

(19) 日本国特許庁(JP)

## (12) 公表特許公報(A)

(11) 特許出願公表番号

特表2016-528811  
(P2016-528811A)

(43) 公表日 平成28年9月15日(2016.9.15)

(51) Int.Cl.

H04S 5/02 (2006.01)  
G10L 19/008 (2013.01)

F 1

H04S 5/02  
G10L 19/008  
G10L 19/008

テーマコード(参考)

Z

5D162  
200  
100

審査請求 有 予備審査請求 有 (全 114 頁)

(21) 出願番号	特願2016-528443 (P2016-528443)
(86) (22) 出願日	平成26年7月17日 (2014.7.17)
(85) 翻訳文提出日	平成28年3月22日 (2016.3.22)
(86) 國際出願番号	PCT/EP2014/065397
(87) 國際公開番号	W02015/011015
(87) 國際公開日	平成27年1月29日 (2015.1.29)
(31) 優先権主張番号	13177374.9
(32) 優先日	平成25年7月22日 (2013.7.22)
(33) 優先権主張国	歐州特許庁 (EP)
(31) 優先権主張番号	13189345.5
(32) 優先日	平成25年10月18日 (2013.10.18)
(33) 優先権主張国	歐州特許庁 (EP)
(31) 優先権主張番号	14161611.0
(32) 優先日	平成26年3月25日 (2014.3.25)
(33) 優先権主張国	歐州特許庁 (EP)

(71) 出願人	591037214 フランホッファーーゲゼルシャフト ツ ア フエルダールング デア アンゲヴァ ンテン フォアシュンク エー. ファオ ドイツ連邦共和国 80686 ミュンヘ ン ハンザシュトラッセ 27ツエー
(74) 代理人	100079577 弁理士 岡田 全啓
(74) 代理人	100167966 弁理士 扇谷 一
(72) 発明者	ディスヒ サッシャ ドイツ連邦共和国 90766 フュルト ヴィルヘルムシュトラーセ 70

最終頁に続く

(54) 【発明の名称】マルチチャネル・オーディオ・デコーダ、マルチチャネル・オーディオ・エンコーダ、レンダリングされたオーディオ信号を使用する方法、コンピュータ・プログラムおよび符号化オーディオ

## (57) 【要約】

符号化表現に基づいて少なくとも 2 つの出力オーディオ信号を供給するためのマルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号を得るために、1 つ以上のレンダリング・パラメータに基づいて、符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするように構成される。マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号から 1 つ以上の非相関化オーディオ信号を導出するように構成され、出力オーディオ信号を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと 1 つ以上の非相関化オーディオ信号とを結合するように構成される。マルチチャネル・オーディオ・エンコーダは、オーディオ・デコーダを制御するための非相関化方法パラメータを供給する。

【選択図】図 1

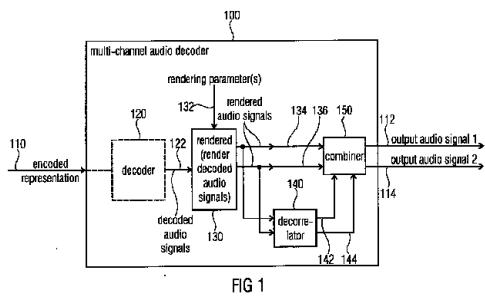


FIG 1

## 【特許請求の範囲】

## 【請求項 1】

符号化表現 (110 ; 710 ; 1516a ; 1516b ; 1518) に基づいて、少なくとも 2 つの出力オーディオ信号 (112, 114 ; 712, 714 ; 1552a - 1552n ; 3012) を供給するためのマルチチャネル・オーディオ・デコーダ (100 ; 700 ; 1550 ; 3000) であって、

ここで、前記マルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号 (134, 136 ; 1582a - 1582n,  $\hat{Z}$ ) を得るために、1 つ以上のレンダリング・パラメータ (132) に基づいて、前記符号化表現に基づいて得られた複数の復号化オーディオ信号 (122 ; 1562a - 1562n,  $\hat{X}$ ) をレンダリングする (130 ; 1580) ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号から 1 つ以上の非相関化オーディオ信号 (142, 144 ; 1592a - 1592n) を導出する (140 ; 1590) ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する (150 ; 1598) ように構成される、マルチチャネル・オーディオ・デコーダ。

## 【請求項 2】

前記マルチチャネル・オーディオ・デコーダは、パラメトリック再構成 (120 ; 1560) を使用して、前記複数のレンダリングされたオーディオ信号を得るためにレンダリングされた前記復号化されたオーディオ信号を得るように構成される、請求項 1 に記載のマルチチャネル・オーディオ・デコーダ。

## 【請求項 3】

前記復号化オーディオ信号は、再構成オブジェクト信号であり、

前記マルチチャネル・オーディオ・デコーダは、サイド情報 (1518) を使用して、1 つ以上のダウンミックス信号 (1516a, 1516b) から前記再構成オブジェクト信号を導出するように構成される、請求項 2 に記載のマルチチャネル・オーディオ・デコーダ。

## 【請求項 4】

前記マルチチャネル・オーディオ・デコーダは、前記サイド情報からアンミキシング係数を導出し、前記アンミキシング係数を使用して、前記 1 つ以上のダウンミックス信号から前記再構成オブジェクト信号を導出するための前記アンミキシング係数を適用するように構成される、請求項 3 に記載のマルチチャネル・オーディオ・デコーダ。

## 【請求項 5】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の所望の相関特性または共分散特性を少なくとも部分的に達成するように、前記レンダリングされたオーディオ信号と前記 1 つ以上の非相関化オーディオ信号とを結合するように構成される、請求項 1 ないし請求項 4 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

## 【請求項 6】

前記マルチチャネル・オーディオ・デコーダは、前記複数のレンダリングされたオーディオ信号を得るためにレンダリングされた前記復号化オーディオ信号 (122 ; 1562a - 1562n) のパラメトリック再構成 (120 ; 1560) の間のエネルギー損失のために少なくとも部分的に補償するように、前記レンダリングされたオーディオ信号と前記 1 つ以上の非相関化オーディオ信号とを結合するように構成される、請求項 1 ないし請求項 5 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

## 【請求項 7】

10

20

30

40

50

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の所望の相関特性または所望の共分散特性を決定するように構成され、

前記得られた出力オーディオ信号の相関特性または共分散特性が、前記所望の相関特性または所望の共分散特性（C）に近似するかまたは等しいように、前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号との結合（150；1598）を調整するように構成される、請求項1ないし請求項6のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項8】

前記マルチチャネル・オーディオ・デコーダは、前記複数のレンダリングされたオーディオ信号（Z）を得るために、前記符号化表現に基づいて得られる前記複数の復号化オーディオ信号（X）をレンダリングすること（130；1560）を記述しているレンダリング情報（R）に基づいて、前記所望の相関特性または所望の共分散特性（C）を決定するように構成される、請求項7に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項9】

前記マルチチャネル・オーディオ・デコーダは、複数のオーディオ・オブジェクトの特性および／または複数のオーディオ・オブジェクトの間ににおける関係を記述しているオブジェクト相関情報またはオブジェクト共分散情報（E<sub>x</sub>）に基づいて、前記所望の相関特性または所望の共分散特性（C）を決定するように構成される、請求項7または請求項8に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項10】

前記マルチチャネル・オーディオ・デコーダは、前記符号化表現に含まれるサイド情報（1518）に基づいて、前記オブジェクト相関情報またはオブジェクト共分散情報（E<sub>x</sub>）を決定するように構成される、請求項9に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項11】

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号および前記1つ以上の非相関化オーディオ信号の現実の相関特性または共分散特性（E<sub>s</sub>）を決定するように、および

前記レンダリングされたオーディオ信号および前記1つ以上の非相関化オーディオ信号の前記現実の相関特性または共分散特性（E<sub>s</sub>）に基づいて、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号との結合（150；1598）を調整するように構成される、請求項7ないし請求項10のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項12】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \hat{\mathbf{P}}\hat{\mathbf{Z}} + \mathbf{M}\mathbf{W}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\hat{\mathbf{z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

10

ここで、 $\mathbf{P}$ は、前記レンダリングされたオーディオ信号 $\hat{\mathbf{z}}$ に適用される混合行列であり、 $\mathbf{M}$ は、前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ に適用される混合行列である、請求項1ないし請求項1-1のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

20

#### 【請求項1-3】

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{\mathbf{Z}}$ の相関特性または共分散特性( $E_{\tilde{\mathbf{Z}}}$ )が、前記所望の相関特性または所望の共分散特性( $C$ )に近似するかまたは等しいように、前記混合行列 $\mathbf{P}$ および前記混合行列 $\mathbf{M}$ の少なくとも1つを調整するように構成される、請求項1-2に記載のマルチチャネル・オーディオ・デコーダ。

30

#### 【請求項1-4】

前記マルチチャネル・オーディオ・デコーダは、前記混合行列 $\mathbf{P}$ および前記混合行列 $\mathbf{M}$ と一緒に算出するように構成される、請求項1-2または請求項1-3に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項1-5】

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{\mathbf{Z}}$ の共分散行列 $E_{\tilde{\mathbf{Z}}}$ が所望の共分散行列 $C$ に近似するかまたは等しいように、

$$\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$$

で示される、結合混合行列 $\mathbf{F}$ を得るように構成される、請求項1-2ないし請求項1-4に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項1-6】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{E}_{\tilde{z}} = \mathbf{F} \mathbf{E}_s \mathbf{F}^H$$

で示される前記共分散行列が、

$$\mathbf{C} = \mathbf{R} \mathbf{E}_x \mathbf{R}^H$$

10

で示される前記所望の共分散行列と等しくなるように、前記結合混合行列 F を決定するよう構成され、

ここで、 $\mathbf{E}_s$  は、次式で定義される、前記レンダリングされたオーディオ信号  $\hat{\mathbf{Z}}$  と前記 1 つ以上の非相関化オーディオ信号  $\mathbf{W}$  とを結合した信号  $\mathbf{S}$  の共分散行列であり、

$$\mathbf{S} = \begin{bmatrix} \hat{\mathbf{Z}} \\ \mathbf{W} \end{bmatrix}$$

20

ここで、 $\mathbf{E}_x$  は、オブジェクト共分散行列である、請求項 1 5 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 1 7】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{A}_{dry} P \hat{\mathbf{Z}} + \mathbf{M} \mathbf{W}$$

または、

$$\tilde{\mathbf{Z}} = P \hat{\mathbf{Z}} + \mathbf{A}_{wet} \mathbf{M} \mathbf{W}$$

30

または、

$$\tilde{\mathbf{Z}} = \mathbf{A}_{dry} P \hat{\mathbf{Z}} + \mathbf{A}_{wet} \mathbf{M} \mathbf{W}$$

で示される前記出力オーディオ信号  $\tilde{\mathbf{Z}}$  を得るために、前記レンダリングされたオーディ

40

オ信号  $\hat{\mathbf{Z}}$  と前記 1 つ以上の非相関化オーディオ信号  $\mathbf{W}$  とを結合するように構成され、

ここで、 $P$  は、前記レンダリングされたオーディオ信号  $\hat{\mathbf{Z}}$  に適応される混合行列であり、

$\mathbf{M}$  は、前記 1 つ以上の非相関化オーディオ信号  $\mathbf{W}$  に適用される混合行列であり、

$\mathbf{A}_{dry}$  は、第 1 の相関マトリックスまたは第 1 の調整マトリックスであり、

$\mathbf{A}_{wet}$  は、第 2 の相関マトリックスまたは第 2 の調整マトリックスである、請求項 1 ないし請求項 1 1 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 1 8】

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{Z}$

または $\tilde{Z}$ とPおよびMを用いるWとのミキシングによって得られたオーディオ信号の相関

特性または共分散特性( $E_{\tilde{Z}}$ )が、前記所望の相関特性または所望の共分散特性(C)に近似するかまたは等しいように、前記混合行列Pおよび前記混合行列Mの少なくとも1つを調整するように構成される、請求項17に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項19】**

前記マルチチャネル・オーディオ・デコーダは、前記混合行列Pおよび前記混合行列Mを一緒に算出するように構成される、請求項17または請求項18に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項20】**

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{Z}$ の共分散行列E<sub>Z</sub>または $\tilde{Z}$ とPおよびMを用いるWとのミキシングによって得られたオーディオ信号の共分散行列が、所望の共分散行列Cに近似するかまたは等しいように、

10

20

$$\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$$

で示される、結合混合行列Fを決定するように構成される、請求項17ないし請求項19に記載のマルチチャンネル・オーディオ・デコーダ。

**【請求項21】**

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{E}_{\tilde{Z}} = \mathbf{F} \mathbf{E}_s \mathbf{F}^H$$

30

で示される前記共分散行列が、

$$\mathbf{C} = \mathbf{R} \mathbf{E}_x \mathbf{R}^H$$

で示される前記所望の共分散行列と等しくなるように、前記結合混合行列Fを決定するように構成され、

ここで、E<sub>s</sub>は、次式で定義される、前記レンダリングされたオーディオ信号 $\tilde{Z}$ と前記1つ以上の非相関化オーディオ信号Wとを結合した信号Sの共分散行列であり、

$$\mathbf{S} = \begin{bmatrix} \hat{\mathbf{Z}} \\ \mathbf{W} \end{bmatrix}$$

ここで、E<sub>x</sub>は、オブジェクト共分散行列である、請求項20に記載のマルチチャンネル・オーディオ・デコーダ。

**【請求項22】**

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記レンダ

40

50

リングされたオーディオ信号の貢献が制限されるように、前記第1の相関マトリックスを決定するように構成され、および／または

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記非相関化オーディオ信号の貢献が制限されるように、前記第2の相関マトリックスを決定するように構成される、請求項17ないし請求項21のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項23】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記レンダリングされたオーディオ信号の貢献が制限されるように、前記レンダリングされたオーディオ信号の特性に基づいて、および／または、前記非相関化オーディオ信号の特性に基づいて、および／または、所望の出力オーディオ信号の特性に基づいて、および／または、混合されレンダリングされたオーディオ信号の推定された特性に基づいて、および／または、混合された非相関化オーディオ信号の推定された特性に基づいて、前記第1の相関マトリックスを決定するように構成され、および／または

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記非相関化オーディオ信号の貢献が制限されるように、前記レンダリングされたオーディオ信号の特性に基づいて、および／または、前記非相関化オーディオ信号の特性に基づいて、および／または、所望の出力オーディオ信号の特性に基づいて、および／または、混合されレンダリングされたオーディオ信号の推定された特性に基づいて、および／または、混合された非相関化オーディオ信号の推定された特性に基づいて、前記第2の相関マトリックスを決定するように構成される、請求項17ないし請求項22のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項24】

前記レンダリングされたオーディオ信号、および／または、前記非相関化オーディオ信号、および／または、前記所望の出力オーディオ信号、および／または、前記混合されレンダリングされたオーディオ信号、および／または、前記混合された非相関化オーディオ信号の前記特性は、エネルギー特性、または、相関特性、または、共分散特性である、請求項23に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項25】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{P}\widehat{\mathbf{Z}} + \mathbf{A}_{\text{wet}} \mathbf{M} \mathbf{W}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\widehat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

ここで、 $\mathbf{A}_{\text{wet}}$ は対角行列であり、 $i$ 番目の出力オーディオ信号において、混合行列 $\mathbf{M}$ を伴って、レンダリングされたオーディオ信号の強度 ( $E_{\mathbf{Y}}^{\text{dry}}(i,i)$ ) と混合された非相関化オーディオ信号の強度 ( $E_{\mathbf{Y}}^{\text{wet}}(i,i)$ ) との比が、閾値よりも小さい場合、標準、すなわち、非減少の対角入力と比較したとき、相関マトリックス $\mathbf{A}_{\text{wet}}$ の入力 $\mathbf{A}_{\text{wet}}(i, i)$ が減少するように、前記相関マトリックス $\mathbf{A}_{\text{wet}}$ を供給するように構成される、請求項1ないし請求項24のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項26】

前記閾値は、予め定められた固定閾値であるか、または、前記閾値は、たとえば、エネルギー特性、相関特性、および／または共分散特性のような信号特性に基づく時間変数、

10

20

30

40

50

および／または周波数変数である、請求項 2 5 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 2 7】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{P}\hat{\mathbf{Z}} + \mathbf{A}_{wet}\mathbf{M}\mathbf{W}$$

で示される前記出力オーディオ信号  $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ 10

信号  $\hat{\mathbf{Z}}$ と前記 1 つ以上の非相関化オーディオ信号  $\mathbf{W}$ とを結合するように構成され、

ここで、 $\mathbf{P} = \mathbf{P}_{dry}$ であり、

$\mathbf{M} = \mathbf{P}_{wet}$ であり、

$$\mathbf{A}_{wet} = \text{matdiag} \left( \min \left( 1, \sqrt{\max \left( 0, \lambda_{Dec} \frac{\mathbf{E}_Y^{dry}(i,i)}{\hat{\mathbf{E}}_Y^{wet}(i,i)} \right)} \right) \right) \text{であり、}$$

$\mathbf{E}_Y^{dry}$ は、前記レンダリングされたオーディオ信号  $\hat{\mathbf{Z}}$ の共分散行列であり、

20

$\hat{\mathbf{E}}_Y^{wet}$ は、前記マトリックス  $\mathbf{P}_{wet}$ が適用された後、前記非相関化オーディオ信号の推定された共分散行列である、請求項 1 ないし請求項 2 6 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

【請求項 2 8】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{F} = (\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)\mathbf{H}(\mathbf{V}\sqrt{\mathbf{Q}^{-1}}\mathbf{V}^H)$$

で示される前記結合混合行列  $\mathbf{F}$  を決定するように構成され、

ここで、マトリックス  $\mathbf{U}$ 、 $\mathbf{T}$ 、 $\mathbf{V}$  および  $\mathbf{Q}$  は、前記共分散行列の特異値分解を使用して決定され、 $\mathbf{E}_s$  および  $\mathbf{C}$  は、

$$\mathbf{C} = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

10

および

$$\mathbf{E}_s = \mathbf{V}\mathbf{Q}\mathbf{V}^H$$

で得られ、

マトリックス  $\mathbf{H}$  は、

$$\mathbf{H} = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 & b_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 & 0 & b_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{N,N} & 0 & 0 & \dots & b_{N,N} \end{bmatrix}$$

20

として定義され、

$a_{i,i}$  および  $b_{i,i}$  は、

$$a_{i,i}^2 + b_{i,i}^2 = 1$$

30

で示すように選択される、請求項 17 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 29】

前記マルチチャネル・オーディオ・デコーダは、単位行列であるか、またはその倍数である混合行列  $\mathbf{P}$  を設定し、そして、前記混合行列  $\mathbf{M}$  を算出するように構成される、請求項 12 または請求項 13 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 30】

前記マルチチャネル・オーディオ・デコーダは、

$$\Delta_E = \mathbf{C} - \mathbf{E}_{\hat{z}}$$

として定義される、前記所望の共分散行列  $\mathbf{C}$  と共分散行列  $\mathbf{E}_{\hat{z}}$  との差  $\Delta_E$  が、共分散である

$$\mathbf{M}\mathbf{E}_w\mathbf{M}^H$$

10

に等しいか、あるいは近似であるように、前記混合行列  $\mathbf{M}$  を決定するように構成され、  
ここで、前記所望の共分散行列  $\mathbf{C}$  は、

$$\mathbf{C} = \mathbf{R}\mathbf{E}_x\mathbf{R}^H$$

として定義され、

$\mathbf{R}$  は、レンダリング行列であり、

$\mathbf{E}_x$  は、オブジェクト共分散行列であり、

$\mathbf{E}_w$  は、前記 1 つ以上の非相関化信号の共分散行列であり、

20

$\mathbf{E}_{\hat{z}}$  は、前記レンダリングされたオーディオ信号の共分散行列である、請求項 2 9 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 1】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{M} = (\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)(\mathbf{V}\sqrt{\mathbf{Q}^{-1}}\mathbf{V}^H)$$

30

で示される前記混合行列  $\mathbf{M}$  を決定するように構成され、

ここで、マトリックス  $\mathbf{U}$ 、 $\mathbf{T}$ 、 $\mathbf{V}$  および  $\mathbf{Q}$  は、前記共分散行列の特異値分解を使用して決定され、 $\Delta_E$  および  $\mathbf{E}_w$  は、

$$\Delta_E = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

および

$$\mathbf{E}_w = \mathbf{V}\mathbf{Q}\mathbf{V}^H$$

40

で得られる、請求項 3 0 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 2】

前記マルチチャネル・オーディオ・デコーダは、所与のレンダリングされたオーディオ信号が、前記所与のレンダリングされたオーディオ信号自身の非相関化バージョンとともに混合されるのみという制約のもとに、前記混合行列  $\mathbf{P}$ 、 $\mathbf{M}$  を決定するように構成される、請求項 1 2 または請求項 1 3 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 3】

相互相関値または相互共分散値が不变である間、レンダリングされたオーディオ信号の自己相関値のみまたは自己共分散値のみが修正されるように、前記マルチチャネル・オーディオ・デコーダは、

50

ディオ・デコーダは、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号とを結合するように構成される、請求項12または請求項13または請求項32に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項34】**

前記マルチチャネル・オーディオ・デコーダは、単位行列であるか、またはその倍数である前記混合行列Pを設定し、そして、Mが対角行列であるという制限のもと、前記混合行列Mを算出するように構成される、請求項12または請求項13または請求項32または請求項33に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項35】**

前記マルチチャネル・オーディオ・デコーダは、

10

$$\tilde{\mathbf{Z}} = \hat{\mathbf{Z}} + \mathbf{MW}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号Wとを結合するように構成され、

ここで、Mは、前記1つ以上の非相関化オーディオ信号Wに適用される対角混合行列であり、そして、

20

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の共分散行列の対角線要素が所望のエネルギーと等しくなるように、前記混合行列Mの対角線要素を算出するように構成される、請求項32または請求項33または請求項34に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項36】**

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{M}(i,j) = \begin{cases} \sqrt{\min\left(\lambda_{Dec}, \max\left(0, \frac{\mathbf{C}(i,i) - \mathbf{E}_{\hat{Z}}(i,i)}{\max(\mathbf{E}_W(i,i), \varepsilon)}\right)\right)} & i = j, \\ 0 & i \neq j. \end{cases}$$

30

で示される前記混合行列Mの前記要素を算出するように構成され、

ここで、前記所望の共分散行列Cは、

$$\mathbf{C} = \mathbf{R}\mathbf{E}_x\mathbf{R}^H$$

として定義され、

Rは、レンダリング行列であり、

40

$\mathbf{E}_x$ は、オブジェクト共分散行列であり、

$\mathbf{E}_w$ は、前記1つ以上の非相関化信号の共分散行列であり、そして、

$\lambda_{Dec}$ は、前記信号に加えられる非相関の量を制限している閾値である、請求項35に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項37】**

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合するための方法を決定する場合、前記マルチチャネル・オーディオ・デコーダは、前記非相関化オーディオ信号の相関特性または共分散特性を考慮するように構成される、請求項1ないし請求項36のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

50

**【請求項 3 8】**

所与の出力オーディオ信号が 2 つ以上のレンダリングされたオーディオ信号および少なくとも 1 つの非相関化オーディオ信号に基づいて供給されるように、前記マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号と非相関化オーディオ信号とを混合するように構成される、請求項 1 ないし請求項 2 8 、または請求項 3 7 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 3 9】**

前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合するための方法を決定するために、異なる制約が適用される場合において、前記マルチチャネル・オーディオ・デコーダは、異なるモードの間で切り替えるように構成される、請求項 1 ないし請求項 3 8 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。  
10

**【請求項 4 0】**

前記マルチチャネル・オーディオ・デコーダは、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第 1 のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第 2 のモード、および  
20

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第 3 のモード、  
の間で切り替えるように構成される、請求項 1 ないし請求項 3 9 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。  
30

**【請求項 4 1】**

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合するための前記 3 つのモードが使用されるかについて指し示す前記符号化表現のビットストリーム要素を評価し、そして、前記ビットストリーム要素に基づいて前記モードを選択するように構成される、請求項 3 9 または請求項 4 0 に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 4 2】**

少なくとも 2 つの入力オーディオ信号 ( 2 1 0 , 2 1 2 ; 1 5 1 2 a - 1 5 1 2 n ; 2 9 1 2 , 2 9 1 4 ) に基づいて符号化表現 ( 2 1 4 ; 1 5 1 6 a , 1 5 1 6 b , 1 5 1 8 ; 2 9 3 2 ) を供給するためのマルチチャネル・オーディオ・エンコーダ ( 2 0 0 ; 1 5 1 0 ; 2 9 0 0 ) であって、  
40

ここで、前記マルチチャネル・オーディオ・エンコーダは、前記少なくとも 2 つの入力オーディオ信号に基づき、1 つ以上のダウンミックス信号 ( 2 2 2 ; 1 5 1 6 a , 1 5 1 6 b ) を供給する ( 2 2 0 ) ように構成され、

前記マルチチャネル・オーディオ・エンコーダは、前記少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータ ( 2 3 2 ; 1 5 1 8 ) を供給する ( 2 3 0 ) ように構成され、そして、

前記マルチチャネル・オーディオ・エンコーダは、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラ  
50

メータ(242；1518)を供給する(240)ように構成される、マルチチャネル・オーディオ・エンコーダ。

#### 【請求項43】

前記マルチチャネル・オーディオ・エンコーダは、前記非相関化方法パラメータを選択的に供給し、オーディオ・デコーダの処理のための、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、

である以上に示す3つのモードの1つの信号を送信するように構成される、請求項42に記載のマルチチャネル・オーディオ・エンコーダ。

#### 【請求項44】

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号が比較的高い相関または比較的低い相関を含むかどうかに基づいて、前記非相関化方法パラメータを選択するように構成される、請求項42または請求項43に記載のマルチチャネル・オーディオ・エンコーダ。

#### 【請求項45】

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号の間の相関が比較的高い場合、前記第1のモードまたは前記第2のモードを指定するための前記非相関化方法パラメータを選択するように構成され、

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号の間の相関が比較的低い場合、前記第3のモードを指定するための前記非相関化方法パラメータを選択するように構成される、請求項43に記載のマルチチャネル・オーディオ・エンコーダ。

#### 【請求項46】

符号化表現に基づく少なくとも2つの出力オーディオ信号を供給するための方法(300)であって、前記方法は、

複数のレンダリングされたオーディオ信号を得るために、1つ以上のレンダリング・パラメータに基づいて、前記符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするステップ(310)と、

前記レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するステップ(320)と、

前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合するステップ(330)と、を含む、方法。

#### 【請求項47】

少なくとも2つの入力オーディオ信号に基づく符号化表現を供給するための方法(400)であって、前記方法は、

前記少なくとも2つの入力オーディオ信号に基づく1つ以上のダウンミックス信号を供給するステップ(410)と、

10

20

30

40

50

前記少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータを供給するステップ(420)と、

オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータを供給するステップ(430)と、を含む、方法。

#### 【請求項 4 8】

コンピュータ・プログラムがコンピュータ上で実行されると、前記コンピュータが請求項 4 6、または請求項 4 7 に記載の方法を実行する、コンピュータ・プログラム。

#### 【請求項 4 9】

符号化オーディオ表現(500)は、

ダウンミックス信号の符号化表現(510)、

前記少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータの符号化表現(520)、および

オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している符号化非相関化方法パラメータ(530)を含む、符号化オーディオ表現。

#### 【発明の詳細な説明】

##### 【技術分野】

##### 【0 0 0 1】

本発明による実施の形態は、符号化表現に基づいて、少なくとも 2 つの出力オーディオ信号を供給するためのマルチチャネル・オーディオ・デコーダに関する。

##### 【0 0 0 2】

本発明による更なる実施の形態は、少なくとも 2 つの入力オーディオ信号に基づいて、符号化表現を供給するためのマルチチャネル・オーディオ・エンコーダに関する。

##### 【0 0 0 3】

本発明による更なる実施の形態は、符号化表現に基づいて、少なくとも 2 つの出力オーディオ信号を供給するための方法に関する。

##### 【0 0 0 4】

本発明による更なる実施の形態は、少なくとも 2 つの入力オーディオ信号に基づいて、符号化表現を供給するための方法に関する。

##### 【0 0 0 5】

本発明による更なる実施の形態は、前記方法のうちの 1 つを実行するためのコンピュータ・プログラムに関する。

##### 【0 0 0 6】

本発明による更なる実施の形態は、符号化オーディオ表現に関する。

##### 【0 0 0 7】

一般的に言って、本発明による実施の形態は、マルチチャネル・ダウンミックス / アップミックスパラメトリックオーディオ・オブジェクト符号化システムのための非相関化コンセプトに関する。

##### 【背景技術】

##### 【0 0 0 8】

近年、オーディオ・コンテンツの記憶および送信の需要は、着実に増加した。さらに、オーディオ・コンテンツの記憶および送信のための良質な要件も、着実に増加した。従って、オーディオ・コンテンツの符号化および復号化のためのコンセプトは、強化されている。

##### 【0 0 0 9】

例えば、いわゆる、開発された「先進的音響符号化(Advanced Audio Coding(AAC))」は、国際標準 ISO / IEC 13818 - 7 : 2003において記述されている。さらに、例えば、国際標準 ISO / IEC 23003 - 1 : 2007において、例えば、記述された、いわゆる「MPEGサラウンド」のように、いく

10

20

30

40

50

つかの空間的な拡張が作成された。さらに、オーディオ信号の空間的な情報の符号化および復号化のためのさらなる改良が、いわゆる「空間オーディオ・オブジェクト符号化（Spatia l Audio Object Coding）」に関する国際標準ISO/IEC 23003-2:2010において記述される。

#### 【0010】

さらに、良好な符号化効率を有する一般のオーディオ信号およびスピーチ信号の両方の符号化して、マルチチャネル・オーディオ信号を扱うという可能性を提供する切り替え可能なオーディオ符号化／復号化のコンセプトは、いわゆる「統一のスピーチおよびオーディオ符号化（Unified Speech and Audio Object Coding）」に関する国際標準ISO/IEC 23003-3:2012において記述される。10

#### 【0011】

さらに、更なる従来の概念が、本記述の最後において言及される参考文献において記述される。

#### 【先行技術文献】

#### 【特許文献】

#### 【0012】

#### 【特許文献1】国際公開第2006/026452号

#### 【非特許文献】

#### 【0013】

【非特許文献1】[BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.20

【非特許文献2】[Blauert] J. Blauert, "Spatial Hearing - The Psychophysics of Human Sound Localization", Revised Edition, The MIT Press, London, 1997.

【非特許文献3】[JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006.

【非特許文献4】[ISS1] M. Parvaix and L. Girin: "Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding", IEEE ICASSP, 2010.30

【非特許文献5】[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: "A watermarking-based method for informed source separation of audio signals with a single sensor", IEEE Transactions on Audio, Speech and Language Processing, 2010.

【非特許文献6】[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: "Informed source separation through spectrogram coding and data embedding", Signal Processing Journal, 2011.

【非特許文献7】[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: "Informed source separation: source coding meets source separation", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011.

【非特許文献8】[ISS5] S. Zhang and L. Girin: "An Informed Source Separation System for Speech Signals", INTERSPEECH, 2011.40

【非特許文献9】[ISS6] L. Girin and J. Pinel: "Informed Audio Source Separation from Compressed Linear Stereo Mixtures", AES 42nd International Conference : Semantic Audio, 2011.

【非特許文献10】[MPS] ISO/IEC, "Information technology - MPEG audio technologies - Part 1: MPEG Surround," ISO/IEC JTC1/SC29/WG11 (MPEG) international Standard 23003-1:2006.

【非特許文献11】[OCD] J. Vilkamo, T. Baeckstroem, and A. Kuntz. "Optimized covariance domain framework for time-frequency processing of spatial audio", Journal of the Audio Engineering Society, 2013. in press.50

【非特許文献12】[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007.

【非特許文献13】[SAOC2] J. Engdegard, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hoelzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: "Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008.

【非特許文献14】[SAOC] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2.

10

#### 【発明の概要】

#### 【発明が解決しようとする課題】

#### 【0014】

しかしながら、3次元オーディオ場面の効果的な符号化および復号化のためのさらなる進歩的な概念を提供することが切望されている。

#### 【課題を解決するための手段】

#### 【0015】

本発明による実施の形態は、符号化表現に基づいて少なくとも2つの出力オーディオ信号を供給するためのマルチチャネル・オーディオ・デコーダをもたらす。マルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号を得るために、1つ以上のレンダリング・パラメータに基づいて、符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするように構成される。マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するように構成される。さらに、マルチチャネル・オーディオ・デコーダは、出力オーディオ信号を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号を結合するように構成される。

20

#### 【0016】

本発明による実施の形態は、オーディオ品質が、複数の復号化オーディオ信号に基づいて得られるレンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出することによって、および、出力オーディオ信号を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合することによって、マルチチャネル・オーディオ・デコーダにおいて改良されうるという知見に基づく。レンダリングの前、またはレンダリングの間に追加の非相関化信号と比較した場合、レンダリングの後、追加の非相関化信号によって出力オーディオ信号の相關特性または共分散特性を調整することがより効率的であることが分かっている。非相関化が、レンダリングの前またはレンダリングの間に実行された場合、より非相関器が必要とされるので、レンダリングされたオーディオ信号よりもレンダリングに入力される復号化オーディオ信号がある一般的のケースにおいて、このコンセプトはより効率的であることが分かっている。さらに、レンダリングは、典型的には、復号化オーディオ信号の結合をもたらすので、非相関化信号は、レンダリングの前に復号化オーディオ信号を追加する場合に、アーティファクトが供給されることが分かっている。従って、本発明の本実施の形態による概念は、非相関化信号が、レンダリングの前に追加されるという点で、従来のアプローチより優れている。例えば、レンダリングされた信号の所望の相關特性または共分散特性を直接的に推定し、実際にレンダリングされた信号に非相関化オーディオ信号の供給を適応させることが可能である。そして、それは、効率とオーディオ品質との間のより良好なトレードオフを結果として得て、同時に、しばしば増加した効率およびより良好な品質を結果として得る。

30

#### 【0017】

好みしい実施の形態において、マルチチャネル・オーディオ・デコーダは、パラメトリック再構成を使用して複数のレンダリングされたオーディオ信号を得るためにレンダリン

40

50

グされる復号化オーディオ信号を得るように構成される。本発明による概念は、オーディオ信号のパラメトリック再構成と結合して効果をもたらすことが分かっており、ここで、例えば、パラメトリック再構成は、オブジェクト信号を記述しているサイド情報および/またはオブジェクト信号の間の関係に基づいている（オブジェクト信号は、復号化オーディオ信号を構成しうる。）。例えば、この種のコンセプトにおける比較的かなりの数のオブジェクト信号（復号化オーディオ信号）があり、レンダリングされたオーディオ信号に基づいて非相関化のアプリケーションが特に効率的であり、この種のシナリオにおいて、アーティファクトを回避することが分かっている。

#### 【0018】

好ましい実施の形態において、復号化オーディオ信号は、再構成オブジェクト信号（例えば、パラメータ的に再構成されたオブジェクト信号）であり、そして、マルチチャネル・オーディオ・デコーダは、サイド情報を使用して1つ以上のダウンミックス信号から再構成オブジェクト信号を導出するように構成される。従って、レンダリングされたオーディオ信号または出力オーディオ信号の数よりも多い比較的かなりの数の再構成オブジェクト信号がある場合であっても、レンダリングされたオーディオ信号に基づくレンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号との結合は、出力オーディオ信号における相関特性または共分散特性の効果的な再構成を許容する。

#### 【0019】

好ましい実施の形態において、マルチチャネル・オーディオ・デコーダは、サイド情報をからアンミキシング係数を導出し、アンミキシング係数を使用して1つ以上のダウンミックス信号から（パラメータ的に）再構成オブジェクト信号を導出するために、アンミキシング係数を適用するように構成される。従って、レンダリングのための入力信号は、オブジェクトに関連したサイド情報（例えば、オブジェクト間の相関情報またはオブジェクト・レベル差情報のように、その結果が、絶対のエネルギーを使用することによって得られる）であるサイド情報から導出される。

#### 【0020】

好ましい実施の形態において、出力オーディオ信号の所望の相関特性または共分散特性を少なくとも部分的に獲得するために、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合するように構成されうる。レンダリングされたオーディオ信号から導出される、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号との結合は、所望の相関特性または共分散特性の調整（または再構成）を許容することが分かっている。さらに、聴覚器官印象に対して、出力オーディオ信号における適当な相関特性または共分散特性を有することが重要であり、そして、これが非相関化オーディオ信号を使用してレンダリングされたオーディオ信号を修正することによって最も達成されうることが分かっている。例えば、レンダリングされたオーディオ信号とレンダリングされたオーディオ信号に基づく非相関化オーディオ信号とを結合する場合、以前の処理ステージにおいて生じる如何なる低下も考慮されうる。

#### 【0021】

好ましい実施の形態において、マルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号を得るためにレンダリングされる復号化オーディオ信号のパラメトリック再構成の間ににおけるエネルギー損失を少なくとも部分的に補償するために、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合するように構成される。非相関化オーディオ信号のポストレンダリングアプリケーションは、例えば、復号化オーディオ信号のパラメトリック再構成によってレンダリング前の処理によって生じる信号欠損を修正することを許容することが分かっている。従って、それは高い精度を伴って、レンダリングに入力される復号化オーディオ信号の相関特性または共分散特性を再構成する必要がない。これは、復号化オーディオ信号の再構成を単純化し、そして、したがって、高効率をもたらす。

#### 【0022】

好ましい実施の形態において、マルチチャネル・オーディオ・デコーダは、出力オーデ

10

20

30

40

50

イオ信号の所望の相関特性または共分散特性を決定するように構成される。さらに、マルチチャネル・オーディオ・デコーダは、得られた出力オーディオ信号の相関特性または共分散特性が所望の相関特性または共分散特性に近似するか、または等しいように、出力オーディオ信号を得るために、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合を調整するように構成される。レンダリングされたオーディオ信号と非相関化オーディオ信号との結合の後に到達される出力オーディオ信号の所望の相関特性または共分散特性を算出（または決定）することによって、次々に比較的正確な再構成を許容する、処理の遅いステージで相関特性または共分散特性を調整することは可能である。従って、出力オーディオ信号の空間的な聴取印象は、所望の聴取印象によく適応される。

10

#### 【0023】

好みの実施の形態において、複数のレンダリングされたオーディオ信号を得るために、符号化表現に基づいて得られる複数の復号化オーディオ信号のレンダリングを記述しているレンダリング情報に基づいて、所望の相関特性または所望の共分散特性を決定するように構成される。所望の相関特性または所望の共分散特性の決定におけるレンダリング処理を考慮することによって、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号との結合を調整するための正確な情報を獲得することは可能である。そして、それは、所望の聴取印象にマッチする出力オーディオ信号を有する可能性をもたらす。

#### 【0024】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、複数のオーディオ・オブジェクトの特性および／または複数のオーディオ・オブジェクトの間の関係を記載しているオブジェクト相関情報またはオブジェクト共分散情報に基づいて、所望の相関特性または所望の共分散特性を決定するように構成される。従って、処理の後、すなわち、レンダリングの後、オーディオ・オブジェクトに適応される相関特性または共分散特性を復元することが可能である。従って、オーディオ・オブジェクトを復号化するための複雑さが低減される。さらに、レンダリングの後のオーディオ・オブジェクトの相関特性または共分散特性を考慮することによって、レンダリングの有害な影響が回避され、相関特性または共分散特性は、良好な精度で再構成される。

20

#### 【0025】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、符号化表現に含まれるサイド情報に基づいて、オブジェクト相関情報またはオブジェクト共分散情報を決定するように構成される。従って、コンセプトは、サイド情報を使用する空間的なオーディオ・オブジェクト符号化アプローチに良好に適している。

30

#### 【0026】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号の現実の相関特性または共分散特性に基づいて、出力オーディオ信号を得るために、レンダリングされたオーディオ信号の現実の相関特性または共分散特性を決定し、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号との結合を調整するように構成される。従って、オーディオ・オブジェクトを再構成するか、またはレンダリングによって欠損が生じた場合、例えば、エネルギー損失のような初期の処理ステージにおいて欠損が渡されうる。このように、現実のレンダリングされたオーディオ信号と非相関化オーディオ信号との結合が所望の特性を結果として得られるように、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号との結合は、ニーズに非常に正確な方法で調整されうる。

40

#### 【0027】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合するように構成され、ここで、レンダリングされたオーディオ信号は、第1の混合行列Pを使用して重み付けされ、1つ以上の非相関化オーディオ信号は、第2の混合行列Mを使用して重み付けされる。これは、レンダリングされたオーディオ信号に適用される混合行列Pによって記述

50

され、1つ以上の非相関化オーディオ信号に適用される混合行列Mによって記述される一次結合が実行される。

【0028】

好みの実施の形態において、得られた出力オーディオ信号の相関特性または共分散特性が所望の相関特性または所望の共分散特性と近似するかまたは等しくなるように、マルチチャネル・オーディオ・デコーダは、混合行列Pおよび混合行列Mの少なくとも1つを調整するように構成される。このように、適度な効率および良好な結果を典型的に可能である1つ以上の混合行列を調整する方法がある。

【0029】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、混合行列Pおよび混合行列Mと一緒に算出するように構成される。従って、得られた出力オーディオ信号の相関特性または共分散特性が所望の相関特性または所望の共分散特性と近似するかまたは等しくなりうるように混合行列を得ることができる。さらに、混合行列Pおよび混合行列Mと一緒に算出した場合、混合行列Pと混合行列Mを前提条件に適合されることが可能であるように、いくつかの自由度は典型的に利用できる。

10

【0030】

好みの実施の形態において、得られた出力オーディオ信号の共分散行列が所望の共分散行列に等しくなるように、マルチチャネル・オーディオ・デコーダは、混合行列Pおよび混合行列Mを含む結合混合行列Fを得るために構成される。

20

【0031】

好みの実施の形態において、結合混合行列は、以下に示される式に従って算出される。

【0032】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号および非相関化オーディオ信号を記述する第1の共分散行列と、出力オーディオ信号の所望の共分散特性を記述する第2の共分散行列との特異値分解を使用して決定される行列を使用して、結合混合行列Fを決定するように構成される。この種の特異値分解を使用することが、結合混合行列を決定する数値的な解決案を構成する。

30

【0033】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、単位行列またはその倍数である混合行列Pを設定し、混合行列Mを算出するように構成される。これは、所望の空間印象を保存するのを助ける異なるレンダリングされたオーディオ信号の混合を回避する。さらに、自由度の数が低減される。

40

【0034】

好みの実施の形態において、所望の共分散行列とレンダリングされたオーディオ信号の共分散行列との間の差が、混合行列Mを混合した後、1つ以上の非相関化信号の共分散と近似するかまたは等しくなるように、マルチチャネル・オーディオ・デコーダは、混合行列Mを決定するように構成される。このように、混合行列Mを得るための計算的にシンプルなコンセプトが与えられる。

【0035】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、所望の共分散行列とレンダリングされたオーディオ信号の共分散行列との差と、1つ以上の非相関化信号の共分散行列との特異値分解を使用して決定される行列を使用して混合行列Mを決定するように構成される。これは、混合行列Mを決定するために計算的に非常に効率的な方法である。

40

【0036】

好みの実施の形態において、所与のレンダリングされたオーディオ信号は、所与のレンダリングされたオーディオ信号自身の非相関化バージョンを混合されるだけの制約のもと、マルチチャネル・オーディオ・デコーダは、混合行列P, Mを決定するように構成さ

50

れる。このコンセプトは、小さい改善（例えば、不完全な非相関化がある場合）に制限し、または相互相関特性もしくは相互共分散特性（例えば、理想的な非相関化がある場合）を防ぎ、そして、従って、認められたオブジェクトポジションの変化を回避するために、場合によっては、価値がある。しかしながら、非理想的な非相関化がある場合には、自己相関値（または自己共分散値）が明確に修正され、そして、交差項における変化は無視される。

#### 【0037】

好みの実施の形態において、相互相関特性または相互共分散特性が小さい値（例えば、不完全な非相関化がある場合）によって変更されないままかまたは修正されるとともに、レンダリングされたオーディオ信号の自己相関値または自己共分散値のみが修正されるように、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合するように構成される。また、オーディオ・オブジェクトの認められた位置の劣化が回避されうる。さらに、複雑な計算量が低減されうる。しかしながら、例えば、自己共分散値がエネルギー（自己相関値）の修正の結果として修正される一方、相互相関値は、修正されないままである（それらは、相互共分散値の規格化されたバージョンを表わす）。

#### 【0038】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、単位行列またはその倍数である混合行列Pを設定し、Mが対角行列であるという制約のもと、混合行列Mを算出するように構成される。このように、相互相関特性または相互共分散特性の修正は、回避されうるか、または小さい値（例えば、不完全な非相関化がある場合）に制限されうる。

#### 【0039】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、出力オーディオ信号を得るために、レンダリングされたオーディオ信号と1つ以上の非相関化オーディオ信号とを結合するように構成され、ここで、対角行列Mは、1つ以上の非相関化オーディオ信号Wに適用される。この場合、出力オーディオ信号の共分散行列の対角要素が所望のエネルギーに等しいように、マルチチャネル・オーディオ・デコーダは、混合行列Mの対角要素を算出するように構成される。従って、レンダリング処理によって、および/または1つ以上のダウンミックス信号および空間サイド情報に基づくオーディオ・オブジェクトの再構成によって、得られうるエネルギー損失が、補償されうる。このように、出力オーディオ信号の適当な強度が獲得される。

#### 【0040】

好みの実施の形態において、マルチチャネル・オーディオ・デコーダは、所望の共分散行列の対角要素、レンダリングされたオーディオ信号の共分散行列の対角要素、および1つ以上の非相関化信号の共分散行列の対角要素に基づいて、混合行列Mの要素を算出するように構成されうる。混合行列Mの対角要素でない要素は、ゼロに設定され、所望の共分散行列は、レンダリング処理およびオブジェクト共分散行列に対して使用されるレンダリング行列に基づいて算出されうる。さらにまた、閾値は、信号に追加される非相関化の量を制限するために使用されうる。このコンセプトは、混合行列Mの要素の非常に計算的に効果的な決定を提供する。

#### 【0041】

好みの実施の形態において、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合するための方法を決定する場合、非相関化されたオーディオ信号の相関特性または共分散特性を考慮するように構成されうる。従って、非相関化の欠損が考慮されうる。

#### 【0042】

好みの実施の形態において、所与の出力オーディオ信号が、2つ以上のレンダリングされたオーディオ信号および少なくとも1つの非相関化オーディオ信号に基づいて供給されるように、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ

10

20

30

40

50

信号と非相関化オーディオ信号とを混合するように構成されうる。このコンセプトを使用することによって、相互相関特性は、（聴覚器官の空間印象を劣化させうる）大量の非相関化信号を導く必要なく、能率的に調整されうる。

#### 【0043】

好ましい実施の形態において、出力オーディオ信号を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合するための方法を決定することを異なる制約が適用される場合において、マルチチャネル・オーディオ・デコーダは、異なるモードの間で切り替えるように構成されうる。従って、複雑さおよび処理の特徴が、処理される信号に調整されうる。

#### 【0044】

好ましい実施の形態において、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第1のモード、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、出力オーディオ信号の相互相関特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、およびレンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、の間で切り替えるために構成されうる。このように、複雑さおよび処理の特徴の両方が、現在のレンダリングされたオーディオ信号のタイプに調整されうる。オーディオ信号の空間的印象がこの種の修正によって劣化する場合、例えば、自己相関特性または自己共分散特性のみを修正し、明確に相互相関特性または相互共分散特性を修正することが有用でありうる。その一方で、にもかかわらず、出力オーディオ信号の強度を調整することは望ましい。一方、出力オーディオ信号の相互相関特性または相互共分散特性を調整することが望ましい場合がある。ここで、記載のマルチチャネル・オーディオ・デコーダは、この種の調整を許容する。ここで、第1のモードにおいて、相互相関特性または相互共分散特性を調整するために必要とされる非相関化信号の要素の量（または強度）が比較的小さいように、レンダリングされたオーディオ信号を結合することは可能である。このように、「ローカライザ可能な」信号要素は、相互相関特性または相互共分散特性を調整するために、第1のモードにおいて使用される。その一方、第2のモードにおいて、非相関化信号は、異なる聴取印象を自然にもたらす相互相関特性または相互共分散特性を調整するために使用される。従って、3つの異なるモードを提供することによって、オーディオ・デコーダは、扱っているオーディオ・コンテンツによく適応されうる。

#### 【0045】

好ましい実施の形態において、マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合するための3つのモードを指示する符号化表現のビットストリーム要素を評価し、そして、ビットストリーム要素に基づいてモードを選択するために構成される。したがって、オーディオ・エンコーダは、オーディオ・コンテンツのその情報に基づいて、適切なモードの信号を送信することができる。このように、出力オーディオ信号の最大の品質が、いかなる環境のもとでも達成されうる。

#### 【0046】

本発明による実施の形態は、少なくとも2つの入力オーディオ信号に基づいて符号化表現を供給するためのマルチチャネル・オーディオ・エンコーダをもたらす。マルチチャネル

10

20

30

40

50

ル・オーディオ・エンコーダは、少なくとも 2 つの入力オーディオ信号に基づき、1 つ以上のダウンミックス信号を供給するように構成される。さらに、マルチチャネル・オーディオ・エンコーダは、少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータを供給するように構成される。さらに、マルチチャネル・オーディオ・エンコーダは、オーディオ・エンコーダのサイドで使用されるべき複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータを供給するように構成される。従って、マルチチャネル・オーディオ・エンコーダは、現在の符号化されるオーディオ信号のタイプに良く適応される適切な非相関化モードを使用するようにオーディオ・デコーダを制御しうる。このように、ここで記載されるマルチチャネル・オーディオ・エンコーダは、以前議論されたマルチチャネル・オーディオ・デコーダと協調するためによく適応される。

10

#### 【0047】

好ましい実施の形態において、マルチチャネル・オーディオ・エンコーダは、非相関化方法パラメータを選択的に供給し、オーディオ・デコーダの処理のための、レンダリングされたオーディオ信号またはそのスケール化バージョンと 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第 1 のモード、レンダリングされたオーディオ信号またはそのスケール化バージョンと 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、出力オーディオ信号の相互相関特性または相互共分散特性を調整するために、所与の非相関化オーディオ信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第 2 のモード、およびレンダリングされたオーディオ信号またはそのスケール化バージョンと 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化オーディオ信号が、所与の非相関化オーディオ信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第 3 のモード、である以上に示す 3 つのモードの 1 つの信号を送信するように構成される。このように、マルチチャネル・オーディオ・エンコーダは、オーディオ・コンテンツに基づいて、上記において議論された 3 つのモードを通してマルチチャネル・オーディオ・デコーダを切り替えることができ、ここで、マルチチャネル・オーディオ・デコーダが処理するモードは、現在の符号化されたオーディオ・コンテンツのタイプにマルチチャネル・オーディオ・エンコーダによってよく適応されうる。しかしながら、いくつかの実施の形態において、オーディオ・デコーダの処理のための上記の 3 つのモードの 1 つか 2 つのみが使用されうる（または利用されうる。）。

20

#### 【0048】

好ましい実施の形態において、マルチチャネル・オーディオ・エンコーダは、入力オーディオ信号が比較的高い相関か比較的低い相関を含むかどうかに基づいて、非相関化方法パラメータを選択するように構成される。このように、デコーダにおいて使用される非相関化の適応は、現在の符号化されるオーディオ信号の重要な特徴に基づいてなされうる。

30

#### 【0049】

好ましい実施の形態において、マルチチャネル・オーディオ・エンコーダは、入力オーディオ信号の間の相関または共分散が比較的高い場合、第 1 のモードまたは第 2 のモードを指定するための非相関化方法パラメータを選択し、入力オーディオ信号の間の相関または共分散が比較的低い場合、第 3 のモードを指定するための非相関化方法パラメータを選択するように構成される。従って、入力オーディオ信号の間の比較的小さい相関または共分散の場合において、相互共分散特性または相互相関特性の修正でない復号化モードが選択される。そのような信号は実質的に独立しているので、これは、相互相関または相互共分散の適応のための必要を取り除く、比較的低い相関（または共分散）を有する信号に大して効果的な選択であることが分かっている。むしろ、（比較的小さい相関または共分散を有する）実質的な独立入力オーディオ信号のための相互相関または相互共分散の調整は

40

50

、典型的には、オーディオの品質を劣化させ、そして、同時に復号化の複雑さを増加させる。このように、このコンセプトは、マルチチャネル・オーディオ・エンコーダに入力される信号に、マルチチャネル・オーディオ・デコーダの合理的な適応を許容する。

#### 【0050】

本発明による実施の形態は、符号化表現に基づく少なくとも2つの出力オーディオ信号を供給するための方法をもたらす。方法は、複数のレンダリングされたオーディオ信号を得るために、1つ以上のレンダリング・パラメータに基づいて、符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするステップを含む。方法は、レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するステップと、出力オーディオ信号を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合するステップも含む。この方法は、上記したマルチチャネル・オーディオ・デコーダと同じ考察に基づく。さらに、方法は、マルチチャネル・オーディオ・デコーダに関して上記について議論された特徴および機能のいずれかによって補充ができる。

10

#### 【0051】

本発明による他の実施の形態は、少なくとも2つの入力オーディオ信号に基づく符号化表現を供給するための方法をもたらす。方法は、少なくとも2つの入力オーディオ信号に基づく1つ以上のダウンミックス信号を供給するステップと、少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータを供給するステップと、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータを供給するステップと、を含む。この方法は、上記したマルチチャネル・オーディオ・エンコーダと同じ考察に基づく。さらに、方法は、マルチチャネル・オーディオ・エンコーダに関して本願明細書において議論された特徴および機能のいずれかによって補充されうる。

20

#### 【0052】

本発明による他の実施の形態は、上記の方法の1つ以上を実行するためのコンピュータ・プログラムをもたらす。

#### 【0053】

本発明による他の実施の形態は、ダウンミックス信号の符号化表現と、少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータの符号化表現と、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している符号化非相関化方法パラメータを含む。この符号化オーディオ表現は、適切な非相関化モードの信号を送信することを許容し、従って、マルチチャネル・オーディオ・エンコーダおよびマルチチャネル・オーディオ・デコーダに関して記載される効果を実装するのに役立つ。

30

#### 【0054】

本発明の好ましい実施の形態は、添付の図面に関してその後述べられる。

#### 【図面の簡単な説明】

#### 【0055】

【図1】図1は、本発明の実施の形態によるマルチチャネル・オーディオ・デコーダのブロック概略図を示す。

40

【図2】図2は、本発明の実施の形態によるマルチチャネル・オーディオ・エンコーダのブロック概略図を示す。

【図3】図3は、本発明の実施の形態による符号化表現に基づいて少なくとも2つの出力オーディオ信号を供給するための方法のフローチャートを示す。

【図4】図4は、本発明の実施の形態による少なくとも2つの入力オーディオ信号に基づいて符号化表現を供給するための方法のフローチャートを示す。

【図5】図5は、本発明の実施の形態による符号化オーディオ表現の概略図を示す。

【図6】図6は、本発明の実施の形態によるマルチチャネル非相関器のブロック概略図を示す。

50

【図 7】図 7 は、本発明の実施の形態によるマルチチャネル・オーディオ・デコーダのブロック概略図を示す。

【図 8】図 8 は、本発明の実施の形態によるマルチチャネル・オーディオ・エンコーダのブロック概略図を示す。

【図 9】図 9 は、本発明の実施の形態による複数の非相関化入力信号に基づいて複数の非相関化信号を供給するための方法のフローチャートを示す。

【図 10】図 10 は、本発明の実施の形態による符号化表現に基づいて少なくとも 2 つの出力オーディオ信号を供給するための方法のフローチャートを示す。

【図 11】図 11 は、本発明の実施の形態による少なくとも 2 つの入力オーディオ信号に基づいて符号化表現を供給するための方法のフローチャートを示す。 10

【図 12】図 12 は、本発明の実施の形態による符号化表現の概略図を示す。

【図 13】図 13 は、パラメトリックダウンミックス / アップミックスのコンセプトに基づく M M S E の概観を提供する概略図である。

【図 14】図 14 は、3 次元空間における直角原理のための幾何学的な表現を示す。

【図 15】図 15 は、本発明の実施の形態によるレンダリングされた出力に適用される非相関化を伴うパラメトリック再構成システムのブロック概略図を示す。 20

【図 16】図 16 は、非相関化装置のブロック概略図を示す。

【図 17】図 17 は、本発明の実施の形態による低減された複雑さの非相関化装置のブロック概略図を示す。

【図 18】図 18 は、本発明の実施の形態によるスピーカ位置のテーブル表現を示す。 20

【図 19 a】図 19 a は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。

【図 19 b】図 19 b は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。

【図 19 c】図 19 c は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。 30

【図 19 d】図 19 d は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。

【図 19 e】図 19 e は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。 30

【図 19 f】図 19 f は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。

【図 19 g】図 19 g は、 $N = 22$  および  $K$  が 5 と 11 との間であるプレミキシング係数のテーブル表現を示す。

【図 20 a】図 20 a は、 $N = 10$  および  $K$  が 2 と 5 との間であるプレミキシング係数のテーブル表現を示す。 40

【図 20 b】図 20 b は、 $N = 10$  および  $K$  が 2 と 5 との間であるプレミキシング係数のテーブル表現を示す。

【図 20 c】図 20 c は、 $N = 10$  および  $K$  が 2 と 5 との間であるプレミキシング係数のテーブル表現を示す。

【図 20 d】図 20 d は、 $N = 10$  および  $K$  が 2 と 5 との間であるプレミキシング係数のテーブル表現を示す。 40

【図 21 a】図 21 a は、 $N = 8$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。

【図 21 b】図 21 b は、 $N = 8$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。

【図 21 c】図 21 c は、 $N = 8$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。

【図 21 d】図 21 d は、 $N = 7$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。 50

【図 21 e】図 21 e は、 $N = 7$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。

【図 21 f】図 21 f は、 $N = 7$  および  $K$  が 2 と 4 との間であるプレミキシング係数のテーブル表現を示す。

【図 22 a】図 22 a は、 $N = 5$  および  $K = 3$  である係数のプレミキシングのテーブル表現を示す。

【図 22 b】図 22 b は、 $N = 5$  および  $K = 2$  であるプレミキシング係数のテーブル表現を示す。

【図 23】図 23 は、 $N = 2$  および  $K = 1$  であるプレミキシング係数のテーブル表現を示す。 10

【図 24】図 24 は、チャネル信号のグループのテーブル表現を示す。

【図 25】図 25 は、SAOC Specific Config() の構文または同等の SAOC3DSpecificConfig() に含まれる付加的なパラメータの構文表現を示す。

【図 26】図 26 は、ビットストリーム変数 bsDecorrrelationMethod のための異なる値のテーブル表現を示す。

【図 27】図 27 は、ビットストリーム変数 bsDecorrlationLevel によって指示される異なる非相関化レベルおよび出力構成のための非相関器の数のテーブル表現を示す。

【図 28】図 28 は、ブロック概略図の形式において、3Dオーディオ・エンコーダの上の概要を示す。 20

【図 29】図 29 は、ブロック概略図の形式において、3Dオーディオ・デコーダの上の概要を示す。

【図 30】図 30 は、フォーマット変換器の構造のブロック概略図を示す。

【図 31】図 31 は、本発明の実施の形態によるダウンミックス・プロセッサのブロック概略図を示す。

【図 32】図 32 は、SAOC ダウンミックスオブジェクトの異なる数のための復号化モードのテーブル表現を示す。

【図 33 a - 1】図 33 a - 1 は、ビットストリーム要素「SAOC3DSpecificConfig」の構文表現を示す。 30

【図 33 a - 2】図 33 a - 2 は、ビットストリーム要素「SAOC3DSpecificConfig」の構文表現を示す。

【図 33 b】図 33 b は、ビットストリーム要素「SAOC3DSpecificConfig」の構文表現を示す。

【発明を実施するための形態】

【0056】

1. 図 1 によるマルチチャネル・オーディオ・デコーダ

図 1 は、本発明の実施の形態によるマルチチャネル・オーディオ・デコーダ 100 のブロック概略図を示す。

【0057】

マルチチャネル・オーディオ・デコーダ 100 は、符号化表現 110 を受信して、それに基づいて、少なくとも 2 つの出力オーディオ信号 112, 114 を供給するように構成される。

【0058】

好ましくは、マルチチャネル・オーディオ・デコーダ 100 は、符号化表現 110 に基づいて復号化オーディオ信号 122 を供給するように構成されるデコーダ 120 を含む。さらに、マルチチャネル・オーディオ・デコーダ 100 は、複数のレンダリングされたオーディオ信号 134, 136 を得るために、1 つ以上のレンダリング・パラメータ 132 に基づいて、(例えば、デコーダ 120 によって) 符号化表現 110 に基づいて得られる複数の復号化オーディオ信号 122 をレンダリングするために構成されるレンダラー 130

10

20

30

40

50

を含む。さらに、マルチチャネル・オーディオ・デコーダ100は、レンダリングされたオーディオ信号134, 136から1つ以上の非相関化オーディオ信号142, 144を導出するように構成される非相関器140を含む。さらに、マルチチャネル・オーディオ・デコーダ100は、出力オーディオ信号112, 114を得るために、レンダリングされたオーディオ信号134, 136またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号142, 144とを結合するように構成されるコンバイナ150を含む。

#### 【0059】

しかしながら、上記の機能が与えられる限り、マルチチャネル・オーディオ・デコーダ100の異なるハードウェア構成が可能である点に留意されたい。

10

#### 【0060】

マルチチャネル・オーディオ・デコーダ100の機能に関して、非相関化オーディオ信号142, 144は、レンダリングされたオーディオ信号134, 136から導出され、非相関化オーディオ信号142, 144は、出力オーディオ信号112, 114を得るために、レンダリングされたオーディオ信号134, 136と結合される点に留意されたい。レンダリングされたオーディオ信号134, 136から非相関化オーディオ信号142, 144を導出することによって、レンダリングされたオーディオ信号134, 136の数は、典型的には、レンダラ130に入力される復号化オーディオ信号122の数から独立しているので、特に効果的な処理が達成されうる。このように、典型的には、非相関化の効果は、実施効率を改良する復号化オーディオ信号122の数から独立している。さらに、レンダリングの後の非相関化を適用することは、非相関化がレンダリングの前に適用される場合において、複数の非相関化信号を結合する場合に、レンダラによって引き起こされるアーティファクトの導入を回避する。さらに、レンダリングされたオーディオ信号の特性は、典型的には、良好な品質の出力オーディオ信号を結果として得る非相関器140によって実行される非相関化において考慮されうる。

20

#### 【0061】

さらに、マルチチャネル・オーディオ・デコーダ100が、本願明細書において記載されている特徴および機能によって補充されうる点に留意されたい。特に、本願明細書において記載されるような個々の改良は、それによる処理の効率化および/または出力オーディオ信号の品質の改良のために、マルチチャネル・オーディオ・デコーダ100に導入されうる点に留意されたい。

30

#### 【0062】

##### 2. 図2によるマルチチャネル・オーディオ・エンコーダ

図2は、本発明の実施の形態によるマルチチャネル・オーディオ・エンコーダ200のプロック概略図を示す。マルチチャネル・オーディオ・エンコーダ200は、2つ以上の入力オーディオ信号210, 212を受信し、それに基づいて符号化表現214を供給するように構成される。マルチチャネル・オーディオ・エンコーダは、少なくとも2つの入力オーディオ信号210, 212に基づいて、1つ以上のダウンミックス信号222を供給するように構成されるダウンミックス信号プロバイダ220を含む。さらに、マルチチャネル・オーディオ・エンコーダ200は、少なくとも2つの入力オーディオ信号210, 214の間の関係（例えば、相互相關、相互共分散、レベル差等）を記述している1つ以上のパラメータ232を供給するように構成されるパラメータ・プロバイダ230を含む。

40

#### 【0063】

さらに、マルチチャネル・オーディオ・エンコーダ200は、オーディオ・デコーダのサイドにおいて使用されるべき複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータ242を供給するように構成される非相関化方法パラメータ・プロバイダ240も含む。1つ以上のダウンミックス信号222、1つ以上のパラメータ232および非相関化方法パラメータ242は、例えば、符号化表現214に、符号化の形式で含まれる。

50

## 【0064】

しかしながら、上記のような機能が満足される限り、マルチチャネル・オーディオ・エンコーダ200のハードウェア構成は異なりうることに留意されたい。換言すれば、個々のロック（例えば、ダウンミックス信号プロバイダ220に、パラメータ・プロバイダ230に、そして非相関化方法パラメータ・プロバイダ240に）マルチチャネル・オーディオ・エンコーダ200の機能の配分は、例として、考慮されるべきである。

## 【0065】

マルチチャネル・オーディオ・エンコーダ200の機能に関して、1つ以上のダウンミックス信号222および1つ以上のパラメータ232が、例えば、S A O Cマルチチャネル・オーディオ・エンコーダまたはU S A Cマルチチャネル・オーディオ・エンコーダにおけるような従来の方法において供給される点に留意されたい。しかしながら、マルチチャネル・オーディオ・エンコーダ200によっても供給され、そして、符号化表現214に含まれる非相関化方法パラメータ242は、入力オーディオ信号210, 212にまたは所望の再生品質に非相関化モードを適用するために使用されうる。従って、非相関化モードは、オーディオ・コンテンツの異なるタイプに適用されうる。例えば、異なる非相関化モードは、入力オーディオ信号210, 212が強く相関しているオーディオ・コンテンツのタイプに対して、および入力オーディオ信号210, 212が独立しているオーディオ・コンテンツのタイプに対して、選択される。さらに、異なる非相関化モードは、空間印象が特に重要であるオーディオ・コンテンツのタイプに対して、および空間印象がより重要でないかまたは下位の重要性（例えば、個々のチャネルの再生と比較した場合）におけるオーディオ・コンテンツのタイプに対して、非相関化モードパラメータ242によって信号を送信されうる。従って、符号化表現214を受信するマルチチャネル・オーディオ・デコーダは、マルチチャネル・オーディオ・エンコーダ200によって制御され、復号化の複雑さと再生品質の間の最良の可能な妥協をもたらす復号化モードを設定されうる。

10

20

30

40

## 【0066】

さらに、マルチチャネル・オーディオ・エンコーダ200は、本願明細書において記載されている特徴および機能のいずれかによって補充されうる点に留意されたい。本願明細書において記載されている可能な付加的な特徴および改良は、それによって、マルチチャネル・オーディオ・エンコーダ200を改良するように（または強化するように）、個々にまたは組み合わせて、マルチチャネル・オーディオ・エンコーダ200に追加されうる点に留意されたい。

## 【0067】

## 3. 図3による少なくとも2つの出力オーディオ信号を供給するための方法

図3は、符号化表現に基づいて少なくとも2つの出力オーディオ信号を供給するための方法300のフローチャートを示す。方法は、複数のレンダリングされたオーディオ信号を得るために、1つ以上のレンダリング・パラメータに基づいて得られた複数の復号化オーディオをレンダリングするステップ310を含む。方法300は、また、レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するステップ320を含む。方法300は、また、出力オーディオ信号332を得るために、レンダリングされたオーディオ信号またはそのスケール化バージョンと1つ以上の非相関化オーディオ信号とを結合するステップ330を含む。

## 【0068】

方法300は、図1によるマルチチャネル・オーディオ・デコーダ100と同じ考察に基づく点に留意されたい。さらに、方法300は、（個々に、または組み合わせて）本願明細書において記載される特徴および機能のいずれかによって補充されうる点に留意されたい。

## 【0069】

## 4. 図4による符号化表現を提供するための方法

図4は、少なくとも2つの入力オーディオ信号に基づく符号化表現を供給する方

50

法 4 0 0 のフロー チャートを示す。方法 4 0 0 は、少なくとも 2 つの入力オーディオ信号に基づく 1 つ以上のダウンミックス信号を供給するステップ 4 1 0 を含む。さらに、方法 4 0 0 は、少なくとも 2 つの入力オーディオ信号 4 1 2 の間の関係を記述している 1 つ以上のパラメータを供給するステップ 4 2 0 と、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータを供給するステップ 4 3 0 とを含む。従って、好ましくは、1 つ以上のダウンミックス信号の符号化表現、少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータ、および非相関化方法パラメータを含む符号化表現 4 3 2 が供給される。

#### 【 0 0 7 0 】

上記の説明も適用されるように、方法 4 0 0 は、マルチチャネル・オーディオ・エンコーダ 2 0 0 と同じ考察に基づく点に留意されたい。10

#### 【 0 0 7 1 】

さらに、ステップ 4 1 0 , 4 2 0 , 4 3 0 の命令は、柔軟に変化することができ、これが、方法 4 0 0 のための実行環境において可能な限り、ステップ 4 1 0 , 4 2 0 , 4 3 0 も、並行して実行可能である点に留意されたい。さらに、方法 4 0 0 は、個々に、または組み合わせて、本願明細書において記載される特徴および機能のいずれかによって補充されうる点に留意されたい。例えば、方法 4 0 0 は、マルチチャネル・オーディオ・エンコーダに関して本願明細書において記載される特徴および機能のいずれかによって補充されうる。しかしながら、符号化表現 4 3 2 を受信する本願明細書に記載されるマルチチャネル・オーディオ・デコーダの特徴および機能に対応する特徴および機能を含むことも可能である。20

#### 【 0 0 7 2 】

#### 5 . 図 5 による符号化オーディオ表現

図 5 は、本発明の実施の形態による符号化オーディオ表現 5 0 0 の概略図を示す。

#### 【 0 0 7 3 】

符号化オーディオ表現 5 0 0 は、ダウンミックス信号の符号化表現 5 1 0 、少なくとも 2 つのオーディオ信号の間の関係を記述している 1 つ以上のパラメータの符号化表現 5 2 0 を含む。符号化オーディオ表現 5 0 0 は、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している符号化非相関化方法パラメータ 5 3 0 も含む。従って、符号化オーディオ表現は、オーディオ・エンコーダからオーディオ・デコーダに非相関化モードの信号を送信することを許容する。従って、符号化オーディオ表現は、( 例えば、1 つ以上のダウンミックス信号の符号化表現 5 1 0 によって、および少なくとも 2 つのオーディオ信号 ( 例えば、1 つ以上のダウンミックス信号の符号化表現 5 1 0 にダウンミックスされている少なくとも 2 つのオーディオ信号 ) ) の間の関係を記述している 1 つ以上のパラメータの符号化表現 5 2 0 によって記述されるオーディオ・コンテンツの特性によく適応される非相関化モードを得ることが可能である。このように、符号化オーディオ表現 5 0 0 は、特に良好な聴覚器官の空間印象および / または特に、聴覚器官の空間印象と復号化の複雑さとの間の良好なトレードオフを伴う符号化オーディオ表現 5 0 0 によって表わされるオーディオ・コンテンツのレンダリングを許容する。30

#### 【 0 0 7 4 】

さらに、符号化表現 5 0 0 が、個々に、または組み合わせて、マルチチャネル・オーディオ・エンコーダおよびマルチチャネル・オーディオ・デコーダに関して記載されている特徴および機能のいずれかによって補充されうる点に留意されたい。40

#### 【 0 0 7 5 】

#### 6 . 図 6 によるマルチチャネル非相関器

図 6 は、本発明の実施の形態によるマルチチャネル非相関器 6 0 0 のプロック概略図を示す。

#### 【 0 0 7 6 】

マルチチャネル非相関器 6 0 0 は、N 個の非相関器入力信号 6 1 0 a ~ 6 1 0 n の第 1

のセットを受信し、それに基づいて、N' 個の非相関器出力信号 612a ~ 612n' の第2のセットを供給するように構成される。換言すると、マルチチャネル非相関器 600 は、非相関器入力信号 610a ~ 610n に基づいて複数の（少なくともおよそ）非相関化信号 612a ~ 612n' を提供するように構成される。

#### 【0077】

マルチチャネル非相関器 600 は、N 個の非相関器入力信号 610a ~ 610n の第1のセットを K 個の非相関器入力信号 622a ~ 622k の第2のセットにプレミックスするように構成される、プレミキサ 620 を含み、K は、N よりも小さい（K および N は整数である）。マルチチャネル非相関器 600 は、K 個の非相関器入力信号 622a ~ 622k の第2のセットに基づいて K' 個の非相関器出力信号 632a ~ 632k' の第1のセットを供給するように構成される非相関化（または非相関器の主要部）も含む。さらに、マルチチャネル非相関器は、K' 個の非相関器出力信号 632a ~ 632k' の第1のセットを N' 個の非相関器出力信号 612a ~ 612n' の第2のセットにアップミックスするように構成されるポスト・ミキサ 640 を含み、N' は K' より大きい（N' および K' は整数である）。

10

#### 【0078】

しかしながら、マルチチャネル非相関器 600 の所与の構成は、例として考慮されるべきであり、本願明細書において記載されている機能が提供される限り、マルチチャネル非相関器 600 を機能的なブロック（例えば、プレミキサ 620、非相関化もしくは非相関器の主要部 630、およびポスト・ミキサ 640）にさらに分割することは必要ない点に留意されたい。

20

#### 【0079】

マルチチャネル非相関器 600 の機能に関して、現実の非相関化が、例えば、直接、N 個の非相関器入力信号が適用されるコンセプトを比較したとき、N 個の非相関器入力信号の第1のセットから K 個の非相関器入力信号の第2のセットを導出するようにプレミックスを実行し、そして、（プレミックスされまたは「ダウンミックされた」）K 個の非相関器入力信号の第2のセットに基づいて非相関化を実行するコンセプトは、複雑さの低減をもたらすことに留意されたい。さらに、N' 個の非相関器出力信号の第2の（アップミックスされた）セットは、アップミキサ 640 によって実行されうるポストミキシングに基づいて、現実の非相関化の結果である非相関器出力信号の第1の（元の）セットに基づいて得られる。このように、マルチチャネル非相関器 600 は、（外側からみられたとき）効果的に N 個の非相関器入力信号を受信し、そして、それに基づいて、N' 個の非相関器出力信号を供給する。その一方で、現実の非相関器の主要部 630 は、より少ない数の信号（すなわち、K 個の非相関器入力信号の第2のセットの K 個のダウンミックスされた非相関器入力信号 622a ~ 622k' のみを処理するだけである。このように、マルチチャネル非相関器 600 の複雑さは、従来の非相関器と比較したとき、非相関化（または非相関器の主要部）630 の入力サイドでの（好ましくは、いかなる非相関化の機能のない線形のプレミキシングである）ダウンミックまたは「プレミキシング」を実行することによって、そして、非相関化（または非相関器の主要部 630）の（元の）出力信号 632a ~ 632k' に基づいて、（例えば、いかなる追加の非相関化の機能のない線形のアップミキシングである）アップミキシングまたは「ポストミキシング」を実行することによって、実質的に低減されうる。

30

#### 【0080】

さらに、マルチチャネル非相関器 600 は、マルチチャネル非相関化に関して、またマルチチャネル・オーディオ・デコーダにも関して、本願明細書において記載される特徴および機能のいずれかによって補充されうる点に留意されたい。本願明細書において記載される特徴は、それによって、マルチチャネル非相関器 600 を改良するかまたは強化するように、個々に、または、組み合わせて、マルチチャネル非相関器 600 に追加されうる点に留意されたい。

40

#### 【0081】

50

複雑さの低減のないマルチチャネル非相関器は、 $K = N$ （そして、おそらく $K' = N$ 、または $K = N = K' = N'$ でさえ）、に対して上記したマルチチャネル非相関器から導出されうる点に留意されたい。

#### 【0082】

##### 7. 図7によるマルチチャネル・オーディオ・デコーダ

図7は本発明の実施の形態によるマルチチャネル・オーディオ・デコーダ700のプロック概略図を示す。

#### 【0083】

マルチチャネル・オーディオ・デコーダ700は、符号化表現710を受信し、それに基づいて、少なくとも2つの出力信号712, 714を供給するように構成される。マルチチャネル・オーディオ・デコーダ700は、図6によるマルチチャネル非相関器600と実質的に同一であるマルチチャネル非相関器720を含む。さらに、マルチチャネル・オーディオ・デコーダ700は、従来技術において当業者により知られている、または他のマルチチャネル・オーディオ・デコーダに関して本願明細書に記載されるようなマルチチャネル・オーディオ・デコーダの特徴または機能のいずれかを含みうる。

10

#### 【0084】

さらに、マルチチャネル・オーディオ・デコーダ700が、高効率のマルチチャネル非相関器720を使用するので、マルチチャネル・オーディオ・デコーダ700は、従来のマルチチャネル・オーディオ・デコーダと比較した場合、特に高い効率を含む。

20

#### 【0085】

##### 8. 図8によるマルチチャネル・オーディオ・エンコーダ

図8は、本発明の実施の形態によるマルチチャネル・オーディオ・エンコーダ800のプロック概略図を示す。マルチチャネル・オーディオ・エンコーダ800は、少なくとも2つの入力オーディオ信号810, 812を受信し、それに基づいて、入力オーディオ信号810, 812によって表わされるオーディオ・コンテンツの符号化表現814を供給するように構成される。

20

#### 【0086】

マルチチャネル・オーディオ・エンコーダ800は、少なくとも2つの入力オーディオ信号810, 812に基づいて1つ以上のダウンミックス信号822を供給するように構成されるダウンミックス信号プロバイダ820を含む。マルチチャネル・オーディオ・エンコーダ800は、1つ以上のパラメータ832（例えば、相互相關パラメータもしくは相互共分散パラメータ、または内部オブジェクト相關パラメータおよび／もしくはオブジェクト・レベル差パラメータ）を供給するように構成されるパラメータ・プロバイダ830も含む。さらに、マルチチャネル・オーディオ・エンコーダ800は、（符号化表現814を受信する）オーディオ・デコーダのサイドで使用される非相関化の複雑さを記述している非相関化複雑さパラメータ842を供給するように構成される非相関化複雑さパラメータ・プロバイダ840を含む。1つ以上のダウンミックス信号822、1つ以上のパラメータ832、および非相関化複雑さパラメータ842が、好ましくは符号化形式において符号化表現814に含まれる。

30

#### 【0087】

しかしながら、マルチチャネル・オーディオ・エンコーダ800（例えば、ダウンミックス信号プロバイダ820、パラメータ・プロバイダ830および非相関化複雑さパラメータ・プロバイダ840の存在）の内部構造は、単なる例示として考慮されるべきである。異なる構成は、本願明細書に記載されている機能が達成する限り、可能である。

40

#### 【0088】

マルチチャネル・オーディオ・エンコーダ800の機能に関して、マルチチャネル・エンコーダは、符号化表現814を供給する点に留意されたい。ここで、1つ以上のダウンミックス信号822および1つ以上のパラメータ832は、（例えば、従来のS A O C オーディオ・エンコーダまたはU S A C オーディオ・エンコーダのような）従来のオーディオ・エンコーダによって供給されるダウンミックス信号およびパラメータに近似するか、

50

または等しい。しかしながら、マルチチャネル・オーディオ・エンコーダ800は、オーディオ・デコーダのサイドにおいて適用される非相関化の複雑さを決定することを許容する、非相関化複雑さパラメータ842も供給するように構成される。従って、非相関化の複雑さは、現在符号化されるオーディオ・コンテンツに適応されうる。例えば、入力オーディオ信号の特性についてエンコーダ側の情報に基づいて、達成可能なオーディオ品質に対応する所望の非相関化の複雑さの信号を送信することが可能である。例えば、空間的な特性がオーディオ信号に対して重要であることが分かっている場合、空間的な特性が重要でないとき、非相関化複雑さパラメータ842を使用して、より高い非相関化の複雑さの信号の送信がされうる。あるいは、高い複雑さの非相関化が、他の理由のためのオーディオ・デコーダのサイドで必要とされるような、オーディオ・コンテンツ、または全てのオーディオ・コンテンツの通過であることが分かっている場合、高い非相関化の複雑さの使用は、非相関化複雑さパラメータ842を使用して、信号の送信がされうる。

10

## 【0089】

要約すると、マルチチャネル・オーディオ・エンコーダ800は、信号特性、またはマルチチャネル・オーディオ・エンコーダ800によって設定されうる所望の再生特性に適用される非相関化の複雑さを使用するために、マルチチャネル・オーディオ・デコーダを制御する可能性を提供する。

20

## 【0090】

さらに、マルチチャネル・オーディオ・エンコーダ800が、個々に、または、組み合わせて、マルチチャネル・オーディオ・エンコーダに関して、本願明細書に記述される特徴および機能のいずれかによって補充されうる点に留意されたい。例えば、マルチチャネル・オーディオ・エンコーダに関して、本願明細書に記載される特徴のいくつかまたは全ては、マルチチャネル・オーディオ・エンコーダ800に追加されうる。さらに、マルチチャネル・オーディオ・エンコーダ800は、本願明細書において記載されるマルチチャネル・オーディオ・デコーダとの協力に対して適応されうる。

20

## 【0091】

9. 図9による複数の非相関器入力信号に基づいて複数の非相関化信号を供給するための方法

30

図9は、複数の非相関器入力信号に基づいて、複数の非相関信号を供給するような方法900のフローチャートを示す。

## 【0092】

方法900は、N個の非相関器入力信号の第1のセットをK個の非相関器入力信号の第2のセットにプレミックスするステップ910を含み、Kは、Nよりも小さい。方法900は、K個の非相関器入力信号の第2のセットに基づいて、K'個の非相関器出力信号の第1のセットを供給するステップ920も含む。例えば、K'個の非相関器出力信号の第1のセットは、非相関化を使用してK個の非相関器入力信号の第2のセットに基づいて供給され、そして、それは、例えば、非相関器の主要部を使用するか、または非相関化アルゴリズムを使用して実行されうる。更に、方法900は、K'個の非相関器出力信号の第1のセットをN'個の非相関器出力信号の第2のセットにポストミックスするステップ930を含み、N'はK'よりも大きい(N'およびK'は整数である)。従って、方法900の出力であるN'個の非相関器出力信号の第2のセットは、方法900に入力されるN個の非相関器入力の第1のセットに基づいて供給されうる。

40

## 【0093】

方法900が、上記のマルチチャネル非相関器と同じ考察に基づく点に留意されたい。さらに、方法900は、個々に、または、組み合わせて、マルチチャネル非相関器に関して(そして、適用できる場合、マルチチャネル・オーディオ・エンコーダに関して)、本願明細書において記載されている特徴および機能のいずれかによって補充されうる点に留意されたい。

## 【0094】

10. 図10による符号化表現に基づいて少なくとも2つの出力オーディオ信号を供給す

50

る方法

図10は、符号化表現に基づいて少なくとも2つの出力オーディオ信号を供給するための方法1000のフローチャートを示す。

**【0095】**

方法1000は、符号化表現1012に基づいて少なくとも2つの出力オーディオ信号1014, 1016を供給するステップ1010を含む。方法1000は、図9による方法900に従って、複数の非相関器入力信号に基づいて複数の非相関化信号を供給するステップ1020を含む。

**【0096】**

方法1000は、図7によるマルチチャネル・オーディオ・デコーダ700と同じ考察に基づく点に留意されたい。10

**【0097】**

また、方法1000は、個々に、または、組み合わせて、マルチチャネル・デコーダについて、本願明細書において記載されている特徴および機能のいずれかによって補充される点に留意されたい。

**【0098】**

11. 図11による少なくとも2つの入力オーディオ信号に基づいて符号化表現を供給する方法

図11は、少なくとも2つの入力オーディオ信号に基づいて符号化表現を供給する方法1100のフローチャートを示す。20

**【0099】**

方法1000は、少なくとも2つの入力オーディオ信号1112, 1114に基づいて1つ以上のダウンミックス信号を供給するステップ1110を含む。方法1100は、少なくとも2つの入力オーディオ信号1112, 1114の間の関係を記述している1つ以上のパラメータを供給するステップ1120も含む。さらに、方法1100は、オーディオ・デコーダのサイドで使用される非相関化の複雑さを記述している非相関化複雑さパラメータを供給するステップ1130を含む。従って、符号化表現1132は、少なくとも2つの入力オーディオ信号1112, 1114に基づいて供給される。ここで、符号化表現は、典型的には、符号化形式において、1つ以上のダウンミックス信号、少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータ、および非相関化複雑さパラメータを含む。30

**【0100】**

ステップ1110, 1120, 1130は、並行して、または本発明によるいくつかの実施の形態における異なる命令において実行されうる。さらに、方法1100は、図8によるマルチチャネル・オーディオ・エンコーダ800として同じ考察に基づき、そして、方法1100は、個々に、または、組み合わせて、マルチチャネル・オーディオ・エンコーダについて、本願明細書において記載されている特徴および機能のいずれかによって補充される点に留意されたい。さらに、方法1100は、マルチチャネル・オーディオ・デコーダおよび本願明細書において記載されている少なくとも2つの出力オーディオ信号を供給するための方法をマッチするように適応されうる点に留意されたい。40

**【0101】**

12. 図12による符号化オーディオ表現

図12は、本発明の実施の形態による符号化オーディオ表現の概略図を示す。符号化オーディオ表現は、ダウンミックス信号の符号化表現1210、少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータの符号化表現1220、およびオーディオ・デコーダのサイドにおいて使用される非相関化の複雑さを記述している符号化非相関化複雑さパラメータ1230を含む。従って、符号化オーディオ表現1200は、改良された復号化効率、および改良されたオーディオ品質または符号化効率とオーディオ品質とのトレードオフの改良の可能性をもたらすマルチチャネル・オーディオ・デコーダによって使用される非相関化の複雑さを調整することを許容する。さらに、符号化オ50

ーディオ表現 1200 は、本願明細書において記載されるようにマルチチャネル・オーディオ・エンコーダによって供給されえ、そして、本願明細書において記載されるようにマルチチャネル・オーディオ・エンコーダによって使用されうる点に留意されたい。従って、符号化オーディオ表現 1200 は、マルチチャネル・オーディオ・エンコーダに関しておよびマルチチャネル・オーディオ・デコーダに関して記載される特徴のいずれかによつて補充されうる。

#### 【0102】

##### 13. 表記法および基本的な検討事項

近年では、複数のオーディオ・オブジェクトを含んでいるオーディオ・シーンのビットレートの効果的な伝送／格納のためのパラメータの技術は、オーディオ符号化の分野（例えば、参考文献 [BCC] [JSC] [SAOC] [SAOC1] [SAOC2] を参照）およびインフォームド（informed）音源分離の分野（例えば、参考文献 [ISS1] [ISS2] [ISS3] [ISS4] [ISS5] [ISS6] を参照）において提案されている。これらの技術は、伝送され／格納されたオーディオ・シーンおよび／またはオーディオ・シーンにおける音源オブジェクトを記述している追加のサイド情報に基づいて、所望の出力オーディオ・シーンまたはオーディオ音源オブジェクトを再構成することを意図する。この再構成は、パラメータのインフォームド音源分離を使用してデコーダにおいて生じる。さらに、参照は、例えば、国際標準 ISO / IEC 23003-1 : 2007において記載されるいわゆる「MPEG Surround (MPEGサラウンド)」のコンセプトにもなされる。さらに、参照は、国際標準 ISO / IEC 23003-2 : 2010において記載されるいわゆる「Spatial Audio Object Coding (空間オーディオ・オブジェクト符号化)」にもなされる。さらに、参照は、国際規格 ISO / IEC 23003-3 : 2012において記載されるいわゆる「Unified Speech and Audio Coding (音声音響統合符号化方式)」にもなされる。これらの標準からのコンセプトは、例えば、本願明細書において記載されるマルチチャネル・オーディオ・エンコーダおよび本願明細書において記載されるマルチチャネル・オーディオ・デコーダにおいて、本発明による実施の形態において使用される。ここで、いくつかの適応は必要とされうる。

#### 【0103】

以下に、いくつかの背景情報が、記載される。特に、パラメータの分離方式における概要は、MPEG 空間オーディオ・オブジェクト符号化 (Spatial Audio Object Coding : SAOC) 技術（例えば、参考文献 [SAOC] を参照）の実施例を使用して、提供される。この方法の数学的プロパティは、考慮される。

#### 【0104】

## 1 3 . 1 表記法および定義

以下の数学的表記法は、現在の明細書において適用される。

$N_{Objects}$  オーディオ・オブジェクト信号の数

$N_{DmxCh}$  ダウンミックス（処理済みの）チャネルの数

$N_{UpmixCh}$  アップミックス（出力）チャネルの数

$N_{Samples}$  処理済のデータサンプルの数

10

D ダウンミックス行列、サイズ  $N_{DmxCh} \times N_{Objects}$

X 入力オーディオ・オブジェクト信号、サイズ  $N_{Objects} \times N_{Samples}$

$E_x$  オブジェクト共分散行列、サイズ  $N_{Objects} \times N_{Objects}$

$E_x = X X^H$  として定義される

Y ダウンミックスオーディオ信号、サイズ  $N_{DmxCh} \times N_{Samples}$

$Y = D X$  として定義される

20

$E_y$  ダウンミックス信号の共分散行列、サイズ  $N_{DmxCh} \times N_{DmxCh}$

$E_y = Y Y^H$  として定義される

G パラメトリック音源推定行列、サイズ  $N_{Objects} \times N_{DmxCh}$

$E_x D^H (D E_x D^H)^{-1}$  を概算する

$\hat{X}$  パラメータ的な再構成オブジェクト信号、サイズ  $N_{Objects} \times N_{Samples}$

Xを概算し、 $\hat{X} = G Y$  として定義される

30

R レンダリング行列（デコーダ側で明示される）、サイズ  $N_{UpmixCh} \times N_{Objects}$

Z 理想的なレンダリングされた出力シーン信号、サイズ  $N_{UpmixCh} \times N_{Samples}$

$Z = R X$  として定義される

$\hat{Z}$  レンダリングされたパラメータ出力、サイズ  $N_{UpmixCh} \times N_{Samples}$

$\hat{Z} = R \hat{X}$  として定義される

C 理想的な出力の共分散行列、サイズ  $N_{UpmixCh} \times N_{UpmixCh}$

$C = R E_x R^H$  として定義される

W 非相関化出力、サイズ  $N_{UpmixCh} \times N_{Samples}$

40

S 結合信号  $S = \begin{bmatrix} \hat{Z} \\ W \end{bmatrix}$ 、サイズ  $2N_{UpmixCh} \times N_{Samples}$

$E_s$  結合信号共分散行列、サイズ  $2N_{UpmixCh} \times 2N_{UpmixCh}$

$E_s = S S^H$  として定義される

$\tilde{Z}$  最後の出力、サイズ  $N_{UpmixCh} \times N_{Samples}$

$(\cdot)^H$  自己共役（エルミート）作用素

$(\cdot)$  の複素共役転置を表わす。表記  $(\cdot)^*$  も使用されうる。

$F_{decorr}(\cdot)$  非相関化関数

$\varepsilon$  ゼロによる除算を回避するために、（例えば、「`maximum`（最大）」動作または「`max`（最大）」動作において使用される）付加的な定数または制限定数である。10

$H = matdiag(M)$  主対角線における行列  $M$  の主対角線からの要素および非対角線位置におけるゼロ値を含む行列である。

#### 【0105】

一般的な損失なしに、方程式の可読性を改良するために、すべての導入される変数に対して、時間および周波数依存を表示しているインデックスは、この明細書において省略される。20

#### 【0106】

##### 13.2 パラメータ分離システム

一般的なパラメータ分離システムは、（例えば、内部チャネル相関値、内部チャネルレベル差値、内部オブジェクト相関値および / またはオブジェクト・レベル差情報のような）補助的なパラメータ情報を使用して信号混合（ダウンミックス）からオーディオ音源の数を推定することを意図する。この作業の典型的な解決は、最小2乗平均誤差（`Minimum Mean Squared Error` : MMSE）推定アルゴリズムのアプリケーションに基づく。SAOC技術は、パラメトリックオーディオ符号化 / 復号化システムのような1つの実施例である。20

#### 【0107】

図13は、SAOCエンコーダ / デコーダ構造の一般的な原理を示す。換言すれば、図13は、パラメトリックダウンミックス / アップミックスのコンセプトに基づくMMSEの概観を、ブロック概略図の形式で示す。30

#### 【0108】

エンコーダ1310は、複数のオブジェクト信号1312a, 1312b ~ 1312nを受信する。さらに、エンコーダ1310は、例えば、ダウンミックス・パラメータであるミキシング・パラメータD, 1314も受信する。エンコーダ1310は、それに基づいて、1つ以上のダウンミックス信号1316a, 1316b等を供給する。さらに、エンコーダは、サイド情報1318を供給する。1つ以上のダウンミックス信号およびサイド情報は、例えば、符号化形式で供給されうる。

#### 【0109】

エンコーダ1310は、典型的には、オブジェクト信号1312a ~ 1312nを受信し、ミキシング・パラメータ1314に基づいて、オブジェクト信号1312a ~ 1312nを結合して（例えば、ダウンミックスして）1つ以上のダウンミックス信号1316a, 1316bを作成するように構成されるミキサ1320を含む。さらに、エンコーダは、オブジェクト信号1312a ~ 1312nからサイド情報1318を導出するように構成されるサイド情報推定器1330を含む。例えば、サイド情報推定器1330は、サイド情報が、例えば、（「内部オブジェクト相関」（`IQC` : `internal-object-correlation`）として指定されうる）オブジェクト信号の間の相互相関および / または（「オブジェクト・レベル差情報」（`OLD` : `object-level-information`）として指定されうる）オブジェクト信号の間のレベル差を記述している情報である、オブジェクト信号の間の関係を記述するようなサイド情報131840

10

20

30

40

50

を導出するように構成されうる。

【0110】

1つ以上のダウンミック信号 1316a, 1316b およびサイド情報 1318 は、参考番号 1340 に示されるように、デコーダ 1350 に格納されおよび / または送信される。

【0111】

デコーダ 1350 は、(例えば、符号化形式で) 1つ以上のダウンミックス信号 1316a, 1316b およびサイド情報 1318 を受信し、そして、それに基づいて、複数の出力オーディオ信号 1352a ~ 1352n を供給する。デコーダ 1350 は、(レンダリング行列を定義しうる) 1つ以上のレンダリング・パラメータを含みうるユーザ相互作用情報 1354 も受信する。デコーダ 1350 は、パラメトリック・オブジェクト・セパレータ 1360、サイド情報プロセッサ 1370 およびレンダラ 1380 を含む。サイド情報プロセッサ 1370 は、サイド情報 1318 を受信し、それに基づいて、パラメトリック・オブジェクト・セパレータ 1360 に対して制御情報 1372 を供給する。パラメトリック・オブジェクト・セパレータ 1360 は、ダウンミックス信号 1360a, 1360b およびサイド情報プロセッサ 1370 によってサイド情報 1318 から導出された制御情報 1372 に基づいて、複数のオブジェクト信号 1362a ~ 1362n を供給する。例えば、オブジェクト・セパレータは、符号化ダウンミックス信号およびオブジェクト分離の復号化を実行しうる。レンダラ 1380 は、それによって出力オーディオ信号 1352a ~ 1352n を得るために、再構成オブジェクト信号 1362a ~ 1362n をレンダリングする。

【0112】

以下に、パラメータ・ダウンミックス / アップミックスのコンセプトに基づく M M S E の機能が述べられる。

【0113】

一般的なパラメトリックダウンミックス / アップミックス処理は、時間 / 周波数の選択的な方法で実行され、以下のステップのシーケンスとして記載されうる。

【0114】

・「エンコーダ」 1310 は、入力された「オーディオ・オブジェクト」 X および「ミキシング・パラメータ」 D を有する。「ミキサ」 1320 は、「ミキシング・パラメータ」 D (例えば、ダウンミックスゲイン) を使用して「オーディオ・オブジェクト」 X をいくつかの「ダウンミックス信号」 Y にダウンミックスする。「サイド情報推定器」は、入力された「オーディオ・オブジェクト」 X (例えば、共分散特性) の特性を記述しているサイド情報 1318 を抽出する。

【0115】

・「ダウンミックス信号」 Y およびサイド情報は、送信されるか、または格納される。これらのダウンミックスオーディオ信号は、さらに、オーディオコーダ (例えば、M P E G - 1 / 2 L a y e r I I または I I I , M P E G - 2 / 4 A d v a n c e d A u d i o C o d i n g (A A C) , M P E G U n i f i e d S p e e c h a n d A u d i o C o d i n g (U S A C) 等) を使用して圧縮されうる。サイド情報は、(例えば、オブジェクトパワーおよびオブジェクト相関係数の無損失性符号化関係として) 効果的に再構成され、符号化されうる。

【0116】

・「デコーダ」1350は、送信されたサイド情報1318を使用して、復号化された「ダウンミックス信号」から元の「オーディオ・オブジェクト」を復元する。「サイド情報プロセッサ」1370は、Xのパラメトリック・オブジェクト再構成を得るために、「パラメトリック・オブジェクト・セパレータ」1360内における「ダウンミックス信号」に適用されるアップミックス係数1372を評価する。再構成された「オーディオ・オブジェクト」1362a～1362nは、「レンダリング・パラメータ」R, 1354を適用されることによって、(マルチチャネルの)ターゲット・シーンにレンダリングされ、出力チャネル $\hat{Z}$ によって表わされる。

10

## 【0117】

さらに、エンコーダ1310およびデコーダ1350に関して記載されている機能は、同様に本願明細書において記載されている他のオーディオ・エンコーダおよびオーディオ・デコーダにおいて使用される点に留意されたい。

## 【0118】

## 13.3 最小2乗平均誤差推定の直交原理

直交原理は、MMSE推定の1つの主な特性である。ベクトル $y_i$ およびベクトル $x \in W$ のセットによって範囲が特定されるVを有する2つのヒルベルト空間WおよびVを考慮する。平均2乗誤差を最小化するうえで、ベクトル $y_i \in V$ の一次結合としてxを近似する推定 $\hat{x} \in V$ を見つけることを望む場合、そのとき、誤りベクトルは、ベクトル $y_i$ によって範囲が特定される空間において、直交する。

20

$$(x - \hat{x})y^H = 0$$

## 【0119】

結果として、推定誤差および推定自身は、直交する。

$$(x - \hat{x})\hat{x}^H = 0$$

30

## 【0120】

幾何学的に、図14に示される例によって、これを視覚化しうる。

## 【0121】

図14は、3次元空間における直角原理のための幾何学的な表現を示す。示されるように、ベクトル空間は、ベクトル $y_1, y_2$ によって範囲が特定される。ベクトル $x$ は、ベクトル $\hat{x}$ と差分ベクトル(誤差ベクトル) $e$ の合計に等しい。示されるように、誤差ベクトル $e$ は、ベクトル $y_1$ および $y_2$ によって範囲が特定されるベクトル空間(または平面)に直交する。従って、ベクトル $\hat{x}$ は、ベクトル空間V内において $x$ に最も近似するよう考へられる。

40

## 【0122】

### 13.4 パラメトリック再構成誤差

N個の信号を含む行列を定義する：Xおよび $\mathbf{X}_{Error}$ を有する推定誤差を示し、以下の同一性が定式化されうる。元の信号は、

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{X}_{Error}$$

に示すように、パラメトリック再構成 $\hat{\mathbf{X}}$ および再構成誤差 $\mathbf{X}_{Error}$ の合計として表わされうる。

【0123】

直角原理のため、元の信号 $\mathbf{E}_X = \mathbf{XX}^H$ の共分散行列は、

$$\begin{aligned}\mathbf{E}_X &= \mathbf{XX}^H = (\hat{\mathbf{X}} + \mathbf{X}_{Error})(\hat{\mathbf{X}} + \mathbf{X}_{Error})^H = \hat{\mathbf{X}}\hat{\mathbf{X}}^H + \mathbf{X}_{Error}\mathbf{X}_{Error}^H + \hat{\mathbf{X}}\mathbf{X}_{Error}^H + \mathbf{X}_{Error}\hat{\mathbf{X}}^H = \\ &= \hat{\mathbf{X}}\hat{\mathbf{X}}^H + \mathbf{X}_{Error}\mathbf{X}_{Error}^H.\end{aligned}$$

に示すように、再構成信号 $\hat{\mathbf{X}}\hat{\mathbf{X}}^H$ と推定誤差の共分散行列 $\mathbf{X}_{Error}\mathbf{X}_{Error}^H$ の合計として定式化されうる。

【0124】

入力オブジェクトXがダウンミックスチャネル（例えば、ダウンミックスチャネルの数が、入力信号の数より少ない）によって測られる空間においてではなく、そして、入力オブジェクトが、ダウンミックスチャネルの一次結合として表わすことができない場合、MSE規範のアルゴリズムは、再構成の不正確さ $\mathbf{X}_{Error}\mathbf{X}_{Error}^H$ を導く。

【0125】

### 13.5 内部オブジェクト相関

聴覚系において、相互共分散（均一性／相関）は、音によって囲まれるように包まれる知覚、および音源の知覚された幅に密接に関係する。例えば、S A O C 規範のシステムにおいて、内部オブジェクト相関（I O C）パラメータは、この特性の特性評価のために使用される。

$$IOC(i,j) = \frac{\mathbf{E}_X(i,j)}{\sqrt{\mathbf{E}_X(i,i)\mathbf{E}_X(j,j)}}$$

【0126】

我々に、2つのオーディオ信号を使用して音源を再生する実施例を考慮させる。I O C 値が1に近い場合、音は、極めて局所化された点の音源として知覚される。I O C 値がゼロに近い場合、音源の知覚された幅は増加し、そして、極端な例では、2つの異なった音源としてとして知覚されることさえある [ Blaueurt, 第3章]。

【0127】

### 13.6 再構成誤りの補償

不完全なパラメトリック再構成の場合、出力信号は、元のオブジェクトと比較して低いエネルギーを示しうる。共分散行列の対角要素の誤差は、（理想的な参照出力と比較して）聞き取れるレベル差、および歪められた空間音像における非対角要素での誤差を結果と

して得る。提案された方法は、この課題を解決するための目的を有する。

#### 【0128】

M P E G S u r r o u n d ( M P S ) において、例えば、この問題は、いくつかの特定のチャネル・ベースの処理シナリオ、すなわち、モノラル / ステレオダウンミックスおよび限られた静的な出力構造（例えば、モノラル、ステレオ、5 . 1 , 7 . 1 等）ためだけに扱われる。S A O C のようにモノラル / ステレオダウンミックスも使用するオブジェクト指向技術において、この課題は、5 . 1 の出力構造だけのための M P S 後処理レンダリングを適用することによって扱われる。

#### 【0129】

既存の解決策は、標準の出力構成および入出力チャネルの定数に限られる。すなわち、それらは、ちょうど「モノラル対ステレオ」（または「ステレオ対3チャネル」）のチャネル非相関を実装しているいくつかのブロックの結果として生じるアプリケーションとして理解される。

10

#### 【0130】

それゆえに、パラメトリック再構成誤りの補償のための一般的な解決策（例えば、エネルギーレベルおよび相関特性修正方法）が所望され、そして、それは柔軟な数のダウンミックス / 出力チャネルおよび任意の出力構成セットアップに対して適用されうる。

20

#### 【0131】

##### 13 . 7 結論

結論として、表記法に関する概要が提供された。さらに、パラメトリック分離システムは、本発明による実施の形態がベースであることが述べられた。さらに、最小2乗平均誤差推定に適用される直交原理が概説された。さらに、再構成誤差  $X_{\text{Error}}$  の存在において適用する共分散行列  $E_x$  の算出のための方程式が提供された。また、例えば、本発明による実施の形態において、（パラメトリックサイド情報において含まれうる）内部オブジェクト相関値から所望の共分散特性（または相関特性）を導出し、そして、おそらくオブジェクト・レベル差を形成するために適用されうる、いわゆる内部オブジェクト相関および共分散行列  $E_x$  の要素との間の関係が提供された。さらに、再構成オブジェクト信号の特性が、不完全な再構成のため、所望の特性と異なることが概説された。さらに、課題を取り扱う既存の解決策が、いくつかの特定の出力構成に限られており、従来の解決策の変更できない標準ブロックの特定の結合に依拠することが概説された。

30

#### 【0132】

##### 14 . 図15による実施の形態

###### 14 . 1 コンセプトの概要

本発明による実施の形態は、任意の数のダウンミックス / アップミックスチャネルのための非相関化の解決策を有するパラメトリックオーディオ分離方式において使用される M M S E パラメトリック再構成方法を拡張する。例えば、発明の装置および発明の方法のような本発明による実施の形態は、パラメトリック再構成の間のエネルギー損失を補償することができ、そして、推定されたオブジェクトの相関特性を復元しうる。

40

#### 【0133】

図15は、統合された非相関化経路を有するパラメトリックダウンミックス / アップミックスのコンセプトの概要を提供する。換言すれば、ブロック概略図の形式において、レンダリングされた出力に適用される非相関化を伴うパラメトリック再構成システムを示す。

50

#### 【0134】

図15に記載のシステムは、図13に記載のエンコーダ1310と実質的に同一であるエンコーダ1510を含む。エンコーダ1510は、複数のオブジェクト信号1512a ~ 1512nを受信し、そして、それに基づいて、1つ以上のダウンミックス信号1516a, 1516bおよびサイド情報1518を供給する。ダウンミックス信号1516a, 1516bは、ダウンミックス信号1316a, 1316bと実質的に同一でありえ、そして、Yによって指定される。サイド情報1518は、サイド情報1318と実質的に

50

同一でありうる。しかしながら、例えば、サイド情報は、非相関化モードパラメータ、または非相関化方法パラメータ、または非相関化複雑さパラメータを含む。さらに、エンコーダ 1510 は、ミキシング・パラメータ 1514 を受信しうる。

#### 【0135】

パラメトリック再構成システムは、1つ以上のダウンミックス信号 1516a, 1516b およびサイド情報 1518 の送信および / または格納も含む。ここで、送信および / または格納は、1540 で指定され、1つ以上のダウンミックス信号 1516a, 1516b および (パラメトリックサイド情報を含みうる) サイド情報 1518 が、符号化される。

#### 【0136】

さらに、図 15 によるパラメトリック再構成システムは、送信されまたは格納された 1 つ以上の (あるいは符号化) ダウンミックス信号 1516a, 1516b および送信されまたは格納された (あるいは符号化) サイド情報 1518 を受信し、そして、それに基づいて、出力オーディオ信号 1552a ~ 1552n を供給するように構成される、デコーダ 1550 を含む。(マルチチャネル・オーディオ・デコーダとして考慮されうる) デコーダ 1550 は、パラメトリック・オーディオ・セパレータ 1560 およびサイド情報プロセッサ 1570 を含む。さらに、デコーダ 1550 は、レンダラ 1580、非相関器 1590 およびミキサ 1598 を含む。

#### 【0137】

パラメトリック・オブジェクト・セパレータ 1560 は、1つ以上のダウンミックス信号 1516a, 1516b、およびサイド情報 1518 に基づいてサイド情報プロセッサ 1570 によって供給される制御情報 1572 を受信し、それに基づいて、復号化オーディオ信号として考慮されえ、 $\hat{X}$  でも指定されるオブジェクト信号 1562a ~ 1562n を供給するように構成される。制御情報 1572 は、再構成オブジェクト信号 (例えば、復号化オーディオ信号 1562a ~ 1562n) を得るためのパラメトリック・オブジェクト・セパレータ内において、ダウンミックス信号 (例えば、符号化ダウンミックス信号 1516a, 1516b から導出された復号化ダウンミックス信号) を供給するためのアップミックス係数を、例えば、含む。レンダラ 1580 は、それによって、複数のレンダリングされたオーディオ信号 1582a ~ 1582n を得るために、(再構成オブジェクト信号であり、そして、例えば、入力オブジェクト信号 1512a ~ 1512n に対応する) 復号化オーディオ信号 1562a ~ 1562n をレンダリングする。例えば、レンダラ 1580 は、例えば、ユーザ相互作用によって供給され、そして、例えば、レンダリング行列として定義されうるレンダリング・パラメータ R を考慮しうる。しかしながら、あるいは、レンダリング・パラメータは、(符号化ダウンミックス信号 1516a, 1516b および符号化サイド情報 1518 を含みうる) 符号化表現から引き出される。

#### 【0138】

非相関器 1590 は、レンダリングされたオーディオ信号 1582a ~ 1582n を受信し、そして、それに基づいて、W でも指定される非相関化オーディオ信号 1592a ~ 1592n を供給する。ミキサ 1598 は、レンダリングされたオーディオ信号 1582a ~ 1582n および非相関化オーディオ信号 1592a ~ 1592n を受信し、そして、レンダリングされたオーディオ信号 1582a ~ 1582n と非相関化オーディオ信号 1592a ~ 1592n とを結合し、それによって、出力オーディオ信号 1552a ~ 1552n を得る。ミキサ 1598 は、後述するように、符号化サイド情報 1518 からサイド情報プロセッサ 1570 によって導出される制御情報 1574 も使用しうる。

#### 【0139】

##### 14.2 非相関器の関数

以下に、非相関器 1590 に関する若干の詳細が記載される。しかしながら、いくつか

10

20

30

40

50

後述されるように、異なる非相関器が使用されうる点に留意されたい。

【0140】

実施の形態において、非相関器の関数  $w = F_{decorr}(\hat{z})$  は、入力信号  $\hat{Z}$  ( $E\{w\hat{z}^H\}=0$ ) に直交する出力信号  $w$  を供給する。出力信号  $w$  は、(入力信号  $\hat{Z}$  に) スペクトルおよび時間的エンベロープ特性 (または、少なくとも類似の特性) に等しい。さらに、信号  $w$  は、同様に知覚され、そして、入力信号  $\hat{Z}$  のように同じ (または類似的な) 主観的品質を有する (例えば、[S A O C 2] を参照)。

10

【0141】

複数の入力信号の場合には、非相関化の関数が、(すべての  $i$  および  $j$  に対して、 $W_i \hat{Z}_j^H = 0$  であり、そして、 $i \neq j$  に対して、 $W_i W_j^H = 0$  であるような、例えば、 $W_i = F_{decorr}(\hat{Z}_i)$ ) 相互に直交する複数の出力を生じる場合、有益である。

【0142】

非相関器の関数の実装のための正確な仕様は、この説明の範囲の外である。例えば、MPEG Surround Standardにおいて特定された非相関器に基づく、いくつかの無限インパルス応答(IIR)フィルタのバンクは、非相関化の目的のために利用されうる([MPS])。

20

【0143】

この説明において記載される一般的な非相関器は、理想的であるとみなされる。これは、(知覚的な要件に加えて) 各非相関器の出力が、すべての他の非相関器のその入力および出力において直交することを意味する。それゆえ、共分散  $\mathbf{E}_{\hat{z}} = \hat{Z}\hat{Z}^H$  を伴う所与の入力  $\hat{Z}$  および出力  $\mathbf{W} = F_{decorr}(\hat{Z})$  に対して、共分散行列の特性が以下の式を保持する。

30

$$\mathbf{E}_W(i,i) = \mathbf{E}_{\hat{z}}(i,i), \mathbf{E}_W(i,j) = 0, \text{for } i \neq j, \hat{Z}\mathbf{W}^H = \mathbf{W}\hat{Z}^H = \mathbf{0}.$$

【0144】

これらの関係から、

$$(\hat{Z} + \mathbf{W})(\hat{Z} + \mathbf{W})^H = \mathbf{E}_{\hat{z}} + \hat{Z}\mathbf{W}^H + \mathbf{W}\hat{Z}^H + \mathbf{E}_W = \mathbf{E}_{\hat{z}} + \mathbf{E}_W$$

を結果として得る。

【0145】

非相関器出力  $\mathbf{W}$  は、入力として予測された信号を使用することによって、(予測誤差が予測信号に対して直交することを記憶している) M M S E 推定器における予測誤りを補償するように使用されうる。

40

【0146】

それは、予測誤差がそれら自身の間において直交する一般的な場合ではない点に留意されたい。このように、結果として得られるミクスチャーの共分散行列 (例えば、出力オーディオ信号 1 5 5 2 a ~ 1 5 5 2 n) が所望の出力の共分散行列に類似することになるように、本発明のコンセプト (例えば、方法) の目的 1 つは、「ドライ」(すなわち、非相関器入力) 信号 (例えば、レンダリングされたオーディオ信号 1 5 8 2 a ~ 1 5 8 2 n) および「ウェット」(すなわち、非相関器出力) 信号 (例えば、非相関化オーディオ信号 1 5 9 2 a ~ 1 5 9 2 n) のミクスチャーを作成することである。

50

【0147】

さらに、以下に詳細に記載され、そして、しかしながら、受け入れられる、非相関化信号のいくつかの欠点をもたらす、非相関化装置のための複雑さの低減が使用される点に留意されたい。

【0148】

14.3 非相関化信号を使用している出力共分散修正

以下に、コンセプトが、合理的に良好な聴覚印象を得るように、出力オーディオ信号1552a～1552nの共分散特性を調整することを記載している。

【0149】

出力共分散の誤差修正のための提案された方法は、パラメータ的に再構成された信号 $\hat{\mathbf{Z}}$   
10  
 (例えば、レンダリングされたオーディオ信号1582a～1582n)とその非相関化部分Wの加重和として、出力信号 $\tilde{\mathbf{Z}}$ (例えば、出力オーディオ信号1552a～1552n)を構成する。この合計は、以下のように表されうる。

$$\tilde{\mathbf{Z}} = \mathbf{P}\hat{\mathbf{Z}} + \mathbf{M}\mathbf{W}$$

【0150】

しかしながら、この方程式は、もっとも一般的な定式化と考えられる点に留意されたい。  
 変化は、本願明細書において記載されているすべての「簡略化された方法」に通用する  
20  
 (または、させ得る)上記の公式に、任意に適用されうる。

【0151】

混合行列Pは、直接信号 $\hat{\mathbf{Z}}$ に適用され、Mは、( $N=N_{UpmixCh}$ を伴い、ここで、 $N_{UpmixCh}$ は、出力オーディオ信号の数に等しいレンダリングされたオーディオ信号の数を示す)  
 )以下の構造を有する非相関化信号Wに適用される。

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,N} \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,N} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N,1} & m_{N,2} & \cdots & m_{N,N} \end{bmatrix}$$
30

【0152】

結合行列 $\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$ および信号 $\mathbf{S} = \begin{bmatrix} \hat{\mathbf{Z}} \\ \mathbf{W} \end{bmatrix}$ のための表記表を適用して、

$$\tilde{\mathbf{Z}} = \mathbf{FS}$$
40

を得る。

【0153】

あるいは、一方、後で詳しく述べるように、以下の方程式

$$\tilde{\mathbf{Z}} = \tilde{\mathbf{FS}}$$

が、適用されうる。

【0154】

この表現を使用して、出力信号  $\tilde{\mathbf{Z}}$  の共分散行列  $\mathbf{E}_{\tilde{z}}$  は、

$$\mathbf{E}_{\tilde{z}} = \mathbf{F} \mathbf{E}_s \mathbf{F}^H$$

として、定義される。

【 0 1 5 5 】

理想的に作成されレンダリングされた出力シーンのターゲット共分散  $C$  は、

10

$$C = R E_x R^H$$

として、定義される。

【 0 1 5 6 】

最終出力の共分散行列  $\mathbf{E}_{\tilde{z}}$  が、

$$\mathbf{E}_{\tilde{z}} \approx \mathbf{C}$$

20

として示される、ターゲット共分散  $C$  に近似するか、等しいように、混合行列  $F$  は計算される。

【 0 1 5 7 】

例えば、混合行列  $F$  は、

$$\mathbf{F} = (\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) \mathbf{H} (\mathbf{V} \sqrt{\mathbf{Q}^{-1}} \mathbf{V}^H)$$

30

として、既知量の関数  $\mathbf{F} = \mathbf{F}(\mathbf{E}_s, \mathbf{E}_x, \mathbf{R})$  として計算される。

ここで、行列  $U$ 、 $T$ 、 $V$ 、 $Q$  は、例えば、共分散行列の特異値分解 (Singular Value Decomposition : SVD) を使用して決定され、 $E_s$  および  $C$  は、

$$\mathbf{C} = \mathbf{U} \mathbf{T} \mathbf{U}^H, \quad \mathbf{E}_s = \mathbf{V} \mathbf{Q} \mathbf{V}^H$$

として得られる。

【 0 1 5 8 】

原型の行列  $H$  は、直接的なおよび非相関信号パスのために所望の加重に従って選択される。

40

【 0 1 5 9 】

例えば、可能な原型行列  $H$  は、

$$\mathbf{H} = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 & b_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 & 0 & b_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{N,N} & 0 & 0 & \dots & b_{N,N} \end{bmatrix}, \text{ここで、 } a_{i,i}^2 + b_{i,i}^2 = 1$$

として、決定されうる。

【0 1 6 0】

以下に、一般の行列  $\mathbf{F}$  の構造のための若干の数学的な導出が提供される。 10

【0 1 6 1】

換言すれば、一般的な解決法のための混合行列  $\mathbf{F}$  の導出は、以下において記載される。

【0 1 6 2】

共分散行列  $\mathbf{E}_s$  および  $\mathbf{C}$  は、

$$\mathbf{E}_s = \mathbf{V} \mathbf{Q} \mathbf{V}^H, \quad \mathbf{C} = \mathbf{U} \mathbf{T} \mathbf{U}^H$$

として、例えば、特異値分解 ( $S V D$ ) を使用して表わされ、 $\mathbf{T}$  および  $\mathbf{Q}$  は、それぞれ  $\mathbf{C}$  および  $\mathbf{E}_s$  の特異値を有する対角行列であり、 $\mathbf{U}$  および  $\mathbf{V}$  は、特異ベクトルに対応するユニタリ行列である。 20

【0 1 6 3】

( $S V D$  の代わりに) シューア三角化または固有値分解のアプリケーションが、類似の結果 (または、対角行列  $\mathbf{Q}$  および  $\mathbf{T}$  が正值に限定される場合、同一の結果) に至ることに留意されたい。

【0 1 6 4】

前提条件  $\mathbf{E}_z \approx \mathbf{C}$  にこの分解を適用して、(少なくとも、近似的に)

$$\mathbf{C} = \mathbf{F} \mathbf{E}_s \mathbf{F}^H,$$

30

$$\mathbf{U} \mathbf{T} \mathbf{U}^H = \mathbf{F} \mathbf{V} \mathbf{Q} \mathbf{V}^H \mathbf{F}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) = \mathbf{F} (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) \mathbf{F}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) = (\mathbf{F} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) \mathbf{F}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)^H = (\mathbf{F} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{F} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H)^H$$

が、得られる。

40

【0 1 6 5】

共分散行列の次元について解決するために、正則化が場合によっては必要である。例えば、 $\mathbf{HH}^H = \mathbf{I}_{N_{UpmixCh}}$  の特性を有するサイズ  $N_{UpmixCh} \times 2N_{UpmixCh}$  の原型行列  $\mathbf{H}$  は、

$$(\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)\mathbf{HH}^H(\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H) = \mathbf{F}(\mathbf{V}\sqrt{\mathbf{Q}}\mathbf{V}^H)(\mathbf{V}\sqrt{\mathbf{Q}}\mathbf{V}^H)\mathbf{F}^H,$$

$$(\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)\mathbf{H} = \mathbf{F}(\mathbf{V}\sqrt{\mathbf{Q}}\mathbf{V}^H)$$

を適用されうる。

10

## 【0 1 6 6】

混合行列  $\mathbf{F}$  は、

$$\mathbf{F} = (\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)\mathbf{H}(\mathbf{V}\sqrt{\mathbf{Q}^{-1}}\mathbf{V}^H)$$

として決定されうるということになる。

20

## 【0 1 6 7】

原型行列  $\mathbf{H}$  は、直接的なおよび非相関化信号パスのための所望の加重によって選択される。例えば、可能な原型行列  $\mathbf{H}$  は、

$$\mathbf{H} = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 & b_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 & 0 & b_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{N,N} & 0 & 0 & \dots & b_{N,N} \end{bmatrix}, \text{ここで、 } a_{i,i}^2 + b_{i,i}^2 = 1$$

として、決定されうる。

30

## 【0 1 6 8】

結合信号の共分散行列  $\mathbf{E}_s$  の条件に応じて、最後の方程式は、いくつかの正則化を含むことを必要としうるが、さもなければ、数値的に安定でなければならない。

## 【0 1 6 9】

結論として、コンセプトは、（マトリックスW、同等地に、ベクトルwによって表わされる）レンダリングされたオーディオ信号に基づいて、（マトリックス $\tilde{Z}$ 、同等地に、ベクトル $\tilde{Z}$ によって表わされる）出力オーディオ信号を導出することが記載されていた。参考されるように、一般的な行列構造の2つの混合行列PおよびMが、通常、決定される。例えば、出力オーディオ信号 $1\ 5\ 5\ 2\ a \sim 1\ 5\ 5\ 2\ n$ の共分散行列 $E_z$ が、所望の（ターゲット共分散としても示される）共分散Cに近似するか、または等しいように、上記で定義されるように、結合行列Fが決定されうる。（例えば、ユーザ相互作用によって供給されうる）レンダリング行列Rの情報に基づいて、および例えば、符号化サイド情報1518に基づいて導出されうるオブジェクト共分散行列 $E_x$ の情報に基づいて、所望の共分散行列Cは、例えば、導出されうる。例えば、オブジェクト共分散行列 $E_x$ は、上記された内部オブジェクト共分散値IOCを使用して導出され、そして、それは、符号化サイド情報1518において含まれうる。このように、例えば、ターゲット共分散行列Cは、情報1574または情報1574の部分として、サイド情報プロセッサ1570によって供給されうる。

10

## 【0170】

しかしながら、あるいは、サイド情報プロセッサ1570は、情報1574として、ミキサ1598に直接的に混合行列Fを供給も、しうる。

20

## 【0171】

さらに、混合行列Fのための計算規則は、特異値分解を使用することを記載されていた。しかしながら、原型行列Hの入力 $a_{i,i}$ および $b_{i,i}$ が選択されうるので、ある自由度がある点に留意されたい。好ましくは、原型行列Hの入力は、およそ0および1の間で選択される。非相関化オーディオ信号のインパクトは、比較的小さく、そして、それは、若干の状況において望ましくもある一方、値 $a_{i,i}$ が1に近づくように選択されうる場合、レンダリングされた出力オーディオ信号の重要な混合がある。しかしながら、レンダリングされたオーディオ信号の間ににおける弱い混合がある一方、若干の他の状況において、非相関化オーディオ信号の比較的大きいインパクトを有することをより望まれうる。この場合、値 $b_{i,i}$ は、 $a_{i,i}$ より大きいように典型的に選択される。このように、デコーダ1550は、原型行列Hの入力を適切に選択することによって、前提条件に適応されうる。

30

## 【0172】

## 14.4 出力共分散修正のための簡略化された方法

このセクションにおいて、前述の混合行列Fのための2つの変形例の構造は、その値を決定するための典型的なアルゴリズムとともに記載される。2つの変形例は、異なる入力コンテンツ（例えば、オーディオ・コンテンツ）のために設計される。

40

- 高い相関コンテンツ（例えば、異なるチャネル対の間の高い相関を伴うチャネル・ベースの入力）のための共分散調整方法

- 独立入力信号（例えば、たいてい独立とみなされる、オブジェクト・ベースの入力）のためのエネルギー補償方法

## 【0173】

#### 14.4.1 共分散調整方法 (A)

信号  $\hat{\mathbf{Z}}$  (例えば、レンダリングされたオーディオ信号 1582a ~ 1582n) が既に MMSE の趣旨において既に最適であることを考慮に入れて、これが、分離特性に影響を及ぼしうるので、出力  $\tilde{\mathbf{Z}}$  の共分散特性を改良するために、パラメトリック再構成 (例えば、出力オーディオ信号 1552a ~ 1552n) を修正することは、通常、望ましくない。

##### 【0174】

非相関化信号  $\mathbf{W}$  の混合が操作されれば、混合行列  $\mathbf{P}$  は、単位行列 (または、その倍数) 10 に低減されうる。このように、この簡略化された方法は、

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,N} \\ m_{2,1} & m_{2,2} & \dots & m_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N,1} & m_{N,2} & \dots & m_{N,N} \end{bmatrix}$$

による設定によって、記載されうる。

##### 【0175】

システムの最後の出力は、

$$\tilde{\mathbf{Z}} = \hat{\mathbf{Z}} + \mathbf{M}\mathbf{W}$$

として、表わされうる。

##### 【0176】

従って、システムの最後の出力共分散は、

$$\mathbf{E}_{\tilde{Z}} = \mathbf{E}_{\hat{Z}} + \mathbf{M}\mathbf{E}_w\mathbf{M}^H$$

として、表わされうる。

##### 【0177】

理想的 (または所望) の出力共分散行列  $\mathbf{C}$  とレンダリングされたパラメトリック再構成 (例えば、レンダリングされたオーディオ信号) の共分散行列  $\mathbf{E}_{\hat{Z}}$  との差  $\Delta_E$  は、

$$\Delta_E = \mathbf{C} - \mathbf{E}_{\hat{Z}}$$

によって、与えられる。

##### 【0178】

従って、混合行列  $\mathbf{M}$  は、

$$\Delta_E \approx \mathbf{M}\mathbf{E}_w\mathbf{M}^H$$

のように、決定される。

##### 【0179】

混合非相関化信号MWの共分散行列は、所望の共分散とドライ信号（例えば、レンダリングされたオーディオ信号の）の共分散との間の共分散差に等しいかまたは近似するよう に、混合行列Mは算出される。従って、最後の出力の共分散は、ターゲット共分散  $E_z \approx C$  に近似する。

$$M = (U\sqrt{T}U^H)(V\sqrt{Q^{-1}}V^H)$$

10

ここで、行列U、T、V、Qは、例えば、共分散行列の特異値分解（S i n g u l a r V a l u e D e c o m p o s i t i o n : S D V ）を使用して決定され、 $\Delta_E$ および $E_w$ は、

$$\Delta_E = U T U^H, \quad E_w = V Q V^H$$

で得られる。

#### 【 0 1 8 0 】

このアプローチは、ドライ出力（例えば、レンダリングされたオーディオ信号 1 5 8 2 a ~ 1 5 8 2 n ）を使用する良好な相互相關再構成の最大化を確実にし、そして、非相関化信号のみの混合の自由を利用する。換言すれば、レンダリングされたオーディオ信号（または、そのスケール化バージョン）と 1 つ以上の非相関化オーディオ信号とを結合する場合、異なるレンダリングされたオーディオ信号の間の混合は許容されない。しかしながら、出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号は、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンを結合することを許容する。ここで、定義されるように、例えば、結合は、行列Mによって定義される。

#### 【 0 1 8 1 】

以下に、制限行列F構造のためのいくつかの数学的な導出が提供される。

20

#### 【 0 1 8 2 】

換言すれば、簡略化された方法「A」のための混合行列Mの導出が説明される。

30

#### 【 0 1 8 3 】

共分散行列  $E_E$  および  $E_w$  は、

$$E_E = U T U^H, \quad E_w = V Q V^H$$

として、例えば、特異値分解（S V D ）を使用して表わされ、TおよびQは、それぞれ  $E_E$  および  $E_w$  の特異値を有する対角行列であり、UおよびVは、対応する特異ベクトルを含むユニタリ行列である。

40

#### 【 0 1 8 4 】

（S V D の代わりに）シューア三角化または固有値分解のアプリケーションが、類似の結果（または、対角行列QおよびTが正值に限定される場合、同一の結果）に至ることに留意されたい。

#### 【 0 1 8 5 】

前提条件  $\mathbf{E}_z \approx \mathbf{C}$  にこの分解を適用して、(少なくとも、近似的に)

$$\Delta_E = \mathbf{M} \mathbf{E}_w \mathbf{M}^H,$$

$$\mathbf{U} \mathbf{T} \mathbf{U}^H = \mathbf{M} \mathbf{V} \mathbf{Q} \mathbf{V}^H \mathbf{M}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) = \mathbf{M} (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) \mathbf{M}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) = (\mathbf{M} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) \mathbf{M}^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H)^H = (\mathbf{M} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H) (\mathbf{M} \mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H)^H,$$

$$(\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) = \mathbf{M} (\mathbf{V} \sqrt{\mathbf{Q}} \mathbf{V}^H)$$

が、得られる。

#### 【0186】

方程式の両側が行列の正方形を表す点に留意して、我々は積算を落とし、全行列  $\mathbf{M}$  のために解決する。

#### 【0187】

混合行列  $\mathbf{M}$  は、

$$\mathbf{M} = (\mathbf{U} \sqrt{\mathbf{T}} \mathbf{U}^H) (\mathbf{V} \sqrt{\mathbf{Q}^{-1}} \mathbf{V}^H)$$

として、決定されうることになる。

#### 【0188】

この方法は、

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 \end{bmatrix}$$

として理解される原型行列  $\mathbf{H}$  を設定することによって、一般的な方法から導出されうる。

#### 【0189】

ウェット信号の共分散行列  $\mathbf{E}_w$  の条件に応じて、最後の方程式は、いくつかの正則化を含むことを必要としうるが、さもなければ、数値的に安定でなければならない。

#### 【0190】

##### 14.4.2 エネルギー補償方法 (B)

時々（アプリケーション・シナリオに応じて）、（例えば、レンダリングされたオーディオ信号の）パラメトリック再構成または非相関化信号の混合を許容することを望まないが、しかし、個々に、各パラメータ的に再構成された信号（例えば、レンダリングされたオーディオ信号）とそれ自身の非相関化信号のみを混合することが望ましい。

#### 【0191】

この前提条件を達成するために、追加の制約は、簡略化方法「A」にもたらされうる。現在、ウェット信号（非相関化信号）の混合行列  $\mathbf{M}$  が、対角形式を有することを必要とする。

10

20

30

40

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_{1,1} & 0 & \dots & 0 \\ 0 & m_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_{N,N} \end{bmatrix}$$

## 【 0 1 9 2 】

このアプローチの主な目標は、パラメトリック再構成（例えば、レンダリングされたオーディオ信号）におけるエネルギーの損失を補償するように、非相関化信号を使用することになっている一方、出力信号の共分散行列の対角線の修正が無視される、すなわち、相互相關の直接的な取り扱いはない。従って、出力オブジェクト／チャネルの間（例えば、レンダリングされたオーディオ信号の間）のクロスリークは、非相関化信号のアプリケーションにおいて導かれない。

10

## 【 0 1 9 3 】

その結果、ターゲット共分散行列（または所望の共分散行列）の主対角線のみに達し、非対角は、パラメトリック再構成および付加的な非相関化信号の精度の軽減にある。この方法は、信号が相関関係のないように考慮されるオブジェクトのみのベースのアプリケーションに最も適している。

## 【 0 1 9 4 】

再構成された信号  $\mathbf{E}_{\hat{z}}(i,i)$  のエネルギーに対する共分散行列の入力が所望のエネルギーの

20

$$\mathbf{E}_{\hat{z}}(i,i) = \mathbf{C}(i,i)$$

と等しくなるように、方法の最終出力（例えば、出力オーディオ信号）は、算出された対角行列  $\mathbf{M}$  を伴う  $\tilde{\mathbf{Z}} = \hat{\mathbf{Z}} + \mathbf{M}\mathbf{W}$  によって与えられる。

## 【 0 1 9 5 】

30

$\mathbf{C}$  は、一般的のケースのために前述したように決定される。

## 【 0 1 9 6 】

例えば、混合行列  $\mathbf{M}$  は、（オーディオ・デコーダによって決定される）非相関化信号のエネルギーを伴う補償信号の所望のエネルギー（（相互共分散行列  $\mathbf{C}$  の対角要素によって記述される）所望のエネルギーと（オーディオ・デコーダによって決定される）パラメトリック再構成のエネルギーとの間の差）を分割することによって直接的に導出される。

$$\mathbf{M}(i,j) = \begin{cases} \sqrt{\min\left(\lambda_{Dec}, \max\left(0, \frac{\mathbf{C}(i,i) - \mathbf{E}_{\hat{z}}(i,i)}{\max(\mathbf{E}_w(i,i), \varepsilon)}\right)\right)} & i = j, \\ 0 & i \neq j. \end{cases}$$

40

ここで、 $\lambda_{Dec}$  は、出力信号（例えば、 $\lambda_{Dec} = 4$ ）に加えられる非相関化の構成要素の量を制限するために用いられる 0 以上の閾値である。

## 【 0 1 9 7 】

エネルギーは、（例えば、O L D, I O C、およびレンダリング係数を使用して）パラメタ的に再構成されると、または、（典型的に、より計算的に価値のある）デコーダによって実質的に算出されると留意されたい。

## 【 0 1 9 8 】

この方法は、

50

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 \end{bmatrix}$$

に示されるような原型行列  $H$  を設定することによって一般的な方法から導出されうる。

**【0199】**

この方法は、明確に、ドライのレンダリングされた出力の使用を最大にする。共分散行列が非対角の入力を有さない場合、方法は、単純化の「A」を伴うものと等価である。10

**【0200】**

この方法は、低減された計算量の複雑さを有する。

**【0201】**

しかしながら、エネルギー補償方法は、相互通関条件が修正されないことを必ずしも意味するものではない点に留意されたい。我々が、理想的な非相關器および非相關化装置のための複雑さの低減でないことを使用する場合だけ、これは保持する。方法の意図は、エネルギーを回復して、交差項における修正を無視することである（交差項における変化は、相関特性を実質的に修正せず、全体の空間印象に影響を及ぼさない）。

**【0202】**

14.5 混合行列  $F$  の前提条件

以下に、混合行列  $F$ 、セクション 14.3 および 14.4 において記載されているその導出は、低下を回避する必要性を満たすと説明される。

**【0203】**

出力における低下を回避するために、パラメトリック再構成誤差のための補償をするいかなる方法が、以下の特性を有する結果をもたらさなければならない。レンダリング行列が、ダウンミックス行列に等しい場合、そのとき、出力チャネルは、ダウンミックスチャネルと等しい（または、少なくとも近似する）べきである。提案されたモデルは、この特性を満足させる。レンダリング行列は、ダウンミックス行列と等しい場合  $R = D$ 、パラメトリック再構成は、30

$$\hat{\mathbf{Z}} = \mathbf{R}\hat{\mathbf{X}} = \mathbf{D}\hat{\mathbf{X}} = \mathbf{D}\mathbf{G}\mathbf{Y} = \mathbf{D}\mathbf{E}\mathbf{D}^H (\mathbf{D}\mathbf{E}\mathbf{D}^H)^{-1}\mathbf{Y} \approx \mathbf{Y}$$

によって、与えられ、そして、所望の共分散行列は、

$$\mathbf{C} = \mathbf{R} \mathbf{E}_x \mathbf{R}^H = \mathbf{D} \mathbf{E}_x \mathbf{D}^H = \mathbf{E}_y$$

である。

**【0204】**

従って、混合行列  $\mathbf{F}$  を得るために解析される方程式は、

$$\mathbf{E}_Y = \mathbf{F} \begin{bmatrix} \mathbf{E}_r & \mathbf{0}_{N_{UpmixCh}} \\ \mathbf{0}_{N_{UpmixCh}} & \mathbf{E}_w \end{bmatrix} \mathbf{F}^H$$

であり、 $\mathbf{0}_{N_{UpmixCh}}$  は、ゼロのサイズ  $N_{UpmixCh} \times N_{UpmixCh}$  の正方行列である。  $\mathbf{F}$  に対する先の方程式を解析するために、それは、

10

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

を得る。

#### 【 0 2 0 5 】

これは、非相関化信号が、合計する際ににおいて、加重ゼロ (zero-weight) を有し、最後の出力は、ドライ信号によって与えられ、そして、それは、ダウンミックス信号と同一である。

20

$$\tilde{\mathbf{Z}} = \mathbf{P}\hat{\mathbf{Z}} + \mathbf{M}\mathbf{W} = \hat{\mathbf{Z}} \approx \mathbf{Y}$$

#### 【 0 2 0 6 】

その結果、このレンダリングシナリオにおいて、ダウンミックス信号に等しいようなシステム出力のための所与の前提条件は満たされる。

#### 【 0 2 0 7 】

##### 1 4 . 6 信号共分散行列 $\mathbf{E}_s$ の推定

30

混合行列  $\mathbf{F}$  を得るために、結合信号  $\mathbf{S}$  の共分散行列  $\mathbf{E}_s$  の情報が必要とされるか、または、少なくとも価値がある。

#### 【 0 2 0 8 】

原則として、利用できる信号から（すなわち、パラメトリック再構成  $\hat{\mathbf{Z}}$  および非相関器出力  $\mathbf{W}$  から）直接的に、共分散行列  $\mathbf{E}_s$  を推定することが可能である。この方法は、より正確な結果につながるけれども、関連する計算の複雑さのため、実際的でありえない。提案された方法は、共分散行列  $\mathbf{E}_s$  のパラメータの近似値を使用する。

#### 【 0 2 0 9 】

共分散行列  $\mathbf{E}_s$  の一般的な構造は、

40

$$\mathbf{E}_s = \begin{bmatrix} \mathbf{E}_{\hat{\mathbf{Z}}} & \mathbf{E}_{\hat{\mathbf{Z}}\mathbf{W}}^H \\ \mathbf{E}_{\hat{\mathbf{Z}}\mathbf{W}} & \mathbf{E}_{\mathbf{W}} \end{bmatrix}$$

として、表される。

ここで、行列  $\mathbf{E}_{\hat{\mathbf{Z}}\mathbf{W}}$  は、直接的な  $\hat{\mathbf{Z}}$  と非相関化信号  $\mathbf{W}$  との間の相互相関である。

#### 【 0 2 1 0 】

非相関器は理想的である（すなわち、エネルギーを保つこと、入力に対して、出力は直

50

交すること、そして、すべての出力は、相互に直交すること)と仮定するならば、共分散行列  $\mathbf{E}_S$  は、

$$\mathbf{E}_S = \begin{bmatrix} \mathbf{E}_{\hat{Z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_W \end{bmatrix}$$

として、簡略化された形式を使用して表されうる。

【0 2 1 1】

パラメータ的に再構成された信号  $\hat{\mathbf{Z}}$  の共分散行列  $\mathbf{E}_{\hat{Z}}$  は

10

$$\mathbf{E}_{\hat{Z}} = \mathbf{R}\mathbf{E}_{\hat{X}}\mathbf{R}^H = \mathbf{R}\mathbf{G}\mathbf{D}\mathbf{E}_X\mathbf{D}^H\mathbf{G}^H\mathbf{R}^H$$

として、パラメータ的に決定されうる。

【0 2 1 2】

非相関化信号  $\mathbf{W}$  の共分散行列  $\mathbf{E}_W$  は、相互の直交特性を満たし、

$$\mathbf{E}_W(i, j) = \begin{cases} \mathbf{E}_{\hat{Z}}(i, i) & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

20

に示すように、そして  $\mathbf{E}_{\hat{Z}}$  の対角要素のみを含むように仮定される。

【0 2 1 3】

相互に直交および / またはエネルギー保存の仮定が妨害される (例えば、利用可能な非相関器の数が、非相関化される信号の数よりも少ないとときのケースにおける) 場合、そのとき、共分散行列  $\mathbf{E}_W$  は、

$$\mathbf{E}_W = \mathbf{M}_{post} \left[ matdiag(\mathbf{M}_{pre} \mathbf{E}_{\hat{Z}} \mathbf{M}_{pre}^H) \right] \mathbf{M}_{post}^H$$

30

として、推定される。

【0 2 1 4】

14.7 任意の改良：非相関化信号およびエネルギー調整装置を使用する出力共分散修正

以下に、特に、有利なコンセプトが記載され、そして、それは、本願明細書において記載される他のコンセプトと組み合わされうる。

【0 2 1 5】

出力共分散誤り訂正のための提案された方法は、パラメータ的に再構成された信号  $\hat{\mathbf{Z}}$  および非相関化の部分  $\hat{\mathbf{Z}}$  の加重和として出力信号を構成する。この合計は、

40

$$\tilde{\mathbf{Z}} = \mathbf{P}\hat{\mathbf{Z}} + \mathbf{M}\mathbf{W} \quad (\text{I } 1)$$

として、表されうる。

【0 2 1 6】

## 結合行列

$$\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$$

および信号

$$\mathbf{S} = \begin{bmatrix} \hat{\mathbf{Z}} \\ \mathbf{W} \end{bmatrix}$$

10

のための表記法を適用して、

$$\tilde{\mathbf{Z}} = \mathbf{FS} \quad (\text{I } 2)$$

を得る。

## 【0 2 1 7】

しかしながら、この方程式は、最も一般的な定式化と考慮されうる点に留意されたい。  
変化は、本願明細書において記載されている全ての「簡略化された方法」に対して有効である上記の公式に、任意に適用されうる。

20

## 【0 2 1 8】

以下に、例えば、エネルギー調整装置によって実行されうる機能性が記載される。

## 【0 2 1 9】

極端な例において、最終の出力におけるアーティファクトの導入を回避するために、異なる制約が、混合行列  $\mathbf{F}$ （または混合行列  $\tilde{\mathbf{F}}$ ）において課されうる。言及される制約は、ターゲットおよび／またはパラメータ的に再構成された信号（例えば、レンダリングされたオーディオ信号）のエネルギーおよび／または相関特性に関して、絶対的な閾値または相対的な閾値によって表されうる。

30

## 【0 2 2 0】

このセクションにおいて記載される方法は、最終の出力混合ブロックにおいてエネルギー調整ステップを追加することによって、これを達成するために提案される。行列  $\mathbf{F}$ （または、そこから導出された「修正された」混合行列  $\tilde{\mathbf{F}}$ ）を伴う混合ステップの後、非相関化（ウェット）信号のエネルギーレベル（例えば、 $A_{wet} MW$ ）、および／またはパラメータ的に再構成された（ドライ）信号のエネルギーレベル（例えば、 $A_{dry} P \tilde{\mathbf{Z}}$ ）および／または最終の出力信号のエネルギーレベル ( $A_{dry} P \tilde{\mathbf{Z}} + A_{wet} MW$ ) が、ある閾値を超えないように、そのような処理ステップの提案は確保できる。

40

## 【0 2 2 1】

この追加の機能は、

$$\tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{A}_{\text{dry}} \mathbf{P} & \mathbf{A}_{\text{wet}} \mathbf{M} \end{bmatrix} \quad (\text{I } 3)$$

である結合混合行列  $\mathbf{F}$  の定義を改良することによって達成される。

ここで、（「エネルギー修正行列」としても言及される）2つの正方（または対角）エネルギー調整行列  $\mathbf{A}_{\text{dry}}$  および  $\mathbf{A}_{\text{wet}}$  は、パラメータ的に再構成された（ドライ）および非相関化された（ウェット）信号のそれぞれの混合加重（例えば、 $\mathbf{P}$  および  $\mathbf{M}$ ）に適用される。その結果、最後の出力は、

$$\begin{aligned} \tilde{\mathbf{Z}} &= \tilde{\mathbf{F}} \mathbf{S} \\ &= \mathbf{A}_{\text{dry}} \hat{\mathbf{P}} \tilde{\mathbf{Z}} + \mathbf{A}_{\text{wet}} \hat{\mathbf{M}} \mathbf{W} \end{aligned} \quad (\text{I } 4)$$

である。

#### 【0 2 2 2】

行列  $\tilde{\mathbf{F}}$  を伴う混合ステップのために、最後の出力信号（例えば、 $\tilde{\mathbf{Z}}$ ）レベルへのドライおよび／またはウェット信号（例えば、 $\tilde{\mathbf{Z}}$  および  $\mathbf{W}$ ）の貢献は、パラメータ的に再構成された信号（例えば、 $\tilde{\mathbf{Z}}$ ）および／または非相関化信号（例えば、 $\mathbf{W}$ ）および／またはターゲット信号に関して、ある相対的な閾値を上回らないように、ドライおよびウェットエネルギー修正行列  $\mathbf{A}_{\text{dry}}$  および  $\mathbf{A}_{\text{wet}}$  は、算出される。換言すれば、一般に、修正行列を算出するという複数の可能性がある。

#### 【0 2 2 3】

ドライおよびウェットエネルギー修正行列  $\mathbf{A}_{\text{dry}}$  および  $\mathbf{A}_{\text{wet}}$  は、例えば、エネルギーの関数、および／または、ドライ信号（例えば、 $\tilde{\mathbf{Z}}$ ）および／またはウェット信号（例えば、 $\mathbf{W}$ ）および／または所望の最後の出力信号の相関および／または共分散の特性、および／または、混合ステップ後のドライおよび／またはウェットおよび／または最後の出力信号の共分散行列の推定として、算出されうる。上述した可能性は、修正行列がどのように得られうるかのいくつかの例を記載する点に留意されたい。

#### 【0 2 2 4】

1つの可能な解決策は、以下の表現によって与えられる。

$$\mathbf{A}_{\text{dry}}(i, j) = \begin{cases} \min \left( 1, \sqrt{\max \left( 0, \lambda_{\text{dry}} \frac{\mathbf{E}_{\hat{z}}(i, i)}{\max(\mathbf{C}_{\text{estim}}(i, i), \varepsilon)} \right)} \right) & i = j, \\ 0 & i \neq j. \end{cases}$$

および、

$$\mathbf{A}_{\text{wet}}(i, j) = \begin{cases} \min \left( 1, \sqrt{\max \left( 0, \lambda_{\text{wet}} \frac{\mathbf{E}_{\hat{z}}(i, i)}{\max(\mathbf{C}_{\text{estim}}(i, i), \varepsilon)} \right)} \right) & i = j, \\ 0 & i \neq j. \end{cases}$$

であり、ここで、 $\lambda_{\text{dry}}$ および $\lambda_{\text{wet}}$ は、信号特性の関数（例えば、エネルギー、相関および／または共分散）としての定数または時間／周波数の変量である、2つの閾値であり、 $\varepsilon$ は、例えば、 $\varepsilon = 10^{-9}$ である、（任意の）小さい0以上の正則化定数であり、 $\mathbf{E}_{\hat{z}}$ は、パラメータ的に再構成された（ドライ）信号の共分散および／またはエネルギー情報を表し、そして、 $\mathbf{C}_{\text{estim}}$ は、行列 $\mathbf{F}$ を伴う混合ステップの後、ドライもしくはウェットの共分散行列の評価、または行列 $\mathbf{F}$ を伴う混合ステップの後、出力信号の共分散行列の推定を表し、そして、それは、現在の発明によって提案されるようなエネルギー調整ステップが適用されない場合、得られる（または、言い換えると、エネルギー調整装置が使用されない場合、得られる。）。

10

20

30

40

### 【0225】

上記の方程式において、例えば、独立変数 $\mathbf{C}_{\text{estim}}(i, i)$ および の最大値を供給する分母における「 $\max(\cdot)$ 」は、ゼロによる除算を回避するために、 または他のメカニズムの追加によって取り替えられる。

### 【0226】

例えば、 $\mathbf{C}_{\text{estim}}$ は、以下によって与えられうる。

$\mathbf{C}_{\text{estim}} = \mathbf{M}\mathbf{E}_w\mathbf{M}^H$  – 行列 $\mathbf{M}$ を伴う混合ステップの後、ウェット信号の共分散行列の推定

### 【0227】

$\mathbf{C}_{\text{estim}} = \mathbf{P}\mathbf{E}_{\hat{z}}\mathbf{P}^H$  – 行列 $\mathbf{P}$ を伴う混合ステップの後、ドライ信号の共分散行列の推定

### 【0228】

$\mathbf{C}_{\text{estim}} = \mathbf{P}\mathbf{E}_{\hat{z}}\mathbf{P}^H + \mathbf{M}\mathbf{E}_w\mathbf{M}^H$  – 行列 $\mathbf{F}$ を伴う混合ステップの後、出力信号の共分散行列の推定

### 【0229】

以下に、いくつかのさらなる簡略化したものが、記載される。換言すれば、出力共分散修正のための簡略化された方法が記載される。

### 【0230】

信号 $\hat{\mathbf{Z}}$ が既にM M S Eの趣旨において既に最適であることを考慮に入れて、これが、分離特性に影響を及ぼしうるので、出力 $\tilde{\mathbf{Z}}$ の共分散特性を改良するために、パラメトリック再構成（ドライ信号） $\hat{\mathbf{Z}}$ を修正することは、通常、望ましくない。

## 【0231】

非相関化（ウェット）信号Wの混合が操作されれば、混合行列Pは、単位行列に低減されうる。この場合、パラメータ的に再構成された（ドライ）信号に対応するエネルギー調整行列も、単位行列に低減されうる。このように、簡略化された方法は、

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \mathbf{A}_{\text{dry}} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

による設定によって、記載されうる。

## 【0232】

システムの最後の出力は、

$$\tilde{\mathbf{Z}} = \hat{\mathbf{Z}} + \mathbf{A}_{\text{wet}} \mathbf{MW}$$

20

として、表されうる。

## 【0233】

## 15. 非相関化装置のための複雑さの低減

以下に、本発明による実施の形態において使用される非相関器の複雑さがどのように低減されうるかが記載される。

## 【0234】

非相関器の機能の実装は、しばしば、計算的に複雑である点に留意されたい。いくつかのアプリケーション（例えば、ポータブル・デコーダ・ソリューション）において、非相関器の数の制限は、制限された計算機のリソースのために導入されることを必要とされうる。このセクションは、適用される非相関器（または非相関化）の数を制御することによって、非相関器の装置の複雑さの低減のための手段の記載を提供する。非相関化装置のインターフェースは、図16および図17において表現される。

30

## 【0235】

図16は、単純な（従来の）非相関化装置のブロック概略図を示す。図16による非相関化装置1600は、例えば、レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ のような、N個の非相関器入力信号1610a～1610nを受信するように構成される。さらに、非相関化装置1600は、N個の非相関器出力信号1612a～1612nを供給する。非相関化装置1600は、例えば、N個の個々の非相関器（または、非相関化の機能）1620a～1620nを含む。例えば、個々の非相関化装置1620a～1620nの各々は、非相関器入力信号1610a～1610nのうちの関連する1つに基づいて、非相関器出力信号1612a～1612nのうちの1つを供給しうる。従って、N個の個々の非相関器、または非相関化の機能1620a～1620nが、N個の非相関器入力信号1610a～1610nに基づいて、N個の非相関化信号1612a～1612nを供給することを必要としうる。

40

## 【0236】

50

一方、図17は、低減された複雑さの非相関化装置1700のブロック概略図を示す。低減された複雑さの非相関化装置1700は、N個の非相関器入力信号1710a～1710nを受信し、それに基づいて、N個の非相関器出力信号1712a～1712nを供給するように構成される。例えば、非相関器入力信号1710a～1710nは、レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ であり、そして、非相関器出力信号1712a～1712nは、非相関化オーディオ信号Wでありうる。

#### 【0237】

非相関器1700は、N個の非相関器入力信号1710a～1710nの第1のセットを受信し、それに基づいて、K個の非相関器入力信号1722a～1722kの第2のセットを供給するように構成される、プレミキサ（または、同等の機能的なプレミキシング）1720を含む。例えば、プレミキサ1720は、N個の非相関入力信号1710a～1710nの第1のセットに基づいて、K個の非相関器入力信号1722a～1722kの第2の設定を導出しうる、いわゆる「プレミックス」または「ダウンミックス」を実行しうる。例えば、K個の非相関化入力信号1722a～1722kの第2のセットのK個の信号は、行列 $\hat{\mathbf{Z}}_{mix}$ を使用して表されうる。非相関化装置（または、同等の、マルチチャネル非相関器）1700は、非相関器入力信号1722a～1722kの第2のセットのK個の信号を受信し、そして、それに基づいて、非相関器出力信号1732a～1732kの第1のセットを構成するK個の非相関器出力信号を供給するように構成される非相関器の主要部1730も含む。例えば、非相関器の主要部1730は、K個の個々の非相関器（または非相関化の機能）を含み、ここで、個々の非相関器（または、非相関化の関数）の各々は、K個の非相関器入力信号1722a～1722kの第2のセットの対応する非相関器入力信号に基づいて、K個の非相関器出力信号1732a～1732kの第1のセットの非相関器出力信号のうちの1つを供給する。あるいは、K個の非相関器出力信号1732a～1732kの第1のセットの非相関器出力信号の各々は、K個の非相関器入力信号1722a～1722kの第2のセットの非相関器入力信号のうちの1つの信号に基づくように、所与の非相関器、または非相関化の機能は、K回、適用されうる。

10

20

30

#### 【0238】

非相関化装置1700は、非相関器出力信号の第1のセットのK個の非相関器出力信号1732a～1732kを受信し、それに基づいて、（「外部の」非相関器出力信号を構成する）非相関器出力信号の第2のセットのN個の信号1712a～1712nを供給するように構成される、ポスト・ミキサ1740を含む。

#### 【0239】

プレミキサ1720は、プレミキシング行列 $M_{pre}$ によって記載される線形混合処理を、好ましくは実行しうる。さらに、ポスト・ミキサ1740は、K個の非相関器出力信号1732a～1732kの第1のセットから（すなわち、非相関器の主要部1730の出力信号から）非相関器出力信号の第2のセットのN個の非相関器出力信号1712a～1712nを導出するために、ポストミキシング行列 $M_{post}$ によって表されうる、線形混合（または、アップミックス）処理を、好ましくは実行する。

40

#### 【0240】

提案された方法と装置の主要な考えは、以下によって、NからKへ、非相関器への（または非相関器の主要部への）入力信号の数を低減することである。

#### 【0241】

・次式を伴う数を下げるために、信号（例えば、レンダリングされたオーディオ信号）をプレミックスする。

$$\hat{\mathbf{Z}}_{mix} = \mathbf{M}_{pre} \hat{\mathbf{Z}}$$

50

## 【0242】

- ・次式を伴う利用可能なK個の非相関器（例えば、非相関器の主要部）を使用して非相関化を適用する。

$$\hat{\mathbf{Z}}_{mix}^{dec} = Decorr(\hat{\mathbf{Z}}_{mix})$$

## 【0243】

- ・次式を伴うN個のチャネルへ非相関化信号をアップミックスする。

$$\mathbf{W} = \mathbf{M}_{post} \hat{\mathbf{Z}}_{mix}^{dec}$$

10

## 【0244】

行列の積( $\mathbf{M}_{pre}\mathbf{M}_{pre}^H$ )が、(反転操作に関して) 良好であるように、プレミキシング行列 $\mathbf{M}_{pre}$ は、ダウンミックス／レンダリング／相関／その他の情報に基づいて、構成されうる。ポストミキシング行列は、

$$\mathbf{M}_{post} \approx \mathbf{M}_{pre}^H (\mathbf{M}_{pre}\mathbf{M}_{pre}^H)^{-1}$$

として、算出されうる。

20

## 【0245】

中間物の非相関化信号 $\tilde{\mathbf{S}}$ （または、 $\hat{\mathbf{Z}}_{mix}^{dec}$ ）の共分散行列は、（理想的な非相関器を仮定する）対角行列であるけれども、この種の処理を使用するときに、最後の非相関化信号Wの共分散行列は、もはや、まったく対角行列ではない。従って、共分散行列は、

$$\mathbf{E}_W = \mathbf{M}_{post} \left[ matdiag(\mathbf{M}_{pre}\mathbf{E}_{\tilde{\mathbf{Z}}} \mathbf{M}_{pre}^H) \right] \mathbf{M}_{post}^H$$

30

として、混合行列を使用して、推定されうる。

## 【0246】

使用される非相関器（または、個々の非相関器）の数Kは、特定されず、所望の計算量の複雑さおよび利用可能な非相関器に依存している。その値は、（最も高い計算量の複雑さの）N個から（最も低い計算量の複雑さの）1個に至るまで変化する。

## 【0247】

非相関器の装置への入力信号の数Nは、任意であり、そして、提案された方法は、システムのレンダリング構造において独立していて、いかなる入力信号の数を支持する。

## 【0248】

例えば、3Dのオーディオ・コンテンツを使用しているアプリケーションにおいて、多数の出力チャネルについては、出力構成に応じて、プレミキシング行列 $\mathbf{M}_{pre}$ のためのある可能な表現が後述される。

40

## 【0249】

以下において、非相関化装置1700がマルチチャネル・オーディオ・デコーダにおいて使用された場合、プレミキサ1720（および、従って、ポスト・ミキサ1740によって実行されるポストミックス）によって実行されるプレミックスが、どのように調整されるかを記載する。ここで、非相関器入力信号の第1のセットの非相関器入力信号1710a～1710nは、オーディオ・シーンの異なる空間的な位置と関連している。

## 【0250】

50

この目的のために、図18は、異なる出力フォーマットのために使用されるスピーカ位置のテーブル表現を示す。

#### 【0251】

図18の表1800において、第1列1810は、スピーカのインデックスナンバーを記載する。第2列1820は、スピーカラベルを記載する。第3列1830は、それぞれのスピーカの方位角の位置を記載し、第4列1832は、スピーカの位置の方位角の許容差を記載する。第5列1840は、それぞれのスピーカの位置の仰角を記載し、第6列1842は、対応する仰角許容差を記載する。第7列1850は、スピーカが出力フォーマットO-2.0のために使用されることを指示する。第8列1860は、スピーカが出力フォーマットO-5.1のために使用されることを示す。第9列1864は、スピーカが出力フォーマットO-7.1のために使用されることを示す。第10列1870は、スピーカが出力フォーマットO-8.1のために使用されることを示し、第11列1880は、スピーカが出力フォーマットO-10.1のために使用されることを示し、そして、第12列1890は、スピーカが出力フォーマットO-22.2のために使用されることを示す。表に示されるように、2個のスピーカが、出力フォーマットO-2.0のために使用され、6個のスピーカが、出力フォーマットO-5.1のために使用され、8個のスピーカが、出力フォーマット7.1のために使用され、9個のスピーカが、出力フォーマットO-8.1のために使用され、11個のスピーカが、出力フォーマットO-10.1に対して使用され、そして、24個のスピーカが、出力フォーマットO-22.2のために使用される。10

#### 【0252】

しかしながら、1つの低音効果のスピーカは、出力フォーマットO-5.1, O-7.1, O-8.1およびO-10.1のために使用され、2つの低音効果のスピーカ(LFE1, LFE2)は、出力フォーマットO-22.2のために使用される点に留意されたい。さらに、好ましい実施の形態において、1つ以上の低音効果のスピーカを除いて、1つのレンダリングされたオーディオ信号(例えば、レンダリングされたオーディオ信号1582a～1582nのうちの1つ)が、スピーカの各々に関連している点に留意されたい。従って、2つのレンダリングされたオーディオ信号は、O-2.1フォーマットにより使用される2個のスピーカに関連しており、O-5.1フォーマットが使用される場合、5つのレンダリングされたオーディオ信号は、5個の低音効果でないスピーカに関連しており、O-7.1フォーマットが使用される場合、7つのレンダリングされたオーディオ信号は、7個の低音効果でないスピーカに関連しており、O-8.1フォーマットが使用される場合、8つのレンダリングされたオーディオ信号は、8個の低音効果でないスピーカに関連しており、O-10.1フォーマットが使用されうる場合、10個のレンダリングされたオーディオ信号が、10個の低音効果でないスピーカに関連しており、そして、O-22.2フォーマットが使用される場合、22個のレンダリングされたオーディオ信号は、22個の低音効果でないスピーカに関連している。30

#### 【0253】

しかしながら、上記したように、たいてい、(個々の)非相関器(または、非相関器の主要部)のより少ない数を使用することが望ましい。以下において、(行列 $\hat{Z}$ またはベクトル $\hat{z}$ によって表されうる)22個のレンダリングされたオーディオ信号1582a～1582nがあるように、O-22.2出力フォーマットが、マルチチャネル・オーディオ・デコーダによって使用される場合、非相関器の数がどのように柔軟に低減されるかについて記載される。40

#### 【0254】

図19a～図19gは、N=22のレンダリングされたオーディオ信号があるという仮定のもとに、レンダリングされたオーディオ信号1582a～1582nをプレミックスするための異なるオプションを表す。例えば、図19aは、プレミキシング行列M<sub>pre</sub>の50

10  
入力のテーブル表現を示す。図19aにおける1~11のラベルを付けられた行は、プレミキシング行列 $M_{pre}$ の行を表し、1~22のラベルを付けられた列は、プレミキシング行列 $M_{pre}$ の列を表す。さらに、プレミキシング行列 $M_{pre}$ の行の各々は、非相関器入力信号の第2のセットのK個の非相関器入力信号1722a~1722kのうちの1つ（すなわち、非相関器の主要部の入力信号）に関連している点に留意されたい。さらに、プレミキシング行列 $M_{pre}$ の列の各々は、非相関器入力信号の第1のセットのN個の非相関器入力信号1710a~1710nのうちの1つに関連しており、そして、その結果、（非相関器入力信号の第1のセットの非相関器入力信号1710a~1710nが、実施の形態におけるレンダリングされたオーディオ信号1582a~1582nに典型的に同一であるので）レンダリングされたオーディオ信号1582a~1582nに関連している。従って、プレミキシング行列 $M_{pre}$ の列の各々は、特定のスピーカに関連し、そして、特定の空間的な位置に関連しているので、その結果、スピーカは空間的な位置に関連している。行1910は、プレミキシング行列 $M_{pre}$ の列が関連しているスピーカ（およびその結果、空間的な位置）を示す（ここで、スピーカのラベルは、表1800の列1820において定義される）。

#### 【0255】

20  
以下において、図19aのプレミキシング $M_{pre}$ によって定義される機能が、さらに詳細に記載される。表に示すように、スピーカ（または、スピーカ位置と同等）「CH\_M\_000」および「CH\_L\_000」に関連するレンダリングされたオーディオ信号は、非相関器入力信号の第2のセットの第1の非相関器入力信号（すなわち、第1のダウンミックスされた非相関器入力信号）を得るために結合され、そして、それは、プレミキシング行列 $M_{pre}$ の第1行の第1列および第2列における「1」の値によって指示される。同様に、スピーカ（または、スピーカ位置と同等）「CH\_U\_000」および「CH\_T\_000」に関連するレンダリングされたオーディオ信号は、第2のダウンミックスされた非相関器入力信号（すなわち、非相関器入力信号の第2のセットの第2の非相関器入力信号）を得るために結合される。さらに、11個のダウンミックスされた非相関器入力信号が、22個のレンダリングされたオーディオ信号から導出されるように、図19aのプレミキシング行列 $M_{pre}$ は、2つのレンダリングされたオーディオ信号の11の組み合わせを定義することが示されうる。（プレミキシング行列の1~4列および1および2行を参照して）2つのダウンミックスされた非相関器入力信号を得るために、4つの中心の信号が結合されることを示す。さらに、他のダウンミックスされた非相関器入力信号が、オーディオ・シーンの同じ側に関連している2つのオーディオ信号を結合することによって各々得られることを示す。例えば、プレミキシング行列の第3行によって表される、第3のダウンミックスされた非相関器入力信号は、+135°。（「CH\_M\_L\_135」；「CH\_U\_L\_135」）の方位角の位置に関連しているレンダリングされたオーディオ信号を結合することによって得られる。さらに、（プレミキシング行列の第4行によって表される）第4の非相関器入力信号は、-135°。（「CH\_M\_R\_135」；「CH\_U\_R\_135」）の方位角の位置に関連しているレンダリングされたオーディオ信号を結合することによって得られる。従って、ダウンミックスされた非相関器入力信号の各々は、その（または同等の）方位角の位置（または、同等の水平位置）に関連している2つのレンダリングされたオーディオ信号を結合することによって得られる。ここで、典型的には、異なる仰角（または、同等の垂直位置）に関連している信号の結合を有する。

#### 【0256】

30  
40  
（プレミキシング行列 $M_{pre}$ の入力）N=22およびK=10に対するプレミキシング係数を示す図19bを現在、参照する。図19bの表の構成は、図19aの表の構成と同一である。しかしながら、表に示されるように、第1行が、チャネルID（または位置）「CH\_M\_000」、「CH\_L\_000」、「CH\_U\_000」および「CH\_T\_000」を有する4つのレンダリングされたオーディオ信号の結合を記載する点において、図19bによるプレミキシング行列 $M_{pre}$ は、図19aのプレミキシング行列 $M_{pre}$ と異なる。換言すれば、垂直に隣接する位置に関連している4つのレンダリングされたオ-

ディオ信号は、必要な非相関器（図19aによる行列のための11個の非相関器の代わりに10個の非相関器）の数を低減するためにプレミキシングにおいて結合される。

#### 【0257】

（プレミキシング行列 $M_{pre}$ の入力） $N = 22$ および $K = 9$ に対するプレミキシング係数を示す図19cを現在、参照し、表に示すように、図19cによるプレミキシング行列 $M_{pre}$ は、9行のみにより構成される。さらに、第2のダウンミックスされた非相関器入力信号（非相関器入力信号の第2のセットの非相関器入力信号）を得るために、チャネルID（または位置）「CH\_M\_L135」、「CH\_U\_L135」、「CH\_M\_R135」および「CH\_U\_R135」に関連するレンダリングされたオーディオ信号が結合される点において、図19cのプレミキシング行列 $M_{pre}$ の第2行から示されうる。  
10 表に示すように、図19aおよび図19bによりプレミキシング行列によって別々にダウンミックスされた非相関器入力信号に結合されているレンダリングされたオーディオ信号は、図19cによる共通のダウンミックスされた非相関器入力信号にダウンミックスされる。さらに、チャネルID「CH\_M\_L135」および「CH\_U\_L135」を有するレンダリングされたオーディオ信号は、オーディオ・シーンの同じ側における同一の水平位置（または方位角位置）および空間的に隣接する垂直位置（または仰角）に関連しており、そして、チャネルID「CH\_M\_R135」および「CH\_U\_R135」を有するレンダリングされたオーディオ信号は、オーディオ・シーンの第2の側における同一の水平位置（または方位角位置）および空間的に隣接する垂直位置（または仰角）に関連している点に留意されたい。さらに、チャネルID「CH\_M\_L135」、「CH\_U\_L135」、「CH\_M\_R135」および「CH\_U\_R135」を有するレンダリングされたオーディオ信号は、左側の位置および右側の位置を含んでいる空間的な位置の水平のペア（または、水平の4つの部分から構成される）に関連している。換言すれば、单一の所与の非相関器を使用して非相関化するように結合される4つのレンダリングされたオーディオ信号の2つが、オーディオ・シーンの左側の空間的な位置に関連しており、同じ所与の非相関器を使用して非相関化するように結合される4つのレンダリングされたオーディオ信号の2つが、オーディオ・シーンの右側の空間的な位置に関連している、図19cのプレミキシング行列 $M_{pre}$ の第2行において示されうる。さらに、レンダリングされたオーディオ信号の「対称な」4つの部分から構成されることとは、单一の（個々の）非相関器を使用して非相関化するようにプレミックスすることによって結合されるよう<sup>20</sup>に、（前記4つのレンダリングされたオーディオ信号の）左側のレンダリングされたオーディオ信号は、（前記4つのレンダリングされたオーディオ信号の）右側のレンダリングされたオーディオ信号に関連している空間的な位置を伴って、オーディオ・シーンの中心面に関して対称的である空間的な位置に関連している。  
30

#### 【0258】

図19d、19e、19fおよび19gを参照して、ますますレンダリングされたオーディオ信号が、（個々の）非相関器の数が減少とともに（すなわち、Kの減少とともに）結合されることが分かる。図19a～図19gに示すように、典型的に、2つに分かれたダウンミックスされた非相関器入力信号にダウンミックスされるレンダリングされたオーディオ信号は、1によって非相関器の数を減少させる場合、結合される。さらに、典型的に、空間的な位置の「対称な4つの部分から構成される」に関連している、この種のレンダリングされたオーディオ信号が、結合される。ここで、非相関器の比較的高い数に対して、等しいか、少なくとも類似の水平位置（または方位角位置）に関連しているレンダリングされたオーディオ信号が結合される。その一方で、非相関器の比較的低い数に対して、オーディオ・シーンの対向側における空間的な位置に関連しているレンダリングされたオーディオ信号も結合される。  
40

#### 【0259】

今、図20a～20d、21a～21c、22a～22bおよび23を参照して、類似のコンセプトが、レンダリングされたオーディオ信号の異なる数に対しても適用されうる。  
50

**【 0 2 6 0 】**

例えば、図 2 0 a ~ 2 0 d は、 $N = 10$  および  $K$  は 2 と 5 の間にに対する、プレミキシング行列  $M_{pre}$  の入力を記載する。

**【 0 2 6 1 】**

同様に、図 2 1 a ~ 2 1 c は、 $N = 8$  および  $K$  は 2 と 4 の間にに対する、プレミキシング行列  $M_{pre}$  の入力を記載する。

**【 0 2 6 2 】**

同様に、図 2 1 d ~ 2 1 f は、 $N = 7$  および  $K$  は 2 と 4 の間にに対する、プレミキシング行列  $M_{pre}$  の入力を記載する。

**【 0 2 6 3 】**

図 2 2 a および 2 2 b は、 $N = 5$  ならびに  $K = 2$  および  $K = 3$  に対する、プレミキシング行列の入力を示す。

**【 0 2 6 4 】**

最後に、図 2 3 は、 $N = 2$  および  $K = 1$  に対する、プレミキシング行列の入力を示す。

**【 0 2 6 5 】**

要約すると、マルチチャネル・オーディオ・デコーダーの一部であるマルチチャネル非相関器において、図 1 9 ~ 2 3 によるプレミキシング行列は、切り替え可能な方法で使用されうる。プレミキシング行列の間の切り替えは、例えば、(  $N$  個のレンダリングされたオーディオ信号を典型的に決定する ) 所望の出力構成に基づき、そして、( 例えば、パラメータ  $K$  を決定し、そして、オーディオ・コンテンツの符号化表現において含まれる複雑さの情報に基づいて調整されうる ) 非相関化の所望の複雑さにも基づいて、実行されうる。

10

20

**【 0 2 6 6 】**

今、図 2 4 を参照して、2 2 . 2 出力フォーマットのための複雑さの低減が、詳細に記載されている。既に、上で概説されるように、プレミキシング行列およびポストミキシング行列を構成するための 1 つの可能な解決法は、一緒に混合されたチャネルを選択するために、再生レイアウトの空間的情報を使用することであり、そして、ミキシング係数を算出する。それらの位置に基づいて、幾何学的に関連したスピーカ ( および、例えば、それに関連しているレンダリングされたオーディオ信号 ) は、図 2 4 において記載されるように垂直および水平の対を与え、グループ化される。換言すれば、図 2 4 は、表の形式において、レンダリングされたオーディオ信号に関連しうるように、スピーカの位置のグループ分けを示す。例えば、第 1 行 2 4 1 0 は、オーディオ・シーンの中央におけるスピーカ位置の第 1 グループを記載する。第 2 行 2 4 1 2 は、空間的に関連しているスピーカ位置の第 2 グループを表す。スピーカ位置「 C H \_ M \_ L 1 3 5 」および「 C H \_ U \_ L 1 3 5 」は、同一の方位角位置 ( または水平位置に同等 ) および隣接する仰角位置 ( または垂直に隣接する位置に同等 ) に関連している。同様に、位置「 C H \_ M \_ R 1 3 5 」および「 C H \_ U \_ R 1 3 5 」は、同一の方位角 ( または同一の水平位置に同等 ) および同様の仰角 ( または垂直に隣接する位置に同等 ) を含む。さらに、位置「 C H \_ M \_ L 1 3 5 」、「 C H \_ U \_ L 1 3 5 」、「 C H \_ M \_ R 1 3 5 」および「 C H \_ U \_ R 1 3 5 」は、位置の 4 つの部分から構成される。ここで、位置「 C H \_ M \_ L 1 3 5 」および「 C H \_ U \_ L 1 3 5 」は、オーディオ・シーンの中心面に関して、位置「 C H \_ M \_ R 1 3 5 」および「 C H \_ U \_ R 1 3 5 」に対称である。さらに、位置「 C H \_ M \_ 1 8 0 」および「 C H \_ U \_ 1 8 0 」も同一の方位角の位置 ( または同一の水平位置に同等 ) および同様の仰角 ( または隣接する垂直位置に同等 ) を含む。

30

40

**【 0 2 6 7 】**

第 3 行 2 4 1 4 は、位置の第 3 グループを表す。位置「 C H \_ M \_ L 0 3 0 」および「 C H \_ L \_ L 0 4 5 」は、空間的に隣接する位置であり、同様な方位角 ( または同様な水平位置に同等 ) および同様な仰角 ( または同様な垂直位置に同等 ) を含む点に留意されたい。同じことは、位置「 C H \_ M \_ R 0 3 0 」および「 C H \_ L \_ R 0 4 5 」に対しても維持する。さらに、位置の第 3 グループの位置は、位置の 4 つの部分から構成されるようになされる。ここで、位置「 C H \_ M \_ L 0 3 0 」および「 C H \_ L \_ 0 4 5 」は、位

50

置「C H \_ M \_ R 0 3 0」および「C H \_ L \_ R 0 4 5」に空間的に隣接しており、オーディオ・シーンの中心面に関して対称である。

#### 【0268】

第4行2416は、4つの追加の位置を表し、第2行の第1の4つの位置と比較したとき、同様の特性を有しており、位置の対称に4つの部分から構成されるように形成する。

#### 【0269】

第5行2418は、対称の位置「C H \_ M \_ L 0 6 0」、「C H \_ U \_ L 0 4 5」、「C H \_ M \_ R 0 6 0」および「C H \_ U \_ R 0 4 5」の他の4つの部分から構成されることを表す。

#### 【0270】

さらに、位置の異なるグループの位置に関連しているレンダリングされたオーディオ信号は、非相関器の数の減少とともにますます結合されうる点に留意されたい。例えば、マルチチャネル非相関器における11個の個々の非相関器において、第1および第2列における位置に関連しているレンダリングされたオーディオ信号は、各グループに対して結合されうる。加えて、第3および第4列において表される位置に関連しているレンダリングされたオーディオ信号は、各グループに対して結合されうる。さらに、第5および第6列において示される位置に関連しているレンダリングされたオーディオ信号は、第2グループに対して結合される。従って、（個々の非相関器に入力される）11個のダウンミックス非相関器入力信号が得られうる。しかしながら、より少ない個々の非相関器を有することが望ましい場合、1～4列において示される位置に関連しているレンダリングされたオーディオ信号は、1つ以上のグループに対して結合されうる。また、個々の非相関器の数をさらに減少することが望ましい場合、第2のグループのすべての位置に関連しているレンダリングされたオーディオ信号が結合されうる。

10

20

30

40

#### 【0271】

要約すると、出力レイアウト（例えば、スピーカに）に供給される信号は、水平および垂直依存があり、非相関化処理の間、維持されなければならない。従って、異なるスピーカのグループに対応するチャネルが一緒に混合されないように、ミキシング係数は算出される。

#### 【0272】

利用可能な非相関器の数、または非相関化の所望のレベルに応じて、各グループにおいて、第1は、（中間層および上層との間、または中間層および下層との間における）垂直のペアと一緒に混合される。第2に、（左と右の間における）水平のペアまたは残りの垂直のペアが一緒に混合される。例えば、グループ3において、最初に、左の垂直のペア（「C H \_ M \_ L 0 3 0」および「C H \_ L \_ L 0 4 5」）と右の垂直のペア（「C H \_ M \_ R 0 3 0」および「C H \_ L \_ R 0 4 5」）におけるチャネルが一緒に混合され、このようにして、4から2まで、このグループのための必要な非相関器の数を減少する。非相関器の数をより減らすことが望ましい場合、得られた水平ペアは、1つのチャネルのみにダウンミックスされ、このグループのための必要な非相関器の数は、4から1まで減らされる。

#### 【0273】

提示された混合規則に基づいて、（例えば、図19～23において示される）前述の表は、所望の非相関器の異なるレベルに対して（または、所望の非相関器の複雑さの異なるレベルに対して）導出される。

#### 【0274】

##### 16. 第2の外部のレンダラ／フォーマット変換器との互換性

S A O C デコーダ（または、さらに一般的にいえば、マルチチャネル・オーディオ・デコーダ）が、外部の第2のレンダラ／フォーマット変換器と共に使用される場合、以下において、提案されたコンセプト（方法または装置）に対する変更が、使用されうる。

#### 【0275】

— (例えば、レンダラの) 内部レンダリング行列Rは、(外部のレンダラが使用される場合) 恒等式  $\mathbf{R} = I_{N_{Objects}}$  に設定されるか、または(外部フォーマット変換器が使用される場合) 中間のレンダリング構造から導出されたミキシング係数に初期化するよう に設定される。

#### 【0276】

- 非相関器の数が、レンダラ / フォーマット変換器から受信されたフィードバック情報に基づいて算出されるプレミキシング行列  $M_{pre}$  を有するセクション 15において記載される方法を使用して低減される(例えば、 $M_{pre} = D_{convert}$ 、ここで、 $D_{convert}$  は、10 フォーマット変換器の内部で使用されるダウンミックス行列である。)。S A O C デコードの外側で一緒に混合されるチャネルは、一緒にプレミックスされ、そして、S A O C デコードの内部の同じ非相関器に供給される。

#### 【0277】

外部フォーマット変換器を用いて、S A O C 内部レンダラは、中間の構造(例えば、最 多数のスピーカの構造)にプレレンダラする。

#### 【0278】

結論として、プレミキシング行列が、外部レンダラに実際に結合される(非相関器入力信号の第 1 のセットの)この種の非相関器入力信号の組み合わせを定義するように、いくつかの実施の形態において、出力オーディオ信号についての情報は、外部レンダラにおいて一緒に混合されるか、またはフォーマット変換器がプレミキシング行列  $M_{pre}$  を決定するように使用される。このように、(マルチチャネル非相関器の出力オーディオ信号を受信する)外部レンダラ / フォーマット変換器から受信された情報は(例えば、マルチチャネル・オーディオ・デコーダの内部レンダリング行列が固有にセットされるか、または、中間のレンダリング構造から導出されたミキシング係数に初期化するよう に設定される場合)、プレミキシング行列を選択もしくは調整するよう に使用され、そして、外部レンダラ / フォーマット変換器は、マルチチャネル・オーディオ・デコーダに関して、上記したよう に出力オーディオ信号を受信するよう に接続される。20

#### 【0279】

##### 17. ビットストリーム

以下において、追加の信号伝達情報がビットストリームにおいて(または、オーディオ・コンテンツの符号化表現において、同等に) 使用されうることが記載される。本発明による実施の形態において、非相関化の方法は、所望の品質レベルを確実にするためのビットストリームに信号を送信されうる。このような方法で、ユーザ(またはオーディオ・エンコーダ)は、コンテンツに基づいて方法を選択するよう に、より多くの柔軟性を有する。この目的のために、M P E G S A O C ビットストリーム構文は、例えば、使用された非相関化の方法を特定するための 2 ビットおよび / または構造(または複雑さ)を特定するための 2 ビットによって延長されうる。30

#### 【0280】

図 25 は、例えば、ビットストリーム部「S A O C S p e c i f i g C o n f i g ()」または「S A O C 3 D S S p e c i f i c C o n f i g ()」に追加されうる、ビットストリーム要素「b s D e c o r r l a t i o n M e t h o d」および「b s D e c o r r e l a t i o n L e v e l」の構文表現を示す。図 25 に示されるように、2 つのビットは、ビットストリーム要素「b s D e c o r r e l a t i o n M e t h o d」のために使用され、そして、2 つのビットは、ビットストリーム要素「b s D e c o r r e l a t i o n L e v e l」のために使用されうる。40

#### 【0281】

図 26 は、表の形式において、ビットストリーム変数「b s D e c o r r e l a t i o n M e t h o d」の値と異なる非相関化方法との間の関連性を示す。例えば、3 つの異なる非相関化の方法が、前記ビットストリーム変数の異なる値によって信号を送信されうる50

。例えば、セクション 14.3において記載されるように、例えば、非相関化信号を使用する出力共分散修正は、オプションのうちの1つとして信号を送信されうる。他のオプションとして、例えば、セクション 14.4.1において記載されるように、共分散調整方法が信号を送信されうる。さらにもう1つのオプションとして、例えば、セクション 14.4.2において記載されるように、エネルギー補償方法が信号を送信されうる。従って、レンダリングされたオーディオ信号および非相関化オーディオ信号に基づいて出力オーディオ信号の信号特性の再構成のための3つの異なる方法は、ビットストリーム変数に依存して選択されうる。

#### 【0282】

エネルギー補償モードは、セクション 14.4.2において記載される方法を使用し、制限された共分散調整モードは、セクション 14.4.1において記載される方法を使用し、一般的な共分散調整モードは、セクション 14.3において記載される方法を使用する。

#### 【0283】

今、表の表現の形式において、異なる非相関化レベルが、ビットストリーム変数「*b s D e c o r r e l a t i o n L e v e l*」によってどのように信号を送信されうるかを示す図 27への参照することで、非相関化の複雑さを選択する方法が記載される。換言すれば、前記変数は、非相関化の複雑さが使用されるかについて決定するように、上記のマルチチャネル非相関器を含むマルチチャネル・オーディオ・デコーダによって評価されうる。例えば、前記ビットストリーム・パラメータは、値 0, 1, 2 および 3 を示されうる異なる非相関化「レベル」の信号を送信しうる。

#### 【0284】

(例えば、非相関化レベルとして示されうる) 非相関化の構成の例が、図 27 の表において与えられる。図 27 は、異なる「レベル」(例えば、非相関化レベル) および出力構造のための非相関器の数の表の表現を示す。換言すれば、図 27 は、マルチチャネル非相関器によって使用される(非相関器入力信号の第 2 のセットの) K 個の非相関器入力信号を示す。図 27 の表において示されるように、ビットストリーム・パラメータ「*b s D e c o r r e l a t i o n L e v e l*」によって信号を送信される「非相関化レベル」に基づいて、マルチチャネル非相関器において使用される(個々の) 非相関器の数は、22.2 の出力構造に対して、11, 9, 7 および 5 の間ににおいて切り替えられる。10.1 の出力構造に対して、選択は、10, 5, 3 および 2 の個々の非相関器の間ににおいてなされ、8.1 の構造に対して、選択は、8, 4, 3 または 2 の個々の非相関器の間ににおいてなされ、そして、7.1 の出力構造に対して、選択は、前記ビットストリーム・パラメータによって信号を送信される「非相関化レベル」に依存する、7, 4, 3 および 2 の間ににおいて切り替えられる。5.1 の出力構造において、個々の非相関器の数、すなわち、5, 3 または 2 のような個々の非相関器の数のための 3 つの有効なオプションのみを有する。2.1 の出力構造に対して、2 つの個々の非相関器(非相関化レベル 0)と 1 つの個々の非相関器(非相関化レベル 1)との間の選択のみである。

#### 【0285】

要約すると、非相関化の方法は、計算機のパワーおよび利用可能な非相関器の数に基づいて、デコーダ側で決定されうる。加えて、非相関器の数の選択は、エンコーダ側でなされ、ビットストリーム・パラメータを使用して信号の送信がなされる。

#### 【0286】

従って、出力オーディオ信号を得るために、非相関化オーディオ信号がどのように適用されるか 2 つの方法が適用され、そして、非相関化信号の供給のための複雑さが、図 25において示されるビットストリームを使用してオーディオ・エンコーダのサイドから制御され、そして、図 26 および 27 においてより詳細に定義される。

#### 【0287】

##### 18. 発明の処理のためのアプリケーションの分野

オーディオ・シーンの人間の認識のためのより大きな重要性である導かれた方法の目的

10

20

30

40

50

のうちの 1 つが、オーディオ・キューを復元することである点に留意されたい。本発明による実施の形態は、エネルギーのレベルおよび相関特性の再構成の正確さを改良し、従って、最後の出力信号の知覚的なオーディオ品質を増加させる。本発明による実施の形態は、任意の数のダウンミックス / アップミックスチャネルのために適用されうる。さらに、本願明細書において記載される方法および装置は、既存のパラメータのソース分離アルゴリズムと結合されうる。本発明による実施の形態は、適用される非相関化の機能の数における設定の制約によってシステムの計算の複雑さを制御することを許容する。本発明による実施の形態は、M P S とトランスコーディング・ステップを取り除くことによって、S A O C のようなオブジェクト・ベースのパラメトリック構造アルゴリズムの簡略化に通じうる。

10

### 【 0 2 8 8 】

#### 1 9 . 符号化 / 復号化の環境

以下において、本発明によるコンセプトにおけるオーディオ符号化 / 復号化の環境が記載される。

### 【 0 2 8 9 】

本発明によるコンセプトが使用されうる 3 D オーディオ・コーデック・システムは、チャネルの符号化のための M P E G - D U S A C コーデックおよびオブジェクトの大量の符号化のための効率を増加させるためのオブジェクト信号に基づく。M P E G - S A O C 技術は、適応されている。レンダラの 3 つのタイプは、チャネルのレンダリングオブジェクト、ヘッドホンへのレンダリングチャネル、または異なるスピーカセットへのレンダリングチャネルのタスクを実行する。オブジェクト信号が、S A O C を使用して明確に送信されるか、またはパラメータ的に符号化される場合に、対応するオブジェクトのメタデータ情報が圧縮され、そして、3 D オーディオ・ストリームに多重化される。

20

### 【 0 2 9 0 】

図 2 8 , 2 9 および 3 0 は、3 D オーディオ・システムの異なるアルゴリズムのブロックに示す。

### 【 0 2 9 1 】

図 2 8 は、この種のオーディオ・エンコーダのブロック概略図を示し、そして、図 2 9 は、この種のオーディオ・デコーダのブロック概略図を示す。換言すれば、図 2 8 および 2 9 は、3 D オーディオ・システムの異なるアルゴリズムのブロックを示す。

30

### 【 0 2 9 2 】

3 D オーディオ・エンコーダ 2 9 0 0 のブロック概略図を示す図 2 8 を参照することで、いくつかの詳細は説明される。エンコーダ 2 9 0 0 は、1 つ以上のチャネル信号 2 9 1 2 および 1 つ以上のオブジェクト信号 2 9 1 4 を受信し、そして、それに基づいて、1 つ以上のチャネル信号 2 9 1 6 および 1 つ以上のオブジェクト信号 2 9 1 8 , 2 9 2 0 を供給するオプションのプレレンダラ / ミキサ 2 9 1 0 を含む。オーディオ・エンコーダは、U S A C エンコーダ 2 9 3 0 および任意に S A O C エンコーダ 2 9 4 0 も含む。S A O C エンコーダ 2 9 4 0 は、1 つ以上の S A O C トランスポート・チャネル 2 9 4 2 および S A O C エンコーダに供給される 1 つ以上のオブジェクト 2 9 2 0 に基づいて S A O C サイド情報 2 9 4 4 を供給するように構成される。さらに、U S A C エンコーダ 2 9 3 0 は、プレレンダラ / ミキサ 2 9 1 0 からチャネルおよびプレレンダリングされたオブジェクトを含むチャネル信号 2 9 1 6 を受信し、プレレンダラ / ミキサ 2 9 1 0 から 1 つ以上のオブジェクト信号 2 9 1 8 を受信し、そして、1 つ以上の S A O C トランスポート・チャネル 2 9 4 2 および S A O C サイド情報 2 9 4 4 を受信し、そして、それに基づいて、符号化表現 2 9 3 2 を供給するように構成される。さらに、オーディオ・エンコーダ 2 9 0 0 は、( プレレンダラ / ミキサ 2 9 1 0 によって評価されうる ) オブジェクト・メタデータ 2 9 5 2 を受信し、符号化オブジェクト・メタデータ 2 9 5 4 を得るためにオブジェクト・メタデータを符号化するように構成される、オブジェクト・メタデータ・エンコーダ 2 9 5 0 も含む。符号化メタデータは、U S A C エンコーダ 2 9 3 0 によっても受信され、符号化表現 2 9 3 2 を供給するために使用される。

40

50

**【0293】**

オーディオ・エンコーダ2900の個々の構成要素に関するいくつかの詳細が以下に説明される。

**【0294】**

図29を参照することで、オーディオ・デコーダ3000が記載される。オーディオ・デコーダ3000は、符号化表現3010を受信し、それに基づいて、マルチチャネル・スピーカ信号3012、ヘッドホン信号3014および/または代替フォーマットにおける（例えば、5.1フォーマットにおける）スピーカ信号3016を供給するように構成される。オーディオ・デコーダ3000は、符号化表現3010に基づいて、1つ以上のチャネル信号3022、1つ以上のプレレンダリングされたオブジェクト信号3024、1つ以上のオブジェクト信号3026、1つ以上のSAOCトランスポート・チャネル3028、SAOCサイド情報3030および圧縮されたオブジェクト・メタデータ情報3032を供給する、USACデコーダ3020を含む。オーディオ・デコーダ3000は、1つ以上のオブジェクト信号3026およびオブジェクト・メタデータ情報3044に基づいて、1つ以上のレンダリングされたオブジェクト信号3042を供給するように構成されるオブジェクト・レンダラ3040も含む。ここで、オブジェクト・メタデータ情報3044は、圧縮されたオブジェクト・メタデータ情報3032に基づいて、オブジェクト・メタデータ・デコーダ3050によって供給される。オーディオ・デコーダ3000は、SAOCトランスポート・チャネル3028およびSAOCサイド情報3030を受信し、それに基づいて、1つ以上のレンダリングされたオブジェクト信号3062を供給するように構成されるSAOCデコーダ3060も任意に含む。オーディオ・デコーダ3000は、チャネル信号3022、プレレンダリングされたオブジェクト信号3024・レンダリングされたオブジェクト信号3042、およびレンダリングされたオブジェクト信号3062を受信し、それに基づいて、例えば、マルチチャネル・スピーカ信号3012を構成する複数の混合チャネル信号3072を供給するように構成されるミキサ3070も含む。オーディオ・デコーダ3000は、例えば、混合チャネル信号3072を受信し、それに基づいて、ヘッドホン信号3014を供給するように構成される、バイノーラル・レンダラ3080も含む。さらに、オーディオ・デコーダ3000は、混合チャネル信号3072および再構成レイアウト情報3092を受信し、それに基づいて、代替のスピーカセットのためにスピーカ信号3016を供給するように構成されるフォーマット変換器3090を含みうる。

10

20

30

**【0295】**

以下において、オーディオ・エンコーダ2900およびオーディオ・デコーダ3000の構成要素に関するいくつかの詳細が、記載されている。

**【0296】****19.1 プレレンダラ／ミキサ**

プレレンダラ／ミキサ2910は、符号化前に、チャネルに加えて、オブジェクト入力シーンをチャネル・シーンに変換するために任意に使用されうる。機能的に、例えば、それは、後述するオブジェクト・レンダラ／ミキサと同一もありうる。

40

**【0297】**

オブジェクトのプレレンダリングは、同時に能動的なオブジェクト信号の数から独立しているエンコーダ入力での決定論的な信号エントロピーを、例えば、確実にしうる。

**【0298】**

オブジェクトのプレレンダリングについて、オブジェクト・メタデータの伝送は、必要とされない。

**【0299】**

個別のオブジェクト信号は、エンコーダが使用するように構成されるチャネル・レイアウトにレンダリングされ、各チャネルに対するオブジェクトの重みは、関連するオブジェクト・メタデータ(OAM)1952から得られる。

**【0300】**

50

## 19.2 U S A C はの主要部のコーデック

スピーカチャネル信号、個々のオブジェクト信号、オブジェクト・ダウンミックス信号およびプレレンダリングされた信号のための主要部のコーデック 2930, 3020 は、M P E G - D U S A C 技術に基づく。それは、入力チャネルおよびオブジェクト割り当ての幾何学的および意味論的な情報に基づくチャネルおよびオブジェクト・マッピング情報による多数の信号の復号化を扱う。このマッピング情報は、どのように、入力チャネルおよびオブジェクトが U S A C チャネル要素 (C P E, S C E, L F E) にマッピングされるか、そして、対応する情報は、デコーダに送信されるかを記載する。

### 【0301】

S A O C データまたはオブジェクト・メタデータのようなすべての追加のペイロードは、拡張要素を通して渡しており、エンコーダにおいて考慮されている。オブジェクトの復号化は、レンダラのためレート / 歪みの前提条件および双方向性の前提条件によって決まる異なる方法で可能である。以下のオブジェクト符号化変数が可能である。

### 【0302】

- ・ プレレンダリングされたオブジェクト：オブジェクト信号は、符号化前にプレレンダリングされ、22.2 チャネル信号に混合される。次の符号化チェーンは、22.2 チャネル信号を参照する。

### 【0303】

- ・ 別々のオブジェクト波形：エンコーダにモノラル波形として適用されるようなオブジェクト。エンコーダは、チャネル信号に加えて、オブジェクトを送信するために単一のチャネル要素 S C E を使用する。復号化オブジェクトは、レンダリングされ、受信側で混合される。圧縮されたオブジェクト・メタデータ情報は、同時に、受信機 / レンダラに送信される。

### 【0304】

- ・ パラメータのオブジェクト波形：オブジェクト特性および各々に対するそれらの関係は、S A O C パラメータによって記載されている。オブジェクト信号のダウンミックスは、U S A C によって符号化される。パラメータ情報は、同時に送信される。ダウンミックスの数は、オブジェクトの数および全体のデータレートに応じて選択される。圧縮されたオブジェクト・メタデータ情報は、S A O C レンダラに送信される。

### 【0305】

#### 19.3. S A O C

オブジェクト信号のための S A O C エンコーダ 2940 および S A O C デコーダ 3060 は、M P E G - S A O C 技術に基づく。システムは、送信されたチャネルおよび付加的なパラメータ・データ（オブジェクト・レベル差 O L D、内部オブジェクト相關 I O C、ダウンミックス・ゲイン D M G）より少ない数に基づいて、オーディオ・オブジェクトの数を再形成し、修正し、そしてレンダリングすることができる。付加的なパラメータのデータは、復号化を非常に効率的にし、個々に、すべてのオブジェクトを送信するための必要であるより著しく低いデータレートを示す。S A O C エンコーダは、モノラル波形としてオブジェクト / チャネル信号を入力し、そして、（3D オーディオ・ビットストリーム 2932, 3010 に圧縮される）パラメトリック情報および（単一のチャネル要素を使用して符号化され、送信される）S A O C トランスポート・チャネルを出力する。S A O C デコーダ 3000 は、復号化 S A O C トランスポート・チャネル 3028 およびパラメータの情報 3030 からオブジェクト / チャネル信号を再構成し、再構成レイアウト、展開されたオブジェクト・メタデータ情報およびユーザ相互作用情報のオプションに基づいて、出力オーディオ・シーンを生成する。

### 【0306】

#### 19.4. オブジェクト・メタデータ・コーデック

オブジェクト毎に、3D 空間ににおけるオブジェクトの幾何学的な位置および量を特定する関連するメタデータは、時間および空間のオブジェクト特性の量子化によって、効率よく符号化される。圧縮されたオブジェクト・メタデータ c O A M 2954, 3032 は、

10

20

30

40

50

サイド情報として受信機に送信される。

【0307】

19.5. オブジェクト・レンダラ／ミキサ

オブジェクト・レンダラは、所与の再構成フォーマットに従ってオブジェクト波形を生成するために、展開されたオブジェクト・メタデータ O A M 3 0 4 4 を利用する。各オブジェクトは、そのメタデータに従って、特定の出力チャネルにレンダリングされる。このブロックの出力は、部分的な結果の合計から結果を得る。

【0308】

コンテンツと個々の / パラメータのオブジェクトに基づく両方のチャネルが復号化される場合、結果として得られる波形を出力する前に（または、バイノーラル・レンダラもしくはスピーカ・レンダラ・モジュールのようなポスト・プロセッサ・モジュールにそれらを供給する前に）、チャネル・ベースの波形とレンダリングされたオブジェクトの波形が混合される。

10

【0309】

19.6. バイノーラル・レンダラ

各入力チャネルが仮想音源によって表されるように、バイノーラル・レンダラ・モジュール 3 0 8 0 は、マルチチャネル・オーディオ材料のバイノーラルのダウンミックスを生じる。処理は、Q M F 領域にフレームごとに実行される。バイノーラル化は、測定されたバイノーラルの部屋のインパルス応答に基づく。

20

【0310】

19.7. スピーカ・レンダラ／フォーマット変換器

スピーカ・レンダラ 3 0 9 0 は、送信されたチャネル構成および所望の再生フォーマットとの間で変換する。以下において、それは、「フォーマット変換器」と呼ばれる。フォーマット変換器は、より少ない数の出力チャネルに変換する、すなわち、ダウンミックスを生成する。システムは、自動的に入出力フォーマットの所与の組み合わせのための最適化されたダウンミックス行列を生成して、ダウンミックス処理においてこれらの行列を適用する。フォーマット変換器は、標準のスピーカの構成と同様に、非標準のスピーカの位置を有するランダムな構成を許容する。

【0311】

図 3 0 は、フォーマット変換器のブロック概略図を示す。換言すれば、図 3 0 は、フォーマット変換器の構造を示す。

30

【0312】

図で示されるように、フォーマット変換器 3 1 0 0 は、ミキサ出力信号 3 1 1 0 、例えば、混合チャネル信号 3 0 7 2 を受信し、スピーカ信号 3 1 1 2 、例えば、スピーカ信号 3 0 1 6 を供給する。フォーマット変換器は、Q M F 領域におけるダウンミックス処理 3 1 2 0 およびダウンミックス・コンフィギュレータ 3 1 3 0 を含む。ここで、ダウンミックス・コンフィギュレータは、ミキサ出力レイアウト情報 3 0 3 2 および再構成レイアウト情報 3 0 3 4 に基づいて、コンフィギュレーション情報をダウンミックス処理 3 0 2 0 に供給する。

40

【0313】

19.8. 概論

さらに、本願明細書において記載されたコンセプト、例えば、オーディオ・デコーダ 1 0 0 、オーディオ・エンコーダ 2 0 0 、マルチチャネル非相関器 6 0 0 、マルチチャネル・オーディオ・デコーダ 7 0 0 、オーディオ・エンコーダ 8 0 0 またはオーディオ・デコーダ 1 5 5 0 が、オーディオ・エンコーダ 2 9 0 0 および / またはオーディオ・デコーダ 3 0 0 0 において使用されうる点に留意されたい。例えば、前述のオーディオ・エンコーダ / デコーダは、S A O C エンコーダ 2 9 4 0 の一部としておよび / または S A O C デコーダ 3 0 6 0 の一部として使用されうる。一方、前述のコンセプトは、3 D オーディオ・デコーダ 3 0 0 0 および / またはオーディオ・エンコーダ 2 9 0 0 の他の位置でも使用されうる。

50

## 【0314】

当然、前述の方法は、図28および29に従って、オーディオ情報を符号化または復号化するためのコンセプトにおいても使用されうる。

## 【0315】

## 20.付加的な実施の形態

## 20.1はじめに

以下において、本発明による他の実施の形態が、記載される。

## 【0316】

図31は、本発明の実施の形態によるダウンミックス・プロセッサのブロック概略図を示す。

10

## 【0317】

ダウンミックス・プロセッサ3100は、アップミキサ3110、レンダラ3120、コンバイナ3130およびマルチチャネル非相関器3140を含む。レンダラは、レンダリングされたオーディオ信号 $Y_{dry}$ をコンバイナ3130およびマルチチャネル非相関器3140に供給する。マルチチャネル非相関器は、(非相関器入力信号の第1のセットとしてみなされうる)レンダリングされたオーディオ信号を受信し、それに基づいて非相関器入力信号のプレミックスされた第2のセットを非相関器の主要部3160に供給する、アップミキサ3150を含む。非相関器の主要部は、ポスト・ミキサ3170によって利用するために、非相関器入力信号の第2のセットに基づいて、非相関器出力信号の第1のセットを供給する。ポスト・ミキサは、コンバイナ3130に供給されるポストミックスされた非相関器出力信号の第2のセットを得るために、非相関器の主要部3160によって供給された非相関器出力信号をポストミックス(または、アップミックス)する。

20

## 【0318】

レンダラ3130は、例えば、レンダリングするための行列Rを適用し、アップミキサは、例えば、プレミックスするための行列 $M_{pre}$ を適用し、ポスト・ミキサは、例えば、ポストミックスするための行列 $M_{post}$ を適用し、そして、コンバイナは、例えば、結合するための行列Pを適用する。

30

## 【0319】

ダウンミックス・プロセッサ3100または個々の構成要素またはその機能は、本願明細書において記載されるオーディオ・デコーダにおいて、使用されうる点に留意されたい。さらに、ダウンミックス・プロセッサは、本願明細書において記載されるいくつかの特徴および機能によって補充されうる点に留意されたい。

30

## 【0320】

## 20.2.SAOC 3D処理

ISO/IEC 23003-1:2007に記載されているハイブリッド・フィルタバンクが適用される。DMG、OLD、IOCパラメータの逆量子化は、ISO/IEC 23003-2:2010の7.1.2において定義されるように同じ規則に従う。

40

## 【0321】

## 20.2.1信号およびパラメータ

オーディオ信号は、あらゆる時間枠nおよびあらゆるハイブリッド・サブバンドkに対して定義される。対応するSAOC 3Dパラメータは、各パラメータ時間枠1および処理バンドmに対して定義される。ハイブリッドおよびパラメータ領域間の次のマッピングは、ISO/IEC 23003-1:2007の表A.3.1によって特定される。それゆえ、すべての計算は、特定の時間/バンド・インデックスに関して実行され、そして、対応する次元は、各導入変数に対して暗に定義される。

## 【0322】

SAOC 3Dデコーダで利用可能なデータは、マルチチャネル・ダウンミックス信号X、共分散行列E、レンダリング行列Rおよびダウンミックス行列Dで構成される。

## 【0323】

## 20.2.1.1 オブジェクト・パラメータ

要素  $e_{i,j}$  を有するサイズ  $N \times N$  の共分散行列 E は、元の信号共分散行列  $\mathbf{E} \approx \mathbf{SS}^*$  の近似を表し、

$$e_{i,j} = \sqrt{OLD_i OLD_j} IOC_{i,j}$$

として、OLD および IOC パラメータから得られる。

10

## 【0324】

ここで、逆量子化オブジェクト・パラメータは、

$$OLD_i = D_{OLD}(i, l, m), IOC_{i,j} = D_{IOC}(i, j, l, m)$$

として、得られる。

## 【0325】

## 20.2.1.3 ダウンミックス行列

入力オーディオ信号 S に適用されるダウンミックス行列 D は、 $X = DS$  としてダウンミックス信号を決定する。サイズ  $N_{dmx} \times N$  のダウンミックス行列 D は、

20

$$\mathbf{D} = \mathbf{D}_{dmx} \mathbf{D}_{premix}$$

として、得られる。

## 【0326】

行列  $\mathbf{D}_{dmx}$  および行列  $\mathbf{D}_{premix}$  は、処理モードに応じて異なるサイズを有する。行列  $\mathbf{D}_{dmx}$  は、

30

$$d_{i,j} = \begin{cases} 0 & , \text{ if no DMG data for (i,j) is present in the bitstream} \\ 10^{0.05DMG_{i,j}} & , \text{ otherwise} \end{cases}$$

として、DMG パラメータから得られる。

## 【0327】

ここで、逆量子化ダウンミックス・パラメータは、

40

$$DMG_{i,j} = D_{DMG}(i, j, l)$$

として、得られる。

## 【0328】

## 20. 2. 1. 3. 1. ダイレクト・モード

ダイレクト・モードの場合において、プレミキシングは使用されない。行列  $\mathbf{D}_{\text{premix}}$  は、サイズ  $N \times N$  を有し、 $\mathbf{D}_{\text{premix}} = \mathbf{I}$  によって与えられる。行列  $\mathbf{D}_{\text{dmx}}$  は、サイズ  $N_{\text{dmx}} \times N$  を有し、20. 2. 1. 3 によるDMGパラメータから得られる。

【0329】

## 20. 2. 1. 3. 2 プレミキシング・モード

プレミキシング・モードの場合において、行列  $\mathbf{D}_{\text{premix}}$  は、サイズ  $(N_{\text{ch}} + N_{\text{premix}}) \times N$  を有し、

10

$$\mathbf{D}_{\text{premix}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{pmatrix}$$

によって与えられる。

ここで、サイズ  $N_{\text{premix}} \times N_{\text{obj}}$  のプレミキシング行列  $\mathbf{A}$  は、オブジェクト・レンダラから 20  
、SAOC 3Dデコーダへの入力として受信される。

【0330】

行列  $\mathbf{D}_{\text{dmx}}$  は、サイズ  $N_{\text{dmx}} \times (N_{\text{ch}} + N_{\text{premix}})$  を有し、20. 2. 1. 3 によるDMGパラ  
メータから得られる。

【0331】

## 20. 2. 1. 4 レンダリング行列

入力オーディオ信号  $\mathbf{S}$  に適用されるレンダリング行列  $\mathbf{R}$  は、 $\mathbf{Y} = \mathbf{R} \mathbf{S}$  としてターゲット  
トレンダリングされた出力を決定する。サイズ  $N_{\text{out}} \times N$  のレンダリング行列  $\mathbf{R}$  は、

30

$$\mathbf{R} = (R_{\text{c}, \text{h}} \quad R_{\text{o}, \text{b}, \text{j}})$$

によって与えられる。

ここで、サイズ  $N_{\text{out}} \times N_{\text{ch}}$  の  $R_{\text{c}, \text{h}}$  は、入力チャネルに関連したレンダリング行列を表  
し、サイズ  $N_{\text{out}} \times N_{\text{obj}}$  の  $R_{\text{o}, \text{b}, \text{j}}$  は、入力オブジェクトに関連したレンダリング行列を表  
す。

40

【0332】

## 20.2.1.4 ターゲット出力共分散行列

要素  $c_{i,j}$  を有するサイズ  $N_{out} \times N_{out}$  の共分散行列  $C$  は、ターゲット出力共分散行列  $C \approx YY^*$  の近似で表し、そして、

$$C = R E R^*$$

として、共分散行列  $E$  およびレンダリング行列  $R$  から得られる。

【0 3 3 3】

### 20.2.2 復号化

S A O C 3 D パラメータおよびレンダリング情報を用いて出力信号を得るための方法が記載される。S A O C 3 D デコーダは、例えば、S A O C 3 D パラメータ・プロセッサおよびS A O C 3 D ダウンミックス・プロセッサから構成されうる。

【0 3 3 4】

### 20.2.2.1 ダウンミックス・プロセッサ

(ハイブリッドQMF領域において表される) ダウンミックス・プロセッサの出力信号は、S A O C 3 D デコーダの最後の出力を得ている、ISO / IEC 23003-1 : 2007において記載されるように、対応する合成フィルタバンクに供給される。ダウンミックス・プロセッサの詳細な構造は、図31において表現される。

【0 3 3 5】

出力信号  $\hat{Y}$  は、

$$\hat{Y} = P_{dry} R U X + P_{wet} M_{post} X_d$$

として、マルチチャネル・ダウンミックス信号  $X$  および非相関化マルチチャネル信号  $X_d$  から算出される。

ここで、 $U$  は、パラメトリック・アンミキシング行列を表し、20.2.2.1.1 および 20.2.2.1.2 において定義される。

【0 3 3 6】

非相関化マルチチャネル信号  $X_d$  は、

$$X_d = decorrFunc(M_{pre} Y_{dry})$$

として、20.2.3 により算出される。

【0 3 3 7】

混合行列  $P = (P_{dry} \quad P_{wet})$  は、20.2.3 において記載される。異なる出力構成のための行列  $M_{pre}$  は、図19～23において与えられ、以下の方程式を使用して得られる。

$$M_{post} = M_{pre}^* (M_{pre} M_{pre}^*)^{-1}$$

【0 3 3 8】

図32において示されるように、復号化モードは、ビットストリーム要素 b s N u m S a o c D m x O b j e s t s によって制御される。

【0 3 3 9】

10

20

20

30

40

50

## 2 0 . 2 . 2 . 1 . 1 結合復号化モード

結合復号化モードの場合において、パラメトリック・アンミキシング行列  $U$  は、

$$U = E D^* J$$

によって与えられる。

【 0 3 4 0 】

サイズ  $N_{\text{dmx}} \times N_{\text{dmx}}$  の行列  $J$  は、 $\Delta = D E D^*$  を伴い  $J \approx \Delta^{-1}$  によって与えられる。

10

【 0 3 4 1 】

## 2 0 . 2 . 2 . 1 . 2 独立復号化モード

独立復号化モードの場合において、アンミキシング行列  $U$  は、

$$U = \begin{pmatrix} U_{\text{ch}} & 0 \\ 0 & U_{\text{obj}} \end{pmatrix}$$

によって与えられる。

ここで、 $U_{\text{ch}} = E_{\text{ch}} D_{\text{ch}}^* J_{\text{ch}}$  および  $U_{\text{obj}} = E_{\text{obj}} D_{\text{obj}}^* J_{\text{obj}}$  である。

20

【 0 3 4 2 】

サイズ  $N_{\text{ch}} \times N_{\text{ch}}$  のチャネル・ベース共分散行列  $E_{\text{ch}}$  およびサイズ  $N_{\text{obj}} \times N_{\text{obj}}$  のオブジェクト・ベース共分散行列  $E_{\text{obj}}$  が、対応する対角上のブロックのみを選択することによって、共分散行列  $E$  から得られる。

$$E = \begin{pmatrix} E_{\text{ch}} & E_{\text{ch,obj}} \\ E_{\text{obj,ch}} & E_{\text{obj}} \end{pmatrix}$$

30

ここで、行列  $E_{\text{ch}, \text{obj}} = (E_{\text{obj}, \text{ch}})^*$  は、入力チャネルと入力オブジェクトとの間における相互共分散行列を表し、算出されることを必要としない。

【 0 3 4 3 】

サイズ  $N_{\text{ch}}^{\text{dmx}} \times N_{\text{ch}}$  のチャネル・ベース・ダウンミックス行列  $D_{\text{ch}}$  およびサイズ  $N_{\text{obj}}^{\text{dmx}} \times N_{\text{obj}}$  のオブジェクト・ベース・ダウンミックス行列  $D_{\text{obj}}$  が、対応する対角上のブロックのみを選択することによって、ダウンミックス行列  $D$  から得られる。

40

$$D = \begin{pmatrix} D_{\text{ch}} & 0 \\ 0 & D_{\text{obj}} \end{pmatrix}$$

【 0 3 4 4 】

サイズ  $N_{\text{ch}}^{\text{dmx}} \times N_{\text{ch}}^{\text{dmx}}$  の行列  $\mathbf{J}_{\text{ch}} \approx (\mathbf{D}_{\text{ch}} \mathbf{E}_{\text{ch}} \mathbf{D}_{\text{ch}}^*)^{-1}$  は、

$$\Delta = \mathbf{D}_{\text{ch}} \mathbf{E}_{\text{ch}} \mathbf{D}_{\text{ch}}^*$$

に対して、20. 2. 2. 1. 4 に従って導出される。

【0 3 4 5】

サイズ  $N_{\text{obj}}^{\text{dmx}} \times N_{\text{obj}}^{\text{dmx}}$  の行列  $\mathbf{J}_{\text{obj}} \approx (\mathbf{D}_{\text{obj}} \mathbf{E}_{\text{obj}} \mathbf{D}_{\text{obj}}^*)^{-1}$  は、

10

$$\Delta = \mathbf{D}_{\text{obj}} \mathbf{E}_{\text{obj}} \mathbf{D}_{\text{obj}}^*$$

に対して、20. 2. 2. 1. 4 により導出される。

【0 3 4 6】

20. 2. 2. 1. 4 行列  $\mathbf{J}$  の算出

行列  $\mathbf{J} \approx \Delta^{-1}$  は、以下の方程式を使用して算出される。

20

$$\mathbf{J} = \mathbf{V} \Lambda^{inv} \mathbf{V}^*$$

【0 3 4 7】

ここで、行列  $\Lambda$  の特異ベクトル  $v$  は、以下の特性方程式を使用して得られる。

$$\mathbf{V} \Lambda \mathbf{V}^* = \Delta$$

【0 3 4 8】

対角特異値行列  $\Lambda$  の正則化の逆  $\Lambda^{inv}$  は、

30

$$\lambda_{i,j}^{inv} = \begin{cases} \frac{1}{\lambda_{i,j}}, & \text{if } i=j \text{ and } \lambda_{i,j} \geq T_{reg}^\Lambda \\ 0, & \text{otherwise} \end{cases}$$

として、算出される。

【0 3 4 9】

相対的な正則化スカラー  $T_{reg}^\Lambda$  は、

40

$$T_{reg}^\Lambda = \max(\lambda_{i,i}) T_{reg}, \quad T_{reg} = 10^{-2}$$

として、絶対の閾値  $T_{reg}$  および  $\Lambda$  の最大値を使用して決定される。

【0 3 5 0】

### 20.2.3. 非相関化

非相関化信号  $\mathbf{X}_d$  は、図 19～24 における表に従って、 $b_s D e c o r r C o n f_i g == 0$  および非相関器インデックス  $X$  を伴って、ISO/IEC 23003-1:2007 の 6.6.2において記載される非相関器から生成される。これゆえ、 $decorrFunc()$  は、非相関化処理を意味する。

$$\mathbf{X}_d = decorrFunc(\mathbf{M}_{pre} \mathbf{Y}_{dry})$$

【0351】

10

### 20.2.4. 混合行列 P - 第 1 のオプション

混合行列  $P = (P_{dry} \quad P_{wet})$  の算出は、ビットストリーム要素  $b_s D e c o r r e l a t i o n M e t h o d$  によって制御される。行列  $P$  は、サイズ  $N_{out} \times 2N_{out}$  を有し、 $P_{dry}$  および  $P_{wet}$  は、いずれもサイズ  $N_{out} \times N_{out}$  を有する。

【0352】

### 20.2.4.1 エネルギー補償モード

エネルギー補償モードは、パラメトリック再構成においてエネルギーの損失に対して補償するために、非相関化信号を使用する。混合行列  $P_{dry}$  および  $P_{wet}$  は、

$$\mathbf{P}_{dry} = \mathbf{I}$$

$$P_{i,j}^{wet} = \begin{cases} \sqrt{\min\left(\lambda_{Dec}, \max\left(0, \frac{\mathbf{C}(i,i) - \mathbf{E}_Y^{dry}(i,i)}{\max(\varepsilon, \mathbf{E}_Y^{wet}(i,i))}\right)\right)} & i = j, \\ 0 & i \neq j. \end{cases}$$

によって、与えられる。

ここで、 $Dec = 4$  は、出力信号に加えられる非相関化要素の量を制限するために使用される定数である。

【0353】

30

## 2 0 . 2 . 4 . 2 制限された共分散調整モード

制限された共分散調整モードは、混合非相関化信号  $\mathbf{P}_{\text{wet}} \mathbf{Y}_{\text{dry}}$  の共分散行列が、異なる共分散行列  $\Delta_E$  :  $\mathbf{P}_{\text{wet}} \mathbf{E}_Y^{\text{wet}} \mathbf{P}_{\text{wet}}^* \approx \Delta_E$  に近似することを確実にする。混合行列  $\mathbf{P}_{\text{dry}}$  および  $\mathbf{P}_{\text{wet}}$  は、以下の方程式を使用して定義される。

$$\mathbf{P}_{\text{dry}} = \mathbf{I}$$

$$\mathbf{P}_{\text{wet}} = (\mathbf{V}_1 \sqrt{\mathbf{Q}_1} \mathbf{V}_1^*) (\mathbf{V}_2 \sqrt{\mathbf{Q}_2^{\text{inv}}} \mathbf{V}_2^*)$$

10

ここで、対角特異値行列  $\mathbf{Q}_2$  の正則化の逆  $\mathbf{Q}_2^{\text{inv}}$  は、

$$\mathbf{Q}_2^{\text{inv}}(i, j) = \begin{cases} \frac{1}{\mathbf{Q}_2(i, j)}, & \text{if } i = j \text{ and } \mathbf{Q}_2(i, j) \geq T_{\text{reg}}^\Lambda, \\ 0, & \text{otherwise,} \end{cases}$$

20

として、算出される。

【 0 3 5 4 】

相対的な正則化スカラー  $T_{\text{reg}}^\Lambda$  は、

$$T_{\text{reg}}^\Lambda = \max(\mathbf{Q}_2^{\text{inv}}(i, i)) T_{\text{reg}}, \quad T_{\text{reg}} = 10^{-2}$$

30

として、絶対の閾値  $T_{\text{reg}}$  および  $\mathbf{Q}_2^{\text{inv}}$  の最大値を使用して決定される。

【 0 3 5 5 】

行列  $\Delta_E$  は、

$$\Delta_E = \mathbf{V}_1 \mathbf{Q}_1 \mathbf{V}_1^*$$

として、特異値非相関化を使用して分解される。

【 0 3 5 6 】

40

非相関化信号  $\mathbf{E}_Y^{\text{wet}}$  の共分散行列は、特異値非相関化を使用しても表現される。

$$\mathbf{E}_Y^{\text{wet}} = \mathbf{V}_2 \mathbf{Q}_2 \mathbf{V}_2^*$$

【 0 3 5 7 】

## 20. 2. 4. 3. 一般の共分散調整モード

一般の共分散調整モードは、最終の出力信号  $\hat{\mathbf{Y}}$  ( $\mathbf{E}_{\hat{\mathbf{Y}}} = \hat{\mathbf{Y}}\hat{\mathbf{Y}}^*$ ) の共分散行列が、ターゲット共分散行列： $\mathbf{E}_{\hat{\mathbf{Y}}} \approx \mathbf{C}$  に近似することを確実にする。混合行列  $\mathbf{P}$  は、以下の方程式を使用して定義される。

$$\mathbf{P} = (\mathbf{V}_1 \sqrt{\mathbf{Q}_1} \mathbf{V}_1^*) \mathbf{H} (\mathbf{V}_2 \sqrt{\mathbf{Q}_2^{inv}} \mathbf{V}_2^*)$$

10

ここで、対角特異値行列  $\mathbf{Q}_2$  の正則化の逆  $\mathbf{Q}_2^{inv}$  は、

$$\mathbf{Q}_2^{inv}(i, j) = \begin{cases} \frac{1}{\mathbf{Q}_2(i, j)}, & \text{if } i = j \text{ and } \mathbf{Q}_2(i, j) \geq T_{reg}^\Lambda, \\ 0, & \text{otherwise,} \end{cases}$$

として、算出される。

20

## 【0358】

相対的な正則化スカラー  $T_{reg}^\Lambda$  は、

$$T_{reg}^\Lambda = \max(\mathbf{Q}_2^{inv}(i, i)) T_{reg}, \quad T_{reg} = 10^{-2}$$

として、絶対の閾値  $T_{reg}$  および  $\mathbf{Q}_2^{inv}$  の最大値を使用して決定される。

30

## 【0359】

ターゲット共分散行列  $\mathbf{C}$  は、

$$\mathbf{C} = \mathbf{V}_1 \mathbf{Q}_1 \mathbf{V}_1^*$$

として、特異値非相關化を使用して分解される。

## 【0360】

結合信号  $\mathbf{E}_{\mathbf{Y}}^{com}$  の共分散行列は、特異値非相關化を使用しても表現される。

40

$$\mathbf{E}_{\mathbf{Y}}^{com} = \mathbf{V}_2 \mathbf{Q}_2 \mathbf{V}_2^*$$

## 【0361】

行列  $\mathbf{H}$  は、サイズ  $(N_{\text{out}} \times 2N_{\text{out}})$  の原型の重み行列を表し、以下の方程式によって与えられる。

$$\mathbf{H} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{2}} & \cdots & 0 & 0 & \frac{1}{\sqrt{2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & \frac{1}{\sqrt{2}} \end{pmatrix}$$

10

【 0 3 6 2 】

20. 2. 4. 4 導入される共分散行列

行列  $\Delta_E$  は、ターゲット出力共分散行列  $\mathbf{C}$  とパラメータ的に再構成信号の共分散行列  $\mathbf{E}_Y^{\text{dry}}$  との間における差を表し、

$$\Delta_E = \mathbf{C} - \mathbf{E}_Y^{\text{dry}}$$

20

によって、得られる。

【 0 3 6 3 】

行列  $\mathbf{E}_Y^{\text{dry}}$  は、パラメータ的に推定された信号  $\mathbf{E}_Y^{\text{dry}} \approx \mathbf{Y}_{\text{dry}} \mathbf{Y}_{\text{dry}}^*$  の共分散行列を表し、以下の方程式を使用して定義される。

$$\mathbf{E}_Y^{\text{dry}} = \mathbf{R} \mathbf{U} \mathbf{E} \mathbf{U}^* \mathbf{R}^*$$

【 0 3 6 4 】

行列  $\mathbf{E}_Y^{\text{wet}}$  は、非相関化信号  $\mathbf{E}_Y^{\text{wet}} \approx \mathbf{Y}_{\text{wet}} \mathbf{Y}_{\text{wet}}^*$  の共分散行列を表し、以下の方程式を使用して定義される。

30

$$\mathbf{E}_Y^{\text{wet}} = \mathbf{M}_{\text{post}} \left[ \text{matdiag}(\mathbf{M}_{\text{pre}} \mathbf{E}_Y^{\text{dry}} \mathbf{M}_{\text{pre}}^*) \right] \mathbf{M}_{\text{post}}^*$$

【 0 3 6 5 】

パラメータ的に推定され、そして非相関化信号の結合から構成される信号  $\mathbf{Y}_{\text{com}}$  を考慮する。

$$\mathbf{Y}_{\text{com}} = \begin{pmatrix} \mathbf{Y}_{\text{dry}} \\ \mathbf{Y}_{\text{wet}} \end{pmatrix}$$

$\mathbf{Y}_{\text{com}}$  の共分散行列は、以下の方程式によって定義される。

10

$$\mathbf{E}_{\mathbf{Y}}^{\text{com}} = \begin{pmatrix} \mathbf{E}_{\mathbf{Y}}^{\text{dry}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\mathbf{Y}}^{\text{wet}} \end{pmatrix}$$

### 【 0 3 6 6 】

混合行列  $\mathbf{P}_{\text{wet}}$  が適用された後、行列  $\hat{\mathbf{E}}_{\mathbf{Y}}^{\text{wet}}$  は、例えば、非相関化信号の推定された共分散行列を表し、以下の方程式を使用して定義される。

20

$$\hat{\mathbf{E}}_{\mathbf{Y}}^{\text{wet}} = \mathbf{P}_{\text{wet}} \mathbf{E}_{\mathbf{Y}}^{\text{wet}} \mathbf{P}_{\text{wet}}^*$$

### 【 0 3 6 7 】

#### 20. 2. 5 混合行列 $\mathbf{P}$ - 第 2 のオプション

混合行列  $\mathbf{P} = [\mathbf{P}_{\text{dry}} \quad \mathbf{A}_{\text{wet}} \mathbf{P}_{\text{wet}}]$  の算出は、ビットストリーム要素  $b_{\text{sD e c o r r e l a t i o n M e t h o d}}$  によって制御される。行列  $\mathbf{P}$  は、サイズ  $N_{\text{out}} \times 2N_{\text{out}}$  を有し、行列  $\mathbf{P}_{\text{dry}}$  および  $\mathbf{P}_{\text{wet}}$  は、いずれもサイズは  $N_{\text{out}} \times N_{\text{out}}$  を有する。サイズ  $N_{\text{out}} \times N_{\text{out}}$  の制限行列  $\mathbf{A}_{\text{wet}}$  は、

30

$$\mathbf{A}_{\text{wet}} = \text{matdiag} \left( \min \left( 1, \sqrt{\max \left( 0, \lambda_{\text{Dec}} \frac{\mathbf{E}_{\mathbf{Y}}^{\text{dry}}(i,i)}{\max(\varepsilon, \hat{\mathbf{E}}_{\mathbf{Y}}^{\text{wet}}(i,i))} \right)} \right) \right)$$

によって、与えられる。

40

ここで、共分散行列  $\mathbf{E}_{\mathbf{Y}}^{\text{dry}}$ 、 $\mathbf{E}_{\mathbf{Y}}^{\text{wet}}$  および  $\hat{\mathbf{E}}_{\mathbf{Y}}^{\text{wet}}$  は、例えば、セクション 20. 2. 4. 4 において与えられ、 $\lambda_{\text{Dec}} = 4$  は、出力信号に加えられる非相関化要素の量を制限するために使用される定数である。

### 【 0 3 6 8 】

## 20.2.5.1 エネルギー補償モード

エネルギー補償モードは、パラメトリック再構成においてエネルギーの損失に対して補償するために、非相関化信号を使用する。混合行列  $\mathbf{P}_{\text{dry}}$  および  $\mathbf{P}_{\text{wet}}$  は、

$$\mathbf{P}_{\text{dry}} = \mathbf{I}$$

$$p_{i,j}^{\text{wet}} = \begin{cases} \sqrt{\max\left(0, \frac{\mathbf{C}(i,i) - \mathbf{E}_Y^{\text{dry}}(i,i)}{\max(\varepsilon, \mathbf{E}_Y^{\text{wet}}(i,i))}\right)} & i=j, \\ 0 & i \neq j. \end{cases}$$
10

によって、与えられる。

### 【0369】

#### 20.2.5.2 更なる概念および詳細

更なる概念および付加的な詳細に関して、参照は、セクション 20.2.4.2 ~ 20.2.4.4 にもなされる。

### 【0370】

#### 20.3 表記法に関する見解

異なる表記法が本出願の範囲内で用いられる点に留意する点に留意されたい。一方、特定の方程式に適用される表記法は、前後関係から明確である。

### 【0371】

例えば、混合行列は、明細書の他の部分において  $\mathbf{P}$  によって示される一方、混合行列は、明細書のいくつかの部分において、 $\mathbf{F}$  または  $\tilde{\mathbf{F}}$  によって示される。

### 【0372】

さらに、ドライ信号（または複数のドライ信号）に適用される混合行列の構成要素は、明細書の一部において  $\mathbf{P}$  によって示され、そして、明細書の他の部分において  $\mathbf{P}_{\text{dry}}$  で示される。同様に、ウェット信号（または複数のウェット信号）に適用される混合行列の構成要素は、明細書の一部において  $\mathbf{M}$  によって示され、明細書の他の部分において  $\mathbf{P}_{\text{wet}}$  で示される。さらに、（行列  $\mathbf{M}$  を伴う混合ステップの前に）ウェット信号の共分散行列  $\mathbf{E}_W$  は、

非相関化信号の共分散行列  $\mathbf{E}_Y^{\text{wet}}$  に等しい。

### 【0373】

#### 21. 実施変形例

いくつかの態様が、装置との関連で記載されるが、これらの態様も、対応する方法の説明を表わすことは明らかであり、ロックあるいは装置は、方法のステップ、または方法のステップの特徴に対応する。類似して、方法のステップとの関連で記載される態様は、装置に対応する、ロック、アイテムまたは特徴の説明を表す。方法のステップのいくつかまたは全ては、たとえば、マイクロプロセッサ、プログラム可能なコンピュータ、または電子回路のようなハードウェアによって（または使用して）実行されうる。いくつかの実施の形態において、最も重要な方法のステップの1つ以上は、この種の装置によって実行されうる。

### 【0374】

本発明の符号化された音声信号は、デジタル記憶媒体に保存されるか、または、ワイヤレス伝送媒体または例えば、インターネットのような有線の伝送媒体のような伝送媒体上

10

20

30

40

50

に送信されうる。

【0375】

特定の実現要求に応じて、本発明の実施の形態は、ハードウェアにおいて、または、ソフトウェアにおいて、実行されうる。その実現態様は、それぞれの方法が実行されるよう、プログラミング可能なコンピュータ・システムと協働するか（または、協働することができる）、そこに格納された電子的に読み込み可能な制御信号を有するデジタル記憶媒体、例えば、フロッピー（登録商標）ディスク、DVD、CD、ROM、EPROM、EEPROM、EEPROM、またはFLASHメモリを使用して実行されうる。従って、デジタル記憶媒体は、コンピュータ読み込み可能でもよい。

【0376】

本発明による若干の実施の形態は、本願明細書において記載される方法のうちの1つが実行されるように、プログラミング可能なコンピュータ・システムと協働することができる電子的に読み込み可能な信号を有するデータキャリアを含む。

【0377】

通常、本発明の実施の形態は、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実施され、コンピュータ・プログラム製品がコンピュータ上で実行する場合、プログラムコードは、方法のうちの1つを実行するために作動される。プログラムコードは、機械可読キャリアに、例えば、格納されうる。

【0378】

他の実施の形態は、機械可読キャリアに格納され、本願明細書において記載される方法のうちの1つを実行するためのコンピュータ・プログラムを含む。

【0379】

換言すれば、従って、コンピュータ・プログラムがコンピュータ上で実行する場合、本発明の方法の実施の形態は、本願明細書において記載される方法のうちの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。

【0380】

従って、本発明の方法の更なる実施の形態は、その上に記録され、本願明細書において記載される方法のうちの1つを実行するためのコンピュータ・プログラムを含むデータキャリア（または、デジタル記憶媒体、またはコンピュータ可読媒体）である。データキャリア、デジタル記憶媒体または記録された媒体は、一般的には、有形でありおよび／または、暫定的である。

【0381】

従って、本発明の方法の更なる実施の形態は、本願明細書において記載される方法のうちの1つを実行するためのコンピュータ・プログラムを表しているデータストリームまたは一連の信号である。例えば、データストリームまたは一連の信号は、データ通信接続、例えば、インターネットを介して転送されるように構成されうる。

【0382】

更なる実施の形態は、本願明細書において記載される方法のうちの1つを実行するため構成され、または適応される処理手段、例えば、コンピュータ、またはプログラミング可能な論理回路を含む。

【0383】

更なる実施の形態は、その上にインストールされ、本願明細書において記載される方法のうちの1つを実行するためのコンピュータ・プログラムを有するコンピュータを含む。

【0384】

本発明による更なる実施の形態は、レシーバに本願明細書に記載される方法のうちの1つを実行するためのコンピュータ・プログラムを（例えば、電子的にまたは光学的に）転送するために構成される装置またはシステムを含む。レシーバは、例えば、コンピュータ、モバイル機器、メモリ素子等でもよい。装置またはシステムは、例えば、レシーバにコンピュータ・プログラムを転送するためのファイルサーバを含む。

【0385】

10

20

30

40

50

いくつかの実施の形態において、プログラミング可能な論理回路（例えば、現場でプログラム可能なゲートアレイ（Field Programmable Gate Array））が、本願明細書において記載されるいくつかまたは全ての機能を実行するために使用される。いくつかの実施の形態において、現場でプログラム可能なゲートアレイは、本願明細書において記載される方法の1つを実行するために、マイクロプロセッサと協働しうる。一般に、方法は、いくつかのハードウェア装置によって、好ましくは実行される。

#### 【0386】

上述した実施の形態は、本発明の原則の例を表すだけである。本願明細書において記載される装置の修正および変更は、他の当業者にとって明らかであるものと理解される。従って、間近に迫った特許請求の範囲だけによってのみ制限され、ならびに、本願発明の記述および説明によって表された明細書の詳細な記載によっては、制限されない。

10

#### 【0387】

##### 参考文献

[BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.

[Blauert] J. Blauert, "Spatial Hearing - The Psychophysics of Human Sound Localization", Revised Edition, The MIT Press, London, 1997.

20

[JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006.

[ISS1] M. Parvaix and L. Girin: "Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding", IEEE ICASSP, 2010.

[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: "A watermarking-based method for informed source separation of audio signals with a single sensor", IEEE Transactions on Audio, Speech and Language Processing, 2010.

30

[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: "Informed source separation through spectrogram coding and data embedding", Signal Processing Journal, 2011.

[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: "Informed source separation : source coding meets source separation", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011.

[ISS5] S. Zhang and L. Girin: "An Informed Source Separation System for Speech Signals", INTERSPEECH, 2011.

40

[ISS6] L. Girin and J. Pinel: "Informed Audio Source Separation from Compressed Linear Stereo Mixtures", AES 42nd International Conference: Semantic Audio, 2011

[MPS] ISO/IEC, "Information technology - MPEG audio technologies - Part 1: MPEG Surround," ISO/IEC JTC1/SC29/WG11 (MPEG) international Standard 23003-1:2006.

[OCD] J. Vilkamo, T. Baeckstroem, and A. Kuntz. "Optimized covariance domain fra

50

mework for time-frequency processing of spatial audio", Journal of the Audio Engineering Society, 2013. in press.

[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007.

[SAOC2] J. Engdegard, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hoelzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: "Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008. 10

[SAOC] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2.

International Patent No. WO/2006/026452, "MULTICHANNEL DECORRELATION IN SPATIAL AUDIO CODING" issued on 9 March 2006.

【図1】

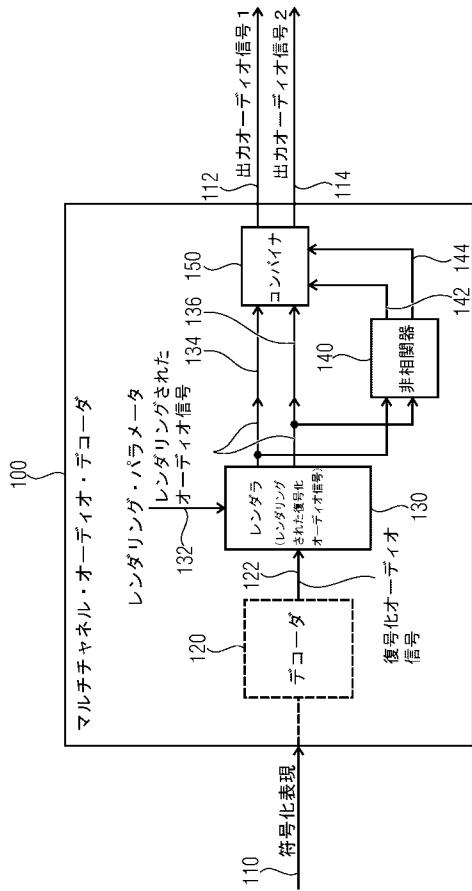


FIG 1

【図2】

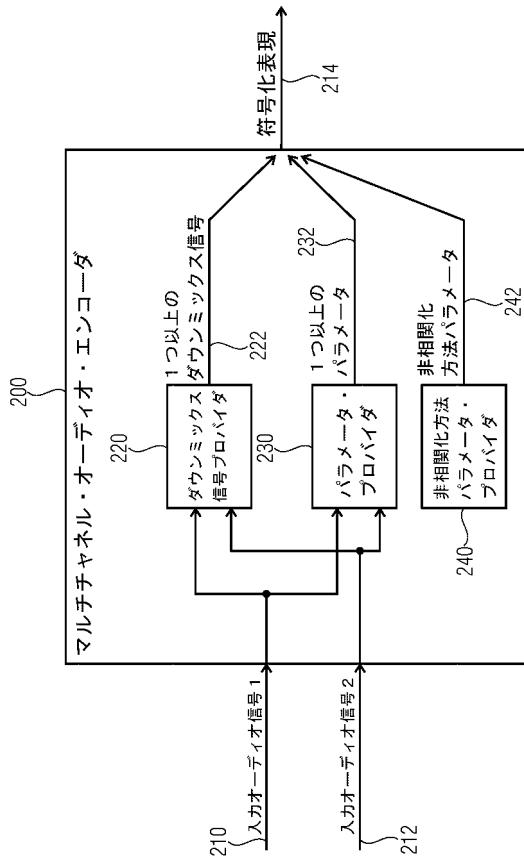


FIG 2

〔 図 3 〕

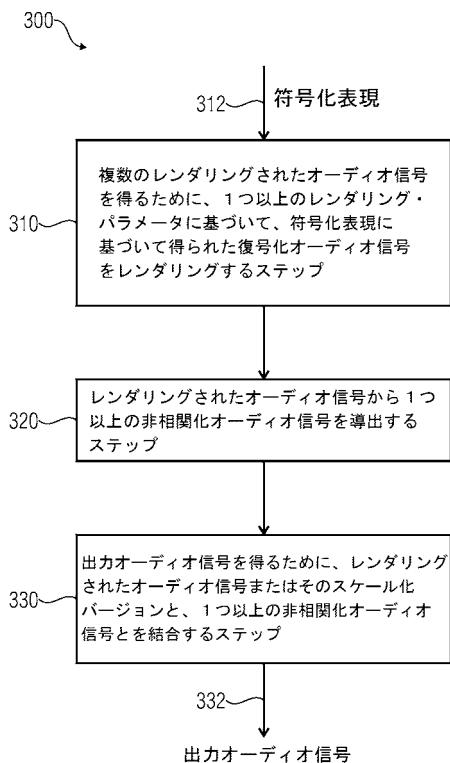


FIG 3

【圖 4】

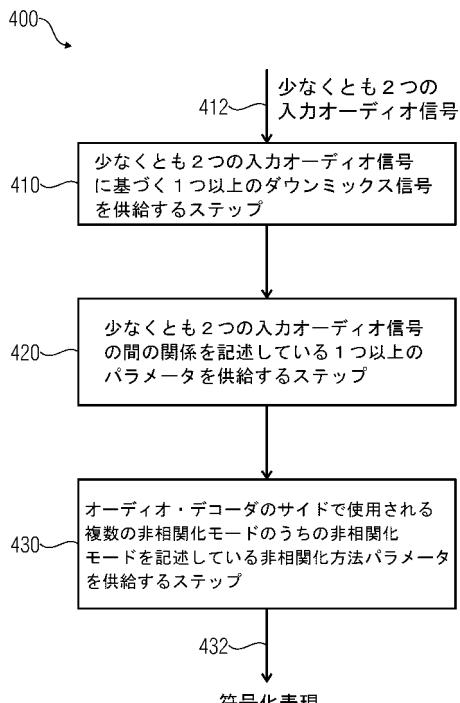


FIG 4

【 四 5 】

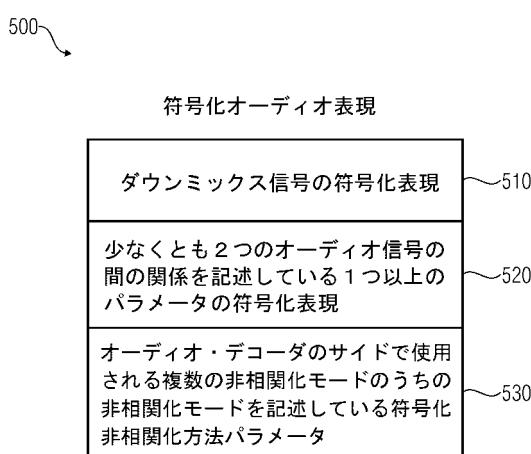


FIG 5

【 四 6 】

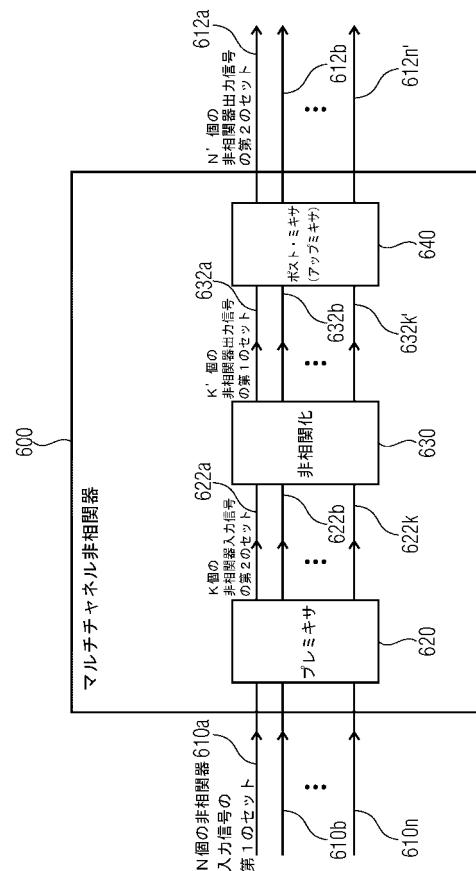


FIG 6

【図 7】

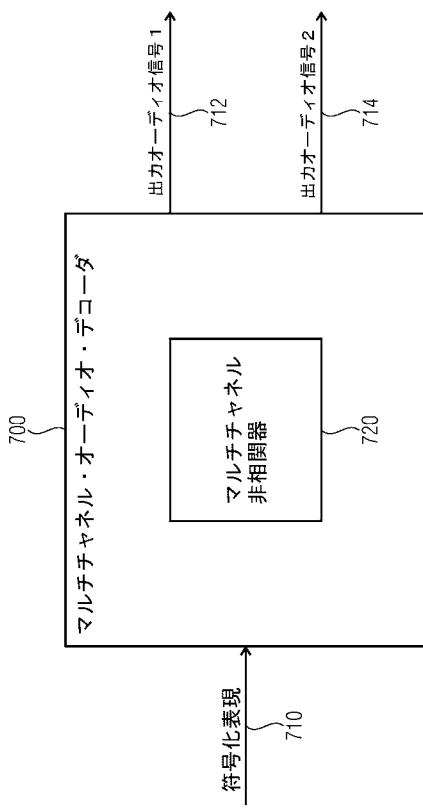


FIG 7

【図 8】

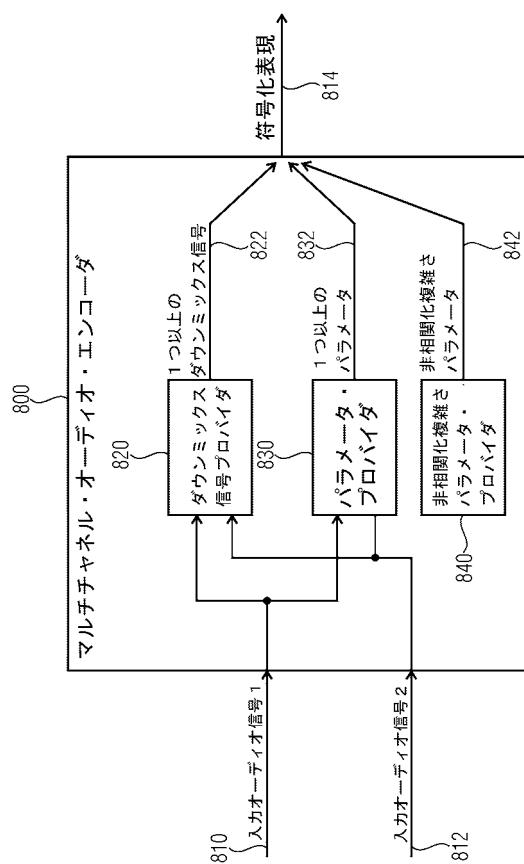


FIG 8

【図 9】

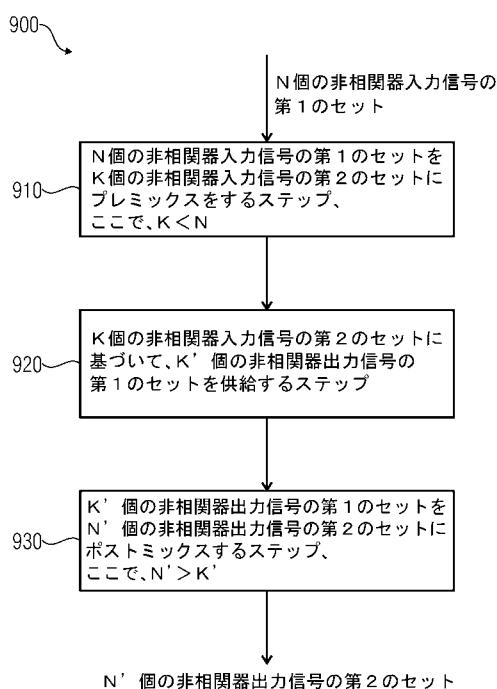


FIG 9

【図 10】

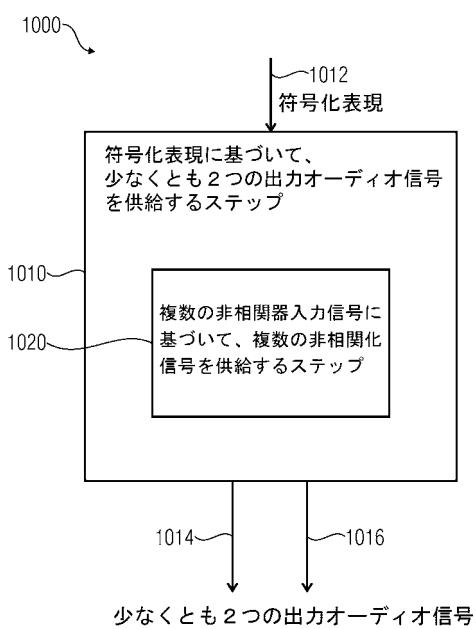


FIG 10

( 四 1 1 )

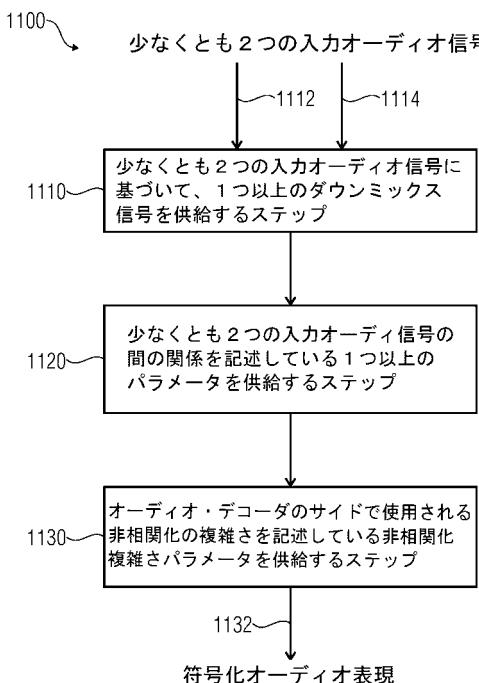


FIG 11

【 义 1 2 】

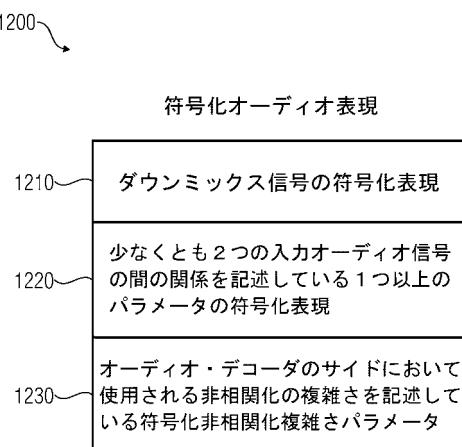


FIG 12

〔 図 1 3 〕

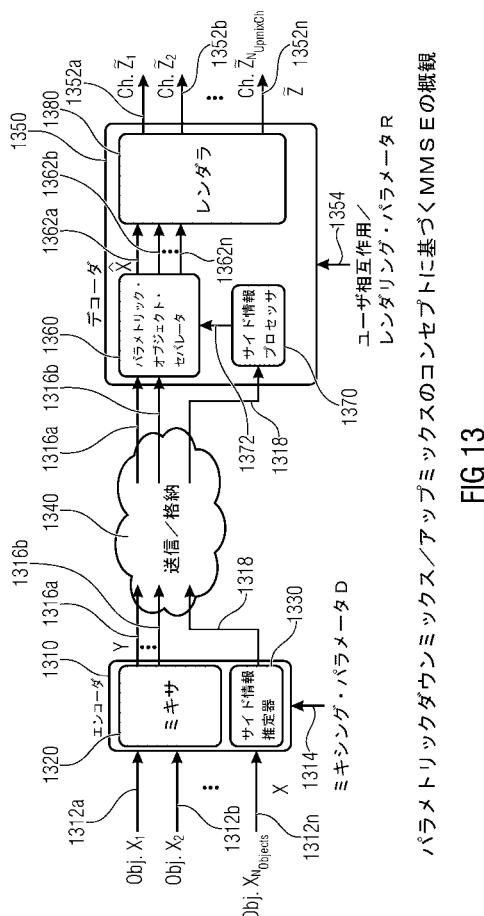
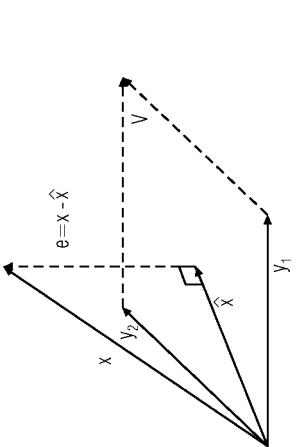


FIG 13

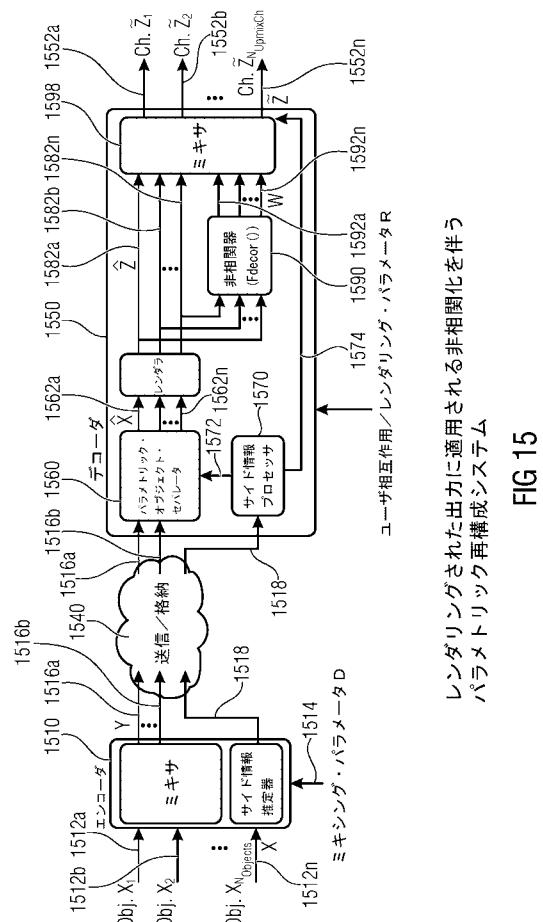
〔 四 1 4 〕



3次元空間における直角原理のための幾何学的な表現

FIG 14

【 図 1 5 】



〔 四 18 〕

#### スピーカ位置および出力フォーマット

No.	LS Label		Az. °	Az. Tol. °	El. °	El. Tol. °	0-2.0	0-5.1	0-7.1	0-8.1	0-10.1	0-22.2
	CH	M										
1	CH_M_000		0	±2	0	±2		3	3		3	3
2	CH_M_L030		30	±2	0	±2	1	1	1	1	1	7
3	CH_M_R030		-30	±2	0	±2	2	2	2	2	2	8
4	CH_M_L060		60	±2	0	±2						1
5	CH_M_R060		-60	±2	0	±2						2
6	CH_M_L090		90	±5	0	±2						11
7	CH_M_R090		-90	±5	0	±2						12
8	CH_M_L110		110	±5	0	±2		5	5	5	5	
9	CH_M_R110		-110	±5	0	±2		6	6	6	6	
10	CH_M_L135		135	±5	0	±2						5
11	CH_M_R135		-135	±5	0	±2						6
12	CH_M_180		180	±5	0	±2						9
13	CH_U_000		0	±2	35	±10			3			15
14	CH_U_L045		45	±5	35	±10						13
15	CH_U_R045		-45	±5	35	±10						14
16	CH_U_L030		30	±5	35	±10		7	7	7		
17	CH_U_R030		-30	±5	35	±10		8	8	8		
18	CH_U_L090		90	±5	35	±10						19
19	CH_U_R090		-90	±5	35	±10						20
20	CH_U_L110		110	±5	35	±10						9
21	CH_U_R110		-110	±5	35	±10						10
22	CH_U_L135		135	±5	35	±10						17
23	CH_U_R135		-135	±5	35	±10						18
24	CH_U_180		180	±5	35	±10						21
25	CH_T_000		0	±2	90	±10					11	16
26	CH_L_000		0	±2	-15	+5-25				9		22
27	CH_L_L045		45	±5	-15	+5-25						23
28	CH_L_R045		-45	±5	-15	+5-25						24
29	CH_LFE1		45	±15	-15	±15	4	4	4	4	4	
30	CH_LFE2		-45	±15	-15	±15						10

FIG 18

【 図 1 6 】

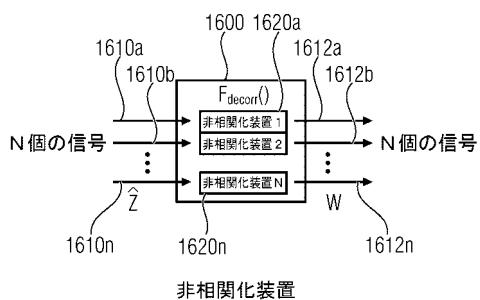
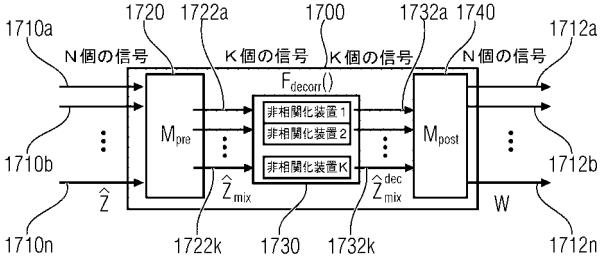


FIG 16

【 図 17 】



### 低減された複雑さの非相関化装置

FIG 17

【 义 1 9 a 】

## 2.2. 2出力フォーマットのための複雑さの低減

【図 1 6 q】

Ch.	ID	$M_{\text{pre}}^{\text{H}}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	

FIG 19B

【図 1 6 p】

Ch.	ID	$M_{\text{pre}}^{\text{H}}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	CH_U_R045	
2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	

FIG 19D

 $N=22$  and  $K=8$ ,  $\text{cond}(M_{\text{pre}} M_{\text{pre}}^{\text{H}}) = 2$ , のためのプレミキシング係数

【図 1 6 c】

Ch.	ID	$M_{\text{pre}}^{\text{H}}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	

FIG 19C

【図 1 6 e】

Ch.	ID	$M_{\text{pre}}^{\text{H}}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	

FIG 19E

【図 19f】

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_000	CH_L_000	CH_U_000	CH_T_000	CH_M_1135	CH_U_1135	CH_M_180	CH_U_180	CH_M_3030	CH_L_3030	CH_M_090	CH_U_090	CH_M_1090	CH_U_1090	CH_M_1060	CH_U_1060	CH_M_045	CH_U_045
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FIG 19F

【図 20a】

10. 1出力フォーマットのための複雑さの低減

N=10 and K=5, cond(M<sub>pre</sub>M<sub>pre</sub><sup>H</sup>)=1. のためのプレミキシング係数

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_030	CH_U_030	CH_M_R030	CH_U_R030	CH_M_000	CH_U_000	CH_M_L110	CH_U_L110	CH_M_R110	CH_U_R110
1	1	1	1	0	0	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0	0	0	0
3	0	0	0	0	1	1	0	0	0	0	0
4	0	0	0	0	0	0	0	0	1	1	0
5	0	0	0	0	0	0	0	0	0	0	1
											1

FIG 20A

【図 20b】

N=10 and K=4, cond(M<sub>pre</sub>M<sub>pre</sub><sup>H</sup>)=2. のためのプレミキシング係数

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_030	CH_U_030	CH_M_R030	CH_U_R030	CH_M_000	CH_U_000	CH_M_L110	CH_U_L110	CH_M_R110	CH_U_R110
1	1	1	1	0	0	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0	0	0	0
3	0	0	0	0	1	1	0	0	0	0	0
4	0	0	0	0	0	0	0	1	1	1	1
											1

FIG 20B

【図 19g】

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_045	CH_U_045	CH_M_060	CH_U_060	CH_M_090	CH_U_090	CH_M_1090	CH_U_1090	CH_M_1045	CH_U_1045	CH_M_1060	CH_U_1060	CH_M_045	CH_U_045
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	1	0	1	1	1	1	1	1	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
														1	1

FIG 19G

【図 20c】

N=10 and K=3, cond(M<sub>pre</sub>M<sub>pre</sub><sup>H</sup>)=2. のためのプレミキシング係数

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_000	CH_U_000								
1	1	1	1	1	1	0	0	0	0	0	0
2	0	0	0	0	0	1	1	1	0	0	0
3	0	0	0	0	0	0	0	0	1	1	1

FIG 20C

【図 20d】

N=10 and K=2, cond(M<sub>pre</sub>M<sub>pre</sub><sup>H</sup>)=1.5. のためのプレミキシング係数

Ch. ID	M <sub>pre</sub> <sup>i,j</sup>	CH_M_030	CH_U_030	CH_M_R030	CH_U_R030	CH_M_000	CH_U_000	CH_M_110	CH_U_110	CH_M_R110	CH_U_R110
1	1	1	1	1	1	1	1	1	0	0	0
2	0	0	0	0	0	0	0	0	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0

FIG 20D

【図 2 1 a】

8. 1出力フォーマットのための複雑さの低減

N=8 and K=4, cond( $M_{pre}M_{pre}^H$ )=1. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_T_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7	8
1	1	1	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0
3	0	0	0	0	1	1	0	0
4	0	0	0	0	0	0	1	1

FIG 21A

【図 2 1 b】

N=8 and K=3, cond( $M_{pre}M_{pre}^H$ )=2. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_T_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7	8
1	1	1	1	1	0	0	0	0
2	0	0	0	0	1	1	0	0
3	0	0	0	0	0	0	1	1

FIG 21B

【図 2 1 e】

N=7 and K=3, cond( $M_{pre}M_{pre}^H$ )=2. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7
1	1	1	1	1	0	0	0
2	0	0	0	0	1	0	0
3	0	0	0	0	0	1	1

FIG 21E

【図 2 1 f】

N=7 and K=2, cond( $M_{pre}M_{pre}^H$ )=2.5. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7
1	1	1	1	1	1	0	0
2	0	0	0	0	0	1	1

FIG 21F

【図 2 1 c】

N=8 and K=2, cond( $M_{pre}M_{pre}^H$ )=3. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_T_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	0
2	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	1	1

FIG 21C

【図 2 1 d】

7. 1出力フォーマットのための複雑さの低減

N=7 and K=4, cond( $M_{pre}M_{pre}^H$ )=1. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_U_L030	CH_M_R030	CH_U_R030	CH_M_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5	6	7
1	1	1	1	0	0	0	0
2	0	0	1	1	0	0	0
3	0	0	0	0	0	1	0
4	0	0	0	0	0	0	1

FIG 21D

【図 2 2 a】

5. 1出力フォーマットのための複雑さの低減

N=5 and K=3, cond( $M_{pre}M_{pre}^H$ )=2. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_M_R030	CH_M_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5
1	1	1	1	0	0
2	0	0	1	0	0
3	0	0	0	1	1

FIG 22A

【図 2 2 b】

N=5 and K=2, cond( $M_{pre}M_{pre}^H$ )=1.5. のためのプレミキシング係数

Ch. ID	CH_M_L030	CH_M_R030	CH_M_000	CH_M_L110	CH_M_R110
$M_{pre}^{ij}$	1	2	3	4	5
1	1	1	1	1	0
2	0	0	0	1	1

FIG 22B

【 図 2 3 】

## 2. O出力フォーマットのための複雑さの低減

N=2 and K=1,  $\text{cond}(M_{\text{pre}} M_{\text{pre}}^H) = 2$  のためのプレミキシング係数

Ch. D	CH_M_L030	CH_M_R030
M <sup>i,j</sup> pre	1	2
1	1	1

FIG 23

【 図 2 4 】

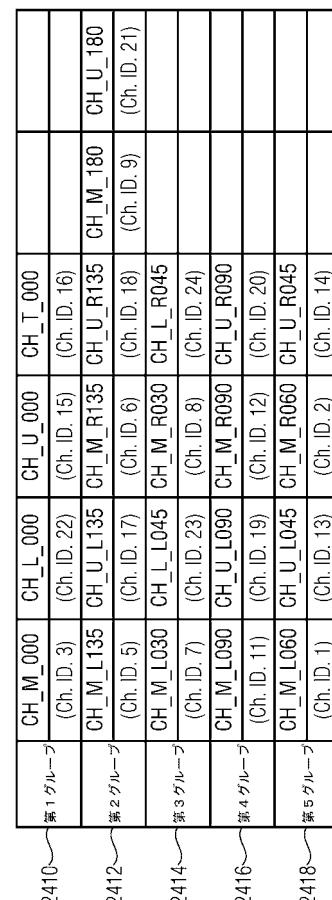


FIG 24

〔 図 25 〕

bsDecorrelationMethod;  
bsDecorrelationLevel:

2 uimsbt  
2 uimsbt

FIG 25

【図28】

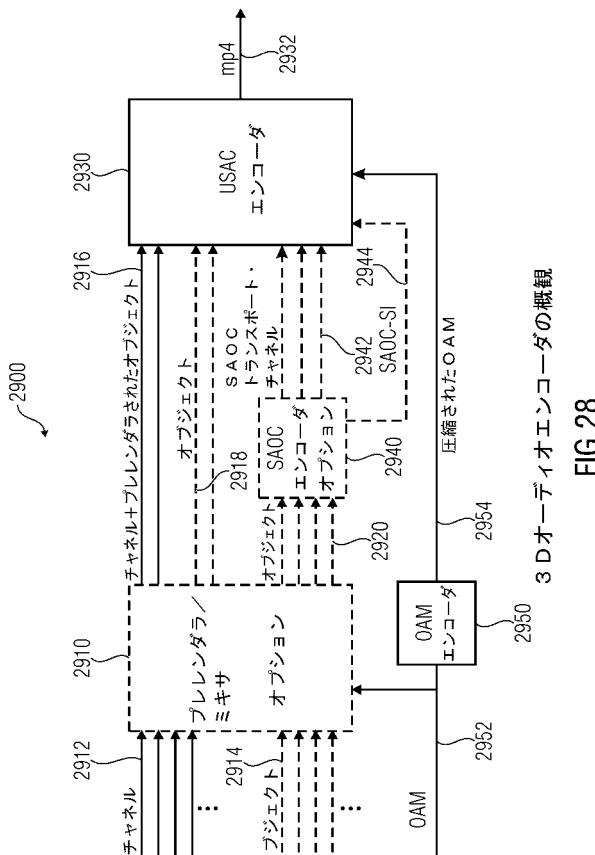


FIG 28

【图 2-7】

hsDecorrelationLevel

以下により、非相関化レベルを定義する

bsDecorrelationMethod	Meaning
0	Energy compensation mode
1	Limited covariance adjustment mode
2	General covariance adjustment mode
3	N/A

FIG 26

【图 2-7】

ational level

以下により、非相関化レベルを定

bsDecorrelationLevel	22.2	10.1	8.1	7.1	5.1	2.1
0	11	10	8	7	5	2
1	9	5	4	4	3	1
2	7	3	3	3	2	0
3	5	2	2	2	0	0

FIG 27

【 図 2 9 】

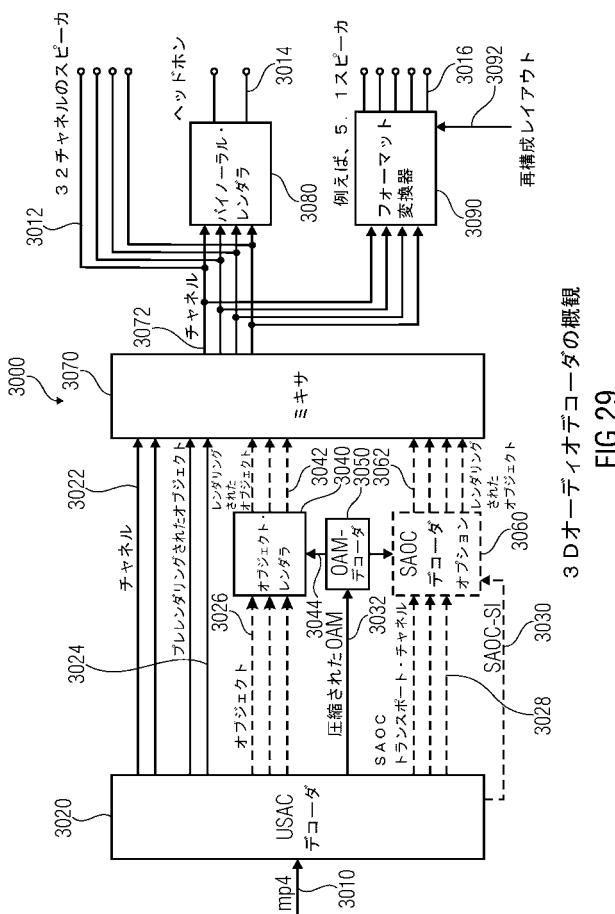


FIG 29

【 図 3 0 】

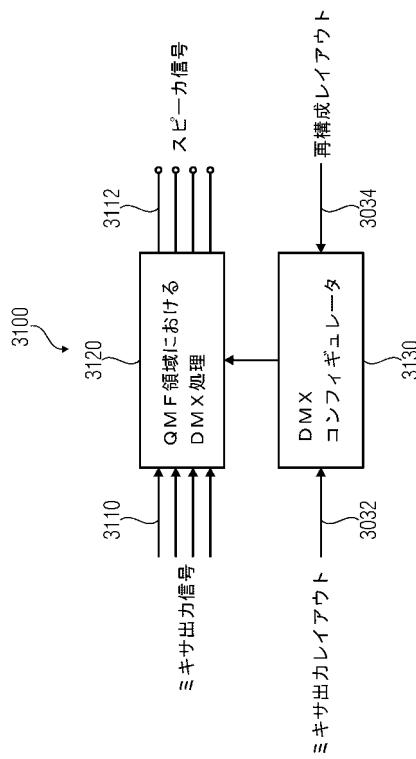


FIG 30 ノオニマット変換器の構造

FIG 30

〔 3 1 〕

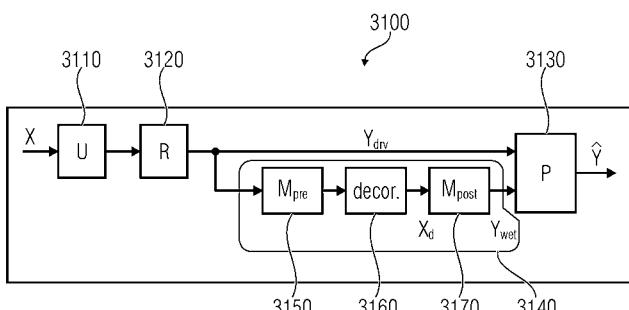


FIG. 31

〔図32〕

bsNumSaocDmxObjects	復号化モード	meaning
0	結合	信号に基づく入力チャネルおよび信号に基づく入力オブジェクトは、 $N_{ch}$ 個のチャネルに一緒にダウンミックスされる
$>=1$	独立	信号に基づく入力チャネルは $N_{ch}$ 個のチャネルにダウンミックスされる 信号に基づく入力オブジェクトは、 $N_{obj}$ 個のチャネルにダウンミックスされる

FIG 32

### 【図 3 3 a - 1】

### SAOC3DSpecificConfig()の構文

構文	ビット数	簡略記憶記号
SAC3DSpecificConfig()		
{		
bsSamplingFrequencyIndex;	4	uimsbf
if (bsSamplingFrequencyIndex == 15) {		
bsSamplingFrequency;	24	uimsbf
}		
bsFreqRes;	3	uimsbf
bsFrameLenth;	7	uimsbf
bsNumSaocDmxChannels;	5	uimsbf
bsNumSaocDmxObjects;	5	uimsbf
bsDecorrelationMethod;	2	uimsbf
NumInputSignals = 0;		
if (bsNumSaocDmxChannels < 0) {		
bsNumSaocChannels;	6	uimsbf
bsNumSaocLFEs;	2	uimsbf
NumInputSigns += bsNumSaocChannels;		
}		
bsNumSaocObjects;	8	uimsbf
bsDecorrelationLevel;	2	uimsbf
NumInputSignals += bsNumSaocObjects;		
for (i=0; i < bsNumSaocChannels; i++) {		
bsRelatedTo[i][j] = 1;		
for (j=i+1; j < bsNumSaocChannels; j++) {		
bsRelatedTo[i][j];	1	uimsbf
bsRelatedTo[j][i] = bsRelatedTo[i][j];		
}		
}		
}		

FIG 33A	FIG 33A-1
	FIG 33A-2

FIG 33A-1

【図 3 3 a - 2】

```

for ( i=bsNumSaocChannels; i<NumInputSignals; i++ ) {
    for ( j=0; j<bsNumSaocChannels; j++ ) {
        bsRelatedTo[i][j] = 0;
        bsRelatedTo[j][i] = 0;
    }
}
for ( i=bsNumSaocChannels; i<NumInputSignals; i++ ) {
    bsRelatedTo[i][i] = 1;
    for ( j=i+1; j<NumInputSignals; j++ ) {
        bsRelatedTo[i][j];
        bsRelatedTo[j][i] = bsRelatedTo[i][i];
    }
}
bsOneLOC;
bsSaocDmxMethod;
if (bsSaocDmxMethod == 15) {
    bsNumPremixedChannels;
}
bsDualMode;
if (bsDualMode) {
    bsBandsLow;
    bsBandsHigh = numBands;
} else {
    bsBandsLow = numBands;
}

```

【図 3 3 b】

}	1	uimsbf
bsDcuFlag;		
if ( bsDcuFlag == 1 ) {	1	uimsbf
bsDcuMandatory;	1	uimsbf
bsDcuDynamic;	1	uimsbf
if ( bsDcuDynamic == 0 ) {		
bsDcuMode;	1	uimsbf
bsDcuParam;	4	uimsbf
}		
} else {		
bsDcuMandatory = 0;		
bsDcuDynamic = 0;		
bsDcuMode = 0;		
bsDcuParam = 0;		
}		
bsSaocReserved;	3	uimsbf
ByteAlign();		
SAOC3DExtensionConfig();		

注釈 1 : numBand は、ISO/IEC 23003-2:2010 における表 3 3において定義される。

FIG 33B

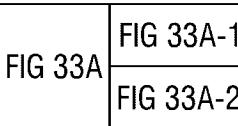


FIG 33A-2

## 【手続補正書】

【提出日】平成27年5月20日(2015.5.20)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

## 【補正の内容】

## 【特許請求の範囲】

## 【請求項1】

符号化表現 ( 1 1 0 ; 7 1 0 ; 1 5 1 6 a ; 1 5 1 6 b ; 1 5 1 8 ) に基づいて、少なくとも 2 つの出力オーディオ信号 ( 1 1 2 , 1 1 4 ; 7 1 2 , 7 1 4 ; 1 5 5 2 a - 1 5 5 2 n ; 3 0 1 2 ) を供給するためのマルチチャネル・オーディオ・デコーダ ( 1 0 0 ; 7 0 0 ; 1 5 5 0 ; 3 0 0 0 ) であって、

ここで、前記マルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号 ( 1 3 4 , 1 3 6 ; 1 5 8 2 a - 1 5 8 2 n ,  $\hat{Z}$  ) を得るために、レンダリング行列を定義する 1 つ以上のレンダリング・パラメータ ( 1 3 2 ) に基づいて、マルチチャネル・ターゲット・シーンに対して、前記符号化表現に基づいて得られた複数の復号化オーディオ信号 ( 1 2 2 ; 1 5 6 2 a - 1 5 6 2 n ,  $\hat{X}$  ) をレンダリングする ( 1 3 0 ; 1 5 8 0 ) ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号から 1 つ以上の非相關化オーディオ信号 ( 1 4 2 , 1 4 4 ; 1 5 9 2 a - 1 5 9 2 n ) を導出する ( 1 4 0 ; 1 5 9 0 ) ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する(150；1598)ように構成され、

前記マルチチャネル・オーディオ・デコーダは、パラメトリック再構成(120；1560)を使用して、前記複数のレンダリングされたオーディオ信号を得るためにレンダリングされた前記復号化されたオーディオ信号を得るように構成され、

前記復号化オーディオ信号は、再構成オブジェクト信号であり、

前記マルチチャネル・オーディオ・デコーダは、サイド情報(1518)を使用して、1つ以上のダウンミックス信号(1516a, 1516b)から前記再構成オブジェクト信号を導出するように構成される、マルチチャネル・オーディオ・デコーダ。

#### 【請求項2】

前記マルチチャネル・オーディオ・デコーダは、前記サイド情報からアンミキシング係数を導出し、前記アンミキシング係数を使用して、前記1つ以上のダウンミックス信号から前記再構成オブジェクト信号を導出するための前記アンミキシング係数を適用するように構成される、請求項1に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項3】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の所望の相関特性または共分散特性を少なくとも部分的に達成するように、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号とを結合するように構成される、請求項1または請求項2に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項4】

前記マルチチャネル・オーディオ・デコーダは、前記複数のレンダリングされたオーディオ信号を得るためにレンダリングされた前記復号化オーディオ信号(122；1562a-1562n)のパラメトリック再構成(120；1560)との間のエネルギー損失のために少なくとも部分的に補償するように、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号とを結合するように構成される、請求項1ないし請求項3のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項5】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の所望の相関特性または所望の共分散特性を決定するように構成され、

前記得られた出力オーディオ信号の相関特性または共分散特性が、前記所望の相関特性または所望の共分散特性(C)に近似するかまたは等しいように、前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号との結合(150；1598)を調整するように構成される、請求項1ないし請求項4のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項6】

前記マルチチャネル・オーディオ・デコーダは、前記複数のレンダリングされたオーディオ信号(2)を得るために、前記符号化表現に基づいて得られる前記複数の復号化オーディオ信号(2̂)をレンダリングすること(130；1560)を記述しているレンダリング情報(R)に基づいて、前記所望の相関特性または所望の共分散特性(C)を決定するように構成される、請求項5に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項7】

前記マルチチャネル・オーディオ・デコーダは、複数のオーディオ・オブジェクトの特性および/または複数のオーディオ・オブジェクトの間ににおける関係を記述しているオブジェクト相関情報またはオブジェクト共分散情報(E<sub>x</sub>)に基づいて、前記所望の相関特性または所望の共分散特性(C)を決定するように構成される、請求項5または請求項6

に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 8】**

前記マルチチャネル・オーディオ・デコーダは、前記符号化表現に含まれるサイド情報(1518)に基づいて、前記オブジェクト相關情報またはオブジェクト共分散情報( $E_x$ )を決定するように構成される、請求項7に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 9】**

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号および前記1つ以上の非相関化オーディオ信号の現実の相關特性または共分散特性( $E_s$ )を決定するように、および

前記レンダリングされたオーディオ信号および前記1つ以上の非相関化オーディオ信号の前記現実の相關特性または共分散特性( $E_s$ )に基づいて、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号との結合(150；1598)を調整するように構成される、請求項5ないし請求項8のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 10】**

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{P}\hat{\mathbf{Z}} + \mathbf{M}\mathbf{W}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ

信号 $\hat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

ここで、 $\mathbf{P}$ は、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ に適用される混合行列であり、 $\mathbf{M}$ は、前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ に適用される混合行列である、請求項1ないし請求項9のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 11】**

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{\mathbf{Z}}$ の相關特性または共分散特性( $E_{\tilde{\mathbf{Z}}}$ )が、前記所望の相關特性または所望の共分散特性( $C$ )に近似するかまたは等しいように、前記混合行列 $\mathbf{P}$ および前記混合行列 $\mathbf{M}$ の少なくとも1つを調整するように構成される、請求項1\_0に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 12】**

前記マルチチャネル・オーディオ・デコーダは、前記混合行列 $\mathbf{P}$ および前記混合行列 $\mathbf{M}$ と一緒に算出するように構成される、請求項1\_0または請求項1\_1に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項 13】**

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\hat{\mathbf{z}}$ の共分散行列 $\mathbf{E}_{\hat{\mathbf{z}}}$ が所望の共分散行列 $\mathbf{C}$ に近似するかまたは等しいように、

$$\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$$

で示される、結合混合行列 $\mathbf{F}$ を得るように構成される、請求項10ないし請求項12に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項14】**

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{E}_{\hat{\mathbf{z}}} = \mathbf{F}\mathbf{E}_s\mathbf{F}^H$$

で示される前記共分散行列が、

$$\mathbf{C} = \mathbf{R}\mathbf{E}_x\mathbf{R}^H$$

で示される前記所望の共分散行列と等しくなるように、前記結合混合行列 $\mathbf{F}$ を決定するよう構成され、

ここで、 $\mathbf{E}_s$ は、次式で定義される、前記レンダリングされたオーディオ信号 $\hat{\mathbf{z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合した信号 $\mathbf{S}$ の共分散行列であり、

$$\mathbf{S} = \begin{bmatrix} \hat{\mathbf{z}} \\ \mathbf{W} \end{bmatrix}$$

ここで、 $\mathbf{E}_x$ は、オブジェクト共分散行列である、請求項13に記載のマルチチャネル・オーディオ・デコーダ。

**【請求項15】**

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{A}_{dry} P \hat{\mathbf{Z}} + \mathbf{M} \mathbf{W}$$

または、

$$\tilde{\mathbf{Z}} = P \hat{\mathbf{Z}} + \mathbf{A}_{wet} \mathbf{M} \mathbf{W}$$

または、

$$\tilde{\mathbf{Z}} = \mathbf{A}_{dry} P \hat{\mathbf{Z}} + \mathbf{A}_{wet} \mathbf{M} \mathbf{W}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

ここで、Pは、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ に適応される混合行列であり、

Mは、前記1つ以上の非相関化オーディオ信号Wに適用される混合行列であり、

$\mathbf{A}_{dry}$ は、第1の相関マトリックスまたは第1の調整マトリックスであり、

$\mathbf{A}_{wet}$ は、第2の相関マトリックスまたは第2の調整マトリックスである、請求項1ないし請求項9に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項16】

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\tilde{\mathbf{Z}}$ または $\hat{\mathbf{Z}}$ とPおよびMを用いるWとのミキシングによって得られたオーディオ信号の相関特性または共分散特性( $E_{\tilde{\mathbf{Z}}}$ )が、前記所望の相関特性または所望の共分散特性(C)に近似するかまたは等しいように、前記混合行列Pおよび前記混合行列Mの少なくとも1つを調整するように構成される、請求項15に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項17】

前記マルチチャネル・オーディオ・デコーダは、前記混合行列Pおよび前記混合行列Mと一緒に算出するように構成される、請求項15または請求項16に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項18】

前記マルチチャネル・オーディオ・デコーダは、前記得られた出力オーディオ信号 $\hat{\mathbf{Z}}$ の共分散行列 $\mathbf{E}_z$ または $\hat{\mathbf{Z}}$ と $\mathbf{P}$ および $\mathbf{M}$ を用いる $\mathbf{W}$ とのミキシングによって得られたオーディオ信号の共分散行列が、所望の共分散行列 $\mathbf{C}$ に近似するかまたは等しいように、

$$\mathbf{F} = [\mathbf{P} \quad \mathbf{M}]$$

で示される、結合混合行列 $\mathbf{F}$ を決定するように構成される、請求項1\_5ないし請求項1\_7に記載のマルチチャンネル・オーディオ・デコーダ。

【請求項19】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{E}_{\hat{\mathbf{Z}}} = \mathbf{F} \mathbf{E}_s \mathbf{F}^H$$

で示される前記共分散行列が、

$$\mathbf{C} = \mathbf{R} \mathbf{E}_x \mathbf{R}^H$$

で示される前記所望の共分散行列と等しくなるように、前記結合混合行列 $\mathbf{F}$ を決定するように構成され、

ここで、 $\mathbf{E}_s$ は、次式で定義される、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合した信号 $\mathbf{S}$ の共分散行列であり、

$$\mathbf{S} = \begin{bmatrix} \hat{\mathbf{Z}} \\ \mathbf{W} \end{bmatrix}$$

ここで、 $\mathbf{E}_x$ は、オブジェクト共分散行列である、請求項1\_8に記載のマルチチャンネル・オーディオ・デコーダ。

【請求項20】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記レンダリングされたオーディオ信号の貢献が制限されるように、前記第1の相関マトリックスを決定するように構成され、および/または

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記非相関化オーディオ信号の貢献が制限されるように、前記第2の相関マトリックスを決定するように構成される、請求項1\_5ないし請求項1\_9のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

【請求項21】

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記レンダリングされたオーディオ信号の貢献が制限されるように、前記レンダリングされたオーディオ信号の特性に基づいて、および/または、前記非相関化オーディオ信号の特性に基づいて、および/または、所望の出力オーディオ信号の特性に基づいて、および/または、混合されレンダリングされたオーディオ信号の推定された特性に基づいて、および/または、混合された非相関化オーディオ信号の推定された特性に基づいて、前記第1の相関マトリックスを決定するように構成され、および/または

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号に前記非相関化オーディオ信号の貢献が制限されるように、前記レンダリングされたオーディオ信号の特性に基づいて、および／または、前記非相関化オーディオ信号の特性に基づいて、および／または、所望の出力オーディオ信号の特性に基づいて、および／または、混合されレンダリングされたオーディオ信号の推定された特性に基づいて、および／または、混合された非相関化オーディオ信号の推定された特性に基づいて、前記第2の相関マトリックスを決定するように構成される、請求項1\_5ないし請求項2\_0のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項22】

前記レンダリングされたオーディオ信号、および／または、前記非相関化オーディオ信号、および／または、前記所望の出力オーディオ信号、および／または、前記混合されレンダリングされたオーディオ信号、および／または、前記混合された非相関化オーディオ信号の前記特性は、エネルギー特性、または、相関特性、または、共分散特性である、請求項2\_1に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項23】

前記マルチチャネル・オーディオ・デコーダは、

$$\widehat{\mathbf{Z}} = \mathbf{P}\widehat{\mathbf{Z}} + \mathbf{A}_{\text{wet}}\mathbf{M}\mathbf{W}$$

で示される前記出力オーディオ信号 $\widehat{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\widehat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

ここで、 $\mathbf{A}_{\text{wet}}$ は対角行列であり、 $i$ 番目の出力オーディオ信号において、混合行列 $\mathbf{M}$ を伴って、レンダリングされたオーディオ信号の強度 ( $E_Y^{\text{dry}}(i,i)$ ) と混合された非相関化オーディオ信号の強度 ( $\widehat{E}_Y^{\text{wet}}(i,i)$ ) との比が、閾値よりも小さい場合、標準、すなわち、非減少の対角入力と比較したとき、相関マトリックス $\mathbf{A}_{\text{wet}}$ の入力 $\mathbf{A}_{\text{wet}}(i,i)$ が減少するように、前記相関マトリックス $\mathbf{A}_{\text{wet}}$ を供給するように構成される、請求項1ないし請求項2\_2のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項24】

前記閾値は、予め定められた固定閾値であるか、または、前記閾値は、たとえば、エネルギー特性、相関特性、および／または共分散特性のような信号特性に基づく時間変数、および／または周波数変数である、請求項2\_3に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項25】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \mathbf{P}\widehat{\mathbf{Z}} + \mathbf{A}_{wet}\mathbf{M}\mathbf{W}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\widehat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号 $\mathbf{W}$ とを結合するように構成され、

ここで、 $\mathbf{P} = \mathbf{P}_{dry}$ であり、

$\mathbf{M} = \mathbf{P}_{wet}$ であり、

$$\mathbf{A}_{wet} = \text{matdiag} \left( \min \left( 1, \sqrt{\max \left( 0, \lambda_{Dec} \frac{\mathbf{E}_Y^{dry}(i,i)}{\widehat{\mathbf{E}}_Y^{wet}(i,i)} \right)} \right) \right) \text{であり、}$$

$\mathbf{E}_Y^{dry}$ は、前記レンダリングされたオーディオ信号 $\widehat{\mathbf{Z}}$ の共分散行列であり、

$\widehat{\mathbf{E}}_Y^{wet}$ は、前記マトリックス $\mathbf{P}_{wet}$ が適用された後、前記非相関化オーディオ信号の推定された共分散行列である、請求項1ないし請求項2 4のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

【請求項2 6】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{F} = (\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)\mathbf{H}(\mathbf{V}\sqrt{\mathbf{Q}^{-1}}\mathbf{V}^H)$$

で示される前記結合混合行列  $\mathbf{F}$  を決定するように構成され、

ここで、マトリックス  $\mathbf{U}$ 、 $\mathbf{T}$ 、 $\mathbf{V}$  および  $\mathbf{Q}$  は、前記共分散行列の特異値分解を使用して決定され、 $\mathbf{E}_s$  および  $\mathbf{C}$  は、

$$\mathbf{C} = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

および

$$\mathbf{E}_s = \mathbf{V}\mathbf{Q}\mathbf{V}^H$$

で得られ、

マトリックス  $\mathbf{H}$  は、

$$\mathbf{H} = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 & b_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 & 0 & b_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{N,N} & 0 & 0 & \dots & b_{N,N} \end{bmatrix}$$

として定義され、

$a_{i,i}$  および  $b_{i,i}$  は、

$$a_{i,i}^2 + b_{i,i}^2 = 1$$

で示すように選択される、請求項 1\_5 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 2 7】

前記マルチチャネル・オーディオ・デコーダは、単位行列であるか、またはその倍数である混合行列  $\mathbf{P}$  を設定し、そして、前記混合行列  $\mathbf{M}$  を算出するように構成される、請求項 1\_0 または請求項 1\_1 に記載のマルチチャネル・オーディオ・デコーダ。

【請求項 2 8】

前記マルチチャネル・オーディオ・デコーダは、

$$\Delta_E = \mathbf{C} - \mathbf{E}_{\hat{Z}}$$

として定義される、前記所望の共分散行列  $\mathbf{C}$  と共分散行列  $\mathbf{E}_{\hat{Z}}$  との差  $\Delta_E$  が、共分散である

$$\mathbf{M}\mathbf{E}_w\mathbf{M}^H$$

に等しいか、あるいは近似であるように、前記混合行列  $\mathbf{M}$  を決定するように構成され、  
ここで、前記所望の共分散行列  $\mathbf{C}$  は、

$$\mathbf{C} = \mathbf{R}\mathbf{E}_x\mathbf{R}^H$$

として定義され、

$\mathbf{R}$  は、レンダリング行列であり、

$\mathbf{E}_x$  は、オブジェクト共分散行列であり、

$\mathbf{E}_w$  は、前記 1 つ以上の非相関化信号の共分散行列であり、

$\mathbf{E}_{\hat{Z}}$  は、前記レンダリングされたオーディオ信号の共分散行列である、請求項 2\_7 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 2\_9】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{M} = (\mathbf{U}\sqrt{\mathbf{T}}\mathbf{U}^H)(\mathbf{V}\sqrt{\mathbf{Q}^{-1}}\mathbf{V}^H)$$

で示される前記混合行列  $\mathbf{M}$  を決定するように構成され、

ここで、マトリックス  $\mathbf{U}$ 、 $\mathbf{T}$ 、 $\mathbf{V}$  および  $\mathbf{Q}$  は、前記共分散行列の特異値分解を使用して

決定され、 $\Delta_E$  および  $\mathbf{E}_w$  は、

$$\Delta_E = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

および

$$\mathbf{E}_w = \mathbf{V}\mathbf{Q}\mathbf{V}^H$$

で得られる、請求項 2\_8 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3\_0】

前記マルチチャネル・オーディオ・デコーダは、所与のレンダリングされたオーディオ信号が、前記所与のレンダリングされたオーディオ信号自身の非相関化バージョンとともに混合されるのみという制約のもとに、前記混合行列  $\mathbf{P}$ 、 $\mathbf{M}$  を決定するように構成される、請求項 1\_0 または請求項 1\_1 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3\_1】

相互相関値または相互共分散値が不变である間、レンダリングされたオーディオ信号の自己相関値のみまたは自己共分散値のみが修正されるように、前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号と前記1つ以上の非相関化オーディオ信号とを結合するように構成される、請求項1\_0または請求項1\_1または請求項3\_0に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項3\_2】

前記マルチチャネル・オーディオ・デコーダは、単位行列であるか、またはその倍数である前記混合行列Pを設定し、そして、Mが対角行列であるという制限のもと、前記混合行列Mを算出するように構成される、請求項1\_0または請求項1\_1または請求項3\_0または請求項3\_1に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項3\_3】

前記マルチチャネル・オーディオ・デコーダは、

$$\tilde{\mathbf{Z}} = \hat{\mathbf{Z}} + \mathbf{MW}$$

で示される前記出力オーディオ信号 $\tilde{\mathbf{Z}}$ を得るために、前記レンダリングされたオーディオ信号 $\hat{\mathbf{Z}}$ と前記1つ以上の非相関化オーディオ信号Wとを結合するように構成され、

ここで、Mは、前記1つ以上の非相関化オーディオ信号Wに適用される対角混合行列であり、そして、

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号の共分散行列の対角線要素が所望のエネルギーと等しくなるように、前記混合行列Mの対角線要素を算出するように構成される、請求項3\_0または請求項3\_1または請求項3\_2に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項3\_4】

前記マルチチャネル・オーディオ・デコーダは、

$$\mathbf{M}(i,j) = \begin{cases} \sqrt{\min\left(\lambda_{Dec}, \max\left(0, \frac{\mathbf{C}(i,i) - \mathbf{E}_{\hat{\mathbf{Z}}}(i,i)}{\max(\mathbf{E}_w(i,i), \varepsilon)}\right)\right)} & i = j, \\ 0 & i \neq j. \end{cases}$$

で示される前記混合行列Mの前記要素を算出するように構成され、

ここで、前記所望の共分散行列Cは、

$$\mathbf{C} = \mathbf{RE}_x \mathbf{R}^H$$

として定義され、

Rは、レンダリング行列であり、

$E_x$ は、オブジェクト共分散行列であり、

$E_w$ は、前記1つ以上の非相関化信号の共分散行列であり、そして、

$\lambda_{Dec}$ は、前記信号に加えられる非相関の量を制限している閾値である、請求項3\_3に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項3\_5】

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合するための方法を決定する場合、前記マルチチャネル・オーディオ・デコーダは、前記非相関化オーディオ信号の相関特性または共分散特性

を考慮するように構成される、請求項 1 ないし請求項 3\_4 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 6】

所与の出力オーディオ信号が 2 つ以上のレンダリングされたオーディオ信号および少なくとも 1 つの非相関化オーディオ信号に基づいて供給されるように、前記マルチチャネル・オーディオ・デコーダは、レンダリングされたオーディオ信号と非相関化オーディオ信号とを混合するように構成される、請求項 1 ないし請求項 2\_6、または請求項 3\_5 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 7】

前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合するための方法を決定するために、異なる制約が適用される場合において、前記マルチチャネル・オーディオ・デコーダは、異なるモードの間で切り替えるように構成される、請求項 1 ないし請求項 3\_6 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 8】

前記マルチチャネル・オーディオ・デコーダは、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第 1 のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第 2 のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第 3 のモード、

の間で切り替えるように構成される、請求項 1 ないし請求項 3\_7 のいずれかに記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 3 9】

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記 1 つ以上の非相関化オーディオ信号とを結合するための前記 3 つのモードが使用されるかについて指示する前記符号化表現のビットストリーム要素を評価し、そして、前記ビットストリーム要素に基づいて前記モードを選択するように構成される、請求項 3\_7 または請求項 3\_8 に記載のマルチチャネル・オーディオ・デコーダ。

#### 【請求項 4 0】

少なくとも 2 つの入力オーディオ信号 ( 2 1 0 , 2 1 2 ; 1 5 1 2 a - 1 5 1 2 n ; 2 9 1 2 , 2 9 1 4 ) に基づいて符号化表現 ( 2 1 4 ; 1 5 1 6 a , 1 5 1 6 b , 1 5 1 8 ; 2 9 3 2 ) を供給するためのマルチチャネル・オーディオ・エンコーダ ( 2 0 0 ; 1 5 1 0 ; 2 9 0 0 ) であって、

ここで、前記マルチチャネル・オーディオ・エンコーダは、前記少なくとも 2 つの入力オーディオ信号に基づき、1 つ以上のダウンミックス信号 ( 2 2 2 ; 1 5 1 6 a , 1 5 1 6 b ) を供給する ( 2 2 0 ) ように構成され、

前記マルチチャネル・オーディオ・エンコーダは、前記少なくとも 2 つの入力オーディオ信号の間の関係を記述している 1 つ以上のパラメータ ( 2 3 2 ; 1 5 1 8 ) を供給する ( 2 3 0 ) ように構成され、そして、

前記マルチチャネル・オーディオ・エンコーダは、オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータ(242；1518)を供給する(240)ように構成され、

前記マルチチャネル・オーディオ・エンコーダは、前記非相関化方法パラメータを選択的に供給し、オーディオ・デコーダの処理のための、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、

である以上に示す3つのモードの1つの信号を送信するように構成される、マルチチャネル・オーディオ・エンコーダ。

#### 【請求項41】

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号が比較的高い相関または比較的低い相関を含むかどうかに基づいて、前記非相関化方法パラメータを選択するように構成される、請求項40に記載のマルチチャネル・オーディオ・エンコーダ。

#### 【請求項42】

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号の間の相関が比較的高い場合、前記第1のモードまたは前記第2のモードを指定するための前記非相関化方法パラメータを選択するように構成され、

前記マルチチャネル・オーディオ・エンコーダは、前記入力オーディオ信号の間の相関が比較的低い場合、前記第3のモードを指定するための前記非相関化方法パラメータを選択するように構成される、請求項40または請求項41に記載のマルチチャネル・オーディオ・エンコーダ。

#### 【請求項43】

符号化表現に基づく少なくとも2つの出力オーディオ信号を供給するための方法(300)であって、前記方法は、

複数のレンダリングされたオーディオ信号を得るために、レンダリング行列を定義する1つ以上のレンダリング・パラメータに基づいて、マルチチャネル・ターゲット・シーンに対して、前記符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするステップ(310)と、

前記レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するステップ(320)と、

前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合するステップ(330)と、を含み、

ここで、前記複数のレンダリングされたオーディオ信号を得るためにレンダリングされた前記復号化オーディオ信号は、パラメトリック再構成(120；1560)を使用して得られ、

前記復号化オーディオ信号は、再構成オブジェクト信号であり、

前記再構成オブジェクト信号は、サイド情報（1518）を使用して、1つ以上のダウンミックス信号（1516a, 1516b）から導出される、方法。

【請求項44】

少なくとも2つの入力オーディオ信号に基づく符号化表現を供給するための方法（400）であって、前記方法は、

前記少なくとも2つの入力オーディオ信号に基づく1つ以上のダウンミックス信号を供給するステップ（410）と、

前記少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータを供給するステップ（420）と、

オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している非相関化方法パラメータを供給するステップ（430）と、を含み、ここで、前記方法は、前記非相関化方法パラメータを選択的に供給するステップを含み、オーディオ・デコーダの処理のための

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間における混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間における混合が許容されず、そして、前記出力オーディオ信号の相互相関特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間における混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、

である以上に示す3つのモードの1つの信号を送信することを含む、方法。

【請求項45】

コンピュータ・プログラムがコンピュータ上で実行されると、前記コンピュータが請求項43、または請求項44に記載の方法を実行する、コンピュータ・プログラム。

【請求項46】

符号化オーディオ表現（500）は、

ダウンミックス信号の符号化表現（510）、

前記少なくとも2つの入力オーディオ信号の間の関係を記述している1つ以上のパラメータの符号化表現（520）、および

オーディオ・デコーダのサイドで使用される複数の非相関化モードのうちの非相関化モードを記述している符号化非相関化方法パラメータ（530）を含み、

ここで、前記非相関化方法パラメータは、オーディオ・デコーダの処理のための、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間における混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間における混合が許容されず、そして、前記出力オーディオ信号の相互相関特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ

信号の間における混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号を結合されることが許容されない、第3のモード、  
である以上に示す3つのモードの1つの信号を送信する、符号化オーディオ表現。

**【請求項47】**

符号化表現(110; 710; 1516a; 1516b; 1518)に基づいて、少なくとも2つの出力オーディオ信号(112, 114; 712, 714; 1552a-1552n; 3012)を供給するためのマルチチャネル・オーディオ・デコーダ(100; 700; 1550; 3000)であって、

ここで、前記マルチチャネル・オーディオ・デコーダは、複数のレンダリングされたオーディオ信号(134, 136; 1582a-1582n,  $\hat{Z}$ )を得るために、1つ以上のレンダリング・パラメータ(132)に基づいて、前記符号化表現に基づいて得られた複数の復号化オーディオ信号(122; 1562a-1562n,  $\hat{X}$ )をレンダリングする(130; 1580)ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号(142, 144; 1592a-1592n)を導出する(140; 1590)ように構成され、

前記マルチチャネル・オーディオ・デコーダは、前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する(150; 1598)ように構成され、

前記マルチチャネル・オーディオ・デコーダは、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相關特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、

の間で切り替えられるように構成される、マルチチャネル・オーディオ・デコーダ。

**【請求項48】**

符号化表現に基づく少なくとも2つの出力オーディオ信号を供給するための方法(300)であって、前記方法は、

複数のレンダリングされたオーディオ信号を得るために、1つ以上のレンダリング・パラメータに基づいて、前記符号化表現に基づいて得られた複数の復号化オーディオ信号をレンダリングするステップ(310)と、

前記レンダリングされたオーディオ信号から1つ以上の非相関化オーディオ信号を導出するステップ(320)と、

前記出力オーディオ信号を得るために、前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合するステップ(330)と、を含み、

前記方法は、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容される、第1のモード、

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、前記出力オーディオ信号の相互相関特性または相互共分散特性を調整するために、所与の非相関化信号が、同じもしくは異なるスケール化、複数のレンダリングされたオーディオ信号、またはそのスケール化バージョンと結合されることが許容される、第2のモード、および

前記レンダリングされたオーディオ信号またはそのスケール化バージョンと前記1つ以上の非相関化オーディオ信号とを結合する場合に、異なるレンダリングされたオーディオ信号の間ににおける混合が許容されず、そして、所与の非相関化信号が、前記所与の非相関化信号から導出されたレンダリングされたオーディオ信号以外のレンダリングされたオーディオ信号と結合されることが許容されない、第3のモード、  
の間で切り替えられるステップを含む、方法。

【請求項49】

コンピュータ・プログラムがコンピュータ上で実行されると、前記コンピュータが請求項48に記載の方法を実行する、コンピュータ・プログラム。

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2014/065397

A. CLASSIFICATION OF SUBJECT MATTER  
INV. H04S3/00 G10L19/008  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
H04S G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, INSPEC, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	"ISO/IEC 23003-1:2006/FCD, MPEG Surround", 75. MPEG MEETING;16-01-2006 - 20-01-2006; BANGKOK; (MOTION PICTUREEXPERT GROUP OR ISO/IEC JTC1/SC29/WG11),, no. N7947, 3 March 2006 (2006-03-03), XP030014439, ISSN: 0000-0341 pages 3,30,58 figures 2,5,10 *Sections 5.2.4, 5.4.1, 5.4.2, 5.5.2, 5.5.3, 5.6, B.1* ----- X WO 2006/026452 A1 (DOLBY LAB LICENSING CORP [US]; SEEFEELDT ALAN JEFFREY [US]; VINTON MARK) 9 March 2006 (2006-03-09) cited in the application page 3, line 24 - page 6, line 16 page 8, line 11 - page 9, line 23 ----- -/-	1-39,41, 42,44, 46-49  1,46,48

Further documents are listed in the continuation of Box C.

See patent family annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier application or patent but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*&\* document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

10 October 2014

20/10/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Bensa, Julien

## INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/065397

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2012/009851 A1 (HUAWEI TECH CO LTD [CN]; CHRISTOF FALLER [CH]; DAVID VIRETTE [DE]; LAN) 26 January 2012 (2012-01-26) page 9, line 16 - page 12, line 8 figure 2 ----- YUE LANG ET AL: "NOVEL LOW COMPLEXITY COHERENCE ESTIMATION AND SYNTHESIS ALGORITHMS FOR PARAMETRIC STEREO CODING", EUSIPCO, 27 August 2012 (2012-08-27), pages 2427-2431, XP055042916, *Sections 1, 2, 4* figure 3 ----- ANONYMOUS: "ISO/IEC FDIS 23003-2: 2010, Spatial Audio Object Coding", 91. MPEG MEETING;18-1-2010 - 22-1-2010; KYOTO; (MOTION PICTURE EXPERTGROUP OR ISO/IEC JTC1/SC29/WG11),, no. N11207, 10 May 2010 (2010-05-10), XP030017704, ISSN: 0000-0030 cited in the application *Section 3.1.1* ----- JURGEN HERRE ET AL: "New Concepts in Parametric Coding of Spatial Audio: From SAC to SAOC", MULTIMEDIA AND EXPO, 2007 IEEE INTERNATIONAL CONFERENCE ON, IEEE, PI, 1 July 2007 (2007-07-01), pages 1894-1897, XP031124020, ISBN: 978-1-4244-1016-3 *Sections 2.1, 2.4, 3.2* figure 6 -----	1-4,46, 48 1-4,46, 48 3 1,46,48

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No
PCT/EP2014/065397

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2006026452 A1 09-03-2006		AT 447756 T AU 2005280041 A1 BR P10514620 A CA 2576739 A1 CN 101010723 A EP 1782417 A1 HK 1099839 A1 IL 181406 A JP 4909272 B2 JP 2008511044 A KR 20070051856 A MY 143850 A TW I393121 B US 2008126104 A1 WO 2006026452 A1	15-11-2009 09-03-2006 17-06-2008 09-03-2006 01-08-2007 09-05-2007 31-12-2009 28-04-2011 04-04-2012 10-04-2008 18-05-2007 15-07-2011 11-04-2013 29-05-2008 09-03-2006
WO 2012009851 A1 26-01-2012		CN 103069481 A EP 2586025 A1 JP 2013536461 A US 2013129096 A1 WO 2012009851 A1	24-04-2013 01-05-2013 19-09-2013 23-05-2013 26-01-2012

---

フロントページの続き

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

(72)発明者 フックス ハラルド

ドイツ連邦共和国 91341 レッテンバッハ アムゼルシュトラーセ 5

(72)発明者 ヘルムート オリヴァー

ドイツ連邦共和国 91054 ブッケンホフ アム ルーシュタイン 29

(72)発明者 ヘレ ユールゲン

ドイツ連邦共和国 91054 エアランゲン ラートシュベルガー シュトラーセ 10 アー

(72)発明者 モルタザ アドリアン

ルーマニア 200082 クラヨバ ストラダ コンスタンティン アルジュトアイア ヌマル  
ル 8 プロック6 スカラ1 アバルタメントウル2

(72)発明者 パウルス ヨウニ

ドイツ連邦共和国 90425 ニュルンベルグ ゼーバルドウスシュトラーセ 8

(72)発明者 リッダーブッシュ フアルコ

ドイツ連邦共和国 86163 アウグスブルク グリューンテンシュトラーセ 45

(72)発明者 テレンチエフ レオン

ドイツ連邦共和国 91056 エアランゲン アム オイローパカナル 36 アバルトマーン  
11

Fターム(参考) 5D162 AA01 AA09 AA11 BA01 BA13 CB12 CC33 CD01 EG02

(54)【発明の名称】マルチチャネル・オーディオ・デコーダ、マルチチャネル・オーディオ・エンコーダ、レンダリングされたオーディオ信号を使用する方法、コンピュータ・プログラムおよび符号化オーディオ表現