

氏 名 井 ^い の ^{うえ} 上 ^{なお} 直 ^み 己
 学位(専攻分野) 博 士 (工 学)
 学 位 記 番 号 論 工 博 第 3355 号
 学位授与の日付 平成 10 年 5 月 25 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 電話サービスのための自由発話認識システムの研究

(主査)

論文調査委員 教授 堂 下 修 司 教授 池 田 克 夫 教授 美 濃 導 彦

論 文 内 容 の 要 旨

音声認識による電話サービスを広く普及させるためには、利用者がシステムと自由に情報伝達が行える必要がある。このような音声認識システムでは、不特定の話者の音声認識だけでなく、連続的に発声された音声の認識、システム設計者が予期していなかったill-formedな発話の認識(自由発話認識)および認識処理の高速化が重要である。しかし、自由発話認識には音声認識処理と自然言語処理を統合した音声言語処理の研究が必要であるが、言語知識の自由発話認識への適用の研究はまだ不十分であった。

本論文では、電話サービスにおける音声認識システムを実用化するため、内線電話受けシステムおよび電話オペレータ作業支援システムを対象に、より高度な言語的制約の利用による連続音声認識手法、ill-formedな発声の認識手法および認識処理の高速化手法に関する研究をまとめたものであり、序論と結論を含めて6章より成っている。

第1章では、本研究の背景および目的について示すとともに、本論文全体を通して用いている隠れマルコフモデル(HMM)による確率的手法と、それに基づく音声認識手法について説明している。

第2章では、内線電話受けシステムを対象に、言語的制約の最も基本的な適用形態である連続音声認識システムにおける音声言語認識処理の全体像を示している。より高度な言語的制約の利用により音声認識性能の向上を図るため、プラン認識モデルに基づく次発話予測機能により認識性能の向上を図る手法を述べ、その有効性を示している。

第3章では、内線電話受けシステムを対象に、自由発話認識の最重要課題として「不要語の出現」とその処理法を取り上げ、従来ではキーワードスポッティングによる小語彙の認識タスクにしか用いられていなかったガーベジモデルを連続音声認識にまで拡張して適用した自由発話認識手法について述べている。単語辞書と構文規則により音韻モデルを駆動する方法を不要語の認識に適用した場合、パーブレキシティをいかに抑えるかが重要である。ガーベジモデルは不要語全体を一つのHMMでモデル化するため、個々の不要語に対して音韻モデルを駆動する必要がなく、わずかな探索空間の増加で不要語を含む文の認識ができる。また、不要語を含む発話の認識では、不要語を含む文に対して高い認識性能を実現するだけでなく、不要語を含まない正常な文に対してその手法を適用しても認識性能が下がらないことも重要である。評価実験の結果、約18%の探索空間の増加で、不要語を含まない文の認識率(約93%)を低下させることなく、不要語を含んだ文の認識率を約48%と大きく改善できたことを述べている。

第4章では、電話サービスとしてオペレータ作業支援システムを取り上げ、認識処理の高速化手法について述べている。従来の音声認識処理では、音声区間検出に伴うマージンとして付加された無音も含んで認識処理しており、このマージン部分の処理に要する分だけ音声認識結果の出力は遅れることとなる。そこで、認識結果出力までの時間遅れを短くする、あるいは無くすために、対象タスクにより規定される発話内容が一意に確定されることを条件とする発話内容依存の認識処理終了条件を考案し、音声区間の途中であっても認識結果出力が可能となる手法を提案し、評価実験の結果、音声区間全てに対

して認識処理を行なう従来法と比較して認識率を低下させることなく、約40%の高速化が可能となることを示している。

第5章では、電話オペレータ作業支援システムを実際のサービスとして提供して収集したフィールド試験データに基づいて、第4章で示した認識手法が自由発話認識にも有効であることを実証している。まず、収集データを分析し、不要語は文末に挿入される場合が多いことを明らかにし、次に文末に不要語が挿入される文の認識に第4章で示す手法を適用して評価した。第4章で示す認識手法は音声区間の途中であっても認識結果の出力が可能であるため、文末に認識対象の表現とは無関係な文が挿入されたill-formedな表現の認識にも有効であり、評価実験の結果、第1位に正解が得られる割合で約29%、第5位までに正解が得られる割合で約15%と認識性能が大きく改善できることを示した。

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

高度情報化社会では、人とコンピュータとの自由な対話の手段として音声認識対話システムに対する必要性が高まりつつある。特に電話サービス系においては、音声対話システムは必須の手段となってきた。本論文は、電話サービスにおける音声認識システムの実用化を目標に、内線電話受付けシステムおよび電話オペレータ作業支援システムを対象として、より高度な言語的制約の利用による認識性能の改善、自由発話文の認識、認識処理の高速化についてまとめたものであり、得られた主な成果は以下のように要約される。

1. 内線電話受付けシステムを対象に、より高度な言語的制約の利用として、プラン認識モデルに基づく次発話予測機能により認識性能の向上をはかる手法を実現するとともに、認識実験により、その有効性を示している。
2. 内線電話受付けシステムを対象に、自由発話認識の最重要問題として「不要語の出現」を取り上げ、従来はキーワードスポッティングによる小語彙の認識タスクにしか用いられていなかったガーベジモデルを連続音声認識にまで拡張して適用する認識手法を提案し、認識実験を行いその有効性を示している。
3. 電話オペレータ作業支援システムを対象に、認識処理の高速化のため、新たな認識処理終了条件を考案し、終話検出前であっても認識結果出力が可能となる手法を提案している。評価実験の結果、本手法により認識処理時間を大幅に改善できることを明らかにしている。
4. 電話オペレータ作業支援システムを対象に、文末に認識対象の表現とは無関係な表現が挿入されたill-formedな表現の認識のためには、第3項に示す手法が有効であることを、実際の電話オペレータサービス業務について収集したデータに基づいて実証している。

以上要するに、本論文は、電話サービスにおける音声認識システムの実用化を目標に、認識性能の向上手法、自由発話文の認識手法、認識処理の高速化手法を提案し、それぞれに対して評価実験によりその有効性を明らかにしたものであり、得られた成果は實際上、学術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10年3月31日、論文内容とそれに関連した試問を行った結果合格と認めた。