

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4162933号  
(P4162933)

(45) 発行日 平成20年10月8日(2008.10.8)

(24) 登録日 平成20年8月1日(2008.8.1)

(51) Int.Cl.

F I

G 1 O L 19/12 (2006.01)

G 1 O L 19/12

B

G 1 O L 19/08 (2006.01)

G 1 O L 19/08

Z

請求項の数 22 外国語出願 (全 19 頁)

(21) 出願番号 特願2002-186971 (P2002-186971)  
 (22) 出願日 平成14年6月26日(2002.6.26)  
 (65) 公開番号 特開2003-122400 (P2003-122400A)  
 (43) 公開日 平成15年4月25日(2003.4.25)  
 審査請求日 平成17年6月13日(2005.6.13)  
 (31) 優先権主張番号 09/896,272  
 (32) 優先日 平成13年6月29日(2001.6.29)  
 (33) 優先権主張国 米国(US)

早期審査対象出願

(73) 特許権者 500046438  
 マイクロソフト コーポレーション  
 アメリカ合衆国 ワシントン州 9805  
 2-6399 レッドモンド ワン マイ  
 クロソフト ウェイ  
 (74) 代理人 100077481  
 弁理士 谷 義一  
 (74) 代理人 100088915  
 弁理士 阿部 和夫  
 (72) 発明者 アジット ブイ、ラオ  
 アメリカ合衆国 98052 ワシントン  
 州 レッドモンド ノースイースト 34  
 コート 166624 ナンバーエスエ  
 ス-302

最終頁に続く

(54) 【発明の名称】 低ビットレート CELP 符号化のための連続タイムワーピングに基づく信号の修正

(57) 【特許請求の範囲】

【請求項 1】

デジタル音声信号のフレームを圧縮するための準備処理をする方法であって、  
 該フレームについての線形予測残差であって不規則な間隔で配置された顕著なピークを  
 有する残差を生成するステップと、

オーバーラップしないで連続する一連のセクションであって各セクションが1つ以下の  
 顕著なピークを含む一連のセクションに、前記残差を分割するステップと、

一連の顕著なピークが、連続する一連のセクションにおいて規則的な間隔で配置されて  
 成る理想化信号を導出するステップと、

前記残差の各セクションを、該理想化信号の対応するセクションと関連付けるステップ  
 と、

各残差セクションについての線形連続ワープ輪郭を、各残差セクションについて最後にな  
 る可能性のあるサンプルラグ値のサブレンジ内で、各残差セクションについて最後にな  
 る可能性のあるサンプルラグ値のサブセットに基づいて、該ワープ輪郭に空間ジャンプま  
 たは不連続がないように計算するステップと、

前記計算されたワープ輪郭を前記残差の前記セクションに適用して、各残差セクション  
 内のいかなる顕著なピークも前記理想化信号の前記対応するセクションにおける前記顕著  
 なピークと揃うように前記残差を修正し、それによって該修正残差の顕著なピッチピーク  
 が規則的な間隔で配置され、該修正残差において前記残差のいかなるセクションの部分も  
 削除されたり繰り返されたりしないステップ

10

20

とを含むことを特徴とする方法。

【請求項 2】

前記生成するステップはさらに、  
前記フレームについての線形予測係数を抽出するステップと、  
前記フレームについての前記線形予測係数を補間して、前記フレームの複数のサブフレームについて線形予測係数を生成するステップと、  
各サブフレームについて予測残差を生成し、それによって前記フレームについての前記予測残差が一組のサブフレーム予測残差を有するステップとを含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記分割するステップはさらに、前記フレームを分析して整数ピッチ周期を確認するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 4】

前記確認するステップはさらに、オープンループ内で相関分析を使用するステップを含むことを特徴とする請求項 3 に記載の方法。

【請求項 5】

前記計算するステップはさらに、  
前記残差セクションの最初のサンプルについて最初のサンプルラグを確立するステップと、

前記残差セクションの最後のサンプルについて最後のサンプルラグの一組の候補を確認するステップであって、該一組の候補が、最後になる可能性のあるすべてのサンプルラグ値のサブレンジ内で、最後になる可能性のあるすべてのサンプルラグ値のサブセットからなるステップと、

各候補の最後のサンプルラグについての前記残差セクションの最初と最後のサンプル間の線形補間を実行し、一組の候補ラグ輪郭を生成するステップと、

各候補ラグ輪郭を前記残差セクションに適用して、一組の候補修正残差を得るステップと、

各候補修正残差と前記理想化信号の前記対応するセクション間の相関強度を計算して、一組の相関強度を生成するステップと、

該一組の相関強度に基づいて前記残差セクションについて最適な最後のサンプルラグを導出するステップと、

最初のサンプルラグから前記残差セクションについて導出された最適な最後のサンプルラグまでのセクション全体を線形補間することで、線形連続ワーブ輪郭を導出するステップとを含むことを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記線形連続ワーブ輪郭を導出するステップはさらに、

前記強度を導出するために使用される前記最後のサンプルラグに応じて、前記一組の相関強度をオーバーラップするサブセクションに分離するステップと、

各サブセクションを曲線として表すステップと、

各曲線の最大値であって、前記サブセクション内で前記強度を導出するために使用される前記最後のサンプルラグを含む可能なラグ値の範囲内で可能性のあるすべてのラグ値からなるグループから選択可能である前記最大値を計算するステップと、

前記サブセクションの前記曲線についての前記最大値に基づいて前記セクションについて前記最大相関強度を計算するステップとを含むことを特徴とする請求項 5 に記載の方法。

【請求項 7】

前記曲線は多項式であることを特徴とする請求項 6 に記載の方法。

【請求項 8】

前記多項式は二次関数であることを特徴とする請求項 7 に記載の方法。

【請求項 9】

計算されたワーブ輪郭の適用時に前記セクション内のいかなるサンプルの最大累積シフ

10

20

30

40

50

とも4つのサンプル位置未満であるように、各残差セクションについて最後になる可能性のあるサンプルラグ値のサブレンジが選択されることを特徴とする請求項1に記載の方法。

【請求項10】

デジタル音声信号のフレームを圧縮するために準備処理をする方法を実行するためのコンピュータ可読命令を有するコンピュータ可読媒体であって、

該フレームについての線形予測残差であって不規則な間隔で配置された顕著なピークを有する残差を生成するステップと、

オーバーラップしないで連続する一連のセクションであって各セクションが1つ以下の顕著なピークを含む一連のセクションに、前記残差を分割するステップと、

一連の顕著なピークが、連続する一連のセクションにおいて規則的な間隔で配置されて成る理想化信号を導出するステップと、

前記残差の各セクションを、該理想化信号の対応するセクションと関連付けるステップと、

各残差セクションについての線形連続ワープ輪郭を、各残差セクションについて最後になる可能性のあるサンプルラグ値のサブレンジ内で、各残差セクションについて最後になる可能性のあるサンプルラグ値のサブセットに基づいて、該ワープ輪郭に空間ジャンプまたは不連続がないように計算するステップと、

前記計算されたワープ輪郭を前記残差の前記セクションに適用して、各残差セクション内のいかなる顕著なピークも前記理想化信号の前記対応するセクションにおける前記顕著なピークと揃うように前記残差を修正し、それによって該修正残差の顕著なピッチピークが規則的な間隔で配置され、該修正残差において前記残差のいかなるセクションの部分も削除されたり繰り返されたりしないステップ

を含むことを特徴とするコンピュータ可読媒体。

【請求項11】

前記生成するステップはさらに、

前記フレームについての線形予測係数を抽出するステップと、

前記フレームについての前記線形予測係数を補間して、前記フレームの複数のサブフレームについて線形予測係数を生成するステップと、

各サブフレームについて予測残差を生成し、それによって前記フレームについての前記予測残差が一組のサブフレーム予測残差を有するステップとを含むことを特徴とする請求項10に記載のコンピュータ可読媒体。

【請求項12】

前記分割するステップはさらに、前記フレームを分析して整数ピッチ周期を確認するステップを含むことを特徴とする請求項10に記載のコンピュータ可読媒体。

【請求項13】

前記確認するステップはさらに、オープンループ内で相関分析を使用するステップを含むことを特徴とする請求項12に記載のコンピュータ可読媒体。

【請求項14】

前記計算するステップはさらに、

前記残差セクションの最初のサンプルについて最初のサンプルラグを確立するステップと、

前記残差セクションの最後のサンプルについて最後のサンプルラグの一組の候補を確認するステップであって、該一組の候補が、最後になる可能性のあるすべてのサンプルラグ値のサブレンジ内で、最後になる可能性のあるすべてのサンプルラグ値のサブセットからなるステップと、

各候補の最後のサンプルラグについての前記残差セクションの最初と最後のサンプル間の線形補間を実行し、一組の候補ラグ輪郭を生成するステップと、

各候補ラグ輪郭を前記残差セクションに適用して、一組の候補修正残差を得るステップと、

各候補修正残差と前記理想化信号の前記対応するセクション間の相関強度を計算して、一組の相関強度を生成するステップと、

該一組の相関強度に基づいて前記残差セクションについて最適な最後のサンプルラグを導出するステップと、

最初のサンプルラグから前記残差セクションについて導出された最適な最後のサンプルラグまでのセクション全体を線形補間することで、線形連続ワープ輪郭を導出するステップとを含むことを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【請求項 15】

前記線形連続ワープ輪郭を導出するステップはさらに、

前記強度を導出するために使用される前記最後のサンプルラグに応じて、前記一組の相関強度をオーバーラップするサブセクションに分離するステップと、

各サブセクションを曲線として表すステップと、

各曲線の最大値であって、前記サブセクション内で前記強度を導出するために使用される前記最後のサンプルラグを含む可能なラグ値の範囲内で可能性のあるすべてのラグ値からなるグループから選択可能である前記最大値を計算するステップと、

前記サブセクションの前記曲線ついての前記最大値に基づいて前記セクションについて前記最大相関強度を計算するステップとを含むことを特徴とする請求項 14 に記載のコンピュータ可読媒体。

【請求項 16】

前記曲線は多項式であることを特徴とする請求項 15 に記載のコンピュータ可読媒体。

【請求項 17】

前記多項式は二次関数であることを特徴とする請求項 16 に記載のコンピュータ可読媒体。

【請求項 18】

計算されたワープ輪郭の適用時に前記セクション内のいかなるサンプルの最大累積シフトも 4 つのサンプル位置未満であるように、各残差セクションについて最後になる可能性のあるサンプルラグ値のサブレンジが選択されることを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【請求項 19】

磁気的に読取り可能なディスク媒体を含むことを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【請求項 20】

光学的に読取り可能なディスク媒体を含むことを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【請求項 21】

被変調データ信号を有することを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【請求項 22】

揮発性のコンピュータ読取り可能なストレージを含むことを特徴とする請求項 10 に記載のコンピュータ可読媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、一般に音声符号化技法に関し、より詳細には、コードブック励起線形予測 (CELP、codebook excited linear prediction) などの低ビットレート符号化技術を介して信号を符号化する際の助けとなる、信号を修正するための技法に関する。

【0002】

【従来の技術】

今日の音声や対話の技術が高度に進んだ環境下においては、ある地点から別の地点へ、時

10

20

30

40

50

にはかなり離れた距離間で、おまけにしばしば限られた帯域のチャンネルを介して、人間の声を電子的に伝送することが必要または望ましい状況がたびたび起きる。例えば、現在、セル式電話のリンクを介してまたはインターネットまたはその他のデジタル電子ネットワークを介しての会話が普通のこととなっている。同様に、コンピュータのハードディスク上またはデジタル記録装置の揮発性または不揮発性メモリ内などに、人間の声をデジタル方式でストアすると便利なのがたびたびある。例えば、デジタル方式でストアした人間の声を、電話応答プロトコルまたは音声によるプレゼンテーションの一部として再生できる。

#### 【 0 0 0 3 】

デジタル音声の伝送および／またはストアに使用できるチャンネルおよび媒体が限られた容量しか有さないことが多く、この状況は日に日に増加の一途を辿っている。例えば、リアルタイムの音声または録音された音声とともに使用される高品質ビデオの到来によって、全部をストリーミングフォーマットで受信したりストアするためにダウンロードすることが可能なものなど、リアルタイムでのデジタルネットワークを介する電話／テレビ会議やリアルタイムではない高品質の音声／映像プレゼンテーションに対する需要が生み出されてきた。映像の内容により様々な伝送チャンネルおよび記憶媒体内の帯域や記憶容量が変位するため、音声および映像の両方を効率的に、かつ適切に圧縮することが緊急課題となっている。また別のシナリオでは、音声の極端な、かつ効果的な圧縮の必要性も増してきている。例えば、ますます輻輳の度合いが増しているセル式電話のリンクでは、チャンネルを介してより多くのユーザを取り扱うことができなければならないが、多数のユーザを取り扱うためのチャンネルの容量が変わっていない。

#### 【 0 0 0 4 】

動機はどうあれ、音声の圧縮が通信技術の重要な分野になってきている。使用可能なデジタル音声符号化技術は、圧縮を用いない非効率的な技術から圧縮率4倍以上を達成する効率的な技術へと、その範囲が広がってきている。一般に、既存の符号器は、波形符号器か音声符号器のどちらかに分類できる。波形符号器は、実際音声波自体を表そうとしており、通常高い圧縮率は得られない。音声符号器、またはボコーダは、単に合成された音声波をマッピングしようとするのではなく、人間の音声の源および癖を対象としており、したがって計算の複雑さが増しはするもののより高い圧縮率が得られるようになる。波形符号器は、一般に、癖のある人間の声、非言語音、高レベルの背景雑音に対して、より頑強である。

#### 【 0 0 0 5 】

広く使用されている多くの音声符号器は、線形予測符号化に基づいた技術を使っている。この線形予測符号化技術は、音声信号の各部分について、特定の信号によって励起した場合、元の音声信号部分によく似た信号を生成するデジタルフィルタが存在することを想定している。特に、線形予測技術を実施している符号器は、通常まず音声信号のスペクトル包絡線またはフォルマントを示す一組の係数を導出する。これらの係数に対応するフィルタが確立され、入力音声信号を予測残差に減らすため使用される。一般に、上述のフィルタは逆合成フィルタで、残差信号に対応する合成フィルタに入力することにより、元の音声信号に極めて近い信号を生成する。

#### 【 0 0 0 6 】

通常、フィルタ係数および残差は、音声信号を後におよび／または遠隔で再合成するために、伝送またはストアされる。フィルタ係数は、少ないストアスペースや、例えば1.5 kbpsの狭い伝送処理帯域しか必要としないが、予測残差は、複雑さの点で元の音声信号と似た高帯域信号である。したがって、音声信号を効果的に圧縮するためには、予測残差を圧縮する必要がある。この圧縮を得るために、コードブック励起線形予測、CELPの技術が使用される。CELPは、特定のベクトルを選択するために使用できる、それぞれが「コードブック」一組からなる、1つ以上のコードブックインデックスを利用している。各コードブックはベクトルの集合である。選択ベクトルは、基準化され合計されると、フィルタの応答を残差自体に最も近づける合成フィルタからの応答を生成するよう選ば

10

20

30

40

50

れる。C E L P復号器はC E L P符号器がアクセスしたのと同じコードブックにアクセスし、したがって符号器コードブックおよび復号器コードブックからの同じベクトルを確認するため、簡単なインデックスが使用できる。

【 0 0 0 7 】

容量または帯域が十分に使用可能な場合は、たとえどんなに複雑であろうとも、元の残差の最良近似を十分に行うことのできるコードブックを有することは困難ではない。しかし、使用可能な容量または帯域が減少してくると、当然C E L Pコードブックの能力も減少する。

【 0 0 0 8 】

残差信号を模倣するのに必要なビット数を減少させるための1方法は、その周期性を増加することである。つまり、元の信号の重複した部分を重複しない機能より、よりコンパクトに表わすことができるようにするということである。この原理を利用している1技術に、R C E L P ( R e l a x a t i o n C o d e b o o k E x c i t e d L i n e a r P r e d i c t i v e c o d i n g ) がある。この技術の一例が、W . B . K l e i j nらによる文献「The R C E L P S p e e c h c o d i n g A l g o r i t h m ( R C E L P 音声符号化アルゴリズム)」、Eur . T r a n s . O n C o m m u n i c a t i o n s 、 v o l . 4 、 n o . 5 、 p p . 5 7 3 - 8 2 ( 1 9 9 4 年 ) であり、これは開示されている全部を参照により本明細書に組み入れてある。特に、この文献は、修正されたピッチ周期の輪郭が合成ピッチ周期の輪郭と一致するように、残差信号の全セグメントを均一に進めるまたは遅らせる方法について記載している。この方法の問題点は、特定のワーピング方法のアーティファクトとして、元の信号の特定部分が削除されたり繰り返されたりする可能性があることである。特に、ある信号の2つの隣接するセグメントが累積圧縮シフトする場合、オーバーラップ部分の近くの元の信号の一部が、修正後の信号内で削除されてしまう可能性がある。同様に、2つの隣接するセグメントが累積拡張シフトする場合、オーバーラップ部分の近くの元の信号の一部が、修正信号内で繰り返されてしまう可能性がある。これらのアーティファクトが、最終的に再生される音声内で可聴ひずみを生成する。

【 0 0 0 9 】

別の技術でも同様の方法が提案されている。例えば、W . B . K l e i j nらによる文献「合成による分析音声符号器におけるピッチ予測子パラメータの補間 ( I n t e r p o l a t i o n o f t h e P i t c h - P r e d i c t o r p a r a m e t e r s i n A n a l y s i s - b y - S y n t h e s i s S p e e c h C o d e r s ) 」、音声処理のI E E E トランザクション ( I E E E T r a n s a c t i o n s o f S p e e c h a n d A u d i o P r o c e s s i n g 、 v o l . 2 、 n o . 1 、 p a r t I ( 1 9 9 4 年 1 月 ) を参照されたい。これは開示されている全部を参照により本明細書に組み入れてある。

【 0 0 1 0 】

過去に提案されたピッチワーピング方法はすべて、セグメントエッジのシフトによる品質低下など、同じような欠点を有しており、この欠点により元の信号の削除や繰り返しが起きていた。音声信号用の伝送ビットレートを減少させるが、信号の繰り返しや削除が起きず、またリアルタイムの通信が不可能となる程度にまで符号化計算を複雑にしたり遅延させたりしないようなフレームワーピング方法を提供することが望ましい。

【 0 0 1 1 】

【 発明が解決しようとする課題 】

本発明は、元の残差信号を合成輪郭と一致するよう修正するために、従来の技術によく見られるエッジシフトの影響を避けることができる、単に区分的連続タイムワープ輪郭ではない連続タイムワープ輪郭を用いる。特に、本発明で用いられるワープ輪郭は連続ワープ輪郭である。すなわち空間ジャンプや不連続がなく、そして隣接フレーム内の隣接端の位置を反転したりまたは過度に広げたりしない。

【 0 0 1 2 】

10

20

30

40

50

符号化アルゴリズムの複雑さを減少して実際上のおよび経済的な実施を可能にするために、最適線形シフトを二次またはその他の近似を使用して導出する。特に、理想化ワープ輪郭を判断するため本発明で利用するアルゴリズムでは、修正信号を合成信号に相互関連させるために、可能性のあるすべてのワープ輪郭を計算し利用する必要がない。一実施形態では、可能な輪郭のサブレンジ全体からの可能な輪郭のサブセットが計算される。次いでこれらの輪郭からの相対相関強度が、二次曲線またはその他のパラメトリック関数曲線上の点としてモデル化される。次いで最適ワープ輪郭が、これは恐らく計算されたサンプル点間のどこかに位置する１点によって表されるが、適切なパラメトリック関数を最大にすることによって計算される。二分法や区分的多項式モデリングなどその他の単純化技術も、本発明で使用することができる。

10

**【 0 0 1 3 】**

本発明の付加的な機能および利点については、以下に例として示す実施形態の詳細な説明および添付図を参照すれば明らかとなるであろう。

**【 0 0 1 4 】****【課題を解決するための手段】**

以下の説明では、それ以外のことを特筆していない限り、１つ以上のコンピュータによって実施されるオペレーションの動作および記号表示を参照しながら本発明について記載する。したがって、時折コンピュータ実行と呼ぶ、この動作 ( a c t s ) およびオペレーションには、構造化形式でデータを表す電気信号のコンピュータの処理ユニットによる操作が含まれることが理解できるであろう。この操作により、コンピュータのメモリシステム内のいろいろな場所でデータが変換または保持され、当業者に良く理解されている方式でコンピュータのオペレーションを再構成または変更する。データが保持されているデータ構造は、そのデータのフォーマットによって定義された特定のプロパティを有するメモリの物理領域である。しかし、本発明を以下に記載していくが、以下に記載する様々な動作およびオペレーションは、ハードウェア内でも実施可能であり、限定的なものではないことが、当業者なら理解できるであろう。

20

**【 0 0 1 5 】**

音声符号器は、高ビットレートの入力デジタル音声信号をより低いビットレート信号に圧縮し、次いでデジタルチャネル、例えばインターネットを介して伝送し、またはデジタルメモリモジュール、例えばハードディスクやＣＤ－Ｒ内にストアするよう動作が可能なソフトウェアモジュールである。伝送またはストアされたビットは、音声復号器によって復号化されたデジタル音声信号に変換される。音声符号器および復号器を、２つ合わせて音声コーデックと呼ぶことが多い。音声コーデックは、特に入力信号が人間の声の場合に、入力音声信号に最も近い可能な再構築を復号器で生成するよう設計されている。音声符号化で使用される最も一般的なパラダイムは、コードブック励起線形予測 ( C E L P ) である。C E L P 音声符号器は、短期予測およびコードブック探索の原理に基づいている。C E L P 符号化のコンセプトおよび関数については、読者の理解の助けとなるよう本明細書にて論じることとする。この論議は、当業界で知られているものとは異なる方式でC E L P 符号化を定義しようとするものではない。

30

**【 0 0 1 6 】**

どの音声符号器においても、人間の音声はその性質上複雑でありまた時間につれて変化するものであり、それを捕らえるために使用可能なビット数が少ないため、低ビットレートにおいて、そのタスクがますます困難かつ複雑になっている。本発明は、ストアまたは伝送のために要するビット数がより少なくて済むように、音声符号器によって入力デジタル音声信号を符号化する前に、それを修正するための新規な方法を提供する。信号の修正の目的は、再構築される信号の知覚品質に悪影響を及ぼさないで、入力音声信号の波形の構造を単純化することである。信号の修正後、修正された入力音声信号は、符号化するため音声符号器に表示される。修正された波形構造が単純なため、音声符号器は、信号符号化のタスクをより上手く、かつ効率的に実施することができる。上述したように、信号の修正は特に低いビットレートで有利である。

40

50

## 【 0 0 1 7 】

本明細書に記載する信号修正技術は、連続タイムワーピングのモデルに基づいている。上記の R C E L P の信号修正技術とは異なり、連続タイムワーピングでは、単なる区分的連続輪郭ではなく連続ワーピング輪郭を使用して、入力信号を修正する。その結果が、その波形が単純な構造を有し、その品質が元の入力信号の品質と実質的に同じである修正音声信号となる。

## 【 0 0 1 8 】

本発明を十分に理解するためには、コーデック技術の C E L P ファミリの基本のファセットを理解することが重要である。当業者には様々な C E L P 技術が周知であるが、読者の便宜のため本明細書に簡単に記載することとする。C E L P 符号化においては、復号化された音声信号が、時間によって変化する合成フィルタを通して励起信号をフィルタリングすることによって生成される。符号器が励起信号および合成フィルタについての情報を復号器に送信する。

## 【 0 0 1 9 】

C E L P とは波形を一致させる方法である。すなわち、提案された合成信号の、残差などのモデル化すべき信号との相関を介して、励起信号の選択を最適化する。したがって、符号器は入力音声信号の短いセグメントを評価し、各セグメントについて最も近いレプリカの生成を試みる。特に、符号器は、まず「コードベクトル」と呼ぶ特定の許容信号を組み合わせることによって一組の励起信号を生成する。このようにして生成された一組内の各励起信号は、合成フィルタを通して通過され、元の音声信号に最も近い、またはレプリカとして複製されるべき別の信号を生成するフィルタリングされた励起信号が選択される。このような探索手順に従って、符号器は、選択された励起信号を生成するため組み合わせられたコードベクトルについての情報と合成フィルタについての情報を復号器に送信する。通常、大部分のビットが、合成フィルタ励起信号の形成のためのコードベクトルについての情報を送信するのに必要であるが、合成フィルタパラメータ自体に必要なビット数は通常  $1.5 \text{ kb/s}$  未満である。したがって、C E L P は、例えば入力音声信号の複雑な性質を表すのに十分なコードベクトルのある、 $4 \text{ kb/s}$  を越える比較的高いビットレートで良く機能する。低いビットレートでは、許容コードベクトルの数が少ないため、再生される信号品質がかなり低下する。

## 【 0 0 2 0 】

音声の知覚的に重要な有声セグメントのための残差信号の主要特性は、おおよそ周期的なスパイクのシーケンスである。これらのスパイクは一般に、ピッチ周期で分けられて、大体均一に間隔をあけているが、これらスパイクの位置の規則性の点で小さなジッタのあることが多い。これらのジッタは、知覚的には重要ではないものの、低ビットレートの波形符号器の限られたビット数の大部分を消費する。

## 【 0 0 2 1 】

既に論じたように、R C E L P は、スパイクの場所を再調整して定期的に起きるようにするため残差信号を非連続的にワーピングすることにより、この変化を除去するよう試みていた。修正信号内のスパイクの場所についての情報を送信するのに必要なビット数が非常に少なく済むため、この方式による信号の修正は、低ビットレートの符号器のタスクを軽減する。残差修正の後、この修正信号は、予測フィルタの反転を通じて通過させることにより、音声領域内に変換し直される。

## 【 0 0 2 2 】

しかし、R C E L P による信号修正により、使用されたワーピング関数のプロパティが最適とは言えないため、音声品質が知覚的に劣化する。特に、R C E L P においては、それぞれが単一のスパイクを含む、元の残差信号のオーバーラップしている可能性のある部分が、修正残差信号を生成するためカットされおおよび一列に繋がってしまうこととなる。このカットされた部分がオーバーラップする可能性があり、また実際オーバーラップすることが多いため、残差信号の一部が修正残差信号に 2 度現れたり、また他の部分は 1 度も現れなくなったりする。



## 【 0 0 2 3 】

本発明は、改良されたワープ輪郭最適化方法を用いた本発明の一実施形態とともに、連続タイムワーピングアルゴリズムを利用することによって、上述の R C E L P 残差修正手順における望ましくないプロパティを克服している。要するに、本発明によるアルゴリズムは、まず R C E L P と同様に、単一スパイクを含む元の残差信号の各部分を確認する。しかし、R C E L P とは異なり、これらの部分はオーバーラップしておらずフレーム全体を覆っている。つまり、カットされた部分が連結している場合でも、元の残差信号が得られ、残差信号の一部が 2 度現れたり、一部が削除されたりしなくなる。

## 【 0 0 2 4 】

本来、R C E L P でのように部分を単にカットしたり動かしたりするのではなく、このアルゴリズムは、連続的および適応的タイムワーピングオペレーションにおいて、各部分を線形加速したり線形減速したりする。各部分をワープする目的は、修正残差信号内のスパイクが定間隔で分離され、それによってスパイクの位置を符号化するのに要するビットレートが減少し、R C E L P と同様の目標を達成し、しかも欠点がないことを確認することにある。後述するように、再生される音声の品質の低下を回避するため、加速または遅延の程度が制限されている。

## 【 0 0 2 5 】

添付の特許請求の範囲は本発明の特徴を詳細に記載しているが、本発明は、その目的および利点とともに、以下の詳細な説明と添付の図面から良く理解されるであろう。

## 【 0 0 2 6 】

## 【発明の実施の形態】

本発明の一般的な側面については上述したが、好ましい実施形態の詳細について、以下にさらに詳しく述べることとする。図 1 を参照すると、本発明の一実施形態による改良された低ビットレートの符号器を実施するためのアーキテクチャを例示してある。このシステムは、すべてカスケードにつないだ、ディジタイザ 1 2 1 と、予測フィルタまたは逆合成フィルタ 1 0 1 と、線形連続残差修正モジュール 1 0 3 と、合成フィルタ 1 0 5 と、C E L P 符号器 1 0 7 などの符号器とを備える。

## 【 0 0 2 7 】

予測フィルタ 1 0 1 は、デジタル化した音声信号 1 0 9 をディジタイザモジュール 1 2 1 から入力として受信する。どの音声をデジタル電気信号に変換するかについては様々な方法があり、当業者には周知である。したがって、このような技術については本明細書では詳細に論じないこととする。予測フィルタ 1 0 1 はまた、逆合成フィルタと呼ぶ場合があるが、L P C 係数および入力信号に基づいて残差信号 1 1 1 を生成するよう動作可能である。逆フィルタおよび残差などの線形予測符号化コンセプトについては、当業者は精通されていることであろう。残差 1 1 1 は、残差修正モジュール 1 0 3 に入力され、以下に詳細に論じる方式で、その信号を修正残差 1 1 3 に変換する。修正残差 1 1 3 は、その後合成フィルタ 1 0 5 に入力され、再生音声信号 1 1 5 を生成する。残差修正モジュール 1 0 3 によって実施される残差修正技術によって、励起または修正残差 1 1 3 が残差 1 1 1 から変化しているにもかかわらず、修正音声信号 1 1 5 が元の信号 1 0 9 と非常に似た音を発する。その後、C E L P 符号器モジュール 1 0 7 が、当業者に良く理解されている方式で、修正音声信号を符号化し、伝送またはストアするために符号化したビット 1 1 7 のストリームを出力する。

## 【 0 0 2 8 】

図 1 に示すモジュールのオペレーションを、図 2 および図 3、図 4 を参照しながら、以下に詳細に記載する。特に、図 2 は、顕著なピッチピーク 2 0 1 を有する単純化された波形 2 0 3、2 0 5、2 0 7、2 0 9、2 1 1 を示す。図 2 に示すピークのシフトを、理解しやすいように誇張してあることに留意されたい。実際のシフト量を制限する必要があるが、このことについては後に論じる。図 3 および図 4 は、音声信号を符号化するため、本発明の一実施形態で実施するステップを示す流れ図である。ステップ 3 0 1 では、アナログ音声信号 1 1 9 がディジタイザ 1 2 1 によって受信される。ステップ 3 0 3 では、ディジ

10

20

30

40

50

タイザ 121 が、8 kHz で信号をサンプリングし、そのサンプリングしたデジタル音声信号  $s(n)$  を得る。その後、ステップ 305 では、信号  $s(n)$  はディジタイザによって 160 サンプル (20 ms) の長さのオーバーラップしていないフレームにグループ化され、そのそれぞれがさらに 80 サンプル (10 ms) の長さの 2 つのオーバーラップしていないサブフレームに細別される。したがって、K 番目のフレーム内の信号が、 $s(160k) \dots s(160k + 159)$  によって与えられる。ステップ 307 で、フレームされサンプリングされた信号 109 が、ディジタイザ 121 から LPC 抽出装置 123 に送られる。

#### 【0029】

LPC 抽出装置 123 は、当業者に周知の方式で入力信号に対応する線形予測係数を計算するよう動作する。特にステップ 309 では、LPC 抽出装置 123 は、相関分析を実施しレビンソン - ダービン (Levinson - Durbin) アルゴリズムを実行することにより、各フレームのための一組の 10 次程度 (tenth order) の線形予測係数を抽出する。K 番目のフレーム  $a_k(j)$ 、 $j = 1, \dots, 10$  内の最適線形予測係数は、ステップ 311 で補間され、各サブフレーム内で一組の LP 係数  $a_{ks}(j)$ 、 $j = 1, \dots, 10$  を生成する。上式で、 $s = 0, 1$  は、それぞれ第 1 のサブフレーム、第 2 のサブフレームに対応する。この補間は、LP 係数をラインスペクトル周波数 (LSF) 領域内に変換し、LSF 領域へ線形補間し、その補間したサブフレーム LSF 係数を LP 係数に変換し直すことによって実施できる。ステップ 313 では、サブフレーム LP 係数  $a_{ks}$  は、予測フィルタ 101 によって使用され、当業者に周知の方式で残差信号 111 を生成する。K 番目のフレーム内の残差 111 は、 $r(n)$ 、 $n = 160k \dots 160k + 159$  で表される。

#### 【0030】

残差信号 111 の顕著な特性を、図 2 に波形 203 で示す。特に、有声セグメントについては、大体周期的だが不規則に配置されているピークまたはスパイク 201 のシーケンスによって、残差 203 が優勢となっている。これらのスパイクは、通常有声音の生成の過程に、音域を励起する声門音の振動 (pulses) を表す。隣接スパイク間の時間間隔は、ピッチ周期に等しい。人間の音声は、通常約 2.5 ms から 18.5 ms の間のピッチ周期を有する。スパイク間の間隔は普通一定ではなく、小さい不規則性またはジッタを示す。

#### 【0031】

ステップ 315 から 333 では、残差修正モジュール 103 のオペレーションについて記載する。ステップ 315 では、残差修正モジュール 103 が残差信号 111 を受信し、現在のフレーム、つまり k 番目のフレームの整数ピッチ周期を判断する。このピッチ周期は、当業界で周知の技術の中のいずれかによって判断できる。本実施形態において使用可能な 1 つの技術は、オープンループ内で相関分析を用いるものである。どのような方法を使用するにしても、ピッチが倍加 (pitch doubling) するなどの望ましくないアーティファクトを避けるため、十分な注意を払う必要がある。

#### 【0032】

ステップ 317 では、フレームピッチ周期のサンプル線形補間によるサンプリングを、以下のように実施する。

$$c(n) = p(k) * ((n - 160k) / 160) + p(k - 1) * (1 - ((n - 160k) / 160)), n = 160k \dots 160k + 159$$

関数  $c(n)$  は、フレームの開始点での  $p(k - 1)$  からフレームの終点での  $p(k)$  までの直線として表される。これは、現在のフレーム内のすべてのサンプルに対する滑らかに変化するピッチ周期 (浮動小数点) を表す。

#### 【0033】

ステップ 319 では、関数  $c(n)$  は、 $c(n)$  の各値を 0.125 に最も近い倍数に丸めることによって形成される。実際、 $c(n)$  は  $1/8$  の倍数であり、したがって  $8 * c(n)$  は 8 倍にオーバーサンプリングされた信号領域内の整数ピッチ周期である。本明

10

20

30

40

50

細書では、 $c(n)$ を望ましいピッチ輪郭と呼ぶ。この理想化輪郭に一致するため残差を修正することによって生じた効率の良さには著しいものがある。例えば、このような輪郭を有するフレームのピッチ周期は、非常に少ないビット数を使用して伝送でき、また復号器がピッチ輪郭を導出するためピッチを使用することができ、次いで現在のフレームのためのピッチスパイクの場所を推定するために、前のフレームからのスパイクの場所とともにピッチ輪郭を使用することができる。

#### 【0034】

つぎのプロセスは、復号器を模倣して、前のフレームのピッチ輪郭および修正残差に基づいて現在のフレーム残差内でスパイクの場所の再構築を試みることである。実際の復号器は、通常前のフレームの修正残差についての情報にアクセスしないが、前のフレームを再構築するために使用される励起信号にアクセスする。したがって、特定のフレームの励起信号内のスパイクが、そのフレームの修正残差内のスパイクと位置を合わせるので、前の励起信号には復号器を使用し、前の修正残差には符号器を使用することが問題とはならない。

#### 【0035】

現在のフレーム内のスパイクの位置を予測するために、残差修正モジュール103が、ピッチ輪郭を使ってステップ321内にある前のフレームの修正残差を遅らせ、修正のためのターゲット信号 $r_1(n)$ を生成する。例示した $r_1(n)$ の波形が、図2の要素211に示してある。このタイムワーピング関数は、ピッチ輪郭 $c(n)$ が0.125の倍数であるため、切り捨て $\text{sinc}(x)$ インパルス応答および90%通過帯域を有する標準補間フィルタを使用して、8倍オーバーサンプリングした領域内で動作する。特に、8倍オーバーサンプリングは、前のフレーム内の修正残差 $r(n)$ の補間されたサンプルを得るため用いられ、以下のようなオーバーサンプリングした信号に到達する。

$$r(n * 0.125), n = 160 * 8 * (k - 1) \dots 160 * 8 * (k - 1) + 1279$$

$r$ のサンプルインデックスは0.125の倍数であり、オーバーサンプリングした状態を表している。次いで、ターゲット信号 $r_1(n)$ を得るため、以下のように遅延線オペレーションを実行する。

$$r_d(n * 0.125) = r_d(n * 0.125) \quad n = 160 * 8 * (k - 1) \dots 160 * 8 * (k - 1) + 1279$$

$$r_d(n * 0.125) = r_d(n * 0.125 - C(\text{INT}(n * 0.125))), \\ n = 160 * 8 * k \dots 160 * 8 * k + 1279$$

$$r_1(n) = r_d(n), n = 160 * k \dots 160 * k + 159$$

上式で、 $\text{INT}(x)$ は、 $x$ 、浮動小数点数に最も近い整数を表し、 $r_d()$ は、中間信号を表す。

#### 【0036】

復号器が、前のフレームの励起信号上で同じ遅延線オペレーションを実行することに留意されたい。

#### 【0037】

ターゲット信号211内に表されている理想化ピッチスパイクの場所を計算した後、符号器は次いで、実際の残差内のスパイクを、 $r_1(n)$ 内のスパイクと一致させるため再配置できる。最初にステップ323で、残差修正モジュール103が、未修正残差信号203を分析し、低エネルギー領域によって取り囲まれている1つの顕著なピークを有する信号の明確なセグメントを確認する。図2の要素205に、結果として生じる波形が例示してある。セグメントに分割された、信号の部分間には隙間のないことが好ましい。言い換えれば、要素205の部分どうしがこの段階で再び一列に繋がる場合、その結果未修正残差203が生じる。好ましくは、この残差203は知覚的に見て目立たない程度に低いエネルギー点でのみカットされる。次いでステップ325で、符号器がターゲット信号の1セクションを未修正残差の適切な部分に関連付ける。

#### 【0038】

ステップ 3 2 7 で、最適なワーピング関数を用いた修正によって、残差 2 0 3 のセグメント内の顕著なスパイクまたはピークが、ターゲット信号 2 1 1 の関連セクション内の顕著なスパイクまたはピークに位置合わせするよう、残差修正モジュール 1 0 3 が、未修正残差の確認されたセクションのための最適ワーピング関数を計算する。図 5 を参照すると、残差の各セクションのための最適ワーピング関数を計算するためにとったステップが例として示してある。特に図 5 は、残差信号 2 0 3 と修正残差 2 0 9 間のサンプル毎の遅延を表すラグ (lag) 輪郭  $l(n)$  の導出を示す。量  $l(n)$  は、修正残差サンプル  $r(m)$  がオーバーサンプリングされた領域内で  $l(m)$  によって遅延される残差信号サンプルと等しくなるような、0.125 の倍数である。つまり、

$$r(m) = r(m - l(m))$$

10

最適ワープ輪郭を見つけ出す場合の問題は、最適ラグ輪郭  $l(n)$  をを見つけ出す問題に減じる。

#### 【0039】

ステップ 4 0 1 では、現在の当該セクションの一番最初のサンプル用のラグ  $l_f$  が、前のセクションの一番最後のサンプル用のラグと等しいように設定され、現在のセクションの最後のサンプルのラグ  $l_l$  用の一組の候補が確認される。特に、最後のサンプルのラグ  $l_l$  用の一組の  $2K + 1$  候補が、 $\{l_f - K, l_f - K + 1, \dots, l_f + K\}$  などの候補範囲内で確認される。K の値は、使用可能な計算能力、音声サンプルの周期性、 $l_f$  の値などのパラメータに基づいて選択される。K の典型的な値は、0、1、2、3、または 4 である。上記の方程式で示す候補範囲は  $l_f$  の周りに対称的に入っているが、必ずしもこ

20

#### 【0040】

残差の少量シフトセクションが、再生された信号の知覚品質に悪影響を与えることはないが、シフト量が大きい場合は知覚的に悪影響となる可能性がある。したがって、前のセクションまたは部分がシフトした結果、累積されたシフトを含む 3 つの元の (オーバーサンプリングされていない) サンプル増分など、少量だけサンプルがシフトするように量を制限することが望ましい。したがって、2 つのサンプル位置の均等物によって、前の部分内の最後のサンプルが遅延した場合は、次いで現在の部分の最後のサンプルは、1 つのサンプル位置の均等物を越えてさらに遅延してはならない。さもないと、全シフト量が元の場所から 3 つのサンプル位置を越えることとなる。この問題点に対する解決法は、シフトが望ましい範囲を越えないように K の値を制限するか、または候補の非対称範囲を使用するかのどちらかである。したがって、上記の例においては、1 つのサンプルを越える遅延は禁じられているが、非対称的に分布している候補ラグ値が使われている場合は、5 つのサンプル位置の加速が認められる。

30

#### 【0041】

可能性のあるすべてのラグ候補を評価するのに必要な計算能力が押さえられているため、候補の組内には可能性のあるすべてのラグ候補よりも少数のラグ候補しかないことに留意されたい。逆に、現在のセクション内の最後のサンプル用に可能なラグ値のサブセットのみが候補として使用される。候補範囲外のラグ値はこの組内に含まれておらず、また候補ラグ値間にある値も同様である。したがって、最後のサンプル用の最適ラグ値 (および結果として生じるラグ輪郭) が、候補の組自体内に含まれてもいない可能性があるが、候補範囲内に入っていることが好ましい。

40

#### 【0042】

次いで、ステップ 4 0 3 では、一組の  $2K + 1$  候補ラグ輪郭を生成するため、符号器が、ステップ 4 0 1 で確認した各候補ラグ値用の現在のセクションの最初と最後のサンプル間の線形補間を実施する。候補ラグ輪郭は、最初と最後の値がそれぞれ  $l_f$ 、 $l_l$  である線形関数を表している。ここで、 $l_l$  とは候補値である。ステップ 4 0 5 では、一組の  $2K + 1$  候補修正残差を得るために、各候補ラグ輪郭が残差信号に適用され、ターゲット信号  $r_1(n)$  2 1 1 と各候補修正残差の間の相関が、ステップ 4 0 7 で計算される。

#### 【0043】

50

ステップ409では、相関強度が最後のサンプルラグ値に応じて二次方程式でモデル化され、最後のサンプル用の最適ラグ値が得られる。特に、各候補修正残差用の相関強度が、図6のグラフ内のプロット点で示すように、関連する最後のサンプルラグ値候補に応じてプロットされる。次いで、そのプロット点が組に分割され、それぞれの組は3点から構成される。隣接する組間に1点のオーバーラップがある。したがって、 $2K+1$ プロット点はこのようにKに分割され、それぞれ3点の組とオーバーラップする。例えば、7点については3組となる。3つの連続するプロット点の各組は、二次関数に従ってモデル化される。例えば図6では、3つの二次モデリング関数が501、503、505として例示してある。関連する3点の最初から最後まで範囲内の各二次関数の最大が得られ、次いでセクション全体の最大が計算される。したがって、正の二次関数、すなわち上向きにくぼんだところ、および点の単調構成については、最大相関値が終点の1つに位置する。一般に、所与の3点の組の最大が必ずしもこの3点のいずれかに位置するとは限らないが、この間のどこかに位置することが多いことに留意されたい。したがって、セクション全体のための最適ラグ値は、ラグ $l_1$ 用の候補の組内に入っていない値である可能性がある。

#### 【0044】

図6のプロットは、本発明の一実施形態によるステップをグラフで示すため使用しているが、本明細書で使用する用語、「プロット」または「プロットイング」では、具体的または目に見えるグラフを作成する必要はない。逆に、これらの用語は、単に量と量との間の関連の創出を暗黙に定義するにすぎない。黙示的なものとしては、使用されている軸が、図6に示す量に対するそれぞれ異なるパラメータであり、明示的なおよび実際のものとしては、グラフィカルプログラムデータ構造の中にあり、また仮想的なものとしては、適切な関係が導出できるメモリ内の一組の数の中にある。したがって、たとえこのような関係が明確であろうとも、これらの用語は、単に表示した量と量との間の関係の創出を表わすにすぎない。

#### 【0045】

現在の相関プロット用のすべての二次方程式の最大は、適切な二次方程式を用いた最後のサンプル用のラグ値に関連付けられており、この値が最適な最後のサンプルラグ値である。点の組をモデル化するために二次関数を使用したり、3点を使用したりする必要はない。例えば、これらの組は3点を越える量を含むことができ、モデリング関数は、受容可能な複雑さのレベルに応じて、どの程度の多項式にもなる可能性がある。点の単調シーケンスについては、一番高い終点が簡単に判断でき、シーケンスの最大を表すため、そのシーケンスを多項式などでモデル化する必要はないことに留意されたい。

#### 【0046】

現在の顕著なピーク含有セクションまたはセグメントの最後のサンプル用の最適ラグ値を判断した後、残差修正モジュール103が、ステップ411で、 $l_f$ からステップ409で計算した最適 $l_1$ までのセクション全体を線形に補間することによって、対応するラグ輪郭を導出する。図4のステップ329では、図5のステップ411で計算されたラグ輪郭が、上記のように残差に適用される。すなわち、 $r(n) = r(l_1(n))$ 。

#### 【0047】

最後に、ステップ331では、現在のフレーム内に分析しシフトすべき部分がこれ以上ないかどうか判断する。もしあれば、オペレーションのフローがステップ325に戻る。そうでない場合は、現在のフレームのためのプロセスは、ステップ333で終了する。図2の要素207には、修正残差209のワープされたセクションが、理解しやすいように別に示してある。波形209として示した修正残差113は、最後に合成フィルタ105への入力として設けられ、元の音声信号の再生を行う。この再生はジッタのあるピッチピークよりむしろ規則的なピッチピークを有する。この観点から、信号は、通常のCELPなどの技術を使用して処理される。しかし、ピッチ構造の周期性が増すため、未修正信号を符号化するのに必要なビットレートと比べて、信号を符号化するのに必要なビットレートが大幅に減少する。

#### 【0048】

フレームを処理した後、後続するフレームに対して処理が開始される。無声のセグメントの場合は、通常ピッチピークがなく、したがって本明細書に記載した方法を適用する必要はない。無声の間隔の間は、アルゴリズム内のすべての量がリセットされる。例えば、累積シフトの表示がゼロにリセットされる。有声の通話が開始されたら、最初の有声フレーム  $k$  は、前のフレームのピッチ値、 $p(k-1)$  がこのフレーム内で知られていないため、特別なケースとして扱われる。この特別フレーム  $k$  のピッチ輪郭は、フレームのピッチ値、 $p(k)$  に等しい一定の関数に設定される。この手順のこれ以降の部分は、規則的なフレームのものと同様である。

#### 【0049】

現在のフレーム内の音声信号の所与のセクションまたは部分用の最適ラグ値  $l_f$  および関連ラグ輪郭を計算するために、多項式モデリング以外の方法も本発明に使用できることに留意されたい。可能なラグ値の実質的なサブセット、例えば可能性のあるすべてのラグ値の半分が相関値を生成するために使用されることが本発明には重要なことである。なぜなら、このことによって、最適ラグ輪郭を見つけ出すのに要する計算コストが大幅に減少するからである。したがって、すべての、または大部分の可能なラグ値を試みることなく、最適ラグ値を見つけ出すのに、二分法などの代替技術を使用することができる。二分法技術では、2つのラグ候補値およびそれらの関連相関強度を確認する必要がある。より高い相関を有するラグ候補および2つのラグ値間にある新しいラグ候補が、二分法プロセスを繰り返すための終点として使用される。このプロセスは、所定の回数を繰り返した後、または所定の閾値を越える相関強度を有するラグ値が見つかった時に、終了することができる。

#### 【0050】

本明細書に記載した方法によって生じる連続線形ワープ輪郭が、図7に示してある。特に、連続線形ワープ輪郭601は黒い実線で示してあり、また従来の技術RCELP技術で使用される不連続輪郭603は破線で示してある。両輪郭とも、元の時間(ワープ前)対修正時間(ワープ後)に応じてプロットされた信号サンプルのための点の組に沿って引いた線で表す。したがって、輪郭601内の各直線セグメントおよび輪郭603の各個々の部分は、それぞれの技術に従ってワープされた元の残差のセクションを表す。RCELP技術では、よくセクションが削除されたりオーバーラップすることが分かる。一方、本発明の連続線形ワープ輪郭では、オーバーラップや削除は起こらない。逆に、連続線形ワープ輪郭601は、傾斜の点で不連続を含むが、位置の点では単なる区分的な連続ではない連続である。特に、領域605は、ワープ輪郭603の2つの部分によって占有されているが、セクション607は同じ輪郭についてのデータを有さない。一方、信号スペース全体は、本発明による輪郭601によってオーバーラップや削除もなく占有されている。

#### 【0051】

隣接セグメント用のワープ輪郭601は、各セグメントに必要な加速または減速に応じて、同じ傾斜または異なる傾斜を有する可能性があることに留意されたい。これと対称的に、RCELP輪郭603の各セクションの傾斜は、一定である。これは、RCELPは信号のセクションをシフトするが、各セクション内の時間変換係数(time scale)を変化させないことによる。したがって、本発明による方法は、セクションエッジで望ましくない時間変換係数の不連続を生成することなく、各セクションのピークが望ましい場所にシフトするように、線形連続方式で各セクション内で時間変換係数をワープすることが分かる。

#### 【0052】

必ずしも必要ではないが、コンピュータで実行するプログラム「モジュール」などの命令を使用して、本発明を実施することができる。一般に、プログラムモジュールは、特定のタスクを実行しまたは特定の抽象データタイプを実施するルーチン、オブジェクト、構成要素、データ構造などを含む。プログラムは、1つ以上のプログラムモジュールを含むことができる。

#### 【0053】

本発明は、セル式電話、パーソナルコンピュータ（ＰＣ）、ハンドヘルド装置、マルチプロセッサシステム、マイクロプロセッサベースのプログラム可能な民生用電子機器、ネットワークＰＣ、ミニコンピュータ、メインフレームコンピュータなどを含む様々なタイプの機械、または本明細書に記載したように音声信号を符号化または復号化し、信号をストア、検索、送信、受信するために使用できるその他の機械で実施できる。本発明は、タスクが、通信ネットワークを通してリンクされている遠隔構成要素によって実行される、分散型コンピューティングシステム内で用いることができる。

#### 【 0 0 5 4 】

図 8 に、本発明の一実施形態を実施するためのシステムを例として示すが、これには、コンピュータ装置 700 などのコンピュータ装置が含まれる。最も基本的な構成では、コンピュータ装置 700 は、通常少なくとも 1 つの処理ユニット 702 およびメモリ 704 を備える。コンピュータ装置の正確な構成およびタイプによって、メモリ 704 は、（ＲＡＭなどの）揮発性、（ＲＯＭ、フラッシュメモリなどの）不揮発性、またはこれら 2 つの組み合わせとなる。この最も基本的な構成は、図 8 の線 706 内に示す。さらに、コンピュータ装置 700 には、追加の特徴／機能も有することができる。例えば、コンピュータ装置 700 に、磁気または光のディスクやテープを含む追加記憶装置（取外し可能および／または取外し不可能）も備えることができるが、必ずしもこれに限定されるものではない。このような追加記憶装置は、図 8 に取外し可能記憶装置 708 および取外し不可能記憶装置 710 で示してある。コンピュータの記憶媒体は、コンピュータ読み取り可能命令、データ構造、プログラムモジュール、その他のデータなどの情報をストアする方法または技術で実施される、揮発性および不揮発性、取外し可能および取外し不可能な媒体を含む。メモリ 704、取外し可能記憶装置 708 および取外し不可能記憶装置 710 は、すべてコンピュータの記憶媒体の例である。コンピュータの記憶媒体は、所望の情報をストアするのに使用しコンピュータ装置 700 からアクセス可能な、ＲＡＭ、ＲＯＭ、ＥＥＰＲＯＭ、フラッシュメモリまたはその他のメモリ技術、ＣＤＲＯＭ、デジタル多目的ディスク（ＤＶＤ）またはその他の光学記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置またはその他の磁気記憶装置、またはその他の媒体を含むが、これらに限定されるものではない。このようなコンピュータの記憶媒体のどれもが、コンピュータ装置 700 の一部となり得る。

#### 【 0 0 5 5 】

コンピュータ装置 700 はまた、他のコンピュータ装置と通信するための 1 つ以上の通信接続 712 も備える。通信接続 712 は、通信媒体の例である。通信媒体は、通常搬送波やその他の転送メカニズムなどの被変調データ信号内のコンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータを実行し、情報配信媒体を含む。用語「被変調データ信号」とは、信号内で情報を符号化する方式で設定または変更した 1 つ以上の特性を有する信号である。通信媒体は、ワイヤードネットワークや直接ワイヤード接続などのワイヤード媒体、および音響、ＲＦ、赤外線などのワイヤレス媒体を含むが、これらは例として示したものにすぎず、限定的なものではない。上述したように、本明細書に使用する用語、コンピュータ可読媒体は、記憶媒体および通信媒体の両方を含む。

#### 【 0 0 5 6 】

コンピュータ装置 700 は、キーボード、マウス、ペン、音声入力装置、タッチ入力装置などの、1 つ以上の入力装置 714 も含む。ディスプレイ、スピーカ、プリンタなどの、1 つ以上の出力装置 716 も含まれる。これらすべての装置は当業界には周知のものであり、本明細書で詳細に論じる必要はない。

#### 【 0 0 5 7 】

本発明の原理を適用することのできる多くの実施形態について、図面に示した本明細書に記載する実施形態は、例示の目的で記載してあるのにすぎず、本発明の範囲を限定するものと解釈されるべきではない。例えば、ソフトウェア内に例として示した実施形態の要素はハードウェア内でも実施でき、またその逆も同様であり、あるいは例示した実施形態は、本発明の趣旨から逸脱することなく、配置上のおよび細部の修正が可能であることを、

10

20

30

40

50

当業者なら理解されるであろう。したがって、本明細書に記載した本発明は、以下の請求項およびその均等物の範囲内に含めることのできる、すべてのこのような実施形態を予想するものである。

【図面の簡単な説明】

【図 1】本発明の一実施形態を実施できる符号器の一例を示す構造図である。

【図 2】本発明の一実施形態における信号のセグメンテーション、タイムワーピング、再構築を示す波形略図である。

【図 3】本発明の一実施形態における信号の修正を行うためのステップを示す流れ図である。

【図 4】本発明の一実施形態における信号の修正を行うためのステップを示す流れ図である。 10

【図 5】本発明の一実施形態における最適ラグ輪郭を計算するためのステップを示す流れ図である。

【図 6】最適な最後のサンプルラグを確認するため本発明の一実施形態において使用する最後のサンプルラグ値に応じて、相関強度のプロットを示す概略特性図である。

【図 7】従来の技術に従って、および本発明の一実施形態に従って、ワープ輪郭を示す特性図である。

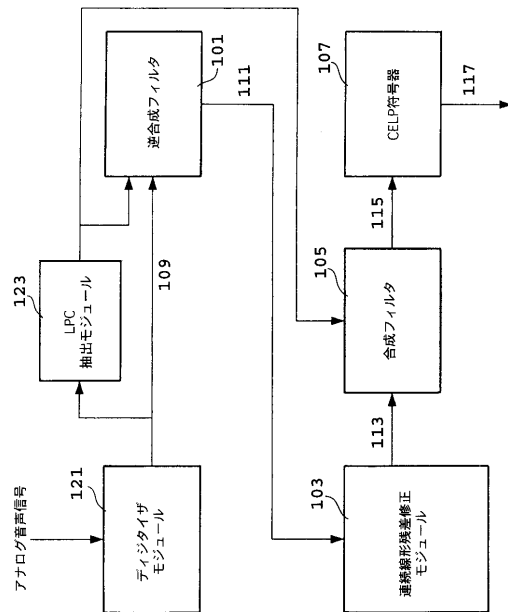
【図 8】本発明の一実施形態を実施できるコンピュータ装置の概略図である。

【符号の説明】

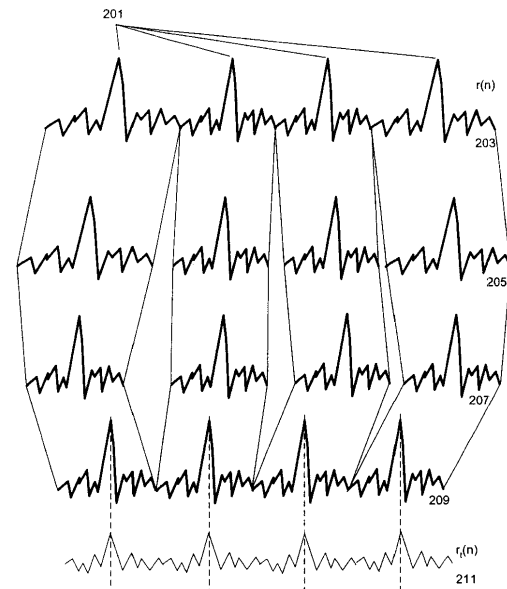
1 0 1	予測フィルタまたは逆合成フィルタ	20
1 0 3	残差修正モジュール	
1 0 5	合成フィルタ	
1 0 7	C E L P 符号器	
1 0 9	音声信号	
1 1 1	残差信号	
1 1 3	修正残差	
1 1 5	修正音声信号	
1 1 7	符号化したビット	
1 1 9	アナログ音声信号	
1 2 1	ディジタイザ	30
1 2 3	L P C 抽出装置	
2 0 1	ピーク	
2 0 3、2 0 5、2 0 7、2 0 9、2 1 1	波形	
5 0 1、5 0 3、5 0 5	モデリング関数	
6 0 1	連続輪郭	
6 0 3	不連続輪郭	
6 0 5	領域	
6 0 7	セクション	
7 0 0	コンピュータ装置	
7 0 2	処理ユニット	40
7 0 4	メモリ	
7 0 8	取外し可能記憶装置	
7 1 0	取外し不能記憶装置	
7 1 2	通信接続	
7 1 4	入力装置	
7 1 6	出力装置	



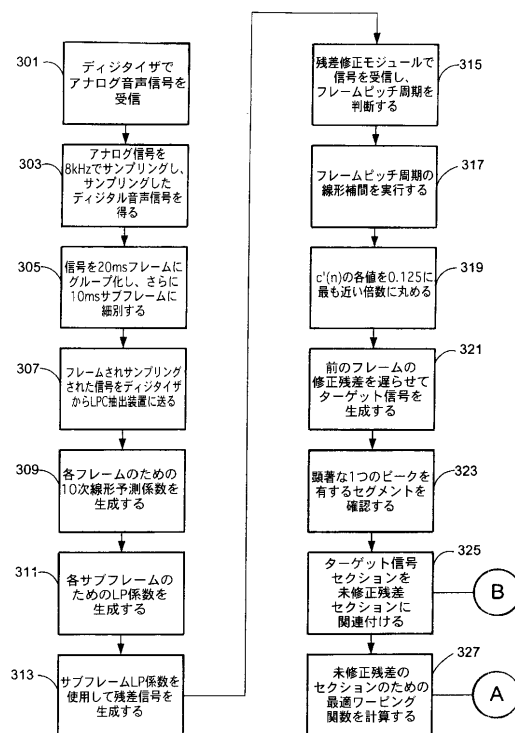
【図 1】



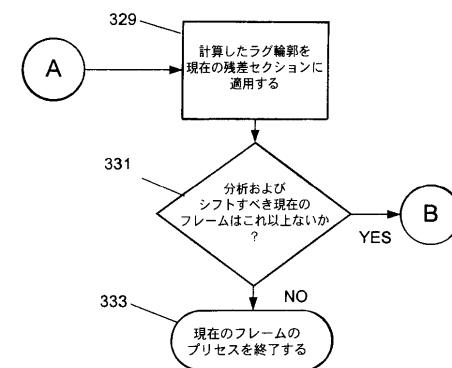
【図 2】



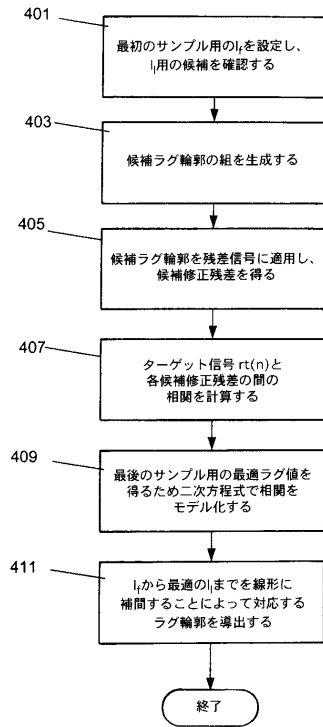
【図 3】



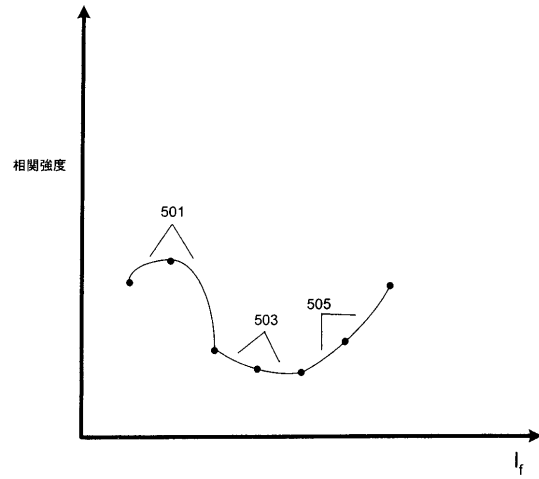
【図 4】



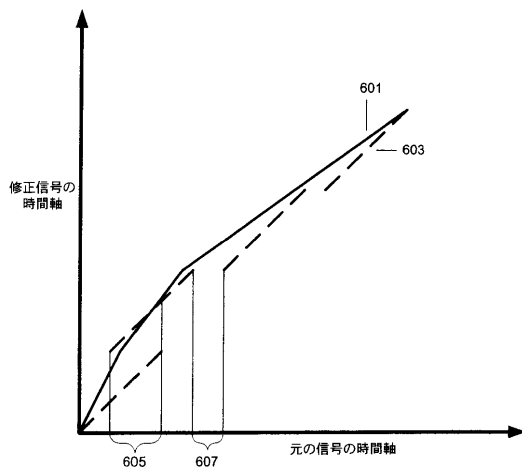
【図 5】



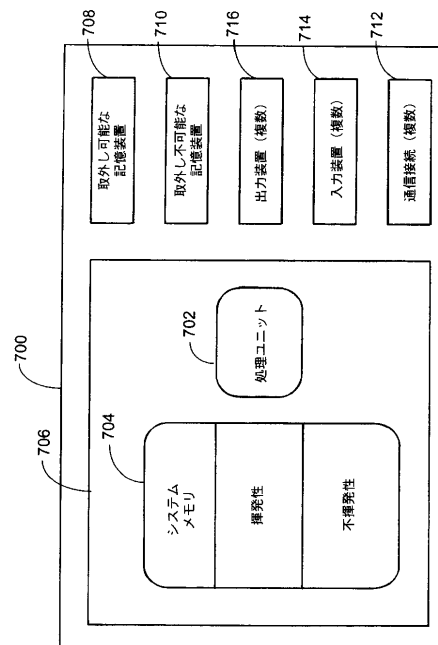
【図 6】



【図 7】



【図 8】



---

フロントページの続き

審査官 山下 剛史

(56)参考文献 特開平 1 0 - 1 4 3 1 9 9 ( J P , A )  
国際公開第 0 0 / 0 4 1 1 6 8 ( W O , A 1 )

(58)調査した分野(Int.Cl. , D B 名)  
G10L19/00-19/14  
H03M 7/30  
H04B 14/04