(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4505597号
(P4505597)
(45)発行日　平成22年7月21日(2010.7.21)　(24)登録日　平成22年5月14日(2010.5.14)

(51)Int.Cl.　　F I
G10L 21/02 (2006.01)　　G10L 21/02 101B
G10L 15/20 (2006.01)　　G10L 15/20 370D

請求項の数 7　(全 12 頁)

(21)出願番号　特願2004-227916(P2004-227916)
(22)出願日　平成16年8月4日(2004.8.4)
(65)公開番号　特開2006-47639(P2006-47639A)
(43)公開日　平成18年2月16日(2006.2.16)
審査請求日　平成19年7月19日(2007.7.19)

（出願人による申告）平成１６年４月１日付け、支出負担行為担当官総務省大臣官房会計課企画官、研究テーマ「ネットワーク・ヒューマン・インターフェースの総合的な研究開発（携帯電話等を用いた多言語自動翻訳システムの研究開発）」に関する委託研究、産業活力再生特別措置法第３０条の適用を受ける特許出願

(73)特許権者　393031586
株式会社国際電気通信基礎技術研究所
京都府相楽郡精華町光台二丁目２番地２
(74)代理人　100099933
弁理士　清水　敏
(72)発明者　黒岩　眞吾
京都府相楽郡精華町光台二丁目２番地２
株式会社国際電気通信基礎技術研究所内
(72)発明者　遠藤　俊樹
京都府相楽郡精華町光台二丁目２番地２
株式会社国際電気通信基礎技術研究所内
(72)発明者　中村　哲
京都府相楽郡精華町光台二丁目２番地２
株式会社国際電気通信基礎技術研究所内

最終頁に続く

(54)【発明の名称】雑音除去装置

(57)【特許請求の範囲】
【請求項１】
周波数帯域で表された複数通りの雑音モデルを記憶するための雑音モデル記憶手段と、
入力される信号をフレームごとに周波数領域に変換するための周波数変換手段と、
所定の第１の周波数帯域において、前記周波数変換手段により周波数領域に変換された前記信号のスペクトル形状に最も近いスペクトル形状を有する雑音モデルを、前記雑音モデル記憶手段に記憶された前記複数の雑音モデルからフレームごとに選択するための雑音モデル選択手段と、
前記信号と、前記雑音モデル選択手段により選択された雑音モデルとの所定の第２の周波数帯域の周波数成分に基づいて、前記選択された雑音モデルのレベルをフレームごとに推定するためのレベル推定手段と、
前記選択された雑音モデルの周波数成分を前記レベル推定手段により推定されたレベルにしたがって変換したものを、前記周波数変換手段により周波数領域に変換された前記信号の周波数成分からフレームごとに減算するための減算手段と、
前記減算手段の出力を周波数帯域から時間領域に逆変換するための時間変換手段とを含む、雑音除去装置。
【請求項２】
前記第２の周波数帯域は、前記第１の周波数帯域よりも広く選ばれている、請求項１に記載の雑音除去装置。
【請求項３】

前記第１の周波数帯域は可変であり、

前記第１の周波数帯域を指定するための帯域指定手段をさらに含む、請求項１または請求項２に記載の雑音除去装置。

【請求項４】

前記周波数変換手段は、前記信号に対し、所定の周波数間隔ごとに周波数成分を算出するための離散的周波数成分算出手段を含み、

前記雑音モデル選択手段は、以下の式にしたがって雑音モデル^N_gを選択するための手段を含む、請求項３に記載の雑音除去装置。

【数１】

$$g = \arg\min_{i}\left\{\sum_{k=h1}^{h2}\left(\left|Y(t,kf_0)\right| - \left|\hat{N}_i(kf_0)\right|\right)^2\right\}$$

ただしｔは時刻、f_0は前記離散的周波数成分算出手段における前記所定の周波数間隔、$Y(t,kf_0)$は時刻ｔにおける、周波数ｋf_0での前記信号の周波数成分、^$N_i(kf_0)$は時刻ｔにおける、周波数ｋf_0での前記雑音モデル^N_iの周波数成分、h_1は$h_1 f_0$＝前記第１の周波数帯域の下限となるような整数、h_2は$h_2 f_0$＝前記第１の周波数帯域の上限となるような整数、をそれぞれ表す。

【請求項５】

前記レベル推定手段は、前記信号$|Y(t,kf_0)|$と前記選択された雑音モデル^$Ng(kf_0)$とから、以下の式にしたがって前記選択された雑音モデルのレベル^$\alpha(t)$をフレームごとに推定するための手段を含む、請求項４に記載の雑音除去装置。

【数２】

$$\hat{\alpha}(t) = \arg\min_{\alpha}\left\{\sum_{k=j1}^{j2}\left(\left|Y(t,kf_0)\right| - \alpha\left|\hat{N}_g(kf_0)\right|\right)^2\right\}$$

ただし、j_1は$j_1 f_0$＝前記第２の周波数帯域の下限となるような整数、j_2は$j_2 f_0$＝前記第２の周波数帯域の上限となるような整数、をそれぞれ表す。

【請求項６】

前記減算手段は、以下の式にしたがって前記信号から雑音をフレームごとに除去した信号^$S(t,kf_0)$を出力する、請求項５に記載の雑音除去装置。

【数３】

$$\left|\hat{S}(t,kf_0)\right| = \left|Y(t,kf_0)\right| - \hat{\alpha}(t)\left|\hat{N}_g(kf_0)\right|$$

ただし^$S(t,kf_0)$は時刻ｔにおける信号^Ｓの周波数ｋf_0における周波数成分を、^$Ng(kf_0)$は前記選択された雑音モデル^Ｎｇの周波数ｋf_0における周波数成分を、それぞれ表す。

【請求項７】

各々が複数通りの雑音モデルを含む複数個の信号源プロフィール情報を記憶するための手段と、

前記複数個の信号源プロフィール情報のうちのいずれかを、ユーザの指定により選択して、当該選択された信号源プロフィール情報に含まれる複数個の雑音モデルを前記雑音モデル記憶手段に格納するための手段とをさらに含む、請求項１～請求項６のいずれかに記載の雑音除去装置。

【発明の詳細な説明】

【技術分野】

【０００１】

この発明は雑音除去装置に関し、特に、風雑音などのように非定常的な雑音を除去するための装置に関する。

【背景技術】

【０００２】

最近の電子機器技術の発達はめざましく、種々の装置が高性能になり、かつ小型化され

た。典型的な例がビデオカメラである。かつてはビデオカメラは、携帯できる形式のものであってもかなりの大きさであったが、最近のビデオカメラは非常に小さく、軽くなっている。また最近のビデオカメラは値段が安くなり、その結果多くの人がビデオカメラを入手し、様々なところにビデオカメラを持っていく機会が増加した。その結果、野外での撮影機会も増加した。

【０００３】

野外での撮影の問題として、風雑音の影響がある。風雑音はマイクに拾われやすく、その結果音質が劣化するという問題がある。

【０００４】

従来、風雑音対策として行なわれていた手法の一つは、マイクに風防を付けるなど，ハードウェアによるものである。しかしそのような手法には限界があり、風雑音を十分効果的に除去することはできない。

【０００５】

一方、風雑音に限らず、音声信号の雑音除去の一般的手法にスペクトルサブトラクション法（ＳＳ法）と呼ばれる手法が存在する。図４を参照して、ＳＳ法の概念を説明する。一般的にはマイクで得られる信号は、目的となる音声信号に雑音信号が重畳されたものとなる。そこで、例えば無音区間の音声信号から雑音信号を推定し、音声信号からこの雑音信号を除去することで、雑音のない音声信号を得る。

【０００６】

ＳＳ法では、まず図４の上段に示されるように雑音を含む音声信号１００の周波数スペクトルを得て、これから、無音区間で推定された雑音１０２の周波数スペクトルを減算し、図４の下段に示す信号１１０を得る。

【０００７】

図５に、従来のＳＳ法を用いる雑音除去装置の構成を示す。図５を参照して、従来の雑音除去装置１２０は、観測信号ｙ(i)に対し短時間ＦＦＴ（Short-Time Fourier Transformation）処理を実行して周波数領域に変換し、短時間音声パワースペクトルを示す、離散的な例えば１２８個の周波数成分$|Y(t,kf_0)|$８２（ｋ＝０～１２７）と位相成分$\varphi_y(kf_0)$８０とを出力するためのＳＴＦＴ処理部６８と、周波数成分８２と、無音区間から推定された雑音成分$|\hat{}N(kf0)|$１３２とから、以下の式にしたがうＳＳ法によって雑音を除去した信号$|\hat{}S(t,kf0)|$１３４を出力するためのＳＳ処理部７２とを含む。

【０００８】

【数１】

$$\left|\hat{S}(t,kf_0)\right| = \left|Y(t,kf_0)\right| - \alpha\left|\hat{N}(kf_0)\right| \qquad (1)$$

$$(\alpha \approx 1.0)$$

雑音除去装置１２０はさらに、ＳＳ処理部７２の出力する信号１３４に対して、ＳＴＦＴ処理部６８の出力する位相成分８０を位相情報として用いてＩＦＦＴ（Inverse FFT）処理を行なうためのＩＦＦＴ処理部７４と、ＩＦＦＴ処理部７４の出力に対しインバースウィンドイング処理を実行するためのインバースウィンドイング処理部７６と、インバースウィンドイング処理部７６の出力に基づいて波形合成を行ない、信号$\hat{}s(i)$を出力するための波形合成処理部７８とを含む。

【０００９】

ＳＴＦＴ処理部６８は、観測信号ｙを所定時間ごとにずらしながら所定長のフレームｙ(i)にデジタル信号化するためのフレーム化部１４０と、フレーム化部１４０から出力される各フレームｙ(i)に対し、所定の時間窓を掛ける処理を行なうためのウィンドイング処理部１４２と、ウィンドイング処理部１４２から出力される各フレームの観測信号に対してＦＦＴ（Fast Fourier Transform）処理を実行し、各フレームの位相成分８０および周波数成分８２を出力するためのＦＦＴ処理部１４４とを含む。

【００１０】

波形合成処理部７８は、インバースウィンドイング処理部７６の出力に対し窓関数を乗

ずるためのウィンドイング処理部１５０と、ウィンドイング処理部１５０の出力に基づいて信号^ｓ(i)を合成するための合成処理部１５２とを含む。

【００１１】

雑音除去装置１２０は概略以下のように動作する。観測信号ｙはフレーム化部１４０によりフレーム化される。フレーム化された観測信号ｙ(i)に対してウィンドイング処理部１４２が窓掛け処理を行なう。窓掛けされた観測信号に対してＦＦＴ処理部１４４がＦＦＴ処理を行ない、位相成分８０および周波数成分８２を出力する。

【００１２】

ＳＳ処理部７２は、周波数成分８２から音声信号の無音区間より推定された雑音成分１３２を減算し、信号１３４としてＩＦＦＴ処理部７４に与える。この処理により、式（１）にしたがって観測信号から雑音成分の推定値が減算される。ＩＦＦＴ処理部７４は、信号１３４に対し位相成分８０を位相情報としてＩＦＦＴ処理を実行し、インバースウィンドイング処理部７６に与える。以下、インバースウィンドイング処理部７６、ウィンドイング処理部１５０、および合成処理部１５２により、雑音の除去された信号^ｓ(i)が合成される。

【発明の開示】

【発明が解決しようとする課題】

【００１３】

ＳＳ法は、簡単なアルゴリズムで観測信号の雑音を効率的に除去できる。しかし、ＳＳ法では雑音成分が定常的であることが仮定されているため、風雑音のように非定常的な雑音下では、雑音成分の予測の誤差が大きく、雑音の引きすぎまたは消し残りが発生する可能性が高いという問題がある。

【００１４】

したがって本発明の目的は、風雑音のような非定常雑音下において、雑音を精度良く除去できる雑音除去装置を提供することである。

【００１５】

本発明の他の目的は、風雑音のような非定常雑音下において、雑音レベルの変化に追従し、雑音を精度良く除去できる雑音除去装置を提供することである。

【００１６】

本発明のさらに他の目的は、風雑音のような非定常雑音下において、雑音レベルの変化に追従し、さらに雑音信号のスペクトル形状を考慮して雑音を精度良く除去できる雑音除去装置を提供することである。

【課題を解決するための手段】

【００１７】

本発明の第１の局面に係る雑音除去装置は、周波数帯域で表された複数通りの雑音モデルを記憶するための雑音モデル記憶手段と、入力される信号をフレームごとに周波数領域に変換するための周波数変換手段と、所定の第１の周波数帯域において、周波数変換手段により周波数領域に変換された信号のスペクトル形状に最も近いスペクトル形状を有する雑音モデルを、雑音モデル記憶手段に記憶された複数の雑音モデルからフレームごとに選択するための雑音モデル選択手段と、信号と、雑音モデル選択手段により選択された雑音モデルとの所定の第２の周波数帯域の周波数成分に基づいて、選択された雑音モデルのレベルをフレームごとに推定するためのレベル推定手段と、選択された雑音モデルの周波数成分をレベル推定手段により推定されたレベルにしたがって変換したものを、周波数変換手段により周波数領域に変換された信号の周波数成分からフレームごとに減算するための減算手段と、減算手段の出力を周波数帯域から時間領域に逆変換するための時間変換手段とを含む。

【００１８】

予め複数通りの雑音モデルを雑音モデル記憶手段に記憶させておく。入力される信号をフレームごとに周波数領域に変換し、その第１の周波数帯域のスペクトル形状に最も近いスペクトル形状を持つ雑音モデルをフレームごとに選択する。さらに、信号と、選択され

た雑音モデルとの第２の周波数帯域における周波数成分に基づいて、雑音モデルのレベルを推定し、雑音モデルを当該推定されたレベルに変換し、もとの信号からフレームごとに除算する。こうした構成により、フレームごとに、信号の第１の周波数帯域のスペクトル形状に最もよく似た雑音モデルを用いてＳＳ法による雑音除去が行なえる。フレームごとに入力信号の雑音と最も良く似た雑音モデルを用いて信号から除算するので、信号に含まれる雑音が非定常なものでもその変化によく追従し、効率的に、かつ精度よく雑音を除去することができる。

【００１９】

好ましくは、第２の周波数帯域は、第１の周波数帯域よりも広く選ばれている。

【００２０】

さらに好ましくは、第１の周波数帯域は可変であり、雑音除去装置は、第１の周波数帯域を指定するための帯域指定手段をさらに含む。

【００２１】

周波数変換手段は、信号に対し、所定の周波数間隔ごとに周波数成分を算出するための離散的周波数成分算出手段を含み、雑音モデル選択手段は、以下の式にしたがって雑音モデル^N_gを選択するための手段を含んでもよい。

【００２２】

【数２】

$$g = \arg\min_i \left\{ \sum_{k=h1}^{h2} \left(|Y(t, kf_0)| - |\hat{N}_i(kf_0)| \right)^2 \right\}$$

ただしｔは時刻、f_0は離散的周波数成分算出手段における所定の周波数間隔、Ｙ(t,kf_0)は時刻ｔにおける、周波数ｋf_0での信号の周波数成分、^N_i(kf_0)は時刻ｔにおける、周波数ｋf_0での雑音モデル^N_iの周波数成分、h_1は$h_1 f_0$＝第１の周波数帯域の下限となるような整数、h_2は$h_2 f_0$＝第１の周波数帯域の上限となるような整数、をそれぞれ表す。

【００２３】

より好ましくは、レベル推定手段は、信号|Ｙ(t,kf_0)|と選択された雑音モデル^Ng(kf_0)とから、以下の式にしたがって選択された雑音モデルのレベル^α(t)をフレームごとに推定するための手段を含む。

【００２４】

【数３】

$$\hat{\alpha}(t) = \arg\min_\alpha \left\{ \sum_{k=j1}^{j2} \left(|Y(t, kf_0)| - \alpha |\hat{N}_g(kf_0)| \right)^2 \right\}$$

ただし、j_1は$j_1 f_0$＝第２の周波数帯域の下限となるような整数、j_2は$j_2 f_0$＝第２の周波数帯域の上限となるような整数、をそれぞれ表す。

【００２５】

減算手段は、以下の式にしたがって信号から雑音をフレームごとに除去した信号^Ｓ(t,kf_0)を出力するようにしてもよい。

【００２６】

【数４】

$$|\hat{S}(t, kf_0)| = |Y(t, kf_0)| - \hat{\alpha}(t) |\hat{N}_g(kf_0)|$$

ただし^Ｓ(t,kf_0)は時刻ｔにおける信号^Ｓの周波数ｋf_0における周波数成分を、^Ng(kf_0)は選択された雑音モデル^Ngの周波数ｋf_0における周波数成分を、それぞれ表す。

【００２７】

好ましくは、雑音除去装置は、各々が複数通りの雑音モデルを含む複数個の信号源プロフィール情報を記憶するための手段と、複数個の信号源プロフィール情報のうちのいずれかを、ユーザの指定により選択して、当該選択された信号源プロフィール情報に含まれる

複数個の雑音モデルを雑音モデル記憶手段に格納するための手段とをさらに含む。

【発明を実施するための最良の形態】

【００２８】

［第１の実施の形態］

－動作の原理－

図１に、本発明の第１の実施の形態に係る雑音除去装置の動作原理を示す。図１上段を参照して、風雑音の場合、雑音２２の成分は周波数スペクトルにおいて比較的低域に集中することが知られている。一方、信号２０の周波数成分はより高域に集中している。そこで、予め特徴的な雑音のスペクトル形状を複数の雑音モデル３０、３２、３４等として準備しておき、観測信号のうち、１点鎖線２４で示される所定のしきい値ＴＨ０以下の周波数成分の形状に最も近いスペクトル形状（スペクトル形状４０、４２または４４）を持つ雑音モデルを選択する。さらに雑音レベルを推定することにより、実際の雑音成分を推定し、観測信号から減算することにより雑音除去を行なう。

【００２９】

なお、雑音モデルは、音声を電気信号に変換するマイクの機種により異なる。したがって、例えばマイク製造者が予め雑音モデルを準備しておき、それを雑音除去装置に取込むような仕組みを設けておくことが望ましい。さらに、上記したように風雑音の場合には、所定のしきい値（例えば１２３Ｈｚ）以下の周波数に周波数成分が集中しているが、他の種類の雑音の場合には、これとは異なる別の帯域に集中していることも考えられる。または、複数の帯域に集中帯域が分散していることも考えられる。したがって、雑音モデルを選択するための帯域を利用者が選択できるようにすることが望ましい。以下に説明する実施の形態に係る雑音除去装置は、そのような仕組みを有している。

【００３０】

－構成－

図２は、本実施の形態に係る雑音除去装置５０の構成を示すブロック図である。図２において、図５に示すものと同じ部品には同じ参照番号を付してある。それらの機能及び名称も同様である。したがって、それらについての詳細な説明は繰返さない。なお、図２に示す雑音除去装置５０は、図５に示す従来の雑音除去装置１２０と同様のＳＳ法による処理も可能であり、いずれを使用するかを選択できる。しかし図２においては、図および説明を分かりやすくするために、図２に示す各部品のうち、従来技術のみに使用される部分は示していない。

【００３１】

図２を参照して、雑音除去装置５０は、上記した複数の雑音モデル及び雑音モデルを選択する信号の帯域を、マイクロフォンごとにマイクプロフィールとして記憶するためのマイクプロフィール記憶部６２と、いわゆるインターネットに接続されると、例えばマイクロフォンメーカが準備したサーバからマイクプロフィールを自動的に取寄せ、取寄せたマイクプロフィールでマイクプロフィール記憶部６２を更新するためのマイクプロフィール更新部６０と、ユーザからの指示に応じてマイクプロフィール記憶部６２に記憶されているマイクプロフィールおよび使用雑音帯域の一覧を表示し、ユーザにいずれかを選択させるためのマイクプロフィール／雑音帯域選択部６６と、マイクプロフィール／雑音帯域選択部６６により選択されたマイクプロフィールから複数の雑音モデルを読出し記憶するための雑音モデル記憶部６４とを含む。

【００３２】

雑音除去装置５０はさらに、観測信号ｙ(i)に対し、ＳＴＦＴ処理を行なって周波数領域に変換し、周波数成分８２と位相成分８０とを出力するためのＳＴＦＴ処理部６８と、ＳＴＦＴ処理部６８の出力する周波数成分８２と、マイクプロフィール／雑音帯域選択部６６から与えられる雑音帯域情報とにしたがって、雑音モデル記憶部６４に記憶されている複数の雑音モデルから例えば以下の式（２）により最適と思われる雑音モデル$|\hat{N}_g(kf_0)|$を選択して出力し、さらに当該雑音モデルと周波数成分８２とに基づいて雑音信号のレベル推定値$\hat{\alpha}(t)$を以下の式（３）により推定するための雑音推定部７０とを含む。

【００３３】
【数５】

$$g = \arg\min_{i}\left\{\sum_{k=0}^{h}\left(\left|Y(t,kf_0)\right|-\left|\hat{N}_g(kf_0)\right|\right)^2\right\} \qquad (2)$$

$$\hat{\alpha}(t) = \arg\min_{\alpha}\left\{\sum_{k=0}^{j}\left(\left|Y(t,kf_0)\right|-\alpha\left|\hat{N}_g(kf_0)\right|\right)^2\right\} \qquad (3)$$

ただし、式（２）におけるｈは雑音モデルの選択に用いるしきい値ＴＨ０（例えば１２３Ｈｚ）に対応するｋの値を示す。また式（３）におけるｊは雑音レベル推定の場合に使用されるしきい値ＴＨ１（例えば１３３Ｈｚ）に対応するｋの値を示す。また、本実施の形態ではｈ≦ｊである。すなわち、雑音モデルの推定に用いる帯域よりも、レベル推定値^α(t)の推定に用いる帯域の方が大きく選ばれている。このようにすることにより、レベル推定がより精度高く行なえる。

【００３４】
なお、使用バンドが異なれば、式（２）（３）において合計の対象となるｋの値の範囲（上限および下限）も異なってくる。上記した式（２）（３）はあくまで風雑音の場合の例である。また上の式におけるｈおよびｊの値はマイクプロフィールに記憶されており、マイクプロフィール／雑音帯域選択部６６により選択されてそれぞれ雑音モデル選択部９０および^α(t)推定部９２に与えられる。

【００３５】
雑音除去装置５０はさらに、ＳＴＦＴ処理部６８の出力する周波数成分$|Y(t,kf_0)|$８２に対し、雑音推定部７０の出力する雑音モデル$|\hat{}N_g(kf_0)|$８４およびレベル推定値^α(t)８６を用いて以下の式（４）にしたがうＳＳ法により雑音除去を行ない、信号$|\hat{}S(t,kf_0)|$８８を出力するためのＳＳ処理部９４を含む。

【００３６】
【数６】

$$\left|\hat{S}(t,kf_0)\right| = \left|Y(t,kf_0)\right| - \hat{\alpha}(t)\left|\hat{N}_g(kf_0)\right| \qquad (4)$$

雑音除去装置５０はさらに、ＳＳ処理部９４の出力する信号$|\hat{}S(t,kf_0)|$８８に対し、ＳＴＦＴ処理部６８の出力する位相成分８０（$\phi_y(kf_0)$）を位相情報として用いてＩＦＦＴ処理を実行するためのＩＦＦＴ処理部７４と、ＩＦＦＴ処理部７４の出力を受けるインバースウィンドイング処理部７６と、インバースウィンドイング処理部７６の出力を受けて波形合成をし、信号^s(i)を出力するための波形合成処理部７８とを含む。ＩＦＦＴ処理部７４、インバースウィンドイング処理部７６および波形合成処理部７８により、雑音成分の除去された信号が周波数領域から時間領域に変換される。

【００３７】
雑音推定部７０は、雑音モデル記憶部６４に記憶された複数の雑音モデルから、マイクプロフィール／雑音帯域選択部６６により指定された雑音帯域の形状が周波数成分８２の当該帯域の形状に最も近いものを上記した式（２）にしたがって選択し、雑音モデル$|\hat{}N_g(kf_0)|$８４を出力するための雑音モデル選択部９０と、雑音モデル選択部９０が出力する雑音モデル$|\hat{}N_g(kf_0)|$８４とＳＴＦＴ処理部６８の出力する周波数成分８２とに基づき、上記した式（３）にしたがって雑音のレベル推定値^α(t)８６を出力するための^α(t)推定部９２とを含む。

【００３８】
－動作－
雑音除去装置５０は以下のように動作する。マイクプロフィール記憶部６２には、雑音除去装置５０が取付けられた機器（例えば携帯ビデオカメラ等）の出荷時に、機器の製造者により、当該機器で使用されているマイクのマイクプロフィールが記憶される。出荷後、マイクプロフィールに修正があったり、新たなマイクプロフィールの追加があったりしたときには、雑音除去装置５０をインターネットに接続することにより、それらマイクプ

ロフィールによりマイクプロフィール記憶部６２に記憶されたマイクプロフィールが自動的に更新される。マイクプロフィールとしては、例えば風雑音除去用のマイクプロフィールがある。

【００３９】

撮影時、通常は図５に示すものと同様の従来のＳＳ法による雑音除去を行なう。野外で、風雑音などがある場合には、ユーザはマイクプロフィール／雑音帯域選択部６６を使用して、どのマイクプロフィールを使用するかを選択する。ここでは風雑音用のマイクプロフィールを選択するものとする。したがってマイクプロフィール／雑音帯域選択部６６は当該帯域を示す情報（ｈ）を雑音推定部７０の雑音モデル選択部９０に与える。マイクプロフィール／雑音帯域選択部６６はまた、^α(t)推定の際に使用される帯域を示す情報（ｊ）を雑音推定部７０の^α(t)推定部９２に与える。

【００４０】

雑音モデル選択部９０は、式（２）にしたがい、周波数成分８２との間の二乗誤差が最小となる雑音モデル$|\hat{N}g(kf_0)|$８４を定め、^α(t)推定部９２およびＳＳ処理部９４に与える。

【００４１】

^α(t)推定部９２は、この雑音モデル|^Ng(kf0)|８４と周波数成分８２とに基づき、式（３）にしたがって雑音のレベル推定値^α(t)８６を推定し、ＳＳ処理部９４に与える。

【００４２】

ＳＳ処理部９４は、ＳＴＦＴ処理部６８からの周波数成分８２に対し、雑音モデル選択部９０からの雑音モデル|^Ng(kf0)|８４、^α(t)推定部９２からの^α(t)８６を用いて式（４）にしたがうＳＳ処理を実行する。ＳＳ処理部９４は、こうして雑音の除去された信号|^S(t,kf0)|８８をＩＦＦＴ処理部７４に与える。

【００４３】

ＩＦＦＴ処理部７４は、信号|^S(t,kf0)|８８に対しＳＴＦＴ処理部６８からの位相成分８０を位相情報として用いてＩＦＦＴ処理を実行し、インバースウィンドイング処理部７６に与える。インバースウィンドイング処理部７６以下で行なわれる処理は、図５に示す従来のものと同様である。

【００４４】

－実験結果－

図３に、雑音除去装置５０を用いて行なった実験によって得られた結果を示す。図３において、棒グラフ１６０、１６２および１６４はそれぞれ、従来のＳＳ法におけるＳＮＲ（Signal-to-Noise Ratio）、上記実施の形態に係るＳＳ法によるＳＮＲ、および理論的なＳＮＲ（いずれもｄＢ）を示す。図３を参照して明らかなように、本実施の形態によれば従来のＳＳ法を用いた場合と比較してはるかに高い精度で効率よく雑音を除去できる。条件にもよるが、図３に示すように上限値に近いＳＮＲも得られる。

【００４５】

以上のように本発明の実施の形態においては、複数の雑音モデルを用意し、観測信号のスペクトルの所定帯域の形状に対応した雑音モデルをフレームごとに選択し、さらにフレームごとに雑音レベルを推定することにより、式（４）に示すＳＳ法にしたがって雑音を除去する。そのため、風雑音などの非定常雑音下でもその変化に追従し、安定して効率よく雑音を除去することができる。

【００４６】

上記した実施の形態では、風雑音の雑音モデル選択のために、観測信号の周波数スペクトルのうち、所定のしきい値より低い周波数成分のみを用いている。しかし本発明はこうした実施の形態に限定されるわけではない。上記したように、雑音の種類によってはこのように雑音が低域ではなく他の帯域に集中することもある。そうした場合、その帯域が分かっていればその帯域の周波数成分を用いて雑音モデルの選択を行なえばよい。また、そのように雑音モデルの選択を行なうための帯域が一つに限定されるわけではなく、集中帯

域が複数の帯域に分散していることもある。その場合には、複数の帯域にわたって上記した最小二乗法により雑音モデルを選択するようにしてもよい。

【００４７】

さらに、上記した帯域の選択を、フレームごとに変化させるようにしてもよい。この場合には何らかの形でフレームごとに雑音が集中している帯域を調べる機構が必要となる。

【００４８】

また、上記実施の形態では、雑音除去装置としてビデオカメラに組込んだ例を説明した。しかし本発明はそのような実施の形態に限定されるわけではない。例えば、テレビジョン受像機のように、映像を再生する装置にこの雑音除去装置を組込んでも良い。さらに、いわゆるパーソナルコンピュータなどで映像の編集を行なうための編集ソフトウェアに対するプラグインの形で、上記した雑音除去装置５０をソフトウェアの形で組込んでもよい。その場合、マイクプロフィール／雑音帯域選択部６６は、編集ソフトウェアのユーザインタフェースにあわせ、パーソナルコンピュータの画面とキーボード、マウスなどの入力装置を用いたＧＵＩ（Graphical User Interface）で実現することが望ましい。

【００４９】

また、上記実施の形態では、もっぱら雑音除去装置のみを示したが、この音声除去装置を音声認識装置の前段に設けることで、音声認識の精度を高めることができる。例えば人間の発声では、母音はハーモニクスを含むため、音声のパワースペクトル上では複数箇所で谷が生ずる。この場合、本実施例での雑音帯域としてそうした谷に対応する領域を用いて雑音モデルを選択できる。この雑音モデルを用いて音声から雑音を除去することで音声認識の精度を高めることができる。

【００５０】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内でのすべての変更を含む。

【図面の簡単な説明】

【００５１】

【図１】本発明の一実施の形態に係る雑音除去装置の動作原理を説明するための図である。

【図２】本発明の一実施の形態に係る雑音除去装置５０のブロック図である。

【図３】本発明の一実施の形態を用いて行なった雑音除去の効果を、従来技術と対比して示すグラフである。

【図４】従来のＳＳ法の原理を説明するための図である。

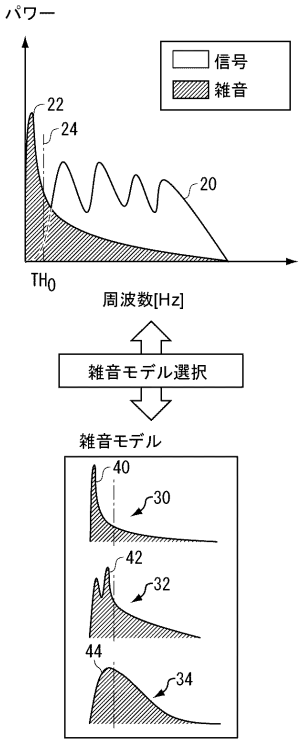
【図５】従来のＳＳ法を用いた雑音除去装置１２０のブロック図である。

【符号の説明】

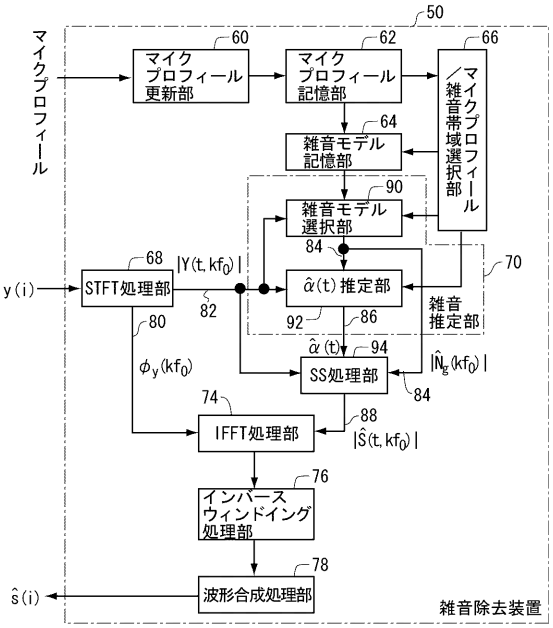
【００５２】

５０，１２０　雑音除去装置、６０　マイクプロフィール更新部、６２　マイクプロフィール記憶部、６４　雑音モデル記憶部、６６　マイクプロフィール／雑音帯域選択部、６８　ＳＴＦＴ処理部、７０　雑音推定部、７２，９４　ＳＳ処理部、７４　ＩＦＦＴ処理部、７６　インバースウィンドイング処理部、７８　波形合成処理部、８０　位相成分、８２　周波数成分、８４　雑音モデル、８６　レベル推定値、８８　信号、９０　雑音モデル選択部、９２　^α(t)推定部

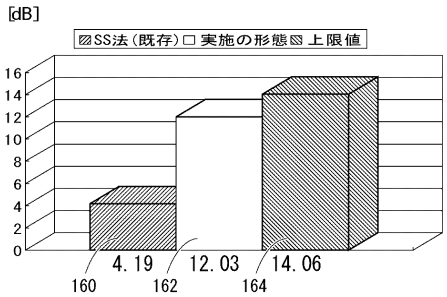
【図１】


パワー
信号
雑音
22
24
20
TH_0
周波数[Hz]
雑音モデル選択
雑音モデル
40
30
42
32
44
34

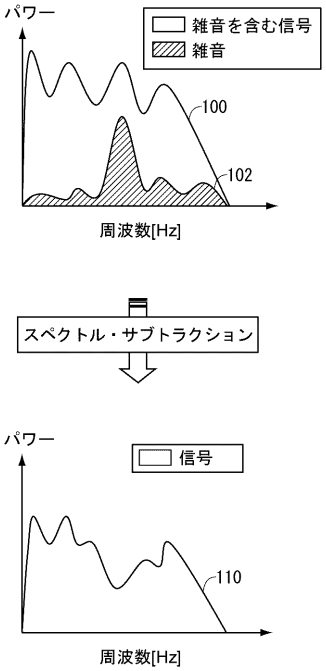
【図２】


50
マイクプロフィール
60
マイクプロフィール更新部
62
マイクプロフィール記憶部
66
マイクプロフィール／雑音帯域選択部
64
雑音モデル記憶部
90
雑音モデル選択部
84
70
雑音推定部
68
$|Y(t,kf_0)|$
y(i)
STFT処理部
82
$\hat{\alpha}(t)$推定部
92
86
80
$\phi_y(kf_0)$
$\hat{\alpha}(t)$
94
$|\hat{N}_g(kf_0)|$
SS処理部
84
74
IFFT処理部
88
$|\hat{S}(t,kf_0)|$
76
インバースウィンドイング処理部
78
$\hat{s}(i)$
波形合成処理部
雑音除去装置

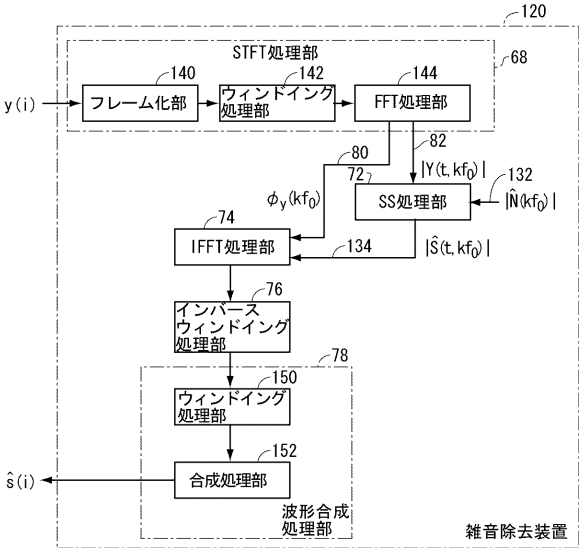
【図３】


[dB]
SS法（既存）
実施の形態
上限値
16
14
12
10
8
6
4
2
0
4.19
12.03
14.06
160
162
164

【図４】


パワー
雑音を含む信号
雑音
100
102
周波数[Hz]
スペクトル・サブトラクション
パワー
信号
110
周波数[Hz]

【図５】


120
STFT処理部
68
140
142
144
y(i)
フレーム化部
ウィンドイング処理部
FFT処理部
82
80
72
$|Y(t,kf_0)|$
132
$\phi_y(kf_0)$
SS処理部
$|\hat{N}(kf_0)|$
74
IFFT処理部
134
$|\hat{S}(t,kf_0)|$
76
インバースウィンドイング処理部
78
150
ウィンドイング処理部
152
$\hat{s}(i)$
合成処理部
波形合成処理部
雑音除去装置

フロントページの続き

審査官　井上　健一

(56)参考文献　特開２００３－３０８０９２（ＪＰ，Ａ）
特開昭５６－０４６３００（ＪＰ，Ａ）
特開平０６－０６７６９０（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　２１／０２
Ｇ１０Ｌ　１５／２０