

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-20346

(P2010-20346A)

(43) 公開日 平成22年1月28日(2010.1.28)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 O L 19/14 (2006.01)	G 1 O L 19/14 4 O O B	5 J O 6 4
H O 3 M 7/30 (2006.01)	H O 3 M 7/30 Z	

審査請求 有 請求項の数 20 O L (全 18 頁)

(21) 出願番号	特願2009-245860 (P2009-245860)	(71) 出願人	500046438
(22) 出願日	平成21年10月26日 (2009.10.26)		マイクロソフト コーポレーション
(62) 分割の表示	特願2002-185213 (P2002-185213) の分割		アメリカ合衆国 ワシントン州 98052-6399 レッドモンド ワン マイクロソフト ウェイ
原出願日	平成14年6月25日 (2002.6.25)	(74) 代理人	100077481
(31) 優先権主張番号	09/892, 105		弁理士 谷 義一
(32) 優先日	平成13年6月26日 (2001.6.26)	(74) 代理人	100088915
(33) 優先権主張国	米国 (US)		弁理士 阿部 和夫
		(72) 発明者	小石田 和人
			アメリカ合衆国 98052 ワシントン州 レッドモンド 146 アベニュー ノースイースト 7507

最終頁に続く

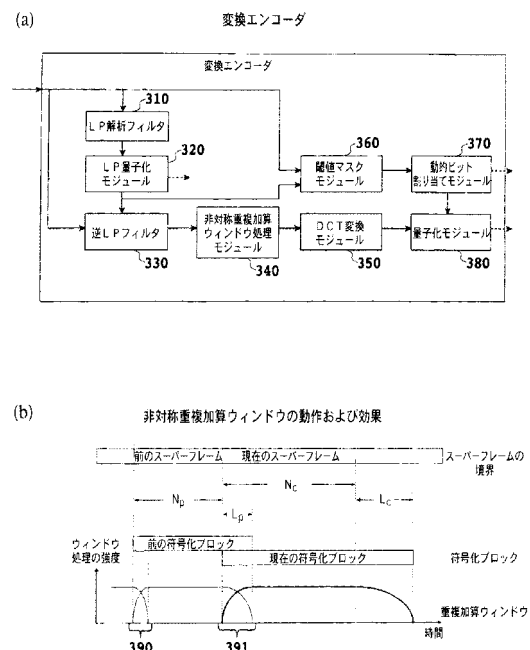
(54) 【発明の名称】 音声信号および音楽信号を符号化する方法

(57) 【要約】

【課題】特に低ビットレート環境で使用するよう適合した、音声および音楽信号の両方を符号化する単純かつ効率的なハイブリッド型の符号化アルゴリズムおよびアーキテクチャを提供すること。

【解決手段】本発明は、ハイブリッドコーデックで使用するのに適した、音楽信号に効率的な変換符号化方法を提供し、この方法では音声および音楽信号の両方に共通の線形予測 (LP) 合成フィルタを用いる。LP 合成フィルタは、音声または音楽信号の符号化に応じて、それぞれ音声励振ジェネレータと変換励振ジェネレータを切り替える。音声信号の符号化には、従来の CELP 技術を使用することができ、一方、音楽信号の符号化には、新規の非対称重複加算変換技術を応用する。共通の LP 合成フィルタリングを行う際に、重複加算操作領域の信号に LP 係数の補間を行う。復号器が音声符号化モードと音楽符号化モードを切り替えるときの滑らかな遷移が可能になる。

【選択図】 図 3



【特許請求の範囲】**【請求項 1】**

音声または音楽を有する信号の部分を符号化する方法であって、該方法は、

前記信号の現在の部分について、コードブック励振線形予測 (C E L P) 符号化モードと変換励振符号化モードとのいずれかを選択することであって、前記信号の現在の部分について前記変換励振符号化モードが選択されることと、

前記信号の前記現在の部分について線形予測解析を行い、線形予測パラメータを決定することと、

前記信号の前記現在の部分について線形予測フィルタリングを行い、前記現在の部分についての励振信号を生成することと、

出力としての符号化変換励振信号を生成する音楽符号化用の変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化することとを有し、前記変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化することは非対称重複加算変換法を適用することとを含み、前記非対称重複加算変換法は、

前の部分と前記現在の部分との間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかを判定し、

前記前の部分と前記現在の部分との間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかに基づいて、当該非対称重複加算変換法の前記現在の部分への適用法を調整することを含む、

ことを特徴とする方法。

【請求項 2】

前記変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化することは、

非対称重複加算変換法のために、非対称重複加算ウィンドウを適用して、ウィンドウ処理した信号を生成することと、

前記ウィンドウ処理した信号に周波数変換を実行して、周波数変換係数のセットを得ることと、

動的ビット割り当て情報を計算することと、

前記動的ビット割り当て情報に従って、前記周波数変換係数を量子化することとを有する、ことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記信号の前記現在の部分は変換符号化に高適合のサイズを有するスーパーフレームであることを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記線形予測フィルタリングの前に、前記線形予測パラメータを量子化したバージョンの補間を行うステップを有することを特徴とする請求項 2 に記載の方法。

【請求項 5】

前記非対称重複加算変換法の一部として、

前記現在の部分についての前記励振信号の非対称重複加算ウィンドウ処理の後に、ウィンドウ処理した信号は、前記現在の部分についての改変したサンプルおよび前記現在の部分の改変していないサンプルを有し、

重複加算処理は、前記現在の部分についての改変したサンプルと前の部分についての改変した重複するサンプルとを組み合わせる、ことを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記非対称重複加算変換法は、前記前の部分の重複の長さの値、前記現在の部分の長さおよび前記現在の部分の重複の長さに依存して変化するウィンドウ関数を使用することを特徴とする請求項 1 に記載の方法。

10

20

30

40

50

【請求項 7】

前記現在の部分のサンプルは、前記前の部分の重複の長さにある第 1 のサンプルおよび前記前の部分の重複の長さの後にある第 2 のサンプルを含み、前記ウィンドウ関数は、前記現在の部分の前記第 1 のサンプルを改変し、前記現在の部分の前記第 2 のサンプルを通過し、前記現在の部分の前記第 2 のサンプルの後にある重複するサンプルを改変する、ことを特徴とする請求項 6 に記載の方法。

【請求項 8】

前記前の部分の重複の長さは、前記現在の部分の前記重複の長さと異なることを特徴とする請求項 6 に記載の方法。

10

【請求項 9】

前記前の部分はコードブック励振線形予測の符号化の部分であって、前記前の部分の前記重複の長さの値はゼロであり、前記現在の部分の前記重複の長さの値はゼロではない、ことを特徴とする請求項 6 に記載の方法。

【請求項 10】

前記信号の次の部分について、前記コードブック励振線形予測 (CELP) 符号化モードと前記変換励振符号化モードとのいずれかを選択することであって、前記次の部分について前記コードブック励振線形予測符号化モードが選択されることと、

前記信号の前記次の部分について線形予測解析を行い、第 2 の線形予測パラメータを決定することと、

20

前記信号の前記次の部分について線形予測フィルタリングを行い、前記次の部分についての励振信号を生成することと、

出力としてのコードブック励振線形予測符号化励振信号を生成する音声符号化用のコードブック励振線形予測励振ジェネレータを用いて前記次の部分についての励振信号を符号化することと

を有することを特徴とする請求項 1 に記載の方法。

【請求項 11】

音声または音楽を有する信号の部分の符号化するステップをコンピュータに実行させるための命令を記憶したコンピュータ可読記憶媒体であって、前記命令は、

前記信号の現在の部分について、コードブック励振線形予測 (CELP) 符号化モードと変換励振符号化モードとのいずれかを選択するステップであって、前記信号の現在の部分について前記変換励振符号化モードが選択される、ステップと、

30

前記信号の前記現在の部分について線形予測解析を行い、線形予測パラメータを決定するステップと、

前記信号の前記現在の部分について線形予測フィルタリングを行い、前記現在の部分についての励振信号を生成するステップと、

出力としての符号化変換励振信号を生成する音楽符号化用の変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化するステップであって、前記変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化することは非対称重複加算変換法を適用することを含み、前記非対称重複加算変換法は、

40

前の部分と前記現在の部分との間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかを判定し、

前記前の部分と前記現在の部分との間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかに基づいて、当該非対称重複加算変換法の前記現在の部分への適用法を調整することを含む、ステップと

をコンピュータに実行させることを特徴とするコンピュータ可読記憶媒体。

【請求項 12】

前記変換励振ジェネレータを用いて前記現在の部分についての励振信号を符号化するス

50

テップは、

非対称重複加算変換法のために、非対称重複加算ウィンドウを適用して、ウィンドウ処理した信号を生成することと、

前記ウィンドウ処理した信号に周波数変換を実行して、周波数変換係数のセットを得ることと、

動的ビット割り当て情報を計算することと、

前記動的ビット割り当て情報に従って、前記周波数変換係数を量子化することとを有することを特徴とする請求項 1 1 に記載のコンピュータ可読記憶媒体。

【請求項 1 3】

前記非対称重複加算変換法の一部として、

前記現在の部分についての前記励振信号の非対称重複加算ウィンドウ処理の後に、ウィンドウ処理した信号は、前記現在の部分についての改変したサンプルおよび前記現在の部分の改変していないサンプルを有し、

重複加算処理は、前記現在の部分についての改変したサンプルと前の部分についての改変した重複するサンプルとを組み合わせる、ことを特徴とする請求項 1 1 に記載のコンピュータ可読記憶媒体。

【請求項 1 4】

前記非対称重複加算変換法は、前記前の部分の重複の長さの値、前記現在の部分の長さおよび前記現在の部分の重複の長さに依存して変化するウィンドウ関数を使用することを特徴とする請求項 1 1 に記載のコンピュータ可読記憶媒体。

【請求項 1 5】

前記前の部分はコードブック励振線形予測の符号化の部分であって、前記前の部分の前記重複の長さの値はゼロであり、前記現在の部分の前記重複の長さの値はゼロではない、ことを特徴とする請求項 1 4 に記載のコンピュータ可読記憶媒体。

【請求項 1 6】

前記命令は、

前記信号の次の部分について、前記コードブック励振線形予測 (C E L P) 符号化モードと前記変換励振符号化モードとのいずれかを選択するステップであって、前記次の部分について前記コードブック励振線形予測符号化モードが選択される、ステップと、

前記信号の前記次の部分について線形予測解析を行い、第 2 の線形予測パラメータを決定するステップと、

前記信号の前記次の部分について線形予測フィルタリングを行い、前記次の部分についての励振信号を生成することと、

出力としてのコードブック励振線形予測符号化励振信号を生成する音声符号化用のコードブック励振線形予測励振ジェネレータを用いて前記次の部分についての励振信号を符号化するステップとを

前記コンピュータに実行させることを特徴とする請求項 1 1 に記載のコンピュータ可読記憶媒体。

【請求項 1 7】

スーパーフレームを符号化する音声 / 音楽符号化装置であって、前記スーパーフレームは音声または音楽を含み、前記装置は、

コードブック励振線形予測 (C E L P) 符号化スーパーフレームである、または変換励振符号化スーパーフレームであるとして現在のスーパーフレームを分類するクラシファイアと、

前記現在のスーパーフレームを解析し、線形予測パラメータのセットを生成する 1 つまたは複数の線形予測解析モジュールと、

前記現在のスーパーフレームの励振信号を生成する 1 つまたは複数の線形予測フィルタリングモジュールと、

前記現在のスーパーフレームがコードブック励振線形予測符号化のスーパーフレームである場合に、前記励振信号を符号化する 1 つまたは複数の音声符号化用の符号化励振 (C

10

20

30

40

50

E L P) 符号化モジュールと、

前記現在のスーパーフレームが変換符号化のスーパーフレームである場合に、前記励振信号を符号化する１つまたは複数の音楽符号化用の変換励振符号化モジュールであって、

前記１つまたは複数の変換励振符号化モジュールを用いて前記励振信号を符号化することは、非対称重複加算変換法を適用することを含み、当該非対称重複加算変換法は、

前のスーパーフレームと前記現在のスーパーフレームとの間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかを判定し、

前記前のスーパーフレームと前記現在のスーパーフレームとの間の移り変わりが、コードブック励振線形予測符号化から変換励振符号化への移り変わりであるか、変換励振符号化から変換励振符号化への移り変わりであるかに基づいて、当該非対称重複加算変換法の前記現在のスーパーフレームへの適用法を調整することを含む、

前記１つまたは複数の変換励振符号化モジュールと
を備えたことを特徴とする装置。

【請求項１８】

前記クラシファイアは、前記現在のスーパーフレームがコードブック励振線形予測符号化のスーパーフレームであるかまたは変換励振符号化のスーパーフレームであるかを表すモードビットを提供することを特徴とする請求項１７に記載の装置。

【請求項１９】

前記１つまたは複数の変換励振符号化モジュールは、

前記励振信号をウィンドウ処理し、ウィンドウ処理した信号を提供する非対称重複加算ウィンドウ処理モジュールと、

前記ウィンドウ処理した信号を周波数変換係数のセットに変換する周波数変換モジュールと、

ビット割り当て情報を提供する動的ビット割り当てモジュールと、

前記ビット割り当て情報に従って、前記周波数変換係数を量子化する周波数変換係数量子化モジュールと

を備えたことを特徴とする請求項１７に記載の装置。

【請求項２０】

前記非対称重複加算変換法は、前記前の部分の重複の長さ、前記現在の部分の長さおよび前記現在の部分の重複の長さの値に依存して変化するウィンドウ関数を使用し、

前記前の部分がコードブック励振線形予測の符号化音声の部分である場合は、前記前の部分の前記重複の長さの値はゼロであり、前記現在の部分の前記重複の長さの値はゼロではないことを特徴とする請求項１７に記載の装置。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、一般には信号を符号化する方法および装置を対象とし、より詳細には音声信号と音楽信号の両方を符号化する方法および装置を対象とする。

【背景技術】

【０００２】

本質的に音声と音楽は大きく異なる信号によって表される。典型的なスペクトルの特徴から見ると、声に出した音声 (s p e e c h) のスペクトルは、一般にピッチの倍音と関連する細かい周期的な構造を持ち、倍音のピークが滑らかなスペクトル包絡線を描くのに対して、音楽のスペクトルは通例はるかに複雑で、複数のピッチの基本波と倍音を示す。スペクトル包絡線もより複雑であると考えられる。この２つの信号モードの符号化技術も非常に異なっており、音声の符号化には、符号励振線形予測 (C E L P) や正弦波符号化などモデルに基づく手法を主に使用し、音楽の符号化には、知覚的なノイズマスキングと合わせて使用する変形重複変換 (M o d i f i e d L a p p e d T r a n s f o r m a t i o n) (M L T) などの変換符号化技術を主に使用する。

【 0 0 0 3 】

近年、インターネットマルチメディア、ＴＶ／ラジオ放送、テレビ会議、あるいは無線媒体といったアプリケーションのために、音声信号と音楽信号の両方を符号化することが増えている。しかし、この２種の信号タイプ向けの符号器（coder）は、異なる技術に最適な形で基づくものなので、音声信号と音楽信号の両方を効率的かつ効果的に再生する汎用コーデックの生産は容易に達成することができない。例えば、ＣＥＬＰのような線形予測ベースの技術は、音声信号については高品質の再生を発揮することができるが、音楽信号の再生の品質は受け入れがたいものである。一方、変換符号化に基づく技術は、音楽信号には良質の再生を提供するが、特に低ビットレートの符号化の場合に、音声信号についての出力が著しく劣化する。

10

【 0 0 0 4 】

可能な方法の１つは、音声信号および音楽信号どちらにも対応することのできるマルチモードの符号器を設計することである。そのような符号器を提供しようとした以前の試みには、例えば、ハイブリッドＡＣＥＬＰ／変換符号化励振符号器、およびマルチモード変換予測符号器（ＭＴＰＣ）がある。残念なことに、これらの符号化アルゴリズムは、音声信号および音楽信号を実用的に符号化するには、あまりにも複雑かつ／または非効率的なものである。

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 5 】

20

特に低ビットレート環境で使用するように適合した、音声信号および音楽信号の両方を符号化する、単純かつ効率的なハイブリッド型の符号化アルゴリズムおよびアーキテクチャを提供することが望まれる。

【 課題を解決するための手段 】

【 0 0 0 6 】

本発明は、音楽信号を効率的に符号化する変換符号化法を提供する。この変換符号化法はハイブリッドコーデックで使用するのに適しており、音声信号および音楽信号両方の再生に、共通の線形予測（ＬＰ）合成フィルタを用いる。ＬＰ合成フィルタの入力は、音声信号または音楽信号の符号化に従って、それぞれ音声励振ジェネレータと変換励振ジェネレータに切り替える。好ましい実施形態では、ＬＰ合成フィルタは、ＬＰ係数の補間を含む。音声信号の符号化には、従来のＣＥＬＰまたはその他のＬＰ技術を使用することができ、一方、音楽信号の符号化には、非対称重複加算変換技術を応用することが好ましい。本発明の潜在的な利点は、コーデックが音声符号化と音楽符号化を切り替える箇所で滑らかな出力推移を可能にすることである。

30

【 0 0 0 7 】

本発明のこの他の特徴および利点は、添付の図面を参照しながら進める以下の例示の実施形態の詳細な説明から明らかになる。

【 0 0 0 8 】

特許請求の範囲に本発明の特徴を詳細に示すが、本発明とその目的および利点は、以下の詳細な説明を添付の図面と合わせて読むことにより、最も明瞭に理解することができよう。

40

【 図面の簡単な説明 】

【 0 0 0 9 】

【 図 1 】 本発明の一実施形態によるネットワークでリンクした例示的なハイブリッド型音声／音楽コーデックの図である。

【 図 2 】 本発明の一実施形態によるハイブリッド型音声／音楽符号変換器の簡略化したアーキテクチャ図である。

【 図 3 】 本発明の一実施形態による変換符号化アルゴリズムの論理図、および、本発明の一実施形態による非対称型の重複加算ウィンドウ操作とその効果を表すタイミング図である。

50

【図４】本発明の一実施形態による変換符号化アルゴリズムのブロック図である。

【図５】本発明の一実施形態により、音声信号および音楽信号の符号化に使用する例示的ステップを表す流れ図である。

【図６】本発明の一実施形態により、音声信号および音楽信号の符号化に使用する例示的ステップを表す流れ図である。

【図７】本発明の一実施形態により、音声信号および音楽信号の復号に使用する例示的ステップを表す流れ図である。

【図８】本発明の一実施形態により、音声信号および音楽信号の復号に使用する例示的ステップを表す流れ図である。

【図９】本発明の一実施形態を実行することが可能な、コンピューティングデバイスによって用いられるコンピューティングデバイスのアーキテクチャを表す簡略図である。

10

【発明を実施するための形態】

【００１０】

本発明は、音楽信号を符号化する効率的な変換符号化法を提供し、この方法はハイブリッドコーデックで使用するのに適しており、音声信号および音楽信号の両方を再生するのに共通の線形予測（ＬＰ）合成フィルタを利用する。概説すると、符号化音声信号を受信したか、あるいは符号化音楽信号を受信したかに応じて、ＬＰ合成フィルタの入力を、それぞれ音声励振ジェネレータと変換励振ジェネレータとの間で動的に切り替える。音声／音楽クラシファイアは、入力された音声／音楽信号が音声であるか音楽であるかを識別し、識別した信号を適切に音声符号変換器（speech encoder）または音楽符号変換器（music encoder）に転送する。音声信号を符号化する際には、従来のＣＥＬＰ技術を使用することができる。しかし、音楽信号の符号化には、新規の非対称重複加算変換技術を応用する。本発明の好ましい実施形態では、共通ＬＰフィルタはＬＰ係数の補間を含み、重複を介して励振が得られる領域の数個のサンプルごとに補間を行う。合成フィルタの出力は切り替えず、合成フィルタの入力だけを切り替えるので、可聴信号の不連続性の原因が回避される。

20

【００１１】

図１を参照して、本発明の一実施形態を実施することが可能な例示的な音声／音楽コーデックの構成を説明する。図示された環境は、雲形で表すネットワーク１００を介して相互に通信するコーデック１１０、１２０を含む。ネットワーク１００は、ルータ、ゲートウェイ、ハブなど多数の周知の構成要素を含むことができ、有線媒体および無線媒体のどちらか、または両方を通じて通信を提供することができる。各コーデックは、少なくとも、符号変換器１１１、１２１、復号器１１２、１２２、および音声／音楽クラシファイア１１３、１２３を含む。

30

【００１２】

本発明の一実施形態では、共通の線形予測合成フィルタを音楽信号および音声信号の両方に使用する。図２を参照すると、本発明を実施することが可能な例示的な音声および音楽コーデックの構造を示している。詳細には、図２は、ハイブリッド音声／音楽符号変換器の高レベル構造を示し、図２は、ハイブリッド音声／音楽復号器の高レベル構造を示す。図２を参照すると、音声／音楽符号変換器は、入力信号を音声信号または音楽信号に分類する音声／音楽クラシファイア２５０を含む。識別された信号は、識別結果に応じてそれぞれ音声符号変換器２６０または音楽符号変換器２７０に送信され、入力信号の音声／音楽特性を特徴化するモードビットが生成される。例えば、ゼロのモードビットは音声信号を表し、１のモードビットは音楽信号を表す。音声符号変換器２６０は、当業者に周知の線形予測の原理に基づいて入力信号を符号化し、符号化した音声ビットストリームを出力する。使用する音声符号化は、例えば、当業者に知られるコードブック励振線形予測（ＣＥＬＰ）技術である。これに対して、音楽符号変換器２７０は、下記で説明する変換符号化法に従って入力音楽信号を符号化し、符号化した音楽ビットストリームを出力する。

40

【００１３】

図２を参照すると、本発明の一実施形態による音声／音楽復号器は、線形予測（ＬＰ）

50

合成フィルタ 240 と、音声励振ジェネレータ 210 と変換励振ジェネレータ 220 を切り替える、フィルタ 240 の入力部に接続された音声 / 音楽スイッチ 230 とを含む。音声励振ジェネレータ 210 は、送信されてきた符号化音声 / 音楽ビットストリームを受信し、音声励振信号を生成する。音楽励振ジェネレータ 220 は、送信されてきた符号化音声 / 音楽信号を受信し、音楽励振信号を生成する。符号器には 2 つのモード、すなわち音声モードと音楽モードがある。現在のフレームまたはスーパーフレームに対する復号器のモードは、送信されるモードビットによって決まる。音声 / 音楽スイッチ 230 は、モードビットに従って励振信号ソースを選択し、したがって音楽モードでは音楽励振信号を選択し、音声モードでは音声励振信号を選択する。次いでスイッチ 230 は、適切な再構築信号を生成するために、選択された励振信号を線形予測合成フィルタ 240 に転送する。音声モードにおける励振または残差は、コード励振線形予測 (CELP) 符号化などの音声最適化技術を使用して符号化し、一方、音楽モードにおける励振は、例えば変換符号化励振 (TCX) などの変換符号化技術によって量子化する。復号器の LP 合成フィルタ 240 は、音楽信号と音声信号の両方に共通である。

【0014】

音声信号または音楽信号を符号化する従来の符号器は、通例フレームと称される 10 ms ~ 40 ms のブロックまたは区分に対して作用する。一般に、変換符号化はフレームサイズが大きい方が効率的なので、一般にこのような 10 ms ~ 40 ms のフレームは、特にビットレートが低い場合には、変換符号器 (transform coder) を整合して許容できる品質を得るには短すぎる。このため、本発明の一実施形態は、整数個の標準的な 20 ms のフレームで構成されるスーパーフレームに対して作用する。一実施形態で使用する標準的なスーパーフレームのサイズは 60 ms である。この結果、音声 / 音楽クラシファイアは、連続したスーパーフレーム 1 つにつき 1 回の分類を行うことが好ましい。

【0015】

音楽信号を符号化する現在の変換符号器と異なり、本発明による符号化プロセスは励振領域で行われる。これは、音声と音楽両タイプの信号の再生に、単一の LP 合成フィルタを使用することの結果である。図 3 (a) を参照すると、本発明の一実施形態による変換符号変換器を示している。線形予測 (LP) 解析フィルタ 310 は、音声 / 音楽クラシファイア 250 から出力される、分類済みの音楽スーパーフレームの音楽信号を解析して、適切な線形予測係数 (LPC) を得る。LP 量子化モジュール 320 は、計算された LPC 係数を量子化する。次いで LPC 係数およびスーパーフレームの音楽信号を入力として音楽信号を得、出力として残差信号を生成する逆フィルタ 330 にかける。

【0016】

一般的なフレームではなくスーパーフレームを使用することは、高品質の変換符号化を得る助けとなる。しかし、スーパーフレームの境界におけるブロッキングひずみによって品質問題が生じる可能性がある。ブロッキングひずみの影響を軽減する好ましい解決法は、例えば、隣接フレームとの重複が 50 % の変形重複変換 (MLT) などの、重複加算ウィンドウ技術に見出される。しかし、CELP では音声符号化にゼロの重複を利用するので、このような解決法を CELP ベースのハイブリッドコーデックに組み込むことは難しいと思われる。この難題を克服し、音楽モードにおけるシステムの高品質の動作を保証するために、本発明の一実施形態は、図 3 (a) の重複加算モジュール 340 によって実施される非対称重複加算ウィンドウ法を提供する。図 3 (b) は、非対称重複加算ウィンドウの動作および効果を表す。図 3 (b) を参照すると、重複加算ウィンドウは、1 つ前のスーパーフレームが、例えばそれぞれ N_p および L_p で表すスーパーフレームの長さおよび重複の長さに異なる値を有し得るという可能性を考慮したものになっている。符号 (designator) N_c および L_c はそれぞれ、現在のスーパーフレームのスーパーフレーム長と重複の長さを表す。現在のスーパーフレームの符号化ブロックは、現在のスーパーフレームのサンプルと重複のサンプルを含む。重複加算のウィンドウ処理は、現在の符号化ブロックの最初の N_p サンプルおよび最後の L_p サンプルで行われる。これに限定しないが、例えば次のように、入力信号 $x(n)$ を重複加算ウィンドウ関数 $w(n)$ で変換して

、ウィンドウ処理した信号 $y(n)$ を得る。

$$y(n) = x(n)w(n), \quad 0 \leq n \leq N_c + L_c - 1 \cdots \cdots \text{(数式 1)}$$

ウィンドウ関数 $w(n)$ は次のように定義される。

【 0 0 1 7 】

【 数 1 】

$$w(n) = \begin{cases} \sin\left(\frac{\pi}{2L_p}(n+0.5)\right), & 0 \leq n \leq L_p - 1 \\ 1, & L_p \leq n \leq N_c - 1 \\ 1 - \sin\left(\frac{\pi}{2L_c}(n - N_c + 0.5)\right), & N_c \leq n \leq N_c + L_c - 1 \end{cases} \cdots \cdots \text{(数式 2)}$$

【 0 0 1 8 】

この場合、 N_c および L_c は、それぞれ、現在のスーパーフレームのスーパーフレーム長と重複の長さである。

【 0 0 1 9 】

図 3 (b) の重複加算ウィンドウの形状から、例えば、重複加算レンジ 3 9 0、3 9 1 が非対称形であり、符号 3 9 0 の領域が符号 3 9 1 の領域と異なり、また重複加算のウィンドウは相互にサイズが異なることが見て取れる。このようなサイズが可変のウィンドウにより、ブロッキングの影響とプリエコーを克服する。また、MLT 技術で利用する 5 0 % の重複と比較すると重複領域が小さいので、この非対称重複加算ウィンドウの方法は、下記で説明するように、CELP ベースの音声符号器 (speech coder) に組み込むことのできる変換符号器に効率的である。

【 0 0 2 0 】

再度図 3 (a) を参照すると、逆 LP フィルタ 3 3 0 から出力される残差信号は、非対称形の重複加算ウィンドウ処理モジュール 3 4 0 によって処理し、ウィンドウ処理した信号を生成する。ウィンドウ処理した信号は次いで離散コサイン変換 (DCT) モジュール 3 5 0 に入力され、ここでウィンドウ処理した信号を周波数領域に変換し、DCT 係数のセットを得る。DCT 変換は次のように定義され、

【 0 0 2 1 】

【 数 2 】

$$Z(k) = \sqrt{\frac{2}{K}} \sum_{i=0}^{K-1} c(i)Z(i) \cos\left(\frac{(i+0.5)k\pi}{K}\right), \quad 0 \leq k \leq K-1 \cdots \cdots \text{(数式 3)}$$

【 0 0 2 2 】

$c(k)$ は次のように定義される。ただし、 K は変換サイズである。

【 0 0 2 3 】

【 数 3 】

$$c(k) = \begin{cases} 1/\sqrt{2}, & k=0 \\ 1, & k=0 \text{ 以外の場合} \end{cases}$$

【 0 0 2 4 】

DCT 変換が好ましいが、変形離散コサイン変換 (MDCT) および高速フーリエ変換 (FFT) を含む技術など、他の変換技術も応用することができる。DCT 係数を効率的に量子化するために、DCT 係数量子化の一部として動的ビット割り当て情報を利用する。動的ビット割り当て情報は、閾値マスキングモジュール 3 6 0 で計算するマスキング閾値に従って、動的ビット割り当てモジュール 3 7 0 から得るが、この閾値マスキングは、入力される信号か、または LPC 解析モジュール 3 1 0 から出力される LPC 係数に基づ

く。動的ビット割り当て情報は、入力音楽信号の解析から得ることもできる。動的ビット割り当て情報を用いて、量子化モジュール 380 で DCT 係数を量子化し、次いで復号器に送出する。

【0025】

本発明の上記の実施形態で用いる符号化アルゴリズムに沿い、変換復号器を図 4 に示す。図 4 を参照すると、変換復号器は、逆動的ビット割り当てモジュール (Inverse Dynamic bit allocation module) 410、逆量子化モジュール 420、DCT 逆変換モジュール 430、非対称重複加算ウィンドウモジュール 440、および重複加算モジュール 450 を含む。逆動的ビット割り当てモジュール 410 は、図 3 (a) の動的ビット割り当てモジュール 370 から出力され、送信されるビット割り当て情報を受け取り、ビット割り当て情報を逆量子化モジュール 420 に提供する。逆量子化モジュール 420 は、送信されてきた音楽ビットストリームとビット割り当て情報を受け取り、ビットストリームに逆量子化を適用して、符号化した DCT 係数を得る。次いで DCT 逆変換モジュール 430 は、符号化した DCT 係数の逆 DCT 変換を実行し、時間領域の信号を生成する。逆 DCT 変換は次のように示すことができ、

【0026】

【数 4】

$$Z(i) = \sqrt{\frac{2}{K}} \sum_{k=0}^{K-1} c(k) Z(k) \cos\left(\frac{(i+0.5)k\pi}{K}\right), 0 \leq i \leq K-1 \quad \dots (数式 4)$$

【0027】

$c(k)$ は次のように定義される。ただし、 K は変換サイズである。

【0028】

【数 5】

$$c(k) = \begin{cases} 1/\sqrt{2}, & k=0 \\ 1, & k=0 \text{ 以外の場合} \end{cases}$$

【0029】

重複加算ウィンドウ処理モジュール 440 は、時間領域の信号に対し、例えば、

【0030】

【数 6】

$$\hat{y}'(n) = w(n) \hat{y}(n)$$

【0031】

など、非対称の重複加算ウィンドウ処理操作を行う。ここで

【0032】

$$\hat{y}(n)$$

【0033】

は時間領域の信号を表す。 $w(n)$ はウィンドウ関数を表す。

【0034】

$$\hat{y}'(n)$$

【0035】

はこの結果得られるウィンドウ処理後の信号である。ウィンドウ処理を行った信号は次いで重複加算モジュール 450 に送られ、ここで重複加算操作を行うことにより励振信号が得られる。これに限定しないが、例として、例示的な重複加算操作は次のようなものである。

【 0 0 3 6 】

【 数 7 】

$$\hat{e}(n) = \begin{cases} w_p(n + N_p) \hat{y}_p(n + N_p) + w_c(n) \hat{y}_c(n), & 0 \leq n \leq L_p - 1 \\ \hat{y}_c(n), & L_p \leq n \leq N_c - 1 \end{cases} \quad \dots\dots\dots \text{(数式 5)}$$

【 0 0 3 7 】

ここで、

【 0 0 3 8 】

 $\hat{e}(n)$

10

【 0 0 3 9 】

は励振信号であり、

【 0 0 4 0 】

 $\hat{y}_p(n)$

【 0 0 4 1 】

および

【 0 0 4 2 】

 $\hat{y}_c(n)$

20

【 0 0 4 3 】

はそれぞれ、

【 0 0 4 4 】

1 つ前と現在の時間領域信号である。関数 $w_p(n)$ および $w_c(n)$ はそれぞれ、以前のスーパーフレームと現在のスーパーフレームについての重複加算ウィンドウ関数である。値 N_p および N_c は、それぞれ 1 つ前のスーパーフレームと現在のスーパーフレームのサイズである。値 L_p は、1 つ前のスーパーフレームの重複加算のサイズである。

【 0 0 4 5 】

生成された励振信号

30

【 0 0 4 6 】

 $\hat{e}(n)$

【 0 0 4 7 】

は次いで、

【 0 0 4 8 】

図 2 に示すように、切り替え可能な形で LP 合成フィルタに送られ、元の音楽信号を再構築する。

【 0 0 4 9 】

励振信号の処理には、補間合成技術を応用することが好ましい。LP 係数は、 $0 \leq n \leq L_p - 1$ の領域で数個のサンプルごとに補間し、重複加算操作を用いて励振を得る。LP 係数の補間は、線スペクトル対 (LSP) 領域で行われ、補間する LSP 係数の値は次の式によって得られる。

【 0 0 5 0 】

【 数 8 】

$$f(i) = (1 - v(i)) \hat{f}_p(i) + v(i) \hat{f}_c(i), \quad 0 \leq i \leq M - 1 \quad \dots\dots\dots \text{(数式 6)}$$

【 0 0 5 1 】

 $\hat{f}_p(i)$

50

【 0 0 5 2 】

および

【 0 0 5 3 】

$\hat{f}_c(i)$

【 0 0 5 4 】

は、

【 0 0 5 5 】

それぞれ、1つ前のスーパーフレームおよび現在のスーパーフレームの量子化 L S P パラメータである。係数 $v(i)$ は補間重み係数であり、値 M は L P 係数の次数である。補間技術を用いた後に、従来の L P 合成技術を励振信号に適用して、再構築された信号を得る。

10

【 0 0 5 6 】

図 5 および図 6 を参照して、本発明の一実施形態により、インタリーブした入力音声信号および音楽信号を符号化する際に従う例示的ステップを説明する。ステップ 5 0 1 で、入力信号を受け取り、スーパーフレームを形成する。ステップ 5 0 3 で、現在のスーパーフレームのタイプ（すなわち音楽 / 音声）がそれまでのスーパーフレームのタイプと異なるかどうかを判定する。スーパーフレームが異なる場合は、現在のスーパーフレームの開始部で「スーパーフレーム遷移」を定義し、動作の流れは分岐してステップ 5 0 5 に進む。ステップ 5 0 5 で、例えば現在のスーパーフレームが音楽であるかどうかを判定することにより、1つ前のスーパーフレームのシーケンスと現在のスーパーフレームを判定する。したがって、例えば、1つ前のスーパーフレームが音声スーパーフレームであり、その後に現在の音楽スーパーフレームが続く場合は、ステップ 5 0 5 の実行の結果は「yes」になる。同様に、1つ前のスーパーフレームが音楽スーパーフレームであり、その後に現在の音声スーパーフレームが続く場合、ステップ 5 0 5 の結果は「no」になる。ステップ 5 0 5 から「yes」の結果に分岐したステップ 5 1 1 で、1つ前の音声スーパーフレームの重複の長さ L_p をゼロにセットし、現在の符号化ブロックの開始部では重複加算ウィンドウを実行しないことを表す。この理由は、C E L P ベースの音声符号器が、隣接するフレームまたはスーパーフレームの重複信号を提供または利用しないためである。ステップ 5 1 1 に続き、ステップ 5 1 3 で音楽スーパーフレームに変換符号化手順を実行する。ステップ 5 0 5 の判定の結果が「no」である場合、動作の流れは分岐してステップ 5 0 9 に進み、ここで1つ前の音楽スーパーフレームの重複サンプルを破棄する。続いて、ステップ 5 1 5 で音声スーパーフレームに C E L P 符号化を実行する。ステップ 5 0 3 から「no」の結果に分岐したステップ 5 0 7 では、現在のスーパーフレームが音楽スーパーフレームか、音声スーパーフレームかを判定する。現在のスーパーフレームが音楽スーパーフレームである場合は、ステップ 5 1 3 で変換符号化を適用し、現在のスーパーフレームが音声である場合は、ステップ 5 1 5 で C E L P 符号化の手順を適用する。ステップ 5 1 3 で変換符号化が完了すると、符号化した音楽ビットストリームが生成される。同様に、ステップ 5 1 5 で C E L P 符号化を実行すると、符号化した音声ビットストリームが生成される。

20

30

40

【 0 0 5 7 】

ステップ 5 1 3 で行われる変換符号化は、図 6 に示す一連のサブステップを含む。ステップ 5 2 3 で、入力信号の L P 係数を計算する。ステップ 5 3 3 で、計算された L P C 係数を量子化する。ステップ 5 4 3 で、受け取ったスーパーフレームおよび計算した L P C 係数に逆フィルタをかけて残差信号 $x(n)$ を生成する。ステップ 5 5 3 で、次のように $x(n)$ にウィンドウ関数 $w(n)$ を乗算することにより、重複加算ウィンドウを残差信号 $x(n)$ に適用する。

$$y(n) = x(n) w(n)$$

この場合、ウィンドウ関数 $w(n)$ は数式 2 と同様に定義される。ステップ 5 6 3 で、ウィンドウ処理した信号 $y(n)$ に D C T 変換を行い、D C T 係数を得る。ステップ 5 8 3

50

で、ステップ 573 で得るマスキング閾値に従って、動的ビット割り当て情報を得る。次いでステップ 593 で、ビット割り当て情報を使用し、DCT 係数を量子化して音楽ビットストリームを生成する。

【0058】

図 5 および図 6 に示す符号化ステップに沿い、図 7 および図 8 は、本発明の一実施形態で合成した信号を提供する際に復号のため採られるステップを示している。図 7 を参照すると、ステップ 601 で、送信されるビットストリームおよびモードビットを受信する。ステップ 603 で、モードビットにより、現在のスーパーフレームが音楽に対応するか、音声に対応するかを判断する。その信号が音楽に対応する場合は、ステップ 607 で変換励振を生成する。ビットストリームが音声に対応する場合は、ステップ 605 を実行して、CEL P 解析の場合と同様に音声励振信号を生成する。ステップ 607 と 605 はどちらもステップ 609 に合流する。ステップ 609 で、LP 合成フィルタが音楽励振信号または音声励振信号を適切に受け取るようにスイッチをセットする。例えば $0 \leq n \leq L_p - 1$ などの領域でスーパーフレームを重複加算するときには、スーパーフレームのこの重複加算領域中の信号の LPC 係数を補間することが好ましい。ステップ 611 で、LPC 係数の補間を実行する。LPC 係数の補間を行うためには、例えば数式 6 を用いることができる。続いてステップ 613 で、当業者にはよく理解される方式で、LPC 合成フィルタを介して元の信号を再構築、すなわち合成する。

【0059】

本発明によると、音声励振ジェネレータは、音声合成に適した任意の励振ジェネレータでよいが、変換励振ジェネレータは、図 8 に示すような特別に適合した方法であることが好ましい。図 8 を参照すると、送信されるビットストリームをステップ 617 で受信した後に、ステップ 627 で逆ビット割り当てを実行してビット割り当て情報を得る。ステップ 637 で、DCT 係数の逆 DCT 量子化を行うことにより、DCT 係数を得る。ステップ 647 で、数式 4 で定義する逆 DCT 変換を DCT 係数に行うことにより、予備的な時間領域の励振信号を再構築する。ステップ 657 で、数式 2 で定義される重複加算ウィンドウを適用することにより、再構築された励振信号をさらに処理する。ステップ 667 で、重複加算操作を行って、数式 5 で定義する音楽励振信号を得る。

【0060】

これは必須ではないが、本発明は、コンピュータで実行されるプログラムモジュールなどの命令を使用して実施することができる。一般に、プログラムモジュールには、特定のタスクを実行するか、または特定の抽象データタイプを実施するルーチン、オブジェクト、コンポーネント、データ構造などが含まれる。ここで使用する用語「プログラム」は、1 つ以上のプログラムモジュールを含む。

【0061】

本発明は、各種タイプのマシンで実施することができるが、これには、携帯電話、パーソナルコンピュータ (PC)、ハンドヘルドデバイス、マルチプロセッサシステム、マイクロプロセッサベースのプログラマブル消費者家電製品、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータなど、あるいは本明細書に述べるようにオーディオ信号を符号化または復号し、また信号の記憶、取り出し、送信、または受信に使用することのできる任意の他のマシンが含まれる。本発明は、通信ネットワークを通じてリンクした遠隔コンポーネントによってタスクを実行する分散型コンピューティングシステムで実施することができる。

【0062】

図 9 を参照すると、本発明の実施形態を実施する例示的なシステムは、コンピューティングデバイス 700 などのコンピューティングデバイスを含む。その最も基本的な構成では、コンピューティングデバイス 700 は、通例少なくとも 1 つの処理装置 702 とメモリ 704 を含む。メモリ 704 は、コンピューティングデバイスの厳密な構成およびタイプに応じて、揮発性 (RAM など)、不揮発性 (ROM、フラッシュメモリなど)、あるいはこの 2 つの組み合わせにすることができる。この最も基本的な構成を、図 9 の線 7

06の中に示している。これに加えて、デバイス700は、追加の装備/機能も有することができる。例えば、デバイス700は、これらに限定しないが磁気ディスクまたは光ディスク、またはテープを含む、追加のストレージ（取り外し可能/取り外し不能）も含むことができる。このような追加ストレージを、取り外し可能ストレージ708および取り外し不能ストレージ710として図9に示している。コンピュータ記憶媒体は、コンピュータ可読命令、データ構造、プログラムモジュール、あるいはその他のデータなどの情報を記憶するための任意の方法または技術に実施された揮発性および不揮発性、取り外し可能および取り外し不能の媒体を含む。メモリ704、取り外し可能ストレージ708、および取り外し不能ストレージ710はすべて、コンピュータ記憶媒体の例である。これらに限定しないが、コンピュータ記憶媒体には、RAM、ROM、EEPROM、フラッシュ

10

【0063】

デバイス700は、デバイスが他のデバイスと通信することを可能にする1つ以上の通信接続712も含むことができる。通信接続712は、通信媒体の一例である。通信媒体は通例、コンピュータ可読命令、データ構造、プログラムモジュール、あるいはその他のデータを搬送波やその他の搬送機構などの変調データ信号に実施し、また任意の情報伝達媒体を含む。用語「変調データ信号」とは、情報を信号中に符号化するような方式で、その特徴の1つ以上を設定または変更した信号を意味する。例として、通信媒体には、有線ネットワークまたは直接配線接続などの有線媒体、および音響、RF、赤外線およびその他の無線媒体などの無線媒体が含まれるが、これらに限定しない。上記で述べたように、本明細書で使用する用語「コンピュータ可読媒体」は、記憶媒体および通信媒体の両方を含む。

20

【0064】

デバイス700は、キーボード、マウス、ペン、音声入力装置、接触式入力装置など、1つ以上の入力装置714も有することができる。ディスプレイ、スピーカ、プリンタなど、1つ以上の出力装置716も含むことができる。こうした装置はいずれも当技術分野で周知のものであり、ここでさらに論じる必要はない。

30

【0065】

音楽信号を符号化するのに効率的であり、かつ共通のLP合成フィルタを用いるハイブリッドコーデックで使用するのに適した、新規で有用な変換符号化方法を提供した。本発明の原理を応用することのできる多数の可能な実施形態を考慮すると、図面の図柄と関連して本明細書で説明した実施形態は、単に例示的なものに過ぎず、発明の範囲を制限するものと解釈すべきでないことは認識されよう。ここに説明した実施形態は、本発明の精神から逸脱することなく、その構成および詳細を変更できることは当業者に認識されよう。したがって、本発明は、DCT変換を利用するものとして説明したが、フーリエ変換や、変形離散コサイン変換など他の変換技術も本発明の範囲内で応用することができる。同様に、ここに説明した他の詳細事項も、本発明の範囲から逸脱せずに、変更または他のものに置き換えることができる。したがって、本明細書に記載した本発明は、そのような実施形態はすべて、頭記の特許請求の範囲およびその同等物の範囲内にあるものと企図する。

40

【符号の説明】

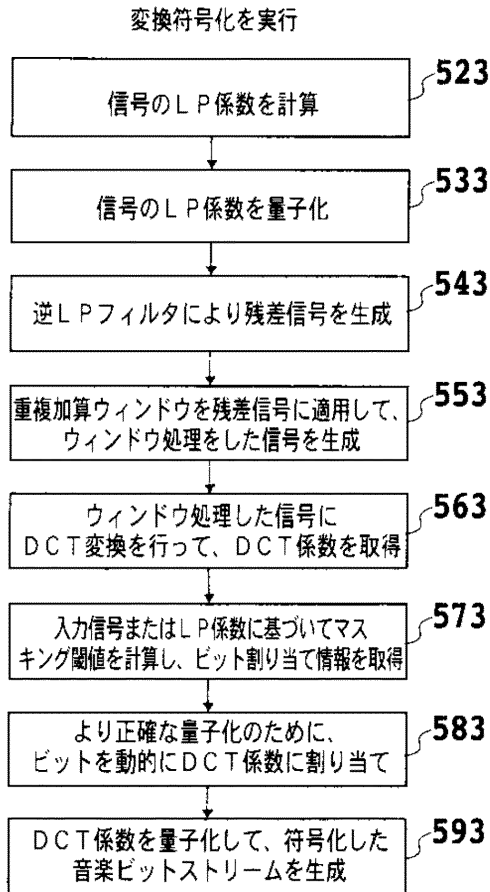
【0066】

- 100 ネットワーク
- 110、120 コーデック
- 111、121 符号変換器
- 112、122 復号器

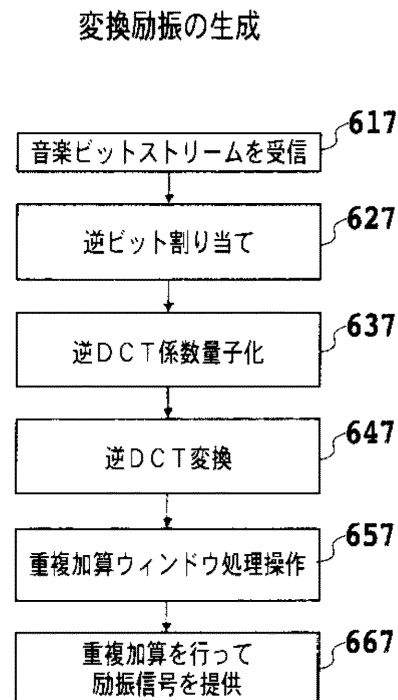
50

1 1 3、1 2 3、2 5 0	音声 / 音楽クラシファイア	
2 1 0	音声励振ジェネレータ	
2 2 0	変換励振ジェネレータ	
2 3 0	音声 / 音楽スイッチ	
2 4 0	線形予測合成フィルタ	
2 6 0	音声符号変換器	
2 7 0	音楽符号変換器	
3 1 0	線形予測解析フィルタ (L P C 解析モジュール)	
3 2 0	線形予測量子化モジュール	
3 3 0	逆線形予測フィルタ	10
3 4 0	重複加算モジュール (重複加算ウィンドウ処理モジュール)	
3 5 0	離散コサイン変換モジュール	
3 6 0	閾値マスキングモジュール	
3 7 0	動的ビット割り当てモジュール	
3 8 0	量子化モジュール	
3 9 0、3 9 1	重複加算レンジ	
4 1 0	逆動的ビット割り当てモジュール	
4 2 0	逆量子化モジュール	
4 3 0	D C T 逆変換モジュール	
4 4 0	非対称重複加算ウィンドウモジュール	20
4 5 0	重複加算モジュール	
7 0 0	コンピューティングデバイス	
7 0 2	処理装置	
7 0 4	メモリ	
7 0 8	取り外し可能ストレージ	
7 1 0	取り外し不能ストレージ	
7 1 2	通信接続	
7 1 4	入力装置	
7 1 6	出力装置	

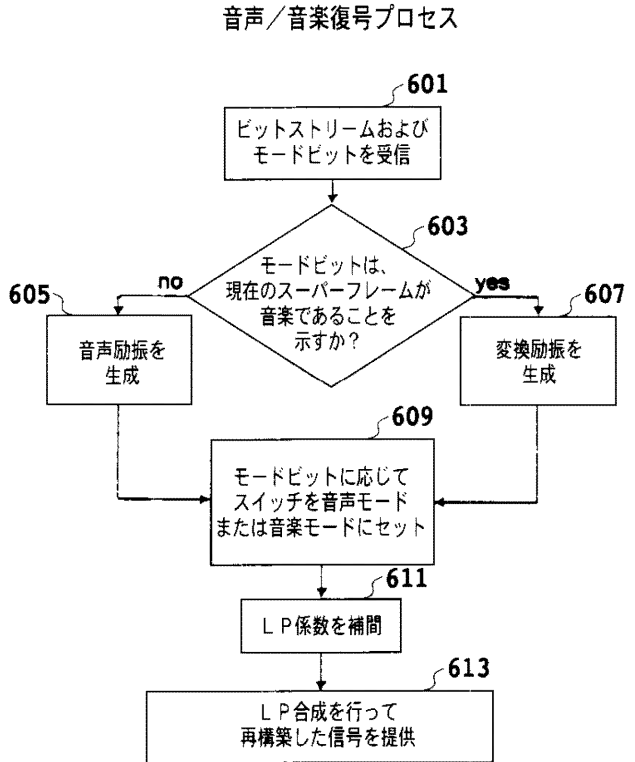
【図 6】



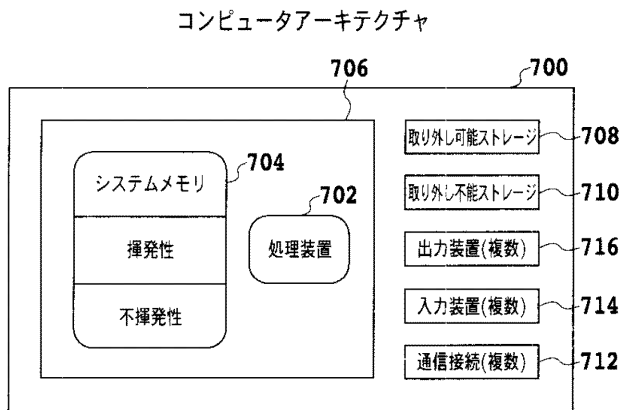
【図 8】



【図 7】



【図 9】



フロントページの続き

(72)発明者 ウラジミール カッパーマン

アメリカ合衆国 9 3 1 1 7 カリフォルニア州 ゴレタ シエロ アベニュー 5 6 3 5

(72)発明者 アミール エイチ・マジディメア

アメリカ合衆国 9 8 0 7 2 ワシントン州 ウッドビル 1 6 4 プレイス ノースイースト
1 4 8 2 4

(72)発明者 アレン ガーショ

アメリカ合衆国 9 3 1 1 1 カリフォルニア州 サンタ バーバラ ピア ジェニータ 4 6 0
4

F ターム(参考) 5J064 AA02 BA16 BB03 BC08 BC11 BC14 BC16 BC25 BD02 BD03