

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4413205号
(P4413205)

(45) 発行日 平成22年2月10日(2010.2.10)

(24) 登録日 平成21年11月27日(2009.11.27)

(51) Int.Cl. F I
H04B 3/20 (2006.01) H04B 3/20

請求項の数 16 (全 20 頁)

(21) 出願番号	特願2006-171045 (P2006-171045)	(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
(22) 出願日	平成18年6月21日(2006.6.21)	(74) 代理人	100121706 弁理士 中尾 直樹
(65) 公開番号	特開2008-5094 (P2008-5094A)	(74) 代理人	100066153 弁理士 草野 卓
(43) 公開日	平成20年1月10日(2008.1.10)	(74) 代理人	100128705 弁理士 中村 幸雄
審査請求日	平成19年3月16日(2007.3.16)	(72) 発明者	小林 和則 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
		(72) 発明者	古家 賢一 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 エコー抑圧方法、装置、エコー抑圧プログラム、記録媒体

(57) 【特許請求の範囲】

【請求項1】

マイク受音信号を2以上のN個の周波数成分から成る周波数領域信号に変換する第1周波数領域変換処理と、

スピーカ出力信号をN個の周波数領域成分から成る周波数領域信号に変換する第2周波数領域変換処理と、

前記周波数領域信号に変換されたマイク受音信号と、前記周波数領域信号に変換されたスピーカ出力信号との比から、スピーカとマイクロホン間の伝達関数の振幅値である音響結合量を周波数領域毎に推定する音響結合量推定処理と、

前記周波数領域信号に変換されたスピーカ出力信号を予め設定した残響時間に対応した時定数で平滑化して時間平滑レベルを求め、この時間平滑レベルに前記音響結合量を乗じてマイク受音信号に含まれるエコー成分のレベルを周波数帯域毎に推定するエコーレベル推定処理と、

前記周波数領域信号に変換されたマイク受音信号の時間平滑レベルを計算して周波数帯域毎にマイクレベルを推定するマイクレベル推定処理と、

前記推定エコーレベルと、前記推定マイクレベルとを比較し、推定エコーレベルが大きい程抑圧量の大きいゲイン値を設定し、推定エコーレベルが充分小さい場合には予め設定した固定ゲイン値を周波数帯域毎に設定するゲイン制御処理と、

前記周波数領域信号に変換されたマイク受音信号に前記ゲイン制御処理で設定したゲイン値を周波数帯域毎に乗算するゲイン乗算処理と、

10

20

前記ゲイン乗算処理で乗算処理された出力信号を逆周波数領域変換し、送話信号として出力する逆周波数領域変換処理と、

を含むことを特徴とするエコー抑圧方法。

【請求項 2】

請求項 1 記載のエコー抑圧方法において、前記周波数帯域毎に施される前記ゲイン制御処理は、前記推定エコーレベルに予め設定した固定係数を乗じて求めた閾値と、前記推定マイクレベルとを比較し、前記推定マイクレベルが前記閾値より小さい場合は予め設定したエコー抑圧ゲインを瞬時利得係数値とし、これ以外の場合には予め設定した固定値を瞬時利得係数値とし、これらの瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧方法。

10

【請求項 3】

請求項 1 に記載のエコー抑圧方法において、前記推定エコーレベルに予め設定した固定係数を乗じて求めた複数の閾値と、前記推定マイクレベルとをそれぞれ比較し、前記推定マイクレベルが何れかの閾値よりも小さい場合は、前記推定マイクレベル以上の値を持つ閾値のうち最も前記推定マイクレベルに近い値を持つ閾値に対応する予め設定されたエコー抑圧ゲインを瞬時利得係数値と定め、これ以外の場合には予め設定した固定値を瞬時利得係数値とし、これら瞬時利得係数値を時間平滑した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧方法。

【請求項 4】

請求項 1 乃至 3 の何れかに記載のエコー抑圧方法において、前記周波数領域信号に変換されたマイク受信信号から、マイク受信信号に含まれる定常雑音のレベルを周波数領域毎に推定するマイク雑音レベル推定処理と、

20

前記周波数領域信号に変換されたスピーカ出力信号から、スピーカ出力信号に含まれる定常雑音のレベルを周波数帯域毎に推定する受話雑音レベル推定処理と、

前記周波数領域信号に変換されたスピーカ出力信号から前記スピーカ出力信号に含まれる定常雑音レベルを減じた信号の時間平滑レベルに前記音響結合量を乗じて、マイク受信信号に含まれるエコー成分の推定エコーレベルを周波数帯域毎に推定するエコーレベル推定処理と、

前記推定エコーレベルに予め設定した固定係数を乗じた値と、前記推定マイクレベルとを比較し、推定エコーレベルが大きい場合は予め設定したエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合は、予め設定した固定値を瞬時エコー抑圧利得係数値とし、前記マイク受信信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを比較し、雑音レベルが大きい場合は予め設定した雑音抑圧ゲイン瞬時雑音抑圧利得係数値とし、これ以外の場合は予め設定した固定値を瞬時雑音抑圧利得係数値とし、前記瞬時エコー抑圧利得係数値と、前記瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧方法。

30

【請求項 5】

請求項 1 乃至 4 の何れかに記載のエコー抑圧方法において、前記周波数帯域毎に施されるゲイン制御処理の出力となるゲイン値を周波数軸で平滑化する周波数特性平滑化処理を付加し、前記周波数帯域毎に施されるゲイン乗算処理は、前記周波数領域に変換されたマイク受信信号に前記周波数特性平滑化処理において平滑化後のゲイン値を、周波数帯域毎に乗算することを特徴とするエコー抑圧方法。

40

【請求項 6】

請求項 1 乃至 4 の何れかに記載のエコー抑圧方法において、前記スピーカ出力信号の周波数領域変換処理で周波数領域信号に変換された N 個の周波数成分のスピーカ信号を、2 以上で N 未満の N' 個の周波数成分のスピーカ出力信号に変換するスピーカ出力信号の帯域分割数変換処理と、

50

前記マイク受信信号の周波数領域変換処理で周波数領域信号に変換されたN個の周波数成分のマイク受信信号をN'個の周波数成分のマイク受信信号に変換するマイク受信信号の帯域分割数変換処理と、

前記ゲイン制御処理の出力であるN'個の周波数帯域のゲイン値をN個の周波数帯域のゲイン値に変換するゲインの帯域分割数変換処理を付加し、

前記N個の周波数帯域毎に施されていた処理を、帯域分割数変換後のN'の周波数帯域毎とすることを特徴とするエコー抑圧方法。

【請求項7】

請求項1乃至6の何れかに記載のエコー抑圧方法において、前記エコーレベル推定処理は、予め設定したM個の残響時間に対応した時定数で時間平滑化してエコーレベルを計算する2以上のM個のエコーレベル計算処理部を有し、

前記周波数帯域ごとに設けられたゲイン制御処理は、前記M個の残響時間に対応した推定エコーレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを各々比較し、推定エコーレベルが大きい場合はあらかじめ残響時間ごとに設定されたエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時エコー抑圧利得係数値としてM個の瞬時エコー抑圧利得係数を求め、前記マイク受信信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを比較し、雑音レベルが大きい場合はあらかじめ設定した雑音抑圧ゲインを瞬時雑音抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時雑音抑圧利得係数値とし、前記M個の瞬時エコー抑圧利得係数値と前記瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域ごとに設定することを特徴とするエコー抑圧方法。

【請求項8】

マイク受信信号を2以上のN個の周波数成分から成る周波数領域信号に変換する第1周波数領域変換部と、

スピーカ出力信号をN個の周波数領域成分から成る周波数領域信号に変換する第2周波数領域変換部と、

前記周波数領域信号に変換されたマイク受信信号と、前記周波数領域信号に変換されたスピーカ出力信号との比から、スピーカとマイクロホン間の伝達関数の振幅値である音響結合量を周波数領域毎に推定する音響結合量推定部と、

前記周波数領域信号に変換されたスピーカ出力信号を予め設定した残響時間に対応した時定数で平滑化して時間平滑レベルを求め、この時間平滑レベルに前記音響結合量を乗じてマイク受信信号に含まれるエコー成分のレベルを周波数帯域毎に推定するエコーレベル推定部と、

前記周波数領域信号に変換されたマイク受信信号の時間平滑レベルを計算して周波数帯域毎にマイクレベルを推定するマイクレベル推定部と、

前記推定エコーレベルと、前記推定マイクレベルとを比較し、推定エコーレベルが大きい程抑圧量の大きいゲイン値を設定し、推定エコーレベルが充分小さい場合には予め設定した固定ゲイン値を周波数帯域毎に設定するゲイン制御部と、

前記周波数領域信号に変換されたマイク受信信号に前記ゲイン制御部で設定したゲイン値を周波数帯域毎に乗算するゲイン乗算部と、

前記ゲイン乗算部で乗算処理された出力信号を逆周波数領域変換し、送話信号として出力する逆周波数領域変換部と、

によって構成することを特徴とするエコー抑圧装置。

【請求項9】

請求項8記載のエコー抑圧装置において、前記周波数帯域毎に設けられた前記ゲイン制御部は、前記推定エコーレベルに予め設定した固定係数を乗じて求めた閾値と、前記推定マイクレベルとを比較し、前記推定マイクレベルが前記閾値より小さい場合は予め設定したエコー抑圧ゲインを瞬時利得係数値とし、これ以外の場合には予め設定した固定値を瞬時利得係数値とし、これらの瞬時利得係数値を時間平滑化した値を前記ゲイン乗算部で用

10

20

30

40

50

いられるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧装置。

【請求項 10】

請求項 8 に記載のエコー抑圧装置において、前記推定エコーレベルに予め設定した値が予め設定した固定値以上の複数個の係数を乗じて求めた複数の閾値と、前記推定マイクレベルとをそれぞれ比較し、前記推定マイクレベルが何れかの閾値よりも小さい場合は、前記推定マイクレベル以上の値を持つ閾値のうち最も前記推定マイクレベルに近い値を持つ閾値に対応する予め設定されたエコー抑圧ゲインを瞬時利得係数と定め、これ以外の場合には予め設定した固定値を瞬時利得係数とし、これら瞬時利得係数を時間平滑した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧装置。

10

【請求項 11】

請求項 8 乃至 10 の何れかに記載のエコー抑圧装置において、前記周波数領域信号に変換されたマイク受信信号から、マイク受信信号に含まれる定常雑音のレベルを周波数領域毎に推定するマイク雑音レベル推定部と、

前記周波数領域信号に変換されたスピーカ出力信号から、スピーカ出力信号に含まれる定常雑音のレベルを周波数帯域毎に推定する受話雑音レベル推定部と、

前記周波数領域信号に変換されたスピーカ出力信号から前記スピーカ出力信号に含まれる定常雑音レベルを減じた信号の時間平滑レベルに前記音響結合量を乗じて、マイク受信信号に含まれるエコー成分の推定エコーレベルを周波数帯域毎に推定するエコーレベル推定部と、

20

前記推定エコーレベルに予め設定した固定係数を乗じた値と、前記推定マイクレベルとを比較し、推定エコーレベルが大きい場合は予め設定したエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合は、予め設定した固定値を瞬時エコー抑圧利得係数値とし、前記マイク受信信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを比較し、雑音レベルが大きい場合は予め設定した雑音抑圧ゲイン瞬時雑音抑圧利得係数とし、これ以外の場合は予め設定した固定値を瞬時雑音抑圧利得係数値とし、前記瞬時エコー抑圧利得係数値と、前記瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とするエコー抑圧装置。

30

【請求項 12】

請求項 8 乃至 11 の何れかに記載のエコー抑圧装置において、前記周波数帯域毎に設けられるゲイン制御部の出力となるゲイン値を周波数軸で平滑化する周波数特性平滑化部を付加し、前記周波数帯域毎に設けられたゲイン乗算部は、前記周波数領域信号に変換されたマイク受信信号に前記周波数特性平滑化部において平滑化後のゲイン値を、周波数帯域毎に乗算することを特徴とするエコー抑圧装置。

【請求項 13】

請求項 8 乃至 12 の何れかに記載のエコー抑圧装置において、前記スピーカ出力信号の周波数領域変換部で周波数領域信号に変換された N 個の周波数成分のスピーカ信号を、2 以上で N 未満の N' 個の周波数成分のスピーカ出力信号に変換するスピーカ出力信号の帯域分割数変換部と、

40

前記マイク受信信号の周波数領域変換部で周波数領域信号に変換された N 個の周波数成分のマイク受信信号を N' 個の周波数成分のマイク受信信号に変換するマイク受信信号の帯域分割数変換部と、

前記ゲイン制御部の出力である N' 個の周波数帯域のゲイン値を N 個の周波数帯域のゲイン値に変換するゲインの帯域分割数変換部を付加し、

前記 N 個の周波数帯域毎に施されていた処理を、帯域分割数変換後の N' の周波数帯域毎とすることを特徴とするエコー抑圧装置。

【請求項 14】

請求項 8 乃至 13 の何れかに記載のエコー抑圧装置において、前記エコーレベル推定部

50

は、予め設定したM個の残響時間に対応した時定数で時間平滑化してエコーレベルを計算する2以上のM個のエコーレベル計算部を有し、

前記周波数帯域ごとに設けられたゲイン制御部は、前記M個の残響時間に対応した推定エコーレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを各々比較し、推定エコーレベルが大きい場合はあらかじめ残響時間ごとに設定されたエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時エコー抑圧利得係数値としてM個の瞬時エコー抑圧利得係数を求め、前記マイク受信信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と前記推定マイクレベルとを比較し、雑音レベルが大きい場合はあらかじめ設定した雑音抑圧ゲインを瞬時雑音抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時雑音抑圧利得係数値とし、前記M個の瞬時エコー抑圧利得係数値と前記瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値を前記ゲイン乗算部で用いるゲイン値として周波数帯域ごとに設定することを特徴とするエコー抑圧装置。

10

【請求項15】

コンピュータが解読可能なプログラム言語によって記述され、コンピュータに請求項1乃至7記載のエコー抑圧方法を実行させるエコー抑圧プログラム。

【請求項16】

コンピュータが読み取り可能な記録媒体によって構成され、この記録媒体に請求項15記載のエコー抑圧プログラムを記録した記録媒体。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明はテレビ会議や音声会議などハンズフリー通信のエコー抑圧方法および装置に関する。

【背景技術】

【0002】

従来技術のエコー抑圧装置について説明する。

図8は特許文献1で開示された従来技術のエコー抑圧装置800の構成図である。従来技術のエコー抑圧装置800は、ゲイン乗算部801と、音響結合量推定部802と、送話判定部803と、受話判定部804と、ゲイン制御部805により構成される。

30

本エコー抑圧装置800は、スピーカとマイクロホンを用いた拡声通話においてマイク受信信号に混入するスピーカ音であるエコーを抑圧する。入力信号は、通話相手からの受信信号である受話信号と、自地点のマイクロホンで受信した信号であるマイク受信信号であり、出力信号は、自地点のスピーカから出力する信号であるスピーカ出力信号と、通話相手への送信信号である送話信号である。本装置は、マイク受信信号にエコーが主に含まれる場合は1よりも小さいゲインをマイク受信信号に乗算し、送話信号とすることで、マイク受信信号に含まれるエコーを抑圧する。

以下に、本装置の詳細について述べる。

音響結合量推定部802は、スピーカとマイクロホン間の伝達特性の振幅値である音響結合量 $A(t)$ を求める。ここでは t は離散化された時刻を表す。スピーカ出力信号とマイク受信信号の時間平均レベルを観測し、その比をとって、音響結合量を求める。スピーカ出力信号を $X(t)$ 、マイク受信信号を $Y(t)$ とした場合、音響結合量 $A(t)$ は式(1)により求められる。

40

$$A(t) = E \{ |Y(t)| / |X(t)| \} \quad \text{式(1)}$$

ただし、 $E\{\cdot\}$ は平均時間をとること、 $|\cdot|$ は絶対値を取ることを表す。

【0003】

送話判定部803は、受話信号 $X(t)$ と音響結合量 $A(t)$ とマイク受信信号 $Y(t)$ より、送話の判定をする。マイク受信信号 $Y(t)$ の時間平滑レベルを観測し、その時間平滑レベルがあらかじめ与えられた固定閾値を越え、さらに、音響結合量 $A(t)$ に受話信号 $X(t)$ の時間平滑レベルを乗じた推定音響エコーレベルを定数倍した閾値を越え

50

た場合に送話と判定する。

受話判定部 804 は、受話信号から、受話の判定をする。受話信号 $X(t)$ の時間平滑レベルを観測し、そのレベルがあらかじめ与えた固定閾値を越えた場合に受話と判定する。

【0004】

ゲイン制御部 805 は、送話判定部 803 と受話判定部 804 の判定結果より、ゲイン値を制御する。受話が検出され、送話が検出されない場合は、マイク受音信号にエコーのみが含まれるので、あらかじめ設定したエコー抑圧ゲインをゲイン値とし、これ以外の場合には、ゲイン値を 1 とする。また、ゲイン値の不連続な変化による音声の劣化を防ぐため、ゲイン値が滑らかに変化するように、時間平滑化が行われる。

10

ゲイン乗算部 801 は、ゲイン制御部 805 で設定されたゲイン値を、マイク受音信号に乘じ、送話信号とする。

以上の処理により、マイク受音信号にエコーのみが含まれる期間のみ、1 未満のゲイン値がマイク受音信号に乘じられ、エコーが抑圧された送話信号となる。マイク受音信号にエコーが含まれないか、近端話者の音声が含まれる場合は、マイク受音信号がそのまま送話信号となり、通話相手に送話音声を伝えることができる。

【特許文献 1】特許第 3082898 号明細書

【発明の開示】

【発明が解決しようとする課題】

【0005】

20

しかし、従来技術のエコー抑圧方法では、受話音声と近端話者の音声が同時に存在するダブルトーク時において問題が生じる。ダブルトーク時に送話が検出された場合には、エコーが抑圧されずに、そのまま送話信号となってしまう、相手側に伝わってしまうために、会話がしづらくなる。もし、送話が検出されなかった場合には、近端話者の音声が抑圧され、小さいレベルの音声となってしまうため、相手側で、音声が聞き取りづらくなる。このように、従来のエコー抑圧技術では双方向通話性能が充分ではない。

本発明の目的は、受話音声と近端話者の音声が同時に存在するダブルトーク時の双方向通話性能を向上するエコー抑圧方法を提供することにある。

【課題を解決するための手段】

【0006】

30

本発明によるエコー抑圧方法は、マイク受音信号を 2 以上の N 個の周波数成分から成る周波数領域信号に変換する第 1 周波数領域変換処理と、スピーカ出力信号を N 個の周波数領域成分から成る周波数領域信号に変換する第 2 周波数領域変換処理と、周波数領域信号に変換されたマイク受音信号と、周波数領域信号に変換されたスピーカ出力信号との比から、スピーカとマイクロホン間の伝達関数の振幅値である音響結合量を周波数領域毎に推定する音響結合量推定処理と、周波数領域信号に変換されたスピーカ出力信号を予め設定した残響時間に対応した時定数で平滑化して時間平滑レベルを求め、この時間平滑レベルに音響結合量を乘じてマイク受音信号に含まれるエコー成分のレベルを周波数帯域毎に推定するエコーレベル推定処理と、周波数領域信号に変換されたマイク受音信号の時間平滑レベルを計算して周波数帯域毎にマイクレベルを推定するマイクレベル推定処理と、推定エコーレベルと、推定マイクレベルとを比較し、推定エコーレベルが大きい程抑圧量の大きいゲイン値を設定し、推定エコーレベルが充分小さい場合には予め設定した固定ゲイン値を周波数帯域毎に設定するゲイン制御処理と、周波数領域信号に変換されたマイク受音信号にゲイン制御処理で設定したゲイン値を周波数帯域毎に乘算するゲイン乗算処理と、ゲイン乗算処理で乘算処理された出力信号を逆周波数領域変換し、送話信号として出力する逆周波数領域変換処理とを含むことを特徴とする。

40

【0007】

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、周波数帯域毎に施される前記ゲイン制御処理は、推定エコーレベルに予め設定した固定係数を乘じて求めた閾値と、推定マイクレベルとを比較し、推定マイクレベルが閾値より小さい場合は予め設

50

定したエコー抑圧ゲインを瞬時利得係数値とし、これ以外の場合には予め設定した固定値を瞬時利得係数値とし、これらの瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とする。

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、推定エコーレベルに予め設定した固定係数を乗じて求めた複数の閾値と、推定マイクレベルとをそれぞれ比較し、推定マイクレベルが何れかの閾値よりも小さい場合は、推定マイクレベル以上の値を持つ閾値のうち最も推定マイクレベルに近い値を持つ閾値に対応する予め設定されたエコー抑圧ゲインを瞬時利得係数値と定め、これ以外の場合には予め設定した固定値を瞬時利得係数値とし、これら瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とする。

10

【 0 0 0 8 】

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、周波数領域信号に変換されたマイク受音信号から、マイク受音信号に含まれる定常雑音のレベルを周波数領域毎に推定するマイク雑音レベル推定処理と、周波数領域信号に変換されたスピーカ出力信号から、スピーカ出力信号に含まれる定常雑音のレベルを周波数帯域毎に推定する受話雑音レベル推定処理と、周波数領域信号に変換されたスピーカ出力信号からスピーカ出力信号に含まれる定常雑音レベルを減じた信号の時間平滑レベルに前記音響結合量を乗じて、マイク受音信号に含まれるエコー成分の推定エコーレベルを周波数帯域毎に推定するエコーレベル推定処理と、推定エコーレベルに予め設定した固定係数を乗じた値と、推定マイクレベルとを比較し、推定エコーレベルが大きい場合は予め設定したエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合は、予め設定した固定値を瞬時エコー抑圧利得係数値とし、マイク受音信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と推定マイクレベルとを比較し、雑音レベルが大きい場合は予め設定した雑音抑圧ゲイン瞬時雑音抑圧利得係数値とし、これ以外の場合は予め設定した固定値を瞬時雑音抑圧利得係数値とし、前記瞬時エコー抑圧係数値と、前記瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値をゲイン乗算処理で用いるゲイン値として周波数帯域毎に設定することを特徴とする。

20

【 0 0 0 9 】

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、周波数帯域毎に施されるゲイン制御処理の出力となるゲイン値を周波数軸で平滑化する周波数特性平滑化処理を付加し、周波数帯域毎に施されるゲイン乗算処理は、周波数領域に変換されたマイク受音信号に周波数特性平滑化処理において平滑化後のゲイン値を、周波数帯域毎に乗算することを特徴とする。

30

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、スピーカ出力信号の周波数領域変換処理で周波数領域信号に変換されたN個の周波数成分のスピーカ信号を、2以上でN未満のN'個の周波数成分のスピーカ出力信号に変換するスピーカ出力信号の帯域分割数変換処理と、マイク受音信号の周波数領域変換処理で周波数領域信号に変換されたN個の周波数成分のマイク受音信号をN'個の周波数成分のマイク受音信号に変換するマイク受音信号の帯域分割数変換処理と、ゲイン制御処理の出力であるN'個の周波数帯域のゲイン値をN個の周波数帯域のゲイン値に変換するゲインの帯域分割数変換処理を付加し、N個の周波数帯域毎に施されていた処理を、帯域分割数変換後のN'の周波数帯域毎とすることを特徴とする。

40

【 0 0 1 0 】

本発明によるエコー抑圧方法は更に、前記エコー抑圧方法において、エコーレベル推定処理は、予め設定したM個の残響時間に対応した時定数で時間平滑化してエコーレベルを計算する2以上のM個のエコーレベル計算処理部を有し、周波数帯域ごとに設けられたゲイン制御処理は、M個の残響時間に対応した推定エコーレベルに予め設定した固定係数を乗じた値と推定マイクレベルとを各々比較し、推定エコーレベルが大きい場合はあらかじめ残響時間ごとに設定されたエコー抑圧ゲインを瞬時エコー抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時エコー抑圧利得係数値としてM個の瞬時エコー抑

50

圧利得係数を求め、マイク受音信号に含まれる定常雑音のレベルに予め設定した固定係数を乗じた値と推定マイクレベルとを比較し、雑音レベルが大きい場合はあらかじめ設定した雑音抑圧ゲインを瞬時雑音抑圧利得係数値とし、これ以外の場合には予め設定した固定値を瞬時雑音抑圧利得係数値とし、M個の瞬時エコー抑圧利得係数値と瞬時雑音抑圧利得係数値のうち最も小さい値を瞬時利得係数値とし、この瞬時利得係数値を時間平滑化した値を前記ゲイン乗算処理で用いるゲイン値として周波数帯域ごとに設定することを特徴とする。

【発明の効果】

【0011】

本発明の以上で説明する第1～5の実施例によれば、マイク受音信号に対し、帯域別にエコー抑圧のためのゲインを与えることで、受話音声と近端話者音声の両方が存在するダブルトーク区間においても、エコーを抑圧しながら、近端話者音声のレベル変動を抑え、高品質な双方向通信を行うことが可能である。

更に本発明の第2の実施例によれば、空調やファンなどの定常雑音を抑圧することができる。本発明の第3の実施例によれば、不連続な周波数特性を軽減し、より自然な音質を実現することができる。本発明の第4の実施例によれば、演算する帯域数を軽減し、演算量を削減することができる。本発明の第5の実施例によれば、残響時間の長い環境においても、十分なエコー抑圧をしながら、近端話者の音声劣化を最小限とすることができる。

【発明を実施するための最良の形態】

【0012】

本発明によるエコー抑圧装置はハードウェアによって構成し、ハードウェアによって構成したエコー抑圧装置により本発明によるエコー抑圧方法を実行させる実施形態も可能であるが、最も簡素に実現するには本発明によるエコー抑圧プログラムをコンピュータにインストールし、コンピュータにエコー抑圧装置として機能させ、本発明によるエコー抑圧方法を実行させる実施形態が最良の実施形態である。

コンピュータにより本発明によるエコー抑圧装置を構成するには、コンピュータにインストールしたエコー抑圧プログラムによりコンピュータには少なくとも、マイク受音信号を2以上のN個の周波数成分から成る周波数領域信号に変換する第1周波数領域変換部と、スピーカ出力信号をN個の周波数領域成分から成る周波数領域信号に変換する第2周波数領域変換部と、周波数領域信号に変換されたマイク受音信号と、周波数領域信号に変換されたスピーカ出力信号との比から、スピーカとマイクロホン間の伝達係数の振幅値である音響結合量を周波数領域毎に推定する音響結合量推定部と、周波数領域信号に変換されたスピーカ出力信号を予め設定した残響時間に対応した時定数で平滑化して時間平滑レベルを求め、この時間平滑レベルに音響結合量を乗じてマイク受音信号に含まれるエコー成分のレベルを周波数帯域毎に推定するエコーレベル推定部と、周波数領域信号に変換されたマイク受音信号の時間平滑レベルを計算して周波数帯域毎にマイクレベルを推定するマイクレベル推定部と、推定エコーレベルと、推定マイクレベルとを比較し、推定エコーレベルが大きい程抑圧量の大きいゲイン値を設定し、推定エコーレベルが充分小さい場合には予め設定した固定ゲイン値を周波数帯域毎に設定するゲイン制御部と、周波数領域信号に変換されたマイク受音信号にゲイン制御部で設定したゲイン値を周波数帯域毎に乗算するゲイン乗算部と、ゲイン乗算部で乗算処理された出力信号を逆周波数領域変換し、送話信号として出力する逆周波数領域変換部とを構築し、エコー抑圧装置として機能させる。

【0013】

本発明の実施例を説明する前に、本発明によるエコー抑圧装置の利用状況の概要を図7を用いて説明する。図7に示す700は本発明の各実施例で説明するエコー抑圧装置を示す。本発明によるエコー抑圧装置700の入力信号は通話相手からの受信信号である受話信号と、自地点のマイクロホンMで受音した信号であるマイク受音信号であり、出力信号は、自地点のスピーカSPから出力する信号であるスピーカ出力信号と、通話相手への送話信号である送話信号である。

図7に示すように、マイク受音信号は、近端話者の音声に加えて、スピーカ出力音がマ

10

20

30

40

50

イクロホンMに回り込んだ信号(エコー)が含まれる。エコー抑圧装置700はこのエコーを抑圧し、会話をしやすくする。また、エコー抑圧装置700の各入力信号は、AD変換により、アナログ信号から離散時間の信号に変換され、各出力信号は、離散時間信号からDA変換によりアナログ信号に変換されている。

本発明では、マイク受音信号をN個の周波数帯域信号に分割し、周波数帯域ごとでエコーを抑圧するためのゲインを信号に乘じ、エコーを抑圧する。ただし、Nは2以上の整数である。エコー成分の多く含まれる帯域はゲインを小さくして抑圧し、近端話者の音声が多く含まれる帯域はゲインを予め定めた固定値、例えば1として通過させるようにゲイン値を制御する。このようにすることで、エコーと近端話者の音声と同時に存在するダブルトーク期間であっても、近端話者の音声のレベルの減衰を抑えながら、エコーを抑圧することが可能となり、双方向通話性能を向上させることができる。

10

【実施例1】

【0014】

図1に、本発明の第1実施例であるエコー抑圧装置100のブロック図を示す。本発明の第1実施例で提案するエコー抑圧装置100は、マイク受音信号を周波数領域信号に変換する第1周波数領域変換部101と、逆周波数領域変換部102と、スピーカ出力信号を周波数領域信号に変換する第2周波数領域変換部103と、ゲイン乗算部104₁~104_Nと、音響結合量推定部105₁~105_Nと、エコーレベル推定部106₁~106_Nと、マイクレベル推定部107₁~107_Nと、ゲイン制御部108₁~108_Nとにより構成される。

20

第1周波数領域変換部101は、マイク受音信号をN個の周波数成分を持つ周波数領域信号に変換する。周波数領域への変換はフーリエ変換や帯域分割フィルタバンクにより行うことができる。変換後のマイク受話信号をY()で表す。ただし、 は周波数領域変換でのN個の周波数帯域の中心周波数 $f_0 \sim f_{N-1}$ に対応した離散的な周波数である。この周波数領域変換は、1サンプル以上の複数サンプルごとに行われる。この周波数領域変換が行われる間隔を1フレームと表す。このような処理により、1フレームごとに、周波数領域に変換された受話信号Y()が得られる。

【0015】

スピーカ出力信号を周波数領域信号に変換する第2周波数領域変換部103は、マイク受音信号の第1周波数領域変換部101と同様の処理によりスピーカ出力信号をN個の周波数成分に変換する。変換後のスピーカ出力信号をX()で表す。

30

周波数帯域ごとに設けられた音響結合量推定部105₁~105_Nは、マイク受音信号の第1周波数領域変換部101の出力Y()とスピーカ出力信号の第2周波数領域変換部103の出力X()から、周波数帯域ごとの音響結合量A()を求める。音響結合量とは、スピーカとマイク間の伝達関数の振幅値であり、周波数領域変換後のスピーカ出力信号とマイク受音信号の絶対値の比で求められる。また、音響結合量の精度を向上するために時間平滑化が行われる。スピーカ出力信号の周波数領域変換信号をX()、マイク受音信号の周波数領域変換信号をY()とした場合、音響結合量A()は式(2)により求められる。

$$A() = E \{ |Y()| / |X()| \} \quad \text{式(2)}$$

40

ただし、E{ }は平均値を取ることを表し、| · |は絶対値をとることを表す。

周波数帯域ごとに設けられたエコーレベル推定部106₁~106_Nは、音響結合量推定部105₁~105_Nで求められた音響結合量A()と、スピーカ出力信号の周波数領域変換部103の出力X()からマイク受音信号に含まれるエコーのレベルを推定する。部屋の反響を無視した場合、エコーレベルは、受話信号の周波数ベクトル|X()|に音響結合量A()を乗じることで推定可能である。しかし、実際には部屋の音響が存在するため、反響成分も含めてエコーを推定する必要がある。通常、部屋の音響成分は時間とともに指数減衰するので、以下の式(3)により、エコーレベルR()の推定を行う。

$$R() = A() \cdot P()$$

50

$$\begin{aligned}
 P(n) &= |X(n)| && \text{for } P'(n) \leq |X(n)| \\
 P(n) &= u \cdot P'(n) + (1-u) \cdot |X(n)| && \text{for } P'(n) > |X(n)| \quad \text{式(3)}
 \end{aligned}$$

ただし、 $P(n)$ は反響に相当する時間平滑を行ったあとのスピーカ出力信号であり、 $P'(n)$ は1フレーム前の $P(n)$ であり、 u は反響の長さ(残響時間)の想定値を調整するための係数でありあらかじめ固定値が設定される。 u は例えば0 $u < 1$ の値をとり、1に近いほど残響時間の長い環境が模擬され、0に近いほど残響時間の短い環境が模擬される。

【0016】

周波数帯域ごとに設けられたマイクレベル推定部107₁ ~ 107_Nは、マイク受信信号の第1周波数領域変換部101の出力 $Y(n)$ からマイク受信信号のレベル(マイクレベル) $S(n)$ を推定する。マイクレベル $S(n)$ は $Y(n)$ の絶対値で求められる。これは、式(4)で表される。

$$S(n) = |Y(n)| \quad \text{式(4)}$$

もしくは、 $Y(n)$ の絶対値を時間平滑化して求めても良い。これは例えば式(5)で表される。

$$S(n) = w \cdot S'(n) + (1-w) \cdot |Y(n)| \quad \text{式(5)}$$

ただし、 $S'(n)$ は1フレーム前のマイクレベルであり、 w は時間平滑化の係数であり、例えば0 $w < 1$ の値をとる。 w が1に近いほど長い時間での平滑化となり、0に近いほど短い時間での時間平滑化となる。

周波数帯域ごとに設けられたゲイン制御108₁ ~ 108_Nは、エコーレベル推定部106₁ ~ 106_Nで推定されたエコーレベル $R(n)$ と、マイクレベル推定部107₁ ~ 107_Nで推定されたマイクレベル $S(n)$ から、マイク受信信号に与える帯域ごとのゲインを制御する。ゲイン制御は、マイク受信信号にエコー成分が多く含まれる場合はエコーを抑圧し、これ以外の場合には近端話者音声を通させるように行われる。

【0017】

まず、マイク受信信号にエコー成分が多く含まれている場合、推定エコーレベル $R(n)$ と推定マイクレベル $S(n)$ が近い値をとるので、推定エコーレベル $R(n)$ に予め設定した固定値、例えば1以上の固定係数 C を乗じた値より、推定マイクレベル $S(n)$ が小さい場合に、エコー成分が多く含まれる期間として検出する。この条件は式(6)で表される。

$$S(n) < C \cdot R(n) \quad \text{式(6)}$$

エコー成分が多く含まれる期間として検出されたら、その帯域の瞬時利得係数 $g(n)$ を、あらかじめ固定値で設定したエコー抑圧量 D に設定する。ただし、エコー抑圧量 D は例えば0 $D < 1$ の値をとり、小さい値にするほどエコー抑圧量が増加するが、ダブルトーク時の近端話者音声の劣化が増加する。次に、エコー成分が多く含まれる期間として検出されなかった場合は、エコー成分が小さいので、瞬時利得係数 $g(n)$ を予め設定した固定値、例えば1に設定し、マイク受信信号をそのまま通過させる。このゲイン制御を式で表せば式(7)となる。

$$\begin{aligned}
 g(n) &= D && \text{for } S(n) < C \cdot R(n) \\
 g(n) &= 1 && \text{for } S(n) > C \cdot R(n)
 \end{aligned}$$

次に、瞬時利得係数 $g(n)$ を時間平滑化して、ゲイン乗算部104₁ ~ 104_Nに出力するゲイン $G(n)$ を求める。時間平滑化することでゲインの急激な変化による音質劣化を抑えることができる。時間平滑化は、例えば式(8)のように行われる。

$$\begin{aligned}
 G(n) &= a \cdot G'(n) + (1-a) \cdot g(n) && \text{for } g(n) \leq G'(n) \\
 G(n) &= b \cdot G'(n) + (1-b) \cdot g(n) && \text{for } g(n) > G'(n) \quad \text{式(8)}
 \end{aligned}$$

ただし、 $G'(n)$ は1フレーム前のゲイン $G(n)$ である。 a はゲイン下降時の平滑化係数、 b はゲイン上昇時の平滑化係数であり、あらかじめ固定値で設定される。 a と b は0から1の間の値をとり、1に近いほど長い時間での平滑化となり、0に近いほど短い時間での時間平滑化となる。

【0018】

10

20

30

40

50

更に、ゲイン制御部 108₁ ~ 108_N で、複数のエコー抑圧量を設定することもできる。この場合のゲイン制御は、マイク受音信号にエコー成分が多く含まれるほど、エコー抑圧量を増加させるように行われ、エコー成分が含まれない場合は近端話者音声を通させるように行われる。具体的な例としては、エコー成分が多く含まれる期間を検出するために固定係数 C を K 個の C₁ ~ C_K とし、推定エコーレベル R () にこれらの係数 C₁ ~ C_K を乗じた値と、推定マイクレベル S () の値を比較して、エコー抑圧量 D₁ ~ D_K を設定する。すなわち、式 (9) により、瞬時利得係数 g () を設定する。

$$\begin{aligned} g() &= D_K && \text{for } S() \geq C_K \cdot R() \\ g() &= D_{K-1} && \text{for } C_K \cdot R() < S() < C_{K-1} \cdot R() \\ &: && : \\ g() &= D_1 && \text{for } C_2 \cdot R() < S() < C_1 \cdot R() \\ g() &= 1 && \text{for } S() \geq C_1 \cdot R() \end{aligned} \quad \text{式 (9)}$$

10

ただし、固定係数 C₁ ~ C_K は、それぞれ予め定めた固定値、例えば 1 以上の値をとり、係数の大小関係は C₁ > C₂ > ... > C_K となるように、あらかじめ固定値を設定する。エコー抑圧量 D₁ ~ D_K は、それぞれ 0 < D_K < 1 の値をとり、その大小関係は D₁ > D₂ > ... > D_K となるようにあらかじめ固定値を設定する。次に、瞬時利得係数 g () を時間平滑化してゲイン乗算部 104₁ ~ 104_N で用いるゲイン G () を求める。時間平滑化は、前述と同様に例えば式 (8) により行われる。このように、複数のエコー抑圧量を設定することで、マイク受音信号に含まれるエコー成分の割合により、エコー抑圧量を変えることができ、抑圧量を従来技術のように 1 段階で設定するよりもエコー抑圧量を上げて、ダブルトーク時の通話品質を保つことが可能となる。

20

【 0 0 1 9 】

周波数帯域ごとに設けられたゲイン乗算部 104₁ ~ 104_N は、ゲイン制御部 108₁ ~ 108_N で計算された周波数帯域ごとの時間平滑化後のゲイン G () を、周波数領域後のマイク受音信号 Y () に、周波数帯域ごとで乗算する。

逆周波数領域変換部 102 は、ゲイン乗算部 104₁ ~ 104_N の出力信号を逆周波数領域変換して、送話信号として出力する。

以上の処理により、本実施例のエコー抑圧装置はマイク受音信号にエコー成分が多く含まれる周波数帯域成分を抑圧し、エコーを抑圧することができる。また、周波数帯域ごとで抑圧するか通過させるかの判定を行っているので、従来のエコー抑圧装置よりも受話音声と近端話者音声と同時に存在するダブルトーク時の近端話者の音声の劣化が小さい。

30

図 6 A、図 6 B、図 6 C に示すダブルトーク時のマイク受音信号と送話音声の周波数スペクトル図から、本発明の効果を説明する。図 6 A はエコー抑圧前のマイク受音信号の周波数スペクトルであり、図 6 B は、従来のエコー抑圧装置通過後の送話音声の周波数スペクトルである。従来のエコー抑圧装置では、フルバンドでロスを挿入するため、図 6 B に示すように、近端話者の音声成分とエコー成分を両方同じだけ抑圧してしまう。このため、ダブルトーク時にエコー抑圧すると送話信号に含まれる近端話者の音声小さくなり聞き取りづらいレベルになってしまう。これに対し、図 6 C に示す本発明のエコー抑圧装置通過後の送話音声の周波数スペクトルでは、エコー成分のレベルが高い周波数成分のみ抑圧され、近端話者音声成分のレベルが高い周波数成分はそのまま出力される。これにより、近端話者音声レベルの減少は、従来のエコー抑圧装置に比べ小さくなる。

40

【 0 0 2 0 】

また、本実施例においてマイク受音信号の第 1 周波数領域変換部 101 と、逆周波数領域変換部 102 と、スピーカ出力信号の第 2 周波数領域変換部 103 を削除して、周波数領域変換を行わずに帯域一括でゲイン制御することもできる。この場合、式 (9) で示したように複数段のエコー抑圧量を設定することで、マイク受音信号に含まれるエコー成分の割合により、エコー抑圧量を変えることができ、従来のエコー抑圧装置のように抑圧量を 1 段階で設定するのに比べ、エコー抑圧量を上げてダブルトーク時の通話品質を保つことができる。

以上示したように、本発明の第 1 の実施例によれば、受話音声と近端話者音声

50

在するダブルトークの場合でも、エコーを抑圧し、かつ音声の切断感の少ない高品質な双方向通信を実現できる。

【実施例 2】

【0021】

図 2 は、本発明の第 2 の実施例であるエコー抑圧装置 200 のブロック図である。本実施例のエコー抑圧装置 200 は、本発明の第 1 の実施例に、マイク雑音レベル推定部 201₁ ~ 201_N と、受話雑音レベル 202₁ ~ 202_N を追加した構成である。本実施例は、雑音レベルを推定することで、エコー抑圧に加えて雑音の抑圧を行うことができる。

周波数帯域ごとに設けられたマイク雑音レベル推定部 201₁ ~ 201_N は、マイク受音信号の第 1 周波数領域変換部 101 の出力信号から、マイク受音信号に含まれる雑音レベル $N_s(\cdot)$ を推定する。ここで推定する雑音は、空調やファンの音など時間的なレベル変動が少ない定常雑音である。雑音レベルの推定は、信号レベルの時間変化の谷を保持するディップホールド処理により行われる。このディップホールドは、推定の雑音レベルよりも信号レベルが大きい場合には長い時間の時定数での平滑化を行い、推定の雑音レベルよりも信号レベルが小さい場合には短い時間の時定数での平滑化を行うことにより実現される。例えば、式 (10) によりマイク受音信号の推定雑音レベル $N_s(\cdot)$ が計算される。

$$\begin{aligned} N_s(\cdot) &= p \cdot N_s'(\cdot) + (1-p) \cdot S(\cdot) && \text{for } N_s'(\cdot) \geq S(\cdot) \\ N_s(\cdot) &= q \cdot N_s'(\cdot) + (1-q) \cdot S(\cdot) && \text{for } N_s'(\cdot) < S(\cdot) \end{aligned} \quad \text{式 (10)}$$

ただし、 $N_s'(\cdot)$ は 1 フレーム前のマイク受音信号の推定雑音レベルであり、 p は推定雑音レベル減少時の平滑化係数であり、 q は推定雑音レベル増加時の平滑化係数であり、それぞれ例えば $0 < p < 1$ 、 $0 < q < 1$ の値をあらかじめ固定値として設定する。また $p < q$ となるように設定し、推定雑音レベル増加時に非常に緩やかな値の上昇とすることで信号をディップホールドして雑音レベルを推定する。

【0022】

周波数帯域ごとに設けられた受話雑音レベル推定部 202₁ ~ 202_N は、マイク雑音レベル推定部 201₁ ~ 201_N と同様の処理により、スピーカ出力信号の第 2 周波数領域変換部 103 の出力信号から、スピーカ出力信号に含まれる雑音である受話雑音レベル N_R を推定する。

周波数帯域ごとに設けられたエコーレベル推定部 106₁ ~ 106_N は、音響結合量推定部 105₁ ~ 105_N で求められた音響結合量 $A(\cdot)$ と、スピーカ出力信号の第 2 周波数領域変換部 103 の出力 $X(\cdot)$ と、受話雑音レベル推定部 202₁ ~ 202_N で推定された受話雑音レベル $N_R(\cdot)$ からマイク受音信号に含まれるエコーのレベルを推定する。ただし、本実施例で推定するエコーレベルは定常雑音成分を除いたエコーレベルである。

【0023】

受話信号の周波数スペクトル $|X(\cdot)|$ から雑音レベル $N_R(\cdot)$ を減算してから、音響結合量 $A(\cdot)$ を乗じてエコーレベル $R(\cdot)$ を推定する。式 (11) により、エコーレベル $R(\cdot)$ の推定を行う。

$$\begin{aligned} R(\cdot) &= A(\cdot) \cdot P(\cdot) \\ P(\cdot) &= u \cdot P'(\cdot) && \text{for } \{|X(\cdot)| - r \cdot N_R(\cdot)\} < 0 \\ P(\cdot) &= |X(\cdot)| && \text{for } P'(\cdot) \geq \{|X(\cdot)| - r \cdot N_R(\cdot)\} \\ P(\cdot) &= u \cdot P'(\cdot) + (1-u) \cdot \{|X(\cdot)| - r \cdot N_R(\cdot)\} && \text{for } P'(\cdot) > \{|X(\cdot)| - r \cdot N_R(\cdot)\} \\ &0 \end{aligned}$$

式 (11)

ただし、 r は雑音レベルの減算係数であり、あらかじめ固定的に与えられる。標準的には例えば $r = 1$ と設定し、より雑音レベルを減算したい場合は 1 以上の値に設定し、雑音レベルの減算を少なめにしたい場合には 1 以下に設定する。この他の記号は、本発明の第 1 の実施例と同様であり、以下のように定義される。 $P(\cdot)$ は反響に相当する時間平滑を行ったあとのスピーカ出力信号であり、 $P'(\cdot)$ は 1 フレーム前の $P(\cdot)$ であり、

10

20

30

40

50

u は反響の長さ（反響時間）の想定値を調整するための係数であり、あらかじめ固定値が設定される。u は例えば $0 < u < 1$ の値をとり、1 に近いほど残響時間の長い環境が模擬され、0 に近いほど残響時間の短い環境が模擬される。

周波数帯域ごとに設けられたゲイン制御部 $108_1 \sim 108_N$ は、エコーレベル推定部 $106_1 \sim 106_N$ で推定されたエコーレベル $R(\)$ と、マイクレベル推定部 $107_1 \sim 107_N$ で推定されたマイクレベル $S(\)$ と、マイク雑音レベル推定部 $201_1 \sim 201_N$ で推定されたマイク雑音レベル $N_S(\)$ から、マイク受信信号に与える帯域ごとのゲインを制御する。ゲイン制御は、マイク受信信号にエコー成分が多く含まれる場合はエコーを抑圧し、雑音が多に含まれる場合は雑音を抑圧し、これ以外の場合には近端話者音声を通させるように行われる。

10

【0024】

まず、マイク受信信号にエコー成分が多く含まれる場合、推定エコーレベル $R(\)$ と推定マイクレベル $S(\)$ が近い値をとるので、推定エコーレベル $R(\)$ に予め定めた固定値、例えば 1 以上の固定係数 C を乗じた値より、推定マイクレベル $S(\)$ が小さい場合に、エコー成分が多く含まれる期間として検出する。この条件は式 (12) で表される。

$$S(\) < C \cdot R(\) \quad \text{式(12)}$$

エコー成分が多く含まれる期間として検出されたら、その帯域の瞬時エコー抑圧ゲイン $g_E(\)$ を、あらかじめ固定値で設定したエコー抑圧量 D に設定する。ただし、エコー抑圧量 D は例えば $0 < D < 1$ の値をとり、小さい値にするほどエコー抑圧量が増加するが、ダブルトーク時の近端話者音声の劣化が増加する。次に、エコー成分が多く含まれる期間として検出されなかった場合は、エコー成分が小さいので、瞬時エコー抑圧利得係数 $g_E(\)$ を例えば 1 に設定し、マイク受信信号をそのまま通過させる。このゲイン制御を式で表せば式 (13) となる。

20

$$\begin{aligned} g_E(\) &= D && \text{for } S(\) < C \cdot R(\) \\ g_E(\) &= 1 && \text{for } S(\) > C \cdot R(\) \end{aligned} \quad \text{式(13)}$$

【0025】

次に、マイク受信信号に雑音成分が多く含まれる場合、推定雑音レベル $N_S(\)$ と推定マイクレベル $S(\)$ が近い値をとるので、推定雑音レベル $N_S(\)$ に予め定めた固定値、例えば 1 以上の固定係数 C_N を乗じた値より、推定マイクレベル $S(\)$ が小さい場合に、エコー成分が多く含まれる期間として検出する。この条件は式 (14) で表される。

30

$$S(\) < C_N \cdot N_S(\) \quad \text{式(14)}$$

雑音成分が多く含まれる期間として検出されたら、その帯域の瞬時雑音抑圧利得係数 $g_N(\)$ を、あらかじめ固定値として設定した雑音抑圧量 D_N に設定する。ただし、雑音抑圧量 D_N は $0 < D_N < 1$ の値をとり、小さい値にするほど雑音抑圧量が増加するが、近端話者音声の劣化が増大する。次に、雑音成分が多く含まれている期間として検出されなかった場合は、雑音成分が小さいので、瞬時雑音抑圧利得係数 $g_N(\)$ を予め定めた固定値、例えば 1 に設定し、マイク受信信号をそのまま通過させる。このゲイン制御を式で表せば式 (15) となる。

40

$$\begin{aligned} g_N(\) &= D_N && \text{for } S(\) < C_N \cdot N_S(\) \\ g_N(\) &= 1 && \text{for } S(\) > C_N \cdot N_S(\) \end{aligned} \quad \text{式(15)}$$

以上により求められた瞬時エコー抑圧利得係数 $g_E(\)$ と瞬時雑音抑圧利得係数 $g_N(\)$ のうち小さい方を瞬時雑音利得係数 $g(\)$ に設定する。これは式 (16) により表される。

$$g(\) = \text{MIN}\{g_E(\), g_N(\)\} \quad \text{式(16)}$$

ただし、 $\text{MIN}\{A, B\}$ は A の値と B の値のうち小さい方を選択するということを意味する。

【0026】

次に、本発明の実施例 1 と同様にして、瞬時利得係数 $g(\)$ を時間平滑化して、ゲイン $G(\)$ を求める。

また、本発明の実施例 1 と同様にして、複数のエコー抑圧量と雑音抑圧量を設定するこ

50

ともできる。エコー成分が多く含まれる期間を検出するための固定係数 C を K 個の $C_1 \sim C_K$ とし、推定エコーレベル $R(\)$ にこれらの係数 $C_1 \sim C_K$ を乗じた値と、マイクロレベル $S(\)$ の値を比較して、エコー抑圧量 $D_1 \sim D_K$ を、瞬時エコー抑圧利得係数 $g_E(\)$ に設定する。これを式で表せば式 (17) となる。

$$\begin{aligned}
 g_E(\) &= D_K && \text{for } S(\) > C_K \cdot R(\) \\
 g_E(\) &= D_{K-1} && \text{for } C_K \cdot R(\) < S(\) < C_{K-1} \cdot R(\) \\
 & \vdots && \vdots \\
 g_E(\) &= D_1 && \text{for } C_2 \cdot R(\) < S(\) < C_1 \cdot R(\) \\
 g_E(\) &= 1 && \text{for } S(\) > C_1 \cdot R(\)
 \end{aligned}
 \tag{17}$$

瞬時雑音抑圧利得係数 $g_N(\)$ も同様に、雑音成分が多く含まれる期間を検出するための固定係数 C_N を J 個の係数 $D_{N,1} \sim D_{N,J}$ を、瞬時エコー抑圧利得係数 $g_N(\)$ に設定する。これを式で表せば式 (18) となる。

$$\begin{aligned}
 g_N(\) &= D_{N,J} && \text{for } S(\) > C_{N,J} \cdot N_S(\) \\
 g_N(\) &= D_{N,J-1} && \text{for } C_{N,J} \cdot N_S(\) < S(\) < C_{N,J-1} \cdot N_S(\) \\
 & \vdots && \vdots \\
 g_N(\) &= D_{N,1} && \text{for } C_{N,2} \cdot N_S(\) < S(\) < C_{N,1} \cdot N_S(\) \\
 g_N(\) &= 1 && \text{for } S(\) > C_{N,1} \cdot N_S(\)
 \end{aligned}
 \tag{18}$$

ただし、固定係数 $C_1 \sim C_K$ および $C_{N,1} \sim C_{N,J}$ は、それぞれ 1 以上の値をとり、係数の大小関係は $C_1 > C_2 > \dots > C_K$ または $C_{N,1} > C_{N,2} > \dots > C_{N,J}$ となるように、あらかじめ固定値を設定する。エコー抑圧量 $D_1 \sim D_K$ と雑音抑圧量 $D_{N,1} \sim D_{N,J}$ は、それぞれ $0 < D_K < 1$ 、 $0 < D_{N,J} < 1$ の値をとり、その大小関係は $D_1 > D_2 > \dots > D_K$ または $D_{N,1} > D_{N,2} > \dots > D_{N,J}$ となるようにあらかじめ固定値を設定する。次に、瞬時エコー抑圧利得係数 $g_E(\)$ のうち小さい方を選択し瞬時利得係数 $g(\)$ とし、この瞬時利得係数 $g(\)$ を時間平均化してゲイン $G(\)$ を求める。時間平滑化は、第 1 の実施例と同様に行われる。このように、複数のエコー抑圧量と雑音抑圧量を設定することで、マイク受信信号に含まれるエコー成分や雑音成分の割合により、エコーまたは雑音抑圧量を変えることができ、抑圧量を従来技術のように 1 段階で設定するよりもエコーや雑音抑圧量を上げて、通話品質を保つことが可能となる。

これら以外の部分については、本発明の実施例 1 と同様であるので説明を省略する。

以上示したように、本実施例 2 によれば、本発明の第 1 実施例の効果に加えて、空調やファンの音などの定常的な雑音を抑圧することができる。

【実施例 3】

【0027】

図 3 は、本発明の第 3 の実施例であるエコー抑圧装置 300 のブロック図である。

本実施例のエコー抑圧装置 300 は、本発明の第 1 または第 2 の実施例に、周波数特性平滑部 301 を追加した構成である。本実施例は、ゲインの周波数特性を周波数軸で平滑化することで、不連続な周波数特性を軽減し、より自然な音質を実現する。

周波数特性平滑部は、ゲイン制御部 108₁ ~ 108_N で設定されたゲイン $G(\)$ を周波数軸で平滑化する。平滑化は例えば式 (19) により行われる。

$$G'(\)_i = \alpha \cdot G'(\)_{j-1} + (1-\alpha) \cdot G(\)_j
 \tag{19}$$

ただし、 $G'(\)_j$ は、平滑化後のゲインであり、 α は平滑係数で、あらかじめ固定値が設定される。 α は例えば $0 < \alpha < 1$ の値をとり、1 に近いほど、より強い平滑化となる。

この周波数特性の平滑化により、音質の不自然さの原因となるゲインの不連続な周波数特性を軽減し、音質を向上させることができる。

これ以外の部分に関しては本発明の第 1 または第 2 の実施例と同様であるので説明を省略する。

以上、示したように本実施例 3 によれば、本発明の第 1 または第 2 の実施例の効果に加えて、不連続な周波数特性を軽減し、より自然な音質を実現することができる。

【実施例 4】

【0028】

10

20

30

40

50

図4は、本発明の第4の実施例であるエコー抑圧装置400のブロック図である。本発明の第4の実施例は、本発明の第1～3の実施例にマイク受信信号の第1帯域分割変換部401と、スピーカ出力信号の帯域分割変換部402と、ゲインの帯域数変換部403が新たに設けられ、音響結合量推定部105₁～105_Nと、エコーレベル推定部106₁～106_Nと、マイクレベル推定部107₁～107_Nと、ゲイン制御部108₁～108_Nと、マイク雑音レベル推定部201₁～201_Nと、受話雑音レベル推定部202₁～202_Nの個数が、変換後の帯域分割数N'となった構成である。

マイク受信信号の帯域分割変換部401は、マイク受信信号の第1周波数領域変換部101の出力信号である周波数領域変換後のマイク受信信号Y()の帯域分割数をNからN'に変換する。ただし、N'はN未満の整数であり、変換後のY()をY'()とする。

10

変換は、変換前の複数の帯域を1つの帯域にまとめることにより行われ、例えば = n, ..., mの範囲のI個の帯域を、1つの帯域'に変換するときには、変換前の複数の帯域の振幅の平均値を、変換後の振幅とする。平均値は、式(20)または式(21)により計算される。

【数1】

$$Y'(\Omega') = \frac{1}{I} \sum_{\Omega=\omega_n}^{\omega_m} |Y(\Omega)| \quad (\text{式20})$$

20

$$Y'(\Omega') = \sqrt{\frac{1}{I} \sum_{\Omega=\omega_n}^{\omega_m} Y(\Omega)^2} \quad (\text{式21})$$

スピーカ出力信号の帯域分割数変換部402は、マイク受信信号の帯域分割数変換部401と同様の処理により、スピーカ出力信号の第2周波数領域変換部103の出力信号である周波数領域変換後のスピーカ出力信号X()の帯域分割数をNからN'に変換する。

音響結合量推定部105₁～105_Nと、エコーレベル推定部106₁～106_Nと、マイクレベル推定部107₁～107_Nと、ゲイン制御部108₁～108_Nと、マイク雑音レベル推定部201₁～201_Nと、受話雑音レベル推定部202₁～202_Nは、変換後の帯域分割数N'と同数あり、それぞれ帯域分割数変換後の信号を入力として、本発明の第1～3の実施例と同様の処理を行う。

30

【0029】

ゲインの帯域分割数変換部403は、ゲイン制御部108₁～108_Nの出力であるN'帯域のゲインを周波数領域変換部の帯域分割数Nのゲインに変換する。変換は変換前の1つの帯域を複数の帯域に分割するようにして行われる。例えば帯域'を、 = n, ..., mのI個の帯域に分割する場合、変換前の帯域'のゲイン値を、変換後のI個の帯域 = n, ..., mのゲイン値とする。

周波数特性平滑化部301を有さない実施例に適用する場合には、ゲイン乗算部104₁～104_Nとは、ゲインの帯域分割数変換部403により帯域分割数に変換されたゲイン値を入力とし、本発明の第1～3の実施例と同様の処理を行う。

40

周波数特性平滑化部301を有する実施例に適用する場合には、周波数特性平滑化部301は、ゲインの帯域分割数変換部403により帯域分割数に変換されたゲイン値を入力とし、本発明の第1～3の実施例と同様の処理を行う。

これら以外の部分については、本発明の第1～3の実施例と同様であるので、説明を省略する。

以上示したように、本実施例によれば、本発明の第1～3の実施例の効果に加えて、帯域分割数を一部の処理で少なくすることで、演算量の削減を行うことができる。

【実施例5】

50

【 0 0 3 0 】

図 5 は、本発明の第 5 の実施例であるエコー抑圧装置のエコーレベル推定部を示す図である。

本実施例のエコー抑圧装置は、本発明の第 1 ~ 第 4 の実施例に含まれるエコーレベル推定部 1 0 6_i (i = 1 , ... , N) が、図 5 に示すような複数個の残響時間に対応したエコーレベル計算部 1 0 6_{i, 1} ~ 1 0 6_{i, M} から構成されることを特徴とする。複数個の残響時間を想定して、複数のエコーレベル R₁ () ~ R_M () を推定することで、残響時間に応じたエコー抑圧量を設定することができる。

エコーレベル推定部 1 0 6_i (i = 1 , ... , N) は本発明の第 1 ~ 第 4 の実施例のエコーレベル推定部 1 0 6₁ ~ 1 0 6_N と同様にしてエコーレベルを推定する。ただし、反響の長さ (残響時間) の想定値を調整するための係数 u の値を M 個 (u₁ ~ u_M) 設定し、各エコーレベル計算部 1 0 6_{i, 1} ~ 1 0 6_{i, M} で、各残響時間を想定した場合のエコーレベル R₁ () ~ R_M () を計算する。

周波数帯域ごとに設けられたゲイン制御部 1 0 8_i ~ 1 0 8_N は、エコーレベル推定部 1 0 6_i ~ 1 0 6_N で推定されたエコーレベル R₁ () ~ R_M () と、マイクレベル推定部 1 0 7_i ~ 1 0 7_N で推定されたマイクレベル S () と、マイク雑音レベル推定部 2 0 1_i ~ 2 0 1_N で推定されたマイク雑音レベル N_S () から、マイクレベル受信信号に与える帯域ごとのゲインを制御する。瞬時雑音抑圧利得係数 g_N () は、本発明の第 1 ~ 第 4 の実施例と同様にして求められる。瞬時エコー抑圧利得係数も本発明の第 1 ~ 第 4 の実施例と同様にして求められるが、複数の残響時間を想定した複数の推定エコーレベル R₁ () ~ R_M () の、各々に対して瞬時エコー抑圧利得係数 g_{E, 1} () ~ g_{E, M} () が求められる。これらの瞬時エコー抑圧利得係数と瞬時雑音抑圧利得係数のうち最も利得値が小さいものを最小利得選択部 1 0 8 - 1 で選択し、選択した利得係数を瞬時抑圧利得係数 g () とし、この瞬時抑圧利得係数 g () を (実施例 1 ~ 4 と同様に) 時間平滑部 1 0 8 - 2 で時間平滑し、ゲイン乗算部 1 0 4₁ ~ 1 0 4_N に出力する。

【 0 0 3 1 】

ここで、エコー抑圧量の設定値 D , D₁ ~ D_K は、想定残響時間ごとに設定される。一般にエコーの大きさは、部屋の残響特性により時間とともに指数関数的に減衰する。このことから、想定残響時間が短いものほど、高いエコー抑圧量を設定し、残響時間がながいほどエコー抑圧量を減らすことで、残響に対して一定の残響エコーレベルとすることができる。これにより、エコーの過度な消去をなくし、近端話者の音声の劣化を最小限とすることができる。

また、本実施例においてマイク受信信号の第 1 周波数領域変換部 1 0 1 と、逆周波数領域変換部 1 0 2 と、スピーカ出力信号の第 2 周波数領域変換部 1 0 3 を削除して、周波数領域変換を行わずに全帯域一括してゲイン制御することもできる。この場合も実施例 1 と同様に、式 (9) で示したように複数のエコー抑圧量を設定することで、マイク受信信号に含まれるエコー成分の割合により、エコー抑圧量を変えることができ、従来のエコー抑圧装置のように抑圧量を 1 段階で設定でするのに比べ、エコー抑圧量を上げてダブルトーク時の通話品質を保つことができる。

【 0 0 3 2 】

これら以外の部分については、本発明の第 1 ~ 第 4 の実施例と同様であるので説明を省略する。

以上示したように、本実施例 5 によれば、本発明の第 1 ~ 第 4 の実施例の効果に加えて、残響時間の長い環境においても、十分なエコー抑圧をしながら、近端話者の音声劣化を最小限とすることができる。

以上各実施例で説明した本発明によるエコー抑圧装置はコンピュータが解読可能なプログラム言語によって記述されたエコー抑圧プログラムをコンピュータにインストールし、コンピュータに備えられた CPU に解読させることにより上述したエコー抑圧装置として機能し、本発明によるエコー抑圧方法が実行される。

10

20

30

40

50

本発明によるエコー抑圧プログラムはコンピュータが読み取り可能な例えば磁気ディスク或いはCD-ROM、半導体メモリ等の記録媒体に記録される。コンピュータにはこれらの記録媒体からインストールするか、又は通信回線を通じてインストールすることができる。

【産業上の利用可能性】

【0033】

本発明によるエコー抑圧方法、装置はテレビ会議システム或いは音声会議システムのようなハンズフリー通信の分野で活用される。

【図面の簡単な説明】

【0034】

【図1】本発明の第1の実施例であるエコー抑圧装置を説明するためのブロック図。

【図2】本発明の第2の実施例であるエコー抑圧装置を説明するためのブロック図。

【図3】本発明の第3の実施例であるエコー抑圧装置を説明するためのブロック図。

【図4】本発明の第4の実施例であるエコー抑圧装置を説明するためのブロック図。

【図5】本発明の第5の実施例であるエコー抑圧装置を説明するためのブロック図。

【図6】Aはエコー抑圧装置通過前の挿受話音声スペクトルを示す図、Bは従来のエコー抑圧装置通過後の送話音声スペクトルを示す図、Cは本発明のエコー抑圧装置通過後の送話音声スペクトルを示す図である。

【図7】本発明によるエコー抑圧装置の使い方を説明するための図。

【図8】従来のエコー抑圧装置を説明するためのブロック図。

【符号の説明】

【0035】

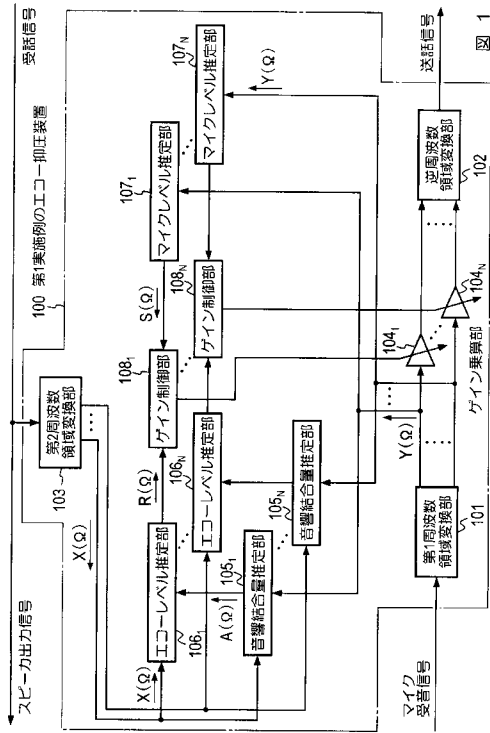
- | | | |
|-------------------------------------|-------------|--|
| 101 | 第1周波数領域変換部 | |
| 102 | 逆周波数領域変換部 | |
| 103 | 第2周波数領域変換部 | |
| 104 ₁ ~ 104 _N | ゲイン乗算部 | |
| 105 ₁ ~ 105 _N | 音響結合量推定部 | |
| 106 ₁ ~ 106 _N | エコーレベル推定部 | |
| 107 ₁ ~ 107 _N | マイクレベル推定部 | |
| 108 ₁ ~ 108 _N | ゲイン制御部 | |
| 201 ₁ ~ 201 _N | マイク雑音レベル推定部 | |
| 202 ₁ ~ 202 _N | 受話雑音レベル推定部 | |
| 301 | 周波数特性平滑化部 | |
| 401 ~ 403 | 帯域分割数変換部 | |

10

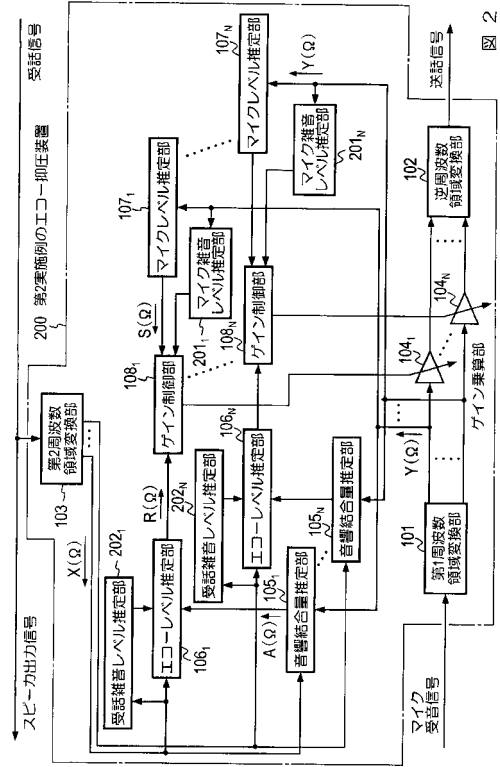
20

30

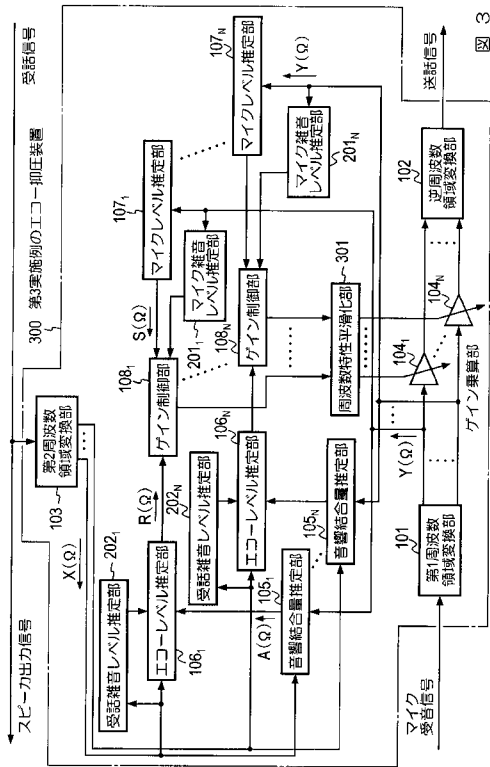
【図1】



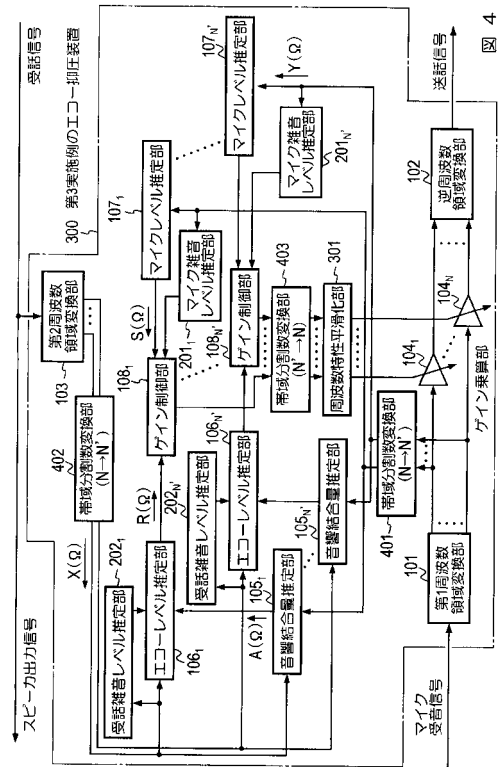
【図2】



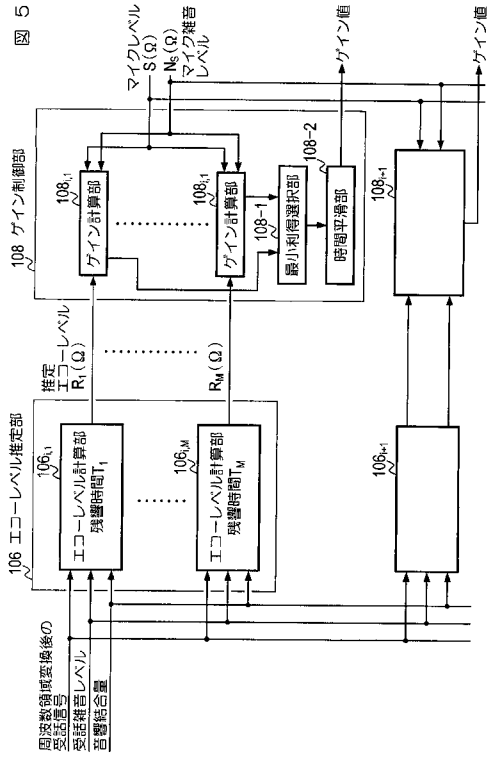
【図3】



【図4】



【 図 5 】



【 図 6 】

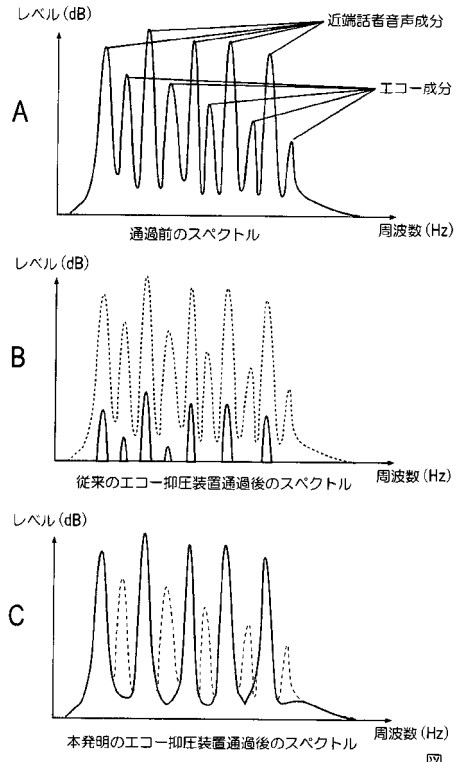


図 6

【 図 7 】

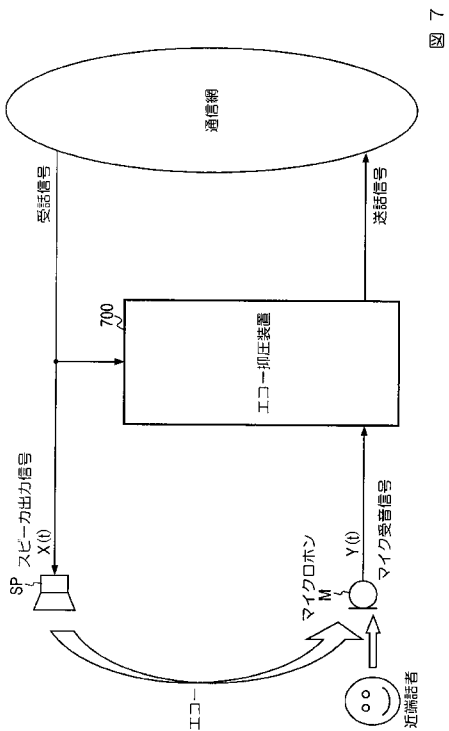


図 7

【 図 8 】

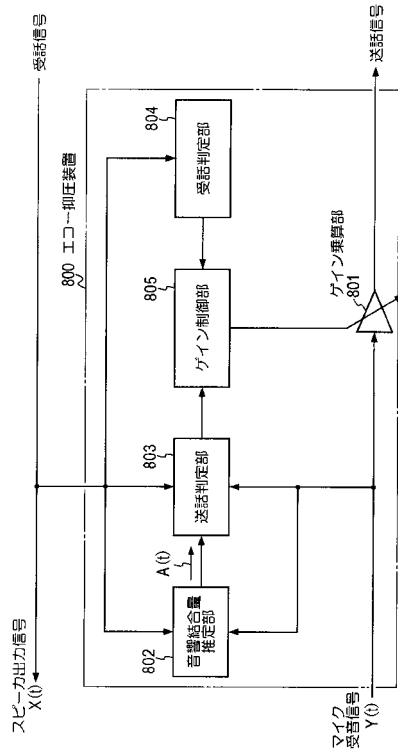


図 8

フロントページの続き

- (72)発明者 羽田 陽一
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 片岡 章俊
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 佐藤 聡史

- (56)参考文献 特開平11-331046(JP,A)
特許第3082898(JP,B2)
特開2004-260491(JP,A)
特表平09-512980(JP,A)

- (58)調査した分野(Int.Cl., DB名)
H04B 1/76 - 3/44