(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3581401号
(P3581401)
(45)発行日　平成16年10月27日(2004.10.27)　(24)登録日　平成16年7月30日(2004.7.30)

(51)Int.Cl.[7]　F I
G10L 15/14　G10L 3/00　535C

請求項の数 4　(全 11 頁)

(21)出願番号　特願平6-244005
(22)出願日　平成6年10月7日(1994.10.7)
(65)公開番号　特開平8-110791
(43)公開日　平成8年4月30日(1996.4.30)
審査請求日　平成13年10月5日(2001.10.5)

(73)特許権者　000001007
キヤノン株式会社
東京都大田区下丸子３丁目３０番２号
(74)代理人　100090538
弁理士　西山　恵三
(74)代理人　100096965
弁理士　内尾　裕一
(72)発明者　小森　康弘
東京都大田区下丸子３丁目３０番２号キヤノン株式会社内
(72)発明者　山田　雅章
東京都大田区下丸子３丁目３０番２号キヤノン株式会社内

最終頁に続く

(54)【発明の名称】音声認識方法

(57)【特許請求の範囲】
【請求項１】
少なくとも２種類以上のＨＭＭセットを有し、１つのＨＭＭセットは計算量の少ない粗いＨＭＭセットであり、その他は該ＨＭＭセットより計算量が多い詳細なＨＭＭセットであり、入力音声に対し前者のＨＭＭセットを用いて尤度を計算し、該尤度計算結果が上位のものについては後者のＨＭＭセットを用いて尤度を計算する計算方法を用いて音声認識を行う音声認識方法であって、
前記２種類以上のＨＭＭセットのＨＭＭは、異なるセット間の同一クラス内の状態構造が同一でかつ複数の状態からなり、ＨＭＭセットごとに、複数の状態が連結された状態で独立して学習されたＨＭＭセットとして作成し、ＨＭＭの状態構造を用いて、同一クラス内における異なるセット間のＨＭＭの状態間の尤度の対応をとり、ＨＭＭの状態の尤度計算を行うことを特徴とする音声認識方法。
【請求項２】
前記粗いＨＭＭセットは分布数の少ないＨＭＭとして学習し、前記詳細なＨＭＭセットは分布数の多いＨＭＭとして学習することを特徴とする請求項１記載の音声認識方法。
【請求項３】
前記粗いＨＭＭセットは音素などでグループ化した音素環境独立ＨＭＭとして学習し、前記詳細なＨＭＭセットは音素環境依存ＨＭＭとして学習することを特徴とする請求項１記載の音声認識方法。
【請求項４】

少なくとも２種類以上のＨＭＭセットを有し、１つのＨＭＭセットは計算量の少ない粗いＨＭＭセットであり、その他は該ＨＭＭセットより計算量が多い詳細なＨＭＭセットであり、入力音声に対し前者のＨＭＭセットを用いて尤度を計算し、該尤度計算結果が上位のものについては後者のＨＭＭセットを用いて尤度を計算する計算方法を用いて音声認識を行う音声認識装置であって、
前記２種類以上のＨＭＭセットのＨＭＭは、異なるセット間の同一クラス内の状態構造が同一でかつ複数の状態からなり、ＨＭＭセットごとに、複数の状態が連結された状態で独立して学習されたＨＭＭセットとして作成し、ＨＭＭの状態構造を用いて、同一クラス内における異なるセット間のＨＭＭの状態間の尤度の対応をとり、ＨＭＭの状態の尤度計算を行うことを特徴とする音声認識装置。

【発明の詳細な説明】
【０００１】
【産業上の利用分野】
本発明は、入力した未知音声を認識する音声認識方法に関するものである。
【０００２】
【従来の技術】
近年、ｗｏｒｋｓｔａｔｉｏｎやｐｅｒｓｏｎａｌ　ｃｏｍｐｕｔｅｒの計算能力が劇的に高速化したのに伴い、高速な音声認識システムがｓｏｆｔｗａｖｅのみで構築できるようになった。これらの音声認識システムの多くは、不特定話者・大語彙・連続音声の認識を目指しており、その多くは、現時点で最も有効と考えられている統計的なモデルであるＨＭＭを採用している。しかしながら、計算機が劇的に高速化したとは言え、不特定話者・大語彙・連続音声の認識を実時間処理する音声認識システムをｓｏｆｔｗａｖｅのみで構築するにはまだ充分とは言えず、音声認識処理の高速化という課題は、いまでも重要である。
【０００３】
【発明が解決しようとする課題】
ＨＭＭを用いた音声認識処理を高速に行なうためには、１）音響処理、２）出力確率計算、３）言語空間探索の３部分の高速化を図る必要がある。特に、最近、盛んに研究されている混合連続出力確率分布型ＨＭＭ（以下、連続ＨＭＭ）を用いた音声認識システムでは、離散出力確率分布型ＨＭＭを用いたものとの比較において、出力確率の計算量が非常に大きくなる。従って、連続ＨＭＭを用いた不特定話者・大語彙・連続音声認識システムを実時間で実用化するには、その出力確率の計算量の削減が、高速化の為には重要な課題となる。
【０００４】
連続ＨＭＭの出力確率の計算量は、総分布数（総分布数＝ＨＭＭの数×ＨＭＭの状態数×状態内混合分布数）に比例する。一般に、連続ＨＭＭを用いた音声認識では、学習データが充分であれば、分布数が多いほど認識率が高いが計算量も多く、逆に、分布数が少ないほど計算量が少ないが認識率が低いことが知られている。従って、認識率を低下させずに、実際に計算する総分布数をいかに削減するかが問題である。
【０００５】
連続ＨＭＭの出力確率計算の高速化を実現するためには「出力確率の値が大きくなるクラスタに対してはできるだけ正確に、小さくなるクラスタに対しては大まかに行なう。」という考え方がこの問題を解決するポイントとなる。
【０００６】
【課題を解決するための手段】
本発明は、小計算量の少数分布のＨＭＭを用いて認識に貢献度の高そうなＨＭＭの状態を推定し、推定された状態のみを高い認識率が得られる多数分布のＨＭＭの状態を用いて出力確率を再計算し、実際に計算する総分布数の削減を図る。
【０００７】
本発明は、同クラスのＨＭＭにおいて、少なくとも２種類のＨＭＭを用いる。

【０００８】
本発明は、その１種が少量計算で認識率もほどほどの簡単なＨＭＭ（例えば、少数の分布からなるＨＭＭ）と、その他の種のＨＭＭが認識率が高い詳細なＨＭＭ（例えば、計算量が大きい多数の分布からなるＨＭＭ）を用いる。本発明は、両ＨＭＭの構造（トポロジー）が同一であり、かつ、各々の状態ごとに対応がとれ、それらの出力確率をとりかかえる。
【０００９】
本発明は、簡単なＨＭＭで出力確率を計算し、音声認識に貢献しそうなクラスやその状態・分布を計算する際、この出力確率の大きい順にＮ個、もしくは最大出力確率より閾値内であるものが、音声認識に貢献しそうなクラスやその状態・分布であると判断する。
【００１０】
本発明は、詳細なＨＭＭを用いて正確に出力確率を再計算する際、上位Ｍ個を音声認識に貢献すると判断し、詳細なＨＭＭで出力確率を計算したり、場合によっては、さらに、その中のＮ個（Ｎ＜Ｍ）を同様にさらに詳細なＨＭＭを用いて、より正確な出力確率を多段で求める。
【００１１】
本発明は、詳細なＨＭＭを用いて正確に出力確率を再計算する際、上位Ｎ個からＭ個（Ｎ＜Ｍ）を比較的詳細なＨＭＭで出力確率を計算し、加えて、上位Ｎ個は同様にさらに詳細なＨＭＭを用いてより正確な出力確率を２段で求める。
【００１２】
本発明は、同音素のＨＭＭにおいて、音素環境型独立ＨＭＭとその音素に属する複数の音素依存型ＨＭＭの２種類のＨＭＭを用いる。
【００１３】
本発明は、それらのＨＭＭの構造（トポロジー）が同一であり、かつ、各々の状態ごとに対応がとれ、それらの出力確率をとりかかえる。
【００１４】
本発明は、音素環境独立型ＨＭＭで出力確率を計算し、音声認識に貢献しそうな音素やその状態・分布を計算する際、この出力確率の大きい順にＮ個、もしくは最大出力確率より閾値内であるものが、音声認識に貢献しそうな音素やその状態・分布であると判断する。
【００１５】
本発明は、音素環境依存型ＨＭＭを用いて正確に出力確率を再計算する際、上位Ｎ個を音声認識に貢献すると判断し、音素環境依存型ＨＭＭで出力確率を計算し、正確な出力確率複数で求める。
【００１６】
【実施例】
（実施例１）
図１は本発明の実施例の音声認識装置の一構成図を示す。
【００１７】
１０１は未知の音声を入力するマイクやＡ／Ｄを含む音声入力部、１０２は音声入力部１０１より入力された音声から音声パラメータを求める音響分析部、１０３は音響処理部１０２で求められた音声パラメータと認識に用いる辞書データとなるＨＭＭを比較することにより出力確率を計算する出力確率計算部、１０４は認識に用いる辞書部となる出力確率を計算するためのＨＭＭ格納部、１０５は入力音声の認識にその音素がどれくらい貢献するかを示す貢献度を予測するための簡単なＨＭＭであるＨＭＭ$_S$を格納したＨＭＭ$_S$格納部、１０６は正確な出力確率を計算する詳細なＨＭＭであるＨＭＭ$_L$を格納したＨＭＭ$_L$格納部、１０７は言語処理を行なう言語探索部、１０８は言語処理に用いる文法を格納した文法辞書、１０９は音声認識の結果を文字列で出力する表示部である。
【００１８】
音声認識装置は、上記の要素により構成され、図２に示す流れにしたがって動作する。２０１（音声入力装置１０１に対応）の音声入力部で切り出された音声は、２０２（１０２

に対応）の音響分析部にて音声パラメータに分析され、２０３（１０３に対応）の出力確率計算部において、２０４（ＨＭＭ格納部１０４に対応）のＨＭＭの中の２０５（$ＨＭＭ_S$格納部１０５に対応）の簡単な$ＨＭＭ_S$を用いて出力確率を計算し、さらに、認識に貢献しそうな音素のＨＭＭの状態を求め、それら認識に貢献すると判断された音素の状態に対応する２０６（$ＨＭＭ_L$格納部１０６に対応）の詳細な$ＨＭＭ_L$を用いて正確な出力確率を再計算し、それらを用いて、２０７（１０７に対応）の言語探索部において、２０８（１０８に対応）の文法辞書を用いて文法的に確認し、２０９（表示部１０９に対応）に認識結果を出力する。

【００１９】

以下、実施例の詳細を示す。

【００２０】

本実施例は、認識に貢献度の高そうなクラスを分布数が少ない$ＨＭＭ_S$（少数分布ＨＭＭ）を用いて推定し、推定されたクラスだけを認識率が高いが分布数が多く、計算量の多い$ＨＭＭ_L$（多数分布ＨＭＭ）を用いて再計算することにより、全音素を$ＨＭＭ_L$で計算する必要がなくなり、実際に計算する総分布数の削減を図り、高速な音声認識を可能とする。

【００２１】

この方法は、ＨＭＭが１状態であれば非常に簡単に適用できる。しかし、実際問題、音声認識に用いられるＨＭＭはモデル内における特徴系列の時系変化を表現するために、１状態のＨＭＭはあまり用いられない。また、１状態のＨＭＭでは良い性能が得られないことも知られている。良い性能が得られる適当な状態数としては、一般には、ＨＭＭは複数状態（３～４程度）のＨＭＭが多く用いられている。ＨＭＭが複数状態である場合、状態出力確率と状態遷移確率で決まるそのクラスの尤度は、各状態の継続時間により変わってしまい、貢献度の高そうなＨＭＭを決定するための比較を均等にすることは簡単でなくなる問題が生じる。そこで、以下に示す仮定を行なうことにより、上述のクラスをＨＭＭの状態とし、考え方はそのまま適用できるようにした。

クラスを音素のような比較的短い単位とし、この短い単位でＨＭＭを学習する場合、ＨＭＭの構造をｌｅｆｔ－ｔｏ－ｒｉｇｈｔに統一し、かつ、状態数が３個程度であれば、分布数の異なるＨＭＭをそれぞれ独立に学習しても、ＨＭＭの各状態が学習している区間がほぼ等しくなり、このため、各状態が表現している特徴はほぼ等しい。この結果、異なる分布数であっても、それぞれのＨＭＭの状態に対応がとれる。

【００２２】

この仮定が成り立つならば、少数分布ＨＭＭを用いて出力確率を計算し、確率が高くて認識に貢献度の高そうなＨＭＭの状態を推定し、推定されたＨＭＭの状態に対応する多数分布のＨＭＭの状態を用いて、出力確率を再計算することが可能となる。また、再計算されない状態については、少数分布ＨＭＭより求めた推定値をそのまま出力確率として利用できる。

【００２３】

具体的なアルゴリズムを以下に示し、その説明のイメージ図を図３に、その認識過程の処理の流れを図４に示す。

【００２４】

学習過程

音声データベース（音素ラベルと特徴パラメータ）を用いて、音素毎に分布数の異なる２種類のＨＭＭ（少数分布$ＨＭＭ_S$と多数分布$ＨＭＭ_L$）をＥＭ－ａｌｇｏｒｉｔｈｍによりそれぞれ学習する。少数分布$ＨＭＭ_S$の分布数をｓ、多数分布$ＨＭＭ_L$の分布数をｌとする（ｓ＜ｌ）。この時の２種類のＨＭＭの構造（例：ｎ状態ｎループ、本実施例においてはｎ＝３）は同一である。

【００２５】

認識過程

時刻ｔにおける特徴パラメータＯ（ｔ）に対して、全音素（例：Ｍ音素）の全状態Ｊ（Ｊ

＝Ｍ・ｎ）について、各状態ｊ（ｊ＝１，２，・・・，Ｊ）毎の出力確率$Ｐ_{j}$（Ｏ（ｔ））を求める（ステップ２０３１）。この際、状態ｊの出力確率$Ｐ_{j}$（Ｏ（ｔ））は、少数分布$ＨＭＭ_{s}$の状態（ｊ）を用いて、出力確率

【００２６】

【外１】

$$P_j^s(O(t))$$

を計算する。

【００２７】

それら出力確率

【００２８】

【外２】

$$P_j^s(O(t))$$

の上位のＮ個が認識に貢献すると仮定し、上位Ｎ個に対応する状態ｊを求める（ステップ２３０２）。

【００２９】

それら状態ｊに対応する多数分布$ＨＭＭ_{L}$の状態ｊのみを用いて出力確率

【００３０】

【外３】

$$P_j^L(O(t))$$

を再計算する（ステップ２３０３）。

【００３１】

多数分布$ＨＭＭ_{L}$より求めたＮ個の出力確率

【００３２】

【外４】

$$P_j^L(O(t))$$

とそれ以外（Ｊ－Ｎ個）の少数分布$ＨＭＭ_{S}$より求めた出力確率

【００３３】

【外５】

$$P_j^s(O(t))$$

を各状態（ｊ）の出力確率$Ｐ_{j}$（Ｏ（ｔ））として用い、認識処理を行なう。従って、必要とされる計算量の総分布数Ｃは、以下の式で決まる。

【００３４】

Ｃ＝Ｍ・ｎ・ｓ＋Ｎ・ｌ

例えば、従来の一種ＨＭＭでの認識方法によれば３状態２分布（総分布数９００、第１位文認識率７９．０％）や３状態１８分布（総分布数１３５０、第１位文認識率８１．０％）であるが、本実施例によれば、３状態４分布で予測し、３状態１８分布で１０状態を再計算した場合にも、ほぼ同程度の認識率が得られるが、総分布数は４８０に、また、３状態６分布で予測し３状態６分布で予測し、３状態２４分布で１０状態を再計算した場合も、ほぼ同程度の認識率が得られるが、総分布数は６９０に、というように、計算量がほぼ半減してもほぼ同程度の認識率が得られる。

【００３５】
尚、本実施例ではＨＭＭの構造を３状態に統一した例について説明したが、先にも述べた様に、ある程度良い認識率と、適度な計算量となる、３～４状態付近であれば、本発明は３状態に限定されるものではない。
【００３６】
尚、本実施例では、分布数を少なくした簡単なＨＭＭと分布数を多くした詳細なＨＭＭの２種類を用いているが、これが多段階に分布数を変化させた複数種のＨＭＭを用いても良いことは勿論である。
【００３７】
尚、本実施例では、簡単なＨＭＭと詳細なＨＭＭの２種類を用いているため、出力確率の計算が二段で行なわれているが、複数種のＨＭＭを用いた場合に、ＨＭＭの種類数に合わせて出力確率の再計算を多段で行なうようにしても良いことは勿論である。
【００３８】
尚、本実施例では、簡単なＨＭＭと詳細なＨＭＭの２種類を用いているため、出力確率の再計算が上位Ｎ個のみ行なわれているが、例えば簡単なＨＭＭと、比較的詳細なＨＭＭと、さらに詳細なＨＭＭというような数種のＨＭＭを用いた場合に、簡単なＨＭＭで全状態の出力確率を計算し、この出力確率の上位Ａ位からＢ位（Ａ＜Ｂ）の（Ｂ－Ａ＋１）個を比較的詳細なＨＭＭで出力確率を再計算し、加えて、上位Ａ－１個は同様にさらに詳細なＨＭＭを用いてさらに正確な出力確率を再計算するようにしても良いことは勿論である。
【００３９】
また、簡単なＨＭＭと、比較的詳細なＨＭＭと、より詳細なＨＭＭというような数種のＨＭＭを用いた場合に、まず簡単なＨＭＭを用いて全状態の出力確率を計算し、その出力確率の上位Ｂ個の状態を比較的詳細なＨＭＭで出力確率を再計算し、この再計算された出力確率の中での上位Ａ個（Ａ＜Ｂ）の状態をより詳細なＨＭＭを用いて出力確率を再々計算し、これらの各ＨＭＭを用いて求めた出力確率を用いて認識処理を行うようにしても良いことは勿論である。
【００４０】
（実施例２）
図５は本実施例音声認識装置の一構成図を示す。先に説明した実施例１と同様の構成は図１と同番号を付し、ここでの説明は省略する。
【００４１】
図１の構成と本実施例における装置の構成と異なるのは、ＨＭＭ格納部１０４内に貢献度を予測するための音素環境独立型ＨＭＭであるＨＭＭ$_I$格納部２１０５と、正確な出力確率を計算する音素環境依存型ＨＭＭであるＨＭＭ$_D$を格納するＨＭＭ$_D$格納部２１０６とを備えている点である。
【００４２】
音声認識装置は、上記の構成を備え、図６に示す流れにしたがって動作する。ここでも、実施例１における図２と同処理を行う構成は同番号を付し、説明は省略する。本実施例においては、２０３（１０３に対応）の出力確率計算部において、２０４（ＨＭＭ格納部１０４に対応）のＨＭＭの中の２０５（ＨＭＭ$_I$格納部２　１０５に対応）の音素環境独立型ＨＭＭ$_I$を用いて出力確率を計算し、さらに、この計算された出力確率の高い、つまりは認識に貢献しそうなＨＭＭの状態を求め、それら状態に対応する２０６（ＨＭＭ$_D$格納部２　１０６に対応）の音素環境依存型ＨＭＭ$_D$を用いて正確な出力確率を再計算し、それらを用いて、２０７（１０７に対応）の言語探索部において、２０８（１０８に対応）の文法・辞書を用い、認識処理を行ない２０９（１０９に対応）の認識結果を出力する。
【００４３】
以下、実施例の詳細を示す。
【００４４】
本実施例は、認識に貢献度の高そうな音素を音素独立型ＨＭＭを用いて推定し、推定された音素だけを認識率の高い音素環境依存型ＨＭＭを用いて再計算し、実際に計算する総分

布数の削減を図り、高速に音声認識を行なうものである。
【００４５】
本実施例においても、実施例１と同様の理由により、３状態のＨＭＭを用いる。
【００４６】
具体的なアルゴリズムを以下に示し、その説明のイメージ図を図７に、その認識過程の処理の流れを図８に示す。
【００４７】
学習過程
音声データベース（音素ラベルと特徴パラメータおよび音素環境依存ラベルと特徴パラメータ）を用いて、音素環境独立型HMM_I（ｍ個）と幾つか音素環境依存型HMM_D（Ｍ個）（ｍ＜＜Ｍ）をＥＭ－ａｌｇｏｒｉｔｈｍによりそれぞれ学習する。この時の２種類のＨＭＭの構造（例：ｎ状態ｎループ、本実施例においてはｎ＝３）は同一である。
【００４８】
認識過程
時刻ｔにおける特徴パラメータＯ（ｔ）に対して、全音素（例：ｍ音素）の全状態ｊ（ｊ＝ｎ・ｍ）について、各状態ｉ（ｉ＝１，２，・・・，ｊ）毎の出力確率P_i（Ｏ（ｔ））を求める（ステップ２０３１）。音素環境依存型ＨＭＭ（Ｍ個）の全状態はｋ＝ｎＭ個（ｊ＜＜ｋ）である。
この際、状態ｉの出力確率P_i（Ｏ（ｔ））は、音素環境独立型HMM_Iの状態（ｉ）を用いて、出力確率
【００４９】
【外６】

$$P_j^I(O(t))$$

を計算する。
【００５０】
それら出力確率
【００５１】
【外７】

$$P_j^I(O(t))$$

の上位のｌ（ｌ≧１）個（ｌ＜ｊ）が認識に貢献すると仮定し、上位ｌ個に対応する状態ｉを求め（ステップ２０３２）、それに対応する音素環境依存型HMM_Dの状態Ｌ（ｐ個、ｋ＞ｐ＞１）を求める（ステップ２０３５）。この際、言語探索処理上で照合する必要のある音素環境依存型ＨＭＭＤのみを考慮する。
【００５３】
音素環境依存型HMM_Dの状態Ｌのみを用いて出力確率
【００５４】
【外９】

$$P_L^D(O(t))$$

を再計算する（ステップ２０３６）。
【００５５】
音素環境依存型HMM_Dより求めたｐ個の出力確率
【００５６】
【外１０】

$P_L^p(O(t))$

とそれ以外（ｋ－ｐ個）の出力確率（これらは、音素環境独立型ＨＭＭ$_I$より求めた出力確率を

【００５７】

【外１１】

$P_i^I(O(t))$

の対応する音素環境依存型ＨＭＭの状態の出力確率に代入する。）を、出力確率Ｐ$_i$（Ｏ（ｔ））として用い、認識処理を行なう。従って、計算される総状態数Ｃは、以下の式で決まる。

【００５８】

Ｃ＝ｉ＋ｐ（Ｃ＜＜ｋ）

【００５９】

【発明の効果】

本発明により、連続ＨＭＭを用いた音声認識方式および装置において、認識率を下げることなく、出力確率計算を簡単に高速化でき、また更に以下のメリットをもつ。

・従来通りの方法でＨＭＭが簡単に学習できる。

・実時間向き処理で、従来のＶｉｔｅｒｂｉ　Ｓｅａｒｃｈ等に簡単に組み込むことができる。

・環境依存型ＨＭＭのみを用いた時と同等の性能が得られる可能性がある。

・状態数の計算量が簡単に決まる。

・その組み合せによっては、大幅に計算する総状態数を削減できる。

・オーバーホッド（出力確率の高い順にｌ個求める計算と音素環境独立型ＨＭＭにより計算したｌ個の推定計算）がほとんど無い。

【図面の簡単な説明】

【図１】本発明による実施例１の構成図

【図２】本発明による実施例１の処理の流れ図

【図３】実施例１のイメージ図

【図４】実施例１の出力計算法の図

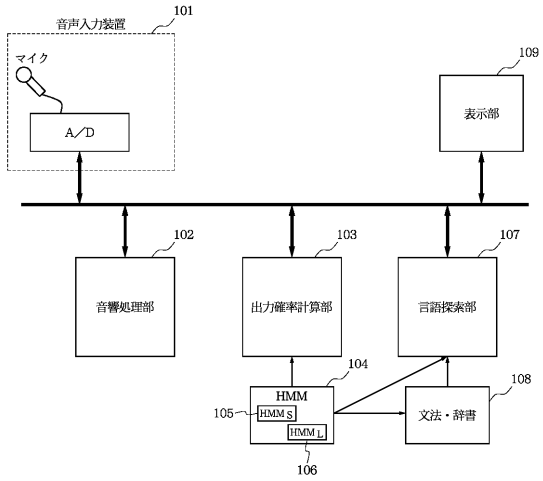
【図５】実施例２の構成図

【図６】実施例２の処理の流れ図

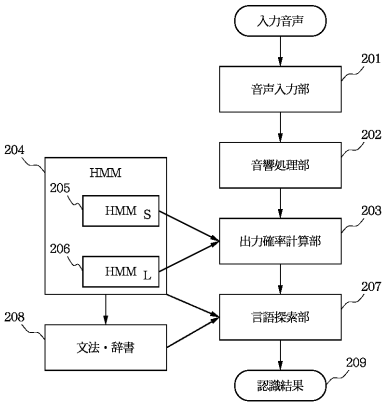
【図７】実施例２のイメージ図

【図８】実施例２の出力計算法の図

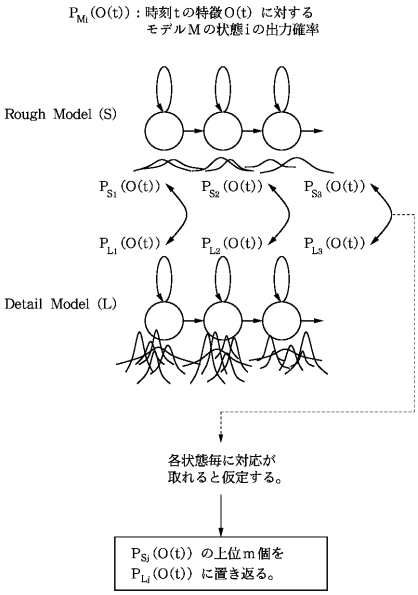
【図１】


101
音声入力装置
マイク
A／D
109
表示部
102
音響処理部
103
出力確率計算部
107
言語探索部
104
HMM
105
HMM S
HMM L
106
108
文法・辞書

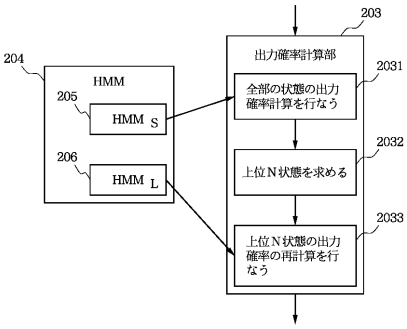
【図２】


入力音声
201
音声入力部
202
音響処理部
204
HMM
205
HMM S
206
HMM L
203
出力確率計算部
207
言語探索部
208
文法・辞書
209
認識結果

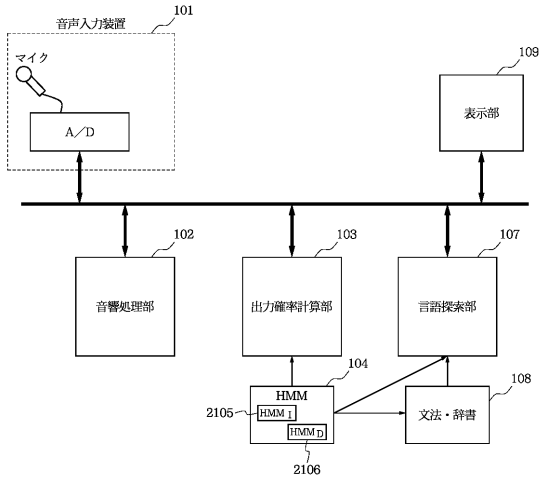
【図３】


$P_{Mi}(O(t))$：時刻tの特徴O(t) に対する
モデルMの状態iの出力確率
Rough Model (S)
$P_{S1}(O(t))$
$P_{S2}(O(t))$
$P_{S3}(O(t))$
$P_{L1}(O(t))$
$P_{L2}(O(t))$
$P_{L3}(O(t))$
Detail Model (L)
各状態毎に対応が
取れると仮定する。
$P_{Sj}(O(t))$ の上位m個を
$P_{Lj}(O(t))$ に置き返る。

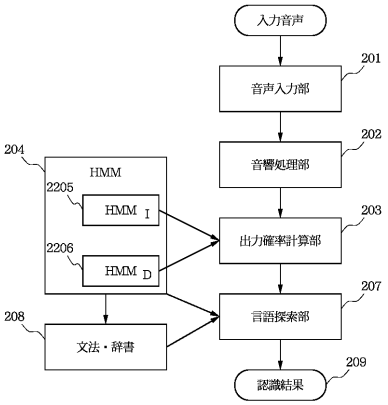
【図４】


203
出力確率計算部
204
HMM
205
HMM S
206
HMM L
2031
全部の状態の出力
確率計算を行なう
2032
上位N状態を求める
2033
上位N状態の出力
確率の再計算を行
なう

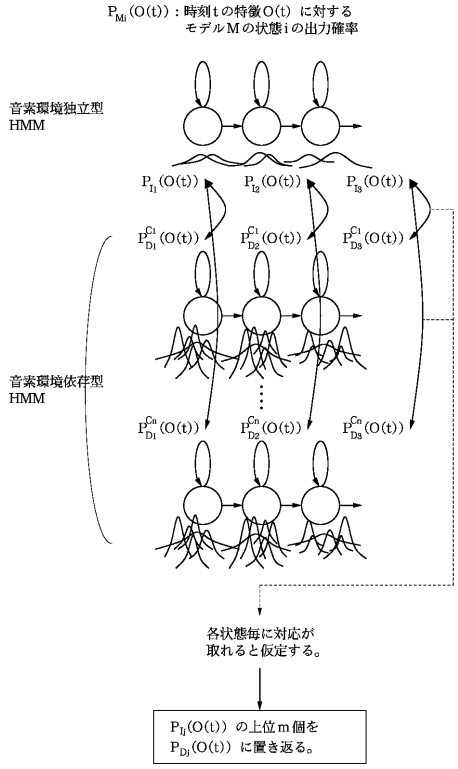
【図５】


音声入力装置
101
マイク
A／D
109
表示部
102
音響処理部
103
出力確率計算部
107
言語探索部
104
HMM
2105
HMM I
HMM D
2106
108
文法・辞書

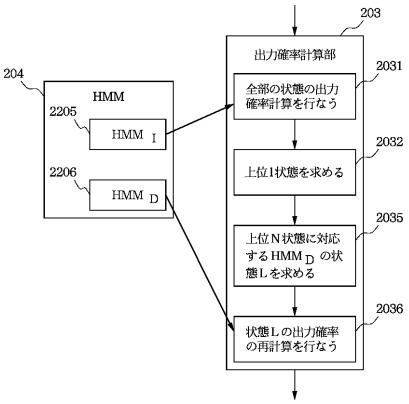
【図６】


入力音声
201
音声入力部
202
音響処理部
204
HMM
2205
HMM I
2206
HMM D
203
出力確率計算部
207
言語探索部
208
文法・辞書
209
認識結果

【図７】


P_Mi(O(t))：時刻tの特徴O(t) に対する
モデルMの状態iの出力確率
音素環境独立型
HMM
音素環境依存型
HMM
各状態毎に対応が
取れると仮定する。
P_Ij(O(t)) の上位m個を
P_Dj(O(t)) に置き換える。

【図８】


203
出力確率計算部
2031
全部の状態の出力
確率計算を行なう
2032
上位1状態を求める
2035
上位N状態に対応
するHMM D の状
態Lを求める
2036
状態Lの出力確率
の再計算を行なう
204
HMM
2205
HMM I
2206
HMM D

フロントページの続き

(72)発明者　大洞　恭則
東京都大田区下丸子３丁目３０番２号キヤノン株式会社内

審査官　渡邊　聡

(56)参考文献　特開平０５－２８１９９４（ＪＰ，Ａ）
特開平０６－１７５６９６（ＪＰ，Ａ）
特開平０７－１６０２８６（ＪＰ，Ａ）
小森　他，少数分布HMMによる出力確率推定に基づいた効率的な混合連続分布HMM音声認識，電子情報通信学会技術研究報告，電子情報通信学会，１９９４年１０月１３日，ＳＰ９４－５２，３１－３８

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G10L 15/14
JICSTファイル(JOIS)