

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6158421号
(P6158421)

(45) 発行日 平成29年7月5日 (2017.7.5)

(24) 登録日 平成29年6月16日 (2017.6.16)

(51) Int. Cl.	F I
G 1 0 L 19/00 (2013.01)	G 1 0 L 19/00 2 5 0
G 1 0 L 19/035 (2013.01)	G 1 0 L 19/035 D
G 1 0 L 19/038 (2013.01)	G 1 0 L 19/038
G 1 0 L 19/028 (2013.01)	G 1 0 L 19/028
H 0 3 M 7/30 (2006.01)	H 0 3 M 7/30 A

請求項の数 20 (全 65 頁)

(21) 出願番号	特願2016-505843 (P2016-505843)	(73) 特許権者	510185767
(86) (22) 出願日	平成26年4月4日 (2014.4.4)		ドルビー・インターナショナル・アーベー
(65) 公表番号	特表2016-519787 (P2016-519787A)		オランダ王国, セーエン アムステルダム
(43) 公表日	平成28年7月7日 (2016.7.7)		ズイドースト 1101, ヘリケルベル
(86) 国際出願番号	PCT/EP2014/056855		グウェグ 1-35, アポロ ビルディング
(87) 国際公開番号	W02014/161994		グ 3エー
(87) 国際公開日	平成26年10月9日 (2014.10.9)	(74) 代理人	100107766
審査請求日	平成27年10月27日 (2015.10.27)		弁理士 伊東 忠重
(31) 優先権主張番号	61/808,673	(74) 代理人	100070150
(32) 優先日	平成25年4月5日 (2013.4.5)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
(31) 優先権主張番号	61/875,817		弁理士 大貫 進介
(32) 優先日	平成25年9月10日 (2013.9.10)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 先進量子化器

(57) 【特許請求の範囲】

【請求項 1】

係数のブロックの第一の係数を量子化するよう構成された量子化ユニットであって、係数の前記ブロックは、複数の対応する周波数ビンについての複数の係数を含み、当該量子化ユニットは、

・一組の量子化器を提供するよう構成されており、前記一組の量子化器は、それぞれSNRと称される異なる信号対雑音比に関連付けられた限られた数の異なる量子化器を含み、前記一組の量子化器の前記異なる量子化器は、そのSNRに従って順序付けられており、前記一組の量子化器は、

・前記第一の係数の値を所定の統計モデルに従って生成されたランダムな値で置き換えることによって、前記第一の係数を量子化するよう構成されているノイズ充填量子化器；

・一つまたは複数のディザリングされる量子化器；および
・一つまたは複数のディザリングされない決定論的な量子化器を含み、
当該量子化ユニットはさらに、

・前記第一の係数に帰されるSNRを示すSNR指示を決定し；
・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択し；
・前記第一の量子化器を使って前記第一の係数を量子化するよう構成されている、
量子化ユニット。

【請求項 2】

・前記ノイズ充填量子化器は、前記異なるSNRのうち相対的に最低のSNRに関連付けられており、

・前記一つまたは複数のディザリングされない決定論的な量子化器は、前記異なるSNRのうち一つまたは複数の相対的に最高のSNRに関連付けられており、

・前記一つまたは複数のディザリングされる量子化器は、前記異なるSNRのうち、前記相対的に最低のSNRより高く、前記一つまたは複数の相対的に最高のSNRより低い、一つまたは複数の中間的なSNRに関連付けられている、

請求項1記載の量子化ユニット。

【請求項3】

前記一組の量子化器は、前記異なる量子化器に関連付けられたSNRの昇順に従って順序付けられている、請求項1記載の量子化ユニット。

10

【請求項4】

・順序付けられた前記一組の量子化器における順序付けにおいて隣接する一対の量子化器に関連付けられたSNRの差によってSNR差が与えられ、

・前記異なる量子化器からの隣接する量子化器のすべての対についてのSNR差が、所定のSNRターゲット差を中心とする所定のSNR差区間内にはいる、

請求項3記載の量子化ユニット。

【請求項5】

前記ノイズ充填量子化器は、

・前記所定の統計モデルに従って乱数を生成するよう構成された乱数発生器を有する；および／または

・本質的には0dBより低いまたは0dBに等しいSNRに関連付けられている、

請求項1記載の量子化ユニット。

20

【請求項6】

前記一つまたは複数のディザリングされる量子化器のうちの特定のディザリングされる量子化器は、

・前記第一の係数にディザ値を適用することによって第一のディザリングされた係数を決定するよう構成されたディザ適用ユニットと；

・前記第一のディザリングされた係数をスカラー量子化器のある区間に割り当てることによって第一の量子化インデックスを決定するよう構成されたスカラー量子化器とを有する

30

請求項1記載の量子化ユニット。

【請求項7】

前記一つまたは複数のディザリングされる量子化器のうちの前記特定のディザリングされる量子化器はさらに、

・前記第一の量子化インデックスに第一の再構成値を割り当てるよう構成された逆スカラー量子化器と；

・前記第一の再構成値から前記ディザ値を除去することによって、第一のディザリング解除された係数を決定するよう構成されたディザ除去ユニットとを有する、

請求項6記載の量子化ユニット。

40

【請求項8】

・前記スカラー量子化器が、所定の量子化器きざみサイズ Δ を有し；

・前記ディザ値は、所定のディザ区間からの値を取り；

・前記所定のディザ区間は、前記所定の量子化器きざみサイズ Δ 以下の幅を有する、

請求項6記載の量子化ユニット。

【請求項9】

・係数の前記ブロックは、スペクトル・ブロック包絡に関連付けられており；

・前記スペクトル・ブロック包絡は前記複数の周波数ビンについて複数のスペクトル・エネルギー値を示し；

・前記SNR指示が前記スペクトル・ブロック包絡に依存する、

50

請求項 1 記載の量子化ユニット。

【請求項 1 0】

- ・係数の前記ブロックの前記複数の係数は、複数の周波数帯域に割り当てられ；
- ・周波数帯域は、一つまたは複数の周波数ビンを含み；
- ・当該量子化ユニットは、同じ周波数帯域に割り当てられる係数が同じ量子化器を使って量子化されるよう、前記複数の周波数帯域のそれぞれについて前記一組の量子化器から量子化器を選択するよう構成されている、

請求項 1 記載の量子化ユニット。

【請求項 1 1】

当該量子化ユニットは、

- ・係数の前記ブロックの属性を示すサイド情報を決定し；
- ・前記サイド情報に依存して量子化器の前記組を生成するよう構成されている、

請求項 1 記載の量子化ユニット。

【請求項 1 2】

量子化インデックスを量子化解除するよう構成された逆量子化ユニットであって、前記量子化インデックスは、複数の対応する周波数ビンについて複数の係数を含む係数のブロックに関連付けられており、当該逆量子化ユニットは、

- ・一組の量子化器を提供するよう構成されており、前記一組の量子化器は、それぞれ SNR と称される異なる信号対雑音比に関連付けられた、限られた数の異なる量子化器を含み、前記一組の量子化器の前記異なる量子化器は、その SNR に従って順序付けられており、前記一組の量子化器は、

・係数の値を所定の統計モデルに従って生成されたランダムな値で置き換えることによって、該係数を量子化するよう構成されているノイズ充填量子化器；

・一つまたは複数のディザリングされる量子化器；および

・一つまたは複数のディザリングされない決定論的な量子化器を含み、

当該逆量子化ユニットはさらに、

- ・係数の前記ブロックからの第一の係数に帰される SNR を示す SNR 指示を決定し；
- ・前記 SNR 指示に基づいて前記一組の量子化器から第一の量子化器を選択し；
- ・前記第一の量子化器を使って前記第一の係数について第一の量子化された係数を決定するよう構成されている、

逆量子化ユニット。

【請求項 1 3】

発話信号をビットストリームにエンコードするよう構成された変換ベースの発話エンコーダであって、

・変換係数の複数の逐次的なブロックを受領するよう構成されたフレーミング・ユニットであって、前記複数の逐次的なブロックは、現在ブロックおよび一つまたは複数の以前のブロックを含み、前記複数の逐次的なブロックは、発話信号のサンプルを示す、フレーミング・ユニットと；

・対応する現在ブロック包絡を使って変換係数の対応する現在ブロックを平坦化することによって、平坦化された変換係数の現在ブロックを決定するよう構成された平坦化ユニットと；

・再構成された変換係数の一つまたは複数の以前のブロックに基づいて、かつ一つまたは複数の予測器パラメータに基づいて、推定された平坦化された変換係数の現在ブロックを決定するよう構成された予測器であって、再構成された変換係数の前記一つまたは複数の以前のブロックは、変換係数の前記一つまたは複数の以前のブロックから導出されたものである、予測器と；

・平坦化された変換係数の現在ブロックに基づいて、かつ推定された平坦化された変換係数の現在ブロックに基づいて、予測誤差係数の現在ブロックを決定するよう構成された差分ユニットと；

- ・予測誤差係数の現在ブロックから導出された係数を量子化するよう構成された、請求項

10

20

30

40

50

1 記載の量子化ユニットとを有しており、前記ビットストリームについての係数データは量子化された係数に関連付けられた量子化インデックスに基づいて決定される、変換ベースの発話エンコーダ。

【請求項 1 4】

・変換係数のブロックがMDCT係数を含む；および／または
・変換係数のブロックが、256個の周波数ビン内の256個の変換係数を含む、
請求項 1 3 記載の変換ベースの発話エンコーダ。

【請求項 1 5】

再スケーリングされた誤差係数の現在ブロックの再スケーリングされた誤差係数の分散が、平均では、予測誤差係数の現在ブロックの予測誤差係数の分散より高くなるよう、一つまたは複数のスケーリング規則を使って予測誤差係数の現在ブロックに基づいて、再スケーリングされた誤差係数の現在ブロックを決定するよう構成されたスケーリング・ユニットをさらに有し、

・予測誤差係数の現在ブロックは、対応する複数の周波数ビンについての複数の予測誤差係数を含み、

・前記一つまたは複数のスケーリング規則に従って前記スケーリング・ユニットによって前記予測誤差係数に適用されるスケーリング利得は、それぞれの予測誤差係数の周波数ビンに依存する、

請求項 1 3 記載の変換ベースの発話エンコーダ。

【請求項 1 6】

・前記予測器は、重み付けされた平均平方誤差基準を使って、推定された平坦化された変換係数の現在ブロックを決定するよう構成されており、

・前記重み付けされた平均平方誤差基準は、現在ブロック包絡を重みとして考慮に入れる、

請求項 1 3 記載の変換ベースの発話エンコーダ。

【請求項 1 7】

・当該変換ベースの発話エンコーダが現在ブロック包絡に基づいて割り当てベクトルを決定するよう構成されたビット割り当てユニットをさらに有しており、

・前記割り当てベクトルは、予測誤差係数の現在ブロックから導出された少なくとも一つの係数を量子化するために使われる前記一組の所定の量子化器からの量子化器を示す、

請求項 1 5 記載の変換ベースの発話エンコーダ。

【請求項 1 8】

ビットストリームをデコードして再構成された発話信号を提供するよう構成された変換ベースの発話デコーダであって、

・再構成された変換係数の一つまたは複数の以前のブロックに基づいて、かつ前記ビットストリームから導出される一つまたは複数の予測器パラメータに基づいて、推定された平坦化された変換係数の現在ブロックを決定するよう構成された予測器と；

・一組の所定の量子化器を使って、前記ビットストリーム内に含まれる係数データに基づいて、量子化された予測誤差係数の現在ブロックを決定するよう構成された、請求項 1 2 記載の逆量子化ユニットと；

・推定された平坦化された変換係数の現在ブロックに基づき、かつ量子化された予測器誤差係数の現在ブロックに基づいて、再構成された平坦化された変換係数の現在ブロックを決定するよう構成された加算ユニットと；

・現在ブロック包絡を使って、再構成された平坦化された変換係数の現在ブロックにスペクトル形状を与えることによって、再構成された変換係数の現在ブロックを決定するよう構成された逆平坦化ユニットとを有しており、

再構成された発話信号は、再構成された変換係数の現在ブロックに基づいて決定される、変換ベースの発話デコーダ。

【請求項 1 9】

係数のブロックの第一の係数を量子化する方法であって、係数の前記ブロックは、複数

10

20

30

40

50

の対応する周波数ビンについての複数の係数を含み、当該方法は、

・一組の量子化器を提供する段階であって、前記一組の量子化器は、それぞれSNRと称される複数の異なる信号対雑音比に関連付けられた複数の異なる量子化器を含み、前記複数の異なる量子化器は、

・前記第一の係数の値を所定の統計モデルに従って生成されたランダムな値で置き換えることによって、前記第一の係数を量子化するよう構成されているノイズ充填量子化器、

・一つまたは複数のディザリングされる量子化器、および

・一つまたは複数のディザリングされない決定論的な量子化器を含む、段階と；

・前記第一の係数に帰されるSNRを示すSNR指示を決定する段階と；

・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択する段階と；

・前記第一の量子化器を使って前記第一の係数を量子化する段階とを含む、

方法。

【請求項 20】

量子化インデックスを量子化解除する方法であって、前記量子化インデックスは、複数の対応する周波数ビンについて複数の係数を含む係数のブロックに関連付けられており、当該方法は、

・一組の量子化器を提供する段階であって、前記一組の量子化器は、それぞれSNRと称される複数の異なる信号対雑音比に関連付けられた複数の異なる量子化器を含み、前記複数の異なる量子化器は、

・係数の値を所定の統計モデルに従って生成されたランダムな値で置き換えることによって、該係数を量子化するよう構成されているノイズ充填量子化器、

・一つまたは複数のディザリングされる量子化器、および

・一つまたは複数のディザリングされない決定論的な量子化器を含む、段階と；

・係数の前記ブロックからの第一の係数に帰されるSNRを示すSNR指示を決定する段階と；

・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択する段階と；

・前記第一の量子化器を使って、前記第一の係数についての第一の量子化された係数を決定する段階とを含む、

方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

本願は、2013年4月5日に出願された米国仮特許出願第61/808,673号および2013年9月10日に出願された米国仮特許出願第61/875,817号の優先権を主張するものである。各出願の内容はここに参照によりその全体において組み込まれる。

【0002】

技術分野

本稿はオーディオ・エンコードおよびデコード・システム（オーディオ・コーデック・システムと称される）に関する。特に、本稿は、音声エンコード／デコードに特に好適な変換ベースのオーディオ・コーデック・システムに関する。

【背景技術】

【0003】

汎用の知覚的オーディオ符号化器は、数十ミリ秒（たとえば20ms）をカバーするサンプルのブロック・サイズをもつ修正離散コサイン変換（MDCT）のような変換を使うことによって、比較的高い符号化利得を達成する。そのような変換ベースのオーディオ・コーデック・システムの例は先進オーディオ符号化（AAC）または高効率（HE）-AACである。しかしながら、そのような変換ベースのオーディオ・コーデック・システムを声信号に使うときは、声信号の品質は、より低ビットレートに向かって、音楽信号の品質より速く劣化する。特に、ドライな（非残響性の）発話信号の場合にそうである。

【0004】

本稿は、発話信号の符号化に特に好適な変換ベースのオーディオ・コーデック・システムを記述する。さらに、本稿は、そのような変換ベースのオーディオ・コーデック・システムにおいて使用される量子化方式を記述する。さまざまな異なる量子化方式が変換ベースのオーディオ・コーデック・システムと関連して使われてもよい。例はベクトル量子化（たとえばツイン・ベクトル量子化）、分布保存量子化、ディザリングされた量子化、ランダム・オフセットをもつスカラー量子化およびノイズ充填と組み合わせられたスカラー量子化である（たとえばUS7447631に記述される量子化器）。これら種々の量子化方式は、下記の属性の一つまたは複数に関し、さまざまな利点および欠点をもつ：

- ・ 演算上の（エンコーダ）複雑さ。これは典型的には量子化およびビットストリーム生成（たとえば、可変長符号化）の計算量を含む。
- ・ 知覚的なパフォーマンス。これは理論的な考察に基づいて（レート - 歪みパフォーマンス）および関連するノイズ充填挙動に基づいて（たとえば、発話の低レートの変換符号化に実際上関連するビットレートで）推定されてもよい。
- ・ 全体的なビットレート制約条件（たとえば最大ビット数）があるときのビットレート割り当てプロセスの複雑さ；および／または
- ・ 種々のデータ・レートおよび種々の歪みレベルを可能にすることに関する柔軟性。

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】L.Schuchman、"Dither signals and their effect on quantization noise"、IEEE TCOM, pp.162-165, Dec. 1964

【発明の概要】

【発明が解決しようとする課題】

【0006】

本稿では、量子化方式は、上述した属性の少なくともいくつかに対処する。特に、上述した属性の一部または全部に関して改善されたパフォーマンスを提供する量子化方式が記述される。

【課題を解決するための手段】

【0007】

ある側面によれば、係数のブロックの第一の係数を量子化するよう構成された量子化ユニット（本稿では係数量子化ユニットとも称される）が記述される。係数のブロックは、予測残差係数のブロック（予測誤差係数のブロックとも称される）に対応していてもよいし、それから導出されてもよい。よって、量子化ユニットは、下記でより詳細に述べるサブバンド予測を利用する変換ベースのオーディオ・エンコーダの一部であってもよい。一般的な言い方では、係数のブロックは複数の対応する周波数ビンについての複数の係数を含んでいてもよい。係数のブロックは、変換係数のブロックから導出されてもよい。ここで、変換係数のブロックは、時間領域から周波数領域への変換（たとえば修正離散コサイン変換、MDCT）を使ってオーディオ信号（たとえば発話信号）を時間領域から周波数領域に変換することによって決定されたものである。

【0008】

係数のブロックの前記第一の係数は、係数のブロックの前記係数の任意の一つまたは複数に対応してもよいことを注意しておくべきである。係数のブロックはK個の係数を含んでいてもよい（ $K > 1$ 、たとえば $K = 256$ ）。前記第一の係数は、 $k = 1, \dots, K$ 周波数係数の任意のものに対応してもよい。以下で概説されるように、前記複数K個の周波数ビンは複数L個の周波数帯域にグループ化されてもよい。ここで、 $1 < L < K$ である。係数の前記ブロックの係数は、前記複数の周波数帯域（ $l = 1, \dots, L$ ）の一つに割り当てられてもよい。 $q = 1, \dots, Q$ および $0 < Q < K$ として、特定の周波数帯域lに割り当てられる諸係数qは、同じ量子化器を使って量子化されてもよい。前記第一の係数は、任意の $q = 1, \dots, Q$ および任意の $l = 1, \dots, L$ について、l番目の周波数帯域のq番目の係数に対応していてもよい。

【0009】

量子化ユニットは、一組の量子化器を提供するよう構成されていてもよい。前記一組の量子化器は、それぞれ複数の異なる信号対雑音比（SNR）または複数の異なる歪みレベルに対応する複数の異なる量子化器を含んでいてもよい。よって、前記一組の量子化器の異なる量子化器は、それぞれのSNRまたは歪みレベルを与えてもよい。前記一組の量子化器内の量子化器は、前記複数の量子化器に関連付けられた前記複数のSNRに従って順序付けられてもよい。特に、前記量子化器は、ある特定の量子化器を使って得られるSNRが、直前の隣接する量子化器を使って得られるSNRより増大するように順序付けられてもよい。

【0010】

前記一組の量子化器は、一組の受け入れ可能な量子化器とも称されてもよい。典型的には、前記一組の量子化器のうちに含まれる量子化器の数は量子化器の数Rに制限される。前記一組の量子化器のうちに含まれる量子化器の数は量子化器の数Rは、前記一組の量子化器によってカバーされるべき全体的なSNR範囲（たとえば約0dBから30dBのSNR範囲）に基づいて選択されてもよい。さらに、量子化器の数Rは典型的には、順序付けられた一組の量子化器内の隣接する量子化器の間のSNRターゲット差に依存する。量子化器の数Rについての典型的な値は10ないし20個の量子化器である。

【0011】

前記複数の異なる量子化器は、ノイズ充填量子化器、一つまたは複数のディザリングされる量子化器および／または一つまたは複数のディザリングされない量子化器を含んでいてもよい。ある好ましい例では、前記複数の異なる量子化器は、単一のノイズ充填量子化器、一つまたは複数のディザリングされる量子化器および／または一つまたは複数のディザリングされない量子化器を含んでいてもよい。本稿で概説されるように、（たとえば大きな量子化きざみサイズをもつディザリングされる量子化器の代わりに）零ビットレート状況についてはノイズ充填量子化器を使うことが有益である。ノイズ充填量子化器は、前記複数のSNRの相対的に最低のSNRに関連付けられており、前記一つまたは複数のディザリングされない量子化器は、前記複数のSNRの一つまたは複数の相対的に最高のSNRと関連付けられてもよい。前記一つまたは複数のディザリングされる量子化器は、前記相対的に最低のSNRより高く、前記複数のSNRのうち前記一つまたは複数の相対的に最高のSNRより低い、一つまたは複数の中間的なSNRに関連付けられてもよい。よって、順序付けられた一組の量子化器は、最低のSNR（たとえば0dB以下）についてのノイズ充填量子化器と、それに続いて中間的なSNRについての一つまたは複数のディザリングされる量子化器と、それに続く、相対的に高いSNRについての一つまたは複数のディザリングされない量子化器とを含んでいてもよい。そうすることにより、（前記一組の量子化器を使って量子化された、量子化係数のブロックから導出される）再構成されるオーディオ信号の知覚的品質が改善されうる。特に、スペクトルの穴によって引き超される可聴アーチファクトが軽減されうる。一方、同時に、量子化ユニットのMSE（平均平方誤差）パフォーマンスは高く保たれる。

【0012】

ノイズ充填量子化器は、所定の統計モデルに従って乱数を生成するよう構成された乱数発生器を有していてもよい。ノイズ充填量子化器の乱数発生器の前記所定の統計モデルは、エンコーダおよび対応するデコーダにおいて利用可能なサイド情報（たとえば、分散保存フラグ（variance preservation flag））に依存してもよい。ノイズ充填量子化器は、前記第一の係数を、乱数発生器によって生成された乱数で置き換えることによって前記第一の係数（または係数の前記ブロックの係数の任意のもの）を量子化するよう構成されていてもよい。量子化ユニットにおいて（たとえばエンコーダ内に含まれるローカル・デコーダにおいて）使用される乱数発生器は、（対応するデコーダにおける）逆量子化ユニットにおける対応する乱数発生器と同期していてもよい。よって、ノイズ充填量子化器の出力は前記第一の係数とは独立であってもよく、よってノイズ充填量子化器の出力はいかなる量子化インデックスの伝送も要求しなくてもよい。ノイズ充填量子化器は、（ほぼまたは実質的に）0dBであるSNRに関連付けられてもよい。換言すれば、ノイズ充填量子化器

は、0dBに近いSNRとともに動作してもよい。レート割り当てプロセスの間、ノイズ充填量子化器は、0dB SNRを提供すると考えられてもよい。ただし、実際上は、そのSNRは0からやや外れてもよい（たとえば、0dBよりやや低くてもよい（入力信号と独立な信号の合成のため））。

【0013】

ノイズ充填量子化器のSNRは、一つまたは複数の追加的パラメータに基づいて調整されてもよい。たとえば、ノイズ充填量子化器の分散は、予測器利得のあらかじめ決定された関数に従って、合成された信号の分散（すなわち、ノイズ充填量子化器を使って量子化された係数の分散）を設定することによって、調整されてもよい。代替的または追加的に、合成された信号の分散は、ビットストリームにおいて伝送されるフラグによって設定されてもよい。特に、ノイズ充填量子化器の分散は、（本稿で後述する）予測器利得 ρ の二つのあらかじめ定義された関数のうちの一つによって調整されてもよい。ここで、これらの関数の一方は、フラグに依存して（たとえば、分散保存フラグに依存して）、合成された信号をレンダリングするために選択されうる。例として、ノイズ充填量子化器によって生成された信号の分散は、ノイズ充填量子化器のSNRが $[-3.0\text{dB}$ ないし $0\text{dB}]$ の範囲内にはいるような仕方では調整されてもよい。0dBにおけるSNRは、典型的には、MMSE（最小平均平方誤差）の観点から有益である。他方、より低いSNR（たとえば -3.0dB まで）を使うときは、知覚的品質は、増大されてもよい。

【0014】

前記一つまたは複数のディザリングされる量子化器は好ましくは、減算的なディザリングされる量子化器である。特に、前記一つまたは複数のディザリングされる量子化器のあるディザリングされる量子化器は、前記第一の係数にディザ値（ディザ数とも称される）を適用することによって第一のディザリングされた係数を決定するよう構成されたディザ適用ユニットを有していてもよい。さらに、前記ディザリングされる量子化器は、前記第一のディザリングされた係数をスカラー量子化器のある区間に割り当てることによって第一の量子化インデックスを決定するよう構成されたスカラー量子化器を有していてもよい。よって、前記ディザリングされる量子化器は、前記第一の係数に基づいて第一の量子化インデックスを生成してもよい。同様の仕方では、係数の前記ブロックの一つまたは複数の他の係数が量子化されうる。

【0015】

前記一つまたは複数のディザリングされる量子化器のあるディザリングされる量子化器はさらに、前記第一の量子化インデックスに第一の再構成値を割り当てるよう構成された逆スカラー量子化器を有していてもよい。さらに、前記ディザリングされる量子化器は、前記第一の再構成された値から前記ディザ値（すなわち前記ディザ適用ユニットによって適用されたのと同じディザ値）を除去することによって、第一のディザリング解除された係数を決定するよう構成されたディザ除去ユニットを有していてもよい。

【0016】

さらに、前記ディザリングされる量子化器は、前記第一のディザリング解除された係数に量子化器事後利得 γ を適用することによって第一の量子化された係数を決定するよう構成された事後利得適用ユニットを有していてもよい。事後利得 γ を前記第一のディザリング解除された係数に適用することにより、ディザリングされる量子化器のMSEパフォーマンスが改善されうる。量子化器事後利得 γ は次式によって与えられてもよい。

【0017】

【数1】

$$\gamma = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}$$

ここで、 $\sigma_x^2 = E\{X^2\}$ は係数の前記ブロックの一つまたは複数の係数の分散であり、 Δ は

前記ディザリングされる量子化器の前記スカラー量子化器の量子化器きざみサイズである。

【0018】

よって、前記ディザリングされる量子化器は、量子化された係数を与えるよう逆量子化を実行するよう構成されていてもよい。これは、閉ループ予測を容易にする、エンコーダの前記ローカルなデコーダにおいて使われてもよい。たとえば、エンコーダにおける予測ループはデコーダにおける予測ループと同期を保たれる。

【0019】

前記ディザ適用ユニットは、前記第一の係数から前記ディザ値を減算するよう構成されていてもよく、前記ディザ除去ユニットは前記第一の再構成値に前記ディザ値を加算するよう構成されていてもよい。あるいはまた、前記ディザ適用ユニットは前記第一の係数に前記ディザ値を加算するよう構成されていてもよく、前記ディザ除去ユニットは前記第一の再構成値から前記ディザ値を減算するよう構成されていてもよい。

10

【0020】

前記量子化ユニットはさらに、ディザ値のブロックを生成するよう構成されたディザ生成器を有していてもよい。前記エンコーダと前記デコーダの間の同期を容易にするために、前記ディザ値は擬似乱数であってもよい。ディザ値の前記ブロックは、それぞれ前記複数の周波数ビンについての複数のディザ値を含んでいてもよい。よって、前記ディザ生成器は、ある特定の係数が前記ディザリングされる量子化器の一つを使って量子化されるべきか否かに関わりなく、量子化されるべき係数の前記ブロックの各係数についてディザ値を生成するよう構成されていてもよい。これは、エンコーダにおいて使用されるディザ生成器と対応するデコーダにおいて使用されるディザ生成器の間の同期を維持するために有益である。

20

【0021】

前記ディザリングされる量子化器の前記スカラー量子化器は、所定の量子化器きざみサイズ Δ を有する。よって、前記ディザリングされる量子化器の前記スカラー量子化器は、一様量子化器であってもよい。前記ディザ値は、所定のディザ区間からの値を取ってもよい。前記所定のディザ区間は、前記所定の量子化器きざみサイズ Δ 以下の幅を有していてもよい。さらに、ディザ値の前記ブロックは、前記所定のディザ区間内に一様に分布したランダム変数の実現 (realizations) から構成されてもよい。たとえば、前記ディザ生成器は、規格化されたディザ区間 (たとえば $[0,1)$ または $[-0.5,0.5)$) から引き出されるディザ値のブロックを生成するよう構成される。よって、規格化されたディザ区間の幅は1であってもよい。ディザ値の前記ブロックは、次いで、その特定のディザリングされる量子化器の前記所定の量子化きざみサイズ Δ を乗算されてもよい。そうすることにより、きざみサイズ Δ をもつ量子化器と一緒に使うのに好適なディザ実現が選られうる。特に、そうすることにより、いわゆるSchuchman条件を満たす量子化器が得られる (非特許文献1)。

30

【0022】

ディザ生成器は、M個の所定のディザ実現のうち一つを選択するよう構成されていてもよい。ここで、Mは1より大きい整数である。さらに、ディザ生成器は、選択されたディザ実現に基づいてディザ値のブロックを生成するよう構成されていてもよい。特に、いくつかの実装では、ディザ実現の数は制限されてもよい。例として、所定のディザ実現の数Mは10、5、4またはそれより少なくてもよい。これは、前記一つまたは複数のディザリングされる量子化器を使って得られた量子化インデックスのその後のエントロピー符号化に関して有益でありうる。特に、ディザ実現の制限された数Mの使用は、量子化インデックスのためのエントロピー符号化器が、ディザ実現の限られた数に基づいてトレーニングされることを可能にする。そうすることにより、算術符号の代わりに瞬時符号 (たとえば多次元ハフマン符号化のような) を使うことができ、これは演算複雑さの点で有利でありうる。

40

【0023】

前記一つまたは複数のディザリングされない量子化器のうちのあるディザリングされな

50

い量子化器は、所定の一様な量子化きざみサイズをもつスカラー量子化器であってもよい。よって、前記一つまたは複数のディザリングされない量子化器は、（擬似）ランダム・ディザを使わない決定論的な量子化器であってもよい。

【0024】

上記で概説したように、前記一組の量子化器は順序付けられてもよい。これは、効率的なビット割り当てプロセスに鑑み有益でありうる。特に、前記一組の量子化器の順序付けは、整数インデックスに基づく前記一組の量子化器からの量子化器の選択を可能にする。前記一組の量子化器は、隣接する量子化器の間のSNRの増大が少なくとも近似的に一定であるよう、順序付けられてもよい。換言すれば、二つの量子化器の間のSNR差が、順序付けられた一組の量子化器からの一対の隣接する量子化器に関連付けられたSNRの差によって与えられてもよい。前記複数の順序付けられた量子化器からの隣接する量子化器のすべての対についてのSNR差が、所定のSNRターゲット差を中心とする所定のSNR差区間内にはいてもよい。所定のSNR差区間の幅は、所定のSNRターゲット差の10%または5%より小さくてもよい。SNRターゲット差は、量子化器の比較的小さな集合が比較的大きな全体的なSNR範囲における動作を与えることができるような仕方 で設定されてもよい。たとえば、典型的な用途では、前記一組の量子化器は、0dBのSNRから30dBのSNRに向けて区間内の動作を容易にしてもよい。所定のSNRターゲット差は1.5dBまたは3dBに設定され、それにより10ないし20個の量子化器の集合を用いて30dBの全体的なSNR範囲がカバーできるようにしてもよい。よって、順序付けられた一組の量子化器の量子化器の整数インデックスの増大は対応するSNRの増大に翻訳される。この一対一の関係は、特定のSNRをもつ量子化器を所与のビットレート制約条件に従って特定の周波数帯域に割り当てる、効率的なビット割り当てプロセスの実装のために有益である。

【0025】

量子化ユニットは、前記第一の係数に帰されるSNRを示すSNR指示を決定するよう構成されていてもよい。前記第一の係数に帰されるSNRは、レート割り当てプロセス（ビット割り当てプロセスとも称される）を使って決定されてもよい。上記のように、前記第一の係数に帰されるSNRは、前記一組の量子化器から量子化器を直接同定してもよい。よって、量子化ユニットは、前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択するよう構成されていてもよい。さらに、量子化ユニットは、前記第一の量子化器を使って前記第一の係数を量子化するよう構成されていてもよい。特に、量子化ユニットは、前記第一の係数についての第一の量子化インデックスを決定するよう構成されていてもよい。第一の量子化インデックスは、エントロピー符号化されてもよく、ビットストリーム内で係数データとして対応する逆量子化ユニット（または対応するデコーダ）に伝送されてもよい。さらに、量子化ユニットは、第一の係数から第一の量子化された係数を決定するよう構成されていてもよい。第一の量子化された係数は、エンコーダの予測器内で使われてもよい。

【0026】

係数のブロックは、スペクトル・ブロック包絡〔エンベロープ〕（たとえば、後述するように、現在の包絡または量子化された現在の包絡）と関連付けられてもよい。特に、係数のブロックは、スペクトル・ブロック包絡を使って（入力オーディオ信号のあるセグメントから導出される）変換係数のブロックを平坦化することによって得られてもよい。スペクトル・ブロック包絡は、前記複数の周波数ビンについて複数のスペクトル・エネルギー値を示してもよい。特に、スペクトル・ブロック包絡は、係数の前記ブロックの係数の相対的な重要性を示してもよい。よって、スペクトル・ブロック包絡（または後述する割り当て包絡のようなスペクトル・ブロック包絡から導出される包絡）は、レート割り当て目的のために使われてもよい。特に、SNR指示は、スペクトル・ブロック包絡に依存してもよい。SNR指示はさらに、スペクトル・ブロック包絡をオフセットさせるためのオフセット・パラメータに依存してもよい。レート割り当てプロセスの間、オフセット・パラメータは、係数の量子化され、エンコードされたブロックから生成される係数データが所定のビットレート制約条件を満たすまで、増減されてもよい（オフセット・パラメータは、

係数のエンコードされたブロックが所定のビット数を超えないように可能な限り大きいよう選択されてもよい)。よって、オフセット・パラメータは、係数のブロックをエンコードするために利用可能な所定のビット数に依存してもよい。

【0027】

前記第一の係数に帰されるSNRを示すSNR指示は、オフセット・パラメータを使って前記第一の係数の周波数ビンに関連付けられたスペクトル・ブロック包絡から導出される値をオフセットさせることによって決定されてもよい。特に、本稿に記載されるビット割り当て公式は、SNR指示を決定するために使用されてもよい。ビット割り当て公式は、スペクトル・ブロック包絡から導出される割り当て包絡およびオフセット・パラメータの関数であってよい。

10

【0028】

よって、SNR指示は、スペクトル・ブロック包絡から導出される割り当て包絡に依存してもよい。割り当て包絡は、割り当て分解能（たとえば、3dBの分解能）を有していてもよい。割り当て分解能は好ましくは、前記一組の量子化器からの隣接する量子化器の間のSNR差に依存してもよい。特に、割り当て分解能およびSNR差は互いに対応していてもよい。対応する割り当て分解能およびSNR差を選択することにより（たとえば、dB領域でSNR差の二倍である割り当て分解能を選択することにより）、ビット割り当てプロセスおよび／または量子化器選択プロセスは（たとえば本稿に記載されるビット割り当て公式を使って）簡略化されてもよい。

【0029】

20

係数の前記ブロックの前記複数の係数は、複数の周波数帯域に割り当てられてもよい。周波数帯域は、一つまたは複数の周波数ビンを含んでいてもよい。よって、前記複数の係数のうちの二つ以上が同じ周波数帯域に割り当てられてもよい。典型的には、周波数帯域当たりの周波数ビンの数は、周波数が増すとともに増大する。特に、周波数帯域構造（たとえば、周波数帯域当たりの周波数ビンの数）は、音響心理学的考察に従ってもよい。量子化ユニットは、同じ周波数帯域に割り当てられる係数が同じ量子化器を使って量子化されるよう、前記複数の周波数帯域のそれぞれについて前記一組の量子化器から量子化器を選択するよう構成されていてもよい。特定の周波数帯域を量子化するために使われる量子化器は、その特定の周波数帯域内のスペクトル・ブロック包絡の前記一つまたは複数のスペクトル・エネルギー値に基づいて決定されてもよい。量子化目的のための周波数帯域構造の使用は、量子化方式の音響心理学的パフォーマンスに関して有益でありうる。

30

【0030】

量子化ユニットは、係数のブロックの属性を示すサイド情報を受領するよう構成されていてもよい。例として、サイド情報が、当該量子化ユニットを有するエンコーダ内に含まれる予測器によって決定された予測器利得を含んでいてもよい。予測器利得は、係数のブロックのトーン性（tonal）内容を示してもよい。代替的または追加的に、サイド情報は、係数のブロックに基づいておよび／またはスペクトル・ブロック包絡に基づいて導出されたスペクトル反射係数を含んでいてもよい。スペクトル反射係数は、係数のブロックの摩擦性（fricative）内容を示していてもよい。量子化ユニットは、当該量子化ユニットおよび対応するデコーダでは対応する逆量子化ユニットを有するエンコーダおよびデコーダの両方において利用可能なデータからサイド情報を抽出するよう構成されていてもよい。よって、エンコーダからデコーダへのサイド情報の伝送は追加的なビットを必要としなくてもよい。

40

【0031】

量子化ユニットは、サイド情報に依存して前記一組の量子化器を決定するよう構成されていてもよい。特に、前記量子化器の組内のディザリングされる量子化器の数はサイド情報に依存してもよい。一層具体的には、前記量子化器の組内に含まれるディザリングされる量子化器の数は、増大する予測器利得とともに減少してもよく、逆も成り立つ。量子化器の前記組をサイド情報に依存させることによって、量子化方式の知覚的パフォーマンス

50

が改善されうる。

【0032】

サイド情報は、分散保存フラグを含んでいてもよい。分散保存フラグは、係数のブロックの分散がどのように調整されるべきかを示してもよい。換言すれば、分散保存フラグは、デコーダによって実行されるべき処理を示していてもよく、それは量子化器によって再構成されるべき係数のブロックの分散に影響をもつ。

【0033】

例として、前記一組の量子化器は、分散保存フラグに依存して決定されてもよい。特に、ノイズ充填量子化器のノイズ利得は分散保存フラグに依存してもよい。代替的または追加的に、前記一つまたは複数のディザリングされる量子化器はあるSNR範囲をカバーしてもよく、該SNR範囲は分散保存フラグに依存して決定されてもよい。さらに、事後利得 γ が分散保存フラグに依存してもよい。代替的または追加的に、ディザリングされる量子化器の事後利得 γ は、予測器利得のあらかじめ決定された関数であるパラメータに依存して決定されてもよい。

10

【0034】

分散保存フラグは、量子化器のノイズ性 (noisiness) の度合いを予測の品質に適應させるために使われてもよい。例として、ディザリングされる量子化器の事後利得 γ は、予測器利得のあらかじめ決定された関数であるパラメータに依存して決定されてもよい。代替的または追加的に、事後利得 γ は、予測器利得のあらかじめ定義された件数によってスケーリングされた分散を保存する事後利得を、平均平方誤差最適事後利得 (mean-squared error optimal post gain) と比較し、それら二つの利得のうちの大きいほうを選択することによって決定されてもよい。特に、予測器利得のあらかじめ決定された関数は、予測器利得がまず二つれて再構成された信号の分散を小さくするものであってもよい。この結果として、コーデックの知覚的品質が改善されうる。

20

【0035】

あるさらなる側面によれば、量子化インデックスのブロックの第一の量子化インデックスを量子化解除するよう構成された逆量子化ユニット (本稿ではスペクトル・デコーダとも称される) が記述される。換言すれば、逆量子化ユニットは、係数データに基づいて (たとえば、量子化インデックスに基づいて) 係数のブロックについての再構成値を決定するよう構成されていてもよい。量子化ユニットのコンテキストにおいて本稿で述べてきたすべての特徴および側面は対応する逆量子化ユニットにも適用可能であることを注意しておくべきである。特に、これは、前記一組の量子化器の構造および設計、前記一組の量子化器のサイド情報への依存性、割り当てプロセスなどに関係する特徴に当てはまる。

30

【0036】

量子化インデックスは、複数の対応するビンについて複数の係数を含む係数のブロックに関連付けられていてもよい。特に、量子化インデックスは、量子化された係数の対応するブロックの量子化された係数 (または再構成値) に関連付けられてもよい。対応する量子化ユニットのコンテキストで概説したように、量子化された係数のブロックは、予測残差係数のブロックに対応してもよく、あるいはそれから導出されてもよい。より一般には、量子化された係数のブロックは、時間領域から周波数領域への変換を使ってオーディオ信号のセグメントから得られた変換係数のブロックから導出されたものであってもよい。

40

【0037】

逆量子化ユニットは、一組の量子化器を提供するよう構成されていてもよい。上記で概説したように、前記一組の量子化器は、逆量子化ユニットおよび対応する量子化ユニットにおいて利用可能なサイド情報に基づいて適應または生成されてもよい。前記一組の量子化器は典型的には、それぞれ複数の異なる信号対雑音比 (SNR) に関連付けられた複数の異なる量子化器を含む。さらに、前記一組の量子化器は、上記で概説したように、増大／減少するSNRに従って順序付けられてもよい。隣り合う量子化器の間でのSNRの増大／減少は実質的に一定であってもよい。

【0038】

50

前記複数の異なる量子化器は、量子化ユニットのノイズ充填量子化器に対応するノイズ充填量子化器を含んでいてもよい。ある好ましい例では、前記複数の異なる量子化器は、単一のノイズ充填量子化器を含んでいてもよい。逆量子化ユニットのノイズ充填量子化器は、所定の統計モデルに従って生成されるランダム変数の実現 (realization) を使って前記第一の係数の再構成を提供するよう構成される。よって、量子化インデックスのブロックは典型的には、ノイズ充填量子化器を使って再構成される係数についての量子化インデックスを含まないことを注意しておくべきである。よって、ノイズ充填量子化器を使って再構成される係数は、0ビットレートに関連付けられる。

【0039】

さらに、前記複数の異なる量子化器は、一つまたは複数のディザリングされる量子化器を含んでいてもよい。前記一つまたは複数のディザリングされる量子化器は、前記第一の量子化インデックスに第一の再構成値を割り当てるよう構成された、一つまたは複数のそれぞれの逆スカラー量子化器を含んでいてもよい。さらに、前記一つまたは複数のディザリングされる量子化器は、前記第一の再構成値から前記ディザ値を除去することによって、第一のディザリング解除された係数を決定するよう構成された一つまたは複数のそれぞれのディザ除去ユニットを有していてもよい。逆量子化ユニットのディザ生成器は、典型的には、量子化器のディザ生成器と同期している。量子化ユニットのコンテキストにおいて概説したように、前記一つまたは複数のディザリングされる量子化器は、好ましくは、前記一つまたは複数のディザリングされる量子化器のMSEパフォーマンスを改善するために、量子化器事後利得を適用する。

【0040】

加えて、前記複数の量子化器は、一つまたは複数のディザリングされない量子化器を含んでいてもよい。前記一つまたは複数のディザリングされない量子化器は、(その後のディザ除去を実行することなくおよび/または量子化器事後利得を適用することなく) 前記第一の量子化インデックスにそれぞれの再構成値を割り当てるよう構成されたそれぞれの一様スカラー量子化器を含んでいてもよい。

【0041】

さらに、逆量子化ユニットは、係数の前記ブロックからの第一の係数に帰される(または量子化された係数の前記ブロックからの第一の量子化された係数に帰される) SNRを示すSNR指示を決定するよう構成されていてもよい。SNR指示は、スペクトル・ブロック包絡(これは典型的には逆量子化ユニットを有するデコーダにおいても利用可能である)に基づいて、かつオフセット・パラメータ(これは典型的にはエンコーダからデコーダに伝送されるビットストリーム中に含められる)に基づいて、決定されてもよい。特に、SNR指示は、前記一組の量子化器から選択される逆量子化器(または量子化器)のインデックス番号を示していてもよい。逆量子化ユニットは、SNR指示に基づいて前記一組の量子化器から第一の量子化器を選択することにおいて進んでもよい。対応する量子化ユニットのコンテキストにおいて概説されるように、この選択プロセスは、順序付けられた一組の量子化器を使うとき、効率的な仕方を実装されてもよい。さらに、逆量子化ユニットは、選択された第一の量子化器を使って前記第一の係数について第一の量子化された係数を決定するよう構成されていてもよい。

【0042】

あるさらなる側面によれば、オーディオ信号をビットストリームにエンコードするよう構成された変換ベースのオーディオ・エンコーダが記述される。エンコーダは、係数のブロックからの複数の係数を量子化することによって複数の量子化インデックスを決定するよう構成された量子化ユニットを有していてもよい。量子化ユニットは、一つまたは複数のディザリングされる量子化器を有していてもよい。量子化ユニットは、本稿に記載される量子化ユニットに関連した特徴の任意のものを有していてもよい。

【0043】

前記複数の係数は、複数の対応する周波数ビンに関連付けられていてもよい。上記で概

10

20

30

40

50

説したように、係数のブロックは、オーディオ信号のセグメントから導出されたものであってもよい。特に、オーディオ信号のセグメントは、時間領域から周波数領域に変換されて変換係数のブロックを与えていてもよい。量子化ユニットによって量子化される係数のブロックは、変換係数の該ブロックから導出されたものであってもよい。

【0044】

エンコーダはさらに、ディザ実現を選択するよう構成されたディザ生成器を有していてもよい。さらに、エンコーダは、変換係数のあらかじめ定義された統計モデルに基づいて符号語を選択するよう構成されたエントロピー符号化器を有していてもよい。ここで、変換係数の統計モデル（すなわち確率分布関数）はさらに、ディザの前記実現を条件としていてもよい。そのような統計モデルは、次いで、量子化インデックスの確率、特に、係数
10 に対応するディザの前記実現を条件とした量子化インデックスの確率を計算するために使われてもよい。量子化インデックスの確率は、この量子化インデックスに関連付けられたバイナリー符号語を生成するために使われてもよい。さらに、量子化インデックスのシーケンスは、それぞれの確率に基づいて合同してエンコードされてもよい。ここで、該それぞれの確率は、前記それぞれのディザ実現を条件としていてもよい。たとえば、量子化インデックスのシーケンスのそのような合同エンコードは、算術符号化または範囲符号化（range coding）によって実装されてもよい。

【0045】

もう一つの側面によれば、エンコーダは、複数の所定のディザ実現の一つを選択するよう構成されたディザ生成器を有していてもよい。前記複数の所定のディザ実現は、M個の
20 異なる所定のディザ実現を含んでいてもよい。さらに、ディザ生成器は、選択されたディザ実現に基づいて前記複数の係数を量子化するための複数のディザ値を生成するよう構成されていてもよい。Mは1より大きい整数であってもよい。特に、所定のディザ実現の数Mは10、5、4またはそれより少なくてもよい。ディザ生成器は、本稿に記述されるディザ生成器に関係した特徴の任意のものを有していてもよい。

【0046】

さらに、エンコーダは、M個の所定のコードブックからコードブックを選択するよう構成されたエントロピー符号化器を有していてもよい。エントロピー符号化器はさらに、選択されたコードブックを使って前記複数の量子化インデックスをエントロピー符号化するよう構成されていてもよい。M個の所定のコードブックは、それぞれM個の所定のディザ実
30 現に関連付けられていてもよい。特に、M個の所定のコードブックは、それぞれM個の所定のディザ実現を使ってトレーニングされたものであってもよい。M個の所定のコードブックは、可変長ハフマン符号語を含んでいてもよい。

【0047】

エントロピー符号化器は、ディザ生成器によって生成されたディザ実現に関連付けられたコードブックを選択するよう構成されていてもよい。換言すれば、エントロピー符号化器は、前記複数の量子化インデックスを生成するために使われたディザ実現に関連付けられた（たとえば該ディザ実現のためにトレーニングされた）エントロピー符号化のための
40 コードブックを選択してもよい。そうすることにより、たとえディザリングされる量子化器を使うときでも、エントロピー符号化器の符号化利得が改善（たとえば最適化）される。たとえ比較的少数M個のディザ実現を使うときでも、ディザリングされる量子化器を使うことの知覚上の恩恵が達成されることが本発明者によって観察された。結果として、最適化されたエントロピー符号化を許容するために比較的少数M個のコードブックが提供されるだけでよい。

【0048】

エントロピー符号化された量子化インデックスを示す係数データは典型的には、対応するデコーダへの伝送または提供のためにビットストリーム中に挿入される。

【0049】

あるさらなる側面によれば、ビットストリームをデコードして再構成されたオーディオ信号を提供するよう構成された変換ベースのオーディオ・デコーダが記述される。対応す
50

るオーディオ・エンコーダのコンテキストにおいて記述された特徴および側面はオーディオ・デコーダにも適用可能であることを注意しておくべきである。特に、限られた数M個のディザ実現および対応する限られた数M個のコードブックの使用に係る諸側面は、オーディオ・デコーダにも適用可能である。

【0050】

オーディオ・デコーダは、M個の所定のディザ実現のうちの一つを選択するよう構成されたディザ生成器を有する。M個の所定のディザ実現は、対応するエンコーダによって使われるM個の所定のディザ実現と同じである。さらに、ディザ生成器は、選択されたディザ実現に基づいて複数のディザ値を生成するよう構成されていてもよい。Mは1より大きい整数であってもよい。例として、Mは10または5の範囲であってもよい。前記複数のディザ値は、対応する複数の量子化インデックスに基づいて対応する複数の量子化された係数を決定するよう構成されている一つまたは複数のディザリングされる量子化器を有する逆量子化ユニットによって使われてもよい。ディザ生成器および逆量子化ユニットは、それぞれ本稿に記載されるディザ生成器に係るおよび逆量子化ユニットに係る特徴の任意のものを有していてもよい。

10

【0051】

さらに、オーディオ・デコーダは、M個の所定のコードブックからコードブックを選択するよう構成されたエントロピー復号器を有していてもよい。M個の所定のコードブックは、対応する符号化器によって使われるコードブックと同じである。加えて、エントロピー復号器は、選択されたコードブックを使ってビットストリームから係数データをエントロピー復号して、前記複数の量子化インデックスを提供するよう構成されていてもよい。M個の所定のコードブックは、それぞれM個の所定のディザ実現と関連付けられていてもよい。エントロピー復号器は、ディザ生成器によって選択されたディザ実現に関連付けられたコードブックを選択するよう構成されていてもよい。再構成されたオーディオ信号は、前記複数の量子化された係数に基づいて決定される。

20

【0052】

あるさらなる側面によれば、発話信号をビットストリームにエンコードするよう構成された変換ベースの発話エンコーダが記述される。すでに上記で示したように、エンコーダは、本稿に記載されるエンコーダ関係の特徴および／または構成要素の任意のものを有していてもよい。特に、エンコーダは、変換係数の複数の逐次的なブロックを受領するよう構成されたフレーミング・ユニットを有していてもよい。前記複数の逐次的なブロックは、現在ブロックおよび一つまたは複数の以前のブロックを含む。さらに、前記複数の逐次的なブロックは、発話信号のサンプルを示す。特に、前記複数の逐次的なブロックは、修正離散コサイン変換(MDCT)のような時間領域から周波数領域への変換を使うことによって決定されたものであってもよい。よって、変換係数のブロックはMDCT係数を含んでいてもよい。変換係数の数は制限されていてもよい。例として、変換係数のブロックは、256個の周波数ビン内の256個の変換係数を含んでいてもよい。

30

【0053】

加えて、発話エンコーダは、対応する現在の(スペクトル)ブロック包絡(たとえば対応する調整された包絡)を使って変換係数の対応する現在のブロックを平坦化することによって、平坦化された変換係数の現在のブロックを決定するよう構成された平坦化ユニットを有していてもよい。さらに、発話エンコーダは、再構成された変換係数の一つまたは複数の以前のブロックに基づいて、かつ一つまたは複数の予測器パラメータに基づいて、推定された平坦化された変換係数の現在のブロックを予測するよう構成された予測器を有していてもよい。さらに、発話エンコーダは、平坦化された変換係数の現在のブロックに基づいて、かつ推定された平坦化された変換係数の現在のブロックに基づいて、予測誤差係数の現在のブロックを決定するよう構成された差分ユニットを有していてもよい。

40

【0054】

予測器は、重み付けされた平均平方誤差基準を使って(たとえば重み付けされた平均平方誤差基準を最小化することによって)推定された平坦化された変換係数の現在のブロッ

50

クを決定するよう構成されていてもよい。重み付けされた平均平方誤差基準は、現在のブロック包絡または現在のブロック包絡の何らかのあらかじめ定義された関数を重みとして考慮に入れてもよい。本稿では、重み付けされた平均平方誤差基準を使って予測器利得を決定するためのさまざまな異なる方法が記述される。

【0055】

さらに、発話エンコーダは、一組の所定の量子化器を使って予測誤差係数の現在のブロックから導出される係数を量子化するよう構成された量子化ユニットを有していてもよい。この量子化ユニットは、本稿に記載される量子化に関する特徴の任意のものを有していてもよい。特に、本量子化ユニットは、量子化された係数に基づいてビットストリームについて係数データを決定するよう構成されていてもよい。よって、係数データは、予測誤差係数の現在ブロックの量子化されたバージョンを示してもよい。

10

【0056】

変換ベースの発話エンコーダはさらに、一つまたは複数のスケーリング規則を使って予測誤差係数の現在ブロックに基づいて再スケーリングされた予測残差係数の現在ブロック（再スケーリングされた誤差係数のブロックとも称される）を決定するよう構成されたスケーリング・ユニットを有していてもよい。再スケーリングされた誤差係数の現在ブロックの決定は、および／または前記一つまたは複数のスケーリング規則は、平均では、再スケーリングされた誤差係数の現在ブロックの再スケーリングされた誤差係数の分散が、予測誤差係数の現在ブロックの予測誤差係数の分散より高いようなものであってもよい。特に、前記一つまたは複数のスケーリング規則は、予測誤差係数の分散がすべての周波数ビンまたは周波数帯域について1により近くなるようなものであってもよい。量子化ユニットは、係数データ（すなわち、係数についての量子化インデックス）を提供するために、再スケーリングされた誤差係数の現在ブロックの再スケーリングされた誤差予測残差係数を量子化するよう構成されていてもよい。

20

【0057】

予測誤差係数の現在ブロックは典型的には、対応する複数の周波数ビンについての複数の予測誤差係数を含む。前記スケーリング規則に従って予測誤差係数にスケーリング・ユニットによって適用されるスケーリング利得は、それぞれの予測誤差係数の周波数ビンに依存しうる。さらに、スケーリング規則は、前記一つまたは複数の予測器パラメータに、たとえば予測器利得に依存してもよい。代替的または追加的に、スケーリング規則は、現在のブロック包絡に依存してもよい。本稿では、周波数ビン・依存のスケーリング規則を決定するためのさまざまな異なる方法が記述される。

30

【0058】

変換ベースの発話エンコーダはさらに、現在ブロック包絡に基づいて割り当てベクトルを決定するよう構成されたビット割り当てユニットを有していてもよい。割り当てベクトルは、予測誤差係数の現在ブロックから導出された第一の係数を量子化するために使われる前記一組の量子化器からの第一の量子化器を示していてもよい。特に、割り当てベクトルは、それぞれ予測誤差係数の現在ブロックから導出された係数全部を量子化するために使われる量子化器を示していてもよい。例として、割り当てベクトルは、各周波数帯域（ $l=1, \dots, L$ ）について使われる異なる量子化器を示していてもよい。

40

【0059】

換言すれば、ビット割り当てユニットは、現在ブロック包絡および所与の最大ビットレート制約条件に基づいて割り当てベクトルを決定するよう構成されていてもよい。ビット割り当てユニットは、前記一つまたは複数のスケーリング規則にも基づいて割り当てベクトルを決定するよう構成されていてもよい。レート割り当てベクトルの次元は典型的には周波数帯域の数 L に等しい。割り当てベクトルのエントリーは、レート割り当てベクトルのそれぞれのエントリーに関連付けられた周波数帯域に属する係数を量子化するために使われるべき、前記一組の量子化器からの量子化器のインデックスを示してもよい。特に、割り当てベクトルは、それぞれ予測誤差係数の現在ブロックから導出される係数すべてを量子化するために使われる量子化器を示してもよい。

50

【0060】

ビット割り当てユニットは、予測誤差係数の現在ブロックについての係数データが所定のビット数を超えないよう割り当てベクトルを決定するよう構成されていてもよい。さらに、ビット割り当てユニットは、現在ブロック包絡から導出される（たとえば、現在の調整された包絡から導出される）割り当て包絡に適用されるべきオフセットを示すオフセット・パラメータを決定するよう構成されていてもよい。オフセット・パラメータは、対応するデコーダが、係数データを決定するために使われた量子化器を同定できるようにするために、ビットストリーム中に含まれてもよい。

【0061】

変換ベースの発話エンコーダはさらに、量子化された係数に関連付けられた量子化インデックスをエントロピー符号化するよう構成されたエントロピー符号化器を有していてもよい。エントロピー符号化器は、算術エンコーダを使って量子化インデックスを符号化するよう構成されていてもよい。あるいはまた、エントロピー符号化器は、（本稿で記述されるように）複数のM個の所定のコードブックを使って量子化インデックスを符号化するよう構成されていてもよい。

【0062】

もう一つの側面によれば、ビットストリームをデコードして再構成された発話信号を提供する変換ベースの発話デコーダが記述される。発話デコーダは、本稿に記述される特徴および／または構成要素の任意のものを有していてもよい。特に、デコーダは、再構成された変換係数の一つまたは複数の以前のブロックに基づいて、かつビットストリームから導出される一つまたは複数の予測器パラメータに基づいて、推定された平坦化された変換係数の現在ブロックを決定するよう構成された予測器を有していてもよい。さらに、発話デコーダは、一組の量子化器を使って、ビットストリーム内に含まれる係数データに基づいて、量子化された予測誤差係数（またはその再スケーリングされたバージョン）の現在ブロックを決定するよう構成された逆量子化ユニットを有していてもよい。特に、逆量子化ユニットは、対応する発話エンコーダによって使われた一組の量子化器に対応する一組の（逆）量子化器を利用してもよい。

【0063】

逆量子化ユニットは、受領されたビットストリームから導出されるサイド情報に依存して前記一組の量子化器（および／または対応する一組の逆量子化器）を決定するよう構成されていてもよい。特に、逆量子化ユニットは、対応する発話エンコーダの量子化ユニットと同じ、前記一組の量子化器についての選択プロセスを、実行してもよい。前記一組の量子化器を前記サイド情報に依存させることによって、再構成される発話信号の知覚的品質が改善されうる。

【0064】

もう一つの側面によれば、係数のブロックの第一の係数を量子化する方法が記述される。係数のブロックは、複数の対応する周波数ビンについての複数の係数を含む。本方法は、一組の量子化器を提供することを含んでいてもよい。前記一組の量子化器は、それぞれ複数の異なる信号対雑音比（SNR）に関連付けられた複数の異なる量子化器を含む。前記複数の異なる量子化器は、ノイズ充填量子化器、一つまたは複数のディザリングされる量子化器および一つまたは複数のディザリングされない量子化器を含んでいてもよい。本方法はさらに、前記第一の係数に帰されるSNRを示すSNR指示を決定することを含んでいてもよい。本方法はさらに、前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択し、前記第一の量子化器を使って前記第一の係数を量子化することを含んでいてもよい。

【0065】

さらなる側面によれば、量子化インデックスを量子化解除する方法が記述される。換言すれば、本方法は、対応する量子化方法を使って量子化された係数のブロックについて、再構成値（量子化された係数とも称される）を決定することに向けられてもよい。再構成値は、量子化インデックスに基づいて決定されてもよい。しかしながら、係数の前記プロ

10

20

30

40

50

ックからの係数のいくつかは、ノイズ充填量子化器を使って量子化されていてもよいことを注意しておくべきである。この場合、これらの係数についての再構成値は、量子化インデックスとは独立に決定されてもよい。

【0066】

上記で概説したように、量子化インデックスは、複数の対応する周波数ビンについて複数の係数を含む係数のブロックに関連付けられる。特に、量子化インデックスは、ノイズ充填量子化器を使って量子化されたのではない係数のブロックの係数と一対一の関係で対応してもよい。本方法は、一組の量子化器（または逆量子化器）を提供することを含んでいてもよい。前記一組の量子化器は、それぞれ複数の異なる信号対雑音比（SNR）に関連付けられた複数の異なる量子化器を含んでいてもよい。前記複数の異なる量子化器は、ノイズ充填量子化器、一つまたは複数のディザリングされる量子化器および一つまたは複数のディザリングされない量子化器を含んでいてもよい。本方法は、係数のブロックの第一の係数に帰されるSNRを示すSNR指示を決定することを含んでいてもよい。本方法は、前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択し、係数のブロックの前記第一の係数についての第一の量子化された係数（すなわち、再構成値）を決定することにおいて進んでもよい。

【0067】

もう一つの側面によれば、オーディオ信号をビットストリームにエンコードする方法が記述される。本方法は、ディザリングされる量子化器を使って係数のブロックからの複数の係数を量子化することによって複数の量子化インデックスを決定することを含んでいてもよい。係数のブロックはオーディオ信号から導出されてもよい。本方法は、M個の所定のディザ実現の一つを選択し、選択されたディザ実現に基づいて前記複数の係数を量子化するための複数のディザ値を生成することを含んでいてもよい。ここで、Mは1より大きい整数である。さらに、本方法は、M個の所定のコードブックからコードブックを選択し、選択されたコードブックを使って前記複数の量子化インデックスをエントロピー符号化することを含んでいてもよい。M個の所定のコードブックは、それぞれM個の所定のディザ実現に関連付けられていてもよく、選択されたコードブックは、選択されたディザ実現に関連付けられていてもよい。さらに、本方法は、エントロピー符号化された量子化インデックスを示す係数データをビットストリーム中に挿入することを含んでいてもよい。

【0068】

あるさらなる側面によれば、ビットストリームをデコードして再構成されたオーディオ信号を提供する方法が記述される。本方法は、M個の所定のディザ実現のうちの一つを選択し、選択されたディザ実現に基づいて複数のディザ値を生成することを含んでいてもよい。ここで、Mは1より大きい整数である。前記複数のディザ値は、対応する複数の量子化インデックスに基づいて対応する複数の量子化された係数を決定するようディザリングされる量子化器を有する逆量子化ユニットによって使われてもよい。よって、本方法は、ディザリングされた（逆）量子化器を使って前記複数の量子化された係数を決定することを含んでいてもよい。加えて、本方法は、M個の所定のコードブックからコードブックを選択し、選択されたコードブックを使ってビットストリームから係数データをエントロピー復号して、前記複数の量子化インデックスを提供することを含んでいてもよい。M個の所定のコードブックは、それぞれM個の所定のディザ実現と関連付けられていてもよく、選択されたコードブックは、選択されたディザ実現に関連付けられていてもよい。さらに、本方法は、前記複数の量子化された係数に基づいて前記再構成されたオーディオ信号を決定することを含んでいてもよい。

【0069】

あるさらなる側面によれば、発話信号をビットストリームにエンコードする方法が記述される。本方法は、現在ブロックおよび一つまたは複数の以前のブロックを含む変換係数の複数の逐次的なブロックを受領することを含んでいてもよい。前記複数の逐次的なブロックは、発話信号のサンプルを示す。さらに、本方法は、再構成された変換係数の一つまたは複数の前のブロックに基づき、かつ予測器パラメータに基づいて、推定された変換係

数の現在ブロックを決定することを含んでいてもよい。再構成された変換係数の前記一つまたは複数の前のブロックは、変換係数の前記一つまたは複数の前のブロックから導出されたものであってもよい。本方法は、変換係数の現在ブロックに基づき、かつ推定された変換係数の現在ブロックに基づいて、予測誤差係数の現在ブロックを決定することにおいて進んでもよい。さらに、本方法は、一組の量子化器を使って、予測誤差係数の現在ブロックから導出される係数を量子化することを含んでいてもよい。前記一組の量子化器は、本稿に記述される特徴の任意のものを示してもよい。さらに、本方法は、量子化された係数に基づいてビットストリームについて係数データを決定することを含んでいてもよい。

【0070】

もう一つの側面によれば、ビットストリームをデコードして再構成された発話信号を提供する方法が記述される。本方法は、再構成された変換係数の一つまたは複数の前のブロックに基づき、かつビットストリームから導出された予測器パラメータに基づいて、推定された変換係数の現在ブロックを決定することを含んでいてもよい。さらに、本方法は、一組の量子化器を使って、ビットストリーム内に含まれる係数データに基づいて量子化された予測残差係数の現在ブロックを決定することを含んでいてもよい。前記一組の量子化器は、本稿に記述される特徴の任意のものを有していてもよい。本方法は、推定された変換係数の現在ブロックに基づき、かつ量子化された予測誤差係数の現在ブロックに基づいて、再構成された変換係数の現在ブロックを決定することにおいて進んでもよい。再構成された発話信号は、再構成された変換係数の現在ブロックに基づいて決定されてもよい。

【0071】

さらなる側面によれば、ソフトウェア・プログラムが記述される。ソフトウェア・プログラムは、プロセッサ上での実行のために、かつ該プロセッサによって実行されたときに本稿において概説される方法段階を実行するために適応されていてもよい。

【0072】

さらなる側面によれば、記憶媒体が記述される。記憶媒体は、プロセッサ上での実行のために、かつ該プロセッサによって実行されたときに本稿において概説される方法段階を実行するために適応されたソフトウェア・プログラムを有していてもよい。

【0073】

さらなる側面によれば、コンピュータ・プログラム・プロダクトが記述される。コンピュータ・プログラムは、コンピュータ上で実行されたときに本稿において概説される方法段階を実行するための実行可能命令を含んでいてもよい。

【0074】

本特許出願において概説される好ましい実施形態を含む方法およびシステムは、単独で、あるいは本稿に開示される他の方法およびシステムとの組み合わせで使われてもよいことを注意しておくべきである。さらに、本特許出願において概説される方法およびシステムのすべての側面は、さまざまな仕方で組み合わせられる。特に、請求項の特徴は、任意の仕方で互いに組み合わせられる。

【図面の簡単な説明】

【0075】

本発明は、付属の図面を参照して例示的な仕方で下記に説明される。

【図1a】一定ビットレートでビットストリームを提供する例示的なオーディオ・エンコーダのブロック図である。

【図1b】可変ビットレートでビットストリームを提供する例示的なオーディオ・エンコーダのブロック図である。

【図2】変換係数の複数のブロックに基づく例示的な包絡の生成を示す図である。

【図3a】変換係数のブロックの例示的な包絡を示す図である。

【図3b】例示的な補間された包絡の決定を示す図である。

【図4】量子化器の諸集合を示す図である。

【図5a】例示的なオーディオ・デコーダのブロック図である。

【図5b】図5aのオーディオ・デコーダの例示的な包絡デコーダのブロック図である。

【図 5 c】図 5 a のオーディオ・デコーダの例示的なサブバンド予測器のブロック図である。

【図 5 d】図 5 a のオーディオ・デコーダの例示的なスペクトル・デコーダのブロック図である。

【図 6 a】受け入れ可能な量子化器の例示的な集合のブロック図である。

【図 6 b】例示的なディザリングされる量子化器のブロック図である。

【図 6 c】変換係数のブロックのスペクトルに基づく量子化器の例示的な選択を示す図である。

【図 7】エンコーダおよび対応するデコーダにおける一組の量子化器を決定する例示的な方式を示す図である。

【図 8】ディザリングされる量子化器を使って決定された、エントロピー符号化された量子化インデックスを復号するための例示的な方式のブロック図である。

【図 9】a ~ c は、例示的な実験結果を示す図である。

【図 10】例示的なビット割り当てプロセスを示す図である。

【発明を実施するための形態】

【0076】

背景セクションにおいて概説したように、発話または声信号について相対的に高い符号化利得を示す変換ベースのオーディオ・コーデックを提供することが望ましい。そのような変換ベースのオーディオ・コーデックは、変換ベースの発話コーデックまたは変換ベースの声コーデックと称されてもよい。変換ベースの発話コーデックは、やはり変換領域で動作するので、AACまたはHE-AACのような一般的な変換ベースのオーディオ・コーデックと便利に組み合わせられる。さらに、入力オーディオ信号のセグメント（たとえばフレーム）の発話または非発話への分類およびその後の一般的なオーディオ・コーデックと特定の発話コーデックとの間の切り換えは、両方のコーデックが変換領域で動作するという事実のため、簡略化される。

【0077】

図 1 a は、例示的な変換ベースの発話エンコーダ 100 のブロック図を示している。エンコーダ 100 は、入力として、変換係数のブロック 131（符号化単位とも称される）を受領する。変換係数のブロック 131 は、入力オーディオ信号のサンプルのシーケンスを時間領域から変換領域に変換するよう構成された変換ユニットによって得られたものであってもよい。変換ユニットは、MDCTを実行するよう構成されていてもよい。変換ユニットは、AACまたはHE-AACのような一般的なオーディオ・コーデックの一部であってもよい。そのような一般的なオーディオ・コーデックは、異なるブロック・サイズ、たとえば長ブロックおよび短ブロックを利用してもよい。例示的なブロック・サイズは長ブロックについては1024サンプル、短ブロックについては256サンプルである。サンプリング・レート44.1kHzおよび50%の重複を想定すると、長ブロックは入力オーディオ信号の約20msをカバーし、短ブロックは入力オーディオ信号の約5msをカバーする。長ブロックは典型的には、入力オーディオ信号の静的セグメントのために使われ、短ブロックは典型的には入力オーディオ信号の過渡的セグメントのために使われる。

【0078】

発話信号は、約20msの時間的セグメントにおいて静的であると考えられてもよい。特に、発話信号のスペクトル包絡は、約20msの時間的セグメントにおいて静的であると考えられてもよい。そのような20msセグメントについて変換領域において意味のある統計量を導出できるためには、変換ベースの発話エンコーダ 100 に、変換係数の（たとえば5msの長さをもつ）諸短ブロック 131 を提供することが有用でありうる。そうすることにより、複数の短ブロック 131 は、たとえば20msの時間セグメント（たとえば長ブロックの時間セグメント）に関して統計を導出するために使用されることができる。さらに、これは、発話信号について十分な時間分解能を提供する利点がある。

【0079】

よって、変換ユニットは、入力オーディオ信号の現在セグメントが発話に分類される場合には、変換係数の短ブロック131を提供するよう構成されていてもよい。エンコーダ100は、ブロック131の集合132と称される変換係数の複数のブロック131を抽出するよう構成されたフレーミング・ユニット101を有していてもよい。ブロックの集合132はフレームと称されてもよい。例として、ブロック131の集合132は、256個の変換係数の四つの短ブロックを含んでいてもよく、それにより入力オーディオ信号の約20msのセグメントをカバーする。

【0080】

ブロックの集合132は、包絡推定ユニット102に提供されてもよい。包絡推定ユニット102は、ブロックの集合132に基づいて包絡133を決定するよう構成されていてもよい。包絡133は、ブロックの集合132内に含まれる複数のブロック131の対応する変換係数の二乗平均平方根（RMS）値に基づいていてもよい。ブロック131は典型的には、対応する複数の周波数ビン301（図3a参照）において複数の変換係数（たとえば256個の変換係数）を与える。複数の周波数ビン301は、複数の周波数帯域302にグループ化されてもよい。複数の周波数帯域302は、音響心理学的考察に基づいて選択されてもよい。例として、周波数ビン301は、対数スケールまたはバーク・スケールに従って周波数帯域302にグループ化されてもよい。ブロックの現在集合132に基づいて決定された包絡134は、それぞれ複数の周波数帯域302についての複数のエネルギー値を含んでいてもよい。特定の周波数帯域302についての特定のエネルギー値は、その特定の周波数帯域302内にはいる周波数ビン301に対応する、集合132の諸ブロック131の変換係数に基づいて決定されてもよい。特定のエネルギー値は、これらの変換係数のRMS値に基づいて決定されてもよい。よって、ブロックの現在の集合132についての包絡133（現在の包絡133とも称される）は、ブロックの現在集合132内に含まれる変換係数の諸ブロック131の平均包絡を示してもよく、あるいは包絡133を決定するために使われる変換係数の諸ブロック132の平均包絡を示してもよい。

【0081】

現在の包絡133が、ブロックの現在集合132に隣接する変換係数の一つまたは複数のさらなるブロック131に基づいて決定されてもよいことを注意しておくべきである。これは図2において示されている。そこでは、現在の包絡133（量子化された現在の包絡134によって示される）がブロックの現在集合132の諸ブロック131に基づき、かつブロックの現在集合132に先行するブロックの集合からのブロック201に基づいて決定される。図示した例では、現在の包絡133は、五つのブロック131に基づいて決定される。現在の包絡133を決定するときに隣接するブロックを考慮に入れることにより、ブロックの隣接する諸集合132の諸包絡の連続性が保証されうる。

【0082】

現在の包絡133を決定するとき、異なるブロック131の変換係数が重み付けされてもよい。特に、現在の包絡133を決定するために考慮に入れられた最も外側のブロック201、202は、残りのブロック131より低い重みを有していてもよい。例として、最も外側のブロック201、202の変換係数は、0.5で重み付けされてもよく、他のブロック131の変換係数は1で重み付けされてもよい。

【0083】

ブロックの先行する集合132の諸ブロック201を考慮するのと同様の仕方で、ブロックの直後の集合132の一つまたは複数のブロック（いわゆる先読みブロック）が、現在の包絡133を決定するために考慮されてもよいことを注意しておくべきである。

【0084】

現在の包絡133のエネルギー値は、対数スケールで（たとえばdBスケールで）表わされてもよい。現在の包絡133は、現在の包絡133のエネルギー値を量子化するよう構成されている包絡量子化ユニット103に提供されてもよい。包絡量子化ユニット103は、所定の量子化器分解能、たとえば3dBの分解能を提供してもよい。包絡133の量子化インデックスは、エンコーダ100によって生成されたビットストリーム内の包絡デー

タ 1 6 1 として提供されてもよい。さらに、量子化された包絡 1 3 4、すなわち包絡 1 3 3 の量子化されたエネルギー値を有する包絡は、補間ユニット 1 0 4 に提供されてもよい。

【0085】

補間ユニット 1 0 4 は、量子化された現在の包絡 1 3 4 に基づき、かつ量子化された以前の包絡 1 3 5（ブロックの現在集合 1 3 2 の直前のブロックの集合 1 3 2 について決定されたもの）に基づいてブロックの現在の集合 1 3 2 の各ブロック 1 3 1 について包絡を決定するよう構成されている。補間ユニット 1 0 4 の動作は図 2、図 3 a および図 3 b に示されている。図 2 は、変換係数の諸ブロック 1 3 1 のシーケンスを示している。ブロック 1 3 1 のシーケンスはブロックの相続く諸集合 1 3 2 にグループ化される。ここで、ブロックの各集合 1 3 2 は、量子化された包絡、たとえば量子化された現在の包絡 1 3 4 および量子化された以前の包絡 1 3 5 を決定するために使われる。図 3 a は、量子化された以前の包絡 1 3 5 および量子化された現在の包絡 1 3 4 の例を示している。上記で示したように、これらの包絡は、スペクトル・エネルギー 3 0 3 を（たとえば dB スケールで）示していてもよい。同じ周波数帯域 3 0 2 についての量子化された以前の包絡 1 3 5 および量子化された現在の包絡 1 3 4 の対応するエネルギー値 3 0 3 が（たとえば線形補間を使って）補間されて、補間された包絡 1 3 6 を決定してもよい。換言すれば、ある特定の周波数帯域 3 0 2 の諸エネルギー値 3 0 3 が補間されて、その特定の周波数帯域 3 0 2 内の補間された包絡 1 3 6 のエネルギー値 3 0 3 を提供してもよい。

【0086】

補間された包絡 1 3 6 が決定され、適用されるブロックの集合は、量子化された現在の包絡 1 3 4 が決定されるもとになったブロックの現在の集合 1 3 2 とは異なることがあることを注意しておくべきである。これは図 2 に示されている。図 2 は、ブロックのシフトされた集合 3 3 2 を示している。これは、ブロックの現在の集合 1 3 2 に比してシフトされており、ブロックの以前の集合 1 3 2 のブロック 3 および 4（それぞれ参照符号 2 0 3 および 2 0 1 によって示されている）およびブロックの現在の集合 1 3 2 のブロック 1 および 2（それぞれ参照符号 2 0 4 および 2 0 5 によって示されている）を含む。実のところ、量子化された現在の包絡 1 3 4 に基づき、かつ量子化された以前の包絡 1 3 5 に基づいて決定された補間された包絡 1 3 6 は、ブロックの現在の集合 1 3 2 のブロックについての関連性に比べ、ブロックのシフトされた集合 3 3 2 のブロックについて増大した関連性を有することがある。

【0087】

よって、図 3 b に示される補間された包絡が、ブロックのシフトされた集合 3 3 2 のブロック 1 3 1 を平坦化するために使われてもよい。これは、図 2 と組み合わせて図 3 b によって示されている。図 3 b の補間された包絡 3 4 1 が図 2 のブロック 2 0 3 に適用されてもよいこと、図 3 b の補間された包絡 3 4 2 が図 2 のブロック 2 0 1 に適用されてもよいこと、図 3 b の補間された包絡 3 4 3 が図 2 のブロック 2 0 4 に適用されてもよいこと、図 3 b の補間された包絡 3 4 4（図示した例ではこれは量子化された現在の包絡 1 3 6 に対応）が図 2 のブロック 2 0 5 に適用されてもよいこと、が見て取れる。よって、量子化された現在の包絡 1 3 4 を決定するためのブロックの集合 1 3 2 は、補間された包絡 1 3 6 がそれについて決定され、補間された包絡 1 3 6 が（平坦化のために）それに適用されるところのブロックのシフトされた集合 3 3 2 とは異なることがある。特に、量子化された現在の包絡 1 3 4 は、ブロックのシフトされた集合 3 3 2 のブロック 2 0 3、2 0 1、2 0 4、2 0 5 に関してある種の先読みを使って決定されてもよい。これらのブロックは、量子化された現在の包絡 1 3 4 を使って平坦化される。これは、連続性の観点から有益である。

【0088】

補間された包絡 1 3 6 を決定するためのエネルギー値 3 0 3 の補間は、図 3 b に示される。量子化された以前の包絡 1 3 5 のエネルギー値から量子化された現在の包絡 1 3 4 の

対応するエネルギー値の間の補間により、補間された包絡 1 3 6 のエネルギー値が、ブロックのシフトされた集合 3 3 2 の諸ブロック 1 3 1 について決定されることが見て取れる。特に、シフトされた集合 3 3 2 の各ブロック 1 3 1 について、補間された包絡 1 3 6 が決定されてもよく、それによりブロックのシフトされた集合 3 3 2 の複数のブロック 2 0 3、2 0 1、2 0 4、2 0 5 について複数の補間された包絡 1 3 6 を提供する。変換係数のあるブロック 1 3 1（たとえば、ブロックのシフトされた集合 3 3 2 のブロック 2 0 3、2 0 1、2 0 4、2 0 5 のうちの任意のもの）の補間された包絡 1 3 6 は、変換係数のブロック 1 3 1 をエンコードするために使われてもよい。現在の包絡 1 3 3 の量子化インデックス 1 6 1 がビットストリーム内の対応するデコーダに提供されることを注意しておくべきである。結果として、対応するデコーダは、エンコーダ 1 0 0 の補間ユニット 1 0 4 と類似の仕方で前記複数の補間された包絡 1 3 6 を決定するよう構成されていてもよい。

10

【0089】

フレーミング・ユニット 1 0 1、包絡推定ユニット 1 0 3、包絡量子化ユニット 1 0 3 および補間ユニット 1 0 4 はブロックの集合（すなわち、ブロックの現在の集合 1 3 2 および／またはブロックのシフトされた集合 3 3 2）に対して動作する。他方、変換係数の実際のエンコードはブロックごとに実行されてもよい。以下では、ブロックのシフトされた集合 3 3 2（または変換ベースの発話エンコーダ 1 0 0 の他の実装において可能性としてはブロックの現在の集合 1 3 2）の複数のブロック 1 3 1 のうちの任意のものであってよい変換係数の現在ブロック 1 3 1 のエンコードが参照される。

20

【0090】

現在ブロック 1 3 1 についての現在の補間された包絡 1 3 6 は、現在ブロック 1 3 1 の変換係数のスペクトル包絡の近似を提供してもよい。エンコーダ 1 0 0 は、事前平坦化ユニット 1 0 5 および包絡利得決定ユニット 1 0 6 を有していてもよい。これらは、現在の補間された包絡 1 3 6 に基づき、かつ現在ブロック 1 3 1 に基づいて、現在ブロック 1 3 1 についての調整された包絡 1 3 9 を決定するよう構成される。特に、現在ブロック 1 3 1 の平坦化された変換係数の分散が調整されるよう、現在ブロック 1 3 1 についての包絡利得が決定されてもよい。 $X(k)$ 、 $k=1, \dots, K$ は現在ブロック 1 3 1 の変換係数であってもよく（たとえば $K=256$ ）、 $E(k)$ 、 $k=1, \dots, K$ は現在の補間された包絡 1 3 6 の平均スペクトル・エネルギー値 3 0 3 であってもよい（同じ周波数帯域 3 0 2 のエネルギー値 $E(k)$ は等しい）。包絡利得 a は、平坦化された変換係数

30

【数 2】

$$\tilde{X}(k) = \frac{X(k)}{a \cdot \sqrt{E(k)}}$$

の分散が調整されるよう決定されてもよい。特に、包絡利得 a は分散が 1 になるよう決定されてもよい。

【0091】

40

包絡利得 a が、変換係数の現在ブロック 1 3 1 の完全な周波数範囲のサブ範囲について決定されてもよいことを注意しておくべきである。換言すれば、包絡利得 a は、周波数ビン 3 0 1 の部分集合のみに基づいておおよび／または周波数帯域 3 0 2 の部分集合のみに基づいて決定されてもよい。例として、包絡利得 a は、開始周波数ビン 3 0 4（開始周波数ビンは 0 または 1 より大きい）より大きい諸周波数ビン 3 0 1 に基づいて決定されてもよい。結果として、現在ブロック 1 3 1 についての調整された包絡 1 3 9 は、包絡利得 a を、開始周波数ビン 3 0 4 より上にある諸周波数ビン 3 0 1 に関連付けられた現在の補間された包絡 1 3 6 の平均スペクトル・エネルギー値 3 0 3 にのみ適用することによって決定されてもよい。よって、現在のブロック 1 3 1 についての調整された包絡 1 3 9 は、開始周波数ビン以下の諸周波数ビン 3 0 1 については現在の補間された包絡 1 3 6 に対応しても

50

よく、開始周波数より上の諸周波数ビン 3 0 1 については現在の補間された包絡 1 3 6 を包絡利得 a によりオフセットしたものに对应してもよい。これは、調整された包絡 3 3 9 によって図 3 a に示されている（破線で示す）。

【0 0 9 2】

包絡利得 a 1 3 7（レベル補正利得とも称される）の、現在の補間された包絡 1 3 6 への適用は、現在の補間された包絡 1 3 6 の調整またはオフセットに対応し、それにより図 3 a に示されるように調整された包絡 1 3 9 が与えられる。包絡利得 a 1 3 7 は、利得データ 1 6 2 として、ビットストリーム中にエンコードされてもよい。

【0 0 9 3】

エンコーダ 1 0 0 はさらに、包絡利得 a 1 3 7 に基づき、かつ現在の補間された包絡 1 3 6 に基づいて、調整された包絡 1 3 9 を決定するよう構成される包絡洗練ユニット 1 0 7 を有していてもよい。調整された包絡 1 3 9 は、変換係数のブロック 1 3 1 の信号処理のために使われてもよい。包絡利得 a 1 3 7 は、（3dB のきざみで量子化されていてもよい）現在の補間された包絡 1 3 6 に比べ、より高い分解能に（たとえば 1dB きざみで）量子化されてもよい。よって、調整された包絡 1 3 9 は、包絡利得 a 1 3 7 の前記より高い分解能まで（たとえば、1dB きざみで）量子化されてもよい。

【0 0 9 4】

さらに、包絡洗練ユニット 1 0 7 は、割り当て包絡 1 3 8 を決定するよう構成されていてもよい。割り当て包絡 1 3 8 は、調整された包絡 1 3 9 の量子化されたバージョン（たとえば 3dB 量子化レベルまで量子化）に対応してもよい。割り当て包絡 1 3 8 は、ビット割り当て目的のために使われてもよい。特に、割り当て包絡 1 3 8 は、——現在ブロック 1 3 1 のある特定の変換係数について——所定の一群の量子化器からある特定の量子化器を決定するために使われてもよい。ここで、前記特定の量子化器が、前記特定の変換係数を量子化するために使われる。

【0 0 9 5】

エンコーダ 1 0 0 は、調整された包絡 1 3 9 を使って現在ブロック 1 3 1 を平坦化し、それにより平坦化された変換係数のブロック 1 4 0 を与えるよう構成されている平坦化ユニット 1 0 8 を有する。平坦化された変換係数のブロック 1 4 0 は、変換領域内で予測ループを使ってエンコードされてもよい。よって、ブロック 1 4 0 は、サブバンド予測器 1 1 7 を使ってエンコードされてもよい。予測ループは、平坦化された変換係数

【数 3】

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

のブロック 1 4 0 に基づき、かつ推定された変換係数

【数 4】

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

のブロック 1 5 0 に基づき、予測誤差係数 $\Delta(k)$ のブロック 1 4 1 を決定するよう構成された差分ユニット 1 1 5 を有する。たとえば、

【数 5】

$$\Delta(k) = \tilde{X}(k) - \hat{X}(k)$$

ブロック 1 4 0 が平坦化された変換係数、すなわち調整された包絡 1 3 9 のエネルギー値 3 0 3 を使って正規化または平坦化された変換係数を含むという事実のため、推定された

10

20

30

40

50

変換係数のブロック150も平坦化された変換係数の推定値を含むことを注意しておくべきである。換言すれば、差分ユニット115はいわゆる平坦化領域（flattened domain）で動作する。結果として、予測誤差係数 $\Delta(k)$ のブロック141は平坦化された領域で表わされる。

【0096】

予測誤差係数 $\Delta(k)$ のブロック141は、1とは異なる分散を示すことがある。エンコーダ100は、予測誤差係数 $\Delta(k)$ を再スケーリングして、再スケーリングされた誤差係数のブロック142を与えるよう構成された再スケーリング・ユニット111を有していてもよい。再スケーリング・ユニット111は、再スケーリングを実行するために一つまたは複数の所定のヒューリスティック規則を利用してもよい。結果として、再スケーリングされた誤差係数のブロック142は、（予測誤差係数のブロック141に比べて）（平均で）1により近い分散を示す。これは、その後の量子化およびエンコードにとって有益であることがある。

10

【0097】

エンコーダ100は、予測誤差係数のブロック141または再スケーリングされた誤差係数のブロック142を量子化するよう構成された係数量子化ユニット112を有する。係数量子化ユニット112は、一組の所定の量子化器を有していてもよく、あるいはそれを利用してもよい。前記一組の所定の量子化器は、異なる精密度または異なる分解能をもつ諸量子化器を提供してもよい。このことは、種々の量子化器321、322、323が示される図4に示されている。種々の量子化器は異なるレベルの精度（異なるdB値によって示される）を提供しうる。前記複数の量子化器321、322、323のうち特定の量子化器が、割り当て包絡138の特定の値に対応してもよい。よって、割り当て包絡138のエネルギー値は、前記複数の量子化器の対応する量子化器をポイントしてもよい。よって、割り当て包絡138の決定は、ある特定の誤差係数について使われるべき量子化器の選択プロセスを簡略化しうる。換言すれば、割り当て包絡138はビット割り当てプロセスを簡略化しうる。

20

【0098】

前記一組の量子化器は、量子化誤差をランダム化するためのディザリングを利用する一つまたは複数の量子化器322を有していてもよい。これは図4に示されている。この図は、ディザリングされる量子化器（dithered quantizer）の部分集合324を含む所定の量子化器の第一の集合326と、ディザリングされる量子化器の部分集合325を含む所定の量子化器の第二の集合327とを示している。よって、係数量子化ユニット112は、所定の量子化器の異なる集合326、327を利用しうる。ここで、係数量子化ユニット112によって使用される所定の量子化器の集合は、予測器117によって提供されるおよび／またはエンコーダにおいておよび対応するデコーダにおいて利用可能な他のサイド情報に基づいて決定される制御パラメータ146に依存してもよい。特に、係数量子化ユニット112は、制御パラメータ146に基づいて、再スケーリングされた誤差係数のブロック142を量子化するための所定の量子化器の集合326、327を選択するよう構成されていてもよい。ここで、制御パラメータ146は、予測器117によって提供される一つまたは複数の予測器パラメータに依存してもよい。前記一つまたは複数の予測器パラメータは、予測器117によって提供される推定される変換係数のブロック150の品質を示してもよい。

30

40

【0099】

量子化された誤差係数は、たとえばハフマン符号を使ってエントロピー符号化されてもよく、それによりエンコーダ100によって生成されるビットストリーム中に含まれる係数データ163を与える。

【0100】

以下では、一組326の量子化器321、322、323の選択または決定に関するさ

50

らなる詳細を述べる。一組 3 2 6 の量子化器は、量子化器の順序付けられた集合 3 2 6 に対応してもよい。量子化器の順序付けられた集合 3 2 6 は N 個の量子化器を含み、各量子化器は異なる歪みレベルに対応してもよい。よって、量子化器の集合 3 2 6 は、 N 個の可能な歪みレベルを提供しうる。集合 3 2 6 の量子化器は、歪みの降順に従って（あるいは等価だが SNR の昇順に従って）順序付けされてもよい。さらに、量子化器は、整数ラベルでラベル付けされてもよい。例として、量子化器は 0, 1, 2 などとラベル付けされてもよい。ここで、整数ラベルの増大は SNR の増大を示してもよい。

【0 1 0 1】

量子化器の集合 3 2 6 は、二つの連続する量子化器の間の SNR ギャップが少なくともほぼ一定であるようなものであってもよい。たとえば、ラベル「1」をもつ量子化器の SNR は 1.5dB であってもよく、ラベル「2」をもつ量子化器の SNR は 3.0dB であってもよい。よって、量子化器の順序付けられた集合 3 2 6 の量子化器は、第一の量子化器から隣接する第二の量子化器に変わることによって、第一および第二の量子化器のすべての対について、SNR（信号対雑音比）が実質的に一定値（たとえば 1.5dB）だけ増大するようなものであってもよい。

【0 1 0 2】

量子化器の集合 3 2 6 は、次の量子化器を含んでいてもよい。

- ・ノイズ充填量子化器 3 2 1。これは、0dB よりやや低いまたは 0dB に等しい SNR を与えうる。該 SNR はレート割り当てプロセスについては 0dB と近似されてもよい。
- ・ N_{dith} 個の量子化器 3 2 2。これは減算的ディザリングを使ってもよく、典型的には中間的な SNR レベルに対応する。（たとえば $N_{\text{dith}} > 0$ ）
- ・ N_{cq} 個の古典的量子化器 3 2 3。これは減算的ディザリングを使わず、典型的には比較的高い SNR レベルに対応する（たとえば $N_{\text{cq}} > 0$ ）。ディザリングされない量子化器 3 2 3 はスカラー量子化器に対応しうる。

【0 1 0 3】

量子化器の総数 N は $N = 1 + N_{\text{dith}} + N_{\text{cq}}$ によって与えられる。

【0 1 0 4】

量子化器集合 3 2 6 の例が図 6 a に示されている。量子化器の集合 3 2 6 のノイズ充填量子化器 3 2 1 はたとえば、あらかじめ定義された統計モデルに従ってランダム変数の実現を出力する乱数発生器を使って実装されてもよい。そのような乱数発生器のある可能な実装は、あらかじめ定義された統計モデルのランダム・サンプルとともに固定したテーブルを使うことおよび可能性としてはその後の再規格化に関わる。エンコーダ 1 0 0 において使われる乱数発生器は、対応するデコーダにおける乱数発生器と同期している。これらの乱数発生器の同期は、共通のシードを使ってこれらの乱数発生器を初期化することによっておよび／または固定した諸時点においてこれらの乱数発生器の状態をリセットすることによって得られてもよい。あるいはまた、これらの発生器は、規定された統計モデルに従って生成されたランダム・データを含んでいるルックアップテーブルとして実装されてもよい。特に、予測器がアクティブであれば、ノイズ充填量子化器 3 2 1 の出力がエンコーダ 1 0 0 と対応するデコーダにおいて同じであることが保証されうる。

【0 1 0 5】

加えて、量子化器の集合 3 2 6 は、一つまたは複数のディザリングされる量子化器 3 2 2 を含んでいてもよい。前記一つまたは複数のディザリングされる量子化器は、図 6 a に示されるように、擬似数ディザ信号 6 0 2 の実現を使って生成されてもよい。擬似数ディザ信号 6 0 2 は、擬似ランダム・ディザ値のブロック 6 0 2 に対応してもよい。ディザ数のブロック 6 0 2 は、量子化されるべき、再スケーリングされた誤差係数のブロック 1 4 2 の次元と同じ次元を有していてもよい。ディザ信号 6 0 2（またはディザ値のブロック 6 0 2）は、ディザ生成器 6 0 1 を使って生成されてもよい。特に、ディザ信号 6 0 2 は、一様に分布したランダム・サンプルを含むルックアップテーブルを使って生成されてもよい。

【0 1 0 6】

図6bのコンテキストにおいて示されるように、ディザ値のブロック602の個々のディザ値632は、量子化されるべき対応する係数に（たとえば、再スケーリングされた誤差係数のブロック142の対応する再スケーリングされた誤差係数に）ディザを適用するために使われる。再スケーリングされた誤差係数のブロック142は、総数K個の再スケーリングされた誤差係数を含んでいてもよい。同様にして、ディザ値のブロック602はK個のディザ値632を含んでいてもよい。ディザ値のブロック602のk番目のディザ値632、 $k=1, \dots, K$ は、再スケーリングされた誤差係数のブロック142のk番目の再スケーリングされた誤差係数に適用されてもよい。

【0107】

上記で示したように、ディザ値のブロック602は、量子化されるべき再スケーリングされた誤差係数のブロック142と同じ次元を有していてもよい。これは量子化器の集合326のすべてのディザリングされる量子化器322についてディザ値の単一のブロック602を使うことを許容するので、有益である。換言すれば、再スケーリングされた誤差係数の所与のブロック142を量子化し、エンコードするために、擬似ランダム・ディザ602は、量子化器のすべての許容可能な集合326、327について、歪みについてのすべての可能な割り当てについて、一回生成されればよい。これは、エンコーダ100と対応するデコーダの間の同期を達成することを容易にする。単一のディザ信号602の使用は対応するデコーダに明示的に信号伝達される必要がないからである。特に、エンコーダ100および対応するデコーダは、再スケーリングされた誤差係数のブロック142についてのディザ値の同じブロック602を生成するよう構成されている同じディザ生成器601を利用してもよい。

【0108】

量子化器の集合326の組成は、好ましくは、音響心理学的な考察に基づく。低レートの変換符号化は、変換係数に適用される通常の量子化方式において行なわれる逆注水（reverse-water filling）プロセスの性質によって引き起こされる、スペクトル上の穴および帯域制限を含むスペクトル・アーチファクトにつながりうる。スペクトルの穴の可聴性は、ノイズを、短期間にわたってたまたま水位より低かった、よって0ビットレートを割り当てられた周波数帯域302にノイズを注入することによって、低減されることができ

【0109】

周波数領域における係数の粗い量子化は、特定の周波数帯域302の係数があるフレームでは0に量子化され（深いスペクトルの穴の場合）、次のフレームでは0でない値に量子化され、その後プロセス全体が数十ミリ秒にわたって繰り返されるときに状況において生成される特定の符号化アーチファクト（たとえば、深いスペクトルの穴、いわゆる「バーディー（birdies）」）につながりうる。量子化器が粗いほど、そのような挙動の生成を起こしやすくなる。この技術的問題は、（たとえば米国特許第7447631号において概説されているように）0レベルでの信号再構成のために使われる量子化インデックスにノイズ充填を適用することによって対処されうる。米国特許第7447631号において記述される解決策は、0レベル量子化に関連付けられた深いスペクトルの穴の可聴性を低減するので、アーチファクトの低減を容易にするが、より浅いスペクトルの穴に関連付けられたアーチファクトは残る。ノイズ充填法を粗い量子化器の量子化インデックスにも適用することができるが、これはこれらの量子化器のMSEパフォーマンスを著しく劣化させる。この欠点は、ディザリングされる量子化器を使うことによって対処できることが、発明者によって観察された。本稿では、MSEパフォーマンス問題に対処するために、低いSNRレベルについて減算的ディザをもつ量子化器322を使うことが提案される。さらに、減算的ディザをもつ量子化器322の使用は、すべての再構成レベルについてノイズ充填属性を容易にする。ディザリングされる量子化器322は、任意のビットレートにおいて解析的に扱うことができるので、高歪みレベル（すなわち低レート）で有用な事後利得614を導出することによって、ディザリングに起因するパフォーマンス損失を低減（たとえば最小化）することが可能である。

【0110】

一般に、ディザリングされる量子化器322を用いて任意に低いビットレートを達成することが可能である。たとえば、スカラーの場合、非常に大きな量子化きざみサイズを使うことを選んでよい。にもかかわらず、0ビットレート動作は実際上は現実的ではない。可変長符号化器と一緒に量子化器の動作を可能にするために必要とされる数値精度に対して、強い要求を課すからである。これは、0dB SNRの歪みレベルに対して、ディザリングされる量子化器322を適用するのではなく、一般的なノイズ充填量子化器321を適用する動機を与える。量子化器の提案される集合326は、ディザリングされる量子化器322が、比較的小さなきざみサイズに関連付けられている歪みレベルについて使われ、可変長符号化が、数値精度を維持することに関する問題に対処する必要なしに実装できるよう、設計される。

10

【0111】

スカラー量子化の場合、減算的ディザリングのある量子化器322は、ほぼ最適なMSEパフォーマンスを提供する事後利得を使って実装されてもよい。減算的にディザリングされるスカラー量子化器322の例が図6bに示されている。ディザリングされる量子化器322は、減算的ディザリング構造内で使われる一様スカラー量子化器Q612を有している。減算的ディザリング構造は、(再スケーリングされた誤差係数のブロック142からの)対応する誤差係数から(ディザ値のブロック602からの)ディザ値632を減算するよう構成されているディザ減算ユニット611を有する。さらに、減算的ディザリング構造は、対応するスカラー量子化された誤差係数に(ディザ値のブロック602からの)ディザ値632を加算するよう構成された対応する加算ユニット613を有する。図示した例では、ディザ減算ユニット611は、スカラー量子化器Q612の上流に置かれ、ディザ加算ユニット613はスカラー量子化器Q612の下流に置かれる。ディザ値のブロック602からのディザ値632は、区間 $[-0.5, 0.5]$ または $[0, 1]$ からの値にスカラー量子化器612のきざみサイズをかけた値を取ってもよい。ディザリングされる量子化器322の代替的な実装では、ディザ減算ユニット611とディザ加算ユニット613は互いに交換されうることを注意しておく。

20

【0112】

減算的ディザリング構造には、量子化器事後利得 γ によって前記量子化された誤差係数を再スケーリングするよう構成されているスケーリング・ユニット614が続いてもよい。量子化された誤差係数のスケーリングのあと、量子化された誤差係数のブロック145が得られる。ディザリングされる量子化器322への入力Xは典型的には、ディザリングされた量子化器322を使って量子化されるべき特定の周波数帯域中にはいる、再スケーリングされた誤差係数のブロック142の係数に対応することを注意しておくべきである。同様にして、ディザリングされる量子化器322の出力は典型的には、その特定の周波数帯域中にはいる、量子化された誤差係数のブロック145の量子化された係数に対応する。

30

【0113】

ディザリングされる量子化器322への入力Xは零平均であり、入力Xの分散 $\sigma_X^2 = E\{X^2\}$ は既知であると想定されてもよい。(たとえば、信号の分散は、信号の包絡から決定されうる。)さらに、ディザ値632を含む擬似ランダム・ディザ・ブロックZ602がエンコーダ100および対応するデコーダに対して利用可能であると想定されてもよい。さらに、ディザ値632は入力Xとは独立であると想定されてもよい。さまざまな異なるディザ602が使われうるが、以下では、ディザZ602は、0と Δ の間に一様に分布していると想定する。それは、 $U(0, \Delta)$ によって表わされてもよい。実際には、いわゆるSchuchman条件を満たすいかなるディザが使用されてもよい(たとえば、 $[-0.5, 0.5]$ かけるスカラー量子化器612のきざみサイズ Δ の間に一様に分布しているディザ602)。量子化器Q612は、格子[ラティス]であってもよく、そのボロノイ・セルの広がり Δ であってもよい。この場合、ディザ信号は使用されるラティスのボロノイ・セルの広がり Δ にわたって一様分布をもつことになる。

40

50

【0114】

量子化器事後利得 γ は、信号の分散および量子化きざみサイズが当てられて導出される。ディザ量子化器は、任意のきざみサイズ（すなわちビットレート）について解析的に扱えるからである。特に、事後利得は、減算的ディザをもつ量子化器のMSEパフォーマンスを改善するよう導出されてもよい。事後利得は次式によって与えられてもよい。

【数6】

$$\gamma = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}$$

10

たとえ事後利得 γ の適用によってディザリングされる量子化器 322 のMSEパフォーマンスが改善されうるとしても、ディザリングされる量子化器 322 は典型的には、ディザリングなしの量子化器より低いMSEパフォーマンスをもつ（このパフォーマンス損失はビットレートが増すと消失するが）。結果として、一般に、ディザリングされる量子化器は、ディザリングされないバージョンよりノイズが多い。よって、ディザリングされる量子化器 322 の使用がディザリングされる量子化器 322 の知覚的に有益なノイズ充填属性によって正当化されるときにのみ、ディザリングされる量子化器 322 を使うことが望ましいことがありうる。

20

【0115】

よって、三つの型の量子化器を含む量子化器の集合 326 が提供されてもよい。順序付けられた量子化器集合 326 は、単一のノイズ充填量子化器 321 と、減算的ディザリングのある一つまたは複数の量子化器 322 と、一つまたは複数の古典的な（ディザリングされない）量子化器 323 を含んでいてもよい。連続する量子化器 321、322、323 はSNRに対して段階的な改善を提供しうる。量子化器の順序付けられた集合 326 の隣り合う一対の量子化器の間の段階的な改善は、隣り合う量子化器の対の一部または全部について実質的に一定であってもよい。

【0116】

量子化器の特定の集合 326 は、ディザリングされる量子化器 322 の数によって、かつ特定の集合 326 内に含まれるディザリングされない量子化器 323 の数によって定義されてもよい。さらに、量子化器の特定の集合 326 は、ディザ信号 602 の特定の實現によって定義されてもよい。集合 326 は、変換係数の知覚的に効率的な量子化を提供するために設計されてもよく、0レート・ノイズ充填（0dBよりわずかに低いまたは0dBに等しいSNRを与える）；中間的な歪みレベル（中間的なSNR）での減算的ディザリングによるノイズ充填；および低歪みレベル（高いSNR）でのノイズ充填の欠如を与える。集合 326 は、レート割り当てプロセスの間に選択されうる一組の許容可能な量子化器を提供する。量子化器の集合 326 からの特定の量子化器の、ある特定の周波数帯域 302 の係数への適用は、レート割り当てプロセスの間に決定される。特定の周波数帯域 302 の係数を量子化するためにどの量子化器が使われるかは典型的には事前には知られていない。しかしながら、典型的には、量子化器の集合 326 の組成がどうなっているかは事前に知られている。

30

40

【0117】

誤差係数のブロック 142 の異なる周波数帯域 302 について異なる型の量子化器を使う側面が図 6c に示されている。ここでは、レート割り当てプロセスの例示的な帰結が示されている。この例では、レート割り当てはいわゆる逆注水原理に従うと想定される。図 6c は、入力信号のスペクトル 625（または量子化されるべき係数のブロックの包絡線）を示している。周波数帯域 623 が比較的高いスペクトル・エネルギーをもち、比較的低い歪みレベルを与える古典的量子化器 323 を使って量子化されることが見て取れる。周波数帯域 622 は水レベル 624 より上のスペクトル・エネルギーを示す。これらの周

50

波数帯域 6 2 2 における係数は、中程度の歪みレベルを与えるディザリングされる量子化器 3 2 2 を使って量子化されてもよい。周波数帯域 6 2 1 は水レベル 6 2 4 より下のスペクトル・エネルギーを示す。これらの周波数帯域 6 2 1 における係数は、0 レートのノイズ充填を使って量子化されてもよい。（スペクトル 6 2 5 によって表わされる）係数の特定のブロックを量子化するために使われる異なる量子化器は、その特定の係数ブロックについて決定された、量子化器の特定の集合 3 2 6 の一部であってもよい。

【0 1 1 8】

よって、三つの異なる型の量子化器 3 2 1、3 2 2、3 2 3 は選択的に適用されてもよい（たとえば、周波数に関して選択的に）。特定の型の量子化器の適用についての決定は、下記で述べるレート割り当て手順のコンテキストにおいて決定されてもよい。レート割り当て手順は、入力信号の RMS 包絡から（またはたとえば信号のパワースペクトル密度から）導出できる知覚的基準を利用してもよい。特定の周波数帯域 3 0 2 において適用される量子化器の型は、対応するデコーダに明示的に信号伝達される必要がない。量子化器の選択された型を信号伝達する必要性がなくなるのは、対応するデコーダが、入力信号のブロックを量子化するために使われた量子化器の特定の集合 3 2 6 を、根底にある知覚的基準（たとえば割り当て包絡 1 3 8）から、量子化器の集合の所定の組成（たとえば、量子化器の種々の集合の所定のセット）からおよび単一のグローバルなレート割り当てパラメータ（オフセット・パラメータとも称される）から決定できるからである。

【0 1 1 9】

エンコーダ 1 0 0 によって使用された量子化器の集合 3 2 6 のデコーダにおける決定は、量子化器の集合 3 2 6 を、量子化器がその歪み（たとえば SNR）に従って順序付けられるように設計することによって容易にされる。集合 3 2 6 の各量子化器は、一定値だけ直前の量子化器の歪みを減少させてもよい（SNR を洗練してもよい）。さらに、量子化器の特定の集合 3 2 6 は、レート割り当てプロセス全体の間、擬似ランダム・ディザ信号 6 0 2 の単一の実現に関連付けられていてもよい。この結果として、レート割り当て手順の帰結は、ディザ信号 6 0 2 の実現に影響しない。これは、レート割り当て手順の収束を保証するために有益である。さらに、これは、デコーダがディザ信号 6 0 2 の単一の実現を知っている場合にデコーダがデコードを実行することを可能にする。デコーダは、エンコーダ 1 0 0 においてと対応するデコーダにおいてとで同じ擬似ランダム・ディザ生成器 6 0 1 を使うことによって、ディザ信号 6 0 2 の実現を知らされてもよい。

【0 1 2 0】

上記で示したように、エンコーダ 1 0 0 は、ビット割り当てプロセスを実行するよう構成されていてもよい。この目的のために、エンコーダ 1 0 0 は、ビット割り当てユニット 1 0 9、1 1 0 を有していてもよい。ビット割り当てユニット 1 0 9 は、再スケーリングされた誤差係数の現在ブロック 1 4 2 をエンコードするために利用可能である総ビット数 1 4 3 を決定するよう構成されていてもよい。総ビット数 1 4 3 は、割り当て包絡 1 3 8 に基づいて決定されてもよい。ビット割り当てユニット 1 1 0 は、割り当て包絡 1 3 8 における対応するエネルギー値に依存して、種々の再スケーリングされた誤差係数へのビットの相対的な割り当てを提供するよう構成されていてもよい。

【0 1 2 1】

ビット割り当てプロセスは、逐次反復的な割り当て手順を利用してもよい。割り当て手順の過程で、割り当て包絡 1 3 8 は、オフセット・パラメータを使ってオフセットされてもよい。それにより、増大／減少した分解能をもつ量子化器が選択される。よって、オフセット・パラメータは、全体的な量子化を洗練するまたは粗くするために使われてもよい。オフセット・パラメータは、オフセット・パラメータおよび割り当て包絡 1 3 8 によって与えられる量子化器を使って得られる係数データ 1 6 3 が、現在ブロック 1 3 1 に割り当てられた総ビット数 1 4 3 に対応する（またはそれを超えない）ビット数を含むように決定されてもよい。現在ブロック 1 3 1 をエンコードするためにエンコーダ 1 0 0 によって使われたオフセット・パラメータは、係数データ 1 6 3 としてビットストリーム中に含まれる。結果として、対応するデコーダは、再スケーリングされた誤差係数のブロック

1 4 2を量子化するために係数量子化ユニット 1 1 2によって使用された量子化器を決定することができるようにされる。

【0 1 2 2】

よって、レート割り当てプロセスは、エンコーダ 1 0 0において実行されてもよく、知覚的モデルに従って利用可能なビット 1 4 3を分配することをねらいとする。知覚的モデルは、変換係数のブロック 1 3 1から導出された割り当て包絡 1 3 8に依存してもよい。レート割り当てアルゴリズムは利用可能なビット 1 4 3を、異なる型の量子化器、すなわち0レートのノイズ充填 3 2 1、前記一つまたは複数のディザリングされる量子化器 3 2 2および前記一つまたは複数の古典的なディザリングされない量子化器 3 2 3の間で分配する。スペクトルの特定の周波数帯域 3 0 2の係数を量子化するために使われる量子化器の型についての最終的な決定は、知覚的信号モデル、擬似ランダム・ディザの実現およびビットレート制約条件に依存しうる。

10

【0 1 2 3】

対応するデコーダでは、ビット割り当て（割り当て包絡 1 3 8およびオフセット・パラメータによって示される）は、無損失のデコードを容易にするために、量子化インデックスの確率を計算するために使われてもよい。量子化インデックスの確率の計算方法であって、フル帯域擬似ランダム・ディザ 6 0 2の実現、単一の包絡 1 3 8およびレート割り当てパラメータ（すなわちオフセット・パラメータ）によってパラメータ化される知覚的モデルの使用を用いるものが使われてもよい。割り当て包絡 1 3 8、オフセット・パラメータおよびディザ値のブロック 6 0 2に関する知識を使って、デコーダにおける量子化器の集合 3 2 6の組成が、エンコーダ 1 0 0において使われた集合 3 2 6と同期しうる。

20

【0 1 2 4】

上記で概説したように、ビットレート制約条件は、フレーム当たりの最大許容されるビット数 1 4 3を用いて指定されてもよい。これはたとえば、たとえばハフマン符号を使ってその後、エントロピー符号化される量子化インデックスに適用される。特に、これは、一時に単一のパラメータが量子化される、逐次的な仕方ではビットストリームが生成される符号化シナリオにおいて適用され、対応する量子化インデックスがバイナリー符号語に変換されてビットストリームにアペンドされる。

【0 1 2 5】

算術符号化（または範囲符号化）が使われている場合、原理は異なる。算術符号化のコンテキストでは、典型的には量子化インデックスの長いシーケンスに単一の符号語が割り当てられる。ビットストリームの特定の部分がある特定のパラメータと厳密に関連付けることは、典型的には可能ではない。特に、算術符号化のコンテキストでは、信号のランダムな実現をエンコードするために必要とされるビット数は典型的には未知である。これは、たとえ信号の統計モデルが既知であったとしても成り立つ。

30

【0 1 2 6】

上述した技術的課題に対処するために、算術符号化器をレート割り当てアルゴリズムの一部とすることが提案される。レート割り当てプロセスの間に、エンコーダは一つまたは複数の周波数帯域 3 0 2の係数の集合を量子化し、エンコードしようとする。そのようなすべての試行について、算術符号化器の状態の変化を観察し、（ビット数を計算する代わりに）ビットストリーム中で進む位置の数を計算することが可能である。最大ビットレート制約条件が設定されている場合、この最大ビットレート制約条件は、レート割り当て手順において使用されてもよい。算術符号の終端ビット（termination bits）のコストは、最後の符号化されたパラメータのコストに含められてもよく、一般に、終端ビットのコストは算術符号化器の状態に依存して変わる。にもかかわらず、ひとたび終端コストが利用可能になれば、前記一つまたは複数の周波数帯域 3 0 2の係数の集合に対応する量子化インデックスをエンコードするために必要とされるビット数を決定することができる。

40

【0 1 2 7】

算術符号化のコンテキストでは、ディザ 6 0 2の単一の実現が（係数の特定のブロック

50

142の) レート割り当てプロセス全体について使用されてもよいことを注意しておくべきである。上記で概説したように、算術符号化器は、レート割り当て手順内で特定の量子化器選択のビットレート・コストを推定するために使われてもよい。算術符号化器の状態の変化が観察されてもよく、状態変化が、量子化を実行するために必要とされるビット数を計算するために使われてもよい。さらに、算術符号の終端のプロセスが、レート割り当てプロセス内で使われてもよい。

【0128】

上記で示したように、量子化インデックスは、算術符号またはエントロピー符号を使ってエンコードされてもよい。量子化インデックスがエントロピー符号化される場合、個々の量子化インデックスまたは量子化インデックスのグループに可変長の符号語を割り当てるために、量子化インデックスの確率分布が考慮に入れられてもよい。ディザリングの使用は、量子化インデックスの確率分布に対する影響をもつことがある。特に、ディザ信号602の特定の実現は、量子化インデックスの確率分布に影響をもつことがある。ディザ信号602の実現の事実上無制限の数のため、一般的な場合において、符号語確率は事前に知られてはおらず、ハフマン符号化を使うことは可能ではない。

【0129】

可能なディザ実現の数を、ディザ信号602の実現の比較的小さい、扱える程度の集合にまで減らすことが可能であることが発明者によって観察された。例として、各周波数帯域302について、ディザ値の限られた集合が提供されてもよい。この目的のために、エンコーダ100（および対応するデコーダ）は、M個の所定のディザ実現のうちのの一つを選択することによってディザ信号602を生成するよう構成された離散的ディザ生成器801を有していてもよい（図8参照）。例として、M個の異なる所定のディザ実現は、すべての周波数帯域302について使用されてもよい。所定のディザ実現の数は $M < 5$ であってもよい（たとえば $M=4$ または $M=3$ ）。

【0130】

ディザ実現の限られた数Mのため、各ディザ実現について（可能性としては多次元の）ハフマン・コードブックをトレーニングすることが可能である。それにより、M個のコードブックの集合603が与えられる。エンコーダ100は、選択されたディザ実現に基づいて、M個の所定のコードブックの集合803のうちのの一つを選択するよう構成されているコードブック選択ユニット802を有していてもよい。そうすることにより、エントロピー符号化がディザ生成と同期していることが保証される。選択されたコードブック811は、選択されたディザ実現を使って量子化された個々の量子化インデックスまたは量子化インデックスのグループをエンコードするために使われてもよい。結果として、ディザリングされる量子化器を使うときエントロピー符号化のパフォーマンスが改善されることができる。

【0131】

所定のコードブックの集合803および離散的ディザ生成器801は、（図8に示されるように）対応するデコーダにおいても使用されてもよい。擬似ランダム・ディザが使われる場合、およびデコーダがエンコーダ100と同期したままである場合、デコードは実現可能である。この場合、デコーダにおいて離散的ディザ生成器801はディザ信号602を生成し、特定のディザ実現はコードブックの集合803から特定のハフマン・コードブック811に一意的に関連付けられている。音響心理学的モデル（たとえば、割り当て包絡138およびレート割り当てパラメータによって表わされる）および選択されたコードブック811を与えられて、デコーダはハフマン・デコーダ551を使ったデコードを実行し、デコードされた量子化インデックス812を与えることができる。

【0132】

よって、算術符号化の代わりに、ハフマン・コードブックの比較的小さな集合803が使われてもよい。ハフマン・コードブックの集合813からの特定のコードブック811の使用は、ディザ信号602の所定の実現に依存してもよい。同時に、M個の所定のディザ実現を形成する許容可能なディザ値の限られた集合が使われてもよい。その際、レート

割り当てプロセスは、ディザリングされない量子化器、ディザリングされる量子化器およびハフマン符号化の使用に関わっていてもよい。

【0133】

再スケーリングされた誤差係数の量子化の結果として、量子化された誤差係数のブロック145が得られる。量子化された誤差係数のブロック145は、対応するデコーダにおいて利用可能な誤差係数のブロックに対応する。結果として、量子化された誤差係数のブロック145は推定された変換係数のブロック150を決定するために使用されうる。エンコーダ100は、再スケーリング・ユニット113によって実行された再スケーリング動作の逆を実行して、それによりスケーリングされた量子化された誤差係数のブロック147を与えるよう構成された逆再スケーリング・ユニット113を有していてもよい。推定された変換係数のブロック150をスケーリングされた量子化された誤差係数のブロック147に加えることによって、再構成された平坦化された係数のブロック148を決定するために、加算ユニット116が使われてもよい。さらに、再構成された平坦化された係数のブロック148に調整された包絡139を適用し、それにより再構成された係数のブロック149を与えるために、逆平坦化ユニット114が使われてもよい。再構成された係数のブロック149は、対応するデコードにおいて利用可能な変換係数のブロック131のバージョンに対応する。結果として、再構成された係数のブロック149は、予測器117において、推定された係数のブロック150を決定するために使われてもよい。

10

【0134】

再構成された係数のブロック149は、平坦化されていない領域で表現されている。すなわち、再構成された係数のブロック149は、現在ブロック131のスペクトル包絡をも表わす。下記で概説するように、これは、予測器117のパフォーマンスにとって有益であることがある。

20

【0135】

予測器117は、再構成された係数の一つまたは複数の以前のブロック149に基づいて、推定された変換係数のブロック150を推定するよう構成されていてもよい。特に、予測器117は、所定の予測誤差基準が低減される（たとえば最小化される）よう一つまたは複数の予測器パラメータを決定するよう構成されていてもよい。例として、前記一つまたは複数の予測器パラメータは、予測誤差係数のブロック141のエネルギーまたは知覚的に重み付けされたエネルギーが低減される（たとえば最小化される）よう決定されてもよい。前記一つまたは複数の予測器パラメータは、予測器データ164として、エンコーダ100によって生成されるビットストリーム中に含められてもよい。

30

【0136】

予測器117は、その内容が参照によって組み込まれる特許出願US61750052およびその優先権を主張する諸特許出願において記述されているような信号モデルを利用してもよい。前記一つまたは複数の予測器パラメータは、信号モデルの一つまたは複数のモデル・パラメータに対応してもよい。

【0137】

図1bは、さらなる例示的な変換ベースの発話エンコーダ170のブロック図を示している。図1bの変換ベースの発話エンコーダ170は、図1aのエンコーダ100のコンポーネントの多くを有するが、図1bの変換ベースの発話エンコーダ170は可変ビットレートをもつビットストリームを生成するよう構成されている。この目的のために、エンコーダ170は、先行する諸ブロック131のためにビットストリームによってすでに使用されたビットレートを追跡するよう構成された平均ビットレート（ABR: Average Bit Rate）状態ユニット172を有する。ビット割り当てユニット171は、変換係数の現在ブロック131をエンコードするために利用可能な総ビット数143を決定するためにこの情報を使う。

40

【0138】

全体的に、変換ベースの発話エンコーダ100、170は、以下を示すまたは以下を含

50

むビットストリームを生成するよう構成される。

- ・量子化された現在の包絡134を示す包絡データ161。量子化された現在の包絡134は、変換係数のブロックの現在の集合132またはシフトされた集合332の諸ブロックの包絡を記述するために使われる。

- ・変換係数の現在ブロック131の補間された包絡136を調整するためのレベル補正利得aを示す利得データ162。典型的には、ブロックの現在の集合132またはシフトされた集合332の各ブロック131について異なる利得aが提供される。

- ・現在ブロック131についての予測誤差係数のブロック141を示す係数データ163。特に、係数データ163は、量子化された誤差係数のブロック145を示す。さらに、係数データ163は、デコーダにおいて逆量子化を実行するための量子化器を決定するために使われてもよいオフセット・パラメータを示してもよい。

- ・再構成された係数の以前のブロック149から、推定された係数のブロック150を決定するために使われるべき一つまたは複数の予測器係数を示す予測器データ164。

【0139】

以下では、対応する変換ベースの発話デコーダ500が図5aないし5dのコンテキストにおいて記述される。図5aは、例示的な変換ベースの発話デコーダ500のブロック図を示している。ブロック図は、再構成された係数のブロック149を変換領域から時間領域に変換し、それによりデコードされたオーディオ信号のサンプルを与えるために使われる合成フィルタバンク504（逆変換ユニットとも称される）を示している。合成フィルタバンク504は、所定のストライド（たとえば、約5msまたは256サンプルのストライド）をもつ逆MDCTを利用してもよい。

【0140】

デコーダ500の主ループは、このストライドの単位で動作する。各ステップは、システムの所定の帯域幅設定に対応する長さまたは次元をもつ変換領域ベクトル（ブロックとも称される）を生成する。合成フィルタバンク504の変換サイズにするゼロ・パディングに際し、変換領域ベクトルは、合成フィルタバンク504の重複／加算プロセスへの所定の長さ（たとえば5ms）の時間領域信号更新を合成するために使われる。

【0141】

上記で示したように、一般的な変換ベースのオーディオ・コーデックは、典型的には、過渡成分の扱いのための5ms範囲の短ブロックのシーケンスをもつ諸フレームを用いる。よって、一般的な変換ベースのオーディオ・コーデックは、短ブロックおよび長ブロックのシームレスな共存のための必要な変換および窓切り換えツールを提供する。したがって、図5aの合成フィルタバンク504を省略することによって定義される声スペクトル・フロントエンドは、追加的な切り換えツールを導入する必要なしに、汎用の変換ベースのオーディオ・コーデックに便利に統合されうる。換言すれば、図5aの変換ベースの発話デコーダ500は、一般的な変換ベースのオーディオ・デコーダと便利に組み合わせられてもよい。特に、図5aの変換ベースの発話デコーダ500は、一般的な変換ベースのオーディオ・デコーダ（たとえばAACまたはHE-AACデコーダ）によって提供される合成フィルタバンク504を利用してもよい。

【0142】

はいってくるビットストリームから（特にビットストリーム内に含まれる包絡データ161からおよび利得データ162から）、包絡デコーダ503によって、信号包絡が決定されてもよい。特に、包絡デコーダ503は、包絡データ161および利得データ162に基づいて、調整された包絡139を決定するよう構成されてもよい。よって、包絡デコーダ503は、エンコーダ100、170の補間ユニット104および包絡洗練ユニット107と同様のタスクを実行してもよい。上記で概説したように、調整された包絡109は、あらかじめ定義された周波数帯域302の集合における信号分散のモデルを表わす。

【0143】

さらに、デコーダ500は、調整された包絡139を、名目上分散1であってもよい要素をもつ平坦化領域ベクトルに適用するよう構成されている逆平坦化ユニット114を有

10

20

30

40

50

する。平坦化領域ベクトルは、エンコーダ100、170のコンテキストにおいて記述された再構成された平坦化された係数のブロック148に対応する。逆平坦化ユニット114の出力において、再構成された係数のブロック149が得られる。再構成された係数のブロック149は、(デコードされたオーディオ信号を生成するための)合成フィルタバンク504およびサブバンド予測器517に与えられる。

【0144】

サブバンド予測器517は、エンコーダ100、170の予測器117と同様の仕方で動作する。特に、サブバンド予測器517は、(ビットストリーム内で信号伝達される前記一つまたは複数の予測器パラメータを使って)再構成された係数の一つまたは複数の以前のブロック149に基づいて、(平坦化された領域における)推定された変換係数のブロック150を決定するよう構成されている。換言すれば、サブバンド予測器517は、予測器ラグおよび予測器利得のような予測器パラメータに基づいて、以前にデコードされた出力ベクトルおよび信号包絡のバッファから、予測された平坦化領域ベクトルを出力するよう構成されている。デコーダ500は、予想器データ164をデコードして前記一つまたは複数の予測器パラメータを決定するよう構成された予測器デコーダ501を有する。

10

【0145】

デコーダ500はさらに、典型的にはビットストリームの最大の部分に基づいて(すなわち、係数データ163に基づいて)予測された平坦化領域ベクトルに加算的補正を備えるよう構成されているスペクトル・デコーダ502を有する。スペクトル・デコード・プロセスは、前記包絡および伝送された割り当て制御パラメータ(オフセット・パラメータとも称される)から導出される割り当てベクトルによって主として制御される。図5aに示されるように、スペクトル・デコーダ502の予測器パラメータ520への直接的な依存性があってもよい。よって、スペクトル・デコーダ502は、受領された係数データ163に基づいてスケーリングされた量子化された誤差係数のブロック147を決定するよう構成されていてもよい。エンコーダ100、170のコンテキストで概説したように、再スケーリングされた誤差係数のブロック142を量子化するために使われる量子化器321、322、323は、典型的には、割り当て包絡138(これは調整された包絡139から導出できる)およびオフセット・パラメータに依存する。さらに、量子化器321、322、323は、予測器117によって提供される制御パラメータ146に依存してもよい。制御パラメータ146は、(エンコーダ100、170と類似の仕方で)予測器パラメータ520を使ってデコーダ500によって導出されてもよい。

20

30

【0146】

上記で示したように、受領されるビットストリームは、包絡データ161および利得データ162を含み、これは調整された包絡139を決定するために使用されうる。特に、包絡デコーダ503のユニット531は、包絡データ161から、量子化された現在の包絡134を決定するよう構成されていてもよい。例として、量子化された現在の包絡134は、(図3aに示されるように)あらかじめ定義された周波数帯域302における3dBの分解能を有していてもよい。量子化された現在の包絡134は、ブロックの集合132、332毎に(たとえば四つの符号化単位、すなわちブロック毎に、あるいは20ms毎に)、特にブロックのシフトされた集合332毎に更新されてもよい。量子化された現在の包絡134の周波数帯域302は、人間の聴覚の属性に適合するために、周波数の関数として、周波数ビン301の増大する数を有していてもよい。

40

【0147】

量子化された現在の包絡134は、ブロックのシフトされた集合332の(あるいは可能性としてはブロックの現在の集合132の)各ブロック131について、量子化された以前の包絡135から補間された包絡136に、線形に補間されてもよい。補間された包絡136は、量子化された3dB領域で決定されてもよい。これは、補間されたエネルギー値303が最も近い3dBレベルに丸められてもよいことを意味する。例示的な補間された

50

包絡136は図3aの点線のグラフによって示されている。各量子化された現在の包絡134について、四つのレベル補正利得a 137（包絡利得とも称される）が利得データ162として提供される。利得デコード・ユニット532は、利得データ162からレベル補正利得a 137を決定するよう構成されていてもよい。レベル補正利得は、1dBきざみで量子化されてもよい。各レベル補正利得は、種々のブロック131について調整された包絡139を提供するために対応する補間された包絡136に適用される。レベル補正利得137の増大した分解能のため、調整された包絡139は増大した分解能（たとえば1dB分解能）をもつことがある。

【0148】

10

図3bは、量子化された以前の包絡135と量子化された現在の包絡134との間の例示的な線形または幾何的補間を示している。包絡135、134は、対数スペクトルの平均レベル部分および形状部分に分離されてもよい。これらの部分は、線形、幾何的または調和的（並列な抵抗器）戦略のような独立な戦略を用いて補間されてもよい。よって、補間された包絡136を決定するために種々の補間方式が使用されうる。デコーダ500によって使われる補間方式は典型的には、エンコーダ100、170によって使われた補間方式に対応する。

【0149】

包絡デコーダ503の包絡洗練ユニット107は、調整された包絡139を（たとえば3dBきざみに）量子化することによって、調整された包絡139から割り当て包絡138を決定するよう構成されていてもよい。割り当て包絡138は、割り当て制御パラメータまたはオフセット・パラメータ（係数データ163内に含まれる）との関連で使われて、スペクトル・デコード、すなわち係数データ163のデコードを制御するために使用される名目整数割り当てベクトルを生成してもよい。特に、名目整数割り当てベクトルは、係数データ163内に含まれる量子化インデックスを逆量子化するための量子化器を決定するために使われてもよい。割り当て包絡138および名目整数割り当てベクトルは、エンコーダ100、170においてとデコーダ500においてとで類似の仕方決定されてもよい。

20

【0150】

図10は、割り当て包絡138に基づく例示的なビット割り当てプロセスを示している。上記で概説したように、割り当て包絡138は、所定の分解能（たとえば3dB分解能）に従って量子化されてもよい。割り当て包絡138の各量子化されたスペクトル・エネルギー値は対応する整数値に割り当てられてもよい。ここで、隣接する整数値は、所定の分解能（たとえば3dB分解能）に対応するスペクトル・エネルギーにおける差を表わしていてもよい。結果として得られる整数の集合は、整数割り当て包絡1004（iEnvと称する）と称されてもよい。整数割り当て包絡1004は、オフセット・パラメータによってオフセットされて、名目整数割り当てベクトル（iAllocと称される）を与えてもよい。このiAllocが、特定の周波数帯域302（周波数帯域インデックスbandIdxによって同定される）の係数を量子化するために使われるべき量子化器の直接的な指示を与える。

30

【0151】

40

図10は、描画1003において、整数割り当て包絡1004を周波数帯域302の関数として示している。周波数帯域1002（bandIdx=7）について、整数割り当て包絡1004が整数値-17を取ることが見て取れる（iEnv[7]=-17）。整数割り当て包絡1004は、ある最大値（iMaxと称される；たとえばiMax=-15）に制限されてもよい。ビット割り当てプロセスは、整数割り当て包絡1004およびオフセット・パラメータ（AllocOffsetと称される）の関数として量子化器インデックス1006（iAlloc[bandIdx]と称される）を与えるビット割り当て公式を利用してよい。上記で概説したように、オフセット・パラメータ（すなわち、AllocOffset）は対応するデコーダ500に伝送され、それにより、デコーダ500がビット割り当て公式を使って量子化器インデックス1006を決定できるようにする。ビット割り当て公式は

50

$iAlloc[bandIdx] = iEnv[bandIdx] - (iMax - CONSTANT_OFFSET) + AllocOffset$

によって与えられてもよい。ここで、CONSTANT_OFFSETは一定のオフセットであってもよく、たとえばCONSTANT_OFFSET=20である。例として、ビット割り当てプロセスが、ビットレート制約条件がオフセット・パラメータAllocOffset=-13を使って達成できると判定したとすると、七番目の周波数帯域の量子化器インデックス1007は $iAlloc[7] = -17 - (-15 - 20) - 13 = 5$ として得られうる。上述したビット割り当て公式をすべての周波数帯域302について使うことによって、すべての周波数帯域302についての量子化器インデックス1006（および結果として量子化器321、322、323）が決定される。0より小さい量子化器インデックスは量子化器インデックス0に丸められてもよい。同様に、最大の利用可能な量子化器インデックスより大きい量子化器インデックスは、最大の利用可能な量子化器インデックスまで丸められてもよい。

10

【0152】

さらに、図10は、本稿に記載される量子化方式を使って達成されうる例示的なノイズ包絡1011を示している。ノイズ包絡1011は、量子化の間に導入される量子化ノイズの包絡を示している。（図10において整数割り当て包絡1004によって表わされる）信号包絡と一緒にプロットされたら、ノイズ包絡1011は、量子化ノイズの分布が信号包絡に関して知覚的に最適化されているという事実を示す。

【0153】

デコーダ500が受領されたビットストリームと同期できるようにするために、種々の型のフレームが伝送されうる。フレームは、ブロックの集合132、332、特にブロックのシフトされたブロック332に対応しうる。特に、以前のフレームに対して相対的な仕方でエンコードされる、いわゆるPフレームが伝送されてもよい。上記において、デコーダ500は量子化された以前の包絡135を知っていることが想定された。量子化された以前の包絡135は以前のフレーム内で与えられてもよく、よって、現在の集合132または対応するシフトされた集合332がPフレームに対応しうる。しかしながら、スタートアップ・シナリオでは、デコーダ500は典型的には量子化された以前の包絡135を知らない。この目的のために、（たとえばスタートアップ時にまたは定期的に）Iフレームが伝送されてもよい。Iフレームは二つの包絡を含んでいてもよく、その一方が量子化された以前の包絡135として使われ、他方が量子化された現在の包絡134として使われる。Iフレームは、声スペクトル・フロントエンドの（すなわち、変換ベースの発話デコーダ500の）スタートアップの場合のために、たとえば異なるオーディオ符号化モードを用いるフレームに続くときに、および／またはオーディオ・ビットストリームの接合点を明示的に可能にするためのツールとして、使われてもよい。

20

30

【0154】

サブバンド予測器517の動作が図5dに示されている。図示した例では、予測器パラメータ520はラグ・パラメータおよび予測器利得パラメータgである。予測器パラメータ520は、ラグ・パラメータおよび予測器利得パラメータについての可能な値の所定のテーブルを使って、予測器データ164から決定されてもよい。これは、予測器パラメータ520のビットレート効率のよい伝送を可能にする。

【0155】

前記一つまたは複数の以前にデコードされた変換係数ベクトル（すなわち、再構成された係数の前記一つまたは複数の以前のブロック149）は、サブバンド（またはMDCT）信号バッファ541内に記憶されてもよい。バッファ541は、ストライドに従って（たとえば5ms毎に）更新されてもよい。予測器抽出器543は、規格化されたラグ・パラメータTに依存してバッファ541に対して作用するよう構成されていてもよい。規格化されたラグ・パラメータTは、ラグ・パラメータ520をストライド単位に（たとえばMDCTストライド単位に）規格化することによって決定されてもよい。ラグ・パラメータTが整数であれば、抽出器543は、T時間単位バッファ541にはいったところの一つまたは複数の以前にデコードされた変換係数ベクトルを取ってきてもよい。換言すれば、ラグ・パラメータTは、再構成された係数の前記一つまたは複数の以前のブロック149のうちの

40

50

どれが推定される変換係数のブロック150を決定するために使われるかを示してもよい。抽出器543の可能な実装に関する詳細な議論は、その内容が参照によって組み込まれる特許出願US61750052およびその優先権を主張する諸特許出願において提供されている。

【0156】

抽出器543は、フル信号包絡を担持するベクトル（またはブロック）に対して作用してもよい。他方、（サブバンド予測器517によって与えられる）推定された変換係数のブロック150は、平坦化された領域で表わされてもよい。結果的に、抽出器543の出力は、平坦化領域ベクトルに整形されてもよい。これは、再構成された係数の前記一つまたは複数の以前のブロック149の調整された包絡139を利用する整形器544を使って達成されてもよい。再構成された係数の前記一つまたは複数の以前のブロック149の調整された包絡139は、包絡バッファ542に記憶されていてもよい。整形器ユニット544は、 T_0 時間単位だけ包絡バッファ542にはいったところから平坦化において使われる遅延された信号包絡を取ってくるよう構成されていてもよい。ここで、 T_0 はTに最も近い整数である。次いで、平坦化領域ベクトルは、利得パラメータgによってスケールングされて、（平坦化領域での）推定された変換係数のブロック150を与えてもよい。

【0157】

代替として、平坦化領域で作用するサブバンド予測器517、たとえば再構成された平坦化された係数のブロック148に対して作用するサブバンド予測器517を使うことによって、整形器544によって実行される遅延された平坦化プロセスは省略されてもよい。しかしながら、平坦化領域ベクトル（またはブロック）のシーケンスが、変換（たとえばMDCT変換）の時間エイリアシングされた（time-aliased）諸側面のため、時間信号にはうまくマップしないことが見出されている。結果として、抽出器543の根底にある信号モデルへのフィットが低減され、より高いレベルの符号化ノイズがこの代替構成から帰結する。換言すれば、サブバンド予測器517によって使用される信号モデル（たとえば正弦波または周期的モデル）は、（平坦化された領域に比して）平坦化されない領域において増大したパフォーマンスを与えることが見出されている。

【0158】

ある代替的な例では、予測器517の出力（すなわち、推定された変換係数のブロック150）は、逆平坦化ユニット114の出力において（すなわち再構成された係数のブロック149に）加えられてもよいことを注意しておくべきである（図5a参照）。その場合、図5cの整形器ユニット544は、遅延された平坦化および逆平坦化の組み合わせされた動作を実行するよう構成されていてもよい。

【0159】

受領されたビットストリーム中の要素が、サブバンド・バッファ541および包絡バッファ541を、たとえばIフレームの最初の符号化単位（すなわち、最初のブロック）の場合に、時折フラッシュすることを制御してもよい。これは、以前のデータを知ることなくIフレームをデコードすることを可能にする。最初の符号化単位は典型的には予測寄与を利用できないが、それでも予測器情報520を伝達するために相対的により少数のビットを使ってもよい。予測利得の喪失は、この最初の符号化単位の予測誤差符号化により多くのビットを割り当てることによって補償されてもよい。典型的には、予測器寄与はIフレームの第二の符号化単位（すなわち第二のブロック）についてやはり実質的である。これらの側面のため、たとえIフレームを非常に頻繁に使ったとしても、比較的小さなビットレート増で品質を維持できる。

【0160】

換言すれば、ブロックの集合132, 332（フレームとも称される）は、予測符号化を使ってエンコードされうる複数のブロック131を含む。Iフレームをエンコードするとき、ブロックの集合332の最初のブロック203だけは、予測エンコーダによって達成される符号化利得を使ってエンコードされることができない。すでに直後のブロック201は予測エンコードの恩恵を利用しうる。つまり、符号化効率に関するIフレームの欠

10

20

30

40

50

点は、フレーム 3 3 2 の変換係数の最初のブロック 2 0 3 のエンコードに限定され、フレーム 3 3 2 の他のブロック 2 0 1、2 0 4、2 0 5 には当てはまらないということである。よって、本稿に記載される変換ベースの発話符号化方式は、符号化効率に対する著しい影響なしに、1 フレームの比較的頻繁な使用を許容する。よって、本稿に記載される変換ベースの発話符号化方式は、比較的高速および／またはデコーダとエンコーダの間の比較的頻繁な同期を必要とする用途に特に好適である。

【0 1 6 1】

図 5 d は、例示的なスペクトル・デコーダ 5 0 2 のブロック図を示している。スペクトル・デコーダ 5 0 2 は、エントロピー符号化された係数データ 1 6 3 をデコードするよう構成されている無損失デコーダ 5 5 1 を有する。さらに、スペクトル・デコーダ 5 0 2 は、係数データ 1 6 3 内に含まれる量子化インデックスに係数値を割り当てるよう構成されている逆量子化器 5 5 2 を有する。エンコーダ 1 0 0、1 7 0 のコンテキストにおいて概説したように、所定の量子化器の集合、たとえばモデル・ベースのスカラー量子化器の有限な集合から選択される異なる量子化器を使って異なる変換係数が量子化されてもよい。図 4 に示されるように、量子化器 3 2 1、3 2 2、3 2 3 の集合は、量子化器の種々の型を含んでいてもよい。量子化器の集合は、(0 ビットレートの場合) ノイズ合成を提供する量子化器 3 2 1、(比較的低い信号対雑音比 SNR のためおよび中間的なビットレートのための) 一つまたは複数のディザリングされる量子化器 3 2 2 および／または(比較的高い SNR および比較的高いビットレートのための) 一つまたは複数の普通の量子化器 3 2 3 を含んでいてもよい。

【0 1 6 2】

包絡洗練ユニット 1 0 7 は、割り当てベクトルを与えるために係数データ 1 6 3 内に含まれるオフセット・パラメータと組み合わせられてもよい割り当て包絡 1 3 8 を提供するように構成されていてもよい。割り当てベクトルは、各周波数帯域 3 0 2 について整数値を含む。特定の周波数帯域 3 0 2 についての整数値は、特定の周波数帯域 3 0 2 の変換係数の逆量子化のために使われるべきレート - 歪み点を指す。換言すれば、特定の周波数帯域 3 0 2 についての整数値は、特定の周波数帯域 3 0 2 の変換係数の逆量子化のために使われるべき量子化器を指す。整数値が 1 増すことは、SNR における 1.5 dB の増加に対応する。ディザリングされる量子化器 3 2 2 および普通の量子化器 3 2 3 について、ラプラシアン確率分布モデルが、算術符号化を用いてもよい無損失符号化において使われてもよい。低ビットレートと高ビットレートの場合の間にシームレスな仕方でギャップを埋めるために、一つまたは複数のディザリングされる量子化器 3 2 2 が使われてもよい。ディザリングされる量子化器 3 2 2 は、静的なノイズ様信号について十分になめらかな出力オーディオ品質を生成することにおいて有益でありうる。

【0 1 6 3】

換言すれば、逆量子化器 5 2 2 は、変換係数の現在のブロック 1 3 1 の係数量子化インデックスを受領するよう構成されていてもよい。特定の周波数帯域 3 0 2 の前記一つまたは複数の係数量子化インデックスは、所定の一组の量子化器からの対応する量子化器を使って決定されている。特定の周波数帯域 3 0 2 についての(オフセット・パラメータを用いて割り当て包絡 1 3 8 をオフセットすることによって決定されうる)割り当てベクトルの値は、特定の周波数帯域 3 0 2 の前記一つまたは複数の係数量子化インデックスを決定するために使われた量子化器を示す。量子化器を同定したら、前記一つまたは複数の係数量子化インデックスは、逆量子化されて、量子化された誤差係数のブロック 1 4 5 を与えてもよい。

【0 1 6 4】

さらに、スペクトル・デコーダ 5 0 2 は、スケーリングされた量子化された誤差係数のブロック 1 4 7 を提供する逆再スケーリング・ユニット 1 1 3 を有していてもよい。図 5 d の無損失デコーダ 5 5 1 および逆量子化器 5 5 2 のまわりの追加的なツールおよび相互接続は、図 5 a に示される全体的なデコーダ 5 0 0 におけるその使用にスペクトル・デコードを適応させるために使われてもよい。ここで、スペクトル・デコーダ 5 0 2 の出力 (

すなわち量子化された誤差係数のブロック145)は、予測された平坦化領域ベクトルに(すなわち、推定された変換係数のブロック150に)加算的補正を提供するために使われる。特に、追加的なツールは、デコーダ500によって実行される処理がエンコーダ100、170によって実行された処理に対応することを保証してもよい。

【0165】

特に、スペクトル・デコーダ502は、ヒューリスティック・スケーリング・ユニット111を有していてもよい。エンコーダ100、170との関連で示したように、ヒューリスティック・スケーリング・ユニット111はビット割り当てに対する影響をもつことがある。エンコーダ100、170では、予測誤差係数の現在ブロック141が、ヒューリスティック規則によって分散1までスケール・アップされてもよい。結果として、デフォルトの割り当ては、ヒューリスティック・スケーリング・ユニット111の最終的なダウンスケーリングされた出力の細かすぎる量子化につながることもある。よって、割り当ては、予測誤差係数の修正と同様の仕方で修正されるべきである。

10

【0166】

しかしながら、下記で概説するように、低周波数ビン(または低周波数帯域)の一つまたは複数についての符号化資源の低減を避けることが有益であることがある。特に、これは、実は有声状況において(すなわち、比較的大きな制御パラメータ146, rfuをもつ信号について)最も顕著であるLF(低周波数)ランブル/ノイズ・アーチファクトに対抗するために有益であることがある。よって、後述する制御パラメータ146に依存したビット割り当て/量子化器選択は、「有声適応LF品質ブースト」と考えられてもよい。

20

【0167】

スペクトル・デコーダは、rfuと名付けられる制御パラメータ146に依存してもよい。rfuは予測器利得gの制限されたバージョンであってもよく、たとえば

$$rfu = \min(1, \max(g, 0))$$

である。制御パラメータ146 rfuを決定するための代替的な諸方法が使われてもよい。特に、制御パラメータ146は、表1において与えられる擬似コードを使って決定されてもよい。

【0168】

【表 1】

```

f_gain = f_pred_gain;

if (f_gain < -1.0)

    f_rfu = 1.0;

else if (f_gain < 0.0)

    f_rfu = -f_gain;

else if (f_gain < 1.0)

    f_rfu = f_gain;

else if (f_gain < 2.0)

    f_rfu = 2.0 - f_gain;

else // f_gain >= 2.0

    f_rfu = 0.0.

```

10

20

表 1

変数f_gainおよびf_predは等しく設定されてもよい。特に変数f_gainは予測器利得gに対応してもよい。制御パラメータ146 rfuは表1ではf_rfuとして言及されている。利得f_gainは実数であってもよい。

【0169】

制御パラメータ146の最初の定義に比較して、(表1による)後者の定義は、1より大きな予測器利得について制御パラメータ146 rfuを低減し、負の予測器利得については制御パラメータ146 rfuを増大させる。

30

【0170】

制御パラメータ146を使って、エンコーダ100、170の係数量子化ユニット112において使われるおよび逆量子化器552において使われる量子化器の集合が適応されてもよい。特に、量子化器の集合のノイズ性が、制御パラメータ146に基づいて適応されてもよい。例として、1に近い制御パラメータ146 rfuの値は、ディザリングされる量子化器を使って割り当てレベルの範囲の制限をトリガーしてもよく、ノイズ合成レベルの分散の低減をトリガーしてもよい。一例では、rfu=0.75におけるディザ決定閾値および1-rfuに等しいノイズ利得が設定されてもよい。ディザ適応は、無損失デコードおよび逆量子化器の両方に影響しうる一方、ノイズ利得適応は典型的には逆量子化器のみに影響する。

40

【0171】

予測器寄与は有声／トーン性状況について実質的であることが想定されてもよい。よって、比較的高い予測器利得g(すなわち、比較的高い制御パラメータ146)は有声またはトーン性の発話信号を示していてもよい。そのような状況では、ディザに関係したまたは明示的な(0割り当ての場合)ノイズの追加は、経験的に、エンコードされた信号の知覚される品質に対して逆効果であることが示されている。結果として、ディザリングされる量子化器322の数および／またはノイズ合成量子化器321のために使われるノイズの型は、予測器利得gに基づいて適応され、それによりエンコードされた発話信号の知覚

50

される品質を改善してもよい。

【0172】

よって、制御パラメータ146は、ディザリングされる量子化器322が使用されるSNRの範囲324、325を修正するために使われてもよい。例として、制御パラメータ146 $\text{rfu} < 0.75$ である場合には、ディザリングされる量子化器の範囲324が使われてもよい。換言すれば、制御パラメータ146が所定の閾値より下であれば、量子化器の第一の集合326が使用されてもよい。他方、制御パラメータ146 $\text{rfu} \geq 0.75$ であれば、ディザリングされる量子化器のための範囲325が使われてもよい。換言すれば、制御パラメータ146が前記所定の閾値以上であれば、量子化器の第二の集合327が使用されてもよい。

10

【0173】

さらに、制御パラメータ146は、分散およびビット割り当ての修正のために使われてもよい。その理由は、典型的には、うまくいった予測では必要とされる補正も小さく、特に0~1kHzの低周波数範囲ではそうであるということである。より高い周波数帯域302に符号化資源を解放するために、単位分散モデルからのこの逸脱を量子化器に明示的に知らせることが有利であることがありうる。このことは、その内容が参照によって組み込まれるW02009/086918の図17cのパネルiiiのコンテキストにおいて記述されている。デコーダ500では、この修正は、(スケーリング・ユニット111を使うこと)によって適用される) ヒューリスティック・スケーリング規則に従って名目割り当てベクトルを修正し、同時に、逆スケーリング・ユニット113を使って逆ヒューリスティック・スケーリング規則に従って逆量子化器552の出力をスケーリングすることによって実装されてもよい。W02009/086918の理論に従い、ヒューリスティック・スケーリング規則および逆ヒューリスティック・スケーリング規則は緊密にマッチされるべきである。しかしながら、有声の信号成分についてLF(低周波数)ノイズに関わる時折の問題に対抗するために、一つまたは複数の最低周波数帯域302については割り当て修正を打ち消すことが有利であることが経験的に見出されている。割り当て修正の打ち消しは、予測器利得gおよび/または制御パラメータ146の値に依存して実行されてもよい。特に、割り当て修正の打ち消しは、制御パラメータ146がディザ決定閾値を超える場合にのみ実行されてもよい。

20

30

【0174】

よって、本稿は、量子化器の集合326の組成(たとえば、ディザリングされない量子化器323の数および/またはディザリングされる量子化器322の数)を、エンコーダ100、170および対応するデコーダ500において利用可能なサイド情報(たとえば制御パラメータ146)に基づいて調整する手段を記述する。量子化器の集合326の組成は、予測器利得gの存在において(たとえば、制御パラメータ146に基づいて)調整されてもよい。特に、予測器利得gが比較的低い場合、ディザリングされる量子化器322の数 N_{dith} が増やされ、ディザリングされない量子化器323の数 N_{cq} が減らされてもよい。さらに、割り当てられたビットの数は、比較的小さい量子化器を選択することによって低減されてもよい。他方、予測器利得gが比較的大きい場合には、ディザリングされる量子化器322の数 N_{dith} が減らされてもよく、ディザリングされる量子化器323の数 N_{cq} が増やされてもよい。さらに、割り当てられたビットの数は、比較的小さい量子化器を選択することによって低減されてもよい。

40

【0175】

代替的または追加的に、入力信号の現在の抜粋のヒス様(hiss-like)属性を示すスペクトル反射係数(spectral reflection coefficient) R_{fc} を決定するための例示的な方式が記述される。スペクトル反射係数 R_{fc} は自己回帰源モデリングのコンテキストにおいて使われる「反射係数」とは異なることを注意しておくべきである。変換係数のブロック131は、L個の周波数帯域302に分割されてもよい。L次元ベクトル B_{w} が定義されてもよい。ここで、ベクトル B_{w} のl番目の要素は、l番目の周波数帯域302 ($l=1, \dots, L$)に

50

属する変換ビン301の数に等しくてもよい。同様に、K次元ベクトルFが定義されてもよい。ここで、1番目の要素は、1番目の周波数帯域302の中心に等しくてもよく、これは、1番目の周波数帯域302に属する、変換ビン301の最小インデックスと変換ビン301の最大インデックスの平均を計算することによって得られる。さらに、L次元ベクトル S_{PSD} が定義されてもよい。ここで、ベクトル S_{PSD} は、信号のパワースペクトル密度の値を含んでいてもよく、これはdBスケールからの包絡に關係した量子化インデックスをもとの線形スケールに変換することによって得られてもよい。さらに、L番目の周波数帯域302に属する最大のビン・インデックスである最大ビン・インデックス N_{core} が定義されてもよい。スカラー反射係数Rfcは次のように決定されてもよい。

【0176】

【数7】

$$Rfc = \frac{\sum_{l=1}^L \frac{-B_w(l)S_{\text{PSD}}(l)\cos\left(\frac{\pi F(l)}{N_{\text{core}}}\right)}{\sum_{l=1}^L B_w(l)S_{\text{PSD}}(l)}}$$

ここで、lはL次元ベクトルのl番目の要素を表わす。

【0177】

一般に、 $Rfc > 0$ は、高周波数部分によって支配されているスペクトルを示し、 $Rfc < 0$ は低周波数部分によって支配されているスペクトルを示す。 Rfc パラメータは、次のように使われてもよい： Rfc 値が低く（すなわち、予測利得が低い場合）、 $Rfc > 0$ である場合には、これは摩擦音（すなわち、無声の歯擦音）に対応するスペクトルを示す。この場合、ディザリングされる量子化器322の相対的に増大した数 N_{dith} が量子化器の集合326、722内で使われてもよい。

【0178】

一般的な言い方では、量子化器（および対応する逆量子化器）の集合326は、エンコーダ100および対応するデコーダ500において利用可能なサイド情報（たとえば、制御パラメータ146および／またはスペクトル反射係数）に基づいて調整されてもよい。サイド情報は、エンコーダ100およびデコーダ500に利用可能なパラメータから抽出されてもよい。上記で概説したように、予測器利得gは、デコーダ500に伝送されてもよく、変換係数の逆量子化前に、逆量子化器の適切な集合326を選択するために使われることができる。代替的または追加的に、反射係数は、デコーダ500に伝送されるスペクトル包絡に基づいて推定または近似されてもよい。

【0179】

図7は、エンコーダ100および対応するデコーダ500において量子化器／逆量子化器の集合326を決定するための例示的な方法のブロック図を示している。関連するサイド情報721（予測器パラメータgおよび／または反射係数など）は、ビットストリームから抽出701されてもよい。サイド情報721は、現在ブロック係数を量子化するためおよび／または対応する量子化インデックスを逆量子化するために使われるべき量子化器の集合722を決定702するために使われてもよい。レート割り当てプロセス703を使って、量子化器の決定された集合722からの特定の量子化器が、特定の周波数帯域302の係数を量子化するためおよび／または対応する量子化インデックスを逆量子化するために使われる。ビット割り当てプロセス703から帰結する量子化器選択723は、量子化インデックスを与えるために量子化プロセス703内で使われるおよび／または量子化された係数を与えるために逆量子化プロセス713内で使われる。

【0180】

図9のaないしcは、本稿に記述される変換ベースのコーデック・システムを使って達成されうる例示的な実験結果を示している。特に、図9のaないしcは、一つまたは複数のディザリングされる量子化器322を含む量子化器の順序付けされた集合326を使う

10

20

30

40

50

ことの恩恵を示している。図9のaは、もとの信号のスペクトログラム901を示している。スペクトログラム901が、白丸によって同定される周波数範囲内にスペクトル内容を有することが見て取れる。図9のbは、もとの信号の量子化されたバージョン（22kpsで量子化）のスペクトログラム902を示している。図9のbの場合、0レート割り当てのためのノイズ充填およびスカラー量子化器が使われた。スペクトログラム902は、浅いスペクトルの穴（いわゆる「バーディー」）に関連付けられている白丸によって同定される周波数範囲における比較的大きなスペクトル・ブロックを示す。これらのブロックは典型的には可聴アーチファクトにつながる。図9のcは、もとの信号のもう一つの量子化されたバージョン（22kpsで量子化）のスペクトログラム903を示している。図9のcの場合には、0レート割り当てのためのノイズ充填、ディザリングされる量子化器およびスカラー量子化器が使われた（本稿に記載されるように）。スペクトログラム903は、白丸によって同定される周波数範囲内にスペクトルの穴に関連付けられた大きなスペクトル・ブロックを示さないことが見て取れる。当業者にはわかることだが、そのような量子化ブロックの不在は、本稿に記載される変換ベースのコーデック・システムの改善された知覚的パフォーマンスを示す。

10

【0181】

以下では、エンコーダ100、170および／またはデコーダ500のさまざまな追加的な側面が記述される。上記で概説したように、エンコーダ100、170および／またはデコーダ500は予測誤差係数 $\Delta(k)$ を再スケーリングして再スケーリングされた誤差係数のブロック142を与えるよう構成されたスケーリング・ユニット111を有していてもよい。再スケーリング・ユニット111は、再スケーリングを実行するために一つまたは複数の所定のヒューリスティック規則を利用してもよい。一例では、再スケーリング・ユニット111は、利得 $d(f)$ 、たとえば

20

$$d(f) = 1 + \frac{7 \cdot rfu^2}{1 + \left(\frac{f}{f_0}\right)^3}$$

30

を含むヒューリスティック・スケーリング規則を利用してもよい。ここで、ブレイク周波数 f_0 はたとえば1000Hzに設定されてもよい。よって、再スケーリング・ユニット111は、予測誤差係数に周波数依存の利得 $d(f)$ を適用して再スケーリングされた誤差係数のブロック142を与えるよう構成されていてもよい。逆再スケーリング・ユニット113は、周波数依存の利得 $d(f)$ の逆を適用するよう構成されていてもよい。周波数依存の利得 $d(f)$ は、制御パラメータ rfu 146に依存していてもよい。上記の例において、利得 $d(f)$ は低域通過特性を示し、よって予測誤差係数は、低周波数より高周波数においてより減衰されるおよび／または予測誤差係数は高周波数より低周波数においてより強調される。上述した利得 $d(f)$ は常に1以上である。よって、ある好ましい実施形態では、ヒューリスティック・スケーリング規則は、予測誤差係数が（周波数に依存して）因数1によってまたはそれ以上強調されるというものである。

40

【0182】

周波数依存利得がパワーまたは分散を示していよいことを注意すべきである。そのような場合、スケーリング規則および逆スケーリング規則は、周波数依存の利得の平方根に基づいて、たとえば $\sqrt{d(f)}$ に基づいて導出されるべきである。

【0183】

強調および／または減衰の度合いは、予測器117によって達成される予測の品質に依存してもよい。予測器利得 g および／または制御パラメータ rfu 146は、予測の品質を示していてもよい。特に、制御パラメータ rfu 146の比較的低い値（比較的0に近い）

50

は、予測の低い品質を示しうる。そのような場合、予測誤差係数がすべての周波数にわたって比較的高い（絶対的な）値を有することが期待される。制御パラメータ $r_{fu} \ 146$ の比較的高い値（比較的1に近い）は、予測の高い品質を示しうる。そのような場合、予測誤差係数が（予測がより難しい）高周波数について比較的高い（絶対的な）値を有することが期待される。よって、再スケーリング・ユニット 111 の出力における単位分散を達成するためには、利得 $d(f)$ は、予測の比較的低い品質の場合に利得 $d(f)$ がすべての周波数について実質的に平坦であり、予測の比較的高い品質の場合には利得 $d(f)$ は低域通過特性をもち、低周波数での分散を増大またはブーストするようなものであってもよい。これは、上述した r_{fu} 依存の利得 $d(f)$ について当てはまる。

10

【0184】

上記で概説したように、ビット割り当てユニット 110 は、割り当て包絡 138 における対応するエネルギー値に依存して、異なる再スケーリングされた誤差係数にビットの相対的な割り当てを提供するよう構成されていてもよい。ビット割り当てユニット 110 は、ヒューリスティック再スケーリング規則を考慮に入れるよう構成されていてもよい。ヒューリスティック再スケーリング規則は、予測の品質に依存してもよい。予測の比較的高い品質の場合には、低周波数での係数のエンコードよりも、高周波数での予測誤差係数（または再スケーリングされた誤差係数のブロック 142 ）のエンコードに、相対的に増大した数のビットを割り当てることが有益でありうる。これは、予測の高い品質の場合、低周波数係数はすでによく予測されており、一方、高周波数係数は典型的にはそれほどよく

20

【0185】

上記の振る舞いは、予測の品質を考慮に入れる割り当て包絡 138 を決定するために、現在の調整された包絡 139 にヒューリスティック規則／利得 $d(f)$ の逆を適用することによって実装されうる。

【0186】

調整された包絡 139 、予測誤差係数および利得 $d(f)$ は、対数またはdB領域で表わされてもよい。そのような場合、予測誤差係数への利得 $d(f)$ の適用は、「加算」演算に対応してもよく、調整された包絡 139 への利得 $d(f)$ の逆の適用は、「減算」演算に対応してもよい。

30

【0187】

ヒューリスティック規則／利得 $d(f)$ のさまざまな変形が可能であることを注意しておくべきである。特に、低域通過特性の固定した周波数依存曲線 $(1 + (f/f_0)^3)^{-1}$ は、包絡データに（たとえば現在ブロック 131 についての調整された包絡 139 に）依存する関数によって置き換えられてもよい。修正されたヒューリスティック規則は、制御パラメータ $r_{fu} \ 146$ および包絡データの両方に依存してもよい。

【0188】

以下では、予測器利得 g に対応しうる予測器利得 ρ を決定するための種々の方法が記述される。予測器利得 ρ は、予測の品質の指示として使われてもよい。予測残差ベクトル（すなわち、予測誤差係数のブロック 141 ） $\underline{z}$ は、 $\underline{z} = \underline{x} - \rho \underline{y}$ によって与えられてもよい。ここで、 $\underline{x}$ はターゲット・ベクトル（たとえば、平坦化された変換係数の現在のブロック 140 または変換係数の現在のブロック 131 ）であり、 $\underline{y}$ は予測のための選ばれた候補を表わすベクトル（たとえば再構成された係数の以前のブロック 149 ）であり、 ρ は（スカラー）予測器利得である。

40

【0189】

$w \geq 0$ が予測器利得 ρ の決定のために使われる重みベクトルであってもよい。いくつかの実施形態では、重みベクトルは信号包絡の関数（たとえば、エンコーダ 100 、 170 において推定され、次いでデコーダ 500 に伝送されてもよい、調整された包絡 139 の関

50

数)である。重みベクトルは典型的には、ターゲット・ベクトルおよび候補ベクトルと同じ次元をもつ。ベクトル x の i 番目の要素は x_i によって表わされてもよい(たとえば $i=1, \dots, K$)。

【0190】

予測器利得 ρ を定義するための種々の仕方がある。ある実施形態では、予測器利得 ρ は、最小平均平方誤差基準に従って定義されるMMSE(最小平均平方誤差)利得である。この場合、予測器利得 ρ は次の公式を使って計算されてもよい。

【0191】

【数9】

10

$$\rho = \frac{\sum_i x_i y_i}{\sum_i y_i^2}$$

そのような予測器利得 ρ は典型的には

【数10】

$$D = \sum_i (x_i - \rho y_i)^2$$

20

として定義される平均平方誤差を最小化する。

【0192】

平均平方誤差 D の定義に重み付けを導入することがしばしば(知覚上)有益である。重み付けは、信号スペクトルの知覚的に重要な部分については x と y の間のマッチの重要性を強調し、比較的重要でない信号スペクトルの部分については x と y の間のマッチの重要性を脱強調するために使われてもよい。そのようなアプローチは、次のような誤差基準を与える：

30

【数11】

$$D = \sum_i (x_i - \rho y_i)^2 w_i$$

これは(重み付けされた平均平方誤差の意味での)最適予測器利得の次の定義につながる：

【数12】

$$\rho = \frac{\sum_i w_i x_i y_i}{\sum_i w_i y_i^2}$$

40

予測器利得の上記の定義は典型的には、制限されない利得を与える。上記で示したように、重みベクトル w の重み w_i は調整された包絡139に基づいて決定されてもよい。たとえば、重みベクトル w は、調整された包絡139のあらかじめ定義された関数を使って決定されてもよい。あらかじめ定義された関数は、エンコーダおよびデコーダにおいて既知であってもよい(これは調整された包絡139についても成り立つ)。よって、重みベクトルは、エンコーダおよびデコーダにおいて同じ仕方で決定されうる。

50

【0193】

もう一つの可能な予測器利得公式は次式によって与えられる：

【数13】

$$\rho = \frac{2C}{E_x + E_y}$$

$$\text{ここで } C = \sum_i w_i x_i y_i, \quad E_x = \sum_i w_i x_i^2, \quad E_y = \sum_i w_i y_i^2$$

予測器利得のこの定義は、常に区間 $[-1,1]$ 内である利得を与える。この公式によって指定される予測器利得の重要な特徴は、予測器利得 ρ がターゲット信号のエネルギー x と残差信号のエネルギー z の間の扱える関係を容易にすることである。LTP残差エネルギーは、

【数14】

$$\sum_i w_i z_i^2 = E_x (1 - \rho^2)$$

と表わされてもよい。

【0194】

制御パラメータrfu 146は、上述した公式を使って予測器利得 g に基づいて決定されてもよい。予測器利得 g は、上述した公式の任意のものを使って決定される予測器利得 ρ に等しくてもよい。

【0195】

上記で概説したように、エンコーダ100、170は、残差ベクトル z （すなわち予測誤差係数のブロック141）を量子化し、エンコードするよう構成されている。量子化プロセスは典型的には信号包絡によって（たとえば割り当て包絡138によって）、根底にある知覚モデルに従って、利用可能なビットを知覚的に意味のある仕方で信号のスペクトル成分の間で分配するために、案内される。レート割り当てのプロセスは、入力信号から（たとえば変換係数のブロック131から）導出される信号包絡によって（たとえば割り当て包絡138によって）案内される。予測器117の動作は典型的には信号包絡を変える。量子化ユニット112は典型的には、単位分散源に対する作用を想定して設計される量子化器を利用する。特に、高品質予測の場合（すなわち、予測器117がうまくいっているとき）、単位分散属性はもはや成り立たないことがあり、すなわち、予測誤差係数のブロック141は単位分散を示さないことがある。

【0196】

予測誤差係数のブロック141の（すなわち残差 z についての）包絡を推定し、この包絡をデコーダに伝送する（そして推定された包絡を使って予測誤差係数のブロック141を再平坦化する）ことは典型的には効率的ではない。その代わりに、エンコーダ100およびデコーダ500は、（上記で概説したように）予測誤差係数のブロック141を再スケーリングするためのヒューリスティック規則を利用してもよい。ヒューリスティック規則は、予測誤差係数のブロック141を再スケーリングするために使われてもよい。それにより、再スケーリングされた係数のブロック142は単位分散に近づく。この結果として、（単位分散を想定する量子化器を使って）量子化結果は改善されうる。

【0197】

さらに、すでに概説したように、ヒューリスティック規則は、ビット割り当てプロセスのために使われる割り当て包絡138を修正するために使われてもよい。割り当て包絡138の修正および予測誤差係数のブロック141の再スケーリングは、典型的にはエンコ

10

20

30

40

50

ーダ100およびデコーダ500によって同じ仕方で（同じヒューリスティック規則を使って）実行される。

【0198】

可能なヒューリスティック規則 $d(f)$ が上記で記載された。以下では、ヒューリスティック規則を決定するための別のアプローチが記載される。重み付けされた領域のエネルギー予測利得の逆が、 $\|z\|_w^2 = p \|x\|_w^2$ となるように $p \in [0, 1]$ によって与えられてもよい。ここで、 $\|z\|_w^2$ は、重み付け領域における残差ベクトル（すなわち、予測誤差係数のブロック141）の平方エネルギーを示し、 $\|x\|_w^2$ は、重み付け領域におけるターゲット・ベクトル（すなわち、平坦化された変換係数のブロック140）の平方エネルギーを示す。

10

【0199】

以下の想定がなされてもよい。

1. ターゲット・ベクトル x の要素は単位分散をもつ。これは、平坦化ユニット108によって実行される平坦化の結果であってもよい。この想定は、平坦化ユニット108によって実行される包絡ベースの平坦化の品質に依存して充足される。
2. 予測残差ベクトル z の要素の分散は、 $i=1, \dots, K$ および何らかの $t \geq 0$ について、 $E\{z^2(i)\} = \min\{t/w(i), 1\}$ の形である。この想定は、最小二乗指向の予測器探索は重み付け領域において均等に分布した誤差寄与につながり、残差ベクトル $(\sqrt{w})z$ は多少なりとも平坦になるというヒューリスティックに基づいている。さらに、予測器候補は平坦に近いことが期待されてもよく、これは合理的な限界 $E\{z^2(i)\} \leq 1$ につながる。この第二の想定のみ

20

【0200】

パラメータ t を推定するために、上述した二つの想定を予測誤差公式（たとえば $D = \sum_i (x_i - p y_i)^2 w_i$ ）に挿入し、それにより「水位型」の次式を与えてもよい。

【0201】

【数15】

$$\sum_i \min\{t, w(i)\} = p \sum_i w(i)$$

30

上記の式には区間 $t \in [0, \max(w(i))]$ 内に解があることを示すことができる。パラメータ t を見出すための方程式は、ソーティング・ルーチンを使って解くことができる。

【0202】

すると、ヒューリスティック規則は $d(i) = \max\{w(i)/t, 1\}$ によって与えられてもよい。ここで、 $i=1, \dots, K$ は周波数ビンを同定する。ヒューリスティック・スケーリング規則の逆が、 $1/d(i) = \min\{t/w(i), 1\}$ によって与えられる。ヒューリスティック・スケーリング規則の逆は、逆再スケーリング・ユニット113によって適用される。周波数依存のスケーリング規則は重み $w(i) = w_i$ に依存する。上記で示したように、重み $w(i)$ は変換係数の現在ブロック131（または調整された包絡139または該調整された包絡139の何らかのあらかじめ定義された関数）に依存していてもよく、あるいはそれに対応していてもよい。

40

【0203】

予測器利得を決定するために公式 $p = 2C/\{E_x + E_y\}$ を使うとき、関係 $p = 1 - \rho^2$ が成り立つことが示せる。

【0204】

よって、ヒューリスティック・スケーリング規則がさまざまな異なる仕方で決定されてもよい。実験的に、上述した二つの想定に基づいて決定されるスケーリング規則（スケーリング方法Bと称される）が固定したスケーリング規則 $d(f)$ に比べて有利であることが示

50

されている。特に、上記二つの想定に基づいて決定されるスケーリング規則は、予測器候補探索の過程で使われる重み付けの効果を考慮に入れてもよい。残差の分散と信号の分散の間に解析的に扱える関係（これは上記で概説したように p の導出を容易にする）のため、スケーリング方法Bは、利得の定義 $\rho = 2C/\{E_x + E_y\}$ と便利に組み合わせられる。

【0205】

以下では、変換ベースのオーディオ符号化器のパフォーマンスを改善するためのさらなる側面が記述される。特に、いわゆる分散保存フラグの使用が提案される。分散保存フラグは、ブロック131毎に決定され、伝送されてもよい。分散保存フラグは、予測の品質を示していてもよい。ある実施形態では、予測の比較的高い品質の場合には分散保存フラグはオフであり、予測の比較的低い品質の場合には分散保存フラグはオンである。分散保存フラグは、エンコーダ100、170によって、たとえば予測器利得 ρ に基づいておよび／または予測器利得 g に基づいて決定されてもよい。例として、分散保存フラグは、予測器利得 ρ または g （またはそれから導出されるパラメータ）が所定の閾値（たとえば2dB）より低い場合に「オン」に設定されてもよい。逆もまたしかりである。上記で概説したように、重み付け領域のエネルギー予測利得の逆 p は、典型的には予測器利得に依存する。たとえば、 $p = 1 - \rho^2$ である。パラメータ p の逆数は、分散保存フラグの値を決定するために使われてもよい。例として、 $1/p$ （たとえばdBで表わされる）は、分散保存フラグの値を決定するために、所定の閾値（たとえば2dB）と比較されてもよい。 $1/p$ が該所定の閾値より大きければ、分散保存フラグは「オフ」に設定されてもよい（予測の比較的高い品質を示す）。逆もまたしかりである。

【0206】

分散保存フラグは、エンコーダ100およびデコーダ500のさまざまな異なる設定を制御するために使われてもよい。特に、分散保存フラグは、複数の量子化器321、322、323のノイズ性の度合いを制御するために使われてもよい。特に、分散保存フラグは、次の設定のうちの一つまたは複数の影響してもよい。

- ・0ビット割り当てのための適応的なノイズ利得。換言すれば、ノイズ合成量子化器321のノイズ利得は分散保存フラグによって影響されてもよい。
- ・ディザリングされる量子化器の範囲。換言すれば、ディザリングされる量子化器322が使われるSNRの範囲324、325が、分散保存フラグによって影響されてもよい。
- ・ディザリングされる量子化器の事後利得。ディザリングされる量子化器の平均平方誤差パフォーマンスに影響するために、ディザリングされる量子化器の出力に対して事後利得が適用されてもよい。事後利得は、分散保存フラグに依存してもよい。
- ・ヒューリスティック・スケーリングの適用。（再スケーリング・ユニット111および逆再スケーリング・ユニット113における）ヒューリスティック・スケーリングの使用が分散保存フラグに依存してもよい。

【0207】

分散保存フラグがエンコーダ100および／またはデコーダ500の一つまたは複数の設定をどのように変えうるかの例を表2に与えておく。

【0208】

【表 2】

設定型	分散保存フラグ	分散保存フラグ
ノイズ利得	$g_N = (1 - rfu)$	$g_N = \sqrt{(1 - rfu^2)}$
ディザリングされる 量子化器の範囲	制御パラメータ rfu に依存	比較的大きな範囲に(たとえば最大可能な範囲に)固定
ディザリングされる量子化 器の事後利得 $\gamma_o = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}; \gamma_1 = \sqrt{\gamma_o}$	$\gamma = \gamma_o$	$\gamma = \max(\gamma_o, g_N \cdot \gamma_1)$
ヒュースティック・スケールリング規則	オン	オフ

表 2

事後利得についての公式において、 $\sigma_x = E\{X^2\}$ は(量子化されるべき)予測誤差係数のブロック 1 4 1 の係数のうち一つまたは複数の係数の分散であり、 Δ は事後利得が適用されるディザリングされる量子化器のスカラー量子化器(6 1 2)の量子化器きざみサイズである。

【0 2 0 9】

表 2 の例から見て取れるように、ノイズ合成量子化器 3 2 1 のノイズ利得 g_N (すなわち、ノイズ合成量子化器 3 2 1 の分散)は分散保存フラグに依存してもよい。上記で概説したように、制御パラメータ rfu 1 4 6 は範囲[0,1]内にあってもよく、rfuの比較的低い値は予測の比較的低い品質を示し、rfuの比較的高い値は予測の比較的高い品質を示す。[0,1]の範囲内の rfu 値について、左の列の公式は右の列の公式より低いノイズ利得 g_N を与える。よって、分散保存フラグがオンであるとき(予測の比較的低い品質を示す)は、分散保存フラグがオフであるとき(予測の比較的高い品質を示す)よりも高いノイズ利得が使われる。実験的に、これが全体的な知覚的品質を改善することが示されている。

【0 2 1 0】

上記で概説したように、ディザリングされる量子化器 3 2 2 の 3 2 4、3 2 5 の SNR 範囲は、制御パラメータ rfu に依存して変わりうる。表 2 によれば、分散保存フラグがオンのとき(予測の比較的低い品質を示す)、ディザリングされる量子化器 3 2 2 の固定した大きな範囲が使われる(たとえば範囲 3 2 4)。他方、分散保存フラグがオフのとき(予測の比較的高い品質を示す)は、制御パラメータ rfu に依存して異なる範囲 3 2 4、3 2 5 が使われる。

【0 2 1 1】

上記で概説したように、量子化された誤差係数のブロック 1 4 5 の決定は、ディザリングされる量子化器 3 2 2 によって量子化された量子化された誤差係数への事後利得 γ の適用に関わってもよい。事後利得 γ は、ディザリングされる量子化器 3 2 2 (たとえば減算的ディザのある量子化器)の MSE パフォーマンスを改善するために導出されてもよい。

【0212】

事後利得を分散保存フラグに依存させるとき、知覚的符号化品質が改善できることが実験的に示されている。上述したMSE最適事後利得は、分散保存フラグがオフのとき（予測の比較的高い品質を示す）に使われる。他方、分散保存フラグがオンのとき（予測の比較的低い品質を示す）は、（表2の右側の公式に従って決定される、）より高い事後利得を使うことが有益であることがある。

【0213】

上記で概説したように、予測誤差係数のブロック141より単位分散属性により近い再スケーリングされた誤差係数のブロック142を与えるために、ヒューリスティック・スケーリングが使われてもよい。ヒューリスティック・スケーリング規則は、制御パラメータ146に依存させられてもよい。換言すれば、ヒューリスティック・スケーリング規則は予測の品質に依存させられてもよい。ヒューリスティック・スケーリングは、予測の比較的高い品質の場合に特に有益であることがある。一方、その恩恵は予測の比較的低い品質の場合には限られていることがある。これに鑑み、分散保存フラグがオフであるとき（予測の比較的高い品質を示す）にのみヒューリスティック・スケーリングを使うことが有益でありうる。

【0214】

本稿では、変換ベースの発話エンコーダ100、170および対応する変換ベースの発話デコーダ500が記述されてきた。変換ベースの発話コーデックは、エンコードされた発話信号の品質を改善することを許容するさまざまな側面を利用しうる。特に、発話コーデックは、古典的な（ディザリングされない）量子化器、減算的ディザリングのある量子化器および「0レート」ノイズ充填を含む量子化器の順序付けられた集合を生成するよう構成されていてもよい。量子化器の順序付けられた集合は、該順序付けられた集合が信号包絡およびレート割り当てパラメータによってパラメータ化される知覚的モデルに従ってレート割り当てプロセスを容易にするような仕方でもよい。量子化器の集合の組成は、量子化方式の知覚的パフォーマンスを改善するためにサイド情報（たとえば予測器利得）の存在において再構成設定されてもよい。デコーダに対する追加的な信号伝達の必要なしに量子化器の順序付けられた集合の使用を容易にするレート割り当てアルゴリズムが使われてもよい。これはたとえば、エンコーダにおいて使われた量子化器の集合の特定の組成に関係するおよび／またはディザリングされる量子化器を実装するために使われたディザ信号に関係する追加的な信号伝達を必要としない。さらに、ビットレート制約条件（たとえば、最大許容ビット数に対する制約条件および／または最大受け入れ可能メッセージ長に対する制約条件）の存在のもとで算術符号化器（または範囲符号化器）の使用を容易にするレート割り当てアルゴリズムが使われてもよい。さらに、量子化器の順序付けられた集合は、特定の周波数帯域には0ビットの割り当てを許容しつつ、ディザリングされる量子化器の使用を容易にする。さらに、ハフマン符号化との関連で量子化器の順序付けられた集合の使用を容易にするレート割り当てアルゴリズムが使用されてもよい。

【0215】

本稿で記述された方法およびシステムは、ソフトウェア、ファームウェアおよび／またはハードウェアとして実装されてもよい。ある種のコンポーネントは、たとえばデジタル信号プロセッサまたはマイクロプロセッサ上で走るソフトウェアとして実装されてもよい。他のコンポーネントは、たとえば、ハードウェアおよび／または特定用途向け集積回路として実装されてもよい。記載される方法およびシステムにおいて遭遇される信号は、ランダム・アクセス・メモリまたは光記憶媒体のような媒体上に記憶されてもよい。それらの信号は、電波ネットワーク、衛星ネットワーク、無線ネットワークまたは有線ネットワーク、たとえばインターネットのようなネットワークを介して転送されてもよい。本稿に記載される方法およびシステムを利用する典型的な装置は、オーディオ信号を記憶および／またはレンダリングするために使われるポータブル電子装置または他の消費者設備である。

いくつかの態様を記載しておく。

10

20

30

40

50

〔態様1〕

係数のブロック（1 4 1）の第一の係数を量子化するよう構成された量子化ユニット（1 1 2）であって、係数の前記ブロックは、複数の対応する周波数ビン（3 0 1）についての複数の係数を含み、当該量子化ユニットは、

・一組（3 2 6、3 2 7）の量子化器を提供するよう構成されており、前記一組の量子化器は、それぞれSNRと称される異なる信号対雑音比に関連付けられた限られた数の異なる量子化器を含み、前記一組の量子化器の前記異なる量子化器は、そのSNRに従って順序付けられており、前記一組の量子化器は、

・ノイズ充填量子化器（3 2 1）；

・一つまたは複数のディザリングされる量子化器（3 2 2）；および

・一つまたは複数のディザリングされない量子化器（3 2 3）を含み、

当該量子化ユニットはさらに、

・前記第一の係数に帰されるSNRを示すSNR指示を決定し；

・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択し；

・前記第一の量子化器を使って前記第一の係数を量子化するよう構成されている、量子化ユニット。

〔態様2〕

・前記ノイズ充填量子化器は、前記異なるSNRのうち相対的に最低のSNRに関連付けられており、

・前記一つまたは複数のディザリングされない量子化器は、前記異なるSNRのうち一つまたは複数の相対的に最高のSNRに関連付けられており、

・前記一つまたは複数のディザリングされる量子化器は、前記異なるSNRのうち、前記相対的に最低のSNRより高く、前記一つまたは複数の相対的に最高のSNRより低い、一つまたは複数の中間的なSNRに関連付けられている、

態様1記載の量子化ユニット。

〔態様3〕

前記一組の量子化器は、前記異なる量子化器に関連付けられたSNRの昇順に従って順序付けられている、態様1または2記載の量子化ユニット。

〔態様4〕

・順序付けられた前記一組の量子化器からの一対の隣接する量子化器に関連付けられたSNRの差によってSNR差が与えられ、

・前記異なる量子化器からの隣接する量子化器のすべての対についてのSNR差が、所定のSNRターゲット差を中心とする所定のSNR差区間内にはいる、

態様3記載の量子化ユニット。

〔態様5〕

前記所定のSNR差区間の幅は、前記所定のSNRターゲット差の所定の割合より小さい、態様4記載の量子化ユニット。

〔態様6〕

前記所定のSNRターゲット差は1.5dBである、態様4または5記載の量子化ユニット。

〔態様7〕

前記ノイズ充填量子化器は、

・所定の統計モデルに従って乱数を生成するよう構成された乱数発生器を有し；

・前記第一の係数の値を前記所定の統計モデルに従って前記乱数発生器によって生成された乱数で置き換えることによって、前記第一の係数を量子化するよう構成されており、

・本質的には0dBより低いまたは0dBに等しいSNRに関連付けられている、

態様1ないし6のうちいずれか一項記載の量子化ユニット。

〔態様8〕

前記一つまたは複数のディザリングされる量子化器のうちの特定のディザリングされる量子化器は、

・前記第一の係数にディザ値を適用することによって第一のディザリングされた係数を決

10

20

30

40

50

定するよう構成されたディザ適用ユニット（６１１）と；

・前記第一のディザリングされた係数をスカラー量子化器のある区間に割り当てることによって第一の量子化インデックスを決定するよう構成されたスカラー量子化器（６１２）とを有する、

態様１ないし７のうちいずれか一項記載の量子化ユニット。

〔態様９〕

前記一つまたは複数のディザリングされる量子化器のうちの前記特定のディザリングされる量子化器はさらに、

・前記第一の量子化インデックスに第一の再構成値を割り当てるよう構成された逆スカラー量子化器（６１２）と；

・前記第一の再構成値から前記ディザ値を除去することによって、第一のディザリング解除された係数を決定するよう構成されたディザ除去ユニット（６１３）とを有する、

態様８記載の量子化ユニット。

〔態様１０〕

・前記ディザ適用ユニットは、前記第一の係数から前記ディザ値を減算するよう構成されており、前記ディザ除去ユニットは前記第一の再構成値に前記ディザ値を加算するよう構成されている；または

・前記ディザ適用ユニットは前記第一の係数に前記ディザ値を加算するよう構成されており、前記ディザ除去ユニットは前記第一の再構成値から前記ディザ値を減算するよう構成されている、

態様９記載の量子化ユニット。

〔態様１１〕

前記一つまたは複数のディザリングされる量子化器のうちの前記特定のディザリングされる量子化器はさらに、

・前記第一のディザリング解除された係数に量子化器事後利得 γ を適用することによって第一の量子化された係数を決定するよう構成された事後利得適用ユニットを有する、

態様９または１０記載の量子化ユニット。

〔態様１２〕

前記量子化器事後利得 γ は

【数１６】

$$\gamma = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}$$

によって与えられ、ここで、 $\sigma_x^2 = E\{X^2\}$ は係数の前記ブロック（１４１）の一つまたは複数の係数の分散であり、 Δ は前記特定のディザリングされる量子化器の前記スカラー量子化器の量子化器きざみサイズである、

態様１１記載の量子化ユニット。

〔態様１３〕

ディザ値のブロック（６０２）を生成するよう構成されたディザ生成器（６０１）をさらに有しており、ディザ値の前記ブロックは、それぞれ前記複数の周波数ビンについての複数のディザ値を含む、態様８ないし１２のうちいずれか一項記載の量子化ユニット。

〔態様１４〕

前記ディザ生成器は、

・ M は整数であるとして、 M 個の所定のディザ実現のうち一つを選択し；

・選択されたディザ実現に基づいてディザ値の前記ブロックを生成するよう構成されている、

態様１３記載の量子化ユニット。

〔態様１５〕

10

20

30

40

50

所定のディザ実現の数Mは10、5、4またはそれより少ない、態様14記載の量子化ユニット。

〔態様16〕

前記ディザ値が擬似乱数である、態様8ないし15のうちいずれか一項記載の量子化ユニット。

〔態様17〕

・前記スカラー量子化器が、所定の量子化器きざみサイズ Δ を有し；
・前記ディザ値は、所定のディザ区間からの値を取り；
・前記所定のディザ区間は、前記所定の量子化器きざみサイズ Δ 以下の幅を有する、
態様8ないし16のうちいずれか一項記載の量子化ユニット。

10

〔態様18〕

ディザ値の前記ブロックは、前記所定のディザ区間内に一様に分布している、態様13を引用する場合の態様17記載の量子化ユニット。

〔態様19〕

前記一つまたは複数のディザリングされる量子化器は減算的なディザリングされる量子化器である、態様1ないし18のうちいずれか一項記載の量子化ユニット。

〔態様20〕

前記一つまたは複数のディザリングされない量子化器のうちのあるディザリングされない量子化器は、所定の一樣な量子化器きざみサイズをもつスカラー量子化器である、態様1ないし19のうちいずれか一項記載の量子化ユニット。

20

〔態様21〕

・係数の前記ブロック(141)は、スペクトル・ブロック包絡(136)に関連付けられており；
・前記スペクトル・ブロック包絡は前記複数の周波数ビンについて複数のスペクトル・エネルギー値(303)を示し；
・前記SNR指示が前記スペクトル・ブロック包絡に依存する、
態様1ないし20のうちいずれか一項記載の量子化ユニット。

〔態様22〕

・前記SNR指示がさらに、前記スペクトル・ブロック包絡をオフセットさせるためのオフセット・パラメータに依存し；
・前記オフセット・パラメータは、係数の前記ブロック(141)をエンコードするために利用可能な所定のビット数に依存する、
態様21記載の量子化ユニット。

30

〔態様23〕

前記第一の係数に帰されるSNRを示す前記SNR指示は、前記オフセット・パラメータを使って前記第一の係数の周波数ビンに関連付けられたスペクトル・ブロック包絡から導出される値をオフセットさせることによって決定される、態様22記載の量子化ユニット。

〔態様24〕

・前記SNR指示は、前記スペクトル・ブロック包絡から導出される割り当て包絡(138)に依存し；
・前記割り当て包絡は、割り当て分解能を有し；
・前記割り当て分解能は、前記一組の量子化器からの隣接する量子化器の間の前記SNR差に依存する、態様21ないし23のうちいずれか一項記載の量子化ユニット。

40

〔態様25〕

・係数の前記ブロック(141)の前記複数の係数は、複数の周波数帯域に割り当てられ；
・周波数帯域は、一つまたは複数の周波数ビンを含み；
・当該量子化ユニットは、同じ周波数帯域に割り当てられる係数が同じ量子化器を使って量子化されるよう、前記複数の周波数帯域のそれぞれについて前記一組の量子化器から量子化器を選択するよう構成されている、

50

態様 1 ないし 2 4 のうちいずれか一項記載の量子化ユニット。

〔態様 2 6〕

周波数帯域当たりの周波数ビンの数は、周波数が増すとともに増大する、態様 2 5 記載の量子化ユニット。

〔態様 2 7〕

当該量子化ユニットは、

- ・係数の前記ブロック (1 4 1) の属性を示すサイド情報 (7 2 1) を決定し (7 0 1) ;
- ・前記サイド情報に依存して量子化器の前記組 (3 2 6、3 2 7) を生成する (7 0 2) よう構成されている、

10

態様 1 ないし 2 6 のうちいずれか一項記載の量子化ユニット。

〔態様 2 8〕

前記ノイズ充填量子化器の前記乱数発生器の前記所定の統計モデルは前記サイド情報に依存する、態様 7 を引用する場合の態様 2 7 記載の量子化ユニット。

〔態様 2 9〕

前記一組の量子化器のうちのディザリングされる量子化器の数が前記サイド情報に依存する、態様 2 7 または 2 8 記載の量子化ユニット。

〔態様 3 0〕

当該量子化ユニットは、当該量子化ユニットを有するエンコーダにおいておおよび対応する逆量子化ユニットを有する対応するデコーダにおいて利用可能なデータから前記サイド情報を抽出するよう構成されている、態様 2 7 ないし 2 9 のうちいずれか一項記載の量子化ユニット。

20

〔態様 3 1〕

前記サイド情報が：

- ・係数の前記ブロック (1 4 1) のトーン性内容を示す、前記エンコーダ内に含まれる予測器 (1 1 7) によって決定された予測器利得；および／または
 - ・係数の前記ブロックの摩擦性内容を示す、係数の前記ブロック (1 4 1) に基づいて導出されたスペクトル反射係数
- のうちの少なくとも一つを含む、態様 3 0 記載の量子化ユニット。

〔態様 3 2〕

30

前記一組の所定の量子化器に含まれるディザリングされる量子化器の数は、予測器利得の増大とともに減少し、予測器利得の減少とともに増大する、態様 3 1 記載の量子化ユニット。

〔態様 3 3〕

- ・前記サイド情報が分散保存フラグを含み；
 - ・前記分散保存フラグは、係数の前記ブロック (1 4 1) の分散がどのように調整されるべきかを示し；
 - ・前記一組の量子化器は、前記分散保存フラグに依存して決定される、
- 態様 2 7 ないし 3 2 のうちいずれか一項記載の量子化ユニット。

〔態様 3 4〕

40

前記ノイズ充填量子化器のノイズ利得が前記分散保存フラグに依存する、態様 3 3 記載の量子化ユニット。

〔態様 3 5〕

前記一つまたは複数のディザリングされる量子化器によってカバーされる SNR 範囲が前記分散保存フラグに依存して決定される、態様 3 3 または 3 4 記載の量子化ユニット。

〔態様 3 6〕

前記事後利得 γ が前記分散保存フラグに依存する、態様 3 3 ないし 3 5 のうちいずれか一項記載の量子化ユニット。

〔態様 3 7〕

量子化インデックスを量子化解除するよう構成された逆量子化ユニット (5 5 2) であ

50

って、前記量子化インデックスは、複数の対応する周波数ビンについて複数の係数を含む係数のブロックに関連付けられており、当該逆量子化ユニットは、

・一組の量子化器を提供するよう構成されており、前記一組の量子化器は、それぞれSNRと称される異なる信号対雑音比に関連付けられた、限られた数の異なる量子化器を含み、前記一組の量子化器の前記異なる量子化器は、そのSNRに従って順序付けられており、前記一組の量子化器は、

・ノイズ充填量子化器；

・一つまたは複数のディザリングされる量子化器；および

・一つまたは複数のディザリングされない量子化器を含み、

当該逆量子化ユニットはさらに、

・係数の前記ブロックからの第一の係数に帰されるSNRを示すSNR指示を決定し；

・前記SNR指示に基づいて前記一組の量子化器から第一の量子化器を選択し；

・前記第一の量子化器を使って前記第一の係数について第一の量子化された係数を決定するよう構成されている、

逆量子化ユニット。

〔態様38〕

オーディオ信号をビットストリームにエンコードするよう構成された変換ベースのオーディオ・エンコーダであって、

・ディザリングされる量子化器を使って、係数のブロック(1 4 1)からの複数の係数を量子化することによって複数の量子化インデックスを決定するよう構成された量子化ユニットを有しており、前記複数の係数は、複数の対応する周波数ビンに関連付けられており、係数の前記ブロックは、前記オーディオ信号から導出され、

当該オーディオ・エンコーダはさらに、

・Mが1より大きな整数であるとして、M個の所定のディザ実現のうちの一つを選択するよう構成されており、選択されたディザ実現に基づいて前記複数の係数を量子化するための複数のディザ値を生成するよう構成されたディザ生成器と；

・M個の所定のコードブックからコードブックを選択するよう構成されており、選択されたコードブックを使って前記複数の量子化インデックスをエントロピー符号化するよう構成されたエントロピー符号化器とを有しており、前記M個の所定のコードブックはそれぞれ前記M個の所定のディザ実現に関連付けられており、前記エントロピー符号化器は、前記ディザ生成器によって選択されたディザ実現に関連付けられたコードブックを選択するよう構成されており、エントロピー符号化された量子化インデックスを示す係数データが前記ビットストリーム中に挿入される、

変換ベースのオーディオ・エンコーダ。

〔態様39〕

所定のディザ実現の数Mが10、5、4またはそれより少ない、態様38記載の変換ベースの発話エンコーダ。

〔態様40〕

前記M個の所定のコードブックが、それぞれ前記M個の所定のディザ実現を使ってトレーニングされたものである、態様38または39記載の変換ベースの発話エンコーダ。

〔態様41〕

前記M個の所定のコードブックが可変長のハフマン符号語を含む、態様38ないし40のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様42〕

ビットストリームをデコードして再構成されたオーディオ信号を提供するよう構成された変換ベースのオーディオ・デコーダであって、

・Mは1より大きな整数であるとして、M個の所定のディザ実現のうちの一つを選択するよう構成され、選択されたディザ実現に基づいて複数のディザ値を生成するよう構成されたディザ生成器を有しており、前記複数のディザ値は、対応する複数の量子化インデックスに基づいて対応する複数の量子化された係数を決定するよう構成されているディザリング

10

20

30

40

50

される量子化器を有する逆量子化ユニットによって使われるものであり、
当該変換ベースのオーディオ・デコーダはさらに、

・M個の所定のコードブックからコードブックを選択するよう構成され、選択されたコードブックを使って前記ビットストリームから係数データ（163）をエントロピー復号するよう構成されたエントロピー復号器を有しており、前記M個の所定のコードブックは、それぞれ前記M個の所定のディザ実現と関連付けられており、前記エントロピー復号器は、前記ディザ生成器によって選択されたディザ実現に関連付けられたコードブックを選択するよう構成されており、再構成されたオーディオ信号は、前記複数の量子化された係数に基づいて決定される、

変換ベースのオーディオ・デコーダ。

〔態様43〕

発話信号をビットストリームにエンコードするよう構成された変換ベースの発話エンコーダであって、

・変換係数の複数の逐次的なブロック（131）を受領するよう構成されたフレーミング・ユニットであって、前記複数の逐次的なブロックは、現在ブロックおよび一つまたは複数の以前のブロックを含み、前記複数の逐次的なブロックは、発話信号のサンプルを示す、フレーミング・ユニットと；

・対応する現在ブロック包絡（136）を使って変換係数の対応する現在ブロック（131）を平坦化することによって、平坦化された変換係数の現在ブロック（140）を決定するよう構成された平坦化ユニットと；

・再構成された変換係数の一つまたは複数の以前のブロック（149）に基づいて、かつ一つまたは複数の予測器パラメータに基づいて、推定された平坦化された変換係数の現在ブロック（150）を決定するよう構成された予測器であって、再構成された変換係数の前記一つまたは複数の以前のブロックは、変換係数の前記一つまたは複数の以前のブロック（131）から導出されたものである、予測器と；

・平坦化された変換係数の現在ブロック（140）に基づいて、かつ推定された平坦化された変換係数の現在ブロック（150）に基づいて、予測誤差係数の現在ブロック（141）を決定するよう構成された差分ユニットと；

・予測誤差係数の現在ブロック（141）から導出された係数を量子化するよう構成された、態様1ないし36のうちいずれか一項記載の量子化ユニットとを有しており、前記ビットストリームについての係数データ（163）は量子化された係数に関連付けられた量子化インデックスに基づいて決定される、

変換ベースの発話エンコーダ。

〔態様44〕

・変換係数のブロック（131）がMDCT係数を含む；および／または

・変換係数のブロック（131）が、256個の周波数ビン内の256個の変換係数を含む、
態様43記載の変換ベースの発話エンコーダ。

〔態様45〕

再スケーリングされた誤差係数の現在ブロック（142）の再スケーリングされた誤差係数の分散が、平均では、予測誤差係数の現在ブロック（141）の予測誤差係数の分散より高くなるよう、一つまたは複数のスケーリング規則を使って予測誤差係数の現在ブロック（141）に基づいて、再スケーリングされた誤差係数の現在ブロック（142）を決定するよう構成されたスケーリング・ユニットをさらに有する、態様43または44記載の変換ベースの発話エンコーダ。

〔態様46〕

・予測誤差係数の現在ブロック（141）は、対応する複数の周波数ビンについての複数の予測誤差係数を含み、

・前記一つまたは複数のスケーリング規則に従って前記スケーリング・ユニットによって前記予測誤差係数に適用されるスケーリング利得は、それぞれの予測誤差係数の周波数ビンに依存する、

10

20

30

40

50

態様 4 5 記載の変換ベースの発話エンコーダ。

〔態様 4 7〕

前記スケーリング規則は、前記一つまたは複数の予測器パラメータに依存する、態様 4 5 または 4 6 記載の変換ベースの発話エンコーダ。

〔態様 4 8〕

前記スケーリング規則は、現在ブロック包絡（1 3 6）に依存する、態様 4 5 ないし 4 7 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 4 9〕

・前記予測器は、重み付けされた平均平方誤差基準を使って、推定された平坦化された変換係数の現在ブロック（1 5 0）を決定するよう構成されており、

・前記重み付けされた平均平方誤差基準は、現在ブロック包絡（1 3 6）を重みとして考慮に入れる、

態様 3 9 ないし 4 8 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 5 0〕

前記係数量子化ユニットは、再スケーリングされた誤差係数の現在ブロック（1 4 2）の再スケーリングされた誤差係数を量子化するよう構成されている、態様 3 9 ないし 4 9 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 5 1〕

・当該変換ベースの発話エンコーダが現在ブロック包絡（1 3 6）に基づいて割り当てベクトルを決定するよう構成されたビット割り当てユニット（1 0 9、1 1 0、1 7 1、1 7 2）をさらに有しており、

・前記割り当てベクトルは、予測誤差係数の現在ブロック（1 4 1）から導出された第一の係数を量子化するために使われる前記一組の所定の量子化器からの第一の量子化器を示す、

態様 3 9 ないし 5 0 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 5 2〕

前記割り当てベクトルが、それぞれ予測誤差係数の現在ブロック（1 4 1）から導出された係数全部について使われる諸量子化器を示す、態様 5 1 記載の変換ベースの発話エンコーダ。

〔態様 5 3〕

前記ビット割り当てユニットは、前記一つまたは複数のスケーリング規則にも基づいて前記割り当てベクトルを決定するよう構成されている、態様 4 5 を引用する場合の態様 5 1 または 5 2 記載の変換ベースの発話エンコーダ。

〔態様 5 4〕

前記ビット割り当てユニットは、

・予測誤差係数の現在ブロック（1 4 1）についての係数データ（1 6 3）が所定のビット数を超えないよう前記割り当てベクトルを決定し；

・現在ブロック包絡（1 3 6）から導出される割り当て包絡（1 3 8）に適用されるべきオフセットを示すオフセット・パラメータを決定するよう構成されており、

前記オフセット・パラメータが、前記ビットストリーム中に含まれる、

態様 5 1 ないし 5 3 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 5 5〕

前記量子化された係数に関連付けられた量子化インデックスをエントロピー符号化するよう構成されたエントロピー符号化器をさらに有する、態様 3 9 ないし 5 4 のうちいずれか一項記載の変換ベースの発話エンコーダ。

〔態様 5 6〕

前記エントロピー符号化器は、算術符号化器を使って前記量子化インデックスを符号化するよう構成されている、態様 5 5 記載の変換ベースの発話エンコーダ。

〔態様 5 7〕

ビットストリームをデコードして再構成された発話信号を提供するよう構成された変換

10

20

30

40

50

ベースの発話デコーダであって、

・再構成された変換係数の一つまたは複数の以前のブロック（１４９）に基づいて、かつ前記ビットストリームから導出される一つまたは複数の予測器パラメータ（５２０）に基づいて、推定された平坦化された変換係数の現在ブロック（１５０）を決定するよう構成された予測器と；

・一組の所定の量子化器を使って、前記ビットストリーム内に含まれる係数データ（１６３）に基づいて、量子化された予測誤差係数の現在ブロック（１４７）を決定するよう構成された、態様３７記載の逆量子化ユニットと；

・推定された平坦化された変換係数の現在ブロック（１５０）に基づき、かつ量子化された予測誤差係数の現在ブロック（１４７）に基づいて、再構成された平坦化された変換係数の現在ブロック（１４８）を決定するよう構成された加算ユニットと；

・現在ブロック包絡（１３６）を使って、再構成された平坦化された変換係数の現在ブロック（１４８）にスペクトル形状を与えることによって、再構成された変換係数の現在ブロック（１４９）を決定するよう構成された逆平坦化ユニットとを有しており、再構成された発話信号は、再構成された変換係数の現在ブロック（１４９）に基づいて決定される、

変換ベースの発話デコーダ。

〔態様５８〕

係数のブロック（１４１）の第一の係数を量子化する方法であって、係数の前記ブロック（１４１）は、複数の対応する周波数ビンについての複数の係数を含み、当該方法は、

・一組の量子化器を提供する段階であって、前記一組の量子化器は、それぞれSNRと称される複数の異なる信号対雑音比に関連付けられた複数の異なる量子化器を含み、前記複数の異なる量子化器は、

・ノイズ充填量子化器、

・一つまたは複数のディザリングされる量子化器、および

・一つまたは複数のディザリングされない量子化器を含む、段階と；

・前記第一の係数に帰されるSNRを示すSNR指示を決定する段階と；

・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択する段階と；

・前記第一の量子化器を使って前記第一の係数を量子化する段階とを含む、

方法。

〔態様５９〕

量子化インデックスを量子化解除する方法であって、前記量子化インデックスは、複数の対応する周波数ビンについて複数の係数を含む係数のブロック（１４１）に関連付けられており、当該方法は、

・一組の量子化器を提供する段階であって、前記一組の量子化器は、それぞれSNRと称される複数の異なる信号対雑音比に関連付けられた複数の異なる量子化器を含み、前記複数の異なる量子化器は、

・ノイズ充填量子化器、

・一つまたは複数のディザリングされる量子化器、および

・一つまたは複数のディザリングされない量子化器を含む、段階と；

・係数の前記ブロック（１４１）からの第一の係数に帰されるSNRを示すSNR指示を決定する段階と；

・前記SNR指示に基づいて、前記一組の量子化器から第一の量子化器を選択する段階と；

・前記第一の量子化器を使って、前記第一の係数についての第一の量子化された係数を決定する段階とを含む、

方法。

〔態様６０〕

オーディオ信号をビットストリームにエンコードする方法であって、

・ディザリングされる量子化器を使って係数のブロック（１４１）からの複数の係数を量子化することによって複数の量子化インデックスを決定する段階であって、前記複数の係

10

20

30

40

50

数は複数の対応する周波数ビンに関連付けられており、係数の前記ブロック（１４１）は前記オーディオ信号から導出される、段階と；

- ・ M個の所定のディザ実現の一つを選択する段階と；
- ・ 選択されたディザ実現に基づいて前記複数の係数を量子化するための複数のディザ値を生成する段階であって、Mは1より大きい整数である、段階と；
- ・ M個の所定のコードブックからコードブックを選択する段階と；
- ・ 選択されたコードブックを使って前記複数の量子化インデックスをエントロピー符号化する段階であって、前記M個の所定のコードブックは、それぞれ前記M個の所定のディザ実現に関連付けられており、選択されたコードブックは、選択されたディザ実現に関連付けられている、段階と；
- ・ エントロピー符号化された量子化インデックスを示す係数データ（１６３）を前記ビットストリーム中に挿入する段階とを含む、

方法。

〔態様 6 1〕

ビットストリームをデコードして再構成されたオーディオ信号を提供する方法であって、

- ・ M個の所定のディザ実現のうちの一つを選択する段階と；
- ・ 選択されたディザ実現に基づいて複数のディザ値を生成する段階であって、Mは1より大きい整数であり、前記複数のディザ値は、対応する複数の量子化インデックスに基づいて対応する複数の量子化された係数を決定する、ディザリングされる量子化器を有する逆量子化ユニットによって使われるものである、段階と；
- ・ M個の所定のコードブックからコードブックを選択する段階と；
- ・ 選択されたコードブックを使って前記ビットストリームから係数データ（１６３）をエントロピー復号して、前記複数の量子化インデックスを提供する段階であって、前記個の所定のコードブックは、それぞれM個の所定のディザ実現と関連付けられており、選択されたコードブックは、選択されたディザ実現に関連付けられている、段階と；
- ・ 前記複数の量子化された係数に基づいて前記再構成されたオーディオ信号を決定する段階とを含む、

方法。

〔態様 6 2〕

- 発話信号をビットストリームにエンコードする方法であって、
- ・ 現在ブロックおよび一つまたは複数の以前のブロックを含む変換係数の複数の逐次的なブロックを受領する段階であって、前記複数の逐次的なブロックは、発話信号のサンプルを示す、段階と；
- ・ 対応する現在ブロック包絡（１３６）を使って変換係数の対応する現在ブロックを平坦化することによって、平坦化された変換係数の現在ブロック（１４０）を決定する段階と；
- ・ 再構成された変換係数の一つまたは複数の以前のブロック（１４９）に基づいて、かつ一つまたは複数の予測器パラメータ（５２０）に基づいて、推定された平坦化された変換係数の現在ブロック（１５０）を決定する段階であって、再構成された変換係数の前記一つまたは複数の以前のブロックは、変換係数の前記一つまたは複数の以前のブロックから導出されたものである、段階と；
- ・ 平坦化された変換係数の現在ブロック（１４０）に基づいて、かつ推定された平坦化された変換係数の現在ブロック（１５０）に基づいて、予測誤差係数の現在ブロック（１４１）を決定する段階と；
- ・ 予測誤差係数の現在ブロック（１４１）から導出された係数を、態様 5 8 記載の方法に従って量子化する段階と；
- ・ 前記ビットストリームについての係数データ（１６３）を、前記量子化された係数に関連付けられた量子化インデックスに基づいて決定する段階とを含む、

方法。

10

20

30

40

50

〔態様 6 3〕

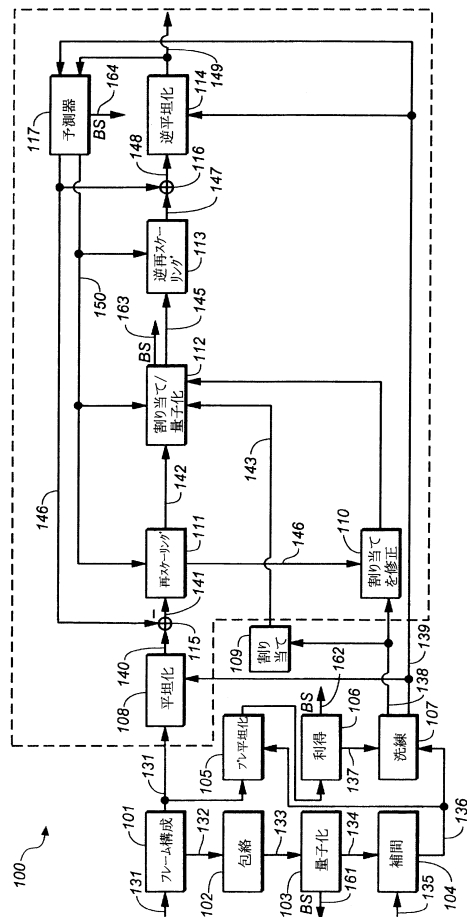
ビットストリームをデコードして、再構成された発話信号を提供する方法であって、

- ・再構成された変換係数の一つまたは複数の以前のブロック（１４９）に基づき、かつ前記ビットストリームから導出された一つまたは複数の予測器パラメータ（５２０）に基づいて、推定された平坦化された変換係数の現在ブロック（１５０）を決定する段階と；
- ・態様５９記載の方法を使って、前記ビットストリーム内に含まれる係数データ（１６３）に基づいて、量子化された予測残差係数の現在ブロック（１４７）を決定する段階と；
- ・推定された平坦化された変換係数の現在ブロック（１５０）に基づき、かつ量子化された予測誤差係数の現在ブロック（１４７）に基づいて、再構成された平坦化された変換係数の現在ブロック（１４８）を決定する段階と；
- ・現在ブロック包絡（１３６）を使って、再構成された平坦化された変換係数の現在ブロック（１４８）にスペクトル形状を与えることによって、再構成された変換係数の現在ブロック（１４９）を決定する段階と；
- ・再構成された変換係数の現在ブロック（１４９）に基づいて再構成された発話信号を決定する段階とを含む、

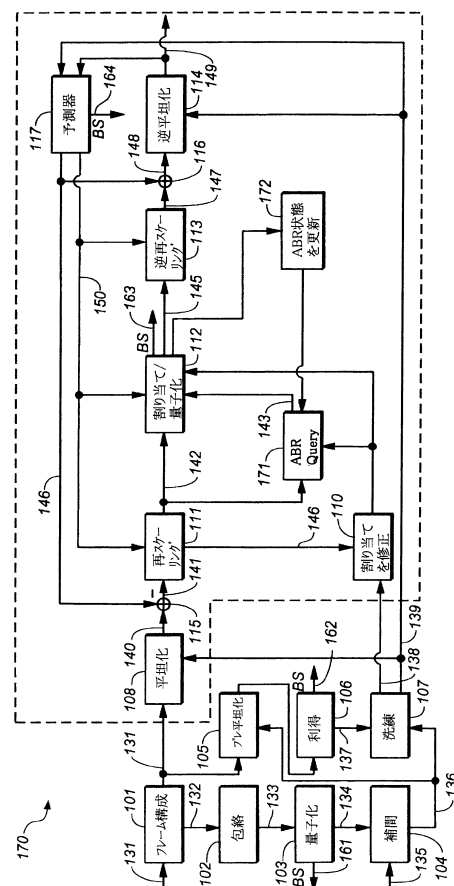
方法。

10

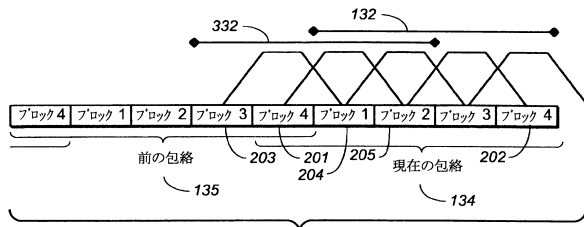
【図 1 a】



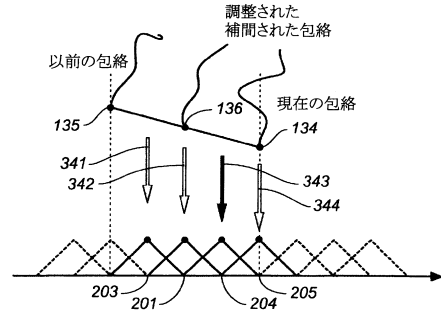
【図 1 b】



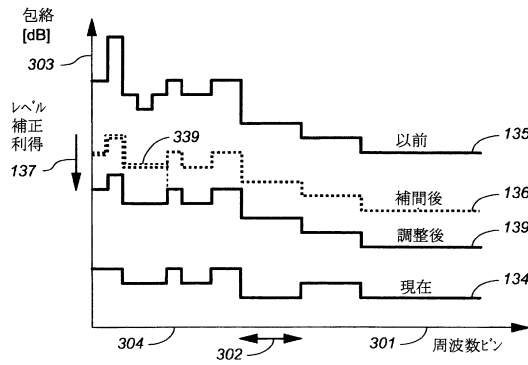
【図 2】



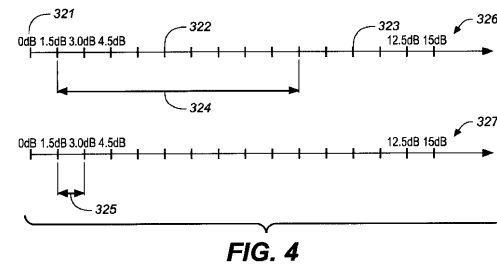
【図 3 b】



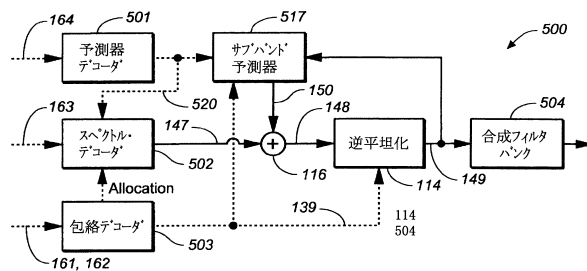
【図 3 a】



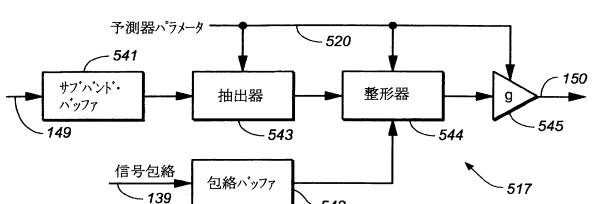
【図 4】



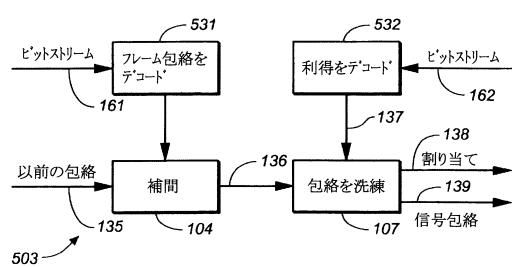
【図 5 a】



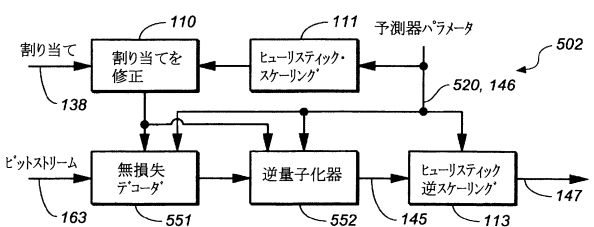
【図 5 c】



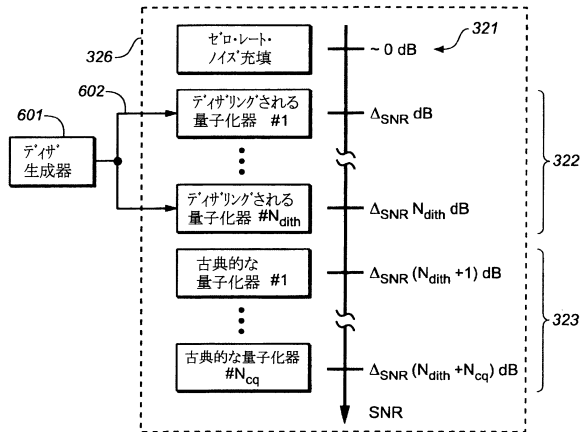
【図 5 b】



【図 5 d】



【図 6 a】



【図 6 b】

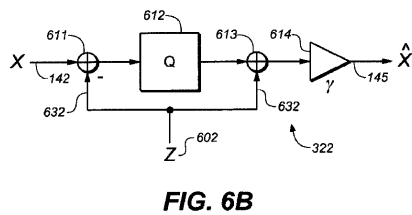


FIG. 6B

【図 6 c】

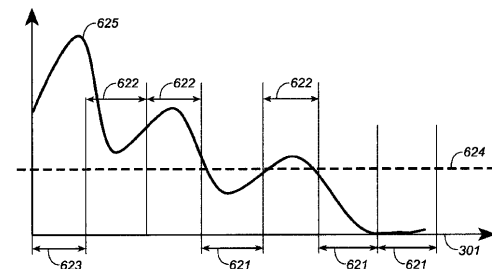
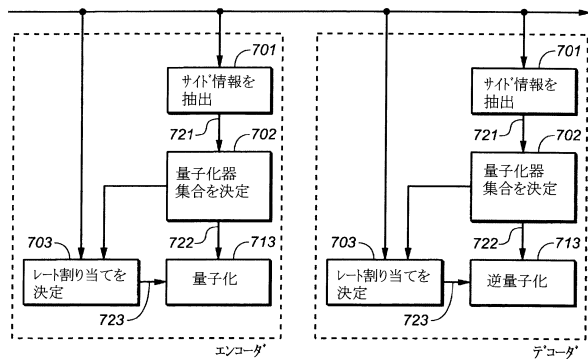
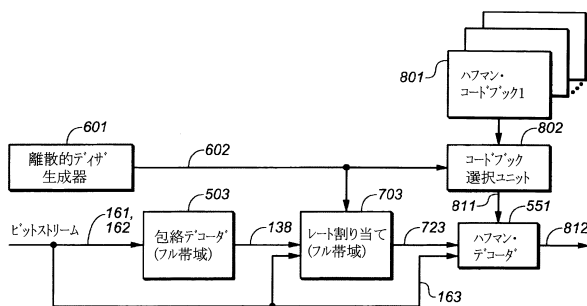


FIG. 6C

【図 7】



【図 8】



【図 9 A】

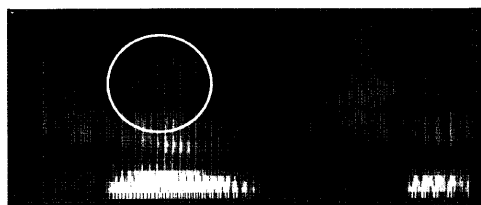


FIG. 9A

【図 9 B】

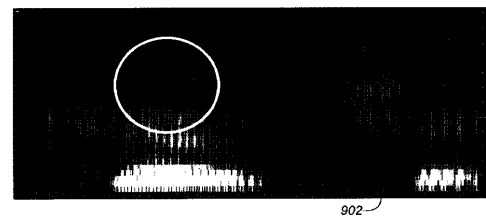


FIG. 9B

【図 9 C】

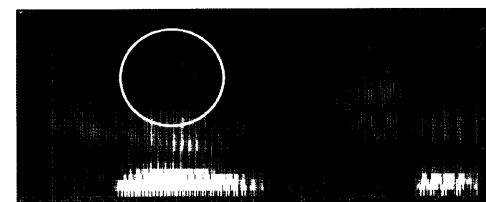
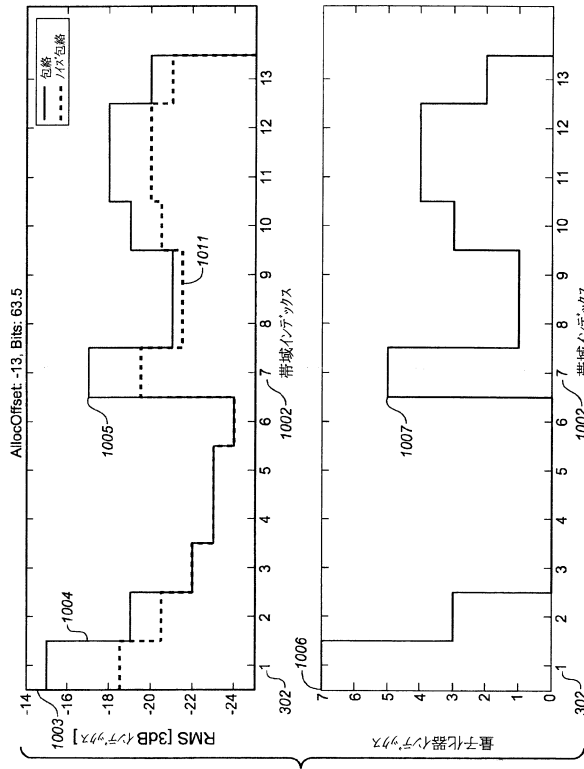


FIG. 9C

【図10】



フロントページの続き

- (72)発明者 クレイサ, ヤヌッシュ
スウェーデン王国, 113 30 ストックホルム, ガブレガタン 12アー ドルビー スウェーデン アーバー内
- (72)発明者 ヴィレモーズ, ラーシュ
スウェーデン王国, 113 30 ストックホルム, ガブレガタン 12アー ドルビー スウェーデン アーバー内
- (72)発明者 ヘデリン, ペール
スウェーデン王国, 113 30 ストックホルム, ガブレガタン 12アー ドルビー スウェーデン アーバー内

審査官 五貫 昭一

(56)参考文献 特表2009-501359 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|--------|
| G10L | 19/00 |
| G10L | 19/028 |
| G10L | 19/035 |
| G10L | 19/038 |
| H03M | 7/30 |