

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

414を遅延させ得る。

【0105】

[0115]オーディオ符号化デバイス20は、遅延した量子化されたVベクトル

【数59】

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

10

408に基づいて、

【数60】

$$\bar{W}_{FG,p} = H_{FG} \hat{V}_{FG,p}^{\#}$$

418を決定し、それはフォアグラウンドHOA係数400と、遅延した量子化されたVベクトル

【数61】

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

20

408の擬似逆行列である、

【数62】

$$\hat{V}_{FG,p}^{\#}$$

との積である。また、オーディオ符号化デバイス200は、

【数63】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

30

420も決定し得、それは、 $W_{FG} = U_{FG}^* S_{FG} 404$ と、

【数64】

$$\bar{W}_{FG,p} = H_{FG} \hat{V}_{FG,p}^{\#}$$

418との積である。

【数65】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

40

420を決定することに少なくとも部分的に基づいて、オーディオ符号化デバイス20は、Vベクトル $V_{FG} 402$ で符号変更406を実行し得る。たとえば、オーディオ符号化デバイス20は、

【数66】

$$\Gamma_{FG} = W_{FG} \bar{W}_{FG,p}^T$$

420の符号が負である場合には、Vベクトル $V_{FG} 402$ に関する符号変更406を実行し得る。

50

【 0 1 0 6 】

[0116] オーディオ符号化デバイス 20 は、量子化された V ベクトル

【 数 6 7 】

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

4 0 8 に少なくとも部分的に基づいて、新たなターゲット 4 2 2

【 数 6 8 】

$$\tilde{W}_{FG} = H_{FG} \hat{V}_{FG}^{\#}$$

10

のような、量子化された V ベクトル

【 数 6 9 】

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

4 0 8 の擬似逆行列である、

【 数 7 0 】

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

と、フォアグラウンド H O A 係数 4 0 0 を表し得る、H_{FG}との積である新たなターゲット

20

4 2 2

【 数 7 1 】

$$\tilde{W}_{FG}$$

を生成し得る。オーディオ符号化デバイス 20 は、量子化されたフォアグラウンド信号 4

3 2

【 数 7 2 】

$$\hat{W}_{FG}$$

30

を生成するように、新たなターゲット 4 2 2

【 数 7 3 】

$$\tilde{W}_{FG} = H_{FG} \hat{V}_{FG}^{\#}$$

を量子化 4 3 0 し得る。オーディオ復号デバイス 24 は、たとえば、オーディオ符号化デバイス 20 から、量子化されたフォアグラウンド V ベクトル

【 数 7 4 】

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

4 1 4 を受信し得、量子化されたフォアグラウンド H O A 係数 4 3 8

40

【 数 7 5 】

$$\hat{H}_{FG} = \hat{W}_{FG} \hat{V}_{FG}^T$$

を生成するように、量子化されたフォアグラウンド V ベクトル

【 数 7 6 】

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

4 1 4 と、量子化されたフォアグラウンド信号 4 3 2

【数 7 7】

$$\hat{W}_{FG}$$

との積を決定し得る。スピーカーが、逆量子化されたフォアグラウンドHOA係数438から分解された逆量子化されたオーディオオブジェクトからレンダリングされたラウドスピーカーフィードを再生できるようにするように、オーディオ復号デバイス24は、フォアグラウンドHOA係数438を分解および逆量子化することができる。

【0107】

[0117]図8は、図3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8の例では、量子化ユニット52は、一様量子化ユニット230と、n bitsユニット232と、予測ユニット234と、予測モードユニット236（「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、遅延ユニット300と、符号変更ユニット302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。n bitsユニット232は、n bitsパラメータまたはn bits値を決定するように構成されるユニットを表す。

10

【0108】

[0118]遅延ユニット300は、一様量子化ユニット230の結果を1フレームだけ遅延させ得、それにより、一様量子化ユニット230が低減されたフォアグラウンドV[k]ベクトル55に作用する前に、符号変更ユニット302が、量子化されたフォアグラウンドV[k]ベクトルに少なくとも部分的に基づいて、低減されたフォアグラウンドV[k]ベクトル55に関する符号変更を実行するか否かを決定し得るようにする。符号変更ユニット302は、言い換えると、低減されたフォアグラウンドV[k]ベクトル55のうちの1つまたは複数のための符号を（正から負に、または負から正に）反転させるように構成されるユニットを表し得る。線形可逆変換の性質を考えると、V[k]ベクトル55は、以前のフレーム（または第(k-1)のフレーム）のV[k-1]ベクトル55のうちの対応する1つまたは複数が符号に関して反転するように、第kのフレームのためのHOA係数11から分解され得る。この点において、フレーム境界を越えて符号を変更する必要があり得る。したがって、現在のフレームの低減されたフォアグラウンドV[k]ベクトル55の符号を変更することが必要であるか否かは、以前のフレームの量子化されたVベクトルに依存し得る。具体的には、量子化ユニット54は、遅延したフォアグラウンド信号を生成するように、第kのフレームのためのHOA係数11と、以前のフレーム(k-1フレーム)の量子化されたVベクトルとを乗算し得る。量子化ユニット54は、遅延したフォアグラウンド信号を、補間されたnFG信号49'を乗算し得る。遅延したフォアグラウンド信号を、補間されたnFG信号49'を乗算した結果が負である場合には、符号変更ユニット302が、低減されたフォアグラウンドV[k]ベクトル55に関する符号変更を実行し得る。

20

30

【0109】

[0119]図9は、図3の例に示されるオーディオ符号化デバイス20の量子化補償ユニット70をより詳細に示すブロック図である。量子化補償ユニット70は、中間補償ユニット304と、遅延ユニット306と、クロスフェードユニット308とを含み得る。中間補償ユニット304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の擬似逆行列を取得するように、コーディングされたフォアグラウンドV[k]ベクトル57に関する擬似逆行列関数を実行し得る。中間補償ユニット304はさらに、中間的な量子化補償されたnFG信号を決定するように、コーディングされたフォアグラウンドV[k]ベクトル57の擬似逆行列と、フォアグラウンドHOA係数との積を計算し得る。一例では、中間補償ユニット304は、補間されたnFG信号49'を、補間されたフォアグ

40

50

ラウンド $V[k]$ ベクトル $5[3]$ との積として、フォアグラウンド H.O.A 係数を決定し得る。

【0110】

[0120] 遅延ユニット $3[0][6]$ は、中間補償ユニット $3[0][4]$ によって生成された中間的な量子化補償された nFG 信号を、1 フレームだけ遅延させ得る。クロスフェードユニット $3[0][8]$ は、量子化補償された nFG 信号 $6[0]$ を生成するように、中間補償ユニット $3[0][4]$ によって出力された現在のフレーム k の中間的な量子化補償された nFG 信号の一部を、遅延ユニット $3[0][6]$ によって出力された以前のフレーム $k - 1$ の中間的な量子化補償された nFG 信号の一部とクロスフェードさせ得る。たとえば、クロスフェードユニット $3[0][8]$ は、 1024×2 のサイズの量子化補償された nFG 信号 $6[0]$ を生成するように、現在のフレーム k の中間的な量子化補償された nFG 信号の最初の 256 サンプルと、以前のフレーム $k - 1$ の中間的な量子化補償された nFG 信号の最後の 256 サンプルとクロスフェードさせ得る。
10

【0111】

[0121] 上記の技法は、任意の数の異なる状況およびオーディオエコシステムに関して実行され得る。いくつかの例示的な状況が以下で説明されるが、本技法はそれらの例示的な状況に限定されるべきではない。1つの例示的なオーディオエコシステムは、オーディオコンテンツと、映画スタジオと、音楽スタジオと、ゲーミングオーディオスタジオと、チャネルベースオーディオコンテンツと、コーディングエンジンと、ゲームオーディオスистемと、ゲームオーディオコーディング／レンダリングエンジンと、配信システムとを含み得る。
20

【0112】

[0122] 映画スタジオ、音楽スタジオ、およびゲーミングオーディオスタジオは、オーディオコンテンツを受信することができる。いくつかの例では、オーディオコンテンツは、獲得物の出力を表し得る。映画スタジオは、デジタルオーディオワークステーション (DAW) を使用することなどによって、(たとえば、2.0、5.1、および 7.1 の) チャネルベースオーディオコンテンツを出力することができる。音楽スタジオは、DAW を使用することなどによって、(たとえば、2.0、および 5.1 の) チャネルベースオーディオコンテンツを出力し得る。いずれの場合も、コーディングエンジンは、配信システムによる出力のために、チャネルベースオーディオコンテンツベースの 1 つまたは複数のコーデック (たとえば、AAC、AC3、Dolby True HD、Dolby Digital Plus、および DTS Master Audio) を受信し符号化し得る。ゲーミングオーディオスタジオは、DAW を使用することなどによって、1 つまたは複数のゲームオーディオスистем を出力し得る。ゲームオーディオコーディング／レンダリングエンジンは、配信システムによる出力のために、オーディオスystem をチャネルベースオーディオコンテンツへとコーディングおよびまたはレンダリングし得る。本技法が実行され得る別の例示的な状況は、放送録音オーディオオブジェクトと、プロフェッショナルオーディオシステムと、消費者向けオンデバイスキャプチャと、H.O.A オーディオフォーマットと、オンデバイスレンダリングと、消費者向けオーディオと、TV、およびアクセサリと、カーオーディオシステムとを含み得る、オーディオエコシステムを備える。
30

【0113】

[0123] 放送録音オーディオオブジェクト、プロフェッショナルオーディオシステム、および消費者向けオンデバイスキャプチャはすべて、H.O.A オーディオフォーマットを使用して、それらの出力をコーディングし得る。このようにして、オーディオコンテンツは、オンデバイスレンダリング、消費者向けオーディオ、TV、およびアクセサリ、ならびにカーオーディオシステムを使用して再生され得る单一の表現へと、H.O.A オーディオフォーマットを使用してコーディングされ得る。言い換れば、オーディオコンテンツの单一の表現は、オーディオ再生システム 16 など、汎用的なオーディオ再生システムにおいて (すなわち、5.1、7.1 などの特定の構成を必要とすることとは対照的に) 再生され得る。
40
50

【0114】

[0124]本技法が実行され得る状況の他の例には、獲得要素と再生要素とを含み得るオーディオエコシステムがある。獲得要素は、有線および／またはワイヤレス獲得デバイス（たとえば、Eig e nマイクロフォン）、オンデバイスサラウンドサウンドキャプチャ、ならびにモバイルデバイス（たとえば、スマートフォンおよびタブレット）を含み得る。いくつかの例では、有線および／またはワイヤレス獲得デバイスは、有線および／またはワイヤレス通信チャネルを介してモバイルデバイスに結合され得る。

【0115】

[0125]本開示の1つまたは複数の技法によれば、モバイルデバイスが音場を獲得するために使用され得る。たとえば、モバイルデバイスは、有線および／もしくはワイヤレス獲得デバイス、ならびに／またはオンデバイスサラウンドサウンドキャプチャ（たとえば、モバイルデバイスに統合された複数のマイクロフォン）を介して、音場を獲得し得る。モバイルデバイスは次いで、再生要素のうちの1つまたは複数による再生のために、獲得された音場をHOA係数へとコーディングし得る。たとえば、モバイルデバイスのユーザは、ライブイベント（たとえば、会合、会議、劇、コンサートなど）を録音し（その音場を獲得し）、録音をHOA係数へとコーディングし得る。

10

【0116】

[0126]モバイルデバイスはまた、HOAコーディングされた音場を再生するために、再生要素のうちの1つまたは複数を利用し得る。たとえば、モバイルデバイスは、HOAコーディングされた音場を復号し、再生要素のうちの1つまたは複数に信号を出力することができ、このことは再生要素のうちの1つまたは複数に音場を再作成させる。一例として、モバイルデバイスは、1つまたは複数のスピーカー（たとえば、スピーカーアレイ、サウンドバーなど）に信号を出力するために、ワイヤレスおよび／またはワイヤレス通信チャネルを利用し得る。別の例として、モバイルデバイスは、1つもしくは複数のドッキングステーションおよび／または1つもしくは複数のドッキングされたスピーカー（たとえば、スマート自動車および／またはスマート住宅の中のサウンドシステム）に信号を出力するために、ドッキング解決手段を利用し得る。別の例として、モバイルデバイスは、ヘッドフォンのセットに信号を出力するために、たとえばリアルなバイノーラルサウンドを作成するために、ヘッドフォンレンダリングを利用し得る。

20

【0117】

[0127]いくつかの例では、特定のモバイルデバイスは、3D音場を獲得することと、より後の時間に同じ3D音場を再生することの両方を行い得る。いくつかの例では、モバイルデバイスは、3D音場を獲得し、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）に送信し得る。

30

【0118】

[0128]本技法が実行され得るまた別の状況は、オーディオコンテンツと、ゲームスタジオと、コーディングされたオーディオコンテンツと、レンダリングエンジンと、配信システムとを含み得る、オーディオエコシステムを含む。いくつかの例では、ゲームスタジオは、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、ゲームスタジオは、HOAをサポートする新しいシステムフォーマットを出力し得る。いずれの場合も、ゲームスタジオは、配信システムによる再生のために音場をレンダリングし得るレンダリングエンジンに、コーディングされたオーディオコンテンツを出力し得る。

40

【0119】

[0129]本技法はまた、例示的なオーディオ獲得デバイスに関して実行され得る。たとえば、本技法は、3D音場を録音するようにまとめて構成される複数のマイクロフォンを含み得る、Eig e nマイクロフォンに関して実行され得る。いくつかの例では、Eig e

50

nマイクロフォン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、マイクロフォンから直接ビットストリーム21を出力するために、Eigenマイクロフォンに統合され得る。

【0120】

[0130]別の例示的なオーディオ獲得状況は、1つまたは複数のEigenマイクロフォンなど、1つまたは複数のマイクロフォンから信号を受信するように構成され得る、製作トラックを含み得る。製作トラックはまた、図3のオーディオ符号化デバイス20などのオーディオエンコーダを含み得る。

【0121】

[0131]モバイルデバイスはまた、いくつかの場合には、3D音場を録音するようにまとめて構成される複数のマイクロフォンを含み得る。言い換えれば、複数のマイクロフォンは、X、Y、Zのダイバーシティを有し得る。いくつかの例では、モバイルデバイスは、モバイルデバイスの1つまたは複数の他のマイクロフォンに関してX、Y、Zのダイバーシティを提供するように回転され得るマイクロフォンを含み得る。モバイルデバイスはまた、図3のオーディオ符号化デバイス20などのオーディオエンコーダを含み得る。

【0122】

[0132]耐衝撃性のビデオキャプチャデバイスは、3D音場を録音するようにさらに構成され得る。いくつかの例では、耐衝撃性のビデオキャプチャデバイスは、ある活動に関するユーザのヘルメットに取り付けられ得る。たとえば、耐衝撃性のビデオキャプチャデバイスは、急流下りをしているユーザのヘルメットに取り付けられ得る。このようにして、耐衝撃性のビデオキャプチャデバイスは、ユーザの周りのすべての活動（たとえば、ユーザの後ろでくだける水、ユーザの前で話している別の乗員など）を表す3D音場をキャプチャし得る。

【0123】

[0133]本技法はまた、アクセサリで増強されたモバイルデバイスに関して実行され得、それは、3D音場を録音するように構成され得る。いくつかの例では、モバイルデバイスは、上記で説明されたモバイルデバイスと同様であり得るが、1つまたは複数のアクセサリが追加されている。たとえば、Eigenマイクロフォンが、アクセサリで増強されたモバイルデバイスを形成するために、上述されたモバイルデバイスに取り付けられ得る。このようにして、アクセサリで増強されたモバイルデバイスは、アクセサリで増強されたモバイルデバイスと一緒にサウンドキャプチャ構成要素をただ使用するよりも高品質なバージョンの3D音場をキャプチャすることができる。

【0124】

[0134]本開示で説明される本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が、以下でさらに説明される。本開示の1つまたは複数の技法によれば、スピーカーおよび/またはサウンドバーは、あらゆる任意の構成で配置され得るが、一方で、依然として3D音場を再生する。その上、いくつかの例では、ヘッドフォン再生デバイスが、有線接続またはワイヤレス接続のいずれかを介してオーディオ復号デバイス24に結合され得る。本開示の1つまたは複数の技法によれば、音場の単一の汎用的な表現が、スピーカー、サウンドバー、およびヘッドフォン再生デバイスの任意の組合せで音場をレンダリングするために利用され得る。

【0125】

[0135]いくつかの異なる例示的なオーディオ再生環境はまた、本開示で説明される技法の様々な態様を実行するために好適であり得る。たとえば、5.1スピーカー再生環境、2.0（たとえば、ステレオ）スピーカー再生環境、フルハイトフロントラウドスピーカーを伴う9.1スピーカー再生環境、22.2スピーカー再生環境、16.0スピーカー再生環境、自動車スピーカー再生環境、およびイヤパッド再生環境を伴うモバイルデバイスは、本開示で説明される技法の様々な態様を実行するために好適な環境であり得る。

【0126】

10

20

30

40

50

[0136]本開示の1つまたは複数の技法によれば、音場の单一の汎用的な表現が、上記の再生環境のいずれかにおいて音場をレンダリングするために利用され得る。加えて、本開示の技法は、レンダードが、上記で説明されたもの以外の再生環境での再生のために、汎用的な表現から音場をレンダリングすることを可能にする。たとえば、設計上の考慮事項が、7.1スピーカー再生環境に従ったスピーカーの適切な配置を妨げる場合（たとえば、右側のサラウンドスピーカーを配置することが可能ではない場合）、本開示の技法は、再生が6.1スピーカー再生環境で達成され得るように、レンダーが他の6つのスピーカーとともに補償することを可能にする。

【0127】

[0137]その上、ユーザは、ヘッドフォンを装着しながらスポーツの試合を見得る。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音場が獲得され得（たとえば、1つまたは複数のEigelnマイクロフォンが野球場の中および／または周りに配置され得）、3D音場に対応するHOA係数が取得されデコーダに送信され得、デコーダがHOA係数に基づいて3D音場を再構成して、再構成された3D音場をレンダラに出力することができ、レンダラが再生環境のタイプ（たとえば、ヘッドフォン）についての指示を取得し、再構成された3D音場を、ヘッドフォンにスポーツの試合の3D音場の表現を出力させる信号へとレンダリングし得る。

10

【0128】

[0138]上記で説明された様々な場合の各々において、オーディオ符号化デバイス20は、ある方法を実行し、またはさもなければ、オーディオ符号化デバイス20が実行するように構成される方法の各ステップを実行するための手段を備え得ることを理解されたい。いくつかの場合には、これらの手段は1つまたは複数のプロセッサを備え得る。いくつかの場合には、1つまたは複数のプロセッサは、非一時的コンピュータ可読記憶媒体に記憶される命令によって構成される、専用のプロセッサを表し得る。言い換えると、符号化の例のセットの各々における本技法の様々な態様は、実行されると、1つまたは複数のプロセッサに、オーディオ符号化デバイス20が実行するように構成されている方法を実行させる命令を記憶した、非一時的コンピュータ可読記憶媒体を提供し得る。

20

【0129】

[0139]1つまたは複数の例において、前述の機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せで実装され得る。ソフトウェアで実装される場合、機能は、コンピュータ可読媒体上の1つまたは複数の命令またはコード上に記憶され、またはこれを介して送信され、ハードウェアベースの処理ユニットによって実行され得る。コンピュータ可読媒体は、データ記憶媒体などの有形媒体に対応するコンピュータ可読記憶媒体を含み得る。データ記憶媒体は、本開示で説明される技法の実装のために命令、コードおよび／またはデータ構造を取り出すために、1つまたは複数のコンピュータあるいは1つまたは複数のプロセッサによってアクセスされ得る任意の利用可能な媒体であり得る。コンピュータプログラム製品は、コンピュータ可読媒体を含み得る。

30

【0130】

[0140]同様に、上記で説明された様々な場合の各々において、オーディオ復号デバイス24は、ある方法を実行し、またはさもなければ、オーディオ復号デバイス24が実行するように構成される方法の各ステップを実行するための手段を備え得ることを理解されたい。いくつかの場合には、これらの手段は1つまたは複数のプロセッサを備え得る。いくつかの場合には、1つまたは複数のプロセッサは、非一時的コンピュータ可読記憶媒体に記憶される命令によって構成される、専用のプロセッサを表し得る。言い換えると、符号化の例のセットの各々における本技法の様々な態様は、実行されると、1つまたは複数のプロセッサに、オーディオ復号デバイス24が実行するように構成されている方法を実行させる命令を記憶した、非一時的コンピュータ可読記憶媒体を提供し得る。

40

【0131】

[0141]限定ではなく例として、そのようなコンピュータ可読記憶媒体は、RAM、ROM、EEPROM（登録商標）、CD-ROMもしくは他の光ディスクストレージ、磁気

50

ディスクストレージ、もしくは他の磁気記憶デバイス、フラッシュメモリ、または命令もしくはデータ構造の形態の所望のプログラムコードを記憶するために使用され得、コンピュータによってアクセスされ得る任意の他の媒体を備えることができる。しかしながら、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時的媒体を含むのではなく、非一時的な有形の記憶媒体を対象とすることを理解されたい。本明細書で使用するディスク (disk) およびディスク (disc) は、コンパクトディスク (disc) (C D)、レーザーディスク (登録商標) (disc)、光ディスク (disc)、デジタル多用途ディスク (disc) (D V D)、フロッピー (登録商標) ディスク (disk) および Blu-ray (登録商標) ディスク (disc) を含み、ここで、ディスク (disk) は、通常、データを磁気的に再生し、一方、ディスク (disc) は、データをレーザーで光学的に再生する。上記の組合せも、コンピュータ可読媒体の範囲の中に含まれるべきである。10

【 0 1 3 2 】

[0142] 命令は、1つもしくは複数のデジタル信号プロセッサ (D S P)、汎用マイクロプロセッサ、特定用途向け集積回路 (A S I C)、フィールドプログラマブルゲートアレイ (F P G A)、あるいは他の同等の集積回路またはディスクリート論理回路などの1つもしくは複数のプロセッサによって実行され得る。したがって、本明細書で使用される「プロセッサ」という用語は、前述の構造、または、本明細書で説明された技法の実装に好適な任意の他の構造のいずれかを指し得る。加えて、いくつかの態様では、本明細書で説明された機能は、符号化および復号のために構成されるか、または複合コーデックに組み込まれる、専用のハードウェアモジュールおよび / またはソフトウェアモジュール内で提供され得る。また、本技法は、1つもしくは複数の回路または論理要素で十分に実装され得る。20

【 0 1 3 3 】

[0143] 本開示の技法は、ワイヤレスハンドセット、集積回路 (I C) もしくは I C のセット (たとえば、チップセット) を含む、多種多様なデバイスまたは装置で実装され得る。本開示では、開示される技法を実行するように構成されたデバイスの機能的態様を強調するために様々な構成要素、モジュール、またはユニットが説明されるが、それらの構成要素、モジュール、またはユニットを、必ずしも異なるハードウェアユニットによって実現する必要があるとは限らない。むしろ、上で説明されたように、様々なユニットが、好適なソフトウェアおよび / またはファームウェアとともに、上記の1つまたは複数のプロセッサを含めて、コーデックハードウェアユニットにおいて組み合わせられるか、または相互動作ハードウェアユニットの集合によって与えられ得る。30

【 0 1 3 4 】

[0144] 本開示の様々な態様が説明された。本技法のこれらおよび他の態様は、以下の特許請求の範囲内に入る。

【図 1】

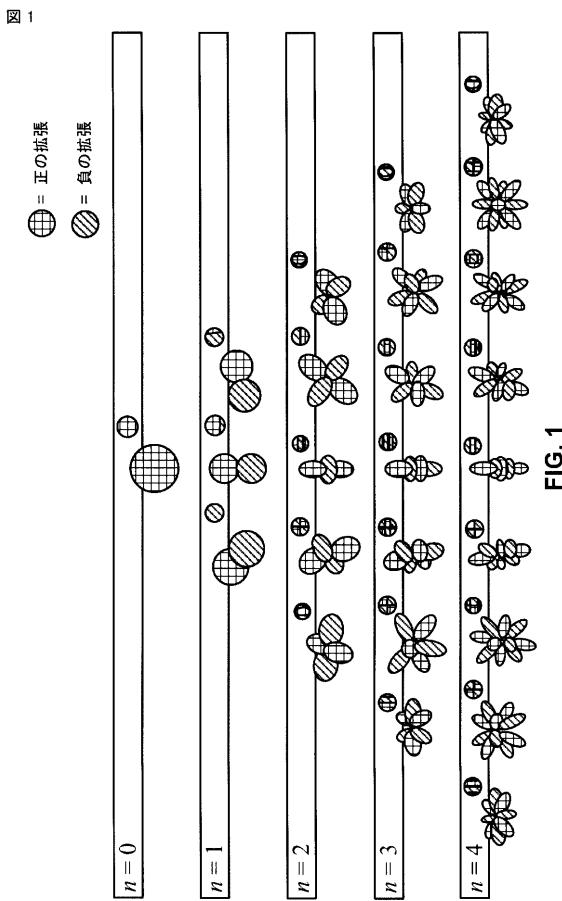


FIG. 1

【図 2】

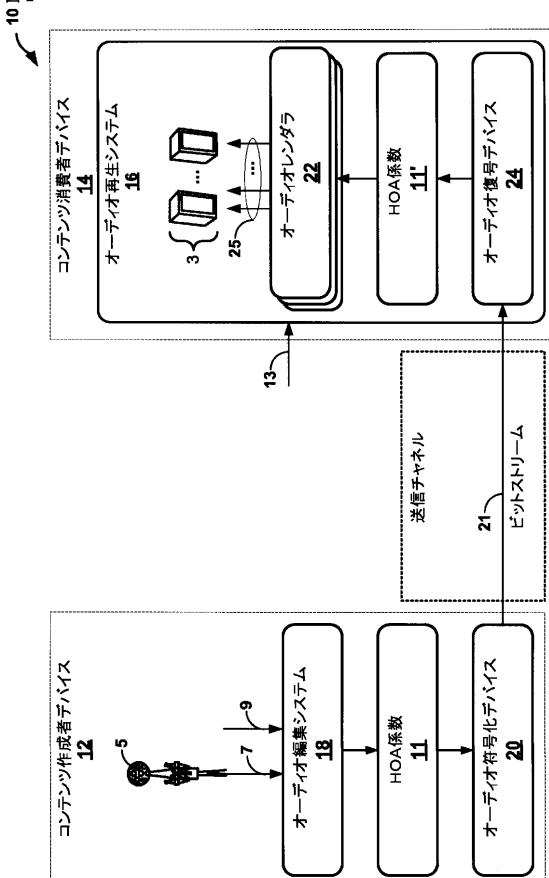


FIG. 2

【図 3】

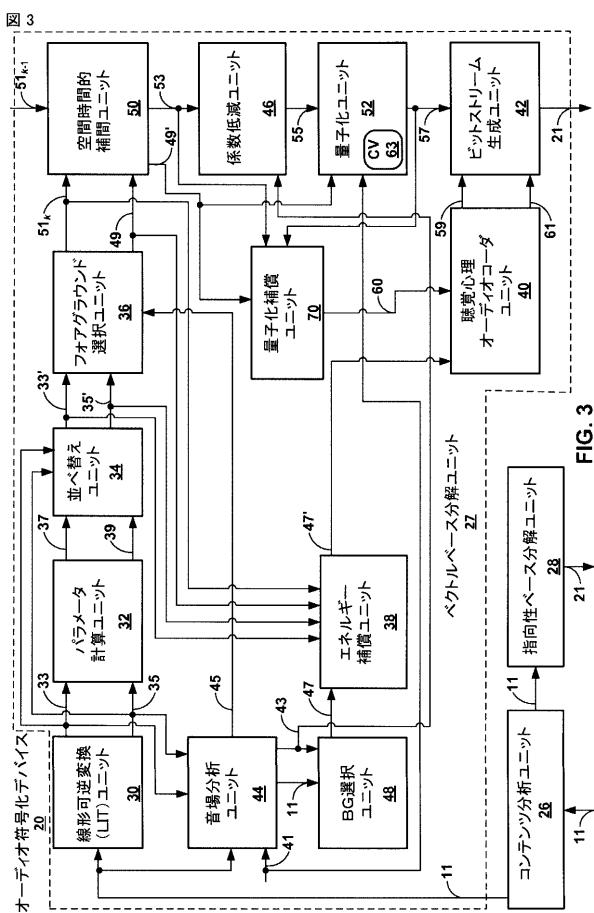


FIG. 3

【図 4】

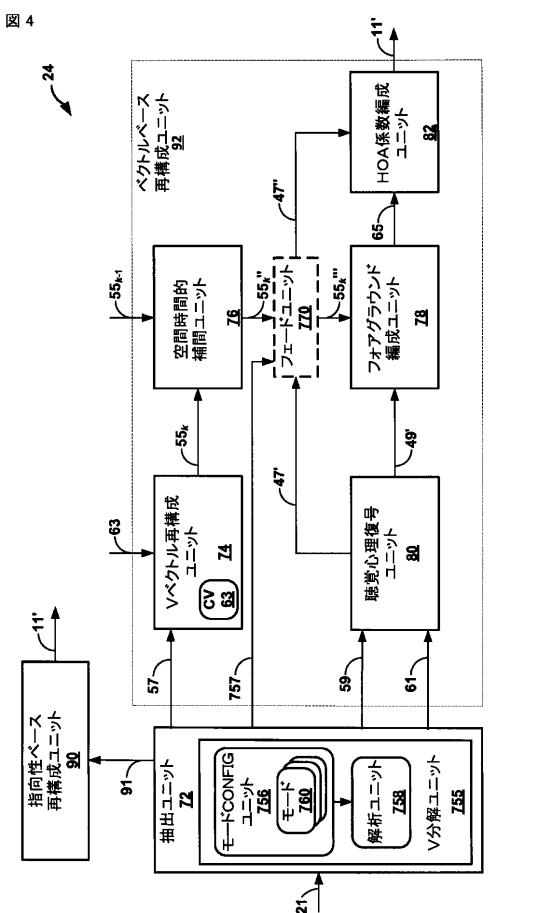
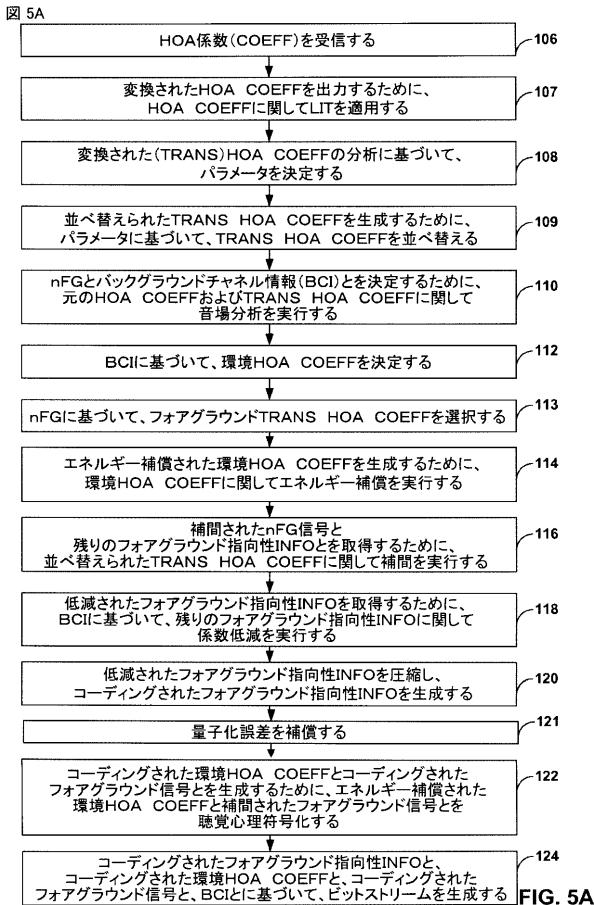


FIG. 4

【図 5A】



【図 5B】

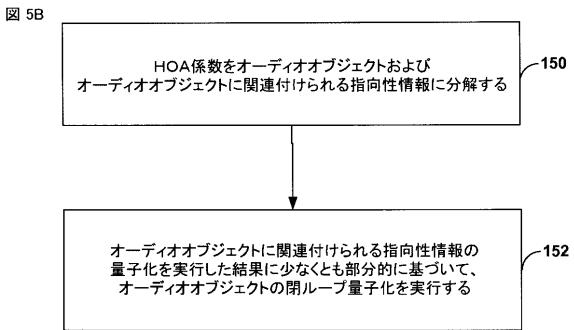
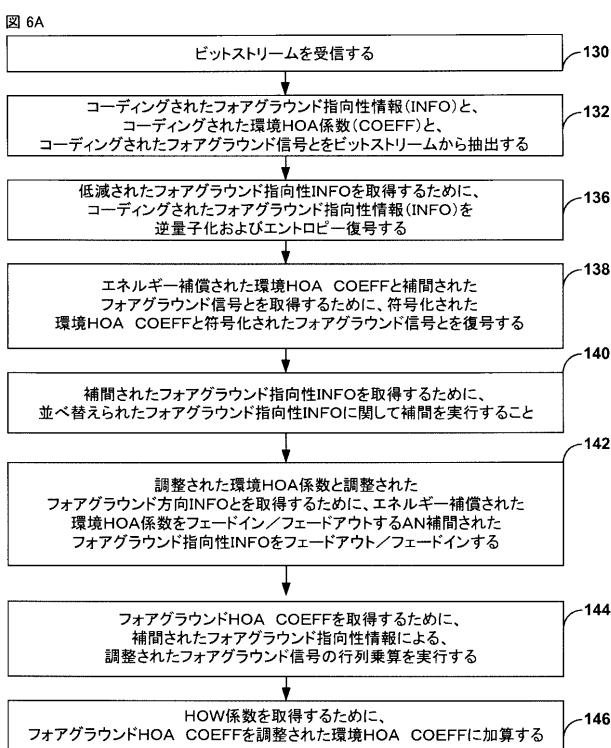


FIG. 5B

【図 6A】



【図 6B】

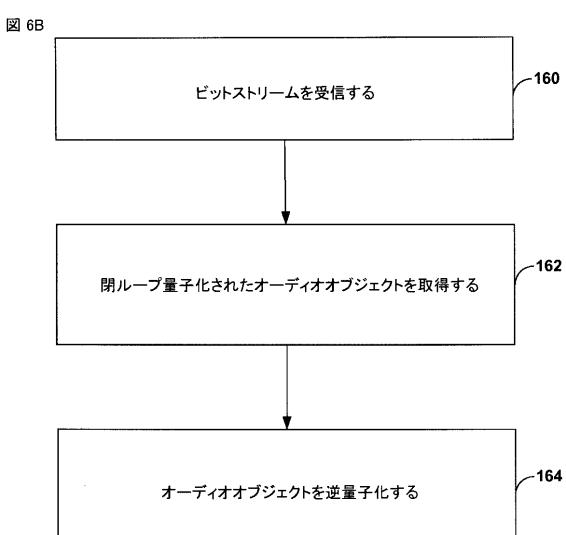


FIG. 6B

【図 7 A】

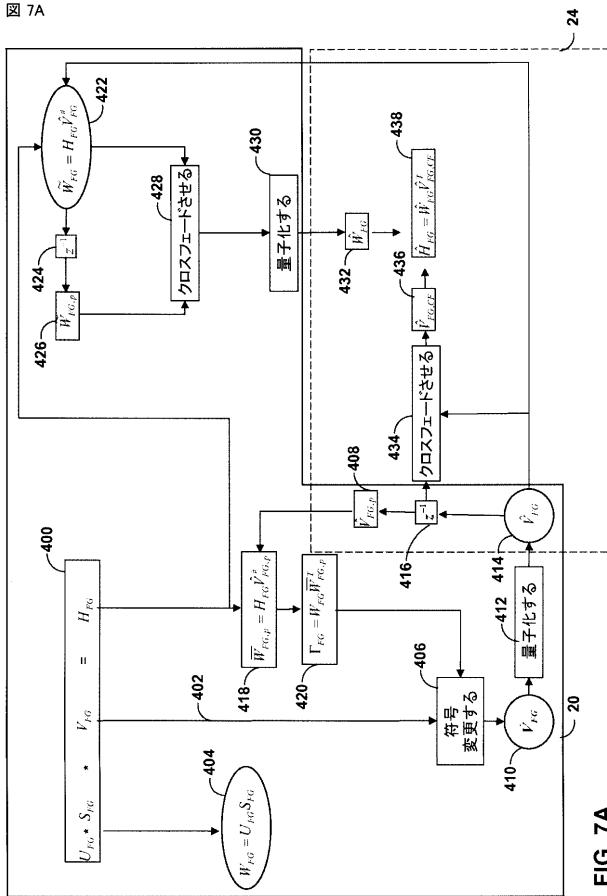


FIG. 7A

【図 7 B】

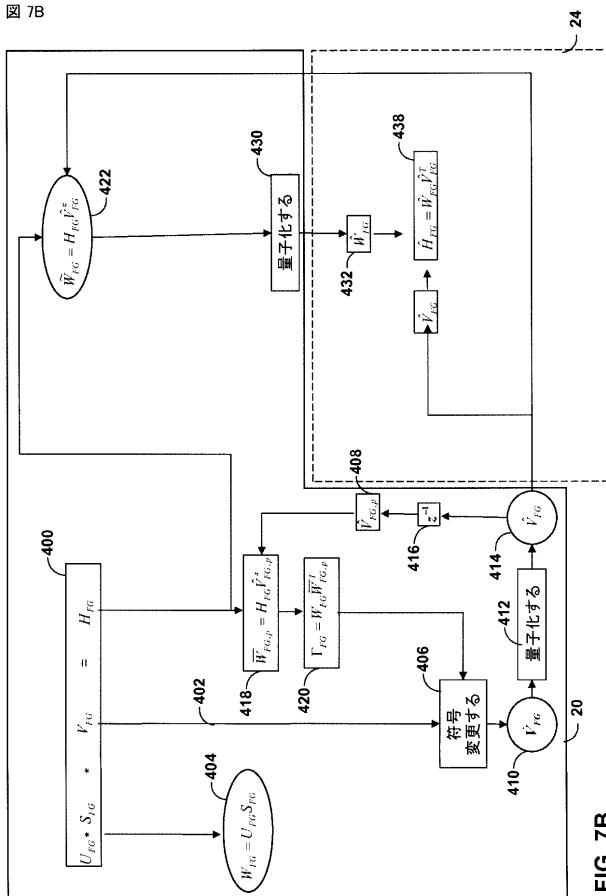


FIG. 7B

【図 8】

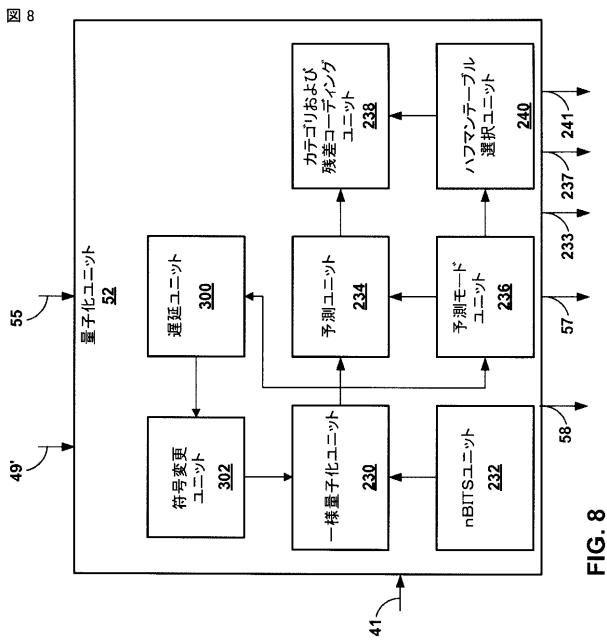


FIG. 8

【図 9】

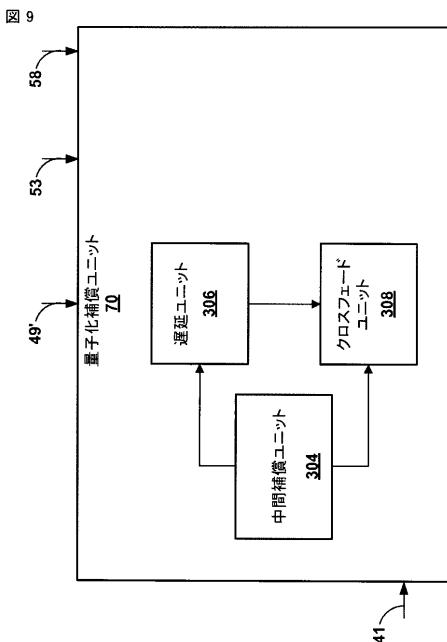


FIG. 9

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No PCT/US2015/031107						
A. CLASSIFICATION OF SUBJECT MATTER INV. G10L19/032 ADD. G10L19/20 G10L19/008								
According to International Patent Classification (IPC) or to both national classification and IPC								
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G10L								
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched								
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data								
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">Category*</th> <th style="text-align: left; padding: 2px;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left; padding: 2px;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px;">A</td> <td style="padding: 2px; vertical-align: top;"> HELLERUD ERIK ET AL: "Quantization of 2D Higher Order Ambisonics", AES CONVENTION 124; MAY 2008, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 May 2008 (2008-05-01), XP040508586, the whole document ----- -/-/ </td> <td style="padding: 2px; text-align: center;">1-30</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	HELLERUD ERIK ET AL: "Quantization of 2D Higher Order Ambisonics", AES CONVENTION 124; MAY 2008, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 May 2008 (2008-05-01), XP040508586, the whole document ----- -/-/	1-30
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.						
A	HELLERUD ERIK ET AL: "Quantization of 2D Higher Order Ambisonics", AES CONVENTION 124; MAY 2008, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 May 2008 (2008-05-01), XP040508586, the whole document ----- -/-/	1-30						
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.						
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed								
Date of the actual completion of the international search 28 July 2015		Date of mailing of the international search report 04/08/2015						
Name and mailing address of the ISA/ European Patent Office, P.B. 5018 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3015		Authorized officer Burchett, Stefanie						

INTERNATIONAL SEARCH REPORT

International application No PCT/US2015/031107

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZOTTER F ET AL: "Energy-Preserving Ambisonic Decoding", ACTA ACUSTICA UNITED WITH ACUSTICA, EUROPEAN ACOUSTICS ASSOCIATION, STUTTGART : HIRZEL, vol. 98, no. 1, 1 January 2012 (2012-01-01), pages 37-47, XP009180661, ISSN: 1436-7947, DOI: 10.3813/AAA.918490 page 38, right-hand column, line 15 - page 40, left-hand column, line 38 -----	1-30
A	AHONEN JUKKA ET AL: "Directional Analysis with Microphone Array Mounted on Rigid Cylinder for Directional Audio Coding", JAES, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, vol. 60, no. 5, 1 May 2012 (2012-05-01), pages 311-324, XP040574648, page 317, right-hand column, line 4 - page 318, left-hand column, line 22 -----	1-30
T	DVB ORGANIZATION: "ISO-IEC 23008-3 (E) (DIS of 3DA).docx", DVB, DIGITAL VIDEO BROADCASTING, C/O EBU - 17A ANCIENNE ROUTE - CH-1218 GRAND SACONNEX, GENEVA - SWITZERLAND, 8 August 2014 (2014-08-08), XP017845411, page 244, line 6 - page 247, line 37; figure 42 -----	

フロントページの続き

(31) 優先権主張番号 14/712,638

(32) 優先日 平成27年5月14日(2015.5.14)

(33) 優先権主張国 米国(US)

(81) 指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,R0,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,D0,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JP,KE,KG,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US

(72) 発明者 キム、モ・ウン

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 ペーターズ、ニ尔斯・ガンザー

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72) 発明者 セン、ディパンジャン

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

F ターム(参考) 5D018 BB23

5D162 AA13 BA14 CA01 CC12 CD01 EE05
5D220 BA06 BC05 DD03