

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6519102号
(P6519102)

(45) 発行日 令和1年5月29日(2019.5.29)

(24) 登録日 令和1年5月10日(2019.5.10)

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

請求項の数 1 (全 11 頁)

(21) 出願番号	特願2014-114463 (P2014-114463)	(73) 特許権者	396020132
(22) 出願日	平成26年6月3日(2014.6.3)		株式会社システック
(65) 公開番号	特開2015-228002 (P2015-228002A)		静岡県浜松市北区新都田 1-9-9
(43) 公開日	平成27年12月17日(2015.12.17)	(72) 発明者	香高 孝之
審査請求日	平成29年5月12日(2017.5.12)		静岡県浜松市北区新都田一丁目9番9号
前置審査			株式会社システック内
		審査官	菊池 智紀

最終頁に続く

(54) 【発明の名称】 複数者間干渉音声雑音を回避した音声入力処理装置

(57) 【特許請求の範囲】

【請求項 1】

音を入力するマイクロフォンと、前記マイクロフォンに入力される音声の発話音の音圧が呼吸音と同程度または、それより小さい音圧領域において、前記マイクロフォンに入力された音声データ内で、前記呼吸音か前記発話音かの音種を判別する呼吸音・発話音判別手段と、前記判別に基づき、前記発話音を選択する音声信号選択手段と、前記音声信号選択手段の電気信号出力を受けて、所望の音声信号処理を行う音声信号処理手段と、これらを制御する制御手段とを有し、

前記マイクロフォンにおいて、前記発話音の音圧が前記呼吸音と同程度または、それより小さい音圧で前記音声入力をする使用条件の場合に、前記呼吸音・発話音判別手段は、前記音声データにおいて、無音区間を基準にして前記無音区間の前後に前記呼吸音又は前記発話音を音種区間として定める手段と、前記音種区間の前記音種を区別するための音圧閾値との比較処理以外の手段を備えることで、前記呼吸音か前記発話音を判別し、同じ音種の連続した複数のデータ集団を特定し、もって、前記音声信号選択手段が、前記データ集団ごとの音種を識別する音種識別手段と、前記音種識別に基づいて、前記発話音を選択する信号選択器を有することで、前記発話音を選択して、前記発話音を前記音声信号処理手段が処理を行うものであり、前記音種識別手段は、前記データ集団に識別可能な音種識別符号を付与する音種識別符号付与手段と、さらに、前記データ集団と音種識別符号を関連づけて格納するためのメモリとを備えているか、又は、前記データ集団は音種が識別可能なように音種ごとに区別して格納されるメモリを備えていることを特徴とする複数者間干

10

20

渉音声雑音を回避した音声入力処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、互いの音声を聞きあえる環境において複数者が音声入力処理装置を使用する場合に、発話音が相互に対して騒音となったり、相互の音声入力装置に意図しない騒音となって入力したりすることを回避する手段を備えた複数者間干渉音声雑音を回避した音声入力処理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60dB付近を中心とする）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。

また、携帯電話などでは、多数の人のいる空間（たとえば、列車内や会議場など）では、うるさいので、使用しないように電源を切っておくか、マナーとして、その空間を出て使用するようになっている。また、相互の発話音が、互いの入力装置に入ってしまう意図しない情報漏えいや騒音になる危険もある。

音声入力による文字や符号、図形への変換装置においても、隣同士の音声の混入は不要な雑音になって、変換結果に誤りが生じ、又、情報漏えいに繋がり、このような場合には、オフィスのような複数者が同居する空間では不都合な状態となる。

このような不都合は避ける必要があった。以下、上記のように、複数者による音声相互への騒音となり、騒音を回避する必要がある音声処理を複数者間干渉音声雑音回避音声入力処理と呼ぶことにする。すなわち、携帯電話や音声入力により文字変換装置のような複数者が使用する環境では、音声入力処理は、このような処理を行わないと上記の不都合が生じて使えない環境として、使われないか、マナーとして使用が規制される。

音声処理装置は、電話機能や音声文字変換機能自体は、機能として有益なものであるが、複数話者の間では、騒音となって使いにくいものであった。しかしながら、今後、これらの装置を複数話者の間で騒音とならずに規制なく自由に使用できるようにすることは、極めて歓迎すべきことである。

【0003】

このような中で、特許文献1には、非発話音と発話音を区別するためにフラクタル次元数（数学的理論に基づいた計算処理を行う）を計算する手法が示されている。非発話を特定して、その結果、非発話音中の溜息を判別し、これを検出して運転士の集中力の低下を検出するものであるが、微小音や極微小音の発話を対象にしたものではない。

又、本出願人が先に出願した特許文献2のものがある。

特許文献2において、発明者は極微小音領域において、音圧の閾値により、呼吸音を除去する装置を提案した。しかしながら、閾値によっては、極微小音領域より音圧の大きな呼吸音が減衰するその裾野に極微小音領域の音圧と同じ程度の呼吸音領域が残存し、呼吸音の完全な除去は難しいことが分かった。かつ、残存部がその後の処理に影響を与える欠点があることが分かった。その後の処理に利用するための発話音と呼吸音の識別も行っていないので、その後の処理は両者を区別して行うことができない欠点がある。

呼吸音域が除かれた時間域は無音区間になり、特に音種の識別がないので、データは順番に格納され、メモリの容量の低減を図ることはできない欠点がある。

そして、その後の処理に利用するための発話音と呼吸音の識別も行っていない点は、特許文献1においても同様な欠点となっている。

【先行技術文献】

【特許文献】

【0004】

10

20

30

40

50

【特許文献 1】特開 2 0 0 7 - 2 6 4 5 6 7

【特許文献 2】特願 2 0 1 2 - 2 0 7 9 0 9

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 5 】

本発明の課題は、複数者間干渉音声雑音を回避した音声入力処理装置において、複数者間の相互への干渉音声雑音を回避した音声入力処理装置を提供する。

【課題を解決するための手段】

【 0 0 0 6 】

複数者間干渉音声雑音を回避した音声入力処理装置であって、

音を入力するマイクロフォンと、マイクロフォンに入力される音声の発話音の音圧が呼吸音と同程度または、それより小さい音圧領域において、マイクロフォンに入力された音声データ内で、無音区間を基準にして前記無音区間の前後に前記呼吸音又は前記発話音を音種区間とすることで、呼吸音が発話音かの音種を判別する呼吸音・発話音判別手段と、判別に基づき、発話音を選択する音声信号選択手段と、音声信号選択手段の電気信号出力を受けて、所望の音声信号処理を行う音声信号処理手段と、これらを制御する制御手段とを有する。

以下詳述する。

【 0 0 0 7 】

請求項 1 記載の発明は、複数者間干渉音声雑音を回避した音声入力処理装置であって、音を入力するマイクロフォンと、前記マイクロフォンに入力される音声の発話音の音圧が呼吸音と同程度または、それより小さい音圧領域において、前記マイクロフォンに入力された音声データ内で、前記呼吸音が前記発話音かの音種を判別する呼吸音・発話音判別手段と、前記判別に基づき、前記発話音を選択する音声信号選択手段と、前記音声信号選択手段の電気信号出力を受けて、所望の音声信号処理を行う音声信号処理手段と、これらを制御する制御手段とを有し、

前記マイクロフォンにおいて、前記発話音の音圧が前記呼吸音と同程度または、それより小さい音圧で前記音声入力をする使用条件の場合に、前記呼吸音・発話音判別手段は、前記音声データにおいて、無音区間を基準にして前記無音区間の前後に前記呼吸音又は前記発話音を音種区間として定める手段と、前記音種区間の前記音種を区別するための音圧閾値との比較処理以外の手段を備えることで、前記呼吸音が前記発話音を判別し、同じ音種の連続した複数のデータ集団を特定し、もって、前記音声信号選択手段が、前記データ集団ごとの音種を識別する音種識別手段と、前記音種識別に基づいて、前記発話音を選択する信号選択器を有することで、前記発話音を選択して、前記発話音を前記音声信号処理手段が処理を行うものであり、前記音種識別手段は、前記データ集団に識別可能な音種識別符号を付与する音種識別符号付与手段と、さらに、前記データ集団と音種識別符号を関連づけて格納するためのメモリとを備えているか、又は、前記データ集団は音種が識別可能なように音種ごとに区別して格納されるメモリを備えていることを特徴とする。

【発明の効果】

【 0 0 0 8 】

以上のように構成されているので本発明の複数者間干渉音声雑音を回避した音声入力処理装置では、

発話音の音圧が呼吸音と同程度または、それより小さい領域において、音声入力を行う使用状態で、無音区間を基準にして前記無音区間の前後に前記呼吸音又は前記発話音を音種区間とすることで、呼吸音と必要な発話音を判別、発話音と呼吸音のいずれか（普通は発話音）を選択してこれにより音声入力処理を行わせたので、他人の発話音に影響されず自らの発話音又は呼吸音のみで音声信号処理手段の本来の処理が可能となり、これらが複数話者の環境で使用可能となる。

【図面の簡単な説明】

【 0 0 0 9 】

10

20

30

40

50

【図 1】本発明の複数者間干渉音声雑音を回避した音声入力処理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】本発明の複数者間干渉音声雑音を回避した音声入力処理装置に入力する音声入力と呼吸音・発話音判別、発話音選択、選択的発話音処理を示す図である。

【図 4】本発明の複数者間干渉音声雑音を回避した音声入力処理装置に使用する音声信号選択手段の他の実施態様と処理フローの一実施態様を示す図である。

【図 5】本発明の複数者間干渉音声雑音を回避した音声入力処理装置に使用される音種識別手段の具体的実施態様を示す図である。

【図 6】本発明の複数者間干渉音声雑音を回避した音声入力処理装置に入力する音声入力
と呼吸音・発話音判別、識別符号付け、選択的発話音処理を示す図である。 10

【発明を実施するための形態】

【0010】

複数者間干渉音声雑音を回避した音声入力処理装置として、必須なことは、発話音の音圧を下げる試みが必要となることである。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、又はそれ以下の音圧に近づく。

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。このような状態の音圧レベルならば、他人からの干渉を受けず又は他人に干渉（うるさがられる）を与えずに音声入力を行える。

雑音とならない代わりに、発話音の音圧を小さい状態で入力したのであるから、これと同程度又はそれより大きな呼吸音がマイクロフォンで拾われる。装置の中では発話音のみ（呼吸音は医療機器の場合のみあり）を処理或いは、増幅して処理しなければならない。そのためには、呼吸音と発話音の判別が必要となる。呼吸音と発話音の判別の方法は、多くの手段があるが、一例では、音圧の閾値で判別する例もあるが、特許文献 1 等で知られている非発話音と発話音を区別するためにフラクタル次元数（数学的理論に基づた計算処理を行う）を計算する手法や、より簡単には、音区間を基準にして前記無音区間の前後に前記呼吸音又は前記発話音を音種区間とし、後述するように呼吸音と発話音の特徴差による判別が可能である。呼吸音と発話音の判別に基づいて発話音のみ選択的に処理を行うことで、従来、複数話者の間では相互干渉雑音となるため使用できなかった装置の機能が複数話者の環境でもできるということが新たに実現可能となった。因みに、特許文献 1 には、
非発話音と発話音を区別するためにフラクタル次元数（数学的理論に基づた計算処理を行う）を計算する手法が示されている。特許文献 1 は、非発話を特定して、その結果、非発話音中の溜息を判別し、これを検出して運転士の集中力の低下を検出するものである。
以下、本発明の複数者間干渉音声雑音を回避した音声入力処理装置を実施例に沿って説明する。 20 30

【0011】

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 60 dB 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 ~ 20 dB と言われている。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。 40

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、他人の耳には聞こえない。そして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。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。

【0012】

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 50

圧より小さいので、ここでも当然、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が取得できる。しかし、所望の発話音より呼吸音が大きいため、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必要とする場合は少ない。これら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、これが邪魔になる。

本願は、微小発話音、極微小発話音を対象にして音声入力を行う使用条件の場合に、呼吸音と発話音を判別し、発話音を選択して処理する複数者間干渉音声雑音を回避した音声入力処理装置である。 10

2 - A には上に述べたことが表にして示している。

【 0 0 1 3 】

図 3 は、本発明の複数者間干渉音声雑音を回避した音声入力処理装置に入力する音声入力と呼吸音・発話音判別、発話音選択、音声信号処理手段への発話音の増幅入力の例を示す図である。

横軸に時間をとって、縦軸に音声の強度（音圧）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、音声入力した音の波形を示す。波形を示すデータは、図 1 の呼吸音・発話音判別手段において、吸気音、呼気音、発話音に判別される。波形を示すデータを判別してゆくには、マイクロフォンからのアナログデータを直接処理する場合と、一旦、A/D 変換器で処理して、デジタルデータにしたのち処理、又は、一旦メモリに記憶したあとに処理など種々の形態がとれる。データとしては、複数のデータを処理してゆくと、無音区間、または、吸気音、呼気音、発話音のうち、他の種別の音に変わる（特に音種の変化の間には無音区間を挟むことが多い）ので、一つの種別の音域の始端データと終端データが検出されるので、始端データと終端データの間のデータ集団が、一つの種別の音域（音種集団）となる。 20

3 - B は、判別された音種により、発話音のみ選択したものである。

更に、3 - C では、発話音のみ選択した中で、発話音のみ増幅（音圧を大きくした）したものである。 30

このようにして、他人への雑音とならない微小音、極微小音でマイクロフォンから音声入力する使用条件において、呼吸音が邪魔になる小さな発話音だったにも関わらず、装置内では、発話音のみが選択され（増幅されて）、音声信号処理手段に入力され、その処理が行われる。

しかして、複数話者のいる環境でも、携帯電話や、音声文字変換装置などの複数者間干渉音声雑音が気になっていた音声信号処理手段の本来の機能が使用制限無く実現される。

【 0 0 1 4 】

図 1 は、本発明の複数者間干渉音声雑音を回避した音声入力処理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、複数者間干渉音声雑音を回避した音声入力処理装置の構成、1 - B には、その処理フローの一例を示す。 40

1 - A には、マイクロフォン 110 と呼吸音・発話音判別手段 120 と音声信号選択手段 130 と音声信号処理手段 140 とこれらを制御する制御手段 150 を示している。

マイクロフォン 110 で音を拾う。音データは、呼吸音・発話音判別手段 120 において、呼吸音（呼気音、吸気音）と発話音に判別される。勿論、この音の間には無音区間があることは当然である。呼吸音・発話音判別手段 120 において、入力されたアナログの音データを直接に判別してもよいが、一旦、A/D 変換器 122 にて A/D 変換し、これをデジタル処理で判断するか、一旦、一時メモリ 123 に格納した後、読みだして判別してもよい。アナログ処理では、A/D 変換器 122 と一時メモリ 123 は、処理の前には必要ではないことは当然である。その代り、アナログ処理後に A/D 変換することが必要になる。 50

一般的には、判別手段 1 2 1 は、最初に A D 変換器 1 2 2 と一時メモリ 1 2 3 を介してからデジタル処理で判別する方が処理が容易である。

音圧の強度から閾値で判別する場合は、アナログ処理でも可能だが、これも含め、他の判別ではデジタル処理の方が容易である。

判別手段 1 2 1 をデジタル処理で行う構成の例としては、音圧の閾値（この場合はデジタル値）での判別以外の手段では、以下のようなことが可能であり、特許文献 1 等に表示されている。これは、フラクタル次元の分析などを行うものである。尚、フーリエ周波数スペクトル分析（F F T 処理）も利用できる。簡単には後述する呼吸音と発話音の特徴差を利用する。

【 0 0 1 5 】

呼吸音・発話音判別手段 1 2 0 において、データ集団の音種が判別されると、音声信号選択手段 1 3 0 において、発話音が選択され、そのまま、或いは増幅された形で、データを出力する。出力先は、複数者間干渉音声雑音を回避したい音声信号処理手段 1 4 0 であり、ここでその装置本来の処理がなされる。この図には、音声信号処理手段 1 4 0 の実例として、携帯電話装置または、音声入力文字・符号・図形変換装置が示されている。勿論、これらへの入力は、音声信号を電気信号の形で受け渡しているのであって、音声で受け渡しているわけではないことは当然である。

制御手段 1 5 0 は、すくなくとも、これらの呼吸音・発話音判別手段 1 2 0、音声信号選択手段 1 3 0、音声信号処理手段 1 4 0 を制御するものである。

【 0 0 1 6 】

尚、マイクロフォン 1 1 0 は、発話音圧が呼吸音とほぼ同じ程度の微小音声、又は呼吸音より小さい極微小音声を取得する場合は、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、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。

【 0 0 1 7 】

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（～20 k H z）に対してサンプリング定理を満足するように、A D 変換器でのデータ取得は、可聴音域（～20 k H z）に対してサンプリング定理を満足するように例えば、50 k H z 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、音種の変化を検出して集団分けし、音種集団毎に同一処理に委ねるということである。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。

【 0 0 1 8 】

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。（吸気で発話する言語は世界中で 1 つ）

- ・ 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なく、また、時間的長さや繰り返し周期もほぼ一定している。発話は、時間的長さは発話文の長さに依存する。

- ・ 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

10

20

30

40

50

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 500 Hz 以上の成分が多い。呼吸音は 500 Hz 以下の成分が多い。

【0019】

1 - B には、1 - A のものの動作フローが示されている。

マイクロフォン 110 から音入力となされる。音入力の後には、

- 1) 呼吸音・発話音判別がなされ、音種集団も確定する。
- 2) 音声信号（発話音信号）が選択される。増幅可能。
- 3) 音声信号処理手段に入力され、その本来の機能で処理される。

【0020】

図1の音声信号選択手段 130 の他の実施態様について以下に説明する。

図4は、本発明の複数者間干渉音声雑音を回避した音声入力処理装置に使用する音声信号選択手段の他の実施態様と処理フローの一実施態様を示す図である。4 - A には、図1の音声信号選択手段 130 を示している。図1の例では、一つの構成であり、呼吸音・発話音判別手段 120 の結果そのもので、選択を行ってしまえば可能である。しかしながら、一旦、メモリに格納するなど後から選択する場合は、特に、選択のための仕掛けがあると好都合である。このように図1とは別の実施態様として、4 - B の構成があげられる。

4 - B において、音声信号選択手段 130 は、音種識別手段 1300 と信号選択器 1350 を備えている。

呼吸音・発話音判別手段 120 において、データ集団の音種が判別されると、音種識別手段 1300 において、各データ集団の音種識別が可能にされメモリに格納される。一方、音種識別に基づいて、発話音又は呼吸音を選択することが可能となる。信号選択器 1350 は、データ集団の音データをこれに関連づいた音種識別を介して予め定められた音種のデータを選択するものである。信号選択器 1350 からの選択後データは、この予定の音種のデータのみが出力され、次の音声信号処理手段に送られる。

4 - C の動作フローでは、

- (1) 音種集団毎に音種識別してメモリに格納する。
- (2) 音種識別に基づいて信号選択、データが出力される。

尚、4 - A、4 - B において、音声信号選択手段 130、信号選択器 1350 に含ませてよい選択機能以外の処理機能例を挙げると、以下のようなことが可能である。

- 1) 残す。
- 2) 送信する。
- 3) 受信側が認識できるように小さい音を増幅する。

例えば、入力された発話音又は呼吸音の音圧強度は、微小音域、極微小音域では、通常の発話に比べ、音圧が極めて小さいので、前記呼吸音を上回るほぼ通常発話の音圧強度に音声信号選択手段 130、信号選択器 1350 において増幅しておき他に伝送する、或いは、他に伝送後に増幅するなど可能であり、使用上好都合である。実際に、伝送された受信側では、ほぼ通常発話音の音圧で聞くなどの行動が必要だからである。すなわち、微小音域、極微小音域では聞き取りにくいし、周りから得る雑音の方が大きいなどと不都合になるため、そのままでは使われないためである。

- 4) フィルタをかける。

【0021】

図4の音種識別手段 1300 の具体的実施態様を次に説明する。

図5は、本発明の複数者間干渉音声雑音を回避した音声入力処理装置に使用される音種識別手段の具体的実施態様を示す図である。

5 - A において、音種識別手段 1300 は、呼吸音・発話音判別器 120 において呼吸音と発話音が判別されたことを受けて、各音種集団毎に音種識別符号付与器 1310 A により音種識別符号を与え、メモリ 1320 中のデータメモリ 1321 に音種識別符号メモリ 1322 に互いのデータを関連付け（ひもづけ）て格納する。したがって、データメモ

10

20

30

40

50

り 1 3 2 1 の各データ集団は、これに関連した音種識別符号により音種がわかるので、信号選択器 1 3 5 0 により選択的に処理される。

5 - B においては、音種識別手段 1 3 0 0 は、呼吸音・発話音判別器 1 2 0 において呼吸音と発話音が判別されたことを受けて、各音種集団毎に音種別データ振分器 1 3 1 0 B により、音種別に分けたメモリ、この図では、発話音データメモリ 1 3 2 3、呼吸音データメモリ 1 3 2 4、必要なら無音データメモリ 1 3 2 5 に各々格納される。信号選択器 1 3 5 0 において、処理を行う場合に、必要な音種集団を順番に読みだし、呼び出した音種集団の始めと終わりの時間位置を音種集団データのデータに付属させて格納しておけば、これによって、データを処理して初期の時間的順番と位置を保持したまま出力できる。この場合には、各音種集団に音種識別符号を与えないが、区別したメモリに格納することで音種識別符号を与えて、これで識別することと同じ機能を果たしている。

尚、無音データメモリ 1 3 2 5 において、「必要なら」としたのは、発話音データと呼吸音データに音の並びの時間的な位置（アドレス）が与えれば、空いた時間範囲は無音区間なので、特に無音データを格納しなくてもよい。また、無音区間の時間位置だけメモリに格納してもよい。同様に、呼吸音データが後で不要な場合は、メモリに格納しなくてもよいことは勿論である。

【 0 0 2 2 】

図 6 は、本発明の複数者間干渉音声雑音を回避した音声入力処理装置に入力する音声入力と呼吸音・発話音判別、識別符号付け、選択的発話音処理を示す図である。図 5 の 5 - A に対応したものである。

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囲で）に示している。

6 - A は、音声入力した音の波形を示す。波形を示すデータは、図 1 の呼吸音・発話音判別手段において、吸気音、呼気音、発話音を判別される。波形を示すデータを判別してゆくには、マイクロフォンからのアナログデータを直接処理する場合と、一旦、A/D 変換器で処理して、デジタルデータにしたのち処理、又は、一旦メモリに記憶したあとに処理など種々の形態がとれる。データとしては、複数のデータを処理してゆくと、無音区間、または、吸気音、呼気音、発話音のうち、他の種別の音に変わるので、一つの種別の音域の始端データと終端データが検出されるので、始端データと終端データの間のデータ集団が、一つの種別の音域（音種集団）となる。したがって、この各音種集団に吸気音、呼気音、発話音、（或いは無音）を示す識別符号を付与するか、又は、各音種ごとに区別して保存などの音種識別をすることが可能となる。例えば、6 - B では、識別符号を付与する例である。

6 - B は、識別符号として、各音種集団の始端データと終端データに吸気音、呼気音、発話音、（或いは無音）のいずれかであることを示す識別符号（たとえば、この例では、K s : 呼気音の始端、K e : 呼気音の終端、B s : 呼気音の始端、B e : 呼気音の終端、H s : 発話音の始端、H e : 発話音の終端を示す）を各データ集団に付与した状態を示す。各データ集団のデータは、これらの識別符号に関連づいた状態でメモリに格納される。

6 - A では、発話音の強度より大きな強度の吸気音と呼気音がある例で書いてある。発話音がない場合は、吸気音と呼気音を繰り返す。発話音がある場合は、通常は、呼気音のところに発話音がなされ、呼気音の場合より時間幅が長く発話音がなされる。発話音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となされる場合も多いことは周知のことである。

6 - C においては、識別符号をもとに選択的に発話音のみを選択していることを示す例である。このように発話音のみを選択して、出力することができる。尚、逆に、発話音ではなく、吸気音と呼気音を処理することも可能である。この応用は、医療分野などで、吸気音と呼気音のみを選択的に処理して、肺や気管の状態を観測することである。

【 0 0 2 3 】

上記の説明では、切替手段には触れなかったが、微小音域、極微小音域の音を扱うのが主体であるが、このような状態に対応する中で、一つの音声入力処理装置のセットにおいて

、部屋内で単独の作業を行うなどもあり得る。この場合には、通常発話の音圧で入力してもよいので、マイクロフォンの感度或いは、その後の増幅手段の感度を、微小音域、極微小音域対応の場合より下げて、上記の処理をしないことも可能であり、１つのセットで対応することは好都合なので、そのために、両者の間の切り替えを行う切替手段を備えると極めて好都合である。

また、発話音の音圧が呼吸音と同程度または、それより小さい音圧領域に対応しようとしているとき、又は、切替手段が低感度状態でない状態にある場合、マイクロフォンから入力した音の音圧が呼吸音の音圧を上回る場合は、周りから意図しない混入雑音（例えば、隣の人の大きな声が入った）があったとして区別することが好ましい。

また、マイクロフォンは、口先直前に配置または、口先とその外部との間で隔絶するための音遮蔽カバーに覆われて配置されることが外部からの雑音混入を避けるために好ましい。

そして、口先直前に配置または、口先とその外部との間で隔絶するための音遮蔽カバーに覆われて配置された状態で使う場合に対して、部屋内で単独の作業を行うなどで、通常発話の音圧で入力したい場合には、例えば、前記マイクロフォンを上記の配置状態以外の配置（例えば、マイクロフォンを口先から離れた状態や、音遮蔽カバーを外した状態、或いは、スイッチ切替による等価的配置状態以外状態）に応じて、切替手段を切り替えると好都合である。以上の切替手段の例は、どれにも可能だが、一例を図１の１－Ａに切替手段１６０として示す。

【産業上の利用可能性】

【００２４】

以上のように構成されているため、本発明の複数者間干渉音声雑音を回避した音声入力処理装置においては、複数話者が発話入力する環境でも、互いに雑音とならず、情報漏えいにもならず音声入力処理の本来の機能が何らの制限がなく行えるので、産業上利用性が極めて大きい。

【符号の説明】

【００２５】

- １１０ マイクロフォン
- １２０ 呼吸音・発話音判別手段
- １２１ 判別手段
- １２２ ＡＤ変換器
- １２３ 一時メモリ
- １３０ 音声信号選択手段
- １３００ 音種識別手段
- １３１０Ａ 音種識別符号付与器
- １３１０Ｂ 音種別データ振分器
- １３２０ メモリ
- １３２１ データメモリ
- １３２２ 音種識別符号メモリ
- １３２３ 発話音データメモリ
- １３２４ 呼吸音データメモリ
- １３２５ 無音データメモリ
- １３５０ 信号選択器
- １４０ 音声信号処理手段
- １５０ 制御手段
- １６０ 切替手段

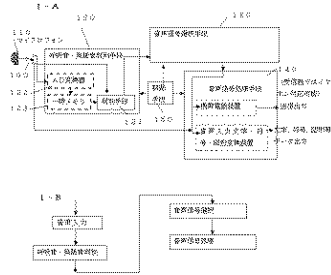
10

20

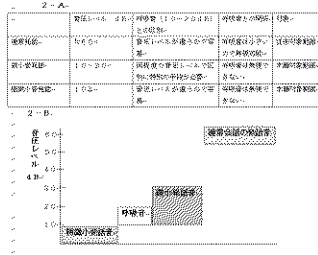
30

40

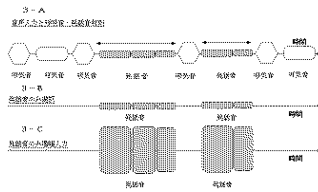
【図 1】



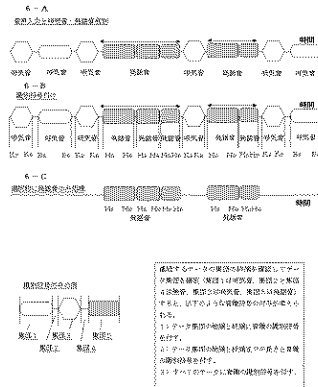
【図 2】



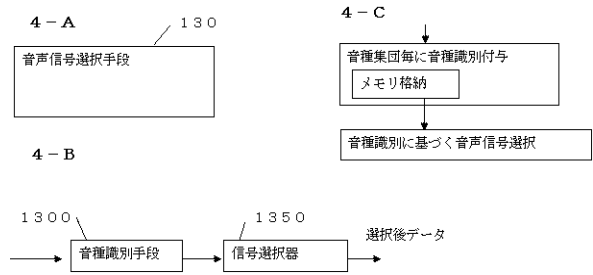
【図 3】



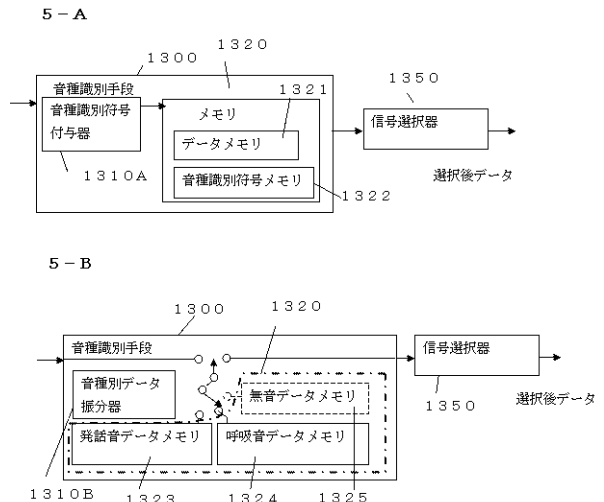
【図 6】



【図 4】



【図 5】



フロントページの続き

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

(58)調査した分野(Int.Cl. , D B 名)

G 1 0 L 1 5 / 0 0 - 1 5 / 3 4 ,
 2 5 / 0 0 - 2 5 / 9 3