

氏 名	趙 チョウ リョク リョク 力
学位(専攻分野)	博 士 (工 学)
学 位 記 番 号	博 甲 第 170 号
学位授与の日付	平成 10 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当
研 究 科・専 攻	工学科学研究科 情報・生産科学専攻
学 位 論 文 題 目	中国語連続音声の認識に関する研究
審 査 委 員	(主査) 教 授 新 美 康 永 教 授 田 村 博 教 授 黒 川 隆 夫

論 文 内 容 の 要 旨

隠れマルコフモデルに代表される統計的手法の導入により、連続音声の認識技術は近年進歩している。しかし、中国語音声の認識に関しては、欧米語や日本語の場合に比較して、未開拓の分野が多かった。本論文では、中国語の連続音声認識で十分に検討されていなかった次の 3 点に関する研究結果を報告している。

- (1) 音声処理と言語処理の密接な統合
- (2) 連続音声の中の四声の認識
- (3) 分節的情報と韻律的情報の統合

本論文は以下に要約する 7 章から構成されている。第 1 ないし第 3 章では、中国語の連続音声認識の背景、連続音声の認識に関する基礎理論、中国語の音声学上および言語学上の特性についてまとめ、次章以降の準備としている。

第 4 章では、上記 (1) の問題、すなわち、音声処理と言語処理の密接な統合について扱っている。認識対象をホテルの予約に関する対話文に限定し、その構文を文脈自由文法を用いて記述し、文脈自由文法の下降型構文解析法の 1 つである Earley 法と、連続音声認識の基本的な技法である One Pass Viterbi 法を組み合わせる計算効率の良いフレーム同期型のアルゴリズムを提案し、音声処理と言語処理の密接な統合を達成した。この方法により 10 名の話者の発声した文音声に対して、平均 63.5 % の文認識率を得たことを報告している。

第 5 章では、上記 (2) の連続音声の中の四声の認識に関する手法と実験結果について報告している。中国語の音調の認識に関しては、従来離散分布型の隠れマルコフモデルが用いられていたが、本研究では連続分布型の隠れマルコフモデルを用いている。また、音調を表現するパラメータとして、音声の基本周波数の 1 次及び 2 次の時間微係数を採用し、さらに音声の振幅と零交差情報を用いて、入力音声をあらかじめセグメンテーションしておくといった新しい手法を導入することにより、従来にない高精度で連続音声の中の四声の認識を行うことが可能となることを実験により示した。

第 6 章では、上記 (3) の問題、すなわち、分節的情報と韻律的情報の統合に関する問題を扱っている。従来の研究では、この両者は個別に用いられ、音素認識と音調認識は独立して行われ、それぞれの結果を統合するのが一般的であった。この研究では、第 4 章で研究した連続音声認識の基本的な手法である One Pass Viterbi 法と下降型構文解析法を組み合わせたフレーム同期型のアルゴリズムの中に両者を組み込むことを検討した。すなわち、音素をモデル化した隠れマルコフモデルの状態、音調をモデル化した隠れマルコフモデルの状態、及び時間で構成される 3 次元空間内で、最適なパスを求める方法を用いた。この 3

次元空間内でのパスに制限を課すことにより、分節的情報と韻律的情報の同期を厳密に単語の終端や音節の終端で取ることが可能になった。この方法を用いて第4章で述べたのと同じ対象に対して、平均文認識率75.9%を達成し、両情報を結合しない場合に比して12.4%向上することを示した。

第7章では以上の結果を要約し、今後の研究の展開の可能性を展望している。

論文審査の結果の要旨

本論文は、中国語の連続音声認識に関する基本的なテーマを扱った論文である。特に言語のもつ文法的な制約を連続音声の認識過程の中に組み込むためのアルゴリズム、連続音声の中の四声を精度良く認識するための新しい方法、分節的情報と韻律的情報の厳密な同期を取りつつ両者を統合するための枠組みなど、今後の中国語音声の認識の研究で十分活用される事が期待でき、その意義は大きい。これらの研究は次の4編の論文として公表されている。

1. 趙力、小林豊、新美康永、"連続分布 HMM を用いた中国語連続音声の音調認識、"日本音響学会誌、53 巻、12 号、pp. 1-8 (1997. 12)
2. 趙力、小林豊、新美康永、"3 次元 Viterbi 法を用いた音素情報と音調情報の統合による中国語連続音声認識、"日本音響学会誌、54 巻、7 号、pp. 497-505 (1997. 7)
3. L. Zhao, Y. Kobayashi and Y. Niimi, "Recognition of Chinese Continuous Speech Based on the Integration of Phonetic and Prosodic Information," Proc. of 7th Int. Conf. on Computer Processing of Oriental Languages, pp. 192-197 (1997. 4).
4. L. Zhao, Y. Kobayashi and Y. Niimi, "HMM Based Recognition of Chinese Tones in Continuous Speech," Proc. of 1997 China-Japan Symposium on Advanced Information Technology, pp.47-51 (1997. 4).