

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-70866

(P2019-70866A)

(43) 公開日 令和1年5月9日(2019.5.9)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 19/005 (2013.01)	G 1 0 L 19/005	
G 1 0 L 19/00 (2013.01)	G 1 0 L 19/00	3 3 0 B
G 1 0 L 19/09 (2013.01)	G 1 0 L 19/09	
	G 1 0 L 19/00	3 3 0 C

審査請求 有 請求項の数 4 O L (全 54 頁)

(21) 出願番号	特願2019-27042 (P2019-27042)	(71) 出願人	392026693
(22) 出願日	平成31年2月19日 (2019.2.19)		株式会社 N T T ドコモ
(62) 分割の表示	特願2017-41456 (P2017-41456) の分割	(74) 代理人	100088155
原出願日	平成25年11月12日 (2013.11.12)		弁理士 長谷川 芳樹
(31) 優先権主張番号	特願2012-251646 (P2012-251646)	(74) 代理人	100113435
(32) 優先日	平成24年11月15日 (2012.11.15)		弁理士 黒木 義樹
(33) 優先権主張国	日本国 (JP)	(74) 代理人	100121980
			弁理士 沖山 隆
		(74) 代理人	100128107
			弁理士 深石 賢治
		(72) 発明者	堤 公孝
			東京都千代田区永田町二丁目11番1号
			株式会社 N T T ドコモ内

最終頁に続く

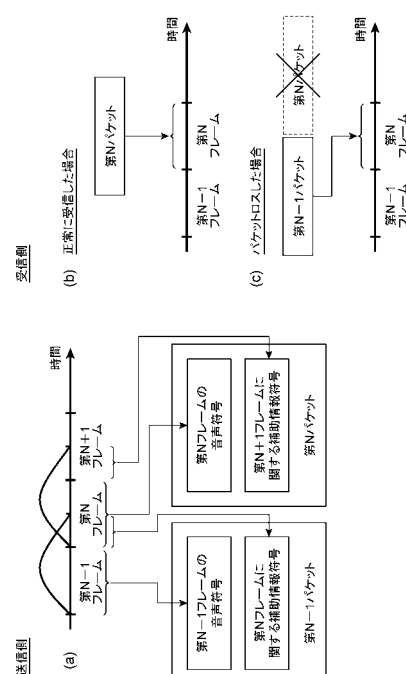
(54) 【発明の名称】 音声符号化装置および音声符号化方法

(57) 【要約】

【課題】 音声符号化におけるパケットロスにおいて、アルゴリズム遅延を増加させずに音声品質を回復することを目的とする。

【解決手段】 本発明は、音声信号を符号化する音声符号化装置による音声符号化方法であって、音声信号を符号化する音声符号化ステップと、先読み信号から、復号音声合成するための音声パラメータの予測値を算出するための補助情報を算出して符号化する補助情報符号化ステップと、を含む。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

音声信号を符号化する音声符号化装置による音声符号化方法であって、
音声信号を符号化する音声符号化ステップと、
先読み信号から、復号音声を作成するための音声パラメータの予測値を算出するための
補助情報を算出して符号化する補助情報符号化ステップと、
を含む音声符号化方法。

【請求項 2】

前記補助情報は、先読み信号におけるピッチラグに関する、請求項 1 に記載の音声符号
化方法。

10

【請求項 3】

前記補助情報は、補助情報の利用可否に関する情報を含む、請求項 1 又は 2 に記載の音
声符号化方法。

【請求項 4】

音声信号を符号化する音声符号化装置であって、
音声信号を符号化する音声符号化部と、
先読み信号から、復号音声を作成するための音声パラメータの予測値を算出するための
補助情報を算出して符号化する補助情報符号化部と、
を備える音声符号化装置。

20

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、音声パケットを、IP 網や移動体通信網経由で伝送する際のエラー隠蔽に関
するものであり、さらに詳しくは、エラー隠蔽を実現する、高精度パケットロス隠蔽信号
生成のための音声符号化装置および音声符号化方法に関する。

【背景技術】**【0002】**

音声・音響信号（以下「音声信号」と総称する）を IP 網や移動体通信において伝送す
る際には、ある一定時間長ごとに音声信号を符号化して音声パケットを生成し、通信網経
由で伝送する。受信側では、通信網を通じて音声パケットを受け取り、受信側のサーバ、
MCU（Multipoint Control Unit）、端末等において復号し、復号音声信号とする。

30

【0003】

音声信号は、通常、デジタル形式で収音される。すなわち、1 秒間あたりサンプリン
グ周波数と同数の数列として計測・蓄積する。この数列の個々の要素をサンプルと呼ぶ。
音声符号化においては、決まったサンプル数の音声信号が、内蔵したバッファに所定量蓄
積するたびにバッファ内の音声信号を符号化する。上記所定のサンプル数をフレーム長と
いい、フレーム長と同数のサンプルの集合をフレームと呼ぶ。例えば、32 kHz のサン
プリング周波数において、フレーム長を 20 ms とした場合、フレーム長は 640 サンプ
ルである。なお、バッファの長さは 1 フレームより長くてもよい。

40

【0004】

通信網を通じて音声パケットを伝送する際には、通信網の輻輳状態等に起因して、一部
の音声パケットが失われる、あるいは音声パケットに書き込まれた情報の一部に誤りが生
じるといった現象（いわゆるパケットロス）が起こりうる。そのような場合には、受信側
において音声パケットを正しく復号することができないため、所望の復号音声信号を得る
ことが出来ない。また、パケットロスが生じた音声パケットに対応する復号音声信号は雑
音として知覚されるため、受聴する人間に対して与える主観品質を著しく損なう。

【0005】

50

上記のような不都合を解消するため、パケットロスにより失われた部分の音声信号を補間する技術として、パケットロス隠蔽技術が用いられる。パケットロス隠蔽技術には、受信側のみでパケットロス隠蔽を行う「補助情報を用いないパケットロス隠蔽技術」と、送信側でパケットロス隠蔽に役立つパラメータを求めた上で、受信側に伝送し、受信側では受け取ったパラメータを利用してパケットロス隠蔽を行う「補助情報を用いるパケットロス隠蔽技術」がある。

【 0 0 0 6 】

このうち「補助情報を用いないパケットロス隠蔽技術」では、例えば、非特許文献 1 の技術のように過去に正常に受信したパケットに含まれていた復号音声信号をピッチ単位でコピーした上で、予め決めた減衰係数を乗算することにより、パケットロスした部分に対応する音声信号を生成する。「補助情報を用いないパケットロス隠蔽技術」は、パケットロスした部分の音声の性質がパケットロスする直前の音声と似ていることを前提としているため、パケットロスした部分がロスする直前の音声と異なる性質を持つ場合や、パワーが急激に変化する場合に、十分な隠蔽効果を発揮することができない。

【 0 0 0 7 】

一方、「補助情報を用いるパケットロス隠蔽技術」には、特許文献 1 のように、送信側でパケットロス隠蔽に必要なパラメータを符号化して伝送し、受信側でのパケットロス隠蔽に利用する技術がある。特許文献 1 では、主要符号化・冗長符号化の 2 つの符号化方式により音声を符号化する。冗長符号化は、主要符号化よりも低いビットレートで、主要符号化が符号化するフレームより 1 つ前のフレームを符号化する（図 1（a）参照）。例えば、第 N 番目のパケットには、第 N フレームを主要符号化により符号化して得た音声符号と、第 N - 1 フレームを冗長符号化により符号化して得た補助情報符号を含めて伝送する。

【 0 0 0 8 】

受信側では、時間的に連続する 2 つ以上のパケットの到着を待ってから、時間的に早い方のパケットを復号して、復号音声信号を得る。例えば、第 N フレームに対応する信号を得る場合、第 N + 1 パケットが到着するのを待ってから復号を行う。第 N パケット、第 N + 1 パケットを正常に受信した場合、第 N パケットに含まれる音声符号を復号することにより第 N フレームの音声信号を得る（図 1（b）参照）。一方、パケットロスした場合（第 N パケットがロスした状況で、第 N + 1 パケットを得た場合）、第 N + 1 パケットに含まれる補助情報符号を復号することにより、第 N フレームの音声信号を得ることができる（図 1（c）参照）。

【 0 0 0 9 】

特許文献 1 の方法では、復号対象となるパケットが到着しても、さらに 1 パケット以上到着するまで、復号を待たなければならず、アルゴリズム遅延が 1 パケット分以上増加する。従って、特許文献 1 の方法では、パケットロス隠蔽による音質向上は期待できるものの、アルゴリズム遅延が増加し、音声通話品質が低下する。

【 0 0 1 0 】

さらに、上記のようなパケットロス隠蔽技術を CELP（Code Excited Linear Prediction、符号励振線形予測）符号化に対して適用する際には、CELP の動作の特徴に起因する別の課題も生じる。CELP は、線形予測に基づく音声モデルであり、音声信号を高精度にかつ高い圧縮率で符号化できるため、多くの国際標準で用いられている。

【 0 0 1 1 】

CELP では、全極型合成フィルタによって励振信号 $e(n)$ をフィルタリングすることにより合成される。すなわち、次式に従い音声信号 $s(n)$ を合成する。

【 0 0 1 2 】

【 数 1 】

$$s(n) = e(n) - \sum_{i=1}^P a(i) \cdot s(n-i)$$

10

20

30

40

50

a (i) は線形予測係数 (L P 係数) であり、次数として例えば $P = 16$ などの値を用いる。

【 0 0 1 3 】

励振信号は、適応符号帳と呼ばれるバッファに蓄積される。あらたなフレームの音声を合成するに当たっては、ピッチラグと呼ばれる位置情報を元に、適応符号帳から読み出された適応符号帳ベクトルと、励振信号の時間変化を表す固定符号帳ベクトルを加算することにより、新たに励振信号を生成する。新たに生成された励振信号は、適応符号帳に蓄積されると同時に、全極型合成フィルタによりフィルタリングされ、復号信号が合成される。

10

【 0 0 1 4 】

C E L P では、全てのフレームについて L P 係数が算出される。L P 係数を算出するに当たっては、10ms 程度の先読み信号が必要である。すなわち、符号化対象フレームに加えて、先読み信号もバッファに蓄積した上で、L P 係数算出およびその後の処理を実施する (図 2 参照)。各フレームは 4 つ程度のサブフレームに分割され、サブフレーム単位で、上記ピッチラグ算出、適応符号帳ベクトル算出、固定符号帳ベクトル算出、適応符号帳更新といった処理を行う。上記サブフレーム単位の処理にあたっては、L P 係数も補間処理を行うことにより、サブフレーム毎に異なる係数に変化させる。また、L P 係数は量子化・補間処理の都合上、L P 係数の等価表現である I S P (Immittance Spectral Pair) パラメータ、I S F (Immittance Spectral Frequency) パラメータに変換した上で符号化される。L P 係数と I S P パラメータ・I S F パラメータの相互変換の処理手順については、非特許文献 2 に記載されている。

20

【 0 0 1 5 】

C E L P 符号化では、符号化側と復号側がそれぞれ適応符号帳をもち、それら適応符号帳が常に同期することを前提として符号化・復号を行う。パケットを正常に受信し、復号が正常に行われる状況では、符号化側の適応符号帳と復号側の適応符号帳は同期しているが、一度パケットロスが起これば、適応符号帳の同期を取ることができなくなる。

【 0 0 1 6 】

例えば、ピッチラグとして用いる値が、符号化側と復号側で異なれば、適応符号帳ベクトルは時間的にずれたものとなる。この適応符号帳ベクトルで適応符号帳の更新を行うため、次のフレームを正常に受信したとしても、符号化側で求まる適応符号帳ベクトルと、復号側で求まる適応符号帳ベクトルは一致せず、適応符号帳の同期が回復することはない。このような適応符号帳の不一致によって、パケットロスしたフレーム以降、数フレームに渡って音質劣化が生じる。

30

【 0 0 1 7 】

C E L P 符号化におけるパケットロス隠蔽において、より高度な技術として特許文献 2 の手法がある。特許文献 2 では、パケットロスによる影響が大きい特定のフレームにおいて、ピッチラグや適応符号帳ゲインの代わりに、遷移モードコードブックのインデックスを送信する。特許文献 2 では、パケットロスによる影響が大きいフレームとして、遷移フレーム (無音の音声セグメントから有音の音声セグメントへの遷移、あるいは 2 つの母音間の遷移) に着目している。この遷移フレームにおいて、遷移モードコードブックを用いた励振信号生成を行うことにより、過去の適応符号帳に非依存の励振信号を生成し、過去のパケットロスによる適応符号帳不一致から回復することができる。

40

【 0 0 1 8 】

特許文献 2 の方法は、例えば長めの母音が継続するフレームなどでは、遷移フレームコードブックを利用しないため、従来と同様、適応符号帳の不一致から回復することができない。さらに、遷移モードコードブックを含むパケットが失われた場合には、これまでと同様に、パケットロス以降のフレームにもロスの影響が残ってしまう。これは、遷移モードコードブックを含むパケットの次のパケットが失われた場合も同様である。

【 0 0 1 9 】

50

遷移フレームコードブックのような過去のフレームに依存しないコードブックを、全てのフレームに適応することもできるが、符号化の効率が著しく落ちるため、低ビットレート・高音質を達成できない。

【先行技術文献】

【特許文献】

【0020】

【特許文献1】特表2003-533916号公報

【特許文献2】特表2010-507818号公報

【非特許文献】

【0021】

【非特許文献1】ITU-T G.711 Appendix I

【非特許文献2】3GPP TS26-191

【非特許文献3】3GPP TS26-190

【非特許文献4】ITU-T G.718

【発明の概要】

【発明が解決しようとする課題】

【0022】

特許文献1の方法を用いると、復号対象のパケットが到着した後、後続のパケット到着を待ってから復号を開始する。そのため、パケットロス隠蔽による音質向上はあるものの、アルゴリズム遅延が増加し、通話品質が低下する。

【0023】

C E L P符号化におけるパケットロス時に、符号化部・復号部間における適応符号帳の不一致が原因で、音声品質の劣化が生じる。特許文献2のような方法により、適応符号帳の不一致から回復することもできるが、遷移フレームの直前以外のフレームがロスした場合に十分な効果が得られない。

【0024】

本発明は、上記課題の解決のためになされたものであり、音声符号化におけるパケットロスにおいて、アルゴリズム遅延を増加させずに音声品質を回復する音声符号化装置および音声符号化方法を提供することを目的とする。

【課題を解決するための手段】

【0025】

上述の課題を解決するために、本発明の音声符号化方法は、本発明は、音声信号を符号化する音声符号化装置による音声符号化方法であって、音声信号を符号化する音声符号化ステップと、先読み信号から、復号音声を作成するための音声パラメータの予測値を算出するための補助情報を算出して符号化する補助情報符号化ステップと、を含む。

【発明の効果】

【0026】

音声符号化におけるパケットロスにおいて、アルゴリズム遅延を増加させずに音声品質を回復することができる。特に、C E L P符号化において、パケットロス時に起こる適応符号帳の劣化を低減し、パケットロス時の音声品質を改善することができる。

【図面の簡単な説明】

【0027】

【図1】特許文献1に記載の先行発明におけるパケットと復号信号の時間的関係を示す図である。

【図2】C E L P符号化におけるL P分析対象信号と先読み信号の時間的関係を示す図である。

【図3】本発明の実施形態におけるパケットと復号信号の時間的関係を示す図である。

【図4】本発明の実施例1における音声信号送信装置の機能構成例を示す図である。

【図5】本発明の実施例1における音声信号受信装置の機能構成例を示す図である。

【図6】本発明の実施例1における音声信号送信装置の処理手順を示す図である。

10

20

30

40

50

- 【図 7】本発明の実施例 1 における音声信号受信装置の処理手順を示す図である。
- 【図 8】本発明の実施例 1 における補助情報符号化部の機能構成例を示す図である。
- 【図 9】本発明の実施例 1 における補助情報符号化部の処理手順を示す図である。
- 【図 10】本発明の実施例 1 における LP 係数算出部の処理手順を示す図である。
- 【図 11】本発明の実施例 1 におけるターゲット信号算出部の処理手順を示す図である。
- 【図 12】本発明の実施例 1 における音声パラメータ紛失処理部の機能構成例を示す図である。
- 【図 13】本発明の実施例 1 における音声パラメータ予測の処理手順を示す図である。
- 【図 14】本発明の実施例 1 の変形例 1 - 1 における励振ベクトル合成部の処理手順を示す図である。
- 【図 15】本発明の実施例 1 における音声合成部の機能構成図を示す図である。
- 【図 16】本発明の実施例 1 における音声合成部の処理手順を示す図である。
- 【図 17】本発明の実施例 1 の変形例 1 - 2 における補助情報符号化部（補助情報出力判定部を設けた場合）の機能構成例を示す図である。
- 【図 18】本発明の実施例 1 の変形例 1 - 2 における補助情報符号化部（補助情報出力判定部を設けた場合）の処理手順を示す図である。
- 【図 19】本発明の実施例 1 の変形例 1 - 2 における音声パラメータ予測の処理手順を示す図である。
- 【図 20】本発明の実施例 2 における音声信号送信装置の機能構成例を示す図である。
- 【図 21】本発明の実施例 2 における主要符号化部の機能構成例を示す図である。
- 【図 22】本発明の実施例 2 における音声信号送信装置の処理手順を示す図である。
- 【図 23】本発明の実施例 2 における音声信号受信装置の機能構成例を示す図である。
- 【図 24】本発明の実施例 2 における音声信号受信装置の処理手順を示す図である。
- 【図 25】本発明の実施例 2 における音声合成部の機能構成図を示す図である。
- 【図 26】本発明の実施例 2 における音声パラメータ復号部の機能構成例を示す図である。
- 【図 27】本発明の実施例 3 における補助情報符号化部の機能構成例を示す図である。
- 【図 28】本発明の実施例 3 における補助情報符号化部の処理手順を示す図である。
- 【図 29】本発明の実施例 3 におけるピッチラグ選定部の処理手順を示す図である。
- 【図 30】本発明の実施例 3 における補助情報復号部の処理手順を示す図である。
- 【図 31】本発明の実施形態に係る音声符号化プログラムの構成を記憶媒体と共に示す図である。
- 【図 32】本発明の実施形態に係る音声復号プログラムの構成を記憶媒体と共に示す図である。
- 【図 33】本発明の実施例 4 における補助情報符号化部の機能構成例を示す図である。
- 【図 34】本発明の実施例 4 における補助情報符号化部の処理手順を示す図である。
- 【図 35】本発明の実施例 4 におけるピッチラグ予測部の処理手順を示す図（その 1）である。
- 【図 36】本発明の実施例 4 におけるピッチラグ予測部の処理手順を示す図（その 2）である。
- 【図 37】本発明の実施例 4 におけるピッチラグ予測部の処理手順を示す図（その 3）である。
- 【図 38】本発明の実施例 4 における適応符号帳算出部の処理手順を示す図である。
- 【図 39】本発明の実施例 5 における補助情報符号化部の機能構成例を示す図である。
- 【図 40】本発明の実施例 5 におけるピッチラグ符号化部の処理手順を示す図である。
- 【図 41】本発明の実施例 5 における補助情報復号部の処理手順を示す図である。
- 【図 42】本発明の実施例 5 におけるピッチラグ予測部の処理手順を示す図である。
- 【図 43】本発明の実施例 5 における適応符号帳算出部の処理手順を示す図である。
- 【発明を実施するための形態】
- 【0028】

10

20

30

40

50

添付図面を参照しながら本発明の実施形態を説明する。可能な場合には、同一の部分には同一の符号を付して、重複する説明を省略する。

【0029】

本発明の実施形態は、符号化器側で算出した補助情報を符号化して伝送し、復号側でのパケットロス隠蔽に利用する「補助情報を用いるパケットロス隠蔽技術」を実現する符号化器、および復号器である。

【0030】

本発明の実施形態では、パケットロス隠蔽に用いる補助情報は1つ前のパケットに含める。パケットに含まれる音声符号と補助情報符号の時間的関係を図3に示す。図3からも明らかなように、本発明の実施形態における補助情報は、CELP符号化における先読み信号について求めたパラメータである（ピッチラグ、適応符号帳ゲイン等）。

10

【0031】

補助情報符号を1つ前のパケットに含めたことにより、復号対象のパケットよりも後続のパケットを待つことなく、復号を行うことができる。パケットロスを検出した際も、隠蔽対象となるフレームに関する補助情報が、直前のパケットにおいて得られているので、後続のパケットを待つことなく、高精度なパケットロス隠蔽を実現することができる。

【0032】

また、補助情報として先読み信号におけるCELP符号化のパラメータを伝送することにより、パケットがロスしても適応符号帳の不一致を軽減することができる。

20

【0033】

本発明の実施形態は、音声信号送信装置（音声符号化装置）、音声信号受信装置（音声復号装置）からなる。音声信号送信装置の機能構成例を図4に示し、処理手順を図6に示す。また、音声信号受信装置の機能構成例を図5に示し、処理手順を図7に示す。

【0034】

音声信号送信装置は、図4に示す通り、音声符号化部111、補助情報符号化部112からなる。音声信号受信装置は、図5に示す通り、音声符号バッファ121、音声パラメータ復号部122、音声パラメータ紛失処理部123、音声合成部124、補助情報復号部125、補助情報蓄積部126からなる。

【0035】

音声信号送信装置は、図6に示す処理手順により、音声信号をフレーム毎に符号化して伝送する。

30

【0036】

音声符号化部111は、符号化対象フレームに対して音声パラメータを算出して、音声符号を出力する（図6ステップS131）。

【0037】

補助情報符号化部112は、先読み信号に対して音声パラメータを算出して、補助情報符号を出力する（図6ステップS132）。

【0038】

音声信号が終了したか否かを判定し、音声信号が終了するまで上記を繰り返す（図6ステップS133）。

40

【0039】

音声信号受信装置は、図7に示す処理手順により、到着する音声パケットを復号して音声信号を出力する。

【0040】

音声符号バッファ121は、音声パケットの到着を待って、音声符号を蓄積する。音声パケットが正常に到着した場合は、処理を音声パラメータ復号部122に切替える。一方、音声パケットが正常に到着しない場合は、処理を音声パラメータ紛失処理部123に切り替える（図7ステップS141）。

【0041】

< 音声パケットを正常に受信した場合 >

50

音声パラメータ復号部 1 2 2 は、音声符号を復号して音声パラメータを出力する（図 7 ステップ S 1 4 2）。

【 0 0 4 2 】

補助情報復号部 1 2 5 は、補助情報符号を復号して、補助情報を出力する。出力した補助情報は、補助情報蓄積部 1 2 6 に送られる（図 7 ステップ S 1 4 3）。

【 0 0 4 3 】

音声合成部 1 2 4 は、音声パラメータ復号部 1 2 2 が出力した音声パラメータから音声信号を合成して出力する（図 7 ステップ S 1 4 4）。

【 0 0 4 4 】

音声パラメータ紛失処理部 1 2 3 は、音声パラメータ復号部 1 2 2 が出力した音声パラメータを、パケットロス時に備えて蓄積する（図 7 ステップ S 1 4 5）。 10

【 0 0 4 5 】

音声符号バッファ 1 2 1 は、音声パケットの送信が終了したかどうかを判断し、音声パケットの送信が終了した場合は処理を停止する。音声パケットの送信が継続する間は、上記ステップ S 1 4 1 - S 1 4 6 を繰り返す（図 7 ステップ S 1 4 7）。

【 0 0 4 6 】

< 音声パケットがロスした場合 >

音声パラメータ紛失処理部 1 2 3 は、補助情報蓄積部 1 2 6 から補助情報を読み出し、補助情報に含まれないパラメータについては予測を行うことで、音声パラメータを出力する（図 7 ステップ S 1 4 6）。 20

【 0 0 4 7 】

音声合成部 1 2 4 は、音声パラメータ紛失処理部 1 2 3 が出力した音声パラメータから音声信号を合成して出力する（図 7 ステップ S 1 4 4）。

【 0 0 4 8 】

音声パラメータ紛失処理部 1 2 3 は、音声パラメータ紛失処理部 1 2 3 が出力した音声パラメータを、パケットロス時に備えて蓄積する（図 7 ステップ S 1 4 5）。

【 0 0 4 9 】

音声符号バッファ 1 2 1 は、音声パケットの送信が終了したかどうかを判断し、音声パケットの送信が終了した場合は処理を停止する。音声パケットの送信が継続する間は、上記ステップ S 1 4 1 - S 1 4 6 を繰り返す（図 7 ステップ S 1 4 7）。 30

【 0 0 5 0 】

[実施例 1]

本実施例では、補助情報としてピッチラグを伝送し、復号側ではパケットロス隠蔽信号の生成に用いる例について記載する。

【 0 0 5 1 】

音声信号送信装置の機能構成例を図 4 に、音声信号受信装置の機能構成例を図 5 にそれぞれ示し、音声信号送信装置の処理手順を図 6 に、音声信号受信装置の処理手順を図 7 にそれぞれ示す。

【 0 0 5 2 】

< 送信側 >

音声信号送信装置において、入力音声信号は、音声符号化部 1 1 1 に送られる。 40

【 0 0 5 3 】

音声符号化部 1 1 1 は、符号化対象フレームを、C E L P 符号化により符号化する（図 6 ステップ 1 3 1）。C E L P 符号化の詳細については、例えば非特許文献 3 に記載の方法を用いる。C E L P 符号化の処理手順の詳細については省略する。なお、C E L P 符号化では符号化側でローカルデコードが行われる。ローカルデコードとは、符号化側でも音声符号を復号して、音声合成に必要なパラメータ（I S P パラメータおよび対応する I S F パラメータ、ピッチラグ、長期予測パラメータ、適応符号帳、適応符号帳ゲイン、固定符号帳ゲイン、固定符号帳ベクトル等）を得ることである。ローカルデコードにより得られたパラメータのうち、少なくとも I S P パラメータおよび I S F パラメータのいずれか 50

又は両方、ピッチラグ、適応符号帳が、補助情報符号化部 112 に送られる。音声符号化部 111 として、非特許文献 4 のような音声符号化を用いる場合には、さらに符号化対象フレームの性質を表すインデックスを送ってもよい。また、音声符号化部 111 として CELP 符号化以外の符号化を用いることもできる。この場合、入力信号あるいはローカルデコードにより得られた復号信号から、少なくとも ISF パラメータおよび ISF パラメータのいずれか又は両方、ピッチラグ、適応符号帳を別途算出し、補助情報符号化部 112 に伝送する。

【0054】

補助情報符号化部 112 は、音声符号化部 111 で算出したパラメータと先読み信号を用いて補助情報符号を算出する（図 6 ステップ 132）。補助情報符号化部 112 は、図 8 に示す通り、LP 係数算出部 151、ターゲット信号算出部 152、ピッチラグ算出部 153、適応符号帳算出部 154、励振ベクトル合成部 155、適応符号帳バッファ 156、合成フィルタ 157、ピッチラグ符号化部 158 からなる。補助情報符号化部の処理手順を図 9 に示す。

10

【0055】

LP 係数算出部 151 は、音声符号化部 111 で算出した ISF パラメータと、過去数フレームにおいて算出した ISF パラメータを用いて LP 係数を算出する（図 9 ステップ 161）。LP 係数算出部 151 の処理手順を図 10 に示す。

【0056】

最初に、音声符号化部 111 から得た ISF パラメータを用いてバッファを更新する（図 10 ステップ 171）。次に、先読み信号における ISF パラメータ

20

【数 2】

$$\dot{\omega}_i$$

を算出する。ISF パラメータ

【数 3】

$$\dot{\omega}_i$$

30

は次式により算出する（図 10 ステップ 172）。

【0057】

【数 4】

$$\dot{\omega}_i = \alpha \omega_i^{(-1)} + (1 - \alpha) \bar{\omega}_i$$

【数 5】

$$\bar{\omega}_i = \beta \omega_i^C + (1 - \beta) \frac{\omega_i^{(-3)} + \omega_i^{(-2)} + \omega_i^{(-1)}}{3}$$

40

ここで、

【数 6】

$$\omega_i^{(-j)}$$

はバッファに格納された j フレーム前の ISF パラメータである。

【0058】

また、

【数 7】

$$\omega_i^C$$

は事前に学習等で求めた発話区間での I S F パラメータである。β は定数であり、例えば 0.75 のような値とすることができるが、これには限られない。また、α も定数であり、0.9 のような値とすることができるが、これに限定されない。

【0059】

【数 8】

$$\omega_i^C, \alpha, \beta$$

10

は、例えば非特許文献 4 記載の I S F コンシールメントのように、符号化対象フレームの性質を表すインデクスにより変化させてもよい。

【0060】

次に、

【数 9】

$$\dot{\omega}_i$$

20

が

【数 10】

$$0 < \dot{\omega}_0 < \dot{\omega}_1 < \dots < \dot{\omega}_{14}$$

を満たすよう i の値を並び替え、隣り合う

【数 11】

$$\dot{\omega}_i$$

30

同士が接近しすぎないように調整する。

【0061】

【数 12】

$$\dot{\omega}_i$$

の値を調整する手順には、例えば非特許文献 4 (式 151) を用いることができる (図 10 ステップ 173)。

【0062】

次に、I S F パラメータ

40

【数 13】

$$\dot{\omega}_i$$

を I S P パラメータに変換した上で、サブフレーム毎に補間する。I S F パラメータから I S P パラメータを算出する方法として非特許文献 4 の 6.4.4 節記載の方法を、補間の方法としては、非特許文献 4 の 6.8.3 節に記載の処理手順を用いることができる (図 10 ステップ 174)。

【0063】

次にサブフレーム毎の I S P パラメータを L P 係数

50

【数 1 4】

$$\dot{a}_i^j (0 < i \leq P, 0 \leq j < M_{la})$$

に変換する。ここで、先読み信号に含まれるサブフレームの数を M_{la} とした。ISP パラメータから LP 係数への変換には、非特許文献 4 の 6 . 4 . 5 節に記載の処理手順を用いることができる（図 10 ステップ 175）。

【0064】

ターゲット信号算出部 152 は、LP 係数

【数 1 5】

$$\dot{a}_i^j$$

10

を用いて、ターゲット信号 $x(n)$ およびインパルス応答 $h(n)$ を算出する（図 9 ステップ 162）。非特許文献 4 の 6 . 8 . 4 . 1 . 3 節に記載の通り、ターゲット信号は線形予測残差信号を聴覚重み付けフィルタに通すことにより得られる（図 11）。

【0065】

まず、LP 係数を用いて先読み信号

【数 1 6】

$$s_{pre}^l(n) (0 \leq n < L')$$

20

の残差信号 $r(n)$ を次式に従い算出する（図 11 ステップ 181）。

【0066】

【数 1 7】

$$r(n) = s_{pre}^l(n) + \sum_{i=1}^P \dot{a}_i^j \cdot s_{pre}^l(n-i)$$

ただし、 L' はサブフレームのサンプル数を表し、 L は符号化対象フレーム $s_{pre}(n)$ ($0 \leq n < L$) のサンプル数を表す。このとき、

30

【数 1 8】

$$s_{pre}^l(n-p) = s_{pre}(n+L-p)$$

である。

【0067】

次に、ターゲット信号 $x(n)$ ($0 \leq n < L'$) を次式に従い算出する（図 11 ステップ 182）。

【0068】

40

【数 1 9】

$$e(n) = r(n) - \sum_{i=1}^P \dot{a}_i^j \cdot e(n-i) (0 \leq n < L')$$

【数 2 0】

$$e(n) = s(n+L-1) - \hat{s}(n+L-1) (-P \leq n < 0)$$

【数 2 1】

$$\dot{e}(n) = r(n) + \sum_{i=1}^P \dot{a}_i^j \cdot \dot{e}(n-i)$$

【数 2 2】

$$x(n) = e(n) + \gamma \cdot e(n-1)$$

ここで、聴覚重み付けフィルタ $\gamma = 0.68$ である。聴覚重み付けフィルタの値は、音声符号化の設計方針により別の値としてもよい。 10

【0 0 6 9】

次に、次式に従いインパルス応答 $h(n)$ ($0 \leq n < L'$) を算出する (図 11 ステップ 183)。

【0 0 7 0】

【数 2 3】

$$\dot{h}(n) = \dot{a}_i^j - \sum_{i=1}^P \dot{a}_i^j \cdot \dot{h}(n-i)$$

20

【数 2 4】

$$h(n) = \dot{h}(n) + \gamma \cdot \dot{h}(n-1)$$

ピッチラグ算出部 153 は、次式を最大化する k を求めることにより、サブフレーム毎にピッチラグを算出する (図 9 ステップ 163)。ここで、演算量削減のために、上記ターゲット信号算出 (図 11 ステップ 182) およびインパルス応答算出 (図 11 ステップ 183) を省略して、残差信号そのものをターゲット信号として用いてもよい。

【0 0 7 1】

$$T_p = \arg \max_k T_k$$

30

【数 2 5】

$$T_k = \frac{\sum_{n=0}^{L'-1} x(n)y_k(n)}{\sqrt{\sum_{n=0}^{L'-1} y_k(n)y_k(n)}},$$

【数 2 6】

$$y_k(n) = \sum_{i=0}^n v'(i) \cdot h(n-i)$$

40

【数 2 7】

$$v'(n) = \sum_{i=-l}^l \text{Int}(i) \cdot u(n + N_{\text{adapt}} - T_p + i)$$

なお、 $y_k(n)$ は線形予測残差にインパルス応答を畳み込むことにより得られる。ここで、 $\text{Int}(i)$ は補間フィルタを表す。補間フィルタの詳細は非特許文献 4 の 6.8.4.1.4.1 節に記載の通りである。当然、補間にフィルタを用いず、 $v'(n) = u(n + N_{\text{adapt}} - T_p + i)$ としてもよい。 50

【 0 0 7 2 】

上記、算出方法により、ピッチラグは整数値で求まるが、上記 T_k を補間することにより、ピッチラグの精度を小数点以下まで上げた場合についても求めることができる。補間により小数点以下のピッチラグを求める処理手順の詳細については、非特許文献 4 の 6 . 8 . 4 . 1 . 4 . 1 節に記載の処理方法を利用できる。

【 0 0 7 3 】

適応符号帳算出部 1 5 4 はピッチラグ T_p 、適応符号帳バッファ 1 5 6 に格納された適応符号帳 $u(n)$ から、次式に従い適応符号帳ベクトル $v'(n)$ および長期予測パラメータを算出する (図 9 ステップ 1 6 4)。

【 0 0 7 4 】

【 数 2 8 】

$$v'(n) = \sum_{i=-l}^l \text{Int}(i) \cdot u(n + N_{\text{adapt}} - T_p + i)$$

10

長期パラメータ算出の詳細な処理手順として非特許文献 3 の 5 . 7 節に記載の方法を用いることができる。

【 0 0 7 5 】

励振ベクトル合成部 1 5 5 は、適応符号帳ベクトル $v'(n)$ に事前に定めた適応符号帳ゲイン

20

【 数 2 9 】

$$g_p^C$$

を乗算して、次式に従い励振信号ベクトルを出力する (図 9 ステップ 1 6 5)。

【 0 0 7 6 】

【 数 3 0 】

$$e(n) = g_p^C \cdot v'(n)$$

30

適応符号帳ゲイン

【 数 3 1 】

$$g_p^C$$

の値には、例えば 1 . 0 などを用いるが、事前の学習により求めた値を用いてもよいし、符号化対象フレームの性質を表すインデクスにより変化させてもよい。

【 0 0 7 7 】

次に、以下の式に従い、励振信号ベクトルにより、適応符号帳バッファ 1 5 6 に格納された適応符号帳 $u(n)$ の状態を更新する (図 9 ステップ 1 6 6)。

40

【 0 0 7 8 】

$$u(n) = u(n + L) \quad (0 \leq n < N - L)$$

$$u(n + N - L) = e(n) \quad (0 \leq n < L)$$

合成フィルタ 1 5 7 は、励振信号ベクトルを励振源とする、線形予測逆フィルタリングにより、次式に従い復号信号を合成する (図 9 ステップ 1 6 7)。

【 0 0 7 9 】

【 数 3 2 】

$$\hat{s}(n) = e(n) - \sum_{i=1}^P \hat{a}_i \cdot \hat{s}(n-i)$$

50

上記図 9 ステップ 1 6 2 ~ ステップ 1 6 7 を先読み信号終了まで、サブフレーム毎に繰り返す (図 9 ステップ 1 6 8) 。

【 0 0 8 0 】

ピッチラグ符号化部 1 5 8 は、先読み信号において算出したピッチラグ
【 数 3 3 】

$$T_p^{(j)} (0 \leq j < M_{la})$$

を符号化する (図 9 のステップ 1 6 9) 。ここで、先読み信号に含まれるサブフレームの
数を M_{la} とした。

【 0 0 8 1 】

符号化の方法としては、例えば以下が考えられるが、符号化の方法には何を用いても良い。

【 0 0 8 2 】

1 . ピッチラグ

【 数 3 4 】

$$T_p^{(j)} (0 \leq j < M_{la})$$

の一部あるいは全部をバイナリ符号化あるいはスカラ量子化あるいはベクトル量子化あるいは算術符号化して伝送する方法。

【 0 0 8 3 】

2 . 直前サブフレームのピッチラグとの差分

【 数 3 5 】

$$T_p^{(j)} - T_p^{(j-1)} (0 \leq j < M_{la})$$

の一部あるいは全部をバイナリ符号化あるいはスカラ量子化あるいはベクトル量子化あるいは算術符号化して伝送する方法。ただし、

【 数 3 6 】

$$T_p^{(-1)}$$

は符号化対象フレームにおける最後のサブフレームのピッチラグである。

【 0 0 8 4 】

3 . ピッチラグ

【 数 3 7 】

$$T_p^{(j)} (0 \leq j < M_{la})$$

の一部あるいは全部と、符号化対象フレームで算出したピッチラグの一部または全部をまとめてベクトル量子化あるいは算術符号化して伝送する方法。

【 0 0 8 5 】

4 . ピッチラグ

【 数 3 8 】

$$T_p^{(j)} (0 \leq j < M_{la})$$

の一部あるいは全部を手がかりに、事前に定めた補間手法のうちの 1 つを選択し、その補

10

20

30

40

50

間手法のインデクスを伝送する方法。この時、過去に音声の合成に用いた複数のサブフレームのピッチラグを、併せて補間手法の選択に用いてもよい。

【 0 0 8 6 】

スカラ量子化およびベクトル量子化には、経験的に定めたコードブックや、学習により事前に算出したコードブックを用いることができる。また、上記ピッチラグにオフセットの値を加算した上で符号化する方法も当然、本発明の実施形態の思想に含まれる。

【 0 0 8 7 】

< 復号側 >

図 5 に示す通り、音声信号受信装置は、音声符号バッファ 1 2 1、音声パラメータ復号部 1 2 2、音声パラメータ紛失処理部 1 2 3、音声合成部 1 2 4、補助情報復号部 1 2 5、補助情報蓄積部 1 2 6 からなる。音声信号受信装置の処理手順は図 7 に示す通りである。

10

【 0 0 8 8 】

音声符号バッファ 1 2 1 は、パケットを正常に受信したか否かを判断し、パケットを正常に受信したと判断した場合には、音声パラメータ復号部 1 2 2 および補助情報復号部 1 2 5 に処理を切替え、パケットを正常に受信できなかったと判断した場合には、音声パラメータ紛失処理部 1 2 3 に処理を切り替える（図 7 ステップ 1 4 1）。

【 0 0 8 9 】

< 正常にパケットを受信した場合 >

音声パラメータ復号部 1 2 2 は、受信した音声符号を復号し、符号化対象フレームに関する音声を合成するのに必要な音声パラメータ（ISP パラメータおよび対応する ISF パラメータ、ピッチラグ、長期予測パラメータ、適応符号帳、適応符号帳ゲイン、固定符号帳ゲイン、固定符号帳ベクトル等）を算出する（図 7 ステップ 1 4 2）。

20

【 0 0 9 0 】

補助情報復号部 1 2 5 は、補助情報符号を復号してピッチラグ

【 数 3 9 】

$$\hat{T}_p^{(j)} (0 \leq j < M_{la})$$

を算出し、補助情報蓄積部 1 2 6 に格納する。補助情報復号部 1 2 5 では、符号化側で用いた符号化方法に対応する復号方法を用いて、補助情報符号を復号する（図 7 ステップ 1 4 3）。

30

【 0 0 9 1 】

音声合成部 1 2 4 は、音声パラメータ復号部 1 2 2 が出力したパラメータから符号化対象フレームに対応する音声信号を合成する（図 7 ステップ 1 4 4）。音声合成部 1 2 4 の機能構成例を図 1 5 に示し、処理手順を図 1 6 に示す。なお、信号の流れを示すために音声パラメータ紛失処理部 1 2 3 を記載しているが、音声パラメータ紛失処理部 1 2 3 は音声合成部 1 2 4 の機能構成には含まれない。

【 0 0 9 2 】

LP 係数算出部 1 1 2 1 は、ISF パラメータを ISP パラメータに変換した上で、補間処理を実施して、サブフレーム毎の ISP 係数を得る。次に ISP 係数を線形予測係数（LP 係数）に変換し、サブフレーム毎の LP 係数を得る（図 1 6 ステップ 1 1 3 0 1）。ISP 係数の補間処理、および ISP - LP 係数については、例えば非特許文献 4 の 6 . 4 . 5 節に記載の方法を用いることができる。これらパラメータ変換処理の手順は本発明の実施形態の本質ではないため、詳細については省略する。

40

【 0 0 9 3 】

適応符号帳算出部 1 1 2 3 は、ピッチラグおよび長期予測パラメータと、適応符号帳 1 1 2 2 を用いて適応符号帳ベクトルを算出する（図 1 6 ステップ 1 1 3 0 2）。ピッチラグ

【数 4 0】

$$\hat{T}_p^{(j)}$$

、適応符号帳 $u(n)$ から、次式に従い、適応符号帳ベクトル $v'(n)$ を算出する。

【0094】

【数 4 1】

$$v'(n) = \sum_{i=-l}^l Int(i) \cdot u(n + N_{adapt} - \hat{T}_p^{(j)} + i) \quad (0 \leq n < L')$$

10

適応符号帳ベクトルは、適応符号帳 $u(n)$ を F I R フィルタ $Int(i)$ により補間することで算出する。ここで、適応符号帳の長さを N_{adapt} とした。補間に用いるフィルタ $Int(i)$ は、(数 2 7) の補間フィルタと同じである。事前に定めた長さ $2l + 1$ の F I R フィルタである。 L' はサブフレームのサンプル数である。エンコーダ側と同様、補間にフィルタを用いなくてもよい。

【0095】

適応符号帳算出部 1123 は、長期予測パラメータの値に応じて、上記、適応符号帳ベクトルに対してフィルタリングを行う (図 16 ステップ 11303)。長期予測パラメータがフィルタリングを指示する値をとる場合は、以下の式により適応符号帳ベクトルにフ

20

【0096】

$$v'(n) = 0.18 v'(n-1) + 0.64 v'(n) + 0.18 v'(n+1)$$

一方、長期予測パラメータがフィルタリングを指示しない値をとる場合は、フィルタリングを行わず、 $v(n) = v'(n)$ とする。

【0097】

励振ベクトル合成部 1124 は、適応符号帳ベクトルに適応符号帳ゲイン g_p を乗算する (図 16 ステップ 11304)。さらに、励振ベクトル合成部 1124 は、固定符号帳ベクトル $c(n)$ に固定符号帳ゲイン g_c を乗算する (図 16 ステップ 11305)。さらに、励振ベクトル合成部 1124 は、適応符号帳ベクトルと固定符号帳ベクトルを加算

30

【0098】

$$e(n) = g_p \cdot v'(n) + g_c \cdot c(n)$$

ポストフィルタ 1125 は、励振信号ベクトルに対して、例えば、ピッチ強調、ノイズ強調、低域強調といった後処理を加える。ピッチ強調、ノイズ強調、低域強調と言った技術の詳細は、非特許文献 3 の 6.1 節に記載の通りである。ポストフィルタにおける処理は、本発明の実施形態の本質との関係が薄いので、詳細を省略する (図 16 ステップ 11307)。

【0099】

適応符号帳 1122 は、以下の式に従い、励振信号ベクトルにより状態を更新する (図 16 ステップ 11308)。

40

【0100】

$$u(n) = u(n+L) \quad (0 \leq n < N-L)$$

$$u(n+N-L) = e(n) \quad (0 \leq n < L)$$

合成フィルタ 1126 は、励振信号ベクトルを励振源とする、線形予測逆フィルタリングにより、次式に従い、復号信号を合成する (図 16 ステップ 11309)。

【0101】

【数 4 2】

$$\hat{s}(n) = e(n) - \sum_{i=1}^P \hat{a}(i) \cdot \hat{s}(n-i)$$

聴覚重み逆フィルタ 1 1 2 7 は、復号信号に対して、次式に従い聴覚重み逆フィルタを適用する（図 1 6 ステップ 1 1 3 1 0）。

【0 1 0 2】

【数 4 3】

$$\hat{s}(n) = \hat{s}(n) + \beta \cdot \hat{s}(n-1)$$

β の値としては典型的には 0 . 6 8 等を用いるが、この値に限定されない。

【0 1 0 3】

音声パラメータ紛失処理部 1 2 3 は、音声合成部 1 2 4 で用いた音声パラメータ（ISF パラメータ、ピッチラグ、適応符号帳ゲイン、固定符号帳ゲイン）をバッファに格納する（図 7 ステップ 1 4 5）。

【0 1 0 4】

< パケットロスを検出した場合 >

音声パラメータ紛失処理部 1 2 3 は、補助情報蓄積部 1 2 6 からピッチラグ

【数 4 4】

$$\hat{T}_p^{(j)} (0 \leq j < M_{la})$$

を読み出し、音声パラメータを予測する。音声パラメータ紛失処理部 1 2 3 の機能構成例を図 1 2 に示し、音声パラメータ予測の処理手順を図 1 3 に示す。

【0 1 0 5】

ISF 予測部 1 9 1 は、直前フレームに関する ISF パラメータと、過去数フレームにおいて算出した ISF パラメータを用いて ISF パラメータを算出する（図 1 3 ステップ 1 1 0 1）。ISF 予測部 1 9 1 の処理手順を図 1 0 に示す。

【0 1 0 6】

最初に、直前フレームの ISF パラメータを用いてバッファを更新する（図 1 0 ステップ 1 7 1）。次に、以下の式に従い ISF パラメータ

【数 4 5】

$$\dot{\omega}_i$$

を算出する（図 1 0 ステップ 1 7 2）。

【0 1 0 7】

【数 4 6】

$$\dot{\omega}_i = \alpha \omega_i^{(-1)} + (1 - \alpha) \vec{\omega}_i$$

【数 4 7】

$$\vec{\omega}_i = \beta \omega_i^C + (1 - \beta) \frac{\omega_i^{(-3)} + \omega_i^{(-2)} + \omega_i^{(-1)}}{3}$$

10

20

30

40

50

ここで、

【数 4 8】

$$\omega_i^{(-j)}$$

はバッファに格納された j フレーム前の I S F パラメータである。また、

【数 4 9】

$$\omega_i^C$$

10

, α , β は、符号化側で用いた値と同様である。

【0 1 0 8】

次に、

【数 5 0】

$$\dot{\omega}_i$$

が

【数 5 1】

$$0 < \dot{\omega}_0 < \dot{\omega}_1 < \dots < \dot{\omega}_{14}$$

20

を満たすよう i の値を並び替え、隣り合う

【数 5 2】

$$\dot{\omega}_i$$

同士が接近しすぎないように調整する。

【0 1 0 9】

【数 5 3】

$$\dot{\omega}_i$$

30

の値を調整手順には、例えば非特許文献 4 (式 1 5 1) を用いることができる (図 1 0 ステップ 1 7 3)。

【0 1 1 0】

ピッチラグ予測部 1 9 2 は、補助情報蓄積部 1 2 6 から補助情報符号を復号してピッチラグ

【数 5 4】

$$\hat{T}_p^{(i)} (0 \leq i < M_{la})$$

40

を得る。さらに、過去に復号に用いたピッチラグ

【数 5 5】

$$\hat{T}_p^{(-j)} (0 \leq j < J)$$

を用いて、ピッチラグ

【数 5 6】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

を出力する。ここで、1フレームに含まれるサブフレームの数はM、補助情報に含まれるピッチラグの数を M_{1a} である。ピッチラグ

【数 5 7】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

10

の予測に当たっては、例えば非特許文献4の7.11.1.3節に記載の処理手順を用いることができる(図13ステップ1102)。

【0111】

適応符号帳ゲイン予測部193は、事前に定めた適応符号帳ゲイン

【数 5 8】

$$g_p^C$$

と、過去に復号に用いた適応符号帳ゲイン

【数 5 9】

20

$$g_p^{(j)} (0 \leq j < J)$$

を用いて、適応符号帳ゲイン

【数 6 0】

$$g_p^{(i)} (M_{la} \leq i < M)$$

を出力する。ここで、1フレームに含まれるサブフレームの数はM、補助情報に含まれるピッチラグの数を M_{1a} である。適応符号帳ゲイン

30

【数 6 1】

$$g_p^{(i)} (M_{la} \leq i < M)$$

の予測に当たっては、例えば、非特許文献4の7.11.2.5.3節記載の処理手順を用いることができる(図13ステップ1103)。

【0112】

固定符号帳ゲイン予測部194は、過去に復号に用いた固定符号帳ゲイン

【数 6 2】

$$g_c^{(j)} (0 \leq j < J)$$

40

を用いて、固定符号帳ゲイン

【数 6 3】

$$g_c^{(i)} (0 \leq i < M)$$

を出力する。ここで、1フレームに含まれるサブフレームの数はMである。固定符号帳ゲイン

【数 6 4】

$$g_c^{(i)} (0 \leq i < M)$$

の予測に当たっては、例えば、非特許文献 4 の 7 . 1 1 . 2 . 6 節に記載の処理手順を用いることができる（図 1 3 ステップ 1 1 0 4）。

【0 1 1 3】

雑音信号生成部 1 9 5 は、長さ L の白色雑音を出力する（図 1 3 ステップ 1 1 0 5）。ここで、1 フレームの長さを L とした。

【0 1 1 4】

音声合成部 1 2 4 は、音声パラメータ紛失処理部 1 2 3 が出力した音声パラメータから復号信号を合成する（図 7 ステップ 1 4 4）。音声合成部 1 2 4 の動作は、＜音声パケットを正常に受信した場合＞の音声合成部の動作と同様であるので、詳細を省略する（図 7 ステップ 1 4 4）。

【0 1 1 5】

音声パラメータ紛失処理部 1 2 3 は、音声合成部 1 2 4 で用いた音声パラメータ（ISF パラメータ、ピッチラグ、適応符号帳ゲイン、固定符号帳ゲイン）をバッファに格納する（図 7 ステップ 1 4 5）。

【0 1 1 6】

上記実施例では、先読み信号に含まれる全てのサブフレームに関する補助情報を符号化して伝送する例を述べたが、特定のサブフレームに関する補助情報のみを伝送する構成としてもよい。

【0 1 1 7】

[変形例 1 - 1]

実施例 1 の変形例として、ピッチゲインを補助情報に追加する例を示す。変形例 1 - 1 と実施例 1 との差分は、励振ベクトル合成部 1 5 5 の動作のみであるので、その他の部分について説明を省略する。

【0 1 1 8】

＜符号化側＞

励振ベクトル合成部 1 5 5 の処理手順を図 1 4 に示す。

【0 1 1 9】

適応符号帳ベクトル $v'(n)$ とターゲット信号 $x(n)$ から、適応符号帳ゲイン

【数 6 5】

$$g_p^C$$

を次式に従い算出する（図 1 4 ステップ 1 1 1 1）。

【0 1 2 0】

【数 6 6】

$$g_p = \frac{\sum_{n=0}^{L'-1} x(n)y(n)}{\sum_{n=0}^{L'-1} y(n)y(n)}, \quad \text{bounded by } 0 \leq g_p \leq 1.2,$$

但し、 $y(n)$ は適応符号帳ベクトルにインパルス応答を畳み込んで得られる信号 $y(n) = v(n) * h(n)$ である。

【0 1 2 1】

算出した適応符号帳ゲインを符号化し、補助情報符号に含める（図 1 4 ステップ 1 1 1 2）。符号化には、事前に学習により求めたコードブックを用いたスカラ量子化を用いることができるが、符号化の手法そのものには何を用いてもよい。

【 0 1 2 2 】

適応符号帳ゲインの符号化において求めた符号を、復号して得られる適応符号帳ゲイン

【 数 6 7 】

$$\hat{g}_p$$

を適応符号帳ベクトルに乗算することにより次式に従い、励振ベクトルを算出する（図 14 ステップ 1 1 1 3）。

【 0 1 2 3 】

【 数 6 8 】

$$e(n) = \hat{g}_p \cdot v'(n)$$

< 復号側 >

励振ベクトル合成部 1 5 5 は、適応符号帳ベクトル $v'(n)$ に補助情報符号を復号することにより得られる適応符号帳ゲイン

【 数 6 9 】

$$\hat{g}_p$$

を乗算して、次式により励振信号ベクトルを出力する（図 9 ステップ 1 6 5）。

【 0 1 2 4 】

【 数 7 0 】

$$e(n) = \hat{g}_p \cdot v'(n)$$

[変形例 1 - 2]

実施例 1 の変形例として、補助情報の利用判断のためのフラグを補助情報に追加する例を示す。

【 0 1 2 5 】

< 符号化側 >

補助情報符号化部の機能構成例を図 1 7 に、補助情報符号化部の処理手順を図 1 8 に示す。実施例 1 との差分は、補助情報出力判定部 1 1 2 8（図 1 8 ステップ 1 1 3 1）のみであるので、その他の部分について説明を省略する。

【 0 1 2 6 】

補助情報出力判定部 1 1 2 8 は、次式に従い復号信号と先読み信号の segmental SNR を算出し、segmental SNR が閾値を越えるときのみフラグの値をオンにセットして補助情報に含める。

【 0 1 2 7 】

【 数 7 1 】

$$segSNR = \frac{\sum_{n=0}^{L'-1} \hat{s}^2(n)}{\sum_{n=0}^{L'-1} (s(n) - \hat{s}(n))^2}$$

一方、segmental SNR が閾値を越えないときは、フラグの値をオフにして補助情報に含める（図 1 8 ステップ 1 1 3 1）。なお、フラグの値がオンのときのみ、ピッチラゲやピッチゲインなどの補助情報をフラグに付加して伝送し、フラグの値がオフのと

10

20

30

40

50

きはフラグの値のみ伝送することにより補助情報のビット量を削減してもよい。

【 0 1 2 8 】

< 復号側 >

補助情報復号部は、補助情報符号に含まれるフラグを復号する。音声パラメータ紛失処理部は、フラグの値がオンの場合は、実施例 1 と同様の処理手順により復号信号を算出する。一方、フラグの値がオフの場合は、補助情報を用いないパケットロス隠蔽手法により復号信号を算出する（図 1 9 ステップ 1 1 5 1 ）。

【 0 1 2 9 】

[実施例 2]

本実施例では、先読み信号部分の復号音声を利用する例について述べる。説明を容易にするため、1 フレームに含まれるサブフレームの数を M サブフレーム、先読み信号の長さを M' サブフレームとする。

【 0 1 3 0 】

< 符号化側 >

音声信号送信装置は、図 2 0 に示す通り、主要符号化部 2 1 1、補助情報符号化部 2 1 2、隠蔽信号蓄積部 2 1 3、誤差信号符号化部 2 1 4 からなる。音声信号送信装置の処理手順を図 2 2 に示す。

【 0 1 3 1 】

誤差信号符号化部 2 1 4 は、隠蔽信号蓄積部 2 1 3 から 1 サブフレーム分の隠蔽信号を読み出し、音声信号から減算して、誤差信号を算出する（図 2 2 ステップ 2 2 1 ）。

【 0 1 3 2 】

誤差信号符号化部 2 1 4 は誤差信号を符号化する。具体的な処理手順として非特許文献 4 の 6 . 8 . 4 . 1 . 5 節に記載の A V Q 等を利用する。誤差信号の符号化にあたって、ローカルデコードを行い、復号誤差信号を出力する（図 2 2 ステップ 2 2 2 ）。

【 0 1 3 3 】

復号誤差信号を隠蔽信号に加算することにより、1 サブフレーム分の復号信号を出力する（図 2 2 ステップ 2 2 3 ）。

【 0 1 3 4 】

上記、ステップ 2 2 1 ~ 2 2 3 を隠蔽信号終了まで M' サブフレーム分繰り返す。

【 0 1 3 5 】

主要符号化部 2 1 1 の機能構成を図 2 1 に示す。主要符号化部 2 1 1 は、I S F 符号化部 2 0 1 1、ターゲット信号算出部 2 0 1 2、ピッチラグ算出部 2 0 1 3、適応符号帳算出部 2 0 1 4、固定符号帳算出部 2 0 1 5、ゲイン算出部 2 0 1 6、励振ベクトル算出部 2 0 1 7、合成フィルタ 2 0 1 8、適応符号帳バッファ 2 0 1 9 からなる。

【 0 1 3 6 】

I S F 符号化部 2 0 1 1 は、符号化対象フレームおよび先読み信号に対してレビンソンダービン法を適用して L P 係数を得る。次に L P 係数を I S F パラメータに変換して符号化する。次に符号を復号して復号 I S F パラメータを得る。最後に復号 I S F パラメータを補間したうえで、サブフレーム毎の復号 L P 係数を得る。レビンソンダービン法、L P 係数からの I S F パラメータ変換の処理手順は実施例 1 と同様である。また、I S F パラメータの符号化には、例えば非特許文献 4 の 6 . 8 . 2 節に記載の処理手順を用いる。I S F 符号化部 2 0 1 1 により、I S F パラメータを符号化したインデックス、復号 I S F パラメータ、および復号 I S F パラメータを L P 係数に変換して得られる復号 L P 係数が得られる（図 2 2 ステップ 2 2 4 ）。

【 0 1 3 7 】

ターゲット信号算出部 2 0 1 2 の詳細な処理手順は実施例 1 の図 9 ステップ 1 6 2 と同様である（図 2 2 ステップ 2 2 5 ）。

【 0 1 3 8 】

ピッチラグ算出部 2 0 1 3 は、適応符号帳バッファを参照して、ターゲット信号を用いてピッチラグ、および長期予測パラメータを算出する。ピッチラグ、および長期予測パラ

10

20

30

40

50

メータ算出の詳細な処理手順は実施例 1 と同様である（図 2 2 ステップ 2 2 6 ）。

【 0 1 3 9 】

適応符号帳算出部 2 0 1 4 は、ピッチラグ算出部 2 0 1 3 で求めたピッチラグおよび長期予測パラメータを用いて適応符号帳ベクトルを算出する。適応符号帳算出部 2 0 1 4 の詳細な処理手順は、実施例 1 と同様である（図 2 2 ステップ 2 2 7 ）。

【 0 1 4 0 】

固定符号帳算出部 2 0 1 5 は、ターゲット信号および適応符号帳ベクトルを用いて、固定符号帳ベクトルおよび固定符号帳ベクトルを符号化して得られるインデックスを算出する。詳細な手順は誤差信号符号化部 2 1 4 で用いた A V Q の処理手順と同様である（図 2 2 ステップ 2 2 8 ）。

10

【 0 1 4 1 】

ゲイン算出部 2 0 1 6 は、ターゲット信号、適応符号帳ベクトル、固定符号帳ベクトルを用いて、適応符号帳ゲイン、固定符号帳ゲイン、およびこれら 2 つのゲインを符号化して得られるインデックスを算出する。詳細な処理手順として非特許文献 4 の 6 . 8 . 4 . 1 . 6 節に記載の処理手順を利用できる（図 2 2 ステップ 2 2 9 ）。

【 0 1 4 2 】

励振ベクトル算出部 2 0 1 7 は、ゲインを適用した適応符号帳ベクトルおよび固定符号帳ベクトルを加算して励振ベクトルを算出する。詳細な処理手順は実施例 1 と同様である。さらに励振ベクトル算出部 2 0 1 7 は、励振ベクトルを用いて適応符号帳バッファ 2 0 1 9 の状態を更新する。詳細な処理手順は実施例 1 と同様である（図 2 2 ステップ 2 2 1 0 ）。

20

【 0 1 4 3 】

合成フィルタ 2 0 1 8 は、復号 L P 係数および励振ベクトルを用いて復号信号を合成する（図 2 2 ステップ 2 2 1 1 ）。

【 0 1 4 4 】

上記、ステップ 2 2 4 ~ 2 2 1 1 を符号化対象フレーム終了まで M - M ' サブフレーム分繰り返す。

【 0 1 4 5 】

補助情報符号化部 2 1 2 は先読み信号 M ' サブフレームに対して、補助情報を算出する。具体的な処理手順は実施例 1 と同様である（図 2 2 ステップ 2 2 1 2 ）。

30

【 0 1 4 6 】

実施例 1 の手順に加えて、実施例 2 では、補助情報符号化部 2 1 2 の合成フィルタ 1 5 7 が出力する復号信号を隠蔽信号蓄積部 2 1 3 に蓄積する（図 2 2 ステップ 2 2 1 3 ）。

【 0 1 4 7 】

< 復号部 >

図 2 3 に示す通り、音声信号受信装置は、音声符号バッファ 2 3 1 、音声パラメータ復号部 2 3 2 、音声パラメータ紛失処理部 2 3 3 、音声合成部 2 3 4 、補助情報復号部 2 3 5 、補助情報蓄積部 2 3 6 、誤差信号復号部 2 3 7 、隠蔽信号蓄積部 2 3 8 からなる。音声信号受信装置の処理手順を図 2 4 に示す。音声合成部 2 3 4 の機能構成を図 2 5 に示す。

【 0 1 4 8 】

40

音声符号バッファ 2 3 1 は、パケットを正常に受信したか否かを判断し、パケットを正常に受信したと判断した場合には、音声パラメータ復号部 2 3 2 、補助情報復号部 2 3 5 、誤差信号復号部 2 3 7 に処理を切替え、パケットを正常に受信できなかったと判断した場合には、音声パラメータ紛失処理部 2 3 3 に処理を切り替える（図 2 4 ステップ 2 4 1 ）。

【 0 1 4 9 】

< 正常にパケットを受信した場合 >

誤差信号復号部 2 3 7 は誤差信号符合を復号して復号誤差信号を得る。具体的な処理手順として非特許文献 4 の 7 . 1 . 2 . 1 . 2 節に記載の A V Q 等、符号化側で用いた方法に対応した復号方法を用いる（図 2 4 ステップ 2 4 2 ）。

50

【 0 1 5 0 】

先読み励振ベクトル合成部 2 3 1 8 は、隠蔽信号蓄積部 2 3 8 から 1 サブフレーム分の隠蔽信号を読み出し、復号誤差信号に加算することにより、1 サブフレーム分の復号信号を出力する（図 2 4 ステップ 2 4 3）。

【 0 1 5 1 】

上記、ステップ 2 4 1 ~ 2 4 3 を隠蔽信号終了まで M' サブフレーム分繰り返す。

【 0 1 5 2 】

音声パラメータ復号部 2 3 2 は、ISF 復号部 2 2 1 1、ピッチラグ復号部 2 2 1 2、ゲイン復号部 2 2 1 3、固定符号帳復号部 2 2 1 4 からなる。音声パラメータ復号部 2 3 2 の機能構成例を図 2 6 に示す。

10

【 0 1 5 3 】

ISF 復号部 2 2 1 1 は、ISF 符号を復号し、LP 係数に変換することにより復号 LP 係数を得る。例えば非特許文献 4 の 7 . 1 . 1 節に記載の処理手順を用いる（図 2 4 ステップ 2 4 4）。

【 0 1 5 4 】

ピッチラグ復号部 2 2 1 2 は、ピッチラグ符号を復号してピッチラグおよび長期予測パラメータを得る（図 2 4 ステップ 2 4 5）。

【 0 1 5 5 】

ゲイン復号部 2 2 1 3 は、ゲイン符号を復号して適応符号帳ゲイン、固定符号帳ゲインを得る。詳細な処理手順は非特許文献 4 の 7 . 1 . 2 . 1 . 3 節に記載の通りである（図 2 4 ステップ 2 4 6）。

20

【 0 1 5 6 】

適応符号帳算出部 2 3 1 3 は、ピッチラグおよび長期予測パラメータを用いて適応符号帳ベクトルを算出する。適応符号帳算出部 2 3 1 3 の詳細な処理手順は、実施例 1 に記載の通りである（図 2 4 ステップ 2 4 7）。

【 0 1 5 7 】

固定符号帳復号部 2 2 1 4 は、固定符号帳符号を復号して、固定符号帳ベクトルを算出する。詳細な手順は非特許文献 4 の 7 . 1 . 2 . 1 . 2 節に記載の通りである（図 2 4 ステップ 2 4 8）。

【 0 1 5 8 】

30

励振ベクトル合成部 2 3 1 4 は、ゲインを適用した適応符号帳ベクトルおよび固定符号帳ベクトルを加算して励振ベクトルを算出する。さらに励振ベクトル算出部は、励振ベクトルを用いて適応符号帳バッファを更新する（図 2 4 ステップ 2 4 9）。詳細な処理手順は実施例 1 と同様である。

【 0 1 5 9 】

合成フィルタ 2 3 1 6 は、復号 LP 係数および励振ベクトルを用いて復号信号を合成する（図 2 4 ステップ 2 4 1 0）。詳細な処理手順は実施例 1 と同様である。

【 0 1 6 0 】

上記、ステップ 2 4 4 ~ 2 4 1 0 を符号化対象フレーム終了まで M - M' サブフレーム分繰り返す。

40

【 0 1 6 1 】

補助情報復号部 2 3 5 の機能構成は実施例 1 と同様である。補助情報復号部 2 3 5 は、補助情報符号を復号してピッチラグを算出する（図 2 4 ステップ 2 4 1 1）。

【 0 1 6 2 】

音声パラメータ紛失処理部 2 3 3 の機能構成は実施例 1 と同様である。

【 0 1 6 3 】

ISF 予測部 1 9 1 は、直前フレームの ISF パラメータを用いて ISF パラメータを予測し、LP 係数に変換する。処理手順は実施例 1 の図 1 0 のステップ 1 7 2、1 7 3、1 7 4 と同様である（図 2 4 ステップ 2 4 1 2）。

【 0 1 6 4 】

50

適応符号帳算出部 2 3 1 3 は、補助情報復号部 2 3 5 が出力したピッチラグと、適応符号帳 2 3 1 2 を用いて適応符号帳ベクトルを算出する（図 2 4 ステップ 2 4 1 3）。処理手順は図 1 6 ステップ 1 1 3 0 1、1 1 3 0 2 と同様である。

【0 1 6 5】

適応符号帳ゲイン予測部 1 9 3 は、適応符号帳ゲインを出力する。具体的な処理手順は図 1 3 ステップ 1 1 0 3 と同様である（図 2 4 ステップ 2 4 1 4）。

【0 1 6 6】

固定符号帳ゲイン予測部 1 9 4 は、固定符号帳ゲインを出力する。具体的な処理手順は図 1 3 ステップ 1 1 0 4 と同様である（図 2 4 ステップ 2 4 1 5）。

【0 1 6 7】

雑音信号生成部 1 9 5 は、白色雑音を出力し、固定符号帳ベクトルとする。処理手中は図 1 3 ステップ 1 1 0 5 と同様である（図 2 4 ステップ 2 4 1 6）。

【0 1 6 8】

励振ベクトル合成部 2 3 1 4 は、適応符号帳ベクトルおよび固定符号帳ベクトルに各々ゲインを適用した上で加算し、励振ベクトルを算出する。また励振ベクトルにより適応符号帳バッファを更新する（図 2 4 ステップ 2 4 1 7）。

【0 1 6 9】

合成フィルタ 2 3 1 6 は、上記 L P 係数と励振ベクトルを用いて復号信号を算出する。算出した復号信号で隠蔽信号蓄積部 2 3 8 を更新する（図 2 4 ステップ 2 4 1 8）。

【0 1 7 0】

上記ステップを M' サブフレーム分繰り返し、復号信号を音声信号として出力する。

【0 1 7 1】

< パケットをロスした場合 >

隠蔽信号蓄積部から 1 サブフレーム分の隠蔽信号を読み出し、復号信号とする（図 2 4 ステップ 2 4 1 9）。

【0 1 7 2】

上記を M' サブフレーム分繰り返す。

【0 1 7 3】

I S F 予測部 1 9 1 は、I S F パラメータを予測する（図 2 4 ステップ 2 4 2 0）。処理手順として、図 1 3 ステップ 1 1 0 1 を用いる。

【0 1 7 4】

ピッチラグ予測部 1 9 2 は、過去に復号に用いたピッチラグを用いて予測ピッチラグを出力する（図 2 4 ステップ 2 4 2 1）。予測に用いる処理手順は、実施例 1 の図 1 3 ステップ 1 1 0 2 と同様である。

【0 1 7 5】

適応符号帳ゲイン予測部 1 9 3、固定符号帳ゲイン予測部 1 9 4、雑音信号生成部 1 9 5、音声合成部 2 3 4 の動作は実施例 1 と同様である（図 2 4 ステップ 2 4 2 2）。

【0 1 7 6】

上記ステップを M サブフレーム分繰り返し、M - M' サブフレーム分の復号信号は音声信号として出力し、残りの M' サブフレーム分の復号信号で隠蔽信号蓄積部 2 3 8 を更新する。

【0 1 7 7】

[実施例 3]

適応符号帳ベクトルの算出に当たって、声門パルス同期を用いる場合について述べる。

【0 1 7 8】

< 符号化側 >

音声信号送信装置の機能構成は実施例 1 と同じである。機能構成および処理手順が異なるのは補助情報符号化部のみであるので、ここでは補助情報符号化部の動作についてのみ述べる。

【0 1 7 9】

10

20

30

40

50

補助情報符号化部は、LP係数算出部311、ピッチラグ予測部312、ピッチラグ選定部313、ピッチラグ符号化部314、適応符号帳バッファ315からなる。補助情報符号化部の機能構成図を図27に、処理手順を図28に示す。

【0180】

LP係数算出部311は、実施例1のLP係数算出部と同じであるので説明を省略する(図28ステップ321)。

【0181】

ピッチラグ予測部312は、音声符号化部から得たピッチラグを用いてピッチラグ予測値

【数72】

$$\hat{T}_p$$

10

を算出する(図28ステップ322)。予測の具体的な処理は実施例1におけるピッチラグ予測部192における、ピッチラグ

【数73】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

の予測と同様である(図13ステップ1102と同じである)。

20

【0182】

次にピッチラグ選定部313は、補助情報として伝送するピッチラグを決定する(図28ステップ323)。ピッチラグ選定部313のより詳細な処理手順を図29に示す。

【0183】

最初に、ピッチラグ予測値

【数74】

$$\hat{T}_p$$

および過去のピッチラグの値

30

【数75】

$$\hat{T}_p^{(-j)} (0 \leq j < J)$$

から、次式に従いピッチラグコードブックを生成する(図29ステップ331)。

【0184】

【数76】

$$\langle \hat{T}_p - \hat{T}_p^{(-1)} \geq 0 \text{ の場合} \rangle$$

40

$$\hat{T}_C^j = \begin{cases} \hat{T}_p & (j=0) \\ \hat{T}_p^{(-1)} - j \cdot \delta_j + \rho & (0 < j < I) \end{cases}$$

【数 7 7】

 $\langle \hat{T}_p - \hat{T}_p^{(-1)} < 0 \text{ の場合} \rangle$

$$\hat{T}_C^j = \begin{cases} \hat{T}_p & (j=0) \\ \hat{T}_p^{(-1)} + j \cdot \delta_j + \rho & (0 < j < I) \end{cases}$$

ここで、1 サブフレーム前のピッチラグの値は

【数 7 8】

$$\hat{T}_p^{(-1)}$$

10

である。またコードブックのインデクス数を I とする。また δ_j は事前に定めたステップ幅であり、 ρ は事前に定めた定数である。

【0 1 8 5】

次に適応符号帳、ピッチラグ予測値

【数 7 9】

$$\hat{T}_p$$

20

を用いて、次式に従い初期励振ベクトル $u_0(n)$ を生成する (図 2 9 ステップ 3 3 2)。

【0 1 8 6】

【数 8 0】

$$u_0(n) = \begin{cases} 0.18u_0(n - \hat{T}_p - 1) + 0.64u_0(n - \hat{T}_p) + 0.18u_0(n - \hat{T}_p + 1) & (0 \leq n < \hat{T}_p) \\ u_0(n - \hat{T}_p) & (\hat{T}_p \leq n < L) \end{cases}$$

30

初期励振ベクトル算出の処理手順は、非特許文献 4 の式 (6 0 7) および式 (6 0 8) と同様である。

【0 1 8 7】

次に初期励振ベクトルに対して、ピッチラグコードブック中の全ての候補ピッチラグ

【数 8 1】

$$\hat{T}_C^j (0 \leq j < J)$$

を用いて声門パルス同期を適用し、候補適応符号帳ベクトル $u^j(n)$ ($0 \leq j < I$) を生成する (図 2 9 ステップ 3 3 3)。声門パルス同期は非特許文献 4 の 7 . 1 1 . 2 . 5 節におけるパルス位置が利用できない場合と同様の処理手順を用いる。ただし、非特許文献 4 における $u(n)$ は本発明の実施形態の $u_0(n)$ に対応し、 $extrapolated\ pitch$ は本発明の実施形態の

40

【数 8 2】

$$\hat{T}_C^j$$

に対応し、 $the\ last\ reliable\ pitch(T_c)$ は本発明の実施形態の

【数 8 3】

$$\hat{T}_p^{(-1)}$$

に対応する。

【0 1 8 8】

候補適応符号帳ベクトル $u^j(n)$ ($0 \leq j < I$) について、評価尺度を計算する (図 29 ステップ 334)。評価尺度として segmental SNR を用いる場合には、LP 係数を用いた逆フィルタリングにより信号を合成し、入力信号との間で次式に従い segmental SNR を算出する。

10

【0 1 8 9】

【数 8 4】

$$\hat{s}_j(n) = u^j(n) - \sum_{i=1}^P \hat{a}(i) \cdot \hat{s}_j(n-i)$$

【数 8 5】

$$segSNR_j = \frac{\sum_{n=0}^{L'-1} \hat{s}_j^2(n)}{\sum_{n=0}^{L'-1} (s(n) - \hat{s}_j(n))^2}$$

20

逆フィルタリングを行う代わりに、次式に従い、残差信号を用いて適応符号帳ベクトルの領域で segmental SNR を算出してもよい。

【0 1 9 0】

【数 8 6】

$$r(n) = s(n) + \sum_{i=1}^P \hat{a}(i) \cdot s(n-i)$$

30

【数 8 7】

$$segSNR_j = \frac{\sum_{n=0}^{L'-1} u^j(n)}{\sum_{n=0}^{L'-1} (r(n) - u^j(n))^2}$$

この場合、LP 係数を用いて先読み信号 $s(n)$ ($0 \leq n < L'$) の残差信号 $r(n)$ を算出する (図 11 ステップ 181)。

40

【0 1 9 1】

ステップ 334 で算出した評価尺度のうち最大のものに対応するインデックスを選択し、当該インデックスに対応するピッチラグを求める。(図 29 ステップ 335)。

【0 1 9 2】

【数 8 8】

$$\arg \max_j [segSNR_j]$$

50

< 復号側 >

音声信号受信装置の機能構成は実施例 1 と同様である。実施例 1 との差分は音声パラメータ紛失処理部 1 2 3、補助情報復号部 1 2 5、補助情報蓄積部 1 2 6 の機能構成と処理手順であるので、これらについてのみ述べる。

【 0 1 9 3 】

< 正常にパケットを受信した場合 >

補助情報復号部 1 2 5 は、補助情報符号を復号してピッチラグ

【 数 8 9 】

$$\hat{T}_C^{idx}$$

10

を算出し、補助情報蓄積部 1 2 6 に格納する。補助情報復号部 1 2 5 の処理手順を図 3 0 に示す。

【 0 1 9 4 】

ピッチラグ算出に当たって、最初にピッチラグ予測部 3 1 2 は、音声復号部から得たピッチラグを用いてピッチラグ予測値

【 数 9 0 】

$$\hat{T}_p$$

20

を算出する (図 3 0 ステップ 3 4 1)。予測の具体的な処理は実施例 3 における図 2 8 ステップ 3 2 2 と同様である。

【 0 1 9 5 】

次に、ピッチラグ予測値

【 数 9 1 】

$$\hat{T}_p$$

および過去のピッチラグの値

【 数 9 2 】

$$\hat{T}_p^{(-j)} (0 \leq j < J)$$

30

から、次式に従いピッチラグコードブックを生成する (図 3 0 ステップ 3 4 2)。

【 0 1 9 6 】

【 数 9 3 】

$$< \hat{T}_p - \hat{T}_p^{(-1)} \geq 0 \text{ の場合} >$$

$$\hat{T}_C^j = \begin{cases} \hat{T}_p & (j=0) \\ \hat{T}_p^{(-1)} - j \cdot \delta_j + \rho & (0 < j < I) \end{cases}$$

40

【 数 9 4 】

$$< \hat{T}_p - \hat{T}_p^{(-1)} < 0 \text{ の場合} >$$

$$\hat{T}_C^j = \begin{cases} \hat{T}_p & (j=0) \\ \hat{T}_p^{(-1)} + j \cdot \delta_j + \rho & (0 < j < I) \end{cases}$$

50

処理手順は図 2 9 ステップ 3 3 1 と同様である。ここで、1 サブフレーム前のピッチラグの値は

【数 9 5】

$$\hat{T}_p^{(-1)}$$

である。またコードブックのインデクス数を I とする。また δ_j は事前に定めたステップ幅であり、 ρ は事前に定めた定数である。

【0 1 9 7】

次に、ピッチラグコードブックを参照し、補助情報として伝送されたインデクス $i d x$ に対応するピッチラグ

【数 9 6】

$$\hat{T}_C^{idx}$$

を求め、補助情報蓄積部 1 2 6 に格納する（図 3 0 ステップ 3 4 3）。

【0 1 9 8】

< パケットロスを検出した場合 >

音声合成部の機能構成も実施例 1 と同様であるが（図 1 5 と同様）、実施例 1 と動作が異なる適応符号帳算出部 1 1 2 3 についてのみ以下に述べる。

【0 1 9 9】

音声パラメータ紛失処理部 1 2 3 は、補助情報蓄積部 1 2 6 からピッチラグを読み出した上で次式に従ってピッチラグ予測値を算出し、ピッチラグ予測部 1 9 2 の出力の代わりに用いる。

【0 2 0 0】

【数 9 7】

$$\hat{T}_p = \hat{T}_p^{(-1)} + \kappa \cdot (\hat{T}_C^{idx} - \hat{T}_p^{(-1)})$$

ここで、

【数 9 8】

$$K$$

は予め定めた定数である。

【0 2 0 1】

次に、適応符号帳、ピッチラグ予測値

【数 9 9】

$$\hat{T}_p$$

を用いて、次式に従い初期励振ベクトル $u_0(n)$ を生成する（図 2 9 ステップ 3 3 2）。

【0 2 0 2】

【数 1 0 0】

$$u_0(n) = \begin{cases} 0.18u_0(n - \hat{T}_p^{(-1)} - 1) + 0.64u_0(n - \hat{T}_p^{(-1)}) + 0.18u_0(n - \hat{T}_p^{(-1)} + 1) & (0 \leq n < \hat{T}_p^{(-1)}) \\ u_0(n - \hat{T}_p^{(-1)}) & (\hat{T}_p^{(-1)} \leq n < L) \end{cases}$$

10

20

30

40

50

次に初期励振ベクトルに対して、ピッチラグ

【数 1 0 1】

$$\hat{T}_C^{idx}$$

を用いて声門パルス同期を適用し、適応符号帳ベクトル $u(n)$ を生成する。声門パルス同期は図 29 ステップ 333 と同様の処理手順を用いる。

【0203】

続いて、上述した一連の音声信号送信装置による処理をコンピュータに実行させるための音声符号化プログラム 70 を説明する。図 31 に示すように、音声符号化プログラム 70 は、コンピュータに挿入されてアクセスされる、あるいはコンピュータが備える記録媒体 60 に形成されたプログラム格納領域 61 内に格納される。

【0204】

音声符号化プログラム 70 は、音声符号化モジュール 700 と、補助情報符号化モジュール 701 とを備えて構成される。音声符号化モジュール 700 と、補助情報符号化モジュール 701 とを実行させることにより実現される機能は、上述した音声信号送信装置の音声符号化部 111 と、補助情報符号化部 112 との機能とそれぞれ同様である。

【0205】

なお、音声符号化プログラム 70 は、その一部若しくは全部が、通信回線等の伝送媒体を介して伝送され、他の機器により受信されて記録（インストールを含む）される構成としてもよい。また、音声符号化プログラム 70 の各モジュールは、1つのコンピュータでなく、複数のコンピュータのいずれかにインストールされてもよい。その場合、当該複数のコンピュータからなるコンピュータシステムによって上述した一連の音声符号化プログラム 70 の処理が行われる。

【0206】

続いて、上述した一連の音声信号受信装置による処理をコンピュータに実行させるための音声復号プログラム 90 を説明する。図 32 に示すように、音声復号プログラム 90 は、コンピュータに挿入されてアクセスされる、あるいはコンピュータが備える記録媒体 80 に形成されたプログラム格納領域 81 内に格納される。

【0207】

音声復号プログラム 90 は、音声符号バッファモジュール 900 と、音声パラメータ復号モジュール 901 と、補助情報復号モジュール 902 と、補助情報蓄積モジュール 903 と、音声パラメータ紛失処理モジュール 904 と、音声合成モジュール 905 とを備えて構成される。音声符号バッファモジュール 900 と、音声パラメータ復号モジュール 901 と、補助情報復号モジュール 902 と、補助情報蓄積モジュール 903 と、音声パラメータ紛失処理モジュール 904 と、音声合成モジュール 905 とを実行させることにより実現される機能は、上述した音声信号受信装置の音声符号バッファ 231 と、音声パラメータ復号部 232 と、補助情報復号部 235 と、補助情報蓄積部 236 と、音声パラメータ紛失処理部 233 と、音声合成部 234 との機能とそれぞれ同様である。

【0208】

なお、音声復号プログラム 90 は、その一部若しくは全部が、通信回線等の伝送媒体を介して伝送され、他の機器により受信されて記録（インストールを含む）される構成としてもよい。また、音声復号プログラム 90 の各モジュールは、1つのコンピュータでなく、複数のコンピュータのいずれかにインストールされてもよい。その場合、当該複数のコンピュータからなるコンピュータシステムによって上述した一連の音声復号プログラム 90 の処理が行われる。

【0209】

[実施例 4]

補助情報を復号側でのピッチラグ予測に用いる例について述べる。

【0210】

10

20

30

40

50

< 符号化側 >

音声信号送信装置の機能構成は実施例 1 と同様である。機能構成および処理手順が異なるのは補助情報符号化部 1 1 2 のみであるので、ここでは補助情報符号化部 1 1 2 の動作についてのみ述べる。

【 0 2 1 1 】

補助情報符号化部 1 1 2 の機能構成図を図 3 3 に、処理手順を図 3 4 に示す。補助情報符号化部 1 1 2 は、L P 係数算出部 5 1 1、残差信号算出部 5 1 2、ピッチラグ算出部 5 1 3、適応符号帳算出部 5 1 4、適応符号帳バッファ 5 1 5、ピッチラグ符号化部 5 1 6 からなる。

【 0 2 1 2 】

L P 係数算出部 5 1 1 は、実施例 1 の図 8 の L P 係数算出部 1 5 1 と同様であるので説明を省略する。

【 0 2 1 3 】

残差信号算出部 5 1 2 は、実施例 1 の図 1 1 のステップ 1 8 1 と同様の処理により、残差信号を算出する。

【 0 2 1 4 】

ピッチラグ算出部 5 1 3 は、次式を最大化する k を求めることにより、サブフレーム毎にピッチラグを算出する（図 3 4 のステップ 1 6 3）。ここで、 $u(n)$ は適応符号帳、 L' は 1 サブフレームに含まれるサンプル数を表す。

【 0 2 1 5 】

$$T_p = \arg \max_k T_k$$

【 数 1 0 2 】

$$T_k = \frac{\sum_{n=0}^{L'-1} r(n)u(n-k)}{\sqrt{\sum_{n=0}^{L'-1} u(n-k)u(n-k)}}$$

適応符号帳算出部 5 1 4 はピッチラグ T_p 、適応符号帳 $u(n)$ から、適応符号帳ベクトル $v'(n)$ を算出する。ここで適応符号帳の長さを N_{adapt} とした。（図 3 4 のステップ 1 6 4）

$$v'(n) = u(n + N_{adapt} - T_p)$$

適応符号帳バッファ 5 1 5 は、適応符号帳ベクトル $v'(n)$ により状態を更新する（図 3 4 のステップ 1 6 6）。

【 0 2 1 6 】

$$u(n) = u(n + L') \quad (0 \leq n < N - L')$$

$$u(n + N - L') = v'(n) \quad (0 \leq n < L')$$

ピッチラグ符号化部 5 1 6 は、実施例 1 と同様なので省略する（図 3 4 のステップ 1 6 9）。

【 0 2 1 7 】

< 復号側 >

音声信号受信装置は、実施例 1 と同様に、音声符号バッファ 1 2 1、音声パラメータ復号部 1 2 2、音声パラメータ紛失処理部 1 2 3、音声合成部 1 2 4、補助情報復号部 1 2 5、補助情報蓄積部 1 2 6 からなる。音声信号受信装置の処理手順は図 7 に示す通りである。

【 0 2 1 8 】

音声符号バッファ 1 2 1 の動作は実施例 1 と同様である。

【 0 2 1 9 】

< 正常にパケットを受信した場合 >

音声パラメータ復号部 1 2 2 の動作は実施例 1 と同様である。

【 0 2 2 0 】

補助情報復号部 1 2 5 は、補助情報符号を復号してピッチラゲ

【数 1 0 3】

$$\hat{T}_p^{(j)} (0 \leq j < M_{la})$$

を算出し、補助情報蓄積部 1 2 6 に格納する。補助情報復号部 1 2 5 では、符号化側で用いた符号化方法に対応する復号方法を用いて、補助情報符号を復号する。

【0 2 2 1】

音声合成部 1 2 4 は、実施例 1 と同様である。

【0 2 2 2】

< パケットロスを検出した場合 >

音声パラメータ紛失処理部 1 2 3 (図 1 2 参照) の I S F 予測部 1 9 1 は、実施例 1 と同様にして I S F パラメータを算出する。

【0 2 2 3】

ピッチラゲ予測部 1 9 2 の処理手順を図 3 5 に示す。ピッチラゲ予測部 1 9 2 は、実施例 1 と同様にして、補助情報蓄積部 1 2 6 から補助情報符号を読み出してピッチラゲ

【数 1 0 4】

$$\hat{T}_p^{(i)} (0 \leq i < M_{la})$$

を得る (図 3 5 のステップ 4 0 5 1) 。さらに、過去に復号に用いたピッチラゲ

【数 1 0 5】

$$\hat{T}_p^{(-j)} (0 \leq j < J)$$

を用いて、ピッチラゲ

【数 1 0 6】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

を出力する (図 3 5 のステップ 4 0 5 2) 。ここで、1 フレームに含まれるサブフレームの数は M 、補助情報に含まれるピッチラゲの数を M_{la} とする。ピッチラゲ

【数 1 0 7】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

の予測に当たっては、非特許文献 4 のような処理手順を用いることができる (図 1 3 のステップ 1 1 0 2) 。

【0 2 2 4】

ここで、ピッチラゲ予測部 1 9 2 は、ピッチラゲ

【数 1 0 8】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

の予測に当たって、過去に復号に用いたピッチラゲ

【数 1 0 9】

$$\hat{T}_p^{(-j)} (1 \leq j \leq J)$$

とピッチラゲ

【数 1 1 0】

$$\hat{T}_p^{(i)} (0 \leq i < M_{la})$$

を用いてピッチラグ

【数 1 1 1】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

を予測してもよい。また、

【数 1 1 2】

$$\hat{T}_p^{(i)} = \hat{T}_p^{(M_{la})}$$

としてもよい。この場合のピッチラグ予測部の処理手順は図 3 6 のようになる。

【0 2 2 5】

さらに、ピッチラグ予測部 1 9 2 は、ピッチラグの予測値について信頼性が低い場合にのみ

【数 1 1 3】

$$\hat{T}_p^{(i)} = \hat{T}_p^{(M_{la})}$$

としてもよい。この場合のピッチラグ予測部 1 9 2 の処理手順を図 3 7 に示す。予測値を用いたか、補助情報により得られたピッチラグ

【数 1 1 4】

$$\hat{T}_p^{(M_{la})}$$

を用いたかに関する指示情報を適応符号帳算出部 1 5 4 に入力するようにしてもよい。

【0 2 2 6】

適応符号帳ゲイン予測部 1 9 3、固定符号帳ゲイン予測部 1 9 4 は実施例 1 と同様である。

【0 2 2 7】

雑音信号生成部 1 9 5 は、実施例 1 と同様である。

【0 2 2 8】

音声合成部 1 2 4 は、音声パラメータ紛失処理部 1 2 3 が出力したパラメータから符号化対象フレームに対応する音声信号を合成する。

【0 2 2 9】

音声合成部 1 2 4 (図 1 5 参照) の L P 係数算出部 1 1 2 1 は、実施例 1 と同様にして L P 係数を得る (図 1 6 のステップ 1 1 3 0 1)。

【0 2 3 0】

適応符号帳算出部 1 1 2 3 は、実施例 1 と同様にして適応符号帳ベクトルを算出する。適応符号帳算出部 1 1 2 3 は、常に適応符号帳ベクトルにフィルタリングを行うようにしてもよいし、常にフィルタリングを行わないようにしてもよい。すなわち以下の式を用いて適応符号帳ベクトルを算出する。ここで、フィルタ係数を f_i とした。

【0 2 3 1】

$$v(n) = f_{-1} v'(n-1) + f_0 v'(n) + f_1 v'(n+1)$$

フィルタリングを指示しない値をとる場合は、 $v(n) = v'(n)$ とする (適応符号帳算出ステップ A)。

【0 2 3 2】

10

20

30

40

50

適応符号帳算出部 1123 は、次の手順で適応符号帳ベクトルを算出してもよい（適応符号帳算出ステップ B）。

【0233】

ピッチラグおよび適応符号帳 1122 を用いて初期適応符号帳ベクトルを算出する。

【0234】

$$v(n) = f_{-1} v'(n-1) + f_0 v'(n) + f_1 v'(n+1)$$

設計方針によって、 $v(n) = v'(n)$ としてもよい。

【0235】

次に初期適応符号帳ベクトルに対して、声門パルス同期を適用する。声門パルス同期は非特許文献 4 の 7.11.2.5 節におけるパルス位置が利用できない場合と同様の処理手順を用いる。ただし、非特許文献 4 における $u(n)$ は本発明の実施形態の $v(n)$ に対応し、*extrapolated pitch* は本発明の実施形態の

10

【数 115】

$$\hat{T}_p^{(M-1)}$$

に対応し、*the last reliable pitch* (T_c) は本発明の実施形態の

【数 116】

$$\hat{T}_p^{(M_{in}-1)}$$

20

に対応する。

【0236】

さらに、適応符号帳算出部 1123 は、ピッチラグ予測部 192 が、上記予測値の指示情報を出している場合に、上記指示情報が、補助情報として送られてきたピッチラグを予測値として用いないことを示す場合（図 38 のステップ 4082：NO）に上記適応符号帳算出ステップ A を用い、それ以外の場合（図 38 のステップ 4082：YES）には上記適応符号帳算出ステップ B を用いるようにしてもよい。この場合の適応符号帳算出部 1123 の処理手順を図 38 に示す。

30

【0237】

励振ベクトル合成部 1124 は、実施例 1 と同様にして、励振信号ベクトルを出力する（図 16 のステップ 11306）。

【0238】

ポストフィルタ 1125 は、実施例 1 と同様にして、合成信号に後処理を加える。

【0239】

適応符号帳 1122 は、実施例 1 と同様にして、励振信号ベクトルにより状態を更新する（図 16 のステップ 11308）。

【0240】

合成フィルタ 1126 は、実施例 1 と同様にして、復号信号を合成する（図 16 のステップ 11309）。

40

【0241】

聴覚重み逆フィルタ 1127 は、実施例 1 と同様にして、聴覚重み逆フィルタを適用する。

【0242】

音声パラメータ紛失処理部 123 は、実施例 1 と同様にして、音声合成部 124 で用いた音声パラメータ（ISF パラメータ、ピッチラグ、適応符号帳ゲイン、固定符号帳ゲイン）をバッファに格納する（図 7 のステップ 145）。

【0243】

[実施例 5]

50

本実施例では、特定のフレームクラスにおいてのみ補助情報としてピッチラグを伝送し、それ以外では、ピッチラグを伝送しない構成について述べる。

【0244】

<送信側>

音声信号送信装置において、入力音声信号は、音声符号化部111に送られる。

【0245】

本実施例における音声符号化部111は、符号化対象フレームの性質を表すインデクスを必ず算出し、補助情報符号化部112に伝送する。それ以外の動作については、実施例1と同様である。

【0246】

補助情報符号化部112において、実施例1-4との違いは、ピッチラグ符号化部158のみであるので、ピッチラグ符号化部158の動作について、以下に述べる。実施例5における補助情報符号化部112の構成図を図39に示す。

【0247】

ピッチラグ符号化部158の処理手順を図40に示す。ピッチラグ符号化部158は、符号化対象フレームの性質を現すインデクスを読み出し(図40のステップ5021)、符号化対象フレームの性質を現すインデクスが、事前に定めた値と等しい場合は、補助情報に割り当てるビット数をBビットとする($B > 1$)。一方、符号化対象フレームの性質を現すインデクスが事前に定めた値と異なる場合は、補助情報に割り当てるビット数を1ビットとする(図40のステップ5022)。

【0248】

補助情報に割り当てるビット数が1ビットの場合(図40のステップ5022:NO)、補助情報インデクスに補助情報を伝送していないことを示す値をセットし、補助情報符号とする(図40のステップ5023)。

【0249】

一方、補助情報に割り当てるビット数がBビットの場合(図40のステップ5022:YES)、補助情報インデクスに補助情報を伝送することを示す値をセットし(図40のステップ5024)、さらにピッチラグを実施例1の方法により符号化して得られるB-1ビットの符号を含めて、補助情報符号とする(図40のステップ5025)。

【0250】

<復号側>

音声信号受信装置は、実施例1と同様に、音声符号バッファ121、音声パラメータ復号部122、音声パラメータ紛失処理部123、音声合成部124、補助情報復号部125、補助情報蓄積部126からなる。音声信号受信装置の処理手順は図7に示す通りである。

【0251】

音声符号バッファ121の動作は実施例1と同様である。

【0252】

<正常にパケットを受信した場合>

音声パラメータ復号部122の動作は実施例1と同様である。

【0253】

補助情報復号部125の処理手順を図41に示す。補助情報復号部125は、最初に補助情報符号に含まれる補助情報インデクスを復号する(図41のステップ5031)。補助情報インデクスが補助情報を伝送していないことを表す場合、それ以上の復号動作を行わない。また、補助情報インデクスの値を補助情報蓄積部126に格納する(図41のステップ5032)。

【0254】

一方、補助情報インデクスが、補助情報を伝送することを表す場合、さらにB-1ビットの復号を行い、ピッチラグ

10

20

30

40

【数 1 1 7】

$$\hat{T}_p^{(j)} (0 \leq j < M_{la})$$

を算出し、補助情報蓄積部 1 2 6 に格納する（図 4 1 のステップ 5 0 3 3）。また、補助情報インデックスの値を補助情報蓄積部 1 2 6 に格納する。なお、B - 1 ビットの補助情報の復号は、実施例 1 の補助情報復号部 1 2 5 と同様の動作である。

【0 2 5 5】

音声合成部 1 2 4 は、実施例 1 と同様である。

【0 2 5 6】

< パケットロスを検出した場合 >

音声パラメータ紛失処理部 1 2 3（図 1 2 参照）の ISF 予測部 1 9 1 は、実施例 1 と同様にして ISF パラメータを算出する。

【0 2 5 7】

ピッチラグ予測部 1 9 2 の処理手順を図 4 2 に示す。ピッチラグ予測部 1 9 2 は、補助情報蓄積部 1 2 6 から補助情報インデックスを読み出して（図 4 2 のステップ 5 0 4 1）、補助情報を伝送することを表す値であるかどうかを調べる（図 4 2 のステップ 5 0 4 2）。

【0 2 5 8】

< 補助情報インデックスが補助情報を伝送することを表す値である場合 >

実施例 1 と同様にして、補助情報蓄積部 1 2 6 から補助情報符号を読み出してピッチラグ

【数 1 1 8】

$$\hat{T}_p^{(i)} (0 \leq i < M_{la})$$

を得る（図 4 2 の 5 0 4 3）。さらに、過去に復号に用いたピッチラグ

【数 1 1 9】

$$\hat{T}_p^{(-j)} (0 \leq j < J)$$

および補助情報として得られた

【数 1 2 0】

$$\hat{T}_p^{(i)} (0 \leq i < M_{la})$$

を用いて、ピッチラグ

【数 1 2 1】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

を出力する（図 4 2 のステップ 5 0 4 4）。ここで、1 フレームに含まれるサブフレームの数は M、補助情報に含まれるピッチラグの数を M_{1a} とする。ピッチラグ

【数 1 2 2】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

の予測に当たっては、非特許文献 4 のような処理手順を用いることができる（図 1 3 のステップ 1 1 0 2）。また、

10

20

30

40

【数 1 2 3】

$$\hat{T}_p^{(i)} = \hat{T}_p^{(M_{la})}$$

としてもよい。

【0 2 5 9】

さらに、ピッチラグ予測部 1 9 2 は、ピッチラグの予測値について信頼性が低い場合にのみ

【数 1 2 4】

$$\hat{T}_p^{(i)} = \hat{T}_p^{(M_{la})}$$

10

とし、それ以外の場合には予測値を

【数 1 2 5】

$$\hat{T}_p^{(i)}$$

とするようにしてもよい（図 4 2 のステップ 5 0 4 6）。また、予測値を用いたか、補助情報により得られたピッチラグ

【数 1 2 6】

20

$$\hat{T}_p^{(M_{la})}$$

を用いたかに関するピッチラグ指示情報を適応符号帳算出部 1 1 2 3 に入力するようにしてもよい。

【0 2 6 0】

< 補助情報インデクスが補助情報を伝送しないことを表す値である場合 >

ピッチラグ予測部 1 9 2 は、ピッチラグ

【数 1 2 7】

$$\hat{T}_p^{(i)} (M_{la} \leq i < M)$$

30

の予測に当たって、過去に復号に用いたピッチラグ

【数 1 2 8】

$$\hat{T}_p^{(-j)} (1 \leq j < J)$$

を用いてピッチラグ

【数 1 2 9】

$$\hat{T}_p^{(i)} (0 \leq i < M)$$

40

を予測する（図 4 2 のステップ 5 0 4 8）。

【0 2 6 1】

さらに、ピッチラグ予測部 1 9 2 は、ピッチラグの予測値について信頼性が低い場合にのみ

【数 1 3 0】

$$\hat{T}_p^{(i)} = \hat{T}_p^{(-1)}$$

50

とし（図４２のステップ５０４９）、それ以外の場合には予測値を

【数１３１】

$$\hat{T}_p^{(i)}$$

とするようにしてもよい。また、予測値を用いたか、過去復号に用いたピッチラグ

【数１３２】

$$\hat{T}_p^{(-1)}$$

10

を用いたかに関するピッチラグ指示情報を適応符号帳算出部１１２３に入力する（図４２のステップ５０５０）。

【０２６２】

適応符号帳ゲイン予測部１９３、固定符号帳ゲイン予測部１９４は実施例１と同様である。

【０２６３】

雑音信号生成部１９５は、実施例１と同様である。

【０２６４】

音声合成部１２４は、音声パラメータ紛失処理部１２３が出力したパラメータから符号化対象フレームに対応する音声信号を合成する。

20

【０２６５】

音声合成部１２４（図１５参照）のＬＰ係数算出部１１２１は、実施例１と同様にしてＬＰ係数を得る（図１６のステップ１１３０１）。

【０２６６】

適応符号帳算出部１１２３の処理手順を図４３に示す。適応符号帳算出部１１２３は、実施例１と同様にして適応符号帳ベクトルを算出する。まず、ピッチラグ指示情報を参照し（図４３のステップ５０５１）、予測値の信頼性が低い場合（図４３のステップ５０５２：ＹＥＳ）は、以下の式を用いて適応符号帳ベクトルを算出する（図４３のステップ５０５５）。ここで、フィルタ係数を f_i とした。

【０２６７】

30

$$v(n) = f_{-1} v'(n-1) + f_0 v'(n) + f_1 v'(n+1)$$

なお設計方針により、 $v(n) = v'(n)$ としてもよい。

【０２６８】

ピッチラグ指示情報を参照し、予測値の信頼性が高い場合（図４３のステップ５０５２：ＮＯ）、適応符号帳算出部１１２３は、次の手順で適応符号帳ベクトルを算出する。

【０２６９】

まず、ピッチラグおよび適応符号帳１１２２を用いて初期適応符号帳ベクトルを算出する（図４３のステップ５０５３）。

【０２７０】

40

$$v(n) = f_{-1} v'(n-1) + f_0 v'(n) + f_1 v'(n+1)$$

設計方針によって、 $v(n) = v'(n)$ としてもよい。

【０２７１】

次に初期適応符号帳ベクトルに対して、声門パルス同期を適用する。声門パルス同期は非特許文献４の７．１１．２．５節におけるパルス位置が利用できない場合と同様の処理手順を用いる（図４３のステップ５０５４）。ただし、非特許文献４における $u(n)$ は本発明の実施形態の $v(n)$ に対応し、*extrapolated pitch*は本発明の実施形態の

【数１３３】

$$\hat{T}_p^{(M-1)}$$

50

に対応し、the last reliable pitch (T_c) は本発明の実施形態の

【数 134】

$$\hat{T}_p^{(-1)}$$

に対応する。

【0272】

励振ベクトル合成部 1124 は、実施例 1 と同様にして、励振信号ベクトルを出力する (図 16 のステップ 11306)。

10

【0273】

ポストフィルタ 1125 は、実施例 1 と同様にして、合成信号に後処理を加える。

【0274】

適応符号帳 1122 は、実施例 1 と同様にして、励振信号ベクトルにより状態を更新する (図 16 のステップ 11308)。

【0275】

合成フィルタ 1126 は、実施例 1 と同様にして、復号信号を合成する (図 16 のステップ 11309)。

【0276】

聴覚重み逆フィルタ 1127 は、実施例 1 と同様にして、聴覚重み逆フィルタを適用する。

20

【0277】

音声パラメータ紛失処理部 123 は、実施例 1 と同様にして、音声合成部 124 で用いた音声パラメータ (ISF パラメータ、ピッチラグ、適応符号帳ゲイン、固定符号帳ゲイン) をバッファに格納する (図 7 のステップ 145)。

【0278】

本発明の一側面に係る音声符号化装置は、音声信号を符号化する音声符号化装置であって、音声信号を符号化する音声符号化部と、先読み信号から補助情報を算出して符号化する補助情報符号化部と、を備える。

30

【0279】

補助情報は、先読み信号におけるピッチラグに関すること、又は、補助情報は、先読み信号におけるピッチゲインに関すること、又は、先読み信号におけるピッチラグおよびピッチゲインに関することとしてもよい。また、補助情報は、補助情報の利用可否に関する情報を含むこととしてもよい。

【0280】

補助情報符号化部は、先読み信号部分について補助情報を算出して符号化し、さらに隠蔽信号を生成し、入力音声信号と補助情報符号化部が出力する隠蔽信号の誤差信号を符号化する誤差信号符号化部と、入力音声信号そのものを符号化する主要符号化部と、をさらに備えることとしてもよい。

40

【0281】

また、本発明の一側面に係る音声復号装置は、音声符号を復号して音声信号を出力する音声復号装置であって、音声パケットの受信状態からパケットロスを検出する音声符号バッファと、音声パケット正常受信時に音声符号を復号する音声パラメータ復号部と、音声パケット正常受信時に補助情報符号を復号する補助情報復号部と、補助情報符号を復号して得られる補助情報を蓄積する補助情報蓄積部と、音声パケットロス検出時に音声パラメータを出力する音声パラメータ紛失処理部と、音声パラメータから復号音声合成する音声合成部と、を備える。

【0282】

補助情報は、先読み信号におけるピッチラグに関すること、又は、先読み信号における

50

ピッチゲインに關すること、又は、先読み信号におけるピッチラグおよびピッチゲインに關することとしてもよい。また、補助情報は、補助情報の利用可否に關する情報を含むこととしてもよい。

【0283】

補助情報復号部は、補助情報符号を復号して補助情報を出し、かつ補助情報を利用して先読み部分に關する隠蔽信号を出し、音声信号と隠蔽信号との誤差信号に關する符号を復号する誤差復号部と、音声信号に關する符号を復号する主要復号部と、補助情報復号部が出力した隠蔽信号を蓄積する隠蔽信号蓄積部と、をさらに備えることとしてもよい。

【0284】

音声パケット正常受信時には、隠蔽信号蓄積部から読み出した隠蔽信号と、誤差復号部が出力した復号誤差信号を加算することにより復号信号の一部を生成し、補助情報復号部が出力した隠蔽信号により隠蔽信号蓄積部を更新する、こととしてもよい。

【0285】

音声パケットロス検出時には、隠蔽信号蓄積部から読み出した隠蔽信号を、復号信号の一部あるいは全部とする、こととしてもよい。

【0286】

音声パケットロス検出時には、音声パラメータ紛失処理部により予測した音声パラメータを用いて復号信号を生成し、その一部を用いて隠蔽信号蓄積部を更新する、こととしてもよい。

【0287】

音声パラメータ紛失処理部は、音声パケットロス検出時に、補助情報蓄積部から読み出した補助情報を、音声パラメータの予測値の一部として利用する、こととしてもよい。

【0288】

音声合成部は、音声パケットロス検出時に、補助情報蓄積部から読み出した補助情報を用いて、音声パラメータの1つである適応符号帳ベクトルを修正する、こととしてもよい。

【0289】

また、本発明の一側面に係る音声符号化方法は、音声信号を符号化する音声符号化装置による音声符号化方法であって、音声信号を符号化する音声符号化ステップと、先読み信号から補助情報を算出して符号化する補助情報符号化ステップと、を含む。

【0290】

また、本発明の一側面に係る音声復号方法は、音声符号を復号して音声信号を出力する音声復号装置による音声復号方法であって、音声パケットの受信状態からパケットロスを検出する音声符号バッファステップと、音声パケット正常受信時に音声符号を復号する音声パラメータ復号ステップと、音声パケット正常受信時に補助情報符号を復号する補助情報復号ステップと、補助情報符号を復号して得られる補助情報を蓄積する補助情報蓄積ステップと、音声パケットロス検出時に音声パラメータを出力する音声パラメータ紛失処理ステップと、音声パラメータから復号音声合成する音声合成ステップと、を含む。

【0291】

また、本発明の一側面に係る音声符号化プログラムは、コンピュータを、音声信号を符号化する音声符号化部と、先読み信号から補助情報を算出して符号化する補助情報符号化部、として機能させる。

【0292】

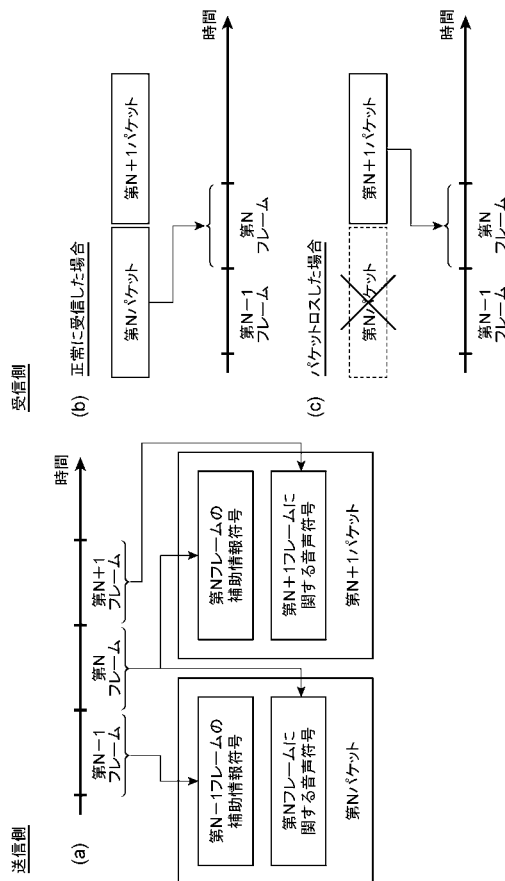
また、本発明の一側面に係る音声復号プログラムは、コンピュータを、音声パケットの受信状態からパケットロスを検出する音声符号バッファと、音声パケット正常受信時に音声符号を復号する音声パラメータ復号部と、音声パケット正常受信時に補助情報符号を復号する補助情報復号部と、補助情報符号を復号して得られる補助情報を蓄積する補助情報蓄積部と、音声パケットロス検出時に音声パラメータを出力する音声パラメータ紛失処理部と、音声パラメータから復号音声合成する音声合成部、として機能させる。

【符号の説明】

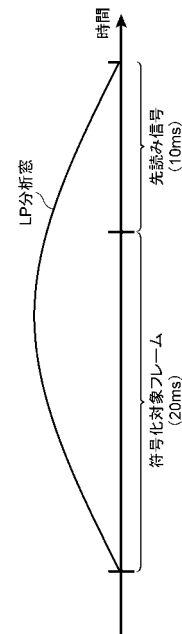
【 0 2 9 3 】

60、80...記憶媒体、61、81...プログラム格納領域、70...音声符号化プログラム、90...音声復号プログラム、111...音声符号化部、112...補助情報符号化部、121、231...音声符号バッファ、122、232...音声パラメータ復号部、123、233...音声パラメータ紛失処理部、124、234...音声合成部、125、235...補助情報復号部、126、236...補助情報蓄積部、151、511、1121...LP係数算出部、152、2012...ターゲット信号算出部、153、513、2013...ピッチラグ算出部、154、1123、514、2014、2313...適応符号帳算出部、155、1124、2314...励振ベクトル合成部、156、315、515、2019...適応符号帳バッファ、157、1126、2018、2316...合成フィルタ、158、516...ピッチラグ符号化部、191...ISF予測部、192...ピッチラグ予測部、193...適応符号帳ゲイン予測部、194...固定符号帳ゲイン予測部、195...雑音信号生成部、211...主要符号化部、212...補助情報符号化部、213、238...隠蔽信号蓄積部、214...誤差信号符号化部、237...誤差信号復号部、311...LP係数算出部、312...ピッチラグ予測部、313...ピッチラグ選定部、314...ピッチラグ符号化部、512...残差信号算出部、700...音声符号化モジュール、701...補助情報符号化モジュール、900...音声パラメータ復号モジュール、901...音声パラメータ紛失処理モジュール、902...音声合成モジュール、903...補助情報復号モジュール、1128...補助情報出力判定部、1122、2312...適応符号帳、1125...ポストフィルタ、1127...聴覚重み逆フィルタ、2011...ISF符号化部、2015...固定符号帳算出部、2016...ゲイン算出部、2017...励振ベクトル算出部、2211...ISF復号部、2212...ピッチラグ復号部、2213...ゲイン復号部、2214...固定符号帳復号部、2318...先読み励振ベクトル合成部。

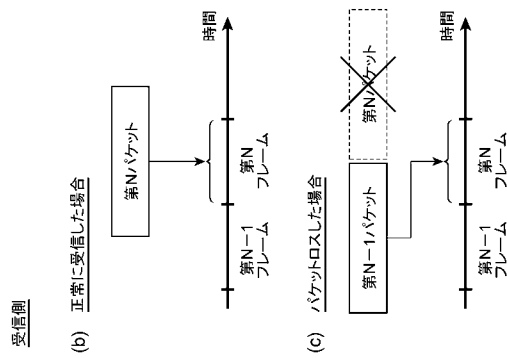
【 図 1 】



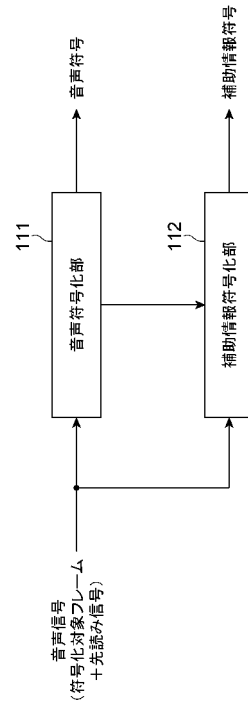
【 図 2 】



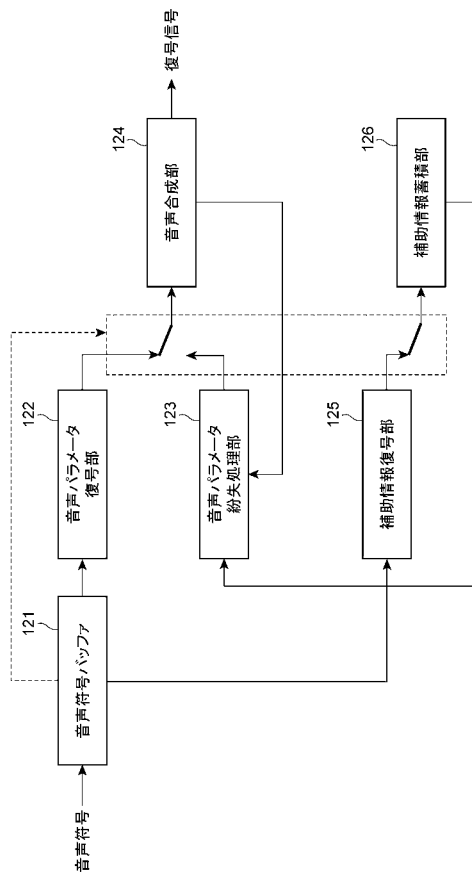
【図 3】



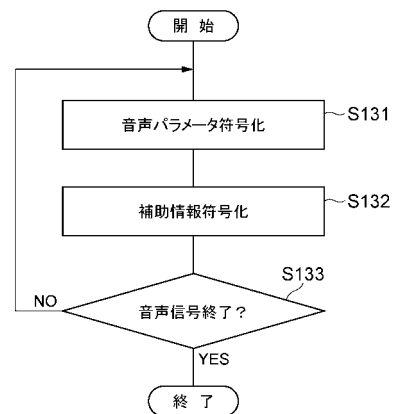
【図 4】



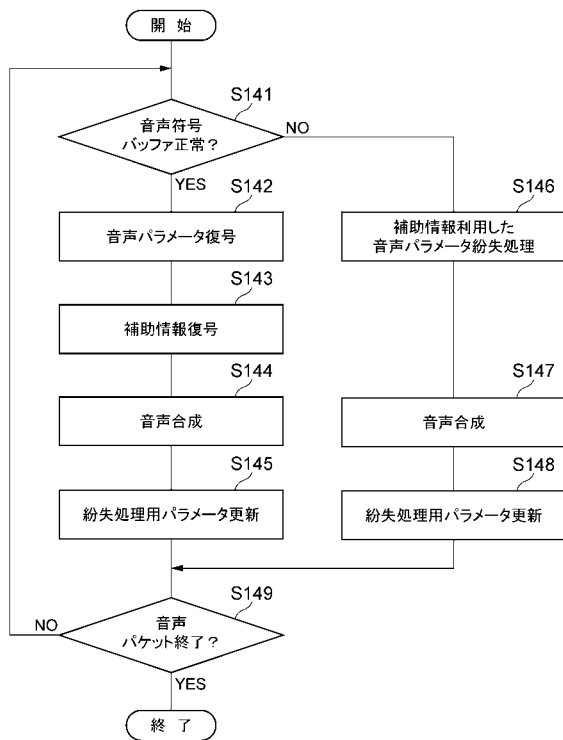
【図 5】



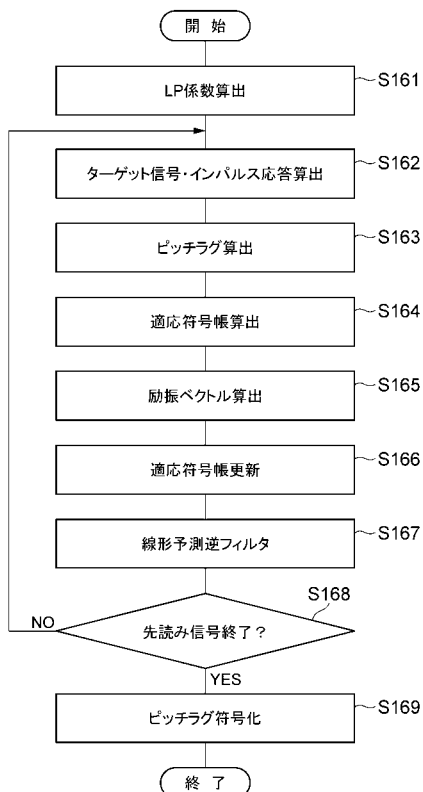
【図 6】



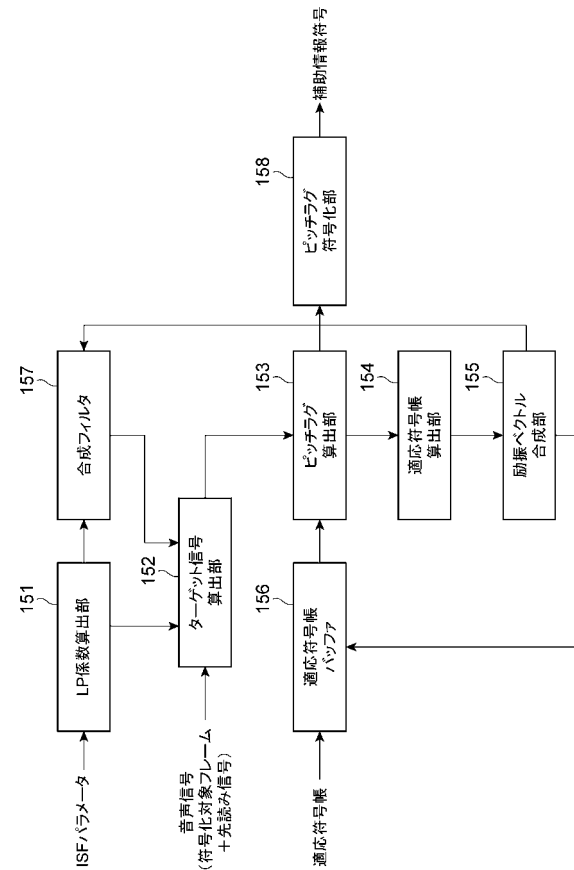
【図 7】



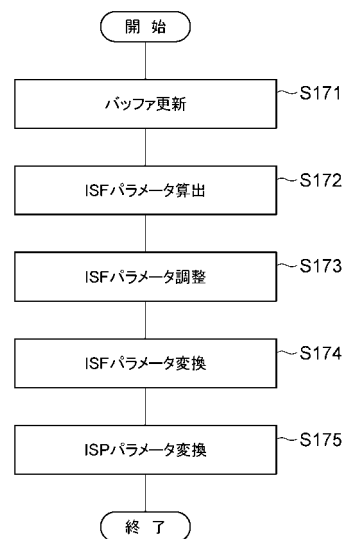
【図 9】



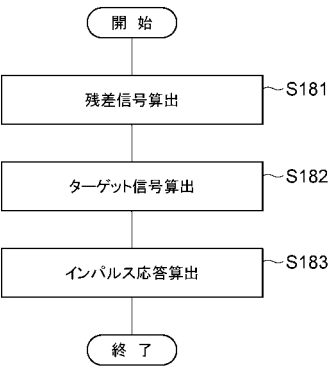
【図 8】



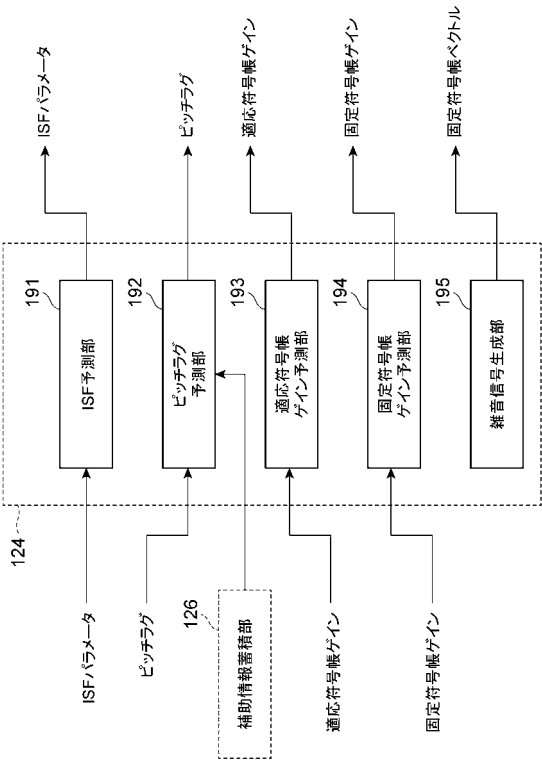
【図 10】



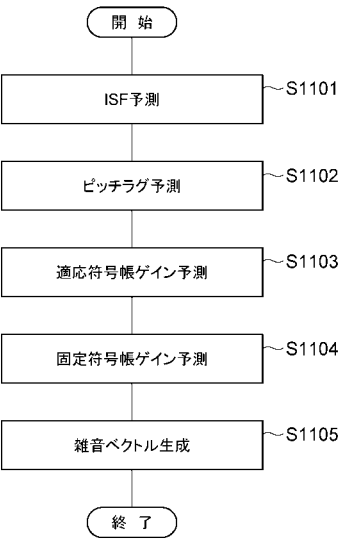
【図 1 1】



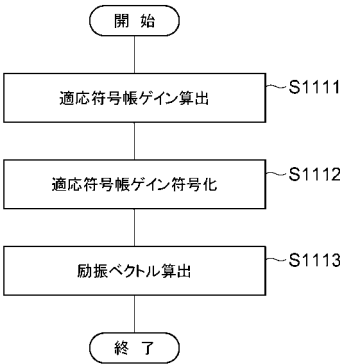
【図 1 2】



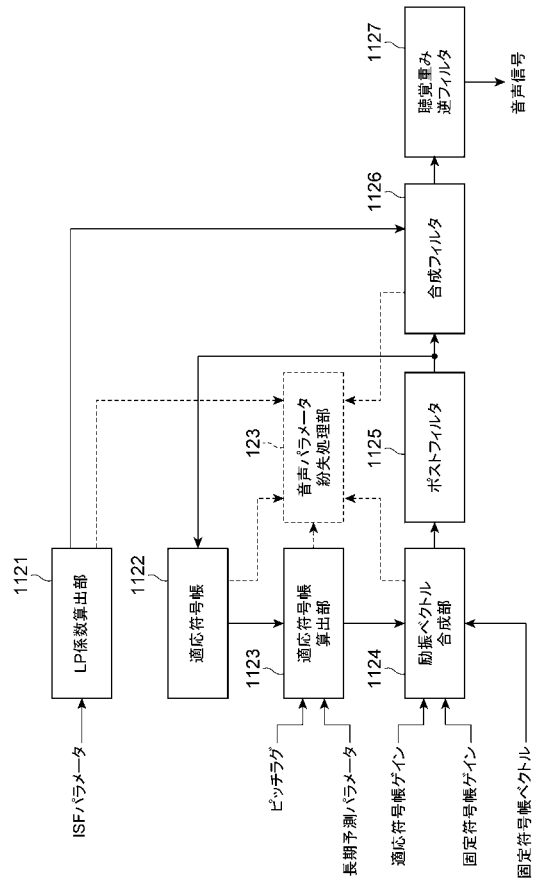
【図 1 3】



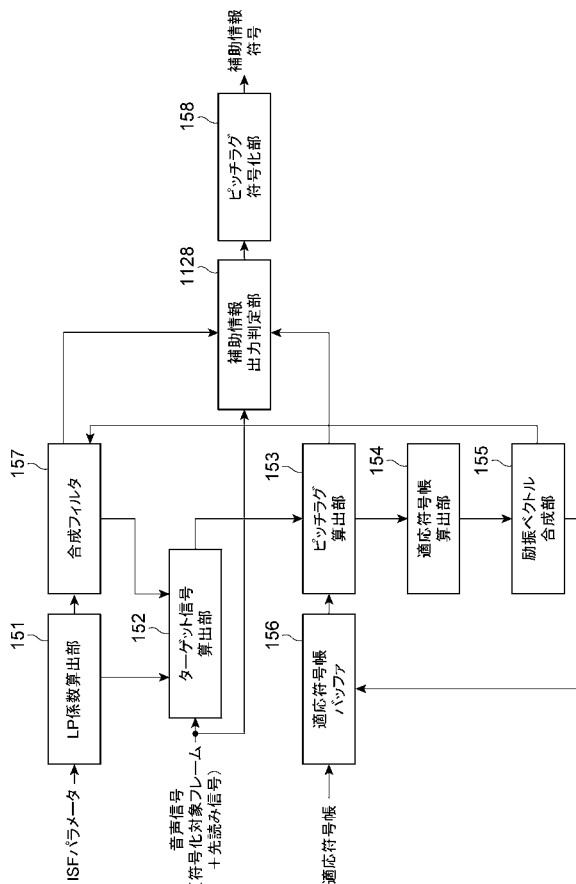
【図 1 4】



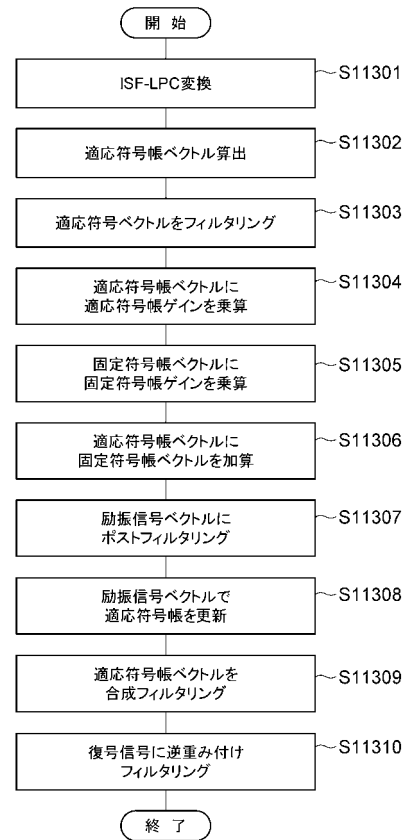
【図 15】



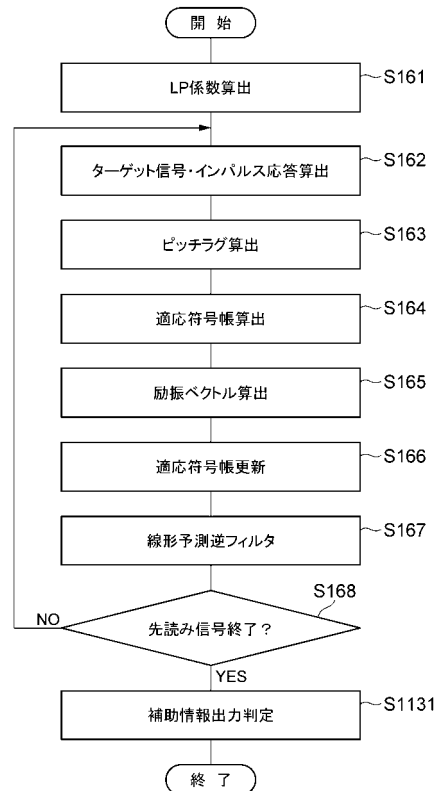
【図 17】



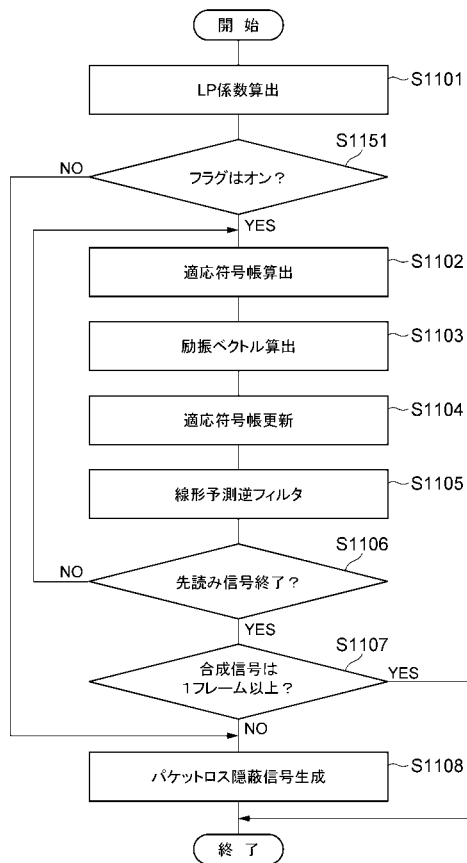
【図 16】



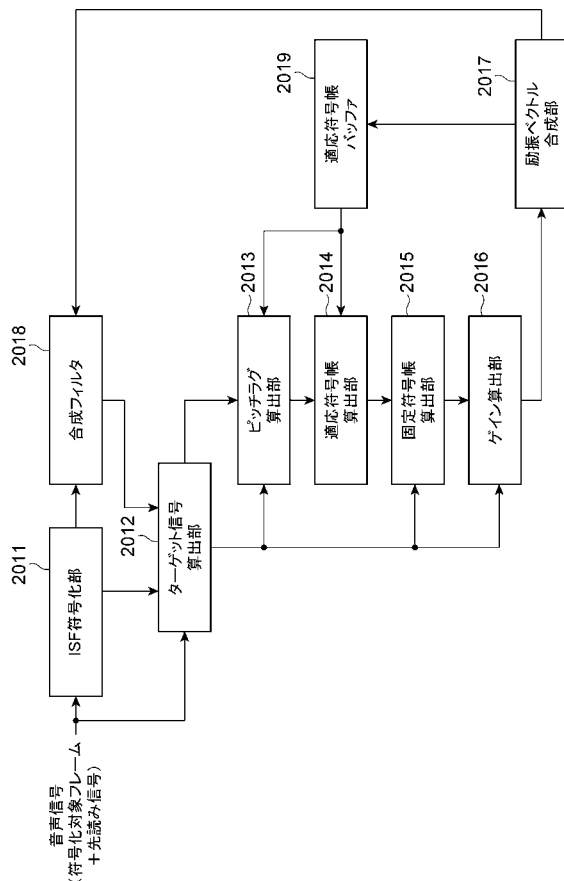
【図 18】



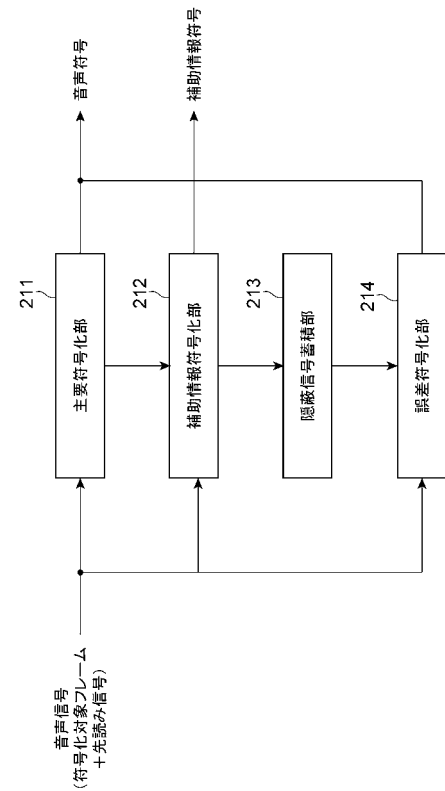
【図 19】



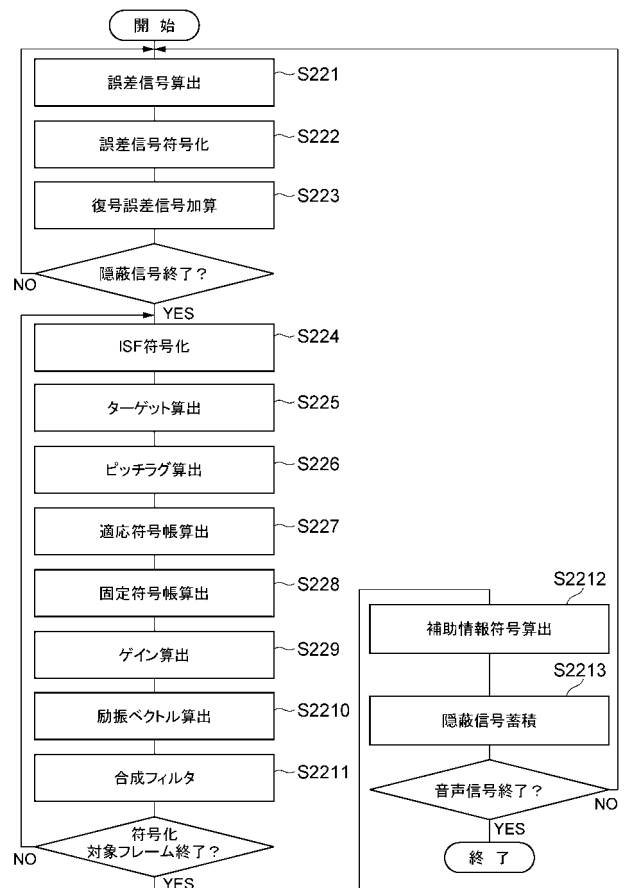
【図 21】



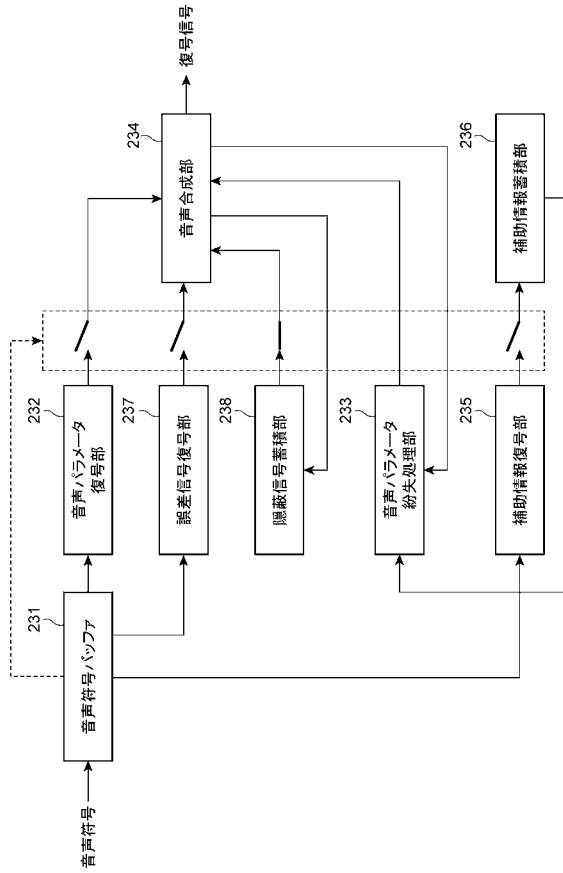
【図 20】



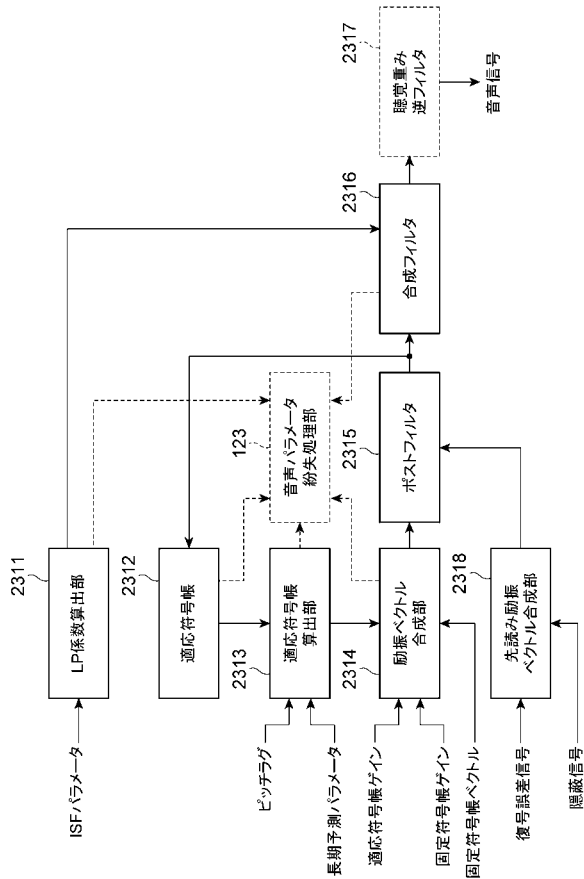
【図 22】



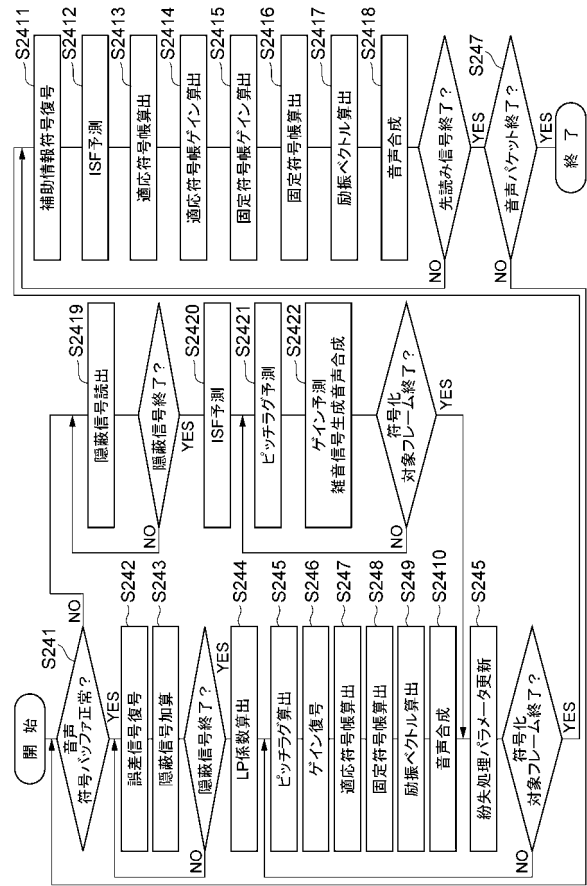
【 図 2 3 】



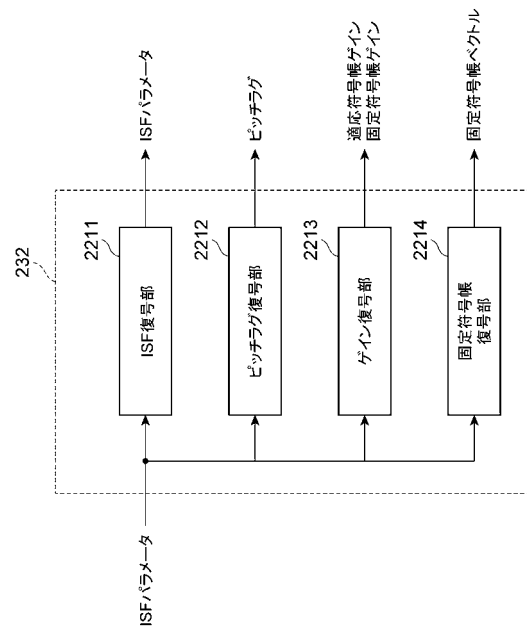
【 図 2 5 】



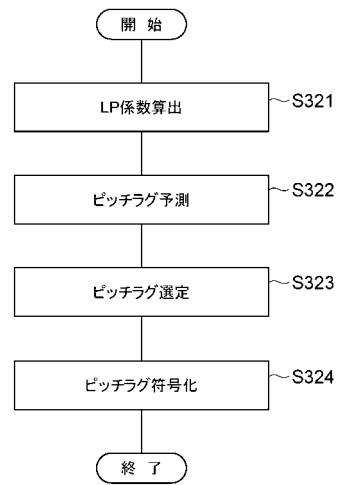
【 図 2 4 】



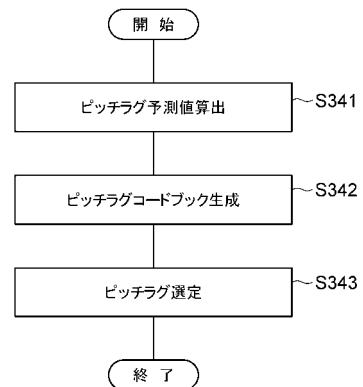
【 図 2 6 】



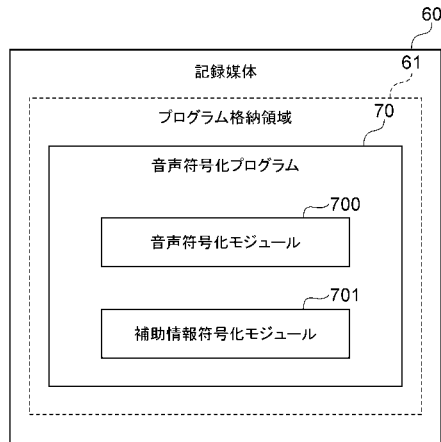
【 圖 2 8 】



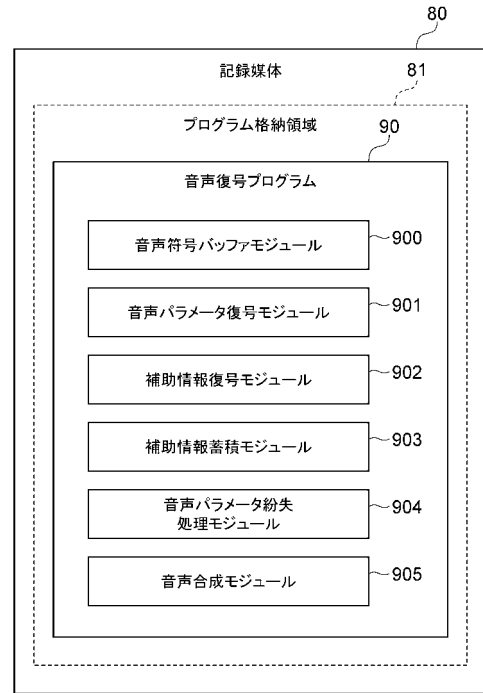
【 ㄨ 3 0 】



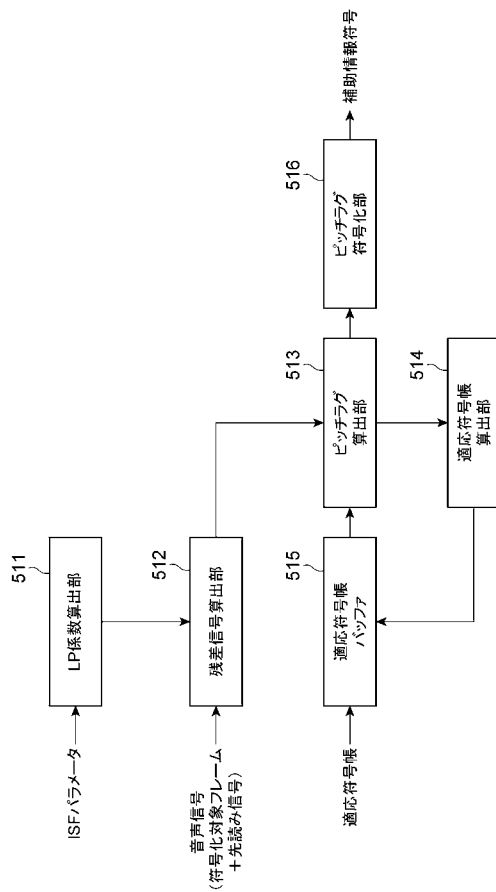
【図 3 1】



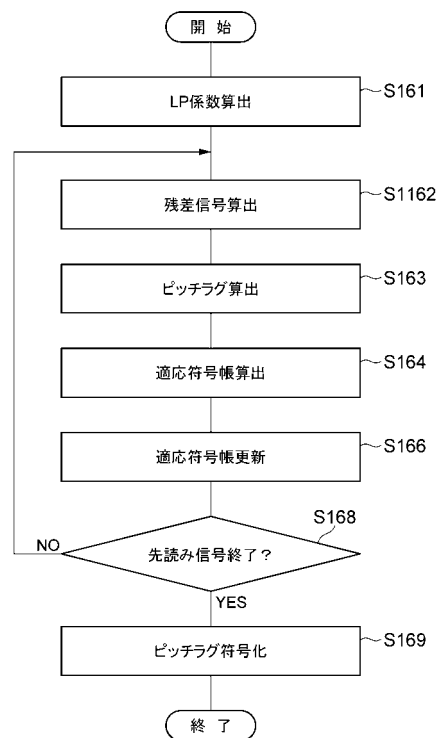
【図 3 2】



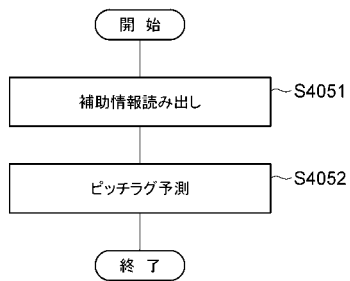
【図 3 3】



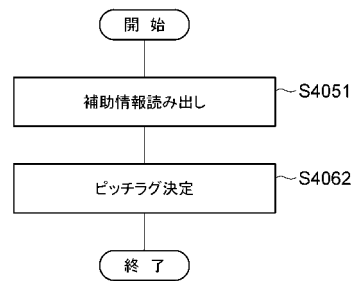
【図 3 4】



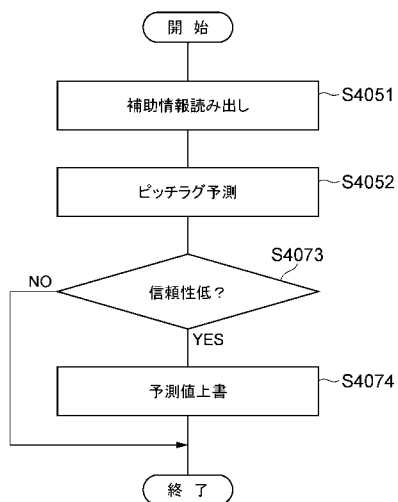
【図 35】



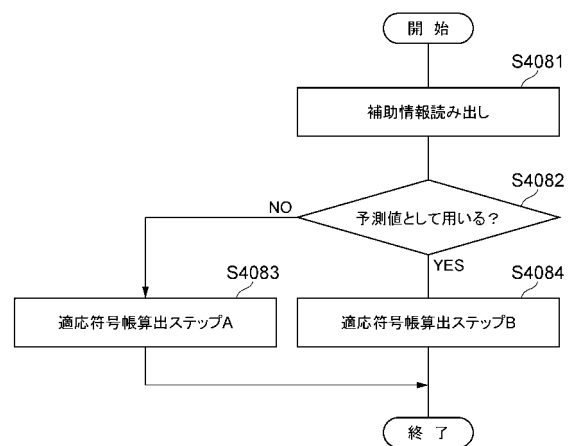
【図 36】



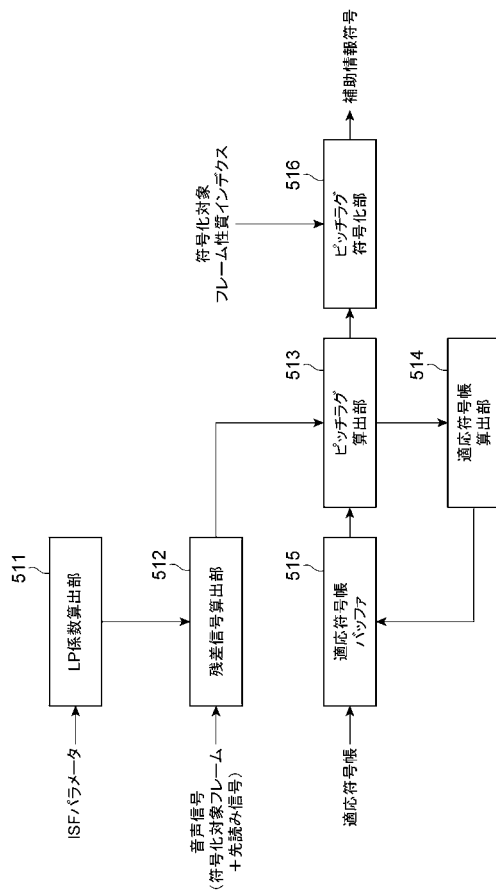
【図 37】



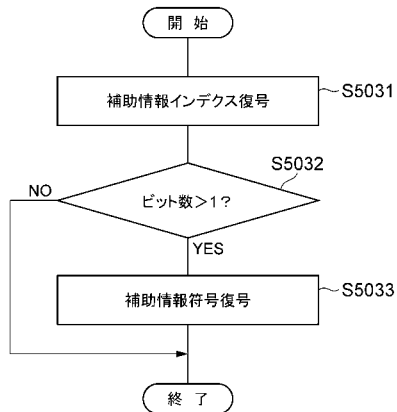
【図 38】



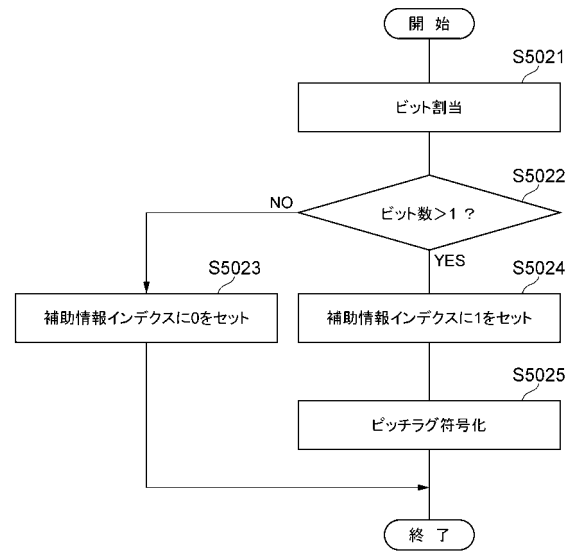
【図 39】



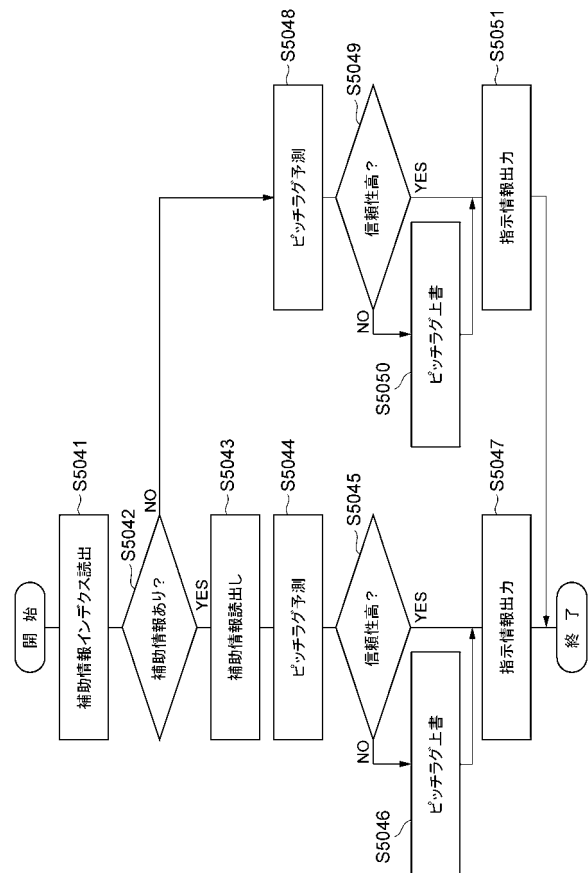
【図 41】



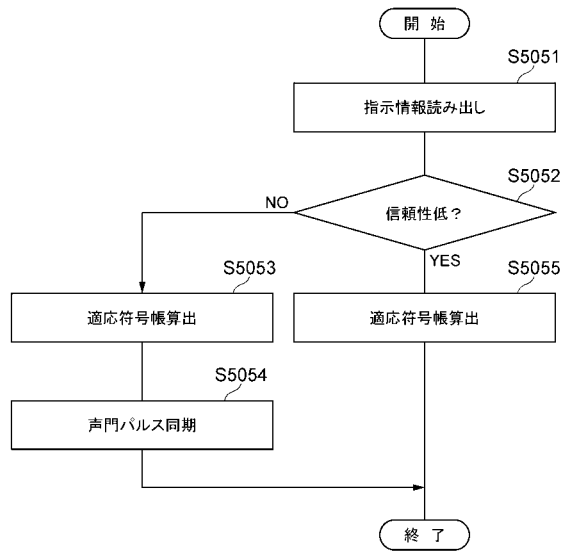
【図 40】



【図 42】



【図 4 3】



フロントページの続き

- (72)発明者 菊入 圭
東京都千代田区永田町二丁目 1 1 番 1 号 株式会社ＮＴＴドコモ内
- (72)発明者 山口 貴史
東京都千代田区永田町二丁目 1 1 番 1 号 株式会社ＮＴＴドコモ内