

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4810044号
(P4810044)

(45) 発行日 平成23年11月9日(2011.11.9)

(24) 登録日 平成23年8月26日(2011.8.26)

(51) Int.Cl. F I
G 1 0 L 15/04 (2006.01) G 1 0 L 15/04 3 0 0 Z
G 1 0 L 11/02 (2006.01) G 1 0 L 15/04 3 0 0 A
 G 1 0 L 11/02

請求項の数 9 (全 11 頁)

(21) 出願番号	特願2001-555084 (P2001-555084)	(73) 特許権者	509199502
(86) (22) 出願日	平成13年1月10日(2001.1.10)		ニュアンス コミュニケーションズ オー
(65) 公表番号	特表2003-521006 (P2003-521006A)		ストリア ゲーエムペーハー
(43) 公表日	平成15年7月8日(2003.7.8)		オーストリア共和国 1101 ウィーン
(86) 国際出願番号	PCT/EP2001/000193		トリースター・シュトラッセ 64
(87) 国際公開番号	W02001/056015	(74) 代理人	100070150
(87) 国際公開日	平成13年8月2日(2001.8.2)		弁理士 伊東 忠彦
審査請求日	平成20年1月9日(2008.1.9)	(72) 発明者	バートオシック ヘインリッチ
(31) 優先権主張番号	00890026.8		オランダ国 5656 アーアー アイン
(32) 優先日	平成12年1月27日(2000.1.27)		ドーフエン プロフ ホルストラーン 6
(33) 優先権主張国	欧州特許庁 (EP)		審査官 菊池 智紀

最終頁に続く

(54) 【発明の名称】 2つのスイッチオフ規準を持つ音声検出装置

(57) 【特許請求の範囲】

【請求項 1】

受信信号における音声信号を検出すると共に、音声タイムスロットを決定する音声検出装置であって、該音声検出装置は、

前記受信信号のエネルギー量が第1エネルギー閾値を越える場合に第1検出情報を送出するスイッチオン閾検出器と、

前記受信信号のエネルギー量が前記第1エネルギー閾値より小さな第2エネルギー閾値に満たない場合に第2検出情報を送出するスイッチオフ閾検出器と、

前記第1検出情報及び第2検出情報を入力及び処理し、第1スイッチオフ期間にわたって中断なく前記第2検出情報が入力された場合に音声タイムスロットを特徴付ける音声検出情報の送出手段と、

を含むような音声検出装置において、

前記情報処理手段は、第3スイッチオフ期間にわたって前記第1検出情報が入力されなかった場合にも前記音声検出情報の送出手段を追加的に終了させるように構成され、前記第3スイッチオフ期間の開始は前記第1検出情報が入力されなくなった後に前記第2検出情報が最初に入力される時として決定されることを特徴とする音声検出装置。

【請求項 2】

請求項 1 に記載の音声検出装置において、前記情報処理手段においては前記第1スイッチオフ期間が前記第3スイッチオフ期間よりも短いことを特徴とする音声検出装置。

【請求項 3】

請求項 1 に記載の音声検出装置において、前記スイッチオン閾検出器は、前記受信信号のエネルギー量が少なくとも 1 つのスイッチオン期間にわたって前記第 1 エネルギー閾値より大きい場合に前記第 1 検出情報を生成するように構成されていることを特徴とする音声検出装置。

【請求項 4】

請求項 1 に記載の音声検出装置において、当該音声検出装置が、前記第 1 エネルギー閾値及び / 又は第 2 エネルギー閾値を前記受信信号に含まれるノイズ信号のエネルギー量に適応化させるように構成されていることを特徴とする音声検出装置。

【請求項 5】

受信信号における音声信号を検出すると共に、音声タイムスロットを決定する音声検出方法であって、該音声検出方法は、

前記受信信号のエネルギー量が第 1 エネルギー閾値を越える場合に第 1 検出情報を送出手続きと、

前記受信信号のエネルギー量が前記第 1 エネルギー閾値より小さな第 2 エネルギー閾値に満たない場合に第 2 検出情報を送出手続きと、

前記第 1 検出情報及び第 2 検出情報を入力及び処理するステップと、

第 1 スwitchオフ期間にわたって中断なく前記第 2 検出情報が入力された場合に音声タイムスロットを特徴付ける音声検出情報の送出手続きを終了させるステップと、
を有するような音声検出方法において、

前記終了させるステップは、第 3 スwitchオフ期間にわたって前記第 1 検出情報が入力されなかった場合にも前記音声検出情報の送出手続きが追加的に終了されるように構成され、前記第 3 スwitchオフ期間の開始は前記第 1 検出情報が入力されなくなった後に前記第 2 検出情報が最初に入力される時として決定されることを特徴とする音声検出方法。

【請求項 6】

請求項 5 に記載の音声検出方法において、前記第 1 検出情報は、前記受信信号のエネルギー量が少なくとも 1 つのスイッチオン期間にわたって前記第 1 エネルギー閾値より大きくなるまで送出手続きされないことを特徴とする音声検出方法。

【請求項 7】

請求項 5 に記載の音声検出方法において、前記第 1 エネルギー閾値及び / 又は第 2 エネルギー閾値が、前記受信信号に含まれるノイズ信号のエネルギー量に適応化されることを特徴とする音声検出方法。

【請求項 8】

デジタルコンピュータの内部メモリに直接ロードすることができると共に、ソフトウェアコード部を含むようなコンピュータプログラムにおいて、該コンピュータプログラムが前記デジタルコンピュータ上で実行された場合に、請求項 5 に記載の音声検出方法の前記各ステップが前記デジタルコンピュータにより実行されることを特徴とするコンピュータプログラム。

【請求項 9】

請求項 8 に記載のコンピュータプログラムが記憶されている、コンピュータにより読み取ることが可能な媒体。

【発明の詳細な説明】

【0001】

【技術分野】

本発明は、請求項 1 の導入部に記載したような音声検出装置、請求項 5 の導入部に記載したような受信信号における十分に良好な信号対雑音比を持つ音声信号を検出する音声検出方法、及び請求項 8 の導入部に記載したようなコンピュータプログラムに関する。

【0002】

【背景技術】

上記のような音声検出装置、上記のような音声検出方法及び上記のようなコンピュータプログラム製品は、1998 年から本出願人により “フリースピーチ 98” と呼ばれるコ

10

20

30

40

50

ンピュータプログラムとして販売されている音声認識装置の一部として知られている。コンピュータが上記コンピュータプログラム“フリースピーチ98”を実行し、ユーザが該コンピュータに接続されたマイクロフォンに対して文を口述すると、上記既知の音声認識装置の音声認識手段により認識された該文が上記コンピュータに接続されたモニタ上に表示される。口述の間、ユーザはマイクロフォンに向かって時には滑らかに話し、また時には短い休止を伴いながら話す。ユーザは、時には、マイクロフォンを口から過度に遠くで保持し、これにより該マイクロフォンにより発生される電気マイクロフォン信号の信号対雑音比が悪くなる。従って、所謂音声タイムスロットの間においては、上記マイクロフォン信号はユーザにより話された文に対応する音声信号を含み、所謂休止タイムスロットの間では上記マイクロフォン信号は音声信号を含まないか又は信号対雑音比が劣る音声信号を含むことになる。

10

【0003】

上記既知の音声認識装置の音声検出装置には、上記マイクロフォンにより受信信号として又は該受信信号を表す受信データとして送出されるマイクロフォン信号を供給することができる。音声検出装置は、受信信号における音声信号の開始及び終了を検出して、対応する音声タイムスロットを決定する。音声検出装置は音声タイムスロットの間において音声認識手段に音声検出情報を供給し、該音声認識手段は音声タイムスロットの間にマイクロフォンから送出されるマイクロフォン信号のみを処理する。

【0004】

受信信号における音声信号を検出するために、上記既知の音声検出装置はスイッチオン閾検出器及びスイッチオフ閾検出器を含み、これら検出器は入力信号のエネルギー量（energy content）を第1及び第2エネルギー閾値と比較するが、ここで第1エネルギー閾値は第2エネルギー閾値より高い。受信信号のエネルギー量が第1エネルギー閾値を越えると、上記スイッチオン閾検出器が第1検出情報を発生し、受信信号のエネルギー量が第2エネルギー閾値に達しないと、スイッチオフ閾検出器が第2検出情報を発生する。

20

【0005】

音声タイムスロットを決定するために、上記音声検出装置は上記検出情報を入力し且つ処理する情報処理手段を含んでいる。音声タイムスロットのスイッチオンの規準として第1検出情報の発生が検出され、その後音声タイムスロットの開始は上記情報処理手段により、上記スイッチオン規準が満たされる240ms前として決定される。第1スイッチオフ期間の間の第2検出情報の中断のない発生は音声タイムスロットのスイッチオフの規準として検出され、その後、音声タイムスロットの終了は上記情報処理手段により上記スイッチオフ規準が満たされた時として決定される。

30

【0006】

上記既知の音声検出装置、既知の音声検出方法及び既知のコンピュータプログラムは、受信信号のスイッチオフ規準が、該受信信号のエネルギー量が第2エネルギー閾値周辺で変化する場合に満たされないという欠点を有している。斯様な受信信号は、例えばユーザが電話の会話のために口述を中断し、マイクロフォンを机の上に置いたような場合に、音声認識装置に供給される。この電話の会話の間にユーザにより又は室内の他の者により、マイクロフォンから遠く離れた位置で話される言葉は、低い信号対雑音比を持つ音声信号を時々含むようなマイクロフォン信号として該マイクロフォンに供給される。この低い信号対雑音比を持つ音声信号を伴う受信信号は、音声認識装置により音声認識に適した音声信号として誤って検出される。何故なら、音声タイムスロットが当該音声検出装置により終了されていないからである。このようにして、認識されるために供給されたものではない音声信号が、音声認識手段により、上記の低い信号対雑音比による当該音声認識装置の悪い認識率で処理され、殆どの場合誤った文が認識される。

40

【0007】**【発明の開示】**

本発明の目的は、上述したような問題を取り除き、冒頭で述べたような形式の音声検出装置、音声検出方法及びコンピュータプログラム製品であって、音声タイムスロットを高信

50

頼度で終了させるために第2スイッチオフ規準が設けられたような音声検出装置、音声検出方法及びコンピュータプログラム製品を提供することにある。この目的は、上記のような音声検出装置によれば請求項1の特徴部の対策により達成され、上記のような音声検出方法によれば請求項5の特徴部の対策により達成され、上記のようなコンピュータプログラム製品によれば請求項8の特徴部の対策により達成される。

【0008】

これによれば、情報処理手段において、第2スイッチオフ期間の間の第1検出情報の中断されない欠落が音声タイムスロットを終了させる第2スイッチオフ規準として検出され、その後、該音声タイムスロットの終了が上記情報処理手段により該第2スイッチオフ規準が満たされるかに応じて決定されるようになる。この第2スイッチオフ規準に加えて、又は該規準の代わりに、上記情報処理手段は第3スイッチオフ規準も検証することができ、該第3スイッチオフ規準によれば、第1検出情報が受信されなくなった後第2検出情報が最初に入力されてから第3スイッチオフ期間の間に該第1検出情報が入力されなかったかが判定される。

10

【0009】

上記第2及び/又は第3スイッチオフ規準に依存して音声タイムスロットを終了させることは、例えば、上述したような作業状況が発生し、受信信号が前記閾の周辺で変化するような場合においても、良好な信号対雑音比を有する1つの音声信号のみが音声認識装置により音声認識のために高信頼度で使用されるという利点を提供する。

【0010】

請求項2に記載の対策によれば高信頼度の第2スイッチオフ規準が得られ、請求項3に記載の対策によれば音声タイムスロット用の高信頼度のスイッチオン規準が得られる。請求項4に記載の対策は、スイッチオン閾検出器及びスイッチオフ閾検出器のエネルギー閾値を受信信号におけるエネルギー量に適應させるので、良好な信号対雑音比を持つ音声信号の検出が改善される。

20

【0011】

【発明を実施するための最良の形態】

以下、本発明を図に示す2つの実施例を参照して説明するが、本発明は、これら実施例に限定されるものではない。

【0012】

図1は、コンピュータプログラム製品をロードすることが可能な内部メモリを有するコンピュータを示し、該プログラム製品はソフトウェアコード部を有すると共に音声認識ソフトウェアにより形成されている。該コンピュータ1が音声認識ソフトウェアを処理する場合、該コンピュータ1は音声信号に割り当てられるべき文章情報を認識する音声認識装置を形成する。

30

【0013】

該コンピュータ1のオーディオポート2にはマイクロフォン3を接続することができ、該マイクロフォンにはユーザが文又は命令を口述することができ、該マイクロフォンによりマイクロフォン信号MSをコンピュータ1に供給することができる。マイクロフォン3に向かって、ユーザは時には文を滑らかに話し、時には短い休止を伴って文を話す。ユーザは時にはマイクロフォン3を口から遠くに離して保持し、その場合には、該マイクロフォンにより送られるマイクロフォン信号MSの信号対雑音比は相対的に悪くなる。従って、所謂音声タイムスロットTSの間ではマイクロフォン信号MSはユーザにより話された文に対応する音声信号SSを含み、所謂休止タイムスロットTPにおいては音声信号SSを含まないか又は音声認識装置により処理されるのには適さない、劣った信号対雑音比の音声信号SSを含むことになる。マイクロフォン3によりオーディオポート2を介してコンピュータ1に供給される斯様なマイクロフォン信号MSは、入力信号としてコンピュータ1へ、従って当該音声認識装置へ処理されるべく供給することができる。図2Aは斯様なマイクロフォン信号を時間の関数として示すが、該信号に関しては後に更に説明する。

40

【0014】

50

コンピュータ1のモニタポート4にはモニタ5を接続することができ、該モニタにより当該音声認識装置により認識された文TXを表示することができる。この目的のため、認識された文を表すテキスト情報TIをモニタポート4からモニタ5に伝送することができる。

【0015】

前記マイクロフォン信号MSはオーディオポート2からA/D変換器6に供給することができる。該A/D変換器6は、通常知られているように、該A/D変換器6に供給される上記マイクロフォン信号MSをデジタル化するように構成されている。A/D変換器6は、マイクロフォン信号MSに含まれるユーザにより話された文の情報を含む受信データEDを生成することができる。

10

【0016】

当該音声認識装置は、更に、A/D変換器6により送られる上記受信データEDを供給することが可能な記憶手段7を含んでいる。コンピュータ1における該記憶装置7は、ハードディスクにより形成され、該ハードディスクに供給される受信データEDを記憶するように構成されている。該記憶手段7に供給された受信データEDは、音声検出情報SDIが入力された場合にのみ永久に記憶されるが、これについては後述する。

【0017】

当該音声認識装置は更に音声検出装置8を含み、該音声検出装置にもA/D変換器6により送られる受信データEDを供給することができる。該音声検出装置8は受信データEDを評価することによりタイムスロットを検出するように構成され、該タイムスロットの間においては上記マイクロフォン信号MSは十分に良好な信号対雑音比を有するような音声信号SSを含む。このようなタイムスロットが検出された場合、音声検出装置8は適切な音声タイムスロットTSを決定するが、これについては後に詳述する。

20

【0018】

更に、当該音声認識装置は上記マイクロフォン信号MSのうちの音声タイムスロットTSの間に受信された部分のみを評価する。何故なら、マイクロフォン信号MSのこれらの部分のみが、ユーザにより話された文の正しく評価することができる情報を含むからである。音声タイムスロットTSを特徴付けるために、音声検出装置8は音声検出情報SDIを記憶手段7に供給し、該記憶手段は結果としてユーザにより話された文の、当該音声認識装置により正しく評価することができる情報を含む受信データEDのみを記憶する。

30

【0019】

コンピュータ1により形成される当該音声認識装置は、更に、音声認識手段9を含み、該音声認識手段により音声認識方法が実行されて、記憶手段7に記憶された受信データEDを評価する。この目的のため、音声認識手段9により記憶手段7に対して活性化情報AIを供給して、記憶手段7に永久的に記憶された受信データEDの送出を可能にする。音声認識手段9のような斯かる音声認識手段の構成及び動作方法、並びに音声認識手段9において実行されるような音声認識方法のステップは、長く知られており、例えば文献国際特許公開公報WO99/35640等に開示されている。

【0020】

ユーザがマイクロフォン3に向かって文を話すと、例えば図2Aに示すようなマイクロフォン信号MSがコンピュータ1により形成された当該音声認識装置に供給される。図2Aに示すマイクロフォン信号MSは、時間区間として、第1音声信号SS1、第2音声信号SS2、第3音声信号SS3及びノイズ信号RSを含んでいる。第3音声信号SS3は、雑音信号RSと比較して、比較的低いエネルギー量を有している。何故なら、ユーザが該文を話した際にマイクロフォン3を口から過度に離して保持したからである。従って、第3音声信号SS3の信号対雑音比は劣り、それ故、該第3音声信号は音声処理手段9で正しく処理するには不適である。

40

【0021】

音声検出装置8の目的は、マイクロフォン信号MSが第1音声信号SS1及び第2音声信号SS2を含むような音声タイムスロットTSを決定して、これら音声信号SS1及びS

50

S 2に含まれる情報を音声認識手段 9 が処理するのを可能にすることである。残りのタイムスロットは音声検出装置 8 により、マイクロフォン信号 M S がノイズ信号 R S 及び第 3 音声信号 S S 3 を含む休止タイムスロット P S として決定されるべきである。音声検出装置 8 により決定された休止タイムスロット P S の間においては、音声検出装置 8 により記憶手段 7 には何の音声検出情報 S D I も供給されない。

【 0 0 2 2 】

この目的を達成するため、音声検出装置 8 は、エネルギー平均化手段 1 0、スイッチオン閾検出器 1 1、スイッチオフ閾検出器 1 2 及び情報処理手段 1 3 を含んでいる。A / D 変換器 6 により送出することができる受信データ E D は、エネルギー決定手段 1 0 に供給することができる。該エネルギー決定手段 1 0 は、評価タイムスロット毎に、受信データ E D の評価によりマイクロフォン M S に含まれるエネルギー量を決定する。ここでは、評価タイムスロットは 2 0 ミリ秒である。受信データ E D はデジタルドメインで評価される。これはアナログドメインではマイクロフォン信号 M S の二乗及び該二乗されたマイクロフォン信号の各評価タイムスロットにわたる積分に相当するであろうのと同様である。専門家は、斯様なデジタルドメインでのデータの評価には長い間馴染みがある。斯様にして決定されたエネルギー情報 E I はエネルギー決定手段 1 0 によりスイッチオン閾検出器 1 1 及びスイッチオフ閾検出器 1 2 に供給することができるが、該情報はマイクロフォン信号 M S のエネルギー量を特徴付けている。

10

【 0 0 2 3 】

図 2 B は、エネルギー決定手段 1 0 により決定された図 2 A に示すマイクロフォン信号 M S のエネルギー情報 E I を時間の関数として示している。マイクロフォン信号 M S に含まれる音声信号 S S 1 及び S S 2 がノイズ信号 R S 及び第 3 音声信号 S S 3 より大きなエネルギー量を含むことを検出することができ、その結果、これら音声信号 S S 1 及び S S 2 の検出がエネルギー情報 E I の評価により可能となる。

20

【 0 0 2 4 】

この目的のため、スイッチオン閾検出器 1 1 は、該スイッチオン閾検出器 1 1 に供給されるエネルギー情報 E I の値を、図 2 B に示すような当該スイッチオン閾検出器 1 1 に記憶された第 1 エネルギー閾値 E S 1 と連続的に比較する。スイッチオン閾検出器 1 1 は、マイクロフォン信号 M S のエネルギー量が第 1 エネルギー閾値 E S 1 より大きい場合に第 1 検出情報 D I 1 を生成するように構成されている。図 2 A に示すマイクロフォン信号 M S が当該音声認識装置により受信された場合に上記スイッチオン閾検出器 1 1 により生成される第 1 検出情報 D I 1 の時間の関数としての波形が、図 2 C に示されている。

30

【 0 0 2 5 】

更に、スイッチオフ閾検出器 1 2 は、該スイッチオフ閾検出器 1 2 に供給されるエネルギー情報 E I の値を、図 2 B に示すような当該スイッチオフ閾検出器 1 2 に記憶された第 2 エネルギー閾値 E S 2 と連続的に比較する。スイッチオフ閾検出器 1 2 は、マイクロフォン信号 M S のエネルギー量が第 2 エネルギー閾値 E S 2 より小さい場合に第 2 検出情報 D I 2 を送出するように構成されている。図 2 A に示すマイクロフォン信号 M S が当該音声認識装置により受信された場合に上記スイッチオフ閾検出器 1 2 により送出される第 2 検出情報 D I 2 の時間の関数としての波形が、図 2 D に示されている。

40

【 0 0 2 6 】

情報処理装置 1 3 には、第 1 検出情報 D I 1 と第 2 検出情報 D I 2 とを供給することができる。該情報処理装置 1 3 は、該装置に供給される上記検出情報 D I 1 及び D I 2 を評価し、音声タイムスロット T S を決定し、及び決定された音声タイムスロット T S の間に音声検出情報 S D I を送出するように構成されている。

【 0 0 2 7 】

以下においては、本発明の第 1 実施例による情報処理手段 1 3 の動作方法が例示として説明される。該例によれば、情報処理手段 1 3 は図 2 C 及び図 2 D に示す検出情報 D I 1 及び D I 2 を評価し、その後該情報処理装置 1 3 により、図 2 E に波形が時間の関数として表されたような音声検出情報 S D I が送出される。

50

【 0 0 2 8 】

時点 t_1 から、情報処理手段 13 は第 1 検出情報 DI_1 を入力し、時点 t_2 において情報処理手段 13 は、第 1 検出情報 DI_1 がスイッチオン期間 TE にわたり入力されたことを確定する。結果として、スイッチオン規準が第 1 音声タイムスロットに関して満足されるが、これが音声検出情報 SDI_1 により特徴付けられている。該第 1 音声タイムスロットの開始は、該情報処理手段 13 により、上記時点 t_1 より進み期間 TV だけ早い時点 t_3 において定められる。

【 0 0 2 9 】

上記スイッチオン期間 TE の間待つことは、例えばマイクロフォン 3 を机の上に置く場合に発生し得るような、短い大きい雑音のマイクロフォン信号 MS の短い大振幅が当該情報処理手段 13 により誤って音声信号 SS として検出されることがないという利点を提供する。第 1 音声タイムスロットの始点を上記進み期間 TV だけ進めて配置することにより、マイクロフォン信号 MS において検出された第 1 音声信号 SS_1 の受信データ ED が、第 1 エネルギー閾値 ES_1 に到達する前に、記憶手段 7 に記憶され、後に音声認識手段 9 により更に処理されるという利点が得られる。これにより、全ての第 1 音声信号 SS_1 の受信データ ED が記憶されると共に、該第 1 音声信号 SS_1 の開始が音声認識手段 9 による処理から失われることがなくなる。上述した 2 つの対策は、当該音声認識装置の認識率を有利に改善する。

【 0 0 3 0 】

スイッチオン規準が満足された後に、上記進み期間 TV 及びスイッチオン期間 TE だけ進められた受信データ ED の記憶に届くよう、記憶手段 7 に供給される受信データ ED は、常に、該記憶手段 7 の受信バッファに記憶される。進み期間 TV 及びスイッチオン期間 TE の間に、予想される受信データ ED は上記受信バッファに短期間記憶することができ、次いで、該データ ED は、スイッチオン規準が満たされる時点 t_2 において記憶手段 7 に永久的に記憶することができる。

【 0 0 3 1 】

情報処理手段 13 は時点 t_4 において該第 1 音声タイムスロットの終了を決定するようになっており、その際、該第 1 音声タイムスロットは音声期間 TS_1 を有するようになる。時点 t_4 において第 1 スイッチオフ規準が満足されるが、該規準によれば第 1 スイッチオフ期間 TA_1 にわたり第 2 検出情報 DI_2 が情報処理手段 13 により中断無く入力されねばならない。図 2 E に示すように、時点 t_3 から時点 t_4 まで、記憶されるべき第 1 音声信号 SS_1 の受信データ ED に対して、音声検出情報 SDI_1 が記憶手段 7 に供給される。

【 0 0 3 2 】

第 1 音声タイムスロットの終了の上述したような方法による決定は、音声信号 SS のエネルギー量が短期間に非常に小さい場合に、第 1 音声タイムスロットが誤って早く終了されて、該第 1 音声信号 SS_1 の最後の部分の受信データ ED が音声認識手段 9 に処理されるべく供給されなくなることがないであろうという利点を提供する。音声信号 SS の斯様な短い非常に小さなエネルギー量は、“t”又は“p”のような子音が発音される場合、及びマイクロフォン信号 MS の短い中断がある場合等に発生し得る。

【 0 0 3 3 】

図 2 の実施例によれば、情報処理手段 13 は、第 1 休止期間 TP_1 の後、第 1 音声タイムスロットに関して上述したのと同様にして、時点 t_5 を第 2 音声タイムスロットの開始と判定する。該第 2 音声タイムスロットの間、マイクロフォン信号 MS は第 2 音声信号 SS_2 を含み、該第 2 音声信号には第 3 音声信号 SS_3 が後続する。第 3 音声信号 SS_3 のエネルギー量は第 2 エネルギー閾値 ES_2 付近で変化し、その際、第 1 スイッチオフ期間 TA_1 より短い期間 TK の間だけ第 2 検出情報 DI_2 が入力される。従って、第 3 音声信号 SS_3 の間では第 1 スイッチオフ規準は満たされず、結果として第 2 音声タイムスロットは情報処理手段 13 により終了されないであろう。

【 0 0 3 4 】

10

20

30

40

50

本発明の第1実施例による情報処理手段13は、ここで、第2スイッチオフ規準が満たされるかを判定するように構成されている。該第2スイッチオフ規準は、第2スイッチオフ期間TA2の間に第1検出情報DI1が入力されなかった場合に満たされる。時点t6以降、情報処理手段13は、最早、第1検出情報DI1を入力することはなく、その結果、情報処理手段13は時点t7において該第2スイッチオフ規準の存在を確定する。図2Eに示すように、時点t5から時点t7までの第2音声タイムスロットTS2の間、第2の音声検出情報SDI2が記憶手段7に供給されて、時点t5以降第2音声信号SS2の受信データEDを記憶する。

【0035】

結果として、ノイズ信号RSのみ又は悪い信号対雑音比を持つ第3音声信号SS3のみを含むマイクロフォン信号MSの受信データEDは音声認識手段9に供給されず、これにより該音声認識手段9による誤った文の認識が防止されるという利点が得られる。

10

【0036】

以下においては、本発明による追加の対策及びそれらの利点を、本発明の第2実施例を参照して説明する。該第2実施例による音声検出装置は、上記第1実施例による図1に示した音声検出装置8に相当するが、該第2実施例による情報処理手段は第1スイッチオフ規準又は第3スイッチオフ規準が満たされるかを検証するように構成されている。第3スイッチオフ規準は、第3スイッチオフ期間TA3の間に第1検出情報DI1が入力されなかった場合に満たされるが、該第3スイッチオフ期間TA3の開始は、第1検出情報DI1が無くなった後、次いで第2検出情報DI2が入力された場合に検出される。

20

【0037】

以下においては、本発明の第2実施例による情報処理手段の動作方法が例示により説明される。この例によれば、図2Aに示すマイクロフォン信号MSが当該音声認識装置に供給され、図2C及び2Dに示す検出情報DI1及びDI2が上記情報処理手段により評価される。該第2実施例に基づく情報処理手段による評価の結果、該情報処理手段は、図2Fに示すようなパターンの音声検出情報SDIを記憶手段7に供給する。

【0038】

当該情報処理手段は、第3音声期間TS3を持つ第3の音声検出情報SDI3により特徴付けられ、本発明の第1実施例による前記第1音声タイムスロットに相当するような第3音声タイムスロットを決定する。この第3音声タイムスロットの開始は前記スイッチオン規準により決定され、該第3音声タイムスロットの終了は前記第1スイッチオフ規準により決定される。第2休止期間TP2の後、該第2実施例による情報処理手段は、上記スイッチオン規準が満たされる時点t5において第4音声タイムスロットの開始を決定する。

30

【0039】

時点t6以降では、当該情報処理手段は、最早、第1検出情報DI1を入力せず、時点t8において、第1検出情報DI1が無くなった後の第2検出情報DI2を入力する。時点t9において、当該情報処理手段は、時点t8から第3スイッチオフ期間TA3にわたって第1検出情報DI1が入力されず、従って第3スイッチオフ規準が満たされたことを確定する。続いて、該時点t9において、当該情報処理手段は音声期間TS4を持つ第4音声タイムスロットの終了を決定する。この第4音声タイムスロットを特徴付けるため、第4の音声検出情報SDI4が記憶手段7に供給される。

40

【0040】

このようにして、該第2実施例による情報処理手段により第3スイッチオフ規準が判定されるという事実は、ノイズ信号RSのみ又は悪い信号対雑音比を持つ第3音声信号SS3のみを含むマイクロフォン信号MSの受信データEDが音声認識手段9に供給されず、これにより該音声認識手段9による誤った文の認識が防止されるという利点を提供する。

【0041】

音声検出情報SDIが前記スイッチオン閾検出器及びスイッチオフ閾検出器に供給されるようにすることが考えられる。その場合、これら閾検出器は、休止タイムスロットTPにおけるエネルギー情報EIを評価して、休止タイムスロットTPの間においてマイクロフォ

50

ン信号MSに含まれるノイズ信号RSのエネルギー量に対する第1及び第2エネルギー閾値を適応化するように構成することができる。

【0042】

これは、当該音声検出装置が、例えば大きな背景雑音の結果として口述の間にノイズ信号RSのエネルギー量が増加するような場合にも、良好な信号対雑音比を持つ音声信号SSのみを検出するようになるという利点を提供することができる。

【0043】

本発明による音声検出装置には、アナログ信号を処理する手段を設けることもできることが分かる。その場合、前記エネルギー決定手段はアナログ受信信号を二乗すると共に該信号を前記評価タイムスロットにわたり積分し、且つ、斯様にして決定されたアナログエネルギー信号を2つの比較器に供給する。この場合、これら比較器は上記スイッチオン閾検出器及びスイッチオフ閾検出器を構成することになる。

10

【0044】

本発明による音声検出装置は、マイクロフォン信号を磁気テープカセット又はハードディスク上に記録する口述記録器（ディクテーションマシン）に組み込み、口述記録の音声により制御される自動起動及び自動停止を可能にするようにすることも考えられる。

【0045】

また、本発明による音声検出装置は、音声入力により起動及び停止される他のマシンに設けることも考えられる。斯様なマシンは、例えば、携帯電話である。

【図面の簡単な説明】

20

【図1】 図1は、マイクロフォン及びモニタが接続されると共に音声認識ソフトウェアが実行されるようなコンピュータであって、該コンピュータが音声検出装置も形成するようなコンピュータをブロック図として示す。

【図2A】 図2Aは、上記音声認識ソフトウェアが第1及び第2実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

【図2B】 図2Bは、上記音声認識ソフトウェアが第1及び第2実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

【図2C】 図2Cは、上記音声認識ソフトウェアが第1及び第2実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

30

【図2D】 図2Dは、上記音声認識ソフトウェアが第1及び第2実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

【図2E】 図2Eは、上記音声認識ソフトウェアが第1実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

【図2F】 図2Fは、上記音声認識ソフトウェアが第2実施例に従い実行される場合に上記コンピュータにおいて発生する信号及び情報の波形を時間の関数として示す。

【 図 1 】

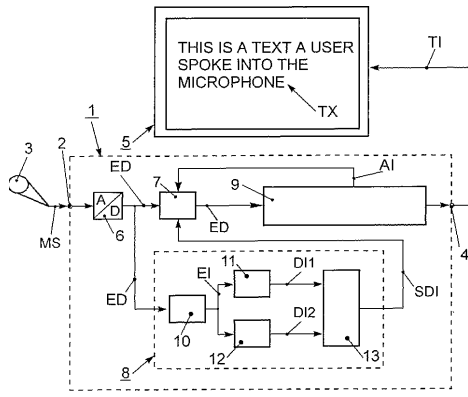


FIG. 1

【 図 2 A - F 】

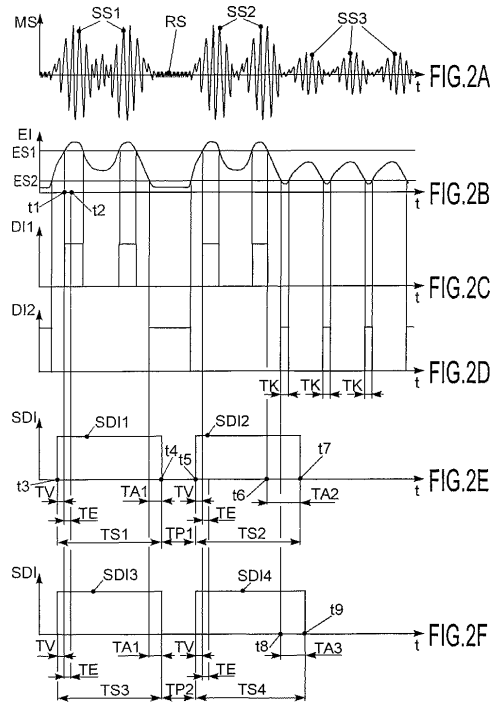


FIG.2A

FIG.2B

FIG.2C

FIG.2D

FIG.2E

FIG.2F

フロントページの続き

- (56)参考文献 特開平 0 3 - 1 8 2 7 9 9 (J P , A)
特開昭 6 3 - 2 2 6 6 9 8 (J P , A)
特開平 1 1 - 0 9 5 7 8 5 (J P , A)
特開昭 6 3 - 2 9 8 2 9 8 (J P , A)
特開昭 6 2 - 2 3 7 4 9 8 (J P , A)
特開昭 6 1 - 2 5 9 2 9 6 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G10L 11/02,15/04
JSTPlus(JDreamII)