

(19)日本国特許庁(JP)

(12)特許公報 (B 2)

(11)特許番号

特許第 3 3 8 4 5 2 3 号

(P 3 3 8 4 5 2 3)

(45)発行日 平成15年3月10日(2003.3.10)

(24)登録日 平成14年12月27日(2002.12.27)

(51)Int. Cl.⁷
G 1 0 L 13/00

識別記号

F I
G 1 0 L 7/00 D

請求項の数 3

(全 6 頁)

(21)出願番号 特願平8-233799

(22)出願日 平成8年9月4日(1996.9.4)

(65)公開番号 特開平10-78797

(43)公開日 平成10年3月24日(1998.3.24)

審査請求日 平成13年5月11日(2001.5.11)

(73)特許権者 000004226

日本電信電話株式会社
東京都千代田区大手町二丁目3番1号

(72)発明者 岩上 直樹

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 池田 和永

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 守谷 健弘

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74)代理人 100066153

弁理士 草野 卓

審査官 渡邊 聡

最終頁に続く

(54)【発明の名称】音響信号処理方法

1

(57)【特許請求の範囲】

【請求項 1】 フレーム単位に、音響信号の周波数特性の概形が取り除かれた周波数領域係数を得る第 1 の段階と、

上記周波数特性の概形を得る第 2 の段階と、
上記周波数特性の概形の各サンプルの値が大きいサンプルは更に値を大きくしてその周波数特性の概形の形状を強調する第 3 の段階と、

上記第 3 の段階で得られた強調された周波数特性の概形を用いて第 1 の段階で得られた周波数領域係数に周波数特性を付加して逆平坦化する第 4 の段階と、を有することを特徴とする音響信号処理方法。

【請求項 2】 フレーム単位に、音響信号の周波数特性の概形が取り除かれた周波数領域係数を得る第 1 の段階と、

2

上記周波数特性の概形を得る第 2 の段階と、
上記周波数特性の概形の各サンプルの値が小さいサンプルは更に値を小さくしてその周波数特性の概形の形状を強調する第 3 の段階と、

上記第 3 の段階で得られた強調された周波数特性の概形を用いて第 1 の段階で得られた周波数領域係数に周波数特性を付加して逆平坦化する第 4 の段階と、を有することを特徴とする音響信号処理方法。

【請求項 3】 上記第 2 の段階は、周波数特性の概形として、パークスケール上で等間隔の周波数分解能を持つスケールファクタを含むことを特徴とする請求項 1 又は 2 記載の音響信号処理方法。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】この発明はオーディオ信号、

特に符号化・復号化された音声信号に含まれる雑音感を低減させる信号処理方法に関する。

【0002】

【従来の技術】この発明は復号化音響信号以外の音響信号に適用できるが、復号化音響信号に適用の際の適用個所の理解のため従来の音響信号変換符号化・復号化法を図3を参照して説明する。符号化器10においては、入力端子11よりの入力音響信号を、時間-周波数変換部12によって周波数領域係数に変換する。この変換の方法としては、MDC T (Modified Discrete Cosine Transformation, 変形離散コサイン変換) や、D C T (Discrete Cosine Transformation, 離散コサイン変換), D F T (Discrete Fourier Transformation, 離散フーリエ変換) などを用いることができる。なお、時間-周波数変換部12では、前処理として、入力信号サンプル列のフレーム分割と窓掛けが必要である。フレーム分割は、MDC Tの場合入力サンプルがN点入力されるごとにこれを含む過去2N点のサンプルを1フレームとして分割する。D C TとD F Tの場合入力サンプルがN点ごとにこれを含む過去N+ 点のサンプルを1フレームとして分割する。窓掛けは従来行われている手法により行い、いずれの変換方法でも、N点の周波数領域係数が得られる。

【0003】概形計算部13で周波数領域係数の概形を抽出する。この概形抽出方法としては、前処理された音響信号を入力として線形予測分析をする方法、周波数領域係数を入力としてスケールファクタを計算する方法、周波数領域係数をリフタリングする方法などを用いることができる。線形予測分析をする方法では、入力信号を線形予測分析し、線形予測係数を求め、この係数のスペクトル振幅の逆数を周波数特性概形とする。線形予測の次数は、20次程度にするのが効果的である。

【0004】スケールファクタを計算する方法では、周波数領域係数を複数の小帯域に分割し、小帯域ごとにスケールファクタを計算し、これを周波数特性概形とする。小帯域に分割する方法は、周波数スケールで等間隔としてもよいし、パークスケールで等間隔(つまり聴覚的に等間隔)としてもよい。小帯域の数は30程度に設定するのが効果的である。スケールファクタは小帯域内のサンプルの振幅の平均値でもよいし、振幅の最大値でもよい。

【0005】周波数領域係数をリフタリングする方法では、周波数領域係数をケプストラム分析し、ケプストラム係数の低次部分のみのスペクトル振幅を周波数特性概形とする。また、周波数領域係数の概形は、上記の方法の併用により求めてもよい。たとえば、線形予測分析とスケールファクタを併用する場合、線形予測分析による線形予測スペクトルを決定した後、これに掛け合わせた際に実際の周波数特性にもっとも近い形状になるようにスケールファクタを決定するなどの方法をとる。

【0006】この周波数特性概形を概形量子化部14で量子化して、そのインデックス I_{n_1} を得る。周波数特性概形を線形予測分析により求めた場合、線形予測係数を線スペクトル対(L S P)に変換し、これを量子化する方法が能率がよい。スケールファクタを量子化する場合、各々のスケールファクタをスカラー量子化してもよいし、いくつかのスケールファクタをまとめてベクトル量子化してもよい。ベクトル量子化をする際、インタリーブベクトル量子化の技術を使うと、能率良く量子化が可能である。ケプストラム係数を量子化する場合、ケプストラム係数をスカラー量子化してもよいし、ベクトル量子化してもよい。

【0007】いずれの方法も、予測量子化を行うとさらに高い能率が得られる。予測の方法としては、AR予測、MA予測などを用いることができる。複数方法で周波数特性概形を求めた場合、用いたすべての方法について量子化を行う。量子化した周波数特性概形を概形再生部15で復号化し、周波数特性概形を再生する。線スペクトル対を量子化した場合、復号化して得られた再生線スペクトル対を再生線形予測係数に変換し、再生線形予測係数のスペクトル振幅の逆数を再生周波数特性概形とする。スケールファクタを量子化した場合、復号化した再生スケールファクタを再生周波数特性概形とする。ケプストラム係数を量子化した場合、復号化された再生ケプストラム係数のスペクトル振幅を再生周波数特性概形とする。

【0008】平坦化部16において周波数領域係数を再生周波数特性概形で平坦化する。ここでは、各々の周波数領域係数をこれに対応する周波数特性概形で割ることによって平坦化周波数領域係数(残差周波数係数)が得られる。この平坦化周波数係数を残差量子化部17でベクトル量子化してインデックス I_{n_2} を得る。この量子化方法として、重み付きベクトル量子化による変換符号化法(TC-WVQ, Transform Coding with Weighted Vector Quantization), 周波数領域重み付けインタリーブベクトル量子化(TWINVQ, Transform-domain Weighted Interleave Vector Quantization)などがある。それぞれの技術については、T. Moriya, H. Suda: "An, 8 kbit/s transform coder for noisy channels," Proc. ICASSP '89 pp 196 - 199 および岩上、守谷、三樹、"周波数領域重みづけインタリーブベクトル量子化(TwinVQ)によるオーディオ符号化、"日本音響学会講演論文集 平成6年10月~11月 pp. 339 - 340 に述べられている。

【0009】復号化器20において、量子化された平坦化された周波数領域係数のインデックス I_{n_2} を再生部21で復号再生する。量子化した周波数特性概形のインデックス I_{n_1} を再生部22で復号化し、再生周波数特性概形を再生する。線スペクトル対を量子化した場合、復号化して得られた再生線スペクトル対を再生線形予測係数

に変換し、再生線形予測係数のスペクトル振幅の逆数を再生周波数特性概形とする。スケールファクタを量子化した場合、復号化した再生スケールファクタを再生周波数特性概形とする。ケプストラム係数を量子化した場合、復号化された再生ケプストラム係数のスペクトル振幅を再生周波数特性概形とする。

【0010】なお、予測量子化を行った場合、同じ予測合成を用いて再生を行う。複数方法の量子化を行った場合、すべての方法について再生を行い、たとえば各々で再生した概形を互いに掛け合わせるなどの方法により再生周波数特性概形を得る。再生された平坦化周波数領域係数を、再生周波数特性概形を用いて逆平坦化部23で逆平坦化する。ここでは、各々の再生された平坦化周波数領域係数と、これに対応する再生周波数特性概形を掛け合わせることによって逆平坦化が行われ、再生周波数領域係数が得られる。

【0011】周波数 - 時間変換部24によって再生周波*

$$f(z) = (1 - \mu z^{-1}) \left\{ (1 + \sum_{i=1}^p \alpha_i \tau_1^i z^{-i}) / (1 + \sum_{i=1}^p \alpha_i \tau_2^i z^{-i}) \right\}$$

(1)

ここで μ はスペクトルの傾斜を補正する定数で例えば0.4、 τ_1 、 τ_2 はスペクトルの山を強調するための1以下の正定数で例えばそれぞれ0.5と0.8である。この手法は畳み込みの処理を必要とするため大きな演算量を必要とする。また詳細なスペクトル強調処理を行うためには線形予測の次数を高くする必要があり、演算量と演算精度の点からも問題がある。

【0014】

【発明が解決しようとする課題】この発明の目的は、オーディオ信号、特に符号化・復号化された音声信号に含まれる雑音感を低減させる信号処理方法を小さな演算量で詳細に実現することを目的とする。

【0015】

【課題を解決するための手段】この発明では、入力信号の周波数特性の概形が取り除かれた周波数領域係数と、そのスペクトルの包絡を求め、そのスペクトル包絡形状を強調し、その強調されたスペクトル包絡により周波数領域係数を逆平坦化する。特にスペクトル包絡形状を求める際パーク尺度（聴覚上で周波数分解能が同一となる）の周波数軸で等分解能をもたせるとより高い能率で処理を行うことができる。

【0016】雑音感の多い音響信号は、スペクトルの大小を強調することにより雑音感を低減することができる。この発明ではこの処理を周波数領域で行うので、少ない演算量で詳細な処理が可能である。この発明を変換符号化方式の復号器に組み込む場合には、この発明の処理過程の一つである周波数 - 時間変換処理を共有できるので演算量の点で特に有利である。

【0017】

*数領域係数を出力音響信号に変換出力する。変換の方法としては、IMDCT (Inverse Modified Discrete Cosine Transformation, 逆変形離散コサイン変換) や、IDCT (Inverse Discrete Cosine Transformation, 逆離散コサイン変換)、IDFT (Inverse Discrete Fourier Transformation, 逆離散フーリエ変換) などを用いることができる。なお、周波数 - 時間変換部では、後処理として、出力信号サンプル列の窓掛けとフレーム結合が必要である。窓掛けは従来手法と同様に行う。

10 【0012】更に符号化音声の雑音感を低減するためにスペクトルの山谷を強調するポストフィルタ25に変換部24よりの復号化音声信号を入力することが知られている。このポストフィルタ25の典型としては線形予測係数に基づく以下の形式がある。

【0013】

【数1】

【発明の実施の形態】図1にこの発明の第1実施例を示す。この実施例では平坦化された周波数領域係数とスペクトル包絡とが端子51、52にそれぞれ入力され、端子53から時間領域信号を出力とする。平坦化周波数領域係数は、例えば図3中の符号化器10で説明したように入力音響信号を時間 - 周波数変換した後、スペクトル包絡を用いて平坦化することによって求めてもよいし、図3中の復号器20に示したように変換符号化方法の復号器において、残差再生部21より再生された平坦化周波数領域係数を用いてもよい。時間 - 周波数変換は、先に述べたように離散フーリエ変換 (Discrete Fourier Transformation, DCT), 離散コサイン変換 (Discrete Cosine Transformation, DCT), 変形離散コサイン変換 (Modified Discrete Cosine Transformation, MDCT) などを用いることができる。これらの変換は、入力Nサンプルごとに行う。Nの値は例えば入力信号のサンプリング周波数が48kHzの場合512ないし4096程度が良好である。

40 【0018】スペクトル包絡は変換符号化方法の復号器において、図3の周波数特性概形再生部22により再生されたスペクトル包絡を用いてもよいし、入力音響信号を時間 - 周波数変換して周波数領域係数を求め、その周波数領域係数の概形を求めてもよい。スペクトル包絡の表現方法として先に述べたように、スケールファクタ、線形予測スペクトルなどを用いることができる。スケールファクタは、周波数領域係数を複数の周波数バンドごとにまとめた各バンドごとの代表値である。代表値はバンド内の係数の振幅の最大値でもよいし平均値でもよい。また各周波数のバンド幅は、線形スケール (Hzスケ

ール)で一定幅でもよいし、非線形スケール(例えばバークスケール)で一定幅としてもよい。特にバークスケールで一定幅とした場合には、聴感的に高効率な処理が可能である。線形予測スペクトルは、線形予測係数を周波数分析し、その逆数を求めることにより与えられる。線形予測係数は入力音響信号を線形予測分析して求めてもよいし、符号化方法の復号器において、再生された線

$$w(i) = w_0 \cdot (w(i) / w_0)^q \quad (2)$$

ここで、 $w(i)$ は変形後のスペクトル包絡、 $w(i)$ は入力スペクトル包絡、 w_0 は変形の基準値、 q は1以上の定数、例えば2~4、 i はスペクトル包絡のサンプル番号である。基準値 w_0 は任意に選ぶことができるが、スペクトル包絡の平均値とすると効果的である。また、式(2)の変形を一律に行うのではなく、基準値 w_0 よりもスペクトル包絡の値 w が小さいときのみ変形を行ってもよい。

$$y(j) = w(j) \times x(j)$$

ただし、 y は逆平坦化して得られた周波数領域係数、 x は平坦化周波数領域係数、 j はサンプル番号である。最後に、逆平坦化して得られた周波数領域係数を変換部56で周波数-時間変換して音響信号出力を得る。周波数-時間変換の方法として、逆離散フーリエ変換(Inverse Discrete Fourier Transformation, IDFT)、逆離散コサイン変換(Inverse Discrete Cosine Transformation, IDCT)、逆変形離散コサイン変換(Inverse Modified Discrete Cosine Transformation, IMDCT)などを用いることができる。

【0022】図2にこの発明の第2実施例を示す。スペクトル包絡強調、逆平坦化の手法は図1に示した第1実施例と同様である。第1実施例との違いは、複数のスペクトル包絡を用い、別々に強調処理を行うことである。端子61よりの微細スペクトル包絡は端子62よりの大局的スペクトル包絡よりもより細かい。例えばそれぞれバークスケール上で等間隔に分割したスケールファクタと線形予測スペクトルなどを用いる。どちらのスペクトル包絡も第1実施例で述べた種類のスペクトル包絡を用いることができる。また微細スペクトル包絡として、ピッチ包絡を用いてもよい。ピッチ包絡は、基本周波数の整数倍ごとに鋭いピークを持つ包絡であり、入力音響信号を分析して求めてもよいし、符号化方法の復号器にお

いて、再生されたピッチ包絡を用いたり、あるいは再生されたピッチ情報からピッチ包絡を用いてもよい。

【0023】大局的スペクトル包絡はスペクトル包絡強調部63で第1実施例と同様に強調処理が行われ、また微細スペクトル包絡もスペクトル包絡強調部64で同様に強調処理が行われる。端子51よりの平坦化周波数領域係数は逆平坦化部65でスペクトル包絡強調部63よ

形予測係数を用いてもよい。端子52に入力されたスペクトル包絡はスペクトル包絡強調部54で強調処理がなされる。この強調処理では、値が大きいたときには更に大きく、値が小さいときには更に小さくすることを行う。例えば式(2)のような変形を行う。

【0019】

$$(2)$$

【0020】次に、端子51に入力された平坦化周波数領域係数を強調されたスペクトル包絡を用いて逆平坦化部55で逆平坦化する。この際、強調されたスペクトル包絡のサンプル点数は平坦化周波数領域係数のサンプル点数と一致している必要がある。一致していない場合には、補間・間引きなどの処理によりサンプル点数を一致させる。逆平坦化は次式(3)に従って行う。

【0021】

$$(3)$$

りの強調処理された大局的スペクトル包絡により逆平坦化処理がなされ、この逆平坦化された周波数領域係数は、逆平坦化部66でスペクトル包絡強調部64で強調処理された微細スペクトル包絡により平坦化処理され、その平坦化処理された周波数領域係数が変換部56で周波数-時間変換がなされて出力される。

【0024】なお、スペクトル包絡強調-逆平坦化の組み合わせは第2実施例のように2つに限定する必要はなく、更に多い組み合わせを用意してもよい。

【0025】

【発明の効果】以上述べたように、この発明によれば周波数領域で音響信号のスペクトル包絡の強調処理を行い、スペクトル包絡の強弱を強調することにより、スペクトルの谷間に歪みがあるために聞こえる雑音感を低減することができる。この処理を時間領域でなく周波数領域で行うことにより、時間領域で必要だった畳み込み演算を行う必要がなく、詳細な処理でも小さな演算量で行うことができる。また、この発明を変換符号化方法の復号器と組み合わせる場合、例えば図3中の逆平坦化部23を、この発明の逆平坦化部53または65と共有することができ、また周波数-時間変換部56を図3中の周波数-時間変換部24と共有でき、演算量とメモリ規模の点で特に有利である。

【図面の簡単な説明】

【図1】この発明の第1実施例の機能構成を示すブロック図。

【図2】この発明の第2実施例の機能構成を示すブロック図。

【図3】従来の音響信号変換符号化・復号化方法における機能構成例を示すブロック図。

【図 1】

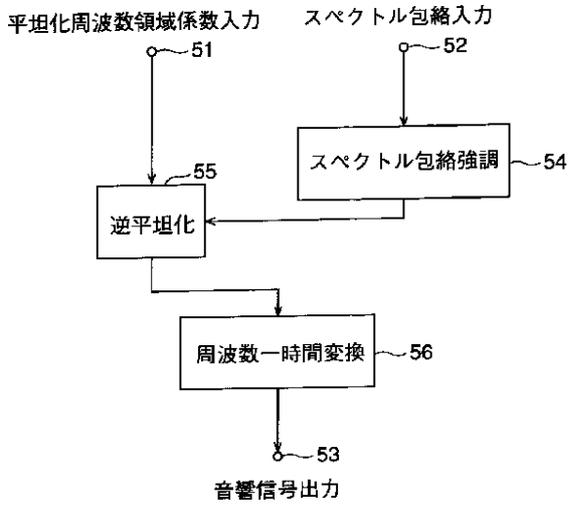


図 1

【図 2】

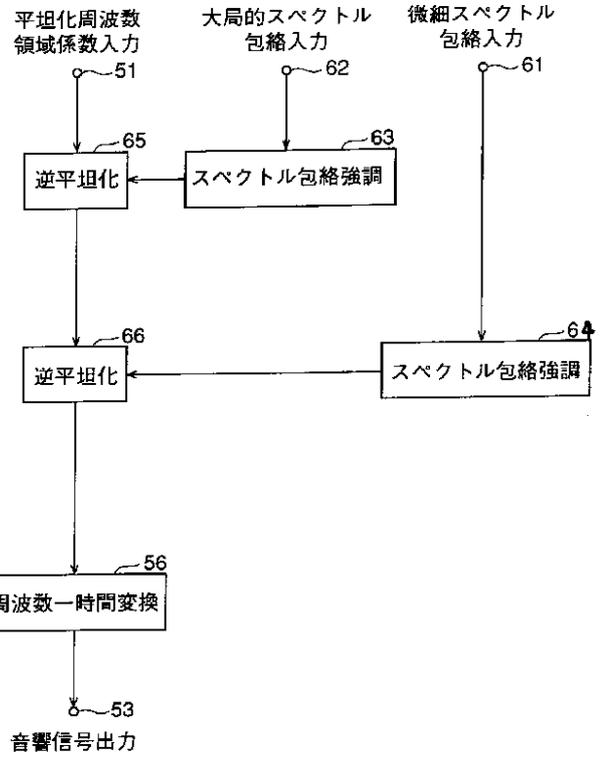


図 2

【図 3】

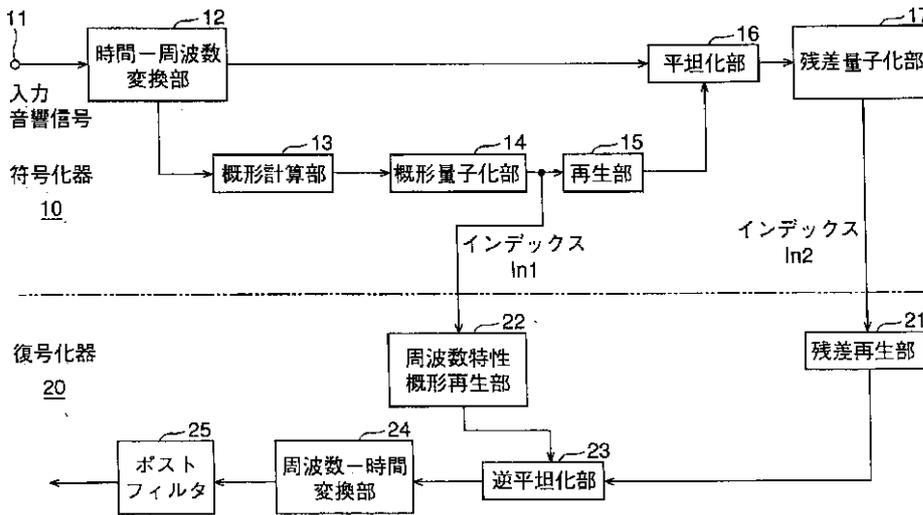


図 3

フロントページの続き

(72)発明者 神 明夫

東京都新宿区西新宿三丁目19番2号 日
本電信電話株式会社内

(56)参考文献 特開 昭61 - 286900 (J P , A)
特開 平 6 - 318876 (J P , A)
特開 平 7 - 107053 (J P , A)
特開 平 7 - 248794 (J P , A)
特開 平 8 - 110796 (J P , A)
特表 昭62 - 500138 (J P , A)

(58)調査した分野(Int.Cl.⁷, D B名)

G10L 13/00

G10L 21/02