

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6189831号

(P6189831)

(45) 発行日 平成29年8月30日 (2017. 8. 30)

(24) 登録日 平成29年8月10日 (2017. 8. 10)

(51) Int. Cl.

G 1 O L 19/035 (2013.01)

F I

G 1 O L 19/035

B

請求項の数 4 (全 32 頁)

(21) 出願番号	特願2014-511291 (P2014-511291)	(73) 特許権者	503447036
(86) (22) 出願日	平成24年5月14日 (2012. 5. 14)		サムスン エレクトロニクス カンパニー
(65) 公表番号	特表2014-514617 (P2014-514617A)		リミテッド
(43) 公表日	平成26年6月19日 (2014. 6. 19)		大韓民国・1 6 6 7 7・キョンギード・ス
(86) 国際出願番号	PCT/KR2012/003777		ウォン・シ・ヨン・ト・ン・ク・サムスン・ロ
(87) 国際公開番号	W02012/157932		・ 1 2 9
(87) 国際公開日	平成24年11月22日 (2012. 11. 22)	(74) 代理人	100107766
審査請求日	平成27年5月14日 (2015. 5. 14)		弁理士 伊東 忠重
(31) 優先権主張番号	61/485, 741	(74) 代理人	100070150
(32) 優先日	平成23年5月13日 (2011. 5. 13)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
(31) 優先権主張番号	61/495, 014		弁理士 大貫 進介
(32) 優先日	平成23年6月9日 (2011. 6. 9)		
(33) 優先権主張国	米国 (US)		
前置審査		最終頁に続く	

(54) 【発明の名称】 ビット割り当て方法及び記録媒体

(57) 【特許請求の範囲】

【請求項 1】

フレームの使用可能なビット数を考慮して、前記フレーム内の各サブバンドに割り当てられるビット数を初期的に小数点単位で推定する段階と、

前記各サブバンドに割り当てられるビット数を最終的に決定するために、前記フレーム内の各サブバンドに割り当てられるビット数の合計が前記使用可能なビット数と同じになるまで、前記各サブバンドに割り当てられるビット数を0でないビット数を有するサブバンドに再分配する段階と、を含み、

前記初期的に推定する段階では、前記各サブバンドに割り当てられるビット数と推定されたビット数が0より小さい場合、0に設定し、

前記各サブバンドに最終的に割り当てられるビット数は、各サブバンドが必要とする最小ビット数と同じか、または大きく、

前記各段階は、少なくとも一つのプロセッサによって行われるビット割り当て方法。

【請求項 2】

前記推定する段階は、各サブバンドのスペクトルエネルギーに基づいて行われる請求項1に記載のビット割り当て方法。

【請求項 3】

前記推定する段階は、下記の数式

【数 4 3】

$$L_b = \frac{1}{C} \left(n_b - \frac{\sum_b N_b n_b - CB}{\sum_b N_b} \right)$$

(ここで、 L_b は、サブバンド b で各サンプルに割り当てられたビット数、 C は、 dB スケール値、 n_b は、サブバンド b でログスケールに量子化された Norm 値、 N_b は、サブバンド b のサンプル数、 B は、与えられたフレームで使用可能な全体ビット数をそれぞれ示す。)

を用いて行われる請求項1に記載のビット割り当て方法。

【請求項4】

請求項1ないし3のうちいずれか一つに記載の方法をコンピュータに実行させるための命令語を記録したコンピュータで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／復号化に係り、さらに具体的には、知覚的に重要な周波数領域にサブバンド単位で効率的にビットを割り当てる方法、オーディオ符号化方法及び装置、オーディオ復号化方法及び装置、記録媒体とこれを採用するマルチメディア機器に関する。

【背景技術】

【0002】

オーディオ信号を符号化または復号化する場合、限定されたビットを効率的に用いることで、該ビット範囲で最上の音質を持つオーディオ信号を復元することが求められる。特に、低いビット率では特定周波数領域にビットが集中せず、知覚的に重要な周波数領域にビットが均一に割り当てられるようにオーディオ信号を符号化及び復号化する技術が必要とする。

【発明の概要】

【発明が解決しようとする課題】

【0003】

本発明が解決しようとする課題は、知覚的に重要な周波数領域にサブバンド単位で効率的にビットを割り当てる方法及び装置、オーディオ符号化方法及び装置、オーディオ復号化方法及び装置、記録媒体とこれを採用するマルチメディア機器を提供することである。

【0004】

本発明が解決しようとする他の課題は、知覚的に重要な周波数領域に低い複雑度でサブバンド単位で効率的にビットを割り当てる方法及び装置、オーディオ符号化方法及び装置、オーディオ復号化方法及び装置、記録媒体とこれを採用するマルチメディア機器を提供することである。

【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ビット割り当て方法は、与えられたフレームに対して使用可能なビット数範囲内で、所定の周波数バンドに存在するスペクトルの SNR を最大化させるように、前記周波数バンド単位で割り当てビット数を小数点単位で定める段階と、前記周波数バンド単位で定められた割り当てビット数を調整する段階と、を含む。

【0006】

10

20

30

40

50

前記課題を解決するための本発明の一実施形態によるビット割り当て装置は、時間ドメインのオーディオ信号を周波数ドメインのオーディオスペクトルに変換する変換部と、前記オーディオスペクトルで与えられたフレームに含まれた周波数バンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定し、スペクトルエネルギーを用いて割り当てビット数を小数点単位で推定し、前記割り当てビット数が前記許容ビット数を超えないように調整するビット割り当て部と、を備える。

【0007】

前記課題を解決するための本発明の一実施形態によるオーディオ符号化装置は、時間ドメインのオーディオ信号を周波数ドメインのオーディオスペクトルに変換する変換部と、前記オーディオスペクトルで与えられたフレームに対して使用可能なビット数範囲内で、所定の周波数バンドに存在するスペクトルのSNRを最大化させるように前記周波数バンド単位で割り当てビット数を小数点単位で定め、前記周波数バンド単位で定められた割り当てビット数を調整するビット割り当て部と、前記周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いて前記オーディオスペクトルを符号化する符号化部と、を備える。

10

【0008】

前記課題を解決するための本発明の一実施形態によるオーディオ符号化装置は、時間ドメインのオーディオ信号を周波数ドメインのオーディオスペクトルに変換する変換部と、前記オーディオスペクトルで与えられたフレームに含まれた周波数バンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定し、スペクトルエネルギーを用いて割り当てビット数を小数点単位で推定し、前記割り当てビット数が前記許容ビット数を超えないように調整するビット割り当て部と、前記周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いて前記オーディオスペクトルを符号化する符号化部と、を備える。

20

【0009】

前記課題を解決するための本発明の一実施形態によるオーディオ復号化装置は、与えられたフレームに対して使用可能なビット数範囲内で、各周波数バンドに存在するスペクトルのSNRを最大化させるように、前記周波数バンド単位で割り当てビット数を小数点単位で定め、前記周波数バンド単位で定められた割り当てビット数を調整するビット割り当て部と、前記周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いてビットストリームに含まれたオーディオスペクトルを復号化する復号化部と、前記復号化されたオーディオスペクトルを時間ドメインのオーディオ信号に変換する逆変換部と、を備える。

30

【0010】

前記課題を解決するための本発明の一実施形態によるオーディオ復号化装置は、与えられたフレームに含まれた周波数バンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定し、スペクトルエネルギーを用いて割り当てビット数を小数点単位で推定し、前記割り当てビット数が前記許容ビット数を超えないように調整するビット割り当て部と、前記周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いてビットストリームに含まれたオーディオスペクトルを復号化する復号化部と、前記復号化されたオーディオスペクトルを時間ドメインのオーディオ信号に変換する逆変換部と、を備える。

40

【発明の効果】

【0011】

本発明によれば、知覚的なモデリングを用いてサブバンド単位で小数点単位の最大許容ビット数を計算し、最大許容ビット数を超えないように制限して他のサブバンドに割り当てられるように調節する。その結果、特定サブバンドに必要以上にビットを使わないように他のサブバンドにビットを再分配することで、さらに効率的なビット割り当てになるようにする。また、サブバンド単位で必要なビット数を数式的に推定することで、低い複雑度で具現でき、小数点単位のビット割り当てが可能になってサブバンド単位で必要とする

50

最適のビットを求める。

【図面の簡単な説明】

【0012】

【図1】本発明の一実施形態によるオーディオ符号化装置の構成を示すブロック図である。

【図2】図1において、本発明の一実施形態によるビット割り当て部の構成を示すブロック図である。

【図3】図1において、本発明の他の実施形態によるビット割り当て部の構成を示すブロック図である。

【図4】図1において、本発明の他の実施形態によるビット割り当て部の構成を示すブロック図である。

10

【図5】図1において、本発明の一実施形態による符号化部の構成を示すブロック図である。

【図6】本発明の他の実施形態によるオーディオ符号化装置の構成を示すブロック図である。

【図7】本発明の一実施形態によるオーディオ復号化装置の構成を示すブロック図である。

【図8】図7において、本発明の一実施形態によるビット割り当て部の構成を示すブロック図である。

【図9】図7において、本発明の一実施形態による復号化部の構成を示すブロック図である。

20

【図10】図7において、本発明の他の実施形態による復号化部の構成を示すブロック図である。

【図11】図7において、本発明の他の実施形態による復号化部の構成を示すブロック図である。

【図12】本発明の他の実施形態によるオーディオ復号化装置の構成を示すブロック図である。

【図13】本発明の他の実施形態によるオーディオ復号化装置の構成を示すブロック図である。

【図14】本発明の一実施形態によるビット割り当て方法の動作を説明するフローチャートである。

30

【図15】本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。

【図16】本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。

【図17】本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。

【図18】本発明の一実施形態による符号化モジュールを備えるマルチメディア機器の構成を示すブロック図である。

【図19】本発明の一実施形態による復号化モジュールを備えるマルチメディア機器の構成を示すブロック図である。

40

【図20】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備えるマルチメディア機器の構成を示すブロック図である。

【発明を実施するための形態】

【0013】

本発明は、多様な変換を加えられ、かつ様々な実施形態を持つことができる、特定の実施形態を図面に例示して詳細な説明に具体的に説明しようとする。しかし、これは本発明を特定の実施形態について限定しようとするものではなく、本発明の技術的思想及び技術範囲に含まれるすべての変換、均等物ないし代替物を含むと理解できる。本発明を説明するにあたって、かかる公知技術についての具体的な説明が本発明の趣旨を不明にす

50

ると判断される場合、その詳細な説明を略する。

【0014】

第1、第2などの用語は、多様な構成要素の説明に使われるが、構成要素が用語によって限定されるものではない。用語は、一つの構成要素を他の構成要素から区別する目的のみで使われる。

【0015】

本発明で使った用語は、単に特定の実施形態を説明するために使われたものであり、本発明を限定しようとする意図ではない。本発明で使った用語は、本発明での機能を考慮してなるべく現在広く使われる一般的な用語を選択したが、これは当業者の意図、判例、または新たな技術の出現などによって変わりうる。また、特定の場合には出願人が任意に選定した用語もあり、この場合、該発明の説明部分で詳細にその意味を記載する。したがって、本発明で使われる用語は単純な用語の名称ではない、その用語が持つ意味及び本発明の全般的な内容に基づいて定義されねばならない。

【0016】

単数の表現は、文脈上明らかに断らない限り、複数の表現を含む。本発明で、“含む”または“持つ”などの用語は、明細書上に記載の特徴、数字、段階、動作、構成要素、部品またはこれらを組み合わせたものの存在を指定しようとするものであり、一つまたはそれ以上の他の特徴や数字、段階、動作、構成要素、部品またはこれらを組み合わせたものの存在または付加可能性を予め排除しないと理解されねばならない。

【0017】

以下、本発明の実施形態を添付図面を参照して詳細に説明し、添付図面を参照して説明するにあたって、同一または対応する構成要素には同じ図面番号をつけ、これについての重なる説明は略する。

【0018】

図1は、本発明の一実施形態によるオーディオ符号化装置100の構成を示すブロック図である。図1に示されたオーディオ符号化装置100は、変換部130、ビット割り当て部150、符号化部170及び多重化部190を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。ここで、オーディオは、オーディオあるいは音声、またはオーディオと音声との混合信号を意味するが、以下では、説明の便宜のためにオーディオと通称する。

【0019】

図1を参照すれば、変換部130は、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してオーディオスペクトルを生成する。この時、時間／周波数ドメイン変換は、DCTなどの公知の多様な方法を使って行う。

【0020】

ビット割り当て部150は、オーディオスペクトルについてスペクトルエネルギーあるいは心理音響モデルを用いて得られるマスキングしきい値と、スペクトルエネルギーを用いて各サブバンド単位で割り当てビット数とを定める。ここで、サブバンドは、オーディオスペクトルのサンプルをグルーピングした単位であり、しきい帯域を反映して均一あるいは不均一な長さを持つ。不均一な場合、1フレームについて最初のサンプルから最後のサンプルに至るまでサブバンドに含まれるサンプルの数が段々増加するようにサブバンドを設定する。ここで、1フレームに含まれるサブバンドの数あるいはサブバンドに含まれるサンプルの数は既定のものでありうる。または、1フレームについて所定数の均一長さのサブバンドに分けた後、スペクトル係数の分布によって長さを調整する。スペクトル係数の分布は、スペクトル平坦度、最大値と最小値との差あるいは最大値の微分値などを用いて定める。

【0021】

一実施形態によれば、ビット割り当て部150は、各サブバンド単位で求められたNorm値、すなわち、平均スペクトルエネルギーを用いて許容ビット数を推定し、平均スペクトルエネルギーを用いてビットを割り当て、割り当てビット数が許容ビット数を超えな

10

20

30

40

50

いように制限する。

【００２２】

他の実施形態によれば、ビット割り当て部１５０は、各サブバンド単位で心理音響モデルを用いて許容ビット数を推定し、平均スペクトルエネルギーを用いてビットを割り当て、割り当てビット数が許容ビット数を超えないように制限する。

【００２３】

符号化部１７０は、各サブバンド単位で最終的に定められた割り当てビット数に基づいてオーディオスペクトルを量子化及び無損失符号化し、符号化されたスペクトルに関する情報を生成する。

【００２４】

多重化部１９０は、ビット割り当て部１５０から提供される符号化されたNorm値と、符号化部１７０に提供される符号化されたスペクトルに関する情報とを多重化してビットストリームを生成する。

【００２５】

一方、オーディオ符号化装置１００は、オプションで与えられたサブバンドについてノイズレベルを生成し、オーディオ復号化装置７００（図７）、１２００（図１２）、１３００（図１３）に提供する。

【００２６】

図２は、図１において、本発明の一実施形態によるビット割り当て部２００の構成を示すブロック図である。図２に示されたビット割り当て部２００は、Norm推定部２１０、Norm符号化部２３０及びビット推定及び割り当て部２５０を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

【００２７】

図２を参照すれば、Norm推定部２１０は、各サブバンド単位で平均スペクトルエネルギーに該当するNorm値を求める。この時、一例として、ITU-T G. 719で適用される下記の数式（１）のようにNorm値を計算できるが、これに限定されるものではない。

【００２８】

【数１】

$$N(p) = \sqrt{\frac{1}{L_p} \sum_{k=s_p}^{e_p} y(k)^2}, \quad p=0, \dots, P-1 \quad (1)$$

ここで、１フレームにＰ個のサブバンドあるいはサブベクトルが存在する場合、 $N(p)$ は、サブバンドあるいはサブベクトル p のNorm値、 L_p は、サブバンドあるいはサブベクトル p の長さ、すなわち、サンプルあるいはスペクトル係数の数、 s_p 及び e_p は、サブバンド p の最初及び最後のサンプル、 $y(k)$ は、サンプルのサイズあるいはスペクトル係数（すなわち、エネルギー）をそれぞれ意味する。

【００２９】

一方、各サブバンドについて求められたNorm値は、符号化部１５０（図１）に提供される。

【００３０】

Norm符号化部２３０は、各サブバンドについて求められたNorm値を量子化及び無損失符号化する。ここで、各サブバンド単位で量子化されたNorm値は、ビット推定及び割り当て部２５０に提供されるか、または各サブバンド単位で再び逆量子化されたN

10

20

30

40

50

orm値は、ビット推定及び割り当て部250に提供される。一方、各サブバンド単位で量子化及び無損失符号化されたNorm値は、多重化部190（図1）に提供される。

【0031】

ビット推定及び割り当て部250は、各サブバンド単位で、Norm値を用いて必要とするビット数を推定して割り当てる。望ましくは、符号化部及び復号化部で同じビット推定及び割り当てプロセスを用いるように、逆量子化されたNorm値を使う。この時、マスキング効果を考慮して調整されたNorm値を使える。Norm値の調整には、一例として、下記の数式（2）のようにITU-T G. 719で適用される心理音響加重（psycho-acoustical weighting）を用いられるが、これに限定されるものではない。

10

【0032】

【数2】

$$\tilde{I}_N^q(p) = I_N^q(p) + WSpe(p) \quad (2)$$

ここで、

【0033】

20

【数3】

$$I_N^q(p)$$

は、サブバンドpの量子化されたNorm値のインデックス、

【0034】

【数4】

30

$$\tilde{I}_N^q(p)$$

は、サブバンドpの調整されたNorm値のインデックス、

【0035】

【数5】

$$WSpe(p)$$

40

は、Norm値調整のためのオフセットスペクトルをそれぞれ示す。

【0036】

ビット推定及び割り当て部250は、各サブバンド単位で、Norm値を用いてマスキングしきい値を計算し、マスキングしきい値を用いて知覚的に必要なビット数を予測する。このために、まず、各サブバンドについて求められたNorm値は、下記の数式（3）のようにdB単位のスเปクトルエネルギーと等価的に表現される。

【0037】

【数 6】

$$2\log_2 \left[\sqrt{\frac{1}{L_p} \sum_{k=1}^{L_p} y(k)^2} \right] = 10\log_{10} \left[\sum_{k=1}^{L_p} y(k)^2 \right] 0.1\log_2 10 - \log_2 (L_p) \quad (3)$$

一方、スペクトルエネルギーを用いてマスキングしきい値を求める方法は、公知の多様な方法を使える。すなわち、マスキングしきい値は、JND (Just Noticeable Distortion) に該当する値であり、量子化ノイズがマスキングしきい値より小さい場合、知覚的なノイズが感じられない。よって、知覚的なノイズを感じられなくするのに必要な最小ビット数を、マスキングしきい値を用いて計算する。一実施形態で、各サブバンド単位で、Norm値とマスキングしきい値との比を用いてSMR (Signal-to-Mask Ratio) を計算し、SMRについて、6.025dB 1ビットの係数を用いてマスキングしきい値を満たすビット数を予測する。ここで、予測されたビット数は、知覚的なノイズを感じられなくするのに必要な最小ビット数であるが、圧縮側面からみれば、予測されたビット数以上に使う必要がないので、サブバンド単位で許容される最大ビット数（以下、許容ビット数と略称）と見なされる。この時、各サブバンドの許容ビット数は、小数点単位で表現される。

【0038】

ビット推定及び割り当て部250は、各サブバンド単位で、Norm値を用いて小数点単位のビット割り当てを行う。この時、Norm値の大きいサブバンドから順次にビットを割り当てるが、各サブバンドのNorm値に対して各サブバンドの知覚的重要度によって加重値を付与することで、知覚的に重要なサブバンドにさらに多いビットが割り当てられるように調整する。知覚的重要度は、一例としてITU-T G.719でのような心理音響加重を通じて定める。

【0039】

具体的に、ビット推定及び割り当て部250は、Norm値の大きいサブバンドから順次にサンプル別にビットを割り当てる。すなわち、優先的に最大Norm値を持つサブバンドについてサンプル当たりビットを割り当て、該サブバンドのNorm値を所定単位ほど低減させて他のサブバンドにビットを割り当てるように優先順位を変更する。このような過程は、与えられたフレームで使用可能な全体ビット数Bがいずれもなくなるまで繰り返して行われる。

【0040】

ビット推定及び割り当て部250は、各サブバンドについて割り当てられたビット数が予測されたビット数、すなわち、許容ビット数を超えないように制限し、最終的に割り当てビット数を定める。すべてのサブバンドについて、割り当てられたビット数と予測されたビット数と比較して、割り当てられたビット数が予測されたビット数より大きい場合には、予測されたビット数に制限する。ビット数制限の結果で得られる、与えられたフレームの全体サブバンドのビット数が、与えられたフレームで使用可能な全体ビット数Bより少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

【0041】

これによれば、各サブバンドの割り当てビット数を小数点単位で定めつつ許容ビット数に制限できるため、与えられたフレームの全体ビット数をさらに効率的に配分する。

【0042】

一方、各サブバンドに必要なビット数を推定及び割り当てる具体的な方法は、次の通りである。これによれば、数回の繰り返しなしに各サブバンド単位で割り当てビット数を一回に定めることができ、複雑度が低くなる。

【0043】

一実施形態で、次の数式（４）に記載の Lagrange function を適用して、量子化歪曲及び各サブバンドに割り当てられるビット数を最適化できる解（solution）を求める。

【0044】

【数 7】

$$L = D + \lambda (\sum N_b L_b - B) \quad (4)$$

10

ここで、L は、Lagrange function を称し、D は、量子化歪曲、B は、与えられたフレームで使用可能な全体ビット数、 N_b は、サブバンド b のサンプル数、 L_b は、サブバンド b で各サンプルに割り当てられたビット数を示す。すなわち、 $N_b L_b$ は、サブバンド b に割り当てられたビット数を示す。ここで、 λ は、最適化係数である Lagrange multiplier を示し、与えられた関数の最小値を探すためのコントロールパラメータである。

【0045】

前記数式（４）を用いれば、量子化歪曲を考慮しつつ、与えられたフレームに含まれた各サブバンドに割り当てられたビット数の総和と、与えられたフレームについての許容ビット数との差が最小になる L_b を定める。

20

【0046】

そして、量子化歪曲 D は、次の数式（５）のように定義する。

【0047】

【数 8】

$$D = \frac{\sum_i (x_i - \tilde{x}_i)^2}{\sum_i x_i^2} \quad (5)$$

30

ここで、

【0048】

【数 9】

$$x_i$$

40

は、入力スペクトル、

【0049】

【数 10】

$$\tilde{x}_i$$

は、復号化されたスペクトルを示す。すなわち、量子化歪曲 D は、任意のフレームで入力スペクトル

50

【 0 0 5 0 】

【数 1 1】

$$x_i$$

及び復号化されたスペクトル

【 0 0 5 1 】

【数 1 2】

10

$$\tilde{x}_i$$

についてのMSE (Mean Square Error) と定義される。

【 0 0 5 2 】

一方、数式 (5) で分母項は、与えられた入力スペクトルによって定められる一定値であり、したがって、最適化に影響しないため、次の数式 (6) のように簡略化される。

【 0 0 5 3 】

【数 1 3】

20

$$L = \sum_i (x_i - \tilde{x}_i)^2 + \lambda (\sum_b N_b L_b - B) \quad (6)$$

入力スペクトル

【 0 0 5 4 】

【数 1 4】

30

$$x_i$$

について任意のサブバンド b の平均スペクトルエネルギーである norm 値

【 0 0 5 5 】

【数 1 5】

$$g_b$$

40

は、次の数式 (7) のように定義され、ログスケールで量子化された norm 値

【 0 0 5 6 】

【数 1 6】

$$n_b$$

は、次の数式 (8) のように定義され、逆量子化された norm 値

50

【 0 0 5 7 】

【 数 1 7 】

$$\tilde{g}_b$$

は、次の数式（ 9 ）のように定義される。

【 0 0 5 8 】

【 数 1 8 】

10

$$g_b = \sqrt{\frac{\sum_{i=s_b}^{e_b} x_i^2}{N_b}} \quad (7)$$

20

$$n_b = \lfloor 2 \log_2 g_b + 0.5 \rfloor \quad (8)$$

$$\tilde{g}_b = 2^{0.5 n_b} \quad (9)$$

30

ここで、 s_b 及び e_b は、それぞれサブバンド b の最初及び最後のサンプルを示す。

【 0 0 5 9 】

次いで、入力スペクトル

【 0 0 6 0 】

【 数 1 9 】

$$x_i$$

は、次の数式（ 1 0 ）でのように、逆量子化された $n o r m$ 値

40

【 0 0 6 1 】

【 数 2 0 】

$$\tilde{g}_b$$

で割って正規化されたスペクトル y_i を生成し、次の数式（ 1 1 ）のように復元された、
正規化されたスペクトル

【 0 0 6 2 】

50

【数 2 1】

$$\tilde{y}_i$$

に逆量子化された n o r m 値

【0 0 6 3】

【数 2 2】

10

$$\tilde{g}_b$$

を掛けて復号化されたスペクトル

【0 0 6 4】

【数 2 3】

$$\tilde{x}_i$$

20

を生成する。

【0 0 6 5】

【数 2 4】

$$y_i = \frac{x_i}{\tilde{g}_b}, \quad i \in [s_b, \dots, e_b] \quad (10)$$

30

$$\tilde{x}_i = \tilde{y}_i \tilde{g}_b, \quad i \in [s_b, \dots, e_b] \quad (11)$$

前記数式（9）ないし（11）を用いて数式（6）の量子化歪曲項をまとめれば、次の数式（12）のように示す。

【0 0 6 6】

【数 2 5】

40

$$\sum_i (x_i - \tilde{x}_i)^2 = \sum_b \tilde{g}_b^2 \sum_{i \in b} (y_i - \tilde{y}_i)^2 = \sum_b 2^{n_b} \sum_{i \in b} (y_i - \tilde{y}_i)^2 \quad (12)$$

通常量子化歪曲と割り当てられたビット数との関係で、サンプル当たり1ビットが加えられる度にSNRが6.02dBほど増加すると定義し、これを用いて正規化されたスペクトルの量子化歪曲を定義すれば、次の数式（13）のように示す。

【0 0 6 7】

【数 2 6】

$$\frac{\sum_{i \in b} (y_i - \tilde{y}_i)^2}{\sum_i y_i^2} = \frac{\sum_{i \in b} (y_i - \tilde{y}_i)^2}{N_b} = 2^{-2L_b} \quad (13)$$

一方、実際オーディオコーディングに適用する場合には、1 b i t / s a m p l e について 6 . 0 2 d B の関係を固定せず、信号の特性によって可変される d B スケール値 C を適用して次の数式 (1 4) のように定義する。

【 0 0 6 8 】

【数 2 7】

$$\sum_{i \in b} (y_i - \tilde{y}_i)^2 = 2^{-CL_b} N_b \quad (14)$$

ここで、C が 2 である場合に 6 . 0 2 d B 、 C が 3 である場合に 9 . 0 3 d B に該当する。

【 0 0 6 9 】

したがって、数式 (6) は、数式 (1 2) 及び数式 (1 4) から次の数式 (1 5) のように示す。

【 0 0 7 0 】

【数 2 8】

$$L = \sum_b 2^{n_b} 2^{-CL_b} N_b + \lambda \left(\sum_b N_b L_b - B \right) \quad (15)$$

前記数式 (1 5) で最適の L_b 及び λ を求めるために、次の数式 (1 6) のように、 L_b 及び λ についてそれぞれ偏微分を行う。

【 0 0 7 1 】

【数 2 9】

$$\begin{aligned} \frac{\partial L}{\partial L_b} &= -C 2^{n_b - CL_b} N_b \ln 2 + \lambda N_b = 0 \\ \frac{\partial L}{\partial \lambda} &= \sum_b N_b L_b - B = 0 \end{aligned} \quad (16)$$

前記数式 (1 6) をまとめれば、 L_b は、次の数式 (1 7) のように示す。

【 0 0 7 2 】

【数 3 0】

$$L_b = \frac{1}{C} \left(n_b - \frac{\sum_b N_b n_b - CB}{\sum_b N_b} \right) \quad (17)$$

前記数式（17）を用いれば、与えられたフレームで使用可能な全体ビット数Bの範囲内で、入力スペクトルのSNRを最大化させられる各サブバンドのサンプル当たり割り当てビット数 L_b を推定する。

10

【0073】

ビット推定及び割り当て部250で、各サブバンド単位で定められた割り当てビット数は、符号化部170（図1）に提供される。

【0074】

図3は、図1において、本発明の他の実施形態によるビット割り当て部300の構成を示すブロック図である。図3に示されたビット割り当て部300は、心理音響モデル310、ビット推定及び割り当て部330、スケールファクタ推定部350及びスケールファクタ符号化部370を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

20

【0075】

図3を参照すれば、心理音響モデル310は、変換部130（図1）から提供されるオーディオスペクトルを入力とし、各サブバンドについてマスキングしきい値を求める。

【0076】

ビット推定及び割り当て部330は、各サブバンド単位でマスキングしきい値を用いて知覚的に必要なビット数を予測する。すなわち、各サブバンド単位でSMRを求めることができ、SMRについて6.025dB 1ビットの関係を用いてマスキングしきい値を満たすビット数を予測する。ここで、予測されたビット数は、知覚的なノイズを感じられなくするのに必要な最小ビット数であるが、圧縮側面からみれば、予測されたビット数以上に使う必要がないので、サブバンド単位で許容される最大ビット数（以下、許容ビット数と略称）と見なされる。この時、各サブバンドの許容ビット数は、小数点単位で表現される。

30

【0077】

ビット推定及び割り当て部330は、各サブバンド単位でスペクトルエネルギーを用いて小数点単位のビット割り当てを行う。この時、例えば、前記数式（4）ないし（17）によるビット割り当て方法を使える。

【0078】

ビット推定及び割り当て部330は、すべてのサブバンドについて、割り当てられたビット数と予測されたビット数とを比較して、割り当てられたビット数が予測されたビット数より大きい場合には予測されたビット数に制限する。ビット数制限の結果で得られる、与えられたフレームの全体サブバンドのビット数が、与えられたフレームで使用可能な全体ビット数Bより少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

40

【0079】

スケールファクタ推定部350は、各サブバンド単位で最終的に定められた割り当てビット数を用いてスケールファクタを推定する。各サブバンド単位で推定されたスケールファクタは、符号化部170（図1）に提供される。

【0080】

スケールファクタ符号化部370は、各サブバンド単位で推定されたスケールファクタ

50

を量子化及び無損失符号化する。サブバンド単位で符号化されたスケールファクタは、多重化部190（図1）に提供される。

【0081】

図4は、図1において、本発明の他の実施形態によるビット割り当て部300の構成を示すブロック図である。図4に示されたビット割り当て部400は、Norm推定部410、ビット推定及び割り当て部430、スケールファクタ推定部450及びスケールファクタ符号化部470を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

【0082】

図4を参照すれば、Norm推定部410は、各サブバンド単位で平均スペクトルエネルギーに該当するNorm値を求める。

10

【0083】

ビット推定及び割り当て部430は、各サブバンド単位でスペクトルエネルギーを用いてマスキングしきい値を求め、マスキングしきい値を用いて知覚的に必要なビット数、すなわち、許容ビット数を予測する。

【0084】

ビット推定及び割り当て部430は、各サブバンド単位で、スペクトルエネルギーを用いて小数点単位のビット割り当てを行う。この時、例えば、前記数式（4）ないし（17）によるビット割り当て方法を使える。

【0085】

20

ビット推定及び割り当て部430は、すべてのサブバンドについて、割り当てられたビット数と予測されたビット数とを比較して、割り当てられたビット数が予測されたビット数より大きい場合には、予測されたビット数に制限する。ビット数の制限の結果で得られる、与えられたフレームの全体サブバンドの割り当てビット数が、与えられたフレームで使用可能な全体ビット数Bより少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

【0086】

スケールファクタ推定部450は、各サブバンド単位で最終的に定められた割り当てビット数を用いてスケールファクタを推定する。各サブバンド単位で推定されたスケールファクタは、符号化部170（図1）に提供される。

30

【0087】

スケールファクタ符号化部470は、各サブバンド単位で推定されたスケールファクタを量子化及び無損失符号化する。サブバンド単位で符号化されたスケールファクタは多重化部190（図1）に提供される。

【0088】

図5は、図1において、本発明の一実施形態による符号化部500の構成を示すブロック図である。

【0089】

図5に示された符号化部500は、スペクトル正規化部510及びスペクトル符号化部530を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

40

【0090】

図5を参照すれば、スペクトル正規化部510は、ビット割り当て部150（図1）から提供される各サブバンドのNorm値を用いてスペクトルを正規化する。

【0091】

スペクトル符号化部530は、正規化されたスペクトルについて各サブバンドの割り当てビット数を用いて量子化を行い、量子化された結果について無損失符号化する。一例として、スペクトル符号化にファクトリアル・パルス・コーディング（F a c t o r i a l P u l s e C o d i n g）を使えるが、これに限定されるものではない。ファクトリアル・パルス・コーディングによれば、割り当てビット数範囲内でパルスの位置、パルス

50

のサイズ、及びパルスの符号などの情報がファクトリアル形式で表現される。

【0092】

スペクトル符号化部530で符号化されたスペクトルに関する情報は、多重化部190(図1)に提供される。

【0093】

図6は、本発明の他の実施形態によるオーディオ符号化装置600の構成を示すブロック図である。図6に示されたオーディオ符号化装置600は、トランジェント検出部610、変換部630、ビット割り当て部650、符号化部670及び多重化部690を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ(図示せず)に具現される。図6のオーディオ符号化装置600は、図1のオーディオ符号化装置100と比較すれば、トランジェント検出部610をさらに備える差があるので、共通の構成要素については具体的な説明を略する。

10

【0094】

図6を参照すれば、トランジェント検出部610は、オーディオ信号を分析してトランジェント特性を示す区間を検出する。トランジェント区間の検出には、公知の多様な方法を使える。トランジェント検出部610に提供されるトランジェントシグナリング情報は、多重化部690を通じてビットストリームに含まれる。

【0095】

変換部630は、トランジェント区間検出結果によって、変換に使われるウィンドウサイズを定め、定められたウィンドウサイズに基づいて時間／周波数ドメイン変換を行う。一例として、トランジェント区間が検出されたサブバンドの場合に短区間ウィンドウ(short window)を、検出されていないサブバンドの場合に長区間ウィンドウ(long window)を適用する。

20

【0096】

ビット割り当て部650は、図2ないし図4に示されたビット割り当て部200、300、400のうちいずれか一つに具現される。

【0097】

符号化部670は、トランジェント区間検出結果によって、変換部630と同様に符号化に使われるウィンドウサイズを定める。

【0098】

一方、オーディオ符号化装置600は、オプションで与えられたサブバンドについてノイズレベルを生成し、オーディオ復号化装置700(図7)、1200(図12)、1300(図13)に提供する。

30

【0099】

図7は、本発明の一実施形態によるオーディオ復号化装置700の構成を示すブロック図である。図7に示されたオーディオ復号化装置700は、逆多重化部710、ビット割り当て部730、復号化部750及び逆変換部770を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ(図示せず)に具現される。

【0100】

図7を参照すれば、逆多重化部710は、ビットストリームを逆多重化して量子化及び無損失符号化されたNorm値と、符号化されたスペクトルに関する情報とを抽出する。

40

【0101】

ビット割り当て部730は、各サブバンド単位で量子化及び無損失符号化されたNorm値から逆量子化されたNorm値を求め、逆量子化されたNorm値を用いて割り当てビット数を定める。ビット割り当て部730は、オーディオ符号化装置100、600のビット割り当て部150、650と実質的に同じく動作する。一方、オーディオ符号化装置100、600で、Norm値が心理音響加重によって調整された場合、オーディオ復号化装置700でも同じく調整される。

【0102】

50

復号化部 750 は、逆多重化部 710 から提供される、符号化されたスペクトルに関する情報を用いて、符号化されたスペクトルを無損失復号化及び逆量子化する。一例として、スペクトル復号化はファクトリアル・パルス・デコーディングを使える。逆変換部 770 は、復号化されたスペクトルを時間ドメインに変換し、復元されたオーディオ信号を生成する。

【0103】

図 8 は、図 7 において、本発明の一実施形態によるビット割り当て部 800 の構成を示すブロック図である。図 8 に示されたビット割り当て部 800 は、Norm 復号化部 810 と、ビット推定及び割り当て部 830 とを備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

10

【0104】

図 8 を参照すれば、Norm 復号化部 810 は、逆多重化部 710（図 7）から提供される量子化及び無損失符号化された Norm 値から、逆量子化された Norm 値を求める。

【0105】

ビット推定及び割り当て部 830 は、逆量子化された Norm 値を用いて割り当てビット数を定める。具体的に、ビット推定及び割り当て部 830 は、各サブバンド単位でスペクトルエネルギー、すなわち、Norm 値を用いてマスキングしきい値を求め、マスキングしきい値を用いて知覚的に必要なビット数、すなわち、許容ビット数を予測する。

【0106】

20

ビット推定及び割り当て部 830 は、各サブバンド単位で、スペクトルエネルギー、すなわち、Norm 値を用いて小数点単位のビット割り当てを行う。この時、例えば、前記数式（4）ないし（17）によるビット割り当て方法を使える。

【0107】

ビット推定及び割り当て部 830 は、すべてのサブバンドについて、割り当てられたビット数と予測されたビット数とを比較して、割り当てられたビット数が予測されたビット数より大きい場合には、予測されたビット数に制限する。ビット数制限の結果で得られる、与えられたフレームの全体サブバンドの割り当てビット数が、与えられたフレームで使用可能な全体ビット数 B より少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

30

【0108】

図 9 は、図 7 において、本発明の一実施形態による復号化部 900 の構成を示すブロック図である。図 9 に示された復号化部 900 は、スペクトル復号化部 910 及びエンベロープ・シェーピング部 930 を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

【0109】

図 9 を参照すれば、スペクトル復号化部 910 は、逆多重化部 710（図 7）から提供される、符号化されたスペクトルに関する情報、及びビット割り当て部 730（図 7）から提供される割り当てビット数を用いて、符号化されたスペクトルを無損失復号化及び逆量子化する。スペクトル復号化部 910 から提供される、復号化されたスペクトルは、正規化されたスペクトルである。

40

【0110】

エンベロープ・シェーピング部 930 は、ビット割り当て部 730（図 7）から提供される、逆量子化された Norm 値を用いて、スペクトル復号化部 910 から提供される、正規化されたスペクトルについてエンベロープ・シェーピングを行って正規化以前のスペクトルに復元する。

【0111】

図 10 は、図 7 において、本発明の他の実施形態による復号化部 1000 の構成を示すブロック図である。図 10 に示された復号化部 1000 は、スペクトル復号化部 1000、エンベロープ・シェーピング部 1030 及びスペクトル・フィーリング部 1050 を備

50

える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。図10の復号化部1000は、図9の復号化部900と比べれば、スペクトル・フィーリング部1050をさらに備える差があるので、共通の構成要素については、具体的な説明を略する。

【0112】

図10を参照すれば、スペクトル・フィーリング部1050は、エンベロープ・シェーピング部1030から提供されるスペクトルで0に逆量子化された部分を含むサブバンドが存在する場合、ノイズ成分で満たすことができる。一実施形態によれば、ノイズ成分はランダムに生成されるか、または0でない値に逆量子化されたサブバンドのスペクトルをコピーして生成する。他の実施形態によれば、0に逆量子化された部分を含むサブバンドについてノイズ成分を生成し、ノイズ成分のエネルギーと、ビット割り当て部730（図7）から提供される逆量子化されたNorm値、すなわち、スペクトルエネルギーとの比を用いてノイズ成分のエネルギーを調節する。他の実施形態によれば、0に逆量子化された部分を含むサブバンドについてノイズ成分を生成し、ノイズ成分の平均エネルギーが1になるように調節する。他の実施形態によれば、オーディオ符号化装置100、600から各サブバンド単位でノイズレベルを受信し、与えられたサブバンドが0に逆量子化された部分を含む場合、与えられたサブバンドについてノイズ成分を生成し、受信されたノイズレベルを用いてノイズ成分のエネルギーを調節する。

10

【0113】

図11は、図7において、本発明の他の実施形態による復号化部1100の構成を示すブロック図である。図11に示された復号化部1100は、スペクトル復号化部1100、スペクトル・フィーリング部1130及びエンベロープ・シェーピング部1150を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。図11の復号化部1100は、図10の復号化部1000と比べれば、スペクトル・フィーリング部1130及びエンベロープ・シェーピング部1150の配置順序が異なる差があるので、共通の構成要素については、具体的な説明を略する。

20

【0114】

図11を参照すれば、スペクトル・フィーリング部1130は、スペクトル復号化部1110から提供される、正規化されたスペクトルで0に逆量子化された部分を含むサブバンドが存在する場合、ノイズ成分で満たすことができる。この時、図10のスペクトル・フィーリング部1050に適用される多様なノイズフィーリング方法を使える。望ましくは、0に逆量子化された部分を含むサブバンドについてノイズ成分を生成し、ノイズ成分の平均エネルギーが1になるように調節する。

30

【0115】

エンベロープ・シェーピング部1150は、ビット割り当て部730（図7）から提供される逆量子化されたNorm値を用いて、ノイズ成分が満たされたサブバンドを含むスペクトルについて正規化以前のスペクトルに復元する。

【0116】

図12は、本発明の他の実施形態によるオーディオ復号化装置1200の構成を示すブロック図である。図12に示されたオーディオ復号化装置1200は、逆多重化部1210、スケールファクタ復号化部1230、スペクトル復号化部1250及び逆変換部1270を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

40

【0117】

図12を参照すれば、逆多重化部1210は、ビットストリームを逆多重化して量子化及び無損失符号化されたスケールファクタと、符号化されたスペクトルに関する情報とを抽出する。

【0118】

スケールファクタ復号化部1230は、各サブバンド単位で量子化及び無損失符号化さ

50

れたスケールファクタを無損失復号化及び逆量子化する。

【0119】

スペクトル復号化部1250は、逆多重化部1210から提供される、符号化されたスペクトルに関する情報及び逆量子化されたスケールファクタを用いて、符号化されたスペクトルを無損失復号化及び逆量子化する。スペクトル復号化部1250は、図10に示された復号化部1000と同じ構成要素を含む。

【0120】

逆変換部1270は、スペクトル復号化部1250で復号化されたスペクトルを時間ドメインに変換して、復元されたオーディオ信号を生成する。

【0121】

図13は、本発明の他の実施形態によるオーディオ復号化装置1300の構成を示すブロック図である。図13に示されたオーディオ復号化装置1300は逆多重化部1310、ビット割り当て部1330、復号化部1350及び逆変換部1370を備える。各構成要素は、少なくとも一つのモジュールに一体化され、少なくとも一つのプロセッサ（図示せず）に具現される。

【0122】

図13に示されたオーディオ復号化装置1300は、図7のオーディオ復号化装置700と比べれば、トランジェントシグナリング情報が復号化部1350及び逆変換部1370に提供されるという差があるので、共通の構成要素については、具体的な説明を略する。

【0123】

図13を参照すれば、復号化部1350では、逆多重化部1310から提供される、符号化されたスペクトルに関する情報を用いてスペクトルを復号化する。この時、トランジェントシグナリング情報によってウィンドウサイズが可変される。

【0124】

逆変換部1370では、復号化されたスペクトルを時間ドメインに変換し、復元されたオーディオ信号を生成する。この時、トランジェントシグナリング情報によってウィンドウサイズが可変される。

【0125】

図14は、本発明の一実施形態によるビット割り当て方法の動作を説明するフローチャートである。図14を参照すれば、1410段階では、各サブバンド単位でスペクトルエネルギーを獲得する。スペクトルエネルギーは、Norm値を使える。

【0126】

1420段階では、各サブバンド単位で、スペクトルエネルギーを用いてマスキングしきい値を獲得する。

【0127】

1430段階では、各サブバンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定する。

【0128】

1440段階では、各サブバンド単位で、スペクトルエネルギーに基づいて小数点単位でビットを割り当てる。

【0129】

1450段階では、各サブバンド単位で、許容ビット数と割り当てビット数とを比較する。

【0130】

1460段階では、1450段階での比較結果、与えられたサブバンドについて割り当てビット数が許容ビット数より大きい場合、割り当てビット数を許容ビット数に制限する。

【0131】

1470段階では、1450段階での比較結果、与えられたサブバンドについて割り当

10

20

30

40

50

てビット数が許容ビット数より小さいか、または同じ場合、割り当てビット数をそのまま使うか、または1460段階で制限された許容ビット数を使って各サブバンドについて最終的な割り当てビット数を定める。

【0132】

一方、図示されていないが、1470段階で与えられたフレームの各サブバンドについて定められた割り当てビット数の総和が、与えられたフレームで使用可能な全体ビット数より多いか、または少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

【0133】

図15は、本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。図15を参照すれば、各サブバンド単位で、逆量子化されたNorm値を受信する。

10

【0134】

1510段階では、各サブバンド単位で、逆量子化されたNorm値を用いてマスキングしきい値を獲得する。

【0135】

1520段階では、各サブバンド単位で、マスキングしきい値を用いてSMRを獲得する。

【0136】

1530段階では、各サブバンド単位で、SMRを用いて許容ビット数を小数点単位で推定する。

20

【0137】

1540段階では、各サブバンド単位で、スペクトルエネルギーあるいは逆量子化されたNorm値に基づいて、小数点単位でビットを割り当てる。

【0138】

1550段階では、各サブバンド単位で、許容ビット数と割り当てビット数とを比較する。1560段階では、1550段階での比較結果、与えられたサブバンドについて割り当てビット数が許容ビット数より大きい場合、割り当てビット数を許容ビット数に制限する。

【0139】

30

1570段階では、1550段階での比較結果、与えられたサブバンドについて割り当てビット数が許容ビット数より小さいか、または同じ場合、割り当てビット数をそのまま使うか、または1560段階で制限された許容ビット数を使って各サブバンドについて最終的な割り当てビット数を定める。

【0140】

一方、図示されていないが、1570段階で与えられたフレームの各サブバンドについて定められた割り当てビット数の総和が、与えられたフレームで使用可能な全体ビット数より多いか、または少ない場合、その差に該当するビット数を全体サブバンドに均一に分配するか、あるいは知覚的重要度によって不均一に分配する。

【0141】

40

図16は、本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。図16を参照すれば、1610段階では初期化を行う。初期化の一例としては、前記数式(17)を使って各サブバンドについて割り当てビット数を推定する場合、すべてのサブバンドについて常に一定の値を持つ

【0142】

【数 3 1】

$$\frac{\sum N_b n_b - CB}{\sum N_b}$$

を計算することで、全体的な複雑度を低減させる。

【0 1 4 3】

1 6 2 0 段階では、前記数式 (4) あるいは数式 (1 7) を使って、各サブバンドについて割り当てビット数を小数点単位で推定する。サンプル当たり割り当てビット数 L_b にサブバンドのサンプル数を掛けて各サブバンドに割り当てられたビット数を計算する。この時、数式 (1 7) を用いて各サブバンドのサンプル当たり割り当てビット数 L_b を計算する場合、 L_b が 0 より小さな値を持つ。この場合、下記の数式 (1 8) のように、0 より小さな値を持つ L_b について 0 を割り当てる。

【0 1 4 4】

【数 3 2】

$$L_b = \max \left(0, \frac{1}{C} \left(n_b - \frac{\sum_b N_b n_b - CB}{\sum_b N_b} \right) \right) \quad (18)$$

その結果、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数 B より大きい。

【0 1 4 5】

1 6 3 0 段階では、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和と、与えられたフレームの使用可能なビット数とを比較する。

【0 1 4 6】

1 6 4 0 段階では、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数と同一になるまで、下記の数式 (1 9) を使って各サブバンドについてビットを再分配する。

【0 1 4 7】

【数 3 3】

$$L_b^k = \max \left(0, L_b^{k-1} - \frac{\sum_b N_b L_b^{k-1} - B}{\sum_b N_b} \right), b \in \{L_b^{k-1} \geq 0\} \quad (19)$$

ここで、

【0 1 4 8】

10

20

30

40

【数 3 4】

$$L_b^{k-1}$$

は、(k-1) 番目の繰り返しによって定められるビット数、

【0 1 4 9】

【数 3 5】

10

$$L_b^k$$

は、k 番目の繰り返しによって定められるビット数を示す。毎繰り返しで定められるビット数は、0 より小さくはならず、したがって、1 6 4 0 段階は、0 より大きいビット数を持つサブバンドについて行われる。

【0 1 5 0】

1 6 5 0 段階では、1 6 3 0 段階での比較結果、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数と同じ場合に各サブバンドの割り当てビット数をそのまま使うか、または 1 6 4 0 段階で再分配結果で得られる各サブバンドの割り当てビット数を使って、各サブバンドについて最終的な割り当てビット数を定める。

20

【0 1 5 1】

図 1 7 は、本発明の他の実施形態によるビット割り当て方法の動作を説明するフローチャートである。図 1 7 を参照すれば、1 7 1 0 段階では、図 1 6 の 1 6 1 0 段階と同様に初期化を行う。1 7 2 0 段階では、図 1 6 の 1 6 2 0 段階と同様に各サブバンドについて割り当てビット数を小数点単位で推定し、各サブバンドのサンプル当たり割り当てビット数 L_b が 0 より小さな場合、前記数式 (1 8) のように 0 より小さな値を持つ L_b について 0 を割り当てる。

30

【0 1 5 2】

1 7 3 0 段階では、S N R 側面で各サブバンドについて必要とする最小ビット数を定義し、1 7 2 0 段階の割り当てビット数が 0 より大きい、最小ビット数よりは少ないサブバンドについては、最小ビット数に制限して割り当てビット数を調整する。このように各サブバンドの割り当てビット数を最小ビット数に制限することで、音質低下の可能性を低減させる。一例として、各サブバンドについて必要とする最小ビット数は、ファクトリアル・パルス・コーディングでパルスコーディングに必要な最小ビット数と定義する。ファクトリアル・パルス・コーディングは、0 ではないパルスの位置 (p o s i t i o n)、パルスのサイズ (m a g n i t u d e)、パルスの符号 (s i g n) のすべての組み合わせを用いて信号を表現する。この時、パルスを表現できるすべての組み合わせ N の場合は、下記の数式 (2 0) のように示す。

40

【0 1 5 3】

【数 3 6】

$$N = \sum_{i=1}^m 2^i F(n, i) D(m, i) \quad (20)$$

ここで、 2^i は、i 個のノン・ゼロ位置にある信号について + / - に表現可能な符号の

50

場合の数を示す。

【0 1 5 4】

数式(20)で、 $F(n, i)$ は、下記の数式(21)のように定義され、与えられた n 個のサンプル、すなわち、位置について i 個のノン・ゼロ位置を選択できる場合の数を示す。

【0 1 5 5】

【数 3 7】

$$F(n, i) = C_i^n = \frac{n!}{i!(n-i)!} \quad (21)$$

10

数式(20)で、 $D(m, i)$ は、下記の数式(22)のように示すことができ、これは、 i 個のノン・ゼロ位置で選択された信号を m 個のサイズに表現できる場合の数を示す。

【0 1 5 6】

【数 3 8】

$$D(m, i) = C_{i-1}^{m-1} = \frac{(m-1)!}{(i-1)!(m-i)!} \quad (22)$$

20

一方、 N 個のすべての組み合わせを表現するために必要なビット数 M は、下記の数式(23)のように示す。

【0 1 5 7】

【数 3 9】

$$M = \lceil \log_2 N \rceil \quad (23)$$

30

結局、与えられたサブバンド b で N_b 個のサンプルについて最小 1 個のパルスを符号化するために必要な最小ビット数 L_{b_min} は、下記の数式(24)のように示す。

【0 1 5 8】

【数 4 0】

$$L_{b_min} = 1 + \log_2 N_b \quad (24)$$

40

この時、ファクトリアル・パルス・コーディングで必要な最小ビット数には、量子化に必要な利得値の伝送に使われるビット数が加えられ、ビット率によって可変される。各サブバンド単位の必要な最小ビット数は、下記の数式(25)のように、ファクトリアル・パルス・コーディングで必要な最小ビット数と、与えられたサブバンドのサンプル数 N_b のうち大きい値と定められる。一例によれば、1 ビット／サンプルで設定する。

【0 1 5 9】

50

【数 4 1】

$$L_{b_min} = \max(N_b, 1 + \log_2 N_b + L_{gain}) \quad (25)$$

一方、1730段階では、ターゲットビット率が小さくて使うビットが不十分な場合、割り当てビット数が0よりは大きい最小ビット数より少ないサブバンドについては、割り当てビット数を回収して割り当てビット数を0に調整する。また、割り当てビット数が数式(24)のビット数より小さな場合に割り当てビット数を回収し、数式(24)のビット数より大きい数式(25)の最小ビット数より少ないサブバンドについては、最小ビット数を割り当てる。

10

【0160】

1740段階では、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和と、与えられたフレームの使用可能なビット数とを比較する。

【0161】

1750段階では、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数と同一になるまで、最小ビット数より多く割り当てられたサブバンドについてビットを再分配する。

【0162】

20

1760段階では、ビット再分配についての以前繰り返しの時、及び現在繰り返しの時に、各サブバンドの割り当てビット数に変動があるかどうかを判断し、ビット再分配についての以前繰り返しの時と現在繰り返しの時とに、各サブバンドの割り当てビット数に変動がないか、または与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数と同一になるまで1740段階ないし1760段階を行う。

【0163】

1770段階では、1760段階での判断結果、ビット再分配についての以前繰り返しの時及び現在繰り返しの時に、各サブバンドの割り当てビット数に変動がない場合、上位サブバンドから下位サブバンドに順次にビットを回収して、与えられたフレームの使用可能なビット数を満たすまで1740段階ないし1760段階を行う。

30

【0164】

すなわち、ビット再分配の時、最小ビット数 N_b より大きい値が割り当てられたサブバンドについて、ビットを低減させつつ可用ビット数を満たすように調整するが、すべてのサブバンドについて割り当てられたビット数が最小ビット数より大きくなく、かつ割り当てられたビット数の総和が依然として可用ビット数より大きい場合には、高周波サブバンドから順次にビットを回収する方式でビット数を調整する。

【0165】

図16及び図17に示されたビット割り当て方法によれば、各サブバンドにビットを割り当てるために、各サブバンドにスペクトルエネルギーあるいは加重されたスペクトルエネルギー順に初期ビットを割り当てた後、再びスペクトルエネルギーあるいは加重されたスペクトルエネルギーを探す動作を数回繰り返す必要なく、各サブバンドが必要とするビット数を一回に予測する。また、前記ビット割り当て方法によれば、与えられたフレームに含まれた各サブバンドについて推定された割り当てビット数の総和が、与えられたフレームの使用可能なビット数と同一になるまでビットを再分配することで、効率的なビット割り当てが可能である。また、前記ビット割り当て方法によれば、任意のサブバンドについて最小ビット数を保証することで、少ないビット数が割り当てられて十分な数のスペクトルサンプルあるいはパルスが符号化できなくてスペクトルホールが発生することを防止する。

40

【0166】

50

図14ないし図17の方法は、プログラミングされ、少なくとも一つのプロセッシングデバイスによって行われる。

【0167】

図18は、本発明の一実施形態による符号化モジュールを備えるマルチメディア機器の構成を示すブロック図である。

【0168】

図18に示されたマルチメディア機器1800は、通信部1810及び符号化モジュール1830を備える。また、符号化結果で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1850をさらに備える。また、マルチメディア機器1800は、マイクロフォン1870をさらに備える。すなわち、保存部1850及びマイクロフォン1870は、オプションで備えられる。一方、図18に示されたマルチメディア機器1800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行う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備える。ここで、符号化モジュール1830は、マルチメディア機器1800に備えられる他の構成要素（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）に具現される。

【0169】

図18を参照すれば、通信部1810は、外部から提供されるオーディオと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するか、または復元されたオーディオと符号化モジュール1830の符号化結果で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する。

【0170】

通信部1810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LAN、ワイファイ（Wi-Fi）、WFD（Wi-Fi Direct）、3G（Generation）、4G（4 Generation）、ブルートゥース、赤外線通信（IrDA、Infrared Data Association）、RFID（Radio Frequency Identification）、UWB（Ultra Wide Band）、ジグビー、NFC（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通じて、外部のマルチメディア機器とデータを送受信するように構成される。

【0171】

符号化モジュール1830は、一実施形態によれば、通信部1810あるいはマイクロフォン1870を通じて提供される時間ドメインのオーディオ信号を、周波数ドメインのオーディオスペクトルに変換し、オーディオスペクトルで与えられたフレームに対して使用可能なビット数範囲内で、所定の周波数バンドに存在するスペクトルのSNRを最大化させるように、周波数バンド単位で割り当てビット数を小数点単位で定め、周波数バンド単位で定められた割り当てビット数を調整し、周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いて、オーディオスペクトルを符号化してビットストリームを生成する。

【0172】

符号化モジュール1830は、他の実施形態によれば、通信部1810あるいはマイクロフォン1870を通じて提供される時間ドメインのオーディオ信号を周波数ドメインのオーディオスペクトルに変換し、オーディオスペクトルで与えられたフレームに含まれた周波数バンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定し、スペクトルエネルギーを用いて割り当てビット数を小数点単位で推定し、割り当てビット数が前記許容ビット数を超えないように調整し、周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いて、オーディオスペクトルを符号化してビットストリームを生成する。

【0173】

保存部1850は、符号化モジュール1830で生成される、符号化されたビットスト

10

20

30

40

50

リームを保存する。一方、保存部１８５０は、マルチメディア機器１８００の運用に必要な多様なプログラムを保存する。

【０１７４】

マイクロフォン１８７０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１８３０に提供する。

【０１７５】

図１９は、本発明の一実施形態による復号化モジュールを備えるマルチメディア機器の構成を示すブロック図である。図１９に示されたマルチメディア機器１８００は、通信部１９１０及び復号化モジュール１９３０を備える。また、復号化結果で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１９５０をさらに備える。また、マルチメディア機器１９００はスピーカー１９７０をさらに備える。すなわち、保存部１９５０及びスピーカー１９７０はオプションで備えられる。一方、図１９に示されたマルチメディア機器１９００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行う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備える。ここで、復号化モジュール１９３０は、マルチメディア機器１９００に備えられる他の構成要素（図示せず）と共に一体化され、少なくとも一つの以上のプロセッサ（図示せず）に具現される。

10

【０１７６】

図１９を参照すれば、通信部１９１０は、外部から提供される、符号化されたビットストリームとオーディオ信号のうち少なくとも一つを受信するか、または復号化モジュール１９３０の復号化結果で得られる復元されたオーディオ信号と、符号化結果で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する。一方、通信部１９１０は、図１８の通信部１８１０と実質的に類似して具現される。

20

【０１７７】

復号化モジュール１９３０は、一実施形態によれば、通信部１９１０を通じて提供されるビットストリームを受信し、与えられたフレームに対して使用可能なビット数範囲内で、各周波数バンドに存在するスペクトルのＳＮＲを最大化させるように周波数バンド単位で割り当てビット数を小数点単位で定め、周波数バンド単位で定められた割り当てビット数を調整し、周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いて、ビットストリームに含まれたオーディオスペクトルを復号化し、復号化されたオーディオスペクトルを時間ドメインのオーディオ信号に変換し、復元されたオーディオ信号を生成する。

30

【０１７８】

復号化モジュール１９３０は、他の実施形態によれば、通信部１９１０を通じて提供されるビットストリームを受信し、与えられたフレームに含まれた周波数バンド単位で、マスキングしきい値を用いて許容ビット数を小数点単位で推定し、スペクトルエネルギーを用いて割り当てビット数を小数点単位で推定し、割り当てビット数が前記許容ビット数を超えないように調整し、周波数バンド単位で調整されたビット数及びスペクトルエネルギーを用いてビットストリームに含まれたオーディオスペクトルを復号化し、復号化されたオーディオスペクトルを時間ドメインのオーディオ信号に変換し、復元されたオーディオ信号を生成する。

40

【０１７９】

保存部１９５０は、復号化モジュール１９３０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１９５０は、マルチメディア機器１９００の運用に必要な多様なプログラムを保存する。

【０１８０】

スピーカー１９７０は、復号化モジュール１９３０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。

【０１８１】

図２０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備える

50

マルチメディア機器の構成を示すブロック図である。図20に示されたマルチメディア機器2000は、通信部2010、符号化モジュール2020と復号化モジュール2030を備える。また、符号化結果で得られるオーディオビットストリーム、あるいは復号化結果で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2040をさらに備える。また、マルチメディア機器2000は、マイクロフォン2050あるいはスピーカー2060をさらに備える。ここで、符号化モジュール2020及び復号化モジュール2030は、マルチメディア機器2000に備えられる他の構成要素（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）に具現される。

【0182】

図20に示された各構成要素は、図18に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に示されたマルチメディア機器1900の構成要素と重なるため、その詳細な説明は略する。

【0183】

図18ないし図20に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、MP3プレーヤなどを含む放送あるいは音楽専用装置、あるいは音声通信専用端末と放送あるいは音楽専用装置の融合端末装置が含まれるが、こ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される変換器として使われる。

【0184】

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、図示されてはいないが、キーパッドなどのユーザ入力部、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備える。また、モバイルフォンは、撮像機能を持つカメラ部と、モバイルフォンで必要とする機能を行う少なくとも一つ以上の構成要素とをさらに備える。

【0185】

一方、マルチメディア機器1800、1900、2000が、例えば、TVである場合、図示されてはいないが、キーパッドなどのユーザ入力部、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備える。また、TVは、TVで必要とする機能を行う少なくとも一つ以上の構成要素をさらに備える。

【0186】

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作成でき、コンピュータで読み取り可能な記録媒体を用い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、前述した本発明の実施形態で使われるデータ構造、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通じて記録される。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られるデータが保存されるすべての保存装置を含む。コンピュータで読み取り可能な記録媒体の例には、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、CD-ROM、DVDのような光記録媒体、フロプティカルディスクのような磁気光媒体、及びROM（Read Only Memory）、RAM、フラッシュメモリなどのプログラム命令を保存して行う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あってもよい。プログラム命令の例には、コンパイラによって作られる同じ機械語コードだけではなく、インタプリタなどを使っ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む。

【0187】

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及び図面によって説明

10

20

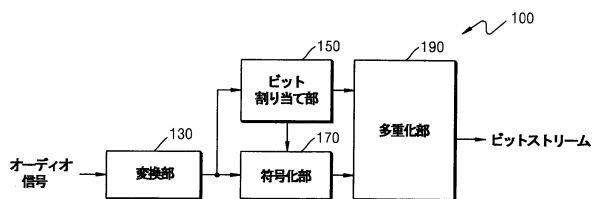
30

40

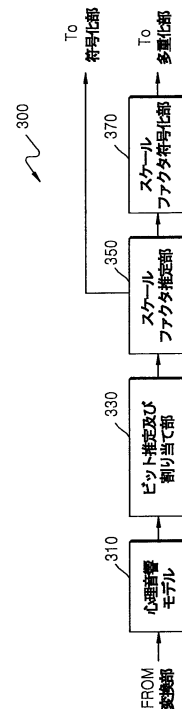
50

されたとしても、本発明の一実施形態は前記の実施形態に限定されるものではなく、当業者ならば、これより多様な修正及び変形が可能であろう。よって、本発明の範囲は前述した説明ではなく特許請求の範囲に示されており、この均等または等価的な変形はいずれも本発明の技術的思想の範ちゅうに属するといえる。

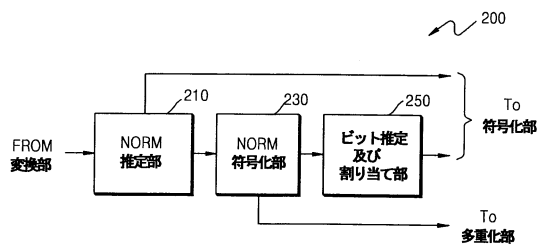
【図 1】



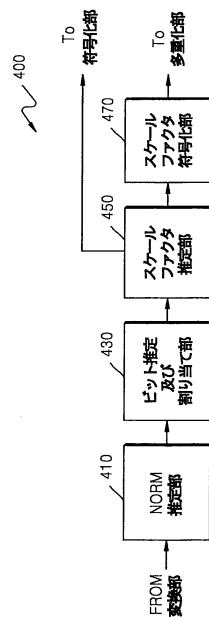
【図 3】



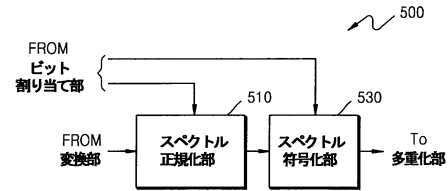
【図 2】



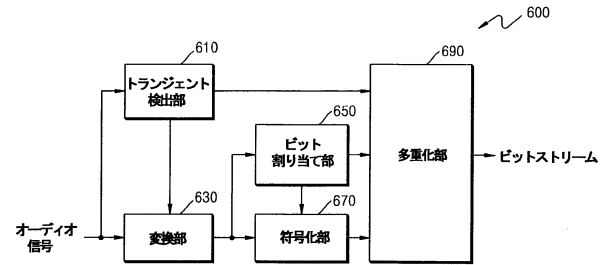
【図 4】



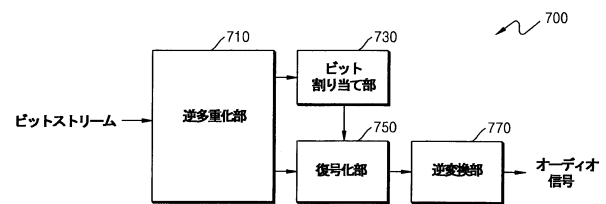
【図 5】



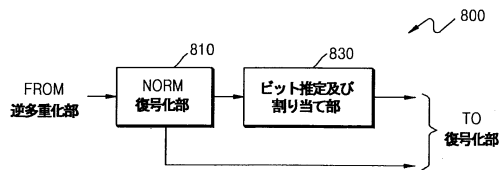
【図 6】



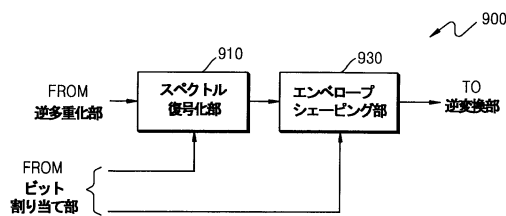
【図 7】



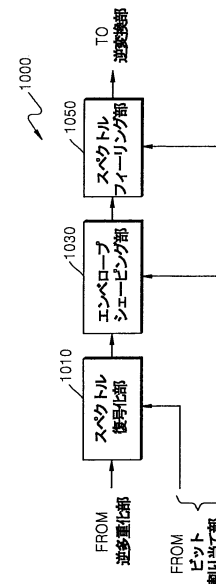
【図 8】



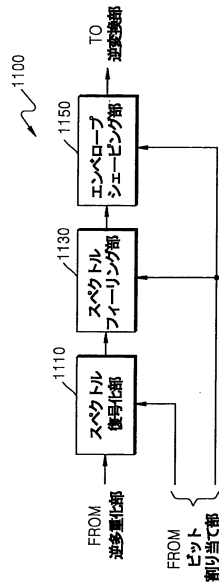
【図 9】



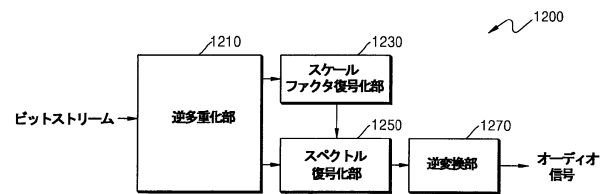
【図 10】



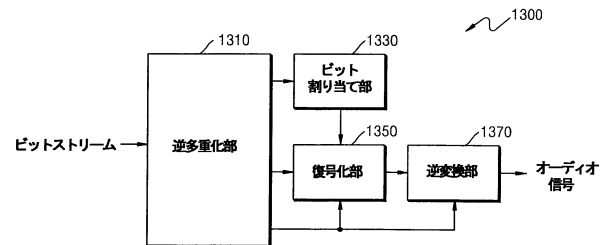
【図 1 1】



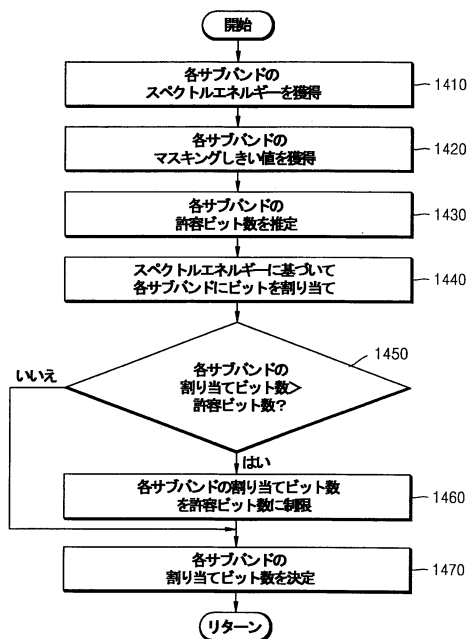
【図 1 2】



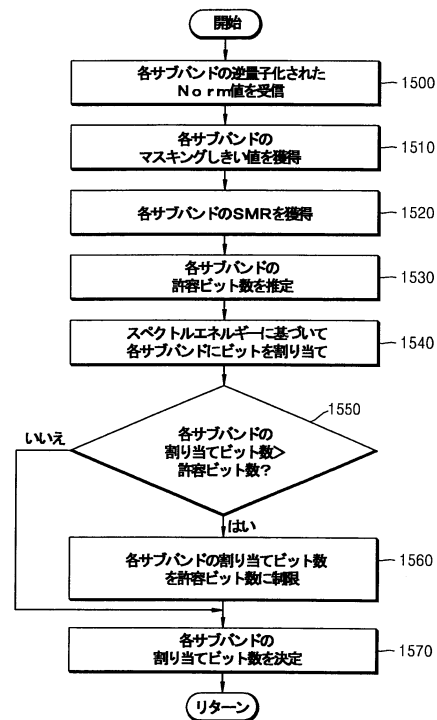
【図 1 3】



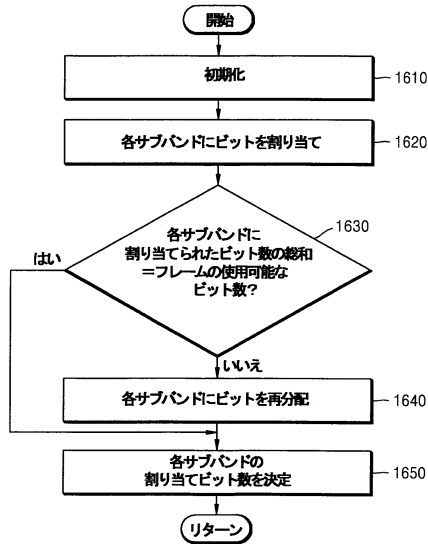
【図 1 4】



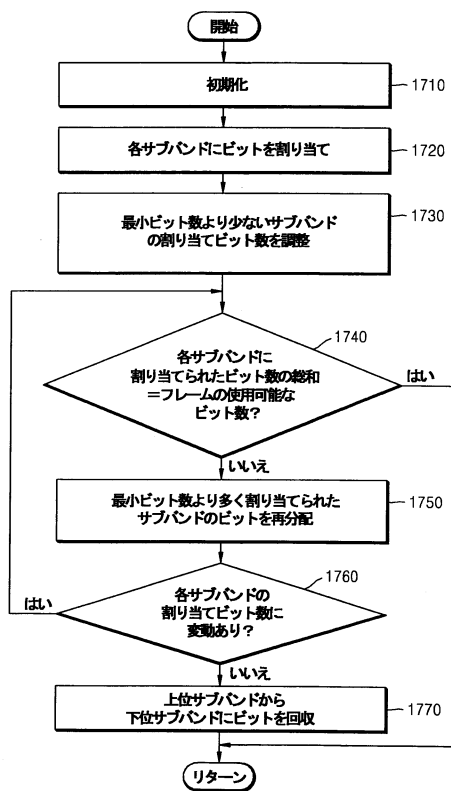
【図 1 5】



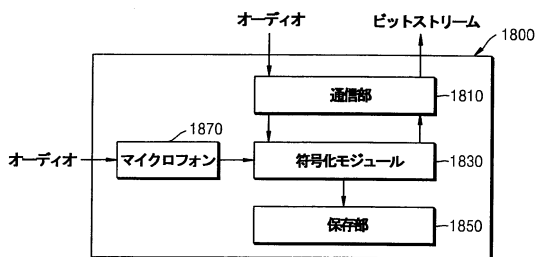
【図 16】



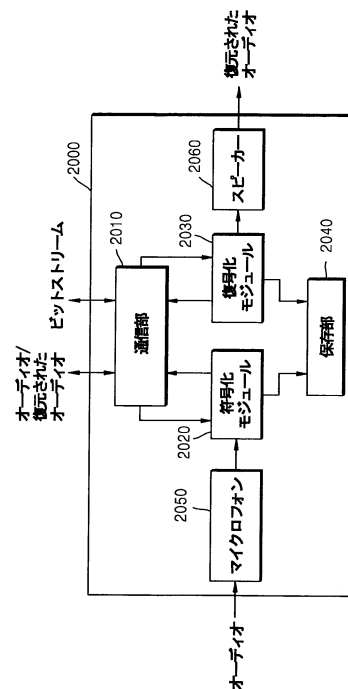
【図 17】



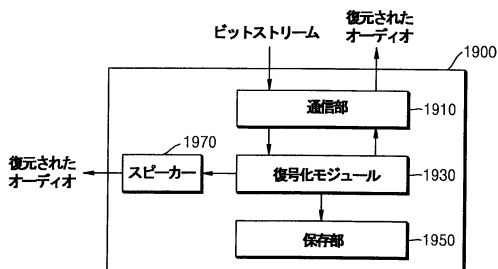
【図 18】



【図 20】



【図 19】



フロントページの続き

(72)発明者 キム, ミーヨン

大韓民国 445-170 キョンギード ファソナーシ ソックードン 53 ウミリンジェイ
ルブンギョンチェ・アパート 117-1603

(72)発明者 ポロフ, アントン

ロシア国 サンクト・ペテルスブルグ 190000 67 ボルシャ モルスカヤ・ストリート
サンクト・ペテルスブルグ・アカデミー・オブ・エアロスペース・インストゥルメンテーション

(72)発明者 オ, ウンミ

大韓民国 135-775 ソウル カンナムーグ テチ 2ードン ミド・アパート 101-
403 (番地なし)

審査官 安田 勇太

(56)参考文献 特開平04-168500 (JP, A)

特開平03-181232 (JP, A)

特開2000-293199 (JP, A)

特開2000-148191 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/02 -19/038