

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号  
特許第6395811号  
(P6395811)

(45) 発行日 平成30年9月26日(2018.9.26)

(24) 登録日 平成30年9月7日(2018.9.7)

(51) Int.Cl.

F I

G 1 O L 19/008 (2013.01)

G 1 O L 25/51 (2013.01)

G 1 O L 19/008 1 O O

G 1 O L 25/51 4 O O

請求項の数 22 (全 36 頁)

(21) 出願番号	特願2016-509473 (P2016-509473)	(73) 特許権者	510185767
(86) (22) 出願日	平成26年4月24日 (2014.4.24)		ドルビー・インターナショナル・アーベー
(65) 公表番号	特表2016-520864 (P2016-520864A)		オランダ王国, セーエン アムステルダム
(43) 公表日	平成28年7月14日 (2016.7.14)		ズイドースト 1101, ヘリケルベル
(86) 国際出願番号	PCT/EP2014/058380		グウェグ 1-35, アポロ ビルディン
(87) 国際公開番号	W02014/177455		グ 3エー
(87) 国際公開日	平成26年11月6日 (2014.11.6)	(74) 代理人	100107766
審査請求日	平成29年3月1日 (2017.3.1)		弁理士 伊東 忠重
(31) 優先権主張番号	13305558.2	(74) 代理人	100070150
(32) 優先日	平成25年4月29日 (2013.4.29)		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】 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】

所定数 ( I ) の知覚符号化処理を使用して音場の H O A と称する高次アンビソニックス表現を H O A 係数列の入力される時間フレーム

$$(C(k), \tilde{C}(k))$$

を用いて圧縮する方法であって、  
前記方法は、フレーム単位で行われ、  
- 現在のフレーム

$$(C(k), \tilde{C}(k))$$

に対して、支配的な方向のセット

$$(\tilde{G}_{\Omega, \text{ACT}}(k))$$

および対応する検出された方向性信号のインデックスのデータセット

$$(\tilde{J}_{\text{DIR}, \text{ACT}}(k))$$

を推定するステップ ( 13 ) と、

- 前記現在のフレームの H O A 係数列を分解するステップ ( 14、15 ) であって、非所定数 ( M ) の方向性信号 ( X<sub>DIR</sub> ( k - 2 ) ) であって、支配的な方向推定値の前記セ

ット

$$(\tilde{G}_{\Omega,ACT}(k))$$

に含まれる各々の方向と前記方向性信号のインデックスの各々の遅延したデータセット

$$(\tilde{J}_{DIR,ACT}(k-2))$$

とを用いた、前記非所定数 (M) が前記所定数 (I) よりも小さい、該非所定数 (M) の方向性信号 ( $X_{DIR}(k-2)$ ) と、前記所定数 (I) と前記非所定数 (M) との差に対応する低減された数の H O A 係数列によって表現される残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) と、対応する前記低減された数の残差のアンビエント H O A 係数列のインデックスのデータセット

10

$$(\tilde{J}_{AMB,ACT}(k-2))$$

と、に分解する、該分解するステップ (14、15) と、

- 前記方向性信号 ( $X_{DIR}(k-2)$ ) および前記残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) の H O A 係数列を前記所定数 (I) に対応する数のチャンネルに割り当てるステップ (16) であって、前記割り当てのために、前記方向性信号のインデックスの前記遅延されたデータセット

$$(\tilde{J}_{DIR,ACT}(k-2))$$

20

および前記低減された数の残差のアンビエント H O A 係数列のインデックスの前記データセット

$$(\tilde{J}_{AMB,ACT}(k-2))$$

が使用される、該割り当てるステップと、

- 関連するフレーム ( $Y(k-2)$ ) の前記チャンネルを知覚符号化するステップ (17) であって、符号化された圧縮されたフレーム

$$(\tilde{Y}(k-2))$$

が得られる、該知覚符号化するステップ (17) と、

30

を含む、前記方法。

【請求項 2】

所定数 (I) の知覚符号化処理を使用して音場の H O A と称する高次アンビソニックス表現を H O A 係数列の入力される時間フレーム

$$(C(k), \tilde{C}(k))$$

を用いて圧縮する装置であって、当該装置は、フレーム単位の処理を実行し、

- 現在のフレーム

$$(C(k), \tilde{C}(k))$$

40

に対して、支配的な方向のセット

$$(\tilde{G}_{\Omega,ACT}(k))$$

および対応する検出された方向性信号のインデックスのデータセット

$$(\tilde{J}_{DIR,ACT}(k))$$

を推定するように構成された手段 (13) と、

- 前記現在のフレームの H O A 係数列を分解するように構成された手段 (14、15) であって、非所定数 (M) の方向性信号 ( $X_{DIR}(k-2)$ ) であって、支配的な方向推定値の前記セット

50

$$(\tilde{G}_{\Omega,ACT}(k))$$

に含まれる各々の方向と、前記方向性信号のインデックスの各々の遅延したデータセット

$$(\tilde{J}_{DIR,ACT}(k-2))$$

とを用いた、前記非所定数 (M) が前記所定数 (I) よりも小さい、該非所定数 (M) の方向性信号 ( $X_{DIR}(k-2)$ ) と、前記所定数 (I) と前記非所定数 (M) との差に対応する低減された数の H O A 係数列によって表現される残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) と、対応する前記低減された数の残差のアンビエント H O A 係数列のインデックスのデータセット

10

$$(\tilde{J}_{AMB,ACT}(k-2))$$

と、に分解するように構成され、割り当てのために、前記方向性信号のインデックスの前記遅延されたデータセット

$$(\tilde{J}_{DIR,ACT}(k-2))$$

および前記低減された数の残差のアンビエント H O A 係数列の前記データセット

$$(\tilde{J}_{AMB,ACT}(k-2))$$

が使用される、該手段 (14、15) と、

20

- 前記方向性信号 ( $X_{DIR}(k-2)$ ) および前記残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) の H O A 係数列を前記所定数 (I) に対応する数のチャンネルに割り当てるように構成された手段 (16) であって、該割り当てにより前記割り当てを記述する選択されたアンビエント H O A 係数列のインデックスのパラメータ

$$(\tilde{J}_{AMB,ACT}(k-2))$$

を取得し、該パラメータが圧縮解除側での対応する再配分に使用可能である、前記手段 (16) と、

- 関連するフレーム ( $Y(k-2)$ ) の前記チャンネルを知覚符号化するように構成された手段 (17) であって、符号化された圧縮されたフレーム

30

$$(\tilde{Y}(k-2))$$

が得られる、該手段 (17) と、

を備える、前記装置。

### 【請求項 3】

前記非所定数 (M) の方向性信号 ( $X_{DIR}(k-2)$ ) は、知覚に関連する基準に従って、

- 前記圧縮のための所定の所与の数のチャンネルを用いることにより、対応して圧縮解除される H O A 表現が最小の知覚可能な誤差をもたらす、前記基準で考慮される誤差は、

- 複数の異なる数の前記方向性信号 ( $X_{DIR}(k-2)$ ) と前記残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) に対して複数の異なる数の H O A 係数列とを使用することから生ずるモデル化誤差と、

40

- 前記方向性信号 ( $X_{DIR}(k-2)$ ) の知覚符号化に伴う量子化雑音と、

- 前記残差のアンビエント H O A 成分 ( $C_{AMB,RED}(k-2)$ ) の個々の H O A 係数列を符号化することに伴う量子化雑音と、であり、

- 前記 3 つの誤差の結果として生ずる合計誤差は、当該合計誤差の知覚可能性に関して、複数のテスト方向および複数の臨界帯域に対して考慮され、

- 前記最小の知覚可能な誤差を達成するために、平均の知覚可能な誤差または最大の知覚可能な誤差を最小化するように前記非所定数 (M) の方向性信号 ( $X_{DIR}(k-2)$ ) が選択される、ように決定される、請求項 1 に記載の方法。

50

## 【請求項 4】

前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) を表現する低減された数のH O A係数列の選択は、3つの場合を区別する基準に従って行われ、前記3つの場合は、

- 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数と同一であるときに、前記前のフレームと同一のH O A係数列が選択される場合と、

- 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数よりも小さいときに、方向性信号によって占められている前記現在のフレーム内のチャンネルに割り当てられた前記前のフレームに存在した、前記前のフレームからのH O A係数列が非アクティブ化される場合と、

- 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数よりも大きいときに、前記前のフレームで選択されていたH O A係数列が前記現在のフレームでも選択され、追加的なH O A係数列が当該追加的なH O A係数列の知覚的な重要性または最も高い平均パワーに従って選択可能である場合と、  
である、請求項 1 または 3 に記載の方法。

10

## 【請求項 5】

前記割り当て ( $16$ ) の実行で、

- アクティブな方向性信号は、前記知覚符号化 ( $17$ ) のための連続的な信号を取得するために、そのアクティブな方向性信号がチャンネル・インデックスを保持するように、所与のチャンネルに割り当てられ、

20

- 前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) のH O A係数列は、最小の数 ( $R E D$ ) の当該H O A係数列が常に対応する数 ( $R E D$ ) の前のチャンネルが含まれるように、割り当てられ、

- 前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) の追加的なH O A係数列を割り当てるために、該追加的なH O A係数列が前記前のフレーム ( $k-1$ ) でも選択されていたかが判定され、

- 前記追加的なH O A係数列が前記前のフレーム ( $k-1$ ) でも選択されていた場合には、前記H O A係数列の知覚符号化 ( $17$ ) されるチャンネルに対する割り当て ( $16$ ) が前記前のフレームに対する割り当てと同一となり、

30

- 前記追加的なH O A係数列が前記前のフレーム ( $k-1$ ) では選択されておらず、H O A係数列が新たに選択されている場合には、前記H O A係数列は、まずインデックスに関して昇順に配列され、当該昇順で、方向性信号によってまだ占められていない知覚符号化 ( $17$ ) されるチャンネルに割り当てられる、

請求項 1、3、および 4 のいずれか 1 項に記載の方法。

## 【請求項 6】

$R E D$  は、前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) を表現するH O A係数列の数であり、前記割り当て ( $16$ ) を記述するパラメータは、前記残差のアンビエントH O A成分を表現する  $R E D$  個のH O A係数列に追加して使用される追加の数のH O A係数列に対応する長さを有するビットアレイに配列され、前記ビットアレイ内で各  $o$  番目のビットは、( $R E D + o$ ) 番目の追加のH O A係数列が前記残差のアンビエントH O A成分を表現するために使用されているかどうかを示す、請求項 1 および 3 ~ 5 のいずれか 1 項に記載の方法。

40

## 【請求項 7】

前記割り当て ( $16$ ) を記述するパラメータは、非アクティブな方向性信号の数に対応する長さを有する割り当てベクトルに配列され、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分の追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかを示す、請求項 1 および 3 ~ 5 のいずれか 1 項に記載の方法。

## 【請求項 8】

50

前記現在のフレームのH O A係数列の分解(14)は、前記方向性信号( $X_{DIR}(k-2)$ )から元のH O A表現の部分を予測するために圧縮解除側で使用可能なパラメータ( $(k-2)$ )をさらに供給する、請求項1および3~7のいずれか1項に記載の方法。

【請求項9】

前記割り当て(16)は、割り当てベクトル( $(k)$ )を供給し、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分に対する追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかについての情報を表現する、請求項5~8のいずれか1項に記載の方法。

【請求項10】

請求項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、

- チャンネルの知覚復号されたフレーム

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

を得るために、現在の符号化圧縮されたフレーム

$$(\tilde{Y}(k-2))$$

を復号するステップ(31)と、

- 方向性信号のインデックスの前記データセット

$$(\tilde{J}_{DIR,ACT}(k))$$

と選択されたアンビエントH O A係数列のインデックスの前記データセット

$$(\tilde{J}_{AMB,ACT}(k-2))$$

を使用して、方向性信号

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

の前記対応するフレームと前記残差のアンビエントH O A成分

$$(\hat{C}_{AMB,RED}(k-2))$$

の前記対応するフレームとを再形成するために、チャンネルの前記知覚復号されたフレーム

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

を再配分するステップ(32)と、

- 検出された方向性信号のインデックスの前記データセット

$$(\tilde{J}_{DIR,ACT}(k))$$

および支配的な方向推定値の前記セット

$$(\tilde{G}_{\Omega,ACT}(k))$$

を使用して、方向性信号

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

の前記フレームと前記残差のアンビエントH O A成分

$$(\hat{C}_{AMB,RED}(k-2))$$

の前記フレームとからH O A表現の現在の圧縮解除されたフレーム

10

20

30

40

$$(\hat{\mathbf{C}}(k-3))$$

を再合成するステップ(33)と、を含み、

均一に分布した方向に対する方向性信号が前記方向性信号

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

から予測され、その後に、前記現在の圧縮解除されたフレーム

$$(\hat{\mathbf{C}}(k-3))$$

は、方向性信号

10

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

の前記フレーム、前記予測された信号、および前記残差のアンビエントH O A成分

$$(\hat{\mathbf{C}}_{\text{AMB,RED}}(k-2))$$

から再合成される、前記方法。

【請求項11】

請求項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、

- チャンネルの知覚復号されたフレーム

20

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

を得るために、現在の符号化圧縮されたフレーム

$$(\tilde{\mathbf{Y}}(k-2))$$

を復号するように構成された手段(31)と、

- 検出された方向性信号のインデックスの前記データセット

$$(\tilde{\mathbf{J}}_{\text{DIR,ACT}}(k))$$

と選択されたアンビエントH O A係数列のインデックスの前記データセット

30

$$(\tilde{\mathbf{J}}_{\text{AMB,ACT}}(k-2))$$

を使用して、方向性信号

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

の前記対応するフレームと前記残差のアンビエントH O A成分

$$(\hat{\mathbf{C}}_{\text{AMB,RED}}(k-2))$$

の前記対応するフレームとを再形成するために、チャンネルの前記知覚復号されたフレーム

40

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

を再配分するように構成された手段(32)と、

- 検出された方向性信号のインデックスの前記データセット

$$(\tilde{\mathbf{J}}_{\text{DIR,ACT}}(k))$$

および支配的な方向推定値の前記セット

$$(\tilde{\mathbf{G}}_{\Omega,\text{ACT}}(k))$$

を使用して、方向性信号

50

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

の前記フレームと前記残差のアンビエントH O A成分

$$(\hat{C}_{\text{AMB,RED}}(k-2))$$

の前記フレームとからH O A表現の現在の圧縮解除されたフレーム

$$(\hat{C}(k-3))$$

を再合成するように構成された手段(33)と、を備え、

均一に分布した方向に対する方向性信号が前記方向性信号

10

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

から予測され、その後、前記現在の圧縮解除されたフレーム

$$(\hat{C}(k-3))$$

は、方向性信号

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

の前記フレーム、前記予測された信号、および前記残差のアンビエントH O A成分

$$(\hat{C}_{\text{AMB,RED}}(k-2))$$

20

から再合成される、前記装置。

【請求項12】

均一に分布した方向に対する方向性信号の前記予測は、前記方向性信号

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

から、前記予測のための前記供給されたパラメータ((k-2))を使用して行われる

、請求項8に記載の方法。

【請求項13】

30

前記再配分(32)において、検出された方向性信号のインデックスの前記データセット

$$(\tilde{J}_{\text{DIR,ACT}}(k))$$

および選択されたアンビエントH O A係数列のインデックスの前記データセット

$$(\mathcal{J}_{\text{AMB,ACT}}(k-2))$$

の代わりに、受信した割り当てベクトル((k))が使用され、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分の追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかについての情報を表現する、請求項10または12に記載の方法。

40

【請求項14】

前記非所定数(M)の方向性信号( $X_{\text{DIR}}(k-2)$ )は、知覚に関連する基準に従って、

- 前記圧縮のための所定の所与の数のチャンネルを用いることにより、対応して圧縮解除されるH O A表現が最小の知覚可能な誤差をもたらし、前記基準で考慮される誤差は、

- 複数の異なる数の前記方向性信号( $X_{\text{DIR}}(k-2)$ )と前記残差のアンビエントH O A成分( $C_{\text{AMB,RED}}(k-2)$ )に対して複数の異なる数のH O A係数列とを使用することから生ずるモデル化誤差と、

- 前記方向性信号( $X_{\text{DIR}}(k-2)$ )の知覚符号化に伴う量子化雑音と、

50

- 前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) の個々のH O A係数列を符号化することに伴う量子化雑音と、であり、
- 前記3つの誤差の結果として生ずる合計誤差は、当該合計誤差の知覚可能性に関して、複数のテスト方向および複数の臨界帯域に対して考慮され、
- 前記最小の知覚可能な誤差を達成するために、平均の知覚可能な誤差または最大の知覚可能な誤差を最小化するように前記非所定数 ( $M$ ) の方向性信号 ( $X_{D I R}(k-2)$ ) が選択される、ように決定される、請求項2に記載の装置。

【請求項15】

前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) を表現する低減された数のH O A係数列の選択は、3つの場合を区別する基準に従って行われ、前記3つの場合は、

- 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数と同一であるときに、前記前のフレームと同一のH O A係数列が選択される場合と、
  - 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数よりも小さいときに、方向性信号によって占められている前記現在のフレーム内のチャンネルに割り当てられた前記前のフレームに存在した、前記前のフレームからのH O A係数列が非アクティブ化される場合と、
  - 前記現在のフレーム ( $k$ ) のH O A係数列の数がその前のフレーム ( $k-1$ ) のH O A係数列の数よりも大きいときに、前記前のフレームで選択されていたH O A係数列が前記現在のフレームでも選択され、追加的なH O A係数列が当該追加的なH O A係数列の知覚的な重要性または最も高い平均パワーに従って選択可能である場合と、
- である、請求項2または14に記載の装置。

【請求項16】

前記割り当て ( $16$ ) の実行で、

- アクティブな方向性信号は、前記知覚符号化 ( $17$ ) のための連続的な信号を取得するために、そのアクティブな方向性信号がチャンネル・インデックスを保持するように、所与のチャンネルに割り当てられ、
  - 前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) のH O A係数列は、最小の数 ( $R E D$ ) の当該H O A係数列が常に対応する数 ( $R E D$ ) の前のチャンネルが含まれるように、割り当てられ、
  - 前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) の追加的なH O A係数列を割り当てるために、該追加的なH O A係数列が前記前のフレーム ( $k-1$ ) でも選択されていたかが判定され、
  - 前記追加的なH O A係数列が前記前のフレーム ( $k-1$ ) でも選択されていた場合には、前記H O A係数列の知覚符号化 ( $17$ ) されるチャンネルに対する割り当て ( $16$ ) が前記前のフレームに対する割り当てと同一となり、
  - 前記追加的なH O A係数列が前記前のフレーム ( $k-1$ ) では選択されておらず、H O A係数列が新たに選択されている場合には、前記H O A係数列は、まずインデックスに関して昇順に配列され、当該昇順で、方向性信号によってまだ占められていない知覚符号化 ( $17$ ) されるチャンネルに割り当てられる、
- 請求項2、14および15のいずれか1項に記載の装置。

【請求項17】

$R E D$  は、前記残差のアンビエントH O A成分 ( $C_{A M B, R E D}(k-2)$ ) を表現するH O A係数列の数であり、前記割り当て ( $16$ ) を記述するパラメータは、前記残差のアンビエントH O A成分を表現する  $R E D$  個のH O A係数列に追加して使用される追加の数のH O A係数列に対応する長さを有するビットアレイに配列され、前記ビットアレイ内で各  $o$  番目のビットは、( $R E D + o$ ) 番目の追加のH O A係数列が前記残差のアンビエントH O A成分を表現するために使用されているかどうかを示す、請求項2および14～16のいずれか1項に記載の装置。

## 【請求項 18】

前記割り当て(16)を記述するパラメータは、非アクティブな方向性信号の数に対応する長さを有する割り当てベクトルに配列され、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分の追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかを示す、請求項2および14～16のいずれか1項に記載の装置。

## 【請求項 19】

前記現在のフレームのH O A係数列の分解(14)は、前記方向性信号( $X_{DIR}(k-2)$ )から元のH O A表現の部分を予測するために圧縮解除側で使用可能なパラメータ( $(k-2)$ )をさらに供給する、請求項2及び14～18のいずれか1項に記載の装置。

10

## 【請求項 20】

前記割り当て(16)は、割り当てベクトル( $(k)$ )を供給し、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分に対する追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかについての情報を表現する、請求項16～19のいずれか1項に記載の装置。

## 【請求項 21】

均一に分布した方向に対する方向性信号の前記予測は、前記方向性信号

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

20

から、前記予測のための前記供給されたパラメータ( $(k-2)$ )を使用して行われる、請求項19に記載の装置。

## 【請求項 22】

前記再配分(32)において、検出された方向性信号のインデックスの前記データセット

$$(\tilde{J}_{DIR,ACT}(k))$$

および選択されたアンビエントH O A係数列のインデックスの前記データセット

$$(J_{AMB,ACT}(k-2))$$

30

の代わりに、受信した割り当てベクトル( $(k)$ )が使用され、該割り当てベクトルの要素は、前記残差のアンビエントH O A成分の追加的なH O A係数列のうちのいずれが非アクティブな方向性信号を有するチャンネルに割り当てられるかについての情報を表現する、請求項11または21に記載の装置。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、方向性信号成分およびアンビエント信号成分を別々に処理することによって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。

40

## 【背景技術】

## 【0002】

高次アンビソニックス(H O A)は、波面合成法(W F S)や2.2.2のようなチャンネルに基づくアプローチといった他の技術が存在する一方で、三次元音声表現する1つの可能性を提供している。チャンネルに基づく方法と対照的に、H O A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H O A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W F Sのアプローチと比較して、H O A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。H O A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

50

要とすることなく同じ表現を利用できる点にある。

#### 【0003】

H O A は、切断球面調和関数 ( S H ) 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 H O A 音場表現は、実際には、“ ” 個の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は同等の意味を有するものとして以下の H O A 係数列または H O A チャンネルを参照する。

#### 【0004】

H O A 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 H O A 表現には、  $= 25$  の個数の H O A ( 展開 ) 係数が必要となる。上記の点を考慮して、H O A 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。したがって、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 H O A 表現を伝送すると、結果として、ビットレートは、  $19.2$  メガビット / 秒となるが、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。

#### 【0005】

H O A 音場表現の圧縮は、欧州特許出願第 1 2 3 0 6 5 6 9 号および欧州特許出願第 1 2 3 0 5 5 3 7 号において提案されている。例えば、E . H e l l e r u d、I . B u r n e t t、A . S o l v a n g および U . P . S v e n s s o n の「AACを用いた高次アンビソニックスの符号化」124回AESコンベンション、アムステルダム、2008年、において行われているような、H O A 係数列を個々に知覚符号化することの代わりに、特に音場分析を行い、所与の H O A 表現を方向性成分および残差アンビエント成分に分解することによって、知覚符号化される信号の数を減少させる試みが行われている。一般的には、方向性成分は、一般的な平面波関数とみなすことができる少数の支配的な方向性信号によって表現されるものとされる。残差のアンビエント H O A 成分の次数が低減される。その理由は、支配的な方向性信号を抽出した後は、より低次の H O A 係数が最も関連する情報を保持していると考えられるからである。

#### 【発明の概要】

#### 【0006】

総括すると、そのような処理を行うことによって、知覚符号化される H O A 係数列の初期数  $(N + 1)^2$  は、 $D$  個の支配的な方向性信号の所定数と、切断次数  $N_{RED} < N$  を用いて残差のアンビエント H O A 成分を表現する  $(N_{RED} + 1)^2$  個の H O A 係数列の数とに低減される。それによって、符号化される信号の数が決まり、すなわち、 $D + (N_{RED} + 1)^2$  となる。特に、この数は、時間フレーム  $k$  におけるアクティブな支配的な方向性音源の実際に検出された数  $D_{ACT}(k)$   $D$  とは独立している。これは、時間フレーム  $k$  において、アクティブな支配的な方向性音源の実際に検出された数  $D_{ACT}(k)$  が方向性信号の最大許容数  $D$  よりも小さい場合、知覚符号化される支配的な方向性信号のいくつかまたは全てさえもが零となることを意味している。つまり、これはこの複数のチャンネルが音場の関連情報を捕捉するために全く使用されないことを意味する。

#### 【0007】

この状況で、欧州特許出願第 1 2 3 0 6 5 6 9 号および欧州特許出願第 1 2 3 0 5 5 3 7 号における処理の別の想定される弱点は、各時間フレーム内の支配的な方向性信号の数を決定するための基準である。その理由は、音場の連続的な知覚符号化に関してアクティブな支配的な方向性信号の最適な数を決定する試みが行われていないからである。例えば、欧州特許出願第 1 2 3 0 5 5 3 7 号においては、支配的な音源の数が単純なパワー基準を使用して、すなわち、最大の固有値に属する係数間の相関行列の部分空間の次元を求め

ることによって推定される。欧州特許出願第 1 2 3 0 6 5 6 9 号においては、支配的な方向性音源のインクリメンタル検出が提案されている。ここで、各々の方向からの平面波関数のパワーが最初の方向性信号に対して十分に高い場合には、方向性音源が支配的であると考慮される。欧州特許出願第 1 2 3 0 6 5 6 9 号および欧州特許出願第 1 2 3 0 5 5 3 7 号の場合のようなパワーに基づく基準を使用すると、音場の知覚符号化に関して最適であるとは云えない方向性 - アンビエント分解となることもある。

#### 【 0 0 0 8 】

本発明によって解決される課題は、現在の H O A オーディオ信号コンテンツに対して、所定の低減された数のチャンネルに、方向性信号およびアンビエント H O A 成分に対する係数をどのように割り当てるかを決定することによって、H O A 圧縮を改善することにある。この課題は、請求項 1 および 3 に開示されたそれぞれの方法によって解決される。これらの方法を利用する装置は、請求項 2 および 4 において開示されている。

#### 【 0 0 0 9 】

本発明は、2つの態様において、欧州特許出願第 1 2 3 0 6 5 6 9 号で提案されている圧縮処理を改善する。第 1 に、知覚符号化される所与の数のチャンネルによってもたらされる帯域幅が良好に利用される。支配的な音源信号が検出されない時間フレームでは、支配的な方向性信号に対して当初より確保されているチャンネルは、アンビエント成分についての追加的な情報を捕捉するために、残差のアンビエント H O A 成分の追加的な H O A 係数列の形式で使用される。第 2 に、所与の H O A 音場表現を知覚符号化するために所与の数のチャンネルを利用するという目的を念頭に置くと、H O A 表現から抽出される方向性信号の数を決定するための基準は、その目的に対して適応化される。方向性信号の数は、復号され再構築された H O A 表現によって知覚される誤差が最も小さくなるように決定される。その基準は、方向性信号を抽出することと残差のアンビエント H O A 成分を記述するために H O A 係数列をより少なく使用することから生ずるモデル化誤差と、方向性信号を抽出することなく、その代わりに残差のアンビエント H O A 成分を記述するために追加的な H O A 係数列を使用することから生ずるモデル化誤差とを比較する。その基準は、さらに、その双方の場合に対して、方向性信号および残差のアンビエント H O A 成分の H O A 係数列の知覚符号化によってもたらされる量子化雑音の空間パワー分布を考慮する。

#### 【 0 0 1 0 】

上述した処理を実施するために、H O A 圧縮を開始する前に、信号（チャンネル）の合計数  $I$  が定められる。この合計数  $I$  は、当初の  $I$  個の H O A 係数列の数と比較して低減させられたものである。アンビエント H O A 成分は、最小の数  $R_{E D}$  個の H O A 係数列によって表現されるものと仮定される。場合によっては、その最小の数が零となることもある。残りの  $D = I - R_{E D}$  個のチャンネルは、方向性信号抽出処理が判定する知覚的に意味のよりあるものに依存して、方向性信号またはアンビエント H O A 成分の追加的な係数列のいずれかを含むものとされる。方向性信号またはアンビエント H O A 成分係数列のいずれかの残りの  $D$  個のチャンネルに対する割り当ては、フレーム単位で変更可能であるものと仮定される。受信機側での音場の再構築のために、この割り当てについての情報は、追加の副情報として送信される。

#### 【 0 0 1 1 】

原理的には、本発明の圧縮方法は、所定数の知覚符号化処理を使用して、H O A と称する音場の高次アンビソニクス表現を H O A 係数列の入力される時間フレームを用いて圧縮するのに適している。この方法は、フレーム単位で行われ、

- 現在のフレームに対して、支配的な方向のセットおよび対応する検出された方向性信号のインデックスのデータセットを推定するステップと、
- 上記現在のフレームの H O A 係数列を分解するステップであって、非所定数の方向性信号であって、支配的な方向推定値の上記セットに含まれる各々の方向と上記方向性信号のインデックスの各々のデータセットとを用いた、上記非所定数が上記所定数よりも小さい、上記非所定数の方向性信号と、上記所定数と上記非所定数との差に対応する低減された

数のH O A係数列によって表現される残差のアンビエントH O A成分と、対応する上記低減された数の残差のアンビエントH O A係数列のインデックスのデータセットと、に分解する、上記分解するステップと、

- 上記方向性信号および上記残差のアンビエントH O A成分のH O A係数列を上記所定数に対応する数のチャンネルに割り当てるステップであって、上記割り当てのために、上記方向性信号のインデックスの上記データセットおよび上記低減された数の残差のアンビエントH O A係数列のインデックスの上記データセットが使用される、上記割り当てるステップと、

- 関連するフレームの上記チャンネルを知覚符号化するステップであって、符号化された圧縮されたフレームが得られる、上記知覚符号化するステップと、を含む。

10

#### 【 0 0 1 2 】

原理的には、本発明の圧縮装置は、所定数の知覚符号化処理を使用して音場のH O Aと称する高次アンビソニックス表現をH O A係数列の入力される時間フレームを用いて圧縮するのに適している。

上記装置は、フレーム単位の処理を実行し、

- 現在のフレームに対して、支配的な方向のセットおよび対応する検出された方向性信号のインデックスのデータセットを推定するように構成された手段と、

- 上記現在のフレームのH O A係数列を分解するように構成された手段であって、非所定数の方向性信号であって、支配的な方向推定値の上記セットに含まれる各々の方向と、上記方向性信号のインデックスの各々のデータセットとを用いた、上記非所定数が上記所定数よりも小さい、上記非所定数の方向性信号と、上記所定数と上記非所定数との差に対応する低減された数のH O A係数列によって表現される残差のアンビエントH O A成分と、対応する上記低減された数の残差のアンビエントH O A係数列のインデックスの対応するデータセットと、に分解するように構成された、上記手段と、

20

- 上記方向性信号および上記残差のアンビエントH O A成分のH O A係数列を上記所定数に対応する数のチャンネルに割り当てるように構成された手段であって、上記割り当てのために、上記方向性信号のインデックスの上記データセットおよび上記低減された数の残差のアンビエントH O A係数列のインデックスの上記データセットが使用される、上記手段と、

- 関連するフレームの上記チャンネルを知覚符号化するように構成された手段であって、符号化された圧縮されたフレームが得られる、上記手段と、を含む。

30

#### 【 0 0 1 3 】

原理的には、本発明の圧縮解除方法は、上述の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を圧縮解除するのに適している。この圧縮解除方法は、

- チャンネルの知覚復号されたフレームを得るために、現在の符号化圧縮されたフレームを復号するステップと、

- 検出された方向性信号のインデックスの上記データセットと上記選択されたアンビエントH O A係数列のインデックスの上記データセットを使用して、方向性信号の上記対応するフレームと残差のアンビエントH O A成分の上記対応するフレームとを再形成するために、チャンネルの上記知覚復号されたフレームを再配分するステップと、

40

- 検出された方向性信号のインデックスの上記データセットおよび支配的な方向性推定値の上記セットを使用して、方向性信号の上記フレームと上記残差のアンビエントH O A成分の上記フレームとからH O A表現の現在の圧縮解除されたフレームを再合成するステップと、を含む、

均一に分布した方向に対する方向性信号が上記方向性信号から予測され、その後に、上記現在の圧縮解除されたフレームは、方向性信号の上記フレーム、上記予測された信号、および上記残差のアンビエントH O A成分から再合成される。

#### 【 0 0 1 4 】

原理的には、本発明の圧縮解除装置は、上述の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を圧縮解除するのに適している。この装置は、

50

- チャンネルの知覚復号されたフレームを得るために、現在の符号化圧縮されたフレームを復号するように構成された手段と、
- 検出された方向性信号のインデックスの上記データセットと選択されたアンビエントH O A係数列のインデックスの上記データセットを使用して、方向性信号の上記対応するフレームと上記残差のアンビエントH O A成分の上記対応するフレームとを再形成するために、チャンネルの上記知覚復号されたフレームを再配分するように構成された手段と、
- 検出された方向性信号のインデックスの上記データセットおよび支配的な方向性推定値の上記セットを使用して、方向性信号の上記フレームと上記残差のアンビエントH O A成分の上記フレームとから、上記H O A表現の現在の圧縮解除されたフレームを再合成するように構成された手段と、を含み、

10

均一に分布した方向に対する方向性信号が上記方向性信号から予測され、その後に、上記現在の圧縮解除されたフレームは、方向性信号の上記フレーム、上記予測された信号、および上記残差のアンビエントH O A成分から再合成される。

【0015】

本発明の追加的な実施形態は、各々の従属請求項に開示されており、有利なものである。

【図面の簡単な説明】

【0016】

【図1】H O A圧縮のブロック図である。

【図2】支配的な音源方向の推定のブロック図である。

20

【図3】H O A圧縮解除のブロック図である。

【図4】球面座標システムを示す図である。

【図5】複数の異なるアンビソニクス次数Nおよび角度  $[0, \quad ]$  に対する正規化された分散関数  $N(\quad)$  を示す図である。

【発明を実施するための形態】

【0017】

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。

A. 改良されたH O A圧縮

本発明に係る圧縮処理は、欧州特許出願第12306569号に基づいており、図1に示されている。ここで、信号処理ブロックは、欧州特許出願第12306569号に対して変更が加えられ、または新たに導入されており、その信号処理ブロックは太字のボックスで示されており、本出願における「

30

G

」(方向推定値とされたもの)および「C」は、それぞれ、欧州特許出願第12306569号の「A」(方向推定値の行列)および「D」に対応する。

【0018】

H O A圧縮のために、長さLのH O A係数列の重複しない入力フレームC(k)を用いた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係数列に関して定義される。

40

【数1】

$$C(k) := [c((kL+1)T_s) \quad c((kL+2)T_s) \quad c((k+1)LT_s)] , \quad (1)$$

ここで、 $T_s$  は、サンプリング期間を表す。

【0019】

図1のステップまたはステージ11/12は、任意に行われ、H O A係数列の重複しないk番目のフレームおよび(k-1)番目のフレームを下記の式に従って連結して長いフレーム

$\tilde{C}(k)$

50

にすることを含む。

【数 2】

$$\tilde{C}(k) := [C(k-1) \ C(k)] \ , \quad (2)$$

この長いフレームは、隣接する長いフレームと 50% 重複し、長いフレームは、支配的な音源方向の推定に連続的に使用される。

$\tilde{C}(k)$

の表記と同様に、チルダ記号は、以下の説明において、各々の量が長い重複するフレームを指すことを示すために使用される。ステップ / ステージ 11 / 12 が存在しない場合には、チルダ記号は特別な意味を持たない。

【0020】

原理的には、支配的な音源の推定ステップまたはステージ 13 は、欧州特許出願第 13305156 号に提案されているように行われるが、重要な変更を有する。この変更は、検出される方向の数の決定、すなわち、何個の方向性信号が H O A 表現から抽出されるかに関する。これは、アンビエント H O A 成分の良好な近似計算のために、追加的な H O A 係数列を使用することよりも方向性信号を抽出することの方が知覚的に関連性が高い場合にのみ、追加的な H O A 係数列を使用する代わりに方向性信号を抽出しようとする考えから成し遂げられるものである。A.2 の項目でこの技術についての詳細な説明を行う。

【0021】

支配的な音源の推定により、検出された方向性信号のインデックスのデータセット

$$\tilde{J}_{\text{DIR,ACT}}(k) \subseteq \{1, \dots, D\}$$

と、対応する方向推定値のセット

$$\tilde{G}_{\Omega, \text{ACT}}(k)$$

とが得られる。D は、H O A 圧縮を開始する前に設定しなければならない方向性信号の最大数を示している。

【0022】

ステップまたはステージ 14 において、H O A 係数列の現在の（長い）フレーム

$\tilde{C}(k)$

が、セット

$$\tilde{G}_{\Omega, \text{ACT}}(k)$$

内に含まれる方向に属する複数の方向性信号  $X_{\text{DIR}}(k-2)$  と、残差のアンビエント H O A 成分  $C_{\text{AMB}}(k-2)$  とに分解される（欧州特許出願第 13305156 号に提案されているように）。滑らかな信号を得るために、重畳加算処理の結果として 2 つのフレーム分の遅延が導入される。 $X_{\text{DIR}}(k-2)$  は、合計 D 個のチャンネルを含むものの、このうち、アクティブな方向性信号に対応するチャンネルのみが零でないと仮定される。このチャンネルを特定するインデックスは、データセット

$$J_{\text{DIR,ACT}}(k-2)$$

内において出力されるものと仮定される。さらに、ステップ / ステージ 14 における分解によって、方向性信号から元の H O A 表現の部分を予測するために圧縮解除側で使用されるいくつかのパラメータ  $(k-2)$  を供給する（より詳細には欧州特許出願第 13305156 号参照）。ステップまたはステージ 15 において、アンビエント H O A 成分  $C_{\text{AMB}}(k-2)$  の係数の数はインテリジェントに低減され、 $R_{\text{ED}} + D - N_{\text{DIR,ACT}}(k-2)$  個の非零の H O A 係数列のみを含むようになる。ここで、

$$N_{\text{DIR,ACT}}(k-2) = |J_{\text{DIR,ACT}}(k-2)|$$

は、データセット

$$J_{\text{DIR,ACT}}(k-2)$$

の組の数、すなわち、フレーム  $k-2$  内のアクティブな方向性信号の数を示す。アンビエント H O A 成分は、最小の数  $R_{\text{ED}}$  個の H O A 係数列によって常に表現されると仮定されるため、この問題は、実際には、想定される  $R_{\text{ED}}$  個の H O A 係数列から残りの  $D - N_{\text{DIR,ACT}}(k-2)$  個の H O A 係数列を選択することに集約される。滑らかな低減されたアンビエント H O A 表現を取得するために、この選択は、前のフレーム  $k-3$  で行った選択と比較して、変更が可能な限り少なくなるように行われる。

10

【0023】

特に、以下の3つの場合を区別すべきである。

【0024】

a)  $N_{\text{DIR,ACT}}(k-2) = N_{\text{DIR,ACT}}(k-3)$  : この場合、フレーム  $k-3$  の場合と同様に、同一の H O A 係数列が選択されるものと想定される。

【0025】

b)  $N_{\text{DIR,ACT}}(k-2) < N_{\text{DIR,ACT}}(k-3)$  : この場合、現在のフレーム内のアンビエント H O A 成分を表現するために、この前のフレーム  $k-3$  よりも多い H O A 係数列を使用することができる。  $k-3$  において選択済のその H O A 係数列は、現在のフレーム内でも選択されるものと仮定される。異なる基準に従って追加的な H O A 係数列を選択可能である。例えば、最高の平均パワーを有する H O A 係数列を  $C_{\text{AMB}}(k-2)$  内で選択するか、あるいは、それぞれの知覚的な重要性に関して H O A 係数列を選択する。

20

【0026】

c)  $N_{\text{DIR,ACT}}(k-2) > N_{\text{DIR,ACT}}(k-3)$  : この場合、現在のフレーム内のアンビエント H O A 成分を表現するために、最後のフレーム  $k-3$  に存在する H O A 係数列よりも少ない H O A 係数列を使用することができる。ここで解決すべき課題は、既に選択済の H O A 係数列のうち、どれを非アクティブ化しなければならないかである。合理的な解決法は、フレーム  $k-3$  で、信号を割り当てるステップまたはステージ 16

30

$$i \in J_{\text{DIR,ACT}}(k-2)$$

に割り当てられた H O A 係数列を非アクティブ化することである。

【0027】

追加的な H O A 係数列がアクティブ化または非アクティブ化されときのフレーム境界での不連続を回避するために、各々の信号を平滑的にフェード・インまたはフェード・アウトさせるとよい。

【0028】

$R_{\text{ED}} + N_{\text{DIR,ACT}}(k-2)$  個の低減された個数の最終的なアンビエント H O A 表現は、 $C_{\text{AMB}, R_{\text{ED}}}(k-2)$  によって示される。選択されたアンビエント係数列のインデックスは、データセット

40

$$J_{\text{AMB,ACT}}(k-2)$$

内に出力される。

【0029】

ステップ/ステージ 16 において、 $X_{\text{DIR}}(k-2)$  に含まれるアクティブな方向性信号および  $C_{\text{AMB}, R_{\text{ED}}}(k-2)$  に含まれる H O A 係数列は、個々の知覚符号化のために  $I$  個のチャンネルのフレーム  $Y(k-2)$  に割り当てられる。より詳細に信号の割り当てを記述すると、フレーム  $X_{\text{DIR}}(k-2)$ 、 $Y(k-2)$  および  $C_{\text{AMB}, R_{\text{ED}}}$

50

$\mathbf{d}(k-2)$  は、下記のように、個々の信号  $x_{\text{DIR},d}(k-2)$  ( $d \in \{1, \dots, D\}$ )、 $y_i(k-2)$  ( $i \in \{1, \dots, I\}$ ) および  $c_{\text{AMB,RED},o}(k-2)$  ( $o \in \{1, \dots, O\}$ ) によって構成されるものと仮定される。

【数 3】

$$\mathbf{X}_{\text{DIR}}(k-2) = \begin{bmatrix} x_{\text{DIR},1}(k-2) \\ x_{\text{DIR},2}(k-2) \\ \vdots \\ x_{\text{DIR},D}(k-2) \end{bmatrix}, \quad \mathbf{C}_{\text{AMB,RED}}(k-2) = \begin{bmatrix} c_{\text{AMB,RED},1}(k-2) \\ c_{\text{AMB,RED},2}(k-2) \\ \vdots \\ c_{\text{AMB,RED},O}(k-2) \end{bmatrix},$$

$$\mathbf{Y}(k-2) = \begin{bmatrix} y_1(k-2) \\ y_2(k-2) \\ \vdots \\ y_I(k-2) \end{bmatrix}. \quad (3)$$

10

【0030】

連続する知覚符号化のために連続した信号を取得するために、それぞれのチャンネルのインデックスを保持するようにアクティブな方向性信号が割り当てられる。これを下記の式のように表すことができる。

【数 4】

$$\text{全ての } d \in \mathcal{I}_{\text{DIR,ACT}}(k-2) \text{ に対して、 } y_d(k-2) = x_{\text{DIR},d}(k-2) \quad (4)$$

20

【0031】

アンビエント成分の H O A 係数列は、最小の数の  $O_{\text{RED}}$  個の係数列が  $\mathbf{Y}(k-2)$  の最後の  $O_{\text{RED}}$  個の信号に常に含まれるように、すなわち、下記の式に従って割り当てられる。

【数 5】

$$1 \leq o \leq O_{\text{RED}} \text{ に対して } y_{D+o}(k-2) = c_{\text{AMB,RED},o}(k-2) \quad (5)$$

【0032】

追加的な  $D - N_{\text{DIR,ACT}}(k-2)$  個のアンビエント成分の H O A 係数列については、これらが前のフレームでも選択されていたかどうかを区別すべきである。

30

a) 追加的な  $D - N_{\text{DIR,ACT}}(k-2)$  個のアンビエント成分の H O A 係数列が送信されるものとして前のフレーム内でも選択されていた場合、すなわち、各々のインデックスもまた、データセット

$$\mathcal{I}_{\text{AMB,ACT}}(k-3)$$

に含まれる場合には、これらの係数列の  $\mathbf{Y}(k-2)$  における信号への割り当ては、前のフレームに対する割り当てと同じである。この処理は、滑らかな信号  $y_i(k-2)$  を確保するものであり、ステップまたはステージ 17 における連続的な知覚符号化にとって好ましいものである。

40

b) そうではなく、いくつかの係数列が新たに選択されている場合、すなわち、これらのインデックスがデータセット

$$\mathcal{I}_{\text{AMB,ACT}}(k-2)$$

に含まれているが、データセット

$$\mathcal{I}_{\text{AMB,ACT}}(k-3)$$

に含まれていない場合には、これらはまず、インデックスに関して昇順に配列され、この順番で方向性信号によってまだ占められていない  $\mathbf{Y}(k-2)$  のチャンネル

$$i \notin \mathcal{I}_{\text{DIR,ACT}}(k-2)$$

に割り当てられる。

【 0 0 3 3 】

この特定の割り当ては、H O A 圧縮解除処理の間に信号の再配分および合成が、どのアンビエントH O A 係数列が  $Y(k-2)$  個のどのチャンネルに含まれているかについての情報無しに行えるようになるという利点を提供する。代わりに、データセット

$$\mathcal{I}_{\text{AMB,ACT}}(k-2)$$

および

$$\tilde{\mathcal{I}}_{\text{DIR,ACT}}(k)$$

の情報のみで、H O A 圧縮解除の間に割り当てを再構築することができる。

【 0 0 3 4 】

この割り当て処理によって、割り当てベクトル

$$\gamma(k) \in \mathbb{R}^{D-N_{\text{DIR,ACT}}(k-2)}$$

ももたらされることが有利である。この要素  $(k)$  ( $= 1, \dots, D - N_{\text{DIR,ACT}}(k-2)$ ) は追加的な  $D - N_{\text{DIR,ACT}}(k-2)$  個のアンビエント成分のH O A 係数列の各々のインデックスを表す。換言すれば、割り当てベクトル  $(k)$  の要素により、追加的な  $-R_{\text{E D}}$  個のアンビエントH O A 成分のH O A 係数列のうちいずれが  $D - N_{\text{DIR,ACT}}(k-2)$  個の非アクティブな方向性信号のチャンネルに割り当てられるかについての情報が得られる。このベクトルは、H O A 圧縮解除のために行われる再配分処理の初期化（項目 B 参照）を可能にするために、追加的に、フレームレートによる送信よりも低い頻度ではあるが送信されることがある。知覚符号化ステップ / ステージ 1 7 は、フレーム  $Y(k-2)$  の  $I$  個のチャンネルを符号化し、符号化されたフレーム

$$\tilde{Y}(k-2)$$

を出力する。

【 0 0 3 5 】

ステップ / ステージ 1 6 でベクトル  $(k)$  が送信されないフレームについては、圧縮解除側で、データ・パラメータ・セット

$$\mathcal{I}_{\text{DIR,ACT}}(k)$$

および

$$\mathcal{I}_{\text{AMB,ACT}}(k-2)$$

がベクトル  $(k)$  の代わりに再配分を行うために使用される。

【 0 0 3 6 】

A . 1 支配的な音源方向の推定

図 1 の支配的な音源方向に対する推定ステップ / ステージ 1 3 が図 2 により詳細に描かれている。これは、本質的に、欧州特許出願第 1 3 3 0 5 1 5 6 号に記載された内容に従って行われるが、決定的な違いがある。その決定的な違いは、支配的な音源の数を決定する手法である。支配的な音源の数は、所与のH O A 表現から抽出される方向性信号の数に対応する。この数は重要であり、その理由は、より多くの方向性信号を使用すること、あるいはその代わりに、より多くのH O A 係数列を使用してアンビエントH O A 成分をより良好にモデル化することのいずれかによって、所与のH O A 表現がより良好に表現されているかを制御するためにこの数が使用されるからである。

【 0 0 3 7 】

支配的な音源方向の推定は、入力される H O A 係数列の長いフレーム

$$\tilde{C}(k)$$

を使用して、支配的な音源方向の予備サーチで、ステップまたはステージ 2 1 において開始する。予備的な方向推定値

$$\tilde{\Omega}_{\text{DOM}}^{(d)}(k), \quad 1 \leq d \leq D,$$

と共に、個々の音源によって形成されるものとされる、予備的な方向推定値に対応する方向性信号

$$\tilde{\chi}_{\text{DOM}}^{(d)}(k)$$

10

および H O A 音場成分

$$\tilde{C}_{\text{DOM,CORR}}^{(d)}(k)$$

を欧州特許出願第 1 3 3 0 5 1 5 6 号に記載された内容に従って算出する。

【 0 0 3 8 】

ステップまたはステージ 2 2 において、予備的な方向推定値、方向性信号、および H O A 音場成分は、抽出される方向性信号の数

$$\tilde{D}(k)$$

20

を決定するために入力される H O A 係数列のフレーム

$$\tilde{C}(k)$$

と共に使用される。結果として、

$$\tilde{D}(k) < d \leq D$$

の方向性推定値

$$\tilde{\Omega}_{\text{DOM}}^{(d)}(k)$$

、これと対応する方向性信号

30

$$\tilde{\chi}_{\text{DOM}}^{(d)}(k)$$

、および H O A 音場成分

$$\tilde{C}_{\text{DOM,CORR}}^{(d)}(k)$$

が破棄される。代わりに、

$$1 \leq d \leq \tilde{D}(k)$$

の方向推定値

$$\tilde{\Omega}_{\text{DOM}}^{(d)}(k)$$

40

のみが、次に、既に見つかっている音源に対して割り当てられる。

【 0 0 3 9 】

ステップまたはステージ 2 3 において、結果として得られる方向軌跡は、音源動きモデルに従ってスムージング（滑らかに）され、音源のいずれがアクティブであるとされるかが決定される（欧州特許出願第 1 3 3 0 5 1 5 6 号参照）。この最後の処理により、アクティブな方向性音源のインデックスのセット

$$\tilde{J}_{\text{DIR,ACT}}(k)$$

とこれに対応する方向推定値のセット

50

$$\tilde{g}_{\Omega, \text{ACT}}(k)$$

とが得られる。

【 0 0 4 0 】

A . 2 抽出される方向性信号の数の決定

ステップ / ステージ 2 2 において方向性信号の数を決定するために、知覚的に最も関連する音場情報を捕捉するために利用される所与の合計数の I 個のチャンネルが存在する状況が想定される。したがって、全体としての H O A 圧縮 / 圧縮解除品質にとって、より多くの方向性信号を使用すること、あるいは、アンビエント H O A 成分のより良好なモデル化のためにより多くの H O A 係数列を使用することのいずれかによって、現在の H O A 表現がより良好に表現されるかというという課題を考慮して、抽出される方向性信号の数が決定される。抽出される方向性音源の数を決定するための基準をステップ / ステージ 2 2 において導出するために、どの基準が人間の知覚に関連しているか、H O A 圧縮が、特に、以下の 2 つの処理によって行われることが考慮される。

- アンビエント H O A 成分を表現するための H O A 係数列の低減（これは、関連するチャンネルの数の低減を意味する）

- 方向性信号およびアンビエント H O A 成分を表現するための H O A 係数列の知覚符号化

【 0 0 4 1 】

抽出された方向性信号の数  $M$  ( 0  $M$  D ) に依存して、1 番目の処理により、下記の式に従って近似計算が行われる。

【 数 6 】

$$\tilde{\mathcal{C}}(k) \approx \tilde{\mathcal{C}}^{(M)}(k) \quad (6)$$

【 数 7 】

$$:= \tilde{\mathcal{C}}_{\text{DIR}}^{(M)}(k) + \tilde{\mathcal{C}}_{\text{AMB,RED}}^{(M)}(k) , \quad (7)$$

ここで、

【 数 8 】

$$\tilde{\mathcal{C}}_{\text{DIR}}^{(M)}(k) := \sum_{d=1}^M \tilde{\mathcal{C}}_{\text{DOM,CORR}}^{(d)}(k) \quad (8)$$

は、 $M$  個の個々に考慮される音源によって形成されとする H O A 音場成分

$$\tilde{\mathcal{C}}_{\text{DOM,CORR}}^{(d)}(k), \quad 1 \leq d \leq M$$

から構成される方向性成分の H O A 表現を示し、

$$\tilde{\mathcal{C}}_{\text{AMB,RED}}^{(M)}(k)$$

は、 $I - M$  個の非零 H O A 係数列のみを用いたアンビエント成分の H O A 表現を示している。

【 0 0 4 2 】

2 番目の処理からの近似計算を下記の式によって表現することができる。

【 数 9 】

$$\tilde{\mathcal{C}}(k) \approx \hat{\mathcal{C}}^{(M)}(k) \quad (9)$$

【 数 1 0 】

$$:= \hat{\mathcal{C}}_{\text{DIR}}^{(M)}(k) + \hat{\mathcal{C}}_{\text{AMB,RED}}^{(M)}(k) \quad (10)$$

ここで、

$$\hat{\mathcal{C}}_{\text{DIR}}^{(M)}(k)$$

および

$$\hat{\tilde{C}}_{\text{AMB,RED}}^{(M)}(k)$$

は、それぞれ、知覚復号処理の後に合成された方向性成分およびアンビエントH O A成分を示している。

【 0 0 4 3 】

基準の形成

抽出される方向性信号の数

$$\tilde{D}(k)$$

は、合計近似誤差（ここで

10

$$M = \tilde{D}(k)$$

である）

【 数 1 1 】

$$\hat{\tilde{E}}^{(M)}(k) := \tilde{C}(k) - \hat{\tilde{C}}^{(M)}(k) \quad (11)$$

が人間の知覚の点で可能な限り顕著とならないように選択される。これを確実にするために、個々のバーク尺度臨界帯域に対する合計誤差の方向性パワー分布は、所定の数Q個のテスト方向  $q$ （ $q = 1, \dots, Q$ ）で考慮される。このテスト方向は、単位球面上でほぼ均一に分布する。より具体的に述べると、 $b$ 番目の臨界帯域（ $b = 1, \dots, B$ ）に

20

対する方向性パワー分布は、下記のベクトルによって表現される。

【 数 1 2 】

$$\hat{\tilde{\mathcal{P}}}^{(M)}(k, b) := \left[ \hat{\tilde{\mathcal{P}}}_1^{(M)}(k, b) \quad \hat{\tilde{\mathcal{P}}}_2^{(M)}(k, b) \quad \dots \quad \hat{\tilde{\mathcal{P}}}_Q^{(M)}(k, b) \right]^T, \quad (12)$$

ベクトルの成分

$$\hat{\tilde{\mathcal{P}}}_q^{(M)}(k, b)$$

は、方向  $q$ 、 $b$ 番目のバーク尺度臨界帯域、および  $k$ 番目のフレームに関連する合計誤差

$$\hat{\tilde{E}}^{(M)}(k)$$

30

のパワーを示す。合計誤差

$$\hat{\tilde{E}}^{(M)}(k)$$

の方向性パワー分布

$$\hat{\tilde{\mathcal{P}}}^{(M)}(k, b)$$

は、元のH O A表現

$$\tilde{C}(k)$$

40

による下記の方向性知覚マスキングパワー分布と比較される。

【 数 1 3 】

$$\tilde{\mathcal{P}}_{\text{MASK}}(k, b) := \left[ \tilde{\mathcal{P}}_{\text{MASK},1}(k, b) \quad \tilde{\mathcal{P}}_{\text{MASK},2}(k, b) \quad \dots \quad \tilde{\mathcal{P}}_{\text{MASK},Q}(k, b) \right]^T \quad (13)$$

次に、各テスト方向  $q$ および臨界帯域  $b$ に対して、合計誤差の知覚レベル

$$\hat{\tilde{\mathcal{L}}}_q^{(M)}(k, b)$$

が算出される。知覚レベルは、ここで、本質的に、合計誤差

$$\hat{\tilde{E}}^{(M)}(k)$$

50

の方向性パワーと方向性マスキングパワーとの比率として下記の式に従って定義される。

【数 1 4】

$$\tilde{\mathcal{L}}_q^{(M)}(k, b) := \max \left( 0, \frac{\hat{\mathcal{P}}_q^{(M)}(k, b)}{\tilde{\mathcal{P}}_{\text{MASK}, q}(k, b)} - 1 \right) . \quad (14)$$

【 0 0 4 4 】

「 1 」を減算し、連続的な最大値を求める処理が行われ、誤差パワーがマスキング閾値未満である限り確実に知覚レベルが零になるようにする。最終的に、抽出される方向性信号の数

$$\tilde{D}(k)$$

10

は、全ての臨界帯域に亘る誤差知覚レベルの最大値の全てのテスト方向に対する平均値が最小になるように、すなわち、下記の式に従って選択される。

【数 1 5】

$$\tilde{D}(k) = \underset{M}{\operatorname{argmin}} \frac{1}{Q} \sum_{q=1}^Q \max_b \tilde{\mathcal{L}}_q^{(M)}(k, b) . \quad (15)$$

【 0 0 4 5 】

なお、代替的には、式 ( 1 5 ) において誤差知覚レベルの最大値を平均化処理によって置き換えることができる。

【 0 0 4 6 】

方向性知覚マスキングパワー分布の算出

元の H O A 表現

$$\tilde{\mathcal{C}}(k)$$

による方向性知覚マスキングパワー分布

$$\tilde{\mathcal{P}}_{\text{MASK}}(k, b)$$

の算出のために、元の H O A 表現

$$\tilde{\mathcal{C}}(k)$$

30

は、テスト方向  $q$  (  $q = 1, \dots, Q$  ) から到来する一般的な平面波

$$\tilde{v}_q(k)$$

によって表現されるようにするために、空間領域に変換される。行列

$$\tilde{V}(k)$$

内の一般的な平面波信号

$$\tilde{v}_q(k)$$

を

40

【数 1 6】

$$\tilde{V}(k) = \begin{bmatrix} \tilde{v}_1(k) \\ \tilde{v}_2(k) \\ \vdots \\ \tilde{v}_Q(k) \end{bmatrix} , \quad (16)$$

のように配列すると、空間領域への変換は、下記の処理によって表現される。

【数 1 7】

$$\tilde{V}(k) = \mathcal{E}^T \tilde{\mathcal{C}}(k) , \quad (17)$$

50

ここで、 $\mathbf{S}_q$  は、テスト方向  $q$  ( $q = 1, \dots, Q$ ) に対して以下の式によって定義されるモード行列を示す。

【数 18】

$$\mathbf{S} := [\mathbf{S}_1 \ \mathbf{S}_2 \ \dots \ \mathbf{S}_Q] \in \mathbb{R}^{O \times Q} \quad (18)$$

ここで、 $\mathbf{S}_q :=$

【数 19】

$$[S_0^0(\boldsymbol{\Omega}_q) \ S_{-1}^{-1}(\boldsymbol{\Omega}_q) \ S_{-1}^0(\boldsymbol{\Omega}_q) \ S_{-1}^1(\boldsymbol{\Omega}_q) \ S_{-2}^{-2}(\boldsymbol{\Omega}_q) \ \dots \ S_N^N(\boldsymbol{\Omega}_q)]^T \in \mathbb{R}^O. \quad (19)$$

元の H O A 表現

10

$$\tilde{\mathbf{C}}(k)$$

による、方向性知覚マスキングパワー分布

$$\tilde{\mathcal{P}}_{\text{MASK}}(k, b)$$

の要素

$$\tilde{\mathcal{P}}_{\text{MASK}}(k, b)$$

は、個々の臨界帯域  $b$  に対する一般的な平面波関数

$$\tilde{v}_q(k)$$

20

のマスキングパワーに対応する。

【0047】

方向性パワー分布の算出

以下の説明において、方向性パワー分布

$$\hat{\mathcal{P}}^{(M)}(k, b)$$

を算出するための以下の2つの代替策が示される。

【0048】

a. 1つの可能性は、項目 A. 2の最初に記載されている2つの処理を実行することによって、所望の H O A 表現

30

$$\tilde{\mathbf{C}}(k)$$

の近似値

$$\hat{\tilde{\mathbf{C}}}^{(M)}(k)$$

を実際に算出することである。次に、合計近似誤差

$$\hat{\tilde{\mathbf{E}}}^{(M)}(k)$$

が式(11)に従って算出される。次に、合計近似誤差

40

$$\hat{\tilde{\mathbf{E}}}^{(M)}(k)$$

が、テスト方向  $q$  ( $q = 1, \dots, Q$ ) から到来する一般的な平面波

$$\hat{\tilde{\mathbf{W}}}^{(M)}_q(k)$$

によって表現されるために、空間領域に変換される。一般的な平面波信号を以下のように表される行列

$$\hat{\tilde{\mathbf{W}}}^{(M)}(k)$$

内に配置すると、

50

【数 2 0】

$$\widehat{\mathbf{W}}^{(M)}(k) = \begin{bmatrix} \widehat{\mathbf{w}}_1^{(M)}(k) \\ \widehat{\mathbf{w}}_2^{(M)}(k) \\ \vdots \\ \widehat{\mathbf{w}}_Q^{(M)}(k) \end{bmatrix}, \quad (20)$$

空間領域への変換は、下記の処理によって表現される。

【数 2 1】

$$\widehat{\mathbf{W}}^{(M)}(k) = \mathbf{E}^T \widehat{\mathbf{E}}^{(M)}(k) \quad (21)$$

10

合計近似誤差

$$\widehat{\mathbf{E}}^{(M)}(k)$$

の方向性パワー分布

$$\widehat{\mathcal{P}}^{(M)}(k, b)$$

の要素

$$\widehat{\mathcal{P}}_q^{(M)}(k, b)$$

20

は、個々の臨界帯域  $b$  内で一般的な平面波関数

$$\widehat{\mathbf{w}}_q^{(M)}(k), \quad q = 1, \dots, Q,$$

のパワーを算出することによって取得される。

【0 0 4 9】

$b$  . 代替的な解決法は、

$$\widehat{\mathcal{C}}^{(M)}(k)$$

の代わりに近似値

$$\widetilde{\mathcal{C}}^{(M)}(k)$$

30

のみを算出することである。この方法には、個々の信号の複雑な知覚符号化を直接行う必要がないという利点がある。この代わりに、個々のパーク尺度臨界帯域内の知覚量子化誤差のパワーを知ることによって十分である。この目的のため、式 ( 1 1 ) に定義された合計近似誤差を、以下の 3 つの近似誤差の合計として記述することができる。

【数 2 2】

$$\widetilde{\mathbf{E}}^{(M)}(k) := \widetilde{\mathcal{C}}(k) - \widetilde{\mathcal{C}}^{(M)}(k) \quad (22)$$

【数 2 3】

$$\widehat{\mathbf{E}}_{\text{DIR}}^{(M)}(k) := \widetilde{\mathcal{C}}_{\text{DIR}}^{(M)}(k) - \widehat{\mathcal{C}}_{\text{DIR}}^{(M)}(k) \quad (23)$$

40

【数 2 4】

$$\widehat{\mathbf{E}}_{\text{AMB,RED}}^{(M)}(k) := \widetilde{\mathcal{C}}_{\text{AMB,RED}}^{(M)}(k) - \widehat{\mathcal{C}}_{\text{AMB,RED}}^{(M)}(k), \quad (24)$$

この 3 つの近似誤差は、互いに独立しているものと仮定することができる。この独立性のため、合計誤差

$$\widehat{\mathbf{E}}^{(M)}(k)$$

の方向性パワー分布は、3 つの個々の誤差

$$\tilde{E}^{(M)}(k)$$

、

$$\hat{\tilde{E}}_{\text{DIR}}^{(M)}(k)$$

、および

$$\hat{\tilde{E}}_{\text{AMB,RED}}^{(M)}(k)$$

の方向性パワー分布の合計として表現することができる。

【 0 0 5 0 】

10

以下、個々のバーク尺度臨界帯域に対する3つの誤差の方向性パワー分布をどのように算出するかについて記載する。

【 0 0 5 1 】

a . 誤差

$$\tilde{E}^{(M)}(k)$$

の方向性パワー分布を算出するために、まず、下記の式によって、空間領域への変換が行われる。

【 数 2 5 】

$$\tilde{W}^{(M)}(k) = \mathcal{E}^T \tilde{E}^{(M)}(k) , \quad (25)$$

20

ここで、近似誤差

$$\tilde{E}^{(M)}(k)$$

は、したがって、テスト方向  $q$  (  $q = 1, \dots, Q$  ) から到来する一般的な平面波

$$\tilde{w}_q^{(M)}(k)$$

によって表現され、これは、下記の式に従って、行列

$$\tilde{W}^{(M)}(k)$$

30

内に配列される。

【 数 2 6 】

$$\tilde{W}^{(M)}(k) = \begin{bmatrix} \tilde{w}_1^{(M)}(k) \\ \tilde{w}_2^{(M)}(k) \\ \vdots \\ \tilde{w}_Q^{(M)}(k) \end{bmatrix} . \quad (26)$$

結果として、近似誤差

$$\tilde{E}^{(M)}(k)$$

40

の方向性パワー分布

$$\tilde{\mathcal{P}}^{(M)}(k, b)$$

の要素

$$\tilde{\mathcal{P}}_q^{(M)}(k, b)$$

は、個々の臨界帯域  $b$  内で、一般的な平面波関数

$$\tilde{w}_q^{(M)}(k), \quad q = 1, \dots, Q,$$

50

のパワーを算出することによって取得される。

【 0 0 5 2 】

b . 誤差

$$\hat{E}_{\text{DIR}}^{(M)}(k)$$

の方向性パワー分布

$$\hat{\mathcal{P}}_{\text{DIR}}^{(M)}(k, b)$$

を算出するために、方向性信号

$$\tilde{\mathbf{x}}_{\text{DOM}}^{(d)}(k), \quad 1 \leq d \leq M$$

10

を知覚符号化することによって、この誤差が方向性 H O A 成分

$$\tilde{\mathcal{C}}_{\text{DIR}}^{(M)}(k)$$

に導入されることに留意すべきである。さらに、方向性 H O A 成分が式 ( 8 ) によって与えられることを考慮すべきである。そして、簡略化のために、H O A 成分

$$\tilde{\mathcal{C}}_{\text{DOM,CORR}}^{(d)}(k)$$

が、空間領域内で、個の一般的な平面波関数

$$\tilde{\mathbf{v}}_{\text{GRID},o}^{(d)}(k)$$

20

によって、等価的に表現されるものと仮定する。これは、単なるスケーリングによって、すなわち、下記の式に従って方向性信号

$$\tilde{\mathbf{x}}_{\text{DOM}}^{(d)}(k)$$

から形成される。

【 数 2 7 】

$$\tilde{\mathbf{v}}_{\text{GRID},o}^{(d)}(k) = \alpha_o^{(d)}(k) \tilde{\mathbf{x}}_{\text{DOM}}^{(d)}(k), \quad (27)$$

ここで、

30

$$\alpha_o^{(d)}(k), \quad o = 1, \dots, O,$$

は、スケーリング・パラメータを示している。各々の平面波方向

$$\tilde{\mathbf{\Omega}}_{\text{ROT},o}^{(d)}(k), \quad o = 1, \dots, O$$

は、単位球面上で均一に分布し、

$$\tilde{\mathbf{\Omega}}_{\text{ROT},1}^{(d)}(k)$$

が方向推定値

$$\tilde{\mathbf{\Omega}}_{\text{DOM}}^{(d)}(k)$$

40

と対応するように、回転されるものと仮定される。したがって、スケーリング・パラメータ

$$\alpha_1^{(d)}(k)$$

は「 1 」である。

【 0 0 5 3 】

回転された方向

$$\tilde{\mathbf{\Omega}}_{\text{ROT},o}^{(d)}(k), \quad o = 1, \dots, O$$

50

に対して

$$\mathbf{\Xi}_{\text{GRID}}^{(d)}(k)$$

をモード行列として定義し、

【数 2 8】

$$\boldsymbol{\alpha}^{(d)}(k) := [1 \quad \alpha_2^{(d)}(k) \quad \alpha_3^{(d)}(k) \quad \dots \quad \alpha_o^{(d)}(k)]^T \in \mathbb{R}^o, \quad (28)$$

に従ってベクトル内の全てのスケーリング・パラメータ

$$\alpha_o^{(d)}(k)$$

10

を配列すると、H O A 成分

$$\tilde{\mathbf{C}}_{\text{DOM,CORR}}^{(d)}(k)$$

を下記の式のように記述することができる。

【数 2 9】

$$\tilde{\mathbf{C}}_{\text{DOM,CORR}}^{(d)}(k) = \mathbf{\Xi}_{\text{GRID}}^{(d)}(k) \boldsymbol{\alpha}^{(d)}(k) \tilde{\mathbf{x}}_{\text{DOM}}^{(d)}(k). \quad (29)$$

【0 0 5 4】

結果として、真の方向性 H O A 成分

【数 3 0】

$$\tilde{\mathbf{C}}_{\text{DIR}}^{(M)}(k) = \sum_{d=1}^M \tilde{\mathbf{C}}_{\text{DOM,CORR}}^{(d)}(k) \quad (30)$$

20

と、

【数 3 1】

$$\hat{\tilde{\mathbf{C}}}_{\text{DIR}}^{(M)}(k) = \sum_{d=1}^M \hat{\tilde{\mathbf{C}}}_{\text{DOM,CORR}}^{(d)}(k) \quad (31)$$

【数 3 2】

$$:= \sum_{d=1}^M \mathbf{\Xi}_{\text{GRID}}^{(d)}(k) \boldsymbol{\alpha}^{(d)}(k) \hat{\tilde{\mathbf{x}}}_{\text{DOM}}^{(d)}(k) \quad (32)$$

によって知覚復号された方向性信号

$$\hat{\tilde{\mathbf{x}}}_{\text{DOM}}^{(d)}(k)$$

30

( d = 1 , ... , M ) が合成されたものとの間の誤差

$$\hat{\tilde{\mathbf{E}}}_{\text{DIR}}^{(M)}(k)$$

( 式 ( 2 3 ) 参照 ) は、下記の式で表される知覚符号化誤差

【数 3 3】

$$\hat{\tilde{\mathbf{e}}}_{\text{DOM}}^{(d)}(k) := \tilde{\mathbf{x}}_{\text{DOM}}^{(d)}(k) - \hat{\tilde{\mathbf{x}}}_{\text{DOM}}^{(d)}(k) \quad (33)$$

の点で個々の方向性信号において下記の式によって表現することができる。

40

【数 3 4】

$$\hat{\tilde{\mathbf{E}}}_{\text{DIR}}^{(M)}(k) = \sum_{d=1}^M \mathbf{\Xi}_{\text{GRID}}^{(d)}(k) \boldsymbol{\alpha}^{(d)}(k) \hat{\tilde{\mathbf{e}}}_{\text{DOM}}^{(d)}(k). \quad (34)$$

【0 0 5 5】

テスト方向  $q$  ( q = 1 , ... , Q ) に対して、空間領域内の誤差

$$\hat{\tilde{\mathbf{E}}}_{\text{DIR}}^{(M)}(k)$$

の表現は、下記の式によって与えられる。

【数 3 5】

$$\widehat{\mathbf{W}}_{\text{DIR},q}^{(M)}(d) = \sum_{d=1}^M \underbrace{\mathbf{E}^T \mathbf{E}_{\text{GRID}}^{(d)}(k) \boldsymbol{\alpha}^{(d)}(k)}_{=: \beta^{(d)}(k)} \widehat{\mathbf{e}}_{\text{DOM}}^{(d)}(k) . \quad (35)$$

【0 0 5 6】

ベクトルの要素  $(d)(k)$  を

$$\beta_q^{(d)}(k), \quad q = 1, \dots, Q$$

と表し、個々の知覚符号化誤差

$$\widehat{\mathbf{e}}_{\text{DOM}}^{(d)}(k), \quad d = 1, \dots, M,$$

10

が互いに独立しているものと仮定することにより、式(35)から、知覚符号化誤差

$$\widehat{\mathbf{E}}_{\text{DIR}}^{(M)}(k)$$

の方向性パワー分布

$$\widehat{\mathcal{P}}_{\text{DIR}}^{(M)}(k, b)$$

の要素

$$\widehat{\mathcal{P}}_{\text{DIR},q}^{(M)}(k, b)$$

20

は、下記の式によって算出することができる。

【数 3 6】

$$\widehat{\mathcal{P}}_{\text{DIR},q}^{(M)}(k, b) = \sum_{d=1}^M \left( \beta_q^{(d)}(k) \right)^2 \widetilde{\sigma}_{\text{DIR},d}^2(k, b) . \quad (36)$$

$$\widetilde{\sigma}_{\text{DIR},d}^2(k, b)$$

は、方向性信号

$$\widehat{\mathbf{x}}_{\text{DOM}}^{(d)}(k)$$

におけるb番目の臨界帯域内の知覚量子化誤差のパワーを表現するように想定されている。  
このパワーは、方向性信号

30

$$\widehat{\mathbf{x}}_{\text{DOM}}^{(d)}(k)$$

の知覚マスキングパワーに対応するものとして算出することができる。

【0 0 5 7】

c. アンビエントH O A成分のH O A係数列の知覚符号化の結果として得られる誤差

$$\widehat{\mathbf{E}}_{\text{AMB,RED}}^{(M)}(k)$$

の方向性パワー分布

40

$$\widehat{\mathcal{P}}_{\text{AMB,RED}}^{(M)}(k, b)$$

を算出するために、各H O A係数列が独立して符号化されるものとする。したがって、各  
バーク尺度臨界帯域内の個々のH O A係数列内に導入される誤差は、相関性がないとす  
ることができる。これは、誤差

$$\widehat{\mathbf{E}}_{\text{AMB,RED}}^{(M)}(k)$$

の係数間相関行列は、各バーク尺度臨界帯域に対して対角である、すなわち、下記の式で  
表される。

## 【数 3 7】

$$\begin{aligned} \tilde{\Sigma}_{\text{AMB,RED}}^{(M)}(k, b) = \\ \text{diag}\left(\tilde{\sigma}_{\text{AMB,RED},1}^{2(M)}(k, b), \tilde{\sigma}_{\text{AMB,RED},2}^{2(M)}(k, b), \dots, \tilde{\sigma}_{\text{AMB,RED},O}^{2(M)}(k, b)\right) \end{aligned} \quad (37)$$

要素

$$\tilde{\sigma}_{\text{AMB,RED},o}^{2(M)}(k, b), \quad o = 1, \dots, O,$$

は、

$$\hat{\tilde{C}}_{\text{AMB,RED}}^{(M)}(k) \quad 10$$

内の  $o$  番目の符号化された H O A 係数列における  $b$  番目の臨界帯域内の知覚量子化誤差のパワーを表現するものとする。これは、 $o$  番目の H O A 係数列

$$\tilde{C}_{\text{AMB,RED}}^{(M)}(k)$$

の知覚マスキングパワーに対応するものと仮定することができる。したがって、知覚符号化誤差

$$\hat{\tilde{E}}_{\text{AMB,RED}}^{(M)}(k)$$

の方向性パワー分布は、下記の式によって算出される。

20

## 【数 3 8】

$$\hat{\tilde{P}}_{\text{AMB,RED}}^{(M)}(k, b) = \text{diag}\left(\mathbf{E}^T \tilde{\Sigma}_{\text{AMB,RED}}^{(M)}(k, b) \mathbf{E}\right). \quad (38)$$

## 【0 0 5 8】

B . 改良された H O A 圧縮解除

対応する H O A 圧縮解除処理が図 3 に示されており、この H O A 圧縮解除処理は、以下のステップまたはステージを含む。

## 【0 0 5 9】

ステップまたはステージ 3 1 において、

$$\hat{\mathbf{Y}}(k-2) \quad 30$$

内の復号された信号を取得するために、

$$\tilde{\mathbf{Y}}(k-2)$$

内に含まれる  $I$  個の信号の知覚復号処理が行われる。

## 【0 0 6 0】

信号再配分ステージまたはステージ 3 2 において、

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

内の知覚復号された信号は、方向性信号のフレーム

40

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

およびアンビエント H O A 成分のフレーム

$$\hat{\tilde{C}}_{\text{AMB,RED}}(k-2)$$

を再形成するために再配分される。インデックスのデータセット

$$\tilde{j}_{\text{DIR,ACT}}(k)$$

および

$$J_{\text{AMB,ACT}}(k-2)$$

を使用して、H O A 圧縮に対して行われる割り当て処理を再現することによって、どのように信号を再配分するかについての情報が取得される。これは、再帰的な処理であるため（項目 A 参照）、例えば、送信に不具合が発生しているような場合に再配分処理を初期化できるようにするために、追加的に送信される割り当てベクトル（ $k$ ）を使用することができる。

【 0 0 6 1 】

合成ステップまたはステージ 3 3 において、（欧州特許出願第 1 2 3 0 6 5 6 9 号の図 2 b および図 4 に関連して記載されている処理に従って、）方向性信号のフレーム

10

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

、対応する方向のセット

$$\tilde{G}_{\Omega,\text{ACT}}(k)$$

と共にアクティブな方向性信号のインデックスのセット

$$\tilde{J}_{\text{DIR,ACT}}(k)$$

、方向性信号からの H O A 表現の部分を予測するためのパラメータ（ $k-2$ ）、および低減されたアンビエント H O A 成分の H O A 係数列のフレーム

20

$$\hat{C}_{\text{AMB,RED}}(k-2)$$

を使用して、所望の合計 H O A 表現の現在のフレーム

$$\hat{C}(k-3)$$

が再合成される。

$$\hat{C}_{\text{AMB,RED}}(k-2)$$

は、欧州特許出願第 1 2 3 0 6 5 6 9 号における

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

30

に対応し、

$$\tilde{G}_{\Omega,\text{ACT}}(k)$$

および

$$\tilde{J}_{\text{DIR,ACT}}(k)$$

は、欧州特許出願第 1 2 3 0 6 5 6 9 号における

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

40

に対応する。ここでアクティブな方向性信号のインデックスは、

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

の行列要素においてマーク付けされる。すなわち、均一に分布する方向に対する方向性信号は、予測のための受信済のパラメータ（（ $k-2$ ））を使用して方向性信号

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

から予測される。その後、現在の圧縮解除されたフレーム

$$(\hat{C}(k-3))$$

50

が、方向性信号

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

のフレーム、予測された部分および低減されたアンビエント H O A 成分

$$(\hat{C}_{\text{AMB,RED}}(k-2))$$

から再合成される。

【 0 0 6 2 】

C . 高次アンビソニックスの基礎

高次アンビソニックス ( H O A ) 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  $x$  での音圧  $p(t, x)$  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 4 に示された球面座標システムに基づいている。使用されている座標システムにおいて、 $x$  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、座標原点への距離)、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。さらに、 $(\cdot)^T$  は、転置を表す。

【 0 0 6 3 】

$F_t(\cdot)$  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 9 】

$$P(\omega, x) = \mathcal{F}_t(p(t, x)) = \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt, \quad (39)$$

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 o u r i e r A c o u s t i c s ( フーリエ・アコースティックス ) ”、応用数理学、第 9 3 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 4 0 】

$$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi). \quad (40)$$

式 ( 4 0 ) において、 $c_s$  は音速を示し、 $k$  は角波数を示し、この角波数  $k$  は  $k = \omega / c_s$  によって角周波数に関連している。さらに、 $j_n(\cdot)$  は、第 1 種球ベッセル関数を表しており、

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

は、C . 1 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。上述した内容において、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、球面調和関数の級数が次数インデックス  $n$  に対して上限  $N$  で打ち切れ、これは、H O A 表現の次数と呼ばれる。

【 0 0 6 4 】

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  $C(\theta, \phi)$  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 . R a f a e l y 著、 “ P l a n e - w a v e D e c o m p o s i t i o n o f t h e S o u n d F i e l d o n a S p h e r e b y S p h e r i c a l C o n v o l u t i o n ( 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 ( 1 1 6 )、2 1 4 9 - 2 1 5 7 頁、2 0 0 4 年参照)。

## 【数 4 1】

$$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi) , \quad (41)$$

ここで、展開係数

$$C_n^m(k)$$

は、展開係数

$$A_n^m(k)$$

と下記の式によって関連する。

10

## 【数 4 2】

$$A_n^m(k) = 4\pi i^n C_n^m(k) . \quad (42)$$

## 【 0 0 6 5 】

個々の係数

$$C_n^m(\omega = kc_s)$$

が角周波数 の関数であると仮定すると、逆フーリエ変換 (

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

20

) によって示される ) を適用することにより、下記の時間領域関数をもたらす。

## 【数 4 3】

$$c_n^m(t) = \mathcal{F}_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega \quad (43)$$

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

## 【数 4 4】

$$c(t) = \quad (44)$$

$$\left[ c_0^0(t) \ c_1^{-1}(t) \ c_1^0(t) \ c_1^1(t) \ c_2^{-2}(t) \ c_2^{-1}(t) \ c_2^0(t) \ c_2^1(t) \ c_2^2(t) \ \dots \ c_N^{N-1}(t) \ c_N^N(t) \right]^T . \quad 30$$

ベクトル  $c(t)$  内の時間領域関数

$$c_n^m(t)$$

の位置インデックスは、 $n(n+1) + 1 + m$  によって与えられる。ベクトル  $c(t)$  内の要素の総計は、 $(N+1)^2$  によって与えられる。

## 【 0 0 6 6 】

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  $c(t)$  のサンプリングされたバージョンをもたらす。

## 【数 4 5】

40

$$\{c(lT_s)\}_{l \in \mathbb{N}} = \{c(T_s), c(2T_s), c(3T_s), c(4T_s), \dots\} \quad (45)$$

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

$$c_n^m(t)$$

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 7 】

C . 1 実数値の球面調和関数の定義

実数値の球面調和関数

50

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

は、下記の式によって与えられる。

【数 4 6】

$$S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_{n,|m|}(\cos\theta) \operatorname{trg}_m(\phi) \quad (46)$$

ここで

【数 4 7】

$$\operatorname{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases} \quad (47)$$

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 8】

$$P_{n,m}(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x), \quad m \geq 0 \quad (48)$$

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 著の文献の場合とは異なり、コンドン-ショートレーの位相項  $(-1)^m$  を用いない。

【0068】

C. 2 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  $x(t)$  は、下記の式によって HOA において表現される。

【数 4 9】

$$c_n^m(t) = x(t) S_n^m(\boldsymbol{\Omega}_0), \quad 0 \leq n \leq N, |m| \leq n. \quad (49)$$

平面波振幅の対応する空間密度

$$c(t, \boldsymbol{\Omega}) := \mathcal{F}_t^{-1}(C(\omega, \boldsymbol{\Omega})) \quad (50)$$

は、下記の式によって与えられる。

【数 5 0】

$$c(t, \boldsymbol{\Omega}) = \sum_{n=0}^N \sum_{m=-n}^n c_n^m(t) S_n^m(\boldsymbol{\Omega}) \quad (50)$$

【数 5 1】

$$= x(t) \underbrace{\left[ \sum_{n=0}^N \sum_{m=-n}^n S_n^m(\boldsymbol{\Omega}_0) S_n^m(\boldsymbol{\Omega}) \right]}_{v_N(\theta)}. \quad (51)$$

【0069】

式(51)から理解されるように、これは、一般的な平面波関数  $x(t)$  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と  $\theta_0$  との間の角度  $\theta$  のみに依存するように示されている。

【数 5 2】

$$\cos \theta = \cos \theta \cos \theta_0 + \cos(\phi - \phi_0) \sin \theta \sin \theta_0. \quad (52)$$

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。

10

20

30

40

【数 5 3】

$$\lim_{N \rightarrow \infty} v_N(\theta) = \frac{\delta(\theta)}{2\pi} . \quad (53)$$

【0 0 7 0】

しかしながら、有限次元  $N$  の場合には、方向  $\theta_0$  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 5 に示されている。

【0 0 7 1】

任意の方向  $\theta$  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、何らかの所定方向  $\theta_1$  および  $\theta_2$  についての関数  $c(t, \theta_1)$  および  $c(t, \theta_2)$  は、高い相関性がある。

10

【0 0 7 2】

C. 3 球面調和関数変換

平面波振幅の空間密度が  $N$  個の空間方向  $\theta_0, \dots, \theta_{N-1}$  で離散化される場合、空間方向  $\theta_0$  は単位球面上でほぼ均一に分布するのだが、 $N$  個の方向性信号  $c(t, \theta_0)$  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、

【数 5 4】

$$\mathbf{c}_{\text{SPAT}}(t) := [c(t, \theta_0) \quad \dots \quad c(t, \theta_{N-1})]^T , \quad (54)$$

20

式 (50) 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 (44) に定義される連続的なアンビソニックス表現  $c(t)$  から計算可能であることを検証できる。

$$\mathbf{c}_{\text{SPAT}}(t) = \mathbf{H}^H \mathbf{c}(t) \quad (55)$$

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 5】

$$\boldsymbol{\Psi} := [S_1 \quad \dots \quad S_0] \quad (56)$$

ここで、

30

【数 5 6】

$$S_o := [S_0^0(\theta_o) \quad S_1^{-1}(\theta_o) \quad S_1^0(\theta_o) \quad S_1^1(\theta_o) \quad \dots \quad S_N^{N-1}(\theta_o) \quad S_N^N(\theta_o)] . \quad (57)$$

【0 0 7 3】

方向  $\theta_0$  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  $c(t, \theta_0)$  から下記の式によって計算することができる。

【数 5 7】

$$\mathbf{c}(t) = \boldsymbol{\Psi}^{-H} \mathbf{c}_{\text{SPAT}}(t) . \quad (58)$$

40

【0 0 7 4】

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。

【0 0 7 5】

なお、方向  $\theta_0$  は単位球面上でほぼ均一に分布するため、近似計算

【数 5 8】

$$\boldsymbol{\Psi}^H \approx \boldsymbol{\Psi}^{-1} \quad (59)$$

が利用可能となり、式 (55) において、 $\boldsymbol{\Psi}^H$  の代わりに  $\boldsymbol{\Psi}^{-1}$  を使用することが正当化

50

される。

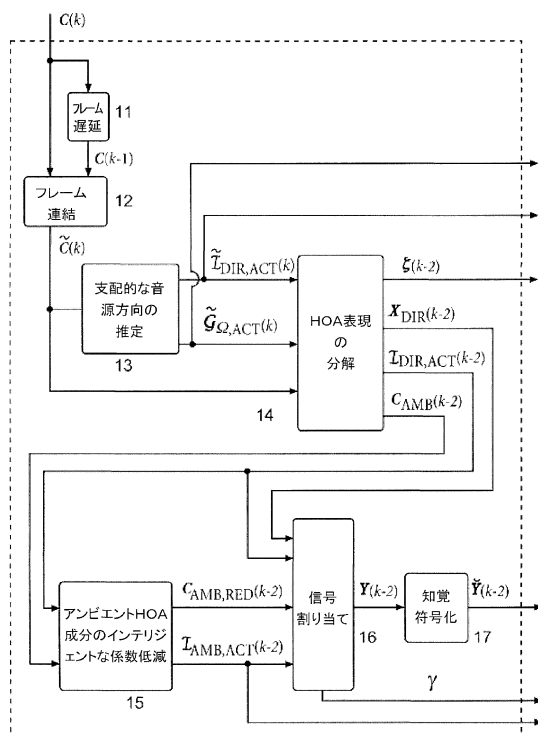
【 0 0 7 6 】

上述した関係の全てが離散時間領域にも有効であることは有利である。

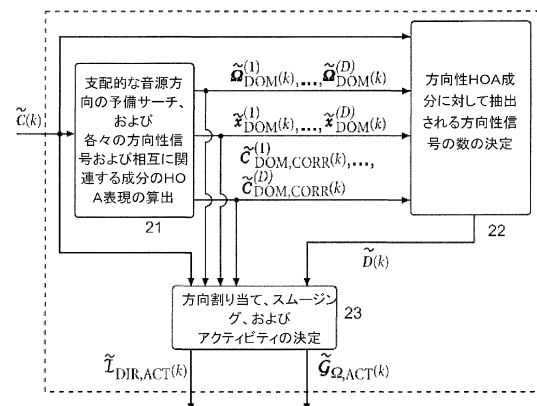
【 0 0 7 7 】

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、並列に動作する複数のプロセッサまたは電子回路、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、複数のプロセッサまたは電子回路で実行することができる。

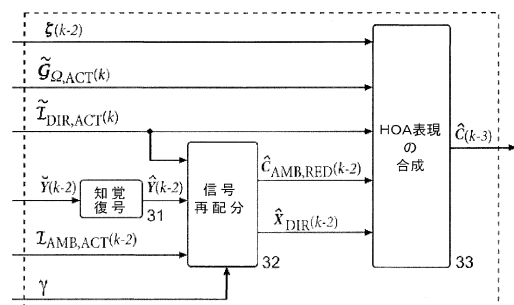
【 図 1 】



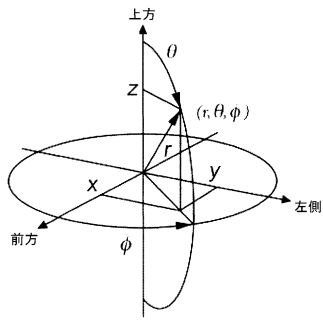
【 図 2 】



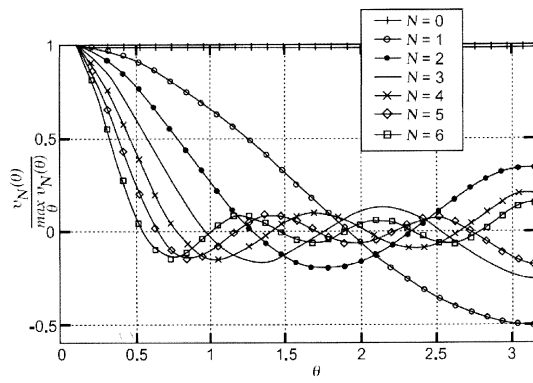
【 図 3 】



【図 4】



【図 5】



---

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

審査官 山下 剛史

- (56)参考文献 特開 2 0 1 2 - 1 3 3 3 6 6 ( J P , A )  
国際公開第 2 0 1 2 / 0 5 9 3 8 5 ( W O , A 1 )  
国際公開第 2 0 1 1 / 1 1 7 3 9 9 ( W O , A 1 )  
欧州特許出願公開第 2 0 9 4 0 3 2 ( E P , A 1 )

- (58)調査した分野(Int.Cl., DB名)
- G 1 0 L 1 9 / 0 0 - 1 9 / 2 6  
H 0 4 S 1 / 0 0 - 7 / 0 0  
A E S E - L i b r a r y