

学位論文題名

音声の特徴抽出のための適応ラティス分析と
スペクトル制御に関する研究

学位論文内容の要旨

現在、音声情報処理の分野では、線形予測分析法が主流となっており、それを技術的な基盤として、音声認識、合成、低ビット符号化などのシステムが構築されている。しかし、音声認識の現状を考えると、特定話者あるいは有限語彙など機能を限定した小規模なシステムが実用化されているにすぎない。これにはいくつかの原因が考えられるが、その一つに一括型 AR モデルに基づく線形予測分析法の分析精度の限界が指摘されている。

本論文は、音声信号から正確でかつ音声認識に適した特徴を抽出する音声分析システムの構築を行い、それらの有効性を実験により検証したものである。

本論文ではまず、従来の一括型 AR モデルに基づく線形予測分析法と比べて高精度な分析が可能なラティス型 ARMA モデルを用いた適応的分析手法の開発を行っている。次に、適応的分析手法の問題点を改善するために、非線形処理を用いた推定スペクトルの平滑化手法の開発を行っている。最後に、実用の音声認識システムを構築するときに問題となる環境雑音の影響を軽減することを目的とした音声認識向きスペクトル制御法を開発し、様々な条件下での実験、評価を行っている。

論文は全 8 章からなる。以下に、各章の概略を述べる。

第 1 章では、本研究の背景および目的、従来の関連する研究の概要と当面する課題、本論文の構成について述べる。

第 2 章と第 3 章では、高精度な音声分析システムを構築することを目的として、従来の一括型 AR モデルによる線形予測分析法を拡張し、ラティス型 ARMA モデルを用いた適応的分析手法を提案する。

第 2 章では、ARMA モデルの性質をラティス型 ARMA モデルを中心にして述べる。はじめに、最小二乗法による推定 ARMA 係数と相関関数は、その対応が一意ではないことを示す。この性質により、ARMA モデルから低次元ノモデルを求めるとき、同じ ARMA モデルから異

なる低次元モデルが得られることを示す。次に、3つの予測モデルに基づくラティス型 ARMA モデルの構造と、分析アルゴリズムを示す。提案するラティス型 ARMA モデルは、規格化により任意次数のラティスモデルを最小個数のパラメータで設計でき、非冗長な特徴表現が可能である。次にラティス型 ARMA モデルの性質を知るため、3つの予測モデルのうち2つの後向き予測モデルで得られる伝達関数を調べ、その結果、後向きモデルの違いにより、モデルの伝達関数が等しくても、係数が異なるラティス型 ARMA モデルが設計されることを示す。

第3章では、第2章で示されたラティス型 ARMA モデルを拡張し、音声信号を適応的に分析する手法を提案する。ARMA モデルを用いて音声信号の時変スペクトルを推定するために必要な規範モデルの入力の推定法とラティス型 ARMA モデルの係数を回帰的に計算する時間更新式を示す。また、従来の適応型ラティス分析では、アルゴリズムの制約上評価関数に導入した重み係数を動的に設定することができず、急峻な変化への追従性に問題があった。提案する手法では、分析に用いる信号に直接重み付けをする新しい重み係数を導入したため、信号の変化に合わせて重み係数を設定することが可能である。重み係数の値は規範モデルの変化の度合に対応して設定する必要がある。本手法では、固定重み係数を用いた AR 分析を同時に行い、それを基準にして規範モデルの変化点を検出する。実音声を用いた分析実験により、本手法が従来の固定重み係数を用いた分析法より高精度な推定スペクトルが得られることを示す。

第4章と第5章では、適応分析手法で得られる推定スペクトルは時間依存性が高く、平滑化処理が必要であることをのべ、非線形処理を用いた平滑化について検討する。

第4章では、従来の線形フィルタによる平滑化では、スペクトルの時間特性が損なわれる問題点があることを述べ、非線形フィルタの一つであるメディアンフィルタを用いた平滑化を検討する。また、特徴パラメータを独立と見なした平滑化では、元のスペクトル時系列中に存在しないスペクトルが得られる。そのため、スペクトルに歪みが生じる可能性があることを述べ、特徴パラメータと平滑化フィルタを組み合わせた実験をもとに、歪みの生じない平滑化について考察する。その結果、ラティス係数を対象とし、メディアンフィルタを用いて平滑化することにより、時間追従性を損なわずに歪みの少ない平滑化スペクトルが得られることを示す。

第5章では、平滑化により失われたスペクトルの時間特性を、非線形時間伸縮することで回復する手法を提案する。平滑化によりスペクトル時系列にある程度の時間遅れが生じたとしても、平滑化後に元のスペクトルの時間パターンを参照して、平滑化スペクトルを時間伸縮すれば時間特性を改善できる。提案する手法では、DP マッチングを利用して、平滑化スペクトル時系列を非線形に時間伸縮する。これにより、平滑化スペクトル時系列は、参照パターンに用いたスペク

トル時系列の時間特性を最も良く近似するように伸縮される。実音声による実験結果から、本手法ではスペクトルに時間的な乱れが生じることなく時間特性が改善されることを示す。

第6章と第7章では、実用的な音声認識システムを構築するとき問題となる環境雑音や伝送路の周波数特性の影響を軽減するため、音声認識に適したスペクトル制御法について述べる。

第6章では、拡張軸外スペクトルを用いたスペクトル制御法を提案する。環境雑音の影響はスペクトルに均一に表れるものではなく、雑音の影響を受け易い部分と受けにくい部分があり、雑音に影響を受けにくい部分を重み付けすることで、雑音の影響を軽減できる。拡張軸外スペクトルは、制御対象となるARモデルをARMAモデルに拡張し、周波数領域での一つのピークを一組の極と零で制御する。そのため、スペクトルピークの鋭さとすそ野の傾きを同時に制御できる。また、実音声に対するモデルの安定性を保証した制御パラメータの設定法に示し、白色雑音を付加した音声や、周波数特性をもつフィルタに通した音声に対する音声認識実験から、提案する手法の有効性を検証する。

第7章では、音声分析の段階で、推定スペクトルから雑音成分を除去する方法を検討する。音声信号をAR過程と仮定し、白色雑音が付加された場合のスペクトルの影響を考察し、白色雑音が付加された音声信号のスペクトルを正確に近似するにはARMAモデルを用いる必要があることを示す。そのとき、推定されたARMAモデルでは、AR部で信号成分をMA部で雑音成分を近似していると考えられる。提案する手法では、第2章で示された、ARMAモデルの性質を利用して、ARMAモデルからARモデルに変換するとき、AR部に重み付けしてリダクションすることにより、雑音成分を除去する。本手法を第6章で提案した拡張軸外スペクトルと組み合わせて使用することにより、低SN比の雑音環境下での認識性能の改善をはかる。

第8章は結論として、本論文に述べた研究を総括し、今後の課題を述べる。

学位論文審査の要旨

主査	教授	梶内香次
副査	教授	小川吉彦
副査	教授	永井信夫
副査	教授	新保勝
副査	助教授	宮永喜一

本論文は、音声の自動認識システムの構築を目的し、その基礎研究として音声信号の高精度な分析法ならびに音声認識に適したスペクトル制御法を提案し、それらの有効性を実験により検証したもので、全8章から構成されている。

第1章では、本研究の背景および目的、本論文の構成について述べている。

第2章では、最小二乗法によるARMAモデルの係数推定法と推定ARMAモデルの性質について述べている。また、最小個数の係数で設計される最小次元実現ラティス型ARMAモデルの構成法を示している。

第3章では、第2章で示されたラティス型ARMAモデルの適応的係数推定法を提案し、音声分析に適用している。このラティス型ARMA分析手法では、分析に用いる信号に直接重み付けをする新しい重み係数を導入しており、評価関数に導入した従来の重み係数ではアルゴリズムの制約上不可能であった重み係数の動的な設定を可能にしている。重み係数の設定では1次の適応AR係数の値を基に有声音声と無声音声の分離を行い、別個のルールを適用して重み係数の値を変化させている。実音声を対象とした分析実験の結果より、本手法が従来の固定重み係数を用いた分析法より高精度な推定スペクトルが得られることが示されている。

第4章では、従来の線形フィルタによるスペクトル平滑化では、スペクトル時系列のもつ時変特性が損なわれる問題点があることを述べ、非線形フィルタの一つであるメディアンフィルタによる平滑化を検討している。また、特徴パラメータを各々独立と見なした平滑化では、スペクトルに歪みが生じる可能性があることを示し、平滑化の対象となる特徴パラメータと平滑化フィルタを組み合わせた実験により、歪みの生じない平滑化について考察を行っている。その結果、ラティス係数を対象とし、メディアンフィルタで平滑化することにより、時間追従性を損なわずに歪みの少ない平滑化スペクトル時系列が得られることを示している。

第5章では、平滑化により損なわれたスペクトル時系列の時変特性を時間伸縮することで回復

する手法を提案している。提案手法では、DP マッチングを利用して参照スペクトル時系列の時間特性を最も良く近似するように平滑化スペクトル時系列の非線形時間伸縮が行われる。時間伸縮は、個々のパラメータ時系列を独立に時間伸縮する方法と、パラメータベクトル毎にまとめて時間伸縮する方法の2つの方法を検討しており、実験により平滑化スペクトルに時間的な乱れが生じることなく、時間特性を改善できることを示している。

第6章では、環境雑音による認識率の低下を防ぐために、音声認識の前処理として拡張軸外モデルを用いたスペクトル制御を提案している。提案する手法では、制御対象となるARモデルをARMAモデルに拡張し、周波数領域での一つのピークを一組の極と零で制御するため、ピークの鋭さとすそ野の傾きを同時に制御できる。更に、実音声に対するモデルの安定性を保証した制御パラメータの設定法を示し、白色雑音を付加した音声や、周波数特性をもつフィルタに通した音声に対する音声認識実験から、提案する手法の有効性を確認している。

第7章では、推定スペクトルから雑音成分を除去する方法を検討している。白色雑音が付加された音声信号のスペクトルを正確に近似するにはARMAモデルを用いる必要があり、その場合AR部で音声信号成分をMA部で雑音成分を近似していることを示している。提案する手法では、その性質を利用してARMAモデルからARモデルに変換するとき、AR部に重み付けをすることにより、雑音成分の除去を行っている。本手法を拡張軸外スペクトルと組み合わせて、低SN比の雑音環境下での認識性能の改善を行っている。

第8章は結論で、本論文に述べた研究を総括し、今後の課題を述べている。

本論文の成果は以下のように要約される。

- (1) 従来の線形予測分析法と比べて高精度な分析が可能なラティス型ARMAモデルを用いた適応分析手法を提案した。
- (2) 適応分析手法の問題点を改善するために、非線形処理を用いたスペクトル平滑化手法を提案した。
- (3) 環境雑音の影響を軽減するために、音声認識向きスペクトル制御法を提案した。

これを要するに本論文は、音声の新しい分析手法の提案を行い、その有効性を理論ならびに実験の両面にわたって示したもので、デジタル信号処理工学ならびに音声情報処理工学に対して貢献するところ大なるものがある。

よって著者は、博士(工学)の学位を授与される資格あるものと認める。