(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6153661号
(P6153661)
(45)発行日　平成29年6月28日(2017.6.28)　(24)登録日　平成29年6月9日(2017.6.9)

(51)Int.Cl.　F I
G10L 19/005 (2013.01)　G10L 19/005

請求項の数 12　(全 68 頁)

(21)出願番号	特願2016-520420 (P2016-520420)
(86)(22)出願日	平成26年6月16日 (2014.6.16)
(65)公表番号	特表2016-523380 (P2016-523380A)
(43)公表日	平成28年8月8日 (2016.8.8)
(86)国際出願番号	PCT/EP2014/062578
(87)国際公開番号	WO2014/202535
(87)国際公開日	平成26年12月24日 (2014.12.24)
審査請求日	平成28年2月18日 (2016.2.18)
(31)優先権主張番号	13173157.2
(32)優先日	平成25年6月21日 (2013.6.21)
(33)優先権主張国	欧州特許庁 (EP)
(31)優先権主張番号	14166995.2
(32)優先日	平成26年5月5日 (2014.5.5)
(33)優先権主張国	欧州特許庁 (EP)

(73)特許権者　597159765
フラウンホーファーゲゼルシャフト　ツール　フォルデルンング　デル　アンゲヴァンテン　フォルシユング　エー．フアー．
ドイツ国，デー－８０６８６　ミュンヘン，ハンサシュトラーセ　２７ツェー
(74)代理人　100112715
弁理士　松山　隆夫
(74)代理人　100104444
弁理士　上羽　秀敏
(74)代理人　100125704
弁理士　坂根　剛
(74)代理人　100120662
弁理士　川上　桂子

最終頁に続く

(54)【発明の名称】改善されたパルス再同期化を採用するＡＣＥＬＰ型封じ込めにおける適応型コードブックの改善された封じ込めのための装置および方法

(57)【特許請求の範囲】
【請求項１】
再構成されたフレームとして音声信号を含むフレームを再構成するための装置であって、前記再構成されたフレームが、１以上の入手可能なフレームと関連し、前記１以上の入手可能なフレームが、再構成されたフレームの１以上の先行フレームおよび再構成されたフレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームが、１以上の入手可能なピッチサイクルとして、１以上のピッチサイクルを含み、装置が
１以上の入手可能なピッチサイクルの１つのサンプルの数と再構成されるべき第１のピッチサイクルのサンプルの数との差を示すサンプル数の差（$\Delta^{p}{}_{0}$；Δ_{i}；$\Delta^{p}{}_{k+1}$）を決定するための決定部（２１０）と、
サンプル数の差（$\Delta^{p}{}_{0}$；Δ_{i}；$\Delta^{p}{}_{k+1}$）および１以上の入手可能なピッチサイクルの前記１つのサンプルに依拠して、第１の再構成ピッチサイクルとして再構成されるべき第１のピッチサイクルを再構成することにより再構成フレームを再構成するためのフレーム再構成部（２２０）とを含み、
フレーム再構成部（２２０）が、再構成フレームを再構成するように構成され、それにより再構成フレームが完全にまたは部分的に第１の再構成ピッチサイクルを含み、再構成フレームが完全にまたは部分的に第２の再構成ピッチサイクルを含み、かつ第１の再構成ピッチサイクルのサンプルの数が第２の再構成ピッチサイクルのサンプルの数とは異なるようになっており、

フレーム再構成部（２２０）が、１以上の入手可能なピッチサイクルの前記１つに依拠して中間フレームを生成するようになっており、

フレーム再構成部（２２０）が、中間フレームが第１の部分中間ピッチサイクル、１以上のさらなる中間ピッチサイクルおよび第２の部分中間ピッチサイクルを含むように、中間フレームを生成するようになっており、

第１の部分中間ピッチサイクルが、１以上の入手可能なピッチサイクルの前記１つのサンプルの１以上に依拠し、１以上のさらなる中間ピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つのサンプルの全部に依拠し、かつ第２の部分中間ピッチサイクルが、前記１以上の入手可能なピッチサイクルの前記１つのサンプルの１以上に依拠し、

決定部（２１０）が、いくつのサンプルを第１の部分中間ピッチサイクルから除去するかまたは加えることになるかを示す開始部差数（Δp_0）を決定するよう構成され、かつフレーム再構成部（２２０）が、開始部差数（Δp_0）に依拠して第１の部分中間ピッチサイクルから１以上の第１のサンプルを除去するよう構成されるかまたは第１の部分中間ピッチサイクルに１以上の第１のサンプルを加えるよう構成され、

決定部（２１０）が、さらなる中間ピッチサイクルの各々について、いくつのサンプルをさらなる中間ピッチサイクルの前記１つから除去するかまたは加えることになるかを示すピッチサイクル差数（Δ_i）を決定するよう構成され、かつフレーム再構成部（２２０）が、前記ピッチサイクル差数（Δ_i）に依拠してさらなる中間ピッチサイクルの前記１つから１以上の第２のサンプルを除去するよう構成されるかまたはさらなる中間ピッチサイクルの前記１つに１以上の第２のサンプルを加えるよう構成され、かつ

決定部（２１０）が、第２の部分中間ピッチサイクルからいくつのサンプルを除去するかまたは加えるかを示す終了部差数（Δp_{k+1}）を決定するよう構成され、かつフレーム再構成部（２２０）が、終了部差数（Δp_{k+1}）に依拠して第２の部分中間ピッチサイクルから１以上の第３のサンプルを除去するよう構成されるかまたは第２の部分中間ピッチサイクルに１以上の第３のサンプルを加えるよう構成される、装置。

【請求項２】

決定部（２１０）が、再構成されるべき複数のピッチサイクルの各々について、サンプル数の差を決定するよう構成され、それにより、ピッチサイクルの各々のサンプル数の差が、１以上の入手可能なピッチサイクルの前記１つのサンプルの数と、再構成されるべき前記ピッチサイクルのサンプルの数との差を示すようになっており、かつ

フレーム再構成部（２２０）が、再構成フレームを再構成するために、再構成されるべき前記ピッチサイクルのサンプル数の差と１以上の入手可能なピッチサイクルの前記１つのサンプルとに依拠して、再構成されるべき複数のピッチサイクルの各サイクルを再構成するよう構成される、請求項１に記載の装置。

【請求項３】

決定部（２１０）が、再構成フレームとして再構成されるべきフレームの音声信号の１以上のパルスの位置を決定するよう構成され、かつ

フレーム再構成部（２２０）が、音声信号の１以上のパルスの位置に依拠して、再構成フレームを再構成するよう構成される、請求項１または請求項２に記載の装置。

【請求項４】

決定部（２１０）が、以下の式に従って、再構成フレームとして再構成されるべきフレームの音声信号の最後のパルスのインデクスｋを決定するよう構成され、

【数１６３】

$$k = \left\lfloor \frac{L - s - T[0]}{T_r} - 1 \right\rfloor$$

ここで、Ｌは再構成フレームのサンプルの数を示し、ｓはフレーム差値を示し、Ｔ［０］は、音声信号の最後のパルスとは異なる、再構成フレームとして再構成されるべきフレ

ームの音声信号のパルスの位置を示し、かつＴ$_r$は、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、

当該装置は、再構成フレームとして再構成されるべきフレームの音声信号の最後のパルスのインデックスｋに依拠して、再構成フレームとして再構成されるべきフレームを再構成するよう構成される、請求項１から請求項３のいずれかに記載の装置。

【請求項５】

決定部（２１０）が、以下の式に基づいて、１以上の入手可能なピッチサイクルの前記１つの丸められた長さＴ$_r$を決定するよう構成され、

【数１６４】

$$T_r = \lfloor T_p + 0.5 \rfloor$$

ここで、Ｔ$_p$が１以上の入手可能なピッチサイクルの前記１つの長さを示し、

当該装置は、１以上の入手可能なピッチサイクルの前記１つの丸められた長さＴ$_r$に依拠して、再構成フレームとして再構成されるべきフレームを再構成するよう構成される、請求項１から請求項４のいずれか１項に記載の装置。

【請求項６】

決定部（２１０）が、以下の式を適用することによりパラメータｓを決定するよう構成され、

【数１６５】

$$s = \delta \frac{L}{T_r} \frac{M+1}{2} - L\left(1 - \frac{T_p}{T_r}\right)$$

ここで、Ｔ$_p$が１以上の入手可能なピッチサイクルの前記１つの長さを示し、Ｔ$_r$が１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、再構成フレームとして再構成されるべきフレームがＭ個のサブフレームを含み、再構成フレームとして再構成されるべきフレームがＬ個のサンプルを含み、かつδが１以上の入手可能なピッチサイクルの前記１つのサンプルの数と再構成されるべき１以上のピッチサイクルの１つのサンプルの数との差を示す実数であり、

当該装置は、パラメータｓに依拠して、再構成フレームとして再構成されるべきフレームを再構成するよう構成される、請求項１から請求項５のいずれか１項に記載の装置。

【請求項７】

当該装置は、以下の式に依拠して、再構成フレームとして再構成されるべきフレームを再構成するよう構成され、

【数１６６】

$$\delta = \frac{T_{ext} - T_p}{M}$$

再構成フレームとして再構成されるべきフレームは、Ｍ個のサブフレームを含み、

Ｔ$_p$は、以上の入手可能なピッチサイクルの前記１つの長さを示し、かつ

Ｔｅｘｔは、再構成フレームとして再構成されるべきフレームの再構成されるべきピッチサイクルの１つの長さを示す、請求項１から請求項６のいずれか１項に記載の装置。

【請求項８】

フレーム再構成部（２２０）は、中間フレームが、１以上のさらなる中間ピッチサイクルとして１つのさらなる中間ピッチサイクルよりも多い第１の部分中間ピッチサイクルと、第２の部分中間ピッチサイクルとを含み、

当該装置は、以下の式に基づいて、１以上のさらなる中間ピッチサイクルの各々から削除すべきまたは追加すべきサンプルの数を計算するよう構成され、

【数１６７】

$$\Delta_i = |T_r - T_{ext}| - (k+1-i)a, 1 \le i \le k$$

Ｔ$_r$は、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、
Ｔ$_{ext}$は、再構成フレームとして再構成されるべきフレームの再構成されるべきピッチサイクルの１つの長さを示し、
ｋは、再構成フレームとして再構成されるべきフレームの音声信号の最後のパルスのインデックスを示し、
ｉは、整数であり、かつ
ａは、連続するピッチサイクル間で追加または削除されるべきサンプルのデルタを示す数である、請求項１に記載の装置。

【請求項９】
当該装置は、以下の式によって数ａを決定するよう構成され、

【数１６８】

$$a = \frac{|T_r - T_{ext}|(L-s) - |s|T_r}{(k+1)\left(T[0] + \frac{k}{2}T_r\right)}$$

Ｌは、再構成フレームのサンプルの数を示し、
ｓは、フレーム差値を示し、
Ｔ［０］は、音声信号の最後のパルスと異なる、再構成フレームとして再構成されるべきフレームの音声信号の１つのパルスの位置を示す、請求項８に記載の装置。

【請求項１０】
当該装置は、以下の式に基づいて、第１の部分中間ピッチサイクルから削除または追加されるべきサンプルの数を計算するよう構成され、

【数１６９】

$$\Delta_0^p = (|T_r - T_{ext}| - (k+1)a)\frac{T[0]}{T_r}$$

当該装置は、以下の式に基づいて、第２の部分中間ピッチサイクルから削除または追加されるべきサンプルの数を計算するよう構成される、請求項９に記載の装置。

【数１７０】

$$\Delta_{k+1}^p = |s| - \Delta_0^p - \sum_{i=1}^{k} \Delta_i$$

【請求項１１】
再構成フレームとして音声信号を含むフレームを再構成するための方法であって、前記再構成フレームが、１以上の入手可能なフレームと関連し、前記１以上の入手可能なフレームが、再構成フレームの１以上の先行フレームおよび再構成フレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームが、１以上の入手可能なピッチサイクルとして１以上のピッチサイクルを含み、方法が、
１以上の入手可能なピッチサイクルの１つのサンプルの数と再構成されるべき第１のピッチサイクルのサンプルの数との差を示すサンプル数の差を決定するステップと、
サンプル数の差および１以上の入手可能なピッチサイクルの前記１つのサンプルに依拠して、第１の再構成ピッチサイクルとして再構成されるべき第１のピッチサイクルを再構成することにより、再構成フレームを再構成するステップとを含み、
再構成フレームを再構成するステップが、再構成フレームが、完全にまたは部分的に第

１の再構成ピッチサイクルを含み、再構成フレームが完全にまたは部分的に第２の再構成ピッチサイクルを含み、かつ第１の再構成ピッチサイクルのサンプルの数が、第２の再構成ピッチサイクルのサンプルの数と異なるようになっており、

当該方法は、１以上の入手可能なピッチサイクルの前記１つに依拠して中間フレームを生成するステップを更に含み、

中間フレームを生成するステップは、中間フレームが第１の部分中間ピッチサイクル、１以上のさらなる中間ピッチサイクルおよび第２の部分中間ピッチサイクルを含むように行われ、

第１の部分中間ピッチサイクルが、１以上の入手可能なピッチサイクルの前記１つのサンプルの１以上に依拠し、１以上のさらなる中間ピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つのサンプルの全部に依拠し、かつ第２の部分中間ピッチサイクルが、前記１以上の入手可能なピッチサイクルの前記１つのサンプルの１以上に依拠し、

当該方法は、いくつのサンプルを第１の部分中間ピッチサイクルから除去するかまたは加えることになるかを示す開始部差数を決定するステップを更に含み、かつ、当該方法は、開始部差数に依拠して第１の部分中間ピッチサイクルから１以上の第１のサンプルを除去するよう構成されるかまたは第１の部分中間ピッチサイクルに１以上の第１のサンプルを加えるよう構成されるステップを更に含み、

当該方法は、さらなる中間ピッチサイクルの各々について、いくつのサンプルをさらなる中間ピッチサイクルの前記１つから除去するかまたは加えることになるかを示すピッチサイクル差数を決定するステップを更に含み、かつ、当該方法は、前記ピッチサイクル差数に依拠してさらなる中間ピッチサイクルの前記１つから１以上の第２のサンプルを除去するかまたはさらなる中間ピッチサイクルの前記１つに１以上の第２のサンプルを加えるよう構成されるステップを更に含み、かつ

当該方法は、第２の部分中間ピッチサイクルからいくつのサンプルを除去するかまたは加えるかを示す終了部差数を決定するステップを更に含み、かつ、当該方法は、フ終了部差数に依拠して第２の部分中間ピッチサイクルから１以上の第３のサンプルを除去するかまたは第２の部分中間ピッチサイクルに１以上の第３のサンプルを加えるよう構成されるステップを更に含む、方法。

【請求項１２】

コンピュータまたは信号プロセッサ上で実行されたときに、請求項１１に記載の方法を実現するためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、オーディオ信号処理、詳細には、音声処理に関し、かつより詳細には、ＡＣＥＬＰ型封じ込め（ＡＣＥＬＰ（Ａｌｇｅｂｒａｉｃ　Ｃｏｄｅ　Ｅｘｃｉｔｅｄ　Ｌｉｎｅａｒ　Ｐｒｅｄｉｃｔｉｏｎ）＝代数符号励振線形予測）における適応型コードブックの改善された封じ込めのための装置および方法に関する。

【背景技術】

【０００２】

オーディオ信号処理は、ますます重要度が増している。オーディオ信号処理の分野では、封じ込め技術が重要な役割を果たす。フレームが失われたり破損された場合、その失われたり破損されたフレームからの失われた情報を置換する必要がある。音声信号処理においては、特に、ＡＣＥＬＰまたはＡＣＥＬＰ型音声コーデックを考慮する場合、ピッチ情報が非常に重要である。ピッチ予測技術およびパルス再同期化技術が必要とされる。

【０００３】

ピッチの再構成に関して、様々なピッチ外挿技術が先行技術には存在する。

【０００４】

これらの技術の１つが、繰り返しによる技術である。前提技術のコーデックのほとんど

が単純な繰り返しによる封じ込めのアプローチを適用しており、これは、良好なフレームが到着して新しいピッチ情報をビットストリームから復号化できるまで、パケット損失前に最後に正確に受信されたピッチ周期を繰り返すことを意味する。あるいは、パケットの損失時よりもう少し前に受信されたピッチ値を選択することによるピッチ安定性論理を適用する。繰り返しによるアプローチに従うコーデックは、例えば、Ｇ．７１９（非特許文献９[ＩＴＵ０８ｂ、８．６]を参照）、Ｇ．７２９（非特許文献１０［ＩＴＵ１２、４．４］を参照）、ＡＭＲ（非特許文献２［３ＧＰ１２ａ、６．２．３．１］、非特許文献４［ＩＴＵ０３］を参照）、ＡＭＲ－ＷＢ（非特許文献３［３ＧＰ１２ｂ、６．２．３．４．２］を参照)およびＡＭＲ－ＷＢ＋（ＡＣＥＬＰおよびＴＣＸ２０（ＡＣＥＬＰ型）封じ込め)、（非特許文献１［３ＧＰ０９］を参照）（ＡＭＲ＝適応型マルチレート（Ａｄａｐｔｉｖｅ　Ｍｕｌｔｉ‐Ｒａｔｅ）、ＡＭＲ‐ＷＢ＝適応型マルチレートワイドバンド（Ａｄａｐｔｉｖｅ　Ｍｕｌｔｉ‐Ｒａｔｅ‐Ｗｉｄｅｂａｎｄ）である。

【０００５】

先行技術の他のピッチ再構成技術は、時間領域からのピッチの生成である。いくつかのコーデックについては、ピッチは、封じ込めのため必要だが、ビットストリームには埋め込まれない。したがって、ピッチ周期を計算するため、前のフレームの時間領域信号に基づいてピッチを計算して、次いでこれを封じ込め中、一定に保つ。このアプローチに従うコーデックは、たとえばＧ．７２２であり、特に、Ｇ．７２２補遺３（非特許文献５[ＩＴＵ０６ａ、III．６．６およびIII．６．７]を参照）およびＧ．７２２補遺４（非特許文献７[ＩＴＵ０７、　ＩＶ．６．１．２．５]を参照）を参照。

【０００６】

先行技術の他のピッチ再構成技術は、外挿によるものである。いくつかの前提技術のコーデックは、ピッチ外挿アプローチを適用し、かつ、応じてパケット損失中に、外挿されたピッチ推定値にピッチを変更する特定のアルゴリズムを実行する。こられのアプローチについては、以下にＧ．７１８およびＧ．７２９．１を参照してより詳細に説明する。

【０００７】

まず、Ｇ．７１８を考察する（非特許文献８［ＩＴＵ０８ａ］を参照）。未来のピッチの推定は、声門パルス再同期化モジュールを支持するために、外挿により実行される。可能な将来のピッチ値についてのこの情報は、封じ込められた励振の声門パルスを同期するために使用される。

【０００８】

最後の良好なフレームが、「無声」ではない場合にのみピッチ外挿が行われる。Ｇ．７１８のピッチ外挿は、エンコーダがスムーズなピッチ輪郭を有するという仮定に基づく。前記外挿は、消失前の最後の７つのサブフレームのピッチラグ$d^{[i]}_{fr}$に基づき実行される。

【０００９】

Ｇ．７１８においては、フレームが正しく受け取られるたびに浮動ピッチ値の履歴更新が行われる。この目的で、ピッチ値は、コアモードが「無声」以外の場合にのみ更新される。損失フレームの場合には、浮動ピッチラグ間の差$\Delta^{[i]}_{dfr}$が以下の式により計算される。

【数１】

$$\Delta^{[i]}_{dfr} = d^{[i]}_{fr} - d^{[i-1]}_{fr} \quad (i = -1, \cdots, -6) \qquad (1)$$

【００１０】

式（１）において、$d^{[-1]}_{fr}$は、前のフレームの最後の（すなわち、第４の）サブフレームのピッチラグを示し、$d^{[-2]}_{fr}$は、前のフレームの第３のサブフレームのピッチラグを示す等である。

【００１１】

Ｇ．７１８によれば、差$d^{[i]}_{fr}$の和は、以下のように計算される。

【数２】

$$s_{\Delta} = \sum_{i=-1}^{-6} \Delta_{dfr}^{[i]} \qquad (2)$$

【００１２】

値$\Delta^{[i]}_{dfr}$は、正または負が可能なので、$\Delta^{[i]}_{dfr}$の符号の反転回数が合計され、かつ第１の反転の位置が、メモリに保存されるパラメータにより示される。

【００１３】

パラメータf_{corr}は、以下の式により得られる。

【数３】

$$f_{corr} = 1 - \frac{\sqrt{\sum_{i=-1}^{-6} \left(\Delta_{dfr}^{[-i]} - s_{\Delta}\right)^2}}{6 \cdot d_{max}} \qquad (3)$$

ここで、d_{max}＝２３１は、最大想定ピッチラグである。

【００１４】

Ｇ．７１８において、最大の絶対差を示す位置i_{max}は、以下の定義により得られる。

【数４】

$$i_{max} = \left\{ max_{i=-1}^{-6} \left(abs\left(\Delta_{dfr}^{[i]}\right)\right)\right\}$$

【００１５】

この最大差についての比は、以下のように計算される。

【数５】

$$r_{max} = \left| \frac{5 \cdot \Delta_{dfr}^{[i_{max}]}}{\left(s_{\Delta} - \Delta_{dfr}^{[i_{max}]}\right)} \right| \qquad (4)$$

【００１６】

この比が５以上の場合、最後に正しく受け取られたフレームの４番目のサブフレームのピッチが、封じ込められるべき全サブフレームについて使用される。この比が５以上の場合、これは、アルゴリズムがこのピッチを外挿するのに十分に確実ではなく、かつ声門パルス再同期化が行われないことを意味する。

【００１７】

r_{max}が、５未満の場合、できる限り良好な外挿が行えるよう、さらなる処理が行われる。未来のピッチを外挿するために３つの異なる方法が利用される。可能なピッチ外挿アルゴリズムから選択を行うため、偏差パラメータ$f_{cor_{r2}}$を計算するが、これは、ファクタf_{cor_r}、および最大ピッチ変化の位置i_{max}に依存する。しかしながら、まず、平均から大きすぎるピッチ差を除去するために、平均浮動ピッチ差を修正する。

【００１８】

f_{cor_r}＜０．９８であり、i_{max}＝３の場合、２つのフレーム間の遷移に関連するピッチ差を除くために、平均分数ピッチ差／Δ_{dfr}が以下の式により決定される。

【数６】

$$\overline{\Delta}_{dfr} = \left(\frac{s_{\Delta} - \Delta_{dfr}^{[-4]} - \Delta_{dfr}^{[-5]}}{3} \right) \qquad (5)$$

【００１９】

$f_{corr} \geqq 0.98$ または $i_{max} \neq 3$ の場合、平均小数ピッチ差／Δ_{dfr}は、以下のとおり計算され、

【数７】

$$\overline{\Delta}_{dfr} = \frac{s_{\Delta} - \Delta_{dfr}^{[i_{max}]}}{6} \qquad (6)$$

かつ最大浮動ピッチ差は、この新しい平均値により置き換えられる。

【数８】

$$\Delta_{dfr}^{[i_{max}]} = \overline{\Delta}_{dfr} \qquad (7)$$

【００２０】

浮動ピッチ差のこの新しい平均で、正規化された偏差f_{corr2}は、以下のとおり計算される。

【数９】

$$f_{corr2} = 1 - \frac{\sqrt{\sum_{i=-1}^{I_{sf}} \left(\Delta_{dfr}^{[i]} - \overline{\Delta}_{dfr} \right)^2}}{I_{sf} \cdot d_{max}} \qquad (8)$$

ここで、I_{sf}は、第１のケースにおいては４であり、第２のケースでは６である。

【００２１】

この新しいパラメータに依拠して、未来のピッチを外挿する３つの方法の中から選択を行う。

【００２２】

・$\Delta^{[i]}_{dfr}$が２回を上回って符号を変え（高いピッチ変化を意味する）、第１の符号反転が、最後の良好なフレーム（$i < 3$について）におけるものであり、かつ$f_{corr2} > 0.945$の場合、外挿されるピッチd_{ext}（外挿されるピッチはT_{ext}とも表す）を以下のとおり計算する。

【数１０】

$$s_y = \sum_{i=-1}^{-4} \Delta_{dfr}^{[i]}$$

$$s_{xy} = \Delta_{dfr}^{[-2]} + 2 \cdot \Delta_{dfr}^{[-3]} + 3 \cdot \Delta_{dfr}^{[-4]}$$

$$d_{est} = round\left[\Delta_{fr}^{[-1]} + \left(\frac{(7 \cdot s_y - 3 \cdot s_{xy})}{10}\right)\right]$$

【００２３】

・０．９４５＜f_{corr2}＜０．９９で、かつ、Δ^{i}_{dfr}が１回以上符号を変える場合には、ピッチを外挿するために分数ピッチ差の重み付き平均が採用される。平均差の重み付けf_Wは、正規化された偏差f_{corr2}に関連し、かつ第１の符号の反転の位置は以下のとおり規定される。

【数１１】

$$f_w = f_{corr2} \cdot \left(\frac{i_{mem}}{7}\right)$$

【００２４】

この式のパラメータi_{mem}は、Δ^{i}_{dfr}の第１の符号反転の位置に依存するので、第１の符号反転が過去のフレームの最後の２つのサブフレーム間で起こっていれば、$i_{mem}=0$となり、第１の符号反転が過去のフレームの第２および第３のサブフレーム間で起こっていれば、$i_{mem}=1$となり、以下同様である。第１の符号反転が、最後のフレームの終りに近ければ、これは、ピッチの変化が損失フレームのすぐ前では、より安定していなかったことを意味する。したがって、平均に適用される重み付けファクタは、０に近くなり、外挿されたピッチd_{ext}は、最後の良好なフレームの第４のサブフレームのピッチに近くなる。

【数１２】

$$d_{ext} = round\left[\Delta_{fr}^{[-1]} + 4 \cdot \overline{\Delta}_{dfr} \cdot f_w\right]$$

【００２５】

・さもなければ、ピッチの展開は、安定していると考えられ、外挿されたピッチd_{ext}は、以下のとおり決定される。

【数１３】

$$d_{ext} = round\left[d_{fr}^{[-1]} + 4 \cdot \overline{\Delta}_{dfr}\right]$$

【００２６】

この処理の後、ピッチラグは、３４から２３１の範囲に制限される（これらの値は、最小および最大許容ピッチラグを示す）。

【００２７】

ここで、ピッチ再構成技術に基づく外挿の他の例を示すため、Ｇ.７２９.１を考える（非特許文献６［ＩＴＵ０６ｂ］を参照）。

【００２８】
Ｇ.７２９．１は、復号化可能な前方誤り封じ込め情報（フェーズ情報等）がない場合のピッチ外挿アプローチ（特許文献１［Ｇａｏ］を参照）を特徴とする。これは、たとえば、２つの連続するフレームが失われた場合に起こる（１つのスーパーフレームが、ＡＣＥＬＰまたはＴＣＸ２０いずれかが可能な４つのフレームからなる）。また、可能なＴＣＸ４０またはＴＣＸ８０フレームおよびそのほとんどすべての組み合わせが存在する。
【００２９】
有声領域で１以上のフレームが失われた場合、つねに前のピッチ情報を使用して現在失われているフレームを再構成する。現在の推定されるピッチの精度は、オリジナル信号の位相整合に直接影響を与える可能性があり、現在の損失フレームおよび損失フレーム後に受信されたフレームの再構成品質には決定的である。前のピッチラグを単にコピーするのではなく、いくつかの過去のピッチラグを使うことで、統計的により良いピッチ推定が得られると考えられる。Ｇ．７２９.１のコーダにおいて、ＦＥＣ（ＦＥＣ＝前方誤り訂正）のためのピッチ外挿は、過去の５つのピッチ値に基づく線形外挿から構成される。過去の５つのピッチ値は、Ｐ（ｉ）（ｉ＝０、１、２、３、４）で、Ｐ（４）が最も最近のピッチ値である。外挿モデルは、以下のとおり規定される。
【数１４】

$$P'(i) = a + i \cdot b \qquad (9)$$

【００３０】
損失フレームにおける、第１のサブフレームについての外挿されたピッチ値は、以下のとおり規定される。
【数１５】

$$P'(5) = a + 5 \cdot b \qquad (10)$$

【００３１】
係数ａおよびｂを決定するために、誤差Ｅを最小化する。誤差Ｅは、以下のとおり規定される。
【数１６】

$$\begin{aligned} E &= \sum_{i=0}^{4} [P'(i) - P(i)]^2 \\ &= \sum_{i=0}^{4} [(a + b \cdot i) - P(i)]^2 \end{aligned} \qquad (11)$$

以下のとおり設定することで、
【数１７】

$$\frac{\delta E}{\delta a} = 0 \quad \text{and} \quad \frac{\delta E}{\delta b} = 0 \qquad (12)$$

ａおよびｂは、以下のとおりになる。

【数１８】

$$a = \frac{3\sum_{i=0}^{4} P(i) - \sum_{i=0}^{4} i \cdot P(i)}{5} \quad \text{and} \quad b = \frac{\sum_{i=0}^{4} i \cdot P(i) - 2\sum_{i=0}^{4} P(i)}{10} \qquad (13)$$

【００３２】

以下では、非特許文献１１（［ＭＣＺ１１］）に提示されるようなＡＭＲ-ＷＢコーデックのための先行技術のフレーム消失封じ込めコンセプトについて説明する。このフレーム消失封じ込めコンセプトは、ピッチおよびゲイン線形予測に基づく。前記論文では、フレームの損失の場合に、最小二乗平均誤差基準（Ｍｉｎｉｍｕｍ　Ｍｅａｎ　Ｓｑｕａｒｅ　Ｅｒｒｏｒ　Ｃｒｉｔｅｒｉｏｎ）に基づいた線形ピッチ内挿／外挿アプローチを提案する。

【００３３】

このフレーム消失封じ込めコンセプトによれば、デコーダで、消失したフレームの前の最後の有効なフレーム（過去のフレーム）のタイプが、消失フレーム後の最も早いフレーム（未来のフレーム）のタイプと同じ場合、ピッチＰ（ｉ）が規定され、ｉ＝－Ｎ，－Ｎ＋１、．．．、０、１、．．．、Ｎ＋４、Ｎ＋５であり、かつ、Ｎは、消失したフレームの過去および未来のサブフレームの数である。Ｐ(１)、Ｐ（２）、Ｐ（３）、Ｐ（４）が、消失したフレームにおける４つのサブフレームの４つのピッチであり、Ｐ（０）、（－１）、．．．Ｐ（－Ｎ）が、過去のサブフレームのピッチであり、かつ、Ｐ（５）、Ｐ（６）、．．．、Ｐ（Ｎ＋５）が未来のサブフレームのピッチである。線形予測モデルＰ’（ｉ）＝ａ＋ｂ・ｉが採用される。ｉ＝１、２、３、４で、Ｐ’（１）、Ｐ’（２）、Ｐ’（３）、Ｐ’（４）は、消失したフレームについての予測ピッチである。ＭＭＳ基準（ＭＭＳ＝最小二乗平均（Ｍｉｎｉｍｕｍ　Ｍｅａｎ　Ｓｑｕａｒｅ））を考慮して、内挿アプローチにより、２つの予測される係数ａおよびｂの値を生成する。このアプローチによれば、誤差Ｅは、以下のとおり規定される。

【数１９】

$$\begin{aligned} E &= \sum_{-N}^{0} [P'(i) - P(i)]^2 + \sum_{5}^{N+5} [P'(i) - P(i)]^2 \\ &= \sum_{-N}^{0} [a + b \cdot i - P(i)]^2 + \sum_{5}^{N+5} [a + b \cdot i - P(i)]^2 \end{aligned} \qquad (14a)$$

【００３４】

次に、係数ａおよびｂは、以下を計算することにより得ることができる。

【数２０】

$$\frac{\delta E}{\delta a}=0 \text{ および } \frac{\delta E}{\delta a}=0 \qquad (14b)$$

$$a = \frac{2\left[\sum_{i=-N}^{0} P(i)+\sum_{i=5}^{N+5} P(i)\right]\cdot\left(N^3+9N^2+38N+1\right)}{(N+1)\cdot(4N^3+36N^2+107N-1)} \qquad (14c)$$

$$b = \frac{9\left[\sum_{i=-N}^{0} P(i)+\sum_{i=5}^{N+5} P(i)\right]}{1-107N-36N^2-4N^3} \qquad (14d)$$

【００３５】
消失フレームの最後の４つのサブフレームについてのピッチラグは、以下のとおり計算できる。

【数２１】

$$P'(1)= a+b\cdot 1;\quad P'(2)= a+b\cdot 2$$
$$P'(3)= a+b\cdot 3;\quad P'(4)= a+b\cdot 4 \qquad (14e)$$

【００３６】
Ｎ＝４で、最良の結果が得られることがわかる。Ｎ＝４とは、過去の５つのサブフレームと未来の５つのサブフレームを内挿に使用することを意味する。

【００３７】
しかしながら、過去のフレームのタイプが未来のフレームのタイプと異なる場合、例えば、過去のフレームが有声で、未来のフレームが無声の場合、上記の外挿アプローチを使用して、消失フレームのピッチを予測するために、過去または未来のフレームの有声のピッチだけが使用される。

【００３８】
ここで、特にＧ．７１８およびＧ．７２９．１を参照して、先行技術のパルス再同期化を考える。パルス再同期化のためのアプローチは、特許文献２（［ＶＪＧＳ１２］）に記載される。

【００３９】
まず、励振の周期的部分を構成することについて説明する。

【００４０】
「無声」以外の正しく受信されたフレームに続く消失したフレームを封じ込めるため、励振の周期的部分を、前のフレームのローパスフィルタ処理した最後のピッチ周期を繰り返すことにより構成する。

【００４１】
周期的部分の構成は、前のフレームの終りから励振信号のローパスフィルタ処理されたセグメントの単純なコピーを使用することによって行う。

【００４２】
ピッチ周期長さは、最も近い整数に丸められる。

【数２２】

$$Tc=round(last_pitch) \qquad (15a)$$

【００４３】
最後のピッチ周期の長さがT_pであると考えれば、コピーされたセグメントの長さT_rは、たとえば以下のように規定され得る。
【数２３】

$$T_r=[T_p+0.5] \qquad (15b)$$

【００４４】
周期的な部分は、１つのフレームおよび１つの追加サブフレームについて構成される。
【００４５】
たとえば、フレームにおいてＭ個のサブフレームがあれば、サブフレームの長さはＬ＿ｓｕｂｆｒ＝Ｌ／Ｍであり、ここで、Ｌは、フレームの長さであり、L_{frame}としても示される（$L=L_{frame}$）。
【００４６】
図３は、音声信号の構成された周期的部分を示す。
【００４７】
Ｔ［０］は、励振の構成された周期的部分における第１の最大パルスの場所である。他のパルスの位置は、以下の式により与えられる。
【数２４】

$$T[i]=T[0]+iTc \qquad (16a)$$

【００４８】
これは、以下の式に対応する。
【数２５】

$$T[i]=T[0]+iTr \qquad (16b)$$

【００４９】
励振の周期的部分の構成の後、損失フレーム（Ｐ）における最後のパルスの推定されるターゲット位置と励振の構成された周期的部分におけるその実際の位置（Ｔ［ｋ］）との間の差を修正するために、声門パルス再同期化が行われる。
【００５０】
ピッチラグ展開は、損失フレームの前の最後の７つのサブフレームのピッチラグに基づいて外挿される。各サブフレームにおける展開ピッチラグは、以下のとおりである。
【数２６】

$$p[i]=round(T_c+(i+1)\delta),\ 0\le i<M \qquad (17a)$$

ここで

【数２７】

$$\delta = \frac{T_{ext} - T_c}{M} \qquad (17b)$$

であり、かつ、T_{ext}（d_{ext}とも呼ぶ）は、d_{ext}についての上に記載する外挿ピッチである。

【００５１】

一定ピッチ（T_c）のピッチサイクル内のサンプルの合計数の和と、展開するピッチｐ［ｉ］のピッチサイクル内のサンプルの合計数の和との間のｄで示す差は、フレーム長さの範囲内でみつかる。ｄを見つける方法について文献には記載がない。

【００５２】

Ｇ．７１８（非特許文献８［ＩＴＵ０８ａ］を参照）のソースコードでは、ｄは、以下のアルゴリズムを用いて見つけられる（ここで、Ｍは、フレームにおけるサブフレームの数）。

```
ftmp = p[0];
i = 1;
while (ftmp < L_frame - pit_min) {
        sect = (short)(ftmp*M/L_frame);
        ftmp += p[sect];
        i++;
}
d = (short)(i*Tc - ftmp);
```

フレーム長さの範囲で構成される周期的部分のパルス＋未来のフレームにおける第１パルスの数はＮである。Ｎを見つける方法について文献には記載がない。

【００５３】

Ｇ．７１８（非特許文献［ＩＴＵ０８ａ］を参照）のソースコードにおいて、Ｎは以下のとおり見つけられる。

【数２８】

$$N = 1 + \left\lfloor \frac{L_frame}{Tc} \right\rfloor \qquad (18a)$$

【００５４】

損失フレームに属する励振の構成された周期的部分における最後のパルスＴ［ｎ］の位置は、以下の式により決定される。

【数２９】

$$n = \begin{cases} N-1 \ , \ T[N-1] < L_frame \\ N-2 \ , \ T[N-1] \geq L_frame \end{cases} \qquad (18b)$$

【００５５】

推定される最後のパルス位置は、

【数３０】

$$P = T[n] + d \qquad (19a)$$

である。

【００５６】

最後のパルス位置の実際の位置Ｔ［ｋ］は、推定されるターゲット位置Ｐに最も近い励振（サーチにおける現在のフレームの後の最初のパルスを含む）の構成された周期的部分のパルスの位置である。

【数３１】

$$\forall i \ |T[k] - P| \leq |T[i] - P|, \ 0 \leq i < N \qquad (19b)$$

【００５７】

声門パルス再同期化は、フルピッチサイクルの最小エネルギ領域においてサンプルを加えたり除いたりすることにより行われる。加えたり除いたりするサンプルの数は以下の差により決定される。

【数３２】

$$diff = P - T[k] \qquad (19c)$$

【００５８】

最小エネルギ領域は、スライドする５サンプルのウィンドウを使用して決定される。最小エネルギ位置は、エネルギが最小のウィンドウの中央に設定される。Ｔ［ｉ］＋Ｔｃ／８～Ｔ［ｉ＋１］－Ｔｃ／４からの２つのピッチパルス間でサーチが行われる。N_{min}＝ｎ-１の最小エネルギ領域が存在する。

【００５９】

N_{min}＝１の場合、最小エネルギ領域は、１つしかなく、ｄｉｆｆサンプルがその位置で挿入されるかまたは削除される。

【００６０】

N_{min}＞１については、最初に加えられるかまたは除かれるサンプルは、より少なく、フレームの終りに向かって多くなる。パルスＴ［ｉ］とＴ［ｉ＋１］との間で除かれるかまたは加えられるサンプルの数は、以下の再帰関係に従って見つけられる。

【数３３】

$$R[i] = round\left(\frac{(i+1)^2}{2} f - \sum_{k=0}^{i-1} R[k]\right) \text{ with } \quad f = \frac{2|diff|}{N_{\min}^2}$$

【００６１】

Ｒ［ｉ］＜Ｒ［ｉ－１］の場合、Ｒ［ｉ］およびＲ［ｉ－１］の値が交換される。

【先行技術文献】

【特許文献】

【００６２】

【特許文献１】ヨーロッパ特許第２００２４２７Ｂ１号（[Gao] Yang Gao, Pitch prediction for packet loss concealment, European Patent 2 002 427 B1）

【特許文献２】米国特許第８２５５２０７Ｂ２号（[VJGS12] Tommy Vaillancourt, Milan Jelinek, Philippe Gournay, and Redwan Salami, Method and device for efficient f

rame erasure concealment in speech codecs, US 8,255,207 B2, 2012)
【非特許文献】
【００６３】
【非特許文献１】[3GP09] 3GPP; Technical Specification Group Services and System Aspects, Extended adaptive multi-rate - wideband (AMR-WB+) codec, 3GPP TS 26.290, 3rd Generation Partnership Project, 2009
【非特許文献２】[3GP12a], Adaptive multi-rate (AMR) speech codec; error concealment of lost frames (release 11), 3GPP TS 26.091, 3rd Generation Partnership Project, Sep 2012
【非特許文献３】[3GP12b], Speech codec speech processing functions; adaptive multi-rate - wideband (AMRWB) speech codec; error concealment of erroneous or lost frames, 3GPP TS 26.191, 3rd Generation Partnership Project, Sep 2012
【非特許文献４】[ITU03] ITU-T, Wideband coding of speech at around 16 kbit/s using adaptive multi-rate wideband (amr-wb), Recommendation ITU-T G.722.2, Telecommunication Standardization Sector of ITU, Jul 2003
【非特許文献５】[ITU06a], G.722 Appendix III: A high-complexity algorithm for packet loss concealment for G.722, ITU-T Recommendation, ITU-T, Nov 2006
【非特許文献６】[ITU06b], G.729.1: G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with g.729, Recommendation ITU-T G.729.1, Telecommunication Standardization Sector of ITU, May 2006
【非特許文献７】[ITU07], G.722 Appendix IV: A low-complexity algorithm for packet loss concealment with G.722, ITU-T Recommendation, ITU-T, Aug 2007
【非特許文献８】[ITU08a], G.718: Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s, Recommendation ITU-T G.718, Telecommunication Standardization Sector of ITU, Jun 2008
【非特許文献９】[ITU08b], G.719: Low-complexity, full-band audio coding for high-quality, conversational applications, Recommendation ITU-T G.719, Telecommunication Standardization Sector of ITU, Jun 2008
【非特許文献１０】[ITU12], G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (cs-acelp), Recommendation ITU-T G.729, Telecommunication Standardization Sector of ITU, June 2012
【非特許文献１１】[MCZ11] Xinwen Mu, Hexin Chen, and Yan Zhao, A frame erasure concealment method based on pitch and gain linear prediction for AMR-WB codec, Consumer Electronics (ICCE), 2011 IEEE International Conference on, Jan 2011, pp. 815－816
【非特許文献１２】[MTTA90] J.S. Marques, I. Trancoso, J.M. Tribolet, and L.B. Almeida, Improved pitch prediction with fractional delays in celp coding, Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on, 1990, pp. 665－668 vol.2
【発明の概要】
【００６４】
本発明の目的は、オーディオ信号処理についての改善されたコンセプトを提供することであり、特に、音声処理についての改善されたコンセプトを提供することであり、かつより詳細には、改善された封じ込めのコンセプトを提供することである。
【００６５】
本発明の目的は、請求項１に記載の装置、請求項１６に記載の方法および請求項１７に記載のコンピュータプログラムにより解決される。
【００６６】
再構成されたフレームとして音声信号を含むフレームを再構成するための装置が提供さ

れ、前記再構成されたフレームが、１以上の入手可能なフレームと関連し、前記１以上の入手可能なフレームが、再構成されたフレームの１以上の先行フレームおよび再構成されたフレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームが、１以上の入手可能なピッチサイクルとして、１以上のピッチサイクルを含む。この装置は、１以上の入手可能なピッチサイクルのうちの１つのサンプルの数と再構成されるべき第１のピッチサイクルのサンプルの数との差を示すサンプル数の差を決定するための決定部を含む。また、この装置は、サンプル数の差と、１以上の入手可能なピッチサイクルのうちの前記１つのサンプルとに依拠して、第１の再構成ピッチサイクルとして再構成されるべき第１のピッチサイクルを再構成することにより再構成フレームを再構成するためのフレーム再構成部を含む。フレーム再構成部は、再構成フレームを再構成するように構成され、それにより再構成フレームが完全にまたは部分的に第１の再構成ピッチサイクルを含み、再構成フレームが完全にまたは部分的に第２の再構成ピッチサイクルを含み、かつ第１の再構成ピッチサイクルのサンプル数が第２の再構成ピッチサイクルのサンプル数とは異なるようになっている。

【００６７】

ある実施形態によれば、決定部は、たとえば、再構成対象の複数のピッチサイクルの各々についてサンプル数の差を決定して、それによりピッチサイクルの各々のサンプル数の差が、１以上の入手可能なピッチサイクルのうちの前記１つのサンプルの数と、再構成対象の前記ピッチサイクルのサンプル数との差を示すようになっている。フレーム再構成部は、たとえば、再構成対象の前記ピッチサイクルのサンプル数の差および１以上の入手可能なピッチサイクルの前記１つのサンプルに依拠して、再構成対象の複数のピッチサイクルの各ピッチサイクルを再構成して、再構成フレームを再構成するよう構成され得る。

【００６８】

ある実施形態においては、フレーム再構成部は、たとえば、１以上の入手可能なピッチサイクルのうちの前記１つに依拠して中間フレームを生成するよう構成され得る。フレーム再構成部は、たとえば、再構成フレームを得るために、中間フレームを修正するよう構成され得る。

【００６９】

実施形態によれば、決定部は、たとえば、いくつのサンプルを中間フレームから取り除くかまたはいくつのサンプルを中間フレームに加えるかを示すフレーム差値（ｄ；ｓ）を決定するよう構成され得る。また、フレーム再構成部は、たとえば、フレーム差値が、第１のサンプルがフレームから除去されることを示す場合、再構成フレームを得るために中間フレームから第１のサンプルを除去するよう構成され得る。さらに、フレーム再構成部は、たとえば、フレーム差値（ｄ；ｓ）が、第２のサンプルがフレームに加えられることを示す場合、再構成フレームを得るために中間フレームに第２のサンプルを加えるよう構成され得る。

【００７０】

ある実施形態において、フレーム再構成部は、たとえば、フレーム差値が、第１のサンプルがフレームから除去されるべきことを示す場合、中間フレームから第１のサンプルを除去するよう構成することが可能で、中間フレームから除去される第１のサンプルの数がフレーム差値により示されるようになっている。また、フレーム再構成部は、たとえば、フレーム差値が、第２のサンプルがフレームに加えられるべきことを示す場合、中間フレームに第２のサンプルを加えるよう構成することが可能で、中間フレームに加えられる第２のサンプルの数がフレーム差値により示されるようになっている。

【００７１】

ある実施形態によれば、決定部は、たとえば、以下の式が真であるように、フレーム差数ｓを決定するように構成され得る。

【数３４】

$$s = \sum_{i=0}^{M-1} (p[i] - T_r) \frac{L}{MT_r}$$

ここで、Ｌは、再構成フレームのサンプルの数を表し、Ｍは、再構成フレームのサブフレームの数を表し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められたピッチ周期長さを示し、ｐ［ｉ］は、再構成フレームのｉ番目のサブフレームの再構成されたピッチサイクルのピッチ周期長さを示す。

【００７２】

ある実施形態において、フレーム再構成部は、たとえば、１以上の入手可能なピッチサイクルの前記１つに依拠して中間フレームを生成するようになっていてもよい。また、フレーム再構成部は、たとえば、中間フレームが、第１の部分中間ピッチサイクル、１以上のさらなる中間ピッチサイクルおよび第２の部分中間ピッチサイクルを含むように中間フレームを生成するようにされてもよい。さらに、第１の部分中間ピッチサイクルは、たとえば、１以上の入手可能なピッチサイクルのうちの前記１つのサンプルのうちの１以上に依拠することが可能で、１以上のさらなる中間ピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つのサンプルの全部に依拠し、かつ第２の部分中間ピッチサイクルが、１以上の入手可能なピッチサイクルの前記１つのサンプルのうちの１以上に依拠する。また、決定部は、たとえば、第１の部分中間ピッチサイクルからいくつのサンプルを除くかまたは加えるかを示す開始部差数を決定するよう構成することができ、かつフレーム再構成部は、開始部差数に依拠して、第１の部分中間ピッチサイクルから１以上の第１のサンプルを除去するよう構成されるか、または第１の部分中間ピッチサイクルに１以上の第１のサンプルを加えるよう構成される。さらに、決定部は、たとえば、さらなる中間ピッチサイクルの前記１つから除くかまたは加えるサンプルの数を表すピッチサイクル差数を、さらなる中間ピッチサイクルの各々について決定するよう構成され得る。また、フレーム再構成部は、たとえば、前記ピッチサイクル差数に依拠して、さらなる中間ピッチサイクルの前記１つから１以上の第２のサンプルを除去するよう構成され得るか、または、さらなる中間ピッチサイクルの前記１つに１以上の第２のサンプルを加えるよう構成され得る。さらに、決定部は、たとえば、第２の部分中間ピッチサイクルから除くかまたは加えるサンプルの数を示す終了部差数を決定するよう構成することができ、かつフレーム再構成部は、終了部差数に依拠して、第２の部分中間ピッチサイクルから１以上の第３のサンプルを除去するよう構成される、かまたは第２の部分中間ピッチサイクルに１以上の第３のサンプルを加えるよう構成される。

【００７３】

ある実施形態によれば、フレーム再構成部は、たとえば、１以上の入手可能なピッチサイクルの前記１つに依拠して中間フレームを生成するよう構成され得る。また、決定部は、たとえば、中間フレームにより含まれる音声信号の１以上の低エネルギ信号部を決定するようにされてもよく、１以上の低エネルギ信号部の各々が、中間フレーム内の音声信号の第１の信号部であり、音声信号のエネルギが、中間フレームにより含まれる音声信号の第２の信号部におけるエネルギより低い。さらに、フレーム再構成部は、たとえば、再構成されたフレームを得るために、音声信号の１以上の低エネルギ信号部の１以上から１以上のサンプルを除去するか、または音声信号の１以上の低エネルギ信号部分の１以上へ１以上のサンプルを加えるよう構成され得る。

【００７４】

特定の実施形態において、フレーム再構成部は、たとえば、中間フレームを生成するよう構成されることが可能で、それにより中間フレームが１以上の再構成されたピッチサイクルを含み、１以上の再構成されたピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つに依拠するようになっている。また、決定部は、たとえば、１以上の

再構成ピッチサイクルの各々から除去するサンプルの数を決定するように構成され得る。さらに、決定部は、たとえば、１以上の低エネルギ信号部の各々について、前記低エネルギ信号部のサンプルの数が、１以上の再構成ピッチサイクルの１つから除去されるサンプル数に依拠するように、１以上の低エネルギ信号部の各々を決定するように構成することが可能で、前記低エネルギ信号部が、１以上の再構成ピッチサイクルの前記１つ内に位置する。

【００７５】

ある実施形態において、決定部は、たとえば、再構成フレームとして再構成されるべきフレームの音声信号の１以上のパルスの位置を決定するよう構成され得る。また、フレーム再構成部は、たとえば、音声信号の１以上のパルスの位置に依拠して再構成フレームを再構成するよう構成され得る。

【００７６】

ある実施形態によれば、決定部は、たとえば、再構成フレームとして再構成されるべきフレームの音声信号の２以上のパルスの位置を決定するよう構成することが可能で、Ｔ［０］は、再構成フレームとして再構成されるべきフレームの音声信号の２以上のパルスのうちの１つの位置であり、かつ決定部は、以下の式に従う音声信号の２以上のパルスのうちのさらなるパルスの位置（Ｔ［ｉ］）を決定するよう構成される。

【数３５】

$$T[i] = T[0] + iT_r$$

ここで、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、かつｉは整数である。

【００７７】

ある実施形態によれば、決定部は、たとえば、以下の式のように、再構成フレームとして再構成されるべきフレームの音声信号の最後のパルスのインデクスｋを決定するよう構成され得る。

【数３６】

$$k = \left\lfloor \frac{L - s - T[0]}{T_r} - 1 \right\rfloor$$

ここで、Ｌは、再構成フレームのサンプルの数を示し、ｓは、フレーム差値を示し、Ｔ［０］は、音声信号の最後のパルスとは異なる再構成フレームとして再構成されるべきフレームの音声信号のパルスの位置を示し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示す。

【００７８】

ある実施形態において、決定部は、たとえば、パラメータδを決定することにより再構成フレームとして再構成されるべきフレームを再構成するよう構成することが可能で、δは以下の式により規定される。

【数３７】

$$\delta = \frac{T_{ext} - T_p}{M}$$

ここで、再構成フレームとして再構成されるべきフレームは、Ｍ個のサブフレームを含み、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示し、T_{ext}は、再構成フレームとして再構成されるべきフレームの再構成されるべきピッチサイクルのうちの１つの長さを示す。

【００７９】
ある実施形態によれば、決定部は、たとえば、以下の式に基づき１以上の入手可能なピッチサイクルの前記１つの丸められた長さT_rを決定することにより再構成フレームを再構成するよう構成され得る。
【数３８】

$$T_r = \lfloor T_p + 0.5 \rfloor$$

ここで、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示す。
【００８０】
ある実施形態において、決定部は、たとえば以下の式を適用することにより再構成フレームを再構成するよう構成され得る。
【数３９】

$$s = \delta \frac{L}{T_r} \frac{M+1}{2} - L\left(1 - \frac{T_p}{T_r}\right)$$

ここで、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、再構成フレームとして再構成されるべきフレームは、Ｍ個のサブフレームを含み、再構成フレームとして再構成されるべきフレームは、Ｌ個のサンプルを含み、δが１以上の入手可能なピッチサイクルのうちの前記１つのサンプルの数と、再構成されるべき１以上のピッチサイクルの１つのサンプルの数との差を表す実数である。
【００８１】
また、音声信号を含むフレームを、再構成されたフレームとして再構成するための方法が提供され、前記再構成されたフレームが、１以上の入手可能なフレームと関連し、前記１以上の入手可能なフレームが、再構成されたフレームの１以上の先行フレームおよび再構成されたフレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームが、１以上の入手可能なピッチサイクルとして、１以上のピッチサイクルを含み、この方法は、以下のステップを含む。
・１以上の入手可能なピッチサイクルのうちの１つのサンプルの数と再構成されるべき第１のピッチサイクルのサンプルの数との差を示すサンプル数の差（$\Delta^p{}_0$；Δ_i；$\Delta^p{}_{k+1}$）を決定するステップ。
・サンプル数の差（$\Delta^p{}_0$；Δ_i；$\Delta^p{}_{k+1}$）および１以上の入手可能なピッチサイクルのうちの前記１つのサンプルに依拠して、第１の再構成ピッチサイクルとして再構成されるべき第１のピッチサイクルを再構成することにより再構成フレームを再構成するステップ。
【００８２】
再構成フレームの再構成が行われ、それにより再構成フレームが完全にまたは部分的に第１の再構成ピッチサイクルを含み、再構成フレームが完全にまたは部分的に第２の再構成ピッチサイクルを含み、かつ第１の再構成ピッチサイクルのサンプル数が第２の再構成ピッチサイクルのサンプル数とは異なるようになっている。
【００８３】
さらに、コンピュータまたは信号プロセッサ上で実行されて、上記の方法を実現するためのコンピュータプログラムが提供される。
【００８４】
また、推定されるピッチラグを決定するための装置が提供される。この装置は、複数のオリジナルピッチラグ値を受信するための入力インターフェースと、推定ピッチラグを推定するためのピッチラグ推定器とを含む。ピッチラグ推推定器は、複数のオリジナルピッ

チラグ値および複数の情報値に依拠して、推定ピッチラグを推定するよう構成され、複数のオリジナルピッチラグ値の各々のオリジナルピッチラグ値について、複数の情報値のうちの１つの情報値が、前記オリジナルピッチラグ値に割り当てられる。

【００８５】

実施形態によれば、ピッチラグ推定器は、たとえば複数のオリジナルピッチラグ値および複数の情報値としての複数のピッチゲイン値に依拠して、推定ピッチラグを推定するよう構成されることが可能で、複数のオリジナルピッチラグ値の各々のオリジナルピッチラグ値について、複数のピッチゲイン値のうちの１つのピッチゲイン値が、前記オリジナルピッチラグ値に割り当てられる。

【００８６】

特定の実施形態において、複数のピッチゲイン値の各々が、たとえば適応型コードブックゲインであり得る。

【００８７】

ある実施形態において、ピッチラグ推定器は、たとえば誤差関数を最小化することにより推定ピッチラグを推定するよう構成され得る。

【００８８】

ある実施形態によれば、ピッチラグ推定器は、たとえば以下の誤差関数を最小化することにより、２つのパラメータａ、ｂを決定して、推定ピッチラグを推定するよう構成されることが可能で、

【数４０】

$$err = \sum_{i=0}^{k} g_p(i) \cdot \left((a + b \cdot i) - P(i)\right)^2$$

ここで、ａは実数であり、ｂは実数であり、ｋは、ｋ≧２の整数であり、Ｐ（ｉ）は、ｉ番目のオリジナルピッチラグ値であり、g_p（ｉ）は、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目のピッチゲイン値である。

【００８９】

ある実施形態において、ピッチラグ推定器は、たとえば以下の誤差関数を最小化することにより２つのパラメータａ、ｂを決定して、推定ピッチラグを推定するよう構成されることが可能で、

【数４１】

$$err = \sum_{i=0}^{4} g_p(i) \cdot \left((a + b \cdot i) - P(i)\right)^2$$

ここで、ａは実数であり、ｂは実数であり、Ｐ（ｉ）は、ｉ番目のオリジナルピッチラグ値であり、g_p（ｉ）は、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目のピッチゲイン値である。

【００９０】

ある実施形態によれば、ピッチラグ推定器は、たとえばｐ＝ａ・ｉ＋ｂによって推定ピッチラグｐを決定するよう構成され得る。

【００９１】

ある実施形態において、ピッチラグ推定器は、たとえば複数のオリジナルピッチラグ値および複数の情報値としての複数の時間値に依拠して、推定ピッチラグを推定するよう構成されることが可能で、複数のオリジナルピッチラグ値の各々のオリジナルピッチラグ値について、複数の時間値のうちの１つの時間値が前記オリジナルピッチラグ値に割り当てられる。

【００９２】

ある実施形態によれば、ピッチラグ推定器は、たとえば誤差関数を最小化することによ

り推定ピッチラグを推定するよう構成され得る。

【００９３】

ある実施形態において、ピッチラグ推定器は、たとえば以下の誤差関数を最小化することにより２つのパラメータａ、ｂを決定して、推定ピッチラグを推定するよう構成されることが可能で、

【数４２】

$$err = \sum_{i=0}^{k} time_{passed}(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは実数であり、ｂは実数であり、ｋはｋ≧２の整数であり、かつｐ（ｉ）はｉ番目のオリジナルピッチラグ値であり、ｔｉｍｅ$_{passed}$（ｉ）は、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられたｉ番目の時間値である。

【００９４】

ある実施形態によれば、ピッチラグ推定器は、たとえば以下の誤差関数を最小化することにより２つのパラメータａ、ｂを決定して、推定ピッチラグを推定するよう構成されることが可能で、

【数４３】

$$err = \sum_{i=0}^{4} time_{passed}(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは実数であり、ｂは実数であり、ｐ（ｉ）はｉ番目のオリジナルピッチラグ値であり、ｔｉｍｅ$_{passed}$（ｉ）が、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目の時間値である。

【００９５】

ある実施形態において、ピッチラグ推定器は、ｐ＝ａ・ｉ＋ｂによって推定ピッチラグｐを決定するよう構成される。

【００９６】

また、推定ピッチラグを決定するための方法が提供される。この方法は、以下のステップを含む。

・複数のオリジナルピッチラグ値を受けるステップ

・推定ピッチラグを推定するステップ。

【００９７】

推定ピッチラグを推定するステップは、複数のオリジナルピッチラグ値および複数の情報値に依拠して行われ、複数のオリジナルピッチラグ値の各々のオリジナルピッチラグ値について、複数の情報値のうちの１つの情報値が、前記オリジナルピッチラグ値に割り当てられる。

【００９８】

さらに、コンピュータまたは信号プロセッサ上で実行されたときに、上記の方法を実現するためのコンピュータプログラムが提供される。

【００９９】

また、音声信号を含むフレームを再構成するためのシステムが提供される。このシステムは、上記および後述の実施形態の１つに従う推定ピッチラグを決定するための装置と、フレームを再構成するための装置とを含み、フレームを再構成するための装置が、推定ピッチラグに依拠してフレームを再構成するよう構成される。推定ピッチラグは、音声信号のピッチラグである。

【０１００】

ある実施形態において、再構成されたフレームが、たとえば、１以上の入手可能なフレームと関連し、前記１以上の入手可能なフレームが、再構成されたフレームの１以上の先

行フレームおよび再構成されたフレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームが、１以上の入手可能なピッチサイクルとして、１以上のピッチサイクルを含む。フレームを再構成するための装置は、たとえば、上記または後述の実施形態の１つに従ってフレームを再構成するための装置であってもよい。

【０１０１】

本発明は、先行技術が大きな欠点を有するとする所見に基づく。Ｇ．７１８（非特許文献８［ＩＴＵ０８ａ］を参照）およびＧ．７２９．１（非特許文献６［ＩＴＵ０６ｂ］を参照）の双方とも、フレーム損失の場合にピッチ外挿を用いる。これが必要なのは、フレーム損失時には、ピッチラグも失われるからである。Ｇ．７１８およびＧ．７２９．１によれば、ピッチは、最後の２つのフレームの間のピッチの展開を考慮に入れることにより外挿される。しかしながら、Ｇ．７１８およびＧ．７２９．１により再構成されるピッチラグは、あまり正確ではなく、たとえば、実際のピッチラグから大きく異なる再構成ピッチラグが得られる場合が多い。

【０１０２】

本発明の実施形態により、より正確なピッチラグ再構成が提供される。この目的で、Ｇ．７１８およびＧ．７２９．１とは対照的に、いくつかの実施形態では、ピッチ情報の信頼性に関する情報を考慮する。

【０１０３】

先行技術では、外挿が基礎とするピッチ情報は、最後に正確に受信された８つのピッチラグを含み、これらについては、符号化モードは「無声」とは異なっていた。しかしながら、先行技術では、低いピッチゲイン（低い予測ゲインに対応する）により示される有声特性が非常に弱いかもしれない。先行技術では、外挿が、異なるピッチゲインを有するピッチラグに基づく場合、外挿が、妥当な結果をもたらさないか、または、全くうまくいかず、再び単純なピッチラグ繰り返しのアプローチに戻ることになる。

【０１０４】

実施形態は、先行技術のこれらの欠点の原因が、エンコーダ側で、適応型コードブックの符号化ゲインを最大化するため、ピッチゲインを最大化することに関してピッチラグを選択するが、音声特性が弱い場合には、音声信号における雑音でピッチラグ推定が不正確になるため、ピッチラグが基本周波数を正確に表示しない可能性があると言う点にあるとする所見に基づく。

【０１０５】

したがって、実施形態によれば、封じ込め中に、ピッチラグ外挿の適用は、この外挿について使用した前に受信したラグの信頼性に依拠して重み付けが行われる。

【０１０６】

いくつかの実施形態によれば、過去の適応型コードブックゲイン（ピッチゲイン）を、信頼性の尺度として採用し得る。

【０１０７】

本発明の他のいくつかの実施形態によれば、ピッチラグがどこまで過去に受信されたかによる重み付けが信頼性の尺度として使用される。たとえば、より最近のラグには高い重み付けがされ、より後に受けられたラグにはより低い重み付けがされる。

【０１０８】

実施形態によれば、重み付けピッチ予測のコンセプトが提供される。先行技術とは対照的に、本発明の実施形態により提供されるピッチ予測は、基礎とするピッチラグの各々についての信頼性の尺度を使用し、予測結果をより有効で安定したものにする。特に、ピッチゲインを信頼性の指標として使用することができる。代替的にまたは付加的に、いくつかの実施形態によれば、たとえば、ピッチラグを正しく受け取った後に経過した時間を、指標として使用することができる。

【０１０９】

パルス再同期化については、本発明は、声門パルス再同期化に関する先行技術の欠点の１つは、ピッチ外挿が封じ込められたフレームにおいて構成すべきパルス（ピッチサイク

ル）の数を考慮しない点にあるとする所見に基づく。

【０１１０】

先行技術によれば、ピッチにおける変化がサブフレームの境界でのみ予測されるようにピッチ外挿が行われる。

【０１１１】

実施形態によれば、声門パルス再同期化を実行する際は、連続するピッチ変化とは異なるピッチ変化を考慮に入れることができる。

【０１１２】

本発明の実施形態は、Ｇ．７１８およびＧ．７２９．１が以下の欠点を有するとする所見に基づく。

【０１１３】

まず、先行技術では、ｄを計算する際に、フレーム内に整数個のピッチサイクルが存在すると仮定する。ｄが、封じ込めフレームにおける最後のパルスの場所を規定するので、非整数のピッチサイクルがフレーム内に存在する場合には、最後のパルスの位置は、正確にならない。これを、図６および図７に示す。図６は、サンプル除去前の音声信号を示す。図７は、サンプル除去後の音声信号を示す。さらに、ｄを計算するために先行技術が採用するアルゴリズムは非効率である。

【０１１４】

また、先行技術の計算では、励振の構成された周期的部分において、パルス数Ｎが必要である。これが、不必要な計算の複雑さを増大させる。

【０１１５】

さらに、先行技術においては、励振の構成された周期的部分におけるパルス数Ｎの計算は、第１のパルスの場所を考慮に入れない。

【０１１６】

図４および図５において提示される信号は、長さT_cの同じピッチ周期を有する。

【０１１７】

図４は、フレーム内に３つのパルスを有する音声信号を示す。

【０１１８】

対照的に、図５は、フレーム内に２つのパルスのみを有する音声信号を示す。

【０１１９】

図４および図５が示すこれらの例は、パルスの数が第１のパルスの位置に依拠することを示す。

【０１２０】

また、先行技術によれば、Ｎが、後続のフレームにおける第１のパルスを含むと規定されても、励振の構成された周期的部分におけるＮ番目のパルスの場所であるＴ［Ｎ-１］が、フレーム長さの範囲内にあるかどうかをチェックする。

【０１２１】

さらに、先行技術によれば、最初のパルスの前および最後のパルスの後には、サンプルが加えられたり除去されたりしない。本発明の実施形態は、このことが、第１のフルピッチサイクルの長さにおける突然の変化が起こる可能性があるという欠点につながり、これが、また、ピッチラグが減少している場合でさえ、最後のパルス後のピッチサイクルの長さが、最後のパルスの前の最後のフルピッチサイクルの長さより大きくなり得るという欠点につながるとする所見に基づく（図６および図７参照）。

【０１２２】

実施形態は、パルスＴ［ｋ］＝Ｐ‐ｄｉｆｆおよびＴ［ｎ］＝Ｐ-ｄが以下の場合等しくないという所見に基づく。

・ｄ＞［T_c／２］の場合。この場合、ｄｉｆｆ＝T_c－ｄであり、かつ除去されたサンプルの数がｄではなくｄｉｆｆになる。

・Ｔ［ｋ］が、未来のフレーム内にあり、かつｄサンプルを除いて初めて現在のフレームに移動する場合。

・Ｔ［ｎ］が、－ｄサンプル（ｄ＜０）を加えた後に未来のフレームに移動する場合。

【０１２３】

これが、封じ込められたフレームにおけるパルスの間違った位置につながる。

【０１２４】

また、実施形態は、先行技術において、ｄの最大値が符号化されたピッチラグの最小許容値に制限されるという所見に基づく。これは、他の問題の発生を制限する制約だが、ピッチにおいて可能な変化も制限し、かつパルス再同期化も制限する。

【０１２５】

さらに、実施形態は、先行技術において、周期的部分が整数ピッチラグを用いて構成され、かつこれが高調波の周波数シフトおよび一定のピッチでの音信号の封じ込めにおける大きな劣化を作り出すとする所見に基づく。この劣化については、丸められたピッチラグを用いるときに再同期化される音声信号の時間－周波数表現を示す図８に見ることができる。

【０１２６】

また、実施形態は、先行技術の問題の殆どが、ｄサンプルが除去される図６および図７の例が示すような状況で発生するという所見に基づく。ここでは、問題をより簡単に可視化するため、ｄの最大値についての制約はないと考える。問題は、ｄに限度があるがあまり明確に可視化されない場合にも生じる。連続して増加するピッチではなく、ピッチが突然増大した後に突然減少することも考えられる。実施形態は、これが、最後のパルスの前後でサンプルが除去されないこと、間接的にはパルスＴ［２］が、ｄサンプル除去の後のフレーム内で移動することを考慮に入れないことにより起こるとする所見に基づく。この例では、Ｎの計算の誤りも発生する。

【０１２７】

実施形態によれば、改善されたパルス再同期化のコンセプトが提供される。実施形態は、音声を含むモノラルの信号の改善された封じ込めを提供し、これは、標準Ｇ．７１８（非特許文献８［ＩＴＵ０８ａ］を参照）およびＧ．７２９．１（非特許文献６［ＩＴＵ０６ｂ］を参照）に記載の既存技術に比べて有利である。本件の実施形態は、一定のピッチの信号およびピッチが変化する信号両方に適している。

【０１２８】

とりわけ、実施形態によれば３つの技術が提供される。

【０１２９】

ある実施形態が提供する第１の技術によれば、Ｇ．７１８およびＧ．７２９．１とは対照的に、Ｎで表す構成された周期的部分におけるパルスの数の計算において、第１のパルスの場所を考慮に入れる、パルスについてのサーチコンセプトが提供される。

【０１３０】

他の実施形態により提供される第２の技術によれば、Ｇ．７１８およびＧ．７２９．１とは対照的に、Ｎで示す、構成された周期的部分におけるパルスの数を必要とせず、第１のパルスの場所を考慮に入れ、かつｋで示す封じ込めフレームにおける最後のパルスインデクスを直接的に計算するパルスをサーチするためのアルゴリズムが提供される。

【０１３１】

他の実施形態により提供される第３の技術によれば、パルスサーチは不要である。この第３の技術によれば、周期的部分の構成とサンプルの除去または付加を組み合わせることにより、以前の技術よりも複雑さが抑えられる。

【０１３２】

付加的または代替的に、いくつかの実施形態は、上記の技術ならびにＧ．７１８およびＧ．７２９．１の技術について以下の変更を提供する。

・ピッチラグの小数部分は、たとえば、一定ピッチの信号について周期的部分を構成するために使用できる。

・封じ込めフレームにおける最後のパルスの予測される場所のオフセットが、たとえば、フレーム内の非整数のピッチサイクルについて計算され得る。

・たとえば、最初のパルスの前と最後のパルスの後にもサンプルを加えたり除いたりできる。
・たとえば、パルスが１つしかない場合にも、サンプルを加えたり除いたりできる。
・除くかまたは加えるべきサンプルの数は、たとえば、ピッチにおける予測線形変化に従って線形に変更できる。

【０１３３】
以下に、図面を参照して本発明の実施形態についてより詳細に説明する。

【図面の簡単な説明】

【０１３４】
【図１】図１は、実施形態による推定ピッチラグを決定するための装置を示す図である。
【図２Ａ】図２Ａは、実施形態による再構成フレームとして音声信号を含むフレームを再構成するための装置を示す図である。
【図２Ｂ】図２Ｂは、複数のパルスを含む音声信号を示す図である。
【図２Ｃ】図２Ｃは、実施形態による音声信号を含むフレームを再構成するためのシステムを示す図である。
【図３】図３は、音声信号の構成された周期的部分を示す図である。
【図４】図４は、フレーム内に３つのパルスを有する音声信号を示す図である。
【図５】図５は、フレーム内に２つのパルスを有する音声信号を示す図である。
【図６】図６は、サンプルの除去前の音声信号を示す図である。
【図７】図７は、サンプルの除去後の図６の音声信号を示す図である。
【図８】図８は、丸められたピッチラグを用いて再同期化された音声信号の時間－周波数表現を示す図である。
【図９】図９は、小数部分を有する非丸めピッチラグを用いて再同期化された音声信号の時間‐周波数表現を示す図である。
【図１０】図１０は、ピッチラグが前提技術のコンセプトを採用して再構成されるピッチラグ図を示す図である。
【図１１】図１１は、実施形態によりピッチラグが再構成されるピッチラグ図を示す図である。
【図１２】図１２は、サンプルを除去する前の音声信号を示す図である。
【図１３】図１３は、Δ_0からΔ_3を付加的に示す図１２の音声信号を示す図である。

【発明を実施するための形態】

【０１３５】
図１は、実施形態による推定ピッチラグを決定するための装置を示す。この装置は、複数のオリジナルピッチラグ値を受けるための入力インターフェース１１０と、推定ピッチラグを推定するためのピッチラグ推定器１２０とを含む。ピッチラグ推定器１２０は、複数のオリジナルピッチラグ値および複数の情報値に依拠して推定ピッチラグを推定するよう構成され、複数のオリジナルピッチラグ値の各オリジナルピッチラグ値について、複数の情報値のうちの１つの情報値が前記オリジナルピッチラグ値に割り当てられる。

【０１３６】
実施形態によれば、ピッチラグ推定器１２０は、たとえば、複数のオリジナルピッチラグ値と、複数の情報値としての複数のピッチゲイン値とに依拠して推定ピッチラグを推定するよう構成することが可能で、複数のオリジナルピッチラグ値の各オリジナルピッチラグ値について、複数のピッチゲイン値のうちの１つのピッチゲイン値が前記オリジナルピッチラグ値に割り当てられる。

【０１３７】
特定の実施形態において、複数のピッチゲイン値の各々は、たとえば、適応型コードブックゲインでもよい。

【０１３８】
ある実施形態において、ピッチラグ推定器１２０は、たとえば、誤差関数を最小化することにより推定ピッチラグを推定するよう構成され得る。

【０１３９】
ある実施形態によれば、ピッチラグ推定器１２０は、たとえば、以下の誤差関数を最小化して、２つのパラメータａ，ｂを決定することにより推定ピッチラグを推定するよう構成され得る。
【数４４】

$$err = \sum_{i=0}^{k} g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは実数であり、ｂは実数であり、ｋはｋ≧２の整数であり、Ｐ（ｉ）はｉ番目のオリジナルピッチラグ値であり、g_p（ｉ）はｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目のピッチゲイン値である。
【０１４０】
ある実施形態において、ピッチラグ推定器１２０は、たとえば、以下の誤差関数を最小化して、２つのパラメータａ、ｂを決定することにより、推定ピッチラグを推定するよう構成され得る。
【数４５】

$$err = \sum_{i=0}^{4} g_p(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは実数であり、ｂは実数であり、Ｐ（ｉ）はｉ番目のオリジナルピッチラグ値であり、g_p（ｉ）はｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目のピッチゲイン値である。
【０１４１】
ある実施形態によれば、ピッチラグ推定器１２０は、たとえば、ｐ＝ａ・ｉ+ｂに従って推定ピッチラグｐを決定するよう構成され得る。
【０１４２】
ある実施形態において、ピッチラグ推定器１２０は、たとえば、複数のオリジナルピッチラグ値と、複数の情報値としての複数の時間値とに依拠して推定ピッチラグを推定するよう構成されることが可能で、複数のオリジナルピッチラグ値のうちの各オリジナルピッチラグ値について、複数の時間値のうちの１つの時間値が前記オリジナルピッチラグ値に割り当てられる。
【０１４３】
ある実施形態によれば、ピッチラグ推定器１２０は、たとえば、誤差関数を最小化することにより推定ピッチラグを推定するよう構成され得る。
【０１４４】
ある実施形態において、ピッチラグ推定器１２０は、たとえば、以下の誤差関数を最小化して、２つのパラメータａ、ｂを決定することにより、推定ピッチラグを推定するよう構成され得る。
【数４６】

$$err = \sum_{i=0}^{k} time_{passed}(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは実数であり、ｂは実数であり、ｋは、ｋ≧２の整数であり、かつＰ（ｉ）はｉ番目のオリジナルピッチラグ値であり、$time_{passed}$（ｉ）は、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目の時間値である。

【０１４５】
ある実施形態によれば、ピッチラグ推定器１２０は、たとえば、以下の誤差関数を最小化して、２つのパラメータａ、ｂを決定することにより、推定ピッチラグを推定するよう構成され得る。
【数４７】

$$err = \sum_{i=0}^{4} time_{passed}(i) \cdot ((a + b \cdot i) - P(i))^2$$

ここで、ａは、実数であり、ｂは実数であり、Ｐ（ｉ）は、ｉ番目のオリジナルピッチラグ値であり、ｔｉｍｅ$_{passed}$（ｉ）は、ｉ番目のピッチラグ値Ｐ（ｉ）に割り当てられるｉ番目の時間値である。
【０１４６】
ある実施形態において、ピッチラグ推定器１２０は、ｐ＝ａ・ｉ+ｂに従って推定ピッチラグｐを決定するよう構成される。
【０１４７】
以下に、重み付けピッチ予測を行う実施形態を、式（２０）から式（２４ｂ）を参照して記載する。
【０１４８】
まず、ピッチゲインによる重み付けを採用する重み付けピッチ予測の実施形態について式（２０）から式（２２ｃ）を参照して説明する。これらの実施形態のいくつかは、先行技術の欠点を克服するため、ピッチラグをピッチゲインで重み付けしてピッチ予測を行う。
【０１４９】
いくつかの実施形態において、ピッチゲインは、標準Ｇ．７２９において規定される適応型コードブックゲインg_pが可能である（非特許文献１０［ＩＴＵ１２］、特に３．７．３章、より詳細には式（４３）を参照）。Ｇ．７２９においては、適応型コードブックゲインは、以下に従って決定される。
【数４８】

$$g_p = \frac{\sum_{n=0}^{39} x(n)\,y(n)}{\sum_{n=0}^{39} y(n)\,y(n)} \quad (0 \leqq g_p \leqq 1.2\text{を境界とする})$$

ここで、ｘ（ｎ）は、ターゲット信号であり、かつｙ（ｎ）は、以下のとおり、ｖ（ｎ）をｈ（ｎ）と畳み込むことにより得られる。
【数４９】

$$y(n) = \sum_{i=0}^{n} v(i)h(n-i) \quad n = 0,\ldots,39$$

【０１５０】
ここで、ｖ（ｎ）は、適応型コードブックベクトルであり、ｙ（ｎ）は、フィルタ化された適応型コードブックベクトルであり、かつｈ（ｎ－ｉ）は、Ｇ．７２９に規定される重み付合成フィルタのインパルス応答である（非特許文献１０［ＩＴＵ１２］を参照）。
【０１５１】
同様に、いくつかの実施形態において、ピッチゲインは、標準Ｇ．７１８において規定される適応型コードブックゲインg_pであることが可能である（非特許文献８［ＩＴＵ０

８ａ］、特に６．８．４．１．４．１章、より詳細には式（１７０）を参照）。Ｇ．７１８においては、適応型コードブックゲインが以下のとおり決定される。

【数５０】

$$C_{CL}=\frac{\sum_{n=0}^{63}x(n)\,y_k(n)}{\sum_{n=0}^{63}y_k(n)\,y_k(n)}$$

ここで、ｘ（ｎ）はターゲット信号であり、かつy_k（ｎ）は、遅延ｋでの過去のフィルタ化された励振である。

【０１５２】

たとえば、定義について、y_k（ｎ）がいかに規定できるかは、非特許文献８（［ＩＴＵ０８ａ］）、６．８.４．１.４．１章、式（１７１）を参照。

【０１５３】

同様に、いくつかの実施形態では、ピッチゲインは、ＡＭＲ標準で規定される適応型コードブックゲインg_p（非特許文献３［３ＧＰ１２ｂ］を参照）が可能で、ピッチゲインとしての適応型コードブックゲインg_pは、以下のとおり規定される。

【数５１】

$$g_p=\frac{\sum_{n=0}^{63}x(n)y(n)}{\sum_{n=0}^{63}y(n)y(n)}\qquad（0\leqq g_p\leqq 1．2を境界とする）$$

ここで、ｙ（ｎ）は、フィルタ化された適応型コードブックベクトルである。

【０１５４】

いくつかの実施形態において、ピッチラグは、たとえば、ピッチ予測を行う前にピッチゲインで重み付けすることが可能である。

【０１５５】

この目的で、ある実施形態によれば、たとえば、ピッチラグと同じサブフレームで取られたピッチゲインを保持する、長さ８の第２のバッファを導入してもよい。ある実施形態では、バッファは、ピッチラグの更新と全く同じルールを使用して更新され得る。１つの可能な実現例は、そのフレームに誤差がないかまたは誤差がありがちかにかかわらず、各フレームの終りに両方のバッファ（最後の８つのサブフレームのピッチラグとピッチゲインを保持）を更新することである。

【０１５６】

先行技術から２つの異なる予測戦略が知られており、重み付けピッチ予測を使用するためにこれらを強化することができる。

【０１５７】

いくつかの実施形態は、Ｇ．７１８標準の予測戦略に対し大きな発明的改善をもたらす。Ｇ．７１８において、パケットが失われる場合において、関連するピッチゲインが高い場合、高いファクタでピッチラグを重み付けし、かつ関連するピッチゲインが低い場合には、低いファクタでこれを重み付けするために、バッファが要素ごとに相互に乗算され得る。その後、Ｇ．７１８に従って、ピッチ予測が通常通り行われる（Ｇ．７１８に関する詳細については、非特許文献８［ＩＴＵ０８ａ、セクション７．１１.１．３］を参照）。

【０１５８】

いくつかの実施形態は、Ｇ．７２９．１標準の予測戦略に対し大きな発明的改善をもたらす。ピッチを予測するためのＧ．７２９.１において使用されるアルゴリズム（Ｇ．７

２９．１に関する詳細については、非特許文献６［ＩＴＵ０６ｂ］を参照）が、重み付け予測を用いるために、実施形態に従って修正される。

【０１５９】

いくつかの実施形態によれば、目標は、以下の誤差関数を最小化することである。

【数５２】

$$err = \sum_{i=0}^{4} g_p(i) \cdot ((a + b \cdot i) - P(i))^2 \qquad (20)$$

ここで、g_p（ｉ）は、過去のサブフレームからのピッチゲインを保持し、かつ、Ｐ（ｉ）は、対応のピッチラグを保持する。

【０１６０】

本発明の式（２０）では、g_p（ｉ）が、重み付けファクタを表す。上記の例では、各g_p（ｉ）が、過去のサブフレームの１つからのピッチゲインを表す。

【０１６１】

以下に、実施形態による等式を記述するが、これらは、ａ+ｉ・ｂ（ｉが予測対象のサブフレームのサブフレーム番号）によってピッチラグを予測するために使用できるファクタａおよびｂを生成する方法を記述する。

【０１６２】

たとえば、最後の５つのサブフレームＰ（０）、．．．、Ｐ（４）に関する予測に基づき第１の予測サブフレームを得るために、予測ピッチ値Ｐ（５）は、以下のようになると考えられる。

【数５３】

Ｐ（５）＝ａ＋５・ｂ

【０１６３】

係数ａおよびｂを生成するために、たとえば、誤差関数を生成し（導き）、かつゼロに設定することができる。

【数５４】

$$\frac{\delta err}{\delta a} = 0 \quad and \quad \frac{\delta err}{\delta b} = 0 \qquad (21a)$$

【０１６４】

先行技術は、実施形態により提供される本発明の重み付けを採用することについて開示してない。特に、先行技術は、重み付けファクタg_p（ｉ）を採用していない。

【０１６５】

このように、重み付けファクタg_p（ｉ）を採用しない先行技術においては、誤差関数を生成して、誤差関数の導関数をゼロに設定すると、以下のようになると考えられる。

【数５５】

$$a = \frac{3\sum_{i=0}^{4} P(i) - \sum_{i=0}^{4} i \cdot P(i)}{5} \quad \text{and} \quad b = \frac{\sum_{i=0}^{4} i \cdot P(i) - 2\sum_{i=0}^{4} P(i)}{10} \qquad (21b)$$

（非特許文献６［ＩＴＵ０６ｂ、７．６．５を参照］）。

【０１６６】
対照的に、実施形態の重み付け予測アプローチ、たとえば、重み付けファクタg_p（ｉ）での式（２０）の重み付け予測アプローチを用いれば、ａおよびｂは、以下のようになる。
【数５６】

$$a = -\frac{A+B+C+D+E}{K} \qquad (22a)$$

$$b = +\frac{F+G+H+I+J}{K} \qquad (22b)$$

【０１６７】
特定の実施形態によれば、Ａ、Ｃ、Ｄ；Ｅ、Ｆ、Ｇ、Ｈ、Ｉ、ＪおよびＫは、たとえば以下の値を有し得る。
【数５７】

$$\begin{aligned}
A &= (3g_{p_3}+4g_{p_2}+3g_{p_1})g_{p_4}\cdot P(4)\\
B &= ((2g_{p_2}+2g_{p_1})g_{p_3}-4g_{p_3}g_{p_4})\cdot P(3)\\
C &= (-8g_{p_2}g_{p_4}-3g_{p_2}g_{p_3}+g_{p_1}g_{p_2})\cdot P(2)\\
D &= (-12g_{p_1}g_{p_4}-6g_{p_1}g_{p_3}-2g_{p_1}g_{p_2})\cdot P(1)\\
E &= (-16g_{p_0}g_{p_4}-9g_{p_0}g_{p_3}-4g_{p_0}g_{p_2}-g_{p_0}g_{p_1})\cdot P(0)\\
F &= (g_{p_3}+2g_{p_2}+3g_{p_1}+4g_{p_0})g_{p_4}\cdot P(4)\\
G &= ((g_{p_2}+2g_{p_1}+3g_{p_0})g_{p_3}-g_{p_3}g_{p_4})\cdot P(3)\\
H &= (-2g_{p_2}g_{p_4}-g_{p_2}g_{p_3}+(g_{p_1}+2g_{p_0})g_{p_2})\cdot P(2)\\
I &= (-3g_{p_1}g_{p_4}-2g_{p_1}g_{p_3}-g_{p_1}g_{p_2}+g_{p_0}g_{p_1})\cdot P(1)\\
J &= (-4g_{p_0}g_{p_4}-3g_{p_0}g_{p_3}-2g_{p_0}g_{p_2}-g_{p_0}g_{p_1})\cdot P(0)\\
K &= (g_{p_3}+4g_{p_2}+9g_{p_1}+16g_{p_0})g_{p_4}+(g_{p_2}+4g_{p_1}+9g_{p_0})g_{p_3}+(g_{p_1}+4g_{p_0})g_{p_2}+g_{p_0}g_{p_1}
\end{aligned} \qquad (22c)$$

【０１６８】
図１０および図１１は、提案されるピッチ外挿のより優れた性能を示す。
【０１６９】
ここで、図１０は、ピッチラグが前提技術のコンセプトを採用して再構成される場合のピッチラグ図を示す。対照的に、図１１は、ピッチラグが実施形態に従って再構成される場合のピッチラグ図を示す。
【０１７０】
詳細には、図１０は、先行技術の標準Ｇ．７１８およびＧ７２９．１の性能を示し、図１１は、実施形態により提供されるコンセプトの性能を示す。
【０１７１】
横軸がサブフレーム番号を表す。実線１０１０が、ビットストリームに埋め込まれ、かつ、グレイのセグメント１０３０の領域で失われるエンコーダピッチラグを示す。左側の縦軸は、ピッチラグ軸を表す。右側の縦軸は、ピッチゲイン軸を表す。実線１０１０は、ピッチラグを示し、破線１０２１、１０２２、１０２３はピッチゲインを示す。
【０１７２】
グレイの矩形１０３０は、フレーム損失を示す。グレイのセグメント１０３０の領域で

生じたフレーム損失のため、この領域におけるピッチラグおよびピッチゲインについての情報は、デコーダ側で入手できず、再構成する必要がある。

【０１７３】

図１０において、Ｇ．７１８標準を使用して封じ込められるピッチラグは、一点鎖線部１０１１により示される。Ｇ．７２９．１標準を使用して封じ込められるピッチラグは、実線部１０１２により示される。提供されるピッチ予測（図１１、実線部１０１３）を使用することは、本質的に、失われたエンコーダピッチラグに対応し、かつ、Ｇ．７１８およびＧ７２９．１の技術により有利であることがはっきりわかる。

【０１７４】

以下では、経過時間に依拠する重み付けを採用する実施形態について、式（２３ａ）から式（２４ｂ）を参照して説明する。

【０１７５】

先行技術の欠点を克服するため、いくつかの実施形態は、ピッチ予測を行う前に、ピッチラグに対し時間重み付けを適用する。時間重み付けの適用は、以下の誤差関数を最小化することにより実行され得る。

【数５８】

$$err = \sum_{i=0}^{4} time_{passed}(i) \cdot ((a + b \cdot i) - P(i))^2 \qquad (23a)$$

ここで、ｔｉｍｅ$_{passed}$（ｉ）は、ピッチラグを正しく受信した後に経過した時間の量の逆数を表し、かつ、Ｐ（ｉ）は、対応するピッチラグを保持する。

【０１７６】

いくつかの実施形態は、たとえば、より最近のラグに対して高い重みを付け、より以前に受信されたラグに対しては低い重みを付け得る。

【０１７７】

次いで、いくつかの実施形態によれば、式（２１ａ）を採用して、ａおよびｂを生成することができる。

【０１７８】

第１の予測されたサブフレームを得るため、いくつかの実施形態では、たとえば、最後の５つのサブフレームＰ（０）、．．．Ｐ（４）に基づいて予測を行い得る。次いで、たとえば、予測ピッチ値Ｐ（５）は、以下のとおり得ることができる。

【数５９】

Ｐ（５）＝ａ＋５・ｂ　　（２３ｂ）

【０１７９】

たとえば、以下のとおりであれば、

【数６０】

$$time_{passed} = [1/5\ 1/4\ 1/3\ 1/2\ 1]$$

（サブフレーム遅延に従う時間重み付け）、以下のようになると考えられる。

【数６１】

$$a = \frac{-3.5833 \cdot P(4) + 1.4167 \cdot P(3) + 3.0833 \cdot P(2) + 3.9167 \cdot P(1) + 4.4167 \cdot P(0)}{9.2500} \quad (24a)$$

$$b = \frac{+2.7167 \cdot P(4) + 0.2167 \cdot P(3) - 0.6167 \cdot P(2) - 1.0333 \cdot P(1) - 1.2833 \cdot P(0)}{9.2500} \quad (24b)$$

【０１８０】
以下では、パルス再同期化を提供する実施形態を説明する。
【０１８１】
図２ａは、実施形態に従う再構成されたフレームとして音声信号を含むフレームを再構成するための装置を示す。前記再構成されたフレームは、１以上の入手可能なフレームに関連し、前記１以上の入手可能なフレームは、再構成されたフレームの１以上の先行フレームおよび再構成されたフレームの１以上の後続のフレームのうちの少なくとも１つであり、１以上の入手可能なフレームが１以上の入手可能なピッチサイクルとして１以上のピッチサイクルを含む。
【０１８２】
装置は、１以上の入手可能なピッチサイクルのうちの１つのサンプルの数と、再構成されるべき第１のピッチサイクルのサンプルの数との差を示すサンプル数の差（Δ^{p}_{0}；Δ_{i}；Δ^{p}_{k+1}）を決定するための決定部２１０を含む。
【０１８３】
また、装置は、サンプル数の差（Δ^{p}_{0}；Δ_{i}；Δ^{p}_{k+1}）および１以上の入手可能なピッチサイクルのうちの前記１つのサンプルに依拠して第１の再構成ピッチサイクルとして再構成されるべき第１のピッチサイクルを再構成することにより再構成フレームを再構成するためのフレーム再構成部を含む。
【０１８４】
フレーム再構成部２２０は、再構成フレームを再構成するよう構成され、それにより再構成されたフレームが、完全にまたは部分的に第１の再構成されたピッチサイクルを含み、再構成されたフレームが、完全にまたは部分的に第２の再構成されたピッチサイクルを含み、かつ第１の再構成されたピッチサイクルのサンプル数が、第２の再構成されたピッチサイクルのサンプル数とは異なるようになっている。
【０１８５】
ピッチサイクルの再構成は、再構成するピッチサイクルのサンプルの一部または全部を再構成することにより行われる。再構成されるべきピッチサイクルが、失われたフレームに完全に含まれる場合には、たとえば、ピッチサイクルのサンプルのすべてを、再構成する必要があるかもしれない。再構成されるべきピッチサイクルが、一部のみ失われたフレームにより含まれ、かつ他のフレームに含まれる場合など、ピッチサイクルのサンプルのいくつかが入手可能な場合、ピッチサイクルを再構成するために、失われたフレームにより含まれるピッチサイクルのサンプルを再構成するだけで十分かもしれない。
【０１８６】
図２ｂは、図２ａの装置の機能性を示す。図２ｂは、特に、パルス２１１、２１２、２１３、２１４、２１５、２１６および２１７を含む音声信号２２２を示す。
【０１８７】
音声信号２２２の第１の部分は、フレームｎ－１により含まれる。音声信号２２２の第２の部分は、フレームｎにより含まれる。音声信号２２２の第３の部分は、フレームｎ＋１により含まれる。

【０１８８】
図２ｂにおいて、フレームｎ－１は、フレームｎに先行し、かつフレームｎ＋１は、フレームｎの後に続く。これは、フレームｎ－１がフレームｎの音声信号の部分に比べて時間的により早く生じた音声信号の部分を含み、かつフレームｎ＋１が、フレームｎの音声信号の部分に比べて時間的により遅く生じた音声信号の部分を含むことを意味する。
【０１８９】
図２ｂの例では、フレームｎが失われたか、または破損されていると仮定するので、フレームｎに先行するフレーム（「先行フレーム」）およびフレームｎに後続するフレーム（「後続フレーム」）のみが入手可能である（「入手可能フレーム」）。
【０１９０】
たとえば、ピッチサイクルを以下のように規定することができる。ピッチサイクルは、音声信号において、パルス２１１、２１２、２１３他のうちの１つで開始され、直後に続くパルスで終了する。たとえば、パルス２１１および２１２がピッチサイクル２０１を規定する。パルス２１２および２１３がピッチサイクル２０２を規定する。パルス２１３および２１４は、ピッチサイクル２０３を規定する等である。
【０１９１】
ピッチサイクルの他の開始および終了点を採用する、当業者に周知の他のピッチサイクルの定義についても、代替的に考慮してもよい。
【０１９２】
図２ｂの例では、フレームｎは、受信部で入手可能ではないか破損されている。したがって、受信部は、フレームｎ－１のパルス２１１および２１２ならびにピッチサイクル２０１について認識する。さらに、受信部は、フレームｎ＋１のパルス２１６および２１７ならびにピッチサイクル２０６についても認識する。しかしながら、パルス２１３、２１４および２１５を含み、ピッチサイクル２０３および２０４を完全に含み、かつ、ピッチサイクル２０４および２０５を部分的に含むフレームｎを再構成する必要がある。
【０１９３】
いくつかの実施形態によれば、フレームｎは、入手可能フレーム（たとえば、先行フレームｎ－１または後続フレームｎ＋１）の１以上のピッチサイクル（「入手可能ピッチサイクル」）のサンプルに依拠して再構成され得る。たとえば、フレームｎ－１のピッチサイクル２０１のサンプルは、失われたか、または破損されたフレームのサンプルを再構成するために周期的にに繰り返しコピーされ得る。ピッチサイクルのサンプルを周期的に繰り返しコピーすることで、ピッチサイクル自体がコピーされ、たとえばピッチサイクルがｃの場合、以下のようになる。
【０１９４】
【数６２】

サンプル（ｘ＋ｉ・ｃ）＝サンプル（ｘ）（ｉは整数）

【０１９５】
実施形態においては、フレームｎ－１の終りからのサンプルがコピーされる。ｎ‐１番目のフレームのコピーされる部分の長さが、ピッチサイクル２０１の長さに等しい（またはほとんど等しい）。しかしながら、２０１および２０２双方からのサンプルがコピーに使用される。これは、ｎ－１番目のフレームにパルスが１つしかない場合には、特に慎重に考慮する必要がある。
【０１９６】
いくつかの実施形態においては、コピーされたサンプルは修正される。
【０１９７】
本発明は、また、失われたフレーム（ｎ）により（完全にまたは部分的に）含まれるピッチサイクル（ピッチサイクル２０２、２０３、２０４および２０５）のサイズが、コピ

ーされた入手可能なピッチサイクル（ここでは、ピッチサイクル２０１）のサイズと異なる場合には、ピッチサイクルのサンプルを周期的に繰り返しコピーすることにより、失われたフレームｎのパルス２１３、２１４および２１５が間違った位置に移動するという所見に基づく。

【０１９８】

たとえば、図２ｂでは、ピッチサイクル２０１とピッチサイクル２０２との差は、Δ_1で示され、ピッチサイクル２０１とピッチサイクル２０３との差は、Δ_2で示され、ピッチサイクル２０１とピッチサイクル２０４との差は、Δ_3で示され、かつピッチサイクル２０１とピッチサイクル２０５との差は、Δ_4で示される。

【０１９９】

図２ｂにおいて、フレームｎ－１のピッチサイクル２０１が、ピッチサイクル２０６よりかなり大きいことがわかる。また、フレームｎに（一部または完全に）含まれるピッチサイクル２０２、２０３、２０４および２０５は、各々ピッチサイクル２０１より小さく、かつ、ピッチサイクル２０６より大きい。さらに、大きいピッチサイクル２０１により近いピッチサイクル（たとえば、ピッチサイクル２０２）は、小さいピッチサイクル２０６により近いピッチサイクル（たとえば、ピッチサイクル２０５）より大きい。

【０２００】

本発明のこれらの所見に基づいて、実施形態によれば、フレーム再構成部２２０は、第１の再構成されたピッチサイクルのサンプル数が、再構成されたフレームに部分的にまたは完全に含まれる第２の再構成されたピッチサイクルのサンプル数と異なるように、再構成フレームを再構成するよう構成される。

【０２０１】

たとえば、いくつかの実施形態によれば、フレームの再構成は、１以上の入手可能なピッチサイクル（ピッチサイクル２０１等）のうちの１つのサンプル数と、再構成される第１のピッチサイクル（ピッチサイクル２０２、２０３、２０４、２０５等）のサンプル数との差を示すサンプル数の差に依拠する。

【０２０２】

たとえば、ある実施形態によれば、ピッチサイクル２０１のサンプルは、たとえば、周期的に繰り返しコピーされ得る。

【０２０３】

そこで、サンプル数の差は、再構成されるべき第１のピッチサイクルに対応する周期的に繰り返されたコピーからいくつのサンプルを削除するか、または再構成されるべき第１のピッチサイクルに対応する周期的に繰り返されたコピーにいくつのサンプルを加えるかを示す。

【０２０４】

図２ｂにおいて、各サンプル数は、周期的に繰り返されたコピーからいくつのサンプルを削除するかを示す。しかしながら、他の例では、サンプル数は、周期的に繰り返されたコピーにいくつのサンプルを加えるかを示し得る。たとえば、いくつかの実施形態では、振幅ゼロのサンプルを対応のピッチサイクルに加えることにより、サンプルを加えることができる。他の実施形態では、たとえば、ピッチサイクルの他のサンプルをコピーすることによって、たとえば、加えるべきサンプルの位置に隣接するサンプルをコピーすることによって、ピッチサイクルにサンプルを加え得る。

【０２０５】

上記では、失われたかまたは破損されたフレームに先行するフレームのピッチサイクルのサンプルが周期的に繰り返しコピーされている実施形態について説明したが、他の実施形態では、失われたかまたは破損されたフレームの後続のフレームのピッチサイクルのサンプルを、周期的に繰り返しコピーして失われたフレームを再構成する。上記および後述の同じ原則が同様に当てはまる。

【０２０６】

このようなサンプル数の差を再構成対象の各ピッチサイクルについて決定し得る。次い

で、各ピッチサイクルのサンプル数の差が、再構成対象の対応のピッチサイクルに対応する周期的に繰り返されるコピーからいくつのサンプルを削除するか、または再構成対象の対応するピッチサイクルに対応する周期的に繰り返されるコピーにいくつのサンプルを加えるかを示す。

【０２０７】

ある実施形態によれば、決定部２１０は、たとえば、再構成対象の複数のピッチサイクルの各々についてサンプル数の差を決定して、それによりピッチサイクルの各々のサンプル数の差が、１以上の入手可能なピッチサイクルのうちの前記１つのサンプルの数と、再構成対象の前記ピッチサイクルのサンプル数との差を示すように構成され得る。フレーム再構成部２２０は、再構成フレームを再構成するために、たとえば、再構成対象の前記ピッチサイクルのサンプル数の差および１以上の入手可能なピッチサイクルの前記１つのサンプルに依拠して、再構成対象の複数のピッチサイクルの各ピッチサイクルを再構成するよう構成され得る。

【０２０８】

ある実施形態においては、フレーム再構成部２２０は、たとえば、１以上の入手可能なピッチサイクルのうちの前記１つに依拠して中間フレームを生成するよう構成され得る。フレーム再構成部２２０は、たとえば、再構成フレームを得るために、中間フレームを修正するよう構成され得る。

【０２０９】

実施形態によれば、決定部２１０は、たとえば、いくつのサンプルを中間フレームから除くか、またはいくつのサンプルを中間フレームに加えるかを示すフレーム差値（ｄ；ｓ）を決定するよう構成され得る。また、フレーム再構成部２２０は、たとえば、フレーム差値が、第１のサンプルがフレームから除去されることを示す場合、再構成フレームを得るために中間フレームから第１のサンプルを除去するよう構成され得る。さらに、フレーム再構成部２２０は、たとえば、フレーム差値（ｄ；ｓ）が、第２のサンプルがフレームに加えられることを示す場合、再構成フレームを得るために中間フレームに第２のサンプルを加えるよう構成され得る。

【０２１０】

ある実施形態において、フレーム再構成部２２０は、たとえば、フレーム差値が、第１のサンプルが中間フレームから除去されるべきであることを示す場合、中間フレームから第１のサンプルを除去するよう構成することが可能で、それにより、中間フレームから除去される第１のサンプルの数がフレーム差値により示されるようになっている。また、フレーム再構成部２２０は、たとえば、フレーム差値が、第２のサンプルがフレームに加えられるべきであることを示す場合、中間フレームに第２のサンプルを加えるよう構成することが可能で、それにより、中間フレームに加えられる第２のサンプルの数がフレーム差値により示されるようになっている。

【０２１１】

ある実施形態によれば、決定部２１０は、たとえば、以下の式が真であるように、フレーム差数ｓを決定するように構成され得る。

【数６３】

$$s = \sum_{i=0}^{M-1} (p[i] - T_r) \frac{L}{MT_r}$$

ここで、Ｌは、再構成フレームのサンプルの数を表し、Ｍは、再構成フレームのサブフレームの数を表し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められたピッチ周期長さを示し、ｐ［ｉ］は、再構成フレームのｉ番目のサブフレームの再構成されたピッチサイクルのピッチ周期長さを示す。

【０２１２】

ある実施形態において、フレーム再構成部２２０は、たとえば、１以上の入手可能なピ

ッチサイクルの前記１つに依拠して中間フレームを生成するようになっていてもよい。また、フレーム再構成部２２０は、たとえば、中間フレームが、第１の部分中間ピッチサイクル、１以上のさらなる中間ピッチサイクルおよび第２の部分中間ピッチサイクルを含むように中間フレームを生成するようにされてもよい。さらに、第１の部分中間ピッチサイクルは、たとえば、１以上の使用可能なピッチサイクルのうちの前記１つのサンプルのうちの１以上に依拠することが可能で、１以上のさらなる中間ピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つのサンプルの全部に依拠し、かつ第２の部分中間ピッチサイクルが、１以上の入手可能なピッチサイクルの前記１つのサンプルのうちの１以上に依拠する。また、決定部２１０は、たとえば、第１の部分中間ピッチサイクルからいくつのサンプルを除去するかまたは加えるかを示す開始部差数を決定するよう構成することができ、かつフレーム再構成部２２０は、開始部差数に依拠して、第１の部分中間ピッチサイクルから１以上の第１のサンプルを除去するよう構成されるかまたは第１の部分中間ピッチサイクルに１以上の第１のサンプルを加えるよう構成される。さらに、決定部２１０は、たとえば、さらなる中間ピッチサイクルの前記１つからいくつのサンプルを除去するまたは加えるかを示すピッチサイクル差数を、さらなる中間ピッチサイクルの各々について決定するよう構成され得る。また、フレーム再構成部２２０は、たとえば、前記ピッチサイクル差数に依拠して、さらなる中間ピッチサイクルの前記１つから１以上の第２のサンプルを除去するよう構成され得るか、または、さらなる中間ピッチサイクルの前記１つに１以上の第２のサンプルを加えるよう構成される。さらに、決定部２１０は、たとえば、第２の部分中間ピッチサイクルからいくつのサンプルを除去するかまたは加えるかを表す終了部差数を決定するよう構成することができ、かつフレーム再構成部２２０は、終了部差数に依拠して、第２の部分中間ピッチサイクルから１以上の第３のサンプルを除去するよう構成されるか、または第２の部分中間ピッチサイクルに１以上の第３のサンプルを加えるよう構成される。

【０２１３】

ある実施形態により、フレーム再構成部２２０は、たとえば、１以上の入手可能なピッチサイクルの前記１つに依拠して中間フレームを生成するよう構成され得る。また、決定部２１０は、たとえば、中間フレームにより含まれる音声信号の１以上の低エネルギ信号部を決定するようにされてもよく、１以上の低エネルギ信号部の各々が、中間フレーム内の音声信号の第１の信号部であり、音声信号のエネルギが、中間フレームにより含まれる音声信号の第２の信号部におけるエネルギより低い。さらに、フレーム再構成部２２０は、たとえば、再構成フレームを得るために、音声信号の１以上の低エネルギ信号部の１以上から１以上のサンプルを除去するか、または音声信号の１以上の低エネルギ信号部分の１以上へ１以上のサンプルを加えるよう構成され得る。

【０２１４】

特定の実施形態において、フレーム再構成部２２０は、たとえば、中間フレームを生成するよう構成されることが可能で、それにより中間フレームが１以上の再構成ピッチサイクルを含み、１以上の再構成ピッチサイクルの各々が、１以上の入手可能なピッチサイクルの前記１つに依拠するようになっている。また、決定部２１０は、たとえば、１以上の再構成ピッチサイクルの各々から除去するサンプルの数を決定するように構成され得る。さらに、決定部２１０は、たとえば、１以上の低エネルギ信号部の各々について、前記低エネルギ信号部のサンプルの数が、１以上の再構成ピッチサイクルの１つから除去されるべきサンプル数に依拠するように、１以上の低エネルギ信号部の各々を決定するように構成することが可能で、前記低エネルギ信号部が、１以上の再構成ピッチサイクルの前記１つ内に位置する。

【０２１５】

ある実施形態において、決定部２１０は、たとえば、再構成フレームとして再構成されるべきフレームの音声信号の１以上のパルスの位置を決定するよう構成され得る。また、フレーム再構成部２２０は、たとえば、音声信号の１以上のパルスの位置に依拠して再構成フレームを再構成するよう構成され得る。

【０２１６】
ある実施形態によれば、決定部２１０は、たとえば、再構成フレームとして再構成されるべきフレームの音声信号の２以上のパルスの位置を決定するよう構成することが可能で、Ｔ［０］は、再構成フレームとして再構成されるべきフレームの音声信号の２以上のパルスのうちの１つの位置であり、かつ決定部２１０は、以下の式に従う音声信号の２以上のパルスのうちのさらなるパルスの位置（Ｔ［ｉ］）を決定するよう構成される。
【数６４】

$$T[i]=T[0]+iT_r$$

ここで、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、かつｉは整数である。
【０２１７】
ある実施形態によれば、決定部２１０は、たとえば、以下の式のように、再構成フレームとして再構成されるべきフレームの音声信号の最後のパルスのインデクスｋを決定するよう構成され得る。
【数６５】

$$k=\left[\frac{L-s-T[0]}{T_r}-1\right]$$

ここで、Ｌは、再構成フレームのサンプルの数を示し、ｓは、フレーム差値を示し、Ｔ［０］は、音声信号の最後のパルスとは異なる、再構成フレームとして再構成されるべきフレームの音声信号のパルスの位置を示し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示す。
【０２１８】
ある実施形態において、決定部２１０は、たとえば、パラメータδを決定することにより再構成フレームとして再構成されるべきフレームを再構成するよう構成することが可能で、δは以下の式により規定される。
【数６６】

$$\delta=\frac{T_{ext}-T_p}{M}$$

ここで、再構成フレームとして再構成されるべきフレームは、Ｍ個のサブフレームを含み、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示し、T_{ext}は、再構成フレームとして再構成されるべきフレームの再構成されるべきピッチサイクルのうちの１つの長さを示す。
【０２１９】
ある実施形態によれば、決定部２１０は、たとえば、以下の式に基づき１以上の入手可能なピッチサイクルの前記１つの丸められた長さT_rを決定することにより再構成フレームを再構成するよう構成され得る。
【数６７】

$$T_r=[T_p+0.5]$$

ここで、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示す。
【０２２０】

ある実施形態において、決定部２１０は、たとえば以下の式を適用することにより再構成フレームを再構成するよう構成され得る。

【数６８】

$$s = \delta \frac{L}{T_r} \frac{M+1}{2} - L\left(1 - \frac{T_p}{T_r}\right)$$

ここで、T_pは、１以上の入手可能なピッチサイクルの前記１つの長さを示し、T_rは、１以上の入手可能なピッチサイクルの前記１つの丸められた長さを示し、再構成フレームとして再構成されるべきフレームは、Ｍ個のサブフレームを含み、再構成フレームとして再構成されるべきフレームは、Ｌ個のサンプルを含み、δが１以上の入手可能なピッチサイクルのうちの前記１つのサンプルの数と、再構成されるべき１以上のピッチサイクルの１つのサンプルの数との差を表す実数である。

【０２２１】

ここで、実施形態についてより詳細に説明する。

【０２２２】

以下では、パルス再同期化の実施形態の第１のグループについて式（２５）から（６３）を参照して説明する。

【０２２３】

これらの実施形態では、ピッチに変化がない場合、最後のピッチラグを、丸めずに小数部分を維持したまま使用する。周期的な部分は、たとえば非特許文献１２（［ＭＴＴＡ９０］）にあるような非整数ピッチおよび内挿を用いて構成される。これにより、丸めたピッチラグを使用する場合に比べて高調波の周波数シフトが減じられるので、一定のピッチの音または有声信号の封じ込めが大きく改善する。

【０２２４】

この効果は、図８および図９により示され、フレームの損失を伴うピッチパイプを表す信号が、それぞれ丸めおよび非丸め小数ピッチラグを用いて封じ込められる。ここで、図８は、丸められたピッチラグを使用する再同期化された音声信号の時間－周波数表現を示す。対照的に、図９は、小数部分を伴う非丸めピッチラグを使用して再同期化された音声信号の時間－周波数表現を示す。

【０２２５】

ピッチの小数部分を使用する場合、計算の複雑さが増大することになる。声門パルス再同期の必要はないので、これが最悪計算量に影響することはないはずである。

【０２２６】

予測されるピッチの変更がない場合には、以下に説明する処理を行う必要はない。

【０２２７】

ピッチの変化が予測される場合、式（２５）から（６３）を参照して説明する実施形態は、一定ピッチ（T_c）を伴うピッチサイクル内の合計サンプル数の和と、展開ピッチｐ［ｉ］を伴うピッチサイクル内の合計サンプル数の和との間の差であるｄを決定するためのコンセプトを提供する。

【０２２８】

以下において、T_cは、式（１５ａ）のように規定される。すなわち、T_c＝ｒｏｕｎｄ（ｌａｓｔ＿ｐｉｔｃｈ）。

【０２２９】

実施形態によれば、以下に説明するとおり、差ｄをより高速でかつより正確なアルゴリズムを使用して決定することができる（ｄを決定するための高速アルゴリズムのアプローチ）。

【０２３０】

このようなアルゴリズムは、たとえば、以下の原則に基づくことができる。

・各サブフレームｉにおいて、（長さT_cの）各ピッチサイクルについて、T_c－ｐ［ｉ］サンプルを除去する必要がある（またはT_c－ｐ［ｉ］＜０の場合、ｐ［ｉ］－T_cを加える必要がある）。

・各サブフレームには、（Ｌ＿ｓｕｂｆｒ）／T_cピッチサイクルが存在する。

・したがって、各サブフレーム（T_c－ｐ［ｉ］）について、（Ｌ＿ｓｕｂｆｒ）／T_cサンプルを除去する必要がある。

【０２３１】

いくつかの実施形態によれば、丸めを行わず、小数ピッチを使用する。その場合、以下のとおりである。

・ｐ［ｉ］＝T_c+（ｉ＋１）δ

・したがって、各サブフレームｉについて、δ＜０の場合、－（ｉ＋１）δ（Ｌ＿ｓｕｂｆｒ）／T_cサンプルを除去する必要がある（またはδ＞０であれば加える必要がある）。

・したがって、$d = -\delta \frac{L_subfr}{T_c} \sum_{i=1}^{M} i$（Ｍは、フレームにおけるサブフレームの数である。）

【０２３２】

いくつかの他の実施形態によれば、丸めが行われる。整数ピッチについては（Ｍはフレームにおけるサブフレームの数である。）、ｄは以下のとおり規定される。

【数６９】

$$d = round\left(\left(MT_c - \sum_{i=0}^{M-1} p[i]\right)\frac{L_subfr}{T_c}\right) \qquad (25)$$

【０２３３】

ある実施形態によれば、応じてｄを計算するためのアルゴリズムが提供される。

```
ftmp = 0;
for (i=0;i <M;i++) {
        ftmp += p[i];
}
d = (short)floor((M*T_c - ftmp)*(float)L_subfr/ T_c +0.5);
```

【０２３４】

他の実施形態では、アルゴリズムの最後の行を以下と置換する。

ｄ＝（ｓｈｏｒｔ）ｆｌｏｏｒ（Ｌ＿ｆｒａｍｅ－ｆｔｍｐ＊（ｆｌｏａｔ）Ｌ＿ｓｕｂｆｒ／Ｔ＿ｃ＋０．５；

【０２３５】

実施形態によれば、最後のパルスＴ［ｎ］は、以下の式に従って見つけられる。

【数７０】

$$n = i \mid T[0] + iT_c < L_frame \ \wedge \ T[0] + (i+1)T_c \geq L_frame \qquad (26)$$

【０２３６】
ある実施形態によれば、Ｎを計算する式が採用される。この式は、以下に従って式（２６）から得られ、
【数７１】

$$N = 1 + \left\lceil \frac{L_frame - T[0]}{T_c} \right\rceil \qquad (27)$$

かつ最後のパルスは、インデクスＮ－１を有する。
【０２３７】
この式によれば、図４および図５に示す例についてＮを計算できる。
【０２３８】
以下において、最後のパルスについて明示的サーチを伴わないが、パルスの位置を考慮に入れるコンセプトについて説明する。このコンセプトは、構成された周期的部分における最後のパルスインデクスのＮを必要としない。
【０２３９】
励振の構成された周期的部分における実際の最後のパルスの位置（Ｔ［ｋ］）がフルピッチサイクルｋの数を決定し、サンプルが除去される（または加えられる）。
【０２４０】
図１２は、ｄ個のサンプルを除去する前の最後のパルスの位置Ｔ［２］を示す。式（２５）から式（６３）を参照して説明する実施形態については、参照番号１２１０がｄを示す。
【０２４１】
図１２の例では、最後のパルスｋのインデクスは２であり、サンプルを除去すべき２つのフルピッチサイクルが存在する。
【０２４２】
信号長さＬ＿ｆｒａｍｅ+ｄの信号からｄ個のサンプルを除去した後は、Ｌ＿ｆｒａｍｅ+ｄ個のサンプルを超えるオリジナル信号からのサンプルは存在しない。したがって、Ｔ［ｋ］は、Ｌ＿ｆｒａｍｅ+ｄサンプルの範囲内であり、かつ、従って、ｋは以下により決定される。
【数７２】

$$k = i \mid T[i] < L_{frame} + d \leq T[i+1] \qquad (28)$$

【０２４３】
式（１７）および式（２８）から、以下のとおりになる。
【数７３】

$$T[0] + kT_c < L_{frame} + d \leq T[0] + (k+1)T_c \qquad (29)$$

すなわち、以下のとおりである。

【数７４】

$$\frac{L_{frame}+d-T[0]}{T_c}-1\leq k<\frac{L_{frame}+d-T[0]}{T_c} \qquad (30)$$

【０２４４】

式（３０）から、以下のとおりになる。

【数７５】

$$k=\left[\frac{L_{frame}+d-T[0]}{T_c}-1\right] \qquad (31)$$

【０２４５】

たとえば、２０ｍｓ以上のフレームを使用するコーデックにおいて、音声の最低基本周波数が、たとえば、４０Ｈｚ以上なら、多くの場合、「無声」以外に、封じ込められたフレームに、１以上のパルスが存在する。

【０２４６】

以下において、式（３２）から式（４６）を参照して、２以上のパルス（ｋ≧１）の場合について説明する。

【０２４７】

パルス間の各フルのｉ番目のピッチサイクルにおいて、Δ_iサンプルが除去されると仮定するが、ここで、Δ_iは、以下のとおり規定される。

【数７６】

$$\Delta_i=\Delta+(i-1)a,\ 1\leq i\leq k \qquad (32)$$

ここで、ａは、既知の変数で表現する必要がある未知の変数である。

【０２４８】

第１のパルスの前でΔ_0サンプルが除去されると仮定するが、ここで、Δ_0は、以下のとおり規定される。

【数７７】

$$\Delta_0=(\Delta-a)\frac{T[0]}{T_c} \qquad (33)$$

【０２４９】

Δ_{k+1}サンプルが、最後のパルスの後に除去されると仮定するが、ここで、Δ_{k+1}は、以下のとおり規定される。

【数７８】

$$\Delta_{k+1}=(\Delta+ka)\frac{L+d-T[k]}{T_c} \qquad (34)$$

【０２５０】

最後の２つの仮定は、部分的な最初と最後のピッチサイクルの長さを考慮に入れる式（３２）と合致する。

【０２５１】

Δ_i値の各々がサンプル数の差である。また、Δ_0は、サンプル数の差である。さらに、Δ_{k+1}は、サンプル数の差である。

【０２５２】

図１３は、図１２の音声信号の図であって、Δ_0からΔ_3を追加して示す。各ピッチサイクルにおいて除去すべきサンプルの数を、図１３の例では模式的に示し、ｋ＝２である。式（２５）から（６３）を参照して記載する実施形態に関しては、参照番号１２１０がｄを示す。

【０２５３】

除去するサンプルの合計数ｄは、以下のとおりΔ_iに関連する。

【数７９】

$$d = \sum_{i=0}^{k+1} \Delta_i \qquad (35)$$

【０２５４】

式（３２）～式（３５）から、ｄは、以下のとおり求めることができる。

【数８０】

$$d = (\Delta - a)\frac{T[0]}{T_c} + (\Delta + ka)\frac{L + d - T[k]}{T_c} + \sum_{i=1}^{k} (\Delta + (i-1)a) \qquad (36)$$

【０２５５】

式（３６）は、以下の式と等価である。

【数８１】

$$d = \Delta\left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k\right) + a\left(k\frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k(k-1)}{2}\right) \qquad (37)$$

【０２５６】

封じ込められたフレームにおける最後のフルピッチサイクルがｐ［Ｍ－１］の長さを有すると仮定する。すなわち、以下のとおりである。

【数８２】

$$\Delta_k = T_c - p[M-1] \qquad (38)$$

【０２５７】

式（３２）および式（３８）から、以下のとおりである。

【数８３】

$$\Delta = T_c - p[M-1] - (k-1)a \qquad (39)$$

【０２５８】

また、式（３７）および式（３９）から、以下のとおりである。
【数８４】

$$d = (T_c - p[M-1] + (1-k)\,a)\left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k\right) +$$

$$+ a\left(k\frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k\,(k-1)}{2}\right) \qquad (40)$$

【０２５９】
式（４０）は、以下の式と等価である。
【数８５】

$$d = (T_c - p[M-1])\left(\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} + k\right)$$

$$+ a\left((1-k)\frac{T[0]}{T_c} + (1-k)\frac{L + d - T[k]}{T_c} +\right.$$

$$\left. + (1-k)\,k + k\frac{L + d - T[k]}{T_c} - \frac{T[0]}{T_c} + \frac{k\,(k-1)}{2}\right) \qquad (41)$$

【０２６０】
式（１７）および式（４１）から、以下のとおりである。
【数８６】

$$d = (T_c - p[M-1])\frac{L + d}{T_c} + a\left(-k\frac{T[0]}{T_c} + \frac{L + d - T[k]}{T_c} - \frac{k\,(k-1)}{2}\right) \qquad (42)$$

【０２６１】
式（４２）は、以下の式と等価である。
【数８７】

$$dT_c = (T_c - p[M-1])\,(L + d) +$$

$$+ a\left(-kT[0] + L + d - T[k] + \frac{k\,(1-k)}{2}T_c\right) \qquad (43)$$

【０２６２】
さらに、式（４３）から、以下のとおりである。

【数８８】

$$a = \frac{dT_c - (T_c - p[M-1])(L+d)}{-kT[0] + L + d - T[k] + \frac{k(1-k)}{2}T_c} \qquad (44)$$

【０２６３】
式（４４）は、以下の式と等価である。
【数８９】

$$a = \frac{p[M-1](L+d) - T_c L}{L + d - (k+1)T[0] - kT_c + \frac{k(1-k)}{2}T_c} \qquad (45)$$

【０２６４】
また、式（４５）は、以下の式と等価である。
【数９０】

$$a = \frac{p[M-1](L+d) - T_c L}{L + d - (k+1)T[0] - \frac{k(1+k)}{2}T_c} \qquad (46)$$

【０２６５】
実施形態によれば、ここで、式（３２）から式（３４）、式（３９）および式（４６）に基づいて、第１のパルスの前および／またはパルスの間および／または最後のパルスの後に除去するかまたは加えるサンプルの数を計算する。
【０２６６】
実施形態において、サンプルは、最小エネルギ領域において除去されるかまたは加えられる。
【０２６７】
実施形態によれば、除去されるサンプルの数は、たとえば、以下を使用して丸めることができる。
【数９１】

$$\Delta_0' = \lfloor \Delta_0 \rfloor$$

$$\Delta_i' = \lfloor \Delta_i + \Delta_{i-1} - \Delta_{i-1}' \rfloor,\ 0 < i \leq k$$

$$\Delta_{k+1} = d - \sum_{i=0}^{k} \Delta_i$$

【０２６８】

以下では、１つのパルス（ｋ＝０）の場合について、式（４７）から（５５）を参照して記載する。

【０２６９】

封じ込められたフレーム内に１つのパルスしかなければ、そのパルスの前のΔ_0のサンプルが除去されることになる。

【数９２】

$$\Delta_0 = (\Delta - a)\frac{T[0]}{T_c} \qquad (47)$$

ここで、Δおよびａは、既知の変数で表現する必要がある未知の変数である。Δ_1個のサンプルが、このパルスの後、除去されることになる。ここで、

【数９３】

$$\Delta_1 = \Delta\frac{L + d - T[0]}{T_c} \qquad (48)$$

である。

【０２７０】

そして、除去されるべきサンプルの合計数は、以下のとおり与えられる。

【数９４】

$$d = \Delta_0 + \Delta_1 \qquad (49)$$

【０２７１】

式（４７）から式（４９）より、以下のとおりである。

【数９５】

$$d = (\Delta - a)\frac{T[0]}{T_c} + \Delta\frac{L + d - T[0]}{T_c} \qquad (50)$$

【０２７２】

式（５０）は、以下の式と等価である。

【数９６】

$$dT_c = \Delta(L + d) - aT[0] \qquad (51)$$

【０２７３】

パルスの前のピッチサイクルのパルス後のピッチサイクルに対する比は、前に受信されたフレームにおける最後のサブフレームと最初のサブフレームにおけるピッチラグの比と同じであると仮定する。

【数９７】

$$\frac{\Delta}{\Delta - a} = \frac{p[-1]}{p[-4]} = r \qquad (52)$$

【０２７４】
式（５２）から、以下のとおりである。
【数９８】

$$a = \Delta\left(1 - \frac{1}{r}\right) \qquad (53)$$

【０２７５】
また、式（５１）と式（５３）から、以下のとおりである。
【数９９】

$$dT_c = \Delta(L + d) - \Delta\left(1 - \frac{1}{r}\right)T[0] \qquad (54)$$

【０２７６】
式（５４）は、以下の式と等価である。
【数１００】

$$\Delta = \frac{dT_c}{L + d + \left(\frac{1}{r} - 1\right)T[0]} \qquad (55)$$

【０２７７】
除去すべきまたは加えるべき［Δ－ａ］個のサンプルが、上記パルスの前の最小エネルギ領域に存在し、同パルスの後にｄ－［Δ－ａ］個のサンプルが存在する。
【０２７８】
以下では、パルス（の場所）のサーチを必要としない、実施形態による簡素化されたコンセプトについて、式（５６）から式（６３）を参照して説明する。
【０２７９】
ｔ［ｉ］は、ｉ番目のピッチサイクルの長さを示す。信号からｄ個のサンプルを除去した後、ｋ個のフルピッチサイクルおよび１つの部分（フルまでの）ピッチサイクルを入手する。したがって、以下のとおりである。
【数１０１】

$$\sum_{i=0}^{k-1} t[i] < L \leq \sum_{i=0}^{k} t[i] \qquad (56)$$

【０２８０】
長さｔ［ｉ］のピッチサイクルを、いくつかのサンプルを除去した後に長さT_Cのピッチサイクルから得て、除去されたサンプルの合計数がｄなので、以下のとおりになる。

【数１０２】

$$kT_c < L + d \leq (k+1)T_c \qquad (57)$$

【０２８１】
したがって、以下のとおりになる。
【数１０３】

$$\frac{L+d}{T_c} - 1 \leq k < \frac{L+d}{T_c} \qquad (58)$$

【０２８２】
また、以下のとおりになる。
【数１０４】

$$k = \left\lceil \frac{L+d}{T_c} \right\rceil - 1 \qquad (59)$$

【０２８３】
実施形態によれば、ピッチラグにおいて線形の変化を想定し得る。
【数１０５】

$$t[i] = T_c - (i+1)\Delta,\ 0 \leq i \leq k$$

【０２８４】
実施形態では、（ｋ＋１）Δ個のサンプルを、ｋ番目のピッチサイクルにおいて除去する。
【０２８５】
実施形態によれば、サンプルを除去した後もフレームにとどまるｋ番目のピッチサイクルの部分において、
【数１０６】

$$\frac{L+d-kT_c}{T_c}(k+1)\Delta$$

個のサンプルが除去される。
【０２８６】
したがって、除去されるサンプルの合計数は、以下のとおりである。
【数１０７】

$$d = \frac{L+d-kT_c}{T_c}(k+1)\Delta + \sum_{i=0}^{k-1}(i+1)\Delta \qquad (60)$$

【０２８７】
式（６０）は以下の式と等価である。

【数１０８】

$$d = \frac{L + d - kT_c}{T_c}(k+1)\Delta + \frac{k(k+1)}{2}\Delta$$

【０２８８】

また、式（６１）は、以下の式と等価である。

【数１０９】

$$\frac{d}{(k+1)} = \left(\frac{L + d - kT_c}{T_c} + \frac{k}{2}\right)\Delta \qquad (62)$$

【０２８９】

さらに、式（６２）は、以下の式と等価である。

【数１１０】

$$\Delta = \frac{2dT_c}{(k+1)(2L + 2d - kT_c)} \qquad (63)$$

【０２９０】

実施形態によれば、（ｉ＋１）Δ個のサンプルが、最小エネルギの位置で除去される。１ピッチサイクルを保持する環状バッファにおいて、最小エネルギ位置のサーチが行われるので、パルスの場所を知る必要はない。

【０２９１】

最小エネルギ位置が、第１のパルスの後であり、かつ第１のパルスの前のサンプルが除去されない場合、ピッチラグが、(T_c＋Δ）、T_c、T_c、（T_c－Δ）、（T_c－２Δ）（最後に受信したフレームにおける２つピッチサイクルおよび封じ込められたフレームにおける３つのピッチサイクル）として展開する状況が発生し得る。したがって、不連続性が存在し得る。同様の不連続性については、最後のパルスの後に生じ得るが、第１のパルスの前に発生する場合と同じ時には生じない。

【０２９２】

他方、パルスが封じ込められたフレームの開始に近いほど、最小エネルギ領域が第１のパルスの後に現れる可能性が高い。第１のパルスが、封じ込められたフレームの開始に近いほど、最後に受信したフレームにおける最後のピッチサイクルがT_cより大きくなる可能性が高くなる。ピッチ変化における不連続性の可能性を減じるため、重み付けを用いてピッチサイクルの開始または終了により近い最小領域を有利にする。

【０２９３】

実施形態によれば、以下の方法ステップの１以上または全部を実現する、提供されるコンセプトの実現例について説明する。

【０２９４】

１．最小エネルギ領域について並列でサーチし、最後に受信したフレームの終わりからローパスフィルタ処理したT_c個のサンプルを一時バッファＢに格納する。一時バッファは、最小エネルギ領域のサーチの際には環状バッファとして考えられる（これは、最小エネルギ領域が、ピッチサイクルの始まりからの数サンプルと終わりからの数サンプルから構成され得るということを意味し得る）。最小エネルギ領域は、たとえば、長さが［（ｋ＋１）Δ］のサンプルのスライディングウィンドウについて最小の場所でもよい。たとえば重み付けを使用して、ピッチサイクルの開始により近い最小領域を有利にすることができる。

【０２９５】
２．最小エネルギ領域の［Δ］個のサンプルをスキップして、フレームに一時バッファＢからのサンプルをコピーする。したがって、長さｔ［０］のピッチサイクルが作られる。$\delta_0 = \Delta - [\Delta]$を設定。
【０２９６】
３．ｉ番目のピッチサイクル（０＜ｉ＜ｋ）について、最小エネルギ領域の［Δ］＋［δ_{i-1}］個のサンプルをスキップして、（ｉ－１）番目のピッチサイクルからのサンプルをコピーする。$\delta_i = \delta_{i-1} - [\delta_{i-1}] + \Delta - [\Delta]$を設定する。このステップをｋ－１回繰り返す。
【０２９７】
４．ｋ番目のピッチサイクルについて、ピッチサイクルの終りに近い最小領域ほど有利になる重み付けを用いて、（ｋ－１）番目のピッチサイクルにおける新たな最小領域をサーチする。次いで、最小エネルギ領域において以下の式で表す個数のサンプルをスキップして、（ｋ－１）番目のピッチサイクルからのサンプルをコピーする。
【数１１１】

$$d - \left\lfloor \frac{k(k+1)}{2}\Delta + \frac{k(k-1)}{2}\Delta \right\rfloor = d - \lfloor k^2\Delta \rfloor$$

【０２９８】
サンプルを加える必要がある場合には、ｄ＜０でかつΔ＜０であり、かつ合計|ｄ|個のサンプルを加えるということを考慮に入れることにより、等価な手順を用いることができ、すなわち（ｋ＋１）|Δ|個のサンプルが、ｋ番目のサイクルにおいて、最小エネルギの位置に加えられる。
【０２９９】
いずれにしても、近似化したピッチサイクル長さを用いるので、「ｄを決定するための高速アルゴリズムアプローチ」に関して、上記のｄを生成するため、サブフレームレベルで小数ピッチを使用することができる。
【０３００】
以下で、パルス再同期化の実施形態の第２のグループについて、式（６４）から（１１３）を参照して説明する。第１のグループのこれらの実施形態は、式（１５ｂ）の定義を採用する。
【数１１２】

$$T_r = \lfloor T_p + 0.5 \rfloor$$

ここで、最後のピッチ周期長さは、T_pであり、かつコピーされたセグメントの長さは、T_rである。
【０３０１】
パルス再同期化実施形態の第２のグループにより使用されるいくつかのパラメータが以下に規定されない場合は、本発明の実施形態は、上に規定したパルス再同期化実施形態の第１のグループに関してこれらのパラメータについて与えられた定義を採用し得る（式（２５）から（６３）を参照）。
【０３０２】
パルス再同期化実施形態の第２のグループの式（６４）から（１１３）のいくつかは、パルス再同期化実施形態の第１のグループに関して既に使用されたパラメータのいくつかを再定義し得る。この場合、与えられる再定義された定義が、第２のパルス再同期化の実施形態に適用される。
【０３０３】

上記のとおり、いくつかの実施形態によれば、周期的部分は、たとえば、１つのフレームおよび１つの追加のサブフレームについて構成することができ、ここで、フレーム長さはＬ＝Ｌ$_{frame}$として示される。

【０３０４】

たとえば、フレームにＭ個のサブフレームがある場合、サブフレームの長さは、Ｌ＿ｓｕｂｆｒ＝Ｌ／Ｍである。

【０３０５】

上記のとおり、Ｔ［０］は、励振の構成された周期的部分における第１の最大パルスの場所である。他のパルスの位置は、以下の式により与えられる。

【数１１３】

$$T[i]=T[0]+iT_r$$

【０３０６】

実施形態によれば、励振の周期的部分の構成に依拠して、たとえば励振の周期的部分の構成後、声門パルス再同期化を行って、失われたフレーム（Ｐ）の最後のパルスの推定目標位置と、励振の構成された周期的部分におけるその実際の位置（Ｔ［ｋ］）との差を訂正する。

【０３０７】

失われたフレーム（Ｐ）における最後のパルスの推定目標位置は、たとえば、ピッチラグ展開の推定により間接的に決定され得る。ピッチラグ展開は、たとえば、失われたフレームの前の最後の７つのサブフレームのピッチラグに基づいて外挿される。各サブフレームにおける展開ピッチラグは、以下のとおりである。

【数１１４】

$$p[i]=T_p+(i+1)\delta,\ 0\leqq i<M \qquad (64)$$

ここで、以下のとおりであり、

【数１１５】

$$\delta=\frac{T_{ext}-T_p}{M} \qquad (65)$$

かつＴ$_{ext}$は、外挿されたピッチであり、かつｉは、サブフレームインデクスである。ピッチ外挿は、たとえば、重み付線形フィッティングまたはＧ．７１８からの方法もしくはＧ．７２９．１からの方法またはたとえば未来のフレームからの１以上のピッチを考慮するピッチ内挿のための他の方法を用いて行うことができる。ピッチ外挿は、非線形でも可能である。実施形態では、Ｔ$_{ext}$は、上記でＴ$_{ext}$が決定されるのと同じ方法で決定され得る。

【０３０８】

展開ピッチ（ｐ［ｉ］）を伴うピッチサイクル内の合計サンプル数の和と、一定ピッチ（Ｔ$_p$）を伴うピッチサイクル内の合計サンプル数の和との間のフレーム長内の差をｓで示す。

【０３０９】

実施形態によれば、Ｔ$_{ext}$＞Ｔ$_p$なら、ｓ個のサンプルをフレームに加える必要があり、かつＴ$_{ext}$＜Ｔ$_p$なら、－ｓ個のサンプルをフレームから除去する必要がある。|ｓ|個のサンプルを追加または除去した後、封じ込められたフレームにおける最後のパルスは、推定目標位置（Ｐ）にあることになる。

【０３１０】

$T_{ext}=T_p$なら、フレーム内にサンプルを加えたり除去したりする必要はない。

【０３１１】

いくつかの実施形態によれば、声門パルス再同期化は、全てのピッチサイクルの最小エネルギ領域において、サンプルを加えるかまたは除去することにより行われる。

【０３１２】

以下では、実施形態によるパラメータｓの計算について、式（６６）から（６９）を参照して説明する。

【０３１３】

いくつかの実施形態によれば、差ｓは、たとえば、以下の原則に基づいて計算され得る。

・各サブフレームｉにおいて、（長さT_rの）ピッチサイクルごとにｐ［ｉ］－T_r個のサンプルを加える必要がある（ｐ［ｉ］－T_r＞０の場合）（さもなくばｐ［ｉ］－T_r＜０の場合、T_r－ｐ［ｉ］個のサンプルを除く必要がある）。

・各サブフレームには、（Ｌ＿ｓｕｂｆｒ）／T_r＝Ｌ／（MT_r）のピッチサイクルが存在する。

・したがって、ｉ番目のサブフレームにおいて（ｐ［ｉ］－T_r）Ｌ／（MT_r） 個のサンプルを除去する必要がある。

【０３１４】

したがって、式（６４）に従って、実施形態により、ｓは、たとえば式（６６）に従って計算され得る。

【数１１６】

$$s=\sum_{i=0}^{M-1}(p[i]-T_r)\frac{L}{MT_r}=\sum_{i=0}^{M-1}\left(T_p+(i+1)\delta-T_r\right)\frac{L}{MT_r}=$$

$$=\frac{L}{MT_r}\sum_{i=0}^{M-1}\left((i+1)\delta+T_p-T_r\right) \qquad (66)$$

【０３１５】

式（６６）は、以下の式と等価である。

【数１１７】

$$s=\frac{L}{MT_r}\left(M\left(T_p-T_r\right)+\delta\sum_{i=0}^{M-1}(i+1)\right)=\frac{L}{MT_r}\left(M\left(T_p-T_r\right)+\delta\frac{M(M+1)}{2}\right) \qquad (67)$$

【０３１６】

ここで、式（６７）は、以下の式と等価である。

【数１１８】

$$s=\frac{L}{T_r}\left(T_p-T_r+\delta\frac{(M+1)}{2}\right)=\frac{L}{T_r}\delta\frac{M+1}{2}+\frac{L}{T_r}\left(T_p-T_r\right) \qquad (68)$$

【０３１７】

式（６８）は、以下の式と等価である。

【数１１９】

$$s = \delta \frac{L}{T_r} \frac{M+1}{2} - L\left(1 - \frac{T_p}{T_r}\right) \qquad (69)$$

【０３１８】

なお、$T_{ext} > T_p$なら、ｓは正であり、サンプルを加える必要があり、$T_{ext} < T_p$なら、ｓは負であり、サンプルを除去する必要がある。したがって、除去または追加するべきサンプルの数は、|ｓ|として示すことができる。

【０３１９】

以下では、実施形態による最後のパルスのインデクスの計算について式（７０）から式（７３）を参照して説明する。

【０３２０】

励振の構成された周期的部分における実際の最後のパルス位置（Ｔ［ｋ］）は、サンプルが除去される（または加えられる）フルピッチサイクルの数ｋを決定する。

【０３２１】

図１２は、サンプルを除去する前の音声信号を示す。

【０３２２】

図１２が示す例においては、最後のパルスｋのインデクスが２であり、サンプルを除去すべき２つのフルピッチサイクルが存在する。式（６４）から（１１３）を参照して説明する実施形態に関しては、参照番号１２１０が|ｓ|を示す。

【０３２３】

|ｓ|個のサンプルを長さＬ－ｓの信号から除去した後（Ｌ＝Ｌ＿ｆｒａｍｅ）または|ｓ|個のサンプルを長さＬ－ｓの信号に加えた後、Ｌ－ｓ個のサンプルを超えるオリジナル信号からのサンプルは存在しない。なお、サンプルが加えられる場合、ｓは正であり、サンプルが除去される場合、ｓは負である。したがって、サンプルが加えられるなら、Ｌ－ｓ＜Ｌであり、かつサンプルが除去されるなら、Ｌ－ｓ＞Ｌである。したがって、Ｔ［ｋ］は、Ｌ－ｓサンプルの範囲内でなければならず、かつｋは以下のとおり決定される。

【数１２０】

$$k = i \mid T[i] < L - s \le T[i+1] \qquad (70)$$

【０３２４】

式（１５ｂ）および式（７０）から、以下のとおりになる。

【数１２１】

$$T[0] + kT_r < L - s \le T[0] + (k+1)T_r \qquad (71)$$

【０３２５】

すなわち、以下のとおりである。

【数１２２】

$$\frac{L - s - T[0]}{T_r} - 1 \le k < \frac{L - s - T[0]}{T_r} \qquad (72)$$

【０３２６】

ある実施形態によれば、ｋは、たとえば式（７２）に基づき以下のとおり決定され得る。

【数１２３】

$$k = \left\lfloor \frac{L - s - T[0]}{T_r} - 1 \right\rfloor \qquad (73)$$

【０３２７】

たとえば、２０ｍｓ以上のフレームおよび４０Ｈｚ以上の音声の最低基本周波数を採用するコーデックにおいては、多くの場合、「無声」以外に封じ込められたフレームにおいて１以上のパルスが存在する。

【０３２８】

以下では、実施形態に従って、最小領域において除去されるべきサンプルの数の計算について、式（７４）から（９９）を参照して説明する。

【０３２９】

たとえば、パルス間の各フルのｉ番目のピッチサイクルにおいてΔ_i個のサンプルを除去（または追加）するものと仮定することができ、ここで、Δ_iは、以下のとおり定義される。

【数１２４】

$$\Delta_i = \Delta + (i-1)a, 1 \le i \le k \qquad (74)$$

ここで、ａは、たとえば既知の変数で表現され得る未知の変数である。

【０３３０】

また、たとえば第１のパルスの前に、Δ^p_0個のサンプルを除去（または追加）すると仮定することができ、ここでΔ^p_0は、以下のとおり規定される。

【数１２５】

$$\Delta^p_0 = \Delta_0 \frac{T[0]}{T_r} = (\Delta - a)\frac{T[0]}{T_r} \qquad (75)$$

【０３３１】

さらに、たとえば、最後のパルスの後にΔ^p_{k+1}個のサンプルを除去（または追加）すると仮定することができ、ここでΔ^p_{k+1}は、以下のとおり規定される。

【数１２６】

$$\Delta^p_{k+1} = \Delta_{k+1} \frac{L - s - T[k]}{T_r} = (\Delta + ka)\frac{L - s - T[k]}{T_r} \qquad (76)$$

【０３３２】

最後の２つの仮定は、部分的な最初と最後のピッチサイクルの長さを考慮に入れる式（７４）に合致する。

【０３３３】

各ピッチサイクルにおいて除去される（または加えられる）サンプルの数を図１３の例に模式的に示し、ここで、ｋ＝２である。図１３は、各ピッチサイクルにおいて除去されるサンプルを模式的に示す図である。式（６４）から（１１３）を参照して説明した実施形態については、参照番号１２１０が|ｓ|を示す。

【０３３４】

除去すべき（または加えるべき）サンプルの合計数ｓは、以下に従ってΔ_iに関連する

。

【数１２７】

$$|s| = \Delta_0^p + \Delta_{k+1}^p + \sum_{i=1}^{k} \Delta_i \qquad (77)$$

【０３３５】

式（７４）～式（７７）から、以下のとおりである。

【数１２８】

$$|s| = (\Delta - a)\frac{T[0]}{T_r} + (\Delta + ka)\frac{L - s - T[k]}{T_r} + \sum_{i=1}^{k}(\Delta + (i-1)a) \qquad (78)$$

【０３３６】

式（７８）は、以下の式と等価である。

【数１２９】

$$|s| = (\Delta - a)\frac{T[0]}{T_r} + (\Delta + ka)\frac{L - s - T[k]}{T_r} + k\Delta + a\sum_{i=1}^{k}(i-1) \qquad (79)$$

【０３３７】

また、式（７９）は、以下の式と等価である。

【数１３０】

$$|s| = (\Delta - a)\frac{T[0]}{T_r} + (\Delta + ka)\frac{L - s - T[k]}{T_r} + k\Delta + a\frac{k(k-1)}{2} \qquad (80)$$

【０３３８】

さらに、式（８０）は、以下の式と等価である。

【数１３１】

$$|s| = \Delta\left(\frac{T[0]}{T_r} + \frac{L - s - T[k]}{T_r} + k\right) + a\left(k\frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2}\right) \qquad (81)$$

【０３３９】

また、式（１６ｂ）を考慮して、式（８１）は、以下の式と等価である。

【数１３２】

$$|s| = \Delta\left(\frac{L - s}{T_r}\right) + a\left(k\frac{L - s - T[k]}{T_r} - \frac{T[0]}{T_r} + \frac{k(k-1)}{2}\right) \qquad (82)$$

【０３４０】

実施形態に従って、最後のパルスの後の完全なピッチサイクルにおいて除去（または追加）するべきサンプルの数が以下の式により与えられると仮定し得る。

【数１３３】

$$\Delta_{k+1}=|T_r-p[M-1]|=|T_r-T_{ext}| \quad (83)$$

【０３４１】
式（７４）および式（８３）から、以下のとおりである。
【数１３４】

$$\Delta=|T_r-T_{ext}|-ka \quad (84)$$

【０３４２】
式（８２）および式（８４）から、以下のとおりである。
【数１３５】

$$|s|=(|T_r-T_{ext}|-ka)\left(\frac{L-s}{T_r}\right)+a\left(k\frac{L-s-T[k]}{T_r}-\frac{T[0]}{T_r}+\frac{k(k-1)}{2}\right) \quad (85)$$

【０３４３】
式（８５）は、以下の式と等価である。
【数１３６】

$$|s|=|T_r-T_{ext}|\left(\frac{L-s}{T_r}\right)+a\left(-k\frac{L-s}{T_r}+k\frac{L-s-T[k]}{T_r}-\frac{T[0]}{T_r}+\frac{k(k-1)}{2}\right) \quad (86)$$

【０３４４】
また、式（８６）は、以下の式と等価である。
【数１３７】

$$|s|=|T_r-T_{ext}|\left(\frac{L-s}{T_r}\right)+a\left(-k\frac{T[k]}{T_r}-\frac{T[0]}{T_r}+\frac{k(k-1)}{2}\right) \quad (87)$$

【０３４５】
さらに、式（８７）は、以下の式と等価である。
【数１３８】

$$|s|T_r=|T_r-T_{ext}|(L-s)+a\left(-kT[k]-T[0]+\frac{k(k-1)}{2}T_r\right) \quad (88)$$

【０３４６】
式（１６ｂ）および式（８８）から、以下のようになる。
【数１３９】

$$|s|T_r=|T_r-T_{ext}|(L-s)+a\left(-kT[0]-k^2T_r-T[0]+\frac{k(k-1)}{2}T_r\right) \quad (89)$$

【０３４７】

式（８９）は、以下の式と等価である。

【数１４０】

$$|s|T_r = |T_r - T_{ext}|(L-s) + a\left(-(k+1)T[0] - \frac{k(k+1)}{2}T_r\right) \quad (90)$$

【０３４８】

また、式（９０）は以下の式と等価である。

【数１４１】

$$|s|T_r - |T_r - T_{ext}|(L-s) = a\left(-(k+1)T[0] - \frac{k(k+1)}{2}T_r\right) \quad (91)$$

【０３４９】

さらに、式（９１）は、以下の式と等価である。

【数１４２】

$$|s|T_r - |T_r - T_{ext}|(L-s) = -(k+1)a\left(T[0] + \frac{k}{2}T_r\right) \quad (92)$$

【０３５０】

また、式（９２）は、以下の式と等価である。

【数１４３】

$$|T_r - T_{ext}|(L-s) - |s|T_r = (k+1)a\left(T[0] + \frac{k}{2}T_r\right) \quad (93)$$

【０３５１】

式（９３）から、以下のとおりである。

【数１４４】

$$a = \frac{|T_r - T_{ext}|(L-s) - |s|T_r}{(k+1)\left(T[0] + \frac{k}{2}T_r\right)} \quad (94)$$

【０３５２】

このように、たとえば、式（９４）に基づき、実施形態に従えば、以下のとおりである。

・第１のパルスの前で除去されるべきかつ／または加えられるべきサンプルの数が計算され、かつ／または

・パルス間で除去されるべきかつ／または加えられるべきサンプルの数が計算されかつ／または

・最後のパルスの後で除去されるべきかつ／または加えられるべきサンプルの数が計算される。

【０３５３】

いくつかの実施形態によれば、サンプルは、たとえば、最小エネルギ領域において除去されるかまたは加えられ得る。

【０３５４】

式（８５）および式（９４）から、以下のとおりになる。

【数１４５】

$$\Delta_0^p = (\Delta - a)\frac{T[0]}{T_r} = (|T_r - T_{ext}| - ka - a)\frac{T[0]}{T_r} \quad (95)$$

【０３５５】
式（９５）は、以下の式と等価である。
【数１４６】

$$\Delta_0^p = (|T_r - T_{ext}| - (k+1)a)\frac{T[0]}{T_r} \quad (96)$$

【０３５６】
また、式（８４）および式（９４）から、以下のとおりである。
【数１４７】

$$\Delta_i = \Delta + (i-1)a = |T_r - T_{ext}| - ka + (i-1)a, 1 \le i \le k \quad (97)$$

【０３５７】
式（９７）は、以下の式と等価である。
【数１４８】

$$\Delta_i = |T_r - T_{ext}| - (k+1-i)a, 1 \le i \le k \quad (98)$$

【０３５８】
ある実施の形態によれば、最後のパルスの後に除去されるべきサンプルの数は、以下の式に従って、式（９７）に基づいて計算することができる。
【数１４９】

$$\Delta_{k+1}^p = |s| - \Delta_0^p - \sum_{i=1}^{k} \Delta_i \quad (99)$$

【０３５９】
なお、実施形態によれば、$\Delta^p{}_0$、Δ_iおよび$\Delta^p{}_{k+1}$は正で、かつｓの符号が、サンプルが加えられるか除去されるかを決定する。
【０３６０】
複雑さを理由に、いくつかの実施形態では、整数個のサンプルを加えるかまたは除去することが望ましく、そのような実施形態においては、$\Delta^p{}_0$、Δ_iおよび$\Delta^p{}_{k+1}$が、たとえば、丸められ得る。他の実施形態では、たとえば波形内挿を用いる他のコンセプトも代替的または付加的に使用して、丸めを回避できるが、複雑さは増大する。
【０３６１】
以下では、実施形態に従うパルス再同期化のためのアルゴリズムについて式（１００）から式（１１３）を参照して説明する。
【０３６２】
実施形態によれば、このようなアルゴリズムの入力パラメータは、たとえば以下のとおりである。
Ｌ　フレーム長さ
Ｍ　サブフレームの数

T_p　最後に受信したフレームの終りのピッチサイクル長さ
T_{ext}　封じ込められたフレームの終りのピッチサイクル長さ
ｓｒｃ＿ｅｘｃ　上記のとおり最後に受信したフレームの終りから励振信号のローパスフィルタ処理された最後のピッチサイクルをコピーして作られた入力励振信号
ｄｓｔ＿ｅｘｃ　パルス再同期化についてここに記載のアルゴリズムを使用してｓｒｃ＿ｅｘｃから作られる出力励振信号。

【０３６３】
実施形態によれば、このようなアルゴリズムは、以下のステップの１以上または全部を含み得る。

【０３６４】
・式（６５）に基づいて、サブフレーム当たりのピッチの変化を計算する。
【数１５０】

$$\delta = \frac{T_{ext} - T_p}{M} \qquad (100)$$

【０３６５】
・式（１５ｂ）に基づき、丸められた開始ピッチを計算する。
【数１５１】

$$T_r = \lfloor T_p + 0.5 \rfloor \qquad (101)$$

【０３６６】
・式（６９）に基づき、加えられるべき（負の場合には除去すべき）サンプルの数を計算する。
【数１５２】

$$s = \delta \frac{L}{T_r} \frac{M+1}{2} - L\left(1 - \frac{T_p}{T_r}\right) \qquad (102)$$

【０３６７】
・励振ｓｒｃ＿ｅｘｃの構成された周期的部分における第１のT_r個のサンプルから、第１の最大パルスＴ［０］の場所を見つける。

【０３６８】
・式（７３）に基づき再同期化されたフレームｄｓｔ＿ｅｘｃにおける最後のパルスのインデクスを取得する。
【数１５３】

$$k = \left\lceil \frac{L - s - T[0]}{T_r} - 1 \right\rceil \qquad (103)$$

【０３６９】
・式（９４）に基づいて、連続するサイクルの間で加えるかまたは除去すべきサンプルのａ－Δを計算する。

【数１５４】

$$a=\frac{|T_r-T_{ext}|(L-s)-|s|T_r}{(k+1)\left(T[0]+\frac{k}{2}T_r\right)} \qquad (104)$$

【０３７０】

・式（９６）に基づいて第１のパルスの前に加えるかまたは除去すべきサンプルの数を計算する。

【数１５５】

$$\Delta_0^p=(|T_r-T_{ext}|-(k+1)a)\frac{T[0]}{T_r} \qquad (105)$$

【０３７１】

・第１のパルスの前に加えるかまたは除去すべきサンプルの数を丸めて、小数部分をメモリに維持する。

【数１５６】

$$\Delta_0'=\lfloor\Delta_0^p\rfloor \qquad (106)$$

$$F=\Delta_0^p-\Delta_0' \qquad (107)$$

【０３７２】

・２つのパルス間の各領域について、式（９８）に基づいて加えるかまたは除去すべきサンプルの数を計算する。

【数１５７】

$$\Delta_i=|T_r-T_{ext}|-(k+1-i)a, 1\le i\le k \qquad (108)$$

【０３７３】

・前回の丸めの時の残余の小数部分を考慮に入れて、２つのパルス間で加えるかまたは除去すべきサンプルの数を丸める。

【数１５８】

$$\Delta_i'=\lfloor\Delta_i+F\rfloor \qquad (109)$$

$$F=\Delta_i-\Delta_i' \qquad (110)$$

【０３７４】

・いくつかのｉについて、加えられたＦにより、$\Delta'_i>\Delta'_{i-1}$となる場合、これらの値をΔ'_iおよびΔ'_{i-1}に交換する。

【０３７５】

・式（９９）に基づいて、最後のパルス後に加えられるかまたは除去されるべきサンプルの数を計算する。

【数１５９】

$$\Delta'_{k+1} = |[s + 0.5]| - \sum_{i=0}^{k} \Delta'_i \qquad (111)$$

【０３７６】

・次いで、最小エネルギ領域の間で加えられるまたは除去されるべきサンプルの最大数を計算する。

【数１６０】

$$\Delta'_{max} = \max_i \Delta'_i = \begin{cases} \Delta'_k, \Delta'_k \geq \Delta'_{k+1} \\ \Delta'_{k+1}, \Delta'_k < \Delta'_{k+1} \end{cases} \qquad (112)$$

【０３７７】

・長さΔ’$_{max}$のｓｒｃ＿ｅｘｃにおける最初の２つのパルスの間の最小エネルギセグメントＰ$_{min}$［１］の場所を見つける。２つのパルスの間のすべての連続する最小エネルギセグメントについて、位置を以下の式により計算する。

【数１６１】

$$P_{min}[i] = P_{min}[1] + (i-1)T_r, 1 < i \leq k \qquad (113)$$

【０３７８】

・Ｐ$_{min}$［１］＞Ｔ$_{r}$なら、Ｐ$_{min}$［０］＝Ｐ$_{min}$［１］－Ｔ$_{r}$を用いて、ｓｒｃ＿ｅｘｃにおける最初のパルスの前の最小エネルギセグメントの場所を計算する。さもなければ、長さΔ’$_{0}$を有するｓｒｃ＿ｅｘｃにおける最初のパルスの前の最小エネルギセグメントＰ$_{min}$［０］の場所を見つける。

【０３７９】

・Ｐ$_{min}$［１］＋ｋＴ$_{r}$＜Ｌ－ｓならば、Ｐ$_{min}$［ｋ＋１］＝Ｐ$_{min}$［１］＋ｋＴ$_{r}$を用いて、ｓｒｃ＿ｅｘｃにおける最後のパルス後の最小エネルギセグメントの場所を計算する。さもなければ、長さΔ'$_{k+1}$を有するｓｒｃ＿ｅｘｃにおける最後のパルス後の最小エネルギセグメントＰ$_{min}$［ｋ＋１］の場所を見つける。

【０３８０】

・封じ込められた励振信号ｄｓｔ＿ｅｘｃにおいてパルスが１つしか存在しない場合には、すなわちｋ＝０の場合、Ｐ$_{min}$［１］のサーチをＬ－ｓに限定する。その場合、Ｐ$_{min}$［１］は、ｓｒｃ＿ｅｘｃにおける最後のパルス後の最小エネルギセグメントの場所を指す。

【０３８１】

ｓ＞０の場合、場所Ｐ$_{min}$［ｉ］（０≦ｉ≦ｋ＋１）で、信号ｓｒｃ＿ｅｘｃに個のサンプルを追加し、それをｄｓｔ＿ｅｘｃに格納し、さもなければ、ｓ＜０の場合、場所Ｐ$_{min}$［ｉ］（０≦ｉ≦ｋ＋１）で個のサンプルを信号ｓｒｃ＿ｅｘｃから除去して、それをｄｓｔ＿ｅｘｔに格納する。サンプルが加えられたり除去されるｋ＋２の領域が存在する。

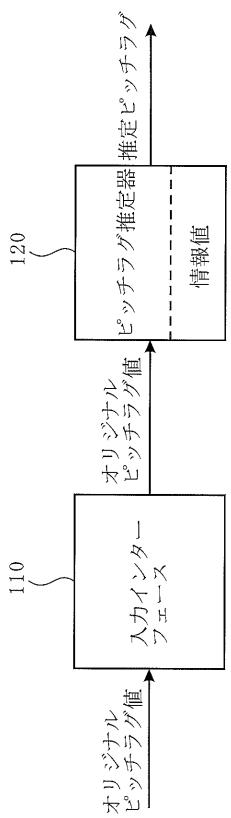
【０３８２】

図２ｃは、実施形態に従って音声信号を含むフレームを再構成するためのシステムを示す図である。このシステムは、上記の実施形態の１つに従って、推定ピッチラグを決定するための装置１００およびフレームを再構成するための装置２００を含み、フレームを再構成するための装置が、推定ピッチラグに依拠してフレームを再構成するよう構成される。推定ピッチラグは、音声信号のピッチラグである。

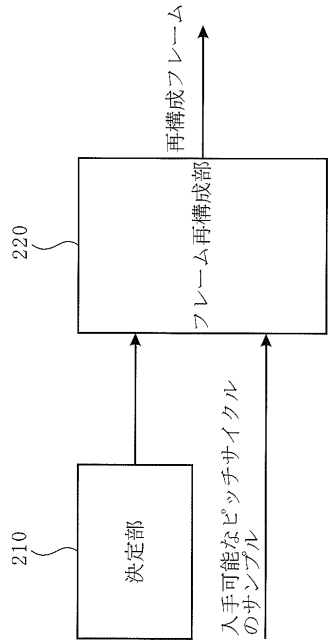
【０３８３】
ある実施形態において、再構成されたフレームは、たとえば、１以上の入手可能なフレームと関連し得るが、前記１以上の入手可能なフレームは、再構成されたフレームの１以上の先行フレームおよび再構成されたフレームの１以上の後続フレームのうちの１以上のフレームであり、１以上の入手可能なフレームは、１以上の入手可能なピッチサイクルとして１以上のピッチサイクルを含む。フレームを再構成するための装置２００は、たとえば、上記実施形態の１つによるフレームを再構成するための装置であり得る。
【０３８４】
いくつかの態様について、装置に関連して説明したが、これらの態様が対応する方法の説明も表すことは明らかであり、その場合、ブロックまたは装置が方法ステップまたは方法ステップの特徴に相当する。同様に、方法ステップに関連して説明した態様も、対応の装置の対応のブロックもしくはアイテムまたは特徴の説明を表す。
【０３８５】
発明の分解された信号は、デジタル記憶媒体に記憶されるかまたは無線送信媒体またはインターネット等の有線送信媒体等の送信媒体で送信され得る。
【０３８６】
特定の実装要件に依拠して、発明の実施形態を、ハードウェアまたはソフトウェアで実装することができる。実装は、それぞれの方法が実行されるように、プログラマブルコンピュータシステムと協働する（または協働することができる）電子的に可読な制御信号を記憶した、フロッピー（登録商標）ディスク、ＤＶＤ，ＣＤ，ＲＯＭ，ＰＲＯＭ，ＥＰＲＯＭ，ＥＥＰＲＯＭまたはフラッシュメモリ等のデジタル記憶媒体を用いて行うことができる。
【０３８７】
発明によるいくつかの実施形態は、本件明細書に記載の方法の１つを実行するように、プログラマブルコンピュータシステムと協働することができる、電子的に可読な制御信号を有する非一時的データキャリアを含む。
【０３８８】
一般に、本発明の実施形態は、プログラムコードを有するコンピュータプログラム製品として実現することが可能で、プログラムコードは、コンピュータプログラム製品がコンピュータ上で実行されると、方法の１つを実行するように動作する。プログラムコードは、たとえば、機械可読なキャリア上に記憶することができる。
【０３８９】
他の実施形態は、機械可読なキャリア上に記憶される、本件明細書に記載の方法の１つを実行するためのコピュータプログラムを含む。
【０３９０】
したがって、言い換えれば、本発明の方法の実施形態は、コンピュータプログラムがコンピュータ上で実行された時に、本件明細書に記載の方法の１つを実行するためのプログラムコードを有するコンピュータプログラムである。
【０３９１】
したがって、本発明の方法の他の実施形態は、本件明細書中に記載の方法の１つを実行するためのコンピュータプログラムを記録するデータキャリア（またはデジタル記憶媒体またはコンピュータ可読媒体）である。
【０３９２】
したがって、本発明の方法の他の実施形態は、本件明細書に記載の方法の１つを実行するためのコンピュータプログラムを表すデータストリームまたは信号のシーケンスである。データストリームまたは信号のシーケンスは、たとえば、インターネットを経由する等データ通信接続を経由して転送されるように構成され得る。
【０３９３】
他の実施形態は、たとえば、本件明細書に記載の方法の１つを実行するよう構成または適合されたコンピュータまたはプログラマブル論理装置等の処理手段を含む。

【０３９４】
他の実施形態は、本件明細書に記載の方法の１つを実行するためのコンピュータプログラムをインストールしたコンピュータを含む。
【０３９５】
いくつかの実施形態において、プログラマブル論理装置（フィールドプログラマブルゲートアレイ等）を使用して、本件明細書に記載の方法の機能の一部または全部を実行することができる。いくつかの実施形態では、フィールドプログラマブルゲートアレイが、本件明細書に記載の方法の１つを実行するために、マイクロプロセッサと協働し得る。一般に、方法は、なんらかのハードウェア装置により実行されることが好ましい。
【０３９６】
上記の実施形態は、本発明の原則を説明するに過ぎない。当業者には、本件明細書に記載の構成および詳細の変形例および修正例が明らかになることは当然である。したがって、特許請求の範囲によってのみ限定され、本件明細書における実施形態の記載および説明により提示される特定の詳細によっては限定されないことを意図する。

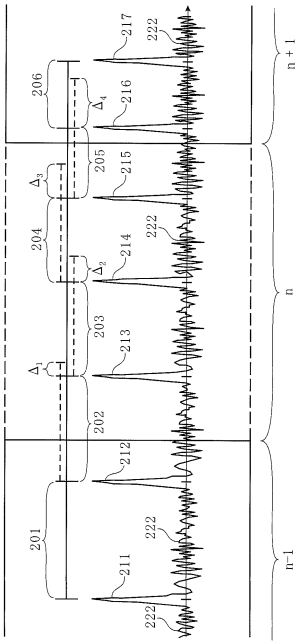
【図１】



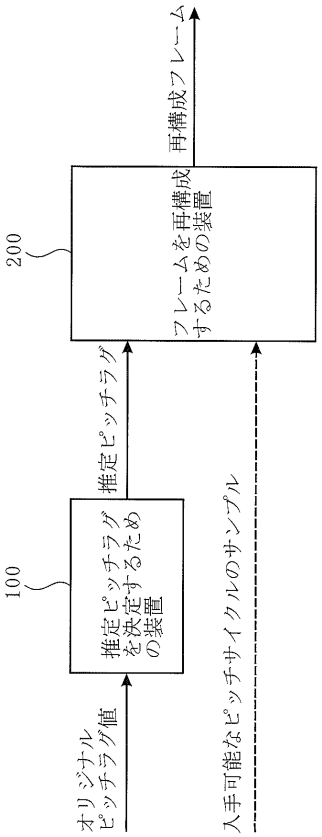
【図２Ａ】



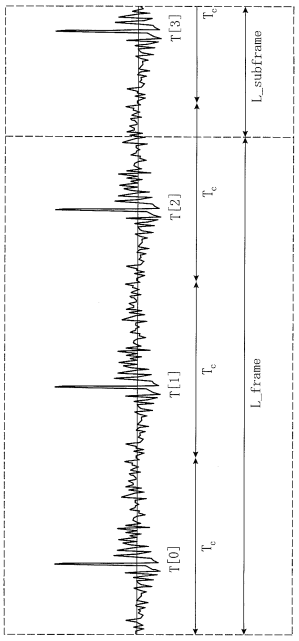
【図２Ｂ】


201
202
203
204
205
206
211
212
213
214
215
216
217
222
Δ1
Δ2
Δ3
Δ4
n-1
n
n+1

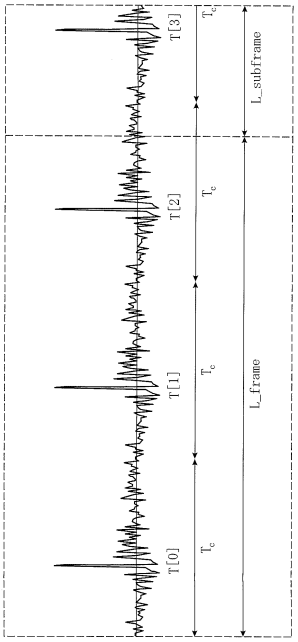
【図２Ｃ】


100
推定ピッチラグを決定するための装置
200
フレームを再構成するための装置
オリジナルピッチラグ値
推定ピッチラグ
再構成フレーム
入手可能なピッチサイクルのサンプル

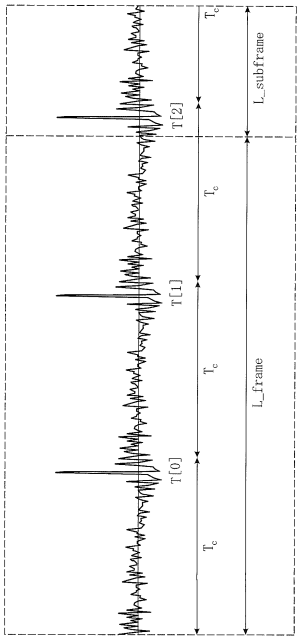
【図３】


T[0]
T[1]
T[2]
T[3]
Tc
L_frame
L_subframe

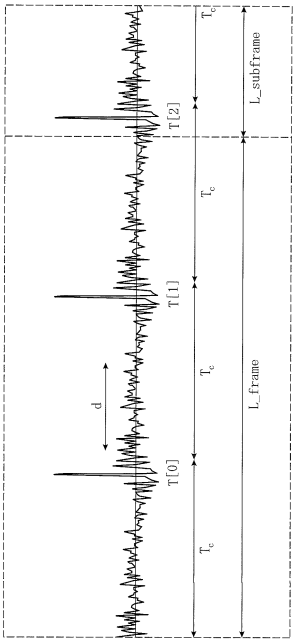
【図４】


T[0]
T[1]
T[2]
T[3]
Tc
L_frame
L_subframe

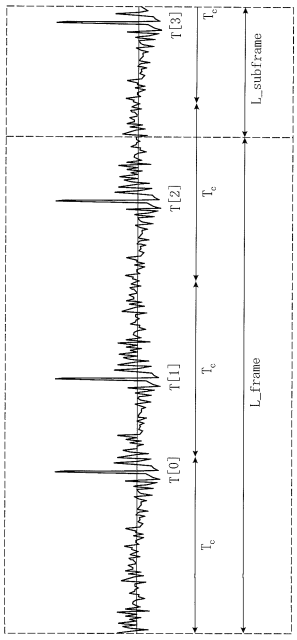
【図５】


T[2]
T[1]
T[0]
T_c
L_subframe
L_frame

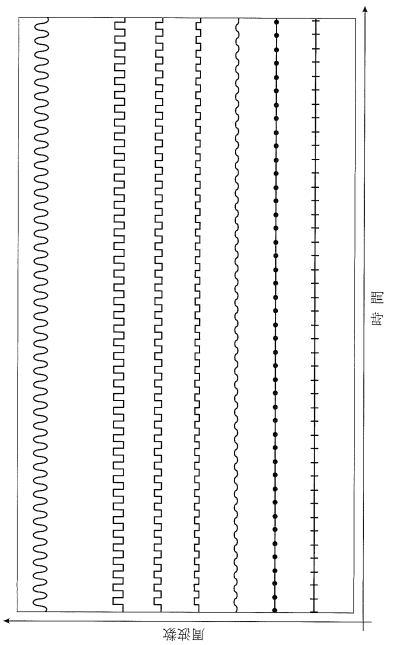
【図６】


T[2]
T[1]
T[0]
T_c
L_subframe
L_frame
d

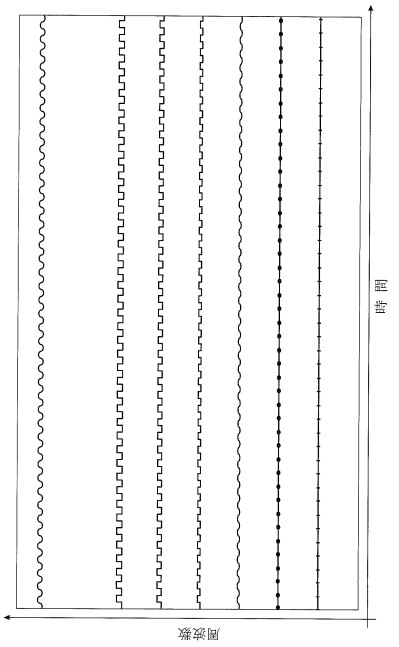
【図７】


T[3]
T[2]
T[1]
T[0]
T_c
L_subframe
L_frame

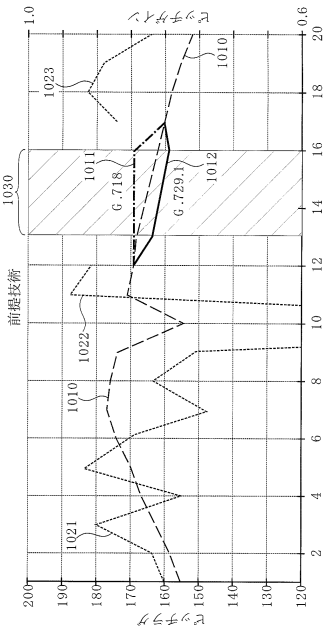
【図８】


周波数
時間

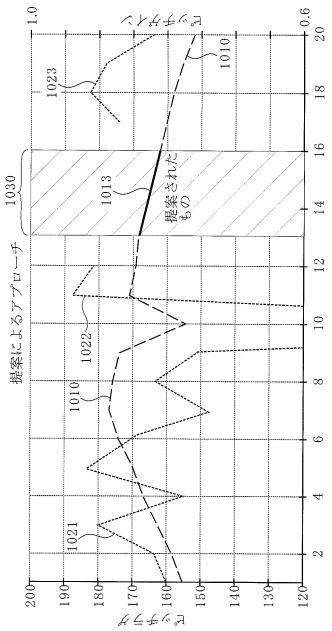
【図９】



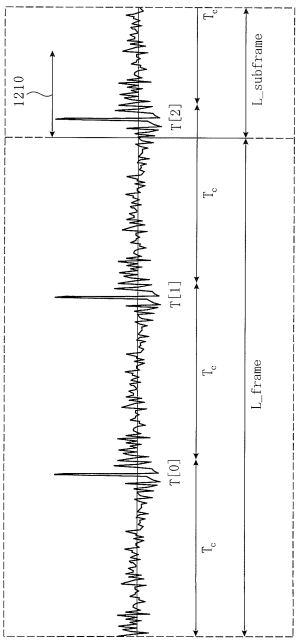
【図１０】



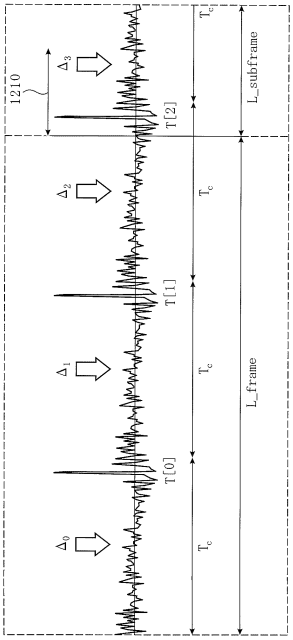
【図１１】



【図１２】



【図１３】


1210
Δ_2
Δ_2
Δ_1
Δ_0
T[2]
T[1]
T[0]
T_c
T_c
T_c
T_c
L_subframe
L_frame

フロントページの続き

(72)発明者　ルコント、　ジェレミー
ドイツ国　９０７６３　フュルト、　トゥルンシュトラーセ　７
(72)発明者　シュナーベル、　ミヒャエル
ドイツ国　９５１７９　ゲーロルツグリューン、　シュナイトベルクヴェーク　４アー
(72)発明者　マールコヴィッチ、　ゴーラン
ドイツ国　９０４２５　ニュルンベルク、　アーヘナー・シュトラーセ　１９
(72)発明者　デイツ、　マルティン
ドイツ国　９０４２９　ニュルンベルク、　ドイチュヘルンシュトラーセ　３７
(72)発明者　ノイゲバウア、　ベルンハルト
ドイツ国　９１０５８　エルランゲン、　フランツェンスバーダー　シュトラーセ　３

審査官　安田　勇太

(56)参考文献　国際公開第２００８／００７６９９（WO，Ａ１）
米国特許出願公開第２００４／００１７８１１（ＵＳ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１９／００５
Ｇ１０Ｌ　１９／００