

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2008年9月4日 (04.09.2008)

PCT

(10) 国際公開番号
WO 2008/105263 A1

- (51) 国際特許分類:
G10L 15/10 (2006.01) G10L 15/18 (2006.01)
G10L 15/08 (2006.01)
- (21) 国際出願番号: PCT/JP2008/052721
- (22) 国際出願日: 2008年2月19日 (19.02.2008)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2007-049975 2007年2月28日 (28.02.2007) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1080014 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 江森 正 (EMORI, Tadashi) [JP/JP]; 〒1080014 東京都港区芝五丁目7番

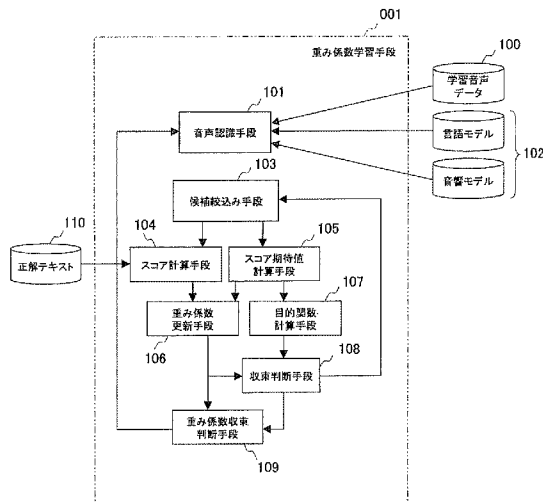
- 1号 日本電気株式会社内 Tokyo (JP). 大西 祥史 (ON-ISHI, Yoshifumi) [JP/JP]; 〒1080014 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 山下 穰平 (YAMASHITA, Johei); 〒1050001 東京都港区虎ノ門五丁目13番1号虎ノ門40MTビル 山下国際特許事務所 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY,

[続葉有]

(54) Title: WEIGHT COEFFICIENT LEARNING SYSTEM AND AUDIO RECOGNITION SYSTEM

(54) 発明の名称: 重み係数学習システム及び音声認識システム

[図1]



- 110 RIGHT ANSWER TEXT
- 001 WEIGHT COEFFICIENT LEARNING MEANS
- 101 AUDIO RECOGNITION MEANS
- 103 CANDIDATE NARROWING MEANS
- 104 SCORE CALCULATION MEANS
- 105 SCORE EXPECTATION VALUE CALCULATION MEANS
- 106 WEIGHT COEFFICIENT UPDATE MEANS
- 107 TARGET FUNCTION CALCULATION MEANS
- 108 CONVERGENCE JUDGMENT MEANS
- 109 WEIGHT COEFFICIENT CONVERGENCE JUDGMENT MEANS
- 100 LEARNING AUDIO DATA
- 102 LANGUAGE MODEL, AUDIO MODEL

(57) Abstract: A weight coefficient learning system includes: audio recognition means for recognizing learning audio data and outputting the recognition result; weight coefficient update means for updating a weight coefficient applied to a score so as to increase the difference between a right answer score calculated by using a right answer text of the learning audio data and the recognition result score with respect to a score obtained by an acoustic model and a language model; convergence judgment means for returning to the weight coefficient update means by using the updated score so as to judge whether to re-update the weight coefficient; and weight coefficient convergence judgment means for returning to the audio recognition means by using the updated score so as to re-execute the processing and judge whether to update the weight coefficient by the weight coefficient update means.

(57) 要約: 重み係数学習システムは、学習音声データを認識しその認識結果を出力する音声認識手段と、音響モデルと言語モデルとから得られるスコアに対し、学習音声データの正解テキストを用いて計算される正解のスコアと、認識結果のスコアとの差が大きくなるようにスコアにかかる重み係数を更新する重み係数更新手段と、更新された後のスコアを用いて、重み係数更新手段に戻り重み係数を再更新するか否かを判断する収束判断手段と、更新された後のスコアを用いて、音声認識手段まで戻りその処理をやり直して重み係数更新手段により重み係数を更新するか否かを判断する重み係数収束判断手段とを備える。

WO 2008/105263 A1



KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

明 細 書

重み係数学習システム及び音声認識システム

技術分野

[0001] 本発明は、音声認識で用いる音響モデルと言語モデルとから得られるスコアの重み係数を学習する重み係数学習システムと、これにより学習された重み係数を用いて音声認識を行う音声認識システムと、これらの方法及びプログラムとに関する。

背景技術

[0002] 大語彙連続音声認識は、音響モデルと、言語モデルとを用いて実現されている。音響モデルは、音声と単語の発音に対する音響的な類似度を表すスコア(以後、「音響スコア」と呼ぶ)を計算するためのものである。言語モデルは、単語の繋がりやすさの程度を表すスコア(以後、「言語スコア」と呼ぶ)を計算するためのものである。理想的な場合、これら音響スコアと言語スコアとの比率は、1:1である。

[0003] しかし、音響モデルで得られる値は、非特許文献1の93ページに示されるように、正規分布のような確率密度分布で近似され、そこから得られる値を用いて近似されている。また、言語モデルは、非特許文献1の192ページに示されるように、先行N-1単語を条件にしたNグラムで近似されている。

[0004] このように、音響モデルと言語モデルとは、共に近似されたモデルを用いていることから、音響スコアと言語スコアとの偏りの整合を取るため、それぞれのスコアに重み係数をかけて実現されている。ここで偏りとは、近似により本来の値よりも大きな値になるなどの現象を指す。音声認識の分野において、これらの重み係数はいくつかの値を用意しておき、テスト用の音声データの認識率を観察しながら、選ぶ方法が取られてきた。このような方法は、音響モデルや言語モデルが1組の場合は問題ないと考えられるが、複数の音響モデルと言語モデルの組や新たなスコアを組み合わせる場合、用意すべきパラメータ数がべき乗のオーダーで大きくなり、計算は不可能であると考ええる。

[0005] このような問題に対し、非特許文献2で示されるように、統計的な機械翻訳の分野において、異なる確率モデルから得られるスコアに対し、最大エントロピー法(以後、「

ME法」と称する)を用いて重み係数を調整する方法が知られている。

[0006] ME法は、非特許文献3の155～174ページで述べられているように、制約条件下でエントロピーを最大にする方法であり、未知データに対し一様な分布関数を推定する学習方式である。この方式では、制約条件として最尤推定を用いた場合、推定される分布関数は式(1)で表されるロジスティック関数になることが知られている。

[0007] [数1]

$$P(w|o) = \frac{\exp\left\{\sum_k \lambda_k f_k(w, o)\right\}}{\sum_w \exp\left\{\sum_k \lambda_k f_k(w, o)\right\}} \quad \text{式(1)}$$

ここで、kはモデルの数(番号)を表す自然数、w、oは出力系列と入力系列であり、非特許文献2では、wを英語の単語系列、oをフランス語の単語系列としている。fk(w, o)は各モデルを用いて計算されるスコアであり、非特許文献2ではf1(w, o)はフランス語の単語から英語の単語が出現する生成確率の対数値、f2(w, o)は英語の単語列の出現する確率値の対数値が用いられている。λkはそれぞれの確率モデルを用いて計算されるスコアの重み係数を表し、正解のwとoの組み合わせにおいて、事後確率P(w | o)が一番大きくなるように最適化される。

[0008] ここで、式(1)の分母は、出力系列wの全ての組み合わせについて足し合わせることを示している。しかし、出力系列wを構成する要素数(非特許文献2の場合、英語の単語の語彙数)が大きくなると、それらの組み合わせは大きくなり、式(1)の分母は計算することができない。非特許文献2のような統計的な機械翻訳の分野では、事前知識として連続して出現しない単語の情報を活用することにより、単語列の組み合わせ数を有限に絞るなどの工夫を行うことで対処している。

非特許文献1:S. Young(ヤング)、他10名著、「The HTK Book for HTK version 3.3(エイチティーケーブック バージョン3.3)」、Cambridge University Engineering Department、April 2005、pp.1-345

非特許文献2:F. J. Och(オーク)、他1名、「Discriminative Training and Maximum Entropy Models for Statistical Machine Translation(統計機械翻訳に対する識別学習と

最大エントロピー法)」、Proc. ACL, July 2002、pp.295-302

非特許文献3:北著、「言語モデルと計算4:確率的言語モデル」、東京大学出版会、1999年

非特許文献4:Lafferty他2著、「Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data」In Proc. of ICML, pp. 282-289, 2001

発明の開示

発明が解決しようとする課題

[0009] 大語彙連続音声認識にME法を適用する場合、式(1)の分母の計算のために全ての単語の組み合わせを計算する場合、語彙数が大きいことに加え、入力系列oに対応する音声の観測時系列の連続量であることから、全ての単語は発声された区間のあらゆる場所に出現することが可能である。このため、同じ単語であっても少しでも出現時刻が違えば同じスコアとはならず、違う単語の組み合わせとして扱うことになるため、その組み合わせは膨大になる。また、膨大な候補から単純に候補を絞った場合、絞られた環境(方法)に依存した重み係数 λ_k が推定されてしまい、本来求めたい重み係数 λ_k に到達しない恐れがある。

[0010] 本発明の目的は、音響モデルと言語モデルとを用いて計算されるスコアにかかる重み係数を、予め範囲などの事前知識が無くとも自動的かつ高速に推定することができ、その重み係数を用いた高精度な音声認識をおこなうシステムを提供することにある。

課題を解決するための手段

[0011] 上記目的を達成するため、本発明に係る重み係数学習システムは、学習音声データを認識しその認識結果を出力する音声認識手段と、音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識手段により出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新手段と、更新された後のスコアを用いて、前記重み係数更新手段に戻り前記重み係数を再更新するか否かを判断する収束判断手段と、前記更新された後のスコアを用いて、前記音声認識手段まで戻りその処理をやり直して前記重み係数更新手段により

前記重み係数を更新するか否かを判断する重み係数収束判断手段と、を備えることを特徴とする。

[0012] 本発明に係る重み係数学習方法は、学習音声データを認識しその認識結果を出力する音声認識ステップと、音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識ステップにより出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新ステップと、更新された後のスコアを用いて、前記重み係数更新ステップに戻り前記重み係数を再更新するか否かを判断する収束判断ステップと、前記更新された後のスコアを用いて、前記音声認識ステップまで戻りその処理をやり直して前記重み係数更新ステップにより前記重み係数を更新するか否かを判断する重み係数収束判断ステップと、を備えることを特徴とする。

[0013] 本発明に係る重み係数学習プログラムは、コンピュータに、学習音声データを認識しその認識結果を出力する音声認識処理と、音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識処理により出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新処理と、更新された後のスコアを用いて、前記重み係数更新処理に戻り前記重み係数を再更新するか否かを判断する収束判断処理と、前記更新された後のスコアを用いて、前記音声認識処理まで戻りその処理をやり直して前記重み係数更新処理により前記重み係数を更新するか否かを判断する重み係数収束判断処理と、を実行させるためのものである。

[0014] 本発明に係る音声認識システムは、上記の重み係数学習システムにより更新された重み係数を用いて音声認識を行うことを特徴とする。

[0015] 本発明に係る音声認識方法は、上記の重み係数学習方法により更新された重み係数を用いて音声認識を行うことを特徴とする。

[0016] 本発明に係る音声認識方法は、コンピュータに、上記の重み係数学習プログラムにより更新された重み係数を用いて音声認識を行う音声認識処理を実行させるための

ものである。

発明の効果

[0017] 本発明によれば、音声認識のための音響モデルと言語モデルとを用いて計算されるスコアにかかる重み係数を、予め範囲などの事前知識が無くても自動的かつ高速に推定することができる。そして、その重み係数を用いた高精度な音声認識をおこなうことができる。

図面の簡単な説明

[0018] [図1]本発明の第1の実施形態に係る重み係数学習システムの構成を示すブロック図である。

[図2]音声認識手段の認識結果として出力される単語グラフを説明するための図である。

[図3]候補絞込み手段が単語グラフから候補を絞り込むためのアルゴリズムであるA*サーチを説明するための図である。

[図4]本発明の第1の実施形態に係る重み係数学習システムの動作を説明するフローチャートである。

[図5]本発明の第2の実施形態に係る音声認識システムの構成を示すブロック図である。

符号の説明

- [0019] 001 重み係数学習手段
100 学習音声データ
101 音声認識手段
102 音響モデル・言語モデル
103 候補絞込み手段
104 スコア計算手段
105 スコア期待値計算手段
106 重み係数更新手段
107 目的関数計算手段
108 収束判断手段

109 重み係数収束判断手段

110 正解テキスト

発明を実施するための最良の形態

[0020] 次に、本発明に係る重み係数学習システム及び音声認識システムを実施するための最良の形態について、図面を参照して詳細に説明する。

[0021] (第1の実施形態)

図1を用いて、音声認識における音響スコアと言語スコアの重み係数学習システムの構成要素について説明する。

[0022] 図1を参照すると、本発明の第1の実施形態に係る重み係数学習システムは、例えばコンピュータシステムを用いて構成されるもので、重み係数学習手段001と、学習音声データ100と、言語モデル・音響モデル102と、正解テキスト110とを備えている。

[0023] 重み係数学習手段001は、音声認識手段101と、候補絞り込み手段103と、スコア計算手段104と、スコア期待値計算手段105と、重み係数更新手段106と、目的関数計算手段107と、収束判断手段108と、重み係数収束判断手段109とを有している。これら各手段は、一例としてコンピュータによるソフトウェアの処理により実現される。この場合、各手段の処理は、個別のプログラムで実現しても、OS(オペレーションシステム)等の他のプログラムと連携して実現してもいずれでもよく、その機能を実現可能なものであればよい。

[0024] 学習音声データ100は、重み係数の値を推定するための音声データで、例えばサンプリング周波数を44.1kHz、1サンプルあたり16ビットでA/D(Analog to Digital)変換されたものである。この学習音声データ100は、例えばハードディスク装置やメモリ等の記憶装置に記憶され、本システムを構成するコンピュータによりアクセス可能となっている。

[0025] 正解テキスト110は、学習音声100の発声内容を書き起こしたテキストデータである。この正解テキスト110は、例えばハードディスク装置やメモリ等の記憶装置に記憶され、本システムを構成するコンピュータによりアクセス可能となっている。

[0026] 音響モデル・言語モデル102は、音声認識に用いる音響モデルと言語モデルであ

る。この音響モデル・言語モデル102は、例えばハードディスク装置やメモリ等の記憶装置に記憶され、本システムを構成するコンピュータによりアクセス可能となっている。

[0027] このうち、音響モデルは、音声の音響的特長を表現した確率モデルであり、例えば非特許文献1の35ページから40ページに書かれている、前後の音素コンテキストを考慮したトライフォンを音素として持つHMM (Hidden Markov Model: 隠れマルコフモデル) が広く使われている。

[0028] 以下、音響モデルの学習について述べる。

[0029] 先ず、音声の音響的特長は、非特許文献1の54ページから64ページに記載されているような、音声データを10msec程度の一定区間を切り出し、プリエンファシス、FFT、フィルタバンクの処理を行った後、コサイン変換を行うことで抽出される。抽出された特徴量に加え、パワーや、前後の時刻の差分を用いることもできる。抽出された特徴とそれに対応する書き起こしテキストを用いて得られたラベルデータを用い、非特許文献1の127ページから130ページに記載されているフォワード・バックワード確率を計算することで、特徴とラベルデータとを対応付ける。

[0030] ここで、ラベルデータとしては、前述のトライフォンなどが考えられる。例えば、「わたくし」と書き起こされた場合、ラベルデータは、「*-w+a w-a+t a-t+a t-a+k a-k+u k-u+s u-s+i s-i+*」ようになる。対応付けられた特徴量とラベルデータとを用いて、音素やさらに細かい状態単位などの区間毎に分布関数のパラメータを計算する。分布関数として正規分布が用いられることが多いが、この場合は、パラメータ特徴量の平均と分散である。

[0031] 言語モデルは、非特許文献3の57ページから62ページに記載されているように、言語モデルをNグラムによる近似で表し、その学習方法は主に最尤推定によって行われる。Nグラムは、言語モデルのモデル化の手法で、履歴のN-1単語を条件としたN個目の単語の出現確率(条件付き確率)を用いて、全単語の出現確率を近似する方法である。出現確率は、最尤推定によると学習コーパスの単語列の頻度を数えることによって計算することができる。例えば、「私」の後に「は」が出現する確率を表す $P(\text{は} | \text{私})$ は、「私」「は」が同時に出現する同時確率 $P(\text{私}, \text{は})$ と、「私」が出現す

る確率 $P(\text{私})$ とを用いて、 $P(\text{は} | \text{私}) = P(\text{私}, \text{は}) / P(\text{私})$ と計算することができる。

- [0032] 音声認識手段101は、学習音声100と音響モデルと言語モデルを用いて、前述のフォワード・バックワード確率の対数値を音響スコアとし、Nグラム対数値を言語スコアとして、両者の和が最も大きいものを認識候補として出力する。出力の形式は、非特許文献1の334～338に掲載されている単語グラフが用いられることが多い。
- [0033] 図2を用いて、単語グラフについて説明する。
- [0034] 図2の左図におけるI1～I5は、単語の始端時刻と終端時刻を定義する点であり、始端時刻を表すものを始端ノード、終端時刻を表すものを終端ノードと呼ぶ。図2の左図にて、単語は、a、b、c、d、e、f、gとしてあらわされており、単語の先頭時刻である始端ノードから単語の終端時刻である終端ノードを結ぶ実線として図示されている。これを単語アークと呼ぶ。始終端ノードの表す時刻がその単語が認識候補として現れた時刻を示す。単語グラフは、図2の左図を右図のように書き下したもので、各始端終端ノード(I1～I5)を用いて単語アークを記述する。記述された単語アークごとに、単語、音響尤度などの情報が記述されている。
- [0035] 候補絞り込み手段103は、音声認識手段101が出力した認識結果である、単語グラフを用いてスコアの高い順にM個の認識候補を選択し、それぞれの認識候補の音響スコアと言語スコア、単語数などの情報を出力する。選択の方法は、例えば、A*サーチ(Aスターサーチ)が使われる。
- [0036] ここで、図2と図3を用いて、候補絞り込み手段103によるA*サーチについて動作を説明する。
- [0037] A*サーチでは、スタックと呼ばれる単語情報を記憶する手段が用意されている。ここで、単語情報とは、単語グラフ上の単語(a～h)とその経路情報および経路のスコアをさす。経路情報とは、単語グラフの終端であるI5のノードに接続する単語、例えばeから、先頭までたどったときに通過した単語 $e \rightarrow c \rightarrow a$ のような経路を表す情報である。
- [0038] 経路のスコアは、この実施例の場合、音響スコアと言語スコアの和とする。経路スコアは、前向きスコアと後ろ向きスコアの和からなる。前向きスコアとは、始端からある単語にたどり着くまでに足された音響スコアと言語スコアの和とし、後ろ向きスコアとは、

終端からある単語スコアにたどり着くまでに足された音響スコアと言語スコアの和とする。例えば、単語aにおける経路スコアを計算する場合、前向きスコアは0で、後ろ向きスコアは単語e、cもしくはg、f、cか、g、dの3通りの経路のいずれかの音響スコアと言語スコアの和として定義される。

[0039] 次に、図3を参照して、A*サーチの動作の詳細を説明する。

[0040] 図3において、Step0では、初期状態として、特に単語の意味を成さない仮の終端(終端)がスタックにつまれている。終端は、図2の単語e、gに繋がっているとす。

[0041] 次いで、Step1では、スタックの先頭を取り出す。ここで、一旦スタックは空になる。次の動作として、終端に繋がる単語e、gをスタックに積む。そのとき、各単語に終端から接続されたとする経路情報と、経路のスコアを付加する。

[0042] 次いで、Step2-1では、スタックを経路のスコアでソートする。図の場合、gが先頭に来たとしている。

[0043] 次いで、Step3-1では、スタックの先頭の単語gを取り出す。このとき一旦、スタックは単語アークeだけになる。単語アークgに繋がる単語f、d、hをスタックに積む。このとき、f、d、hの経路情報として、「g←終端」のように、終端からgを経て到達したことを示す経路情報を付加する。更に、経路スコアを付加する。

[0044] 次いで、Step2-2、Step3-2では、上記と同様なことを繰り返す。Step2-3の処理が終了後に先頭に來ている単語の前に新たな単語が無い場合(単語aのように)、第1位候補として、単語で表されていた単語g、d、aを認識結果として出力する。

[0045] 以後、Step2-n、Step3-n(nは整数)を繰り返し行い、スタックの先頭に來た単語アークの接続先が無い場合、認識結果として単語列を出力する動作を行うことで、第M位候補となる単語列を認識結果として出力することができる。

[0046] スコア計算手段104は、学習音声データ100と正解テキスト110を用いて、後述する式(2)における

[0047] [数2]

$$\sum_i f_k(w_i, o_i)$$

を計算する。

[0048] ここで、 $k=1$ の場合は音響スコアで、 $k=2$ の場合は言語スコアとし、 o_i は i 番目の学習音声データ、 w_i は i 番目の学習音声データに対応する単語系列である。ここで計算される音響スコアは、個々の学習音声データとその単語系列セット o_i と w_i を用いて得られる、前述のフォワード・バックワード確率の対数値である。同様に言語スコアは、単語系列のNグラム確率の対数値である。

[0049] スコア期待値計算手段105は、候補絞込み手段103によって絞られた認識候補と、音響スコアと、言語スコアとを用いて、

[0050] [数3]

$$\sum_i \sum_w p(w|o) f_k(w_i, o_i)$$

を計算する。

[0051] ここで、 $p(w | o_i)$ は、学習音声データ o_i と任意の単語列 w の事後確率で、式(1)で計算することができる。ただし、式(1)の分母は、単語列 w を全ての組み合わせについて計算する必要があるが、スコア期待値計算手段105では、候補絞込み手段103で選択された M 個の候補のみについて計算する。

[0052] 重み係数更新手段106は、音響モデルと言語モデルとから得られるスコアに対し、学習音声データ100の正解テキスト110を用いて計算される正解のスコアと、音声認識手段101により出力される認識結果のスコアとの差が大きくなるようにスコアにかかる重み係数を更新する。具体的には、下記の式(2)を用いて、各モデルのスコアの重み係数を更新する。

[0053] [数4]

$$\lambda_k \leftarrow \lambda_k + \eta \sum_i \left\{ f_k(w_i, o_i) - \sum_w p(w|o) f_k(w_i, o_i) \right\} \quad \text{式(2)}$$

ここで、式(2)の

[0054] [数5]

$$\sum_i f_k(w_i, o_i)$$

は、スコア計算手段104で計算される値であり、

[0055] [数6]

$$\sum_i \sum_w p(w|o) f_k(w_i, o_i)$$

は、スコア期待値計算手段105によって計算される値である。 η は予め適当に設定された値とする。例えば、 $\eta = 0.1$ 程度に設定しておくものとする。

[0056] 目的関数計算手段107は、スコア期待値計算手段105で計算された事後確率 $p(w|o)$ のうち、学習音声データ o に対応する単語系列 w に当たるものを積算する。

[0057] 収束判断手段108は、目的関数計算手段107により積算された事後確率に対し、1回前の更新時に計算された値からの変化量が、ある一定の閾値以下の場合、更新終了の判断を行う。一方、閾値よりも大きい場合、再度候補選択手段103に戻って処理を続ける。ここで変化量は、前回からの差分を今回の値で割ったものであるとする。

[0058] 重み係数収束判断手段109は、重み係数更新手段106により更新された重み係数の値に対し、1回前の更新時で得られた値との差分の比率が閾値以下になった場合、終了の判断を行う。また、閾値よりも大きい場合、再度、音声認識手段101に戻り処理を繰り返すことを判断する。

[0059] 次に、図4を参照して、本実施形態の重み係数学習手段001の動作を説明する。

[0060] まず、音声認識手段101の処理を実行する(ステップSt11)。この処理では、音響モデル・言語モデル102を用いて学習音声データ100の認識を行う。この場合、音響モデルと言語モデルのそれぞれの重み係数 λ_1 、 λ_2 は、初期値として任意の値を用いるが、理想的な場合を想定し、 $\lambda_1 = 1$ 、 $\lambda_2 = 1$ とする。認識を行った後に、認識結果を図2に示される形式の単語グラフを出力する。

[0061] 次いで、単語絞込み手段103の処理を実行する(ステップSt12)。この処理では、単語グラフをA*サーチを用いて探索を行い、M個の単語列候補と、それに付随する音響スコア、言語スコアとを出力する。

[0062] 次いで、スコア計算手段104の処理を実行する(ステップSt13)。この処理では、学習音声データ100と正解テキスト110を用いて、式(2)における

[0063] [数7]

$$\sum_i f_k(w_i, o_i)$$

を計算する。これは、全ての学習音声データ100に対して行われるとする。

[0064] 次いで、スコア期待値計算手段105の処理を実行する(ステップSt14)。この処理では、M個の単語列候補と音響スコアと言語スコアを用いて、

[0065] [数8]

$$\sum_i \sum_w p(w|o) f_k(w_i, o_i)$$

を計算する。これは、全ての学習音声データ100に対して計算されるとする。すなわち、学習音声データ数×M回の計算が行われる。

[0066] 次いで、重み係数更新手段106の処理を実行する(ステップSt15)。この処理では、 $\eta = 0.1$ としてスコア計算手段104で計算された

[0067] [数9]

$$\sum_i f_k(w_i, o_i)$$

と、スコア期待値計算手段105で計算された

[0068] [数10]

$$\sum_i \sum_w p(w|o) f_k(w_i, o_i)$$

とを、式(2)へ適用し、重み係数 λ_k を更新する。

[0069] 次いで、目的関数計算手段107の処理を実行する(ステップSt16)。この処理では、式(1)と、単語列候補、および正解テキスト110を用いて各学習音声データの事後確率を計算し、それらを積算する。

[0070] 次いで、収束判断手段108の処理を実行する(ステップSt17)。この処理では、積算された事後確率に対し、変化率がある一定の閾値、たとえば 1.0×10^{-4} 以下であるか否かを判断する。その結果、閾値よりも大きい場合(NO)は、候補絞り込み手段103の処理(ステップSt12)に戻る。この場合、候補絞り込み手段103は、重み係数更新手

段106で更新された重み係数の値を用いて候補を絞り込む。一方、閾値以下の場合 (YES)は、収束したものとして次段の重み係数収束判断手段109に処理を進める。

[0071] 次いで、重み係数収束判断手段109の処理を実行する(ステップSt18)。この処理では、更新された重み係数の変化率がある一定の閾値以下であるか否かを判断する。その結果、閾値よりも大きい場合 (NO)は、音声認識手段101の処理(ステップSt11)に戻る。この場合、音声認識手段101は、重み係数の値を更新して、再度認識処理を行う。一方、閾値以下の場合 (YES)は、収束したものとして終了の判断を行う。

[0072] 従って、本実施形態によれば、音響モデルや言語モデルなどから計算されるスコアにかかる重み係数の調整において、自動的にそれらの最適な値を調整することができる重み係数学習システムを提供することができる。

[0073] すなわち、本実施形態では、各モデルにより計算されるスコアの重み係数の計算において、任意の初期値から推定を始めても、最適値を推定することができる。このため、それらの重み係数の値を用いることで高精度な音声認識が可能になる。

[0074] その理由は、次のとおりである。式(1)を最大にする場合、正確に分母の計算を行う必要があるが、全ての単語などの出力系列の組み合わせを計算する代わりに、自動的にその組み合わせの選別を行い、その選別された中で最適値を計算する。そのことにより、初期値よりも適した重み係数の値(準最適値と呼ぶ)が推定されることになる。そしてさらに、この準最適値を用いて更に認識処理などを行い、候補を選別し直すことで、より推定に適した選別が行われることになる。その中で最適値を再度推定することで、前回の準最適値よりもより適当な値になっていることが期待できる。

[0075] なお、上記実施形態では、スコア期待値計算手段105において、計算される値

[0076] [数11]

$$\sum_i \sum_w p(w|o) f_k(w_i, o_i)$$

の計算方法として、候補絞り込み手段103で選択されたM個の候補を用いて計算する方法をあげたが、その他の例として、非特許文献4で記述されているフォワード・バックワードアルゴリズムを用いることで、単語グラフから直接計算することも可能である。

[0077] (第2の実施形態)

次に、本発明の第2の実施形態について説明する。本実施形態は、第1の実施形態における音響スコアと言語スコアの重み係数学習システムを用いた音声認識システムに適用したものである。

[0078] 図5を参照すると、本実施形態の音声認識システムは、重み係数学習手段001と、学習音声データ100と、音響モデル・言語モデル102と、音声認識手段103とを有している。それぞれの構成は、第1の実施形態と同様である(音声認識手段103は、重み係数学習手段001内の音声認識手段101の構成と同様である)。

[0079] 次に、本実施形態における動作を説明する。

[0080] まず、重み係数学習手段001により、前述したように、学習音声データ100と、音響モデル・言語モデル102とを用いて、音響スコアと言語スコアにかかる重み係数 $\lambda 1$ 、 $\lambda 2$ を推定する。

[0081] 次に、音声認識手段103により、音響モデル・言語モデル102を用いて、学習音声データ100とは別の認識対象となる音声データの認識を行う。認識の際に、前述と同様の音響スコアと言語スコアが計算されるが、それにかかる重み係数は、重み係数学習手段001が推定した値 $\lambda 1$ 、 $\lambda 2$ を用いる。

[0082] 従って、本実施形態によれば、音響モデルや言語モデルなどから計算されるスコアにかかる重み係数の調整において、自動的にそれらの最適な値を調整することができる重み係数学習システムを用いることにより、調整された重み係数を用いた高精度な音声認識システムを提供することができる。

[0083] (変形例)

以上、本発明の実施形態を詳細に説明したが、本発明は、代表的に例示した上述の実施形態に限定されるものではなく、当業者であれば、特許請求の範囲の記載内容に基づき、本発明の要旨を逸脱しない範囲内で種々の態様に変形、変更することができる。これらの変形例や変更例も本発明の権利範囲に属するものである。

[0084] 例えば、前記の実施形態では、音響モデルと言語モデルの重み係数の推定について記述したが、モデルの数は2つに限らず、複数の音響モデルや複数の言語モデルでも同様の処理でそれぞれのモデルで得られるスコアにかかる重み係数を推定することが可能である。その場合、式(1)や式(2)の k がモデルの数に応じて増えていく

ことになる。また、音響スコアや言語スコアだけに限らず、任意のスコアが導入されても、それにかかる重み係数を設定した上で適用が可能である。

- [0085] 候補絞込み手段103は、M個の認識候補の決め方として、Mが予め定められた一定値にする場合が考えられる。また、同様にM個の決め方として、音響スコアや言語スコア、更に音響スコアと言語スコアの和が閾値よりも大きいものだけを選択することもできる。また、Mの決め方として、音響スコアや言語スコア、または音響スコアと言語スコアの和が1位候補のそれとの差分や比率が閾値以上のものを選択することができる。
- [0086] スコア期待値計算手段105は、M個の候補を用いて期待値を計算する方法だけではなく、単語グラフを用いて期待値を計算する方法に置き換えることも可能である。
- [0087] 収束判断手段108は、重み係数が更新される前のスコアと、重み係数が更新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、重み係数更新手段108に戻り処理するようにしてもよい。
- [0088] また、収束判断手段108は、積算された事後確率の値だけでなく、重み係数更新手段108により更新された重み係数の値を用いて収束の判断を行ってもよい。
- [0089] 例えば、収束判断手段108は、重み係数が更新される前と更新された後との重み係数の差分、および、当該差分を更新される前の値で割った重み係数の変化率が、予め定められた閾値よりも大きい場合に、重み係数更新手段108に戻り処理するようにしてもよい。
- [0090] 重み係数収束判断手段109は、重み係数が更新される前と更新された後との重み係数の差分、および、当該差分を更新される前の値で割った重み係数の変化率が、予め定められた閾値よりも大きい場合に、音声認識手段101に戻り処理するようにしてもよい。
- [0091] また、重み係数収束判断手段109は、重み係数の値だけではなく、目的関数計算手段107によって計算された積算された事後確率を用いて収束の判断をしてもよい。
- [0092] 例えば、重み係数収束判断手段109は、重み係数が更新される前のスコアと、更

新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、音声認識手段101に戻り処理するようにしてもよい。

[0093] また、上記の重み係数学習システムと音声認識システムは、そのハードウェア及びソフトウェア構成は特に限定されるものではなく、上述した各手段の機能を実現可能なものであれば、いずれのものでも適用可能である。例えば、各部の機能毎に回路を独立させて構成したものでも、複数の機能を1つの回路にまとめて一体に構成したものでも、いずれのものであってもよい。或いは、全ての機能を主にソフトウェアの処理で実現するものでもあってもよい。

[0094] また、上記の重み係数学習システムと、音声認識システムとを構成する各手段の少なくとも一部の機能を、プログラムコードを用いて実現する場合、かかるプログラムコード及びこれを記録する記録媒体は、本発明の範疇に含まれる。この場合、OS (Operating System) やアプリケーションソフト等の他のソフトウェアと連携して上記機能が実現される場合は、それらのプログラムコードも含まれる。記録媒体には、コンピュータ内に配置されるROM (Read Only Memory) 等のメモリやハードディスクのほか、CD-ROM (Compact Disk Read Only Memory)、DVD-ROM (Digital Versatile Disk Read Only Memory) 等のディスク型記録媒体や、磁気テープ等のテープ型記録媒体、或いは持ち運び可能なIC (Integrated Circuit) メモリ等の可搬型記録媒体も含まれる。また、他の実施形態としては、インターネット等の通信ネットワークを介して、サーバ等の他のコンピュータ機の記録媒体上に格納されたプログラムコードをダウンロードしたりする態様も含まれる。

[0095] この出願は、2007年2月28日に出願された日本出願特願2007-049975号を基礎とする優先権を主張し、その開示の全てをここに取り込む。

産業上の利用可能性

[0096] 本発明によれば、音声認識を行う音声認識装置や、音声認識装置をコンピュータに実現させるためのプログラムといった用途に適用できる。

請求の範囲

- [1] 学習音声データを認識しその認識結果を出力する音声認識手段と、
音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識手段により出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新手段と、
更新された後のスコアを用いて、前記重み係数更新手段に戻り前記重み係数を再更新するか否かを判断する収束判断手段と、
前記更新された後のスコアを用いて、前記音声認識手段まで戻りその処理をやり直して前記重み係数更新手段により前記重み係数を更新するか否かを判断する重み係数収束判断手段と、
を備えることを特徴とする重み係数学習システム。
- [2] 前記認識結果は、単語グラフであることを特徴とする、請求項1記載の重み係数学習システム。
- [3] 前記認識結果は、音響スコアと言語スコアとの和のうち、和の大きいものから予め定められた数の単語列であることを特徴とする、請求項1記載の重み係数学習システム。
- [4] 前記認識結果は、音響スコアと、言語スコアと、それらの和とが、予め定められた閾値より大きい単語列であることを特徴とする、請求項1記載の重み係数学習システム。
- [5] 前記認識結果は、音響スコアと、言語スコアと、それらの和とが、1番大きなスコアからの差分が、予め定められた閾値より小さい単語列であることを特徴とする、請求項1記載の重み係数学習システム。
- [6] 前記収束判断手段は、前記重み係数が更新される前のスコアと、前記重み係数が更新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、前記重み係数更新手段に戻り処理することを特徴とする、請求項1乃至5のいずれか1項に記載の重み係数学習システム。

- [7] 前記収束判断手段は、前記重み係数が更新される前と更新された後との前記重み係数の差分、および、当該差分を更新される前の値で割った前記重み係数の変化率が、予め定められた閾値よりも大きい場合に、前記重み係数更新手段に戻り処理することを特徴とする、請求項1乃至5のいずれか1項に記載の重み係数学習システム。
- [8] 前記重み係数収束判断手段は、前記重み係数が更新される前のスコアと、更新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、前記音声認識手段に戻り処理することを特徴とする、請求項1乃至7のいずれか1項に記載の重み係数学習システム。
- [9] 前記重み係数収束判断手段は、前記重み係数が更新される前と更新された後との前記重み係数の差分、および、当該差分を更新される前の値で割った前記重み係数の変化率が、予め定められた閾値よりも大きい場合に、前記音声認識手段に戻り処理することを特徴とする、請求項1乃至7のいずれか1項に記載の重み係数学習システム。
- [10] 学習音声データを認識しその認識結果を出力する音声認識ステップと、
音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識ステップにより出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新ステップと、
更新された後のスコアを用いて、前記重み係数更新ステップに戻り前記重み係数を再更新するか否かを判断する収束判断ステップと、
前記更新された後のスコアを用いて、前記音声認識ステップまで戻りその処理をやり直して前記重み係数更新ステップにより前記重み係数を更新するか否かを判断する重み係数収束判断ステップと、
を備えることを特徴とする重み係数学習方法。
- [11] 前記認識結果は、単語グラフであることを特徴とする、請求項10に記載の重み係数学習方法。
- [12] 前記認識結果は、音響スコアと言語スコアとの和のうち、和の大きいものから予め定

- められた数の単語列であることを特徴とする、請求項10記載の重み係数学習方法。
- [13] 前記認識結果は、音響スコアと、言語スコアと、それらの和とが、予め定められた閾値より大きい単語列であることを特徴とする、請求項10記載の重み係数学習方法。
- [14] 前記認識結果は、音響スコアと、言語スコアと、それらの和とが、1番大きなスコアからの差分が、予め定められた閾値より小さい単語列であることを特徴とする、請求項10記載の重み係数学習方法。
- [15] 前記収束判断ステップは、前記重み係数が更新される前のスコアと、前記重み係数が更新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、前記重み係数更新ステップに戻り処理することを特徴とする、請求項10乃至14のいずれか1項に記載の重み係数学習方法。
- [16] 前記収束判断ステップは、前記重み係数が更新される前と更新された後との前記重み係数の差分、および、当該差分を更新される前の値で割った前記重み係数の変化率が、予め定められた閾値よりも大きい場合に、前記重み係数更新ステップに戻り処理することを特徴とする、請求項10乃至14のいずれか1項に記載の重み係数学習方法。
- [17] 前記重み係数収束判断ステップは、前記重み係数が更新される前のスコアと、更新された後のスコアとの差分、および、当該差分を更新される前のスコアで割った変化率が、予め定められた閾値よりも大きい場合に、前記音声認識ステップに戻り処理することを特徴とする、請求項10乃至16のいずれか1項に記載の重み係数学習方法。
- [18] 前記重み係数収束判断ステップは、前記重み係数が更新される前と更新された後との前記重み係数の差分、および、当該差分を更新される前の値で割った前記重み係数の変化率が、予め定められた閾値よりも大きい場合に、前記音声認識ステップに戻り処理することを特徴とする、請求項10乃至16のいずれか1項に記載の重み係数学習方法。
- [19] コンピュータに、
学習音声データを認識しその認識結果を出力する音声認識処理と、

音声認識で用いる音響モデルと言語モデルとから得られるスコアに対し、前記学習音声データの正解テキストを用いて計算される正解のスコアと、前記音声認識処理により出力される認識結果のスコアとの差が大きくなるように前記スコアにかかる重み係数を更新する重み係数更新処理と、

更新された後のスコアを用いて、前記重み係数更新処理に戻り前記重み係数を再更新するか否かを判断する収束判断処理と、

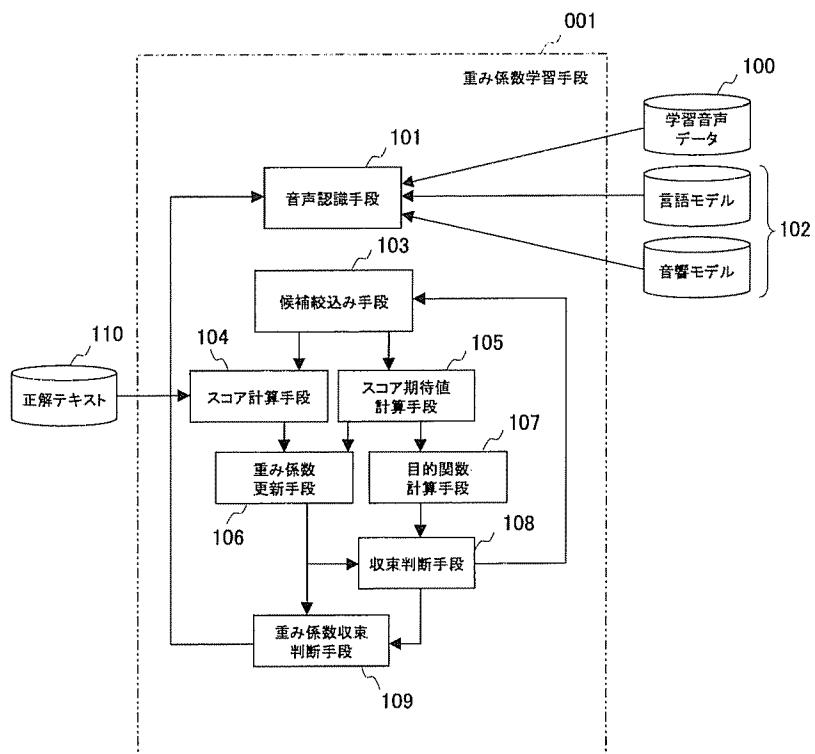
前記更新された後のスコアを用いて、前記音声認識処理まで戻りその処理をやり直して前記重み係数更新処理により前記重み係数を更新するか否かを判断する重み係数収束判断処理と、

を実行させるための重み係数学習プログラム。

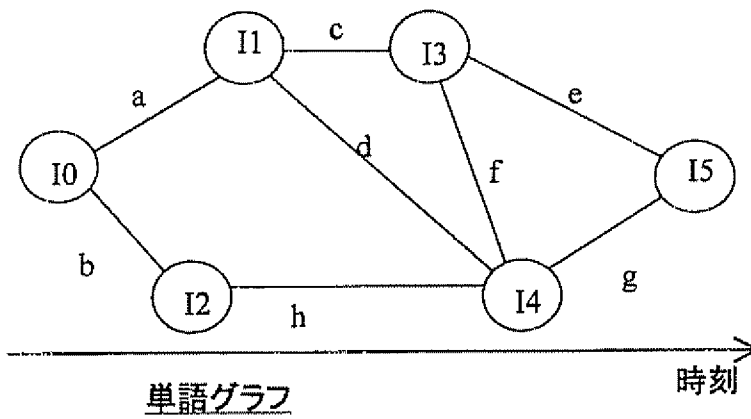
- [20] 請求項1乃至9のいずれか1項に記載の重み係数学習システムにより更新された重み係数を用いて音声認識を行うことを特徴とする音声認識システム。
- [21] 請求項10乃至18のいずれか1項に記載の重み係数学習方法により更新された重み係数を用いて音声認識を行うことを特徴とする音声認識方法。
- [22] コンピュータに、請求項19に記載の重み係数学習プログラムにより更新された重み係数を用いて音声認識を行う音声認識処理を実行させるための音声認識プログラム

。

[図1]



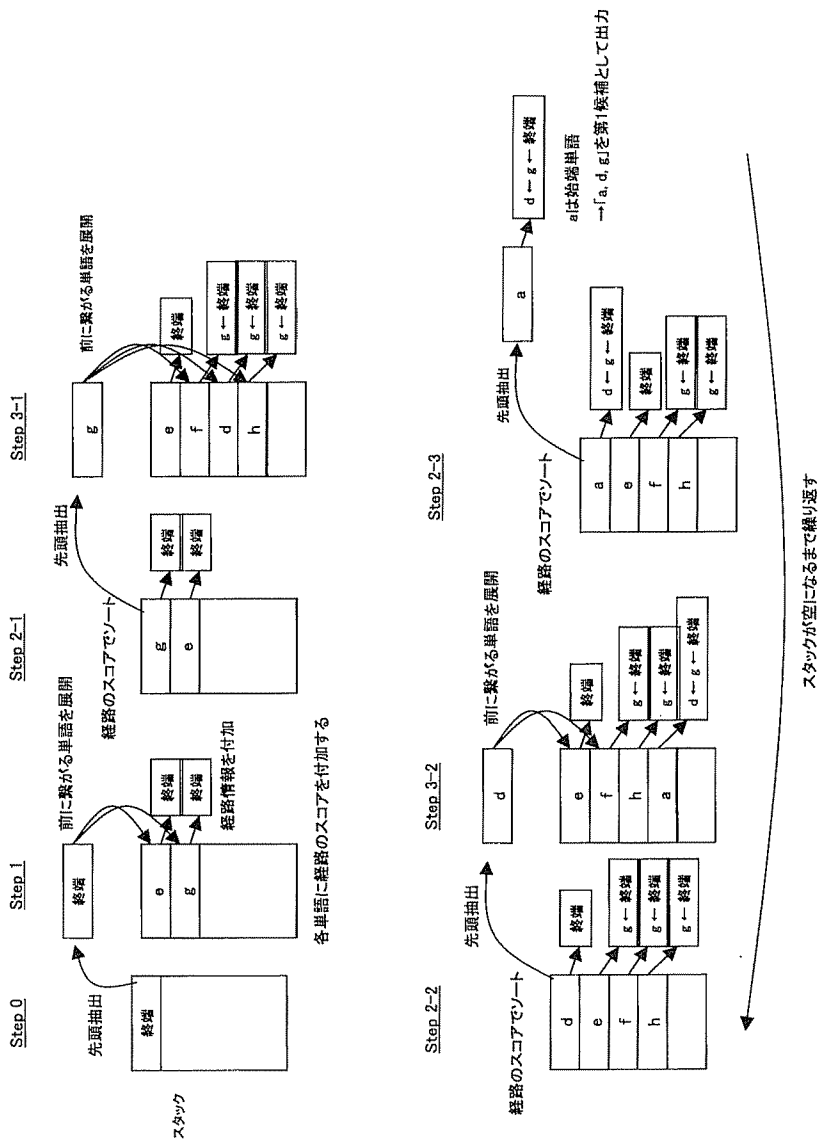
[図2]



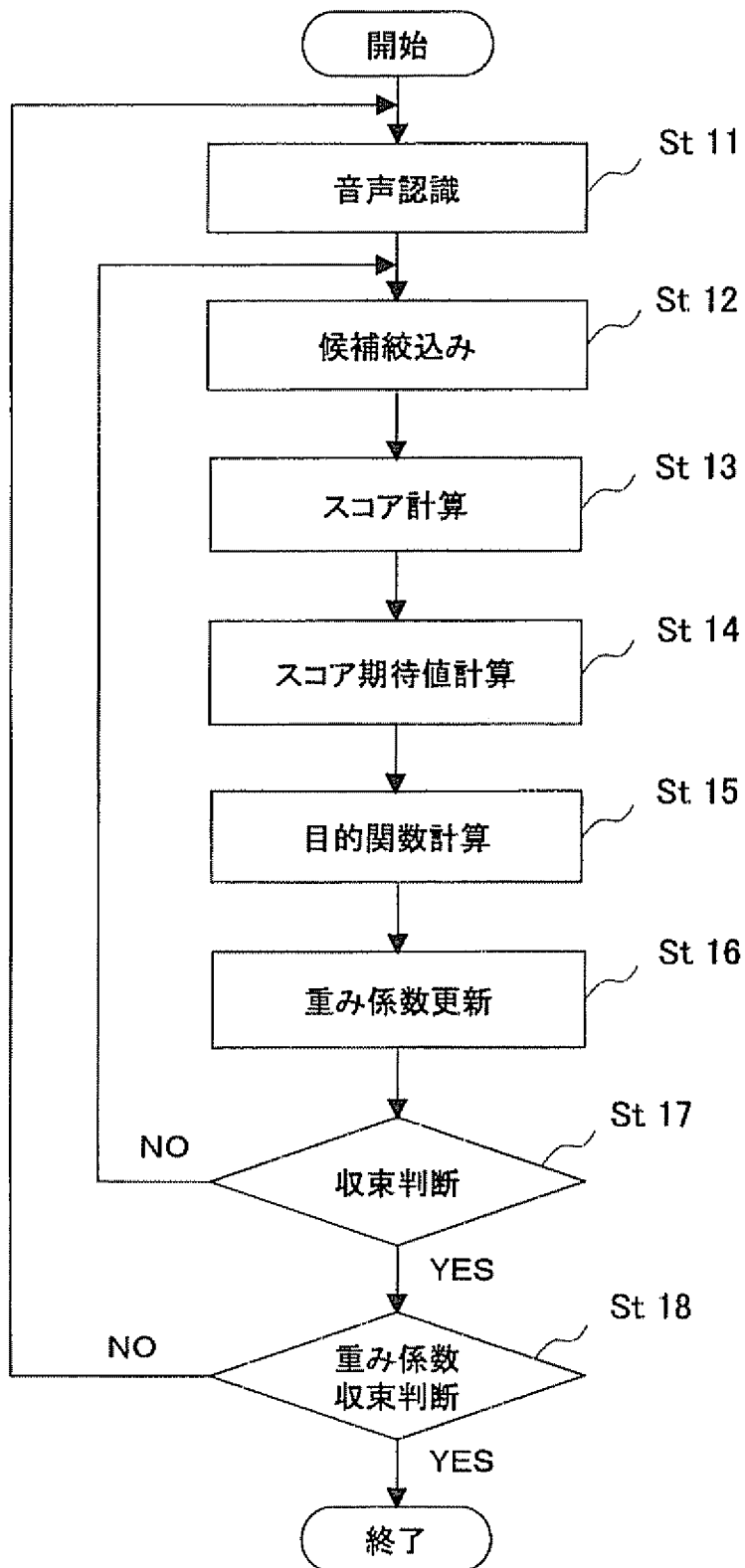
ノード	時刻
I0	= 1.0秒
I1	= 2.0秒
I2	= 2.5秒
I3	= 3.0秒
I4	= 3.5秒
I5	= 4.0秒

始端	終端	単語	音響尤度
I0	I1	a	-10.0
I0	I2	b	-12.7
I1	I3	c	-15.1
I1	I4	d	-20.4
I2	I4	h	-18.5
I3	I4	f	-13.6
I3	I5	e	-16.0
I4	I5	g	-10.2

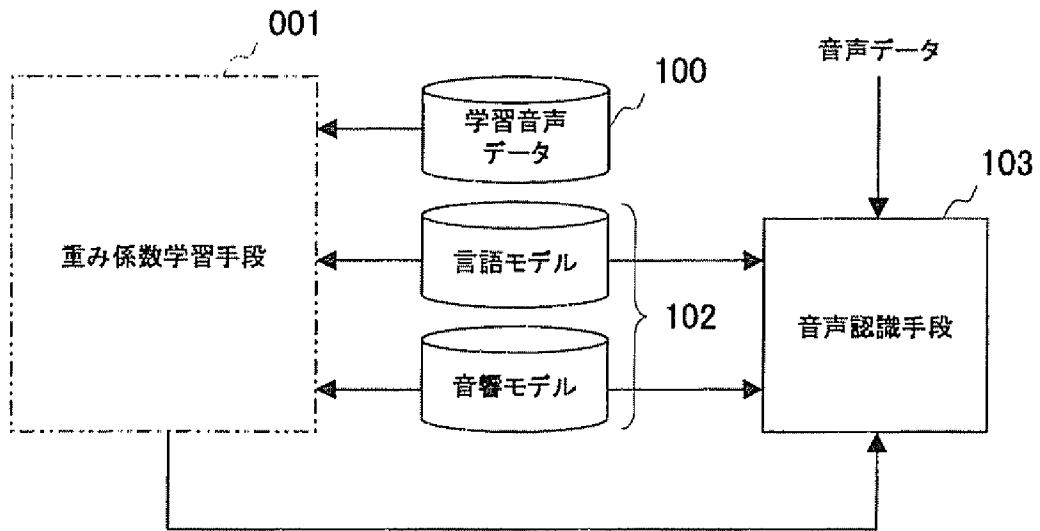
[図3]



[図4]



[図5]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2008/052721

A. CLASSIFICATION OF SUBJECT MATTER
G10L15/10(2006.01) i, G10L15/08(2006.01) i, G10L15/18(2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L15/00-15/28

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2008
Kokai Jitsuyo Shinan Koho	1971-2008	Toroku Jitsuyo Shinan Koho	1994-2008

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007-17548 A (Advanced Telecommunications Research Institute International), 25 January, 2007 (25.01.07), Full text; Figs. 1 to 8 (Family: none)	1-22
A	JP 2002-278578 A (Koninklijke Philips Electronics N.V.), 27 September, 2002 (27.09.02), Full text; Figs. 1 to 4 & EP 1231596 A1 & US 2002/0156628 A1	1-22
A	JP 9-81182 A (ATR Interpreting Telecommunications Research Laboratories), 28 March, 1997 (28.03.97), Full text; Figs. 1 to 10 (Family: none)	1-22

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 23 May, 2008 (23.05.08)	Date of mailing of the international search report 03 June, 2008 (03.06.08)
--	--

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/052721

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 11-85186 A (ATR Interpreting Telecommunications Research Laboratories), 30 March, 1999 (30.03.99), Full text; Figs. 1 to 6 (Family: none)	1-22
A	JP 7-152394 A (AT & T Corp.), 16 June, 1995 (16.06.95), Full text; Figs. 1 to 3 & EP 635820 A1 & US 5606644 A	1-22
A	JP 2000-352993 A (Oki Electric Industry Co., Ltd.), 19 December, 2000 (19.12.00), Full text; Figs. 1 to 4 (Family: none)	1-22
A	JP 2004-333738 A (NEC Corp.), 25 November, 2004 (25.11.04), Full text; Figs. 1 to 17 (Family: none)	1-22

A. 発明の属する分野の分類 (国際特許分類 (IPC))
 Int.Cl. G10L15/10(2006.01)i, G10L15/08(2006.01)i, G10L15/18(2006.01)i

B. 調査を行った分野
 調査を行った最小限資料 (国際特許分類 (IPC))
 Int.Cl. G10L15/00-15/28

最小限資料以外の資料で調査を行った分野に含まれるもの
 日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2008年
 日本国実用新案登録公報 1996-2008年
 日本国登録実用新案公報 1994-2008年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)
 WPI

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2007-17548 A (株式会社国際電気通信基礎技術研究所) 2007.01.25, 全文, 図1-8 (ファミリーなし)	1-22
A	JP 2002-278578 A (コーニンクレッカ フィリップス エレクトロ ニクス エヌ ヴィ) 2002.09.27, 全文, 図1-4 & EP 1231596 A1 & US 2002/0156628 A1	1-22
A	JP 9-81182 A (株式会社エイ・ティ・アール音声翻訳通信研究所) 1997.03.28, 全文, 図1-10 (ファミリーなし)	1-22

C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
 「A」特に関連のある文献ではなく、一般的な技術水準を示すもの
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」口頭による開示、使用、展示等に言及する文献
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献
 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」同一パテントファミリー文献

国際調査を完了した日 23.05.2008	国際調査報告の発送日 03.06.2008
--------------------------	--------------------------

国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 山下 剛史	5 Z	8 9 4 6
	電話番号 03-3581-1101 内線 3541		

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 11-85186 A (株式会社エイ・ティ・アール音声翻訳通信研究所) 1999.03.30, 全文, 図1-6 (ファミリーなし)	1-22
A	JP 7-152394 A (エイ・ティ・アンド・ティ・コーポレーション) 1995.06.16, 全文, 図1-3 & EP 635820 A1 & US 5606644 A	1-22
A	JP 2000-352993 A (沖電気工業株式会社) 2000.12.19, 全文, 図1-4 (ファミリーなし)	1-22
A	JP 2004-333738 A (日本電気株式会社) 2004.11.25, 全文, 図1-17 (ファミリーなし)	1-22