(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3782103号
(P3782103)
(45)発行日　平成18年6月7日(2006.6.7)　(24)登録日　平成18年3月17日(2006.3.17)

(51)Int.Cl.　F I
G10L 19/02 (2006.01)　G10L 19/02　142Z

請求項の数 10　(全 12 頁)

(21)出願番号	特願平7-517872
(86)(22)出願日	平成6年12月13日(1994.12.13)
(65)公表番号	特表平8-507391
(43)公表日	平成8年8月6日(1996.8.6)
(86)国際出願番号	PCT/IB1994/000418
(87)国際公開番号	WO1995/018523
(87)国際公開日	平成7年7月6日(1995.7.6)
審査請求日	平成13年12月10日(2001.12.10)
審判番号	不服2004-17045(P2004-17045/J1)
審判請求日	平成16年8月16日(2004.8.16)
(31)優先権主張番号	93203652.8
(32)優先日	平成5年12月23日(1993.12.23)
(33)優先権主張国	欧州特許庁(EP)

(73)特許権者　590000248
コーニンクレッカ　フィリップス　エレクトロニクス　エヌ　ヴィ
Ｋｏｎｉｎｋｌｉｊｋｅ　Ｐｈｉｌｉｐｓ　Ｅｌｅｃｔｒｏｎｉｃｓ　Ｎ．Ｖ．
オランダ国　５６２１　ベーアー　アインドーフェン　フルーネヴァウツウェッハ　１
Ｇｒｏｅｎｅｗｏｕｄｓｅｗｅｇ　１，５６２１　ＢＡ　Ｅｉｎｄｈｏｖｅｎ，Ｔｈｅ　Ｎｅｔｈｅｒｌａｎｄｓ
(74)代理人　100072051
弁理士　杉村　興作

最終頁に続く

(54)【発明の名称】アダプティブディザを減算し、埋没チャンネルビットを挿入し、フィルタリングすることによりマルチビット符号ディジタル音声を符号化する方法及び装置、及びこの方法のための符号化及び

(57)【特許請求の範囲】
【請求項１】
マルチビット入力サンプルのストリームとして受信される音声表現を符号化する方法において、
前記入力サンプルの有限長系列を構成し、該系列内で瞬時不均一可聴しきい値対周波数特性を決定するステップと、
前記入力サンプルから整形雑音信号を減算して処理サンプルを発生させるステップと、
埋没チャネルビットとしてのディザ信号（ｖ）を減算し次いで前記特性に関連するビットレベルより下位の可変数のｂビットにより量子化することにより前記ビットレベルより下位の処理サンプルビットをダイナミックに無視するが、前記ビットレベルより上位の少なくとも全ての処理サンプルビットは保持するステップと、
このように無視した処理サンプルビットを前記ディザ信号と置き換えることにより、１処理サンプルにつき調整数のｂビットづつ埋没チャンネルビットを埋め込むステップと、
無視されてない処理サンプルビットと埋没チャンネルビットを具えるアップグレードサンプル（ｙ）を出力するステップとを具え、
前記整形雑音信号は、前記アップグレードサンプルと前記処理サンプルとの間のディザ決定差に、ノイズシェイピングフィルタにより整形フィルタリングを施すことにより発生させ、この整形フィルタリングは受信差信号を前記瞬時しきい値対周波数特性に従って修正するものであることを特徴とする符号化方法。
【請求項２】

前記ノイズシェイピングフィルタは白色雑音特性に近い差信号を受信することを特徴とする請求項１記載の方法。
【請求項３】
前記整形雑音信号の発生は可変特性を有するノイズシェイピングフィルタにより行うことを特徴とする請求項１又は２記載の方法。
【請求項４】
個々の時間インターバルにおいて音楽遷移部を検出し、この検出時に当該インターバルにおいてｂの値を音楽遷移部のない隣接時間インターバルから少なくとも共通に決定される値にセットすることを特徴とする請求項１、２、又は３記載の方法。
【請求項５】
前記置き換えステップ前に埋没チャンネルデータとして使用するデータをバッファ内で一時的にバッファし、時間変化レート制御により前記バッファの充填度を時間的に等しく維持することを特徴とする請求項１～４のいずれかに記載の方法。
【請求項６】
前記埋没チャンネルデータをＭＰＥＧオーディオチャンネルとして受信することを特徴とする請求項１～５のいずれかに記載の方法。
【請求項７】
前記整形フィルタリングは、それぞれ独自の格子周波数θk^cを有するとともに総合スペクトルの局部的パワースペクトル密度を近似する幅Δｋを有する一連の基本フィルタ曲線から合成された総合フィルタ曲線を有するフィルタで行うことを特徴とする請求項１～６記載のいずれかに記載の方法。
【請求項８】
前記格子周波数は不均一間隔を有することを特徴とする請求項１～７のいずれかに記載の方法。
【請求項９】
前記音声表現の周波数スペクトル内に共存する各別の周波数帯域ごとに符号化することを特徴とする請求項１～８のいずれかに記載の方法。
【請求項１０】
請求項１～９のいずれかに記載された方法を実行する符号化装置。
【発明の詳細な説明】
発明の技術分野
本発明はマルチビットサンプルのストリームとして受信される音声を埋没チャンネルビットを挿入しながら符号化する方法に関するものである。このような方法はM.A.Gerzon及びP.G.Graven,"A High-Rate Buried Channel for Audio CD",preprint3551 of the Audio Eng. Soc. Conv.,Berlin,March 1993,に記載されている。同一著者の国際特許出願ＷＯ９４／０３９８８，優先日１９９４年８月５日，公開日１９９４年２月１７日，も参照されたい。第１の引用文献は減算式ディザ処理した雑音整形量子化により埋没チャンネルを付加する比較的簡単な方法に基づくものである。付加された特徴は埋没チャンネルにより高品質の伝送音声を向上しうるが、本願発明者は、この文献は”減算原理”の潜在能力を十分に利用しておらず、埋没チャンネルの最適伝送容量より低い容量、或いは本来高品質の音声、例えばＣＤ音声の予定の残留知覚品質より低い品質をもたらすことを発見した。
発明の概要
従って、本発明の目的は、特に、埋没チャンネルの特性を向上させ、残留知覚品質及び埋没音声チャンネルの伝送容量を両方とも最適にすることにある。
本発明の特徴の一つは、マルチビット入力サンプルのストリームとして受信される音声表現を符号化する方法において、
前記入力サンプルの有限長系列を構成し、該系列内で瞬時不均一可聴しきい値対周波数特性を決定するステップと、
前記入力サンプルから整形雑音信号を減算して処理サンプルを発生させるステップと、
ディザ信号（ｖ）を減算し次いで前記特性に関連するビットレベルより下位の可変数のｂ

ビットにより量子化することにより前記ビットレベルより下位の処理サンプルビットをダイナミックに無視するが、前記ビットレベルより上位の少なくとも全ての処理サンプルビットは保持するステップと、
このように無視した処理サンプルビットを前記ディザ信号と、１処理サンプルあたり可変数のｂビットを特徴とする埋没チャンネルビットとして置き換えるステップと、
無視されてない処理サンプルビットと埋没チャンネルビットを具えるアップグレードサンプル（ｙ）を出力するステップとを具え、
前記整形雑音信号は、前記アップグレードサンプルと前記処理サンプルとの間のディザ決定差に、ノイズシェイピングフィルタにより整形フィルタリングを施すことにより発生させ、この整形フィルタリングは受信差信号を前記瞬時しきい値対周波数特性に従って修正するものである点にある。
特に、前記差を瞬時しきい値対周波数特性に従って整形フィルタリングすると、人間の聴覚系が比較的鈍感な周波数に”スペクトル空間”を生成することができる。その結果、かなり大きな長さのサンプルに対し１サンプルにつき均等数の埋没チャンネルビットを注入することができる。次の系列に対し、埋没チャンネルビットの数を再び決定する必要がある。本発明の他の特徴は、メインチャンネルの特性の整形に続いて埋没チャンネルの特性も同様に整形して転送容量を最適にする。特に、ディザ信号は整形する必要がある。これはシステムをある程度複雑にする。
ノイズシェイピングフィルタには白色雑音に近い差信号を受信させるのが有利である。多くの場合、この差信号に対する要件はメインチャンネルの品質に与えるその影響については高くないため、ディザは任意の内容を有することができる。他の状態においては、前記差はメインチャンネルの情報と相関しなくてももよく、また自己相関しなくてもよい。後者の状態では、ディザを前処理して相関の適切な欠落を達成する。このようなランダム化手段自体は公知である。
前記整形雑音信号の発生は可変特性を有する量子化フィルタにより行うのが有利である。このようにすると、前記瞬時聴度への適合の改善及び主観的音質の改善が達成される。
前記整形フィルタリングは、それぞれ独自の格子周波数θ_k^cを有するとともに総合スペクトルの局部的パワースペクトル密度を近似する幅Δ_kを有する一連の基本フィルタ曲線から合成された総合フィルタ曲線を有するフィルタで行うのが有利である。これは、パワースペクトル密度をモデル化する極めて直接的な方法であり、前記格子周波数を不均等間隔のものにすることができる。一般に、これは精度対使用格子周波数の数を改善し、従って計算を高速化する。
本発明は上述した符号化方法を実行する符号化装置及びこの方法を実行することにより得られた信号を復号する復号化装置にも関するものである。種々の他の特徴は請求の範囲の従属項に記載されている。

【図面の簡単な説明】

本発明のこれらの特徴及び他の特徴及び利点を特に添付図面を参照して好適実施例につき以下に詳細に説明する。
図１は本発明装置の全体ブロック図であり、
図２は本発明により改善された埋没チャンネルエンコーダの基礎として使用される減算型ディザ量子化を示し、
図３は典型的な音声スペクトルによりマスクされたしきい値の周波数依存性を示し、
図４は埋没チャンネル形成の第１の典型的な例を示し、
図５は埋没チャンネル形成の第２の典型的な例を示し、
図６はピッチ予測のない簡略ＣＥＬＰエンコーダを示し、
図７はノイズシェピング量子化器を示し、
図８は基本フィルタ曲線を示す。

原理の簡単な説明

埋没チャンネル技術は、オーディオ又は音声信号が所定の長さのサンプルビット列により、知覚音声品質の点からすると実際上高すぎる精度でしばしば表現される事実を利用する

。即ち、情報の量を減少させて伝送容量を追加の情報サービスのために開放することができる。追加の情報をメイン信号の下位ビット部分に挿入する。慣例の受信機に対しては、メイン信号のこの変更は、人間聴取者がその差を認識し得なければ問題ない。しかし、高品位受信システムは追加の情報を取り出し、この情報を別の出力端子に発生する。本発明では、受信信号を実際のしきい値対周波数特性に従って修正するために差スペクトルを整形フィルタ処理する。これにより人間の聴覚系があまり感じない周波数に”スペクトル空間”を生成しうる。前記差が白色雑音状である場合には、聴取者は付加チャンネルを殆ど知覚できない。

好適実施例の詳細な説明

図１は本発明装置の全体ブロック図である。ブロック２０は、例えば各サンプルが１６ビットからなる約４４ｋＨｚで反復するサンプルストリームとして符号化されたディジタル音声の音声源である。この音声は予め規定された帯域幅を有し、音楽、会話、その他の任意の内容を有することができる。ブロック２２はこれらのサンプルの有限長系列を構成し、これらの系列が所定の時間インターバル、例えば１０２４サンプル＝０．０２秒を占めるようにするとともに、ここで可聴しきい値対周波数特性を決定する。聴度は入音声信号の限定数のキー特性に基づいて決定することができる。聴度は瞬時周波数に依存し、一つの周波数帯域が他の周波数帯域によりマスキング作用を受ける時には、音声の全体的大きさ又は局部的大きさに依存する。また、聴度は聴取者ごとに相違しうるがこの相違は一般に無視される。可聴しきい値は既知の方法で決定することができ、その結果を後に示す。更に、可聴しきい値の決定後に、サンプルをその下位のｂビットを無視することにより量子化する。ブロック２４は埋没チャンネルビット源である。埋没チャンネルの性質は、メインチャンネル音声への追加のコメント、例えば表示可能なサブタイトル又はテキスト、マルチチャンネル音声再生における同一又はそれぞれ異なる品質レベルの追加の音声チャンネル、多国語サービス、又はカラオケ又はビデオのような任意のものとすることができる。何の関係もないサービスも考えられる。しかし、特に有利な使用法は埋没チャンネルをＭＰＥＧオーディオチャンネルとして規定するものである。この標準規格は高品質のオーディオ伝送を中ビットレートで提供するのに有用であることが証明されている。更に、埋没チャンネル自体は機能的に無関係であるが相まってアップグレードサンプルの埋没部分を構成する２以上のサブチャンネルで構成することができる。ブロック２６において、音声源２０からの無視されたサンプルビット又はその一部分を、その下位レベルから出発して、埋没チャンネル源２４からのビットと置き換える。更に、アップグレードサンプルのストリームの相互離間位置において、１サンプル当たりの無視ビット数に関する指示を埋没チャンネルに挿入し、次の指示が発生するとき適用しうる。効率の良い動作のために、これらの指示の間隔は最適値にセットする必要がある。この間隔が小さすぎると、オーバヘッドが増大する。この間隔が大きすぎると、無視ビットの数が個々のサンプルから見て少なくなりすぎる。伝送用又は記憶用のチャンネル２８はアップグレードサンプルを受信機３０に供給する。受信機３０は、受信したこれらの指示に基づいて、埋没チャンネルビットからサンプルの標準部分を分離することができる。サンプルの標準部分は音声を標準方法で再生するデコーダ３２に供給され、置換ビットがサブ可聴雑音を表すものとして維持される。埋没チャンネルビットはこれを正しく処理するようプログラムされた補助デコーダに供給される。他の構成例では、埋没チャンネルデコーダがチャンネル受信機のディジタル出力端子に発生する全サンプルを受信するとともにこれから標準チャンネルを無視して埋没チャンネル情報を抽出するようにする。他方、ノーマル非アップグレード受信機がチャンネルに対し設けられている場合には、この受信機がアップグレードサンプルを、これらのサンプルがノーマルサンプルであるかのように処理する。この受信機はある種のスピーカを附勢するアナログオーディオ増幅器で終わる。この場合にはメインチャンネルビットに相関関係を示さない埋没チャンネルビットは、予定の可聴しきい値より低く維持されているある種の雑音を直接表す。

図２は、前記文献に記載され、本発明の特徴により改善された埋没チャンネルエンコーダ用の減算式ディザ量子化器を示す。ｂビットのディザ信号ｖが後述する方法で入力端子５

２に入力される。素子５４、５８、６０、６２は適切なビット幅のディジタル加算器－減算器である。素子５６は１６ビット受信信号ｘの精度を打ち切りにより１６ビットより少数ビットに減少する量子化器である。このような特徴自体は、S.P.Lipshiz et al,"Quantization and Dither: A theoretical Survey", J. Audio Eng. Soc. Vol.40, no.5, pp.355-375, May, 1992,に記載されている。ｂビットディザ信号ｖの使用により、メインチャンネルとの相関関係の欠落が十分であれば、量子化誤差ｅがスペクトル的に白色雑音に維持され、統計的に入力信号ｘとほぼ無関係になり、知覚上好適である。ディザ信号は埋没チャンネル信号をランダム処理したものとすることができ、情報を付加したり、犠牲にしたりしない。このようなランダム化は再同期化の必要なしにもとにもどすことができる。ランダム特性はメインチャンネルに対し、埋没チャンネル自体内で指定される。埋没チャンネル自体に関しその信号が良好に構成されている場合には、この信号を慣例の方法でランダム化することができる。同一のディザ信号ｖを素子５４において加算して出力端子６６に蓄積用、伝送用、又は他の処理のためのコンパチブル出力信号ｙを形成する。ブロック６４はノイズシェイピングフィルタであり、コンパチブル出力信号ｙとディザ信号ｖの導入前の入力信号との差を減算器６２から受信する。ノイズシェイピングフィルタ６４の出力信号は原オーディオ信号ｘも受信する減算器６０に帰還される。雑音の大きさを９次ＦＩＲ（有限インパルス応答）フィルタのみで約１６ｄＢ減少させることができることが確かめられた。このアプローチは匹敵する大きさの埋没チャンネル信号の２－３ビット／サンプルの雑音レベルをＣＤ信号の固有の雑音フロアにする。フィルタ特性Ｈ（Ｚ）は、

$$y = x + |1 - H(Z)|^2 \cdot \Delta^2 / 12$$

で定義されるｙをｘに対し、聴取者に主観的に重要でない量だけ変化させるものとする必要がある。$\Delta = 2^b$は単位ステップサイズである。この場合、埋没チャンネルの伝送レートは再量子化処理の粗さに依存する。この点については、前記のGerzon等の引用文献の第２頁第４０行～第１３頁の第５章に、得られる誤差が高レベルメイン音声信号によりマスクされるときは再量子化をもっと粗くすることがでる旨記載されている。他方、本願発明者は、埋没チャンネルの伝送レートを増大するもっと有効な方法、即ち人間の聴覚系の周波数依存感度を用いる方法を発見した。図２の構成の他の特徴は埋没チャンネル６８により供給されるデータを一時的に記憶しうるバッファ５０にある。出力端子６６の埋没チャンネルの転送速度は不均一であるため、このバッファはその充填度をほぼ一定に維持するある種の帰還構成を有するものとすることができる。バッファが空になりすぎる場合には、埋没チャンネルの全容量を部分的に放棄することができる。バッファが満杯になりすぎる場合には、種々のストラテジィがあり、一つはチャンネル６８からの送り速度を低減する。２以上の埋没サブチャンネルが存在する場合のもっと強烈なストラテジィは最も重要でないサブチャンネルを放棄する。サブチャンネルが動画ビデオを表す場合には、これを一時的に静止画の系列に減少させることができる。図２の構成の種々の変更が自明であり、例えば、所定の加算器を減算器に変更することができる。この場合には借りの伝搬を必要とするためハードウエアが僅かに複雑になる。しかし、特に所定のサンプル表記システムに対してはインパクトが最小になる。

この点に関し、図３に第１の典型的な音声スペクトルによりマスクされたしきい値の周波数依存特性を示す。図４はゼロ音声入力に基づく第１の典型的な埋没チャンネルの形成を示す。

図４において、標準化された、又は推定された人間の聴覚系において、曲線４は単一周波数に基づく可聴レベルを示す。計算上の理由から、曲線４はシミュレートされているが（その直線部分に注意）、これは自然現象に密接に追従する。約４０００Ｈｚでこのしきい値は最低であり、これにより高い周波数又は低い周波数で著しく高くなる。曲線１はフラットなＣＤ雑音フロアを示し、このレベルは$10^{10}\log(1/12\times 22050)$ｄＢで与えられる。曲線４は単一周波数に対する可聴しきい値を与えるが、雑音に対しては可聴作用はもっと高く、その周波数特性は曲線４より遙に低くなる。曲線２は曲線４のしきい値とほぼ同一の周波数依存性になるように整形した曲線１のフラット雑音のスペクトルである。正確な近似からの数ｄＢの偏差も、約１５ｋＨｚ以上の平坦形状も、総合特性に

悪影響を及ぼさないことが実験的に確かめられた。曲線３は曲線２に一致するが、ｂ×６ｄＢの距離だけ上方に相対的にシフトされ、図４ではｂ＝２である。このことは埋没チャンネルが２ビット／サンプルの幅であることを意味する。曲線３と曲線４との距離は付加された情報を聞こえないままに維持する。曲線１と４との最小距離は$10^{10}log(660)$であり、約４ｋＨｚの臨界帯域幅に対応する。設計は、曲線３のスペクトルの積分パワーを示す補助曲線を用い、この曲線が曲線４に接触しないように行うのが好ましい。図を簡単にするためこの補助曲線は図示してない。

図３は実際の音声スペクトルが可聴しきい値の形状に及ぼす影響を示す。曲線Ａは実際の音声スペクトルである。曲線Ｂは単音に対する可聴しきい値を示す。図４の曲線４は山－谷特性が消去されていること明らかである。

図５は第２の典型的な埋没チャンネルの形成を示す。本例では、図３の曲線Ａと異なる形状のシミュレート音声スペクトルによって、４０ｄＢの近似レベルに位置するとともに１３ｋＨｚの位置に約１０ｄＢの浅いへこみを有するマスクされた音声しきい値を発生する。明瞭のために、音声自体のスペクトルは示してない。図３から明らかなように、各スペクトルピークはマスクされたしきい値を臨界周波数程度の周波数幅に亘って上昇し、一般にしきい値の滑らかな外観を生ずる。ここでも、１６ｋＨｚ以上の高周波数範囲の影響は無視されている。曲線１は図４のものと同一である。曲線２は曲線４のマスクされたしきい値とほぼ同一の形状になるように整形されたフラット雑音のスペクトルであり、本例では、曲線２の全体的な形状は図４のものより遙に平坦である。曲線３は曲線２に一致するが、ｂ×６ｄＢの距離だけ上方に相対的にシフトされ、図５ではｂ＝６である。これは埋没チャンネルが６ビット／サンプルの幅であることを意味する。更に、音声スペクトルの形状は多くの場合に重要である。図４の曲線２の形状を図５に適用すると、改善は図５の場合より小さくなる。これらの図のアプローチは全周波数スペクトルに基づいている。所定のオーディオシステムでは、スペクトルがサブバンドに分配され、それぞれのサブバンドが互いにほぼ無関係のオーディオ信号を搬送する。このような状態では、本発明の方法及び装置を任意のサブバンド又はサブバンドの組み合わせに別々に適用するすることができる。この場合、構成が幾分複雑になるが、埋没チャンネルのデータ伝送レートが更に上昇する。

マスクされた目標レベルからのノイズウェイティングフィルタ及びノイズシェイピングフィルタの計算

以下に、図２のフィルタ６４のための複合フィルタ曲線を計算する好適実施例を示す。この分野の関連する刊行物は、E.Ordentlich及びY.Shoham,"Low-delay code-excited linear-predictive coding of wideband speech at 32 kbps", Proc.ICASSP-91, pp.9-12, Toronto,1991である。これに示された技術はＭＰＥ，ＲＰＥ及びＣＥＬＰのような種々の分野の種々の用途に好適である。ここで、駆動系列（ＭＰＥ，ＲＰＥ）又は駆動ベクトル（ＣＥＬＰ＝codebook excited linear prediction）は加重平均二乗誤差基準に基づいて選択される。

このようなコーダでは、短い出力系列が多数の駆動系列又はベクトルから発生される。発生された出力系列は原入力系列と比較される。比較の基準は加重平均二乗誤差である。これは、入力と発生された出力との差をノイズウェイティングフィルタに通すことを意味する。次に、フィルタ処理された差系列のパワーを推定する。このパワーを加重平均二乗誤差という。最小の加重平均二乗誤差を発生する駆動系列を選択する。

図６はウェイティングフィルタがＣＥＬＰコーダにどのように使用されるかを示す。ＬＰＣ分析ボックスが入音声信号ｘ［ｉ］から予測係数$a_1, \ldots, a_p$、ウェイティングフィルタの係数及び利得を計算する。コードブックは多数の駆動ベクトルを含む。ベクトル長はＮである。選択中に全てのベクトルに利得を乗算するとともに分析フィルタに通す。この結果$\underline{x}$［ｉ］で示すＮ個のサンプルが生ずる。誤差系列はＮ個のサンプル$\underline{x}$［ｉ］からＮ個の入力サンプルｘ［ｉ］を減算することにより形成される。誤差系列はウェイティングフィルタに通す。加重誤差系列の短時間パワーである加重平均二乗誤差を計算する。選択ボックスは最低の加重平均二乗誤差を発生する符号ベクトルを選択する。利得、予測係

数及び選択されたベクトルのインデックスをデコーダに伝送する。
この点に関し、駆動ベクトルは駆動系列とみなすことができるので、駆動系列表現のみを使用する。
慣用されているウェイティングフィルタは音声信号の予測係数（ＬＰＣ係数）a_1，．．．，a_nに基づくものである。このフィルタの可能な例は、

$$W(z) = \frac{1+\sum_{k=1}^{p} a_k\gamma_1^k z^{-k}}{1+\sum_{k=1}^{p} a_k\gamma_2^k z^{-k}} \cdot \frac{1}{1+\sum_{k=1}^{2} p_k\delta^k z^{-k}} \qquad (1)$$

である。係数p_1及びp_2は、ＬＰＣ分析を系列の最初の３つの自己相関遅延に適用することにより得られる。係数δ、γ_1及びγ_2はホルマントの位置における加重の大きさを制御する係数である。これらの係数は良好な知覚性能が得られるように調整される。有利な値は、
δ＝０．７、γ_1＝０．９５、γ_2＝０．８
である。
他の類似の例も同様に有用である。適切なコードブックに対し、符号化誤差のパワースペクトル密度関数は、

$$\left|\frac{1}{W(\exp(j\theta))}\right|^2. \qquad (2)$$

に比例する。

ノイズシェイパ

ノイズシェイパの機能は量子化雑音にスペクトル整形を与えることにある。図７はノイズシェイパの基本図を示す。デコーディング後の量子化雑音のパワースペクトル密度関数は、

$$\hat{S}(\exp(j\theta)) = |1 + F(\exp(j\theta))|^2 \frac{1}{12}\Delta^2, \qquad (3)$$

により与えられる。ここで、Δは量子化ステップサイズである。音声の線形予測符号化に慣用されているフィルタＦ（ｚ）は、

$$F(z) = \sum_{k=1}^{p} a_k\gamma^k z^{-k}, \qquad (4)$$

である（ここで、γ＜１）。この場合、量子化雑音のパワースペクトル密度関数は、

$$S(\exp(j\theta)) = \frac{|1+F(\exp(j\theta))|^2}{|1+\sum_{k=1}^{p} a_k\exp(-j\theta)|^2}\frac{1}{12}\Delta^2 = \frac{|1+\sum_{k=1}^{p} a_k\gamma^k\exp(-j\theta)|^2}{|1+\sum_{k=1}^{p} a_k\exp(-j\theta)|^2}\frac{1}{12}\Delta^2, \qquad (5)$$

で与えられる。
ウェイティングフィルタ及びノイズシェイパのねらいは、符号化誤差を歪みが知覚的に最低の妨害になるようにスペクトル整形することにある。音楽符号化においては、幾つかの方法で種々の周波数帯域内の量子化雑音のマスキングレベルを推定する。このレベルはマスクされた目標レベルという。これらの方法の目的も知覚的に最小の妨害を生ずる歪みを得ることにある。しかし、これらの方法は（１）のウェイティングフィルタ又は（４）のノイズシェイパより良好な精神音響学的基礎を有し、従ってマスクされた雑音のパワースペクトル密度関数の一層良好な近似を生ずる。
以下においてウェイティングフィルタ技術又はノイズシェピング技術をマスクされた目標レベルの計算と組み合わせる。このようにして得られたウェイティングフィルタ又はノイズシェイパは、それらの伝達関数がマスクされた雑音のスペクトル形状に一層良好に対応するため、既知のものより良好である。

目標レベル及び周波数帯域

一組の目標レベルt_1，．．，t_mが、例えば図３、４、５につき述べた測定法により予め計算されているものとする。これらの目標レベルは種々の周波数帯域内のマスキングしき

い値における雑音パワーを示す。これらの周波数帯域は互いに隣接するとともに零からサンプリング周波数の１／２までの範囲をカバーする必要がある。このため、正規化周波数θ：

$-\pi \leqq \theta \leqq \pi$

を使用する。対応する音声正規化周波数ｆは、

$f = (\theta / 2\pi) f_s$

になり、ここでf_sはサンプリング周波数である。

帯域幅は任意に選択することができる。実際上、臨界的帯域又は等間隔帯域を使用する。マスクされた目標レベルt_kに対応する周波数帯域の下端、上端及び中心周波数をそれぞれθ_k^l、θ_k^u及びθ_k^cで示す。

パワースペクトル密度の再構成

なめらかなパワースペクトル密度関数（ｐｓｄ）：Ｓ（exp（ｊθ））は、ｐｓｄ－再構成関数S_k（θ）をｋ番周波数帯域と関連させることにより、マスクされた目標レベルから導くことができる。この場合、ｐｓｄは、

$$S(\exp(j\theta)) = \sum_{k=1}^{m} t_k S_k(\theta). \qquad (6)$$

となる。ｐｓｄ－再構成関数には幾つかの制約がある。第１に、ｐｓｄは非負でなければならない。これは、

$$S_k(\theta) \geq 0, \; k = 1 \ldots, m \qquad (7)$$

であることを意味する。第２に、パワー保存、即ち

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} S(\exp(j\theta)) d\theta = \sum_{k=1}^{m} t_k, \qquad (8)$$

が必要とされる場合には、

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} S_k(\theta) \, d\theta = 1, \; k = 1, \ldots, m. \qquad (9)$$

である。

ｐｓｄ－再構成関数：

$$S_k(\theta) = \frac{\pi}{\Delta_k}, \; \theta_k^l \leq |\theta| < \theta_k^u, \;\; 0, \; otherwise, \qquad (10)$$

を選択することにより、ｐｓｄの段階的近似が得られる。ここで、$\Delta_k = \theta_k^u - \theta_k^l$。段階的近似はなめらかなｐｓｄを生じない。このため、raised-cosine ｐｓｄ－再構成関数のほうが好適である。この関数は、

$$S_k(\theta) = \frac{\pi}{2\Delta_k}\left(1+\cos\left(\frac{\pi}{\Delta_k}|\theta| - \theta_k^c\right)\right), \; \theta_k^c - \Delta_k \leq |\theta| < \theta_k^c + \Delta_k, \; 0, \; otherwise \qquad (11)$$

により与えられる。

図８はこのような基本フィルタ曲線を示す。raised-cosine ｐｓｄ－再構成関数も制限を有する。例えば、周波数帯域が均等に離間していない場合には、この関数は平坦なｐｓｄを再構成することができない。これは異なる上側スロープ及び下側スロープを用いることにより改善することができる。ｐｓｄ－再構成関数の選択はマスクされた雑音の所望のスペクトル動作により決定する。

パワースペクトル密度の近似

再構成されたｐｓｄ　Ｓ(exp（ｊθ））は全極型スペクトルで次のように近似される。

$$\hat{S}(\exp(J\theta)) = \frac{\sigma^2}{\left|1 + \sum_{k=1}^{q} b_k \exp(-jk\theta)\right|^2}, \qquad (12)$$

ここで、ｑは全極型スペクトルの次数である。この近似は次式の伝達関数を有するウェイティングフィルタを生ずる。

$$B(z) = 1 + \sum_{k=1}^{l} b_k z^{-k}. \qquad (13)$$

このウェイティングフィルタは、（１）のフィルタと異なり、ＦＩＲフィルタである。以下において、$b_1, \ldots, b_n$を$t_1, \ldots, t_m$から、

$$Q(b_1,\ldots,b_q) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S(\exp(j\theta))\,|B(\exp(j\theta))|^2\, d\theta \qquad (14)$$

を$b_1, \ldots, b_q$の関数として最小にすることにより計算する。式（１４）において、Ｓ（exp（ｊθ））は式（６）に従う。要するに、スペクトルＳ（θ）を近似する必要がある（ここで、θは正規化周波数である）。ここで、関数ＢはＦ（式２３）の逆関数であり、Ｑは定数である。導関数：

$$\frac{\partial Q(b_1, \ldots, b_q)}{\partial b_n}, \; n = 1, \ldots, q,$$

を計算し、これらの導関数を零にセットして次の一組の方程式：

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{k=1}^{m} t_k S_k(\theta)\left[1+\sum_{l=1}^{q} b_l \exp(-jl\theta)\right]\exp(jn\theta)d\theta = 0, \; n=1,\ldots,q. \qquad (15)$$

又は

$$\sum_{l=1}^{q} b_l \sum_{k=1}^{m} t_k \frac{1}{2\pi}\int_{-\pi}^{\pi} S_k(\theta)\exp(j\theta(n-l))d\theta = -\sum_{k=1}^{m} t_k \frac{1}{2\pi}\int_{-\pi}^{\pi} S_k(\theta)\exp(jn\theta)d\theta, \; n=1,\ldots,q \qquad (16)$$

を導く。

$$g_{k,n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_k(\theta)\exp(jn\theta)d\theta, \; k=1,\ldots,m, \; n=1,\ldots,q \qquad (17)$$

及び

$$p_n = \sum_{k=1}^{m} t_k g_{k,n}. \qquad (18)$$

と定義する。$g_{k,n}$はｐｓｄ－再構成関数から予め計算し、ｍ×ｑマトリクスに記憶することができる。これらの結果を式（１６）に代入して、次の一組の方程式：

$$\sum_{l=1}^{q} p_{n-l} b_l = -p_n, \; n = 1,\ldots,q. \qquad (19)$$

を得る。これは対称な正限定のToeplitz系であり、線形予測符号化から既知のYule-Walker方程式に一致する。ｑ×ｑ行列Ｒを

$r_{ij} = P_{i-j}$, ｉ，ｊ＝１，．．，ｑ

と定義し、ｑ個のベクトルｒを

$r_i = p_i$, ｉ＝１，．．，ｑ

と定義する。その結果として、

Ｒｂ＝－ｒ (20)

になる。ここで、ｑ個のベクトル（ｂ）は係数$b_1, \ldots, b_q$を含む。方程式群（１９）又は（２０）は既知のレビンソン－ダービンアルゴリズムにより容易に解くことができる。

$g_{k,n}$の例

Ｓ（exp（ｊθ））の段階的近似に対し、$g_{k,n}$は、

$$g_{k,n} = \cos(n\theta_k^S)\frac{\sin(n\frac{\Delta_k}{2})}{n\frac{\Delta_k}{2}}. \qquad (21)$$

で与えられる。
S（exp（jθ））のraised-cosine近似に対し、$g_{k.n}$は、

$$g_{k,n} = \cos(n\theta_k^S)\frac{\sin(n\Delta_k)}{n\Delta_k}\frac{(\frac{\pi}{\Delta_k})^2}{(\frac{\pi}{\Delta_k})^2 - n^2}. \qquad (22)$$

で与えられる。
係数b_1，．．，b_qは図７に示すようなウェイティングフィルタW（z）に直接供給することができる。この場合、W（z）＝B（z）で、B（z）は式（１３）で定義される。ノイズシェイパの場合には、F（z）は、

$$1 - F(z) = \frac{1}{B(z)}, \qquad (23)$$

となり、従って

$$f(z) = \frac{\sum_{k=1}^{q} b_k z^{-k}}{1 + \sum_{k=1}^{q} b_k z^{-k}}. \qquad (24)$$

となる。

【図１】

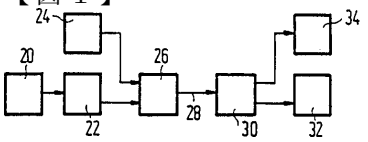

FIG.1

【図２】

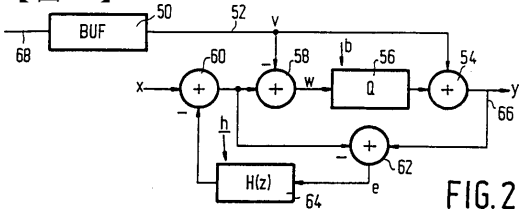

FIG.2

【図３】

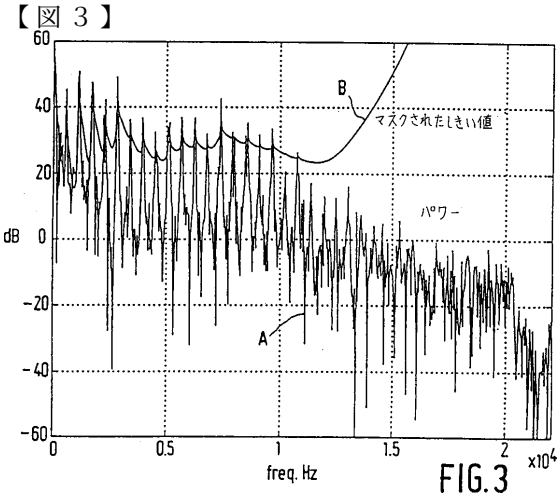

FIG.3

【図４】

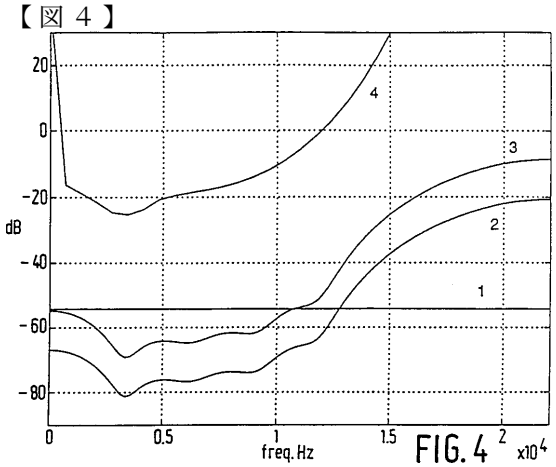

FIG.4

【図５】

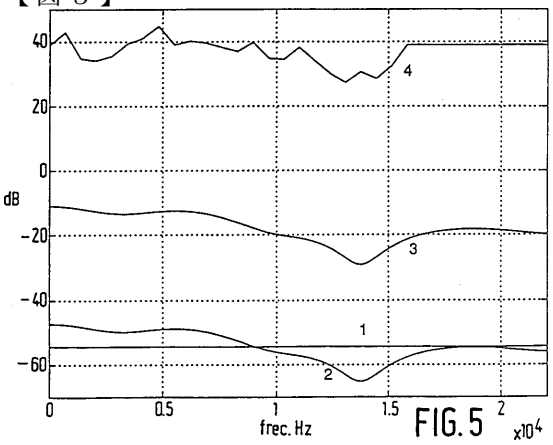

FIG.5

【図６】

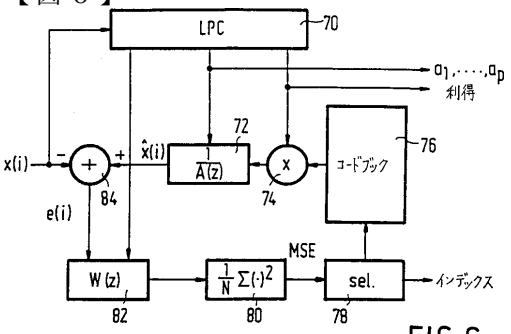

FIG.6

【図７】

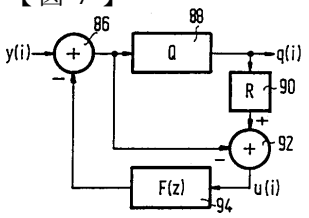

FIG.7

【図８】

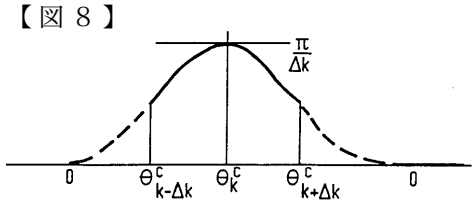

FIG.8

フロントページの続き

(72)発明者 オーメン アーノルダス ウェルナー ヨハネス
オランダ国 ５６２１ ベーアー アインドーフェン フルーネヴァウツウェッハ 1
(72)発明者 フロウネウェヘン マルク エドウィン
オランダ国 ５６２１ ベーアー アインドーフェン フルーネヴァウツウェッハ 1
(72)発明者 ファン デル ウァル ロベルト ヘルブランド
オランダ国 ５６２１ ベーアー アインドーフェン フルーネヴァウツウェッハ 1
(72)発明者 フェルドゥイス レイモンド ニコラス ヨハン
オランダ国 ５６２１ ベーアー アインドーフェン フルーネヴァウツウェッハ 1

合議体
審判長 藤内 光武
審判官 清水 正一
審判官 堀井 啓明

(56)参考文献 欧州特許出願公開第５４０３３０（EP, A2）
国際公開第９４／３９８８（WO, A2）
国際公開第９４／１８７６２（WO, A1）
特開平3－１６７９２７（JP, A）
特表平6－５０９２０１（JP, A）
M. A. Gerzon and P. G. Graven, "A High-Rate Buried-Data Channel for Audio CD", J. Audio Eng. Soc., Vol. 43, No 1/2, 1995 January/February, p. 3-22（1995/01-1995/02）
S. P. Lipshitz et al, "Quantization and Dither: A theoretical Survey", J. Audio Eng. Soc., Vol. 40, no. 5, 1992 May, p. 355-375（1992/05）

(58)調査した分野(Int.Cl., DB名)
G10L 19/00-19/14

(54)【発明の名称】アダプティブディザを減算し、埋没チャンネルビットを挿入し、フィルタリングすることによりマルチビット符号ディジタル音声を符号化する方法及び装置、及びこの方法のための符号化及び復号化装置。