

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7675295号
(P7675295)

(45)発行日 令和7年5月12日(2025.5.12)

(24)登録日 令和7年4月30日(2025.4.30)

(51)国際特許分類

FI

G 1 0 L 21/0388(2013.01)

G 1 0 L 21/0388 1 0 0

G 1 0 L 19/00 (2013.01)

G 1 0 L 19/00 3 3 0 B

請求項の数 9 外国語出願 (全41頁)

(21)出願番号	特願2025-16475(P2025-16475)	(73)特許権者	510185767
(22)出願日	令和7年2月4日(2025.2.4)		ドルビー・インターナショナル・アーベ
(62)分割の表示	特願2024-81435(P2024-81435)の分割		ー
原出願日	平成31年4月25日(2019.4.25)		アイルランド ディー02 ヴィーケイ6
(65)公開番号	特開2025-66171(P2025-66171A)		0 ダブリン グランド カナル ドックラ
(43)公開日	令和7年4月22日(2025.4.22)		ンズ サー ジョン ロジャーソンズ キー
審査請求日	令和7年2月7日(2025.2.7)	(74)代理人	77 ブロック シー
(31)優先権主張番号	62/662,296		110004381
(32)優先日	平成30年4月25日(2018.4.25)	(72)発明者	弁理士法人ITOH
(33)優先権主張国・地域又は機関	米国(US)		ショエルリング, クリストフェル
早期審査対象出願			アメリカ合衆国 94103 カリフォル
			ニア州 サンフランシスコ マーケット
			ストリート 1275 ドルビー ラボラ
		(72)発明者	トリーズ インコーポレイテッド 内
			ヴィレモエス, ラース
			最終頁に続く

(54)【発明の名称】 後処理遅延低減との高周波再構成技術の統合

(57)【特許請求の範囲】

【請求項1】

オーディオ信号の高周波再構成を実行する方法であって、当該方法は、

符号化されたオーディオビットストリームを受信し、該符号化されたオーディオビットストリームは、前記オーディオ信号の低帯域部分を表すオーディオデータと、高周波再構成メタデータとを含み、該高周波再構成メタデータは、エンベロープスケールファクタを含み、

前記オーディオデータを復号して、復号された低帯域オーディオ信号を生成し、

前記符号化されたオーディオビットストリームから前記高周波再構成メタデータを抽出し、前記高周波再構成メタデータは、高周波再構成プロセスのための動作パラメータを含み、該動作パラメータは、前記符号化されたオーディオビットストリームの後方互換拡張コンテナ内に置かれたパッチングモードパラメータを含み、該パッチングモードパラメータの第1の値は、スペクトル変換を指し示し、該パッチングモードパラメータの第2の値は、位相ボコーダ周波数拡散による高調波トランスポジションを指し示し、

前記復号された低帯域オーディオ信号をフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成し、

前記フィルタリングされた低帯域オーディオ信号及び前記高周波再構成メタデータを用いて、前記オーディオ信号の高帯域部分を再生成し、当該再生成することは、前記パッチングモードパラメータが前記第1の値である場合にスペクトル変換を含み、当該再生成することは、前記パッチングモードパラメータが前記第2の値である場合に位相ボコーダ周

波数拡散による高調波トランスポジションを含む、
ことを有し、

前記フィルタリングすること及び前記再生成することは、オーディオチャンネル当たり
3010 サンプルの遅延で後処理動作として実行され、前記スペクトル変換は、適応逆
フィルタリングによって、音調成分と雑音ライク成分との間の比を維持することを有する、
方法。

【請求項2】

前記後方互換拡張コンテナは更に、前記パッチングモードパラメータが前記第1の値に
等しいときに、前記高帯域部分のスペクトルエンベロープの形状における不連続を回避す
るために追加の前処理が使用されるかを指し示すフラグを含み、該フラグの第1の値は、
前記追加の前処理をイネーブルし、該フラグの第2の値は、前記追加の前処理をディセー
ブルする、請求項1に記載の方法。

10

【請求項3】

前記追加の前処理は、線形予測フィルタ係数を用いてプリゲイン曲線を計算すること
を含む、請求項2に記載の方法。

【請求項4】

前記後方互換拡張コンテナは更に、前記パッチングモードパラメータが前記第2の値に
等しいときに、信号適応周波数ドメインオーバーサンプリングが適用されるべきかを指し
示すフラグを含み、該フラグの第1の値は、前記信号適応周波数ドメインオーバーサン
プリングをイネーブルし、該フラグの第2の値は、前記信号適応周波数ドメインオーバー
サンプリングをディセーブルする、請求項1に記載の方法。

20

【請求項5】

前記信号適応周波数ドメインオーバーサンプリングは、過渡信号を含むフレームに対
してのみ適用される、請求項4に記載の方法。

【請求項6】

位相ボコーダ周波数拡散による前記高調波トランスポジションは、毎秒450万演算以
下及び3kワード以下のメモリの推定複雑度で実行される、請求項1に記載の方法。

【請求項7】

命令を含んだ非一時的なコンピュータ読み取り可能媒体であって、前記命令は、プロセ
ッサによって実行されるときに請求項1に記載の方法を実行する、非一時的なコンピュ
ータ読み取り可能媒体。

30

【請求項8】

命令を有した、非一時的なコンピュータ読み取り可能媒体に格納されたコンピュータ
プログラムであって、前記命令は、コンピューティング装置又はシステムによって実行さ
れるときに、該コンピューティング装置又はシステムに、請求項1に記載の方法を実行さ
せる、コンピュータプログラム。

【請求項9】

オーディオ信号の高周波再構成を実行するオーディオ処理ユニットであって、当該オー
ディオ処理ユニットは、

符号化されたオーディオビットストリームを受信する入力インタフェースであり、前記
符号化されたオーディオビットストリームは、前記オーディオ信号の低帯域部分を表すオ
ーディオデータと、高周波再構成メタデータとを含み、該高周波再構成メタデータは、エ
ンベロープスケールファクタを含む、入力インタフェースと、

40

前記オーディオデータを復号して、復号された低帯域オーディオ信号を生成するコア
オーディオデコーダと、

前記符号化されたオーディオビットストリームから前記高周波再構成メタデータを抽出
するデフォーマッタであり、前記高周波再構成メタデータは、高周波再構成プロセスのた
めの動作パラメータを含み、該動作パラメータは、前記符号化されたオーディオビット
ストリームの後方互換拡張コンテナ内に置かれたパッチングモードパラメータを含み、該
パッチングモードパラメータの第1の値は、スペクトル変換を指し示し、該パッチングモ

50

ドパラメータの第2の値は、位相ボコーダ周波数拡散による高調波トランスポジションを指し示す、デフォーマッタと、

前記復号された低帯域オーディオ信号をフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成する分析フィルタバンクと、

前記フィルタリングされた低帯域オーディオ信号及び前記高周波再構成メタデータを用いて、前記オーディオ信号の高帯域部分を再構成する高周波リジェネレータであり、前記再構成することは、前記パッチングモードパラメータが前記第1の値である場合にスペクトル変換を含み、前記再構成することは、前記パッチングモードパラメータが前記第2の値である場合に位相ボコーダ周波数拡散による高調波トランスポジションを含む、高周波リジェネレータと、

を有し、

前記分析フィルタバンク及び前記高周波リジェネレータは、オーディオチャンネル当たり3010サンプルの遅延でポストプロセッサにて実行され、前記スペクトル変換は、適応逆フィルタリングによって、音調成分と雑音ライク成分との間の比を維持することを有する、

オーディオ処理ユニット。

【発明の詳細な説明】

【技術分野】

【0001】

この出願は、2018年4月25日に出願された米国仮特許出願第62／662，296号に対する優先権の利益を主張するものであり、その全体をここに援用する。

【0002】

実施形態は、オーディオ信号処理に関し、より具体的には、基本形式の高周波再構成（high frequency reconstruction；“HFR”）又は強化形式のHFRのいずれがオーディオデータに対して実行されるべきかを指定する制御データを有するオーディオビットストリームの符号化、復号、又はトランスコーディングに関する。

【背景技術】

【0003】

典型的なオーディオビットストリームは、オーディオコンテンツの1つ以上のチャンネルを示すオーディオデータ（例えば、符号化されたオーディオデータ）と、オーディオデータ又はオーディオコンテンツの少なくとも1つの特徴を示すメタデータとの双方を含んでいる。符号化されたオーディオビットストリームを生成するためのよく知られた1つのフォーマットは、MPEG規格ISO／IEC 14496－3：2009に記載された、MPEG－4 Advanced Audio Coding（AAC）フォーマットである。MPEG－4規格において、AACは“Advanced Audio Coding”を表し、HE－AACは“High－Efficiency Advanced Audio Coding”を表す。

【0004】

MPEG－4 AAC規格は、準拠したエンコーダ又はデコーダにどのオブジェクト及び符号化ツールが存在するかを決定するものである幾つかのオーディオプロファイルを規定している。それらのオーディオプロファイルのうちの3つは、（1）AACプロファイル、（2）HE－AACプロファイル、及び（3）HE－AAC v2プロファイルである。AACプロファイルは、AAC low complexity（すなわち、“AAC－LC”）オブジェクトタイプを含む。AAC－LCオブジェクトは、若干の調整を加えられた、MPEG－2 AAC low complexityプロファイルに対応するものであり、スペクトルバンド複製（“SBR”）オブジェクトタイプ及びパラメトリックステレオ（“PS”）オブジェクトタイプのいずれも含まない。HE－AACプロファイルは、AACプロファイルの上位集合であり、SBRオブジェクトタイプを更に含む。HE－AAC v2プロファイルは、HE－AACプロファイルの上位集合であり、PSオブジェクトタイプを更に含む。

10

20

30

40

50

【0005】

SBRオブジェクトタイプはスペクトルバンド複製ツールを含み、これは、知覚的オーディオコーデックの圧縮効率を有意に改善する重要な高周波再構成（“HFR”）符号化ツールである。SBRは、受信器側（例えば、デコーダ内）でオーディオ信号の高周波成分を再構成する。従って、エンコーダは、低周波成分を符号化して送信することを必要とするのみであり、低いデータレートで遥かに高いオーディオ品質を可能にする。SBRは、エンコーダから得られた制御データ及び利用可能な限られた帯域幅の信号からの、データレートを低減させるために以前に切り捨てられた高調波のシーケンスの複製に基づく。音調（tonal）成分と雑音ライク（noise-like）成分との間の比が、適応逆フィルタリングと、オプションでの雑音及び正弦波の付加とによって維持される。MPEG-4 AAC規格において、SBRツールは、スペクトルパッチング（線形変換又はスペクトル変換とも呼ばれる）を実行し、それにおいて、オーディオ信号の送信された低帯域部分からオーディオ信号の高帯域部分に多数の連続した直交ミラーフィルタ（QMF）サブバンドが複製（又は“パッチ”）され、それがデコーダ内で生成される。

10

【0006】

スペクトルパッチング又は線形変換は、例えば比較的低いクロスオーバー周波数を持つ音楽コンテンツなどの、ある一定のオーディオタイプには理想的ではないことがある。従って、スペクトルバンド複製を改善する技術が望まれる。

【発明の概要】

【0007】

20

第1のクラスの実施形態は、符号化されたオーディオビットストリームを復号する方法に関する。当該方法は、符号化されたオーディオビットストリームを受信し、オーディオデータを復号して、復号された低帯域オーディオ信号を生成することを含む。当該方法は更に、高周波再構成メタデータを抽出し、復号された低帯域オーディオ信号を分析フィルタバンクでフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成することを含む。当該方法は更に、オーディオデータに対してスペクトル変換又は高調波トランスポジションのいずれが実行されるべきかを指し示すフラグを抽出し、該フラグに従って、フィルタリングされた低帯域オーディオ信号及び高周波再構成メタデータを用いてオーディオ信号の高帯域部分を再生成することを含む。最後に、当該方法は、フィルタリングされた低帯域オーディオ信号と再生成された高帯域部分と組み合わせて、広帯域オーディオ信号を形成することを含む。

30

【0008】

第2のクラスの実施形態は、符号化されたオーディオビットストリームを復号するオーディオデコーダに関する。当該デコーダは、符号化されたオーディオビットストリームを受信する入力インタフェースであり、符号化されたオーディオビットストリームは、オーディオ信号の低帯域部分を表すオーディオデータを含む、入力インタフェースと、オーディオデータを復号して、復号された低帯域オーディオ信号を生成するコアデコーダと、を含む。当該デコーダはまた、符号化されたオーディオビットストリームから高周波再構成メタデータを抽出するデマルチプレクサであり、高周波再構成メタデータは、オーディオ信号の低帯域部分からオーディオ信号の高帯域部分へと、連続数のサブバンドを線形変換する高周波再構成プロセスのための動作パラメータを含む、デマルチプレクサと、復号された低帯域オーディオ信号をフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成する分析フィルタバンクと、を含む。当該デコーダは更に、符号化されたオーディオビットストリームから、オーディオデータに対して線形変換又は高調波トランスポジションのいずれが実行されるべきかを指し示すフラグを抽出するデマルチプレクサと、該フラグに従って、フィルタリングされた低帯域オーディオ信号及び高周波再構成メタデータを用いて、オーディオ信号の高帯域部分を再生成する高周波リジェネレータと、を含む。最後に、当該デコーダは、フィルタリングされた低帯域オーディオ信号と再生成された高帯域部分と組み合わせて、広帯域オーディオ信号を形成する合成フィルタバンクを含む。

40

50

【0009】

他のクラスの実施形態は、エンハンストスペクトルバンド複製（enhanced spectral band replication；e S B R）処理が実行されるべきかを特定するメタデータを含むオーディオビットストリームを符号化及びトランスコーディングすることに関する。

【図面の簡単な説明】

【0010】

【図1】発明方法の一実施形態を実行するように構成され得るシステムの一実施形態のブロック図である。

【図2】発明オーディオ処理ユニットの一実施形態であるエンコーダのブロック図である。

【図3】発明オーディオ処理ユニットの一実施形態であるデコーダを含み、及びオプションで、それに結合されたポストプロセッサを含むシステムのブロック図である。

【図4】発明オーディオ処理ユニットの一実施形態であるデコーダのブロック図である。

【図5】発明オーディオ処理ユニットの他の一実施形態であるデコーダのブロック図である。

【図6】発明オーディオ処理ユニットの他の一実施形態のブロック図である。

【図7】MPEG-4 AACビットストリーム（それが分割されるセグメントを含む）のブロックの図である。

【発明を実施するための形態】

【0011】

表記及び用語体系

特許請求の範囲中を含め、この開示全体を通して、信号又はデータに“対して”処理を実行するという表現（例えば、信号又はデータをフィルタリングする、スケーリングする、変換する、又はそれに利得を適用する）は、信号又はデータに対して直接的に、あるいは号又はデータの処理されたバージョンに対して（例えば、処理実行前の予備的なフィルタリング又は前処理を受けた信号のバージョンに対して）、処理を実行することを表すよう、広い意味で使用される。

【0012】

特許請求の範囲中を含め、この開示全体を通して、“オーディオ処理ユニット”又は“オーディオプロセッサ”という表現は、オーディオデータを処理するように構成されたシステム、デバイス、又は装置を示すよう、広い意味で使用される。オーディオ処理ユニットの例は、以下に限られないが、エンコーダ、トランスコーダ、デコーダ、コーデック、前処理システム、後処理システム、及びビットストリーム処理システム（ビットストリーム処理ツールとして参照されることもある）を含む。例えば携帯電話、テレビジョン、ラップトップ、及びタブレットコンピュータなど、ほぼ全ての家電製品が、オーディオ処理ユニット又はオーディオプロセッサを含んでいる。

【0013】

特許請求の範囲中を含め、この開示全体を通して、用語“結合する”又は“結合される”は、直接的又は間接的のいずれの接続も意味するよう、広い意味で使用される。従って、第1のデバイスが第2のデバイスに結合する場合、その接続は、直接的な接続を介してであってもよいし、あるいは、他のデバイス及び接続を介する間接的な接続を介してであってもよい。また、他のコンポーネントに一体化された又は他のコンポーネントと一体化されたコンポーネントも互いに結合されている。

【0014】

発明の実施形態の詳細な説明

MPEG-4 AAC規格は、符号化されたMPEG-4 AACビットストリームが、以下のメタデータ、すなわち、ビットストリームのオーディオコンテンツを復号するために（もし適用されるべきであれば）デコーダによって適用されるべき高周波再構成（“HFR”）処理の各タイプを示す、及び／又はそのようなHFR処理を制御する、及び／又はビットストリームのオーディオコンテンツを復号するために使用されるべき少なくとも1つのHFRツールの少なくとも1つの特性若しくはパラメータを示すメタデータ、を含むこ

とを企図している。ここでは、スペクトルバンド複製（“SBR”）での使用に関して MPEG-4 AAC 規格で記述又は言及されているこのタイプのメタデータを表すために、“SBRメタデータ”という表現を使用する。当業者によって理解されるように、SBRはHFRの一形式である。

【0015】

SBRは、好ましくは、デュアルレートシステムとして使用され、基礎となるコーデックが、元のサンプリングレートの半分で動作する一方で、SBRは、元のサンプリングレートで動作する。SBRエンコーダは、より高いサンプリングレートではあるが、基礎となるコアコーデックと並列に動作する。SBRは主に、デコーダにおける後処理であるが、デコーダにおける最も正確な高周波再構成を確保するために、重要なパラメータがエンコーダで抽出される。エンコーダは、現在の入力信号セグメント特性に適した時間及び周波数レンジ／解像度に関して、SBRレンジのスペクトルエンベロープを推定する。スペクトルエンベロープは、複素QMF解析とその後のエネルギー計算によって推定される。スペクトルエンベロープの時間及び周波数解像度は、所与の入力セグメントに最も適した時間周波数解像度を確保するために、高い自由度で選択されることができる。エンベロープ推定は、元の、主に高周波領域に位置した、過渡成分（例えば、ハイハット）が、エンベロープ調整前のSBR生成された高帯域において僅かに存在することを考慮する必要がある。何故なら、デコーダにおける高帯域は、高帯域と比較して過渡成分が遥かに目立たない低帯域に基づくからである。この側面は、他のオーディオ符号化アルゴリズムで使用される通常のスペクトルエンベロープ推定と比較して異なる要件を、スペクトルエンベロープデータの時間周波数解像度に関して課す。

【0016】

スペクトルエンベロープはさておき、異なる時間及び周波数領域の入力信号のスペクトル特性を表す幾つかの更なるパラメータが抽出される。エンコーダは当然に、元の信号だけでなく、特定の制御パラメータセットを所与としてデコーダ内のSBRユニットがどのようにして高帯域を作成するかについての情報へのアクセスを有するので、システムが以下の状況を取り扱うことが可能であり、すなわち、低帯域が強い高調波系列を構成し、再作成される高帯域が主にランダム信号成分を構成する状況、及び、高帯域領域が基礎とする低帯域内には対応物がない強い音調成分が元の高帯域内に存在する状況を取り扱うことが可能である。さらに、SBRエンコーダは、所与の時点においてどの周波数レンジがSBRによってカバーされるべきかを調べるために、基礎となるコアコーデックと密接に関わって動作する。SBRデータは、エントロピー符号化、及びステレオ信号の場合に制御データのチャンネル依存性、を利用することによって、伝送前に効率的に符号化される。

【0017】

制御パラメータ抽出アルゴリズムは典型的に、所与のビットレート及び所与のサンプリングレートで、基礎となるコーデックに合わせて注意深く調整される必要がある。これは、より低いビットレートは、通常、より高いビットレートと比較して大きいSBRレンジを意味し、且つ異なるサンプリングレートは、SBRフレームの異なる時間解像度に対応する、という事実による。

【0018】

SBRデコーダは、典型的に、幾つかの異なるパーツを含む。それは、ビットストリーム復号モジュール、高周波再構成（HFR）モジュール、追加の高周波成分モジュール、及びエンベロープ調整モジュールを含む。システムは、複素数値のQMFフィルタバンク（高品質SBR用）又は実数値のQMFフィルタバンク（低電力SBR用）に基づく。発明の実施形態は、高品質SBR及び低電力SBRの双方に適用可能である。ビットストリーム抽出モジュールにて、制御データがビットストリームから読み出されて復号される。ビットストリームからエンベロープデータを読み取る前に、現在フレーム用に時間周波数グリッドが取得される。基礎となるコアデコーダが、（低い方のサンプリングレートではあるが）現在フレームのオーディオ信号を復号して、時間ドメインオーディオサンプルを生成する。結果として得られた、オーディオデータのフレームが、HFRモジュールによ

る高周波再構成に使用される。次いで、復号された低帯域信号が、QMFフィルタバンクを用いて解析される。続いて、QMFフィルタバンクのサブバンドサンプルに対して高周波再構成及びエンベロープ調整が実行される。高周波は、所与の制御パラメータに基づいて、柔軟なやり方で低帯域から再構成される。さらに、再構成された高帯域は、所与の時間／周波数領域の適切なスペクトル特性を保証するために、制御データに従ってサブバンドチャンネルベースで適応的にフィルタリングされる。

【0019】

MPEG-4 AACビットストリームの最上位レベルは、データブロック（“raw_data_block”要素）のシーケンスであり、その各々が、オーディオデータ（典型的に1024又は960サンプルの期間にわたる）並びに関連情報及び／又は他のデータを含むデータのセグメント（ここでは、“ブロック”として参照する）である。ここでは、用語“ブロック”を、1つの（1つより多くない）“raw_data_block”要素を決める又は示すオーディオデータ（並びに、対応するメタデータ、及びオプションで他の関連データも）を有するMPEG-4 AACビットストリームのセグメント表すために使用する。

【0020】

MPEG-4 AACビットストリームの各ブロックは、ある数の構文要素を含むことができる（それらの各々も、データのセグメントとしてビットストリーム内に具現化される）。そのような構文要素の7つのタイプが、MPEG-4 AAC規格で定義されている。各構文要素は、データ要素“id_syn_ele”の異なる値によって識別される。構文要素の例は、“single_channel_element()”、“channel_pair_element()”、“fill_element()”を含む。単一のチャンネル要素は、単一のオーディオチャンネル（モノラルオーディオ信号）のオーディオデータを含むコンテナである。チャンネルペア要素は、2つのオーディオチャンネル（すなわち、ステレオオーディオ信号）のオーディオデータを含む。

【0021】

充填要素は、識別子（例えば、上記の要素“id_syn_ele”の値）と、それに続くデータ（“充填データ”として参照する）とを含む情報のコンテナである。充填要素は、歴史的に、一定速度のチャンネル上で伝送されるべきビットストリームの瞬時的なビットレートを調整するために使用されてきた。各ブロックに適量の充填データを加えることによって、一定のデータレートが達成され得る。

【0022】

発明の実施形態によれば、充填データは、ビットストリームで伝送されることが可能なデータ（例えば、メタデータ）のタイプを拡張する1つ以上の拡張ペイロードを含み得る。新たなタイプのデータを含む充填データを有するビットストリームを受信するデコーダは、オプションで、装置の機能を拡張するためにこのビットストリームを受信する装置（例えば、デコーダ）によって使用され得る。従って、当業者によって理解され得るように、充填要素は、特殊なタイプのデータ構造であり、オーディオデータ（例えば、チャンネルデータを含んだオーディオペイロード）を伝送するのに典型的に使用されるデータ構造とは異なる。

【0023】

発明の一部の実施形態において、充填要素を識別するのに使用される識別子は、0x6の値を持った、最上位ビット（“uimsbf”）が先に伝送される3ビット符号なし整数で構成され得る。1つのブロック内で、同じタイプの構文要素の幾つかのインスタンス（例えば、幾つかの充填要素）が発生してもよい。

【0024】

オーディオビットストリームを符号化するための別の標準は、MPEG-USAC（統一オーディオ及びオーディオ符号化（Unified Speech and Audio Coding）規格（ISO/IEC 23003-3:2012））である。MPEG-USAC規格は、スペクトルバンド複製処理（MPEG-4 AAC規格に記載されているSBR処理を含むと

10

20

30

40

50

ともに、他の強化された形式のスペクトルバンド複製処理も含む）を用いたオーディオコンテンツの符号化及び復号を記述している。この処理は、MPEG-4 AAC規格に記載されているSBRツールセットの拡張・強化バージョンのスペクトルバンド複製ツール（ここでは“エンハンストSBRツール”又は“eSBRツール”として参照することもある）を適用する。従って、eSBR（USAC規格で定義されている）は、SBR（MPEG-4 AAC規格で定義されている）の改良である。

【0025】

ここでは、“エンハンストSBR処理”（又は“eSBR処理”）という表現を、MPEG-4 AAC規格では記述又は言及されていない少なくとも1つのeSBRツール（例えば、MPEG USAC規格で記述又は言及されている少なくとも1つのeSBRツール）を用いたスペクトルバンド複製処理を表すために使用する。このようなeSBRツールの例は、高調波（ハーモニック）トランスポジション並びにQMFパッチングによる追加の前処理又は“プレフラット化（pre-flattening）”である。

【0026】

整数次数Tの高調波トランスポーザは、信号持続時間を維持しながら、周波数 ω の正弦波を周波数 $T\omega$ の正弦波へとマッピングする。可能な最小のトランスポジション次数を用いて所望の出力周波数レンジの各部分を生成するために、典型的に、 $T=2, 3, 4$ の3つの次数が順に使用される。4次より上のトランスポジションレンジの出力が必要とされる場合、それは周波数シフトによって生成され得る。可能であるとき、計算の複雑さを最小化する処理のために、略クリティカルにサンプリングされたベースバンド時間ドメインが作成される。

【0027】

高調波トランスポーザは、QMFベース又はDFTベースのいずれであってもよい。QMFベースの高調波トランスポーザを使用するとき、コアコーダ時間ドメイン信号の帯域幅拡張が、改良位相ボコーダ構造を用いてQMFドメイン内で完全に実行され、全てのQMFサブバンドに対してデシメーションとそれに続く時間伸長を実行する。幾つかのトランスポジションファクタ（例えば、 $T=2, 3, 4$ ）を用いるトランスポジションが、共通のQMF分析／合成変換ステージで実行される。QMFベースの高調波トランスポーザは信号適応周波数ドメインオーバーサンプリングを特徴としないので、ビットストリーム内の対応するフラグ（`sbrOversamplingFlag[ch]`）は無視され得る。

【0028】

DFTベースの高調波トランスポーザを使用するとき、好ましくは、複雑さを低減させるために、ファクタ3及び4のトランスポーザ（3次及び4次のトランスポーザ）が、補間によってファクタ2のトランスポーザ（2次のトランスポーザ）に統合される。各フレーム（`coreCoderFrameLength`コアコーダサンプルに対応する）に対して、まず、公称“フルサイズ”の変換サイズの変換サイズのトランスポーザが、ビットストリーム内の信号適応周波数ドメインオーバーサンプリングフラグ（`sbrOverSamplingFlag[ch]`）によって決定される。

【0029】

`sbrPitchingMode==1`であるとき、高帯域を生成するために線形トランスポジションが使用されるべきことを指し示しており、後続のエンベロープ調整器に入力される高周波信号のスペクトルエンベロープの形状における不連続を回避するために追加のステップが導入され得る。これは、続くエンベロープ調整ステージの処理を改善し、より安定しているように感じられる高帯域信号をもたらす。この追加の前処理の動作は、高周波再構成に使用される低帯域信号の粗いスペクトルエンベロープが大きいレベル変動を示す信号タイプにとって有益である。しかしながら、ビットストリーム要素の値は、何らかの種類の信号依存分類を適用することによってエンコーダで決定され得る。この追加の前処理は、好ましくは、1ビットのビットストリーム要素である`bs_sbr_preprocessing`によってアクティブにされる。`bs_sbr_preprocessing`

10

20

30

40

50

`s i n g` が 1 に設定されるときに、この追加処理がイネーブルされる。`b s _ s b r _ p r e p r o c e s s i n g` がゼロに設定されるときには、この追加の前処理はディセーブルされる。この追加処理は好ましくは、各パッチについて低帯域 $X_{L o w}$ をスケーリングするために高周波発生器によって使用される `p r e G a i n` (プリゲイン) 曲線を利用する。例えば、`p r e G a i n` 曲線は、

【数 1】

$$preGain(k) = 10^{(meanNrg - lowEnvSlope(k))/20}, 0 \leq k < k_0$$

に従って計算されることができ、ただし、 k_0 は、マスター周波数帯域テーブルの最初の QMF サブバンドであり、`lowEnvSlope` は、例えば `polyfit()` など、(最小二乗で) 最もフィットする多項式の係数を計算する関数を使用して計算される。例えば、

【数 2】

$$polyfit(3, k_0, x_lowband, lowEnv, lowEnvSlope);$$

を、(三次多項式を用いて) 使用することができ、ここで、

【数 3】

$$lowEnv(k) = 10 \log_{10} \frac{\varphi_k(0,0)}{numTimeSlots \cdot RATE + 6}, 0 \leq k < k_0$$

であり、ただし、`x_lowband(k) = [0 . . . k_0 - 1]` であり、`numTimeSlot` は、フレーム内に存在する SBR エンベロープタイムスロットの数であり、`RATE` は、タイムスロット当たりの QMF サブバンドサンプルの数を指し示す定数 (例えば、2) であり、 φ_k は、(場合により共分散法から取得され得る) 線形予測フィルタ係数であり、ここで、

【数 4】

$$meanNrg = \frac{\sum_{k=0}^{k_0-1} lowEnv(k)}{k_0}$$

である。

【0030】

MPEG USAC 規格に従って生成されるビットストリーム (ここでは “USAC” ビットストリームとして参照することもある) は、符号化されたオーディオコンテンツを含むとともに、典型的に、USAC ビットストリームのオーディオコンテンツを復号するためにデコーダによって適用される各タイプのスペクトルバンド複製処理を示すメタデータ、及び／又は、そのようなスペクトルバンド複製処理を制御し、且つ／或いは USAC ビットストリームのオーディオコンテンツを復号するのに使用される少なくとも 1 つの SBR ツール及び／又は eSBR ツールの少なくとも 1 つの特性又はパラメータを示す、メタデータ、を含む。

【0031】

ここでは、“エンハンスト SBR メタデータ” (又は “eSBR メタデータ”) という表現を、符号化されたオーディオビットストリーム (例えば、USAC ビットストリーム) のオーディオコンテンツを復号するためにデコーダによって適用される各タイプのスペクトルバンド複製処理を示す、及び／又は、そのようなスペクトルバンド複製処理を制御し、且つ／或いはそのようなオーディオコンテンツを復号するのに使用される少なくとも 1 つの SBR ツール及び／又は eSBR ツールの少なくとも 1 つの特性又はパラメータを示すが、MPEG-4 AAC 規格で記述又は言及されていないメタデータを表すために使用

10

20

30

40

50

する。e S B Rメタデータの一例は、M P E G - 4 A A C規格では記述又は言及されていないがM P E G U S A C規格では記述又は言及されているメタデータ（スペクトルバンド複製処理を指し示す、又はそれを制御する）である。従って、e S B Rメタデータは、ここでは、S B Rメタデータではないメタデータを表し、S B Rメタデータは、ここでは、e S B Rメタデータではないメタデータを表す。

【0032】

U S A Cビットストリームは、S B Rメタデータ及びe S B Rメタデータの双方を含み得る。より具体的には、U S A Cビットストリームは、デコーダによるe S B R処理の実行を制御するe S B Rメタデータと、デコーダによるS B R処理の実行を制御するS B Rメタデータとを含み得る。本発明の典型的な実施形態によれば、e S B Rメタデータ（例えば、e S B R固有の構成データ）は、（本発明に従って）M P E G - 4 A A Cビットストリーム（例えば、S B Rペイロードの終端のs b r _ e x t e n s i o n () コンテナ内）に含まれる。

【0033】

デコーダによる、（少なくとも1つのe S B Rツールを有する）e S B Rツールセットを用いた、符号化されたビットストリームの復号中の、e S B R処理の実行は、符号化中に切り捨てられた高調波のシーケンスの複製に基づいて、オーディオ信号の高周波帯域を再生する。このようなe S B R処理は典型的に、元のオーディオ信号のスペクトル特性を再現するために、生成される高周波帯域のスペクトルエンベロープを調整し、逆フィルタリングを適用し、ノイズ成分及び正弦波成分を加える。

【0034】

発明の典型的な実施形態によれば、e S B Rメタデータは、他のセグメント（オーディオデータセグメント）内に符号化されたオーディオデータも含む符号化されたオーディオビットストリーム（例えばM P E G - 4 A A Cビットストリーム）の複数のメタデータセグメントのうちの1つ以上に含められる（例えば、e S B Rメタデータである少数の制御ビットが含まれる）。典型的に、ビットストリームの各ブロックの少なくとも1つのそのようなメタデータセグメントは、充填要素（充填要素の始まりを指し示す識別子を含む）であり（又は含み）、e S B Rメタデータは、該識別子の後の充填要素に含められる。

【0035】

図1は、システムの要素のうちの1つ以上が本発明の一実施形態に従って構成され得る例示的なオーディオ処理チェーン（オーディオデータ処理システム）のブロック図である。このシステムは、図示のように共に結合される以下の要素、すなわち、エンコーダ1、送達サブシステム2、デコーダ3、及び後処理ユニット4を含んでいる。図示のシステムのバリエーションでは、これらの要素のうちの1つ以上が省略され、あるいは追加のオーディオデータ処理ユニットが含まれる。

【0036】

一部の実装において、エンコーダ1（これはオプションで前処理ユニットを含む）は、オーディオコンテンツを有するP C M（時間ドメイン）サンプルを入力として受け入れ、オーディオコンテンツを示す符号化されたオーディオビットストリーム（M P E G - 4 A A C規格に準拠したフォーマットを持つ）を出力するように構成される。ビットストリームのうち、オーディオコンテンツを示すデータを、ここでは、“オーディオデータ”又は“符号化されたオーディオデータ”として参照することがある。エンコーダが本発明の典型的な実施形態に従って構成される場合、エンコーダから出力されるオーディオビットストリームは、オーディオデータ並びにe S B Rメタデータ（及び典型的に、他のメタデータも）を含む。

【0037】

エンコーダ1から出力される1つ以上の符号化されたオーディオビットストリームは、符号化オーディオ送達サブシステム2にアサートされ得る。サブシステム2は、エンコーダ1から出力された符号化されたビットストリーム各々を格納及び／又は送達するように構成される。エンコーダ1から出力された符号化されたオーディオビットストリームは、

サブシステム 2 によって格納され（例えば、DVD 又は Blu-ray（登録商標）ディスクの形態で）、若しくはサブシステム 2 によって伝送されることができ、又はサブシステム 2 によって格納されるとともに伝送され得る。

【0038】

デコーダ 3 は、サブシステム 2 を介して受信した（エンコーダ 1 によって生成された）符号化された MPEG-4 AAC オーディオビットストリームを復号するように構成される。一部の実施形態において、デコーダ 3 は、ビットストリームの各ブロックから eSBR メタデータを抽出し、そして、ビットストリームをデコードする（抽出した eSBR メタデータを用いて eSBR 処理を実行することによって、を含む）ことで、復号されたオーディオデータ（例えば、復号された PCM オーディオサンプルのストリーム）を生成するように構成される。一部の実施形態において、デコーダ 3 は、ビットストリームから SBR メタデータを抽出し（しかし、ビットストリームに含まれる eSBR メタデータを無視し）、ビットストリームを復号する（抽出した SBR メタデータを用いて SBR 処理を実行することによって、を含む）ことで、復号されたオーディオデータ（例えば、復号された PCM オーディオサンプルのストリーム）を生成するように構成される。典型的に、デコーダ 3 は、サブシステム 2 から受信した符号化されたオーディオビットストリームのセグメントを（例えば、非一時的に）格納するバッファを含む。

【0039】

図 1 の後処理ユニット 4 は、デコーダ 3 からの復号されたオーディオデータのストリーム（例えば、復号された PCM オーディオサンプル）を受け入れ、それに対して後処理を実行するように構成される。後処理ユニットはまた、後処理されたオーディオコンテンツ（又はデコーダ 3 から受信した復号されたオーディオ）を、1 つ以上のスピーカによる再生のためにレンダリングするように構成され得る。

【0040】

図 2 は、発明オーディオ処理ユニットの一実施形態であるエンコーダ（100）のブロック図である。エンコーダ 100 のコンポーネント又は要素のいずれも、1 つ以上のプロセス及び／又は 1 つ以上の回路（例えば、ASIC、FPGA、又は他の集積回路）として、ハードウェアにて、ソフトウェアにて、あるいはハードウェアとソフトウェアとの組み合わせにて実装され得る。エンコーダ 100 は、図示のように接続された、エンコーダ 105、スタッファ／フォーマッタステージ 107、メタデータ生成ステージ 106、及びバッファメモリ 109 を含んでいる。典型的に、エンコーダ 100 は、他のプロセッシング要素（図示せず）も含む。エンコーダ 100 は、入力オーディオビットストリームを、符号化された出力 MPEG-4 AAC ビットストリームに変換するように構成される。

【0041】

メタデータ生成器 106 は、エンコーダ 100 から出力される符号化ビットストリームに、ステージ 107 によって含められるべきメタデータ（eSBR メタデータ及び SBR メタデータを含む）を生成する（及び／又はステージ 107 へと渡す）ように結合及び構成される。

【0042】

エンコーダ 105 は、入力オーディオデータを（例えば、それに対して圧縮を実行することによって）符号化し、得られた符号化されたオーディオを、ステージ 107 から出力される符号化されたビットストリームに含めるために、ステージ 107 にアサートするように結合及び構成される。

【0043】

ステージ 107 は、エンコーダ 105 からの符号化されたオーディオと、生成器 106 からのメタデータ（eSBR メタデータ及び SBR メタデータを含む）とを多重化して、好ましくは、符号化されたビットストリームが、本発明の実施形態のうちの 1 つによって指定されるフォーマットを有するように、ステージ 107 から出力される符号化されたビットストリームを生成するように構成される。

【0044】

バッファメモリ 109 は、ステージ 107 から出力された符号化オーディオビットストリームの少なくとも 1 つのブロックを（例えば、非一時的に）格納するように構成され、そして、符号化されたオーディオビットストリームの一連のブロックが、エンコーダ 100 から送達システムへの出力としてバッファメモリ 109 からアサートされる。

【0045】

図 3 は、発明オーディオ処理ユニットの一実施形態であるデコーダ（200）を含み、及びオプションで、それに結合されたポストプロセッサ（300）を含むシステムのブロック図である。デコーダ 200 及びポストプロセッサ 300 のコンポーネント又は要素のいずれも、1 つ以上のプロセス及び／又は 1 つ以上の回路（例えば、ASIC、FPGA、又は他の集積回路）として、ハードウェアにて、ソフトウェアにて、あるいはハードウェアとソフトウェアとの組み合わせにて実装され得る。デコーダ 200 は、図示のように接続された、バッファメモリ 201、ビットストリームペイロードデフォーマッタ（パーサ）205、オーディオ復号サブシステム 202（“コア”復号ステージ又は“コア”復号サブシステムとして参照することもある）、eSBR 処理ステージ 203、及び制御ビット生成ステージ 204 を有している。典型的に、デコーダ 200 は、他のプロセッシング要素（図示せず）も含む。

【0046】

バッファメモリ（バッファ）201 は、デコーダ 200 によって受信された符号化された MPEG-4 AAC オーディオビットストリームの少なくとも 1 つのブロックを（例えば、非一時的に）格納する。デコーダ 200 の動作にて、ビットストリームの一連のブロックが、バッファ 201 からデフォーマッタ 205 にアサートされる。

【0047】

図 3 の実施形態（又は後述する図 4 の実施形態）についてのバリエーションでは、デコーダではない APU（例えば、図 6 の APU 500）が、図 3 又は図 4 のバッファ 201 によって受信されるのと同じタイプの符号化されたオーディオビットストリーム（すなわち、eSBR メタデータを含む符号化されたオーディオビットストリーム）（例えば、MPEG-4 AAC オーディオビットストリーム）の少なくとも 1 つのブロックを（例えば、非一時的に）格納するバッファメモリ（例えば、バッファ 201 と同じバッファメモリ）を含む。

【0048】

図 3 を再び参照するに、デフォーマッタ 205 は、ビットストリームの各ブロックを逆多重化して、それから SBR メタデータ（量子化されたエンベロープデータを含む）及び eSBR メタデータ（及び典型的に他のメタデータも）を抽出し、少なくとも eSBR メタデータ及び SBR メタデータを eSBR 処理ステージ 203 にアサートし、また典型的に、抽出した他のメタデータを復号サブシステム 202（及びオプションで、制御ビット生成器 204 も）にアサートするように結合及び構成される。デフォーマッタ 205 はまた、ビットストリームの各ブロックからオーディオデータを抽出し、抽出したオーディオデータを復号サブシステム（復号ステージ）202 にアサートするように結合及び構成される。

【0049】

図 3 のシステムはまた、オプションでポストプロセッサ 300 も含む。ポストプロセッサ 300 は、バッファメモリ（バッファ）301 と、バッファ 301 に結合された少なくとも 1 つのプロセッシング要素を含む他のプロセッシング要素（図示せず）とを含む。バッファ 301 は、デコーダ 200 からポストプロセッサ 300 によって受信された復号されたオーディオデータの少なくとも 1 つのブロック（又はフレーム）を格納する。ポストプロセッサ 300 のプロセッシング要素は、バッファ 301 から出力される復号されたオーディオの一連のブロック（又はフレーム）を受信し、それを、復号サブシステム 202（及び／又はデフォーマッタ 205）から出力されるメタデータ及び／又はデコーダ 200 のステージ 204 から出力される制御ビットを用いて適応的に処理するように結合及び構成される。

【0050】

デコーダ200のオーディオ復号サブシステム202は、パーサ205によって抽出されたオーディオデータを復号して（このような復号は“コア”復号処理として参照され得る）、復号されたオーディオデータを生成し、そして、復号されたオーディオデータをeSBR処理ステージ203にアサートするように構成される。この復号は周波数ドメインで実行され、典型的に、逆量子化とそれに続くスペクトル処理とを含む。典型的に、サブシステム202の出力が、時間ドメインの復号されたオーディオデータであるように、サブシステム202における処理の最終ステージが、復号された周波数ドメインのオーディオデータに対して、周波数ドメインー時間ドメイン変換を適用する。ステージ203は、復号されたオーディオデータに、（パーサ205によって抽出された）SBRメタデータ及びeSBRメタデータによって指し示されるSBRツール及びeSBRツールを適用して（すなわち、SBR及びeSBRメタデータを使用して、復号サブシステム202の出力に対してSBR及びeSBR処理を実行して）、デコーダ200から（例えばポストプロセッサ300に）出力される完全に復号されたオーディオデータを生成する。典型的に、デコーダ200は、デフォーマッタ205から出力されるデフォーマットされたオーディオデータ及びメタデータを格納するメモリ（サブシステム202及びステージ203によってアクセス可能）を含み、ステージ203は、SBR及びeSBR処理中に必要に応じてオーディオデータ及びメタデータ（SBRメタデータ及びeSBRメタデータを含む）にアクセスするように構成される。ステージ203におけるSBR処理及びeSBR処理は、コア復号サブシステム202の出力に対する後処理であるとみなされ得る。オプションで、デコーダ200はまた、ステージ203の出力に対してアップミキシングを実行して、デコーダ200から出力される完全に復号され、アップミキシングされたオーディオを生成するように結合及び構成された最終アップミキシングサブシステム（これは、デフォーマッタ205によって抽出されるPSメタデータ及び／又はサブシステム204で生成される制御ビットを用いて、MPEG-4 AAC規格で規定されたパラメトリックステレオ（“PS”）ツールを適用し得る）を含む。あるいは、ポストプロセッサ300が、デコーダ200の出力に対してアップミキシングを実行するように構成される（例えば、デフォーマッタ205によって抽出されるPSメタデータ及び／又はサブシステム204で生成される制御ビットを用いる）。

【0051】

デフォーマッタ205によって抽出されたメタデータに応答して、制御ビット生成器204は制御データを生成することができ、該制御データが、デコーダ200内で（例えば、最終アップミキシングサブシステムにおいて）使用され及び／又はデコーダ200の出力として（例えば、後処理での使用のためにポストプロセッサ300に）アサートされ得る。入力ビットストリームから抽出されたメタデータに応答して（及びオプションで制御データにも応答して）、ステージ204は、eSBR処理ステージ203から出力される復号されたオーディオデータが特定タイプの後処理を受けるべきであることを指し示す制御ビットを生成（及びポストプロセッサ300にアサート）し得る。一部の実装において、デコーダ200は、入力ビットストリームからデフォーマッタ205によって抽出されたメタデータをポストプロセッサ300にアサートするように構成され、そして、ポストプロセッサ300は、メタデータを使用して、デコーダ200から出力される復号されたオーディオデータに対して後処理を実行するように構成される。

【0052】

図4は、発明オーディオ処理ユニットの他の一実施形態であるオーディオ処理ユニット（audio processing unit；“APU”）（210）のブロック図である。APU210は、eSBR処理を実行するようには構成されないレガシーデコーダである。APU210のコンポーネント又は要素のいずれも、1つ以上のプロセス及び／又は1つ以上の回路（例えば、ASIC、FPGA、又は他の集積回路）として、ハードウェアにて、ソフトウェアにて、あるいはハードウェアとソフトウェアとの組み合わせにて実装され得る。APU210は、図示のように接続された、バッファメモリ201、ビットストリームペイ

ロードデフォーマッタ（パーサ）２１５、オーディオ復号サブシステム２０２（“コア”復号ステージ又は“コア”復号サブシステムとして参照することもある）、及びＳＢＲ処理ステージ２１３を有している。典型的に、ＡＰＵ２１０は、他のプロセッシング要素（図示せず）も含む。ＡＰＵ２１０は、例えば、オーディオエンコーダ、デコーダ又はトランスコーダを表し得る。

【００５３】

ＡＰＵ２１０の要素２０１及び２０２は、（図３の）デコーダ２００の同じ番号の要素と同じであり、上でのそれらの説明を繰り返すことはしない。ＡＰＵ２１０の動作にて、ＡＰＵ２１０によって受信された符号化されたオーディオビットストリーム（ＭＰＥＧ－４ ＡＡＣビットストリーム）の一連のブロックが、バッファ２０１からデフォーマッタ２０５にアサートされる。

10

【００５４】

デフォーマッタ２１５は、ビットストリームの各ブロックを逆多重化して、それからＳＢＲメタデータ（量子化されたエンベロープデータを含む）を抽出し及び典型的に他のメタデータも抽出するが、本発明の任意の実施形態に従ってビットストリームに含められ得るｅＳＢＲメタデータは無視するように結合及び構成される。デフォーマッタ２１５は、少なくともＳＢＲメタデータをＳＢＲ処理ステージ２１３にアサートするように構成される。デフォーマッタ２１５はまた、ビットストリームの各ブロックからオーディオデータを抽出し、抽出したオーディオデータを復号サブシステム（復号ステージ）２０２にアサートするように結合及び構成される。

20

【００５５】

ＡＰＵ２１０のオーディオ復号サブシステム２０２は、デフォーマッタ２１５によって抽出されたオーディオデータを復号して（このような復号は“コア”復号処理として参照され得る）、復号されたオーディオデータを生成し、そして、復号されたオーディオデータをＳＢＲ処理ステージ２１３にアサートするように構成される。この復号は周波数ドメインで実行される。典型的に、サブシステム２０２の出力が、時間ドメインの復号されたオーディオデータであるように、サブシステム２０２における処理の最終ステージが、復号された周波数ドメインのオーディオデータに対して、周波数ドメイン－時間ドメイン変換を適用する。ステージ２１３は、復号されたオーディオデータに、（デフォーマッタ２１５によって抽出された）ＳＢＲメタデータによって指し示されるＳＢＲツールを適用して（ｅＳＢＲツールは適用せずに）（すなわち、ＳＢＲメタデータを使用して、復号サブシステム２０２の出力に対してＳＢＲ処理を実行して）、ＡＰＵ２１０から（例えばポストプロセッサ３００に）出力される完全に復号されたオーディオデータを生成する。典型的に、ＡＰＵ２１０は、デフォーマッタ２１５から出力されるデフォーマットされたオーディオデータ及びメタデータを格納するメモリ（サブシステム２０２及びステージ２１３によってアクセス可能）を含み、ステージ２１３は、ＳＢＲ処理中に必要に応じてオーディオデータ及びメタデータ（ＳＢＲメタデータを含む）にアクセスするように構成される。ステージ２１３におけるＳＢＲ処理は、コア復号サブシステム２０２の出力に対する後処理であるとみなされ得る。オプションで、ＡＰＵ２１０はまた、ステージ２１３の出力に対してアップミキシングを実行して、ＡＰＵ２１０から出力される完全に復号され、アップミキシングされたオーディオを生成するように結合及び構成された最終アップミキシングサブシステム（これは、デフォーマッタ２０５によって抽出されるＰＳメタデータを用いて、ＭＰＥＧ－４ ＡＡＣ規格で規定されたパラメトリックステレオ（“ＰＳ”）ツールを適用し得る）を含む。あるいは、ＡＰＵ２１０の出力に対してアップミキシングを実行する（例えば、デフォーマッタ２１５によって抽出されるＰＳメタデータ及び／又はＡＰＵ２１０で生成される制御ビットを用いる）ように、ポストプロセッサが構成される。

30

40

【００５６】

発明方法の異なる実施形態を実行するように、エンコーダ１００、デコーダ２００、及びＡＰＵ２１０の様々な実装が構成される。

【００５７】

50

一部の実施形態によれば、符号化されたオーディオビットストリーム（例えば、MPEG-4 AACビットストリーム）にeSBRメタデータが含まれる（例えば、eSBRメタデータである少数の制御ビットが含まれる）が、レガシーデコーダ（これは、eSBRメタデータを構文解析（パース）したり、eSBRメタデータが関係するeSBRツールを使用したりするようには構成されない）が、eSBRメタデータを無視することができ、それにもかかわらず、eSBRメタデータ又はeSBRメタデータが関係するeSBRツールの使用なしで、典型的には復号オーディオ品質における重大なペナルティなしで、ビットストリームを可能な範囲で復号することができるようにされる。一方で、ビットストリームを構文解析してeSBRメタデータを識別し、そして、eSBRメタデータに応答して少なくとも1つのeSBRツールを使用するように構成されたeSBRデコーダは、少なくとも1つのそのようなeSBRツールを使用することの利益を享受することになる。従って、発明の実施形態は、エンハンストスペクトルバンド複製（eSBR）制御データ又はメタデータを後方互換性のある方法で効率的に伝送する手段を提供する。

10

【0058】

典型的に、ビットストリーム内のeSBRメタデータは、以下のeSBRツール（これらは、MPEG USAC規格にて記述されており、ビットストリームの生成中にエンコーダによって適用されたり適用されなかったりし得る）のうちの1つ以上を示す（例えば、その少なくとも1つの特性又はパラメータを示す）：

- ・高調波トランスポジション、及び
- ・QMFパッチングによる追加の前処理（プレフラット化）。

20

【0059】

例えば、ビットストリームに含まれるeSBRメタデータは、sbrPatchingMode[ch]、sbrOversamplingFlag[ch]、sbrPitchInBins[ch]、及びbs_sbr_preprocessingなるパラメータ（MPEG USAC規格及び本開示に記載されている）の値を示し得る。

【0060】

ここでは、Xは何らかのパラメータであるとして、X[ch]という表記は、そのパラメータが、復号されるべき符号化されたビットストリームのオーディオコンテンツのチャンネル（“ch”）に関係することを表す。単純化のため、[ch]という表現を省略することがあり、また、該当パラメータがオーディオコンテンツのチャンネルに関係すると仮定することがある。

30

【0061】

ここでは、Xは何らかのパラメータであるとして、X[ch][env]という表記は、そのパラメータが、復号されるべき符号化されたビットストリームのオーディオコンテンツのチャンネル（“ch”）のSBRエンベロープ（“env”）に関係することを表す。単純化のため、[env]及び[ch]という表現を省略することがあり、また、該当パラメータがオーディオコンテンツのチャンネルのSBRエンベロープに関係すると仮定することがある。

【0062】

符号化されたビットストリームの復号において、（ビットストリームによって示されるオーディオコンテンツの各チャンネル“ch”の）復号のeSBR処理ステージ中の高調波トランスポジションの実行は、sbrPatchingMode[ch]、sbrOversamplingFlag[ch]、sbrPitchInBinsFlag[ch]、及びsbrPitchInBins[ch]というeSBRメタデータパラメータによって制御される。

40

【0063】

値“sbrPatchingMode[ch]”は、eSBRで使用されるトランスポータタイプを指し示し、sbrPatchingMode[ch]=1は、MPEG-4 AAC規格のセクション4.6.18に記載されている線形トランスポジションパッチン

50

グ（高品質 SBR 又は低電力 SBR のいずれとも使用される）を指し示し、`sbrPatchingMode[ch] = 0` は、MPEG USAC 規格のセクション 7.5.3 又は 7.5.4 に記載されている高調波 SBR パッチングを指し示す。

【0064】

値“`sbrOversamplingFlag[ch]`”は、MPEG USAC 規格のセクション 7.5.3 に記載されている DFT ベースの高調波 SBR パッチングと組み合わせての、eSBR における信号適応周波数ドメインオーバーサンプリングの使用を指し示す。このフラグは、トランスポータで使用する DFT のサイズを制御し、1 は、MPEG USAC 規格のセクション 7.5.3.1 に記載されるように信号適応周波数ドメインオーバーサンプリングがイネーブルされることを指し示し、0 は、MPEG USAC 規格のセクション 7.5.3.1 に記載されるように信号適応周波数ドメインオーバーサンプリングがディセーブルされることを指し示す。

10

【0065】

値“`sbrPitchInBinsFlag[ch]`”は、`sbrPitchInBins[ch]` パラメータの解釈を制御し、1 は、`sbrPitchInBins[ch]` の値が有効であってゼロより大きいことを指し示し、0 は、`sbrPitchInBins[ch]` の値がゼロに設定されることを指し示す。

【0066】

値“`sbrPitchInBins[ch]`”は、SBR 高調波トランスポータにおける外積項の追加を制御する。値 `sbrPitchInBins[ch]` は、範囲 [0, 127] 内の整数値であり、コアコードのサンプリング周波数に作用する 1536 ライン DFT の周波数ビンで測定される距離を表す。

20

【0067】

MPEG-4 AAC ビットストリームが、（単一の SBR チャンネルではなく）それらのチャンネルが結合されない SBR チャンネルペアを示す場合、そのビットストリームは、`sbr__channel__pair__element()` の各チャンネルに対して 1 つずつの、（高調波トランスポジション又は非高調波トランスポジションに関する）上の構文の 2 つのインスタンスを示す。

【0068】

eSBR ツールの高調波トランスポジションは、典型的に、比較的低いクロスオーバー周波数にある復号された音楽信号の品質を改善する。非高調波トランスポジション（すなわち、レガシースペクトルパッチング）は典型的に音声（スピーチ）信号を改善する。従って、特定のオーディオコンテンツを符号化するのにどちらのタイプのトランスポジションが好ましいかに関する決定における出発点は、音楽コンテンツには高調波トランスポジションが使用され、音声コンテンツにはスペクトルパッチングが使用されるとして、音声／音楽検出に応じてトランスポジション方法を選択することである。

30

【0069】

eSBR 処理におけるプレフラット化の実行は、“`bs__sbr__preprocessing`”として知られる 1 ビットの eSBR メタデータパラメータの値によって（この 1 ビットの値に応じてプレフラット化が実行されるか実行されないかのいずれかであるという意味で）制御される。MPEG-4 AAC 規格のセクション 4.6.18.6.3 に記載されている SBR QMF パッチングアルゴリズムが使用されるとき、後プレフラット化のステップは、続くエンベロープ調整器（エンベロープ調整器は eSBR 処理の別のステージを実行する）に入力される高周波信号のスペクトルエンベロープの形状における不連続を回避する努力の一環として、（“`bs__sbr__preprocessing`”パラメータによって指し示されるときに）実行され得る。プレフラット化は、典型的に、続くエンベロープ調整ステージの処理を改善し、より安定しているように感じられる高帯域信号をもたらす。

40

【0070】

上述の eSBR ツール（高調波トランスポジション及びプレフラット化）を示す eSBR

50

RメタデータをMPEG-4 AACビットストリームに含めるための全体的なビットレート要求は、発明の一部の実施形態によれば、eSBR処理を実行するために必要とされる差分の制御データのみが伝送されるので、数百ビット／秒のオーダーであると期待される。この情報は（後述するように）後方互換的に含められるので、レガシーデコーダはこの情報を無視することができる。従って、eSBRメタデータを含めることに伴うビットレートへの悪影響は、以下を含む複数の理由から無視できるものである：

- ・ eSBR処理を実行するために必要とされる差分の制御データのみが伝送される（SBR制御データの同時伝送ではない）ので、（eSBRメタデータを含めることによる）ビットレートペナルティは、ビットレート全体のごく一部である、及び

- ・ SBR関係の制御情報の調整（チューニング）は、典型的に、トランスポジションの詳細に依存しない。制御データがトランスポザの動作に依存する場合の例については、この出願中で後述する。

【0071】

従って、発明の実施形態は、エンハンストスペクトルバンド複製（eSBR）制御データ又はメタデータを後方互換性のある方法で効率的に伝送する手段を提供する。eSBR制御データのこの効率的な伝送は、ビットレートに対する目に見える悪影響を有することなく、発明の態様を採用するデコーダ、エンコーダ、及びトランスコーダにおけるメモリ要求を低減させる。さらに、発明の実施形態に従ってeSBRを実行することに関連する複雑さ及び処理要件も低減される。何故なら、SBRデータは、（eSBRが、後方互換的にMPEG-4 AACコーデックに統合される代わりに、MPEG-4 AACにおける完全に別個のオブジェクトタイプとして扱われる、とした場合にそうであるように同時伝送されずに）一度だけ処理されればよいからである。

【0072】

次に、図7を参照して、本発明の一部の実施形態に従ってeSBRメタデータが含まれるMPEG-4 AACビットストリームのブロック（“raw_data_block”）の要素を記述する。図7は、そのセグメントの一部を示すMPEG-4 AACビットストリームのブロック（“raw_data_block”）の図である。

【0073】

MPEG-4 AACビットストリームのブロックは、オーディオプログラムのオーディオデータを含んだ、少なくとも1つの“single_channel_element（）”（例えば、図7に示す単一チャンネル要素）及び／又は少なくとも1つの“channel_pair_element（）”（図7には特に示していないが、存在してもよい）を含み得る。このブロックはまた、そのプログラムに関係するデータ（例えば、メタデータ）を含む複数の“充填要素”（例えば、図7の充填要素1及び／又は充填要素2）を含み得る。各“single_channel_element（）”は、単一チャンネル要素の始まりを示す識別子（例えば、図7の“ID1”）を含むとともに、マルチチャンネルオーディオプログラムのうちの異なるチャンネルを示すオーディオデータを含むことができる。各“channel_pair_element（）”は、チャンネルペア要素の始まりを示す識別子（図7に示されず）を含むとともに、プログラムの2つのチャンネルを示すオーディオデータを含むことができる。

【0074】

MPEG-4 AACビットストリームのfill_element（ここでは充填要素として参照する）は、充填要素の始まりを示す識別子（図7の“ID2”）と、該識別子の後の充填データとを含む。識別子ID2は、0x6の値を持った、最上位ビット（“uimsbf”）が先に伝送される3ビット符号なし整数で構成され得る。充填データは、extension_payload（）要素（ここでは拡張ペイロードとして参照することもある）を含むことができ、その構文は、MPEG-4 AAC規格の表4.57に示されている。幾つかのタイプの拡張ペイロードが存在し、最上位ビット（“uimsbf”）が先に伝送される4ビット符号なし整数である“extension_type”パラメータを介して識別される。

10

20

30

40

50

【0075】

充填データ（例えば、その拡張ペイロード）は、SBRオブジェクトを示す充填データのセグメントを示すヘッダ又は識別子（例えば、図7の“ヘッダ1”）を含むことができる（すなわち、ヘッダが、MPEG-4 AAC規格において`sbr_extension_data()`として参照される“SBRオブジェクト”タイプを開始する）。例えば、ヘッダ内の`extension_type`フィールドの‘1101’又は‘1110’の値で、スペクトルバンド複製（SBR）拡張ペイロードが特定され、識別子‘1101’が、SBRデータを有する拡張ペイロードを特定し、‘1110’が、SBRデータの正確性を検証する周期的冗長検査（CRC）を備えたSBRデータを有する拡張ペイロードを特定する。

【0076】

ヘッダ（例えば、`extension_type`フィールド）がSBRオブジェクトタイプを開始するとき、SBRメタデータ（MPEG-4 AAC規格では“`sbr_data()`”と呼ばれており、ここでは“スペクトルバンド複製データ”として参照することがある）がヘッダに続き、そして、少なくとも1つのスペクトルバンド複製拡張要素（例えば、図7の充填要素1の“SBR拡張要素”）がSBRメタデータに続くことができる。このようなスペクトルバンド複製拡張要素（ビットストリームの一セグメント）は、MPEG-4 AAC規格では“`sbr_extension()`”コンテナと呼ばれている。スペクトルバンド複製拡張要素は、オプションで、ヘッダ（例えば、図7の充填要素1の“SBR拡張ヘッダ”）を含む。

【0077】

MPEG-4 AAC規格は、スペクトルバンド複製拡張要素が、プログラムのオーディオデータに関するPS（パラメトリックステレオ）データを含むことができることを企図している。MPEG-4 AAC規格は、（図7の“ヘッダ1”がそうであるように）充填要素のヘッダ（例えば、その拡張ペイロードのヘッダ）がSBRオブジェクトタイプを開始し、充填要素のスペクトルバンド複製拡張要素がPSデータを含むときに、充填要素（例えば、その拡張ペイロード）が、スペクトルバンド複製データと、PSデータが充填要素のスペクトルバンド複製拡張要素に含まれることを指し示す値（すなわち、`bs_extension_id=2`）を有する“`bs_extension_id`”パラメータとを含むことを企図している。

【0078】

本発明の一部の実施形態によれば、eSBRメタデータ（例えば、ブロックのオーディオコンテンツに対してエンハンストスペクトルバンド複製（eSBR）処理が実行されるべきかを指し示すフラグ）が、充填要素のスペクトルバンド複製拡張要素に含められる。例えば、このようなフラグは、図7の充填要素1に示されており、図7では、該フラグは、充填要素1の“SBR拡張要素”のヘッダ（充填要素1の“SBR拡張ヘッダ”）の後に生じている。オプションで、このようなフラグ及び追加のeSBRメタデータは、スペクトルバンド複製拡張要素のヘッダの後のスペクトルバンド複製拡張要素（例えば、図7の、SBR拡張ヘッダの後の、充填要素1のSBR拡張要素）に含められる。本発明の一部の実施形態によれば、eSBRメタデータを含む充填要素はまた、充填要素にeSBRメタデータが含まれること及び該当ブロックのオーディオコンテンツに対してeSBR処理が

【0079】

本発明の一部の実施形態によれば、eSBRメタデータは、MPEG-4 AACビットストリームのうち、充填要素のスペクトルバンド複製拡張要素（SBR拡張要素）以外の充填要素（例えば、図7の充填要素2）に含められる。これは何故なら、SBRデータ又はCRCを備えたSBRデータを有する`extension_payload()`を含む充填要素は、他の拡張タイプの如何なる他の拡張ペイロードも含まないからである。従って、eSBRメタデータがそれ自身の拡張ペイロードを格納される実施形態において、eSBRメタデータを格納するために別個の充填要素が使用される。そのような充填要素

は、充填要素の始まりを示す識別子（例えば、図7の“ID2”）と、該識別子の後の充填データとを含む。充填データは、`extension_payload()`要素（ここでは拡張ペイロードとして参照することがある）を含むことができ、その構文は、MPEG-4 AAC規格の表4.57に示されている。充填データ（例えば、その拡張ペイロード）は、eSBRオブジェクトを示すヘッダ（例えば、図7の充填要素2の“ヘッダ2”）を含み（すなわち、このヘッダがエンハンススペクトルバンド複製（eSBR）オブジェクトタイプを開始する）、充填データ（例えば、その拡張ペイロード）は、該ヘッダの後にeSBRメタデータを含む。例えば、図7の充填要素2は、そのようなヘッダ（“ヘッダ2”）を含むとともに、該ヘッダの後に、eSBRメタデータ（すなわち、ブロックのオーディオコンテンツに対してエンハンススペクトルバンド複製（eSBR）処理が実行されるべきかを指し示すものである、充填要素2内の“フラグ”）を含んでいる。オプションで、追加のeSBRメタデータも、ヘッダ2の後で、図7の充填要素2の充填データに含められる。本段落で記述している実施形態において、ヘッダ（例えば、図7のヘッダ2）は、MPEG-4 AAC規格の表4.57に規定されている従来の値のうちの1つではない識別値を持ち、代わりに、eSBR拡張ペイロードを指し示す（充填データがeSBRメタデータを含むことをヘッダの`extension_type`フィールドが指し示すようにする）。

【0080】

第1のクラスの実施形態において、発明はオーディオ処理ユニット（例えば、デコーダ）であり、当該オーディオ処理ユニットは、

符号化されたオーディオビットストリームの少なくとも1つのブロック（例えば、MPEG-4 AACビットストリームの少なくとも1つのブロック）を格納するように構成されたメモリ（例えば、図3又は図4のバッファ201）と、

メモリに結合され、ビットストリームの上記ブロックの少なくとも1つの部分を逆多重化するように構成されたビットストリームペイロードデフォーマッタ（例えば、図3の要素205、又は図4の要素215）と、

ビットストリームの上記ブロックのオーディオコンテンツの少なくとも1つの部分を復号するように結合及び構成された復号サブシステム（例えば、図3の要素202及び203、又は図4の要素202及び213）と、を有し、ブロックは、

充填要素であり、当該充填要素の始まりを示す識別子（例えば、MPEG-4 AAC規格の表4.85の値0x6を持つ“`id_syn_ele`”識別子）と、該識別子の後の充填データと、を含む充填要素と、

該ブロックのオーディオコンテンツに対してエンハンススペクトルバンド複製（eSBR）処理が実行される（例えば、該ブロックに含められたスペクトルバンド複製データ及びeSBRメタデータを使用して）べきかを特定する少なくとも1つのフラグと、を含む。

【0081】

このフラグはeSBRメタデータであり、フラグの例は`sbrPatchingMode`フラグである。フラグの他の一例は、`harmonicSBR`フラグである。これらのフラグはどちらも、ブロックのオーディオデータに対して基本形式のスペクトルバンド複製が実行されるべきか、それとも強化形式のスペクトルバンド複製が実行されるべきかを指し示す。基本形式のスペクトルバンド複製はスペクトルパッチングであり、強化形式のスペクトルバンド複製は高調波トランスポジションである。

【0082】

一部の実施形態において、充填データはまた、追加のeSBRメタデータ（すなわち、上記フラグ以外のeSBRメタデータ）を含む。

【0083】

メモリは、符号化されたオーディオビットストリームの少なくとも1つのブロックを（例えば、非一時的に）格納するバッファメモリ（例えば、図4のバッファ201の実装）とし得る。

10

20

30

40

50

【0084】

推定されることには、eSBRメタデータ（これらのeSBRツールを指し示す）を含むMPEG-4 AACビットストリームの復号中のeSBRデコーダによるeSBR処理（eSBR高調波トランスポジション及びプレフラット化を用いる）の実行の複雑さは、（指し示されるパラメータを用いた典型的な復号に関して）以下：

- ・高調波トランスポジション（16 kbps、14400/28800 Hz）

○ DFTベース：3.68 WMOPS (weighted million operations per second)

○ QMFベース：0.98 WMOPS

- ・QMFパッチング前処理（プレフラット化）：0.1 WMOPS

ようになる。知られることには、DFTベースのトランスポジションは、典型的に、過渡信号に関してQMFベースのトランスポジションよりも良好に機能する。

【0085】

本発明の一部の実施形態によれば、eSBRメタデータを含む（符号化されたオーディオビットストリームの）充填要素はまた、その値が充填要素にeSBRメタデータが含まれること及び該当ブロックのオーディオコンテンツに対してeSBR処理が実行されるべきことをシグナリングする値（例えば、bs_extension_id=3）を持つパラメータ（例えば、“bs_extension_id”パラメータ）、及び／又は、充填要素のsbr_extension()コンテナがPSデータを含むことをシグナリングする値（例えば、bs_extension_id=2）を持つパラメータ（例えば、同じ“bs_extension_id”パラメータ）を含む。例えば、下の表1に示されるように、このようなパラメータがbs_extension_id=2なる値を持つことが、充填要素のsbr_extension()コンテナがPSデータを含むことをシグナリングし得るとともに、のようなパラメータがbs_extension_id=3なる値を持つことが、充填要素のsbr_extension()コンテナがeSBRメタデータを含むことをシグナリングし得る。

【表1】

表1

bs_extension_id	意味
0	Reserved
1	Reserved
2	EXTENSION_ID_PS
3	EXTENSION_ID_ESBR

【0086】

発明の一部の実施形態によれば、eSBRメタデータ及び／又はPSデータを含む各スペクトルバンド複製拡張要素の構文は、下の表2に示す通りである（“sbr_extension()”は、スペクトルバンド複製拡張要素であるコンテナを表し、“bs_extension_id”は、上の表1に記載される通りであり、“ps_data”は、PSデータを表し、そして、“esbr_data”は、eSBRメタデータを表す）。

10

20

30

40

50

【表 2】

表 2

sbr_extension(bs_extension_id, num_bits_left)	
{	
switch (bs_extension_id) {	
case EXTENSION_ID_PS:	
num_bits_left -= ps_data();	注 1
break;	
case EXTENSION_ID_ESBR:	
num_bits_left -= esbr_data();	注 2
break;	
default:	
bs_fill_bits;	
num_bits_left = 0;	
break;	
}	
}	
注 1 : ps_data()は、読み込んだビット数を返す。	
注 2 : esbr_data()は、読み込んだビット数を返す。	

例示的な一実施形態において、上の表 2 で参照されている `esbr_data()` は、以下のメタデータパラメータの値を指し示す：

1. 1ビットメタデータパラメータ“`bs_sbr_processing`”、及び
2. 復号されるべき符号化されたビットストリームのオーディオコンテンツの各チャンネル（“`ch`”）についての、上述のパラメータ“`sbrPatchingMode[ch]`”、“`sbrOversamplingFlag[ch]`”、“`sbrPitchInBinsFlag[ch]`”、及び“`sbrPitchInBins[ch]`”の各々。

【0087】

例えば、一部の実施形態において、`esbr_data()` は、これらのメタデータパラメータを指し示すために、表 3 に示される構文を持ち得る。

【表 3 - 1】

表 3

構文	ビット数
esbr_data(id_aac, bs_coupling)	
{	
bs_sbr_preprocessing;	1
if (id_aac == ID_SCE) {	
if (sbrPatchingMode[0] == 0) {	1
sbrOversamplingFlag[0];	1
if (sbrPitchInBinsFlag[0])	1
sbrPitchInBins[0];	7
else	
sbrPitchInBins[0] = 0;	
} else {	
sbrOversamplingFlag[0] = 0;	
sbrPitchInBins[0] = 0;	
}	
} else if (id_aac == ID_CPE) {	
if (bs_coupling) {	
if (sbrPatchingMode[0,1] == 0) {	1
sbrOversamplingFlag[0,1];	1
if (sbrPitchInBinsFlag[0,1])	1
sbrPitchInBins[0,1];	7
else	
sbrPitchInBins[0,1] = 0;	
} else {	
sbrOversamplingFlag[0,1] = 0;	
sbrPitchInBins[0,1] = 0;	
}	
} else { /* bs_coupling == 0 */	
if (sbrPatchingMode[0] == 0) {	1

10

20

30

40

50

【表 3 - 2】

sbrOversamplingFlag[0];	1
if (sbrPitchInBinsFlag[0])	1
sbrPitchInBins[0];	7
else	
sbrPitchInBins[0] = 0;	
} else {	
sbrOversamplingFlag[0] = 0;	
sbrPitchInBins[0] = 0;	
}	
if (sbrPatchingMode[1] == 0) {	1
sbrOversamplingFlag[1];	1
if (sbrPitchInBinsFlag[1])	1
sbrPitchInBins[1];	7
else	
sbrPitchInBins[1] = 0;	
} else {	
sbrOversamplingFlag[1] = 0;	
sbrPitchInBins[1] = 0;	
}	
}	
}	
}	
注 : <i>bs_sbr_preprocessing</i> は、ISO/IEC 23003-3:2012 の第 6.2.12 節に記載のように定義される。 <i>sbrPatchingMode[ch]</i> 、 <i>sbrOversamplingFlag[ch]</i> 、 <i>sbrPitchInBinsFlag[ch]</i> 、及び <i>sbrPitchInBins[ch]</i> は、ISO/IEC 23003-3:2012 の第 7.5 節に記載のように定義される。	

【 0 0 8 8 】

上の構文は、レガシーデコーダへの拡張として、例えば高調波トランスポジションなどの強化形式のスペクトルバンド複製の効率的な実装を可能にする。具体的には、表 3 の e S B R データは、ビットストリームにて既にサポートされているものでもなければ、ビットストリームにて既にサポートされているパラメータから直接的に導出可能なものでもない強化形式のスペクトルバンド複製を実行するために必要なパラメータのみを含む。強化形式のスペクトルバンド複製を実行するために必要な他の全てのパラメータ及び処理データは、ビットストリーム内の既定の位置に前もって存在するパラメータから抽出される。

【 0 0 8 9 】

例えば、M P E G - 4 H E - A A C 又は H E - A A C v 2 に準拠したデコーダは、例えば高調波トランスポジションなどの強化形式のスペクトルバンド複製を含むように拡張

され得る。この強化形式のスペクトルバンド複製は、デコーダによって既にサポートされている基本形式のスペクトルバンド複製に加えてのものである。MPEG-4 HE-AAC又はHE-AAC v2に準拠したデコーダの文脈において、この基本形式のスペクトルバンド複製は、MPEG-4 AAC規格のセクション4.6.18に規定されるQMFSPECTRUMPATCHINGSBRツールである。

【0090】

強化形式のスペクトルバンド複製を実行するとき、拡張HE-AACデコーダは、ビットストリームのSBR拡張ペイロードに既に含まれているビットストリームパラメータの多くを再使用し得る。再使用され得る具体的なパラメータは、例えば、マスター周波数帯域テーブルを決定する様々なパラメータを含む。それらのパラメータは、`bs_start_freq`（マスター周波数テーブルパラメータの始まりを特定するパラメータ）、`bs_stop_freq`（マスター周波数テーブルの終わりを特定するパラメータ）、`bs_freq_scale`（オクターブ当たりの周波数帯域数を特定するパラメータ）、`bs_alter_scale`（周波数帯域のスケールを変更するパラメータ）を含む。再使用され得るパラメータはまた、ノイズ帯域テーブル（`bs_noise_bands`）及びリミッタ帯域テーブル（`bs_limiter_bands`）を決定するパラメータを含む。従って、様々な実施形態において、USAC規格で規定されるのと等価なパラメータのうちの少なくとも一部がビットストリームから省略され、それによってビットストリームにおける制御オーバーヘッドが低減される。典型的に、AAC規格で規定されるパラメータが、USAC規格で規定される等価なパラメータを持つ場合、USAC規格で規定される等価なパラメータは、AAC規格で規定されるパラメータと同じ名前、例えば、`envelope_scale_factorEOri gMapped`を持つ。しかしながら、USAC規格で規定される等価なパラメータは典型的に、AAC規格で規定されるSBR処理に対してではなく、USAC規格で規定されるエンハンストSBR処理に対して“チューン”されたものである異なる値を持つ。

【0091】

特に低ビットレートで高調波周波数構造及び強い音調特性を有するオーディオコンテンツの主観的品質を改善するために、エンハンストSBRの起動が推奨される。それらのツールを制御する対応するビットストリーム要素（すなわち、`esbr_data()`）の値は、信号依存分類メカニズムを適用することによって、エンコーダにて決定され得る。一般に、非常に低いビットレートで音楽信号を符号化するには高調波パッチング法（`sbrPatchingMode==1`）の使用が好ましく、その場合、コアコーデックは、オーディオ帯域幅において相当に制限され得る。これは、特に、これらの信号が顕著な高調波構造を含む場合に当てはまる。対照的に、音声信号及び混合信号に対しては、通常のSBRパッチング法の使用が好ましい。何故なら、それは、音声における時間的構造のいっそう良好な保存を提供するからである。

【0092】

高調波トランスポータの性能を改善するために、後続のエンベロープ調整器に入る信号のスペクトル不連続の導入を回避することを目指す前処理ステップ（`bs_sbr_preprocessing==1`）を起動することができる。このツールの動作は、高周波再構成のために低帯域信号の粗いスペクトルエンベロープを使用することが大きいレベル変動を示す信号タイプに有益である。

【0093】

高調波SBRパッチングの過渡応答を改善するために、信号適応周波数ドメインオーバーサンプリング（`sbrOversamplingFlag==1`）を適用することができる。信号適応周波数ドメインオーバーサンプリングはトランスポータの計算の複雑さを増加させるが、過渡成分を含むフレームに対してのみ利益をもたらすので、このツールの使用は、独立SBRチャンネル当たり及びフレーム当たり1回伝送されるものであるビットストリーム要素によって制御される。

【0094】

提案するエンハンスト SBR モードで動作するデコーダは、典型的に、レガシー SBR パッチングとエンハンスト SBR パッチングとの間で切り換わることができる必要がある。従って、デコーダ設定に応じて、1つのコアオーディオフレームの継続時間ほどの長さとし得る遅延が導入され得る。典型的に、この遅延は、レガシー SBR パッチング及びエンハンスト SBR パッチングの双方で同等となる。

【0095】

これら数多くのパラメータに加えて、他のデータ要素も、発明の実施形態に従って強化形式のスペクトルバンド複製を実行するときに拡張 HE-AAC デコーダによって再使用され得る。例えば、エンベロープデータ及びノイズフロアデータも、`bs__data__env` (エンベロープスケールファクタ) 及び `bs__noise__env` (ノイズフロアスケールファクタ) データから抽出されて、強化形式のスペクトルバンド複製の間に使用され得る。

【0096】

本質的に、これらの実施形態は、SBR 拡張ペイロード内のレガシー HE-AAC 又は HE-AAC v2 デコーダによって既にサポートされている構成パラメータ及びエンベロープデータを利用して、可能な限り追加の伝送データを必要しない強化形式のスペクトルバンド複製を可能にする。メタデータは、もともと、基本形式の HFR (例えば、SBR のスペクトル変換動作) に対してチューンされたものであるが、実施形態に従って、強化形式の HFR (例えば、eSBR の高調波トランスポジション) に使用される。前述したように、メタデータは概して、基本形式の HFR (例えば、線形スペクトル変換) で使用されるように意図及びチューンされた動作パラメータ (例えば、エンベロープスケールファクタ、ノイズフロアスケールファクタ、時間/周波数グリッドパラメータ、正弦波加算情報、可変クロスオーバー周波数/帯域、逆フィルタリングモード、エンベロープ解像度、平滑化モード、周波数補間モード) を表す。しかしながら、このメタデータが、強化形式の HFR (例えば、高調波トランスポジション) に特有の追加のメタデータパラメータと組み合わせられて、強化形式の HFR を使用してオーディオデータを効率的かつ効果的に処理するために使用され得る。

【0097】

従って、既に規定されているビットストリーム要素 (例えば、SBR 拡張ペイロード内のもの) を当てにするとともに、強化形式のスペクトルバンド複製をサポートするために必要なパラメータのみを追加することによって、強化形式のスペクトルバンド複製をサポートする拡張デコーダを非常に効率的に作り出し得る。新たに追加されるパラメータを例えば拡張コンテナなどの保留データフィールドに置くことと組み合わせると、このデータ削減フィーチャは、強化形式のスペクトルバンド複製をサポートしていないレガシーデコーダに対してビットストリームが後方互換であることを保証することによって、強化形式のスペクトルバンド複製をサポートするデコーダを作成することに対する障壁を実質的に低減させる。

【0098】

表 3 において、右列内の数字は、左列内の対応するパラメータのビット数を示している。

【0099】

一部の実施形態において、MPEG-4 AAC で規定される SBR オブジェクトタイプが、SBR 拡張要素 (`bs__extension__id==EXTENSION__ID__ESBR`) にてシグナリングされる SBR-Tool 及びエンハンスト SBR (eSBR) ツールの態様を含むように更新される。デコーダがこの SBR 拡張要素をサポートしていてそれを検出すると、該デコーダは、シグナリングされたエンハンスト SBR ツールの態様を使用する。このようにして更新された SBR オブジェクトタイプを、SBR エンハンスメントとして参照する。

【0100】

一部の実施形態において、発明は、オーディオデータを符号化して、符号化されたビットストリーム (例えば、MPEG-4 AAC ビットストリーム) を生成するステップを

10

20

30

40

50

含む方法であり、符号化されたビットストリームの少なくとも1つのブロックの少なくとも1つのセグメントにe S B Rメタデータを含め、且つ該ブロックの少なくとも1つの他のセグメントにオーディオデータを含めることによって、を含む。典型的な実施形態において、当該方法は、符号化されたビットストリームの各ブロックでオーディオデータをe S B Rメタデータと多重化するステップを含む。e S B Rデコーダにおける符号化されたビットストリームの典型的な復号において、デコーダは、ビットストリームからe S B Rメタデータを抽出し（e S B Rメタデータ及びオーディオデータの解析及び逆多重化することによって、を含む）、e S B Rメタデータを用いてオーディオデータを処理して、復号されたオーディオデータのストリームを生成する。

【0101】

発明の他の一態様は、e S B Rメタデータを含まない符号化されたオーディオビットストリーム（例えば、M P E G - 4 A A Cビットストリーム）の復号中にe S B R処理を実行する（例えば、高調波トランスポジション又はプレフラット化として知られるe S B Rツールのうちの少なくとも1つを使用する）ように構成されたe S B Rデコーダである。そのようなデコーダの一例を、図5を参照して説明する。

【0102】

図5のe S B Rデコーダ（400）は、デコーダ200は、図示のように接続された、バッファメモリ201（図3及び図4のメモリ201と同じである）、ビットストリームペイロードデフォーマッタ215（図4のデフォーマッタ215と同じである）、オーディオ復号サブシステム202（“コア”復号ステージ又は“コア”復号サブシステムとして参照することもあり、図3のコア復号サブシステム202と同じである）、e S B R制御データ生成サブシステム401、及びe S B R処理ステージ203（図3のステージ203と同じである）を含んでいる。典型的に、デコーダ400は、他のプロセッシング要素（図示せず）も含む。

【0103】

デコーダ400の動作において、デコーダ400によって受信された符号化されたオーディオビットストリーム（M P E G - 4 A A Cビットストリーム）の一連のブロックが、バッファ201からデフォーマッタ215にアサートされる。

【0104】

デフォーマッタ215は、ビットストリームの各ブロックを逆多重化して、それからS B Rメタデータ（量子化されたエンベロープデータを含む）を抽出するとともに典型的に他のメタデータも抽出する。デフォーマッタ215は、少なくともS B Rメタデータをe S B R処理ステージ203にアサートするように構成される。デフォーマッタ215はまた、ビットストリームの各ブロックからオーディオデータを抽出し、抽出したオーディオデータを復号サブシステム（復号ステージ）202にアサートするように結合及び構成される。

【0105】

デコーダ400のオーディオ復号サブシステム202は、デフォーマッタ215によって抽出されたオーディオデータを復号して（このような復号は“コア”復号処理として参照され得る）、復号されたオーディオデータを生成し、そして、復号されたオーディオデータをe S B R処理ステージ203にアサートするように構成される。この復号は周波数ドメインで実行される。典型的に、サブシステム202の出力が、時間ドメインの復号されたオーディオデータであるように、サブシステム202における処理の最終ステージが、復号された周波数ドメインのオーディオデータに対して、周波数ドメインー時間ドメイン変換を適用する。ステージ203は、復号されたオーディオデータに、（デフォーマッタ215によって抽出された）S B Rメタデータによって及びサブシステム401にて生成されるe S B Rメタデータによって指し示されるS B Rツール（及びe S B Rツール）を適用して（すなわち、S B R及びe S B Rメタデータを使用して、復号サブシステム202の出力に対してS B R及びe S B R処理を実行して）、デコーダ400から出力される完全に復号されたオーディオデータを生成する。典型的に、デコーダ400は、デフォー

10

20

30

40

50

マッタ 2 1 5（及びオプションでサブシステム 4 0 1 も）から出力されるデフォーマットされたオーディオデータ及びメタデータを格納するメモリ（サブシステム 2 0 2 及びステージ 2 0 3 によってアクセス可能）を含み、ステージ 2 0 3 は、S B R 及び e S B R 処理中に必要に応じてオーディオデータ及びメタデータにアクセスするように構成される。ステージ 2 0 3 における S B R 処理は、コア復号サブシステム 2 0 2 の出力に対する後処理であるとみなされ得る。オプションで、デコーダ 4 0 0 はまた、ステージ 2 0 3 の出力に対してアップミキシングを実行して、デコーダ 4 0 0 から出力される完全に復号され、アップミキシングされたオーディオを生成するように結合及び構成された最終アップミキシングサブシステム（これは、デフォーマッタ 2 1 5 によって抽出される P S メタデータを用いて、M P E G - 4 A A C 規格で規定されたパラメトリックステレオ（“ P S ”）ツールを適用し得る）を含む。

10

【 0 1 0 6 】

パラメトリックステレオは、ステレオ信号の左チャンネル及び右チャンネルの線形ダウンミキシングと、ステレオイメージを記述する空間パラメータのセットとを用いてステレオ信号を表す符号化ツールである。パラメトリックステレオは、典型的に、（１）チャンネル間の強度差を記述するチャンネル間強度差（inter-channel intensity difference s ; I I D）、（２）チャンネル間の位相差を記述するチャンネル間位相差（inter-channel phase differences ; I P D）、及び（３）チャンネル間のコヒーレンス（又は類似性）を記述するチャンネル間コヒーレンス（inter-channel coherence ; I C C）という３つのタイプの空間パラメータを使用する。コヒーレンスは、時間又は位相の関数としての相互相関の最大として測定され得る。これら３つのパラメータは概して、ステレオイメージの高品質再構成を可能にする。しかしながら、I P D パラメータは、ステレオ入力信号のチャンネル間の相対的位相差を記述するのみであり、左チャンネル及び右チャンネルにわたるこれら位相差の分布を示さない。従って、全体的な位相オフセット又は全体的な位相差を記述する第４のタイプのパラメータが、追加で使用され得る。ステレオ再構成プロセスにおいて、受信ダウンミキシング信号 $s[n]$ と受信ダウンミキシングの相関解除バージョン $d[n]$ との双方の連続したウィンドウセグメントが、空間パラメータと共に処理され、

20

$$l_k(n) = H_{11}(k, n) s_k(n) + H_{21}(k, n) d_k(n)$$

$$r_k(n) = H_{12}(k, n) s_k(n) + H_{22}(k, n) d_k(n)$$

30

に従って、左再構成信号（ $l_k(n)$ ）及び右再構成信号（ $r_k(n)$ ）が生成され、ここで、 H_{11} 、 H_{12} 、 H_{21} 及び H_{22} は、ステレオパラメータによって規定されるものである。信号 $l_k(n)$ 及び信号 $r_k(n)$ は、最終的に周波数－時間変換によって時間ドメインに変換され返す。

【 0 1 0 7 】

図 5 の制御データ生成サブシステム 4 0 1 は、復号されるべき符号化されたオーディオビットストリームの少なくとも１つの特性を検出し、検出ステップの少なくとも１つの結果に応答して e S B R 制御データ（これは、発明の他の実施形態に従って符号化されたオーディオビットストリームに含められるタイプのうちのいずれかの e S B R メタデータであり又はそれを含み得る）を生成するように結合及び構成される。e S B R 制御データはステージ 2 0 3 にアサートされ、ビットストリームの特定の特性（又は複数の特性の組み合わせ）を検出したことを受けて個々の e S B R ツール又は e S B R ツールの組み合わせの適用をトリガし、及び／又はそのような e S B R ツールの適用を制御する。例えば、高調波トランスポジションを用いた e S B R 処理の実行を制御するために、制御データ生成サブシステム 4 0 1 の一部の実施形態は、ビットストリームが音楽を示すか否かを検出することに応答して $sbrPatchingMode[ch]$ パラメータを設定する（及び、設定したパラメータをステージ 2 0 3 にアサートする）ミュージック検出器、ビットストリームによって示されるオーディオコンテンツにおける過渡成分の存在又は不存在を検出したことに応答して $sbrOversamplingFlag[ch]$ パラメータを設定する（及び、設定したパラメータをステージ 2 0 3 にアサートする）トランジェント検

40

50

出器、及び／又は、ビットストリームによって示されるオーディオコンテンツのピッチを検出したことに応答して `sbrPitchInsFlag[ch]` 及び `sbrPitchIns[ch]` パラメータを設定する（及び、設定したパラメータをステージ203にアサートする）ピッチ検出器を含み得る。発明の他の態様は、本段落及び前段落に記載した発明デコーダのいずれかの実施形態によって実行されるオーディオビットストリーム復号方法である。

【0108】

発明の態様は、発明APU、システム又は装置のいずれかの実施形態が実行するように構成される（例えば、プログラムされる）タイプの符号化又は復号方法を含む。発明の他の態様は、発明方法のいずれかの実施形態を実行するように構成される（例えば、プログラムされる）システム又は装置、並びに、発明方法のいずれかの実施形態又はそのステップを実装するためのコードを（例えば、非一時的に）格納するコンピュータ読み取り可能媒体（例えば、ディスク）を含む。例えば、発明システムは、発明方法の一実施形態又はそのステップを含め、多様な処理のうちのいずれかをデータに対して実行するようにソフトウェア又はファームウェアでプログラミングされた又はその他の方法で構成された、プログラム可能な汎用プロセッサ、デジタル信号プロセッサ、又はマイクロプロセッサであるか、それを含むかであることができる。そのような汎用プロセッサは、入力装置と、メモリと、それに対してアサートされるデータに응答して発明方法の一実施形態（又はそのステップ）を実行するようにプログラムされる（及び／又はその他の方法で構成される）プロセッシング回路と、を含むコンピュータシステムであるか、それを含むかであるとし得る。

【0109】

本発明の実施形態は、ハードウェア、ファームウェア、若しくはソフトウェア、又は双方の組み合わせ（例えば、プログラマブル論理アレイ）にて実装され得る。別段の断りがない限り、発明の一部として含まれるアルゴリズム又はプロセスは、特定のコンピュータ又は他の装置に本質的には関係付けられない。特に、ここでの教示に従って記述されたプログラムと共に種々の汎用マシンを使用することができ、あるいは、必要な方法ステップを実行するように、いっそう特殊化された装置（例えば、集積回路）を構築する方がいっそう好都合なこともある。従って、発明は、各々が、少なくとも1つのプロセッサと、少なくとも1つのデータストレージシステム（揮発性及び不揮発性のメモリ及び／又は記憶素子を含む）と、少なくとも1つの入力装置若しくはポートと、少なくとも1つの出力装置若しくはポートと、を有する1つ以上のプログラム可能なコンピュータシステム（例えば、図1の要素のうちのいずれかを実装したもの、又は図2のエンコーダ100（又はその要素）、又は図3のデコーダ200（又はその要素）、又は図4のデコーダ210（又はその要素）、又は図5のデコーダ400（又はその要素））の上で実行する1つ以上のコンピュータプログラムにて実装され得る。プログラムコードが入力データに適用されて、ここに記載された機能が実行され、出力情報が生成される。その出力情報が、知られたやり方で1つ以上の出力装置に与えられる。

【0110】

このようなプログラムは各々、コンピュータシステムと通信するために、望ましい任意のコンピュータ言語（機械語、アセンブリ言語、又はハイレベルの手続き型、論理型、又はオブジェクト指向型のプログラミング言語を含む）で実装され得る。いずれにしても、言語は、コンパイル型言語であってもよいし、インタープリタ型言語であってもよい。

【0111】

例えば、コンピュータソフトウェア命令シーケンスによって実装されるとき、発明の実施形態の様々な機能及びステップは、適切なデジタル信号処理ハードウェア上で走るマルチスレッド化ソフトウェア命令シーケンスによって実装されてもよく、その場合、実施形態の様々な装置、ステップ及び機能は、ソフトウェア命令の一部に対応し得る。

【0112】

そのようなコンピュータプログラムは各々、好ましくは、汎用又は専用のプログラマブ

ルコンピュータによって読み取り可能な記憶媒体又は記憶装置（例えば、ソリッドステートメモリ若しくは媒体、又は磁気媒体若しくは光学媒体）に格納又はダウンロードされ、該記憶媒体又は記憶装置がコンピュータシステムによって読み取られるときに、ここに記載された手順を実行するようにコンピュータを構成して動作させる。発明システムはまた、コンピュータプログラムを備えて（すなわち、格納して）構成された、コンピュータ読み取り可能記憶媒体として実装されてもよく、そのように構成された記憶媒体は、コンピュータシステムに、ここに記載された機能を実行するよう、特定の予め定められたように動作させる。

【0113】

発明の数多くの実施形態を説明してきた。とはいえ、理解されることには、発明の精神及び範囲から逸脱することなく様々な変更が為され得る。上での教示に照らして、本発明の数多くの変更及び変形が可能である。例えば、効率的な実装を支援するために、複素QMF分析及び合成フィルタバンクと組み合わせて位相シフトを使用してもよい。分析フィルタバンクは、コアダコーダによって生成された時間ドメイン低帯域信号を複数のサブバンド（例えば、QMFサブバンド）へとフィルタリングすることを担う。合成フィルタバンクは、選択されたHFR技術（受信されるsbr Patching Modeパラメータによって指し示される）によって生成された再生成高帯域を、復号された低帯域と組み合わせて、広帯域出力オーディオ信号を生成することを担う。例えば通常のデュアルレート動作又はダウンサンプリングSBRモードといった特定のサンプルレートモードで動作する所与のフィルタバンク実装は、しかしながら、ビットストリームに依存する位相シフトを持つべきでない。SBRで使用されるQMFバンクは、余弦変調フィルタバンクの理論の複素指数関数拡張である。示され得ることには、複素指数関数変調を用いて余弦変調フィルタバンクを拡張するとき、エイリアス相殺制約が使われないものとなる。従って、SBR QMFバンクでは、分析フィルタ $h_k(n)$ 及び合成フィルタ $f_k(n)$ の双方を、

【数5】

$$h_k(n) = f_k(n) = p_0(n) \exp \left\{ i \frac{\pi}{M} \left(k + \frac{1}{2} \right) \left(n - \frac{N}{2} \right) \right\}, 0 \leq n \leq N; 0 \leq k < M \quad (1)$$

によって規定することができ、ここで、 $p_0(n)$ は実数値の対称又は非対称プロトタイプフィルタ（典型的に、低域通過プロトタイプフィルタ）であり、 M はチャンネル数を表し、 N はプロトタイプフィルタ次数である。分析フィルタバンクで使用されるチャンネルの数は、合成フィルタバンクで使用されるチャンネルの数と異なり得る。例えば、分析フィルタバンクは32チャンネルを有し、合成フィルタバンクは64チャンネルを有し得る。ダウンサンプリングモードで合成フィルタバンクを動作させるとき、合成フィルタバンクは32チャンネルのみを有し得る。フィルタバンクからのサブバンドサンプルは複素数の値であるので、追加のチャンネル依存であり得る位相シフトステップが、分析フィルタバンクに付加され得る。これらの追加の位相シフトは、合成フィルタバンクの前に補償される必要がある。原理的に位相シフト項はQMF分析／合成チェーンの動作を破壊することなく任意の値とすることができるが、それらはまた、適合性検証のために特定の値に制約されてもよい。SBR信号は位相ファクタの選択によって影響されることになるが、コアダコーダから来る低域通過信号は影響されない。出力信号の音質は影響を受けない。

【0114】

プロトタイプフィルタの係数 $p_0(n)$ の係数は、下の表4に示すように、640の長さLで規定され得る。

10

20

30

40

50

【表 4 - 1】

表 4

n	p ₀ (n)	n	p ₀ (n)	n	p ₀ (n)
0	0.0000000000	214	0.0019765601	428	0.0117623832
1	-0.0005525286	215	-0.0032086896	429	0.0163701258
2	-0.0005617692	216	-0.0085711749	430	0.0207997072
3	-0.0004947518	217	-0.0141288827	431	0.0250307561
4	-0.0004875227	218	-0.0198834129	432	0.0290824006
5	-0.0004893791	219	-0.0258227288	433	0.0329583930
6	-0.0005040714	220	-0.0319531274	434	0.0366418116
7	-0.0005226564	221	-0.0382776572	435	0.0401458278
8	-0.0005466565	222	-0.0447806821	436	0.0434768782
9	-0.0005677802	223	-0.0514804176	437	0.0466303305
10	-0.0005870930	224	-0.0583705326	438	0.0495978676
11	-0.0006132747	225	-0.0654409853	439	0.0524093821
12	-0.0006312493	226	-0.0726943300	440	0.0550460034
13	-0.0006540333	227	-0.0801372934	441	0.0575152691
14	-0.0006777690	228	-0.0877547536	442	0.0598166570
15	-0.0006941614	229	-0.0955533352	443	0.0619602779
16	-0.0007157736	230	-0.1035329531	444	0.0639444805
17	-0.0007255043	231	-0.1116826931	445	0.0657690668
18	-0.0007440941	232	-0.1200077984	446	0.0674525021
19	-0.0007490598	233	-0.1285002850	447	0.0689664013
20	-0.0007681371	234	-0.1371551761	448	0.0703533073
21	-0.0007724848	235	-0.1459766491	449	0.0715826364
22	-0.0007834332	236	-0.1549607071	450	0.0726774642
23	-0.0007779869	237	-0.1640958855	451	0.0736406005
24	-0.0007803664	238	-0.1733808172	452	0.0744664394
25	-0.0007801449	239	-0.1828172548	453	0.0751576255
26	-0.0007757977	240	-0.1923966745	454	0.0757305756
27	-0.0007630793	241	-0.2021250176	455	0.0761748321
28	-0.0007530001	242	-0.2119735853	456	0.0765050718
29	-0.0007319357	243	-0.2219652696	457	0.0767204924
30	-0.0007215391	244	-0.2320690870	458	0.0768230011
31	-0.0006917937	245	-0.2423016884	459	0.0768173975
32	-0.0006650415	246	-0.2526480309	460	0.0767093490
33	-0.0006341594	247	-0.2631053299	461	0.0764992170
34	-0.0005946118	248	-0.2736634040	462	0.0761992479
35	-0.0005564576	249	-0.2843214189	463	0.0758008358
36	-0.0005145572	250	-0.2950716717	464	0.0753137336
37	-0.0004606325	251	-0.3059098575	465	0.0747452558
38	-0.0004095121	252	-0.3168278913	466	0.0741003642
39	-0.0003501175	253	-0.3278113727	467	0.0733620255
40	-0.0002896981	254	-0.3388722693	468	0.0725682583
41	-0.0002098337	255	-0.3499914122	469	0.0717002673
42	-0.0001446380	256	0.3611589903	470	0.0707628710
43	-0.0000617334	257	0.3723795546	471	0.0697630244
44	0.0000134949	258	0.3836350013	472	0.0687043828
45	0.0001094383	259	0.3949211761	473	0.0676075985
46	0.0002043017	260	0.4062317676	474	0.0664367512
47	0.0002949531	261	0.4175696896	475	0.0652247106
48	0.0004026540	262	0.4289119920	476	0.0639715898

10

20

30

40

50

【表 4 - 2】

49	0.0005107388	263	0.4402553754	477	0.0626857808
50	0.0006239376	264	0.4515996535	478	0.0613455171
51	0.0007458025	265	0.4629308085	479	0.0599837480
52	0.0008608443	266	0.4742453214	480	0.0585915683
53	0.0009885988	267	0.4855253091	481	0.0571616450
54	0.0011250155	268	0.4967708254	482	0.0557173648
55	0.0012577884	269	0.5079817500	483	0.0542452768
56	0.0013902494	270	0.5191234970	484	0.0527630746
57	0.0015443219	271	0.5302240895	485	0.0512556155
58	0.0016868083	272	0.5412553448	486	0.0497385755
59	0.0018348265	273	0.5522051258	487	0.0482165720
60	0.0019841140	274	0.5630789140	488	0.0466843027
61	0.0021461583	275	0.5738524131	489	0.0451488405
62	0.0023017254	276	0.5845403235	490	0.0436097542
63	0.0024625616	277	0.5951123086	491	0.0420649094
64	0.0026201758	278	0.6055783538	492	0.0405349170
65	0.0027870464	279	0.6159109932	493	0.0390053679
66	0.0029469447	280	0.6261242695	494	0.0374812850
67	0.0031125420	281	0.6361980107	495	0.0359697560
68	0.0032739613	282	0.6461269695	496	0.0344620948
69	0.0034418874	283	0.6559016302	497	0.0329754081
70	0.0036008268	284	0.6655139880	498	0.0315017608
71	0.0037603922	285	0.6749663190	499	0.0300502657
72	0.0039207432	286	0.6842353293	500	0.0286072173
73	0.0040819753	287	0.6933282376	501	0.0271859429
74	0.0042264269	288	0.7022388719	502	0.0257875847
75	0.0043730719	289	0.7109410426	503	0.0244160992
76	0.0045209852	290	0.7194462634	504	0.0230680169
77	0.0046606460	291	0.7277448900	505	0.0217467550
78	0.0047932560	292	0.7358211758	506	0.0204531793
79	0.0049137603	293	0.7436827863	507	0.0191872431
80	0.0050393022	294	0.7513137456	508	0.0179433381
81	0.0051407353	295	0.7587080760	509	0.0167324712
82	0.0052461166	296	0.7658674865	510	0.0155405553
83	0.0053471681	297	0.7727780881	511	0.0143904666
84	0.0054196775	298	0.7794287519	512	-0.0132718220
85	0.0054876040	299	0.7858353120	513	-0.0121849995
86	0.0055475714	300	0.7919735841	514	-0.0111315548
87	0.0055938023	301	0.7978466413	515	-0.0101150215
88	0.0056220643	302	0.8034485751	516	-0.0091325329
89	0.0056455196	303	0.8087695004	517	-0.0081798233
90	0.0056389199	304	0.8138191270	518	-0.0072615816
91	0.0056266114	305	0.8185776004	519	-0.0063792293
92	0.0055917128	306	0.8230419890	520	-0.0055337211
93	0.0055404363	307	0.8272275347	521	-0.0047222596
94	0.0054753783	308	0.8311038457	522	-0.0039401124
95	0.0053838975	309	0.8346937361	523	-0.0031933778
96	0.0052715758	310	0.8379717337	524	-0.0024826723
97	0.0051382275	311	0.8409541392	525	-0.0018039472
98	0.0049839687	312	0.8436238281	526	-0.0011568135
99	0.0048109469	313	0.8459818469	527	-0.0005464280
100	0.0046039530	314	0.8480315777	528	0.0000276045
101	0.0043801861	315	0.8497805198	529	0.0005832264
102	0.0041251642	316	0.8511971524	530	0.0010902329

10

20

30

40

50

【表 4 - 3】

103	0.0038456408	317	0.8523047035	531	0.0015784682
104	0.0035401246	318	0.8531020949	532	0.0020274176
105	0.0032091885	319	0.8535720573	533	0.0024508540
106	0.0028446757	320	0.8537385600	534	0.0028446757
107	0.0024508540	321	0.8535720573	535	0.0032091885
108	0.0020274176	322	0.8531020949	536	0.0035401246
109	0.0015784682	323	0.8523047035	537	0.0038456408
110	0.0010902329	324	0.8511971524	538	0.0041251642
111	0.0005832264	325	0.8497805198	539	0.0043801861
112	0.0000276045	326	0.8480315777	540	0.0046039530
113	-0.0005464280	327	0.8459818469	541	0.0048109469
114	-0.0011568135	328	0.8436238281	542	0.0049839687
115	-0.0018039472	329	0.8409541392	543	0.0051382275
116	-0.0024826723	330	0.8379717337	544	0.0052715758
117	-0.0031933778	331	0.8346937361	545	0.0053838975
118	-0.0039401124	332	0.8311038457	546	0.0054753783
119	-0.0047222596	333	0.8272275347	547	0.0055404363
120	-0.0055337211	334	0.8230419890	548	0.0055917128
121	-0.0063792293	335	0.8185776004	549	0.0056266114
122	-0.0072615816	336	0.8138191270	550	0.0056389199
123	-0.0081798233	337	0.8087695004	551	0.0056455196
124	-0.0091325329	338	0.8034485751	552	0.0056220643
125	-0.0101150215	339	0.7978466413	553	0.0055938023
126	-0.0111315548	340	0.7919735841	554	0.0055475714
127	-0.0121849995	341	0.7858353120	555	0.0054876040
128	0.0132718220	342	0.7794287519	556	0.0054196775
129	0.0143904666	343	0.7727780881	557	0.0053471681
130	0.0155405553	344	0.7658674865	558	0.0052461166
131	0.0167324712	345	0.7587080760	559	0.0051407353
132	0.0179433381	346	0.7513137456	560	0.0050393022
133	0.0191872431	347	0.7436827863	561	0.0049137603
134	0.0204531793	348	0.7358211758	562	0.0047932560
135	0.0217467550	349	0.7277448900	563	0.0046606460
136	0.0230680169	350	0.7194462634	564	0.0045209852
137	0.0244160992	351	0.7109410426	565	0.0043730719
138	0.0257875847	352	0.7022388719	566	0.0042264269
139	0.0271859429	353	0.6933282376	567	0.0040819753
140	0.0286072173	354	0.6842353293	568	0.0039207432
141	0.0300502657	355	0.6749663190	569	0.0037603922
142	0.0315017608	356	0.6655139880	570	0.0036008268
143	0.0329754081	357	0.6559016302	571	0.0034418874
144	0.0344620948	358	0.6461269695	572	0.0032739613
145	0.0359697560	359	0.6361980107	573	0.0031125420
146	0.0374812850	360	0.6261242695	574	0.0029469447
147	0.0390053679	361	0.6159109932	575	0.0027870464
148	0.0405349170	362	0.6055783538	576	0.0026201758
149	0.0420649094	363	0.5951123086	577	0.0024625616
150	0.0436097542	364	0.5845403235	578	0.0023017254
151	0.0451488405	365	0.5738524131	579	0.0021461583
152	0.0466843027	366	0.5630789140	580	0.0019841140
153	0.0482165720	367	0.5522051258	581	0.0018348265
154	0.0497385755	368	0.5412553448	582	0.0016868083
155	0.0512556155	369	0.5302240895	583	0.0015443219
156	0.0527630746	370	0.5191234970	584	0.0013902494

10

20

30

40

50

【表 4 - 4】

157	0.0542452768	371	0.5079817500	585	0.0012577884
158	0.0557173648	372	0.4967708254	586	0.0011250155
159	0.0571616450	373	0.4855253091	587	0.0009885988
160	0.0585915683	374	0.4742453214	588	0.0008608443
161	0.0599837480	375	0.4629308085	589	0.0007458025
162	0.0613455171	376	0.4515996535	590	0.0006239376
163	0.0626857808	377	0.4402553754	591	0.0005107388
164	0.0639715898	378	0.4289119920	592	0.0004026540
165	0.0652247106	379	0.4175696896	593	0.0002949531
166	0.0664367512	380	0.4062317676	594	0.0002043017
167	0.0676075985	381	0.3949211761	595	0.0001094383
168	0.0687043828	382	0.3836350013	596	0.0000134949
169	0.0697630244	383	0.3723795546	597	-0.0000617334
170	0.0707628710	384	-0.3611589903	598	-0.0001446380
171	0.0717002673	385	-0.3499914122	599	-0.0002098337
172	0.0725682583	386	-0.3388722693	600	-0.0002896981
173	0.0733620255	387	-0.3278113727	601	-0.0003501175
174	0.0741003642	388	-0.3168278913	602	-0.0004095121
175	0.0747452558	389	-0.3059098575	603	-0.0004606325
176	0.0753137336	390	-0.2950716717	604	-0.0005145572
177	0.0758008358	391	-0.2843214189	605	-0.0005564576
178	0.0761992479	392	-0.2736634040	606	-0.0005946118
179	0.0764992170	393	-0.2631053299	607	-0.0006341594
180	0.0767093490	394	-0.2526480309	608	-0.0006650415
181	0.0768173975	395	-0.2423016884	609	-0.0006917937
182	0.0768230011	396	-0.2320690870	610	-0.0007215391
183	0.0767204924	397	-0.2219652696	611	-0.0007319357
184	0.0765050718	398	-0.2119735853	612	-0.0007530001
185	0.0761748321	399	-0.2021250176	613	-0.0007630793
186	0.0757305756	400	-0.1923966745	614	-0.0007757977
187	0.0751576255	401	-0.1828172548	615	-0.0007801449
188	0.0744664394	402	-0.1733808172	616	-0.0007803664
189	0.0736406005	403	-0.1640958855	617	-0.0007779869
190	0.0726774642	404	-0.1549607071	618	-0.0007834332
191	0.0715826364	405	-0.1459766491	619	-0.0007724848
192	0.0703533073	406	-0.1371551761	620	-0.0007681371
193	0.0689664013	407	-0.1285002850	621	-0.0007490598
194	0.0674525021	408	-0.1200077984	622	-0.0007440941
195	0.0657690668	409	-0.1116826931	623	-0.0007255043
196	0.0639444805	410	-0.1035329531	624	-0.0007157736
197	0.0619602779	411	-0.0955533352	625	-0.0006941614
198	0.0598166570	412	-0.0877547536	626	-0.0006777690
199	0.0575152691	413	-0.0801372934	627	-0.0006540333
200	0.0550460034	414	-0.0726943300	628	-0.0006312493
201	0.0524093821	415	-0.0654409853	629	-0.0006132747
202	0.0495978676	416	-0.0583705326	630	-0.0005870930
203	0.0466303305	417	-0.0514804176	631	-0.0005677802
204	0.0434768782	418	-0.0447806821	632	-0.0005466565
205	0.0401458278	419	-0.0382776572	633	-0.0005226564
206	0.0366418116	420	-0.0319531274	634	-0.0005040714
207	0.0329583930	421	-0.0258227288	635	-0.0004893791
208	0.0290824006	422	-0.0198834129	636	-0.0004875227
209	0.0250307561	423	-0.0141288827	637	-0.0004947518
210	0.0207997072	424	-0.0085711749	638	-0.0005617692

【表 4 - 5】

211	0.0163701258	425	-0.0032086896	639	-0.0005525280
212	0.0117623832	426	0.0019765601		
213	0.0069636862	427	0.0069636862		

プロトタイプフィルタ $p_0(n)$ はまた、例えば丸め、サブサンプリング、補間、及び間引きなどの1つ以上の数学演算によって、表 4 から導出されてもよい。

【0115】

10

20

30

40

50

SBR関係の制御情報のチューニングは、典型的には（先述のように）トランスポジションの詳細に依存しないが、一部の実施形態では、再生成される信号の品質を改善するために、制御データのうちの特定の要素が、eSBR拡張コンテナ（`bs__extension__id==EXTENSION_ID_ESBR`）内で同時伝送されてもよい。同時伝送される要素の一部は、ノイズフロアデータ（例えば、ノイズフロアスケールファクタ、及び各ノイズフロアに対するデルタコーディングの周波数方向又は時間方向のいずれかでの方向を指し示すパラメータ）、逆フィルタリングデータ（例えば、逆フィルタリングなし、低いレベルの逆フィルタリング、中間レベルの逆フィルタリング、及び強いレベルの逆フィルタリングから選択される逆フィルタリングモードを指し示すパラメータ）、及び欠落高調波データ（例えば、再生成される高帯域の特定の周波数帯域に正弦波を加えるべきかを指し示すパラメータ）を含み得る。これらの要素は全て、エンコーダで実行されるデコーダのトランスポーザの合成エミュレーションを当てにしており、従って、選択されたトランスポーザに対して適切に調整される場合に再生成信号の品質を高め得る。

【0116】

具体的には、一部の実施形態において、欠落高調波及び逆フィルタリング制御データが、eSBR拡張コンテナ内で（表3の他のビットストリームパラメータとともに）伝送され、eSBRの高調波トランスポーザに対して調整される。eSBRの高調波トランスポーザのためにこれらの2つのクラスのメタデータを伝送するのに必要とされる追加のビットレートは比較的低い。従って、調整された欠落高調波及び／又は逆フィルタリング制御データをeSBR拡張コンテナで送ることは、ビットレートに最小限の影響しか与えずに、トランスポーザによって生成されるオーディオの品質を高めることになる。レガシーデコーダとの後方互換性を確保するために、SBRのスペクトル変換処理に対して調整されたパラメータも、暗黙的又は明示的のいずれかのシグナリングを用いてSBR制御データの一部としてビットストリームで送られ得る。

【0117】

この出願に記載されるSBRエンハンスメントを有するデコーダの複雑さは、実装したものの全体的な計算の複雑さを著しく増加させないように制限されなければならない。好ましくは、eSBRツールを使用するとき、SBRオブジェクトタイプのPCU（MOP）は4.5以下であり、eSBRツールを使用するとき、SBRオブジェクトタイプのRCUは3以下である。近似による処理能力は、整数のMOPS数で規定されるプロセッサ複雑度単位（Processor Complexity Units；PCU）で与えられる。近似によるRAM使用量は、整数のkWords（1000ワード）数で規定されるRAM複雑度単位（RAM Complexity Units；RCU）で与えられる。RCU数は、異なるオブジェクト及び／又はチャンネルの間で共されることが可能な作業バッファを含まない。また、PCUはサンプリング周波数に比例する。PCU値は、チャンネル当たりのMOPS（Million Operations per Second）で与えられ、RCU値はチャンネル当たりのkWordsで与えられる。

【0118】

異なるデコーダ構成によって復号されることが出来るものである、HE-AAC符号化オーディオのような、圧縮されたデータでは、特別な注意が必要である。この場合、復号は、後方互換的（AACのみ）及び強化的（AAC+SBR）に行われることができる。圧縮されたデータが、後方互換性のある復号及び強化された復号の双方を許す場合であって、且つデコーダが、幾分の追加遅延を挿入するポストプロセッサ（例えば、HE-AACにおけるSBRポストプロセッサ）を使用しているように、強化的に動作している場合、対応するnの値によって記述される、後方互換モードに対して生じるこの追加の時間遅延が、合成ユニットを提示するときに考慮に入れられることを保証しなければならない。

（オーディオが他のメディアと同期したままであるように）合成タイムスタンプが正しく扱われることを確保するために、出力サンプルレートでの（オーディオチャンネル当たりの）サンプル数で与えられる後処理によって導入される追加遅延は、デコーダ動作モードがこの出願に記載されるSBRエンハンスメント（eSBRを含む）を含むときに、30

10

20

30

40

50

10である。従って、オーディオ合成ユニットにおいて、デコーダ動作モードがこの出願に記載されるSBRエンハンスメントを含むとき、その合成時間が合成ユニット内の3011番目のオーディオサンプルに適用される。

【0119】

特に低ビットレートで高調波周波数構造及び強い音調特性を有するオーディオコンテンツの主観的品質を改善するには、エンハンストSBRがアクティブにされるべきである。それらのツールを制御する対応するビットストリーム要素（すなわち、`esbr_data()`）の値は、信号依存分類メカニズムを適用することによって、エンコーダにて決定され得る。

【0120】

一般に、非常に低いビットレートで音楽信号を符号化するには高調波パッチング法（`sbrPatchingMode==0`）の使用が好ましく、その場合、コアコーデックは、オーディオ帯域幅において相当に制限され得る。これは、特に、これらの信号が顕著な高調波構造を含む場合に当てはまる。対照的に、音声信号及び混合信号に対しては、通常のSBRパッチング法の使用が好ましい。何故なら、それは、音声における時間的構造のいっそう良好な保存を提供するからである。

【0121】

高調波トランスポートの性能を改善するために、後続のエンベロープ調整器に入る信号のスペクトル不連続の導入を回避する前処理ステップ（`bs_sbr_preprocess==1`）をアクティブにすることができる。このツールの動作は、高周波再構成のために低帯域信号の粗いスペクトルエンベロープを使用することが大きいレベル変動を示す信号タイプに有益である。

【0122】

高調波SBRパッチング（`sbrPatchingMode==0`）の過渡応答を改善するために、信号適応周波数ドメインオーバーサンプリング（`sbrOversamplingFlag==1`）を適用することができる。信号適応周波数ドメインオーバーサンプリングはトランスポートの計算の複雑さを増加させるが、過渡成分を含むフレームに対してのみ利益をもたらすので、このツールの使用は、独立SBRチャンネル当たり及びフレーム当たり1回伝送されるものであるビットストリーム要素によって制御される。

【0123】

SBRエンハンスメント（すなわち、eSBRツールの高調波トランスポートをイネーブルすること）を備えたHE-AAC v2の典型的なビットレート設定推奨は、44.1 kHz又は48 kHzのいずれかのサンプリングレートのステレオオーディオコンテンツに対して20-32 kbpsに相当する。SBRエンハンスメントの相対的な主観的品質利得は、低い側のビットレート境界に向かって増加し、適切に構成されたエンコーダは、この範囲をいっそう低いビットレートまで拡張することを可能にする。上で提示したビットレートは推奨に過ぎず、特定のサービス要求に合わせて適応され得る。

【0124】

提案するエンハンストSBRモードで動作するデコーダは、典型的に、レガシーSBRパッチングとエンハンストSBRパッチングとの間で切り換えることができる必要がある。従って、デコーダ設定に応じて、1つのコアオーディオフレームの継続時間ほどの長さとし得る遅延が導入され得る。典型的に、この遅延は、レガシーSBRパッチング及びエンハンストSBRパッチングの双方で同等となる。

【0125】

理解されるべきことには、添付の請求項の範囲内で、ここに具体的に記載されたのとは異なるように発明が実施され得る。以下の請求項に含まれる如何なる参照符号も、単に例示目的でのものであり、いかようにも請求項を解釈又は限定するために使用されるべきではない。

【0126】

本発明の様々な態様が、以下の列挙実施形態例（enumerated example embodiments）

10

20

30

40

50

t; E E E) から理解され得る。

【0127】

E E E 1. オーディオ信号の高周波再構成を実行する方法であって、当該方法は、符号化されたオーディオビットストリームを受信し、該符号化されたオーディオビットストリームは、前記オーディオ信号の低帯域部分を表すオーディオデータと、高周波再構成メタデータとを含み、

前記オーディオデータを復号して、復号された低帯域オーディオ信号を生成し、

前記符号化されたオーディオビットストリームから前記高周波再構成メタデータを抽出し、前記高周波再構成メタデータは、高周波再構成プロセスのための動作パラメータを含み、該動作パラメータは、前記符号化されたオーディオビットストリームの後方互換拡張コンテナ内に置かれたパッチングモードパラメータを含み、該パッチングモードパラメータの第1の値は、スペクトル変換を指し示し、該パッチングモードパラメータの第2の値は、位相ボコーダ周波数拡散による高調波トランスポジションを指し示し、

前記復号された低帯域オーディオ信号をフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成し、

前記フィルタリングされた低帯域オーディオ信号及び前記高周波再構成メタデータを用いて、前記オーディオ信号の高帯域部分を再生成し、当該再生成することは、前記パッチングモードパラメータが前記第1の値である場合にスペクトル変換を含み、当該再生成することは、前記パッチングモードパラメータが前記第2の値である場合に位相ボコーダ周波数拡散による高調波トランスポジションを含み、

前記フィルタリングされた低帯域オーディオ信号を前記再生成された高帯域部分と組み合わせて、広帯域オーディオ信号を形成する、

ことを有し、

前記フィルタリングすること、前記再生成すること、及び前記組み合わせることは、オーディオチャンネル当たり3010サンプル以下の遅延で後処理動作として実行され、前記スペクトル変換は、適応逆フィルタリングによって、音調成分と雑音ライク成分との間の比を維持することを有する、

方法。

【0128】

E E E 2. 前記符号化されたオーディオビットストリームは更に充填要素を含み、該充填要素は、該充填要素の始まりを指し示す識別子と、該識別子の後の充填データとを有し、該充填データが前記後方互換拡張コンテナを含む、E E E 1の方法。

【0129】

E E E 3. 前記識別子は、最上位ビットが先に伝送され且つ0x6の値を持つ3ビット符号なし整数である、E E E 2の方法。

【0130】

E E E 4. 前記充填データは拡張ペイロードを含み、該拡張ペイロードはスペクトルバンド複製拡張データを含み、前記拡張ペイロードは、最上位ビットが先頭に送信され且つ‘1101’又は‘1110’の値を持つ4ビット符号なし整数で識別され、

オプションで、前記スペクトルバンド複製拡張データは、

オプションのスペクトルバンド複製ヘッダと、

前記ヘッダの後のスペクトルバンド複製データと、

前記スペクトルバンド複製データの後のスペクトルバンド複製拡張要素であり、フラグが含まれているスペクトルバンド複製拡張要素と、

を含む、

E E E 2又は3の方法。

【0131】

E E E 5. 前記高周波再構成メタデータは、エンベロープスケールファクタ、ノイズフロアスケールファクタ、時間／周波数グリッド情報、又はクロスオーバー周波数を指し示すパラメータを含む、E E E 1乃至4のいずれか一の方法。

【0132】

EEE 6. 前記後方互換拡張コンテナは更に、前記パッチングモードパラメータが前記第1の値に等しいときに、前記高帯域部分のスペクトルエンベロープの形状における不連続を回避するために追加の前処理が使用されるかを指し示すフラグを含み、該フラグの第1の値は、前記追加の前処理をイネーブルし、該フラグの第2の値は、前記追加の前処理をディセーブルする、EEE 1乃至5のいずれか一の方法。

【0133】

EEE 7. 前記追加の前処理は、線形予測フィルタ係数を用いてプリゲイン曲線を計算することを含む、EEE 6の方法。

【0134】

EEE 8. 前記後方互換拡張コンテナは更に、前記パッチングモードパラメータが前記第2の値に等しいときに、信号適応周波数ドメインオーバーサンプリングが適用されるべきかを指し示すフラグを含み、該フラグの第1の値は、前記信号適応周波数ドメインオーバーサンプリングをイネーブルし、該フラグの第2の値は、前記信号適応周波数ドメインオーバーサンプリングをディセーブルする、EEE 1乃至5のいずれか一の方法。

【0135】

EEE 9. 前記信号適応周波数ドメインオーバーサンプリングは、過渡信号を含むフレームに対してのみ適用される、EEE 8の方法。

【0136】

EEE 10. 位相ボコーダ周波数拡散による前記高調波トランスポジションは、毎秒450万演算及び3kワードのメモリの又はそれよりも低い推定複雑度で実行される、EEE 1乃至9のいずれか一の方法。

【0137】

EEE 11. プロセッサによって実行されるときにEEE 1乃至10のいずれか一の方法を実行する命令を含んだ非一時的なコンピュータ読み取り可能媒体。

【0138】

EEE 12. 命令を有するコンピュータプログラムプロダクトであって、前記命令は、コンピューティング装置又はシステムによって実行されるときに、該コンピューティング装置又はシステムに、EEE 1乃至10のいずれか一の方法を実行させる、コンピュータプログラムプロダクト。

【0139】

EEE 13. オーディオ信号の高周波再構成を実行するオーディオ処理ユニットであって、当該オーディオ処理ユニットは、

符号化されたオーディオビットストリームを受信する入力インタフェースであり、前記符号化されたオーディオビットストリームは、前記オーディオ信号の低帯域部分を表すオーディオデータと、高周波再構成メタデータとを含む、入力インタフェースと、

前記オーディオデータを復号して、復号された低帯域オーディオ信号を生成するコアオーディオデコーダと、

前記符号化されたオーディオビットストリームから前記高周波再構成メタデータを抽出するデフォーマッタであり、前記高周波再構成メタデータは、高周波再構成プロセスのための動作パラメータを含み、該動作パラメータは、前記符号化されたオーディオビットストリームの後方互換拡張コンテナ内に置かれたパッチングモードパラメータを含み、該パッチングモードパラメータの第1の値は、スペクトル変換を指し示し、該パッチングモードパラメータの第2の値は、位相ボコーダ周波数拡散による高調波トランスポジションを指し示す、デフォーマッタと、

前記復号された低帯域オーディオ信号をフィルタリングして、フィルタリングされた低帯域オーディオ信号を生成する分析フィルタバンクと、

前記フィルタリングされた低帯域オーディオ信号及び前記高周波再構成メタデータを用いて、前記オーディオ信号の高帯域部分を再構成する高周波リジェネレータであり、前記再構成することは、前記パッチングモードパラメータが前記第1の値である場合にスペク

10

20

30

40

50

トル変換を含み、前記再構成することは、前記パッチングモードパラメータが前記第 2 の値である場合に位相ボコーダ周波数拡散による高調波トランスポジションを含む、高周波リジェネレータと、

前記フィルタリングされた低帯域オーディオ信号を前記再生成された高帯域部分と組み合わせて、広帯域オーディオ信号を形成する合成フィルタバンクと、

を有し、

前記分析フィルタバンク、前記高周波リジェネレータ、及び前記合成フィルタバンクは、オーディオチャンネル当たり 3010 サンプル以下の遅延でポストプロセッサにて実行され、前記スペクトル変換は、適応逆フィルタリングによって、音調成分と雑音ライク成分との間の比を維持することを有する、

オーディオ処理ユニット。

【0140】

EEE14. 位相ボコーダ周波数拡散による前記高調波トランスポジションは、毎秒 450 万演算及び 3 k ワードのメモリの又はそれよりも低い推定複雑度で実行される、EEE13 のオーディオ処理ユニット。

10

20

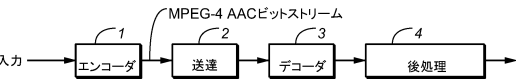
30

40

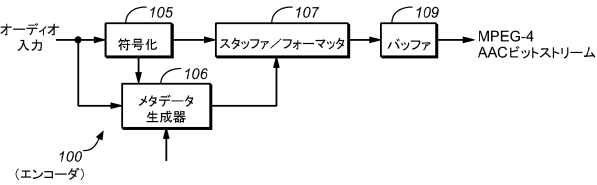
50

【図面】

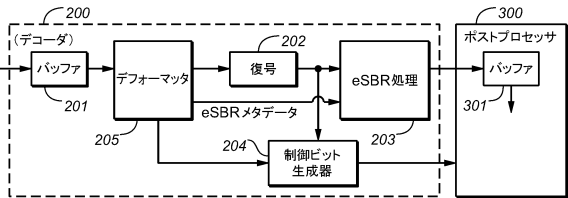
【図 1】



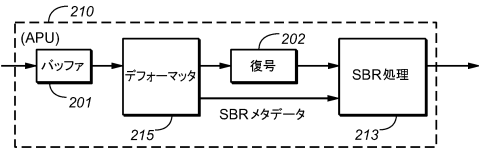
【図 2】



【図 3】

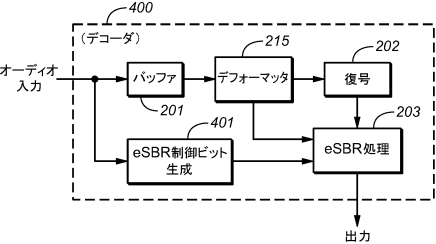


【図 4】

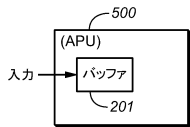


10

【図 5】



【図 6】

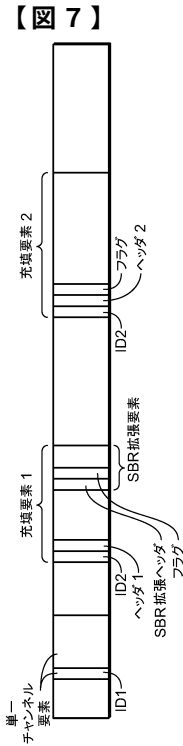


20

30

40

50



10

20

30

40

50

フロントページの続き

- アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内
- (72)発明者 ブルンハーゲン, ヘイコ
アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内
- (72)発明者 エクストランド, ペール
アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内
- 審査官 山下 剛史
- (56)参考文献 特開 2 0 1 0 - 2 0 2 5 1 (J P , A)
特表 2 0 1 1 - 5 2 7 4 4 7 (J P , A)
特表 2 0 1 2 - 5 3 1 6 3 2 (J P , A)
特表 2 0 1 3 - 5 1 6 6 5 2 (J P , A)
特表 2 0 1 3 - 5 2 1 5 3 6 (J P , A)
特表 2 0 1 3 - 5 2 1 5 3 8 (J P , A)
則松武志他, 音声と楽音を統合した音響信号符号化, 日本音響学会誌, 2012年03月, 第68
巻, 第3号, pp.123-128
Frederik NAGEL, et al., A HARMONIC BANDWIDTH EXTENSION METHOD FOR AUDIO CO
DECS, ICASSP 2009, IEEE, 2009年01月, pp.145-148
- (58)調査した分野 (Int.Cl., D B 名)
G 1 0 L 1 9 / 0 0 - 2 5 / 9 3
I E E E X p l o r e