

(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2012-163682

(P2012-163682A)

(43) 公開日 平成24年8月30日(2012.8.30)

(51) Int. Cl.		F I			テーマコード (参考)	
G 1 O L	21/02	(2006.01)	G 1 O L	21/02	3 O 1 B	5 K 1 2 7
G 1 O L	11/00	(2006.01)	G 1 O L	11/00	1 O 1 E	
H O 4 M	1/00	(2006.01)	H O 4 M	1/00	H	

審査請求 未請求 請求項の数 6 O L (全 12 頁)

(21) 出願番号	特願2011-22715 (P2011-22715)	(71) 出願人	310006855
(22) 出願日	平成23年2月4日 (2011.2.4)		NECカシオモバイルコミュニケーションズ株式会社
			神奈川県川崎市中原区下沼部1753番地
		(74) 代理人	100123788
			弁理士 宮崎 昭夫
		(74) 代理人	100106138
			弁理士 石橋 政幸
		(74) 代理人	100127454
			弁理士 緒方 雅昭
		(72) 発明者	嶋田 修
			神奈川県川崎市中原区下沼部1753番地
			NECカシオモバイルコミュニケーションズ株式会社内

最終頁に続く

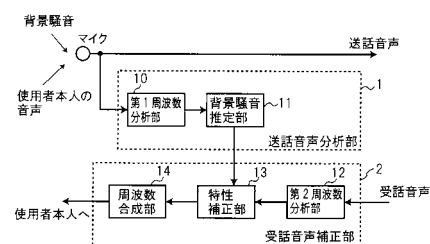
(54) 【発明の名称】 音声処理装置及び方法

(57) 【要約】

【課題】 使用者本人の送話中であっても、受話音声の明瞭度を向上させることができる音声処理装置及び方法を提供する。

【解決手段】 マイクから入力された音声信号を所定の周波数帯域の複数の音声信号に分割し、該周波数帯域毎の音声信号に対して信号対雑音比に応じた重みづけを行い、該音声信号に含まれる、分割された周波数帯域毎の背景騒音量を推定する。そして、受話音声信号を所定の周波数帯域毎の複数の音声信号に分割し、推定された周波数帯域毎の背景騒音量に基づいて、背景騒音量が大きいほど周波数帯域毎の音声信号が大きくなるように、該周波数帯域毎の音声信号に適用するゲインを設定し、該ゲインを周波数帯域毎の音声信号に乗算することで受話音声信号を補正する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

マイクから入力された音声信号を所定の周波数帯域の複数の第 1 音声信号に分割する第 1 周波数分析部と、

前記第 1 周波数分析部で分割された周波数帯域毎の第 1 音声信号に対して信号対雑音比に応じた重みづけを行い、前記音声信号に含まれる前記周波数帯域毎の背景騒音量を推定する背景騒音推定部と、

受話音声信号を所定の周波数帯域毎の複数の第 2 音声信号に分割する第 2 周波数分析部と、

前記背景騒音推定部で推定された前記周波数帯域毎の背景騒音量に基づいて、前記背景騒音量が大きいほど前記第 2 周波数分析部から出力された第 2 音声信号が大きくなるように、前記周波数帯域毎の第 2 音声信号に適用するゲインを設定し、該ゲインを対応する前記周波数帯域の第 2 音声信号に乗算することで前記受話音声信号を補正する特性補正部と、

10

前記特性補正部から出力された補正後の前記周波数帯域毎の第 2 音声信号を周波数合成し、補正後の前記受話音声信号を再生する周波数合成部を備えた受話音声補正部と、を有する音声処理装置。

【請求項 2】

前記特性補正部は、

前記周波数帯域毎のゲインを所定の上限値以下に制限するゲインリミッタ部を有する請求項 1 記載の音声処理装置。

20

【請求項 3】

前記特性補正部は、

背景雑音内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の背景雑音量を算出する背景雑音マスキング算出部と、

受話音声内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の受話音声量を算出する受話音声マスキング算出部と、を有し、

前記ゲイン生成部は、

前記背景雑音マスキング算出部から出力された背景雑音量と、前記受話音声マスキング算出部から出力された受話音声量とを周波数帯域毎に比較し、前記受話音声量よりも前記背景雑音量が大きい場合に、対応する周波数帯域の前記第 2 音声信号に適用するゲインを 1 以上に設定する請求項 1 または 2 記載の音声処理装置。

30

【請求項 4】

マイクから入力された音声信号を所定の周波数帯域の複数の第 1 音声信号に分割し、前記周波数帯域毎の第 1 音声信号に対して信号対雑音比に応じた重みづけを行い、前記音声信号に含まれる前記分割された周波数帯域毎の背景騒音量を推定し、

受話音声信号を所定の周波数帯域毎の複数の第 2 音声信号に分割し、前記推定された前記周波数帯域毎の背景騒音量に基づいて、前記背景騒音量が大きいほど対応する第 2 音声信号が大きくなるように、前記周波数帯域毎の第 2 音声信号に適用するゲインを設定し、

40

該ゲインを対応する前記周波数帯域の第 2 音声信号に乗算することで前記受話音声信号を補正し、

該補正後の前記周波数帯域毎の第 2 音声信号を周波数合成し、補正後の受話音声信号を再生する音声処理方法。

【請求項 5】

前記周波数帯域毎のゲインを所定の上限値以下に制限する請求項 4 記載の音声処理方法。

【請求項 6】

背景雑音内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の背景雑音量を算出し、

50

受話音声内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の受話音声量を算出し、

該背景雑音量と該受話音声量とを周波数帯域毎に比較し、該受話音声量よりも該背景雑音量が大きい場合に、対応する周波数帯域の第2音声信号に適用するゲインを1以上に設定する請求項4または5記載の音声処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は背景騒音に応じて音声信号を補正するための音声処理装置及び方法に関する。

【背景技術】

【0002】

例えば、携帯電話機は、屋外等で使用することが多いため、周囲の交通騒音や広告騒音等の背景騒音により受話音声聞き難くなることがしばしば起こる。これは背景騒音によって受話音声の一部の周波数成分がマスキングされることで、受話音声の明瞭度が低下することによる。そのため、騒音下でも受話音声聞き取り易いように、背景騒音に応じて音声信号（受話音声）を補正する様々な技術が提案されている。

【0003】

例えば、特許文献1では、受話音声の音声スペクトルを目標スペクトル（例えば子音等の比較的信号レベルが小さい周波数成分を大きくするような音声スペクトル）へ整形することで、受話音声の明瞭度を向上させる技術が記載されている。さらに、特許文献1では、使用者本人の発話による送話音声背景騒音と判定して受話音声を極端に増幅するのを防止するため、マイクから入力された音声信号の音声スペクトルを解析して該音声信号が使用者本人の送話音声であるか否かを判定し、使用者本人の送話音声であると判定した場合に以下の（1）～（3）で示す3種類の処理を実行することが記載されている。

（1）使用者本人の送話音声と判定した期間において、上記受話音声の音声スペクトルを目標スペクトルに整形するためのフィルタ部のフィルタ係数を初期値（例えば、受話音声をそのまま出力させるための値）に設定する（以下、第1背景処理と称す）。

（2）使用者本人の送話音声と判定した期間において、上記フィルタ係数を予め設定した最大値以下に抑制する（以下、第2背景処理と称す）。

（3）使用者本人の送話音声と判定した期間では上記フィルタ係数の更新を停止する。すなわち、使用者本人の送話音声と判定する直前のフィルタ係数を用いる（以下、第3背景処理と称す）。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2004-061617号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

上述した特許文献1に記載された技術では、使用者本人が発話中であるときに受話音声の明瞭度を向上させることができない問題がある。

【0006】

例えば、第1背景処理では、使用者本人が発話中は、上記フィルタ係数を、例えば初期値で固定することで、受話音声補正されないことになる。そのため、使用者本人が発話中は受話音声の明瞭度を向上させることができない。

【0007】

第2背景処理では、使用者本人が発話中であっても受話音声補正されるが、上記フィルタ係数が最大値以下に制限されることで、受話音声を必ずしも目標スペクトルに整形できるとは限らない。そのため、使用者本人が発話中は受話音声の明瞭度を向上させることができない場合がある。

10

20

30

40

50

【0008】

また、第3背景処理では、使用者本人の送話音声と判定される直前の状態（上記フィルタ係数の値）を維持するため、使用者本人が発話中は背景騒音の変化に追従して受話音声を補正することができない。そのため、第2背景処理と同様に、使用者本人が発話中は受話音声の明瞭度を向上させることができない場合がある。

【0009】

本発明は上述したような背景技術が有する問題点を解決するためになされたものであり、使用者本人の送話中であっても、受話音声の明瞭度を向上させることができる音声処理装置及び方法を提供することを目的とする。

【課題を解決するための手段】

【0010】

上記目的を達成するため本発明の音声処理装置は、マイクから入力された音声信号を所定の周波数帯域の複数の第1音声信号に分割する第1周波数分析部と、

前記第1周波数分析部で分割された周波数帯域毎の第1音声信号に対して信号対雑音比に応じた重みづけを行い、前記音声信号に含まれる前記周波数帯域毎の背景騒音量を推定する背景騒音推定部と、

受話音声信号を所定の周波数帯域毎の複数の第2音声信号に分割する第2周波数分析部と、

前記背景騒音推定部で推定された前記周波数帯域毎の背景騒音量に基づいて、前記背景騒音量が大きいほど前記第2周波数分析部から出力された第2音声信号が大きくなるように、前記周波数帯域毎の第2音声信号に適用するゲインを設定し、該ゲインを対応する前記周波数帯域の第2音声信号に乗算することで前記受話音声信号を補正する特性補正部と、

前記特性補正部から出力された補正後の前記周波数帯域毎の第2音声信号を周波数合成し、補正後の前記受話音声信号を再生する周波数合成部を備えた受話音声補正部と、を有する。

【0011】

一方、本発明の音声処理方法は、マイクから入力された音声信号を所定の周波数帯域の複数の第1音声信号に分割し、前記周波数帯域毎の第1音声信号に対して信号対雑音比に応じた重みづけを行い、前記音声信号に含まれる前記分割された周波数帯域毎の背景騒音量を推定し、

受話音声信号を所定の周波数帯域毎の複数の第2音声信号に分割し、前記推定された前記周波数帯域毎の背景騒音量に基づいて、前記背景騒音量が大きいほど対応する第2音声信号が大きくなるように、前記周波数帯域毎の第2音声信号に適用するゲインを設定し、

該ゲインを対応する前記周波数帯域の第2音声信号に乗算することで前記受話音声信号を補正し、

該補正後の前記周波数帯域毎の第2音声信号を周波数合成し、補正後の受話音声信号を再生する方法である。

【発明の効果】

【0012】

本発明によれば、使用者本人の送話中であっても、受話音声の明瞭度を向上させることができる。

【図面の簡単な説明】

【0013】

【図1】第1の実施の形態の音声処理装置の一構成例を示すブロック図である。

【図2】図1に示した背景騒音推定部の一構成例を示すブロック図である。

【図3】第1の実施の形態の特性補正部の一構成例を示すブロック図である。

【図4】第2の実施の形態の音声処理装置の一構成例を示すブロック図である。

【図5】第2の実施の形態の特性補正部の一構成例を示すブロック図である。

【図6】図5に示したゲインリミッタ部の処理例を示す模式図である。

10

20

30

40

50

【図 7】第 3 の実施の形態の特性補正部の一構成例を示すブロック図である。

【発明を実施するための形態】

【0014】

次に本発明について図面を用いて説明する。

(第 1 の実施の形態)

図 1 は、第 1 の実施の形態の音声処理装置の一構成例を示すブロック図である。

【0015】

図 1 に示すように、第 1 の実施の形態の音声処理装置は、送話音声分析部 1 と受話音声補正部 2 とを有する構成である。

【0016】

送話音声分析部 1 は、第 1 周波数分析部 10 及び背景騒音推定部 11 を備えている。受話音声補正部 2 は、第 2 周波数分析部 12、特性補正部 13 及び周波数合成部 14 を備えている。

【0017】

図 1 に示す音声処理装置は、例えば、マイクから入力された音声信号（アナログ信号）をデジタル信号に変換する A/D 変換器、受話音声補正部 2 から出力される音声信号（デジタル信号）をアナログ信号に変換する D/A 変換器、プログラムにしたがって音声信号に対する処理を実行する CPU、演算処理を実行する DSP、プログラムや処理に必要な各種データが格納されるメモリ、各種の論理回路等を含む、周知の信号処理回路で実現できる。

【0018】

第 1 周波数分析部 10 は、DFT (Discrete Fourier Transform: 離散フーリエ変換) 等の処理によりマイクから入力された音声信号を複数の周波数帯域の音声信号（第 1 音声信号）に分割する。第 1 周波数分析部 10 は、IIR (Infinite Impulse Response: 無限インパルス応答) フィルタ等の帯域分割フィルタで実現してもよい。また、第 1 周波数分析部 10 は、マイクから入力された音声信号を一定の帯域幅で分割してもよく、人の聴覚特性を考慮して、例えば高い周波数ほど帯域幅が広くなるように分割してもよい。

【0019】

背景騒音推定部 11 は、マイクから入力された音声信号に含まれる、第 1 周波数分析部 10 で分割された周波数帯域毎の背景騒音量を推定する。このとき、音声信号に使用者本人の送話音声の信号（以下、送話音声信号と称す）が含まれている場合は、該送話音声信号をできるだけ除いた残りの背景騒音量を算出する。

【0020】

図 2 は、図 1 に示した背景騒音推定部の一構成例を示すブロック図である。

【0021】

図 2 に示すように、背景騒音推定部 11 は重みつき音声計算部 100 及び推定雑音計算部 101 を備えている。

【0022】

重みつき音声計算部 100 は、背景騒音や使用者本人の送話音声を含むマイクから入力された周波数帯域毎の音声信号と推定雑音計算部 101 で推定された周波数帯域毎の背景騒音量とを用いて周波数帯域毎の SNR (Signal-to-Noise Ratio: 信号対雑音比) を算出し、マイクから入力された周波数帯域毎の音声信号に対して該 SNR に応じた重みづけを行う。

【0023】

推定雑音計算部 101 は、マイクから入力された周波数帯域毎の音声信号及び重みつき音声計算部 100 から出力された周波数帯域毎の重みつき音声信号から周波数帯域毎の背景騒音量を推定する。推定雑音計算部 101 は、重みつき音声計算部 100 から出力された周波数帯域毎の重みつき音声信号に基づき、背景騒音と推定された所定数のフレームの音声信号から周波数帯域毎の背景騒音量を算出しつつ、その値を更新する。

【0024】

10

20

30

40

50

図2に示す背景騒音推定部11による背景騒音量の具体的な算出方法は、例えば特開2008-216721号公報に記載されている。

【0025】

このように、推定された背景騒音量の更新に用いる、マイクから入力された周波数帯域毎の音声信号に対してSNRに応じた重みづけを行い、背景騒音量を推定することで、マイクから入力された音声信号に含まれる使用者本人の送話音声の影響を低減できるため、背景騒音量をより精度よく推定できる。

【0026】

第2周波数分析部12は、DFT等の処理により受話音声の信号（以下、受話音声信号と称す）を、第1周波数分析部10と同様に複数の周波数帯域の音声信号（第2音声信号）に分割する。

10

【0027】

特性補正部13は、例えば図3に示す構成で実現できる。

【0028】

図3は、第1の実施の形態の特性補正部の一構成例を示すブロック図である。

【0029】

図3に示すように、特性補正部13は、平滑化部200、ゲイン生成部201、ゲインMatrix部202及び補正部203を備えている。

【0030】

平滑化部200は、背景騒音推定部11で推定された周波数帯域毎の背景雑音量を時間軸方向または周波数軸方向で平滑化する。平滑化部200は、各周波数帯域の背景雑音量を時間軸方向と周波数軸方向の両方でそれぞれ平滑化してもよい。例えば音声処理装置が所定のフレーム単位で音声信号に対する処理を実行する場合、時間軸方向に隣接する複数のフレーム毎の背景雑音量を平滑化すればよい。または、周波数軸方向に隣接する複数のフレーム毎に背景雑音量を平滑化すればよい。あるいは、時間軸方向及び周波数軸方向に隣接する複数のフレーム毎の背景雑音量を平滑化すればよい。平滑化部200を備えることで、背景騒音量が急激に変化した場合でも、受話音声が不自然に変化するのを抑制できる。

20

【0031】

ゲイン生成部201は、平滑化部200で平滑化された各周波数帯域の背景騒音量から周波数帯域毎の受話音声信号に適用する増幅率（ゲイン）をそれぞれ算出する。ゲイン生成部201は、例えば背景雑音量が大きいほどゲインが大きくなるように、周波数帯域毎の受話音声信号のゲインをそれぞれ設定すればよい。ゲイン生成部201は、周波数帯域毎の受話音声信号のゲインを、例えば、背景雑音量の一次式で求めてもよく、二次以上の高次の方程式で求めてもよい。

30

【0032】

ゲインMatrix部202は、例えば下記式(1)を用いて、ゲイン生成部201で算出された各周波数帯域のゲインを混合して平滑化する。式(1)は、受話音声信号をN（Nは正数）個の周波数帯域（周波数が低い方から f_1 , f_2 , ..., f_N ）に分割したときのゲインMatrix部202による計算式例を示している。式(1)の左辺がゲインMatrix部202による処理後の周波数帯域（ f_1 , f_2 , ..., f_N ）毎のゲインとなる。

40

【0033】

【数 1】

$$\begin{bmatrix} \text{処理後のゲイン}_{f1} \\ \text{処理後のゲイン}_{f2} \\ \vdots \\ \text{処理後のゲイン}_{fN} \end{bmatrix} = \begin{bmatrix} \text{混合係数}_{11} & \text{混合係数}_{12} & \cdots & \text{混合係数}_{1N} \\ \text{混合係数}_{21} & \text{混合係数}_{22} & \cdots & \text{混合係数}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \text{混合係数}_{N1} & \text{混合係数}_{N2} & \cdots & \text{混合係数}_{NN} \end{bmatrix} \cdot \begin{bmatrix} \text{ゲイン}_{f1} \\ \text{ゲイン}_{f2} \\ \vdots \\ \text{ゲイン}_{fN} \end{bmatrix}$$

・・・(1)

【0034】

10

ゲインMatrix部202は、ゲイン生成部201で算出された周波数帯域毎のゲインを周波数軸方向に平滑化する。混合係数は、基本的に隣接する周波数帯域ほど値が大きくなるように設定する。各混合係数は、予め設定しておいてもよく、ゲイン生成部201で算出された各ゲインの分布状況から動的に決定してもよい。例えば、各ゲインの分散が大きい場合は、より多くの周波数帯域のゲインを用いて平滑化すればよい。このようなゲインMatrix部202による処理を行うことで、周波数合成部14による周波数合成後の受話音声の不自然になるのを抑制できる。なお、混合係数は、全ての値を設定する必要はなく、一部の混合係数を「0」とすることも可能である。その場合、「0」に設定した混合係数が多いほど、演算量を低減することが可能であり、混合係数を保持するメモリを節約できる。但し、混合係数は、ゲインMatrix部202による処理後の各周波数帯域のゲインが「0」とならないように設定する必要がある。例えば、各周波数帯域に対応する混合係数を「1」とし、他の混合係数を「0」とすることで、ゲイン生成部201で算出された各周波数帯域のゲインをそのまま処理後のゲインとして出力させることが考えられる。

20

【0035】

補正部203は、ゲインMatrix部202から出力された処理後の周波数帯域毎のゲインを、対応する周波数帯域の音声信号に乗算して出力する。

【0036】

周波数合成部14は、第2周波数分析部12による処理（DFT等）の逆変換を実行することで、特性補正部13から出力された周波数帯域毎の音声信号を周波数合成し、受話音声信号（補正後）を再生する。

30

【0037】

第1の実施の形態の音声処理装置によれば、マイクから入力された音声信号に対して周波数帯域毎のSNRに応じた重みづけを行い、背景騒音量を推定することで、マイクから入力された音声信号に含まれる使用者本人の送話音声の影響を低減できるため、背景騒音量をより精度よく推定できる。そのため、推定される背景騒音量では、使用者本人の送話音声の影響が低減されているため、該背景騒音量に基づいて受話音声を補正すれば、使用者本人による送話中であっても受話音声を補正することが可能であり、受話音声の明瞭度を向上させることができる。

(第2の実施の形態)

40

図4は、第2の実施の形態の音声処理装置の一構成例を示すブロック図であり、図5は、第2の実施の形態の特性補正部の一構成例を示すブロック図である。

【0038】

図4に示すように、第2の実施の形態の音声処理装置は、第1の実施の形態で示した特性補正部13に代えて特性補正部15を備えた構成である。図5に示すように、第2の実施の形態の特性補正部15は、第1の実施の形態で示した特性補正部13にゲインリミッタ部204を追加した構成である。音声処理装置のその他の構成及び動作は、第1の実施の形態の音声処理装置と同様であるため、その説明は省略する。

【0039】

第1の実施の形態で示した特性補正部13では、ゲイン生成部201で算出されたゲイ

50

ンをそのまま用いているため、周波数合成部 1 4 による周波数合成後の受話音声信号でデジタルクリップが発生する可能性がある。ゲインリミッタ部 2 0 4 は、ゲイン生成部 2 0 1 で算出された周波数帯域毎のゲインを抑制し、周波数合成後の受話音声信号でデジタルクリップが発生するのを抑制する。

【0040】

図 6 は、図 5 に示したゲインリミッタ部の処理例を示す模式図である。図 6 の横軸は時間軸（フレーム）を示し、縦軸はゲインリミッタ部 2 0 4 で設定する周波数帯域毎のゲインの上限値（ゲインリミット値）を示している。

【0041】

ゲインリミッタ部 2 0 4 は、まず受話音声信号の振幅（デジタル値）から、音声信号を処理するフレーム毎に、許容される最大ゲインの値（実線）を算出する。最大ゲインの値は、デジタル信号処理で扱える最大値を、フレーム内の受話音声信号の最大値（絶対値）で除算することで求める。最大ゲインの値は、受話音声信号の振幅が小さいほど大きくなり、受話音声信号の振幅が大きいほど小さくなる。

【0042】

ゲインリミッタ部 2 0 4 は、ゲインリミッタ値（点線）を算出した最大ゲインの値以下となるように設定する。また、ゲインリミッタ部 2 0 4 は、最大ゲインの値が急激に上昇した場合、予め設定されたフレーム数（H o l d フレーム数）の期間はゲインリミッタ値を変更しない。例えば、図 6 に示す時点 T 1 以降では、最大ゲインの値が上昇するため、周波数帯域毎のゲインも大きくできる。しかしながら、ゲインリミッタ部 2 0 4 は、上記 H o l d フレーム数の期間はゲインリミッタ値を上昇させない。これは、ゲインリミッタ値を急激に上昇させて周波数帯域毎のゲインも急激に大きくすると、周波数合成後の受話音声の音量が不自然に大きくなるため、そのような不自然な音量増大を防止するためである。

【0043】

ゲインリミッタ部 2 0 4 は、上記 H o l d フレーム数の期間が経過した後（図 6 の時点 T 2）、徐々にゲインリミッタ値を上昇させる。このとき、ゲインリミッタ部 2 0 4 は、ゲインリミッタ値を予め設定した割合で上昇させてもよく、背景雑音量や最大ゲインの値に応じて上昇割合を変化させてもよい。

【0044】

ゲインリミッタ部 2 0 4 は、上昇させたゲインリミッタ値が最大ゲインの値と等しくなった場合（図 6 の時点 T 3）、その後、最大ゲインの値が上昇する場合は、上述した T 1 以降と同様の処理を実行する。また、最大ゲインの値が下降する場合は、ゲインリミッタ値を最大ゲインの値に合わせて下降させる。

【0045】

第 2 の実施の形態の音声処理装置によれば、ゲインリミッタ部 2 0 4 により受話音声信号の振幅に応じて受話音声の各周波数帯域のゲインを上限値以下に制限するため、補正後の受話音声信号がデジタル信号処理で扱える最大値を越えることがない。そのため、第 1 の実施の形態の音声処理装置と比べて、音質の劣化を招くことなく、受話音声を明瞭化できる。

（第 3 の実施の形態）

図 7 は、第 3 の実施の形態の特性補正部の一構成例を示すブロック図である。

【0046】

図 7 に示すように、第 3 の実施の形態の音声処理装置は、第 1 の実施の形態で示した特性補正部 1 3 に、背景雑音量マスキング算出部 2 0 5 及び受話音声マスキング算出部 2 0 6 を追加した構成である。音声処理装置のその他の構成及び動作は、第 1 の実施の形態の音声処理装置と同様であるため、その説明は省略する。

【0047】

本実施形態の特性補正部は、背景騒音推定部 1 1 で推定された背景雑音量から背景雑音によってマスキングされる受話音声信号の周波数帯域のゲインのみを増大させる。

【0048】

背景雑音量マスキング算出部205は、平滑化部200で平滑化された背景雑音量から分割した周波数帯域毎に周知のマスキング閾値を算出し、該マスキング閾値を用いて人が聴取可能な周波数帯域毎の背景雑音量を算出する。

【0049】

受話音声マスキング算出部206は、受話音声信号から分割した周波数帯域毎に周知のマスキング閾値を算出し、該マスキング閾値を用いて人が聴取可能な周波数帯域毎の受話音声量を算出する。

【0050】

マスキング閾値は、ある所望音を聞き取る際、それをマスキングする他の音がある場合に、該所望音が聞き取れる限界の音圧レベルを指す。通常、受話音声や背景騒音には様々な周波数成分を含んでいるため、受話音声であっても、該受話音声のある周波数成分が受話音声の他の周波数成分をマスキングすることがある。背景雑音量マスキング算出部205は、周波数帯域毎にマスキング閾値と背景雑音量とを比較し、背景雑音内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の背景雑音量を算出する。同様に、受話音声マスキング算出部206は、周波数帯域毎にマスキング閾値と受話音声量とを比較し、受話音声量内の他の周波数成分でマスキングされない、人が聴取可能な周波数帯域毎の受話音声量を算出する。なお、マスキング閾値の算出方法やマスキング閾値を用いた音声信号の補正方法等については、例えば特開2009-175420に開示されている。

【0051】

ゲイン生成部207は、背景雑音量マスキング算出部205から出力された背景雑音量と、受話音声マスキング算出部206から出力された受話音声量とを周波数帯域毎に比較し、受話音声量よりも背景雑音量が大きい場合に、対応する周波数帯域の音声信号（第2音声信号）のゲインを「1」以上に設定する。

【0052】

ここで、受話音声量よりも背景雑音量が大きい周波数帯域のゲインを背景雑音量／受話音声量に設定すれば、該周波数帯域における背景雑音量と補正処理後の受話音声量とは同等となる。

【0053】

受話音声量よりも背景雑音量が大きい周波数帯域において、補正処理後の受話音声量を背景雑音量よりも常に大きくしたい場合、ゲイン生成部207は、該周波数帯域の音声信号（受話音声）のゲインを、（背景雑音量／受話音声量）以上の値に設定すればよい。例えば、対応する周波数帯域の音声信号（受話音声）のゲインを、（背景雑音量／受話音声量）× α （ $\alpha > 1.0$ ）、あるいは（背景雑音量／受話音声量）+ α （ $\alpha > 0.0$ ）に設定すればよい。

【0054】

ゲイン生成部207は、背景雑音量よりも受話音声量が大きい場合、対応する周波数帯域の音声信号（受話音声）のゲインを「1」に設定し、増幅させないようにする。

【0055】

第3の実施の形態の音声処理装置によれば、受話音声量よりも背景雑音量が大きい周波数帯域でのみ受話音声信号が増幅されるため、受話音声が不必要に増大することがない。そのため、第1の実施の形態の音声処理装置よりも高品質な受話音声再生される。なお、図7に示した背景雑音量マスキング算出部205及び受話音声マスキング算出部206は、第2の実施の形態の特性補正部15に備えることも可能である。

【符号の説明】

【0056】

- 1 送話音声分析部
- 2 受話音声補正部
- 10 第1周波数分析部

10

20

30

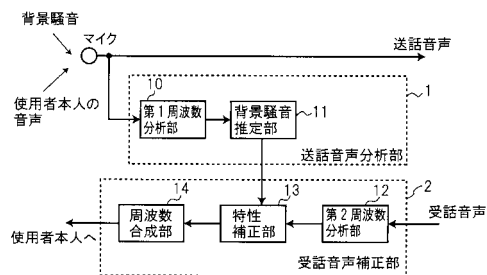
40

50

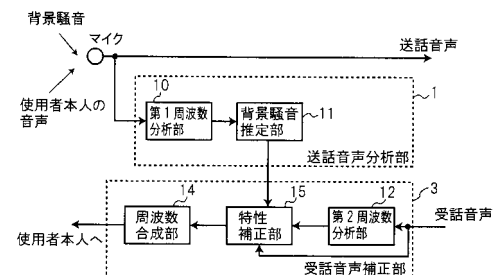
- 1 1 背景騒音推定部
- 1 2 第2周波数分析部
- 1 3、1 5 特性補正部
- 1 4 周波数合成部
- 1 0 0 重みつき音声計算部
- 1 0 1 推定雑音計算部
- 2 0 0 平滑化部
- 2 0 1 ゲイン生成部
- 2 0 2 ゲインMatrix部
- 2 0 3 補正部
- 2 0 4 ゲインリミッタ部
- 2 0 5 背景雑音量マスキング算出部
- 2 0 6 受話音声マスキング算出部

10

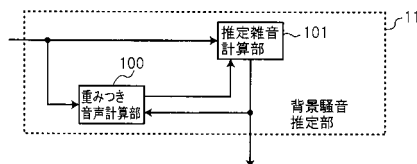
【図1】



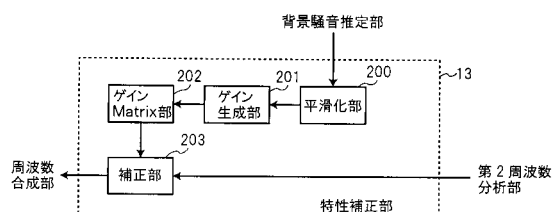
【図4】



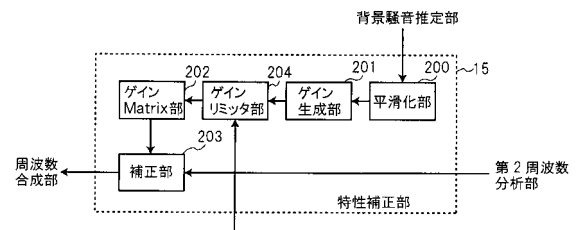
【図2】



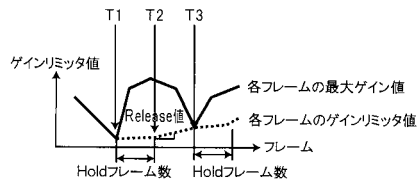
【図3】



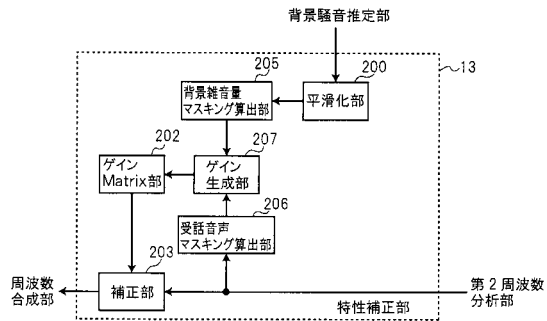
【図5】



【図 6】



【図 7】



フロントページの続き

- (72)発明者 細川 知志
神奈川県川崎市中原区下沼部1 7 5 3 番地 NECカシオモバイルコミュニケーションズ株式会社
内
- (72)発明者 黒田 淳
神奈川県川崎市中原区下沼部1 7 5 3 番地 NECカシオモバイルコミュニケーションズ株式会社
内
- (72)発明者 岸波 雄一郎
神奈川県川崎市中原区下沼部1 7 5 3 番地 NECカシオモバイルコミュニケーションズ株式会社
内
- (72)発明者 上野 寿之
神奈川県川崎市中原区下沼部1 7 5 3 番地 NECカシオモバイルコミュニケーションズ株式会社
内
- (72)発明者 遠藤 次夫
神奈川県川崎市中原区下沼部1 7 5 3 番地 NECカシオモバイルコミュニケーションズ株式会社
内

Fターム(参考) 5K127 AA03 BA03 CA25 HA11 HA24 JA02 JA06 MA05 MA08 MA35