

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

**特許第3774580号  
(P3774580)**

(45) 発行日 平成18年5月17日(2006.5.17)

(24) 登録日 平成18年2月24日(2006.2.24)

(51) Int. Cl.

**G 1 O L 15/20 (2006.01)**

F I

G 1 O L 15/20 3 7 O E

請求項の数 2 (全 9 頁)

(21) 出願番号	特願平10-338454	(73) 特許権者	000101732
(22) 出願日	平成10年11月12日(1998.11.12)		アルパイン株式会社
(65) 公開番号	特開2000-148200(P2000-148200A)		東京都品川区西五反田1丁目1番8号
(43) 公開日	平成12年5月26日(2000.5.26)	(74) 代理人	100103171
審査請求日	平成15年3月28日(2003.3.28)		弁理士 雨貝 正彦
前置審査		(72) 発明者	木内 真吾
			東京都品川区西五反田1丁目1番8号 ア
			ルパイン株式会社内
		(72) 発明者	斉藤 望
			東京都品川区西五反田1丁目1番8号 ア
			ルパイン株式会社内
		(72) 発明者	中田 孝一
			東京都品川区西五反田1丁目1番8号 ア
			ルパイン株式会社内

最終頁に続く

(54) 【発明の名称】 音声入力装置

(57) 【特許請求の範囲】

【請求項1】

集音手段から出力される音声信号に含まれる周辺雑音成分を適応フィルタを用いて除去する周辺雑音除去手段と、前記周辺雑音除去手段によって前記周辺雑音が除去された後の音声信号からスピーカの出力成分を除去するスピーカ音声除去手段と、前記集音手段に対する音声入力のタイミングを指示する入力タイミング指示手段とを備える音声入力装置において、

前記集音手段は、第1および第2のマイクロホンを有し、

前記周辺雑音除去手段は、前記第1のマイクロホンから出力されるノイズ信号を前記第2のマイクロホンから出力されるノイズ信号を用いて模擬する前記適応フィルタと、これら2つのノイズ信号の差分を計算する第1の演算手段と、前記適応フィルタのフィルタ係数の更新処理を行うとともに、前記入力タイミング指示手段によって前記音声入力のタイミングが指示されたときに、前記適応フィルタのフィルタ係数の更新動作を停止するフィルタ処理手段と、前記第1のマイクロホンと前記第1の演算手段との間に挿入されて伝達特性 $Z^{-m}$ で表される遅延素子とを有し、

前記スピーカ音声除去手段は、前記スピーカおよび前記周辺雑音除去手段を介した伝達系の特性を模擬するフィルタと、前記周辺雑音除去手段を介した信号と前記フィルタを介した信号との差分を計算する第2の演算手段と、前記入力タイミング指示手段によって前記音声入力のタイミングが指示されたときに、前記スピーカから前記第1および第2のマイクロホンのそれぞれまでの伝達特性 $C S 1$ 、 $C S 2$ 、更新動作が停止された前記適応フ

10

20

フィルタのフィルタ係数 $W_1$ を用いて、 $W_2 = CS_1 \cdot Z^{-m} - CS_2 \cdot W_1$ となるように前記フィルタのフィルタ係数 $W_2$ を設定するフィルタ係数計算手段とを有することを特徴とする音声入力装置。

【請求項 2】

請求項 1 において、

前記入力タイミング指示手段によって前記音声入力のタイミングが指示されて前記適応フィルタのフィルタ係数の更新動作が停止されたときに、このフィルタ係数を格納するフィルタ係数格納部をさらに備え、

前記フィルタ係数計算手段は、前記フィルタ係数格納部に格納されたフィルタ係数を用いて、前記フィルタのフィルタ係数の計算を行うことを特徴とする音声入力装置。

10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、マイクロホンによって集音した音声の中から各種の雑音を除去する音声入力システムに関する。

【0002】

【従来の技術】

最近の車載用機器、例えばナビゲーション装置やオーディオ機器に各種の操作指示を与える方法としては、利用者が操作パネルやリモートコントロール（リモコン）ユニットに備わった各種のキーを押下する方法の他に、利用者によって発せられた操作音声の内容を音声認識することによって行う方法がある。音声認識装置を用いて操作指示を行う場合には、操作キーの配置等を覚える必要がなく、しかも走行中に車両が振動した状態でキーの操作を行わないですむため、操作の簡略化が可能であり、特に最近ではプロセッサの高速化等に伴って車載用機器についても用いられることが多い。

20

【0003】

このような音声認識装置による認識率を低下させる要因の代表的なものには、ロードノイズやエンジンノイズ等の走行に伴って生じる車室内自動車ノイズと、車室内のオーディオ装置から出力されるオーディオ音がある。利用者の音声にこれらのノイズやオーディオ音が重畳されると、音声認識装置ではそれらを区別して利用者の音声に対してのみ音声認識処理を行うことができないため、正確な音声認識処理を行うことが困難となる。このため、従来は、適応マイクロホンアレイ技術を用いてロードノイズを低減したり、トークスイッチが押下されたときにオーディオ音の出力を中断あるいは音量を下げたりして、音声認識の対象となる音声信号に重畳される各種のノイズやオーディオ音を低減する工夫を行っている。

30

【0004】

【発明が解決しようとする課題】

ところで、上述したようにトークスイッチを押下したときにオーディオ音の出力を中断したり、オーディオ音の音量を下げた場合には、頻繁に音声認識処理を行ったときにオーディオ音が断続することになるため、利用者による聴取を妨げるという問題がある。

40

【0005】

特に、音声入力によって車載用機器の操作を行っている利用者以外に搭乗者は、この操作とは無関係にオーディオ音を聴取している場合もあるため、聴取しているオーディオ音が頻繁に断続されると不快であり、認識対象となる入力音声以外の発生音を断続することなく入力音声を抽出することができる手法が望まれている。

【0006】

本発明は、このような点に鑑みて創作されたものであり、その目的は、入力音声以外の発生音を断続することなく入力音声を抽出することができる音声入力システムを提供することにある。

【0007】

50

**【課題を解決するための手段】**

上述した課題を解決するために、本発明の音声入力装置は、集音手段から出力される音声信号に含まれる周辺雑音成分を、適応フィルタを用いて除去する周辺雑音除去手段と、この周辺雑音が除去された後の音声信号からスピーカの出力成分を除去するスピーカ音声除去手段とを有しており、音声信号から周辺雑音成分とスピーカ音成分（例えばスピーカから出力されるオーディオ音声）を除去して、利用者が発声した音声のみを抽出することができ、その際にスピーカの音量等を下げる等の操作を行う必要がない。

**【0008】**

具体的には、上述した集音手段は第1および第2のマイクロホンを有しており、周辺雑音除去手段は、第1のマイクロホンから出力されたノイズ信号を第2のマイクロホンから出力されたノイズ信号により模擬する適応フィルタと、これら2つのノイズ信号の差分を計算する第1の演算手段とを有している。適応フィルタを用いることにより、2つのマイクロホンから出力されるノイズ信号を一致させることができるため、これら2つの信号の差分を求めることにより、周辺雑音成分をキャンセルすることができ、この周辺雑音成分が除去された音声信号を得ることができる。

10

**【0009】**

また、上述したスピーカ音声除去手段は、スピーカおよび周辺雑音除去手段を介した伝達系の特性を模擬するフィルタと、周辺雑音除去手段を介した信号とフィルタを介した信号との差分を計算する第2の演算手段とを有している。スピーカから出力されるオーディオ音声等の伝達系の特性がフィルタによって模擬されるため、スピーカから実際に出力される音声に対応する信号と等価な信号をフィルタによって生成することができるため、スピーカの出力音成分をキャンセルすることができ、この出力音成分が除去された音声信号を得ることができる。

20

**【0010】**

また、集音手段に対する音声入力のタイミングが指示された場合に、適応フィルタのフィルタ係数の更新動作を停止するとともに、この更新動作が停止された適応フィルタのフィルタ係数に基づいてフィルタのフィルタ係数を計算することが好ましい。適応フィルタのフィルタ係数が固定されると、周辺雑音除去手段を介した音響伝達系の特性が固定化されるため、この特性を模擬するフィルタのフィルタ係数を計算により求めることができる。

**【0011】**

30

**【発明の実施の形態】**

以下、本発明を適用した一実施形態の音声入力装置について、図面を参照しながら説明する。

図1は、本実施形態の音声入力装置の構成を示す図である。同図に示す本実施形態の音声入力装置1は、2つのマイクロホン10、12、周辺雑音除去部20、スピーカ音声除去部40、トークスイッチ50を含んで構成されており、利用者がマイクロホン10、12に向かって車載用のナビゲーション装置（図示せず）等に対する操作音声を発声したときに、マイクロホン10、12から出力される音声信号に含まれる各種の雑音を除去して音声認識装置90に入力する。マイクロホン10、12から出力される音声信号に含まれる各種の雑音としては、エンジンの回転音や道路上を走行するために生じる走行雑音等からなる周辺雑音と、オーディオ機器100に接続されたスピーカ102から車室内に放出されマイクロホン10、12によって集音されるオーディオ音とを考慮するものとする。

40

**【0012】**

マイクロホン10、12は、利用者が発声した操作音声を集音可能な車室内の所定位置に設置されている。例えば、主に運転者が操作音声を発するものとする、運転席前方のサンバイザーなどに所定間隔（例えば15cm間隔）で設置される。

**【0013】**

周辺雑音除去部20は、マイクロホン10、12から出力される音声信号に含まれる周辺雑音を除去するためのものであり、遅延素子22、適応フィルタ24、LMSアルゴリズム処理部26、演算部28、フィルタ係数格納部30を含んで構成されている。

50

## 【 0 0 1 4 】

遅延素子 2 2 は、一方のマイクロホン 1 0 から出力される音声信号 d 1 を所定時間遅延した信号 Y 1 を出力する。例えば、伝達特性  $Z^{-m}$  を有する F I R (Finite Impulse Response) 型のデジタルフィルタであり、遅延時間 t に対応するタップ係数を 1 に設定し、それ以外のタップ係数を 0 に設定することにより実現される。

## 【 0 0 1 5 】

また、適応フィルタ 2 4 は、フィルタ係数 (タップ係数) W 1 を有する F I R 型のデジタルフィルタであり、他方のマイクロホン 1 2 から出力される音声信号 d 2 に対して所定の適応等化処理を行う。このフィルタ係数 W 1 は、L M S (Least Mean Square) アルゴリズム処理部 2 6 によって更新される。

10

## 【 0 0 1 6 】

L M S アルゴリズム処理部 2 6 は、マイクロホン 1 2 から出力される音声信号と演算部 2 8 から出力される差分信号 Y 3 (後述する) が入力されており、L M S アルゴリズムを用いることにより、演算部 2 8 から出力される差分信号 Y 3 (後述する) のパワーが最小となるように、上述した適応フィルタ 2 4 のフィルタ係数 W 1 を更新する。

## 【 0 0 1 7 】

演算部 2 8 は、一方のマイクロホン 1 0 から出力される音声信号 d 1 と他方のマイクロホン 1 2 から出力される音声信号 d 2 との差分を計算して、差分信号 Y 3 を出力する。上述した L M S アルゴリズム処理部 2 6 によってこの差分信号 Y 3 のパワーが最小となるように適応フィルタ 2 4 のフィルタ係数 W 1 が設定されるため、遅延素子 2 2 から出力される遅延後の信号 Y 1 を打ち消すように適応フィルタ 2 4 の出力信号 Y 2 の振幅、位相が制御される。演算部 2 8 から出力される差分信号 Y 3 は、周辺雑音除去部 2 0 の出力信号として後段のスピーカ音声除去部 4 0 に入力される。

20

## 【 0 0 1 8 】

フィルタ係数格納部 3 0 は、所定のタイミングにおける適応フィルタ 2 4 のフィルタ係数 W 1 を格納する。具体的には、利用者によってトークスイッチ 5 0 が押下されると、L M S アルゴリズム処理部 2 6 による適応フィルタ 2 4 のフィルタ係数 W 1 の更新動作が所定時間中断され、このときのフィルタ係数 W 1 がフィルタ係数格納部 3 0 に格納される。また、フィルタ係数格納部 3 0 には、スピーカ 1 0 2 から一方のマイクロホン 1 0 までの伝達特性 C S 1 と、他方のマイクロホン 1 2 までの伝達特性 C S 2 が格納されている。これらの伝達特性 C S 1、C S 2 は、あらかじめ測定されたものが用意されている。例えば、車両購入時に利用者の操作によって伝達特性 C S 1、C S 2 が測定され、車種別にあらかじめ用意されている伝達特性 C S 1、C S 2 の中から本実施形態の音声入力装置 1 が搭載された車種に適合したものが選択され、フィルタ係数格納部 3 0 に格納される。

30

## 【 0 0 1 9 】

また、図 1 に示したスピーカ音声除去部 4 0 は、周辺雑音除去部 2 0 の出力信号 Y 3 に含まれるオーディオ音を除去するためのものであり、フィルタ 4 2、フィルタ係数計算部 4 4、演算部 4 6 を含んで構成されている。

## 【 0 0 2 0 】

フィルタ 4 2 は、フィルタ係数 W 2 を有する F I R 型のデジタルフィルタであり、オーディオ機器 1 0 0 から出力されるオーディオ信号 X に対して所定の処理を行う。フィルタ係数計算部 4 4 は、周辺雑音除去部 2 0 内のフィルタ係数格納部 3 0 に格納されている各種のデータ (適応フィルタ 2 4 のフィルタ係数 W 1、スピーカ 1 0 2 から各マイクロホン 1 0、1 2 までの伝達特性 C S 1、C S 2) に基づいて、フィルタ 4 2 のフィルタ係数 W 2 を計算によって求める。このフィルタ係数 W 2 は、スピーカ 1 0 2 から出力されたオーディオ音が周辺雑音除去部 2 0 を通ってスピーカ音声除去部 4 0 内の演算部 4 6 に到達するまでの伝達特性を示しており、具体的な計算方法については後述する。

40

## 【 0 0 2 1 】

演算部 4 6 は、周辺雑音除去部 2 0 の出力信号 Y 3 とフィルタ W 2 の出力信号 Y 4 との差分を計算して、差分信号 Y 5 を出力する。スピーカ音声除去部 4 0 内の演算部 4 6 から出

50

力される差分信号 Y 5 は、音声入力装置 1 の出力信号として後段の音声認識装置 9 0 に入力される。

#### 【 0 0 2 2 】

また、図 1 に示したトークスイッチ 5 0 は、利用者が操作音声を発声する前に操作されるものであり、操作状況が周辺雑音除去部 2 0 内の L M S アルゴリズム処理部 2 6 とスピーカ音声除去部 4 0 内のフィルタ係数計算部 4 4 のそれぞれに送られる。L M S アルゴリズム処理部 2 6 は、トークスイッチ 5 0 が押下されたときに、適応フィルタ 2 4 のフィルタ係数 W 1 を更新する処理を一旦停止し、このときのフィルタ係数 W 1 をフィルタ係数格納部 3 0 に格納する。また、フィルタ係数計算部 4 4 は、トークスイッチ 5 0 が押下された後にフィルタ係数格納部 3 0 に格納されたフィルタ係数 W 1 と、あらかじめ格納されている伝達特性 C S 1、C S 2 を読み出して、フィルタ 4 2 のフィルタ係数 W 2 を計算する。

10

#### 【 0 0 2 3 】

上述したマイクロホン 1 0、1 2 が集音手段に、周辺雑音除去部 2 0 が周辺雑音除去手段に、L M S アルゴリズム処理部 2 6 がフィルタ処理手段に、演算部 2 8 が第 1 の演算手段に、スピーカ音声除去部 4 0 がスピーカ音声除去手段に、フィルタ係数計算部 4 4 がフィルタ係数計算手段に、演算部 4 6 が第 2 の演算手段にそれぞれ対応する。

#### 【 0 0 2 4 】

本実施形態の音声入力装置 1 はこのような構成をしており、次にその動作を説明する。図 2 は、本実施形態の音声入力装置 1 に含まれる各フィルタの係数更新手順を示す流れ図である。

20

#### 【 0 0 2 5 】

L M S アルゴリズム処理部 2 6 には、車両の走行状態に関係なく、常にトークスイッチ 5 0 が利用者によって押下されたか否かを判定する機能も付加されており（ステップ 2 0 0）、トークスイッチ 5 0 が押下されない場合には、演算部 2 8 から出力される差分信号 Y 3 のパワーが最小になるように適応フィルタ 2 4 のフィルタ係数 W 1 を計算して更新する（ステップ 2 0 1）。このようにして適応フィルタ 2 4 のフィルタ係数 W 1 が設定されるため、周辺雑音の発生源（エンジンやタイヤ等）から一方のマイクロホン 1 0 および遅延素子 2 2 を経由して演算部 2 8 に至るまでの周辺雑音の一方の音響伝達系に対応する伝達特性と、周辺雑音の発生源から他方のマイクロホン 1 2 および適応フィルタ 2 4 を経由して演算部 2 8 に至るまでの周辺雑音の他方の音響伝達系に対応する伝達特性を同じにすることができる。したがって、これら 2 つの音響伝達系を通過して演算部 2 8 に到達する周辺雑音は、位相および振幅が同じになるため、演算部 2 8 によって 2 つの入力信号 Y 1、Y 2 の差分を求めることにより、この周辺雑音を除去することができる。

30

#### 【 0 0 2 6 】

このようにして行われる周辺雑音の除去動作と並行して上述したステップ 2 0 0 の判定動作（トークスイッチが押下されたか否かの判定動作）が行われており、利用者によってトークスイッチ 5 0 が押下されると、次に、L M S アルゴリズム処理部 2 6 は、適応フィルタ 2 4 のフィルタ係数 W 1 の更新処理を停止し（ステップ 2 0 2）、その時点で適応フィルタ 2 4 に設定されているフィルタ係数 W 1 をフィルタ係数格納部 3 0 に格納する（ステップ 2 0 3）。なお、上述したように、このフィルタ係数格納部 3 0 には、あらかじめ求められたスピーカ 1 0 2 から各マイクロホン 1 0、1 2 までの伝達特性 C S 1、C S 2 と、遅延素子 2 2 の伝達特性  $Z^{-m}$  が格納されている。

40

#### 【 0 0 2 7 】

次に、スピーカ音声除去部 4 0 内のフィルタ係数計算部 4 4 は、フィルタ係数格納部 3 0 に格納されているフィルタ係数 W 1、伝達特性 C S 1、C S 2、 $Z^{-m}$  を読み出して、これらに基づいてフィルタ 4 2 に設定するフィルタ係数 W 2 を計算する（ステップ 2 0 4）。

#### 【 0 0 2 8 】

ところで、オーディオ機器 1 0 0 の出力信号 X と、周辺雑音除去部 2 0 から出力されてスピーカ音声除去部 4 0 内の演算部 4 6 に入力される信号 Y 3 との関係は、周波数領域で、 $Y 3 = (C S 1 \cdot Z^{-m} - C S 2 \cdot W 1) X \dots (1)$

50

と表すことができる。また、オーディオ機器 100 の出力信号 X と、スピーカ音声除去部 40 内のフィルタ 42 から出力されて演算部 46 に入力される信号 Y4 との関係は、

$$Y4 = W2 \cdot X \quad \dots (2)$$

と表すことができる。したがって、 $W2 = CS1 \cdot Z^{-m} - CS2 \cdot W1$  を満たすようにフィルタ 42 のフィルタ係数 W2 を設定することにより、オーディオ機器 100 から出力されて周辺雑音除去部 20 を通って演算部 46 に入力される信号 Y3 と、オーディオ機器 100 から出力されてスピーカ 102 を介さずに直接フィルタ 42 を通って演算部 46 に入力される信号 Y4 とを同じにすることができ、演算部 46 の出力信号 Y5 に含まれるオーディオ音声成分を除去することができる。特に、トークスイッチ 50 が押下されたときにフィルタ係数 W1 の更新処理が停止されているため、フィルタ係数 W1 および伝達特性 CS1、CS2、 $Z^{-m}$  の全てが既知の値であり、計算によって容易にフィルタ 42 のフィルタ係数 W2 を求めることができる。

10

#### 【0029】

上述した動作は、トークスイッチ 50 が押下されてから一定時間（例えば数秒間）行われる。LMS アルゴリズム処理部 26 には、トークスイッチ 50 が押下されたか否かを判定する部分も内蔵されており（ステップ 200）、一定時間が経過すると（ステップ 205）、ステップ 201 に戻ってフィルタ係数 W1 の計算および更新処理以降の動作を繰り返す。

#### 【0030】

このように、本実施形態の音声入力装置 1 では、利用者がトークスイッチ 50 を押下した後にマイクロホン 10、12 に向かって何らかの操作音声を発声したときに、マイクロホン 10、12 の出力信号に含まれる周辺雑音成分を前段の周辺雑音除去部 20 によって除去し、オーディオ機器 100 から出力されるオーディオ音声成分を後段のスピーカ音声除去部 40 によって除去しており、これらの各種の雑音が除去された後の操作音声のみが含まれる音声信号が音声認識装置 90 に入力される。したがって、マイクロホン 10、12 によって集音された音声に周辺雑音成分やオーディオ音声成分が含まれている場合であっても、これらを除去した操作音声のみが音声認識装置 90 に送られるため、音声認識装置 90 において所定の音声認識処理を実行したときの認識率を高めることができる。また、利用者が操作音声を発声する際に、スピーカ 102 から出力されるオーディオ音声を遮断したり、その音量を小さくした状態で、利用者による操作音声の入力を行うことができ、オーディオ音声を断続させた場合に聴取者が感じる違和感、不快感をなくすることができる。

20

30

#### 【0031】

特に、トークスイッチ 50 が押下される直前まで適応フィルタ 24 におけるフィルタ係数 W1 の更新が行われているため、実状に即した周辺雑音の除去を行うことができる。

#### 【0032】

なお、本発明は上記実施形態に限定されるものでなく、本発明の要旨の範囲内で種々の変形実施が可能である。例えば、上述した実施形態では、ナビゲーション装置に対する操作音声を利用者が発声する場合を例にとって説明したが、ナビゲーション装置以外の車載用機器（例えばオーディオ機器）に対する操作音声を利用者が発声する場合にも本発明を適用することができる。また、音声認識の対象となる音声を入力する場合の他に、例えば車室内で移動体電話を用いる場合に、会話に含まれる周辺雑音やオーディオ音声を除去する場合であってもよい。また、車室内で使用される場合の他に、屋内に設置された音声認識装置やその他の装置に対して入力する音声に含まれる周辺雑音成分やスピーカ出力音声（オーディオ音声に限定されず、テレビジョン放送の音声等であってもよい）成分を除去するようにしてもよい。

40

#### 【0033】

また、上述した実施形態では、2つのマイクロホン 10、12 を用いたが、3つ以上のマイクロホンを用いるようにしてもよい。また、利用者は、トークスイッチ 50 を押下した後に操作音声を発声するようにしたが、操作音声の先頭部分を検出する回路を設けて、こ

50

の検出信号を L M S アルゴリズム処理部 2 6 やフィルタ係数計算部 4 4 に送るようにしてもよい。

【 0 0 3 4 】

【 発明の効果 】

上述したように、本発明によれば、集音手段から出力される音声信号に含まれる周辺雑音成分を、適応フィルタを用いて除去するとともに、この周辺雑音が除去された後の音声信号からスピーカの出力音成分を除去しており、利用者が発声した音声のみを抽出することができるため、その際にスピーカの音量等を下げる等の操作を行う必要がない。

【 図面の簡単な説明 】

【 図 1 】 本実施形態の音声入力装置の構成を示す図である。

10

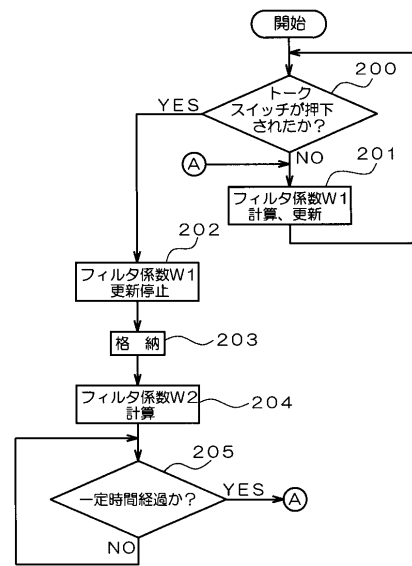
【 図 2 】 本実施形態の音声入力装置の動作手順を示す流れ図である。

【 符号の説明 】

- 1 音声入力装置
- 1 0、1 2 マイクホン
- 2 0 周辺雑音除去部
- 2 2 遅延素子
- 2 4 適応フィルタ
- 2 6 L M S アルゴリズム処理部
- 2 8、4 6 演算部
- 3 0 フィルタ係数格納部
- 4 0 スピーカ音声除去部
- 4 2 フィルタ
- 4 4 フィルタ係数計算部
- 5 0 トークスイッチ
- 9 0 音声認識装置
- 1 0 0 オーディオ機器
- 1 0 2 スピーカ

20

【 図 2 】





---

フロントページの続き

審査官 山下 剛史

(56)参考文献 特開平02-244099(JP,A)  
特開昭61-194913(JP,A)

(58)調査した分野(Int.Cl., DB名)  
G10L 15/20