

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4645869号
(P4645869)

(45) 発行日 平成23年3月9日(2011.3.9)

(24) 登録日 平成22年12月17日(2010.12.17)

(51) Int.Cl.	F 1		
G 1 O L 19/00	(2006.01)	G 1 O L 19/00	600
H 03 M 7/38	(2006.01)	H 03 M 7/38	
H 04 B 14/04	(2006.01)	H 04 B 14/04	Z

請求項の数 20 (全 20 頁)

(21) 出願番号	特願2000-238898 (P2000-238898)
(22) 出願日	平成12年8月2日 (2000.8.2)
(65) 公開番号	特開2002-49399 (P2002-49399A)
(43) 公開日	平成14年2月15日 (2002.2.15)
審査請求日	平成19年2月20日 (2007.2.20)

(73) 特許権者	000002185 ソニー株式会社 東京都港区港南1丁目7番1号
(74) 代理人	100082740 弁理士 田辺 恵基
(72) 発明者	近藤 哲二郎 東京都品川区北品川6丁目7番35号ソニ 一株式会社内
(72) 発明者	服部 正明 東京都品川区北品川6丁目7番35号ソニ 一株式会社内
(72) 発明者	渡辺 勉 東京都品川区北品川6丁目7番35号ソニ 一株式会社内

最終頁に続く

(54) 【発明の名称】デジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体

(57) 【特許請求の範囲】

【請求項 1】

デジタルオーディオ信号からパワースペクトルデータを算出する周波数分析ステップと、

上記パワースペクトルデータを最大値幅で正規化する正規化ステップと、

正規化されたパワースペクトルデータに基づいてそのクラスを分類するクラス分類ステップと、

分類されたクラスに対応した予測方式で上記デジタルオーディオ信号を演算することにより上記デジタルオーディオ信号を変換する予測演算ステップと
を有するデジタル信号処理方法。

10

【請求項 2】

上記周波数分析ステップで算出されるパワースペクトルデータの一部を抽出するスペクトルデータ抽出ステップ

をさらに有し、

上記正規化ステップは、

上記スペクトルデータ抽出ステップで抽出されるパワースペクトルデータを最大値幅で正規化する

請求項 1 に記載のデジタル信号処理方法。

【請求項 3】

上記スペクトルデータ抽出ステップでは、

20

直流成分のパワースペクトルデータが抽出対象として除外される
請求項 1に記載のディジタル信号処理方法。

【請求項 4】

上記予測演算ステップでは、

予め所望とするディジタルオーディオ信号に基づいて学習により生成されている予測係数が用いられる

請求項 1に記載のディジタル信号処理方法。

【請求項 5】

上記スペクトルデータ抽出ステップでは、

N 個のパワースペクトルデータのうち、ゼロから N / 2 個まで又は N / 2 個から N までのパワースペクトルデータが抽出対象とされる

請求項 1に記載のディジタル信号処理方法。

【請求項 6】

ディジタルオーディオ信号からパワースペクトルデータを算出する周波数分析手段と、
 上記パワースペクトルデータを最大値幅で正規化して正規化する正規化手段と、
正規化されたパワースペクトルデータに基づいてそのクラスを分類するクラス分類手段と、

分類されたクラスに対応した予測方式で上記ディジタルオーディオ信号を演算すること
 により上記ディジタルオーディオ信号を変換する予測演算手段と
 を具えるディジタル信号処理装置。

【請求項 7】

上記周波数分析ステップで算出されるパワースペクトルデータの一部を抽出するスペクトルデータ抽出手段

をさらに具え、

上記正規化手段は、

上記スペクトルデータ抽出手段で抽出されるパワースペクトルデータを最大値幅で正規化する

請求項 6に記載のディジタル信号処理装置。

【請求項 8】

上記スペクトルデータ抽出手段は、

直流成分のパワースペクトルデータを抽出対象として除く

請求項 6に記載のディジタル信号処理装置。

【請求項 9】

上記予測演算手段は、

予め所望とするディジタルオーディオ信号に基づいて学習により生成されている予測係数を用いる

請求項 6に記載のディジタル信号処理装置。

【請求項 10】

上記スペクトルデータ抽出手段は、

N 個のパワースペクトルデータのうち、ゼロから N / 2 個まで又は N / 2 個から N までのパワースペクトルデータを抽出対象とする

請求項 6に記載のディジタル信号処理装置。

【請求項 11】

コンピュータに対して、

ディジタルオーディオ信号からパワースペクトルデータを算出すること、

上記パワースペクトルデータを最大値幅で正規化すること、

正規化されたパワースペクトルデータに基づいてそのクラスを分類すること、

分類されたクラスに対応した予測方式で上記ディジタルオーディオ信号を演算すること
 により上記ディジタルオーディオ信号を変換すること

を実行させるプログラムが格納されるプログラム格納媒体。

10

20

30

40

50

【請求項 1 2】

ディジタルオーディオ信号を変換する変換処理の予測に用いられる予測係数を生成する学習方法において、

ディジタルオーディオ信号から当該ディジタルオーディオ信号を劣化させた生徒ディジタルオーディオ信号を生成する生徒ディジタルオーディオ信号生成ステップと、

上記生徒ディジタルオーディオ信号からパワースペクトルデータを算出する周波数分析ステップと、

上記パワースペクトルデータを最大値幅で正規化する正規化ステップと、

正規化されたパワースペクトルデータに基づいてそのクラスを分類するクラス分類ステップと、

上記ディジタルオーディオ信号と上記生徒ディジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップと

を有する学習方法。

【請求項 1 3】

上記周波数分析ステップで算出されるパワースペクトルデータの一部を抽出するスペクトルデータ抽出ステップ

をさらに有し、

上記正規化ステップは、

上記スペクトルデータ抽出ステップで抽出されるパワースペクトルデータを最大値幅で正規化する

請求項 1 2 に記載の学習方法。

【請求項 1 4】

上記スペクトルデータ抽出ステップでは、

直流成分のパワースペクトルデータが抽出対象として除外される

請求項 1 2 に記載の学習方法。

【請求項 1 5】

上記スペクトルデータ抽出ステップでは、

N 個のパワースペクトルデータのうち、ゼロから N / 2 個まで又は N / 2 個から N までのパワースペクトルデータが抽出対象とされる

請求項 1 2 に記載の学習方法。

【請求項 1 6】

ディジタルオーディオ信号を変換する変換処理の予測演算に用いられる予測係数を生成する学習装置において、

所望とするディジタルオーディオ信号から当該ディジタルオーディオ信号を劣化させた生徒ディジタルオーディオ信号を生成する生徒ディジタルオーディオ信号生成手段と、

上記生徒ディジタルオーディオ信号からパワースペクトルデータを算出する周波数分析手段と、

上記パワースペクトルデータを最大値幅で正規化する正規化手段と、

正規化されたパワースペクトルデータに基づいてそのクラスを分類するクラス分類手段と、

上記ディジタルオーディオ信号と上記生徒ディジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出手段と

を具える学習装置。

【請求項 1 7】

上記周波数分析ステップで算出されるパワースペクトルデータの一部を抽出するスペクトルデータ抽出手段

をさらに具え、

上記正規化手段は、

上記スペクトルデータ抽出手段で抽出されるパワースペクトルデータを最大値幅で正規化する

10

50

請求項 1 6 に記載の学習装置。

【請求項 1 8】

上記スペクトルデータ抽出手段は、

直流成分のパワースペクトルデータを抽出対象として除く

請求項 1 6 に記載の学習装置。

【請求項 1 9】

上記スペクトルデータ抽出手段は、

N 個のパワースペクトルデータのうち、ゼロから N / 2 個まで又は N / 2 個から N までのパワースペクトルデータを抽出対象とする

請求項 1 6 に記載の学習装置。

10

【請求項 2 0】

コンピュータに対して、

ディジタルオーディオ信号から当該ディジタルオーディオ信号を劣化させた生徒ディジタルオーディオ信号を生成すること、

上記生徒ディジタルオーディオ信号からパワースペクトルデータを算出すること、

上記パワースペクトルデータを最大値幅で正規化すること、

正規化されたパワースペクトルデータに基づいてそのクラスを分類すること、

上記ディジタルオーディオ信号と上記生徒ディジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出すること

を実行させるプログラムが格納されるプログラム格納媒体。

20

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はディジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に関し、レートコンバータ又はPCM(Pulse Code Modulation)復号装置等においてディジタル信号に対してデータの補間処理を行うディジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に適用して好適なものである。

【0002】

【従来の技術】

従来、ディジタルオーディオ信号をディジタル／アナログコンバータに入力する前に、サンプリング周波数を元の値の数倍に変換するオーバサンプリング処理を行っている。これにより、ディジタル／アナログコンバータから出力されたディジタルオーディオ信号はアナログ・アンチ・エイリアス・フィルタの位相特性が可聴周波数高域で一定に保たれ、また、サンプリングに伴うディジタル系のイメージ雑音の影響が排除されるようになされている。

30

【0003】

かかるオーバサンプリング処理では、通常、線形一次(直線)補間方式のディジタルフィルタが用いられている。このようなディジタルフィルタは、サンプリングレートが変わったりデータが欠落した場合等に、複数の既存データの平均値を求めて直線的な補間データを生成するものである。

40

【0004】

【発明が解決しようとする課題】

ところが、オーバサンプリング処理後のディジタルオーディオ信号は、線形一次補間によって時間軸方向に対してデータ量が数倍に緻密になっているものの、オーバサンプリング処理後のディジタルオーディオ信号の周波数帯域は変換前とあまり変わらず、音質そのものは向上していない。さらに、補間されたデータは必ずしもA/D変換前のアナログオーディオ信号の波形に基づいて生成されたのではないため、波形再現性もほとんど向上していない。

【0005】

また、サンプリング周波数の異なるディジタルオーディオ信号をダビングする場合におい

50

て、サンプリング・レート・コンバータを用いて周波数を変換しているが、かかる場合でも線形一次ディジタルフィルタによって直線的なデータの補間しか行うことができず、音質や波形再現性を向上することが困難であった。さらに、ディジタルオーディオ信号のデータサンプルが欠落した場合において同様である。

【0006】

本発明は以上の点を考慮してなされたもので、ディジタルオーディオ信号の波形再現性を一段と向上し得るディジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体を提案しようとするものである。

【0007】

【課題を解決するための手段】

10

かかる課題を解決するため本発明においては、ディジタルオーディオ信号からパワースペクトルデータを算出し、算出されたパワースペクトルデータを最大値幅で正規化して正規化データを算出し、算出された正規化データに基づいてそのクラスを分類し、分類されたクラスに対応した予測方式でディジタルオーディオ信号を変換するようにしたことにより、一段とディジタルオーディオ信号の特徴に適応した変換を行うことができる。

【0008】

【発明の実施の形態】

以下図面について、本発明の一実施の形態を詳述する。

【0009】

図1においてオーディオ信号処理装置10は、ディジタルオーディオ信号（以下これをオーディオデータと呼ぶ）のサンプリングレートを上げたり、オーディオデータを補間する際に、真値に近いオーディオデータをクラス分類適用処理によって生成するようになされている。

20

【0010】

因みに、この実施の形態におけるオーディオデータとは、人間の声や楽器の音等を表す楽音データ、さらにはその他種々の音を表すデータである。

【0011】

すなわち、オーディオ信号処理装置10において、スペクトル処理部11は入力端子T_{IN}から供給された入力オーディオデータD10を所定時間毎の領域（この実施の形態の場合、例えば6サンプル毎とする）に切り出した時間軸波形データであるクラスタップを構築した後、当該構築したクラスタップについて、後述する対数データ算出方法により、入力手段18から供給される制御データD18に応じて対数データを算出する。

30

【0012】

スペクトル処理部11は入力オーディオデータD10のこのとき構築されたクラスタップについて、対数データ算出方法による算出結果であってクラス分類しようとする対数データD11を算出し、これをクラス分類部14に供給する。

【0013】

クラス分類部13は、スペクトル処理部11から供給された対数データD11について、当該対数データD11を圧縮して圧縮データパターンを生成するADR(C(Adaptive Dynamic Range Coding)回路部と、対数データD11の属するクラスコードを発生するクラスコード発生回路部とを有する。

40

【0014】

ADR(C)回路部は対数データD11に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADR(C)回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

【0015】

具体的には、6つの8ビットのデータ（対数データ）をクラス分類しようとする場合、2⁴⁸という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、

50

この実施の形態のクラス分類部 14 ではその内部に設けられた A D R C 回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば 6 つの対数データに対して 1 ビットの量子化を実行すると、6 つの対数データを 6 ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

【0016】

ここで、A D R C 回路部は、切り出された領域内のダイナミックレンジを D R 、ビット割り当てを m、各対数データのデータレベルを L、量子化コードを Q とすると、次式、

【0017】

【数 1】

10

$$D R = M A X - M I N + 1$$

$$Q = \{ (L - M I N + 0.5) \times 2^m / D R \} \quad \dots \dots (1)$$

【0018】

に従って、領域内の最大値 M A X と最小値 M I N との間を指定されたビット長で均等に分割して量子化を行う。なお、(1) 式において { } は小数点以下の切り捨て処理を意味する。かくしてスペクトル処理部 11 において算出された 6 つの対数データが、それぞれ例えば 8 ビット ($m = 8$) で構成されているとすると、これらは A D R C 回路部においてそれぞれが 2 ビットに圧縮される。

20

【0019】

このようにして圧縮された対数データをそれぞれ q_n ($n = 1 \sim 6$) とすると、クラス分類部 14 に設けられたクラスコード発生回路部は、圧縮された対数データ q_n に基づいて、次式、

【0020】

【数 2】

$$class = \sum_{i=1}^n q_i (2^P)^i \quad \dots \dots (2)$$

30

【0021】

に示す演算を実行することにより、そのブロック ($q_1 \sim q_6$) が属するクラスを示すクラスコード class を算出し、当該算出されたクラスコード class を表すクラスコードデータ D 14 を予測係数メモリ 15 に供給する。このクラスコード class は、予測係数メモリ 15 から予測係数を読み出す際の読み出しアドレスを示す。因みに (2) 式において、n は圧縮された対数データ q_n の数を表し、この実施の形態の場合 n = 6 であり、また P はビット割り当てを表し、この実施の形態の場合 P = 2 である。

【0022】

このようにして、クラス分類部 14 は入力オーディオデータ D 10 から算出された対数データ D 11 のクラスコードデータ D 14 を生成し、これを予測係数メモリ 15 に供給する。

40

【0023】

予測係数メモリ 15 には、各クラスコードに対応する予測係数のセットがクラスコードに対応するアドレスにそれぞれ記憶されており、クラス分類部 14 から供給されるクラスコードデータ D 14 に基づいて、当該クラスコードに対応するアドレスに記憶されている予測係数のセット $W_1 \sim W_n$ が読み出され、予測演算部 16 に供給される。

【0024】

予測演算部 16 は、予測演算部抽出部 13 において入力オーディオデータ D 10 から時間軸領域で切り出された予測演算しようとするオーディオ波形データ (予測タップ) D 13

50

($X_1 \sim X_n$) と、予測係数 $w_1 \sim w_n$ に対して、次式

【0025】

【数3】

$$y' = w_1 X_1 + w_2 X_2 + \dots + w_n X_n \quad \dots \quad (3)$$

【0026】

に示す積和演算を行うことにより、予測結果 y' を得る。この予測値 y' が、音質が改善されたオーディオデータ D16 として予測演算部 16 から出力される。

10

【0027】

なお、オーディオ信号処理装置 10 の構成として図 1 について上述した機能ブロックを示したが、この機能ブロックを構成する具体的構成として、この実施の形態においては図 2 に示すコンピュータ構成の装置を用いる。すなわち、図 2 において、オーディオ信号処理装置 10 は、バス B U S を介して C P U 21、R O M (Read Only Memory) 22、予測係数メモリ 15 を構成する R A M (Random Access Memory) 15、及び各回路部がそれぞれ接続された構成を有し、C P U 11 は R O M 22 に格納されている種々のプログラムを実行することにより、図 1 について上述した各機能ブロック（スペクトル処理部 11、予測演算部抽出部 13、クラス分類部 14 及び予測演算部 16）として動作するようになされている。

20

【0028】

また、オーディオ信号処理装置 10 にはネットワークとの間で通信を行う通信インターフェース 24、フロッピディスクや光磁気ディスク等の外部記憶媒体から情報を読み出すリムーバブルドライブ 28 を有し、ネットワーク経由又は外部記憶媒体から図 1 について上述したクラス分類適用処理を行うための各プログラムをハードディスク装置 25 のハードディスクに読み込み、当該読み込まれたプログラムに従ってクラス分類適応処理を行うこともできる。

【0029】

ユーザは、キーボードやマウス等の入力手段 18 を介して種々のコマンドを入力することにより、C P U 21 に対して図 1 について上述したクラス分類処理を実行させる。この場合、オーディオ信号処理装置 10 はデータ入出力部 27 を介して音質を向上させようとするオーディオデータ（入力オーディオデータ）D10 を入力し、当該入力オーディオデータ D10 に対してクラス分類適用処理を施した後、音質が向上したオーディオデータ D16 をデータ入出力部 27 を介して外部に出力し得るようになされている。

30

【0030】

因みに、図 3 はオーディオ信号処理装置 10 におけるクラス分類適応処理の処理手順を示し、オーディオ信号処理装置 10 はステップ S P 101 から当該処理手順に入ると、続くステップ S P 102 において入力オーディオデータ D10 の対数データ D11 をスペクトル処理部 11 において算出する。

【0031】

この算出された対数データ D11 は入力オーディオデータ D10 の特徴を表すものであり、オーディオ信号処理装置 10 は、ステップ S P 103 に移ってクラス分類部 14 により対数データ D11 に基づいてクラスを分類する。そしてオーディオ信号処理装置 10 はクラス分類の結果得られたクラスコードを用いて予測係数メモリ 15 から予測係数を読み出す。この予測係数は予め学習によりクラス毎に対応して格納されており、オーディオ信号処理装置 10 はクラスコードに対応した予測係数を読み出すことにより、このときの対数データ D11 の特徴に合致した予測係数を用いることができる。

40

【0032】

予測係数メモリ 15 から読み出された予測係数は、ステップ S P 104 において予測演算部 16 の予測演算に用いられる。これにより、入力オーディオデータ D10 はその対数デ

50

ータ D 1 1 の特徴に適応した予測演算により、所望とするオーディオデータ D 1 6 に変換される。かくして入力オーディオデータ D 1 0 はその音質が改善されたオーディオデータ D 1 6 に変換され、オーディオ信号処理装置 1 0 はステップ S P 1 0 5 に移って当該処理手順を終了する。

【 0 0 3 3 】

次に、オーディオ信号処理装置 1 0 のスペクトル処理部 1 1 における入力オーディオデータ D 1 0 の対数データ D 1 1 の算出方法について説明する。

【 0 0 3 4 】

すなわち、図 4 はスペクトル処理部 1 1 における対数データ算出方法の対数データ算出処理手順を示し、スペクトル処理部 1 1 はステップ S P 1 から当該処理手順に入ると、続くステップ S P 2 において入力オーディオデータ D 1 0 を所定時間毎の領域に切り出した時間軸波形データであるクラスタップを構築し、ステップ S P 3 に移る。10

【 0 0 3 5 】

ステップ S P 3 において、スペクトル処理部 1 1 はクラスタップに対して、窓関数を「 W (k) 」とすると、次式、

【 0 0 3 6 】

【 数 4 】

$$W [k] = 0.54 + 0.46 * \cos(\pi * k / N) \quad \langle k = 0, \dots, N - 1 \rangle$$

20

..... (4)

【 0 0 3 7 】

に示すハミング窓に従って、乗算データを算出し、ステップ S P 4 に移る。因みに、この窓関数の乗算処理においては、続くステップ S P 4 において行われる周波数分析の精度を向上させるために、このとき構築されたそれぞれのクラスタップの最初の値と最後の値を等しくするようになされている。また、(1) 式において、「 N 」はハミング窓のサンプル数を表しており、「 k 」は何番目のサンプルデータであるかを表している。30

【 0 0 3 8 】

ステップ S P 4 において、スペクトル処理部 1 1 は乗算データに対して、高速フーリエ変換 (F F T : Fast Fourier Transform) を行うことにより、図 5 に示すようなパワースペクトルデータを算出し、ステップ S P 5 に移る。

【 0 0 3 9 】

ステップ S P 5 において、スペクトル処理部 1 1 はパワースペクトルデータから有意であるパワースペクトルデータのみを抽出するようになされている。

【 0 0 4 0 】

この抽出処理において、N 個の乗算データから算出したパワースペクトルデータのうち、N / 2 から右側のパワースペクトルデータ群 A R 2 (図 5) は、ゼロ値から N / 2 までの左側のパワースペクトルデータ群 A R 1 (図 5) とほぼ同じ成分になる (すなわち、左右対称となる) 。このことは、N 個の乗算データの周波数帯域内で、両端から等距離にある 2 個の周波数点におけるパワースペクトルデータの成分が互いに共役であることを示している。従って、スペクトル処理部 1 1 は、ゼロ値から N / 2 までの左側のパワースペクトルデータ群 A R 1 (図 5) のみを抽出対象とする。40

【 0 0 4 1 】

そしてスペクトル処理部 1 1 は、このとき抽出対象としたパワースペクトルデータ群 A R 1 のうち、予めユーザが入力手段 1 8 (図 1 及び図 2) を介して選択設定した以外の m 個のパワースペクトルデータを除いて抽出する。

【 0 0 4 2 】

50

具体的には、ユーザが入力手段 18 を介して例えば人間の声を一段と高音質にするように選択設定を行った場合、当該選択操作に応じた制御データ D18 が入力手段 18 からスペクトル処理部 11 に出力され(図 1 及び図 2)、これによりスペクトル処理部 11 は、このとき抽出したパワースペクトルデータ群 A R 1 (図 5) から、人間の声において有意となる 500 Hz から 4 kHz 付近のパワースペクトルデータのみを抽出する(すなわち 500 Hz から 4 kHz 付近以外のパワースペクトルデータが、除くべき m 個のパワースペクトルデータである)。

【0043】

また、ユーザが入力手段 18 (図 1 及び図 2) を介して例えば音楽を一段と高音質にするように選択を行った場合には、当該選択操作に応じた制御データ D18 が入力手段 18 からスペクトル処理部 11 に出力され、これによりスペクトル処理部 11 は、このとき抽出したパワースペクトルデータ群 A R 1 (図 5) から、音楽において有意となる 20 Hz から 20 kHz 付近のパワースペクトルデータのみを抽出する(すなわち 20 Hz から 20 kHz 付近以外のパワースペクトルデータが、除くべき m 個のパワースペクトルデータである)。

10

【0044】

このように入力手段 18 (図 1 及び図 2) から出力される制御データ D18 は、有意なパワースペクトルデータとして抽出する周波数成分を決定づけるようになされており、入力手段 18 (図 1 及び図 2) を介して手動で選択操作するユーザの意図を反映している。

20

【0045】

従って、制御データ D18 に応じてパワースペクトルデータを抽出するスペクトル処理部 11 は、ユーザが高音質での出力を希望する特定のオーディオ成分の周波数成分を有意なパワースペクトルデータとして抽出することとなる。

【0046】

因みに、スペクトル処理部 11 は、抽出対象としたパワースペクトルデータ群 A R 1 のうち、もとの波形の音程を表すため、有意な特徴をもたない直流成分のパワースペクトルデータをも除いて抽出するようになされている。

【0047】

このように、ステップ S P 5 において、スペクトル処理部 11 は制御データ D18 に応じて、パワースペクトルデータ群 A R 1 (図 5) から m 個のパワースペクトルデータを除くと共に、直流成分のパワースペクトルデータも除いてなる必要最小限のパワースペクトルデータ、すなわち有意なパワースペクトルデータのみを抽出し、続くステップ S P 6 に移る。

30

【0048】

ステップ S P 6 において、スペクトル処理部 11 は抽出されたパワースペクトルデータに対して、次式、

【0049】

【数 5】

$$p_s _max = max(p_s [k])$$

..... (5)

40

【0050】

に従って、このとき抽出されたパワースペクトルデータ (p_s [k]) の最大値 (p_s _max) を算出し、次式、

【0051】

【数 6】

$$psn[k] = ps[k] / ps_max \quad \dots\dots (6)$$

【0052】

に従って、このとき抽出されたパワースペクトルデータ($ps[k]$)の最大値(ps_max)での正規化(除算)し、このとき得られた基準値($psn[k]$)に対して、次式、

【0053】

【数7】

10

$$psl[k] = 10.0^* \log(ps[n[k]) \quad \dots\dots (7)$$

【0054】

に従って、対数(デシベル値)変換を行うようになされている。

【0055】

因みに(7)式において、 \log は常用対数である。また対数変換においては、任意の基準値によって、小さな波形をもデシベル値(音圧レベル)として表し得る。従って、例えば大きな波形付近に有意である小さな波形が存在するオーディオデータをスペクトル処理部11が対数変換しなかった場合、当該オーディオデータは一般的に16ビット等の大きなビット数で量子化されることにより、有意である小さな波形部分が大きな波形にマスキングされてしまう。

20

【0056】

このため、スペクトル処理部11は、特徴部分(有意である小さな波形部分)を見い出せないことになる。従って、スペクトル処理部11は、対数変換を行うことにより、特徴部分(有意である小さな波形部分)をも見い出すようになされている。

【0057】

また、音感等の刺激に対する人間の感覚は、ほぼその強さの対数に比例するため、対数変換にて表した量(すなわち、デシベル値)は、感覚の度合いを表すことになる。従って、スペクトル処理部11は、対数変換を行うことにより、結果として、音声を聞く対象である人間が心地よく聞き得るようにする。

30

【0058】

このように、ステップSP6において、スペクトル処理部11は最大振幅で正規化及び振幅の対数変換を行うことにより、特徴部分(有意である小さな波形部分)をも見い出すと共に、結果として、音声を聞く対象である人間が心地よく聞き得るようにする対数データD11を算出し、続くステップSP7に移って対数データ算出処理手順を終了する。

【0059】

このようにして、スペクトル処理部11は対数データ算出方法の対数データ算出処理手順によって、入力オーディオデータD10で表される信号波形の特徴を一段と見い出した対数データD11を算出することができる。

40

【0060】

次に、図1について上述した予測係数メモリ15に記憶するクラス毎の予測係数のセットを予め学習によって得るために学習回路について説明する。

【0061】

図6において、学習回路30は、高音質の教師オーディオデータD30を生徒信号生成フィルタ37に受ける。生徒信号生成フィルタ37は、間引き率設定信号D39により設定された間引き率で教師オーディオデータD30を所定時間ごとに所定サンプル間引くようになされている。

【0062】

この場合、生徒信号生成フィルタ37における間引き率によって、生成される予測係数が

50

異なり、これに応じて上述のオーディオ信号処理装置10で再現されるオーディオデータも異なる。例えば、上述のオーディオ信号処理装置10においてサンプリング周波数を高くすることでオーディオデータの音質を向上しようとする場合、生徒信号生成フィルタ37ではサンプリング周波数を減らす間引き処理を行う。また、これに対して上述のオーディオ信号処理装置10において入力オーディオデータD10の欠落したデータサンプルを補うことで音質の向上を図る場合には、これに応じて、生徒信号生成フィルタ37ではデータサンプルを欠落させる間引き処理を行うようになされている。

【0063】

かくして、生徒信号生成フィルタ37は教師オーディオデータ30から所定の間引き処理により生徒オーディオデータD37を生成し、これをスペクトル処理部31及び予測演算部抽出部33にそれぞれ供給する。10

【0064】

スペクトル処理部31は生徒信号生成フィルタ37から供給された生徒オーディオデータD37を所定時間毎の領域（この実施の形態の場合、例えば6サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、図4について上述した対数データ算出方法による算出結果であってクラス分類しようとする対数データD31を算出し、これをクラス分類部34に供給する。

【0065】

クラス分類部34は、スペクトル処理部31から供給された対数データD31について、当該対数データD31を圧縮して圧縮データパターンを生成するADR回路部と、対数データD31の属するクラスコードを発生するクラスコード発生回路部とを有する。20

【0066】

ADR回路部は対数データD31に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADR回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

【0067】

具体的には、6つの8ビットのデータ（対数データ）をクラス分類しようとする場合、 2^{48} という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部34ではその内部に設けられたADR回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つの対数データに対して1ビットの量子化を実行すると、6つの対数データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。30

【0068】

ここで、ADR回路部は、切り出された領域内のダイナミックレンジをDR、ビット割り当てをm、各対数データのデータレベルをL、量子化コードをQとして、上述の（1）式と同様の演算により、領域内の最大値MAXと最小値MINとの間を指定されたビット長で均等に分割して量子化を行う。かくしてスペクトル処理部31において算出された6つの対数データが、それぞれ例えば8ビット（m=8）で構成されているとすると、これらはADR回路部においてそれが2ビットに圧縮される。40

【0069】

このようにして圧縮された対数データをそれぞれ q_n （n=1～6）とすると、クラス分類部34に設けられたクラスコード発生回路部は、圧縮された対数データ q_n に基づいて、上述の（2）式と同様の演算を実行することにより、そのブロック（ $q_1 \sim q_6$ ）が属するクラスを示すクラスコードclassを算出し、当該算出されたクラスコードclassを表すクラスコードデータD34を予測係数算出部36に供給する。因みに（2）式において、nは圧縮された対数データ q_n の数を表し、この実施の形態の場合n=6であり、またPはビット割り当てを表し、この実施の形態の場合P=2である。

【0070】

10

20

30

40

50

このようにして、クラス分類部34はスペクトル処理部31から供給された対数データD31のクラスコードデータD34を生成し、これを予測係数算出部36に供給する。また、予測係数算出部36には、クラスコードデータD34に対応した時間軸領域のオーディオ波形データD33(x_1, x_2, \dots, x_n)が予測演算部抽出部33において切り出されて供給される。

【0071】

予測係数算出部36は、クラス分類部34から供給されたクラスコードclassと、各クラスコードclass毎に切り出されたオーディオ波形データD33と、入力端T_{IN}から供給された高音質の教師オーディオデータD30とを用いて、正規方程式を立てる。

【0072】

すなわち、生徒オーディオデータD37のnサンプルのレベルをそれぞれ x_1, x_2, \dots, x_n として、それぞれにpビットのADRCを行った結果の量子化データを q_1, \dots, q_n とする。このとき、この領域のクラスコードclassを上述の(2)式のように定義する。そして、上述のように生徒オーディオデータD37のレベルをそれぞれ、 x_1, x_2, \dots, x_n とし、高音質の教師オーディオデータD30のレベルをyとしたとき、クラスコード毎に、予測係数 w_1, w_2, \dots, w_n によるnタップの線形推定式を設定する。これを次式、

【0073】

【数8】

$$y = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad \dots \quad (8)$$

【0074】

とする。学習前は、 w_n が未定係数である。

【0075】

学習回路30では、クラスコード毎に、複数のオーディオデータに対して学習を行う。データサンプル数がMの場合、上述の(8)式に従って、次式、

【0076】

【数9】

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn} \quad \dots \quad (9)$$

【0077】

が設定される。但し $k = 1, 2, \dots, M$ である。

【0078】

$M > n$ の場合、予測係数 w_1, \dots, w_n は一意的に決まらないので、誤差ベクトルeの要素を次式、

【0079】

【数10】

$$e_k = y_k - \{w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn}\} \quad \dots \quad (10)$$

【0080】

によって定義し(但し、 $k = 1, 2, \dots, M$)、次式、

【0081】

【数11】

10

20

30

40

$$e^2 = \sum_{k=0}^M e_k^2 \quad \dots\dots (11)$$

【0082】

を最小にする予測係数を求める。いわゆる、最小自乗法による解法である。

【0083】

ここで、(11)式による w_n の偏微分係数を求める。この場合、次式、

【0084】

【数12】

$$\frac{\partial e^2}{\partial w_i} = \sum_{k=0}^M 2 \left(\frac{\partial e_k}{\partial w_i} \right) e_k = \sum_{k=0}^M 2 x_{ki} \cdot e_k$$

$$= \sum_{k=0}^M 2 x_{ki} \cdot e_k \quad (i = 1, 2 \dots n) \quad \dots\dots (12)$$

10

20

【0085】

を「0」にするように、各 w_n ($n = 1 \sim 6$) を求めれば良い。

【0086】

そして、次式、

【0087】

【数13】

$$X_{ij} = \sum_{p=0}^M X_{pi} \cdot X_{pj} \quad \dots\dots (13)$$

30

【0088】

【数14】

$$Y_i = \sum_{k=0}^M X_{ki} \cdot y_k \quad \dots\dots (14)$$

【0089】

のように、 X_{ij} 、 Y_i を定義すると、(12)式は行列を用いて次式、

【0090】

【数15】

40

$$\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \dots \dots (15)$$

【0091】

10

として表される。

【0092】

この方程式は、一般に正規方程式と呼ばれている。なお、ここでは $n = 6$ である。

【0093】

全ての学習用データ（教師オーディオデータ D30、クラスコード class、オーディオ波形データ D33）の入力が完了した後、予測係数算出部 36 は各クラスコード class に上述の(15)式に示した正規方程式を立てて、この正規方程式を掃き出し法等の一般的な行列解法を用いて、各 w_n について解き、各クラスコード毎に、予測係数を算出する。予測係数算出部 36 は、算出された各予測係数 (D36) を予測係数メモリ 15 に書き込む。

20

【0094】

このような学習を行った結果、予測係数メモリ 15 には、量子化データ q_1, \dots, q_6 で規定されるパターン毎に、高音質のオーディオデータ y を推定するための予測係数が、各クラスコード毎に格納される。この予測係数メモリ 15 は、図 1 について上述したオーディオ信号処理装置 10 において用いられる。かかる処理により、線形推定式に従って通常のオーディオデータから高音質のオーディオデータを作成するための予測係数の学習が終了する。

【0095】

30

このように、学習回路 30 は、オーディオ信号処理装置 10 において補間処理を行う程度を考慮して、生徒信号生成フィルタ 37 で高音質の教師オーディオデータの間引き処理を行うことにより、オーディオ信号処理装置 10 における補間処理のための予測係数を生成することができる。

【0096】

以上の構成において、オーディオ信号処理装置 10 は、入力オーディオデータ D10 に対して高速フーリエ変換を行うことにより、周波数軸上にパワースペクトルを算出する。周波数分析（高速フーリエ変換）は、時間軸波形データからでは知りえない微妙な違いを見出すことが可能であることにより、オーディオ信号処理装置 10 は、時間軸領域に特徴を見い出せない微妙な特徴を見い出しえるようになる。

【0097】

40

微妙な特徴を見い出しえる状態（すなわち、パワースペクトルを算出した状態）において、オーディオ信号処理装置 10 は、選択範囲設定手段（ユーザが入力手段 18 から手動で行う選択設定）に応じて、有意とされるパワースペクトルデータのみを抽出（すなわち、 $N/2 - m$ 個）する。

【0098】

これによりオーディオ信号処理装置 10 は、処理負担を一段と軽減することができ、かつ処理速度を向上させることができる。

【0099】

さらに、オーディオ信号処理装置 10 は、有意とされた必要最小限のパワースペクトルデータに対して、最大振幅で正規化及び振幅の対数変換を行うことにより、対数データを生成する。この対数変換においては、特徴部分（有意である小さな波形部分）を見い出す

50

と共に、結果として、音声を聞く対象である人間が心地よく聞き得るようにする対数データを生成する。

【0100】

このように、オーディオ信号処理装置10は、周波数分析を行うことにより、微妙な特徴を見い出し得るようになされたパワースペクトルデータから有意とされるパワースペクトルデータのみを抽出し、さらに、抽出したパワースペクトルデータに対して、最大振幅で正規化及び振幅の対数変換を行うことにより得られる対数データに基づいて、そのクラスを特定する。

【0101】

そしてオーディオ信号処理装置10は、抽出した有意なパワースペクトルデータに基づいて特定したクラスに基づく予測係数を用いて入力オーディオデータD10を予測演算することにより、当該入力オーディオデータD10を一段と高音質のオーディオデータD16に変換することができる。 10

【0102】

また、クラス毎の予測係数を生成する学習時において、位相の異なる多数の教師オーディオデータについてそれぞれに対応した予測係数を求めておくことにより、オーディオ信号処理装置10における入力オーディオデータD10のクラス分類適応処理時に位相変動が生じても、位相変動に対応した処理を行うことができる。

【0103】

以上の構成によれば、周波数分析を行うことにより、微妙な特徴を見い出し得るようになされたパワースペクトルデータから有意とされるパワースペクトルデータのみを抽出し、さらに、抽出したパワースペクトルデータに対して最大振幅で正規化及び振幅の対数変換を行うことにより得た対数データをクラス分類した結果に基づく予測係数を用いて入力オーディオデータD10を予測演算するようにしたことにより、入力オーディオデータD10を一段と高音質のオーディオデータD16に変換することができる。 20

【0104】

なお上述の実施の形態においては、窓関数としてハミング窓を用いて乗算する場合について述べたが、本発明はこれに限らず、ハミング窓に代えて、例えばハニング窓やブラックマン窓等、他の種々の窓関数によって乗算する、又はスペクトル処理部において予め各種窓関数（ハミング窓、ハニング窓及びブラックマン窓等）を用いて乗算し得るようにしておき、入力されるデジタルオーディオ信号の周波数特性に応じて、スペクトル処理部が所望の窓関数を用いて乗算するようにしても良い。 30

【0105】

因みに、スペクトル処理部がハミング窓を用いて乗算する場合、スペクトル処理部は、切り出し部から供給されたクラスタップに対して、次式、

【0106】

【数16】

$$W[k] = 0.50 + 0.50 * \cos(\pi * k / N) \quad \langle k = 0, \dots, N-1 \rangle$$

40

..... (16)

【0107】

からなるハミング窓を乗算して乗算データを算出する。

【0108】

また、スペクトル処理部がブラックマン窓を使用して乗算する場合、スペクトル処理部は、切り出し部から供給されたクラスタップに対して、次式、

【0109】

【数17】

50

$$W(k) = 0.42 + 0.50^* \cos(\pi^* k/N) + 0.08^* \cos(2\pi^* k/N)$$

$\langle k = 0, \dots, N-1 \rangle$

..... (17)

【0110】

10

からなるブラックマン窓を乗算して乗算データを算出する。

【0111】

また上述の実施の形態においては、高速フーリエ変換を用いる場合について述べたが、本発明はこれに限らず、例えば離散フーリエ変換（DFT：Discrete Fourier Transformer）や離散コサイン変換（DCT：Discrete Cosine Transform）や最大エントロピー法、さらには線形予測分析による方法等、他の種々の周波数分析手段を適用することができる。

【0112】

さらに上述の実施の形態においては、スペクトル処理部11がゼロ値からN/2までの左側のパワースペクトルデータ群AR1（図5）のみを抽出対象とする場合について述べたが、本発明はこれに限らず、右側のパワースペクトルデータ群AR2（図5）のみを抽出対象とするようにしても良い。

20

【0113】

この場合、オーディオ信号処理装置10の処理負担を一段と軽減することができ、処理速度を一段と向上させることができる。

【0114】

さらに上述の実施の形態においては、圧縮データパターンを生成するパターン生成手段として、ADRCを行う場合について述べたが、本発明はこれに限らず、例えば可逆符号化（DPCM：Differential Pulse Code Modulation）やベクトル量子化（VQ：Vector Quantize）等の圧縮手段を用いるようにしても良い。要は、信号波形のパターンを少ないクラスで表現し得るような圧縮手段であれば良い。

30

【0115】

さらに上述の実施の形態においては、ユーザが手動で選択操作し得る選択範囲設定手段として、人間の声及び音声を選択（すなわち、抽出する周波数成分として500Hz～4kHz又は20Hz～20kHz）する場合について述べたが、本発明はこれに限らず、例えば図7に示すように、高域（UPP）、中域（MID）及び低域（LOW）のいづれかの周波数成分を選択する、又は図8に示すように、まばらに周波数成分を選択する、さらには図9に示すように、不均一に帯域を周波数成分する等、他の種々の選択範囲設定手段を適用し得る。

【0116】

40

この場合、オーディオ信号処理装置には、新たに設けられた選択範囲設定手段に対応するプログラムを作成してハードディスクドライブやROM等、所定の記憶手段に格納させる。これにより、ユーザが手動で入力手段18を介して新たに設けられた選択範囲設定手段を選択操作した場合においても、このとき選択された選択範囲設定手段に応じた制御データが入力手段からスペクトル処理部に出力され、これによりスペクトル処理部は、新たに設けられた選択範囲設定手段に対応するプログラムによって、所望の周波数成分からパワースペクトルデータの抽出を行う。

【0117】

このようにすれば、他の種々の選択範囲設定手段を適用することができ、ユーザの意図に応じた有意なパワースペクトルデータを抽出することができる。

50

【0118】

さらに上述の実施の形態においては、オーディオ信号処理装置10(図2)がプログラムによってクラスコード生成処理手順を実行する場合について述べたが、本発明はこれに限らず、ハードウェア構成によってこれらの機能を実現して種々のデジタル信号処理装置(例えば、レートコンバータ、オーバーサンプリング処理装置、B S(Broadcasting Sate llite)放送等に用いられているP C M(Pulse Code Modulation)デジタル音声エラー訂正を行うP C Mエラー修正装置等)内に設けたり、又は各機能を実現するプログラムを格納したプログラム格納媒体(フロッピーディスク、光ディスク等)からこれらのプログラムを種々のデジタル信号処理装置にロードして各機能部を実現するようにしても良い。

【0119】

10

【発明の効果】

上述のように本発明によれば、デジタルオーディオ信号からパワースペクトルデータを算出し、算出されたパワースペクトルデータを最大値幅で正規化して正規化データを算出し、算出された正規化データに基づいてそのクラスを分類し、分類されたクラスに対応した予測方式でデジタルオーディオ信号を変換するようにしたことにより、一段とデジタルオーディオ信号の特徴に適応した変換を行うことができ、かくして、デジタルオーディオ信号の波形再現性を一段と向上した高音質のデジタルオーディオ信号への変換を行うことができる。

【図面の簡単な説明】

【図1】本発明によるオーディオ信号処理装置を示す機能ブロック図である。

20

【図2】本発明によるオーディオ信号処理装置を示すブロック図である。

【図3】オーディオデータの変換処理手順を示すフローチャートである。

【図4】対数データ算出処理手順を示すフローチャートである。

【図5】パワースペクトルデータ算出例を示す略線図である。

【図6】学習回路の構成を示すブロック図である。

【図7】パワースペクトルデータ選択例を示す略線図である。

【図8】パワースペクトルデータ選択例を示す略線図である。

【図9】パワースペクトルデータ選択例を示す略線図である。

【符号の説明】

10 オーディオ信号処理装置、 11 スペクトル処理部、 22 R O M、 15 ...
... R A M、 24 通信インターフェース、 25 ハードディスクドライブ、 26
入力手段、 27 データ入出力部、 28 リムーバブルドライブ。

30

【図1】

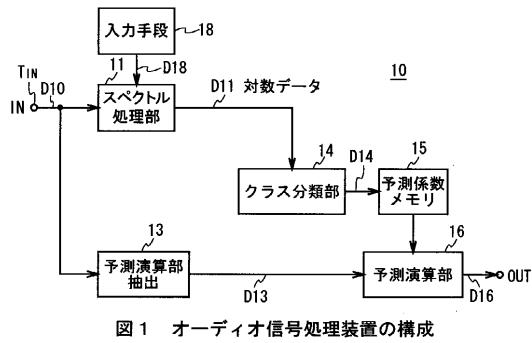


図1 オーディオ信号処理装置の構成

【図2】

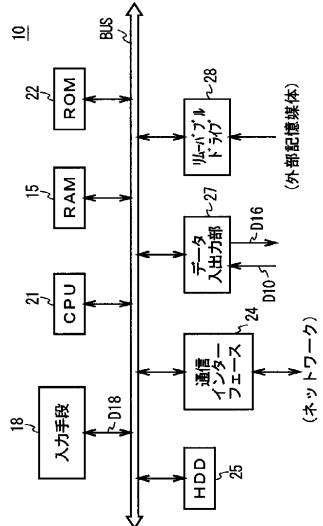


図2 オーディオ信号処理装置

【図3】

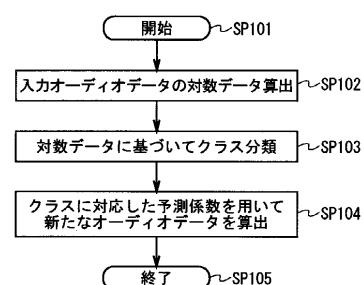


図3 オーディオデータの変換処理手順

【図4】

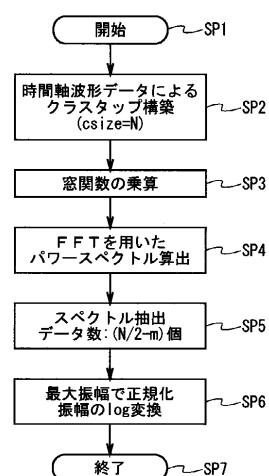


図4 対数データ算出処理手順

【図5】

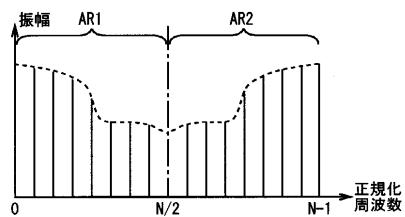


図5 パワースペクトルデータ算出例

【図7】

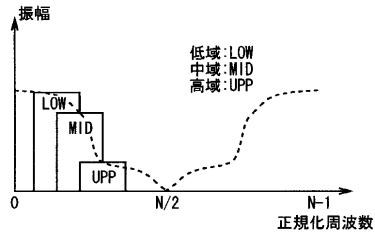


図7 パワースペクトルデータ選択例

【図6】

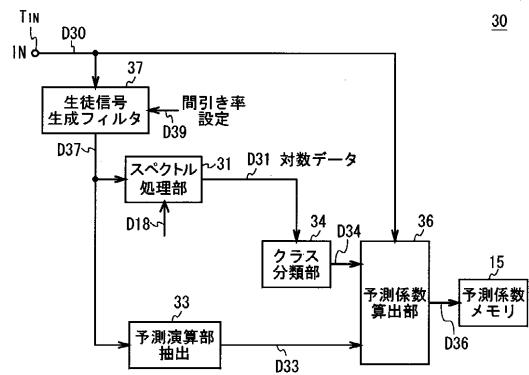


図6 学習回路の構成

【図8】

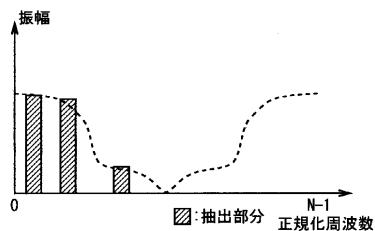


図8 パワースペクトルデータ選択例

【図9】

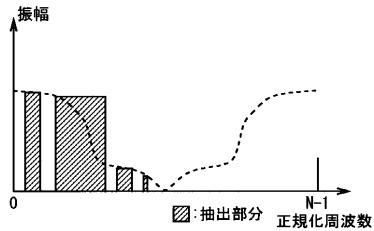


図9 パワースペクトルデータ選択例

フロントページの続き

(72)発明者 木村 裕人
東京都品川区北品川6丁目7番35号ソニー株式会社内

審査官 鈴木 聰一郎

(56)参考文献 特開平10-313251(JP,A)
特開平07-147566(JP,A)
特開平01-204538(JP,A)
特開平08-265711(JP,A)
特開平08-275120(JP,A)
特開平09-167240(JP,A)
特開平07-321662(JP,A)
特開平11-003094(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 11/00- 21/06
G10H 1/00- 7/12
H03M 7/38
H04B 14/04