

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6285458号  
(P6285458)

(45) 発行日 平成30年2月28日 (2018. 2. 28)

(24) 登録日 平成30年2月9日 (2018. 2. 9)

(51) Int. Cl.

F I

G 1 0 L 19/008 (2013.01)

G 1 0 L 19/008 1 0 0

請求項の数 14 (全 32 頁)

(21) 出願番号	特願2015-546945 (P2015-546945)	(73) 特許権者	510185767
(86) (22) 出願日	平成25年12月4日 (2013. 12. 4)		ドルビー・インターナショナル・アーベー
(65) 公表番号	特表2015-537256 (P2015-537256A)		オランダ王国, セーエン アムステルダム
(43) 公表日	平成27年12月24日 (2015. 12. 24)		ズイドースト 1101, ヘリケルベル
(86) 国際出願番号	PCT/EP2013/075559		グウェグ 1-35, アポロ ビルディン
(87) 国際公開番号	W02014/090660		グ 3エー
(87) 国際公開日	平成26年6月19日 (2014. 6. 19)	(74) 代理人	100107766
審査請求日	平成28年11月22日 (2016. 11. 22)		弁理士 伊東 忠重
(31) 優先権主張番号	12306569.0	(74) 代理人	100070150
(32) 優先日	平成24年12月12日 (2012. 12. 12)		弁理士 伊東 忠彦
(33) 優先権主張国	欧州特許庁 (EP)	(74) 代理人	100091214
			弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 音場のための高次アンビソニックス表現を圧縮および圧縮解除する方法および装置

(57) 【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 (H O A) 表現を圧縮する方法であって、

- 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ステップと、
- 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 成分とに分解するステップであって、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 H O A 領域に再び変換される、ステップと、
- 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ステップであって、結果として、低次化された残差の H O A 成分が得られる、ステップと、
- 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 H O A 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 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

10

20

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 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 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 H O A 成分時間領域信号の合同の圧縮を含む、請求項 1 に記載の方法。

【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

10

20

30

40

50

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 に記載の装置。

【請求項 10】

前記分解することは、

- 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 6 に記載の装置。

【請求項 11】

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 10 に記載の装置。

【請求項 12】

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 10 に記載の装置。

【請求項 13】

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 14】

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

10

20

30

40

50

ＯＡ成分と、推定された支配的な音源方向とを使用して、ＨＯＡ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。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、音場のための高次アンビソニックス表現を圧縮および圧縮解除する方法および装置に関する。

【背景技術】

【０００２】

ＨＯＡと称する高次アンビソニックス表現は、三次元音声表現する１つの方法である。他の技術は波面合成法（ＷＦＳ）や２．２のようなチャンネルに基づく方法である。チャンネルに基づく方法と比較して、ＨＯＡ表現には、特定のラウドスピーカの設定とは独立しているという利点がある。しかしながら、この柔軟性を得るためには特定のラウドスピーカの設定でＨＯＡ表現を再生するための復号処理が必要となる。通常、必要なラウドスピーカの数が大変多くなるＷＦＳのアプローチと比較して、ＨＯＡは極めて少ない数のラウドスピーカのみで構成される設定にすることができる。ＨＯＡのさらなる利点は、ヘッドフォンへのバイノーラル・レンダリングにも変更を必要とすることなく同じ表現を利用することができる点にある。

【０００３】

ＨＯＡは、切断球面調和関数（ＳＨ）展開による複素調和平面波振幅の空間密度の表現に基づいている。各展開係数は角周波数の関数であり、これを時間領域関数によって同等に表現することができる。したがって、一般性を失うことなく、完全なＨＯＡ音場表現は、実際には、“ ” 時間領域関数から構成されるものと考えることができる。ここで、は、展開係数の数を表している。これらの時間領域関数と同等の意味を有するものとして、以下のＨＯＡ係数列を参照する。

【０００４】

ＨＯＡ表現の空間解像度は、展開の最大次数 $N$ の増加とともに向上する。残念ながら、展開係数の数“ ” は、次数 $N$ に対して二乗的に増加し、特に  $= (N + 1)^2$  となる。例えば、次数 $N = 4$ を使用した一般的なＨＯＡ表現には、 $= 25$ の個数のＨＯＡ（展開）係数が必要となる。上記の点を考慮して、ＨＯＡ表現の伝送のための合計ビットレートは、所望の単一チャンネルのサンプリング・レート $f_s$ およびサンプル毎のビットの数 $N_b$ が与えられると、 $\cdot f_s \cdot N_b$ によって求めることができる。サンプル毎に $N_b = 16$ の個数のビットを使用して $f_s = 48 \text{ kHz}$ のサンプリング・レートでの次数 $N = 4$ のＨＯＡ表現を伝送すると、結果として、ビットレートは、 $19.2 \text{ Mbps}$ となるが、これは、多くの実用的なアプリケーション、例えば、ストリーミングでは極めて高いビットレートである。したがって、ＨＯＡ表現を圧縮することが大いに望まれている。

【発明の概要】

【０００５】

１次よりも高いＨＯＡ表現の圧縮を取り扱う既存の方法は殆ど存在しない。E. Hellerud、I. Burnett、A. Solvang、およびU. P. Svenssonによって探究されている最も直接的なアプローチ「Encoding Higher Order Ambisonics with AAC（AACを用いた高次アンビソニックスの符号化）」第１２４回AESコンベンション、アムステルダム、２００８年は、知覚符号化アルゴリズムである、AAC（Advanced Audio Coding）を用いて個々のＨＯＡ係数列の直接的な符号化を行うものである。しかしながら、この手法に伴う固有の問題は、全く聴かれることのない信号の知覚符号化である。再構築された再生信号は、通常、ＨＯＡ係数列の加重和によって得られ、特定のラウドスピーカの設定で圧縮解除されたＨＯＡ表現がレンダリングされる場合には、知覚符号化ノイズをマス

10

20

30

40

50

ク除去する可能性が高い。知覚符号化ノイズのマスク除去の抱える主要な問題は、個々のH O A係数列間の高い相互相関である。個々のH O A係数列における符号化ノイズ信号は、互いに相関していないため、知覚符号化ノイズの構造的な重畳が発生することがあり、それと同時に、その重畳でノイズのないH O A係数列がキャンセルされてしまう。別の問題は、これらの相互相関が知覚符号化器の効率の低下につながる点である。

#### 【 0 0 0 6 】

双方の影響の程度を最小限にするために、欧州特許出願第2 4 6 9 7 4 2号 ( E P 2 4 6 9 7 4 2 A 2 ) では、H O A表現を知覚符号化の前に離散空間領域において、等価な表現に変換することが提案されている。形式的には、離散空間領域は、何らかの離散方向でサンプリングされる、複素調和平面波振幅の空間密度と等価な時間領域である。したがって、離散空間領域は、“ ” 個の従来の時間領域信号によって表現される。この信号は、サンプリング方向から到来する一般的な平面波として解釈することができ、空間領域変換に対して想定されるものと厳密に同じ方向にラウドスピーカが位置しているのであれば、ラウドスピーカ信号に対応するであろう。

10

#### 【 0 0 0 7 】

離散空間領域への変換により、個々の空間領域信号間の相互相関が低減するが、これらの相互相関は、完全には除去されない。比較的に高い相互相関の例は、空間領域信号によって包含される複数の隣接した方向の間を方向とする方向性信号である。

#### 【 0 0 0 8 】

双方のアプローチの主な欠点は、知覚符号化される信号の数が  $(N + 1)^2$  であり、圧縮されたH O A表現のデータ・レートがアンビソニックスの次数Nの二乗で増加することである。

20

#### 【 0 0 0 9 】

知覚符号化される信号の数を減少させるために、欧州特許出願公開第2 6 6 5 2 0 8号は、H O A表現を所与の最大数の支配的な方向性信号と残差のアンビエント成分とに分解することを提案している。知覚符号化されるべき信号の数の減少は、残差のアンビエント成分の次数を減少させることによって成し遂げることができる。この手法の背景にある理論的根拠は、支配的な方向性信号に関して高い空間解像度を維持する一方で、より低い次数のH O A表現によって十分な精度で残差を表現することにある。

#### 【 0 0 1 0 】

30

このアプローチは、音場に関する仮定が満たされる限り、すなわち、音場が少ない数の支配的な方向性信号 (これは、完全な次数Nで符号化された一般的な平面波関数を表現するものである。) と、方向性を有しない残差のアンビエント成分とからなるという仮定が満たされる限り、大変良好に機能する。しかしながら、分解の後、残差のアンビエント成分が依然として幾らかの支配的な方向性成分を含んでいる場合には、低次元化によって、分解の後のレンダリングの際に顕著に知覚される誤りが生じる。その仮定が満たされない場合のH O A表現の一般的な例は、Nよりも低い次数で符号化される一般的な平面波である。このようなNよりも低い次数の一般的な平面波は、音源の範囲が広がりをも有するよう感じられるようにする芸術的な創作の結果として生ずることがあり、球形マイクロフォンによるH O A音場表現の収録に伴って生ずることもある。双方の例において、音場は、多数の相関性の高い空間領域信号によって表現される (説明については、高次アンビソニックスの空間解像度の項目を参照されたい。 ) 。

40

#### 【 0 0 1 1 】

本発明によって解決される課題は、欧州特許出願公開第2 6 6 5 2 0 8号に記載された処理の結果として生ずる不都合を解消することによって、他の従来技術の上述した不都合を回避することにある。この課題は、請求項1および3に開示されている方法によって解決される。これらの方法を利用する対応する装置は、請求項2および4に開示されている。

#### 【 0 0 1 2 】

本発明は、欧州特許出願公開第2 6 6 5 2 0 8号に記載されたH O A音場表現圧縮処理

50

を改良する。まず、欧州特許出願公開第2665208号と同様に、H O A表現が支配的な音源の存在に対して分析され、その方向が推定される。支配的な音源の方向の情報を用いて、H O A表現は一般的な平面波を表現する複数の支配的な方向性信号と残差の成分とに分解される。しかしながら、この残差のH O A成分の次数を直ちに減少させる代わりに、残差のH O A成分を表現する均一なサンプリング方向における一般的な平面波関数を取得するために、この残差のH O A成分が離散空間領域へ変換される。この後、これらの平面波関数が支配的な方向性信号から予測される。この処理を行う理由は、残差のH O A成分の部分が支配的な方向性信号と高い相関性を有している場合があるからである。

【0013】

その予測は、少量の副情報のみを生み出すといった単純なものとすることができる。最も単純な場合では、予測は適切なスケールリングおよび遅延からなる。最終的に、予測誤りは再びH O A領域に変換され、低次元化が行われる残差のアンビエントH O A成分とされる。

【0014】

有利には、残差のH O A成分から予測可能な信号を差し引く効果は、その全体の次数および支配的な方向性信号の残量を減少させることであり、このようにして、低次元化の結果として生じる分解誤りを低減することにある。

【0015】

原理的には、本発明の圧縮方法は、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のに適している。この方法は、

- 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ステップと、
- 上記H O A係数および上記支配的な音源方向に依存して、上記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成分とに分解するステップであって、上記残差のH O A成分を表現する均一なサンプリング方向において平面波関数を取得するために、上記残差のH O A成分が離散空間領域に変換され、上記平面波関数が上記支配的な方向性信号から予測されることによって、上記予測を記述するパラメータがもたらされ、対応する予測誤りが上記H O Aの領域に再び変換される、上記分解するステップと、
- 上記残差のH O A成分の現在の次数をより低い次数に低減するステップであって、結果として、低次元化された残差のH O A成分が得られる、上記低減するステップと、
- 上記低次元化された残差の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、上記支配的な方向性信号および上記残差のH O A成分時間領域信号を知覚符号化するステップと、を含む。

【0016】

原理的には、本発明の圧縮装置は、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の圧縮に適している。この装置は、

- 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ように構成された手段と、
- 上記H O A係数および上記支配的な音源方向に依存して、上記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成分とに分解するように構成された手段であって、上記残差の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、上記残差のH O A成分が離散空間領域に変換され、上記平面波関数が上記支配的な方向性信号から予測されることによって、上記予測を記述するパラメータが供給され、対応する予測誤りが上記H O Aの領域に再び変換される、上記手段と、
- 上記残差のH O A成分の現在の次数をより低い次数に低減するように構成された手段であって、結果として、低次元化された残差のH O A成分が生成される、上記手段と、
- 上記低次元化された残差の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ように構成された手段と、
- 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、上

10

20

30

40

50

記支配的な方向性信号および上記残差のH O A成分時間領域信号を知覚符号化するように構成された手段と、を含む。

【0017】

原理的には、本発明の圧縮解除方法は、上述した圧縮方法に従って圧縮された高次アンビソニックス表現の圧縮解除に適している。この方法は、

- 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、上記圧縮された支配的な方向性信号および上記圧縮された残差の成分信号を知覚復号するステップと、
- 上記圧縮解除された時間領域信号を再相関させて、対応する低次元化された残差のH O A成分を取得するステップと、
- 上記低次元化された残差の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、対応する圧縮解除された残差のH O A成分を供給する、上記拡張するステップと、
- 上記圧縮解除された支配的な方向性信号と、上記当初の次数の圧縮解除された残差のH O A成分と、上記推定された支配的な音源方向と、上記予測を記述する上記パラメータとを使用して、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合成するステップと、を含む。

10

【0018】

原理的には、本発明の圧縮解除装置は、上述した圧縮方法に従って圧縮された高次アンビソニックス表現の圧縮解除に適している。この装置は、

- 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、上記圧縮された支配的な方向性信号および上記圧縮された残差の成分信号を知覚復号するように構成された手段と、
- 上記圧縮解除された時間領域信号を再相関させるように構成された手段であって、対応する低次元化された残差のH O A成分を取得する、上記手段と、
- 上記低次元化された残差のH O A成分の次数を当初の次数に拡張するように構成された手段であって、対応する圧縮解除された残差のH O A成分を供給する、上記手段と、
- 上記圧縮解除された支配的な方向性信号と、上記当初の次数の圧縮解除された残差のH O A成分と、上記推定された支配的な音源方向と、上記予測を記述する上記パラメータとを使用することによって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合成するように構成された手段と、を含む。

20

30

【0019】

本発明の有利な追加的な実施形態は、各々の従属請求項に開示されている。

【0020】

本発明の例示的な実施形態は、添付図面を参照して説明される。

【図面の簡単な説明】

【0021】

【図1a】圧縮ステップ1：H O A信号の複数の支配的な方向性信号、残差のアンビエントH O A成分、および副情報への分解を示す図である。

【図1b】圧縮ステップ2：アンビエントH O A成分の低次元化および相関除去および双方の成分の知覚符号化を示す図である。

40

【図2a】圧縮解除ステップ1：時間領域信号の知覚復号、残差のアンビエントH O A成分を表現する信号の再相関、および次数拡張を示す図である。

【図2b】圧縮解除ステップ2：全てのH O A表現の合成を示す図である。

【図3】H O A分解を示す図である。

【図4】H O A合成を示す図である。

【図5】球面座標系を示す図である。

【図6】Nの複数の異なる値に対する正規化された関数  $N(\quad)$  のプロットを示す図である。

【発明を実施するための形態】

【0022】

50

## 圧縮処理

本発明に係る圧縮処理は、図 1 a および図 1 b の各々に例示されたステップである 2 つの連続するステップを含む。個々の信号の正確な定義は、H O A 分解および再合成の詳細な説明の項目に記載されている。長さ B の H O A 係数列の重複しない入力フレーム D ( k ) を用いた圧縮のためのフレーム単位の処理が使用される。ここで、k は、フレームのインデックスを表す。フレームは、下記の式 ( 1 ) に特定された H O A 係数列に関して規定される。

【数 1】

$$D(k) := [d((kB + 1)T_s) \quad d((kB + 2)T_s) \quad \dots \quad d((kB + B)T_s)] \quad (1)$$

10

ここで、 $T_s$  は、サンプリング期間を表す。

【0 0 2 3】

図 1 a において、H O A 係数列のフレーム D ( k ) は、支配的な音源方向推定ステップまたはステージ 1 1 に入力され、このステップ 1 1 で、支配的な方向性信号の存在に対して H O A 表現が分析され、その方向が推定される。その方向の推定が行われ、例えば、欧州特許出願公開第 2 6 6 5 2 0 8 号に記載された処理によって行うことができる。その推定された方向は、

$$\hat{\mathbf{n}}_{\text{DOM},1}(k), \dots, \hat{\mathbf{n}}_{\text{DOM},D}(k)$$

によって表される。ここで、添字 D は方向推定値の個数を表す。方向推定値は行列

20

$$\mathbf{A}_{\hat{\mathbf{n}}}(k)$$

に、下記のように配列されるものと仮定される。

【数 2】

$$\mathbf{A}_{\hat{\mathbf{n}}}(k) := [\hat{\mathbf{n}}_{\text{DOM},1}(k) \quad \dots \quad \hat{\mathbf{n}}_{\text{DOM},D}(k)] \quad (2)$$

【0 0 2 4】

暗黙的に、方向推定値は、これらを従前のフレームからの方向推定値に割り当てることによって適切に順序付けられるものと仮定される。したがって、個々の方向推定値の時間的な列は、支配的な音源の方向軌跡を記述するものと仮定される。特に、d 番目の支配的な音源がアクティブでないと想定される場合には、

30

$$\hat{\mathbf{n}}_{\text{DOM},d}(k)$$

に無効値を割り当てることによってこれを示すことができる。そして、

$$\mathbf{A}_{\hat{\mathbf{n}}}(k)$$

において推定された方向を利用して、H O A 表現は、分解ステップまたはステージ 1 2 において最大の数 D の支配的な方向性信号  $X_{\text{DIR}}(k - 1)$  と、支配的な方向性信号からの残差の H O A 成分の空間領域信号の予測を記述する幾らかのパラメータ  $(k - 1)$  と、予測誤りを表すアンビエント H O A 成分  $D_A(k - 2)$  とに分解される。H O A 分解の項目でこの分解についての詳細な説明を行う。

40

【0 0 2 5】

図 1 b において、方向性信号  $X_{\text{DIR}}(k - 1)$  の知覚符号化、および残差のアンビエント H O A 成分  $D_A(k - 2)$  の知覚符号化が示されている。方向性信号  $X_{\text{DIR}}(k - 1)$  は、従来の時間領域信号であり、この信号は、任意の既存の知覚圧縮技術を使用して個々に圧縮することができる。アンビエント H O A 領域成分  $D_A(k - 2)$  の圧縮は、2 つの連続したステップまたはステージで実行することができる。低次元化ステップまたはステージ 1 3 において、アンビソニックス次数  $N_{\text{RED}}$  の低減が行われる。ここで、例えば  $N_{\text{RED}} = 1$  である。結果として、アンビエント H O A 成分  $D_{A, \text{RED}}(k - 2)$  が得られる。このような低次元化は、 $D_A(k - 2)$  において、 $N_{\text{RED}}$  H O A 係数のみを

50



保持し、他の係数を破棄することによって行われる。復号器側では、以下に説明するように、省略された値に対して対応する零値が付加される。

【 0 0 2 6 】

なお、欧州特許出願公開第 2 6 6 5 2 0 8 号のアプローチと比較して、低減された次数  $N_{RED}$  は、一般的には、小さくなるように選択されることがある。この理由は、全体の次数、さらに、残差のアンビエント H O A 成分の方向性の残量が小さくなるからである。したがって、低次元化により、欧州特許出願公開第 2 6 6 5 2 0 8 号の場合と比較して誤りが小さくなる。

【 0 0 2 7 】

以下の相関除去ステップまたはステージ 1 4 において、低次元化されたアンビエント H O A 成分  $D_{A, RED}(k-2)$  を表現する H O A 係数列は相関除去され、時間領域信号  $W_{A, RED}(k-2)$  が得られる。この時間領域信号は、任意の知覚圧縮技術によって動作する（バンクの）パラレル知覚符号化器またはコンプレッサ 1 5 に入力される。この相関除去は、圧縮解除した後に H O A 表現をレンダリングする際に知覚符号化ノイズのマスク除去を回避するために行われる（説明については、欧州特許出願第 1 2 3 0 5 8 6 0 号参照）。近似的な相関除去は、欧州特許出願公開第 2 4 6 9 7 4 2 号に記載されているように、球面調和変換を適用して  $D_{A, RED}(k-2)$  を空間領域内の  $RED$  等価信号に変換することによって成し遂げることができる。

【 0 0 2 8 】

代替的には、欧州特許出願第 1 2 3 0 5 8 6 1 号において提案されている適応的球面調和変換を使用できる。ここでは、最大限の相関除去効果を得るためにサンプリング方向のグリッドを回転させる。別の代替的な相関解除技術は、欧州特許出願第 1 2 3 0 5 8 6 0 号に記載されているカルーネンレーベ変換（KLT）である。なお、これらの最後の 2 つのタイプの相関除去のために、H O A 圧縮解除ステージでの相関除去の逆処理を可能にするべく、 $(k-2)$  で表される何らかの副情報が供給される。

【 0 0 2 9 】

一実施形態においては、符号化効率を改善するために、全ての時間領域信号  $X_{DIR}(k-1)$  および  $W_{A, RED}(k-2)$  の知覚圧縮が共に行われる。

【 0 0 3 0 】

知覚符号化の出力は、圧縮された方向性信号

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

および圧縮されたアンビエント時間領域信号

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

である。

【 0 0 3 1 】

圧縮解除処理

圧縮解除処理は図 2 a および図 2 b に示されている。圧縮処理の場合と同様に、圧縮解除処理は 2 つの連続したステップからなる。図 2 a において、方向性信号

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

および残差のアンビエント H O A 成分を表現する時間領域信号

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

の知覚圧縮解除が、知覚復号または知覚圧縮解除のステップまたはステージ 2 1 において行われる。結果として得られる知覚圧縮解除された時間領域信号

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

は次数  $N_{RED}$  の残差の成分の H O A 表現

10

20

30

40

50

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

を供給するために、再相関ステップまたはステージ 2 2 において再相関される。必要に応じて、この再相関は、ステップ / ステージ 1 4 に記載された 2 つの代替的な処理に対して記載されたのとは逆の手順で実行することができ、使用された相関解除方法に依存して送信あるいは格納されたパラメータ (  $k-2$  ) が使用される。その後、次数拡張によって、次数拡張ステップまたはステージ 2 3 において、

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

から、次数  $N$  の適切な H O A 表現

10

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

が推定される。次数拡張は、対応する「零」値の列を

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

に付加することによって行われ、これにより、より高い次数に関し、H O A 係数が零値を有するものと仮定する。

【 0 0 3 2 】

図 2 b において、全ての H O A 表現は、圧縮解除された支配的な方向性信号

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

20

が対応する方向

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

および予測パラメータ (  $k-1$  ) とから、さらに、残差のアンビエント H O A 成分

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

から、合成ステップまたはステージ 2 4 において再合成される。結果として、H O A 係数の圧縮解除され再合成されたフレーム

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

30

となる。

【 0 0 3 3 】

符号化効率を改善するために、全ての時間領域信号  $X_{DIR}(k-1)$  および  $W_{A,RED}(k-2)$  の知覚圧縮が共に行われた場合には、圧縮された方向性信号

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

および圧縮された時間領域信号

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

40

の知覚圧縮解除もまた、対応する方法で共に行われる。

【 0 0 3 4 】

再合成の詳細な説明は、H O A 再合成の項目に存在する。

【 0 0 3 5 】

H O A 分解

H O A 分解のために実行される処理を例示するブロック図が図 3 に与えられている。この処理を以下のように要約する。最初に、平滑化された支配的な方向性信号  $X_{DIR}(k-1)$  は計算され、知覚圧縮のために出力される。次に、支配的な方向性信号の H O A 表現  $D_{DIR}(k-1)$  と当初の H O A 表現  $D(k-1)$  との間の残差は、“ ” 個の数の方向性信号

50

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

によって表現される。これは、均一に分布した方向からの一般的な平面波と考えることができる。これらの方向性信号は、支配的な方向性信号  $X_{\text{DIR}}(k-1)$  から予測される。ここで、予測パラメータ  $(k-1)$  が出力される。最終的に、当初の H O A 表現  $D(k-2)$  と支配的な方向性信号の H O A 表現  $D_{\text{DIR}}(k-1)$  との間の残差  $D_A(k-2)$  が均一に分布した方向からの予測された方向性信号の H O A 表現

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

と共に計算され、出力される。

10

【 0 0 3 6 】

詳細について述べる前に、連続するフレームの間の方向の変化が合成の間の全ての計算された信号に不連続を生じさせることがある点について述べる。したがって、まず、2 B の長さを有する重複するフレームの各々の信号の瞬時推定値が計算される。第 2 に、連続する重複するフレームの結果が適切な窓関数を使用して平滑化される。しかしながら、各平滑化は、1 フレーム分の待ち時間を伴う。

【 0 0 3 7 】

瞬時支配的な方向性信号の計算

H O A 係数列の現在のフレーム  $D(k)$  に対する

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

20

内の推定された音源方向からの、ステップまたはステージ 3 0 での瞬時支配的な方向信号の計算は、M . A . P o l e t t i 著、" T h r e e - D i m e n s i o n a l S u r r o u n d S o u n d S y s t e m s B a s e d o n S p e h r i c a l H a r m o n i c s ( 球面調和関数に基づく 3 次元サラウンド・サウンド・システム ) "、アメリカ音響学会誌、5 3 ( 1 1 )、1 0 0 4 ~ 1 0 2 5 頁、2 0 0 5 年、に記載されたモード・マッチングに基づいている。特に、所与の H O A 信号の最も良い近似となる H O A 表現の方向性信号がサーチされる。

【 0 0 3 8 】

さらに、一般性を失うことなく、下記の式に従って、傾斜角  $\theta_{\text{DOM},d}(k)$  [ 0 , ] および方位角  $\phi_{\text{DOM},d}(k)$  [ 0 , 2 ] ( 図 5 に示す内容を参照されたい ) のベクトルによって、アクティブな支配的な音源の各方向の推定値

30

$$\hat{\Omega}_{\text{DOM},d}(k)$$

を明確に特定できるものと仮定する。

【 数 3 】

$$\hat{\Omega}_{\text{DOM},d}(k) := \left( \hat{\theta}_{\text{DOM},d}(k), \hat{\phi}_{\text{DOM},d}(k) \right)^T \quad (3)$$

【 0 0 3 9 】

まず、アクティブ音源の方向推定値に基づくモード行列は、下記の式に従って計算され

40

【 数 4 】

$$\Xi_{\text{ACT}}(k) :=$$

$$\left[ S_{\text{DOM},d_{\text{ACT},1}}(k) \quad S_{\text{DOM},d_{\text{ACT},2}}(k) \quad \dots \quad S_{\text{DOM},d_{\text{ACT},D_{\text{ACT}}(k)}}(k) \right] \in \mathbb{R}^{O \times D_{\text{ACT}}(k)} \quad (4)$$

ここで、

【数 5】

$$\mathbf{S}_{\text{DOM},d}(k) :=$$

$$\left[ S_0^0(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), S_1^{-1}(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), S_1^0(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), \dots, S_N^N(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)) \right]^T \in \mathbb{R}^O \quad (5)$$

式 (4) において、 $\mathbf{D}_{\text{ACT}}(k)$  は、 $k$  番目のフレームに対するアクティブな方向の数を表しており、 $d_{\text{ACT},j}(k)$ ,  $1 \leq j \leq \mathbf{D}_{\text{ACT}}(k)$  は、それらの添え字を示している。また、

$$S_n^m(\cdot)$$

10

は、実数値の球面調和関数を示しており、これは、実数値の球面調和関数の定義の項目で定義されている。

【0040】

第2に、行列

$$\tilde{\mathbf{X}}_{\text{DIR}}(k) \in \mathbb{R}^{D \times 2B}$$

が下記の式にしたがって計算され、これは、 $(k-1)$  番目および  $k$  番目のフレームに対する全ての支配的な方向性信号の瞬時推定値を含む。

【数 6】

$$\tilde{\mathbf{X}}_{\text{DIR}}(k) := [\tilde{x}_{\text{DIR}}(k, 1) \quad \tilde{x}_{\text{DIR}}(k, 2) \quad \dots \quad \tilde{x}_{\text{DIR}}(k, 2B)] \quad (6)$$

20

ここで、

【数 7】

$$\tilde{x}_{\text{DIR}}(k, l) := [\tilde{x}_{\text{DIR},1}(k, l), \tilde{x}_{\text{DIR},2}(k, l), \dots, \tilde{x}_{\text{DIR},D}(k, l)]^T \in \mathbb{R}^D, 1 \leq l \leq 2B \quad (7)$$

この計算は、2つのステップで行うことができる。第1のステップにおいては、アクティブでない方向に対応する列の方向性信号サンプルが零に設定され、すなわち、以下のようになる。

【数 8】

$$\tilde{x}_{\text{DIR},d}(k, l) = 0 \quad \forall 1 \leq l \leq 2B, \quad \text{if } d \notin \mathcal{M}_{\text{ACT}}(k), \quad (8)$$

30

ここで、 $\mathcal{M}_{\text{ACT}}(k)$  は、アクティブな方向の組である。第2のステップにおいて、アクティブな方向に対応する方向性信号サンプルは、まず、これらを下記に従った行列に配列することによって取得できる。

【数 9】

$$\tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) := \begin{bmatrix} \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k, 1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k, 2B) \\ \vdots & \ddots & \vdots \\ \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(k, 1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(k, 2B) \end{bmatrix} \quad (9)$$

この行列は、次に、下記の誤りのユークリッドノルムを最小にするように計算される。

40

【数 10】

$$\mathbf{E}_{\text{ACT}}(k) \tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) - [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (10)$$

この解は、下記の式によって与えられる。

【数 11】

$$\tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) = [\mathbf{E}_{\text{ACT}}^T(k) \mathbf{E}_{\text{ACT}}(k)]^{-1} \mathbf{E}_{\text{ACT}}^T(k) [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (11)$$

【0041】

時間的平滑化

ステップまたはステージ 31 に関しては、方向性信号

50

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

についてのみ平滑化を説明する。その理由は、信号の他のタイプの平滑化は、完全に類似の方法で行うことができるからである。式 ( 6 ) に従った行列

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

にサンプルが含まれる方向性信号の推定値

$$\tilde{x}_{\text{DIR},d}(k,l), \quad 1 \leq d \leq \mathcal{D}$$

は、適切な窓関数  $w(1)$  によって窓を掛けられる。

10

【数 1 2】

$$\tilde{x}_{\text{DIR},\text{WIN},d}(k,l) := \tilde{x}_{\text{DIR},d}(k,l) \cdot w(l), \quad 1 \leq l \leq 2B \quad (12)$$

この窓関数は、重複領域においてシフトされたバージョンを用いて (  $B$  個のサンプルのシフトがあると仮定する )、合計で「 1 」となる条件を満たさなければならない。

【数 1 3】

$$w(l) + w(B+l) = 1 \quad \forall 1 \leq l \leq B \quad (13)$$

このような窓関数の例は、下記の式によって定義されるハン窓 ( H a n n w i n d o w ) によって与えられる。

20

【数 1 4】

$$w(l) := 0.5 \left[ 1 - \cos \left( \frac{2\pi(l-1)}{2B} \right) \right] \quad \text{for } 1 \leq l \leq 2B \quad (14)$$

(  $k - 1$  ) 番目のフレームに対する平滑化された方向性信号は、下記の式に従って窓を掛けられた瞬時推定値の適切な重ね合わせによって計算される。

【数 1 5】

$$x_{\text{DIR},d}((k-1)B+l) = \tilde{x}_{\text{DIR},\text{WIN},d}(k-1, B+l) + \tilde{x}_{\text{DIR},\text{WIN},d}(k,l) \quad (15)$$

(  $k - 1$  ) 番目のフレームに対する全ての平滑化された方向性信号のサンプルは、下記の行列  $\mathbf{X}_{\text{DIR}}(k-1)$  に配列される。

30

【数 1 6】

$$\mathbf{X}_{\text{DIR}}(k-1) :=$$

$$[x_{\text{DIR}}((k-1)B+1) \quad x_{\text{DIR}}((k-1)B+2) \quad \dots \quad x_{\text{DIR}}((k-1)B+B)] \in \mathbb{R}^{\mathcal{D} \times B} \quad (16)$$

ここで、

【数 1 7】

$$\mathbf{x}_{\text{DIR}}(l) = [x_{\text{DIR},1}(l), x_{\text{DIR},2}(l), \dots, x_{\text{DIR},\mathcal{D}}(l)]^T \in \mathbb{R}^{\mathcal{D}} \quad (17)$$

平滑化された支配的な方向性信号  $\mathbf{x}_{\text{DIR},d}(l)$  は連続した信号であると想定され、これらの信号は知覚符号化器に順次入力される。

40

【0 0 4 2】

平滑化された支配的な方向性信号の H O A 表現の計算

$\mathbf{X}_{\text{DIR}}(k-1)$  および

$$\mathbf{A}_{\hat{\Omega}}(k)$$

から、ステップまたはステージ 3 2 において、連続的な信号  $\mathbf{x}_{\text{DIR},d}(l)$  に依存して、H O A 合成のために行われる処理と同様の処理を真似るために、平滑化された支配的な方向性信号の H O A 表現が計算される。連続するフレーム間の方向推定値の変化が不連続を生じさせることがあるため、長さ  $2B$  の重複するフレームの瞬時 H O A 表現が再び計

50

算され、連続して重複するフレームの結果が適切な窓関数を使用することによって平滑化される。よって、H O A 表現  $D_{DIR}(k-1)$  は、以下の式によって取得される。

【数 18】

$$D_{DIR}(k-1) = \mathcal{E}_{ACT}(k)X_{DIR,ACT,WIN1}(k-1) + \mathcal{E}_{ACT}(k-1)X_{DIR,ACT,WIN2}(k-1) \quad (18)$$

ここで、

【数 19】

$$X_{DIR,ACT,WIN1}(k-1) := \begin{bmatrix} x_{DIR,d_{ACT,1}(k)}((k-1)B+1) \cdot w(1) & \dots & x_{DIR,d_{ACT,1}(k)}(kB) \cdot w(B) \\ x_{DIR,d_{ACT,2}(k)}((k-1)B+1) \cdot w(1) & & x_{DIR,d_{ACT,2}(k)}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ x_{DIR,d_{ACT,D_{ACT}(k)}(k)}((k-1)B+1) \cdot w(1) & \dots & x_{DIR,d_{ACT,D_{ACT}(k)}(k)}(kB) \cdot w(B) \end{bmatrix} \quad (19) \quad 10$$

さらに、

【数 20】

$$X_{DIR,ACT,WIN2}(k-1) := \begin{bmatrix} x_{DIR,d_{ACT,1}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & x_{DIR,d_{ACT,1}(k-1)}(kB) \cdot w(2B) \\ x_{DIR,d_{ACT,2}(k-1)}((k-1)B+1) \cdot w(B+1) & & x_{DIR,d_{ACT,2}(k-1)}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ x_{DIR,d_{ACT,D_{ACT}(k-1)}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & x_{DIR,d_{ACT,D_{ACT}(k-1)}(k-1)}(kB) \cdot w(2B) \end{bmatrix} \quad (20) \quad 20$$

【0043】

均一なグリッド上の方向性信号によって残差H O A 表現を表現すること

$D_{DIR}(k-1)$  および  $D(k-1)$  (すなわち、フレーム遅延 381 によって遅延された  $D(k)$ ) から、均一なグリッド上の方向性信号による残差H O A 表現がステップまたはステージ 33 で計算される。この処理の目的は、残差  $[D(k-2) D(k-1)] - [D_{DIR}(k-2) D_{DIR}(k-1)]$  を表すために、何らかの固定された、ほぼ均一に分布する方向

$$\hat{\mathbf{n}}_{GRID,o}, \quad 1 \leq o \leq O$$

(グリッド方向とも称する) から到来する方向性信号 (すなわち、一般的な平面波関数) を取得することにある。

【0044】

最初に、グリッド方向に関して、モード行列  $\mathbf{S}_{GRID}$  が下式のように計算される。

【数 21】

$$\mathbf{S}_{GRID} := [\mathbf{S}_{GRID,1} \quad \mathbf{S}_{GRID,2} \quad \dots \quad \mathbf{S}_{GRID,O}] \in \mathbb{R}^{O \times O} \quad (21)$$

ここで、

【数 22】

$$\mathbf{S}_{GRID,o} := [S_0^o(\hat{\mathbf{n}}_{GRID,o}), S_1^{-1}(\hat{\mathbf{n}}_{GRID,o}), S_1^0(\hat{\mathbf{n}}_{GRID,o}), \dots, S_N^N(\hat{\mathbf{n}}_{GRID,o})]^T \in \mathbb{R}^O \quad (22)$$

圧縮処理全体の間、グリッド方向は固定されているためモード行列  $\mathbf{S}_{GRID}$  の計算が必要となるのは一度のみである。

【0045】

各グリッド上の方向性信号は、下記の式によって取得される。

【数 23】

$$\tilde{\mathbf{X}}_{GRID,DIR}(k-1) = \mathbf{S}_{GRID}^{-1}([\mathbf{D}(k-2) \quad \mathbf{D}(k-1)] - [\mathbf{D}_{DIR}(k-2) \quad \mathbf{D}_{DIR}(k-1)]) \quad (23)$$

## 【 0 0 4 6 】

支配的な方向性信号からの均一なグリッド上の方向性信号の予測

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

および  $X_{\text{DIR}}(k-1)$  から、ステップまたはステージ 3 4 で均一なグリッド上の方向性信号が予測される。方向性信号からのグリッド方向

$$\tilde{\Omega}_{\text{GRID},o}, \quad 1 \leq o \leq O$$

から構成される均一なグリッド上の方向性信号の予測は、平滑化の目的で、2つの連続したフレームに基づく、すなわち、(長さ 2B の)グリッド信号

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

の拡張されたフレームは、平滑化された支配的な方向性信号の拡張されたフレームから下記のように予測される。

## 【 数 2 4 】

$$\tilde{X}_{\text{DIR,EXT}}(k-1) := [X_{\text{DIR}}(k-3) \quad X_{\text{DIR}}(k-2) \quad X_{\text{DIR}}(k-1)] \quad (24)$$

## 【 0 0 4 7 】

最初に、

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

に含まれる各グリッド信号

$$\tilde{x}_{\text{GRID,DIR},o}(k-1,l), \quad 1 \leq o \leq O$$

が

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

に含まれる支配的な方向性信号

$$\tilde{x}_{\text{DIR,EXT},d}(k-1,l), \quad 1 \leq d \leq D$$

に割り当てられる。この割り当ては、グリッド信号と全ての支配的な方向性信号との間の正規化された相互相関関数の計算に基づくことができる。特に、その支配的な方向性信号はグリッド信号に割り当てられ、これは正規化された相互相関関数の最も高い値をもたらすグリッド。この割り当ての結果は、 $f_{A,k-1}(\cdot)$  番目の支配的な方向性信号に割り当てる割り当て関数

$$f_{A,k-1}: \{1, \dots, O\} \rightarrow \{1, \dots, D\}$$

によって定式化することができる。

## 【 0 0 4 8 】

次に、各グリッド信号

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

は、割り当てられた支配的な方向性信号

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

から予測される。予測されたグリッド信号

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

は、割り当てられた支配的な方向性信号

10

20

30

40

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

からの遅延およびスケーリングによって、以下のように計算することができる。

【数 2 5】

$$\hat{x}_{\text{GRID,DIR},o}(k-1,l) = K_o(k-1) \cdot \tilde{x}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l-\Delta_o(k-1)) \quad (25)$$

ここで、 $K_o(k-1)$  は、スケーリング係数であり、 $\Delta_o(k-1)$  は、サンプル遅延を示している。これらのパラメータは、予測誤りを最小にするように選択される。

【0049】

予測誤りの次数がグリッド信号自体のものよりも大きい場合には、予測が失敗していると想定される。そして、各予測パラメータを任意の無効値に設定することができる。

10

【0050】

なお、予測を他のタイプにすることも可能である。例えば、全帯域のスケーリング係数を計算するかわりに、知覚指向の周波数帯域に対するスケーリング係数を求めることも合理的である。しかしながら、この処理では、予測が改善するものの、副情報の量が増えてしまう。

【0051】

全ての予測パラメータは、下記のように、パラメータ行列に配列させることができる。

【数 2 6】

$$\zeta(k-1) := \begin{bmatrix} f_{\mathcal{A},k-1}(1) & K_1(k-1) & \Delta_1(k-1) \\ f_{\mathcal{A},k-1}(2) & K_2(k-1) & \Delta_2(k-1) \\ \vdots & \vdots & \vdots \\ f_{\mathcal{A},k-1}(O) & K_O(k-1) & \Delta_O(k-1) \end{bmatrix} \quad (26)$$

20

全ての予測された信号

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

は、行列

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

30

に配列されていると仮定される。

【0052】

均一なグリッド上の予測された方向性信号の H O A 表現の計算

予測されたグリッド信号の H O A 表現は、ステップまたはステージ 3 5 において、下記の式に従って

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

から計算される。

【数 2 7】

$$\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (27)$$

40

【0053】

残差のアンビエント音場成分の H O A 表現の計算

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

の (ステップ / ステージ 3 6 における) 時間的平滑化されたバージョンである

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

と、 $\mathbf{D}(k)$  の 2 フレーム遅延されたバージョンである (遅延 3 8 1 および 3 8 3)  $\mathbf{D}(k-2)$  と、 $\mathbf{D}_{\text{DIR}}(k-1)$  の 1 フレーム遅延されたバージョン (遅延 3 8 2) であ

50



る  $D_{DIR}(k-2)$  とから、残差のアンビエント音場成分の H O A 表現がステップまたはステージ 37 において、下記の式によって計算される。

【数 28】

$$D_A(k-2) = D(k-2) - \hat{D}_{GRID,DIR}(k-2) - D_{DIR}(k-2) \quad (28)$$

【0054】

H O A 再合成

図 4 における個々のステップまたはステージの処理について詳細に説明する前に、概要について述べる。均一に分布した方向に対して方向性信号

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

10

は、予測パラメータ

$$\hat{\zeta}(k-1)$$

を使用して、復号された支配的な方向性信号

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

から予測される。次に、支配的な方向性信号の H O A 表現

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

20

と、予測された方向性信号の H O A 表現

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

と、残差のアンビエント H O A 成分

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

とから、全体の H O A 表現

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

が合成される。

30

【0055】

支配的な方向性信号の H O A 表現の計算

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

および

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

は、支配的な方向性信号の H O A 表現を求めるために、ステップまたはステージ 41 に入力される。モード行列

$$E_{ACT}(k)$$

40

および

$$E_{ACT}(k-1)$$

を k 番目および (k-1) 番目のフレームに対するアクティブな音源の方向推定値に基づいて方向推定値

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

および

$$\mathbf{A}_{\hat{\Omega}}(k-1)$$

から計算した後、支配的な方向性信号

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

のH O A表現は、下記のように取得される。

【数 2 9】

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

$$\mathbf{E}_{\text{ACT}}(k)\mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) + \mathbf{E}_{\text{ACT}}(k-1)\mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) \quad (29)$$

10

ここで、

【数 3 0】

$$\mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) :=$$

$$\begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}(k)}((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}(k)}(kB) \cdot w(B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}(k)}((k-1)B+1) \cdot w(1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}(k)}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(kB) \cdot w(B) \end{bmatrix} \quad (30)$$

20

並びに、

【数 3 1】

$$\mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) :=$$

$$\begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}(k-1)}(kB) \cdot w(2B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}(k-1)}((k-1)B+1) \cdot w(B+1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}(k-1)}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}(k-1)}(kB) \cdot w(2B) \end{bmatrix} \quad (31)$$

【0 0 5 6】

30

支配的な方向性信号から均一なグリッド上の方向性信号の予測

$$\hat{\zeta}(k-1)$$

および

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

は、支配的な方向性信号から均一なグリッド上の方向性信号を予測するため

に、ステップまたはステージ4 3に入力される。均一なグリッド上の予測された方向性信号の拡張フレームは、下記の式に従って要素

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

40

から構成される。

【数 3 2】

$$\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) = \begin{bmatrix} \hat{x}_{\text{GRID,DIR},1}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},1}(k-1,2B) \\ \hat{x}_{\text{GRID,DIR},2}(k-1,1) & & \hat{x}_{\text{GRID,DIR},2}(k-1,2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{GRID,DIR},O}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},O}(k-1,2B) \end{bmatrix} \quad (32)$$

これは、下記の式によって支配的な方向性信号から予測される。

50

【数 3 3】

$$\hat{x}_{\text{GRID,DIR},o}(k-1, l) = K_o(k-1) \cdot \hat{x}_{\text{DIR},f_{\mathcal{A},k-1}(o)}((k-1)B + l - \Delta_o(k-1)) \quad (33)$$

【0 0 5 7】

均一なグリッド上の予測された方向性信号の H O A 表現の計算

均一なグリッド上の予測された方向性信号の H O A 表現を計算するステップまたはステージ 4 4 において、予測されたグリッド方向性信号の H O A 表現は、下記の式によって取得される。

【数 3 4】

$$\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (34)$$

ここで、 $\mathbf{E}_{\text{GRID}}$  は、所定のグリッド方向に対するモード行列を表す（定義については、等式 (2 1) を参照。）。

【0 0 5 8】

H O A 音場表現の合成

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

（すなわち、フレーム遅延 4 2 によって遅延された

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

）と、

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

（ステップ / ステージ 4 5 において、

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

の時間的平滑化されたバージョン）と、

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

とから、ステップまたはステージ 4 6 において全体の音場表現が最終的に下記のように合成される。

【数 3 5】

$$\hat{\mathbf{D}}(k-2) = \hat{\mathbf{D}}_{\text{DIR}}(k-2) + \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2) + \hat{\mathbf{D}}_{\text{A}}(k-2) \quad (35)$$

【0 0 5 9】

高次アンビソニックスの基礎

高次アンビソニックスは注目されるコンパクトな領域内の音場の記述に基づいており、音源が存在しないものと仮定される。その場合、注目領域内の時間  $t$  および位置  $x$  で音圧  $p(t, x)$  の空間時間的な挙動は、均質媒質の波動方程式によって物理的に完全に求められる。以下の内容は、図 5 に示された球面座標システムに基づいている。 $x$  軸は、前方の位置を指し、 $y$  軸は、左側を指し、 $z$  軸は上方を指す。空間内の位置  $x = (r, \theta, \phi)^T$  は、半径  $r > 0$ （すなわち、座標原点へ距離）、極軸  $z$  から測定される傾斜角  $\theta \in [0, \pi]$ 、さらに、 $x$  軸からの、 $x - y$  平面内で反時計周りに測定される、方位角  $\phi \in [0, 2\pi]$  によって表される。 $(\cdot)^T$  は、転置を表す。

【0 0 6 0】

$\mathbf{F}_t(\cdot)$  によって表される時間に対する音圧のフーリエ変換、すなわち、

【数 3 6】

$$P(\omega, x) = \mathcal{F}_t(p(t, x)) = \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt \quad (36)$$

は下記の式に従った一連の球面調和関数に拡張される（E. G. W i l l i a m s 著 “ F

10

20

30

40

50

ourier Acoustics (フーリエ・アコースティクス))”、応用数理学、第93巻、アカデミックプレス社、1999年参照)。ここで、 $k$  は角周波数を表し、 $i$  は虚数単位を表す。

【数37】

$$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi) \quad (37)$$

ここで、 $c_s$  は音速を示し、 $k$  は角波数を示し、この角波数  $k$  は  $k = \omega / c_s$  によって角周波数  $\omega$  に関連している。 $j_n(\cdot)$  は、第1種球ベッセル関数を表しており、

$$S_n^m(\theta, \phi)$$

10

は、実数値の球面調和関数の定義の項目で定義されている次数  $n$  および位数  $m$  の実数値の球面調和関数を示している。展開係数

$$A_n^m(k)$$

は、角波数  $k$  のみに依存する。なお、音圧は、空間的に帯域制限されているものと暗黙的に仮定されている。したがって、級数が次数インデックス  $n$  に対して上限  $N$  で打ち切られ、これは、HOA表現の次数と呼ばれる。

【0061】

音場が相異なる角周波数の調和平面波の無限個の重ね合わせによって表現され、角の組  $(\theta, \phi)$  によって特定される全ての想定可能な方向から到来する場合には、各々の平面波複素振幅関数  $D(\theta, \phi, k)$  は、下記の球面調和展開によって表すことができることが分かる(B. Rafaely 著、“Plane-wave Decomposition of the Sound Field on a Sphere by Spherical Convolution (球面畳み込みによる球面上の音場の平面波分解)”、米国音響学会誌4(116)、2149-2157頁、2004年参照)。

20

【数38】

$$D(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n D_n^m(k) S_n^m(\theta, \phi) \quad (38)$$

ここで、展開係数

$$D_n^m(k)$$

30

は、

$$A_n^m(k)$$

と下記の式によって関連する。

【数39】

$$A_n^m(k) = 4\pi i^n D_n^m(k) \quad (39)$$

【0062】

個々の係数

40

$$D_n^m(k = \omega/c_s)$$

が角周波数  $\omega$  の関数であると仮定すると、逆フーリエ変換

$$\mathcal{F}_t^{-1}(\cdot)$$

によって示される)を適用することにより、各次数  $n$  および位数  $m$  に対し、下記の時間領域関数をもたらす。

【数 4 0】

$$d_n^m(t) = \mathcal{F}_t^{-1} \left( D_n^m \left( \frac{\omega}{c_s} \right) \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} D_n^m \left( \frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (40)$$

これは、次数  $n$  および位数  $m$  に対して、下記の単一のベクトルにまとめられる。

【数 4 1】

$$\mathbf{d}(t) = \begin{bmatrix} d_0^0(t) & d_1^{-1}(t) & d_1^0(t) & d_1^1(t) & d_2^{-2}(t) & d_2^{-1}(t) & d_2^0(t) & d_2^1(t) & d_2^2(t) & \dots \\ d_N^{N-1}(t) & d_N^N(t) & & & & & & & & \end{bmatrix}^T \quad (41)$$

10

ベクトル  $\mathbf{d}(t)$  内の時間領域関数

$$d_n^m(t)$$

の位置インデックスは、 $n(n+1) + 1 + m$ によって与えられる。

【0 0 6 3】

最終的なアンビソニックス形式は、サンプリング周波数  $f_s$  を使用して、下記の  $\mathbf{d}(t)$  のサンプリングされたバージョンをもたらす。

【数 4 2】

$$\{\mathbf{d}(lT_s)\}_{l \in \mathbb{N}} = \{\mathbf{d}(T_s), \mathbf{d}(2T_s), \mathbf{d}(3T_s), \mathbf{d}(4T_s), \dots\} \quad (42)$$

20

ここで、 $T_s = 1 / f_s$  は、サンプリング期間を示す。 $\mathbf{d}(lT_s)$  の要素は、アンビソニックス係数として参照される。なお、時間領域信号、

$$d_n^m(t)$$

は、実数値であり、したがって、アンビソニックス係数は、実数値である。

【0 0 6 4】

実数値の球面調和関数の定義

実数値の球面調和関数

$$S_n^m(\theta, \phi) \quad (43)$$

30

は、下記の式によって与えられる。

【数 4 3】

$$S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_{n,|m|}(\cos \theta) \text{trg}_m(\phi) \quad (43)$$

ここで

【数 4 4】

$$\text{trg}_m(\phi) = \begin{cases} \sqrt{2} \cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2} \sin(m\phi) & m < 0 \end{cases} \quad (44)$$

40

関連するルジャンドル関数  $P_{n,m}(x)$  は、下記の式で定義される。

【数 4 5】

$$P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0 \quad (45)$$

ここで、ルジャンドル多項式  $P_n(x)$  を用い、上述した、E. G. Williams 著のテキストブックの場合とは異なり、コンドン-ショートレーの位相項  $(-1)^m$  を用いない。

50

【 0 0 6 5 】

高次アンビソニックスの空間解像度

方向  $\mathbf{0} = (\mathbf{0}, \mathbf{0})^T$  から到来する一般的な平面波関数  $x(\mathbf{t})$  は、下記の式によって H O A において表現される。

【数 4 6】

$$d_n^m(t) = x(t) S_n^m(\boldsymbol{\Omega}_0), \quad 0 \leq n \leq N, |m| \leq n \quad (46)$$

平面波振幅の対応する空間密度

$$d(t, \boldsymbol{\Omega}) := \mathcal{F}_t^{-1}(D(\omega, \boldsymbol{\Omega}))$$

10

は、下記の式によって与えられる。

【数 4 7】

$$d(t, \boldsymbol{\Omega}) = \sum_{n=0}^N \sum_{m=-n}^n d_n^m(t) S_n^m(\boldsymbol{\Omega}) \quad (47)$$

$$= x(t) \underbrace{\left[ \sum_{n=0}^N \sum_{m=-n}^n S_n^m(\boldsymbol{\Omega}_0) S_n^m(\boldsymbol{\Omega}) \right]}_{v_N(\theta)} \quad (48)$$

式 (48) から理解されるように、これは、一般的な平面波関数  $x(\mathbf{t})$  と空間分散関数  $v_N(\theta)$  との積であり、空間分散関数  $v_N(\theta)$  は、下記の式の特性を有する  $\theta_0$  と  $\theta$  の間の角度  $\theta$  のみに依存するように示されている。

20

【数 4 8】

$$\cos\theta = \cos\theta\cos\theta_0 + \cos(\phi - \phi_0)\sin\theta\sin\theta_0 \quad (49)$$

想定のとおり、無限次元の極限、つまり、 $N \rightarrow \infty$  である場合において、空間分散関数はディラックのデルタ関数  $\delta(\cdot)$ 、すなわち、下記のように変化する。

【数 4 9】

$$\lim_{N \rightarrow \infty} v_N(\theta) = \frac{\delta(\theta)}{2\pi} \quad (50)$$

しかしながら、有限次元  $N$  の場合には、方向  $\mathbf{0}$  からの一般的な平面波の寄与は、近隣の方向ににじみ、このにじみの度合いは次数の増加に伴い減少する。 $N$  の複数の異なる値に対する正規化された関数  $v_N(\theta)$  のプロットが図 6 に示されている。任意の方向  $\mathbf{0}$  での平面波振幅の空間密度の時間領域の挙動は、他の任意の方向での平面波振幅の空間密度の時間領域の挙動の倍数となることが指摘される。特に、時間  $t$  に対して、何らかの固定方向  $\mathbf{1}$  および  $\mathbf{2}$  についての関数  $d(t, \mathbf{1})$  および  $d(t, \mathbf{2})$  は、高い相関性がある。

30

【 0 0 6 6 】

離散空間領域

平面波振幅の空間密度が  $M$  個の空間方向  $\mathbf{0}$  ( $M=1$  で離散化される場合、空間方向  $\mathbf{0}$  は単位球面上でほぼ均一に分布するのだが、 $M$  個の方向性信号  $d(t, \mathbf{0}_i)$  が取得される。これらの信号をベクトルにまとめると、下記の式で表され、

40

【数 5 0】

$$\mathbf{d}_{\text{SPAT}}(t) := [d(t, \boldsymbol{\Omega}_1) \quad \dots \quad d(t, \boldsymbol{\Omega}_M)]^T \quad (51)$$

式 (47) を使用してこのベクトルを、下記のような単純な行列乗算によって式 (41) に定義される連続的なアンビソニックス表現  $d(\mathbf{t})$  から計算することができることを検証できる。

$$\mathbf{d}_{\text{SPAT}}(\mathbf{t}) = \mathbf{H} \mathbf{d}(\mathbf{t}) \quad (52)$$

ここで、 $(\cdot)^H$  は、複素共役転置を示し、 $\mathbf{H}$  は、下記の式によって定義されるモード行列を表す。

【数 5 1】

$$\Psi = [S_1 \quad \dots \quad S_0] \quad (53)$$

ここで、

【数 5 2】

$$S_o = [S_o^0(\Omega_o) \quad S_1^{-1}(\Omega_o) \quad S_1^0(\Omega_o) \quad S_1^1(\Omega_o) \quad \dots \quad S_N^{N-1}(\Omega_o) \quad S_N^N(\Omega_o)] \quad (54)$$

方向  $\Omega_o$  は単位球面上にほぼ均一に分布しているため、一般的には、モード行列は可逆である。したがって、連続的なアンビソニックス表現は、方向性信号  $d(t, \Omega_o)$  から下記の式によって計算することができる。

$$d(t) = \Psi^H d_{SPAT}(t) \quad (55)$$

双方の式は、アンビソニックス表現と空間領域との間の変換および逆変換を構成する。本願において、これらの変換は、球面調和関数変換および逆球面調和関数変換と呼ばれる。

【0067】

方向  $\Omega_o$  は単位球面上でほぼ均一に分布するため、

【数 5 3】

$$\Psi^H \approx \Psi^{-1} \quad (56)$$

となり、式 (52) において、 $\Psi^H$  の代わりに  $\Psi^{-1}$  を使用することが正当化される。有利には、上述した関係の全ては離散時間領域にも有効である。

【0068】

符号化側、さらに復号側においても、本発明の処理を単一のプロセッサまたは電子回路、または、並列に動作する、および/または、本発明の処理の複数の異なる部分に対して動作する、幾つかのプロセッサまたは電子回路で実行することができる。

【0069】

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 ( $\Omega_o$ )

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

) を推定するステップ (11) と、

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

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

) 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 ( $X_{DTR}(k-1)$ ) と残差の H O A 成分 ( $D_A(k-2)$ ) とに分解するステップ (12) 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 (33)、前記平面波関数が前記支配的な方向性信号 ( $X_{DTR}(k-1)$ ) 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 ( $\Omega_{REF}(k-1)$ ) 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 (35)、該ステップ (12) と、

- 前記残差の H O A 成分 ( $D_A(k-2)$ ) の現在の次数 ( $N$ ) をより低い次数 ( $N_{RED}$ ) に低減するステップ (13) であって、結果として、低次元化された残差の H O A 成分 ( $D_{A,RED}(k-2)$ ) が得られる、該ステップ (13) と、

10

20

30

40

50

- 前記低次元化された残差のH O A成分 ( $D_{A,RED}(k-2)$ ) 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 ( $W_{A,RED}(k-2)$ ) を取得するステップ (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 (

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

10

) 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H O A成分時間領域信号 ( $W_{A,RED}(k-2)$ ) を知覚符号化するステップ (15) 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 (

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

20

) を推定するように構成された手段 (11) と、

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

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

) 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 ( $X_{DIR}(k-1)$ ) と残差のH O A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 (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H O Aの領域に再び変換される (35)、前記手段 (12) と、

30

- 前記残差のH O A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H O A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H O A成分 ( $D_{A,RED}(k-2)$ ) 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 ( $W_{A,RED}(k-2)$ ) 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 (

40

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

) 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H O A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、

50



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 (

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

) 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 (

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

10

) を供給するように、前記圧縮された支配的な方向性信号 (

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

) および前記圧縮された残差の成分信号 (

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

20

) を知覚復号するステップ ( 2 1 ) と、

- 前記圧縮解除された時間領域信号 (

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

) 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 (

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

) 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 (

30

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

) の次数 (  $N_{\text{RED}}$  ) 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 (

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

) 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 (

40

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

) と、前記当初の次数の圧縮解除された残差の H O A 成分 (

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

) と、前記推定された ( 1 1 ) 支配的な音源方向 (

50

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

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

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

を合成するステップ（24）と、

を含む、前記方法。

10

〔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（

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

）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（

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

20

）を供給するように、前記圧縮された支配的な方向性信号（

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

）および前記圧縮された残差の成分信号（

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

）を知覚復号するように構成された手段（21）と、

30

- 前記圧縮解除された時間領域信号（

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

）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（

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

）を取得するように構成された手段（22）と、

40

- 前記低次元化された残差のH O A成分（

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

）の次数（ $N_{\text{RED}}$ ）を当初の次数（ $N$ ）に拡張するように構成された手段（23）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（

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

）を供給する、該手段（23）と、

- 前記圧縮解除された支配的な方向性信号（

50

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

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 (

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

10

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 (  $D_{A, \text{RED}}(k-2)$  ) 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。

20

〔態様 6〕

前記低次元化された残差の H O A 成分 (  $D_{A, \text{RED}}(k-2)$  ) 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、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。

30

〔態様 7〕

前記支配的な方向性信号 (  $X_{\text{DIR}}(k-1)$  ) および前記残差の H O A 成分時間領域信号 (  $W_{A, \text{RED}}(k-2)$  ) の知覚圧縮 ( 1 5 ) が共に行われ、前記圧縮された方向性信号 (

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

および前記圧縮された時間領域信号 (

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

40

の前記知覚圧縮 ( 2 1 ) 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 ( 1 2 ) は、  
- H O A 係数の現在のフレーム (  $D(k)$  ) に対して (

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

50

における推定された音源方向から支配的な方向性信号 (

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

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 ( $X_{\text{DIR}}(k-1)$ ) が取得される、該ステップと、  
- (

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

10

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 ( $X_{\text{DIR}}(k-1)$ ) から平滑化された支配的な方向性信号 ( $D_{\text{DIR}}(k-1)$ ) 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 (

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

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 ( $X_{\text{DIR}}(k-1)$ ) および方向性信号 (

20

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

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 (

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

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の前記平滑化された予測された方向性信号 (

30

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

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 (

40

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

と、前記圧縮解除された支配的な方向性信号 (

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

とから、支配的な方向性信号 (

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

の H O A 表現を計算するステップ ( 4 1 ) と、  
前記圧縮解除された支配的な方向性信号 (

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

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

10

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

を予測するステップ ( 4 3 ) 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

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

を計算するステップ ( 4 4 ) であって、その後、時間的平滑化を行う

20

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

、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

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

の前記平滑化された H O A 表現と、支配的な方向性信号 (

30

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

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 (

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

とから、H O A 音場表現 (

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

40

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 (

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

の前記予測 ( 3 4 ) において、予測されたグリッド信号 (

50

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

）が、割り当てられた支配的な方向性信号（

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

）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（

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

）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。

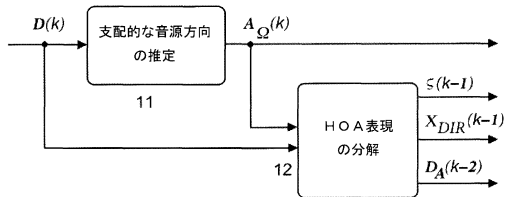
〔態様 1 2〕

態様 1、5～8、10、および 11 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。

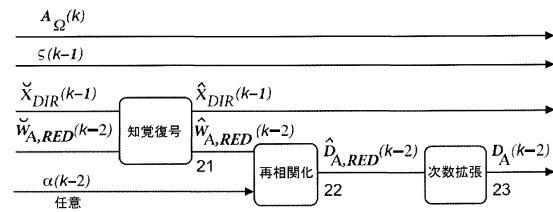
10

20

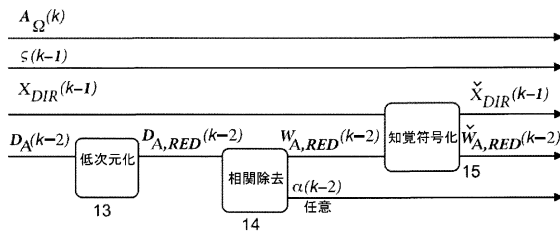
【図 1 a】



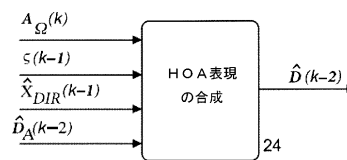
【図 2 a】



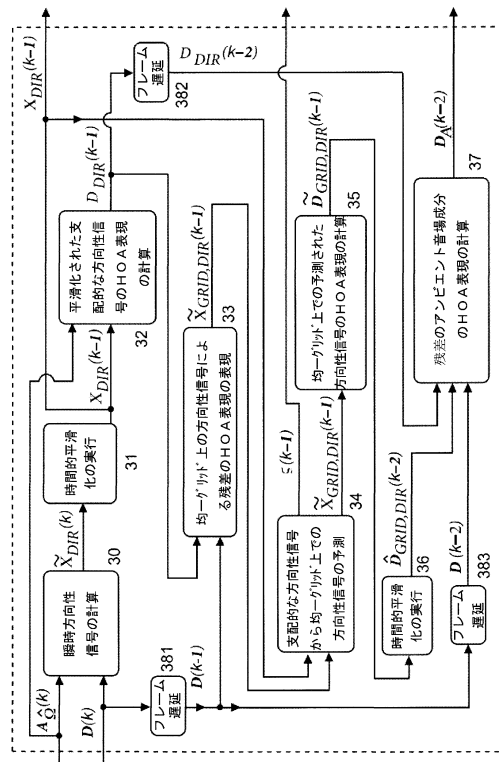
【図 1 b】



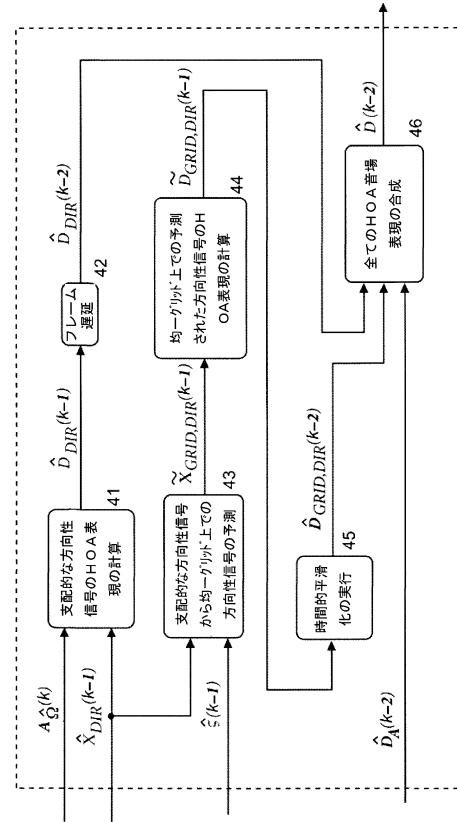
【図 2 b】



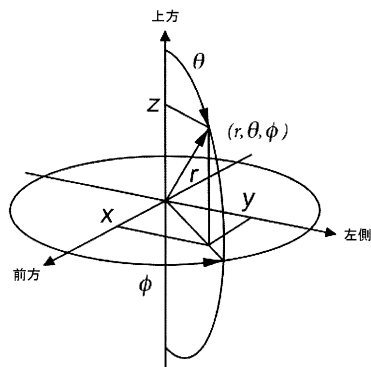
【図 3】



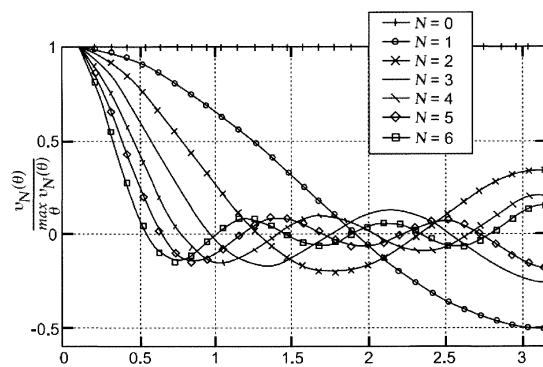
【図 4】



【図 5】



【図 6】



## フロントページの続き

- (72)発明者 クルーガー, アレクサンダー  
ドイツ国 3 0 6 2 5 ハノーファー カール - ヴィーヘルト - アレー 7 4 ドイチェ・トムソ  
ン・オーハーゲー リサーチ・アンド・イノベーション
- (72)発明者 コルドン, スフエン  
ドイツ国 3 0 6 2 5 ハノーファー カール - ヴィーヘルト - アレー 7 4 ドイチェ・トムソ  
ン・オーハーゲー リサーチ・アンド・イノベーション
- (72)発明者 ベーム, ヨハネス  
ドイツ国 3 0 6 2 5 ハノーファー カール - ヴィーヘルト - アレー 7 4 ドイチェ・トムソ  
ン・オーハーゲー リサーチ・アンド・イノベーション

審査官 富澤 直樹

- (56)参考文献 欧州特許出願公開第 0 2 6 6 5 2 0 8 ( E P , A 1 )  
特開 2 0 1 2 - 1 3 3 3 6 6 ( J P , A )  
米国特許出願公開第 2 0 1 2 / 0 1 5 5 6 5 3 ( U S , A 1 )

- (58)調査した分野(Int.Cl. , D B 名)
- |         |                       |
|---------|-----------------------|
| G 1 0 L | 1 9 / 0 0 - 1 9 / 2 6 |
| H 0 4 S | 1 / 0 0 - 7 / 0 0     |