

(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2010-74531

(P2010-74531A)

(43) 公開日 平成22年4月2日(2010.4.2)

(51) Int.Cl.	F I	テーマコード (参考)
H04R 3/00 (2006.01)	H04R 3/00 320	5D020
G10L 21/02 (2006.01)	G10L 21/02 101A	
	G10L 21/02 102A	

審査請求 有 請求項の数 10 O L (全 21 頁)

(21) 出願番号	特願2008-239637 (P2008-239637)	(71) 出願人	000004226
(22) 出願日	平成20年9月18日 (2008.9.18)		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
		(74) 代理人	100121706
			弁理士 中尾 直樹
		(74) 代理人	100066153
			弁理士 草野 卓
		(74) 代理人	100128705
			弁理士 中村 幸雄
		(72) 発明者	木下 慶介
			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	中谷 智広
			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】 残響除去装置、残響除去方法、残響除去プログラム、記録媒体

(57) 【要約】

【課題】 振幅の大きな収音信号の直後、及び、目的音の終端部分に残響成分が残留しやすいという問題があった。本発明は、振幅の大きな収音信号の直後、及び、目的音の終端部分に残留する残留成分を抑圧することができる残響除去装置等を提供することを目的とする。

【解決手段】 本発明の残響除去装置は、M個のセンサによってそれぞれ観測されたM個のチャネルmの収音信号をそれぞれ複数の時点でサンプリングして得られた収音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力部と、収音信号 $x_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離部を有する。

【選択図】 図3

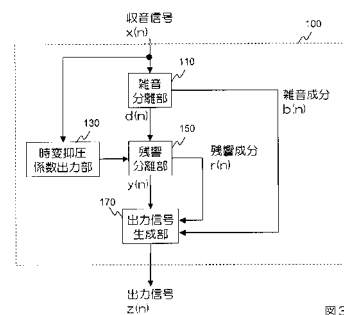


図3

【特許請求の範囲】

【請求項 1】

M ($M - 1$) 個のセンサによってそれぞれ観測された M 個のチャネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力部と、

前記收音信号 $x_m(n)$ 及び前記時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離部と、
を有する残響除去装置。

10

【請求項 2】

M ($M - 1$) 個のセンサによってそれぞれ観測された M 個のチャネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) を用いて、雑音除去信号 $d_m(n)$ を算出する雑音分離部と、

前記收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力部と、

前記雑音除去信号 $d_m(n)$ 及び前記時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離部と、
を有する残響除去装置。

20

【請求項 3】

請求項 2 記載の残響除去装置であって、

前記雑音分離部は、前記收音信号 $x_m(n)$ から前記雑音除去信号 $d_m(n)$ を差し引いて雑音成分 $b_m(n)$ を求め、

前記残響分離部は、前記残響除去信号 $y_m(n)$ に加え、残響成分 $r_m(n)$ を求め、
残響除去信号、前記雑音成分 $b_m(n)$ 及び前記残響成分 $r_m(n)$ を重み付け加算して、出力信号を算出する出力信号生成部を有すること、
を特徴とする残響除去装置。

【請求項 4】

請求項 1 から 3 記載の何れかの残響除去装置であって、

前記時変抑圧係数出力部は、前記收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ を計算するパワー時間包絡値計算部と、

前記パワー時間包絡値 $e_m(t)$ を用いて、目的音の終端部分の時刻 A_m を検出する目的音終端時刻検出部と、

前記パワー時間包絡値 $e_m(t)$ を用いて、高パワーを有する前記收音信号 $x_m(n)$ が発生した直後の時刻 B_m を検出する高パワー時刻検出部と、

前記時刻 A_m または B_m を用いて、予め定められた抑圧係数を出力する抑圧係数出力部を有すること

を特徴とする残響除去装置。

30

【請求項 5】

請求項 2 から 4 記載の何れかの残響除去装置であって、

前記残響分離部は、

前記雑音除去信号 $d_m(n)$ を疑似白色化する疑似白色化部と、

前記疑似白色化した雑音除去信号 $d'_m(n)$ を用いて、マルチステップ線形予測モデルの各線形予測係数を求める線形予測係数算出部と、

前記疑似白色化した雑音除去信号 $d'_m(n)$ と各線形予測係数から後部残響信号 $r_m(n)$ を予測する後部残響予測部と、

前記疑似白色化した雑音除去信号 $d'_m(n)$ 及び前記後部残響信号 $r_m(n)$ を周波数領域に変換し、それぞれの振幅スペクトルと位相情報を抽出し、出力する周波数領域変換部と、

40

50

前記周波数領域の雑音除去信号のパワー値から後部残響信号のパワー値を差し引いた値に、前記時変抑圧係数 $\lambda_m(t)$ を掛け合わせた値を残響除去信号の振幅スペクトル $|Y_m(\omega, t)|$ として算出する後部残響除去部と、

前記振幅スペクトル $|Y_m(\omega, t)|$ と前記位相情報を用いて時間領域の残響除去信号 $y_m(n)$ を算出する時間領域変換部と、

前記雑音除去信号 $d_m(n)$ から残響除去信号を差し引くことによって残響成分 $r_m(n)$ を求める残響抽出部を有すること

を特徴とする残響除去装置。

【請求項 6】

$M(M-1)$ 個のセンサによってそれぞれ観測された M 個のチャンネル $m(m=1, \dots, M)$ の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力ステップと、

前記收音信号 $x_m(n)$ 及び前記時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離ステップと、

を有する残響除去方法。

【請求項 7】

$M(M-1)$ 個のセンサによってそれぞれ観測された M 個のチャンネル $m(m=1, \dots, M)$ の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) を用いて、雑音除去信号 $d_m(n)$ を算出する雑音分離ステップと、

前記收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力ステップと、

前記雑音除去信号 $d_m(n)$ 及び前記時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離ステップと、

を有する残響除去方法。

【請求項 8】

請求項 7 記載の残響除去方法であって、

前記雑音分離ステップは、前記收音信号 $x_m(n)$ から前記雑音除去信号 $d_m(n)$ を差し引いて雑音成分 $b_m(n)$ を求め、

前記残響分離ステップは、前記残響除去信号 $y_m(n)$ に加え、残響成分 $r_m(n)$ を求め、

残響除去信号、前記雑音成分 $b_m(n)$ 及び前記残響成分 $r_m(n)$ を重み付け加算して、出力信号を算出する出力信号生成ステップを有すること、

を特徴とする残響除去方法。

【請求項 9】

請求項 1 から 5 の何れかに記載の残響除去装置としてコンピュータを機能させるためのプログラム。

【請求項 10】

請求項 9 記載のプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、残響を含む收音信号から残響成分を除去する残響除去装置、残響除去方法、残響除去プログラム、残響除去記録媒体に関する。

【背景技術】

【0002】

残響のある環境で收音された收音信号は、本来の目的音に残響が重畳された信号となる

10

20

30

40

50

。その目的音が音声等の場合、重畳した残響成分により音声の明瞭性が大きく低下する。残響を含む収音信号の残響除去装置として特許文献 1 記載の装置がある。図 1 は特許文献 1 記載の残響除去装置 50 の構成例を示す。センサで観測された収音信号が所定の標準化周波数でサンプリングされ、収音信号 $x(n)$ が生成される。収音信号 $x(n)$ は、メモリ 10a に格納される。なお、残響除去を行う全時間区間の収音信号 $x(n)$ を事前 to 取得し、メモリ 10a に格納しておき、分析フレーム毎に、以下の各過程を実行する。

【0003】

まず、疑似白色化 (Pre-whitening) 部 51 が、入力された収音信号 $x(n)$ が有する短時間区間での自己相関成分を抑制し、疑似白色化した収音信号 x' を生成して出力する。次に、第一線形予測係数算出部 52 が、収音信号 $x'(n)$ を用い、マルチステップ線形予測モデルの線形予測係数 $\alpha(p)$ を算出する。なお、 $\alpha(p)$ は、 $x(n)$ に対応する線形予測項の p 番目の線形予測係数とする。次に、後部残響予測部 53 に、疑似白色化部 51 で生成された疑似白色化後の収音信号 $x'(n)$ と、第一線形予測係数算出部 52 で算出された線形予測係数 $\alpha(p)$ とが入力される。そして、後部残響予測部 53 は、線形予測係数 $\alpha(p)$ と疑似白色化後の収音信号 $x'(n)$ とをマルチステップ線形予測モデルの線形予測項に代入して得られた線形予測値を、後部残響予測値 $r(n)$ として算出し出力する。次に、周波数領域変換部 55 に、疑似白色化部 51 で生成された疑似白色化後の収音信号 $x'(n)$ と、後部残響予測部 53 で算出された後部残響予測値 $r(n)$ とが入力される。周波数領域変換部 55 は、入力された疑似白色化後の収音信号 $x'(n)$ を周波数領域の収音信号 $X'(\omega, t)$ に変換し、後部残響予測値 $r(n)$ を周波数領域の後部残響予測値 $R(\omega, t)$ に変換する。周波数領域変換部 55 は、これらの処理により、周波数領域の収音信号 $X'(\omega, t)$ の振幅スペクトル $|X'(\omega, t)|$ と位相情報 $\arg[X'(\omega, t)]$ 、及び、周波数領域の後部残響予測値 $R(\omega, t)$ の振幅スペクトル $|R(\omega, t)|$ と位相情報 $\arg[R(\omega, t)]$ とを抽出し、出力する。次に、後部残響除去部 59 に、周波数領域の収音信号の振幅スペクトル $|X'(\omega, t)|$ と、周波数領域の後部残響予測値の振幅スペクトル $|R(\omega, t)|$ とが入力される。そして、後部残響除去部 59 は、以下の式 (1) を用いて、周波数領域の収音信号の振幅スペクトル $|X'(\omega, t)|$ と、周波数領域の後部残響予測値の振幅スペクトル $|R(\omega, t)|$ との相対値を求め、当該相対値を後部残響除去信号の振幅スペクトル予測値 $|Y(\omega, t)|$ として出力する。

【0004】

$$|Y(\omega, t)| = (|X'(\omega, t)|^k - |R(\omega, t)|^k \cdot \text{const})^{1/k} \quad (1)$$

なお、const は定数を示し、 k は自然数を示す。時間領域変換部 61 は、後部残響除去部 59 から出力された後部残響除去信号の振幅スペクトル予測値 $|Y(\omega, t)|$ と、周波数領域変換部 55 から出力された周波数領域の収音信号の位相情報 $\arg[X'(\omega, t)]$ とが入力される。これらの情報を用い、後部残響除去信号の複素スペクトル予測値 $Y(\omega, t)$ を算出し、さらに、時間領域に変換した残響除去信号 $y(n)$ を算出して出力する。

【特許文献 1】国際公開第 WO 2007/100137 A1 号パンフレット

【発明の開示】

【発明が解決しようとする課題】

【0005】

図 2 (i) は、特許文献 1 記載の残響除去装置の出力信号を示す。特許文献 1 記載の残響除去装置を含む従来技術では、振幅の大きな収音信号の直後 (B)、及び、目的音 (例えば、音声等) の終端部分 (A) に残響成分が残留しやすいという問題があった。本発明は、振幅の大きな収音信号の直後、及び、目的音の終端部分に残留する残留成分を抑圧することができる残留除去装置、残留除去方法、残留除去プログラム、記録媒体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】

10

20

30

40

50

請求項 1 記載の残響除去装置は、 M ($M \geq 1$) 個のセンサによってそれぞれ観測された M 個のチャンネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力部と、收音信号 $x_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離部を有する。

【0007】

請求項 2 記載の残響除去装置は、 M ($M \geq 1$) 個のセンサによってそれぞれ観測された M 個のチャンネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) を用いて、雑音除去信号 $d_m(n)$ を算出する雑音分離部と、收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力部と、雑音除去信号 $d_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離部を有する。

10

【0008】

請求項 6 記載の残響除去方法は、 M ($M \geq 1$) 個のセンサによってそれぞれ観測された M 個のチャンネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力ステップと、收音信号 $x_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離ステップを有する。

20

【0009】

請求項 7 記載の残響除去方法は、 M ($M \geq 1$) 個のセンサによってそれぞれ観測された M 個のチャンネル m ($m = 1, \dots, M$) の收音信号をそれぞれ複数の時点でサンプリングして得られた收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) を用いて、雑音除去信号 $d_m(n)$ を算出する雑音分離ステップと、收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する時変抑圧係数出力ステップと、雑音除去信号 $d_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める残響分離ステップを有する。

30

【発明の効果】

【0010】

振幅の大きな收音信号の直後、及び、目的音の終端部分に残留する残響成分を抑圧することができる。

【発明を実施するための最良の形態】

【0011】

ここで、本発明の実施例について述べる。

【実施例 1】

【0012】

図 3 は、実施例 1 の残響除去装置 100 の構成例を、図 4 は、実施例 1 の処理の流れを示す。以下、同様の構成要素については、同一の符号を付す。実施例 1 の残響除去装置 100 は、雑音分離部 110 と、時変抑圧係数出力部 130 と、残響分離部 150 と、出力信号生成部 170 と、を有する。センサで観測された收音信号 $x(u)$ (但し、 u は、連続時間を表す) が所定の標本化周波数でサンプリングされ、收音信号 $x(n)$ (但し、 n は、離散時間を表す) が生成される。なお、收音信号 $x(n)$ には、目的音 (例えば、音声等)、雑音成分 $b(n)$ 及び残響成分 $r(n)$ が含まれる。收音信号 $x(n)$ は、メモリ等に格納される。なお、残響除去を行う全時間区間の收音信号 $x(n)$ を事前に取得し、メモリ等に格納しておき、分析フレーム毎に、以下の各過程を実行する。また、分析フレームとは、複数の離散時間 n を含む時間区間を意味する。なお、收音信号 $x(n)$ の取得をリアルタイムで行いつつ、以下の各過程を実行してもよい。

40

【0013】

50

雑音分離部 110 は、收音信号 $x(n)$ を用いて、雑音除去信号 $d(n)$ を算出し、收音信号 $x(n)$ から雑音除去信号 $d(n)$ を差し引いて雑音成分 $b(n)$ を求める (s110)。図 5 は、雑音分離部 110 の構成例を、図 6 は、雑音分離部 110 の構成例の流れを示す。例えば、雑音分離部 110 は、雑音除去部 111 と、雑音抽出部 113 を有し、收音信号 $x(n)$ は、雑音除去部 111 と、雑音抽出部 113 に入力される。雑音除去部 111 は、收音信号 $x(n)$ を用いて、雑音除去信号 $d(n)$ を算出する (s111)。雑音を除去する従来技術として、例えば、Nathalie Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System", IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, MARCH 1999, VOL.7, NO.2, p.126-137 (以下「参考文献 1」という) 等がある。但し、雑音除去方法を参考文献 1 に限定するものではなく、他の雑音除去方法を用いてもよい。雑音除去部 111 は、算出した雑音除去信号 $d(n)$ を残響分離部 150 及び雑音抽出部 113 へ出力する。図 7 は、雑音除去部 111 の構成例を示す。雑音除去部 111 は、周波数領域変換部 111a、雑音推定部 111b、減算部 111c、閾値計算部 111d、係数取得部 111e、利得関数計算部 111f、時間領域変換部 111g を有する。例えば、雑音除去部 111 の有する周波数領域変換部 111a は、收音信号 $x(n)$ を入力され、短時間フーリエ変換等によって收音信号 $x(n)$ を周波数領域の收音信号 $X(\omega, t)$ に変換して、振幅スペクトル $|X(\omega, t)|$ と位相情報 $\arg[X(\omega, t)]$ を出力する。なお、短時間フーリエ変換によってこの処理を行う場合は、例えば以下の式 (2) を用いる。また、 $F[\cdot]$ は短時間フーリエ変換関数を示し、 $\text{Log}[\cdot]$ は対数関数を示す。

$$X(\omega, t) = \text{Log}[F[x(n)]] \quad (2)$$

【0014】

雑音推定部 111b は、收音信号 $x(n)$ を入力される。例えば、音声等の目的音が存在しない状態において、統計的仮定に基づいて、雑音成分を検出し、推定雑音成分 $b'(n)$ を求め、さらに、式 (2) により周波数領域の推定雑音成分 $B'(\omega, t)$ に変換し、振幅スペクトル $|B'(\omega, t)|$ を求め、出力する。なお、雑音成分が同じ状態を維持したり、非常にゆっくりと変化する場合には、以下の式により振幅スペクトル $|B'(\omega, t)|$ を更新してもよい。

$$|B'(\omega, t)|^y = \lambda_D \cdot |B'(\omega, t-1)|^y + (1 - \lambda_D) \cdot |X(\omega, t-1)|^y$$

但し、 λ_D は、忘却因子であり、 $0.5 \leq \lambda_D \leq 0.9$ である。 y は指数パラメータである。

【0015】

減算部 111c は、收音信号の振幅スペクトル $|X(\omega, t)|$ と推定雑音成分の振幅スペクトル $|B'(\omega, t)|$ を入力され、以下の式 (3) により推定雑音除去信号の振幅スペクトル $|D'(\omega, t)|$ を求め、出力する。

$$|D'(\omega, t)| = \{|X(\omega, t)|^2 - |B'(\omega, t)|^2\}^{1/2} \quad (3)$$

【0016】

閾値計算部 111d では、聴覚をモデルとして閾値 $T(\omega, t)$ を取得する。閾値 $T(\omega, t)$ を求める方法としては、例えば、Nathalie Virag, "Speech Enhancement Based on Masking Properties of the Auditory System", in Proc. IEEE ICASSP, Detroit, MI, MAY 1995, p.796-799 (以下「参考文献 2」という) 等がある。

【0017】

係数取得部 111e は、閾値 $T(\omega, t)$ を入力され、分析フレーム毎に、利得関数を求める際に用いるパラメータ $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ を取得する。利得関数計算部 111f は、パラメータ $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ 、收音信号の振幅スペクトル $|X(\omega, t)|$ と推定雑音成分の振幅スペクトル $|B'(\omega, t)|$ を入力され、以下の式によって、利得関数 $G(\omega, t)$ を計算し、出力する。

【0018】

10

20

30

40

【数 1】

$$G(\omega, t) = \begin{cases} \left(1 - \alpha(\omega, t) \cdot \left[\frac{|B'(\omega, t)|}{|X(\omega, t)|} \right]^{\gamma_1} \right)^{\gamma_2}, & \text{if } \left[\frac{|B'(\omega, t)|}{|X(\omega, t)|} \right]^{\gamma_1} < \frac{1}{\alpha(\omega, t) + \beta(\omega, t)} \\ \beta(\omega, t) \cdot \left[\frac{|B'(\omega, t)|}{|X(\omega, t)|} \right]^{\gamma_1}, & \text{otherwise} \end{cases}$$

【0019】

10

γ_1 、 γ_2 は指数パラメータであり、 $\gamma_2 = 1 / \gamma_1$ である。 $\alpha(\omega, t)$ の値を大きくすると、残留ノイズを減少させることができるが、音声歪みが増加する。 $\beta(\omega, t)$ の値を大きくすると、残留ノイズを減少させることができるが、雑音成分が増加する。トレードオフの関係にある。残留ノイズを減少させる適切な $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ を計算するために、係数取得部 111e において、以下のように求める。

$$\alpha(\omega, t) = F_\alpha [\alpha_{\min}, \alpha_{\max}, T(\omega, t)]$$

$$\beta(\omega, t) = F_\beta [\beta_{\min}, \beta_{\max}, T(\omega, t)]$$

【0020】

20

なお、 $\alpha_{\min}$ 、 $\beta_{\min}$ はそれぞれ $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ の最小値を、 $\alpha_{\max}$ 、 $\beta_{\max}$ はそれぞれ $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ の最大値を表す。 F_α 、 F_β は、 $T(\omega, t)$ が最大値の時に各パラメータの最小値 $\alpha_{\min}$ 、 $\beta_{\min}$ を返し、 $T(\omega, t)$ が最小値の時に各パラメータの最大値 $\alpha_{\max}$ 、 $\beta_{\max}$ を返し、 $T(\omega, t)$ が最小値と最大値の間の場合には、その値に応じて補間した $\alpha(\omega, t)$ 、 $\beta(\omega, t)$ を返す関数である。例えば、 $\alpha_{\min} = 1$ 、 $\alpha_{\max} = 6$ 、 $\beta_{\min} = 0$ 、 $\beta_{\max} = 0.02$ 、 $\gamma = \gamma_1 = 2$ 、 $\gamma_2 = (1/\gamma_1) = 0.5$ とすることが考えられる。詳しくは、参考文献 1 に述べられている。

【0021】

時間領域変換部 111g は、利得関数 $G(\omega, t)$ と收音信号の振幅スペクトル $|X(\omega, t)|$ と位相情報 $\arg[X(\omega, t)]$ を入力される。雑音除去信号の振幅スペクトル $|D(\omega, t)|$ を以下のように求め、

$$|D(\omega, t)| = G(\omega, t) \cdot |X(\omega, t)|$$

30

さらに、位相情報 $\arg[X(\omega, t)]$ を用いて、複素スペクトル予測値 $D(\omega, t)$ を算出し、さらに、時間領域に変換した雑音除去信号 $d(n)$ を算出して出力する。

【0022】

雑音抽出部 113 は、以下のように、收音信号 $x(n)$ から雑音除去信号 $d(n)$ を差し引いて雑音成分 $b(n)$ を求める (s113)。

$$b(n) = x(n) - d(n)$$

求めた雑音成分 $b(n)$ を出力信号生成部 170 へ出力する。

【0023】

40

特許文献 1 記載の残響除去装置は、收音信号 $x(n)$ には、残響成分だけではなく、雑音成分も含まれるため、高精度な残響成分の推定、除去は困難であった。加えて、雑音成分の特徴を大きく変化させるといった問題もあった。しかし、本実施例では、残響分離部 150 において (詳細は後述)、雑音成分を除去した雑音除去信号 $d(n)$ を用いて、残響除去を行うため、高精度な残響成分の推定、除去ができる。加えて、出力信号生成部 170 において (詳細は後述)、出力信号に雑音成分を加えることができるため、雑音成分の特徴を変化させずに、残響を除去することができるという効果がある。

【0024】

図 8 は時変抑圧係数出力部 130 の構成例を、図 9 は時変抑圧係数出力部 130 の処理の流れを示す。時変抑圧係数出力部 130 は、收音信号 $x(n)$ のパワー時間包絡値 $e(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda(t)$ を出力する (s130)。

50

【 0 0 2 5 】

例えば、時変抑圧係数出力部 1 3 0 は、パワー時間包絡値計算部 1 3 1 と、目的音終端時刻検出部 1 3 3 と、高パワー時刻検出部 1 3 5 と、抑圧係数出力部 1 3 7 を有する。

【 0 0 2 6 】

パワー時間包絡値計算部 1 3 1 は、收音信号 $x(n)$ のパワー時間包絡値 $e(t)$ を計算する (s 1 3 1)。パワー時間包絡値計算部 1 3 1 は、收音信号 $x(n)$ が入力され、收音信号 $x(n)$ を数十ミリ秒程度の短時間窓 (ハミング窓等) で分割し、窓内のパワー時間包絡値 $e(t)$ を求め、これを目的音終端時刻検出部 1 3 3 と、高パワー時刻検出部 1 3 5 へ出力する。例えば、以下のように、パワー時間包絡値を求めることができる。

【 0 0 2 7 】

【 数 2 】

$$e(t) = \sum_{k=1}^N wx(t, k)$$

【 0 0 2 8 】

但し、 N は、ハミング窓等の窓関数 $w(k)$ の窓長を表し、 $k = 1, 2 \dots N$ である。さらに、オーバーラップ処理を行う際のフレーム周期を N/j とすると、 $wx(t, k)$ は以下のように表される。

【 0 0 2 9 】

【 数 3 】

$$wx(t, k) = w(k)x(t \times (N/j) + k) \quad (t=1, 2, \dots, \quad k=1, 2, \dots, N)$$

【 0 0 3 0 】

但し、 $n = t \times (N/j) + k$ であり、 $k = N$ となる度に、パワー時間包絡値 $e(t)$ を出力する。

【 0 0 3 1 】

図 1 0 は、目的音終端時刻検出部及び高パワー時刻検出部における検出例の概要を示す。目的音終端時刻検出部 1 3 3 は、パワー時間包絡値 $e(t)$ を入力され、目的音 (例えば、音声等) の終端部分の時刻 A を検出する (s 1 3 3)。例えば、以下のような検出方法が考えられる。パワー時間包絡値 $e(t)$ の値が閾値 T_A 以上の値から以下の値となった時刻を検出し、目的音の終端部分の時刻 A とする。

【 0 0 3 2 】

高パワー時刻検出部 1 3 5 は、パワー時間包絡値 $e(t)$ を入力され、高パワーを有する收音信号 $x(n)$ が発生した直後の時刻 B を検出する (s 1 3 5)。例えば、以下のような検出方法が考えられる。パワー時間包絡値 $e(t)$ の値が閾値 T_B 以上の値から以下の値となった時刻を検出し、高パワーを有する收音信号 $x_m(n)$ が発生した直後の時刻 B とする。このような方法により時刻 A, B を検出することで、図 2 (i) の (A)、(B) の時刻、あるいは、その近傍の時刻を検出することができる。なお、目的音終端時刻 A 及び高パワー時刻を検出する方法は、上記方法に限られるものではなく、例えば、L.R. Rabiner and M.R. Sambur, "An algorithm for determining the endpoints of isolated utterances," Bell Syst. Tech. J., vol.54, no.2, pp.297-315, Feb. 1975. やその他従来からある発話区間検出技術及び利用者による目視等で判定してもよい。

【 0 0 3 3 】

抑圧係数出力部 1 3 7 は、時刻 A または B を入力されたときに、予め定められた時変抑圧係数を出力する (s 1 3 7)。例えば、目的音の終端部分の時刻 A が入力された場合には、時変抑圧係数として、1 を基準とし、10ms 毎に a 倍した係数 $a^0, a^1, a^2, a^3, a^4, a^5, a^6, a^7, a^8, a^9$ (例えば、 $a = 0.3162$) を出力し、高パワーを有する收音信号 $x(n)$ が発生した直後の時刻 B が入力された場合には、時変抑圧係数として、10ms 毎に以下の係数 0.6, 0.4, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 を出力するように予め定めてもよい。なお、時変抑圧係数を出力するタイミングや個数や値は、適宜設定することが

10

20

30

40

50

できる。

【 0 0 3 4 】

図 1 1 は、残響分離部 1 5 0 の構成例を、図 1 2 は、残響分離部 1 5 0 の処理の流れを示す。残響分離部 1 5 0 は、雑音除去信号 $d(n)$ 及び時変抑圧係数 $\lambda(t)$ を用いて、残響除去信号 $y(n)$ 及び残響成分 $r(n)$ を求める (s 1 5 0)。残響分離部 1 5 0 は、残響除去部 1 6 2、及び残響抽出部 1 6 3 を有し、残響除去部 1 6 2 は、疑似白色化部 1 5 1 と、線形予測係数算出部 1 5 2 と、後部残響予測部 1 5 3 と、周波数領域変換部 1 5 5 と、後部残響除去部 1 5 9 及び時間領域変換部 1 6 1 を有する。

【 0 0 3 5 】

図 1 3 は、疑似白色化 (Pre-whitening) 部 1 5 1 の構成例を示す。疑似白色化部 1 5 1 は、雑音除去信号 $d(n)$ を用いて、疑似白色化 (Pre-whitening) を施した信号 $d'(n)$ を求める (s 1 5 1)。

10

【 0 0 3 6 】

具体的には、疑似白色化 (Pre-whitening) 部 1 5 1 は、周波数領域変換部 1 5 1 1 と、時間平均化部 1 5 1 2 と、減算部 1 5 1 3 と、時間領域変換部 1 5 1 4 を有する。一例として、Cepstral Mean Subtraction (CMS) を用いて雑音除去信号の疑似白色化を行う。まず、疑似白色化部 1 5 1 の周波数領域変換部 1 5 1 1 が、メモリ等から収音信号 1 分析フレーム分の雑音除去信号 $d(n)$ を読み込む。そして、周波数領域変換部 1 5 1 1 は、短時間フーリエ変換等によって雑音除去信号 $d(n)$ を周波数領域の雑音除去信号 $D(\omega, t)$ に変換して出力する。なお、短時間フーリエ変換によってこの処理を行う場合は、以

20

$$D(\omega, t) = \text{Log}[F[d(n)]]$$

【 0 0 3 7 】

次に、時間平均化部 1 5 1 2 に周波数領域の雑音除去信号 $D(\omega, t)$ が読み込まれ、時間平均化部 1 5 1 2 は、以下の式によって、周波数領域の雑音除去信号 $D(\omega, t)$ の時間平均 $E\{D(\omega, t)\}$ を求め、出力する。

【 0 0 3 8 】

【 数 4 】

$$E\{D(\omega, t)\} = \frac{1}{T} \sum_{t=1}^T D(\omega, t)$$

30

【 0 0 3 9 】

次に、減算部 1 5 1 3 に周波数領域の雑音除去信号 $D(\omega, t)$ とその時間平均 $E\{D(\omega, t)\}$ とが読み込まれ、減算部 1 5 1 3 は、以下の式によって $D'(\omega, t)$ を算出し、出力する。

$$D'(\omega, t) = D(\omega, t) - E\{D(\omega, t)\}$$

【 0 0 4 0 】

次に、時間領域変換部 1 5 1 4 に $D'(\omega, t)$ が読み込まれ、時間領域変換部 1 5 1 4 は、逆フーリエ変換等によってこれらを時間領域に変換し、疑似白色化された雑音除去信号 $d'(n)$ を算出し、出力する。なお、逆フーリエ変換によってこの処理を行う場合は、例えば以下の式を用いる。また、 $\text{invF}[\cdot]$ は逆フーリエ変換関数を、 $\exp[\cdot]$ はネイピア数を底とした指数関数を表す。

40

$$d'(n) = \text{invF}[\exp[D'(\omega, t)]]$$

【 0 0 4 1 】

なお、上述した短時間フーリエ変換関数 $F[\cdot]$ や逆フーリエ変換関数 $\text{invF}[\cdot]$ において窓長 2 5 m s の窓関数を用いた場合、2 5 m s 以内の初期反射成分及び短時間相関を取り除くことができる。

【 0 0 4 2 】

また、本実施例では、時間領域変換部 1 5 1 4 で生成された疑似白色化後の雑音除去信号 $d'(n)$ は、後部残響予測部 1 5 3、周波数領域変換部 1 5 5 及び残響抽出部 1 6 3

50

にも転送される。各離散時間の上記雑音除去信号 $d(n)$ から、当該離散時間 n 直前の短時間区間内の各雑音除去信号と自己相関性を持つ自己相関成分を抑制し、疑似白色化した雑音除去信号 $d'(n)$ を生成する。このような短期的な自己相関を雑音除去信号 $d(n)$ から抑制することは、後部残響を精度良く推定する上で望ましい。

【0043】

第一線形予測係数算出部 152 は、疑似白色化した雑音除去信号 $d'(n)$ を用いて、以下の式 (4) の関係を有する、マルチステップ線形予測モデルの各線形予測係数 $\alpha(p)$ を求める (s152)。詳しくは特許文献 1 に述べられている。

【0044】

【数 5】

10

$$d'(n) = \sum_{p=1}^N \alpha(p) \cdot d'(n-p-D) + e(n) \quad (4)$$

【0045】

なお、 n を離散時間とし、 $d'(n)$ を離散時間 n に対応する雑音除去信号とし、 $\alpha(p)$ を線形予測係数とし、 N を線形予測係数の数とし、 D をステップサイズ (遅延) とし、 $e(n)$ を離散時間 n に対応する予測誤差とする。また、各線形予測係数 $\alpha(p)$ を算出する方法としては、自己相関法 (correlation method) や共分散法 (covariance method) を例示できる。また、MATLAB (登録商標) 等を利用してこの処理を行ってもよい。

【0046】

20

後部残響予測部 153 は、前記疑似白色化した雑音除去信号 $d'(n)$ と各線形予測係数 $\alpha(p)$ から後部残響信号 $r(n)$ を予測する (s153)。後部残響予測部 153 は、線形予測係数 $\alpha(p)$ と疑似白色化後の雑音除去信号 $d'(n)$ をマルチステップ線形予測モデルの線形予測項に代入して得られた線形予測値を、後部残響予測値 $r(n)$ として算出し出力する。本実施例では、式 (4) のマルチステップ線形予測モデルを用いているため、後部残響予測部 153 は、以下の式 (5) に従って後部残響予測値 $r(n)$ を求めて出力する。

【0047】

【数 6】

30

$$r(n) = \sum_{p=1}^N \alpha(p) \cdot d'(n-p-D) \quad (5)$$

【0048】

周波数領域変換部 155 は、疑似白色化した雑音除去信号 $d'(n)$ 及び後部残響信号 $r(n)$ を周波数領域の信号 $D'(\omega, t)$ 及び $R(\omega, t)$ に変換し、周波数領域の雑音除去信号の振幅スペクトル $|D'(\omega, t)|$ と位相情報 $\arg[D'(\omega, t)]$ 、及び、後部残響信号の振幅スペクトル $|R(\omega, t)|$ と位相情報 $\arg[R(\omega, t)]$ を抽出し、出力する (s155)。なお、短時間フーリエ変換によってこの処理を行う場合には、例えば以下の式を用いる。

$$D'(\omega, t) = \text{Log}[F[d'(n)]]$$

$$R(\omega, t) = \text{Log}[F[r(n)]]$$

40

【0049】

後部残響除去部 159 は、周波数領域の雑音除去信号のパワー値から後部残響信号のパワー値を差し引いた値に、前記時変抑圧係数 $\lambda(t)$ を掛け合わせた値を周波数領域の残響除去信号の振幅スペクトル $|Y(\omega, t)|$ として算出する (s159)。例えば、

$$|D'(\omega, t)|^2 - |R(\omega, t)|^2 \cdot \text{const} > 0 \quad (6)$$

但し、const は定数 (例えば const=1) を示す。式 (6) の関係を満たすと判断された場合、後部残響除去部 159 は、以下の式により周波数領域の残響除去信号の振幅スペクトル $|Y(\omega, t)|$ を算出して出力する。

$$|Y(\omega, t)| = \lambda(t) \{ (|D'(\omega, t)|^2 - |R(\omega, t)|^2 \cdot \text{const})^{1/2} \} \quad (7)$$

50

【 0 0 5 0 】

一方、式 (6) の関係を満たさないと判断された場合、後部残響除去部 5 9 は、 $|Y(\omega, t)|$ として 0 又は十分小さい値を出力する。このように、時変抑圧係数 $\lambda(t)$ を用いることで、図 2 (i) における目的音の終端部分の時刻 A や高パワーを有する前記收音信号 $x_m(n)$ が発生した直後の時刻 B においても、効果的に残響成分を抑圧することができ、聴覚上も残響が除去されたように感じられるという効果がある。

【 0 0 5 1 】

時間領域変換部 1 6 1 は、残響除去信号の振幅スペクトル $|Y(\omega, t)|$ と疑似白色化された雑音除去信号の位相情報 $\arg[D'(\omega, t)]$ を用いて時間領域の残響除去信号 $y(n)$ を算出する (s 1 6 1)。例えば、周波数領域の残響除去信号の振幅スペクトル $|Y(\omega, t)|$ から、以下の式 (8) により残響除去信号の複素スペクトル予測値 $Y(\omega, t)$ を求める。なお、 $\exp(\cdot)$ は、ネイピア数を底とした指数関数であり、 j は虚数単位である。

$$Y(\omega, t) = |Y(\omega, t)| \cdot \exp(j \cdot \arg[D'(\omega, t)]) \quad (8)$$

【 0 0 5 2 】

そして、複素スペクトル予測値 $Y(\omega, t)$ を時間領域に変換した残響除去信号 $y(n)$ を算出し、出力する。時間領域への変換は、例えば、以下の式を用いて逆フーリエ変換によって行う。

$$y(n) = \text{invF}[\exp[Y(\omega, t)]]$$

このようにして求めた残響除去信号 $y(n)$ を信号生成部 1 7 0 及び残響抽出部 1 6 3 へ出力する。

【 0 0 5 3 】

残響抽出部 1 6 3 は、雑音除去信号 $d(n)$ から残響除去信号 $y(n)$ を差し引くことによって残響成分 $r(n)$ を求める。

$$r(n) = d(n) - y(n)$$

求めた残響成分 $r(n)$ を出力信号生成部 1 7 0 へ出力する。

【 0 0 5 4 】

出力信号生成部 1 7 0 は、残響除去信号 $y(n)$ 、雑音成分 $b(n)$ 及び残響成分 $r(n)$ を重み付け加算して、出力信号 $z(n)$ を算出する。例えば、以下の式のように算出する。

$$z(n) = y(n) + \alpha \times b(n) + \beta \times r(n)$$

【 0 0 5 5 】

但し、 α 、 β は、予め利用者によって設定されるパラメータであり、 $0 \leq \alpha \leq 1$ 、 $0 \leq \beta \leq 1$ とする。例えば、 $\alpha = 1$ と設定することで雑音成分を維持することができ、また、 $\alpha = 0$ と設定することで、雑音成分を除去することもできる。また α には、残響除去前と、残響除去後の目的音と雑音の比を一定に保つために、以下のような値を自動的に設定することが考えられる。

【 0 0 5 6 】

【 数 7 】

$$\alpha = \sqrt{\frac{E\{y(n)^2\}}{E\{d(n)^2\}}}$$

【 0 0 5 7 】

また、 β を 0 とすることもできるが、例えば目的音が音声等の場合には、実環境における音声にも残響が含まれることを考慮して、より実環境における音声に近い出力信号とするために、 β に小さな値 (例えば、 0.01) を設定することが考えられる。このように出力信号生成部 1 7 0 を設け、残響除去信号に雑音成分を加えることで、收音信号の雑音成分の特徴を変化させずに、残響成分を除去することができる。さらに、残響成分を加えることで、より実環境における音声等に近い出力信号とすることもできる。

【 0 0 5 8 】

但し、出力信号生成部 170 を設けなくとも、本発明の効果（振幅の大きな収音信号の直後、及び、目的音の終端部分に残留する残響成分を抑圧するという効果）を得ることができる。その場合、雑音分離部 110 には雑音抽出部 113 を、残響分離部 150 には、残響抽出部 163 を設けなくともよい。また、残響分離部 150 が出力する残響除去信号 $y(n)$ を出力信号 $z(n)$ として出力する。この場合にも、残響分離部 150 において、雑音成分を除去した雑音除去信号 $d(n)$ を用いて、残響除去を行うため、高精度な残響成分の推定、除去ができる。

【0059】

[比較実験の結果]

従来技術（特許文献 1）と本実施例で残響除去処理の比較実験の結果に基づき、本発明の効果の説明する。図 2（i）は、特許文献 1 記載の残響除去装置を使って残響除去処理を行った結果を、図 2（ii）は、本実施例の残響除去装置を使って残響除去処理を行った結果を示す。なお、本実施例では、目的音終端時刻検出部 133 にて検出した時刻 B から 100ms 後までエネルギーを抑圧した。時変抑圧係数は、1 を基準として、10ms 毎に 0.3162 倍した係数 $0.3162^0, 0.3162^1, 0.3162^2, 0.3162^3, 0.3162^4, 0.3162^5, 0.3162^6, 0.3162^7, 0.3162^8, 0.3162^9$ を設定した。また、高パワー時刻検出部 135 に検出した時刻 A から 80ms 後までエネルギーを抑圧した。時変抑圧係数として、10ms 毎に以下の係数 0.6, 0.4, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 を設定した。本実施例の残響除去装置を用いた場合、より効果的に大きな音の直後、及び、目的音の終端部分の残響が抑圧されていることを見て取ることができる。

【0060】

図 14（A）は、雑音と残響を含む収音信号のスペクトログラムを、図 14（B）は、特許文献 1 の出力信号のスペクトログラムを、図 14（C）は、本実施例の出力信号のスペクトログラムを示す。なお、本実施例では、出力信号生成部 170 におけるパラメータは、 $\alpha = 1$ 、 $\beta = 0$ と設定した。また、雑音除去部 111 では、参考文献 1 記載の雑音除去方法を用いた。図 14（A）において、雑音は白色性であるため、周波数特性が一様であることが見て取れる。一方、図 14（B）において、図 14（A）と比べて処理後は雑音の特性が著しく変化していることがわかる。図 14（C）において、雑音の特性が変化していないことが分かる。

【0061】

また、特許文献 1 の出力信号と本実施例の出力信号の残響除去性能を比較するため、2 名の被験者により聴取実験を行った。実験結果として、本実施例の出力信号は、雑音の特徴に大きな変化を加えずに、目的音の終端、及び高パワーを有する収音信号が発生した直後において残響成分が抑圧され、より品質のよい残響除去信号が得られていることが分かった。

【0062】

[変形例 1]

実施例 1 と異なる部分のみ説明する。図 15 は変形例 1 の残響除去装置 200 の構成例を、図 16 は変形例 1 の処理の流れを示す。残響除去装置 200 は、残響分離部 250 及び時変抑圧部 130 を有する。

【0063】

残響分離部 250 は、収音信号 $x(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める（S250）。つまり、雑音除去信号 $d(n)$ に代えて、収音信号 $x(n)$ を利用して残響成分を除去する。この場合、残響除去装置 200 は、雑音分離部、出力信号生成部を有さなくともよい。また、残響分離部 250 は、残響抽出部を有さなくともよい。このような構成においても、振幅の大きな収音信号の直後、及び、目的音の終端部分に残留する残響成分を抑圧するという効果を得ることができる。雑音が少ない場合等には、演算量等を減らすことができ有効である。なお、このとき、残響分離部 250 は残響除去信号 $y(n)$ を出力信号 $z(n)$ として出力し、残響成分 $r(n)$ は出力しない。

10

20

30

40

50

【実施例 2】

【0064】

図 17 は、実施例 2 の残響除去装置 300 の構成例を示す。以下、同様の構成要素については、同一の符号を付す。実施例 2 の残響除去装置 300 は、雑音分離部 310 と、時変抑圧係数出力部 330 と、残響分離部 350 と、出力信号生成部 370 と、を有する。M 個 (M ≥ 2) のセンサで観測された M 個のチャネル m ($m = 1, \dots, M$) の收音信号 $x_m(u)$ (但し、 u は、連続時間を表す) が所定の標本化周波数でサンプリングされ、收音信号 $x_m(n)$ (但し、 n は、離散時間を表す) が生成される。收音信号 $x_m(n)$ は、それぞれメモリ等に格納される。

【0065】

雑音分離部 310 は、收音信号 $x_m(n)$ を用いて、雑音除去信号 $d_m(n)$ を算出し、收音信号 $x_m(n)$ から雑音除去信号 $d_m(n)$ を差し引いて雑音成分 $b_m(n)$ を求める。各チャネル m において、実施例 1 と同様の処理を行う。

【0066】

時変抑圧係数出力部 330 は、收音信号 $x_m(n)$ のパワー時間包絡値 $e_m(t)$ (但し、 t はフレーム番号を表す) の変化に依存して、時変抑圧係数 $\lambda_m(t)$ を出力する。各チャネル m において、実施例 1 と同様の処理を行う。

【0067】

残響分離部 350 は、雑音除去信号 $d_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y_m(n)$ を求める。各チャネル m において、実施例 1 と同様の処理を行う。

【0068】

出力信号生成部 370 は、残響除去信号 $y_m(n)$ 、雑音成分 $b_m(n)$ 及び残響成分 $r_m(n)$ を重み付け加算して、出力信号 $z_m(n)$ を算出する。各チャネル m において、実施例 1 と同様の処理を行う。

$$z_m(n) = y_m(n) + \alpha \times b_m(n) + \beta \times r_m(n) \quad (9)$$

但し、 α 、 β は、予め利用者によって設定されるパラメータであり、 $0 \leq \alpha \leq 1$ 、 $0 \leq \beta \leq 1$ とする。このような構成とすることによって、各チャネルからの出力信号は、実施例 1 と同様の効果が得られる。なお、出力信号生成部 370、または、雑音分離部 310 と出力信号生成部 370 を有さない場合であっても、実施例 1、または、変形例 1 と同様の効果が得られる。

【0069】

[変形例 1]

実施例 2 と異なる部分についてのみ説明する。図 18 は、変形例 1 の構成例を示す。変形例 1 では、残響分離部 350' 及び出力信号生成部 370' の処理が異なる。図 19 は、残響分離部 350' の構成例を示す。残響分離部 350' は、雑音除去信号 $d_m(n)$ 及び時変抑圧係数 $\lambda_m(t)$ を用いて、残響除去信号 $y(n)$ 及び残響成分 $r_m(n)$ を求める。残響分離部 350' は、残響除去部 362、残響抽出部 363、遅延量算出部 365 及び遅延調節部 367 を有する。残響除去部 362 では、各チャネル m において、実施例 1 と同様の処理を行い、 m 個の残響除去信号 $y_m(n)$ を出力する。

【0070】

遅延量算出部 365 に、各チャネルの残響除去信号 $y_1(n) \dots y_M(n)$ が入力される。そして、遅延量算出部 365 は、残響除去信号 $y_1(n) \dots y_M(n)$ のチャネル間相互相関を極大にする残響除去信号の遅延量 $\tau_1 \dots \tau_M$ を、各チャネルについて決定する。以下にこの具体例を示す。まず、遅延量算出部 365 は、入力された分析フレーム内の残響除去信号 $y_1(n) \dots y_M(n)$ に対し、以下の式のようなチャネル間相関関数 $A_m(\tau)$ の関数値を求める。なお、 $E\{\cdot\}$ は平均演算子である。

$$A_m(\tau) = E\{y_1(n) \cdot y_m(n+\tau)\}$$

【0071】

次に、遅延量算出部 365 は、各 m について、チャネル間相関関数 $A_m(\tau)$ を極大 (例えば最大) とする τ を τ_m として求める。例えば、チャネル間相関関数 $A_m(\tau)$ を最大とす

10

20

30

40

50

る τ を τ_m とする場合には、遅延量算出部 365 は、

$$\tau_m = \max\{A_m(\tau)\}$$

を算出して出力する。なお、 $\max\{\cdot\}$ は $\cdot$ の最大値を検出する。また、 τ_m は、チャネル m の残響除去信号の遅延量であり、遅延量には $\tau_m = 0$ も含む。

【0072】

各遅延量 $\tau_1 \dots \tau_M$ と、残響除去信号 $y_1(n) \dots y_M(n)$ とが、遅延調節部 367 に入力される。そして、遅延調節部 367 の遅延部 368 は、各チャネルの残響除去信号 $y_1(n) \dots y_M(n)$ を、それぞれ遅延量 $\tau_1 \dots \tau_M$ だけ遅延させて、 $y_1(n + \tau_1) \dots y_M(n + \tau_M)$ を算出して出力する。

【0073】

次に、 $y_1(n + \tau_1) \dots y_M(n + \tau_M)$ が、遅延補正部 369 に入力され、遅延補正部 369 は、以下の式に従い、 $y_1(n + \tau_1) \dots y_M(n + \tau_M)$ の平均を算出する。複数チャネルの場合には、雑音分離部 350' が、この平均を新たに残響除去信号 $y(n)$ として出力する。各チャネルの後部残響信号に含まれる誤差成分は統計的に独立であると過程した場合、この操作により誤差を抑圧できることになる。

【0074】

【数8】

$$y(n) = \frac{1}{M} \sum_{m=1}^M y_m(n + \tau_m)$$

【0075】

残響抽出部 363 は、雑音除去信号 $d_m(n)$ から残響除去信号 $y(n)$ を差し引くことによって残響成分 $r_m(n)$ を求める。

$$r_m(n) = d_m(n) - y(n)$$

求めた残響成分 $r_m(n)$ を出力信号生成部 370' へ出力する。

【0076】

出力信号生成部 370' は、残響除去信号 $y(n)$ 、雑音成分 $b_m(n)$ 及び残響成分 $r_m(n)$ を重み付け加算して、出力信号 $z_m(n)$ を算出する。例えば、以下の式のように算出する。

$$z_m(n) = y(n) + \alpha \times b_m(n) + \beta \times r_m(n) \quad (10)$$

但し、 α 、 β は、予め利用者によって設定されるパラメータである。このような構成とすることによって、実施例 1 と同様の効果が得られる。また、式 (10) に代えて、以下の式を用いることにより、 m チャネルの収音信号から 1 チャネルの出力信号 $z(n)$ を得ることができる。

【0077】

【数9】

$$z(n) = y(n) + \frac{1}{M} \times \alpha \times \sum_{m=1}^M b_m(n) + \frac{1}{M} \times \beta \times \sum_{m=1}^M r_m(n)$$

【0078】

なお、出力信号生成部 370'、または、雑音分離部 310 と出力信号生成部 370' を有さない場合であっても、実施例 1、または、変形例 1 と同様の効果が得られる。

【0079】

上述した残響除去装置は、コンピュータにより機能させることもできる。この場合はコンピュータに、目的とする装置（各種実施例で図に示した機能構成をもつ装置）として機能させるためのプログラム、又はその処理手順（各実施例で示したもの）の各過程をコンピュータに実行させるためのプログラムを、CD-ROM、磁気ディスク、半導体記憶装置などの記録媒体から、あるいは通信回線を介してそのコンピュータ内にダウンロードし、そのプログラムを実行させればよい。

【産業上の利用可能性】

10

20

30

40

50

【 0 0 8 0 】

本発明に係る残響除去装置は、さまざまな音響信号処理システムの要素技術として利用することが可能である。例えば、発話された音声信号の残響除去処理が要素技術として性能向上に寄与できるような音響信号処理システムには、残響環境下で收音した音声を整音するためのシステムや、残響環境下で残響を除去することで聞き取り易さを向上させる補聴器や、残響除去により音声の明瞭性を向上させるＴＶ会議システム等の通信システム等が挙げられる。実環境で收音された信号には、常に残響（反射音）が含まれるが、本発明に係る残響除去装置は、そのような環境で用いられることを想定している。

【図面の簡単な説明】

【 0 0 8 1 】

10

【図 1】特許文献 1 記載の残響除去装置 5 0 の構成例を示す。

【図 2】図 2（ i ）は、特許文献 1 記載の残響除去装置を使って残響除去処理を行った結果を、図 2（ ii ）は、本実施例の残響除去装置を使って残響除去処理を行った結果を示す。

【図 3】実施例 1 の残響除去装置 1 0 0 の構成例を示す。

【図 4】実施例 1 の処理の流れを示す。

【図 5】雑音分離部 1 1 0 の構成例を示す。

【図 6】雑音分離部 1 1 0 の構成例の流れを示す。

【図 7】雑音除去部 1 1 1 の構成例を示す。

【図 8】時変抑圧係数出力部 1 3 0 の構成例を示す。

20

【図 9】時変抑圧係数出力部 1 3 0 の処理の流れを示す。

【図 1 0】目的音終端時刻検出部及び高パワー時刻検出部における検出例の概要を示す。

【図 1 1】残響分離部 1 5 0 の構成例を示す。

【図 1 2】残響分離部 1 5 0 の処理の流れを示す。

【図 1 3】疑似白色化（Pre-whitening）部 1 5 1 の構成例を示す。

【図 1 4】図 1 4（ A ）は雑音と残響を含む收音信号のスペクトログラムを、図 1 4（ B ）は特許文献 1 の出力信号のスペクトログラムを、図 1 4（ C ）は本実施例の出力信号のスペクトログラムを示す。

【図 1 5】実施例 1 の変形例 1 の残響除去装置 2 0 0 の構成例を示す。

【図 1 6】実施例 1 の変形例 1 の処理の流れを示す。

30

【図 1 7】実施例 2 の残響除去装置 3 0 0 の構成例を示す。

【図 1 8】実施例 2 の変形例 1 の残響除去装置 3 0 0 ' の構成例を示す。

【図 1 9】実施例 2 の変形例 1 の残響分離部 3 5 0 ' の構成例を示す。

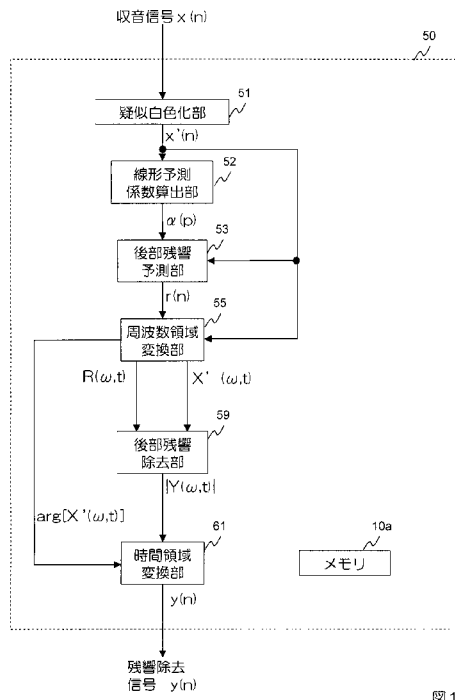
【符号の説明】

【 0 0 8 2 】

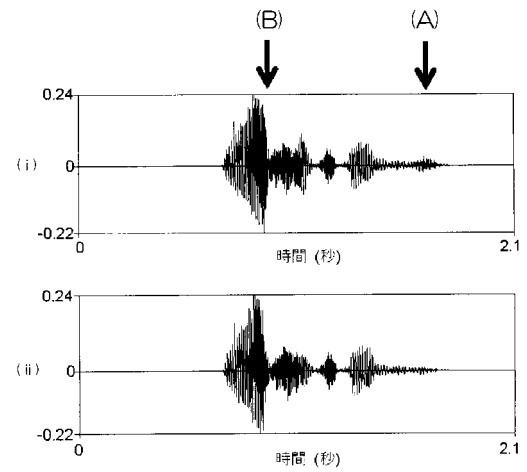
1 0 0 , 2 0 0 , 3 0 0	残響除去装置	1 1 0 , 3 1 0	雑音分離部
1 3 0 , 2 3 0 , 3 3 0	時変抑圧係数出力部		
1 5 0 , 2 5 0 , 3 5 0 , 3 5 0 '	残響分離部		
1 7 0 , 3 7 0 , 3 7 0 '	出力信号生成部		
1 3 1	パワー時間包絡値計算部	1 3 3	目的音終端時刻検出部
1 3 5	高パワー時刻検出部	1 3 7	抑圧係数出力部

40

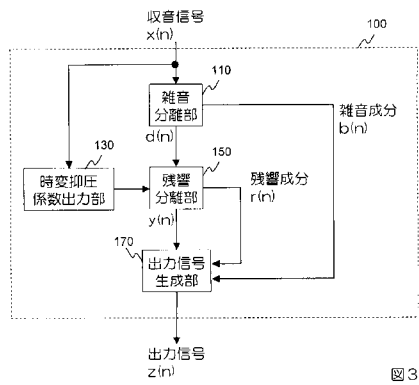
【 図 1 】



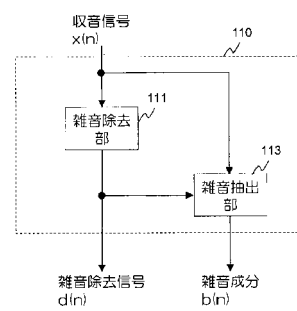
【 図 2 】



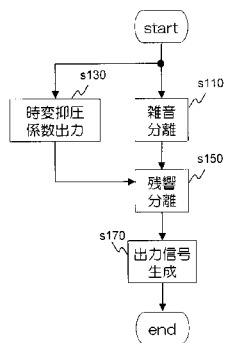
【 図 3 】



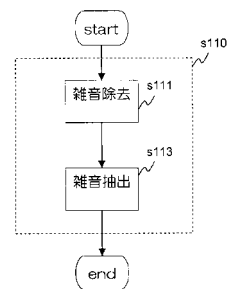
【 図 5 】



【 図 4 】



【 図 6 】



【図 7】

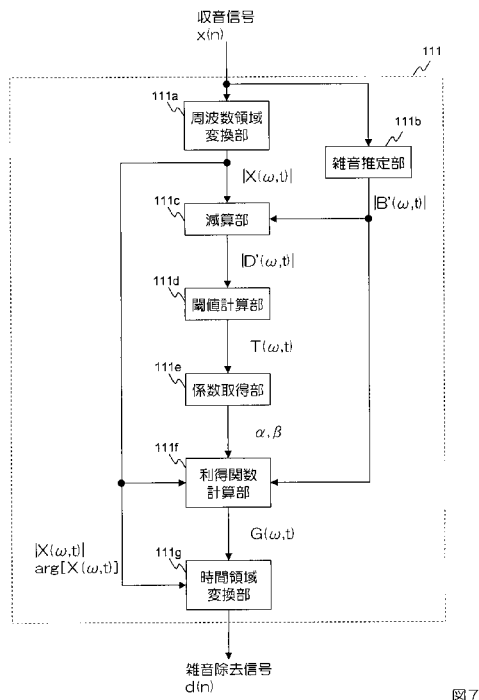


図 7

【図 8】

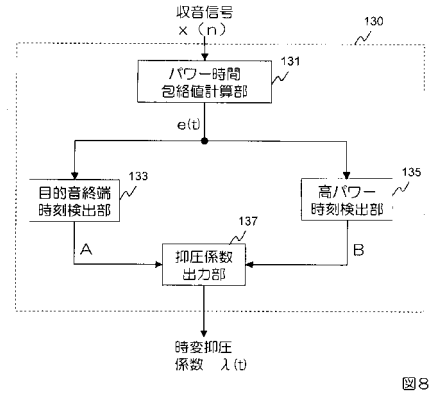


図 8

【図 9】

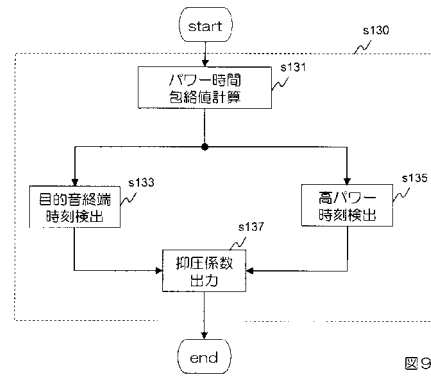


図 9

【図 10】

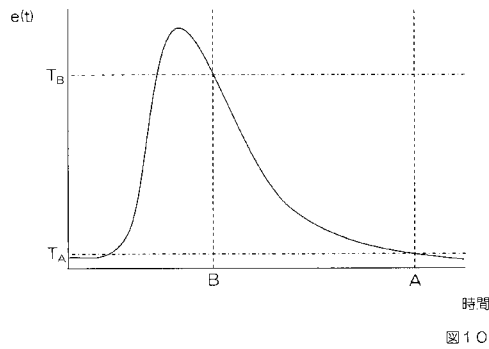


図 10

【図 11】

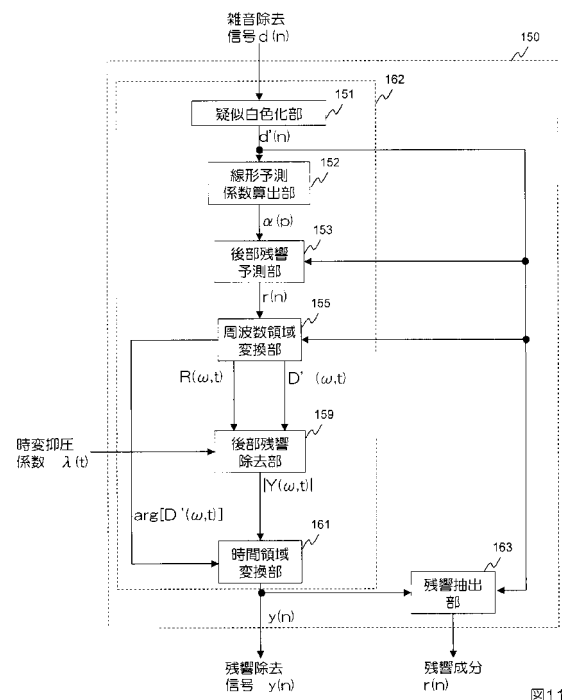


図 11

【図 1 2】

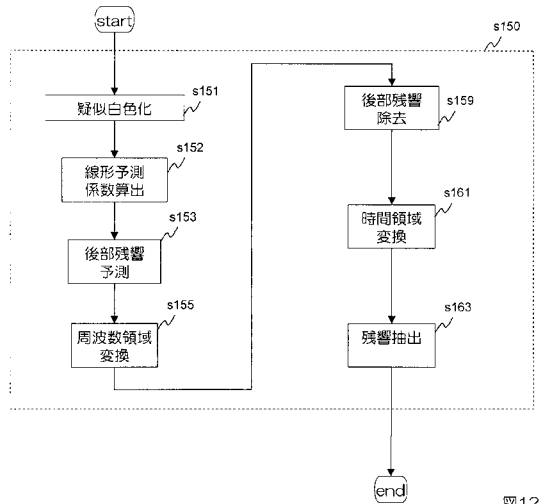


図12

【図 1 3】

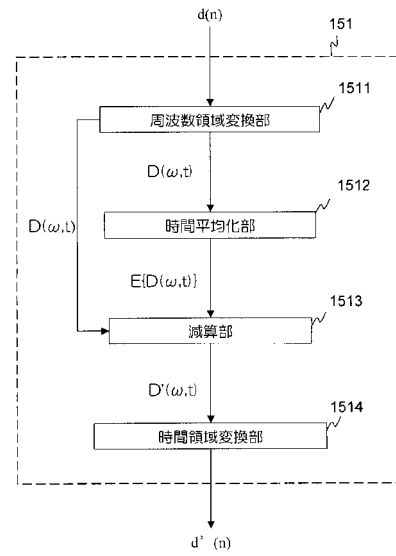


図13

【図 1 5】

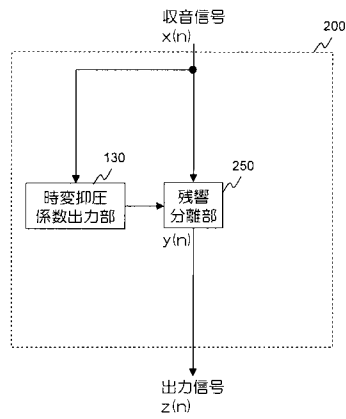


図 15

【図 1 7】

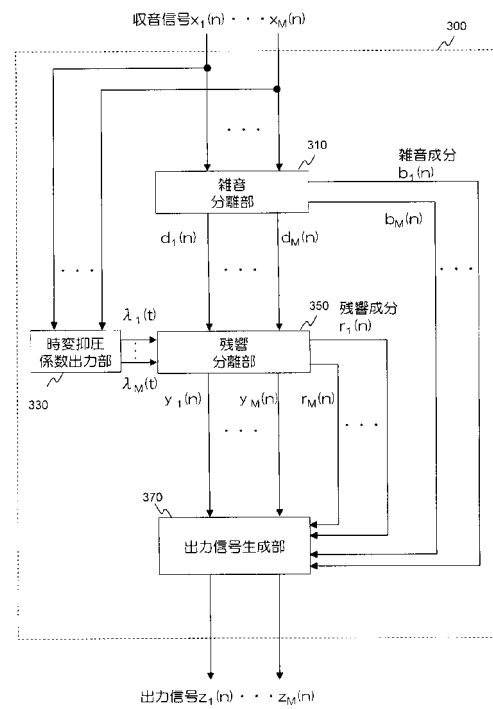


図17

【図 1 6】

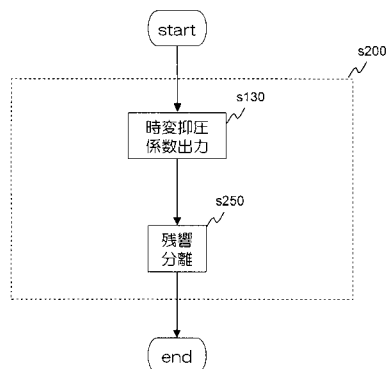


図 16

【図 18】

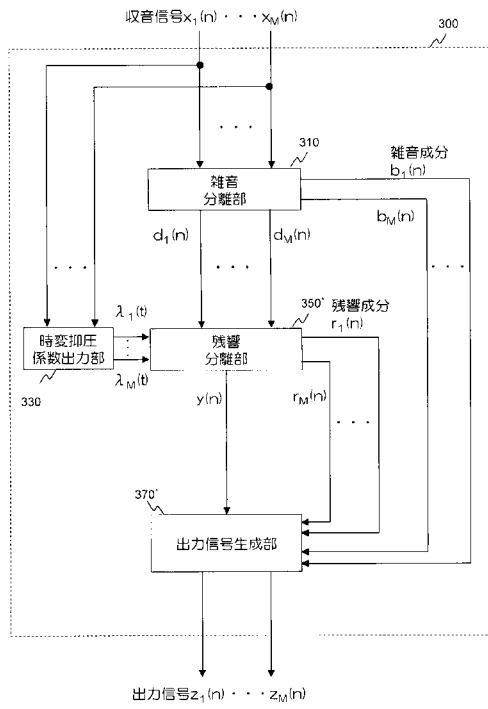


図18

【図 19】

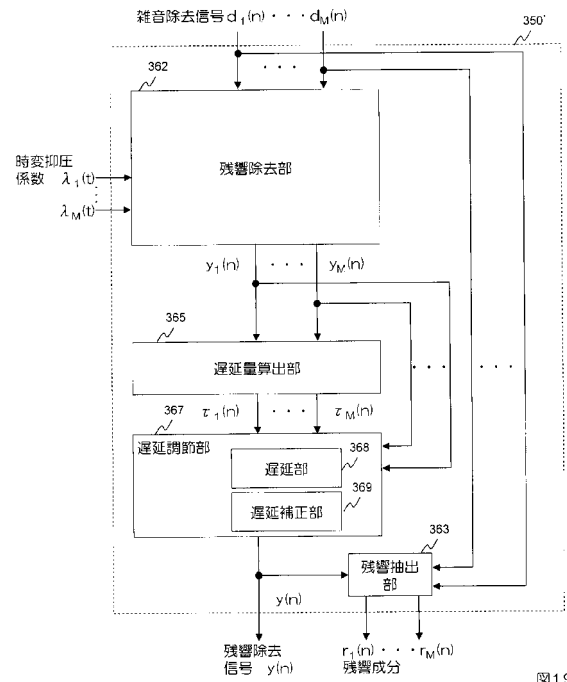


図19

【 図 1 4 】

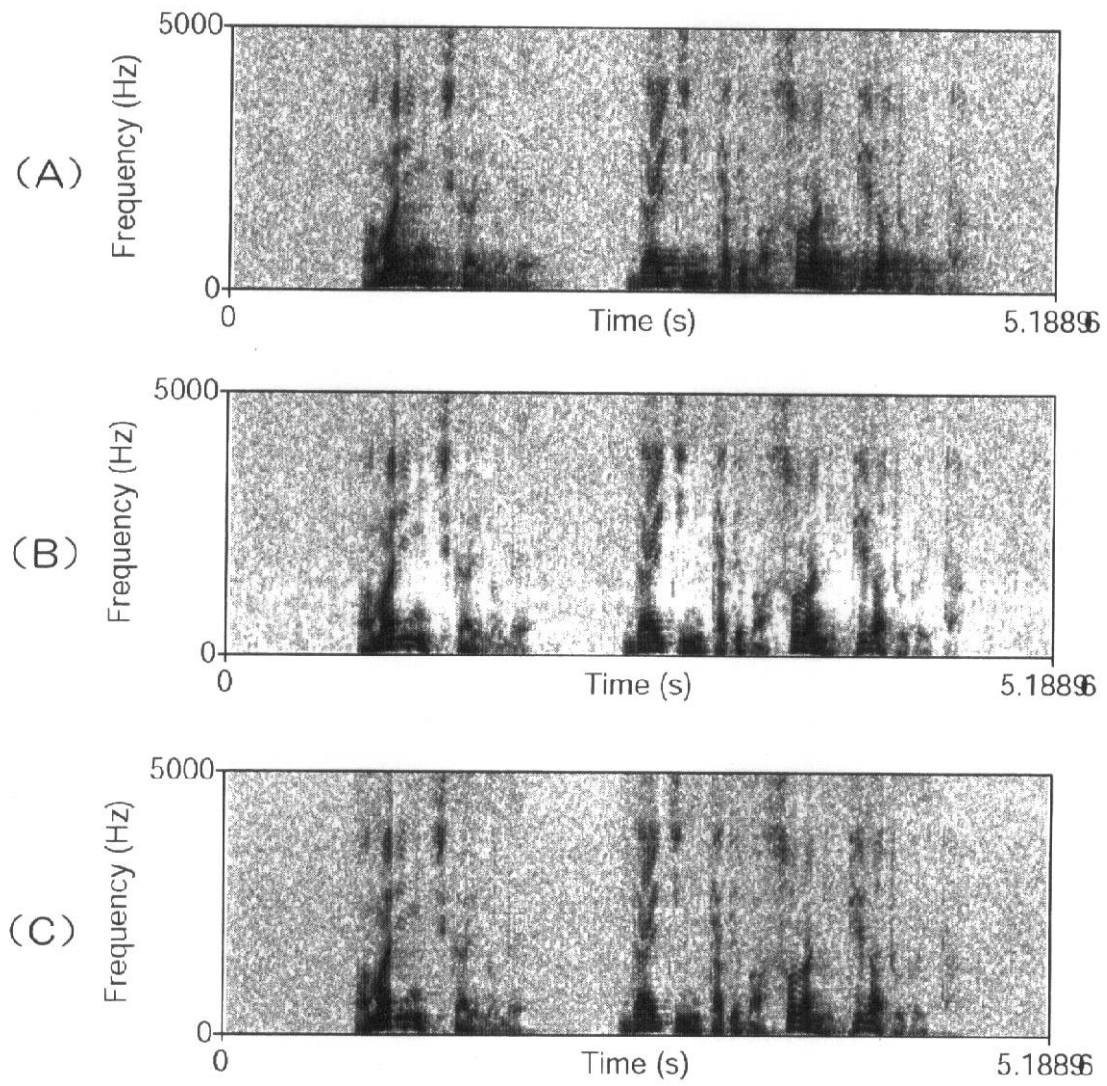


図 1 4

フロントページの続き

(72)発明者 三好 正人

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

Fターム(参考) 5D020 BB07