

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成29年1月12日(2017.1.12)

【公表番号】特表2015-537256(P2015-537256A)

【公表日】平成27年12月24日(2015.12.24)

【年通号数】公開・登録公報2015-081

【出願番号】特願2015-546945(P2015-546945)

【国際特許分類】

G 10 L 19/008 (2013.01)

【F I】

G 10 L 19/008 100

【手続補正書】

【提出日】平成28年11月22日(2016.11.22)

【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】

音場に対する高次アンビソニックス(HOA)表現を圧縮する方法であって、  
 - HOA係数の現在の時間フレームから支配的な音源方向を推定するステップと、  
 - 前記HOA表現を時間領域内の支配的な方向性信号と残差のHOA成分とに分解するステップであって、該残差のHOA成分を表現する均一なサンプリング方向で平面波関数を取得するために前記残差のHOA成分が離散空間領域に変換され、前記平面波関数が前記支配的な方向性信号から予測され、それにより、前記予測を記述するパラメータが与えられ、前記予測からの対応する予測誤りがHOA領域に再び変換される、ステップと、  
 - 前記残差のHOA成分の現在の次数をより低い次数に低減するステップであって、結果として、低次化された残差のHOA成分が得られる、ステップと、  
 - 前記低次化された残差のHOA成分を相関除去して対応する残差のHOA成分時間領域信号を取得するステップと、  
 - 圧縮された支配的な方向性信号および圧縮された残差の成分信号を供給するように、前記支配的な方向性信号および前記残差のHOA成分時間領域信号を知覚符号化するステップと、  
 を含む、方法。

【請求項2】

前記低次化された残差のHOA成分の前記相関除去は、球面調和関数変換を使用して、前記低次化された残差のHOA成分を空間領域内で対応する次数の等価信号に変換することによって行われる、請求項1に記載の方法。

【請求項3】

前記低次化された残差のHOA成分の前記相関除去は、均一なサンプリング方向のグリッドを回転させる球面調和関数変換を使用して、前記低次化された残差のHOA成分を空間領域内で対応する次数の等価信号に変換し、前記相関除去の反転を可能にする副情報を提供することによって行われる、請求項1に記載の方法。

【請求項4】

前記知覚符号化は、前記支配的な方向性信号および前記残差のHOA成分時間領域信号の合同の圧縮を含む、請求項1記載の方法。

【請求項5】

前記分解するステップは、

- H O A 係数の現在のフレームについての推定された支配的な音源方向から、支配的な方向性信号を計算するステップであって、その後の時間的平滑化によって平滑化された支配的な方向性信号が取得される、ステップと、
  - 前記推定された支配的な音源方向および前記平滑化された支配的な方向性信号から、平滑化された支配的な方向性信号の H O A 表現を計算するステップと、
  - 均一なグリッド上の方向性信号による対応する残差の H O A 表現を表現するステップと、
  - 前記平滑化された支配的な方向性信号および方向性信号による前記残差の H O A 表現から、均一なグリッド上の方向性信号を予測し、該予測から均一なグリッド上の予測された方向性信号の H O A 表現を計算し、その後、時間的平滑化を行う、ステップと、
  - 均一なグリッド上での前記平滑化された予測された方向性信号と、H O A 係数の前記現在のフレームの 2 フレーム遅延したバージョンと、前記平滑化された支配的な方向性信号の 1 フレーム遅延したバージョンとから、残差のアンビエント音場成分の H O A 表現を計算するステップと、
- を含む、請求項 1 に記載の方法。

#### 【請求項 6】

音場に対する高次アンビソニックス ( H O A ) 表現を圧縮する装置であって、

- H O A 係数の現在の時間フレームから支配的な音源方向を推定する推定器と、
  - 前記 H O A 表現を時間領域内の支配的な方向性信号と残差の H O A 成分とに分解する分解器であって、該残差の H O A 成分を表現する均一なサンプリング方向で平面波関数を取得するために前記残差の H O A 成分が離散空間領域に変換され、前記平面波関数が前記支配的な方向性信号から予測され、それにより前記予測を記述するパラメータが与えられ、前記予測からの対応する予測誤りが前記 H O A の領域に再び変換される、分解器と、
  - 前記残差の H O A 成分の現在の次数をより低い次数に低減する次数低減器であって、結果として、低次化された残差の H O A 成分を生成する、次数低減器と、
  - 前記低次化された残差の H O A 成分を相関除去して、対応する残差の H O A 成分時間領域信号を取得する相関除去器と、
  - 圧縮された支配的な方向性信号および圧縮された残差の成分信号を供給するように、前記支配的な方向性信号および前記残差の H O A 成分時間領域信号を知覚符号化する符号化器と、
- を備える、装置。

#### 【請求項 7】

前記低次化された残差の H O A 成分の前記相関除去は、球面調和関数変換を使用して、前記低次化された残差の H O A 成分を空間領域内で対応する次数の等価信号に変換することによって行われる、請求項 6 に記載の装置。

#### 【請求項 8】

前記低次化された残差の H O A 成分の前記相関除去は、均一なサンプリング方向のグリッドを回転させる球面調和関数変換を使用して、前記低次化された残差の H O A 成分を空間領域内で対応する次数の等価信号に変換し、前記相関除去の反転を可能にする副情報を提供することによって行われる、請求項 6 に記載の装置。

#### 【請求項 9】

前記支配的な方向性信号および前記残差の H O A 成分時間領域信号の前記知覚符号化は、合同して行なわれる、請求項 6 記載の装置。

#### 【請求項 10】

前記分解することとは、

- H O A 係数の現在のフレームについての推定された支配的な音源方向から、支配的な方向性信号を計算するステップであって、その後の時間的平滑化によって平滑化された支配的な方向性信号が取得される、ステップと、
- 前記推定された支配的な音源方向および前記平滑化された支配的な方向性信号から、平

滑化された支配的な方向性信号の H O A 表現を計算するステップと、

- 均一なグリッド上の方向性信号による対応する残差の H O A 表現を表現するステップと、

- 前記平滑化された支配的な方向性信号および方向性信号による前記残差の H O A 表現から、均一なグリッド上の方向性信号を予測し、該予測から均一なグリッド上の予測された方向性信号の H O A 表現を計算し、その後、時間的平滑化を行う、ステップと、

- 均一なグリッド上で前記平滑化された予測された方向性信号と、 H O A 係数の前記現在のフレームの 2 フレーム遅延したバージョンと、前記平滑化された支配的な方向性信号の 1 フレーム遅延したバージョンとから、残差のアンビエント音場成分の H O A 表現を計算するステップと、

を含む、請求項 6 に記載の装置。

【請求項 1 1】

均一なグリッド上の方向性信号の前記予測が、割り当てられた支配的な方向性信号からの遅延および全帯域スケーリングによって計算される、請求項 1 0 に記載の装置。

【請求項 1 2】

均一なグリッド上の方向性信号の前記予測において、知覚指向の周波数帯域に対するスケーリング係数が求められる、請求項 1 0 に記載の装置。

【請求項 1 3】

高次アンビソニックス ( H O A ) 表現を圧縮解除する方法であって、

- 圧縮解除された支配的な方向性信号および空間領域内の残差の H O A 成分を表現する圧縮解除された時間領域信号を供給するように、圧縮された支配的な方向性信号および圧縮された残差の成分信号を知覚復号するステップと、

- 前記圧縮解除された時間領域信号を再相関させて、対応する低次化された残差の H O A 成分を取得するステップと、

- 前記低次化された残差の H O A 成分の次数を当初の次数に拡張するステップであって、それによって当初の次数の圧縮解除された残差の H O A 成分を供給する、ステップと、

- 前記圧縮解除された支配的な方向性信号と、前記当初の次数の圧縮解除された残差の H O A 成分と、推定された支配的な音源方向とを使用して、 H O A 係数の対応する圧縮解除され、再合成されたフレームを生成するステップと、

を含む、方法。

【請求項 1 4】

圧縮された高次アンビソニックス ( H O A ) 表現を圧縮解除する装置であって、

- 圧縮解除された支配的な方向性信号および空間領域内の残差の H O A 成分を表現する圧縮解除された時間領域信号を供給するように、圧縮された支配的な方向性信号および圧縮された残差の成分信号を知覚復号する復号器と、

- 前記圧縮解除された時間領域信号を再相関させて、対応する低次化された残差の H O A 成分を取得する再相関器と、

- 前記低次化された残差の H O A 成分の次数を当初の次数に拡張する次数拡張器であって、それによって当初の次数の圧縮解除された H O A 成分を供給する、次数拡張器と、

- 前記圧縮解除された支配的な方向性信号と、前記当初の次数の圧縮解除された残差の H O A 成分と、推定された支配的な音源方向とを使用して、 H O A 係数の圧縮解除され、再合成されたフレームを生成する合成器と、

を備える、装置。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 6 9

【補正方法】変更

【補正の内容】

【0 0 6 9】

本発明は、家庭環境におけるラウドスピーカ構成上で、または、劇場におけるラウドスピーカ構成上でレンダリングおよび再生が可能な音声信号に対応する処理に適用することができる。

いくつかの態様を記載しておく。

[態様 1]

音場に対する H O A と称する高次アンビソニックス表現を圧縮する方法であって、

- H O A 係数 ( D ( k ) ) の現在の時間フレームから支配的な音源方向 (

$A_{\hat{a}}(k)$

) を推定するステップ ( 1 1 ) と、

- 前記 H O A 係数 ( D ( k ) ) および前記支配的な音源方向 (

$A_{\hat{a}}(k)$

) に依存して、前記 H O A 表現を時間領域内の支配的な方向性信号 ( X<sub>DTR</sub> ( k - 1 ) ) と残差の H O A 成分 ( D<sub>A</sub> ( k - 2 ) ) とに分解するステップ ( 1 2 ) であって、該残差の H O A 成分を表現する均一なサンプリング方向で平面波関数を取得するために前記残差の H O A 成分が離散空間領域に変換され ( 3 3 ) 、前記平面波関数が前記支配的な方向性信号 ( X<sub>DTR</sub> ( k - 1 ) ) から予測されること ( 3 4 ) によって、前記予測を記述するパラメータ ( ( k - 1 ) ) がもたらされ、対応する予測誤りが前記 H O A の領域に再び変換される ( 3 5 ) 、該ステップ ( 1 2 ) と、

- 前記残差の H O A 成分 ( D<sub>A</sub> ( k - 2 ) ) の現在の次数 ( N ) をより低い次数 ( N<sub>RED</sub> ) に低減するステップ ( 1 3 ) であって、結果として、低次元化された残差の H O A 成分 ( D<sub>A,RED</sub> ( k - 2 ) ) が得られる、該ステップ ( 1 3 ) と、

- 前記低次元化された残差の H O A 成分 ( D<sub>A,RED</sub> ( k - 2 ) ) を相関除去して対応する残差の H O A 成分時間領域信号 ( W<sub>A,RED</sub> ( k - 2 ) ) を取得するステップ ( 1 4 ) と、

- 圧縮された支配的な方向性信号 (

$\tilde{X}_{DIR}(k - 1)$

) および圧縮された残差の成分信号 (

$W_{A,RED}(k - 2)$

) を供給するように、前記支配的な方向性信号 ( X<sub>DTR</sub> ( k - 1 ) ) および前記残差の H O A 成分時間領域信号 ( W<sub>A,RED</sub> ( k - 2 ) ) を知覚符号化するステップ ( 1 5 ) と、

を含む、前記方法。

[態様 2]

音場に対する H O A と称する高次アンビソニックス表現を圧縮する装置であって、

- H O A 係数 ( D ( k ) ) の現在の時間フレームから支配的な音源方向 (

$A_{\hat{a}}(k)$

) を推定するように構成された手段 ( 1 1 ) と、

- 前記 H O A 係数 ( D ( k ) ) および前記支配的な音源方向 (

## $A_{\hat{a}}(k)$

)に依存して、前記HOA表現を時間領域内の支配的な方向性信号( $X_{DIR}(k-1)$ )と残差のHOA成分( $D_A(k-2)$ )とに分解するように構成された手段(12)であって、該残差のHOA成分を表現する均一なサンプリング方向で平面波関数を取得するために前記残差のHOA成分が離散空間領域に変換され(33)、前記平面波関数が前記支配的な方向性信号( $X_{DIR}(k-1)$ )から予測されること(34)によって前記予測を記述するパラメータ(( $k-1$ ))がもたらされ、対応する予測誤りが前記HOAの領域に再び変換される(35)、前記手段(12)と、

- 前記残差のHOA成分( $D_A(k-2)$ )の現在の次数( $N$ )をより低い次数( $N_{RED}$ )に低減するように構成された手段(13)であって、結果として、低次元化された残差のHOA成分( $D_{A,RED}(k-2)$ )を生成する、該手段(13)と、
- 前記低次元化された残差のHOA成分( $D_{A,RED}(k-2)$ )を相関除去して、対応する残差のHOA成分時間領域信号( $W_{A,RED}(k-2)$ )を取得するように構成された手段(14)と、
- 圧縮された支配的な方向性信号(

## $\tilde{X}_{DIR}(k-1)$

)および圧縮された残差の成分信号(

## $\tilde{W}_{A,RED}(k-2)$

)を供給するように、前記支配的な方向性信号( $X_{DIR}(k-1)$ および前記残差のHOA成分時間領域信号( $W_{A,RED}(k-2)$ )を知覚符号化するように構成された手段と、

を備える、前記装置。

[態様3]

態様1に記載の方法に従って圧縮された高次アンビソニックス表現を圧縮解除する方法であって、

- 圧縮解除された支配的な方向性信号(

## $\hat{X}_{DIR}(k-1)$

)および空間領域内の残差のHOA成分を表現する圧縮解除された時間領域信号(

## $\hat{W}_{A,RED}(k-2)$

)を供給するように、前記圧縮された支配的な方向性信号(

## $\tilde{X}_{DIR}(k-1)$

)および前記圧縮された残差の成分信号(

## $\tilde{W}_{A,RED}(k-2)$

)を知覚復号するステップ(21)と、

- 前記圧縮解除された時間領域信号(

$\widehat{W}_{A,RED}(k-2)$

) を再相関させて、対応する低次元化された残差の H O A 成分 (

$\widehat{D}_{A,RED}(k-2)$

) を取得するステップ ( 2 2 ) と、

- 前記低次元化された残差の H O A 成分 (

$\widehat{D}_{A,RED}(k-2)$

) の次数 (  $N_{RED}$  ) を当初の次数 (  $N$  ) に拡張するステップ ( 2 3 ) であって、それに  
よって対応する圧縮解除された残差の H O A 成分 (

$\widehat{D}_A(k-2)$

) を供給する、該ステップ ( 2 3 ) と、

- 前記圧縮解除された支配的な方向性信号 (

$\widehat{X}_{DIR}(k-1)$

) と、前記当初の次数の圧縮解除された残差の H O A 成分 (

$\widehat{D}_A(k-2)$

) と、前記推定された ( 1 1 ) 支配的な音源方向 (

$A_{\hat{a}}(k)$

) と、前記予測を記述する前記パラメータ ( (  $k-1$  ) ) とを使用して、 H O A 係数の  
対応する圧縮解除され、再合成されたフレーム

$\widehat{D}(k-2)$

を合成するステップ ( 2 4 ) と、

を含む、前記方法。

[ 態様 4 ]

態様 1 に記載の方法に従って圧縮された高次アンビソニックス表現を圧縮解除する装置  
であって、

- 圧縮解除された支配的な方向性信号 (

$\widehat{X}_{DIR}(k-1)$

) および空間領域内の残差の H O A 成分を表現する圧縮解除された時間領域信号 (

$\widehat{W}_{A,RED}(k-2)$

) を供給するように、前記圧縮された支配的な方向性信号 (

$\tilde{X}_{DIR}(k-1)$

) および前記圧縮された残差の成分信号 (

$\tilde{W}_{A,RED}(k-2)$

) を知覚復号するように構成された手段 ( 2 1 ) と、

- 前記圧縮解除された時間領域信号 (

$\tilde{W}_{A,RED}(k-2)$

) を再相関させて、対応する低次元化された残差の H O A 成分 (

$\hat{D}_{A,RED}(k-2)$

) を取得するように構成された手段 ( 2 2 ) と、

- 前記低次元化された残差の H O A 成分 (

$\hat{D}_{A,RED}(k-2)$

) の次数 (  $N_{RED}$  ) を当初の次数 (  $N$  ) に拡張するように構成された手段 ( 2 3 ) であ  
って、それによって対応する圧縮解除された H O A 成分 (

$\hat{D}_A(k-2)$

) を供給する、該手段 ( 2 3 ) と、

- 前記圧縮解除された支配的な方向性信号 (

$\hat{X}_{DIR}(k-1)$

) と、前記当初の次数の圧縮解除された残差の H O A 成分 (

$\hat{D}_A(k-2)$

) と、前記推定された ( 1 1 ) 支配的な音源方向 (

$A_{\hat{a}}(k)$

) と、前記予測を記述する前記パラメータ ( (  $k - 1$  ) ) とを使用して、 H O A 係数の  
対応する圧縮解除され、再合成されたフレーム (

$\hat{D}(k-2)$

) を合成するように構成された手段 ( 2 4 ) と、

を備える、前記装置。

[ 態様 5 ]

前記低次元化された残差の H O A 成分 (  $D_{A,RED}(k-2)$  ) の前記相関除去 ( 1  
4 ) は、球面調和関数変換を使用して、前記低次元化された残差の H O A 成分を空間領域  
内で対応する次数の等価信号に変換することによって行われる、態様 1 に記載の方法、ま

たは態様 2 に記載の装置。

[ 態様 6 ]

前記低次元化された残差の H O A 成分 ( $D_{A,RED}(k-2)$ ) の前記相関除去 (14) は、球面調和関数変換を使用して、前記低次元化された残差の H O A 成分を空間領域内で対応する次数の等価信号に変換することによって行われ、前記相関除去の反転を可能にする副情報 ( $(k-2)$ ) を提供することによって、サンプリング方向のグリッドが回転されて最大限の相関除去効果を得る、態様 1 に記載の方法、または態様 2 に記載の装置。

[ 態様 7 ]

前記支配的な方向性信号 ( $X_{DIR}(k-1)$ ) および前記残差の H O A 成分時間領域信号 ( $W_{A,RED}(k-2)$ ) の知覚圧縮 (15) が共に行われ、前記圧縮された方向性信号 (

$$\tilde{X}_{DIR}(k-1)$$

) および前記圧縮された時間領域信号 (

$$\tilde{W}_{A,RED}(k-2)$$

) の前記知覚圧縮 (21) が対応する方法で共に行われる、態様 1、3、5、および 6 のいずれか 1 項に記載の方法、または態様 2 および 4 ~ 6 のいずれか 1 項に記載の装置に従った方法。

[ 態様 8 ]

前記分解するステップ (12) は、  
- H O A 係数の現在のフレーム ( $D(k)$ ) に対して (

$$A_{\hat{\theta}}(k)$$

) における推定された音源方向から支配的な方向性信号 (

$$\tilde{X}_{DIR}(k)$$

) を計算するステップ (30) であって、その後の時間的平滑化 (31) によって平滑化された支配的な方向性信号 ( $X_{DIR}(k-1)$ ) が取得される、該ステップと、  
- (

$$A_{\hat{\theta}}(k)$$

) における前記推定された音源方向および前記平滑化された支配的な方向性信号 ( $X_{DIR}(k-1)$ ) から平滑化された支配的な方向性信号 ( $D_{DIR}(k-1)$ ) の H O A 表現を計算するステップ (32) と、  
- 均一なグリッド上の方向性信号 (

$$\tilde{X}_{GRID,DIR}(k-1)$$

) による対応する残差の H O A 表現を表現するステップ (33) と、

- 前記平滑化された支配的な方向性信号 ( $X_{DIR}(k-1)$ ) および方向性信号 (

$$\tilde{X}_{GRID,DIR}(k-1)$$

)による前記残差の H O A 表現から、均一なグリッド上の方向性信号(

$$\widehat{X}_{\text{GRID,DIR}}(k-1)$$

)を予測し(34)、該予測から均一なグリッド上の予測された方向性信号の H O A 表現を計算し(35)、その後、時間的平滑化を行う(36)、ステップと、  
- 均一なグリッド上で前記平滑化された予測された方向性信号(

$$\widehat{D}_{\text{GRID,DIR}}(k-2)$$

)と、H O A 係数の前記現在のフレーム(  $D(k)$  )の2フレーム遅延したバージョンと、前記平滑化された支配的な方向性信号(  $X_{\text{DIR}}(k-1)$  )の1フレーム遅延したバージョンとから、残差のアンビエント音場成分の H O A 表現(  $D_A(k-2)$  )を計算するステップと、

を含む、態様1および5~7のいずれか1項に記載の方法に従った方法、または態様2および5~7のいずれか1項に記載の装置に従った装置。

[態様9]

前記合成するステップ(24)は、

- H O A 係数の現在のフレーム(  $D(k)$  )に対して前記推定された音源方向(

$$A_{\hat{n}}(k)$$

)と、前記圧縮解除された支配的な方向性信号(

$$\widehat{X}_{\text{DIR}}(k-1)$$

)とから、支配的な方向性信号(

$$\widehat{D}_{\text{DIR}}(k-1)$$

)の H O A 表現を計算するステップ(41)と、

前記圧縮解除された支配的な方向性信号(

$$\widehat{X}_{\text{DIR}}(k-1)$$

)と、前記予測を記述した前記パラメータ( (  $k-1$  ) )とから、均一なグリッド上の方向性信号

$$(\widehat{\widehat{X}}_{\text{GRID,DIR}}(k))$$

を予測するステップ(43)と、当該予測から、均一なグリッド上の予測された方向性信号の H O A 表現

$$(\widehat{\widehat{D}}_{\text{GRID,DIR}}(k))$$

を計算するステップ(44)であって、その後に、時間的平滑化を行う

$$(45, \widehat{D}_{\text{GRID,DIR}}(k-1))$$

、該ステップと、

- 均一なグリッド上の予測された方向性信号

$(\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1))$

の前記平滑化された H O A 表現と、支配的な方向性信号 (

$\hat{\mathbf{D}}_{\text{DIR}}(k-1)$

) の前記 H O A 表現の 1 フレーム遅延された ( 4 2 ) バージョンと、前記圧縮解除された残差の H O A 成分 (

$\hat{\mathbf{D}}_{\text{A}}(k-2)$

) とから、H O A 音場表現 (

$\hat{\mathbf{D}}(k-2)$

) を合成するステップ ( 4 6 ) と、

を含む、態様 3 または 7 に記載の方法に従った方法、または態様 4 または 7 に記載の装置に従った装置。

[ 態様 1 0 ]

均一なグリッド上の方向性信号 (

$\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$

) の前記予測 ( 3 4 ) において、予測されたグリッド信号 (

$\hat{\mathbf{x}}_{\text{GRID,DIR},o}(k-1, l)$

) が、割り当てられた支配的な方向性信号 (

$\tilde{x}_{\text{DIR,EXT,f}_{\mathcal{A},k-1}(o)}(k-1, l)$

) からの遅延および全帯域スケーリングによって計算される、態様 8 に記載の方法に従った方法、または態様 8 に記載の装置に従った装置。

[ 態様 1 1 ]

均一なグリッド上の方向性信号 (

$\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$

) の前記予測 ( 3 4 ) において、知覚指向の周波数帯域に対するスケーリング係数が求められる、態様 8 に記載の方法に従った方法、または態様 8 に記載の装置に従った装置。

[ 態様 1 2 ]

態様 1 、 5 ~ 8 、 1 0 、および 1 1 のいずれか 1 項に記載の方法に従って符号化されるデジタル・オーディオ信号。