

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3776004号

(P3776004)

(45) 発行日 平成18年5月17日(2006.5.17)

(24) 登録日 平成18年3月3日(2006.3.3)

(51) Int. Cl.

G 1 0 L 19/02 (2006.01)

F I

G 1 0 L 19/02 1 4 2 B

G 1 0 L 19/02 1 4 2 D

請求項の数 2 (全 14 頁)

(21) 出願番号 特願2001-158767 (P2001-158767)  
 (22) 出願日 平成13年5月28日(2001.5.28)  
 (65) 公開番号 特開2002-351500 (P2002-351500A)  
 (43) 公開日 平成14年12月6日(2002.12.6)  
 審査請求日 平成15年2月7日(2003.2.7)

(73) 特許権者 000005049  
 シャープ株式会社  
 大阪府大阪市阿倍野区長池町22番22号  
 (74) 代理人 100085501  
 弁理士 佐野 静夫  
 (72) 発明者 藤井 修  
 大阪府大阪市阿倍野区長池町22番22号  
 シャープ株式会社内

審査官 山下 剛史

最終頁に続く

(54) 【発明の名称】 デジタルデータの符号化方法

(57) 【特許請求の範囲】

【請求項1】

楽音や音声等のデジタルデータを周波数領域に変換する手段を有するデジタルデータの符号化装置において、

前記周波数領域を複数の周波数帯域に分割する手段と、

前記各周波数帯域のパワーまたはエネルギーを算出する手段と、

前記算出された各周波数帯域のパワーまたはエネルギーの最大値及び平均値に基づき前記デジタルデータの純音性を判定する純音性判定手段と、

聴覚心理特性を反映した基準マスキング特性に基づきマスキング閾値を決定する第一のマスキング算出手段と、

周波数に重み付けを行っていない平坦なマスキング特性に基づきマスキング閾値を決定する第二のマスキング算出手段と、

前記純音性判定手段の判定結果に基づき前記第一のマスキング算出手段と前記第二のマスキング算出手段とを切替える切替手段と、

前記算出されたパワーまたはエネルギーと前記マスキング閾値とに基づきマスキング閾値対雑音比を算出する手段と、

前記マスキング閾値対雑音比に基づき前記各周波数帯域にビットを割り当てる手段と、を備えることを特徴とする符号化装置。

【請求項2】

楽音や音声等のデジタルデータを複数のサブバンドフレームに分割する手段と、

10

20

前記サブバンドフレーム単位毎に周波数領域に変換する手段と、  
 前記周波数領域を複数の周波数帯域に分割する手段と、  
 前記各周波数帯域のパワーを算出する手段と、  
 マスキング閾値を決定する手段と、

前記算出されたパワーと前記マスキング閾値とに基づきマスキング閾値対雑音比を算出する手段と、

前記マスキング閾値対雑音比に基づき前記各周波数帯域にビット割り当てを行う割り当て手段と、を備える符号化装置において、

前記算出されたパワーの最大値と平均値の差分値が所定値以上であり、かつ、前記パワーの最大値が存在する周波数が所定周波数以下である場合、前記パワーの最大値が存在するサブバンドフレームの全ての周波数帯域に、少なくとも最低量子化ビット数以上のビットを割り当てるよう前記割り当て手段に指示を行う手段を備えることを特徴とする符号化装置。

10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ミニディスク等の記録媒体に楽音や音声等のデジタルデータを記録するにあたって、前記楽音や音声等に適応して各周波数帯域のスペクトルに対するビット割り当てを行い、データ量を圧縮する符号化方法に関する。

【0002】

20

【従来の技術】

楽音や音声等のデジタルデータを高能率で圧縮符号化する方法として、ミニディスクで用いられているATRAC (Adaptive TRansform Acoustic Coding) が挙げられる。このATRACでは、デジタルデータを高能率で圧縮するために、入力デジタルデータは複数の周波数帯域(以下、適宜サブバンドフレームと呼ぶ)に分割され、可変長の単位時間でブロック化される。ブロック化されたデジタルデータはMDCT (Modified Discrete Cosine Transform) 処理によってスペクトル信号に変換され、さらに聴覚心理特性を利用して割り当てられたビット数で各スペクトル信号がそれぞれ符号化される。

【0003】

上記の圧縮符号化に適用することができる聴覚心理特性には、等ラウドネス特性やマスキング効果が挙げられる。等ラウドネス特性とは、同じ音圧レベルの音であっても、人間が感じ取る音の大きさが周波数によって変化することを表すものである。従って、人間が感じ取ることのできる音の大きさを示す最小可聴限がその音の周波数によって変化することを表している。

30

【0004】

一方、マスキング効果には同時マスキング効果と経時マスキング効果がある。同時マスキング効果とは、複数の周波数成分の音が同時に発生しているときに、ある音が別の音を聞き取り難くさせる現象を言う。また、経時マスキング効果とは、大きな音の時間軸方向の前後では、別の音が聞き取り難くなる現象を言う。

【0005】

40

このような聴覚心理特性を利用したビット割り当て法、例えば反復法と呼ばれる割り当て法では、入力されたデジタルデータに適応した実際のビット割り当てを、次のようにして行っている。

【0006】

まず、各周波数帯域のパワー $S$ を求め、そのパワー $S$ による他の周波数帯域に対するマスキング閾値 $M$ を求める。次に、このマスキング閾値 $M$ と、各周波数帯域を $n$ ビットで量子化したときの量子化雑音パワー $N(n)$ とから、マスキング閾値対雑音比 $MNR(n) = M/N(n)$ を求める。続いて、そのマスキング閾値対雑音比 $MNR(n)$ が最小となる周波数帯域にビットの割り当てを行った後、該マスキング閾値対雑音比 $MNR(n)$ を更新し、再び最小の周波数帯域にビットの割り当てが行われる。

50

## 【0007】

## 【発明が解決しようとする課題】

確かに、上記で説明した従来の符号化方法によれば、楽音や音声等のデジタルデータを高能率で圧縮符号化することができる。

## 【0008】

しかしながら、正弦波のように純音性の高いデジタルデータに対してマスキング閾値対雑音比MNR(n)を用いたビット割り当てを行うと、自身のパワー或いはエネルギーによって自身がマスキングの影響を受けてしまうため、信号対雑音比SNR(n) = S/N(n)を用いたビット割り当てを行った場合と比べて、符号化・復号化時の歪み率、S/N特性、及びダイナミックレンジといったオーディオ特性が悪化するという課題があった。

10

## 【0009】

また、超低域や超高域の正弦波に対してマスキング閾値対雑音比MNR(n)を用いたビット割り当てを行うと、最小可聴限もオーディオ特性の悪化要因となるおそれがあった。なお、超低域の正弦波に対して信号対雑音比SNR(n)を用いたビット割り当てを行った場合には、隣接するサブバンドフレーム間の分析窓のクロスポイントで、知覚可能な量子化誤差が生じるおそれがあった。

## 【0010】

一方、ホワイトノイズのように純音性の低いデジタルデータに対してマスキング閾値対雑音比MNR(n)を用いたビット割り当てを行うと、自身のパワー或いはエネルギーによって、マスキング閾値対雑音比MNR(n)が広帯域でフラットにならないため、信号対雑音比SNR(n)を用いたビット割り当てを行った場合と比べて、符号化・復号化時の音質が悪化するという課題があった。また、最小可聴限も音質の悪化要因となるおそれがあった。

20

## 【0011】

この点、本件出願人は、特開平10-207489号公報において、純音性の高いデジタルデータ或いは純音性の低いデジタルデータを符号化する場合、各周波数帯域の相互に隣接するスペクトルのパワーS(或いはエネルギー)の差から求めたピーク及びローカルピークとマスキング閾値Mとの関係に対応して、マスキング閾値対雑音比MNR(n)を用いたビット割り当てを行うビットレートと、信号対雑音比SNR(n)を用いたビット

30

## 【0012】

確かに、上記の符号化方法によれば、正弦波のように狭帯域なデジタルデータから、ホワイトノイズのように広帯域なデジタルデータに至るまで、最適なビット割り当てを自動的に行うことができるので、マスキング閾値対雑音比MNR(n)等の同時マスキングを利用したビット割り当てに不向きな楽音に対しても音質の劣化を防止することができる。しかしながら、マスキング閾値対雑音比MNR(n)と信号対雑音比SNR(n)とを併用する上記の符号化方法では、アルゴリズムが複雑なものとなっていた。

## 【0013】

本発明は上記の問題点に鑑み、アルゴリズムを複雑化することなく、純音性の高いデジタルデータから純音性の低いデジタルデータまで、高忠実に符号化することが可能なデジタルデータの符号化方法を提供することを第1の目的とする。また、本発明は、超低域であってかつ純音性の高いデジタルデータを符号化する際であっても、隣接するサブバンドフレーム間の分析窓のクロスポイントで、知覚可能な量子化誤差が生じるおそれの少ないデジタルデータの符号化方法を提供することを第2の目的とする。

40

## 【0014】

## 【課題を解決するための手段】

上記した第1の目的を達成するために、本発明に係るデジタルデータの符号化装置は、楽音や音声等のデジタルデータを周波数領域に変換する手段を有するデジタルデー

50

タの符号化装置において、

前記周波数領域を複数の周波数帯域に分割する手段と、

前記各周波数帯域のパワーまたはエネルギーを算出する手段と、

前記算出された各周波数帯域のパワーまたはエネルギーの最大値及び平均値に基づき前記デジタルデータの純音性を判定する純音性判定手段と、

聴覚心理特性を反映した基準マスキング特性に基づきマスキング閾値を決定する第一のマスキング算出手段と、

周波数に重み付けを行っていない平坦なマスキング特性に基づきマスキング閾値を決定する第二のマスキング算出手段と、

前記純音性判定手段の判定結果に基づき前記第一のマスキング算出手段と前記第二のマスキング算出手段とを切替える切替手段と、 10

前記算出されたパワーまたはエネルギーと前記マスキング閾値とに基づきマスキング閾値対雑音比を算出する手段と、

前記マスキング閾値対雑音比に基づき前記各周波数帯域にビットを割り当てる手段と、を備えることを特徴とする。

【 0 0 1 6 】

また、上記した第2の目的を達成するために、本発明に係るデジタルデータの符号化装置は、楽音や音声等のデジタルデータを複数のサブバンドフレームに分割する手段と、

前記サブバンドフレーム単位毎に周波数領域に変換する手段と、 20

前記周波数領域を複数の周波数帯域に分割する手段と、

前記各周波数帯域のパワーを算出する手段と、

マスキング閾値を決定する手段と、

前記算出されたパワーと前記マスキング閾値とに基づきマスキング閾値対雑音比を算出する手段と、

前記マスキング閾値対雑音比に基づき前記各周波数帯域にビット割り当てを行う割り当て手段と、を備える符号化装置において、

前記算出されたパワーの最大値と平均値の差分値が所定値以上であり、かつ、前記パワーの最大値が存在する周波数が所定周波数以下である場合、前記パワーの最大値が存在するサブバンドフレームの全ての周波数帯域に、少なくとも最低量子化ビット数以上のビットを割り当てるよう前記割り当て手段に指示を行う手段を備えることを特徴とする。 30

【 0 0 1 7 】

【発明の実施の形態】

本発明に係るデジタルデータの符号化方法を採用したデジタル録音再生装置として、ここでは、ミニディスク録音再生装置を例に挙げて説明を行う。図1は本発明に係るデジタルデータの符号化方法を採用したミニディスク録音再生装置の一構成例を示すブロック図である。

【 0 0 1 8 】

本図に示すミニディスク録音再生装置1に設けられた入力端子2には、コンパクトディスク再生装置や衛星放送受信装置などのデジタル音声信号源から出力されたデジタル音声データが、例えば光信号としてシリアル入力される。入力端子2に入力された光信号は、光電素子3によって電気信号に変換された後、デジタルPLL(Phase-Locked-Loop)回路4に入力される。 40

【 0 0 1 9 】

デジタルPLL回路4は、入力されたデジタル音声データからクロックの抽出を行うとともに、サンプリング周波数および量子化ビット数に対応したマルチビットデータを再現する。このマルチビットデータは信号源毎に異なるサンプリングレート(コンパクトディスク; 44.1kHz、デジタルオーディオテープレコーダ; 48kHz、衛星放送(Aモード); 32kHzなど)で標本化されたデジタルデータである。そこで、デジタルPLL回路4から出力されたマルチビットデータは、周波数変換回路5によってそ 50

のサンプリングレートをミニディスクの規格に対応した44.1kHzに変換される。

【0020】

音声圧縮回路6は、ATRAC(Adaptive Transform Acoustic Coding)方式によって入力されたデジタル音声データの圧縮符号化を行い、符号化されたデジタル音声データをショックブーフメモリコントローラ7を介して信号処理回路8に送出する。なお、音声圧縮回路6におけるデジタルデータの符号化方法については、後ほど詳細に説明を行う。

【0021】

ショックブーフメモリコントローラ7で制御されるショックブーフメモリ9は、音声圧縮回路6から出力されるデジタル音声データの転送速度と、信号処理回路8に入力されるデジタル音声データの転送速度との差を吸収するとともに、再生時における振動等の外乱による再生信号の中断を補間し、デジタル音声データを保護するためのものである。

10

【0022】

信号処理回路8は、エンコーダおよびデコーダとしての機能を備えており、入力されたデジタル音声データをシリアル磁界変調信号にエンコードしてヘッド駆動回路10に与える。ヘッド駆動回路10は、記録ヘッド11をミニディスク12上の所定記録位置に移動させるとともに、前記磁界変調信号に対応した磁界を発生させる。このとき、ミニディスク12上の所定記録位置には、光ピックアップ13からレーザ光が照射されており、これによって前記磁界に対応した磁化パターンがミニディスク12上に形成される。

20

【0023】

一方、光ピックアップ13は、ミニディスク12から前記磁化パターンに対応したシリアル信号を再生する。再生されたシリアル信号は高周波アンプ14(以下、RFアンプ14と呼ぶ)で増幅された後、信号処理回路8によってデジタル音声データにデコードされる。デコードされたデジタル音声データは、ショックブーフメモリコントローラ7及びショックブーフメモリ9で外乱による影響を除去された後、音声伸長回路15に送出される。

【0024】

音声伸長回路15は、ATRAC方式による圧縮符号化の逆変換処理を行い、フルビットのデジタル音声データを復調する。復調されたデジタル音声データは、デジタル/アナログ変換回路16(以下、D/A変換回路16と呼ぶ)によってアナログ音声データに変換され、出力端子17から外部へ出力される。

30

【0025】

なお、RFアンプ14で増幅されたシリアル信号は、サーボ回路18にも入力されている。サーボ回路18は、再生されたシリアル信号に応じてドライバ回路19に制御信号を送出し、該ドライバ回路19を介してスピンドルモータ20の回転速度をフィードバック制御する。このようなフィードバック制御により、ミニディスク12を線速度一定で回転させることができる。

【0026】

また、サーボ回路18は、ドライバ回路19を介して送りモータ21の回転速度もフィードバック制御している。このようなフィードバック制御により、ミニディスク12の半径方向に対する光ピックアップ13の変移制御、すなわちトラッキング制御を行うことができる。さらに、サーボ回路18は、ドライバ回路19を介して光ピックアップ13のフォーカシング制御も行っている。

40

【0027】

上記した信号処理回路8、光ピックアップ13、RFアンプ14、サーボ回路18、及びドライバ回路19等には、図示しない電源回路から電力供給が行われるが、このような電力供給動作や後述する信号処理動作は、全てシステムコントロールマイコン22によって集中管理されている。なお、システムコントロールマイコン22には、曲名入力や選曲操作、或いは音質調整動作等を行うための入力装置23が接続されている。

50

## 【0028】

続いて、上記した音声圧縮回路6におけるデジタルデータ符号化処理の第1実施形態について説明する。図2は音声圧縮回路6の第1実施形態を示すブロック図であり、特に、スペクトル変換部に続くビット割当処理部を模式化したものである。

## 【0029】

本図に示すビット割当処理部の入力端には、その前段に設けられたスペクトル変換部（図示せず）で得られたMDC T係数（デジタル音声データを構成する周波数成分（スペクトル））が入力される。なお、スペクトル変換部は、周波数変換回路5から入力されたデジタル音声データ（44.1kHz）を、帯域分割フィルタであるQMF（Quadrature Mirror Filter）によって複数の周波数帯域（サブバンドフレーム）に分割し、そのサブバンドフレーム単位毎にMDC T（Modified Discrete Cosine Transform）処理を施すことで、デジタル音声データのスペクトル変換を行っている。

10

## 【0030】

パワー算出部31は、入力されたMDC T係数をさらに*i*個の周波数帯域（臨界帯域等）に分割し、各周波数帯域に属するMDC T係数の2乗和から、各周波数帯域のスペクトルパワー $S_i$ （ $i = 1, 2, \dots, I$ 、例えば $I = 25$ ）を算出する。なお、臨界帯域とは、周波数選択性・マスキング閾値等の特定の音響心理学的規則性が有効な広帯域オーディオスペクトルの特性的部分のことである。

## 【0031】

純音性判定部32は、パワー算出部31で算出されたスペクトルパワー $S_i$ の最大値 $S_{max}$ と平均値 $S_{av}$ （ $= S_i / I$ ）との差分値（ $S_{max} - S_{av}$ ）を求めるとともに、該差分値の大小からデジタル音声データの純音性の高低を判定し、その判定結果に基づいて切換部33の切換制御を行う。

20

## 【0032】

図3はパワー算出部31で算出されたスペクトルパワー $S_i$ の一例を示す図である。本図中（a）に示すように、スペクトルパワー $S_i$ の最大値 $S_{max}$ と平均値 $S_{av}$ との差分値（ $S_{max} - S_{av}$ ）が非常に大きい場合、例えば $S_{max} - S_{av} > 40$  dBを満たす場合、純音性判定部32は、入力されたデジタル音声データの純音性が高いと判定して、平坦マスキング算出部35を選択するように切換部33の切換制御を行う。

## 【0033】

また、本図中（b）に示すように、スペクトルパワー $S_i$ の最大値 $S_{max}$ と平均値 $S_{av}$ との差分値（ $S_{max} - S_{av}$ ）が非常に小さい場合、例えば $S_{max} - S_{av} < 6$  dBを満たす場合、純音性判定部32は、入力されたデジタル音声データの純音性が低いと判定して、上記と同様、平坦マスキング算出部35を選択するように切換部33の切換制御を行う。

30

## 【0034】

一方、スペクトルパワー $S_i$ の最大値 $S_{max}$ と平均値 $S_{av}$ との差分値（ $S_{max} - S_{av}$ ）が上記のいずれにも該当しない場合、例えば $6$  dB  $< S_{max} - S_{av} < 40$  dBを満たす場合、純音性判定部32は、入力されたデジタル音声データに対する聴覚心理、すなわちマスキング効果が有効であると判断して、基準マスキング算出部34を選択するように切換部33の切換制御を行う。

40

## 【0035】

上記の純音性判定動作により、基準マスキング算出部34が選択された場合、最小可聴限合成部36は、音声圧縮部6のテーブルROM（図示せず）に予め格納されている基準マスキング特性と最小可聴限特性とを合成することで、最終的なマスキング閾値 $M_i$ を決定する。一方、平坦マスキング算出部35が選択された場合、最小可聴限合成部36は、周波数に重み付けを行っていない平坦なマスキング特性と最小可聴限特性とを合成することで、最終的なマスキング閾値 $M_i$ を決定する。

## 【0036】

SMR算出部37は、各周波数帯域のインデックスを前記*i*とするとき、パワー算出部31で算出されたスペクトルパワー $S_i$ と、最小可聴限合成部36で決定された各周波数帯

50

域のマスクング閾値  $M_i$  との比  $SMR_i (= S_i / M_i)$  を、全ての周波数帯域に亘って計算する。

【0037】

MNR 算出部 38 は、まず各周波数帯域のスペクトルパワー  $S_i$  と、該スペクトルパワー  $S_i$  を  $n$  ビットで量子化したときに生じる量子化雑音パワー  $N_i(n)$  との比、すなわち信号対雑音比  $SNR_i(n) (= S_i / N_i(n))$  を求める。なお、この信号対雑音比  $SNR_i(n)$  は、統計的には信号特性に応じた定数となるので、予め統計処理によって求めておいてもよい。さらに、MNR 算出部 38 は、この信号対雑音比  $SNR_i(n)$  と、SMR 算出部 37 で得られた比  $SMR_i$  から、マスクング閾値  $M_i$  と量子化雑音パワー  $N_i$  との比、すなわちマスクング閾値対雑音比  $MNR_i(n) (= SNR_i(n) / SMR_i)$  を算出する。

10

【0038】

量子化ビット数算出部 39 は、各周波数帯域の量子化ビット数  $n$  を 0 から順に大きくしていき、その都度、各周波数帯域のマスクング閾値対雑音比  $MNR_i(n)$  を計算する。そして、マスクング閾値対雑音比  $MNR_i(n)$  が最小となる周波数帯域から順にビットを割り当てていく。その後、量子化ビット数  $n$  を更新する度毎に、マスクング閾値対雑音比  $MNR_i(n)$  が最小となる周波数帯域に対する同様のビット割り当てを行う。所定の割り当て可能ビット数となるまで割り当てを行うと、各周波数帯域の語長が決定されて出力が行われる。すなわち、スペクトルパワー  $S_i$  の絶対値が、マスクング閾値  $M_i$  を超えた部分の長さが最も長い周波数帯域から順にビット割り当てが行われることになる。

20

【0039】

上記したデジタルデータの符号化方法であれば、正弦波のように純音性の高いデジタルデータ、或いはホワイトノイズのように純音性の低いデジタルデータに対して、マスクング閾値対雑音比  $MNR_i(n)$  のみを用いてビット割り当てを行った場合であっても、信号対雑音比  $SNR_i(n)$  を用いてビット割り当てを行った場合と同等のオーディオ特性及び音質を得ることができる。

【0040】

また、本実施形態におけるデジタルデータの符号化方法であれば、楽音や音声のように、聴覚心理を利用した方が好ましい音源には、通常のマスキング閾値対雑音比  $MNR_i(n)$  を用いたビット割り当てを行うので、信号対雑音比  $SNR_i(n)$  を用いてビット割り当てを行うよりも、聴覚的に優れた音質を得ることができる。さらに、マスクング閾値対雑音比  $MNR_i(n)$  と信号対雑音比  $SNR_i(n)$  とを併用する従来の符号化方法に比べて、アルゴリズムを容易に実現することが可能である。

30

【0041】

続いて、音声圧縮回路 6 におけるデジタルデータ符号化処理の第 2 実施形態について説明する。図 4 は音声圧縮回路 6 の第 2 実施形態を示すブロック図であり、説明の理解を深めるために音声伸長回路 15 も合わせて示している。

【0042】

本図に示す音声圧縮回路 6 の入力端には、周波数変換回路 5 で得られたデジタル音声データ (44 . 1 kHz) が入力される。音声圧縮回路 6 の最前段に設けられた周波数帯域分割部 41 は、入力されたデジタル音声データを複数の周波数帯域 (サブバンドフレーム) に分割する。

40

【0043】

時間周波数変換部 42 は、周波数帯域分割部 41 で得られたサブバンドフレーム単位毎に MDCT 処理を施すことで、デジタル音声データを MDCT 係数に変換する。このときの MDCT 処理によって得られる変換データ  $X_m(k)$  は、次の (1) 式で示される。

【数 1】

$$X_m(k) = \sum_{i=0}^{2M-1} x_m(i) \cdot h(i) \cdot \cos \left\{ \frac{\pi}{M} \cdot \left( k + \frac{1}{2} \right) \cdot \left( i + \frac{M}{2} + \frac{1}{2} \right) \right\} \quad \cdots (1)$$

(k=0, ..., M-1)

## 【0044】

なお、上式中の変数  $m$  はブロック番号を表しており、関数  $x_m(i)$  は入力信号を表している。また、関数  $h(i)$  は順変換用窓関数を表している。図5は順変換窓関数  $h(i)$  の時間特性の一例を示す概念図であり、図6は順変換窓関数  $h(i)$  の周波数特性の一例を示す概念図である。

10

## 【0045】

パワー算出部43は、時間周波数変換部42で得られたMDC T係数をさらに  $i$  個の周波数帯域（臨界帯域等）に分割し、各周波数帯域に属するMDC T係数の2乗和から、各周波数帯域のスペクトルパワー  $S_i$  ( $i = 1, 2, \dots, I$ 、例えば  $I = 25$ ) を算出する。

## 【0046】

純音性判定部44は、パワー算出部43で算出されたスペクトルパワー  $S_i$  の最大値  $S_{max}$  と平均値  $S_{av}$  ( $= S_i / I$ ) との差分値 ( $S_{max} - S_{av}$ ) を求めるとともに、該差分値の大小からデジタル音声データの純音性の高低を判定し、その判定結果に基づいて量子化ビット数算出部49における量子化ビット数の割り当て制御を行う。

20

## 【0047】

図7は本実施形態における量子化ビット数の割り当て制御を説明するための図であり、パワー算出部31で算出されたスペクトルパワー  $S_i$  の一例 (a) と、その際に割り当てられる量子化ビット数の一例 (b) と、を示している。なお、本図では、入力されたデジタル音声データが4つのサブバンドフレーム  $SB1 \sim SB4$  に分割されている場合を例に挙げて説明を行う。

## 【0048】

本図中 (a) に示すように、スペクトルパワー  $S_i$  の最大値  $S_{max}$  と平均値  $S_{av}$  との差分値 ( $S_{max} - S_{av}$ ) が非常に大きく (例えば  $S_{max} - S_{av} = 40 \text{ dB}$ )、かつスペクトルパワー  $S_i$  の最大値  $S_{max}$  が存在する周波数が所定周波数 (例えば  $100 \text{ Hz}$ ) 以下である場合、純音性判定部44は、入力されたデジタル音声データが超低域であるとともに純音性が高いと判定して、本図中 (b) に示すように、スペクトルパワー  $S_i$  の最大値  $S_{max}$  が存在するサブバンドフレーム  $SB1$  に、少なくとも最低の量子化ビット数を割り当てるよう、量子化ビット数算出部49に対する指示を行う。

30

## 【0049】

このようなビット割り当てを行うことにより、特定周波数のノイズを低減することが可能となる。従って、超低域であるとともに純音性が高いデジタル音声データ (例えば、超低域の正弦波) を符号化する場合であっても、隣接するサブバンドフレーム間の分析窓のクロスポイントで、知覚可能な量子化誤差が生じるおそれが少なくなる。

## 【0050】

パワー算出部43の後段に接続されたマスキング算出部45、最小可聴限合成部46、SMR算出部47、MNR算出部48、及び量子化ビット数算出部49は、前述の第1実施形態と同様、マスキング閾値対雑音比  $MNR_i(n)$  を用いたビット割り当てを行い、量子化ビット数を決定する。

40

## 【0051】

量子化部50及びパッキング部51は、量子化ビット数算出部49で得られた量子化ビット数に従って、入力されたデジタル音声データを圧縮符号化する。このようにして圧縮符号化されたデジタル音声データは、信号処理回路8などを介してミニディスク12に記録される。

## 【0052】

50

一方、ミニディスク12を再生する際、音声伸長回路15のアンパッキング部52及び逆量子化部53は、圧縮符号化されたデジタル音声データを元のMDC T係数に復元する。

【0053】

周波数時間変換部54は、復元されたMDC T係数に対して、サブバンドフレーム単位毎にIMDC T (Inverse Modified Discrete Cosine Transform) 処理を施す。このときのIMDC T処理によって得られる復元信号 $\hat{x}_m(i)$ は、次の(2)式で示される。

【数2】

$$\left. \begin{aligned} \hat{x}_m(i) &= y_{m-1}(i+M) \cdot f(i+M) + y_m(i) \cdot f(i) \\ &\quad (i=0, \dots, M-1) \\ y_m(i) &= \frac{2}{M} \sum_{k=0}^{M-1} X_m(k) \cdot \cos \left\{ \frac{\pi}{M} \cdot \left( k + \frac{1}{2} \right) \cdot \left( i + \frac{M}{2} + \frac{1}{2} \right) \right\} \\ &\quad (k=0, \dots, 2M-1) \end{aligned} \right\} \dots (2)$$

【0054】

なお、上式中の変数 $m$ はブロック番号を表しており、関数 $X_m(k)$ は変換データ(復元されたMDC T係数)を表している。また、関数 $y_m(i)$ は逆変換信号を表しており、関数 $f(i)$ は逆変換用窓関数を表している。

【0055】

続く周波数帯域合成部55は、周波数時間変換部54によって得られた復元信号 $\hat{x}_m(i)$ を合成することで元のデジタル音声データを復元し、該デジタル音声データを次段のD/A変換回路16に送出する。

【0056】

なお、変換データ $X_m(k)$ が量子化による影響を受けることなく、復元信号 $\hat{x}_m(i)$ から元のデジタル音声データを復元できるように、上記した時間周波数変換部42及び周波数時間変換部54は、次の(3)式を満たすように設計されるべきである。本条件はCAS90-10やDSP90-14等により既に公知とされている。

【数3】

$$\left. \begin{aligned} h(i) \cdot f(i) + h(i+M) \cdot f(i+M) &= 1 \\ h(2M-i-1) \cdot f(i+M) &= h(M-i-1) \\ &\quad (i=0, \dots, M-1) \end{aligned} \right\} \dots (3)$$

【0057】

一方、変換データ $X_m(k)$ が量子化による影響を受けた場合について、参考までに説明する。図8は変換データ $X_m(k)$ が量子化による影響を受けた場合を説明する図であり、周波数帯域合成部55から出力されるデジタル音声データの一例(a)と、該デジタル音声データの符号化時におけるビット割り当ての一例(b)と、を示している。

【0058】

本図中(b)に示すように、サブバンドフレームSB1の低域から1/3程度の周波数成分にのみビット割り当てを行った場合、復元信号 $\hat{x}_m(i)$ の周波数特性は本図中(a)のようになり、窓関数の周波数特性のノイズが発生する。このような量子化ノイズは、サブバンドフレームSB1のどの周波数にでも少なからず発生するが、この例ではサブバンドフレームSB1の低域から1/3程度に集中してビットを割り当てているため、残りの2/3の周波数成分における量子化ノイズが知覚されやすい。従って、前述した通り、スペクトルパワー $S_i$ の最大値 $S_{max}$ が存在するサブバンドフレームSB1に、少なくとも最低の量子化ビット数を割り当てれば、特定周波数のノイズを低減することが可能となる

10

20

30

40

50

## 【 0 0 5 9 】

なお、上記の実施形態では、本発明に係るデジタルデータの符号化方法をミニディスク録音再生装置に適用した例を挙げて説明を行ったが、本発明の適用範囲がこれに限定されないことは言うまでもない。

## 【 0 0 6 0 】

## 【 発明の効果 】

本発明に係るデジタルデータの符号化装置は、楽音や音声等のデジタルデータを周波数領域に変換する手段を有するデジタルデータの符号化装置において、

前記周波数領域を複数の周波数帯域に分割する手段と、

前記各周波数帯域のパワーまたはエネルギーを算出する手段と、

前記算出された各周波数帯域のパワーまたはエネルギーの最大値及び平均値に基づき前記デジタルデータの純音性を判定する純音性判定手段と、

聴覚心理特性を反映した基準マスキング特性に基づきマスキング閾値を決定する第一のマスキング算出手段と、

周波数に重み付けを行っていない平坦なマスキング特性に基づきマスキング閾値を決定する第二のマスキング算出手段と、

前記純音性判定手段の判定結果に基づき前記第一のマスキング算出手段と前記第二のマスキング算出手段とを切替える切替手段と、

前記算出されたパワーまたはエネルギーと前記マスキング閾値とに基づきマスキング閾値対雑音比を算出する手段と、

前記マスキング閾値対雑音比に基づき前記各周波数帯域にビットを割り当てる手段と、を備える。

## 【 0 0 6 2 】

このようなデジタルデータの符号化方法であれば、正弦波のように純音性の高いデジタルデータ、或いはホワイトノイズのように純音性の低いデジタルデータに対して、マスキング閾値対雑音比のみを用いてビット割り当てを行った場合であっても、信号対雑音比を用いてビット割り当てを行った場合と同等のオーディオ特性及び音質を得ることができる。

## 【 0 0 6 3 】

また、本発明に係るデジタルデータの符号化方法であれば、楽音や音声のように、聴覚心理を利用した方が好ましい音源には、通常のマスキング閾値対雑音比を用いたビット割り当てを行うので、信号対雑音比を用いてビット割り当てを行うよりも、聴覚的に優れた音質を得ることができる。さらに、マスキング閾値対雑音比と信号対雑音比とを併用する従来の符号化方法に比べて、アルゴリズムを容易に実現することが可能である。

## 【 0 0 6 4 】

また、本発明に係るデジタルデータの符号化装置は、楽音や音声等のデジタルデータを複数のサブバンドフレームに分割する手段と、

前記サブバンドフレーム単位毎に周波数領域に変換する手段と、

前記周波数領域を複数の周波数帯域に分割する手段と、

前記各周波数帯域のパワーを算出する手段と、

前記各周波数帯域にビット割り当てを行う割り当て手段と、を備える符号化装置において、

前記算出されたパワーの最大値と平均値の差分値が所定値以上であり、かつ、前記パワーの最大値が存在する周波数が所定周波数以下である場合、前記パワーの最大値が存在するサブバンドフレームの全ての周波数帯域に、少なくとも最低量子化ビット数以上のビットを割り当てるよう前記割り当て手段に指示を行う手段を備える。

## 【 0 0 6 5 】

このようなデジタルデータの符号化方法であれば、特定周波数のノイズを低減することが可能となる。従って、超低域であるとともに純音性が高いデジタル音声データ（例え

10

20

30

40

50

ば、超低域の正弦波)を符号化する場合であっても、隣接するサブバンドフレーム間の分析窓のクロスポイントで、知覚可能な量子化誤差が生じるおそれが少なくなる。

【図面の簡単な説明】

【図1】 本発明に係るデジタルデータの符号化方法を採用したミニディスク録音再生装置の一構成例を示すブロック図である。

【図2】 音声圧縮回路6の第1実施形態を示すブロック図である。

【図3】 パワー算出部31で算出されたスペクトルパワー $S_i$ の一例を示す図である。

【図4】 音声圧縮回路6の第2実施形態を示すブロック図である。

【図5】 順変換窓関数 $h(i)$ の時間特性の一例を示す概念図である。

【図6】 順変換窓関数 $h(i)$ の周波数特性の一例を示す概念図である。

10

【図7】 第2実施形態における量子化ビット数の割り当て制御を説明するための図である。

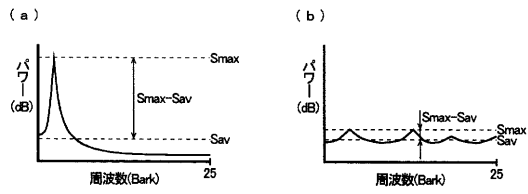
【図8】 変換データ $X_m(k)$ が量子化による影響を受けた場合を説明する図である。

【符号の説明】

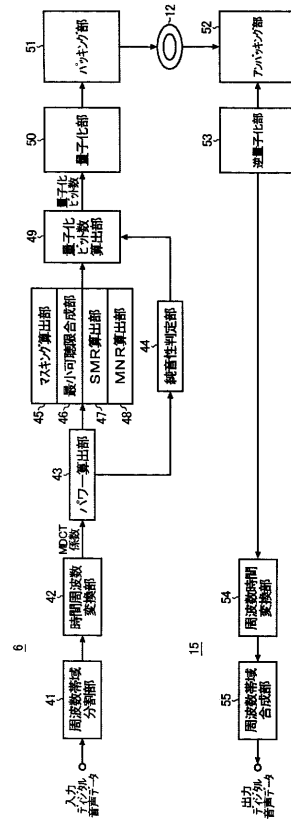
- |    |                        |    |
|----|------------------------|----|
| 1  | ミニディスク録音再生装置           |    |
| 2  | 入力端子                   |    |
| 3  | 光電素子                   |    |
| 4  | デジタルPLL回路              |    |
| 5  | 周波数変換回路                |    |
| 6  | 音声圧縮回路                 | 20 |
| 7  | ショックプルーフメモリコントローラ      |    |
| 8  | 信号処理回路                 |    |
| 9  | ショックプルーフメモリ            |    |
| 10 | 記録ヘッド駆動回路              |    |
| 11 | 記録ヘッド                  |    |
| 12 | ミニディスク                 |    |
| 13 | 光ピックアップ                |    |
| 14 | 高周波アンプ(RFアンプ)          |    |
| 15 | 音声伸長回路                 |    |
| 16 | デジタル/アナログ変換回路(D/A変換回路) | 30 |
| 17 | 出力端子                   |    |
| 18 | サーボ回路                  |    |
| 19 | ドライバ回路                 |    |
| 20 | スピンドルモータ               |    |
| 21 | 送りモータ                  |    |
| 22 | システムコントロールマイコン         |    |
| 23 | 入力装置                   |    |
| 31 | パワー算出部                 |    |
| 32 | 純音性判定部                 |    |
| 33 | 切換部                    | 40 |
| 34 | 基準マスキング算出部             |    |
| 35 | 平坦マスキング算出部             |    |
| 36 | 最小可聴限合成部               |    |
| 37 | SMR算出部                 |    |
| 38 | MNR算出部                 |    |
| 39 | 量子化ビット数算出部             |    |
| 41 | 周波数帯域分割部               |    |
| 42 | 時間周波数変換部               |    |
| 43 | パワー算出部                 |    |
| 44 | 純音性判定部                 | 50 |



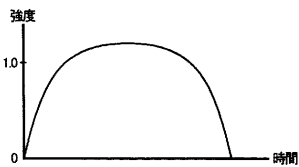
【図3】



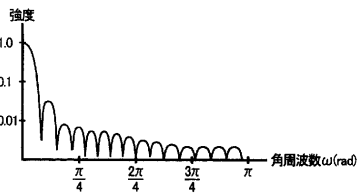
【図4】



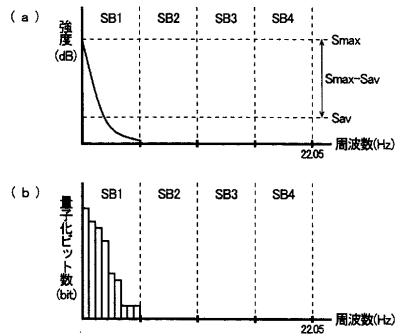
【図5】



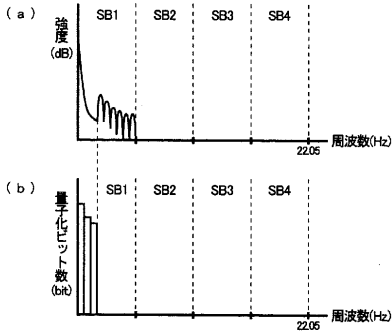
【図6】



【図7】



【図8】



---

フロントページの続き

- (56)参考文献 国際公開第94/028633(WO, A1)  
特開平08-211899(JP, A)  
特開平08-272400(JP, A)  
特開平07-295594(JP, A)

- (58)調査した分野(Int.Cl., DB名)  
G10L 19/00-19/14  
G11B 20/10  
H03M 7/30