

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5627843号
(P5627843)

(45) 発行日 平成26年11月19日(2014.11.19)

(24) 登録日 平成26年10月10日(2014.10.10)

(51) Int.Cl.

G 1 O L 19/025 (2013.01)

F 1

G 1 O L 19/025

請求項の数 12 外国語出願 (全 19 頁)

(21) 出願番号 特願2008-154011 (P2008-154011)
 (22) 出願日 平成20年6月12日 (2008.6.12)
 (65) 公開番号 特開2008-310327 (P2008-310327A)
 (43) 公開日 平成20年12月25日 (2008.12.25)
 審査請求日 平成23年6月6日 (2011.6.6)
 (31) 優先権主張番号 07110289.1
 (32) 優先日 平成19年6月14日 (2007.6.14)
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 501263810
 トムソン ライセンシング
 Thomson Licensing
 フランス国, 92130 イッシー レ
 ムーリノー, ル ジヤンヌ ダルク,
 1-5
 1-5, rue Jeanne d' Arc, 92130 ISSY LES
 MOULINEAUX, France
 (74) 代理人 110001243
 特許業務法人 谷・阿部特許事務所
 (74) 復代理人 100115624
 弁理士 濱中 淳宏
 (74) 復代理人 100115635
 弁理士 署田 郁大

最終頁に続く

(54) 【発明の名称】スペクトル領域において適応切り替え式時間分解を使用して音声信号を符号化及び復号化する方法及び装置

(57) 【特許請求の範囲】

【請求項 1】

音声入力信号を、前記音声入力信号の第一の長さの区間に適用される周波数領域への第一のM D C T 又は整数M D C T 又はD C T - 4 変換を使用し、時間分解能の適応切り替えを使用し、その後、結果的に生じた周波数領域ビンの値の量子化及びエントロピ符号化を使用して符号化する方法であって、

前記切り替え、量子化、及び / 又はエントロピ符号化の制御は、前記音声入力信号の心理音響分析に由来し、

前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換に続いて、前記変換済みの第一の長さの区間内の第二の長さの区間に適用される第二のM D C T 又は整数M D C T 又はD C T - 4 変換を実行することで、前記時間分解能の適応的な制御を行うステップであり、

前記第二の長さは、前記第一の長さより短く、前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換の出力値、又は前記第二のM D C T 又は整数M D C T 又はD C T - 4 変換の出力値及び前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換の残りの出力値、の何れかを前記量子化及びエントロピ符号化において処理し、

前記第一および第二の変換の各々の前に、前記第一の長さ又は前記第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ又は第二の長さの区間にに対する重複加算処理が加えられ、前記第一の長さ及び第二の長さの区間の間の境界における遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記第二の長さの区間にに対して、

前記加重のために開始及び停止窓関数が使用される、前記時間分解能の適応的な制御を行うステップと、

符号化出力信号に対して、対応する時間分解能制御情報を副次的情報として添付するステップと、

を含む、前記方法。

【請求項 2】

入力信号の第一の長さの区間に適用される周波数領域への第一のM D C T 又は整数M D C T 又はD C T - 4 変換を使用して符号化された符号化音声信号を復号化し、前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換に続いて、前記変換済みの第一の長さの区間内の第二の長さの区間に適用される第二のM D C T 又は整数M D C T 又はD C T - 4 変換を実行することで時間分解能を適応的に切り替え、前記第二の長さは、前記第一の長さより短く、前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換の出力値、又は前記第二のM D C T 又は整数M D C T 又はD C T - 4 変換の出力値及び前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換の残りの出力値、の何れかを量子化及びエントロピ符号化において処理し、

前記第一及び第二の変換の各々の前に、前記第一の長さ又は前記第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ又は第二の長さの区間にに対する重複加算処理が加えられ、前記第一の長さ及び第二の長さの区間の間の境界における遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記第二の長さの区間にに対して、前記加重のために開始及び停止窓関数が使用され、前記切り替え、量子化、及び / 又はエントロピ符号化の制御は、前記入力信号の心理音響分析に由来し、対応する時間分解能制御情報を符号化出力信号に対して副次的情報として添付した符号化信号を復号化する方法であって、

前記符号化信号から前記副次的情報を提供するステップと、

前記符号化信号を逆量子化及びエントロピ復号化するステップと、

前記副次的情報に対応して、前記逆量子化及びエントロピ復号化信号の第一の長さの信号区間に作用し、復号化信号を提供する、時間領域への第一の逆M D C T 又は逆整数M D C T 又は逆D C T - 4 変換を実行するか、或いは、前記第一の逆M D C T 又は逆整数M D C T 又は逆D C T - 4 変換を実行する前に、第二の逆M D C T 又は逆整数M D C T 又は逆D C T - 4 変換において、前記逆量子化及びエントロピ復号化信号の第二の長さの区間を処理するステップと、

を含み、

前記第一及び第二の逆変換後に、前記第一の長さ及び前記第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ及び第二の長さの区間にに対する重複加算処理が加えられ、遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記第二の長さの区間にに対して、前記加重のために開始及び停止窓関数が使用される、前記復号化方法。

【請求項 3】

二つ以上の異なる第二の長さが使用される場合、利用する異なる第二の長さのトポロジを信号伝達するために、変化する時間分解能の領域を示す幾つかの指標、或いは復号化側でアクセス可能な対応するコードブックにおける一致するエントリを示す指標番号が、前記副次的情報に含まれる、請求項 1 又は 2 に記載の方法。

【請求項 4】

二つ以上の異なる第二の長さが連続して使用される場合、長さは、低周波数線を表す周波数ビンを開始点として増加する、請求項 1 乃至 3 の何れかに記載の方法。

【請求項 5】

前記トポロジは、

選択周波数帯域に対して、変換ビンのスペクトル電力を決定し、前記スペクトル電力の算術平均値をその幾何平均値により割ることで、前記第一のM D C T 又は整数M D C T 又はD C T - 4 変換を使用してスペクトル平坦性基準S F Mを実行するステップと、

10

20

30

40

50

非加重入力信号区間をサブセグメント化し、 m 個のサブセクションに対して加重及び短変換を実行し、こうした変換の周波数分解能は、前記選択周波数帯域に対応するステップと、

m 個の変換セグメントからなる各周波数線に対して、スペクトル電力を決定し、 m 個のセグメントの幾何平均により割った算術平均を決定することで時間平坦性基準T F Mを計算するステップと、

S F M値を使用することで、音調又はノイズ周波数帯域を決定するステップと、

こうした帯域内での時間的変化を認識するために、T F M値を使用し、閾値を使用して、前記決定されたノイズ周波数帯域のために、更に細かな時間分解能への切り替えを行うステップと、により決定される、請求項3記載の方法。

10

【請求項6】

音声入力信号を符号化する装置であって、

前記音声入力信号の第一の長さの区間を周波数領域へ変換するのに適した第一のM D C T又は整数M D C T又はD C T - 4変換手段と、

前記変換済みの第一の長さの区間の、前記第一の長さより短い第二の長さの区間を変換するのに適した第二のM D C T又は整数M D C T又はD C T - 4変換手段と、

前記第一のM D C T又は整数M D C T又はD C T - 4変換手段の出力値、又は前記第二のM D C T又は整数M D C T又はD C T - 4変換手段の出力値及び前記第一のM D C T又は整数M D C T又はD C T - 4変換の残りの出力値、を量子化及びエントロピ符号化するのに適した手段と、

20

前記量子化及び／又はエントロピ符号化の制御を行うと共に、前記第一のM D C T又は整数M D C T又はD C T - 4変換手段の出力値、又は前記第二のM D C T又は整数M D C T又はD C T - 4変換手段の出力値及び前記第一のM D C T又は整数M D C T又はD C T - 4変換の残りの出力値、が前記量子化及びエントロピ符号化手段において処理されるかについて、適応的に制御を行うのに適した手段であって、前記制御は、前記音声入力信号の心理音響分析に由来し、

前記第一および第二の各々の変換の前に、前記第一の長さ又は第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ又は第二の長さの区間にに対する重複加算処理が加えられ、前記第一の長さ及び第二の長さの区間の間の境界における遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記加重のために前記第二の長さの区間にに対する、開始及び停止窓関数が使用される、前記制御を行うのに適した手段と、

30

符号化装置出力信号に対して、対応する時間分解能制御情報を副次的情報として添付するのに適した手段と、

を含む、前記装置。

【請求項7】

入力信号の第一の長さの区間に適用される周波数領域への第一のM D C T又は整数M D C T又はD C T - 4変換を使用して符号化された符号化音声信号を復号化し、前記第一のM D C T又は整数M D C T又はD C T - 4変換に続いて、前記変換済みの第一の長さの区間内の第二の長さの区間に適用される第二のM D C T又は整数M D C T又はD C T - 4変換を実行することで時間分解能を適応的に切り替え、前記第二の長さは、前記第一の長さより短く、前記第一のM D C T又は整数M D C T又はD C T - 4変換の出力値、又は前記第二のM D C T又は整数M D C T又はD C T - 4変換の出力値及び前記第一のM D C T又は整数M D C T又はD C T - 4変換の残りの出力値、の何れかを量子化及びエントロピ符号化において処理し、前記切り替え、量子化、及び／又はエントロピ符号化の制御は、前記入力信号の心理音響分析に由来し、対応する時間分解能制御情報を符号化出力信号に対して副次的情報として添付し、前記第一及び第二の変換の各々の前に、前記第一の長さ又は前記第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ又は第二の長さの区間にに対する重複加算処理が加えられ、前記第一の長さ及び第二の長さの区間の間の境界における遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記第二の長さの区間にに対する、開始及び停止窓関数が使用された、符号化信

40

50

号を復号化する装置であって、

前記符号化信号から前記副次的情報を提供すること及び前記符号化信号を逆量子化及びエントロピ復号化することに適した手段と、

前記副次的情報に対応して、前記逆量子化及びエントロピ復号化信号の第一の長さの信号区間に作用し、復号化信号を提供する、時間領域への第一の逆M D C T又は逆整数M D C T又は逆D C T - 4変換を実行するか、或いは、前記第一の逆M D C T又は逆整数M D C T又は逆D C T - 4変換を実行する前に、第二の逆M D C T又は逆整数M D C T又は逆D C T - 4変換において、前記逆量子化及びエントロピ復号化信号の第二の長さの区間を処理するのに適した手段と、

を含み、

10

前記第一及び第二の逆変換後に、前記第一の長さ及び前記第二の長さの区間の振幅値は、窓関数を使用して加重され、前記第一の長さ及び第二の長さの区間にに対する重複加算処理が加えられ、遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記加重のために前記第二の長さの区間にに対して、開始及び停止窓関数が使用される、前記装置。

【請求項 8】

二つ以上の異なる第二の長さが使用される場合、利用する異なる第二の長さのトポロジを信号伝達するために、変化する時間分解能の領域を示す幾つかの指數、或いは復号化側でアクセス可能な対応するコードブックにおける一致するエントリを示す指數番号が、前記副次的情報に含まれる、請求項6又は7記載の装置。

【請求項 9】

20

二つ以上の異なる第二の長さが連続して使用される場合、長さは、低周波数線を表す周波数ビンを開始点として増加する、請求項6乃至8の何れかに記載の装置。

【請求項 10】

前記トポロジは、

選択周波数帯域に対して、変換ビンのスペクトル電力を決定し、前記スペクトル電力の算術平均値をその幾何平均値により割ることで、前記第一のM D C T又は整数M D C T又はD C T - 4変換を使用してスペクトル平坦性基準S F Mを実行するステップと、

非加重入力信号区間をサブセグメント化し、m個のサブセクションに対して加重及び短変換を実行し、こうした変換の周波数分解能は、前記選択周波数帯域に対応するステップと、

30

m個の変換セグメントからなる各周波数線に対して、スペクトル電力を決定し、m個のセグメントの幾何平均により割った算術平均を決定することで時間平坦性基準T F Mを計算するステップと、

S F M値を使用することで、音調又はノイズ周波数帯域を決定するステップと、

こうした帯域内の時間的変化を認識するために、T F M値を使用し、閾値を使用して、前記決定されたノイズ周波数帯域のために、更に細かな時間分解能への切り替えを行うステップと、により決定される、請求項8記載の装置。

【請求項 11】

請求項1に記載の方法により符号化された符号化デジタル音声信号が副次的情報と共に記録された、記憶媒体。

40

【請求項 12】

請求項1に記載の方法であって、音声信号を符号化するために透かし埋め込み器により実行され、透かしの埋め込みにおいて透かしが埋め込まれる、前記方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、変換符号化及びスペクトル領域における時間分解の適応切り替えを使用して音声信号を符号化及び復号化する方法及び装置に関する。

【背景技術】

【0002】

50

知覚的音声コーデックでは、音声信号のコンパクトな表現、即ち、冗長性の低減を達成し、元の音声信号との無関係性を減少できるように、フィルタ・バンク及びM D C T (Modified Discrete Cosine Transform: 修正離散コサイン変換、順変換)を使用する。音声信号の準定常部分中は、高いコーディング・ゲインを達成するために、フィルタ・バンクの高周波又はスペクトル分解能が有利となるが、こうした高周波分解能は、遷移信号部分中に問題となる粗い時間分解に結びつく。周知の結果は、可聴のプリエコー効果である。

【0003】

B.エドラー (Edler) 氏, "Codierung von Audiosignalen mit ueberlappender Transformation und adaptiven Fensterfunktionen", Frequenz, Vol.43, No.9, p.252-256, September 1989 では、時間領域における適応的窓切り替え、及び / 又は、長さの異なる二種類の窓関数を交互に使用することによる二種類の分解能の切り替えである変換長切り替えを開示している。米国特許第6,029,126号公報では、行列乗算を使用してスペクトル帯域を組み合わせることで時間分解能を増加させる長変換について開示している。異なる固定分解能間の切り替えは、時間領域での窓切り替えを避けるために実行される。これを利用して、二種類の分解能を有する不均一なフィルタ・バンクを形成できる。

10

【0004】

国際公開第03/019,532号パンフレットでは、多相フィルタ・バンクの構築に適した非常に複雑なフィルタ設計法である、コサイン変調フィルタ・バンクでのサブ帯域の結合について開示している。

20

【発明の開示】

【発明が解決しようとする課題】

【0005】

エドラー氏が開示した上記の窓及び / 又は変換長切り替えは、先読みが長いことによる遅延の長さと、無関係性の最適な減少に十分な分解能の提供を妨げる、短ブロックの周波数分解能の低さとのため、最適なものとは言えない。

【0006】

本発明により解決する課題は、遷移音声信号部分に対して、高周波分解及び高時間分解を適用して、改良されたコーディング / デコーディング・ゲインを提供することである。

30

【課題を解決するための手段】

【0007】

この課題は、請求項1及び3において開示した方法により解決される。こうした方法を利用する装置は、請求項2及び4において開示される。

【0008】

本発明は、第一のフィルタ・バンクの出力の上に、第二の不均一フィルタ・バンク、即ち、カスケードM D C Tを適用することで、改良されたコーディング / デコーディング品質を達成する。本発明のコーデックは、遷移又は高速変化音声信号区間中に、時間周波数表現を再編成するために、付加的な延長フィルタ・バンク (又は多重分解フィルタ・バンク)への切り替えを使用する。

40

【0009】

対応する切り替え制御を適用することで、プリエコー効果を回避し、高いコーディング・ゲインが達成される。本発明のコーデックは、低いコーディング遅延 (先読み無し) という利点を有する。

【0010】

原則的に、本発明の符号化方法は、入力信号、例えば、音声信号を、入力信号の第一の長さの区間に適用される周波数領域への第一の順変換を使用し、時間分解能の適応切り替えを使用し、その後、結果的に生じた周波数領域ビンの値の量子化及びエントロピ符号化を使用して符号化するのに適しており、切り替え、量子化、及び / 又はエントロピ符号化の制御は、入力信号の心理音響分析に由来し、

50

- ・第一の順変換に続いて、変換済みの第一の長さの区間の第二の長さの区間に適用される第二の順変換を実行することで達成される時間分解能を適応的に制御するステップであり、第二の長さは、第一の長さより短く、第一の順変換の出力値又は第二の順変換の出力値の何れかを量子化及びエントロピ符号化において処理する、時間分解能を適応的に制御するステップと、
- ・符号化出力信号に対して、対応する時間分解能制御情報を副次的情報として添付するステップと、を含む方法である。

【0011】

原則的に、本発明の符号化装置は、入力信号、例えば、音声信号を符号化するのに適しており、

10

- ・入力信号の第一の長さの区間を周波数領域へ変換するのに適した第一の順変換手段と、
- ・変換済みの第一の長さの区間の、第一の長さより短い第二の長さの区間を変換するのに適した第二の順変換手段と、
- ・第一の順変換手段の出力値又は第二の順変換手段の出力値を量子化及びエントロピ符号化するのに適した手段と、
- ・量子化及び／又はエントロピ符号化を制御すると共に、第一の順変換手段の出力値又は第二の順変換手段の出力値が量子化及びエントロピ符号化手段において処理されるかについて、適応的に、入力信号の心理音響分析に由来する制御を行うのに適した手段と、
- ・符号化装置出力信号に対して、対応する時間分解能制御情報を副次的情報として添付するのに適した手段と、を含む装置である。

20

【0012】

原則的に、本発明の復号化方法は、入力信号の第一の長さの区間に適用される周波数領域への第一の順変換を使用して符号化された符号化信号、例えば、音声信号であり、第一の順変換に続いて、変換済みの第一の長さの区間の第二の長さの区間に適用される第二の順変換を実行することで時間分解能を適応的に切り替え、第二の長さは、第一の長さより短く、第一の順変換の出力値又は第二の順変換の出力値の何れかを量子化及びエントロピ符号化において処理し、切り替え、量子化、及び／又はエントロピ符号化の制御は、入力信号の心理音響分析に由来し、対応する時間分解能制御情報を符号化出力信号に対して副次的情報として添付した符号化信号を復号化するのに適しており、

30

- ・符号化信号から副次的情報を提供するステップと、
- ・符号化信号を逆量子化及びエントロピ復号化するステップと、
- ・副次的情報に対応して、逆量子化及びエントロピ復号化信号の第一の長さの信号区間に作用し、復号化信号を提供する、時間領域への第一の順逆変換を実行するか、或いは、
- ・第一の順逆変換を実行する前に、第二の順逆変換において、逆量子化及びエントロピ復号化信号の第二の長さの区間を処理するステップと、を含む復号化方法である。

【0013】

原則的に、本発明の復号化装置は、入力信号の第一の長さの区間に適用される周波数領域への第一の順変換を使用して符号化された符号化信号、例えば、音声信号であり、第一の順変換に続いて、変換済みの第一の長さの区間の第二の長さの区間に適用される第二の順変換を実行することで時間分解能を適応的に切り替え、第二の長さは、第一の長さより短く、第一の順変換の出力値又は第二の順変換の出力値の何れかを量子化及びエントロピ符号化において処理し、切り替え、量子化、及び／又はエントロピ符号化の制御は、入力信号の心理音響分析に由来し、対応する時間分解能制御情報を符号化出力信号に対して副次的情報として添付した符号化信号を復号化するのに適しており、

40

- ・副次的情報を提供すること及び符号化信号を逆量子化及びエントロピ復号化することに適した手段と、
- ・副次的情報に対応して、逆量子化及びエントロピ復号化信号の第一の長さの信号区間に作用し、復号化信号を提供する、時間領域への第一の順逆変換を実行するか、或いは、
- ・第一の順逆変換を実行する前に、第二の順逆変換において、逆量子化及びエントロピ復号化信号の第二の長さの区間を処理するステップと、を含む装置である。

50

【0014】

本発明の有利な追加実施形態は、それぞれの従属請求項において開示する。

【発明を実施するための最良の形態】

【0015】

以下、本発明の実施形態例を添付図面を参照して説明する。

【0016】

図1において、コーダ入力音声信号CISの各連続重複ブロック又はセグメント又は区間の大きさの値は、窓関数により加重され、対応する変換係数又は周波数ビンを提供する、長(即ち、高周波分解能)MDCTフィルタ・バンク又は変換段階又はステップMDCT-1において変換される。遷移音声信号区間中には、周波数及び時間フィルタ分解能を変更するために、短固定変換長を備えるか、或いは、異なる短変換長を有する多重分解能MDCTフィルタ・バンクであることが好ましい第二のMDCTフィルタ・バンク又は変換段階又はステップMDCT-2を、第一の順変換(即ち同一ブロック上)の周波数ビンに適用し、即ち、一連の不均一MDCTを周波数データに適用し、これにより、不均一な時間/周波数表現を生成する。第一の順変換の周波数ビンの各連続重複区間の振幅値は、第二段階の変換の前に窓関数により加重する。加重に使用される窓関数は、図4乃至図7及び式(3)及び(4)に関連して説明する。MDCT又は整数MDCT変換の場合、区間は、50%重複する。様々な変換が使用される場合、重複の度合いは異なる可能性がある。

【0017】

段階またはステップMDCT-2に二種類の異なる長さのみが使用される場合、そのステップ又は段階は、単独で考えた場合、上記のエドラー氏のコーデックと同様である。

【0018】

第二のMDCTフィルタ・バンクMDCT-2のオン又はオフの切り替えは、第一及び第二のスイッチSW1及びSW2を使用して実行可能であり、心理音響分析器段階又はステップPSYMと統合された、あるいは並行して動作する、フィルタ・バンク制御部またはステップFBCTLにより制御され、PSYMとFBCTLとは共に信号CISを受信する。段階又はステップPSYMは、入力信号CISからの時間又は空間情報を使用する。第二段階フィルタMDCT-2のトポロジ又は状態は、副次的情報としてコード化され、コーダ出力ビットストリームCOSとなる。スイッチSW2から出力された周波数データは、心理音響分析器PSYM、特に、量子化ステップのサイズにより制御される、量子化器及びエントロピ符号化段階又はステップQUCODにおいて量子化及びエントロピ符号化される。段階QUCODからの出力(符号化周波数ビン)及びFBCTLからの出力(トポロジ又は状態情報又は時間分解能制御情報又は切り替え情報SWI又は副次的情報)は、ストリーム・パッカー・ステップ又は段階STRPCKにおいて結合され、出力ビット・ストリームCOSを形成する。

【0019】

量子化は、ひずみ信号の挿入に置き換えることができる。図2において、デコーダ側では、デコーダ入力ビット・ストリームDISのデパッキングと、これに対応する復号化及び逆「量子化」(又は再量子化)とが、デパッキング、復号化、及び再量子化段階又はステップDPCRQUにおいて行われ、これによりDPCRQUは、復号化周波数ビン及び切り替え情報SWIを提供する。対応する逆不均一MDCTステップ又は段階iMDCT-2は、切り替え情報SWIを介してビットストリームにより信号が送られた場合、例えば、スイッチSW3及びSW4を使用して、こうした符号化復号化周波数ビンに適用される。逆変換値の各連続区間の振幅値は、ステップ又は段階iMDCT-2での変換に続いて、窓関数により加重され、加重後、重複加算処理が行われる。信号は、復号化周波数ビンに対して、或いは、ステップ又は段階iMDCT-2の出力に対して、対応する逆高分解能MDCTステップ又は段階iMDCT-1を適用することで再構築される。逆変換値の各連続区間の振幅値は、ステップ又は段階iMDCT-1での変換に続いて、窓関数により加重され、加重後、重複加算処理が行われる。その後、PCM音声デコーダ出力信号

10

20

30

40

50

DOSとなる。復号化側で適用された変換長は、対応する移動長を鏡写ししたものとなり、即ち、受領した値の同じブロックが二度逆変換される。加重に使用される窓関数は、図4乃至図7及び式(3)及び(4)に関連して説明する。逆MDCT又は逆整数MDCT変換の場合、区間は、50%重複する。様々な逆変換が使用される場合、重複の度合いは異なる可能性がある。

【0020】

図3は、上記の処理、即ち、第一及び第二の段階のフィルタ・バンクの適用を示している。左側では、時間領域サンプルのブロックは、長MDCTにおいて、窓処理され、周波数領域へ変換される。遷移音声信号区間中には、一連の不均一MDCTが周波数データに適用され、図3の右側に示した不均一な時間／周波数表現が生成される。時間／周波数表現は、灰色又は斜線で表示する。

10

【0021】

第一の段階の変換又はフィルタ・バンクMDCT-1の時間／周波数表現（左側）は、定常信号区間を符号化するのに最適な高周波又はスペクトル分解能を提供する。フィルタ・バンクMDCT-1及びiMDCT-1は、50%の重複ブロックを有する一定サイズのMDCT及びiMDCTのペアを表す。フィルタ・バンクiMDCT-1においてオーバレイ及び加算(OLA)を使用して、時間領域のエイリアスをキャンセルする。そのため、フィルタ・バンクペアMDCT-1及びiMDCT-1は、理論上完全な再構築が可能となる。高速で変化する信号区間、特に遷移信号は、人間の知覚に一致する分解能、あるいは時間／周波数に同調させた最大信号圧縮を表す分解能により、時間／周波数において良好に表現される。これは、第一の順変換フィルタ・バンクMDCT-1において選択された周波数ビンのブロックに対して、第二の変換フィルタ・バンクMDCT-2を適用することで達成される。

20

【0022】

第二の順変換は、図3の中間部分に示したように、50%重複したサイズの異なる窓を使用すること、及び、あるサイズから別のサイズへ切り替える時に、遷移窓関数（即ち、非対称の傾斜をそれぞれ有する「エドラー窓関数」）を使用することを特徴とする。窓サイズは、長さ4から始まり、長さ 2^n まであり、nは、2より大きな整数である。窓サイズ「4」は、二個の周波数ビンと、二倍の時間分解能を組み合わせ、窓サイズ 2^n は、 $2^{(n-1)}$ 個の周波数ビンを組み合わせ、時間分解能を $2^{(n-1)}$ 倍させる。MDCTの連続の先頭及び終わりでは、特殊な開始及び停止窓関数（遷移窓）を使用する。復号化側において、フィルタ・バンクiMDCT-2は、OLAを含む逆変換を適用する。これにより、フィルタ・バンクペアMDCT-2/iMDCT-2は、理論上完全な再構築が可能となる。フィルタ・バンクMDCT-2の出力データは、フィルタ・バンクMDCT-2を適用する時には含まれていなかった、フィルタ・バンクMDCT-1の単一分解能のビンと組み合わせる。

30

【0023】

フィルタ・バンクMDCT-2の各変換又はMDCTの出力は、第一の順変換の結合周波数ビンの時間反転させた時間サンプルであると解釈できる。これにより、図3の右側に示したような不均一な時間／周波数表現の構築が実現可能になるという利点が生じる。

40

【0024】

フィルタ・バンク制御部又はステップFBCTLは、心理音響分析器段階又はステップPSYMにおける心理音響モデルからの時間データ及び励起パターンを使用して、実際の処理ブロックの信号分析を実行する。簡略化した実施形態において、FBCTLは、遷移信号区間中、フィルタ・バンクMDCT-2の固定フィルタトポロジに切り替わり、このフィルタ・バンクは、人間の知覚の時間／周波数分解能を利用し得る。所望のフィルタ・バンクMDCT-2のトポロジをコードブックのエントリとして復号化側へ信号で伝えるには、ほんの僅かなビット数の副次的情報のみで済むという利点がある。

【0025】

更に複雑な実施形態において、フィルタ・バンク制御部又はステップFBCTLは、入

50

力信号 C I S のスペクトル及び時間平坦性を評価し、柔軟性のあるフィルタ・バンク M D C T - 2 のフィルタトポロジを決定する。本実施形態では、フィルタ・バンク i M D C T - 2 の構築を可能にするためには、開始窓、遷移窓、及び停止窓位置のコード化開始位置をデコーダに対して送信することで十分となる。

【0026】

心理音響モデルは、フィルタ・バンク M D C T - 1 の分解能に等しい高スペクトル分解能を利用すると同時に、スペクトルは粗いが、時間分解能の高い信号分析を利用する。この第二の分解能は、フィルタ・バンク M D C T - 2 のもっとも粗い周波数分解能に一致させることができる。

【0027】

別の方法として、心理音響モデルは、フィルタ・バンク M D C T - 1 の出力により直接駆動可能であり、遷移信号区間中は、フィルタ・バンク M D C T - 2 の適用に続いて、図 3 の右側に示した時間 / 周波数表現により駆動できる。

【0028】

以下、更に詳細なシステムの説明を行う。

【0029】

M D C T

修正離散コサイン変換 (M D C T) 及び逆 M D C T (i M D C T : i n v e r s e M D C T) は、臨界サンプリング・フィルタ・バンクを表すものと考えることができる。M D C T は、当初、J.P. プリンセン (Princen) 氏及びA.B. ブラッドレー (Bradley) 氏により、"Analysis/synthesis filter bank design based on time domain aliasing cancellation" , IEEE Transactions on Acoust. Speech Sig. Proc. ASSP-34 (5) , pp.1153-1161 , 1986 において「Oddly-stacked time domain alias cancellation transform」と命名された。

【0030】

H.S. マルバー (Malvar) 氏, "Signal processing with lapped transform" , Artech House Inc. , Norwood , 1992 、及び M. テメリナック (Temerinac) 氏 , B. エドラー (Edler) 氏, "A unified approach to lapped orthogonal transforms" , IEEE Transactions on Image Processing , Vol.1, No.1, pp.111-116 , January 1992 では「Modulated Lapped Transform (MLT) 」と呼び、重複直交変換との全般的な関係を示すと共に、Q M F フィルタ・バンクの特殊なケースであることを証明した。

【0031】

変換及び逆変換の式を式 (1) 及び (2) に記載する。

【数1】

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

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

【0032】

こうした変換において、50% 重複ブロックが処理される。符号化側では、何れの場合も、N 個のサンプルのブロックが窓処理され、大きさの値が窓関数 h (n) により加重され、その後、K = N / 2 個の周波数ビンに変換され、ここで N は整数となる。復号化側では、逆変換により、M 個の周波数ビンが N 個の時間サンプルに、その都度変換され、その後、大きさの値が窓関数 h (n) により加重され、ここで N 及び M は整数となる。その後のオーバレイ加算手順により、時間エイリアスをキャンセルする。窓関数 h (n) は、完全な再構築を可能にするために、幾つかの制約を満たす必要がある。式 (3) 及び (4) を参照されたい。

10

20

30

40

50

【数2】

$$h^2(n + N/2) + h^2(n) = 1 \quad (3)$$

$$h(n) = h(N - n - 1) \quad (4)$$

【0033】

分析及び合成窓関数は異なるものであってよいが、復号化において使用される逆変換長は、符号化で使用された変換長に対応する。しかしながら、この選択肢は、ここでは考慮しない。適切な窓関数は、(5)に記載のサイン窓関数である。

【数3】

10

$$h_{\sin}(n) = \sin\left(\pi \cdot \frac{n+0.5}{N}\right), \quad n=0 \dots N-1 \quad (5)$$

【0034】

上記の論文において、エドラー氏は、遷移窓を使用してM D C T時間周波数分解能を切り替えることを示した。遷移窓1、10を使用した長変換から八個の短変換への切り替え(遷移条件により生じる)の例を、図4の下部に示しており、ここでは、窓関数のゲインGを垂直方向に、時間、即ち、入力信号サンプルを水平方向に示している。図4の上部は、定常状態の状況において適用される三つの連続する基本窓関数A、B、及びCを示している。

20

【0035】

遷移窓関数は、長変換の長さ N_L を有する。小さい窓側の端部には、 r 個のゼロ振幅窓関数サンプルが存在する。 $N_L/2$ に位置する窓関数の中心に向かい、小変換(N_{short} サンプルの長さを有する)用の鏡写しの半分の窓関数が続き、更に、「1」(又は「单一」定数)の値を有する r 個の窓関数サンプルが続く。図5の左側では、短い窓へ向かう遷移について、図5の右側では短い窓からの遷移について、原理を示している。値 r は、次のように表される。

【数4】

$$r = (N_L - N_{short})/4 \quad (6)$$

30

【0036】

多重分解能フィルタ・バンク

第一段階のフィルタ・バンクM D C T-1、i M D C T-1は、例えば15乃至25Hzのサブバンド・フィルタ帯域幅を有する高分解能M D C Tフィルタ・バンクである。例えば32乃至48kHzの音声サンプリング・レートに対して、 N_L の通常の長さは2048サンプルとなる。窓関数 $h(n)$ は、式(3)及び(4)を満たす。フィルタM D C T-1の適用後、好適な実施形態では、1024個の周波数ビンが存在する。定常入力信号区間において、こうしたビンは、心理音響的な考慮に従って量子化される。

【0037】

40

高速変化する遷移入力信号区間は、第一のM D C Tのビンに適用される追加のM D C Tにより処理する。この追加ステップ又は段階では、図3の右側に示したように、2、4、8、16、又はそれより多くのサブバンドを融合し、これにより時間分解能を増加させる。

【0038】

図6は、周波数領域内での第二段階のM D C T用に加えた一連の窓処理の例を示している。したがって、水平軸線は、 f / ビンに関連する。遷移窓関数は、時間領域と同様に、図5及び式(6)にしたがって設計されている。特殊な開始窓関数S T W及び停止窓関数S P Wは、変換信号の開始及び終了区画、即ち、最初及び最後のM D C Tを扱う。こうした開始及び停止窓関数の設計原理を図7に示す。こうした窓関数の一方の半分は、通常又

50

は標準の窓関数 NW 、例えば、式(5)によるサイン窓関数の半分の窓関数を鏡写しにしたものである。こうした窓関数の他方の半分では、隣接する半分が「1」(又は「单一」定数)の連続的なゲインを有し、他の半分はゼロのゲインを有する。

【0039】

M D C T の特性から、M D C T - 2 の実行は、部分的な逆変換と見なすこともできる。第二段階のM D C T の順M D C T を適用する時、こうした新たなM D C T (M D C T - 2)のそれぞれは、元の窓処理済みビンを結合させた新たな周波数線(ビン)とみなすことが可能であり、その新たなM D C T の時間反転出力は、新たな時間ブロックと見なすことができる。図8及び図9の提示は、この仮定又は条件に基づいている。

【0040】

図6の指數 k_i は、変化する時間分解能の領域を示す。位置ゼロから始まる位置 $k_1 - 1$ までの周波数ビンは、单一の時間分解能に対応する、第一の順変換 (M D C T - 1) からコピーされる(即ち、第一の順変換を表す)。

【0041】

指數 $k_1 - 1$ から指數 k_2 のビンは、 g_1 の周波数線に変換される。 g_1 は、実行された変換数に等しい(重複する窓の数に対応する数であり、第二又は上位の変換レベルM D C T - 2 における周波数ビン数とみなすことができる)。図6において、指數 k_1 は第一の順変換の第二のサンプルとして選択されているため、開始指數は、ビン $k_1 - 1$ となる(第一のサンプルは振幅ゼロを有する。図10aを参照)。

【0042】

$g_1 = (\text{窓処理済みビン数}) / (N / 2) - 1 = (k_2 - k_1 + 1) / 2 - 1$ であり、例えば、標準の窓サイズ N が4ビンであれば、このサイズにより、二倍の時間分解能を備えた区間が形成される。

【0043】

指數 $k_2 - 3$ から指數 $k_3 + 4$ のビンは、 g_2 の周波数線(変換)に結合され、即ち、 $g_2 = (k_3 - k_2 + 2) / 4 - 1$ となる。例えば、標準の窓サイズが8ビンであれば、このサイズにより、四倍の時間分解能を備えた区間が生じる。

【0044】

図6において、次の区間は、例えば16ビンに渡る窓(変換長)により変換され、このサイズにより、八倍の時間分解能を有する区間が生じる。窓処理は、ビン $k_3 - 5$ で開始される。これが選択された最後の解像度である場合(図6では当てはまる)、窓処理はビン $k_4 + 4$ で終了し、そうでない場合はビン k_4 で終了する。

【0045】

低周波数線に対応する周波数ビンから始まる連続変換ブロックにおいて、第二段階の変換の順序(即ち、長さ)が可変である場合、第一の第二段階M D C T は、小さな順位数で開始し、後続の第二段階M D C T は、高い順位数を有する。完全な再構築の特性を満たす遷移窓を使用する。

【0046】

図6による処理を図10において更に説明する。図10は、より良好な時間分解能を達成する第二の(即ち、カスケード)変換(M D C T - 2)の範囲を示す周波数指數のサンプル精度の割り当てを示している。円は、第一又は初期変換(M D C T - 1)のビン位置、即ち、周波数線を表す。

【0047】

図10aは、二倍の時間分解能を提供するために使用された四点の第二段階M D C T の範囲を示している。図示した五つのM D C T 区間は、五本の新たなスペクトル線を形成する。図10bは、四倍の時間分解能を提供するために使用された八点の第二段階M D C T の範囲を示している。三つのM D C T 区間を示している。図10cは、八倍の時間分解能を提供するために使用された16点の第二段階M D C T の範囲を示している。四つのM D C T 区間を示している。

【0048】

10

20

30

40

50

デコーダ側において、定常信号は、フィルタ・バンク $i M D C T - 1$ 、即ち、時間エイリアスをキャンセルするためにオーバレイ加算手順 (OLA) を含む長変換ブロックの $i M D C T$ を使用して復元される。ビット・ストリームにおいて信号で指示された場合、復号化又はデコーダは、それぞれ、フィルタ・バンク $i M D C T - 1$ を適用する前に、信号伝達されたトポロジ (OLA を含む) に従った一連の $i M D C T$ を適用することで、多重分解能フィルタ・バンク $i M D C T - 2$ へ切り替わる。

【0049】

フィルタ・バンク・トポロジのデコーダへの信号伝達

最も単純な実施形態では、フィルタ・バンク $M D C T - 2 / i M D C T - 2$ 用の単一の固定トポロジを利用し、伝送ビット・ストリーム中の単一ビットにより、これを信号伝達する。トポロジの固定セットを更に使用する場合、対応するビット数を使用して、現在使用されているトポロジの一つを信号伝達する。更に進んだ実施形態では、固定コードブック・トポロジのセットから最善のものを取り出し、対応するコードブック・エントリをビット・ストリーム内部で信号伝達する。

【0050】

第二段階の変換のフィルタ・トポロジが固定ではない実施形態において、対応する副次的情報が符号化出力ビット・ストリームにおいて送信される。好ましくは、指数 $k_1, k_2, k_3, k_4, \dots, k_{end}$ が送信される。

【0051】

四倍の分解能で開始され、 k_2 は、ゼロに等しい k_1 と同じ値で送信される。最大の時間分解能より粗い時間分解能で終了するトポロジにおいて、 k_{end} で送信される値は、 k_4, k_3, \dots にコピーされる。

【0052】

次の表では、これを幾つかの例により示している。 b_i は、値として周波数ビンの代わりになるものである。

【表1】

トポロジ	トポロジを信号伝達する指数				
	k_1	k_2	k_3	k_4	k_{end}
1倍、2倍、4倍、8倍、16倍の時間分解能を備えたトポロジ	$b_1 > 1$	b_2	b_3	b_4	b_5
1倍、2倍、4倍、8倍の時間分解能を備えたトポロジ (図6と同様)	$b_1 > 1$	b_2	b_3	b_4	b_4
8倍の時間分解能を備えたトポロジ	0	0	0	b_{max}	b_{max}
4倍、8倍、16倍の時間分解能を備えたトポロジ	0	0	b_2	b_3	b_{max}

【0053】

人間の聴覚系の時間的心理音響特性のため、信号伝達は、周波数と共に増加する時間分解能を備えたトポロジに限定しても十分となる。

10

20

30

40

50

【0054】

フィルタ・バンク・トポロジの例

図8及び9は、第二段階のフィルタ・バンクの多重分解能T/F(時間/周波数)エネルギープロット二例を示している。図8は、「8倍時間分解能のみ」のトポロジを示している。図8aにおける時間領域信号の遷移は、経時的な振幅として示している(時間はサンプル単位で表現)。図8bは、第一段階のMDCTの対応するT/Fエネルギー・プロット(一変換ブロックに対応する標準化時間に渡るビン内での周波数)、図8cは、第二段階のMDCTの対応するT/Fプロット(8×128時間周波数タイル)を示している。

【0055】

図9は、「1倍、2倍、4倍、8倍のトポロジ」を示す。図9aにおける時間領域信号の遷移は、経時的な振幅として示している(時間はサンプル単位で表現)。図9bは、第二段階のMDCTの対応するT/Fプロットを示しており、これにより、低帯域部分の周波数分解能は、人間の聴覚系における知覚の帯域幅(臨界帯域)に比例して選択され、合計1024の係数について、 $b_{N1} = 16$ 、 $b_{N2} = 16$ 、 $b_{N4} = 16$ 、 $b_{N8} = 114$ である(これらの数は以下の意味を有する。16本の周波数線が单一の時間分解能を有し、16本の周波数線が二倍、16本の周波数線が四倍、114本の周波数線が八倍の時間分解能を有する)。低周波数には単一の区画が存在し、その後、二区画、四区画となり、約 $f = 50$ より上では、八区画となる。

【0056】

フィルタ・バンクの制御

最も単純な実施形態では、任意の最新の遷移検出器を使用して、人間の知覚のT/F分解能に一致する或いは近づく固定トポロジへの切り替えを行う。好適な実施形態では、更に進んだ制御処理を使用する。

- スペクトル平坦性基準SFMを、例えば式(7)により、電力スペクトル密度PmのM周波数線(f_{block})の選択帯域に渡って、 N_L 個のサンプル、即ち、MDCT-1の長さを有する長変換ブロックの窓処理信号の離散型フーリエ変換(DET)を使用して計算する(選択帯域は臨界帯域に比例する)。

- N_L 個のサンプルの分析ブロックを、S-8の重複ブロックに分割し、S回の窓処理DFTをサブブロックに加える。S列(時間分解能、 t_{block})と、各DFTの周波数線数に従った行数とを有する行列に結果を配置し、Sは整数とする。

- SのスペクトログラムPs、例えば、電力スペクトル密度又は心理音響的に成形したスペクトログラム(又は励起パターン)を計算する。

- 各周波数線について、式(8)により時間平坦性基準(TFM)を決定する。

- SFMベクトルを使用して、音調又はノイズ帯域を決定し、TFMベクトルを使用して、こうした帯域内での時間的变化を認識する。閾値を使用して、多重分解能フィルタ・バンクへの切り替えを行うか、どのトポロジを選び出すかを決定する。

$$SFM = \text{算術平均値} [f_{block}] / \text{幾何平均値} [f_{block}]$$

【数5】

$$= \frac{1}{M} \cdot \sum_m Pm \left/ \left(\prod_M Pm \right) \right. \frac{1}{M} \quad (7)$$

$$TFM = \text{算術平均値} [t_{block}] / \text{幾何平均値} [t_{block}]$$

【数6】

$$= \frac{1}{S} \cdot \sum_s Ps \left/ \left(\prod_s Ps \right) \right. \frac{1}{S} \quad (8)$$

【0057】

異なる実施形態において、トポロジは、以下のステップにより決定される。

10

20

30

40

50

- 選択周波数帯域に対して、変換ビンのスペクトル電力を決定し、スペクトル電力値の算術平均値をその幾何平均値により割ることで、第一の順変換を使用してスペクトル平坦性基準 S F M を実行する。

- 非加重入力信号区間をサブセグメント化し、m 個のサブセクションに対して加重及び短変換を実行する。ここで、こうした変換の周波数分解能は、選択周波数帯域に対応する。

- m 個の変換セグメントからなる各周波数線に対して、スペクトル電力を決定し、m 個のセグメントの幾何平均により割った算術平均を決定することで時間平坦性基準 T F M を計算する。

- S F M ベクトルを使用することで、音調又はノイズ帯域を決定する。 10

- こうした帯域内での時間的变化を認識するために、T F M 値を使用する。閾値を使用して、指示されたノイズ周波数帯域のために、更に細かな時間分解能への切り替えを行う。

【 0 0 5 8 】

M D C T は、D C T 、特にD C T - 4 に置き換えることができる。本発明は、音声信号に適用する代わりに、対応する形で映像信号にも適用され、その場合、心理音響分析器 P S Y M は、人間の視覚系の特性を考慮する分析器に置き換える。

【 0 0 5 9 】

本発明は、透かし埋め込み器において使用可能である。直接的な埋め込みと比較して、本発明の多重分解能フィルタ・バンクを使用して音声又は映像信号にデジタル透かし情報埋め込むことの利点は、透かし情報の伝送及び受信側での透かし情報の検出における堅牢性の増加である。 20

【 0 0 6 0 】

本発明の一実施形態において、カスケード・フィルタ・バンクは、音声透かしシステムで使用される。透かしエンコーダでは、第一の（整数）M D C T を実行する。第一の透かしは、心理音響的に制御された埋め込み処理を使用して、ビン 0 乃至 k 1 - 1 に挿入される。この透かしの目的は、透かしデコーダでのフレーム同期にすることができる。第二段階の可変サイズ（整数）M D C T は、上記のように、ビン指數 k 1 から開始するビンに適用される。この第二の段階の出力は、出力を時間反転時間ブロックとして解釈し、各第二段階 M D C T を新たな周波数線（ビン）として解釈することで、時間周波数表現を得るために再分類される。第二の透かし信号は、心理音響的な考慮により制御される減衰係数を使用することで、こうした新たな周波数線のそれぞれに追加される。データは再分類され、逆（整数）M D C T （上記の第二段階 M D C T に関する）は、窓処理及びオーバレイ／加算を含め、上記の実施形態（デコーダ）について説明したように実行される。第一の順変換に関するスペクトル全体が復元される。そのデータ、窓処理、及びオーバレイ／加算に対して実行されるフルサイズの逆（整数）M D C T により、透かしが埋め込まれた時間信号が復元される。 30

【 0 0 6 1 】

多重分解能フィルタ・バンクは、透かしデコーダ内でも使用される。ここでは、第二段階 M D C T のトポロジは、アプリケーションにより固定される。 40

本発明は以下の態様を含む。

（付記 1 ）

入力信号（C I S）、例えば、音声信号を、前記入力信号の第一の長さ（N_L）の区間に適用される周波数領域への第一の順変換（M D C T - 1）を使用し、時間分解能の適応切り替えを使用し、その後、結果的に生じた周波数領域ビンの値の量子化及びエントロピ符号化（Q U C O D）を使用して符号化する方法であって、

前記切り替え、量子化、及び／又はエントロピ符号化の制御（P S Y M、F B C T L）は、前記入力信号の心理音響分析に由来し、

前記第一の順変換（M D C T - 1）に続いて、前記変換済みの第一の長さの区間の第二の長さ（N_{short}）の区間に適用される第二の順変換（M D C T - 2）を実行することで 50

、前記時間分解能の適応的な制御 (S W 1 、 S W 1 、 S W I) を行うステップであり、前記第二の長さは、前記第一の長さ (N_L) より短く、前記第一の順変換の出力値又は前記第二の順変換の出力値の何れかを前記量子化及びエントロピ符号化 (Q U C O D) において処理する、時間分解能を適応的に制御するステップと、

符号化出力信号 (C O S) に対して、対応する時間分解能制御情報 (S W I) を副次的情報として添付 (S T R P C K) するステップと、
を含む、前記方法。

(付記 2)

入力信号 (C I S) 、例えば、音声信号を符号化する装置であって、
前記入力信号の第一の長さ (N_L) の区間を周波数領域へ変換するのに適した第一の順 10
変換手段 (M D C T - 1) と、

前記変換済みの第一の長さの区間の、前記第一の長さ (N_L) より短い第二の長さ (N_{short}) の区間を変換するのに適した第二の順変換手段 (M D C T - 2) と、

前記第一の順変換手段の出力値又は前記第二の順変換手段の出力値を量子化及びエントロピ符号化するのに適した手段 (Q U C O D) と、

前記量子化及び / 又はエントロピ符号化を制御すると共に、前記第一の順変換手段の出力値又は前記第二の順変換手段の出力値が前記量子化及びエントロピ符号化手段において処理されるかについて、適応的に、前記入力信号の心理音響分析に由来する制御を行うのに適した手段と、

符号化装置出力信号 (C O S) に対して、対応する時間分解能制御情報 (S W I) を副次的情報として添付するのに適した手段 (S T R P C K) と、
を含む、前記装置。

(付記 3)

入力信号の第一の長さ (N_L) の区間に適用される周波数領域への第一の順変換 (M D C T - 1) を使用して符号化された符号化信号 (D I S) 、例えば、音声信号であり、前記第一の順変換 (M D C T - 1) に続いて、前記変換済みの第一の長さの区間の第二の長さ (N_{short}) の区間に適用される第二の順変換 (M D C T - 2) を実行することで時間分解能を適応的に切り替え (S W 1 、 S W 2) 、前記第二の長さは、前記第一の長さ (N_L) より短く、前記第一の順変換の出力値又は前記第二の順変換の出力値の何れかを量子化及びエントロピ符号化 (Q U C O D) において処理し、前記切り替え、量子化、及び / 30
又はエントロピ符号化の制御 (P S Y M 、 F B C T L) は、前記入力信号の心理音響分析に由来し、対応する時間分解能制御情報 (S W I) を符号化出力信号 (C O S) に対して副次的情報として添付した符号化信号を復号化する方法であって、

前記符号化信号 (D I S) から前記副次的情報 (S W I) を提供 (D P C R Q U) するステップと、

前記符号化信号 (D I S) を逆量子化及びエントロピ復号化 (D P C R Q U) するステップと、

前記副次的情報に対応して、前記逆量子化及びエントロピ復号化信号の第一の長さ (N_L) の信号区間に作用し、復号化信号 (D O S) を提供する、時間領域への第一の順逆変換 (i M D C T - 1) を実行するか、或いは、前記第一の順逆変換 (i M D C T - 1) を実行する前に、第二の順逆変換 (i M D C T - 2) において、前記逆量子化及びエントロピ復号化信号の第二の長さ (N_{short}) の区間を処理するステップと、
を含む、前記復号化方法。

(付記 4)

入力信号の第一の長さ (N_L) の区間に適用される周波数領域への第一の順変換 (M D C T - 1) を使用して符号化された符号化信号 (D I S) 、例えば、音声信号であり、前記第一の順変換 (M D C T - 1) に続いて、前記変換済みの第一の長さの区間の第二の長さ (N_{short}) の区間に適用される第二の順変換 (M D C T - 2) を実行することで時間分解能を適応的に切り替え (S W 1 、 S W 2) 、前記第二の長さは、前記第一の長さ (N_L) より短く、前記第一の順変換の出力値又は前記第二の順変換の出力値の何れかを量子 50

化及びエントロピ符号化 (Q U C O D) において処理し、前記切り替え、量子化、及び / 又はエントロピ符号化の制御 (P S Y M 、 F B C T L) は、前記入力信号の心理音響分析に由来し、対応する時間分解能制御情報 (S W I) を符号化出力信号 (C O S) に対して副次的情報として添付した符号化信号を復号化する装置であって、

前記符号化信号 (D I S) から前記副次的情報 (S W I) を提供すること及び前記符号化信号を逆量子化及びエントロピ復号化することに適した手段 (D P C R Q U) と、

前記副次的情報に対応して、前記逆量子化及びエントロピ復号化信号の第一の長さ (N_L) の信号区間に作用し、復号化信号 (D O S) を提供する、時間領域への第一の順逆変換を実行するか、或いは、前記第一の順逆変換を実行する前に、第二の順逆変換において、前記逆量子化及びエントロピ復号化信号の第二の長さ (N_{short}) の区間を処理するのに適した手段 (i M D C T - 1 、 i M D C T - 2 、 S W 3 、 S W 4) と、
10 を含む、前記装置。

(付記 5)

前記第一及び第二の順変換は、 M D C T 又は整数 M D C T 又は D C T - 4 又は D C T 変換であり、前記第一及び第二の順変換は、それぞれ逆 M D C T 又は逆整数 M D C T 又は逆 D C T - 4 又は逆 D C T 変換である、付記 1 又は 3 記載の方法又は付記 2 又は 4 記載の装置。

(付記 6)

符号化側での前記変換の前、及び復号化側での前記変換後に、前記第一の長さ及び前記第二の長さの振幅値は、窓関数を使用して加重され、前記第一の長さ及び第二の長さの区間にに対する重複加算処理が加えられ、遷移窓に対して、振幅値は非対称窓関数を使用して加重され、前記第二の長さの区間にに対して、開始及び停止窓関数が使用される、付記 1 、
20 3 、又は 5 記載の方法又は付記 2 、 4 、又は 5 記載の装置。

(付記 7)

二つ以上の異なる第二の長さが使用される場合、利用する異なる第二の長さのトポロジを信号伝達するために、変化する時間分解能の領域を示す幾つかの指数、或いは復号化側でアクセス可能な対応するコードブックにおける一致するエントリを示す指標番号が、前記副次的情報に含まれる、付記 1 、 3 、 5 、又は 6 記載の方法又は付記 2 及び 4 乃至 6 の何れかに記載の装置。

(付記 8)

二つ以上の異なる第二の長さが連続して使用される場合、長さは、低周波数線を表す周波数ビンを開始点として増加する、付記 1 、 3 、及び 5 乃至 7 の何れかに記載の方法又は付記 2 及び 4 乃至 7 の何れかに記載の装置。

(付記 9)

前記トポロジは、

選択周波数帯域に対して、変換ビンのスペクトル電力を決定し、前記スペクトル電力値の算術平均値をその幾何平均値により割ることで、前記第一の順変換を使用してスペクトル平坦性基準 S F M を実行するステップと、

非加重入力信号区間をサブセグメント化し、m 個のサブセクションに対して加重及び短変換を実行し、こうした変換の周波数分解能は、選択周波数帯域に対応するステップと、
40

m 個の変換セグメントからなる各周波数線に対して、スペクトル電力を決定し、m 個のセグメントの幾何平均により割った算術平均を決定することで時間平坦性基準 T F M を計算するステップと、

S F M ベクトルを使用することで、音調又はノイズ周波数帯域を決定するステップと、こうした帯域内での時間的变化を認識するために、T F M 値を使用し、閾値を使用して、前記確認されたノイズ周波数帯域のために、更に細かな時間分解能への切り替えを行うステップと、により決定される、付記 7 又は 8 記載の方法又は装置。

(付記 10)

付記 1 及び 5 乃至 9 の何れかに記載の方法により符号化されたデジタル映像信号。

(付記 11)

付記 10 記載のデジタル映像信号を収容又は格納する、或いは付記 10 記載のデジタル映像信号が記録された、記憶媒体、例えば光ディスク。

(付記 12)

透かし埋め込み器における、付記 1 及び 5 乃至 9 のいずれかに記載の方法の使用。

【図面の簡単な説明】

【0062】

【図 1】本発明のエンコーダを示す図である。

【図 2】本発明のデコーダを示す図である。

【図 3】長 M D C T により窓処理及び変換した音声サンプルのブロックと、周波数データに適用した一連の不均一 M D C T とを示す図である。

【図 4】M D C T のブロック長を変化させることで時間 - 周波数分解能を変化させることを示す図である。

【図 5】遷移窓を示す図である。

【図 6】第二段階の M D C T に対する窓の連続の例を示す図である。

【図 7】第一及び最後の M D C T に対する開始及び停止窓を示す図である。

【図 8】遷移の時間領域信号と、第一の M D C T 段階の T / F プロットと、八倍の時間分解能トポロジを備えた第二段階の M D C T の T / F プロットとを示す図である。

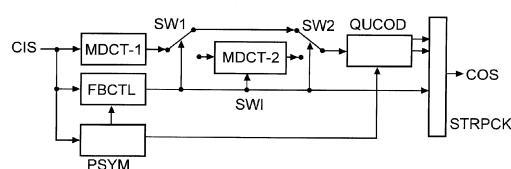
【図 9】遷移の時間領域信号と、単一、二倍、四倍、及び八倍の時間分解能トポロジの第二段階のフィルタ・バンク T / F プロットとを示す図である。

【図 10】図 6 による窓処理を更に詳細に示す図である。

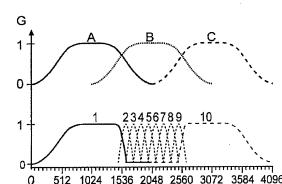
10

20

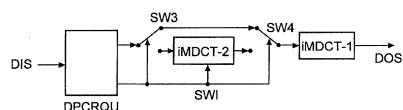
【図 1】



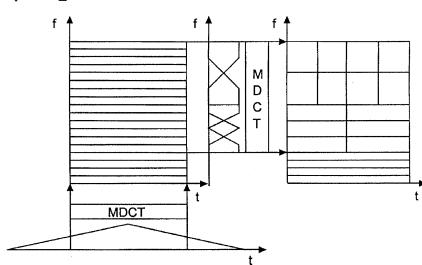
【図 4】



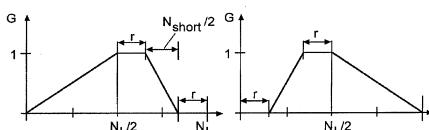
【図 2】



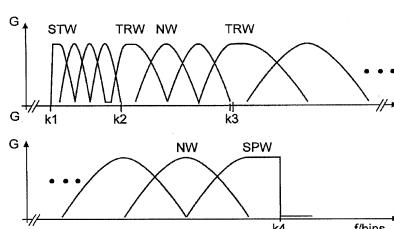
【図 3】



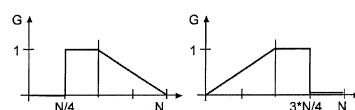
【図 5】



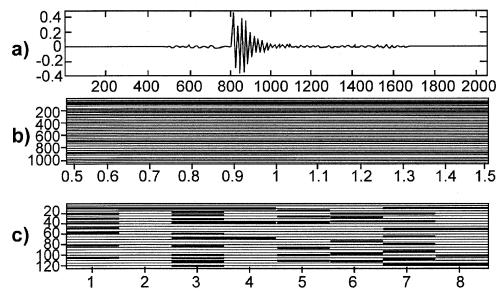
【図 6】



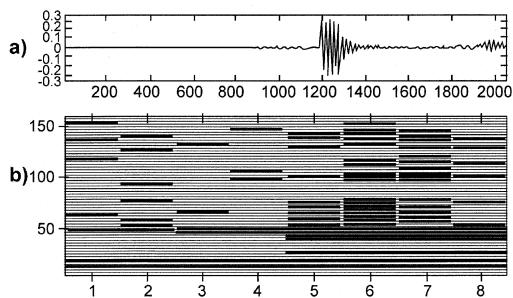
【図 7】



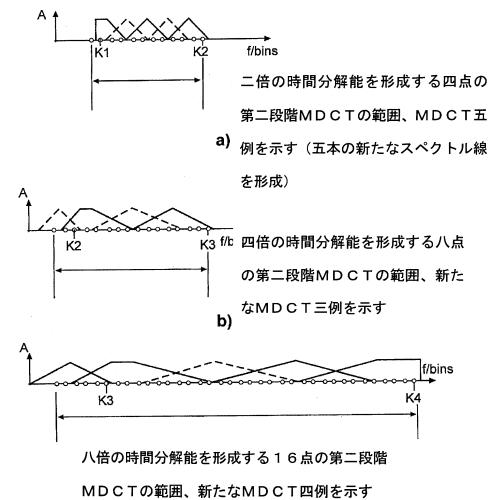
【図8】



【図9】



【図10】



フロントページの続き

(72)発明者 ヨハンネス ベーム
ドイツ国 37081 ゲツティング シーベルベーグ 35

(72)発明者 スベン コルドン
ドイツ国 30173 ハノーファー ヒルデスハイマー・シュトラッセ 117

審査官 上田 雄

(56)参考文献 国際公開第05/083682 (WO, A1)
特表2007-526691 (JP, A)
特開平10-039897 (JP, A)
特開平08-162964 (JP, A)
特開2001-265394 (JP, A)

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