

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4678440号
(P4678440)

(45) 発行日 平成23年4月27日(2011.4.27)

(24) 登録日 平成23年2月10日(2011.2.10)

(51) Int. Cl. F I
G 1 O L 19/00 (2006.01) G 1 O L 19/00 3 3 O E
G 1 O L 19/14 (2006.01) G 1 O L 19/14 4 0 O E

請求項の数 1 (全 15 頁)

<p>(21) 出願番号 特願2008-526756 (P2008-526756) (86) (22) 出願日 平成19年7月23日(2007.7.23) (86) 国際出願番号 PCT/JP2007/064421 (87) 国際公開番号 W02008/013135 (87) 国際公開日 平成20年1月31日(2008.1.31) 審査請求日 平成20年11月20日(2008.11.20) (31) 優先権主張番号 特願2006-204781 (P2006-204781) (32) 優先日 平成18年7月27日(2006.7.27) (33) 優先権主張国 日本国(JP)</p>	<p>(73) 特許権者 000004237 日本電気株式会社 東京都港区芝五丁目7番1号 (74) 代理人 100102864 弁理士 工藤 実 (72) 発明者 伊藤 博紀 東京都港区芝五丁目7番1号 日本電気株式会社社内 (72) 発明者 小澤 一範 東京都港区芝五丁目7番1号 日本電気株式会社社内 審査官 菊池 智紀</p>
---	--

最終頁に続く

(54) 【発明の名称】 音声データ復号装置

(57) 【特許請求の範囲】

【請求項1】

音声データ中にロスがあるかを検出するロスディテクタと、
 前記音声データの前記ロス直前の部分を復号して第一復号音声信号を生成する音声データデコーダと、

前記第一復号音声信号から第一パラメータを抽出する音声データアナライザと、前記第一パラメータは、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号のすべて、又は、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号ゲインのすべてを含み、

乱数を加える又はゲインを小さくしていくことで前記第一パラメータを修正するパラメータ修正部と、

修正された前記第一パラメータを用いて第一合成音声信号を生成する音声合成部と、
音声信号出力部と

を具備し、

前記ロスディテクタは、前記音声信号出力部が前記ロスを補間する音声信号を出力する前に前記音声データの前記ロスの次の部分を受信したかを検出し、

前記音声データデコーダは、前記次の部分を復号して第二復号音声信号を生成し、

前記音声データアナライザは、前記第二復号音声信号の時間を反転させて第二パラメータを抽出し、

前記第二パラメータは、スペクトルパラメータ、遅延パラメータ、適応コードブックゲ

10

20

イン、正規化残差信号のすべて、又は、スペクトルパラメータ、遅延パラメータ、適応ロードブックゲイン、正規化残差信号ゲインのすべてを含み、

前記パラメータ修正部は、乱数を加える又はゲインを小さくしていくことで前記第二パラメータを修正し、

前記音声合成部は、修正された前記第二パラメータを用いて第二合成音声信号を生成し

、前記音声信号出力部は、前記第一合成音声信号と前記第二合成音声信号の時間反転信号とを含む音声信号を前記ロスを補間する音声信号として前記第一合成音声信号の強度の前記第二合成音声信号の強度に対する比率が時間とともに低下するように出力する

波形符号化方式による音声データ復号装置。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声データの復号装置、音声データの変換装置、及び誤り補償方法に関する。

【背景技術】

【0002】

回線交換網又はパケット網を使って音声データを伝送する際、音声データを符号化、復号を行うことで音声信号の授受を行っている。この音声圧縮の方式としては、例えば、ITU-T(International Telecommunication Union Telecommunication Standardization Sector)勧告G.711方式、及びCELP(Code-Excited Linear Prediction)方式が知られている。

20

【0003】

これらの圧縮方式で符号化された音声データを伝送すると、無線誤り又はネットワークの輻輳等により、音声データの一部が欠落することがある。この欠落部に対する誤り補償として、欠落部より前の音声データの部分の情報に基づいて、欠落部に対する音声信号の生成を行う。

【0004】

このような誤り補償においては、音質が劣化することがある。特開2002-268697号公報は、音質の劣化を低減する方法を開示している。この方法においては、遅れて受信したパケットに含まれる音声フレームデータを用いて、フィルタメモリ値を更新する。すなわち、ロスしたパケットを遅れて受信した場合、このパケットに含まれる音声フレームデータを用いて、ピッチフィルタ、またはスペクトル概形を表すフィルタで使用するフィルタメモリ値を更新する。

30

【0005】

また、特開2005-274917号公報は、ADPCM(Adaptive Differential Pulse Code Modulation)符号化に関連する技術を開示している。この技術は、符号化側と復号化側の予測器の状態不一致により不快な異常音を出力するという課題を解決することを可能とする。この課題は、符号化データの欠落後に正しい符号化データを受け取っても発生することがある。すなわち、パケット損失が「検出」から「非検出」へ遷移してから所定時間、検出状態制御部が過去の音声データを基に生成した補間信号の強度を徐々に減少させ、時間が経つにつれて符号化側と復号化側との予測器の状態が次第に一致して音声信号が正常になっていくので、音声信号の強度を徐々に増大させる。その結果、この技術は、符号化データの欠落状態から復旧した直後においても異常音を出力しないという効果を奏する。

40

【0006】

さらに、特開平11-305797号公報では、音声信号から線形予測計数を算出し、この線形予測計数から音声信号を生成する方法が開示されている。

【発明の開示】

50

【 0 0 0 7 】

従来の音声データに対する誤り補償方式は、過去の音声波形を繰り返す単純な方式であるため、上記のような技術が開示されているものの、音質に依然、改善の余地が残されていた。

【 0 0 0 8 】

本発明の目的は、音質の劣化を防止しなら音声データの誤りを補償することである。

【 0 0 0 9 】

波形符号化方式による音声データ復号装置は、ロスディテクタと、音声データデコーダと、音声データアナライザと、パラメータ修正部と、音声合成部を備える。ロスディテクタは、音声データ中にロスがあるかを検出する。音声データデコーダは、音声データを復号して第一復号音声信号を生成する。音声データアナライザは、第一復号音声信号から第一パラメータを抽出する。パラメータ修正部は、ロス検出の結果に基づいて第一パラメータを修正する。音声合成部は、修正された第一パラメータを用いて第一合成音声信号を生成する。

10

【 0 0 1 0 】

本発明によれば、音質の劣化を防止しながら音声データの誤りが補償される。

【 図面の簡単な説明 】

【 0 0 1 1 】

【 図 1 】 本発明の実施例 1 の音声データ復号装置の構成を示す概略図である。

【 図 2 】 本発明の実施例 1 の音声データ復号装置の動作を示す流れ図である。

20

【 図 3 】 本発明の実施例 2 の音声データ復号装置の構成を示す概略図である。

【 図 4 】 本発明の実施例 2 の音声データ復号装置の動作を示す流れ図である。

【 図 5 】 本発明の実施例 3 の音声データ復号装置の構成を示す概略図である。

【 図 6 】 本発明の実施例 3 の音声データ復号装置の動作を示す流れ図である。

【 図 7 】 本発明の実施例 4 の音声データ復号装置の構成を示す概略図である。

【 図 8 】 本発明の実施例 4 の音声データ復号装置の動作を示す流れ図である。

【 図 9 】 本発明の実施例 5 の音声データ変換装置の構成を示す概略図である。

【 図 1 0 】 本発明の実施例 5 の音声データ変換装置の動作を示す流れ図である。

【 発明を実施するための最良の形態 】

【 0 0 1 2 】

30

本発明の実施の形態について図面を参照しながら説明する。しかしながら、係る形態は本発明の技術的範囲を限定するものではない。

【 0 0 1 3 】

本発明の実施例 1 について、図 1 及び図 2 を参照しながら以下に説明する。

【 0 0 1 4 】

図 1 は、G . 7 1 1 方式に代表される波形符号化方式で符号化された音声データに対する復号装置の構成を示す。実施例 1 の音声データ復号装置は、ロスディテクタ 1 0 1、音声データデコーダ 1 0 2、音声データアナライザ 1 0 3、パラメータ修正部 1 0 4、音声合成部 1 0 5 及び音声信号出力部 1 0 6 を備える。ここで、音声データとは、ある一連の音声データを符号化したデータをいい、また、少なくとも 1 つの音声フレームを含む音声データのことをいう。

40

【 0 0 1 5 】

ロスディテクタ 1 0 1 は、受信した音声データを音声データデコーダ 1 0 2 に出力するとともに、受信した音声データがロスしたかを検出し、ロス検出結果を音声データデコーダ 1 0 2 とパラメータ修正部 1 0 4 と音声信号出力部 1 0 6 に出力する。

【 0 0 1 6 】

音声データデコーダ 1 0 2 は、ロスディテクタ 1 0 1 から入力された音声データを復号して、復号音声信号を音声データ出力部 1 0 6 と音声データアナライザ 1 0 3 に出力する。

【 0 0 1 7 】

50

音声データアナライザ103は、復号音声信号をフレーム毎に分割し、分割した信号に対して線形予測分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。各フレームの長さは、例えば、20msである。次に、音声データアナライザ103は、分割した音声信号をサブフレームに分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータとして、ピッチ周期に対応する遅延パラメータと適応コードブックゲインを抽出する。各サブフレームの長さは、例えば5msである。また、音声データアナライザ103は、適応コードブックにより該当するサブフレームの音声信号をピッチ予測する。さらに、音声データアナライザ103は、ピッチ予測して求めた残差信号を正規化して、正規化残差信号と正規化残差信号ゲインを抽出する。そして、抽出したスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲイン（これらはパラメータと呼ばれる場合がある）をパラメータ修正部104に出力する。音声データアナライザ103は、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号及び正規化残差信号ゲインのうちから2つ以上を抽出することが好ましい。

10

【0018】

パラメータ修正部104は、ロスディテクタ101から入力されたロス検出結果に基づいて、音声データアナライザ103から入力されたスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを修正しない、又は±1%の乱数を加える、或いはゲインを小さくしていくなどの修正をする。さらに、パラメータ修正部104は、修正した又は修正していない値を音声合成部105に出力する。これらの値を修正する理由は、繰り返しにより不自然な音声信号が生成されることを避けるためである。

20

【0019】

音声合成部105は、パラメータ修正部104から入力されたスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを使って合成音声信号を生成し、音声信号出力部106に出力する。

【0020】

音声信号出力部106は、ロスディテクタ101から入力されたロス検出結果に基づいて、音声データデコーダ102から入力された復号音声信号、音声合成部105から入力された合成音声信号、又は復号音声信号と合成音声信号とをある比率で混合した信号のいずれかを出力する。

30

【0021】

次に、図2を参照しながら、実施例1の音声データ復号装置の動作を説明する。

【0022】

まず、ロスディテクタ101は、受信した音声データがロスしているかを検出する（ステップS601）。ロスディテクタ101は、無線網におけるビット誤りをCRC（Cyclic Redundancy Check）符号を用いて検出した場合に音声データがロスしたとして検出する方法、又はIP（Internet Protocol）網におけるロスをRFC3550 RTP（A Transport Protocol for Real-Time Applications）ヘッダのシーケンス番号の抜けにより検出した場合に音声データがロスしたとして検出する方法を用いることができる。

40

【0023】

ロスディテクタ101が音声データのロスを検出しなかったならば、音声データアナライザ102が受信した音声データを復号し、音声信号出力部へ出力する（ステップS602）。

【0024】

ロスディテクタ101が音声データのロスを検出したならば、音声データアナライザ103が、音声データのロス直前の部分に対応する復号音声信号に基づいて、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを抽出する（ステップS603）。ここで、復号音声信号の分析は、音声デー

50

タのロス直前の部分に対応する復号音声信号に対して行なってもよいし、全ての復号音声信号に対して行ってもよい。次に、パラメータ修正部104はロス検出結果に基づいて、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを修正しない、或いは±1%の乱数を加える等して修正する(ステップS604)。音声合成部105は、これらの値を使って、合成音声信号を生成する(ステップS605)。

【0025】

そして、音声信号出力部106は、ロス検出結果に基づいて、音声データデコーダ102から入力された復号音声信号、音声合成部105から入力された合成音声信号又は復号音声信号と合成音声信号とをある比率で混合した信号のいずれかを出力する(ステップS606)。具体的には、前フレームと現フレームでロスが検出されていない場合は、音声信号出力部106は復号音声信号を出力する。ロスが検出された場合は、音声信号出力部106は合成音声信号を出力する。ロスが検出された次のフレームでは、最初は、合成音声信号の比が大きく、時間が経過するにつれて復号音声信号の比が大きくなるように音声信号を加算することにより、音声信号出力部106から出力される音声信号が不連続になることを避ける。

10

【0026】

実施例1の音声データ復号装置は、パラメータを抽出し、これらの値を、音声データのロスを補間する信号に利用することで、ロスを補間する音声の音質を向上させることができる。従来、G.711方式においてはパラメータを抽出していなかった。

20

【0027】

実施例2について、図3及び図4を参照しながら説明する。実施例2と実施例1との異なる点は、音声データのロスを検出した際、ロス部分を補間する音声信号を出力する前に、ロス後の次の音声データを受信しているかを検出する。そして、次の音声データを検出した場合、ロスした音声データに対する音声信号を生成するのに、実施例1の動作に加え、次の音声データの情報をも用いる点である。

【0028】

図3は、G.711方式に代表される波形符号化方式で符号化された音声データに対する復号装置の構成を示す。実施例2の音声データ復号装置は、ロスディテクタ201、音声データデコーダ202、音声データアナライザ203、パラメータ修正部204、音声合成部205及び音声信号出力部206を含む。ここで、音声データデコーダ202、パラメータ修正部204及び音声合成部205は、実施例1の音声データデコーダ102、パラメータ修正部104及び音声合成部105と同じ動作をする。

30

【0029】

ロスディテクタ201は、ロスディテクタ101と同じ動作を実行する。音声データのロスを検出した場合、ロスディテクタ201は、音声信号出力部206がロス部分を補間する音声信号を出力する前に、ロス後の次の音声データを受信しているかを検出する。さらに、ロスディテクタ201は、この検出結果を音声データデコーダ202と音声データアナライザ203とパラメータ修正部204と音声信号出力部206に出力する。

【0030】

音声データアナライザ203は、音声データアナライザ103と同じ動作を実行する。音声データアナライザ203は、ロスディテクタ201からの検出結果に基づいて、ロスを検出した次の音声データに対する音声信号の時間を反転させた信号を生成する。そして、この信号について実施例1と同様の手順で分析を行い、抽出したスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインをパラメータ修正部204に出力する。

40

【0031】

音声信号出力部206は、ロスディテクタ201から入力されたロス検出結果に基づいて、音声データデコーダ202から入力された復号音声信号、或いは最初はロスが検出された前の音声データのパラメータにより生成された合成音声信号の比率が高く、最後はロ

50

スが検出された次の音声データのパラメータにより生成された合成音声信号の時間を反転させた信号の比率が高くなるように加算した信号のいずれかを出力する。

【0032】

次に、図4を参照しながら、実施例2の音声データ復号装置の動作を説明する。

【0033】

まず、ロスディテクタ201は、受信した音声データがロスしているかを検出する(ステップS701)。ロスディテクタ201が音声データのロスを検出しなかったならば、ステップS602と同様の動作を行う(ステップ702)。

【0034】

ロスディテクタ201が音声データのロスを検出したならば、ロスディテクタ201が、音声信号出力部206がロス部分を補間する音声信号を出力する前にロス後の次の音声データを受信しているか、検出する(ステップS703)。次の音声データを受信していないならば、ステップS603乃至S605と同様の動作を行う(ステップS704乃至S706)。次の音声データを受信したならば、音声データデコーダ202が次の音声データを復号する(ステップS707)。この復号した次の音声データを基に、音声データアナライザ203がスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを抽出する(ステップS708)。次に、パラメータ修正部204はロス検出結果に基づいて、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを修正しない、或いは±1%の乱数を加える等して修正する(ステップS709)。音声合成部205は、これらの値を使って、合成音声信号を生成する(ステップS710)。

【0035】

そして、音声信号出力部206は、ロスディテクタ201から入力されたロス検出結果に基づいて、音声データデコーダ202から入力された復号音声信号、または最初はロスが検出された前の音声データのパラメータにより生成された合成音声信号の比率が高く、最後はロスが検出された次の音声データのパラメータにより生成された合成音声信号の時間を反転させた信号の比率が高くなるように加算した信号を出力する(ステップS711)。

【0036】

近年、急速に普及しているVoIP(Voice over IP)では、音声データの到着時間の揺らぎを吸収するために、受信した音声データのバッファリングを行っている。実施例2によれば、ロスした部分の音声信号を補間する際に、バッファに存在しているロスした次の音声データを用いることで、補間信号の音質を向上させることができる。

【0037】

実施例3について、図5及び図6を参照しながら説明する。本実施例では、CELP方式で符号化された音声データの復号に関して、音声データのロスを検出した場合に、実施例2と同様に、第一音声データデコーダ302がロス部分を補間する音声信号を出力する前にロス後の次の音声データを受信していれば、ロスした音声データに対する音声信号を生成する際に次の音声データの情報を用いる。

【0038】

図5は、CELP方式で符号化された音声データに対する復号装置の構成を示す。実施例3の音声データ復号装置は、ロスディテクタ301、第一音声データデコーダ302、パラメータ補間部304、第二音声データデコーダ303及び音声信号出力部305を備える。

【0039】

ロスディテクタ301は、受信した音声データを第一音声データデコーダ302と第二音声データデコーダ303に出力するとともに、受信した音声データがロスしているかを検出する。ロスを検出した場合に、第一音声データデコーダ302がロス部分を補間する音声信号を出力する前に次の音声データを受信しているかを検出し、検出結果を第一音声データデコーダ302と第二音声データデコーダ303に出力する。

【0040】

第一音声データデコーダ302は、ロスが検出されなかった場合、ロスディテクタ301から入力された音声データを復号して、復号音声信号を音声データ出力部に出力し、復号時のスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインをパラメータ補間部303に出力する。また、第一音声データデコーダ302は、ロスを検出し、次の音声データを受信していない場合、過去の音声データの情報を用いてロス部分を補間する音声信号を生成する。第一音声データデコーダ302は、特開2002-268697号公報に記載されている方法を用いて音声信号を生成することができる。さらに、第一音声データデコーダ302は、パラメータ補間部304から入力されたパラメータを用いてロスした音声データに対する音声信号を生成し、音声信号出力部305に出力する。

10

【0041】

第二音声データデコーダ303は、ロスを検出し、第一音声データデコーダ302がロス部分を補間する音声信号を出力する前に次の音声データを受信している場合、ロスした音声データに対する音声信号を過去の音声データの情報を用いて生成する。そして、第二音声データデコーダ303は、生成した音声データを使って次の音声データを復号し、復号に用いるスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを抽出し、パラメータ補間部304に出力する。

【0042】

パラメータ補間部304は、第一音声データデコーダ302から入力されたパラメータと第二音声データデコーダ303から入力されたパラメータを用いて、ロスした音声データに対するパラメータを生成し、第一音声データデコーダ302に出力する。

20

【0043】

音声信号出力部305は、音声データデコーダ302から入力された復号音声信号を出力する。

【0044】

次に、図6を参照しながら、実施例3の音声データ復号装置の動作を説明する。

【0045】

まず、ロスディテクタ301が受信した音声データがロスしているかを検出する(ステップS801)。ロスしていないならば、第一音声データデコーダ302が、ロスディテクタ301から入力された音声データを復号し、復号時のスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインをパラメータ補間部304に出力する(ステップS802及びS803)。

30

【0046】

ロスしているならば、ロスディテクタ301が第一音声データデコーダ302がロス部分を補間する音声信号を出力する前にロス後の次の音声データを受信しているか、検出する(ステップS804)。次の音声データを受信していないならば、第一音声データデコーダ302が、過去の音声データの情報を用いてロス部分を補間する音声信号を生成する(ステップS805)。

【0047】

次の音声データを受信しているならば、第二音声データデコーダ303が、ロスした音声データに対する音声信号を過去の音声データの情報を用いて生成する(ステップS806)。第二音声データデコーダ303は、生成した音声信号を使って次の音声データを復号し、復号時のスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号または正規化残差信号ゲインを生成し、パラメータ補間部303に出力する(ステップS807)。次に、パラメータ補間部304が、第一音声データデコーダ302から入力されたパラメータと第二音声データデコーダ303から入力されたパラメータを用いて、ロスした音声データに対するパラメータを生成する(ステップS808)。そして、第一音声データデコーダ302は、パラメータ補間部304が生成したパラメータを用いて、ロスした音声データに対する音声信号を生成し、音声信号出力部305に出力す

40

50

る(ステップS809)。

【0048】

第一音声データデコーダ302はそれぞれの場合で生成した音声信号を音声信号出力部305へ出力し、音声信号出力部305が復号音声信号を出力する(ステップS810)。

【0049】

近年、急速に普及しているVoIPでは、音声データの到着時間の揺らぎを吸収するために、受信した音声データのバッファリングを行っている。実施例3によれば、CELP方式においてロスした部分の音声信号を補間する際に、バッファに存在しているロスした次の音声データを用いることで、補間信号の音質を向上させることができる。

10

【0050】

実施例4について、図7及び図8を参照しながら説明する。CELP方式において、音声データのロスが生じたときに補間信号を用いると、ロスした部分は補うことができるものの、補間信号は正しい音声データから生成したわけではないので、その後受信した音声データの音質を低下させてしまう。そこで、実施例4は、実施例3に加えて、音声データのロスの部分に対する補間音声信号を出力した後に、ロスした部分の音声データが遅れて届いた場合、この音声データを用いることにより、ロスした次の音声データの音声信号の品質を向上させる。

【0051】

図7は、CELP方式で符号化された音声データに対する復号装置の構成を示す。実施例4の音声データ復号装置は、ロスディテクタ401、第一音声データデコーダ402、第二音声データデコーダ403、メモリ蓄積部404及び音声信号出力部405を備える。

20

【0052】

ロスディテクタ401は、受信した音声データを第一音声データデコーダ402と第二音声データデコーダ403に出力する。また、ロスディテクタ401は、受信した音声データがロスしたかを検出する。ロスを検出した場合には、次の音声データを受信しているかを検出し、検出結果を第一音声データデコーダ402、第二音声データデコーダ403及び音声信号出力部405に出力する。さらに、ロスディテクタ401は、ロスした音声データが遅れて受信したかどうかを検出する。

30

【0053】

第一音声データデコーダ402は、ロスが検出されなかった場合、ロスディテクタ401から入力された音声データを復号する。また、第一音声データデコーダ402は、ロスが検出された場合、過去の音声データの情報を用いて音声信号を生成して、音声データ出力部405に出力する。第一音声データデコーダ402は、特開2002-268697号後方に記載されている方法を用いて音声信号を生成することができる。さらに、第一音声データデコーダ402は、合成フィルタ等のメモリをメモリ蓄積部404に出力する。

【0054】

第二音声データデコーダ403は、ロス部分の音声データが遅れて到着した場合、遅れて到着した音声データを、メモリ蓄積部404に蓄積されているロス検出直前パケットの合成フィルタ等のメモリを使って復号し、復号信号を音声信号出力部405に出力する。

40

【0055】

音声信号出力部405は、ロスディテクタ401から入力されたロス検出結果に基づいて、第一音声データデコーダ402から入力された復号音声信号、第二音声データデコーダ403から入力された復号音声信号または前記二つの信号をある比率で加算した音声信号を出力する。

【0056】

次に、図8を参照しながら、実施例4の音声データ復号装置の動作を説明する。

【0057】

まず、音声データ復号装置は、ステップS801乃至S810の動作を行い、ロスした

50

音声データを補間する音声信号を出力する。ここで、ステップS805及びS806のときに、過去の音声データより音声信号を生成したときに、合成フィルタ等のメモリをメモリ蓄積部404に出力する(ステップS903及びS904)。そして、ロスディテクタ401が、ロスしていた音声データを遅れて受信したのかを検出する(ステップS905)。ロスディテクタ401が検出していないならば、実施例3で生成した音声信号を出力する。ロスディテクタ401が検出したならば、第二音声データデコーダ403が、遅れて到着した音声データを、メモリ蓄積部404に蓄積されているロス検出直前パケットの合成フィルタ等のメモリを使って復号する(ステップS906)。

【0058】

そして、音声出力部405が、ロスディテクタ401から入力されたロス検出結果に基づいて、第一音声データデコーダ402から入力された復号音声信号、第二音声データデコーダ403から入力された復号音声信号または前記二つの信号をある比率で加算した音声信号を出力する(ステップS907)。具体的には、ロスを検出し、音声データが遅れて到着した場合、音声信号出力部405は、ロスした音声データの次の音声データに対する音声信号として、最初は、第一音声データデコーダ402から入力された復号音声信号の比を大きくする。そして、時間が経過するにつれて、音声信号出力部405は、第二音声データデコーダ403から入力された復号音声信号の比を大きくするように加算した音声信号を出力する。

10

【0059】

実施例4によれば、遅れて届いたロス部分の音声データを用いて合成フィルタ等のメモリを書き換えることで、正しい復号音声信号を生成することができる。また、この正しい復号音声信号を、あえてすぐに出しせず、ある比率で加算した音声信号を出力することで、音声の不連続になることを防止することができる。さらに、ロスした部分に補間信号を用いたとしても、遅れて届いたロス部分の音声データで合成フィルタ等のメモリを書きかえて復号音声信号を生成することで、補間信号後の音質を向上させることができる。

20

【0060】

ここで、実施例4は、実施例3の変形例として説明したが、他の実施例の変形例であってもよい。

【0061】

実施例5の音声データ変換装置について、図9及び図10を参照しながら説明する。

30

【0062】

図9は、ある音声符号化方式で符号化された音声信号を、別の音声符号化方式に変換する音声データ変換装置の構成を示している。音声データ変換装置は、例えば、G.711で代表される波形符号化方式で符号化された音声データを、CELP方式で符号化された音声データに変換する。実施例5の音声データ変換装置は、ロスディテクタ501、音声データデコーダ502、音声データエンコーダ503、パラメータ修正部504及び音声データ出力部505を備える。

【0063】

ロスディテクタ501は、受信した音声データを音声データデコーダ502に出力する。また、ロスディテクタ501は、受信した音声データがロスしているかを検出し、検出結果を音声データデコーダ502と音声データエンコーダ503とパラメータ修正部504と音声データ出力部505に出力する。

40

【0064】

音声データデコーダ502は、ロスが検出されなかった場合、ロスディテクタ501から入力された音声データを復号し、復号音声信号を音声データエンコーダ503に出力する。

【0065】

音声データエンコーダ503は、ロスが検出されなかった場合、音声データデコーダ502から入力された復号音声信号を符号化し、符号化した音声データを音声データ出力部505に出力する。また、音声データエンコーダ503は、符号化時のパラメータである

50

スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、残差信号または残差信号ゲインをパラメータ修正部504に出力する。さらに、音声データエンコーダ503は、ロスが検出された場合、パラメータ修正部504から入力されパラメータを受け取る。そして、音声データエンコーダ503は、パラメータ抽出に用いるフィルタ(図示せず)を保持しており、パラメータ修正部504から受け取ったパラメータを符号化して、音声データを生成する。その際に、音声データエンコーダ503はフィルタ等のメモリを更新する。ここで、音声データエンコーダ503は、符号化時に生じる量子化誤差により、符号化後のパラメータ値がパラメータ修正部504から入力された値と同じ値にならない場合、符号化後のパラメータ値がパラメータ修正部504から入力された値に最も近い値となるように選択する。また、通信相手の無線通信装置が保持するフィルタのメモリとの齟齬が生じることを避けるために、音声データエンコーダ503は、音声データを生成する際に、パラメータ抽出などに用いるフィルタが持つメモリ(図示せず)を更新する。さらに、音声データエンコーダ503は、生成した音声データを音声データ出力部505に出力する。

10

【0066】

パラメータ修正部504は、音声データエンコーダ503から符号化時のパラメータであるスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、残差信号または残差信号ゲインを受け取り、保存する。また、パラメータ修正部504は、保持していたロス検出前のパラメータを修正しないで、又は所定の修正をし、ロスディテクタ501から入力されるロス検出結果に基づいて、音声データエンコーダ503へ出力する。

20

【0067】

音声データ出力部505は、ロスディテクタ501から受け取ったロス検出結果に基づいて、音声データエンコーダ503から受け取った音声信号を出力する。

【0068】

次に、図10を参照しながら、実施例5の音声データ変換装置を説明する。

【0069】

まず、ロスディテクタ501が、受信した音声データがロスしているかを検出する(ステップS1001)。ロスディテクタ501がロスを検出しなかったなら、音声データエンコーダ502が受信した音声データを基に復号音声信号を生成する(ステップS1002)。そして、音声データエンコーダ503が、復号音声信号を符号化し、符号化時のパラメータであるスペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、残差信号または残差信号ゲインを出力する(ステップS1003)。

30

【0070】

ロスディテクタ501がロスを検出したなら、パラメータ修正部504が、保持しているロス前のパラメータを修正しないで、または所定の修正をして、音声データエンコーダ503へ出力する。このパラメータを受信した音声データエンコーダ503は、パラメータを抽出するためのフィルタが持つメモリを更新する(ステップS1004)。さらに、音声データエンコーダ503が、ロスする直前のパラメータを基に音声信号を生成する(ステップS1005)。

【0071】

そして、音声データ出力部505が、ロス検出結果に基づいて、音声データエンコーダ503から受け取った音声信号を出力する(ステップS1006)。

40

【0072】

実施例5により、例えばゲートウェイなどのようなデータを変換する装置において、音声データのロスに対する補間信号を波形符号化方式で生成せず、パラメータなどを用いてロス部分を補間することで、補間信号の音質を向上させることができる。また、音声データのロスに対する補間信号を波形符号化方式で生成せず、パラメータなどを用いてロス部分を補間することで、演算量を少なくすることができる。

【0073】

ここで、実施例5ではG.711で代表される波形符号化方式で符号化された音声デー

50

タをC E L P方式で符号化された音声データに変換する形態を示したが、C E L P方式で符号化された音声データを別のC E L P方式で符号化された音声データに変換する形態でもよい。

【0074】

上記実施例に係る装置のうちのあるものは、例えば、以下のようにまとめることが可能である。

【0075】

波形符号化方式による音声データ復号装置は、ロスディテクタと、音声データデコーダと、音声データアナライザと、パラメータ修正部と、音声合成部と、音声信号出力部を備える。ロスディテクタは、音声データ中にロスを検出し、音声信号出力部がロスを補間する音声信号を出力する前にロス後の音声フレームを受信したかを検出する。音声データデコーダは、音声フレームを復号して復号音声信号を生成する。音声データアナライザは、復号音声信号の時間を反転させてパラメータを抽出する。パラメータ修正部は、パラメータに所定の修正を行う。音声合成部は、修正されたパラメータを用いて合成音声信号を生成する。

【0076】

C E L P方式(C o d e - E x c i t e d L i n e a r P r e d i c t i o n)による音声データ復号装置は、ロスディテクタと、第一音声データデコーダと、第二音声データデコーダと、パラメータ補間部と、音声信号出力部を備える。ロスディテクタは、音声データ中にロスがあるかを検出し、第一音声データデコーダが第一音声信号を出力する前にロス後の音声フレームを受信したかを検出する。第一音声データデコーダは、ロス検出の結果に基づいて、音声データを復号して音声信号を生成する。第二音声データデコーダは、ロス検出の結果に基づいて、音声フレームに対応する音声信号を生成する。パラメータ補間部は、第一及び第二パラメータを用いてロスに対応する第三パラメータを生成して第一音声データデコーダに出力する。音声信号出力部は、第一音声データデコーダから入力された音声信号を出力する。第一音声データデコーダは、ロスが検出されなかった場合、音声データを復号して音声信号を生成し、この復号時に抽出した第一パラメータをパラメータ補間部に出力する。第一音声データデコーダは、ロスが検出された場合、音声データのロスの前の部分を用いてロスに対応する第一音声信号を生成する。第二音声データデコーダは、ロスが検出され、かつ第一音声データデコーダが第一音声信号を出力する前に音声フレームが検出された場合、音声データのロスの前の部分を用いてロスに対応する第二音声信号を生成し、第二音声信号を用いて音声フレームを復号し、この復号時に抽出した第二パラメータをパラメータ補間部に出力する。第一音声データデコーダは、パラメータ補間部から入力された第三パラメータを用いてロスに対応する第三音声信号を生成する。

【0077】

C E L P方式により、音声データ中のロスを補間する補間信号を出力する音声データ復号装置は、ロスディテクタと、音声データデコーダと、音声信号出力部を備える。ロスディテクタは、ロスを検出し、音声データのロス部分を遅れて受信したことを検出する。ロス部分はロスに対応する。音声データデコーダは、メモリ蓄積部に蓄積されている音声データのロスの前の部分を使ってロス部分を復号して復号音声信号を生成する。音声信号出力部は、復号音声信号を含む音声信号を復号音声信号の強度の音声信号の強度に対する比率が変化するように出力する。

【0078】

第一音声符号化方式の第一音声データを第二音声符号化方式の第二音声データに変換する音声データ変換装置は、ロスディテクタと、音声データデコーダと、音声データエンコーダと、パラメータ修正部を備える。ロスディテクタは、第一音声データ中のロスを検出する。音声データデコーダは、第一音声データを復号して復号音声信号を生成する。音声データエンコーダは、パラメータを抽出するフィルタを備え、復号音声信号を第二音声符号化方式で符号化する。パラメータ修正部は、音声データエンコーダからパラメータを受

10

20

30

40

50

け取って保持する。パラメータ修正部は、パラメータに所定の修正を行い、又は行わずに、ロス検出の結果に基づいて、音声データエンコーダに出力する。音声データエンコーダは、ロスが検出されなかった場合、復号音声信号を第二音声符号化方式で符号化し、この符号化時に抽出したパラメータをパラメータ修正部に出力する。音声データエンコーダは、ロスが検出された場合、パラメータ修正部から入力されるパラメータに基づいて音声信号を生成し、フィルタが持つメモリを更新する。

【0079】

第一音声符号化方式が波形符号化方式であり、第二音声符号化方式がC E L P方式であることが好ましい。

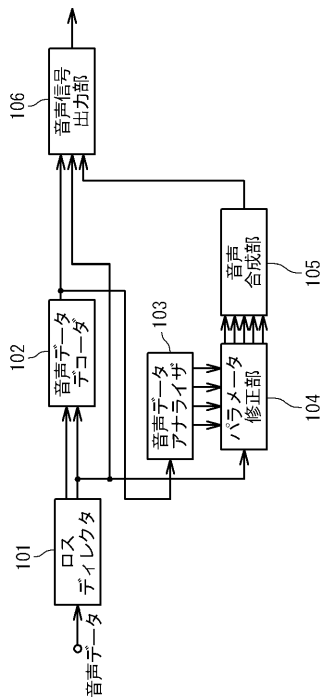
【0080】

パラメータが、スペクトルパラメータ、遅延パラメータ、適応コードブックゲイン、正規化残差信号、または正規化残差信号ゲインであることが好ましい。

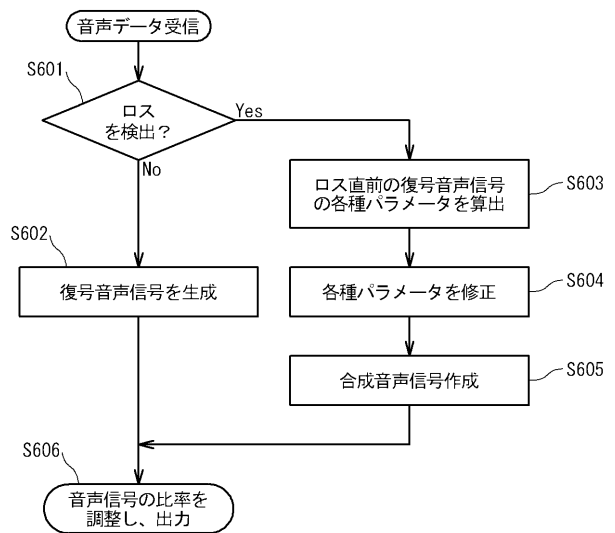
【0081】

当業者は上記実施例の様々な変形を容易に実施することができる。したがって、本発明は上記実施例に限定されることはなく、請求項やその均等物によって参酌される最も広い範囲で解釈される。

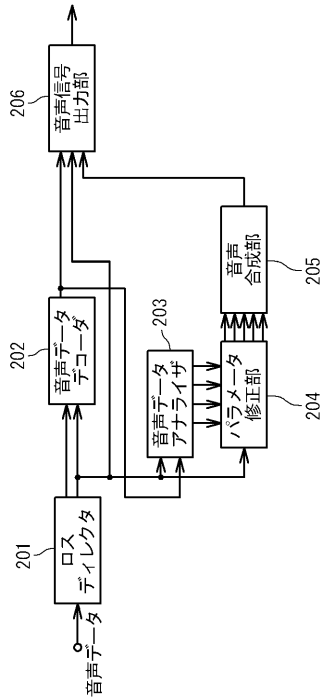
【図1】



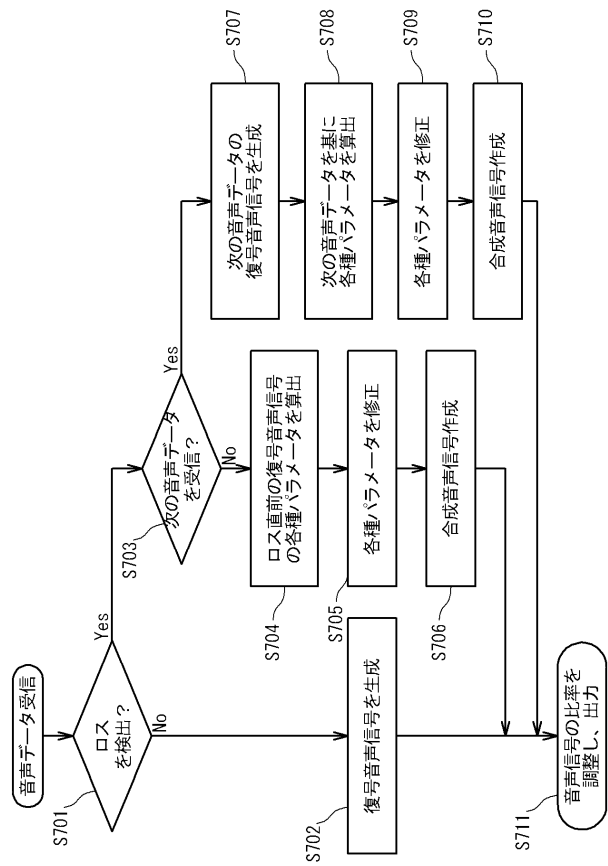
【図2】



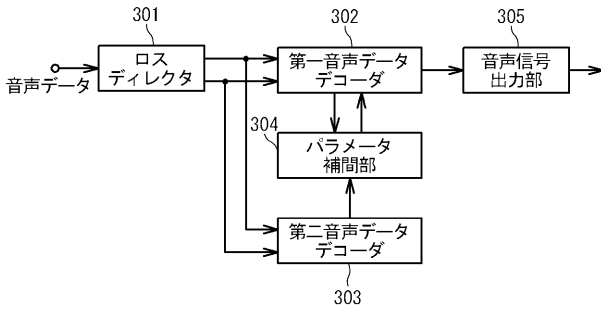
【図3】



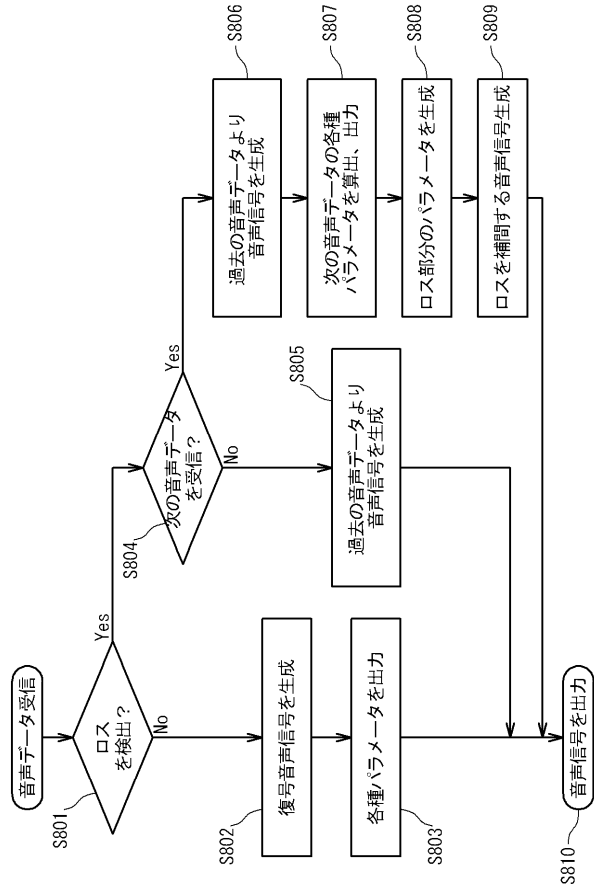
【図4】



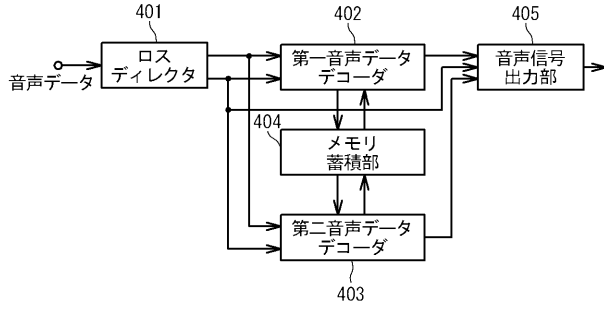
【図5】



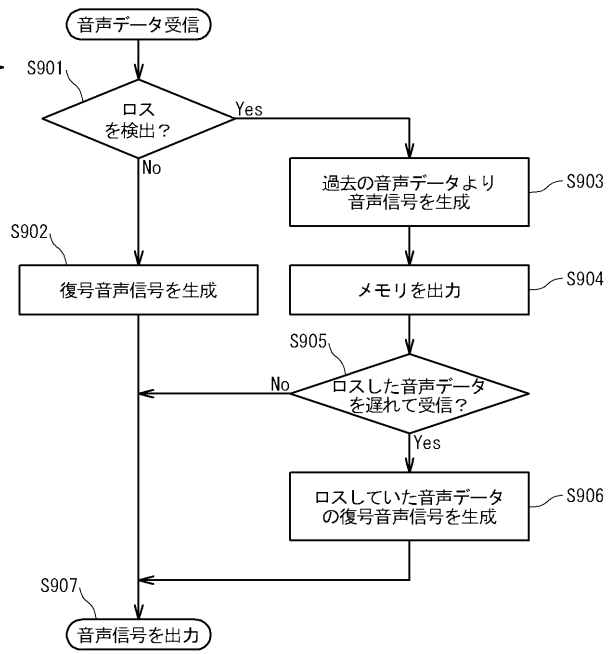
【図6】



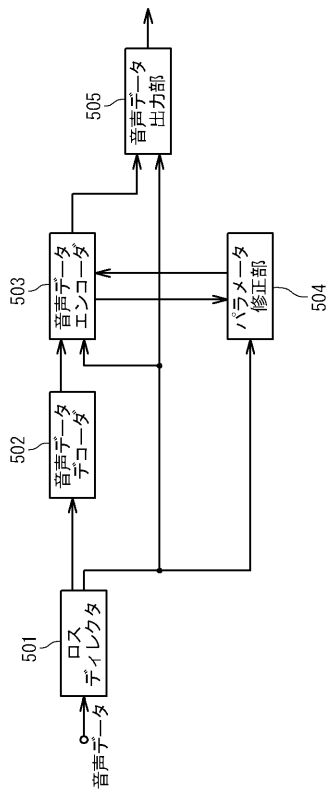
【図7】



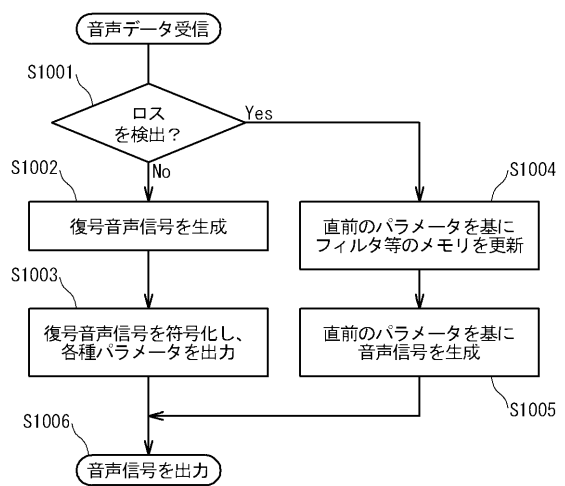
【図8】



【図9】



【図10】



フロントページの続き

- (56)参考文献 特開平09-321783(JP,A)
特開平08-110798(JP,A)
特開2005-077889(JP,A)
舟木慶一 他, "G.711パケットロス隠蔽方式の改良に関する検討", 日本音響学会2004年秋季研究発表会講演論文集 - I -, 2004年 9月21日, p.383-384
森永徹 他, "広帯域IP網におけるパケット消失に耐性のある音声符号化", 電子情報通信学会2001年総合大会講演論文集 通信2, 2001年 3月 7日, p.377
芹沢昌宏 他, "遅延パケットを用いたフィルタメモリ修復によるCELP復号音質改善方法", 電子情報通信学会2001年情報・システムソサイエティ大会講演論文集, 2001年 8月29日, p.234
近藤和弘 他, "線形予測を用いた音声パケット欠落補間方式の一検討", 日本音響学会2003年秋季研究発表会講演論文集 - I -, 2003年 9月17日, p.187-188

- (58)調査した分野(Int.Cl., DB名)
G10L 19/00-19/14
JSTPlus(JDreamII)