

電子・情報工学専攻	学籍番号	963425
申請者氏名	西崎 博光	

指導教官氏名	中川 聖一 永持 仁 梅村 恭司 宇津呂 武仁
--------	----------------------------------

論文要旨 (博士)

論文題目	音声文書を対象とした音声入力型 情報検索システムに関する研究
------	-----------------------------------

(要旨 1,200 字程度)

近年、音声認識技術の性能が向上してきており、ニュースの原稿や新聞の読み上げ音声に関しては、ほぼ誤認識することなく話し手の発話をそのまま書き起こすことができるようになってきた。実際に音声認識ソフトウェアが実用化されるようになってきており、たとえば音声のディクテーションソフトなどがある。

本研究では大語彙連続音声認識システムの精度向上に関する研究と音声認識システムを用いた音声文書検索に関する研究を行った。まず、この言語モデルの構築方法についての検討を行った。通常、N-gram 言語モデルでは単語単位にモデル構築を行うが、これを高頻度の単語の連鎖を一単位としてモデルを構築することで、モデルのパープレキシティを大幅に削減することができた。また、音声を使ったアプリケーションなどでは、認識された単語に対してその単語がどれだけ信頼できるかという指標を表す信頼度付与が行われることがある。本論文では、音声認識結果の単語に対する信頼度付与に関する研究も行った。これは、複数の認識システムの出力結果で、共通に出力された単語は信頼度が高いのではないかという考え方である。実際に、新聞の読み上げ音声を用いた実験では、共通で出力された単語の約 98% は正解単語であったことから、認識単語の信頼度付与には非常に有効であることを示した。

音声認識技術の応用として、音声文書の要約・検索が有望と考えられる。今後はテキストばかりでなく動画などを含むマルチメディアデータを扱うことが必要となることが予想できる。そこで、テレビニュースを対象とした検索システムの開発を行った。テレビニュースには当然ながらテキストデータというものも存在しないため、ニュース音声の大語彙音声認識の出力結果を用いてインデキシングを行う。自動的に音声を書き起こしインデックス化するため、できるだけ誤認識を回避する必要がある。そこで複数の大語彙認識システムを用いて、いろいろな認識率を持ったインデックスを複数作成し、認識率によって検索性能がどのように変わるかということ調べ、挿入誤りが大幅に増加してもできるだけ正解の単語を多く含んでいるインデックスから良い検索結果を得られることが実証された。この結果を踏まえ、複数の認識システムの出力結果を組み合わせたインデックスを構築することで、検索性能を改善することができた。

音声文書検索で誤認識と並ぶ大きな問題として未知語問題がある。この問題を解決する方法として、音声データに対して未知語検出処理を施すことで未知語区間を推定し、未知語区間を音節系列として書き起こした。未知語は音節列でインデキシングすることで、入力クエリーが未知語であれば、音節列表記とマッチングを行い、登録語であれば単語同士でマッチングを行うことで、未知語を含んだ音声文書でも検索性能が改善することを実験的に示した。

キーボードなどの入力装置がない携帯端末から検索システムを使用するのに音声によるクエリー入力は非常に有効である。音声クエリーの複数候補の認識結果から最適な結果を選択することでクエリーの認識率をあげることができるが、その反面、クエリーと全く関係のない単語の湧き出しも発生する。正解単語、無関係単語が混ざっている状態から検索に必要な候補だけを選択する方法として、単語間の相互情報量を用いたクエリー選択手法（グルーピングと呼んでいる）を提案し、その効果を検索実験で示した。また、検索要求の入力に対して、検索語を区切って入力するよりも、自然な文として入力する方が検索性能が良いことを実験的に示した。

最後に、音声文書検索の発展として、音声入力を用いた質問応答システムの開発を行った。質問応答システムでは、文書単位の回答を行うのではなく、ずばりユーザの質問に対する答えを提示するシステムである。そこで、単純な音声質問応答システムとそれを実現するのに必要な要素技術の一手法を提案し、実際の音声質問を用いた応答実験について述べる。