

## 論文内容の要旨

博士論文題目 音声対話インタフェースのための雑音に頑健な音声入力の研究

氏名 伊田 政樹

### (論文内容の要旨)

音声認識を利用する音声インタフェースは、誰でもが容易に使える便利なインタフェースである可能性が大きい。近年、音声認識の研究開発が進展して、さまざまなアプリケーションで音声認識技術が一般ユーザに利用されるようになった。しかしながら、実験室の環境では、高い認識率を達成しているが、実環境下では、まだ、さまざまな問題を抱えており、十分な性能を達成していない。この論文では、実環境で音声認識を利用する際に必要となる次の2つの技術の改善について述べている。

(1) 変化する雑音環境下でも頑健に動作する音声認識性能の実現と、(2) ユーザがマイクを意識しなくても自由に音声入力できるハンズフリー音声入力インタフェースの実現について述べている。

第一の課題にたいして、雑音 GMM (Gaussian mixture model) を利用した環境雑音適応化と SN 比別マルチパスモデルを用いた HMM 合成法を提案している。HMM 合成法は環境雑音のモデル化を行う際に適応データとして雑音データのみを用いる方法である。HMM 合成法の欠点は多量の適応データ量を必要とする点と雑音の変動に対して対応できない問題がある。提案法は、さまざまな雑音データを用いて準備した初期雑音 GMM を選択適応することで、1秒程度の雑音データでの雑音適応を実現している。

第二のハンズフリー音声入力課題にたいしては、情報提供端末向けにハンズフリー音声入力インタフェースを、マイクロホンアレーを用いて開発している。実際のアプリケーションとして券売機や情報提供端末を取り上げ、遅延和形マイクロホンアレー処理とスペクトル減算による雑音除去を組み合わせた音声入力インタフェースを試作している。実環境下における認識性能評価実験の結果、216 単語の孤立単語認識で 91.6% の認識性能を達成し、認識性能の改善が達成できている。

以上のように、本研究では実環境下での音声認識システムの利用における二つの課題について検討し、成果を得ている。

氏名	伊田 政樹
----	-------

(論文審査結果の要旨)

音声認識を利用する音声インタフェースは、誰でもが容易に使える便利なインタフェースである可能性が大きい。しかしながら、実環境下では、まだ、さまざまな問題を抱えており、十分な性能を達成していない。この論文では、実環境で音声認識を利用する際に必要となる2つの技術の改善について研究を行い、有効な手法を提供している。

(1) 変化する雑音環境下でも頑健に動作する音声認識性能の実現と、(2) ユーザがマイクを意識しなくても自由に音声入力できるハンズフリー音声入力インタフェースの実現について述べている。

(1) に対して、雑音 GMM (Gaussian mixture model) を利用した環境雑音適応化と SN 比別マルチパスモデルを用いた HMM 合成法を提案している。この手法により、環境雑音の変動が常に起こっているような環境でも、HMM 合成法が利用できることを示した。この成果は、PDA のような移動体、カーナビなどでの実環境での音声対話システムの頑健化に役立つことが期待できる。この成果は、電子情報通信学会の論文誌、国際会議などにも採録されている。

(2) のハンズフリー音声入力課題に対しては、情報提供端末向けにハンズフリー音声入力インタフェースを、遅延和形マイクロホンアレー処理とスペクトル減算による雑音除去を組み合わせで開発している。これらの技術の評価は、実環境で行われた例はなく、実際のアプリケーションとして、券売機や情報提供端末を取り上げ、音声入力インタフェースを試作して評価を行った点は高く評価される。この成果は、電子情報通信学会のシステム開発論文として発表されている。

以上のように、本研究では実環境下での音声認識システムの利用における二つの課題について検討し、成果を得ている。現在、伊田君はオムロンにおいて、これらの成果を活用して音声対話システムの研究開発を行っており、実環境、商用の観点からの音声対話システムに対する知見も備えている。また、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、上記の観点を踏まえて、本論文は、博士(工学)として十分な価値があるものと判断した。