

特開2008-129318

言語モデル作成装置及び言語モデル作成プログラム並びに音声認識装置及び音声認識プログラム

特徴と利用分野

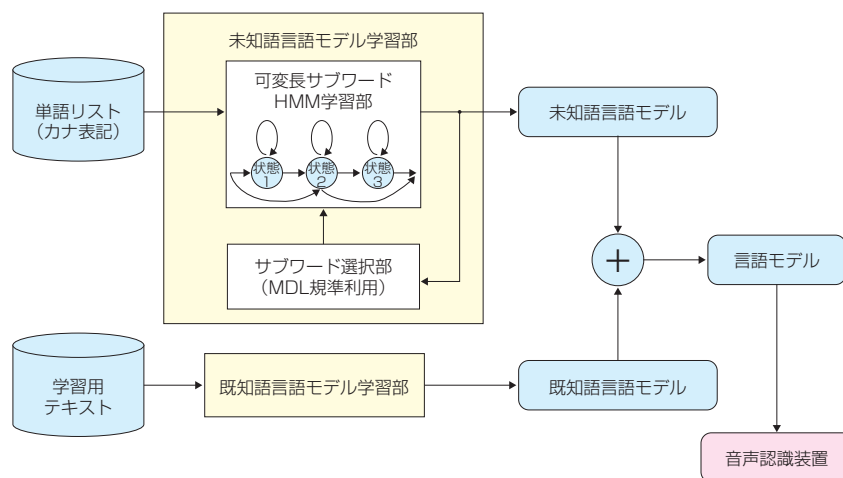
本発明は、音声認識装置に登録されていない単語（未知語）をカタカナの文字列で出力するための技術である。これまでの音声認識装置では、未知語を含む音声区間で認識誤りを必然的に生じていたが、本発明を適用することで、固有名詞などの未知語をカナ文字列として出力できるようになるので、認識誤りの削減効果が期待できる。

技術概要

本発明の構成を1図に示す。本装置の言語モデル学習部は未知語言語モデル学習部と既知語言語モデル学習部で構成され、未知語言語モデルと既知語言語モデルを統合した認識率の高い言語モデルを音声認識装置で用いる。言語モデルの語彙に「ア」「キャ」のような日本語の「子音+母音（音節）」の単位を登録しても認識精度は少ししか改善できないので、音節よりも長く、単語よりも短い「単語の一部（サブワード）」を語彙に登録する。例えば、単語「アイウ」を未知語とした場合、[ア][イ][ウ]、[ア][イウ]、[アイ][ウ]、[アイウ]の4通りのサブワード系列が考えられるが、未知語言語モデル学習部では、このようなすべてのパターンのサブワード系列の出現確率を、カナで表記した大量の単語リストから、可変長の音節列を出力可能な隠れマルコフモデル（可変長サブワードHMM）で学習する。ただし、学習に用いる単語に含まれるすべてのサブワードの出力確率を求めると、サブワードのパターン数が膨大となるので、MDL（Minimum Description Length：最小記述長）規準*を用いて、サブワードを削減する。このように学習した未知語言語モデルと、大量の学習用テキストで学習した既知語言語モデルを統合することによって、認識率の高い言語モデルが得られる。

（発明者：本間真一，今井亨）

* 複数のモデルの中から、最適なパラメーターの次数を持つモデルを選択するための規準。データを記述するのに必要な符号長が最小となるモデルは予測能力が高いという考えに基づく。



1図 言語モデル作成装置の構成