

(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2010-85913

(P2010-85913A)

(43) 公開日 平成22年4月15日(2010.4.15)

(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 A
H O 4 R	3/04	(2006.01)	H O 4 R	3/04	5 D O 2 O

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2008-257471 (P2008-257471)	(71) 出願人	000003078
(22) 出願日	平成20年10月2日 (2008.10.2)		株式会社東芝
			東京都港区芝浦一丁目1番1号
		(74) 代理人	100109900
			弁理士 堀口 浩
		(72) 発明者	長田 将高
			東京都港区芝浦一丁目1番1号 株式会社
			東芝内
		(72) 発明者	三関 公生
			東京都港区芝浦一丁目1番1号 株式会社
			東芝内
		Fターム(参考)	5D020 AC05 CE02

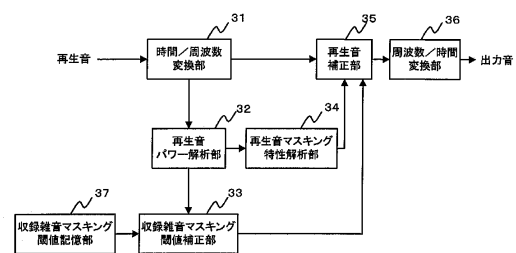
(54) 【発明の名称】 音補正装置

(57) 【要約】

【課題】 収録された雑音信号のマスクング閾値を用いて再生音の補正を行い、実際の周囲の雑音の大小にかかわらず聞き取りやすさを向上する音補正装置を提供する。

【解決手段】 再生音の単位区間のパワーの平均値と予め定めた閾値とを比較することによって、当該区間の再生音のレベルを複数の段階のいずれかに当てはめ、そのレベルの段階に応じた方法によって、予め収録された雑音信号のマスクング閾値を補正し、補正した収録雑音のマスクング閾値を用いて再生音の補正を行う。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

再生音の周波数成分に対する補正係数を算出し、再生音の補正を行う音補正装置であって、

予め収録された収録雑音のマスキング閾値を記憶する収録雑音マスキング閾値記憶手段と、

再生音の所定の区間のパワーの平均値を算出し、この再生音の区間ごとのパワーの平均値に応じて前記収録雑音マスキング閾値記憶手段から読み出した収録雑音のマスキング閾値を補正し、補正後の収録雑音マスキング閾値未満のパワーを有する再生音の周波数成分を増幅させるよう補正する補正手段を有することを特徴とする音補正装置。

10

【請求項 2】

前記補正手段は、前記収録雑音のマスキング閾値記憶手段から読み出した収録雑音のマスキング閾値を補正する際、前記再生音の所定の区間ごとのパワーの平均値が所定の閾値未満である場合は、収録雑音のマスキング閾値を前記所定の区間の再生音よりも大きくなるよう補正することを特徴とする請求項 1 に記載の音補正装置。

【請求項 3】

前記補正手段は、前記収録雑音のマスキング閾値記憶手段から読み出した収録雑音のマスキング閾値を補正する際、前記再生音の所定の区間ごとのパワーの平均値が第 1 の閾値未満である場合には、前記収録雑音のマスキング閾値を前記所定の区間の再生音よりも大きくなるよう補正し、第 1 の閾値よりも大きい第 2 の閾値未満である場合には、前記収録雑音のマスキング閾値を前記所定の区間の再生音のレベルと同じになるよう補正することを特徴とする請求項 1 に記載の音補正装置。

20

【請求項 4】

前記補正手段は、補正後の収録雑音マスキング閾値未満のパワーを有する再生音の周波数成分を増幅させるときに、前記再生音の所定の区間ごとのパワーの平均値が所定の閾値未満であった場合には、前記再生音の所定の区間ごとのパワーの平均値が所定の閾値以上であった場合よりも増幅率を高く設定することを特徴とする請求項 1 に記載の音補正装置。

【請求項 5】

ユーザからの再生音の音量増減操作を受ける操作手段を更に有し、

30

前記補正手段は、前記操作手段によって設定された再生音の音量増減設定値を用いて再生音の所定の区間ごとのパワーの平均値を補正し、補正後の再生音のパワーの平均値に応じて前記収録雑音マスキング閾値記憶手段から読み出した収録雑音のマスキング閾値を補正し、補正後の収録雑音マスキング閾値未満のパワーを有する再生音の周波数成分を増幅させるよう補正することを特徴とする請求項 1 に記載の音補正装置。

【請求項 6】

予め記憶されていた再生音または受信した再生音を補正する音補正装置であって、

収録雑音のマスキング閾値を記憶する収録雑音マスキング閾値記憶手段と、

再生音の所定の区間のパワーの平均値を算出し、この平均値が所定の閾値未満ならば、前記収録雑音マスキング閾値記憶手段から読み出される収録雑音のマスキング閾値を増幅させ、増幅後の収録雑音のマスキング閾値に基づいて前記再生音を補正し、再生音の所定の区間のパワーの平均値が所定の閾値以上ならば前記収録雑音マスキング閾値記憶手段から読み出された収録雑音マスキング閾値に基づいて前記再生音を補正する補正手段を有することを特徴とする音補正装置。

40

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は音補正装置に関する。

【背景技術】**【0002】**

50

テレビやラジオの放送受信再生装置、音楽プレイヤー、携帯電話機などの音声・音楽を再生する機器は、電車の中や屋外や車の中など周囲に雑音がある場所で使用される場合がある。この場合、機器によって再生する音（以降、再生音と称する）と収録雑音との周波数やパワーの関係によっては、再生音が収録雑音によってマスクされ、音の明瞭度が低下する場合がある。多くの再生機器は再生音量をユーザの操作によって調整することができるが、再生音の周波数成分ごとに音量調整ができるわけではないため、音量を上げたとしても音の明瞭度が向上するとは限らない。また、再生音量を上げた場合には、再生音の全帯域のパワーが増幅されるため、音が歪んでしまい、かえって音質が悪化することもある。更に、音量を上げすぎると、聴覚に対してダメージを与えるという問題が起こる可能性がある。

10

【0003】

そこで、収録雑音のある環境下での音声通話において、サブバンド毎の騒音レベルを予め測定し、この騒音レベルによって決定したゲインに基づいて、受話音声信号に対するフィルタ処理を行うことにより、収録雑音によってマスクされていた音も聞き取れるレベルにまで増幅する受話音声処理装置が提案されている（例えば、特許文献1参照。）。

【特許文献1】特開2001-188599号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

特許文献1に記載される発明では、固定的な騒音信号の長時間周波数特性を利用してゲインを決定するため、補正の必要が無い音量の大きな区間の信号を補正したり、逆に補正が必要な音量が低レベルの区間の信号に対する補正が不足したりする場合があるという問題点がある。

20

【0005】

そこで本発明は、収録された雑音信号のマスキング閾値を用いて周囲の環境に適合した再生音の補正を行う音補正装置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】

【0006】

上記目的を達成するために、本発明による音補正装置は、再生音の周波数成分に対する補正係数を算出し、再生音の補正を行う音補正装置であって、予め収録された収録雑音のマスキング閾値を記憶する収録雑音マスキング閾値記憶手段と、再生音の所定の区間ごとのパワーの平均値を算出し、この再生音の区間ごとのパワーの平均値に応じて前記収録雑音マスキング閾値記憶手段から読み出した収録雑音のマスキング閾値を補正し、補正後の収録雑音マスキング閾値未満のパワーを有する再生音の周波数成分を増幅させるよう補正する補正手段を有することを特徴としている。

30

【発明の効果】

【0007】

本発明によれば、収録された雑音信号のマスキング閾値を用いて周囲の環境に適合した再生音の補正を行う音補正装置を提供することができる。

【発明を実施するための最良の形態】

40

【0008】

以下、本発明の一例である実施形態について図面を参照して説明する。

【実施例1】

【0009】

本発明の音補正装置は、携帯電話機、PC、ポータブルオーディオ機器などに実装される。ここでは、携帯電話機に実装した場合を例に説明する。

【0010】

図1は本発明にかかる携帯電話機の構成図である。この携帯電話機は、全体の制御を行う制御部11を含み、制御部11には、送受信部12、放送受信部13、信号処理部14、操作部15、記憶部16、表示部17、音声入出力部18が接続されている。

50

【 0 0 1 1 】

送受信部 1 2 は、図示しない基地局との間で情報の送受信を行う。送受信部 1 2 には、アンテナが接続されており、アンテナによって電波に変換した情報を基地局へ送信する送信機能と、基地局から電波を受信し、電気信号へ変換する受信機能を有する。

【 0 0 1 2 】

放送受信部 1 3 は、T V 放送受信用のアンテナが接続されている。放送受信部 1 3 は、T V 放送受信用のアンテナによって入力された電波のうち、選局された物理チャンネルの信号を取得する。

【 0 0 1 3 】

信号処理部 1 4 は、映像信号や音声信号、オーディオ信号などのデジタル信号を処理する。信号処理部 1 4 は、再生音の補正処理を行う補正処理部 3 0 を有しており、送受信部 1 2 によって受信した電話やテレビ電話などの通話音声や、放送受信部 1 3 によって受信したテレビ放送やラジオ放送の音データや、記憶部 1 6 に記憶されている音楽データなどを再生するときの再生音を明瞭化するように補正処理を行う。

【 0 0 1 4 】

操作部 1 5 は、入力キーなどによって構成され、ユーザからの操作入力手段として用いられる。記憶部 1 6 は、アプリケーションソフトウェアや、音楽データや映像データなどが格納される。表示部 1 7 は、液晶ディスプレイや有機 E L ディスプレイなどから成る。

【 0 0 1 5 】

表示部 1 7 は、携帯電話機の動作状態に合わせた画像を表示する。

【 0 0 1 6 】

音声入出力部 1 8 は、マイクロホンやスピーカから構成される。スピーカによって、T V 放送の音声や通話機能使用時の受話音声、着信時の鳴動音などを出力する。また、マイクロホンによって音声信号が携帯電話機へ入力される。

【 0 0 1 7 】

以下では、補正処理部 3 0 について説明する。

図 2 は、補正処理部 3 0 の詳細を示す構成図である。補正処理部 3 0 には、補正処理の対象となる再生音が所定の単位区間ごとに入力される。前述のとおり、再生音は、通信によって得た音でも良いし、記憶部 1 6 に記憶されている音でも良い。例えば、放送受信部 1 3 を介して受信したテレビ放送波やラジオ放送波から得られる音や、送受信部 1 2 を介して受信した音声通話の音声や、記憶部 1 6 に記憶されたオーディオなどが考えられる。

【 0 0 1 8 】

補正処理部 3 0 に入力された所定の単位区間ごとの再生音は、時間 / 周波数変換部 3 1 に入力される。

【 0 0 1 9 】

時間 / 周波数変換部 3 1 は、所定の単位区間の再生音を時間域の信号から周波数域の信号に変換する。時間域と周波数域との変換には、例えば、F F T (Fast Fourier Transform) や、M D C T (Modified Discrete Cosine Transform) などの手法を用いることができる。このように周波数域に変換された再生音は、再生音パワー解析部 3 2 と再生音補正部 3 5 に入力される。

【 0 0 2 0 】

再生音パワー解析部 3 2 では、時間 / 周波数変換部 3 1 から入力される周波数域の再生音パワー $\text{signal_power}[i]$ を算出する。ここで、 $\text{signal_power}[i]$ は、再生音の周波数成分の実部の二乗と虚部の二乗とを加算することによって算出することができる。さらに再生音パワー解析部 3 2 では、 $\text{signal_power}[i]$ の単位区間での平均値である frame_power を算出する。

【 0 0 2 1 】

再生音マスキング特性解析部 3 4 では、再生音パワー解析部 3 2 で算出された再生音の周波数域パワー $\text{signal_power}[i]$ を用いて再生音のマスキング閾値 $\text{signal_thr}[i]$ を算出する。マスキング閾値の算出方法としては、例えば $\text{spreading function}$ と呼ばれる関数を再

10

20

30

40

50

生音の周波数域パワーに畳み込むことで算出する方法がある。spreading functionは、ISO/IEC13818-7, ITU-R1387, 3GPP TS 26.403といった文献によって説明されている。

【0022】

一方、収録雑音マスキング閾値記憶部37には、雑音のマスキング閾値が記憶されている。雑音のマスキング閾値は、予め収録された長時間の雑音の有するマスキング閾値を算出したものである。収録雑音マスキング閾値記憶部37からは、収録雑音のマスキング閾値noise_thr[i]が読み出され、収録雑音マスキング閾値補正部33に入力される。なお、収録雑音マスキング閾値記憶部37に記憶される収録雑音のマスキング閾値は、1種類でも良いし、複数種類でも良い。様々な環境下で収録された複数種類の収録雑音のマスキング閾値を収録雑音マスキング閾値記憶部37に記憶させておくと、携帯電話機が様々な環境下で使用されるような場合であっても、周囲の雑音に合わせた環境の選択操作が操作部15からなされることによって、適切な収録雑音のマスキング閾値によって再生音の補正処理を行うことができる。以降の説明では、収録雑音マスキング閾値記憶部37に記憶されている収録雑音のマスキング閾値は、1種類であるとして説明する。

【0023】

収録雑音マスキング閾値補正部33では、後述の再生音パワー解析部32から出力される再生音の周波数域平均パワーframe_powerを用いて、収録雑音マスキング閾値記憶部37から読み出された収録雑音のマスキング閾値noise_thr[i]に対する補正処理を行う。なお、この補正処理は、後ほど詳細に説明する。

【0024】

再生音補正部35には、収録雑音マスキング閾値補正部33から収録雑音マスキング閾値が入力され、再生音パワー解析部32から再生音の周波数域パワーsignal_power[i]が入力され、再生音マスキング特性解析部34から再生音のマスキング特性signal_thr[i]が入力され、時間/周波数変換部31から再生音の周波数スペクトルが入力される。再生音補正部35は、これらの値を用いて、収録雑音のマスキング閾値によってマスクされている周波数帯域であって、再生音自身によってマスクされていない周波数帯域を増幅するよう再生音の補正を行う。再生音補正部35の詳細な処理については、後述する。

【0025】

周波数/時間変換部36は、再生音補正部35から出力された再生音の周波数スペクトルを時間域の信号に変換する。これによって、補正された再生音が得られる。この再生音は音声入出力部18のスピーカから出力することができる。

【0026】

次に、以上のような構成を有する補正処理部30の収録雑音マスキング閾値補正部33と再生音補正部35について詳細に説明する。

【0027】

収録雑音マスキング閾値補正部33は、再生音パワー解析部32から入力される再生音の単位区間ごとの周波数域平均パワーframe_powerと所定の閾値TH1およびTH2(ただしTH1<TH2)とを比較する。閾値TH1およびTH2は、予め設定された値であって、再生音の周波数域平均パワーframe_powerがTH1よりも小さい場合には、低レベルの再生音と判定でき、frame_powerがTH2よりも大きい場合には、再生音が十分に大きく、収録雑音にかかわらず十分に知覚できるような大レベルの再生音と判定できるように設定される。

【0028】

収録雑音マスキング閾値補正部33での周波数域平均パワーframe_powerと所定の閾値TH1およびTH2との比較結果としては、3つの場合がある。

【0029】

第1に、収録雑音マスキング閾値補正部33が、再生音の周波数域平均パワーframe_powerはTH1よりも小さい低レベルの再生音であるとの判定する場合がある。この場合は、当該区間の再生音は低レベルであるため、音量増加を含めてより明瞭度を向上させる必要がある。そこで、まず収録雑音のマスキング閾値noise_thr[i]の平均値が再生音の周波

10

20

30

40

50

数域パワーの平均値frame_powerよりも大きくなるよう補正する。図3は、再生音の周波数域平均パワーframe_powerがTH1よりも小さいときに、収録雑音マスキング閾値noise_thr[i]を補正するときの処理の概念図である。図3(a)のような収録雑音マスキング閾値が収録雑音マスキング閾値記憶部37から読み出されている状態で、図3(b)のような信号特性を有する再生音の単位区間が補正処理部30に入力されたとする。収録雑音マスキング閾値補正部33が、この再生音の周波数域平均パワーframe_powerはTH1よりも低いと判定すると、図3(c)のように、収録雑音のマスキング閾値の平均値を再生音の周波数域パワーの平均値frame_powerよりも大きくなるよう、収録雑音のマスキング閾値を補正する。そして、収録雑音マスキング閾値補正部33は、補正処理を施したマスキング閾値を出力する。

10

【0030】

第2に、収録雑音マスキング閾値補正部33が、再生音の周波数域平均パワーはTH1以上であり、TH2未満であると判定する場合がある。この場合は、中レベルの再生音の音量増加を抑制しつつ聴感上の明瞭度を向上させるために、収録雑音マスキング閾値を再生音パワーに合わせて正規化する。つまり、収録雑音のマスキング閾値noise_thr[i]の平均値が再生音の周波数域パワー平均値frame_powerと等しくなるよう補正する。図4は、再生音の周波数域平均パワーframe_powerがTH1以上であり、TH2未満であるときに、収録雑音マスキング閾値noise_thr[i]を補正するときの処理の概念図である。図4(a)のような収録雑音マスキング閾値が収録雑音マスキング閾値記憶部37から読み出されている状態で、図4(b)のような信号特性を有する再生音の単位区間が補正処理部30に入力されたとする。収録雑音マスキング閾値補正部33が、この再生音の周波数域平均パワーframe_powerはTH1以上であり、TH2未満であると判定すると、図4(c)のように、収録雑音のマスキング閾値の平均値を再生音の周波数域パワーの平均値frame_powerと等しくなるよう、全帯域に対して同じレベル分だけ増幅させる。そして、収録雑音マスキング閾値補正部33は、補正処理を施したマスキング閾値を出力する。

20

【0031】

第3に、収録雑音マスキング閾値補正部33が、再生音の周波数域平均パワーframe_powerはTH2よりも大きい高レベルの再生音であると判定する場合がある。この場合、再生音は、収録雑音にかかわらず十分に近くされるレベルであるため、収録雑音マスキング閾値noise_thr[i]に対してとくに補正処理を行わず、そのままnoise_thr[i]を出力する。

30

【0032】

図5は、この収録雑音マスキング閾値補正部33によって補正された後の収録雑音マスキング閾値パワーnoise_thr[i]の平均値と再生音の周波数域平均パワーframe_powerとの関係を示す図である。図5で表されるように、再生音の周波数域平均パワーframe_powerがTH1よりも小さい場合には、収録雑音マスキング閾値の平均値を再生音の周波数域平均パワーよりも大きくなるように設定している。一方、再生音の周波数域平均パワーがTH1以上TH2未満の場合には、収録雑音マスキング閾値の平均値を再生音の周波数域平均パワーと同じになるよう設定している。

【0033】

なお、収録雑音マスキング閾値パワーを、図6のように再生音の周波数域平均パワーに対して連続的に変化する特性を用いて調整しても良い。この場合、図6のような特性を持つ関数を用いることによって、低レベルほど大きくなるように補正する。この関数は、例えば、式(1)のような関数である。式(1)中の閾値THは、任意の値が設定され、 $x = TH$ を満たすときに、式(1)が傾き1となるように設定される。

40

【数1】

$$y = \frac{1}{2 \cdot TH} \cdot x^2 + \frac{TH}{2} \quad (x \leq TH)$$

$$y = x \quad (x > TH) \quad \cdots (1)$$

50

このように、収録雑音マスキング閾値補正部 33 では、再生音の周波数域平均パワーに応じて収録雑音マスキング閾値のレベルを補正する。補正後の収録雑音マスキング閾値 $noise_thr[i]$ は、後述の再生音補正部 35 によって算出される再生音の増幅率に影響するため、再生音の周波数域平均パワーが小さいほど収録雑音マスキング閾値のレベルが高くなるよう補正することは、再生音の周波数域平均パワーが小さいほど再生音の周波数帯域に対する増幅率を高くすることを意味する。なお、上記説明においては、 $frame_power$ を再生音の周波数域パワーとして既定の閾値 $TH1$ 、 $TH2$ と比較しているが、単位区間のパワーの大小関係を判定することが目的であり、時間 / 周波数変換前の時間域平均パワーを用いても良い。

【0034】

10

再生音補正部 35 は、収録雑音マスキング閾値補正部 33 から出力された補正後のマスキング閾値 ($noise_thr_new[i]$ と称する) を用いて再生音の補正処理を行う。図 7 は、再生音補正部 35 を詳細に説明した図である。再生音補正部 35 は、再生音マスキング判定部 35a、パワースムージング部 35b、補正係数算出部 35c、補正係数スムージング部 35d、補正演算部 35e を含み、再生音マスキング判定部 35a から補正係数スムージング部 35d の処理を行って得られた補正係数を用いて、補正演算部 35e によって再生音の補正処理を行う構成となっている。以下、それぞれの処理について説明する。

【0035】

再生音マスキング判定部 35a は、再生音マスキング特性解析部 34 から入力された周波数成分ごとの再生音の周波数域パワーと再生音のマスキング閾値を用いて、再生音の他の周波数成分によってマスクされる周波数成分とマスクされない周波数成分とに分ける。

20

【0036】

再生音自身によってマスクされるか否かを判定するために、周波数成分ごとに、再生音の周波数域パワー $signal_power[i]$ と再生音のマスキング閾値 $signal_thr[i]$ との比較を行い、再生音のパワーが再生音のマスキング閾値以上ならば、その周波数成分は、再生音の他の周波数成分にマスクされないという情報を記憶する。また、再生音のパワーが再生音のマスキング閾値未満であるならば、その周波数成分は、再生音の他の周波数成分にマスクされるという情報を記憶する。

【0037】

パワースムージング部 35b は、補正係数算出部 35c の前段階の処理として、再生音の周波数域パワー $signal_power[i]$ のスムージングを行う。再生音のパワーをスムージングする理由は、補正係数の算出には収録雑音のマスキング閾値と再生音のパワーとの比が用いられるため、再生音のパワーをスムージングさせないで補正係数を求め、この補正係数を用いて補正を行った場合、再生音の微細な構造が崩れてしまい、聴感が悪くなるためである。再生音のパワーのスムージングは、例えば、加重移動平均を用いる方法が考えられる。

30

【0038】

補正係数算出部 35c では、再生音の補正を行うための補正係数 $tmp_coef[i]$ を算出する。この補正係数 $tmp_coef[i]$ の算出には、パワースムージング部 35b でスムージングされた再生音の周波数成分ごとのパワーと、収録雑音マスキング閾値補正部 33 から出力された収録雑音のマスキング閾値 $noise_thr_new[i]$ の値を用いる。

40

【0039】

ここで、再生音のマスキングについて説明する。図 8 は、収録雑音および再生音自身によるマスキングを模式的に表している。この図に示すように、収録雑音によってマスクされる周波数成分は、再生音自身によってマスクされる周波数成分と、再生音にはマスクされない周波数成分とがある。再生音自身によってマスクされる周波数成分 ($signal_thr[i] > signal_power[i]$ となる周波数成分) は、収録雑音が無かったとしても聞こえない周波数成分であるため、増幅させないように補正係数 $tmp_coef[i]$ を 1 または 1 以下の値に設定する。

【0040】

50

それに対して、収録雑音によってマスクされ、かつ、再生音自身にはマスクされない周波数成分（ $\text{signal_thr}[i]$ $\text{signal_power}[i]$ かつ $\text{noise_thr_new} > \text{signal_power}$ となる周波数成分）は、収録雑音が無ければ知覚できるにも関わらず、収録雑音によってマスクされていると判断できる。そこで、この周波数成分を増幅するように補正係数を設定する。

【0041】

このときの補正係数の算出は、式（2）によって行われる。

【0042】

$$\text{tmp_coef}[i] = F(\text{noise_thr_new}[i] / \text{signal_power}[i]) \quad \dots (2)$$

ただし、 $F()$ は補正係数 $\text{tmp_coef}[i]$ を算出する関数であり、例えば、式（3）や式（4）のような関数である。

【0043】

$$F(\text{noise_thr_new}[i] / \text{signal_power}[i]) = \text{noise_thr_new}[i] / \text{signal_power}[i] + \alpha \quad \dots (3)$$

$$F(\text{noise_thr_new}[i] / \text{signal_power}[i]) = \text{noise_thr_new}[i] / \text{signal_power}[i] \times \alpha \quad \dots (4)$$

なお、式（3）や式（4）中の α は補正係数を調整するための値である。この補正係数を調整するための値は、収録雑音マスキング閾値補正部33で再生音の周波数域平均パワーがTH1未満であると判定された場合であるか、収録雑音マスキング閾値補正部33で再生音の周波数域平均パワーがTH1以上でありTH2未満であると判定された場合であるかに応じて、この値を変えても良い。

【0044】

例えば、収録雑音マスキング閾値補正部33で再生音の周波数域平均パワーがTH1未満であると判定された場合には、式（3）や式（4）の関数によって補正係数 $\text{tmp_coef}[i]$ を算出するのに対して、収録雑音マスキング閾値補正部33で再生音の周波数域平均パワーがTH1以上でありTH2未満であると判定された場合には、式（5）や式（6）の関数を用いて補正係数 $\text{tmp_coef}[i]$ を算出とする。ただし、 α と β との関係は $\alpha > \beta$ である。

【0045】

$$F(\text{noise_thr_new}[i] / \text{signal_power}[i]) = \text{noise_thr_new}[i] / \text{signal_power}[i] + \beta \quad \dots (5)$$

$$F(\text{noise_thr_new}[i] / \text{signal_power}[i]) = \text{noise_thr_new}[i] / \text{signal_power}[i] \times \beta \quad \dots (6)$$

このように、再生音の周波数域平均パワーがTH1未満であると判定された場合（再生音が低レベルの場合）と再生音の周波数域平均パワーがTH1以上でありTH2未満であると判定された場合（再生音が中レベルの場合）とに応じて、補正係数を調整するために用いる値を変えることによって、もともと聞き取りにくい低レベルの区間に対して重みをつけて補正することができ、さらなる明瞭度の向上をはかることができる。

【0046】

補正係数算出部35cは、以上のような処理を行って補正係数 $\text{tmp_coef}[i]$ を算出し、出力する。出力された補正係数 $\text{tmp_coef}[i]$ は、補正係数スムージング部35dに入力される。

【0047】

補正係数スムージング部35dでは、補正係数算出部35cまでの処理で算出された補正係数 $\text{tmp_coef}[i]$ に対してスムージングを行い、スムージングされた補正係数 $\text{coef}[i]$ を出力する。補正係数 $\text{tmp_coef}[i]$ は、隣接する周波数成分に対する補正係数 $\text{tmp_coef}[i+1]$ や $\text{tmp_coef}[i-1]$ と不連続な場合がある。特に、再生音マスキング判定部35aで再生音自身にマスクされると判定された周波数成分に対する補正係数と、再生音自身にマスクされないと判定された周波数成分に対する補正係数とは算出方法が異なるため、隣接していた場合、不連続になりやすい。そこで、この不連続性を緩和するために、補正係数のスムージングを行い、再生音の品質劣化を抑制する。補正係数のスムージングは、例えば加重

10

20

30

40

50

移動平均によって行う。

【 0 0 4 8 】

なお、補正係数のスムージングは全周波数成分に対して行っても良いが、再生音自身にマスクされる周波数成分とマスクされない周波数成分との境界周辺に限定してスムージングを行っても良い。

【 0 0 4 9 】

補正演算部 3 5 e には、再生音のスペクトル $signal[i]$ と、補正係数スムージング部 3 5 d によってスムージングされた補正係数 $coef[i]$ とが入力される。補正演算部 3 5 e は、入力された補正係数 $coef[i]$ と再生音のスペクトル $signal[i]$ とを式 (7) のようにかけ合わせて補正後の再生音のスペクトル $signal_new[i]$ を得て、この値を出力する。

10

【 0 0 5 0 】

$$signal_new[i] = coef[i] \times signal[i] \quad \cdots (7)$$

なお、補正演算部 3 5 e によって再生音の補正を行うときに、低域信号 (例えば、1 0 0 H z 以下の信号) は補正を行わないという条件や、低域信号を増幅させるときには所定の閾値以下の補正係数とするという条件などを付しても良い。

【 0 0 5 1 】

このように、再生音補正部 3 5 では、収録雑音によってマスクされた再生音の周波数成分を補正するときに、再生音自身によってマスクされる周波数成分の信号は増幅させないことにより、再生音の音量の増幅をできるだけ抑えつつ、再生音の明瞭化を図る。その結果、再生音補正部 3 5 によって再生音の補正処理を行うと、図 3 (d) や図 4 (d) のように、再生音を雑音によってマスクされないよう増幅することができる。

20

【 0 0 5 2 】

以上のように、本実施形態の補正処理部では、収録雑音のマスキング閾値に基づいて再生音の補正係数を算出する前に、再生音のパワーに応じて収録雑音のマスキング閾値を補正している。これによって、実際の周囲の雑音レベルの大小に係わらず、聞き取りやすさを向上することができる。

【実施例 2】

【 0 0 5 3 】

本発明の第 2 の実施形態を説明する。実施例 2 でも、実施例 1 と同様に携帯電話機に実装した場合を例にして説明する。なお、携帯電話機の構成は、実施例 1 と同様であるため、説明を省略する。

30

【 0 0 5 4 】

第 2 の実施形態の補正処理部の構成を図 9 に示す。第 1 の実施形態の補正処理部と同じ構成要素は図 9 でも図 2 と同じ記号で示し、詳細な説明を省略する。第 2 の実施形態の補正処理部では、収録雑音マスキング閾値の補正に対して、操作部 1 5 を用いてユーザに設定された音量増減の設定値 (以降、ユーザボリューム値と称する) を考慮する。そのため、再生音の周波数域平均パワーをユーザボリューム値に応じて補正するユーザボリューム適用部 3 8 が設けられている。

【 0 0 5 5 】

ユーザボリューム適用部 3 8 には、再生音パワー解析部 3 2 から再生音の周波数域平均パワー $frame_power$ および、操作部 1 5 を用いてユーザに設定されたユーザボリューム値が入力される。ユーザボリューム適用部 3 8 では、入力された再生信号パワー情報をユーザボリューム値に応じて式 (8) のように補正を行う。ただし、式 (8) 中の ΔV は、基準となるボリューム値 (0 d B) とユーザによって設定されたボリューム値との差分の値である。

40

【 0 0 5 6 】

$$frame_power_new = 10^{\Delta V / 20} \cdot frame_power \quad \cdots (8)$$

また、再生音の周波数域パワー $signal_power[i]$ も同様にユーザボリューム値に応じて式 (9) のように補正を行う。

【 0 0 5 7 】

50

$$\text{signal_power_new}[i] = 10^{\Delta V/20} \cdot \text{signal_power}[i] \quad \dots (9)$$

このようにユーザボリューム適用部 38 では、再生音の周波数域平均パワー frame_power および再生音の周波数パワー $\text{signal_power}[i]$ を補正することによって、収録雑音マスキング閾値補正部 33 における再生音のレベル（低レベルか中レベルか大レベルか）の判定をより正確に行うことができ、聞き取りやすさを向上することができる。なお、実施例 1 と同様、上記説明においては、 frame_power を再生音の周波数域パワーとしているが、単位区間のパワーの大小関係を判定することが目的であり、時間 / 周波数変換前の時間域平均パワーを用いても良い。

【0058】

なお、上記実施形態に限定されることはなく、本発明の要旨を逸脱しない範囲において、適宜変更しても良い。

【図面の簡単な説明】

【0059】

【図 1】本発明の第 1 の実施形態に係る携帯電話機の構成を示すブロック図。

【図 2】本発明の第 1 の実施形態に係る携帯電話機の補正処理部の構成を示す図。

【図 3】再生音が低レベルの場合に、収録雑音マスキング閾値を補正する処理を示す概念図。

【図 4】再生音が中レベルの場合に、収録雑音マスキング閾値を補正する処理を示す概念図。

【図 5】本発明の第 1 の実施形態に係る携帯電話機の収録雑音マスキング閾値補正部によって補正された後の収録雑音マスキング閾値パワーの平均値と再生音の周波数域平均パワーとの関係を示す図。

【図 6】本発明の第 1 の実施形態に係る携帯電話機の収録雑音マスキング閾値補正部によって収録雑音マスキング閾値パワーの平均値を補正する関数の一例。

【図 7】本発明の第 1 の実施形態に係る携帯電話機の再生音補正部を詳細に表した図。

【図 8】周囲雑音によってマスクされる周波数成分を表した図。

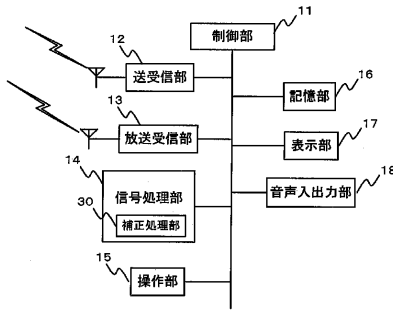
【図 9】本発明の第 2 の実施形態に係る携帯電話機の補正処理部の構成を示す図。

【符号の説明】

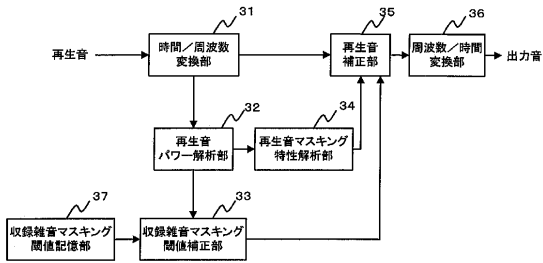
【0060】

11 制御部、12 送受信部、13 放送受信部、14 信号処理部、15 操作部、16 記憶部、17 表示部、18 音声入出力部、30 補正処理部、31 時間 / 周波数変換部、32 再生音パワー解析部、33 収録雑音マスキング閾値補正部、34 再生音マスキング特性解析部、35 再生音補正部、36 周波数 / 時間変換部、37 収録雑音マスキング閾値記憶部、35a 再生音マスキング判定部、35b パワースムージング部、35c 補正係数算出部、35d 補正係数スムージング部、35e 補正演算部、38 ユーザボリューム適用部

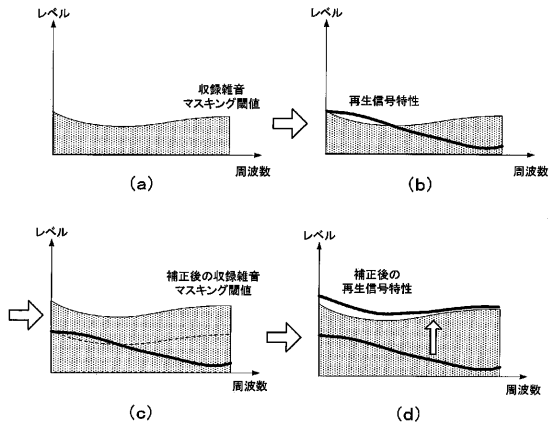
【図 1】



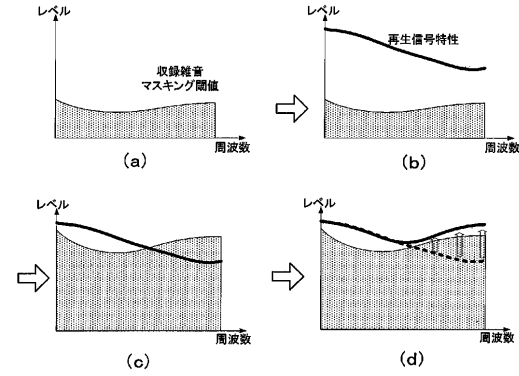
【図 2】



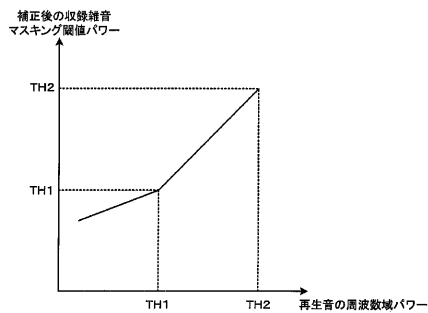
【図 3】



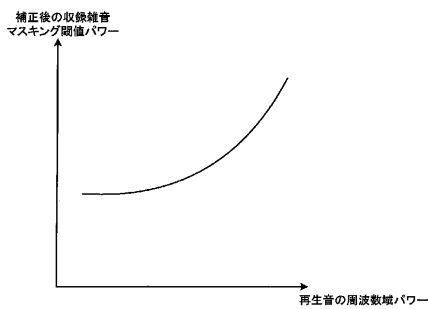
【図 4】



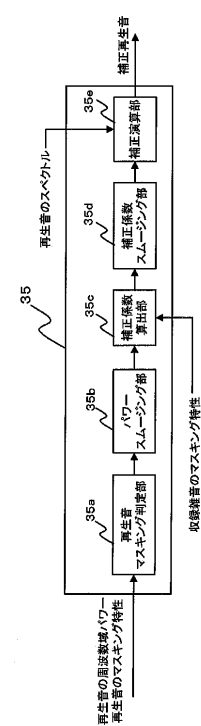
【図 5】



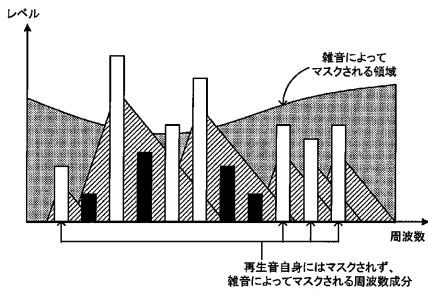
【図 6】



【図 7】



【図 8】



【図 9】

