

専攻	システム情報	学籍番号	907453	指導教官氏名	中川 聖一
申請者氏名	周 旻				湯浅 太一
					井上 克巳

論 文 要 旨

論文題目	A Study on Stochastic Models for Spoken Language (音声言語の確率モデルに関する研究)
------	---

(要旨 和文 1,200字程度)

(1)

コンピュータの発達と計算機の処理能力の向上に伴い、自然言語の音声認識が実現できるようになってきた。特に大量のデータベースが利用でき、色々な知識情報を組み込んで、音声認識の研究が初期よりかなり発展してきた。しかしながら、現在のところ不特定話者の制限のない自然発話の音声認識はまだ難しい。そのため、話者・語彙数・発声内容・発声方法・環境などに対する制限の下で研究を進めている場合が多い。それらの制限をできるだけ排除し、自然発話の音声をより正確に認識することは現在の音声認識分野の重要な課題となっている。

連続音声認識システムの構築では音響モデルと言語モデルが二つの重要な柱になっている。特に大語彙のためのシステムの構築には、(1)音韻、音節、単語を精度よく識別するための音響モデルの利用、(2)音声認識システムの探索空間を縮小させ、音響的な特徴のみによる認識結果を補正するための言語モデルの利用、が重要である。

本研究では大語彙連続音声認識システムの実現を目指し、確率モデルを用いた音響モデルと言語モデルの研究を行ってきた。主として以下のテーマについて検討を行ってきた。

音響モデルのレベル: (1) Inside-Outside アルゴリズムを用いた確率文脈自由文法 (SCFG) の学習方法を音声認識に適用した。SCFG のパラメータの推定では、スケールリング方法を導入して計算時におけるアンダーフローの問題を解決した。(2) 確率文脈自由文法を単語モデルとして孤立音声認識に適用し、現在の主流技術となっている隠れマルコフモデル(HMM)と比較した。確率文脈自由文法を音声認識に応用する場合の有効性や問題点と、

HMM よりも優位な結果になる条件を明らかにした。(3)誤り訂正法をSC FG に適用し、最尤推定より高い認識能力があることを示した。また、この方法を日本語の孤立音声認識のための音響モデルの学習に適用し、HM Mと比較した。

5 言語のモデル化のレベル: (1) 確率文脈自由文法、隠れマルコフモデル、確率文節文法、n重マルコフモデル及び新しく提案した隠れマルコフモデルと単純マルコフモデルの結合法を用いて、英語及び日本語のテキストデータに基づいて言語モデルを構築し、エントロピーによる相互比較で優劣関係を調べた。(2)任意の文の先頭からn単語まで既知として、後続品詞や後続単語の予測可能性の基準で評価し、それらの確率モデルによる自然言語のモデル化の有効性を調べた。(3) CKY構文解析法に基づく SCFG の学習法を Inside-Outside アルゴリズムと比較し、各々の特徴及び適用範囲を明らかにした。また、文法のサイズと有効な学習データ量との関係を調べた。

連続音声認識への適用: 従来の連続音声認識手法 — One-Pass DPアルゴリズムを拡張し、確率言語モデルの利用に基づく大語彙の連続音声認識システムの構築を行なった。語彙数が5000単語の日本語観光案内タスクを用いて、連続音声認識の評価実験を行ない、実際の音声認識実験を行なう時に出てくる問題点を検討した。