

論文内容の要旨

博士論文題目

Statistical Acoustic Model Adaptation for Robust Speech Recognition
in Noisy Reverberant Environments

(雑音・残響環境下でのロバストな音声認識のためのモデル適応化法)

氏名

滝口 哲也

(論文内容の要旨)

(1, 200字程度)

現在の音声認識システムは、パソコンでの動作も可能となっており、実用化段階にある。しかしながら、これまでの音声認識システムを使うには、比較的静かな環境下で使用しないといけない。更に、ユーザはマイクロフォンに向かって（マイクロフンの前で）発話する必要がある。なぜなら、もし雑音及び残響環境下においてマイクロフォンから離れて発話した場合、音声認識精度が著しく劣化してしまうためである。

雑音及び残響環境下で発話された音声は、周囲の雑音及び残響の影響を受けてしまい、音声モデル作成時の学習データとの間に mismatches が生じてしまう。それらの影響に対処するために、本論文では従来の音声 HMM (Hidden Markov Model) と雑音 HMM の合成に加えて、音響伝達特性 HMM の合成を試みる。この HMM の各状態を音源位置に対応させることにより、ユーザの自由な場所移動にも対処することが可能になり、ユーザインタフェースの向上が実現される。

音響伝達特性 HMM を作成するためには、認識を行なう前にあらかじめ各場所からの音響伝達特性を測定しておく必要があった。しかしながら実際の環境において、あらかじめ音響伝達特性を測定しておくのは非現実的である。そこで本論文では、更に HMM 分解法による音響伝達特性 HMM の推定方法を提案する。HMM 分解法では、ユーザの場所が既知である必要はなく、任意の場所から発話された観測データを用いて音響伝達特性 HMM の推定が可能である。雑音・残響環境下では、この HMM 分解法が 2 回適用される。まず、周波数領域において雑音 HMM からの分解を最尤推定にもとづいて行ない、更に領域変換を行ないケプストラム領域において、音響伝達特性 HMM を最尤推定にもとづいて分解する。この HMM 分解法により各々の場所からの音響伝達特性を推定し、HMM 合成法により雑音 HMM 及び音声 HMM と組み合わせることにより、対象とする雑音及び残響環境下での音声モデルを作成し、音声認識を行なう。

提案手法の評価実験を実際に雑音及び残響環境下にてマイクロフォンから約 2m から 3m 離れた場所から発話された音声に対して行ない、その有効性を確認した。しかしながら、まだこの技術だけでは、実際の様々な環境下においてハンズフリー音声認識を可能にするレベルまで達してはいない。今後は、マイクロフォンアレイとの統合なども含め他の手法と併用することにより更に効果が期待できる。本論文により、今後の研究の発展と伴わせて、ハンズフリー音声認識の実現へ向けて大きな可能性を示せた。

氏名	滝口 哲也
----	-------

(論文審査結果の要旨)

音声認識技術は、確率モデルとデータベースの利用により著しい進歩を見せているが、現状の技術では、接話マイクロホンや静かな環境でしか十分な性能を得ることができない。この原因は、環境における雑音などの加法性雑音と部屋の音響特性に起因する乗法性歪みが原因である。マイクロホンから離れて発話すると、これらの加法性雑音と乗法性歪みにより致命的な音声認識性能の劣化が生じる。本論文では、このような加法性雑音および乗法性歪みに対し、音声認識の確率モデルである隠れマルコフモデル (HMM) のモデルを観測データに適合させるモデル適応化により対処する方法の基礎研究について述べている。

本論文の成果は、次のように要約できる。

(1) 実環境における雑音などの加法性雑音および部屋の音響伝達特性などの乗法性歪みを確率モデルである HMM によりモデル化し、HMM 合成法に基づいて、これら音声モデル、雑音モデル、部屋の伝達特性モデルの HMM の合成を行い、観測信号に一致した 1 つの HMM モデルを生成する方法を提案している。

(2) 複数の HMM モデルを合成する HMM 合成法では、あらかじめそれぞれの要因を HMM として学習しておく必要が有る。本論文では、観測される環境における信号に基づいてそれぞれの HMM を推定する HMM 分解法を提案している。HMM 分解法は、HMM 合成法の逆のプロセスであり、1 つの既知の HMM を 1 つの既知の HMM と未知の HMM に最尤推定法に基づき分解する方法として導出されることが示されている。あらかじめ測定することが困難である部屋の音響伝達特性の推定に適用し、HMM 分解法が有効であることを示している。複数の要因を HMM で表現し合成する際の、HMM パラメータの推定・合成の一貫した枠組みの完成に成功したことについて、極めて高く評価できる。

(3) HMM 合成・分解法を加法性雑音、乗法性歪みを含んだ実環境音声で評価し有効性を確認している。論文では、以下の評価が行われている。

- ・雑音・残響の存在する実際の環境においてマイクロホンから離れて移動しながら発話下音声の認識に対する効果の評価
- ・電話回線の音声認識。通常電話回線による乗法性歪みに加えて、コードレス電話の使用による乗法性歪みへの効果の評価

いずれの場合においても、提案する HMM 分解・合成法により大きな改善が得られることが示されている。

以上のように、本論文は実環境下において発話された音声の認識を目的に、HMM 分解・合成法に基づくモデル適応化手法を提案し、有効性を示した。これらの研究成果は、情報科学、音情報処理学の分野における学術的な貢献として高く評価できる。平成 10 年 12 月 25 日開催の公聴会の結果を参考に、本博士論文の審査を行なった結果、本論文は博士(工学)の学位論文として十分に価値のあるものと判断する。