

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6400702号
(P6400702)

(45) 発行日 平成30年10月3日 (2018. 10. 3)

(24) 登録日 平成30年9月14日 (2018. 9. 14)

(51) Int. Cl.

F I

G 1 0 L 19/02 (2013. 01)

G 1 0 L 19/02 1 5 0

G 1 0 L 21/0388 (2013. 01)

G 1 0 L 21/0388

請求項の数 14 (全 43 頁)

(21) 出願番号 特願2016-528417 (P2016-528417)
 (86) (22) 出願日 平成26年7月15日 (2014. 7. 15)
 (65) 公表番号 特表2016-527557 (P2016-527557A)
 (43) 公表日 平成28年9月8日 (2016. 9. 8)
 (86) 国際出願番号 PCT/EP2014/065118
 (87) 国際公開番号 W02015/010953
 (87) 国際公開日 平成27年1月29日 (2015. 1. 29)
 審査請求日 平成28年3月18日 (2016. 3. 18)
 (31) 優先権主張番号 13177346.7
 (32) 優先日 平成25年7月22日 (2013. 7. 22)
 (33) 優先権主張国 欧州特許庁 (EP)
 (31) 優先権主張番号 13177350.9
 (32) 優先日 平成25年7月22日 (2013. 7. 22)
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 500341779
 フラウンホーファー・ゲゼルシャフト・ツ
 ール・フェルデルング・デル・アンゲヴァ
 ンテン・フォルシュング・アインゲトラ
 ゲネル・フェライン
 ドイツ連邦共和国, 80686 ミュンヘ
 ン, ハンザシュトラッセ 27 ツェー
 (74) 代理人 110000121
 アイアット国際特許業務法人
 (72) 発明者 ディッシュ, ザッシャ
 ドイツ連邦共和国, 90766 フュルト
 , ヴィルヘルムシュトラッセ 70
 (72) 発明者 ガイガー, ラルフ
 ドイツ連邦共和国, 91052 エアラン
 ゲン, ヤコブ・ヘルツ・ヴェク 36
 最終頁に続く

(54) 【発明の名称】 符号化オーディオ信号の復号装置、方法およびコンピュータプログラム

(57) 【特許請求の範囲】

【請求項 1】

符号化コア信号とパラメトリックデータとを含む符号化オーディオ信号を復号する復号装置において、

前記符号化コア信号を復号して復号化コア信号を得るコアデコーダ (600) と、

周波数再生動作の前または後に前記復号化コア信号を解析して解析結果 (603) を得るアナライザ (602, 614) と、

前記復号化コア信号のスペクトル部、前記パラメトリックデータ (605) および前記解析結果 (603) を使用して、前記復号化コア信号に含まれないスペクトル部を再生成する周波数再生器 (604) と、

を備え、

前記アナライザ (602, 614) は、

前記復号化コア信号のスペクトル部内の前記復号化コア信号の周波数境界でのピーク部の分離、

再生信号内の、前記復号化コア信号の周波数境界、または前記復号化コア信号の同じまたは異なる部分を用いて生成された2つの再生スペクトル部の間の周波数境界でのピーク部の分離、あるいは、

前記再生信号の最大周波数境界でのピーク部の分離

を検出する構成であり、

前記周波数再生器 (604) は、

10

20

前記復号化コア信号と前記再生信号との間の周波数境界、
前記復号化コア信号の同じまたは異なる部分を用いて生成された2つの再生スペクトル部の間の周波数境界、または、
最大周波数境界
を変更して、前記ピーク部の分離を削減または除去する構成であることを特徴とする復号装置。

【請求項2】

請求項1に記載の復号装置において、
前記アナライザ(614)は、周波数再生の前または後の前記復号化コア信号を解析して、前記復号化コア信号内に1以上の局所スペクトル極小を見つけ、その局所スペクトル極小を識別する解析結果(603)を提供する構成であり、
前記周波数再生器(604)は、スペクトル部を再生し、前記復号化コア信号または前記再生スペクトル部内の1以上の周波数境界が1以上のスペクトル極小に設定される構成であることを特徴とする復号装置。

10

【請求項3】

請求項1または2に記載の復号装置において、
前記周波数再生器(604)は、前記復号化コア信号の少なくとも1つのスペクトル部を用いて前記復号化コア信号内に含まれないスペクトル部を有する予備再生信号(703)を再生して、再生スペクトル部を得る構成であり、
前記周波数再生器(604)は、前記復号化コア信号と前記再生スペクトル部との間の周波数境界の近くの、または前記周波数再生器(604)を用いて生成された2つの再生スペクトル部または前記復号化コア信号の異なるスペクトル部の間の周波数境界の近くの、アーティファクトが作り出した信号部を検出する構成であり、
前記周波数再生器(604)はさらに、前記予備再生信号を生成するために使用する制御データとは異なる操作された制御データを使用して再生信号を新たに生成するため、前記予備再生信号または制御データを操作するマニピュレータ(824)を有することを特徴とする復号装置。

20

【請求項4】

請求項3に記載の復号装置において、
前記周波数再生器(604)は、前記復号化コア信号の1以上のスペクトル部を用いて前記スペクトル部を引き出し、生のスペクトル部を得る周波数タイル生成器(820)を有し、
前記マニピュレータ(824)は、前記生のスペクトル部または前記周波数タイル生成器(820)を操作して、操作されたスペクトル部を得るように構成され、
前記周波数再生器(604)はさらに、前記パラメトリックデータ(605)を用いて前記操作されたスペクトル部の包絡線調整を行うように構成されたスペクトル包絡線調整器(826)を有することを特徴とする復号装置。

30

【請求項5】

請求項1から4のいずれか1項に記載の復号装置において、
前記アナライザ(602)は、周波数検出範囲内にあるトーン信号部を検出する構成であり、その周波数検出範囲は、再構築範囲の周波数境界から所定の帯域幅だけ延びる、または再構築範囲内の隣接周波数タイルの間の、所定の帯域幅の範囲であり、この所定の帯域幅は、ソース周波数レンジまたは再構築周波数レンジの帯域幅の20%未満、または1バークである
ことを特徴とする復号装置。

40

【請求項6】

請求項3または4に記載の復号装置において、
前記アナライザ(602)は、周波数検出範囲内にあるトーン信号部を検出する構成で

50

あり、その周波数検出範囲は、再構築範囲の周波数境界から所定の帯域幅だけ延びる、または再構築範囲内の隣接周波数タイルの間の、所定の帯域幅の範囲であり、この所定の帯域幅は、ソース周波数レンジまたは再構築周波数レンジの帯域幅の20%未満、または1バークであり、

前記マニピュレータ(824)は、前記所定の帯域幅内で、前記再生信号内のトーン部を含むスペクトル部を減衰させる、または取り除く(708)ように構成されていることを特徴とする復号装置。

【請求項7】

請求項6に記載の復号装置において、前記マニピュレータ(824)は、トーン部(802)の開始周波数にある開始スペクトル部と、前記トーン部(802)の終了周波数にある終了スペクトル部と、を判定し、前記開始周波数と前記終了周波数との間に内挿(804)して内挿信号部(806)を取得し、前記開始周波数と前記終了周波数との間のトーン部(802)を前記内挿信号部(806)で置き替えることを特徴とする復号装置。

【請求項8】

請求項6に記載の復号装置において、前記マニピュレータ(824)は、ランダムに、または非ランダムに、前記復号化コア信号の非トーン信号部または前記再生スペクトル部(810)の非トーン信号部により決定されるエネルギーを有するスペクトル線(808)を生成することを特徴とする復号装置。

【請求項9】

請求項3から8のいずれか1項に記載の復号装置において、前記アナライザ(602, 614)は、ある特定の周波数のアーティファクトが作り出した信号部を検出する構成であり、

前記マニピュレータ(824)は、周波数タイル生成器(820)が前記復号化コア信号のスペクトル部の周波数境界または再生スペクトル部の周波数境界を変化させるように前記周波数タイル生成器(820)を制御して、アーティファクトが作り出した信号部がアーティファクトをほとんど、または全く、引き起こさないようにする

ことを特徴とする復号装置。

【請求項10】

請求項1から9のいずれか1項に記載の復号装置において、前記周波数再生器(604)は、前記復号化コア信号の同じまたは異なるスペクトル部を用いて、前記復号化コア信号の第1スペクトル部に対する第1周波数タイルおよび第2スペクトル部に対する第2周波数タイルを生成する周波数タイル生成器(820)を有し、

前記第2周波数タイルの下側周波数境界は前記第1周波数タイルの上側境界は一致しており、

前記アナライザ(602)は、ピークスペクトル部が前記第2周波数タイルの下側周波数境界によりクリップされているか、あるいは前記第1周波数タイルの上側周波数境界によりクリップされているか、あるいは前記第1周波数タイルと前記復号化コア信号の所定のギャップ充填開始周波数(309)との下側周波数境界によりクリップされているかを検出する構成であり、

前記マニピュレータ(824)は、クリッピングが削減または除去されるように修正された開始または停止周波数境界を有する修正周波数タイルを生成するように、前記周波数タイル生成器(820)を制御する構成である

ことを特徴とする復号装置。

【請求項11】

請求項1から10のいずれか1項に記載の復号装置において、前記コアデコーダ(600)は、ゼロ表現とは異なるスペクトル値により表現される第1セットの第1スペクトル部を含む周波数領域復号化スペクトル部を得る構成であり、第2セットの第2スペクトル部はスペクトル値に対してゼロ表現により表現され、パラメトリック情報が前記第2セットの前記第2スペクトル部に対して提供され、

前記周波数再生生成器（604）は、前記第1セットのスペクトル部から復号化されたスペクトル部を用いて、前記第1セットのスペクトル部に含まれない再構築帯域内のスペクトル部を再生するように構成され、

前記復号装置はさらに、再生されたスペクトル部と前記復号化コア信号のスペクトル部とを時間表現に変換する周波数時間変換器（828）を備える

ことを特徴とする復号装置。

【請求項12】

請求項11に記載の復号装置において、

前記コアデコーダ（600）は、修正離散コサイン変換（MDC T）スペクトル値を出力する構成であり、

前記周波数時間変換器（828）は、オーバーラップ加算処理をその後に得られるMDC Tフレームに加える逆MDC T変換（512，514，516）を実行するためのプロセッサを備える

ことを特徴とする復号装置。

【請求項13】

請求項1から12のいずれか1項に記載の復号装置において、

前記周波数再生生成器（604）は、予備再生された信号を生成（702）する構成であり、

前記周波数再生生成器（604）は、前記予備再生された信号内のトーン成分を検出（704）する構成であり、

前記周波数再生生成器（604）は、ソースレンジと再構成レンジとの間、または再構成されたレンジ内の隣接周波数タイル間の遷移周波数を、前記検出（704）の結果に基づいて調整し、再生信号を生成する構成であり、

前記周波数再生生成器（604）はさらに、前記遷移周波数の周りの検出レンジ内にあるトーン成分を除去する（708）構成であり、

前記周波数再生生成器（604）はさらに、前記パラメトリックデータ（605）を用いるクロスフィルタの結果をスペクトル包絡線整形するスペクトル波形整形器（712）を有する

ことを特徴とする復号装置。

【請求項14】

符号化コア信号とパラメトリックデータとを含む符号化オーディオ信号を復号する方法において、

前記符号化コア信号を復号（600）して復号化コア信号を取得し、

周波数再生操作を実行する前または後に前記復号化コア信号を解析（602）して解析結果（603）を提供し、

前記復号化コア信号のスペクトル部、前記パラメトリックデータ（605）、および前記解析結果（603）を用いて、前記復号化コア信号に含まれないスペクトル部を再生（604）し、

前記解析（602）では、前記復号化コア信号のスペクトル部内の前記復号化コア信号の周波数境界でのピーク部の分離、再生信号内の、前記復号化コア信号の周波数境界、または前記復号化コア信号の同じまたは異なる部分を用いて生成された2つの再生スペクトル部の間の周波数境界でのピーク部の分離、あるいは、前記再生信号の最大周波数境界でのピーク部の分離を検出し、

前記再生（604）では、前記復号化コア信号と前記再生信号との間の周波数境界、前記復号化コア信号の同じまたは異なる部分を用いて生成された2つの再生スペクトル部の間の周波数境界、または、最大周波数境界を変更して、前記ピーク部の分離を削減または除去する

復号方法。

【発明の詳細な説明】

【技術分野】

10

20

30

40

50

【 0 0 0 1 】

本発明は、オーディオ符号化および復号化に関し、特に、I G F (I n t e l l i g e n t g a p f i l l i n g) を用いるオーディオ符号化に関する。

【 背景技術 】

【 0 0 0 2 】

オーディオ符号化は信号圧縮の一分野であり、音響心理学の知識を用いて、オーディオ信号における冗長性および非関連性を利用する処理を行う。今日において、オーディオコーデックは、典型的には、ほぼ全ての種のオーディオ信号の知覚的に透明な符号化のために、チャンネルあたり 6 0 k b p s 程度を要する。さらに新しいコーデックは、帯域幅拡張 (B W E) などの技術を用いて、信号におけるスペクトルの類似性を活用して、符号化ビットレートの削減を試みている。ある B W E スキームにおいては、オーディオ信号の高周波数 (H F) 成分を示すよう設定された低ビットレートパラメータが用いられる。H F スペクトルは、低周波数 (L F) 域からのスペクトルコンテンツによって充足され、スペクトルの形状、傾きおよび時間的連続性が、オリジナル信号の音質および色彩を保つよう調整される。そのような B W E 手法によって、オーディオコーデックは、チャンネルあたり 2 4 k b p s 程度の低ビットレートであっても、良好な品質を保つことができる。

10

【 0 0 0 3 】

独創的なオーディオ符号化システムは、広範囲のビットレートにより、任意のオーディオ信号を効率的に符号化する。ここで、高ビットレートにおいては、独創的なシステムは、透明性に集中する一方、低ビットレートにおいては、知覚的な不快感が最小限化される。したがって、利用可能なビットレートの主要な部分は、エンコーダにおいて、知覚的にもっとも適切な信号構造に波形符号化することに用いられる。そして、その結果としてのスペクトルギャップは、デコーダにおいて、元来のスペクトルに大まかに近接した信号コンテンツによって充填される。エンコーダからデコーダに対して送信される専用のサイド情報によって、相当に限られたビット配分のみが、パラメータ駆動されるいわゆるスペクトルインテリジェントギャップ充填 (I G F) を制御するのに消費される。

20

【 0 0 0 4 】

オーディオ信号の蓄積または送信については、厳格なビットレート制限が適用されることがよくある。過去においては、極めて低ビットレートしか利用可能でない場合には、符号化する側において、強制的に、送信オーディオ帯域幅を大幅に低減させられていた。

30

【 0 0 0 5 】

今日における近代オーディオコーデックは、帯域幅拡張 (B W E) 法を用いて、広帯域信号を符号化できる (非特許文献 1)。これらのアルゴリズムは、高周波数コンテンツ (H F) のパラメータ表現に依拠している。このパラメータ表現は、復号信号の波形符号化された低周波数部 (L F) から、H F スペクトル領域に移転 (「パッチング」) させて、パラメータ駆動の後処理を行うことにより、生成される。B W E スキームにおいては、所与のいわゆるクロスオーバー周波数の上への H F スペクトル領域の再構築が、しばしば、スペクトルパッチングに基づいて行われる。一般的に、H F 領域は、複数の近接パッチから構成され、これらのパッチのそれぞれは、所与のクロスオーバー周波数より下の L F スペクトルのバンドパス (B P) 領域から供給される。最新のシステムは、例えば、近接するサブバンド係数一式をソースから対象領域にコピーすることによって、直交ミラーフィルタバンク (Q M F) などのフィルタバンク表現内において、パッチングを効率的に実施する。

40

【 0 0 0 6 】

今日のオーディオコーデックに見られる、圧縮効率を向上させかつ低ビットレートでオーディオ帯域幅を拡張する別の技術としては、オーディオスペクトルの適宜な部分を、パラメータ駆動により合成置換する技術が挙げられる。例えば、オリジナルのオーディオ信号におけるノイズ様の信号部分は、デコーダ内で生成され副情報パラメータによりスケールリングされた人工ノイズによって、主観的品質を実質的に損なうことなく置換される。一例としては、M P E G - 4 アドバンストオーディオコーディング (A A C) に含まれる、

50

知覚的ノイズ置換（PNS）ツールが挙げられる（非特許文献5）。

【0007】

低ビットレートでオーディオ帯域幅を拡張する他の規定としては、MPEG-Dの音声音響統合符号化方式（USAC）に含まれる、ノイズ充填技術が挙げられる（非特許文献7）。粗すぎる量子化による量子化装置のデッドゾーンから推測されるスペクトルギャップ（ゼロ）は、デコーダにおける人工ノイズにより事後に充填され、パラメータによって決定される事後処理によってスケージングされる。

【0008】

別の最新システムは、精密スペクトル置換（ASR）と称される（非特許文献2-4）。波形コーデックに加えて、ASRは、専用の信号合成段階を採用し、この信号合成段階によって、信号内の知覚的に重要な正弦部分が、デコーダにより復元される。また、非特許文献5に記載のシステムは、波形符号器のHF領域における正弦モデリングに依拠して、低ビットレートでも相当の知覚品質を有する拡張オーディオ帯域幅を実現している。これらの方法は全て、変形離散コサイン変換（MDCT）とは別の第2領域へのデータ変換を含み、また、HF正弦成分の保存のためのかなり複雑な解析/合成段階を含む。

【0009】

図13Aは、例えば、高効率アドバンストオーディオコーディング（HE-AAC）において使用されている帯域幅拡張技術用の、オーディオエンコーダの概略図である。ライン1300上のオーディオ信号は、ローパス1302およびハイパス1304からなるフィルタシステムに入力される。ハイパスフィルタ1304から出力された信号は、パラメータ抽出器/符号器1306に入力される。パラメータ抽出器/符号器1306は、例えば、スペクトル包絡線パラメータやノイズ付加パラメータ、欠落高調波パラメータ、逆フィルタリングパラメータなどのパラメータを計算し、符号化するように構成される。これらの抽出パラメータは、ビットストリーム多重変換器1308に入力される。ローパス出力信号は、ダウンサンプラ1310およびコア符号器1312の機能を一般的に有する処理器に入力される。ローパス1302は、帯域幅を制限して、ライン1300上のオリジナル入力オーディオ信号よりも、相当に狭い帯域幅に符号化されるようにする。コア符号器において生じる全ての機能が、削減された帯域の信号上で機能すれば良いという事実により、かなりの符号化利得が得られる。例えば、ライン1300上のオーディオ信号の帯域幅が20キロヘルツであり、かつローパスフィルタ1302が仮に4キロヘルツの帯域幅を有しているとする、サンプリング定理を満たすためには、ダウンサンプラの後段の信号が8キロヘルツのサンプリング周波数を有することで理論上は十分であり、これは、オーディオ信号1300において必要となる少なくとも40キロヘルツのサンプリングレートの、実質的削減である。

【0010】

図13Bは、対応する帯域幅拡張デコーダの概略図である。デコーダは、ビットストリーム多重変換器1320を有する。ビットストリーム多重分離器1320は、コアデコーダ1322用の入力信号、およびパラメトリックデコーダ1324用の入力信号を抽出する。コアデコーダ出力信号は、上述の例においては、8キロヘルツのサンプリングレートを有し、したがって、4キロヘルツの帯域幅を有する。一方、完全な帯域幅再構築のためには、高周波数再構築器1330の出力信号は、少なくとも40キロヘルツのサンプリングレートを要する20キロヘルツでなければならない。これを可能にするためには、デコーダ処理器が、アップサンプラ1325およびフィルタバンク1326の機能を有することが必要となる。そして高周波数再構築器1330は、フィルタバンク1326より出力され周波数解析された低周波数信号を受信して、図13Aのハイパスフィルタ1304によって定義された周波数範囲を、周波数帯域のパラメータ表現を用いて再構築する。高周波数再構築器1330は、低周波数範囲におけるソース範囲を利用して上位の周波数範囲を再生成したり、スペクトル包絡線を調整したり、ノイズを付与するなど、様々な機能を有している。また、高周波数再構築器1330には、欠落高調波を上位の周波数範囲に導入する機能や、図13Aの符号器において適用され算出された場合には逆フィルタを行い

10

20

30

40

50

、高周波数範囲が一般的には低周波数範囲ほど音調がないという事実に対処する機能がある。H E - A A Cにおいては、デコーダ側において欠落高調波が再合成され、再構築帯域のちょうど真ん中に配置される。したがって、所定の再構築帯域において決定された全ての欠落高調波線は、オリジナル信号で配置されていた周波数値には配置されない。その代わりに、これらの欠落高調波線は、所定帯域の中心周波数に配置される。よって、オリジナル信号中の欠落高調波線が、オリジナル信号中の再構築帯域の境界に非常に近接して配置されると、この欠落高調波線を再構築信号において帯域の中心に配置することにより生じる周波数エラーは、個別の再構築帯域の50%近くになり、これについてパラメータが生成され送られる。

【0011】

10

さらに、一般的なオーディオコア符号器は、スペクトル領域で機能するにも関わらず、コアデコーダはなお時間領域信号を生成し、この時間領域信号が、再びフィルタバンク1326機能によりスペクトル領域に変換される。これによって追加的な処理遅延が発生し、またまずスペクトル領域から周波数領域に変形し、そして一般的には再び異なる周波数領域へと変形する直列処理によって、アーティファクトが発生する恐れがある。もちろん、これによって、計算が相当程度複雑になり、相当量の電力が必要となる。この電力の問題は、帯域幅拡張技術が携帯電話やタブレット、ノートパソコンなどの携帯機器に適用された場合、特に問題となる。

【0012】

20

現在のオーディオコーデックは、符号化スキームの不可分の一部として、B W Eを用いて低ビットレートオーディオ符号化を実施する。しかしながら、B W E技術は、高周波数(H F)コンテンツのみを置換することに限定されている。さらに、B W E技術は、所定のクロスオーバー周波数より上の、知覚的に重要なコンテンツを、波形符号化することはできない。したがって、現代のオーディオコーデックは、信号の音の高調波の正確な配列がほとんどのシステムにおいて考慮されないため、B W Eが実行されると、H Fの詳細または音質のいずれかが喪失される。

【0013】

B W Eシステムの現代技術に関する別の課題としては、オーディオ信号を、B W Eを実行するための新たな領域に変換する必要があることである(例: M D C TからQ M F領域への変換)。これによって、同期性の煩雑化、追加的計算の煩雑性およびメモリ要件の増加などにつながる。

30

【0014】

オーディオ信号の蓄積または送信については、厳格なビットレート制限が適用されることが多い。過去において、符号器は、相当な低ビットレートしか利用可能でない場合には、送信オーディオ帯域を徹底的に削減することを強制されていた。今日において、近代的なオーディオコーデックは、帯域幅拡張(B W E)方法を使って、広帯域信号を符号化することができる(非特許文献1, 2)。これらのアルゴリズムは、H Fスペクトル領域への転換(「パッチング」)およびパラメータ駆動の後処理によって復号信号の波形符号化された低周波数部分(L F)から生成される、高周波数コンテンツ(H F)のパラメータ表現に依存している。

40

【0015】

B W Eスキームにおいて、いわゆる所定のクロスオーバー周波数より上のH Fスペクトル領域は、スペクトルパッチングに基づき再構築されることが多い。スペクトルギャップを充填する他のスキーム、例えばインテリジェントギャップ充填(I G F)は、隣接するいわゆるスペクトルタイルを使用して、オーディオ信号H Fスペクトル部分を再生する。典型的に、H F領域は、複数の近接パッチから構成され、これらの各パッチは、所定のクロスオーバー周波数よりも下のL Fスペクトルのバンドパス(B P)から供給される。従来のシステムは、近接するサブバンド係数のセットをソースから対象領域にコピーすることにより、フィルタバンク表現内において、効率的にパッチング、すなわちタイル張り、を実施する。ただし、信号コンテンツによっては、L F帯域と隣接するH F帯域内のパッ

50

チから再構成された信号の組合せが、ピーティング、不協和音と、および聴覚的粗さを引き起こすことがある。

【 0 0 1 6 】

そこで、特許文献 1 には、不協和音ガード帯域フィルタリングのコンセプトが、フィルタバンクを基礎とする B W E システムの文脈で提示されている。L F と B W E 再生された H F との間のクロスオーバー周波数で約 1 バーク帯域のノッチフィルタを有効に適用して、不協和音の可能性を除去し、スペクトルコンテンツをゼロまたは雑音に置き換えることが示唆されている。

【 0 0 1 7 】

しかし、特許文献 1 に提案された解法は、いくつかの欠点を有している。第一に、ゼロまたは雑音でスペクトルコンテンツの厳格な置き換えは、信号の知覚品質も損なう可能性がある。さらに、提案された処理は、信号適応的ではなく、そのため、ある場合には知覚品質を害する可能性がある。例えば、信号コンテンツが遷移する場合、これが、前または後エコーを引き起こす可能性がある。

【 0 0 1 8 】

第二に、不協和音はまた、連続的な H F パッチの間の遷移でも生じる可能性がある。特許文献 1 で提案された解決策は、単に機能的に、L F と B W E 再生された H F との間のクロスオーバー周波数で発生する不協和音を是正するだけである。

【 0 0 1 9 】

最後に、特許文献 1 に提案されたようなフィルタバンクを基本とするシステムとは反対に、B W E システムは、修正離散コサイン変換 (M D C T) のような変換に基づく実装でも実現される。M D C T のような変換は、「さえずり w a r b l i n g 」(非特許文献 1 9) あるいはリングングといったアーティファクト (人工的な音) を生じる傾向がある。このようなアーティファクトは、スペクトル係数の帯域通過領域がコピーされる場合、またはスペクトル係数が特許文献 1 のようにゼロに設定される場合に生じるものである。

【 0 0 2 0 】

特に、特許文献 1 は、フィルタバンクを基本とする変換または畳み込みにおいて、ゼロに設定された 1 または幾つかのサブバンドチャンネルで構成されて挿入される「ガード帯域」の使用を開示している。ある数のフィルタバンクチャンネルがガード帯域として使用され、ひとつのガード帯域の帯域幅は 0 . 5 バークである。これらの不協和音ガード帯域は、ランダム白雑音信号を用いて、すなわちサブバンドにゼロの代わりに白雑音を供給して、部分的に再構成される。ガード帯域は、処理される現在の信号に関わりなく挿入される。

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 2 1 】

本発明は、符号化オーディオ信号を復号する改善された概念を提供することを目的とする。

【 課題を解決するための手段 】

【 0 0 2 2 】

この目的は、請求項 1 に記載の符号化オーディオ信号を復号する復号装置、請求項 1 6 に記載の符号化オーディオ信号を復号する復号方法または請求項 1 9 に記載のコンピュータプログラムによって達成される。

【 0 0 2 3 】

本発明によると、アナライザを用いるデコーダ側の信号解析が、周波数再生の実行の前または後に、復号化コア信号を解析するために行われ、解析結果が提供される。このとき、この解析結果は、周波数再生器により、復号化コア信号に含まれないスペクトル部を再生するために用いられる。

【 0 0 2 4 】

したがって、パッチングまたは周波数タイリングが固定されて実行される場合、すなわち、ある特定のソースレンジがコア信号からとり出され、ある特定の固定された周波数境

10

20

30

40

50

界が、ソースレンジと再構築レンジとの間の周波数の設定、または再構築レンジ内の2つの隣接する周波数パッチまたはタイルの間の周波数境界の設定のいずれかに適用される固定されたデコーダ設定とは異なり、信号依存パッチングまたはタイリングが実行され、このとき、コアレンジは、コアレンジの周波数境界がコア信号スペクトルの局所極小に一致するように選択される。

【0025】

これとは別に、あるいは追加的に、信号解析が、予備再生信号または予備周波数パッチまたはタイルされた信号で実行されることもできる。ここで、予備周波数再生手続きの後で、コアレンジと再構築レンジの境界が解析され、互いに非常に近接して再構築されたときにうなりアーティファクトを生じる問題のあるトーン部のような、あらゆるアーティファクトが作り出した信号部を検出する。これとは別に、あるいは追加的に、境界を解析して、トーン部の途中でのクリッピングを検出することもできる。このようなトーン部のクリッピングは、そのまま再構築されたときに、アーティファクトを作り出すかもしれない。これらの手続きを避けるため、再構築レンジおよび/またはソースレンジおよび/または再構築レンジ内の2つの個別の周波数タイルまたはパッチの間の周波数境界を信号マニピュレータにより修正し、再度、新しく設定された境界で再構築を実行することもできる。

【0026】

追加的に、または代替として、周波数再生が解析結果に基づく再生成であり、周波数境界がそのまま残り、ソースレンジと再構築レンジとの間、または再構築レンジ内の2つの個別の周波数タイルまたはパッチの間の周波数境界に近接する問題のあるトーン部の除去または少なくとも減衰が行われる。このようなトーン部は近接したトーン部であり、アーティファクトを生じるか、あるいは途中でクリップされたトーン部である。

【0027】

特に、MDC Tのような非エネルギー保存変換を用いる場合、単一トーンは、直接には単一スペクトル線にマッピングされない。その代わりに、単一トーンは、トーンの位相に依存するある特定強度の一群のスペクトル線にマッピングされる。パッチング動作がこのトーン部をクリップすると、そのとき、MDC T再構築器内のような完全な再構築が適用されたても、再構築後にアーティファクトが生じる。これは、MDC T再構築器が、最終的に正しいトーンを再構築するためには、そのトーンに対する完全なトーンパターンを必要とするからである。クリッピングが前に行われていることから、完全なトーンパターンをもちや可能ではなく、したがって、時間変化するさえずりアーティファクトが作り出される。本発明による解析に基づいて、周波数再生器は、アーティファクトを作り出す完全なトーン部を減衰させることにより、あるいは、上述したように、対応する境界周波数を変更することにより、またはそれらの双方を適用することにより、またはそのようなトーンパターンについての特定の予見に基づいてクリップされた部分を再構築することによって、上述のような状況を取り除く。

【0028】

付加的に、または代替として、クロスオーバーフィルタリングを、符号化コア信号と、ギャップ充填周波数から最初のタイル停止周波数に延びる周波数を有する最初の周波数タイルとを、スペクトル上でクロスオーバーフィルタリングするため、あるいは最初の周波数タイルと第2の周波数タイルとをスペクトル上でクロスオーバーフィルタリングするため用いることができる。

【0029】

このクロスオーバーフィルタリングは、いわゆるフィルタリングの削減に有用である。

【0030】

本発明のアプローチは、主に、MDC Tのような変換に基づくBWEで適用される。

【0031】

本発明のアプローチは、聴覚的な粗さ、うねりおよび不調和は、遷移点（クロスオーバ

10

20

30

40

50

一周波数またはパッチ境界のような)に近接しているスペクトル領域の信号内容がまさしくトーンである場合にのみ生じる、という観測に基づいている。したがって、従来見出されていた問題点に対して、ここで提案された解決手段は、遷移領域のトーン成分の信号適応的検出と、それに続くそれらの成分の減衰または除去とからなる。それらの成分の減衰または除去は、望ましくは、そのような成分の足から足へのスペクトル内挿により、または代替的にはゼロまたは雑音の内挿により、完遂される。これとは別に、遷移アーティファクトが最小化されるように、遷移のスペクトル位置を信号適応的に選択することもできる。

【0032】

さらに、この技術は、フィルタリングの減少または除去に使用できる。特に過渡的な信号に対して、リングングは可聴の耳に障るアーティファクトである。フィルタリングアーティファクトは、遷移帯域のいわゆるレンガ壁特性(カットオフ周波数における通過帯域から阻止帯域への急峻な遷移)により生じる。このようなフィルタは、時間周波数変換の周波数領域で1つの係数または係数のグループをゼロに設定することにより、効果的に実現される。そのため、BWEの場合、出願人は、リングング効果を削減するため、クロスオーバーフィルタを、パッチ間、またはコア帯域と最初のパッチとの間の各遷移周波数に適用することを提案する。クロスオーバーフィルタは、適切な利得関数を採用して、変換領域にスペクトル重み付けを行うことにより実現される。

【0033】

本発明のさらなる見地によると、符号化オーディオ信号を復号する復号装置は、コアデコーダと、符号化コア信号のスペクトル部を用いて、符号化コア信号に含まれない周波数をもつ1以上のスペクトルタイルを生成するタイル生成器と、符号化コア信号と、ギャップ充填周波数から最初のタイル停止周波数に延びる周波数を有する最初の周波数タイルとを、スペクトル上でクロスオーバーフィルタリングするため、あるいは、ひとつの周波数タイルと、その周波数タイルの上側境界周波数に周波数上で隣接する下側境界周波数を有するさらなる周波数タイルとを、スペクトル上でクロスオーバーフィルタリングするクロスオーバーフィルタと備える。

【0034】

望ましくは、この手続きは、MDCTのような変換に基づく帯域幅拡張で適用することが意図されている。しかしながら、本発明は、一般に適用でき、特に、交ミラーフィルタバンク(QMF)に頼る帯域幅拡張シナリオ、特に、システムが決定的に標準化される場合、例えば時間周波数変換または周波数時間変換として実値QMF表現がある場合に、適用できる。

【0035】

本発明の実施は、過渡的な信号に対して特に有用である。その理由は、そのような過渡的な信号に対しては、リングングが可聴で耳障りなアーティファクトだからである。フィルタリングアーティファクトは、遷移帯域におけるフィルタのいわゆるレンガ壁特性、すなわちカットオフ周波数での通過帯域から阻止帯域への急峻な遷移により生じる。このようなフィルタは、時間周波数変換の周波数領域で1つの係数または係数のグループをゼロに設定することにより、効果的に実現される。したがって、本発明は、このリングングアーティファクトを削減するため、パッチ/タイル間、またはコア帯域と最初のパッチ/タイルとの間の各遷移周波数のクロスオーバーフィルタに依存する。クロスオーバーフィルタは、変換領域での適当な利得関数を採用するスペクトル重み付けにより実現されることが望ましい。

【0036】

望ましくは、クロスオーバーフィルタは、信号適応的であり、2つのフィルタ、すなわち、低スペクトル領域に適用されるフェードアウトフィルタと、高スペクトル領域に適用されるフェードインフィルタとを含む。

【0037】

さらなる態様として、周波数タイルまたは周波数パッチは、クロスオーバーフルタリン

10

20

30

40

50

グを目的とするものだけでなく、タイル生成器は、望ましくは、クロスオーバーフィルタリングを実行する前に、周波数境界のスペクトル極小への設定を含むパッチ適応、および遷移周波数の周囲の遷移領域に残るトーン部の除去または減衰を実行する。

【0038】

本発明の好適な実施形態を、以下、添付図面を参照して説明する。

【図面の簡単な説明】

【0039】

【図1A】図1Aは、オーディオ信号を符号化する装置を示す図である。

【図1B】図1Bは、図1Aのエンコーダと整合する符号化されたオーディオ信号を復号するデコーダを示す図である。

10

【図2A】図2Aは、デコーダの好適な実施形態を示す図である。

【図2B】図2Bは、エンコーダの好適な実施形態を示す図である。

【図3A】図3Aは、図1Bのスペクトル領域デコーダによって生成されたスペクトル表現を示す概略図である。

【図3B】図3Bは、換算係数帯域用の換算係数と、再構築帯域用のエネルギーと、ノイズ充填帯域用のノイズ充填情報との関係を示す表である。

【図4A】図4Aは、スペクトル部の選択を第1および第2セットのスペクトル部に対して適用するスペクトル領域エンコーダの機能を示す図である。

【図4B】図4Bは、図4Aの機能の実施形態を示す図である。

【図5A】図5Aは、MDCTエンコーダの機能を示す図である。

20

【図5B】図5Bは、MDCT技術とデコーダの機能を示す図である。

【図5C】図5Cは、周波数再生器の実施形態を示す図である。

【図6A】図6Aは、本発明の一実施形態により符号化オーディオ信号復号装置を示す図である。

【図6B】図6Bは、符号化オーディオ信号復号装置のさらなる実施形態を示す図である。

【図7A】図7Aは、図6Aまたは図6Bの周波数再生器の望ましい実施形態を示す図である。

【図7B】図7Bは、アナライザと周波数再生器の協働のさらなる実施形態を示す図である。

30

【図8A】図8Aは、周波数再生器のさらなる実施形態を示す図である。

【図8B】図8Bは、本発明のさらなる実施形態を示す図である。

【図9A】図9Aは、エネルギー値を再生周波数範囲に使用する周波数生成技術を有するデコーダを示す図である。

【図9B】図9Bは、図9Aの周波数再生器のさらに詳細なる実施形態を示す図である。

【図9C】図9Cは、図9Bの機能を概略的に示す図である。

【図9D】図9Dは、図9Aのデコーダのさらなる実施形態を示す図である。

【図10A】図10Aは、図9Aのデコーダと整合するエンコーダのブロック図である。

【図10B】図10Bは、図10Aのパラメータ計算器のさらなる機能の実施形態を示すブロック図である。

40

【図10C】図10Cは、図10Aのパラメータ計算器のさらなる機能の実施形態を示すブロック図である。

【図10D】図10Dは、図10Aのパラメータ計算器のさらなる機能の実施形態を示すブロック図である。

【図11A】図11Aは、遷移の周囲のフィルタリングのスペクトル示す図である。

【図11B】図11Bは、帯域幅拡張を適用した後の遷移のスペクトル図を示す。

【図11C】図11Cは、フィルタリング削減を伴う帯域幅拡張を適用した後の遷移のスペクトル図を示す。

【図12A】図12Aは、符号化オーディオ信号復号装置のブロック図である。

【図12B】図12Bは、トーン信号、パッチ/タイル適用なしのコピー、変化した周波

50

数境界のコピー、およびアーティファクトが作り出したトーン部の付加的除去の強度スペクトル（様式化された）を示す図である。

【図 1 2 C】図 1 2 C は、クロスフェード関数の一例を示す図である。

【図 1 3 A】図 1 3 A は、帯域幅拡張を有する従来の符号器を示す図である。

【図 1 3 B】図 1 3 B は、帯域幅拡張を有する従来のデコーダを示す図である。

【図 1 4 A】図 1 4 A は、クロスオーバーフィルタを用いるさらなる符号化オーディオ信号復号装置を示す図である。

【図 1 4 B】図 1 4 B は、典型的クロスオーバーフィルタのより詳細を説明する図である。

【発明を実施するための形態】

10

【0040】

図 6 A は、符号化コア信号とパラメトリックデータとを含む符号化オーディオ信号を復号する装置を示す。この装置は、符号化コア信号を復号して復号化コア信号を得るコアデコーダ 600 と、周波数再生動作を実行する前または後に復号化コア信号を解析するアナライザ 602 とを備える。アナライザ 602 は、解析結果 603 を出力するように構成される。周波数再生器 604 は、復号化コア信号のスペクトル部分、消失したスペクトル部に対する包絡線データ 605 および解析結果 603 を使用して、復号化コア信号に含まれないスペクトル部分を再生成する。したがって、従来のものと比較すると、周波数再生成が、デコーダ側の信号に依存せずに実行されるのではなく、信号依存で実行される。これは、何も問題無い場合には、周波数再生成がそのまま実行されるが、問題となる信号部

20

【0041】

望ましくは、コアデコーダ 600 は、図 6 B に示すように、エントロピー（例えばハフマンまたは算術デコーダ）復号化・逆量子化段 612 として実現される。コアデコーダ 600 はコア信号スペクトルを出力し、このスペクトルは、スペクトルアナライザ 614 により解析される。図 6 A のアナライザ 602 は、時間領域の信号も解析できるアナライザであるが、そのような任意のアナライザではなくスペクトルアナライザであってもよく、その場合は、スペクトルアナライザ 614 は図 6 A のアナライザ 602 と全く同じになる。図 6 B の実施形態では、スペクトルアナライザがスペクトル信号を解析するように構成され、これにより、ソース帯域および/またはターゲット帯域内の、すなわち周波数パッチまたは周波数タイル内の、局所極小が判定される。周波数再生器 604 は、参照番号 616 として示すように周波数再生成を実行し、パッチ境界が、ソース帯域および/またはターゲット帯域の極小に置かれる。

30

【0042】

続いて、図 7 A を参照して、図 6 A の周波数再生器 604 の望ましい実施例を説明する。予備信号再生器 702 は、入力として、ソース帯域からのソースデータと、追加的に、予備境界周波数のような予備パッチ情報とを受け取る。その結果、予備再生成信号 703 が生成され、これが、予備再生成信号 703 内のトーン成分を検出するための検出器 704 により検出される。これとは別に、あるいは付加的に、ソースデータ 705 は、図 6 A のアナライザ 602 に相当する検出器により解析されてもよい。このとき、予備信号再生成ステップは必要ないかもしれない。ソースデータから再構築データへの明確に定義されたマッピングがある場合には、ソースデータのみを考慮するだけで、極小またはトーンが、コアレンジの上側境界の近くにあるか、または図 1 2 B を参照して後述する 2 つの個別に生成された周波数タイル領域の間の周波数境界にあるか、を検出できる。

40

【0043】

50

問題のあるトーン成分が周波数境界の近くで発見された場合には、遷移周波数調整器 706 が、コア帯域と再構築帯域との間、または再構築帯域の全く同一のソースデータにより生成された別個の周波数部の間の遷移周波数またはクロスオーバー周波数またはギャップ充填開始周波数のような、遷移周波数の調整を実行する。ブロック 706 の出力信号は、境界のトーン成分を除去する除去器 708 に送られる。除去器は、ブロック 706 による遷移周波数調整の後にもそこに残っている残留トーン成分を除去するように構成されている。除去器 708 の結果は、続いて、フィルタリング問題に対処するためのクロスオーバーフィルタ 701 に送られ、クロスオーバーフィルタ 701 の結果が、再構築帯域内でスペクトル包絡線整形を行う包絡線整形ブロック 712 に入力される。

【0044】

図 7A を参照して説明したように、ブロック 704 内でのトーン成分の検出は、ソースデータ 705 と予備再構築信号 703 のどちらでも実行される。この実施形態は図 7B に示したもので、予備再構築信号が、ブロック 718 で示されるように作り出される。図 7A の信号成分が続いて検出器 720 に送られ、検出器 720 は、アーティファクトが作り出した成分を検出する。検出器 720 は図 7A に示すように、周波数境界でトーン成分を検出するための検出器として構成されるが、検出器が他のアーティファクトが作り出した成分を検出するような構成とすることもできる。そのようなスペクトル成分は、トーン成分以外の成分でも良く、アーティファクトが作り出されているかの検出が、異なる再構築を試み、この異なる再構築結果を比較して、どちらがアーティファクトが作り出した成分かを見出すことにより実行されてもよい。

【0045】

検出器 720 はここで、信号すなわち予備再生成信号を操作するためのマニピュレータ 722 を制御する。この制御は、線 723 により予備再生成信号を実際に処理することにより行われてもよく、例えば線 724 により示される修正された遷移周波数を用いた新たな再生成を実行することにより行われてもよい。

【0046】

操作手続きのひとつの実施形態では、遷移周波数が、図 7A に 706 で示すように調整される。さらなる実施形態が、図 8A に示される。この実施形態は、図 7A のブロック 706 あるいはブロック 706 を含む構成の代わりに実行される。検出器 802 が、問題のあるトーン部の開始および終了周波数を検出するために設けられる。内挿器 804 が、スペクトルレンジ内のトーン部の開始および終了の間に、内挿、望ましくは複素内挿するように構成される。このとき、図 8A にブロック 806 で示すように、トーン部は、内挿結果により置き換えられる。

【0047】

別の実施形態を、図 8A に、ブロック 808, 810 で示す。内挿を行う代わりに、トーン部の開始と終了との間で、スペクトル線 808 のランダム生成が実行される。続いて、ランダム生成されたスペクトル線のエネルギー調整が、810 で示されるように行われ、ランダム生成されたスペクトル線のエネルギーが、隣接する非トーンスペクトル部と同じになるように設定される。続いて、トーン部が、包絡線調整されたランダムに生成されたスペクトル線により置き換えられる。スペクトル線は、可能な限りアーティファクトの無い置き換え信号を提供するために、ランダムに、あるいは擬似ランダムに生成される。

【0048】

さらなる実施形態を、図 8B に示す。図 6A の周波数再生成器 604 内に配置される周波数タイル生成器を、ブロック 820 で示す。このとき、アナライザは、周波数タイル生成器により生成された信号を解析する。周波数タイル生成器 820 は、望ましくは、多タiling (タイル貼り) を実行して、多数の周波数タイルを生成する構成である。続いて、図 8B のマニピュレータ 824 は、周波数タイル生成器の結果を、アナライザ 822 による解析結果出力にしたがって操作する。この操作は、周波数境界の変更、または個々の部分の減衰である。続いて、スペクトル包絡線調整器 826 が、図 6A を参照して説明したように、パラメトリック情報 605 を用いて、スペクトル包絡線調整を実行する。

10

20

30

40

50

【 0 0 4 9 】

続いて、ブロック 8 2 6 によるスペクトル調整された信号出力は、周波数時間変換器 8 2 8 に入力される。周波数時間変換器 8 2 8 は、加えて、第 1 周波数部、すなわちコアデコード 6 0 0 の出力信号のスペクトル表現を受け取る。周波数時間変換器 8 2 8 の出力は、蓄積のために使用されるか、あるいは、オーディオレンダリングのためラウドスピーカに送られる。

【 0 0 5 0 】

本発明は、図 1 3 A , 1 3 B に示すような既知の周波数再生手続きにも適用でき、望ましくは、図 1 A から 5 B , 9 A から 1 0 D を参照して後述するインテリジェントギャップ充填のいずれにも適用できる。

10

【 0 0 5 1 】

図 1 A は、オーディオ信号 9 9 を符号化する装置を示す。オーディオ信号 9 9 は、あるサンプリングレートをもつオーディオ信号をスペクトル表現 1 0 1 へと変換する時間スペクトル変換器 1 0 0 に入力され、変換されたスペクトル表現 1 0 1 が、時間スペクトル変換器により出力される。スペクトル 1 0 1 は、スペクトル表現 1 0 1 を解析するスペクトル解析器 1 0 2 に入力される。スペクトル解析器 1 0 1 は、第 1 スペクトル解像度により符号化されるべき第 1 セットの第 1 スペクトル部 1 0 3 (第 1 スペクトル部の集合)と、第 2 スペクトル解像度により符号化されるべき別の第 2 セットの第 2 スペクトル部 1 0 5 (第 2 スペクトル部の集合)とを決定するよう構成される。第 2 スペクトル解像度は、第 1 スペクトル解像度よりも小さい。第 2 セットの第 2 スペクトル部 1 0 5 は、第 2 スペクトル解像度を有するスペクトル包絡線情報を計算するパラメータ計算器すなわちパラメトリック符号器 1 0 4 に入力される。さらに、第 1 スペクトル解像度を有する第 1 セットの第 1 スペクトル部の第 1 符号化表現 1 0 7 を生成するために、スペクトル領域オーディオコード 1 0 6 が設けられる。また、パラメータ計算器すなわちパラメトリック符号器 1 0 4 は、第 2 セットの第 2 スペクトル部の第 2 符号化表現 1 0 9 を生成するよう構成される。第 1 符号化表現 1 0 7 と第 2 符号化表現 1 0 9 とは、ビットストリーム多重変換装置またはビットストリーム形成器 1 0 8 に入力され、ビットストリーム形成器 1 0 8 は、最終的に、送信または蓄積装置における蓄積のために、符号化されたオーディオ信号を出力する。

20

【 0 0 5 2 】

一般的に、図 3 A の 3 0 6 などの第 1 スペクトル部は、3 0 7 a や 3 0 7 b のような 2 つの第 2 スペクトル部によって取り囲まれている。コア符号化周波数範囲が帯域的に制限されている H E A A C の場合には、この限りではない。

30

【 0 0 5 3 】

図 1 B は、図 1 A のエンコーダと整合するデコーダを示す。第 1 符号化表現 1 0 7 は、第 1 セットの第 1 スペクトル部の第 1 復号化表現を生成するスペクトル領域オーディオデコード 1 1 2 に入力される。さらに、第 2 符号化表現 1 0 9 は、第 1 スペクトル解像度よりも低い第 2 スペクトル解像度を有する第 2 セットの第 2 スペクトル部の第 2 復号化表現を生成するパラメトリックデコード 1 1 4 に入力される。

【 0 0 5 4 】

デコーダは、第 1 解像度を有する再構築第 2 スペクトル部を、第 1 スペクトル部を使って再生する周波数再生器 1 1 6 をさらに有する。周波数再生器 1 1 6 は、タイル充填手順を実行する。つまり、第 1 セットの第 1 スペクトル部のタイルまたは部分を使って、第 1 セットの第 1 スペクトル部を、第 2 スペクトル部を有する再構築レンジまたは再構築帯域にコピーする。そして、一般的には、パラメトリックデコード 1 1 4 により出力された復号された第 2 表現により示されるように、すなわち、第 2 セットの第 2 スペクトル部に關する情報を使って、スペクトル包絡線整形処理、またはその他の手順を実行する。復号化された第 1 セットの第 1 スペクトル部と、線 1 1 7 上の周波数再生器 1 1 6 の出力に示される再構築された第 2 セットのスペクトル部とは、スペクトル時間変換器 1 1 8 に入力される。このスペクトル時間変換器 1 1 8 は、第 1 復号表現と再構築された第 2 スペクトル

40

50

部とを、所定の高サンプリングレートを有する時間表現 1 1 9 に変換するよう構成される。

【 0 0 5 5 】

図 2 B は、図 1 A のエンコーダの実施形態を示す。オーディオ入力信号 9 9 は、図 1 A の時間スペクトル変換器 1 0 0 に対応する解析フィルタバンク 2 2 0 に入力される。そして、T N S ブロック 2 2 2 において、時間ノイズ整形処理が実行される。したがって、図 2 B のブロック音調マスク 2 2 6 に対応する図 1 A のスペクトル解析器 1 0 2 は、時間ノイズ整形 / 時間タイル整形処理が適用されないときには、全スペクトル値となり、図 2 B のブロック 2 2 2 に示す T N S 処理が適用されるときには、スペクトル残差値となる。2 チャンネル信号または多チャンネル信号の場合、ジョイントチャンネル符号化 2 2 8 が追 10 加的に実施されてもよく、図 1 A のスペクトル領域エンコーダ 1 0 6 には、ジョイントチャンネル符号化ブロック 2 2 8 が含まれてもよい。さらに、可逆データ圧縮を実施するエントロピー符号器 2 3 2 が設けられ、これは、図 1 A のスペクトル領域エンコーダ 1 0 6 の一部となる。

【 0 0 5 6 】

スペクトル解析器 / 音調マスク 2 2 6 は、T N S ブロック 2 2 2 の出力を、コア帯域と、第 1 セットの第 1 スペクトル部 1 0 3 に対応する音調成分と、図 1 A の第 2 セットの第 2 スペクトル部 1 0 5 に対応する残留成分とに分別する。I G F パラメータ抽出符号化手順として表示されるブロック 2 2 4 は、図 1 A のパラメトリック符号器 1 0 4 に対応し、 20 ビットストリーム多重変換装置 2 3 0 は、図 1 A のビットストリーム多重変換装置 1 0 8 に対応する。

【 0 0 5 7 】

解析フィルタバンク 2 2 2 が M D C T (変更離散コサイン変換フィルタバンク) として実施され、M D C T を使って、変更離散コサイン変換を周波数解析ツールとして機能させて、信号 9 9 を時間 - 周波数領域に変換することが好ましい。

【 0 0 5 8 】

スペクトル解析器 2 2 6 は、音調マスクを適用することが好ましい。この音調マスク推定段階は、信号中のノイズ様成分から音調成分を分別するのに用いられる。これによって、コア符号器 2 2 8 が全ての音調成分を心理音響モジュールによって符号化することができる。音調マスク推定段階は、いくつもの異なる方法によって実施されてもよく、音声 / オー 30 ディオ符号化 (非特許文献 8 , 9) 用の正弦およびノイズモデリング、または非特許文献 1 0 に記載される H I L N モデル系オーディオコードにおいて用いられる正弦トラック推定段階と機能的に同様に実施されるのが好ましい。実施形態としては、生死軌跡を維持する必要なく簡単に実施できることが好ましいが、その他の音調またはノイズ検知器を用いてもよい。

【 0 0 5 9 】

I G F モジュールは、ソース領域と対象領域との間に存在する類似性を計算する。対象領域は、ソース領域からのスペクトルにより表現される。ソース領域と対象領域との類似点を測ることは、クロス相関性アプローチを用いて行われる。対象領域は、非オーバーラップ周波数タイル $n T a r$ に分解される。対象領域における全てのタイルについて、固定 40 開始周波数からソーススタイル $n S r c$ が整形される。これらのソーススタイルは、0 および 1 の間の要素によってオーバーラップし、ここで 0 とは 0 % のオーバーラップを、1 とは 1 0 0 % のオーバーラップを意味する。これらのソーススタイルのそれぞれは、対象タイルに最もマッチするソーススタイルを見つけるための様々な遅延において、対象タイルと相関性を有する。最もマッチするタイルの数は、 $t i l e N u m [i d x _ { t a r }]$ に記憶され、対象と最も相関性を有する遅延は、 $x c o r r _ { l a g } [i d x _ { t a r }] [i d x _ { s r c }]$ に記憶され、相関性のサインは、 $x c o r r _ { s i g n } [i d x _ { t a r }] [i d x _ { s r c }]$ に記憶される。相関性が極めて否定的な場合には、ソーススタイルは、デコードにおけるタイル充填処理の前に、- 1 を乗じる必要がある。I G F モジュールはまた、音調成分が音調マスクを使って保存されていないため、スペクトルにおける音調成分を上書きしない 50

よう注意する必要がある。帯域側のエネルギーパラメータは、対象領域のエネルギーを蓄積するのに用いられ、これによってスペクトルを正確に再構築することができる。

【0060】

この方法は、正弦曲線間のギャップがソース領域からの最も整合する「整形ノイズ」により充填される一方、マルチ音調信号の倍音グリットがコア符号器によって保存されるという点で、伝統的なSBR（非特許文献1）に対して利点がある。ASR（精密スペクトル置換）（非特許文献2-4）と比べて、このシステムの別の利点としては、デコーダにおいて信号の重要部分を作り出す信号合成段がないということが挙げられる。その代わり、この作業はコア符号器によって実行される。よって、スペクトルの重要性分を保存できる。提案システムの別の利点としては、その機能が提案する継続的なスケーラビリティが挙げられる。全てのタイルについて単に $tileNum[idx_{tar}]$ と $xcorr_{lag} = 0$ とを使用することは、総粒度マッチングと称され、低ビットレートに使用することができる。一方、可変的な $xcorr_{lag}$ を全てのタイルに使用すると、対象スペクトルとソーススペクトルの整合性がより良くなる。

【0061】

さらに、トリリングやミュージカルノイズなどの周波数領域アーティファクトを除去するタイル選択安定化技術が提案されている。

【0062】

ステレオチャンネルのペアの場合には、追加的ジョイントステレオ処理が適用される。この処理は必要であり、というのも、ある目標レンジにおいては、信号が高い相関性を有する定位音源である場合があるためである。この特定の領域について選択された音源領域があまり相関を有しない場合には、目標領域についてエネルギーが整合していても、相関しない音源領域によって空間像が悪化する場合がある。エンコーダは、一般的にはスペクトル値のクロス相関処理をして、各目標領域のエネルギー帯域を解析して、もし所定の閾値を超過している場合には、このエネルギー帯域にジョイントフラグを設定する。デコーダにおいて、左右チャンネルのエネルギー帯域は、このジョイントステレオフラグが設定されていない場合、個別に処理される。ジョイントステレオフラグが設定されている場合、両方のエネルギーおよびパッチングがジョイントステレオ領域において実行される。IGF領域におけるジョイントステレオ情報は、コア符号処理におけるジョイントステレオ情報と同様に信号化され、予測の場合には、予測の方向性がダウンミックスからの残差なのか、それともその逆なのか、を示すフラグが含まれる。

【0063】

エネルギーは、L/R領域の送信済エネルギーから計算することができる。

$$midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

ここで、kは変形領域における周波数指数である。

【0064】

その他の解決としては、ジョイントステレオが積極的な帯域については、ジョイントステレオ領域において直接エネルギーを計算および送信し、デコーダ側にける追加的エネルギー変形を不要とすることが挙げられる。

【0065】

ソースタイルは、常に、ミッド/サイドマトリック：

$$midTile[k] = 0.5 \cdot (leftTile[k] + rightTile[k])$$

$$sideTile[k] = 0.5 \cdot (leftTile[k] - rightTile[k])$$

エネルギー調整：

$$midTile[k] = midTile[k] * midNrg[k];$$

$$sideTile[k] = sideTile[k] * sideNrg[k];$$

ジョイントステレオ LR変換：

10

20

30

40

50

何ら追加的予測パラメータが符号化されていない場合：

$$\text{leftTile}[k] = \text{midTile}[k] + \text{sideTile}[k]$$

$$\text{rightTile}[k] = \text{midTile}[k] - \text{sideTile}[k]$$

追加的な予測パラメータが必要な場合で、信号化された方向性がミッドからサイドの場合：

$$\text{sideTile}[k] = \text{sideTile}[k] - \text{predictionCoef} \cdot \text{midTile}[k]$$

$$\text{leftTile}[k] = \text{midTile}[k] + \text{sideTile}[k]$$

$$\text{rightTile}[k] = \text{midTile}[k] - \text{sideTile}[k]$$

信号化された方向性がサイドからミッドの場合：

$$\text{midTile1}[k] = \text{midTile}[k] - \text{predictionCoef} \cdot \text{sideTile}[k]$$

$$\text{leftTile}[k] = \text{midTile1}[k] - \text{sideTile}[k]$$

$$\text{rightTile}[k] = \text{midTile1}[k] + \text{sideTile}[k]$$

に従って生成される。

【0066】

この処理によって、相関性の高い目標領域と定位目標領域とを再生するのに用いられたタイルから生じた左右チャンネルが、たとえ音源領域が相関を有しないとしても、当該領域のステレオイメージを保存し、相関性を有する定位音源を示すよう確保できる。

【0067】

つまり、ビットストリームにおいては、例えばL/RまたはM/Sのいずれを総合ジョイントステレオ符号化処理に用いるべきかを示すジョイントステレオフラグが送信される。デコーダにおいては、まずコア帯域についてジョイントステレオフラグの示す通り、コア信号が復号される。第二に、コア信号は、L/RおよびM/S表現の両方に蓄積される。IGFタイル充填について、ソースタイル表現が選択され、IGF帯域に関するジョイントステレオ情報の示す対象タイル表現に合うようされる。

【0068】

時間ノイズ整形(TNS)は、標準技術であり、AAC(非特許文献11-13)の一部である。TNSは、知覚符号化の基本スキームを延長したものと考えられ、フィルタバンクと量子化段の間に、オプションとしての処理工程を挿入する。TNSモジュールのメイン作業は、過渡的な信号の時間的マスキング領域において、生成された量子化ノイズを隠すことであり、これによって、効率的な符号化スキームとなる。まず、TNSは、例えばMDCTのような変形領域において、「前方予測」を用いて予測係数セットを算出する。そして、これらの係数は、信号の時間的包絡線をフラット化するのに用いられる。量子化によってTNSフィルタスペクトルが影響を受けるため、量子化ノイズも一時的にフラットである。デコーダ側において、逆TNSフィルタリングを適用することで、量子化ノイズがTNSフィルタの時間的包絡線に従って整形され、よって量子化ノイズがその過渡性によりマスクされる。

【0069】

IGFは、MDCT表現に基づく。効率的符号化のためには、好ましくは、約20msの長ブロックを用いなければならない。このような長ブロック内における信号が過渡を含む場合、IGFスペクトル帯域において、タイル充填により、可聴のプリ・ポストエコーが発生する。図7Cは、過渡開始前の、IGFによる一般的なプリエコー効果を示す。左側において、オリジナル信号のスペクトル写真が示され、右側において、TNSフィルタリングなしで拡張された帯域幅のスペクトル写真が示される。

【0070】

このプリエコー効果は、IGF環境においてTNSを用いることにより低減される。ここでTNSは、デコーダにおけるスペクトル再生がTNS残差信号に対して実行されるため、時間タイル整形ツール(TTS)として用いられている。必要となるTTS予測係数は、通常通り、エンコーダ側における全スペクトルを用いて計算され、適用される。TN

10

20

30

40

50

S / T T S 開始および停止周波数は、I G F ツールの I G F 開始周波数 $f_{IGFs_{rart}}$ によって影響されない。伝統的な T N S と比較して、T T S 停止周波数は、 $f_{IGFs_{rart}}$ よりも高い I G F ツールの停止周波数まで増加される。デコーダ側において、T N S / T T S 係数が再び全スペクトルにて適用される。つまり、コアスペクトル + 再生スペクトル + 音調マップからの音調成分である（図 7 e を参照）。T T S の適用は、再生スペクトルの時間包絡線がオリジナル信号の包絡線と整合するよう形成するのに必要である。したがって、示されたプリエコーは低減される。さらに、T N S には通常のことではあるが、それによって $f_{IGFs_{rart}}$ より下の信号における量子化ノイズが依然として形成される。

【 0 0 7 1 】

10

伝統的なデコーダにおいては、オーディオ信号に対するスペクトルパッチングは、パッチ境界におけるスペクトル相関を害する。それによって、分散が生じ、オーディオ信号の時間包絡線が害される。よって、残差信号に対して I G F タイル充填を実施する別のメリットとしては、整形フィルタを適用した後、タイル境界がシームレスに相関し、信号をより忠実に時間的に再生することができることが挙げられる。

【 0 0 7 2 】

発明に係るエンコーダにおいては、T N S / T T S フィルタリング処理、音調マスク処理および I G F パラメータ予測処理を経たスペクトルは、音調成分を除き、I G F 開始周波数より上の信号は存在しない。まばらなスペクトルが今、算術符号化および予測符号化の原則によって、コア符号器により符号化される。これらの符号化成分は、信号のビットとともに、オーディオのビットストリームを形成する。

20

【 0 0 7 3 】

図 2 A は、対応するデコーダの実施形態を示す。符号化オーディオ信号に相当する図 2 A のビットストリームは、図 1 B であればブロック 1 1 2 および 1 1 4 に接続される多重分離装置 / デコーダに入力される。ビットストリーム多重分離装置は、入力オーディオ信号を、図 1 B の第 1 符号化表現 1 0 7 と図 1 B の第 2 符号化表現 1 0 9 とに分離する。第 1 セットの第 1 スペクトル部を有する第 1 符号化表現は、図 1 B のスペクトル領域デコーダ 1 1 2 に対応するジョイントチャンネル復号ブロック 2 0 4 に入力される。第 2 符号化表現は、図 2 A に図示されないパラメトリックデコーダ 1 1 4 に入力され、そして図 1 B の周波数再生器 1 1 6 に対応する I G F ブロック 2 0 2 に入力される。周波数再生に必要な第 1 セットの第 1 スペクトル部は、線 2 0 3 を経て I G F ブロック 2 0 2 に入力される。さらに、ジョイントチャンネル復号処理 2 0 4 の後、音調マスクブロック 2 0 6 において、音調マスク 2 0 6 の出力がスペクトル領域符号器 1 1 2 の出力と対応するよう、特定のコア復号処理が適用される、そして、合成器 2 0 8 による合成が実行され、合成器 2 0 8 の出力がフルレンジのスペクトルを有するが、それでいて、T N S / T T S フィルタ領域内にあるというフレーム構成が作成される。そして、ブロック 2 1 0 において、逆 T N S / T T S 処理が、線 1 0 9 を介して供給される T N S / T T S フィルタ情報を用いて実行される。つまり、好ましくは、T T S サイド情報が、スペクトル領域エンコーダ 1 0 6 によって生成される第 1 符号化表現に含まれる。スペクトル領域エンコーダ 1 0 6 は、例えば、直接的な A A C または U S A C コアエンコーダであってもよく、あるいは第 2 符号化表現に含まれてもよい。ブロック 2 1 0 の出力において、最大周波数までの全スペクトルを得ることができ、これはオリジナル入力信号のサンプリングレートによって定められるフルレンジの周波数である。そして、スペクトル / 時間変換が合成フィルタバンク 2 1 2 において実行され、オーディオ出力信号が最終的に得られる。

30

40

【 0 0 7 4 】

図 3 A は、スペクトル表現を表す概略図である。スペクトルは、スケールファクタ帯域 S C B に分割され、図 3 A における図示例においては、7 つのスケールファクタ帯域 S C B 1 ~ S C B 7 がある。スケールファクタ帯域は、A A C 規格に定められる A A C スケールファクタ帯域であってもよく、図 3 A にその概略が図示される通り、上側の周波数まで増加する帯域を有してもよい。スペクトルの当初から（つまり低周波数から）インテリジ

50

ェントギャップ充填を実行するのではなく、I G F 手順を 3 0 9 にて図示する I G F 開始周波数で開始することが好ましい。したがって、コア周波数帯域は、最低周波数から I G F 開始周波数まで存在する。I G F 開始周波数より上においては、スペクトル解析を適用して、高解像度のスペクトル成分 3 0 4、3 0 5、3 0 6、3 0 7（第 1 セットの第 1 スペクトル部）を、第 2 セットの第 2 スペクトル部の示す低解像度成分から分離する。図 3 A は、スペクトル領域エンコーダ 1 0 6 またはジョイントチャンネル符号器 2 2 8 に例示的に入力されるスペクトルを示し、つまりコアエンコーダがフルレンジで機能しているが、相当量のスペクトル値 0 を符号化している。これらのスペクトル値 0 は、ゼロに量子化されるか、または量子化の前または後にゼロに設定される。いずれにせよ、エンコーダは、フルレンジで動作し、スペクトルがその通り表現されているかのように、すなわち、い
10
ずれのインテリジェントギャップ充填についても、第 2 セットの第 2 スペクトル部の符号化が低スペクトル解像度で符号化されていることも、知る必要なしに動作する。

【 0 0 7 5 】

好適には、高解像度は、M D C T 線のようなスペクトル線の線単位の符号化によって定義づけられ、一方、第 2 解像度または低解像度は、例えばいくつもの周波数線を対象とするスケールファクタ帯域について、そのスケールファクタ帯域あたりの単一スペクトル値のみを計算することにより定義づけられる。したがって、第 2 低解像度は、スペクトル解像度については、一般に A A C や U S A C コアエンコーダなどのコアエンコーダによって適用される線単位符号化によって定義される第 1 または高解像度よりも、ずっと低い。

【 0 0 7 6 】

スケールファクタまたはエネルギー計算について、その状況が図 3 B に図示される。エンコーダがコアエンコーダであるという事実、および各帯域において第 1 セットのスペクトル部の成分が含まれ得る（ただし、必ずしも含まれる必要はない）という事実に伴い、コアエンコーダは、I G F 開始周波数 3 0 9 より下のコア範囲のみならず、I G F 開始周波数より上の最大周波数 $f_{I G F \text{ stop}}$ （これはサンプリング周波数、つまり $f_s / 2$ の半分以下である）に達するまでの帯域について、I G F 開始周波数各帯域についてスケールファクタを計算する。よって、本実施形態において、図 3 A における符号化音調部 3 0 2、3 0 4、3 0 5、3 0 6、3 0 7 は、スケールファクタ S C B 1 ~ S C B 7 とともに、高解像度スペクトルデータに対応する。低解像度スペクトルデータは、I G F 開始周波数から開始して計算され、スケールファクタ S F 4 ~ S F 7 とともに送信される、エネルギー情報値 E_1 、 E_2 、 E_3 、 E_4 に対応する。
20
30

【 0 0 7 7 】

特に、コアエンコーダが低ビットレート状況下にある場合、コア帯域、つまり I G F 開始周波数よりも低い周波数（スケールファクタ帯域 S C B 1 ~ S C B 3）において、追加的ノイズ充填手順が適用されてもよい。ノイズ充填においては、ゼロに量子化されたいくつもの近接するスペクトル線が存在する。デコーダ側においては、これらのゼロに量子化されたスペクトル値は再度合成され、この合成されたスペクトル値について、図 3 B に 3 0 8 として示される N F₂ などのノイズ充填エネルギーを用いて、その大きさが調整される。特に、U S A C のように、スケールファクタについて絶対ベースまたは相対ベースで設定され得るノイズ充填エネルギーは、ゼロに量子化されたスペクトル値のセットのエネルギー
40
に対応する。これらのノイズ充填スペクトル線はまた、第 3 セットの第 3 スペクトル部と考えられ、直接的なノイズ充填合成により、ソースレンジおよびエネルギー情報値 E_1 、 E_2 、 E_3 、 E_4 からのスペクトル値を用いた周波数タイルを再構築するためのその他の周波数からの周波数タイルを用いた周波数再生に依拠する周波数タイル I G F 手順を適用せずに、再生される。

【 0 0 7 8 】

エネルギー情報が計算される帯域は、スケールファクタ帯域と一致することが好ましい。別の実施形態においては、例えばスケールファクタ帯域 4 および 5 について単一のエネルギー情報値のみが送信されるよう、エネルギー情報値のグループ分けが適用されるが、この実施形態においても、グループ分けされた再構築帯域は、スケールファクタ帯域の境界と一
50

致する。もし異なる帯域分別が適用された場合には、所定の再計算または同期化計算を適用してもよく、所定の実施形態によってはこれが合理的である。

【 0 0 7 9 】

図 1 A のスペクトル領域エンコーダ 1 0 6 は、図 4 A に図示されるような心理音響的に駆動されるエンコーダであることが好ましい。一般的には、例えば M P E G 2 / 4 の A A C 規格または M P E G 1 / 2 レイヤ 3 規格に示される通り、スペクトル範囲（図 4 A の 4 0 1 ）に変換された後の符号化対象のオーディオ信号は、スケールファクタ算出器 4 0 0 に転送される。スケールファクタ算出器は、心理音響モデルによる制御により、量子化対象のオーディオ信号をさらに受信するか、または M P E G 1 / 2 レイヤ 3 規格または M P E G A A C 規格の通り、オーディオ信号の復号スペクトル表現を受信する。心理音響モデルは、各スケールファクタ帯域について、心理音響閾値を表現するスケールファクタを算出する。さらに、スケールファクタは、公知の内側および外側反復ループの連携により、またはその他適切なエンコーディング手順により、所定のビットレート条件が満たされるよう調整される。そして、量子化対象スペクトル値と算出スケールファクタとが量子化処理器 4 0 4 に入力される。直接的なオーディオ符号化動作においては、量子化対象スペクトル値は、スケールファクタによって重み付けされ、この重み付けされたスペクトル値が、一般的に上位の振幅範囲まで圧縮する機能を有する固定量子化装置に入力される。そして、量子化処理器の出力においては、エントロピーエンコーダに転送される量子化インデックスが確かに存在する。一般的に、このエントロピーエンコーダは、近接する周波数値のためのゼロ量子化インデックスまたは当該分野におけるいわゆるゼロ値「ラン」のために、特定のかつ相当に効率のよい符号機能を有する。

【 0 0 8 0 】

しかしながら、図 1 A のオーディオエンコーダにおいては、量子化処理器は、一般的に、スペクトル解析器から第 2 スペクトル部に関する情報を受信する。したがって、量子化処理器 4 0 4 は、その出力において、スペクトル解析器 1 0 2 によって識別された第 2 スペクトル部がゼロであるか、またはエンコーダもしくはデコーダによってゼロ表現であると認められた表現（特にスペクトルにゼロ値の「ラン」が存在する場合に、相当に効率的に符号化できる表現）を有するよう確保する。

【 0 0 8 1 】

図 4 B は、量子化処理器の実施形態を図示する。M D C T スペクトル値は、ゼロ設定ブロック 4 1 0 に入力することができる。そして、ブロック 4 1 2 におけるスケールファクタによる重み付けが実施される前に、第 2 スペクトル部が既にゼロに設定される。追加の実施形態においては、ブロック 4 1 0 が設けられず、重み付けブロック 4 1 2 の後に、ブロック 4 1 8 においてゼロ設定の連携が実施される。さらに別の実施形態においては、量子化ブロック 4 2 0 における量子化の後に、ゼロ設定ブロック 4 2 2 においてゼロ設定手順が実施されてもよい。この実施形態においては、ブロック 4 1 0 および 4 1 8 は存在しない。概して、ブロック 4 1 0、4 1 8 および 4 2 2 のうち少なくとも 1 つが、具体的な実施形態に応じて設けられる。

【 0 0 8 2 】

そして、ブロック 4 2 2 の出力において、図 3 A に図示されるものに対応する量子化スペクトルが得られる。そして、この量子化スペクトルは、図 2 B における 2 3 2 のようなエントロピー符号器に入力される。このエントロピー符号器は、例えば U S A C 規格に定義されるハフマン符号器または演算符号器であってもよい。

【 0 0 8 3 】

ゼロ設定ブロック 4 1 0、4 1 8 および 4 2 2 は、互いに交互に、または並列に設けられるが、スペクトル解析器 4 2 4 によって制御される。スペクトル解析器は、公知の音調検知器を有することが好ましく、またはスペクトルを高解像度で符号化される成分と低解像度で符号化される成分とに分別することができる異なる種の検知器を有することが好ましい。その他スペクトル解析器において実行されるアルゴリズムは、異なるスペクトル部に関する解像度要件を、スペクトル情報または関連メタデータに応じて決定する音声活動

検出器、ノイズ検出器、音声検出器、その他検出器であってもよい。

【 0 0 8 4 】

図 5 A は、例えば A A C や U S A C において実施される、図 1 A の時間スペクトル変換器 1 0 0 の好ましい実施形態を図示する。時間スペクトル変換器 1 0 0 は、過渡検知器 5 0 4 によって制御される窓回路 5 0 2 を有する。過渡検知器 5 0 4 が過渡を検知すると、長い窓から短い窓への切り替えが窓回路 5 0 2 に対して指示される。そして、窓回路 5 0 2 は、オーバーラップしているブロックについて、各窓フレームが一般的に 2 0 4 8 値のような 2 つの N 値を有する場合には、窓フレームを計算する。その後、ブロック変換器 5 0 6 内での変換が実施され、このブロック変換器が追加的デシメーションを提供する。そして、デシメーション / 変換の結合が実行され、M D C T スペクトル値などの N 値を有するスペクトルフレームが得られる。そのため、長窓動作については、ブロック 5 0 6 の入力におけるフレームが、例えば 2 0 4 8 値のような 2 つの N 値を有し、そしてスペクトルフレームが 1 0 2 4 値を有する。しかしながら、その後、長窓と比較して各短ブロックが 1 / 8 の窓時間領域値しか有さず、かつ各スペクトルブロックが長ブロックと比較して 1 / 8 のスペクトル値しか有さない場合に 8 つの短ブロックが実行されたときには、短ブロックに対する切替が実行される。そのため、このデシメーションが窓回路の 5 0 % オーバーラップ動作と結合したとき、スペクトルは、時間領域オーディオ信号 9 9 のクリティカルサンプルバージョンである。

【 0 0 8 5 】

続いて、図 1 B の周波数再生器 1 1 6 およびスペクトル時間変換器 1 1 8 の特定の実施形態、または図 2 A のブロック 2 0 8 および 2 1 2 の結合動作の具体的実施形態を示す図 5 B を参照する。図 5 B において、特定の再構築帯域が、図 3 A のスケールファクタ帯域 6 のように考慮される。この再構築帯域における第 1 スペクトル部（つまり図 3 A の第 1 スペクトル部 3 0 6 ）は、フレーム構築 / 調整ブロック 5 1 0 に入力される。その上、スケールファクタ帯域 6 のための再構築第 2 スペクトル部も、フレーム構築器 / 調整器 5 1 0 に入力される。また、図 3 B における E_3 のような、スケールファクタ帯域 6 のためのエネルギー情報もまた、ブロック 5 1 0 に入力される。再構築帯域における再構築第 2 スペクトル部は、ソースレンジを使って周波数タイル充填により既に生成されており、この再構築帯域が対象レンジに対応する。ここで、フレームのエネルギー調整を実行し、例えば、図 2 A の結合器 2 0 8 の出力において得られる、N 値を有する完全な再構築フレームを最終的に得る。そして、ブロック 5 1 2 において、逆ブロック変換 / 補間処理を実行して、例えば、ブロック 5 1 2 の入力における 1 2 4 スペクトル値について、2 4 8 時間領域値を得る。その後、ブロック 5 1 4 において合成窓処理が実行され、これもまた、符号化されたオーディオ信号においてサイド情報として送信される長窓 / 短窓指示により制御される。そして、ブロック 5 1 6 において、前回時間フレームとのオーバーラップ / 加算処理が実行される。各新規の時間フレームが 2 N 値の場合、N の時間領域値が最終的に出力されるよう、M D C T は、5 0 % オーバーラップ処理を適用することが好ましい。5 0 % オーバーラップ処理によってクリティカルサンプルが提供され、あるフレームから次のフレームへの継続的クロスオーバーが、ブロック 5 1 6 におけるオーバーラップ / 加算処理により提供されるという事実からすると、5 0 % オーバーラップ処理が特に好ましい。

【 0 0 8 6 】

図 3 A の 3 0 1 に示す通り、ノイズ充填処理は、I G F 開始周波数より下で追加的に適用されるのみならず、図 3 A のスケールファクタ帯域 6 と合致する対象再構築帯域などのような、I G F 開始周波数より上においても追加的に適用されてもよい。そして、ノイズ充填スペクトル値も、フレーム構築器 / 調整器 5 1 0 に入力されてもよく、ノイズ充填スペクトル値がこのブロック内にて調整されてもよい。あるいは、ノイズ充填スペクトル値は、フレーム構築器 / 調整器 5 1 0 に入力されるよりも前に、ノイズ充填エネルギーを用いて既に調整されていてもよい。

【 0 0 8 7 】

好ましくは、I G F 処理（つまりその他の部分からのスペクトル値を使った周波数タイ

10

20

30

40

50

ル充填処理)は、全スペクトルに対して適用されてもよい。したがって、スペクトルタイル充填処理は、I G F 開始周波数より上の高帯域のみならず、低帯域においても適用されてもよい。また、周波数タイル充填によらないノイズ充填もまた、I G F 開始周波数より下のみならず、I G F 開始周波数より上でも適用されてもよい。但し、ノイズ充填処理がI G F 開始周波数よりも下に限定され、かつ図3Aに示される通り、周波数タイル充填処理がI G F 開始周波数よりも上に限定されているときに、高品質および高効率のオーディオエンコーディングが得られることが判明している。

【0088】

対象タイル(T T)(I G F 開始周波数より大きい周波数を有する)は、フルレート符号器のスケールファクタ帯域境界に結合させることが好ましい。ソースタイル(S T)は、ここから情報を得られるが、I G F 開始周波数よりも低周波数であり、スケールファクタ帯域境界と結合していない。S Tの大きさは、関連するT Tの大きさに対応すべきである。これを以下の例を使って示す。T T[0]は、10MDC Tピンの長さを有する。これはまさに、2つの後続のS C Bの長さ(例えば4+6)に相当する。そして、T T[0]と相関性を有する可能性のある全てのS Tもまた、10ピンの長さを有する。T T[0]に近接する第2対象タイルT T[1]は、15ピンの長さを有する(7+8の長さを有するS C B)。そして、そのS Tは、T T[0]用の10ピンではなく、15ピンの長さを有する。

【0089】

対象タイルの長さを有するS Tについて、T Tを見つけれない場合が生じたとき(例えば、T Tの長さが利用可能なソースレンジよりも大きいとき)には、相関性は計算されず、対象タイルT Tが完全に充填されるまで、ソースレンジが何度もこのT Tにコピーされる(このコピーは、第2コピーの最も低い周波数を有する周波数線が、第1コピーの最も高い周波数を有する周波数線の直後に周波数的に来るよう、次から次になされる。)。

【0090】

そして、図1Bの周波数生成器116または図2AのI G F ブロック202のさらなる好適実施形態を示す図5Cを参照する。ブロック522は、対象帯域I Dのみならず、さらにソース帯域I Dをも受信する周波数タイル生成器である。例示的に、エンコーダ側において、図3Aのスケールファクタ帯域3がスケールファクタ帯域7の再構築にかなり適していると決定されている。したがって、ソース帯域I Dは2であり、対象帯域I Dは7である。この情報に基づき、周波数タイル生成器522は、スペクトル成分523の生の第2部を生成するために、コピー処理や倍音タイル充填処理、その他タイル充填処理を適用する。スペクトル成分の生の第2部は、第1セットの第1スペクトル部に含まれる周波数解像度と同じ周波数解像度を有している。

【0091】

そして、図3Aの307などのような再構築帯域の第1スペクトル部は、フレーム構築器524に入力され、生の第2部523もフレーム構築器524に入力される。その後、再構築されたフレームは、ゲイン係数算出器528によって算出された再構築帯域のゲイン係数を利用して、調整器526により調整される。ただし、重要な点としては、フレームにおける第1スペクトル部は、調整器526による影響を受けず、再構築フレームのための生の第2部のみが調整器526の影響を受けることである。このため、ゲイン係数算出器528は、ソース帯域または生の第2部523を解析し、さらに、スケールファクタ帯域7が想定される場合には、調整器526によって出力される調整後フレームのエネルギーがエネルギー E_4 を有するよう、再構築帯域における第1スペクトル部を追加的に解析して、正しいゲイン係数527を最終的に見出す。

【0092】

この意味では、H E - A A Cと比べて、本発明による周波数再構築の正確性が高いことを評価することが非常に重要である。これは、図3Aにおけるスケールファクタ帯域7について説明される。図13Aに図示される従来のエンコーダは、高解像度で符号化されるべきスペクトル部307を「欠落倍音」として検知する。そして、このスペクトル成分の

10

20

30

40

50

エネルギーが、スケールファクタ帯域 7 などの再構築帯域に関するスペクトル包線情報とともに、デコーダに送信される。その後、デコーダが欠落倍音を再生する。しかしながら、欠落倍音 307 が図 13A に示す従来のデコーダによって再構築されるスペクトル値が、再構築周波数 390 によって示される周波数における帯域 7 の中央に存在する。よって、本発明は、図 13D の従来のデコーダによって生じる周波数エラー 391 を避けることができる。

【0093】

一実施形態において、スペクトル解析器を実行して、第 1 スペクトル部と第 2 スペクトル部の類似性を算出し、算出された類似性に基づき、再構築範囲の第 2 スペクトル部について、この第 2 スペクトル部とできる限りマッチする第 1 スペクトル部を決定する。そして、可変的ソース範囲 / 対象範囲の実行において、パラメトリックコードは、第 2 符号化表現において、各対象範囲について整合する対象範囲を表現する整合情報を追加的に導入する。デコーダ側において、この情報は、図 5C の周波数タイル生成器 522 によって用いられる。図 5C は、ソース帯域 ID および対象帯域 ID に基づく生の第 2 部 523 の生成を図示する。

10

【0094】

さらに、図 3A に示す通り、スペクトル解析器は、最大限の解析周波数まで、スペクトル表現を解析するよう構成される。この最大限の解析周波数は、サンプリング周波数の半分を少しだけ下回り、少なくともサンプリング周波数の $1/4$ かまたは一般的にはこれより高い方が好ましい。

20

【0095】

図示される通り、エンコーダは、低解像度処理をすることなく動作し、またデコーダは、高解像度処理をすることなく動作する。つまり、スペクトル領域オーディオコードは、オリジナルの入力オーディオ信号のサンプリングレートにより定義されるナイキスト周波数を有するスペクトル表現を生成するよう構成される。

【0096】

さらに、図 3A に図示される通り、スペクトル解析器は、ギャップ充填開始周波数から開始し、スペクトル表現に含まれる最大周波数に示される最大周波数で終了するスペクトル表現を解析するよう構成される。そして、最大周波数からギャップ充填開始周波数までのスペクトル部が第 1 セットのスペクトル部に属し、例えば 304、305、306、307 などのギャップ充填周波数より上の周波数値を有するさらなるスペクトル部が、第 1 セットの第 1 スペクトル部に追加的に含まれている。

30

【0097】

上述の通り、スペクトル領域オーディオデコーダ 112 は、第 1 復号表現のスペクトル値によって示される最大周波数が、サンプリングレートを有する時間表現に含まれる最大周波数と等しくなるよう構成され、第 1 セットの第 1 スペクトル部における最大周波数のスペクトル値が、ゼロまたはゼロとは異なる値である。いずれにせよ、スペクトル成分の第 1 セットにおけるこの最大周波数については、スケールファクタ帯域のスケールファクタが存在する。このスケールファクタは、図 3A および 3B の関係で議論された通り、このスケールファクタ帯域における全てのスペクトル値がゼロに設定されたか否かを問わず、生成され送信される。

40

【0098】

したがって、本発明は、ノイズ置換やノイズ充填（これらの技術は、ローカル信号コンテンツのようなノイズの効率的表現に特化している）などのような圧縮効率を向上させるその他のパラメータ技術について、本発明によると音調成分の正確な周波数再生が可能であるという利点を有する。今日まで、いずれの従来技術も、低帯域（LF）と高帯域（HF）における固定された演繹分割の制限のない、スペクトルギャップ充填による任意の信号コンテンツの効率的パラメトリック表現について、提案していない。

【0099】

本発明のシステムの実施形態は、従来技術のアプローチを改善し、高圧縮効率を実現し

50

、知覚的な不快感をほとんどなくし、低ビットレートであっても全オーディオ帯域を提供する。

【0100】

一般的なシステムは、下記から構成される。

- ・全帯域コア符号化
- ・インテリジェントギャップ充填（タイル充填またはノイズ充填）
- ・音調マスクにより選択されたコア内におけるまばらな音調部分
- ・充填を含む、全帯域のためのジョイントステレオペア符号化
- ・タイル上のTNS
- ・IGFレンジにおけるスペクトル白色化

10

【0101】

より効率的なシステムに向かった第1ステップとしては、スペクトルデータを、コア符号器の1つとは異なる第2変形領域に変換する必要性を除去することである。例えばAACなどの多くのオーディオコーデックは、MDCTを基本的な変換として用いているので、MDCT領域においてBWEも実行できると便利である。BWEシステムの第2要件としては、音調グリッドを保存する必要がある。これによって、HF音調成分が保存され、符号化オーディオの質が既存のシステムよりも優れたものになる。BWEスキームについて、上記の要件の両方に対処するために、インテリジェントギャップ充填（IGF）と称される新システムが提案されている。図2Bは、エンコーダ側における提案システムのブロック図を示し、図2Aは、デコーダ側におけるシステムを示す。

20

【0102】

図9Aは、符号化オーディオ信号を復号する装置を示す。このオーディオ信号は、第1セットの第1スペクトル部の符号化表現、および第2セットの第2スペクトル部についてのスペクトルエネルギーを示すパラメトリックデータの符号化表現を含んでいる。第1セットの第1スペクトル部は、図9Aにおいて901aで示され、パラメトリックデータの符号化表現は、図9Aの901bで示される。オーディオデコーダ900は、第1セットの第1スペクトル部の符号化表現901aを復号して、復号化された第1セットの第1スペクトル部904を得るよう設けられる。また、オーディオデコーダ900は、パラメトリックデータの符号化表現を復号して、個別の再構築帯域について個別のエネルギーを示す第2セットの第2スペクトル部についての復号パラメトリックデータ902を得る。ここで、第2スペクトル部は、再構築帯域内に存在している。さらに、周波数再生器906は、第2スペクトル部を有する再構築帯域のスペクトル値を再構築するために設けられる。さらに、周波数再生器906は、第1セットの第1スペクトル部に含まれる第1スペクトル部と、再構築帯域に係る個別のエネルギー情報を用いる。ここで、再構築帯域には、第1スペクトル部と第2スペクトル部が含まれる。周波数再生器906は、残存エネルギー情報を決定する算出器912を有する。この残存エネルギー情報には、再構築帯域内に周波数を有する第1スペクトル部の累積エネルギーが含まれる。さらに、周波数再生器906は、再構築帯域のさらなるスペクトル部のタイルエネルギー情報を決定し、第1スペクトル部とは異なる周波数値のための計算器918を有する。これらの周波数値は、再構築帯域に周波数を有するものであり、さらなるスペクトル部は、再構築帯域の第1スペクトル部とは異なる第1スペクトル部を用いて、周波数再構築により生成されるものである。

30

40

【0103】

周波数再生器906は、再構築帯域における欠落エネルギーのための計算器914をさらに有する。この計算器914は、再構築帯域についての個別エネルギーおよびブロック912により生成された残存エネルギーを用いて動作する。さらに、周波数再生器906は、欠落エネルギー情報、およびブロック918により生成されたタイルエネルギー情報に基づき、再構築帯域におけるさらなるスペクトル部を調整するスペクトル包絡線調整器916を有する。

【0104】

再構築帯域920の一例を示す図9Cを参照する。再構築帯域は、図3Aにおける第1

50

スペクトル部 306 などの再構築帯域における第 1 スペクトル部を有し、これは 921 としてその概略が示される。さらに、再構築帯域 920 における残りのスペクトル値は、例えば、図 3 A のインテリジェントギャップ充填開始周波数 309 より下のスケールファクタ帯域 1、2、3 からのソース領域を用いて生成されることとなる。周波数生成器 906 は、第 2 スペクトル部 922 および 923 のための生スペクトル値を生成するよう構成される。そして、図 9 C に示される通り、周波数帯域 922 および 923 におけるスペクトル値を最終的に調整するために、そして再構築帯域 920 における再構築および調整第 2 スペクトル部を得るために、ゲインファクタ g が計算される。これらの第 2 スペクトル部は、第 1 スペクトル部 921 と同じスペクトル解像度、つまり同じ線距離を有している。図 9 C において 921 で示される再構築帯域における第 1 スペクトル部は、オーディオデコード 900 によって復号され、図 9 B のブロック 916 の実施する包絡線調整によって影響を受けないことを理解することが重要である。その代わり、921 で示される再構築帯域における第 1 スペクトル部は、線 904 を介して、全帯域または全レートのオーディオデコード 900 により出力されるため、そのままに放置される。

10

【0105】

以下において、実数を用いた一例について説明する。ブロック 912 により算出された残りの残存エネルギーは、例えば 5 エネルギー単位だとする。このエネルギーは、第 1 スペクトル部 921 において例示的に示された 4 つのスペクトル線のエネルギーである。

【0106】

さらに、図 3 B または図 3 A のスケールファクタ帯域 6 に対応する再構築帯域に係るエネルギー数 E_3 は、10 単位に等しい。重要なこととしては、エネルギー値には、スペクトル部 922 および 923 のエネルギーのみならず、エンコード側で算出された再構築帯域 920 の全エネルギー（つまり、例えば、音調マスクを用いてスペクトル解析を実行する前のもの）も含まれる。したがって、10 エネルギー単位は、再構築帯域における第 1 および第 2 スペクトル部を対象とする。そして、ブロック 922、923 のソースレンジデータのエネルギーまたはブロック 922、923 用の生の対象レンジデータのエネルギーは、8 エネルギー単位であるという前提とする。よって、5 単位の欠落エネルギーが計算される。

20

【0107】

タイルエネルギー tE_k により分割された欠落エネルギーに基づき、0.79 のゲインファクタが計算される。そして、第 2 スペクトル部 922、923 の生のスペクトルラインに対して、算出されたゲインファクタを乗じる。よって、第 2 スペクトル部 922、923 のスペクトル値のみが調整され、第 1 スペクトル部 921 のスペクトル線は、この包絡線調整によって影響を受けない。第 2 スペクトル部 922、923 について生のスペクトル値を乗じた後、完全なる再構築帯域が計算される。これは、再構築帯域における第 1 スペクトル部から構成され、再構築帯域 920 における第 2 スペクトル部 922、923 のスペクトル線から構成される。

30

【0108】

帯域 922、923 の生のスペクトルデータを生成するためのソースレンジは、周波数について、IGF 開始周波数 309 より下であり、再構築帯域 920 は IGF 開始周波数 309 よりも上であることが好ましい。

40

【0109】

さらに、再構築帯域境界がスケールファクタ帯域境界と一致することが好ましい。よって、一実施形態において、再構築帯域は、コアオーディオデコードの対応スケールファクタ帯域の大きさを有するか、またはエネルギーペアリングが適用されたとき、再構築帯域のエネルギー値が整数 2 以上のスケールファクタ帯域のエネルギーを提供する大きさを有する。したがって、スケールファクタ帯域 4、スケールファクタ帯域 5 およびスケールファクタ帯域 6 についてエネルギー蓄積が実施されるとの前提の場合には、再構築帯域 920 の低周波数境界は、スケールファクタ帯域 4 の下限境界と等しくなり、再構築帯域 920 の高周波数境界は、スケールファクタ帯域 6 の上限境界と一致する。

【0110】

50

以下において、図 9 A のデコーダのさらなる機能を示すために、図 9 D について説明する。オーディオデコーダ 9 0 0 は、第 1 セットのスペクトル部の第 1 スペクトル部に対応する逆量子化スペクトル値を受信する。そして、図 3 B に示すようなスケールファクタ帯域のためのスケールファクタが、追加的に逆スケールブロック 9 4 0 に提供される。逆スケールブロック 9 4 0 は、図 3 A の I G F 開始周波数 3 0 9 より下の全ての第 1 セットの第 1 スペクトル部を提供し、さらに I G F 開始周波数より上の第 1 スペクトル部（つまり、図 3 A の第 1 スペクトル部 3 0 4、3 0 5、3 0 6、3 0 7 であり、これらは全て図 9 D の 9 4 1 に示す通り、再構築帯域内に位置している）を提供する。さらに、再構築帯域における周波数タイル充填に用いられるソース帯域の第 1 スペクトル部は、包絡線調整器 / 計算器 9 4 2 に提供され、このブロックは、図 9 D において 9 4 3 で示す通り、符号化されたオーディオ信号に対してパラメータサイド情報として提供された再構築帯域のエネルギー情報を追加的に受信する。そして、包絡線調整器 / 計算器 9 4 2 は、図 9 B および 9 C の機能を提供し、再構築帯域における第 2 スペクトル部のために、調整されたスペクトル値を最終的に出力する。再構築帯域における第 2 スペクトル部に係るこれらの調整スペクトル値 9 2 2、9 2 3、並びに図 9 D において線 9 4 1 により示される再構築帯域の第 1 スペクトル部 9 2 1 は、統合すると、再構築帯域の全スペクトル表現をなしている。

【 0 1 1 1 】

以下において、符号化されたオーディオ信号を提供または生成するために、オーディオ信号を符号化するオーディオエンコーダの好適実施形態を説明する図 1 0 A および 1 0 B を参照する。このエンコーダは、スペクトル解析器 1 0 0 4 に供給する時間 / スペクトル変換器 1 0 0 2 を有し、このスペクトル解析器 1 0 0 4 は、パラメータ計算器 1 0 0 6 に接続する一方、オーディオエンコーダ 1 0 0 8 にも他方で接続する。オーディオエンコーダ 1 0 0 8 は、第 1 セットの第 1 スペクトル部の符号化表現を提供し、第 2 セットの第 2 スペクトル部を対象としない。一方で、パラメータ計算器 1 0 0 6 は、第 1 および第 2 スペクトル部を対象とする再構築帯域のエネルギー情報を提供する。さらに、オーディオエンコーダ 1 0 0 8 は、第 1 スペクトル解像度を有する第 1 スペクトル部の第 1 セットの第 1 符号化表現を生成するよう構成される。ここで、オーディオエンコーダ 1 0 0 8 は、ブロック 1 0 0 2 によって生成されるスペクトル表現の全ての帯域に係るスケールファクタを提供する。これに加えて、図 3 B に示す通り、エンコーダは、図 3 A に示す I G F 開始周波数 3 0 9 よりも上の周波数に位置する再構築帯域については少なくとも、エネルギー情報を提供する。したがって、再構築帯域がスケールファクタ帯域またはスケールファクタ帯域群と好適に一致しているので、オーディオエンコーダ 1 0 0 8 からの対応スケールファクタと、パラメータ計算器 1 0 0 6 によって出力されるエネルギー情報という 2 つの値が与えられる。

【 0 1 1 2 】

オーディオエンコーダは、異なる周波数帯域幅を有する（つまり異なる数のスペクトル値を有する）スケールファクタ帯域を有することが好ましい。したがって、パラメータ計算器は、特定の再構築帯域の帯域幅について、異なる帯域幅のためのエネルギーを正規化する正規化器 1 0 1 2 を有する。このため、正規化器 1 0 1 2 は、入力として、帯域におけるエネルギーと、帯域におけるいくつものスペクトル値を受信し、再構築 / スケールファクタ帯域毎に正規化されたエネルギーを出力する。

【 0 1 1 3 】

さらに、図 1 0 A のパラメータ計算器 1 0 0 6 a は、図 1 0 A における線 1 0 0 7 により示されるコアまたはオーディオエンコーダ 1 0 0 8 から制御情報を受信するエネルギー値計算器を有する。この制御信号には、オーディオエンコーダによって使用される長 / 短ブロックに関する情報、および / またはグループ化情報が含まれる場合がある。したがって、長 / 短ブロックに関する情報および短窓に関するグループ化情報が「時間」グループ分けに関するものである一方、グループ化情報はまた、スペクトルグループ分け（つまり、2 つのスケールファクタ帯域を単一の再構築帯域にグルーピングすること）に言及してもよい。よって、エネルギー値計算器 1 0 1 4 は、スペクトル部のみがグループ分けされてい

るとき、第1予備第2スペクトル部を対象とするグループ分けされた各帯域について、単一のエネルギー値を出力する。

【0114】

図10Dは、スペクトルグループ分けを実施するためのさらなる実施形態を示す。このため、ブロック1016は、2つの近接帯域のエネルギー値を計算するよう構成される。そして、ブロック1018において、近接する帯域のエネルギー値を比較し、エネルギー値がそれほど異ならない場合、または例えば、閾値で規定したほど異ならない場合には、ブロック1020に示すように、両方の帯域について、単一の（正規化された）値が生成される。線1019に示す通り、ブロック1018を迂回してもよい。さらに、ブロック1020により実行される2つ以上の帯域に係る単一値の生成は、エンコードビットレート制御1024により制御されてもよい。このため、ビットレートを低減する必要がある場合には、ブロック1018における比較において、エネルギー情報値のグループ分けが許されない場合であったとしても、符号化されたビットレート制御処理1024により、ブロック1020を制御して、2つ以上の帯域について単一の正規化値を生成させる。

【0115】

オーディオエンコーダが2つ以上の短窓のグループ分けを行っているときには、このグループ分けがエネルギー情報についても適用される。コアエンコーダが2つ以上の短ブロックのグループ分けを行うときには、これら2つ以上のブロックについて、スケールファクタの単一セットのみを計算し送出する。デコーダ側においては、オーディオデコーダは、スケールファクタの同じセットを両方のグループ分けされた窓に適用する。

【0116】

エネルギー情報計算に関して、再構築帯域におけるスペクトル値が2つ以上の短窓に対して蓄積される。つまり、短ブロックおよび後続の短ブロックについての所定の再構築帯域におけるスペクトル値は一緒に蓄積され、2つの短ブロックを対象とするこの再構築帯域については、単一のエネルギー情報値しか送信されないことを意味する。そして、デコーダ側においては、図9Aないし9Dに関して説明された包絡線調整は、各短ブロックについて個別には実施されず、グループ分けされた短窓のセットについて一緒に実施される。

【0117】

周波数におけるグループ分けまたは時間におけるグループ分けが実施された場合であっても、正規化によって、デコーダ側のエネルギー情報計算について、一方のエネルギー情報値と再構築帯域またはグループ分けされた再構築帯域のセットにおけるスペクトルライン量のみが簡単に知られるよう、対応する正規化が再び実施される。

【0118】

さらに、スペクトルエネルギーの情報、個別エネルギーの情報または個別のエネルギー情報、残存エネルギーの情報または残存エネルギー情報、タイルエネルギーの情報またはタイルエネルギー情報、あるいは欠落エネルギーの情報または欠落エネルギー情報は、エネルギー値だけでなく、振幅値（例えば絶対値）、レベル値、あるいは他の値であって、そこから最終的なエネルギー値を導き出すことのできるどのような値も含むことができる。したがって、エネルギーの情報は、例えば、エネルギー値それ自身、および/または、レベルおよび/または振幅および/または絶対振幅の値を含むことができる。

【0119】

図12Aは、復号装置のさらなる実施形態を示す。ビットストリームを、例えばAACデコーダのようなコアデコーダ1200により受け取る。結果は、例えば周波数再生器604に相当する帯域幅拡張パッチングまたはタイリング1202の実行段に送られる。そして、パッチ/タイル適応および後処理が実行され、パッチ適応が実行されると、周波数再生器1202が、さらなる周波数再生を実行するように制御されるが、今回は、例えば調整された周波数境界を用いる。さらに、パッチ処理がトーン線の除去または減衰などにより実行されるとき、その結果が、続いてブロック1206に送られ、例えばブロック712または826を参照して説明したように、パラメータ駆動帯域包絡線整形が実行される。その結果は、続いて、例えば図12Aに示すPCM出力ドメインのような最終

10

20

30

40

50

出力ドメインへの変換を行う合成変換ブロック 1208 に送られる。

【0120】

本発明の実施形態の主な特徴は、以下の通りである。

【0121】

望ましい実施形態は M D C T に基づくものであり、M D C T は、トーンスペクトル領域がクロスオーバー周波数および / またはセパッチマージンの不運な選択により切り取られた場合、またはトーン成分がパッチ境界近傍に近接しすぎて配置された場合に、上述したさえずりアーティファクトを生じる。

【0122】

図 12B は、新たに提案される技術が、最新の B W E 法で生じるアーティファクトをいかに削減するかを示す。図 12B のパネル 2 において、従来の B W E 法の出力の定型化されたマグニチュードスペクトルを示す。この例では、信号は、2 つの隣接トーンによるうなり (b e a t i n g) により、また、トーンの分離により、知覚的に劣化している。双方の問題のあるスペクトル領域は、それぞれ円により印付けられている。

【0123】

これらの問題を解決するため、この新しい技術は、最初に、信号に含まれるトーン成分のスペクトル位置を検出する。次に、本発明のひとつの見地によると、L F とすべてのパッチとの間の遷移周波数を、トーン成分の分離またはうなりが最小化されるように、個別のシフト (所定の限度内で) により調整しようとする。この目的のため、遷移周波数が、局所スペクトルの極小に一致することが望ましい。このステップは、図 12B のパネル 2 およびパネル 3 に示し、ここで、遷移周波数 f_{x2} がより高い周波数に向かってシフトさせられ、 f'_{x2} となっている。

【0124】

本発明の他の見地によると、遷移領域に問題のあるスペクトルコンテンツが残っている場合には、少なくともひとつの間違って置かれたトーン成分を除去して、遷移周波数のうなりアーティファクトまたはさえずりのいずれかを削減する。これは、図 2 パネル 3 に示すように、スペクトル外挿または内挿 / フィルタリングを介して行われる。トーン成分は、これにより、足点から足点、すなわちその左局所極小から右局所極小に取り除かれる。本発明の技術を適用した後の結果スペクトルを図 12B パネル 4 に示す。

【0125】

言い換えると、図 12B は、左上欄、すなわちパネル 1 に、オリジナルの信号を示す。右上欄、すなわちパネル 2 に、長円 1220, 1221 で示された問題のある領域をもつ対照帯域幅拡張信号が示される。左下欄、すなわちパネル 3 に、2 つの望ましいパッチまたは周波数タイル処理が示される。トーン部の分離が周波数境界 f_{x2} を増加させることにより対処され、対応するトーン部のクリッピングが、もはやそこには存在しなくなる。さらに、トーン部 1031 を除去するための利得関数 1030, 1032 が適用され、あるいは、代替的に、1033 で示される内挿が示される。最後に、図 12B の右下欄、すなわちパネル 4 は、一方でのタイル / パッチ周波数調整と問題のあるトーン部の除去または少なくとも減衰との組み合わせから得られる改善された信号を示す。

【0126】

図 12B のパネル 1 は、前述したように、オリジナルのスペクトルを示す。このオリジナルのスペクトルは、クロスオーバーまたはギャップ充填の開始周波数 f_{x1} までのコア周波数を有する。

【0127】

したがって、周波数 f_{x1} は、ソースレンジ 1252 と、境界周波数 1250 とナイキスト周波数 $f_{Nyquist}$ 以下最大周波数との間に延びる再構築レンジ 1254 との間の境界周波数 1250 を示す。エンコード側では、信号は f_{x1} で帯域制限されていると仮定される、あるいは、インテリジェントギャップ充填についての技術が適用される場合には、 f_{x1} が図 3A のギャップ充填開始周波数 309 に相当していると仮定される。この技術に依存して、 f_{x1} を超える再構築レンジは、空となるか (図 13A, 13B の実

10

20

30

40

50

施形態の場合)、あるいは、図3Aを参照して説明するような高解像度で符号化されるべきある第1周波数部からなる。

【0128】

図12B、パネル2は、予備再生成された信号を示す。この信号は、例えば図7Aのブロック702により生成された信号であり、2つの問題のある部分を含む。問題のある部分のひとつを、1220で示す。1220aで示すコア領域内のトーン部と1220bで示す周波数タイルの開始のトーン部との周波数距離は小さすぎ、このため、うなりアーティファクトが作り出される。さらなる問題は、1225で示す最初のパッチ操作または周波数タイル操作により生成された第1周波数タイルの上側境界が、途中でクリップされた、分離したトーン部1226であることである。このトーン部1226を図12Bの他のトーン部と比較するとき、幅が典型的なトーン部の幅より小さいことが明らかであり、これが、このトーン部が、第1周波数タイル1225と第2周波数タイル1227との間の周波数境界が、ソースレンジ1252内の間違った場所に設定されることにより、このトーン部が分離されていることを意味する。この発生に対処するため、図12Bのパネル3に示すように、境界周波数 f_{x2} を修正して少し大きくする。これにより、トーン部のクリッピングは生じない。

10

【0129】

一方、 f'_{x2} が変化させられるこの手続きは、うなりの問題には有効には対処していない。この問題は、したがって、図7Aのブロック708を参照して説明したフィルタリングによるトーン成分の除去または内挿または他のいずれかの手続きにより対処される。図12Bは、遷移周波数調整706と、708で示す境界でのトーン成分の除去との順次適用を示す。

20

【0130】

他の選択肢として、遷移周波数 f_{x1} を少し低く設定して、トーン部1220aがもはやコアレンジには無いようにすることもできる。このとき、トーン部1220aは、遷移周波数 f_{x1} を少し低く設定することにより、除去される。

【0131】

この手続きはまた、問題のあるトーン成分1032の発生への対処にも作用する。 f'_{x2} をさらに高く設定することで、トーン部1032が置かれているスペクトル部が、第1パッチング動作1225内で再生成され、したがって、2つの隣接するトーン部が発生することはない。

30

【0132】

基本的にうなりの問題は、隣接するトーン部の周波数の強度と距離に依存する。検出器704、720、あるいはより一般的にはアナライザ602が、どのトーン成分も配置されるため、 f_{x1} 、 f_{x2} 、 f_{x2} のような遷移周波数より救い周波数に配置される低周波数部を解析するように構成されることが望ましい。さらに、遷移周波数より上のスペクトルレンジも、トーン成分を検出するために解析される。2つのトーン成分、ひとつは周波数に対して遷移周波数の左、もうひとつは右(高くなる周波数)が検出されたとき、図7Aに708で示す境界のトーン成分の除去器が動作する。トーン成分の検出が、ある検出範囲、対応する帯域の帯域幅に対して遷移周波数から両方向に少なくとも20%延びる範囲、望ましくは、対応する帯域の帯域幅に対して遷移周波数の左の低域、および遷移周波数の右に高域に10%だけ延びる範囲、すなわち、一方ではソースレンジの帯域幅で、他方では再構築レンジ、または、遷移周波数が2つの周波数タイル1225、1227の間の遷移周波数の場合には、対応する周波数タイルの10%の範囲で行われる。さらなる実施形態では、あらかじめ定められた検出帯域幅が1パークである。パッチ境界の周囲の1パークの範囲内でトーン部を除去することができ、これにより、完全な検出範囲が2パーク、すなわち、低帯域に1パーク、高帯域に1パークとなる。ここで、低帯域の1パークは、高帯域の1パークに直に隣接する。

40

【0133】

本発明の他の見地によると、フィルタリングアーティファクトの削減のため、周波

50

数領域のクロスオーバーフィルタが、2つの連続するスペクトル領域、すなわちコア帯域と第1パッチとの間、または2つのパッチの間に適用される。クロスオーバーフィルタは、信号適応的であることが望ましい。

【0134】

クロスオーバーフィルタは、2つのフィルタ、低スペクトル領域周に適用されるフェードアウトフィルタ h_{out} と、高スペクトル領域に適用されるフェードインフィルタ h_{in} とからなる。

【0135】

各フィルタは、長さ N を有する。

【0136】

加えて、双方のフィルタの傾斜は、クロスオーバーフィルタのノッチ特性を決定する信号適応値 X_{bias}

により特徴付けられる。ここで、 $0 \leq X_{bias} \leq N$ であり、

$X_{bias} = 0$ のとき、双方のフィルタの合計が1に等しく、そのフィルタにはノッチフィルタ特性はない。

$X_{bias} = N$ のとき、双方のフィルタは完全に零である。

【0137】

【数1】

$$\begin{aligned} h_{out}(k) &= h_{in}(N-1-k), \forall X_{bias} \\ h_{out}(k) + h_{in}(k) &= 1, X_{bias} = 0 \end{aligned}$$

10

20

ここで、 $k = 0, 1, \dots, N-1$ は周波数インデックスである。図12Cは、そのようなクロスオーバーフィルタの一例を示す。

【0138】

この例において、フィルタ h_{out} を生成するために、以下の式を使用する。

【数2】

$$h_{out}(k) = 0.5 + 0.5 \cdot \cos\left(\frac{k}{N-1-X_{bias}} \cdot \pi\right), k = 0, 1, \dots, N-1-X_{bias}$$

【0139】

以下の式が、フィルタ h_{in} 、 h_{out} がどのように適用されるかを説明する。

【数3】

$$\begin{aligned} Y(k_t - (N-1) + k) = & LF(k_t - (N-1) + k) \cdot h_{out}(k) + \\ & HF(k_t - (N-1) + k) \cdot h_{in}(k), \quad k = 0, 1, \dots, N-1 \end{aligned}$$

ここで、 Y は、組み立てられたスペクトル、 k_t は遷移周波数、 LF は低周波数コンテンツ、 HF は高周波数コンテンツである。

【0140】

次に、この技術が役に立つことの説明を提示する。以下の例のオリジナル信号は過渡信号であり、特に、カットオフ周波数 2.2 kHz で低域通過濾波されたものである。最初に、この過渡信号が、変換領域で 6 kHz に帯域制限される。続いて、低域通過濾波されたオリジナル信号の帯域が、 2.4 kHz に拡張される。帯域の拡張は、 LF 帯域を3回コピーすることにより達成され、変換により 6 kHz を超えて得られる周波数レンジを完全に充填する。

【0141】

図11Aは、この信号のスペクトルを示す。このスペクトルは、上述した変換のレンガ壁特性（音声ピーク 1100 ）による、過渡信号をスペクトル的に取り囲むフィルタリングアーティファクトの典型的なスペクトルと考えられる。本発明のアプローチを適用することにより、フィルタリングが、各過渡周波数毎に、ほぼ 20 dB 削減される（音声ピークが削減される）。

40

50

【0142】

同じ効果が、異なる図面の図11B, 11Cにも示される。図11Bは、過渡信号のスペクトル図を示す。この過渡信号は、どのようなフィルタリング削減もしない、上述のBWEを適用した後の、過渡に一時的に前後する上述したフィルタリングアーティファクトを有する。水平線はそれぞれ、連続するパッチの間の遷移周波数におけるフルタリングを示す。図6は、BWEに本発明のアプローチを適用した後の、同じ信号を示す。フィルタリングは、前の図に示した信号に比較し、ほぼ20dB削減されている。

【0143】

次に、図14A, 14Bを参照して、クアナライザ特性について既に説明した、クロスオーバーフィルタの発明の見地についてさらに説明する。しかしながら、クロスオーバーフィルタ710は、図6Aから図7Bを参照して説明した本発明とは独立に実施することもできる。

【0144】

図14Aは、符号化されたコア信号とパラメトリックデータの情報とを含む符号化オーディオ信号を復号する復号装置を示す。この復号装置は、符号化されたコア信号を復号して復号コア信号を得るコアデコーダ1400を備える。復号コア信号が、帯域幅が図13A、図13Bを参照した説明したように帯域制限されていてよく、また、コアデコーダが、図1から図5Cまたは図9Aから図10Dを参照して説明した全周波数範囲すなわちフルレートの符号器であってもよい。

【0145】

さらに、タイル生成器1404が、復号コア信号のスペクトル部を用いて、復号コア信号に含まれない周波数を有する1以上のスペクトルタイルを再生する。タイルは、例えば図3Aを参照して説明したような、再構築帯域に再構築された第2スペクトル部でもよく、または、高解像度で再構築される第1スペクトル部を含むこともできるが、代替的に、エンコーダが図13Aに示すような強い帯域制限を実行するときには、スペクトルタイルは完全に空の周波数帯域でもよい。

【0146】

さらに、クロスオーバーフィルタ1406は、復号コア信号と、ギャップ充填周波数309から第1タイル停止周波数に延びる周波数を有する第1周波数タイルと、を周波数上でクロスオーバーフィルタリングするために、あるいは、第1周波数タイル1225と、第1周波数タイル1225の上側境界周波数に周波数隣接する下側境界周波数を有する第2周波数タイルと、を周波数上でクロスオーバーフィルタリングするために、設けられる。

【0147】

さらなる実施形態では、クロスオーバーフィルタ1406の出力信号が、包絡線調整器1408に入力される。包絡線調整器1408は、復号オーディオ信号に含まれるパラメトリックスペクトル包絡線情報をパラメトリック付帯情報として適用し、最終的に包絡線調整された再生信号を得る。要素1404, 1406, 1408は、例えば図13B、図1Bまたは図6Aに示した周波数再生器として実現される。

【0148】

図14Bは、クロスオーバーフィルタ1406のさらなる実現形態を示す。クロスオーバーフィルタ1406は、第1入力信号IN1を受け取るフェードアウトサブフィルタ1420と、第2入力信号IN2を受け取るフェードインサブフィルタ1422と、を備える。双方のフィルタ1420, 1422の結果、すなわち出力は、結合器1424に供給される。結合器1424は、例えば加算器である。結合器1424は、周波数ビンのための周波数値を出力する。図12Cは、フェードアウトサブフィルタ特性1420aと、フェードインサブフィルタ特性1422aと、を含むクロスフェード関数の一例を示す。2つのフィルタは、ある周波数オーバーラップを有する。図12Cの例では、この周波数オーバーラップは、12に等しい、すなわち $N = 12$ である。したがって、例えばソースレン

ジ 1 2 5 2 などの他の周波数値は、影響を受けない。ソースレンジ 1 2 5 2 の最も高い周波数ピン 2 1 だけが、フェードアウト関数 1 4 2 0 a の影響を受ける。

【 0 1 4 9 】

一方、周波数タイル 1 2 2 5 の最も低い 2 1 周波数線だけは、フェードイン関数 1 4 2 2 の影響を受ける。

【 0 1 5 0 】

さらに、クロスフェード関数から、9 と 1 3 の間の周波数線が影響を受けるが、フェードイン関数は実際には 1 と 9 の間の周波数には影響せず、フェードアウト関数 1 4 2 0 a は 1 3 と 2 1 の間の周波数線に影響しないことが明らかである。これは、オーバーラップが周波数線 9 および 1 3 の間に必要なだけで、 $f_{x,1}$ のようなクロスオーバー周波数は、周波数標本すなわち周波数ピン 1 1 に配置されることになる。したがって、ソースレンジと最初の周波数タイルとの間の 2 つの周波数ピンまたは周波数値のオーバーラップは、クロスオーバーまたはクロスフェード関数を実現するために必要なだけである。

【 0 1 5 1 】

特定の実施形態に依存して、より高いまたはより低いオーバーラップが適用でき、付加的に、コサイン関数ではない他のフェージング関数を用いることもできる。さらに、図 1 2 C に示すように、ある特定のノッチをクロスオーバーレンジ内に適用することが望ましい。別の言い方をすれば、境界領域のエネルギーは、ノッチの無いクロスフェード関数の場合と同様に、双方のフィルタ関数が 1 に加算されることがないという事実により削減される。この周波数タイルすなわち最初の周波数タイルの境界に対するエネルギー損失は、低側境界および高側境界で減衰する。しかしながら、クロスオーバーフィルタによる処理に続いてスペクトル包絡線調整を行うことで、全体の周波数には触れられず、図 3 A を参照して説明した対応するスケールファクタのような、スペクトル包絡線データにより定義される。言い換えると、図 9 B の計算器 9 1 8 がこのとき、「既に生成された生ターゲットレンジ」を計算し、これがクロスオーバーフィルタの出力となる。さらに、内挿によるトン部の除去によるエネルギー損失は、この除去が次に低側タイルエネルギーを引き起こし、完全再構築帯域に対する利得ファクタが高くなるという事実により相殺される。その一方で、しかしながら、クロスオーバー周波数はさらなるエネルギー集中を主端数タイルの中間に引き起こし、これが最終的に、特に図 1 1 A から図 1 1 C を参照して説明した過渡により生じるアーティファクトを有効に削減する。

【 0 1 5 2 】

図 1 4 B は、異なる入力組み合わせを示す。ソース周波数レンジと周波数タイルとの間の境界のフィルタリングのため、入力 1 はコアレレンジの上側スペクトル部であり、入力 2 は、最初の周波数タイル、または単一周波数タイルだけが存在する場合にはそのタイルの、下側スペクトル部である。さらに、入力は最初の周波数タイルでもよく、遷移周波数は最初のタイルの上側周波数境界でもよく、サブフィルタ 1 4 2 2 への入力は、2 番目の周波数タイルの低側部とする。付加的な 3 番目の周波数タイルが存在する場合には、さらなる遷移周波数が、2 番目の周波数タイルと 3 番目の周波数タイルとの間の周波数境界となり、フェードアウトサブフィルタ 1 4 2 0 への入力は、図 1 2 C の特性を使用する場合、フィルタパラメータにより定義されるように、2 番目の周波数タイルの上側周波数レンジであり、フェードインサブフィルタ 1 4 2 2 への入力は、3 番目の周波数タイルの低側部で、図 1 2 C の例では、最も下側の 2 1 スペクトル線となる。

【 0 1 5 3 】

図 1 2 A に示すように、フェードアウトサブフィルタおよびフェードインサブフィルタに対して、等しいパラメータ N を有することが望ましい。しかしながら、これは必須ではない。 N の値は変化させることができ、そのとき、結果として、フィルタ「ノッチ」が下側と上側のレンジ間で非対称となる。さらに、フェードイン/フェードアウト関数は、必ずしも図 1 2 A のように同じ特性である必要はない。代わりに、非対称特性を使用することもできる。

【 0 1 5 4 】

さらに、クロスオーバーフィルタ特性を、信号適応的とすることが望ましい。したがって、信号解析に基づいて、フィルタ特性が適応化される。クロスオーバーフィルタは特に過渡信号に有効であることから、過渡信号が発生しているかが検出される。過渡信号が発生しているときには、図12Cに示すようなフィルタ特性を用いることができる。しかし、非過渡信号が検出されたときには、フィルタ特性を変更して、クロスオーバーフィルタの影響を削減することが望ましい。これは、例えば、Nをゼロに設定するか、あるいはX_{bias}をゼロに設定し、これにより、双方のフィルタの合計を1にする、すなわち、結果として得られるフィルタを無ノッチフィルタ特性とする。しかしながら、望ましくは、急激なフィルタ特性の変化によるアーティファクトを防止するために、パラメータN、X_{bias}を変更して比較的遅い変化特性とすることがよい。さらに、信号がある特定の過渡/トーン性検出器により検出されるようなより急速に変化する場合でも、比較的小さなフィルタ特性変化を許容するために、低域通過フィルタが望ましい。検出器は、図14Aに1405で示される。検出器1405は、タイル生成器1404への入力信号またはタイル生成器1404の出力信号を受け取ることもでき、また、AAC復号化からのショートブロックインディケーションのような過渡/非過渡情報を受け取るため、コアデコーダ1400に接続されることもできる。当然、図12Cに示す物以外のどのようなクロスオーバーフィルタでも、同様に使用することができる。

10

【0155】

このように、過渡検出に基づいて、またはトーン性検出あるいは他のいずれかの信号特性検出に基づいて、クロスオーバーフィルタ1406の特性が、上述のように変更される。

20

【0156】

一部の側面については、エンコーディングまたは復号化の装置の文脈において説明されたが、これらの側面は、対応する方法の記載も示すことは明らかであり、ブロックや装置は、方法ステップまたは方法ステップの特徴に対応する。同様に、方法の観点から説明された側面もまた、対応するブロックもしくは物品または対応する装置の特徴の説明としても機能するものである。一部のまたは全ての方法ステップは、例えば、マイクロプロセッサ、プログラム可能なコンピュータ、または電子回路などのハードウェア装置によって（またはこれを用いて）実行されてもよい。一部の実施形態において、最も重要な方法ステップの1つ以上がこのような装置によって実行されてもよい。

30

【0157】

所定の実施に係る要件次第では、発明の実施形態は、ハードウェアで実行されてもよいし、ソフトウェアで実行されてもよい。実施は、例えばフレキシブルディスク、ハードディスクドライブ(HDD)、DVD、ブルーレイ、CD、ROM、PROM、EPROM、EEPROM（登録商標）、またはフラッシュメモリなどのような、電子的に読み取り可能な制御信号が蓄積されたデジタル蓄積媒体を用いてすることができ、当該方法が実行されるようこれらのデジタル蓄積媒体がプログラム可能なコンピュータシステムと連携する（または連携することできる）。したがって、デジタル蓄積媒体は、コンピュータに読み取り可能である。

【0158】

本発明による一部の実施形態では、電子的に読み取り可能な制御信号を有するデータキャリアを含み、当該キャリアは、開示される方法のいずれかが実施されるよう、プログラム可能なコンピュータシステムと連携することができる。

40

【0159】

概して、本発明の実施形態は、プログラムコードを有するコンピュータプログラム製品として実行されてもよく、このようなプログラムコードは、コンピュータプログラム製品がコンピュータにおいて実行されるとき、上記方法のいずれかを実施するよう動作する。このプログラムコードは、例えば機械で読み取り可能なキャリアに蓄積されてもよい。

【0160】

その他の実施形態においては、開示されるいずれかの方法を実行する、機械で読み取り

50

可能なキャリアに蓄積されたコンピュータプログラムを有する。

【0161】

つまり、本発明に係る方法は、その一実施形態においては、コンピュータプログラムがコンピュータで実行されたとき、開示されるいずれかの方法を実行するプログラムコードを有するコンピュータプログラムとして構成される。

【0162】

本発明に係る方法のさらなる実施形態は、したがって、開示される方法のいずれかを実施するコンピュータプログラムが蓄積されたデータキャリア（またはデジタル蓄積媒体またはコンピュータに読み取り可能な媒体）として構成される。データキャリア、デジタル蓄積媒体、または蓄積媒体は、一般的には有形および/または固定的である。

10

【0163】

本発明に係る方法のさらなる実施形態は、したがって、開示される方法のいずれかを実施するコンピュータプログラムを示すデータストリームまたは信号シーケンスとして構成される。このデータストリームまたは信号シーケンスは、例えば、データコミュニケーション接続（例えばインターネットなど）を介して送られるよう構成されてもよい。

【0164】

その他の実施形態においては、開示されるいずれかの方法を実行するよう構成または適応された処理手段、例えばコンピュータ、プログラム可能な論理機構を有する。

【0165】

その他の実施形態においては、開示されるいずれかの方法を実行するコンピュータプログラムをインストールしたコンピュータを有する。

20

【0166】

本発明によるその他の実施形態においては、開示されるいずれかの方法を実行するコンピュータプログラムを、受信器に対して移動（例えば電子的にまたは選択的に）させるよう構成された装置またはシステムを有する。受信器は、例えばコンピュータや携帯機器、メモリ機器などであってもよい。装置またはシステムは、例えば、コンピュータプログラムを受信器側に移動させるファイルサーバを有してもよい。

【0167】

一部の実施形態においては、開示される方法の機能の一部または全部を実行するために、プログラム可能な論理機構（例えば、フィールドプログラマブルゲートアレイ）を用いてもよい。一部の実施形態においては、開示される方法のいずれかを実行するために、フィールドプログラマブルゲートアレイとマイクロプロセッサとを連携させてもよい。一般に、これらの方法はハードウェア装置によって実行されることが好ましい。

30

【0168】

上述の実施形態は、本発明の原理を単に示すに過ぎない。この明細書に開示する構成や詳細に対して変更または変形を施すことができることは、当業者にとって明らかであると理解される。したがって、次に続く特許請求の範囲によってのみ制限される趣旨であり、実施形態の記載または説明によってなされた特定の詳細によっては、何ら限定されない。

【先行技術文献】

【特許文献】

40

【0169】

【特許文献1】US 8,412,365 B:Liljeryd, Lars; Ekstrand, Per; Henn, Fredrik; Kjoerling, Kristofer: Spectral translation/folding in the subband domain, April 2, 2013.

【非特許文献】

【0170】

【非特許文献1】Dietz, L. Liljeryd, K. Kjoerling and O. Kunz, "Spectral Band Replication, a novel approach in audio coding," in

50

112th AES Convention, Munich, May 2002.

【非特許文献2】Ferreira, D. Sinha, "Accurate Spectral Replacement", Audio Engineering Society Convention, Barcelona, Spain 2005.

【非特許文献3】D. Sinha, A. Ferreira¹ and E. Harinarayanan, "A Novel Integrated Audio Bandwidth Extension Toolkit (ABET)", Audio Engineering Society Convention, Paris, France 2006.

10

【非特許文献4】R. Annadana, E. Harinarayanan, A. Ferreira and D. Sinha, "New Results in Low Bit Rate Speech Coding and Bandwidth Extension", Audio Engineering Society Convention, San Francisco, USA 2006.

【非特許文献5】T. Zernicki, M. Bartkowiak, "Audio bandwidth extension by frequency scaling of sinusoidal partials", Audio Engineering Society Convention, San Francisco, USA 2008.

20

【非特許文献6】J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, 104th AES Convention, Amsterdam, 1998, Preprint 4720.

【非特許文献7】M. Neuendorf, M. Multus, N. Retelbach, et al., MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types, 132nd AES Convention, Budapest, Hungary, April, 2012.

30

【非特許文献8】McAulay, Robert J., Quatieri, Thomas F. "Speech Analysis/Synthesis Based on a Sinusoidal Representation". IEEE Transactions on Acoustics, Speech, And Signal Processing, Vol 34(4), August 1986.

【非特許文献9】Smith, J.O., Serra, X. "PARSHL: An analysis/synthesis program for non-harmonic sounds based on a sinusoidal representation", Proceedings of the International Computer Music Conference, 1987.

40

【非特許文献10】Purnhagen, H.; Meine, Nikolaus, "HILN - the MPEG-4 parametric audio coding tools", Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on, vol.3, no., pp.201,204 vol.3, 2000

【非特許文献11】International Standard ISO/IEC 13818-3, "Generic Coding of Moving Pictures and Associated Audio: Audio", Geneva, 1998.

50

【非特許文献12】M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, Oikawa: "MPEG-2 Advanced Audio Coding", 101st AES Convention, Los Angeles 1996

【非特許文献13】J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, August 1999

10

【非特許文献14】J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, August 1999

【非特許文献15】International Standard ISO/IEC 23001-3:2010, Unified speech and audio coding Audio, Geneva, 2010.

20

【非特許文献16】International Standard ISO/IEC 14496-3:2005, Information technology - Coding of audio-visual objects - Part 3: Audio, Geneva, 2005.

【非特許文献17】P. Ekstrand, "Bandwidth Extension of Audio Signals by Spectral Band Replication", in Proceedings of 1st IEEE Benelux Workshop on MPCA, Leuven, November 2002

【非特許文献18】F. Nagel, S. Disch, S. Wilde, A continuous modulated single sideband bandwidth extension, ICASSP International Conference on Acoustics, Speech and Signal Processing, Dallas, Texas (USA), April 2010

30

【非特許文献19】Daudet, L.; Sandler, M.; "MDCT analysis of sinusoids: exact results and applications to coding artifacts reduction," Speech and Audio Processing, IEEE Transactions on, vol.12, no.3, pp. 302 - 312, May 2004.

40

【図 1 A】

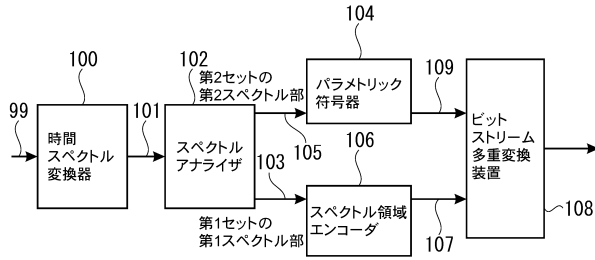


FIG 1A

【図 1 B】

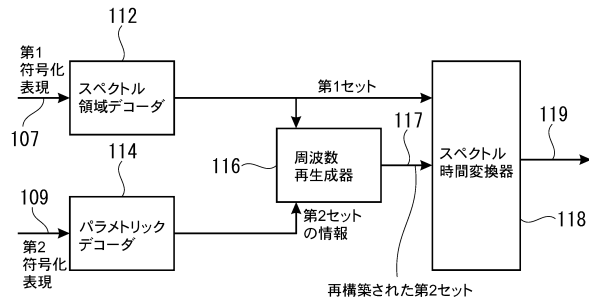


FIG 1B

【図 2 A】

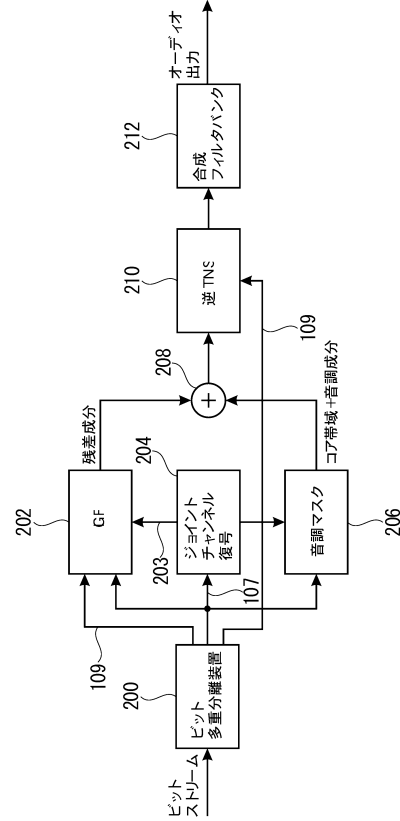


FIG 2A

【図 2 B】

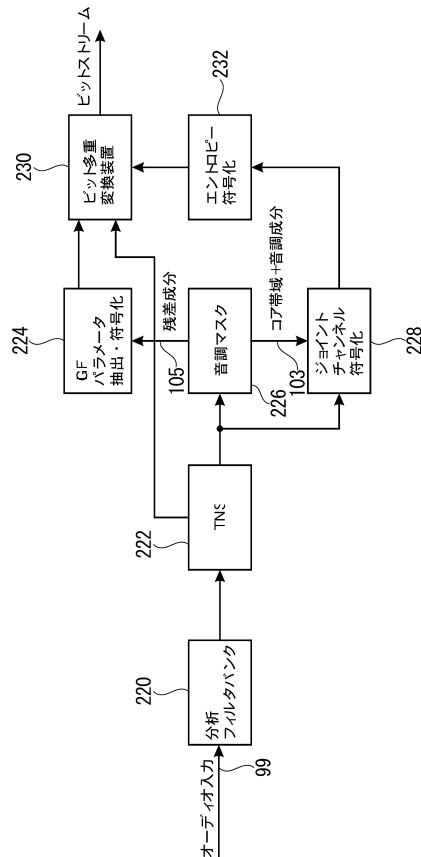


FIG 2B

【図 3 A】

- 第1セットの包絡線の第1解像度（高解像度）（線単位符号化）
- 第2セットの包絡線の第2解像度（低解像度）（SCB毎のスケールファクタ）

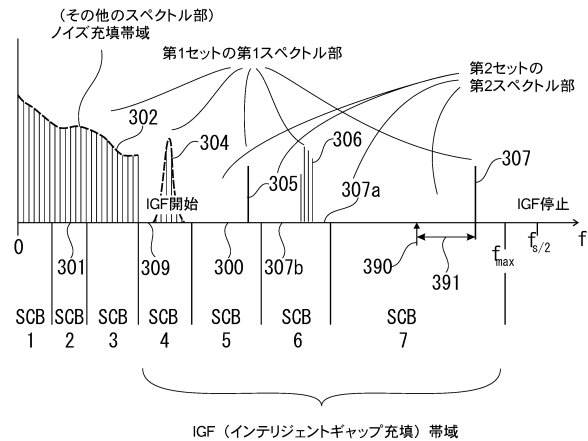


FIG 3A

【図3B】

SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E_1	E_2	E_3	E_4
	NF_2					

308
310
312

FIG 3B

【図4A】

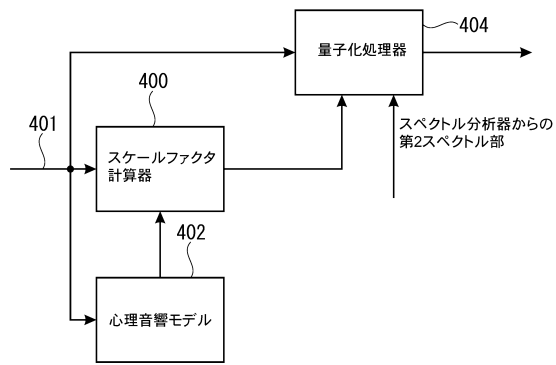
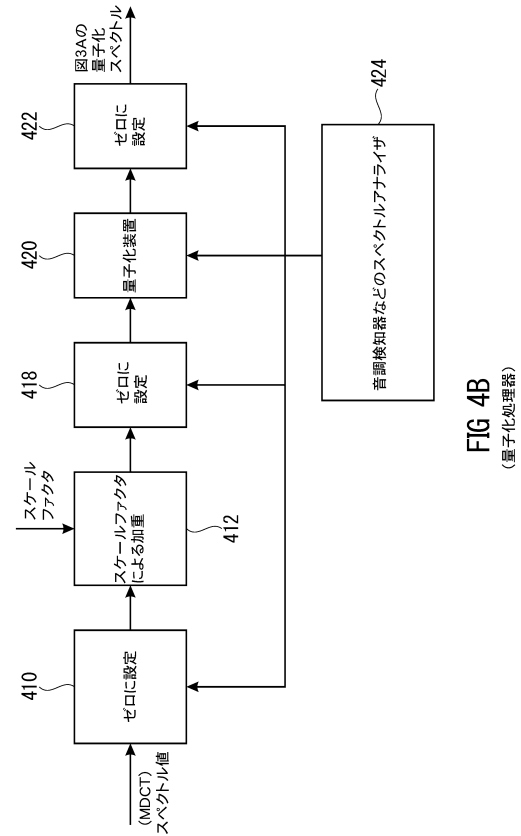
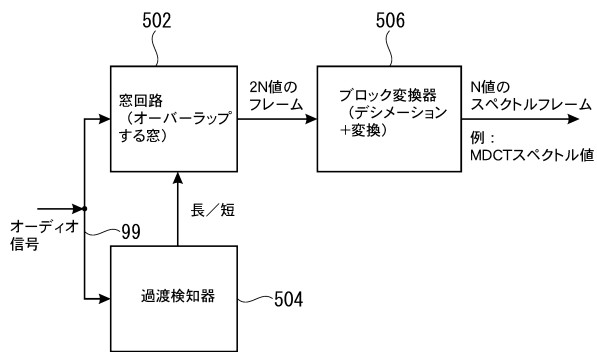


FIG 4A

【図4B】

FIG 4B
(量子化処理器)

【図5A】

FIG 5A
(その他のスペクトル部)

【図5B】

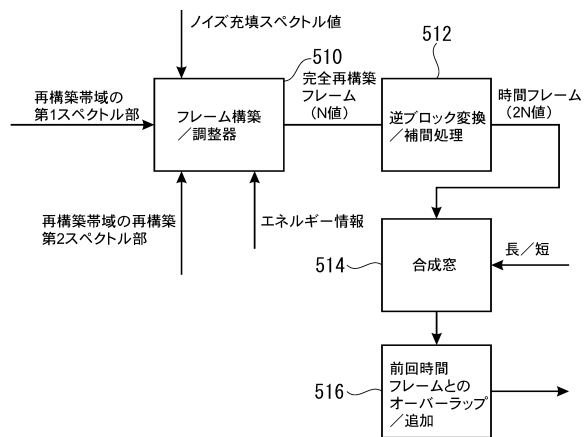


FIG 5B

【図 5 C】

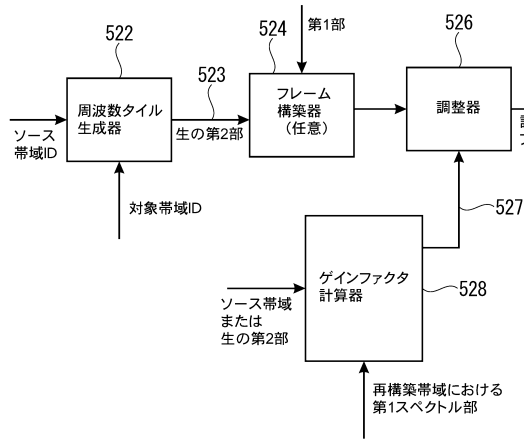


FIG 5C

【図 6 A】

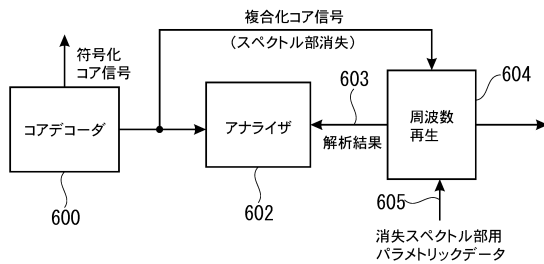


FIG 6A

【図 7 B】

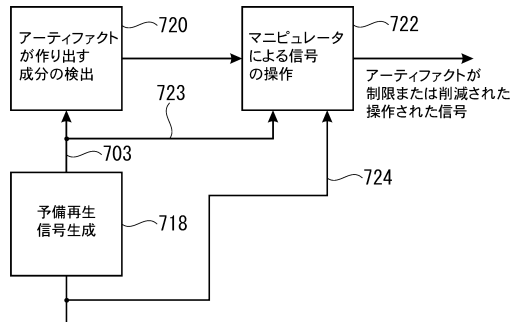


FIG 7B

【図 8 A】

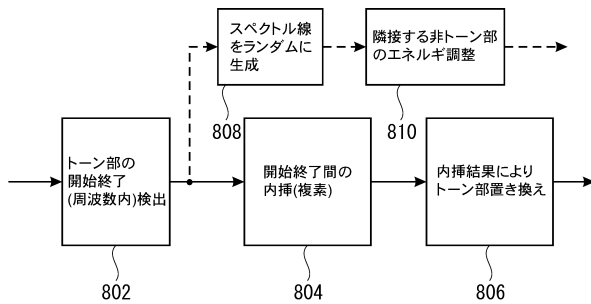


FIG 8A

【図 6 B】

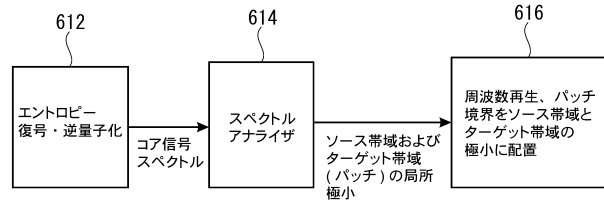


FIG 6B

【図 7 A】

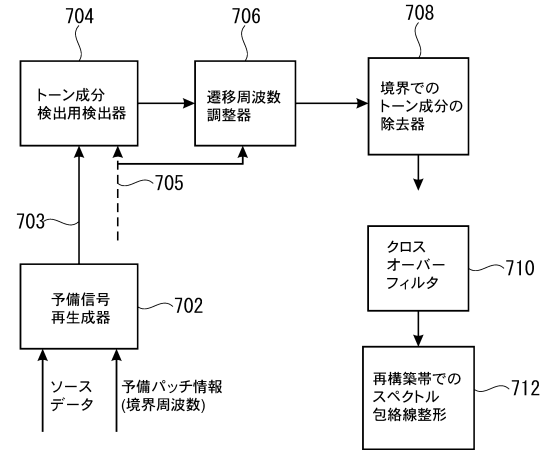


FIG 7A

【図 8 B】

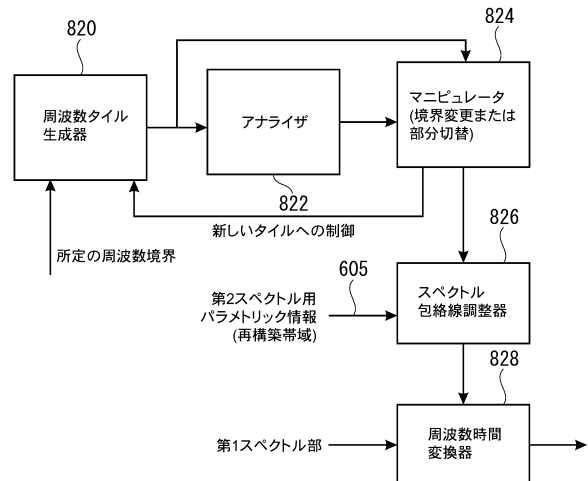


FIG 8B

【図 9 A】

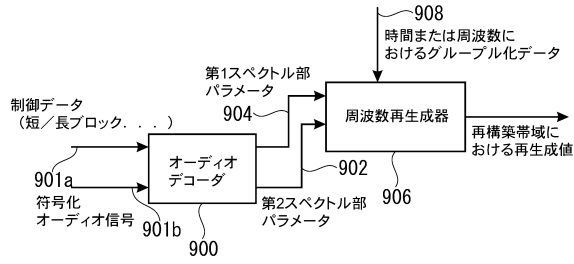


FIG 9A

【図 9 B】

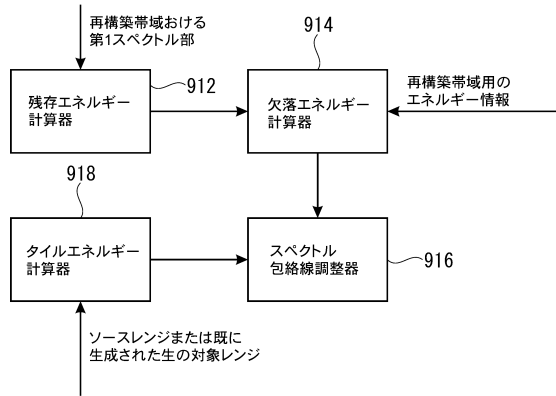


FIG 9B

【図 9 D】

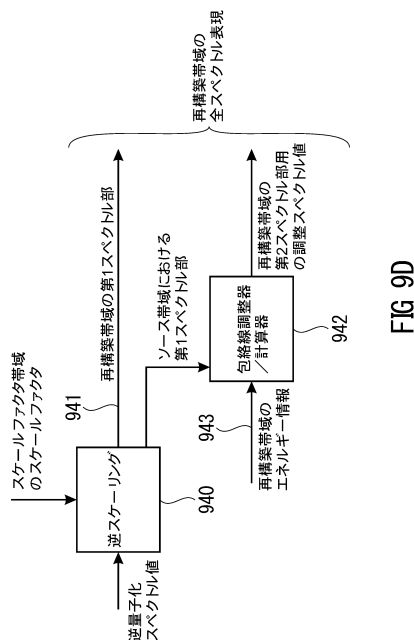
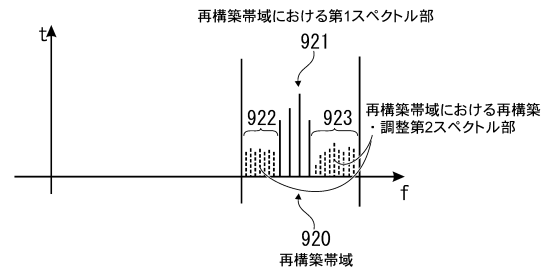


FIG 9D

【図 9 C】



例：● 残存エネルギー

5単位

● 再構築帯域用のエネルギー値

10単位（再構築帯域の第1、第2スペクトル部を覆い隠す）

● ソースレンジデータまたは生の対象レンジデータのエネルギー

8単位

● 欠落エネルギー

5単位

● ゲインファクタ $g := \sqrt{\frac{mE_k}{pE_k}} = 0.79$

→第2スペクトル部のスペクトル値のみが調整される

→第1スペクトル部は包絡線調整による影響を受けない

FIG 9C

【図 10 A】

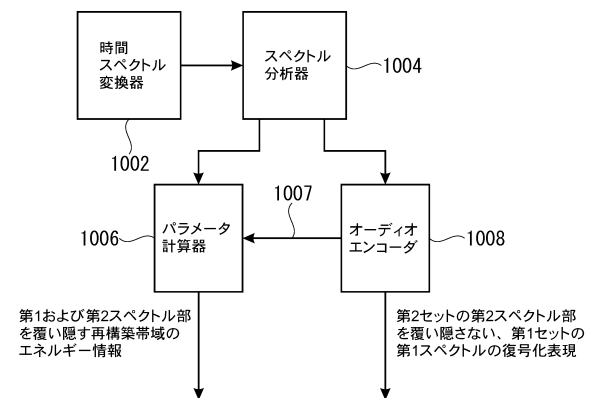


FIG 10A

【図 10 B】

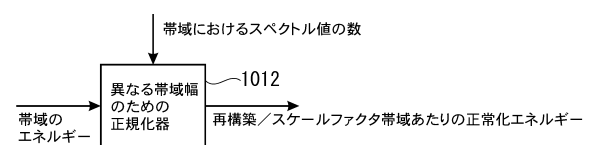


FIG 10B

【図 10 C】

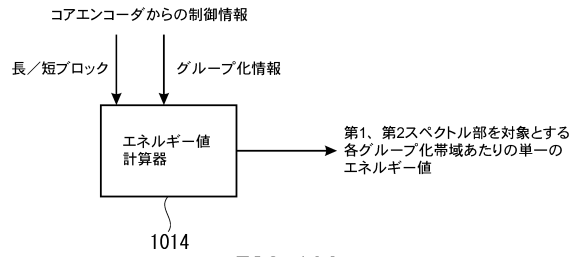


FIG 10C

【図 10 D】

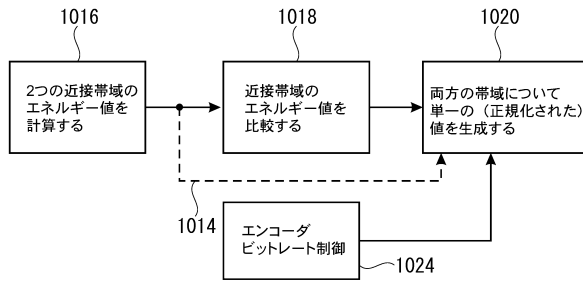


FIG 10D

【図 11 A】

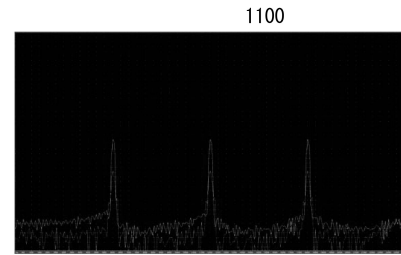
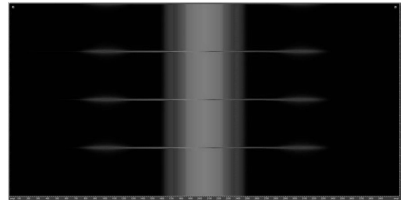


FIG 11A

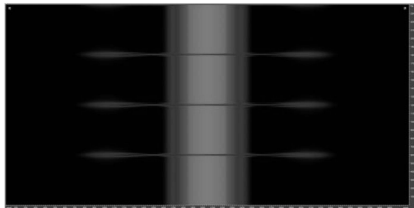
【図 11 B】



BWE適用後の過渡のスペクトル図x軸は時間、y軸は周波数を表す

FIG 11B

【図 11 C】



BWE適用後の過渡のスペクトル図
x軸は時間、y軸は周波数を表す
フィルタリング削減の適用により、フィルタリングが20dB削減されている

FIG 11C

【図 12 B】

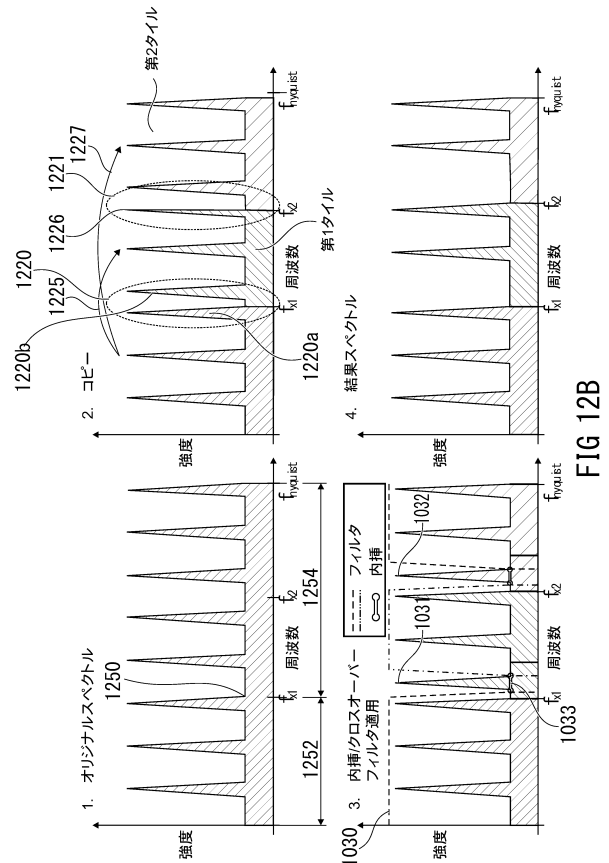


FIG 12B

【図 12 A】

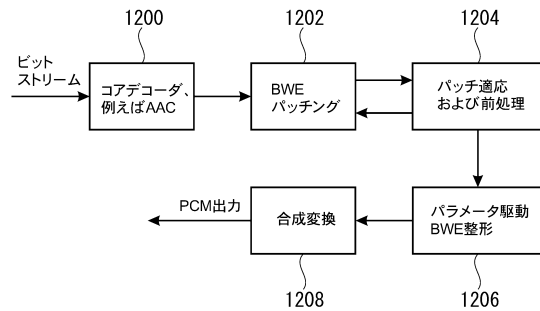
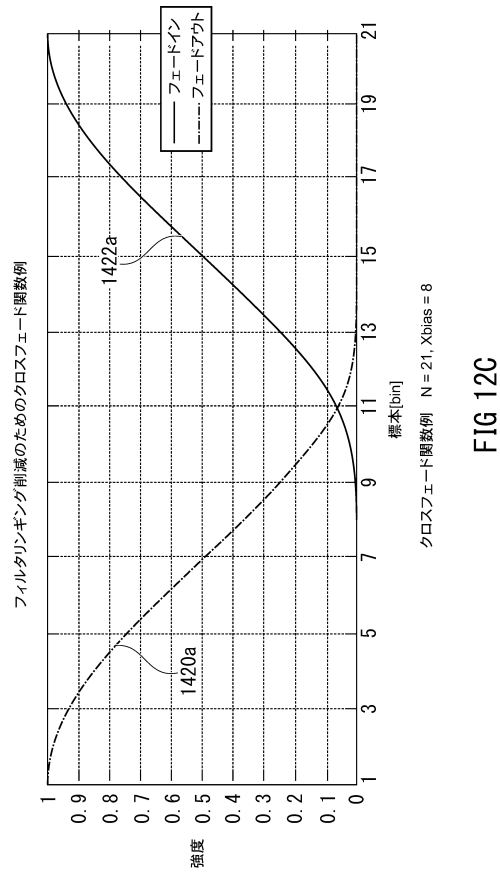
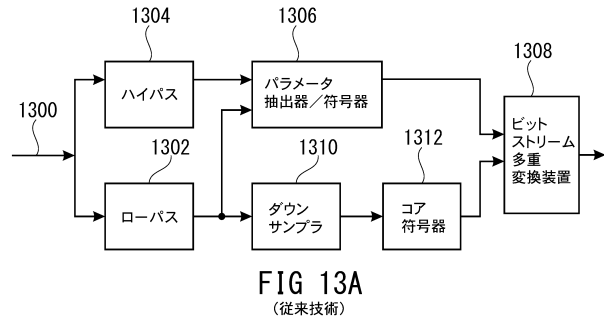


FIG 12A

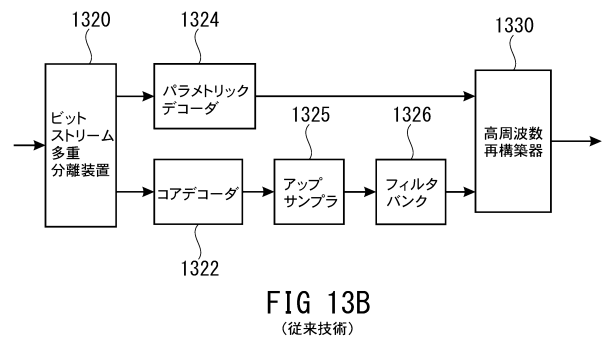
【図 1 2 C】



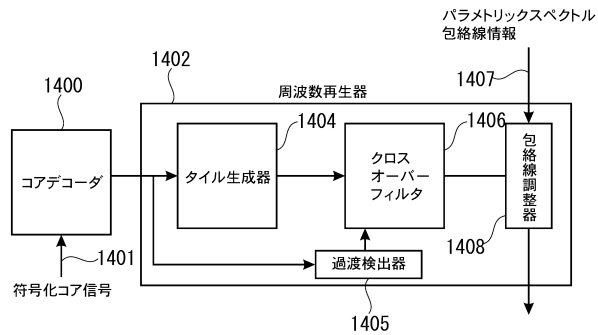
【図 1 3 A】



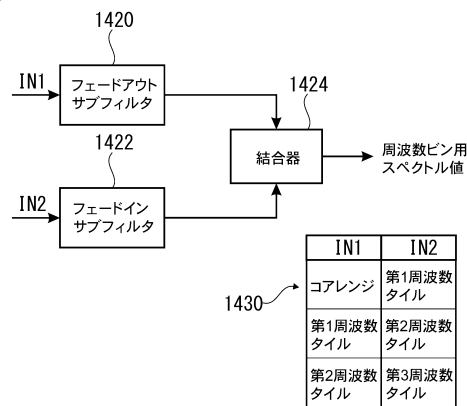
【図 1 3 B】



【図 1 4 A】



【図 1 4 B】



フロントページの続き

- (31)優先権主張番号 13177353.3
(32)優先日 平成25年7月22日(2013.7.22)
(33)優先権主張国 欧州特許庁(EP)
(31)優先権主張番号 13177348.3
(32)優先日 平成25年7月22日(2013.7.22)
(33)優先権主張国 欧州特許庁(EP)
(31)優先権主張番号 13189382.8
(32)優先日 平成25年10月18日(2013.10.18)
(33)優先権主張国 欧州特許庁(EP)

- (72)発明者 ヘルムリッヒ, クリスティアン
ドイツ連邦共和国, 9 1 0 5 4 エアランゲン, ハオプトシュトラーセ 6 8
(72)発明者 ナゲル, フレデリク
ドイツ連邦共和国, 9 0 4 2 5 ニュルンベルク, ヴィルヘルムシャベナー・シュトラーセ 7 2
(72)発明者 ノイカム, クリスティアン
ドイツ連邦共和国, 9 0 5 6 2 カルヒロイト, ヴェイスガーセ 2 4
(72)発明者 シュミット, コンスタンティン
ドイツ連邦共和国, 9 0 4 8 9 ニュルンベルク, ヘアーヴァーゲンシュトラーセ 2 1
(72)発明者 フィッシャー, ミヒャエル
ドイツ連邦共和国, 9 1 0 5 4 エアランゲン, ハークシュトラーセ 2 4

審査官 菊池 智紀

- (56)参考文献 特開2003-108197(JP, A)
特表2013-521538(JP, A)
特表2010-538318(JP, A)
国際公開第2010/070770(WO, A1)
米国特許出願公開第2004/0028244(US, A1)

- (58)調査した分野(Int.Cl., DB名)
G10L 19/00 - 19/26,
21/038 - 21/0388