

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6306565号  
(P6306565)

(45) 発行日 平成30年4月4日(2018.4.4)

(24) 登録日 平成30年3月16日(2018.3.16)

(51) Int.Cl. F I  
**G 1 O L 19/22 (2013.01)** G 1 O L 19/22  
**G 1 O L 21/0388 (2013.01)** G 1 O L 21/0388 1 0 0

請求項の数 5 (全 30 頁)

(21) 出願番号	特願2015-501583 (P2015-501583)	(73) 特許権者	503447036
(86) (22) 出願日	平成25年3月21日 (2013. 3. 21)		サムスン エレクトロニクス カンパニー リミテッド
(65) 公表番号	特表2015-512528 (P2015-512528A)		大韓民国・1 6 6 7 7・キョンギード・ス ウォン-シ・ヨントン-ク・サムスン-ロ ・1 2 9
(43) 公表日	平成27年4月27日 (2015. 4. 27)	(74) 代理人	100107766
(86) 国際出願番号	PCT/KR2013/002372		弁理士 伊東 忠重
(87) 国際公開番号	W02013/141638	(74) 代理人	100070150
(87) 国際公開日	平成25年9月26日 (2013. 9. 26)		弁理士 伊東 忠彦
審査請求日	平成28年3月22日 (2016. 3. 22)	(74) 代理人	100091214
(31) 優先権主張番号	61/613, 610		弁理士 大貫 進介
(32) 優先日	平成24年3月21日 (2012. 3. 21)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	61/719, 799		
(32) 優先日	平成24年10月29日 (2012. 10. 29)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 帯域幅拡張のための高周波数符号化／復号化方法及びその装置

(57) 【特許請求の範囲】

【請求項 1】

信号分類結果に基づき、現在フレームが音声信号に当るか否かを決定する段階と、  
前記現在フレームが前記音声信号に当る場合、前記現在フレームの第1励起クラスの情報  
を生成する段階と、

前記現在フレームが前記音声信号に当たらない場合、前記現在フレームのトータル特性を  
取得する段階と、

前記トータル特性を閾値(threshold value)と比較して、前記現在フレ  
ームの第2励起クラスの情報を生成する段階と、

を含む励起クラス生成方法。

【請求項 2】

前記第1励起クラスの情報  
は前記現在フレームの励起クラスが音声クラスであることを  
示し、前記第2励起クラスの情報  
は前記現在フレームの励起クラスが第1非音声クラスある  
いは第2非音声クラスであることを示す請求項1に記載の方法。

【請求項 3】

前記第1非音声クラスは前記現在フレームがトータル(tonal)信号である場合に  
割り当てられ、前記第2非音声クラス  
は前記現在フレームがノイジー(noisy)信号  
である場合に割り当てられている請求項2に記載の方法。

【請求項 4】

前記第1励起クラスの情報と前記第2励起クラスの情報  
 はフレーム単位で生成される請

求項 1 に記載の方法。

【請求項 5】

請求項 1 ないし 4 のいずれかに記載の方法を実行できるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オーディオ符号化及び復号化に係り、さらに詳細には、帯域幅拡張のための高周波数符号化/復号化方法及びその装置に関する。

10

【背景技術】

【0002】

G.719 のコーディング・スキームは、テレカンファレンシングを目的として、開発及び標準化されたものであり、MDC T (modified discrete cosine transform) を行い、周波数ドメイン変換を行い、ステーションナリー (stationary) フレームである場合には、MDC T スペクトルを直ちにコーディングする。ノンステーションナリー (non-stationary) フレームは、時間ドメインエイリアシング順序 (time domain aliasing order) を変更することにより、時間的な特性を考慮するように変更する。ノンステーションナリー・フレームについて得られたスペクトルは、ステーションナリー・フレームと同一のフレームワークによって、コーデックを構成するためにインターリーブを行い、ステーションナリー・フレームと類似した形態によって構成される。かように構成されたスペクトルのエネルギーを求めて正規化を行った後、量子化を行う。一般的にエネルギーは、RMS (root mean square) 値で表現され、正規化されたスペクトルは、エネルギー基盤のビット割り当てを介して、バンド別に必要なビットを生成し、バンド別ビット割り当て情報を基に、量子化及び無損失符号化を介して、ビットストリームを生成する。

20

【0003】

G.719 のデコーディング・スキームによれば、コーディング方式の逆過程で、ビットストリームからエネルギーを逆量子化し、逆量子化されたエネルギーを基に、ビット割り当て情報を生成し、スペクトルの逆量子化を行って正規化された逆量子化されたスペクトルを生成する。このとき、ビットが不足している場合、特定バンドには、逆量子化したスペクトルがなくなる。かような特定バンドに対してノイズを生成するために、低周波数の逆量子化されたスペクトルを基に、ノイズコードブックを生成し、伝送されたノイズレベルに合わせてノイズを生成するノイズフィードバック方式が適用される。一方、特定周波数以上のバンドについては、低周波数信号をフォールディングして高周波数信号を生成する帯域幅拡張技法が適用される。

30

【発明の概要】

【発明が解決しようとする課題】

【0004】

本発明が解決しようとする課題は、復元音質を向上させることができる帯域幅拡張のための高周波数符号化/復号化方法及びその装置、並びにそれを採用するマルチメディア機器を提供するところにある。

40

【課題を解決するための手段】

【0005】

前記課題を解決するための本発明の一実施形態による帯域幅拡張のための高周波数符号化方法は、復号化端で高周波数励起信号を生成するのに適用される加重値を推定するためのフレーム別励起タイプ情報を生成する段階と、前記フレーム別励起タイプ情報を含むビットストリームを生成する段階と、を含んでもよい。

【0006】

前記課題を解決するための本発明の一実施形態による帯域幅拡張のための高周波数復号化方法は、加重値を推定する段階と、ランダムノイズと、復号化された低周波数スペクト

50

ルとの間に、前記加重値を適用し、高周波数励起信号を生成する段階と、を含んでもよい。

【発明の効果】

【0007】

本発明による帯域幅拡張のための高周波数符号化／復号化方法及びその装置によれば、複雑度の増大なしに、復元音質を向上させることができる。

【図面の簡単な説明】

【0008】

【図1】一実施形態によって、低周波数信号のバンド及び高周波数信号のバンドを構成する例について説明する図面である。

10

【図2A】一実施形態によって、R0領域及びR1領域が選択されたコーディング方式に対応し、R2及びR3、並びにR4及びR5に区分した図面である。

【図2B】一実施形態によって、R0領域及びR1領域が選択されたコーディング方式に対応し、R2及びR3、並びにR4及びR5に区分した図面である。

【図2C】一実施形態によって、R0領域及びR1領域が選択されたコーディング方式に対応し、R2及びR3、並びにR4及びR5に区分した図面である。

【図3】一実施形態によるオーディオ符号化装置の構成を示したブロック図である。

【図4】一実施形態によって、BWE領域R1において、R2及びR3を決定する方法について説明するフローチャートである。

【図5】一実施形態によって、BWEパラメータを決定する方法について説明するフローチャートである。

20

【図6】他の実施形態によるオーディオ符号化装置の構成を示したブロック図である。

【図7】一実施形態によって、BWEパラメータ符号化部の構成を示したブロック図である。

【図8】一実施形態によるオーディオ復号化装置の構成を示したブロック図である。

【図9】一実施形態による励起信号生成部の細部的な構成を示すブロック図である。

【図10】他の実施形態による励起信号生成部の細部的な構成を示すブロック図である。

【図11】さらに他の実施形態による励起信号生成部の細部的な構成を示すブロック図である。

【図12】バンド境界において、加重値に係わるスムージング処理について説明するための図面である。

30

【図13】一実施形態によって、オーバーラッピング領域に存在するスペクトルを再構成するために使用される寄与分である加重値について説明する図面である。

【図14】一実施形態による、スイッチング構造のオーディオ符号化装置の構成を示したブロック図である。

【図15】他の実施形態による、スイッチング構造のオーディオ符号化装置の構成を示したブロック図である。

【図16】一実施形態による、スイッチング構造のオーディオ復号化装置の構成を示したブロック図である。

【図17】他の実施形態による、スイッチング構造のオーディオ復号化装置の構成を示したブロック図である。

40

【図18】一実施形態による、符号化モジュールを含むマルチメディア機器の構成を示したブロック図である。

【図19】一実施形態による、復号化モジュールを含むマルチメディア機器の構成を示したブロック図である。

【図20】一実施形態による、符号化モジュール及び復号化モジュールを含むマルチメディア機器の構成を示したブロック図である。

【発明を実施するための形態】

【0009】

本発明は、多様な変換を加えることができ、さまざまな実施形態を有することができる

50

が、特定実施形態を図面に例示し、詳細な説明において具体的に説明する。しかし、それは、本発明を特定の実施形態について限定するものではなく、本発明の技術的思想及び技術範囲に含まれる全ての変換、均等物ないし代替物を含むと理解される。本発明について説明するにおいて、関連公知技術に係わる具体的な説明が、本発明の要旨を不明瞭にする判断される場合、その詳細な説明を省略する。

【0010】

第1、第2のような用語は、多様な構成要素について説明するのに使用されるが、構成要素は、用語によって限定されるものではない。用語は、1つの構成要素を他の構成要素から区別する目的だけに使用される。

【0011】

本発明で使用した用語は、ただ特定の実施形態について説明するために使用されたものであり、本発明を限定する意図ではない。本発明で使用した用語は、本発明における機能を考慮しながら、可能な限り現在汎用される一般的な用語を選択したが、それは当分野に携わる技術者の意図、判例または新たな技術の出現などによって異なりもする。また、特定の場合は、出願人が任意に選定した用語もあり、その場合、該当する発明の説明部分において、詳細にその意味を記載する。従って、本発明で使用される用語は、単純な用語の名称ではない、その用語が有する意味と、本発明の全般にわたった内容とを基に定義されなければならない。

【0012】

単数の表現は、文脈上明白に異なって意味しない限り、複数の表現を含む。本発明において、「含む」または「有する」というような用語は、明細書上に記載された特徴、数字、段階、動作、構成要素、部品、またはそれらを組み合わせが存在することを指定するものであり、一つまたはそれ以上の他の特徴、数字、段階、動作、構成要素、部品、またはそれら組み合わせの存在または付加の可能性をあらかじめ排除するものではないということが理解されなければならない。

【0013】

以下、本発明の実施形態について、添付図面を参照して詳細に説明するが、添付図面を参照して説明するにおいて、同一であるか、あるいは対応する構成要素は、同一の図面番号を付し、それに係わる重複説明は省略する。

【0014】

図1は、低周波数信号のバンド及び高周波数信号のバンドを構成する例について説明する図面である。実施形態によれば、サンプリングレートは、32kHzであり、640個のMDCT(modified discrete cosine transform)スペクトル係数を、22個のバンドによって構成し、具体的には、低周波数信号について、17個のバンド、高周波数信号について、5個のバンドによって構成される。高周波数信号の開始周波数は、241番目のスペクトル係数であり、0~240までのスペクトル係数は、低周波数コーディング方式でコーディングされる領域であり、R0と定義する。また、241~639までのスペクトル係数は、BWE(bandwidth extension)が行われる領域であり、R1と定義する。一方、R1領域には、低周波数コーディング方式でコーディングされるバンドも存在する。

【0015】

図2Aないし図2Cは、図1のR0領域及びR1領域を、選択されたコーディング方式によって、R2、R3、R4、R5に区分した図面である。まず、BWE領域であるR1領域は、R2及びR3に、低周波数コーディング領域であるR0領域は、R4及びR5に区分される。R2は、低周波数コーディング方式、例えば、周波数ドメインコーディング方式で、量子化及び無損失符号化がなされる信号を含んでいるバンドを示し、R3は、低周波数コーディング方式でコーディングされる信号がないバンドを示す。一方、R2が低周波数コーディング方式でコーディングされるために、ビット割り当てを行うように定義した場合であっても、ビットが不足して、R3と同一方式でバンドが生成される。R5は、ビットが割り当てられ、低周波数コーディング方式でコーディングが行われる

10

20

30

40

50

バンドを示し、R 4 は、ビット余裕分がなく、低周波数信号にもかかわらず、コーディングされないか、あるいはビットが少なく割り当てられ、ノイズを付加しなければならないバンドを示す。従って、R 4 及び R 5 の区分は、ノイズ付加いかんによって判断され、それは、低周波数コーディングされたバンド内スペクトル個数の比率によって決定され、または F P C (factorial pulse coding) を使用した場合には、バンド内パルス割り当て情報に基づいて決定する。R 4 バンド及び R 5 バンドは、復号化過程においてノイズを付加するときに区分されるために、符号化過程においては、明確に区分されるものではない。R 2 バンド～ R 5 バンドは、符号化される情報が互いに異なるだけでなく、デコーディング方式が異なって適用されるもする。

#### 【 0 0 1 6 】

図 2 A に図示された例の場合、低周波数コーディング領域 R 0 において、1 7 0 ~ 2 4 0 までの 2 個バンドが、ノイズを付加する R 4 であり、B W E 領域 R 1 において、2 4 1 ~ 3 5 0 までの 2 個バンド、及び 4 2 7 ~ 6 3 9 までの 2 個バンドが、低周波数コーディング方式でコーディングされる R 2 である。図 2 B に図示された例の場合、低周波数コーディング領域 R 0 において、2 0 2 ~ 2 4 0 までの 1 個バンドが、ノイズを付加する R 4 であり、B W E 領域 R 1 において、2 4 1 ~ 6 3 9 までの 5 個バンドが、いずれも低周波数コーディング方式でコーディングされる R 2 である。図 2 C に図示された例の場合、低周波数コーディング領域 R 0 において、1 4 4 ~ 2 4 0 までの 3 個バンドが、ノイズを付加する R 4 であり、B W E 領域 R 1 において、R 2 は存在しない。低周波数コーディング領域 R 0 において、R 4 は、一般的に高周波数部分に分布されるが、B W E 領域 R 1 において、R 2 は、特定周波数部分に制限されない。

#### 【 0 0 1 7 】

図 3 は、一実施形態によるオーディオ符号化装置の構成を示したブロック図である。図 3 に、図示されたオーディオ符号化装置は、トランジェント検出部 3 1 0、変換部 3 2 0、エネルギー抽出部 3 3 0、エネルギー符号化部 3 4 0、トナリティ算出部 3 5 0、コーディングバンド選択部 3 6 0、スペクトル符号化部 3 7 0、B W E パラメータ符号化部 3 8 0 及び多重化部 3 9 0 を含んでもよい。各構成要素は、少なくとも 1 つのモジュールに一体化され、少なくとも 1 つのプロセッサ (図示せず) によって具現されるもする。ここで、入力信号は、音楽あるいは音声、あるいは音楽と音声との混合信号を意味し、音声信号と、それ以外野一般的な信号とに大別されるもする。以下では、説明の便宜のために、オーディオ信号と総称する。

#### 【 0 0 1 8 】

図 3 を参照すれば、トランジェント検出部 3 1 0 は、時間ドメインのオーディオ信号について、トランジェント信号あるいはアタック信号が存在するか否かということを検出する。そのために、公知された多様な方法を適用することができ、一例として、時間ドメインのオーディオ信号のエネルギー変化を利用することが可能である。現在フレームからトランジェント信号あるいはアタック信号が検出されれば、現在フレームをトランジェント・フレームと定義し、そうではない場合、ノンランジェント・フレーム、例えば、ステーションナリー (stationary) ・フレームと定義する。

#### 【 0 0 1 9 】

変換部 3 2 0 は、トランジェント検出部 3 1 0 での検出結果に基づいて、時間ドメインのオーディオ信号を周波数ドメインに変換する。変換方式の一例として、M D C T が適用されるが、それに限定されるものではない。トランジェント・フレームとステーションナリー・フレームとの各変換処理、及びインターリーブ処理は、G . 7 1 9 だと同一に行われるが、それに限定されるものではない。

#### 【 0 0 2 0 】

エネルギー抽出部 3 3 0 は、変換部 3 2 0 から提供される周波数ドメインのスペクトルについてエネルギーを抽出する。周波数ドメインのスペクトルは、バンド単位で構成され、バンド長は、均一でもあり、あるいは不均一でもある。エネルギーは、各バンドの平均エネルギー、平均電力、エンベロープあるいは norm を意味する。各バンドについて抽出さ

10

20

30

40

50

れたエネルギーは、エネルギー符号化部 3 4 0 及びスペクトル符号化部 3 7 0 に提供される。

【 0 0 2 1 】

エネルギー符号化部 3 4 0 は、エネルギー抽出部 3 3 0 から提供される各バンドのエネルギーについて、量子化及び無損失符号化を行う。エネルギー量子化は、均一スカラー量子化器 (uniform scalar quantizer)、非均一スカラー量子化器 (non-uniform scalar quantizer) あるいはベクトル量子化器 (vector quantizer) など多様な方式を利用して行われる。エネルギー無損失符号化は、算術コーディング (arithmetic coding) あるいはホフマンコーディング (Huffman coding) など多様な方式を利用して行われる。

10

【 0 0 2 2 】

トナリティ算出部 3 5 0 は、変換部 3 2 0 から提供される周波数ドメインのスペクトルについて、トナリティを算出する。各バンドについてトナリティを算出することにより、現在バンドがトーン性 (tone-like characteristic) を有する否かということ、あるいはノイズ性 (noise-like characteristic) を有するか否かということを判断する。トナリティは、SFM (spectral flatness measurement) に基づいて算出されるか、あるいは下記数式 (1) のように、平均振幅に対するピークの比率と定義される。

【 0 0 2 3 】

【数 1】

$$T(b) = \frac{\max[S(k)*S(k)]}{\frac{1}{N} \sum S(k)*S(k)} \quad (1)$$

20

ここで、 $T(b)$  は、バンド  $b$  のトナリティ、 $N$  は、バンド長、 $S(k)$  は、バンド  $b$  のスペクトル係数を示す。 $T(b)$  は、 $d$   $b$  値に変更されて使用される。

【 0 0 2 4 】

一方、トナリティは、以前フレームの当該バンドのトナリティ、及び現在フレームの当該バンドのトナリティに係わる加重和 (weighted sum) として算出される。その場合、バンド  $b$  のトナリティ  $T(b)$  は、下記数式 (2) のように定義される。

30

【 0 0 2 5 】

【数 2】

$$T(b) = a_0 * T(b, n-1) + (1-a_0) * T(b, n) \quad (2)$$

ここで、 $T(b, n)$  は、フレーム  $n$  のバンド  $b$  でのトナリティを示し、 $a_0$  は、加重値であり、実験的に、あるいはシミュレーションを介して、事前に最適値に設定される。

40

【 0 0 2 6 】

トナリティは、高周波数信号を構成するバンド、例えば、図 1 の  $R_1$  領域のバンドについて算出されるが、必要により、低周波数信号を構成するバンド、例えば、図 1 の  $R_0$  領域のバンドについても算出される。一方、バンド内のスペクトル長が過度に長い場合は、トナリティ算出時、エラーが発生することができるために、バンドを分離して算出した後、その平均値あるいは最大値により、そのバンドを代表するトナリティとして設定することができる。

【 0 0 2 7 】

コーディングバンド選択部 3 6 0 は、各バンドのトナリティを基にして、コーディング

50

バンドを選択する。一実施形態によれば、図1のBWE領域R1について、R2及びR3を決定する。一方、図1の低周波数コーディング領域R0のR4及びR5は、割り当てることができるビットを考慮して決定することが可能である。

【0028】

具体的には、低周波数コーディング領域R0でのコーディングバンド選択処理について説明する。

【0029】

R5は、周波数ドメインコーディング方式によって、ビットを割り当ててコーディングを行う。一実施形態によれば、周波数ドメインコーディング方式でコーディングを行うために、バンド別ビット割り当て情報によって割り当てられたビットを基にパルスをコーディングするファクトリアル・パルスコーディング (factorial pulse coding) 方式を適用する。ビット割り当て情報としては、エネルギーを使用することができ、エネルギーが大きいバンドには、多くのビットが割り当てられ、エネルギーが小さいバンドには、少ないビットが割り当てられるように設計する。割り当てることができるビットは、ターゲットビット率によって制限され、かような制限条件下で、ビットを割り当てるために、ターゲットビット率が低い場合、R5とR4とのバンド区分がさらに意味があり得る。ところで、トランジェント・フレームである場合には、ステーションナリー・フレームとは異なる方式でビット割り当てを行う。一実施形態によれば、トランジェント・フレームである場合、高周波数信号のバンドについては、ビット割り当てを強制的に行わないように設定する。すなわち、トランジェント・フレームにおいて、特定周波数以後のバンドについては、ビットを0に割り当てることにより、低周波数信号を良好に表現するようにすれば、低いターゲットビット率において音質改善を得ることができる。一方、ステーションナリー・フレームにおいて、特定周波数以後のバンドについて、ビットを0に割り当てる。また、ステーションナリー・フレームにおいて、高周波数信号のバンドにおいて、で所定臨界値を超えるエネルギーが含まれたバンドについては、ビット割り当てを行う。かようなビット割り当て処理は、エネルギー情報及び周波数情報を基にして行われ、符号化部及び復号化部において、同一方式を適用するために、追加する付加情報をビットストリームに含める必要がない。一実施形態によれば、量子化された後でさらに逆量子化されたエネルギーを利用して、ビット割り当てを行うことができる。

【0030】

図4は、一実施形態によって、BWE領域R1において、R2及びR3を選択する方法について説明するフローチャートである。ここで、R2は、周波数ドメインコーディング方式でコーディングされた信号を含んでいるバンドであり、R3は、周波数ドメインコーディング方式でコーディングされた信号を含んでいないバンドである。BWE領域R0において、R2に該当するバンドがいずれも選択されれば、残りのバンドがR3に該当する。R2は、トーン性を持ったバンドであるために、大きい値のトナリティを有する。一方、トナリティの代わりに、ノイズネス (noisiness) は、小さい値を有する。

【0031】

図4を参照すれば、410段階では、各バンドについてトナリティを算出し、420段階では、算出されたトナリティを所定臨界値  $T_{th0}$  と比較する。

【0032】

430段階では、420段階での比較結果、算出されたトナリティが所定臨界値より大きい値を有するバンドをR2に割り当て、 $f\_flag(b)$  を1に設定する。

【0033】

440段階では、420段階での比較結果、算出されたトナリティが所定臨界値より小さい値を有するバンドをR3に割り当て、 $f\_flag(b)$  を0に設定する。

【0034】

BWE領域R0に含まれた各バンドについて設定された  $f\_flag(b)$  は、コーディングバンド選択情報として定義され、ビットストリームに含められる。コーディングバンド選択情報は、ビットストリームに含められない。

10

20

30

40

50

## 【 0 0 3 5 】

再び図3に戻り、スペクトル符号化部370は、コーディングバンド選択部360で生成されたコーディングバンド選択情報に基づいて、低周波数信号のバンド、及びf\_flag(b)が1に設定されたR2バンドについて、スペクトル係数の周波数ドメインコーディングを行う。周波数ドメインコーディングは、量子化及び無損失符号化を含み、一実施形態によれば、ファクトリアル・パルスコーディング(FPC)方式を使用することができる。FPC方式は、コーディングされたスペクトル係数の位置、大きさ及び符号情報をパルスで表現する方式である。

## 【 0 0 3 6 】

スペクトル符号化部370は、エネルギー抽出部330から提供される各バンド別エネルギーを基に、ビット割り当て情報を生成し、各バンド別に割り当てられたビットに基づいて、FPCのためのパルス個数を計算し、パルス個数をコーディングする。そのとき、ビット不足現象によって、低周波数信号の一部バンドがコーディングされないか、あるいは、非常に少ないビットでコーディングが行われ、復号化端でノイズを付加する必要があるバンドが存在する。かような低周波数信号のバンドがR4に定義される。一方、十分な個数のパルスでコーディングが行われるバンドの場合には、復号化端でノイズを付加する必要がなく、かような低周波数信号のバンドがR5に定義される。符号化端では、低周波数信号に係わるR4及びR5の区分に意味がないので、別途のコーディングバンド選択情報を生成する必要がない。ただし、与えられた全体ビット内において、各バンド別に割り当てられたビットに基づいてパルス個数を計算し、パルス個数に対するコーディングを行う。

## 【 0 0 3 7 】

BWEパラメータ符号化部380は、低周波数信号のバンドのうち、R4バンドがノイズを付加する必要があるバンドであるということを示す情報(lf\_att\_flag)を含み、高周波数帯域幅拡張に必要なBWEパラメータを生成する。ここで、復号化端において、高周波数帯域幅拡張に必要なBWEパラメータは、低周波数信号及びランダムノイズに対して適切に加重値を付加して生成する。他の実施形態では、低周波信号をホワイトニングした信号及びランダムノイズに対して適切に加重値を付加して生成する。

## 【 0 0 3 8 】

そのとき、BWEパラメータは、現在フレームの全ての高周波数信号生成のために、ランダムノイズをさらに強く付加しなければならないという情報(all\_noise)、低周波数信号をさらに強調しなければならないという情報(all\_lf)によって構成される。lf\_att\_flag情報、all\_noise情報、all\_lf情報は、フレームごとに1度伝送され、各情報別で1ビットずつ割り当てられて伝送される。必要によってはバンド別に分離して伝送される。

## 【 0 0 3 9 】

図5は、一実施形態によって、BWEパラメータを決定する方法について説明するフローチャートである。そのために、図2の例において、241~290までバンドをPbと、521~639までバンドをEbと、すなわち、BWE領域R1の開始バンドと、最後のバンドとをそれぞれPb及びEbと定義する。

## 【 0 0 4 0 】

図5を参照すれば、510段階では、BWE領域R1の平均トナリティTa0を算出し、520段階では、平均トナリティTa0を臨界値Th1と比較する。

## 【 0 0 4 1 】

525段階では、520段階での比較結果、平均トナリティTa0が臨界値Th1より小さければ、all\_noiseを1に設定する一方、all\_lfとlf\_att\_flagは、いずれも0に設定して伝送しない。

## 【 0 0 4 2 】

530段階では、520段階での比較結果、平均トナリティTa0が臨界値Th1より大きいか、あるいはそれと同じである、ならばall\_noiseを0に設定する一方、all\_

10

20

30

40

50



1 f と 1 f \_\_ a t t \_\_ f l a g とを下記のように決定して伝送する。

【 0 0 4 3 】

一方、5 4 0 段階では、平均トナリティ T a 0 を臨界値 T t h 2 と比較する。ここで、臨界値 T t h 2 は、臨界値 T t h 1 より小さい値であることが望ましい。

【 0 0 4 4 】

5 4 5 段階では、5 4 0 段階での比較結果、平均トナリティ T a 0 が臨界値 T t h 2 より大きければ、all \_\_ i f を 1 に設定する一方、1 f \_\_ a t t \_\_ f l a g は、0 に設定して伝送しない。

【 0 0 4 5 】

5 5 0 段階では、5 4 0 段階での比較結果、平均トナリティ T a 0 が臨界値 T t h 2 より小さいが、あるいはそれと同じであるならば、all \_\_ i f を 0 に設定する一方、1 f \_\_ a t t \_\_ f l a g を下記のように決定して伝送する。

【 0 0 4 6 】

5 6 0 段階では、P b 以前バンドの平均トナリティ T a 1 を算出する。一実施形態によれば、1 つの以前バンドないし 5 つの以前バンドを考慮する。

【 0 0 4 7 】

5 7 0 段階では、以前フレームと係わりなく、平均トナリティ T a 1 を臨界値 T t h 3 と比較するか、あるいは以前フレームの 1 f \_\_ a t t \_\_ f l a g、すなわち、p \_\_ 1 f \_\_ a t t \_\_ f l a g を考慮する場合、平均トナリティ T a 1 を臨界値 T t h 4 と比較する。

【 0 0 4 8 】

5 8 0 段階では、5 7 0 段階での比較結果、平均トナリティ T a 1 が臨界値 T t h 3 より大きければ、1 f \_\_ a t t \_\_ f l a g を 1 に設定し、5 9 0 段階では、5 7 0 段階での比較結果、平均トナリティ T a 1 が臨界値 T t h 3 より小さいか、あるいはそれと同じであるならば、1 f \_\_ a t t \_\_ f l a g を 0 に設定する。

【 0 0 4 9 】

一方、5 8 0 段階では、p \_\_ 1 f \_\_ a t t \_\_ f l a g が 1 に設定された場合、平均トナリティ T a 1 が臨界値 T t h 4 より大きければ、1 f \_\_ a t t \_\_ f l a g を 1 に設定する。そのとき、以前フレームがトランジェント・フレームである場合、p \_\_ 1 f \_\_ a t t \_\_ f l a g は、0 に設定される。5 9 0 段階では、p \_\_ 1 f \_\_ a t t \_\_ f l a g が 1 に設定された場合、平均トナリティ T a 1 が臨界値 T t h 4 より小さいか、あるいはそれと同じであるならば、1 f \_\_ a t t \_\_ f l a g を 0 に設定する。ここで、臨界値 T t h 3 は、臨界値 T t h 4 より大きい値であることが望ましい。

【 0 0 5 0 】

一方、高周波数信号のバンドのうち、f l a g ( b ) が 1 に設定されたバンドが一つでも存在する場合、all \_\_ n o i s e は、0 に設定される。その理由は、高周波数信号にトーン性を有したバンドが存在するということを意味するために、all \_\_ n o i s e を 1 に設定することができないからである。その場合、all \_\_ n o i s e は、0 で伝送しながら、前記 5 4 0 段階ないし 5 9 0 段階を遂行し、all \_\_ 1 f と 1 f \_\_ a t t \_\_ f l a g とに係わる情報を生成する。

【 0 0 5 1 】

以下の表 1 は、図 5 を介して生成された B W E パラメータの伝送関係を表示したものである。ここで、数字は、当該 B W E パラメータの伝送に必要なビットを意味し、X と表記した場合には、当該 B W E パラメータを伝送しないことを意味する。B W E パラメータ、すなわち、all \_\_ n o i s e、all \_\_ 1 f、1 f \_\_ a t t \_\_ f l a g は、コーディングバンド選択部 3 6 0 で生成されたコーディングバンド選択情報である f \_\_ f l a g ( b ) と相関関係を有する。例えば、表 1 のように、all \_\_ n o i s e が 1 に設定された場合には、f \_\_ f l a g、all \_\_ 1 f、1 f \_\_ a t t \_\_ f l a g を伝送する必要がない。一方、all \_\_ n o i s e が 0 に設定された場合には、f \_\_ f l a g ( b ) を伝送しなければならず、B W E 領域 R 1 に属したバンド個数ほどの情報を伝達しなければならない。

【 0 0 5 2 】

all \_\_ 1 f 値が 0 に設定された場合には、1 f \_\_ a t t \_\_ f l a g 値は、0 に設定されて伝

10

20

30

40

50

送されない。all\_lf 値が 1 に設定された場合には、lf\_att\_flag の伝送を必要とする。かような相関関係によって、従属的に伝送されもし、コーデック構造簡素化のために、従属的な相関関係なしにも、伝送も可能である。結果として、スペクトル符号化部 370 では、全体許容ビットで伝送される BWE パラメータ及びコーディングバンド選択情報のために使用されるビットを除いて残った残余ビットを利用して、バンド別ビット割り当て及びコーディングを行う。

【0053】

【表1】

表 1

10

all_noise	f_flag	all_lf	lf_att_flag	使用ビット数
1	X	X	1X	1
0	# of BWE bands	1	1	3 + # of bands in R1
0	# of BWE bands	1	0	3 + # of bands in R1
0	# of BWE bands	0	X	2 + # of bands in R1

20

再び図3に戻り、多重化部390は、エネルギー符号化部340から提供される各バンド別エネルギー、コーディングバンド選択部360から提供されるBWE領域R1のコーディングバンド選択情報、スペクトル符号化部370から提供される、低周波数コーディング領域R0と、BWE領域R1とのうち、R2バンドの周波数ドメインコーディング結果、BWEパラメータ符号化部380から提供される、BWEパラメータを含むビットストリームを生成し、所定の記録媒体に保存するか、あるいは復号化端に伝送する。

30

【0054】

図6は、他の実施形態によるオーディオ符号化装置の構成を示したブロック図である。図6に図示されたオーディオ符号化装置は、基本的には、復号化端において、高周波数励起信号を生成するのに適用される加重値を推定するためのフレーム別励起タイプ情報を生成する構成要素と、フレーム別励起タイプ情報を含むビットストリームを生成する構成要素とからなる。残りの構成要素は、オプションとしてさらに追加される。

【0055】

図6に図示されたオーディオ符号化装置は、トランジェント検出部610、変換部620、エネルギー抽出部630、エネルギー符号化部640、スペクトル符号化部650、トナリティ算出部660、BWEパラメータ符号化部670及び多重化部680を含んでもよい。各構成要素は、少なくとも1つのモジュールに一体化され、少なくとも1つのプロセッサ(図示せず)によって具現される。ここでは、図3の符号化装置と同一の構成要素に係わる説明は省略する。

40

【0056】

図6において、スペクトル符号化部650は、変換部620から提供される低周波数信号のバンドについて、スペクトル係数の周波数ドメインコーディングを行う。残りの動作は、スペクトル符号化部370と同一である。

【0057】

トナリティ算出部660は、フレーム単位で、BWE領域R1のトナリティを算出する。

50

## 【 0 0 5 8 】

BWEパラメータ符号化部670は、トナリティ算出部660から提供されるBWE領域R1のトナリティを利用して、BWE励起タイプ情報あるいは励起クラス情報を生成して符号化する。一実施形態によれば、入力信号のモード情報をまず考慮し、BWE励起タイプを決定する。BWE励起タイプ情報は、フレーム別に伝送される。例えば、BWE励起タイプ情報が2ビットで構成される場合、0～3までの値を有する。0に行くほど、ランダムノイズに付加する加重値が大きくなり、3に行くほど、ランダムノイズに付加する加重値が小さくなる方式で割り当てる。一実施形態によれば、トナリティが高いほど、3に近い値を有するように設定し、トナリティが低いほど、0に近い値を有するように設定する。

10

## 【 0 0 5 9 】

図7は、一実施形態によって、BWEパラメータ符号化部の構成を示したブロック図である。図7に図示されたBWEパラメータ符号化部は、信号分類部710と、励起タイプ決定部730とを含んでもよい。

## 【 0 0 6 0 】

周波数ドメインのBWE方式は、時間ドメインコーディング・パートと結合されて適用される。時間ドメインコーディングには、主にCELP (code excited linear prediction) 方式が使用され、CELP方式で低周波帯域をコーディングし、周波数ドメインでのBWEではない時間ドメインでのBWE方式と結合されるように具現される。かような場合、全体的に、時間ドメインコーディングと、周波数ドメインコーディングとの間の適応的コーディング方式決定に基づいて、コーディング方式を選択的に適用することができる。適切なコーディング方式を選択するために信号分類を必要とし、一実施形態によれば、信号分類結果をさらに活用し、バンド別加重値が割り当てられる。

20

## 【 0 0 6 1 】

図7を参照すれば、信号分類部710においては、入力信号の特性をフレーム単位で分析し、現在フレームが音声信号であるか否かということ进行分类し、分類結果により、BWE励起タイプを決定する。信号分類処理は、公知された多様な方法、例えば、短区間特性及び/または長区間特性を利用して行われる。現在フレームが、時間ドメインコーディングが適切な方式である音声信号として分類される場合、高周波数信号の特性に基づいた方式より、固定された形態の加重値を付加する方式が音質向上に役に立つ。ところで、後述する図14及び図15のスイッチング構造の符号化装置に使用される通常的信号分類部1410, 1510は、複数個の以前フレームの結果と、現在フレームの結果とを組み合わせ、現在フレームの信号进行分类する。従って、中間結果として現在フレームだけの信号分類結果を活用して、たとえ最終的には、周波数ドメインコーディングが適用されたとしても、現在フレームが、時間ドメインコーディングが適切な方式であると出力された場合には、固定された加重値を設定して行う。例えば、かように現在フレームが、時間ドメインコーディングが適切な音声信号として分類される場合、BWE励起タイプは、例えば、2に設定される。

30

## 【 0 0 6 2 】

一方、信号分類部710の分類結果、現在フレームが音声信号として分類されない場合には、複数個の臨界値を利用して、BWE励起タイプを決定する。

40

## 【 0 0 6 3 】

励起タイプ決定部730は、3個の臨界値を設定し、トナリティの平均値の領域を4個に区分することにより、音声信号ではないと分類された現在フレームの4種BWE励起タイプを生成する。常に4種BWE励起タイプを限定するものではなく、場合により、3種あるいは2種である場合を使用することもでき、BWE励起タイプの個数に対応して使用される臨界値の個数及び値も調整される。かようなBWE励起タイプ情報に対応し、フレーム別加重値が割り当てられる。他の実施形態としては、フレーム別加重値は、さらに多くのビットを割り当てることができる場合には、バンド別加重値情報を抽出して伝送することもできる。

50

## 【 0 0 6 4 】

図 8 は、一実施形態によるオーディオ復号化装置の構成を示したブロック図である。図 8 に図示されたオーディオ復号化装置は、基本的には、フレーム単位で受信される励起タイプ情報を利用して、加重値を推定する構成要素、及びランダムノイズと、復号化された低周波数スペクトルとの間に加重値を適用し、高周波数励起信号を生成する構成要素からなる。残りの構成要素は、オプションとしてさらに追加される。

## 【 0 0 6 5 】

図 8 に図示されたオーディオ復号化装置は、逆多重化部 8 1 0、エネルギー復号化部 8 2 0、BWE パラメータ復号化部 8 3 0、スペクトル復号化部 8 4 0、第 1 逆正規化部 8 5 0、ノイズ付加部 8 6 0、励起信号生成部 8 7 0、第 2 逆正規化部 8 8 0 及び逆変換部 8 9 0 を含んでもよい。各構成要素は、少なくとも 1 つのモジュールに一体化され、少なくとも 1 つのプロセッサ（図示せず）によって具現される。

10

## 【 0 0 6 6 】

図 8 を参照すれば、逆多重化部 8 1 0 は、ビットストリームをパーズングし、符号化されたバンド別エネルギー、低周波数コーディング領域 R 0 と、BWE 領域 R 1 とのうち、R 2 バンドの周波数ドメインコーディング結果、BWE パラメータを抽出する。そのとき、コーディングバンド選択情報と、BWE パラメータとの相関関係により、コーディングバンド選択情報が、逆多重化部 8 1 0 からパーズングされるか、あるいは BWE パラメータ復号化部 8 3 0 からパーズングされる。

## 【 0 0 6 7 】

エネルギー復号化部 8 2 0 は、逆多重化部 8 1 0 から提供される符号化されたバンド別エネルギーを復号化し、バンド別逆量子化されたエネルギーを生成する。バンド別逆量子化されたエネルギーは、第 1 逆正規化部 8 5 0 及び第 2 逆正規化部 8 8 0 に提供される。また、バンド別に逆量子化されたエネルギーは、符号化端においてと同様に、ビット割り当てのために、スペクトル復号化部 8 4 0 に提供される。

20

## 【 0 0 6 8 】

BWE パラメータ復号化部 8 3 0 は、逆多重化部 8 1 0 から提供される BWE パラメータを復号化する。そのとき、コーディングバンド選択情報である  $f\_flag(b)$  が、BWE パラメータ、例えば、 $all\_noise$  と相関関係がある場合には、BWE パラメータ復号化部 8 3 0 において、BWE パラメータと共に復号化が行われる。一実施形態によれば、 $all\_noise$  情報、 $f\_flag$  情報、 $all\_lf$  情報、 $lf\_att\_flag$  情報が、表 1 でのような相関関係がある場合、順次に復号化を行う。かような相関関係は、他の方式に変更されも、変更時には、それに相応しい方式で、順次に復号化を行う。表 1 を例として挙げれば、 $all\_noise$  をまずパーズングし、1 であるか、あるいは 0 であるかということを確認する。もし  $all\_noise$  が 1 である場合には、 $f\_flag$  情報、 $all\_lf$  情報、 $lf\_att\_flag$  情報は、いずれも 0 に設定する。一方、 $all\_noise$  が 0 である場合には、 $f\_flag$  情報を、BWE 領域 R 1 に属したバンドの個数ほどパーズングし、次の  $all\_lf$  情報をパーズングする。もし  $all\_lf$  情報が 0 である場合には、 $lf\_att\_flag$  を 0 に設定し、1 である場合には、 $lf\_att\_flag$  情報をパーズングする。

30

## 【 0 0 6 9 】

一方、コーディングバンド選択情報である  $f\_flag(b)$  が BWE パラメータと相関関係がない場合には、逆多重化部 8 1 0 において、ビットストリームとしてパーズングされ、低周波数コーディング領域 R 0 と、BWE 領域 R 1 とのうち、R 2 バンドの周波数ドメインコーディング結果と共に、スペクトル復号化部 8 4 0 に提供される。

40

## 【 0 0 7 0 】

スペクトル復号化部 8 4 0 は、低周波数コーディング領域 R 0 の周波数ドメインコーディング結果を復号化する一方、コーディングバンド選択情報に対応して、WE 領域 R 1 のうち R 2 バンドの周波数ドメインコーディング結果を復号化する。そのために、エネルギー復号化部 8 2 0 から提供されるバンド別逆量子化されたエネルギーを利用して、全体許容ビットにおいて、パーズングされた BWE パラメータと、コーディングバンド選択情報の

50

ために使用されたビットとを除いて残った残余ビットを利用して、バンド別ビット割り当てを行う。スペクトル復号化のために、無損失復号化及び逆量子化が行われ、一実施形態によれば、FPCが使用される。すなわち、スペクトル復号化は、符号化端でのスペクトル符号化に使用されたものと同じの方式を使用して行われる。

【0071】

一方、BWE領域R1において、 $f\_flag(b)$ が1に設定されてビットが割り当てられ、実際パルスが割り当てられたバンドは、R2バンドに分類され、 $f\_flag(b)$ が0に設定され、ビット割り当てられていないバンドは、R3バンドに分類される。ところで、BWE領域R1において、 $f\_flag(b)$ が1に設定されており、スペクトル復号化を行うバンドであるにもかかわらず、ビット割り当てを行うことができず、FPCでコーディングされたパルス個数が0であるバンドが存在する。かように周波数ドメインコーディングを行うと設定されたR2バンドであるにもかかわらず、コーディングを行うことができないバンドは、R2バンドではないR3バンドに分類され、 $f\_flag(b)$ が0に設定された場合と同じ方式で処理される。

10

【0072】

第1逆正規化部850は、エネルギー復号化部820から提供されるバンド別逆量子化されたエネルギーを利用して、スペクトル復号化部840から提供される周波数ドメインデコーディング結果に対して逆正規化を行う。かような逆正規化処理は、復号化されたスペクトルのエネルギーを、各バンド別エネルギーにマッチングさせる過程に該当する。一実施形態によれば、逆正規化処理は、低周波数コーディング領域R0と、BWE領域R1

20

【0073】

ノイズ付加部860は、低周波数コーディング領域R0の復号化されたスペクトルの各バンドをチェックし、R4バンド及びR5バンドのうち一つに分離する。そのとき、R5に分離するバンドについては、ノイズを付加せず、R4に分離するバンドについて、ノイズを付加する。一実施形態によれば、ノイズを付加するときに使用されるノイズレベルは、バンド内に存在するパルスの密度を基に決定される。すなわち、ノイズレベルは、コーディングされたパルスのエネルギーを基に決定され、ノイズレベルを利用して、ランダムエネルギーを生成する。他の実施形態によれば、ノイズレベルは、符号化端から伝送される。一方、ノイズレベルは、 $lf\_att\_flag$ 情報を基に調整される。一実施形態によれば、下記のように、所定条件が満足されれば、ノイズレベルN1を、 $Att\_factor$ ほど修正する。

30

【0074】

```
if (all_noise==0 && all_lf==1 && lf_att_flag==1)
{
    ni_gain = ni_coef * NI * Att_factor;
}
else
{
    ni_gain = ni_coef * Ni;
}
```

40

ここで、 $ni\_gain$ は、最終ノイズに適用するゲインであり、 $ni\_coef$ は、ランダムシード(random seed)であり、 $Att\_factor$ は、調節定数である。

【0075】

励起信号生成部870は、BWE領域R1に属した各バンドについて、コーディングバンド選択情報に対応し、ノイズ付加部880から提供される復号化された低周波数スペクトルを利用して、高周波数励起信号を生成する。

【0076】

第2逆正規化部880は、エネルギー復号化部820から提供されるバンド別逆量子化されたエネルギーを利用して、励起信号生成部870から提供される高周波数励起信号に

50

ついて逆正規化を行い、高周波数スペクトルを生成する。かような逆正規化処理は、BWE領域R1のエネルギーを各バンド別エネルギーにマッチングさせる過程に該当する。

【0077】

逆変換部890は、第2逆正規化部880から提供される高周波数スペクトルについて逆変換を行い、時間ドメインの復号化された信号を生成する。

【0078】

図9は、一実施形態による励起信号生成部の細部的な構成を示すブロック図であり、BWE領域R1のR3バンド、すなわち、ビット割り当てがなされていないバンドに係わる励起信号生成を担当する。図9に図示された励起信号生成部は、加重値割当て部910、ノイズ信号生成部930及び演算部950を含んでもよい。各構成要素は、少なくとも1つのモジュールに一体化され、少なくとも1つのプロセッサ(図示せず)によって具現される。

10

【0079】

図9を参照すれば、加重値割当て部910は、バンド別に加重値を推定して割り当てる。ここで、加重値は、復号化された低周波数信号及びランダムノイズを基に生成された高周波数ノイズ信号とランダムノイズとを混合する比率を意味する。具体的には、HF(high frequency)励起信号 $H_e(f, k)$ は、下記数式(3)のように示すことができる。

【0080】

$$H_e(f, k) = (1 - W_s(f, k)) * H_n(f, k) + W_s(f, k) * R_n(f, k) \quad (3)$$

20

ここで、 $W_s(f, k)$ は、加重値を示し、 $f$ は、周波数インデックスを、 $k$ は、バンドインデックスを示す。 $H_n$ は、高周波数ノイズ信号を、 $R_n$ は、ランダムノイズをそれぞれ示す。

【0081】

一方、加重値 $W_s(f, k)$ は、1つのバンド内では、同一の値を有するが、バンド境界では、隣接バンドの加重値により、スムージングされるように処理される。

【0082】

加重値割当て部910では、BWEパラメータ、及びコーディングバンド選択情報、例えば、all\_noise情報、all\_lf情報、lf\_att\_flag情報、f\_flag情報を利用して、バンド別加重値を割り当てる。具体的には、all\_noiseが1であるならば、 $W_s(k) = w_0$ (全ての $k$ に対して)と割り当てられる。一方、all\_noiseが0であるならば、R2バンドについては、 $W_s(k) = w_4$ と割り当てる。all\_noiseが0であるならば、R3バンドについては、all\_lf = 1であり、lf\_att\_flag = 1であるならば、 $W_s(k) = w_3$ と割り当て、all\_lf = 1であり、lf\_att\_flag = 0であるならば、 $W_s(k) = w_2$ と割り当て、それ以外の場合には、 $W_s(k) = w_1$ と決定する。一実施形態によれば、 $w_0 = 1$ 、 $w_1 = 0.65$ 、 $w_2 = 0.55$ 、 $w_3 = 0.4$ 、 $w_4 = 0$ と割り当てる。望ましくは、 $w_0$ から $w_4$ に行くほど、小さい値を有するように設定する。

30

【0083】

加重値割当て部910は、推定されたバンド別加重値 $W_s(k)$ について、隣接バンドの加重値 $W_s(k-1)$ 、 $W_s(k+1)$ を考慮してスムージングを行う。スムージング結果、バンド $k$ について、周波数 $f$ によって、互いに異なる値を有する加重値 $W_s(f, k)$ が決定される。

40

【0084】

図12は、バンド境界において、加重値に係わるスムージング処理について説明するための図面である。図12を参照すれば、 $(K+2)$ バンドの加重値と、 $(K+1)$ バンドの加重値とが互いに異なるために、バンド境界でスムージングを行う必要がある。図10の例においては、 $(K+1)$ バンドは、スムージングを行わず、 $(K+2)$ バンドでのみスムージングを行う。その理由は、 $(K+1)$ バンドでの加重値 $W_s(K+1)$ が0であ

50

るために、 $(K + 1)$  バンドでスムージングを行えば、 $(K + 1)$  バンドでの加重値  $W_s$  ( $K + 1$ ) が 0 ではない値を有することになり、 $(K + 1)$  バンドにおいて、ランダムノイズまで考慮しなければならないからである。すなわち、加重値が 0 であるということは、当該バンドでは、高周波数励起信号の生成時、ランダムノイズを考慮しないということを示す。それは、極端なトーン信号である場合に該当し、ランダムノイズによって、ハーモニック信号のバレー区間にノイズが挿入され、ノイズ発生を防ぐためのものである。

【0085】

加重値割当て部 910 で決定された加重値  $W_s(f, k)$  は、高周波数ノイズ信号  $H_n$  と、ランダムノイズ  $R_n$  とに適用させるために、演算部 950 に提供される。

【0086】

ノイズ信号生成部 930 は、高周波数ノイズ信号を生成するためのものであり、ホワイトニング部 931 と、HF ノイズ生成部 933 とを含んでもよい。

【0087】

ホワイトニング部 931 は、逆量子化された低周波数スペクトルについて、ホワイトニングを行う。ホワイトニング処理は、公知された多様な方式を適用することができ、一例を挙げれば、逆量子化された低周波数スペクトルを、均一な複数のブロックに分け、ブロック別に、スペクトル係数の絶対値平均を求め、ブロックに属したスペクトル係数を平均して分ける方式が適用される。

【0088】

HF ノイズ生成部 933 は、ホワイトニング部 931 から提供される低周波数スペクトルを、高周波数、すなわち、BWE 領域  $R_1$  に輻射し、ランダムノイズとレベルをマッチングさせ、高周波数ノイズ信号を生成する。高周波数への輻射処理は、符号化端と復号化端とのあらかじめ設定された規則、パッチング、フォールディングあるいはコピーングによって行われ、ビット率によって選択的に適用する。レベルマッチング処理は、BWE 領域  $R_1$  の全体バンドについて、ランダムノイズの平均と、ホワイトニング処理された信号を高周波数に輻射した信号の平均とをマッチングさせることを意味する。一実施形態によれば、ホワイトニング処理された信号を高周波数に輻射した信号の平均が、ランダムノイズの平均より若干大きいように設定することもできる。その理由は、ランダムノイズは、ランダムな信号であるために、フラットな特性を有していると見られる、LF (low frequency) 信号は、相対的にダイナミックレンジが大きくなるので、大きさの平均をマッ

【0089】

演算部 950 は、ランダムノイズ及び高周波数ノイズ信号に対して加重値を適用し、バンド別高周波数励起信号を生成するためのものであり、第 1 乗算器 951 及び第 2 乗算器 953 と、加算器 955 とを含んでもよい。ここで、ランダムノイズ  $R_n$  は、公知された多様な方式で生成され、一例を挙げれば、ランダムシード (random seed) を利用して生成される。

【0090】

第 1 乗算器 951 は、ランダムノイズに第 1 加重値  $W_s(k)$  を乗算し、第 2 乗算器 953 は、高周波数ノイズ信号に第 2 加重値  $1 - W_s(k)$  を乗算し、加算器 955 は、第 1 乗算器 951 の乗算結果と、第 2 乗算器 953 の乗算結果とを加算し、バンド別高周波数励起信号を生成する。

【0091】

図 10 は、他の実施形態による励起信号生成部の細部的な構成を示すブロック図であり、BWE 領域  $R_1$  の  $R_2$  バンド、すなわち、ビット割り当てがなされているバンドに係わる励起信号生成処理を担当する。図 10 に図示された励起信号生成部は、調整パラメータ算出部 1010、ノイズ信号生成部 1030、レベル調整部 1050 及び演算部 1060 を含んでもよい。各構成要素は、少なくとも 1 つのモジュールに一体化され、少なくとも 1 つのプロセッサ (図示せず) によって具現される。

【0092】

10

20

30

40

50

図10を参照すれば、R2バンドは、FPCでコーディングされたパルスが存在するために、加重値を利用して高周波数励起信号を生成する処理に、レベル調整処理をさらに必要とする。周波数ドメイン符号化が行われたR2バンドの場合には、ランダムノイズは、付加しない。図10では、加重値 $W_s(k)$ が0である場合を例として挙げたものであり、加重値 $W_s(k)$ が0ではない場合には、図9のように、ノイズ信号生成部930において同一方式で、高周波数ノイズ信号を生成し、生成された高周波数ノイズ信号は、図10のノイズ信号生成部1030の出力にマッピングされる。すなわち、図10のノイズ信号生成部1030の出力は、図9のノイズ信号生成部1030の出力と同様になる。

【0093】

調整パラメータ算出部1010は、レベル調整に使用されるパラメータを算出するためのものである。まず、R2バンドについて逆量子化されたFPC信号を、 $C(k)$ と定義する場合、 $C(k)$ において、絶対値の最大値を選択し、選択された値を $A_p$ と定義し、FPCコーディング結果、0ではない値の位置は、 $CP_s$ と定義する。 $CP_s$ を除いた他の位置において、 $N(k)$ (ノイズ信号生成部830の出力)信号のエネルギーを求め、そのエネルギーを $E_n$ と定義する。 $E_n$ 値、 $A_p$ 値、及び符号化時に、 $f\_flag(b)$ 値を設定するために使用した $T_{th0}$ を基に、調整パラメータを、下記数式(4)のように求める。

【0094】

【数3】

$$\gamma = \sqrt{\frac{A_p^2}{E_n}} * 10^{-T_{th0}} * Att\ factor \quad (4)$$

ここで、 $Att\_factor$ は、調整定数である。

【0095】

演算部1060は、調整パラメータを、ノイズ信号生成部1030から提供されるノイズ信号 $N(k)$ に乘算し、高周波数励起信号を生成する。

【0096】

図11は、一実施形態による励起信号生成部の細部的な構成を示すブロック図であり、BWE領域R1の全体バンドに係わる励起信号生成を担当する。図11に図示された励起信号生成部は、加重値割当て部1110、ノイズ信号生成部1130及び演算部1150を含んでもよい。各構成要素は、少なくとも1つのモジュールに一体化され、少なくとも1つのプロセッサ(図示せず)によって具現される。ここで、ノイズ信号生成部1130及び演算部1150は、図9のノイズ信号生成部930及び演算部950と同一であるので、その説明を省略する。

【0097】

図11を参照すれば、加重値割当て部1110は、フレーム別に加重値を推定して割り当てる。ここで、加重値は、復号化された低周波数信号及びランダムノイズを基に生成された高周波数ノイズ信号及びランダムノイズを混合する比率を意味する。

【0098】

加重値割当て部1110は、ビットストリームからパーズングされたBWE励起タイプ情報を受信する。加重値割当て部1110には、BWE励起タイプが0であるならば、 $W_s(k) = w_{00}$ (全てのkに対して)に設定し、BWE励起タイプが1であるならば、 $W_s(k) = w_{01}$ (全てのkに対して)に設定し、BWE励起タイプが2であるならば、 $W_s(k) = w_{02}$ (全てのkに対して)に設定し、BWE励起タイプが3であるならば、 $W_s(k) = w_{03}$ (全てのkに対して)に設定する。一実施形態によれば、 $w_{00}$

10

20

30

40

50



$w_{00} = 0.8$ 、 $w_{01} = 0.5$ 、 $w_{02} = 0.25$ 、 $w_{03} = 0.05$ と割り当てる。 $w_{00}$ から $w_{03}$ に行くほど、小さくなるように設定する。

【0099】

一方、BWE領域R1において、特定周波数以後のバンドについては、BWE励起タイプ情報と係わりなく、同一の加重値を適用することもできる。一実施形態によれば、BWE領域R1において、特定周波数以後で最後のバンドを含む複数個のバンドについては、常に同一の加重値を使用して、特定周波数以下のバンドについては、BWE励起タイプ情報に基づいて加重値を生成する。例えば、12kHz以上の周波数が属するバンドである場合には、 $W_s(k)$ 値をいずれも $w_{02}$ に割り当てる。その結果、符号化端において、BWE励起タイプを決定するために、トナリティの平均値を求めるバンドの領域は、BWE領域R1内においても、特定周波数以下、すなわち、低周波数部分に限定されるために、演算の複雑度を低減させる。一実施形態によれば、BWE領域R1内において、特定周波数以下、すなわち、低周波数部分についてトナリティの平均を求めて励起タイプを決定し、決定された励起タイプを、そのままBWE領域R1内において、特定周波数以上、すなわち、高周波数部分に適用する。すなわち、フレーム単位に励起クラス情報を1個だけ伝送するために、励起クラス情報を推定する領域を狭く持って行けば、それほど正確度はさ、らに高くなり、復元音質の向上を図ることができる。一方、BWE領域R1において、高周波部分については、低周波数部分におけるところと同一の励起クラスを適用したとしても、音質劣化が起こる可能性は低くなる。また、BWE励起タイプ情報をバンド別に伝送する場合には、BWE励起タイプ情報を表示するために使用されるビットを節減することが可能である。

10

20

【0100】

次に、高周波数のエネルギーを、低周波数のエネルギー伝送方式とは異なる方式で、例えば、VQ (vector quantization) のような方式を適用すれば、低周波数のエネルギーは、スカラー量子化後、無損失符号化を使用して伝送し、高周波数のエネルギーは、他の方式で量子化を行って伝送される。かように処理する場合、低周波数コーディング領域R0の最後のバンドと、BWE領域R1の開始バンドとをオーバーラッピングする方式で構成する。また、BWE領域R1のバンド構成は、他の方式で構成し、さらに稠密なバンド割り当て構造を有する。

【0101】

例えば、低周波数コーディング領域R0の最後のバンドは、8.2kHzまで構成され、BWE領域R1の開始バンドは、8kHzから始まるように構成する。その場合、低周波数コーディング領域R0と、BWE領域R1との間にオーバーラッピング領域が生じる。その結果、オーバーラッピング領域には、2つの復号化されたスペクトルを生成する。一つは、低周波数の復号化方式を適用して生成したスペクトルであり、他の一つは、高周波数の復号化方式で生成したスペクトルである。2つのスペクトル、すなわち、低周波の復号化スペクトルと、高周波の復号化スペクトルとの遷移 (transition) がさらにスムージングになるように、オーバーラップアド (overlap add) 方式を適用する。すなわち、2つのスペクトルを同時に活用しながら、オーバーラッピングされた領域のうち低周波数側に近いスペクトルは、低周波方式で生成されたスペクトルの寄与分 (contribution) を高め、高周波数側に近いスペクトルは、高周波方式で生成されたスペクトルの寄与分を高め、オーバーラッピングされた領域を再構成する。

30

40

【0102】

例えば、低周波数コーディング領域R0の最後のバンドは、8.2kHzまで、BWE領域R1の開始バンドは、8kHzから始まる場合、32kHzサンプリングレートとして、640サンプルのスペクトルを構成すれば、320~327まで8個のスペクトルがオーバーラップされ、8個のスペクトルについては、下記数式(5)のように生成する。

【0103】

【数 4】

$$\bar{S}(k) = \bar{S}_l(k) \times w_o(k - L0) + (1 - w_o(k - L0)) \times \bar{S}_h(k) \quad (5)$$

ここで、

【0104】

【数 5】

$$\bar{S}_l(k)$$

10

は、低周波方式で復号化されたスペクトルを、

【0105】

【数 6】

$$\bar{S}_h(k)$$

20

は、高周波方式で復号化されたスペクトルを、L0は、高周波の開始スペクトル位置を、L0～L1は、オーバーラッピングされた領域を、w0は、寄与分をそれぞれ示す。

【0106】

図13は、一実施形態によって、復号化端でBWE処理した後、オーバーラッピング領域に存在するスペクトルを再構成するために使用される寄与分について説明する図面である。

【0107】

図13を参照すれば、 $w_0(k)$ は、 $w_{00}(k)$ 及び $w_{01}(k)$ を選択的に適用することができるが、 $w_{00}(k)$ は、低周波数と高周波数との復号化方式に、同一の加重値を適用するものであり、 $w_{01}(k)$ は、高周波数の復号化方式に、さらに大きい加重値を加える方式である。2つの $w_0(k)$ に係わる選択基準は、低周波数のオーバーラッピングバンドにおいて、FPCを使用したパルスが存在したか否かということの有無である。低周波数のオーバーラッピングバンドで、パルスが選択されてコーディングされた場合には、 $w_{00}(k)$ を活用し、低周波数で生成したスペクトルに係わる寄与分をL1近くまで有効にさせ、高周波数の寄与分を低減させる。基本的には、BWEを介して生成された信号のスペクトルよりは、実際コーディング方式によって生成されたスペクトルが、原信号との近接性側面において、さらに高くなる。それを活用して、オーバーラッピングバンドにおいて、原信号にさらに近接したスペクトルの寄与分を高める方式を適用することができ、従って、スムージング効果及び音質向上を図ることが可能である。

30

【0108】

図14は、一実施形態による、スイッチング構造のオーディオ符号化装置の構成を示したブロック図である。図14に図示された符号化装置は、信号分類部1410、TD(time domain)符号化部1420、TD拡張符号化部1430、FD(frequency domain)符号化部1440及びFD拡張符号化部1450を含んでもよい。

40

【0109】

信号分類部1410は、入力信号の特性を参照し、入力信号の符号化モードを決定する。信号分類部1410は、入力信号の時間ドメイン特性と、周波数ドメイン特性とを考慮し、入力信号の符号化モードを決定する。また、信号分類部1410は、入力信号の特性が、音声信号に該当する場合、入力信号に対して、TD符号化が行われるように決定し、入力信号の特性が、音声信号ではないオーディオ信号に該当する場合、入力信号に対して

50

、FD符号化が行われるように決定する。

【0110】

信号分類部1410に入力される入力信号は、ダウンサンプリング部(図示せず)によってダウンサンプリングされた信号になる。実施形態によれば、入力信号は、32kHzまたは48kHzのサンプリングレートを有する信号をリサンプリング(re-sampling)することにより、12.8kHzまたは16kHzのサンプリングレートを有する信号になる。そのとき、リサンプリングは、ダウンサンプリングになる。ここで、32kHzのサンプリングレートを有する信号は、SWB(super wide band)信号になり、そのとき、SWB信号は、FB(full band)信号になる。また、16kHzのサンプリングレートを有する信号は、WB(wide band)信号になる。

10

【0111】

それにより、信号分類部1410は、入力信号の低周波数領域に存在する低周波数信号の特性を参照し、低周波数信号の符号化モードをTDモードまたはFDモードのうちいずれか一つに決定する。

【0112】

TD符号化部1420は、入力信号の符号化モードがTDモードに決定されれば、入力信号について、CELP(code excited linear prediction)符号化を行う。TD符号化部1420は、入力信号から励起信号(excitation signal)を抽出し、抽出された励起信号を、ピッチ(pitch)情報に該当するadaptive codebook contribution及びfixed codebook contributionそれぞれを考慮して量子化する。

20

【0113】

他の実施形態によれば、TD符号化部1420は、入力信号から線形予測係数(LPC: linear prediction coefficient)を抽出し、抽出された線形予測係数を量子化し、量子化された線形予測係数を利用して、励起信号を抽出する過程をさらに含んでもよい。

【0114】

また、TD符号化部1420は、入力信号の特性による多様な符号化モードによって、CELP符号化を行う。例えば、CELP符号化部1420は、有声音符号化モード(voiced coding mode)、無声音符号化モード(unvoiced coding mode)、トランジション符号化モード(transition coding mode)または一般的な符号化モード(generic coding mode)のうちいずれか一つの符号化モードで、入力信号についてCELP符号化を行う。

30

【0115】

TD拡張符号化部1430は、入力信号の低周波信号についてCELP符号化が行われれば、入力信号の高周波信号について、拡張符号化を行う。例えば、TD拡張符号化部1430は、入力信号の高周波領域に対応する高周波信号の線形予測係数を量子化する。そのとき、TD拡張符号化部1430は、入力信号の高周波信号の線形予測係数を抽出し、抽出された線形予測係数を量子化することもできる。実施形態によれば、TD拡張符号化部1430は、入力信号の低周波信号の励起信号を使用して、入力信号の高周波信号の線形予測係数を生成することもできる。

【0116】

40

FD符号化部1440は、入力信号の符号化モードがFDモードに決定されれば、入力信号についてFD符号化を行う。そのために、入力信号について、MDCT(modified discrete cosine transform)などを利用して、周波数ドメインに変換し、変換された周波数スペクトルについて、量子化及び無損失符号化を行う。実施形態によれば、FPCを適用する。

【0117】

FD拡張符号化部1450は、入力信号の高周波数信号について、拡張符号化を行う。実施形態によれば、FD拡張符号化部1450は、低周波数スペクトルを利用して、高周波数拡張を行う。

【0118】

50

図15は、他の実施形態による、スイッチング構造のオーディオ符号化装置の構成を示したブロック図である。図15に図示された符号化装置は、信号分類部1510、LPC符号化部1520、TD符号化部1530、TD拡張符号化部1540、オーディオ符号化部1550及びオーディオ拡張符号化部1560を含んでもよい。

【0119】

図15を参照すれば、信号分類部1510は、入力信号の特性を参照し、入力信号の符号化モードを決定する。信号分類部1510は、入力信号の時間ドメイン特性と、周波数ドメイン特性とを考慮し、入力信号の符号化モードを決定する。信号分類部1510は、入力信号の特性が音声信号に該当する場合、入力信号について、TD符号化が行われるように決定し、入力信号の特性が音声信号ではないオーディオ信号に該当する場合、入力信号について、オーディオ符号化が行われるように決定する。

10

【0120】

LPC符号化部1520は、入力信号の低周波信号から、線形予測係数(LPC)を抽出し、抽出された線形予測係数を量子化する。実施形態によれば、LPC符号化部1520は、TCQ(trellis coded quantization)方式、MSVQ(multi-stage vector quantization)方式、LVQ(lattice vector quantization)方式などを使用して、線形予測係数を量子化することができるが、それらに限定されるものではない。

【0121】

具体的には、LPC符号化部1520は、32kHzまたは48kHzのサンプリングレートを有する入力信号をリサンプリングすることにより、12.8kHzまたは16kHzのサンプリングレートを有する入力信号の低周波信号から、線形予測係数を抽出する。LPC符号化部1520は、量子化された線形予測係数を利用して、LPC励起信号を抽出する過程をさらに含んでもよい。

20

【0122】

TD符号化部1530は、入力信号の符号化モードがTDモードに決定されれば、線形予測係数を利用して抽出されたLPC励起信号について、CELP符号化を行う。例えば、TD符号化部1530は、LPC励起信号について、ピッチ情報に該当するadaptive codebook contribution及びfixed codebook contributionそれぞれを考慮して量子化する。そのとき、LPC励起信号は、LPC符号化部1520、TD符号化部1530、及びそれらのうち少なくともいずれか一つにおいて生成される。

30

【0123】

TD拡張符号化部1540は、入力信号の低周波信号のLPC励起信号について、CELP符号化が行われれば、入力信号の高周波信号について、拡張符号化を行う。例えば、TD拡張符号化部1540は、入力信号の高周波信号の線形予測係数を量子化する。実施形態によれば、TD拡張符号化部1540は、入力信号の低周波信号のLPC励起信号を使用して、入力信号の高周波信号の線形予測係数を抽出することもできる。

【0124】

オーディオ符号化部1550は、入力信号の符号化モードが、オーディオモードに決定されれば、線形予測係数を利用して抽出されたLPC励起信号について、オーディオ符号化を行う。例えば、オーディオ符号化部1550は、線形予測係数を利用して抽出されたLPC励起信号を、周波数ドメインに変換し、変換されたLPC励起信号を量子化する。オーディオ符号化部1550は、周波数ドメインに変換された励起スペクトルについて、FPC方式またはlattice VQ(LVQ)方式による量子化を行うこともできる。

40

【0125】

さらに、オーディオ符号化部1550は、LPC励起信号について、量子化を行うにあたり、ビットの余裕がある場合、adaptive codebook contribution及びfixed codebook contributionのTDコーディング情報をさらに考慮して量子化することもできる。

【0126】

FD拡張符号化部1560は、入力信号の低周波信号のLPC励起信号について、オーディオ符号化が行われれば、入力信号の高周波信号について、拡張符号化を行う。すなわ

50

ち、FD拡張符号化部1560は、低周波数スペクトルを利用して、高周波数拡張を行う。

【0127】

図14及び図15に図示されたFD拡張符号化部1450, 1560は、図3及び図6の符号化装置でもって具現される。

【0128】

図16は、一実施形態による、スイッチング構造のオーディオ復号化装置の構成を示したブロック図である。図16を参照すれば、復号化装置は、モード情報検査部1610、TD復号化部1620、TD拡張復号化部1630、FD復号化部1640及びFD拡張復号化部1650を含んでもよい。

10

【0129】

モード情報検査部1610は、ビットストリームに含まれたフレームそれぞれに係わるモード情報を検査する。モード情報検査部1610は、ビットストリームから、モード情報をパーズングし、パーズング結果による現在フレームの符号化モードによって、TD復号化モードまたはFD復号化モードのうちいずれか1つの復号化モードで、スイッチング作業を行う。

【0130】

具体的には、モード情報検査部1610は、ビットストリームに含まれたフレームそれぞれについて、TDモードで符号化されたフレームは、CELP復号化が行われるようにスイッチングし、FDモードで符号化されたフレームは、FD復号化が行われるようにス

20

【0131】

TD復号化部1620は、検査結果によって、CELP符号化されたフレームについてCELP復号化を行う。例えば、TD復号化部1620は、ビットストリームに含まれた線形予測係数を復号化し、adaptive codebook contribution及びfixed codebook contributionに係わる復号化を行い、復号化遂行結果を合成し、低周波数に係わる復号化信号である低周波信号を生成する。

【0132】

TD拡張復号化部1630は、CELP復号化が行われた結果、及び低周波信号の励起信号のうち少なくとも一つを利用して、高周波数に係わる復号化信号を生成する。そのとき、低周波信号の励起信号は、ビットストリームに含まれる。また、TD拡張復号化部1630は、高周波数に係わる復号化信号である高周波信号を生成するために、ビットストリームに含まれた高周波信号に係わる線形予測係数情報を活用する。

30

【0133】

実施形態によれば、TD拡張復号化部1630は、生成された高周波信号を、TD復号化部1620で生成された低周波信号と合成し、復号化された信号を生成する。そのとき、TD拡張復号化部1620は、復号化された信号を生成するために、低周波信号及び高周波信号のサンプリングレートが同一になるように変換する作業をさらに行う。

【0134】

FD復号化部1640は、検査結果によって、FD符号化されたフレームについて、FD復号化を行う。実施形態によるFD復号化部1640は、ビットストリームに含まれた以前フレームのモード情報を参照し、無損失復号化及び逆量子化を行うこともできる。そのとき、FPC復号化が適用され、FPC復号化が行われた結果、所定周波数バンドにノイズを付加する。

40

【0135】

FD拡張復号化部1650は、FD復号化部1640において、FPC復号化及び/またはノイズフィーリングが行われた結果を利用して、高周波数拡張復号化を行う。FD拡張復号化部1650は、低周波帯域について復号化された周波数スペクトルのエネルギーを逆量子化し、高周波帯域幅拡張の多様なモードによって、低周波信号を利用して、高周波信号の励起信号を生成し、生成された励起信号のエネルギーが逆量子化されたエネルギ

50

ーに対称になるようにゲインを適用することにより、復号化された高周波信号を生成する。例えば、高周波帯域幅拡張の多様なモードは、ノルマル(normal)モード、ハーモニック(harmonic)モードまたはノイズ(noise)モードのうちいずれか1つのモードになる。

【0136】

図17は、他の実施形態による、スイッチング構造のオーディオ復号化装置の構成を示したブロック図である。図17を参照すれば、復号化装置は、モード情報検査部1710、LPC復号化部1720、TD復号化部1730、TD拡張復号化部1740、オーディオ復号化部1750及びFD拡張復号化部1760を含んでもよい。

【0137】

モード情報検査部1710は、ビットストリームに含まれたフレームそれぞれに係わるモード情報を検査する。例えば、モード情報検査部1710は、符号化されたビットストリームから、モード情報をパーズングし、パーズング結果による現在フレームの符号化モードによって、TD復号化モードまたはオーディオ復号化モードのうちいずれか1つの復号化モードで、スイッチング作業を行う。

【0138】

具体的には、モード情報検査部1710は、ビットストリームに含まれたフレームそれぞれについて、TDモードで符号化されたフレームは、CELP復号化が行われるようにスイッチングし、オーディオ符号化モードで符号化されたフレームは、オーディオ復号化が行われるようにスイッチングする。

【0139】

LPC復号化部1720は、ビットストリームに含まれたフレームについて、LPC復号化を行う。

【0140】

TD復号化部1730は、検査結果によって、CELP符号化されたフレームについて、CELP復号化を行う。例を挙げて説明すれば、TD復号化部1730は、adaptive codebook contribution及びfixed codebook contributionに係わる復号化を行い、復号化遂行結果を合成し、低周波数に係わる復号化信号である低周波信号を生成する。

【0141】

TD拡張復号化部1740は、CELP復号化が行われた結果、及び低周波信号の励起信号のうち少なくとも一つを利用して、高周波数に係わる復号化信号を生成する。そのとき、低周波信号の励起信号は、ビットストリームに含まれる。また、TD拡張復号化部1740は、高周波数に係わる復号化信号である高周波信号を生成するために、LPC復号化部1720で復号化された線形予測係数情報を利用する。

【0142】

また、実施形態によればTD拡張復号化部1740は、生成された高周波信号を、TD復号化部1730で生成された低周波信号と合成し、復号化された信号を生成する。そのとき、TD拡張復号化部1740は、復号化された信号を生成するために、低周波信号及び高周波信号のサンプリングレートが同一になるように変換する作業をさらに行う。

【0143】

オーディオ復号化部1750は、検査結果によって、オーディオ符号化されたフレームについて、オーディオ復号化を行う。例えば、オーディオ復号化部1750は、ビットストリームを参照し、時間ドメイン寄与分が存在する場合、時間ドメイン寄与分及び周波数ドメイン寄与分を考慮して復号化を行い、時間ドメイン寄与分が存在しない場合、周波数ドメイン寄与分を考慮して復号化を行う。

【0144】

また、オーディオ復号化部1750は、FPCまたはLVQで量子化された信号について、IDCTなどを利用して、時間ドメインに変換して復号化された低周波数励起信号を生成し、生成された励起信号を、逆量子化されたLPC係数と合成し、復号化された低周波数信号を生成する。

10

20

30

40

50

## 【 0 1 4 5 】

F D 拡張復号化部 1 7 6 0 は、オーディオ復号化が行われた結果を利用して、拡張復号化を行う。例えば、F D 拡張復号化部 1 7 6 0 は、復号化された低周波数信号を、高周波数拡張復号化に適するサンプリングレートに変換し、変換された信号について、M D C T のような周波数変換を行う。F D 拡張復号化部 1 7 6 0 は、変換された低周波数スペクトルのエネルギーを逆量子化し、高周波帯域幅拡張の多様なモードによって、低周波信号を利用して、高周波信号の励起信号を生成し、生成された励起信号のエネルギーが、逆量子化されたエネルギーに対称になるようにゲインを適用することにより、復号化された高周波信号を生成する。例えば、高周波帯域幅拡張の多様なモードは、ノルマルモード、転移モード、ハーモニックモード、またはノイズモードのうちいずれか 1 つのモードになる。

10

## 【 0 1 4 6 】

また、F D 拡張復号化部 1 7 6 0 は、復号化された高周波信号について、inverse M D C T を利用して、時間ドメインに変換し、時間ドメインに変換された信号について、オーディオ復号化部 1 7 5 0 で生成された低周波信号とサンプリングレートを合わせるための変換作業を行った後、低周波信号と、変換作業が行われた信号とを合成する。

## 【 0 1 4 7 】

図 1 6 及び図 1 7 に図示された F D 拡張復号化部 1 6 5 0 , 1 7 6 0 は、図 8 の復号化装置でもって具現される。

## 【 0 1 4 8 】

図 1 8 は、本発明の一実施形態による、符号化モジュールを含むマルチメディア機器の構成を示したブロック図である。図 1 8 に図示されたマルチメディア機器 1 8 0 0 は、通信部 1 8 1 0 及び符号化モジュール 1 8 3 0 を含んでもよい。また、符号化の結果として得られるオーディオビットストリームの用途によって、オーディオビットストリームを保存する保存部 1 8 5 0 をさらに含んでもよい。また、マルチメディア機器 1 8 0 0 は、マイクロフォン 1 8 7 0 をさらに含んでもよい。すなわち、保存部 1 8 5 0 とマイクロフォン 1 8 7 0 は、オプションとして具備される。一方、図 1 8 に図示されたマルチメディア機器 1 8 0 0 は、任意の復号化モジュール（図示せず）、例えば、一般的な復号化機能を遂行する復号化モジュール、あるいは本発明の一実施形態による復号化モジュールをさらに含んでもよい。ここで、符号化モジュール 1 8 3 0 は、マルチメディア機器 1 8 0 0 に具備される他の構成要素（図示せず）と共に一体化され、少なくとも一つ以上のプロセッサ（図示せず）によって具現される。

20

30

## 【 0 1 4 9 】

図 1 8 を参照すれば、通信部 1 8 1 0 は、外部から提供されるオーディオ及び符号化されたビットストリームのうち少なくとも一つを受信したり、あるいは復元されたオーディオ、及び符号化モジュール 1 8 3 0 の符号化結果として得られるオーディオビットストリームのうち少なくとも一つを送信したりする。

## 【 0 1 5 0 】

通信部 1 8 1 0 は、無線インターネット、無線イントラネット、無線電話網、無線 L A N (local area network)、W i - F i (wireless fidelity)、W F D (W i - F i direct)、3 G (generation)、4 G (4 generation)、ブルートゥース、赤外線通信 (I r D A : infrared data association)、R F I D (radio frequency identification)、U W B (ultra-wideband)、ジグビー（（登録商標）Zigbee）、N F C (near field communication) のような無線ネットワーク、または有線電話網、有線インターネットのような有線ネットワークを介して、外部のマルチメディア機器とデータを送受信することができるように構成される。

40

## 【 0 1 5 1 】

符号化モジュール 1 8 3 0 は、一実施形態によれば、通信部 1 8 1 0 あるいはマイクロフォン 1 8 7 0 を介して提供される時間ドメインのオーディオ信号について、図 1 4 あるいは図 1 5 の符号化装置を利用した符号化を行う。また、F D 拡張符号化は、図 3 あるいは図 6 の符号化装置を利用する。

50

## 【 0 1 5 2 】

保存部 1 8 5 0 は、符号化モジュール 1 8 3 0 で生成される符号化されたビットストリームを保存する。一方、保存部 1 8 5 0 は、マルチメディア機器 1 8 0 0 の運用に必要な多様なプログラムを保存する。

## 【 0 1 5 3 】

マイクロフォン 1 8 7 0 は、ユーザあるいは外部のオーディオ信号を、符号化モジュール 1 8 3 0 に提供する。

## 【 0 1 5 4 】

図 1 9 は、本発明の一実施形態による、復号化モジュールを含むマルチメディア機器の構成を示したブロック図である。図 1 9 に図示されたマルチメディア機器 1 9 0 0 は、通信部 1 9 1 0 と復号化モジュール 1 9 3 0 とを含んでもよい。また、復号化の結果として得られる復元されたオーディオ信号の用途によって、復元されたオーディオ信号を保存する保存部 1 9 5 0 をさらに含んでもよい。また、マルチメディア機器 1 9 0 0 は、スピーカ 1 9 7 0 をさらに含んでもよい。すなわち、保存部 1 9 5 0 とスピーカ 1 9 7 0 は、オプションとして具備される。一方、図 1 9 に図示されたマルチメディア機器 1 9 0 0 は、任意の符号化モジュール（図示せず）、例えば、一般的な符号化機能を遂行する符号化モジュール、あるいは本発明の一実施形態による、符号化モジュールをさらに含んでもよい。ここで、復号化モジュール 1 9 3 0 は、マルチメディア機器 1 9 0 0 に具備される他の構成要素（図示せず）と共に一体化され、少なくとも 1 つの以上のプロセッサ（図示せず）によって具現される。

## 【 0 1 5 5 】

図 1 9 を参照すれば、通信部 1 9 1 0 は、外部から提供される符号化されたビットストリーム及びオーディオ信号のうち少なくとも一つを受信したり、あるいは復号化モジュール 1 9 3 0 の復号化結果として得られる復元されたオーディオ信号、及び符号化の結果として得られるオーディオビットストリームのうち少なくとも一つを送信したりする。一方、通信部 1 9 1 0 は、図 1 8 の通信部 1 8 1 0 と実質的に類似して具現される。

## 【 0 1 5 6 】

復号化モジュール 1 9 3 0 は、一実施形態によれば、通信部 1 9 1 0 を介して提供されるビットストリームを受信し、ビットストリームに含まれたオーディオスペクトルについて、図 1 6 あるいは図 1 7 の復号化装置を利用した復号化を行う。また、FD 拡張復号化は、図 8 の復号化装置を利用することができ、具体的には、図 9 ないし図 1 1 に図示された高周波数励起信号生成部を利用する。

## 【 0 1 5 7 】

保存部 1 9 5 0 は、復号化モジュール 1 9 3 0 で生成される復元されたオーディオ信号を保存する。一方、保存部 1 9 5 0 は、マルチメディア機器 1 9 0 0 の運用に必要な多様なプログラムを保存する。

## 【 0 1 5 8 】

スピーカ 1 9 7 0 は、復号化モジュール 1 9 3 0 で生成される復元されたオーディオ信号を外部に出力する。

## 【 0 1 5 9 】

図 2 0 は、本発明の一実施形態による、符号化モジュール及び復号化モジュールを含むマルチメディア機器の構成を示したブロック図である。

## 【 0 1 6 0 】

図 2 0 に図示されたマルチメディア機器 2 0 0 0 は、通信部 2 0 1 0、符号化モジュール 2 0 2 0 及び復号化モジュール 2 0 3 0 を含んでもよい。また、符号化の結果として得られるオーディオビットストリーム、あるいは復号化の結果として得られる復元されたオーディオ信号の用途によって、オーディオビットストリームあるいは復元されたオーディオ信号を保存する保存部 2 0 4 0 をさらに含んでもよい。また、マルチメディア機器 2 0 0 0 は、マイクロフォン 2 0 5 0 あるいはスピーカ 2 0 6 0 をさらに含んでもよい。ここで、符号化モジュール 2 0 2 0 と復号化モジュール 2 0 3 0 は、マルチメディア機器 2 0

10

20

30

40

50



00に具備される他の構成要素(図示せず)と共に一体化され、少なくとも一つ以上のプロセッサ(図示せず)によって具現される。

【0161】

図20に図示された各構成要素は、図18に図示されたマルチメディア機器1800の構成要素、あるいは図19に図示されたマルチメディア機器1900の構成要素と重複するので、その詳細な説明は省略する。

【0162】

図18ないし図20に図示されたマルチメディア機器1800, 1900, 2000には、電話、モバイルフォンなどを含む音声通信専用端末; TV (television)、MP3プレーヤなどを含む放送専用装置または音楽専用装置、あるいは音声通信専用端末と、放送専用装置あるいは音楽専用装置との融合端末装置が含まれるが、それらに限定されるものではない。また、マルチメディア機器1800, 1900, 2000は、クライアント、サーバ、あるいはクライアントとサーバとの間に配置される変換器として使用される。

【0163】

一方、マルチメディア機器1800, 1900, 2000が、例えば、モバイルフォンである場合、図示されていないが、キーパッドのようなユーザ入力部、ユーザ・インターフェースあるいはモバイルフォンで処理される情報をディスプレイするディスプレイ部、モバイルフォンの全般的な機能を制御するプロセッサをさらに含んでもよい。また、モバイルフォンは、撮像機能を有するカメラ部と、モバイルフォンで必要とする機能を遂行する少なくとも一つ以上の構成要素とをさらに含んでもよい。

【0164】

一方、マルチメディア機器1800, 1900, 2000が、例えば、TVである場合、図示されていないが、キーパッドのようなユーザ入力部、受信された放送情報をディスプレイするディスプレイ部、TVの全般的な機能を制御するプロセッサをさらに含んでもよい。また、TVは、TVで必要とする機能を遂行する少なくとも一つ以上の構成要素をさらに含んでもよい。

【0165】

前記実施形態による方法は、コンピュータで実行されるプログラムでもって作成可能であり、コンピュータで読み取り可能な記録媒体を利用して、前記プログラムを動作させる汎用デジタル・コンピュータで具現される。また、前述の本発明の実施形態で使用されるデータ構造、プログラム命令あるいはデータファイルは、コンピュータで読み取り可能な記録媒体に多様な手段を介して記録される。コンピュータで読み取り可能な記録媒体は、コンピュータシステムによって読み取り可能なデータが保存される全種の保存装置を含む。コンピュータで読み取り可能な記録媒体の例としては、ハードディスク、フロッピー(登録商標)ディスク及び磁気テープのような磁気媒体(magnetic media); CD(compact disc) - ROM(read-only memory)、DVD(digital versatile disc)のような光記録媒体(optical media); フロプティカルディスク(floptical disk)のような磁気・光媒体(magneto-optical media); 及びROM、RAM(random-access memory)、フラッシュメモリのようなプログラム命令を保存して遂行するように特別に構成されたハードウェア装置; が含まれる。また、コンピュータで読み取り可能な記録媒体は、プログラム命令、データ構造などを指定する信号を伝送する伝送媒体でもある。プログラム命令の例としては、コンパイラによって作られるような機械語コードだけでなく、インタープリタなどを使用して、コンピュータによって実行される高級言語コードを含んでもよい。

【0166】

以上のように、本発明の一実施形態は、たとえ限定された実施形態と図面とによって説明されたにしても、本発明の一実施形態は、前述の実施形態に限定されるものではなく、それは、本発明が属する分野で当業者であるならば、かような記載から、多様な修正及び変形が可能であろう。従って、本発明のスコープは、前述の説明ではなく、特許請求の範囲に示されており、その均等または等価的変形は、いずれも本発明技術的思想の範疇に属

10

20

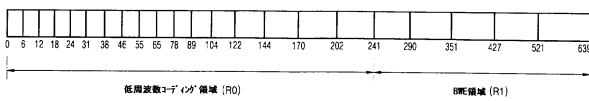
30

40

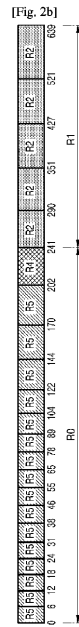
50

するものである。

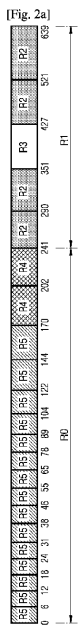
【図 1】



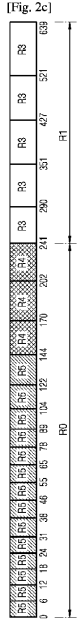
【図 2 b】



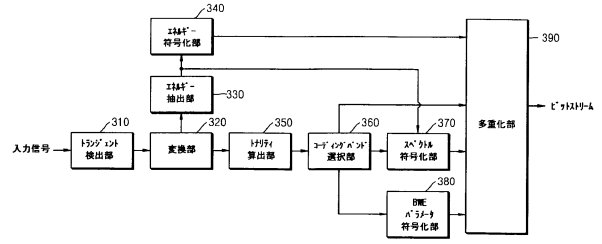
【図 2 a】



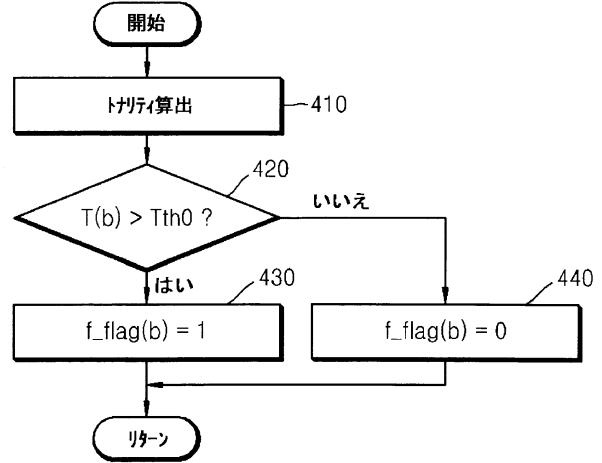
【図 2c】



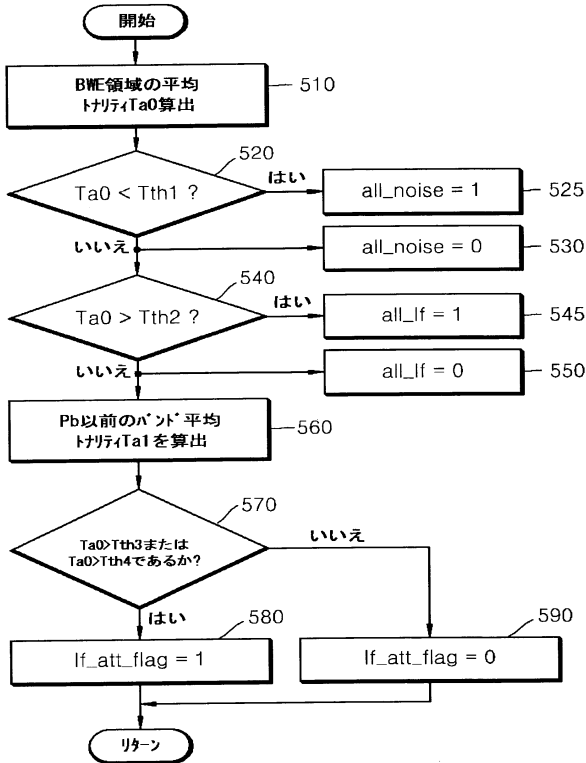
【図 3】



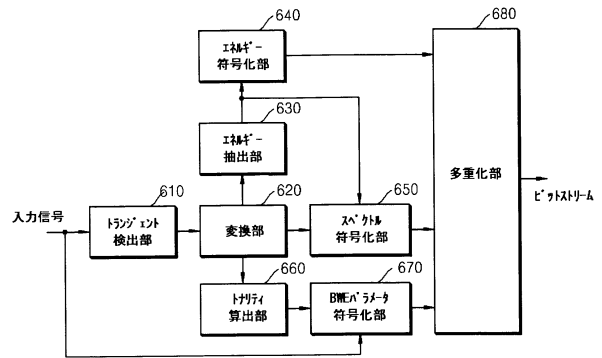
【図 4】



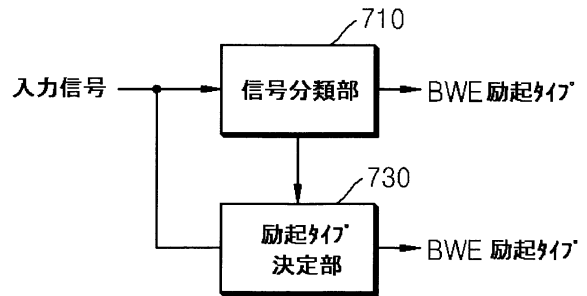
【図 5】



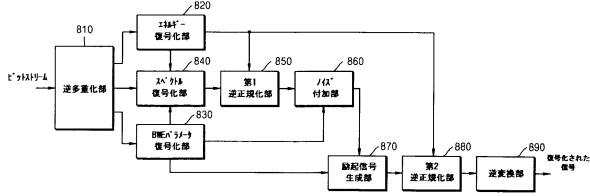
【図 6】



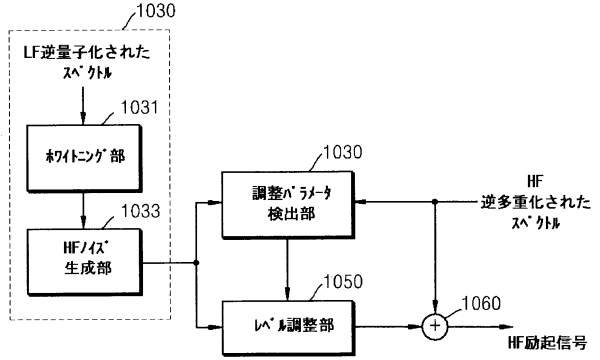
【図 7】



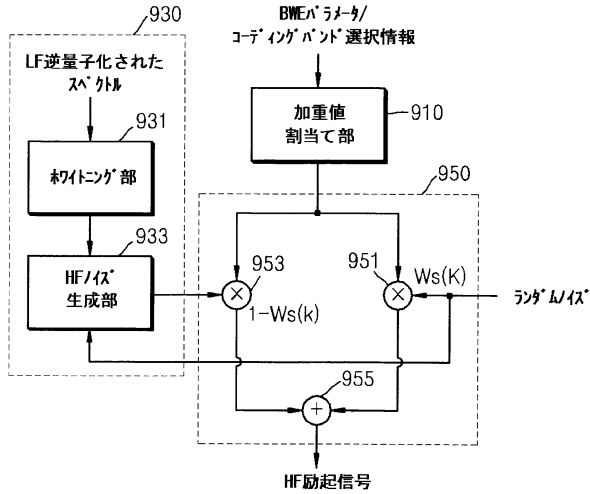
【図 8】



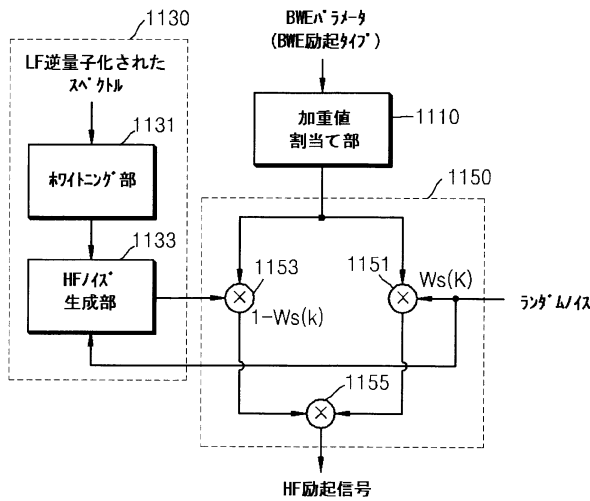
【図 10】



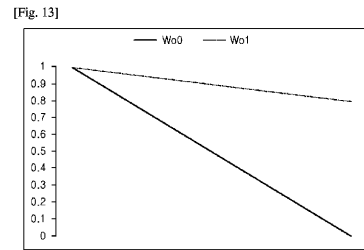
【図 9】



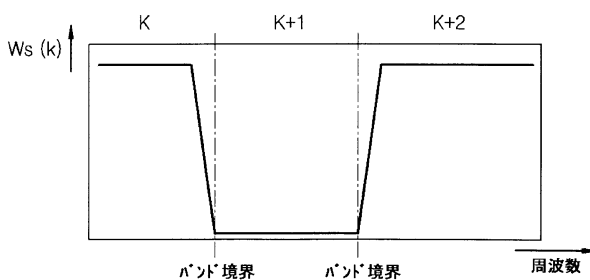
【図 11】



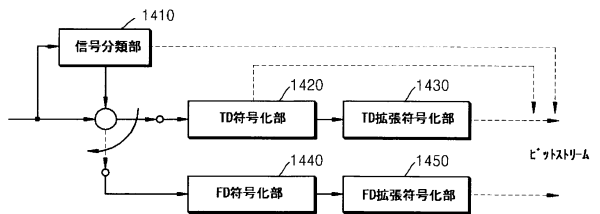
【図 13】



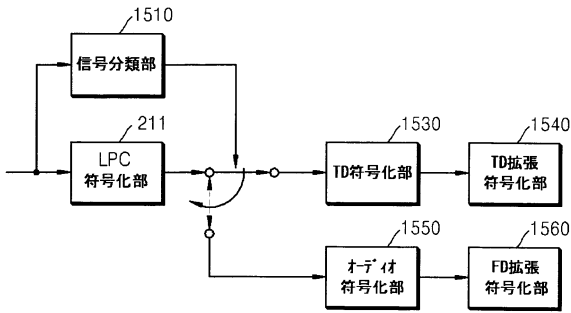
【図 12】



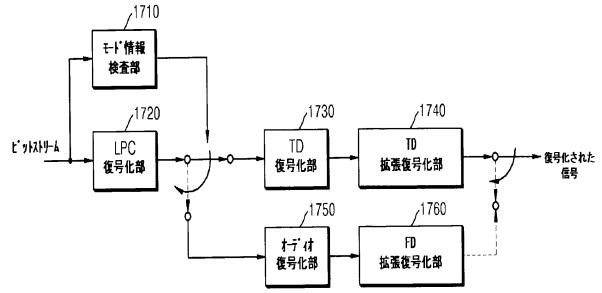
【図 14】



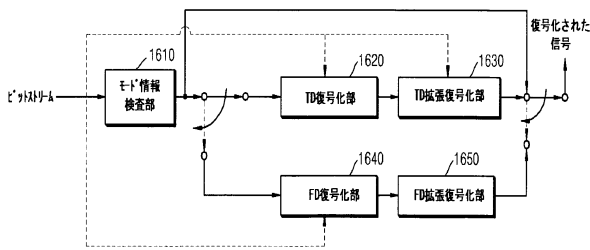
【図15】



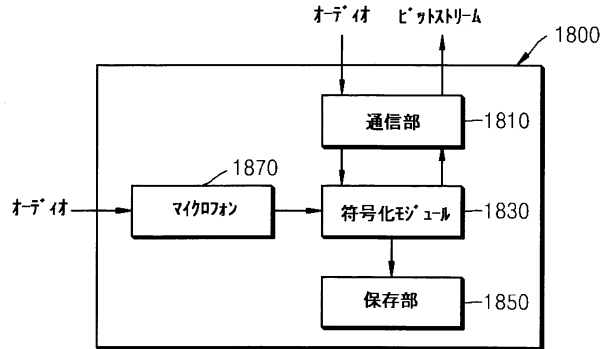
【図17】



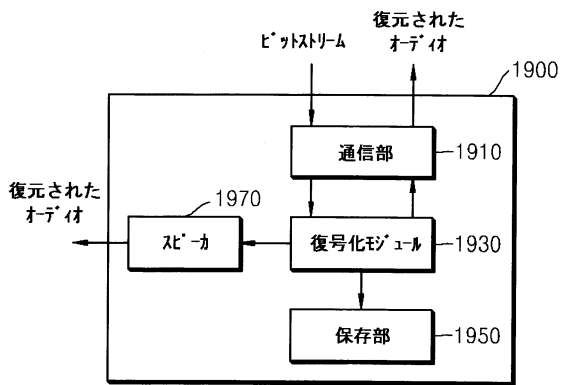
【図16】



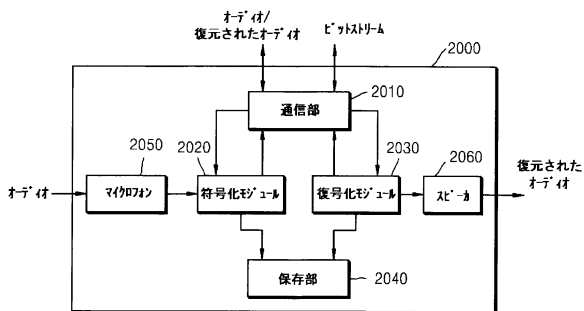
【図18】



【図19】



【図20】



---

フロントページの続き

(72)発明者 ジュ,キ-ヒョン

大韓民国 143-192 ソウル クァンジン-グ ジャヤン 2-ドン ハニャンアパート  
6-705 (番地なし)

審査官 上田 雄

(56)参考文献 欧州特許出願公開第02273493 (EP, A1)

国際公開第2012/020828 (WO, A1)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00 - 19/26

G10L 21/00 - 21/18