

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6239109号  
(P6239109)

(45) 発行日 平成29年11月29日(2017.11.29)

(24) 登録日 平成29年11月10日(2017.11.10)

(51) Int.Cl.

F 1

G 1 O L 19/008 (2013.01)

G 1 O L 19/008 19/008 100

G 1 O L 19/00 (2013.01)

G 1 O L 19/008 200

G 1 O L 19/00 330B

請求項の数 15 (全 37 頁)

(21) 出願番号 特願2016-528434 (P2016-528434)  
 (86) (22) 出願日 平成26年7月16日 (2014.7.16)  
 (65) 公表番号 特表2016-525714 (P2016-525714A)  
 (43) 公表日 平成28年8月25日 (2016.8.25)  
 (86) 國際出願番号 PCT/EP2014/065283  
 (87) 國際公開番号 WO2015/010996  
 (87) 國際公開日 平成27年1月29日 (2015.1.29)  
 審査請求日 平成28年2月23日 (2016.2.23)  
 (31) 優先権主張番号 EP13177365  
 (32) 優先日 平成25年7月22日 (2013.7.22)  
 (33) 優先権主張国 欧州特許庁 (EP)  
 (31) 優先権主張番号 EP13177367  
 (32) 優先日 平成25年7月22日 (2013.7.22)  
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 500341779  
 フラウンホーファーゲゼルシャフト・ツ  
 ル・フェルデルング・デル・アンゲヴァ  
 ンテン・フォルシュング・アインゲトラー  
 ゲネル・フェライン  
 ドイツ連邦共和国, 80686 ミュンヘ  
 ン, ハンザシュトラーセ 27ツエー  
 (74) 代理人 100085497  
 弁理士 筒井 秀隆  
 (72) 発明者 ボルス, クリストチャン  
 ドイツ連邦共和国 91058 エルラン  
 ゲン エリーゼースペースストラーセ  
 8

最終頁に続く

(54) 【発明の名称】低遅延オブジェクト・メタデータ符号化の装置と方法

## (57) 【特許請求の範囲】

## 【請求項 1】

1つ以上のオーディオチャネルを生成する装置(100)であって、  
 制御信号(b)に依存して、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )から1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するメタデータ復号器(110；901)であって、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々は1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示し、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々について複数の再生メタデータサンプル( $x_1'(n), \dots, x_N'(n)$ )を決定することにより、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するよう構成されている、メタデータ復号器(110；901)と、  
 10

前記1つ以上のオーディオオブジェクト信号と前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )とに依存して、前記1つ以上のオーディオチャネルを生成するオーディオチャネル生成部(120)と、を含み、

前記メタデータ復号器(110；901)は、前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各々の複数の処理済みメタデータサンプル( $z_1(n), \dots, z_N(n)$ )を受信するよう構成され、

前記メタデータ復号器(110；901)は、前記制御信号(b)を受信するよう構成され、

前記メタデータ復号器(110；901)は、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各再生メタデータ信号( $x_i'$ )の複数の再生メタデータサンプル( $x_i'(1), \dots, x_i'($  20

$n-1$ ),  $x_i'(n)$ )の各再生メタデータサンプル( $x_i'(n)$ )を、前記制御信号(b)が第1状態( $b(n)=0$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が前記1つ以上の処理済みメタデータ信号の1つ( $z_i$ )における処理済みメタデータサンプルの1つ( $z_i(n)$ )と前記再生メタデータ信号( $x_i'$ )の他の1つの既に生成された再生メタデータサンプル( $x_i'(n-1)$ )との合計であり、前記制御信号が前記第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の前記1つ( $z_i$ )における前記処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の前記1つ( $z_i(n)$ )であるように、決定するよう構成されている、装置。

【請求項2】

請求項1に記載の装置(100)であって、

10

前記メタデータ復号器(110; 901)は、2つ以上の前記処理済みメタデータ信号( $z_1, \dots, z_N$ )を受信して、2つ以上の前記再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するよう構成され

前記メタデータ復号器(110; 901)は、2つ以上のメタデータ復号器サブユニット(911, ..., 91N)を含み、

前記2つ以上のメタデータ復号器サブユニット(911, ..., 91N)の各々(91i, ..., 91i')は、加算部(910)及び選択部(930)を含むよう構成され、

前記2つ以上のメタデータ復号器サブユニット(911, ..., 91N)の各々(91i, ..., 91i')は、前記2つ以上の前記処理済みメタデータ信号( $z_1, \dots, z_N$ )の1つ( $z_i$ )における前記複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )を受信し、前記2つ以上の前記再生メタデータ信号( $x_1', \dots, x_N'$ )の1つ( $x_i'$ )を生成するよう構成され、

20

前記メタデータ復号器サブユニット(911, ..., 91N)の加算部(910)は、前記2つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の前記1つ( $z_i$ )における前記処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の1つ( $z_i(n)$ )と、前記2つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の1つ( $x_i'$ )における他の1つの既に生成された再生メタデータサンプル( $x_i'(n-1)$ )とを加算して、合計値( $s_i(n)$ )を得るよう構成され、

前記メタデータ復号器サブユニット(911, ..., 91N)の選択部(930)は、前記処理済みメタデータサンプルの前記1つ( $z_i(n)$ )と前記合計値( $s_i(n)$ )と前記制御信号とを受信するよう構成され、前記選択部(930)は、前記制御信号(b)が前記第1状態( $b(n)=0$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が前記合計値( $s_i(n)$ )であり、前記制御信号が前記第2状態( $b(n)=1$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が前記処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の前記1つ( $z_i(n)$ )であるように、前記再生メタデータ信号( $x_i'$ )の前記複数の再生メタデータサンプル( $x_i'(1), \dots, x_i'(n-1), x_i'(n)$ )の1つを決定するよう構成されている、装置。

30

【請求項3】

請求項1又は2に記載の装置(100)であって、

前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の少なくとも1つは、前記1つ以上のオーディオオブジェクト信号の1つにおける位置情報を示し、

前記オーディオチャネル生成部(120)は、前記1つ以上のオーディオオブジェクト信号の前記1つと前記位置情報とに依存して、前記1つ以上のオーディオチャネルの少なくとも1つを生成するよう構成されている、装置。

40

【請求項4】

請求項1乃至3のいずれか1項に記載の装置(100)であって、

前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の少なくとも1つは、前記1つ以上のオーディオオブジェクト信号の1つにおける音量を示し、

前記オーディオチャネル生成部(120)は、前記1つ以上のオーディオオブジェクト信号の前記1つと前記音量とに依存して、前記1つ以上のオーディオチャネルの少なくとも1つを生成するよう構成されている、装置。

【請求項5】

50

符号化済みオーディオデータを復号化する装置であって、

符号化済みオーディオデータを受信する入力インターフェイス(1100)であって、前記符号化済みオーディオデータが、複数の符号化済みチャネル、複数の符号化済みオブジェクト、及び前記複数の符号化済みオブジェクトに関する圧縮済みメタデータを含む、入力インターフェイスと、

請求項1乃至4のいずれか1項に記載の装置(100)と、を備え、

請求項1乃至4のいずれか1項に記載の前記装置(100)の前記メタデータ復号器(110；901)は、前記圧縮済みメタデータを解凍するメタデータ解凍部(1400)であり、

請求項1乃至4のいずれか1項に記載の前記装置(100)の前記オーディオチャネル生成部(120)は、前記複数の符号化済みチャネルと前記複数の符号化済みオブジェクトとを復号化するコア復号器(1300)を備え、

前記オーディオチャネル生成部(120)は、複数の復号化済みオブジェクトを解凍済みメタデータを使用して処理し、前記復号化済みオブジェクト及び復号化済みチャネルからオーディオデータを含む幾つかの出力チャネル(1205)を取得する、オブジェクト処理部(1200)を更に備え、

前記オーディオチャネル生成部(120)は、前記幾つかの出力チャネル(1205)を出力フォーマットへ変換する後処理部(1700)を更に備える、装置。

#### 【請求項6】

1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成する装置(250)であって、

1つ以上のオリジナル・メタデータ信号を受信し、前記1つ以上の処理済みメタデータ信号を決定するメタデータ符号器(210；801；802)であって、前記1つ以上のオリジナル・メタデータ信号の各々は複数のオリジナル・メタデータサンプルを含み、前記1つ以上のオリジナル・メタデータ信号の各々の前記オリジナル・メタデータサンプルは1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関する情報を示す、メタデータ符号器(210；801；802)と、

前記1つ以上のオーディオオブジェクト信号を符号化して前記1つ以上の符号化済みオーディオ信号を取得するオーディオ符号器(220)と、を備え、

前記メタデータ符号器(210；801；802)は、前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の各処理済みメタデータサンプル( $z_i(n)$ )を、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )における複数のオリジナル・メタデータサンプルの1つ( $x_i(n)$ )と前記処理済みメタデータ信号( $z_i$ )の他の1つの既に生成された処理済みメタデータサンプルとの差又は量子化された差を示し、前記制御信号(b)が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記1つ以上のオリジナル・メタデータ信号の前記1つ( $x_i$ )におけるオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記1つ( $x_i(n)$ )か又は前記オリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記1つ( $x_i(n)$ )における量子化済み表現( $q_i(n)$ )であるように、決定するよう構成されている、装置。

#### 【請求項7】

請求項6に記載の装置(250)であって、

前記メタデータ符号器(210；801；802)は、前記オリジナル・メタデータ信号( $x_1, \dots, x_N$ )の2つ以上を受信し、前記処理済みメタデータ信号( $z_1, \dots, z_N$ )の2つ以上を生成するよう構成され、

前記メタデータ符号器(210；801；802)は2つ以上のDPCM符号器(811, ..., 81N)を含み、

前記2つ以上のDPCM符号器(811, ..., 81N)の各々は、前記2つ以上のオリジナル・メタデータ信号( $x_1, \dots, x_N$ )の1つ( $x_i$ )におけるオリジナル・メタデータサンプル

10

20

30

40

50

$(x_i(1), \dots, x_i(n))$  の 1 つ( $x_i(n)$ )と、前記 2 つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の 1 つ( $z_i$ )における他の 1 つの既に生成された処理済みメタデータサンプルとの差又は量子化された差を決定して、差分サンプル( $y_i(n)$ )を取得するよう構成され、

前記メタデータ符号器(210；801；802)は、前記処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の 1 つを、前記制御信号( $b$ )が前記第 1 状態( $b(n)=0$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記差分サンプル( $y_i(n)$ )であり、前記制御信号( $b$ )が前記第 2 状態( $b(n)=1$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記オリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記 1 つ( $x_i(n)$ )か、又は前記オリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記 1 つ( $x_i(n)$ )における量子化済み表現( $q_i(n)$ )であるように決定するよう構成された選択部(830)をさらに含む、装置。  
10

#### 【請求項 8】

請求項 6 又は 7 に記載の装置(250)であって、

前記 1 つ以上のオリジナル・メタデータ信号の少なくとも 1 つは、前記 1 つ以上のオーディオオブジェクト信号の 1 つにおける位置情報を示し、

前記メタデータ符号器(210；801；802)は、前記位置情報を示す前記 1 つ以上のオリジナル・メタデータ信号の前記少なくとも 1 つに依存して、前記 1 つ以上の処理済みメタデータ信号の少なくとも 1 つを生成するよう構成されている、装置。

#### 【請求項 9】

請求項 6 乃至 8 のいずれか 1 項に記載の装置(250)であって、  
20

前記 1 つ以上のオリジナル・メタデータ信号の少なくとも 1 つは、前記 1 つ以上のオーディオオブジェクト信号の 1 つにおける音量を示し、

前記メタデータ符号器(210；801；802)は、前記音量を示す前記 1 つ以上のオリジナル・メタデータ信号の前記少なくとも 1 つに依存して、前記 1 つ以上の処理済みメタデータ信号の少なくとも 1 つを生成するよう構成されている、装置。

#### 【請求項 10】

請求項 6 乃至 9 のいずれか 1 項に記載の装置(250)であって、

前記メタデータ符号器(210；801；802)は、前記 1 つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の 1 つ( $z_i$ )における前記処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の各々を、前記制御信号が前記第 1 状態( $b(n)=0$ )を示すときには第 1 のビット数で符号化するよう構成され、前記制御信号が前記第 2 状態( $b(n)=1$ )を示すときには第 2 のビット数で符号化するよう構成され、前記第 1 のビット数は前記第 2 のビット数より小さい、装置。  
30

#### 【請求項 11】

オーディオ入力データ(101)を符号化してオーディオ出力データ(501)を取得する装置であって、

複数のオーディオチャネルと複数のオーディオオブジェクトと前記複数のオーディオオブジェクトの 1 つ又は複数に関連するメタデータとを受信する、入力インターフェイス(1100)と、

前記複数のオーディオオブジェクトと前記複数のオーディオチャネルとをミキシングして、複数のプレミクス済みチャネルを取得するミキサー(200)であって、各プレミクス済みチャネルは 1 つのオーディオチャネルのオーディオデータと少なくとも 1 つのオーディオオブジェクトのオーディオデータとを含む、ミキサー(200)と、  
40

請求項 6 乃至 9 のいずれか 1 項に記載の装置(250)と、を備え、

請求項 6 乃至 9 のいずれか 1 項に記載の装置(250)の前記オーディオ符号器(220)は、コア符号器入力データをコア符号化するコア符号器(300)であり、

請求項 6 乃至 9 のいずれか 1 項に記載の装置(250)の前記メタデータ符号器(210；801；802)は、前記複数のオーディオオブジェクトの 1 つ以上に関連する前記メタデータを圧縮するメタデータ圧縮部(400)である、装置。

#### 【請求項 12】

10

20

30

40

50

1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成するための請求項6乃至10のいずれか1項に記載の装置(250)と、

前記1つ以上の符号化済みオーディオ信号と前記1つ以上の処理済みメタデータ信号とを受信し、前記1つ以上の符号化済みオーディオ信号と前記1つ以上の処理済みメタデータ信号とに依存して1つ以上のオーディオチャネルを生成するための、請求項1乃至4のいずれか1項に記載の装置(100)と、

を備えるシステム。

【請求項13】

1つ以上のオーディオチャネルを生成する方法であって、

制御信号(b)に依存して、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )から1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップであって、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々は1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示し、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップは、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々について、複数の再生メタデータサンプル( $x_1'(n), \dots, x_N'(n)$ )を決定することにより実行される、ステップと、

前記1つ以上のオーディオオブジェクト信号と前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )とに依存して、前記1つ以上のオーディオチャネルを生成するステップと、を備え、

前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップは、前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各々の複数の処理済みメタデータサンプル( $z_1(n), \dots, z_N(n)$ )を受信し、かつ前記制御信号(b)を受信して、前記1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各再生メタデータ信号( $x_i'$ )の複数の再生メタデータサンプル( $x_i'(1), \dots, x_i'(n-1), x_i'(n)$ )の各再生メタデータサンプル( $x_i'(n)$ )を、前記制御信号(b)が第1状態( $b(n)=0$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が1つ以上の処理済みメタデータ信号の1つ( $z_i$ )における処理済みメタデータサンプルの1つ( $z_i(n)$ )と前記再生メタデータ信号( $x_i'$ )の他の1つの既に生成された再生メタデータサンプル( $x_i'(n-1)$ )との合計であり、前記制御信号が前記第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、前記再生メタデータサンプル( $x_i'(n)$ )が前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の前記1つ( $z_i$ )における前記処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の前記1つ( $z_i(n)$ )であるように、決定することにより実行される、方法。

【請求項14】

1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成する方法であって、

1つ以上のオリジナル・メタデータ信号を受信するステップと、

前記1つ以上の処理済みメタデータ信号を決定するステップと、

1つ以上のオーディオオブジェクト信号を符号化して前記1つ以上の符号化済みオーディオ信号を取得するステップと、を備え、

前記1つ以上のオリジナル・メタデータ信号の各々は複数のオリジナル・メタデータサンプルを含み、前記1つ以上のオリジナル・メタデータ信号の各々の前記オリジナル・メタデータサンプルは1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示しており、

前記1つ以上の処理済みメタデータ信号を決定するステップは、前記1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の各処理済みメタデータサンプル( $z_i(n)$ )を、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )における複数のオリジナル・メタデータサンプルの1つ( $x_i(n)$ )と前記処理済みメタデータ信号( $z_i$ )の他の1つの既に生成された処理済みメタデータサンプルとの差又は量子化された差を示し、前記制御信号(b)

10

20

30

40

50

が前記第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、前記処理済みメタデータサンプル( $z_i(n)$ )が前記1つ以上のオリジナル・メタデータ信号の前記1つ( $x_i$ )における前記オリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記1つ( $x_i(n)$ )又はオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の前記1つ( $x_i(n)$ )における量子化済み表現( $q_i(n)$ )であるように、決定することを含む、方法。

【請求項15】

コンピュータ又は信号プロセッサ上で作動されたとき、請求項13又は14に記載の方法を実行するためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明はオーディオ符号化/復号化に関し、特に空間オーディオ符号化及び空間オーディオオブジェクト符号化に関し、更に詳しくは効率的なオブジェクト・メタデータ符号化の装置と方法に関する。

【背景技術】

【0002】

空間オーディオ符号化ツールは当該技術において公知であり、例えばMPEGサラウンド標準で標準化されている。空間オーディオ符号化は、5個又は7個のチャネルなどのオリジナル入力チャネルから開始し、それらチャネルは再生設定におけるそれらの配置によって識別される。即ち、左チャネル、中央チャネル、右チャネル、左サラウンドチャネル、右サラウンドチャネル、及び低周波数強化チャネルである。空間オーディオ符号器は、典型的にはオリジナルチャネルから1つ以上のダウンミクスチャネルを導出し、加えて空間的キューに関連するパラメトリックデータを導出してあり、その空間的キューにはチャネルコヒーレンス値におけるチャネル間レベル差(interchannel level differences)、チャネル間位相差(interchannel phase differences)、チャネル間時間差(interchannel time differences)などがある。1つ以上のダウンミクスチャネルは、空間的キューを示すパラメトリックサイド情報と一緒に空間オーディオ復号器へと伝送され、その復号器は、ダウンミクスチャネルとその関連するパラメトリックデータとを復号化して、オリジナル入力チャネルの近似されたバージョンである出力チャネルを最終的に取得する。出力設定におけるチャネルの配置は典型的には固定されており、例えば5.1フォーマット、7.1フォーマットなどである。

20

【0003】

そのようなチャネルベースのオーディオフォーマットは、多チャネルオーディオコンテンツを記憶又は伝送するために広く使用されており、その場合、各チャネルは所与の位置にある特異なラウドスピーカに関連している。このような種類のフォーマットを忠実に再生するためには、オーディオ信号の生成時に使用されたスピーカー位置と同じ位置にスピーカが配置されているような、ラウドスピーカ設定が要求される。ラウドスピーカの個数を増加させると、真に音に浸りこむような改善された3Dオーディオシーンの再生が可能になる一方で、特にリビングルームのような家庭的な環境では、そのような要求を満たすことはますます困難になる。

30

【0004】

特異なラウドスピーカ設定を有することの必要性は、ラウドスピーカ信号が再生設定のために特異的にレンダリングされるオブジェクト・ベースの手法によって克服され得る。

【0005】

例えば空間オーディオオブジェクト符号化ツールは、当該技術において公知であり、MPEG SAOC標準(SAOC=空間オーディオオブジェクト符号化)において標準化されている。オリジナルチャネルから開始する空間オーディオ符号化とは対照的に、空間オーディオオブジェクト符号化はオーディオオブジェクトから開始し、それらオブジェクトはあるレンダリング再生設定に対して自動的に専用となる訳ではない。代わりに、再生シーン内におけるオーディオオブジェクトの配置には柔軟性があり、あるレンダリング情

40

50

報を空間オーディオオブジェクト符号化・復号器へと入力することによりユーザーが決定することもできる。代替的又は追加的に、レンダリング情報、即ち再生設定におけるどの位置に、あるオーディオオブジェクトが典型的には時間にわたって配置されるべきかという情報は、追加的サイド情報又はメタデータとして伝送され得る。あるデータ圧縮を得るために幾つかのオーディオオブジェクトがS A O C 符号器によって符号化され、その符号器は、あるダウンミクス情報に従ってオブジェクトをダウンミクスすることで入力オブジェクトから1つ以上の転送チャネルを計算する。更に、S A O C 符号器は、オブジェクトレベル差(O L D)、オブジェクトコヒーレンス値などのオブジェクト間キューを表現しているパラメトリックサイド情報を計算する。S A C (S A C = 空間オーディオ符号化)においては、オブジェクト間のパラメトリックデータが個別の時間／周波数タイルについて計算される。即ち、例えば1024個又は2048個のサンプルを有するオーディオ信号のあるフレームについて、最終的に各フレーム及び各周波数帯域に対してパラメトリックデータが存在するように、24個、32個又は64個などの周波数帯域が考慮される。一例として、あるオーディオピースが20フレームを有し、各フレームが32個の周波数帯域へと分割される場合、時間／周波数タイルの数は640個となる。

#### 【0006】

オブジェクト・ベースの手法において、音場は離散的なオーディオオブジェクトによって記述される。そのため、とりわけ3D空間における各音源の時間変化する位置を記述するオブジェクト・メタデータが必要になる。

#### 【0007】

先行技術における第1のメタデータ符号化概念は、空間サウンド記述インターチェンジフォーマット(S p a t D I F)であり、これは未だ開発中のオーディオシーン記述フォーマットである(非特許文献1)。そのフォーマットは、オブジェクト・ベースのサウンドシーンのためのインターチェンジフォーマットとして設計され、オブジェクト軌跡についての圧縮方法を何も提供してはいない。S p a t D I Fは、オブジェクト・メタデータを構築するために、テキスト・ベースのオープンサウンド制御(O S C)フォーマットを使用する(非特許文献2)。しかしながら、単純なテキスト・ベースの表現はオブジェクト軌跡の圧縮された伝送のための選択肢にはならない。

#### 【0008】

先行技術における他のメタデータ概念はオーディオシーン記述フォーマット(A S D F)(非特許文献3)であり、同様の欠点を持つテキスト・ベースの解決策である。そのデータは、拡張可能なマーク付け言語(Extensible Markup Language: X M L)(非特許文献4、非特許文献5)の部分集合である、同期されたマルチメディア統合言語(Synchronized Multimedia Integration Language: SMIL)の拡張によって構築される。

#### 【0009】

先行技術における更なるメタデータ概念は、シーンのためのオーディオバイナリフォーマット(A u d i o B I F S)であり、M P E G - 4仕様(非特許文献6、非特許文献7)の一部であるバイナリフォーマットである。そのフォーマットは、視聴覚の3Dシーンや双方向仮想現実アプリケーション(非特許文献8)のために開発されたX M Lベースの仮想現実モデリング言語(Virtual Reality Modeling Language: V R M L)に深く関連している。複雑なA u d i o B I F S仕様は、オブジェクトの動きの経路を特定するためにシーディングラフを使用する。A u d i o B I F Sの主たる欠点は、制限されたシステム遅延及びデータストリームへのランダムアクセスが要件となるようなリアルタイム操作のために設計されていないという点である。更に、オブジェクト位置の符号化は人間のリスナーの制限された定位性能を活用していない。視聴覚シーン内の固定的なリスナー位置に対しては、オブジェクトデータは遙かに少数のビットで量子化され得る(非特許文献9)。よって、A u d i o B I F Sの中で適用されるオブジェクト・メタデータの符号化は、データ圧縮に関して効率的でない。

#### 【0010】

そこで、改善された効率的なオブジェクト・メタデータ符号化の概念が提供されれば、

10

20

30

40

50

高く評価されるであろう。

【先行技術文献】

【特許文献】

【0011】

【特許文献1】[10] Cutler, C. C. (1950), "Differential Quantization of Communication Signals", US Patent US2605361, Jul. 1952.

【非特許文献】

【0012】

【非特許文献1】[1] Peters, N., Lossius, T. and Schacher J. C., "SpatDIF: Principles, Specification, and Examples", 9th Sound and Music Computing Conference, Copenhagen, Denmark, Jul. 2012.

10

【非特許文献2】[2] Wright, M., Freed, A., "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers", International Computer Music Conference, Thessaloniki, Greece, 1997.

【非特許文献3】[3] Matthias Geier, Jens Ahrens, and Sascha Spors. (2010), "Object-based audio reproduction and the audio scene description format", Org. Sound, Vol. 15, No. 3, pp. 219-227, December 2010.

【非特許文献4】[4] W3C, "Synchronized Multimedia Integration Language (SMIL 3.0)", Dec. 2008.

【非特許文献5】[5] W3C, "Extensible Markup Language (XML) 1.0 (Fifth Edition)", Nov. 2008.

20

【非特許文献6】[6] MPEG, "ISO/IEC International Standard 14496-3- Coding of a audio-visual objects, Part 3 Audio", 2009.

【非特許文献7】[7] Schmidt, J.; Schroeder, E. F. (2004), "New and Advanced Features for Audio Presentation in the MPEG-4 Standard", 116th AES Convention, Berlin, Germany, May 2004

【非特許文献8】[8] Web3D, "International Standard ISO/IEC 14772-1:1997-The Virtual Reality Modeling Language (VRML), Part 1: Functional specification and UTF-8 encoding", 1997.

【非特許文献9】[9] Sporer, T. (2012), "Codierung raeumlicher Audiosignale mit leichtgewichtigen Audio-Objekten", Proc. Annual Meeting of the German Audiological Society (DGA), Erlangen, Germany, Mar. 2012.

30

【非特許文献10】[11] Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning" ; J. Audio Eng. Soc., Volume 45, Issue 6, pp. 456-466, June 1997.

【発明の概要】

【発明が解決しようとする課題】

【0013】

本発明の目的は、オブジェクト・メタデータ符号化のための改善された概念を提供することである。本発明の目的は、請求項1に記載の装置と、請求項6に記載の装置と、請求項12に記載のシステムと、請求項13に記載の方法と、請求項14に記載の方法と、請求項15に記載のコンピュータプログラムとによって達成される。

40

【課題を解決するための手段】

【0014】

1つ以上のオーディオチャネルを生成する装置が提供される。この装置は、制御信号(b)に依存して、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )から1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するメタデータ復号器を含み、それら1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々は1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関する情報を示し、メタデータ復号器は、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々に対して複数の再生メタデータサンプル( $x_1'(n), \dots, x_N'(n)$ )

50

を決定することにより、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するよう構成されている。更に、この装置は、1つ以上のオーディオオブジェクト信号に依存しつつ1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )に依存して、1つ以上のオーディオチャネルを生成するオーディオチャネル生成部を含む。メタデータ復号器は、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各々の複数の処理済みメタデータサンプル( $z_1(n), \dots, z_N(n)$ )を受信するよう構成されている。更に、メタデータ復号器は制御信号(b)を受信するよう構成されている。更に、メタデータ復号器は、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各再生メタデータ信号( $x_i'$ )の複数の再生メタデータサンプル( $x_i'(1), \dots, x_i'(n-1), x_i'(n)$ )の各再生メタデータサンプル( $x_i'(n)$ )を以下のように決定する。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、再生メタデータサンプル( $x_i'(n)$ )が、1つ以上の処理済みメタデータ信号のうちの1つ( $z_i$ )の処理済みメタデータサンプルの1つ( $z_i(n)$ )と再生メタデータ信号( $x_i'$ )の他の1つの既に生成された再生メタデータサンプル( $x_i'(n-1)$ )との合計であり、制御信号が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、再生メタデータサンプル( $x_i'(n)$ )が、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の1つ( $z_i$ )の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の1つ( $z_i(n)$ )であるように、決定する。  
10

#### 【0015】

更に、1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成する装置が提供される。この装置は、1つ以上のオリジナル・メタデータ信号を受信して1つ以上の処理済みメタデータ信号を決定する、メタデータ符号器を含む。1つ以上のオリジナル・メタデータ信号の各々は複数のオリジナル・メタデータサンプルを含み、1つ以上のオリジナル・メタデータ信号の各々のオリジナル・メタデータサンプルは1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示す。  
20

#### 【0016】

更にこの装置は、1つ以上のオーディオオブジェクト信号を符号化して1つ以上の符号化済みオーディオ信号を取得するオーディオ符号器を含む。

#### 【0017】

メタデータ符号器は、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の各処理済みメタデータサンプル( $z_i(n)$ )を、以下のように決定する。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )の複数のオリジナル・メタデータサンプルの1つ( $x_i(n)$ )と、処理済みメタデータ信号( $z_i$ )の他の1つの既に生成された処理済みメタデータサンプルとの差又は量子化された差を示し、制御信号(b)が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )のオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )、又はオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )の量子化済み表現( $q_i(n)$ )となるように、決定する。  
30

#### 【0018】

一実施形態によれば、制限されたデータレートでチャネルを伝送するための効率的な圧縮メカニズムを達成する、オブジェクト・メタデータのためのデータ圧縮概念が提供される。符号器及び復号器によってそれぞれ導入される追加的な遅延は全くない。更に、純粋な方位角変化、例えばカメラ回転に対する良好な圧縮レートが達成される。更にこの提案の概念は、例えば位置的ジャンプのような不連続的な軌跡をサポートする。更に、複雑性の低い復号化を実現できる。更に、制限された再初期化時間でランダムアクセスを達成できる。

#### 【0019】

1つ以上のオーディオチャネルを生成する方法が提供される。その方法は、

- 制御信号(b)に依存して、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )から1つ以上

50

の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップであって、それら1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々は1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示し、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップは、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々に対し複数の再生メタデータサンプル( $x_1'(n), \dots, x_N'(n)$ )を決定することにより実行される、ステップと、

- 1つ以上のオーディオオブジェクト信号に依存しあつ1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )に依存して、1つ以上のオーディオチャネルを生成するステップと、  
を備える。

#### 【0020】

1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するステップは、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各々の複数の処理済みメタデータサンプル( $z_1(n), \dots, z_N(n)$ )を受信すること、制御信号(b)を受信すること、及び1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各再生メタデータ信号( $x_i'$ )の複数の再生メタデータサンプル( $x_i'(1), \dots, x_i'(n-1), x_i'(n)$ )の各再生メタデータサンプル( $x_i'(n)$ )を以下のように決定すること、により実行される。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、再生メタデータサンプル( $x_i'(n)$ )が、1つ以上の処理済みメタデータ信号のうちの1つ( $z_i$ )の処理済みメタデータサンプルの1つ( $z_i(n)$ )と前記再生メタデータ信号( $x_i'$ )の他の1つ既に生成された再生メタデータサンプル( $x_i'(n-1)$ )との合計となり、制御信号が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、再生メタデータサンプル( $x_i'(n)$ )が1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の前記1つ( $z_i$ )の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の前記1つ( $z_i(n)$ )となるように、決定する。

#### 【0021】

更に、1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成する方法が提供される。その方法は、

- 1つ以上のオリジナル・メタデータ信号を受信するステップと、  
- 1つ以上の処理済みメタデータ信号を決定するステップと、  
- 1つ以上のオーディオオブジェクト信号を符号化して1つ以上の符号化済みオーディオ信号を取得するステップと、  
を備える。

#### 【0022】

1つ以上のオリジナル・メタデータ信号の各々は複数のオリジナル・メタデータサンプルを含み、1つ以上のオリジナル・メタデータ信号の各々のオリジナル・メタデータサンプルは1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示している。1つ以上の処理済みメタデータ信号を決定するステップは、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の各処理済みメタデータサンプル( $z_i(n)$ )を、以下のように決定することを含む。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )の複数のオリジナル・メタデータサンプルの1つ( $x_i(n)$ )と処理済みメタデータ信号( $z_i$ )の他の1つ既に生成された処理済みメタデータサンプルとの差又は量子化された差を示し、制御信号(b)が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、1つ以上のオリジナル・メタデータ信号の前記1つ( $x_i$ )のオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )、又はオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )の量子化済み表現( $q_i(n)$ )となるように、決定する。

#### 【0023】

更に、コンピュータ又は信号プロセッサ上で作動されたときに、上述した方法を実行するコンピュータプログラムが提供される。

#### 【0024】

10

20

30

40

50

以下に、本発明の実施形態を図面を参照しながらより詳細に説明する。

【図面の簡単な説明】

【0025】

【図1】1つ以上のオーディオチャネルを生成する、一実施形態に係る装置を示す。

【図2】符号化済みオーディオ情報を生成する、一実施形態に係る装置を示す。

【図3】一実施形態に係るシステムを示す。

【図4】方位角、仰角及び半径により表現された、原点からの3次元空間におけるオーディオオブジェクトの位置を示す。

【図5】オーディオチャネル生成部により想定されたオーディオオブジェクトの位置とラウドスピーカ設定とを示す。 10

【図6】差分パルス符号変調符号器を示す。

【図7】差分パルス符号変調復号器を示す。

【図8a】一実施形態に係るメタデータ符号器を示す。

【図8b】他の実施形態に係るメタデータ符号器を示す。

【図9a】一実施形態に係るメタデータ復号器を示す。

【図9b】一実施形態に係るメタデータ復号器サブユニットを示す。

【図10】3Dオーディオ符号器の第1実施形態を示す。

【図11】3Dオーディオ復号器の第1実施形態を示す。

【図12】3Dオーディオ符号器の第2実施形態を示す。

【図13】3Dオーディオ復号器の第2実施形態を示す。 20

【図14】3Dオーディオ符号器の第3実施形態を示す。

【図15】3Dオーディオ復号器の第3実施形態を示す。

【発明を実施するための形態】

【0026】

図2は、1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成するための一実施形態に係る装置250を示す。

【0027】

装置250は、1つ以上のオリジナル・メタデータ信号を受信し、1つ以上の処理済みメタデータ信号を決定する、メタデータ符号器210を含む。1つ以上のオリジナル・メタデータ信号の各々は複数のオリジナル・メタデータサンプルを含み、1つ以上のオリジナル・メタデータ信号の各々のオリジナル・メタデータサンプルは、1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示す。 30

【0028】

更に、装置250は、1つ以上のオーディオオブジェクト信号を符号化して1つ以上の符号化済みオーディオ信号を取得する、オーディオ符号器220を含む。

【0029】

メタデータ符号器210は、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各処理済みメタデータ信号( $z_i$ )の複数の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n-1), z_i(n)$ )の各処理済みメタデータサンプル( $z_i(n)$ )を、以下のように決定するよう構成されている。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、前記1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )の複数のオリジナル・メタデータサンプルの1つ( $x_i(n)$ )と処理済みメタデータ信号( $z_i$ )の他の1つの既に生成された処理済みメタデータサンプルとの差又は量子化された差を示し、制御信号(b)が第1状態とは異なる第2状態( $b(n)=1$ )を示すときには、処理済みメタデータサンプル( $z_i(n)$ )が、1つ以上のオリジナル・メタデータ信号の1つ( $x_i$ )のオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )、又はオリジナル・メタデータサンプル( $x_i(1), \dots, x_i(n)$ )の1つ( $x_i(n)$ )の量子化済み表現( $q_i(n)$ )となるように、決定する。 40

【0030】

図1は、1つ以上のオーディオチャネルを生成する一実施形態に係る装置100を示す。 50

## 【0031】

装置100は、制御信号(b)に依存して、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )から1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するメタデータ復号器110を含み、それら1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々は、1つ以上のオーディオオブジェクト信号の1つのオーディオオブジェクト信号に関連する情報を示し、メタデータ復号器110は、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各々に対し複数の再生メタデータサンプル( $x_1'(n), \dots, x_N'(n)$ )を決定することにより、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )を生成するよう構成されている。

## 【0032】

更に装置100は、1つ以上のオーディオオブジェクト信号に依存しかつ1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )に依存して、1つ以上のオーディオチャネルを生成するオーディオチャネル生成部120を含む。10

## 【0033】

メタデータ復号器110は、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の各々の複数の処理済みメタデータサンプル( $z_1(n), \dots, z_N(n)$ )を受信するよう構成されている。更にメタデータ復号器110は、制御信号(b)を受信するよう構成されている。

## 【0034】

更に、メタデータ復号器110は、1つ以上の再生メタデータ信号( $x_1', \dots, x_N'$ )の各再生メタデータ信号( $x_i'$ )の複数の再生メタデータサンプル( $x_i'(n), \dots, x_i'(n-1), x_i'(n)$ )の各再生メタデータサンプル( $x_i'(n)$ )を、以下のように決定するよう構成されている。即ち、制御信号(b)が第1状態( $b(n)=0$ )を示すとき、再生メタデータサンプル( $x_i'(n)$ )が、1つ以上の処理済みメタデータ信号の1つ( $z_i$ )の処理済みメタデータサンプルの1つ( $z_i(n)$ )と再生メタデータ信号( $x_i'$ )の他の1つの既に生成された再生メタデータサンプル( $x_i'(n-1)$ )との合計であり、制御信号が第1状態とは異なる第2状態( $b(n)=1$ )を示すとき、再生メタデータサンプル( $x_i'(n)$ )が、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の前記1つ( $z_i$ )の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の前記1つ( $z_i(n)$ )であるよう<sup>20</sup>に、決定する。

## 【0035】

メタデータサンプルについて言及する場合には、1つのメタデータサンプルは、そのメタデータサンプル値によって特徴付けられるだけでなく、そのメタデータサンプルが関連する時点によっても特徴付けられるという点にも留意すべきである。例えば、そのような時点とは、オーディオシーケンスの開始点又はそれと同様な点に対して相対的であってよい。例えば、インデックスn又はkはメタデータ信号内のメタデータサンプルの位置を識別していくてもよく、これにより、(開始時点に関連する)(相対的)時点が示されてもよい。注意すべきは、2つのメタデータサンプルが異なる時点に関連する場合、(時々起こり得ることであるが)たとえこれらのメタデータサンプル値が同一であったとしても、それら2つのメタデータサンプルは異なるメタデータサンプルであるということである。30

## 【0036】

上述の実施形態は、オーディオオブジェクト信号と関連する(メタデータ信号に含まれる)メタデータ情報がゆっくりと変化する場合が多い、という知見に基づいている。40

## 【0037】

例えば、メタデータ信号は、オーディオオブジェクトについての位置情報(例えばオーディオオブジェクトの位置を定義する方位角、仰角又は半径)を示してもよい。殆どの時点でオーディオオブジェクトの位置は変化しないか又はゆっくりとだけ変化する、と想定されてもよい。

## 【0038】

あるいは、メタデータ信号は、例えばオーディオオブジェクトの音量(例えばゲイン)を示してもよく、殆どの時点でオーディオオブジェクトの音量はゆっくりと変化すると想定されてもよい。

## 【0039】

10

20

30

40

50

このような理由により、全ての時点における（完全な）メタデータ情報を伝送する必要はない。

【0040】

その代わり、幾つかの実施形態によれば、（完全な）メタデータ情報が例えはある時点においてだけ伝送されてもよく、例えばN番目の時点毎に周期的に、例えば時点0, N, 2N, 3N等において伝送されてもよい。

【0041】

例えば、実施形態においては、3個のメタデータ信号が3D空間におけるオーディオオブジェクトの位置を特定する。メタデータ信号の1番目は、例えばオーディオオブジェクトの位置の方位角を特定してもよい。メタデータ信号の2番目は、例えばオーディオオブジェクトの位置の仰角を特定してもよい。メタデータ信号の3番目は、例えばオーディオオブジェクトの距離に関係する半径を特定してもよい。

【0042】

方位角と仰角と半径とは、3D空間におけるオーディオオブジェクトの原点からの位置を明確に定義する。これについては図4を参照しながら説明する。

【0043】

図4は、三次元（3D）空間におけるオーディオオブジェクトの原点400からの位置410を、方位角と仰角と半径とで示す。

【0044】

仰角は、例えば、原点からオブジェクト位置までの直線と、この直線のx y平面（x軸とy軸とによって定義される平面）への垂直投影線との角度を特定する。方位角は、例えばx軸と前記垂直投影線との角度を定義する。方位角と仰角とを特定することで、原点400とオーディオオブジェクトの位置410とを通る直線415が定義され得る。更に半径を特定することで、オーディオオブジェクトの正確な位置410が定義され得る。

【0045】

一実施形態において、方位角は-180° < 方位角 180° の範囲で定義され、仰角は-90°。仰角 90° の範囲で定義され、半径は例えばメートル[m]（0m以上である）で定義され得る。

【0046】

例えばx y z座標系におけるオーディオオブジェクト位置の全てのx値がゼロ以上であると想定され得るような他の実施形態においては、方位角は-90°。方位角 90° の範囲で定義され、仰角は-90°。仰角 90° の範囲で定義され、半径は例えばメートル[m]で定義され得る。

【0047】

更なる実施形態において、方位角が-128° < 方位角 128° の範囲で定義され、仰角が-32°。仰角 32° の範囲で定義され、半径が例えば対数スケールで定義され得るよう、メタデータ信号はスケールされてもよい。幾つかの実施形態において、オリジナル・メタデータ信号、処理済みメタデータ信号、及び再生メタデータ信号は、それぞれ、1つ以上のオーディオオブジェクト信号の1つの位置情報のスケールされた表現及び/又は音量のスケールされた表現を含んでもよい。

【0048】

オーディオチャネル生成部120は、例えば、1つ以上のオーディオオブジェクト信号に依存しつつ再生メタデータ信号に依存して、1つ以上のオーディオチャネルを生成するよう構成されてもよく、その再生メタデータ信号は、例えばオーディオオブジェクトの位置を示してもよい。

【0049】

図5は、オーディオオブジェクトの位置と、オーディオチャネル生成部により想定されるラウドスピーカ設定とを示す。x y z座標系の原点500が示されている。更に、第1オーディオオブジェクトの位置510と、第2オーディオオブジェクトの位置520とが示されている。更に、図5は、オーディオチャネル生成部120が4個のラウドスピーカ

10

20

30

40

50

のための4個のオーディオチャネルを生成するシナリオを示す。オーディオチャネル生成部120は、4個のラウドスピーカ511, 512, 513, 514が図5に示す位置に配置されていると想定している。

#### 【0050】

図5において、第1オーディオオブジェクトは、ラウドスピーカ511と512の想定位置に近い位置510に配置されており、ラウドスピーカ513と514からは遠い位置に配置されている。従って、オーディオチャネル生成部120は、第1オーディオオブジェクト510がラウドスピーカ511及び512により再生され、ラウドスピーカ513及び514では再生されないように、4個のオーディオチャネルを生成してもよい。

#### 【0051】

他の実施形態において、オーディオチャネル生成部120は、第1オーディオオブジェクト510がラウドスピーカ511及び512により高い音量で再生され、ラウドスピーカ513及び514により低い音量で再生されるように、4個のオーディオチャネルを生成してもよい。

#### 【0052】

更に、第2オーディオオブジェクトは、ラウドスピーカ513と514の想定位置に近い位置520に配置されており、ラウドスピーカ511と512からは遠い位置に配置されている。従って、オーディオチャネル生成部120は、第2オーディオオブジェクト520がラウドスピーカ513及び514により再生され、ラウドスピーカ511及び512では再生されないように、4個のオーディオチャネルを生成してもよい。

10

#### 【0053】

他の実施形態において、オーディオチャネル生成部120は、第2オーディオオブジェクト520がラウドスピーカ513及び514により高い音量で再生され、ラウドスピーカ511及び512により低い音量で再生されるように、4個のオーディオチャネルを生成してもよい。

#### 【0054】

代替的な実施形態において、オーディオオブジェクトの位置を特定するために、2個のメタデータ信号だけが使用されてもよい。例えば、全てのオーディオオブジェクトが単一平面に配置されていると想定される場合には、例えば方位角と半径だけが特定されてもよい。

20

#### 【0055】

更に他の実施形態においては、各オーディオオブジェクトのために、単一のメタデータ信号だけが位置情報として符号化されかつ伝送される。例えば、あるオーディオオブジェクトについて、方位角だけが位置情報として特定されてもよい（例えば全てのオーディオオブジェクトが同一平面上に配置され、中心点から同一距離を持ち、従って同一半径を有すると想定される場合など）。方位角情報は、例えば、オーディオオブジェクトが左のラウドスピーカに近く、右のラウドスピーカからは遠いと判定することで十分であってもよい。そのような状況において、オーディオチャネル生成部120は、例えばオーディオオブジェクトが左のラウドスピーカによって再生されるが、右のラウドスピーカでは再生されないように、1つ以上のオーディオチャネルを生成してもよい。

30

#### 【0056】

例えば、ラウドスピーカのオーディオチャネルの各々の中におけるオーディオオブジェクト信号の重みを決定するために、ベクトルベース振幅パニング（Vector Base Amplitude Panning (V B A P)）が使用されてもよい（例えば非特許文献10を参照）。例えば、V B A Pに関しては、オーディオオブジェクトが仮想音源に関連すると想定されている。

40

#### 【0057】

実施形態において、各オーディオオブジェクトについて、更なるメタデータ信号が音量、例えばゲイン（例えばデシベル [d B] で表現された）を特定してもよい。

#### 【0058】

50

例えば図5において、第1ゲイン値は、位置510に配置された第1オーディオオブジェクトのための更なるメタデータ信号により特定されてもよく、その値は、位置520に配置された第2オーディオオブジェクトのための別の更なるメタデータ信号によって特定される第2ゲイン値よりも高い。そのような状況において、ラウドスピーカ511及び512は第1オーディオオブジェクトを、ラウドスピーカ513及び514が第2オーディオオブジェクトを再生する音量よりも高い音量で再生してもよい。

【0059】

実施形態はまた、オーディオオブジェクトのそのようなゲイン値がゆっくりと変化する場合が多いと想定している。従って、そのようなメタデータ情報を全ての時点において伝送する必要はない。代わりに、メタデータ情報は、ある時点において伝送されるだけである。中間の時点においては、メタデータ情報は、例えば伝送された先行するメタデータサンプルと後続のメタデータサンプルとを使用して近似されてもよい。例えば、中間値の近似のために線形補間が使用されてもよい。例えば、オーディオオブジェクトの各々のゲイン、方位角、仰角及び/又は半径が、そのようなメタデータが伝送されなかつた時点のために近似されてもよい。

10

【0060】

そのような手法により、メタデータの伝送レートにおける相当な節約を達成し得る。

【0061】

図3は、一実施形態に従うシステムを示す。

【0062】

20

このシステムは、1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを含む符号化済みオーディオ情報を生成する、上述のような装置250を備える。

【0063】

更に、そのシステムは、1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とを受信し、その1つ以上の符号化済みオーディオ信号と1つ以上の処理済みメタデータ信号とに依存して、上述のように1つ以上のオーディオチャネルを生成する装置100を備える。

【0064】

30

例えば、1つ以上のオーディオオブジェクトを符号化するための符号化装置250がS A O C符号器を使用した場合には、1つ以上の符号化済みオーディオ信号が、現状技術に係るS A O C復号器を使用して1つ以上のオーディオチャネルを生成する装置100によって復号化されて、1つ以上のオーディオオブジェクト信号が取得されてもよい。

【0065】

実施形態は、差分パルス符号変調の概念が拡張可能であり、そのような拡張された概念はオーディオオブジェクトのためのメタデータ信号を符号化するのに適している、という知見に基づいている。

【0066】

差分パルス符号変調(DPCM)の方法は、量子化を介して不適切さ(irrelevance)を低減し、差分伝送を介して冗長性を低減するような、低速変化する時間信号のための確立された方法である(特許文献1)。あるDPCM符号器を図6で示す。

40

【0067】

図6のDPCM符号器において、入力信号xの実際の入力サンプルx(n)が減算ユニット610へと入力される。減算ユニットの他の入力では、別の値が減算ユニットへと入力される。この別の値は、以前に受信されたサンプルx(n-1)であると想定されてもよい。しかし、量子化誤差又は他の誤差により、他の入力における値が以前のサンプルx(n-1)と正確に同一ではないという結果をもたらしている可能性もある。そのようなx(n-1)からの可能性のあるずれに起因して、減算部の他の入力はx\*(n-1)と称されてもよい。減算ユニットは、x(n)からx\*(n-1)を減算して差分値d(n)を取得する。

【0068】

50

次に、 $d(n)$ は量子化部 620 内で量子化されて、出力信号  $y$  の別の出力サンプル  $y(n)$  が取得される。一般的に、 $y(n)$  は  $d(n)$  と等しいか又は  $d(n)$  に近い値である。

【0069】

更に、 $y(n)$  は加算部 630 へと入力される。更に、 $x^*(n-1)$  も加算部 630 へと入力される。 $d(n)$  が減算  $d(n)=x(n)-x^*(n-1)$  からもたらされ、 $y(n)$  が  $d(n)$  と等しい値又は少なくとも近い値であるため、加算部 630 の出力  $x^*(n)$  は、 $x(n)$  と等しいか又は少なくとも  $x(n)$  に近い。

【0070】

$x^*(n)$  はサンプリング期間中、ユニット 640 において保持され、次に、処理は次のサンプル  $x(n+1)$  で継続される。

10

【0071】

図 7 は、対応する DPCM 復号器を示す。

【0072】

図 7 において、DPCM 符号器からの出力信号  $y$  のサンプル  $y(n)$  は加算部 710 へと入力される。 $y(n)$  は、再生されるべき信号  $x(n)$  の差分値を表す。加算部 710 の他の入力においては、以前に再生されたサンプル  $x'(n-1)$  が入力される。加算部の出力  $x'(n)$  は加算  $x'(n)=x'(n-1)+y(n)$  からもたらされる。 $x'(n-1)$  は、一般的に  $x(n-1)$  と等しいか又は少なくとも近く、また、 $y(n)$  は、 $x(n)-x(n-1)$  と等しいか又は近いので、加算部 710 の出力  $x'(n)$  は、一般的に  $x(n)$  と等しいか又は近い。

【0073】

$x'(n)$  はサンプリング期間中、ユニット 740 において保持され、次に、処理は次のサンプル  $y(n+1)$  で継続される。

20

【0074】

DPCM 圧縮方法は前述した要求される特徴の殆どを満足するが、その圧縮方法はランダムアクセスを許可しない。

【0075】

図 8 a は、一実施形態に従うメタデータ符号器 801 を示す。

【0076】

図 8 a のメタデータ符号器 801 により使用される符号化方法は、従来の DPCM 符号化方法の拡張である。

30

【0077】

図 8 a のメタデータ符号器 801 は、1 つ以上の DPCM 符号器 811, ..., 81N を含む。例えば、メタデータ符号器 801 が N 個のオリジナル・メタデータ信号を受信するよう構成されている場合、メタデータ符号器 801 は、例えば、正に N 個の DPCM 符号器を含んでもよい。一実施形態において、N 個の DPCM 符号器の各々は図 6 に関して説明したように構成されている。

【0078】

一実施形態において、N 個の DPCM 符号器の各々は、N 個のオリジナル・メタデータ信号  $x_1, \dots, x_N$  の 1 つのメタデータサンプル  $x_i(n)$  を受信し、かつ、前記 DPCM 符号器に入力された前記オリジナル・メタデータ信号  $x_i$  のメタデータサンプル  $x_i(n)$  の各々のために、メタデータ差分信号  $y_i$  の差分サンプル  $y_i(n)$  としての差分値を生成するよう構成されている。一実施形態において、差分サンプル  $y_i(n)$  を生成するステップは、例えば図 6 を参照しながら説明したように実行されてもよい。

40

【0079】

図 8 a のメタデータ符号器 801 は、制御信号  $b(n)$  を受信するよう構成された選択部 830 ("A") を更に含む。

【0080】

選択部 830 は、更に、N 個のメタデータ差分信号  $y_1, \dots, y_N$  を受信するよう構成されている。

【0081】

50

更に図8 aの実施形態において、メタデータ符号器801は、N個のオリジナル・メタデータ信号 $x_1, \dots, x_N$ を量子化してN個の量子化済みメタデータ信号 $q_1, \dots, q_N$ を取得する、量子化部820を含む。そのような実施形態において、量子化部は、N個の量子化済みメタデータ信号を選択部830へと供給するよう構成されてもよい。

【0082】

選択部830は、制御信号 $b(n)$ に依存して、量子化済みメタデータ信号 $q_i$ とDPCM符号化された差分メタデータ信号 $y_i$ とから、処理済みメタデータ信号 $z_i$ を生成するよう構成されてもよい。

【0083】

例えば、制御信号 $b$ が第1状態（例えば $b(n)=0$ ）にあるとき、選択部830は、処理済みメタデータ信号 $z_i$ のメタデータサンプル $z_i(n)$ として、メタデータ差分信号 $y_i$ の差分サンプル $y_i(n)$ を出力するよう構成されてもよい。

10

【0084】

制御信号 $b$ が第1状態とは異なる第2状態（例えば $b(n)=1$ ）にあるとき、選択部830は、処理済みメタデータ信号 $z_i$ のメタデータサンプル $z_i(n)$ として、量子化済みメタデータ信号 $q_i$ のメタデータサンプル $q_i(n)$ を出力するよう構成されてもよい。

【0085】

図8 bは、他の実施形態に係るメタデータ符号器802を示す。

【0086】

図8 bの実施形態において、メタデータ符号器802は量子化部820を含んでおらず、N個の量子化済みメタデータ信号 $q_1, \dots, q_N$ の代わりに、N個のオリジナル・メタデータ信号 $x_1, \dots, x_N$ が選択部830に対して直接的に供給される。

20

【0087】

そのような実施形態において、例えば制御信号 $b$ が第1状態（例えば $b(n)=0$ ）にあるとき、選択部830は、処理済みメタデータ信号 $z_i$ のメタデータサンプル $z_i(n)$ として、メタデータ差分信号 $y_i$ の差分サンプル $y_i(n)$ を出力するよう構成されてもよい。

【0088】

制御信号 $b$ が第1状態とは異なる第2状態（例えば $b(n)=1$ ）にあるとき、選択部830は、処理済みメタデータ信号 $z_i$ のメタデータサンプル $z_i(n)$ として、オリジナル・メタデータ信号 $x_i$ のメタデータサンプル $x_i(n)$ を出力するよう構成されてもよい。

30

【0089】

図9 aは、一実施形態に係るメタデータ復号器901を示す。図9 aに係るメタデータ復号器は、図8 a及び図8 bのメタデータ符号器と対応している。

【0090】

図9 aのメタデータ復号器901は、1つ以上のメタデータ復号器サブユニット911, ..., 91Nを含む。メタデータ復号器901は、1つ以上の処理済みメタデータ信号 $z_1, \dots, z_N$ を受信するよう構成されている。更に、メタデータ復号器901は、制御信号 $b$ を受信するよう構成されている。メタデータ復号器は、制御信号 $b$ に依存して、1つ以上の処理済みメタデータ信号 $z_1, \dots, z_N$ から1つ以上の再生メタデータ信号 $x_1', \dots, x_N'$ を生成するよう構成されている。

40

【0091】

一実施形態において、N個の処理済みメタデータ信号 $z_1, \dots, z_N$ の各々が、メタデータ復号器サブユニット911, ..., 91Nの異なる1つへと供給される。更に、一実施形態によれば、制御信号 $b$ がメタデータ復号器サブユニット911, ..., 91Nの各々に供給される。一実施形態によれば、メタデータ復号器サブユニット911, ..., 91Nの個数は、メタデータ復号器901によって受信される処理済みメタデータ信号 $z_1, \dots, z_N$ の個数と等しい。

【0092】

図9 bは、一実施形態に係る、図9 aのメタデータ復号器サブユニット911, ..., 91Nの1つのメタデータ復号器サブユニット(91i)を示す。メタデータ復号器サブユニ

50

ット91iは、単一の処理済みメタデータ信号 $z_i$ のために復号化を実行するよう構成されている。メタデータ復号器サブユニット91iは、選択部930("B")と加算部910とを含む。

【0093】

メタデータ復号器サブユニット91iは、制御信号 $b(n)$ に依存して、受信された処理済みメタデータ信号 $z_i$ から再生メタデータ信号 $x_i'$ を生成するよう構成されている。

【0094】

この操作は、例えば以下のようにして実現され得る。

【0095】

再生メタデータ信号 $x_i'$ の最後の再生メタデータサンプル $x_i'(n-1)$ が加算部910へと10  
入力される。更に、処理済みメタデータ信号 $z_i$ の実際のメタデータサンプル $z_i(n)$ もまた  
加算部910へと入力される。加算部は、最後の再生メタデータサンプル $x_i'(n-1)$ と実際  
のメタデータサンプル $z_i(n)$ とを加算して合計値 $s_i(n)$ を取得するよう構成されており、そ  
の合計値は選択部930へと入力される。

【0096】

更に、実際のメタデータサンプル $z_i(n)$ もまた、選択部930へと入力される。

【0097】

選択部930は、制御信号 $b$ に依存して、加算部910からの合計値 $s_i(n)$ 、又は実際  
のメタデータサンプル $z_i(n)$ のいずれかを、再生メタデータ信号 $x_i'(n)$ の実際のメタデータ  
サンプル $x_i'(n)$ として選択するよう構成されている。20

【0098】

例えば制御信号 $b$ が第1状態(例えば $b(n)=0$ )にあるとき、制御信号 $b$ は実際のメタデータサンプル $z_i(n)$ が差分値であることを示しており、よって、合計値 $s_i(n)$ が再生メタデータ信号 $x_i'$ の正確な実際のメタデータサンプル $x_i'(n)$ である。選択部930は、制御信号が第1状態にあるとき( $b(n)=0$ のとき)、合計値 $s_i(n)$ を、再生メタデータ信号 $x_i'$ の実際のメタデータサンプル $x_i'(n)$ として選択するよう構成されている。

【0099】

制御信号 $b$ が第1状態とは異なる第2状態(例えば $b(n)=1$ )にあるとき、制御信号 $b$ は実際のメタデータサンプル $z_i(n)$ が差分値でないことを示しており、よって、実際のメタデータサンプル $z_i(n)$ が再生メタデータ信号 $x_i'$ の正確な実際のメタデータサンプル $x_i'(n)$ である。選択部930は、制御信号が第2状態にあるとき( $b(n)=1$ のとき)、実際のメタデータサンプル $z_i(n)$ を、再生メタデータ信号 $x_i'$ の実際のメタデータサンプル $x_i'(n)$ として選択するよう構成されている。30

【0100】

実施形態によれば、メタデータ復号器サブユニット91i'は、さらにユニット920を  
含む。ユニット920は、再生メタデータ信号の実際のメタデータサンプル $x_i'(n)$ を、サンプリング期間の持続時間中、保持するよう構成されている。一実施形態において、これ  
により、 $x_i'(n)$ が生成されつつあるとき、生成された $x'(n)$ のフィードバックが早くなり  
過ぎないことが確保され、 $z_i(n)$ が差分値である場合に、 $x_i'(n)$ が現実に $x_i'(n-1)$ に基づいて40  
生成されるようになる。

【0101】

図9bの実施形態において、選択部930は、制御信号 $b(n)$ に依存して、受信された信号成分 $z_i(n)$ 及び、遅延された出力成分(再生メタデータ信号の既に生成されたメタデータサンプル)と受信された信号成分 $z_i(n)$ との線形結合から、メタデータサンプル $x_i'(n)$ を生成してもよい。

【0102】

以下において、DPCM符号化済み信号は $y_i(n)$ で表され、Bの第2入力信号(合計信号)は $s_i(n)$ で表される。対応する入力成分だけに依存する出力成分について、符号器及び復号器の出力は以下のように表される。

$$z_i(n) = A(x_i(n), y_i(n), b(n))$$

10

20

30

40

50

$$x_i'(n) = B(z_i(n), s_i(n), b(n))$$

## 【0103】

上述した一般的な手法に係る一実施形態に従う解決策は、 $b(n)$ を用いて、DPCM符号化済み信号と量子化済み入力信号との間を切り替えることである。簡素化を目的として、時間インデックス $n$ を省略すると、機能ブロックAとBとは以下のように表すことができる。

## 【0104】

メタデータ符号器801, 802において、選択部830(A)は次のように選択する。

$$A : z_i(x_i, y_i, b) = y_i, \quad b = 0 \text{ のとき } (z_i \text{ は差分値を示す})$$

$$A : z_i(x_i, y_i, b) = x_i, \quad b = 1 \text{ のとき } (z_i \text{ は差分値を示さない})$$

## 【0105】

メタデータ復号器サブユニット91i, 91i'において、選択部930(B)は次のように選択する。

$$B : x_i'(z_i, s_i, b) = s_i, \quad b = 0 \text{ のとき } (z_i \text{ は差分値を示す})$$

$$B : x_i'(z_i, s_i, b) = z_i, \quad b = 1 \text{ のとき } (z_i \text{ は差分値を示さない})$$

## 【0106】

これにより、 $b(n)$ が1と等しいときは常に量子化済み入力信号を伝送し、 $b(n)$ が0のときは常にDPCM信号を伝送することが可能になる。後者の場合、復号器はDPCM復号器となる。

## 【0107】

オブジェクト・メタデータの伝送に適用される場合、このメカニズムは、復号器によってランダムアクセスのために使用され得る、未圧縮のオブジェクト位置を規則正しく伝送するために使用される。

## 【0108】

好ましい実施形態において、メタデータサンプルを符号化するために使用されるビットの数よりも少数のビットが差分値を符号化するために使用される。これらの実施形態は、(例えば $N$ 個の)連続するメタデータサンプルは、多くの場合、僅かしか変化しないという知見に基づいている。例えば、一種類のメタデータサンプルが例えば8ビットで符号化される場合、これらメタデータサンプルは256個の異なる値の中から1つをとり得る。一般的に、(例えば $N$ 個の)連続するメタデータ値のその僅かな変化に起因して、例えば5ビットだけで差分値を符号化することは十分と考えられ得る。従って、差分値が伝送される場合でも、伝送されるビット数は低減され得る。

## 【0109】

一実施形態において、メタデータ符号器210は、1つ以上の処理済みメタデータ信号( $z_1, \dots, z_N$ )の1つ( $z_i$ )の処理済みメタデータサンプル( $z_i(1), \dots, z_i(n)$ )の各々を、制御信号が第1状態( $b(n)=0$ )を示すときには第1のビット数を用いて符号化し、制御信号が第2状態( $b(n)=1$ )を示すときには第2のビット数を用いて符号化するよう構成されており、第1のビット数は第2のビット数よりも少ない。

## 【0110】

好ましい実施形態においては、1つ以上の差分値が伝送され、1つ以上の差分値の各々はメタデータサンプルの各々よりも少ないビットを用いて符号化され、差分値の各々は整数値である。

## 【0111】

一実施形態によれば、メタデータ符号器110は、1つ以上の処理済みメタデータ信号の内の1つの1つ以上のメタデータサンプルを第1のビット数を用いて符号化するよう構成されており、ここで、1つ以上の処理済みメタデータ信号の1つの1つ以上のメタデータサンプルの各々は整数を示す。更に、メタデータ符号器(110)は、1つ以上の差分値を第2のビット数を用いて符号化するよう構成されており、ここで1つ以上の差分値の各々は整数を示し、第2のビット数は第1のビット数よりも少ない。

10

20

30

40

50

## 【0112】

例えば一実施形態において、メタデータサンプルが8ビットで符号化された方位角を表現できると考慮されたい。例えば、その方位角は-90°方位角90°の整数であってもよい。従って、その方位角は181個の異なる値をとり得る。しかし、(例えばN個の)後続の方位角サンプルは、例えば±15以下しか変化しないと想定することができ、その場合、差分値を符号化するために5ビット( $2^5 = 32$ )で十分となり得る。差分値が整数として表現される場合、その差分値を決定することは、伝送されるべき追加的な値を適切な値領域へと自動的に変換することになる。

## 【0113】

例えば、第1オーディオオブジェクトの第1方位角値が60°であり、その後続の値が45°から75°まで変化する場合を考慮されたい。さらに、第2オーディオオブジェクトの第2方位角値が-30°であり、その後続の値が-45°から-15°まで変化する場合を考慮されたい。第1オーディオオブジェクトの両方の後続の値についての差分値、及び第2オーディオオブジェクトの両方の後続の値についての差分値を決定すると、第1方位角値及び第2方位角値の差分値は両方とも-15°から+15°までの値領域内にある。よって、差分値の各々を符号化するために5ビットで十分となり、差分値を符号化するビットシーケンスは、第1方位角の差分値と第2方位角の差分値とに対して同じ意味を持つ。

## 【0114】

次に、実施形態に係るオブジェクト・メタデータ・フレームと、実施形態に係るシンボル表現とについて説明する。

## 【0115】

符号化済みオブジェクト・メタデータは、フレーム内で伝送される。これらのオブジェクト・メタデータ・フレームは、イントラ符号化されたオブジェクトデータ、又はダイナミック・オブジェクトデータの何れかを含むことができ、後者の場合は、伝送された最後のフレームからの変化を含んでいる。

## 【0116】

例えば、オブジェクト・メタデータ・フレームについて、以下のシンタックスの一部又は全部が使用されてもよい。

## 【0117】

## 【表1】

	No. of bits	Mnemonic
<pre>object_metadata() {     has_intracoded_object_metadata;     if (has_intracoded_object_metadata) {         intracoded_object_metadata ();     }     else {         dynamic_object_metadata();     } }</pre>	1	bslbf

## 【0118】

以下に、一実施形態に係るイントラ符号化されたオブジェクトデータについて説明する。

## 【0119】

符号化済みオブジェクト・メタデータのランダムアクセスが、イントラ符号化されたオブジェクトデータを介して実現され(「Iフレーム」)、イントラ符号化されたオブジェクトデータは規則的な格子上で(例えば長さ1024の32フレーム毎に)サンプリング

10

20

30

40

50

された量子化済み値を含む。これら I フレームは、例えば次のようなシンタックスを有し、その中で、position\_azimuth, position\_elevation, position\_radius 及び gain\_factor は現時点の量子化済み値を特定する。

【0 1 2 0】

【表 2 - 1】

	No. of bits	Mnemonic	
intracoded_object_metadata()			
{			
if (num_objects>1) {			
fixed_azimuth;	1	bslbf	10
if (fixed_azimuth) {			
default_azimuth;	8	tcimsbf	
}			
else {			
common_azimuth;	1	bslbf	
if (common_azimuth) {			
default_azimuth;	8	tcimsbf	
}			
else {			
for (o=1:num_objects) {			
position_azimuth[o];	8	tcimsbf	20
}			
}			
}			
fixed_elevation;	1	bslbf	
if (fixed_azimuth) {			
default_elevation;	6	tcimsbf	
}			
else {			
common_elevation;	1	bslbf	
if (common_azimuth) {			
default_elevation;	6	tcimsbf	30
}			
else {			
for (o=1:num_objects) {			
position_azimuth[o];	6	tcimsbf	
}			
}			
}			
fixed_radius;	1	bslbf	
if (fixed_azimuth) {			
default_radius;	4	tcimsbf	
}			40
else {			
common_radius;	1	bslbf	
if (common_azimuth) {			
default_radius;	4	tcimsbf	
}			
else {			
for (o=1:num_objects) {			
position_radius [o];	4	tcimsbf	
}			
}			

【表 2 - 2】

```

    }
    fixed_gain;
    if (fixed_azimuth) {
        default_gain;
    }
    else {
        common_gain;
        if (common_azimuth) {
            default_gain;
        }
        else {
            for (o=1:num_objects) {
                gain_factor [o];
            }
        }
    }
}
else {
    position_azimuth;
    position_elevation;
    position_radius;
    gain_factor;
}
}

```

【 0 1 2 1 】

以下に、一実施形態に係るダイナミック・オブジェクトデータについて説明する。

【 0 1 2 2 】

DPCMデータは、例えば以下のシンタックスを有するダイナミック・オブジェクト・フレーム内で伝送される。

【 0 1 2 3 】

【表3】

	No. of bits	Mnemonic
<pre>dynamic_object_metadata() {     flag_absolute;     for (o=1:num_objects) {         has_object_metadata;         if (has_object_metadata) {             single_dynamic_object_metadata( flag_absolute );         }     } }</pre>	1 1	bslbf bslbf

【 0 1 2 4 】

【表 4】

	No. of bits	Mnemonic	
single_dynamic_object_metadata ( flag_absolute ) {			
if ( flag_absolute ) {			
if (!fixed_azimuth*) {			
position_azimuth;	8	tcimsbf	
}			
if (!fixed_elevation*) {			
position_elevation;	6	tcimsbf	
}			
if (!fixed_radius*) {			10
position_radius;	4	tcimsbf	
}			
if (!fixed_gain*) {			
gain_factor;	7	tcimsbf	
}			
}			
else {			
nbits;	3	uimsbf	
if (!fixed_azimuth*) {			
flag_azimuth;	1	bslbf	20
if (flag_azimuth) {			
position_azimuth_difference ;	num_bits	tcimsbf	
}			
}			
if (!fixed_elevation*) {			
flag_elevation;	1	bslbf	
if (flag_elevation) {			
position_elevation_difference ;	min(num_bits,7)	tcimsbf	
}			
}			
if (!fixed_radius*) {			30
flag_radius;	1	bslbf	
if (flag_radius) {			
position_radius_difference ;	min(num_bits,5)	tcimsbf	
}			
}			
if (!fixed_gain*) {			
flag_gain;	1	bslbf	
if (flag_gain) {			
gain_factor_difference ;	min(num_bits,8)	tcimsbf	
}			
}			
}			
Note: num_bits = nbits + 2;			
Footnote *: Given by the preceding			
intracoded_object_data()-frame			

【0 1 2 5】

特に、一実施形態において、上述のマクロは例えば以下のようない意味を有してもよい。

【0 1 2 6】

一実施形態に係るobject\_data() payloadsの定義：

has\_intracoded\_object\_metadata : そのフレームがイントラ符号化されたか又は差分符号化されたかを示す

【0 1 2 7】

一実施形態に係る `intracoded_object_metadata()` payloads の定義 :

`fixed_azimuth` : 全てのオブジェクトについて方位角値が固定か否かを示すフラグであり、`dynamic_object_metadata()` の場合には伝送されない

`default_azimuth` 固定又は共通の方位角の値を定義する

`common_azimuth` : 全てのオブジェクトについて共通の方位角が使用されるか否かを示す

`position_azimuth` : 共通の方位角値がない場合、各オブジェクトのための値が伝送される

`fixed_elevation` : 全てのオブジェクトについて仰角値が固定か否かを示すフラグであり、`dynamic_object_metadata()` の場合には伝送されない

`default_elevation` : 固定又は共通の仰角の値を定義する

`common_elevation` : 全てのオブジェクトについて共通の仰角が使用されるか否かを示す

10

`position_elevation` : 共通の仰角値がない場合、各オブジェクトのための値が伝送される

`fixed_radius` : 全てのオブジェクトについて半径が固定か否かを示すフラグであり、`dynamic_object_metadata()` の場合には伝送されない

`default_radius` : 共通の半径の値を定義する

`common_radius` : 全てのオブジェクトについて共通の半径値が使用されるか否かを示す

`position_radius` : 共通の半径値がない場合、各オブジェクトのための値が伝送される

`fixed_gain` : 全てのオブジェクトについてゲインファクタが固定か否かを示すフラグであり、`dynamic_object_metadata()` の場合には伝送されない

`default_gain` : 固定又は共通のゲインファクタの値を定義する

`common_gain` : 全てのオブジェクトについて共通のゲイン値が使用されるか否かを示す

20

`gain_factor` : 共通のゲイン値がない場合、各オブジェクトのための値が伝送される

`position_azimuth` : 単一のオブジェクトだけがある場合、その方位角である

`position_elevation` : 単一のオブジェクトだけがある場合、その仰角である

`position_radius` : 単一のオブジェクトだけがある場合、その半径である

`gain_factor` : 単一のオブジェクトだけがある場合、そのゲインファクタである

【 0 1 2 8 】

一実施形態に係る `dynamic_object_metadata()` payloads の定義 :

`flag_absolute` : 構成要素の値が差分的に伝送されるか又は絶対値で伝送されるかを示す

`has_object_metadata` : ビットストリーム内にオブジェクトデータが存在するか否かを示す

30

【 0 1 2 9 】

一実施形態に係る `single_dynamic_object_metadata()` payloads の定義 :

`position_azimuth` : 値が固定でない場合の方位角の絶対値

`position_elevation` : 値が固定でない場合の仰角の絶対値

`position_radius` : 値が固定でない場合の半径の絶対値

`gain_factor` : 値が固定でない場合のゲインファクタの絶対値

`nbits` : 差分値を表現するために必要なビットの数

`flag_azimuth` : 方位角値が変化するか否かを示すオブジェクト毎のフラグ

`position_azimuth_difference` : 以前の値と活性値との間の差

`flag_elevation` : 仰角値が変化するか否かを示すオブジェクト毎のフラグ

40

`position_elevation_difference` : 以前の値と活性値との間の差の値

`flag_radius` : 半径が変化するか否かを示すオブジェクト毎のフラグ

`position_radius_difference` : 以前の値と活性値との間の差

`flag_gain` : ゲインが変化するか否かを示すオブジェクト毎のフラグ

`gain_factor_difference` : 以前の値と活性値との間の差

【 0 1 3 0 】

先行技術においては、低ビットレートで許容可能なオーディオ品質が得られるように、チャネル符号化と一方としオブジェクト符号化を他方として結合するような、柔軟性のある技術は存在しない。

【 0 1 3 1 】

50

この制約は 3D オーディオコーデックシステムにより克服できる。以下に、3D オーディオコーデックシステムについて説明する。

#### 【0132】

図 10 は、本発明の一実施形態に係る 3D オーディオ符号器を示す。この 3D オーディオ符号器は、オーディオ入力データ 101 を符号化してオーディオ出力データ 501 を取得するよう構成されている。3D オーディオ符号器は、CH で示された複数のオーディオチャネルと OBJ で示された複数のオーディオオブジェクトとを受信する入力インターフェイスを備える。更に、図 10 に示すように、入力インターフェイス 1100 は、複数のオーディオオブジェクト OBJ の 1 つ以上に関連するメタデータを追加的に受信する。更に、3D オーディオ符号器は、複数のオブジェクトと複数のチャネルとをミキシングして複数のプレミクス済みチャネルを取得するミキサー 200 を備え、各プレミクス済みチャネルは 1 つのチャネルのオーディオデータと少なくとも 1 つのオブジェクトのオーディオデータとを含む。  
10

#### 【0133】

更に、3D オーディオ符号器は、コア符号器入力データをコア符号化するコア符号器 300 と、複数のオーディオオブジェクトの 1 つ以上に関連するメタデータを圧縮するメタデータ圧縮部 400 とを備える。

#### 【0134】

更に、3D オーディオ符号器は、ミキサーとコア符号器及び / 又は出力インターフェイス 500 を複数の動作モードの 1 つで制御する、モード制御部 600 を備え、第 1 モードでは、コア符号器は、入力インターフェイス 1100 により受信された複数のオーディオチャネル及び複数のオーディオオブジェクトを、ミキサーによる相互作用なしに、即ちミキサー 200 によるミキシングなしに、符号化するよう構成されている。しかし、ミキサー 200 が活性化していた第 2 モードでは、コア符号器は、複数のミクス済みチャネル、即ちブロック 200 により生成された出力を符号化する。後者の場合、それ以上のオブジェクトデータを符号化しないことが好ましい。代わりに、オーディオオブジェクトの位置を示すメタデータは、メタデータによって示された通りにオブジェクトをチャネル上へとレンダリングするために、ミキサー 200 によって既に使用されている。換言すれば、ミキサー 200 は複数のオーディオオブジェクトに関連するメタデータを、オーディオオブジェクトをプレレンダリングするために使用し、次に、プレレンダリングされたオーディオオブジェクトはチャネルとミキシングされて、ミキサーの出力においてミクス済みチャネルが得られる。この実施形態では、如何なるオブジェクトも必ずしも伝送される必要がなく、このことは、ブロック 400 により出力される圧縮済みメタデータにも当てはまる。しかし、インターフェイス 1100 に入力された全てのオブジェクトがミキシングされる訳でなく、所定量のオブジェクトだけがミキシングされる場合には、ミキシングされていない残りのオブジェクト及び関連するメタデータだけが、コア符号化 300 又はメタデータ圧縮部 400 へとそれぞれ伝送される。  
20  
30

#### 【0135】

図 10 において、メタデータ圧縮部 400 は、上述した実施形態の 1 つに従う、符号化済みオーディオ情報を生成する装置 250 のメタデータ符号器 210 である。更に、図 10 において、ミキサー 200 及びコア符号器 300 は一緒に、上述した実施形態の 1 つに従う符号化済みオーディオ情報を生成する装置 250 のオーディオ符号器 220 を形成する。  
40

#### 【0136】

図 12 は、S A O C 符号器 800 を追加的に含む、3D オーディオ符号器の更なる実施形態を示す。S A O C 符号器 800 は、空間オーディオオブジェクト符号器入力データから、1 つ以上の転送チャネル及びパラメトリックデータを生成するよう構成されている。図 12 に示すように、空間オーディオオブジェクト符号器入力データは、プレレンダラー / ミキサーによって処理されなかったオブジェクトである。代替的に、個別のチャネル / オブジェクト符号化が活性化しているモード 1 におけるように、プレレンダラー / ミキサ  
50

ーが迂回されていたと仮定すると、入力インターフェイス 1100 に入力された全てのオブジェクトは SAOC 符号器 800 により符号化される。

【0137】

更に、図 12 に示すように、コア符号器 300 は好ましくは、USAC 符号器、即ち MPEG-USAC 標準 (USAC = 統合されたスピーチ及びオーディオ符号化) の中で定義されかつ標準化されているような符号器として構成される。図 12 に示す全体的な 3D オーディオ符号器の出力は、個別のデータタイプについてコンテナ状構造を有している MPEG4 データストリームである。更に、メタデータは「OAM」データとして示され、図 10 におけるメタデータ圧縮部 400 は、圧縮済み OAM データを取得する OAM 符号器 400 に対応し、その圧縮済み OAM データは USAC 符号器 300 へ入力され、USAC 符号器 300 は、図 12 に示すように、MP4 出力データストリームを取得するための出力インターフェイスを追加的に含み、その MP4 出力データストリームは符号化済みチャネル / オブジェクトデータだけでなく圧縮済み OAM データをも有する。

【0138】

図 12 において、OAM 符号器 400 は、上述した実施形態の 1 つに従う、符号化済みオーディオ情報を生成する装置 250 のメタデータ符号器 210 である。更に、図 12 では、SAOC 符号器 800 と USAC 符号器 300 とは一緒に、上述した実施形態の 1 つに従う、符号化済みオーディオ情報を生成する装置 250 のオーディオ符号器 220 を形成する。

【0139】

図 14 は、3D オーディオ符号器の更なる実施形態を示し、ここでは、図 12 とは対照的に、SAOC 符号器は SAOC 符号化アルゴリズムを用いて、このモードにおいて活性化していないプレレンダラー / ミキサー 200 により提供されたチャネルを符号化するか、又は代替的に、プレレンダリング済みチャネル + オブジェクトを SAOC 符号化するか、の何れかを実行するよう構成されている。従って、図 14 においては、SAOC 符号器 800 は 3 種類の異なる入力データ、即ち、プレレンダリング済みオブジェクトを持たないチャネル、チャネル及びプレレンダリング済みオブジェクト、又はオブジェクトのみ、に対して作動できる。更に、追加的な OAM 復号器 420 を図 14 に設け、SAOC 符号器 800 がその処理のために復号器側と同じデータを使用できるように、即ち、オリジナル OAM データよりも寧ろ損失の多い圧縮により得られたデータを使用できるようにすることが好ましい。

【0140】

図 14 の 3D オーディオ符号器は、複数の個別モードにおいて作動できる。

【0141】

図 10 の文脈の中で説明した第 1 モード及び第 2 モードに加え、図 14 の 3D オーディオ符号器は追加的に第 3 モードでも作動でき、そのモードでは、プレレンダラー / ミキサー 200 が活性化していなかった場合、コア符号器が個別のオブジェクトから 1 つ以上の転送チャネルを生成する。代替的又は追加的に、この第 3 モードにおいて、図 10 のミキサー 200 に対応するプレレンダラー / ミキサー 200 が活性化していなかった場合、SAOC 符号器 800 はオリジナルチャネルから 1 つ以上の代替的又は追加的な転送チャネルを生成することができる。

【0142】

最後に、3D オーディオ符号器が第 4 モードで構成されている場合、SAOC 符号器 800 は、プレレンダラー / ミキサー によって生成されたチャネル + プレレンダリング済みオブジェクトを符号化することができる。そのため、第 4 モードにおいては、最低ビットレートのアプリケーションが次のような事実により良好な品質を提供できる。なぜなら、チャネルとオブジェクトとが、個別の SAOC 転送チャネル及び図 12 と図 14 の中で「SAOC - SI」として示すような関連するサイド情報へと完全に変換されており、加えて、この第 4 モードでは如何なる圧縮済みメタデータも伝送される必要がないからである。

10

20

30

40

50

【 0 1 4 3 】

図14において、OAM符号器400は、上述した実施形態の1つに従う符号化済みオーディオ情報を生成する装置250のメタデータ符号器210である。更に、図14において、SAOC符号器800とUSAC符号器300とは一緒に、上述した実施形態の1つに従う符号化済みオーディオ情報を生成する装置250のオーディオ符号器220を形成する。

【 0 1 4 4 】

—実施形態によれば、オーディオ入力データ101を符号化してオーディオ出力データ501を取得する装置が提供される。そのオーディオ入力データ101を符号化する装置は、

- 複数のオーディオチャネルと複数のオーディオオブジェクトと複数のオーディオオブジ

エクトの1つ以上に関連するメタデータとを愛信する、入カインターフェイス1100と

1

- 複数のオブジェクトと複数のチャネルとをミキシングして、複数のプレミクス済みチャネルを取得するミキサー 200 であって、各プレミクス済みチャネルが 1 つのチャネルのオーディオデータと少なくとも 1 つのオブジェクトのオーディオデータとを含む、ミキサー 200 と、

- 上述したようなメタデータ符号器とオーディオ符号器とを含む、符号化済みオーディオ情報を生成する装置 250 と、

を備える。

【 0 1 4 5 】

符号化済みオーディオ情報を生成する装置 250 のオーディオ符号器 220 は、コア符号器入力データを符号化するコア符号器 (300) である。

【 0 1 4 6 】

符号化済みオーディオ情報を生成する装置 250 のメタデータ符号器 210 は、複数のオーディオオブジェクトの 1 つ以上に関連するメタデータを圧縮するメタデータ圧縮部 400 である。

〔 0 1 4 7 〕

図11は本発明の一実施形態に係る3Dオーディオ復号器を示す。その3Dオーディオ復号器は、入力として符号化済みオーディオデータ、即ち図10のデータ501を受信する。

[ 0 1 4 8 ]

3 D オーディオ復号器は、メタデータ解凍部 1400 と、コア復号器 1300 と、オブジェクト処理部 1200 と、干渉制御部 1600 と、後処理部 1700 とを備える。

[ 0 1 4 9 ]

具体的には、その 3D オーディオ復号器は符号化済みオーディオデータを復号化するよう構成されており、入力インターフェイスは符号化済みオーディオデータを受信するよう構成されており、符号化済みオーディオデータは、複数の符号化済みチャネルと、複数の符号化済みオブジェクトと、あるモードにおいて複数のオブジェクトに関連する圧縮済みメタデータとを含む

[0 1 5 0 1]

更に、コア復号器 1300 は複数の符号化済みチャネルと複数の符号化済みオブジェクトとを復号化するよう構成されており、追加的に、メタデータ解凍部は圧縮済みメタデータを解凍するよう構成されている。

【0 1 5 1】

更に、オブジェクト処理部 1200 は、コア復号器 1300 により生成された複数の復号化済みオブジェクトを解凍済みメタデータを使用して処理し、オブジェクトデータと復号化済みチャネルとを含む所定数の出力チャネルを得るよう構成されている。符号 1205 で示されたこれら出力チャネルは、次に後処理部 1700 へと入力される。後処理部 1700 は、出力チャネル 1205 の数を、バイナリ出力フォーマット又は 5-1 や 7-

10

20

30

40

50

. 1などの出力フォーマットのようなラウドスピーカ出力フォーマットであり得る、ある出力フォーマットへと変換するよう構成されている。

【0152】

好ましくは、3Dオーディオ復号器は、符号化済みデータを分析してモード指示を検出するよう構成された、モード制御部1600を備える。従って、そのモード制御部1600が図11の入力インターフェイス1100に接続されている。しかし、代替的に、モード制御部が必ずしも存在する必要はない。代わりに、柔軟性のあるオーディオ復号器は、ユーザー入力や任意の他の制御のような、他の如何なる種類の制御データによってもプリセットされ得る。好ましくはモード制御部1600により制御される図11の3Dオーディオ復号器は、その一方では、オブジェクト処理部を迂回して複数の復号化済みチャネルを後処理部1700へと供給するよう構成される。これは、図10の3Dオーディオ符号器においてモード2が適用されていた場合のモード2における作動であり、即ちプレレンダリング済みチャネルだけが受信される場合である。代替的に、3Dオーディオ符号器においてモード1が適用されていた場合、即ち3Dオーディオ符号器が個別のチャネル／オブジェクト符号化を実行していた場合、オブジェクト処理部1200は迂回されず、複数の復号化済みチャネルと複数の復号化済みオブジェクトとが、メタデータ解凍部1400によって生成された解凍済みメタデータと一緒にオブジェクト処理部1200へと供給される。

10

【0153】

好ましくは、モード1又はモード2が適用されるべきかどうかの指示は符号化済みオーディオデータの中に含まれており、よって、モード制御部1600がモード指示を検出するために符号化済みデータを分析する。符号化済みオーディオデータは符号化済みチャネルと符号化済みオブジェクトとを含むとモード指示が示す場合には、モード1が使用され、他方、符号化済みオーディオデータはオーディオオブジェクトを何も含まない、即ち、図10の3Dオーディオ符号器のモード2によって得られたプレレンダリング済みチャネルだけを含むとモード指示が示す場合には、モード2が適用される。

20

【0154】

図11において、メタデータ解凍部1400は、上述の実施形態の1つに従い1つ以上のオーディオチャネルを生成する装置100のメタデータ復号器110である。更に図11において、コア復号器1300とオブジェクト処理部1200と後処理部1700とは一緒に、上述の実施形態の1つに従い1つ以上のオーディオチャネルを生成する装置100のオーディオ復号器120を形成する。

30

【0155】

図13は、図11の3Dオーディオ復号器と比較した好適な実施形態を示し、図13の実施形態は図12のオーディオ符号器に対応する。図11の3Dオーディオ復号器の構成に加えて、図13の3Dオーディオ復号器はS A O C復号器1800を含む。更に、図11のオブジェクト処理部1200は、別個のオブジェクトレンダラー1210とミキサー1220として構成されているが、モードに依存して、オブジェクトレンダラー1210の機能はS A O C復号器1800によっても実行され得る。

【0156】

40

更に、後処理部1700は、バイノーラルレンダラー1710又はフォーマット変換部1720として構成され得る。代替的に、図11のデータ1205の直接的な出力もまた、1730で示されるように構成され得る。従って、より小さなフォーマットが要求される場合には、柔軟性を持ち、かつ次に後処理するために、復号器内の処理は22.2や32などの最大数のチャネルに対して実行することが好ましい。しかしながら、5.1フォーマットのような小さなフォーマットだけが要求されることが正に最初から明白になる場合には、図15におけるショートカット1727で示すように、不要なアップミクス操作及び後続のダウンミクス操作を防止するための、S A O C復号器及び／又はU S A C復号器に対するある制御を適用し得ることが望ましい。

【0157】

50

本発明の好適な実施形態において、オブジェクト処理部1200はS A O C復号器1800を含み、そのS A O C復号器は、コア復号器により出力される1つ以上の転送チャネル及び関連するパラメトリックデータを復号化し、かつ解凍済みメタデータを使用して、複数のレンダリング済みオーディオオブジェクトを取得するよう構成されている。この目的で、O A M出力がボックス1800に接続されている。

#### 【0158】

更に、オブジェクト処理部1200は、コア復号器により出力された復号化済みオブジェクトをレンダリングするよう構成されており、そのオブジェクトはS A O C転送チャネルの中で符号化されたものではなく、オブジェクトレンダラー1210により示されるよう、典型的には単一チャネル化された構成要素の中で個別に符号化されたものである。更に、復号器は、ミキサーの出力をラウドスピーカへと出力するための、出力1730に対応する出力インターフェイスを備える。

10

#### 【0159】

更なる実施形態において、オブジェクト処理部1200は、符号化済みオーディオ信号又は符号化済みオーディオチャネルを表現している1つ以上の転送チャネル及び関連するパラメトリックサイド情報を復号化する、空間オーディオオブジェクト符号化・復号器1800を含む。その空間オーディオオブジェクト符号化・復号器は、関連するパラメトリック情報及び解凍済みメタデータを、出力フォーマットを直接的にレンダリングするためには使用可能な、例えばS A O Cの初期バージョンで定義されているような、符号変換済みパラメトリックサイド情報をヘッド符号変換するよう構成されている。後処理部1700は、復号化済み転送チャネルと符号変換済みパラメトリックサイド情報を使用して、出力フォーマットのオーディオチャネルを計算するよう構成されている。後処理部により実行される処理は、M P E Gサラウンド処理と類似していてもよく、又はB C C処理などのような他の如何なる処理であってもよい。

20

#### 【0160】

更なる一実施形態において、オブジェクト処理部1200は、(コア復号器による)復号化済み転送チャネルとパラメトリックサイド情報を使用して、出力フォーマットのためのチャネル信号を直接的にアップミクス及びレンダリングするよう構成された、空間オーディオオブジェクト符号化・復号器1800を含む。

#### 【0161】

30

更にかつ重要なことに、図11のオブジェクト処理部1200は、チャネルとミキシングされたプレレンダリング済みオブジェクトが存在する場合、即ち図10のミキサー200が活性化していた場合、入力としてU S A C復号器1300により出力されたデータを直接的に受信する、ミキサー1220をさらに備える。加えて、ミキサー1220は、S A O C復号化を用いずにオブジェクトレンダリングを実行しているオブジェクトレンダラーからのデータを受信する。更にミキサーは、S A O C復号器出力データ、即ちS A O Cレンダリング済みオブジェクトを受信する。

#### 【0162】

ミキサー1220は、出力インターフェイス1730とバイノーラルレンダラー1710とフォーマット変換部1720とに接続されている。バイノーラルレンダラー1710は、頭部関連伝達関数又はバイノーラル室内インパルス応答(B R I R)を使用して、出力チャネルを2つのバイノーラルチャネルへとレンダリングするよう構成されている。フォーマット変換部1720は、出力チャネルを、ミキサーの出力チャネル1205よりも少数のチャネルを有する出力フォーマットへと変換するよう構成されており、そのフォーマット変換部1720は、5.1スピーカなどの再生レイアウトについての情報を要求する。

40

#### 【0163】

図13において、O A M復号器1400は、上述した実施形態の1つに従って1つ以上のオーディオチャネルを生成する装置100のメタデータ復号器110である。更に、図13において、オブジェクトレンダラー1210とU S A C復号器1300とミキサー1

50

220とは一緒に、上述した実施形態の1つに従って1つ以上のオーディオチャネルを生成する装置100のオーディオ復号器120を形成する。

【0164】

図15の3Dオーディオ復号器は、図13の3Dオーディオ復号器とは以下の点で異なる。即ち、SAOC復号器は、レンダリング済みオブジェクトだけでなくレンダリング済みチャネルをも生成しており、このことは、図14の3Dオーディオ符号器が使用され、チャネル/プレレンダリング済みオブジェクトとSAOC符号器800の入力インターフェイスとの間の接続900が活性化している場合であるという点である。

【0165】

更に、ベクトル方式振幅パニング(VBAP)ステージ1810は、SAOC復号器から再生レイアウトについての情報を受信し、かつSAOC復号器に対してレンダリング行列を出力するよう構成され、その結果、SAOC復号器が、ミキサーの更なる動作を必要とせずに、レンダリング済みチャネルを高いチャネルフォーマット1205で、即ち32個のラウドスピーカに提供できるようになる。

10

【0166】

VBAPブロックは、好適には復号化済みOAMデータを受信してレンダリング行列を導出する。より一般的には、VBAPブロックは、再生レイアウトの幾何学的情報だけでなく、その再生レイアウト上で入力信号がレンダリングされるべき位置の幾何学的情報をも要求することが好ましい。この幾何学的入力データは、オブジェクトについてのOAMデータであってもよく、又は、SAOCを用いて伝送されたチャネルについてのチャネル位置情報であってもよい。

20

【0167】

しかしながら、ある特異な出力インターフェイスだけが要求される場合、VBAPステージ1810は、例えば5.1出力について要求されたレンダリング行列を既に供給することができる。その場合、SAOC復号器1800は、SAOC転送チャネルと関連するパラメトリックデータと解凍済みメタデータとから、直接的レンダリング、即ちミキサー1220の相互作用を何も受けずに、要求された出力フォーマットへの直接的なレンダリングを実行する。しかしながら、モード間のあるミキシングが適用される場合、即ち、複数のチャネルがSAOC符号化されているが、全てのチャネルがSAOC符号化されてはいない場合、複数のオブジェクトがSAOC符号化されているが、全てのオブジェクトがSAOC符号化されてはいない場合、又は、プレレンダリング済みオブジェクトとチャネルとのある量だけがSAOC復号化され、残りのチャネルがSAOC処理されない場合には、ミキサーは、個別の入力部分からのデータ、即ちコア復号器1300とオブジェクトレンダラー1210とSAOC復号器1800とからの直接的なデータを、結合するであろう。

30

【0168】

図15において、OAM復号器1400は、上述した実施形態の1つに従って1つ以上のオーディオチャネルを生成する装置100のメタデータ復号器110である。更に図15において、オブジェクトレンダラー1210とUSAC復号器1300とミキサー1220とは一緒に、上述した実施形態の1つに従って1つ以上のオーディオチャネルを生成する装置100のオーディオ復号器120を形成する。

40

【0169】

符号化済みオーディオデータを復号化する装置が提供される。その符号化済みオーディオデータを復号化する装置は、

- 符号化済みオーディオデータを受信する入力インターフェイス1100であって、符号化済みオーディオデータは、複数の符号化済みチャネル、複数の符号化済みオブジェクト、又は複数のオブジェクトに関連する圧縮済みメタデータを含む、インターフェイス1100と、
- メタデータ復号器110と、上述したように1つ以上のオーディオチャネルを生成するオーディオチャネル生成部120とを含む、装置100と、

50

を備える。

【0170】

1つ以上のオーディオチャネルを生成する装置100のメタデータ復号器110は、圧縮済みメタデータを解凍するメタデータ解凍部400である。

【0171】

1つ以上のオーディオチャネルを生成する装置100のオーディオチャネル生成部120は、複数の符号化済みチャネルと複数の符号化済みオブジェクトとを復号化する、コア復号器1300を備える。

【0172】

更に、オーディオチャネル生成部120は、複数の復号化済みオブジェクトを解凍済みメタデータを使用して処理し、オブジェクト及び復号化済みチャネルからオーディオデータを含む幾つかの出力チャネル1205を取得する、オブジェクト処理部1200を更に備える。

【0173】

更に、オーディオチャネル生成部120は、幾つかの出力チャネル1205を出力フォーマットへと変換する後処理部1700を更に備える。

【0174】

これまで装置の文脈で幾つかの態様を示してきたが、これらの態様は対応する方法の説明をも表しており、1つのブロック又は装置が1つの方法ステップ又は方法ステップの特徴に対応することは明らかである。同様に、方法ステップを説明する文脈で示した態様もまた、対応する装置の対応するブロックもしくは項目又は特徴を表している。

【0175】

本発明の分解された信号は、デジタル記憶媒体に記憶されることができ、又は、インターネットのような無線伝送媒体もしくは有線伝送媒体などの伝送媒体を介して伝送されることもできる。

【0176】

所定の構成要件にもよるが、本発明の実施形態は、ハードウェア又はソフトウェアにおいて構成可能である。この構成は、その中に格納される電子的に読み取り可能な制御信号を有し、本発明の各方法が実行されるようにプログラム可能なコンピュータシステムと協働する（又は協働可能な）、デジタル記憶媒体、例えばフレキシブルディスク、DVD、CD、ROM、PROM、EPROM、EEPROM、フラッシュメモリなどのデジタル記憶媒体を使用して実行することができる。

【0177】

本発明に従う幾つかの実施形態は、上述した方法の1つを実行するようプログラム可能なコンピュータシステムと協働可能で、電子的に読み取り可能な制御信号を有する非一時的なデータキャリアを含む。

【0178】

一般的に、本発明の実施例は、プログラムコードを有するコンピュータプログラム製品として構成することができ、そのプログラムコードは当該コンピュータプログラム製品がコンピュータ上で作動するときに、本発明の方法の一つを実行するよう作動可能である。そのプログラムコードは例えば機械読み取り可能なキャリアに記憶されていても良い。

【0179】

本発明の他の実施形態は、上述した方法の1つを実行するための、機械読み取り可能なキャリアに格納されたコンピュータプログラムを含む。

【0180】

換言すれば、本発明の方法のある実施形態は、そのコンピュータプログラムがコンピュータ上で作動するときに、上述した方法の1つを実行するためのプログラムコードを有するコンピュータプログラムである。

【0181】

本発明の他の実施形態は、上述した方法の1つを実行するために記録されたコンピュ-

10

20

30

40

50

タプログラムを含む、データキャリア（又はデジタル記憶媒体、又はコンピュータ読み取り可能な媒体）である。

【0182】

本発明の他の実施形態は、上述した方法の1つを実行するためのコンピュータプログラムを表現するデータストリーム又は信号列である。そのデータストリーム又は信号列は、例えばインターネットのようなデータ通信接続を介して伝送されるよう構成されても良い。

【0183】

他の実施形態は、上述した方法の1つを実行するように構成又は適応された、例えばコンピュータ又はプログラム可能な論理デバイスのような処理手段を含む。

10

【0184】

他の実施形態は、上述した方法の1つを実行するためのコンピュータプログラムがインストールされたコンピュータを含む。

【0185】

幾つかの実施形態においては、（例えば書換え可能ゲートアレイのような）プログラム可能な論理デバイスが、上述した方法の幾つか又は全ての機能を実行するために使用されても良い。幾つかの実施形態では、書換え可能ゲートアレイは、上述した方法の1つを実行するためにマイクロプロセッサと協働しても良い。一般的に、そのような方法は、好適には任意のハードウェア装置によって実行される。

【0186】

上述した実施形態は、本発明の原理を単に例示的に示したに過ぎない。本明細書に記載した構成及び詳細について修正及び変更が可能であることは、当業者にとって明らかである。従って、本発明は、本明細書に実施形態の説明及び解説の目的で提示した具体的詳細によって限定されるものではなく、添付した特許請求の範囲によってのみ限定されるべきである。

20

【図1】

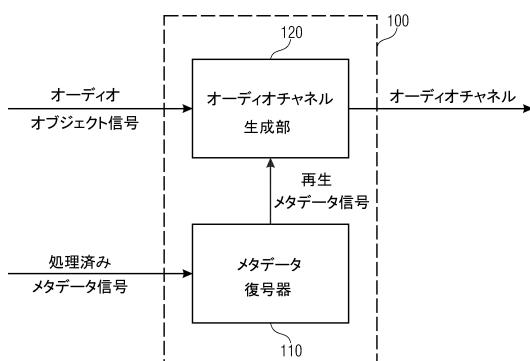


FIGURE 1

【図2】

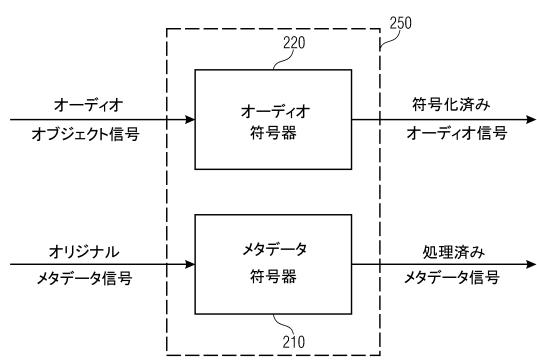


FIGURE 2

【図3】

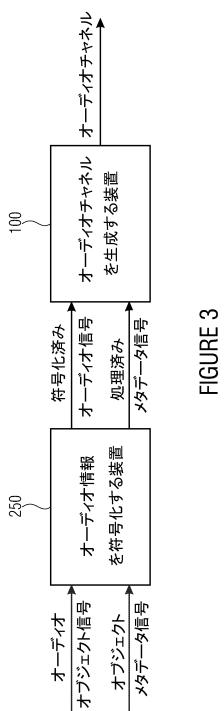


FIGURE 3

【図4】

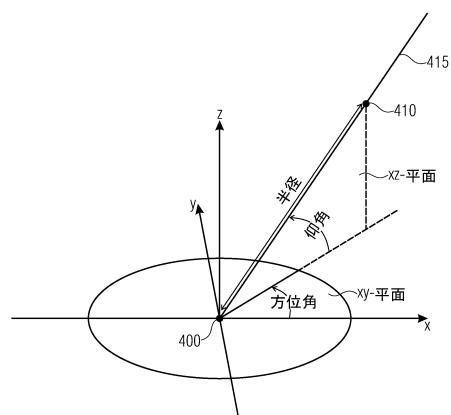


FIGURE 4

【図5】

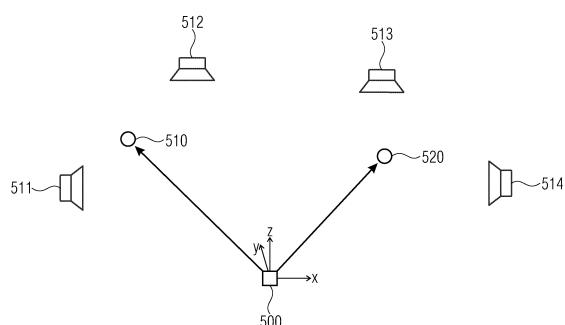


FIGURE 5

【図6】

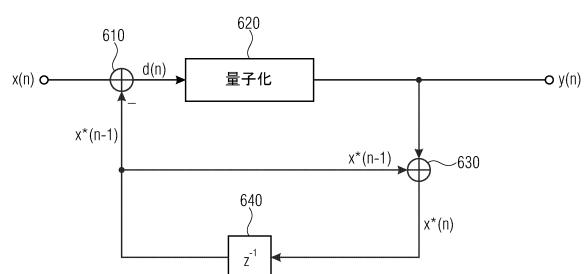


FIGURE 6

【図7】

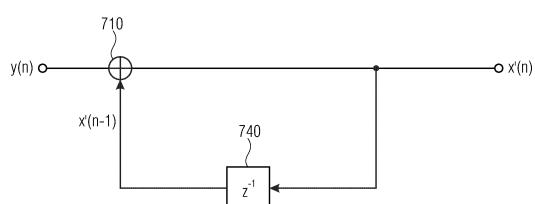


FIGURE 7

【図 8 a】

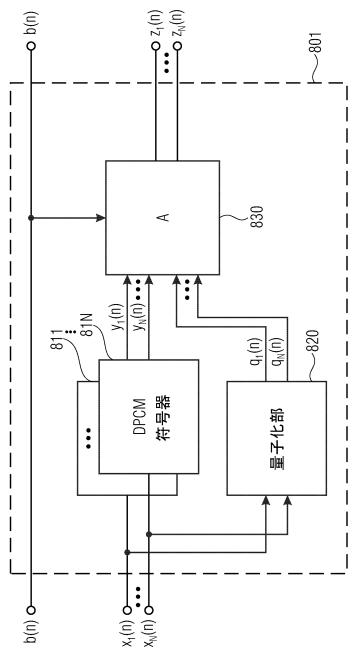


FIGURE 8A

【図 8 b】

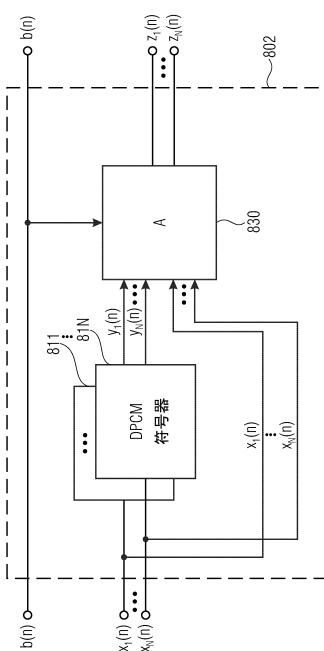


FIGURE 8B

【図 9 a】

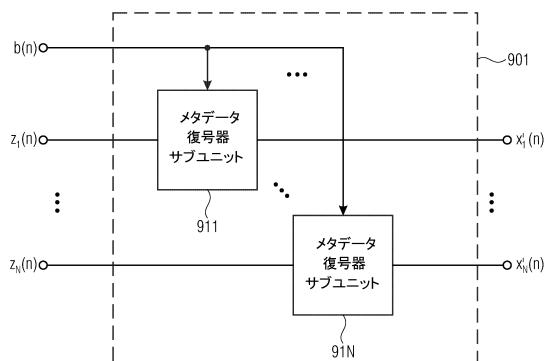


FIGURE 9A

【図 9 b】

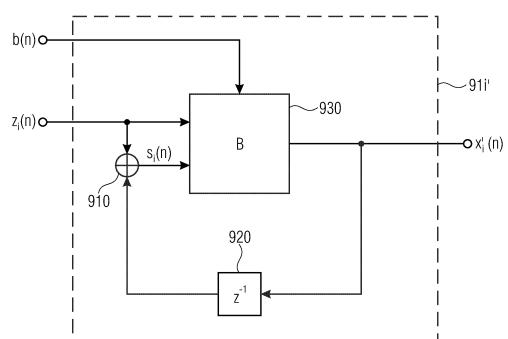


FIGURE 9B

【図10】

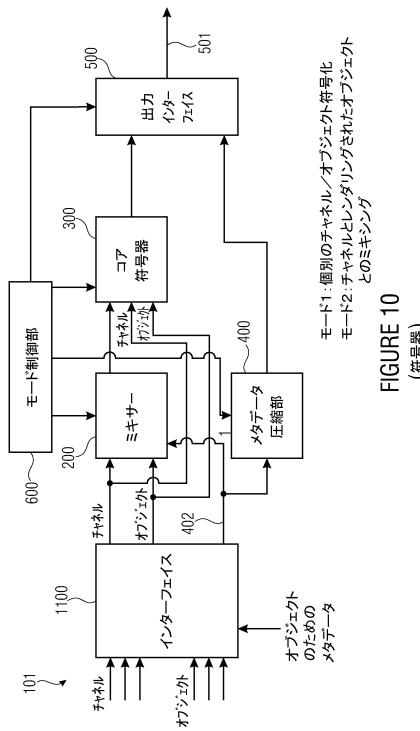


FIGURE 10  
(符号器)

### 【図11】

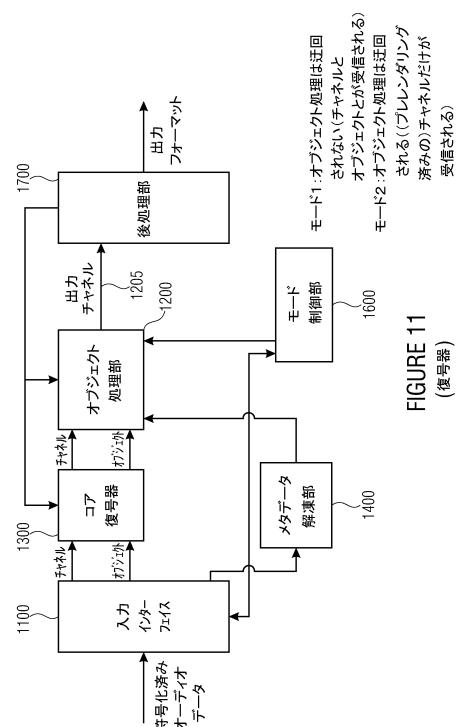


FIGURE 11 (復号器)

【図12】

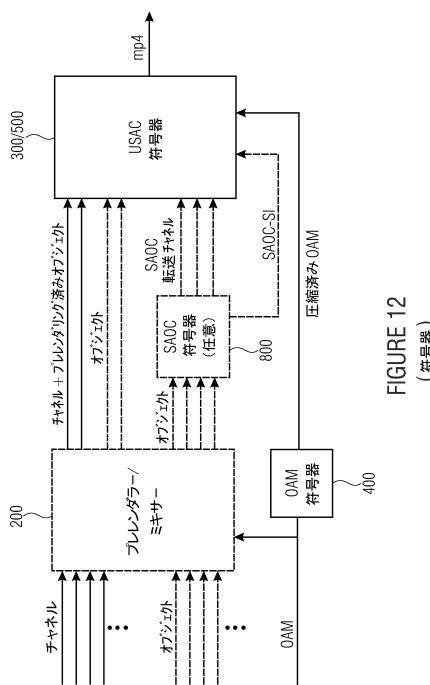


FIGURE 12  
(符號器)

【図13】

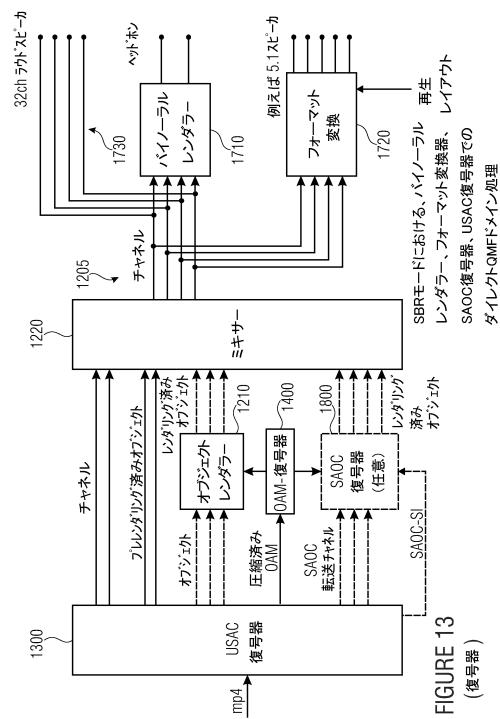
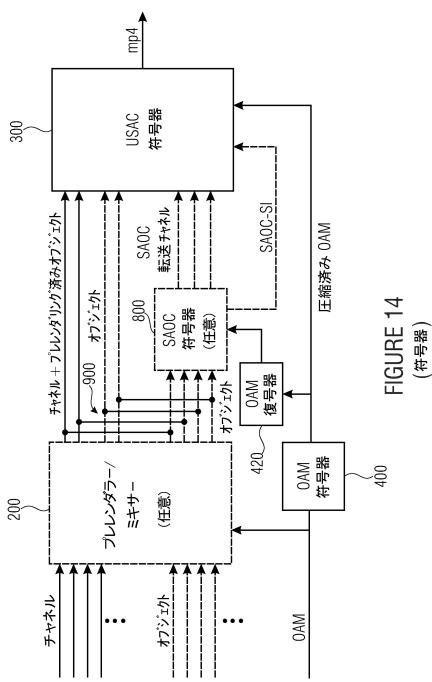
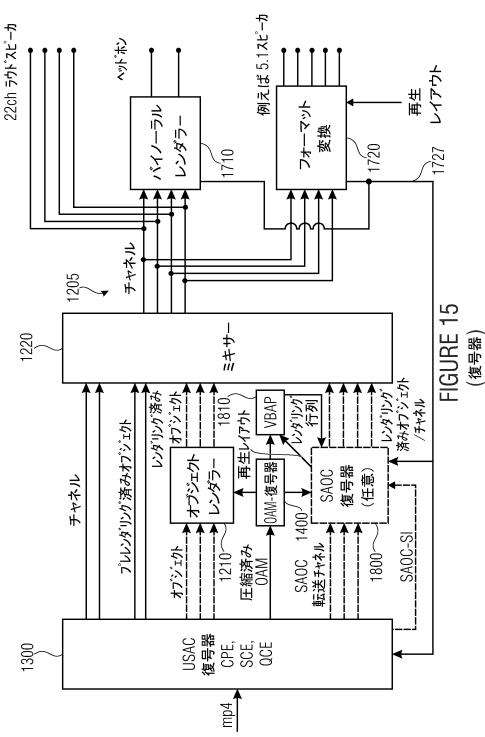


FIGURE 13  
(復号器)

【図 14】

FIGURE 14  
(符号器)

【図 15】

FIGURE 15  
(復号器)

---

フロントページの続き

(31)優先権主張番号 EP13177378  
(32)優先日 平成25年7月22日(2013.7.22)  
(33)優先権主張国 欧州特許庁(EP)  
(31)優先権主張番号 EP13189279  
(32)優先日 平成25年10月18日(2013.10.18)  
(33)優先権主張国 欧州特許庁(EP)

(72)発明者 エルテル, ク里斯チャン  
ドイツ連邦共和国 90542 エッケンタル ニュルンベルゲル ストラーーゼ 24  
(72)発明者 ヒルペルト, ヨハネス  
ドイツ連邦共和国 90411 ニュルンベルク ツィーゲルシュタインストラーーゼ 183

審査官 五貫 昭一

(56)参考文献 特表2014-520491(JP, A)  
特表2014-522155(JP, A)  
特許第5129888(JP, B2)

(58)調査した分野(Int.Cl., DB名)  
G10L 19/008  
G10L 19/00