

論文内容の要旨

博士論文題目 Selective Training for Cost-Effective Development of
Real-Environment Speech Recognition Applications
(実環境音声認識システムの開発コスト削減のための選択学習法)

氏名 Tobias Cincarek

(論文内容の要旨)

音声によるコミュニケーションは人にとって自然であり、音声認識技術は、多くの応用分野で期待されている。しかしながら、現状の音声認識技術では、タスクに依存した音韻モデルや言語モデルを必要としている。通常音声認識システムの構築には、タスクでの実会話音声を収集して、音韻モデルと言語モデルを作成する必要があるが、これが音声認識システム構築のコストを大きくしている。とくに、音韻モデルの学習には、実会話音声の書き起こし作業が必要となっている。

この論文では、タスクからの少量の書き起こし音声データをもとに、既存の多様な音声データベースから自動的にタスクに合った音声データを選択するアルゴリズム(Selective Training)を提案している。これは、音声データベースを一発話単位にまで細分化しておいて、提示したタスクの音声データの尤度が大きくなるように選択するアルゴリズムである。学習の高速化のために隠れマルコフ(HMM)音韻モデルの十分統計量を利用している。このHMM十分統計量に基づく選択学習アルゴリズムの評価を、大量の音声データを収集することが難しい幼児と高齢者の音韻モデルの構築に適用して、有効性を実証した。

さらに、実際に音声情報案内システム「たけまるくん」「キタちゃん」「キタロボ」を運用して、実環境音声認識システムにおける音声データの書き起こしコストの削減を、音韻モデルの学習、言語モデルの学習、質問応答データベースの構築の観点から実証的なデータを示した。とくに、書き起こしを必要としない教師なし学習の導入により大幅な書き起こしコストの削減ができることを実証した。

これらの成果は、音声認識の実用化や運用に大きな貢献となっている。

氏名	Tobias Cincarek
----	-----------------

(論文審査結果の要旨)

音声認識システムの実用化の問題のひとつは、タスクに合った音声データを集めて、モデルの学習のための音声の書き起しのコストである。この論文では、少量のタスクに合った音声データを既存の多様な音声データベースを一発声ごとに分解した上で、タスクに合った音声を自動的に選択するアルゴリズムを考案した。このアルゴリズムは、HMM音韻モデルの十分統計量を利用したアルゴリズムで、尤度最大の評価基準であり、かつ、実際的な時間で実行できる画期的なアルゴリズムである。さらに、自ら中心となって、3台の音声情報システムを実環境で長期間にわたり運用して、実環境での音声認識システムの認識性能、ポータビリティ、運用コストを示したことも高く評価できる。

これらの成果を、電子情報通信学会などの3編の学术论文、ICASSPなどの重要な国際会議などで11編の論文として発表しており、国際的にも高い評価を得ている。理論だけでなく、実際の音声対話システムを自ら構築して、運用できる力をもっており、非常に貴重な人材である。また、日本語も読み書きおよび会話も日本人以上であり、また、英語でのコミュニケーション能力もTOEICが900点以上であることからわかるように非常に優れています。

音声認識の分野を深く理解した研究者として、国際的な舞台での活躍ができる研究者に育っており、世界レベルでの活躍が期待できます。

平成20年1月15日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は、博士論文(工学)として十分な価値があるものと判断した。