

## 明 細 書

### 発明の名称：

音声符号化装置、音声復号装置、音声符号化方法、音声復号方法、音声符号化プログラム及び音声復号プログラム

### 技術分野

[0001] 本発明は、音声符号化装置、音声復号装置、音声符号化方法、音声復号方法、音声符号化プログラム及び音声復号プログラムに関する。

### 背景技術

[0002] 聴覚心理を利用して人間の知覚に不必要な情報を取り除くことにより信号のデータ量を数十分の一に圧縮する音声音響符号化技術は、信号の伝送・蓄積において極めて重要な技術である。広く利用されている知覚的オーディオ符号化技術の例として、“ISO/IEC MPEG”で標準化された“MPEG4 AAC”などを挙げることができる。

[0003] 音声符号化の性能をさらに向上させ、低いビットレートで高い音声品質を得る方法として、音声の低周波成分を用いて高周波成分を生成する帯域拡張技術が近年広く用いられるようになった。帯域拡張技術の代表的な例は“MPEG4 AAC”で利用されるSBR (Spectral Band Replication) 技術である。SBRでは、QMF (Quadrature Mirror Filter) フィルタバンクによって周波数領域に変換された信号に対し、低周波帯域から高周波帯域へのスペクトル係数の複写を行うことにより高周波成分を生成した後、複写された係数のスペクトル包絡とトーンリティを調整することによって高周波成分の調整を行う。帯域拡張技術を利用した音声符号化方式は、信号の高周波成分を少量の補助情報のみを用いて再生することができるため、音声符号化の低ビットレート化のために有効である。

[0004] SBRに代表される周波数領域での帯域拡張技術は、周波数領域で表現されたスペクトル係数に対してスペクトル包絡とトーンリティの調整を、スペクトル係数に対するゲインの調整、時間方向の線形予測逆フィルタ処理、ノ

イズの重畳によって行う。この調整処理により、スピーチ信号や拍手、カスタネットのような時間エンベロープの変化の大きい信号を符号化した際には復号信号においてプリエコー又はポストエコーと呼ばれる残響状の雑音が知覚される場合がある。この問題は、調整処理の過程で高周波成分の時間エンベロープが変形し、多くの場合は調整前より平坦な形状になることに起因する。調整処理により平坦になった高周波成分の時間エンベロープは符号前の原信号における高周波成分の時間エンベロープと一致せず、プリエコー・ポストエコーの原因となる。

- [0005] 同様のプリエコー・ポストエコーの問題は、“MPEG Surround” およびパラメトリックステレオに代表される、パラメトリック処理を用いたマルチチャネル音響符号化においても発生する。マルチチャネル音響符号化における復号器は復号信号に残響フィルタによる無相関化処理を施す手段を含むが、無相関化処理の過程において信号の時間エンベロープが変形し、プリエコー・ポストエコーと同様の再生信号の劣化が生じる。この課題に対する解決法として、TES (Temporal Envelope Shaping) 技術が存在する (特許文献 1) 。TES 技術では、QMF 領域で表現された無相関化処理前の信号に対し周波数方向に線形予測分析を行い、線形予測係数を得た後、得られた線形予測係数を用いて無相関化処理後の信号に対し周波数方向に線形予測合成フィルタ処理を行う。この処理により、TES 技術は無相関化処理前の信号の持つ時間エンベロープを抽出し、それに合わせて無相関化処理後の信号の時間エンベロープを調整する。無相関化処理前の信号は歪の少ない時間エンベロープを持つため、以上の処理により、無相関化処理後の信号の時間エンベロープを歪の少ない形状に調整し、プリエコー・ポストエコーの改善された再生信号を得ることができる。

## 先行技術文献

## 特許文献

- [0006] 特許文献1：米国特許出願公開第 2 0 0 6 / 0 2 3 9 4 7 3 号明細書

## 発明の概要

## 発明が解決しようとする課題

- [0007] 以上に示したＴＥＳ技術は、無相関化処理前の信号が歪の少ない時間エンベロープを持つことを利用したものである。しかし、ＳＢＲ復号器では信号の高周波成分を低周波成分からの信号複写によって複製するため、高周波成分に関する歪の少ない時間エンベロープを得ることができない。この問題に対する解決法の一つとして、ＳＢＲ符号器において入力信号の高周波成分を分析し、分析の結果得られた線形予測係数を量子化し、ビットストリームに多重化して伝送する方法が考えられる。これにより、ＳＢＲ復号器において高周波成分の時間エンベロープに関する歪の少ない情報を含む線形予測係数を得ることができる。しかし、この場合、量子化された線形予測係数の伝送に多くの情報量が必要となり、符号化ビットストリーム全体のビットレートが著しく増大してしまうという問題を伴う。そこで、本発明の目的は、ＳＢＲに代表される周波数領域での帯域拡張技術において、ビットレートを著しく増大させることなく、発生するプリエコー・ポストエコーを軽減し復号信号の主観的品質を向上させることである。

## 課題を解決するための手段

- [0008] 本発明の音声符号化装置は、音声信号を符号化する音声符号化装置であって、前記音声信号の低周波成分を符号化するコア符号化手段と、前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出手段と、少なくとも、前記コア符号化手段によって符号化された前記低周波成分と、前記時間エンベロープ補助情報算出手段によって算出された前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化手段と、を備える、ことを特徴とする。
- [0009] 本発明の音声符号化装置では、前記時間エンベロープ補助情報は、所定の解析区間内において前記音声信号の高周波成分における時間エンベロープの変化の急峻さを示すパラメータを表すのが好ましい。

- [0010] 本発明の音声符号化装置では、前記音声信号を周波数領域に変換する周波数変換手段を更に備え、前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って取得された高周波線形予測係数に基づいて、前記時間エンベロープ補助情報を算出するのが好ましい。
- [0011] 本発明の音声符号化装置では、前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の低周波側係数に対し周波数方向に線形予測分析を行って低周波線形予測係数を取得し、該低周波線形予測係数と前記高周波線形予測係数とに基づいて前記時間エンベロープ補助情報を算出するのが好ましい。
- [0012] 本発明の音声符号化装置では、前記時間エンベロープ補助情報算出手段は、前記低周波線形予測係数及び前記高周波線形予測係数のそれぞれから予測ゲインを取得し、当該二つの予測ゲインの大小に基づいて前記時間エンベロープ補助情報を算出するのが好ましい。
- [0013] 本発明の音声符号化装置では、前記時間エンベロープ補助情報算出手段は、前記音声信号から高周波成分を分離し、時間領域で表現された時間エンベロープ情報を当該高周波成分から取得し、当該時間エンベロープ情報の時間的変化の大きさに基づいて前記時間エンベロープ補助情報を算出するのが好ましい。
- [0014] 本発明の音声符号化装置では、前記時間エンベロープ補助情報は、前記音声信号の低周波成分に対し周波数方向への線形予測分析を行って得られる低周波線形予測係数を用いて高周波線形予測係数を取得するための差分情報を含むのが好ましい。
- [0015] 本発明の音声符号化装置では、前記音声信号を周波数領域に変換する周波数変換手段を更に備え、前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の低周波成分及び高周波側係数のそれぞれに対し周波数方向に線形予測分析を行って低周波線形予測係数と高周波線形予測係数とを取得し、当該低周波線形予測係数及

び高周波線形予測係数の差分を取得することによって前記差分情報を取得するのが好ましい。

[0016] 本発明の音声符号化装置では、前記差分情報は、L S P (Linear Spectrum Pair)、I S P (Immittance Spectrum Pair)、L S F (Linear Spectrum Frequency)、I S F (Immittance Spectrum Frequency)、P A R C O R 係数のいずれかの領域における線形予測係数の差分を表すのが好ましい。

[0017] 本発明の音声符号化装置は、音声信号を符号化する音声符号化装置であって、前記音声信号の低周波成分を符号化するコア符号化手段と、前記音声信号を周波数領域に変換する周波数変換手段と、前記周波数変換手段によって周波数領域に変換された前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析手段と、前記線形予測分析手段によって取得された前記高周波線形予測係数を時間方向に間引く予測係数間引き手段と、前記予測係数間引き手段によって間引きされた後の前記高周波線形予測係数を量子化する予測係数量子化手段と、少なくとも前記コア符号化手段による符号化後の前記低周波成分と前記予測係数量子化手段による量子化後の前記高周波線形予測係数とが多重化されたビットストリームを生成するビットストリーム多重化手段と、を備える、ことを特徴とする。

[0018] 本発明の音声復号装置は、符号化された音声信号を復号する音声復号装置であって、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと時間エンベロープ補助情報とに分離するビットストリーム分離手段と、前記ビットストリーム分離手段によって分離された前記符号化ビットストリームを復号して低周波成分を得るコア復号手段と、前記コア復号手段によって得られた前記低周波成分を周波数領域に変換する周波数変換手段と、前記周波数変換手段によって周波数領域に変換された前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成手段と、前記周波数変換手段によって周波数領域に変換された前記低周波成分を分析して時間エンベロープ情報を取得する低周

波時間エンベロープ分析手段と、前記低周波時間エンベロープ分析手段によって取得された前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整手段と、前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を用いて、前記高周波生成手段によって生成された前記高周波成分の時間エンベロープを変形する時間エンベロープ変形手段と、を備えることを特徴とする。

[0019] 本発明の音声復号装置では、前記高周波成分を調整する高周波調整手段を更に備え、前記周波数変換手段は、実数又は複素数の係数を持つ64分割QMFフィルタバンクであり、前記周波数変換手段、前記高周波生成手段、前記高周波調整手段は“ISO/IEC 14496-3”に規定される“MPEG4 AAC”におけるSBR復号器（SBR：Spectral Band Replication）に準拠した動作をするのが好ましい。

[0020] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分に周波数方向の線形予測分析を行って低周波線形予測係数を取得し、前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整し、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し前記時間エンベロープ調整手段によって調整された線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形するのが好ましい。

[0021] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分の時間スロットごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に前記調整後の時間エンベロープ情報を重畳することにより高周波成分の時間

エンベロープを変形するのが好ましい。

[0022] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分のQMFサブバンドサンプルごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に前記調整後の時間エンベロープ情報を乗算することにより高周波成分の時間エンベロープを変形するのが好ましい。

[0023] 本発明の音声復号装置では、前記時間エンベロープ補助情報は、線形予測係数の強度の調整に用いるためのフィルタ強度パラメータを表すのが好ましい。

[0024] 本発明の音声復号装置では、前記時間エンベロープ補助情報は、前記時間エンベロープ情報の時間変化の大きさを示すパラメータを表すのが好ましい。

[0025] 本発明の音声復号装置では、前記時間エンベロープ補助情報は、前記低周波線形予測係数に対する線形予測係数の差分情報を含むのが好ましい。

[0026] 本発明の音声復号装置では、前記差分情報は、LSP (Linear Spectrum Pair)、ISP (Immittance Spectrum Pair)、LSF (Linear Spectrum Frequency)、ISF (Immittance Spectrum Frequency)、PARCOR係数のいずれかの領域における線形予測係数の差分を表すのが好ましい。

[0027] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分に対し周波数方向の線形予測分析を行って前記低周波線形予測係数を取得するとともに、当該周波数領域の前記低周波成分の時間スロットごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整するとともに前記時間エンベロープ補助情報を用いて前記時間

エンベロープ情報を調整し、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に対し前記時間エンベロープ調整手段によって調整された線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形するとともに当該周波数領域の前記高周波成分に前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を重量することにより前記高周波成分の時間エンベロープを変形するのが好ましい。

[0028] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分に対し周波数方向の線形予測分析を行って前記低周波線形予測係数を取得するとともに、当該周波数領域の前記低周波成分のQMFサブバンドごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整するとともに前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に対し前記時間エンベロープ調整手段による調整後の線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形するとともに当該周波数領域の前記高周波成分に前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を乗算することにより前記高周波成分の時間エンベロープを変形するのが好ましい。

[0029] 本発明の音声復号装置では、前記時間エンベロープ補助情報は、線形予測係数のフィルタ強度と、前記時間エンベロープ情報の時間変化の大きさとの両方を示すパラメータを表すのが好ましい。

[0030] 本発明の音声復号装置は、符号化された音声信号を復号する音声復号装置であって、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと線形予測係数とに分離するビットストリーム分離手段と、前記線形予測係数を時間方向に補間又は補外する線形予測係数補

間・補外手段と、前記線形予測係数補間・補外手段によって補間又は補外された線形予測係数を用いて周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する時間エンベロープ変形手段と、を備える、ことを特徴とする。

[0031] 本発明の音声符号化方法は、音声信号を符号化する音声符号化装置を用いた音声符号化方法であって、前記音声符号化装置が、前記音声信号の低周波成分を符号化するコア符号化ステップと、前記音声符号化装置が、前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出ステップと、前記音声符号化装置が、少なくとも、前記コア符号化ステップにおいて符号化した前記低周波成分と、前記時間エンベロープ補助情報算出ステップにおいて算出した前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化ステップと、を備える、ことを特徴とする。

[0032] 本発明の音声符号化方法は、音声信号を符号化する音声符号化装置を用いた音声符号化方法であって、前記音声符号化装置が、前記音声信号の低周波成分を符号化するコア符号化ステップと、前記音声符号化装置が、前記音声信号を周波数領域に変換する周波数変換ステップと、前記音声符号化装置が、前記周波数変換ステップにおいて周波数領域に変換した前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析ステップと、前記音声符号化装置が、前記線形予測分析手段ステップにおいて取得した前記高周波線形予測係数を時間方向に間引く予測係数間引きステップと、前記音声符号化装置が、前記予測係数間引き手段ステップにおける間引き後の前記高周波線形予測係数を量子化する予測係数量子化ステップと、前記音声符号化装置が、少なくとも前記コア符号化ステップにおける符号化後の前記低周波成分と前記予測係数量子化ステップにおける量子化後の前記高周波線形予測係数とが多重化されたビットストリームを生成するビットストリーム多重化ステップと、を備える、ことを特徴

とする。

[0033] 本発明の音声復号方法は、符号化された音声信号を復号する音声復号装置を用いた音声復号方法であって、前記音声復号装置が、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと時間エンベロープ補助情報とに分離するビットストリーム分離ステップと、前記音声復号装置が、前記ビットストリーム分離ステップにおいて分離した前記符号化ビットストリームを復号して低周波成分を得るコア復号ステップと、前記音声復号装置が、前記コア復号ステップにおいて得た前記低周波成分を周波数領域に変換する周波数変換ステップと、前記音声復号装置が、前記周波数変換ステップにおいて周波数領域に変換した前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成ステップと、前記音声復号装置が、前記周波数変換ステップにおいて周波数領域に変換した前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析ステップと、前記音声復号装置が、前記低周波時間エンベロープ分析ステップにおいて取得した前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整ステップと、前記音声復号装置が、前記時間エンベロープ調整ステップにおける調整後の前記時間エンベロープ情報を用いて、前記高周波生成ステップにおいて生成した前記高周波成分の時間エンベロープを変形する時間エンベロープ変形ステップと、を備えることを特徴とする。

[0034] 本発明の音声復号方法は、符号化された音声信号を復号する音声復号装置を用いた音声復号方法であって、前記音声復号装置が、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと線形予測係数とに分離するビットストリーム分離ステップと、前記音声復号装置が、前記線形予測係数を時間方向に補間又は補外する線形予測係数補間・補外ステップと、前記音声復号装置が、前記線形予測係数補間・補外ステップにおいて補間又は補外した前記線形予測係数を用いて、周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の

時間エンベロープを変形する時間エンベロープ変形ステップと、を備える、ことを特徴とする。

[0035] 本発明の音声符号化プログラムは、音声信号を符号化するために、コンピュータ装置を、前記音声信号の低周波成分を符号化するコア符号化手段、前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出手段、及び、少なくとも、前記コア符号化手段によって符号化された前記低周波成分と、前記時間エンベロープ補助情報算出手段によって算出された前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化手段、として機能させることを特徴とする。

[0036] 本発明の音声符号化プログラムは、音声信号を符号化するために、コンピュータ装置を、前記音声信号の低周波成分を符号化するコア符号化手段、前記音声信号を周波数領域に変換する周波数変換手段、前記周波数変換手段によって周波数領域に変換された前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析手段、前記線形予測分析手段によって取得された前記高周波線形予測係数を時間方向に間引く予測係数間引き手段、前記予測係数間引き手段によって間引きされた後の前記高周波線形予測係数を量子化する予測係数量子化手段、及び、少なくとも前記コア符号化手段による符号化後の前記低周波成分と前記予測係数量子化手段による量子化後の前記高周波線形予測係数とが多重化されたビットストリームを生成するビットストリーム多重化手段、として機能させることを特徴とする。

[0037] 本発明の音声復号プログラムは、符号化された音声信号を復号するために、コンピュータ装置を、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと時間エンベロープ補助情報とに分離するビットストリーム分離手段、前記ビットストリーム分離手段によって分離された前記符号化ビットストリームを復号して低周波成分を得るコア復

号手段、前記コア復号手段によって得られた前記低周波成分を周波数領域に変換する周波数変換手段、前記周波数変換手段によって周波数領域に変換された前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成手段、前記周波数変換手段によって周波数領域に変換された前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析手段、前記低周波時間エンベロープ分析手段によって取得された前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整手段、及び、前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を用いて、前記高周波生成手段によって生成された前記高周波成分の時間エンベロープを変形する時間エンベロープ変形手段、として機能させることを特徴とする。

[0038] 本発明の音声復号プログラムは、符号化された音声信号を復号するために、コンピュータ装置を、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと線形予測係数とに分離するビットストリーム分離手段、前記線形予測係数を時間方向に補間又は補外する線形予測係数補間・補外手段、及び、前記線形予測係数補間・補外手段によって補間又は補外された線形予測係数を用いて周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する時間エンベロープ変形手段、として機能させることを特徴とする。

[0039] 本発明の音声復号装置では、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し周波数方向の線形予測フィルタ処理を行った後、前記線形予測フィルタ処理の結果得られた高周波成分の電力を前記線形予測フィルタ処理前と等しい値に調整するのが好ましい。

[0040] 本発明の音声復号装置では、前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し周波数方向の線形予測フィルタ処理を行った後、前記線形予測フィルタ処理の結果得

られた高周波成分の任意の周波数範囲内の電力を前記線形予測フィルタ処理前と等しい値に調整するのが好ましい。

[0041] 本発明の音声復号装置では、前記時間エンベロープ補助情報は、前記調整後の前記時間エンベロープ情報における最小値と平均値の比率であるのが好ましい。

[0042] 本発明の音声復号装置では、前記時間エンベロープ変形手段は、前記周波数領域の高周波成分のSBRエンベロープ時間セグメント内での電力が時間エンベロープの変形の前後で等しくなるように前記調整後の時間エンベロープの利得を制御した後に、前記周波数領域の高周波成分に前記利得制御された時間エンベロープを乗算することにより高周波成分の時間エンベロープを変形するのが好ましい。

[0043] 本発明の音声復号装置では、前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分のQMFサブバンドサンプルごとの電力を取得し、さらにSBRエンベロープ時間セグメント内での平均電力を用いて前記QMFサブバンドサンプルごとの電力を正規化することによって、各QMFサブバンドサンプルへ乗算されるべきゲイン係数として表現された時間エンベロープ情報を取得するのが好ましい。

[0044] 本発明の音声復号装置は、符号化された音声信号を復号する音声復号装置であって、前記符号化された音声信号を含む外部からのビットストリームを復号して低周波成分を得るコア復号手段と、前記コア復号手段によって得られた前記低周波成分を周波数領域に変換する周波数変換手段と、前記周波数変換手段によって周波数領域に変換された前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成手段と、前記周波数変換手段によって周波数領域に変換された前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析手段と、前記ビットストリームを分析して時間エンベロープ補助情報を生成する時間エンベロープ補助情報生成部と、前記低周波時間エンベロープ分析手段に

よって取得された前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整手段と、前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を用いて、前記高周波生成手段によって生成された前記高周波成分の時間エンベロープを変形する時間エンベロープ変形手段と、を備える、ことを特徴とする。

[0045] 本発明の音声復号装置では、前記高周波調整手段に相当する、一次高周波調整手段と、二次高周波調整手段とを具備し、前記一次高周波調整手段は、前記高周波調整手段に相当する処理の一部を含む処理を実行し、前記時間エンベロープ変形手段は、前記一次高周波調整手段の出力信号に対し時間エンベロープの変形を行い、前記二次高周波調整手段は、前記時間エンベロープ変形手段の出力信号に対して、前記高周波調整手段に相当する処理のうち前記一次高周波調整手段で実行されない処理を実行するのが好ましく、前記二次高周波調整手段は、SBRの復号過程における正弦波の付加処理であるのが好ましい。

### 発明の効果

[0046] 本発明によれば、SBRに代表される周波数領域での帯域拡張技術において、ビットレートを著しく増大させることなく、発生するプリエコー・ポストエコーを軽減し復号信号の主観的品質を向上できる。

### 図面の簡単な説明

[0047] [図1]第1の実施形態に係る音声符号化装置の構成を示す図である。  
[図2]第1の実施形態に係る音声符号化装置の動作を説明するためのフローチャートである。  
[図3]第1の実施形態に係る音声復号装置の構成を示す図である。  
[図4]第1の実施形態に係る音声復号装置の動作を説明するためのフローチャートである。  
[図5]第1の実施形態の変形例1に係る音声符号化装置の構成を示す図である。  
[図6]第2の実施形態に係る音声符号化装置の構成を示す図である。

[図7]第2の実施形態に係る音声符号化装置の動作を説明するためのフローチャートである。

[図8]第2の実施形態に係る音声復号装置の構成を示す図である。

[図9]第2の実施形態に係る音声復号装置の動作を説明するためのフローチャートである。

[図10]第3の実施形態に係る音声符号化装置の構成を示す図である。

[図11]第3の実施形態に係る音声符号化装置の動作を説明するためのフローチャートである。

[図12]第3の実施形態に係る音声復号装置の構成を示す図である。

[図13]第3の実施形態に係る音声復号装置の動作を説明するためのフローチャートである。

[図14]第4の実施形態に係る音声復号装置の構成を示す図である。

[図15]第4の実施形態の変形例に係る音声復号装置の構成を示す図である。

[図16]第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図17]第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図18]第1の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図19]第1の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図20]第1の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図21]第1の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図22]第2の実施形態の変形例に係る音声復号装置の構成を示す図である。

[図23]第2の実施形態の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図24] 第2の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図25] 第2の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図26] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図27] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図28] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図29] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図30] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図31] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図32] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図33] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図34] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図35] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図36] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図37] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図38] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図39] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図40] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図41] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図42] 第4の実施形態の他の変形例に係る音声復号装置の構成を示す図である。

[図43] 第4の実施形態の他の変形例に係る音声復号装置の動作を説明するためのフローチャートである。

[図44] 第1の実施形態の他の変形例に係る音声符号化装置の構成を示す図である。

[図45] 第1の実施形態の他の変形例に係る音声符号化装置の構成を示す図である。

[図46] 第2の実施形態の変形例に係る音声符号化装置の構成を示す図である。

[図47] 第2の実施形態の他の変形例に係る音声符号化装置の構成を示す図である。

[図48] 第4の実施形態に係る音声符号化装置の構成を示す図である。

[図49] 第4の実施形態の他の変形例に係る音声符号化装置の構成を示す図である。

[図50] 第4の実施形態の他の変形例に係る音声符号化装置の構成を示す図である。

### 発明を実施するための形態

[0048] 以下、図面を参照して、本発明に係る好適な実施形態について詳細に説明する。なお、図面の説明において、可能な場合には、同一要素には同一符号

を付し、重複する説明を省略する。

[0049] (第1の実施形態)

図1は、第1の実施形態に係る音声符号化装置11の構成を示す図である。音声符号化装置11は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声符号化装置11の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図2のフローチャートに示す処理を行うためのコンピュータプログラム）をRAMにロードして実行することによって音声符号化装置11を統括的に制御する。音声符号化装置11の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。

[0050] 音声符号化装置11は、機能的には、周波数変換部1a（周波数変換手段）、周波数逆変換部1b、コアコーデック符号化部1c（コア符号化手段）、SBR符号化部1d、線形予測分析部1e（時間エンベロープ補助情報算出手段）、フィルタ強度パラメータ算出部1f（時間エンベロープ補助情報算出手段）及びビットストリーム多重化部1g（ビットストリーム多重化手段）を備える。図1に示す音声符号化装置11の周波数変換部1a～ビットストリーム多重化部1gは、音声符号化装置11のCPUが音声符号化装置11の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。音声符号化装置11のCPUは、このコンピュータプログラムを実行することによって（図1に示す周波数変換部1a～ビットストリーム多重化部1gを用いて）、図2のフローチャートに示す処理（ステップS a 1～ステップS a 7の処理）を順次実行する。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声符号化装置11のROMやRAM等の内蔵メモリに格納されるものとする。

[0051] 周波数変換部1aは、音声符号化装置11の通信装置を介して受信された外部からの入力信号を多分割QMFフィルタバンクにより分析し、QMF領域の信号 $q(k, r)$ を得る（ステップS a 1の処理）。ただし、 $k$  ( $0 \leq$

$k \leq 63$ ) は周波数方向のインデックスであり、 $r$  は時間スロットを示すインデックスである。周波数逆変換部 1 b は、周波数変換部 1 a から得られた QMF 領域の信号のうち、低周波側の半数の係数を QMF フィルタバンクにより合成し、入力信号の低周波成分のみを含むダウンサンプルされた時間領域信号を得る (ステップ S a 2 の処理)。コアコーデック符号化部 1 c は、ダウンサンプルされた時間領域信号を符号化し、符号化ビットストリームを得る (ステップ S a 3 の処理)。コアコーデック符号化部 1 c における符号化は CELP 方式に代表される音声符号化方式に基づいてもよく、また AAC に代表される変換符号化や TCX (Transform Coded Excitation) 方式などの音響符号化に基づいてもよい。

[0052] SBR 符号化部 1 d は、周波数変換部 1 a から QMF 領域の信号を受け取り、高周波成分の電力・信号変化・トナリティ等の分析に基づいて SBR 符号化を行い、SBR 補助情報を得る (ステップ S a 4 の処理)。周波数変換部 1 a における QMF 分析の方法および SBR 符号化部 1 d における SBR 符号化の方法は、例えば文献 “3GPP TS 26.404; Enhanced aacPlus encoder SBR part” に詳述されている。

[0053] 線形予測分析部 1 e は、周波数変換部 1 a から QMF 領域の信号を受け取り、この信号の高周波成分に対し周波数方向に線形予測分析を行って高周波線形予測係数  $a_H(n, r)$  ( $1 \leq n \leq N$ ) を取得する (ステップ S a 5 の処理)。ただし  $N$  は線形予測次数である。また、インデックス  $r$  は、QMF 領域の信号のサブサンプルに関する時間方向のインデックスである。信号線形予測分析には、共分散法又は自己相関法を用いることができる。 $a_H(n, r)$  を取得する際の線形予測分析は、 $q(k, r)$  のうち  $k_x < k \leq 63$  をみたく高周波成分に対して行う。ただし  $k_x$  はコアコーデック符号化部 1 c によって符号化される周波数帯域の上限周波数に対応する周波数インデックスである。また、線形予測分析部 1 e は、 $a_H(n, r)$  を取得する際に分析したのとは別の低周波成分に対して線形予測分析を行い、 $a_H(n, r)$  とは別の低周波線形予測係数  $a_L(n, r)$  を取得してもよい (このような低周波成分に

係る線形予測係数は時間エンベロープ情報に対応しており、以下、第1の実施形態においては同様)。  $a_L(n, r)$  を取得する際の線形予測分析は、 $0 \leq k < k_x$  をみたす低周波成分に対するものである。また、この線形予測分析は  $0 \leq k < k_x$  の区間に含まれる一部の周波数帯域に対するものであってもよい。

- [0054] フィルタ強度パラメータ算出部1fは、例えば、線形予測分析部1eによって取得された線形予測係数を用いてフィルタ強度パラメータ（フィルタ強度パラメータは時間エンベロープ補助情報に対応しており、以下、第1の実施形態においては同様）を算出する（ステップS a 6の処理）。まず、 $a_H(n, r)$  から予測ゲイン  $G_H(r)$  が算出される。予測ゲインの算出方法は、たとえば“音声符号化、守谷健弘著、電子情報通信学会編”に詳述されている。さらに、 $a_L(n, r)$  が算出されている場合には同様に予測ゲイン  $G_L(r)$  が算出される。フィルタ強度パラメータ  $K(r)$  は、 $G_H(r)$  が大きいほど大きくなるパラメータであり、例えば次の数式(1)に従って取得することができる。ただし、 $\max(a, b)$  は  $a$  と  $b$  の最大値、 $\min(a, b)$  は  $a$  と  $b$  の最小値を示す。

[数1]

$$K(r) = \max(0, \min(1, G_H(r) - 1))$$

- [0055] また、 $G_L(r)$  が算出されている場合には、 $K(r)$  は  $G_H(r)$  が大きいほど大きくなり、 $G_L(r)$  が大きくなるほど小さくなるパラメータとして取得することができる。この場合の  $K$  は例えば次の数式(2)に従って取得することができる。

[数2]

$$K(r) = \max(0, \min(1, G_H(r)/G_L(r) - 1))$$

- [0056]  $K(r)$  は、SBR復号時に高周波成分の時間エンベロープを調整する強度を示すパラメータである。周波数方向の線形予測係数に対する予測ゲインは、分析区間の信号の時間エンベロープが急峻な変化を示すほど大きな値と

なる。 $K(r)$ は、その値が大きいほど、SBRによって生成された高周波成分の時間エンベロープの変化を急峻にする処理を強めるよう復号器に指示するためのパラメータである。なお、 $K(r)$ は、その値が小さいほど、SBRによって生成された高周波成分の時間エンベロープを急峻にする処理を弱めるよう復号器（例えば、音声復号装置21等）に指示するためのパラメータであってもよく、時間エンベロープを急峻にする処理を実行しないことを示す値を含んでも良い。また、各時間スロットの $K(r)$ を伝送せずに、複数の時間スロットに対して代表する $K(r)$ を伝送しても良い。同一の $K(r)$ の値を共有する時間スロットの区間を決定するためには、SBR補助情報に含まれるSBRエンベロープの時間境界（SBR envelope time border）情報を用いることが望ましい。

[0057]  $K(r)$ は、量子化された後にビットストリーム多重化部1gに送信される。量子化の前に複数の時間スロット $r$ について例えば $K(r)$ の平均をとることにより、複数の時間スロットに対して代表する $K(r)$ を計算することが望ましい。また、複数の時間スロットを代表する $K(r)$ を伝送する場合には、 $K(r)$ の算出を数式（2）のように個々の時間スロットを分析した結果から独立に行うのではなく、複数の時間スロットからなる区間全体の分析結果からそれらを代表する $K(r)$ を取得してもよい。この場合の $K(r)$ の算出は例えば次の数式（3）に従って行うことができる。ただし、 $\text{mean}(\cdot)$ は、 $K(r)$ によって代表される時間スロットの区間内における平均値を示す。

[数3]

$$K(r) = \max(0, \min(1, \text{mean}(G_H(r))/\text{mean}(G_L(r)) - 1)))$$

[0058] なお、 $K(r)$ を伝送する際には、“ISO/IEC 14496-3 subpart 4 General Audio Coding”に記載のSBR補助情報に含まれる逆フィルタモード情報と排他的に伝送しても良い。すなわち、SBR補助情報の逆フィルタモード情報を伝送する時間スロットに対しては $K(r)$ を伝送せず、 $K(r)$ を伝送

する時間スロットに対してはSBR補助情報の逆フィルタモード情報（“ISO/IEC 14496-3 subpart 4 General Audio Coding”におけるbs#invf#mode）を送信しなくてもよい。なお、K（r）又はSBR補助情報に含まれる逆フィルタモード情報のいずれを送信するかを示す情報を付加してもよい。また、K（r）とSBR補助情報に含まれる逆フィルタモード情報とを組み合わせ、一つのベクトル情報として取り扱い、このベクトルをエントロピー符号化してもよい。この際、K（r）と、SBR補助情報に含まれる逆フィルタモード情報との値の組み合わせに制約を加えてもよい。

[0059] ビットストリーム多重化部1gは、コアコーデック符号化部1cによって算出された符号化ビットストリームと、SBR符号化部1dによって算出されたSBR補助情報と、フィルタ強度パラメータ算出部1fによって算出されたK（r）と、を多重化し、多重化ビットストリーム（符号化された多重化ビットストリーム）を、音声符号化装置11の通信装置を介して出力する（ステップS a 7の処理）。

[0060] 図3は、第1の実施形態に係る音声復号装置21の構成を示す図である。音声復号装置21は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置21の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図4のフローチャートに示す処理を行うためのコンピュータプログラム）をRAMにロードして実行することによって音声復号装置21を統括的に制御する。音声復号装置21の通信装置は、音声符号化装置11、後述の変形例1の音声符号化装置11a、又は、後述の変形例2の音声符号化装置から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置21は、図3に示すように、機能的には、ビットストリーム分離部2a（ビットストリーム分離手段）、コアコーデック復号部2b（コア復号手段）、周波数変換部2c（周波数変換手段）、低周波線形予測分析部2d（低周波時間エンベロープ分析手段）、信号変化検出部2e、フィルタ強度調整部2f（時間エンベロープ調整手段）、高周波生成部2g（

高周波生成手段)、高周波線形予測分析部2h、線形予測逆フィルタ部2i、高周波調整部2j(高周波調整手段)、線形予測フィルタ部2k(時間エンベロープ変形手段)、係数加算部2m及び周波数逆変換部2nを備える。図3に示す音声復号装置21のビットストリーム分離部2a~エンベロープ形状パラメータ算出部1nは、音声復号装置21のCPUが音声復号装置21の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。音声復号装置21のCPUは、このコンピュータプログラムを実行することによって(図3に示すビットストリーム分離部2a~エンベロープ形状パラメータ算出部1nを用いて)、図4のフローチャートに示す処理(ステップSb1~ステップSb11の処理)を順次実行する。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声復号装置21のROMやRAM等の内蔵メモリに格納されるものとする。

[0061] ビットストリーム分離部2aは、音声復号装置21の通信装置を介して入力された多重化ビットストリームを、フィルタ強度パラメータと、SBR補助情報と、符号化ビットストリームとに分離する。コアコーデック復号部2bは、ビットストリーム分離部2aから与えられた符号化ビットストリームを復号し、低周波成分のみを含む復号信号を得る(ステップSb1の処理)。この際、復号の方式は、CELP方式に代表される音声符号化方式に基づいてもよく、またAACやTCX(Transform Coded Excitation)方式などの音響符号化に基づいてもよい。

[0062] 周波数変換部2cは、コアコーデック復号部2bから与えられた復号信号を多分割QMFフィルタバンクにより分析し、QMF領域の信号 $q_{\text{dec}}(k, r)$ を得る(ステップSb2の処理)。ただし、 $k$ ( $0 \leq k \leq 63$ )は周波数方向のインデックスであり、 $r$ はQMF領域の信号のサブサンプルに関する時間方向のインデックスを示すインデックスである。

[0063] 低周波線形予測分析部2dは、周波数変換部2cから得られた $q_{\text{dec}}(k, r)$ を時間スロット $r$ の各々に関して周波数方向に線形予測分析し、低周波線

形予測係数  $a_{dec}(n, r)$  を取得する（ステップ S b 3 の処理）。線形予測分析は、コアコーデック復号部 2 b から得られた復号信号の信号帯域に対応する  $0 \leq k < k_x$  の範囲に対して行う。また、この線形予測分析は  $0 \leq k < k_x$  の区間に含まれる一部の周波数帯域に対するものであってもよい。

[0064] 信号変化検出部 2 e は、周波数変換部 2 c から得られた QMF 領域の信号の時間変化を検出し、検出結果  $T(r)$  として出力する。信号変化の検出は、例えば以下に示す方法によって行うことができる。

1. 時間スロット  $r$  における信号の短時間電力  $p(r)$  を次の数式 (4) によって取得する。

[数4]

$$p(r) = \sum_{k=0}^{63} |q_{dec}(k, r)|^2$$

2.  $p(r)$  を平滑化したエンベロープ  $p_{env}(r)$  を次の数式 (5) によって取得する。ただし  $\alpha$  は  $0 < \alpha < 1$  を満たす定数である。

[数5]

$$p_{env}(r) = \alpha \cdot p_{env}(r-1) + (1-\alpha) \cdot p(r)$$

3.  $p(r)$  と  $p_{env}(r)$  とを用いて  $T(r)$  を次の数式 (6) に従って取得する。ただし  $\beta$  は定数である。

[数6]

$$T(r) = \max(1, p(r)/(\beta \cdot p_{env}(r)))$$

以上に示した方法は電力の変化に基づく信号変化検出の単純な例であり、他のもっと洗練された方法により信号変化検出を行ってもよい。また、信号変化検出部 2 e は省略してもよい。

[0065] フィルタ強度調整部 2 f は、低周波線形予測分析部 2 d から得られた  $a_{dec}(n, r)$  に対してフィルタ強度の調整を行い、調整された線形予測係数  $a_{adj}(n, r)$  を得る（ステップ S b 4 の処理）。フィルタ強度の調整は、ビット

ストリーム分離部 2 a を介して受信されたフィルタ強度パラメータ  $K$  を用いて、たとえば次の数式 (7) に従って行うことができる。

[数7]

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N)$$

さらに、信号変化検出部 2 e の出力  $T(r)$  が得られる場合には、強度の調整は次の数式 (8) に従って行ってもよい。

[数8]

$$a_{adj}(n, r) = a_{dec}(n, r) \cdot (K(r) \cdot T(r))^n \quad (1 \leq n \leq N)$$

[0066] 高周波生成部 2 g は、周波数変換部 2 c から得られた QMF 領域の信号を低周波帯域から高周波帯域に複写し、高周波成分の QMF 領域の信号  $q_{exp}(k, r)$  を生成する (ステップ S b 5 の処理)。高周波の生成は、“MPEG4 AAC” の SBR における HF generation の方法に従って行う (“ISO/IEC 14496-3 subpart 4 General Audio Coding”)。

[0067] 高周波線形予測分析部 2 h は、高周波生成部 2 g によって生成された  $q_{exp}(k, r)$  を時間スロット  $r$  の各々に関して周波数方向に線形予測分析し、高周波線形予測係数  $a_{exp}(n, r)$  を取得する (ステップ S b 6 の処理)。線形予測分析は、高周波生成部 2 g によって生成された高周波成分に対応する  $k_x \leq k \leq 63$  の範囲に対して行う。

[0068] 線形予測逆フィルタ部 2 i は、高周波生成部 2 g によって生成された高周波帯域の QMF 領域の信号を対象とし、周波数方向に  $a_{exp}(n, r)$  を係数とする線形予測逆フィルタ処理を行う (ステップ S b 7 の処理)。線形予測逆フィルタの伝達関数は次の数式 (9) の通りである。

[数9]

$$f(z) = 1 + \sum_{n=1}^N a_{exp}(n, r) z^{-n}$$

この線形予測逆フィルタ処理は、低周波側の係数から高周波側の係数に向

かって行われてもよいし、その逆でもよい。線形予測逆フィルタ処理は、後段において時間エンベロープ変形を行う前に高周波成分の時間エンベロープを一旦平坦化しておくための処理であり、線形予測逆フィルタ部 2 i は省略されてもよい。また、高周波生成部 2 g からの出力に対して高周波成分への線形予測分析と逆フィルタ処理を行うかわりに、後述する高周波調整部 2 j からの出力に対して高周波線形予測分析部 2 h による線形予測分析と線形予測逆フィルタ部 2 i による逆フィルタ処理とを行ってもよい。さらに、線形予測逆フィルタ処理に用いる線形予測係数は、 $a_{\text{exp}}(n, r)$  ではなく、 $a_{\text{dec}}(n, r)$  又は  $a_{\text{adj}}(n, r)$  であってもよい。また、線形予測逆フィルタ処理に用いられる線形予測係数は、 $a_{\text{exp}}(n, r)$  に対してフィルタ強度調整を行って取得される線形予測係数  $a_{\text{exp, adj}}(n, r)$  であってもよい。強度調整は、 $a_{\text{adj}}(n, r)$  を取得する際と同様、例えば、次の数式 (10) に従って行われる。

[数10]

$$a_{\text{exp, adj}}(n, r) = a_{\text{exp}}(n, r) \cdot K(r)^n \quad (1 \leq n \leq N)$$

[0069] 高周波調整部 2 j は、線形予測逆フィルタ部 2 i からの出力に対して高周波成分の周波数特性およびトーナリティの調整を行う（ステップ S b 8 の処理）。この調整はビットストリーム分離部 2 a から与えられた S B R 補助情報に従って行われる。高周波調整部 2 j による処理は、“MPEG4 AAC” の S B R における “HF adjustment” ステップに従って行われるものであり、高周波帯域の Q M F 領域の信号に対し、時間方向の線形予測逆フィルタ処理、ゲインの調整及びノイズの重畳を行うことによる調整である。以上のステップにおける処理の詳細については “ISO/IEC 14496-3 subpart 4 General Audio Coding” に詳述されている。なお、上記したように、周波数変換部 2 c、高周波生成部 2 g 及び高周波調整部 2 j は、全て、“ISO/IEC 14496-3” に規定される “MPEG4 AAC” における S B R 復号器に準拠した動作をする。

[0070] 線形予測フィルタ部 2 k は、高周波調整部 2 j から出力された Q M F 領域

の信号の高周波成分  $q_{adj}(n, r)$  に対し、フィルタ強度調整部 2 f から得られた  $a_{adj}(n, r)$  を用いて周波数方向に線形予測合成フィルタ処理を行う（ステップ S b 9 の処理）。線形予測合成フィルタ処理における伝達関数は次の数式（11）の通りである。

[数11]

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_{adj}(n, r) z^{-n}}$$

この線形予測合成フィルタ処理によって、線形予測フィルタ部 2 k は、SBR に基づいて生成された高周波成分の時間エンベロープを変形する。

[0071] 係数加算部 2 m は、周波数変換部 2 c から出力された低周波成分を含む QMF 領域の信号と、線形予測フィルタ部 2 k から出力された高周波成分を含む QMF 領域の信号とを加算し、低周波成分と高周波成分の双方を含む QMF 領域の信号を出力する（ステップ S b 10 の処理）。

[0072] 周波数逆変換部 2 n は、係数加算部 2 m から得られた QMF 領域の信号を QMF 合成フィルタバンクによって処理する。これによって、コアコーデックの復号によって得られた低周波成分と、SBR によって生成され線形予測フィルタによって時間エンベロープが変形された高周波成分との双方を含む時間領域の復号した音声信号を取得し、この取得した音声信号を、内蔵する通信装置を介して外部に出力する（ステップ S b 11 の処理）。なお、周波数逆変換部 2 n は、 $K(r)$  と “ISO/IEC 14496-3 subpart 4 General Audio Coding” に記載の SBR 補助情報の逆フィルタモード情報とが排他的に伝送された場合、 $K(r)$  が伝送され SBR 補助情報の逆フィルタモード情報の伝送されない時間スロットに対しては、当該時間スロットの前後における時間スロットのうちの少なくとも一つの時間スロットに対する SBR 補助情報の逆フィルタモード情報を用いて、当該時間スロットの SBR 補助情報の逆フィルタモード情報を生成しても良く、当該時間スロットの SBR 補助情報の逆フィルタモード情報をあらかじめ決められた所定のモードに設定しても

良い。一方、周波数逆変換部 2 n は、S B R 補助情報の逆フィルタデータが伝送され K ( r ) の伝送されない時間スロットに対しては、当該時間スロットの前後における時間スロットのうちの少なくとも一つの時間スロットに対する K ( r ) を用いて、当該時間スロットの K ( r ) を生成しても良く、当該時間スロットの K ( r ) を予め決められた所定の値に設定しても良い。なお、周波数逆変換部 2 n は、K ( r ) 又は S B R 補助情報の逆フィルタモード情報のいずれを伝送したかを示す情報に基づき、伝送された情報が、K ( r ) か、S B R 補助情報の逆フィルタモード情報か、を判断しても良い。

[0073] (第 1 の実施形態の変形例 1)

図 5 は、第 1 の実施形態に係る音声符号化装置の変形例（音声符号化装置 1 1 a）の構成を示す図である。音声符号化装置 1 1 a は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 1 a の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 1 a を統括的に制御する。音声符号化装置 1 1 a の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。

[0074] 音声符号化装置 1 1 a は、図 5 に示すように、機能的には、音声符号化装置 1 1 の線形予測分析部 1 e、フィルタ強度パラメータ算出部 1 f 及びビットストリーム多重化部 1 g にかえて、高周波周波数逆変換部 1 h、短時間電力算出部 1 i（時間エンベロープ補助情報算出手段）、フィルタ強度パラメータ算出部 1 f 1（時間エンベロープ補助情報算出手段）及びビットストリーム多重化部 1 g 1（ビットストリーム多重化手段）を備える。ビットストリーム多重化部 1 g 1 は 1 G と同様の機能を有する。図 5 に示す音声符号化装置 1 1 a の周波数変換部 1 a ~ S B R 符号化部 1 d、高周波周波数逆変換部 1 h、短時間電力算出部 1 i、フィルタ強度パラメータ算出部 1 f 1 及びビットストリーム多重化部 1 g 1 は、音声符号化装置 1 1 a の CPU が音声符号化装置 1 1 a の内蔵メモリに格納されたコンピュータプログラムを実行

することによって実現される機能である。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声符号化装置 11a の ROM や RAM 等の内蔵メモリに格納されるものとする。

[0075] 高周波周波数逆変換部 1h は、周波数変換部 1a から得られた QMF 領域の信号のうち、コアコーデック符号化部 1c によって符号化される低周波成分に対応する係数を “0” に置き換えた後に QMF 合成フィルタバンクを用いて処理し、高周波成分のみが含まれた時間領域信号を得る。短時間電力算出部 1i は、高周波周波数逆変換部 1h から得られた時間領域の高周波成分を短区間に区切ってその電力を算出し、 $p(r)$  を算出する。なお、代替的な方法として、QMF 領域の信号を用いて次の数式 (12) に従って短時間電力を算出してもよい。

[数12]

$$p(r) = \sum_{k=0}^{63} |q(k, r)|^2$$

[0076] フィルタ強度パラメータ算出部 1f1 は、 $p(r)$  の変化部分を検出し、変化が大きいほど  $K(r)$  が大きくなるよう、 $K(r)$  の値を決定する。 $K(r)$  の値は、例えば、音声復号装置 21 の信号変化検出部 2e における  $T(r)$  の算出と同一の方法で行ってもよい。また、他のもっと洗練された方法により信号変化検出を行ってもよい。また、フィルタ強度パラメータ算出部 1f1 は、低周波成分と高周波成分の各々について短時間電力を取得した後に音声復号装置 21 の信号変化検出部 2e における  $T(r)$  の算出と同一の方法によって低周波成分及び高周波成分各々の信号変化  $T_r(r)$ 、 $T_h(r)$  を取得し、これらを用いて  $K(r)$  の値を決定してもよい。この場合、 $K(r)$  は例えば次の数式 (13) に従って取得することができる。ただし、 $\varepsilon$  は、例えば 3.0 などの定数である。

[数13]

$$K(r)=\max(0, \varepsilon \cdot (Th(r)-Tr(r)))$$

[0077] (第1の実施形態の変形例2)

第1の実施形態の変形例2の音声符号化装置（不図示）は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の変形例2の音声符号化装置の内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって変形例2の音声符号化装置を統括的に制御する。変形例2の音声符号化装置の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。

[0078] 変形例2の音声符号化装置は、機能的には、音声符号化装置11のフィルタ強度パラメータ算出部1f及びビットストリーム多重化部1gにかえて、図示しない線形予測係数差分符号化部（時間エンベロープ補助情報算出手段）と、この線形予測係数差分符号化部からの出力を受けるビットストリーム多重化部（ビットストリーム多重化手段）とを備える。変形例2の音声符号化装置の周波数変換部1a～線形予測分析部1e、線形予測係数差分符号化部、及び、ビットストリーム多重化部は、変形例2の音声符号化装置のCPUが変形例2の音声符号化装置の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、変形例2の音声符号化装置のROMやRAM等の内蔵メモリに格納されるものとする。

[0079] 線形予測係数差分符号化部は、入力信号の $a_H(n, r)$ と入力信号の $a_L(n, r)$ を用い、次の数式(14)に従って線形予測係数の差分値 $a_D(n, r)$ を算出する。

[数14]

$$a_D(n,r)=a_H(n,r)-a_L(n,r) \quad (1 \leq n \leq N)$$

[0080] 線形予測係数差分符号化部は、さらに  $a_D(n, r)$  を量子化し、ビットストリーム多重化部（ビットストリーム多重化部 1 g に対応する構成）へ送信する。このビットストリーム多重化部は、 $K(r)$  に代わり  $a_D(n, r)$  をビットストリームに多重化し、この多重化ビットストリームを内蔵する通信装置を介して外部に出力する。

[0081] 第 1 の実施形態の変形例 2 の音声復号装置（不図示）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の変形例 2 の音声復号装置の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって変形例 2 の音声復号装置を統括的に制御する。変形例 2 の音声復号装置の通信装置は、音声符号化装置 1 1、変形例 1 に係る音声符号化装置 1 1 a、又は、変形例 2 に係る音声符号化装置から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。

[0082] 変形例 2 の音声復号装置は、機能的には、音声復号装置 2 1 のフィルタ強度調整部 2 f にかえて、図示しない線形予測係数差分復号部を備える。変形例 2 の音声復号装置のビットストリーム分離部 2 a ～信号変化検出部 2 e、線形予測係数差分復号部、及び、高周波生成部 2 g ～周波数逆変換部 2 n は、変形例 2 の音声復号装置の CPU が変形例 2 の音声復号装置の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、変形例 2 の音声復号装置の ROM や RAM 等の内蔵メモリに格納されるものとする。

[0083] 線形予測係数差分復号部は、低周波線形予測分析部 2 d から得られた  $a_L(n, r)$  とビットストリーム分離部 2 a から与えられた  $a_D(n, r)$  を利用し、次の数式（15）に従って差分復号された  $a_{adj}(n, r)$  を得る。

[数15]

$$a_{adj}(n,r)=a_{dec}(n,r)+a_D(n,r), 1 \leq n \leq N$$

[0084] 線形予測係数差分復号部は、このようにして差分復号された  $a_{adj}(n, r)$  を線形予測フィルタ部 2 k に送信する。 $a_D(n, r)$  は、数式 (14) に示すように予測係数の領域での差分値であってもよいが、予測係数を LSP (Linear Spectrum Pair)、ISP (Immittance Spectrum Pair)、LSF (Linear Spectrum Frequency)、ISF (Immittance Spectrum Frequency)、PARCOR 係数などの別の表現形式に変換した後に差分をとった値であってもよい。この場合、差分復号も同じこの表現形式と同様となる。

[0085] (第 2 の実施形態)

図 6 は、第 2 の実施形態に係る音声符号化装置 12 の構成を示す図である。音声符号化装置 12 は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 12 の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図 7 のフローチャートに示す処理を行うためのコンピュータプログラム）を RAM にロードして実行することによって音声符号化装置 12 を統括的に制御する。音声符号化装置 12 の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。

[0086] 音声符号化装置 12 は、機能的には、音声符号化装置 11 のフィルタ強度パラメータ算出部 1 f 及びビットストリーム多重化部 1 g にかえて、線形予測係数間引き部 1 j（予測係数間引き手段）、線形予測係数量子化部 1 k（予測係数量子化手段）及びビットストリーム多重化部 1 g 2（ビットストリーム多重化手段）を備える。図 6 に示す音声符号化装置 12 の周波数変換部 1 a ～線形予測分析部 1 e（線形予測分析手段）、線形予測係数間引き部 1 j、線形予測係数量子化部 1 k 及びビットストリーム多重化部 1 g 2 は、音声符号化装置 12 の CPU が音声符号化装置 12 の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。音声符号化装置 12 の CPU は、このコンピュータプログラムを実行することによって（図 6 に示す音声符号化装置 12 の周波数変換部 1 a ～線形予測分析部 1 e、線形予測係数間引き部 1 j、線形予測係数量子化部 1 k 及びビッ

トストリーム多重化部 1 g 2 を用いて)、図 7 のフローチャートに示す処理 (ステップ S a 1 ~ ステップ S a 5、及び、ステップ S c 1 ~ ステップ S c 3 の処理) を順次実行する。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声符号化装置 1 2 の ROM や RAM 等の内蔵メモリに格納されるものとする。

[0087] 線形予測係数間引き部 1 j は、線形予測分析部 1 e から得られた  $a_H(n, r)$  を時間方向に間引き、 $a_H(n, r)$  のうち一部の時間スロット  $r_i$  に対する値と、対応する  $r_i$  の値を線形予測係数量子化部 1 k に送信する (ステップ S c 1 の処理)。ただし、 $0 \leq i < N_{ts}$  であり、 $N_{ts}$  はフレーム中で  $a_H(n, r)$  の伝送が行われる時間スロットの数である。線形予測係数の間引きは、一定の時間間隔によるものであってもよく、また、 $a_H(n, r)$  の性質に基づく不等時間間隔の間引きであってもよい。例えば、ある長さを持つフレームの中で  $a_H(n, r)$  の  $G_H(r)$  を比較し、 $G_H(r)$  が一定の値を超えた場合に  $a_H(n, r)$  を量子化の対象とするなどの方法が考えられる。線形予測係数の間引き間隔を  $a_H(n, r)$  の性質によらず一定の間隔とする場合には、伝送の対象とならない時間スロットに対しては  $a_H(n, r)$  を算出する必要がない。

[0088] 線形予測係数量子化部 1 k は、線形予測係数間引き部 1 j から与えられた間引き後の高周波線形予測係数  $a_H(n, r_i)$  と、対応する時間スロットのインデックス  $r_i$  を量子化し、ビットストリーム多重化部 1 g 2 に送信する (ステップ S c 2 の処理)。なお、代替的な構成として、 $a_H(n, r_i)$  を量子化するかわりに、第 1 の実施形態の変形例 2 に係る音声符号化装置と同様に、線形予測係数の差分値  $a_D(n, r_i)$  を量子化の対象としてもよい。

[0089] ビットストリーム多重化部 1 g 2 は、コアコーデック符号化部 1 c で算出された符号化ビットストリームと、SBR 符号化部 1 d で算出された SBR 補助情報と、線形予測係数量子化部 1 k から与えられた量子化後の  $a_H(n, r_i)$  に対応する時間スロットのインデックス  $\{r_i\}$  とをビットストリームに

多重化し、この多重化ビットストリームを、音声符号化装置 12 の通信装置を介して出力する（ステップ S c 3 の処理）。

[0090] 図 8 は、第 2 の実施形態に係る音声復号装置 22 の構成を示す図である。音声復号装置 22 は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 22 の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図 9 のフローチャートに示す処理を行うためのコンピュータプログラム）を RAM にロードして実行することによって音声復号装置 22 を統括的に制御する。音声復号装置 22 の通信装置は、音声符号化装置 12 から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。

[0091] 音声復号装置 22 は、機能的には、音声復号装置 21 のビットストリーム分離部 2 a、低周波線形予測分析部 2 d、信号変化検出部 2 e、フィルタ強度調整部 2 f 及び線形予測フィルタ部 2 k にかえて、ビットストリーム分離部 2 a 1（ビットストリーム分離手段）、線形予測係数補間・補外部 2 p（線形予測係数補間・補外手段）及び線形予測フィルタ部 2 k 1（時間エンベロープ変形手段）を備える。図 8 に示す音声復号装置 22 のビットストリーム分離部 2 a 1、コアコーデック復号部 2 b、周波数変換部 2 c、高周波生成部 2 g～高周波調整部 2 j、線形予測フィルタ部 2 k 1、係数加算部 2 m、周波数逆変換部 2 n、及び、線形予測係数補間・補外部 2 p は、音声符号化装置 12 の CPU が音声符号化装置 12 の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。音声復号装置 22 の CPU は、このコンピュータプログラムを実行することによって（図 8 に示すビットストリーム分離部 2 a 1、コアコーデック復号部 2 b、周波数変換部 2 c、高周波生成部 2 g～高周波調整部 2 j、線形予測フィルタ部 2 k 1、係数加算部 2 m、周波数逆変換部 2 n、及び、線形予測係数補間・補外部 2 p を用いて）、図 9 のフローチャートに示す処理（ステップ S b 1～ステップ S b 2、ステップ S d 1、ステップ S b 5～ステップ S b 8、ステップ S d 2、及び、ステップ S b 10～ステップ S b 11 の処理）を

順次実行する。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声復号装置 22 の ROM や RAM 等の内蔵メモリに格納されるものとする。

[0092] 音声復号装置 22 は、音声復号装置 22 のビットストリーム分離部 2 a、低周波線形予測分析部 2 d、信号変化検出部 2 e、フィルタ強度調整部 2 f 及び線形予測フィルタ部 2 k にかえて、ビットストリーム分離部 2 a 1、線形予測係数補間・補外部 2 p 及び線形予測フィルタ部 2 k 1 を備える。

[0093] ビットストリーム分離部 2 a 1 は、音声復号装置 22 の通信装置を介して入力された多重化ビットストリームを、量子化された  $a_H(n, r_i)$  に対応する時間スロットのインデックス  $r_i$  と、SBR 補助情報と、符号化ビットストリームとに分離する。

[0094] 線形予測係数補間・補外部 2 p は、量子化された  $a_H(n, r_i)$  に対応する時間スロットのインデックス  $r_i$  をビットストリーム分離部 2 a 1 から受け取り、線形予測係数の伝送されていない時間スロットに対応する  $a_H(n, r)$  を、補間又は補外により取得する（ステップ S d 1 の処理）。線形予測係数補間・補外部 2 p は、線形予測係数の補外を、例えば次の数式（16）に従って行うことができる。

[数16]

$$a_H(n, r) = \delta^{|r-r_{i0}|} a_H(n, r_{i0}) \quad (1 \leq n \leq N)$$

ただし、 $r_{i0}$  は線形予測係数が伝送されている時間スロット  $\{r_i\}$  のうち  $r$  に最も近いものとする。また、 $\delta$  は  $0 < \delta < 1$  を満たす定数である。

[0095] また、線形予測係数補間・補外部 2 p は、線形予測係数の補間を、例えば次の数式（17）に従って行うことができる。ただし、 $r_{i0} < r < r_{i0+1}$  を満たす。

[数17]

$$a_H(n, r) = \frac{r_{i0+1} - r}{r_{i0+1} - r_i} \cdot a_H(n, r_i) + \frac{r - r_{i0}}{r_{i0+1} - r_{i0}} \cdot a_H(n, r_{i0+1}) \quad (1 \leq n \leq N)$$

[0096] なお、線形予測係数補間・補外部 2 p は、線形予測係数を L S P (Linear Spectrum Pair)、I S P (Immittance Spectrum Pair)、L S F (Linear Spectrum Frequency)、I S F (Immittance Spectrum Frequency)、P A R C O R 係数などの別の表現形式に変換した後に補間・補外し、得られた値を線形予測係数に変換して用いても良い。補間又は補外後の  $a_H(n, r)$  は線形予測フィルタ部 2 k 1 に送信され、線形予測合成フィルタ処理における線形予測係数として利用されるが、線形予測逆フィルタ部 2 i における線形予測係数として用いられてもよい。ビットストリームに  $a_H(n, r)$  ではなく  $a_D(n, r_i)$  が多重化されている場合、線形予測係数補間・補外部 2 p は、上記の補間又は補外処理に先立ち、第 1 の実施形態の変形例 2 に係る音声復号装置と同様の差分復号処理を行う。

[0097] 線形予測フィルタ部 2 k 1 は、高周波調整部 2 j から出力された  $q_{adj}(n, r)$  に対し、線形予測係数補間・補外部 2 p から得られた、補間又は補外された  $a_H(n, r)$  を用いて周波数方向に線形予測合成フィルタ処理を行う（ステップ S d 2 の処理）。線形予測フィルタ部 2 k 1 の伝達関数は次の数式（18）の通りである。線形予測フィルタ部 2 k 1 は、音声復号装置 2 1 の線形予測フィルタ部 2 k と同様に、線形予測合成フィルタ処理を行うことによって、S B R により生成された高周波成分の時間エンベロープを変形する。

[数18]

$$g(z) = \frac{1}{1 + \sum_{n=1}^N a_H(n, r) z^{-n}}$$

[0098] （第 3 の実施形態）

図10は、第3の実施形態に係る音声符号化装置13の構成を示す図である。音声符号化装置13は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声符号化装置13の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図11のフローチャートに示す処理を行うためのコンピュータプログラム）をRAMにロードして実行することによって音声符号化装置13を統括的に制御する。音声符号化装置13の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。

[0099] 音声符号化装置13は、機能的には、音声符号化装置11の線形予測分析部1e、フィルタ強度パラメータ算出部1f及びビットストリーム多重化部1gにかえて、時間エンベロープ算出部1m（時間エンベロープ補助情報算出手段）、エンベロープ形状パラメータ算出部1n（時間エンベロープ補助情報算出手段）及びビットストリーム多重化部1g3（ビットストリーム多重化手段）を備える。図10に示す音声符号化装置13の周波数変換部1a～SBR符号化部1d、時間エンベロープ算出部1m、エンベロープ形状パラメータ算出部1n、及び、ビットストリーム多重化部1g3は、音声符号化装置12のCPUが音声符号化装置12の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。音声符号化装置13のCPUは、このコンピュータプログラムを実行することによって（図10に示す音声符号化装置13の周波数変換部1a～SBR符号化部1d、時間エンベロープ算出部1m、エンベロープ形状パラメータ算出部1n、及び、ビットストリーム多重化部1g3を用いて）、図11のフローチャートに示す処理（ステップSa1～ステップSa4、及び、ステップSe1～ステップSe3の処理）を順次実行する。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声符号化装置13のROMやRAM等の内蔵メモリに格納されるものとする。

[0100] 時間エンベロープ算出部1mは、 $q(k, r)$ を受け取り、例えば、 $q$ （

$k, r$ ) の時間スロットごとの電力を取得することによって、信号の高周波成分の時間エンベロープ情報  $e(r)$  を取得する (ステップ S e 1 の処理)。この場合、 $e(r)$  は次の数式 (19) に従って取得される。

[数19]

$$e(r) = \sqrt{\sum_{k=k_x}^{63} |q(k, r)|^2}$$

[0101] エンベロープ形状パラメータ算出部 1 n は、時間エンベロープ算出部 1 m から  $e(r)$  を受け取り、さらに SBR 符号化部 1 d から SBR エンベロープの時間境界  $\{b_i\}$  を受け取る。ただし、 $0 \leq i \leq N_e$  であり、 $N_e$  は符号化フレーム内の SBR エンベロープの数である。エンベロープ形状パラメータ算出部 1 n は、符号化フレーム内の SBR エンベロープの各々について、例えば次の数式 (20) に従ってエンベロープ形状パラメータ  $s(i)$  ( $0 \leq i < N_e$ ) を取得する (ステップ S e 2 の処理)。なお、エンベロープ形状パラメータ  $s(i)$  は時間エンベロープ補助情報に対応しており、第 3 の実施形態において同様とする。

[数20]

$$s(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2$$

ただし、

[数21]

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i}$$

上記の数式における  $s(i)$  は  $b_i \leq r < b_{i+1}$  を満たす  $i$  番目の SBR エンベロープ内における  $e(r)$  の変化の大きさを示すパラメータであり、時間

エンベロープの変化が大きいほど  $e(r)$  は大きい値をとる。上記の数式 (20) 及び (21) は、 $s(i)$  の算出方法の一例であり、例えば  $e(r)$  の SMF (Spectral Flatness Measure) や、最大値と最小値の比等、を用いて  $s(i)$  を取得してもよい。この後、 $s(i)$  は量子化され、ビットストリーム多重化部 1 g 3 に伝送される。

[0102] ビットストリーム多重化部 1 g 3 は、コアコーデック符号化部 1 c によって算出された符号化ビットストリームと、SBR符号化部 1 d によって算出された SBR 補助情報と、 $s(i)$  とをビットストリームに多重化し、この多重化したビットストリームを、音声符号化装置 1 3 の通信装置を介して出力する (ステップ S e 3 の処理)。

[0103] 図 1 2 は、第 3 の実施形態に係る音声復号装置 2 3 の構成を示す図である。音声復号装置 2 3 は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 3 の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 1 3 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 2 3 を統括的に制御する。音声復号装置 2 3 の通信装置は、音声符号化装置 1 3 から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。

[0104] 音声復号装置 2 3 は、機能的には、音声復号装置 2 1 のビットストリーム分離部 2 a、低周波線形予測分析部 2 d、信号変化検出部 2 e、フィルタ強度調整部 2 f、高周波線形予測分析部 2 h、線形予測逆フィルタ部 2 i 及び線形予測フィルタ部 2 k にかえて、ビットストリーム分離部 2 a 2 (ビットストリーム分離手段)、低周波時間エンベロープ算出部 2 r (低周波時間エンベロープ分析手段)、エンベロープ形状調整部 2 s (時間エンベロープ調整手段)、高周波時間エンベロープ算出部 2 t、時間エンベロープ平坦化部 2 u 及び時間エンベロープ変形部 2 v (時間エンベロープ変形手段) を備える。図 1 2 に示す音声復号装置 2 3 のビットストリーム分離部 2 a 2、コア

$$\sqrt[k]{\sum_{\ell=0}^k |q^{\ell} p^{\ell}|} = e(r)$$

。この  $e(r)$  に対する調整は、例えば次の数式 (23) ~ (25) に従って行うことができる。

[数23]

$$e_{adj}(r) = \overline{e(i)} + \sqrt{s(i) - v(i)} \cdot (e(r) - \overline{e(i)}) \quad (s(i) > v(i))$$

$$e_{adj}(r) = e(r) \quad (\text{otherwise})$$

ただし、

[数24]

$$\overline{e(i)} = \frac{\sum_{r=b_i}^{b_{i+1}-1} e(r)}{b_{i+1} - b_i}$$

[数25]

$$v(i) = \frac{1}{b_{i+1} - b_i - 1} \sum_{r=b_i}^{b_{i+1}-1} (\overline{e(i)} - e(r))^2$$

である。

[0107] 上記の数式 (23) ~ (25) は調整方法の一例であり、 $e_{adj}(r)$  の形状が  $s(i)$  によって示される形状に近づくような他の調整方法を用いてもよい。

[0108] 高周波時間エンベロープ算出部 2 t は、高周波生成部 2 g から得られた  $q_{exp}(k, r)$  を用いて時間エンベロープ  $e_{exp}(r)$  を次の数式 (26) に従って算出する (ステップ S f 3 の処理)。

[数26]

$$e_{exp}(r) = \sqrt{\sum_{k=k_x}^{63} |q_{exp}(k, r)|^2}$$

[0109] 時間エンベロープ平坦化部 2 u は、高周波生成部 2 g から得られた  $q_{exp}(k$

、 $r$ ）の時間エンベロープを次の数式（27）に従って平坦化し、得られた QMF 領域の信号  $q_{flat}(k, r)$  を高周波調整部 2 j に送信する（ステップ S f 4 の処理）。

[数27]

$$q_{flat}(k, r) = \frac{q_{exp}(k, r)}{e_{exp}(r)} \quad (k_x \leq k \leq 63)$$

[0110] 時間エンベロープ平坦化部 2 u における時間エンベロープの平坦化は省略

されてもよい。また、高周波生成部 2 g からの出力に対して、高周波成分の時間エンベロープ算出と時間エンベロープの平坦化処理とを行うかわりに、高周波調整部 2 j からの出力に対して、高周波成分の時間エンベロープ算出と時間エンベロープの平坦化処理とを行ってもよい。さらに、時間エンベロープ平坦化部 2 u において用いる時間エンベロープは、高周波時間エンベロープ算出部 2 t から得られた  $e_{exp}(r)$  ではなく、エンベロープ形状調整部 2 s から得られた  $e_{adj}(r)$  であってもよい。

[0111] 時間エンベロープ変形部 2 v は、高周波調整部 2 j から得られた  $q_{adj}(k, r)$  を時間エンベロープ変形部 2 v から得られた  $e_{adj}(r)$  を用いて変形し、時間エンベロープが変形された QMF 領域の信号  $q_{envadj}(k, r)$  を取得する（ステップ S f 5 の処理）。この変形は、次の数式（28）に従って行われる。  $q_{envadj}(k, r)$  は高周波成分に対応する QMF 領域の信号として係数加算部 2 m に送信される。

[数28]

$$q_{envadj}(k, r) = q_{adj}(k, r) \cdot e_{adj}(r) \quad (k_x \leq k \leq 63)$$

[0112] （第 4 の実施形態）

図 1 4 は、第 4 の実施形態に係る音声復号装置 2 4 の構成を示す図である。音声復号装置 2 4 は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 の内蔵メモ

りに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声復号装置24を統括的に制御する。音声復号装置24の通信装置は、音声符号化装置11又は音声符号化装置13から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。

[0113] 音声復号装置23は、機能的には、音声復号装置21の構成（コアコーデック復号部2b、周波数変換部2c、低周波線形予測分析部2d、信号変化検出部2e、フィルタ強度調整部2f、高周波生成部2g、高周波線形予測分析部2h、線形予測逆フィルタ部2i、高周波調整部2j、線形予測フィルタ部2k、係数加算部2m及び周波数逆変換部2n）と、音声復号装置24の構成（低周波時間エンベロープ算出部2r、エンベロープ形状調整部2s及び時間エンベロープ変形部2v）とを備える。更に、音声復号装置24は、ビットストリーム分離部2a3（ビットストリーム分離手段）及び補助情報変換部2wを備える。線形予測フィルタ部2kと時間エンベロープ変形部2vの順序は図14に示すものと逆であってもよい。なお、音声復号装置24は、音声符号化装置11又は音声符号化装置13によって符号化されたビットストリームを入力とすることが望ましい。図14に示す音声復号装置24の構成は、音声復号装置24のCPUが音声復号装置24の内蔵メモリに格納されたコンピュータプログラムを実行することによって実現される機能である。このコンピュータプログラムの実行に必要な各種データ、及び、このコンピュータプログラムの実行によって生成された各種データは、全て、音声復号装置24のROMやRAM等の内蔵メモリに格納されるものとする。

[0114] ビットストリーム分離部2a3は、音声復号装置24の通信装置を介して入力された多重化ビットストリームを、時間エンベロープ補助情報と、SBR補助情報と、符号化ビットストリームとに分離する。時間エンベロープ補助情報は、第1の実施形態において説明した $K(r)$ 、又は、第3の実施形態において説明した $s(i)$ 、であってもよい。また、 $K(r)$ 、 $s(i)$

のいずれでもない他のパラメータ  $X(r)$  であってもよい。

- [0115] 補助情報変換部 2w は、入力された時間エンベロープ補助情報を変換し、 $K(r)$  と  $s(i)$  とを得る。時間エンベロープ補助情報が  $K(r)$  の場合、補助情報変換部 2w は、 $K(r)$  を  $s(i)$  に変換する。補助情報変換部 2w は、この変換を、例えば  $b_i \leq r < b_{i+1}$  の区間内での  $K(r)$  の平均値 [数29]

$$\overline{K(i)}$$

を取得した後に、所定のテーブルを用いて、この数式 (29) に示す平均値を  $s(i)$  に変換することによって行ってもよい。また、時間エンベロープ補助情報が  $s(i)$  の場合、補助情報変換部 2w は、 $s(i)$  を  $K(r)$  に変換する。補助情報変換部 2w は、この変換を、例えば所定のテーブルを用いて  $s(i)$  を  $K(r)$  に変換することによって行ってもよい。ただし、 $i$  と  $r$  は  $b_i \leq r < b_{i+1}$  の関係を満たすよう対応づけられるものとする。

- [0116] 時間エンベロープ補助情報が  $s(i)$  でも  $K(r)$  でもないパラメータ  $X(r)$  の場合、補助情報変換部 2w は、 $X(r)$  を、 $K(r)$  と  $s(i)$  とに変換する。補助情報変換部 2w は、この変換を、例えば所定のテーブルを用いて  $X(r)$  を  $K(r)$  および  $s(i)$  に変換することによって行うのが望ましい。また、補助情報変換部 2w は、 $X(r)$  を SBR エンベロープ毎に 1 つの代表値を伝送するのが望ましい。 $X(r)$  を  $K(r)$  および  $s(i)$  に変換するテーブルは互いに異なってもよい。

- [0117] (第 1 の実施形態の変形例 3)

第 1 の実施形態の音声復号装置 21 において、音声復号装置 21 の線形予測フィルタ部 2k は、自動利得制御処理を含むことができる。この自動利得制御処理は、線形予測フィルタ部 2k の出力の QMF 領域の信号の電力を入力された QMF 領域の信号電力に合わせる処理である。利得制御後の QMF 領域信号  $q_{\text{syn, pow}}(n, r)$  は、一般的には、次式により実現される。

[数30]

$$q_{syn,pow}(n,r) = q_{syn}(n,r) \cdot \sqrt{\frac{P_0(r)}{P_1(r)}}$$

ここで、 $P_0(r)$ 、 $P_1(r)$  はそれぞれ以下の数式（31）及び数式（32）で表される。

[数31]

$$P_0(r) = \sum_{n=k_x}^{63} |q_{adj}(n,r)|^2$$

[数32]

$$P_1(r) = \sum_{n=k_x}^{63} |q_{syn}(n,r)|^2$$

この自動利得制御処理により、線形予測フィルタ部2kの出力信号の高周波成分の電力は線形予測フィルタ処理前と等しい値に調整される。その結果、SBRに基づいて生成された高周波成分の時間エンベロップを変形した線形予測フィルタ部2kの出力信号において、高周波調整部2jにおいて行われた高周波信号の電力の調整の効果が保たれる。なお、この自動利得制御処理は、QMF領域の信号の任意の周波数範囲に対して個別に行うことも可能である。個々の周波数範囲に対する処理は、それぞれ、数式（30）、数式（31）、数式（32）のnをある周波数範囲に限定することで実現できる。例えばi番目の周波数範囲は $F_i \leq n < F_{i+1}$ と表すことができる（この場合のiは、QMF領域の信号の任意の周波数範囲の番号を示すインデックスである）。 $F_i$ は周波数範囲の境界を示し、“MPEG4 AAC”のSBRにおいて規定されるエンベロップスケールファクタの周波数境界テーブルであることが望ましい。周波数境界テーブルは“MPEG4 AAC”のSBRの規定に従い、高周波生成部2gにおいて決定される。この自動利得制御処理により、線形予測フ

フィルタ部 2 k の出力信号の高周波成分の任意の周波数範囲内の電力は線形予測フィルタ処理前と等しい値に調整される。その結果、SBRに基づいて生成された高周波成分の時間エンベロープを変形した線形予測フィルタ部 2 k の出力信号で、高周波調整部 2 j において行われた高周波信号の電力の調整の効果が周波数範囲の単位で保たれる。また、第 1 の実施形態の本変形例 3 と同様の変更を第 4 の実施形態における線形予測フィルタ部 2 k に加えてもよい。

[0118] (第 3 の実施形態の変形例 1)

第 3 の実施形態の音声符号化装置 1 3 におけるエンベロープ形状パラメータ算出部 1 n は、以下のような処理で実現することもできる。エンベロープ形状パラメータ算出部 1 n は、符号化フレーム内の SBR エンベロープのの各々について、次の数式 (33) に従ってエンベロープ形状パラメータ  $s(i)$  ( $0 \leq i < N_e$ ) を取得する。

[数33]

$$s(i) = 1 - \min\left(\frac{e(r)}{e(i)}\right)$$

ただし、

[数34]

$$\overline{e(i)}$$

は  $e(r)$  の SBR エンベロープ内での平均値であり、その算出方法は数式 (21) に従う。ただし、SBR エンベロープとは、 $b_i \leq r < b_{i+1}$  を満たす時間範囲を示す。また、 $\{b_i\}$  は、SBR 補助情報に情報として含まれている、SBR エンベロープの時間境界であり、任意の時間範囲、任意の周波数範囲の平均信号エネルギーを表す SBR エンベロープスケールファクタが対象とする時間範囲の境界である。また、 $\min(\cdot)$  は  $b_i \leq r < b_{i+1}$  の範囲における最小値を表す。従って、この場合には、エンベロープ形状パラメータ

$s(i)$  は、調整後の時間エンベロープ情報の SBR エンベロープ内での最小値と平均値の比率を指示するパラメータである。また、第 3 の実施形態の音声復号装置 23 におけるエンベロープ形状調整部 2s は、以下のような処理で実現することもできる。エンベロープ形状調整部 2s は、 $s(i)$  を用いて  $e(r)$  を調整し、調整後の時間エンベロープ情報  $e_{adj}(r)$  を取得する。調整の方法は次の数式 (35) 又は数式 (36) に従う。

[数35]

$$e_{adj}(r) = e(i) \left( 1 + s(i) \frac{e(i) - \min(e(r))}{e(r) - e(i)} \right)$$

[数36]

$$e_{adj}(r) = e(i) \left( 1 + s(i) \frac{e(i)}{e(r) - e(i)} \right)$$

数式 35 は、調整後の時間エンベロープ情報  $e_{adj}(r)$  の SBR エンベロープ内での最小値と平均値の比率が、エンベロープ形状パラメータ  $s(i)$  の値と等しくなるようエンベロープ形状を調整するものである。また、上記した第 3 の実施形態の本変形例 1 と同様の変更を第 4 の実施形態に加えてもよい。

[0119] (第 3 の実施形態の変形例 2)

時間エンベロープ変形部 2v は、数式 (28) に代わり、次の数式を利用することもできる。数式 (37) に示すとおり、 $e_{adj, scaled}(r)$  は、 $q_{adj}(k, r)$  と  $q_{envadj}(k, r)$  の SBR エンベロープ内での電力が等しくなるよう調整後の時間エンベロープ情報  $e_{adj}(r)$  の利得を制御したものである。また、数式 (38) に示すとおり、第 3 の実施形態の本変形例 2 では、 $e_{adj}(r)$  ではなく  $e_{adj, scaled}(r)$  を QMF 領域の信号  $q_{adj}(k, r)$  に乗算して  $q_{envadj}$

( $k, r$ ) を得る。従って、時間エンベロープ変形部 2 v は、SBR エンベロープ内での信号電力が時間エンベロープの変形の前後で等しくなるように QMF 領域の信号  $q_{adj}(k, r)$  の時間エンベロープの変形を行うことができる。ただし、SBR エンベロープとは、 $b_i \leq r < b_{i+1}$  を満たす時間範囲を示す。また、 $\{b_i\}$  は、SBR 補助情報に情報として含まれている、SBR エンベロープの時間境界であり、任意の時間範囲、任意の周波数範囲の平均信号エネルギーを表す SBR エンベロープスケールファクタが対象とする時間範囲の境界である。また、本発明の実施例中における用語 “SBR エンベロープ” は、“ISO/IEC 14496-3” に規定される “MPEG4 AAC” における用語 “SBR エンベロープ時間セグメント” に相当し、実施例全体を通して “SBR エンベロープ” は “SBR エンベロープ時間セグメント” と同一の内容を意味する。

[数37]

$$e_{adj,scaled}(r) = e_{adj}(r) \cdot \sqrt{\frac{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k, r)|^2}{\sum_{k=k_x}^{63} \sum_{r=b_i}^{b_{i+1}-1} |q_{adj}(k, r) \cdot e_{adj}(r)|^2}}$$

$$(k_x \leq k \leq 63, b_i \leq r < b_{i+1})$$

[数38]

$$q_{envadj}(k, r) = q_{adj}(k, r) \cdot e_{adj,scaled}(r)$$

$$(k_x \leq k \leq 63, b_i \leq r < b_{i+1})$$

また、上記した第 3 の実施形態の本変形例 2 と同様の変更を第 4 の実施形態に加えてもよい。

[0120] (第 3 の実施形態の変形例 3)

数式 ( 1 9 ) は下記の数式 ( 3 9 ) であつてもよい。

[数39]

$$e(r) = \sqrt{\frac{\sum_{k=0}^{b_{i+1}-1} \sum_{r=b_i, k=0}^{r=b_i, k=0} |q(k, r)|^2}{\sum_{k=0}^{b_{i+1}-1} |q(k, r)|^2}}$$

数式 ( 2 2 ) は下記の数式 ( 4 0 ) であつてもよい。

[数40]

$$e(r) = \sqrt{\frac{\sum_{k=0}^{b_{i+1}-1} \sum_{r=b_i, k=0}^{r=b_i, k=0} |q_{dec}(k, r)|^2}{\sum_{k=0}^{b_{i+1}-1} |q_{dec}(k, r)|^2}}$$

数式 ( 2 6 ) は下記の数式 ( 4 1 ) であつてもよい。

[数41]

$$e_{exp}(r) = \sqrt{\frac{\sum_{k=k_x}^{b_{i+1}-1} \sum_{r=b_i, k=k_x}^{r=b_i, k=k_x} |q_{exp}(k, r)|^2}{\sum_{k=k_x}^{b_{i+1}-1} |q_{exp}(k, r)|^2}}$$

数式 ( 3 9 ) 及び数式 ( 4 0 ) にしたがつた場合、時間エンベロープ情報  $e(r)$  は、QMFサブバンド信号ごとの電力をSBRエンベロープ内の平均電力で正規化し、さらに平方根をとつたものとなる。ただし、QMFサブバンド信号は、QMF領域信号において、同一の時間インデックス“ $r$ ”に対応する信号ベクトルであり、QMF領域における一つのサブバンドを意味する。また、本発明の実施形態全体において、用語“時間スロット”は“QMFサブバンド信号”と同一の内容を意味する。この場合、時

間エンベロープ情報  $e(r)$  は、各 QMF サブバンド サンプル へ乗算されるべきゲイン係数を意味することとなり、調整後の時間エンベロープ情報  $e_{adj}(r)$  も同様である。

[0121] (第 4 の実施形態の変形例 1)

第 4 の実施形態の変形例 1 の音声復号装置 24 a (不図示) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 24 a の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声復号装置 24 a を統括的に制御する。音声復号装置 24 a の通信装置は、音声符号化装置 11 又は音声符号化装置 13 から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 24 a は、機能的には、音声復号装置 24 のビットストリーム分離部 2 a 3 に代わり、ビットストリーム分離部 2 a 4 (不図示) を備え、さらに、補助情報変換部 2 w に代わり、時間エンベロープ補助情報生成部 2 y (不図示) を備える。ビットストリーム分離部 2 a 4 は、多重化ビットストリームを、SBR 補助情報と、符号化ビットストリームとに分離する。時間エンベロープ補助情報生成部 2 y は、符号化ビットストリームおよび SBR 補助情報に含まれる情報に基づいて、時間エンベロープ補助情報を生成する。

[0122] ある SBR エンベロープにおける時間エンベロープ補助情報の生成には、例えば当該 SBR エンベロープの時間幅 ( $b_{i+1} - b_i$ )、フレームクラス、逆フィルタの強度パラメータ、ノイズフロア、高周波電力の大きさ、高周波電力と低周波電力の比率、QMF 領域で表現された低周波信号を周波数方向に線形予測分析した結果の自己相関係数または予測ゲインなどを用いることができる。これらのパラメータの一つ、または複数の値に基づいて  $K(r)$  または  $s(i)$  を決定することで、時間エンベロープ補助情報を生成することができる。例えば SBR エンベロープの時間幅 ( $b_{i+1} - b_i$ ) が広いほど  $K(r)$  または  $s(i)$  が小さくなるよう、または SBR エンベロープの時間幅 ( $b_{i+1} - b_i$ ) が広いほど  $K(r)$  または  $s(i)$  が大きくなるよう ( $b_{i+1} -$

b<sub>j</sub>)に基づいてK(r)またはs(i)を決定することで、時間エンベロープ補助情報を生成することができる。また、同様の変更を第1の実施形態及び第3の実施形態に加えてもよい。

[0123] (第4の実施形態の変形例2)

第4の実施形態の変形例2の音声復号装置24b(図15参照)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置24bの内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声復号装置24bを統括的に制御する。音声復号装置24bの通信装置は、音声符号化装置11又は音声符号化装置13から出力される符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置24bは、図15に示すとおり、高周波調整部2jにかえて、一次高周波調整部2j1と二次高周波調整部2j2とを備える。

[0124] ここで、一次高周波調整部2j1は、“MPEG4 AAC”のSBRにおける“HF adjustment”ステップにある、高周波帯域のQMF領域の信号に対する時間方向の線形予測逆フィルタ処理、ゲインの調整及びノイズの重畳処理による調整を行う。このとき、一次高周波調整部2j1の出力信号は、“ISO/IEC 14496-3:2005”の“SBR tool”内、4.6.18.7.6節“Assembling HF signals”の記述内における信号W<sub>2</sub>に相当するものとなる。線形予測フィルタ部2k(又は、線形予測フィルタ部2k1)および時間エンベロープ変形部2vは、一次高周波調整部の出力信号を対象に時間エンベロープの変形を行う。二次高周波調整部2j2は、時間エンベロープ変形部2vから出力されたQMF領域の信号に対し、“MPEG4 AAC”のSBRにおける“HF adjustment”ステップにある正弦波の付加処理を行う。二次高周波調整部の処理は、“ISO/IEC 14496-3:2005”の“SBR tool”内、4.6.18.7.6節“Assembling HF signals”の記述内における、信号W<sub>2</sub>から信号Yを生成する処理において、信号W<sub>2</sub>を時間エンベロープ変形部2vの出力信号に置き換えた処理に相当する。

[0125] なお、上記の説明では正弦波付加処理のみを二次高周波調整部2j2の処

理としたが、“HF adjustment”ステップにある処理のいずれかを二次高周波調整部 2 j 2 の処理としてよい。また、同様な変形は、第 1 の実施形態、第 2 の実施形態、第 3 の実施形態に加えてもよい。この際、第 1 の実施形態および第 2 の実施形態は線形予測フィルタ部（線形予測フィルタ部 2 k, 2 k 1）を備え、時間エンベロープ変形部を備えないため、一次高周波調整部 2 j 1 の出力信号に対して線形予測フィルタ部での処理を行った後、線形予測フィルタ部の出力信号を対象に二次高周波調整部 2 j 2 での処理を行う。

[0126] また、第 3 の実施形態は時間エンベロープ変形部 2 v を備え、線形予測フィルタ部を備えないため、一次高周波調整部 2 j 1 の出力信号に対して時間エンベロープ変形部 2 v での処理を行った後、時間エンベロープ変形部 2 v の出力信号を対象に二次高周波調整部での処理を行う。

[0127] また、第 4 の実施形態の音声復号装置（音声復号装置 2 4, 2 4 a, 2 4 b）において、線形予測フィルタ部 2 k と時間エンベロープ変形部 2 v の処理の順序は逆でもよい。すなわち、高周波調整部 2 j または一次高周波調整部 2 j 1 の出力信号に対して、時間エンベロープ変形部 2 v の処理を先に行い、次に、時間エンベロープ変形部 2 v の出力信号に対して線形予測フィルタ部 2 k の処理を行ってもよい。

[0128] また、時間エンベロープ補助情報は線形予測フィルタ部 2 k または時間エンベロープ変形部 2 v での処理を行うか否かを指示する 2 値の制御情報を含み、この制御情報が線形予測フィルタ部 2 k または時間エンベロープ変形部 2 v での処理を行うことを指示している場合に限って、フィルタ強度パラメータ  $K(r)$ 、エンベロープ形状パラメータ  $s(i)$ 、または  $K(r)$  と  $s(i)$  の双方を決定するパラメータである  $X(r)$  のいずれか一つ以上をさらに情報として含む形式をとってもよい。

[0129] （第 4 の実施形態の変形例 3）

第 4 の実施形態の変形例 3 の音声復号装置 2 4 c（図 16 参照）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 c の内蔵メモリに格納された所定のコン

ピュータプログラム（例えば、図 17 のフローチャートに示す処理を行うためのコンピュータプログラム）を RAM にロードして実行することによって音声復号装置 24c を統括的に制御する。音声復号装置 24c の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 24c は、図 16 に示すとおり、高周波調整部 2j にかえて、一次高周波調整部 2j3 と二次高周波調整部 2j4 とを備え、さらに線形予測フィルタ部 2k と時間エンベロープ変形部 2v に代えて個別信号成分調整部 2z1, 2z2, 2z3 を備える（個別信号成分調整部は、時間エンベロープ変形手段に相当する）。

[0130] 一次高周波調整部 2j3 は、高周波帯域の QMF 領域の信号を、複写信号成分として出力する。一次高周波調整部 2j3 は、高周波帯域の QMF 領域の信号に対して、ビットストリーム分離部 2a3 から与えられる SBR 補助情報を利用して時間方向の線形予測逆フィルタ処理及びゲインの調整（周波数特性の調整）の少なくとも一方を行った信号を複写信号成分として出力してもよい。さらに、一次高周波調整部 2j3 は、ビットストリーム分離部 2a3 から与えられる SBR 補助情報を利用してノイズ信号成分および正弦波信号成分を生成し、複写信号成分、ノイズ信号成分および正弦波信号成分を分離された形で各々出力する（ステップ Sg1 の処理）。ノイズ信号成分および正弦波信号成分は、SBR 補助情報の内容に依存し、生成されない場合があってもよい。

[0131] 個別信号成分調整部 2z1, 2z2, 2z3 は、前記一次高周波調整手段の出力に含まれる複数の信号成分の各々に対し処理を行う（ステップ Sg2 の処理）。個別信号成分調整部 2z1, 2z2, 2z3 における処理は、線形予測フィルタ部 2k と同様の、フィルタ強度調整部 2f から得られた線形予測係数を用いた周波数方向の線形予測合成フィルタ処理であってもよい（処理 1）。また、個別信号成分調整部 2z1, 2z2, 2z3 における処理は、時間エンベロープ変形部 2v と同様の、エンベロープ形状調整部 2s から得られた時間エンベロープを用いて各 QMF サブバンドサンプルヘゲイン

係数を乗算する処理であつてもよい(処理2)。また、個別信号成分調整部221、222、223における処理は、入力信号に対して線形予測フィルタ221、222、223と同様の、フィルタ強度調整部24から得られた線形予測係数を用いた周波数方向の線形予測合成フィルタ処理を行った後、その出力信号に対して線形予測フィルタ221、222、223における処理は、入力信号に対して線形予測フィルタ221、222、223と同様の、エンベロープ形状調整部25から得られた時間エンベロープを用いて各QMFサブバンドサンプリング係数を乗算する処理を行った後、その出力信号に対してさらに線形予測フィルタ221、222、223と同様の、フィルタ強度調整部24から得られた線形予測係数を用いた周波数方向の線形予測合成フィルタ処理を行うことであつてもよい(処理4)。また、個別信号成分調整部221、222、223は入力信号に対して時間エンベロープ変形部26と同様の、エンベロープ形状調整部25から得られた時間エンベロープを用いて各QMFサブバンドサンプリング係数を乗算する処理を行うことであつてもよい(処理3)。また、個別信号成分調整部221、222、223における処理は、入力信号に対して線形予測フィルタ221、222、223と同様の、エンベロープ形状調整部25から得られた時間エンベロープを用いて各QMFサブバンドサンプリング係数を乗算する処理を行った後、その出力信号に対して線形予測フィルタ221、222、223における処理は、入力信号の時間エンベロープを3における処理は、処理1〜5以外の方法で入力信号の時間エンベロープを変形するための何らかの処理を加えるものであつてもよい(処理6)。また、個別信号成分調整部221、222、223における処理は、処理1〜6のうちの複数の処理を任意の順序で組み合わせた処理であつてもよい(処理7)。

[0132] 個別信号成分調整部221、222、223における処理は互いに同じでもよいが、個別信号成分調整部221、222、223は、一次高周波調整手段の出力に含まれる複数の信号成分の各々に対し互いに異なる方法で時間エンベロープの変形を行つてもよい。例えば個別信号成分調整部221は入力された複写信号に対して処理2を行い、個別信号成分調整部222は入力されたノイズ信号成分に対して処理3を行い、個別信号成分調整部223は入力された正弦波信号に対して処理5を行うといったように、複写信号、ノイ

ズ信号、正弦波信号の各々に対して互いに異なる処理を行ってよい。また、この際、フィルタ強度調整部 2 f とエンベロープ形状調整部 2 s は、個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 の各々に対して互いに同じ線形予測係数や時間エンベロープを送信してもよいが、互いに異なる線形予測係数や時間エンベロープを送信してもよく、また個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 のいずれか 2 つ以上に対して同一の線形予測係数や時間エンベロープを送信してもよい。個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 の 1 つ以上は、時間エンベロープ変形処理を行わず、入力信号をそのまま出力するもの（処理 5）であってもよいため、個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 は全体として、一次高周波調整部 2 j 3 から出力された複数の信号成分の少なくとも一つに対し時間エンベロープ処理を行うものである（個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 の全てが処理 5 である場合は、いずれの信号成分に対しても時間エンベロープ変形処理が行われないため、本発明の効果を有さない）。

[0133] 個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 のそれぞれにおける処理は、処理 1 から処理 7 のいずれかに固定されていてもよいが、外部から与えられる制御情報に基づいて、処理 1 から処理 7 のいずれを行うかが動的に決定されてもよい。この際、上記制御情報は多重化ビットストリームに含まれることが望ましい。また、上記制御情報は、特定の SBR エンベロープ時間セグメント、符号化フレーム、またはその他の時間範囲において処理 1 から処理 7 のいずれを行うかを指示するものであってもよく、また、制御の時間範囲を特定せず、処理 1 から処理 7 のいずれを行うかを指示するものであってもよい。

[0134] 二次高周波調整部 2 j 4 は、個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 から出力された処理後の信号成分を足し合わせ、係数加算部へ出力する（ステップ S g 3 の処理）。また、二次高周波調整部 2 j 4 は、複写信号成分に対して、ビットストリーム分離部 2 a 3 から与えられる SBR 補助情報を利用して時間方向の線形予測逆フィルタ処理及びゲインの調整（周波数特性の

調整)の少なくとも一方を行ってもよい。

[0135] 個別信号成分調整部は  $2 \geq 1$ ,  $2 \geq 2$ ,  $2 \geq 3$  は互いに協調して動作し、処理 1 ~ 7 のいずれかの処理を行った後の 2 つ以上の信号成分を互いに足し合わせ、足し合わされた信号に対してさらに処理 1 ~ 7 のいずれかの処理を加えて途中段階の出力信号を生成してもよい。この際には、二次高周波調整部  $2 \geq 4$  は、前記途中段階の出力信号と、前記途中段階の出力信号にまだ足しあわされていない信号成分を足し合わせ、係数加算部へ出力する。具体的には、複写信号成分に処理 5 を行い、雑音成分に処理 1 を加えた後にこれら 2 つの信号成分を互いに足し合わせ、足しあわされた信号に対してさらに処理 2 を加えて途中段階の出力信号を生成することが望ましい。この際には、二次高周波調整部  $2 \geq 4$  は、前記途中段階の出力信号に正弦波信号成分を足し合わせ、係数加算部へ出力する。

[0136] 一次高周波調整部  $2 \geq 3$  は、複写信号成分、ノイズ信号成分、正弦波信号成分の 3 つの信号成分に限らず、任意の複数の信号成分を互いに分離された形で出力してもよい。この場合の信号成分は、複写信号成分、ノイズ信号成分、正弦波信号成分のうち 2 つ以上を足し合わせたものであってもよい。また、複写信号成分、ノイズ信号成分、正弦波信号成分のいずれかを帯域分割した信号であってもよい。信号成分の数は 3 以外であってもよく、この場合には個別信号成分調整部の数は 3 以外であってもよい。

[0137] SBR によって生成される高周波信号は、低周波帯域を高周波帯域に複写して得られた複写信号成分と、ノイズ信号、正弦波信号の 3 つの要素から構成される。複写信号、ノイズ信号、正弦波信号の各々は、互いに異なる時間エンベロープを持つため、本変形例の個別信号成分調整部が行うように、各々の信号成分に対して互いに異なる方法で時間エンベロープの変形を行うことにより、本発明の他の実施例と比較し、復号信号の主観品質をさらに向上させることができる。特に、ノイズ信号は一般に平坦な時間エンベロープを持ち、複写信号は低周波帯域の信号に近い時間エンベロープを持つため、これらを分離して扱い、互いに異なる処理を加えることにより、複写信号とノ

イズ信号の時間エンベロープを独立に制御することが可能となり、これは復号信号の主観品質向上に有効である。具体的には、ノイズ信号に対しては時間エンベロープを変形させる処理（処理 3 または処理 4）を行い、複写信号に対しては、ノイズ信号に対するものとは異なる処理（処理 1 または処理 2）を行い、さらに、正弦波信号に対しては、処理 5 を行う（すなわち、時間エンベロープ変形処理を行わない）ことが好ましい。または、ノイズ信号に対しては時間エンベロープの変形処理（処理 3 または処理 4）を行い、複写信号と正弦波信号に対しては、処理 5 を行う（すなわち、時間エンベロープ変形処理を行わない）ことが好ましい。

[0138] （第 1 の実施形態の変形例 4）

第 1 の実施形態の変形例 4 の音声符号化装置 1 1 b（図 4 4）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 1 b の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 1 b を統括的に制御する。音声符号化装置 1 1 b の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 1 1 b は、音声符号化装置 1 1 の線形予測分析部 1 e にかえて線形予測分析部 1 e 1 を備え、時間スロット選択部 1 p をさらに備える。

[0139] 時間スロット選択部 1 p は、周波数変換部 1 a から QMF 領域の信号を受け取り、線形予測分析部 1 e 1 での線形予測分析処理を施す時間スロットを選択する。線形予測分析部 1 e 1 は、時間スロット選択部 1 p より通知された選択結果に基づき、選択された時間スロットの QMF 領域信号を線形予測分析部 1 e と同様に線形予測分析し、高周波線形予測係数、低周波線形予測係数のうち少なくともひとつを取得する。フィルタ強度パラメータ算出部 1 f は、線形予測分析部 1 e 1 において得られた、時間スロット選択部 1 p で選択された時間スロットの線形予測係数を用いてフィルタ強度パラメータを算出する。時間スロット選択部 1 p での時間スロットの選択では、例えば後に記

載の本変形例の復号装置 2 1 a における時間スロット選択部 3 a と同様の高周波成分の Q M F 領域信号の信号電力を用いた選択方法のうち少なくともひとつを用いてもよい。その際、時間スロット選択部 1 p における高周波成分の Q M F 領域信号は、周波数変換部 1 a から受け取る Q M F 領域の信号のうち、S B R 符号化部 1 d において符号化される周波数成分であることが望ましい。時間スロットの選択方法は、前記の方法を少なくともひとつ用いてもよく、さらには前記とは異なる方法を少なくともひとつ用いてもよく、さらにはそれらを組み合わせて用いてもよい。

[0140] 第 1 の実施形態の変形例 4 の音声復号装置 2 1 a (図 1 8 参照) は、物理的には図示しない C P U、R O M、R A M 及び通信装置等を備え、この C P U は、R O M 等の音声復号装置 2 1 a の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 1 9 のフローチャートに示す処理を行うためのコンピュータプログラム) を R A M にロードして実行することによって音声復号装置 2 1 a を統括的に制御する。音声復号装置 2 1 a の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 1 a は、図 1 8 に示すとおり、音声復号装置 2 1 の低周波線形予測分析部 2 d、信号変化検出部 2 e、高周波線形予測分析部 2 h、及び線形予測逆フィルタ部 2 i、及び線形予測フィルタ部 2 k にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、及び線形予測フィルタ部 2 k 3 を備え、時間スロット選択部 3 a をさらに備える。

[0141] 時間スロット選択部 3 a は、高周波生成部 2 g にて生成された時間スロット  $r$  の高周波成分の Q M F 領域の信号  $q_{exp}(k, r)$  に対して、線形予測フィルタ部 2 k において線形予測合成フィルタ処理を施すか否かを判断し、線形予測合成フィルタ処理を施す時間スロットを選択する (ステップ S h 1 の処理)。時間スロット選択部 3 a は、時間スロットの選択結果を、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、線形予測フィルタ部 2 k 3 に通知する。低周

波線形予測分析部 2 d 1 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロット  $r_1$  の QMF 領域信号を、低周波線形予測分析部 2 d と同様に線形予測分析し、低周波線形予測係数を取得する（ステップ S h 2 の処理）。信号変化検出部 2 e 1 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロットの QMF 領域信号の時間変化を、信号変化検出部 2 e と同様に検出し、検出結果  $T(r_1)$  を出力する。

[0142] フィルタ強度調整部 2 f では、低周波線形予測分析部 2 d 1 において得られた、時間スロット選択部 3 a で選択された時間スロットの低周波線形予測係数に対してフィルタ強度調整を行い、調整された線形予測係数  $a_{dec}(n, r_1)$  を得る。高周波線形予測分析部 2 h 1 では、高周波生成部 2 g によって生成された高周波成分の QMF 領域信号を、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロット  $r_1$  に関して、高周波線形予測分析部 2 k と同様に、周波数方向に線形予測分析し、高周波線形予測係数  $a_{exp}(n, r_1)$  を取得する（ステップ S h 3 の処理）。線形予測逆フィルタ部 2 i 1 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロット  $r_1$  の高周波成分の QMF 領域の信号  $q_{exp}(k, r)$  を、線形予測逆フィルタ部 2 i と同様に周波数方向に  $a_{exp}(n, r_1)$  を係数とする線形予測逆フィルタ処理を行う（ステップ S h 4 の処理）。

[0143] 線形予測フィルタ部 2 k 3 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロット  $r_1$  の高周波調整部 2 j から出力された高周波成分の QMF 領域の信号  $q_{adj}(k, r_1)$  に対し、線形予測フィルタ部 2 k と同様に、フィルタ強度調整部 2 f から得られた  $a_{adj}(n, r_1)$  を用いて、周波数方向に線形予測合成フィルタ処理を行う（ステップ S h 5 の処理）。また、変形例 3 に記載の線形予測フィルタ部 2 k への変更を、線形予測フィルタ部 2 k 3 に加えてもよい。時間スロット選択部 3 a での線形予測合成フィルタ処理を施す時間スロットの選択では、例えば高周波成分の QMF 領域信号  $q_{exp}(k, r)$  の信号電力が所定の値  $P_{exp, Th}$  よりも大きい時

間スロット  $r$  をひとつ以上選択してもよい。 $q_{\text{exp}}(k, r)$  の信号電力は次の数式で求めることが望ましい。

[数42]

$$P_{\text{exp}}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{\text{exp}}(k, r)|^2$$

ただし、 $M$  は高周波生成部 2 g によって生成される高周波成分の下限周波数  $k_x$  より高い周波数の範囲を表す値であり、さらには高周波生成部 2 g によって生成される高周波成分の周波数範囲を  $k_x \leq k < k_x + M$  のように表してもよい。また、所定の値  $P_{\text{exp, Th}}$  は時間スロット  $r$  を含む所定の時間幅の  $P_{\text{exp}}(r)$  の平均値でもよい。さらに所定の時間幅は SBR エンベロープでもよい。

[0144] また、高周波成分の QMF 領域信号の信号電力がピークになる時間スロットが含まれるように選択してもよい。信号電力のピークは、例えば信号電力の移動平均値

[数43]

$$P_{\text{exp, MA}}(r)$$

について

[数44]

$$P_{\text{exp, MA}}(r+1) - P_{\text{exp, MA}}(r)$$

が正の値から負の値に変わる時間スロット  $r$  の高周波成分の QMF 領域の信号電力をピークとしてもよい。信号電力の移動平均値

[数45]

$$P_{\text{exp, MA}}(r)$$

は、例えば次の式で求めることができる。

[数46]

$$P_{exp,MA}(r) = \frac{1}{c} \sum_{r'=r-\frac{c}{2}}^{r+\frac{c}{2}-1} P_{exp}(r')$$

ただし、 $c$ は平均値を求める範囲を定める所定の値である。また信号電力のピークは、前記の方法で求めてもよく、異なる方法により求めてもよい。

[0145] さらに、高周波成分のQMF領域信号の信号電力の変動が小さい定常状態から変動の大きい過渡状態になるまでの時間幅 $t$ が所定の値 $t_{th}$ よりも小さく、当該時間幅に含まれる時間スロットを少なくともひとつ選択してもよい。

さらに、高周波成分のQMF領域信号の信号電力の変動が大きい過渡状態から変動の小さい定常状態になるまでの時間幅 $t$ が所定の値 $t_{th}$ よりも小さく、当該時間幅に含まれる時間スロットを少なくともひとつ選択してもよい。

$|P_{exp}(r+1) - P_{exp}(r)|$ が所定の値よりも小さい（または、所定の値と等しいまたは小さい）時間スロット $r$ を前記定常状態とし、 $|P_{exp}(r+1) - P_{exp}(r)|$ が所定の値と等しいまたは大きい（または、所定の値よりも大きい）時間スロット $r$ を前記過渡状態としてもよく、 $|P_{exp,MA}(r+1) - P_{exp,MA}(r)|$ が所定の値よりも小さい（または、所定の値と等しいまたは小さい）時間スロット $r$ を前記定常状態とし、 $|P_{exp,MA}(r+1) - P_{exp,MA}(r)|$ が所定の値と等しいまたは大きい（または、所定の値よりも大きい）時間スロット $r$ を前記過渡状態としてもよい。また過渡状態、定常状態は前記の方法で定義してもよく、異なる方法で定義してもよい。時間スロットの選択方法は、前記の方法を少なくともひとつ用いてもよく、さらには前記とは異なる方法を少なくともひとつ用いてもよく、さらにはそれらを組み合わせてもよい。

[0146] （第1の実施形態の変形例5）

第1の実施形態の変形例5の音声符号化装置11c(図45)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは

、ROM等の音声符号化装置 11c の内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声符号化装置 11c を統括的に制御する。音声符号化装置 11c の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 11c は、変形例 4 の音声符号化装置 11b の時間スロット選択部 1p、及びビットストリーム多重化部 1g にかえて、時間スロット選択部 1p1、及びビットストリーム多重化部 1g4 を備える。

[0147] 時間スロット選択部 1p1 は、第1の実施形態の変形例 4 に記載の時間スロット選択部 1p と同様に時間スロットを選択し、時間スロット選択情報をビットストリーム多重化部 1g4 へ送る。ビットストリーム多重化部 1g4 は、コアコーデック符号化部 1c によって算出された符号化ビットストリームと、SBR符号化部 1d によって算出されたSBR補助情報と、フィルタ強度パラメータ算出部 1f によって算出されたフィルタ強度パラメータとを、ビットストリーム多重化部 1g と同様に多重化し、さらに時間スロット選択部 1p1 から受け取った時間スロット選択情報とを多重化し、多重化ビットストリームを、音声符号化装置 11c の通信装置を介して出力する。前記時間スロット選択情報は、後に記載の音声復号装置 21b における時間スロット選択部 3a1 が受け取る時間スロット選択情報であり、例えば選択する時間スロットのインデックス r1 を含んでいてもよい。さらに、例えば時間スロット選択部 3a1 の時間スロット選択方法に利用されるパラメータでもよい。第1の実施形態の変形例 5 の音声復号装置 21b (図 20 参照) は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置 21b の内蔵メモリに格納された所定のコンピュータプログラム(例えば、図 21 のフローチャートに示す処理を行うためのコンピュータプログラム)をRAMにロードして実行することによって音声復号装置 21b を統括的に制御する。音声復号装置 21b の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信

号を外部に出力する。

[0148] 音声復号装置 2 1 b は、図 2 0 に示すとおり、変形例 4 の音声復号装置 2 1 a のビットストリーム分離部 2 a、及び時間スロット選択部 3 a にかえて、ビットストリーム分離部 2 a 5、及び時間スロット選択部 3 a 1 を備え、時間スロット選択部 3 a 1 に時間スロット選択情報が入力される。ビットストリーム分離部 2 a 5 では、多重化ビットストリームを、ビットストリーム分離部 2 a と同様に、フィルタ強度パラメータと、SBR 補助情報と、符号化ビットストリームとに分離し、時間スロット選択情報をさらに分離する。時間スロット選択部 3 a 1 では、ビットストリーム分離部 2 a 5 から送られた時間スロット選択情報に基づいて時間スロットを選択する（ステップ S i 1 の処理）。時間スロット選択情報は、時間スロットの選択に用いる情報であり、例えば選択する時間スロットのインデックス  $r_1$  を含んでいてもよい。さらに、例えば変形例 4 に記載の時間スロット選択方法に利用されるパラメータでもよい。この場合、時間スロット選択部 3 a 1 には、時間スロット選択情報に加えて、図示されていないが高周波信号生成部 2 g にて生成された高周波成分の QMF 領域信号も入力される。前記パラメータは、例えば前記時間スロットの選択のために用いる所定の値（例えば、 $P_{exp, Th}$ 、 $t_{Th}$  など）でもよい。

[0149] （第 1 の実施形態の変形例 6）

第 1 の実施形態の変形例 6 の音声符号化装置 1 1 d（不図示）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 1 d の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 1 d を統括的に制御する。音声符号化装置 1 1 d の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 1 1 d は、変形例 1 の音声符号化装置 1 1 a の短時間電力算出部 1 i にかえて、図示しない短時間電力算出部 1 i 1 を備え、時間スロット選択部 1 p 2 をさらに備える。

[0150] 時間スロット選択部 1 p 2 は、周波数変換部 1 a から Q M F 領域の信号を受け取り、短時間電力算出部 1 i での短時間電力算出処理を施す時間区間に対応する時間スロットを選択する。短時間電力算出部 1 i 1 は、時間スロット選択部 1 p 2 より通知された選択結果に基づき、選択された時間スロットに対応する時間区間の短時間電力を、変形例 1 の音声符号化装置 1 1 a の短時間電力算出部 1 i と同様に算出する。

[0151] (第 1 の実施形態の変形例 7)

第 1 の実施形態の変形例 7 の音声符号化装置 1 1 e (不図示) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 1 e の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 1 e を統括的に制御する。音声符号化装置 1 1 e の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 1 1 e は、変形例 6 の音声符号化装置 1 1 d の時間スロット選択部 1 p 2 にかえて、図示しない時間スロット選択部 1 p 3 を備える。さらに、ビットストリーム多重化部 1 g 1 にかえて、時間スロット選択部 1 p 3 からの出力をさらに受けるビットストリーム多重化部を備える。時間スロット選択部 1 p 3 は、第 1 の実施形態の変形例 6 に記載の時間スロット選択部 1 p 2 と同様に時間スロットを選択し、時間スロット選択情報をビットストリーム多重化部へ送る。

[0152] (第 1 の実施形態の変形例 8)

第 1 の実施形態の変形例 8 の音声符号化装置 (不図示) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の変形例 8 の音声符号化装置の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって変形例 8 の音声符号化装置を統括的に制御する。変形例 8 の音声符号化装置の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。変形例 8 の音声符号化装置は、変形

例 2 に記載の音声符号化装置に加え、時間スロット選択部 1 p をさらに備える。

[0153] 第 1 の実施形態の変形例 8 の音声復号装置（不図示）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の変形例 8 の音声復号装置の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって変形例 8 の音声復号装置を統括的に制御する。変形例 8 の音声復号装置の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。変形例 8 の音声復号装置は、変形例 2 に記載の音声復号装置の低周波線形予測分析部 2 d、信号変化検出部 2 e、高周波線形予測分析部 2 h、及び線形予測逆フィルタ部 2 i、及び線形予測フィルタ部 2 k にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、及び線形予測フィルタ部 2 k 3 を備え、時間スロット選択部 3 a をさらに備える。

[0154] （第 1 の実施形態の変形例 9）

第 1 の実施形態の変形例 9 の音声符号化装置（不図示）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の変形例 9 の音声符号化装置の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって変形例 9 の音声符号化装置を統括的に制御する。変形例 9 の音声符号化装置の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。変形例 9 の音声符号化装置は、変形例 8 に記載の音声符号化装置の時間スロット選択部 1 p にかえて、時間スロット選択部 1 p 1 を備える。さらに、変形例 8 に記載のビットストリーム多重化部にかえて、変形例 8 に記載のビットストリーム多重化部への入力に加えて時間スロット選択部 1 p 1 からの出力をさらに受けるビットストリーム多重化部を備える。

[0155] 第 1 の実施形態の変形例 9 の音声復号装置（不図示）は、物理的には図示

しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の変形例9の音声復号装置の内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって変形例9の音声復号装置を統括的に制御する。変形例9の音声復号装置の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。変形例9の音声復号装置は、変形例8に記載の音声復号装置の時間スロット選択部3aにかえて、時間スロット選択部3a1を備える。さらに、ビットストリーム分離部2aにかえて、ビットストリーム分離部2a5のフィルタ強度パラメータにかえて前記変形例2に記載の $a_D(n, r)$ を分離するビットストリーム分離部を備える。

[0156] (第2の実施形態の変形例1)

第2の実施形態の変形例1の音声符号化装置12a(図46)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声符号化装置12aの内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声符号化装置12aを統括的に制御する。音声符号化装置12aの通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置12aは、音声符号化装置12の線形予測分析部1eにかえて、線形予測分析部1e1を備え、時間スロット選択部1pをさらに備える。

[0157] 第2の実施形態の変形例1の音声復号装置22a(図22参照)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置22aの内蔵メモリに格納された所定のコンピュータプログラム(例えば、図23のフローチャートに示す処理を行うためのコンピュータプログラム)をRAMにロードして実行することによって音声復号装置22aを統括的に制御する。音声復号装置22aの通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置22aは、図22に示すとおり、第2の実

施形態の音声復号装置 2 2 の高周波線形予測分析部 2 h、線形予測逆フィルタ部 2 i、線形予測フィルタ部 2 k 1、及び線形予測補間・補外部 2 p にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、線形予測フィルタ部 2 k 2、及び線形予測補間・補外部 2 p 1 を備え、時間スロット選択部 3 a をさらに備える。

[0158] 時間スロット選択部 3 a は、時間スロットの選択結果を、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、線形予測フィルタ部 2 k 2、線形予測係数補間・補外部 2 p 1 に通知する。線形予測係数補間・補外部 2 p 1 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロットであり線形予測係数の伝送されていない時間スロット  $r$  1 に対応する  $a_H(n, r)$  を、線形予測係数補間・補外部 2 p と同様に、補間又は補外により取得する（ステップ S j 1 の処理）。線形予測フィルタ部 2 k 2 では、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロット  $r$  1 に関して、高周波調整部 2 j から出力された  $q_{adj}(n, r 1)$  に対し、線形予測係数補間・補外部 2 p 1 から得られた、補間又は補外された  $a_H(n, r 1)$  を用いて、線形予測フィルタ部 2 k 1 と同様に、周波数方向に線形予測合成フィルタ処理を行う（ステップ S j 2 の処理）。また、第 1 の実施形態の変形例 3 に記載の線形予測フィルタ部 2 k への変更を、線形予測フィルタ部 2 k 2 に加えてもよい。

[0159] （第 2 の実施形態の変形例 2）

第 2 の実施形態の変形例 2 の音声符号化装置 1 2 b（図 4 7）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 2 b の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 1 b を統括的に制御する。音声符号化装置 1 2 b の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 1 2 b は、変形例 1 の音声

符号化装置 12a の時間スロット選択部 1p、及びビットストリーム多重化部 1g2 にかえて、時間スロット選択部 1p1、及びビットストリーム多重化部 1g5 を備える。ビットストリーム多重化部 1g5 は、ビットストリーム多重化部 1g2 と同様に、コアコーデック符号化部 1c で算出された符号化ビットストリームと、SBR 符号化部 1d で算出された SBR 補助情報と、線形予測係数量子化部 1k から与えられた量子化後の線形予測係数に対応する時間スロットのインデックスとを多重化し、さらに時間スロット選択部 1p1 から受け取る時間スロット選択情報をビットストリームに多重化し、多重化ビットストリームを、音声符号化装置 12b の通信装置を介して出力する。

[0160] 第 2 の実施形態の変形例 2 の音声復号装置 22b (図 24 参照) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 22b の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 25 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 22b を統括的に制御する。音声復号装置 22b の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 22b は、図 24 に示すとおり、変形例 1 に記載の音声復号装置 22a のビットストリーム分離部 2a1、及び時間スロット選択部 3a、にかえて、ビットストリーム分離部 2a6、及び時間スロット選択部 3a1 を備え、時間スロット選択部 3a1 に時間スロット選択情報が入力される。ビットストリーム分離部 2a6 では、ビットストリーム分離部 2a1 と同様に、多重化ビットストリームを、量子化された  $a_H(n, r_i)$  と、これに対応する時間スロットのインデックス  $r_i$  と、SBR 補助情報と、符号化ビットストリームとに分離し、時間スロット選択情報をさらに分離する。

[0161] (第 3 の実施形態の変形例 4)

第 3 の実施形態の変形例 1 に記載の

[数47]  $\overline{e(t)}$

は、 $e(r)$  のSBRエンベロープ内の平均値であつてもよく、さらに別に定める値であつてもよい。

[0162] (第3の実施形態の変形例5)

エンベロープ形状調整部2sは、前記第3の実施形態の変形例3に記載のとおり、調整後の時間エンベロープ $e_{adj}(r)$ が例えば数式(28)、数式(37)及び(38)のとおり、QMFサブバンドサンプルへ乗算されるゲイン係数であることを鑑み、 $e_{adj}(r)$ を所定の値 $e_{adj,th}(r)$ により以下のうに制限することが望ましい。

[数48]

$$e_{adj}(r) \geq e_{adj,th}$$

[0163] (第4の実施形態)

第4の実施形態の音声符号化装置14(図48)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声符号化装置14の内蔵メモリに格納された所定のコンピュータプログラムのRAMにロードして実行することによって音声符号化装置14を統括的に制御する。音声符号化装置14の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置14は、第1の実施形態の変形例4の音声符号化装置11bのビットストリーム多重化部1gにかえて、ビットストリーム多重化部1g7を備え、さらに音声符号化装置13の時間エンベロープ算出部1m、及びエンベロープパラメータ算出部1nを備える。

[0164]

ビットストリーム多重化部1g7は、ビットストリーム多重化部1gと同様に、コアコーディング符号化部1cによって算出された符号化ビットストリームと、SBR符号化部1dによって算出されたSBR補助情報とを多重化し、さらに、フィルタ強度パラメータ算出部によって算出されたフィルタ強

度パラメータと、エンベロープ形状パラメータ算出部 1 n によって算出されたエンベロープ形状パラメータとを時間エンベロープ補助情報に変換して多重化し、多重化ビットストリーム（符号化された多重化ビットストリーム）を、音声符号化装置 1 4 の通信装置を介して出力する。

[0165] （第 4 の実施形態の変形例 4）

第 4 の実施形態の変形例 4 の音声符号化装置 1 4 a（図 4 9）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声符号化装置 1 4 a の内蔵メモリに格納された所定のコンピュータプログラムを RAM にロードして実行することによって音声符号化装置 1 4 a を統括的に制御する。音声符号化装置 1 4 a の通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置 1 4 a は、第 4 の実施形態の音声符号化装置 1 4 の線形予測分析部 1 e にかえて、線形予測分析部 1 e 1 を備え、時間スロット選択部 1 p をさらに備える。

[0166] 第 4 の実施形態の変形例 4 の音声復号装置 2 4 d（図 2 6 参照）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 d の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図 2 7 のフローチャートに示す処理を行うためのコンピュータプログラム）を RAM にロードして実行することによって音声復号装置 2 4 d を統括的に制御する。音声復号装置 2 4 d の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 4 d は、図 2 6 に示すとおり、音声復号装置 2 4 の低周波線形予測分析部 2 d、信号変化検出部 2 e、高周波線形予測分析部 2 h、及び線形予測逆フィルタ部 2 i、及び線形予測フィルタ部 2 k にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、及び線形予測フィルタ部 2 k 3 を備え、時間スロット選択部 3 a をさらに備える。時間エンベロープ変形部 2 v は、線形予測フィルタ部 2 k 3 から得られた QMF 領域の信号を

、エンベロープ形状調整部 2s から得られた時間エンベロープ情報を用いて、第 3 の実施形態、第 4 の実施形態、及びそれらの変形例の時間エンベロープ変形部 2v と同様に変形する (ステップ S k 1 の処理)。

[0167] (第 4 の実施形態の変形例 5)

第 4 の実施形態の変形例 5 の音声復号装置 24e (図 28 参照) は、物理

的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 24e の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 29 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって

音声復号装置 24e を統括的に制御する。音声復号装置 24e の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号

を外部に出力する。音声復号装置 24e は、図 28 に示すとおり、変形例 5

においては、第 1 の実施形態と同様に第 4 の実施形態の全体を通して省略可

能である、変形例 4 に記載の音声復号装置 24d の高周波線形予測分析部 2

h1 と、線形予測逆フィルタ部 2i1 を省略し、音声復号装置 24d の時間

スロット選択部 3a、及び時間エンベロープ変形部 2v にかえて、時間スロ

ット選択部 3a2、及び時間エンベロープ変形部 2v1 を備える。さらに、

第 4 の実施形態の全体を通して処理順序を入れ替え可能である線形予測フイ

ルタ部 2k3 の線形予測合成フィルタ処理と時間エンベロープ変形部 2v1

での時間エンベロープの変形処理の順序を入れ替える。

[0168] 時間エンベロープ変形部 2v1 は、時間エンベロープ変形部 2v と同様に

、高周波調整部 2j から得られた  $q_{adj}(k, r)$  をエンベロープ形状調整部 2

s から得られた  $e_{adj}(r)$  を用いて変形し、時間エンベロープが変形された  $Q$

MF 領域の信号  $q_{envadj}(k, r)$  を取得する。さらに、時間エンベロープ変形

処理時に得られたパラメータ、または少なくとも時間エンベロープ変形処理

時に得られたパラメータを用いて算出したパラメータを時間スロット選択情報

報として、時間スロット選択部 3a2 に通知する。時間スロット選択情報と

しては、数式 (22)、数式 (40) の  $e(r)$  またはその算出過程にて平

方根演算をしない  $|e(r)|^2$  でもよく、さらにある複数時間スロット区間  
(例えば SBR エンベロープ)

[数49]

$$b_i \leq r < b_{i+1}$$

でのそれらの平均値である数式 (24) の

[数50]

$$\overline{e(i)}, |\overline{e(i)}|^2$$

もあわせて時間スロット選択情報としてもよい。ただし、

[数51]

$$|\overline{e(i)}|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e(r)|^2}{b_{i+1} - b_i}$$

である。

[0169] さらに時間スロット選択情報としては、数式 (26)、数式 (41) の  $e_{\text{exp}}(r)$  またはその算出過程にて平方根演算をしない  $|e_{\text{exp}}(r)|^2$  でもよく、さらにある複数時間スロット区間 (例えば SBR エンベロープ)

[数52]

$$b_i \leq r < b_{i+1}$$

でのそれらの平均値である

[数53]

$$\overline{e}_{\text{exp}}(i), |\overline{e}_{\text{exp}}(i)|^2$$

もあわせて時間スロット選択情報としてもよい。ただし、

[数54]

$$\bar{e}_{exp}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{exp}(r)}{b_{i+1} - b_i}$$

[数55]

$$|\bar{e}_{exp}(i)|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} |e_{exp}(r)|^2}{b_{i+1} - b_i}$$

である。さらに時間スロット選択情報としては、数式（23）、数式（35）、数式（36）の  $e_{adj}(r)$  またはその算出過程にて平方根演算をしない  $|e_{adj}(r)|^2$  でもよく、さらにある複数時間スロット区間（例えばSBRエンベロープ）

[数56]

$$b_i \leq r < b_{i+1}$$

でのそれらの平均値である

[数57]

$$\bar{e}_{adj}(i), \quad |\bar{e}_{adj}(i)|^2$$

もあわせて時間スロット選択情報としてもよい。ただし、

[数58]

$$\bar{e}_{adj}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{adj}(r)}{b_{i+1} - b_i}$$

[数59]

$$\left| \bar{e}_{adj}(i) \right|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} \left| e_{adj}(r) \right|^2}{b_{i+1} - b_i}$$

である。さらに時間スロット選択情報としては、数式（37）の  $e_{adj, scaled}(r)$  またはその算出過程にて平方根演算をしない  $\left| e_{adj, scaled}(r) \right|^2$  でもよく、さらにある複数時間スロット区間（例えばSBRエンベロープ）

[数60]

$$b_i \leq r < b_{i+1}$$

でのそれらの平均値である

[数61]

$$\bar{e}_{adj, scaled}(i), \quad \left| \bar{e}_{adj, scaled}(i) \right|^2$$

もあわせて時間スロット選択情報としてもよい。ただし、

[数62]

$$\bar{e}_{adj, scaled}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} e_{adj, scaled}(r)}{b_{i+1} - b_i}$$

[数63]

$$\left| \bar{e}_{adj, scaled}(i) \right|^2 = \frac{\sum_{r=b_i}^{b_{i+1}-1} \left| e_{adj, scaled}(r) \right|^2}{b_{i+1} - b_i}$$

である。さらに時間スロット選択情報としては、時間エンベロープが変形された高周波成分に対応するQMF領域信号の時間スロット  $r$  の信号電力  $P_{envadj}$

( $r$ ) またはその平方根演算をした信号振幅値

[数64]

$$\sqrt{P_{envadj}(r)}$$

でもよく、さらにある複数時間スロット区間（例えばSBRエンベロープ）

[数65]

$$b_i \leq r < b_{i+1}$$

でのそれらの平均値である

[数66]

$$\bar{P}_{envadj}(i), \quad \sqrt{\bar{P}_{envadj}(i)}$$

もあわせて時間スロット選択情報としてもよい。ただし、

[数67]

$$P_{envadj}(r) = \sum_{k=k_x}^{k_x+M-1} |q_{envadj}(k, r)|^2$$

[数68]

$$\bar{P}_{envadj}(i) = \frac{\sum_{r=b_i}^{b_{i+1}-1} P_{envadj}(r)}{b_{i+1} - b_i}$$

である。ただし、 $M$ は高周波生成部2gによって生成される高周波成分の下  
限周波数 $k_x$ より高い周波数の範囲を表す値であり、さらには高周波生成部2  
gによって生成される高周波成分の周波数範囲を $k_x \leq k < k_x + M$ のように表  
してもよい。

[0170] 時間スロット選択部3a2は、時間エンベロープ変形部2v1から通知され  
た時間スロット選択情報に基づいて、時間エンベロープ変形部2v1にて時間

エンベロープを変形された時間スロット  $r$  の高周波成分の QMF 領域の信号  $q_{\text{envadj}}(k, r)$  に対して、線形予測フィルタ部 2 k において線形予測合成フィルタ処理を施すか否かを判断し、線形予測合成フィルタ処理を施す時間スロットを選択する(ステップ S p 1 の処理)。

[0171] 本変形例における時間スロット選択部 3 a 2 での線形予測合成フィルタ処理を施す時間スロットの選択では、時間エンベロープ変形部 2 v 1 から通知された時間スロット選択情報に含まれるパラメータ  $u(r)$  が所定の値  $u_{\text{Th}}$  よりも大きい時間スロット  $r$  をひとつ以上選択してもよく、 $u(r)$  が所定の値  $u_{\text{Th}}$  よりも大きい等しい時間スロット  $r$  をひとつ以上選択してもよい。 $u(r)$  は、上記  $e(r)$ 、 $|e(r)|^2$ 、 $e_{\text{exp}}(r)$ 、 $|e_{\text{exp}}(r)|^2$ 、 $e_{\text{adj}}(r)$ 、 $|e_{\text{adj}}(r)|^2$ 、 $e_{\text{adj, scaled}}(r)$ 、 $|e_{\text{adj, scaled}}(r)|^2$ 、 $P_{\text{envadj}}(r)$ 、そして、

[数69]

$$\sqrt{P_{\text{envadj}}(r)}$$

のうち少なくともひとつを含んでもよく、 $u_{\text{Th}}$  は、上記

[数70]

$$\begin{aligned} & \overline{e(i)}, |\overline{e(i)}|^2, e_{\text{exp}}(i), \\ & |\overline{e_{\text{exp}}(i)}|^2, \overline{e_{\text{adj}}(i)}, |\overline{e_{\text{adj}}(i)}|^2, \\ & \overline{e_{\text{adj, scaled}}(i)}, |\overline{e_{\text{adj, scaled}}(i)}|^2, \\ & \overline{P_{\text{envadj}}(i)}, \sqrt{\overline{P_{\text{envadj}}(i)}}, \end{aligned}$$

のうち少なくともひとつを含んでもよい。また  $u_{\text{Th}}$  は、時間スロット  $r$  を含む所定の時間幅(例えば SBR エンベロープ)の  $u(r)$  の平均値でもよい。

さらに、 $u(r)$  がピークになる時間スロットが含まれるように選択してもよい。 $u(r)$  のピークは、前記第 1 の実施形態の変形例 4 における高周波成分の QMF 領域信号の信号電力のピークの算出と同様に算出できる。さらに、前記第 1 の実施形態の変形例 4 における定常状態と過渡状態を、 $u(r)$

）を用いて前記第 1 の実施形態の変形例 4 と同様に判断し、それに基づいて時間スロットを選択してもよい。時間スロットの選択方法は、前記の方法を少なくともひとつ用いてもよく、さらには前記とは異なる方法を少なくともひとつ用いてもよく、さらにはそれらを組み合わせてもよい。

[0172] （第 4 の実施形態の変形例 6）

第 4 の実施形態の変形例 6 の音声復号装置 2 4 f（図 3 0 参照）は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 e の内蔵メモリに格納された所定のコンピュータプログラム（例えば、図 2 9 のフローチャートに示す処理を行うためのコンピュータプログラム）を RAM にロードして実行することによって音声復号装置 2 4 f を統括的に制御する。音声復号装置 2 4 f の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 4 f は、図 3 0 に示すとおり、変形例 6 においては、第 1 の実施形態と同様に第 4 の実施形態の全体を通して省略可能である、変形例 4 に記載の音声復号装置 2 4 d の信号変化検出部 2 e 1 と、高周波線形予測分析部 2 h 1 と、線形予測逆フィルタ部 2 i 1 を省略し、音声復号装置 2 4 d の時間スロット選択部 3 a、及び時間エンベロープ変形部 2 v にかえて、時間スロット選択部 3 a 2、及び時間エンベロープ変形部 2 v 1 を備える。さらに、第 4 の実施形態の全体を通して処理順序を入れ替え可能である線形予測フィルタ部 2 k 3 の線形予測合成フィルタ処理と時間エンベロープ変形部 2 v 1 での時間エンベロープの変形処理の順序を入れ替える。

[0173] 時間スロット選択部 3 a 2 は、時間エンベロープ変形部 2 v 1 から通知された時間スロット選択情報に基づいて、時間エンベロープ変形部 2 v 1 にて時間エンベロープを変形された時間スロット  $r$  の高周波成分の QMF 領域の信号  $q_{\text{envadj}}(k, r)$  に対して、線形予測フィルタ部 2 k 3 において線形予測合成フィルタ処理を施すか否かを判断し、線形予測合成フィルタ処理を施す時間スロットを選択し、選択した時間スロットを低周波線形予測分析部 2 d

1と線形予測フィルタ部2k3に通知する。

[0174] (第4の実施形態の変形例7)

第4の実施形態の変形例7の音声符号化装置14b(図50)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声符号化装置14bの内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声符号化装置14bを統括的に制御する。音声符号化装置14bの通信装置は、符号化の対象となる音声信号を外部から受信し、更に、符号化された多重化ビットストリームを外部に出力する。音声符号化装置14bは、変形例4の音声符号化装置14aのビットストリーム多重化部1g7、及び時間スロット選択部1pにかえて、ビットストリーム多重化部1g6、および時間スロット選択部1p1を備える。

[0175] ビットストリーム多重化部1g6は、ビットストリーム多重化部1g7と同様に、コアコーデック符号化部1cによって算出された符号化ビットストリームと、SBR符号化部1dによって算出されたSBR補助情報と、フィルタ強度パラメータ算出部によって算出されたフィルタ強度パラメータとエンベロープ形状パラメータ算出部1nによって算出されたエンベロープ形状パラメータとを変換した時間エンベロープ補助情報とを多重化し、さらに時間スロット選択部1p1より受け取った時間スロット選択情報を多重化し、多重化ビットストリーム(符号化された多重化ビットストリーム)を、音声符号化装置14bの通信装置を介して出力する。

[0176] 第4の実施形態の変形例7の音声復号装置24g(図31参照)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置24gの内蔵メモリに格納された所定のコンピュータプログラム(例えば、図32のフローチャートに示す処理を行うためのコンピュータプログラム)をRAMにロードして実行することによって音声復号装置24gを統括的に制御する。音声復号装置24gの通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を

外部に出力する。音声復号装置 2 4 g は、図 3 1 に示すとおり、変形例 4 に記載の音声復号装置 2 d のビットストリーム分離部 2 a 3、及び時間スロット選択部 3 a にかえて、ビットストリーム分離部 2 a 7、及び時間スロット選択部 3 a 1 を備える。

[0177] ビットストリーム分離部 2 a 7 は、音声復号装置 2 4 g の通信装置を介して入力された多重化ビットストリームを、ビットストリーム分離部 2 a 3 と同様に、時間エンベロープ補助情報と、SBR補助情報と、符号化ビットストリームと、に分離し、さらに時間スロット選択情報とに分離する。

[0178] (第 4 の実施形態の変形例 8)

第 4 の実施形態の変形例 8 の音声復号装置 2 4 h (図 3 3 参照) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 h の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 3 4 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 2 4 h を統括的に制御する。音声復号装置 2 4 h の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 4 h は、図 3 3 に示すとおり、変形例 2 の音声復号装置 2 4 b の低周波線形予測分析部 2 d、信号変化検出部 2 e、高周波線形予測分析部 2 h、線形予測逆フィルタ部 2 i、及び線形予測フィルタ部 2 k にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、及び線形予測フィルタ部 2 k 3 を備え、時間スロット選択部 3 a をさらに備える。一次高調波調整部 2 j 1 は、第 4 の実施形態の変形例 2 における一次高調波調整部 2 j 1 と同様に、前記 “MPEG-4 AAC” の SBR における “HF Adjustment “ステップにある処理のいずれか一つ以上を行う (ステップ S m 1 の処理)。二次高調波調整部 2 j 2 は、第 4 の実施形態の変形例 2 における二次高調波調整部 2 j 2 と同様に、前記 “MPEG-4 AAC” の SBR における “HF Adjustment “ステップにある処理のいずれか一つ以上を行う (ステップ S m 2 の処理)。二次

高調波調整部 2 j 2 で行う処理は、前記 “MPEG-4 AAC” の SBR における “HF A  
adjustment “ステップにある処理のうち、一次高調波調整部 2 j 1 で行われな  
かった処理とすることが望ましい。

[0179] (第 4 の実施形態の変形例 9)

第 4 の実施形態の変形例 9 の音声復号装置 2 4 i (図 3 5 参照) は、物理  
的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU  
は、ROM 等の音声復号装置 2 4 i の内蔵メモリに格納された所定のコン  
ピュータプログラム (例えば、図 3 6 のフローチャートに示す処理を行うた  
めのコンピュータプログラム) を RAM にロードして実行することによって  
音声復号装置 2 4 i を統括的に制御する。音声復号装置 2 4 i の通信装置は  
、符号化された多重化ビットストリームを受信し、更に、復号した音声信号  
を外部に出力する。音声復号装置 2 4 i は、図 3 5 に示すとおり、第 1 の実  
施形態と同様に第 4 の実施形態の全体を通して省略可能である、変形例 8 の音  
声復号装置 2 4 h の高周波線形予測分析部 2 h 1、及び線形予測逆フィルタ  
部 2 i 1 を省略し、変形例 8 の音声復号装置 2 4 h の時間エンベロープ変形  
部 2 v、及び時間スロット選択部 3 a にかえて、時間エンベロープ変形部 2  
v 1、及び時間スロット選択部 3 a 2 を備える。さらに、第 4 の実施形態の  
全体を通して処理順序を入れ替え可能である線形予測フィルタ部 2 k 3 の線  
形予測合成フィルタ処理と時間エンベロープ変形部 2 v 1 での時間エンベロ  
ープの変形処理の順序を入れ替える。

[0180] (第 4 の実施形態の変形例 10)

第 4 の実施形態の変形例 10 の音声復号装置 2 4 j (図 3 7 参照) は、物  
理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU  
は、ROM 等の音声復号装置 2 4 j の内蔵メモリに格納された所定のコン  
ピュータプログラム (例えば、図 3 6 のフローチャートに示す処理を行う  
ためのコンピュータプログラム) を RAM にロードして実行することによっ  
て音声復号装置 2 4 j を統括的に制御する。音声復号装置 2 4 j の通信装置  
は、符号化された多重化ビットストリームを受信し、更に、復号した音声信

号を外部に出力する。音声復号装置 2 4 j は、図 3 7 に示すとおり、第 1 の実施形態と同様に第 4 の実施形態の全体を通して省略可能である、変形例 8 の音声復号装置 2 4 h の信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、及び線形予測逆フィルタ部 2 i 1 を省略し、変形例 8 の音声復号装置 2 4 h の時間エンベロープ変形部 2 v、及び時間スロット選択部 3 a にかえて、時間エンベロープ変形部 2 v 1、及び時間スロット選択部 3 a 2 を備える。さらに、第 4 の実施形態の全体を通して処理順序を入れ替え可能である線形予測フィルタ部 2 k 3 の線形予測合成フィルタ処理と時間エンベロープ変形部 2 v 1 での時間エンベロープの変形処理の順序を入れ替える。

[0181] (第 4 の実施形態の変形例 1 1)

第 4 の実施形態の変形例 1 1 の音声復号装置 2 4 k (図 3 8 参照) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 k の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 3 9 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 2 4 k を統括的に制御する。音声復号装置 2 4 k の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 4 k は、図 3 8 に示すとおり、変形例 8 の音声復号装置 2 4 h のビットストリーム分離部 2 a 3、及び時間スロット選択部 3 a にかえて、ビットストリーム分離部 2 a 7、及び時間スロット選択部 3 a 1 を備える。

[0182] (第 4 の実施形態の変形例 1 2)

第 4 の実施形態の変形例 1 2 の音声復号装置 2 4 q (図 4 0 参照) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 q の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 4 1 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 2 4 q を統括的に制御する。音声復号装置 2 4 q の通信装置

は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 24 q は、図 40 に示すとおり、変形例 3 の音声復号装置 24 c の低周波線形予測分析部 2 d、信号変化検出部 2 e、高周波線形予測分析部 2 h、線形予測逆フィルタ部 2 i、及び個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 にかえて、低周波線形予測分析部 2 d 1、信号変化検出部 2 e 1、高周波線形予測分析部 2 h 1、線形予測逆フィルタ部 2 i 1、及び個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 を備え（個別信号成分調整部は、時間エンベロープ変形手段に相当する）、時間スロット選択部 3a をさらに備える。

[0183] 個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 のうち少なくともひとつは、前記一次高周波調整手段の出力に含まれる信号成分に関して、時間スロット選択部 3 a より通知された選択結果に基づき、選択された時間スロットの QMF 領域信号に対して、個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 と同様に、処理を行う（ステップ S n 1 の処理）。時間スロット選択情報を用いて行う処理は、前記第 4 の実施形態の変形例 3 に記載の個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 における処理のうち、周波数方向の線形予測合成フィルタ処理を含む処理のうち少なくともひとつを含むのが望ましい。

[0184] 個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 における処理は、前記第 4 の実施形態の変形例 3 に記載の個別信号成分調整部 2 z 1, 2 z 2, 2 z 3 の処理と同様に、互いに同じでもよいが、個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 は、一次高周波調整手段の出力に含まれる複数の信号成分の各々に対し互いに異なる方法で時間エンベロープの変形を行ってもよい。（個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 の全てが時間スロット選択部 3 a より通知された選択結果に基づいて処理しない場合は、本発明の第 4 の実施形態の変形例 3 と同等になる）。

[0185] 時間スロット選択部 3 a から個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 のそれぞれに通知される時間スロットの選択結果は、必ずしも全てが同じである必要はなく、全てまたは一部が異なってもよい。

[0186] さらに、図 40 ではひとつの時間スロット選択部 3 a から個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 のそれぞれに時間スロットの選択結果を通知する構成になっているが、個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 のそれぞれ、または一部に対して異なる時間スロットの選択結果を通知する時間スロット選択部を複数有してもよい。またその際に、個別信号成分調整部 2 z 4, 2 z 5, 2 z 6 のうち、第 4 の実施形態の変形例 3 に記載の処理 4 (入力信号に対して時間エンベロープ変形部 2 v と同様の、エンベロープ形状調整部 2 s から得られた時間エンベロープを用いて各 QMF サブバンドサンプルヘゲイン係数を乗算する処理を行った後、その出力信号に対してさらに線形予測フィルタ部 2 k と同様の、フィルタ強度調整部 2 f から得られた線形予測係数を用いた周波数方向の線形予測合成フィルタ処理) を行う個別信号成分調整部に対する時間スロット選択部は、時間エンベロープ変形部から時間スロット選択情報を入力されて時間スロットの選択処理を行ってもよい。

[0187] (第 4 の実施形態の変形例 1 3)

第 4 の実施形態の変形例 1 3 の音声復号装置 2 4 m (図 4 2 参照) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU は、ROM 等の音声復号装置 2 4 m の内蔵メモリに格納された所定のコンピュータプログラム (例えば、図 4 3 のフローチャートに示す処理を行うためのコンピュータプログラム) を RAM にロードして実行することによって音声復号装置 2 4 m を統括的に制御する。音声復号装置 2 4 m の通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置 2 4 m は、図 4 2 に示すとおり、変形例 1 2 の音声復号装置 2 4 q のビットストリーム分離部 2 a 3、及び時間スロット選択部 3 a にかえて、ビットストリーム分離部 2 a 7、及び時間スロット選択部 3 a 1 を備える。

[0188] (第 4 の実施形態の変形例 1 4)

第 4 の実施形態の変形例 1 4 の音声復号装置 2 4 n (不図示) は、物理的には図示しない CPU、ROM、RAM 及び通信装置等を備え、この CPU

は、ROM等の音声復号装置24nの内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声復号装置24nを統括的に制御する。音声復号装置24nの通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置24nは、機能的には、変形例1の音声復号装置24aの低周波線形予測分析部2d、信号変化検出部2e、高周波線形予測分析部2h、線形予測逆フィルタ部2i、及び線形予測フィルタ部2kにかえて、低周波線形予測分析部2d1、信号変化検出部2e1、高周波線形予測分析部2h1、線形予測逆フィルタ部2i1、及び線形予測フィルタ部2k3を備え、時間スロット選択部3aをさらに備える。

[0189] (第4の実施形態の変形例15)

第4の実施形態の変形例15の音声復号装置24p(不図示)は、物理的には図示しないCPU、ROM、RAM及び通信装置等を備え、このCPUは、ROM等の音声復号装置24pの内蔵メモリに格納された所定のコンピュータプログラムをRAMにロードして実行することによって音声復号装置24pを統括的に制御する。音声復号装置24pの通信装置は、符号化された多重化ビットストリームを受信し、更に、復号した音声信号を外部に出力する。音声復号装置24pは、機能的には、変形例14の音声復号装置24nの時間スロット選択部3aにかえて、時間スロット選択部3a1を備える。さらに、ビットストリーム分離部2a4にかえて、ビットストリーム分離部2a8(不図示)を備える。

[0190] ビットストリーム分離部2a8は、ビットストリーム分離部2a4と同様に、多重化ビットストリームを、SBR補助情報と、符号化ビットストリームとに分離し、さらに時間スロット選択情報とに分離する。

### 産業上の利用可能性

[0191] SBRに代表される周波数領域での帯域拡張技術において適用される技術であって、ビットレートを著しく増大させることなく、発生するプリエコー・ポストエコーを軽減し復号信号の主観的品質を向上させるための技術に利

用できる。

## 符号の説明

[0192] 1 1, 1 1 a, 1 1 b, 1 1 c, 1 2, 1 2 a, 1 2 b, 1 3, 1 4, 1 4 a, 1 4 b…音声符号化装置、1 a…周波数変換部、1 b…周波数逆変換部、1 c…コアコーデック符号化部、1 d…SBR符号化部、1 e, 1 e 1…線形予測分析部、1 f…フィルタ強度パラメータ算出部、1 f 1…フィルタ強度パラメータ算出部、1 g, 1 g 1, 1 g 2, 1 g 3, 1 g 4, 1 g 5, 1 g 6, 1 g 7…ビットストリーム多重化部、1 h…高周波周波数逆変換部、1 i…短時間電力算出部、1 j…線形予測係数間引き部、1 k…線形予測係数量子化部、1 m…時間エンベロープ算出部、1 n…エンベロープ形状パラメータ算出部、1 p, 1 p 1…時間スロット選択部、2 1, 2 2, 2 3, 2 4, 2 4 b, 2 4 c…音声復号装置、2 a, 2 a 1, 2 a 2, 2 a 3, 2 a 5, 2 a 6, 2 a 7…ビットストリーム分離部、2 b…コアコーデック復号部、2 c…周波数変換部、2 d, 2 d 1…低周波線形予測分析部、2 e, 2 e 1…信号変化検出部、2 f…フィルタ強度調整部、2 g…高周波生成部、2 h, 2 h 1…高周波線形予測分析部、2 i, 2 i 1…線形予測逆フィルタ部、2 j, 2 j 1, 2 j 2, 2 j 3, 2 j 4…高周波調整部、2 k, 2 k 1, 2 k 2, 2 k 3…線形予測フィルタ部、2 m…係数加算部、2 n…周波数逆変換部、2 p, 2 p 1…線形予測係数補間・補外部、2 r…低周波時間エンベロープ計算部、2 s…エンベロープ形状調整部、2 t…高周波時間エンベロープ算出部、2 u…時間エンベロープ平坦化部、2 v, 2 v 1…時間エンベロープ変形部、2 w…補助情報変換部、2 z 1, 2 z 2, 2 z 3, 2 z 4, 2 z 5, 2 z 6…個別信号成分調整部、3 a, 3 a 1, 3 a 2…時間スロット選択部

## 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であって、  
前記音声信号の低周波成分を符号化するコア符号化手段と、  
前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出手段と、  
少なくとも、前記コア符号化手段によって符号化された前記低周波成分と、前記時間エンベロープ補助情報算出手段によって算出された前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化手段と、  
を備える、ことを特徴とする音声符号化装置。
- [請求項2] 前記時間エンベロープ補助情報は、所定の解析区間内において前記音声信号の高周波成分における時間エンベロープの変化の急峻さを示すパラメータを表す、ことを特徴とする請求項1に記載の音声符号化装置。
- [請求項3] 前記音声信号を周波数領域に変換する周波数変換手段を更に備え、  
前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って取得された高周波線形予測係数に基づいて、前記時間エンベロープ補助情報を算出する、ことを特徴とする請求項2に記載の音声符号化装置。
- [請求項4] 前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の低周波側係数に対し周波数方向に線形予測分析を行って低周波線形予測係数を取得し、該低周波線形予測係数と前記高周波線形予測係数とに基づいて前記時間エンベロープ補助情報を算出する、ことを特徴とする請求項3に記載の音声符号化装置。
- [請求項5] 前記時間エンベロープ補助情報算出手段は、前記低周波線形予測係

数及び前記高周波線形予測係数のそれぞれから予測ゲインを取得し、当該二つの予測ゲインの大小に基づいて前記時間エンベロープ補助情報を算出する、ことを特徴とする請求項 4 に記載の音声符号化装置。

[請求項6] 前記時間エンベロープ補助情報算出手段は、前記音声信号から高周波成分を分離し、時間領域で表現された時間エンベロープ情報を当該高周波成分から取得し、当該時間エンベロープ情報の時間的变化の大きさに基づいて前記時間エンベロープ補助情報を算出する、ことを特徴とする請求項 2 に記載の音声符号化装置。

[請求項7] 前記時間エンベロープ補助情報は、前記音声信号の低周波成分に対し周波数方向への線形予測分析を行って得られる低周波線形予測係数を用いて高周波線形予測係数を取得するための差分情報を含む、ことを特徴とする請求項 1 に記載の音声符号化装置。

[請求項8] 前記音声信号を周波数領域に変換する周波数変換手段を更に備え、前記時間エンベロープ補助情報算出手段は、前記周波数変換手段によって周波数領域に変換された前記音声信号の低周波成分及び高周波側係数のそれぞれに対し周波数方向に線形予測分析を行って低周波線形予測係数と高周波線形予測係数とを取得し、当該低周波線形予測係数及び高周波線形予測係数の差分を取得することによって前記差分情報を取得する、ことを特徴とする請求項 7 に記載の音声符号化装置。

[請求項9] 前記差分情報は、L S P (Linear Spectrum Pair)、I S P (Immittance Spectrum Pair)、L S F (Linear Spectrum Frequency)、I S F (Immittance Spectrum Frequency)、P A R C O R 係数のいずれかの領域における線形予測係数の差分を表す、ことを特徴とする請求項 8 に記載の音声符号化装置。

[請求項10] 音声信号を符号化する音声符号化装置であって、前記音声信号の低周波成分を符号化するコア符号化手段と、前記音声信号を周波数領域に変換する周波数変換手段と、前記周波数変換手段によって周波数領域に変換された前記音声信号

の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析手段と、

前記線形予測分析手段によって取得された前記高周波線形予測係数を時間方向に間引く予測係数間引き手段と、

前記予測係数間引き手段によって間引きされた後の前記高周波線形予測係数を量子化する予測係数量子化手段と、

少なくとも前記コア符号化手段による符号化後の前記低周波成分と前記予測係数量子化手段による量子化後の前記高周波線形予測係数とが多重化されたビットストリームを生成するビットストリーム多重化手段と、

を備える、ことを特徴とする音声符号化装置。

[請求項11]

符号化された音声信号を復号する音声復号装置であって、

前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと時間エンベロープ補助情報とに分離するビットストリーム分離手段と、

前記ビットストリーム分離手段によって分離された前記符号化ビットストリームを復号して低周波成分を得るコア復号手段と、

前記コア復号手段によって得られた前記低周波成分を周波数領域に変換する周波数変換手段と、

前記周波数変換手段によって周波数領域に変換された前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成手段と、

前記周波数変換手段によって周波数領域に変換された前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析手段と、

前記低周波時間エンベロープ分析手段によって取得された前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整手段と、

前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を用いて、前記高周波生成手段によって生成された前記高周波成分の時間エンベロープを変形する時間エンベロープ変形手段と、  
を備える、ことを特徴とする音声復号装置。

[請求項12]

前記高周波成分を調整する高周波調整手段を更に備え、  
前記周波数変換手段は、実数又は複素数の係数を持つ64分割QMFフィルタバンクであり、  
前記周波数変換手段、前記高周波生成手段、前記高周波調整手段は“ISO/IEC 14496-3”に規定される“MPEG4 AAC”におけるSBR復号器（SBR：Spectral Band Replication）に準拠した動作をする、  
ことを特徴とする請求項11の音声復号装置。

[請求項13]

前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分に周波数方向の線形予測分析を行って低周波線形予測係数を取得し、  
前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整し、  
前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し前記時間エンベロープ調整手段による調整後の線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する、ことを特徴とする請求項11又は12に記載の音声復号装置。

[請求項14]

前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分の時間スロットごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、  
前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、  
前記時間エンベロープ変形手段は、前記高周波生成手段によって生

成された周波数領域の高周波成分に前記調整後の時間エンベロープ情報を重畳することにより高周波成分の時間エンベロープを変形する、ことを特徴とする請求項 11 又は 12 記載の音声復号装置。

[請求項15] 前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分の QMF サブバンドサンプルごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、

前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、

前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に前記調整後の時間エンベロープ情報を乗算することにより高周波成分の時間エンベロープを変形する、ことを特徴とする請求項 11 又は 12 記載の音声復号装置。

[請求項16] 前記時間エンベロープ補助情報は、線形予測係数の強度の調整に用いるためのフィルタ強度パラメータを表す、ことを特徴とする請求項 13 記載の音声復号装置。

[請求項17] 前記時間エンベロープ補助情報は、前記時間エンベロープ情報の時間変化の大きさを示すパラメータを表す、ことを特徴とする請求項 14 又は 15 に記載の音声復号装置。

[請求項18] 前記時間エンベロープ補助情報は、前記低周波線形予測係数に対する線形予測係数の差分情報を含む、ことを特徴とする請求項 13 記載の音声復号装置。

[請求項19] 前記差分情報は、LSP (Linear Spectrum Pair)、ISP (Impedance Spectrum Pair)、LSF (Linear Spectrum Frequency)、ISF (Impedance Spectrum Frequency)、PARCOR 係数のいずれかの領域における線形予測係数の差分を表す、ことを特徴とする請求項 18 に記載の音声復号装置。

[請求項20] 前記低周波時間エンベロープ分析手段は、前記周波数変換手段によ

って周波数領域に変換された前記低周波成分に対し周波数方向の線形予測分析を行って前記低周波線形予測係数を取得するとともに、当該周波数領域の前記低周波成分の時間スロットごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、

前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整するとともに前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、

前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に対し前記時間エンベロープ調整手段による調整後の線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形するとともに当該周波数領域の前記高周波成分に前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を重畳することにより前記高周波成分の時間エンベロープを変形する、ことを特徴とする請求項 11 又は 12 に記載の音声復号装置。

[請求項21]

前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分に対し周波数方向の線形予測分析を行って前記低周波線形予測係数を取得するとともに、当該周波数領域の前記低周波成分の QMF サブバンドサンプルごとの電力を取得することによって音声信号の時間エンベロープ情報を取得し、

前記時間エンベロープ調整手段は、前記時間エンベロープ補助情報を用いて前記低周波線形予測係数を調整するとともに前記時間エンベロープ補助情報を用いて前記時間エンベロープ情報を調整し、

前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の高周波成分に対し前記時間エンベロープ調整手段による調整後の線形予測係数を用いて周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形するとともに当該周波数領域の前記高周波成分に前記時間エンベロープ調整手段による

調整後の前記時間エンベロープ情報を乗算することにより前記高周波成分の時間エンベロープを変形する、ことを特徴とする請求項 11 又は 12 に記載の音声復号装置。

[請求項22] 前記時間エンベロープ補助情報は、線形予測係数のフィルタ強度と、前記時間エンベロープ情報の時間変化の大きさとの両方を示すパラメータを表す、ことを特徴とする請求項 20 又は 21 に記載の音声復号装置。

[請求項23] 符号化された音声信号を復号する音声復号装置であって、  
前記符号化された音声信号を含む外部からのビットストリームを、  
符号化ビットストリームと線形予測係数とに分離するビットストリーム分離手段と、

前記線形予測係数を時間方向に補間又は補外する線形予測係数補間・補外手段と、

前記線形予測係数補間・補外手段によって補間又は補外された線形予測係数を用いて周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する時間エンベロープ変形手段と、

を備える、ことを特徴とする音声復号装置。

[請求項24] 音声信号を符号化する音声符号化装置を用いた音声符号化方法であって、

前記音声符号化装置が、前記音声信号の低周波成分を符号化するコア符号化ステップと、

前記音声符号化装置が、前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出ステップと、

前記音声符号化装置が、少なくとも、前記コア符号化ステップにおいて符号化した前記低周波成分と、前記時間エンベロープ補助情報算

出ステップにおいて算出した前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化ステップと、

を備える、ことを特徴とする音声符号化方法。

[請求項25]

音声信号を符号化する音声符号化装置を用いた音声符号化方法であって、

前記音声符号化装置が、前記音声信号の低周波成分を符号化するコア符号化ステップと、

前記音声符号化装置が、前記音声信号を周波数領域に変換する周波数変換ステップと、

前記音声符号化装置が、前記周波数変換ステップにおいて周波数領域に変換した前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析ステップと、

前記音声符号化装置が、前記線形予測分析手段ステップにおいて取得した前記高周波線形予測係数を時間方向に間引く予測係数間引きステップと、

前記音声符号化装置が、前記予測係数間引き手段ステップにおける間引き後の前記高周波線形予測係数を量子化する予測係数量子化ステップと、

前記音声符号化装置が、少なくとも前記コア符号化ステップにおける符号化後の前記低周波成分と前記予測係数量子化ステップにおける量子化後の前記高周波線形予測係数とが多重化されたビットストリームを生成するビットストリーム多重化ステップと、

を備える、ことを特徴とする音声符号化方法。

[請求項26]

符号化された音声信号を復号する音声復号装置を用いた音声復号方法であって、

前記音声復号装置が、前記符号化された音声信号を含む外部からの

ビットストリームを、符号化ビットストリームと時間エンベロープ補助情報とに分離するビットストリーム分離ステップと、

前記音声復号装置が、前記ビットストリーム分離ステップにおいて分離した前記符号化ビットストリームを復号して低周波成分を得るコア復号ステップと、

前記音声復号装置が、前記コア復号ステップにおいて得た前記低周波成分を周波数領域に変換する周波数変換ステップと、

前記音声復号装置が、前記周波数変換ステップにおいて周波数領域に変換した前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成ステップと、

前記音声復号装置が、前記周波数変換ステップにおいて周波数領域に変換した前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析ステップと、

前記音声復号装置が、前記低周波時間エンベロープ分析ステップにおいて取得した前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整ステップと、

前記音声復号装置が、前記時間エンベロープ調整ステップにおける調整後の前記時間エンベロープ情報を用いて、前記高周波生成ステップにおいて生成した前記高周波成分の時間エンベロープを変形する時間エンベロープ変形ステップと、

を備える、ことを特徴とする音声復号方法。

[請求項27]

符号化された音声信号を復号する音声復号装置を用いた音声復号方法であって、

前記音声復号装置が、前記符号化された音声信号を含む外部からのビットストリームを、符号化ビットストリームと線形予測係数とに分離するビットストリーム分離ステップと、

前記音声復号装置が、前記線形予測係数を時間方向に補間又は補外する線形予測係数補間・補外ステップと、

前記音声復号装置が、前記線形予測係数補間・補外スツツにおいて補間又は補外した前記線形予測係数を用いて、周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する時間エンベロープ変形スツツと、を備える、ことを特徴とする音声復号方法。

[請求項28]

音声信号を符号化するために、コンピュータ装置を、

前記音声信号の低周波成分を符号化するコア符号化手段、  
前記音声信号の低周波成分の時間エンベロープを用いて、前記音声信号の高周波成分の時間エンベロープの近似を得るための時間エンベロープ補助情報を算出する時間エンベロープ補助情報算出手段、及び

少なくとも、前記コア符号化手段によって符号化された前記低周波成分と、前記時間エンベロープ補助情報算出手段によって算出された前記時間エンベロープ補助情報とが多重化されたビットストリームを生成するビットストリーム多重化手段、

として機能させる、ことを特徴とする音声符号化プログラム。

[請求項29]

音声信号を符号化するために、コンピュータ装置を、

前記音声信号の低周波成分を符号化するコア符号化手段、  
前記音声信号を周波数領域に変換する周波数変換手段、  
前記周波数変換手段によって周波数領域に変換された前記音声信号の高周波側係数に対し周波数方向に線形予測分析を行って高周波線形予測係数を取得する線形予測分析手段、

前記線形予測分析手段によって取得された前記高周波線形予測係数を時間方向に間引く予測係数間引き手段、

前記予測係数間引き手段によって間引きされた後の前記高周波線形予測係数を量子化する予測係数量子化手段、及び、

少なくとも前記コア符号化手段による符号化後の前記低周波成分と前記予測係数量子化手段による量子化後の前記高周波線形予測係数と

が多重化されたビットストリームを生成するビットストリーム多重化手段、

[請求項30]

として機能させる、ことを特徴とする音声符号化プログラム。  
 符号化された音声信号を復号するために、コンピュータ装置を、  
 前記符号化された音声信号を含む外部からのビットストリームを、  
 符号化ビットストリームと時間エンベロープ補助情報とに分離するビ  
 ッストリーム分離手段、

前記ビットストリーム分離手段によって分離された前記符号化ビッ  
 トストリームを復号して低周波成分を得るコア復号手段、  
 前記コア復号手段によって得られた前記低周波成分を周波数領域に  
 変換する周波数変換手段、

前記周波数変換手段によって周波数領域に変換された前記低周波成  
 分を低周波帯域から高周波帯域に複写することによって高周波成分  
 を生成する高周波生成手段、  
 前記周波数変換手段によって周波数領域に変換された前記低周波成  
 分を分析して時間エンベロープ情報を取得する低周波時間エンベロー  
 プ分析手段、

前記低周波時間エンベロープ分析手段によって取得された前記時間  
 エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整す  
 る時間エンベロープ調整手段、及び、

前記時間エンベロープ調整手段による調整後の前記時間エンベロー  
 プ情報を用いて、前記高周波生成手段によって生成された前記高周波  
 成分の時間エンベロープを変形する時間エンベロープ変形手段、  
 として機能させる、ことを特徴とする音声復号プログラム。

[請求項31]

符号化された音声信号を復号するために、コンピュータ装置を、  
 前記符号化された音声信号を含む外部からのビットストリームを、  
 符号化ビットストリームと線形予測係数とに分離するビットストリー  
 ム分離手段、

前記線形予測係数を時間方向に補間又は補外する線形予測係数補間・補外手段、及び、

前記線形予測係数補間・補外手段によって補間又は補外された線形予測係数を用いて周波数領域で表現された高周波成分に周波数方向の線形予測フィルタ処理を行って音声信号の時間エンベロープを変形する時間エンベロープ変形手段、

として機能させる、ことを特徴とする音声復号プログラム。

[請求項32] 前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し周波数方向の線形予測フィルタ処理を行った後、前記線形予測フィルタ処理の結果得られた高周波成分の電力を前記線形予測フィルタ処理前と等しい値に調整する、ことを特徴とする請求項13、20、21のうち何れか一項に記載の音声復号装置。

[請求項33] 前記時間エンベロープ変形手段は、前記高周波生成手段によって生成された周波数領域の前記高周波成分に対し周波数方向の線形予測フィルタ処理を行った後、前記線形予測フィルタ処理の結果得られた高周波成分の任意の周波数範囲内の電力を前記線形予測フィルタ処理前と等しい値に調整する、ことを特徴とする請求項13、20、21のうち何れか一項に記載の音声復号装置。

[請求項34] 前記時間エンベロープ補助情報は、前記調整後の前記時間エンベロープ情報における最小値と平均値の比率であることを特徴とする請求項14、15、20、21、32、33のうち何れか一項に記載の音声復号装置。

[請求項35] 前記時間エンベロープ変形手段は、前記周波数領域の高周波成分のSBRエンベロープ時間セグメント内での電力が時間エンベロープの変形の前と後で等しくなるように前記調整後の時間エンベロープの利得を制御した後に、前記周波数領域の高周波成分に前記利得制御された時間エンベロープを乗算することにより高周波成分の時間エンベロ

[請求項36]

前記低周波時間エンベロープ分析手段は、前記周波数変換手段によって周波数領域に変換された前記低周波成分のQMFサブバンドフィルタごとの電力を取得し、さらにSBRエンベロープ時間セグメント内の平均電力を用いて前記QMFサブバンドフィルタごとの電力を正規化することによって、各QMFサブバンドフィルタへ乗算されるべきゲイン係数として表現された時間エンベロープ情報を取得するこ

とを特徴とする請求項12、14、15、17、20、21、32～35のうち何れか一項に記載の音声復号装置。

[請求項37]

符号化された音声信号を復号する音声復号装置であって、前記符号化された音声信号を含む外部からのビットストリームを復号して低周波成分を得るコア復号手段と、

前記コア復号手段によって得られた前記低周波成分を周波数領域に変換する周波数変換手段と、

前記周波数変換手段によって周波数領域に変換された前記低周波成分を低周波帯域から高周波帯域に複写することによって高周波成分を生成する高周波生成手段と、

前記周波数変換手段によって周波数領域に変換された前記低周波成分を分析して時間エンベロープ情報を取得する低周波時間エンベロープ分析手段と、

前記ビットストリームを分析して時間エンベロープ補助情報を生成する時間エンベロープ補助情報生成部と、

前記低周波時間エンベロープ分析手段によって取得された前記時間エンベロープ情報を、前記時間エンベロープ補助情報を用いて調整する時間エンベロープ調整手段と、

前記時間エンベロープ調整手段による調整後の前記時間エンベロープ情報を用いて、前記高周波生成手段によって生成された前記高周波

成分の時間エンベロープを変形する時間エンベロープ変形手段と、  
を備える、ことを特徴とする音声復号装置。

[請求項38] 前記高周波調整手段に相当する、一次高周波調整手段と、二次高周波調整手段とを具備し、

前記一次高周波調整手段は、前記高周波調整手段に相当する処理の一部を含む処理を実行し、

前記時間エンベロープ変形手段は、前記一次高周波調整手段の出力信号に対し時間エンベロープの変形を行い、

前記二次高周波調整手段は、前記時間エンベロープ変形手段の出力信号に対して、前記高周波調整手段に相当する処理のうち前記一次高周波調整手段で実行されない処理を実行する、ことを特徴とする請求項 1 1 ～ 2 2, 3 2 ～ 3 7 のうち何れか一項に記載の音声復号装置。

[請求項39] 前記二次高周波調整手段は、S B R の復号過程における正弦波の付加処理であることを特徴とする請求項 3 8 に記載の音声復号装置。