

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6061069号
(P6061069)

(45) 発行日 平成29年1月18日(2017.1.18)

(24) 登録日 平成28年12月22日(2016.12.22)

(51) Int.Cl. F I
G 1 0 L 25/84 (2013.01) G 1 0 L 25/84

請求項の数 4 (全 8 頁)

<p>(21) 出願番号 特願2012-207909 (P2012-207909) (22) 出願日 平成24年9月21日 (2012.9.21) (65) 公開番号 特開2014-63018 (P2014-63018A) (43) 公開日 平成26年4月10日 (2014.4.10) 審査請求日 平成27年9月7日 (2015.9.7)</p>	<p>(73) 特許権者 396020132 株式会社システック 静岡県浜松市北区新都田1-9-9 (72) 発明者 香高 孝之 静岡県浜松市北区新都田一丁目9番9号 株式会社システック内 審査官 安田 勇太</p>
--	---

最終頁に続く

(54) 【発明の名称】 極微小音声入力装置

(57) 【特許請求の範囲】

【請求項1】

多数の話者が音声入力を行っても他人に聞こえず、他人の入力の妨げにならない音声入力装置であって、

呼吸音より小さい音圧の話者が発する極微小音声を含む音声を取得するために、口先又は口腔内又は口の周りの閉じた閉区間に設置されるマイクロフォンと前記マイクロフォンで取得した前記音声をアナログデジタル変換するAD変換器と、前記AD変換器の出力する音声データを記憶するメモリと、前記呼吸音と前記極微小音声の間に設定した比較判別のための音圧閾値と比較することで、前記音声の音圧が前記音圧閾値より大きい場合は、前記音声が、前記呼吸音であるとして除去し、前記音声の音圧が前記音圧閾値より小さい場合は、前記音声が前記極微小音声の発話であるとして残す呼吸音除去手段と、前記AD変換器と前記メモリと前記呼吸音除去手段を制御する中央演算ユニットとを有することを特徴とする極微小音声入力装置。

【請求項2】

前記呼吸音除去手段は、前記音声の音圧が前記音圧閾値より大きい場合には、前記呼吸音と話者が通常発する通常発話の音圧である通常音圧(60dB)の間に設定した比較判別のための第2の音圧閾値と比較することで、前記音声の音圧が前記第2の音圧閾値より小さい場合は、前記音声が、前記呼吸音であるとして除去し、前記音声の音圧が前記第2の音圧閾値より大きい場合は、前記音声が前記通常発話であるとして除去し、前記極微小音声での再度の音声入力を指示する判断を行うことを特徴とする請求項1記載の極微小音声

入力装置。

【請求項 3】

前記呼吸音除去手段、又は前記中央演算ユニットは、前記メモリから前記音声データを前記呼吸音の呼気又は吸気の 1 つの時間区間の単位より小さい時間間隔で集団として前記音声データを取得し、二つの隣り合う集団が呼気又は吸気、又は発話として連続している場合には同一集団とし、1 つの集団内で連続しない場合は異なる集団に分け、同一集団ごとに同じ処理を行うことを特徴とする請求項 1 又は請求項 2 記載の極微小音声入力装置。

【請求項 4】

前記話者が単独での前記通常発話を行う場合と前記極微小音声での発話を行う場合の両者の間の切り替えを行うために、マイクロフォンの感度を低く、前記呼吸音除去手段が作動しない前記通常発話対応モードとマイクロフォンの感度が高く、前記呼吸音除去手段が作動する前記前記極微小音声発話モードの間で切り替えを行う入力モード切替手段を有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 1 つに記載の極微小音声入力装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、キーボードなどでのデータ入力の代わりに音声で入力する装置に関するものである。

【背景技術】

【0002】

従来、音声入力する場合は、マイクロフォンから発話音を入力するが、通常行われているように、発話音は、通常耳にする音圧レベル（約 60 dB）のものであり、周辺からの雑音の音圧レベルはこれより小さいので問題がない。しかし、一人で個室での作業では問題にならないが、多数の作業者が同室で発話するのは、極めてうるさく、他人の発話は雑音になりやすい。そのために、発話音の音圧を下げる試みが必要となる。発話音の音圧を下げると、今まで気にならなかった呼吸音と同程度の音圧に近づく。呼吸音は、呼吸している本人が気にして聞こうとして聞かないと感じない程度なので、少し離れた（隣の）他人には聞こえない。このような状態では、他人からの影響を受けずに音声入力を行えるが、発話音の音圧と同程度の呼吸音がマイクロフォンで拾われるので、呼吸音の判別と除去処理が必要になる。音圧が同程度の場合は、音圧で分けることができないので、特別な手段が必要となる。呼吸音を除去する目的ではないが、非発話音と発話音を区別するための方法として、フラクタル次元数を計算する手法が特許文献 1 に示されている。又、周辺からの雑音と発話音を区別して、周辺の雑音を除く目的でカオス分析し、リアプノフ指数を算出する装置が特許文献 2 に示されている。両者とも数学的理論に基づき計算処理を行うが、計算処理が重いのが欠点となっている。

尚、上記の音声入力は、通常考えられるようにマイクロフォンを口先または特別の場合は、口腔内において行うものである。これとは全く別の分野になるが、一般に NAM マイクロフォンと呼ばれるものがある。これは、喉の奥で生ずる音を耳下の皮膚に接触したマイクロフォンで捉えるものであるが、喉から身体の骨、肉を伝達してくる音であり、更に未だ

口腔や鼻などでも共振を受けてない音であるので、これらの微小音は、本願の対象範囲のものではない。

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2007 - 264567

【特許文献 2】特開 2003 - 99094

【発明の概要】

【発明が解決しようとする課題】

【0004】

10

20

30

40

50

本発明の課題は、今まで必要な信号より雑音が多いとして注目されず使われることが無かった発話音の音圧が呼吸音のそれより小さい領域（極微小音領域）での簡単な呼吸音判別と除去の手段を備えた音声入力装置を提供することである。

【課題を解決するための手段】

【0005】

本発明の極微小音声入力装置は、呼吸音とこれより音圧の小さい極微小音声の間に設定した音圧閾値を基に音声は呼吸音であるか極微小音声であるかを判別し、呼吸音を除去するものであり以下詳述する。

【0006】

請求項1記載の発明は、極微小音声入力装置であって、多数の話者が音声入力を行っても他人に聞こえず、他人の入力の妨げにならない音声入力装置であって、呼吸音より小さい音圧の話者が発する極微小音声を含む音声を取得するために、口先又は口腔内又は口の周りの閉じた閉区間に設置されるマイクロフォンと前記マイクロフォンで取得した前記音声をアナログデジタル変換するAD変換器と、前記AD変換器の出力する音声データを記憶するメモリと、前記呼吸音と前記極微小音声の間に設定した比較判別のための音圧閾値と比較することで、前記音声の音圧が前記音圧閾値より大きい場合は、前記音声は、前記呼吸音であるとして除去し、前記音声の音圧が前記音圧閾値より小さい場合は、前記音声は前記極微小音声の発話であるとして残す呼吸音除去手段と、前記AD変換器と前記メモリと前記呼吸音除去手段を制御する中央演算ユニットとを有することを特徴とする。

10

20

【0007】

請求項2記載の発明は、請求項1記載の極微小音声入力装置において、前記呼吸音除去手段は、前記音声の音圧が前記音圧閾値より大きい場合には、前記呼吸音と話者が通常発する通常発話の音圧である通常音圧（60dB）の間に設定した比較判別のための第2の音圧閾値と比較することで、前記音声の音圧が前記第2の音圧閾値より小さい場合は、前記音声は、前記呼吸音であるとして除去し、前記音声の音圧が前記第2の音圧閾値より大きい場合は、前記音声は前記通常発話であるとして除去し、前記極微小音声での再度の音声入力を指示する判断を行うことを特徴とする。

【0008】

請求項3記載の発明は、請求項1又は請求項2記載の極微小音声入力装置において、前記呼吸音除去手段、又は前記中央演算ユニットは、前記メモリから前記音声データを前記呼吸音の呼気又は吸気の1つの時間区間の単位より小さい時間間隔で集団として前記音声データを取得し、二つの隣り合う集団が呼気又は吸気、又は発話として連続している場合には同一集団とし、1つの集団内で連続しない場合は異なる集団に分け、同一集団ごとに同じ処理を行うことを特徴とする。

30

【0009】

請求項4記載の発明は、請求項1から請求項3のいずれか1つに記載の極微小音声入力装置において、前記話者が単独での前記通常発話を行う場合と前記極微小音声での発話を行う場合の両者の間の切り替えを行うために、マイクロフォンの感度を低く、前記呼吸音除去手段が作動しない前記通常発話対応モードとマイクロフォンの感度が高く、前記呼吸音除去手段が作動する前記前記極微小音声発話モードの間で切り替えを行う入力モード切替手段を有することを特徴とする。

40

【発明の効果】

【0010】

以上のように構成されているので本発明の極微小音声入力装置では、多数の作業者が音声入力を行っても他人に聞こえず、他人の入力の雑音にならず、呼吸音を除去して、モニタするに好都合であり、データ処理に好都合な処理が可能となる。

【図面の簡単な説明】

【0011】

50

【図1】本発明の極微小音声入力装置の構成と処理フローの一実施態様を示す図である。

【図2】音声入力の発話音と呼吸音の音圧レベルを示す図である。

【図3】本発明の極微小音声入力装置に入力する音声入力と呼吸音除去後の発話音圧（強度）の時間推移を示す図である。

【発明を実施するための形態】

【0012】

本発明の極微小音声入力装置は、多数の話者が音声入力を行っても他人に聞こえず、他人の入力の妨げにならない音声入力装置であって、

呼吸音より小さい音圧の話者が発する極微小音声を含む音声を取得するために、口先又は口腔内又は口の周りの閉じた閉区間に設置されるマイクロフォンとマイクロフォンで取得した音声をアナログデジタル変換するAD変換器と、AD変換器の出力する音声データを記憶するメモリと、呼吸音と極微小音声の間に設定した比較判別のための音圧閾値と比較することで、音声の音圧が音圧閾値より大きい場合は、音声は呼吸音であるとして除去し、音声の音圧が音圧閾値より小さい場合は、音声は極微小音声の発話であるとして残す呼吸音除去手段と、AD変換器とメモリと呼吸音除去手段を制御する中央演算ユニットとを有する。以下、実施例に沿って説明する。

【0013】

図2は、音声入力の発話音と呼吸音の音圧レベルを示す図である。マイクロフォン（特に都合がよいのは、単一指向性のマイクロフォン）を話者の口先又は特別の場合は、口腔内、又は口の周りに閉じた閉区間に設置して発話する。2-Bに示すように、通常会話の発話音圧は60dB程度で行われている。これに対して、人間が呼吸をしているときの呼吸音圧は、10～20dBとされている。通常会話の発話音圧では、呼吸音圧が極めて小さいので気にならず、又、音圧での区別が容易である。

通常会話の音圧から、どんどん音圧を下げてゆく（小さい声にしてゆく）と、微小発話音として示したように発話音圧が呼吸音とほぼ同じかやや大きい程度になり、音圧のみでの両者の区別は難しくなる。ここでの区別は、カオス解析などの数学的アプローチの手段が従来行われるが、計算処理が複雑で重たい処理であることは、従来例で述べた。

処理を簡単にする手段として、発話音と呼吸音の空気の流れの揺らぎを含む特性の違いに基づいて区別することが提案できる。

【0014】

この微小発話音より更に音圧を下げていった場合、極微小発話音と記述したように呼吸音圧より小さいので、当然、隣にいる他人の耳には聞こえない。且つ、呼吸音よりは小さいが、マイクロフォンには発話音を取得できる。しかし、所望の発話音より呼吸音が大きいので、自分の入力した内容をモニタする場合、呼吸音が聞くに堪えない大きさである。データ処理としても呼吸音は不要なので除去したい。この音圧領域は通常は常識外の範囲なので使用されることが無かったが、多人数で行っても他者の迷惑にならない静かな音声入力にこだわると、他者に聞こえないという利点を享受できる。しかしながら弱点もある。たまたま、他人の大きな発話などが入った場合は、これも除去する必要がある。

本願は、極微小音発話を対象にして音声入力を行う装置である。

2-Aには上に述べたことが表にして示している。

【0015】

図3は、本発明の極微小音声入力装置に入力する音声入力と呼吸音除去後の発話音圧（強度）の時間推移を示す図である。3-A、3-Bでは、横軸に時間をとって、縦軸に音声の強度変化を概念的（実際の波形ではなく、波形の存在する範囲を枠で）に示している。3-Aは、音声入力後、3-Bは望ましい呼吸音除去が行われた後に対応している。3-Aでは、発話音の強度より大きな強度の吸気と呼気がある、発話がない場合は、吸気と呼気を繰り返す。発話がある場合は、通常は、呼気のところに発話がなされ、呼気の場合より時間幅が長く発話がなされる。発話は話す言葉の内容により長さが異なり、息継ぎをしてから追加発話がなされる場合も多いことは周知のことである。

【0016】

図1は、本発明の極微小音声入力装置の構成と処理フローの一実施態様を示す図である。
 1 - Aには、ハードウェア構成、1 - Bには、呼吸音除去手段の処理フローの一例を示す。
 1 - Aにおいて、マイクロフォン101とADC（アナログデジタル変換器）102とCPU（中央演算処理ユニット）103とメモリ104と呼吸音除去手段105を有している。尚、呼吸音除去手段105は、判断・処理を行うので、ソフトウェアで処理する場合には、CPU103、メモリ104（ハードディスクを含んでもよい）に内在するという見方も可能である。

【0017】

マイクロフォン101は、呼吸音より小さい極微小音声を取得するので、口先又は特別の場合は口腔内又は、口の周りに閉じた閉区間に設置される。又、周囲からの音を拾わないように単一指向性のマイクロフォンが好都合である。口腔内又は、口の周りに閉じた閉区間では単一指向性のマイクロフォンでなくとも使用できる。

マイクロフォン101で取得したアナログの音声データは、ADC102に入力し、デジタル量に変換され、メモリ104に記憶される。CPU103は、ADC102の動作とメモリ104の記憶動作を制御する。呼吸音除去手段105は、メモリ104から所望の時間間隔ごとのデータを読み取り、判断と処理を行って、呼吸音を除去したデータをメモリ104（他のメモリでもよい）又は、記述されていない他所へデータを送付する。

【0018】

データの読み取り時間間隔は、呼吸の時間間隔（1～2秒程度）を目安にそれより少ない時間間隔であることが好ましい。例えば、可聴音域（～20kHz）に対してサンプリング定理を満足するように、ADCでのデータ取得は、可聴音域（～20kHz）に対してサンプリング定理を満足するように例えば、50kHz等の可聴音域の2倍以上の取得周波数（サンプリング周波数）で行うが、1つの吸気、呼気、発話の集団を特定するには、取得したデータを1～2秒程度又はそれより小さい読み取り時間間隔で読み取り集団付けることを行う。時間間隔ごとに読まれたデータは、連続する音圧時間区間を、必要に応じて同一の音に属するとして集団分けし、同一処理に委ねるということである。なぜなら、呼吸音も吸気、呼気は、その間は無音区間で隔てられ、それぞれ一つの集団と扱ってよい。又、発話音は、呼気の区間に発せられ、複数の語が連続しているか、又は、1語1語の間が無音区間で隔てられているかのいずれかである。連続している場合は、その区間は、同じ集団として同じ処理が可能である。又、1つの集団内で連続しない場合は、異なる集団に分けることも可能である。

【0019】

尚、集団分けや処理の判断には、以下の発話の特性を考慮して行うことも可能である。

< 音声の特性 >

・ 発話は呼気（空気の流れは外向き）の始めからが多い。呼気の途中から始まるのは意図的な場合以外は極めて少ない。

呼気の最初のパターンが現われたら、そこから呼気と次の吸気の区間までは、呼吸音と判断できる。呼気の最初のパターンが現われずに発話になった場合は、発話区間になる。

- ・ 吸気（空気の流れは内向き）では発話しない。
- ・ 呼吸音は、空気の流れが1方向であり、発話に比べ揺らぎが無い。
- ・ 呼吸音の音圧は、同一個人ではいつも同じような呼吸をしていることで、発話音に比べてばらつきが少ない。
- ・ 鼻からの呼気では発話しない。
- ・ 発話音は500Hz以上の成分が多い。呼吸音は500Hz以下の成分が多い。

【0020】

呼吸音除去手段105の動作の例を、1 - Bで説明する。

マイクロフォン101から発話入力となされる。発話入力の後には、メモリ104へのデータの格納を介して、又は直接に、1) 音圧が閾値1より大きいかを判断する。閾値1は、呼吸音と極微小音の間の音圧値に設定される。例えば、図2の2 - Bを参照すると、10dB周辺が適当である。この値は、図3の音声入力の測定値を実際に測定し、呼吸音と

10

20

30

40

50

発話音の間の適度な値に設定してもよい。音圧が閾値 1 より小さい場合は、極微小発話音と判断してデータを除去せず存続する。音圧が閾値 1 より大きい場合は、2) 音圧が閾値 2 より大きいかを判断する。閾値 2 は、呼吸音と通常発話音の間の音圧値に設定される。例えば 30 dB 等が適当である。

音圧が閾値 2 より小さい場合は、呼吸音と判断して除去され、この時間区間は、無音化される。音圧が閾値 2 より大きい場合は、通常発話音が混入したと判断し除去され、且つ、「適正レベルで再度入力してください」という指示を出す。時間の進行に沿って、音声入力に対してこの判断処理を繰り返し行う。

【0021】

尚、音圧が閾値 2 より大きいかの判断以降の処理は、必須ではないが、現実の中では、うっかり大きな声で入力してしまったり、他人の大きな声を拾ってしまう場合があるので、このような判断処理があると都合がよい。又、多人数の入力では、このような呼吸音除去手段を稼働させるが、部屋内で単独の作業の場合には、通常発話の音圧で入力してもよいので、この場合は、呼吸音除去手段の稼働・不稼働は特に気にせずに、マイクロフォンの感度を下げて入力すればよいので、両者の間の切り替えを行う入力モード切替手段を備えると好都合である。

以上のような処理によって、図 3 の 3 - A の音声入力は、3 - B の呼吸音除去後の音声データを得ることができる。

【産業上の利用可能性】

【0022】

以上のように構成されているため、本発明の極微小音声入力装置は、多数の作業者が音声入力を行っても他人聞こえず、他人の入力の雑音にならず、呼吸音を除去して、モニタするに好都合であり、データ処理に提供されるので、産業上利用性が極めて大きい。

【符号の説明】

【0023】

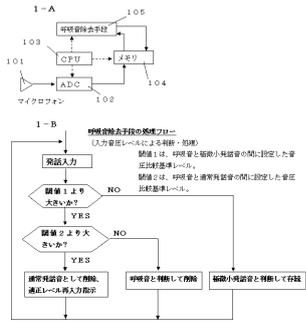
- 101 マイクロフォン
- 102 ADC
- 103 CPU
- 104 メモリ
- 105 呼吸音除去手段

10

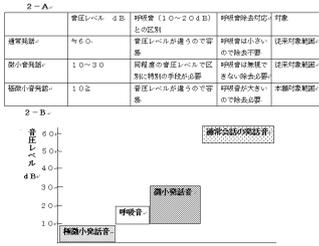
20

30

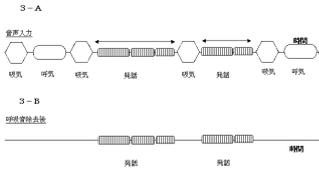
【図1】



【図2】



【図3】



フロントページの続き

- (56)参考文献 特開2007-094002(JP,A)
登録実用新案第3069249(JP,U)
特開2000-276191(JP,A)
特開平7-312795(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 25/84

G10L 15/04