

【公報種別】特許法第 17 条の 2 の規定による補正の掲載

【部門区分】第 6 部門第 2 区分

【発行日】平成20年10月9日(2008.10.9)

【公表番号】特表2008-511024(P2008-511024A)

【公表日】平成20年4月10日(2008.4.10)

【年通号数】公開・登録公報2008-014

【出願番号】特願2007-528000(P2007-528000)

【国際特許分類】

G 1 0 L 15/10 (2006.01)

G 1 0 L 15/00 (2006.01)

G 1 0 L 15/18 (2006.01)

G 0 6 F 3/16 (2006.01)

【F I】

G 1 0 L 15/10 5 0 0 T

G 1 0 L 15/00 2 0 0 B

G 1 0 L 15/18 2 0 0 B

G 1 0 L 15/18 2 0 0 D

G 1 0 L 15/18 2 0 0 E

G 1 0 L 15/18 4 0 0

G 0 6 F 3/16 3 4 0 C

【手続補正書】

【提出日】平成20年8月15日(2008.8.15)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

(A) 文書の複数の部分構造に関連する複数の確率的言語モデルを含む確率的言語モデルを識別するステップと、

(B) 口語音声ストリームに前記確率的言語モデルを適用するために、音声認識デコーダを使用して、前記複数の部分構造に構成された内容を含む文書を生成するステップと、を含む方法であって、前記複数の部分構造のそれぞれにおける内容は、前記部分構造に関連する前記確率的言語モデルを使用して、音声を認識することによって生成される方法。

【請求項 2】

前記ステップ(B)は、

(B)(1)口語の音声ストリームの複数のセグメントSのそれぞれに対して、

(a) セグメントSに対する少なくとも2つの候補内容を識別するために、前記複数の確率的言語モデルのうちの少なくとも2つによってセグメントSを認識するステップと、

(b) セグメントSに対応する最終内容として、前記少なくとも2つの候補内容のうちの1つを選択するステップと、

(c) ステップ(B)(1)(b)で選択した前記候補内容を生成した前記確率的言語モデルに関連する前記文書の部分構造に、セグメントSに対する前記最終内容を挿入するステップと、

を行うステップを含む請求項1に記載の方法。

【請求項 3】

前記ステップ(B)(1)(b)は、

( i ) 前記少なくとも 2 つの候補内容に対するフィットネススコアを生成するために、前記少なくとも 2 つの候補内容にメトリックを適用するステップであって、前記フィットネススコアは、前記候補内容が前記口語音声ストリームを表す可能性を示すステップと、  
( i i ) 最も高いフィットネススコアを有する候補内容を選択するステップと、  
を含む請求項 2 に記載の方法。

【請求項 4】

前記ステップ ( B ) ( 1 ) ( b ) は、

( i i i ) ステップ ( B ) ( 1 ) ( b ) ( i i ) で選択された前記候補内容を生成するために使用された前記確率的言語モデルを選択するステップをさらに含む、請求項 3 に記載の方法。

【請求項 5】

前記文書の前記複数の部分構造は複数の前記文書の一部を含む請求項 1 に記載の方法。

【請求項 6】

前記複数の部分構造は意味概念を表す部分構造を含む請求項 1 に記載の方法。

【請求項 7】

前記意味概念はコンピュータ可読形式で前記文書内に表される請求項 6 に記載の方法。

【請求項 8】

前記複数の確率的言語モデルは、階層に構成され、前記ステップ ( B ) は、

( B ) ( 1 ) 前記階層を介して経路を識別するステップと、

( B ) ( 2 ) ステップ ( B ) ( 1 ) で識別された前記経路に対応する構造を有する文書を生成するステップと、

を含む請求項 1 に記載の方法。

【請求項 9】

前記ステップ ( B ) ( 1 ) は、前記口語音声ストリームを認識するために前記音声認識デコードによって適用された場合に、前記複数の確率的言語モデルの前記階層に関する最適認識結果を生成する前記階層を介して経路を識別するステップを含む、請求項 8 に記載の方法。

【請求項 10】

ステップ ( B ) ( 1 ) は、

( B ) ( 1 ) ( a ) 前記階層を介して複数の経路を識別するステップと、

( B ) ( 1 ) ( b ) 前記複数の経路 P のそれぞれに対して、経路 P 上の前記言語モデルを使用して前記口語音声ストリームを認識するために、前記音声認識デコードを使用して前記口語音声ストリームに対する構造化文書の候補を生成するステップと、

( B ) ( 1 ) ( c ) 前記複数の構造化文書の候補に対する複数のフィットネススコアを生成するために、ステップ ( B ) ( 1 ) ( b ) で生成された前記複数の構造化文書の候補にメトリックを適用するステップと、

( B ) ( 1 ) ( d ) 最も高いフィットネススコアを有する前記構造化文書の候補を生成する前記経路を選択するステップと、

を含む請求項 8 に記載の方法。

【請求項 11】

前記音声認識デコードは、複数の音声認識デコードを含み、

前記ステップ ( B ) は、

( B ) ( 1 ) 口語音声ストリームのセグメントを識別するステップと、

( B ) ( 2 ) 前記複数の確率的言語モデルのうちの 1 つを識別するステップと、

( B ) ( 3 ) 前記複数の確率的言語モデルのうちの識別された 1 つに関連する、前記複数の音声認識デコードのうちの 1 つを識別するステップと、

( B ) ( 4 ) 前記識別されたセグメントに、前記識別された確率的言語モデルを適用するために、前記識別された音声認識デコードを使用して内容を生成するステップと、

を含む請求項 1 に記載の方法。

【請求項 12】

ステップ ( B ) は、

( B ) ( 1 ) 前記音声ストリーム内で、前記複数の確率的言語モデルと複数のセグメントとの間のマッピングを識別するステップと、

( B ) ( 2 ) 前記複数のセグメントのそれぞれに対して、

( B ) ( 2 ) ( a ) 前記マッピングを使用して、前記複数の確率的言語モデルのうちの対応する 1 つを識別するステップと、

( B ) ( 2 ) ( b ) 前記識別された確率的言語モデルに関連する前記複数の部分構造のうちの 1 つを識別するステップと、

( B ) ( 3 ) ( b ) 前記確率的言語モデルのうちの識別された 1 つを使用して前記セグメントを認識するために、前記音声認識デコーダを使用して前記識別された部分構造内に内容を生成するステップと、

を行うステップと、

を含む請求項 1 に記載の方法。

【請求項 1 3】

文書の複数の部分構造に関連する複数の確率的言語モデルを含む確率的言語モデルを識別するための第 1 の識別手段と、

口語音声ストリームに前記確率的言語モデルを適用するために音声認識デコーダを使用して、前記複数の部分構造に構成された内容を含む文書を生成するための文書生成手段と、  
を含む装置であって、前記複数の部分構造のそれぞれにおける内容は、前記部分構造に関連する前記確率的言語モデルを使用して、音声を認識することによって生成される装置。

【請求項 1 4】

前記文書生成手段は、

口語の音声ストリームの複数のセグメント S のそれぞれのための繰り返し手段を備え、該繰り返し手段は、

セグメント S に対する少なくとも 2 つの候補内容を識別するために、前記複数の確率的言語モデルのうちの少なくとも 2 つによってセグメント S を認識するための認識手段と、  
セグメント S に対応する最終内容として、前記少なくとも 2 つの候補内容のうちの 1 つを選択するための第 1 の選択手段と、

前記選択手段で選択した前記候補内容を生成した前記確率的言語モデルに関連する前記文書の部分構造に、セグメント S に対する前記最終内容を挿入するための挿入手段と、  
を備える請求項 1 3 に記載の装置。

【請求項 1 5】

前記文書の前記複数の部分構造は複数の前記文書の部分を含む請求項 1 3 に記載の装置。

【請求項 1 6】

前記複数の確率的言語モデルは、階層に構成され、前記文書生成手段は、

前記階層を介して経路を識別するための第 2 の識別手段と、

前記第 2 の識別手段で識別された前記経路に対応する構造を有する文書を生成するための手段と、

を備える請求項 1 3 に記載の装置。

【請求項 1 7】

請求項 1 のステップ ( A ) で識別された言語モデルを含むデータ構造。

【請求項 1 8】

複数の言語モデルが階層に論理的に構成され、前記複数の言語モデルが第 1 の言語モデルおよび第 2 の言語モデルを含むデータ構造であって、

前記第 1 の言語モデルは、前記階層において第 2 の言語モデルの親であり、

前記第 1 の言語モデルは、文書の部分構造に関連する第 1 の概念を表す音声の認識に好適であり、

前記第 2 の言語モデルは、前記文書の部分構造のサブセットに関連する第 2 の概念を表す音声の認識に好適である、

データ構造。

**【請求項 19】**

(A) 第1の階層内に論理的に構成された複数の概念に関連する複数の確率的言語モデルを含む確率的言語モデルを識別するステップと、

(B) 口語音声ストリームに前記確率的言語モデルを適用するために音声認識デコーダを使用して、前記第1の階層を介した経路によって定義される論理構造を有する第2の階層に論理的に構成された複数の部分構造に構成された内容を含む文書を生成するステップと

を含む方法。

**【請求項 20】**

前記ステップ(B)は、前記文書を生成するために、前記第1の階層を介して前記経路をトラバースするステップを含む、請求項19に記載の方法。

**【請求項 21】**

前記ステップ(B)は、

(B)(1) 口語の音声ストリームの複数のセグメントSのそれぞれに対して、

(a) セグメントSに対する少なくとも2つの候補内容を識別するために、前記複数の確率的言語モデルのうちの少なくとも2つによってセグメントSを認識するステップと、

(b) セグメントSに対応する最終内容として、前記少なくとも2つの候補内容のうちの1つを選択するステップと、

(c) ステップ(B)(1)(b)で選択した前記候補内容を生成した前記確率的言語モデルに関連する前記文書の部分構造に、セグメントSに対する前記最終内容を挿入するステップと、

を行うステップを含む請求項19に記載の方法。

**【請求項 22】**

前記複数の部分構造は意味概念を表す部分構造を含む請求項19に記載の方法。

**【請求項 23】**

前記ステップ(B)は、

(B)(1) 前記階層を介して経路を識別するステップと、

(B)(2) ステップ(B)(1)で識別された前記経路に対応する構造を有する文書を生成するステップと、を含む請求項19に記載の方法。

**【請求項 24】**

第1の階層内に論理的に構成される複数の概念に関連する複数の確率的言語モデルを含む確率的言語モデルを識別するための識別手段と、

口語音声ストリームに前記確率的言語モデルを適用するために音声認識デコーダを使用して、前記第1の階層を介した経路によって定義される論理構造を有する第2の階層に論理的に構成された複数の部分構造に構成された内容を含む文書を生成するための文書生成手段と、

を備える装置。

**【請求項 25】**

前記文書生成手段は、前記文書を生成するために、前記第1の階層を介して前記経路をトラバースする手段を備える請求項24に記載の装置。