

氏 名	かわ 河 原 達 也
学位(専攻分野)	博 士 (工 学)
学位記番号	論 工 博 第 2983 号
学位授与の日付	平 成 7 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 2 項 該 当
学位論文題目	Studies on Speech Recognition based on Discriminative Statistical Models and Heuristic Search Strategies (識別的統計モデルと発見的探索戦略に基づく音声認識に関する研究)
論文調査委員	(主 査) 教 授 堂 下 修 司 教 授 池 田 克 夫 教 授 長 尾 真

論 文 内 容 の 要 旨

会話音声の自動認識は、より自然な対話によるヒューマン・マシンインタフェイスの実現のために重要である。音声理解・対話システムにおいては、不特定の話者に対応できること、自然な連続発声、特に非定型な発話が扱えることが必要である。また、音声認識は、パターン認識・言語解析・知識表現・探索などの問題を包含する総合的なシステムとして扱われなければならない。本論文は、より高精度で頑健な会話音声認識を実現するために、音声パターンの認識における識別的な統計モデルと、この音響モデルと言語モデルを統合したヒューリスティックな探索戦略に関する研究をまとめたもので、6章からなっている。

第1章は、本研究の位置付けとアプローチについて説明している。

第2章では、音声情報処理研究の最終目標である音声対話システムの概要を示し、その中での本会話音声認識の位置付けを示し、信号処理・音響モデル・言語モデルなどからなる処理の概要と、それらを制御する探索アルゴリズムについて、基本的な考え方と手法を述べている。また、本研究で用いたタスクとデータベースについても説明している。

第3章では、音声認識の基盤となる音響モデルとして、識別的統計モデルである対判別にに基づく隠れマルコフモデル(HMM)を提案している。音素間の分離度を向上させるために、音素の状態/シンボルの対毎に、判別分析を行ない、ベイズ識別器を構築する。個々の識別器では、2つのクラスのランクづけ、及び確率の比の計算を行う。連続型HMMでは、それらの結果を組み合わせ、正規化することにより、出力確率を求める。離散型HMMでは、対判別結果を組み合わせ、シンボルを決定する。実験の結果、不特定話者の音素認識率で86.3%、653単語の認識率で85.0%を得た。これらは、従来のHMMより10%以上高く、本方式の有効性が確認された。

第4章では、第3章のHMMを用いた連続音声のパーズングにおけるヒューリスティックな探索戦略に

ついて論じている。%LR (文脈自由) パージングにおいて、未探索部分のヒューリスティックスを評価値に加えながら、best-first に仮説を展開する A* 探索を実現した。ヒューリスティックスとしては、計算量も小さくなく、適格性を満たし、制約としてもかなり強い単語対制約を提案している。実験の結果、従来広く用いられているビームサーチより、少ない計算量で高い文認識率 (63.0%) を得た。さらに、この枠組において、種々の言語的知識の統合を行なった。対話の現在の状態から次発話の構文を予測する機構を実現し、増進的な解析・検証を行なう意味解析器と相互作用を行なうことにより、文の意味理解率を 5~10% 向上させ、最高で 75.5% を得た。

第 5 章では、ヒューリスティックな単語・文節のスポッティングの方法について述べている。単語部分のマッチングスコアに、発話の残りの部分の言語らしさを加えた評価関数を提案した。ヒューリスティックな言語モデルとして、音節連鎖・単語連鎖などを用いて、219 単語のスポッティング実験で、従来の始端フリーマッチングより高い抽出率を得た。次に、文節を単位とするスポッティングを提案した。文節は、構文的・意味的なまとまりを形成し、非定型的な発話の理解に適している。ヒューリスティックな言語的制約を文節構文の部分集合として、文節スポッティングを A* 探索として実現し、抽出率を向上させた。さらに、文節スポッティングに基づく文認識の探索戦略に関する考察を行ない、非文法的な発話に対して頑健な理解を可能とした。

以上の各要素技術を統合して、会話音声認識システムを実現した。これを対話管理部と結合することにより、音声対話システムを構築し、本研究の有効性が確認された。

第 6 章では、結論と今後の課題を述べている。

論文審査の結果の要旨

本論文は、自然な連続発声による非定型な発話に対して、より高精度で頑健な会話音声認識を実現するために、音声パターンの認識における識別的統計モデルと、この音響モデルと言語モデルを統合するためのヒューリスティックな探索戦略に関する研究をまとめたもので、得られた主な成果は以下の通りである。

1. 音響モデルとして、識別的統計モデルである対判別にに基づく隠れマルコフモデル (HMM) を提案し、連続型 HMM・離散型 HMM で実現し、良好な音素認識率と単語認識率を得た。
2. HMM を用いた連続発声のパージングにおいて、未探索部分をヒューリスティックに評価する A* 探索を実現し、単語対制約を用いることにより、少ない計算量で高い文認識率を得た。
3. 上記の方式において、対話の次発話の構文を予測する機構、増進的に解析・検証を行なう意味解析との相互作用を導入し、文の意味理解率を向上させた。
4. 単語のスポッティングにおいて、対象単語のスコア以外に周辺部分における音節や単語連鎖を評価関数として用いる方式を提案し、従来より高い抽出率を得た。また、文節単位のスポッティング方式を提案し、それを A* 探索として実現し、非文法的発話に対して高い理解率を得た。
5. 以上の各要素技術を統合して会話音声認識システムを実現し、さらに対話管理部と結合して音声対話システムを構築し、本方式の有効性を確認した。

以上のように、本論文は、音声対話システムのための頑健な音声認識・理解方式に新たな知見を加えたものであり、学術上・実際上寄与するところが少なくない。よって、本論文は博士（工学）の学位論文として価値あるものと認める。

また、平成7年2月6日、論文内容とそれに関連した事項について試問を行なった結果、合格と認めた。