

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4950210号
(P4950210)

(45) 発行日 平成24年6月13日(2012.6.13)

(24) 登録日 平成24年3月16日(2012.3.16)

(51) Int.Cl. F I
G 1 0 L 19/02 (2006.01) G 1 0 L 19/02 1 5 0

請求項の数 24 (全 27 頁)

(21) 出願番号	特願2008-538430 (P2008-538430)	(73) 特許権者	398012616
(86) (22) 出願日	平成17年11月4日(2005.11.4)		ノキア コーポレイション
(65) 公表番号	特表2009-515212 (P2009-515212A)		フィンランド エフイーエンー02150
(43) 公表日	平成21年4月9日(2009.4.9)		エスプー ケイララーデンティエ 4
(86) 国際出願番号	PCT/IB2005/003293	(74) 代理人	100099759
(87) 国際公開番号	W02007/052088		弁理士 青木 篤
(87) 国際公開日	平成19年5月10日(2007.5.10)	(74) 代理人	100092624
審査請求日	平成20年5月2日(2008.5.2)		弁理士 鶴田 準一
		(74) 代理人	100102819
			弁理士 島田 哲郎
		(74) 代理人	100108383
			弁理士 下道 晶久
		(74) 代理人	100114018
			弁理士 南山 知広

最終頁に続く

(54) 【発明の名称】 オーディオ圧縮

(57) 【特許請求の範囲】

【請求項1】

オーディオ信号をエンコーディングする方法において、
 入力オーディオ信号を受信する段階と、
 前記オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割する段階と、
 前記高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割する段階と、
低周波数信号サンプルを前記受信入力オーディオ信号の高周波数サブ帯域と比較することにより、低周波数帯域信号セクションと前記高周波数サブ帯域信号との間の類似性尺度を判定する段階と、
低周波数帯域信号セクションを量子化及びコーディングする段階と、
前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域信号セクションを少なくとも参照するパラメータを生成する段階と、
前記入力信号を時間的に連続したフレームに分割する段階と、
前記入力信号の2つの連続したフレーム内において調性セクションを検出する段階と、
少なくとも1つの正弦波の位置及び振幅を表すパラメータを調性セクションのパラメータに追加する段階と、
前記コーディング済みの低周波数信号セクションと前記パラメータとを有するビットストリームを作成する段階と、

10

20

を有する方法。

【請求項 2】

パラメータを生成する段階は、前記低周波数帯域信号セクションをスケールリングするための少なくとも 1 つのスケールリングファクタを生成する段階を更に有する、請求項 1 記載の方法。

【請求項 3】

前記スケールリングファクタは、前記パラメータを使用して前記高周波数サブ帯域信号に転位される前記低周波数信号セクションのエンベロープが、前記受信信号の前記高周波数サブ帯域信号のエンベロープと連続するように生成される、請求項 2 記載の方法。

【請求項 4】

スケールリングファクタを生成する段階は、前記スペクトル内の少なくとも振幅ピークにマッチングするように、線形ドメイン内においてスケールリングファクタを生成する段階を有する、請求項 2 記載の方法。

【請求項 5】

スケールリングファクタを生成する段階は、前記スペクトルの少なくともエネルギー及び/又は形状にマッチングするように、対数ドメイン内においてスケールリングファクタを生成する段階を有する、請求項 2 記載の方法。

【請求項 6】

パラメータを生成する段階は、前記対応する高周波数サブ帯域信号を表す低周波数信号セクションに対するリンクを生成する段階を有する、請求項 1 記載の方法。

【請求項 7】

前記低周波数帯域信号セクションの中において高周波数サブ帯域信号に最良にマッチングしているものを判定する段階は、

A) 正規化された相関と、

B) ユークリッド距離と、

の中の少なくとも 1 つを使用する段階を有する、請求項 1 記載の方法。

【請求項 8】

少なくとも前記低周波数信号セクションのサンプルは、修正離散コサイン変換を使用することにより生成される、請求項 1 記載の方法。

【請求項 9】

前記高周波数サブ帯域信号の前記エンベロープを正規化する段階を更に有する、請求項 1 記載の方法。

【請求項 10】

前記低周波数信号のサンプルを量子化する段階と、前記少なくとも 1 つのスケールリングファクタを量子化する段階と、を更に有する、請求項 2 記載の方法。

【請求項 11】

調性セクションを検出する段階は、シフト離散フーリエ変換を使用する段階を有する、請求項 1 記載の方法。

【請求項 12】

調性セクション用の前記高周波数サブ帯域の数を増大させる段階を更に有する、請求項 1 記載の方法。

【請求項 13】

オーディオ信号をデコーディングする方法において、

エンコーディング済みのビットストリームを受信する段階と、

前記ビットストリームから、少なくとも 1 つの低周波数信号と、低周波数信号セクション並びに少なくとも 1 つの正弦波の位置及び振幅を参照する少なくともパラメータと、をデコーディングする段階であって、前記パラメータは、前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域信号セクションを少なくとも参照するものである、段階と、

10

20

30

40

50

前記低周波数信号のサンプルと、前記低周波数信号セクション並びに正弦波の位置及び振幅を参照する前記パラメータと、を利用することにより、少なくとも2つの高周波数サブ帯域信号を再構築する段階と、

前記少なくとも1つの低周波数信号及び前記少なくとも2つの高周波数サブ帯域信号を有する出力信号を出力する段階であって、前記少なくとも2つの高周波数サブ帯域が、前記少なくとも1つのデコーディング済みの低周波数信号と前記パラメータとから再構築されている、段階と、

を有する方法。

【請求項14】

オーディオ信号をエンコーディングするエンコーダにおいて、

入力オーディオ信号を受信するべく構成された受信手段と、

前記オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割し、前記高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割するべく構成されており、更に、前記入力信号を時間的に連続したフレームに分割し、前記入力信号の2つの連続したフレーム内において調性セクションを検出するべく構成されたフィルタリング手段と、

低周波数信号サンプルを前記受信入力オーディオ信号の高周波数サブ帯域と比較することにより、低周波数帯域信号セクションと前記高周波数サブ帯域信号との間の類似性尺度を判定し、前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域信号セクションを少なくとも参照するパラメータを生成し、少なくとも1つの正弦波の位置及び振幅を表すパラメータを調性セクションのパラメータに追加するべく構成され、更に、低周波数帯域信号セクションを量子化及びコーディングするべく構成されたコーディング手段と、

前記コーディング済みの低周波数信号セクションと前記パラメータとを有するビットストリームを作成する出力手段と、

を有するエンコーダ。

【請求項15】

前記コーディング手段は、前記低周波数帯域信号セクションをスケールリングするための少なくとも1つのスケールリングファクタを生成するべく構成されている、請求項14記載のエンコーダ。

【請求項16】

前記コーディング手段は、前記パラメータを使用して高周波数サブ帯域信号に転位される前記低周波数信号セクションのエンベロープが、前記受信信号の前記高周波数サブ帯域信号のエンベロープと連続するように、前記スケールリングファクタを生成するべく構成されている、請求項14記載のエンコーダ。

【請求項17】

前記フィルタリング手段は、シフト離散フーリエ変換を使用して調性セクションを検出するべく構成されている、請求項14記載のエンコーダ。

【請求項18】

前記コーディング手段は、調性セクション用の前記高周波数サブ帯域の数を増大させるべく構成されている、請求項14記載のエンコーダ。

【請求項19】

オーディオ信号をデコーディングするデコーダにおいて、

エンコーディング済みのビットストリームを受信するべく構成された受信手段と、

前記ビットストリームから、少なくとも1つの低周波数信号と、低周波数信号セクション並びに少なくとも1つの正弦波の位置及び振幅を参照する少なくともパラメータと、をデコーディングするべく構成されたデコーディング手段であって、前記パラメータは、前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域

10

20

30

40

50

信号セクションを少なくとも参照するものである、デコーディング手段と、

前記低周波数信号のサンプルと、前記低周波数信号セクション並びに正弦波の位置及び振幅を参照する前記パラメータと、を利用して少なくとも2つの高周波数サブ帯域信号を再構築するべく構成された生成手段であって、前記少なくとも2つの高周波数サブ帯域が、前記少なくとも1つのデコーディング済みの低周波数信号と前記パラメータとから再構築されている、生成手段と、

を有するデコーダ。

【請求項20】

請求項19記載のデコーダと、請求項14記載のエンコーダと、を有するデジタルオーディオ圧縮用のシステム。

【請求項21】

オーディオ信号をエンコーディングするためのコンピュータプログラムにおいて、前記プログラムは、

入力オーディオ信号を受信する段階と、

前記オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割する段階と、

前記高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割する段階と、

低周波数信号サンプルを前記受信入力オーディオ信号の高周波数サブ帯域と比較することにより、低周波数帯域信号セクションと前記高周波数サブ帯域信号との間の類似性尺度を判定する段階と、

低周波数帯域信号セクションを量子化及びコーディングする段階と、

前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域信号セクションを少なくとも参照するパラメータを生成する段階と、

前記入力信号を時間的に連続したフレームに分割する段階と、

前記入力信号の2つの連続したフレーム内において調性セクションを検出する段階と、少なくとも1つの正弦波の位置及び振幅を表すパラメータを調性セクションのパラメータに追加する段階と、

前記コーディング済みの低周波数信号セクションと前記パラメータとを有するビットストリームを作成する段階と、

をプロセッサに実行させるべく動作可能である命令を有する、コンピュータプログラム

。【請求項22】

シフト離散フーリエ変換を使用して調性セクションを検出する段階をプロセッサに実行させるべく動作可能である、請求項21記載のコンピュータプログラム。

【請求項23】

調性セクション用の前記高周波数サブ帯域の数を増大させる段階をプロセッサに実行させるべく動作可能である、請求項21記載のコンピュータプログラム。

【請求項24】

ビットストリームをデコーディングするためのコンピュータプログラムにおいて、前記プログラムは、

エンコーディング済みのビットストリームを受信する段階と、

前記ビットストリームから、少なくとも1つの低周波数信号と、低周波数信号セクション並びに少なくとも1つの正弦波の位置及び振幅を参照する少なくともパラメータと、をデコーディングする段階であって、前記パラメータは、前記低周波数帯域信号セクションを高周波数に転位して前記高周波数サブ帯域を再構築するために、前記高周波数サブ帯域信号に対して最高の類似性尺度を有する前記低周波数帯域信号セクションを少なくとも参照するものである、段階と、

前記低周波数信号のサンプルと、前記低周波数信号セクション並びに正弦波の位置及び振幅を参照する前記パラメータと、を利用することにより、少なくとも2つの高周波数サ

10

20

30

40

50

ブ帯域信号を再構築する段階と、

前記少なくとも1つの低周波数信号及び前記少なくとも2つの高周波数サブ帯域信号を有する出力信号を出力する段階であって、前記少なくとも2つの高周波数サブ帯域が、前記少なくとも1つのデコーディング済みの低周波数信号と前記パラメータとから再構築されている、段階と、

をプロセッサに実行させるべく動作可能である命令を有する、コンピュータプログラム

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、一般に、オーディオ圧縮に関するものである。

【背景技術】

【0002】

オーディオ圧縮は、一般に、デジタルオーディオ信号を保存又は伝送するべく、最近の消費者装置において採用されている。消費者装置は、通信装置、ビデオ装置、オーディオプレーヤー、無線装置、及びその他の消費者装置であってよい。高圧縮比により、相対的に大きなストレージ容量、或いは、通信チャネル、即ち、無線通信チャネル又は有線通信チャネルを介した相対的に効率的な伝送が実現される。但し、圧縮比と同時に、圧縮された信号の品質をハイレベルに維持する必要がある。オーディオコーディングの目的は、一般に、所与の圧縮比、即ち、ビットレートとの関係において、オーディオ品質を極大化させることにある。

【0003】

過去数十年の間に、多数のオーディオコーディング法が開発されている。先進的なオーディオコーディングシステムは、人間の耳の特性を効果的に利用している。主要な概念は、知覚品質に影響を及ぼすことが最も少ない信号の領域にコーディング雑音を配置可能であり、この結果、可聴歪みを導入することなしにデータレートを低減可能であるというものである。従って、音響心理学の理論が最近のオーディオコーディングの重要な部分を構成している。

【0004】

既知のオーディオエンコーダにおいては、入力信号を限られた数のサブ帯域に分割している。これらのサブ帯域信号のそれぞれを量子化可能である。スペクトル内の最高周波数は、低い周波数よりも、知覚的な重要性が低いことが音響心理学の理論から判明している。これは、コーダ内において、低周波数サブ帯域の場合よりも、相対的に少ない数のビットを高周波数サブ帯域の量子化に割り当てることにより、ある程度考慮可能である。

【0005】

更に高度なオーディオコーディングにおいては、大部分の場合に、オーディオ信号の低周波数領域と高周波数領域の間には、大きな依存性が存在している、即ち、スペクトルの上半分は、一般に、下半分と非常に類似している、という事実を利用している。低周波数領域とは、オーディオスペクトルの下半分である見なすことが可能であり、高周波数は、オーディオスペクトルの上半分である見なすことができる。低周波数と高周波数の間の境界は、固定されてはいないが、2kHz～15kHzの間に位置可能であり、場合によっては、これらの境界を超過することも可能であることを理解されたい。

【0006】

高周波数領域をコーディングするための現在の方法は、SBR(Spectral-Band-Replication)と呼ばれている。この技法については、M. Dietz、L. Liljeryd、K. Kjorling、及びO. Kunzによる「Spectral Band Replication, a novel approach in audio coding」(112th AES Convention, Munich, Germany, 2002年5月)及びP. Ekstrandによる「Bandwidth extension of audio signals b

10

20

30

40

50

y spectral band replication」(1st IEEE Benelux Workshop on Model Based Processing and Coding of Audio、Leuven, Belgium、2002年11月)に記述されている。この記述されている方法は、例えば、AAC又はMP3(MPEG-1 Layer III)コーデなどの一般的なオーディオコーデ、並びに、多くのその他の最新技術のコーデに適用可能である。

【0007】

当技術分野における方法の欠点は、低周波数帯域の高周波数帯域への単純な転位が、オリジナルの高周波数と、転位された低周波数を利用したそれらの再構築との間の不同性に結び付く可能性を有しているという点にある。別の欠点は、雑音及び正弦波を既知の方法に従って周波数スペクトルに追加する必要があるという点にある。

10

【発明の開示】

【発明が解決しようとする課題】

【0008】

従って、改善されたオーディオコーディング法を提供することが本出願の目的である。適度に低いビットレートによって入力信号を更に正確に表現するコーディング法を提供することが本出願の更なる目的である。

【課題を解決するための手段】

【0009】

前述の欠点を克服するべく、本出願は、一態様によれば、入力オーディオ信号を受信する段階と、オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割する段階と、高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割する段階と、低周波数帯域信号セクションの中において高周波数サブ帯域信号と最良にマッチングしているものを判定する段階と、高周波数サブ帯域信号と最良にマッチングしている低周波数帯域信号セクションを少なくとも参照するパラメータを生成する段階と、を有するオーディオ信号をエンコーディングする方法を提供している。

20

【0010】

本出願は、入力信号の高周波数領域をコーディングする新しい方法を提供している。入力信号を時間的に連続したフレームに分割可能である。フレームのそれぞれは、入力信号の時間的なインスタンスを表している。それぞれのフレーム内においては、そのスペクトル成分によって入力信号を表現可能である。これらのスペクトル成分、即ち、サンプルは、入力信号内における周波数を表している。

30

【0011】

低周波数領域を高周波数領域に盲目的に転位させる代わりに、本出願は、オリジナルの及びコーディング済みの高周波数スペクトル成分の間の類似性を極大化させている。本出願によれば、信号の既にコーディング済みの低周波数領域を利用することにより、高周波数領域を形成している。

【0012】

低周波数信号サンプルを受信信号の高周波数サブ帯域と比較することにより、実際の高周波数サブ帯域と最良にマッチングしている低周波数における信号セクションを見出すことが可能である。本出願は、低周波数スペクトルサンプルの全体において、高周波数サブ帯域に最良に類似している信号セクションについてサンプルごとにサーチする段階を提供している。信号セクションは、サンプルシーケンスに対応していることから、本出願は、換言すれば、高周波数サブ帯域と最良にマッチングしているサンプルシーケンスを見出す段階を提供している。このサンプルシーケンスは、低周波数帯域内における最後の検討された開始点は、低周波数帯域内の最後のサンプルからマッチング対象である高周波数サブ帯域の長さを減算したものである必要があるという点を除いて、低周波数帯域内の任意の場所において開始可能である。

40

【0013】

実際の高周波数サブ帯域に最良にマッチングしている低周波数信号セクションに対する

50

インデックス又はリンクを使用することにより、高周波数サブ帯域をモデル化可能である。対応した高周波数サブ帯域の表現の受信側における回復を実現するには、このインデックス又はリンクをエンコードし、保存又は伝送するだけでよい。

【0014】

実施例によれば、信号セクションと高周波数サブ帯域の最も類似したマッチング、即ち、最も類似したスペクトル形状を低周波数帯域内においてサーチしている。高周波数サブ帯域に最も類似していると判明した信号セクションを少なくとも参照するパラメータをエンコーダ内において生成している。これらのパラメータは、検出されたセクションを高周波数帯域内にスケールリングするスケールリングファクタを有することができる。デコーダ側においては、これらのパラメータを使用して対応する低周波数信号セクションを高周波数領域に転位させることにより、高周波数サブ帯域を再構築している。

10

【0015】

スケールリングファクタを使用することにより、複製された低周波数信号セクションに対してスケールリングを適用可能である。実施例によれば、エンコードする必要があるのは、スケールリングファクタ及び低周波数信号セクションに対するリンクのみである。

【0016】

高周波数サブ帯域を再生成するべく最良にマッチングしている低周波数信号セクションを使用した際には、高周波数領域の形状は、既知の方法によるものよりも、オリジナルの高周波数スペクトルに更に密接に類似している。振幅、形状、及び周波数の位置が、オリジナルの信号に対して更に類似しているため、知覚的に重要なスペクトルピークを更に正確にモデル化可能である。モデル化された高周波数サブ帯域をオリジナルの高周波数サブ帯域と比較可能であることから、消失したスペクトル成分、即ち、正弦波又は雑音を容易に検出し、次いで、これらを追加可能である。

20

【0017】

エンベロープ整形を実現するべく、実施例は、パラメータを使用して低周波数信号サンプルを高周波数サブ帯域信号に転位させることによって低周波数信号セクションを利用する段階を提供しており、この場合に、これらのパラメータは、転位した低周波数信号セクションのエンベロープが、受信信号の高周波数サブ帯域信号のエンベロープと連続するように、スケールリングファクタを有している。スケールリングファクタにより、複製された低周波数信号セクションのエネルギー及び形状を調節し、実際の高周波数サブ帯域に更に良好にマッチングさせることが可能である。

30

【0018】

パラメータは、実施例による対応した高周波数サブ帯域信号を表すための低周波数信号セクションに対するリンクを有することが可能である。これらのリンクは、低周波数信号セクションに対するポインタ又はインデックスであってよい。この情報により、高周波数サブ帯域を構築する際に低周波数信号セクションを参照可能である。

【0019】

量子化ビットの数を低減するべく、高周波数サブ帯域信号のエンベロープを正規化可能である。正規化により、低周波数帯域と高周波数帯域の両方が、正規化された振幅レンジ内に存在することになる。これにより、スケールリングファクタの量子化に必要なビットの数が低減される。正規化に使用した情報をエンコーダから提供することにより、デコーダ内において高周波数サブ帯域の表現を構築しなければならない。実施例は、線形予測コーディングによるエンベロープの正規化を提供している。又、ケプストラムモデル化を利用することにより、エンベロープを正規化することも可能である。ケプストラムモデル化は、信号のパワースペクトルの対数の逆フーリエ変換を使用している。

40

【0020】

スケールリングファクタの生成段階は、スペクトル内の少なくとも振幅ピークにマッチングするべく線形ドメイン内においてスケールリングファクタを生成する段階を有することができる。又、スケールリングファクタを生成する段階は、実施例に従って、対数ドメイン内において少なくともスペクトルのエネルギー及びノイズ又は形状をマッチングさせる段階を有

50

することも可能である。

【0021】

実施例は、修正離散コサイン変換(MDCT: Modified Discrete cosine transformation)を使用して低周波数帯域及び/又は高周波数帯域内において信号サンプルを生成する段階を提供している。MDCT変換は、好ましくは、実数としてスペクトル係数を提供している。実施例によるMDCT変換は、任意の適切なフレームサイズ、具体的には、通常フレームの場合には、2048個のサンプルの、そして、過渡フレームの場合には、256個のサンプルのフレームサイズにおいて、使用可能であるが、この間の任意のその他の値を使用することも可能である。

【0022】

対応した高周波数サブ帯域信号に最良にマッチングしている低周波数信号セクションを取得するべく、実施例は、正規化された相関又はユークリッド距離を使用して類似性尺度を算出する段階を提供している。

【0023】

入力信号をエンコーディングするべく、実施例は、低周波数信号サンプルを量子化する段階と、少なくともスケールングファクタを量子化する段階と、を提供している。低周波数信号セクションに対するリンクは、整数であってよい。

【0024】

追加の正弦波を追加することにより、高周波数信号の品質を改善可能である。このような正弦波に対応するべく、実施例は、入力信号を時間的に連続したフレームに分割する段階と、入力信号の2つの連続したフレーム内において調性セクション(tonal section)を検出する段階と、を提供している。調性セクションは、追加の正弦波を追加することにより、機能強化可能である。高度な調性セクションは、更に、対応する高周波数領域内における高周波数サブ帯域の数を増大させることにより、更に機能強化可能である。入力信号は、例えば、非調性(not tonal)、調性(tonal)、及び強力な調性(strongly tonal)などの異なる調性グループ(tonality group)に分割可能である。

【0025】

調性セクションを検出する段階は、シフト離散フーリエ変換(SDFT: Shifted Discrete Fourier Transformation)を使用する段階を有することができる。SDFTの結果をエンコーダ内において利用することにより、MDCT変換を提供可能である。

【0026】

本出願の別の態様は、エンコーディング済みのビットストリームを受信する段階と、このビットストリームから少なくとも1つの低周波数信号及び低周波数信号セクションを参照する少なくともパラメータをデコーディングする段階と、低周波数信号サンプル及び低周波数信号セクションを参照するパラメータを利用して少なくとも2つの高周波数サブ帯域信号を再構築する段階と、少なくとも1つの低周波数信号及び少なくとも2つの高周波数サブ帯域信号を有する出力信号を出力する段階と、を有するオーディオ信号をデコーディングする方法である。

【0027】

本出願の更なる態様は、入力オーディオ信号を受信するべく構成された受信手段と、オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割するべく構成されており、且つ、高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割するべく更に構成されたフィルタリング手段と、高周波数サブ帯域信号に最良にマッチングしている低周波数帯域信号セクションを少なくとも参照するパラメータを生成するべく構成されたコーディング手段と、を有するオーディオ信号をエンコーディングするエンコーダである。

【0028】

更には、本出願の更なる態様は、エンコーディング済みのビットストリームを受信する

10

20

30

40

50

べく構成された受信手段と、このビットストリームから少なくとも1つの低周波数信号及び低周波数信号セクションを参照する少なくともパラメータをデコーディングするべく構成されたデコーディング手段と、低周波数信号のサンプル及び低周波数信号セクションを参照するパラメータを利用して少なくとも2つの高周波数サブ帯域信号を再構築するべく構成された生成手段と、を有するオーディオ信号をデコーディングするデコーダである。

【0029】

本出願の更なる態様は、前述のデコーダと、前述のエンコーダと、を有するデジタルオーディオ圧縮用のシステムである。

【0030】

更には、本出願の更なる態様は、オーディオ信号をエンコーディングするコンピュータプログラムプロダクトに関するものであり、このプログラムは、入力オーディオ信号を受信する段階と、オーディオ信号を少なくとも1つの低周波数帯域及び1つの高周波数帯域に分割する段階と、高周波数帯域を少なくとも2つの高周波数サブ帯域信号に分割する段階と、高周波数サブ帯域信号に最良にマッチングしている低周波数帯域信号セクションを少なくとも参照するパラメータを生成する段階と、をプロセッサに実行させるべく動作可能である命令を有している。

【0031】

又、ビットストリームをデコーディングするコンピュータプログラムプロダクトであって、プログラムは、エンコーディング済みのビットストリームを受信する段階と、このビットストリームから少なくとも1つの低周波数信号及び低周波数信号セクションを参照する少なくともパラメータをデコーディングする段階と、低周波数信号のサンプル及び低周波数信号セクションを参照するパラメータを利用して少なくとも2つの高周波数サブ帯域信号を再構築する段階と、少なくとも1つの低周波数信号及び少なくとも2つの高周波数サブ帯域信号を有する出力信号を出力する段階と、をプロセッサに実行させるべく動作可能である命令を有している。

【発明を実施するための最良の形態】

【0032】

一般的なオーディオコーディングシステムは、図1に概略的に示されているように、エンコーダ及びデコーダから構成されている。図示されているのは、エンコーダ4、ストレージ又は媒体チャンネル6、及びデコーダ8を有するコーディングシステム2である。

【0033】

エンコーダ4は、入力オーディオ信号10を圧縮することにより、ビットストリーム12を生成しており、このビットストリームは、保存されるか、又は媒体チャンネル6を通じて伝送される。ビットストリーム12は、デコーダ8内において受信可能である。デコーダ8は、ビットストリーム12を圧縮解除することにより、出力オーディオ信号14を生成している。ビットストリーム12のビットレートと、入力信号10との関係における出力オーディオ信号14の品質が主要な特徴であり、コーディングシステム2の性能は、これらによって定義されている。

【0034】

最近のオーディオエンコーダ4の代表的な構造が図2に概略的に示されている。入力信号10を分析フィルタバンク構造16を使用することにより、サブ帯域に分割している。音響心理学モデル20から供給される情報を利用することにより、コーディング手段18内において、それぞれのサブ帯域を量子化及びコーディング可能である。コーディングは、Huffmanコーディングであってよい。量子化の設定とコーディングの方式は、音響心理学モデル18によって決定可能である。量子化及びコーディング済みの情報をビットストリームフォーマッタ22内において使用することにより、ビットストリーム12を生成している。

【0035】

ビットストリーム12は、図3に概略的に示されているように、デコーダ8内においてデコーディング可能である。デコーダ8は、ビットストリームアンパッキング手段24、

10

20

30

40

50

サブ帯域再構築手段 26、及び合成フィルタバンク 28 を有することができる。

【0036】

デコーダ 8 は、エンコーダ 4 の逆演算を実行しており、変換によってビットストリーム 12 を出力オーディオ信号 14 に戻している。デコーディングプロセスにおいては、サブ帯域再構築手段 26 内において、ビットストリーム 12 をサブ帯域信号に逆量子化している。これらのサブ帯域信号を合成フィルタバンク 28 に供給し、この合成フィルタバンクが、サブ帯域信号からオーディオ信号を合成し、出力信号 14 を生成している。

【0037】

多くの場合に、低周波数領域及び限られた量の追加制御情報のみを使用することにより、高周波数領域を効率的且つ知覚的に正確に合成可能である。最適には、高周波数部分のコーディングに必要なとされるのは、少数の制御パラメータのみである。スペクトルの上部部分の全体を少量の情報によって合成可能であるため、合計ビットレートにおける相当の節約を実現可能である。

10

【0038】

MP3pro などの現在のコーディングは、音響心理学的なコーディングに加えて、SBR コーディング方式を導入することにより、これらの特性をオーディオ信号内において利用している。SBR においては、図 4 及び図 5 に概略的に示されているように、コーディング済みの低周波数領域を利用することにより、高周波数領域を別個に生成可能である。

【0039】

図 4 は、エンコーダ 4 を概略的に示している。エンコーダ 4 は、低域通過フィルタリング手段 30、コーディング手段 31、SBR 手段 32、エンベロープ抽出手段 34、及びビットストリームフォーマッタ 22 を有している。

20

【0040】

低域通過フィルタ 30 は、まず、入力信号 10 をフィルタリングするためのカットオフ周波数を定義している。この効果が図 6 a に示されている。カットオフ周波数 36 未満の周波数のみがフィルタを通過している。

【0041】

コーディング手段 31 は、32 個の低周波数サブ帯域を伴う量子化及び Huffman コーディングを実行している。コーディング手段 31 内において、低周波数コンテンツを QMF ドメインに変換している。この低周波数コンテンツをコーダ 31 の出力に基づいて転位させている。この転位は、SBR 手段 32 内において実行している。低周波数の高周波数への転位の効果が図 6 b に示されている。この転位は、低周波数サブ帯域サンプルが高周波数サブ帯域サンプル内にそのまま複製されるように、盲目的に実行されている。これは、入力信号のすべてのフレーム内において同様に、且つ、入力信号の特性とは無関係に、実行されている。

30

【0042】

SBR 手段 32 内においては、高周波数サブ帯域を追加情報に基づいて調節可能である。これは、合成された高周波数領域の特定の特徴をオリジナルのものに類似したものにすべく実行されている。正弦波又は雑音などの追加成分を高周波数領域に追加することにより、オリジナルの高周波数領域との類似性を増大させることが可能である。最終的に、オリジナルの高周波数スペクトルのエンベロープと連続するように、エンベロープ抽出手段 34 内において、エンベロープを調節している。この効果を図 6 c において観察可能であり、この場合には、入力信号の実際の高周波数成分により近接したものとなるように、高周波数成分をスケールアップしている。

40

【0043】

ビットストリーム 12 内において、コーディング済みの低周波数信号は、スケールアップ及びエンベロープ調節パラメータと共に存在している。図 5 に示されているように、このビットストリーム 12 は、デコーダ内においてデコーディング可能である。

【0044】

50

図5は、アンパッキング手段24、低周波数デコーダ38、高周波数再構築手段40、成分調節手段42、及びエンベロープ調節手段44を有するデコーダ8を示している。デコーダ38内において、低周波数サブ帯域を再構築している。高周波数再構築手段40内において、これらの低周波数サブ帯域から高周波数サブ帯域を静的に再構築している。成分調節手段42内において、正弦波を追加可能であり、エンベロープ調節手段44内において、エンベロープを調節可能である。

【0045】

本出願によれば、低周波数信号サンプルの高周波数サブ帯域への転位は、動的に実行されており、例えば、どの低周波数信号セクションが高周波数サブ帯域に最良にマッチングしているかをチェックしている。対応した低周波数信号セクションに対するインデックスを生成している。このインデックスをエンコードし、デコーダ内において使用することにより、低周波数信号から高周波数サブ帯域を構築している。

10

【0046】

図7は、エンコーダ4と、デコーダ8と、を有するコーディングシステムを示している。エンコーダ4は、高周波数コーディング手段50、低周波数コーダ52、及びビットストリームフォーマッタ22から構成されている。エンコーダ4は、更に複雑なオーディオコーディング体系の一部であってよい。本出願は、低ビットレートにおいて良好な品質を目的としている大部分の任意のオーディオコーダ内において使用可能である。例えば、実際の低ビットレートオーディオコーダとは完全に別個に本出願を使用可能であり、例えば、AACやMP3などの音響心理学的なコーダの前段に配置可能である。

20

【0047】

高周波数領域は、通常、低周波数領域と類似したスペクトル形状を含んでことから、一般的に、良好なコーディング性能が実現される。デコーダに伝送する必要があるのは、複写されたスペクトルのインデックス及びスケールリングファクタのみであることから、これは、相対的に低い合計ビットレートによって実現される。

【0048】

低周波数コーダ22内において、低周波数サンプル $X_L(k)$ をコーディングしている。高周波数コーダ50内においては、後程詳述するように、コーディングのために、変換、スケールリング、及びエンベロープ形成を表すパラメータ α_1 、 α_2 、 i を生成している。

【0049】

まず、高周波数スペクトルを n_b 個のサブ帯域に分割している。それぞれのサブ帯域ごとに、最も類似したマッチング(即ち、最も類似したスペクトル形状)を低周波数領域からサーチしている。

30

【0050】

この方法は、MDCT(Modified Discrete Cosine transformation)ドメイン内において動作可能である。その良好な特性(クリティカルサンプリングとの50%のオーバーラップや柔軟なウィンドウ切り替えなど)に起因し、MDCTドメインは、大部分の最新技術のオーディオコーダ内において使用されている。MDCT変換は、次式のように実行している。

【0051】

【数1】

$$X(k) = \sum_{n=0}^{2N-1} h(n)x(n) \cos \left[\frac{2\pi}{N} \left(k + \frac{1}{2} \right) \left(n + \frac{1}{2} + \frac{N}{2} \right) \right] \quad (1)$$

40

【0052】

ここで、 $x(n)$ は、入力信号であり、 $h(n)$ は、長さ $2N$ を有する時間分析ウィンドウであって、 $0 \leq k < N$ である。通常、オーディオコーディングにおいては、 N は、1024個(通常フレーム)又は128個(過渡)のサンプルである。スペクトル係数 $X(k)$

50

k) は、実数であってよい。前述のフレームサイズ、並びに、任意のその他のフレームサイズが可能である。

【 0 0 5 3 】

高周波数サブ帯域を表すパラメータを生成するべく、高周波数コーダ 5 0 内において、高周波数サブ帯域に最もマッチングしている低周波数信号セクションを見出す必要がある。高周波数コーダ 5 0 及び低周波数コーダ 5 2 は、 N 個の M D C T コーディング済みの成分を生成可能であり、ここで、 $X_L(k)$ は、低周波数成分を表しており、 $X_H(k)$ は、高周波数成分を表している。

【 0 0 5 4 】

低周波数コーダ 5 2 により、 N_L 個の低周波数 M D C T 係数

10

【 0 0 5 5 】

【 数 2 】

$$\hat{X}_L(k)$$

【 0 0 5 6 】

をコーディング可能であり、ここで、 $0 \leq k < N_L$ である。通常は、 $N_L = N / 2$ であるが、その他の選択も実行可能である。

20

【 0 0 5 7 】

オリジナルのスペクトル $X(k)$ 及び

【 0 0 5 8 】

【 数 3 】

$$\hat{X}_L(k)$$

【 0 0 5 9 】

を利用している目的は、使用する尺度により、オリジナルの高周波数信号 $X_H(k) = X(N_L + k)$ に最大限に類似した高周波数成分

30

【 0 0 6 0 】

【 数 4 】

$$\hat{X}_H(k)$$

【 0 0 6 1 】

を生成することであり、ここで、 $0 \leq k < N - N_L$ である。

40

【 0 0 6 2 】

【 数 5 】

$$\hat{X}_L(k)$$

【 0 0 6 3 】

及び

50

【 0 0 6 4 】

【 数 6 】

$$\hat{X}_H(k)$$

【 0 0 6 5 】

が、協働し、次のように、合成されたスペクトル

【 0 0 6 6 】

【 数 7 】

10

$$\hat{X}(k)$$

【 0 0 6 7 】

を形成している。

【 0 0 6 8 】

【 数 8 】

20

$$\hat{X}(k) = \begin{cases} \hat{X}_L(k), & 0 \leq k \leq N_L \\ \hat{X}_H(k), & N_L \leq k < N \end{cases} \quad (2)$$

【 0 0 6 9 】

オリジナルの高周波数スペクトル $X_H(k)$ を n_b 個のオーバーラップしていない帯域に分割している。原則的に、帯域の数と帯域の幅は、任意に選択可能である。例えば、 N が 1024 個のサンプルに等しい際には、8 個の等しい幅の周波数帯域を使用可能である。別の適度な選択肢は、人間の聴力の知覚的特性に基づいて帯域を選択するというものである。例えば、Bark 又は ERB (Equivalent Rectangular Bandwidth) スケールを利用することにより、帯域の数とそれらの幅を選択可能である。

30

【 0 0 7 0 】

高周波数コーダ内において、高周波数信号及び低周波数成分の間の類似性尺度を算出可能である。

【 0 0 7 1 】

w_j 個のサンプルの長さを有する $X_H(k)$ の j 番目の帯域を含む列ベクトルを

【 0 0 7 2 】

【 数 9 】

40

$$X_H^j$$

【 0 0 7 3 】

としよう。

【 0 0 7 4 】

【数 1 0】

$$X_H^j$$

【0 0 7 5】

を、次のように、コーディング済みの低周波数スペクトル

【0 0 7 6】

【数 1 1】

10

$$\hat{X}_L(k)$$

【0 0 7 7】

と比較可能である。

【0 0 7 8】

【数 1 2】

20

$$\max_{i(j)} (S(\hat{X}_L^{i(j)}, X_H^j)), 0 \leq i(j) < N_L - w_j \quad (3)$$

【0 0 7 9】

ここで、 $S(a, b)$ は、ベクトル a 及び b の間の類似性尺度であり、

【0 0 8 0】

【数 1 3】

30

$$\hat{X}_L^{i(j)}$$

【0 0 8 1】

は、コーディング済みの低周波数スペクトル

【0 0 8 2】

【数 1 4】

40

$$\hat{X}_L(k)$$

【0 0 8 3】

のインデックス $i(j)$ $k < i(j) + w_j$ を含むベクトルである。望ましい低周波数信号セクションの長さは、現在の高周波数サブ帯域の長さと同じであり、従って、基本的に、必要な唯一の情報は、インデックス $i(j)$ であり、これは、個々の低周波数信号セクションがどこで始まっているのかを示している。

【0 0 8 4】

類似性尺度を使用することにより、最高の類似性を提供するインデックス $i(j)$ を選択可能である。類似性尺度を使用することにより、ベクトルの形状がどれほど類似してい

50

るかを表しており、それらの相対的な振幅は、重要ではない。類似性尺度には、多数の選択肢が存在している。1つの可能な実装は、次のように、正規化された相関であってよい。

【 0 0 8 5 】

【 数 1 5 】

$$S(a,b) = \frac{|b^T a|}{\sqrt{a^T a}} \quad (4)$$

10

【 0 0 8 6 】

これは、a 及び b の振幅の影響を受けにくい尺度を提供している。別の適度な代替肢は、次のように、ユークリッド距離に基づいた類似性尺度である。

【 0 0 8 7 】

【 数 1 6 】

$$S(a,b) = \frac{1}{\|a-b\|} \quad (5)$$

20

【 0 0 8 8 】

同様に、多くのその他の類似性尺度を利用可能である。

【 0 0 8 9 】

低周波数信号サンプル内のこれらの最も類似したセクションを高周波数サブ帯域に複写すると共に、特定のスケールリングファクタを使用することにより、スケールリング可能である。これらのスケールリングファクタは、コーディング済みの高周波数スペクトルのエンベロープが、オリジナルのスペクトルのエンベロープと連続するように、機能している。

【 0 0 9 0 】

インデックス $i(j)$ を使用することにより、

【 0 0 9 1 】

【 数 1 7 】

$$X_H^j$$

30

【 0 0 9 2 】

に形状において最も類似している選択されたベクトル

40

【 0 0 9 3 】

【 数 1 8 】

$$\hat{X}_L^{i(j)}$$

【 0 0 9 4 】

を

【 0 0 9 5 】

50

【数 19】

$$X_H^j$$

【0096】

と同一の振幅にスケールリングしなければならない。スケールリングには、多数の異なる技法が存在している。例えば、スケールリングは、まず、線形ドメインにおいてスペクトル内の高振幅ピークをマッチングさせた後に、対数ドメインにおいて、エネルギー及び形状をマ

10

【0097】

【数 20】

$$\hat{X}_L^{i(j)}$$

【0098】

をスケールリングすることにより、結果的に、コーディング済みの周波数成分

20

【0099】

【数 21】

$$\hat{X}_H^j$$

【0100】

が得られる。

30

【0101】

線形ドメインスケールリングは、単純に次のように実行される。

【0102】

【数 22】

$$\hat{X}_H^j = \alpha_1(j) \hat{X}_L^{i(j)} \quad (6)$$

【0103】

ここで、 $\alpha_1(j)$ は、次式から得られる。

40

【0104】

【数 23】

$$\alpha_1(j) = \frac{(\hat{X}_L^{i(j)})^T X_H^j}{(\hat{X}_L^{i(j)})^T \hat{X}_L^{i(j)}} \quad (7)$$

【0105】

50

$\alpha_1(j)$ が正及び負の両方の値を取得可能であることに留意されたい。対数スケールリングの前に、ベクトルサンプルの符号と

【 0 1 0 6 】

【 数 2 4 】

$$\hat{X}_H^j$$

【 0 1 0 7 】

の最大対数値を次のように保存している。

【 0 1 0 8 】

【 数 2 5 】

$$K_{\hat{X}_H^j} = \frac{\hat{X}_H^j}{|\hat{X}_H^j|} \quad (8)$$

$$M_{\hat{X}_H^j} = \max(\log_{10} |\hat{X}_H^j|) \quad (9)$$

【 0 1 0 9 】

この段階で、対数スケールリングを実行可能であり、

【 0 1 1 0 】

【 数 2 6 】

$$\hat{X}_H^j$$

【 0 1 1 1 】

を次のように更新する。

【 0 1 1 2 】

【 数 2 7 】

$$V_{\hat{X}_H^j} = \alpha_2(j)(\log_{10}(|\hat{X}_H^j|) - M_{\hat{X}_H^j}) + M_{\hat{X}_H^j} \quad (10)$$

$$\hat{X}_H^j = 10^{V_{\hat{X}_H^j}} (K_{\hat{X}_H^j})^T \quad (11)$$

【 0 1 1 3 】

ここで、スケールリングファクタ $\alpha_2(j)$ は、次式から得られる。

【 0 1 1 4 】

10

20

30

40

【数 2 8】

$$\alpha_2(j) = \frac{\left(\log_{10}(|\hat{X}_H^j|) - M_{\hat{X}_H^j}\right)^T \left(\log_{10}(|X_H^j|) - M_{\hat{X}_H^j}\right)}{\left(\log_{10}(|\hat{X}_H^j|) - M_{\hat{X}_H^j}\right)^T \left(\log_{10}(|X_H^j|) - M_{\hat{X}_H^j}\right)} \quad (12)$$

【0 1 1 5】

このスケーリングファクタは、対数ドメインにおける波形の間の類似性を極大化している。或いは、この代わりに、エネルギーが略等しいレベルに設定されるように、次のように、 $\alpha_2(j)$ を選択することも可能である。

【0 1 1 6】

【数 2 9】

$$\alpha_2(j) = \frac{\left\| \log_{10}(X_H^j) - M_{\hat{X}_H^j} \right\|}{\left\| \log_{10}(\hat{X}_H^j) - M_{\hat{X}_H^j} \right\|} \quad (13)$$

20

【0 1 1 7】

上記の式において、変数

【0 1 1 8】

【数 3 0】

$$M_{\hat{X}_H^j}$$

30

【0 1 1 9】

の目的は、

【0 1 2 0】

【数 3 1】

$$\hat{X}_H^j$$

【0 1 2 1】

内の最大の値の振幅（即ち、スペクトルピーク）が過大にスケーリングされないようにすることにある（第 1 スケーリングファクタ $\alpha_1(j)$ が、これらを既に正しいレベルに設定している）。変数

40

【0 1 2 2】

【数 3 2】

$$K_{\hat{X}_H^j}$$

50

【 0 1 2 3 】

を使用することにより、オリジナルサンプルの符号を保存しているが、この理由は、この情報が、対数ドメインへの変換の際に失われるためである。

【 0 1 2 4 】

帯域のスケーリングが完了した後に、ベクトル

【 0 1 2 5 】

【 数 3 3 】

$$\hat{X}_H^j$$

10

【 0 1 2 6 】

を合成することにより、合成済みの高周波数スペクトル

【 0 1 2 7 】

【 数 3 4 】

$$\hat{X}_H(k)$$

20

【 0 1 2 8 】

を取得可能であり、ここで、 $j = 0, 1, \dots, n_b - 1$ である。

【 0 1 2 9 】

パラメータの選択が完了した後に、高周波数領域再構築情報をデコーダ 8 に伝送するべく、これらのパラメータを量子化する必要がある。

【 0 1 3 0 】

デコーダ 8 内において

【 0 1 3 1 】

【 数 3 5 】

30

$$\hat{X}_H(k)$$

【 0 1 3 2 】

を再構築するには、それぞれの帯域について、パラメータ $i(j)$ 、 ${}_1(j)$ 、及び ${}_2(j)$ が必要である。デコーダ 8 内において、高周波数生成手段 5 4 がこれらのパラメータを利用している。インデックス $i(j)$ は、整数であるため、そのまま提出可能である。 ${}_1(j)$ 及び ${}_2(j)$ は、例えば、スカラー又はベクトル量子化を使用することにより、量子化可能である。

40

【 0 1 3 3 】

これらのパラメータの量子化されたバージョン

【 0 1 3 4 】

【 数 3 6 】

$$\hat{a}_i(j)$$

50

【 0 1 3 5 】

及び

【 0 1 3 6 】

【 数 3 7 】

$$\hat{a}_2(j)$$

【 0 1 3 7 】

を高周波数生成手段 5 4 内において使用することにより、式 (6) 及び式 (1 0) に従って

【 0 1 3 8 】

【 数 3 8 】

$$\hat{X}_H(k)$$

【 0 1 3 9 】

を構築している。

【 0 1 4 0 】

低周波数デコーディング手段 5 6 は、低周波数信号をデコーディングし、且つ、再構築された高周波数サブ帯域と共に、式 2 に従って出力信号 1 4 を形成している。

【 0 1 4 1 】

図 7 に示されているシステムは、エンベロープ正規化手段によって更に機能強化可能である。図 8 に示されているシステムは、図 7 に示されているシステムに加えて、エンベロープ正規化手段 5 8 及びエンベロープ合成手段 6 0 を有している。

【 0 1 4 2 】

このシステムにおいては、高周波数コーディング法を使用し、且つ、エンコーダ 4 内のエンベロープ正規化手段 5 8 を使用することにより、エンベロープ正規化スペクトルを生成している。実際のエンベロープの合成は、デコーダ 8 内の別個のエンベロープ合成手段 6 0 内において実行されている。

【 0 1 4 3 】

エンベロープの正規化は、例えば、LPC 分析又はケプストラムモデル化を使用することにより、実行可能である。図 8 に示されているように、エンベロープの正規化に伴い、オリジナルの高周波数スペクトルエンベロープを表すエンベロープパラメータをデコーダに提出しなければならないことに留意されたい。

【 0 1 4 4 】

SBR において、追加の正弦波及び雑音成分を高周波数領域に追加している。前述のアプリケーション内においても、同様の内容を実行可能である。必要に応じて、追加成分を容易に追加可能である。これは、この記述されている方法においては、オリジナルの及び合成されたスペクトルの間の差を計測可能であり、且つ、従って、スペクトル形状において大きな差が存在している場所を見出すことが可能であるためである。例えば、一般的な BWE コーダにおいては、スペクトル形状がオリジナルのスペクトルから大きく異なっているため、通常、追加の正弦波又は雑音成分を追加するべきかどうかを決定することが相対的に困難である。

【 0 1 4 5 】

いくつかの場合において、入力信号が顕著な調性を有している際には、オリジナルと比較した際に、コーディング済みの信号の品質が低下可能であることが判明した。これは、

10

20

30

40

50

コーディング済みの高周波数領域が、オリジナルの信号ほどには、フレーム間における周期性を有していないためである。いくつかの周期的（正弦波的）成分が消失可能であるため、又は既存の周期的成分の振幅がフレーム間において過剰に変化しているため、周期性が失われている。

【0146】

高周波数サブ帯域を再構築するべく使用される低周波数信号サンプルが完全な正弦波を表していない際にも調性セクションを包含するべく、2つの更なる段階を提供可能である。

【0147】

第1段階において、可能な品質の劣化を伴う調性信号セクションを検出可能である。調性セクションは、SDF T (S h i f t e d D i s c r e t e F o u r i e r T r a n s f o r m) ドメインにおいて2つの連続したフレーム間の類似性を比較することにより、検出可能である。SDF Tは、位相情報をも含んでいるため、この目的に有用な変換であるが、これは、コーダのその他の部分において使用されているMDC T変換と依然として密接に関連している。

10

【0148】

調性検出は、過渡検出の直後、又は実際の高周波数領域コーディングの初期化の前に実行可能である。過渡フレームは、一般に、調性成分を含んでいないため、調性検出は、現在の及び以前のフレームの両方が通常の長いフレーム（例えば、2048個のサンプル）である際にのみ、適用可能である。

20

【0149】

調性検出は、前述のように、SDF T (S h i f t e d D i s c r e t e F o u r i e r T r a n s f o r m) に基づいており、これは、2N個のサンプルの長いフレームについて、次のように定義可能である。

【0150】

【数39】

$$Y(k) = \sum_{n=0}^{2N-1} h(n)x(n)\exp(i2\pi(n+u)(k+v)/2N) \quad (14)$$

30

【0151】

ここで、 $h(n)$ は、ウィンドウであり、 $x(n)$ は、入力信号であり、 u 及び v は、それぞれ、時間及び周波数ドメインシフトを表している。 $X(k) = \text{real}(Y(k))$ であるため、 $u = (N+1)/2$ 及び $v = 1/2$ となるように、これらのドメインシフトを選択可能である。

【0152】

従って、SDF T及びMDC T変換を別個に演算する代わりに、まず、調性分析のために、SDF T変換を演算可能であり、この結果、SDF T係数の実数部分として、MDC T変換が容易に得られる。従って、調性検出によって演算の複雑性が大幅に増大することにはならない。

40

【0153】

それぞれ、現在の及び以前のフレームのSDF T変換を表している $Y(k)_b$ 及び $Y(k)_{b-1}$ により、フレーム間の類似性を次式を使用して計測可能である。

【0154】

【数 4 0】

$$S = \frac{\sum_{k=N_L+1}^N (|Y_b(k)| - |Y_{b-1}(k)|)^2}{\sum_{k=N_L+1}^N (|Y_b(k)|)^2} \quad (15)$$

【0155】

10

ここで、 $N_L + 1$ は、高周波数コーディング用の限界周波数に対応している。パラメータ S が小さくなるほど、高周波数スペクトルの類似性が増大する。 S の値に基づいて、フレームを次のように分類可能である。

【0156】

【数 4 1】

$$TONALITY = \begin{cases} STRONGLY TONAL, & 0 \leq S < S_{lim1} \\ TONAL, & S_{lim1} \leq S < S_{lim2} \\ NOT TONAL, & S_{lim2} \leq S \end{cases} \quad (16)$$

20

【0157】

制限ファクタ s_{lim1} 及び s_{lim2} の良好な選択肢は、それぞれ、0.02 及び 0.2 である。但し、その他の選択肢も選択可能である。更には、異なる変形を使用可能であり、例えば、これらの分類の中の 1 つを完全に除去することも可能である。

【0158】

図 10 に示されているように、前述の調性検出 (62) は、入力信号 10 に基づいて実行可能である。

30

【0159】

図 10 に示されているように、調性検出 (62) に基づいて、入力フレームを、非調性 (64)、調性 (66)、及び強力な調性 (66) という 3 つのグループに分割可能である。

【0160】

調性検出 (62) の後に、第 2 段階において、前述のように、追加の正弦波を高周波数領域に追加すると共に、恐らくは、高周波数領域の生成に使用される高周波数サブ帯域の数を増大させることにより、調性セクションの品質を改善可能である。

【0161】

最も一般的なケースは、信号が非調性 (64) であるというものであり、この場合には、コーディングは、前述のように継続される。

40

【0162】

入力信号が調性 (66) に分類される場合には、前述のように、コーディングを適用した後に、追加の正弦波を高周波数スペクトルに追加可能である。固定数の正弦波を MDCT ドメインスペクトルに追加可能である。オリジナルの及びコーディング済みのスペクトルの間の絶対差が最大である周波数に対して正弦波を単純に追加可能である。正弦波の位置及び振幅を量子化し、デコーダに提出している。

【0163】

フレームが調性 (又は、強力な調性) であると検出された際には、正弦波をスペクトルの高周波数領域に追加可能である。それぞれ、オリジナルの及びコーディング済みの高周

50

波数サブ帯域成分を表している $X_H(k)$ 及び

【 0 1 6 4 】

【 数 4 2 】

$$\hat{X}_H(k)$$

【 0 1 6 5 】

により、第 1 正弦波をインデックス k_1 に追加可能であり、これは、次式から取得可能である。

10

【 0 1 6 6 】

【 数 4 3 】

$$\max_{k_i} |X_H(k_i) - \hat{X}_H(k_i)| \quad (17)$$

【 0 1 6 7 】

正弦波の（その符号を含む）振幅は、次のように定義可能である。

20

【 0 1 6 8 】

【 数 4 4 】

$$A_i = X_H(k_i) - \hat{X}_H(k_i) \quad (18)$$

【 0 1 6 9 】

最後に、

【 0 1 7 0 】

30

【 数 4 5 】

$$\hat{X}_H(k)$$

【 0 1 7 1 】

を次のように更新可能である。

【 0 1 7 2 】

【 数 4 6 】

40

$$\hat{X}_H(k_i) = \hat{X}_H(k_i) + A_i \quad (19)$$

【 0 1 7 3 】

所望の数の正弦波の追加が完了する時点まで、式 (17) ~ 式 (19) を反復可能である。通常は、4 つの追加の正弦波により、調性セクションにおいて明瞭に改善された結果を得ることが可能である。正弦波の振幅 A_i を量子化し、デコーダ 8 に提出可能である。正弦波の位置 k_i も提出可能である。更には、現在のフレームが調性であることをデコー

50

ダ 8 に通知可能である。

【 0 1 7 4 】

調性セクションにおいては、第 2 スケーリングファクタ α_2 は、品質を改善不可能であり、従って、これを除去可能であることが判明した。

【 0 1 7 5 】

強力な調性セクション (6 8) が検出された際には、現在のセクションの高周波数領域のコーディングが特に困難であることがわかる。従って、正弦波の追加だけでは、不十分であろう。高周波数コーディングの精度を増大させることにより、品質を更に改善可能である。これは、高周波数領域の生成に使用される帯域の数を追加することにより、実行可能である。

10

【 0 1 7 6 】

強力な調性セクションにおいては、高周波数サブ帯域は、フレーム間において非常に類似した状態に留まっている。この類似性をコーディング済みの信号においても維持するべく、特殊な操作を適用可能である。特に、高周波数サブ帯域の数 n_b が相対的に小さい (即ち、8 以下である) 場合には、高周波数サブ帯域の数を更に大きなレートに増大可能である。例えば、16 個の高周波数サブ帯域は、一般に、更に高い精度の性能を提供している。

【 0 1 7 7 】

多数の帯域に加えて、多数の正弦波を追加することも可能である。一般に、良好な解決策は、「通常」の調性セクションにおけるものの 2 倍の数の正弦波を使用するというものである。

20

【 0 1 7 8 】

高周波数サブ帯域の数を増大させると共に、正弦波の数を増大させれば、「通常」のフレームと比べた場合に、強力な調性セクションのビットレートは、容易に倍増することになる。しかしながら、強力な調性セクションは、非常に特殊なケースであって、非常にまれにしか発生せず、従って、平均ビットレートの増大は、非常にわずかである。

【 図面の簡単な説明 】

【 0 1 7 9 】

【 図 1 】 当技術分野におけるオーディオ信号をコーディングするシステムである。

【 図 2 】 当技術分野におけるエンコーダである。

30

【 図 3 】 当技術分野におけるデコーダである。

【 図 4 】 S B R エンコーダである。

【 図 5 】 S B R デコーダである。

【 図 6 】 異なる段階におけるオーディオ信号のスペクトル表現である。

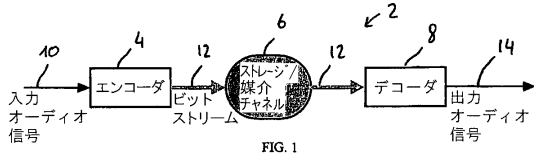
【 図 7 】 第 1 実施例によるシステムである。

【 図 8 】 第 2 実施例によるシステムである。

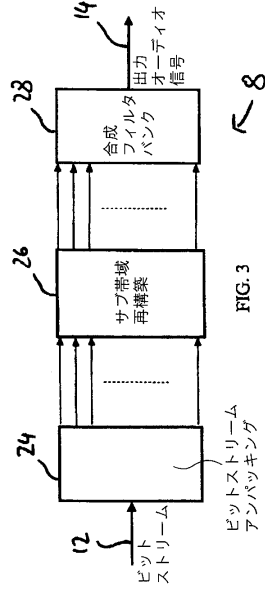
【 図 9 】 エンベロープの正規化を伴う周波数スペクトルである。

【 図 1 0 】 調性検出を使用したコーディングの機能強化である。

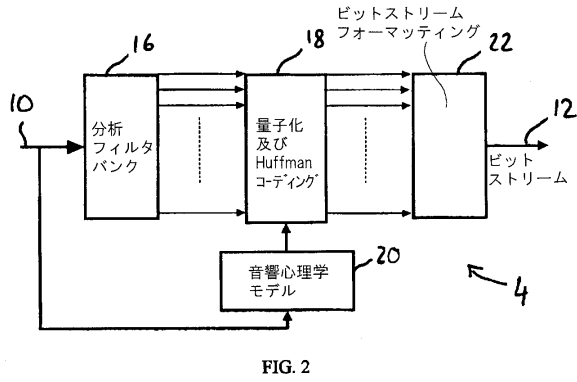
【図1】



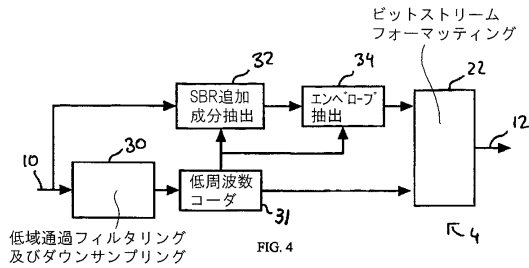
【図3】



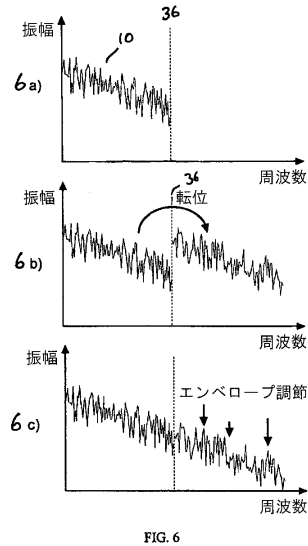
【図2】



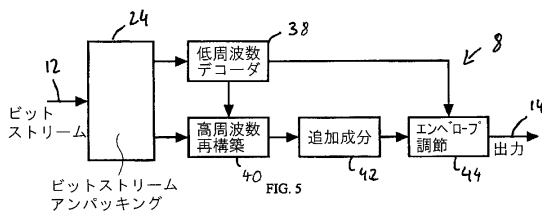
【図4】



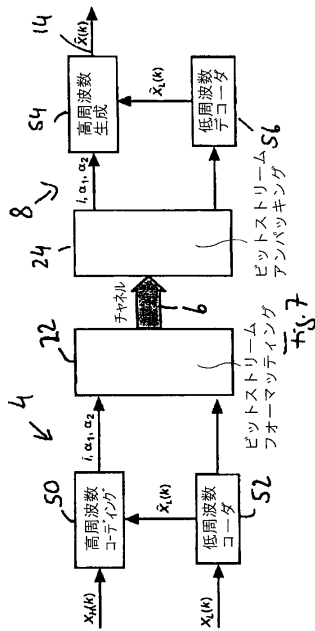
【図6】



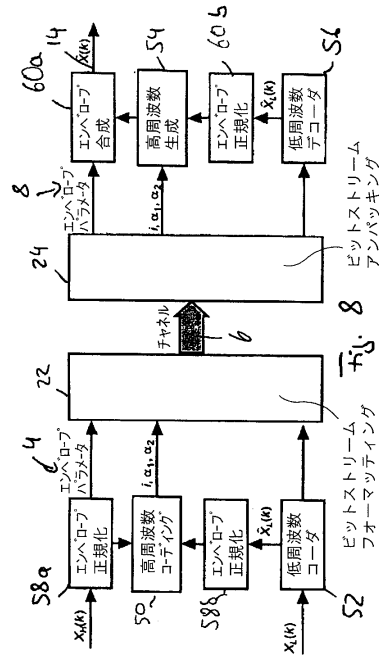
【図5】



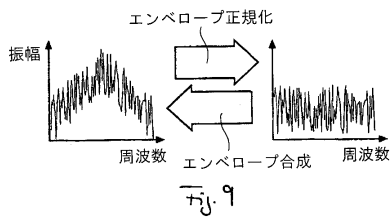
【 図 7 】



【 図 8 】



【 図 9 】



【 図 10 】

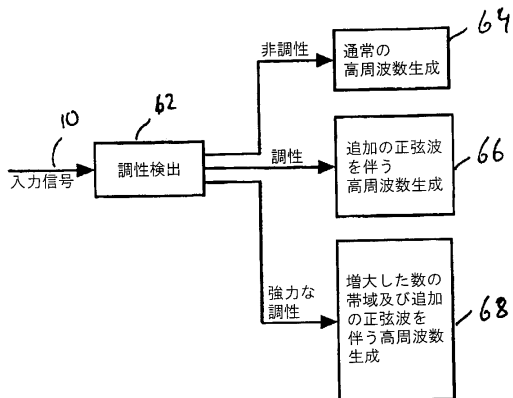


FIG. 10

フロントページの続き

(74)代理人 100122965

弁理士 水谷 好男

(72)発明者 タンミ, ミッコ

フィンランド国, エフイー - 3 3 7 2 0 タンペレ, ケミアンカトゥ 9 エー 5 1

審査官 井上 健一

(56)参考文献 特開2004 - 206129 (JP, A)

特表2001 - 521648 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00-19/14