

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4556028号
(P4556028)

(45) 発行日 平成22年10月6日(2010.10.6)

(24) 登録日 平成22年7月30日(2010.7.30)

(51) Int.Cl.	F I
G 1 0 L 17/00 (2006.01)	G 1 0 L 17/00 2 0 0 C
G 1 0 L 15/02 (2006.01)	G 1 0 L 15/02 2 0 0 A
	G 1 0 L 15/02 3 0 0 D

請求項の数 8 (全 18 頁)

(21) 出願番号	特願2005-321351 (P2005-321351)	(73) 特許権者	393031586
(22) 出願日	平成17年11月4日(2005.11.4)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2007-127891 (P2007-127891A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成19年5月24日(2007.5.24)	(74) 代理人	100099933
審査請求日	平成19年10月26日(2007.10.26)		弁理士 清水 敏
(出願人による申告)平成17年度独立行政法人情報通信研究機構、研究テーマ「人間情報コミュニケーションの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願		(72) 発明者	パーハム・モクタリ
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	北村 達也
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	竹本 浩典
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内

最終頁に続く

(54) 【発明の名称】 発話主体同定装置及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

所定の音素を発話中の発話主体の声道形状を表す所定の声道形状パラメータと、前記所定の音素を発話中の前記発話主体の音声データと、当該所定の音素を発話中の同定対象者の音声データとに基づいて、同定処理対象者を同定するための発話主体同定装置であって、

複数の発話主体について、前記声道形状パラメータに対する主成分分析を行ない、前記声道形状パラメータの平均値及び前記声道形状パラメータの発話主体による変動を説明する主成分ベクトルのうちの上位の所定個数を算出するための主成分分析手段と、

前記複数の発話主体の各々について、前記所定の音素を発話中の音声データに対する音響分析を行ない、所定の複数の音響特徴量を算出するための音響特徴量算出手段と、

前記複数の発話主体の各々の前記声道形状パラメータについて、前記所定個数の主成分ベクトルの線形和で近似するための主成分係数を算出するための主成分係数算出手段と、

前記複数の発話主体の各々に対し、前記音響特徴量算出手段により算出された前記複数の音響特徴量の線形和で、当該発話主体について前記主成分係数算出手段により算出された前記主成分係数を推定するための係数系列を重回帰分析により決定するための重回帰分析手段と、

前記所定の音素を発話中の同定処理対象者の音声データから前記音響特徴量算出手段により算出されるものと同じ前記複数の音響特徴量を算出し、当該複数の音響特徴量に対し前記重回帰分析手段により算出された係数系列を適用した線形和により、前記同定処理対象者の前記声道形状パラメータを算出するための主成分係数を推定するための推定手段と、

10

20

前記推定手段により推定された主成分係数と前記主成分係数算出手段により前記複数の発話主体の各々に対し算出された主成分係数とを比較することにより、前記同定処理対象者の同定を行なうための同定手段とを含む、発話主体同定装置。

【請求項 2】

前記音響特徴量算出手段は、前記複数の発話主体の各々について、前記所定の音素を発話中の音声データに対するケプストラム分析を行ない、所定の次数までのケプストラム係数を前記音響特徴量として算出するためのケプストラム算出手段を含む、請求項 1 に記載の発話主体同定装置。

【請求項 3】

前記音響特徴量算出手段は、

10

前記複数の発話主体の各々について、前記所定の音素を発話中の音声データに対するケプストラム分析を行ない、所定の次数までのケプストラム係数を算出するためのケプストラム算出手段と、

前記ケプストラム算出手段により前記複数の発話主体の各々について算出された前記所定の次数までのケプストラム係数に対し主成分分析を行ない、前記所定の次数までのケプストラム係数の、発話主体による変動を説明するための主成分ベクトルを算出し、前記複数の発話主体の各々について当該主成分ベクトルにより前記所定の次数までのケプストラム係数を近似するための主成分係数を前記音響特徴量として算出するための手段とを含む、請求項 1 に記載の発話主体同定装置。

【請求項 4】

20

前記声道形状パラメータは、前記所定の音素を発話中の発話主体の声道断面関数を含む、請求項 1 ～請求項 3 のいずれかに記載の発話主体同定装置。

【請求項 5】

前記声道形状パラメータは、前記所定の音素を発話中の発話主体の声道断面幅関数を含む、請求項 1 ～請求項 3 のいずれかに記載の発話主体同定装置。

【請求項 6】

前記同定手段は、

前記推定手段により推定された主成分係数と、前記複数の発話主体の各々に対し前記主成分係数算出手段により算出された主成分係数との間に定義される所定の距離関数の値を算出するための距離算出手段と、

30

前記複数の発話主体のうち、前記距離算出手段により算出された距離が最も小さい発話主体として前記同定処理対象者を同定するための手段とを含む、請求項 1 ～請求項 5 のいずれかに記載の発話主体同定装置。

【請求項 7】

前記同定処理対象者を同定するための手段は、

前記距離算出手段により算出された距離のうちの最短距離を特定するための手段と、

前記最短距離が所定のしきい値に対し所定の関係を充足するか否かを判定するための手段と、

前記判定するための手段により前記最短距離が前記しきい値以下であると判定されたことに応答して、前記最短距離を与える発話主体として前記同定処理対象者を同定するための手段と、

40

前記判定するための手段により前記最短距離が前記しきい値より大きいと判定されたことに応答して、同定ができないことを示す処理を実行するための手段とを含む、請求項 1 ～請求項 5 のいずれかに記載の発話主体同定装置。

【請求項 8】

コンピュータにより実行されると、当該コンピュータを請求項 1 ～請求項 7 のいずれかに記載の発話主体同定装置として動作させる、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

50

この発明は発話主体の音声に基づいて発話主体を同定又は認証（以下単に「同定」と呼ぶ。）する発話主体同定技術に関し、特に、複数の候補者の中から最も可能性の高い者を発話主体として同定する発話主体同定装置に関する。

【背景技術】

【0002】

社会生活において情報の持つ意味が大きくなるに従い、情報の窃取、携帯電話の不正使用等、情報に関連する問題が多発するようになってきている。こうした問題に対処するために、コンピュータ、銀行の全自動現金受払機、携帯電話等の機器において、使用者が正当な使用者か否か、使用者が誰かを判定する技術、すなわち個人同定技術が求められている。

【0003】

使用者が正当な者か否かを判定する技術として、最近バイオメトリクス情報を用いたものが増えている。例えば指紋、手のひらにおける静脈の分布、虹彩の文様などがその典型的なものである。こうした情報は、人により異なり、かついわゆる「なりすまし」を行なうことがむずかしいという特徴がある。

【0004】

しかし、こうした情報により個人を同定するためには、専用のハードウェアが必要になることが多い。従って、システムが高価になりやすいという欠点がある。

【0005】

バイオメトリクス情報の一つに、音声情報がある。音声情報は、個人の同定に從來から使用されている。音声は、個人の発話器官の形状的特徴をよくあらわしており、従って個人の同定には好適であると思われる。

【0006】

音声を用いた従来の技術として、特許文献1に開示の声紋と呼ばれる情報を用いて個人を識別する声紋認証システムがある。この声紋認証システムでは、予めユーザの声紋を登録しておき、認証時には、入力された音声の声紋と、登録されていた声紋とを照合して認証すべきか否かを判定する。従ってこの場合の声紋認証は、声紋による個人の識別技術であるということもできる。

【特許文献1】特開2004-258845号

【発明の開示】

【発明が解決しようとする課題】

【0007】

声紋は、基本的には音声信号に含まれる周波数分布（スペクトル）の時間的変化をグラフ化したものである。そのため、声紋を照合する作業を機械化する場合、同定処理対象の音声のスペクトルを時系列で算出し、それらスペクトルから算出される所定の音響特徴量の時系列を照合作業に用いることになる。

【0008】

このような手法は、音響特徴量の個人間の差異と発話器官の個人間の差異とが一对一の関係であるという前提に立つ限り、機能するものと考えられる。しかし、そのような前提が常に成立することについての保証はない。むしろ、発話器官としての声道について、異なる声道形状から同様の音響特徴を持つ音声が生じられることもあることが知られている。そうした場合、同定を正しく行なうことができない可能性がある。また、スペクトルの時系列を扱うので、処理するデータ量が多くなり、処理に時間を要するという問題もある。

【0009】

従って、個人の発話器官の形状をより確実に反映した形で、発話主体の同定処理を短時間に行なうことができる発話主体同定装置が求められている。

【0010】

それゆえに本発明の目的は、個人の発話器官の形状をより確実に反映した形で、個人の同定を短時間に行なうことができる発話主体同定装置を提供することである。

【課題を解決するための手段】

【0011】

本発明の第1の局面に係る発話主体同定装置は、所定の音素を発話中の発話主体の声道形状を表す所定の声道形状パラメータ及び当該音素の発話時の発話主体の音声データと、当該所定の音素を発話中の同定処理対象者の音声データとに基づいて、同定処理対象者を同定するための発話主体同定装置であって、複数の発話主体について、声道形状パラメータに対する主成分分析を行ない、声道形状パラメータの平均値及び声道形状パラメータの発話主体による変動を説明する主成分ベクトルのうちの上位の所定個数を算出するための主成分分析手段と、複数の発話主体の各々について、所定の音素を発話中の音声データに対する音響分析を行ない、所定の複数の音響特徴量を算出するための音響特徴量算出手段と、複数の発話主体の各々の声道形状パラメータについて、所定個数の主成分ベクトルの線形和で近似するための主成分係数を算出するための主成分係数算出手段と、複数の発話主体の各々に対し、音響特徴量算出手段により算出された複数の音響特徴量の線形和で、当該発話主体について主成分係数算出手段により算出された主成分係数を推定するための係数系列を重回帰分析により決定するための重回帰分析手段と、所定の音素を発話中の同定処理対象者の音声データから音響特徴量算出手段により算出されるものと同じ複数の音響特徴量を算出し、当該複数の音響特徴量に対し回帰分析手段により算出された係数系列を適用した線形和により、同定処理対象者の声道形状パラメータを算出するための主成分係数を推定するための推定手段と、推定手段により推定された主成分係数と主成分係数算出手段により複数の発話主体の各々に対し算出された主成分係数とを比較することにより、同定処理対象者の同定を行なうための同定手段とを含む。

10

20

【0012】

主成分分析手段は、複数の発話主体の所定の音素の発話時の声道形状を表す声道形状パラメータに対する主成分分析を行なう。各発話主体の声道形状は、その結果得られる上位の所定個数の主成分ベクトルの線形和で近似できる。そのための主成分係数が主成分係数算出手段により算出される。音響特徴量算出手段は、複数の発話主体について、声道形状パラメータを得たときと同じ音素について発話した際の音声データに対する音響分析を行なう。その結果、所定の複数の音響特徴量が得られる。重回帰分析手段は、各発話主体に対して算出された音響特徴量の線形和で、当該発話主体に対して算出された主成分係数を推定するための係数系列を重回帰分析により決定する。同定処理時には、同定処理対象者の音声データに対し、音響特徴量算出手段により音響特徴量を算出する。この音響特徴量に対し、重回帰分析手段により決定された係数系列による線形和を適用することにより、推定手段が同定処理対象者の声道形状パラメータを主成分ベクトルで近似する際の主成分係数を推定する。同定手段は、この主成分係数と、主成分係数算出手段により算出された主成分係数とを比較することにより、同定処理対象者を複数の発話主体の中から同定する。

30

【0013】

同定処理時には同定処理対象者の音声データのみが必要であり、声道形状パラメータは必要とされない。この音声データにより、声道形状パラメータを算出するための主成分係数を算出し、その結果で同定処理対象者が同定される。音声データのみにより同定を行なう場合と異なり、声道形状を推定した結果によって同定を行なうため、音声と声道形状との間の対応関係のあいまいさが削減され、従来よりも信頼性の高い同定処理が行なえる。またスペクトルそのものでなく、主成分係数で同定処理を行なうため、処理に要する時間が削減できる。その結果、個人の発話器官の形状をより確実に反映した形で、個人の同定を短時間に行なうことができる発話主体同定装置を提供できる。

40

【0014】

好ましくは、音響特徴量算出手段は、複数の発話主体の各々について、所定の音素を発話中の音声データに対するケプストラム分析を行ない、所定の次数までのケプストラム係数を音響特徴量として算出するためのケプストラム算出手段を含む。

【0015】

ケプストラム係数は音響処理に多用される、音響的特徴をよく現すものとして知られた

50

音響特徴量であり、その算出手法は確立されている。そのためのツール（コンピュータ用プログラム群）も容易に入手できる。従って、容易に信頼性の高い同定処理を実現できる。

【0016】

より好ましくは、音響特徴量算出手段は、複数の発話主体の各々について、所定の音素を発話中の音声データに対するケプストラム分析を行ない、所定の次数までのケプストラム係数を算出するためのケプストラム算出手段と、ケプストラム算出手段により複数の発話主体の各々について算出された所定の次数までのケプストラム係数に対し主成分分析を行ない、所定の次数までのケプストラム係数の、発話主体による変動を説明するための主成分ベクトルを算出し、複数の発話主体の各々について当該主成分ベクトルにより所定の

10

【0017】

ケプストラム係数は、前述のように音響処理で多用された信頼性の高い特徴量である。さらにこれらに対する主成分分析によってケプストラム係数を近似するための主成分係数を算出する。上位の限定された個数の主成分係数により、発話主体によるケプストラム係数の変動がよく説明されるため、重回帰分析処理で扱うデータの個数が減少でき、処理が簡略化される。

【0018】

声道形状パラメータは、所定の音素を発話中の発話主体の声道断面関数を含んでもよいし、所定の音素を発話中の発話主体の声道断面幅関数を含んでもよい。

20

【0019】

声道形状パラメータとして、声道断面積関数を用いると、発話者の発話時の声道形状を精度高く近似できる。声道断面幅関数を用いると、使用するデータ量が削減でき、処理を簡略化できる。

【0020】

好ましくは、同定手段は、推定手段により算出された主成分係数と、複数の発話主体の各々に対し主成分係数算出手段により算出された主成分係数との間に定義される所定の距離関数の値を算出するための距離算出手段と、複数の発話主体のうち、距離算出手段により算出された距離が最も小さい発話主体として同定処理対象者を同定するための手段とを含む。

30

【0021】

主成分係数により主成分ベクトル空間内で特定される点の間の距離は、それら点に対応する声道形状の間の相違量に対応すると考えられる。その結果、この距離が最も小さいものを用いて同定を行なうことにより、同定処理対象者の声道形状と最もよく似た声道形状を持つ発話主体を特定できる。

【0022】

より好ましくは、同定処理対象者を同定するための手段は、距離算出手段により算出された距離のうちの最短距離を特定するための手段と、最短距離が所定のしきい値に対し所定の関係を充足するか否かを判定するための手段と、判定するための手段により最短距離がしきい値以下であると判定されたことに応答して、最短距離を与える発話主体として同定処理対象者を同定するための手段と、判定するための手段により最短距離がしきい値より大きいと判定されたことに応答して、同定ができないことを示す処理を実行するための手段とを含む。

40

【0023】

最短距離が所定のしきい値より大きい場合には、同定処理対象者の声道形状と十分類似した声道形状を持った発話主体がないことになる。この場合には同定ができないことを示す処理を実行することにより、誤った同定を行なう危険性が小さくなる。

【0024】

本発明の第2の局面に係るコンピュータプログラムは、コンピュータにより実行される

50

と、当該コンピュータを上記したいずれかの発話主体同定装置として動作させるコンピュータプログラムである。

【0025】

このコンピュータプログラムをコンピュータで実行することにより、上記した発話主体同定装置を実現できる。その結果、上記した発話主体同定装置による効果と同様の効果を得ることができる。

【発明を実施するための最良の形態】

【0026】

以下に説明する実施の形態は、MRI (magnetic resonance imaging: 核磁気共鳴画像法) により得られた発話中の発話者の声道形状の特徴量と、発話の音声信号から得られた音響特徴量との関係を用いて、音響特徴量から声道形状の特徴量を推定するための関係式を求めるというものである。同定時には、同定処理対象者の音声から得た音響特徴量に基づき、当該関係式を用いて同定処理対象者の声道形状の特徴量を推定し、予め登録しておいた発話者の声道形状の特徴量と比較することにより発話者を識別し、同定するというものである。

【0027】

なお、声道形状を現すものとしては、例えば声道断面積関数又は声道断面幅関数を用いることができる。声道断面積関数とは、声道の中心線を考え、声道中心線上のある点における声道の断面積を、その点の、声門からの距離の関数として表したものである。声道断面積関数により、声道形状が具体的に規定されるので、同定処理の精度が高くなることが期待される。声道断面幅関数とは、人体の正中矢状断面において、声道中心線上のある点における声道の幅（声道中心線と直交する方向の声道の幅）を、声紋からその点までの距離の関数として表したものである。声道形状を現すものとして、声道断面積関数と比較するとデータ量は少なく済み、必要な処理も簡略となる。

【0028】

また、以下の実施の形態では、声道形状の特徴量を求めるためにMRI 画像データを用いる。しかし、本発明はそのような実施の形態には限定されず、声道形状を知ることができる手段であればどのようなものを用いてもよい。例えばCT (Computed Tomography: コンピュータ断層法) など、生体内の3次元形状を撮像可能な装置によるデータを用いることも可能である。

【0029】

〔第1の実施の形態〕

<構成>

図1に、本発明の第1の実施の形態に係る音声による発話者同定システム20のブロック図を示す。図1を参照して、発話者同定システム20は、複数の登録対象者の各々について、所定の音素の発話時の声道形状を撮影したMRI データを記憶するためのMRI データ記憶部30と、MRI データの撮影時と近接したとき（可能なら同時）に、撮影時の発話から得られた音声データを記憶するための音声データ記憶部32と、MRI データ記憶部30に記憶された複数の登録対象者のMRI データと、音声データ記憶部32に記憶された各登録対象者の音声データとに基づいて、これら複数の登録対象者を識別するための参照データと、同定処理時の同定処理対象者の音声の音響特徴量から当該同定処理対象者の声道形状関数を推定する際に使用する回帰係数とを求め、それぞれ出力するための同定用データ作成装置34とを含む。

【0030】

本実施の形態では、同定用データ作成装置34は、複数の登録対象者のMRI データから声道形状を表す関数（以下「声道形状関数」と呼ぶ。）を求め、それらに対する主成分分析（以下「PCA」）により、登録対象者の個人的差異に基づく声道形状の変動に主として貢献している主成分を求める。MRI データ記憶部30は各登録対象者の声道形状をこれら主成分の線形和として表すための係数をさらに算出する。この係数が、各登録対象者を同定し認証するための参照データとして用いられる。本実施の形態では第1主成分（

PC I) 及び第2主成分(PC II)を用いる。

【0031】

同定時には、同定処理対象者のMRIを撮影することはできない。従って、本実施の形態では、より簡便に利用できる音声を用いる。参照データは、各登録対象者の声道形状関数を主成分の線形和で表すための係数である。従って、発話者の音声の音響特徴量から何らかの形でこの係数を推定する必要がある。本実施の形態では、同定用データ作成装置34は、登録対象者の音声からケプストラム係数を算出し、登録対象者の参照データである主成分係数を、このケプストラム係数の線形和として表すための係数を、複数の登録対象者から得たデータを用いた重回帰分析により求める。ケプストラム係数は音響処理に多用される、音響的特徴をよく現すものとして知られた音響特徴量であり、その算出手法は確立されている。そのためのツールも容易に入手できる。従って、以下の同定処理の信頼性を高くすることができ、装置の作成も簡略となる。

10

【0032】

発話者同定システム20はさらに、同定用データ作成装置34が算出した回帰係数を記憶するための回帰係数記憶部36と、同定用データ作成装置34が作成した参照用データを記憶するための参照データベース（以下「参照DB」と呼ぶ。）38とを含む。同定処理対象者の音声に対してケプストラム分析を行なってケプストラム係数を算出し、このケプストラム係数に対し、回帰係数記憶部36に記憶された回帰係数を係数とする線形和を計算することにより、同定処理対象者の同定（認証）発話時の声道形状関数を主成分で表すための主成分係数を推定できる。この主成分係数を、参照データベース38に記憶された複数の登録対象者の主成分係数と比較することにより、登録対象者の中から同定処理対象者40を同定することができる。

20

【0033】

本実施の形態では、発話者同定システム20はさらに、同定処理対象者40の音声に対し上記したようなケプストラム分析及び声道形状関数のための主成分係数の算出処理を行ない、さらに、算出された主成分係数を参照データベース38に記憶された参照データと比較することにより、複数の登録対象者の中から同定処理対象者40を同定し、同定結果44を出力するための同定処理装置42を含む。本実施の形態に係る発話者同定システム20は、実際には単一のコンピュータでも実現できるし、同定処理装置42のみを独立のコンピュータとすることもできる。

30

【0034】

図2に、同定用データ作成装置34の詳細なブロック図を示す。図2を参照して、同定用データ作成装置34は、MRIデータ記憶部30に記憶されている複数の登録対象者のMRIデータ（声道形状データ）に対しPCAを行ない、声道形状データに関する平均値とPC I及びPC IIとを出力するためのPCA分析部60と、PCA分析部60により得られた平均値とPC I及びPC IIとを記憶するための主成分記憶部62と、主成分記憶部62に記憶された平均値とPC I及びPC IIとを用い、MRIデータ記憶部30に記憶された各登録対象者の声道形状データをPC I及びPC IIの線形和で表すための第1及び第2の主成分係数を算出し参照データベース38に格納するための主成分係数算出部67とを含む。

40

【0035】

同定用データ作成装置34はさらに、音声データ記憶部32に記憶された各登録対象者の音声データに対しケプストラム分析を行ない、所定次数までのケプストラム係数を出力するためのケプストラム算出部64と、ケプストラム算出部64により各登録対象者について算出されたケプストラム係数を記憶するためのケプストラム係数記憶部66と、ケプストラム係数記憶部66に記憶された各登録対象者のケプストラム係数の線形和で、主成分係数算出部67により算出された、対応する登録対象者の主成分係数を算出するための係数を重回帰分析により求め、回帰係数記憶部36に格納させるための重回帰分析処理部68とを含む。

【0036】

50

ここで、登録対象者が m 人、算出されるケプストラム係数が n 次までとすると、重回帰分析処理部68の行なう処理は、次の回帰式から係数 α_j 及び β_j ($0 \leq j \leq n$)を決定することである。

【0037】

【数1】

$$PC\ I_1 = \alpha_0 + \alpha_1 C_{1,1} + \alpha_2 C_{1,2} + \cdots + \alpha_n C_{1,n}$$

$$PC\ II_1 = \beta_0 + \beta_1 C_{1,1} + \beta_2 C_{1,2} + \cdots + \beta_n C_{1,n}$$

$$PC\ I_2 = \alpha_0 + \alpha_1 C_{2,1} + \alpha_2 C_{2,2} + \cdots + \alpha_n C_{2,n}$$

$$PC\ II_2 = \beta_0 + \beta_1 C_{2,1} + \beta_2 C_{2,2} + \cdots + \beta_n C_{2,n}$$

⋮

$$PC\ I_m = \alpha_0 + \alpha_1 C_{m,1} + \alpha_2 C_{m,2} + \cdots + \alpha_n C_{m,n}$$

$$PC\ II_m = \beta_0 + \beta_1 C_{m,1} + \beta_2 C_{m,2} + \cdots + \beta_n C_{m,n}$$

ただし $PC\ I_i$ 及び $PC\ II_i$ はそれぞれ、 i 番目の登録対象者の声道形状関数を第1主成分と第2主成分との線形和で表す際の第1主成分及び第2主成分の主成分係数を表す。 $C_{i,j}$ は i 番目の登録対象者の音声から得た第 j 次ケプストラム係数である。

【0038】

図3に、同定処理装置42の詳細なブロック図を示す。図3を参照して、同定処理装置42は、同定処理対象者40の同定処理時に入力される音声データ80から、第 n 次までのケプストラム係数を算出するためのケプストラム算出部82と、ケプストラム算出部82から出力される第 n 次までのケプストラム係数に対し、回帰係数記憶部36に記憶された回帰係数による線形和を適用し、同定処理対象者40の発話時の声道形状関数を主成分 $PC\ I$ 及び $PC\ II$ で表すための主成分係数を第1主成分 $PC\ I$ 及び第2主成分 $PC\ II$ に対し推定するための主成分係数推定部84とを含む。

【0039】

同定処理装置42はさらに、主成分係数 $PC\ I$ 及び $PC\ II$ が張る空間内において、主成分係数推定部84により算出された主成分係数の組により定まる点と、参照データベース38に記憶された各登録対象者のための主成分係数の組により定まる点との間のユークリッド距離を各登録対象者に対し算出するための距離算出部86と、距離算出部86により算出された距離のうち、最短のものがどれかを判定し、その距離に対応する登録対象者を同定処理対象者40として同定し同定結果44を出力するための最短距離判定部88とを含む。なお、図3には図示していないが、最短距離判定部88は、距離算出部86により算出された距離のうち最短のものが、所定のしきい値より大きい場合には、同定処理対象者40に該当する登録対象者はいないものとみなし、失敗を示す同定結果44を出力するものとする。

【0040】

上記実施の形態では、同定処理には母音の発話を用いる。母音を発話しているときの声道形状は一般的に安定していることが知られており、声道形状が不安定であるために生ずる同定誤りが少なくなることが期待されるためである。MRIデータを撮影しているときには、所定の母音を発話し、そのときの声道形状を撮影する。音声データ記憶部32には、このときの音声を記憶する。周知のようにMRI画像を撮影するときにはMRI機器から大きな音が発生する。従ってMRI画像と全く同時にそのときの発話音声を良好に録音することは難しい。しかし、MRI画像の撮影の直前から登録対象者が発声を始めるようにし、撮影が終了した後にも所定時間発声を継続し、その間に音声を録音することにより、MRI画像撮影時の声道形状とほぼ一致する声道形状のときの良好な音声データを得ることができる。

【0041】

10

20

30

40

50

上記した母音は、全登録対象者に対し共通の一つの母音である。さもないと同じ条件での照合が行なえない。しかし、全登録対象者によって複数の母音による発話時のMR I データと音声データとを取得し、母音ごとに区別して同定処理を行なうこともできる。

【0042】

例えば、3人の登録対象者（話者）の各々から、3つの母音について上記した処理で音響データを取得したときについて図4を参照して説明する。図4では、図面及び説明を簡略にするために、3人の話者に3つの母音を発声させて得た音声から第2次までのケプストラム係数を算出し、それらケプストラム係数を主成分分析した結果の例を示している。図4において、横軸は第1次のケプストラム係数、縦軸は第2次のケプストラム係数を示す。

10

【0043】

図4を参照して、一般的に、異なる母音についてはケプストラムの値は互いに異なる。従って、同じ母音について複数の話者から得たケプストラム係数は図4に示す領域100、102及び104に示すように比較的小さな領域にそれぞれまとまるが、異なる母音についてのこれら領域は互いに大きく離れることになる。これら領域100、102及び104の各々について主成分分析を行なうことにより、異なる母音について、独立に上記した同定処理のためのデータを準備できる。そこで、例えば同定処理時にはランダムに母音を選択し、その母音を同定処理対象者に発声させ、その結果と、当該母音について予め準備しておいた回帰係数記憶部36及び参照データベース38を用いて同定処理を行なうことができる。

20

【0044】

図5に、図3に示す距離算出部86での距離算出と、最短距離判定部88での最短距離判定との原理を概略的に示す。図5においても、図面及び説明を簡略にするために、登録対象者が第1の話者から第3の話者までの3人、使用するケプストラム係数は第2次までとしている。

【0045】

図5を参照して、第1～第3の話者のある音声の発話から得た第1次及び第2次のケプストラム係数により、それぞれ点120、122及び124がこのケプストラム係数空間で定まる。これらに対する主成分分析を行なうことにより、第1主成分に対応する第1主成分軸110と、第2主成分軸112とがこの空間内に規定される。

30

【0046】

同定処理対象者の音声データから得たケプストラム係数により、点130が定まる。距離算出部86での処理は単純である。すなわち、距離算出部86は、この点130と、点120、122及び124との間の距離L1、L2及びL3をそれぞれ算出する。最短距離判定部88は、これら距離L1、L2及びL3を互いに比較し、最も値の小さな距離を選択する。図5に示す例では $L2 < L3 < L1$ であるから、最短距離判定部88はL2を選択する。最短距離判定部88はさらに、同定処理対象者がこの距離L2に対応する話者、すなわち第2の話者である、と判定し判定結果を出力する。ただし前述の通り、この最短距離があるしきい値より大きいと同定処理失敗という判定を行なう。

40

【0047】

なお、音声データからケプストラム係数を算出する際には、音声データのうちで最もよい性能が得られるような帯域を予め選択しておく必要がある。そのために、予め次のような処理を行なう。サンプリング周波数の範囲内で複数の周波数帯域を用いて予めテスト用データを用いてケプストラム係数及び主成分係数を算出する。算出された主成分係数を用い、テスト用の同定処理データから主成分を推定する。推定された主成分とテスト用のMR I データから得られた主成分との間の誤差を算出する。こうして複数の周波数帯域を用いて算出された誤差のうち、最も小さな誤差が得られた周波数帯域を実際の同定でも用いるようにする。

【0048】

<コンピュータによる実現>

50

上記した実施の形態に係る発話者同定システム20は、既に述べたようにコンピュータ及び当該コンピュータ上で実行されるプログラムにより実現できる。ただしMRIデータと、MRIデータ取得時の登録対象者の発話の音声データとは与えられるものとする。また、同定処理には音声データの取得を行なうことが必要であるが、市販されているコンピュータでは音声関係の処理をするための機能が実装されていることが大半であるから、市販のコンピュータを用いることができる。

【0049】

図6に、上記した同定用データ作成装置34を実現するためのプログラムのフローチャートを示す。図6を参照して、ステップ140で繰返制御変数*i*に0が代入される。これは初期化処理の一部である。

10

【0050】

ステップ142で、変数*i*の値が、予め定められた最大繰返回数MAX_{*i*}より大きいかが判定される。結果がYESであればステップ154に進む。NOであればステップ144に進む。なお、最大繰返回数MAX_{*i*}は、処理対象として選択された周波数帯域の数を示す。

【0051】

ステップ144では、予めテストで使用する複数の周波数帯域の下限及び上限を格納した配列FLOW及びFHIGHの*i*番目の要素（これらをそれぞれFLOW[*i*]及びFHIGH[*i*]とする。）が読出され、周波数帯域FLOW[*i*]－FHIGH[*i*]で、テスト用データのケプストラム係数が算出される。

20

【0052】

ステップ146では、算出されたケプストラム係数を用い、既知のMRIデータから得られた声道形状関数の主成分をケプストラム係数の線形和で求めるための係数を重回帰分析により求める。こうして得られたケプストラム係数とステップ144で算出されたケプストラム係数とに基づき、ステップ148でケプストラム係数の線形和によって主成分係数を算出し、元の主成分係数との間の誤差を求める。ステップ150で、こうして得られた誤差を誤差記憶用の配列の*i*番目の要素として記憶する。

【0053】

ステップ152では変数*i*の値を1インクリメントする。制御はステップ142に戻る。

30

【0054】

ステップ142での判定の結果、変数*i*の値が最大繰返回数MAX_{*i*}より大きいと、ステップ154において今まで算出され配列に記憶された誤差のうち、最小のものがどれかが判定される。そして、最小の誤差を与える回帰係数が出力される。

【0055】

こうすることにより、最も精度の高い結果が得られる周波数帯域を用いて同定処理を行なうことができる。

【0056】

図7に、同定処理時の同定処理装置42を実現するためのプログラムのフローチャートを示す。図7を参照して、ステップ180でまず乱数を発生させる。発生した乱数に基づき、予め定められた複数の母音のうちどの母音を同定処理に用いるかを決定する。

40

【0057】

ステップ182において、決定された母音の参照データを参照データベース38（図1参照）から取得する。続いてステップ184において、決定された母音に対応する回帰係数を回帰係数記憶部36（図1参照）より取得する。

【0058】

ステップ186で、発声すべき母音を同定処理対象者に提示する。この提示は、画面へのメッセージの表示又は音声による提示のいずれでもよい。

【0059】

ステップ188で、同定処理対象者による、指定された母音の発話の音声データを取得

50

する。ステップ190で、この音声からケプストラム係数を算出する。ステップ192で、このケプストラム係数と、ステップ184で取得された回帰係数とにより、ケプストラム係数の線形和を計算する。この計算により、同定処理対象者の母音発声時の声道形状関数を主成分の線形和で表すための主成分係数が推定される。ステップ194で、こうして推定された主成分係数により、主成分ベクトルが張る空間内で定まる点と、参照データにより定まる点との間の距離が算出される。

【0060】

ステップ196では、こうして算出された距離のうち最短距離を決定し、さらにこの最短距離が所定のしきい値より小さいか否かが判定される。最短距離がしきい値より小さければステップ198に進み、さもなければステップ200に進む。

10

【0061】

ステップ198では、同定処理対象者は、最短距離を与える点に対応する登録対象者であるものとして、同定成功時の処理が実行される。一方ステップ200では、登録対象者のうち、同定処理対象者に相当する者はいないものとして、同定失敗を示す処理、例えば認証失敗を示すメッセージを出力する処理を行なう。

【0062】

＜コンピュータハードウェア＞

図8に、本実施の形態に係る同定処理装置42を実現するためのコンピュータシステム210の外観図を、図9にコンピュータシステム210のブロック図を、それぞれ示す。なおここに示すコンピュータシステム210はあくまで一例であり、この他にも種々の構成が可能である。

20

【0063】

図8を参照して、コンピュータシステム210は、ノート型のパーソナルコンピュータ（以下「ノートパソコン」と呼ぶ。）212と、このノートパソコン212の音声入出力のための端子に接続されたマイクロフォン284とを含む。ノートパソコン212は、キーボード286、及びカーソル移動用の静電感知パッド288を有する第1の筐体220と、第1の筐体220の上面の一辺を中心として回動可能に第1の筐体220に取付けられた、液晶表示装置282を持つ第2の筐体222とを含む。ノートパソコン212にはさらに、DVD（Digital Versatile Disc）302を装着可能なDVDドライブ290が内蔵されている。

30

【0064】

図9を参照して、ノートパソコン212はさらに、DVDドライブ290に接続されたバス306と、いずれもバス306に接続された中央演算装置（Central Processing Unit：CPU）296、ノートパソコン212のブートアッププログラムなどを記憶したROM（Read-Only Memory）298、CPU296が使用する作業エリア及びCPU296により実行されるプログラムの格納エリアを提供するRAM（Random Access Memory）300、及び前述した回帰係数記憶部36、参照データベース38、MRIデータ記憶部30、音声データ記憶部32、主成分記憶部62、ケプストラム係数記憶部66等として動作するハードディスク294と、バス306に接続され、半導体メモリ304が装着されるメモリドライブ292と、マイクロフォン284に接続され、音声アナログ信号をデジタル化する処理等を行なうためのサウンドボード308とを含む。

40

【0065】

上記した実施の形態のシステムを実現するソフトウェアは、たとえば、DVD302又は半導体メモリ304のような記録媒体上に記録されて流通し、DVDドライブ290又はメモリドライブ292のような読取装置を介してノートパソコン212に読込まれ、ハードディスク294に格納される。CPU296がこのプログラムを実行する際には、ハードディスク294からこのプログラムを読出してRAM300に格納し、図示しないプログラムカウンタによって指定されるアドレスから命令を読出して実行する。CPU296は、処理対象のデータをハードディスク294又は半導体メモリ304のような記憶装

50

置から読出し、処理結果を同じくハードディスク 294 又は半導体メモリ 304 のような記憶装置に格納する。

【0066】

ノートパソコン 212 の動作自体は周知であるので、ここではその詳細については繰返さない。

【0067】

なお、ソフトウェアの流通形態は上記したように記憶媒体に固定された形には限定されない。たとえば、ネットワークを通じて接続された他のコンピュータからデータを受取る形で流通することもあり得る。また、ソフトウェアの一部が予めハードディスク 294 中に格納されており、ソフトウェアの残りの部分をネットワーク経由でハードディスク 294 に取込んで実行時に統合するような形の流通形態もあり得る。

10

【0068】

一般的に、現代のプログラムはコンピュータのオペレーティングシステム（OS）又はいわゆるサードパーティ等によって提供される汎用の機能を利用し、それらを所望の目的に従って組織化した形態で実行することにより前記した所望の目的を達成する。従って、以下に述べる本実施の形態の各機能のうち、OS 又はサードパーティが提供する汎用的な機能を含まず、それら汎用的な機能の実行順序の組合せだけを指定するプログラム（群）であっても、それらを利用して全体的として所望の目的を達成する制御構造を有するプログラム（群）である限り、それらが本発明の技術的範囲に含まれることは明らかである。

【0069】

20

<動作>

上記した発話者同定システム 20 は以下のように動作する。発話者同定システム 20 の動作には二つのフェーズがある。第 1 のフェーズでは MRI データと音声データから回帰係数を算出し、同時に参照データベース 38 を作成する。第 2 のフェーズでは、第 1 のフェーズで作成された回帰係数と参照データベース 38 とを用い、同定処理対象者 40 の同定を行なう。

【0070】

第 1 のフェーズでは発話者同定システム 20 は以下のように動作する。なお、MRI データ記憶部 30 及び対応する音声データ記憶部 32 は予め準備されているものとする。

【0071】

30

図 2 を参照して、PCA 分析部 60 が、MRI データ記憶部 30 に格納された各 MRI データから、各登録対象者の声道形状関数を算出し、それらに対する主成分分析を行なう。その結果、声道形状関数の平均値と第 1 及び第 2 主成分とが得られる。これらは主成分記憶部 62 に記憶される。

【0072】

主成分係数算出部 67 が、主成分記憶部 62 に記憶された平均値及び主成分、並びに MRI データ記憶部 30 に記憶された MRI データに基づき、各登録対象者の声道形状関数を主成分の線形和で表すための主成分係数を算出し、参照データベース 38 に参照データとして格納する。

【0073】

40

一方、ケプストラム算出部 64 は、音声データ記憶部 32 に格納された各登録対象者の音声データの各々から、所定次数までのケプストラム係数を算出する。なお、ここでは詳細は繰返さないが、実際には複数の周波数帯域についてこの処理を繰返し、最終的に最もよい結果を与える周波数帯域で算出されたケプストラム係数を用いるようにするとよい。

【0074】

各登録対象者に対し算出されたケプストラム係数は、ケプストラム係数記憶部 66 に記憶される。

【0075】

重回帰分析処理部 68 は、ケプストラム係数記憶部 66 に記憶された各登録対象者に対するケプストラム係数の線形和で主成分係数算出部 67 に記憶された各登録対象者のため

50

の主成分係数を得るための係数を、重回帰分析により算出する。算出された回帰係数は回帰係数記憶部 36 に記憶される。

【0076】

以上で第1のフェーズは終了である。

【0077】

第2のフェーズは同定処理フェーズである。図3を参照して、同定処理対象者40から、所定の母音の音声データ80が与えられると、ケプストラム算出部82はその音声データ80からケプストラム算出部64と同じ周波数帯域を用いて同じ次数までのケプストラム係数を算出する。このケプストラム係数は主成分係数推定部84に与えられる。

【0078】

主成分係数推定部84は、回帰係数記憶部36に記憶された回帰係数を用い、ケプストラム算出部82から与えられたケプストラム係数の線形和を計算し、同定処理対象者の主成分係数を推定する。推定された主成分係数は距離算出部86に与えられる。

【0079】

距離算出部86は、主成分係数推定部84から与えられた主成分係数により、主成分ベクトルが張る空間内に定められる点と、参照データベース38に記憶された各登録対象者の主成分係数により定められる点との間のユークリッド距離を算出する。距離算出部86は、算出されたユークリッド距離に、対応する登録対象者を特定する情報を付して最短距離判定部88に与える。

【0080】

最短距離判定部88は、与えられた距離のうちの最短距離で、かつ所定のしきい値以下のものがあれば、その距離に対応する登録対象者を特定する情報を同定結果44として出力する。もし上記した条件を満足する距離がなければ、最短距離判定部88は同定失敗を示す同定結果44を出力する。

【0081】

以上が第2のフェーズにおける発話者同定システム20の動作である。

【0082】

以上のように、本実施の形態によれば、同定処理には基本的に声道形状関数を用いる。従って、音声のみを用いて話者の同定処理をする場合と比較して、音声と声道形状との間の対応関係のあいまいさに起因する識別処理のあいまいさが小さくなり、より確実な同定処理を行なうことが可能になる。また、同定には同定処理対象者の音声を用いるので、特別な機器は必要なく、通常の音声入出力を持つコンピュータで実現できる。同定には主成分係数を用いるので、声道面積関数を用いる場合と比較して回帰係数記憶部36及び参照データベース38に記憶させるべきデータ量も少なく済み、簡略なシステムで利用者の同定処理を行なうことができる。

【0083】

なお、上記実施の形態では、音声データから得たケプストラム係数を用い、声道形状関数の主成分係数を得るための重回帰分析を行なっている。しかし本発明はそのような実施の形態には限定されず、音声データから得られる別の音響特徴量を用いてもよい。また、算出されたケプストラム係数をそのまま重回帰分析に用いるのではなく、一旦ケプストラム係数に対し主成分分析を行ない、ケプストラム係数の平均値とケプストラム係数の主成分とを求め、さらに各話者から得たケプストラム係数を主成分の線形和で近似するための主成分係数に変換してから重回帰分析を行なうようにしてもよい。この場合、同定処理時にも同定処理対象者の音声に対し同様の処理をする必要がある。

【0084】

ケプストラム係数に対する主成分分析を行なう場合、照合処理には主成分係数を使用するだけである。従って、ケプストラム係数をそのまま使用する場合と比較して、処理が簡略化される。

【0085】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制

10

20

30

40

50

限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味及び範囲内のすべての変更を含む。

【図面の簡単な説明】

【0086】

【図1】本発明の一実施の形態に係る発話者同定システム20の概略構成を示すブロック図である。

【図2】図1に示す同定用データ作成装置34のより詳細なブロック図である。

【図3】図1に示す同定処理装置42のより詳細なブロック図である。

【図4】複数の母音と複数の話者とに対する声道形状関数の主成分分析結果を示す図である。 10

【図5】同定処理の原理を示す図である。

【図6】同定用データ作成装置34をコンピュータで実現するためのプログラムのフローチャートである。

【図7】同定処理装置42をコンピュータで実現するためのプログラムのフローチャートである。

【図8】発話者同定システム20を実現するコンピュータシステム210の概観を示す図である。

【図9】図8に示すコンピュータシステム210の構成を示すブロック図である。 20

【符号の説明】

【0087】

20 発話者同定システム

32 音声データ記憶部

34 同定用データ作成装置

36 回帰係数記憶部

38 参照データベース

40 同定処理対象者

42 同定処理装置

44 同定結果

60 PCA分析部 30

62 主成分記憶部

64 ケプストラム算出部

66 ケプストラム係数記憶部

68 重回帰分析処理部

80 音声データ

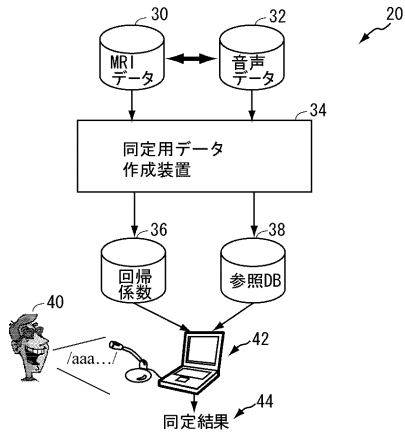
82 ケプストラム算出部

84 主成分係数推定部

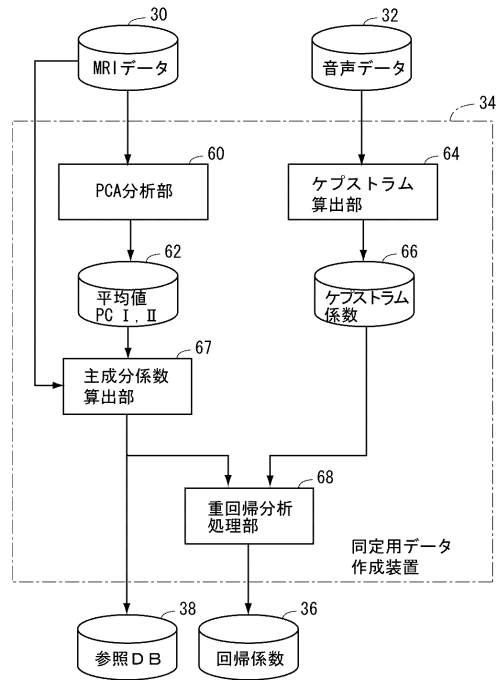
86 距離算出部

88 最短距離判定部 40

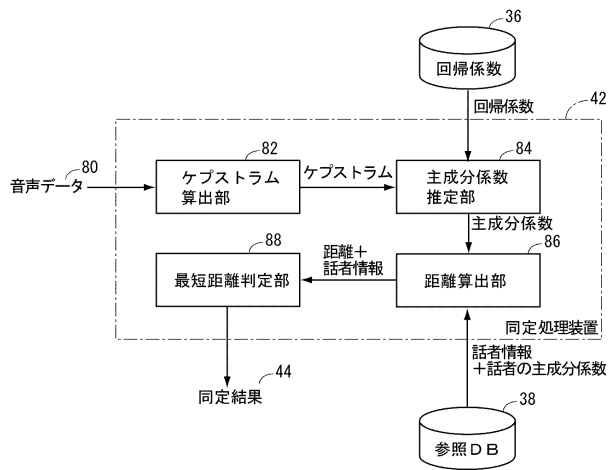
【図 1】



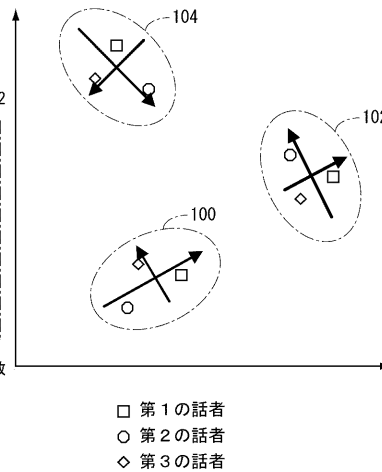
【図 2】



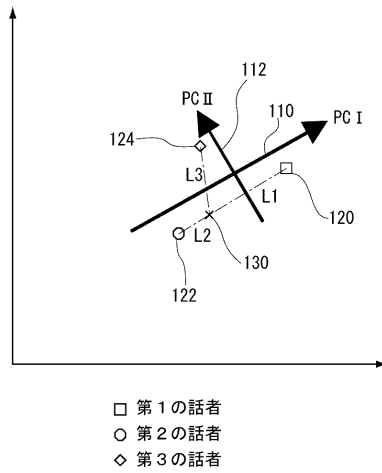
【図 3】



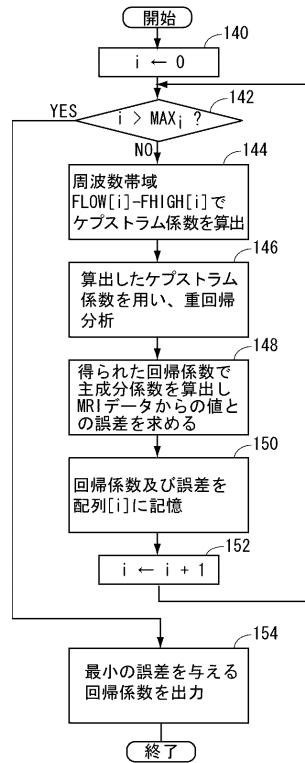
【図 4】



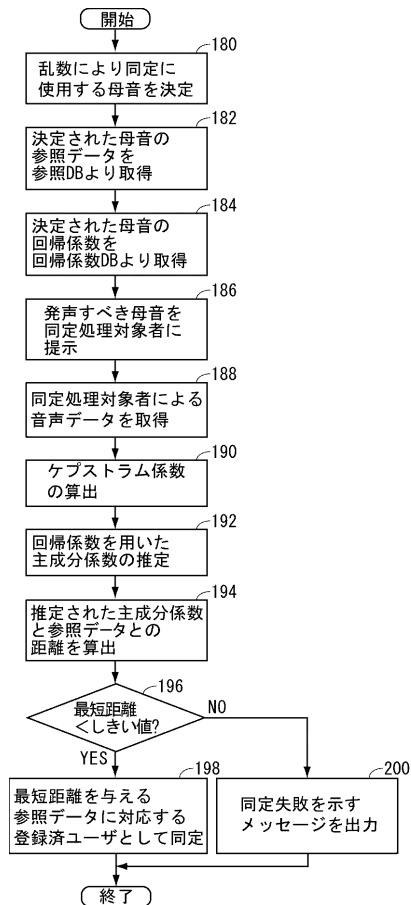
【図 5】



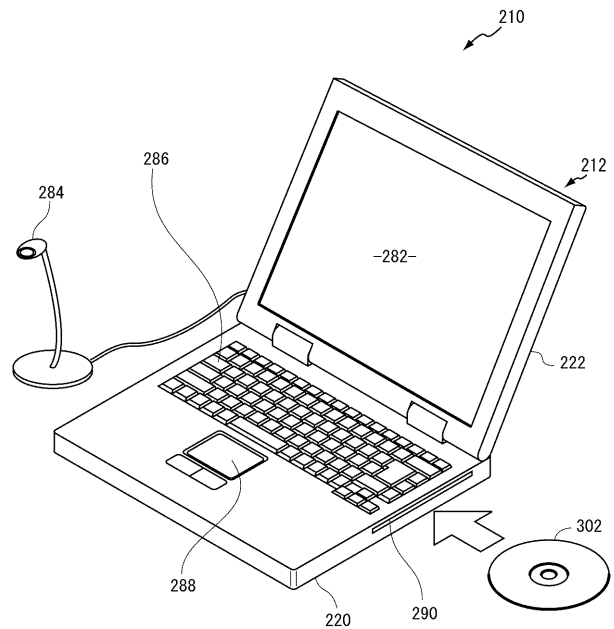
【図 6】



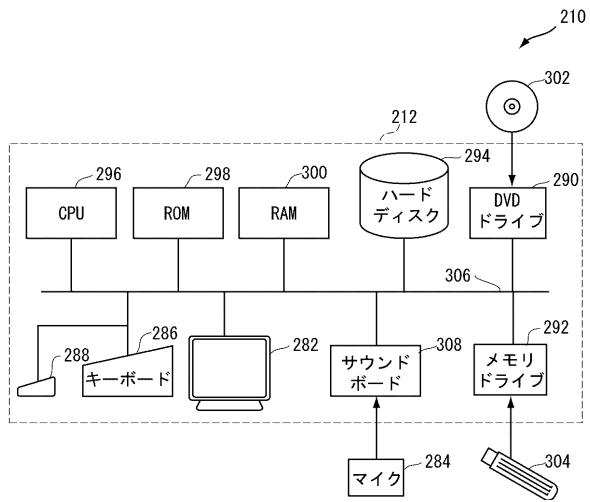
【図 7】



【図 8】



【図 9】



フロントページの続き

(72)発明者 足立 整治

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

(72)発明者 本多 清志

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 毛利 太郎

(56)参考文献 特開2005-148640(JP, A)

特表平05-508242(JP, A)

特開2001-249675(JP, A)

北村 達也 Tatsuya Kitamura, 母音発声時の声道断面積関数の個人差について Individualities in vocal tract area functions during vowel production., 日本音響学会2004年春季研究発表会講演論文集-I- THE 2004 SPRING MEETING OF THE ACOUSTICAL SOCIETY OF JAPAN, 日本, 社団法人日本音響学会, 2004年 3月17日, p.285-286

本多 清志 Kiyoshi Honda, 下咽頭腔の共鳴特性 Resonance characteristics of hypopharyngeal cavities, 日本音響学会2004年秋季研究発表会講演論文集-I- THE 2004 AUTUMN MEETING OF THE ACOUSTICAL SOCIETY OF JAPAN, 日本, 社団法人日本音響学会, 2004年 9月21日, p.235-236

坂寄 洋介 Y. Sakayori, 声道形状の非一様性を考慮した女声・男声変換に関する検討 Female-to-male voice conversion taking non-uniformity of vocal tract shape into account., 日本音響学会2004年秋季研究発表会講演論文集-I- THE 2004 AUTUMN MEETING OF THE ACOUSTICAL SOCIETY OF JAPAN, 日本, 社団法人日本音響学会, 2004年 9月, p.293-294

(58)調査した分野(Int.Cl., DB名)

G10L 17/00

G10L 15/00-15/28