

論文内容の要旨

博士論文題目 大語彙連続音声認識を基盤とする実用指向
音声インタフェースに関する研究

氏名 西村 竜一

(論文内容の要旨)

音声認識の進歩とともに機械との対話を自由に行える音声インタフェースへの期待が高まっている。しかしながら、音声インタフェースが日常社会に普及した例は少ない。本研究では、最新の大語彙連続音声認識技術を用いて音声情報案内システムを研究開発し、実環境での運用を行った。実環境で、本格的な大語彙連続音声認識技術を用いた音声対話システムを運用した功績は大きなものである。本論文の成果を下記にまとめる。

(1) タスク適応言語モデルの構築法。

大語彙連続音声認識では、必要な認識精度を得るのに認識対象と合致した単語N-gram言語モデルが必要である。半自動に言語モデルを構築できるアルゴリズムを開発した。

(2) 音声対話受付案内ロボットASKA

ASKAは大学の受付案内が仕事であり、合成音声と手と頭のジェスチャにより来訪者に対応することができる。ASKAの音声認識、音声応答部を構築した。大語彙連続音声認識を用いたロボット音声対話システムが可能であることを示した。

(3) 音声情報案内システム「たけまるくん」の実環境への設置

一問一答形式の音声インタフェースを持ち、誰でも気軽に利用できる受付案内サービスを提供する音声情報案内システムを生駒市の北コミュニティセンターに常設し、1年以上にわたって利用者の発話を収録した。実環境での音声対話システムが運用できることを示した。大人で86%の単語認識率と76%の応答正解率を得ることができた。

(4) 音声情報案内システム「たけまるくん」の改良

音声情報案内システム「たけまるくん」では、子供の利用者が多いにもかかわらず、子供話者に対する性能不足も明らかになった。従来のシステムでは不十分であった子供に対する利便性の改善を目指して、年齢層に即した柔軟な対話を可能にする音声インタフェースを、収録音声データを活用して構築し、子供に対する性能を大幅に向上させた。

氏名	西村 竜一
----	-------

(論文審査結果の要旨)

大語彙連続音声認識システムは、静かな環境で丁寧に読み上げた音声に対しては、数万語彙でも95%以上の単語認識率を達成している。しかしながら、実環境での音声対話システムに適用して、長期間にわたり運用された実例は少ない。現在の音声認識技術の最も大きな課題は、実際のシステムに導入して、広く使われ、IT社会の基盤とすることである。

このような状況の中で、西村竜一君は、大学にしながら、社会で広く利用される音声対話システムを構築するという課題に積極的に挑戦し、生駒市の北コミュニティセンターに、音声情報案内システム「たけまるくん」を設置し、1年以上、ユーザ数も減少することなく、運用をなすとげた。この実行力、技術の活用方法などの先見性は、賞賛に値する。このシステムに対しては、情報処理学会から「第16回全国大会大会奨励賞」が授与されている。さらに、このシステムを1年以上運用し、従来音声認識が難しいとされてきた子供の音声データを大量に収集して、音韻モデル、言語モデルを大幅に改良し、子供の単語音声認識率を60%から80%にまで向上させたの特筆に価する。これらの結果は、電子通信学会の論文誌に採録になった。また、音声関連の権威のある国際会議 ICASSP2004にも採録となっている。

音声情報案内システム「たけまるくん」の前には、学内の共同プロジェクトとして、対話ロボット「ASKA」の開発を、ロボティクス講座と共同で行い、音声認識モジュール、音声合成モジュールの実装で主体的な役割を果たした。この結果は、ロボット関連の権威ある国際会議 IROS2002に採録され、発表を行った。

在学中に、3件の学術論文、6件の国際会議の発表を行った。それ以上に、実際の音声対話システムを構築できる技術力、さらに、実システムで収録した音声データベースの構築力など、研究からシステム開発までバランスのとれた研究者となっている。

平成15年12月25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。