

2001年6月21日

申請者氏名	安藤彰男
-------	------

紹介教官氏名	中川聖一教授
--------	--------

論文要旨(博士)

論文題目	リアルタイム字幕制作のための音声認識技術に関する研究
------	----------------------------

(要旨 1,200字程度)

最近の計算機技術の進展に伴い、音声認識も実用化への期待が高まっている。しかしながら、音声認識は、まだ研究途上の技術であり、自由に発話された音声を認識できるレベルには到達していない。本研究では、現状の音声認識技術でも、その用途をうまく制限すれば実用化可能であること、特に、テレビジョン放送における字幕のリアルタイム制作に利用できることを示す。

本論文で扱うのは、大相撲番組の字幕スーパーのための連続音声認識の研究、およびニュース番組の字幕放送実現のための連続音声認識の研究である。

大相撲番組字幕スーパー装置では、大相撲における場内アナウンス音声を認識対象としている。場内アナウンスは、基本的には文節単位で区切って発声されている。本研究は、文節内文法のみで発話内容を記述可能な対象を選ぶことにより、認識性能の向上を図ったものである。音声認識方式としては、各音節が母音で終わるという日本語の特徴を利用するため、ボトムアップ的な母音認識結果によって、トップダウン処理を制御する方式を採用した。さらに、認識性能向上のため、母音標準パターンの新しい学習法を開発した。この学習法は、識別学習を行う方法であるが、誤認識の程度を考慮した関数を最小化することにより、認識精度の向上と学習データへの過度な適合の回避を実現している。本論文では、従来から提案されている標準パターン学習法との比較実験により、その有効性を示す。

ニュース番組の字幕放送実現のための研究では、アナウンサーが原稿を読む部分を対象とした音声認識方式について述べる。この研究は、原稿読み部分という、どちらかといえば書き言葉に近い認識対象を設定することにより、音声認識の実用化を図ったものである。特に、ニュース番組の特徴を最大限利用することにより、アナウンサーがスタジオ内で発声した読み上げ音声であれば、高い性能で認識できることを示す。また、今後の課題として、図表などを用いた解説や、スポーツ、現場レポートなどの部分を行う際の問題を整理した。

アナウンサーの読み上げ音声を対象とした、ニュース音声認識の研究が、実用化レベルに到達したことから、音声認識を利用した字幕原稿制作システムを開発した。この際、誤りのない字幕放送を実現するため、音声認識結果中の認識誤りは人間が即座に修正することとした。NHKでは、このシステムを利用したニュース番組の字幕放送を、平成12年3月27日から実施している。これは、音声認識を利用してニュース字幕放送を行った、世界での初めての例である。さらに、音声認識技術の信頼性を示す実例として、「ニュース7」字幕放送の実施状況も紹介する。

今後は、ニュース番組全体の字幕化を目指して、解説、スポーツ、現場レポート部分の音声認識のほか、アナウンサーと記者との対談部分などの音声認識にも取り組んでいく。これらの研究テーマが、音声認識の研究の流れの中では、実用化レベルにある読み上げ音声の認識から、将来の研究課題である話し言葉音声の認識への掛け橋として位置づけられる重要な課題であることを指摘して、論文を締めくくる。