

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6474874号
(P6474874)

(45) 発行日 平成31年2月27日 (2019. 2. 27)

(24) 登録日 平成31年2月8日 (2019. 2. 8)

(51) Int. Cl.

F I

G 1 O L 21/0388 (2013.01)

G 1 O L 21/0388 1 0 0

請求項の数 3 外国語出願 (全 17 頁)

(21) 出願番号	特願2017-195350 (P2017-195350)	(73) 特許権者	598036300
(22) 出願日	平成29年10月5日 (2017. 10. 5)		テレフオンアクチーボラゲット エルエム
(62) 分割の表示	特願2016-107734 (P2016-107734)		エリクソン (パブル)
原出願日	平成24年12月21日 (2012. 12. 21)		スウェーデン国 スtockホルム エスー
(65) 公開番号	特開2018-41088 (P2018-41088A)	(74) 代理人	100076428
(43) 公開日	平成30年3月15日 (2018. 3. 15)		弁理士 大塚 康德
審査請求日	平成29年11月2日 (2017. 11. 2)	(74) 代理人	100115071
(31) 優先権主張番号	61/617, 175		弁理士 大塚 康弘
(32) 優先日	平成24年3月29日 (2012. 3. 29)	(74) 代理人	100112508
(33) 優先権主張国	米国 (US)		弁理士 高柳 司郎
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治

最終頁に続く

(54) 【発明の名称】 ハーモニックオーディオ信号の帯域幅拡張

(57) 【特許請求の範囲】

【請求項 1】

ハーモニックオーディオ信号の帯域幅拡張 B W E をサポートするオーディオデコーダ (5 0 1) を含むユーザ装置であって、

前記オーディオデコーダは、

周波数帯域 b 及び該周波数帯域 b に隣接する複数の周波数帯域に関連付けられた複数のゲイン値を受信する受信部と、

帯域幅拡張された周波数領域の再構成された対応する周波数帯域 b' にスペクトルピークが含まれるか否かを判定する判定部 (5 0 4) と、

前記再構成された周波数帯域 b' に少なくとも1つのスペクトルピークが含まれる場合、前記受信した複数のゲイン値に基づいて、前記再構成された周波数帯域 b' に関連付けられたゲイン値を、前記受信した複数のゲイン値の荷重和である第1の値に設定し、

前記再構成された周波数帯域 b' に全くスペクトルピークが含まれない場合、前記受信した複数のゲイン値に基づいて、前記再構成された周波数帯域 b' に関連付けられたゲイン値を、前記第1の値以下の値であって、前記受信した複数のゲイン値のうちの最小のゲイン値である第2の値に設定するゲイン適用部 (5 0 8) と、

を有することを特徴とするユーザ装置。

【請求項 2】

前記荷重和は、前記受信した複数のゲイン値の平均値であることを特徴とする請求項1に記載のユーザ装置。

10

20

【請求項 3】

前記オーディオデコーダは、原信号の高周波部分の少なくとも一部の区間のピークエネルギーとノイズフロアのエネルギーとの関係を反映する係数 α を受信するように更に構成され、

前記オーディオデコーダは、

前記受信した係数 α に基づいて、対応する再構成された高周波部分の変換係数をノイズと混合することにより、前記原信号の高周波部分のノイズ特性の再構成を可能にするノイズ混合部（510）

を更に有することを特徴とする請求項1又は2に記載のユーザ装置。

【発明の詳細な説明】

10

【技術分野】

【0001】

[01]提案する技術は、オーディオ信号の符号化、復号化に関し、特に、ハーモニックオーディオ信号の帯域幅拡張（BWE）をサポートするものに関する。

【背景技術】

【0002】

[02]変換ベースの符号化は、今日の音声圧縮/伝送システムで最も一般的に使用される方式である。この方式における主要な工程は、まず、DFT（離散フーリエ変換）、DCT（離散コサイン変換）、又はMDCT（修正離散コサイン変換）などの好適な変換によって信号波形のショートブロックを周波数領域に変換することである。変換係数は、量子化され、送信または保存され、その後、オーディオ信号を再構成するために使用される。この手法は一般的なオーディオ信号に適しているが、変換係数の十分に良好な表現を作成するためには十分に高いビットレートを必要とする。以下、このような変換領域符号化方式について詳細に説明する。

20

【0003】

[03]符号化される波形は、ブロックごとに、周波数領域に変換される。この目的のために一般に用いられる変換の一つが、離散コサイン変換（MDCT）である。得られた周波数領域ベクトルは、スペクトル包絡（ゆっくりと変化するエネルギー）とスペクトル残差に分割される。スペクトル残差は、得られた周波数領域ベクトルをスペクトル包絡で正規化することによって得られる。スペクトル包絡が量子化され、デコーダには量子化インデックスが送信される。次に、量子化されたスペクトル包絡は、ビット配分アルゴリズムへの入力として使用され、残差ベクトルの符号化のためのビットはスペクトル包絡の特性に基づいて分配される。このステップの結果として、一定数のビットが、残差（残差ベクトルまたは"サブベクトル"）の異なる部分に割り当てられる。いくつかの残差ベクトルは、ビットを受信せず、ノイズ・フィル又は帯域幅拡張する必要がある。典型的には、残差ベクトルの符号化は、まず、ベクトル要素の振幅を符号化し、次に、非ゼロ要素の符号（フーリエ変換等に係る「位相」と混同すべきではない）を符号化する、という2段階の手順を含む。残差の振幅及び符号の量子化インデックスはデコーダに送信され、そこで残差とスペクトル包絡とが結合され、最終的に時間領域に変換される。

30

【0004】

40

[04]通信ネットワークにおける容量は増加し続けている。しかし、容量が増加しているにもかかわらず、依然として、通信チャネルごとに必要な帯域幅を制限することが強く要請されている。移動体ネットワークにおいて、各呼のための送信帯域幅が小さいほど、移動体装置とそれにサービスを提供する基地局との両方の電力消費が少なく済む。これは、移動体通信業者にとってはエネルギー及びコストの削減、ひいてはエンドユーザが長時間のバッテリー寿命及び通話時間の増加をもたらすことにつながる。また、ユーザごとに消費される帯域幅が小さくなれば、移動体ネットワークにより（並行して）より多くのユーザにサービスを提供することができる。

【0005】

[05]低ビットレート又は中ビットレートで伝送されるオーディオ信号の品質を向上させ

50

る一つの方法は、利用可能なビットを、オーディオ信号の低周波域を高精度に表すために、集中させることである。BWE技術は、低周波数に基づいて、少ないビット数で高周波数をモデル化するために使用することができる。これらの技法は、人間の聴覚の感度は周波数に依存すること、具体的には、人間の聴覚すなわち私たちの聞こえ方は高い周波数にはあまり正確ではないこと、を利用している。

【0006】

[06]典型的な周波数領域BWE方式において、高周波変換係数が帯域ごとにグループ化される。各帯域のゲイン（エネルギー）が計算され、量子化され、（信号のデコーダに）送信される。デコーダでは、受信した低周波係数の反転もしくは変換されエネルギー正規化されたバージョンが、高周波ゲインでスケールされる。スペクトルエネルギーは、対象信号の高周波帯域のそれに近似しているため、少なくともこの点において、BWEは全く「不明」なわけではない。

10

【0007】

[07]しかし、特定のオーディオ信号のBWEは、聴取者に耳障りな欠陥を含んだオーディオ信号になってしまうことがある。

【発明の概要】

【0008】

[08]本明細書では、ハーモニックオーディオ信号のBWEをサポートし性能を向上させるための技術を提案する。

【0009】

20

[09]第一の側面によれば、変換オーディオデコーダにおける方法が提案される。この方法は、ハーモニックオーディオ信号の帯域幅拡張BWEをサポートするためのものである。提案する方法は、周波数帯域 b 及び該周波数帯域 b に隣接する複数の周波数帯域に関連付けられた複数のゲイン値を受信するステップを含みうる。提案する方法はさらに、帯域幅拡張された周波数領域の再構成された対応する周波数帯域 b' に少なくとも1つのスペクトルピークが含まれるか否かを判定するステップを含む。さらに、その帯域に少なくとも1つのスペクトルピークが含まれる場合、方法は、受信した複数のゲイン値に基づいて、周波数帯域 b' に関連付けられたゲイン値 G_b を第1の値に設定するステップを含む。その帯域にスペクトルピークが含まれていない場合、方法は、受信した複数のゲイン値に基づいて、周波数帯域 b' に関連付けられたゲイン値 G_b を第2の値に設定するステップを含む。これにより、ゲイン値をスペクトルの帯域幅拡張された部分のピーク位置に応じた値にすることができる。

30

【0010】

[010]さらに、本方法は、原信号の高周波部分の少なくとも一部の区間のピークエネルギーとノイズフロアエネルギーとの関係を反映するパラメータ又は係数 α を受信するステップを含みうる。方法はさらに、受信した係数 α に基づいて、対応する再構成された高周波部分の変換係数をノイズと混合するステップを含みうる。これにより、原信号の高周波部分のノイズ特性の再構成／エミュレーションが可能になる。

【0011】

[011]第2の側面によれば、ハーモニックオーディオ信号の帯域幅拡張BWEをサポートする、オーディオデコーダあるいはコーデックが提案される。変換オーディオコーデックは、上述の動作を実行するように構成された機能部を含みうる。さらに、変換オーディオデコーダに提供されるときに、本明細書に記載のノイズ混合を可能にする1又は2以上のパラメータを導出し提供するように構成された機能部を含むオーディオエンコーダ又はコーデックが提案される。

40

【0012】

[012]第3の側面によれば、第2の側面に係る変換オーディオコーデックを含むユーザ端末が提案される。ユーザ端末は、モバイル端末、タブレットコンピュータ、スマートフォン等の装置でありうる。

【図面の簡単な説明】

50

【0013】

[013]以下、添付図面を参照して、実施形態により、提案技術を詳細に説明する。

【図1】ハーモニックオーディオスペクトル、すなわち、ハーモニックオーディオ信号のスペクトルを示す図。このタイプのスペクトルは、典型的には、例えば、単一の楽器音、ボーカル音などである。

【図2】帯域幅拡張されたハーモニックオーディオスペクトルを示す図。

【図3a】デコーダによって受信され、対応するBWE帯域ゲイン $\wedge G_b$ でスケールリングされたBWEスペクトル（図2にも示される）を示す図。スペクトルのBWE部分がかなり歪んでいる。

【図3b】本明細書で示唆するように、修正BWE帯域ゲイン $\wedge G^{mod}_b$ でスケールリングされたBWEスペクトルを示す。この場合、スペクトルのBWE部分は所望の形状を得る。

【図4a】、

【図4b】実施形態に係る変換オーディオデコーダにおける動作手順を示すフローチャート。

【図5】実施形態に係る変換オーディオデコーダのブロック図。

【図6】実施形態に係る変換オーディオエンコーダにおける動作手順を示すフローチャート。

【図7】実施形態に係る変換オーディオエンコーダのブロック図。

【図8】実施形態に係る変換オーディオデコーダの構成を示すブロック図。

【発明を実施するための形態】

【0014】

[014]ハーモニックオーディオ信号(harmonic audio signals)の帯域幅拡張は、上記で示したようないくつかの問題を伴う。低周波数帯域、すなわち符号化され伝送され復号化される周波数帯域の一部、が反転(flipped)又は変換(translated)されて高周波数帯域が形成される場合、原信号又は「真の」高周波数帯域におけるスペクトルピークと同じ帯域にスペクトルピークが現れるかは不確かである。原信号ではピークがない帯域に、低域から生成されたスペクトルピークが現れてしまう可能性がある。また、その逆に、原信号ではピークを有する帯域に、(反転又は変換後に)ピークがない低域信号の一部が現れてしまう可能性がある。図1に高調波スペクトルの例を示し、BWEの概念を図2に示す。以下、詳細に説明する。

【0015】

[015]上記の効果は、主に高調波成分を有する信号に重大な品質劣化を引き起こす可能性がある。その理由は、ピークとゲインの位置の不整合によって、不必要なピーク減衰、又は、2つのスペクトルピークの間の低エネルギースペクトル係数の増幅、が引き起こされるためである。

【0016】

[016]本明細書に記載される解決策は、ピークの位置に関する情報に基づいて帯域幅拡張領域における帯域ゲインを制御する新規な方法に関する。さらに、本明細書で提案されるBWEアルゴリズムは、送信されるノイズ混合レベルによって「スペクトルピーク対ノイズフロア」比を制御することができる。これは、拡張された高周波数域の構造の量を保存するBWEとなる。

【0017】

[017]本明細書中に記載される解決策は、ハーモニックオーディオ信号に使用するのに適している。図1は、ハーモニックオーディオ信号の周波数スペクトルを示す図であり、これは高調波スペクトルを示している。図からわかるように、スペクトルはピークを含む。この種のスペクトルは、フルートなどの単一の楽器音やボーカル音などが典型的である。

【0018】

[018]ここで、ハーモニックオーディオ信号のスペクトルの2つの部分について説明する。1つは、低い周波数を含む低域であり、ここで「低」とは、帯域幅拡張に供される部

10

20

30

40

50

分よりも低いことを示す。もう1つは、低域よりも高い周波数を含む高域である。本明細書において、「低域 (lower part)」、「低い周波数 (low/lower frequencies)」という表現は、BWEクロスオーバー周波数以下のハーモニックオーディオスペクトルの一部を指す (図2参照)。同様に、「高域 (upper part)」、「高い周波数 (high/higher frequencies)」という表現は、BWEクロスオーバー周波数より上のハーモニックオーディオスペクトルの一部を指す (図2参照)。

【0019】

[019]図2は、ハーモニック音声信号のスペクトルを示す。ここには、以下で説明する、BWEクロスオーバー周波数の左側の低域と、BWEクロスオーバー周波数の右側の高域の2つの部分との2つの部分を見ることができる。図2において、原スペクトル、すなわち (エンコーダ側で見られるような) 原オーディオ信号のスペクトル、はライトグレーで示されている。スペクトルの帯域幅拡張された部分は、ダークグレーで示されている。前述したように、スペクトルの帯域幅拡張部分は、エンコーダによって符号化されるのではなく、スペクトルの受信した低域部分を使用することにより、デコーダで再生成される。図2においては、比較のため、原スペクトル (ライトグレー) およびBWEスペクトル (ダークグレー) の両方が、高い周波数で見ることができる。高い周波数の原スペクトルは、各BWE帯域 (すなわち高周波帯域) のゲイン値を除き、デコーダには不明である。BWE帯域は、図2の破線によって分割される。

【0020】

[020]図3aは、スペクトルの帯域幅拡張部のゲイン値とピーク位置との間の不整合の問題をよりよく理解できるように示したものである。帯域302aにおいて、原スペクトルはピークを有するのにも、生成されたBWEスペクトルにはピークがない。これは、図2における帯域202からわかる。ピークを含む原帯域に対して計算されるゲインがピークを含まないBWE帯域に適用されると、帯域302aに見られるように、BWE帯域における低エネルギースペクトル係数が増幅されてしまう。

【0021】

[021]図3aの帯域304aは反対の状況を呈し、原スペクトルの対応する帯域はピークを含まないが、生成されたBWEスペクトルの対応する帯域はピークを含んでいる。したがって、(エンコーダから受信した) その帯域に対して得られたゲインは、低エネルギー帯域のために計算される。このゲインがピークを有する対応する帯域に適用されると、図3aの帯域304aに見られるように、結果は、減衰されたピークとなる。知覚又は心理音響の観点からすると、帯域302aに示されている状況は、様々な理由で、帯域304aの状況よりも聴取者にとって悪い状況である。簡単に説明すると、聴取者が音声成分の異常な欠如よりも音の成分の異常な存在を体験することのほうが一般には不快である。

【0022】

[022]以下、本明細書の概念を表す新規なBWEアルゴリズムの例を説明する。

【0023】

[023]BWE領域における変換係数の集合 (高周波変換係数) を、 $Y(k)$ とする。これらの変換係数は以下で示されるB個の帯域にグループ化される。

【数1】

$$\{Y_b\}_{b=1}^B$$

帯域のサイズ M_b は一定であってもよいし、周波数が高くなるに従い増加させてもよい。一例として、帯域を8次元で均等分割 (すなわち、全て $M_b=8$) とする場合には、 $Y_1=\{Y(1) \dots Y(8)\}$, $Y_2=\{Y(9) \dots Y(16)\}$ となる。

【0024】

[024]BWEアルゴリズムの第1のステップは、全ての帯域のゲインを計算することである。

10

20

30

40

【数 2】

$$G_b = \sqrt{\frac{\mathbf{Y}_b^T \mathbf{Y}_b}{M_b}} \quad (1)$$

【0025】

[025] これらのゲインは、 $\hat{G}_b = Q(G_b)$ として量子化され、デコーダに伝送される。

【0026】

[026] BWE アルゴリズムにおける第2のステップ（オプションである）は、例えば BWE スペクトルの平均ピークエネルギー E_p 及び平均ノイズフロアエネルギー E_{nf} の関数であるノイズ混合パラメータ又は係数 α を次式のように計算することである。

10

【数 3】

$$\alpha = f\left(\frac{\overline{E}_{nf}}{\overline{E}_p}\right) \quad (2)$$

ここで、パラメータ α は、下記（3）式に従って導出される。ただし、使用される正確な表現は、使用するコーデックや量子化器に対して何が適切かを考慮するなどして、様々な方法で選択することができる。

【数 4】

20

$$\alpha = \left(10 \frac{\overline{E}_{nf}}{\overline{E}_p}\right)^3 \quad (3)$$

【0027】

[027] ピークとノイズフロアエネルギーは、例えばスペクトルエネルギーの最大値及び最小値のそれぞれを追跡することによって計算されうる。

【0028】

[028] ノイズ混合パラメータ α は、少ないビット数で量子化されうる。例えば、 α は 2 ビットで量子化される。ノイズ混合パラメータ α を量子化すると、パラメータ $\hat{\alpha} = Q(\alpha)$ が得られる。パラメータ $\hat{\alpha}$ はデコーダへと伝送される。BWE 領域は 2 つ以上のセクション S に分割され、ノイズ混合パラメータ α_s が、これらのセクションの各々において、独立に計算されうる。このような場合には、エンコーダはノイズ混合パラメータのセットを、例えばセクション当たり 1 つずつ、デコーダに送信することができる。

30

【0029】

<デコーダの動作>

[029] デコーダは、ビットストリームから、（各帯域毎に）計算された量子化ゲイン $\hat{G}_b$ のセットと 1 つ以上の量子化ノイズ混合パラメータ又は係数 $\hat{\alpha}$ を抽出する。デコーダはまた、スペクトルの低周波数部分、すなわち、帯域幅拡張される高周波部分とは対照的に符号化された（ハーモニックオーディオ信号の）スペクトルの一部、の量子化された変換係数を受信する。

40

【0030】

[030] エネルギー正規化され量子化された低周波係数の集合を、 $\hat{X}_b$ とする。これらの係数は、ノイズ、例えば予め生成されノイズコードブック N_b に記憶されたノイズ、と混合される。予め生成され記憶されたノイズを使用することで、ノイズの品質を確保することができる。すなわち、意図しない不一致や偏差が含まれないようにすることができる。しかし、ノイズは代わりに、必要に応じ、「オンザフライ (on the fly)」で生成することができる。係数 $\hat{X}_b$ は、次式のように、ノイズコードブック N_b のノイズと混合されうる。

【数 5】

$$\hat{X}_b^{\text{mod}} = (1 - \hat{\alpha}) \hat{X}_b + \hat{\alpha} N_b \quad (4)$$

【0031】

[031] ノイズ混合パラメータまたは係数の範囲は、様々な方法で設定することができる。例えば、ノイズ混合係数の範囲は次のように設定される。

【数 6】

$$\alpha \in [0, 0.4)$$

10

この範囲は、例えば、ノイズの影響を完全に無視できる場合（ $\alpha = 0$ ）、及び、この範囲を使用したときの最大の寄与である、ノイズコードブックが混合ベクトルにおいて40%寄与する場合（ $\alpha = 0.4$ ）、を意味する。この種のノイズ混合を導入する結果得られるベクトルが原音の低域の構造の例えば60%から100%を含むのは、スペクトルの低周波部分よりも高周波部分のほうが一般にノイズ成分が多いからである。そのため、上述のノイズ混合動作は、スペクトルのBWE高周波数領域を低周波領域の反転または変換された成分から作るものと比較して、より良い、原信号のスペクトルの高周波部分の統計的性質がよく似たベクトルを生成する。ノイズ混合動作は、例えば複数のノイズ混合係数（ α ）が提供され受信される場合、BWE領域のそれぞれの区間ごとに独立に実行されうる。

20

【0032】

[032] 従来技術の解決策では、受信された量子化ゲイン $\hat{G}_b$ のセットはBWE領域内の対応する帯域に直接使用される。しかし、本実施形態では、適切である場合には、BWEスペクトルのピーク位置に関する情報に基づき、これらの受信された量子化ゲイン $\hat{G}_b$ が最初に修正される。必要なピークの位置に関する情報は、ビットストリーム中の低周波領域情報から抽出することができ、あるいは、低域の量子化変換係数（またはBWE帯域の導出された係数）のピークピッキングアルゴリズムによって推定される。低周波領域でのピークに関する情報は、高周波（BWE）領域に変換することができる。高域（BWE）信号が低域信号から導出される場合には、アルゴリズムは、（BWE領域の）スペクトルピークが位置する帯域を登録することができる。

30

【0033】

[033] 例えば、フラグ $f_p(b)$ は、BWE領域における帯域 b に移動（反転または変換）された低周波係数がピークを含むかどうかを示すために用いることができる。例えば、 $f_p(b) = 1$ は、帯域 b に少なくとも1つのピークが含まれることを示し、 $f_p(b) = 0$ は、帯域 b にはピークが含まれないことを示す、とすることができる。前述したように、BWE領域の各帯域 b は、原信号の対応する帯域に含まれるピークの数およびサイズに依存するゲイン $\hat{G}_b$ と関連している。BWE領域における各帯域の実際のピーク内容にゲインを一致させるために、ゲインが適応化されるべきである。ゲイン修正は、例えば次式に従い、各帯域ごとに行われる。

【数 7】

40

$$\hat{G}_b^{\text{mod}} = \begin{cases} \frac{1}{3}(\hat{G}_{b-1} + \hat{G}_b + \hat{G}_{b+1}) & \text{if } f_p(b) = 1 \\ \min\{\hat{G}_{b-1}, \hat{G}_b, \hat{G}_{b+1}\} & \text{if } f_p(b) = 0 \end{cases} \quad (5a)$$

ゲインを修正するのは以下の理由による。（BWE）帯域にピークが含まれる場合（ $f_p(b) = 1$ ）、そのピークが減衰されることを避けるべく、対応するゲインがピークのない（原信号の）帯域からのものである場合、この帯域のゲインは、現在の帯域とそれに隣接する2つの帯域のゲインの荷重和に修正される。上記の式（5a）において、重みは1／

50

3である。これは、修正ゲインが現在の帯域のゲインと隣接する2つの帯域のゲインとの平均値であることを意味する。

別のゲイン修正を、例えば次式に従って行うことができる。

【数8】

$$\hat{G}_b^{\text{mod}} = \begin{cases} (0.1\hat{G}_{b-1} + 0.8\hat{G}_b + 0.1\hat{G}_{b+1}) & \text{if } f_p(b) = 1 \\ \min\{\hat{G}_{b-1}, \hat{G}_b, \hat{G}_{b+1}\} & \text{if } f_p(b) = 0 \end{cases} \quad (5b)$$

ピークが含まれない場合 ($f_p(b) = 0$)、1つ以上のピークを含む原信号帯域から計算された大きなゲインが適用されてしまうことでこの帯域におけるノイズ様構造が増幅されてしまうのは望ましくない。これを避けるために、この帯域のゲインが、例えば現在の帯域のゲインと隣接する2つの帯域のゲインの最小値となるように選択される。あるいは、ピークを含む帯域のゲインは、3つ以上の帯域、例えば5又は7帯域の平均、といった荷重和として選択又は計算されるか、または、3, 5, もしくは7帯域の中央値として選択される。平均や中央値などの荷重和を用いることで、ピークは「真の」ゲインを用いた場合よりもわずかに減衰されることになる。しかし、「真の」ゲインと比べたときの減衰は、その反対よりは有益であるといえる。それは、知覚の観点からは適度な減衰が好ましく、増幅によってオーディオ成分が増大してしまうことに比べれば有益なためである。

【0034】

[034] ピーク mismatches のために、ゲインを修正する理由は、事前定義されたグリッド上にスペクトル帯域が配置されることであるが、ピーク位置および（反転または低周波数係数を変換した後の）ピークは、時間とともに変化する。これはピークが制御されない方法で帯域の中に入る又は外に出ることに起因する。したがって、スペクトルの BWE 部のピーク位置は、必ずしも原信号のピーク位置と一致せず、したがって、帯域に関連付けられたゲインとその帯域のピークの内容との間の不一致はありうることである。未修正のゲインでスケールリングされる例は図3aに示され、修正されたゲインでスケールリングされる例は図3bに示される。

【0035】

[035] 本実施形態で説明したように修正されたゲインを用いた結果が、図3bに示される。帯域302aにおいて、低エネルギースペクトル係数は、図3aの帯域302aのように増幅されず、より適切な帯域ゲインでスケールリングされる。さらに、帯域304bのピークは、図3aの帯域304aのピークのように減衰もされない。多くの場合、図3bに示されたスペクトルは、図3aのスペクトルに対応するオーディオ信号よりも聴取者に対してより快適である音声信号に対応する。

【0036】

[036] したがって、BWE アルゴリズムは、スペクトルの高周波数部分を作成することができる。（帯域幅の節約上の理由などにより）高周波係数 Y_b のセットがデコーダで利用できないため、高周波係数

【数9】

$$\tilde{Y}_b$$

が代わりに再構成され形成される。これは反転（又は変換）された低周波係数（ノイズ混合後でもよい）を修正量子化ゲインでスケールリングすることで得られる。

【数10】

$$\tilde{Y}_b = \hat{G}_b^{\text{mod}} \hat{X}_b^{\text{mod}} \quad (6)$$

この変換係数のセット

10

20

30

40

【数 1 1】

 $\tilde{Y}_b$

は、オーディオ信号波形の高周波部分を再構築するために使用される。

【0037】

[037]本実施形態に記載される解決策は、一般に変換領域オーディオ符号化において使用される、BWEの概念に対する改善である。提示されたアルゴリズムは、BWE領域におけるピーク構造（ピーク対ノイズフロア比）を維持し、これにより、再構成された信号のオーディオ品質の向上を提供する。

10

【0038】

[038]「変換オーディオコーデック（transform audio codec）」または「変換コーデック（transform codec）」の用語は、エンコーダとデコーダのペアを包含し、そして当該分野で通常使用される用語である。本開示において、変換コーデックの機能／部分を別々に記述するため、「変換オーディオエンコーダ」又は「エンコーダ」と、「変換オーディオデコーダ」または「デコーダ」の用語が使用される。「変換オーディオエンコーダ」／「エンコーダ」、ならびに、「変換オーディオデコーダ」／「デコーダ」の用語は、「変換オーディオコーデック」／「変換コーデック」の用語と交換可能である。

【0039】

<デコーダにおける手順の例（図4a，4b）>

20

[039]以下、デコーダにおける、ハーモニックオーディオ信号の帯域幅拡張BWEをサポートするための手順の例を、図4aを参照して説明する。手順は、例えばMDCTエンコーダまたはその他エンコーダのような変換オーディオ符号化における使用に適している。オーディオ信号は、主に音楽を含むことが想定されるが、その代わりまたはそれに加えて、音声が入り込んでもよい。

【0040】

[040]ステップ401aにおいて、周波数帯域bに関連付けられたゲイン値（原周波数帯域）と周波数帯域bに隣接する複数の他の周波数帯域に関連付けられたゲイン値が受信される。次に、ステップ404aにおいて、BWE領域の再構成された対応する周波数帯域b'にスペクトルピークが含まれるか否かが判定される。再構成された周波数帯域b'に少なくとも一つのスペクトルピークが含まれる場合は、ステップ406a：1で、受信した複数のゲイン値に基づいて、再構成された周波数帯域b'に関連付けられたゲイン値が第1の値に設定される。再構成された周波数帯域b'にスペクトルピークが含まれない場合は、ステップ406a：2で、受信した複数のゲイン値に基づいて、再構成された周波数帯域b'に関連付けられたゲイン値が第2の値に設定される。第2の値は第2の値以下である。

30

【0041】

[041]図4bには、図4aに示した手順と少し異なり、より拡張されたかたちで、例えば図示されている前述のノイズ混入に関連する追加オプションの動作を持つ方法が示される。以下、図4bについて説明する。

40

【0042】

[042]ステップ401bにおいて、周波数スペクトルの高域に関連付けられたゲイン値が受信される。変換係数やゲイン値などの、周波数スペクトルの低い部分に関連する情報も、いずれかの時点で受信されているものとする（図4a，4bには示されない）。また、前述したように、低域スペクトルを反転又は変換することにより高域スペクトルが作成される帯域幅拡張はがいずれかの時点で行われるものとする。

【0043】

[043]ステップ402bにおいて、1つ以上のノイズ混合係数が受信される。これら受信された1つ以上ノイズ混合係数は、原高域スペクトルのエネルギー分布に基づいて、エンコーダにおいて計算されたものである。（同じくオプションである）ステップ403b

50

において、例えば上述の式（４）に従い、ノイズ混合係数が、高域領域の係数とノイズを混合するために使用されうる。したがって、帯域幅拡張領域のスペクトルは「ノイズネス」またはノイズの内容の点で原高域スペクトルにより良く対応することになる。

【００４４】

[044]次に、ステップ４０４ｂにおいて、作成されたＢＷＥ領域の帯域にピークが含まれるか否かが判定される。例えば、帯域にピークが含まれる場合、当該帯域に係る指標が１に設定される。他の帯域にピークが含まれない場合は、当該帯域に係る指標が０に設定される。ステップ４０５ｂでは、当該帯域に関連付けられたゲインが修正されうる。帯域のゲインを修正する場合、前述のように、所望の結果を達成するため、隣接する帯域のゲインも考慮される。このようにゲインを修正することにより、ＢＷＥスペクトルが改善される。ステップ４０６ｂにおいて、修正されたゲインがＢＷＥスペクトルの各帯域に適用される。

10

【００４５】

＜デコーダの例＞

[045]以下、ハーモニックオーディオ信号の帯域幅拡張ＢＷＥをサポートする上記の手順を実行するように構成された変換オーディオデコーダの例を、図５を参照して説明する。変換オーディオデコーダは例えば、ＭＤＣＴデコーダその他のデコーダでありうる。

【００４６】

[046]変換オーディオデコーダ５０１は通信部５０２を介して他のエンティティと通信するものとして示されている。上記した手順を実行可能な変換オーディオデコーダの一部は、破線で囲まれた構成５００として示されている。変換オーディオデコーダはまた、例えば通常のデコーダやＢＷＥ機能を提供する他の機能部５１６を含み、更には、１つ以上の記憶部５１４を含みうる。

20

【００４７】

[047]変換オーディオデコーダ５０１及び構成５００の少なくともいずれかは、例えばプロセッサ又はマイクロプロセッサ、適当なソフトウェア及びそれを格納するための記憶装置、プログラマブルロジックデバイス（ＰＬＤ）その他の電子部品のうちの１つ以上によって実装されうる。

【００４８】

[048]変換オーディオデコーダは、符号化エンティティから提供された適当なパラメータを取得するための機能部を有することが想定される。ノイズ混合係数は、従来技術に対して、取得するための新規なパラメータである。したがって、デコーダは、この機能が所望されるときに、１つ以上のノイズ混合係数が取得されるように構成されるべきである。オーディオデコーダは、受信部を有するように実装され、受信部は、周波数帯域ｂ及び当該周波数帯域ｂに隣接する複数の周波数帯域に関連付けられた複数のゲイン値、及び、可能なノイズ混合係数を受信する。ただし、このような受信部は図５に示されていない。

30

【００４９】

[049]変換オーディオデコーダは、ＢＷＥスペクトル領域のどの帯域にピークが含まれるどの帯域にピークが含まれないかを判定する判定部５０４あるいはピーク検出部を有する。判定部は、帯域幅拡張領域における再構成された対応する周波数帯域ｂ'にスペクトルピークが含まれるか否かを判定する。また、変換オーディオデコーダは、帯域にピークが含まれるか否かに依存してその帯域に係るゲインを修正するゲイン修正部５０６を含みうる。その帯域にピークが含まれる場合、修正ゲインは、例えば、注目帯域に隣接する複数の帯域の（修正前の）ゲインの平均値又は中央値などの、荷重和として計算される。

40

【００５０】

[050]変換オーディオデコーダは、更に、修正ゲインをＢＷＥスペクトルの適当な帯域に適用又は設定するゲイン適用部５０８を含みうる。ゲイン適用部は、再構成された周波数帯域ｂ'に少なくとも１つのスペクトルピークが含まれる場合、受信した複数のゲイン値に基づいて、再構成された周波数帯域ｂ'に係るゲイン値を第１の値に設定し、再構成された周波数帯域ｂ'に１つもスペクトルピークが含まれない場合は、受信した複数のゲ

50

イン値に基づいて、再構成された周波数帯域 b' に係るゲイン値を第 1 の値以下の第 2 の値に設定する。これにより、ゲイン値を帯域幅拡張領域におけるピーク位置に合わせて決めることが可能になる。

【0051】

[051]あるいは、修正なしに可能であれば、適用されるゲインが元のゲインではなく修正されたゲインであるということだけで、機能を適用することが（通常の）他の機能部 516 によって提供されてもよい。また、変換オーディオデコーダは、ノイズ混合部 510 を含む。ノイズ混合部 510 は、オーディオ信号のエンコーダによって提供される 1 つ以上のノイズ係数又はパラメータに基づいて、スペクトルの BWE 部分の係数を、例えばコードブックからのノイズと混合する。

10

【0052】

＜エンコーダの手順の例＞

[052]以下、図 6 を参照して、エンコーダにおける、ハーモニックオーディオ信号の帯域幅拡張 BWE をサポートするための手順の例を説明する。手順は、MDC T エンコーダその他のエンコーダなどの変換オーディオエンコーダでの使用に適している。前述したとおり、オーディオ信号は主に、音楽及び／又は音声が想定される。

【0053】

[053]以下の手順は、変換エンコーダを用いたハーモニックオーディオ信号の従来の符号化とは異なる符号化手順の部分に関するものである。したがって、以下で説明する動作は、スペクトルの低周波部分に対する変換係数やゲイン等を導出することに加えて、スペクトルの高周波部分の帯域（デコーダ側の BWE によって生成される部分）のゲインを導出するオプションとして説明される。

20

【0054】

[054]ステップ 602 で、周波数スペクトルの高周波部分のピークエネルギーが求められる。次に、ステップ 603 で、当該周波数スペクトルの高周波部分に係るノイズフロアエネルギーが求められる。例えば、前述のような、BWE スペクトルの 1 つ以上の区間の平均ピークエネルギー E_p 及び平均ノイズフロアエネルギー E_{nf} が計算されうる。次に、ステップ 604 で、BWE スペクトルのある区間に係るノイズ混合係数とその区間のノイズの量又は「ノイズネス」を反映するように、例えば前述の式 (3) に従い、ノイズ混合係数が計算される。ステップ 606 で、1 つ以上のノイズ混合係数が、復号化エンティティ又は記憶部に、エンコーダによって提供された従来の情報とともに、提供される。この提供は、計算されたノイズ混合係数の出力先への単純な出力、及び、計算されたノイズ混合係数のデコーダへの送信のうち少なくともいずれかであってもよい。ノイズ混合係数は、前述したように提供する前に量子化されうる。

30

【0055】

＜エンコーダの例＞

[055]以下、図 7 を参照して、ハーモニックオーディオ信号の帯域幅拡張 BWE をサポートするための上述の手順を実行する変換オーディオエンコーダの例を説明する。変換オーディオエンコーダは、MDC T エンコーダその他のエンコーダでありうる。

【0056】

[056]変換オーディオエンコーダ 701 は、通信部 702 を介して他のエンティティと通信するものとして示されている。上記した手順を実行可能な変換オーディオエンコーダの一部は、破線で囲まれた構成 700 として示されている。変換オーディオエンコーダはまた、例えば通常のエンコーダや BWE 機能を提供する他の機能部 712 を含み、更には、1 つ以上の記憶部 710 を含む。

40

【0057】

[057]変換オーディオエンコーダ 701 及び構成 700 の少なくともいずれかは、例えばプロセッサ又はマイクロプロセッサ、適当なソフトウェア及びそれを格納するための記憶装置、プログラマブルロジックデバイス (PLD) その他の電子部品のうちの 1 つ以上によって実装されうる。

50

【0058】

[058]変換オーディオエンコーダは、スペクトルの高周波部分のピークエネルギー及びノイズフロアエネルギーを判定する判定部704を有する。また、変換オーディオエンコーダは、スペクトルの高周波部分の全部又は一部に対する1つ以上のノイズ混合係数を計算するノイズ係数部706を有する。変換オーディオエンコーダは、更に、デコーダで使用されるために、計算したノイズ混合係数を提供する提供部708を有する。この提供は、計算されたノイズ混合係数の出力先への単純な出力、及び、計算されたノイズ混合係数のデコーダへの送信のうち少なくともいずれかであってもよい。

【0059】

<構成例>

[059]図8は、変換オーディオデコーダにおける使用に適した装置800の例を示す図であり、図5に示した変換オーディオデコーダにおける使用のための構成の例の代替案にもなりうるものである。装置800は、処理部806を有し、これはDSP（デジタルシグナルプロセッサ）でありうる。処理部806は、単一のユニットで構成されてもよいし、本明細書に記載した手順の異なるステップを実行する複数のユニットで構成されてもよい。装置800はまた、符号化スペクトルの低周波部分、スペクトル全体のゲイン、ノイズ混合係数といった信号（参照：エンコーダの場合、高調波スペクトルの高周波部分）を受信する入力部802と、修正ゲイン及び全体のスペクトルの少なくともいずれかといった信号（参照：エンコーダの場合、ノイズ混合係数）を出力する出力部804を有する。入力部802及び出力部804は、当該装置のハードウェアの1つとして構成されうる。

【0060】

[060]また、装置800は、EEPROM、フラッシュメモリ、ハードドライブなどの不揮発性または揮発性メモリの形態で、少なくとも1つのコンピュータプログラム製品808を有する。コンピュータプログラム製品808は、コンピュータプログラム810を含む。このコンピュータプログラム810は、装置800の処理部806により実行されると、当該装置及び変換オーディオエンコーダの少なくともいずれかに図4を参照して前述した手順の動作を実行させるためのコードを含む。

【0061】

[061]そして、実施形態では、装置800のコンピュータプログラム810のコードは、オーディオスペクトルの低周波部分に係る情報及びオーディオスペクトル全体に係るゲインを得るための取得モジュール810aを含む。さらに、オーディオスペクトルの高周波部分に関連するノイズ係数も取得することができる。コンピュータプログラムは、帯域幅拡張された周波数領域の再構成された帯域bの帯域にスペクトルピークが含まれるか否かを検出し指示するための検出モジュール810bを含みうる。コンピュータプログラム810はさらに、スペクトルの高周波部分の再構成された一部の帯域に関連付けられたゲインを修正するためのゲイン修正モジュール810cを含みうる。コンピュータプログラム810はさらに、スペクトルの高周波部分の対応する帯域に修正されたゲインを適用するためのゲイン適用モジュール810dを含みうる。また、コンピュータプログラム810は、受信したノイズ混合係数に基づいてスペクトルの高周波部分をノイズで混合するノイズ混合モジュール810dを含みうる。

【0062】

[062]コンピュータプログラム810は、コンピュータプログラムモジュールで構成されたコンピュータプログラムコードの形態である。モジュール810a～810dは、図4aまたは図4bに示すフローの動作を実行して図5に示す構成500をエミュレートする。すなわち、モジュール810a～dが処理部806で実行されると、それらは少なくとも図5の504～510の各部に対応する。

【0063】

[063]図8に関連して上述した実施形態におけるコードは、処理部によって実行されると、前記構成及び変換オーディオエンコーダの少なくともいずれかに上述の各図を用いて説明したステップを実行させるが、代替の実施形態においては、コードの一部がハードウ

エア回路として少なくとも一部として実装されてもよい。

【0064】

[064]同様に、コンピュータプログラムモジュールを有する実施形態は、図7に示した変換オーディオエンコーダにおける構成に対応するように記述することができる。

【0065】

[065]特定の例示的な実施形態を参照して提案技術を説明したが、それらの説明は、一般的に概念を説明することだけを意図したものであって、本明細書に記載された解決策の範囲を限定するものとして解釈されるべきではない。要件や好みに応じて、上記の例示の実施形態の異なる特徴を様々な方法で組み合わせることも可能である。

【0066】

[066]上述の解決策は、携帯端末、タブレットコンピュータ、スマートフォンなどの装置における、オーディオコーデックが適用されるあらゆる物に使用可能である。

【0067】

[067]ユニットとモジュールとの相互作用ならびにユニットの名称の選択は一例にすぎず、また、上述の方法を実行するのに適したノードは、提案した処理動作を実行することができるようにする他の複数の方法で構成されうることが、理解されるべきである。

【0068】

[068]また、本開示に記載のユニット又はモジュールは、別個の物理エンティティである必要はなく、論理エンティティとして考えることができる。上記の説明は多くの具体例を含むが、それらは、本発明の範囲を限定するものであると解釈されるべきではなく、現時点で好適な実施形態のいくつかの例示を提供するにすぎないと解釈されるべきである。従って、本発明の範囲は当業者には明らかになるだろう他の実施形態を完全に含むこと、並びにそのため範囲は限定されないことが理解されるだろう。単数形の要素は、「1つ及び1つだけ」と明示しない限りそのように解釈すべきではなく、むしろ「1つ以上」と解釈すべきである。当業者には既知である上述の好適な実施形態の要素に対する全ての構造的、機能的な均等物は、本発明に明確に組み込まれ、本発明に含まれることを意図する。更に、装置又は方法は、本発明に含まれるために本明細書中に記載されるか又は現在の技術により解決することを求められる全ての問題を解決する必要はない。

【0069】

[069]上述の説明では、説明及び限定の目的で、提案技術の完全な理解を提供するために、特定のアーキテクチャ、インターフェース、技術が詳しく記載されている。しかし、提案技術は、これらの特定の詳細から外れる他の実施形態においても実施されうることが、当業者には明らかであろう。つまり、提案技術の原理を明示的に記載、または本明細書に示され、具体化されていなくても、当業者は、その様々な構成を思い付くことができるであろう。不必要に詳しく説明することで提案技術の説明が不明瞭にならないように、いくつかの事例では、周知の装置、回路及び方法の詳細な説明が省略される。すべての記述は、その構造的および機能的な均等物を包含することが意図されている原理、態様を提示技術の実施形態と同様に、特定の例を列挙する。さらに、そのような均等物は、現在公知の均等物だけでなく、例えば、構造に関わらず、同じ機能を実行する構成要素などの、将来開発される均等物を含むことが意図される。

【0070】

[070]したがって、例えば、本明細書のブロック図は、技術の原理を具体化する回路その他の機能ユニットの概念を表すものであることは、当業者には理解されよう。同様に、あらゆるフローチャート、状態遷移図、擬似コードなどは、実質的にコンピュータ可読媒体で実現され、したがって、コンピュータやプロセッサが明示されているか否かにかかわらず、コンピュータまたはプロセッサによって実行することができることは、当業者に理解されよう。

【0071】

[071]「機能部」、「プロセッサ」、「制御部」として記述されたさまざまな機能要素は、その機能ブロックが特定の物に限定されるものではなく、回路ハードウェア及びコン

10

20

30

40

50

コンピュータ読み取り可能な記憶媒体に格納された命令のかたちのソフトウェアの少なくともいずれかで実現されうる。したがって、これらの機能や機能ブロックはハードウェア実装、コンピュータ実装、機械実装の少なくともいずれかで実現されうる。

【0072】

[072]ハードウェア実装において、機能ブロックは、デジタルシグナルプロセッサ（DSP）ハードウェア、縮小命令セットプロセッサ、特定用途向け集積回路（ASIC）を含むハードウェア（例えば、デジタル、アナログ）回路、これらの機能を実行可能な（適当な）状態マシンなどによって実現可能であるが、これらに限定されない。

【0073】

（略語）

BWE Bandwidth Extension

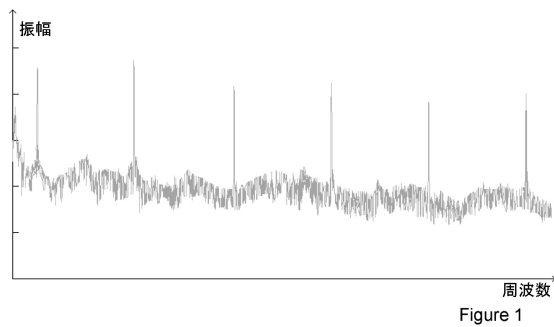
DFT Discrete Fourier Transform

DCT Discrete Cosine Transform

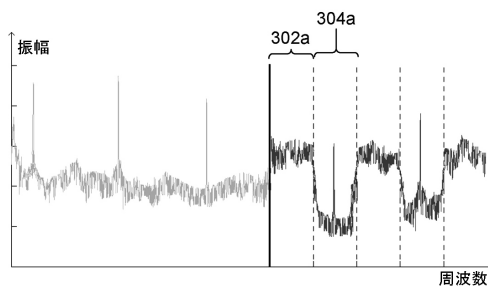
MDCT Modified Discrete Cosine Transform

10

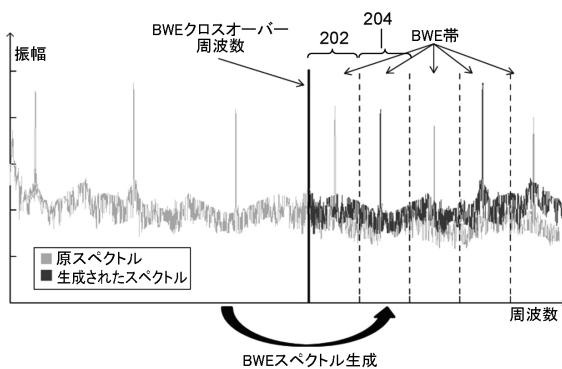
【図 1】



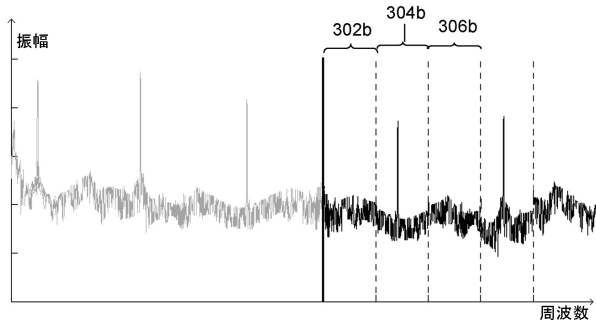
【図 3 a】



【図 2】



【図 3 b】



【図 4 a】

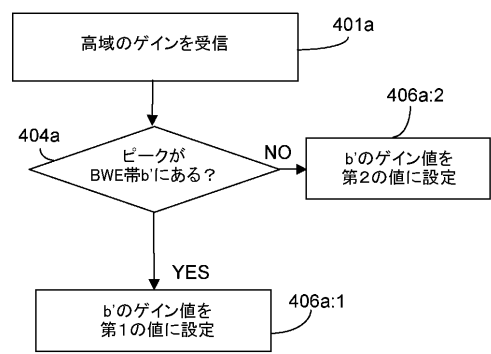


Figure 4a

【図 4 b】

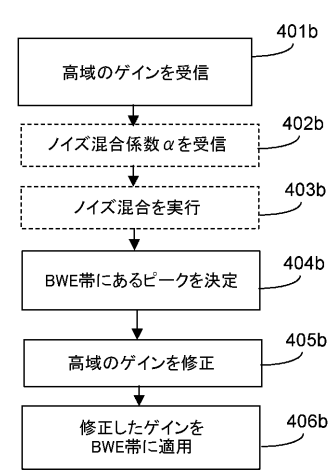


Figure 4b

【図 5】

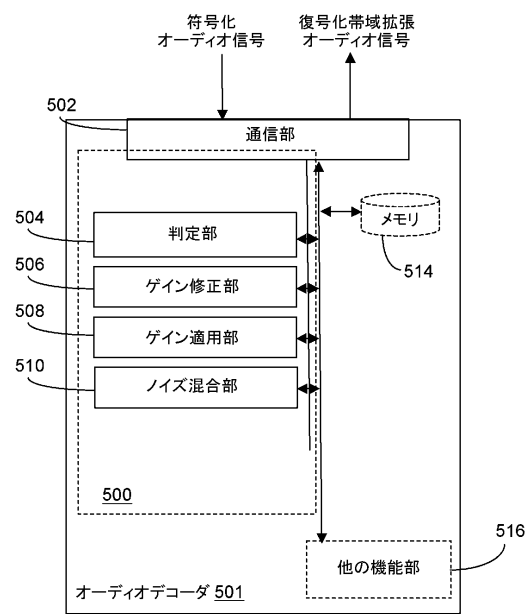


Figure 5

【図 6】

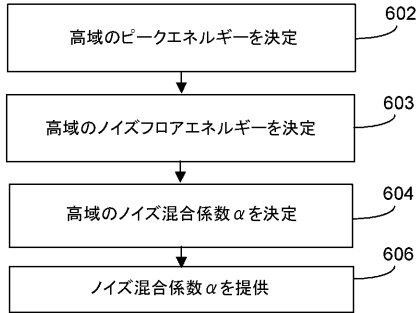


Figure 6

【図 7】

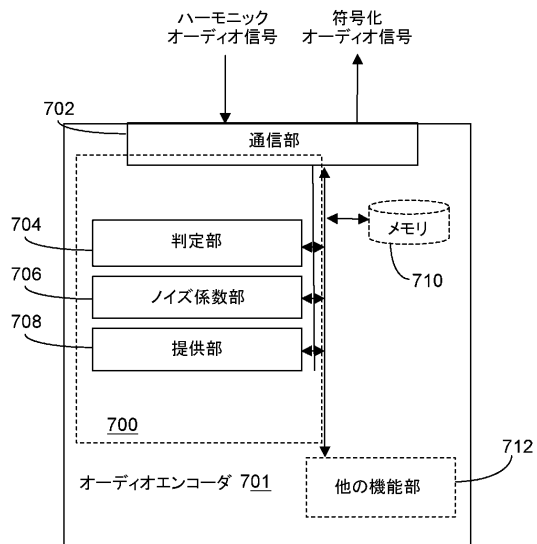


Figure 7

【図 8】

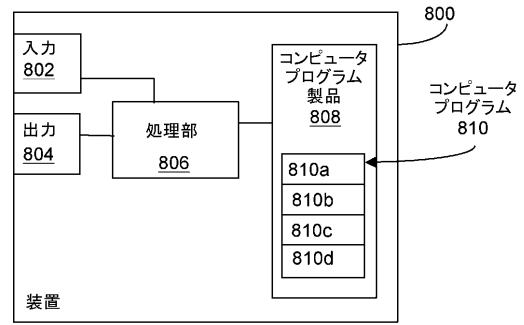


Figure 8

フロントページの続き

- (72)発明者 ナズランド, セバスチャン
スウェーデン国 ソルナ 169 33, ボクヴェーゲン 13
- (72)発明者 グランチャロブ, ヴォロージャ
スウェーデン国 ソルナ エス-171 67, アンクダムスガタン 29
- (72)発明者 ヤンソン, トフトガード, トマス
スウェーデン国 ウプサラ エス-754 35, アペルガタン 20エー

審査官 大野 弘

- (56)参考文献 特表2002-536679 (JP, A)
特表2005-510772 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
G10L 21/0388