(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5097219号
(P5097219)
(45)発行日　平成24年12月12日(2012.12.12)　(24)登録日　平成24年9月28日(2012.9.28)

(51)Int.Cl.　F I
G10L 19/14 (2006.01)　G10L 19/14 520B

請求項の数 19 (全 19 頁)

(21)出願番号	特願2009-551967 (P2009-551967)
(86)(22)出願日	平成19年12月14日 (2007.12.14)
(65)公表番号	特表2010-520505 (P2010-520505A)
(43)公表日	平成22年6月10日 (2010.6.10)
(86)国際出願番号	PCT/SE2007/051000
(87)国際公開番号	WO2008/108702
(87)国際公開日	平成20年9月12日 (2008.9.12)
審査請求日	平成22年11月12日 (2010.11.12)
(31)優先権主張番号	60/892,667
(32)優先日	平成19年3月2日 (2007.3.2)
(33)優先権主張国	米国 (US)

(73)特許権者　598036300
テレフオンアクチーボラゲット　エル　エム　エリクソン (パブル)
スウェーデン国　ストックホルム　エス－１６４　８３
(74)代理人　100076428
弁理士　大塚　康徳
(74)代理人　100112508
弁理士　高柳　司郎
(74)代理人　100115071
弁理士　大塚　康弘
(74)代理人　100116894
弁理士　木村　秀二
(74)代理人　100130409
弁理士　下山　治

最終頁に続く

(54)【発明の名称】非因果性ポストフィルタ

(57)【特許請求の範囲】
【請求項１】
フレームごとに符号化信号のパラメータ（４）を入力する受信機入力部（４０）と、
前記受信機入力部（４０）に接続され、前記パラメータに基づいて復号化オーディオ信号（５；５４）のフレームを出力するデコーダ（２０）と、
前記デコーダ（２０）の出力に接続され、前記復号化オーディオ信号（５；５４）のフレームに基づいて出力信号（６）を出力するポストフィルタ（３０；３０Ａ，３０Ｂ）と、
前記出力信号（６）の出力部（６０）と、
を有するデコーダ装置であって、
前記受信機入力部（４０）及び前記デコーダ（２０）の少なくともいずれか一方は、第１のフレームのパラメータが前記受信機入力部（４０）で利用可能になる時と前記第１のフレームの復号化オーディオ信号が前記デコーダ（２０）の出力で利用可能になる時との間に少なくとも１つのフレームに対応する時間差が生じるように構成され、
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、前記受信機入力部（４０）に接続され、
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、各後続フレームの前記パラメータ（４）に応じて前記復号化オーディオ信号（５；５４）のフレームをフィルタリングして前記出力信号（６）を得るように構成される
ことを特徴とするデコーダ装置。

【請求項２】
前記受信機入力部（４０）は、連続する少なくとも２つのフレームのパラメータを記憶する記憶部（４１）を含み、前記デコーダ（２０）は、第１のフレームのパラメータ（４Ａ）を前記記憶部（４１）から受信し、前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、後続する第２のフレームのパラメータ（４Ｂ）にアクセスすることを特徴とする請求項１に記載のデコーダ装置。
【請求項３】
前記デコーダ（２０）は、前記ポストフィルタ（３０；３０Ａ，３０Ｂ）に出力する前に前記復号化オーディオ信号のフレームを遅延させる手段（５１：５３）を含むことを特徴とする請求項１に記載のデコーダ装置。
【請求項４】
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、ピッチポストフィルタを含み、前記ピッチポストフィルタにおいて使用されるピッチパラメータは、前記後続フレームのピッチパラメータに基づくものであることを特徴とする請求項１乃至３のいずれか１項に記載のデコーダ装置。
【請求項５】
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）の前記ピッチポストフィルタは、後続のサブフレームごとに、当該後続のサブフレームのピッチ値を差し引いた時間インデックスの値を求め、前記求めた値が現在の時間インデックス以上である場合、前記後続のサブフレームの前記ピッチ値を、現在のフレームのピッチ進みパラメータとして利用することを特徴とする請求項４に記載のデコーダ装置。
【請求項６】
出力が前記ポストフィルタ（３０；３０Ａ，３０Ｂ）に接続されるオーディオ特性検出器を更に有し、
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、前記ピッチポストフィルタにおいて使用される前記ピッチパラメータが、現在のフレームと先行フレーム及び後続フレームの少なくともいずれか一方との間のオーディオ特性の変化の有意性の基準に依存して、前記後続フレームの前記ピッチパラメータに基づいて適応化されるように構成されることを特徴とする請求項４又は５記載のデコーダ装置。
【請求項７】
前記オーディオ特性検出器は、音声アクティビティ検出器及び有声音検出器のうちの少なくともいずれか一方であり、前記ポストフィルタは、有声音の開始が検出された場合に前記ピッチポストフィルタにおいて使用されるピッチパラメータを前記後続フレームのピッチパラメータに基づくものとするように構成されることを特徴とする請求項６に記載のデコーダ装置。
【請求項８】
前記ポストフィルタ（３０；３０Ａ，３０Ｂ）は、前記後続フレームの復号化信号にもアクセスするように構成されることを特徴とする請求項１乃至７のいずれか１項に記載のデコーダ装置。
【請求項９】
前記デコーダ（２０）は、スケーラブルデコーダ（１２０）又はスケーラブルデコーダの一部であり、前記スケーラブルデコーダの２次デコーダ（２５）は、前記スケーラブルデコーダの１次デコーダ（２１）よりも高遅延であることを特徴とする請求項１乃至８のいずれか１項に記載のデコーダ装置。
【請求項１０】
前記デコーダ（２０）はスケーラブルデコーダ（１２０）であって、該スケーラブルデコーダは、
前記受信機入力部（４０）に接続され、前記パラメータ（４）に基づいて１次復号化信号（２３）を出力する１次デコーダ（２１）と、
前記受信機入力部（４０）に接続され、前記パラメータ（４）に基づいて前記１次復号

化信号（２３）とは異なる２次復号化信号（２６）を出力する２次デコーダ（２５）と、
を含むことを特徴とする請求項１乃至８のいずれか１項に記載のデコーダ装置。
【請求項１１】
フレームごとに符号化信号のパラメータを受信する受信ステップ（２１０）と、
前記パラメータを復号化して復号化オーディオ信号を得る復号化ステップ（２１２）と、を有し、
前記受信ステップ及び前記復号化ステップの少なくともいずれか一方は、第１のフレームのパラメータが受信後に利用可能になった時と前記第１のフレームの復号化オーディオ信号が復号化後に利用可能になる時との間に少なくとも１つのフレームに対応する時間差を生じさせ、
更に、
各後続フレームの前記パラメータに応じて前記復号化オーディオ信号のフレームをポストフィルタリングして出力信号を得るポストフィルタリングステップ（２１４）と、
前記出力信号を出力する出力ステップ（２１６）と、
を有することを特徴とする復号化方法。
【請求項１２】
各時点において連続する少なくとも２つのフレームのパラメータを記憶する記憶ステップを更に有し、前記復号化ステップは、第１のフレームのパラメータを使用して実行され、前記ポストフィルタリングは後続する第２のフレームのパラメータにアクセスして実行されることを特徴とする請求項１１に記載の復号化方法。
【請求項１３】
前記ポストフィルタリングステップを実行する前に前記復号化オーディオ信号のフレームを遅延させるステップを更に有することを特徴とする請求項１１に記載の復号化方法。
【請求項１４】
前記ポストフィルタリングステップ（２１４）はピッチポストフィルタリングを行うステップを含み、前記ピッチポストフィルタリングにおいて使用されるピッチパラメータは、前記後続フレームのピッチパラメータに基づくものであることを特徴とする請求項１１乃至１３のいずれか１項に記載の復号化方法。
【請求項１５】
前記ポストフィルタリングステップ（２１４）における前記ピッチポストフィルタリングは、
後続のサブフレームごとに、当該後続のサブフレームのピッチ値を差し引いた時間インデックスの値を求めるステップ（２２４）と、
前記求めた値が現在の時間インデックス以上である場合、前記後続のサブフレームの前記ピッチ値を、現在のフレームのピッチ進みパラメータとして利用するステップ（２２６）と、
を含むことを特徴とする請求項１４に記載の復号化方法。
【請求項１６】
前記フレームごとの符号化信号のオーディオ特性を検出する検出ステップを更に有し、
前記ポストフィルタリングステップは、前記ピッチパラメータを、現在のフレームと先行フレーム及び後続フレームの少なくともいずれか一方との間のオーディオ特性の変化の有意性の基準に依存して、前記後続フレームの前記ピッチパラメータに基づいて適応化させる
ことを特徴とする請求項１４又は１５に記載の復号化方法。
【請求項１７】
前記検出ステップは、音声アクティビティ及び有声音の少なくともいずれか一方を検出するステップを含み、前記ポストフィルタリングステップは、有声音の開始が検出された場合にのみ、前記ピッチパラメータを前記後続フレームのピッチパラメータに基づくものとすることを特徴とする請求項１６に記載の復号化方法。
【請求項１８】

前記ポストフィルタリングステップ（２１４）は、各後続フレームの復号化信号にも応じて実行されることを特徴とする請求項１１乃至１７のいずれか１項に記載の復号化方法。

【請求項１９】

前記復号化ステップ（２１２）は、スケーラブルデコーダにおいて復号化を行うステップであり、前記スケーラブルデコーダの２次復号化は、前記スケーラブルデコーダの１次復号化よりも高遅延であることを特徴とする請求項１１乃至１８のいずれか１項に記載の復号化方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、一般に、オーディオ及び／又は音声信号の符号化及び復号化に関し、特に、符号化ノイズを低減することに関する。

【背景技術】

【０００２】

一般に、オーディオ符号化（audio coding）、特に音声符号化（speech coding）は、符号化領域においてアナログ入力オーディオ信号又は音声信号をデジタル表現にマッピングし、再びアナログ出力オーディオ信号又は音声信号に戻す。デジタル表現は、オーディオ又は音声を表す値又はパラメータの量子化又は離散化に必要である。量子化又は離散化は、符号化ノイズにより本来の値又はパラメータを乱すと考えられる。オーディオ符号化又は音声符号化の技術は、所与のビットレートの復号化音声における符号化ノイズの影響が可能な限り小さくなるように符号化を行う。しかし、音声が符号化される際に与えられるビットレートによって、符号化ノイズが最も低減される理論上の限度が規定される。符号化ノイズを少なくとも可能な限り除去することが目的となる。

【０００３】

符号化ノイズに対する適切な考えは、その符号化ノイズが付加的なホワイトノイズ又はカラーノイズであると仮定することである。デコーダでオーディオ信号又は音声信号の復号化した後に、符号化ノイズをより低減するように修正し、その結果、オーディオ信号又は音声の品質を向上させるある種のエンハンスメント方法が存在する。そのような技術は、一般に「ポストフィルタリング」と呼ばれる。これは、実際のデコーダの後の後処理において、改善されたオーディオ信号又は音声信号が得られることを意味する。ポストフィルタによる音質の改善に関する文献は多く存在する。最も基本的な文献の一部は非特許文献１乃至４である。

【０００４】

ピッチポストフィルタの基本的な動作原理は、有声音声の高調波間のスペクトルの谷に入り込む符号化ノイズの少なくとも一部を除去することである。これは、一般に、復号化音声信号を、その復号化音声信号をタイムシフトした信号に重み付き重ね合わせすることにより達成される。ここで、タイムシフトは音声のピッチラグ又はピッチ周期に対応する。その結果、特に音声高調波間において所望の音声信号に対して相関性のない符号化ノイズが減衰する。上述の効果は、非再帰フィルタ構造及び再帰フィルタ構造の双方により得られる。実際には、非再帰フィルタ構造が好ましい。

【０００５】

本発明の説明は、ピッチポストフィルタ又は微細構造ポストフィルタに関する。それらの基本的な動作原理は、有声音声の高調波間のスペクトルの谷に入り込む符号化ノイズの少なくとも一部を除去することである。これは、一般に、復号化音声信号を、その復号化音声信号をタイムシフトした信号に重み付き重ね合わせすることにより達成される。ここで、タイムシフトは音声のピッチラグ又はピッチ期間に対応する。また、タイムシフトされた信号が後続の音声信号サンプルに含まれることが好ましい。より最近の１つの非再帰ピッチポストフィルタ方法は、特許文献１において説明される。ここで、信号符号化のピッチパラメータは、対応する信号サンプルのポストフィルタリングにおいて再利用される

。特許文献１の非再帰ピッチポストフィルタ方法は、3GPPのAMR-WB+オーディオ及び音声符号化標準規格3GPP TS 26.290, "Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions"、及び、3GPPのVMR-WB［3GPP2 C.S0052-A, "Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB), Service Options 62 and 63 for Spread Spectrum Systems"］においても適用される。特許文献２は、１つのピッチポストフィルタ方法を開示する。同文献は、１つの同一フレーム内の過去及び後続の合成音声を使用することについて記載している。

【先行技術文献】

【非特許文献】

【０００６】

【非特許文献１】P. Kroon、B. Atal, "Quantization procedures for 4.8 kbps CELP coders", in Proc IEEE ICASSP, pp. 1650-1654, 1987.

【非特許文献２】V. Ramamoorthy, N.S. Jayant, "Enhancement of ADPCM speech by adaptive postfiltering", AT&T Bell Labs Tech. J., pp. 1465-1475, 1984.

【非特許文献３】V. Ramamoorthy, N.S., Jayant, R. Cox, M. Sondhi, "Enhancement of ADPCM speech coding with backward-adaptive algorithms for postfiltering and noise feed-back", IEEE J. on Selected Areas in Communications, vol. SAC-6, pp. 364-382, 1988.

【非特許文献４】J. H. Chen, A. Gersho, "Adaptive postfiltering for quality enhancements of coded speech", IEEE Trans. Speech Audio Process., vol. 3, no. 1, 1995

【特許文献】

【０００７】

【特許文献１】米国特許出願公開第２００５／０１６５６０３Ａ１号

【特許文献２】欧州特許第０８０７３０７Ｂ１号

【特許文献３】欧州特許第１０５００４０Ｂ１号

【発明の概要】

【発明が解決しようとする課題】

【０００８】

後続の音声信号を評価するピッチポストフィルタの１つの問題は、後続の１ピッチ周期の復号化オーディオ信号又は音声信号へのアクセスを必要とすることである。一般に、この後続の信号をポストフィルタで利用可能にすることは、復号化オーディオ信号又は音声信号をバッファリングすることにより可能である。しかし、オーディオ又は音声コーデックの従来のアプリケーションにおいては、これはコーデックのアルゴリズム的な遅延が増加するものであり、通信品質及び特に対話性に影響を与えるため、望ましくない。

【０００９】

本発明の目的は、デコーダ装置によりオーディオ又は音声の品質を向上することである。本発明の更なる目的は、オーディオ信号又は音声信号の遅延増加の原因とならないスケーラブルなデコーダ装置用の効果的なポストフィルタの構成を提供することである。

【課題を解決するための手段】

【００１０】

上記目的は、添付の請求の範囲に係る装置及び方法により達成される。第１の側面によれば、デコーダ装置は、フレームごとに符号化信号のパラメータを入力する受信機入力部と、前記受信機入力部に接続され、前記パラメータに基づいて復号化オーディオ信号のフレームを出力するデコーダとを含む。前記受信機入力部及び前記デコーダの少なくともいずれか一方は、第１のフレームのパラメータが前記受信機入力部で利用可能になった時と、前記第１のフレームの復号化オーディオ信号が前記デコーダの出力で利用可能になる時との間に少なくとも１つのフレームに対応する時間差が生じるように構成される。ポストフィルタは、前記デコーダの出力及び前記受信機入力部に接続される。前記ポストフィルタは、各後続フレームの前記パラメータに応じて前記復号化オーディオ信号のフレームを

フィルタリングして前記出力信号を得るように構成される。デコーダ装置は、前記ポストフィルタに接続された前記出力信号の出力部も有する。

【００１１】

第２の側面によれば、復号化方法は、フレームごとに符号化信号のパラメータを受信する受信ステップと、前記パラメータを復号化して復号化オーディオ信号を得る復号化ステップとを有する。前記受信ステップ及び前記復号化ステップの少なくともいずれか一方は、第１のフレームのパラメータが受信後に利用可能になった時と、前記第１のフレームの復号化オーディオ信号が復号化後に利用可能になる時との間に少なくとも１つのフレームに対応する時間差を生じさせる。前記復号化オーディオ信号のフレームは、各後続フレームの前記パラメータに応じてポストフィルタリングされて出力信号が得られる。前記方法は、前記出力信号を出力するステップも有する。

【発明の効果】

【００１２】

本発明の１つの利点は、音声及びオーディオコーデックの再構成信号の品質を向上できることである。例えば、コーデックがスケーラブル音声及びオーディオコーデックである場合、あるいはコーデックが受信端末のジッタバッファと共にＶｏＩＰアプリケーションにおいて使用される場合、追加の遅延という不利益を伴わずに再構成信号品質を向上させることができる。特に、例えば音声の開始のような過渡音における改善が可能である。

【図面の簡単な説明】

【００１３】

【図１】ポストフィルタを有するオーディオ又は音声コーデックの基本構造を示す図である。

【図２】本発明に係るデコーダ装置の一実施形態を示すブロック図である。

【図３】本発明に係るデコーダ装置の別の実施形態を示すブロック図である。

【図４】一般的なスケーラブルオーディオ又は音声コーデックを示すブロック図である。

【図５】上位レイヤが非音声オーディオ信号の符号化をサポートする場合の別のスケーラブルオーディオコーデックを示すブロック図である。

【図６】本発明に係る方法の一実施形態の手順を示すフローチャートである。

【図７】本発明に係るスケーラブルデコーダ装置の一実施形態を示すブロック図である。

【図８】本発明に係るスケーラブルデコーダ装置の別の実施形態を示すブロック図である。

【図９】本発明に係るスケーラブルデコーダ装置の更に別の実施形態を示すブロック図である。

【図１０】本発明に係るスケーラブルデコーダ装置の別の実施形態を示すブロック図である。

【図１１】本発明に係る改良されたピッチ進みパラメータ計算を説明する図である。

【図１２】。

【発明を実施するための形態】

【００１４】

本開示において、各図面及び実施形態における同等の又は直接対応する機能は同一の符号で示される。

【００１５】

詳細な説明の十分な理解を提供するため、いくつかの用語については混乱を避けるために、明示的に定義をしておく必要があろう。本開示において、用語「パラメータ」は一般名称として使用され、ビット又はビットストリームを含む任意の種類の信号の表現を表す。

【００１６】

本発明により達成される利点を理解するために、詳細な説明は、一般的なポストフィルタリングの簡単な説明から始める。図１は、ポストフィルタを含むオーディオ又は音声コーデックの基本構造を示す。送信機１は、入力オーディオ又は音声信号３をパラメータ４

のストリームに符号化するエンコーダ１０を含む。一般に、パラメータ４は符号化され、受信機２に転送される。受信機２はデコーダ２０を含み、デコーダ２０は、元のオーディオ又は音声信号３を表すパラメータ４を受信し、それらのパラメータ４を復号化オーディオ又は音声信号５に復号化する。復号化オーディオ又は音声信号５は、可能な限り元のオーディオ又は音声信号３と類似することが意図される。しかし、復号化オーディオ又は音声信号５は多少の符号化ノイズを常に含む。受信機２はポストフィルタ３０を更に含み、ポストフィルタ３０は、復号化オーディオ又は音声信号５をデコーダ２０から受信し、ポストフィルタリング手順を実行し、そしてポストフィルタ復号化オーディオ又は音声信号６を出力する。

【００１７】

ポストフィルタの基本概念は、符号化ノイズがより低減されるように符号化ノイズのスペクトル形状を形成することであり、これは実質的に人間の聴知覚特性を活用する。一般にこれは、音声信号が相対的に高い電力（スペクトルピーク）を有する知覚感度の低い周波数領域にノイズが移動し、音声信号が低い電力（スペクトル谷）を有する領域からノイズが除去されるように行われる。２つの基本的なポストフィルタ手法として、フォルマントポストフィルタ、ピッチポストフィルタ、微細構造ポストフィルタとも呼ばれる短期ポストフィルタ及び長期ポストフィルタがある。適切な性能を得るために、適応ポストフィルタがよく使用される。

【００１８】

上述のように、ピッチポストフィルタ又は微細構造ポストフィルタは本発明で有用である。復号化音声信号のタイムシフトされた信号への復号化音声信号の重ね合わせの結果、特に音声高調波間において所望の音声信号に対して相関性のない符号化ノイズが減衰する。上述の効果は、非再帰型フィルタ構造及び再帰型フィルタ構造の双方により得られる。非特許文献４で説明される１つのそのような一般的な形式は以下の式で与えられる。

【００１９】

$$H(z) = \frac{1 + \alpha z^{-T}}{1 - \beta z^{-T}}$$

ただし、Tは音声のピッチ周期に対応する。

【００２０】

実際には、非再帰型フィルタ構造が好ましい。最近の１つの非再帰型ピッチポストフィルタ方法は、米国特許出願公開第２００５／０１６５６０３号（特許文献１）において説明される。これは、オーディオ及び音声符号化標準規格である、３ＧＰＰ（第３世代パートナーシッププロジェクト）のAMR-WB+（拡張適応マルチレート広帯域コーデック）［3GPP TS 26.290］、及び、３ＧＰＰ２のVMR-WB（可変レートマルチモード広帯域（VMR-WB）コーデック）［3GPP2 C.S0052-A: "Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB), Service Options 62 and 63 for Spread Spectrum Systems"］において適用される。ここで、基本概念は、第１に以下の関係により符号化ノイズ推定値r(n)を計算することである。

【００２１】

$$r(n) = y(n) - y_p(n)$$

ただし、y(n)は復号化オーディオ又は音声信号、y_pは以下のように計算される予測信号である。

$$y_p = 0.5 \cdot (y(n-T)) + y(n+T)) \qquad (1)$$

【００２２】

第２に、ノイズ推定値をローパス（又はバンドパス）フィルタリングし、係数αで重み付けした値を音声信号から減算し、その結果、エンハンスメントオーディオ又は音声信号を得る。

$$y_{enh}(n) = y(n) - \alpha \cdot LP\{r(n)\} \qquad (2)$$

【００２３】

符号が反転した場合のローパスフィルタリングされたノイズ信号の適切な解釈は、その信号を符号化ノイズの低周波数部分を補償するエンハンスメント信号として見ることである。係数αは、予測信号及び復号化音声信号の相関性、予測信号のエネルギ、並びに音声信号及び予測信号の差のエネルギのある時間平均に応じて適応化される。

【００２４】

上述のように、上記定義式 $y_p = 0.5 \cdot (y(n-T)) + y(n+T))$ を評価する従来技術のピッチポストフィルタの１つの問題は、それらのポストフィルタが後続の１ピッチ周期の復号化音声信号y(n+T)が必要となり、その結果、アルゴリズム的な遅延が増加することである。AMR-WB+及びVMR-WBでは、得られた復号化オーディオ又は音声信号に基づいて、復号化オーディオ又は音声信号を後方に延ばすことにより、また、そのオーディオ又は音声信号はピッチ周期Tで周期的に延びていくと仮定することにより、その問題を解決している。復号化オーディオ又は音声信号が時間インデックスn^+までのみ利用可能であるという仮定の下、後続のピッチ周期は以下の式に従って計算される。

【００２５】

$$\hat{y}(n+T) = \begin{cases} y(n+T) & n+T < n^+ \\ y(n) & n+T \geq n^+ \end{cases}$$

【００２６】

この拡張は、単なる近似であるため、本来の後続の復号化音声信号を使用した場合に得られる品質と比較すると、品質に関しては妥協したものとなる。なお、特許文献２も、この問題に対しては望ましい解決策を提供するものではない。特許文献２は、エンハンスされるサブフレームに後続するサブフレームが利用可能であるという条件の下、現フレームにおける後続の合成音声データを用いたポストフィルタリングのみが行われることを記載している。これに対し本明細書は、後続のフレームではなく、現在の音声フレームまでの音声フレームを利用することを想定している。

【００２７】

別のポストフィルタ方法が特許文献３に開示されるが、本発明の説明においては関連性は低い。同文献は、ポストフィルタの強度が平均ビットレートに応じて制御される可変レート音声コーデックのためのポストフィルタ方法を説明している。

【００２８】

従来のポストフィルタ（例えば、フォルマントポストフィルタ／ピッチポストフィルタ）は、コーデックの遅延を最小限に維持するために、いかなる遅延も生じない。これは、符号化遅延配分が例えば将来のためにエンコーダにおいてより効果的に費やされるためである。このため、ポストフィルタのエンハンスメント能力を低下させる以下の問題が発生する。

【００２９】

なお、時間延長は、特に音声信号のピッチ期間が一定でない場合に問題となる。これは、とりわけ有声音声開始の場合に当てはまる。更に一般的には、音声の過渡状態に対する従来のポストフィルタの性能は、それらのパラメータの信頼性が同程度に低いため最適ではないことが示されている。

【００３０】

従って、本発明の基本概念の重要な部分は、将来のフレームからの情報を利用することによってポストフィルタ性能を向上することである。そのために、受信動作及び復号化動作における固有の時間遅延が利用される。本発明は、後続フレームのパラメータが利用可能になる時に関連して又はそれより後に、フレームの復号化信号が利用可能になるという状況に基づく。換言すると、受信機入力及びデコーダにより構成される集合体は、第１のフレームnに後続するフレームn+1のパラメータx(n+1)とほぼ同時に第１のフレームnの復号化信号y(n)を出力するように構成される。復号化音声フレームy(n)は、エンハンス出力音声フレーム$y_{out}(n)$を生成するポストフィルタに供給される。本発明によれば、ポストフィルタの動作は、ポストフィルタに少なくとも１つの後続フレームn+1のパラメータx(n

+1)へのアクセスを提供することによって改善される。信号遅延は受信動作及び復号化動作において固有であるため、追加の信号遅延が発生することはない。

【００３１】

一実施形態は、少なくともフレーム長Lだけ出力の遅延を生じるアルゴリズムに従って動作するデコーダを含む。デコーダが復号化音声フレームy(n)を出力し、それがポストフィルタリングで使用されるようになったときに、インデックスn+1の符号化音声フレームが受信機において利用可能になる。そのような遅延は、種々のデコーダ装置において生じうる。図２は、本発明に係るデコーダ装置のそのような一実施形態を示すブロック図である。受信機２は受信機入力部４０を有し、一般には符号化音声又はオーディオ信号であるフレームに基づく符号化信号x(n+1)を表すパラメータ４を受信するように構成される。デコーダ２０は受信機入力部４０に接続され、パラメータ４に基づいて復号化オーディオ信号５のフレームy(n)を出力するように構成される。デコーダ２０は、第１のフレームのパラメータ４が受信機入力部４０で利用可能になった時と、第１のフレームの復号化オーディオ信号がデコーダ２０の出力で利用可能になる時との間に時間差が生じるように構成される。この時間差は、少なくとも１つのフレームに対応する。本実施形態において、復号化動作は１フレーム分の信号の遅延５１を発生させる。デコーダ２０及び受信機入力部４０の集合体５０は、復号化信号y(n)を出力すると同時に後続フレームx(n+1)のパラメータを入力する。

【００３２】

ポストフィルタ３０は、デコーダ２０の出力及び受信機入力部４０に接続される。ポストフィルタ３０は、後続フレームのパラメータx(n+1)に応じて復号化オーディオ信号のフレーム５に基づいて出力信号６を提供するように構成される。それにより、将来の信号フレームの情報をポストフィルタリング処理において利用可能となるが、追加の復号化遅延を生じることはない。受信機出力部６０はポストフィルタ３０と接続され、出力信号６を出力する。

【００３３】

ＶｏＩＰシステムの１つの重要な要素は、受信端末におけるジッタバッファである。その目的は、パケットに含まれる受信した符号化音声フレームの非同期ストリームを同期ストリームに変換することである。同期ストリームは、その後音声デコーダにより復号化される。従って、ジッタバッファは上述の概念に従ってパラメータバッファとして動作可能である。換言すると、本発明の一実施形態は、ＶｏＩＰアプリケーションにおいて有利に適用可能であり、受信端末のジッタバッファは、バッファが空でないという条件の下、将来のフレームへのアクセスを容易に提供する。

【００３４】

従って、本発明の別の実施形態は、パラメータバッファを含む受信機入力を含む。パラメータバッファは受信した少なくとも２つの符号化音声フレームを格納する。デコーダは、バッファリングされたフレームnを復号化し、復号化音声フレームy(n)を出力する。それと同時に、インデックスn+1の符号化音声フレームがパラメータバッファにおいて入手可能であり、ポストフィルタリングで使用可能である。図３は、本発明に係るデコーダ装置のそのような一実施形態を示すブロック図である。受信機２は受信機入力部４０を含み、フレームに基づく符号化信号を表すパラメータ４を受信するように構成される。受信機入力部４０は、少なくとも２つのフレームのパラメータに対する格納位置４２Ａ、４２Ｂを有するジッタバッファ４１を含む。

【００３５】

デコーダ２０は、ジッタバッファ４１の第１の位置４２Ａに接続され、第１のフレームx(n)のパラメータ４Ａを受信する。デコーダ２０は、パラメータ４Ａに基づいて復号化オーディオ信号５のフレームy(n)を出力するように構成される。受信機入力部４０は、ジッタバッファ４１のために、特定のフレームのパラメータ４Ｂが受信機入力部４０において利用可能になった時と、同一フレームの復号化オーディオ信号５がデコーダ２０の出力で利用可能になる時との間に時間差を生じさせる。その時間差は、少なくとも１つのフレー

ムに対応する。本実施形態において、ジッタの動作は、少なくとも１フレーム分の信号の遅延を発生させる。デコーダ２０及び受信機入力部４０の集合体５０は、復号化信号y(n)を出力すると同時に後続フレームx(n+1)のパラメータを入力する。ポストフィルタ３０は、図２と同様に構成される。

【００３６】

図４は、本発明に係る方法の一実施形態の手順を示すフローチャートである。復号化方法は、ステップ２００で開始する。ステップ２１０において、フレームに基づく符号化信号のパラメータが受信される。ステップ２１２において、パラメータは復号化オーディオ信号のフレームに復号化される。ステップ２１０及び２１２の少なくともいずれか一方は、第１のフレームのパラメータが受信後に利用可能になった時と、第１のフレームの復号化オーディオ信号が復号化後に利用可能になる時との間に時間差を生じさせる。この時間差は、少なくとも１つのフレームに対応する。ステップ２１４において、復号化オーディオ信号のフレームは、各後続フレームのパラメータに応じてポストフィルタリングされて出力信号が得られる。ステップ２１６において、出力信号が出力される。処理は、ステップ２９９で終了する。

【００３７】

固有の遅延を有するコーデックの一般的な例は、スケーラブルコーデック又はエンベデッドコーデックである。従って、スケーラブルコーデックの簡単な説明を以下に提示する。図５は、一般的なスケーラブルオーディオ又は音声コーデックシステムを示すブロック図である。ここでは、送信機１は、入力オーディオ又は音声信号３をパラメータ４のストリームに符号化するエンコーダ１０、この例においてはスケーラブルエンコーダ１１０を含む。符号化全体は、２つのレイヤ、送信機において１次エンコーダ１１を含む下位レイヤ７及び送信機において２次エンコーダ１５を含む少なくとも１つの上位レイヤ８において行われる。スケーラブルコーデック装置は追加のレイヤを有してもよいが、ここでは２レイヤのデコーダシステムがモデルシステムとして使用される。しかし、本発明の原理は、３つ以上のレイヤを含むスケーラブルコーデックにも適用可能である。

【００３８】

１次エンコーダ１１は、入力オーディオ又は音声信号３を受信し、それを１次パラメータ１２のストリームに符号化する。更に１次エンコーダは、１次パラメータ１２を推定１次信号１３に復号化する。推定１次信号１３は、デコーダ側で１次パラメータ１２から取得される信号に対応するのが理想的である。推定１次信号１３は、比較器１４、この場合は減算器、において、元の入力オーディオ又は音声信号３と比較される。従って、差信号が、１次エンコーダ１１の１次符号化ノイズ信号１６である。１次符号化ノイズ信号１６は２次エンコーダに供給され、２次エンコーダはその信号を２次パラメータ１７のストリームに符号化する。それらの２次パラメータ１７は、１次パラメータ１２から復号化可能な信号の好適なエンハンスメントパラメータとして考えられる。それと共に、１次パラメータ１２及び２次パラメータ１７は、入力オーディオ又は音声信号３のパラメータ４の一般的なストリームを形成する。

【００３９】

一般に、パラメータ４は符号化されて受信機２に転送される。受信機２はデコーダ２０、この例においてはスケーラブルデコーダ１２０、を含み、デコーダは、元のオーディオ又は音声信号３を表すパラメータ４を受信し、それらのパラメータ４を復号化オーディオ又は音声信号５に復号化する。復号化全体は、２つのレイヤ、すなわち下位レイヤ７及び上位レイヤ８において行われる。受信機において、下位レイヤ７は１次デコーダ２１を含む。同様に、上位レイヤ８は受信機において２次デコーダ２５を含む。１次デコーダ２１は、パラメータ４のストリームの入力１次パラメータ２２を受信する。それらのパラメータはエンコーダ１０において作成されるパラメータと同一であるのが理想的であるが、送信ノイズはパラメータを歪ませている場合がある。１次デコーダ２１は、入力１次パラメータ２２を復号化１次オーディオ又は音声信号２３に復号化する。２次デコーダ２５は、同様に、パラメータ４のストリームの入力２次パラメータ２７を受信する。それらのパラ

メータはエンコーダ１０において作成されるパラメータと同一であるのが理想的であるが、この場合も送信ノイズはパラメータを歪ませている場合がある。２次デコーダ２５は、入力２次パラメータ２７を復号化エンハンスメントオーディオ又は音声信号２６に復号化する。この復号化エンハンスメントオーディオ又は音声信号２６は、可能な限り正確に１次エンコーダ１１の符号化ノイズに対応し、それにより１次デコーダ２１から結果として得られる符号化ノイズと類似することが意図される。復号化１次オーディオ又は音声信号２３及び復号化エンハンスメントオーディオ又は音声信号２６は、加算器２４で加算され、最終的な出力信号５が出力される。

【００４０】

１次パラメータ２２のみが受信機２において受信される場合、受信機が１次復号化のみをサポートする場合、あるいは、何らかの理由により２次復号化を実行しないと決定した場合は、得られる復号化エンハンスメントオーディオ又は音声信号２６はゼロとなり、出力信号５は復号化１次オーディオ又は音声信号２３と同一になる。これは、スケーラブルコーデックシステムの概念のフレキシビリティである。従来技術によれば、一般に、ポストフィルタリングは出力信号５に対して実行される。

【００４１】

今日、最も使用されるスケーラブル音声圧縮アルゴリズムは、１９９８年１１月のＩＴＵ－Ｔ勧告Ｇ．７１１「音声周波数のパルス符号変調（ＰＣＭ）（Pulse code modulation (PCM) of voice frequencies）」による６４ｋｂｐｓのA/U-law対数ＰＣＭコーデックである。８ｋＨｚサンプリングのＧ．７１１コーデックは、１２ビット又は１３ビットリニアＰＣＭ（パルス符号変調）サンプルを８ビット対数サンプルに変換する。対数サンプルのビット表現は、Ｇ．７１１ビットストリームの最下位ビット（ＬＳＢ）スチールを可能にし、Ｇ．７１１コーデックは実際には４８、５６及び６４ｋｂｐｓの間でＳＮＲ（信号対雑音比）スケーラブルとなる。このＧ．７１１コーデックのスケーラビリティは、帯域内制御信号の目的で回線交換通信網において使用される。このＧ．７１１のスケーラビリティの使用の最近の例は、従来の６４ｋｂｐｓのＰＣＭリンクを介する広帯域音声の設定及び転送を可能にする３ＧＰＰ－ＴＦＯプロトコル（３ＧＰＰのTS28.062によると、ＴＦＯ＝Tandem Free Operation）である。元の６４ｋｂｐｓのＧ．７１１ストリームのうちの８ｋｂｐｓは、狭帯域サービス品質にそれ程影響を及ぼさずに広帯域音声サービスの呼設定を可能にするためにまず使用される。呼設定の後、広帯域音声は、６４ｋｂｐｓのＧ．７１１ストリームのうち１６ｋｂｐｓを使用する。オープンループ・スケーラビリティをサポートする他の従来の音声符号化標準規格としては、１９９０年１２月のＩＴＵ－Ｔ勧告Ｇ．７２７「5-, 4-, 3-, and 2-bit/sample embedded adaptive differential pulse code modulation (ADPCM)」や、Ｇ．７２２（サブバンドＡＤＰＣＭ）がある。

【００４２】

スケーラブル音声符号化技術における更なる最近の進歩は、ＭＰＥＧ－４（ＭＰＥＧ＝Moving Picture Experts Group)ＣＥＬＰにスケーラビリティを提供するＭＰＥＧ－４規格（ＩＳＯ／ＩＥＣ－１４４９６）である。ＭＰＥ基本レイヤは、追加のフィルタパラメータ情報又は追加の新しいパラメータ情報の送信により拡張されうる。国際電気通信連合の標準化部門であるＩＴＵ－Ｔは、近年、Ｇ．７２９．ＥＶと呼ばれるＩＴＵ－Ｔ勧告Ｇ．７２９．１「G.729 based Embedded Variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729」（２００６年５月）に係る新たなスケーラブルコーデックの標準化を終了した。このスケーラブル音声コーデックのビットレートの範囲は、８ｋｂｐｓ～３２ｋｂｐｓである。このコーデックの主な使用例は、いくつかのＶｏＩＰ（Voice over IP（インターネットプロトコル））呼び出しの間の共有ｘＤＳＬ６４／１２８ｋｂｐｓ（ＤＳＬ＝デジタル加入者回線、ｘＤＳＬ＝種々の特定のＤＳＬ方法の一般的名称）アップリンク等のホーム又はオフィスゲートウェイにおける制限のある帯域幅リソースの効率的な共有を可能にすることである。

【００４３】

スケーラブル音声符号化の最近の１つの傾向は、音楽等の非音声オーディオ信号の符号

化のサポートを上位レイヤに提供することである。１つのそのような方法を図６に示す。そのようなコーデックにおいて、下位レイヤ７は、例えばＣＥＬＰ（符号励振線形予測）が周知の例である合成による分析（ＡｂＳ）パラダイムに従う単なる従来の音声符号化を採用する。本実施形態において、１次エンコーダ１１はＣＥＬＰエンコーダ１８であり、１次デコーダ２１はＣＥＬＰデコーダ２８である。そのような符号化は音声にのみよく適しており、音楽等の非音声オーディオ信号にはあまり適さないため、上位レイヤ８はオーディオコーデックにおいて使用される符号化パラダイムに従って動作する。従って、本実施形態において、２次エンコーダはオーディオエンコーダ１９であり、２次デコーダはオーディオデコーダ２９である。本実施形態において、一般に上位レイヤ８の符号化は下位レイヤの符号化の符号化エラーに対して動作する。

【００４４】

下位レイヤが１次デコーダ２１において１次復号化信号y_pへの１次復号化を実行し、上位レイヤが２次デコーダ２５において２次エンハンスメント信号y_sへの２次復号化を実行するスケーラブル音声／オーディオデコーダ１２０の応用例において、本発明の特定の一実施形態を図７に示す。２次エンハンスメント信号y_sは、１次復号化信号y_pをエンハンスメント復号化信号y_eにエンハンスする。本実施形態において、デコーダ２０は例えば２０ｍｓの長さの音声フレームに対して動作し、１次デコーダ２１は少なくとも１つのフレームの２次デコーダ２５よりも低遅延であると仮定する。すなわち、本質的な遅延５１は２次デコーダ２５内に存在する。

【００４５】

いくつかの特殊なコーデックシステムにおいて、２次コーデックは１次コーデックと異なるフレーム長で動作してもよい。例えば、２次コーデックは１次コーデックと比較して半分のフレーム長を有してもよく、従って２次コーデックは、１次デコーダが１つのフレームを復号化する間に２つの２次フレームを復号化する。設計によって、２次デコーダが有する遅延は、１次デコーダのフレーム長又は２次デコーダのフレーム長となる。

【００４６】

詳細には図７に示すように、１次デコーダ２１は特定の遅延なしで、すなわちフレームインデックスn+1の対応する受信符号化音声フレームデータx(n+1)に基づいて、n+1番目の音声フレームx(n+1)を１次復号化信号２３の出力フレームy_p(n+1)に復号化できると仮定する。これに対して、２次デコーダ２５は次の符号化フレームデータを必要とする。従って、２次デコーダ２５は、インデックスn+1の利用可能なフレームx(n+1)を使用して復号化２次エンハンスメント信号２６の復号化フレームy_s(n)を出力する。復号化２次エンハンスメント信号２６を１次復号化信号２３と適切に合成するために、１次復号化信号２３は１フレーム分遅延させる必要がある。これは遅延フィルタ５３において実行され、遅延復号化１次信号５４を与える。

【００４７】

これにより、デコーダにおける望ましくない遅延が更に増加するという不利益なしに本発明を適用できる。受信ビットストリームがエンハンスメントレイヤ情報を含む場合、復号化２次エンハンスメント信号２６のフレームy_s(n)が生成可能である。この信号２６は、遅延１次復号化信号のフレームy_p(n)と合成され、エンハンスメント復号化信号のフレームy_e(n)を形成する。このフレームy_e(n)は、パラメータのフレームx(n+1)が集合体５０Ｂから入手可能になった時に利用可能になる。その後、フレームy_e(n)は非因果性２次ポストフィルタ３０Ｂに供給され、ポストフィルタ３０Ｂは上述のように本発明から利益を得られる。それらの概念によれば、ポストフィルタ３０Ｂの動作は、フレームn+1の符号化パラメータを利用することにより改善可能である。また、このポストフィルタ３０Ｂは、まだ利用可能でない将来のフレームy_e(n+1)の近似を構成する１次復号化信号２３の次のフレームy_p(n+1)を使用することから更に利益を得られる。従って本実施形態において、ポストフィルタ３０Ｂは、将来のフレームのパラメータに基づくだけでなく将来のフレームの実際の信号の非常に適切な近似からも信号の品質を改善できる。それにより、２次ポストフィルタ３０Ｂは、デコーダ装置からの出力信号６としてポストフィルタエンハン

スメント信号５６を提供する。

【００４８】

図８は、本発明に係るスケーラブルデコーダ装置の別の実施形態を示すブロック図である。本実施形態において、１次ポストフィルタ３０Ａが提供され、遅延フィルタ５３からの出力に接続される。すなわち、１次ポストフィルタ３０Ａは遅延復号化１次信号５４に対して動作する。本実施形態において、集合体５０Ａは、受信機入力部４０、１次デコーダ２１及び遅延フィルタ５３を含む。本発明によると、１次ポストフィルタ３０Ａは後続フレームのパラメータにアクセス可能に動作する。本実施形態において、後続フレームの復号化１次信号２３も利用可能であり、１次ポストフィルタ３０Ａにおいて有利に使用される。換言すると、遅延復号化１次信号５４の音声フレーム$y_p(n)$は非因果性１次ポストフィルタ３０Ａによりエンハンス可能であり、ポストフィルタ３０Ａは復号化１次信号２３の音声フレーム$y_p(n+1)$及びフレームn+1のパラメータ４へのアクセスから利益を得る。

【００４９】

ポストフィルタ３０Ａからの出力信号５５、すなわち$y_p^*(n)$は、最終的な出力信号を生成するために２次エンハンスメント信号２６と合成するのに使用される。しかし、状況によっては、２次エンハンスメント信号２６により提供されるエンハンスメントは１次ポストフィルタ３０Ａにより得られるエンハンスメントと類似し、その結果、符号化ノイズの過剰補償になる場合がある。そのような場合、ポストフィルタ３０Ａは、２次符号化に対するパラメータが受信機入力部４０において利用可能であるかを判定するように構成されるのが有利である。２次パラメータが利用可能である場合、ポストフィルタの動作はＯＦＦにされて１次ポストフィルタ３０Ａからの出力として元の復号化１次信号を与えるか、あるいは少なくとも２次エンハンスメント信号の動作を干渉しないようにポストフィルタリングの原理を変更する。

【００５０】

図９は、本発明に係るスケーラブルデコーダ装置の更に別の実施形態を示すブロック図である。本実施形態において、図７のように２次デコーダ２５の後に２次ポストフィルタ３０Ｂが存在するが、１次ポストフィルタ３０Ａも提供される。そのような実施形態において、２次デコーダ２５からエンハンスされた出力信号は、２次ポストフィルタ３０Ｂを使用して更に改善される。この場合も、２次ポストフィルタ３０Ｂの動作は後続フレームのパラメータに基づくことができる。このポストフィルタ３０Ｂはエンハンスメントデコーダ出力５の将来のフレーム$y_e(n+1)$にアクセスできないが、その一方で、ポストフィルタ３０Ｂの動作は１次復号化信号の将来のフレーム$y_p(n+1)$に基づくことができる。１次集合体５０Ａが受信機入力部４０、１次デコーダ２１及び遅延フィルタ５３を含む一方で、２次集合体５０Ｂは受信機入力部４０、スケーラブルデコーダ全体１２０及び１次ポストフィルタ３０Ａを含む。

【００５１】

図１０は、本発明に係るスケーラブルデコーダ装置の更なる実施形態を示すブロック図である。ここで、ポストフィルタリングされていない遅延復号化１次信号５４は加算器２４に提供され、２次エンハンスメント信号２６と合成される。これにより、１次ポストフィルタ３０Ａの符号化ノイズ修正及び２次デコーダ２５からのエンハンスメントの混合を回避する。その代わり、出力部６０は、セレクタ６１として構成され、ポストフィルタ復号化１次信号５５又はポストフィルタエンハンスメント信号５６をデコーダ装置からの出力信号として出力するように構成される。セレクタ６１は、破線矢印６２により示されるように、入力信号に応じて動作されるのが好ましい。より多くのこれらの可能性について、以下に更に説明する。

【００５２】

上述したように、本発明の更なる部分の側面は、音声又はオーディオ信号の特性に依存してポストフィルタの非因果性エンハンスメントを適用することである。特に、そのような適用は、音声の過渡状態に有益である。この音声の過渡状態とは、例えば相対的に固定又は静止している１つの音素（音声要素）から別の音素に遷移する期間をいう。一般的な

そのような過渡状態においては、信号が静止しておらず、音声エンコーダにより行われるパラメータ推定の信頼性が安定した音声の期間より低い。ポストフィルタがそのような信頼性の低いパラメータに基づく場合、ポストフィルタの性能は低い可能性が高い。本発明によると、そのような過渡状態でのポストフィルタ性能は、パラメータ及び好ましくは将来のフレームの合成音声を利用することにより向上される。将来のフレームの間の音声がより安定し、より信頼性のあるパラメータ推定を可能にするため、ポストフィルタ性能は向上する。

【００５３】

本実施形態は、特定の非因果性ポストフィルタ動作が可能になる過渡状態の検出に依存する。そのような検出は、音声分類器により行われる。単純な例において、音声分類器は音声アクティビティ検出器（VAD：voice activity detector）であってもよいが、あるいはより一般的には、基本的な音声／非音声識別とは異なる、有声音、無声音、音声開始等の種々の音声を区別できる音声検出器（sound detector）であってもよい。そのような検出は、エネルギ又はＬＰＣパラメータ等の特定の信号パラメータの時間変化の評価に基づくことができ、それらのパラメータが急激に変化する音声又はオーディオ信号の部分を、過渡状態として識別することができる。過渡状態検出器は、エンコーダ又はデコーダで実現されてもよく、前者の場合、検出情報を受信機に送信することが必要になる。オーディオ特性の変化は、有意度（significance degree）で定量化されて測定され、ポストフィルタの動作を制御するために使用される。特に本発明に係るポストフィルタは、ピッチポストフィルタにおいて使用されるピッチパラメータが後続フレームのピッチパラメータに基づくように適応するように構成されてもよい。その適応化は、現在のフレームと先行フレーム又は後続フレームとの間のオーディオ特性の変化の有意性の基準に依存して実行される。

【００５４】

ポストフィルタ性能が向上する１つの特定の好適な実施形態は、無音期間後の有声音開始に対する応用例である。ここでは特に、ポストフィルタはピッチポストフィルタであり、そのポストフィルタにおいて使用される将来のフレームからのパラメータは現在のフレームに後続するフレームに属するサブフレームピッチパラメータである。

【００５５】

ピッチポストフィルタの改善に対処する本発明の更なる好適な実施形態によると、ピッチパラメータは新しいより正確な方法で処理される。上述のように、最新のピッチポストフィルタは式（１）及び（２）に基づく表現を評価する。ここで、合成音声の過去及び将来のセグメントは現在の音声セグメントと合成される。セグメントは、サブフレームあるいはピッチ周期等の単位であってもよい。ピッチパラメータ値Ｔを使用すると、過去のセグメントは現在のセグメントに対して遅れているし、将来のセグメントは現在のセグメントに対して進んでいることになる。過去の音声セグメントに対して遅れパラメータとしてＴを使用することは、遅延したセグメントと現在の音声セグメントとの相関性を最大にする遅れ値としてＴを計算する一般的なAbS（analysis-by-synthesis）音声コーデックの適応コードブック探索パラダイムと一致するため概念上は適切である。

【００５６】

しかし、一般に、ピッチ遅れパラメータが将来のセグメントに対しても一定のままであると仮定されるため、将来のセグメントに対してＴを進みパラメータとして使用することは的確でない。これは、特にピッチが大きく変化する可能性のある過渡状態において問題である。特許文献２は、セグメント間の相関性の計算に基づいて追加の遅れ及び進み判定器を特定することによりその問題に対する解決策を提供する。しかし、これは演算量の点で不利である。

【００５７】

図１１を参照すると、本発明に係る問題に対する解決策は以下の通りである。ピッチポストフィルタは、現在のフレームn及び少なくとも１つの将来のフレームn+1に対するサブフレームピッチパラメータのベクトルにアクセスできると仮定する。一般に、各フレーム

は４つのサブフレームを含む。T[0]...T[3]は現在のフレームの４つのサブフレームピッチパラメータを示し、T[4]...T[7]は将来のフレームの４つのサブフレームピッチパラメータを示す。所定のセグメントに対する進みパラメータは、現在のセグメントへの時間的遅れのあるサブフレーム位置に関連するサブフレームピッチパラメータを探索することにより見つけられると仮定する。所定の現在のセグメント１００に対する図１１の例によると、これはサブフレームピッチ値T[4]の場合である。図から分かるように、進みパラメータとして現在のセグメントのピッチパラメータ値T[1]を使用することは、ピッチがより小さな値に変更しているため不的確である。

【００５８】

図１２を参照すると、所定のセグメントに対する進みパラメータが見つけられる時に従うアルゴリズムの好適な例は以下の通りである。図４のステップ２１４の一部である手順はステップ２２０で開始する。ステップ２２２において、現在のセグメントに後続する第１のサブフレームが選択される。現在のセグメントに後続するこの第１のサブフレームから開始して、ステップ２２４において、対応するサブフレームのピッチ値を差し引いたサブフレーム時間インデックスが現在のセグメントの時間インデックス以上であるかがチェックされる。現在のセグメントの時間インデックス以上である場合、ステップ２２６において、サブフレームのピッチ値は現在のセグメントに対するピッチ進みパラメータとして利用され、アルゴリズムはステップ２９９で終了する。現在のセグメントの時間インデックス以上でない場合は、次のサブフレームに対してチェックが繰り返される。ステップ２２８において、利用可能なサブフレームが更に存在するかがチェックされる。存在しない場合、手順はステップ２９９で終了する。存在する場合、ステップ２３０で新しいサブフレームが選択され、ステップ２２４のチェックが繰り返される。このアルゴリズムにおいて、サブフレーム時間インデックスは、例えばサブフレームの開始時間インデックス又は中間時間インデックスであってもよい。なお、相関性の計算が実行される必要のある範囲を限定することにより複雑さを軽減するのを助長できるため、特許文献２で説明される進み判定器が使用される場合、このアルゴリズムはある利得を伴って使用される。

【００５９】

上述の実施形態は、本発明のいくつかの例として理解されよう。本発明の範囲から逸脱することなく、実施形態に対する種々の変形、組み合わせ、変更が行われうることは、当業者には理解されるだろう。特に、技術的に可能であれば、種々の実施形態における種々の部分的な解決策は他の構成と組み合わせ可能である。本発明の範囲は添付の特許請求の範囲によって定義される。

【図１】

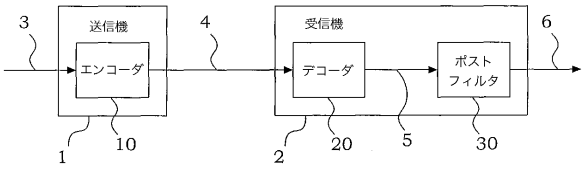


Fig. 1

【図２】

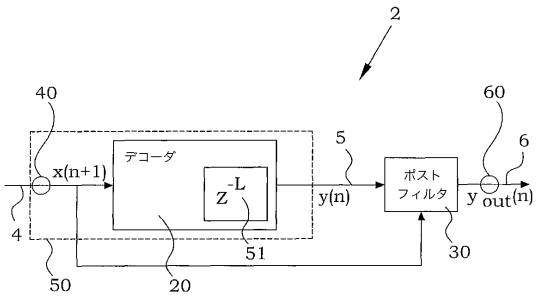


Fig. 2

【図３】

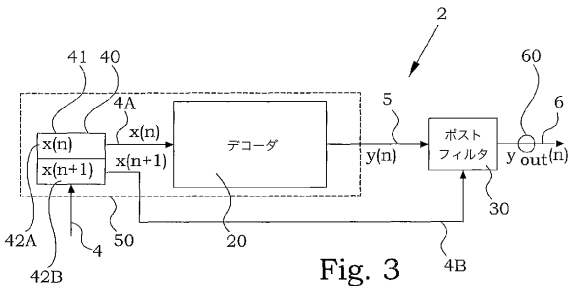


Fig. 3

【図４】

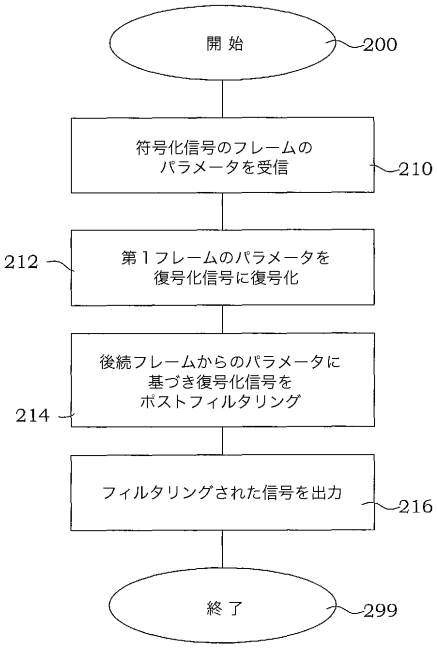


Fig. 4

【図５】

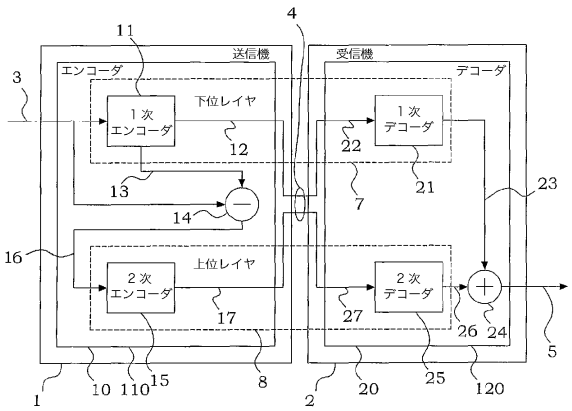


Fig. 5

【図６】

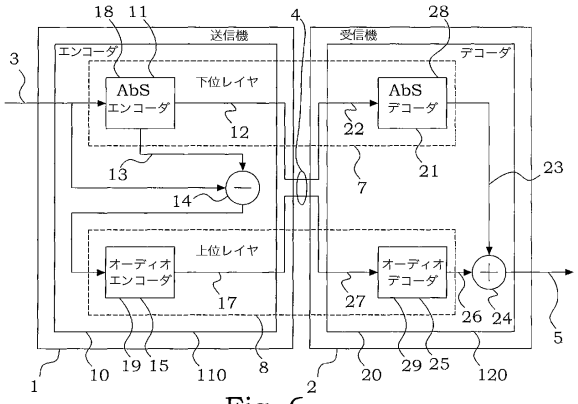


Fig. 6

【図７】

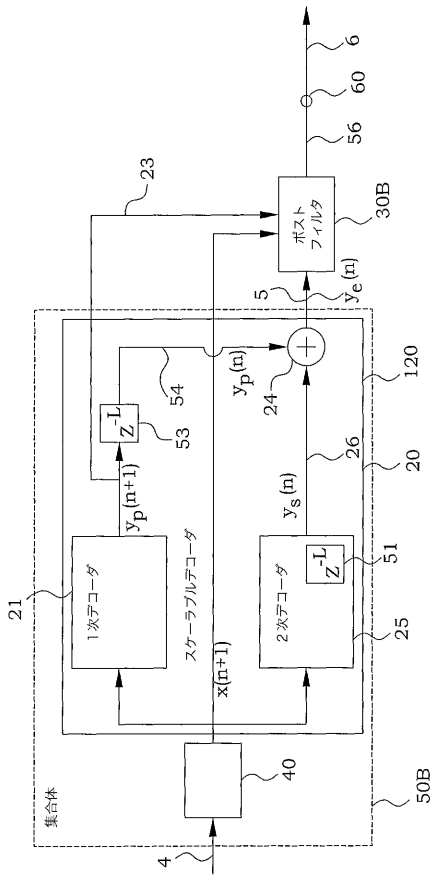


Fig. 7

【図８】

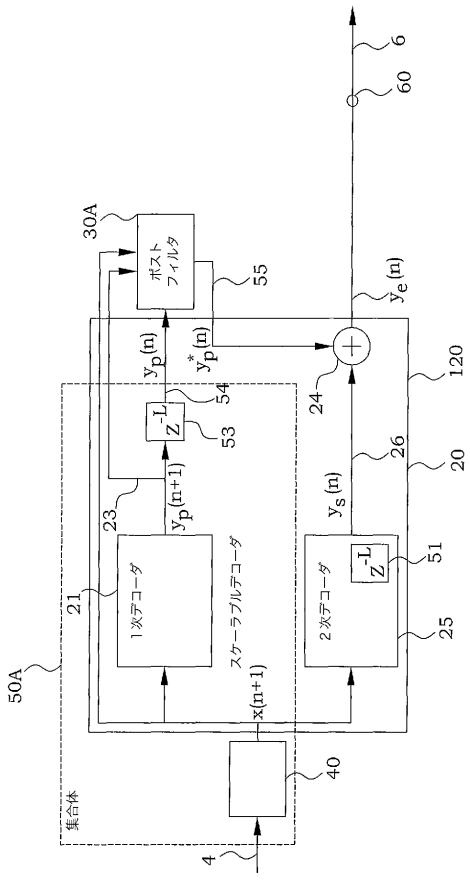


Fig. 8

【図９】

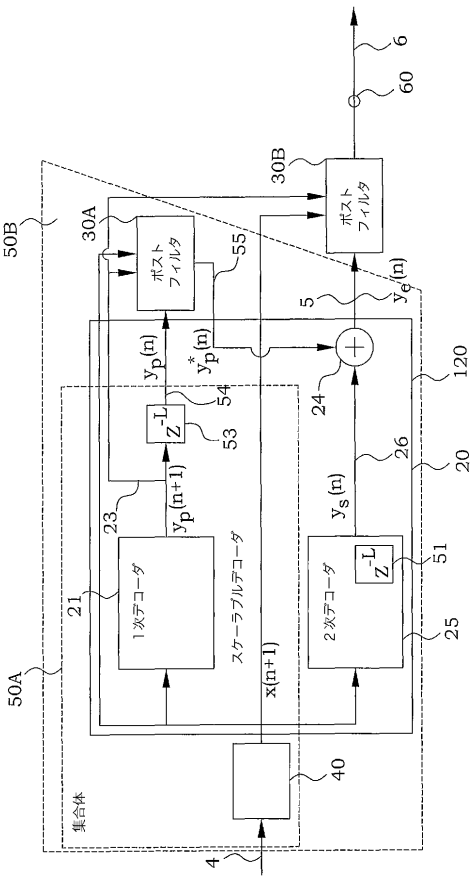


Fig. 9

【図１０】

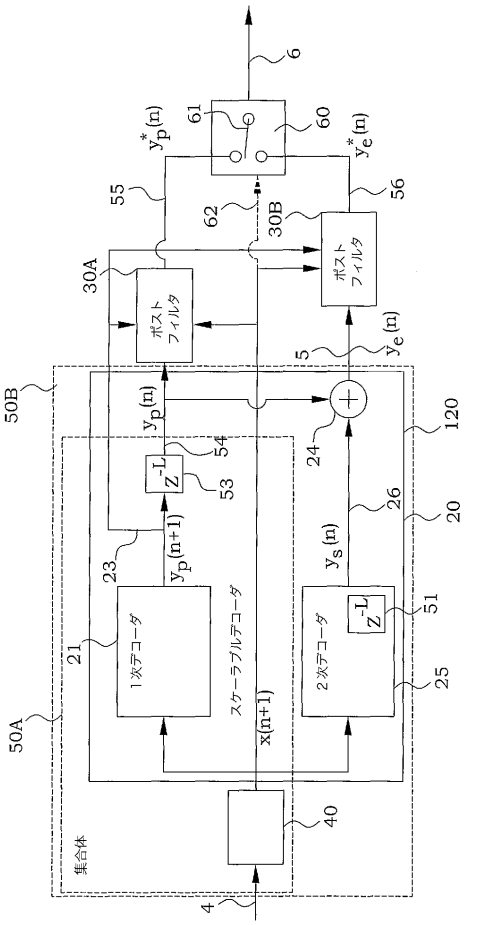


Fig. 10

【図１１】

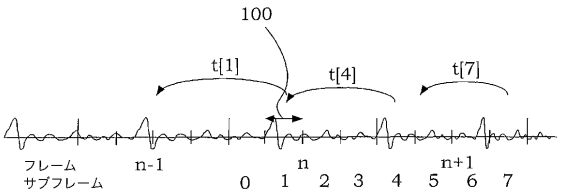


Fig. 11

【図１２】

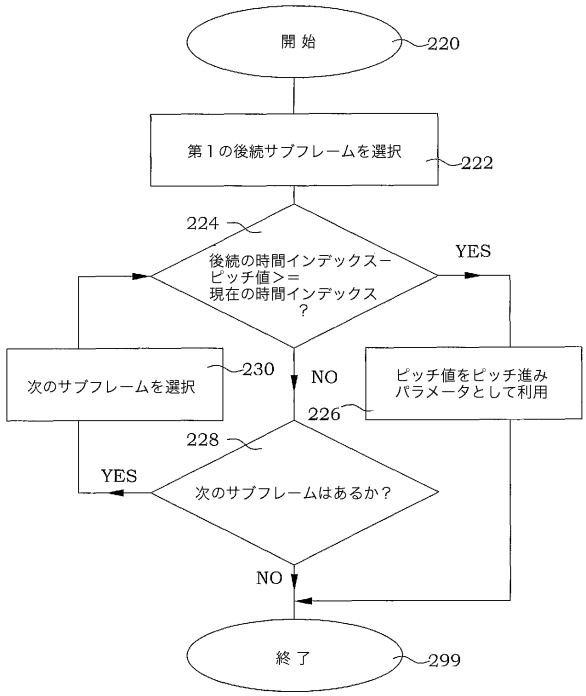


Fig. 12

フロントページの続き

(72)発明者　ブルーン，　ステファン
スウェーデン国　ソレンチュナ　エス－１９２　６７，　ヴァスヴェーゲン　１１ビー

審査官　五貫　昭一

(56)参考文献　特開平２－８２７１０（ＪＰ，Ａ）
特開平９－６３９７（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
G10L　19/14