

(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2009-31809

(P2009-31809A)

(43) 公開日 平成21年2月12日(2009.2.12)

(51) Int.Cl.		F I			テーマコード (参考)
G 1 0 L 15/20	(2006.01)	G 1 0 L 15/20	3 7 0 D		5 D 0 1 5
G 1 0 L 15/28	(2006.01)	G 1 0 L 15/20	3 7 0 F		
		G 1 0 L 15/28	2 3 0 K		

審査請求 有 請求項の数 4 O L (全 13 頁)

(21) 出願番号	特願2008-240988 (P2008-240988)	(71) 出願人	000004260
(22) 出願日	平成20年9月19日 (2008.9.19)		株式会社デンソー
(62) 分割の表示	特願2002-368441 (P2002-368441) の分割	(74) 代理人	110000578 名古屋国際特許業務法人
原出願日	平成14年12月19日 (2002.12.19)	(72) 発明者	大野 宏 愛知県刈谷市昭和町1丁目1番地 株式会 社デンソー内
		Fターム(参考)	5D015 EE05

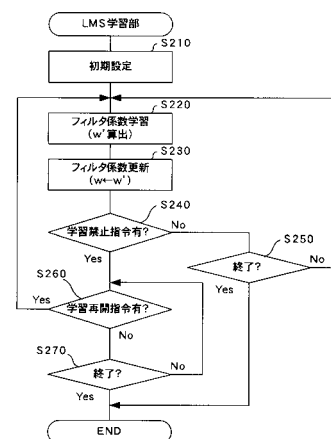
(54) 【発明の名称】 音声認識装置

(57) 【要約】

【課題】 音声信号に含まれる雑音成分を適切に除去可能にして、音声認識の精度を向上させること。

【解決手段】 音声認識装置は、マイクロフォンから入力される音声信号から雑音成分を除去するための適応フィルタを備える。LMS学習部は、繰り返し、LMS法に基づきフィルタ係数を学習し(S220)、その結果得られたフィルタ係数を適応フィルタに設定する(S230)。また、音声認識の開始と共に制御部から学習禁止指令が入力されると、フィルタ係数の学習を停止する。その後、音声認識に必要な音声信号が音声認識部に与えられると制御部から入力される学習再開指令(S260)に従って、再び、フィルタ係数を学習し更新する。

【選択図】 図3



【特許請求の範囲】

【請求項 1】

雑音源から入力される参照信号を、予め設定されたフィルタ係数に従い濾波することで、雑音を除去するための雑音除去信号を生成する雑音除去信号生成手段と、

該雑音除去信号生成手段が生成した前記雑音除去信号を用いて、マイクロフォンから入力される音声信号に含まれる雑音成分を除去し、その雑音除去後の音声信号を出力する雑音除去手段と、

該雑音除去手段が出力する音声信号に基づいて、前記雑音除去信号生成手段に設定すべきフィルタ係数を学習し、学習した結果得たフィルタ係数を、前記雑音除去信号生成手段に設定する係数更新手段と、

外部から動作開始指令が入力されると、所定期間、前記雑音除去手段から出力される音声信号を取得して、その音声信号に基づき、前記マイクロフォンに入力された音声を確認する音声認識手段と、

該音声認識手段の非動作時に、第一の学習速度で、前記係数更新手段にフィルタ係数を学習させ、前記音声認識手段が前記雑音除去手段から出力される音声信号を取得している期間には、前記第一の学習速度より低い第二の学習速度で、前記係数更新手段にフィルタ係数を学習させる学習速度切替手段と、

を備えることを特徴とする音声認識装置。

【請求項 2】

前記学習速度切替手段は、前記動作開始指令が前記音声認識手段に入力されると同時に、前記音声認識手段が前記音声信号の取得を終了するまでの期間、前記第二の学習速度で、前記係数更新手段にフィルタ係数を学習させることを特徴とする請求項 1 に記載の音声認識装置。

【請求項 3】

前記係数更新手段は、前記雑音除去手段が出力する音声信号に基づき、LMS法を用いて、前記雑音除去信号生成手段に設定すべきフィルタ係数を学習することを特徴とする請求項 1 又は請求項 2 に記載の音声認識装置。

【請求項 4】

雑音源がオーディオ機器である請求項 1～請求項 3 のいずれかに記載の音声認識装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、マイクロフォンから得た音声信号に基づいて、利用者がマイクロフォンに入力した音声を認識する音声認識装置に関する。

【背景技術】

【0002】

従来より、利用者から発せられた音声をマイクロフォンで集音し、これを予め認識語として記憶された音声のパターンと比較し、一致度の高い認識語を利用者が発声した語彙であると認識する音声認識装置が知られている。この種の音声認識装置は、例えばカーナビゲーション装置などに組み込まれている。

【0003】

このような音声認識装置の音声認識率（音声認識の正解率）は、マイクロフォンから入力される音声信号に含まれる雑音量によって左右されることがよく知られているが、特に自動車などの車両内においては、車載オーディオ機器の動作時に、再生される音楽等が雑音として利用者の音声と共にマイクロフォンで集音されてしまう問題がある。

【0004】

この問題に対し、従来では、車載オーディオ機器と、音声認識装置とを連動させ、音声認識処理の際に、車載オーディオ機器にて再生される音楽等のボリュームを調節する（例えば車載オーディオ機器をミュートに設定する）ことで、再生される音楽等がマイクロフォンに入力されないようにし、一定程度以上の音声認識率を確保するようにしていた。尚、

10

20

30

40

50

このような先行技術は周知公用のため、関連文献を非開示とする。

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、従来の音声認識装置では、車載オーディオ機器にて再生される音楽等のボリュームを調節するので、一時的に音楽等が利用者に聞こえなくなってしまう、その事が原因で利用者に不満が及ぶ可能性があった。

【0006】

そこで、本発明者らは、雑音源（車載オーディオ機器）から得られる参照信号に基づき、マイクロフォンから得られる音声信号に含まれる雑音成分を学習し、その音声信号から学習した雑音成分を除去する雑音除去部を音声認識装置に設けることにした。

10

【0007】

しかしながら、最小二乗平均（Least Mean Square：LMS）法などの周知の学習法では、雑音除去後の信号が小さくなる方向に学習を繰り返すため、利用者がマイクロフォンに音声を入力している間に雑音成分の学習が繰り返されると、その利用者による発話の影響を受けて雑音除去部が誤学習をし、音声信号に含まれる雑音成分が適切に除去できなくなってしまう問題があった。したがって、このような雑音除去部を音声認識装置に導入しても、音声認識の精度向上には限界があった。

【0008】

本発明はこうした問題に鑑みなされたものであり、音声認識対象の音声信号に含まれる雑音成分を適切に除去して高精度に音声認識可能な音声認識装置を提供することを目的とする。

20

【課題を解決するための手段】

【0009】

かかる目的を達成するためになされた請求項1に記載の音声認識装置によれば、予め設定されたフィルタ係数に従い雑音源から入力される参照信号を濾波することで雑音除去信号生成手段が生成した雑音除去信号を用いて、雑音除去手段が、マイクロフォンから入力される音声信号に含まれる雑音成分を除去し、雑音除去後の音声信号を出力する。また、この音声認識装置は、係数更新手段を備えており、雑音除去手段から出力される音声信号に基づき、係数更新手段にて、雑音除去信号生成手段に設定すべきフィルタ係数を学習し、その結果得たフィルタ係数を、雑音除去信号生成手段に対して設定する。

30

【0010】

一方、音声認識手段は、外部（例えば、PTTスイッチ等の操作スイッチ）から動作開始指令が入力されると、所定期間、雑音除去手段から出力される音声信号を取得し、その音声信号に基づき、マイクロフォンに入力された音声を認識する。

【0011】

また、この音声認識装置では、学習速度切替手段が、音声認識手段の非動作時に、第一の学習速度で、係数更新手段にフィルタ係数を学習させ、音声認識手段が雑音除去手段から出力される音声信号を取得している間には、第一の学習速度より低い第二の学習速度で、係数更新手段にフィルタ係数を学習させる。

40

【0012】

マイクロフォンに入力される利用者の音声は、定常及び準定常的な音とは異なり、突発的に発生する非定常的な音声であることから、音声認識手段が雑音除去手段から出力される音声信号を取得している期間、フィルタ係数の学習速度を遅くすれば、フィルタ係数の学習時に利用者の音声を与える影響を抑えることができ、係数更新手段によるフィルタ係数の誤学習を抑制することができる。

【0013】

即ち、請求項1に記載の音声認識装置によれば、従来と比較して、適切にフィルタ係数の学習を係数更新手段に実行させることができ、雑音除去の精度を向上させることができる。したがって、本発明によれば、音声認識装置における音声認識の精度を向上させるこ

50

とができる。

【0014】

尚、学習速度切替手段は、少なくとも音声認識手段が雑音除去手段から音声信号の取得を開始した時点から終了する時点まで、第二の学習速度で、係数更新手段にフィルタ係数の学習を行わせる構成にされていればよく、例えば、音声信号の取得完了後、音声認識手段による音声の認識が完了し音声認識手段の動作が停止するまで、第二の学習速度で、フィルタ係数の学習を係数更新手段に実行させても構わない。

【0015】

また、学習速度切替手段は、少なくとも音声認識手段の非動作時に第一の学習速度で係数更新手段にフィルタ係数を学習させる構成にされていればよく、音声認識手段の非動作時に加え、音声認識手段の動作時であって音声認識手段が雑音除去手段から出力される音声信号を取得していない期間に、第一の学習速度で係数更新手段にフィルタ係数を学習させる構成にされていてもよい。即ち、学習速度切替手段は、音声認識手段が音声の認識を行っているか否かにかかわらず、音声認識手段が音声信号を取得した直後から第一の学習速度で係数更新手段にフィルタ係数の学習を行わせる構成にされていてもよい。

【0016】

その他、本発明の音声認識装置における音声認識手段は、外部から動作開始指令が入力された後、利用者による発声となされた発声期間に限定して、雑音除去手段から出力される音声信号を取得する構成にされていると良い。音声認識手段をこのような構成とすれば、利用者による発話内容が含まれない雑音区間の音声信号を、音声認識の際に用いずに済み、音声認識の精度が向上する。

【0017】

また、このように利用者による発声期間の音声信号を選択的に音声認識手段に取得させる場合には、雑音除去手段から出力される音声信号に基づいて、利用者による発声となされた発声期間を判別し、雑音除去手段から出力される音声信号の内、その発声期間に該当する音声信号のみを選択的に、音声認識手段に取得させる取得制御手段を装置内に設ければ良い。

【0018】

ところで、動作開始指令と同時に音声認識手段が雑音除去手段から音声信号を取得しない場合には、音声認識手段が雑音除去手段から出力される音声信号を取得している期間のみ第二の学習速度でフィルタ係数が学習されるようにすると、装置構成が煩雑になる可能性がある。

【0019】

したがって、上述の音声認識装置においては、動作開始指令が音声認識手段に入力されると同時に、第二の学習速度で、係数更新手段にフィルタ係数を学習させ、音声認識手段が音声信号の取得を終了するまでの期間は、第二の学習速度によるフィルタ係数の学習を係数更新手段に継続させるように、学習速度切替手段を構成するとよい。

【0020】

このような構成にされた請求項2に記載の音声認識装置によれば、音声認識手段に外部から動作開始指令が入力されたか否かを監視する程度で、音声認識手段が雑音除去手段から音声信号を取得する際には、係数更新手段に第二の学習速度でフィルタ係数を学習させることができる。つまり、この音声認識装置によれば、簡単な装置構成（制御）で、係数更新手段の学習速度を適切に切り替えることができる。

【0021】

その他、上述した発明は、請求項3に記載のように、係数更新手段が、LMS法を用いて、雑音除去信号生成手段に設定すべきフィルタ係数を学習する音声認識装置に適用される良い。

【0022】

LMS法を用いる場合には、マイクロフォンに入力される音声に、雑音源以外の音源から発生した音声（即ち、利用者の音声）が含まれると、フィルタ係数の誤学習を引き起こ

10

20

30

40

50

しやすい。したがって、請求項 3 に記載のように、LMS 法を用いて学習を行う音声認識装置に、本発明（請求項 1 又は請求項 2）を適用すれば、音声認識の精度を効果的に向上させることができる。

【0023】

また、請求項 1～請求項 3 に記載の発明は、請求項 4 に記載のように、雑音源がオーディオ機器である音声認識装置に適用されるとよい。

請求項 4 に記載の音声認識装置によれば、オーディオ機器の動作によりスピーカから再生される音楽等のボリュームを調節しなくても、高精度に音声認識を行うことが可能であるので、便利である。

【発明を実施するための最良の形態】

【0024】

以下に本発明の実施例について、図面とともに説明する。尚、図 1 は、音声認識装置 1 の概略構成を表すブロック図である。

図 1 に示す本実施例の音声認識装置 1 は、カーナビゲーション装置 3 に接続されており、マイクロフォン 5 に入力された利用者の音声認識し、その音声に従う操作信号をカーナビゲーション装置 3 に入力することで、利用者の音声に従った操作をカーナビゲーション装置 3 に対して施す。

【0025】

この音声認識装置 1 は、主に、マイクロフォン 5 及び車載オーディオ機器 7 にアナログ→デジタル変換器（ADC）11、13 を介して接続されたオーディオキャンセラ部 20 と、音声抽出部 31 と、音声認識部 33 と、PTT（Push to Talk）スイッチ 35 と、制御部 37 と、音声合成部 39 と、から構成されている。

【0026】

オーディオキャンセラ部 20 は、主に、適応フィルタ 21 と、減算部 23 と、LMS 学習部 25 と、から構成されており、マイクロフォン 5 から ADC 11 を介して入力される音声信号 $y(t)$ を減算部 23 に入力すると共に、車載オーディオ機器 7 からスピーカ 9 に入力されるオーディオ信号 $x(t)$ を ADC 13 から取得し、そのオーディオ信号 $x(t)$ を適応フィルタ 21 に入力する。

【0027】

適応フィルタ 21 は、フィルタ係数 w を記憶する図示しないレジスタ等を備えている。

$$w = (w[0], w[1], \dots, w[J])^T \quad \cdots \text{式 (1)}$$

尚、上付き記号 T は、転置行列を意味する。また、パラメータ $J+1$ は、タップ長を表す。

【0028】

この適応フィルタ 21 は、LMS 学習部 25 の動作（詳細後述）により予めレジスタに設定されたフィルタ係数 w と、雑音源としての車載オーディオ機器 7 から参照信号として得た上記オーディオ信号 $x(t)$ とを、次式に代入し演算することで、オーディオ信号 $x(t)$ をフィルタ係数 w に従い濾波し、音声信号 $y(t)$ から雑音成分を除去するための雑音除去信号 $c(t)$ を生成する。そして、雑音除去信号 $c(t)$ を減算部 23 に入力する。

【0029】

$$c(t) = x^T \cdot w \quad \cdots \text{式 (2)}$$

但し、パラメータ x は、下式で表されるオーディオ信号 $x(t)$ の時系列ベクトルである。また、ここでいうパラメータ t は、サンプリング周期を単位とする時間パラメータである。

【0030】

$$x = (x(t), x(t-1), x(t-2), \dots, x(t-J))^T \quad \cdots \text{式 (3)}$$

一方、減算部 23 は、ADC 11 を介してマイクロフォン 5 より入力される音声信号 $y(t)$ から雑音除去信号 $c(t)$ を減算することで、音声信号 $y(t)$ に含まれる雑音成

10

20

30

40

50

分（即ち、車載オーディオ機器 7 の動作によりスピーカ 9 から再生される音声成分）を除去し、雑音除去後の音声信号 $z(t)$ を得る。

【0031】

$$z(t) = y(t) - c(t) \quad \cdots \text{式(4)}$$

また、減算部 23 は、減算した結果得られた雑音除去後の音声信号 $z(t)$ を、音声抽出部 31 に入力する。

【0032】

音声抽出部 31 は、制御部 37 からの動作開始指令を受けて動作を開始する構成にされており、動作を開始すると、オーディオキャンセラ部 20 から入力された雑音除去後の音声信号 $z(t)$ が、音声区間（即ち、利用者による発声になされた発声期間）の信号であるか、それとも、利用者の音声が含まず音声区間に属さない雑音区間の信号であるのかを判定し、音声区間の信号であると判定した場合には、その音声信号 $z(t)$ を音声認識部 33 に入力する。そして音声区間が終了すると、動作を停止する。

【0033】

尚、判定方法としては、例えば入力信号の短時間パワーを一定時間毎に抽出していき、所定の閾値以上の短時間パワーが一定以上継続したか否かによって音声区間であるか雑音区間であるかを判定する手法がよく採用されている。

【0034】

一方、音声認識部 33 は、制御部 37 から入力される動作開始指令に従い動作を開始し、音声抽出部 31 から出力される音声信号 $z(t)$ を取得することにより、音声抽出部 31 を介して、減算部 23 から音声区間の信号 $z(t)$ を選択的に取得する。また、音声認識部 33 は、音声信号 $z(t)$ の取得後に、その音声信号 $z(t)$ を音響分析し、音声信号 $z(t)$ から特徴量（例えばケプストラム）を抽出することで、特徴量の時系列データを得る。

【0035】

その後、音声認識部 33 は、特徴量の時系列データを、周知の技法を用いて、自身が備える図示しない音声辞書に登録された音声パターンと比較し、一致度の高い音声パターンに対応する語彙を、利用者が発声した語彙であると認識して、その認識結果を制御部 37 に入力し、この後動作を停止する。

【0036】

制御部 37 は、P T T スイッチ 35 が押されたタイミングや戻されたタイミングを監視する構成にされており、P T T スイッチ 35 が押され、P T T スイッチ 35 から動作開始指令信号が入力されたと判断すると（S 100 で Y e s）、オーディオキャンセラ部 20 の L M S 学習部 25 に対して学習禁止指令を入力し（S 110）、その後に音声認識部 33 及び音声抽出部 31 に動作開始指令を入力することで、音声認識部 33 及び音声抽出部 31 を作動させて、音声認識を開始する（S 120）。尚、図 2 は、制御部 37 の処理動作を表すフローチャートである。

【0037】

その後、制御部 37 は、音声区間が終了し音声認識部 33 による音声信号の取得が完了したか否かを、音声抽出部 31 の動作状態に基づき判断し（S 130）、音声区間が終了したと判断すると（S 130 で Y e s）、L M S 学習部 25 に学習再開指令を入力する（S 140）と共に、音声認識部 33 から認識結果を取得する（S 150）。そして認識結果が正しいか否かを利用者に問い合わせるために、認識結果をトークバックする（S 160）。

【0038】

即ち、制御部 37 は、音声合成部 39 を制御して、音声合成部 39 に、認識結果に従う音声信号を生成させ、その音声信号をスピーカ 9 に入力させる。尚、音声合成部 39 は、図示しない波形データベース内に格納されている音声波形を用い、制御部 37 からの音声の出力指示に基づく音声信号を合成してスピーカ 9 に出力する。従って、S 160 においては、認識結果が音声で利用者に通知される。

10

20

30

40

50

【0039】

この後、制御部37は、利用者の操作によりPTTスイッチ35等の操作スイッチから認識結果が正しいことを表す認識結果確定信号が入力されたか否か判断し（S170）、認識結果確定信号が入力されたと判断すると（S170でYes）、確定後処理を実行する（S180）。一方、認識結果確定信号が入力されなかったと判断すると（S170でNo）、確定後処理を実行せずに、当該処理を終了する。

【0040】

尚、S180で行われる確定後処理において、制御部37は、認識結果に従う操作信号をカーナビゲーション装置3に入力する。このような確定後処理は、周知の技術を用いたものであるので、詳細な説明を省略する。

10

【0041】

次に、オーディオキャンセラ部20のLMS学習部25の処理動作について図3を用いて説明する。図3は、音声認識装置1に電源が投入されると同時に、LMS学習部25が実行する学習処理を表すフローチャートである。

【0042】

LMS学習部25は、学習処理の実行を開始すると、まず最初に、適応フィルタ21に対して初期設定を施す（S210）。即ち、LMS学習部25は、予め定められた所定のフィルタ係数（初期値）を適応フィルタ21に設定する。

【0043】

その後、LMS学習部25は、減算部23から出力される音声信号 $z(t)$ を用い、LMS法に基づく次式に従い係数 w' を算出することで、次に適応フィルタ21に設定すべきフィルタ係数 w' を学習する（S220）。

20

【0044】

【数1】

$$w' = w \times \alpha + \mu \times \frac{z(t) \cdot x}{\|x\|^2 + \beta} \quad \cdots \text{式(5)}$$

ここで、代入する係数 w は、既に適応フィルタ21に設定したフィルタ係数 w の値である。また、 α は、係数 w' が発散するのを防止するための忘却係数であり、 β は、除数がゼロになるのを防止するための正の定数である。その他 μ は、ステップサイズパラメータと呼ばれるものであり、フィルタ係数の学習速度に対応するパラメータである。

30

【0045】

S220におけるフィルタ係数 w' の計算が完了すると、LMS学習部25は、S220で算出したフィルタ係数 w' を、新たなフィルタ係数 w として、適応フィルタ21に設定する（S230）。

【0046】

この後、LMS学習部25は、学習禁止指令が制御部37より入力されているか否か判断し（S240）、入力されていなければ（S240でNo）、当該装置の電源オフやエラー等により学習処理の終了指令が制御部37から入力されているか否か判断する（S250）。そして、終了指令が入力されていれば（S250でYes）、当該処理を終了し、終了指令が入力されていなければ（S250でNo）、処理をS220に戻して、フィルタ係数 w' を学習し、その後フィルタ係数を更新する（S230）。

40

【0047】

また、S240において、学習禁止指令が制御部37より入力されていると判断すると（S240でYes）、LMS学習部25は、処理をS260に移して、学習再開指令が制御部37から入力されているか否か判断する。そして学習再開指令が入力されていなければ（S260でNo）、続くS270にて終了指令が入力されているか否か判断し、終了指令が入力されていれば（S270でYes）、当該処理を終了し、終了指令が入力されていなければ（S270でNo）、処理をS260に戻して、学習再開指令が制御部37より入力されるまで待機する。

50

【0048】

そして、学習再開指令が入力されたと判断すると（S260でYes）、処理をS220に戻して、フィルタ係数 w' を学習し、その結果得られたフィルタ係数 w' を、新たなフィルタ係数 w として、適応フィルタ21に設定する（S230）。

【0049】

LMS学習部25は、このような動作を繰り返すことによって、図4に示すように、PTTスイッチ35が押下（オン）されてから音声区間が終了するまでの間、フィルタ係数の学習動作を停止する。また、音声区間が終了して学習再開指令が入力されると、再び、次の学習禁止指令が入力されるまで、フィルタ係数の学習を継続する。尚、図4は、LMS学習部25の動作切替タイミングを表すタイムチャートである。

10

【0050】

以上、本実施例の音声認識装置1について説明したが、この音声認識装置1では、制御部37の動作により、音声認識部33が音声抽出部31を介してオーディオキャンセラ部20から音声信号を取得している間、LMS学習部25によるフィルタ係数の学習が禁止されるので、音声認識のために利用者が発した音声がマイクロフォン5に入力される際に、フィルタ係数 w の学習更新が行われるのを防止することができる。

【0051】

したがって、この音声認識装置1によれば、音声認識部33による音声信号取得の際に、マイクロフォン5に入力される利用者の音声の影響によって、フィルタ係数が不適切に学習更新されるのを防止することができ、音声認識の対象となる音声信号から精度よく雑音成分を取り除くことができる。結果、本実施例によれば、音声認識装置1における音声認識の精度を高めることができ、高い音声認識率を実現することが可能である。

20

【0052】

その他、本実施例では、オーディオキャンセラ部20から出力される音声信号 $z(t)$ に基づき、音声抽出部31にて、利用者による発声となされた発声期間を判別し、オーディオキャンセラ部20から出力される音声信号の内、その発声期間に該当する音声信号のみを選択的に音声認識部33に入力するようにしているので、利用者による発話内容が含まれない雑音区間の音声信号を、音声認識部33に入力せずに済み、雑音に影響されず、音声認識部33に正確な音声認識を行わせることができる。また、本実施例では、音声抽出部31が、自動で発声期間を判別するので、利用者に発声期間に関する情報を操作スイッチから入力させなくて済み便利である。

30

【0053】

また、本実施例では、PTTスイッチ35から動作開始指令が入力されると同時に、その時点から音声抽出部31が音声区間の検出を終了して音声認識部33が音声信号の取得を終了するまでの期間、LMS学習部25によるフィルタ係数の学習を禁止するように制御部37を構成しているので、簡単な制御で、利用者の発声期間には、フィルタ係数の学習を停止することができる。

【0054】

さて、上記実施例ではフィルタ係数の学習を禁止することにより音声認識装置1の高性能化を実現したが、利用者の発声期間中に、フィルタ係数の学習速度を遅くすることで、従来と比較して高精度に音声認識を行えるようにすることも可能である。

40

【0055】

次には、このような構成にされた変形例の音声認識装置について説明することにする。尚、変形例の音声認識装置は、制御部37及びLMS学習部25の一部処理動作が異なる程度の構成であり、その他の装置内各部の構成は上述の音声認識装置1と同一である。したがって、以下では、上述の音声認識装置1と同一構成の各部の説明を省略することにし、図5及び図6を用いて、制御部37及びLMS学習部25の動作を説明する程度に留める。

【0056】

図5は、変形例の音声認識装置における制御部37の処理動作を表すフローチャートで

50

ある。図 5 に示すように、制御部 37 は、P T T スイッチ 35 から動作開始指令信号が入力されたと判断すると (S 300 で Y e s)、フィルタ係数の学習速度を遅くするための低速学習指令をオーディオキャンセラ部 20 の L M S 学習部 25 に入力し (S 310)、その後音声認識部 33 及び音声抽出部 31 を作動させて音声認識を開始する (S 320)。

【0057】

その後、制御部 37 は、音声区間が終了し音声抽出部 31 から音声認識部 33 への音声信号入力完了したか否かを、音声抽出部 31 の動作状態に基づき判断し (S 330)、音声区間が終了したと判断すると (S 330 で Y e s)、フィルタ係数の学習速度を通常の学習速度に変更するための通常学習指令を L M S 学習部 25 に入力する (S 340)。

10

また、同時に、音声認識部 33 から認識結果を取得する (S 350)。そして認識結果が正しいか否かを利用者に問い合わせるために、認識結果をトークバックする (S 360)。

【0058】

この後、制御部 37 は、利用者の操作により P T T スイッチ 35 等の操作スイッチから認識結果が正しいことを表す認識結果確定信号が入力されたか否か判断し (S 370)、認識結果確定信号が入力されたと判断すると (S 370 で Y e s)、確定後処理を実行する (S 380)。一方、認識結果確定信号が入力されなかったと判断すると (S 370 で N o)、確定後処理を実行せずに、当該処理を終了する。

【0059】

20

次に、変形例の音声認識装置における L M S 学習部 25 の処理動作について図 6 を用いて説明する。図 6 は、音声認識装置に電源が投入されると同時に、変形例の L M S 学習部 25 が実行する学習処理を表すフローチャートである。

【0060】

L M S 学習部 25 は、学習処理を開始すると S 410 で、初期設定として、予め定められた所定のフィルタ係数 (初期値) を適応フィルタ 21 に対して設定すると共に、フィルタ係数 w' 算出の際に用いる式 (5) のパラメータ μ を、初期値 μ_H に設定する ($\mu = \mu_H$)。

【0061】

この後、L M S 学習部 25 は、減算部 23 から出力される音声信号 $z(t)$ を用いて、L M S 法に基づく式 (5) に従い係数 w' を算出する (S 420)。この動作により L M S 学習部 25 は、次に適応フィルタ 21 に設定すべきフィルタ係数 w' を学習し、S 430 にて、フィルタ係数 w' を、新たなフィルタ係数 w として、適応フィルタ 21 に設定する。

30

【0062】

続いて、L M S 学習部 25 は、低速学習指令が制御部 37 より入力されているか否か判断し (S 440)、入力されていると判断すると (S 440 で Y e s)、S 450 にて、学習速度を表すパラメータ μ に、予め定められた値 μ_L を設定する ($\mu = \mu_L$)。尚、値 μ_L 及び値 μ_H には、不等式 $\mu_L < \mu_H$ の関係が成立する。

【0063】

40

式 (5) を見れば理解できるように、パラメータ μ の値を小さくすると、フィルタ係数 w' の変化量を、小さくすることができる。つまり、パラメータ μ を小さくすることで、フィルタ係数 w' が収束するまでの時間を長期化することができ、学習速度を抑えることができる。L M S 学習部 25 は、このようにパラメータ μ を通常より小さい値 μ_L に設定することで、フィルタ係数の学習速度を低くしているのである。

【0064】

この後、制御部 37 は、処理を S 420 に移して、 $\mu = \mu_L$ である式 (5) に従い、フィルタ係数 w' を算出し、その後フィルタ係数 w を更新する (S 430)。

一方、制御部 37 は、S 440 にて、低速学習指令が入力されていないと判断すると (S 440 で N o)、S 460 にて、制御部 37 から通常学習指令が入力されているか否か

50

判断する。

【0065】

ここで、通常学習指令が入力されていると判断すると（S460でYes）、制御部37は、S470にて、学習速度を表すパラメータ μ を、 μ_H に変更する（ $\mu = \mu_H$ ）。そして、再び処理をS420に移し、 $\mu = \mu_H$ である式（5）に従い、フィルタ係数 w' を算出し、その後フィルタ係数 w を更新する（S430）。

【0066】

また、制御部37は、S440及びS460でNoと判断すると、S480にて、当該学習処理の終了指令が制御部37から入力されているか否か判断する。そして、終了指令が入力されていないと判断すると（S480でNo）、処理をS420に戻して、フィルタ係数 w' を学習し、その後フィルタ係数を更新する（S430）。一方、終了指令が入力されていると判断すると（S480でYes）、当該学習処理を終了する。

10

【0067】

LMS学習部25は、このような処理を実行することによって、図7に示すように、PTTスイッチ35が押下（オン）されてから、音声区間が終了するまでの間は、フィルタ係数の学習速度を低くする。また、音声区間が終了して通常学習指令が入力されると、再び、次の低速学習指令が入力されるまで、通常の学習速度でフィルタ係数の学習を行う。尚、図7は、学習速度の切替タイミングを表すタイムチャートである。

【0068】

以上、変形例について説明したが、変形例の音声認識装置では、音声認識部33及び音声抽出部31の作動と共にLMS学習部25に低速学習指令を入力することで、音声認識部33が音声抽出部31を介してオーディオキャンセラ部20から音声信号を取得している間、通常より低学習速度で、LMS学習部25に、フィルタ係数を学習させているので、その期間においてフィルタ係数の学習動作に及ぶ利用者の音声の影響を抑えることができ、LMS学習部25におけるフィルタ係数の誤学習を抑制することができる。

20

【0069】

この結果、変形例の音声認識装置によれば、LMS学習部25に適切にフィルタ係数の学習を行わせることができ、オーディオキャンセラ部20における雑音除去の精度を向上させることができる。したがって、変形例によれば、高精度に音声認識可能な音声認識装置を提供することができる。

30

【0070】

また、変形例では、制御部37が、音声認識部33の非動作時に加え、音声認識部33が音声信号を取得した直後（即ち音声区間が終了した直後）から通常の学習速度でLMS学習部25にフィルタ係数の学習を行わせる構成にされているので、連続してPTTスイッチ35から動作開始指令信号が入力され音声認識部33が動作する場合にも、オーディオキャンセラ部20にて適切な雑音除去が可能である。

【0071】

その他、変形例においても、音声抽出部31が、利用者による発声になされた発声期間に該当する音声信号のみを選択的に音声認識部33に入力するので、利用者による発話内容が含まれない雑音区間の音声信号を、音声認識部33に入力せずに済み、雑音に影響されず、音声認識部33にて正確な音声認識を行うことができる。

40

【0072】

また、変形例の音声認識装置においては、PTTスイッチ35から動作開始指令信号が入力されると同時に、低学習速度でLMS学習部25にフィルタ係数を学習させる手法を採用しているので、簡単な制御で確実に、フィルタ係数の誤学習を抑制することができる。

【0073】

その他、上記実施例の音声認識装置によれば、車載オーディオ機器7の動作によりスピーカ9から再生される音楽等のボリュームを調節しなくても、高精度に音声認識を行うことができるので、ボリューム調整などによって利用者に不満が及ぶといった従来問題を解

50

消することができる。

【0074】

以上、本発明の実施例について説明したが、本発明の雑音除去信号生成手段は、本実施例の適応フィルタ21に相当し、本発明の雑音除去手段は、減算部23に相当する。また、係数更新手段は、LMS学習部25に相当し、音声認識手段は、音声区間における音声信号 $z(t)$ を取得して音声認識を行う音声認識部33に相当する。

【0075】

その他、学習速度切替手段は、制御部37が図5に示す処理に従うタイミングで低速学習指令及び通常学習指令をLMS学習部25に入力する動作にて実現されている。尚、学習速度切替手段が、第二の学習速度で係数更新手段にフィルタ係数を学習させる動作は、本実施例において、LMS学習部25に、第二の学習速度に対応するパラメータ $\mu = \mu_L$ でフィルタ係数 w' の演算を行わせる動作にて実現されている。また、学習速度切替手段が、第一の学習速度で係数更新手段にフィルタ係数を学習させる動作は、本実施例において、LMS学習部25に、第一の学習速度に対応するパラメータ $\mu = \mu_H$ でフィルタ係数 w' の演算を行わせる動作にて実現されている。

【0076】

また、本発明の音声認識装置は、上記実施例に限定されるものではなく、種々の態様を採ることができる。

例えば、制御部37は、少なくとも音声認識部33による音声信号の取得期間において、LMS学習部25の動作を禁止する、若しくは、LMS学習部25におけるフィルタ係数の学習速度を低速度化する構成にされていればよく、音声区間の終了後音声認識部33における音声認識が完了し、音声認識の結果が得られるまで、LMS学習部25によるフィルタ係数の学習を禁止してもよいし、低学習速度でLMS学習部25を動作させてもよい。

【0077】

また、上記実施例では、フィルタ係数の学習方法としてLMS法、具体的にはNormalized LMS (NLMS) アルゴリズムが採用された音声認識装置に、本発明を適用した例を示したが、その他の学習方法でフィルタ係数を学習する音声認識装置に本発明を適用しても構わない。尚、本発明を適用可能な学習方法としては、上述した適応アルゴリズム以外に、例えば、複素LMSアルゴリズム、Fast LMS (FLMS) アルゴリズム、射影アルゴリズム、RLS (Recursive Least Square) アルゴリズム、SHARF (Simple Hyperstable Adaptive Recursive Filter) アルゴリズム、DCT (Discrete Cosine Transform) を用いた適応フィルタ、SAN (Single Frequency Adaptive Notch) フィルタ、ニューラルネットワーク、遺伝的アルゴリズム等が挙げられる。

【図面の簡単な説明】

【0078】

【図1】本実施例の音声認識装置1の概略構成を表すブロック図である。

【図2】制御部37の処理動作を表すフローチャートである。

【図3】LMS学習部25が実行する学習処理を表すフローチャートである。

【図4】LMS学習部25の動作切替タイミングを表すタイムチャートである。

【図5】変形例の制御部37における処理動作を表すフローチャートである。

【図6】変形例のLMS学習部25が実行する学習処理を表すフローチャートである。

【図7】学習速度の切替タイミングを表すタイムチャートである。

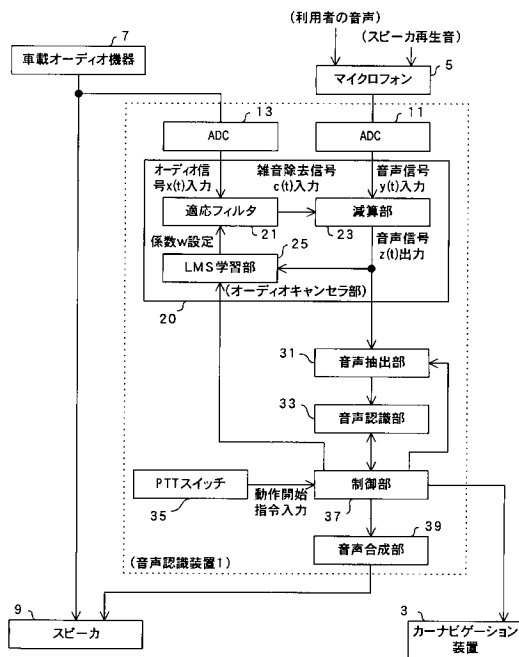
【符号の説明】

【0079】

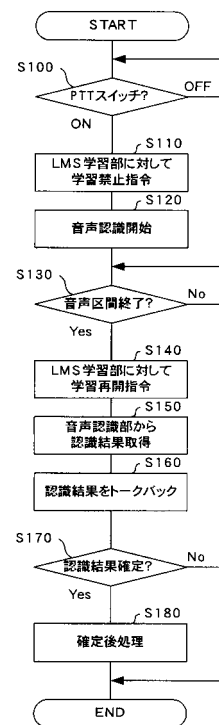
1…音声認識装置、3…カーナビゲーション装置、5…マイクロフォン、7…車載オーディオ機器、9…スピーカ、11, 13…ADC、20…オーディオキャンセラ部、21…適応フィルタ、23…減算部、25…LMS学習部、31…音声抽出部、33…音声認識

部、35…PTTスイッチ、37…制御部、39…音声合成部

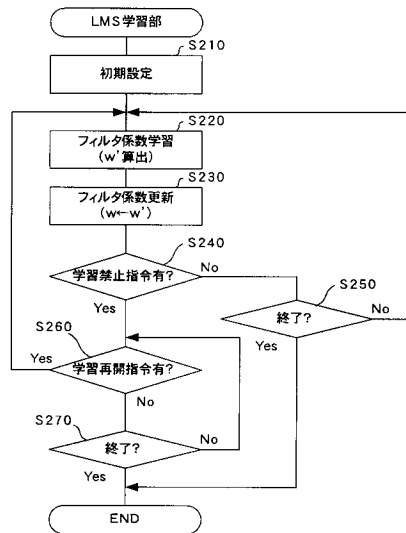
【図1】



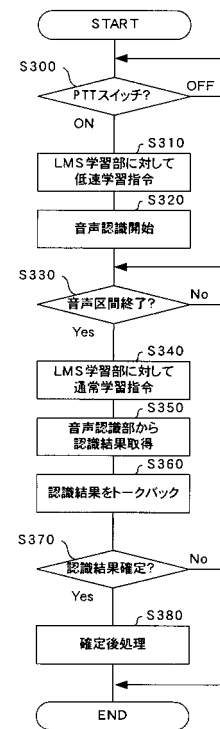
【図2】



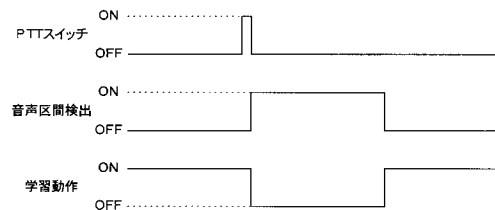
【図 3】



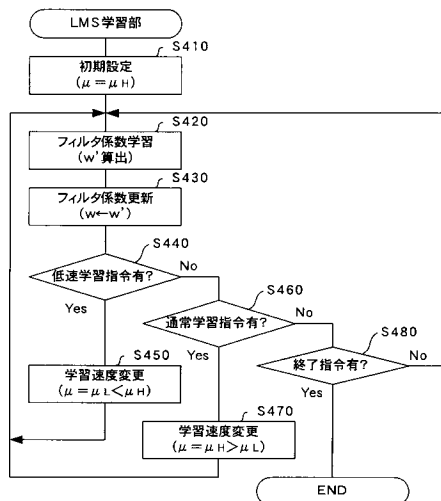
【図 5】



【図 4】



【図 6】



【図 7】

