

(論文内容の要旨)

本論文は、雑音下における音声発話区間の自動検出を対象としたオンライン教師なし学習に関する研究をまとめたものである。音声認識などのパターン認識技術の多くは、大規模なデータベースを用いた統計モデルの教師あり機械学習によって成功を収めているが、あらゆる環境の雑音を事前に収集して学習することは事実上不可能であり、また刻々と変化する雑音環境にも対応できる必要がある。そこで本論文では、まず雑音に頑健な音響特徴量に関する検討を行い、音声区間検出を音声と雑音の2クラス分類問題ととらえた上で、教師なしでオンラインに学習する枠組みについて述べている。

第1章は序論で、本研究で扱う音声区間検出の問題について述べた後、従来の手法、特に教師あり学習やモデル適応の問題点について指摘した上で、本研究で採用するオンライン教師なし学習のアプローチについて述べている。

第2章では、音声区間検出の従来手法、特に信号処理に基づく手法を概観し、実装において多くのヒューリスティックやパラメータ調整が必要となる問題点を指摘している。

第3章では、雑音に頑健な音声区間検出のための高次統計量に基づく特徴量を提案している。線形予測分析の予測残差は音声の調波性を表す特徴量であるが、そのキュムラントと自己相関関数を組み合わせたスカラー量を考案し、ガウス性雑音と突発性雑音の両方に対応できることを述べている。実際に、オープンハウスにおいて収録された会話音声を対象として評価を行った結果、本特徴量を用いることにより、音声のパワーやキュムラントのみを用いる場合に比べて有意に検出精度が高くなることを示した。また、音声区間検出の評価のための標準的なデータベースである CENSEREC-1-C を用いた実験においても有効性を確認した。

第4章では、音声と雑音の2クラス分類をオンライン教師なし学習で行う方法を述べている。両者の特徴量の分布をそれぞれ正規分布と仮定し、2混合正規分布を推定する問題として定式化した上で、オンライン EM アルゴリズムを適用している。これにより、事前学習を一切必要とせず、音声の入力環境に応じて自動的にパラメータ推定ができる。前章と同じ評価データを用いた実験において、本手法は、全データを用いて一括学習する通常の EM アルゴリズムと同程度の精度を、オンラインで実現できることを示している。また、事前に多くの種類の雑音を学習した GMM と比較して、有意に高い精度が得られることを示し、入力音声に特化してオンラインで学習することの効果を確認している。

第5章では、前章の教師なしのオンライン EM アルゴリズムが、無音（雑音）区間が長く続く場合には想定外の動作に陥ることを指摘し、ベイズ学習の枠組みを導入することで、教師なしモデル学習の信頼度を推定することを提案している。具体的には、変分ベイズ (VB) 法によって定式化し、自由エネルギーの確率的な近似計算を行う。このオンライン VB-EM アルゴリズムにより、音声及び雑音が混合している場合と雑音のみの場合について、正規分布のパラメータ推定とそれぞれの信頼度計算を同時に行うことが可能になる。この改善により、前章のオンライン EM アルゴリズムと比べて、音声検出誤りが約半分に削減できることを示した。CENSEREC-1-C データベースを用いた評価において、最高水準の既存手法と比較して、接話マイクでは同等、遠隔マイクで2割以上低い誤り率を実現した。また、様々な雑音の種類や SN 比に対して安定に動作することも示した。

第6章では、本研究の結論を述べており、提案手法の音声区間検出以外のタスクへの適用可能性について議論している。

(論文審査の結果の要旨)

本論文は、雑音下における音声発話区間の自動検出を対象としたオンライン教師なし学習に関する研究をまとめたものであり、得られた主な成果は次の通りである。

1. 雑音に頑健な音声区間検出のために、線形予測分析の予測残差の高次統計量に自己相関関数を組み合わせた特徴量を提案し、2種類のデータベースを用いた評価実験において有効性を示した。
2. 音声区間検出を2混合正規分布の教師なし学習の問題としてとらえた上で、オンライン EM アルゴリズムにより実現できること、特にオンライン EM 推定がオフラインの EM 推定と同様の精度で実現できることを示した。本手法は、事前学習を一切必要とせず、音声の入力環境に応じてパラメータ推定ができる点で優れている。
3. 上記のオンライン EM 推定を変分ベイズ学習の枠組みで実現し、混合分布数の最適な数も同時に推定することにより、雑音のみからなる区間を棄却できる枠組みを導入した。この最終的な提案手法は、大規模なデータベースを用いた評価実験において、最高水準の既存手法と比べて、高精度でかつ頑健な音声区間検出を実現した。

以上のように本論文は、事前学習や予測が困難な雑音に対応する上で、教師なしオンライン学習が有効であることを実証し、ベイズ学習の新たな適用可能性を示したもので、学術上・實際上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。

また、平成 21 年 6 月 23 日実施した論文内容とそれに関連した試問の結果合格と認めた。