

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7589883号
(P7589883)

(45)発行日 令和6年11月26日(2024.11.26)

(24)登録日 令和6年11月18日(2024.11.18)

(51)国際特許分類

FI

G 1 0 L 19/008 (2013.01) G 1 0 L 19/008 1 0 0

G 1 0 L 19/00 (2013.01) G 1 0 L 19/00 3 3 0 B

請求項の数 45 (全74頁)

(21)出願番号	特願2023-532525(P2023-532525)	(73)特許権者	503433420
(86)(22)出願日	令和3年5月28日(2021.5.28)		華為技術有限公司
(65)公表番号	特表2023-551016(P2023-551016 A)		HUAWEI TECHNOLOGIES CO., LTD.
(43)公表日	令和5年12月6日(2023.12.6)		中華人民共和國 5 1 8 1 2 9 広東省深
(86)国際出願番号	PCT/CN2021/096839		▲チェン▼市龍崗区坂田 華為總部▲ベ
(87)国際公開番号	WO2022/110722		ン▼公樓
(87)国際公開日	令和4年6月2日(2022.6.2)		Huawei Administrat
審査請求日	令和5年6月29日(2023.6.29)		ion Building, Banti
(31)優先権主張番号	202011377433.0		an, Longgang Distri
(32)優先日	令和2年11月30日(2020.11.30)		ct, Shenzhen, Guang
(33)優先権主張国・地域又は機関	中国(CN)	(74)代理人	dong 5 1 8 1 2 9, P. R. C
			hina
			100107766
			弁理士 伊東 忠重
			最終頁に続く

(54)【発明の名称】 オーディオ符号化及び復号方法並びに装置

(57)【特許請求の範囲】

【請求項1】

オーディオ符号化方法であって、
第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップと、
前記第1のシーンオーディオ信号及び前記第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成するステップと、
前記第1のターゲット仮想スピーカーの前記属性情報及び前記第1の仮想スピーカー信号を使用することにより、第2のシーンオーディオ信号を取得するステップと、
前記第1のシーンオーディオ信号及び前記第2のシーンオーディオ信号に基づいて残差信号を生成するステップと、
前記第1の仮想スピーカー信号及び前記残差信号を符号化し、符号化された信号をビットストリームに書き込むステップと
を含む方法。

【請求項2】

当該方法は、
前記仮想スピーカーセットに基づいて前記第1のシーンオーディオ信号から主要音場成分を取得するステップを更に含み、
第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップは、

前記主要音場成分に基づいて前記仮想スピーカーセットから前記第1のターゲット仮想スピーカーを選択するステップを含む、請求項1に記載の方法。

【請求項3】

前記主要音場成分に基づいて前記仮想スピーカーセットから前記第1のターゲット仮想スピーカーを選択するステップは、

前記主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから前記主要音場成分についてのHOA係数を選択するステップであり、前記HOA係数セット内のHOA係数は、前記仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にある、ステップと、

前記仮想スピーカーセットの中で前記主要音場成分についての前記HOA係数に対応する仮想スピーカーを前記第1のターゲット仮想スピーカーとして決定するステップとを含む、請求項2に記載の方法。

【請求項4】

前記主要音場成分に基づいて前記仮想スピーカーセットから前記第1のターゲット仮想スピーカーを選択するステップは、

前記主要音場成分に基づいて前記第1のターゲット仮想スピーカーの構成パラメータを取得するステップと、

前記第1のターゲット仮想スピーカーの前記構成パラメータに基づいて前記第1のターゲット仮想スピーカーについてのHOA係数を生成するステップと、

前記仮想スピーカーセットの中で前記第1のターゲット仮想スピーカーについての前記HOA係数に対応する仮想スピーカーを前記第1のターゲット仮想スピーカーとして決定するステップとを含む、請求項2に記載の方法。

【請求項5】

前記主要音場成分に基づいて前記第1のターゲット仮想スピーカーの構成パラメータを取得するステップは、

オーディオエンコーダの構成情報に基づいて前記仮想スピーカーセット内の複数の仮想スピーカーの構成パラメータを決定するステップと、

前記主要音場成分に基づいて前記複数の仮想スピーカーの前記構成パラメータから前記第1のターゲット仮想スピーカーの前記構成パラメータを選択するステップと

を含む、請求項4に記載の方法。

【請求項6】

前記第1のターゲット仮想スピーカーの前記構成パラメータは、前記第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含み、

前記第1のターゲット仮想スピーカーの前記構成パラメータに基づいて前記第1のターゲット仮想スピーカーについてのHOA係数を生成するステップは、

前記第1のターゲット仮想スピーカーの前記位置情報及び前記HOAオーダー情報に基づいて前記第1のターゲット仮想スピーカーについての前記HOA係数を決定するステップを含む、請求項4又は5に記載の方法。

【請求項7】

当該方法は、

前記第1のターゲット仮想スピーカーの前記属性情報を符号化し、符号化された情報を前記ビットストリームに書き込むステップを更に含む、請求項1乃至6のうちいずれか1項に記載の方法。

【請求項8】

前記第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、前記第1のターゲット仮想スピーカーの前記属性情報は、前記第1のターゲット仮想スピーカーについてのHOA係数を含み、

前記第1のシーンオーディオ信号及び前記第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成するステップは、

10

20

30

40

50

符号化されるべき前記HOA信号及び前記第1のターゲット仮想スピーカーについての前記HOA係数に対して線形結合を実行して、前記第1の仮想スピーカー信号を取得するステップを含む、請求項1乃至7のうちいずれか1項に記載の方法。

【請求項9】

前記第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、前記第1のターゲット仮想スピーカーの前記属性情報は、前記第1のターゲット仮想スピーカーの位置情報を含み、

前記第1のシーンオーディオ信号及び前記第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成するステップは、

前記第1のターゲット仮想スピーカーの前記位置情報に基づいて前記第1のターゲット仮想スピーカーについてのHOA係数を取得するステップと、

符号化されるべき前記HOA信号及び前記第1のターゲット仮想スピーカーについての前記HOA係数に対して線形結合を実行して、前記第1の仮想スピーカー信号を取得するステップと

を含む、請求項1乃至7のうちいずれか1項に記載の方法。

【請求項10】

当該方法は、

前記第1のシーンオーディオ信号に基づいて前記仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するステップと、

前記第1のシーンオーディオ信号及び前記第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するステップと、

前記第2の仮想スピーカー信号を符号化し、符号化された信号を前記ビットストリームに書き込むステップと

を更に含み、

前記第1のターゲット仮想スピーカーの前記属性情報及び前記第1の仮想スピーカー信号を使用することにより、第2のシーンオーディオ信号を取得するステップは、

前記第1のターゲット仮想スピーカーの前記属性情報、前記第1の仮想スピーカー信号、前記第2のターゲット仮想スピーカーの前記属性情報及び前記第2の仮想スピーカー信号に基づいて前記第2のシーンオーディオ信号を取得するステップを含む、請求項1乃至9のうちいずれか1項に記載の方法。

【請求項11】

当該方法は、

前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するステップを更に含み、

前記第2の仮想スピーカー信号を符号化することは、

前記整列された第2の仮想スピーカー信号を符号化することを含み、

前記第1の仮想スピーカー信号及び前記残差信号を符号化することは、

前記整列された第1の仮想スピーカー信号及び前記残差信号を符号化することを含む、請求項10に記載の方法。

【請求項12】

当該方法は、

前記第1のシーンオーディオ信号に基づいて前記仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するステップと、

前記第1のシーンオーディオ信号及び前記第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するステップと

を更に含み、

前記第1の仮想スピーカー信号及び前記残差信号を符号化することは、

前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得し、ここで、前記第1のサイド情報は前記第

10

20

30

40

50

1の仮想スピーカ信号と前記第2の仮想スピーカ信号との間の関係を示すことと、
前記ダウンミキシングされた信号、前記第1のサイド情報及び前記残差信号を符号化することと

を含む、請求項1乃至9のうちいずれか1項に記載の方法。

【請求項13】

当該方法は、

前記第1の仮想スピーカ信号及び前記第2の仮想スピーカ信号を整列させて、整列された第1の仮想スピーカ信号及び整列された第2の仮想スピーカ信号を取得するステップを更に含み、

前記第1の仮想スピーカ信号及び前記第2の仮想スピーカ信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得することは、

前記整列された第1の仮想スピーカ信号及び前記整列された第2の仮想スピーカ信号に基づいて前記ダウンミキシングされた信号及び前記第1のサイド情報を取得することを含み、

前記第1のサイド情報は、前記整列された第1の仮想スピーカ信号と前記整列された第2の仮想スピーカ信号との間の関係を示す、請求項12に記載の方法。

【請求項14】

前記第1のシーンオーディオ信号に基づいて前記仮想スピーカセットから第2のターゲット仮想スピーカを選択する前に、当該方法は、

前記第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、前記第1のターゲット仮想スピーカ以外のターゲット仮想スピーカが取得される必要があるか否かを決定するステップと、

前記第1のターゲット仮想スピーカ以外の前記ターゲット仮想スピーカが取得される必要がある場合にのみ、前記第1のシーンオーディオ信号に基づいて前記仮想スピーカセットから前記第2のターゲット仮想スピーカを選択するステップと

を更に含む、請求項10乃至13のうちいずれか1項に記載の方法。

【請求項15】

前記残差信号は、少なくとも2つのサウンドチャンネル上の残差サブ信号を含み、当該方法は、

オーディオエンコーダの構成情報及び/又は前記第1のシーンオーディオ信号の信号クラス情報に基づいて、前記少なくとも2つのサウンドチャンネル上の前記残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定するステップを更に含み、

前記第1の仮想スピーカ信号及び前記残差信号を符号化することは、

前記第1の仮想スピーカ信号と、符号化される必要があり且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号とを符号化することを含む、請求項1乃至14のうちいずれか1項に記載の方法。

【請求項16】

前記少なくとも2つのサウンドチャンネル上の前記残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、当該方法は、

第2のサイド情報を取得するステップであり、前記第2のサイド情報は、符号化される必要があり且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号と、符号化される必要がなく且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号との間の関係を示す、ステップと、

前記第2のサイド情報を前記ビットストリームに書き込むステップと

を更に含む、請求項15に記載の方法。

【請求項17】

オーディオ復号方法であって、

ビットストリームを受信するステップと、

10

20

30

40

50

前記ビットストリームを復号して、仮想スピーカー信号、残差信号及びターゲット仮想スピーカーの属性情報を取得するステップと、

前記ターゲット仮想スピーカーの前記属性情報、前記残差信号及び前記仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップと

を含み、

前記ターゲット仮想スピーカーの前記属性情報は、前記ターゲット仮想スピーカーの位置情報を含み、

前記ターゲット仮想スピーカーの前記属性情報、前記残差信号及び前記仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

前記ターゲット仮想スピーカーの前記位置情報に基づいて前記ターゲット仮想スピーカーについてのHOA係数を決定するステップであり、前記ターゲット仮想スピーカーについての前記HOA係数は行列であり、前記行列のサイズは $(M \times C)$ であり、 C はターゲット仮想スピーカーの数であり、 M は N 次HOA係数のサウンドチャンネルの数である、ステップと、

前記仮想スピーカー信号及び前記ターゲット仮想スピーカーについての前記HOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得するステップと、

前記残差信号を使用することにより、前記合成されたシーンオーディオ信号を調整して、前記再構成されたシーンオーディオ信号を取得するステップと

を含む、方法。

【請求項 18】

前記仮想スピーカー信号は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号をダウンミキシングすることにより取得されたダウンミキシングされた信号であり、当該方法は、

前記ビットストリームを復号して、第1のサイド情報を取得するステップであり、前記第1のサイド情報は、前記第1の仮想スピーカー信号と前記第2の仮想スピーカー信号との間の関係を示す、ステップと、

前記第1のサイド情報及び前記ダウンミキシングされた信号に基づいて前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号を取得するステップと

を更に含み、

ターゲット仮想スピーカーの属性情報、前記残差信号及び前記仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

前記ターゲット仮想スピーカーの前記属性情報、前記残差信号、前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号に基づいて前記再構成されたシーンオーディオ信号を取得するステップを含む、請求項17に記載の方法。

【請求項 19】

前記残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該方法は、

前記ビットストリームを復号して、第2のサイド情報を取得するステップであり、前記第2のサイド情報は、前記第1のサウンドチャンネル上の前記残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示す、ステップと、

前記第2のサイド情報及び前記第1のサウンドチャンネル上の前記残差サブ信号に基づいて前記第2のサウンドチャンネル上の前記残差サブ信号を取得するステップと

を更に含み、

ターゲット仮想スピーカーの属性情報、前記残差信号及び前記仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

前記ターゲット仮想スピーカーの前記属性情報、前記第1のサウンドチャンネル上の前記残差サブ信号、前記第2のサウンドチャンネル上の前記残差サブ信号及び前記仮想スピーカー信号に基づいて前記再構成されたシーンオーディオ信号を取得するステップを含む、請求項17又は18に記載の方法。

【請求項 20】

前記残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該方法は、

前記ビットストリームを復号して、第2のサイド情報を取得するステップであり、前記

10

20

30

40

50

第2のサイド情報は、前記第1のサウンドチャンネル上の前記残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示す、ステップと、

前記第2のサイド情報及び前記第1のサウンドチャンネル上の前記残差サブ信号に基づいて前記第3のサウンドチャンネル上の前記残差サブ信号及び前記第1のサウンドチャンネル上の更新された残差サブ信号を取得するステップと

を更に含み、

ターゲット仮想スピーカの属性情報、前記残差信号及び前記仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

前記ターゲット仮想スピーカの前記属性情報、前記第1のサウンドチャンネル上の前記更新された残差サブ信号、前記第3のサウンドチャンネル上の前記残差サブ信号及び前記仮想スピーカ信号に基づいて前記再構成されたシーンオーディオ信号を取得するステップを含む、請求項 1 7 又は 1 8 に記載の方法。

【請求項 2 1】

オーディオ符号化装置であって、

第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカセットから第1のターゲット仮想スピーカを選択するように構成された取得モジュールと、

前記第1のシーンオーディオ信号及び前記第1のターゲット仮想スピーカの属性情報に基づいて第1の仮想スピーカ信号を生成するように構成された信号生成モジュールであり、

前記信号生成モジュールは、前記第1のターゲット仮想スピーカの前記属性情報及び前記第1の仮想スピーカ信号を使用することにより、第2のシーンオーディオ信号を取得するように構成され、

前記信号生成モジュールは、前記第1のシーンオーディオ信号及び前記第2のシーンオーディオ信号に基づいて残差信号を生成するように構成される、信号生成モジュールと、

前記第1の仮想スピーカ信号及び前記残差信号を符号化して、ビットストリームを取得するように構成された符号化モジュールと

を含む装置。

【請求項 2 2】

前記取得モジュールは、前記仮想スピーカセットに基づいて前記第1のシーンオーディオ信号から主要音場成分を取得し、前記主要音場成分に基づいて前記仮想スピーカセットから前記第1のターゲット仮想スピーカを選択するように構成される、請求項 2 1 に記載の装置。

【請求項 2 3】

前記取得モジュールは、前記主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから前記主要音場成分についてのHOA係数を選択するように構成され、ここで、前記HOA係数セット内のHOA係数は、前記仮想スピーカセット内の仮想スピーカと1対1の対応関係にあり、前記仮想スピーカセットの中で前記主要音場成分についての前記HOA係数に対応する仮想スピーカを前記第1のターゲット仮想スピーカとして決定するように構成される、請求項 2 2 に記載の装置。

【請求項 2 4】

前記取得モジュールは、前記主要音場成分に基づいて前記第1のターゲット仮想スピーカの構成パラメータを取得し、前記第1のターゲット仮想スピーカの前記構成パラメータに基づいて前記第1のターゲット仮想スピーカについてのHOA係数を生成し、前記仮想スピーカセットの中で前記第1のターゲット仮想スピーカについての前記HOA係数に対応する仮想スピーカを前記第1のターゲット仮想スピーカとして決定するように構成される、請求項 2 2 に記載の装置。

【請求項 2 5】

前記取得モジュールは、オーディオエンコーダの構成情報に基づいて前記仮想スピーカセット内の複数の仮想スピーカの構成パラメータを決定し、前記主要音場成分に基づいて前記複数の仮想スピーカの前記構成パラメータから前記第1のターゲット仮想スピー

10

20

30

40

50

ーカーの前記構成パラメータを選択するように構成される、請求項 2 4 に記載の装置。

【請求項 2 6】

前記第1のターゲット仮想スピーカーの前記構成パラメータは、前記第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含み、

前記取得モジュールは、前記第1のターゲット仮想スピーカーの前記位置情報及び前記HOAオーダー情報に基づいて前記第1のターゲット仮想スピーカーについての前記HOA係数を決定するように構成される、請求項 2 4 又は 2 5 に記載の装置。

【請求項 2 7】

前記符号化モジュールは、前記第1のターゲット仮想スピーカーの前記属性情報を符号化し、符号化された情報を前記ビットストリームに書き込むように更に構成される、請求項 2 1 乃至 2 6 のうちいずれか1項に記載の装置。

【請求項 2 8】

前記第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、前記第1のターゲット仮想スピーカーの前記属性情報は、前記第1のターゲット仮想スピーカーについてのHOA係数を含み、

前記信号生成モジュールは、符号化されるべき前記HOA信号及び前記第1のターゲット仮想スピーカーについての前記HOA係数に対して線形結合を実行して、前記第1の仮想スピーカー信号を取得するように構成される、請求項 2 1 乃至 2 7 のうちいずれか1項に記載の装置。

【請求項 2 9】

前記第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、前記第1のターゲット仮想スピーカーの前記属性情報は、前記第1のターゲット仮想スピーカーの位置情報を含み、

前記信号生成モジュールは、前記第1のターゲット仮想スピーカーの前記位置情報に基づいて前記第1のターゲット仮想スピーカーについてのHOA係数を取得し、符号化されるべき前記HOA信号及び前記第1のターゲット仮想スピーカーについての前記HOA係数に対して線形結合を実行して、前記第1の仮想スピーカー信号を取得するように構成される、請求項 2 1 乃至 2 7 のうちいずれか1項に記載の装置。

【請求項 3 0】

前記取得モジュールは、前記第1のシーンオーディオ信号に基づいて前記仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するように構成され、

前記信号生成モジュールは、前記第1のシーンオーディオ信号及び前記第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するように構成され、

前記符号化モジュールは、前記第2の仮想スピーカー信号を符号化し、符号化された信号を前記ビットストリームに書き込むように構成され、

前記信号生成モジュールは、前記第1のターゲット仮想スピーカーの前記属性情報、前記第1の仮想スピーカー信号、前記第2のターゲット仮想スピーカーの前記属性情報及び前記第2の仮想スピーカー信号に基づいて前記第2のシーンオーディオ信号を取得するように構成される、請求項 2 1 乃至 2 9 のうちいずれか1項に記載の装置。

【請求項 3 1】

前記信号生成モジュールは、前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するように構成され、

前記符号化モジュールは、前記整列された第2の仮想スピーカー信号を符号化するように構成され、

前記符号化モジュールは、前記整列された第1の仮想スピーカー信号及び前記残差信号を符号化するように構成される、請求項 3 0 に記載の装置。

【請求項 3 2】

前記取得モジュールは、前記第1のシーンオーディオ信号に基づいて前記仮想スピーカ

10

20

30

40

50

ーセットから第2のターゲット仮想スピーカーを選択するように構成され、

前記信号生成モジュールは、前記第1のシーンオーディオ信号及び前記第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するように構成され、

前記符号化モジュールは、前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得するように構成され、前記第1のサイド情報は前記第1の仮想スピーカー信号と前記第2の仮想スピーカー信号との間の関係を示し、

前記符号化モジュールは、前記ダウンミキシングされた信号、前記第1のサイド情報及び前記残差信号を符号化するように構成される、請求項 2.1 乃至 2.9 のうちいずれか 1 項に記載の装置。

10

【請求項 3.3】

前記信号生成モジュールは、前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するように構成され、

前記符号化モジュールは、前記整列された第1の仮想スピーカー信号及び前記整列された第2の仮想スピーカー信号に基づいて前記ダウンミキシングされた信号及び前記第1のサイド情報を取得するように構成され、

前記第1のサイド情報は、前記整列された第1の仮想スピーカー信号と前記整列された第2の仮想スピーカー信号との間の関係を示す、請求項 3.2 に記載の装置。

20

【請求項 3.4】

前記取得モジュールは、前記第1のシーンオーディオ信号に基づいて前記仮想スピーカーセットから前記第2のターゲット仮想スピーカーを選択する前に、前記第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、前記第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要があるか否かを決定し、前記第1のターゲット仮想スピーカー以外の前記ターゲット仮想スピーカーが取得される必要がある場合にのみ、前記第1のシーンオーディオ信号に基づいて前記仮想スピーカーセットから前記第2のターゲット仮想スピーカーを選択するように構成される、請求項 3.0 乃至 3.3 のうちいずれか 1 項に記載の装置。

【請求項 3.5】

前記残差信号は、少なくとも2つのサウンドチャンネル上の残差サブ信号を含み、

前記信号生成モジュールは、オーディオエンコーダの構成情報及び/又は前記第1のシーンオーディオ信号の信号クラス情報に基づいて、前記少なくとも2つのサウンドチャンネル上の前記残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定するように構成され、

前記符号化モジュールは、前記第1の仮想スピーカー信号と、符号化される必要があり且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号とを符号化するように構成される、請求項 2.1 乃至 3.4 のうちいずれか 1 項に記載の装置。

30

【請求項 3.6】

前記取得モジュールは、前記少なくとも2つのサウンドチャンネル上の前記残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、第2のサイド情報を取得するように構成され、前記第2のサイド情報は、符号化される必要があり且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号と、符号化される必要がなく且つ前記少なくとも1つのサウンドチャンネル上にある前記残差サブ信号との間の関係を示し、

40

前記符号化モジュールは、前記第2のサイド情報を前記ビットストリームに書き込むように構成される、請求項 3.5 に記載の装置。

【請求項 3.7】

オーディオ復号装置であって、

ビットストリームを受信するように構成された受信モジュールと、

50

前記ビットストリームを復号して、仮想スピーカー信号、残差信号及びターゲット仮想スピーカーの属性情報を取得するように構成された復号モジュールと、
前記ターゲット仮想スピーカーの前記属性情報、前記残差信号及び前記仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するように構成された再構成モジュールと

を含み、

前記ターゲット仮想スピーカーの前記属性情報は、前記ターゲット仮想スピーカーの位置情報を含み、

前記再構成モジュールは、前記ターゲット仮想スピーカーの前記位置情報に基づいて前記ターゲット仮想スピーカーについてのHOA係数を決定し、前記仮想スピーカー信号及び前記ターゲット仮想スピーカーについての前記HOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、前記残差信号を使用することにより、前記合成されたシーンオーディオ信号を調整して、前記再構成されたシーンオーディオ信号を取得するように構成され、前記ターゲット仮想スピーカーについての前記HOA係数は行列であり、前記行列のサイズは $(M \times C)$ であり、 C はターゲット仮想スピーカーの数であり、 M は N 次HOA係数のサウンドチャンネルの数である、装置。

【請求項 38】

前記仮想スピーカー信号は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号をダウンミキシングすることにより取得されたダウンミキシングされた信号であり、当該装置は第1の信号補償モジュールを更に含み、

前記復号モジュールは、前記ビットストリームを復号して、第1のサイド情報を取得するように構成され、前記第1のサイド情報は、前記第1の仮想スピーカー信号と前記第2の仮想スピーカー信号との間の関係を示し、

前記第1の信号補償モジュールは、前記第1のサイド情報及び前記ダウンミキシングされた信号に基づいて前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号を取得するように構成され、

前記再構成モジュールは、前記ターゲット仮想スピーカーの前記属性情報、前記残差信号、前記第1の仮想スピーカー信号及び前記第2の仮想スピーカー信号に基づいて前記再構成されたシーンオーディオ信号を取得するように構成される、請求項 37 に記載の装置。

【請求項 39】

前記残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該装置は第2の信号補償モジュールを更に含み、

前記復号モジュールは、前記ビットストリームを復号して、第2のサイド情報を取得するように構成され、前記第2のサイド情報は、前記第1のサウンドチャンネル上の前記残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示し、

前記第2の信号補償モジュールは、前記第2のサイド情報及び前記第1のサウンドチャンネル上の前記残差サブ信号に基づいて前記第2のサウンドチャンネル上の前記残差サブ信号を取得するように構成され、

前記再構成モジュールは、前記ターゲット仮想スピーカーの前記属性情報、前記第1のサウンドチャンネル上の前記残差サブ信号、前記第2のサウンドチャンネル上の前記残差サブ信号及び前記仮想スピーカー信号に基づいて前記再構成されたシーンオーディオ信号を取得するように構成される、請求項 37 又は 38 に記載の装置。

【請求項 40】

前記残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該装置は第3の信号補償モジュールを更に含み、

前記復号モジュールは、前記ビットストリームを復号して、第2のサイド情報を取得するように構成され、前記第2のサイド情報は、前記第1のサウンドチャンネル上の前記残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示し、

前記第3の信号補償モジュールは、前記第2のサイド情報及び前記第1のサウンドチャンネル上の前記残差サブ信号に基づいて前記第3のサウンドチャンネル上の前記残差サブ信号及

10

20

30

40

50

び前記第1のサウンドチャンネル上の更新された残差サブ信号を取得するように構成され、
前記再構成モジュールは、前記ターゲット仮想スピーカの前記属性情報、前記第1のサウンドチャンネル上の前記更新された残差サブ信号、前記第3のサウンドチャンネル上の前記残差サブ信号及び前記仮想スピーカ信号に基づいて前記再構成されたシーンオーディオ信号を取得するように構成される、請求項37又は38に記載の装置。

【請求項41】

オーディオ符号化装置であって、

当該オーディオ符号化装置は少なくとも1つのプロセッサを含み、

前記少なくとも1つのプロセッサは、メモリに結合され、前記メモリ内の命令を読み取って実行して、請求項1乃至16のうちいずれか1項に記載の方法を実現するように構成される、オーディオ符号化装置。

10

【請求項42】

当該オーディオ符号化装置は前記メモリを更に含む、請求項41に記載のオーディオ符号化装置。

【請求項43】

オーディオ復号装置であって、

当該オーディオ復号装置は少なくとも1つのプロセッサを含み、

前記少なくとも1つのプロセッサは、メモリに結合され、前記メモリ内の命令を読み取って実行して、請求項17乃至20のうちいずれか1項に記載の方法を実現するように構成される、オーディオ復号装置。

20

【請求項44】

当該オーディオ復号装置は前記メモリを更に含む、請求項43に記載のオーディオ復号装置。

【請求項45】

命令を含むコンピュータ読み取り可能記憶媒体であって、

前記命令がコンピュータ上で実行されたとき、前記コンピュータは、請求項1乃至16のうちいずれか1項に記載の方法又は請求項17乃至20のうちいずれか1項に記載の方法を実行することが可能になる、コンピュータ読み取り可能記憶媒体。

【発明の詳細な説明】

【技術分野】

30

【0001】

〔関連出願への相互参照〕

この出願は、2020年11月30日に中国国家知識産権局に出願された「AUDIO ENCODING AND DECODING METHOD AND APPARATUS」という名称の中国特許出願第202011377433.0号に対する優先権を主張し、その全内容を参照により援用する。

【0002】

〔技術分野〕

この出願は、オーディオ符号化及び復号技術の分野に関し、特に、オーディオ符号化及び復号方法並びに装置に関する。

【背景技術】

40

【0003】

3次元オーディオ技術は、現実世界におけるサウンドイベント及び3次元音場情報を取得、処理、伝送、レンダリング及び再生するために使用されるオーディオ技術である。3次元オーディオ技術は、強い空間感覚、包囲感覚、没入感覚をサウンドに与え、「真に迫る」並外れた聴覚体験を人々に与える。高次アンビソニックス(higher order ambisonics, HOA)技術は、録音、符号化及び再生の段階においてスピーカレイアウトに依存しない特性と、HOAフォーマットでデータを回転再生する特性とを有し、3次元オーディオ再生においてより高い柔軟性を有し、したがって、より注目されて研究されている。

【0004】

より良いオーディオ聴覚効果を達成するために、HOA技術は、サウンドシーンに関する

50

より詳細な情報を記録するために大量のデータを必要とする。3次元オーディオ信号のシーンベースのサンプリング及び記憶は、オーディオ信号の空間情報の記憶及び伝送をより促しているが、HOAオーダーが増加するとより多くのデータが生成され、大量のデータは伝送及び記憶に課題を引き起こす。したがって、HOA信号が符号化及び復号される必要がある。

【0005】

現在、マルチチャネルデータを符号化及び復号するための方法が存在し、以下を含む。エンコーダのコアエンコーダ(例えば、16チャネルエンコーダ)は、元のシーンにおけるオーディオ信号の各サウンドチャネルを直接符号化し、次いで、ビットストリームを出力する。デコーダのコアデコーダ(例えば、16チャネルデコーダ)は、ビットストリームを復号して、復号シーンにおけるオーディオ信号の各サウンドチャネルを取得する。

10

【0006】

上記のマルチチャネル符号化及び復号方法では、対応するエンコーダ及びデコーダが、元のシーンにおけるオーディオ信号のサウンドチャネルの数に基づいて適応される必要がある。さらに、サウンドチャネルの数が増加すると、ビットストリーム圧縮中の大きいデータ量及び高い帯域幅占有率という問題が存在する。

【発明の概要】

【0007】

この出願の実施形態は、符号化及び復号効率を改善するために、符号化及び復号されるデータの量を低減するためのオーディオ符号化及び復号方法並びに装置を提供する。

20

【0008】

上記の技術的問題を解決するために、この出願の実施形態は、以下の技術的解決策を提供する。

【0009】

第1の態様によれば、この出願の実施形態は、オーディオ符号化方法を提供し、

第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップと、

第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成するステップと、

第1のターゲット仮想スピーカーの属性情報及び第1の仮想スピーカー信号を使用することにより、第2のシーンオーディオ信号を取得するステップと、

30

第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて残差信号を生成するステップと、

第1の仮想スピーカー信号及び残差信号を符号化し、符号化された信号をビットストリームに書き込むステップと

を含む。

【0010】

この出願の実施形態では、まず、第1のターゲット仮想スピーカーは、第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから選択される。第1の仮想スピーカー信号は、第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて生成される。次いで、第2のシーンオーディオ信号は、第1のターゲット仮想スピーカーの属性情報及び第1の仮想スピーカー信号を使用することにより取得される。残差信号は、第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて生成される。最後に、第1の仮想スピーカー信号及び残差信号は符号化され、ビットストリームに書き込まれる。この出願のこの実施形態では、第1の仮想スピーカー信号は、第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて生成できる。さらに、オーディオエンコーダは、第1の仮想スピーカー信号及び第1のターゲット仮想スピーカーの属性情報に基づいて残差信号を更に取得できる。オーディオエンコーダは、第1のシーンオーディオ信号を直接符号化する代わりに、第1の仮想スピーカー信号及び残差信号を符号化する。この出願のこの実施形態では、第1のターゲット仮想スピー

40

50

カーは、第1のシーンオーディオ信号に基づいて選択され、第1のターゲット仮想スピーカーに基づいて生成された第1の仮想スピーカー信号は、空間内のリスナーの位置における音場を表すことができる。当該位置における音場は、第1のシーンオーディオ信号が記録されるとき元の音場にできるだけ近くなり、それにより、オーディオエンコーダの符号化品質を確保する。さらに、第1の仮想スピーカー信号及び残差信号は、ビットストリームを取得するために符号化され、第1の仮想スピーカー信号の符号化データの量が第1のターゲット仮想スピーカーに関連し、第1のシーンオーディオ信号のサウンドチャンネルの数に関連せず、それにより、符号化データの量が低減され、符号化効率が改善されるようにする。

【0011】

可能な実現方式では、当該方法は、

仮想スピーカーセットに基づいて第1のシーンオーディオ信号から主要音場成分を取得するステップを更に含み、

第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップは、

主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップを含む。

【0012】

上記の解決策では、仮想スピーカーセット内の各仮想スピーカーは1つの音場成分に対応し、第1のターゲット仮想スピーカーは、主要音場成分に基づいて仮想スピーカーセットから選択される。例えば、主要音場成分に対応する仮想スピーカーは、エンコーダにより選択された第1のターゲット仮想スピーカーである。この出願のこの実施形態では、エンコーダは、主要音場成分に基づいて第1のターゲット仮想スピーカーを選択して、エンコーダが第1のターゲット仮想スピーカーを決定する必要があるという問題を解決できる。

【0013】

可能な実現方式では、主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップは、

主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから主要音場成分についてのHOA係数を選択するステップであり、HOA係数セット内のHOA係数は、仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にある、ステップと、

仮想スピーカーセットの中で主要音場成分についてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するステップとを含む。

【0014】

上記の解決策では、エンコーダは仮想スピーカーセットに基づいてHOA係数セットを予め構成し、HOA係数セット内のHOA係数と仮想スピーカーセット内の仮想スピーカーとの間に1対1の対応関係が存在する。したがって、HOA係数が主要音場成分に基づいて選択された後に、1対1の対応関係に基づいて、主要音場成分についてのHOA係数に対応するターゲット仮想スピーカーを求めて仮想スピーカーセットが検索され、見つかったターゲット仮想スピーカーが第1のターゲット仮想スピーカーである。これは、エンコーダが第1のターゲット仮想スピーカーを決定する必要があるという問題を解決する。

【0015】

可能な実現方式では、主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するステップは、

主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを取得するステップと、

第1のターゲット仮想スピーカーの構成パラメータに基づいて第1のターゲット仮想スピーカーについてのHOA係数を生成するステップと、

仮想スピーカーセットの中で第1のターゲット仮想スピーカーについてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するステップと

10

20

30

40

50

を含む。

【0016】

上記の解決策では、主要音場成分を取得した後に、エンコーダは主要音場成分に基づいて第1のターゲット仮想スピーカースの構成パラメータを決定できる。例えば、主要音場成分は、複数の音場成分の中で最も大きい値を有する1つ以上の音場成分であるか、或いは、主要音場成分は、複数の音場成分の中で支配的な方向を有する1つ以上の音場成分でもよい。主要音場成分は、第1のシーンオーディオ信号に一致する第1のターゲット仮想スピーカースを決定するために使用でき、対応する属性情報は、第1のターゲット仮想スピーカースについて構成され、第1のターゲット仮想スピーカースについてのHOA係数は、第1のターゲット仮想スピーカースの設定構成パラメータに基づいて生成できる。HOA係数を生成するプロセスは、HOAアルゴリズムを使用することにより実現でき、詳細はここでは再び説明しない。仮想スピーカースセット内の各仮想スピーカースは、HOA係数に対応する。したがって、第1のターゲット仮想スピーカースは、各仮想スピーカースについてのHOA係数に基づいて仮想スピーカースセットから選択され、エンコーダが第1のターゲット仮想スピーカースを決定する必要があるという問題を解決できる。

10

【0017】

可能な実現方式では、主要音場成分に基づいて第1のターゲット仮想スピーカースの構成パラメータを取得するステップは、

オーディオエンコーダの構成情報に基づいて仮想スピーカースセット内の複数の仮想スピーカースの構成パラメータを決定するステップと、

20

主要音場成分に基づいて複数の仮想スピーカースの構成パラメータから第1のターゲット仮想スピーカースの構成パラメータを選択するステップとを含む。

【0018】

上記の解決策では、エンコーダは、仮想スピーカースセットから複数の仮想スピーカースの構成パラメータを取得する。仮想スピーカース毎に、対応する仮想スピーカース構成パラメータが存在し、各仮想スピーカース構成パラメータは、仮想スピーカースのHOAオーダー及び仮想スピーカースの位置座標のような情報を含むが、これらに限定されない。各仮想スピーカースの構成パラメータは、仮想スピーカースについてのHOA係数を生成するために使用できる。HOA係数を生成するプロセスは、HOAアルゴリズムを使用することにより実現でき、詳細はここでは再び説明しない。仮想スピーカースセット内の仮想スピーカース毎にHOA係数が生成され、仮想スピーカースセット内の全ての仮想スピーカースにそれぞれ構成されたHOA係数がHOA係数セットを形成して、エンコーダが仮想スピーカースセット内の各仮想スピーカースについてのHOA係数を決定する必要があるという問題を解決する。

30

【0019】

可能な実現方式では、第1のターゲット仮想スピーカースの構成パラメータは、第1のターゲット仮想スピーカースの位置情報及びHOAオーダー情報を含み、

第1のターゲット仮想スピーカースの構成パラメータに基づいて第1のターゲット仮想スピーカースについてのHOA係数を生成するステップは、

第1のターゲット仮想スピーカースの位置情報及びHOAオーダー情報に基づいて第1のターゲット仮想スピーカースについてのHOA係数を決定するステップを含む。

40

【0020】

上記の解決策では、仮想スピーカースセット内の各仮想スピーカースの構成パラメータは、仮想スピーカースの位置情報及び仮想スピーカースのHOAオーダー情報を含んでもよい。同様に、第1のターゲット仮想スピーカースの構成パラメータは、第1のターゲット仮想スピーカースの位置情報及びHOAオーダー情報を含む。例えば、仮想スピーカースセット内の各仮想スピーカースの位置情報は、局所等距離仮想スピーカース空間分布方式に従って決定できる。局所等距離仮想スピーカース空間分布方式は、複数の仮想スピーカースが局所的な等距離の方式で空間内に分布することを意味する。例えば、局所的な等距離の方式は、均等分布又は不均等分布を含んでもよい。各仮想スピーカースの位置情報及びHOAオーダー情報の双方は

50

、仮想スピーカについてのHOA係数を生成するために使用できる。HOA係数を生成するプロセスは、HOAアルゴリズムを使用することにより実現できる。これは、エンコーダが第1のターゲット仮想スピーカについてのHOA係数を決定する必要があるという問題を解決する。

【0021】

可能な実現方式では、当該方法は、

第1のターゲット仮想スピーカの属性情報を符号化し、符号化された情報をビットストリームに書き込むステップを更に含む。

【0022】

上記の解決策では、仮想スピーカを符号化することに加えて、エンコーダはまた、第1のターゲット仮想スピーカの属性情報を符号化し、第1のターゲット仮想スピーカの符号化された属性情報をビットストリームに書き込むことができる。この場合、取得されたビットストリームは、符号化された仮想スピーカと、第1のターゲット仮想スピーカの符号化された属性情報とを含んでもよい。この出願のこの実施形態では、ビットストリームは、第1のターゲット仮想スピーカの符号化された属性情報を搬送でき、それにより、デコーダがビットストリームを復号することにより第1のターゲット仮想スピーカの属性情報を決定して、デコーダによるオーディオ復号を容易にできるようにする。

【0023】

可能な実現方式では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカの属性情報は、第1のターゲット仮想スピーカについてのHOA係数を含み、

第1のシーンオーディオ信号及び第1のターゲット仮想スピーカの属性情報に基づいて第1の仮想スピーカ信号を生成するステップは、

符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカ信号を取得するステップを含む。

【0024】

上記の解決策では、第1のシーンオーディオ信号が符号化されるべきHOA信号である例が使用される。まず、エンコーダは、第1のターゲット仮想スピーカについてのHOA係数を決定する。例えば、エンコーダは、主要音場成分に基づいてHOA係数セットからHOA係数を選択し、選択されたHOA係数は第1のターゲット仮想スピーカについてのHOA係数である。エンコーダが符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数を取得した後に、第1の仮想スピーカ信号は、符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数に基づいて生成できる。符号化されるべきHOA信号は、第1のターゲット仮想スピーカについてのHOA係数を使用することにより線形結合を実行することで取得でき、第1の仮想スピーカ信号の解決が線形結合の解決に変換できる。

【0025】

可能な実現方式では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカの属性情報は、第1のターゲット仮想スピーカの位置情報を含み、

第1のシーンオーディオ信号及び第1のターゲット仮想スピーカの属性情報に基づいて第1の仮想スピーカ信号を生成するステップは、

第1のターゲット仮想スピーカの位置情報に基づいて第1のターゲット仮想スピーカについてのHOA係数を取得するステップと、

符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカ信号を取得するステップとを含む。

【0026】

上記の解決策では、エンコーダが符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数を取得した後に、エンコーダは、符号化されるべきHOA信

10

20

30

40

50

号及び第1のターゲット仮想スピーカについてのHOA係数に対して線形結合を実行する。言い換えると、エンコーダは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数を組み合わせて線形結合行列を取得する。次いで、エンコーダは、線形結合行列の最適解を取得でき、取得された最適解は第1の仮想スピーカ信号である。

【0027】

可能な実現方式では、当該方法は、

第1のシーンオーディオ信号に基づいて仮想スピーカセットから第2のターゲット仮想スピーカを選択するステップと、

第1のシーンオーディオ信号及び第2のターゲット仮想スピーカの属性情報に基づいて第2の仮想スピーカ信号を生成するステップと、

第2の仮想スピーカ信号を符号化し、符号化された信号をビットストリームに書き込むステップと

を更に含み、

対応して、第1のターゲット仮想スピーカの属性情報及び第1の仮想スピーカ信号を使用することにより、第2のシーンオーディオ信号を取得するステップは、

第1のターゲット仮想スピーカの属性情報、第1の仮想スピーカ信号、第2のターゲット仮想スピーカの属性情報及び第2の仮想スピーカ信号に基づいて第2のシーンオーディオ信号を取得するステップを含む。

【0028】

上記の解決策では、エンコーダは、第1のターゲット仮想スピーカの属性情報を取得でき、第1のターゲット仮想スピーカは、仮想スピーカセット内にあり且つ第1の仮想スピーカ信号を再生するために使用される仮想スピーカである。エンコーダは、第2のターゲット仮想スピーカの属性情報を取得でき、第2のターゲット仮想スピーカは、仮想スピーカセット内にあり且つ第2の仮想スピーカ信号を再生するために使用される仮想スピーカである。第1のターゲット仮想スピーカの属性情報は、第1のターゲット仮想スピーカの位置情報と、第1のターゲット仮想スピーカについてのHOA係数とを含んでもよい。第2のターゲット仮想スピーカの属性情報は、第2のターゲット仮想スピーカの位置情報と、第2のターゲット仮想スピーカについてのHOA係数とを含んでもよい。エンコーダが第1の仮想スピーカ信号及び第2の仮想スピーカ信号を取得した後に、エンコーダは、第1のターゲット仮想スピーカの属性情報及び第2のターゲット仮想スピーカの属性情報に基づいて信号再構成を実行し、信号再構成を通じて第2のシーンオーディオ信号を取得できる。

【0029】

可能な実現方式では、当該方法は、

第1の仮想スピーカ信号及び第2の仮想スピーカ信号を整列させて、整列された第1の仮想スピーカ信号及び整列された第2の仮想スピーカ信号を取得するステップを更に含み、

対応して、第2の仮想スピーカ信号を符号化することは、

整列された第2の仮想スピーカ信号を符号化することを含み、

対応して、第1の仮想スピーカ信号及び残差信号を符号化することは、

整列された第1の仮想スピーカ信号及び残差信号を符号化することを含む。

【0030】

上記の解決策では、整列された第1の仮想スピーカ信号を取得した後に、エンコーダは、整列された第1の仮想スピーカ信号及び残差信号を符号化できる。この出願の実施形態では、第1の仮想スピーカ信号のサウンドチャンネルを再び調整して整列させることにより、チャンネル間相関が強化されて、コアエンコーダによる第1の仮想スピーカ信号の符号化処理を容易にする。

【0031】

可能な実現方式では、当該方法は、

第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するステップと、

第1のシーンオーディオ信号及び第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するステップと

を更に含み、

対応して、第1の仮想スピーカー信号及び残差信号を符号化することは、

第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得し、ここで、第1のサイド情報は第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示すことと、

ダウンミキシングされた信号、第1のサイド情報及び残差信号を符号化することとを含む。

10

【0032】

上記の解決策では、エンコーダが第1の仮想スピーカー信号及び第2の仮想スピーカー信号を取得した後に、エンコーダは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングを更に実行して、ダウンミキシングされた信号を生成でき、例えば、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に対して振幅ダウンミキシングを実行して、ダウンミキシングされた信号を取得できる。さらに、第1のサイド情報は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいて更に生成できる。第1のサイド情報は、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示し、当該関係は複数の実現方式を有する。第1のサイド情報は、デコーダにより、ダウンミキシングされた信号をアップミキシングし、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元するために使用できる。例えば、第1のサイド情報は信号情報ロス分析パラメータを含み、それにより、デコーダは信号情報ロス分析パラメータを使用することにより第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元するようにする。他の例では、第1のサイド情報は、具体的には、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の相関パラメータでもよく、例えば、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間のエネルギー比率パラメータでもよい。したがって、デコーダは、相関パラメータ又はエネルギー比率パラメータを使用することにより、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元する。

20

【0033】

可能な実現方式では、当該方法は、

第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するステップを更に含み、

対応して、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得することは、

整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得することを含む。

30

【0034】

対応して、第1のサイド情報は、整列された第1の仮想スピーカー信号と整列された第2の仮想スピーカー信号との間の関係を示す。

40

【0035】

上記の解決方法では、ダウンミキシングされた信号を生成する前に、まず、エンコーダは、仮想スピーカー信号に対して整列動作を実行し、整列動作を完了した後に、ダウンミキシングされた信号及び第1のサイド情報を生成できる。この出願のこの実施形態では、第1の仮想スピーカー信号及び第2の仮想スピーカー信号のサウンドチャンネルを再び調整して整列させることにより、チャンネル間相関が強化されて、コアエンコーダによる第1の仮想スピーカー信号の符号化処理を容易にする。

【0036】

可能な実現方式では、第1のシーンオーディオ信号に基づいて仮想スピーカーセットが

50

ら第2のターゲット仮想スピーカーを選択する前に、当該方法は、

第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要があるか否かを決定するステップと、

第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要がある場合にのみ、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するステップと

を更に含む。

【0037】

上記の解決策では、エンコーダは、信号を更に選択して、第2のターゲット仮想スピーカーが取得される必要があるか否かを決定できる。第2のターゲット仮想スピーカーが取得される必要があるとき、エンコーダは第2の仮想スピーカー信号を生成してもよい。第2のターゲット仮想スピーカーが取得される必要がないとき、エンコーダは第2の仮想スピーカー信号を生成しなくてもよい。エンコーダは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、第1のターゲット仮想スピーカーに加えて他のターゲット仮想スピーカーが選択される必要があるか否かを決定できる。例えば、符号化率が予め設定された閾値よりも高い場合、2つの主要音場成分に対応するターゲット仮想スピーカーが取得される必要があると決定され、第1のターゲット仮想スピーカーが決定されることに加えて、第2のターゲット仮想スピーカーが更に決定されてもよい。他の例では、第1のシーンオーディオ信号の信号クラス情報に基づいて、支配的な音源方向を含む2つの主要音場成分に対応するターゲット仮想スピーカーが取得される必要があると決定された場合、第1のターゲット仮想スピーカーが決定されることに加えて、第2のターゲット仮想スピーカーが更に決定されてもよい。逆に、第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、1つのターゲット仮想スピーカーのみが取得される必要があると決定された場合、第1のターゲット仮想スピーカーが決定された後に、第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得されないと決定される。この出願のこの実施形態では、信号が選択され、それにより、エンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。

【0038】

可能な実現方式では、残差信号は、少なくとも2つのサウンドチャンネル上の残差サブ信号を含み、当該方法は、

オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、少なくとも2つのサウンドチャンネル上の残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定するステップを更に含む、

対応して、第1の仮想スピーカー信号及び残差信号を符号化することは、

第1の仮想スピーカー信号と、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号とを符号化することを含む。

【0039】

上記の解決策では、エンコーダは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて残差信号に対する決定を行うことができる。例えば、残差信号が少なくとも2つのサウンドチャンネル上の残差サブ信号を含む場合、エンコーダは、残差サブ信号が符号化される必要があるサウンドチャンネル又は複数のサウンドチャンネルと、残差サブ信号が符号化される必要がないサウンドチャンネル又は複数のサウンドチャンネルとを選択できる。例えば、残差信号において支配的なエネルギーを有する残差サブ信号は、符号化するためにオーディオエンコーダの構成情報に基づいて選択される。他の例では、残差信号における低次H0Aサウンドチャンネルによる計算を通じて取得された残差サブ信号は、符号化するために第1のシーンオーディオ信号の信号クラス情報に基づいて選択される。残差信号についてサウンドチャンネルが選択され、それにより、エ

10

20

30

40

50

ンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。

【0040】

可能な実現方式では、少なくとも2つのサウンドチャンネル上の残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、当該方法は、

第2のサイド情報を取得するステップであり、第2のサイド情報は、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号と、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号との間の関係を示す、ステップと、

第2のサイド情報をビットストリームに書き込むステップとを更に含む。

【0041】

上記の解決策では、信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。この出願のこの実施形態では、符号化される必要がある残差サブ信号が符号化され、符号化される必要がない残差サブ信号が符号化されず、それにより、エンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。エンコーダが信号を選択するときに情報ロスが発生するので、伝送されない残差サブ信号に対して信号補償が実行される必要がある。信号補償は、情報ロス分析、エネルギー補償、エンベロープ補償及びノイズ補償でもよく、これらに限定されない。補償方法は、線形補償、非線形補償等でもよい。信号補償の後に、第2のサイド情報が生成されてもよく、第2のサイド情報がビットストリームに書き込まれてもよい。第2のサイド情報は、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間の関係を示す。当該関係は複数の実現方式を有する。例えば、第2のサイド情報は信号情報ロス分析パラメータを含み、それにより、デコーダが信号情報ロス分析パラメータを使用することにより、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号とを復元するようにする。他の例では、第2のサイド情報は、具体的には、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間の相関パラメータでもよく、例えば、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間のエネルギー比率パラメータでもよい。したがって、デコーダは、相関パラメータ又はエネルギー比率パラメータを使用することにより、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号とを復元する。この出願のこの実施形態では、デコーダは、ビットストリームを使用することにより第2のサイド情報を取得でき、デコーダは、第2のサイド情報に基づいて信号補償を実行して、デコーダの復号信号の品質を改善できる。

【0042】

第2の態様によれば、この出願の実施形態は、オーディオ復号方法を更に提供し、ビットストリームを受信するステップと、ビットストリームを復号して、仮想スピーカー信号及び残差信号を取得するステップと、ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップとを含む。

【0043】

この出願の実施形態では、まず、ビットストリームが受信され、次いで、ビットストリームが復号されて仮想スピーカー信号及び残差信号を取得し、最後に、再構成されたシーンオーディオ信号は、ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて取得される。この出願のこの実施形態では、オーディオデコーダは、オーディオエンコーダによる符号化プロセスとは逆の復号プロセスを実行し、復号を通じてビットストリームから仮想スピーカー信号及び残差信号を取得し、ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号を使用することにより、再構成された

10

20

30

40

50

シーンオーディオ信号を取得できる。この出願のこの実施形態では、取得されたビットストリームは、仮想スピーカー信号及び残差信号を搬送し、復号されるデータの量を低減し、復号効率を改善する。

【0044】

可能な実現方式では、当該方法は、

ビットストリームを復号して、ターゲット仮想スピーカーの属性情報を取得するステップを更に含む。

【0045】

上記の解決策では、仮想スピーカーを符号化することに加えて、エンコーダはまた、ターゲット仮想スピーカーの属性情報を符号化し、ターゲット仮想スピーカーの符号化された属性情報をビットストリームに書き込むことができる。例えば、第1のターゲット仮想スピーカーの属性情報は、ビットストリームを使用することにより取得できる。この出願のこの実施形態では、ビットストリームは、第1のターゲット仮想スピーカーの符号化された属性情報を搬送でき、それにより、デコーダがビットストリームを復号することにより第1のターゲット仮想スピーカーの属性情報を決定して、デコーダによるオーディオ復号を容易にできるようにする。

【0046】

可能な実現方式では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーについての高次アンビソニクス(HOA)係数を含み、

ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得するステップと、

残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するステップとを含む。

【0047】

上記の解決策では、まず、デコーダは、ターゲット仮想スピーカーについてのHOA係数を決定する。例えば、デコーダは、ターゲット仮想スピーカーについてのHOA係数を予め記憶してもよい。仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数を取得した後に、デコーダは、仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に基づいて合成されたシーンオーディオ信号を取得できる。最後に、合成されたシーンオーディオ信号を調整するために残差信号が使用されて、再構成されたシーンオーディオ信号の品質を改善する。

【0048】

可能な実現方式では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーの位置情報を含み、

ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

ターゲット仮想スピーカーの位置情報に基づいてターゲット仮想スピーカーについてのHOA係数を決定するステップと、

仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得するステップと、

残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するステップと

を含む。

【0049】

上記の解決策では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーの位置情報を含んでもよい。デコーダは、仮想スピーカーセット内の各仮想スピーカーについてのHOA係数を予め記憶し、デコーダは、各仮想スピーカーの位置情報を更に記憶

10

20

30

40

50

する。例えば、デコーダは、仮想スピーカークの位置情報と仮想スピーカークについてのHOA係数との間の対応関係に基づいて、ターゲット仮想スピーカークの位置情報についてのHOA係数を決定でき、或いは、デコーダは、ターゲット仮想スピーカークの位置情報に基づいてターゲット仮想スピーカークについてのHOA係数を計算できる。したがって、デコーダは、ターゲット仮想スピーカークの位置情報に基づいてターゲット仮想スピーカークについてのHOA係数を決定できる。これは、デコーダがターゲット仮想スピーカークについてのHOA係数を決定する必要があるという問題を解決する。

【0050】

可能な実現方式では、仮想スピーカーク信号は、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号をダウンミキシングすることにより取得されたダウンミキシングされた信号であり、当該方法は、

10

ビットストリームを復号して、第1のサイド情報を取得するステップであり、第1のサイド情報は、第1の仮想スピーカーク信号と第2の仮想スピーカーク信号との間の関係を示す、ステップと、

第1のサイド情報及びダウンミキシングされた信号に基づいて第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号を取得するステップと

を更に含み、

対応して、ターゲット仮想スピーカークの属性情報、残差信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

ターゲット仮想スピーカークの属性情報、残差信号、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するステップを含む。

20

【0051】

上記の解決策では、エンコーダは、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号に基づいてダウンミキシングを実行するとき、ダウンミキシングされた信号を生成し、エンコーダは、ダウンミキシングされた信号に対して信号補償を更に実行して、第1のサイド情報を生成できる。第1のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第1のサイド情報を取得できる。デコーダは、第1のサイド情報に基づいて信号補償を実行して、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号を取得できる。したがって、信号再構成中に、第1の仮想スピーカーク信号、第2の仮想スピーカーク信号、ターゲット仮想スピーカークの属性情報及び残差信号が使用されて、デコーダの復号信号の品質を改善できる。

30

【0052】

可能な実現方式では、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該方法は、

ビットストリームを復号して、第2のサイド情報を取得するステップであり、第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示す、ステップと、

第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第2のサウンドチャンネル上の残差サブ信号を取得するステップと

40

を更に含み、

対応して、ターゲット仮想スピーカークの属性情報、残差信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

ターゲット仮想スピーカークの属性情報、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するステップを含む。

【0053】

上記の解決策では、信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。エンコーダが信号を選択するときに情報ロスが発生するので、エンコーダは第2のサイド情報を生成する。第2

50

のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第2のサイド情報を取得できる。ビットストリームで搬送される残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、デコーダは、第2のサイド情報に基づいて信号補償を実行して、第2のサウンドチャンネル上の残差サブ信号を取得できる。例えば、デコーダは、第1のサウンドチャンネル上の残差サブ信号及び第2のサイド情報を使用することにより、第2のサウンドチャンネル上の残差サブ信号を復元する。第2のサウンドチャンネルは、第1のサウンドチャンネルから独立している。したがって、信号再構成中に、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号、ターゲット仮想スピーカーの属性情報及び仮想スピーカー信号が使用されて、デコーダの復号信号の品質を改善できる。

10

【0054】

可能な実現方式では、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含み、当該方法は、

ビットストリームを復号して、第2のサイド情報を取得するステップであり、第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示す、ステップと、

第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を取得するステップと

を更に含み、

20

対応して、ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップは、

ターゲット仮想スピーカーの属性情報、第1のサウンドチャンネル上の更新された残差サブ信号、第3のサウンドチャンネル上の残差サブ信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するステップを含む。

【0055】

上記の解決策では、信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。エンコーダが信号を選択するときに情報ロスが発生するので、エンコーダは第2のサイド情報を生成する。第2のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第2のサイド情報を取得できる。ビットストリームで搬送される残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、デコーダは、第2のサイド情報に基づいて信号補償を実行して、第2のサウンドチャンネル上の残差サブ信号を取得できる。第3のサウンドチャンネル上の残差サブ信号は、第1のサウンドチャンネル上の残差サブ信号とは異なる。第3のサウンドチャンネル上の残差サブ信号が第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて取得されるとき、第1のサウンドチャンネル上の残差サブ信号は、第1のサウンドチャンネル上の更新された残差サブ信号を取得するために更新される必要がある。例えば、デコーダは、第1のサウンドチャンネル上の残差サブ信号及び第2のサイド情報を使用することにより、第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を生成する。したがって、信号再構成中に、第3のサウンドチャンネル上の残差サブ信号、第1のサウンドチャンネル上の更新された残差サブ信号、ターゲット仮想スピーカーの属性情報及び仮想スピーカー信号が使用されて、デコーダの復号信号の品質を改善できる。

30

40

【0056】

第3の態様によれば、この出願の実施形態は、オーディオ符号化装置を提供し、

第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカースセットから第1のターゲット仮想スピーカーを選択するように構成された取得モジュールと、

第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて仮想スピーカー信号を生成するように構成された信号生成モジュールであり、

信号生成モジュールは、第1のターゲット仮想スピーカーの属性情報及び第1の仮想スピー

50

ーカー信号を使用することにより、第2のシーンオーディオ信号を取得するように構成され、

信号生成モジュールは、第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて残差信号を生成するように構成される、信号生成モジュールと、

仮想スピーカー信号及び残差信号を符号化して、ビットストリームを取得するように構成された符号化モジュールと

を含む。

【0057】

可能な実現方式では、取得モジュールは、仮想スピーカーセットに基づいて第1のシーンオーディオ信号から主要音場成分を取得し、主要音場成分に基づいて仮想スピーカーセ

10

【0058】

可能な実現方式では、取得モジュールは、主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから主要音場成分についてのHOA係数を選択するように構成され、ここで、HOA係数セット内のHOA係数は、仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にあり、仮想スピーカーセットの中で主要音場成分についてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するように構成される。

【0059】

可能な実現方式では、取得モジュールは、主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを取得し、第1のターゲット仮想スピーカーの構成パラメータに基づいて第1のターゲット仮想スピーカーについてのHOA係数を生成し、仮想スピーカーセットの中で第1のターゲット仮想スピーカーについてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するように構成される。

20

【0060】

可能な実現方式では、取得モジュールは、オーディオエンコーダの構成情報に基づいて仮想スピーカーセット内の複数の仮想スピーカーの構成パラメータを決定し、主要音場成分に基づいて複数の仮想スピーカーの構成パラメータから第1のターゲット仮想スピーカーの構成パラメータを選択するように構成される。

【0061】

可能な実現方式では、第1のターゲット仮想スピーカーの構成パラメータは、第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含む。

30

【0062】

取得モジュールは、第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報に基づいて第1のターゲット仮想スピーカーについてのHOA係数を決定するように構成される。

【0063】

可能な実現方式では、符号化モジュールは、第1のターゲット仮想スピーカーの属性情報を符号化し、符号化された情報をビットストリームに書き込むように更に構成される。

【0064】

可能な実現方式では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカーの属性情報は、第1のターゲット仮想スピーカーについてのHOA係数を含む。

40

【0065】

信号生成モジュールは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカーについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカー信号を取得するように構成される。

【0066】

可能な実現方式では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカーの属性情報は、第1のターゲ

50

ット仮想スピーカーの位置情報を含む。

【0067】

信号生成モジュールは、第1のターゲット仮想スピーカーの位置情報に基づいて第1のターゲット仮想スピーカーについてのHOA係数を取得し、符号化されるべきHOA信号及び第1のターゲット仮想スピーカーについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカー信号を取得するように構成される。

【0068】

可能な実現方式では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するように構成される。

【0069】

信号生成モジュールは、第1のシーンオーディオ信号及び第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するように構成される。

【0070】

符号化モジュールは、第2の仮想スピーカー信号を符号化し、符号化された信号をビットストリームに書き込むように構成される。

【0071】

対応して、信号生成モジュールは、第1のターゲット仮想スピーカーの属性情報、第1の仮想スピーカー信号、第2のターゲット仮想スピーカーの属性情報及び第2の仮想スピーカー信号に基づいて第2のシーンオーディオ信号を取得するように構成される。

【0072】

可能な実現方式では、信号生成モジュールは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するように構成される。

【0073】

対応して、符号化モジュールは、整列された第2の仮想スピーカー信号を符号化するように構成される。

【0074】

対応して、符号化モジュールは、整列された第1の仮想スピーカー信号及び残差信号を符号化するように構成される。

【0075】

可能な実現方式では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するように構成される。

【0076】

信号生成モジュールは、第1のシーンオーディオ信号及び第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成するように構成される。

【0077】

対応して、符号化モジュールは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得するように構成される。第1のサイド情報は第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示す。

【0078】

対応して、符号化モジュールは、ダウンミキシングされた信号、第1のサイド情報及び残差信号を符号化するように構成される。

【0079】

可能な実現方式では、信号生成モジュールは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得するように構成される。

【0080】

符号化モジュールは、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得するよ

10

20

30

40

50

うに構成される。

【0081】

対応して、第1のサイド情報は、整列された第1の仮想スピーカー信号と整列された第2の仮想スピーカー信号との間の関係を示す。

【0082】

可能な実現方式では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択する前に、第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要があるか否かを決定し、第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要がある場合にのみ、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択するように構成される。

10

【0083】

可能な実現方式では、残差信号は、少なくとも2つのサウンドチャンネル上の残差サブ信号を含む。

【0084】

信号生成モジュールは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、少なくとも2つのサウンドチャンネル上の残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定するように構成される。

20

【0085】

対応して、符号化モジュールは、第1の仮想スピーカー信号と、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号とを符号化するように構成される。

【0086】

可能な実現方式では、取得モジュールは、少なくとも2つのサウンドチャンネル上の残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、第2のサイド情報を取得するように構成される。第2のサイド情報は、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号と、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号との間の関係を示す。

30

【0087】

対応して、符号化モジュールは、第2のサイド情報をビットストリームに書き込むように構成される。

【0088】

この出願の第3の態様では、オーディオ符号化装置の構成モジュールは、第1の態様及び可能な実現方式に記載されるステップを更に実行してもよい。詳細については、第1の態様及び可能な実現方式における説明を参照する。

【0089】

第4の態様によれば、この出願の実施形態は、オーディオ復号装置を更に提供し、

40

ビットストリームを受信するように構成された受信モジュールと、

ビットストリームを復号して、仮想スピーカー信号及び残差信号を取得するように構成された復号モジュールと、

ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するように構成された再構成モジュールとを含む。

【0090】

可能な実現方式では、復号モジュールは、ビットストリームを復号して、ターゲット仮想スピーカーの属性情報を取得するように更に構成される。

【0091】

50

可能な実現方式では、ターゲット仮想スピーカークの属性情報は、ターゲット仮想スピーカークについての高次アンビソニックス(HOA)係数を含む。

【0092】

再構成モジュールは、仮想スピーカーク信号及びターゲット仮想スピーカークについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するように構成される。

【0093】

可能な実現方式では、ターゲット仮想スピーカークの属性情報は、ターゲット仮想スピーカークの位置情報を含む。

【0094】

再構成モジュールは、ターゲット仮想スピーカークの位置情報に基づいてターゲット仮想スピーカークについてのHOA係数を決定し、仮想スピーカーク信号及びターゲット仮想スピーカークについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するように構成される。

【0095】

可能な実現方式では、仮想スピーカーク信号は、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号をダウンミキシングすることにより取得されたダウンミキシングされた信号である。当該装置は第1の信号補償モジュールを更に含む。

【0096】

復号モジュールは、ビットストリームを復号して、第1のサイド情報を取得するように構成される。第1のサイド情報は、第1の仮想スピーカーク信号と第2の仮想スピーカーク信号との間の関係を示す。

【0097】

第1の信号補償モジュールは、第1のサイド情報及びダウンミキシングされた信号に基づいて第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号を取得するように構成される。

【0098】

対応して、再構成モジュールは、ターゲット仮想スピーカークの属性情報、残差信号、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0099】

可能な実現方式では、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含む。当該装置は第2の信号補償モジュールを更に含む。

【0100】

復号モジュールは、ビットストリームを復号して、第2のサイド情報を取得するように構成される。第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示す。

【0101】

第2の信号補償モジュールは、第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第2のサウンドチャンネル上の残差サブ信号を取得するように構成される。

【0102】

対応して、再構成モジュールは、ターゲット仮想スピーカークの属性情報、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0103】

可能な実現方式では、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含む。当該装置は第3の信号補償モジュールを更に含む。

【0104】

10

20

30

40

50

復号モジュールは、ビットストリームを復号して、第2のサイド情報を取得するように構成される。第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示す。

【0105】

第3の信号補償モジュールは、第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を取得するように構成される。

【0106】

対応して、再構成モジュールは、ターゲット仮想スピーカ属性情報、第1のサウンドチャンネル上の更新された残差サブ信号、第3のサウンドチャンネル上の残差サブ信号及び仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

10

【0107】

この出願の第4の態様では、オーディオ復号装置の構成モジュールは、第2の態様及び可能な実現方式に記載されるステップを更に実行してもよい。詳細については、第2の態様及び可能な実現方式における説明を参照する。

【0108】

第5の態様によれば、この出願の実施形態は、コンピュータ読み取り可能記憶媒体を提供する。コンピュータ読み取り可能記憶媒体は命令を記憶する。命令がコンピュータ上で実行されたとき、コンピュータは第1の態様又は第2の態様による方法を実行することが可能になる。

20

【0109】

第6の態様によれば、この出願の実施形態は、命令を含むコンピュータプログラム製品を提供する。コンピュータプログラム製品がコンピュータ上で実行したとき、コンピュータは第1の態様又は第2の態様による方法を実行することが可能になる。

【0110】

第7の態様によれば、この出願の実施形態は通信装置を提供する。通信装置は、端末デバイス又はチップのようなエンティティを含んでもよい。通信装置はプロセッサを含む。任意選択で、通信装置はメモリを更に含む。メモリは命令を記憶するように構成される。プロセッサはメモリ内の命令を実行するように構成され、それにより、通信装置は第1の態様又は第2の態様のいずれか1つによる方法を実行するようにする。

30

【0111】

第8の態様によれば、この出願はチップシステムを提供する。チップシステムはプロセッサを含み、上記の態様における機能を実現する際に、オーディオ符号化装置又はオーディオ復号装置をサポートするように構成され、例えば、上記の方法においてデータ及び/又は情報を送信又は処理する。可能な設計では、チップシステムはメモリを更に含み、メモリはオーディオ符号化装置又はオーディオ復号装置に必要なプログラム命令及びデータを記憶するように構成される。チップシステムはチップを含んでもよく、或いは、チップ及び他のディスクリットデバイスを含んでもよい。

【0112】

第9の態様によれば、この出願は、コンピュータ読み取り可能記憶媒体を提供し、第1の態様のいずれか1つによる方法において生成されたビットストリームを含む。

40

【図面の簡単な説明】

【0113】

【図1】 この出願の実施形態によるオーディオ処理システムの構成構造の概略図である。

【図2a】 この出願の実施形態による、オーディオエンコーダ及びオーディオデコーダが使用される端末デバイスの概略図である。

【図2b】 この出願の実施形態による、オーディオエンコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。

【図2c】 この出願の実施形態による、オーディオデコーダが使用される無線デバイス又

50

はコアネットワークデバイスの概略図である。

【図3a】この出願の実施形態による、マルチチャネルエンコーダ及びマルチチャネルデコーダが使用される端末デバイスの概略図である。

【図3b】この出願の実施形態による、マルチチャネルオーディオエンコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。

【図3c】この出願の実施形態による、マルチチャネルオーディオデコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。

【図4】この出願の実施形態によるオーディオ符号化装置とオーディオ復号装置との間の相互作用の概略フローチャートである。

【図5】この出願の実施形態によるエンコーダの構造の概略図である。

10

【図6】この出願の実施形態によるデコーダの構造の概略図である。

【図7】この出願の実施形態による他のエンコーダの構造の概略図である。

【図8】この出願の実施形態による球上にほぼ均等に分布した仮想スピーカースの概略図である。

【図9】この出願の実施形態による他のエンコーダの構造の概略図である。

【図10】この出願の実施形態によるオーディオ符号化装置の構成構造の概略図である。

【図11】この出願の実施形態によるオーディオ復号装置の構成構造の概略図である。

【図12】この出願の実施形態による他のオーディオ符号化装置の構成構造の概略図である。

【図13】この出願の実施形態による他のオーディオ復号装置の構成構造の概略図である。

20

【発明を実施するための形態】

【0114】

この出願の実施形態は、符号化及び復号されるデータの量を低減し、符号化及び復号効率を改善するためのオーディオ符号化及び復号方法並びに装置を提供する。

【0115】

以下に、添付図面を参照してこの出願の実施形態について説明する。

【0116】

この出願の明細書、特許請求の範囲及び添付図面において、「第1」、「第2」等の用語は、同様の対象を区別することを意図しているが、必ずしも特定の順序又は系列を示すとは限らない。このように使用される用語は、適切な状況において相互に交換可能であり、これは、この出願の実施形態において同じ属性を有する対象を記述するとき使用される単なる識別方式であることが理解されるべきである。さらに、「含む(include)」、「含む(contain)」という用語及びいずれかの他の変形は、非排他的包含をカバーすることを意味しており、それにより、一連のユニットを含むプロセス、方法、システム、製品又はデバイスは、必ずしもこれらのユニットに限定されず、このようなプロセス、方法、システム、製品又はデバイスに明示的に列挙されていないか或いは固有である他のユニットを含んでもよい。

30

【0117】

この出願の実施形態における技術的解決策は、様々なオーディオ処理システムに適用されてもよい。図1は、この出願の実施形態によるオーディオ処理システムの構成構造の概略図である。オーディオ処理システム100は、オーディオ符号化装置101及びオーディオ復号装置102を含んでもよい。オーディオ符号化装置101は、ビットストリームを生成するように構成されてもよく、次いで、オーディオ符号化されたビットストリームは、オーディオ伝送チャネルを通じてオーディオ復号装置102に伝送されてもよい。オーディオ復号装置102は、ビットストリームを受信し、次いで、オーディオ復号装置102のオーディオ復号機能を実行して、最終的に再構成された信号を取得してもよい。

40

【0118】

この出願の実施形態では、オーディオ符号化装置は、オーディオ通信を必要とする様々な端末デバイス、並びにトランスコーディングを必要とする無線デバイス及びコアネットワークデバイスで使用されてもよい。例えば、オーディオ符号化装置は、上記の端末デバ

50

イス、無線デバイス又はコアネットワークデバイスのオーディオエンコーダでもよい。同様に、オーディオ復号装置は、オーディオ通信を必要とする様々な端末デバイス、並びにトランスコーディングを必要とする無線デバイス及びコアネットワークデバイスで使用されてもよい。例えば、オーディオ復号装置は、上記の端末デバイス、無線デバイス又はコアネットワークデバイスのオーディオデコーダでもよい。例えば、オーディオエンコーダは、無線アクセスネットワーク、コアネットワークのメディアゲートウェイ、トランスコーディングデバイス、メディアリソースサーバ、モバイル端末及び固定ネットワーク端末を含んでもよい。さらに、オーディオエンコーダは、仮想現実(virtual reality, VR)ストリーミング(streaming)メディアサービスに適用されるオーディオコーデックでもよい。

【0119】

この出願のこの実施形態では、仮想現実ストリーミング(VR streaming)メディアサービスに適用可能なオーディオ符号化及び復号モジュール(audio encoding and audio decoding)が例として使用される。エンドツーエンドのオーディオ信号処理手順は以下を含む。オーディオ信号Aが獲得モジュール(acquisition)を通過した後に、オーディオ信号Aに対して前処理操作(audio preprocessing)を実行し、ここで、前処理操作は信号の低周波数部分をフィルタリング除去することを含み、20Hz又は50Hzを境界点として使用することにより信号から方向情報を抽出することでもよく、次いで、符号化(audio encoding)及びカプセル化(file/segment encapsulation)を実行し、次いで、カプセル化された信号をデコーダに送信(delivery)し、ここで、デコーダはまずカプセル化解除(file/segment decapsulation)を実行し、次いで復号(audio decoding)を実行し、復号された信号に対してバイノーラルレンダリング(audio rendering)を実行し、レンダリングされた信号をリスナーのヘッドセット(headphones)にマッピングし、ヘッドセットは独立したヘッドセットでもよく、或いは、メガネデバイス上のヘッドセットでよい。

【0120】

図2aは、この出願の実施形態による、オーディオエンコーダ及びオーディオデコーダが使用される端末デバイスの概略図である。各端末デバイスは、オーディオエンコーダと、チャンネルエンコーダと、オーディオデコーダと、チャンネルデコーダとを含んでもよい。具体的には、チャンネルエンコーダは、オーディオ信号に対してチャンネル符号化を実行するように構成され、チャンネルデコーダは、オーディオ信号に対してチャンネル復号を実行するように構成される。例えば、第1の端末デバイス20は、第1のオーディオエンコーダ201と、第1のチャンネルエンコーダ202と、第1のオーディオデコーダ203と、第1のチャンネルデコーダ204とを含んでもよい。第2の端末デバイス21は、第2のオーディオデコーダ211と、第2のチャンネルデコーダ212と、第2のオーディオエンコーダ213と、第2のチャンネルエンコーダ214とを含んでもよい。第1の端末デバイス20は無線又は有線の第1のネットワーク通信デバイス22に接続され、第1のネットワーク通信デバイス22はデジタルチャンネルを通じて無線又は有線の第2のネットワーク通信デバイス23に接続され、第2の端末デバイス21は無線又は有線の第2のネットワーク通信デバイス23に接続される。無線又は有線のネットワーク通信デバイスは、一般的に信号伝送デバイス、例えば、通信基地局又はデータ交換デバイスでもよい。

【0121】

オーディオ通信では、送信機として機能する端末デバイスは、まずオーディオ獲得を実行し、獲得されたオーディオ信号に対してオーディオ符号化を実行し、次いで、チャンネル符号化を実行し、無線ネットワーク又はコアネットワークを使用することによりデジタルチャンネル上で符号化されたオーディオ信号を送信する。受信機として機能する端末デバイスは、受信した信号に基づいてチャンネル復号を実行して、ビットストリームを取得し、次いで、オーディオ復号を通じてオーディオ信号を復元する。受信機として機能する端末デバイスはオーディオ再生を実行する。

【0122】

図2bは、この出願の実施形態による、オーディオエンコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。無線デバイス又はコアネットワークデ

10

20

30

40

50

バイス25は、チャンネルデコーダ251と、他のオーディオデコーダ252と、この出願のこの実施形態において提供されるオーディオエンコーダ253と、チャンネルエンコーダ254とを含む。他のオーディオデコーダ252は、オーディオデコーダ以外のオーディオデコーダである。無線デバイス又はコアネットワークデバイス25では、チャンネルデコーダ252が、まずデバイスに入る信号に対してチャンネル復号を実行し、次いで、他のオーディオデコーダ252がオーディオ復号を実行し、次いで、この出願の実施形態において提供されるオーディオエンコーダ253がオーディオ符号化を実行し、最後にチャンネルエンコーダ254がオーディオ信号に対してチャンネル符号化を実行する。チャンネル符号化が完了すると、チャンネル符号化されたオーディオ信号が伝送される。他のオーディオデコーダ252は、チャンネルデコーダ251により復号されたビットストリームに対してオーディオ復号を実行する。

10

【0123】

図2cは、この出願の実施形態による、オーディオデコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。無線デバイス又はコアネットワークデバイス25は、チャンネルデコーダ251と、この出願のこの実施形態において提供されるオーディオデコーダ255と、他のオーディオエンコーダ256と、チャンネルエンコーダ254とを含む。他のオーディオエンコーダ256は、オーディオエンコーダ以外のオーディオエンコーダである。無線デバイス又はコアネットワークデバイス25では、チャンネルデコーダ251が、まずデバイスに入る信号に対してチャンネル復号を実行し、次いで、オーディオデコーダ255が受信したオーディオ符号化されたビットストリームを復号し、次いで、他のオーディオエンコーダ256がオーディオ符号化を実行し、最後にチャンネルエンコーダ254がオーディオ信号に対してチャンネル符号化を実行する。チャンネル符号化が完了した後に、チャンネル符号化されたオーディオ信号が伝送される。無線デバイス又はコアネットワークデバイスでは、トランスコーディングが実現される必要がある場合、対応するオーディオ符号化及び復号処理が実行される必要がある。無線デバイスは通信における無線周波数関連デバイスであり、コアネットワークデバイスは通信におけるコアネットワーク関連デバイスである。

20

【0124】

この出願のいくつかの実施形態では、オーディオ符号化装置は、オーディオ通信を必要とする様々な端末デバイス、並びにトランスコーディングを必要とする無線デバイス及びコアネットワークデバイスで使用されてもよい。例えば、オーディオ符号化装置は、上記の端末デバイス、無線デバイス又はコアネットワークデバイスのマルチチャンネルエンコーダでもよい。同様に、オーディオ復号装置は、オーディオ通信を必要とする様々な端末デバイス、並びにトランスコーディングを必要とする無線デバイス及びコアネットワークデバイスで使用されてもよい。例えば、オーディオ復号装置は、上記の端末デバイス、無線デバイス又はコアネットワークデバイスのマルチチャンネルデコーダでもよい。

30

【0125】

図3aは、この出願の実施形態による、マルチチャンネルエンコーダ及びマルチチャンネルデコーダが使用される端末デバイスの概略図である。各端末デバイスは、マルチチャンネルエンコーダと、チャンネルエンコーダと、マルチチャンネルデコーダと、チャンネルデコーダとを含んでもよい。マルチチャンネルエンコーダは、この出願の実施形態において提供されるオーディオ符号化方法を実行してもよく、マルチチャンネルデコーダは、この出願の実施形態において提供されるオーディオ復号方法を実行してもよい。具体的には、チャンネルエンコーダは、マルチチャンネル信号に対してチャンネル符号化を実行するために使用され、チャンネルデコーダは、マルチチャンネル信号に対してチャンネル復号を実行するために使用される。例えば、第1の端末デバイス30は、第1のマルチチャンネルエンコーダ301と、第1のチャンネルエンコーダ302と、第1のマルチチャンネルデコーダ303と、第1のチャンネルデコーダ304とを含んでもよい。第2の端末デバイス31は、第2のマルチチャンネルデコーダ311と、第2のチャンネルデコーダ312と、第2のマルチチャンネルエンコーダ313と、第2のチャンネルエンコーダ314とを含んでもよい。第1の端末デバイス30は無線又は有線の第1のネットワーク通信デバイス32に接続され、第1のネットワーク通信デバイス32はデジタルチャネ

40

50

ルを通じて無線又は有線の第2のネットワーク通信デバイス33に接続され、第2の端末デバイス31は無線又は有線の第2のネットワーク通信デバイス33に接続される。無線又は有線のネットワーク通信デバイスは、一般的に信号伝送デバイス、例えば、通信基地局又はデータ交換デバイスでもよい。オーディオ通信では、送信機として機能する端末デバイスは、獲得されたマルチチャネル信号に対してマルチチャネル符号化を実行し、次いで、チャネル符号化を実行し、無線ネットワーク又はコアネットワークを使用することによりデジタルチャネル上で符号化されたマルチチャネル信号を伝送する。受信機として機能する端末デバイスは、受信した信号に基づいてチャネル復号を実行して、ビットストリームに符号化されたマルチチャネル信号を取得し、次いで、マルチチャネル復号を通じてマルチチャネル信号を復元する。受信機として機能する端末デバイスは再生を実行する。

10

【0126】

図3bは、この出願の実施形態による、マルチチャネルエンコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。無線デバイス又はコアネットワークデバイス35は、チャネルデコーダ351と、他のオーディオデコーダ352と、マルチチャネルエンコーダ353と、チャネルエンコーダ354とを含む。図3bは図2bと同様であり、詳細はここでは再び説明しない。

【0127】

図3cは、この出願の実施形態による、マルチチャネルデコーダが使用される無線デバイス又はコアネットワークデバイスの概略図である。無線デバイス又はコアネットワークデバイス35は、チャネルデコーダ351と、マルチチャネルデコーダ355と、他のオーディオエンコーダ356と、チャネルエンコーダ354とを含む。図3cは図2cと同様であり、詳細はここでは再び説明しない。

20

【0128】

オーディオ符号化処理はマルチチャネルエンコーダの一部でもよく、オーディオ復号処理はマルチチャネルデコーダの一部でもよい。例えば、獲得されたマルチチャネル信号に対してマルチチャネル符号化を実行することは、獲得されたマルチチャネル信号を処理してオーディオ信号を取得し、次いで、この出願の実施形態において提供される方法に従って、取得されたオーディオ信号を符号化することでもよい。デコーダは、マルチチャネル信号符号化されたビットストリームに基づいて復号してオーディオ信号を取得し、アップミキシングの後にマルチチャネル信号を復元する。したがって、この出願の実施形態はまた、端末デバイス、無線デバイス又はコアネットワークデバイスにおけるマルチチャネルエンコーダ及びマルチチャネルデコーダにも適用されてもよい。無線デバイス又はコアネットワークデバイスでは、トランスコーディングが実現される必要がある場合、対応するマルチチャネル符号化及び復号処理が行される必要がある。

30

【0129】

この出願の実施形態において提供されるオーディオ符号化及び復号方法は、オーディオ符号化方法及びオーディオ復号方法を含んでもよい。オーディオ符号化方法はオーディオ符号化装置により実行され、オーディオ復号方法はオーディオ復号装置により実行される。オーディオ符号化装置及びオーディオ復号装置は、相互に通信してもよい。以下に、上記のシステムアーキテクチャ、オーディオ符号化装置及びオーディオ復号装置に基づいて、この出願の実施形態において提供されるオーディオ符号化方法及びオーディオ復号方法について説明する。図4は、この出願の実施形態によるオーディオ符号化装置とオーディオ復号装置との相互作用の概略フローチャートである。以下のステップ401~403はオーディオ符号化装置(エンコーダと呼ばれる)により実行されてもよく、以下のステップ411~413はオーディオ復号装置(デコーダと呼ばれる)により実行されてもよい。主に以下のプロセスが含まれる。

40

【0130】

401:第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択する。

【0131】

50

エンコーダは第1のシーンオーディオ信号を取得する。第1のシーンオーディオ信号は、空間内のマイクの位置における音場から獲得されたオーディオ信号であり、第1のシーンオーディオ信号はまた、元のシーンにおけるオーディオ信号と呼ばれてもよい。例えば、第1のシーンオーディオ信号は、高次アンビソニックス(higher order ambisonics, HOA)技術を使用することにより取得されたオーディオ信号でもよい。

【0132】

この出願のこの実施形態では、仮想スピーカーセットがエンコーダについて予め構成できる。仮想スピーカーセットは、複数の仮想スピーカーを含んでもよい。実際の再生中には、シーンオーディオ信号がヘッドセットを使用することにより再生されてもよく、或いは、室内に配置された複数のスピーカーを使用することにより再生されてもよい。スピーカーが再生に使用されるとき、基本的な方法は、複数のスピーカーの信号を重ね合わせることであり、それにより、空間内の或る点(リスナーの位置)における音場が、シーンオーディオ信号が記録されるとき標準上の元の音場にできるだけ近づくようにする。この出願のこの実施形態では、仮想スピーカーがシーンオーディオ信号に対応する再生信号を計算するために使用され、再生信号が伝送信号として使用され、圧縮信号が生成される。仮想スピーカーは空間内の音場に存在するスピーカーを仮想的に表し、仮想スピーカーはエンコーダにおいてシーンオーディオ信号の再生を実現できる。

【0133】

この出願の実施形態では、仮想スピーカーセットは複数の仮想スピーカーを含み、複数の仮想スピーカーのそれぞれが仮想スピーカー構成パラメータ(略称、構成パラメータ)に対応する。仮想スピーカー構成パラメータは、仮想スピーカーの数、仮想スピーカーのHOAオーダー及び仮想スピーカーの位置座標のような情報を含むが、これらに限定されない。仮想スピーカーセットを取得した後に、エンコーダは、第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択する。第1のシーンオーディオ信号は、元のシーンにおける符号化対象のオーディオ信号であり、第1のターゲット仮想スピーカーは、仮想スピーカーセット内の仮想スピーカーでもよい。例えば、第1のターゲット仮想スピーカーは、予め構成されたターゲット仮想スピーカー選択ポリシーに従って予め設定された仮想スピーカーセットから選択できる。ターゲット仮想スピーカー選択ポリシーは、仮想スピーカーセットから第1のシーンオーディオ信号に一致するターゲット仮想スピーカーを選択するポリシーであり、例えば、第1のシーンオーディオ信号から各仮想スピーカーにより取得された音場成分に基づいて第1のターゲット仮想スピーカーを選択するポリシーである。他の例では、第1のターゲット仮想スピーカーは、各仮想スピーカーの位置情報に基づいて第1のシーンオーディオ信号から選択される。第1のターゲット仮想スピーカーは、仮想スピーカーセット内にあり且つ第1のシーンオーディオ信号を再生するために使用される仮想スピーカーであり、すなわち、エンコーダは、仮想スピーカーセットから第1のシーンオーディオ信号を再生できるターゲット仮想エンコーダを選択できる。

【0134】

この出願のこの実施形態では、401において第1のターゲット仮想スピーカーが選択された後に、第1のターゲット仮想スピーカーについての後続の処理プロセス、例えば、後続のステップ402~405が実行されてもよい。これは限定されない。この出願の実施形態では、第1のターゲット仮想スピーカーが選択できるだけでなく、より多くのターゲット仮想スピーカーも選択できる。例えば、第2のターゲット仮想スピーカーが選択されてもよい。第2のターゲット仮想スピーカーについても、後続のステップ402~405と同様のプロセスが実行される必要がある。詳細については、以降の実施形態における説明を参照する。

【0135】

この出願の実施形態では、エンコーダが第1のターゲット仮想スピーカーを選択した後に、エンコーダは、第1のターゲット仮想スピーカーの属性情報を更に取得できる。第1のターゲット仮想スピーカーの属性情報は、第1のターゲット仮想スピーカーの属性に関連

10

20

30

40

50

する情報を含む。属性情報は、特定の適用シナリオに依存して設定されてもよい。例えば、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報又は第1のターゲット仮想スピーカークについてのHOA係数を含む。第1のターゲット仮想スピーカークの位置情報は、空間内の第1のターゲット仮想スピーカークの分布位置に関する情報でもよく、或いは、他の仮想スピーカークに対する仮想スピーカークセット内の第1のターゲット仮想スピーカークの位置に関する情報でもよい。ここでは具体的に限定されない。仮想スピーカークセット内の各仮想スピーカークはHOA係数に対応し、HOA係数はまたアンビソニック係数と呼ばれてもよい。以下に、仮想スピーカークについてのHOA係数について説明する。

【0136】

例えば、HOAオーダーはオーダー2～10のうち1つでもよい。オーディオ信号が記録されるとき、信号サンプリングレートは48～192キロヘルツ(kHz)であり、サンプリング深度は16又は24ビット(bit)である。HOA信号は、シーンオーディオ信号及び仮想スピーカークについてのHOA係数に基づいて生成されてもよい。HOA信号は、音場を有する空間に関する情報を特徴とし、HOA信号は空間内の或る点における音場信号の特定の精度を記述する情報である。したがって、他の表現形式が位置点の音場信号を記述するために使用されることが考えられることができる。この記述方法では、空間内の位置点の信号がより少ないデータ量を使用することにより同じ精度で記述されて、信号圧縮の目的を達成できる。空間内の音場は複数の平面波の重ね合わせに分解できる。したがって、理論的には、HOA信号により表現される音場は複数の平面波の重ね合わせを使用することにより表現でき、各平面波は1つのサウンドチャンネル上のオーディオ信号及び方向ベクトルを使用することにより表される。重ね合わされた平面波の表現形式は、より少ないサウンドチャンネルを使用することにより元の音場を正確に表現して、信号圧縮の目的を達成できる。

【0137】

この出願のいくつかの実施形態では、エンコーダにより401を実行することに加えて、この出願のこの実施形態において提供されるオーディオ符号化方法は、以下のステップを更に含む。

A1:仮想スピーカークセットに基づいて第1のシーンオーディオ信号から主要音場成分を取得する。

【0138】

A1における主要音場成分はまた、第1の主要音場成分と呼ばれてもよい。

【0139】

A1が実行されるとき、401において第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカークセットから第1のターゲット仮想スピーカークを選択することは、以下を含む。

B1:主要音場成分に基づいて仮想スピーカークセットから第1のターゲット仮想スピーカークを選択する。

【0140】

エンコーダは、仮想スピーカークセットを取得し、エンコーダは、仮想スピーカークセットを使用することにより第1のシーンオーディオ信号に対して信号分解を実行して、第1のシーンオーディオ信号に対応する主要音場成分を取得する。主要音場成分は、第1のシーンオーディオ信号内の主要音場に対応するオーディオ信号を表す。例えば、仮想スピーカークセットは複数の仮想スピーカークを含み、複数の音場成分は、複数の仮想スピーカークに基づいて第1のシーンオーディオ信号から取得されてもよく、すなわち、各仮想スピーカークは、第1のシーンオーディオ信号から1つの音場成分を取得してもよく、次いで、主要音場成分が複数の音場成分から選択される。例えば、主要音場成分は、複数の音場成分の中で最大値を有する1つ以上の音場成分でもよく、主要音場成分は、代替として、複数の音場成分の中で支配的な方向を有する1つ以上の音場成分でもよい。仮想スピーカークセット内の各仮想スピーカークは、音場成分に対応し、第1のターゲット仮想スピーカークは、主要音場成分に基づいて仮想スピーカークセットから選択される。例えば、主要音場成分に対応する

仮想スピーカーは、エンコーダにより選択された第1のターゲット仮想スピーカーである。この出願のこの実施形態では、エンコーダは、主要音場成分に基づいて第1のターゲット仮想スピーカーを選択して、エンコーダが第1のターゲット仮想スピーカーを決定する必要があるという問題を解決できる。

【0141】

この出願のこの実施形態では、エンコーダは複数の方式で第1のターゲット仮想スピーカーを選択できる。例えば、エンコーダは、指定の位置における仮想スピーカーを第1のターゲット仮想スピーカーとして予め設定してもよく、すなわち、仮想スピーカーセット内の各仮想スピーカーの位置に基づいて、指定の位置を満たす仮想スピーカーを第1のターゲット仮想スピーカーとして選択してもよい。これは限定されない。

10

【0142】

この出願のいくつかの実施形態では、B1において主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択することは、以下を含む。

主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから主要音場成分についてのHOA係数を選択し、ここで、HOA係数セット内のHOA係数は、仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にあり、

仮想スピーカーセットの中で主要音場成分についてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定する。

【0143】

エンコーダは仮想スピーカーセットに基づいてHOA係数セットを予め構成し、HOA係数セット内のHOA係数と仮想スピーカーセット内の仮想スピーカーとの間に1対1の対応関係が存在する。したがって、HOA係数が主要音場成分に基づいて選択された後に、1対1の対応関係に基づいて、主要音場成分についてのHOA係数に対応するターゲット仮想スピーカーを求めて仮想スピーカーセットが検索され、見つかったターゲット仮想スピーカーが第1のターゲット仮想スピーカーである。これは、エンコーダが第1のターゲット仮想スピーカーを決定する必要があるという問題を解決する。例えば、HOA係数セットはHOA係数1、HOA係数2及びHOA係数3を含み、仮想スピーカーセットは仮想スピーカー1、仮想スピーカー2及び仮想スピーカー3を含む。HOA係数セット内のHOA係数は、仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にある。例えば、HOA係数1は仮想スピーカー1に対応し、HOA係数2は仮想スピーカー2に対応し、HOA係数3は仮想スピーカー3に対応する。HOA係数3が主要音場成分に基づいてHOA係数セットから選択された場合、第1のターゲット仮想スピーカーが仮想スピーカー3であると決定できる。

20

30

【0144】

この出願のいくつかの実施形態では、B1において主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択することは、以下を更に含む。

C1:主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを取得する。

C2:第1のターゲット仮想スピーカーの構成パラメータに基づいて第1のターゲット仮想スピーカーについてのHOA係数を生成する。

C3:仮想スピーカーセットの中で第1のターゲット仮想スピーカーについてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定する。

40

【0145】

主要音場成分を取得した後に、エンコーダは主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを決定できる。例えば、主要音場成分は、複数の音場成分の中で最も大きい値を有する1つ以上の音場成分であるか、或いは、主要音場成分は、複数の音場成分の中で支配的な方向を有する1つ以上の音場成分でもよい。主要音場成分は、第1のシーンオーディオ信号に一致する第1のターゲット仮想スピーカーを決定するために使用でき、対応する属性情報は、第1のターゲット仮想スピーカーについて構成され、第1のターゲット仮想スピーカーについてのHOA係数は、第1のターゲット仮想スピーカーの設定構成パラメータに基づいて生成できる。HOA係数を生成するプロセスは、HOAア

50

ルゴリズムを使用することにより実現でき、詳細はここでは再び説明しない。仮想スピーカーセット内の各仮想スピーカーは、HOA係数に対応する。したがって、第1のターゲット仮想スピーカーは、各仮想スピーカーについてのHOA係数に基づいて仮想スピーカーセットから選択され、エンコーダが第1のターゲット仮想スピーカーを決定する必要があるという問題を解決できる。

【0146】

この出願のいくつかの実施形態では、C1において主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを取得することは、以下を含む。

オーディオエンコーダの構成情報に基づいて仮想スピーカーセット内の複数の仮想スピーカーの構成パラメータを決定し、

主要音場成分に基づいて複数の仮想スピーカーの構成パラメータから第1のターゲット仮想スピーカーの構成パラメータを選択する。

【0147】

オーディオエンコーダは、複数の仮想スピーカーの構成パラメータを予め記憶してもよく、各仮想スピーカーの構成パラメータはオーディオエンコーダの構成情報を使用することにより決定されてもよい。オーディオエンコーダは上記のエンコーダを示し、オーディオエンコーダの構成情報は、HOAオーダー及び符号化ビットレートを含むが、これらに限定されない。オーディオエンコーダの構成情報は、仮想スピーカーの数及び各仮想スピーカーの位置パラメータを決定するために使用されて、エンコーダが仮想スピーカーの構成パラメータを決定する必要があるという問題を解決してもよい。例えば、符号化ビットレートが低い場合、少数の仮想スピーカーが構成されてもよく、或いは、符号化ビットレートが高い場合、多数の仮想スピーカーが構成されてもよい。他の例では、仮想スピーカーのHOAオーダーはオーディオエンコーダのHOAオーダーと等しくてもよい。この出願の実施形態では、オーディオエンコーダの構成情報を使用することにより複数の仮想スピーカーの構成パラメータを決定することに加えて、複数の仮想スピーカーの構成パラメータはユーザ定義情報に基づいて更に決定できる。例えば、ユーザは仮想スピーカーの位置、HOAオーダー及び仮想スピーカーの数を定義できる。これは限定されない。

【0148】

エンコーダは、仮想スピーカーセットから複数の仮想スピーカーの構成パラメータを取得する。仮想スピーカー毎に、対応する仮想スピーカー構成パラメータが存在し、各仮想スピーカー構成パラメータは、仮想スピーカーのHOAオーダー及び仮想スピーカーの位置座標のような情報を含むが、これらに限定されない。各仮想スピーカーの構成パラメータは、仮想スピーカーについてのHOA係数を生成するために使用できる。HOA係数を生成するプロセスは、HOAアルゴリズムを使用することにより実現でき、詳細はここでは再び説明しない。仮想スピーカーセット内の仮想スピーカー毎にHOA係数が生成され、仮想スピーカーセット内の全ての仮想スピーカーにそれぞれ構成されたHOA係数がHOA係数セットを形成して、エンコーダが仮想スピーカーセット内の各仮想スピーカーについてのHOA係数を決定する必要があるという問題を解決する。

【0149】

この出願のいくつかの実施形態では、第1のターゲット仮想スピーカーの構成パラメータは、第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含む。

【0150】

C2において第1のターゲット仮想スピーカーの構成パラメータに基づいて第1のターゲット仮想スピーカーについてのHOA係数を生成することは、以下を含む。

第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報に基づいて第1のターゲット仮想スピーカーについてのHOA係数を決定する。

【0151】

仮想スピーカーセット内の各仮想スピーカーの構成パラメータは、仮想スピーカーの位置情報及び仮想スピーカーのHOAオーダー情報を含んでもよい。同様に、第1のターゲット仮想スピーカーの構成パラメータは、第1のターゲット仮想スピーカーの位置情報及びH

10

20

30

40

50

OAオーダー情報を含む。例えば、仮想スピーカーセット内の各仮想スピーカーの位置情報は、局所等距離仮想スピーカー空間分布方式に従って決定できる。局所等距離仮想スピーカー空間分布方式は、複数の仮想スピーカーが局所的な等距離の方式で空間内に分布することを意味する。例えば、局所的な等距離の方式は、均等分布又は不均等分布を含んでもよい。各仮想スピーカーの位置情報及びHOAオーダー情報の双方は、仮想スピーカーについてのHOA係数を生成するために使用できる。HOA係数を生成するプロセスは、HOAアルゴリズムを使用することにより実現できる。これは、エンコーダが第1のターゲット仮想スピーカーについてのHOA係数を決定する必要があるという問題を解決する。

【0152】

さらに、この出願のこの実施形態では、仮想スピーカーセット内の仮想スピーカー毎にHOA係数のグループが生成され、複数のHOA係数のグループが上記のHOA係数セットを形成する。仮想スピーカーセット内の全ての仮想スピーカーについてそれぞれ構成されたHOA係数はHOA係数セットを形成し、エンコーダが仮想スピーカーセット内の各仮想スピーカーについてのHOA係数を決定する必要があるという問題を解決する。

【0153】

402:第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成する。

【0154】

エンコーダが第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報を取得した後に、エンコーダは第1のシーンオーディオ信号を再生してもよく、エンコーダは第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて第1の仮想スピーカー信号を生成する。第1の仮想スピーカー信号は、第1のシーンオーディオ信号の再生信号である。第1のターゲット仮想スピーカーの属性情報は、第1のターゲット仮想スピーカーの属性に関連する情報を記述する。第1のターゲット仮想スピーカーは、エンコーダにより選択され且つ第1のシーンオーディオ信号を再生できる仮想スピーカーである。したがって、第1のシーンオーディオ信号は、第1のターゲット仮想スピーカーの属性情報を使用することにより再生されて、第1の仮想スピーカー信号を取得する。第1の仮想スピーカー信号のデータ量は、第1のシーンオーディオ信号のサウンドチャンネルの数に関連せず、第1の仮想スピーカー信号のデータ量は、第1のターゲット仮想スピーカーに関連する。例えば、この出願のこの実施形態では、第1のシーンのオーディオ信号と比較して、第1の仮想スピーカー信号はより少ないサウンドチャンネルを使用することにより表される。例えば、第1のシーンオーディオ信号は3次HOA信号であり、HOA信号は16個のサウンドチャンネルを有する。この出願のこの実施形態では、16個のサウンドチャンネルは4つのサウンドチャンネルに圧縮できる。4つのサウンドチャンネルは、エンコーダにより生成された仮想スピーカー信号により占有される2つのサウンドチャンネルと、残差信号により占有される2つのサウンドチャンネルとを含む。例えば、エンコーダにより生成された仮想スピーカー信号は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を含んでもよく、エンコーダにより生成された仮想スピーカー信号のサウンドチャンネルの数は、第1のシーンオーディオ信号のサウンドチャンネルの数に関連しない。後続のステップにおける説明から、ビットストリームが2つのサウンドチャンネル上で仮想スピーカー信号を搬送し、2つのサウンドチャンネル上で残差信号を搬送してもよいことが分かる。対応して、デコーダはビットストリームを受信し、ビットストリームを復号して、2つのサウンドチャンネル上の仮想スピーカー信号と、2つのサウンドチャンネル上の残差信号とを取得する。デコーダは、2つのサウンドチャンネル上の仮想スピーカー信号及び2つのサウンドチャンネル上の残差信号を使用することにより、16個のサウンドチャンネル上のシーンオーディオ信号を再構成できる。これは、再構成されたシーンオーディオ信号が、元のシーンにおけるオーディオ信号と比較されたときに、同等の主観的及び客観的品質を有することを確保する。

【0155】

上記のステップ401及び402は、空間エンコーダ、例えば、動画専門家グループ(movin

10

20

30

40

50

g picture experts group, MPEG)空間エンコーダを使用することにより具体的に実現されてもよいことが理解され得る。

【0156】

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべきHOA信号を含んでもよく、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークについてのHOA係数を含む。

【0157】

402において第1のシーンオーディオ信号及び第1のターゲット仮想スピーカークの属性情報に基づいて第1の仮想スピーカーク信号を生成することは、以下を含む。

符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカーク信号を取得する。

10

【0158】

第1のシーンオーディオ信号が符号化されるべきHOA信号である例が使用される。まず、エンコーダは、第1のターゲット仮想スピーカークについてのHOA係数を決定する。例えば、エンコーダは、主要音場成分に基づいてHOA係数セットからHOA係数を選択し、選択されたHOA係数は第1のターゲット仮想スピーカークについてのHOA係数である。エンコーダが符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数を取得した後に、第1の仮想スピーカーク信号は、符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に基づいて生成できる。符号化されるべきHOA信号は、第1のターゲット仮想スピーカークについてのHOA係数を使用することにより線形結合を実行することで取得でき、第1の仮想スピーカーク信号の解決が線形結合の解決に変換できる。

20

【0159】

例えば、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークについてのHOA係数を含んでもよい。エンコーダは、第1のターゲット仮想スピーカークの属性情報を復号することにより、第1のターゲット仮想スピーカークについてのHOA係数を取得できる。エンコーダは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行する。言い換えると、エンコーダは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数を一緒に組み合わせて線形結合行列を取得する。次いで、エンコーダは、線形結合行列の最適解を取得でき、取得された最適解は第1の仮想スピーカーク信号である。最適解は、線形結合行列を解くために使用されるアルゴリズムに関連する。この出願のこの実施形態は、エンコーダが第1の仮想スピーカーク信号を生成する必要があるという問題を解決する。

30

【0160】

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報を含む。

【0161】

402において第1のシーンオーディオ信号及び第1のターゲット仮想スピーカークの属性情報に基づいて第1の仮想スピーカーク信号を生成することは、以下を含む。

40

第1のターゲット仮想スピーカークの位置情報に基づいて第1のターゲット仮想スピーカークについてのHOA係数を取得し、

符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカーク信号を取得する。

【0162】

第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報を含んでもよい。エンコーダは、仮想スピーカークセット内の各仮想スピーカークについてのHOA係数を予め記憶する。エンコーダは、各仮想スピーカークの位置情報を更に記憶する。仮想スピーカークの位置情報と仮想スピーカークについてのHOA係数との間に対応関係が存在する。したがって、エンコーダは、第1のターゲット仮想スピーカークの位置情報に基

50

づいて第1のターゲット仮想スピーカについてのHOA係数を決定できる。属性情報がHOA係数を含む場合、エンコーダは、第1のターゲット仮想スピーカの属性情報を復号することにより、第1のターゲット仮想スピーカについてのHOA係数を取得できる。

【0163】

エンコーダが符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数を取得した後に、エンコーダは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数に対して線形結合を実行する。言い換えると、エンコーダは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカについてのHOA係数を組み合わせて線形結合行列を取得する。次いで、エンコーダは、線形結合行列の最適解を取得でき、取得された最適解は第1の仮想スピーカ信号である。

10

【0164】

例えば、第1の仮想スピーカについてのHOA係数は行列Aにより表され、符号化されるべきHOA信号は、行列Aを使用することにより線形結合を通じて取得できる。理論的な最適解w、すなわち、第1の仮想スピーカ信号は、最小二乗法を使用することにより取得できる。例えば、以下の計算式が使用されてもよい。

$$w = A^{-1}X$$

ここで、 A^{-1} は行列Aの逆行列であり、行列Aのサイズは $(M \times C)$ であり、Cは第1のターゲット仮想スピーカの数であり、MはN次HOA係数のサウンドチャンネルの数であり、aは第1のターゲット仮想スピーカについてのHOA係数を表す。例えば、

【数1】

20

$$A = \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1C} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{M1} & \cdot & \cdot & \cdot & a_{MC} \end{bmatrix}$$

である。

【0165】

30

Xは符号化されるべきHOA信号を表し、行列Xのサイズは $(M \times L)$ であり、MはN次HOA係数のサウンドチャンネルの数であり、Lはサンプリング点の数であり、xは符号化されるべきHOA信号についての係数を表す。例えば、

【数2】

$$X = \begin{bmatrix} x_{11} & \cdot & \cdot & \cdot & x_{1L} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{M1} & \cdot & \cdot & \cdot & x_{ML} \end{bmatrix}$$

40

である。

【0166】

この出願のこの実施形態では、デコーダがエンコーダから第1の仮想スピーカ信号を正確に取得できるために、エンコーダは以下のステップ403及び404を更に実行して、残差信号を生成してもよい。

【0167】

403:第1のターゲット仮想スピーカの属性情報及び第1の仮想スピーカ信号を使用することにより、第2のシーンオーディオ信号を取得する。

50

【0168】

エンコーダは、第1のターゲット仮想スピーカークの属性情報を取得でき、第1のターゲット仮想スピーカークは、仮想スピーカークセット内にあり且つデコーダにおいて第1の仮想スピーカーク信号を再生するために使用される仮想スピーカークでもよい。第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報と、第1のターゲット仮想スピーカークについてのHOA係数とを含んでもよい。エンコーダが第1の仮想スピーカーク信号を取得した後に、エンコーダは、第1のターゲット仮想スピーカークの属性情報に基づいて信号再構成を実行し、信号再構成を通じて第2のシーンオーディオ信号を取得できる。

【0169】

この出願のいくつかの実施形態では、403において第1のターゲット仮想スピーカークの属性情報及び第1の仮想スピーカーク信号を使用することにより、第2のシーンオーディオ信号を取得することは、以下を含む。

第1のターゲット仮想スピーカークについてのHOA係数を決定し、

第1の仮想スピーカーク信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して合成処理を実行する。

【0170】

エンコーダは、まず第1のターゲット仮想スピーカークについてのHOA係数を決定する。例えば、エンコーダは、第1のターゲット仮想スピーカークについてのHOA係数を予め記憶してもよい。第1の仮想スピーカーク信号及び第1のターゲット仮想スピーカークについてのHOA係数を取得した後に、エンコーダは、第1の仮想スピーカーク信号及び第1のターゲット仮想スピーカークについてのHOA係数に基づいて再構成されたシーンオーディオ信号を生成できる。

【0171】

例えば、第1のターゲット仮想スピーカークについてのHOA係数は行列Aにより表され、行列Aのサイズは $(M \times C)$ であり、Cは第1のターゲット仮想スピーカークの数であり、MはN次HOA係数のサウンドチャネルの数である。第1の仮想スピーカーク信号は行列Wにより表され、行列Wのサイズは $(C \times L)$ であり、Lは信号サンプリング点の数を表す。再構成されたHOA信号は、以下の式を使用することにより取得される。

$$T = AW$$

【0172】

上記の計算式を使用することにより取得されたTは、第2のシーンオーディオ信号である。

【0173】

404:第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて残差信号を生成する。

【0174】

この出願のこの実施形態では、エンコーダは信号再構成(ローカル復号とも呼ばれてもよい)を通じて第2のシーンオーディオ信号を取得する。第1のシーンオーディオ信号は、元のシーンにおけるオーディオ信号である。したがって、第1のシーンオーディオ信号及び第2のシーンオーディオ信号について残差が計算されて、残差信号を生成できる。残差信号は、第1のターゲット仮想スピーカークを使用することにより生成された第2のシーンオーディオ信号と元のシーンにおけるオーディオ信号(すなわち、第1のシーンオーディオ信号)との間の差を表すことができる。

【0175】

この出願のいくつかの実施形態では、第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて残差信号を生成することは、以下を含む。

第1のシーンオーディオ信号及び第2のシーンオーディオ信号に対して差分計算を実行して残差信号を取得する。

【0176】

第1のシーンオーディオ信号及び第2のシーンオーディオ信号の双方は行列形式で表されることができ、2つのシーンオーディオ信号にそれぞれ対応する行列に対して差分計算を実行することにより残差信号が取得できる。

【0177】

405:第1の仮想スピーカー信号及び残差信号を符号化して、ビットストリームを取得する。

【0178】

この出願のこの実施形態では、エンコーダが第1の仮想スピーカー信号及び残差信号を生成した後に、エンコーダは、第1の仮想スピーカー信号及び残差信号を符号化して、ビットストリームを取得できる。例えば、エンコーダは具体的にはコアエンコーダでもよく、コアエンコーダは、第1の仮想スピーカー信号を符号化してビットストリームを取得する。ビットストリームはまた、オーディオ信号符号化されたビットストリームと呼ばれてもよい。この出願のこの実施形態では、エンコーダは第1の仮想スピーカー信号及び残差信号を符号化するが、シーンオーディオ信号を符号化しない。第1のターゲット仮想スピーカーが選択され、それにより、空間内のリスナーの位置における音場は、シーンオーディオ信号を記録されるとき元の音場にできるだけ近くなり、エンコーダの符号化品質を確保するようにする。さらに、第1の仮想スピーカー信号の符号化データの量は、シーンオーディオ信号のオーディオチャネルの数に関連せず、したがって、符号化されたシーンオーディオ信号のデータの量を低減し、符号化及び復号効率を改善する。

【0179】

この出願のいくつかの実施形態では、エンコーダが上記のステップ401～405を実行した後に、この出願の実施形態において提供されるオーディオ符号化方法は、以下のステップを更に含む。

第1のターゲット仮想スピーカーの属性情報を符号化し、符号化された情報をビットストリームに書き込む。

【0180】

仮想スピーカーを符号化することに加えて、エンコーダはまた、第1のターゲット仮想スピーカーの属性情報を符号化し、第1のターゲット仮想スピーカーの符号化された属性情報をビットストリームに書き込むことができる。この場合、取得されたビットストリームは、符号化された仮想スピーカーと、第1のターゲット仮想スピーカーの符号化された属性情報とを含んでもよい。この出願のこの実施形態では、ビットストリームは、第1のターゲット仮想スピーカーの符号化された属性情報を搬送でき、それにより、デコーダがビットストリームを復号することにより第1のターゲット仮想スピーカーの属性情報を決定して、デコーダによるオーディオ復号を容易にできるようにする。

【0181】

上記のステップ401～405は、第1のターゲット仮想スピーカーが仮想スピーカーセットから選択されるとき、第1のターゲット仮想スピーカーに基づいて第1の仮想スピーカー信号を生成し、第1の仮想スピーカーに基づいて信号再構成、残差信号生成及び信号符号化を実行するプロセスを記載している点に留意すべきである。この出願の実施形態では、エンコーダは第1のターゲット仮想スピーカーを選択するだけでなく、より多くのターゲット仮想スピーカーも選択できる。例えば、エンコーダは第2のターゲット仮想スピーカーを更に選択してもよい。これは限定されない。第2のターゲット仮想スピーカーについても、上記のステップ402～405と同様のプロセスが実行される必要がある。詳細は以下に説明する。

【0182】

この出願のいくつかの実施形態では、エンコーダにより上記のステップを実行することに加えて、この出願のこの実施形態において提供されるオーディオ符号化方法は、以下を更に含む。

D1:第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択する。

D2:第1のシーンオーディオ信号及び第2のターゲット仮想スピーカ属性情報に基づいて第2の仮想スピーカ信号を生成する。

D3:第2の仮想スピーカ信号を符号化し、符号化された信号をビットストリームに書き込む。

【0183】

D1の実現方式は401の実現方式と同様である。第2のターゲット仮想スピーカは、エンコーダにより選択され且つ第1のターゲット仮想エンコーダとは異なる他のターゲット仮想スピーカである。第1のシーンオーディオ信号は元のシーンにおける符号化対象のオーディオ信号であり、第2のターゲット仮想スピーカは仮想スピーカセット内の仮想スピーカでもよい。例えば、第2のターゲット仮想スピーカは予め構成されたターゲット仮想スピーカ選択ポリシーに従って予め設定された仮想スピーカセットから選択できる。ターゲット仮想スピーカ選択ポリシーは、仮想スピーカセットから第1のシーンオーディオ信号に一致するターゲット仮想スピーカを選択するポリシーであり、例えば、第1のシーンオーディオ信号から各仮想スピーカにより取得された音場成分に基づいて第2のターゲット仮想スピーカを選択するポリシーである。

【0184】

この出願のいくつかの実施形態では、この出願のこの実施形態において提供されるオーディオ符号化方法は、以下のステップを更に含む。

E1:仮想スピーカセットに基づいて第1のシーンオーディオ信号から第2の主要音場成分を取得する。

【0185】

E1が実行されるとき、D1において第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカセットから第2のターゲット仮想スピーカを選択することは、以下を含む。

F1:第2の主要音場成分に基づいて仮想スピーカセットから第2のターゲット仮想スピーカを選択する。

【0186】

エンコーダは、仮想スピーカセットを取得し、エンコーダは、仮想スピーカセットを使用することにより第1のシーンオーディオ信号に対して信号分解を実行して、第1のシーンオーディオ信号に対応する第2の主要音場成分を取得する。第2の主要音場成分は、第1のシーンオーディオ信号内の主要音場に対応するオーディオ信号を表す。例えば、仮想スピーカセットは複数の仮想スピーカを含み、複数の音場成分は、複数の仮想スピーカに基づいて第1のシーンオーディオ信号から取得されてもよく、すなわち、各仮想スピーカは、第1のシーンオーディオ信号から1つの音場成分を取得してもよく、次いで、第2の主要音場成分が複数の音場成分から選択される。例えば、第2の主要音場成分は、複数の音場成分の中で最大値を有する1つ以上の音場成分でもよく、代替として、第2の主要音場成分は、複数の音場成分の中で支配的な方向を有する1つ以上の音場成分でもよい。第2のターゲット仮想スピーカは、第2の主要音場成分に基づいて仮想スピーカセットから選択される。例えば、第2の主要音場成分に対応する仮想スピーカは、エンコーダにより選択された第2のターゲット仮想スピーカである。この出願のこの実施形態では、エンコーダは、主要音場成分を使用することにより第2のターゲット仮想スピーカを選択して、エンコーダが第2のターゲット仮想スピーカを決定する必要があるという問題を解決できる。

【0187】

この出願のいくつかの実施形態では、F1において第2の主要音場成分に基づいて仮想スピーカセットから第2のターゲット仮想スピーカを選択することは、以下を含む。

第2の主要音場成分に基づいてHOA係数セットから第2の主要音場成分についてのHOA係数を選択し、ここで、HOA係数セット内のHOA係数は、仮想スピーカセット内の仮想スピーカと1対1の対応関係にあり、

仮想スピーカセットの中で第2の主要音場成分についてのHOA係数に対応する仮想ス

10

20

30

40

50

ピーカーを第2のターゲット仮想スピーカーとして決定する。

【0188】

上記の実現方式は、上記の実施形態における第1のターゲット仮想スピーカーを決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0189】

この出願のいくつかの実施形態では、F1において第2の主要音場成分に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択することは、以下を更に含む。

G1:第2の主要音場成分に基づいて第2のターゲット仮想スピーカーの構成パラメータを取得する。

G2:第2のターゲット仮想スピーカーの構成パラメータに基づいて第2のターゲット仮想スピーカーについてのHOA係数を生成する。

G3:仮想スピーカーセットの中で第2のターゲット仮想スピーカーについてのHOA係数に対応する仮想スピーカーを第2のターゲット仮想スピーカーとして決定する。

【0190】

上記の実現方式は、上記の実施形態における第1のターゲット仮想スピーカーを決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0191】

上記の実現方式は、上記の実施形態における第1のターゲット仮想スピーカーを決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0192】

この出願のいくつかの実施形態では、G1において第2の主要音場成分に基づいて第2のターゲット仮想スピーカーの構成パラメータを取得することは、以下を含む。

オーディオエンコーダの構成情報に基づいて仮想スピーカーセット内の複数の仮想スピーカーの構成パラメータを決定し、

第2の主要音場成分に基づいて複数の仮想スピーカーの構成パラメータから第2のターゲット仮想スピーカーの構成パラメータを選択する。

【0193】

上記の実現方式は、上記の実施形態における第1のターゲット仮想スピーカーの構成パラメータを決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0194】

この出願のいくつかの実施形態では、第2のターゲット仮想スピーカーの構成パラメータは、第2のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含む。

【0195】

G2において第2のターゲット仮想スピーカーの構成パラメータに基づいて第2のターゲット仮想スピーカーについてのHOA係数を生成することは、以下を含む。

第2のターゲット仮想スピーカーの位置情報及びHOAオーダー情報に基づいて第2のターゲット仮想スピーカーについてのHOA係数を決定する。

【0196】

上記の実現方式は、上記の実施形態における第1のターゲット仮想スピーカーについてのHOA係数を決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0197】

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべきHOA信号を含み、第2のターゲット仮想スピーカーの属性情報は、第2のターゲット仮想スピーカーについてのHOA係数を含む。

【0198】

D2において第1のシーンオーディオ信号及び第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成することは、以下を含む。

符号化されるべきHOA信号及び第2のターゲット仮想スピーカーについてのHOA係数に対して線形結合を実行して、第2の仮想スピーカー信号を取得する。

【0199】

10

20

30

40

50

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第2のターゲット仮想スピーカークの属性情報は、第2のターゲット仮想スピーカークの位置情報を含む。

【0200】

D2において第1のシーンオーディオ信号及び第2のターゲット仮想スピーカークの属性情報に基づいて第2の仮想スピーカーク信号を生成することは、以下を含む。

第2のターゲット仮想スピーカークの位置情報に基づいて第2のターゲット仮想スピーカークについてのHOA係数を取得し、

符号化されるべきHOA信号及び第2のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行して、第2の仮想スピーカーク信号を取得する。

【0201】

上記の実現方式は、上記の実施形態における第1の仮想スピーカーク信号を決定するプロセスと同様であり、詳細はここでは再び説明しない。

【0202】

この出願のこの実施形態では、エンコーダが第2の仮想スピーカーク信号を生成した後に、エンコーダは、D3を更に実行して、第2の仮想スピーカーク信号を符号化し、符号化された信号をビットストリームに書き込んでもよい。エンコーダにより使用される符号化方法は405と同様であり、それにより、ビットストリームが第2の仮想スピーカーク信号の符号化結果を搬送できるようにする。

【0203】

対応して、上記のステップD1～D3が実行される実現シーンにおいて、403において第1のターゲット仮想スピーカークの属性情報及び第1の仮想スピーカーク信号を使用することにより、第2のシーンオーディオ信号を取得することは、以下を含む。

H1:第1のターゲット仮想スピーカークの属性情報、第1の仮想スピーカーク信号、第2のターゲット仮想スピーカークの属性情報及び第2の仮想スピーカーク信号に基づいて第2のシーンオーディオ信号を取得する。

【0204】

エンコーダは、第1のターゲット仮想スピーカークの属性情報を取得でき、第1のターゲット仮想スピーカークは、仮想スピーカークセット内にあり且つ第1の仮想スピーカーク信号を再生するために使用される仮想スピーカークである。エンコーダは、第2のターゲット仮想スピーカークの属性情報を取得でき、第2のターゲット仮想スピーカークは、仮想スピーカークセット内にあり且つ第2の仮想スピーカーク信号を再生するために使用される仮想スピーカークである。第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報と、第1のターゲット仮想スピーカークについてのHOA係数とを含んでもよい。第2のターゲット仮想スピーカークの属性情報は、第2のターゲット仮想スピーカークの位置情報と、第2のターゲット仮想スピーカークについてのHOA係数とを含んでもよい。エンコーダが第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号を取得した後に、エンコーダは、第1のターゲット仮想スピーカークの属性情報及び第2のターゲット仮想スピーカークの属性情報に基づいて信号再構成を実行し、信号再構成を通じて第2のシーンオーディオ信号を取得できる。

【0205】

この出願のいくつかの実施形態では、H1において第1のターゲット仮想スピーカークの属性情報、第1の仮想スピーカーク信号、第2のターゲット仮想スピーカークの属性情報及び第2の仮想スピーカーク信号に基づいて第2のシーンオーディオ信号を取得することは、以下を含む。

第1のターゲット仮想スピーカークについてのHOA係数及び第2のターゲット仮想スピーカークについてのHOA係数を決定し、

第1の仮想スピーカーク信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して合成処理を実行し、第2の仮想スピーカーク信号及び第2のターゲット仮想スピーカークについてのHOA係数に対して合成処理を実行する。

【0206】

エンコーダは、まず第1のターゲット仮想スピーカーについてのHOA係数を決定する。例えば、エンコーダは、第1のターゲット仮想スピーカーについてのHOA係数を予め記憶してもよく、エンコーダは、第2のターゲット仮想スピーカーについてのHOA係数を決定する。例えば、エンコーダは、第2のターゲット仮想スピーカーについてのHOA係数を予め記憶してもよく、エンコーダは、第1の仮想スピーカー信号、第1のターゲット仮想スピーカーについてのHOA係数、第2の仮想スピーカー信号及び第2のターゲット仮想スピーカーについてのHOA係数に基づいて再構成されたシーンオーディオ信号を生成する。

【0207】

この出願のいくつかの実施形態では、エンコーダにより実行されるオーディオ符号化方法は、以下のステップを更に含んでもよい。

I1:第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得する。

【0208】

I1が実行されるとき、対応して、D3において第2の仮想スピーカー信号を符号化することは、以下を含む。

整列された第2の仮想スピーカー信号を符号化する。

【0209】

対応して、405において第1の仮想スピーカー信号及び残差信号を符号化することは、以下を含む。

整列された第1の仮想スピーカー信号及び残差信号を符号化する。

【0210】

エンコーダは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を生成でき、エンコーダは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得できる。例えば、2つの仮想スピーカー信号が存在し、現在のフレームの仮想スピーカー信号のサウンドチャンネル系列が、ターゲット仮想スピーカーP1及びP2により生成された仮想スピーカー信号にそれぞれ対応する1及び2であり、以前のフレームの仮想スピーカー信号のサウンドチャンネル系列が、ターゲット仮想スピーカーP2及びP1により生成された仮想スピーカー信号にそれぞれ対応する1及び2である場合、現在のフレームの仮想スピーカー信号のサウンドチャンネル系列は、以前のフレームのターゲット仮想スピーカーの系列に基づいて調整できる。例えば、現在のフレームの仮想スピーカー信号のサウンドチャンネル系列は2及び1に調整され、それにより、同じターゲット仮想スピーカーにより生成された仮想スピーカー信号が同じサウンドチャンネル上にあるようにする。

【0211】

整列された第1の仮想スピーカー信号を取得した後に、エンコーダは、整列された第1の仮想スピーカー信号及び残差信号を符号化できる。この出願の実施形態では、第1の仮想スピーカー信号のサウンドチャンネルを再び調整して整列させることにより、チャンネル間相関が強化されて、コアエンコーダによる第1の仮想スピーカー信号の符号化処理を容易にする。

【0212】

この出願のいくつかの実施形態では、エンコーダにより上記のステップを実行することに加えて、この出願のこの実施形態において提供されるオーディオ符号化方法は、以下を更に含む。

D1:第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択する。

D2:第1のシーンオーディオ信号及び第2のターゲット仮想スピーカーの属性情報に基づいて第2の仮想スピーカー信号を生成する。

【0213】

対応して、エンコーダがD1及びD2を実行するとき、405において第1の仮想スピーカー

10

20

30

40

50

信号及び残差信号を符号化することは、以下のステップを含む。

【0214】

J1:第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得し、ここで、第1のサイド情報は第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示す。

【0215】

本発明のこの実施形態では、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係は、直接的な関係又は間接的な関係でもよい。例えば、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係が直接的な関係であるとき、第1のサイド情報は、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の相関パラメータを含んでもよく、例えば、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間のエネルギー比率パラメータでもよい。例えば、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係が間接的な関係であるとき、第1のサイド情報は、第1の仮想スピーカー信号とダウンミキシングされた信号との間の相関パラメータと、第2の仮想スピーカー信号とダウンミキシングされた信号との間の相関パラメータとを含んでもよく、例えば、第1の仮想スピーカー信号とダウンミキシングされた信号との間のエネルギー比率パラメータと、第2の仮想スピーカー信号とダウンミキシングされた信号との間のエネルギー比率パラメータとを含んでもよい。

【0216】

第1の仮想スピーカー信号と第2の仮想スピーカー信号との関係が直接的な関係でもよいとき、デコーダは、ダウンミキシングされた信号、ダウンミキシングされた信号を取得するための方式及び直接的な関係に基づいて、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を決定できる。第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係が間接的な関係でもよいとき、デコーダは、ダウンミキシングされた信号及び間接的な関係に基づいて、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を決定できる。

【0217】

J2:ダウンミキシングされた信号、第1のサイド情報及び残差信号を符号化する。

【0218】

エンコーダが第1の仮想スピーカー信号及び第2の仮想スピーカー信号を取得した後に、エンコーダは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングを更に実行して、ダウンミキシングされた信号を生成でき、例えば、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に対して振幅ダウンミキシングを実行して、ダウンミキシングされた信号を取得できる。さらに、第1のサイド情報は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいて更に生成できる。第1のサイド情報は、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示し、当該関係は複数の実現方式を有する。第1のサイド情報は、デコーダにより、ダウンミキシングされた信号をアップミキシングし、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元するために使用できる。例えば、第1のサイド情報は信号情報ロス分析パラメータを含み、それにより、デコーダは信号情報ロス分析パラメータを使用することにより第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元するようにする。他の例では、第1のサイド情報は、具体的には、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の相関パラメータでもよく、例えば、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間のエネルギー比率パラメータでもよい。したがって、デコーダは、相関パラメータ又はエネルギー比率パラメータを使用することにより、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を復元する。

【0219】

この出願のいくつかの実施形態では、エンコーダがD1及びD2を実行するとき、エンコーダは、以下のステップを更に実行してもよい。

I1:第1の仮想スピーカー信号及び第2の仮想スピーカー信号を整列させて、整列された

第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号を取得する。

【0220】

I1が実行される時、対応して、J1において第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得することは、以下を含む。

整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得する。

【0221】

対応して、第1のサイド情報は、整列された第1の仮想スピーカー信号と整列された第2の仮想スピーカー信号との間の関係を示す。

【0222】

ダウンミキシングされた信号を生成する前に、まず、エンコーダは、仮想スピーカー信号に対して整列動作を実行し、整列動作を完了した後に、ダウンミキシングされた信号及び第1のサイド情報を生成できる。この出願のこの実施形態では、第1の仮想スピーカー信号及び第2の仮想スピーカー信号のサウンドチャンネルを再び調整して整列させることにより、チャンネル間相関が強化されて、コアエンコーダによる第1の仮想スピーカー信号の符号化処理を容易にする。

【0223】

この出願の上記の実施形態では、第2のシーンオーディオ信号は、整列前の第1の仮想スピーカー信号及び整列前の第2の仮想スピーカー信号に基づいて取得でき、或いは、整列された第1の仮想スピーカー信号及び整列された第2の仮想スピーカー信号に基づいて取得できる点に留意すべきである。具体的な実現方式は適用シーンに依存し、ここでは限定されない。

【0224】

この出願のいくつかの実施形態では、D1において第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択する前に、この出願のこの実施形態において提供されるオーディオ信号符号化方法は、以下を更に含む。

K1:第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要があるか否かを決定する。

K2:第1のターゲット仮想スピーカー以外のターゲット仮想スピーカーが取得される必要がある場合にのみ、第1のシーンオーディオ信号に基づいて仮想スピーカーセットから第2のターゲット仮想スピーカーを選択する。

【0225】

エンコーダは、信号を更に選択して、第2のターゲット仮想スピーカーが取得される必要があるか否かを決定できる。第2のターゲット仮想スピーカーが取得される必要があるとき、エンコーダは第2の仮想スピーカー信号を生成してもよい。第2のターゲット仮想スピーカーが取得される必要がないとき、エンコーダは第2の仮想スピーカー信号を生成しなくてもよい。エンコーダは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、第1のターゲット仮想スピーカーに加えて他のターゲット仮想スピーカーが選択される必要があるか否かを決定できる。例えば、符号化率が予め設定された閾値よりも高い場合、2つの主要音場成分に対応するターゲット仮想スピーカーが取得される必要があると決定され、第1のターゲット仮想スピーカーが決定されることに加えて、第2のターゲット仮想スピーカーが更に決定されてもよい。他の例では、第1のシーンオーディオ信号の信号クラス情報に基づいて、支配的な音源方向を含む2つの主要音場成分に対応するターゲット仮想スピーカーが取得される必要があると決定された場合、第1のターゲット仮想スピーカーが決定されることに加えて、第2のターゲット仮想スピーカーが更に決定されてもよい。逆に、第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、1つのターゲット仮想スピーカーのみが取得される必要があると決定された場合、第1のターゲット仮想スピーカーが決定された後に

10

20

30

40

50

、第1のターゲット仮想スピーカ以外のターゲット仮想スピーカが取得されないと決定される。この出願のこの実施形態では、信号が選択され、それにより、エンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。

【0226】

信号を選択するとき、エンコーダは、第2の仮想スピーカ信号が生成される必要があるか否かを決定できる。エンコーダが信号を選択するときに情報ロスが発生するので、伝送されない仮想スピーカ信号に対して信号補償が実行される必要がある。信号補償は、情報ロス分析、エネルギー補償、エンベロープ補償及びノイズ補償でもよく、これらに限定されない。補償方法は、線形補償、非線形補償等でもよい。信号補償の後に、第1のサイド情報が生成でき、第1のサイド情報がビットストリームに書き込まれることができ、それにより、デコーダはビットストリームを使用することにより第1のサイド情報を取得でき、デコーダは第1のサイド情報に基づいて信号補償を実行して、デコーダの復号信号の品質を改善できるようにする。

【0227】

この出願のいくつかの実施形態では、信号選択のために、第2の仮想スピーカ信号が生成される必要があるか否かを選択することに加えて、エンコーダは、残差信号についての信号選択を更に実行して、残差信号の中のどの残差サブ信号が伝送されるかを決定してもよい。例えば、残差信号は少なくとも2つのサウンドチャンネル上の残差サブ信号を含み、この出願のこの実施形態において提供されるオーディオ信号符号化方法は、以下を更に含む。

L1:オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、少なくとも2つのサウンドチャンネル上の残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定する。

【0228】

L1が実行される実現シーンでは、対応して、405において第1の仮想スピーカ信号及び残差信号を符号化することは、以下を含む。

第1の仮想スピーカ信号と、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号とを符号化する。

【0229】

エンコーダは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて残差信号に対する決定を行うことができる。例えば、残差信号が少なくとも2つのサウンドチャンネル上の残差サブ信号を含む場合、エンコーダは、残差サブ信号が符号化される必要があるサウンドチャンネル又は複数のサウンドチャンネルと、残差サブ信号が符号化される必要がないサウンドチャンネル又は複数のサウンドチャンネルとを選択できる。例えば、残差信号において支配的なエネルギーを有する残差サブ信号は、符号化するためにオーディオエンコーダの構成情報に基づいて選択される。他の例では、残差信号における低次HOAサウンドチャンネルによる計算を通じて取得された残差サブ信号は、符号化するために第1のシーンオーディオ信号の信号クラス情報に基づいて選択される。残差信号についてサウンドチャンネルが選択され、それにより、エンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。

【0230】

この出願のいくつかの実施形態では、少なくとも2つのサウンドチャンネル上の残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、この出願のこの実施形態において提供されるオーディオ信号符号化方法は、以下を更に含む。

第2のサイド情報を取得し、ここで、第2のサイド情報は、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号と、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号との間の関係を示し、

第2のサイド情報をビットストリームに書き込む。

【0231】

10

20

30

40

50

信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。この出願のこの実施形態では、符号化される必要がある残差サブ信号が符号化され、符号化される必要がない残差サブ信号が符号化されず、それにより、エンコーダにより符号化されるデータの量が低減されて、符号化効率を改善できるようにする。エンコーダが信号を選択するときに情報ロスが発生するので、伝送されない残差サブ信号に対して信号補償が実行される必要がある。信号補償は、情報ロス分析、エネルギー補償、エンベロープ補償及びノイズ補償でもよく、これらに限定されない。補償方法は、線形補償、非線形補償等でもよい。信号補償の後に、第2のサイド情報が生成されてもよく、第2のサイド情報がビットストリームに書き込まれてもよい。第2のサイド情報は、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間の関係を示す。当該関係は複数の実現方式を有する。例えば、第2のサイド情報は信号情報ロス分析パラメータを含み、それにより、デコーダが信号情報ロス分析パラメータを使用することにより、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号とを復元するようにする。他の例では、第2のサイド情報は、具体的には、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間の相関パラメータでもよく、例えば、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号との間のエネルギー比率パラメータでもよい。したがって、デコーダは、相関パラメータ又はエネルギー比率パラメータを使用することにより、符号化される必要がある残差サブ信号と符号化される必要がない残差サブ信号とを復元する。この出願のこの実施形態では、デコーダは、ビットストリームを使用することにより第2のサイド情報を取得でき、デコーダは、第2のサイド情報に基づいて信号補償を実行して、デコーダの復号信号の品質を改善できる。

【0232】

上記の実施形態における例示的な説明によれば、この出願の実施形態では、第1のシーンオーディオ信号について第1のターゲット仮想スピーカーが構成できる。さらに、オーディオエンコーダは、第1の仮想スピーカー信号及び第1のターゲット仮想スピーカーの属性情報に基づいて残差信号を更に取得できる。オーディオエンコーダは、第1のシーンオーディオ信号を直接符号化する代わりに、第1の仮想スピーカー信号及び残差信号を符号化する。この出願のこの実施形態では、第1のターゲット仮想スピーカーは、第1のシーンオーディオ信号に基づいて選択され、第1のターゲット仮想スピーカーに基づいて生成された第1の仮想スピーカー信号は、空間内のリスナーの位置における音場を表すことができる。当該位置における音場は、第1のシーンオーディオ信号が記録されときの元の音場にできるだけ近くなり、それにより、オーディオエンコーダの符号化品質を確保する。さらに、第1の仮想スピーカー信号及び残差信号は、ビットストリームを取得するために符号化され、第1の仮想スピーカー信号の符号化データの量が第1のターゲット仮想スピーカーに関連し、第1のシーンオーディオ信号のサウンドチャンネルの数に関連せず、それにより、符号化データの量が低減され、符号化効率が改善されるようにする。

【0233】

この出願のこの実施形態では、エンコーダは、第1の仮想スピーカー信号及び残差信号を符号化して、ビットストリームを生成する。次いで、エンコーダはビットストリームを出力し、オーディオ伝送チャネルを通じてビットストリームをデコーダに送信できる。デコーダは後続のステップ411~413を実行する。

【0234】

411:ビットストリームを受信する。

【0235】

デコーダは、エンコーダからビットストリームを受信する。ビットストリームは、符号化された第1の仮想スピーカー信号及び符号化された残差信号を搬送できる。ビットストリームは、第1のターゲット仮想スピーカーの符号化された属性情報を更に搬送してもよい。これは限定されない。ビットストリームは、第1のターゲット仮想スピーカーの属性情報を搬送しなくてもよい点に留意すべきである。この場合、デコーダは事前構成を通じ

て第1のターゲット仮想スピーカークの属性情報を決定できる。

【0236】

さらに、この出願のいくつかの実施形態では、エンコーダが第2の仮想スピーカーク信号を生成するとき、ビットストリームは第2の仮想スピーカーク信号を更に搬送してもよい。ビットストリームは、第2のターゲット仮想スピーカークの符号化された属性情報を更に搬送してもよい。これは限定されない。ビットストリームは、第2のターゲット仮想スピーカークの属性情報を搬送しなくてもよい点に留意すべきである。この場合、デコーダは、事前構成を通じて第2のターゲット仮想スピーカークの属性情報を決定できる。

【0237】

412:ビットストリームを復号して、仮想スピーカーク信号及び残差信号を取得する。

10

【0238】

エンコーダからビットストリームを受信した後に、デコーダはビットストリームを復号し、ビットストリームから仮想スピーカーク信号及び残差信号を取得する。

【0239】

仮想スピーカーク信号は、具体的には第1の仮想スピーカーク信号でもよく、或いは、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号でもよく、これはここでは限定されない点に留意すべきである。

【0240】

この出願のいくつかの実施形態では、デコーダが411及び412を実行した後に、この出願のこの実施形態において提供されるオーディオ復号方法は、以下のステップを更に含む。

20

ビットストリームを復号して、ターゲット仮想スピーカークの属性情報を取得する。

【0241】

仮想スピーカークを符号化することに加えて、エンコーダはまた、ターゲット仮想スピーカークの属性情報を符号化し、ターゲット仮想スピーカークの符号化された属性情報をビットストリームに書き込むことができる。例えば、第1のターゲット仮想スピーカークの属性情報は、ビットストリームを使用することにより取得できる。この出願のこの実施形態では、ビットストリームは、第1のターゲット仮想スピーカークの符号化された属性情報を搬送でき、それにより、デコーダがビットストリームを復号することにより第1のターゲット仮想スピーカークの属性情報を決定して、デコーダによるオーディオ復号を容易にできるようにする。

30

【0242】

413:ターゲット仮想スピーカークの属性情報、残差信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得する。

【0243】

デコーダは、ターゲット仮想スピーカークの属性情報及び残差信号を取得できる。ターゲット仮想スピーカークは、仮想スピーカークセット内にあり且つ再構成されたシーンオーディオ信号を再生するために使用される仮想スピーカークである。ターゲット仮想スピーカークの属性情報は、ターゲット仮想スピーカークの位置情報と、ターゲット仮想スピーカークのHOA係数とを含んでもよい。仮想スピーカーク信号を取得した後に、デコーダは、ターゲット仮想スピーカークの属性情報及び残差信号に基づいて信号再構成を実行し、信号再構成を通じて再構成されたシーンオーディオ信号を出力できる。仮想スピーカーク信号はシーンオーディオ信号内の主要音場成分を再構成するために使用され、残差信号は再構成されたシーンオーディオ信号内の無指向性成分を補償する。残差信号は再構成されたシーンオーディオ信号の品質を改善できる。

40

【0244】

この出願のいくつかの実施形態では、ターゲット仮想スピーカークの属性情報は、ターゲット仮想スピーカークについてのHOA係数を含む。

【0245】

413においてターゲット仮想スピーカークの属性情報、残差信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得することは、以下を含む。

50

仮想スピーカ信号及びターゲット仮想スピーカについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、

残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得する。

【0246】

まず、デコーダは、ターゲット仮想スピーカについてのHOA係数を決定する。例えば、デコーダは、ターゲット仮想スピーカについてのHOA係数を予め記憶してもよい。仮想スピーカ信号及びターゲット仮想スピーカについてのHOA係数を取得した後に、デコーダは、仮想スピーカ信号及びターゲット仮想スピーカについてのHOA係数に基づいて合成されたシーンオーディオ信号を取得できる。最後に、合成されたシーンオーディオ信号を調整するために残差信号が使用されて、再構成されたシーンオーディオ信号の品質を改善する。

【0247】

例えば、ターゲット仮想スピーカについてのHOA係数は行列 A' により表され、行列 A' のサイズは $(M \times C)$ であり、 C はターゲット仮想スピーカの数であり、 M は N 次HOA係数のサウンドチャンネルの数である。仮想スピーカ信号は行列 W' により表され、行列 W' のサイズは $(C \times L)$ であり、 L は信号サンプリング点の数を表す。再構成されたHOA信号は、以下の式を使用することにより取得される。

$$H = A'W'$$

【0248】

上記の計算式を使用することにより取得された H は、再構成されたHOA信号である。

【0249】

上記の再構成されたHOA信号が取得された後に、合成されたシーンオーディオ信号を調整するために残差信号が更に使用されて、再構成されたシーンオーディオ信号の品質を改善できる。

【0250】

この出願のいくつかの実施形態では、ターゲット仮想スピーカの属性情報は、ターゲット仮想スピーカの位置情報を含む。

【0251】

413においてターゲット仮想スピーカの属性情報、残差信号及び仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得することは、以下を含む。

ターゲット仮想スピーカの位置情報に基づいてターゲット仮想スピーカについてのHOA係数を決定し、

仮想スピーカ信号及びターゲット仮想スピーカについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、

残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得する。

【0252】

ターゲット仮想スピーカの属性情報は、ターゲット仮想スピーカの位置情報を含んでもよい。デコーダは、仮想スピーカセット内の各仮想スピーカについてのHOA係数を予め記憶し、デコーダは、各仮想スピーカの位置情報を更に記憶する。例えば、デコーダは、仮想スピーカの位置情報と仮想スピーカについてのHOA係数との間の対応関係に基づいて、ターゲット仮想スピーカの位置情報についてのHOA係数を決定でき、或いは、デコーダは、ターゲット仮想スピーカの位置情報に基づいてターゲット仮想スピーカについてのHOA係数を計算できる。したがって、デコーダは、ターゲット仮想スピーカの位置情報に基づいてターゲット仮想スピーカについてのHOA係数を決定できる。これは、デコーダがターゲット仮想スピーカについてのHOA係数を決定する必要があるという問題を解決する。

【0253】

この出願のいくつかの実施形態では、エンコーダの方法の説明から、仮想スピーカ信

10

20

30

40

50

号が、第1の仮想スピーカー信号及び第2の仮想スピーカー信号をダウンミキシングすることにより取得されたダウンミキシングされた信号であることが分かる。この実現シーンにおいて、この出願のこの実施形態において提供されるオーディオ復号方法は、以下を更に含む。

ビットストリームを復号して、第1のサイド情報を取得し、ここで、第1のサイド情報は、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示し、

第1のサイド情報及びダウンミキシングされた信号に基づいて第1の仮想スピーカー信号及び第2の仮想スピーカー信号を取得する。

【0254】

対応して、413においてターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得することは、以下を含む。

ターゲット仮想スピーカーの属性情報、残差信号、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得する。

【0255】

エンコーダは、第1の仮想スピーカー信号及び第2の仮想スピーカー信号に基づいてダウンミキシングを実行するとき、ダウンミキシングされた信号を生成し、エンコーダは、ダウンミキシングされた信号に対して信号補償を更に実行して、第1のサイド情報を生成できる。第1のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第1のサイド情報を取得できる。デコーダは、第1のサイド情報に基づいて信号補償を実行して、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を取得できる。したがって、信号再構成中に、第1の仮想スピーカー信号、第2の仮想スピーカー信号、ターゲット仮想スピーカーの属性情報及び残差信号が使用されて、デコーダの復号信号の品質を改善できる。

【0256】

この出願のいくつかの実施形態では、エンコーダの方法の説明から、エンコーダが残差信号のための信号選択を実行し、第2のサイド情報をビットストリームに追加することが分かる。この実現シーンにおいて、残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、この出願のこの実施形態において提供されるオーディオ復号方法は、以下を更に含む。

ビットストリームを復号して、第2のサイド情報を取得し、ここで、第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示し、

第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第2のサウンドチャンネル上の残差サブ信号を取得する。

【0257】

対応して、413においてターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得することは、以下を含む。

ターゲット仮想スピーカーの属性情報、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得する。

【0258】

信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。エンコーダが信号を選択するときに情報ロスが発生するので、エンコーダは第2のサイド情報を生成する。第2のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第2のサイド情報を取得できる。ビットストリームで搬送される残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、デコーダは、第2のサイド情報に基づいて信号補償を実行して、第2のサウンドチャンネル上の残差サブ信号を取得できる。例えば、デコーダは、第1のサウンドチャンネル上の残差サブ信号及び第2のサイド情報を使用することにより、第2のサウンドチャンネル上の残差サブ信号を復元する。第2のサウンド

10

20

30

40

50

チャンネルは、第1のサウンドチャンネルから独立している。したがって、信号再構成中に、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号、ターゲット仮想スピーカークの属性情報及び仮想スピーカーク信号が使用されて、デコーダの復号信号の品質を改善できる。例えば、シーンオーディオ信号は合計で16個のサウンドチャンネルを含む。16個のサウンドチャンネルには、4つの第1のサウンドチャンネル、例えば、サウンドチャンネル1、3、5及び7が存在し、第2のサイド情報は、サウンドチャンネル1、3、5及び7上の残差サブ信号と他のサウンドチャンネル上の残差サブ信号との間の関係を記述する。したがって、デコーダは、第1のサウンドチャンネル上の残差サブ信号及び第2のサイド情報に基づいて、16個のサウンドチャンネル内の他の12個のサウンドチャンネル上の残差サブ信号を取得できる。他の例では、シーンオーディオ信号は合計で16個のサウンドチャンネルを含む。第1のサウンドチャンネルは16個のサウンドチャンネル内の第3のサウンドチャンネルであり、第2のサウンドチャンネルは16個のサウンドチャンネル内の第8のサウンドチャンネルであり、第2のサイド情報は、第3のサウンドチャンネル上の残差サブ信号と第8のサウンドチャンネル上の残差サブ信号との間の関係を記述する。したがって、デコーダは第3のサウンドチャンネル上の残差サブ信号及び第2のサイド情報に基づいて第8のサウンドチャンネル上の残差サブ信号を取得できる。

【0259】

この出願のいくつかの実施形態では、エンコーダの方法の説明から、エンコーダが残差信号のための信号選択を実行し、第2のサイド情報をビットストリームに追加することが分かる。この実現シーンにおいて、残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、この出願のこの実施形態において提供されるオーディオ復号方法は、以下を含む。

ビットストリームを復号して、第2のサイド情報を取得し、ここで、第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示し、

第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を取得する。

【0260】

対応して、413においてターゲット仮想スピーカークの属性情報、残差信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得することは、以下を含む。

ターゲット仮想スピーカークの属性情報、第1のサウンドチャンネル上の更新された残差サブ信号、第3のサウンドチャンネル上の残差サブ信号及び仮想スピーカーク信号に基づいて再構成されたシーンオーディオ信号を取得する。

【0261】

1つ以上の第1のサウンドチャンネルが存在してもよく、1つ以上の第2のサウンドチャンネルが存在してもよく、或いは、1つ以上の第3のサウンドチャンネルが存在してもよい。

【0262】

信号を選択するとき、エンコーダは、符号化される必要がある残差サブ信号と、符号化される必要がない残差サブ信号とを決定できる。エンコーダが信号を選択するときに情報ロスが発生するので、エンコーダは第2のサイド情報を生成する。第2のサイド情報はビットストリームに書き込まれることができる。デコーダは、ビットストリームを使用することにより、第2のサイド情報を取得できる。ビットストリームで搬送される残差信号が第1のサウンドチャンネル上の残差サブ信号を含むと仮定し、デコーダは、第2のサイド情報に基づいて信号補償を実行して、第2のサウンドチャンネル上の残差サブ信号を取得できる。第3のサウンドチャンネル上の残差サブ信号は、第1のサウンドチャンネル上の残差サブ信号とは異なる。第3のサウンドチャンネル上の残差サブ信号が第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて取得されるとき、第1のサウンドチャンネル上の残差サブ信号は、第1のサウンドチャンネル上の更新された残差サブ信号を取得するために更新される必要がある。例えば、デコーダは、第1のサウンドチャンネル上の残差サブ信号及

び第2のサイド情報を使用することにより、第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を生成する。したがって、信号再構成中に、第3のサウンドチャンネル上の残差サブ信号、第1のサウンドチャンネル上の更新された残差サブ信号、ターゲット仮想スピーカ属性情報及び仮想スピーカー信号が使用されて、デコーダの復号信号の品質を改善できる。例えば、シーンオーディオ信号は合計で16個のサウンドチャンネルを含む。16個のサウンドチャンネルには、4つの第1のサウンドチャンネル、例えば、サウンドチャンネル1、3、5及び7が存在し、第2のサイド情報は、サウンドチャンネル1、3、5及び7上の残差サブ信号と他のサウンドチャンネル上の残差サブ信号との間の関係を記述する。したがって、デコーダは、第1のサウンドチャンネル上の残差サブ信号及び第2のサイド情報に基づいて、16個のサウンドチャンネル上の残差サブ信号を取得でき、16個のサウンドチャンネル上の残差サブ信号は、サウンドチャンネル1、3、5及び7上の更新された残差サブ信号を含む。他の例では、シーンオーディオ信号は合計で16個のサウンドチャンネルを含む。第1のサウンドチャンネルは16個のサウンドチャンネル内の第3のサウンドチャンネルであり、第2のサウンドチャンネルは16個のサウンドチャンネル内の第8のサウンドチャンネルであり、第2のサイド情報は、第3のサウンドチャンネル上の残差サブ信号と第8のサウンドチャンネル上の残差サブ信号との間の関係を記述する。したがって、デコーダは第3のサウンドチャンネル上の残差サブ信号及び第2のサイド情報に基づいて第8のサウンドチャンネル上の残差サブ信号及び第3のサウンドチャンネル上の更新された残差サブ信号を取得できる。

10

【0263】

20

この出願のいくつかの実施形態では、エンコーダの方法の説明から、エンコーダにより生成されたビットストリームが第1のサイド情報及び第2のサイド情報の双方を搬送してもよいことが分かる。この場合、デコーダは、ビットストリームを復号して第1のサイド情報及び第2のサイド情報を取得する必要がある。デコーダは、第1のサイド情報を使用して信号補償を実行する必要がある。さらに、第2のサイド情報を使用して信号補償を実行する必要がある。言い換えると、デコーダは、第1のサイド情報及び第2のサイド情報に基づいて信号補償を実行して、信号補償された仮想スピーカー信号及び信号補償された残差信号を取得してもよい。したがって、信号再構成中に、信号補償された仮想スピーカー信号及び信号補償された残差信号が使用されて、デコーダの復号信号の品質を改善できる。

【0264】

30

上記の実施形態における例の説明では、まず、ビットストリームが受信され、次いで、復号されて仮想スピーカー信号及び残差信号を取得し、最後に、再構成されたシーンオーディオ信号は、ターゲット仮想スピーカ属性情報、残差信号及び仮想スピーカー信号に基づいて取得される。この出願のこの実施形態では、オーディオデコーダは、オーディオエンコーダによる符号化プロセスとは逆の復号プロセスを実行し、復号を通じてビットストリームから仮想スピーカー信号及び残差信号を取得し、ターゲット仮想スピーカ属性情報、残差信号及び仮想スピーカー信号を使用することにより、再構成されたシーンオーディオ信号を取得できる。この出願のこの実施形態では、取得されたビットストリームは、仮想スピーカー信号及び残差信号を搬送し、復号されるデータの量を低減し、復号効率を改善する。

40

【0265】

例えば、この出願のこの実施形態では、第1のシーンのオーディオ信号と比較して、第1の仮想スピーカー信号はより少ないサウンドチャンネルを使用することにより表される。例えば、第1のシーンオーディオ信号は3次HOA信号であり、HOA信号は16個のサウンドチャンネルを有する。この出願のこの実施形態では、16個のサウンドチャンネルは4つのサウンドチャンネルに圧縮できる。4つのサウンドチャンネルは、エンコーダにより生成された仮想スピーカー信号により占有される2つのサウンドチャンネルと、残差信号により占有される2つのサウンドチャンネルとを含む。例えば、エンコーダにより生成された仮想スピーカー信号は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号を含んでもよく、エンコーダにより生成された仮想スピーカー信号のサウンドチャンネルの数は、第1のシーンオーデ

50

ィオ信号のサウンドチャンネルの数に関連しない。後続のステップにおける説明から、ビットストリームが2つのサウンドチャンネル上で仮想スピーカー信号を搬送し、2つのサウンドチャンネル上で残差信号を搬送してもよいことが分かる。対応して、デコーダはビットストリームを受信し、ビットストリームを復号して、2つのサウンドチャンネル上の仮想スピーカー信号と、2つのサウンドチャンネル上の残差信号とを取得する。デコーダは、2つのサウンドチャンネル上の仮想スピーカー信号及び2つのサウンドチャンネル上の残差信号を使用することにより、16個のサウンドチャンネル上のシーンオーディオ信号を再構成できる。これは、再構成されたシーンオーディオ信号が、元のシーンにおけるオーディオ信号と比較されたときに、同等の主観的及び客観的品質を有することを確保する。

【0266】

この出願のこの実施形態における上記の解決策のより良い理解及び実現のために、対応する適用シーンを例として使用することにより、具体的な説明が以下に提供される。

【0267】

この出願のこの実施形態では、シーンオーディオ信号がHOA信号である例が使用される。音波は理想的な媒体で伝播され、波の数は $k=w/c$ であり、角周波数は $w=2\pi f$ であり、 f は音波の周波数であり、 c は音速である。この場合、音圧 p は次の計算式を満たし、ここで、 ∇^2 はラプラス演算子である。

$$\nabla^2 p + k^2 p = 0$$

【0268】

上記の式が球座標の下で解かれる。受動球面領域では、式の解は以下のようになる。

【数3】

$$p(r, \theta, \varphi, k) = s \sum_{m=0}^{\infty} (2m+1) j_m^{kr} j_m^{kr}(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} Y_{m,n}^{\sigma}(\theta_s, \varphi_s) Y_{m,n}^{\sigma}(\theta, \varphi)$$

【0269】

上記の計算式において、 r は球面半径を表し、 θ は水平角を表し、 φ は仰角を表し、 k は波の数を表し、 s は理想的な平面波の振幅であり、 m はHOAオーダーの系列番号であり、 $j_m^{kr} j_m^{kr}(kr)$ は球面ベッセル関数であり、ラジアル基底関数とも呼ばれ、ここで、最初の j は虚数単位である。 $(2m+1) j_m^{kr} j_m^{kr}(kr)$ は角度によって変化しない。 $Y_{m,n}^{\sigma}(\theta, \varphi)$ は θ, φ の方向の球面調和関数であり、 $Y_{m,n}^{\sigma}(\theta_s, \varphi_s)$ は音源の方向の球面調和関数である。

【0270】

HOA係数は、 $B_{(m,n)}^{\sigma} = s \cdot Y_{m,n}^{\sigma}(\theta_s, \varphi_s)$ として表現されてもよい。

【0271】

以下の計算式が提供される。

【数4】

$$p(r, \theta, \varphi, k) = \sum_{m=0}^{\infty} j_m^{kr} j_m^{kr}(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta, \varphi)$$

【0272】

上記の計算式は、音場が球面調和関数に従って球面上で展開され、係数 $B_{m,n}^{\sigma}$ を使用することにより表現できることを示している。代替として、係数 $B_{m,n}^{\sigma}$ が既知である場合、音場が再構成できる。上記の式は第 N 項まで切り捨てられ、係数 $B_{m,n}^{\sigma}$ は音場の近似的な記述として使用され、 N 次HOA係数と呼ばれる。HOA係数はまた、アンビソニック係数と呼ばれてもよい。 N 次HOA係数は、合計で $(N+1)^2$ 個のサウンドチャンネルを有する。1次よりも大きいアンビソニック信号はまた、HOA信号とも呼ばれる。HOA信号のサンプリング

10

20

30

40

50

点についての係数に従って球面調和関数を重ね合わせるにより、サンプリング点に対応する時点での空間的音場が再構成できる。

【0273】

例えば、構成において、HOAオーダーは2～6でもよく、シーン内のオーディオが記録されるとき、信号サンプリングレートは48kHz～192kHzであり、サンプリング深度は16ビット又は24ビットである。HOA信号は、音場の空間情報を特徴とし、空間内の或る点における音場信号の特定の精度の記述である。したがって、他の表現形式が当該点における音場信号を記述するために使用されると考えられることができる。当該点における信号を同じ精度で記述するために、この記述方法がより少ないデータ量を使用できる場合、信号圧縮の目的が達成できる。

10

【0274】

空間内の音場は、複数の平面波の重ね合わせに分解できる。したがって、HOA信号により表現される音場は、複数の平面波の重ね合わせを使用することにより表現でき、各平面波は、1つのサウンドチャンネル上のオーディオ信号及び方向ベクトルを使用することにより表現される。重ね合わせられた平面波の表現形式が、より少ないサウンドチャンネルを使用することにより元の音場をより良く表現できる場合、信号圧縮が実現できる。

【0275】

実際の再生中に、HOA信号はヘッドセットを使用することにより再生されてもよく、或いは、室内に配置された複数のスピーカーを使用することにより再生されてもよい。スピーカーが再生に使用されるとき、基本的な方法は、複数のスピーカーの音場を重ね合わせることであり、それにより、空間内の或る点(リスナーの位置)における音場が、HOA信号が記録されるとき標準上の元の音場ができるだけ近くなるようにする。この出願の実施形態では、仮想スピーカーレイが使用されると仮定する。次いで、仮想スピーカーレイの再生信号が計算され、再生信号が伝送信号として使用され、圧縮信号が生成される。デコーダは、ビットストリームを復号して再生信号を取得し、再生信号を使用することにより、シーンオーディオ信号を再構成する。

20

【0276】

この出願の実施形態は、シーンオーディオ信号の符号化に適用可能なエンコーダと、シーンオーディオ信号の復号に適用可能なデコーダとを提供する。エンコーダは、元のHOA信号を圧縮されたビットストリームに符号化し、エンコーダは、圧縮されたビットストリームをデコーダに送信し、次いで、デコーダは、圧縮されたビットストリームを再構成されたHOA信号に復元する。この出願のこの実施形態では、エンコーダにより実行された圧縮の後に取得されるデータの量ができるだけ小さくなるか、或いは、同じビットレートでデコーダにより実行された再構成後に取得されるHOA信号の品質がより高くなる。

30

【0277】

この出願のこの実施形態では、HOA信号の符号化中の大きいデータ量、高い帯域幅占有率、低い圧縮効率及び低い符号化品質という問題が解決できる。 N 次HOA信号は $(N+1)^2$ 個のサウンドチャンネルを有するので、HOA信号を直接伝送するには高い帯域幅が消費される必要がある。したがって、効果的なマルチチャンネル符号化方式が必要とされる。

【0278】

40

この出願のこの実施形態では、異なるサウンドチャンネル抽出方法が使用され、この出願の実施形態では音源の仮定は限定されず、時間周波数領域における単一音源の仮定に依存せず、それにより、複数の音源の信号のような複雑なシーンがより効果的に処理できるようにする。この出願のこの実施形態におけるエンコーダ及びデコーダは、元のHOA信号を示すためにより少ないサウンドチャンネルが使用される空間符号化及び復号方法を提供する。図5は、この出願のこの実施形態によるエンコーダの構造の概略図である。エンコーダは、空間エンコーダ及びコアエンコーダを含む。空間エンコーダは、仮想スピーカー信号を生成するために符号化されるべきHOA信号に対してサウンドチャンネル抽出を実行してもよい。コアエンコーダは、仮想スピーカー信号を符号化して、ビットストリームを取得してもよい。エンコーダはビットストリームをデコーダに送信する。図6は、この出願のこ

50

の実施形態によるデコーダの構造の概略図である。デコーダはコアデコーダ及び空間デコーダを含む。コアデコーダは、まずエンコーダからビットストリームを受信し、次いで、ビットストリームを復号して仮想スピーカー信号を取得する。次いで、空間デコーダは、仮想スピーカー信号を再構成して、再構成されたHOA信号を取得する。

【0279】

以下に、エンコーダ及びデコーダからの例を別々に説明する。

【0280】

図7に示すように、まず、この出願のこの実施形態において提供されるエンコーダについて説明する。エンコーダは、仮想スピーカー構成ユニットと、符号化分析ユニットと、仮想スピーカーセット生成ユニットと、仮想スピーカー選択ユニットと、仮想スピーカー信号生成ユニットと、コアエンコーダ処理ユニットと、信号再構成ユニットと、残差信号生成ユニットと、選択ユニットと、信号補償ユニットとを含んでもよい。以下に、エンコーダの各コンポーネントユニットの機能について別々に説明する。この出願のこの実施形態では、図7に示すエンコーダは、1つの仮想スピーカー信号を生成してもよく、或いは、複数の仮想スピーカー信号を生成してもよい。複数の仮想スピーカー信号を生成するプロセスは、図7に示すエンコーダ構造に従って複数回の生成を実行することにより実現されてもよい。以下に、1つの仮想スピーカー信号を生成するプロセスを例として使用する。

【0281】

仮想スピーカー構成ユニットは、仮想スピーカーセット内の仮想スピーカーを構成して、複数の仮想スピーカーを取得するように構成される。

【0282】

仮想スピーカー構成ユニットは、エンコーダの構成情報に基づいて仮想スピーカー構成パラメータを出力する。エンコーダの構成情報は、HOAオーダー、符号化ビットレート及びユーザ定義情報を含むが、これらに限定されない。仮想スピーカー構成パラメータは、仮想スピーカーの数、仮想スピーカーのHOAオーダー及び仮想スピーカーの位置座標を含むが、これらに限定されない。

【0283】

仮想スピーカー構成ユニットにより出力された仮想スピーカー構成パラメータは、仮想スピーカーセット生成ユニットの入力として使用される。

【0284】

符号化分析ユニットは、符号化されるべきHOA信号に対して符号化分析を実行し、例えば、符号化されるべきHOA信号の音源の数、指向性及び分散のような特性を含む、符号化されるべきHOA信号の音場分布を分析するように構成され、これらはターゲット仮想スピーカーをどのように選択するかを決定するための決定条件の1つとして使用される。

【0285】

この出願のこの実施形態では、エンコーダは符号化分析ユニットを含まなくてもよく、すなわち、エンコーダは入力信号を分析しなくてもよく、ターゲット仮想スピーカーをどのように選択するかを決定するためにデフォルト構成が使用される。これは限定されない。

【0286】

エンコーダは、符号化されるべきHOA信号を取得し、例えば、実際の獲得デバイスから記録されたHOA信号、又は人工オーディオオブジェクトを使用することにより合成されたHOA信号を、エンコーダの入力として使用してもよく、エンコーダにより入力された符号化されるべきHOA信号は、時間領域HOA信号又は周波数領域HOA信号でもよい。

【0287】

仮想スピーカーセット生成ユニットは、仮想スピーカーセットを生成するように構成される。仮想スピーカーセットは、複数の仮想スピーカーを含んでもよく、仮想スピーカーセット内の仮想スピーカーはまた、「候補仮想スピーカー」と呼ばれてもよい。

【0288】

仮想スピーカーセット生成ユニットは、指定の候補仮想スピーカーについてのHOA係数を生成する。候補仮想スピーカーについてのHOA係数を生成することは、候補仮想スピー

10

20

30

40

50

カーの座標(すなわち、位置座標又は位置情報)と候補仮想スピーカーのHOAオーダーとを必要とする。候補仮想スピーカーの座標を決定するための方法は、等距離規則に従ってK個の仮想スピーカーを生成し、聴覚知覚原理に従って均等に分布していないK個の候補仮想スピーカーを生成することを含むが、これに限定されない。以下に、均等に分布した固定量の仮想スピーカーを生成するための方法の例を提供する。

【0289】

均等に分布した候補仮想スピーカーの座標は、候補仮想スピーカーの数に基づいて生成され、例えば、数値反復計算法を使用することにより、ほぼ均一なスピーカー配置が提供される。図8は、球上にほぼ均等に分布した仮想スピーカーの概略図である。いくつかの物質粒子が単位球上に分布していると仮定し、これらの物質粒子の間に二次の反比例の反発力が設定され、これは同じ電荷の間での静電反発力と同様である。これらの物質粒子は反発力の下で自由に移動することが可能であり、物質粒子が定常状態に達したとき、物質粒子の分布は均一であることが想定される。計算では、実際の物理法則が簡略化され、物質粒子の運動距離は応力に直接等しくなる。したがって、第iの物質粒子について、反復計算のステップにおける物質粒子の運動距離、すなわち、応力を受けた仮想力は、以下の式を使用することにより計算される。

【数5】

$$\vec{D} = \vec{F} = \sum_{j=1, j \neq i}^N \frac{k}{r_{ij}^2} \vec{d}_{ij}.$$

【0290】

【数6】

$$\vec{D}$$

は変位ベクトルを表し、

【数7】

$$\vec{F}$$

は力ベクトルを表し、 r_{ij} は第iの物質粒子と第jの物質粒子との間の距離を表し、

【数8】

$$\vec{d}_{ij}$$

は第jの物質粒子から第iの物質粒子へ方向ベクトルを表す。パラメータkは単一のステップのサイズを制御する。物質粒子の初期位置はランダムに指定される。

【0291】

変位ベクトル

【数9】

$$\vec{D}$$

に従って移動した後に、通常では物質粒子は単位球から逸脱する。次の反復の前に、物質粒子と球の中心との間の距離が正規化され、物質粒子は単位球に戻される。したがって、図8に示す仮想スピーカーの分布の概略図が取得されてもよく、ここで、複数の仮想スピーカーが球上にほぼ均等に分布している。

【0292】

次に、候補仮想スピーカーについてのHOA係数が生成される。理想的な平面波が球面調和関数により展開された後に、振幅がsであり且つスピーカーの位置座標が (θ_s, ϕ_s) である理想的な平面波の形は、以下の計算式となる。

【数10】

10

20

30

40

50

$$p(r, \theta, \varphi, k) = \sum_{m=0}^{\infty} (2m+1) j_m^{kr} (kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} Y_{m,n}^{\sigma}(\theta_s, \varphi_s) Y_{m,n}^{\sigma}(\theta, \varphi)$$

【0293】

平面波についてのHOA係数は $B_{m,n}^{\sigma}$ であり、以下の計算式を満たす。

$$B_{m,n}^{\sigma} = s \cdot Y_{m,n}^{\sigma}(\theta_s, \varphi_s)$$

【0294】

仮想スピーカーセット生成ユニットにより出力された候補仮想スピーカーのHOA係数は、仮想スピーカー選択ユニットの入力として使用される。

【0295】

仮想スピーカー選択ユニットは、符号化されるべきHOA信号に基づいて仮想スピーカーセット内の複数の候補仮想スピーカーからターゲット仮想スピーカーを選択するように構成される。ターゲット仮想スピーカーは、「符号化されるべきHOA信号に一致する仮想スピーカー」と呼ばれてもよく、或いは、略して一致する仮想スピーカーと呼ばれてもよい。

【0296】

仮想スピーカー選択ユニットは、符号化されるべきHOA信号と仮想スピーカーセット生成ユニットにより出力された候補仮想スピーカーのHOA係数とを照合し、指定の一致する仮想スピーカーを選択する。

【0297】

以下に、仮想スピーカーを選択するための方法について例を使用することにより説明する。実施形態では、候補仮想スピーカーが取得された後に、符号化されるべきHOA信号は、仮想スピーカーセット生成ユニットにより出力された候補仮想スピーカーのHOA係数と照合されて、候補仮想スピーカーに対して符号化されるべきHOA信号の最適な一致を見つけ、候補仮想スピーカーのHOA係数に基づいて、符号化されるべきHOA信号を照合して組み合わせることを目的とする。実施形態では、候補仮想スピーカーのHOA係数と符号化されるべきHOA信号との間で内積が実行され、内積の最大の絶対値を有する候補仮想スピーカーがターゲット仮想スピーカー、すなわち、一致する仮想スピーカーとして選択され、候補仮想スピーカーに対して符号化されるべきHOA信号の射影が、候補仮想スピーカーのHOA係数の線形結合に重ね合わされ、次いで、符号化されるべきHOA信号から射影ベクトルが減算されて差を取得する。差について上記のプロセスが繰り返されて反復計算を実現し、反復のたびに一致する仮想スピーカーが生成され、一致する仮想スピーカーの座標及びターゲット仮想スピーカーのHOA係数が出力される。複数の一致する仮想スピーカーが選択され、反復のたびに1つの一致する仮想スピーカーが生成されることが理解され得る。

【0298】

仮想スピーカー選択ユニットにより出力されたターゲット仮想スピーカーの座標及びターゲット仮想スピーカーについてのHOA係数は、仮想スピーカー信号生成ユニットの入力として使用される。

【0299】

この出願のいくつかの実施形態では、図7に示す構成ユニットに加えて、エンコーダはサイド情報生成ユニットを更に含んでもよい。エンコーダは、サイド情報生成ユニットを含まなくてもよく、これはここでは単なる例である。これは限定されない。

【0300】

仮想スピーカー選択ユニットにより出力されたターゲット仮想スピーカーの座標及び/又はターゲット仮想スピーカーについてのHOA係数は、サイド情報生成ユニットの入力として使用される。

【0301】

サイド情報生成ユニットは、ターゲット仮想スピーカーについてのHOA係数又はターゲ

10

20

30

40

50

ット仮想スピーカークの座標をサイド情報に変換し、これは、コアエンコーダによる処理及び伝送を容易にする。

【0302】

サイド情報生成ユニットの出力は、コアエンコーダ処理ユニットの入力として使用される。

【0303】

仮想スピーカーク信号生成ユニットは、符号化されるべきHOA信号及びターゲット仮想スピーカークの属性情報に基づいて仮想スピーカーク信号を生成するように構成される。

【0304】

仮想スピーカーク信号生成ユニットは、符号化されるべきHOA信号及びターゲット仮想スピーカークについてのHOA係数を使用することにより仮想スピーカーク信号を計算する。

【0305】

ターゲット仮想スピーカークについてのHOA係数は行列Aにより表され、符号化されるべきHOA信号は、行列Aを使用することにより線形結合を通じて取得できる。理論的な最適解w、すなわち、仮想スピーカーク信号は、最小二乗法を使用することにより取得できる。例えば、以下の計算式が使用されてもよい。

$$w = A^{-1}X$$

ここで、 A^{-1} は行列Aの逆行列であり、行列Aのサイズは $(M \times C)$ であり、Cはターゲット仮想スピーカークの数であり、MはN次HOA係数のサウンドチャネルの数であり、aはターゲット仮想スピーカークについてのHOA係数を表す。例えば、

【数11】

$$A = \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1C} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ a_{M1} & \cdot & \cdot & \cdot & a_{MC} \end{bmatrix}$$

である。

【0306】

Xは符号化されるべきHOA信号を表し、行列Xのサイズは $(M \times L)$ であり、MはN次HOA係数のサウンドチャネルの数であり、Lはサンプリング点の数であり、xは符号化されるべきHOA信号についての係数を表す。例えば、

【数12】

$$X = \begin{bmatrix} x_{11} & \cdot & \cdot & \cdot & x_{1L} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ x_{M1} & \cdot & \cdot & \cdot & x_{ML} \end{bmatrix}$$

である。

【0307】

仮想スピーカーク信号生成ユニットにより出力された仮想スピーカーク信号は、コアエンコーダ処理ユニットの入力として使用される。

【0308】

この出願のいくつかの実施形態では、図7に示す構成ユニットに加えて、エンコーダは信号整列ユニットを更に含んでもよい。エンコーダは信号整列ユニットを含まなくてもよ

く、これはここでは単なる例である。これは限定されない。

【0309】

仮想スピーカ信号生成ユニットにより出力された仮想スピーカ信号は、信号整列ユニットの入力として使用される。

【0310】

信号整列ユニットは、仮想スピーカ信号のサウンドチャンネルを再調整して、チャンネル間相関を強化し、コアエンコーダによる処理を容易にするように構成される。

【0311】

信号整列ユニットにより出力された整列された仮想スピーカ信号は、コアエンコーダ処理ユニットの入力である。

【0312】

信号再構成ユニットは、仮想スピーカ信号及びターゲット仮想スピーカについてのHOA係数を使用することにより、HOA信号を再構成するように構成される。

【0313】

ターゲット仮想スピーカについてのHOA係数の構成は行列Aにより表され、行列Aのサイズは $(M \times C)$ であり、行列は、Cが一致する仮想スピーカの数であり、MがN次HOA係数のサウンドチャンネルの数であることによって示される。仮想スピーカ信号は行列Wにより表され、行列Wのサイズは $(C \times L)$ であり、Lは信号サンプリング点の数を表す。したがって、再構成されたHOA信号は、

$$T = AW$$

である。

【0314】

信号再構成ユニットにより出力された再構成されたHOA信号は、残差信号生成ユニットの入力である。

【0315】

残差信号生成ユニットは、符号化されるべきHOA信号及び信号再構成ユニットにより出力された再構成されたHOA信号を使用することにより、残差信号を計算するように構成される。例えば、計算方法は、符号化されるべきHOA信号と、信号再構成ユニットにより出力された再構成されたHOA信号に対応するサウンドチャンネル内の対応するサンプリング点との間の差を取得することである。

【0316】

残差信号生成ユニットにより出力された残差信号は、信号補償ユニット及び選択ユニットの入力である。

【0317】

選択ユニットは、エンコーダの構成情報及び信号クラス情報に基づいて仮想スピーカ信号及び/又は残差信号を選択するように構成され、例えば、選択は仮想スピーカ信号の選択及び残差信号の選択を含む。

【0318】

例えば、サウンドチャンネルの数を低減するために、M個未満のサウンドチャンネルを有する残差信号が、符号化されるべき残差信号として選択されてもよい。低次の残差信号が、符号化されるべき残差信号として選択されてもよく、或いは、高エネルギーを有する残差信号が、符号化されるべき残差信号として選択されてもよい。

【0319】

選択ユニットにより出力された残差信号は、コアエンコーダ処理ユニットの入力及び信号補償ユニットの入力である。

【0320】

信号補償ユニットは、M個のサウンドチャンネルを有する残差信号が符号化されるべき残差信号として機能することに比較して、M個未満のサウンドチャンネルを有する残差信号が符号化されるべき残差信号として選択されるときに信号ロスが発生するので、伝送されない残差信号に対して信号補償を実行するように構成される。信号補償は、情報ロス分析、

10

20

30

40

50

エネルギー補償、エンベロープ補償及びノイズ補償でもよいが、これらに限定されない。補償方法は、線形補償、非線形補償等でもよい。信号補償ユニットは、信号補償のためのサイド情報を生成する。

【0321】

コアエンコーダ処理ユニットは、サイド情報及び整列された仮想スピーカ信号に対してコアエンコーダ処理を実行して、伝送のためにビットストリームを取得するように構成される。

【0322】

コアエンコーダ処理は、変換、量子化、心理音響モデル及びビットストリーム生成を含むが、これらに限定されず、周波数領域のサウンドチャンネル又は時間領域のサウンドチャンネルを処理してもよく、これはここでは限定されない。

【0323】

図9に示すように、この出願のこの実施形態において提供されるデコーダは、コアデコーダ処理ユニットと、HOA信号再構成ユニットとを含んでもよい。

【0324】

コアデコーダ処理ユニットは、伝送のためにビットストリームに対してコアデコーダ処理を実行して、仮想スピーカ信号及び残差信号を取得するように構成される。

【0325】

エンコーダがサイド情報をビットストリームに追加する場合、デコーダはサイド情報復号ユニットを更に含む必要がある。これは限定されない。

【0326】

サイド情報復号ユニットは、コアデコーダ処理ユニットにより出力された復号対象のサイド情報を復号して、復号されたサイド情報を取得するように構成される。

【0327】

コアデコーダ処理は、変換、ビットストリーム分析及び量子化解除を含み、周波数領域のサウンドチャンネル又は時間領域のサウンドチャンネルを処理してもよく、これはここでは限定されない。

【0328】

コアデコーダ処理ユニットにより出力された仮想スピーカ信号及び残差信号は、HOA信号再構成ユニットの入力として使用され、コアデコーダ処理ユニットにより出力された復号されたサイド情報は、サイド情報復号ユニットの入力である。

【0329】

サイド情報復号ユニットは、復号されたサイド情報をターゲット仮想スピーカについてのHOA係数に変換する。

【0330】

サイド情報復号ユニットにより出力されたターゲット仮想スピーカについてのHOA係数は、HOA信号再構成ユニットの入力である。

【0331】

HOA信号再構成ユニットは、残差信号及びターゲット仮想スピーカについてのHOA係数を使用することにより仮想スピーカ信号を再構成して、再構成されたHOA信号を取得するように構成される。

【0332】

ターゲット仮想スピーカについてのHOA係数は行列 A' により表される。行列 A' のサイズは $(M \times C)$ であり、行列は A' により表され、 C はターゲット仮想スピーカの数であり、 M は N 次HOA係数のサウンドチャンネルの数である。仮想スピーカ信号の構成は行列 W' により示される $(C \times L)$ 行列であり、 L は信号サンプリング点の数を表す。再構成されたHOA信号 H は、以下の式を使用することにより取得される。

$$H = A'W'$$

ここで、信号再構成ユニットにより出力された再構成されたHOA信号は、デコーダの出力である。

10

20

30

40

50

【0333】

この出願のいくつかの実施形態では、エンコーダのビットストリームが信号補償に使用されるサイド情報を更に搬送する場合、デコーダは、

再構成されたHOA信号及び残差信号を合成して、合成されたHOA信号を取得するように構成された信号補償ユニットを更に含んでもよい。合成されたHOA信号は、信号補償に使用されるサイド情報を使用することにより調整されて、再構成されたHOA係数を取得する。

【0334】

この出願のこの実施形態では、エンコーダは、より少ないサウンドチャンネルを使用することにより元のHOA信号を表すために空間エンコーダを使用してもよい。例えば、元の3次HOA信号について、この出願のこの実施形態における空間エンコーダは、16個のサウンドチャンネルを4つのサウンドチャンネルに圧縮し、主観的なリスニングが明らかに異なることを確保できる。主観的なリスニングテストは、オーディオ符号化及び復号における評価基準である。明らかな違いがないことは主観的な評価のレベルである。

【0335】

この出願のいくつかの他の実施形態では、エンコーダの仮想スピーカー選択ユニットは、仮想スピーカーセットからターゲット仮想スピーカーを選択するか、或いは、指定の方向及び位置における仮想スピーカーをターゲット仮想スピーカーとして使用してもよく、仮想スピーカー信号生成ユニットは、各ターゲット仮想スピーカーに対して投影を直接実行して、仮想スピーカー信号を取得する。

【0336】

上記のように、指定の方向及び位置における仮想スピーカーは、ターゲット仮想スピーカーとして使用される。これは、仮想スピーカーの選択プロセスを簡略化し、符号化及び復号速度を改善できる。

【0337】

この出願のいくつかの他の実施形態では、エンコーダは、信号整列ユニットを含まなくてもよい。この場合、仮想スピーカー信号生成ユニットの出力は、コアエンコーダにより直接符号化される。上記の方式は、信号整列処理を低減し、エンコーダの複雑さが低減される。

【0338】

上記の例の説明から、この出願の実施形態では、選択されたターゲット仮想スピーカーがHOA信号の符号化及び復号に適用されることが分かる。この出願の実施形態では、HOA信号の音源の正確な位置が取得でき、HOA信号を再構成するための方向がより正確になり、符号化効率がより高くなり、デコーダの複雑さが非常に低くなる。これはモバイル端末での適用に有益であり、符号化及び復号のパフォーマンスを改善できる。

【0339】

簡単な説明のために、上記の方法の実施形態は一連のアクションとして表される点に留意すべきである。しかし、この出願によれば、いくつかのステップは他の順序で或いは同時に実行されてもよいので、この出願は記載のアクションの順序に限定されないことを当業者は認識すべきである。この明細書に記載されている実施形態は全て例示的な実施形態に属しており、関与するアクション及びモジュールは必ずしもこの出願で必要とされないことは、当業者により更に認識されるべきである。

【0340】

この出願の実施形態の解決策をより良く実現するために、解決策を実現するための関連装置が以下に更に提供される。

【0341】

図10に示すように、この出願の実施形態において提供されるオーディオ符号化装置1000は、取得モジュール1001と、信号生成モジュール1002と、符号化モジュール1003とを含んでもよい。

【0342】

取得モジュールは、第1のシーンオーディオ信号に基づいて予め設定された仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するように構成される。

【0343】

信号生成モジュールは、第1のシーンオーディオ信号及び第1のターゲット仮想スピーカーの属性情報に基づいて仮想スピーカー信号を生成するように構成される。

【0344】

信号生成モジュールは、第1のターゲット仮想スピーカーの属性情報及び第1の仮想スピーカー信号を使用することにより、第2のシーンオーディオ信号を取得するように構成される。

【0345】

信号生成モジュールは、第1のシーンオーディオ信号及び第2のシーンオーディオ信号に基づいて残差信号を生成するように構成される。

【0346】

符号化モジュールは、仮想スピーカー信号及び残差信号を符号化して、ビットストリームを取得するように構成される。

【0347】

この出願のいくつかの実施形態では、取得モジュールは、仮想スピーカーセットに基づいて第1のシーンオーディオ信号から主要音場成分を取得し、主要音場成分に基づいて仮想スピーカーセットから第1のターゲット仮想スピーカーを選択するように構成される。

【0348】

この出願のいくつかの実施形態では、取得モジュールは、主要音場成分に基づいて高次アンビソニックス(HOA)係数セットから主要音場成分についてのHOA係数を選択するように構成され、ここで、HOA係数セット内のHOA係数は、仮想スピーカーセット内の仮想スピーカーと1対1の対応関係にあり、仮想スピーカーセットの中で主要音場成分についてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するように構成される。

【0349】

この出願のいくつかの実施形態では、取得モジュールは、主要音場成分に基づいて第1のターゲット仮想スピーカーの構成パラメータを取得し、第1のターゲット仮想スピーカーの構成パラメータに基づいて第1のターゲット仮想スピーカーについてのHOA係数を生成し、仮想スピーカーセットの中で第1のターゲット仮想スピーカーについてのHOA係数に対応する仮想スピーカーを第1のターゲット仮想スピーカーとして決定するように構成される。

【0350】

この出願のいくつかの実施形態では、取得モジュールは、オーディオエンコーダの構成情報に基づいて仮想スピーカーセット内の複数の仮想スピーカーの構成パラメータを決定し、主要音場成分に基づいて複数の仮想スピーカーの構成パラメータから第1のターゲット仮想スピーカーの構成パラメータを選択するように構成される。

【0351】

この出願のいくつかの実施形態では、第1のターゲット仮想スピーカーの構成パラメータは、第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報を含む。

【0352】

取得モジュールは、第1のターゲット仮想スピーカーの位置情報及びHOAオーダー情報に基づいて第1のターゲット仮想スピーカーについてのHOA係数を決定するように構成される。

【0353】

この出願のいくつかの実施形態では、符号化モジュールは、第1のターゲット仮想スピーカーの属性情報を符号化し、符号化された情報をビットストリームに書き込むように更に構成される。

【0354】

10

20

30

40

50

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークについてのHOA係数を含む。

【0355】

信号生成モジュールは、符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカーク信号を取得するように構成される。

【0356】

この出願のいくつかの実施形態では、第1のシーンオーディオ信号は、符号化されるべき高次アンビソニックス(HOA)信号を含み、第1のターゲット仮想スピーカークの属性情報は、第1のターゲット仮想スピーカークの位置情報を含む。

10

【0357】

信号生成モジュールは、第1のターゲット仮想スピーカークの位置情報に基づいて第1のターゲット仮想スピーカークについてのHOA係数を取得し、符号化されるべきHOA信号及び第1のターゲット仮想スピーカークについてのHOA係数に対して線形結合を実行して、第1の仮想スピーカーク信号を取得するように構成される。

【0358】

この出願のいくつかの実施形態では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカークセットから第2のターゲット仮想スピーカークを選択するように構成される。

20

【0359】

信号生成モジュールは、第1のシーンオーディオ信号及び第2のターゲット仮想スピーカークの属性情報に基づいて第2の仮想スピーカーク信号を生成するように構成される。

【0360】

符号化モジュールは、第2の仮想スピーカーク信号を符号化し、符号化された信号をビットストリームに書き込むように構成される。

【0361】

対応して、信号生成モジュールは、第1のターゲット仮想スピーカークの属性情報、第1の仮想スピーカーク信号、第2のターゲット仮想スピーカークの属性情報及び第2の仮想スピーカーク信号に基づいて第2のシーンオーディオ信号を取得するように構成される。

30

【0362】

この出願のいくつかの実施形態では、信号生成モジュールは、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号を整列させて、整列された第1の仮想スピーカーク信号及び整列された第2の仮想スピーカーク信号を取得するように構成される。

【0363】

対応して、符号化モジュールは、整列された第2の仮想スピーカーク信号を符号化するように構成される。

【0364】

対応して、符号化モジュールは、整列された第1の仮想スピーカーク信号及び残差信号を符号化するように構成される。

40

【0365】

この出願のいくつかの実施形態では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカークセットから第2のターゲット仮想スピーカークを選択するように構成される。

【0366】

信号生成モジュールは、第1のシーンオーディオ信号及び第2のターゲット仮想スピーカークの属性情報に基づいて第2の仮想スピーカーク信号を生成するように構成される。

【0367】

対応して、符号化モジュールは、第1の仮想スピーカーク信号及び第2の仮想スピーカーク信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得するように構成さ

50

れる。第1のサイド情報は第1の仮想スピーカ信号と第2の仮想スピーカ信号との間の関係を示す。

【0368】

対応して、符号化モジュールは、ダウンミキシングされた信号、第1のサイド情報及び残差信号を符号化するように構成される。

【0369】

この出願のいくつかの実施形態では、信号生成モジュールは、第1の仮想スピーカ信号及び第2の仮想スピーカ信号を整列させて、整列された第1の仮想スピーカ信号及び整列された第2の仮想スピーカ信号を取得するように構成される。

【0370】

符号化モジュールは、整列された第1の仮想スピーカ信号及び整列された第2の仮想スピーカ信号に基づいてダウンミキシングされた信号及び第1のサイド情報を取得するように構成される。

【0371】

対応して、第1のサイド情報は、整列された第1の仮想スピーカ信号と整列された第2の仮想スピーカ信号との間の関係を示す。

【0372】

この出願のいくつかの実施形態では、取得モジュールは、第1のシーンオーディオ信号に基づいて仮想スピーカセットから第2のターゲット仮想スピーカを選択する前に、第1のシーンオーディオ信号の符号化率及び/又は信号クラス情報に基づいて、第1のターゲット仮想スピーカ以外のターゲット仮想スピーカが取得される必要があるか否かを決定し、第1のターゲット仮想スピーカ以外のターゲット仮想スピーカが取得される必要がある場合にのみ、第1のシーンオーディオ信号に基づいて仮想スピーカセットから第2のターゲット仮想スピーカを選択するように構成される。

【0373】

この出願のいくつかの実施形態では、残差信号は、少なくとも2つのサウンドチャンネル上の残差サブ信号を含む。

【0374】

信号生成モジュールは、オーディオエンコーダの構成情報及び/又は第1のシーンオーディオ信号の信号クラス情報に基づいて、少なくとも2つのサウンドチャンネル上の残差サブ信号から、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を決定するように構成される。

【0375】

対応して、符号化モジュールは、第1の仮想スピーカ信号と、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号とを符号化するように構成される。

【0376】

この出願のいくつかの実施形態では、取得モジュールは、少なくとも2つのサウンドチャンネル上の残差サブ信号が、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号を含む場合、第2のサイド情報を取得するように構成される。第2のサイド情報は、符号化される必要があり且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号と、符号化される必要がなく且つ少なくとも1つのサウンドチャンネル上にある残差サブ信号との間の関係を示す。

【0377】

対応して、符号化モジュールは、第2のサイド情報をビットストリームに書き込むように構成される。

【0378】

図11に示すように、この出願の実施形態において提供されるオーディオ復号装置1100は、受信モジュール1101と、復号モジュール1102と、再構成モジュール1103とを含んでもよい。

10

20

30

40

50

【0379】

受信モジュールは、ビットストリームを受信するように構成される。

【0380】

復号モジュールは、ビットストリームを復号して、仮想スピーカー信号及び残差信号を取得するように構成される。

【0381】

再構成モジュールは、ターゲット仮想スピーカーの属性情報、残差信号及び仮想スピーカー信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0382】

この出願のいくつかの実施形態では、復号モジュールは、ビットストリームを復号して、ターゲット仮想スピーカーの属性情報を取得するように更に構成される。

10

【0383】

この出願のいくつかの実施形態では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーについての高次アンビソニックス(HOA)係数を含む。

【0384】

再構成モジュールは、仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するように構成される。

【0385】

この出願のいくつかの実施形態では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーの位置情報を含む。

20

【0386】

再構成モジュールは、仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するように構成される。

【0387】

この出願のいくつかの実施形態では、ターゲット仮想スピーカーの属性情報は、ターゲット仮想スピーカーの位置情報を含む。

30

【0388】

再構成モジュールは、ターゲット仮想スピーカーの位置情報に基づいてターゲット仮想スピーカーについてのHOA係数を決定し、仮想スピーカー信号及びターゲット仮想スピーカーについてのHOA係数に対して合成処理を実行して、合成されたシーンオーディオ信号を取得し、残差信号を使用することにより、合成されたシーンオーディオ信号を調整して、再構成されたシーンオーディオ信号を取得するように構成される。

【0389】

この出願のいくつかの実施形態では、図11に示すように、仮想スピーカー信号は、第1の仮想スピーカー信号及び第2の仮想スピーカー信号をダウンミキシングすることにより取得されたダウンミキシングされた信号である。当該装置1100は第1の信号補償モジュール1104を更に含む。

40

【0390】

復号モジュールは、ビットストリームを復号して、第1のサイド情報を取得するように構成される。第1のサイド情報は、第1の仮想スピーカー信号と第2の仮想スピーカー信号との間の関係を示す。

【0391】

第1の信号補償モジュールは、第1のサイド情報及びダウンミキシングされた信号に基づいて第1の仮想スピーカー信号及び第2の仮想スピーカー信号を取得するように構成される。

【0392】

50

対応して、再構成モジュールは、ターゲット仮想スピーカ属性情報、残差信号、第1の仮想スピーカ信号及び第2の仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0393】

この出願のいくつかの実施形態では、図11に示すように、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含む。当該装置1100は第2の信号補償モジュール1105を更に含む。

【0394】

復号モジュールは、ビットストリームを復号して、第2のサイド情報を取得するように構成される。第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第2のサウンドチャンネル上の残差サブ信号との間の関係を示す。

【0395】

第2の信号補償モジュールは、第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第2のサウンドチャンネル上の残差サブ信号を取得するように構成される。

【0396】

対応して、再構成モジュールは、ターゲット仮想スピーカ属性情報、第1のサウンドチャンネル上の残差サブ信号、第2のサウンドチャンネル上の残差サブ信号及び仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0397】

この出願のいくつかの実施形態では、図11に示すように、残差信号は、第1のサウンドチャンネル上の残差サブ信号を含む。当該装置1100は第3の信号補償モジュール1106を更に含む。

【0398】

復号モジュールは、ビットストリームを復号して、第2のサイド情報を取得するように構成される。第2のサイド情報は、第1のサウンドチャンネル上の残差サブ信号と第3のサウンドチャンネル上の残差サブ信号との間の関係を示す。

【0399】

第3の信号補償モジュールは、第2のサイド情報及び第1のサウンドチャンネル上の残差サブ信号に基づいて第3のサウンドチャンネル上の残差サブ信号及び第1のサウンドチャンネル上の更新された残差サブ信号を取得するように構成される。

【0400】

対応して、再構成モジュールは、ターゲット仮想スピーカ属性情報、第1のサウンドチャンネル上の更新された残差サブ信号、第3のサウンドチャンネル上の残差サブ信号及び仮想スピーカ信号に基づいて再構成されたシーンオーディオ信号を取得するように構成される。

【0401】

装置のモジュール/ユニットの間の情報交換及びその実行プロセスのような内容は、この出願の方法の実施形態と同じ概念に基づいており、この出願の方法の実施形態と同じ技術的效果を生じる点に留意すべきである。具体的な内容については、この出願の方法の実施形態における上記の説明を参照し、詳細はここでは再び説明しない。

【0402】

この出願の実施形態は、コンピュータ記憶媒体を更に提供する。コンピュータ記憶媒体はプログラムを記憶し、プログラムは上記の方法の実施形態において記載されるステップの一部又は全部を実行する。

【0403】

以下に、この出願の実施形態において提供される他のオーディオ符号化装置について説明する。図12に示すように、オーディオ符号化装置1200は、受信機1201と、送信機1202と、プロセッサ1203と、メモリ1204とを含む(オーディオ符号化装置1200には1つ以上のプロセッサ1203が存在してもよく、1つのプロセッサが図12における例として使用される)。この出願のいくつかの実施形態では、受信機1201、送信機1202、プロセッ

10

20

30

40

50

サ1203及びメモリ1204は、バスを通じて或いは他の方式で接続されてもよい。図12では、バスを通じた接続が例として使用される。

【0404】

メモリ1204は、読み取り専用メモリ及びランダムアクセスメモリを含み、命令及びデータをプロセッサ1203に提供してもよい。メモリ1204の一部は、不揮発性ランダムアクセスメモリ(non-volatile random access memory, NVRAM)を更に含んでもよい。メモリ1204は、オペレーティングシステム及び動作命令、実行可能モジュール若しくはデータ構造、又はこれらのサブセット、又はこれらの拡張セットを記憶する。動作命令は、様々な動作を実行するために使用される様々な動作命令を含んでもよい。オペレーティングシステムは、様々な基本サービスを実現し、ハードウェアベースのタスクを処理するための様々なシステムプログラムを含んでもよい。

10

【0405】

プロセッサ1203はオーディオ符号化装置の動作を制御し、プロセッサ1203はまた、中央処理装置(central processing unit, CPU)とも呼ばれてもよい。特定の用途では、オーディオ符号化装置のコンポーネントはバスシステムを通じて一緒に結合される。データバスに加えて、バスシステムは、電力バス、制御バス、状態信号バス等を更に含んでもよい。しかし、明確な説明のために、図面における様々なタイプのバスがバスシステムとして記されている。

【0406】

この出願の実施形態において開示される方法は、プロセッサ1203に適用されてもよく、或いは、プロセッサ1203を使用することにより実現されてもよい。プロセッサ1203は集積回路チップでもよく、信号処理能力を有する。実現プロセスでは、上記の方法におけるステップは、プロセッサ1203内のハードウェアの集積論理回路又はソフトウェアの形式の命令を使用することにより完了されてもよい。プロセッサ1203は、汎用プロセッサ、デジタルシグナルプロセッサ(digital signal processor, DSP)、特定用途向け集積回路(application-specific integrated circuit, ASIC)、フィールドプログラマブルゲートアレイ(field-programmable gate array, FPGA)若しくは他のプログラマブルロジックデバイス、ディスクリートゲート若しくはトランジスタロジックデバイス、又はディスクリートハードウェアコンポーネントでもよい。これは、この出願の実施形態で開示される方法、ステップ及び論理ブロック図を実現又は実行してもよい。汎用プロセッサはマイクロプロセッサでもよく、或いは、プロセッサは、代替としていずれかの従来のプロセッサ等でもよい。この出願の実施形態を参照して開示される方法のステップは、ハードウェアデコーディングプロセッサにより直接実行されて達成されてもよく、或いは、デコーディングプロセッサにおいてハードウェアとソフトウェアモジュールとの組み合わせを使用することにより実行されて達成されてもよい。ソフトウェアモジュールは、ランダムアクセスメモリ、フラッシュメモリ、読み取り専用メモリ、プログラム可能読み取り専用メモリ、電氣的消去可能プログラム可能メモリ又はレジスタのような、当技術分野における成熟した記憶媒体に位置してもよい。記憶媒体はメモリ1204に位置し、プロセッサ1203はメモリ1204内の情報を読み取り、プロセッサのハードウェアと組み合わせて上記の方法のステップを完了する。

20

30

40

【0407】

受信機1201は、入力デジタル又は文字情報を受信し、オーディオ符号化装置の関連する設定及び機能制御に関連する信号入力を生成するように構成されてもよい。送信機1202は、表示画面のような表示デバイスを含んでもよく、送信機1202は、外部インタフェースを通じてデジタル又は文字情報を出力するように構成されてもよい。

【0408】

この出願のこの実施形態では、プロセッサ1203は、図4に示す上記の実施形態においてオーディオ符号化装置により実行されるオーディオ符号化方法を実行するように構成される。

【0409】

50

以下に、この出願の実施形態において提供される他のオーディオ復号装置について説明する。図13に示すように、オーディオ復号装置1300は、受信機1301と、送信機1302と、プロセッサ1303と、メモリ1304とを含む(オーディオ復号装置1300には1つ以上のプロセッサ1303が存在してもよく、1つのプロセッサが図13における例として使用される)。この出願のいくつかの実施形態では、受信機1301、送信機1302、プロセッサ1303及びメモリ1304は、バスを通じて或いは他の方式で接続されてもよい。図13では、バスを通じた接続が例として使用される。

【0410】

メモリ1304は、読み取り専用メモリ及びランダムアクセスメモリを含み、命令及びデータをプロセッサ1303に提供してもよい。メモリ1304の一部は、NVRAMを更に含んでもよい。メモリ1304は、オペレーティングシステム及び動作命令、実行可能モジュール若しくはデータ構造、又はこれらのサブセット、又はこれらの拡張セットを記憶する。動作命令は、様々な動作を実行するために使用される様々な動作命令を含んでもよい。オペレーティングシステムは、様々な基本サービスを実現し、ハードウェアベースのタスクを処理するための様々なシステムプログラムを含んでもよい。

【0411】

プロセッサ1303はオーディオ復号装置の動作を制御し、プロセッサ1303はまた、CPUとも呼ばれてもよい。特定の用途では、オーディオ復号装置のコンポーネントはバスシステムを通じて一緒に結合される。データバスに加えて、バスシステムは、電力バス、制御バス、状態信号バス等を更に含んでもよい。しかし、明確な説明のために、図面における様々なタイプのバスがバスシステムとして記されている。

【0412】

この出願の実施形態において開示される方法は、プロセッサ1303に適用されてもよく、或いは、プロセッサ1303を使用することにより実現されてもよい。プロセッサ1303は集積回路チップでもよく、信号処理能力を有する。実現プロセスでは、上記の方法におけるステップは、プロセッサ1303内のハードウェアの集積論理回路又はソフトウェアの形式の命令を使用することにより完了されてもよい。プロセッサ1303は、汎用プロセッサ、DSP、ASIC、FPGA若しくは他のプログラマブルロジックデバイス、ディスクリートゲート若しくはトランジスタロジックデバイス、又はディスクリートハードウェアコンポーネントでもよい。これは、この出願の実施形態で開示される方法、ステップ及び論理ブロック図を実現又は実行してもよい。汎用プロセッサはマイクロプロセッサでもよく、或いは、プロセッサは、代替としていずれかの従来のプロセッサ等でもよい。この出願の実施形態を参照して開示される方法のステップは、ハードウェアデコーディングプロセッサにより直接実行されて達成されてもよく、或いは、デコーディングプロセッサにおいてハードウェアとソフトウェアモジュールとの組み合わせを使用することにより実行されて達成されてもよい。ソフトウェアモジュールは、ランダムアクセスメモリ、フラッシュメモリ、読み取り専用メモリ、プログラム可能読み取り専用メモリ、電氣的消去可能プログラム可能メモリ又はレジスタのような、当技術分野における成熟した記憶媒体に位置してもよい。記憶媒体はメモリ1304に位置し、プロセッサ1303はメモリ1304内の情報を読み取り、プロセッサのハードウェアと組み合わせて上記の方法のステップを完了する。

【0413】

この出願のこの実施形態では、プロセッサ1303は、図4に示す上記の実施形態においてオーディオ復号装置により実行されるオーディオ復号方法を実行するように構成される。

【0414】

他の可能な設計では、オーディオ符号化装置又はオーディオ復号装置が端末内のチップである場合、チップは処理ユニット及び通信ユニットを含む。処理ユニットは、例えばプロセッサでもよい。通信ユニットは、例えば、入出力インタフェース、ピン又は回路でもよい。処理ユニットは、記憶ユニットに記憶されたコンピュータ実行可能命令を実行して、端末内のチップが第1の態様のいずれかにおけるオーディオ符号化方法、又は第2の態様のいずれかにおけるオーディオ復号方法を実行することを可能にしてもよい。任意選択で

、記憶ユニットはチップ内の記憶ユニット、例えば、レジスタ又はキャッシュである。代替として、記憶ユニットは端末内にあり且つチップ外にある記憶ユニット、例えば、読み取り専用メモリ(read-only memory, ROM)、静的な情報及び命令を記憶できる他のタイプの静的記憶装置、又はランダムアクセスメモリ(random access memory, RAM)でもよい。

【0415】

上記のいずれかに記載されるプロセッサは、汎用中央処理装置、マイクロプロセッサ、ASIC、又は第1の態様若しくは第2の態様における方法のプログラム実行を制御するように構成された1つ以上の集積回路でもよい。

【0416】

さらに、記載の装置の実施形態は単なる例である点に留意すべきである。別々の部分として記載されるユニットは、物理的に分離していてもよく或いは分離してなくてもよく、ユニットとして表示される部分は、物理的なユニットでもよく或いは物理的なユニットでなくてもよく、1つの場所に位置してもよく、或いは、複数のネットワークユニットに分散されてもよい。モジュールの一部又は全部は、実施形態における解決策の目的を達成するために、実際のニーズに基づいて選択されてもよい。さらに、この出願により提供される装置の実施形態の添付図面では、モジュールの間の接続関係は、モジュールが相互に通信接続を有することを示しており、これは1つ以上の通信バス又は信号ケーブルとして具体的に実現されてもよい。

【0417】

上記の実現方式の説明に基づいて、当業者は、この出願が必要なユニバーサルハードウェアに加えてソフトウェアにより実現されてもよく、或いは、専用集積回路、専用CPU、専用メモリ、専用コンポーネント等を含む専用ハードウェアにより実現されてもよいことを明確に理解し得る。一般的に、コンピュータプログラムにより実行できるいずれかの機能は、対応するハードウェアを使用することにより容易に実現できる。さらに、同じ機能を達成するために使用される特定のハードウェア構造は様々な形式になってもよく、例えば、アナログ回路、デジタル回路又は専用回路の形式になってもよい。しかし、この出願に関しては、ほとんどの場合、ソフトウェアプログラムの実現方式がより良い実現方式である。このような理解に基づいて、本質的にこの出願の技術的解決策又は従来技術に寄与する部分はソフトウェア製品の形式で実現されてもよい。コンピュータソフトウェア製品は、読み取り可能記憶媒体、例えば、コンピュータのフロッピーディスク、USBフラッシュドライブ、取り外し可能ハードディスク、ROM、RAM、磁気ディスク又は光ディスクに記憶され、コンピュータデバイス(パーソナルコンピュータ、サーバ、ネットワークデバイス等でもよい)にこの出願の実施形態に記載される方法を実行するように命令するためのいくつかの命令を含む。

【0418】

上記の実施形態の全部又は一部は、ソフトウェア、ハードウェア、ファームウェア又はこれらのいずれかの組み合わせを使用することにより実現されてもよい。ソフトウェアが実施形態を実現するために使用されるとき、実施形態の全部又は一部は、コンピュータプログラム製品の形式で実現されてもよい。

【0419】

コンピュータプログラム製品は、1つ以上のコンピュータ命令を含む。コンピュータプログラム命令がコンピュータにロードされて実行されたとき、この出願の実施形態による手順又は機能が全部又は一部生成される。コンピュータは、汎用コンピュータ、専用コンピュータ、コンピュータネットワーク又は他のプログラム可能装置でもよい。コンピュータ命令は、コンピュータ読み取り可能記憶媒体に記憶されてもよく、或いは、コンピュータ読み取り可能記憶媒体から他のコンピュータ読み取り可能記憶媒体に伝送されてもよい。例えば、コンピュータ命令は、有線(例えば、同軸ケーブル、光ファイバ又はデジタル加入者線(DSL))又は無線(例えば、赤外線、無線又はマイクロ波)方式で、ウェブサイト、コンピュータ、サーバ又はデータセンタから他のウェブサイト、コンピュータ、サーバ又は

10

20

30

40

50

【図 3 a】

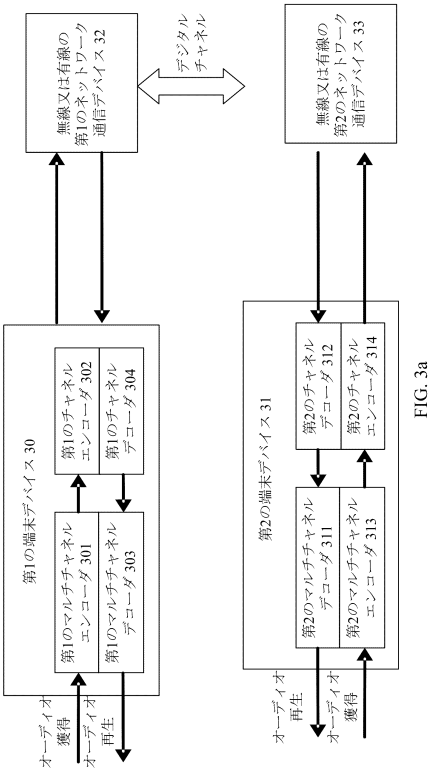


FIG. 3a

【図 3 b】

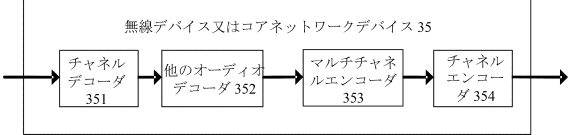


FIG. 3b

【図 3 c】

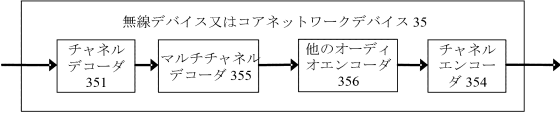


FIG. 3c

【図 4】

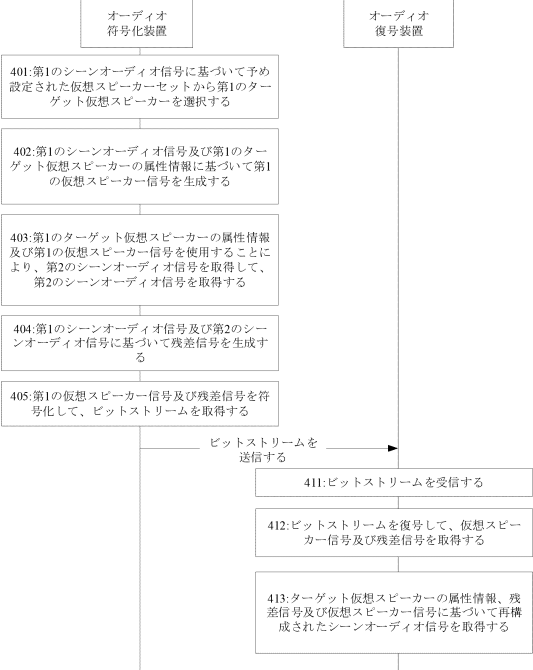
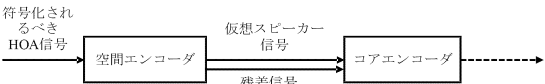
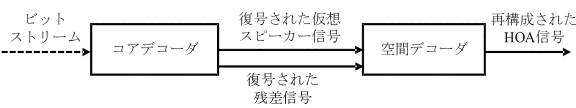


FIG. 4

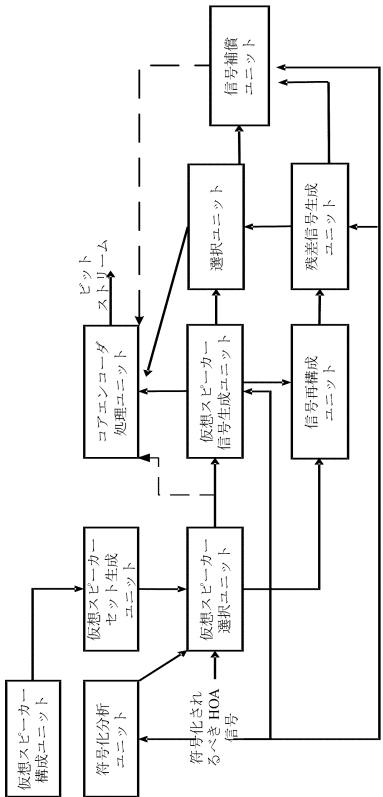
【図 5】



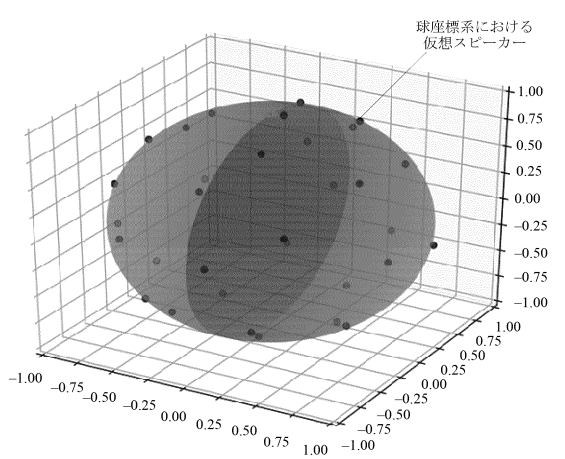
【図 6】



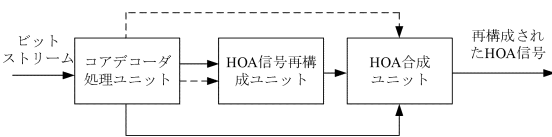
【図 7】



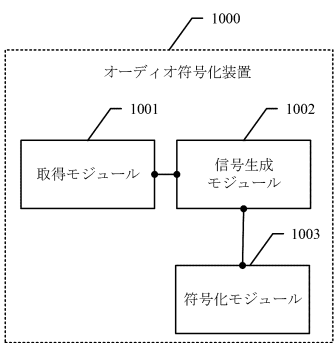
【図 8】



【図 9】



【図 10】



10

20

30

40

50

【図 1 1】

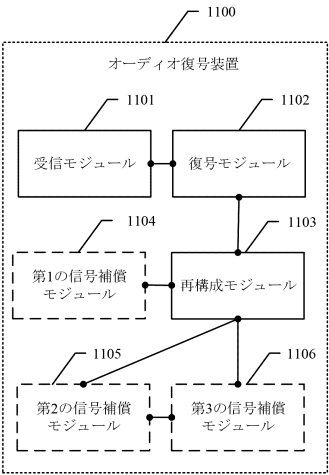


FIG. 11

【図 1 2】

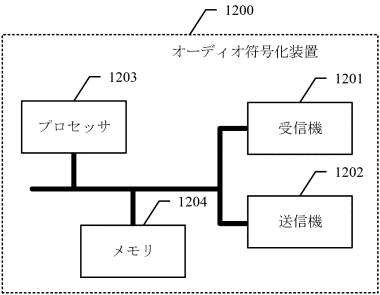


FIG. 12

【図 1 3】

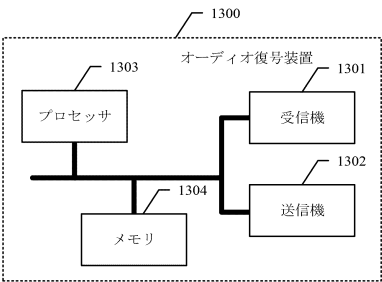


FIG. 13

10

20

30

40

50

フロントページの続き

- (74)代理人 100070150
弁理士 伊東 忠彦
- (74)代理人 100135079
弁理士 宮崎 修
- (72)発明者 ガオ, ユエン
中国 5 1 8 1 2 9 グァンドン シェンチェン ロンガン・ディストリクト バンティエン ホアウ
エイ・アドミニストレーション・ビルディング
- (72)発明者 リウ, シュワイ
中国 5 1 8 1 2 9 グァンドン シェンチェン ロンガン・ディストリクト バンティエン ホアウ
エイ・アドミニストレーション・ビルディング
- (72)発明者 ワーン, ビン
中国 5 1 8 1 2 9 グァンドン シェンチェン ロンガン・ディストリクト バンティエン ホアウ
エイ・アドミニストレーション・ビルディング
- (72)発明者 ワーン, ジョーア
中国 5 1 8 1 2 9 グァンドン シェンチェン ロンガン・ディストリクト バンティエン ホアウ
エイ・アドミニストレーション・ビルディング
- (72)発明者 チュイ, ティエンシュウ
中国 1 0 0 8 7 1 베이징 하이디엔・ディストリクト イーホーアユエン・ロード ナン
バー・5
- (72)発明者 シュイ, ジアハオ
中国 1 0 0 8 7 1 베이징 하이디엔・ディストリクト イーホーアユエン・ロード ナン
バー・5
- 審査官 土井 悠生
- (56)参考文献 特表 2 0 1 5 - 5 3 2 0 6 2 (J P , A)
米国特許出願公開第 2 0 1 9 / 0 2 5 3 8 2 6 (U S , A 1)
特表 2 0 1 6 - 5 2 0 8 6 4 (J P , A)
特表 2 0 2 0 - 5 2 7 7 4 5 (J P , A)
特表 2 0 1 6 - 5 2 3 3 7 7 (J P , A)
特表 2 0 1 5 - 5 3 7 2 5 6 (J P , A)
- (58)調査した分野 (Int.Cl., D B 名)
G 1 0 L 1 3 / 0 0 - 1 3 / 1 0
G 1 0 L 1 9 / 0 0 - 9 9 / 0 0