

(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2017-37328  
(P2017-37328A)

(43) 公開日 平成29年2月16日(2017.2.16)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 O L 19/00 (2013.01)	G 1 O L 19/00 4 0 0 A	
G 1 O L 25/90 (2013.01)	G 1 O L 25/90	
G 1 O L 25/93 (2013.01)	G 1 O L 25/93	

審査請求 有 請求項の数 20 O L (全 27 頁)

(21) 出願番号 特願2016-197150 (P2016-197150)  
 (22) 出願日 平成28年10月5日 (2016.10.5)  
 (62) 分割の表示 特願2015-94729 (P2015-94729) の分割  
 原出願日 平成23年6月23日 (2011.6.23)  
 (31) 優先権主張番号 61/361, 237  
 (32) 優先日 平成22年7月2日 (2010.7.2)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 510185767  
 ドルビー・インターナショナル・アーベー  
 オランダ王国, セーエン アムステルダム  
 ズイドーフト 1101, ヘリケルベル  
 グウェグ 1-35, アポロ ビルディン  
 グ 3エー  
 (74) 代理人 100107766  
 弁理士 伊東 忠重  
 (74) 代理人 100070150  
 弁理士 伊東 忠彦  
 (74) 代理人 100091214  
 弁理士 大貫 進介

最終頁に続く

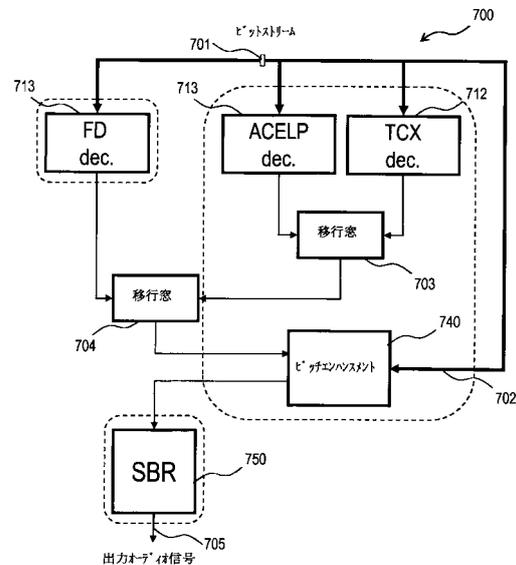
(54) 【発明の名称】 オーディオデコーダ及び復号方法

(57) 【要約】 (修正有)

【課題】異なる音源の成分を混合した信号のオーディオ符号化及び復号化の方法と装置を提供する。

【解決手段】オーディオデコーダは、第1の符号化モードで動作するように適応されたACELPモジュールと、第2の符号化モードで動作するように適応された、ACELPモジュールとは異なるTCXモジュールと、第1の符号化モードまたは第2の符号化モードのうちどちらかに含まれる、ACELPモジュールまたはTCXモジュールにより生成された予備オーディオ信号をフィルタするように適応されたピッチフィルタ740とを有する。ピッチフィルタはオーディオビットストリームにエンコードされた第1のパラメータの値に基づいて選択的にイネーブルまたはディスエーブルされ、第1のパラメータはオーディオビットストリームにエンコードされた第2のパラメータと異なり、第2のパラメータはオーディオデコーダの現在の符号化モードを指定する。

【選択図】 図7



**【特許請求の範囲】****【請求項 1】**

オーディオエンコーダにより生成されたオーディオビットストリームを復号するオーディオデコーダであって、

第 1 の符号化モードで動作するように適応された第 1 の復号モジュールと、

第 2 の符号化モードで動作するように適応された、前記第 1 の復号モジュールとは異なる第 2 の復号モジュールと、

前記第 1 の符号化モードまたは前記第 2 の符号化モードのうちどちらかに含まれるピッチフィルタであって、前記第 1 の復号モジュールまたは前記第 2 の復号モジュールにより生成された予備オーディオ信号をフィルタするように適応されたピッチフィルタとを有し

10

、  
前記ピッチフィルタは前記オーディオビットストリームにエンコードされた第 1 のパラメータの値に基づいて選択的にイネーブルまたはディスエーブルされ、前記第 1 のパラメータは前記オーディオビットストリームにエンコードされた第 2 のパラメータとは異なり、前記第 2 のパラメータは前記オーディオデコーダの現在の符号化モードを指定するものである、オーディオデコーダ。

**【請求項 2】**

前記ピッチフィルタに関連するピッチ情報は前記オーディオビットストリーム中の第 3 のパラメータから決定され、前記ピッチフィルタに関連するゲインは前記オーディオビットストリーム中の第 4 のパラメータから決定され、前記第 1 のパラメータ、第 2 のパラメータ、第 3 のパラメータ、及び第 4 のパラメータは異なる、  
請求項 1 に記載のオーディオデコーダ。

20

**【請求項 3】**

前記第 1 のパラメータは長さが 1 ビットであり、前記第 1 のパラメータの第 1 の値は前記ピッチフィルタをイネーブルし、前記第 1 のパラメータの第 2 の値は前記ピッチフィルタをディスエーブルする、  
請求項 1 に記載のオーディオデコーダ。

**【請求項 4】**

前記オーディオビットストリームはオーディオコンテンツのフレームに分割され、前記第 1 のパラメータはフレームタイプを示し、前記第 1 のパラメータの 1 以上の第 1 の値は前記ピッチフィルタをイネーブルし、前記第 1 のパラメータの 1 以上の第 2 の値は前記ピッチフィルタをディスエーブルする、請求項 1 に記載のオーディオデコーダ。

30

**【請求項 5】**

前記フレームタイプは、各フレームが有聲コンテンツを含む又は含みそうか、または前記各フレームが無声コンテンツを含む又は含みそうかを示す、  
請求項 4 に記載のオーディオデコーダ。

**【請求項 6】**

前記ピッチフィルタは、( i ) 前記ゲインパラメータをゼロに設定すること、( i i ) 前記ピッチフィルタをパススルーモードで動作させること、( i i i ) 前記ピッチフィルタをオフに切り替えること、または( i v ) 前記ピッチフィルタを非アクティブ化することによりディスエーブルされる、  
請求項 1 に記載のオーディオデコーダ。

40

**【請求項 7】**

前記ピッチフィルタの 1 以上のフィルタタップは前記オーディオビットストリーム中の第 5 のパラメータにより決定される、請求項 2 に記載のオーディオデコーダ。

**【請求項 8】**

前記ピッチフィルタはポストフィルタまたはピッチエンハンスメントフィルタである、  
請求項 1 に記載のオーディオデコーダ。

**【請求項 9】**

前記ポストフィルタ及びピッチエンハンスメントフィルタは、ハーモニクス間の信号成

50

分を減衰、またはスペクトルの谷を減衰するように適応されている、請求項 8 に記載のオーディオデコーダ。

【請求項 10】

前記ポストフィルタとピッチエンハンスメントフィルタは、前記予備オーディオ信号の周期的成分を回復するように適応されている、請求項 8 に記載のオーディオデコーダ。

【請求項 11】

前記第 1 の符号化モードは周波数領域符号化または変換符号化を含み、前記第 2 の符号化モードは線形予測符号化を含む、請求項 1 に記載のオーディオデコーダ。

【請求項 12】

前記ピッチフィルタは低周波数特性を有する、請求項 1 に記載のオーディオデコーダ。

10

【請求項 13】

符号化されたオーディオ信号のフレームを復号する方法であって、前記符号化されたオーディオ信号はオーディオエンコーダにより生成され、前記方法は、

前記符号化されたオーディオ信号から、前記フレームの符号化モードを表す第 1 のパラメータを取り出すステップと、

前記符号化されたオーディオ信号から第 2 のパラメータを取り出すステップと、

前記第 1 のパラメータの値に基づいて第 1 の復号モードまたは第 2 の復号モードのうちどちらかに入るステップと、

前記第 1 の復号モードまたは前記第 2 の復号モードで動作している間に、前記符号化されたオーディオ信号から予備オーディオ信号を生成するステップと、

20

前記予備オーディオ信号をピッチフィルタでフィルタするステップと、

前記フィルタするステップは、前記第 2 のパラメータを用いて前記ピッチフィルタを選択的にイネーブルまたはディスエーブルし、前記第 2 のパラメータは前記第 1 のパラメータとは異なる、方法。

【請求項 14】

前記符号化されたオーディオ信号から、前記ピッチフィルタのピッチ情報を表す第 3 のパラメータを取り出すステップと、

前記符号化されたオーディオ信号から、前記ピッチフィルタに関連するゲインを表す第 4 のパラメータを取り出すステップとをさらに含む、

請求項 13 に記載の方法。

30

【請求項 15】

前記第 1 の符号化モードは周波数領域符号化または変換符号化であり、前記第 2 の符号化モードは線形予測符号化である、

請求項 13 に記載の方法。

【請求項 16】

前記フィルタするステップは、前記予備オーディオ信号を長期的フィルタでフィルタし、前記長期的フィルタの出力を短期的フィルタでフィルタするステップを含む、

請求項 13 に記載の方法。

【請求項 17】

オーディオ信号とそれに関連するパラメータとを符号化して、オーディオビットストリームを生成するオーディオエンコーダであって、

40

第 1 の符号化モードで動作するように適応された第 1 の符号化モジュールと、

前記第 1 の符号化モードとは異なる第 2 の符号化モードで動作するように適応された第 2 の符号化モジュールと、

前記オーディオ信号が前記第 1 の符号化モジュールまたは前記第 2 の符号化モジュールのうちどちらで符号化すべきか決定する決定モジュールと、

前記第 1 の符号化モジュールまたは前記第 2 の符号化モジュールのどちらかから出力される予備オーディオ信号を符号化するように適応され、前記オーディオビットストリームを生成する第 3 の符号化モジュールとを有し、

前記第 3 の符号化モジュールは、さらに、第 1 のパラメータと第 2 のパラメータとを符

50

号化するように適応され、前記第 1 のパラメータはピッチフィルタを選択的にイネーブルまたはディスエーブルするのに用いられ、前記第 2 のパラメータは前記エンコーダの現在の符号化モードを示し、前記第 1 のパラメータは前記第 2 のパラメータとは異なる、オーディオエンコーダ。

【請求項 18】

前記第 3 の符号化モジュールは、さらに、第 3 のパラメータと第 4 のパラメータとをオーディオビットストリームに符号化するように適応され、前記第 3 のパラメータは前記ピッチフィルタに関連するピッチを表し、前記第 4 のパラメータは前記ピッチフィルタに関連するゲインパラメータを表す、  
請求項 17 に記載のオーディオエンコーダ。

10

【請求項 19】

前記第 1 のパラメータは長さが 1 ビットであり、前記第 1 のパラメータの第 1 の値は前記ピッチフィルタをイネーブルし、前記第 1 のパラメータの第 2 の値は前記ピッチフィルタをディスエーブルする、  
請求項 17 に記載のオーディオエンコーダ。

【請求項 20】

前記オーディオビットストリームはオーディオコンテンツのフレームに分割され、前記第 1 のパラメータはフレームタイプを示し、前記第 1 のパラメータの 1 以上の第 1 の値は前記ピッチフィルタがイネーブルされること、前記第 1 のパラメータの 1 以上の第 2 の値は前記ピッチフィルタがディスエーブルされることを示す、  
請求項 17 に記載のオーディオエンコーダ。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、概してデジタルオーディオコーディングに関し、より詳しくは異なるキャラクタの成分を含むオーディオ信号のコーディング手法に関する。

【背景技術】

【0002】

広く普及した、スピーチや歌唱を含むオーディオ信号のコーディング方法のクラスには、特に音楽に適合した周波数ドメインコーディング方法や一般的性質の方法を含む複数の異なるコーディング方法を時間的に切り換えて用いて、オーディオ信号の連続した期間の間のキャラクタの変化に対応する CELP (code excited linear prediction) がある。例えば、単純化された MPEG (Moving Pictures Experts Group) USAC (Unified Speech and Audio Coding; ISO/IEC 23003-3 標準を参照) デコーダは、図 2 の上部に示したように、少なくとも 3 つの符号化モードで、すなわち AAC (Advanced Audio Coding; ISO/IEC 13818-7 標準を参照)、ACELP (algebraic CELP)、及び TCX (transform-coded excitation) で動作する。

30

【0003】

CELP の様々な実施形態が、人間の発話器官の特性に、及び場合によっては人間の聴覚に適合されている。本出願では、CELP は、可能性のあるすべての実施形態や変形例を指し、ACELP、wide- and narrow-band CELP、SB-CELP (sub-band CELP)、low- and high-rate CELP、RCELP (relaxed CELP)、LD- CELP (low-delay CELP)、CS-CELP (conjugate-structure CELP)、CS-ACELP (conjugate-structure ACELP)、PSI-CELP (pitch-synchronous innovation CELP) 及び VSELP (vector sum excited linear prediction) を含むが、これらに限定されない。CELP の原理は、非特許文献 1 で説明されており、そのいくつかの応用は非特許文献 2 で引用された参考文献 25 - 29 に説明されている。前者の論文に詳しく説明されているように、CELP デコーダ (又は、同様に CELP スピーチシンセサイザ) は、ピッチ予測器とパルスコードブックとを含む。ピッチ予測器は、符号化されたスピーチ信号の周期的成分を回復する。パルスコードブックからは、イノベーションシーケンスが付加される。ピッチ予測器は、ピッチを回復する

40

50

長遅延予測器と、スペクトルエンベロープ形成によりフォルマントを回復する短遅延予測器とを含む。ここでは、ピッチは概して、声帯により生じ、さらに声道の共鳴部分により色づけされる調性音の基本周波数として理解される。この周波数はその高調波とともにスピーチや歌唱を特色付ける。一般的に言って、CELP法は、ソロ又は1つのパートの歌唱であって、ピッチ周波数が明確に定義され、決定が比較的容易なものの処理に最も適している。

【0004】

CELP符号化したスピーチの知覚品質を改善するため、CELP法をポストフィルタリング（あるいは、言い換えるとピッチエンハンスメント）と組み合わせることがよく行われる。特許文献1と非特許文献2のセクション2には、かかるポストフィルタの望ましい特性が、すなわち検出されたボイスピッチ（長期部分；セクション4を参照）の高調波間にあるノイズ成分を抑制する機能が開示されている。このノイズの重要部分はスペクトルエンベロープ形成に由来すると信じられている。図1と図2に示すように、単純なポストフィルタの長期部分は、伝達関数

【数1】

$$H_E(z) = 1 + \alpha \left( \frac{z^T + z^{-T}}{2} - 1 \right)$$

を有するように設計される。ここで、Tはサンプル数に対する推定ピッチ期間であり、 $\alpha$ はポストフィルタのゲインである。かかるフィルタは、コムフィルタと同様に、周波数  $1/(2T)$ 、 $3/(2T)$ 、 $5/(2T)$ （これらはピッチ周波数の高調波間の間にある）と隣接周波数とを減衰する。減衰はゲイン  $\alpha$  の値に依存する。もう少し高度なポストフィルタは、この減衰を、ノイズが最も知覚される低周波数だけに適用する - そのため、一般的にはパスポストフィルタと言う。これは、上記の伝達関数  $H_E$  とローパスフィルタ  $H_{LP}$  をカスケードすることにより表せる。よって、ポストフィルタにより出力される後処理した符号化  $S_E$  は、変換領域では

【数2】

$$S_E(z) = S(z) - \alpha S(z) P_{LT}(z) H_{LP}(z)$$

により与えられる。ここで、

【数3】

$$P_{LT}(z) = 1 - \frac{z^T + z^{-T}}{2}$$

であり、Sはポストフィルタへの入力として供給される符号化信号である。図3は、これらの特徴を有するポストフィルタの一実施形態である。このポストフィルタは、非特許文献3のセクション6.1.3でさらに説明されている。この図が示唆するように、ピッチ情報はビットストリーム信号中のパラメータとして符号化され、 $P_{LT}$ により表される演算を実行する長期予測フィルタに通信可能に接続されたピッチトラッキングモジュールにより読み出される。

【0005】

前出の параグラフで説明した長期部分のみを用いてもよい。あるいは、フォルマントに対応する周波数区間の成分を保存し、他のスペクトル領域（短期部分；セクション3を参照）の、すなわちフォルマントエンベロープの「スペクトルの谷」のノイズを減衰するノイズ成形フィルタと直列に構成される。可能性のある他の変形例として、このフィルタ集

合は、短期部分のスペクトル傾斜による知覚される劣化を低減する漸進的ハイパスタイプフィルタによりさらに補完される。

【0006】

調性、非調性、ボーカル、楽器、非音楽など音源が異なる成分がミックスされたオーディオ信号は、必ずしも今あるデジタルコーディング技術で満足がいく再生をできるとは限らない。より詳しく言うと、今ある技術では、このような不均質なオーディオ素材を処理するには不足であり、ある成分には適していても、他の成分には適していないことがある。具体的には、歌と1つ又は複数の楽器又はコーラスパートとがある音楽であって、上記の性質の方法で符号化されたものを復号すると、知覚でき、リスニング体験の一部を損ねるアーティファクトが生じることが多い。

10

【先行技術文献】

【特許文献】

【0007】

【特許文献1】米国特許第4969192号

【非特許文献】

【0008】

【非特許文献1】R. Schroeder and S. Atal in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 10, pp. 937-940, 1985

【非特許文献2】Chen and Gersho, IEEE Transactions on Speech and Audio Processing, vol. 3, no. 1, 1995

20

【非特許文献3】Technical Specification ETSI TS 126 290, version 6.3.0, release 6

【発明の概要】

【発明が解決しようとする課題】

【0009】

前のセクションで説明した欠点の少なくとも一部を緩和するために、本発明の目的は、異なる音源の成分を混合した信号のオーディオ符号化及び復号をするように構成された方法と装置を提供することである。具体的な目的として、本発明は、コーディング効率又は（知覚される）再生フィデリティ又はその両方の観点から適した方法と装置を提供しようとするものである。

30

【課題を解決するための手段】

【0010】

本発明は、独立請求項に規定したように、エンコーダシステム、デコーダシステム、符号化方法、復号方法、及びこれらの各方法を実行するコンピュータプログラム製品を提供することにより、これらの目的のうち少なくとも一を達成する。従属項は本発明の実施形態を規定している。

【0011】

発明者は、不均一音源 (non-homogeneous origin) の符号化オーディオ信号において知覚される一部のアーティファクトは、複数の符号化モードであってそのうちの少なくとも一がデコーダにポストフィルタリングを含み、少なくとも一は含まない複数のコーディングモードの不適切なスイッチングにより生じること気づいた。より正確には、利用可能なポストフィルタが高調波ノイズ（及び、場合によっては、スペクトルの谷のノイズ）だけでなく、「所望の」性質を有する楽器やボーカル伴奏その他の素材を表す信号成分も除去する。(Ghitza and Goldstein, IEEE Trans. Acoust, Speech, Signal Processing, vol. ASSP-4, pp. 697-708, 1986に記載されているように) スペクトルの谷の丁度可知差異は10dBであることが、多くの設計者によりこれらの周波数帯域を強くフィルタリングする正当化として考えられている。しかし、高調波間（及びスペクトルの谷の）減衰自体による音質低下は、スイッチングの場合の音質低下より重要度は低い。ポストフィルタをオンにスイッチングすると、歌っている声のバックグラウ

40

50

ンドが急に小さくなり、フィルタを非アクティブにすると、バックグラウンドがすぐに大きくなる。スイッチングが頻繁に行われると、オーディオ信号の性質により、又はコーディング装置の構成により、スイッチングアーティファクトが生じる。一例として、USACデコーダは、ポストフィルタリングと組み合わせたACELPモードで、又はポストフィルタリング無しのTCXモードで動作できる。ACELPモードは、支配的なボーカル成分がある場合に使われる。よって、ACELPモードへのスイッチングは、新しい音楽フレーズの始めなど歌の始めで、新しいパースの始めで、又はボーカル成分が目立たないという意味で伴奏が歌声をかき消した後などにトリガーされる。実験により確認したところ、TCXコーディングを通して使う代替的ソリューションでは、又はむしろ問題の迂回では、反響状のアーティファクトが現れるので、問題の解決にならない。

10

**【0012】**

したがって、第1と第2の態様では、本発明は、符号化方法により出力されるビットストリームを復号する装置が高調波間ノイズの減衰を含むポストフィルタリングを適用すべきか決定することを特徴とするオーディオ符号化方法（及び対応する機能を有するオーディオ符号化システム）を提供する。決定の結果はビットストリームにエンコードされ、復号装置にとってアクセス可能である。

**【0013】**

本発明により、ポストフィルタを使うかの決定は、最も適したコーディングモードに関する決定とは別に行われる。これにより、スイッチングがリスナをわずらわさないだけの時間にポストフィルタリング状態を維持可能になる。よって、本符号化方法により、従来はフィルタがアクティブであったコーディングモードにスイッチングされても、ポストフィルタを非アクティブに保つように指示する。

20

**【0014】**

留意点として、ポストフィルタリングを適用するかの決定は、通常、フレームごとに行われる。よって、第1に、ポストフィルタリングは一度に一フレーム未満には適用されない。第2に、ポストフィルタリングをディスエーブルするかの決定は、現在のフレーム中においてのみ有効であり、後続のフレームにおいてはそのまま維持されても、再評価されてもよい。主フレームフォーマットと、通常のフォーマットの一部（例えば、1/8の長さ）である短いフォーマットとを許容するコーディングフォーマットでは、個々の短いフレームに対してポストフィルタリングの決定をする必要はない。代わりに、まとめると通常フレームになる複数の短いフレームを考え、その平均又はメジアンを計算して、フィルタリング決定に関するパラメータを求めてもよい。

30

**【0015】**

本発明の第3と第4の態様では、高調波間ノイズ減衰を含み、ビットストリーム信号にエンコードされたポストフィルタリング情報によりポストフィルタをディスエーブルするステップを有する、ポストフィルタリング段階が続く復号ステップを有するオーディオ復号方法（及び対応する機能を有するオーディオ復号システム）が提供される。

**【0016】**

これらの特徴を有する復号方法は、ポストフィルタリング情報のみにより、そのため現在のコーディングモードなどのファクタとは独立して、ポストフィルタを非アクティブ化する機能があるので、混合音源オーディオ信号のコーディングに適している。従来のようにポストフィルタ機能が特定のコーディングモードに付随していたコーディング方法に適用する場合、ポストフィルタリングをディスエーブルする機能により、新しい動作モード、すなわち従来のフィルタリングされた復号モードをフィルタリングをせずに適用するモードが可能となる。

40

**【0017】**

さらに別の態様では、本発明は、上記の方法の一を実行するコンピュータプログラム製品も提供される。さらにまた、本発明により、ポストフィルタに供給されるポストフィルタリング信号により示される、アクティブモード又はパススルーモードで動作可能な、高調波間ノイズを減衰するポストフィルタが提供される。このポストフィルタは、ポストフ

50

フィルタリング動作を自律的に制御する決定部を有していてもよい。

【0018】

当業者には言うまでもないが、デコーダと協働するように構成されたエンコーダには、機能的に等価なモジュールが備えられており、符号化された信号の忠実な再生が可能になっている。かかる等価なモジュールは、同じ又は類似したモジュールか、同じ又は類似した伝達特性を有するモジュールである。特に、エンコーダとデコーダのモジュールは、それぞれ、等価な数学的演算を実行するコンピュータプログラムを実行する類似した、または非類似のユニットであってもよい。

【0019】

一実施形態では、本方法の符号化は、(フォルマントエンベロープ(上記参照)の)スペクトルの谷の減衰をさらに含むポストフィルタに関する意思決定を含む。これはポストフィルタの短期部分に対応する。決定が基づく基準をポストフィルタの性質に適合させると都合がよい。

【0020】

一実施形態は、特にスピーチ符号化に適合したエンコーダに関する。本発明の動機付けとなる問題は、ポータルとその他の成分の混合を符号化する時に見られるので、本発明によるスピーチコーディングとポストフィルタリングに関する独立した意思決定との組合せは、特に都合がよい。具体的に、かかるデコーダは、CELP符号化モジュールを含み得る。

【0021】

一実施形態では、エンコーダによる決定は、支配的基本周波数(ピッチ)を有する信号成分と、基本周波数より下にある他の信号成分が同時に存在することの検出に基づく。また、検出は、支配的基本周波数を有する成分と、この基本周波数の高調波間にエネルギーを有する他の成分との共起(co-occurrence)の発見を目的としてもよい。これは、問題のタイプのアーティファクトが頻繁に生じる場合である。よって、このように同時に存在することが分かったら、エンコーダは、ポストフィルタリングが適していないと決定する。これはビットストリーム中に含まれるポストフィルタリング情報により示される。

【0022】

一実施形態では、その検出基準として、ピッチ周波数より下のオーディオ時間信号中の全信号パワーコンテンツを用いる。ピッチ周波数は、場合によっては、エンコーダの長期予測により推定されたピッチ周波数である。これが所定閾値より大きいとき、ピッチ成分(高調波を含む)以外の成分があると考えられる。そのため、ポストフィルタはディスエーブルされる。

【0023】

CELPモジュールを有するエンコーダでは、かかるモジュールはオーディオ時間信号のピッチ周波数を推定することを用いることができる。そして、詳細は上記したように、さらに別の検出基準は、この周波数の高調波間の、又はそれより下のエネルギーコンテンツをチェックすることである。

【0024】

CELPモジュールを含む前述の実施形態のさらなる発展として、決定は、CELPコード化(すなわち、符号化及び復号)された時のオーディオ信号の推定パワー間と、CELPコード化及びポストフィルタされた時のオーディオ信号の推定パワーとの比較を含み得る。パワー差が閾値より大きく場合、これは信号の非ノイズ成分が失われることを示し、エンコーダはポストフィルタをディスエーブルする決定をする。

【0025】

有利な実施形態では、エンコーダはCELPモジュールとTCXモジュールとを有する。本技術分野では知られているように、TCXコーディングはある種の信号、特に非ポータル信号には有利である。TCXコーディングした信号にポストフィルタリングをかけるのは実務上一般的ではない。よって、エンコーダは、TCXコーディング、ポストフィルタリングを伴うCELPコーディング、又はポストフィルタリングを伴わないCELPコ

10

20

30

40

50

ーディングのいずれかを選択して、広い範囲の信号タイプをカバーできる。

【0026】

前述の実施形態のさらなる発展として、3つのコーディングモード間の決定は、レート・歪み基準に基づき、すなわち本技術分野でそれ自体知られた最適化手順を適用して行う。

【0027】

前述の実施形態の他のさらなる発展では、エンコーダはさらに、AAC (Advanced Audio Coding) コーダを有する。AAC コーダもある種の信号に対して特に適していることが知られている。好ましくは、AAC (周波数領域) コーディングを適用するかの決定は、他のどの(線形予測)モードを用いるかに関する決定とは別に行われる。このように、エンコーダは、2つのスーパーモード、すなわちAAC又はTCX CELPで動作するととらえることができ、後者では、エンコーダはTCX、ポストフィルタリングを伴うCELP、又はフィルタリングを伴わないCELPを選択する。この実施形態により、さらに広い範囲のオーディオ信号タイプの処理が可能となる。

【0028】

一実施形態では、エンコーダは、復号におけるポストフィルタリングを徐々に適用することを、すなわちゲインを徐々に上げることを決定できる。同様に、ポストフィルタリングを徐々に除去することを決定できる。このように徐々に適用したり除去したりすることによって、ポストフィルタリングの有無の切替が知覚しづらくなる。一例として、ポストフィルタリングを伴うCELPコーディングが好適である歌唱部分には、その前に、TCXコーディングが最適なインストルメンタル部分があるとする。本発明によるデコーダは、歌唱部分の始め又はその近くでポストフィルタリングを徐々にかけて、煩わしいスイッチングアーティファクトを回避しつつ、ポストフィルタリングの利益を保つようにできる。

【0029】

一実施形態では、ポストフィルタリングを適用するかの決定は、近似差分信号に基づき行われる。この信号は、ポストフィルタにより復号される信号から除かれる信号成分を近似するものである。一オプションとして、近似差分信号を、オーディオ時間信号と、(シミュレーションされた)ポストフィルタリングをかけたオーディオ時間信号との間の差分として、計算する。他の一オプションとして、符号化部は、中間復号信号を取り出す。これにより、オーディオ時間信号と、ポストフィルタリングをかけた中間復号信号との間の差分として、近似差分信号を計算できる。中間復号信号はエンコーダの長期予測バッファに格納できる。これは信号の励起を表し、さらなる合成フィルタリング(ボークトラクト、共鳴)を適用して、最終的な復号信号を得る必要があることを示唆する。中間復号信号を用いるポイントは、中間復号信号がコーディング方法の特殊性、特に弱点の一部を捉え、それによりポストフィルタの効果の推定がより現実的になることにある。第3のオプションとして、復号部は、中間復号信号を取り出す。これにより、中間復号信号と、ポストフィルタリングをかけた中間復号信号との間の差分として、近似差分信号を計算できる。この手順による推定は、おそらく、最初の2つのオプションによる推定より信頼性は高くないが、他方で、デコーダによりスタンドアロンで実行できる。

【0030】

このように求めた近似差分信号は、次の基準:

- a) 近似差分信号のパワーが所定の閾値を越え、信号の大部分がポストフィルタにより除去されるか、
- b) 前記近似差分信号の特徴がノイズ状でなく音調的であるか;
- c) 近似差分信号の強度周波数スペクトルとオーディオ時間信号の強度周波数スペクトルとの間の差分が、周波数について不均一に分散しており、ノイズではなく、人間のリスナにとって意味を成す信号であるか、
- d) 近似差分信号の強度周波数スペクトルは、処理されるタイプの信号から期待できる、所定の関連エンベロープ内の周波数区間にあるか、

10

20

30

40

50

e) 前記近似差分信号の強度周波数スペクトルが、所定のスケールファクタによりダウンスケールされた最大信号成分の強度により前記オーディオ時間信号の強度周波数スペクトルに閾値を設けることにより得られた関連エンベロープ内の周波数区間にあるか；のーについて評価される。肯定的に評価されると、ポストフィルタをディスエーブルする決定がなされる。

#### 【0031】

基準e)を評価する時、強度スペクトルにピークトラッキングを適用すると有利である。すなわち、ノイズではなく通常は音調的成分に関連するピーク形状を有する部分を区別すると有利である。本技術分野でそれ自体知られている何らかのアルゴリズムにより行われるピークトラッキングにより特定される成分は、さらに、ピークの高さに閾値を適用することによりソートされる。残りの成分はある強度を有する音調素材である。かかる成分は、通常は、ノイズではなく関連信号コンテンツを表し、これはポストフィルタをディスエーブルする決定の動機付けとなる。

#### 【0032】

デコーダとしての本発明の一実施形態では、ポストフィルタをディスエーブルする決定は、制御部により制御可能であり、回路においてポストフィルタをバイパスできるスイッチにより実行される。他の一実施形態では、ポストフィルタは制御部により又はその中のゲインコントローラにより制御可能な可変ゲインを有し、ディスエーブルする決定は、ポストフィルタゲイン(前のセクションを参照)をゼロに設定することにより、又はその絶対値を所定閾値より下に設定することにより、行われる。

#### 【0033】

一実施形態では、本発明による復号は、復号されるビットストリーム信号からのポストフィルタリング情報の取り出しを含む。より正確には、ポストフィルタリング情報は、伝送に適したフォーマットの、少なくとも1ビットを含むデータフィールドにエンコードされている。有利にも、データフィールドは、適用可能な標準により画定されたが使われていない既存のフィールドであり、ポストフィルタリング情報は伝送されるペイロードを増加させない。

#### 【0034】

留意点として、このセクションで開示した方法と装置は、通常の実験を含む当業者の能力の範囲内で適当な修正をして、ステレオチャンネルなどの異なる複数のチャンネルに対応する、複数の成分を有する信号のコーディングに適用できる。本願を通して、ピッチエンハンスメントとポストフィルタリングは同意語として用いられている。さらに留意点として、AACは、周波数領域のコーディング方法の代表例として説明されている。実際、本発明を、AAC以外の周波数領域コーディングモードで動作可能なデコーダやエンコーダに適用するには、必要であれ、当業者の能力範囲内の小さな修正を加えるだけでよい。同様に、TCXは、重み付け線形予測変換コーディング及び変換コーディング一般の例として説明したものである。

#### 【0035】

さらに別の実施形態では、上記の2つ以上の実施形態の特徴を、明らかに補完的でない限り、組み合わせられる。2つの特徴が異なるクレームに記載されていても、それらを組み合わせられないと言うわけではない。同様に、さらに別の実施形態では、所望の目的に対して必要でない、または本質的でない特徴を省略してもよい。

#### 【図面の簡単な説明】

#### 【0036】

ここで添付した図面を参照して、本発明の実施形態を詳細に説明する。

【図1】ポストフィルタを伴う従来のデコーダを示すブロック図である。

【図2】AAC、ACELP、及びTCXモードで動作可能であり、ACELPモジュールの下流に接続されたポストフィルタを含む従来のデコーダを示すブロック図である。

【図3】ポストフィルタの構成を示すブロック図である。

【図4】本発明によるデコーダを示すブロック図である。

10

20

30

40

50

【図 5】本発明による他のデコーダを示すブロック図である。

【図 6】従来のデコーダを示すブロック図である。

【図 7】本発明によるデコーダを示すブロック図である。図 6 と図 7 により、従来のデコーダと本発明によるデコーダとの相違点を示す。

【図 8】本発明によるエンコーダを示すブロック図である。

【図 9】従来のデコーダを示すブロック図である。

【図 10】本発明によるデコーダを示すブロック図である。図 9 と図 10 により、従来のデコーダと本発明によるデコーダとの相違点を示す。

【図 11】選択的にアクティブにしたり非アクティブにできる自律的ポストフィルタを示すブロック図である。

10

【発明を実施するための形態】

【0037】

図 4 は、ビットストリーム信号が入力され、オーディオ信号が出力される本発明の一実施形態によるデコーダシステム 400 を示す図である。図 1 に示した従来のデコーダのように、ポストフィルタ 440 は、復号モジュール 410 の後段に配置され、スイッチ 442 の動作により復号経路に加えられたりはされたりできる。ポストフィルタは、図に示したスイッチ位置でイネーブルされる。ポストフィルタがディスエーブルされるのは、スイッチが反対位置に設定され、それにより復号モジュール 410 からの信号がバイパスライン 444 に流れたときである。発明的貢献として、スイッチ 442 はビットストリーム信号に含まれるポストフィルタリング情報により制御可能であり、復号モジュール 410 の現在の状態にかかわらずポストフィルタリングを適用したり除いたりできるようになっている。ポストフィルタ 440 は多少遅れて動作するので（例えば、図 3 に示したポストフィルタは少なくともピッチ期間 T の遅延を生じる）、補償遅延モジュール 443 をバイパスライン 444 に配置し、スイッチング時にモジュールを同期状態に維持する。遅延モジュール 443 は、ポストフィルタ 440 と同じ時間だけ信号を遅延させるが、信号の処理はしない。切替時間を最小にするため、補償遅延モジュール 443 は、常にポストフィルタ 440 と同じ信号を受け取る。ポストフィルタ 440 をゼロ遅延ポストフィルタ（例えば、2 つのタップを有し、将来の信号値には依存しない因果性フィルタ (causal filter) など）と置き換えた実施形態では、補償遅延モジュール 443 は省略できる。

20

【0038】

図 5 は、図 2 のトリプルモードデコーダシステム 500 の発明の教示をさらに発展させたものである。ACELP 復号モジュール 511 は、TCX 復号モジュール 512 及び AAC 復号モジュール 513 と並列に配置されている。ACELP 復号モジュール 511 と直列に、ノイズを減衰するポストフィルタ 540 が配置されている。ポストフィルタ 540 は、具体的には、デコーダシステム 500 が合わせられたビットストリーム信号から直接的又は間接的に求まるピッチ周波数の高調波間にあるノイズを減衰するものである。また、ビットストリーム信号は、図 4 に示したように、上側スイッチ 541 の設定 (positions) を制御するポストフィルタリング情報をエンコードしている。この上側スイッチ 541 は、ポストフィルタ 540 を処理経路からはずすスイッチをし、補正遅延 543 と置き換えるように動作する。下側スイッチ 542 は、異なる復号モード間のスイッチをするのに用いられる。このように構成されているので、TCX モジュール 512 又は AAC モジュール 513 の一方を用いる場合、上側スイッチ 541 の設定 (position) は重要ではない。よって、ポストフィルタリング情報は、ACELP モードを除き、必ずしもこの設定 (position) を示さない。現在どんな復号モードが使われていても、信号は、下側スイッチ 542 のダウンストリーム接続点からスペクトル帯域複製 (SBR) モジュール 550 に供給され、SBR モジュール 550 はオーディオ信号を出力する。当業者には言うまでもないが、図面は概念的なものであり、図に示したスイッチは可動接触手段を有する物理的には別々の実体として示した。デコーダシステムの実際の実施形態では、スイッチやその他のモジュールはコンピュータ読み取り可能命令により実施される。

40

【0039】

50

図6と図7も、ACELP、TCX又は周波数領域復号モードで動作可能なトリプルモードデコーダシステムを示すブロック図である。図7を参照して、本発明の一実施形態を示すが、ビームストリーム信号は、入力点701に供給される。入力点701は、それぞれの分岐を介して復号モジュール711、712、713に永続的に接続されている。また、入力点701は、ピッチエンハンスメントモジュール740への接続分岐702（これは、図6の従来の復号システムには無い）を有する。このピッチエンハンスメントモジュール740は、上記の一般的なタイプのポストフィルタとして機能する。本技術分野における一般的な実務として、第1の移行窓モジュール703がACELPモジュール711とTCXモジュール712の下流に配置されていて、復号モジュール間の移行を行う。第2の移行モジュール704は、周波数領域復号モジュール713と第1の移行窓モジュール703の下流に配置され、2つのスーパーモード間の移行を行う。さらに、SBRモジュール750は、出力点705のすぐ上流に設けられている。明らかに、ビットストリーム信号は、すべての復号モジュール711、712、713及びピッチエンハンスメントモジュール740に直接的に（又は、場合によっては逆多重後に）供給される。ビットストリームに含まれる情報が、どの復号モジュールをアクティブ化するか制御する。しかし、本発明により、ピッチエンハンスメントモジュール740は類似の自己起動を行う。これは、ビットストリーム中のポストフィルタリング情報に応じて、ポストフィルタ又は単なるパススルーとして機能する。これは、例えば、ピッチエンハンスメントモジュール740に制御部（図示せず）を設け、その制御部によりポストフィルタリング機能をオン・オフすることにより実現できる。デコーダシステムが周波数領域復号モード又はTCX復号モードで動作している時、厳密に言えば、ポストフィルタリング情報は不要であり、ピッチエンハンスメントモジュール740は、常にパススルーモードにある。言うまでもなく、本発明の貢献の部分構成せず、当業者にはあることが自明なモジュールは、例えばデマルチプレクサは、明りょうかのために図7その他の同様の図面では省略した。

10

20

30

40

50

【0040】

バリエーションとして、図7のデコーダシステムには、合成による分析アプローチを用いてポストフィルタリングを適用するか決定する制御モジュール（図示せず）を設けてもよい。かかる制御モジュールは、ピッチエンハンスメントモジュール740とACELPモジュール711に通信可能に接続され、復号プロセスの中間段階を表し、好ましくは信号の励起に対応したものを表す中間復号信号  $s_{i\_DEC}(n)$  を取り出す。検出モジュールは、伝達関数  $P_{LT}(z)$  と  $H_{LP}(z)$ （背景技術欄及び図3を参照）、又は等価であるがこれらのフィルタインパルス応答  $p_{LT}(z)$  と  $h_{LP}(n)$  で画定されるピッチエンハンスメントモジュール740の機能をシミュレーションするのに必要な情報を有する。背景技術欄で説明したように、ポストフィルタリングで減算される成分は、

[外1]

$$[(s_{i\_DEC} * p_{LT}) * h_{LP}](n)$$

に比例する近似差分信号  $s_{AD}(n)$  により推定できる。ここで、\*は離散畳み込みを示す。これは、元のオーディオ信号とポストフィルタリングされた復号信号との間の真の差分の近似である。すなわち、

【数4】

$$s_{ORIG}(n) - s_E(n) = s_{ORIG}(n) - \alpha[s_{DEC} * p_{LT} * h_{LP}](n),$$

であり、ここで  $\alpha$  はポストフィルタゲインである。発明の概要欄及び特許請求の範囲に開示したように、この信号の全エネルギー、低周波帯域エネルギー、実強度スペクトル、又は過去の強度スペクトルが分かれば、制御部はピッチエンハンスメントモジュール740

をアクティブ化又は非アクティブ化するか決定する基本を見つけることができる。

【 0 0 4 1 】

図 8 は、本発明の一実施形態によるエンコーダシステム 8 0 0 を示す図である。エンコーダシステム 8 0 0 は、デジタルオーディオ信号を処理するように構成され、そのデジタルオーディオ信号は一般的にマイクロホンで音波をキャプチャし、その音波をアナログの電気信号に変換することにより取得される。電気信号は、サンプリングされて、エンコーダシステム 8 0 0 に好適なフォーマットで提供され得るデジタル信号にされる。このシステムは、一般的には、符号化モジュール 8 1 0、決定モジュール 8 2 0、及びマルチプレクサ 8 3 0 よりなる。(象徴的に表された)スイッチ 8 1 4、8 1 5 により、符号化モジュール 8 1 0 は、モジュール 8 1 1、8 1 2、8 1 3 を選択的にアクティブ化することにより、CELP、TCX、又は AAC モードのいずれかにより動作可能である。決定モジュール 8 2 0 は、所定の基準を適用して、符号化システム 8 0 0 により生成されるビットストリーム信号の復号中にポストフィルタリングをディスエーブルしてオーディオ信号をエンコードするか決定する。この目的のため、決定モジュール 8 2 0 は、オーディオ信号を直接調べて、又は接続ライン 8 1 6 を介して符号化モジュール 8 1 0 からデータを受け取る。決定モジュール 8 2 0 によりなされる決定を示す信号は、符号化モジュール 8 1 0 からの符号化オーディオ信号とともに、マルチプレクサ 8 3 0 に供給され、マルチプレクサ 8 3 0 はこれらの信号を連結して、エンコーダシステム 8 0 0 の出力を構成するビットストリームにする。

10

【 0 0 4 2 】

好ましくは、決定モジュール 8 2 0 は、中間復号信号  $s_{i\_DEC}$  から計算した近似差分信号に基づいて決定をする。この中間復号信号  $s_{i\_DEC}$  は符号化モジュール 8 1 0 から取得できる。前述の通り、中間復号信号は復号プロセスの中間段階を表すが、符号化プロセスの対応する段階から得ることもできる。しかし、エンコーダシステム 8 0 0 では、有利にも近似差分信号が

20

【 数 5 】

$$s_{ORIG}(n) - (s_{i\_DEC}(n) - \alpha [(s_{i\_DEC} * p_{LT}) * h_{LP}](n))$$

30

となるように、元のオーディオ信号  $s_{ORIG}$  を得ることができる。近似した点は、最終的復号信号の替わりに中間復号信号を用いている。これにより、復号でポストフィルタが除去する成分の性質を評価でき、発明の概要欄で説明した基準の 1 つを適用することにより、決定モジュール 8 2 0 はポストフィルタリングをディスエーブルするか決定をすることができる。

【 0 0 4 3 】

このバリエーションとして、決定モジュール 8 2 0 は、中間復号信号の替わりに元の信号を用いて、近似差分信号が

[ 外 2 ]

$$[(s_{i\_DEC} * p_{LT}) * h_{LP}](n)$$

40

となるようにしても良い。これは忠実さが低い近似となるが、他方では、決定モジュール 8 2 0 と符号化モジュール 8 1 0 との間の接続ライン 8 1 6 の在否が任意的となる。

【 0 0 4 4 】

決定モジュール 8 2 0 がオーディオ信号を直接調べる (study) する本実施形態の他のバリエーションでは、以下の基準の 1 つ又は複数が適用できる：

・オーディオ信号は基本周波数が支配的な成分と、基本周波数よりも低い成分とを両方とも含むか？ (基本周波数は符号化モジュール 8 1 0 の副作用として供給される。)

50

・オーディオ信号は基本周波数が支配的な成分と、基本周波数の高調波間にある成分とを両方とも含むか？

・オーディオ信号は基本周波数より低い大きな信号エネルギーを含むか？

・ポストフィルタリングした復号は、レート歪み最適性の点でフィルタリングしない復号よりも好ましいか（好ましそうか）？

図 8 に示したエンコーダ構成について説明したすべてのバリエーションでは、すなわち基礎となる決定基準にはかかわらずに、決定部 8 2 0 は、スムーズな移行を実現するように、ポストフィルタリングを徐々にかけること又は除くことを決定できる。ポストフィルタのゲインを調性することにより、ポストフィルタリングを徐々にかけたり除いたりできる。

#### 【 0 0 4 5 】

図 9 は、デコーダに供給されるビットストリーム信号に応じて周波数復号モードと C E L P 復号モードで動作できる従来のデコーダを示す。C E L P 復号モードが選択されているときはいつもポストフィルタリングがかけられる。このデコーダの改良を図 1 0 に示す。図 1 0 は本発明の一実施形態によるデコーダ 1 0 0 0 を示す。このデコーダは、周波数領域復号モジュール 1 0 1 3 がアクティブである周波数領域ベースの復号モードと、C E L P 復号モジュール 1 0 1 1 とポストフィルタ 1 0 4 0 がアクティブであるフィルタリングされた C E L P 復号モードとにおいて動作可能であるだけでなく、C E L P モジュール 1 0 1 1 がその信号をバイパスライン 1 0 4 4 を介して補正遅延モジュール 1 0 4 3 に供給する、フィルタリングされていない C E L P モードでも動作可能である。スイッチ 1 0 4 2 は、デコーダ 1 0 0 0 に提供されるビットストリーム信号に含まれるポストフィルタリング情報に応じて、今どの復号モードを使うか制御する。このデコーダと図 9 のデコーダでは、最後の処理ステップは S B R モジュール 1 0 5 0 により行われ、S B R 1 0 5 0 から最終的なオーディオ信号が出力される。

#### 【 0 0 4 6 】

図 1 1 は、デコーダ 1 1 9 9 の下流に配置するのに好適なポストフィルタ 1 1 0 0 を示す。フィルタ 1 1 0 0 はポストフィルタリングモジュール 1 1 4 0 を含む。ポストフィルタリングモジュール 1 1 4 0 は、ポストフィルタ 1 1 0 0 内の決定モジュール 1 1 2 0 から受け取るポストフィルタリング信号に応じて制御モジュール（図示せず）により、特にバイナリ又は非バイナリのゲインコントローラにより、イネーブル又はディスエーブルされる。決定モジュールは、デコーダから得られる信号に 1 つ以上のテストをして、ポストフィルタリングモジュール 1 1 4 0 をアクティブにするか否か決定する。決定は、図 8 の決定モジュール 8 2 0 の機能のある時点で行われる。決定モジュール 8 2 0 は、元の信号及び / 又は中間復号信号を用いて、ポストフィルタの動作を予測する。また、決定モジュール 1 1 2 0 の決定は、決定モジュールが中間復号信号が構成される上記の実施形態で用いるのと同様の情報に基づく。一例として、決定モジュール 1 1 2 0 は、（ビットストリーム信号からピッチ周波数を容易に得られない限り）ピッチ周波数を推定し、そのピッチ周波数より下の信号とその高調波の間の信号のエネルギーコンテンツを計算する。このエネルギーコンテンツが大きければ、ノイズよりも信号成分を表すだろう。これはポストフィルタリングモジュール 1 1 4 0 をディスエーブルする決定をするモチベーションとなる。

#### 【 0 0 4 7 】

6 人の被験者によるリスニングテストを行い、その間に、本発明により符号化及び復号をした音楽サンプルを、符号化された同じ音楽を含む基準サンプルと比較した。ポストフィルタリングは従来通りかけたが、他のパラメータはすべて変更していない。結果として音質が改善されることが確認された。

#### 【 0 0 4 8 】

本発明のさらなる実施形態は、上記の説明を読めば、当業者には明らかになるだろう。本明細書と図面は実施形態と実施例を開示しているが、本発明はこれらの具体的な例に制約されない。添付した特許請求の範囲で規定した本発明の範囲から逸脱することなく、多

10

20

30

40

50

数の修正や変形をすることができる。

【0049】

ここに開示したシステムと方法は、ソフトウェア、ファームウェア、ハードウェアまたはこれらの組み合わせとして実施できる。一部または全部のコンポーネントは、デジタルシグナルプロセッサやマイクロプロセッサにより実行されるソフトウェアとして実施でき、またはハードウェアまたは特定目的集積回路として実施できる。かかるソフトウェアは、コンピュータ読み取り可能媒体で配布可能である。コンピュータ読み取り可能媒体は、コンピュータ記憶媒体（すなわち非一時的媒体）と通信媒体（すなわち一時的媒体）を含む。当業者には周知だが、コンピュータ記憶媒体には、コンピュータ読み取り可能命令、データ構造、プログラムモジュール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や技術で実施された、揮発性および不揮発性、取り外し可能および取り外し不可能媒体を含む。コンピュータ記憶媒体は、RAM、ROM、EEPROM、フラッシュメモリその他のメモリ技術、CD-ROM、デジタルバーサタイルディスク（DVD）その他の光ディスク記憶媒体、磁気カセット、磁気テープ、磁気ディスク記憶その他の磁気記憶デバイス、またはその他の、所望の情報の記憶に使える任意の媒体を含むが、これらに限定されない。さらに、当業者には周知であるように、通信媒体は、一般的に、コンピュータ読み取り可能命令、データ構造、プログラムモジュール、その他の搬送波その他の伝送メカニズムなどの変調データ信号中のデータを化体し、任意の情報配信媒体を含む。

10

【0050】

実施形態のリスト

20

1. オーディオ時間信号としてビットストリーム信号を復号するデコーダシステム（400；500；700；1000）であって：  
ビットストリーム信号を予備的オーディオ時間信号として復号する復号部（410；511，512，513；711，712，713；1011，1013）と；  
前記予備的オーディオ時間信号をフィルタリングしてオーディオ時間信号を取得する高調波間ノイズ減衰ポストフィルタ（440；540；740；1040）とを有し、  
前記ビットストリーム信号にエンコードされたポストフィルタリング情報に応じて前記ポストフィルタをディスエーブルするように構成された制御部を有し、  
前記予備的オーディオ時間信号が前記オーディオ時間信号として出力されることを特徴とする、デコーダシステム。

30

【0051】

2. 前記ポストフィルタはさらにスペクトルの谷にあるノイズを減衰するように構成されている、実施形態1に記載のデコーダシステム。

【0052】

3. 前記制御部は、前記デコーダシステムの信号処理経路から前記ポストフィルタを選択的に除外するスイッチ（442；541；1042）を含み、これにより前記ポストフィルタはディスエーブルされる、実施形態1に記載のデコーダシステム。

【0053】

4. 前記ポストフィルタは前記高調波間減衰を決定する可変ゲインを有し、前記制御部は前記ゲインの絶対値を所定閾値より下に設定するように動作可能なゲインコントローラを含む、実施形態1に記載のデコーダシステム。

40

【0054】

5. 前記復号部はスピーチ復号モジュールを含む、実施形態1に記載のデコーダシステム。

【0055】

6. 前記復号部は、コード励起線形予測、CELP、復号モジュール（511；711；1011）を含む、実施形態1に記載のデコーダシステム。

【0056】

7. 前記エンコーダの長期予測部により推定されたピッチ周波数が前記ビットストリーム信号にエンコードされる、実施形態5に記載のデコーダシステム。

50

## 【 0 0 5 7 】

8. 前記ポストフィルタは、前記ピッチ周波数の高調波間にあるスペクトル成分を減衰するように構成されている、実施形態 7 に記載のデコーダシステム。

## 【 0 0 5 8 】

9. 前記ビットストリーム信号はピッチ周波数の表示を含み、前記ポストフィルタは前記ピッチ周波数の高調波間にあるスペクトル成分を減衰するように構成されている、実施形態 1 に記載のデコーダシステム。

## 【 0 0 5 9 】

10. 前記ポストフィルタは、所定のカットオフ周波数より下にあるスペクトル成分のみを減衰するように構成されている、実施形態 8 または 9 に記載のデコーダシステム。

10

## 【 0 0 6 0 】

11. 前記復号部は、ビットストリーム信号をオーディオ時間信号として復号する T C X (transform-coded excitation) 復号モジュール (5 1 2 ; 7 1 2) をさらに有し、  
前記制御部は少なくとも次のモード：

a) 前記 T C X モジュールがイネーブルされ、前記ポストフィルタがディスエーブルされる；

b) 前記 C E L P モジュールと前記ポストフィルタがイネーブルされている；及び

c) 前記 C E L P モジュールがイネーブルされ、前記ポストフィルタがディスエーブルされており、前記呼びオーディオ時間信号と前記オーディオ時間信号が同時である、  
で前記デコーダシステムを動作させるように構成されている、実施形態 6 に記載のデコーダシステム。

20

## 【 0 0 6 1 】

12. 前記復号部は、ビットストリーム信号をオーディオ時間信号として復号する A A C (Advanced Audio Coding) 復号モジュール (5 1 3 ; 7 1 3) をさらに有し、  
前記制御部は少なくとも次のモード：

d) 前記 A A C モジュールがイネーブルされ、前記ポストフィルタがディスエーブルされる

で前記デコーダシステムを動作させるように構成されている、実施形態 10 に記載のデコーダシステム。

30

## 【 0 0 6 2 】

13. 前記ビットストリーム信号は時間フレームに分割され、前記制御部は一時間フレーム全体又は一連の時間フレームをディスエーブルするように構成されている、  
実施形態 1 に記載のデコーダシステム。

14. 前記制御部は、さらに M P E G (Moving Pictures Experts Group) ビットストリーム中の各時間フレームに対して、この時間フレームに関連するデータフィールドを受け取るように構成され、前記データフィールドの値に応じて、前記ポストフィルタをディスエーブルするように動作可能である、実施形態 13 に記載のデコーダシステム。

## 【 0 0 6 3 】

15. 前記制御部は前記ポストフィルタのゲインを徐々に下げる及び / 又は上げるように構成されている、実施形態 4 に記載のデコーダシステム。

40

## 【 0 0 6 4 】

16. デコーダシステム (4 0 0 ; 5 0 0 ; 7 0 0 ; 1 0 0 0) であって：  
ビットストリーム信号を予備的オーディオ時間信号として復号する復号部 (4 1 0 ; 5 1 1 , 5 1 2 , 5 1 3 ; 7 1 1 , 7 1 2 , 7 1 3 ; 1 0 1 1 , 1 0 1 3) と；  
前記予備的オーディオ時間信号をフィルタリングしてオーディオ時間信号を取得する高調波間ノイズ減衰ポストフィルタ (4 4 0 ; 5 4 0 ; 7 4 0 ; 1 0 4 0) とを有し、  
前記復号部は励起を表す中間復号信号を生成し、それを前記制御部に提供するように構成されており；

前記制御部は、前記中間復号信号と前記ポストフィルタリングをかけ、次の基準：

50

- a) 前記近似差分信号のパワーが所定閾値を越えている；
  - b) 前記近似差分信号の特徴が音調的である；
  - c) 前記近似差分信号の強さ・周波数スペクトルと前記オーディオ信号の強さ・周波数スペクトルとの間の差分が、周波数に対して不均一に分布している；
  - d) 前記近似差分信号の強度周波数スペクトルは所定の関連エンベロープ内の周波数区間にあるか、
  - e) 前記近似差分信号の強度周波数スペクトルが、所定のスケールファクタによりダウンスケールされた最大信号成分の強度により前記オーディオ時間信号の強度周波数スペクトルに閾値を設けることにより得られた関連エンベロープ内の周波数区間にあるか；
- のうち少なくとも一を評価した前記中間復号信号との間の差分として、前記ポストフィルタにより前記復号信号から除去される信号成分を近似する近似差分信号を計算するよう構成され、及び、
- 肯定的な決定に応じて、前記ポストフィルタをディスエーブルし、それにより前記予備的オーディオ時間信号が前記オーディオ時間信号として出力される、デコーダシステム。

10

## 【0065】

17. 予備的オーディオ信号を含む入力信号を受け取り、出力オーディオ信号を供給するように構成された高調波間ノイズ減衰ポストフィルタ(440；550；740；1040；1140)であって、

ポストフィルタリング信号の値により、次のモード：

- i) 前記予備的オーディオ信号をフィルタリングしてフィルタリングされた信号を求め、これを出力オーディオ信号として供給するフィルタリングモードと、
- ii) 前記予備的オーディオ信号を出力オーディオ信号として供給するパススルーモード

20

、  
 の一で前記ポストフィルタを選択的に動作させる制御部を有することを特徴とする、ポストフィルタ。

## 【0066】

18. 前記ポストフィルタリング信号は前記入力信号に含まれる、  
 実施形態17に記載のポストフィルタ。

## 【0067】

19. 前記予備的オーディオ信号のピッチ周波数を推定し、次の基準：

30

- a) 前記ピッチ周波数より下のスペクトル成分のパワーが所定閾値を越えているか；
- b) 前記ピッチ周波数より下のスペクトル成分は音調的であるか；
- c) 前記ピッチ周波数の高調波間のスペクトル成分のパワーが所定閾値を越えているか；
- d) 前記ピッチ周波数の高調波間のスペクトル成分は音調的であるか；

のうち少なくとも一を評価するように構成された決定モジュール(1120)を更に有する、及び、

肯定的な決定に応じて、前記ポストフィルタをディスエーブルする負のポストフィルタリング信号を生成する決定をする、

実施形態17に記載のポストフィルタ。

## 【0068】

40

20. ビットストリーム信号をオーディオ時間信号として復号する方法であって、ビットストリーム信号を予備的オーディオ時間信号として復号するステップと；高調波間ノイズを減衰することにより前記予備的オーディオ時間信号をポストフィルタリングして、それによりオーディオ時間信号を求めるステップとを有し、前記ポストフィルタリングするステップは、前記ビットストリーム信号にエンコードされたポストフィルタリング情報に応じて選択的に省略されることを特徴とする、方法。

## 【0069】

21. 前記ポストフィルタリングするステップは、さらにスペクトルの谷にあるノイズを減衰するステップを含む、実施形態20に記載の方法。

## 【0070】

50

22. 前記復号するステップは、スピーチコーディング用のコーディング方法を適用するステップを含む、実施形態20に記載の方法。

【0071】

23. 前記復号するステップは、CELP (code-excited linear prediction) 復号を適用するステップを含む、実施形態20に記載の方法。

【0072】

24. 前記ポストフィルタリングするステップは、前記ピッチ周波数の高調波間にあるスペクトル成分を減衰するステップを含み、前記ピッチ周波数は前記ビットストリーム信号から取り出されるか、又は前記復号ステップで推定される、実施形態22又は23に記載の方法。

10

【0073】

25. 前記ポストフィルタリングするステップは、所定のカットオフ周波数より下にあるスペクトル成分のみを減衰するステップを含む、実施形態20に記載の方法。

【0074】

26. 前記復号するステップとポストフィルタリングするステップは、次の：

- a) TCX復号；
- b) ポストフィルタリング付きのCELP復号と；及び
- c) ポストフィルタリング無しのCELP復号；

のうちの一を選択的に実行する、実施形態23に記載の方法。

20

【0075】

27. 前記復号するステップとポストフィルタリングするステップは、モードa)、b)、c)、及びd) AAC (Advanced Audio Coding) 復号のうちの一を選択的に実行する、実施形態26に記載の方法。

【0076】

28. 前記ビットストリーム信号は時間フレームに分割され、前記ポストフィルタリングするステップは一時間フレーム全体又は一連の時間フレームの間省略される、実施形態20に記載の方法。

【0077】

30

29. 前記ビットストリーム信号はMP EG (Moving Pictures Experts Group) ビットストリームであり、各時間フレームに対して、関連するデータフィールドを含み、前記ポストフィルタリングするステップは、前記関連するデータフィールドの値に応じて時間フレームにおいて省略される、実施形態28に記載の方法。

【0078】

30. 前記ポストフィルタリングの省略は、次の：

- 減衰の完全省略、
  - 減衰の部分的省略、
  - 減衰を徐々に大きくすること、及び
  - 減衰を徐々に小さくすること、
- うちの一を含む、実施形態20に記載の方法。

40

【0079】

31. ビットストリーム信号をオーディオ時間信号として復号する方法であって、ビットストリーム信号を予備的オーディオ時間信号として復号するステップと；高調波間ノイズを減衰することにより前記予備的オーディオ時間信号をポストフィルタリングして、それによりオーディオ時間信号を求めるステップとを有し、前記復号するステップは、励起を表す中間復号信号を取り出すステップと；前記ポストフィルタにより前記復号信号から除去される信号成分を近似する近似的差分信号を、前記中間復号信号と、ポストフィルタリングをかけられた中間復号信号との差分と

50

して計算するステップと；

次の基準：

- a) 前記近似差分信号のパワーが所定閾値を越えている；
- b) 前記近似差分信号の特徴が音調的である；
- c) 前記近似差分信号の強さ・周波数スペクトルと前記オーディオ信号の強さ・周波数スペクトルとの間の差分が、周波数に対して不均一に分布している；
- d) 前記近似差分信号の強度周波数スペクトルは所定の関連エンベロープ内の周波数区間にあるか、

e) 前記近似差分信号の強度周波数スペクトルが、所定のスケールファクタによりダウンスケールされた最大信号成分の強度により前記オーディオ時間信号の強度周波数スペクトルに閾値を設けることにより得られた関連エンベロープ内の周波数区間にあるか；

のうち少なくとも一を評価するステップと；及び、

肯定的な決定に応じて、前記ポストフィルタをディスエーブルし、それにより前記予備的オーディオ時間信号が前記オーディオ時間信号として出力される、エンコーダシステム。

【0080】

32. オーディオ時間信号をビットストリーム信号として符号化するエンコーダシステム(800)であって、

オーディオ時間信号をビットストリーム信号として符号化する符号化部(810)を含み、

前記ビットストリームの復号において、高調波間ノイズの減衰を含むポストフィルタリングをディスエーブルするか決定し、この決定をポストフィルタリング情報として前記ビットストリーム信号にエンコードするように構成された決定部(820)を有することを特徴とする、エンコーダシステム。

【0081】

33. 前記決定部は、スペクトルの谷にあるノイズの減衰をさらに含むポストフィルタリングをディスエーブルするか決定するように構成されている、実施形態32に記載のエンコーダシステム。

【0082】

34. 前記符号化部はスピーチコーディングモジュールを含む、実施形態32に記載のエンコーダシステム。

【0083】

35. 前記符号化部は、C E L P (code-excited linear prediction) 符号化モジュールを含む、実施形態32に記載のエンコーダシステム。

【0084】

36. 前記決定部は：

支配的基本周波数を有する信号成分と、前記基本周波数より下にあり、及び任意的にその高調波の間にある信号成分とが共に存在することを検出し、

それに応じて、ディスエーブルする決定を行う、

実施形態32に記載のエンコーダシステム。

【0085】

37. 前記C E L P 符号化モジュールは前記オーディオ時間信号中のピッチ周波数を推定するように構成され、

前記決定部は、前記推定されたピッチ周波数より下にあるスペクトル成分を検出し、それに応じて、ディスエーブルする決定をするように構成されている、

実施形態35に記載のエンコーダシステム。

【0086】

38. 前記決定部は：

C E L P 符号化された時の前記オーディオ時間信号の予測パワーと、C E L P 符号化されたポストフィルタリングされた時の前記オーディオ時間信号の予測パワーとの間の差分を計算し、

10

20

30

40

50

この差分が所定の閾値を越えるのに応じて、ディスエーブルする決定をする、実施形態 35 に記載のエンコーダシステム。

【0087】

39. 前記符号化部はさらに T C X (transform-coded excitation) 符号化モジュールを含み、

前記決定部は、次の符号化モード：

- a) T C X 符号化；
- b) ポストフィルタリング付きの C E L P コーディングと；及び
- c) ポストフィルタリング無しの C E L P コーディング；

のうちの一を選択するように構成されている、実施形態 35 に記載のエンコーダシステム。

10

【0088】

40. 次のスーパーモード：

i) 前記決定部がディスエーブルされる A A C (Advanced Audio Coding) コーディングと；

i i) 前記決定部はコーディングモード a)、b) 及び c) のうちの一を選択できる T C X / C E L P コーディング；

のうちの一を選択するように構成されたコーディングセレクタ (814) を更に有する、実施形態 39 に記載のエンコーダシステム。

20

【0089】

41. 前記決定部は、レート・歪み最適化に基づきどのモードを用いるか決定するように構成されている、実施形態 39 に記載のエンコーダシステム。

【0090】

42. 前記ビットストリーム信号を時間フレームに分割するように構成され、前記決定部は全フレームよりなる時間セグメントにおいて前記ポストフィルタをディスエーブルする決定をするように構成されている、実施形態 32 に記載のエンコーダシステム。

【0091】

43. 前記決定部は前記ポストフィルタの減衰を徐々に下げる及び / 又は上げることを決定するように構成されている、実施形態 32 に記載のエンコーダシステム。

【0092】

44. 前記決定部は：

推定されたピッチ周波数より下のオーディオ時間信号のパワーを計算し、このパワーが所定の閾値を越えるのに応じて、ディスエーブルする決定をする、実施形態 32 に記載のエンコーダシステム。

30

【0093】

45. 前記決定部は：

前記オーディオ時間信号から、前記ポストフィルタによりさらなる復号信号から除去される信号成分を近似する近似差分信号を求め、

次の基準：

- a) 前記近似差分信号のパワーが所定閾値を越えている；
- b) 前記近似差分信号の特徴が音調的である；
- c) 前記近似差分信号の強さ・周波数スペクトルと前記オーディオ信号の強さ・周波数スペクトルとの間の差分が、周波数に対して不均一に分布している；
- d) 前記近似差分信号の強度周波数スペクトルは所定の関連エンベロープ内の周波数区間にあるか、
- e) 前記近似差分信号の強度周波数スペクトルが、所定のスケールファクタによりダウンスケールされた最大信号成分の強度により前記オーディオ時間信号の強度周波数スペクトルに閾値を設けることにより得られた関連エンベロープ内の周波数区間にあるか；

40

のうち少なくとも一を評価する；及び、

肯定的な決定に応じて、前記ポストフィルタをディスエーブルする決定をする、

50

実施形態 3 2 に記載のエンコーダシステム。

【 0 0 9 4 】

4 6 . 前記決定部は、前記近似差分信号を、前記オーディオ時間信号と、ポストフィルタリングをかけたオーディオ時間信号との間の差分として、計算するように構成されている、

実施形態 4 5 に記載のエンコーダシステム。

【 0 0 9 5 】

4 7 . 前記符号化部は励起を表す中間復号信号を取り出し、それを前記決定部に提供するように構成されており、

前記決定部は、前記近似差分信号を、前記オーディオ時間信号と、ポストフィルタリングをかけた中間復号信号との間の差分として、計算するように構成されている、

実施形態 4 5 に記載のエンコーダシステム。

【 0 0 9 6 】

4 8 . オーディオ時間信号をビットストリーム信号として符号化する方法であって、オーディオ時間信号をビットストリーム信号として符号化するステップを含み、前記ビットストリームの復号において、高調波間ノイズの減衰を含むポストフィルタリングをディスエーブルするか決定し、この決定をポストフィルタリング情報として前記ビットストリーム信号にエンコードするステップをさらに有することを特徴とする、方法。

【 0 0 9 7 】

4 9 . 前記決定するステップは、さらにスペクトルの谷にあるノイズを減衰するステップを含むポストフィルタリングに関する、実施形態 4 8 に記載の方法。

【 0 0 9 8 】

5 0 . 前記符号化するステップは、スピーチコーディング用のコーディング方法を適用するステップを含む、実施形態 4 8 に記載の方法。

【 0 0 9 9 】

5 1 . 前記符号化するステップは、C E L P (code-excited linear prediction) コーディングを適用するステップを含む、実施形態 4 8 に記載の方法。

【 0 1 0 0 】

5 2 . 支配的基本周波数を有する信号成分と、前記基本周波数より下にあり、及び任意的にその高調波の間にある信号成分とが共に存在することを検出するステップをさらに有し、

肯定的決定の場合に、ポストフィルタリングをディスエーブルする決定が行われる、実施形態 4 8 に記載の方法。

【 0 1 0 1 】

5 3 . 前記 C E L P コーディングするステップは、前記オーディオ時間信号中のピッチ周波数を推定するステップを含み、決定するステップは、前記推定されたピッチ周波数より下にあるスペクトル成分を検出するステップを含み、肯定的な検出結果の場合に、ポストフィルタリングをディスエーブルする決定が行われる、

実施形態 5 1 に記載の方法。

【 0 1 0 2 】

5 4 . C E L P 符号化された時の前記オーディオ時間信号の予測パワーと、C E L P 符号化されポストフィルタリングされた時の前記オーディオ時間信号の予測パワーとの間の差分を計算するステップをさらに含み、この差分が所定の閾値を越えた場合、ポストフィルタリングをディスエーブルする決定が行われる、実施形態 5 1 に記載の方法。

【 0 1 0 3 】

5 5 . 符号化するステップは、C E L P コーディング又は T C X コーディングのいずれかを選択的に適用するステップを含み、

ポストフィルタリングをディスエーブルするか決定するステップは、C E L P コーディン

10

20

30

40

50

グが適用されている時にのみ行われる、  
実施形態 5 1 に記載の方法。

5 6. 前記決定するステップは、レート・歪み最適化に基づき、次の動作モード：

- a) T C X コーディング；
- b) ポストフィルタリング付きの C E L P コーディングと；及び
- c) ポストフィルタリング無しの C E L P コーディング；

のうちの一を選択するステップを含む、実施形態 5 5 に記載の方法。

【 0 1 0 4 】

5 7. 前記決定するステップは、レート・歪み最適化に基づき、次の動作モード：

- a) T C X コーディング；
- b) ポストフィルタリング付きの C E L P コーディングと；及び
- c) ポストフィルタリング無しの C E L P コーディング；及び
- d) A A C (Advanced Audio Coding) コーディング

のうちの一を選択するステップを含む、実施形態 5 5 に記載の方法。

【 0 1 0 5 】

5 8. 符号化するステップは、前記オーディオ時間信号を時間フレームに分割し、対応する時間フレームを有するビットストリーム信号を構成するステップを含み、  
ポストフィルタリングをディスエーブルすることを決定するステップは、時間フレームごと  
に一度行われる、

実施形態 4 8 に記載の方法。

【 0 1 0 6 】

5 9.

ポストフィルタリングをディスエーブルすることを決定するステップの結果は、次から：  
減衰無し、

全減衰、

部分的減衰、

減衰を徐々に大きくすること、及び

減衰を徐々に小さくすること、

選択される、実施形態 4 8 に記載の方法。

【 0 1 0 7 】

6 0. 決定するステップは、推定されたピッチ周波数より下のオーディオ時間信号の  
パワーを計算し、そのパワーが所定の閾値を越えるのに応じて、前記ポストフィルタをデ  
ィスエーブルするステップを含む、

実施形態 4 8 に記載の方法。

【 0 1 0 8 】

6 1. 符号化するステップは、前記オーディオ時間信号から、前記ポストフィルタに  
よりさらなる復号信号から除去される信号成分を近似する近似差分信号を求めるステップ  
を含む、

決定するステップは、次の基準：

- a) 前記近似差分信号のパワーが所定閾値を越えている；
- b) 前記近似差分信号の特徴が音調的である；
- c) 前記近似差分信号の強さ・周波数スペクトルと前記オーディオ信号の強さ・周波数ス  
ペクトルとの間の差分が、周波数に対して不均一に分布している；
- d) 前記近似差分信号の強度周波数スペクトルは所定の関連エンベロープ内の周波数区  
間にあるか、
- e) 前記近似差分信号の強度周波数スペクトルが、所定のスケールファクタによりダウ  
ンスケールされた最大信号成分の強度により前記オーディオ時間信号の強度周波数ス  
ペクトルに閾値を設けることにより得られた関連エンベロープ内の周波数区間にあるか；

のうち少なくとも一を評価するステップを含む、及び、

少なくとも肯定的な決定に応じて、前記ポストフィルタをディスエーブルするステップ

10

20

30

40

50

実施形態 4 8 に記載の方法。

【 0 1 0 9 】

6 2 . 前記近似差分信号を、前記オーディオ時間信号と、ポストフィルタリングをかけたオーディオ時間信号との間の差分として、計算する、  
実施形態 6 1 に記載の方法。

【 0 1 1 0 】

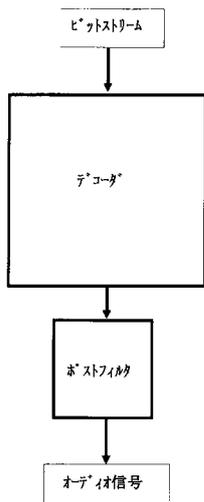
6 3 . 符号化するステップは、励起を表す中間復号信号を取り出すステップを含み、決定するステップは、前記近似差分信号を、前記オーディオ時間信号と、ポストフィルタリングをかけた中間復号信号との間の差分として、計算するステップを含む、  
実施形態 6 1 に記載の方法。

10

【 0 1 1 1 】

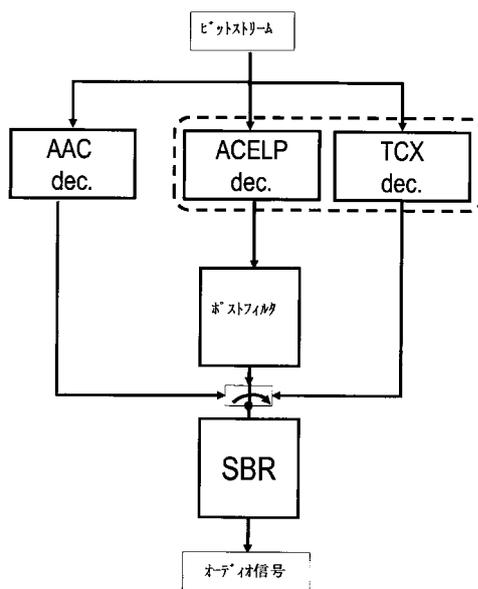
6 4 . 実施形態 2 0 乃至 3 1、及び 4 8 乃至 6 3 いずれか一項に記載の方法を実行する命令を記憶したデータ担体を含むコンピュータ読み取り可能媒体。

【 図 1 】



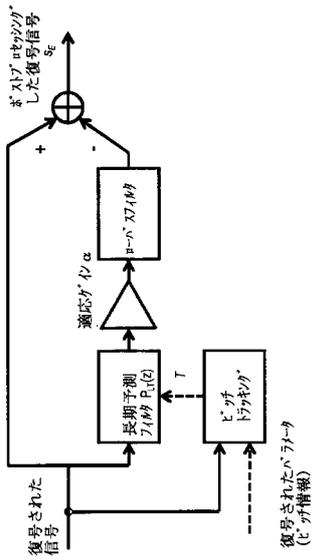
(従来技術)

【 図 2 】

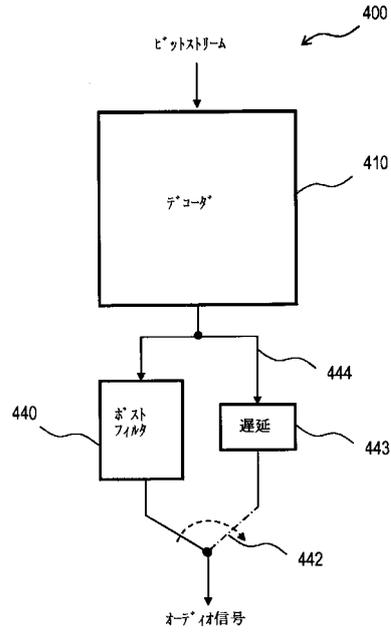


(従来技術)

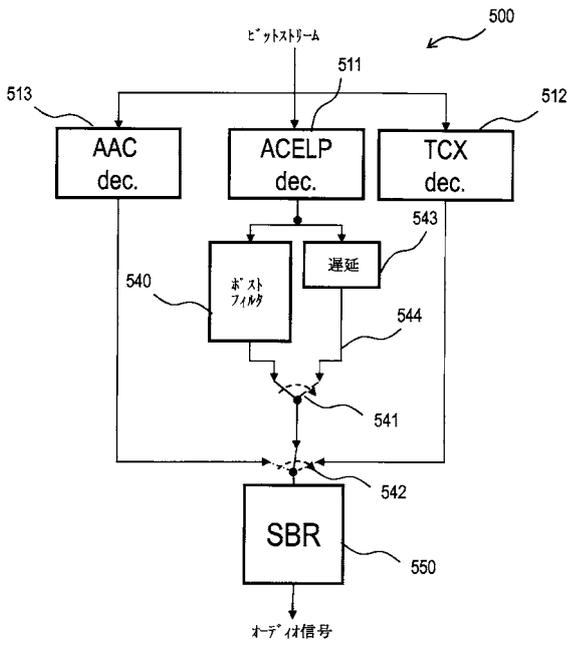
【 図 3 】



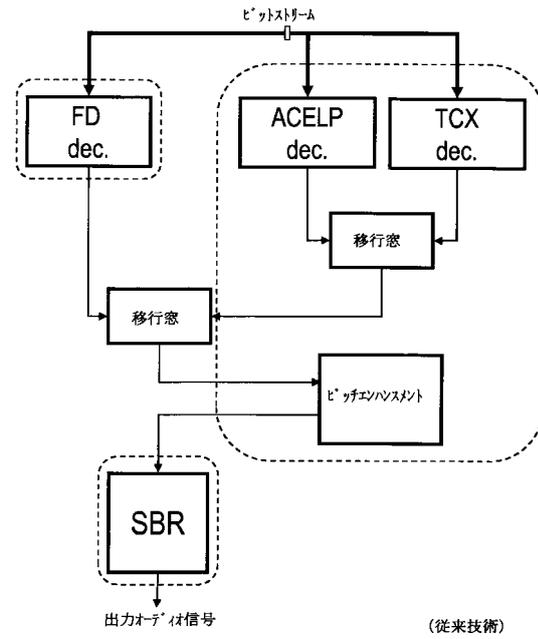
【 図 4 】



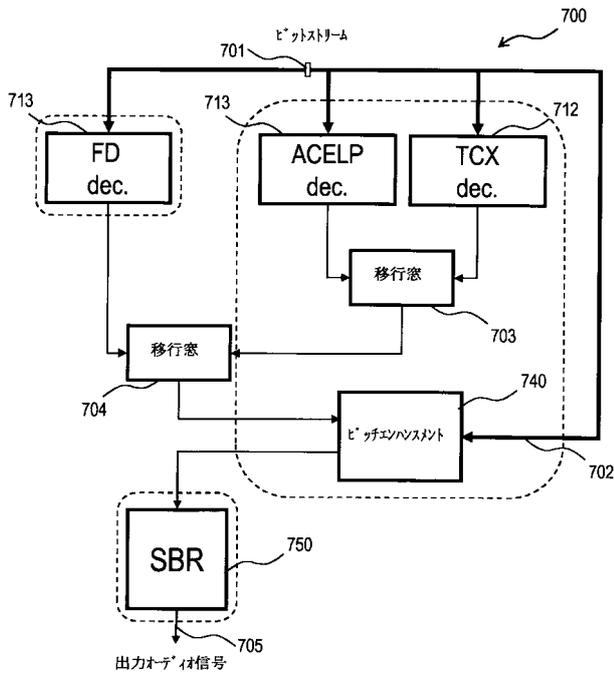
【 図 5 】



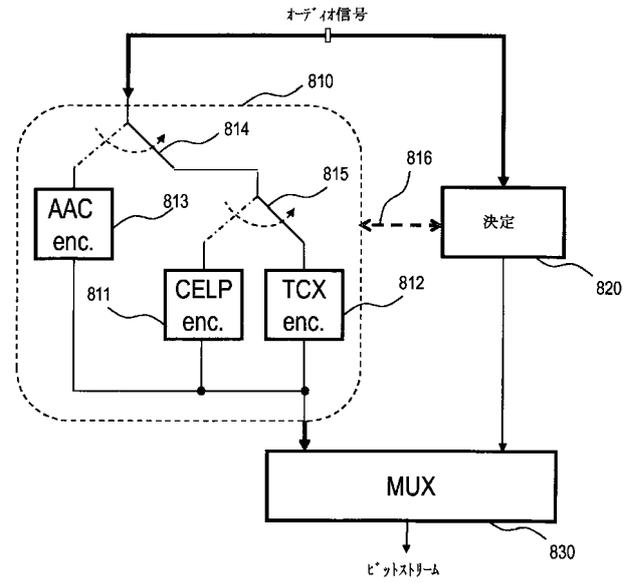
【 図 6 】



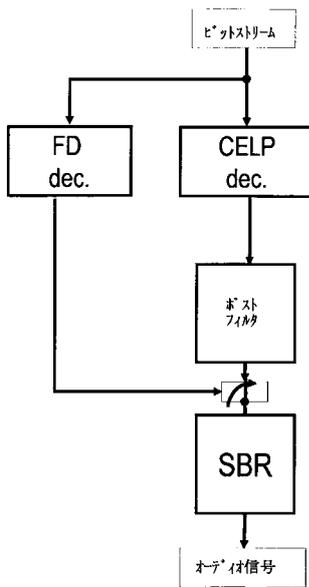
【図7】



【図8】

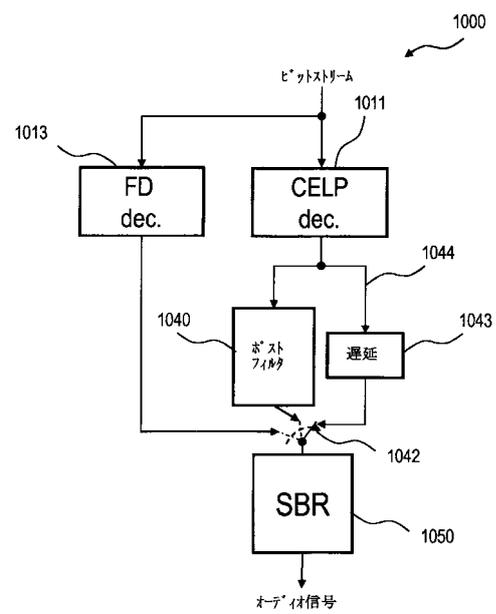


【図9】

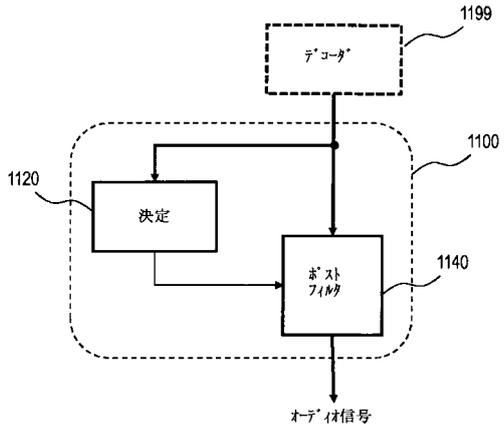


(従来技術)

【図10】



【 図 1 1 】



---

フロントページの続き

- (72)発明者 レスク,バルバラ  
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴルガタン 1 2アー, ドルビー スウェーデン アーペー内
- (72)発明者 クヨエルリン, クリストフェル  
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴルガタン 1 2アー, ドルビー スウェーデン アーペー内
- (72)発明者 ヴィレモーズ, ラーシュ  
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴレガタン 1 2アー, ドルビー スウェーデン アーペー内