

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-35839

(P2019-35839A)

(43) 公開日 平成31年3月7日(2019.3.7)

(51) Int.Cl.		F I	テーマコード (参考)	
G 1 O L	19/03	(2013.01)	G 1 O L	19/03
G 1 O L	19/035	(2013.01)	G 1 O L	19/035
				Z

審査請求 未請求 請求項の数 7 O L (全 15 頁)

(21) 出願番号	特願2017-156486 (P2017-156486)	(71) 出願人	398034168
(22) 出願日	平成29年8月14日 (2017.8.14)		株式会社アクセル
			東京都千代田区外神田四丁目14番1号
		(74) 代理人	110000279
			特許業務法人ウィルフォート国際特許事務所
			所
		(72) 発明者	道上 将志
			東京都千代田区外神田四丁目14番1号
			株式会社アクセル内

(54) 【発明の名称】 音声処理装置、音声処理方法、及び音声処理プログラム

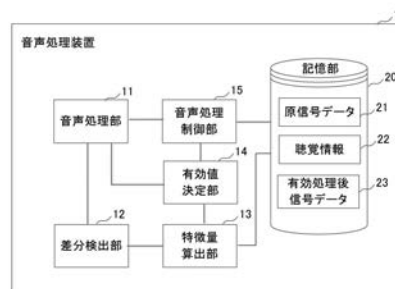
(57) 【要約】

【課題】再生信号の品質を向上させることができるようにする。

【解決手段】対象信号に対して音声処理を行って処理後信号を生成する音声処理部11を有する音声処理装置1において、音声処理部11のパラメータの値を複数の値に変更させて、音声処理部11に音声処理を実行させる音声処理制御部15と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、対象信号のエネルギーと、処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出する差分検出部12と、検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらを合計した特徴量を算出する特徴量算出部13と、特徴量算出部13により算出される複数の特徴量に基づいて、音声処理の利用に適したパラメータの値である有効値を決定する有効値決定部14と、を備えるようにする。

【選択図】 図1

図1



【特許請求の範囲】

【請求項 1】

対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理部を有する音声処理装置であって、

前記音声処理部による前記音声処理における 1 以上のパラメータの値を複数の値に変更させて、前記音声処理部に前記音声処理を実行させる音声処理制御部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記対象信号のエネルギーと、前記処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出する差分検出部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記差分検出部により検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらに基づく特徴量を算出する特徴量算出部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記特徴量算出部により算出される複数の前記特徴量に基づいて、前記音声処理の利用に適したパラメータの値である有効値を決定する有効値決定部と、
を備える音声処理装置。

10

【請求項 2】

前記音声処理における前記パラメータの値を、前記有効値検出部により検出された前記有効値に設定して、前記音声処理を実行させ、前記音声処理により得られた処理後信号を有効な処理後信号である有効処理後信号として記憶部に格納させる有効信号生成制御部をさらに有する

請求項 1 に記載の音声処理装置。

20

【請求項 3】

前記音声処理部は、前記対象信号における量子化雑音の周波数特性を変化させるノイズシェーピングを行うノイズシェーピング部を含み、

前記パラメータは、前記ノイズシェーピング部におけるノイズシェーピングに関わるパラメータである

請求項 1 又は請求項 2 に記載の音声処理装置。

【請求項 4】

前記差分検出部は、前記対象信号の所定のサイズのブロックを単位として前記差分を検出し、

前記特徴量算出部は、前記ブロックを単位として前記特徴量を算出し、

前記有効値決定部は、前記ブロックを単位として、前記有効値を決定する

請求項 3 に記載の音声処理装置。

30

【請求項 5】

前記人間の聴覚の周波数成分に対応する感度に関する情報を記憶する聴覚情報記憶部をさらに備え、

前記特徴量算出部は、前記聴覚情報記憶部の前記感度に関する情報に基づいて、前記重みを決定する

請求項 1 乃至請求項 4 のいずれか一項に記載の音声処理装置。

40

【請求項 6】

対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理部を有する音声処理装置による音声処理方法であって、

前記音声処理部による前記音声処理における 1 以上のパラメータの値を複数の値に変更させて、前記音声処理部に前記音声処理を実行させ、

前記パラメータを複数の値のそれぞれに変更させた際に、前記対象信号のエネルギーと、前記処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出し、検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらを合計した特徴量を算出し、算出された複数の前記特徴量に基づいて、前記音声処理の利用に適したパラメータの値である有効値を

50

決定する
音声処理方法。

【請求項 7】

対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理装置を構成するコンピュータに実行させる音声処理プログラムであって、

前記音声処理プログラムは、

前記コンピュータを、

前記音声処理部による前記音声処理における 1 以上のパラメータの値を複数の値に変更させて、前記音声処理部に前記音声処理を実行させる音声処理制御部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記対象信号のエネルギーと、前記処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出する差分検出部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記差分検出部により検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらに基づく特徴量を算出する特徴量算出部と、

前記パラメータの値を複数の値に変更させた場合のそれぞれにおける、前記特徴量算出部により算出される複数の前記特徴量に基づいて、前記音声処理の利用に適したパラメータの値である有効値を決定する有効値決定部と、

して機能させる音声処理プログラム。

10

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理装置等に関する。

【背景技術】

【0002】

信号処理技術として、ADPCM (adaptive differential pulse code modulation) が知られている。ADPCMは、ハードウェア実装の回路規模が小さく、かつ高速なデコードが可能であるため、複数の音声信号を同時に再生するアミューズメント分野などで重要な技術である。

30

【0003】

しかし、ADPCMは、同程度の圧縮率を実現できる他の技術、例えば、MP3等の変換符号化方式のコーデック技術よりも再生される音質が劣ることがある。

【0004】

これに対して、ADPCMによって発生する量子化雑音と、ADPCMのエンコード対象である原信号との周波数スペクトルの傾向を考慮し、再生信号の品質を向上させる緩やかな遮蔽周波数を持つポストフィルタを用いる技術が知られている。

40

【0005】

関連する技術として、複数の時系列信号によって形成される音声信号を生成するための量子化処理部と、量子化処理部において量子化を行った際に発生する量子化ノイズについてノイズシェーピングを行うためのノイズシェーピング手段とを備えた音声信号処理装置が知られている。ノイズシェーピング手段は、量子化された音声信号のノイズシェーピングを実現できる特性を有するポストフィルタと、ポストフィルタの逆特性を有するプリフィルタとを備え、ポストフィルタを量子化処理部の後段に、プリフィルタを量子化処理部の前段にそれぞれ設ける技術が開示されている（例えば、特許文献1参照）。

【0006】

また、関連する他の技術として、量子化誤差を低減するノイズシェーピングフィルタの係数設定方法において、所定のサンプル数の周波数分析結果に基づいたパワースペクトル

50

形状の逆数をとった後、逆直交変換して求められた自己相関係数に基づいてフィルタの係数を演算する技術が知られている（例えば、特許文献2）。これによると、通常のいわゆるDSP（digital signal processor）によって量子化誤差低減の実時間処理が可能となるノイズシェーピングフィルタの係数設定方法が提供できる。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2016-213683号公報

【特許文献2】特開平4-72907号公報

【発明の概要】

10

【発明が解決しようとする課題】

【0008】

例えば、上記したポストフィルタを用いて再生信号の品質を向上させる音声処理技術は、高周波帯域の特性を改善するものであり、低・中周波数帯域の量子化雑音は低減されない。

【0009】

これに対して、高周波帯域だけでなく、低・中周波数帯域の量子化雑音も低減するために、量子化雑音のエネルギーを最小とする方法が考えられる。量子化雑音のエネルギーを最小にすると、量子化雑音のスペクトル包絡は全周波数帯域で平坦となる。

【0010】

20

しかしながら、このように量子化雑音のスペクトル包絡を全周波数帯域で平坦としたとしても、人間にとって、再生信号の品質が向上しているとは限らない。

【0011】

一方、原信号の周波数特性を考慮してノイズシェーピングを施し、量子化雑音のスペクトルを変形し、原信号のエネルギーの強いところに量子化雑音を集中させることにより、原信号のエネルギーの弱いところの量子化雑音を軽減させる方法が考えられる。しかしながら、ノイズシェーピングの強度を固定してしまうと、量子化雑音の総量を増加させてしまつて再生信号の品質が逆に劣化してしまう問題がある。

【0012】

本発明は、上記事情に鑑みなされたものであり、その目的は、再生信号の品質を向上させることのできる技術を提供することにある。

30

【課題を解決するための手段】

【0013】

上記目的を達成するため、第1の観点に係る音声処理装置は、対象信号に対して所定の音声処理を行つて処理後信号を生成する音声処理部を有する音声処理装置であつて、音声処理部による音声処理における1以上のパラメータの値を複数の値に変更させて、音声処理部に音声処理を実行させる音声処理制御部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、対象信号のエネルギーと、処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出する差分検出部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、差分検出部により検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらに基づいた特徴量を算出する特徴量算出部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、特徴量算出部により算出される複数の特徴量に基づいて、音声処理の利用に適したパラメータの値である有効値を決定する有効値決定部と、を備える。

40

【0014】

上記音声処理装置において、音声処理におけるパラメータの値を、有効値検出部により検出された有効値に設定して、音声処理を実行させ、音声処理により得られた処理後信号を有効な処理後信号である有効処理後信号として記憶部に格納させる有効信号生成制御部をさらに有するようにしてもよい。

50

【0015】

また、上記音声処理装置において、音声処理部は、対象信号における量子化雑音の周波数特性を変化させるノイズシェーピングを行うノイズシェーピング部を含み、パラメータは、ノイズシェーピング部におけるノイズシェーピングに関わるパラメータであってもよい。

【0016】

また、上記音声処理装置において、差分検出部は、対象信号の所定サイズのブロックを単位として差分を検出し、特徴量算出部は、ブロックを単位として特徴量を算出し、有効値決定部は、ブロックを単位として、有効値を決定するようにしてもよい。

【0017】

また、上記音声処理装置において、人間の聴覚の周波数成分に対応する感度に関する情報を記憶する聴覚情報記憶部をさらに備え、特徴量算出部は、聴覚情報記憶部の感度に関する情報に基づいて、重みを決定するようにしてもよい。

10

【0018】

また、上記目的を達成するため、第2の観点に係る音声処理方法は、対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理部を有する音声処理装置による音声処理方法であって、音声処理部による音声処理における1以上のパラメータの値を複数の値に変更させて、音声処理部に音声処理を実行させ、パラメータを複数の値のそれぞれに変更させた際に、原信号のエネルギーと、処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出し、検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらに基づいた特徴量を算出し、算出された複数の特徴量に基づいて、音声処理の利用に適したパラメータの値である有効値を決定する。

20

【0019】

また、上記目的を達成するため、第3の観点に係る音声処理プログラムは、対象信号に対して所定の音声処理を行って処理後信号を生成する音声処理装置を構成するコンピュータに実行させる音声処理プログラムであって、音声処理プログラムは、コンピュータを、音声処理部による音声処理における1以上のパラメータの値を複数の値に変更させて、音声処理部に音声処理を実行させる音声処理制御部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、対象信号のエネルギーと、処理後信号に基づいて生成される再生信号のエネルギーとの所定の周波数成分ごとの差分を検出する差分検出部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、差分検出部により検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらに基づいた特徴量を算出する特徴量算出部と、パラメータの値を複数の値に変更させた場合のそれぞれにおける、特徴量算出部により算出される複数の特徴量に基づいて、音声処理の利用に適したパラメータの値である有効値を決定する有効値決定部として機能させる。

30

【発明の効果】

【0020】

本発明によれば、再生信号の品質を向上させることができる。

40

【図面の簡単な説明】

【0021】

【図1】図1は、一実施形態に係る音声処理装置の機能ブロック図である。

【図2】図2は、一実施形態に係る音声処理装置の音声処理部の機能ブロック図である。

【図3】図3は、一実施形態に係る人間の聴覚の周波数に対する感度を示すAカーブを示す図である。

【図4】図4は、一実施形態に係る音声処理装置のハードウェア構成図である。

【図5】図5は、一実施形態に係る音声生成処理のフローチャートである。

【図6】図6は、一実施形態に係る特徴量算出処理のフローチャートである。

【発明を実施するための形態】

50

【 0 0 2 2 】

実施形態について、図面を参照して説明する。なお、以下に説明する実施形態は特許請求の範囲に係る発明を限定するものではなく、また実施形態の中で説明されている諸要素及びその組み合わせの全てが発明の解決手段に必須であるとは限らない。

【 0 0 2 3 】

まず、一実施形態に係る音声処理装置について説明する。

【 0 0 2 4 】

図 1 は、一実施形態に係る音声処理装置の機能ブロック図である。図 2 は、一実施形態に係る音声処理装置の音声処理部の機能ブロック図である。

【 0 0 2 5 】

音声処理装置 1 は、音声処理部 1 1 と、差分検出部 1 2 と、特徴量算出部 1 3 と、有効値決定部 1 4 と、有効信号生成制御部の一例としての音声処理制御部 1 5 と、聴覚情報記憶部の一例としての記憶部 2 0 とを備える。

【 0 0 2 6 】

音声処理部 1 1 は、対象信号の一例としての原信号 $s(n)$ に対して、所定の音声処理（例えば、ADPCM のエンコード処理）を行う。音声処理部 1 1 は、例えば、図 2 に示すように、量子化部の一例としての適応量子化部 3 1 と、逆量子化部の一例としての逆適応量子化部 3 2 と、演算部 3 3、3 4 と、ノイズシェーピング部 3 5 と、演算部 3 6 と、適応予測部 3 7 と、演算部 3 8 とを備える。

【 0 0 2 7 】

演算部 3 3 は、原信号 $s(n)$ と、適応予測部 3 7 から出力される予測信号との差分である差分信号を出力する。演算部 3 4 は、演算部 3 3 から出力された差分信号と、ノイズシェーピング部 3 5 により出力された信号とを加算する。

【 0 0 2 8 】

適応量子化部 3 1 は、演算部 3 4 から出力された信号（ノイズシェーピング後の差分信号）に対して、適応量子化処理（デコード処理）を実行することにより、処理後信号 $x(n)$ を出力する。逆適応量子化部 3 2 は、適応量子化部 3 1 により出力された処理後信号 $x(n)$ に対して、逆適応量子化処理（デコード処理）を実行する。

【 0 0 2 9 】

演算部 3 6 は、逆適応量子化部 3 2 から出力された信号と、演算部 3 4 から出力された信号との差分の信号を出力する。演算部 3 6 から出力された信号は、適応量子化部 3 1 に入力された信号に対する量子化雑音の信号である。

【 0 0 3 0 】

ノイズシェーピング部 3 5 は、演算部 3 6 から出力された信号を入力として、量子化雑音の信号の形状を変形させるノイズシェーピングを行う。

【 0 0 3 1 】

ここで、量子化雑音の伝達関数（雑音伝達関数） $H(z)$ について、ノイズシェーピングの強度を変更するためのパラメータ（強度（0 以上 1 以下））を導入した伝達関数 $H(\gamma z)$ とすると、この伝達関数は、以下の式（1）に示すように表すことができる。式（1）において、 $H(\gamma z)$ （なお、本明細書では、「 H 」は、 H の上に「 γ 」を表すこととする。）は、ノイズシェーピング部 3 5 のフィードバック要素に相当する。

【 0 0 3 2 】

【数 1】

$$H(\gamma z) = 1 + \overline{H}(\gamma z) \quad \cdot \cdot \cdot \quad (1)$$

【 0 0 3 3 】

本実施形態では、雑音伝達関数 $H(\gamma z)$ を、再生信号 $y(n)$ を自己回帰モデルとし

10

20

30

40

50

た場合における伝達関数 $A(z)$ に対して上記したパラメータである強度 α_i を導入した伝達関数 $A(z)$ としている。このように、雑音伝達関数 $H(z)$ を、伝達関数 $A(z)$ とすることにより、原信号と同様なスペクトル包絡によるノイズシェーピングを行うことができ、所謂聴覚マスキングによる雑音の隠蔽が期待できる。また、後述するように、強度 α_i の値を再生信号の品質の向上に有効な値に設定することができるので、雑音の総エネルギーを抑制し、再生信号による音質を向上することができる。

【0034】

ここで、自己回帰モデルの伝達関数 $A(z)$ について説明する。

【0035】

まず、時刻 n の再生信号 $y(n)$ の予測値を、過去 k 個 (k は所定の数) のサンプルと、 k 個の予測係数 α_i とを用いた線形和で予測することとすると、時刻 n の再生信号の予測値を $\hat{y}_k(n)$ (なお、本明細書では、「 \hat{y} 」は、 y の真上に「 $\hat{}$ 」を表すこととする。) とすると、以下の式 (2) に示すように表される。

10

【0036】

【数2】

$$\hat{y}_k(n) = -\sum_{i=1}^k \alpha_i y(n-i) \quad \cdot \cdot \cdot \quad (2)$$

20

【0037】

予測値 $\hat{y}_k(n)$ と、実際の値 $y(n)$ との差分信号を $d(n)$ とすると、差分信号 $d(n)$ は、以下の式 (3) に示すように表される。

【0038】

【数3】

$$d(n) = y(n) - \hat{y}_k(n) \quad \cdot \cdot \cdot \quad (3)$$

【0039】

この式 (3) に、式 (2) を代入すると、以下の式 (4) に示す関係が得られる。この式 (4) は、自己回帰モデルと呼ばれる。

30

【0040】

【数4】

$$d(n) = y(n) + \sum_{i=1}^k \alpha_i y(n-i) \quad \cdot \cdot \cdot \quad (4)$$

【0041】

この式 (4) に対して、 z 変換することにより、以下の式 (5) に示すように、式 (4) の伝達関数 $A(z)$ が得られる。

40

【0042】

【数5】

$$A(z) = \frac{1}{1 + \sum_{i=1}^k \alpha_i z^{-i}} \quad \cdot \cdot \cdot \quad (5)$$

【0043】

なお、予測係数 α_i は、Levinson-Durbin のアルゴリズムを用いること

50

により求めることができる。

【0044】

$H(z) = A(z)$ と、式(5)及び式(1)によると、フィードバック要素 $H(z)$ は、以下の式(6)に示すように表される。

【0045】

【数6】

$$\overline{H}(yz) = A(yz) - 1 = - \frac{\sum_{i=1}^k \alpha_i yz^{-i}}{1 + \sum_{i=1}^k \alpha_i yz^{-i}} \quad \dots \quad (6)$$

10

【0046】

演算部38は、逆適応量子化部32の出力した信号と、適応予測部37が出力した予測信号とを加算した再生信号 $y(n)$ を出力する。

【0047】

適応予測部37は、再生信号 $y(n)$ を入力として、原信号 $s(n)$ を予測する予測信号を出力する。

【0048】

図1の説明に戻り、記憶部20は、原信号データ21と、聴覚情報22と、有効処理後信号データ23とを記憶する。

20

【0049】

原信号データ21は、音声処理部11の処理対象となるデジタルデータ化された原信号 $s(n)$ である。原信号 $s(n)$ は、例えば、アナログの原信号を、所定のサンプリング間隔ごとに、各時点におけるアナログ値を、そのアナログ値に対応するデジタル値に変換したものである。

【0050】

聴覚情報22は、人間の聴覚の周波数成分に対応する感度に関する情報(聴覚情報)である。聴覚情報としては、例えば、人間が感覚的に同じ音の大きさであると認識する音圧レベルの周波数変化を示す等ラウドネス曲線のミラー特性を持つ曲線に対応する情報や、JIS C 1509-1:2005のAカーブやCカーブに対応する情報である。

30

【0051】

ここで、Aカーブについて説明する。

【0052】

図3は、一実施形態に係る人間の聴覚の周波数に対する感度を示すAカーブを示す図である。図3において、横軸は周波数 [kHz] であり、縦軸は、音圧 [dB] である。縦軸については、周波数 1.0 [kHz] の音圧を 0 [dB] としている。

【0053】

人間の聴覚は、図3のAカーブに示すように、周波数に応じて、同一の音圧と感ずる値が異なっている、すなわち、感度が異なっている。

40

【0054】

図1の説明に戻り、有効処理後信号データ23は、ノイズシェーピング部35の強度に対して、有効値決定部14により決定された有効値が設定されている場合において、音声処理部11により生成される処理後信号 $x(n)$ である。この有効処理後信号データ23を、デコードすることにより、高品質の音声を再現できる再生信号を生成することができる。

【0055】

差分検出部12は、ノイズシェーピング部35のパラメータの値を複数の値に変更させた場合のそれぞれにおいて、原信号 $s(n)$ の所定の単位(ブロック)を処理対象として

50

、そのブロック（処理対象ブロック）についての対数パワースペクトルと、原信号 $s(n)$ の処理対象ブロックを適応量子化部 31 でエンコードし、逆適応量子化部 32 でデコードした後に得られる再生信号 $y(n)$ （図 2 では、演算部 38 から出力される信号）についての対数パワースペクトルとを算出し、求めた 2 つの対数パワースペクトルの差の絶対値を周波数成分ごとに求める。具体的には、差分検出部 12 は、以下の式（7）により、原信号 $s(n)$ のブロックについての対数パワースペクトル $S(f)$ を算出する。

【0056】

【数 7】

$$S(f) = 20 \log |FFT(s(n))| \quad \cdot \cdot \cdot \quad (7)$$

10

【0057】

また、差分検出部 12 は、以下の式（8）により、再生信号 $y(n)$ のブロックについての対数パワースペクトル $Y(f)$ を算出する。

【0058】

【数 8】

$$Y(f) = 20 \log |FFT(y(n))| \quad \cdot \cdot \cdot \quad (8)$$

20

【0059】

次いで、差分検出部 12 は、以下の式（9）により、原信号 $s(n)$ の対数パワースペクトル $S(f)$ と、再生信号 $y(n)$ のブロックについての対数パワースペクトル $Y(f)$ との周波数成分ごとの絶対値の差分である差分パワースペクトル $D(f)$ を算出する。

【0060】

【数 9】

$$D(f) = |S(f) - Y(f)| \quad \cdot \cdot \cdot \quad (9)$$

30

【0061】

特徴量算出部 13 は、ノイズシェーピング部 35 のパラメータの値を複数の値に変更させた場合のそれぞれにおいて、差分パワースペクトル $D(f)$ の各周波数成分に対して、人間の聴力に基づく重み $w(f)$ を乗算し、得られた値のすべてに基づいて（例えば、すべてを加算することにより）、特徴量（FDD: Frequency Domain Difference）を算出する。

【0062】

ここで、聴覚情報 22 を、図 3 に示す A カーブが示す周波数と音圧との対応関係としている場合には、各周波数成分に対する重み $w(f)$ は、周波数に対する縦軸の音圧の値を $C_A(f)$ とすると、以下の式（10）により算出される。

【0063】

【数 10】

$$w(f) = 10^{\frac{C_A(f)}{20}} \quad \cdot \cdot \cdot \quad (10)$$

50

【 0 0 6 4 】

特徴量算出部 1 3 は、特徴量 F D D を、以下の式 (1 1) により算出する。

【 0 0 6 5 】

【 数 1 1 】

$$FDD = \sum_f D(f)w(f) \quad \cdot \cdot \cdot \quad (11)$$

【 0 0 6 6 】

10

ここで、特徴量 F D D は、各周波数の原信号 $s(n)$ と再生信号 $y(n)$ とのエネルギーの差分に対して人間の聴覚の特性に応じた重みをつけて合計したものである。このことから、特徴量 F D D が小さいと、人間の聴覚に対しては、再生信号 $y(n)$ における量子化雑音の影響が少ない、すなわち、再生信号 $y(n)$ による音声の音質が良いことを示している。

【 0 0 6 7 】

音声処理制御部 1 5 は、原信号 $s(n)$ のブロックのそれぞれを対象（処理対象ブロック）として、音声処理部 1 1 による音声処理における 1 以上のパラメータの値を複数の値に変更させて、それぞれの値を設定したそれぞれの状態で音声処理部 1 1 に音声処理を実行させ、それぞれの音声処理に対して、差分検出部 1 2 及び特徴量算出部 1 3 による処理
20

【 0 0 6 8 】

有効値決定部 1 4 は、音声処理制御部 1 5 によって変更されたパラメータの各値に対応する特徴量算出部 1 3 により算出された特徴量 F D D に基づいて、劣化の少ない音声処理
30

【 0 0 6 9 】

次に、音声処理装置 1 のハードウェア構成について詳細に説明する。

【 0 0 7 0 】

図 4 は、一実施形態に係る音声処理装置のハードウェア構成図である。

【 0 0 7 1 】

音声処理装置 1 は、制御回路 1 0 1 と、記憶装置 1 0 2 と、リーダライタ 1 0 3 と、通信
40

【 0 0 7 2 】

制御回路 1 0 1 は、例えば、プロセッサであり、音声処理装置 1 の全体を統括制御する。制御回路 1 0 1 は、記憶装置 1 0 2 に格納されたプログラムを実行することにより各種処理を実行する。本実施形態では、制御回路 1 0 1 は、記憶装置 1 0 2 に格納された音声処理プログラムを実行することにより、音声処理部 1 1、差分検出部 1 2、特徴量算出部
50

【 0 0 7 3 】

記憶装置 102 は、例えば、HDD (Hard Disk Drive)、SSD (Solid State Drive)、RAM、ROM 等であり、制御回路 101 に実行されるプログラム (音声処理プログラム等) や、各種情報を記憶する。記録装置 102 は、図 1 に示す記憶部 20 を構成する。

【0074】

リーダライタ 103 は、記録媒体 110 を着脱可能であり、記録媒体 110 からのデータの読み出し、及び記録媒体 110 へのデータの書き込みを行う。記録媒体 110 としては、例えば、SD メモリーカード、FD (フロッピーディスク：登録商標)、CD、DVD、BD (登録商標)、フラッシュメモリ等の非一時的記録媒体 (不揮発性記録媒体) がある。本実施形態においては、記録媒体 110 に、音声処理プログラム、処理に使用する原信号 $s(n)$ 、聴覚情報等を格納させておき、リーダライタ 103 により、これらを読み出して、利用するようにしてもよい。

10

【0075】

通信 I/F 104 は、ネットワーク 111 に接続されており、ネットワーク 111 に接続された他の装置との間でのデータの送受信を行う。

【0076】

次に、本実施形態に係る音声処理装置 1 における音声生成処理の動作について説明する。

【0077】

図 5 は、一実施形態に係る音声生成処理のフローチャートである。

20

【0078】

音声処理制御部 15 は、記憶部 20 から原信号 $s(n)$ の処理範囲内の未処理のブロック (処理対象ブロック) を抽出する (ステップ S101)。なお、原信号 $s(n)$ の処理範囲は、例えば、原信号 $s(n)$ の全体としてもよいし、原信号 $s(n)$ のうちのユーザから指定された一部の範囲としてもよい。

【0079】

次いで、音声処理制御部 15 は、音声処理部 11 のノイズシェーピング部 35 の強度に 0 を設定し (ステップ S102)、抽出した処理対象ブロックを音声処理部 11 に入力して音声処理を実行させる (ステップ S103)。

【0080】

次いで、音声処理制御部 15 は、差分検出部 12 及び特徴量算出部 13 に、音声処理部 11 による音声処理に用いた原信号 $s(n)$ と、音声処理部 11 により出力される再生信号 $y(n)$ とを用いて、特徴量 FDD を算出する処理 (特徴量算出処理) を実行させる (ステップ S104)。

30

【0081】

次いで、音声処理制御部 15 は、音声処理部 11 の強度 α を変更する (ステップ S105)。例えば、音声処理制御部 15 は、強度 α の値に所定値 (例えば、0.01) を加算する。

【0082】

次いで、音声処理制御部 15 は、強度 α が 1 より大きいかが否かを判定する (ステップ S106)。この結果、強度 α が 1 より大きくない場合 (ステップ S106: No) には、この強度 α による特徴量 FDD を算出する必要があるため、音声処理制御部 15 は、処理をステップ S103 に進める。一方、強度 α が 1 より大きい場合 (ステップ S106: Yes) には、強度 α を十分に小さく変更させて必要な特徴量 FDD を算出したことを意味するので、音声処理制御部 15 は、処理をステップ S107 に進める。

40

【0083】

ステップ S107 では、有効値決定部 14 が、ステップ S104 で算出された複数の特徴量 FDD に基づいて、量子化雑音が少ない再生信号を得ることのできる処理後信号を生成することのできる強度 α の値 (有効値) を決定し、音声処理制御部 15 が、決定した有効値を音声処理部 11 の強度 α に設定する。本実施形態では、有効値決定部 14 は、特徴

50

量 F D D が最も小さくなる強度 の値を、処理対象ブロックに対する有効値として決定する。

【 0 0 8 4 】

次いで、音声処理制御部 1 5 は、処理対象ブロックを音声処理部 1 1 に入力して音声処理を実行させ、音声処理部 1 1 により生成された処理後信号 $x(n)$ を有効処理後信号データとして記憶部 2 0 に格納する（ステップ S 1 0 8 ）。

【 0 0 8 5 】

次いで、音声処理制御部 1 5 は、原信号 $s(n)$ の処理範囲の全てのブロックに対して処理を行ったか否かを判定する（ステップ S 1 0 9 ）。この結果、処理範囲の全てのブロックに対して処理を行っていない場合（ステップ S 1 0 9 : N o ）には、音声処理制御部 1 5 は、処理をステップ S 1 0 1 に進めて、次のブロックを対象に処理を行う一方、処理範囲の全てのブロックに対して処理を行った場合（ステップ S 1 0 9 : Y e s ）には、音声処理制御部 1 5 は、音声生成処理を終了する。

10

【 0 0 8 6 】

次に、図 5 のステップ S 1 0 4 に示す特徴量算出処理について説明する。

【 0 0 8 7 】

図 6 は、一実施形態に係る特徴量算出処理のフローチャートである。

【 0 0 8 8 】

差分検出部 1 2 は、式（ 7 ）により、原信号 $s(n)$ の対数パワースペクトル $S(f)$ を算出する（ステップ S 2 0 1 ）。さらに、差分検出部 1 2 は、式（ 8 ）により、再生信号 $y(n)$ の対数パワースペクトル $Y(f)$ を算出する（ステップ S 2 0 2 ）。そして、差分検出部 1 2 は、式（ 9 ）により、対数パワースペクトル $S(f)$ と対数パワースペクトル $Y(f)$ との絶対値の差分である差分パワースペクトル $D(f)$ を周波数毎に算出する（ステップ S 2 0 3 ）。次いで、特徴量算出部 1 3 は、各差分パワースペクトル $D(f)$ に対して、対応する周波数に応じた人間の聴覚特性に応じた重み $w(f)$ を付与する。すなわち、特徴量算出部 1 3 は、式（ 1 0 ）により、各差分パワースペクトル $D(f)$ に対して対応する重み $w(f)$ を乗算する（ステップ S 2 0 4 ）。次いで、特徴量算出部 1 3 は、式（ 1 1 ）により、重みが付与された周波数ごとの差分パワースペクトル $D(f)$ を合計することにより、特徴量 F D D を算出する（ステップ S 2 0 5 ）。

20

【 0 0 8 9 】

以上説明したように、本実施形態に係る音声処理装置 1 によると、音声処理における 1 以上のパラメータの値を複数の値に変更させて、音声処理部 1 1 に音声処理を実行させ、原信号 $s(n)$ のエネルギーと、再生信号 $y(n)$ のエネルギーとの所定の周波数成分ごとの差分を検出し、検出された周波数成分ごとの差分に対して、人間の聴覚の周波数成分に対応する感度に応じた重み付けを行い、それらを合計した特徴量を算出し、パラメータを複数の値のそれぞれに変更させた際に特徴量算出部 1 3 により算出される複数の特徴量に基づいて、量子化誤差の少ない再生信号が得られる処理後信号の生成に適したパラメータの有効値を決定するようにしたので、音声処理によって生成される処理後信号により生成される再生信号による音声の品質を向上することができる。

30

【 0 0 9 0 】

なお、本発明は、上述の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲で、適宜変形して実施することが可能である。

40

【 0 0 9 1 】

例えば、上記実施形態においては、原信号の各ブロックに対して、ノイズシェーピング部 3 5 の強度 の値を常に所定の範囲（ 0 以上 1 以下の範囲）内で変更して特徴量 F D D を算出するようにしていたが、本発明はこれに限られず、所定数前（例えば、直前）のブロックに対する強度 の有効値を基準に、その有効値の周辺の限られた範囲内で強度 の値を変更するようにしてもよい。ブロックに対する強度 の有効値は、時間的に近い範囲（例えば、直前）のブロックの強度 の有効値と近い値となる傾向が高く、このようにすることにより、有効値を適切に決定できると共に、有効値を決定するために必要な処理負

50

荷を低減することができる。

【0092】

また、上記実施形態においては、各ブロックのそれぞれに対して特徴量を算出して強度の有効値を決定するようにしていたが、本発明はこれに限られず、例えば、複数のブロックのうちの一つのブロックに対して強度を変更して有効値を決定し、それら複数のブロックに対する強度として決定した1つの有効値を用いて音声処理をするようにしてもよい。このようにすると、ブロックごとに特徴量を算出して強度の有効値を決定する場合に比して、処理負荷を低減することができ、短時間で比較的品質の良い再生信号を生成することのできる処理後信号を生成することができる。

【0093】

また、上記実施形態では、ノイズシェーピング部35の強度の値を変化させた場合の特徴量を算出し、特徴量に基づいて、強度の有効値を決定するようにしていたが、本発明はこれに限られず、音声処理における他のパラメータの値（例えば、音声処理部にLPF（ローパスフィルタ）があれば、そのフィルタの遮断周波数、ノイズシェーピング部35の雑音伝達関数として用いる自己回帰モデルの予測係数の個数、量子化する際のブロックのサイズ等の少なくともいずれか1つのパラメータ）を変化させて、各パラメータ値での特徴量を算出し、その特徴量に基づいて、そのパラメータの有効値を決定するように、有効処理後信号を生成する際に用いるようにしてもよい。

【0094】

また、上記実施形態では、共通の聴覚情報22に基づいて、重み付けを決定し、特徴量を算出するようにしていたが、本発明はこれに限られず、例えば、処理後信号から生成される再生信号に基づく音声を聞く対象となる利用者の年代別（例えば、10代、20代、30代、・・・等）に、その年代に対応する聴覚情報を用意し、それら聴覚情報に基づいて重み付けを決定し、特徴量を算出するようにしてもよい。この場合には、音声を聞く主な年代の聴覚情報22を利用して特徴量を算出して、パラメータの有効度を決定して、処理後信号を生成するようにしてもよく、年代ごとの特徴量を算出し、年代ごとのパラメータの有効度を決定し、各年代に適した処理後信号を生成するようにしてもよい。

【0095】

また、上記実施形態では、聴覚情報22として、主に等ラウドネス曲線に対応する情報の例を示していたが、本発明はこれに限られず、例えば、等ラウドネス曲線に対応する情報に代えて、又は等ラウドネス曲線に対応する情報に加えて、時間的に前の音によりその後の音がかき消されてしまうという現象である時間マスキングに対応する情報や、或る音が発生している際に他の音がかき消されてしまう現象であるスペクトルマスキングに対応する情報を用いるようにしてもよい。

【0096】

また、上記実施形態では、音声処理部による音声処理を、ADPCMとしていたが、本発明はこれに限られず、例えば、他のDPCMや他の方式（例えば、MP3）のコーデックであってもよい。音声処理部による音声処理がMP3である場合、例えば、音声処理におけるカットする音の周波数（パラメータ）を変えた場合のそれぞれにおいて、聴覚特性（例えば、等ラウドネス曲線とマスキングに対応する情報とのいずれか1以上に基づく聴覚特性）に応じた重みを用いて特徴量FDDを算出し、特徴量FDDを閾値以下とすることができ、且つ圧縮率を高くすることのできるような（例えば、カットする音の周波数を多くすることができるような）パラメータを選択するようにすればよい。このようにすると、MP3では、聴覚特性に応じて、音質を一定以上に維持しつつ、音質に影響のない、又は少ない音のカットにより圧縮率を高くすることができる。

【0097】

また、上記実施形態において、制御回路101が行っていた処理の一部又は全部を、ハードウェア回路で行うようにしてもよい。例えば、図1に示す各機能部11～15の少なくともいずれか1つをハードウェア回路で構成してもよい。また、音声処理部11の各構成の少なくとも一部をハードウェア回路で構成してもよい。

10

20

30

40

50

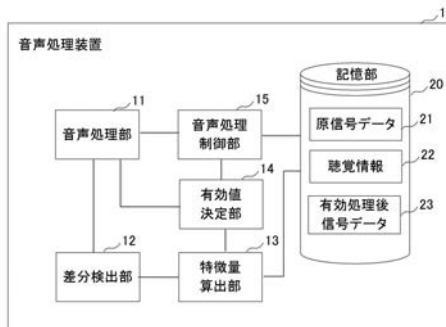
【符号の説明】

【0098】

1 ... 音声処理装置、 11 ... 音声処理部、 12 ... 差分検出部、 13 ... 特徴量算出部、 14 ... 有効値決定部、 15 ... 音声処理制御部、 20 ... 記憶部、 21 ... 原信号データ、 22 ... 聴覚情報、 23 ... 有効処理後信号データ

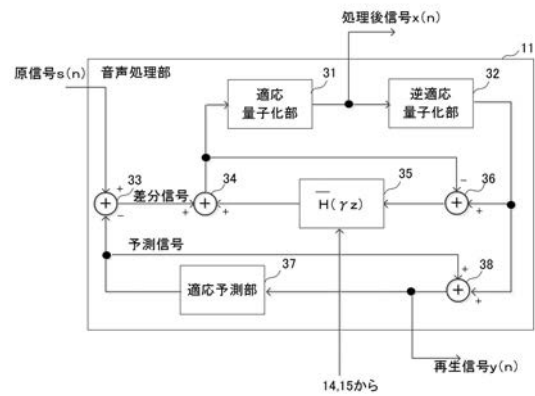
【図1】

図1



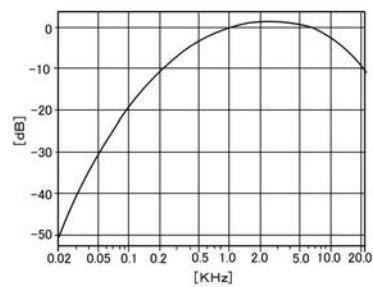
【図2】

図2



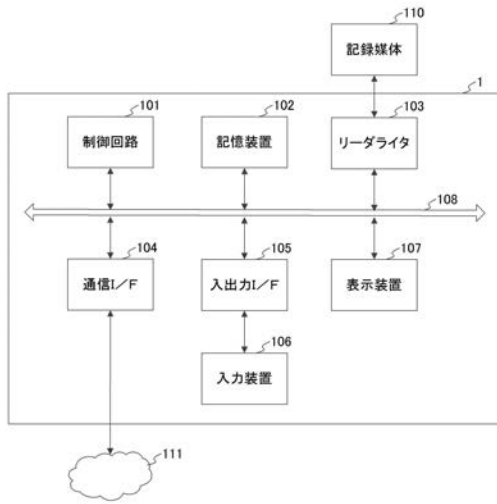
【図3】

図3



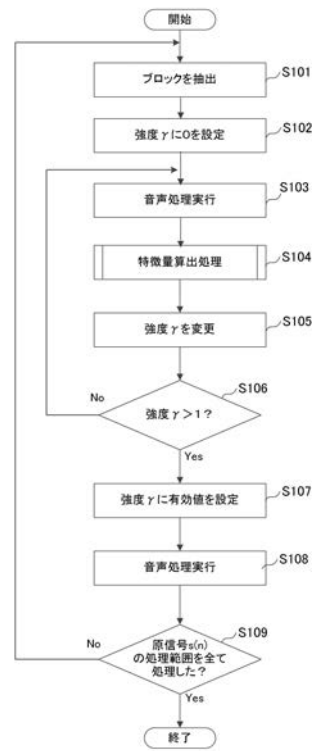
【 図 4 】

図4



【 図 5 】

図5



【 図 6 】

図6

