

審査の結果の要旨

氏名 加藤 恒夫

本論文は「携帯電話向け HMM 音声認識の高精度化と高速化に関する研究」と題し、携帯電話による音声認識を対象として、HMM 音素モデルの高度化、ビーム探索の高速化などを実現したもので、全 8 章からなる。

第 1 章は「序論」であって、携帯電話端末を用いた音声認識の歴史を概観した上で、分散型音声認識システム、ローカル音声認識エンジンの高精度化と高速化を達成する際に必要な項目を整理し、本論文の意義、目的と提案手法の概要を述べている。また章の構成が示されている。

第 2 章は「HMM に基づく音声認識」と題し、HMM 音素モデルと N-gram 言語モデルに基づく音声認識手法を概説した上で、音素モデルの高度化と探索の高速化に向けた従来の試みを整理し、本論文の提案手法の導入としている。

第 3 章は「混合分布 HMM における決定木に基づく状態クラスタリング」と題し、音素 HMM の学習コーパスサイズの制約に対処する手法として一般的な決定木に基づく状態クラスタリングについて述べ、本章での提案手法が、リーフノードの出力確率密度分布の単一正規分布表現から混合正規分布表現への拡張であるとしている。連結学習を含めた評価として、音節タイプライタあるいは単語連続の認識実験を行い、従来の単一正規分布表現による手法と比較して、提案手法が認識精度、演算時間の双方で優れているとしている。

第 4 章は「コーデック適応音響モデルおよび雑音モデルによる認識精度の改善」と題し、携帯電話のコーデックによる音声歪みと背景雑音に対処する手法として、コーデック毎に音声モデルと非定常雑音モデルを用意し、最尤基準に従ってモデル選択を行うものを開発している。孤立単語音声認識実験を行い、単語認識率の大幅な向上を達成している。

第 5 章は「木構造辞書における到達可能単語数を利用した探索高速化」と題し、木構造辞書での到達可能単語数がノードからリーフに向けて単調減少することに着目した枝刈り（ビーム探索）手法を提案している。到達可能単語数に従って、仮説の尤度に一定値を加算するもので、孤立単語音声認識、短文認識、メール読み取りの各場合について認識実験を行い、演算時間の大幅な短縮を達成できることを示している。

第 6 章は「音声区間検出の情報と木構造辞書における深さを利用した探索高速化」と題し、前章での提案に加え、音声区間の始端（の短時間前）部分の仮説の広がり、“非音声らしさ”によって絞り込む手法を提案している。孤立単語音声認識で 1 割弱の演算時間の短縮を達成している。

第 7 章は「実用システムの構築」と題し、3～6 章での提案を実装した、乗換検索などの種々のアプリケーションについての概要と性能評価実験について、まとめている。

第 8 章は、「結論」であって、各章の概要を述べた上で、今後の課題・展望を整理している。

以上を要するに、本論文は、携帯電話を用いた音声認識を対象として、混合分布 HMM の状態クラスタリング手法、コーデックに対応した音響・雑音モデルの選択手法、木構造辞書における探索高速化手法などを開発して認識の高精度化と高速化を達成するとともに、音声認識システムとして実用化したものであって、音声認識の発展に大きく寄与したものであり、電子情報学に貢献するところが少なくない。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。