

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3960834号
(P3960834)

(45) 発行日 平成19年8月15日(2007.8.15)

(24) 登録日 平成19年5月25日(2007.5.25)

(51) Int. Cl. F I
G 1 O L 21/02 (2006.01) G 1 O L 21/02 1 O 1 B
G 1 O L 11/04 (2006.01) G 1 O L 21/02 1 O 2 B
 G 1 O L 11/04

請求項の数 16 (全 32 頁)

<p>(21) 出願番号 特願2002-77327 (P2002-77327) (22) 出願日 平成14年3月19日 (2002.3.19) (65) 公開番号 特開2003-280696 (P2003-280696A) (43) 公開日 平成15年10月2日 (2003.10.2) 審査請求日 平成16年12月2日 (2004.12.2)</p>	<p>(73) 特許権者 000005821 松下電器産業株式会社 大阪府門真市大字門真1006番地 (74) 代理人 100105050 弁理士 鷺田 公一 (72) 発明者 王 幼華 石川県金沢市西念一丁目1番3号 株式会 社 松下通信金沢研究所内 (72) 発明者 吉田 幸司 神奈川県横浜市港北区綱島東四丁目3番1 号 松下通信工業株式会社内 審査官 山下 剛史</p>
--	--

最終頁に続く

(54) 【発明の名称】 音声強調装置及び音声強調方法

(57) 【特許請求の範囲】

【請求項1】

入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割手段と、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成手段と、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成手段と、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定手段と、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復手段と、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正手段と、前記コムフィルタ修正手段において修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧手段と、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成手段と、を具備することを特徴とする音声強調装置。

10

【請求項2】

第一コムフィルタと第二コムフィルタとから音声スペクトルに音声成分が含まれているか否か判定する音声ノイズフレーム検出手段を具備し、コムフィルタ修正手段は、前記音声ノイズフレーム検出手段の判定が音声成分を含まない結果である場合、前記第一コムフィルタに各周波数成分で信号を減衰する修正をすることを特徴とする請求項1に記載の音

20

声強調装置。

【請求項 3】

前記音声ノイズフレーム検出手段は、第一コムフィルタの通過域における入力信号のパワスペクトルの和と第一コムフィルタの阻止域における入力信号のパワスペクトルの和の比を第一結果とし、第二コムフィルタの通過域における入力信号のパワスペクトルの和と第二コムフィルタの阻止域における入力信号のパワスペクトルの和の比を第二結果とし、第一結果と第二結果を加算した結果が所定の閾値より大きい場合、前記加算結果を用い、前記加算結果が所定の閾値以下である場合、第二結果を用いて音声スペクトルに音声が含まれるか否か判定することを特徴とする請求項 2 に記載の音声強調装置。

【請求項 4】

第一コムフィルタ生成手段は、入力信号のパワスペクトルの極小値から所定の範囲を阻止域とする第一コムフィルタを作成することを特徴とする請求項 1 から請求項 3 のいずれかに記載の音声強調装置。

【請求項 5】

入力信号のパワスペクトルからノイズベースを減算し、減算結果の自己相関関数を用いて音声ピッチを推定するピッチ推定手段を具備し、コムフィルタ修正手段は、前記ピッチ推定手段において推定した音声ピッチで前記第一コムフィルタに含まれる音声ピッチ調波構造を修正することを特徴とする請求項 1 から請求項 4 のいずれかに記載の音声強調装置。

【請求項 6】

入力信号のパワスペクトルからノイズベースを減算した結果に、直流成分に所定のパワを持つ擬似ピークを生成する直流成分生成手段を具備し、ピッチ推定手段は、前記直流成分生成手段において擬似ピークを生成した前記パワスペクトルから音声ピッチを推定することを特徴とする請求項 5 に記載の音声強調装置。

【請求項 7】

入力信号のパワスペクトルからノイズベースを減算した結果において、パワが所定の閾値以上である周波数領域の数の移動平均を算出するノイズ特性推定手段を具備し、第二コムフィルタ作成手段は、前記移動平均から入力信号に音声が含まれるか否か判断した結果より第二コムフィルタを作成することを特徴とする請求項 1 から請求項 6 のいずれかに記載の音声強調装置。

【請求項 8】

入力信号のパワスペクトルからノイズベースを減算した結果において、パワが所定の閾値以上である周波数領域の数の移動平均を算出するノイズ特性推定手段を具備し、第二コムフィルタ作成手段は、ノイズ特性推定手段において算出された移動平均が所定の値以下である場合、所定の周波数領域を阻止域とする第二コムフィルタを作成することを特徴とする請求項 1 から請求項 7 のいずれかに記載の音声強調装置。

【請求項 9】

入力信号のパワスペクトルとノイズベースから信号対雑音比を算出する SNR 推定手段を具備し、抑圧手段は、前記信号対雑音比から周波数分割スペクトルのノイズの抑圧量を決定することを特徴とする請求項 1 から請求項 8 のいずれかに記載の音声強調装置。

【請求項 10】

SNR 推定手段は、音声成分のレベルを入力信号のパワスペクトルの移動平均値より算出し、雑音成分のレベルをノイズベースの推定値に各周波数成分別に重み計数を乗算した値より算出し、前記音声成分のレベルと前記雑音成分のレベルの比から信号対雑音比を算出することを特徴とする請求項 9 に記載の音声強調装置。

【請求項 11】

信号対雑音比と前記信号対雑音比の移動平均値との偏差を算出し、前記偏差を用いて前記信号対雑音比の移動平均値を更新する変動抑圧手段を具備し、抑圧手段は、変動抑圧手段において更新された前記信号対雑音比の移動平均値から周波数分割スペクトルのノイズの抑圧量を決定することを特徴とする請求項 9 または請求項 10 に記載の音声強調装置。

10

20

30

40

50

【請求項 1 2】

所定の時間単位で更新速度の異なる二つのノイズベースの移動平均値を算出し、第一移動平均値より更新速度の速い第二移動平均値で第一移動平均値の更新条件を変更し、第一移動平均値をノイズベース推定値として出力するノイズベース更新手段を具備することを特徴とする請求項 1 から請求項 1 1 のいずれかに記載の音声強調装置。

【請求項 1 3】

請求項 1 から請求項 1 2 のいずれかに記載の雑音抑圧装置を具備することを特徴とする無線通信装置。

【請求項 1 4】

入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、を具備することを特徴とする音声強調方法

【請求項 1 5】

入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、をコンピュータに実行させることを特徴とする音声強調プログラム。

【請求項 1 6】

入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、をコンピュータに実行させることを特徴とする音声強調プログラムを記憶し、要求に応じて前記音声強調

プログラムを出力することを特徴とするサーバ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声強調装置及び音声強調方法に関し、特に通信に用いて好適な音声強調装置及び音声強調方法に関する。

【0002】

【従来の技術】

従来の音声符号化通信において、低ビットレートで音声を符号化して通信を行うと、背景雑音のない音声に対しては高品質な音声での通話を提供することができるが、背景雑音が含まれた音声に対しては低ビットレート符号化特有の耳障りな歪みが生じ、音質が劣化する。

10

【0003】

この音質が劣化する問題の対処するため雑音抑圧処理を行うことがある。この雑音抑圧処理に関する音声強調技術としてスペクトルサブトラクション法、およびコムフィルタ法がある。

【0004】

スペクトルサブトラクション法（SS法）は、雑音情報に着目して無音区間で雑音の性質を推定して雑音を含む音声信号の短時間パワースペクトルから雑音の短時間パワースペクトルを減算する、または減衰係数を乗算することにより音声信号のパワースペクトルを推定して雑音を抑圧する方法である。SS法は、例えば、文献1（S.Boll, Suppression of acoustic noise in speech using spectral subtraction, IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP-27, pp.113-120, 1979）、文献2（R.J. McAulay, M.L. Malpass, Speech enhancement using a soft-decision noise suppression filter, IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP-28, pp.137-145, 1980）に記載されているものがある。

20

【0005】

しかしながら、SS法では、音声からノイズを差し引いた後の残留ノイズとして、音声ピッチ間の残留ノイズに起因する歪を生じる。

【0006】

例えば、音声強調方法として用いられる文献1に示したSS法では、ノイズ情報のみに着目し、短時間のノイズ特性を定常と見なして、音声とノイズを区別せず、一律にノイズベース（推定されたノイズのスペクトル特性）を差し引く。しかし、SS法では、音声の情報（例えば、音声のピッチ）は利用していない。実際の通信では、ノイズの特性が定常でないため、差し引かれた後の残留ノイズ、特に音声ピッチ間の残留ノイズにより、いわゆる「ミュージカルノイズ」と呼ばれる不自然な歪を生じる。

30

【0007】

この不自然な歪みの改善法として、音声パワ対ノイズパワの比（SNR）に基づき、減衰係数を乗じてノイズを減衰する方法、例えば、特許第2714656号及び特願平9-518820号に開示されたものが提案されている。

40

【0008】

これらに開示された方法では、音声の大きい（SNRが大きい）帯域とノイズの大きい（SNRが小さい）帯域を区別して異なる減衰係数を用いるため、ミュージカルノイズを抑制し、音質を向上させた。

【0009】

しかし、特許第2714656号及び特願平9-518820号に開示された方法では、処理する周波数チャネル数（16チャネル）は十分でないので、音声ピッチ調波情報を雑音から分離し抽出することがむずかしい。

【0010】

また、音声とノイズ両方の帯域に減衰係数を用いるため、互いに影響を及ぼし合う結果、

50

減衰係数は大きくすることができない。例えば、減衰係数を大きくすると、SNR推定の誤りによって、音声の歪みを生じる可能性がある。この結果、ノイズの減衰が不十分となる。

【0011】

一方、コムフィルタ法は、音声情報に着目し、音声ピッチにコムフィルタをかけることにより雑音減衰を行う方法である。コムフィルタとは、周波数領域単位で入力された信号を所定の比率で減衰させ、または減衰させずに信号を出力するフィルタであり、櫛状の減衰特性をもつ。デジタルデータ処理でコムフィルタ法を実現する場合、コムフィルタ法では、コムフィルタの減衰特性を周波数領域毎に減衰特性のデータを作成し、周波数毎に音声スペクトルを乗算することにより雑音を抑圧できる。

10

【0012】

コムフィルタ法に関する文献として、文献3(J.S.Lim etc., Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition, IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP26, pp. 354-358, 1978)に示すものがある。

【0013】

従来のコムフィルタ法では、基本周波数であるピッチに推定誤差があると、その高調波では誤差分が拡大し、本来の高調波成分がその通過帯域からはずれる可能性がより大きくなる。また、準周期性を持つ音声とそうでない音声を判別する必要があるため、実現性に問題がある。また、移動体通信において、単に雑音を抑圧することによって、自然感のある

20

【0014】

通話の違和感を解消する方法として、音声と雑音を分離し、それぞれ符号化と復号化を行うことによって良好な音声と周囲騒音を再生する技術がある。例えば、文献4(三関 公生, 押切 正浩, 音声・背景雑音分離にもとづく低レート音声符号化, 日本音響学会講演論文集, pp. 235-236, 平成10年3月)に示した方法がある。

【0015】

文献4に示した方法は、SS法を用いて音声強調を行い、雑音抑圧された入力信号を音声成分とし、入力信号と音声成分を減算した結果を雑音成分とする方法である。基本的にはSS法と同じ方法であるため、本方法による分離方法では良好な音声と雑音特性を得ることが難しい。

30

【0016】

【発明が解決しようとする課題】

このように、従来の装置においては、音声の歪みが少なくかつ雑音を十分に除去することが難しいという問題がある。

【0017】

発明は、かかる点に鑑みてなされたものであり、音声の歪みが少なくかつ雑音を十分に除去することができる音声強調装置を提供するとともに、良好な音声と雑音の特性を得られる音声強調装置及び音声強調方法を提供することを目的とする。

【0018】

【課題を解決するための手段】

本発明の音声強調装置は、入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割手段と、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成手段と、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成手段と、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定手段と、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復手段と、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフ

40

50

フィルタ修正手段と、前記コムフィルタ修正手段において修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧手段と、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成手段と、を具備する構成を採る。

【0019】

この構成によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声ピッチ情報を取得し、また、できる限り多くの音声情報を取り出してノイズ信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いてノイズ信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

10

【0020】

本発明の音声強調装置は、第一コムフィルタと第二コムフィルタとから音声スペクトルに音声成分が含まれているか否か判定する音声/雑音フレーム検出手段を具備し、コムフィルタ修正手段は、前記音声/雑音フレーム検出手段の判定が音声成分を含まない結果である場合、前記第一コムフィルタに各周波数成分で信号を減衰する修正をする構成を採る。

【0021】

本発明の音声強調装置は、前記音声/雑音フレーム検出手段は、第一コムフィルタの通過域における入力信号のパワスペクトルの和と第一コムフィルタの阻止域における入力信号のパワスペクトルの和の比を第一結果とし、第二コムフィルタの通過域における入力信号のパワスペクトルの和と第二コムフィルタの阻止域における入力信号のパワスペクトルの和の比を第二結果とし、第一結果と第二結果を加算した結果が所定の閾値より大きい場合、前記加算結果を用い、前記加算結果が所定の閾値以下である場合、第二結果を用いて音声スペクトルに音声が含まれるか否か判定する構成を採る。

20

【0022】

これらの構成によれば、第一コムフィルタと第二コムフィルタとから音声スペクトルに音声成分が含まれているか否か判定し、この判定が音声成分を含まない結果である場合、前記第一コムフィルタに各周波数成分で信号を減衰することにより、突発的に発生するノイズを抑圧することができる。

30

【0023】

本発明の音声強調装置は、第一コムフィルタ生成手段は、入力信号のパワスペクトルの極小値から所定の範囲を阻止域とする第一コムフィルタを作成する構成を採る。

【0024】

この構成によれば、入力信号のパワスペクトルの極小値から所定の範囲を阻止域とするコムフィルタを作成し、このコムフィルタを用いてノイズ信号を抑圧することにより、音声とノイズのレベル差が少ない場合でも、音声ピッチ調波構造を抽出して修復して音声歪を低減することができる。

【0025】

本発明の音声強調装置は、入力信号のパワスペクトルからノイズベースを減算し、減算結果の自己相関関数を用いて音声ピッチを推定するピッチ推定手段を具備し、コムフィルタ修正手段は、前記ピッチ推定手段において推定した音声ピッチで前記第一コムフィルタに含まれる音声ピッチ調波構造を修正する構成を採る。

40

【0026】

この構成によれば、入力信号のパワスペクトルからノイズベースを減算し、減算結果の自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチでコムフィルタに含まれる音声ピッチ調波構造を修正することにより、コムフィルタのピッチ調波構造の修復を行うことができ、音声歪みの少ない音声強調を行うことができる。

【0027】

本発明の音声強調装置は、入力信号のパワスペクトルからノイズベースを減算した結果に

50

、直流成分に所定のパワーを持つ擬似ピークを生成する直流成分生成手段を具備し、ピッチ推定手段は、前記直流成分生成手段において擬似ピークを生成した前記パワースペクトルから音声ピッチを推定する構成を採る。

【0028】

この構成によれば、入力信号のパワースペクトルからノイズベースを減算した結果に直流成分に所定のパワーを持つ擬似ピークを生成し、直流成分を生成したスペクトルの自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチでコムフィルタに含まれる音声ピッチ調波構造を修正することにより、音声スペクトルの調波ピークが少ない場合でも、ピッチ情報を得てコムフィルタのピッチ調波構造の修復を行うことができ、音声歪みの少ない音声強調を行うことができる。

10

【0029】

本発明の音声強調装置は、入力信号のパワースペクトルからノイズベースを減算した結果において、パワーが所定の閾値以上である周波数領域の数の移動平均を算出するノイズ特性推定手段を具備し、第二コムフィルタ作成手段は、前記移動平均から入力信号に音声が含まれるか否か判断した結果より第二コムフィルタを作成する構成を採る。

【0030】

この構成によれば、入力信号のノイズレベルの分布を検出し、この分布に基づいて音声スペクトルからコムフィルタを生成する基準を決定し、作成したコムフィルタからピッチ情報を取得することにより、雑音の状態に応じたピッチ情報を取得してコムフィルタを作成することができ、音声歪みの少ない音声強調を行うことができる。

20

【0031】

本発明の音声強調装置は、入力信号のパワースペクトルからノイズベースを減算した結果において、パワーが所定の閾値以上である周波数領域の数の移動平均を算出するノイズ特性推定手段を具備し、第二コムフィルタ作成手段は、ノイズ特性推定手段において算出された移動平均が所定の値以下である場合、所定の周波数領域を阻止域とする第二コムフィルタを作成する構成を採る。

【0032】

この構成によれば、ノイズ特性の推定結果に基づいて周波数成分を選択し、第二コムフィルタにおいて、選択された周波数領域をすべて阻止域に変換することによって、分散値の大きいノイズにより生成した偽のピッチ調波を減少し、偽のピッチ調波を生じにくい低周波数領域におけるピッチ調波を基準に、ピッチ調波構造の修復を行えば、正確にピッチ調波構造を修復することができる。

30

【0033】

本発明の音声強調装置は、入力信号のパワースペクトルとノイズベースから信号対雑音比を算出するSNR推定手段を具備し、抑圧手段は、前記信号対雑音比から周波数分割スペクトルのノイズの抑圧量を決定する構成を採る。

【0034】

この構成によれば、修正コムフィルタの通過域と阻止域において、入力音声パワースペクトルからノイズベースを減算する量とノイズ減衰の度合いをSNR推定値の大きさに応じて調整することによって、異なるSNRの環境下でも適切なノイズ減衰を行い、音声歪と残留ノイズの少ない音声強調を実現できる。

40

【0035】

本発明の音声強調装置は、SNR推定手段は、音声成分のレベルを入力信号のパワースペクトルの移動平均値より算出し、雑音成分のレベルをノイズベースの推定値に各周波数成分別に重み計数を乗算した値より算出し、前記音声成分のレベルと前記雑音成分のレベルの比から信号対雑音比を算出する構成を採る。

【0036】

この構成によれば、入力音声パワースペクトルの移動平均値からノイズベースを減算して音声レベルの計算することにより、ノイズの影響を減少し、低SNRの環境下でも正確な音声レベルを計算することができる。また、ノイズベース推定値の各周波数成分に重み係数

50

を乗算して計算することにより、異なるノイズに対して適切な減衰を行い、音声歪を低減することができる。

【0037】

本発明の音声強調装置は、信号対雑音比と前記信号対雑音比の移動平均値との偏差を算出し、前記偏差を用いて前記信号対雑音比の移動平均値を更新する変動抑圧手段を具備し、抑圧手段は、変動抑圧手段において更新された前記信号対雑音比の移動平均値から周波数分割スペクトルのノイズの抑圧量を決定する構成を採る。

【0038】

この構成によれば、SNR推定値とSNR推定値の長期移動平均値の偏差を計算し、SNR推定値の長期移動平均値と前記偏差の一部を加算してSNR推定値として用いることによつて、SNRの変動を有効に抑制し、安定的にSNRの大きさに応じてノイズ減衰のレベル調整を行うことができる。

10

【0039】

本発明の音声強調装置は、所定の時間単位で更新速度の異なる二つのノイズベースの移動平均値を算出し、第一移動平均値より更新速度の速い第二移動平均値で第一移動平均値の更新条件を変更し、第一移動平均値をノイズベース推定値として出力するノイズベース更新手段を具備する構成を採る。

【0040】

この構成によれば、更新速度の速い移動平均係数を用いてノイズベースの推定を行うことによつて、音声区間においても雑音レベルの急激な変動を追跡することができる。また、更新速度の遅いノイズベースの更新は、更新速度の速いノイズベースに基づいて行うことによつて、正確にノイズベースの推定を行うことができ、雑音レベルの急激な変動によるノイズベース更新の停止を防止することができる。

20

【0041】

本発明の無線通信装置は、上記いずれかに記載の音声強調装置を具備する構成を採る。

【0042】

この構成によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声ピッチ情報を取得し、また、できる限り多くの音声情報を取り出して力信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いてノイズ信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

30

【0053】

本発明の音声強調方法は、入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、を具備するようにした。

40

【0054】

この方法によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声

50

ピッチ情報を取得し、また、できる限り多くの音声情報を取り出して力信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いて音声信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

【0057】

本発明の音声強調プログラムは、入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、をコンピュータに実行させる構成を採る。

10

【0058】

この構成によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声ピッチ情報を取得し、また、できる限り多くの音声情報を取り出してノイズ信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いてノイズ信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

20

【0061】

本発明のサーバ装置は、入力信号のスペクトルを所定の周波数単位で分割した周波数分割スペクトルを出力する周波数分割ステップと、周波数分割スペクトルに基づいて無音と判定した周波数領域の信号を減衰する第一コムフィルタを作成する第一コムフィルタ作成ステップと、周波数分割スペクトルに基づいて第一コムフィルタより多くの雑音のピークを取り除いたフィルタである第二コムフィルタを作成する第二コムフィルタ作成ステップと、前記周波数分割スペクトルから音声ピッチを推定するピッチ推定ステップと、前記第二コムフィルタと前記音声ピッチに基づいてピッチ調波構造を表すピッチ調波修復コムフィルタを作成するピッチ調波構造修復ステップと、前記ピッチ調波修復コムフィルタの通過域と重複する第一コムフィルタの通過域のみを通過域とするように前記第一コムフィルタを修正するコムフィルタ修正ステップと、前記修正された第一コムフィルタを用いて前記周波数分割スペクトルのノイズを抑圧する抑圧ステップと、ノイズを抑圧した前記周波数分割スペクトルを周波数領域で連続したスペクトル信号に合成する音声周波数合成ステップと、をコンピュータに実行させることを特徴とする音声強調プログラムを記憶し、要求に応じて前記音声強調プログラムを出力する構成を採る。

30

40

【0062】

この構成によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声ピッチ情報を取得し、また、できる限り多くの音声情報を取り出して力信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いてノイズ信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

50

【 0 0 6 5 】

【 発明の実施の形態 】

本発明の骨子は、音声信号の周波数分割スペクトルに基づいて音声抑圧に用いるコムフィルタより多くの雑音のピークを取り除いたコムフィルタを生成し、このコムフィルタを用いて音声信号のピッチ情報を取得し、コムフィルタの音声ピッチを補うことである。

【 0 0 6 6 】

以下、本発明の実施の形態について図面を参照して詳細に説明する。

(実施の形態 1)

図 1 は、本発明の実施の形態 1 に係る音声強調装置の構成を示すブロック図である。図 1 において、音声強調装置 100 は、時間分割部 101 と、窓掛け部 102 と、FFT 部 103 と、周波数分割部 104 と、ノイズベース推定部 105 と、第一音声/非音声識別部 106 と、第二音声/非音声識別部 107 と、第一コムフィルタ生成部 108 と、第二コムフィルタ生成部 109 と、有声/無声判別部 110 と、ピッチ推定部 111 と、ピッチ調波構造修復部 112 と、コムフィルタ修正部 113 と、音声分離係数計算部 114 と、乗算部 115 と、音声周波数合成部 116 と、IFFT 部 117 と、から主に構成される。

10

【 0 0 6 7 】

時間分割部 101 は、入力された音声信号から所定時間単位で区切られたフレームを構成し、窓掛け部 102 に出力する窓掛け部 102 は、時間分割部 101 から出力されたフレームにハニングウインドウ等を利用した窓掛け処理を行って FFT 部 103 に出力する。FFT 部 103 は、窓掛け部 102 から出力された音声信号に FFT (Fast Fourier Transform) を行い、音声スペクトル信号を周波数分割部 104 に出力する。

20

【 0 0 6 8 】

周波数分割部 104 は、FFT 部 103 から出力された音声スペクトルを各周波数成分に分割して、周波数成分毎に分割された音声スペクトル $S_f(k)$ (ただし、 k は周波数成分を特定する番号) をノイズベース推定部 105、第一音声/非音声識別部 106、第二音声/非音声識別部 107、及び乗算部 115 に出力する。なお、周波数成分は、所定の周波数単位で分割された音声スペクトルの最小単位を示すものである。 $S_f(k)$ は式 (1) で示される。

【 0 0 6 9 】

【 数 1 】

$$S_f(k) = \sqrt{\operatorname{Re}\{D_f(k)\}^2 + \operatorname{Im}\{D_f(k)\}^2} \quad \dots (1)$$

ここで、 $\operatorname{Re}\{D_f(k)\}^2$ は、FFT 変換後の入力音声信号のスペクトルの実数部を示し、 $\operatorname{Im}\{D_f(k)\}^2$ は、FFT 変換後の入力音声信号のスペクトルの虚数部を示す。

【 0 0 7 0 】

ノイズベース推定部 105 は、フレームに音声成分が含まれていない判定結果が出力された場合、周波数分割部 104 から出力された音声スペクトルの周波数成分毎の短時間パワースペクトルを用いてノイズベースを更新する。

40

【 0 0 7 1 】

具体的には、式 (2) を用いて各周波数成分におけるノイズベースを推定し、推定されたノイズベースを音声/非音声識別部 106 と音声/非音声識別部 107 に出力する。

【 0 0 7 2 】

【 数 2 】

$$P_{base}(n, k) = \begin{cases} P_{base}(n-1, k) & S_f(k) > \Theta_{base} P_{base}(n-1, k) \\ (1-\alpha) \cdot P_{base}(n-1, k) + \alpha \cdot S_f(k) & S_f(k) \leq \Theta_{base} P_{base}(n-1, k) \end{cases} \quad \dots (2)$$

ここで、 $P_{base}(n-1, k)$ はノイズベース、 n は処理を行うフレームを特定する番号

50

、 k は周波数成分を特定する番号を示す。また、 θ_{base} は音声とノイズを判別する閾値、 α は移動平均係数を示す。

【0073】

第一音声/非音声識別部106と第二音声/非音声識別部107は、周波数分割部104から出力された音声スペクトル信号とノイズベース推定部105から出力されるノイズベースの値の差が所定の閾値以上である場合、音声成分を含む有音部分と判定し、それ以外の場合、音声成分を含まない雑音のみの無音部分であると判定する。

【0074】

そして、第一音声/非音声識別部106は、判定結果を第一コムフィルタ生成部108に出力し、第二音声/非音声識別部107は、判定結果を第二コムフィルタ生成部109に出力する。

10

【0075】

第一コムフィルタ生成部108は、音声ピッチ調波情報を多く抽出するように、第一音声/非音声識別部106の閾値を低く設定して、各周波数成分における音声成分の有無に基づいて音声ピッチ調波構造を強調するコムフィルタを生成して、このコムフィルタ結果をコムフィルタ修正部113に出力する。

【0076】

具体的には、以下の式(3)を用いて第一コムフィルタ $COMB_low(k)$ を生成する。

【0077】

【数3】

20

$$COMB_low(k) = \begin{cases} 1 & S_f(k) > \theta_{low} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \\ 0 & S_f(k) \leq \theta_{low} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \end{cases} \quad \dots (3)$$

ここで θ_{low} は第1コムフィルタ用閾値である。また、 HB は、FFT変換長つまり高速フーリエ変換を行うデータ数であり、例えば $HB = 512$ とする。

【0078】

第二コムフィルタ生成部109は、雑音情報に影響されないように、第二音声/非音声識別部107の閾値を高く設定して、各周波数成分における音声成分の有無に基づいて音声ピッチ調波構造を修復する基準コムフィルタとして生成して、このコムフィルタの結果を有声/無声判別部110およびピッチ調波構造修復部112に出力する。具体的には、以下の式(4)を用いて第二コムフィルタを生成する。

30

【数4】

$$COMB_high(k) = \begin{cases} 1 & S_f(k) > \theta_{high} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \\ 0 & S_f(k) \leq \theta_{high} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \end{cases} \quad \dots (4)$$

ここで θ_{high} は第2コムフィルタ用閾値であり、 θ_{high} は θ_{low} より大きい値である。

【0079】

有声/無声判別部110は、第二コムフィルタ生成部109から出力された結果に基づいて有声と無声を判別し、判別結果をピッチ推定部111に出力する。

40

【0080】

具体的には、以下の式(5)と式(6)を用いて入力音声信号の低周波数領域と高周波数領域に分けて、第二コムフィルタの通過域(すなわち、 $COMB_high(k)=1$)の周波数成分の個数を加算する。

【0081】

【数5】

$$VOICE_low(n) = \sum_{k=0}^{k_L} comb_high(k) \quad \dots (5)$$

【 0 0 8 2 】

【 数 6 】

$$VOICE_high(n) = \sum_{k=k_L}^{HB/2} comb_high(k) \quad \dots (6)$$

ここで、式(5)と式(6)共に設定された閾値より大きい場合、または式(5)は設定された閾値より大きく、かつ式(6)は設定された閾値より小さい場合は有声と判別し、それ以外の場合は無声と判別する。

【 0 0 8 3 】

有声と判別した場合は音声ピッチの推定と音声ピッチ調波構造の修復を行い、無声と判別した場合は音声ピッチの推定と音声ピッチ調波構造の修復を行わない。

【 0 0 8 4 】

ピッチ推定部 1 1 1 は、周波数分割部 1 0 4 から出力された音声スペクトルから音声ピッチ周期を推定し、推定結果をピッチ調波構造修復部 1 1 2 に出力する。

【 0 0 8 5 】

ピッチ調波構造修復部 1 1 2 は、ピッチ推定部 1 1 1 から出力された推定結果と第二コムフィルタ生成部 1 0 9 の結果に基づいてピッチの修復を行い、その結果をコムフィルタ修正部 1 1 3 に出力する。

【 0 0 8 6 】

具体的には、音声ピッチ調波構造の修復は以下のステップで行う。第1のステップでは、第二コムフィルタ $COMB_high(k)$ の通過域毎の音声スペクトルのパワのピークを抽出し、全ての通過域からピッチ調波構造修復の基準となるピッチ基準コムフィルタ $COMB_int(k)$ を生成する。

【 0 0 8 7 】

第2のステップでは、ピッチ基準コムフィルタのピークとピークの間隔を計算し、所定の閾値(例えば1.5倍のピッチ周期)を超えたら、ピッチ推定の結果に基づいて欠落したピッチ調波の挿入を行い、ピッチ調波挿入コムフィルタ $COMB_rec(k)$ を生成する。

【 0 0 8 8 】

第3のステップでは、ピッチ周期の値に応じてピッチ調波挿入コムフィルタの櫛、すなわち通過域の幅を広くするピッチ調波修復コムフィルタ $COMB_ext(k)$ を生成する。

【 0 0 8 9 】

コムフィルタ修正部 1 1 3 は、ピッチ調波構造修復部 1 1 2 から出力された推定結果と第一コムフィルタ生成部 1 0 8 から出力された結果を組み合わせるコムフィルタの修正を行い、その結果を音声分離係数計算部 1 1 4 に出力する。

【 0 0 9 0 】

具体的には、ピッチ調波修復コムフィルタ $COMB_ext(k)$ の通過域と第一コムフィルタ $COMB_low(k)$ の通過域を比較して、重複する部分があれば、第一コムフィルタの通過域を修正後のコムフィルタの通過域とし、それ以外の部分は修正後のコムフィルタの阻止域とすることによって修正後のコムフィルタ $COMB_res(k)$ を生成する。

【 0 0 9 1 】

音声分離係数計算部 1 1 4 は、コムフィルタ修正部 1 1 3 において生成されたコムフィルタに、周波数特性に基づいた分離係数を乗算して、周波数成分毎に入力信号の分離係数の設定を行い、各周波数成分の分離係数を乗算部 1 1 5 に出力する。

【 0 0 9 2 】

例えば、以下の式(7)から分離係数 $seps(k)$ を算出して入力信号に乗算することもでき

10

20

30

40

50

る。

【0093】

【数7】

$$seps(k) = \begin{cases} |S_f(k) - \gamma \cdot P_{base}(n, k)| / S_f(k) & COMB_res(k) = 1 \quad 0 \leq k < HB/2 \\ gc \cdot P_{max}(n) / P_{base}(n, k) & COMB_res(k) = 0 \quad 0 \leq k < HB/2 \end{cases} \dots (7)$$

ここで gc は定数、 k は周波数成分を特定する変数、 γ はノイズベースの減算量を調整する係数である。また、 $P_{MAX}(n)$ は、 $P_{base}(n, k)$ の最大値を示す。また、 $gc \cdot P_{MAX}(n) / P_{base}(n, k)$ はフレーム毎にノイズベース推定値の正規化を行い、その逆数を利用した減衰係数であり、 $COMB_res(k)$ はコムフィルタの修正結果である。 10

【0094】

図2及び図3は本発明の音声強調結果の一例であり、上記のコムフィルタの生成、ピッチ調波構造の修復、コムフィルタの修正の各過程および音声分離係数(減衰係数)を示す。図2は、本実施の形態にかかる音声強調装置で作成されるコムフィルタの例を示す図である。図2において、縦軸はスペクトルのパワ及び、フィルタの減衰度を示し、横軸は周波数を示す。

【0095】

コムフィルタは、 $S1$ に示す減衰特性を持ち、減衰特性は、周波数成分毎に設定される。第一コムフィルタ生成部108は、音声成分を含まない周波数領域の信号を減衰し、音声信号を含む周波数領域の信号を減衰しない減衰特性のコムフィルタを作成する。 20

【0096】

雑音成分を含む音声スペクトル $S2$ は、 $S1$ の減衰特性を持つコムフィルタをかけることにより、雑音成分を含む周波数領域の信号が減衰されてパワが小さくなり、音声信号を含む部分は減衰されずパワが変化しない。得られた音声スペクトルは、雑音成分の周波数領域がより低くなりピークが失われずに強調されたスペクトル形状となり、ピッチ調波情報が失われない雑音を抑圧した音声スペクトル $S3$ が出力される。

【0097】

図3に、本実施の形態にかかる音声処理装置におけるコムフィルタの修復の例を示す。図3において、縦軸は減衰度を示し、横軸は、周波数成分を示す。具体的には、横軸には、256の周波数成分があり、0kHzから4kHzの領域を示す。 30

【0098】

$C1$ は生成されたコムフィルタを、 $C2$ はコムフィルタ $C1$ にピッチの修復を行ったコムフィルタを、 $C3$ は、コムフィルタ $C2$ にピッチの幅を修正したコムフィルタを示す。

【0099】

コムフィルタ $C1$ は、100から140までの周波数成分でピッチ情報が失われている。ピッチ調波構造修復部112は、ピッチ推定部111において推定されたピッチ周期情報に基づいてコムフィルタ $C1$ の100から140までの周波数成分にあるピッチ調波情報を補う。これによりコムフィルタ $C2$ が得られる。

【0100】

次に、ピッチ調波構造修復部112は、周波数分割部104から出力された音声スペクトルに基づいてコムフィルタ $C2$ のピッチ調波の幅を修正する。これによりコムフィルタ $C3$ が得られる。 40

【0101】

乗算部115は、周波数分割部104から出力された音声スペクトルに音声分離係数計算部114から出力された分離係数を周波数成分毎に乗算する。そして、乗算した結果を音声周波数合成部116に出力する。

【0102】

音声周波数合成部116は、乗算部115から出力された各周波数成分のスペクトルを所定の処理時間単位で周波数領域の連続する音声スペクトルに合成してIFFT部117に 50

出力する。

【0103】

IFFT部117は、音声周波数合成部116から出力された音声スペクトルにIFFT (Inverse Fast Fourier Transform) を行って音声信号に変換した信号を出力する。

【0104】

このように、本実施の形態の音声強調装置によれば、音声信号のスペクトルから音声ピークである可能性の高いピークを取り出して音声ピッチを推定するコムフィルタを作成し、このコムフィルタから正確な音声ピッチ情報を取得し、また、できる限り多くの音声情報を取り出してノイズ信号の抑圧を行うコムフィルタを作成し、このコムフィルタを用いて雑音に埋もれた音声スペクトルのピークを抑圧しないことにより、正しい音声ピッチ情報に基づいて欠落した音声ピッチ調波構造を補ったコムフィルタを作成することができ、このコムフィルタを用いてノイズ信号を抑圧することにより、音声歪みの少ない音声強調を行うことができる。

10

【0105】

具体的には、本発明の実施の形態の音声強調装置によれば、周波数成分毎のノイズベースの推定及び音声/非音声識別を行うことによって周波数領域において、音声ピッチ調波情報を抽出する第一コムフィルタを生成することができる。

【0106】

また、第二コムフィルタはピッチ調波の基本構造を提供することによって、ピッチ推定誤差による音声の高調波成分がコムフィルタの通過域からはずれることはない。

20

【0107】

また、第二コムフィルタの生成結果に基づいて有声/無声の判別を行い、ピッチ調波構造を持つ有声の場合のみ音声ピッチの推定を行い、その結果に基づいて欠落したピッチ調波構造を修復することにより、雑音に埋もれた音声情報の復元が可能となり、音声ピッチ調波欠落による音声歪を減少することができる。また、修正後のコムフィルタの結果に基づいて周波数成分毎に音声スペクトルの減衰を行うかどうかを決めるので、減衰を大きくしても音声歪の少ない音声強調を行うことができる。

【0108】

また、第一音声/非音声識別閾値を低く設定して第一コムフィルタを生成することによって、音声情報をより多く抽出することができる。一方、第二音声/非音声識別閾値を高く設定して第二コムフィルタを生成することによって、雑音情報の影響を受け難いコムフィルタを生成することができる。その結果に基づいて正確に音声ピッチ調波構造の修復を行うことができる。

30

【0109】

また、第二コムフィルタの生成結果に基づいて有声と無声の判別を行うことによって、少ない演算量で簡単に有声と無声を判別することができる。また、無声区間の場合は音声ピッチ推定と音声ピッチ調波構造の修復を行わないことによって、ピッチ調波構造を持たない無声区間にも対応できる。

【0110】

また、音声ピッチ推定結果に基づいてピッチ調波を挿入することによって、音声ピッチ調波を修復することができる。また、ピッチ調波の幅はピッチの推定結果によって自動的に調整することによって、音声ピッチ推定誤差による影響を低減し、より確実に音声ピッチ調波構造を修復することができる。更に、音声ピッチ調波構造修復の結果と第一コムフィルタの結果を比較し、重複する部分があれば、第一コムフィルタの通過域を修正後のコムフィルタの通過域とし、それ以外は修正後のコムフィルタの阻止域とすることによって、音声ピッチ調波情報のみ抽出し、ピッチ調波間のノイズ情報を抑圧することができる。

40

【0111】

(実施の形態2)

図4は実施の形態2にかかる音声強調装置の構成の例を示すブロック図である。但し、図1と共通する構成については図1と同一番号を付し、詳しい説明を省略する。

50

【 0 1 1 2 】

図 4 の音声強調装置 3 0 0 は、音声 / 雑音フレーム検出部 3 0 1 を具備し、第一フィルタと第二フィルタとから音声スペクトルに音声成分が含まれているか否か判定し、この判定が音声成分を含まない結果である場合、前記第一フィルタに各周波数成分で信号を減衰する修正をする点が図 1 の音声強調装置と異なる。

【 0 1 1 3 】

具体的には、図 4 の音声強調装置 3 0 0 は、第一フィルタの通過域における入力音声パワースペクトルの和と第一フィルタの阻止域における入力音声パワースペクトルの和の比を第 1 結果とし、第二フィルタの通過域における入力音声パワースペクトルの和と第二フィルタの阻止域における入力音声パワースペクトルの和の比を第 2 結果として、それが所定の閾値より大きい場合は第 1 結果と第 2 結果を加算し、所定の閾値より小さい場合は第 2 結果を用いることによって音声 / 雑音フレームを検出する点が、図 1 の音声強調装置と異なる。

10

【 0 1 1 4 】

図 4 において、第一フィルタ生成部 1 0 8 と第二フィルタ生成部 1 0 9 から出力された結果および入力音声パワースペクトルを音声 / 雑音フレーム検出部 3 0 1 に入力し、音声 / 雑音フレーム検出部 3 0 1 で計算した音声 / 雑音フレーム検出結果をフィルタ修正部 1 1 3 に出力する。

【 0 1 1 5 】

具体的には、まず、以下の式 (8) と式 (9) を用いて第一フィルタと第二フィルタに基づく音声と雑音の S N 比を計算する。

20

【 0 1 1 6 】

【 数 8 】

$$SN_low(n) = \frac{\sum_{k=0}^{HB/2} \{S_f(k) \cdot COMB_low(k)\}}{\sum_{k=0}^{HB/2} \{S_f(k) \cdot (1 - COMB_low(k))\}} \quad \dots (8)$$

【 0 1 1 7 】

【 数 9 】

$$SN_high(n) = \frac{\sum_{k=0}^{HB/2} \{S_f(k) \cdot COMB_high(k)\}}{\sum_{k=0}^{HB/2} \{S_f(k) \cdot (1 - COMB_high(k))\}} \quad \dots (9)$$

30

次に、以下の式 (1 0) により当該フレームの S N 比 (SNR__frame(n)) を算出する。

【 0 1 1 8 】

【 数 1 0 】

$$SN_frame(n) = \begin{cases} SN_low(n) + SN_high(n) & SN_frame(n-1) > \Theta_{sn} \\ SN_high(n) & SN_frame(n-1) \leq \Theta_{sn} \end{cases} \quad \dots (10)$$

40

ここで、 Θ_{sn} は閾値である。そして、SNR__frame(n) と Θ_{sn} の比較により音声 / 雑音フレームの検出を行う。音声 / 雑音フレームの検出結果が雑音フレーム (すなわち、SNR__frame(n) < Θ_{sn}) であれば、修正フィルタ COMB__res(k) の各周波数成分をすべて阻止域にする。

【 0 1 1 9 】

このように、本実施の形態の音声強調装置によれば、第一フィルタと第二フィルタとから音声スペクトルに音声成分が含まれているか否か判定し、この判定が音声成分を含まない結果である場合、前記第一フィルタに各周波数成分で信号を減衰する修正をすることにより、突発的に発生するノイズを抑圧することができる。

【 0 1 2 0 】

50

具体的には、本実施の形態の音声強調装置によれば、音声/非音声識別閾値の低い第一コムフィルタの結果に基づいて音声とノイズのSN比を計算することによって、音声と雑音をより検出しやすくなる。一方、音声/非音声識別閾値の高い第二コムフィルタの結果に基づいて音声と雑音のSN比を計算することによって、突発ノイズによる誤検出を少なくすることができる。上記音声/雑音フレーム検出を用いれば、両方の長所を活かすことができ、より確実に音声/雑音フレームを検出し、突発ノイズによる音声/雑音フレームの検出への影響を最小限にすることができる。

【0121】

(実施の形態3)

図5は実施の形態3にかかる音声強調装置の構成の例を示すブロック図である。但し、図1と共通する構成については図1と同一番号を付し、詳しい説明を省略する。

10

【0122】

図5の音声強調装置400は、局部最小値計算部401を具備し、入力信号のパワスペクトルの極小値から所定の範囲を阻止域とするコムフィルタを作成する点が図1の音声強調装置と異なる。

【0123】

具体的には、図5の音声強調装置400は、所定の周波数領域において、入力音声パワスペクトルの局部最小値を第一コムフィルタの阻止域とし、それ以外の周波数成分を通過域とすることによって第一コムフィルタを生成する点が、図1の音声強調装置と異なる。

【0124】

図5において、入力音声スペクトルを局部最小値計算部401に入力し、局部最小値計算部401の出力は第一音声/非音声識別部106により設定された閾値と共に第一コムフィルタ生成部108に入力し、第一コムフィルタ生成部108の結果をコムフィルタ修正部113に出力する。

20

【0125】

局部最小値計算部401は入力された音声パワスペクトルに基づいて、所定の周波数領域における局部最小値に位置する周波数成分を第一コムフィルタの通過域と阻止域の変換点の周波数成分とする。具体的には、所定の周波数領域において、以下の手順で第一コムフィルタを生成する。

【0126】

周波数成分毎に分割された音声スペクトル $S_f(k)$ (ただし、 k は周波数成分を特定する番号)について、隣接する周波数成分 $S_f(k-1)$ 及び $S_f(k+1)$ の両方のパワーより小さい場合、第一コムフィルタを阻止域とし、上記条件を満たさない場合、第一コムフィルタを通過域とする。所定の周波数領域以外の周波数領域に対しては、実施の形態1と同じ手段で第一コムフィルタを生成する。

30

【0127】

このように、本実施の形態の音声強調装置によれば、入力信号のパワスペクトルの極小値から所定の範囲を阻止域とするコムフィルタを作成し、このコムフィルタを用いて入力信号を抑圧することにより、音声とノイズのレベル差が少ない場合でも、音声ピッチ調波構造を抽出して修復して音声歪を低減することができる。

40

【0128】

具体的には、本実施の形態の音声強調装置によれば、所定の周波数領域(特に、低域の周波数帯域帯域)において、局部最小値を抽出し、局部最小値の周波数成分を第一コムフィルタの阻止域とし、それ以外の周波数成分を通過域とすることによって、音声ノイズに埋もれやすい低SN比の環境下でも音声ピッチ調波構造をより確実に抽出して修復ことができ、音声ピッチ調波構造の欠落による音声歪を低減することができる。

【0129】

(実施の形態4)

図6は、実施の形態4にかかる音声強調装置の構成の例を示すブロック図である。但し、図1と共通する構成については図1と同一番号を付し、詳しい説明を省略する。

50

【 0 1 3 0 】

図 6 の音声強調装置 5 0 0 は、ノイズベース減算部 5 0 1 を具備し、入力信号のパワスペクトルからノイズベースを減算し、減算結果の自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチで第一コムフィルタに含まれる音声ピッチ調波構造を修正する点が図 1 の音声強調装置と異なる。

【 0 1 3 1 】

具体的には、図 6 の音声強調装置 5 0 0 は、入力音声のパワスペクトルからノイズベースを減算して自己相関関数を計算し、周波数領域で入力音声のパワスペクトルの自己相関関数に基づいてピッチ推定値を計算する方法を採る点が、図 1 の音声強調装置と異なる。

【 0 1 3 2 】

図 6 において、入力音声パワスペクトルとノイズベース推定部 1 0 5 により推定されたノイズベースの推定値をノイズベース減算部 5 0 1 に入力し、ノイズベース減算部 5 0 1 は入力音声パワスペクトルからノイズベースの推定値を減算し、その結果をピッチ推定部 1 1 1 に入力する。また、有声/無声判別部 1 1 0 から出力されたピッチ推定を行うかどうかの信号もピッチ推定部 1 1 1 に入力する。ピッチ推定部 1 1 1 で推定されたピッチ周期はピッチ調波構造修復部 1 1 2 へ出力する。

【 0 1 3 3 】

ピッチ推定部 1 1 1 は入力音声パワスペクトルからノイズベースを減算した結果を用いて自己相関関数を計算し、自己相関関数の最大値に対応する遅延をピッチ周期とする。

【 0 1 3 4 】

具体的には、以下の式 (1 1) を用いて入力音声パワスペクトルからノイズベースを減算し、式 (1 2) を用いて自己相関関数を計算する。

【 0 1 3 5 】

【 数 1 1 】

$$S'_f(k) = S_f(k) - P_{base}(n, k) \quad \dots (11)$$

【 0 1 3 6 】

【 数 1 2 】

$$r(\tau) = \sum_{k=0}^{k_M} S'_f(k) \cdot S'_f(k + \tau) \quad \dots (12)$$

ここで k_M は周波数の上限である。式 (1 2) で計算された自己相関関数の最大値に対応する τ をピッチ周期とする。

【 0 1 3 7 】

このように、本実施の形態の音声強調装置によれば、入力信号のパワスペクトルからノイズベースを減算し、減算結果の自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチで第一コムフィルタに含まれる音声ピッチ調波構造を修正することにより、ピッチ調波構造の修復を行うことができ、音声歪みの少ない音声強調を行うことができる。

【 0 1 3 8 】

具体的には、本実施の形態の音声強調装置によれば、ピッチ推定部 1 1 1 は入力音声パワスペクトルからノイズベースを減算した結果を用いて自己相関関数を計算することによって、ノイズによるピッチ推定誤差を減少し、より正確にピッチ調波構造の修復を行うことができる。

【 0 1 3 9 】

(実施の形態 5)

図 7 は、実施の形態 5 にかかる音声強調装置の構成の例を示すブロック図である。但し、図 1 と共通する構成については図 1 と同一番号を付し、詳しい説明を省略する。

【 0 1 4 0 】

10

20

30

40

50

図7の音声強調装置600は、入力信号のパワスペクトルからノイズベースを減算した結果に直流成分に所定のパワを持つ擬似ピークを生成し、直流成分を生成したスペクトルの自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチでコムフィルタの音声ピッチを修正する点が図1の音声強調装置と異なる。

【0141】

具体的には、図7の音声強調装置600は、直流成分生成部601を具備し、自己相関関数を計算するとき、適当なエネルギーを持つ擬似パワスペクトルを直流成分として生成し、それに基づいて自己相関関数を計算する点が図1の音声強調装置と異なる。

【0142】

図7において、直流成分生成部601は直流成分に適当なエネルギーを持つパワスペクトルを生成し、ピッチ推定部111に入力する。また、有声/無声判別部110によりピッチ推定を行うかどうかの信号もピッチ推定部111に入力する。ピッチ推定部111で推定されたピッチ周期はピッチ調波構造修復部112に出力する。

10

【0143】

具体的には、第一音声ピッチ調波のパワスペクトルと同じエネルギーを持つ擬似パワスペクトルを直流成分として付加した入力スペクトルを用いて自己相関関数を計算し、その結果に基づいてピッチ周期を推定する。

【0144】

このように、本実施の形態の音声強調装置によれば、入力信号のパワスペクトルからノイズベースを減算した結果に直流成分に所定のパワを持つ擬似ピークを生成し、直流成分を生成したスペクトルの自己相関関数に基づいて音声ピッチを推定し、推定した音声ピッチでピッチ調波構造を修復することにより、ピッチ調波の一部がノイズに埋もれても、ピッチ情報を得てピッチ調波構造の修復を行うことができ、音声歪みの少ない音声強調を行うことができる。

20

【0145】

具体的には、本実施の形態の音声強調装置によれば、入力音声信号は直流成分が含まなくても、周波数領域において、直流成分はピッチ調波の基点として利用することができる。直流成分に適当なエネルギーを持つ擬似パワスペクトルを生成することによって、参照となるピッチ調波は一つが増える。それを利用して自己相関関数を計算すれば、より正確にピッチ推定を行うことができる。特に第一音声ピッチ調波のエネルギーが大きく高次ピッチ調波のエネルギーが小さい場合、あるいは低周波数領域における雑音レベルが高いときにピッチ推定の誤差を減少するために非常に有効である。

30

【0146】

なお、実施の形態5は、実施の形態4と組み合わせることができる。すなわち、図7の音声強調装置に実施の形態4に記載のノイズベース減算部501を用いれば、実施の形態4の効果も得ることができる。

【0147】

(実施の形態6)

図8は実施の形態6にかかる音声強調装置の構成の例を示すブロック図である。但し、図1または図4と共通する構成については図1または図4と同一番号を付し、詳しい説明を省略する。

40

【0148】

図8の音声強調装置700は、第三音声/非音声識別部701と、第三コムフィルタ生成部702と、ノイズ特性推定部703とを具備し、入力信号のノイズ分散値を推定し、第三コムフィルタの通過域における周波数成分の個数の移動平均を計算し、その値が大きい場合はノイズの分散値が大きいと判断し、逆にノイズの分散値が小さいと判断し、その結果に基づいて第二コムフィルタを生成するときの音声/非音声識別閾値を設定する点が、図1の音声強調装置と異なる。

【0149】

具体的には、図8の音声強調装置700は、ノイズの特性を推定する第三コムフィルタを

50

周波数領域で生成し、雑音フレームにおいて、第三コムフィルタの通過域の周波数成分の個数を加算し、その移動平均値に基づいて第二音声/非音声の識別閾値を決定する点が、図1の音声強調装置と異なる。

【0150】

第三音声/非音声識別部701は、周波数分割部104から出力された音声スペクトル信号とノイズベース推定部105から出力されるノイズベースの値の差が所定の閾値以上である場合、通過域と判定し、それ以外の場合、阻止域と判定する。判定結果を第三コムフィルタ生成部702に出力する。

【0151】

第三コムフィルタ生成部702は、第三音声/非音声識別部701から出力された音声/非音声識別結果に基づいてコムフィルタの通過域/阻止域を生成し、その結果をノイズ特性推定部703に出力する。ノイズ特性推定部703は、音声/雑音フレーム検出部301から検出された雑音フレームにおいて、第三コムフィルタの通過域における周波数成分の個数を加算し、所定のフレーム数に渡って平均値を計算し、その結果を第二音声/非音声識別部107に出力する。具体的には、以下の式(13)を用いてノイズ特性の推定を行う。

【0152】

【数13】

$$NS_var(n) = (1 - \alpha_v) \cdot NS_var(n-1) + \alpha_v \cdot \sum_{k=0}^{HB/2} COMB_var(k) \quad \dots (13)$$

ここで、COMB_var(k)は第三コムフィルタ、NS_var(n)はノイズ特性の推定結果、 α_v は移動平均係数である。

【0153】

そして、第二音声/非音声識別閾値をNS_var(n)で適応的に制御するようにし、NS_var(n)の値が大きければ、雑音の特性として分散が大きいと判断し、第二音声/非音声識別閾値を高く設定し、逆に、その値が小さければ、雑音の特性として分散が小さいと判断し、第二音声/非音声識別閾値を低く設定する。

【0154】

このように、本実施の形態の音声強調装置によれば、入力信号のノイズレベルの分布を検出し、この分布に基づいて音声スペクトルからコムフィルタを生成する基準を決定し、ノイズの種類に対応して雑音を抑圧することができ、音声歪みの少ない音声強調を行うことができる。

【0155】

具体的には、本実施の形態の音声強調装置によれば、ノイズの特性を推定するための第三コムフィルタを設けることによって、簡単な計算で間接的にノイズの分散値を推定することができ、その結果に基づいて第二音声/非音声識別閾値を設定すれば、第二コムフィルタを生成するとき、分散値の大きいノイズによる偽のピッチ調波の混入を減少することができる。また、分散値の小さいノイズの場合は、より多くの音声ピッチ調波情報を残すことができる。

【0156】

(実施の形態7)

図9は、実施の形態7にかかる音声強調装置の構成の例を示すブロック図である。但し、図1及び図4と共通する構成については図1及び図4と同一番号を付し、詳しい説明を省略する。

【0157】

図9の音声強調装置800は、周波数領域選択部801を具備し、音声ピッチ調波構造の修復を行うとき、第三コムフィルタの結果に基づいて第二コムフィルタの所定の周波数領域における周波数成分をすべて阻止域に変換する点が、図1の音声強調装置と異なる。

【 0 1 5 8 】

図 9 において、ノイズ特性推定部 7 0 3 は、第三コムフィルタ生成部 7 0 2 から出力された結果に基づいてノイズの特性を推定し、その結果を周波数領域選択部 8 0 1 に出力する。周波数領域選択部 8 0 1 は、ノイズ特性の推定結果に基づいて第二コムフィルタを阻止域にする中間周波数領域を決定し、その結果を第二コムフィルタ生成部 1 0 9 に出力する。

【 0 1 5 9 】

具体的には、ノイズ特性推定部 7 0 3 により計算されたノイズ特性の移動平均値がある閾値を越えたら、分散値の大きいノイズと判断し、第二コムフィルタの中間周波数領域、例えば 1 kHz ~ 2 kHz 間の周波数成分をすべて阻止域に変換する。

10

【 0 1 6 0 】

このように、本実施の形態の音声強調装置によれば、ノイズ特性の推定結果に基づいて周波数成分を選択し、第二コムフィルタにおいて、選択された周波数領域をすべて阻止域に変換することによって、分散値の大きいノイズにより生成した偽のピッチ調波を減少し、偽のピッチ調波を生じにくい低周波数領域におけるピッチ調波を基準に、実施の形態 1 に記載したピッチ調波構造の修復を行えば、正確にピッチ調波構造を修復することができる。

【 0 1 6 1 】

なお、実施の形態 7 は、実施の形態 6 と組み合わせることができる。すなわち、図 9 の音声強調装置に実施の形態 6 に記載のノイズ特性推定部 7 0 3 を用いれば、実施の形態 6 の効果も得ることができる。

20

【 0 1 6 2 】

(実施の形態 8)

図 10 は、実施の形態 8 にかかる音声強調装置の構成の例を示すブロック図である。但し、図 1 と共通する構成については図 1 と同一番号を付し、詳しい説明を省略する。

【 0 1 6 3 】

図 10 の音声強調装置 9 0 0 は、SNR 推定部 9 0 1 を具備し、音声分離係数計算手段において、ノイズの減衰量を SNR 推定値の大きさに応じて調整する点が、図 1 の音声強調装置と異なる。

【 0 1 6 4 】

図 10 において、SNR 推定部 9 0 1 は入力音声パワースペクトルおよびノイズベース推定部 1 0 5 から出力されたノイズベースの推定値に基づいて音声レベルと雑音レベルの比を計算し、その結果を音声分離係数計算部 1 1 4 に出力する。音声分離係数計算部 1 1 4 は SNR 推定値の大きさに応じて周波数成分毎のノイズの減衰量を計算し、乗算部 1 1 5 で入力音声スペクトルと乗算する。具体的には、以下の式 (14) ~ (16) を用いて SNR を計算する。

30

【 0 1 6 5 】

【数 1 4】

$$S_p(n) = (1 - \alpha_p) \cdot S_p(n-1) + \alpha_p \cdot \sum_{k=0}^{HB/2} S_f(k) \quad \dots (14)$$

40

【 0 1 6 6 】

【数 1 5】

$$N_s(n) = \sum_{k=0}^{HB/2} P_{base}(n, k) \quad \dots (15)$$

【 0 1 6 7 】

【数 1 6】

50

$$SNR(n) = \frac{S_p(n)}{N_s(n)} \quad \dots (16)$$

ここで、 $S_p(n)$ は音声レベルの移動平均値を、 $N_s(n)$ はノイズの移動平均値を、 s は移動平均係数を示す。

【 0 1 6 8 】

SNRの値に応じて、修正コムフィルタの通過域と阻止域において、音声分離係数(ノイズの減衰量)は以下の式(17)を用いて計算する。

【 0 1 6 9 】

【 数 1 7 】

$$seps(k) = \begin{cases} |S_f(k) - \gamma(n) \cdot P_{base}(n,k)| / S_f(k) & COMB_res(k) = 1 \quad 0 \leq k < HB/2 \\ gc(n) \cdot P_{max}(n) / P_{base}(n,k) & COMB_res(k) = 0 \quad 0 \leq k < HB/2 \end{cases} \quad \dots (17)$$

ここで、 $\gamma(n)$ はノイズベースを減算する量を示す係数で、 $gc(n)$ はノイズ減衰の度合い係数である。異なるSNRの環境に対応できるように、 $\gamma(n)$ と $gc(n)$ の値をSNRの推定値より自動的に調整できるようにする。例えば、 $\gamma(n)$ の大きさはSNR(n)の値に正比例し、 $gc(n)$ の大きさはSNR(n)の値に反比例するように自動的に調整する。

【 0 1 7 0 】

このように、本実施の形態の音声分離装置によれば、修正コムフィルタの通過域と阻止域において、入力音声パワースペクトルからノイズベースを減算する量とノイズ減衰の度合いをSNR推定値の大きさに応じて調整することによって、異なるSNRの環境下でも適切なノイズ減衰を行い、音声歪と残留ノイズの少ない音声強調を実現できる。

【 0 1 7 1 】

(実施の形態9)

図11は、実施の形態9にかかる音声強調装置の構成の例を示すブロック図である。但し、図1及び図10と共通する構成については図1及び図10と同一番号を付し、詳しい説明を省略する。

【 0 1 7 2 】

図11の音声強調装置1000は、ノイズベース減算部1001と、重み係数計算部1002とを具備し、音声成分のレベルを入力信号のパワースペクトルの移動平均値より算出し、雑音成分のレベルをノイズベースの推定値に各周波数成分別に重み計数を乗算した値より算出し、前記音声成分のレベルと前記雑音成分のレベルの比から信号対雑音比を算出する点が、図1の音声強調装置と異なる。

【 0 1 7 3 】

図11において、ノイズベース推定部105はノイズベースを推定し、その結果をノイズベース減算部1001と重み係数計算部1002に出力する。ノイズベース減算部1001は入力音声パワースペクトルの移動平均値を計算し、移動平均値からノイズベースの推定値を減算してSNR推定部901に出力する。重み係数計算部1002はノイズベース推定値の各周波数成分に重み係数を計算し、その結果をSNR推定部901に出力する。SNR推定部901は音声レベルと雑音レベルの比を計算し、その結果を音声分離係数計算部114に出力する。具体的には、以下の式(18)~式(20)を用いてSNRを計算する。

【 0 1 7 4 】

【 数 1 8 】

10

20

30

40

$$S'_p(n) = S_p(n) - \beta \cdot N_s(n) \quad \dots (18)$$

【 0 1 7 5 】

【 数 1 9 】

$$N'_s(n) = \sum_{k=0}^{HB/2} \delta(k) \cdot P_{base}(n, k) \quad \dots (19)$$

【 0 1 7 6 】

【 数 2 0 】

$$SNR'(n) = \frac{S'_p(n)}{N'_s(n)} \quad \dots (20)$$

ここで、 β はノイズベースを減算する量を示す係数で、 $\delta(k)$ は重み係数である。重み係数 $\delta(k)$ は音声の特性を利用して設定する。例えば、中間周波数領域における音声スペクトルのエネルギーが小さいが音声の明瞭性に与える影響が大きいため、中間周波数領域におけるノイズに対して重み係数の値を増やして雑音レベルを計算すれば、異なるノイズに対して適切な減衰を行うことができる。

【 0 1 7 7 】

このように、本実施の形態の音声強調装置によれば、入力音声パワースペクトルの移動平均値からノイズベースを減算して音声レベルの計算することにより、ノイズの影響を減少し、低 SNR の環境下でも正確な音声レベルを計算することができる。また、ノイズベース推定値の各周波数成分に重み係数を乗算して計算することにより、異なるノイズに対して適切な減衰を行い、音声歪を低減することができる。

【 0 1 7 8 】

なお、実施の形態 9 は、実施の形態 8 と組み合わせることができる。すなわち、図 11 の音声強調装置に実施の形態 8 に記載の SNR 推定部 901 を用いれば、実施の形態 8 の効果も得ることができる。

【 0 1 7 9 】

(実施の形態 10)

図 12 は、実施の形態 10 にかかる音声強調装置の構成の例を示すブロック図である。但し、図 1、図 10 及び図 11 と共通する構成については図 1、図 10 及び図 11 と同一番号を付し、詳しい説明を省略する。

【 0 1 8 0 】

図 12 の音声強調装置 1100 は、SNR の変動を抑える SNR 変動抑圧部 1101 を具備し、SNR 変動抑圧部 1101 は、SNR 推定値と SNR 推定値の長期移動平均値の結果に基づいて SNR の変動を抑圧する点が、図 1 の音声強調装置と異なる。

【 0 1 8 1 】

図 12 において、SNR 推定部 901 は音声レベルと雑音レベルの比を計算し、その結果を SNR 変動抑圧部 1101 に出力する。SNR 変動抑圧部 1101 は SNR 推定値に基づいて SNR の長期移動平均値を計算し、その結果と SNR 推定値の偏差を計算し、SNR 推定値の長期移動平均値と前記偏差の一部を加算して SNR 変動抑圧結果として用いる。そして、変動が抑圧された SNR 推定値を音声分離係数計算部 114 に出力する。

【 0 1 8 2 】

具体的には、以下の式 (21) を用いて SNR の長期移動平均値を計算し、式 (22) を用いて SNR の変動を抑圧する SNR の推定値を計算する。

【 0 1 8 3 】

【 数 2 1 】

10

20

30

40

50

$$\overline{SNR}'(n) = (1 - \alpha_r) \cdot \overline{SNR}'(n-1) + \alpha_r \cdot SNR'(n) \quad \dots (21)$$

【 0 1 8 4 】

【 数 2 2 】

$$\overline{SNR}(n) = \overline{SNR}'(n) + \left(SNR'(n) - \overline{SNR}'(n) \right) / \mu \quad \dots (22)$$

ここで、 α_r は移動平均係数で、 μ は加算する偏差の大きさを決める係数である。

【 0 1 8 5 】

このように、本実施の形態の音声強調装置によれば、SNR推定値とSNR推定値の長期移動平均値の偏差を計算し、SNR推定値の長期移動平均値と前記偏差の一部を加算してSNR推定値として用いることによって、SNRの変動を有効に抑制し、安定的にSNRの大きさに応じてノイズ減衰のレベル調整を行うことができる。

【 0 1 8 6 】

なお、実施の形態10は、実施の形態8あるいは実施の形態9と組み合わせることができる。すなわち、図12の音声強調装置に実施の形態8に記載のSNR推定部901を用いれば、実施の形態8の効果も得ることができ、図12の音声強調装置に実施の形態9に記載のノイズベース減算と重み係数計算手段を用いてSNRの推定を行えば、実施の形態9の効果も得ることができる。

【 0 1 8 7 】

(実施の形態11)

図13は、実施の形態11にかかる音声強調装置の構成の例を示すブロック図である。但し、図1と共通する構成については図1と同一番号を付し、詳しい説明を省略する。

【 0 1 8 8 】

図13の音声強調装置1200は、更新速度の速いノイズベース更新部1201を具備し、音声区間においてもノイズベースの追跡を行うことができる点が、図1の音声強調装置と異なる。

【 0 1 8 9 】

図13において、ノイズベース更新部1201は、入力音声パワースペクトルに基づいて更新速度の速い移動平均係数を用いてノイズベースを推定し、その結果をノイズベース推定部105に出力する。ノイズベース推定部105は、更新速度の遅い移動平均係数を用いてノイズベースを推定し、その結果を第一音声/非音声識別部106および第二音声/非音声識別部107に出力する。

【 0 1 9 0 】

具体的には、以下の式(23)と式(24)を用いて更新速度の速いノイズベースと更新速度の遅いノイズベースを推定する。

【 0 1 9 1 】

【 数 2 3 】

$$P_{fast}(n, k) = \begin{cases} P_{fast}(n-1, k) & S_f(k) > \Theta_{fast} P_{fast}(n-1, k) \\ (1 - \alpha_f) \cdot P_{fast}(n-1, k) + \alpha_f \cdot S_f(k) & S_f(k) \leq \Theta_{fast} P_{fast}(n-1, k) \end{cases} \quad \dots (23)$$

【 0 1 9 2 】

【 数 2 4 】

10

20

30

40

$$P_{slow}(n, k) = \begin{cases} P_{slow}(n-1, k) & S_f(k) > \Theta_{fast} P_{fast}(n-1, k) \\ (1-\alpha_s) \cdot P_{slow}(n-1, k) + \alpha_s \cdot S_f(k) & S_f(k) \leq \Theta_{fast} P_{fast}(n-1, k) \end{cases} \dots (24)$$

ここで、 α_f と α_s はそれぞれ速い更新係数と遅い更新係数であり、 Θ_{fast} は音声と雑音を識別する閾値である。

【0193】

このように、本実施の形態の音声強調装置によれば、更新速度の速い移動平均係数を用いてノイズベースの推定を行うことによって、音声区間においても雑音レベルの急激な変動を追跡することができる。また、更新速度の遅いノイズベースの更新は、更新速度の速いノイズベースに基づいて行うことによって、正確にノイズベースの推定を行うことができ、雑音レベルの急激な変動によるノイズベース更新の停止を防止することができる。

10

【0194】

(実施の形態12)

図14は、本発明の実施の形態12に係る雑音分離装置の構成を示すブロック図である。本実施の形態の雑音分離装置1300は、雑音を含む音声信号から雑音信号を分離、抽出するものである。

【0195】

図14において、雑音分離装置1300は、時間分割部101と、窓掛け部102と、FFT部103と、周波数分割部104と、ノイズベース推定部105と、音声/非音声識別部1301と、雑音コムフィルタ生成部1302と、実数虚数分離部1303と、雑音分離係数計算部1304と、乗算部1305と、雑音周波数合成部1306と、IFFT部1307と、から主に構成される。

20

【0196】

ここで、但し、図1と共通する構成については図1と同一番号を付し、詳しい説明を省略する。

【0197】

音声/非音声識別部1301は、周波数分割部104から出力された音声スペクトル信号とノイズベース推定部105から出力されたノイズベースの推定値の差が所定の閾値以上である場合、音声成分を含む有音部分と判定し、それ以外の場合、音声成分を含まない雑音のみの無音部分であると判定し、その結果を雑音コムフィルタ生成部1302に出力する。雑音コムフィルタ生成部1302は、音声/非音声識別部1301の結果に基づいて雑音分離コムフィルタを生成して、このコムフィルタを実数虚数分離部1303に出力する。

30

【0198】

具体的には、雑音コムフィルタ生成部1302は、音声情報を抑制するように音声/非音声識別の閾値を低く設定して以下の式(25)を用いて雑音分離コムフィルタを生成する。

【0199】

【数25】

$$COMB_nos(k) = \begin{cases} 1 & S_f(k) > \Theta_{nos} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \\ 0 & S_f(k) \leq \Theta_{nos} \cdot P_{base}(n, k) & 0 \leq k < HB/2 \end{cases} \dots (25)$$

40

ここで Θ_{nos} は雑音分離用閾値である。

【0200】

実数虚数分離部1303は、入力音声スペクトルの実数部と虚数部を分離し、その結果を雑音分離係数計算部1304に出力する。雑音分離係数計算部1304は、雑音分離コムフィルタの通過域と阻止域に対して、周波数成分毎に別々の分離係数を計算し、その結果

50

を乗算部 1305 に出力する。

【0201】

具体的には、以下の式(26)と式(27)を用いて、雑音分離コムフィルタの阻止域において、雑音分離係数は1とし、雑音分離コムフィルタの通過域において、雑音分離係数は入力音声スペクトルの実数部と虚数部に対して、別々の乱数とノイズベースの推定値を乗算する。

【0202】

【数26】

$$sepn_{re}(k) = \begin{cases} 1 & COMB_nos(k) = 0 \quad 0 \leq k < HB/2 \\ rd_{re}(i) \cdot P_{base}(n, k) / S_f(k) & COMB_nos(k) = 1 \quad 0 \leq k < HB/2 \\ \dots & \dots \end{cases} \quad (26)$$

【0203】

【数27】

$$sepn_{im}(k) = \begin{cases} 1 & COMB_nos(k) = 0 \quad 0 \leq k < HB/2 \\ rd_{im}(i) \cdot P_{base}(n, k) / S_f(k) & COMB_nos(k) = 1 \quad 0 \leq k < HB/2 \end{cases} \quad \dots \quad (27)$$

ここで $rd_{re}(i)$ は、実数部に用いられるランダム関数で均一分布の乱数により構成され、 $rd_{im}(i)$ は、虚数部に用いられるランダム関数で均一分布の乱数により構成される。

【0204】

乗算部 1305 は、周波数分割部 104 から出力された音声スペクトルに雑音分離係数計算部 1304 から出力された分離係数を周波数成分毎に乗算する。そして、乗算の結果得られた雑音スペクトルを雑音周波数合成部 1306 に出力する。

【0205】

雑音周波数合成部 1306 は、乗算部 1305 から出力された周波数成分のスペクトルを所定の処理時間単位で、周波数領域で連続する雑音スペクトルに合成して IFFT 部 1307 に出力する。

【0206】

IFFT 部 1307 は、雑音周波数合成部 1306 から出力された雑音スペクトルに IFFT (Inverse Fast Fourier Transform) を行って雑音信号に変換した信号を出力する。

【0207】

このように、本実施の形態の雑音分離装置によれば、雑音専用コムフィルタを生成することにより、雑音の特性を最大限に抽出することができる。また、雑音分離コムフィルタの阻止域において、雑音成分を減衰せず、雑音分離コムフィルタの通過域において、入力音声スペクトルの実数部と虚数部に対して、別々の乱数とノイズベースの推定値を乗算することによって、雑音成分の実数部と虚数部の振幅と位相はすべてランダム化され、良好な雑音分離特性を得ることができる。

【0208】

(実施の形態 13)

図 15 は、実施の形態 13 に係る雑音分離装置の構成の例を示すブロック図である。但し、図 1 及び図 14 と共通する構成については図 1 及び図 14 と同一番号を付し、詳しい説明を省略する。

【0209】

図 15 の雑音分離装置は、雑音成分保存部 1401 を具備し、雑音分離用コムフィルタの阻止域における入力音声のスペクトル成分をメモリに保存し、その値を雑音分離用コムフィルタの通過域に用いる点が、図 14 と異なる。

10

20

30

40

50

【 0 2 1 0 】

図 1 5 において、雑音成分保存部 1 4 0 1 は、雑音コムフィルタ生成部 1 3 0 2 から出力された雑音コムフィルタの阻止域において入力音声スペクトルを保存し、雑音コムフィルタの通過域において雑音成分保存結果を雑音分離係数計算部 1 3 0 4 に出力する。

【 0 2 1 1 】

具体的には、例えば所定数のメモリを用意して雑音コムフィルタの阻止域における入力音声スペクトルを低周波数領域から高周波数領域まで順次に保存し、雑音コムフィルタの通過域において同じ低周波数領域から高周波数領域までの順で、最新の保存されたデータかつもっとも近い周波数成分を有するものを選択して雑音コムフィルタの通過域における入力音声スペクトルとして用いる。

10

【 0 2 1 2 】

このように、本実施の形態の雑音分離装置によれば、雑音分離用コムフィルタの阻止域における入力音声のスペクトル成分をメモリに保存し、その値を雑音分離用コムフィルタの通過域に用いることにより、実際の雑音と特性の近い擬似雑音を再構成することができ、良好な雑音分離特性を得ることができる。

【 0 2 1 3 】

なお、実施の形態 1 3 は、実施の形態 1 2 と組み合わせることができる。すなわち、図 1 5 の雑音分離装置に実施の形態 1 2 に雑音分離係数計算部 1 3 0 4 を用いれば、実施の形態 1 2 の効果も得ることができる。

【 0 2 1 4 】

なお、本発明は上記実施の形態に限定されず、複数の実施の形態を組み合わせる、または種々変更して実施することが可能である。例えば、上記実施の形態では、音声強調装置または雑音抑圧装置として行う場合について説明しているが、これに限られるものではなく、この音声強調方法または雑音抑圧方法をソフトウェアとして行うことも可能である。

20

【 0 2 1 5 】

例えば、上記音声強調方法または雑音抑圧方法を実行するプログラムを予め R O M (Read Only Memory) に格納しておき、そのプログラムを C P U (Central Processor Unit) によって動作させるようにしても良い。

【 0 2 1 6 】

また、上記音声強調方法または雑音抑圧方法を実行するプログラムをコンピュータで読み取り可能な記憶媒体に格納し、記憶媒体に格納されたプログラムをコンピュータの R A M (Random Access memory) に記録して、コンピュータをそのプログラムにしたがって動作させるようにしても良い。

30

【 0 2 1 7 】

また、上記音声強調または雑音抑圧を行うプログラムをサーバに格納し、サーバに格納されたプログラムをクライアントに転送して、クライアント上でそのプログラムを実行させてもよい。このような場合においても、上記実施の形態と同様の作用及び効果を呈する。

【 0 2 1 8 】

また、上記いずれかの実施の形態に係る音声強調装置または雑音抑圧装置は、無線通信装置、通信端末、基地局装置等に搭載することもできる。この結果、通信時の音声を音声強調または雑音抽出できる。

40

【 0 2 1 9 】**【 発明の効果 】**

以上説明したように、本発明の音声強調装置及び音声強調方法によれば、音声信号の周波数分割スペクトルに基づいて音声抑圧に用いるコムフィルタより多くの雑音のピークを取り除いたコムフィルタを生成し、このコムフィルタを用いて音声信号のピッチ情報を取得し、コムフィルタの音声ピッチを補うことにより、音声の歪みが少なくかつ雑音を十分に除去することができる。

【 図面の簡単な説明 】

【 図 1 】 本発明の実施の形態 1 に係る音声強調装置の構成を示すブロック図

50

- 【図2】上記実施の形態にかかる音声強調装置で作成されるコムフィルタの例を示す図
 【図3】上記実施の形態にかかる音声処理装置におけるコムフィルタの修復の例を示す図
 【図4】本発明の実施の形態2に係る音声強調装置の構成を示すブロック図
 【図5】本発明の実施の形態3に係る音声強調装置の構成を示すブロック図
 【図6】本発明の実施の形態4に係る音声強調装置の構成を示すブロック図
 【図7】本発明の実施の形態5に係る音声強調装置の構成を示すブロック図
 【図8】本発明の実施の形態6に係る音声強調装置の構成を示すブロック図
 【図9】本発明の実施の形態7に係る音声強調装置の構成を示すブロック図
 【図10】本発明の実施の形態8に係る音声強調装置の構成を示すブロック図
 【図11】本発明の実施の形態9に係る音声強調装置の構成を示すブロック図
 【図12】本発明の実施の形態10に係る音声強調装置の構成を示すブロック図
 【図13】本発明の実施の形態11に係る音声強調装置の構成を示すブロック図
 【図14】本発明の実施の形態12に係る雑音分離装置の構成を示すブロック図
 【図15】本発明の実施の形態13に係る雑音分離装置の構成を示すブロック図

10

【符号の説明】

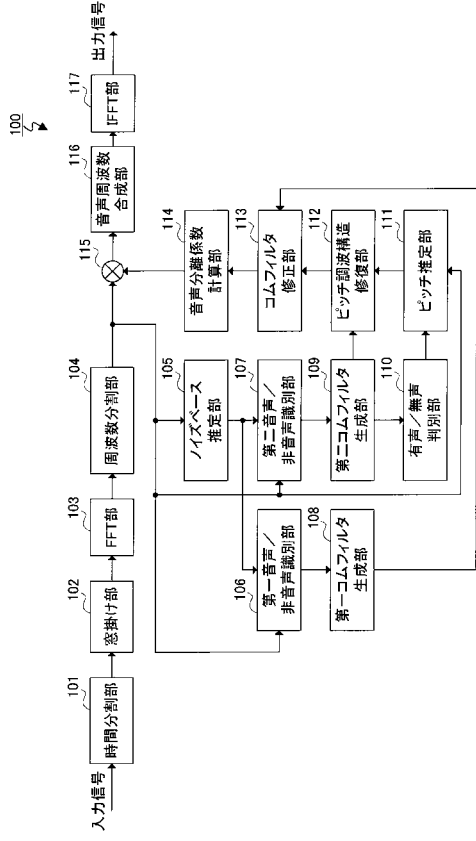
- 103 F F T部
 104 周波数分割部
 105 ノイズベース推定部
 106 第一音声/非音声識別部
 107 第二音声/非音声識別部
 108 第一コムフィルタ生成部
 109 第二コムフィルタ生成部
 110 有音/無声判別部
 111 ピッチ推定部
 112 ピッチ調波構造修復部
 113 コムフィルタ修正部
 114 音声分離係数計算部
 115、1305 乗算部
 116 音声周波数合成部
 117 I F F T部
 301 音声/雑音フレーム検出部
 401 局部最小値計算部
 501、1001 ノイズベース減算部
 601 直流成分生成部
 701 第3音声/非音声識別部
 702 第三コムフィルタ生成部
 703 ノイズ特性推定部
 801 周波数領域選択部
 901 S N R推定部
 1002 重み係数計算部
 1101 S N R変動抑圧部
 1201 ノイズベース更新部
 1301 音声/非音声識別部
 1302 雑音コムフィルタ生成部
 1303 実数虚数分離部
 1304 雑音分離係数計算部
 1306 雑音周波数合成部
 1401 雑音成分保存部

20

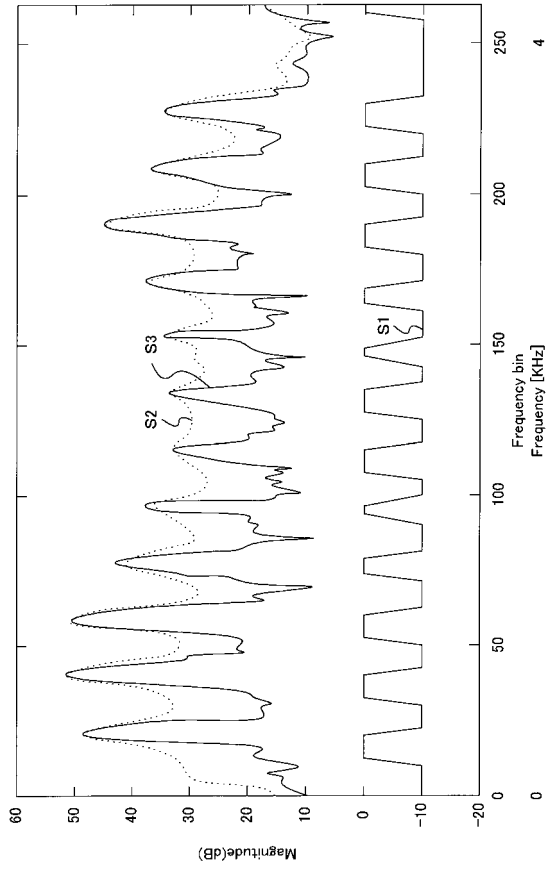
30

40

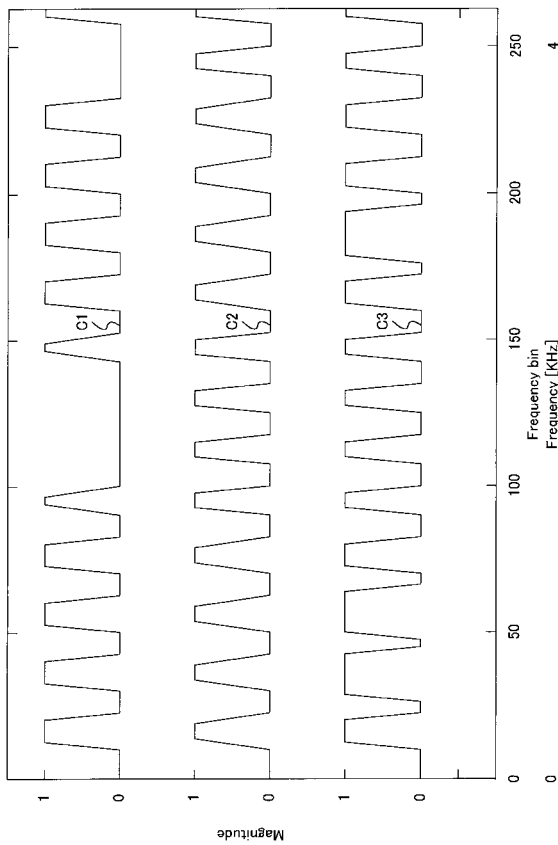
【図1】



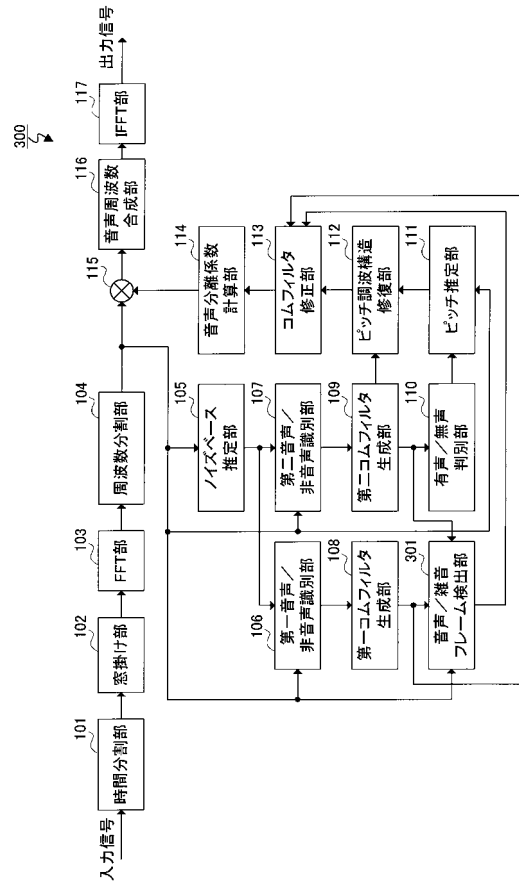
【図2】



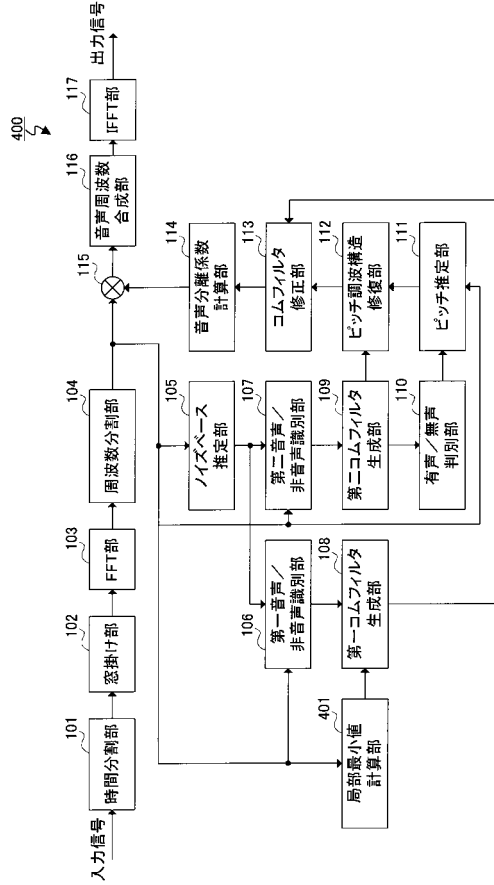
【図3】



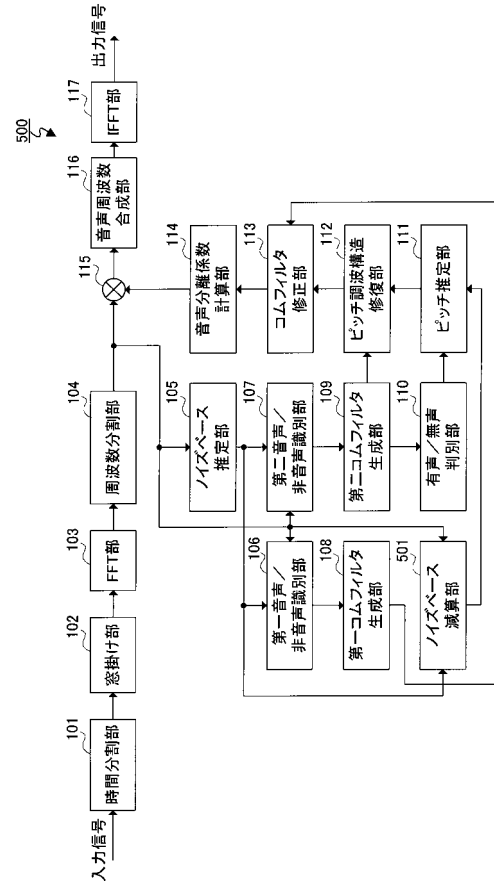
【図4】



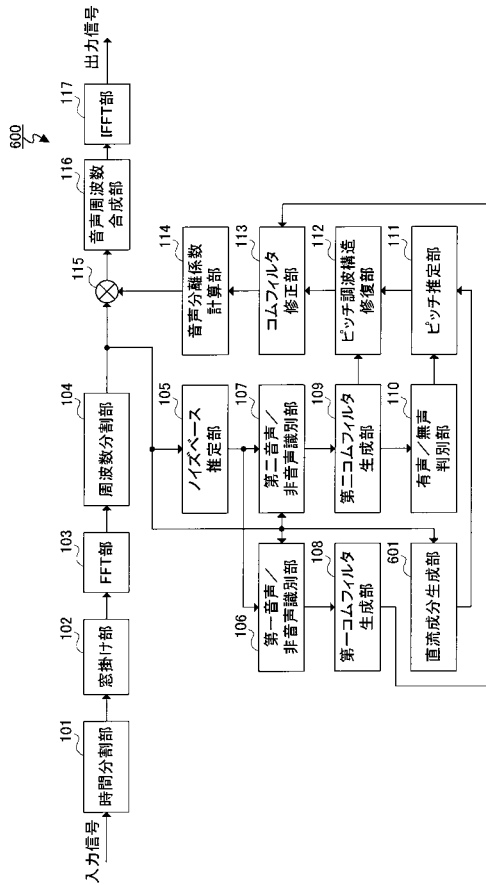
【図5】



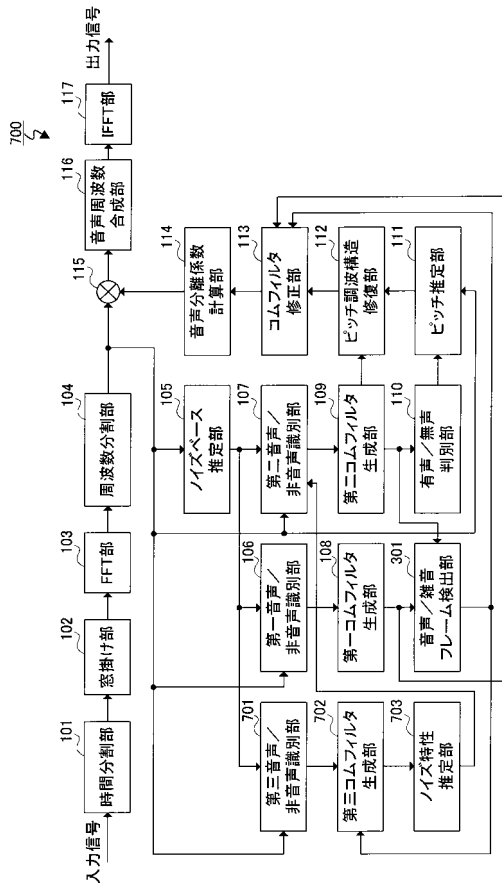
【図6】



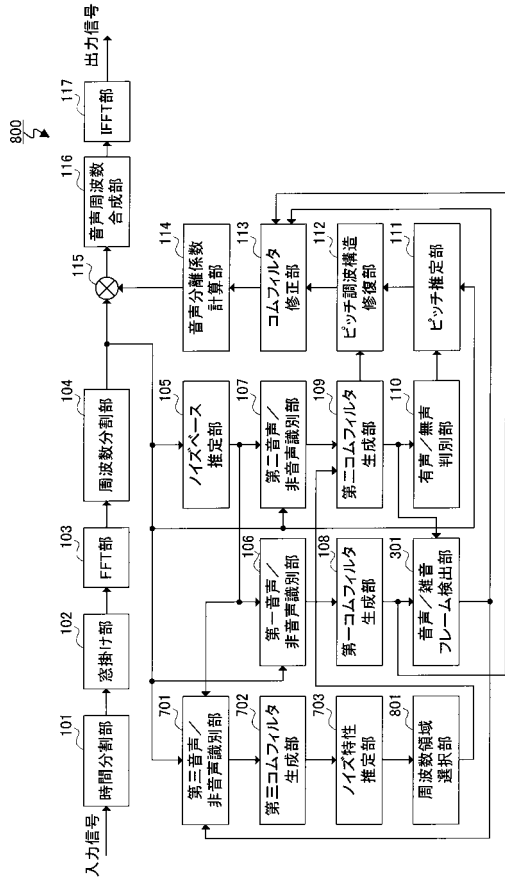
【図7】



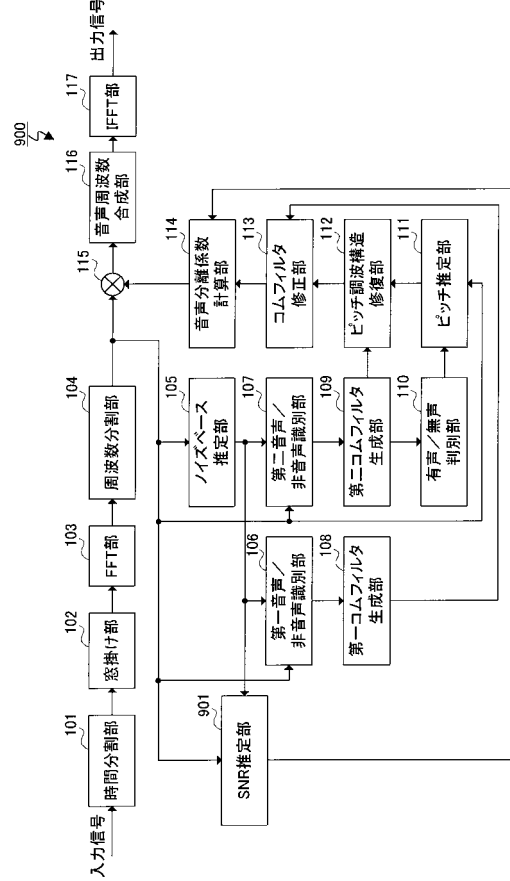
【図8】



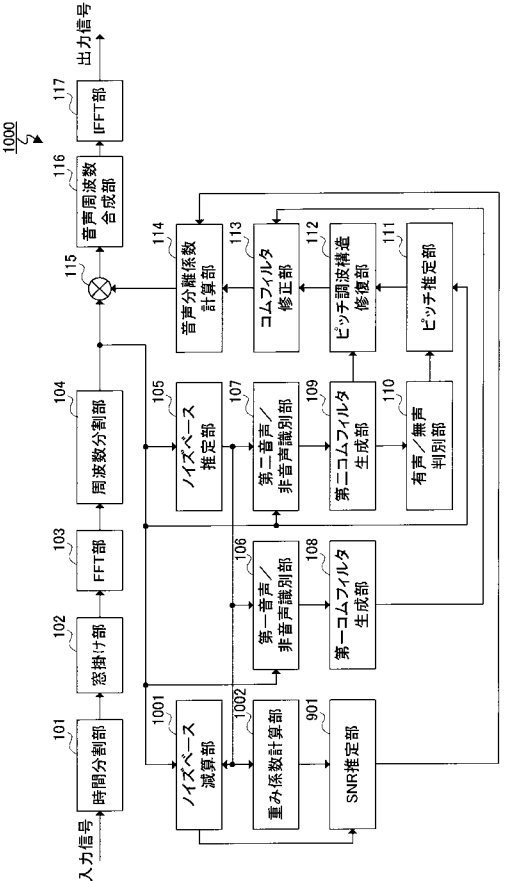
【図9】



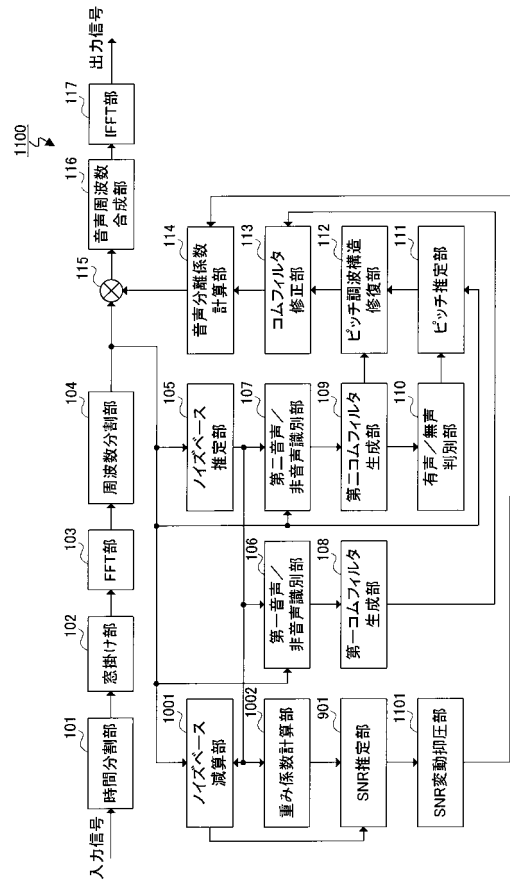
【図10】



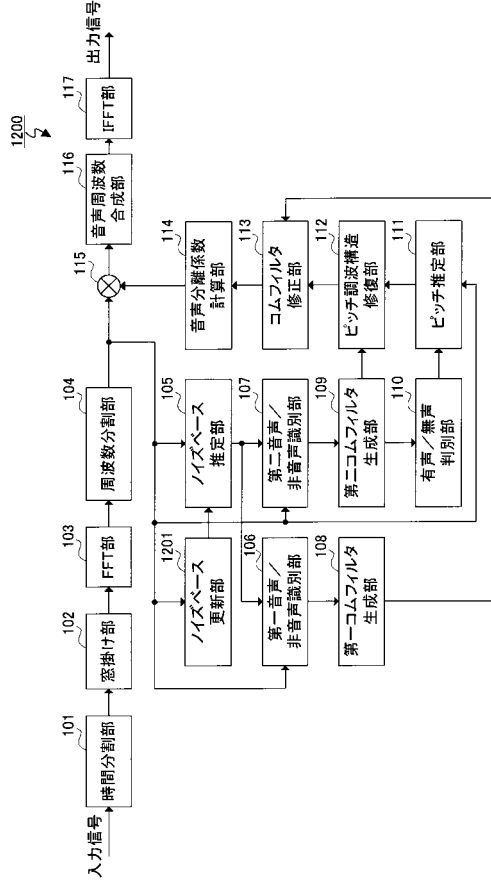
【図11】



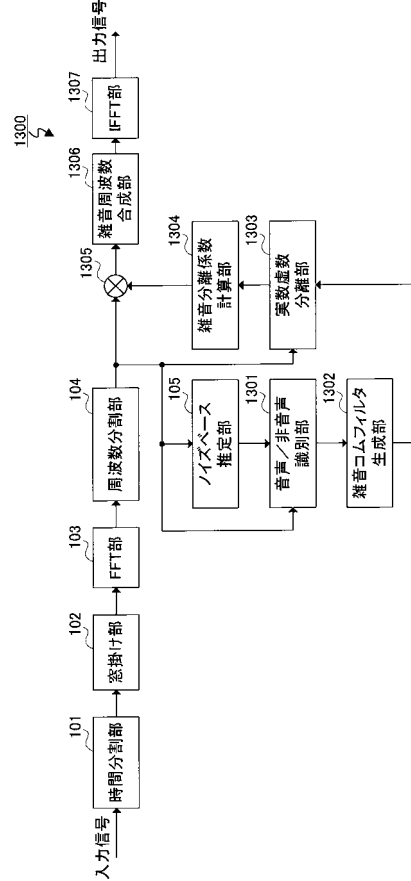
【図12】



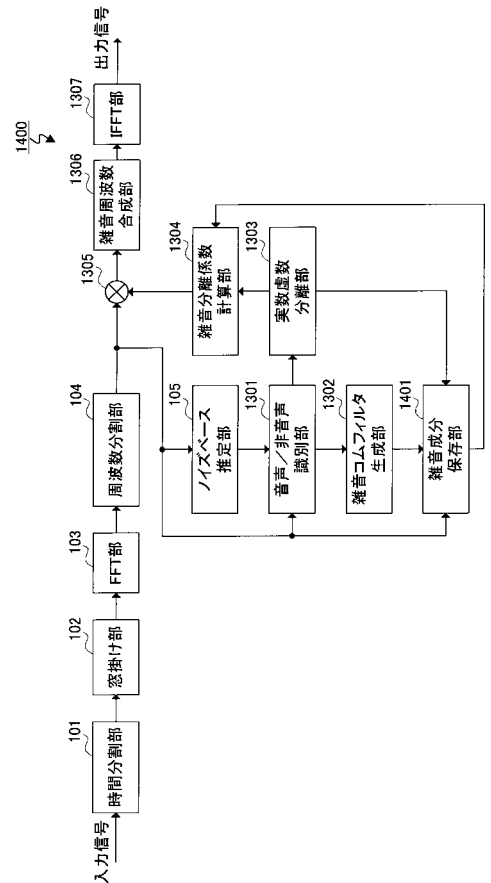
【図 13】



【図 14】



【図 15】



フロントページの続き

(56)参考文献 特開2001-249676(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 11/04,21/02

H04B 1/10