

Title	音声認識における特徴量の非同期性と音素環境依存性のモデル化に関する研究
Author(s)	松田, 繁樹
Citation	
Issue Date	2003-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/937
Rights	
Description	Supervisor: 下平 博, 情報科学研究科, 博士

音声認識における特徴量の非同期性と音素環境依存性の モデル化に関する研究

松田 繁樹

北陸先端科学技術大学院大学 情報科学研究科

2003年2月10日

あらまし

本研究では、音響特徴ベクトル時系列における個々の音響特徴量の振舞いに着目した音声認識性能の改善に関する検討を行う。音声認識システムの音響モデルとして広く用いられている隠れマルコフモデル(HMM)は、音声の観測量がベクトルであることを仮定している。本研究は、このような常識を覆し、「個別特徴量の集合」として捉えることにより、従来になかった次に述べる2つの仮説について検討を行った。

第1の仮説として、「個々の特徴量の時間非同期性」に着目した。音響特徴ベクトル時系列の個々の音響特徴量の値は必ずしも同じタイミングで変化していない。従来型HMMは、音響特徴ベクトルを構成しているすべての特徴量の値がHMMの状態遷移と同じタイミングで変化することを仮定したモデルである。従って、このような信号を従来型HMMでモデル化した場合、大量の時間方向状態数が必要となる。しかし、大量の時間方向状態数は、モデルパラメータの多大な増加を招くため、モデルの統計的信頼性の低下に繋がる。そこで本検討では、個々の特徴量の値がお互いに異なるタイミングで変化するモデルの検討を行う。本検討の結果、特定話者音声認識において、本手法の有効性を確認した。

第2の仮説として、「個々の特徴量の音素環境依存性」に着目した。従来提案されたパラメータ共有法の幾つかは、音素環境依存性を考慮したクラスタリングを行なっている。この従来のパラメータ共有法は、全体的特徴量に対して共通のパラメータ共有構造を割り当てる手法である。しかし、音声の観測量である音響特徴ベクトルは、お互いに異なった振舞いを持つ音響特徴量の集合であり、お互いに異なった複雑性や音素環境依存性を持つと考えられる。そこで本検討では、個々の特徴量に依存したパラメータ共有構造を、音素環境依存性を考慮して決定することによる、音声認識性能の改善を検証した。本検討の結果、比較的少ないパラメータ数を持つモデルで、その有効性を確認した。

以上のように、従来の音声認識においては「特徴ベクトル」という概念のもとに特徴量が一括して扱われて来たのに対し、本研究では、個別特徴量の扱いを提唱し、時間特性と環境依存性の2つの側面から検証し、モデルや学習及び、認識アルゴリズムを提案し、実験を通して有効性を実証した。

キーワード: 音声認識, 音響モデル, 非同期遷移型隠れマルコフモデル, 時間方向共有法,
特徴量依存音素環境クラスタリング, 特徴量依存逐次状態分割法